Signal & Image Processing

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

- 11th International Conference on Signal & Image Processing (SIP 2022)
- 2nd International Conference on Cryptography and Blockchain (CRBL 2022)
- 9th International Conference on Wireless and Mobile Network (WiMNeT 2022)
- 2nd International Conference on Big Data, IoT and Machine Learning (BIOM 2022)
- International Conference on Education, Pedagogy and Technology (EDUPT 2022)
- 2nd International Conference on NLP & Data Mining (NLDM 2022)
- International Conference on Software Engineering Advances and Formal Methods (SOFTFM 2022)
- 6th International Conference on Computer Science and Information Technology (COMIT 2022)

Published By

AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai (Eds),
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403
ISBN: 978-1-925953-77-0
DOI: 10.5121/csit.2022.121701 - 10.5121/csit.2022.121718

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

11th International Conference on Signal & Image Processing (SIP 2022), October 22 ~ 23, 2022, Sydney, Australia, 2nd International Conference on Cryptography and Blockchain (CRBL 2022), 9th International Conference on Wireless and Mobile Network (WiMNeT 2022), 2nd International Conference on Big Data, IoT and Machine Learning (BIOM 2022), International Conference on Education, Pedagogy and Technology (EDUPT 2022), 2nd International Conference on NLP & Data Mining (NLDM 2022), International Conference on Software Engineering Advances and Formal Methods (SOFTFM 2022), 6th International Conference on Computer Science and Information Technology (COMIT 2022) was collocated with 11th International Conference on Signal & Image Processing (SIP 2022). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The SIP 2022, CRBL 2022, WiMNeT 2022, BIOM 2022, EDUPT 2022, NLDM 2022, SOFTFM 2022 and COMIT 2022. Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, SIP 2022, CRBL 2022, WiMNeT 2022, BIOM 2022, EDUPT 2022, NLDM 2022, SOFTFM 2022 and COMIT 2022 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SIP 2022, CRBL 2022, WiMNeT 2022, BIOM 2022, EDUPT 2022, NLDM 2022, SOFTFM 2022 and COMIT 2022

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld,
Dhinaharan Nagamalai (Eds)

General Chair

Organization

David C. Wyld, Southeastern Louisiana University, USA
Dhinaharan Nagamalai (Eds) Wireilla Net Solutions, Australia

Program Committee Members

Abbas Al-Thaedan, Al-Muthanna University, Iraq
Abdel-Badeeh M. Salem, Ain Shams University, Egypt
Abdelhadi Assir, Hassan 1st University, Morocco
Abderrahmane EZ-Zahout, Mohammed V University, Morocco
Addisson Salazar, Universitat Politècnica de València, Spain
Afaq Ahmad, Sultan Qaboos University, Oman
Ajay Anil Gurjar, Sipna College of Engineering and Technology, India
Alexander Gelbukh, Instituto Politécnico Nacional, Mexico
AlirezaValipourBaboli, University Technical and Vocational, Iran
Altynbek Sharipbay, Eurasian National University, Kazakhstan
Álvaro Rocha, University of Coimbra, Portugal
Aman Malhotra, SRM Institute of Science and Technology, India
Amina El murabet, Abdelmalek Essaadi University, Morocco
Ana Leal, University of Macau, China
Ana Luísa Varani Leal, University of Macau, China
Anand Nayyar, Duy Tan University, Viet Nam
Anastasios Doulamis, National technical University of Athens, Greece
Anirban Banik, National Institute of Technology Sikkim, India
Anita Yadav, Harcourt Butler Technical University, India
Anouar Abtoy, Abdelmalek Essaadi University, Morocco
António Abreu, ISEL - InstitutoPolitécnico de Lisboa, Portugal
Archit Yajnik, Sikkim Manipal University, India
Ashkan Ebadi, Concordia University, Canada
Ashraf Elnagar, University of Sharjah, UAE
Asimi Ahmed, University Ibn Zohr, Morocco
Assem Abdel Hamied Moussa, chief Eng, Egypt
B Nandini, Telangana University, India
Badir Hassan, Abdelmalek Essaadi University, Morocco
Bahaeddin Turkoglu, Konya Techical University, Turkey
Bandu B. Meshram, VeermataJijabai Technological Institute (VJTI), India
Benyamin Ahmadnia, University of California, United States
Beshair Alsiddiq, Riyadbank, Saudi Arabia
Bhagyashree SR, ATMECE, Mysore
Bogdan Wiszniewski, Gdansk University of Technology, Poland
Brahami Menaouer, National Polytechnic School of Oran, Algeria
Brahim Lejdel, University of El-Oued, Algeria
Bunil Kumar Balabantaray, National Institute of Technology, India
Cagdas Hakan Aladag, Hacettepe University, Turkey
Carla Osthoff, National Laboratory for Scientific Computing, Brazil
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Casalino Gabriella, University of Bari, Italy
Chang-Yong Lee, Kongju National University, South Korea
Cheng Siong Chin, Newcastle University in Singapore

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

11th International Conference on Signal & Image Processing (SIP 2022)

2nd International Conference on Cryptography and Blockchain (CRBL 2022)

9th International Conference on Wireless and Mobile Network (WiMNeT 2022)

2nd International Conference on Big Data, IoT and Machine Learning (BIOM 2022)

International Conference on Education, Pedagogy and Technology (EDUPT 2022)

2nd International Conference on NLP & Data Mining (NLDM 2022)

International Conference on Software Engineering Advances and Formal Methods (SOFTFM 2022)

6th International Conference on Computer Science and Information Technology (COMIT 2022)

HANDWRITTEN DIGIT RECOGNITION SYSTEM BASED ON CNN AND SVM

Yousra Berrich and Zouhair Guennoun

Smart Communications Research Team - ERSC,
Mohammadia School of Engineering, Rabat, Morocco

ABSTRACT

The recognition of handwritten digits has aroused the interest of the scientific community and is the subject of a large number of research works thanks to its various applications. The objective of this paper is to develop a system capable of recognizing handwritten digits using a convolutional neural network (CNN) combined with machine learning approaches to ensure diversity in automatic classification tools. In this work, we propose a classification method based on deep learning, in particular the convolutional neural network for feature extraction, it is a powerful tool that has had great success in image classification, followed by the support vector machine (SVM) for higher performance. We used the dataset (MNIST), and the results obtained showed that the combination of CNN with SVM improves the performance of the model as well as the classification accuracy with a rate of 99.12%.

KEYWORDS

Classification, feature extraction, convolutional neural network, support vector machine, MNIST.

1. INTRODUCTION

The recognition of handwritten digits (RCM) is part of the Optical character recognition (OCR) among its applications: the postal sorting where every day thousands of items are automatically sorted, moreover there is the reading of the numerical amount of bank checks.

Even after several works in this direction, the recognition performance remains far from that of the human eye due to the similarity between handwritten digits or the difference between writing styles, so it is necessary to develop a system able to acquire an image of an input digit and recognize the digit present in the image. The objective of this work is to create an automatic system capable of recognizing the number of manuscripts in the MNIST database with resolution (28*28) pixels [01].

In this proposed model, we decide to verify its efficiency in recognizing the handwritten digit while using the SVM classifier at the last layer of the CNN neural network; this proposed method shows an excellent performance with a high accuracy.

The structure of this paper is articulated around three parts: the first part is devoted to give a general overview on some previous works in the direction of handwritten digit recognition, then in the second part we will present the process of the handwritten digit recognition system and at the end we will elaborate an implementation for handwritten digit recognition using a

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 01-06, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121701

classification approach based on convolutional neural network (CNN) combined with SVM and we will present our results obtained in the form of graphs and in a summary table.

2. SOME PREVIOUS WORKS

The recognition of the writing is known under the name of O.C.R (Optical Character Recognition). The first works date back to the 1900s by TYURIN, during which the scanning scanner for television was invented and in the year 1912 by ALBE and in the year 1925 by Thomas, which mimicked the human interpretation of visual computing, a point of transformation appeared with the invention of the first computer in 1946 by MAUCHLY and ECKERT, a few years later the first experiments in character recognition could be carried out during the sixties, and seventies, the first systems of automatic writing of the printed text were born. [02]

In addition, in 1975 the Japanese were using readers that decipher the postal code written by hand or typed. In the same period, the French company CONTER built a system of automatic reading of printed text intended for the blind, then the American company KURZWELL improved the previous system by proposing reading machines for the blind formulating the text aloud by vocal synthesis. [03]

During this phase the researchers encountered many difficulties in addition to the complexity of the problem of recognition due to the great variability of handwriting, the non-availability of memory and computing power for the realization of concrete systems operating in real time, on the other hand since 1980 when the recent electronic progress and more particularly the advent of powerful computers at low cost made it possible to solve this type of problem and the research in handwriting recognition multiplied in a spectacular way, and many new techniques were born.

3. CHARACTERIZATION OF THE HANDWRITTEN DIGIT RECOGNITION SYSTEM

Handwritten digit recognition systems are generally based on the following main steps:
Signal acquisition, pre-processing, feature extraction, classification.

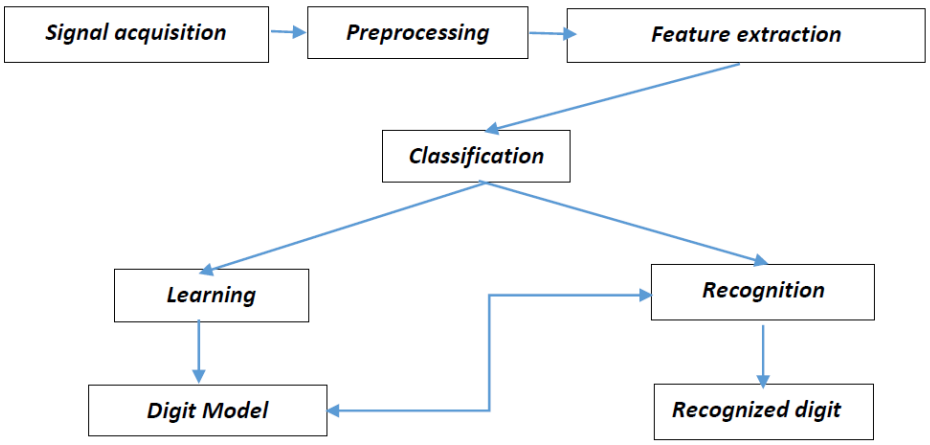


Figure 1. The process of the handwritten digit recognition system

3.1. Signal Acquisition

This step consists in converting a paper document into a digital image format with a minimum of degradation. However, despite all this, some noise may appear and cause a loss of quality due to the work area and its lighting [04].

3.2. Preprocessing

The purpose of pre-processing is to reduce the noise superimposed on the data and to normalize the image size while keeping the significant information of the presented shape. The pre-processing operations generally used are binarization, smoothing (noise), normalization (size normalization, straightening normalization), Tinning [05].

3.3. Feature Extraction

In a handwritten digit recognition system, the feature extraction phase consists of obtaining the most relevant volume of information that will be provided to the classifier later on.

3.4. Classification

Classification allows to transform the attributes characterizing the shapes into class membership (passage from the coding space to the decision space).

4. IMPLEMENTATION AND RESULTS

The goal of our work is to develop a cipher recognition system while relying on CNNs and SVM as a classifier. We developed a pre-trained model by adding the SVM classifier to the last layer of the CNN neural network, and then we added a set of images to the real input set.

4.1. Dataset

The **MNIST** database (**Modified National Institute of Standards and Technology** database) is a large collection of handwritten digits. It has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger NIST Special Database 3 (digits written by employees of the United States Census Bureau) and Special Database 1 (digits written by high school students) which contain monochrome images of handwritten digits [06].

The original black-and-white (bi-level) NIST images were normalized in size with digits centered in the middle to fit in a 20x20-pixel box while preserving their aspect ratio. The resulting images were centered in a 28x28 image keeping the aspect ratio of the image.

4.2. Augmentation Data

We adopted an image augmentation technique that consists of increasing the size of a training dataset by creating new modified versions of the images from the available training images by ensuring a variation of images that makes the model able to generalize what it has learned to new images and that strongly improves the performance of this model.

4.3. Libraries

We start by importing our libraries and dataset. The first library we import is Tensorflow [07]. It is an open source library, developed by Google and released in 2015, which is very popular in Machine Learning.

In Tensorflow, we import Keras, which is a programming interface and can decrease the development time of a neural network prototype by 30%). [08]

4.4. SVM (Support Vector Machine)

SVMs are a family of machine learning algorithms that can be used to solve classification, regression and anomaly detection problems. The SVM classifier is an algorithm that maximizes the margin between classes of the problem to be solved by a hyper plane and reduces the classification error.

The SVM classifier is a machine-learning algorithm that maximizes the margin between classes of the problem to be solved by a hyper plane and reduces the classification error. The calculation of this hyper plane is based on maximizing the margin between the closest learning examples that belong to different classes [09, 10, 11].

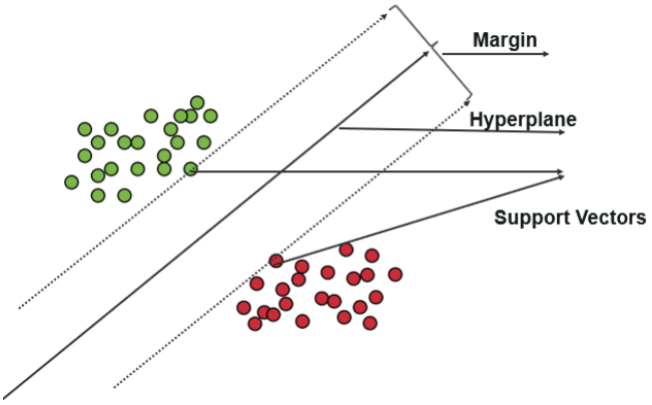


Figure 2. The classifier Support Vector Machine

4.5. Proposed Architecture

Figure 3 shows the proposed architecture principle.

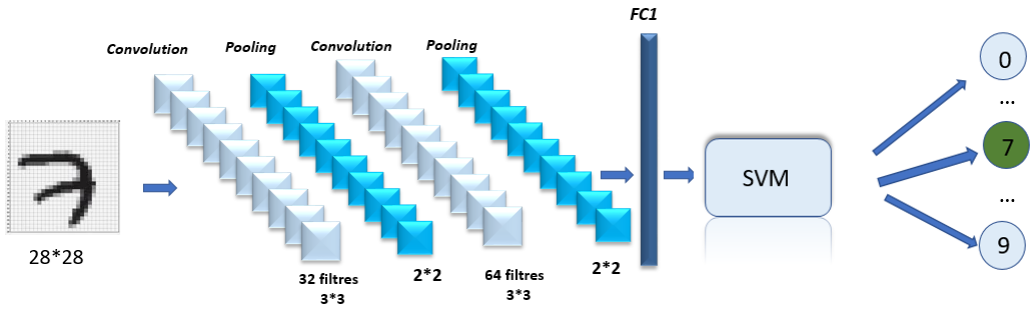


Figure 3. The proposed Architecture

The model that we present in figure 3 is composed of two convolution layers, two Maxpooling layers [12], and a fully connected layer, finally an output layer. The input image is of size 28*28, it passes to the first convolution layer which is composed of 32 filters of size 3*3, after this convolution, 32 feature maps of size 28*28 will be created. Then we apply Maxpooling to reduce the size of the image and the amount of parameters and calculation. At the output of this layer, we will have 32 feature maps of size 13*13. We repeat the same thing with the second layer of convolution this layer is composed of 64 filters, the activation function ELU [13] is applied always on each convolution. A Maxpooling layer is applied after the second convolution layer. At the output of this layer, we will have 64 feature maps of size 3*3.

In order not to fall into the problem of overlearning, we must use the dropout instruction, which is very effective for neural networks. It allows us to deactivate a number of neurons according to our configuration, which will also be used at the output of the fully connected layer [14]. The feature vector resulting from the convolutions has a dimension of 1600.

Finally, at the output layer, we added the SVM classifier by replacing the "Softmax" activation function by the linear activation function, which allows computing the probability distribution of the 10 classes (number of classes in the MNIST database).

4.6. Results and Discussion

After 100 iterations, the results in terms of accuracy and error are illustrated in the following.

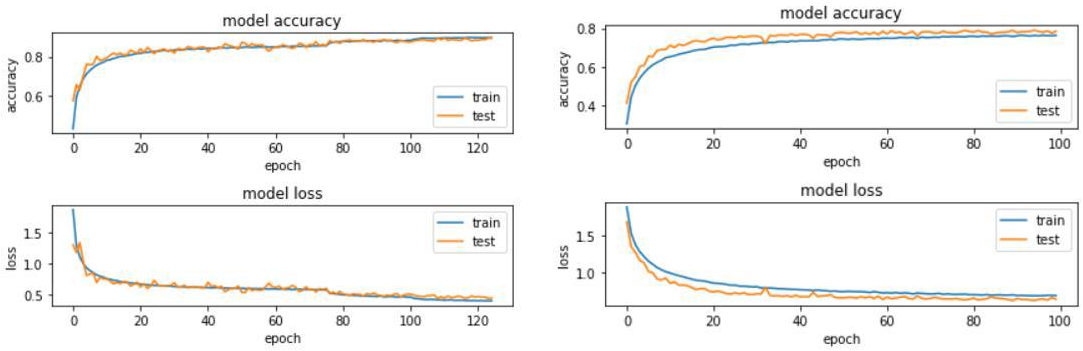


Figure 4. a) Proposed model b) Based model

From Figure 4, the learning and testing accuracy increases with the number of epochs, this reflects that at each epoch the model learns more information. On the other hand, we find that the learning and validation error decreases with increasing number of epochs.

Table 1. Comparison between the results obtained by the basic model and the proposed model

Model	accuracy	loss	Execution time	Number of iterations
Based model	98.970%	0.0571	00 :34 :47	100
Proposed model	99.129%	0.0386	00 :50 :21	100

Summarizing the results obtained in the table 1, we notice an error rate of 3.86%, which means that there are 386 images out of 10000 misclassified with an accuracy rate of 99.129%, which largely exceeds that of the basic model.

5. CONCLUSION

In the end, we developed an implementation for handwritten digit recognition using a classification approach based on convolutional neural network combined with SVM, so we established relationships between error and accuracy via graphs, and for this, we obtained very good results with an accuracy of 99.129% with an error that is almost zero.

The comparison between the results found showed that the increase of the data set, the use of the ELU function, the addition of a Maxpooling layer after the first convolution layer and so on are important players in obtaining better results.

REFERENCES

[1] http://yann.lecun.com/exdb/mnist/
[2] Nedjem. Eddine. Ayat, sélection de modèle automatique des machines à vecteurs de suports : application à la reconnaissance d'images de chiffres manuscrits, thèse doctorat en génie. P. H. D. Université du Québec, Montréal, Canada, 2004.
[3] Matougui. Farida, système interactif pour la reconnaissance des Caractères arabes Manuscrits, SIRCAM, Thèse de Magistère, C, D, T, A, 1990.
[4] Jean-Pierre – "reconnaissance de l'écriture manuscrite", Thèse, Université de Rennes 1.
[5] A. Gattal – "Segmentation automatique pour la reconnaissance numérique des chéques bancaires algériens", Thèse, Thèse de MAGISTER, centre Universitaire de Khanchela, 2011.
[6] https://www.nist.gov/srd/shop/special-database-catalog
[7] www.tensorflow.org
[8] https://keras.io/
[9] C.-L. Huang et C.-J. Wang – "A ga-based feature selection and parameters optimization for support vector machines", Expert Systems with applications 31 (2006), no. 2, p. 231–240.
[10] I. Guyon, J. Weston, S. Barnhill et V. Vapnik – "Gene selection for cancer classification using support vector machines", Machine learning 46 (2002), no. 1, p. 389–422.
[11] D. Abdelhakim – "La reconnaissance des chiffres manuscrits par les machines à vecteurs de support (svms)", Thèse, Thèse de Master, Université de Tébessa, 2011.
[12] P. Buyssens – "Fusion de différents modes de capture pour la reconnaissance du visage appliquée aux e_transactions", Thèse, Université de Caen, 2011
[13] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). arXiv preprint arXiv :1511.07289, 2015
[14] T. C. Faltemier – "Flexible and robust 3d face recognition", Thèse, Citeseer, 2007.

BLOCKCHAIN IN INSURANCE INDUSTRY: TURNING THREAT INTO INNOVATIVE OPPORTUNITIES

Wadnerson Boileau

USA

ABSTRACT

Insurance has been around for more than centuries. This risk mitigation strategy has been utilized in maritime commerce as early thousand years ago, where Asian merchant seafarers were pooling together their wares in collective funds to pay for damages of individual's capsized ship. In 2018, insurance industry made up 6% of global GDP while financial industry amounted to about 7-9% of the US GDP. In 2020, the industry net premiums written totaled $1.28 trillion, created 2.9 million jobs, and recorded $2.0 trillion investments. Despite of growing reform, the insurance market is dominated by intermediaries assisting people to match their insurance needs. While many predictions focused on artificial intelligence, cloud computing, blockchain stands out as the most disruptive technology that can change the driving forces underlying the global economy. We will focus on presenting blockchain use cases in insurance, demonstrating how the sector can turn blockchain threat into innovative opportunities.

KEYWORDS

Blockchain, insurance, risk management, innovation.

1. INTRODUCTION

Blockchain in recent years becomes one of the largest and most popular technology in the academic sector and financial industry. Even though blockchain appears to be new and speedily evolving technology, it has the potentiality to digitally transform processes inside and outside organizations and influences coordination as a mechanism to rule companies. Financial sector in many instances rely on blockchain to find durable solutions to some of the common problems faced in the insurance industry including decreasing operational cost and false claims, increasing efficiency and service delivery through technological innovations, and growing the customer base through trust. The paper aims to provide a clear-eyed view of how blockchain can be used in the insurance industry for process innovation. Throughout this paper, we focus on business use cases showing where the technology can bring efficiencies which enables the sector to turn the disruptive power of blockchain into opportunities.

2. LITERATURE REVIEW

2.1. Blockchain – How it Works

Blockchain technology is a publicly verifiable, shared, immutable distributed ledger used for recording the history of transactions. This is a chain of blocks that contains information inside a

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 07-15, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121702

block, and each block is associated with a hash of its previous and subsequent blocks to create a chain. It consists of nodes where each node maintains its local copy of the chain and relates to peer- to-peer(P2P) connections. Blockchain is a combination of tree existing technologies which are digital signature, cryptographic hash, and distributed consensus mechanisms. It is the basic infrastructure, or the underlying digital foundation that supports applications such as Bitcoin, Ethereum, Litecoin, etc. This emerging technology enhances the process of storing transactions and tracing assets in a cooperate network.

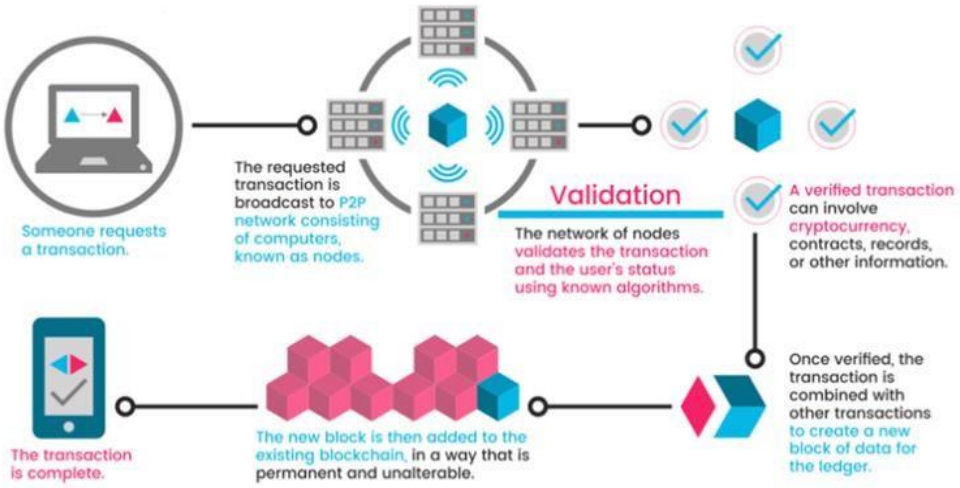


Figure 1. How blockchain technology works

2.2. Smart Contract

A conventional contract is an agreement between two or more parties that is enforceable by law; a smart contract is an agreement between two or more parties that lives on a blockchain and is enforceable by code. Permissioned blockchain are more suitable for data protection regulation compliance its abilities to restrict participants involved in the consensus mechanism on its network.

2.3. Immutability

In a blockchain network, once a transaction is fully processed, it cannot be altered. No falsification could be made because each transaction is unique and recorded solely in the digital ledger.

2.4. Distributed

All transactions can be seen on a distributed database; they can be validated by every partaking node contingent on the uniform rules (Martino, 2016). Cryptography and consensus algorithms are fundamentally utilized to validate transactions, update the ledger, keep the network and the ledger secure.

2.5. Distributed Ledger

In accounting, a ledger is simply a book or computer file that records transactions. Today, most of computing services that we utilize run on centralized networks, in which a central server stores and distributes data to other computers on the network called clients. In contrast, blockchain systems run on peer-to-peer (P2P) networks in which all nodes have equal status and concurrently operate as both client and server to one another. Every node participates in the consensus process saves a copy in real time of the ledger, which is why blockchain is referred to as a “single source of truth”.

2.6. Auditability

In the distributed ledger, every transaction is "validated and recorded with a timestamp”, this enables users to verify and trace the previous records through accessing any node in the distributed network. It improves the traceability and the transparency of the data stored in the blockchain.

2.7. Trust

Once a transaction has been completed in a blockchain network, the immutability of blockchain makes it nearly impossible to make changes, which increases confidence in data integrity and reduces opportunities for fraud.

3. CHALLENGES

Insurance industry encounters numerous challenges, and they keep growing over the time to become more consistent and complex. More recently, many insurance experts turn their attention to the potential of blockchain to address long-standing challenges related to the industry future and blockchain technology adoption. There are plenty of obstacles identified, but global regulations, trust and lack of accounting standards are eminent challenges for the world facing blockchain adoption [8].

3.1. Insurance industry challenges

Nowadays, when it comes to support, it’s not a secret for anyone in service industries that customer needs and expectations are increasing and getting higher than ever before. High customer expectations lead to new challenges and create needs for new standards of services by insurance companies. The central finding of this document is that blockchain solutions have the potential to increase efficiency and improve outcomes, that can force the sector to find out ways to address the following challenges: (1) Insurance organizations and reinsurers retaining fragmented records about insureds, assets, providers, and policyholders on various systems simultaneously, and the will of implementing blockchain based innovation is not an easy decision to pursue. (2) Insurance applications and systems lacking to provide robust security measures to secure participants information. (3) Lack of visibility when information is sharing between insurers, peers, providers, claimants. (4) Carriers collecting and assembling pertinent information and documents from multiple sources. (5) Multiple handoffs between systems contributors increasing time, cost, and risk of fraud. Currently, it is very challenging for a Certified Public Accountant

(CPA) to assess management’s accounting policies for digital assets and liabilities, which are presently not directly addressed in global financial reporting standards or in U.S. generally

accepted accounting principles. A financial audit involves an assessment that recorded transactions are supported by evidence that is relevant, reliable, objective, accurate, and verifiable. The nature of a transaction recorded in blockchain system may or may not provide proper audit evidence.

3.2. Challenges for Blockchain Adoption

Like every technology, there are upsides and downsides, blockchain technology comes with multiple form of challenges for its mass adoption. One of the major concerns in this adoption is in terms of security. Scalability of blockchain technology constitutes another major challenge for its adoption and implementation; transactions in blockchain system are validated through consensus mechanisms, the continuous replication, and the immutability of this technology lead to ever-growing amount of stored data. Industry professionals still want to see how well blockchain solution can perform in terms of integration with legacy systems to reduce total cost of its implementation. Other obstacle that needs to take into consideration is the lack of legal and regulatory framework since organizations have complex business rules and regulatory obligation to comply with.

3.3. Trust in Digital Economy

The key denominator of any economic or financial exchange is TRUST. Digital economy is an economy based on electronic goods and services and created by an electronic business model. Today behind each transaction that is taking place requires the intervention of a trusted third party, even to claim ownership of an asset we rely on government or central authorities to verify and confirm our property rights [6]. Since last decade, we have noticed a phenomenon called "Erosion of Trust" which is expressed as: distrust of central authority, the desire for freedom, the desire for privacy and anonymity, the distrust of intermediaries, and the distrust of corporations. In addition to this new phenomenon, there are various reasons why we may feel not comfortable to rely on third parties to provide these operations. First, and most obvious, are the fees that intermediaries charge for their services, which can be quite high. Third parties sometimes bring inefficiencies. Relying on third party also implies cybersecurity risks, as retaining sensitive data on centralized servers creates a single point of failure for bad actors. Finally, public confidence in financial establishments significantly deteriorated during the global financial crisis, and it may be more than mere coincidence that the Bitcoin protocol, which attempted to provide an alternative to the traditional financial system, was introduced in October 2008, as the global financial crisis was taking hold.

4. INSURANCE INDUSTRY AND THE DRIVERS

In recent years, many factors are shaping the future of the insurance market. Emerging technologies such as blockchain and artificial intelligence (AI), consumer expectations have become predominant drivers that have inflamed the competitive environment. According to Gartner.com, "by 2023 35% of enterprise blockchain applications will integrate with decentralized applications and services" [7]. The adoption of distributed ledger technologies by Payment networks along with decentralized finance (DeFi) applications are among of key drivers of the blockchain technology. The biggest concern remains the ability of blockchain implementation to eliminate the middlemen like brokers and agents sooner or later from the business process model of the insurance industry.

4.1. Emerging Technologies

The evolution of emerging technologies such as blockchain, artificial intelligence (AI), and machine learning offer a great opportunity for the current insurance system into a type of digital insurance platform.

This approach will lead to create new value, and new concept like “InsurTech” that stands for Insurance and Technology which is an insurance version of FinTech. The objective is to leverage technology power to increase productivity and efficiency in the insurance industry. To have an idea of how this technology could impact the global economy, on October 13th, 2020, London-UK, a report produced by PwC states Blockchain technology has the potential to boost global gross domestic product (GDP) by US$1.76 trillion over the next decade [5].

4.2. Customer Expectations

Based on various studies, customer expectations are highest that it has never been before. Insured, claimants, and other participants expect personalized and fast service at their finger. This new and ongoing trend of social and technology are shaking in many ways the regular business patterns in the sector. Insurance agencies need to ensure customers their information remains private and have adequate measures in place to secure while sharing data with other entities. Prospects like options” and they are generally looking for advises and recommendations from the industry’s Professionals. By establishing a consistent communication with their customers and assisting them with the shopping process, are the effective ways for insurers to elevate customer service experiences.

4.3. Innovation and Disruption

Innovation and disruption are frequently walk hand-to-hand, throughout this paper, these terms have more connection how the insurance industry is now developing and will continue to advance into future. Conventional insurance agencies and reinsurers are utilizing emerging technology to bring certain efficiencies into the industry which could potentially lead to higher financial gain or a more favorable competitive edge. It may also result in a competitive drawback due to their long-established business models and their high-value investments in infrastructure. Any variation to the current business model could be greatly challenging for some agencies since a new business model may necessitate a reduction of the sales of the insurer’s existing business. Despite of those pain points, new start-up companies can take these opportunities to build a unique or different business model designed to exploit weaknesses in old-fashioned insurance’s operations [9]. This advantage is not a guarantee for new start-up to survive in the marketplace. Other factors such as time, workforce mindset, customer satisfaction, and user acceptance play a role to their success or failure.

4.4. Key Players

Among Fintech initiatives, there are two major players capturing our attention:
B3i - Launched in 2016, Blockchain Insurance Industry Initiative (https://b3i.tech/) which is an international industry blockchain technology consortium that is owned by 21 major insurers and involves more than 40 companies as shareholders, customers, and community members, its main purpose is to improve successes in data exchange among insurance and reinsurance enterprises [2].

The Institutes RiskStream Collaborative (https://web.theinstitutes.org/riskstream-collaborative) is the risk management and insurance industry's largest enterprise-level blockchain consortium that connects experts and developers to advance insurance specific use cases. Founded in 2017 and located in Malvern, State of Pennsylvania (USA), The Institutes RiskStream Collaborative states its vision is to advance blockchain technologies and its capabilities to streamlining and bringing efficiencies in all areas of the risk management and insurance industries.

5. BLOCKCHAIN IN INSURANCE: USE CASES

Insurance entails a risk transfer mechanism that involves a financial institution(insurer) taking a potential risk from a customer in form of a loss or damage caused by events beyond the control of the customer (insured party) and in return the insured agrees to pay a fee called premium. The following use cases address improvements in insurance market operational functions as well as dealings with providers, intermediaries, claimants, insured and policyholders, thereby improving the customer experience, enhancing product value, and laying the groundwork for greater consumer choice in the financial industry. Blockchain technology can be implemented throughout the insurance sector and across many lines of business including:

1. Know-Your-Customer and Anti-Money Laundering (KYC-AML)
2. Loss adjustment / Claim management (Property and Casualty insurance)

5.1. Know-Your-Customer and Anti-Money Laundering (KYC-AML)

Know-Your-Customer is one of the benefits of the blockchain technology which providing a secure digital infrastructure for verifying identity. Insurance companies, reinsurers, and brokers frequently conduct auditing activities to know their customers, it is proactive approach on preventing money laundering usually participates multiple entities like individuals and legal personnel [4]. In May 2017, Singapore government along with several financial institutions have launched a successful KYC project leveraging blockchain technology to provide a registry for customer due diligence and verifying a valid identity in order to reduce money laundering activities. The following use case shows how blockchain can help to reshape outdated KYC processes by allowing for the effective outsourcing and decentralizing of personal data, while also enabling the owner of the data to maintain full control over their data. Blockchain solution will establish a tamper-proof repository using mutual distributed ledger technology which allows several parties to add, certify and exchange KYC and AML documentation. By doing so, customer’s data can be securely shared between organizations, providers, and customers.

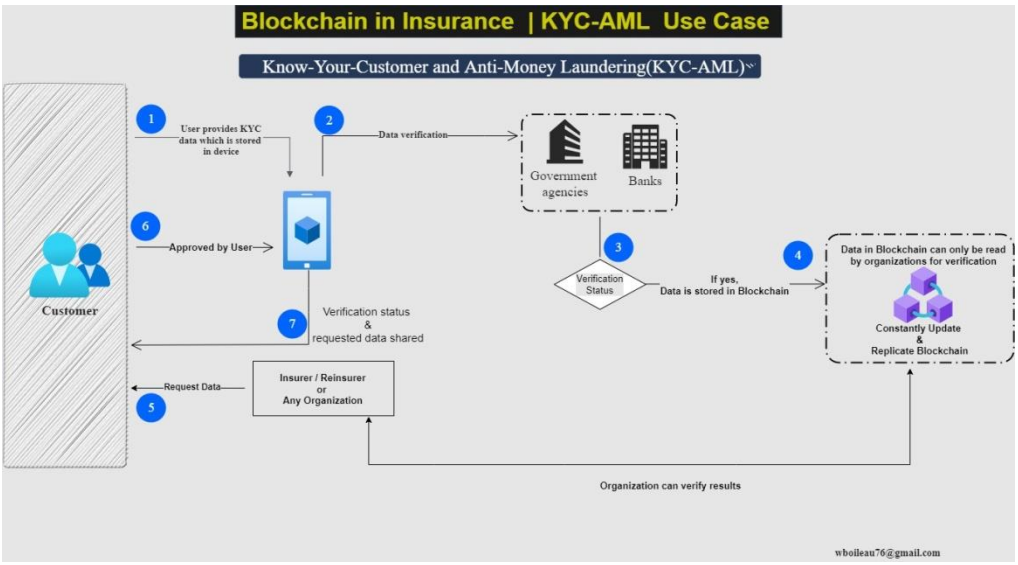


Figure 2. Know Your Customer – Anti-Money Laundering use case

5.2. Loss Adjustment/Claim Management

Majorly, there are two types of insurance product: life insurance and general insurance. We encounter two main categories of life insurance – permanent and term [3]. General insurance in other hand deals more with other form insurance related to valuables including home, car, goods, and other hazard like fire. It is divided into 6 major segments namely Motor insurance, Health insurance, Combined, Comprehensive and Package policies, Property insurance, Pecuniary insurance, and Casualty insurance.

Our use case is focusing on property and casualty insurance also known as P&C insurance. Property insurance and casualty insurance are types of coverage that help protect you and your property while Casualty insurance provides liability coverage to protect if you're found legally responsible for an accident that causes injuries to another individual or damage to another entity's belongings. One of the big challenges with insurers is collecting the proper data to assess and process claims. Most insurance systems are an errorprone process using excessive human involvement in terms of manual data entry, paper works and coordination between different parties. Blockchain technology can be utilized to automate claim applications and processing. In the event a natural disaster is confirmed by the mandated state institution with the magnitude and level of impact, a smart contract can be triggered to make payments to the affected communities.

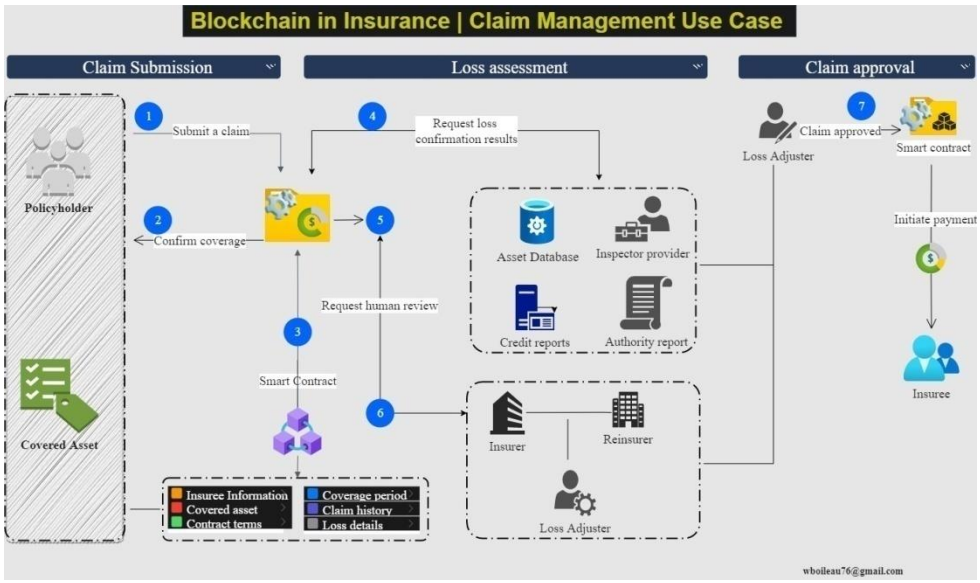


Figure 3. Claim management use case

6. CONCLUSIONS

As has been noted, in section 3, we provide challenges for the insurance industry based on blockchain technology, section 5 explains how blockchain can bring value and enable innovation in the value chain of the industry invariable giving rise to a new business model. As has been demonstrated, insurance industry can leverage blockchain technology to enrich crucial processes like claim submission and processing, antimoney laundering, fraud detection and prevention. Firstly, while blockchain technology is making its baby steps, there are several promising use cases and applications for it in the insurance industry. We have already seen areas blockchain can lower operational cost and automate redundant process. Secondly, from the industry perspective, we discovered there is a great need for insurance companies to align around standards and processes within blockchain. While blockchain technology is developing better tools and framework for this market to collaborating and sharing data, insurance agencies themselves and other entities part of this eco-system like regulators, financial institutions, and government entities must be willing to work with each other. The technology must go further in terms of development to address privacy, security specially about public blockchain where everyone has access to each transaction in the ledger. Finally, since the insurance is highly regulated sector, legislators need to play their partition by providing legal and regulatory frameworks for Insurtech, in addition laws and procedures currently in place, there is an urgent need to expand and furnish a clear guidance for blockchain technology to succeed.

REFERENCES

[1] https://www.iii.org/publications/a-firm-foundation-how-insurance-supports-theeconomy/introduction/insurance-industry-at-a-glance

[2] Nicholson, E Jack, Cole, Cassandra and McCullough, Kathleen - Challenges for the Insurance Industry in the Future. National Association of Insurance Commissioners (NAIC), Journal of Insurance Regulation, Vol 38 No. 6, 2019.

[3] Akande, Adefisayo - Disruptive Power of Blockchain on the Insurance Industry. Tatu University Library 2018.
[4] Acheampong, Anokye AMPONSAH, WEYORI, Benjamin Asubam & ADEKOYA, Adebayo Felix-Blockchain in Insurance: Exploratory Analysis of Prospects and Threats. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 1, 2021.
[5] Blockchain technologies couldboost the global economy US$1.76 trillion by 2030 (pwc.com)
[6] Michael Pisa and Matt Juden. 2017. "Blockchain and Economic Development: Hype vs. Reality." CGD Policy Paper. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/blockchain-andeconomic-development-hype-vs-reality
[7] Pwc, "Insurance 2020:Turning change into opportunity," *PwC Insur. 2020*, no. January, p. 24, 2012.
[8] https://blogs.gartner.com/avivah-litan/2021/07/14/hype-cycle-for-blockchain-2021-more-action-than-hype/
[9] R. Hans, H. Zuber, A. Rizk, and R. Steinmetz, "Blockchain and Smart Contracts: Disruptive Technologies for the Insurance Market," *Proc. Twenty-third Am. Conf. Inf. Syst. (AMCIS 2017)*, no. August, pp. 1–10, 2017.
[10] H. Kim, M. Mehar, ―Blockchain in Commercial Insurance: Achieving and Learning Towards Insurance That Keeps Pace in a Digitally Transformed Business Landscape, ‖ SSRN Electron J 2019.

AUTHOR

Wadnerson Boileau has a Master of Science in Business Analytics & Information Systems (2021) from University of South Florida (USF), and a bachelor's degree in Electronic Engineering from State University of Haiti (Université d'Etat d'Haïti - 2001). He started his career as Analyst Programmer/Developer where he built software for small businesses. In 2003, he became SAS Certified while implementing SAS Times Series Project for Central Bank of Haiti. He joined an Oracle Partner Company providing support in processes and documentation. In 2014, after getting his Florida Licensed insurance, he started working with multiple Fortune 500 companies as Financial Consultant. In 2019, he earned two Blockchain certifications- Certified Blockchain Business Foundations and Certified Blockchain Solution Architect.

MEASUREMENT STUDY ON 5G NSA ARCHITECTURE OVER FADING CHANNEL

Bruno S. da Silva and Iury da S. Batalha

Sidia Institute of Science and Technology, Manaus, Brazil

ABSTRACT

The 5G NR network with the Non-Standalone (NSA) architecture aims to advance with regard to throughput. When compared to fourth-generation mobile communication (4G LTE), the 5G has a higher data exchange capability through the gNB and the UE (User Equipment). For evaluation and optimization, it is necessary to carry out practical studies on the behaviour of the system in different environmental conditions, subject to attenuation processes, such as large-scale fading (Shading) and small-scale fading (Multipath propagation). This work has analysed the effect of the MCS (Modulation and Coding Scheme) variation on Throughput/BLER for, initially, a channel degraded by default AWGN, then the analysis extends to the multipath fading effect, which emulates more realistically a mobile communication network. The analysis confirmed the need for robust decision process algorithms in terms of MCS switching to maintain adequate data rates according to the requirement of each scenario with specific QoS (Quality of service), considering both 64 QAM and 256 QAM. The throughput degradation effect was more evident in higher-order modulations due to the higher probability of error inherent in the symbol arrangement. This study can be a key for understanding and developing huge modulation and coding schemes for fifth generation communications.

KEYWORDS

5G, NSA, Fading, MCS, Throughput, Modulation, Coding scheme, BLER, Signal-Noise Ratio.

1. INTRODUCTION

The exchange of data at high transmission rates is a fundamental requirement of today's mobile networks [1], [2], encompassing ultra-high resolution video streaming, online gaming, among other services that require high demand for both downlink and uplink. The fifth-generation mobile technology, 5G New Radio (5G NR), has been widely implemented around the world to meet new requirements, naturally imposed by the advancement of telecommunications services. 5G emerges as the foundation for Enhanced Mobile Broadband (eMBB), Massive Machine-Type Communication (mMTC) and Ultra Reliable Low Latency Communication (URLLC) scenarios [3]. The initial implementation of the fifth generation network has been established initially following the NSA (Non-Standalone) architecture, which makes use of the LTE network as an interface for connecting 5G devices, adding greater throughput potential to the network [4]. This solution aims to save time and reduce costs for operators, since the entire LTE architecture already deployed serves as the core for the New Radio system. Fig. 1 illustrates the 5G Non-standalone (5G NSA) connection.

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 17-27, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121703

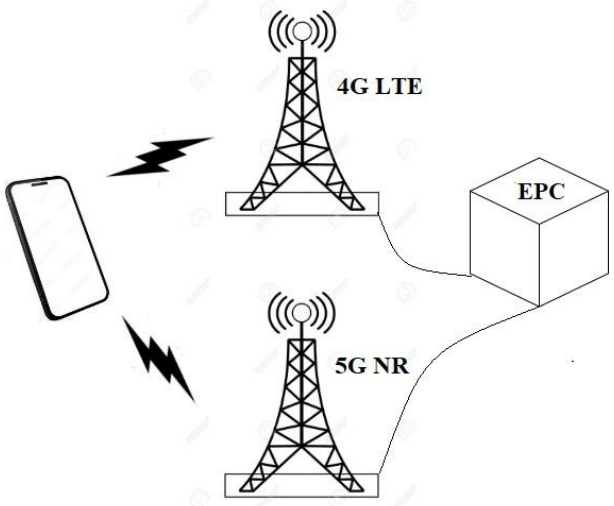


Figure 1. NSA Architecture

As standardized by the 3GPP (3rd Generation Partnership Project), in this type of configuration, the device (User Equipment) connects to both the LTE and NR networks, in the so-called Dual Connectivity (DC), in order to improve the total network coverage and increase the data rate. This mechanism is accomplished through the carry aggregation (CA) technique, in which two or more frequency bands are allocated on a single data channel to improve the network capacity. When Dual Connectivity is realized by joining an LTE band (sub 6 GHz) and an NR FR2 band (mmWave band), transmission rates can reach Multi-Gbps, combining the coverage power of the LTE band (lower frequency) and the high bandwidth capability FR2 (up to 400 MHz) [5], [6].

A more realistic analysis of a mobile wireless communication system is given by considering the Fading effect, which characterizes attenuations in the power level received by the mobile device as a consequence of the temporal variation inherent to mobility, causing harmful effects to the data transmission rate (throughput) due to the increase in block error rate (BLER) [7]. This work analyses the performance of a 5G NSA network under the effect of multi-path fading, characterized by an urban propagation scenario, varying basic parameters such as SNR levels and modulation and coding schemes (MCS). The data are obtained from measurements in a professional 5G laboratory, that provides a controlled environment, which certifies and validates the analysis.

The work continues with the following division: Section II makes a brief theoretical discussion about AWGN and Fading and the effect on data transmission. Section III presents the lab setup and the measurement methodology. The collected data is analysed, compared and discussed in Section IV. Section V provides concluding discussions and perspectives on future work.

2. WIRELESS COMMUNICATION AND FADING CHANNELS

In wireless and mobile electromagnetic communication, the transmitted signal suffers several external interferences as it propagates through the medium. A basic and constant effect is the noise addition, which is well modelled using the AWGN model, which adds noise equally to all frequency components of the signal, simulating various random processes seen in nature. Another mechanism that is part of the propagation of electromagnetic waves is Fading, which considers the effect of multiple path copies of the originally transmitted signal [8].

2.1. AWGN Channel

The AWGN is a basic model of thermal noise that occurs naturally as the signal travels through the air. Its addition causes uncertainties in decision-making of digital signal demodulation. It is considered *white* because it equally affects the signal in its band (flat power spectral density). It is *Gaussian*, because its samples have a Gaussian distribution. Figure 2 compares the constellations of a random 1000-bit stream sent by 8-QPSK and 4-QAM modulated signals, showing the results with the presence of AWGN (Fig.2 (b) and (d) The simulations were performed using Python language.

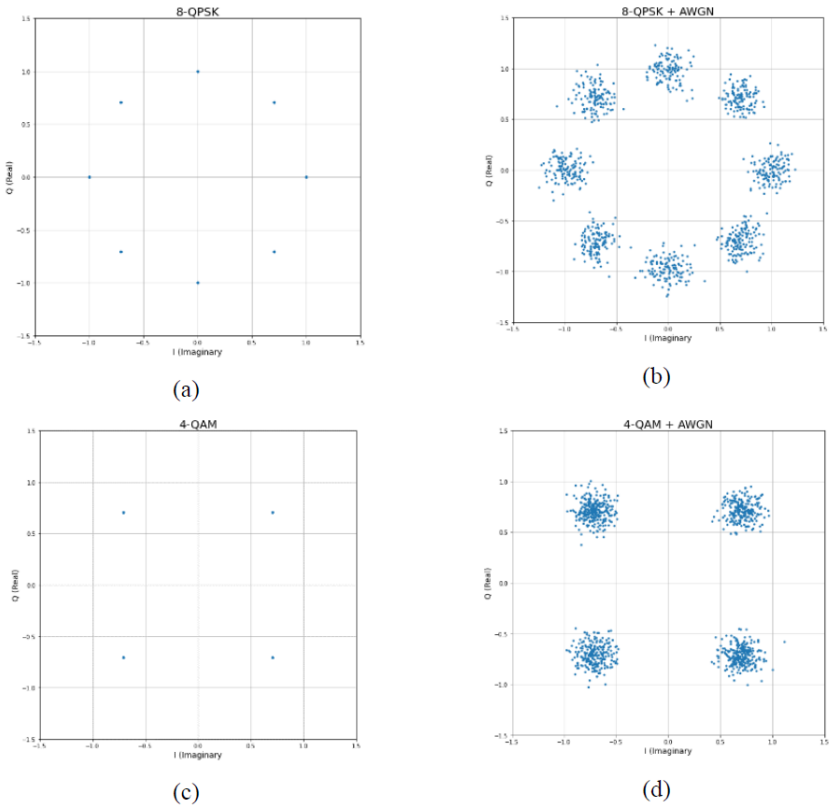


Figure 2. Constellations under AWGN

When analysing figures 2(b) and 2(d), it can be seen that the uncertainty in the bit detection decision is increased as noise is added to the transmitted signal. High power of additive white Gaussian noise compromise data exchange, leading to low throughput efficiency. This situation can be minimized by increasing the transmission power or by using lower-order modulations, which have a greater spacing between the elements of their constellation, facilitating the correct bits demodulation.

2.2. Fading Channel

The receiver obtains and demodulates multiple copies of the signal sent by the transmitter. This process occurs by reason of the complexity of the propagation environment, which is time-varying and with the presence of several reflective and diffracting objects. Such copies can be received with relative time delays and with different spatial orientations (arrival angles at the receiver) and amplitudes. As a consequence, there will be random modulations in frequency, due to the Doppler Effect of each signal variant, time dispersion caused by the multipath propagation

components and rapid variations (fluctuations) of the intensity of the final received signal, as this is the result of constructive and destructive combinations of copies [9], as illustrated in Fig. 3.

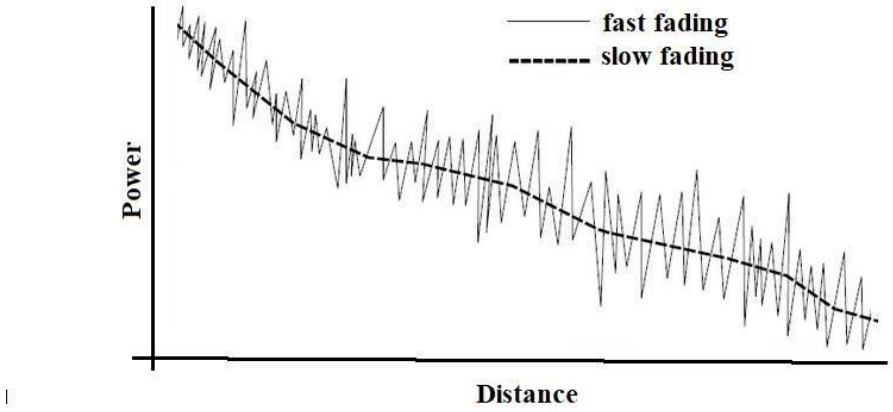


Figure 3. Fading effect

In the literature, statistical approaches are found that aim to model the fading effect on the channel according to propagation scenario. For mobile communications, the most adopted model is the Rayleigh model, applied to estimate propagation from multipath, when there is no direct line of sight (LOS) between transmitter and receiver. In this case, the response of the channel g on the multipath effect considering L paths is given according to Equation 1.

$$g = \sum_{i=1}^{L} \sqrt{\alpha_i}\, e^{\frac{-j2\pi(d_i - d)}{\lambda}} \qquad (1)$$

Where α is the channel gain of path i and the exponential term refers to the phase shift of each path. When L is large, a statistical model must be implemented, because of the random channel gains and phase shifts. As the variables of Equation 1 are independent and identically distributed, a Gaussian distribution is a good approximation, according to the Central Limit Theorem, as shown in Equation 2.

$$g \approx N_c(0, \beta) \qquad (2)$$

The N term indicates a complex distribution with mean equal to zero and variance β. The absolute value |g| has a Rayleigh distribution. The Rayleigh fading model can be applied to analyse radio propagation with a statistical basis. It operates best under conditions when there is no dominant signal (direct line of sight between transmitter and receiver). In many instances, mobile phones being used in a dense urban environment fall into this category.

3. SETUP AND MEASUREMENT METHODOLOGY

3.1. Setup

The Anritsu Measurement Setup used in this work makes it possible analysis with actual parameters standardized by 3GPP. By varying specific system variables, important KPIs (Key Performance Indicators) of the 5G can be verified [10]. The Anritsu module MT8000A builds and controls the 5G network and emulates the Fading effect over the NR, supported by a Server loaded with many fading profiles (3GPP standard). As a NSA architecture is been considered for

this study, the 4G base stations and EPC (Evolved Packet Core) are controlled by a second MT8000A module. The RTD (Rapid Test Design) software is the graphic interface where the campaigns are developed and executed. The setup design is illustrated in Fig. 4.

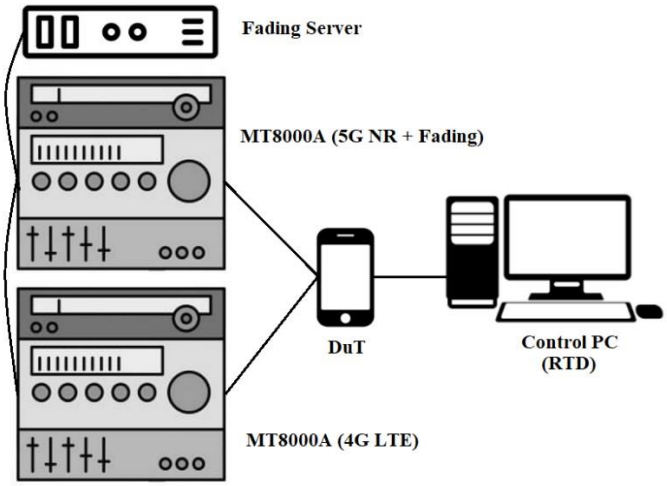


Figure 4. Measurement Setup

3.2. Methodology

In order to obtain samples, a measurement methodology consisting of a volume of data collected for statistical validation [11], then the treatment of such data is done. In this context, measurement campaigns were carried out, considering a NSA architecture. 5G NR supports QPSK, 16-QAM, 64-QAM and 256-QAM modulations, for this study, the last two are considered due to greater throughput capabilities. The campaign was defined in LTE band 1 (2100 MHz) with 20 MHz bandwidth and the n78 5G band (3500 MHz), considering 64/256-QAM modulation. The objective was to analyse the MCS variation effect in the 5G NR cell over throughput, with 100 MHz bandwidth, 30 kHz subcarrier spacing, 273 RBs (Resources Blocks) allocated and with a SISO scheme. Table I summarizes the parameters analysed in this article.

TABLE I. Campaign Parameters

Parameter	Value
Architecture	NSA
NR band/bandwidth	band n78 (3500 MHz)/100 MHz
LTE band/bandwidth	band 1 (2100 MHz)/20 MHz 11
LTE MCS	(64-QAM/fixed)
Resource Blocks	273
Modulation	64/256-QAM SISO
Transmission Mode	(1x1)
Fading profile	Suburban (static)

Data was collected according to the received power variation (SNR), from a bad quality channel (12 dB) to a good one (30 dB).

4. RESULTS ANALYSIS

Based on the measurement campaigns and the data obtained, it is possible to determine some analysis on the throughput metrics. There is a range of received power according to the network configurations, the objective is to measure how Throughput/BLER behave with the channel degradation, which can be used for estimating a threshold for specific QoS (Quality of Service).

Block Error Rate (BLER) is defined as the number of erroneous received blocks divided by the total number of blocks that was transmitted. The target BLER is typically about 10 percent for many application scenarios. If the BLER is larger than 10 percent (due to signal losses or interferences), the MCS must be switched to a lower value, providing a reliability communication, by making the BLER value stable (equal to or less than 10 per cent, in accordance with the application requirements).

The MCS is related to modulation and coding scheme. Digital modulation specifies how many bits can be transmitted in a symbol, which is defined as Resource Element (RE) and MCS defines how many useful bits can be transmitted per Resource Element. The higher order is the modulation, the more data is able to be transmitted, in the other hand, higher order modulation increases the probability of error. However, a Coding scheme is implemented according to the channel conditions, by adding redundant bits on data stream to enable error detection. A trade-off is that there is less useful bits in the transmitted blocks.

There are about 32 MCS Indexes (0 to 31). MCS Index 29,30 and 31 are reserved and used for re-transmission. The 3GPP Specification 38.214 has given three tables for MCS: 64-QAM Table, 256-QAM Table and Low Spectral Efficiency (Low SE) 64-QAM Table. The gNB instructs the UE to select a specific MCS table using a combination of RRC and Physical layer signalling.

64-QAM table may be used when gNB or UE does not support 256 QAM or in poor radio condition, where 256-QAM table is not decoded successfully. The 256-QAM table is allocated in good radio conditions. A specific table (Low SE 64-QAM) is used considering the Ultra reliable and Low-latency communications (URLLC) category, when applications need reliable data transfer, by reducing coding rate (increasing channel coding redundancy).

For this article, the campaigns were carried out considering the 256-QAM Table, which has 28 (0 to 27) defined indexes.

4.1. Analysis for AWGN Channel

First analysis is considered by applying the AWGN effect to the signal, emulating many random processes that occur in nature, such as the constant thermal noise, which is added to the signal in the propagation phenomenon by the motion of charged particles in conducted media. The AWGN effect is a good reference for the next section (Section B), where Fading will be introduced in the transmitted signal. Table II presents the statistical metrics related to measurements of Block error rate (BLER), MCS and SNR for campaigns with 64-QAM and 256-QAM modulations.

TABLE II. BLER *versus* MCS under AWGN

64-QAM		
MCS	BLER (SNR 15 dB)	BLER (SNR 25 dB)
11	0%	0%
15	20%	0%
19	50%	4%
256-QAM		
MCS	BLER (SNR 15 dB)	BLER (SNR 25 dB)
20	50%	3%
24	65%	32%
27	70%	50%

From Table II we can see that a higher order modulation produces higher BLER, mainly when a bad channel condition is considered. Note that a lower order modulation (64-QAM), due to the greater spacing between symbols, it provides a lower block error rate when compared to a higher order modulation (256-QAM), assuming similar channel qualities. This expected effect is better illustrated in Figure 5, which considers Throughput capability values (%) according to SNR variation for 64-QAM modulation.

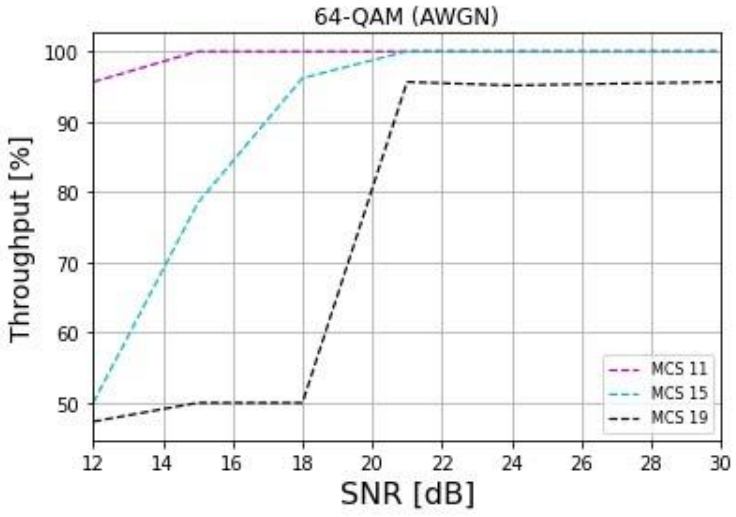


Figure 5. Throughput efficiency for 64-QAM (AWGN)

Note that throughput efficiency is directly related to the channel condition and the MCS index for the downlink. In this case, considering a low MCS (MCS 11), even under severe channel noise conditions, a throughput rate of more than 95 % can be achieved. The better the quality of channel, the less redundant bits need to be add for avoiding error detection.

As the SNR value increases, the probability of error reduces. Therefore, in good SNR conditions, a higher-order modulation can be scheduled for the user, increasing the throughput data rate.

Figure 6 shows the throughput behaviour for a 5G NSA network operating in 256-QAM modulation over an AWGN channel.

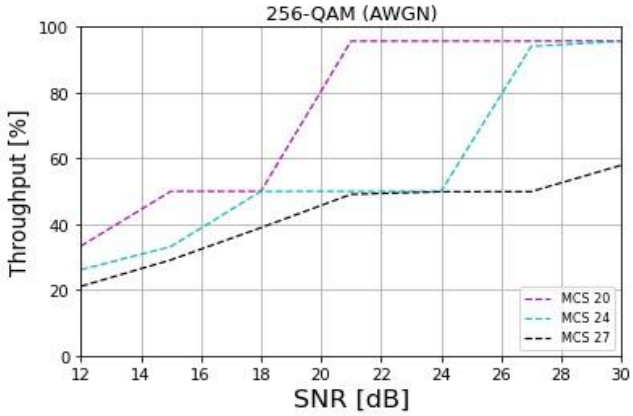


Figure 6. Throughput efficiency for 256-QAM (AWGN)

By analysing Fig.5 and Fig. 6, it can be noticed that the 256-QAM modulation, even providing a better throughput rate, is more susceptible to errors, when compared to the 64-QAM modulation scheme. For 256-QAM, even with good channel condition (SNR > 25), the MCS 27 is unable to deliver an acceptable BLER value (approximately 40%, in this case), which means that the coding scheme must be more robust, by allocating more redundant bits on the data stream to improve the throughput efficiency.

Next section the Fading effect will be studied over a 5G NSA network. The throughput response is analysed over the MCS variation.

4.2. Analysis for Fading Channel

The addition of Fading on the transmitted signal increases the error probability due to the small scale power oscillations. This causes poor performance in a communication system because it can result in a loss of signal power without reducing the noise power. This signal loss can be over a part of or all of the signal bandwidth. This can cause problems such as phase distortions and inter symbol interference in data transmission. The consequence is a higher BLER compared with the analysis made in Section A. This section shows the strong degradation over the signal considering an suburban fading profile.

Table III presents the campaign results for a 5G signal weakened for fading losses for 64QAM/256-QAM modulations and variable coding scheme (MCS).

TABLE III. BLER *versus* MCS under AWGN + Fading

64-QAM		
MCS	BLER (SNR 15 dB)	BLER (SNR 25 dB)
11	7.5%	0%
15	50%	1%
19	53.5%	24%

256-QAM		
MCS	BLER (SNR 15 dB)	BLER (SNR 25 dB)
20	61%	20%
24	70%	50%
27	77.5%	50.5%

By observing Table III it is clear the degradation of the transmission compared to the results presented in Table II, where only the AWGN was assumed.

Figure 7 shows the throughput rate efficiency over the fading channel for the 64-QAM modulation. While in the AWGN case (Figure 5). with SNR equal to 21 dB the 10 per cent BLER criteria was achieved, when fading was added the criteria is achieved only for about SNR around 30 dB. These values make clear the strong effect of fading over the signal propagation.

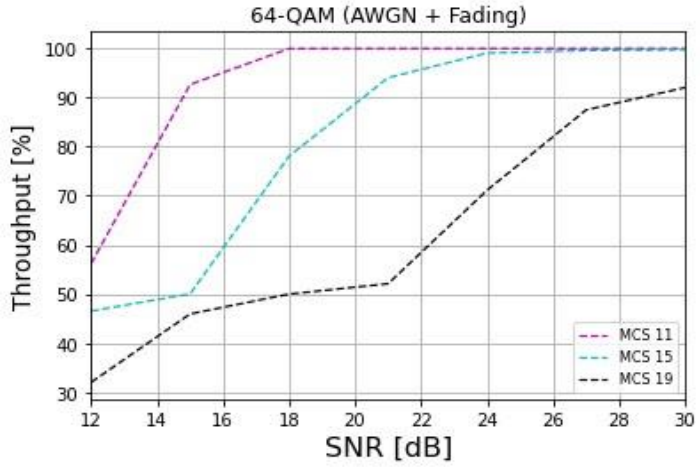


Figure 7. Throughput efficiency for 64-QAM (AWGN + Fading)

Similar to what was done in Section 4.1, campaign results for a higher-order modulation were analysed. The MCS variation comparison is illustrated in Figure 8. Because of its dense bits arrangement, the 256-QAM modulation over the considered fading must operate in very good SNR for providing an adequate BLER. Thus, for achieving a more efficient throughput, the MCS must be switched for a lower value regarding an acceptable BLER, according to the required QoS.

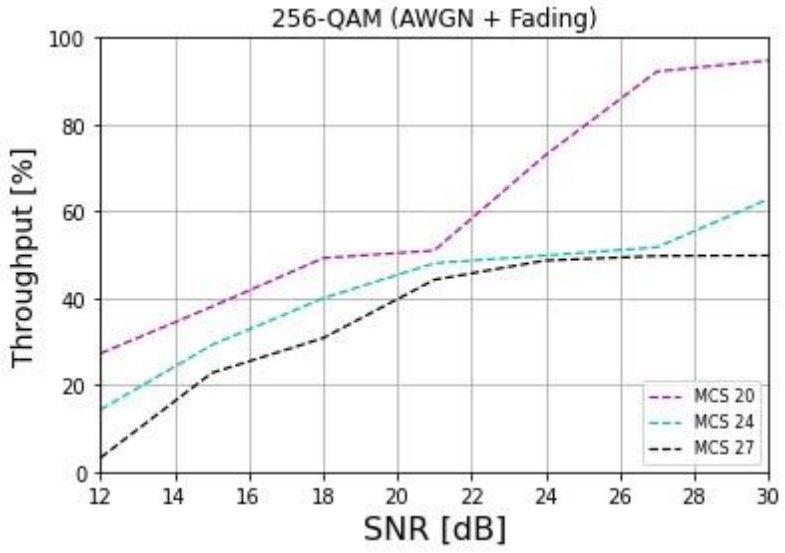


Figure 8. Throughput efficiency for 256-QAM (AWGN + Fading)

5. CONCLUSIONS

The present work aimed to study laboratory measurements about the influence of modulation and coding scheme over the throughput of an 5G NSA network, considering a theoretical scenario, with the effect of AWGN, and further analysis, in which effect of fading was added to the signal. A great difference was observed in the degradation of the signal even in good condition for the reception. Therefore, an efficient and well-modelled dynamically MCS switching is necessary for the better functioning of the network. Although generally, a BLER of 10\% or less is a target, certain conditions can show good functionality under higher values. It has been observed that users in good radio conditions or with high traffic utilization perform better with lower BLER Targets. Nevertheless, users in poor radio conditions or with small packets can perform satisfactorily with higher BLER. To deal with that, new MCS allocation models must be developed to meet specific throughput demands, according to each scenario.

In conclusion, for future work, other frequency bands in 5G, including the FR2 Band (mmWave band) will be studied in order to understand the behaviour of the network under fading and how strategies can be applied to optimize the system, such as the implementation of MIMO (Multiple Input Multiple Output) and proposals of MCS allocation algorithms according to specific applications.

ACKNOWLEDGEMENTS

The authors would like to thank Samsung Electronics da Amazônia Ltda. for partially supporting this work.

REFERENCES

[1] ITU-R M. 2412-0, “Guidelines for evaluation of radio interface technologies for IMT-2020,” Oct. 2017.
[2] 3GPP TR 38.912 version 15.0.0 Release 15.

[3] C. H. T. Arteaga, A. Ordoñez and O. M. C. Rendon, "Scalability and Performance Analysis in 5G Core Network Slicing," in IEEE Access, vol. 8, pp. 142086-142100, 2020, doi: 10.1109/ACCESS.2020.3013597.
[4] A. El Rhayour and T. Mazri, "5G Architecture: Deployment scenarios and options," 2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 2019, pp. 1-6, doi: 10.1109/ISAECT47714.2019.9069723.
[5] A. Benjebbour, et al., "5G radio access technology," NTT DOCOMO Technical Journal, vol. 17, no. 4, pp. 16-28, Apr. 2016.
[6] R. Mohamed, S. Zemouri and C. Verikoukis, "Performance Evaluation and Comparison between SA and NSA 5G Networks in Indoor Environment," 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), 2021, pp. 112-116, doi: 10.1109/MeditCom49071.2021.9647621.
[7] T. Lam, "Fading model for wireless 5G mmW transmission," 2016 IEEE International Symposium on Antennas and Propagation (APSURSI), 2016, pp. 2067-2068, doi: 10.1109/APS.2016.7696740.
[8] L. Gopal and M. L. Sim, "Performance Analysis of Signal-to-Noise Ratio Estimators in AWGN and Fading Channels," 2008 6th National Conference on Telecommunication Technologies and 2008 2nd Malaysia Conference on Photonics, 2008, pp. 300-304, doi: 10.1109/NCTT.2008.4814291.
[9] V. Vasić, "Performance of 5G New Radio Low-Density Parity-Check Codes in Nakagami-m fading Channel," 2021 29th Telecommunications Forum (TELFOR), 2021, pp. 1-4, doi: 10.1109/TELFOR52709.2021.9653235.
[10] R. Tafazolli, "IET 5G Radio Technology Seminar," 5G Radio Technology Seminar. Exploring Technical Challenges in the Emerging 5G Ecosystem, 2015, pp. 1-14, doi: 10.1049/ic.2015.0027.
[11] Y. -J. Cho, Y. -S. Kim, S. Kim, D. Sim, D. Kwak and J. Lee, "AI-Enabled Wireless KPI Monitoring and Diagnosis System for 5G Cellular Networks," 2019 International Conference on Information and Communication Technology Convergence (ICTC), 2019, pp. 899-901, doi: 10.1109/ICTC46691.2019.8939988.

AUTHORS

Bruno Sátiro da Silva is PhD in Applied Electromagnetism in the University Federal of Rio Grande do Norte. His research interests cover 5G communications, frequency selective surfaces, antennas and electromagnetic propagation. R&D engineer on Sidia/Samsung Brazil, focused on 4G/5G protocols.

Iury da Silva Batalha. His current research interests include electrical engineering and computing, especially in the areas of telecommunications and applied computing, working on the following topics: Mobile networks, propagation, IoT and bio-inspired algorithms. Researcher linked to PPGEE/UFPA, SIMA and INCT, focused on research on IoT, propagation and 5G protocol.

COMPUTERBANK: A COMMUNITY-BASED COMPUTER DONATION PLATFORM USING MACHINE LEARNING AND NFT

Eric Gong[1] and Yu Sun[2]

[1]University High School, 4771 Campus Drive, Irvine, CA 92612
[2]California State Polytechnic University, Pomona, CA, 91768

ABSTRACT

Many people donate money to fund organizations, but very rarely do those donors have information about where those donations go. Donation platforms are both non-transparent and also leave a large portion of potential donors unnoticed: gamers [1]. This paper explores the concept of utilizing blockchain technology and its existence as a web3 token-based platform in order to provide transparency for donation routes, showing donors and other companies exactly where donations are coming from and where that money is going. Our application utilizes HTTP requests in order to greatly increase compatibility, and also uses multiple private key encryptions in order to ensure that any user data or information and monetary transactions are kept secure and private [2].

KEYWORDS

web3, donation, platform, decentralized.

1. INTRODUCTION

In recent years, the popularity of cryptocurrency has soared through the roof due to new accessibility and promises of striking it rich [3]. According to a study done by Insider Intelligence, around 34 million US adults own cryptocurrency [4]. Nowadays, it would be rare to find a person who has yet to hear of cryptocurrency. Moreover, in 2020, after the advent of the Corona virus pandemic as well as the racial justice movements of the time, Americans gave a record $471 billion in donations. Video games are played by hundreds of millions around the world, from the gamers who'll dedicate hours at a time to their hobby, to people who might just play on their phone every once in a while to pass a train journey. Video games are big business and have the potential to offer a lot of support to good causes [5]. According to a study done by the Charities Aid Foundation, around 58% of players were interested in donating while playing; 59% would be more likely to pay to remove adverts if some of the cost went to charity; and 63% would use funds from online wallets to donate [6]. Despite such promising numbers, the industry of video games has surprisingly not overlapped too much with the idea of donations. Instead of utilizing the traditional methods of campaigning our platform uses online wallets and the large plethora of gamers willing to donate in order to make money-raising much quicker and more transparent.

Existing donation platforms exist, sites such as GoFundMe and DonorBox allow for patrons from all over the world to donate money to their favorite charities. They allow for organizations to create a fundraising campaign, with a description and the option for anyone to donate. The main

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 29-42, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121704

way that they gain traction is through affiliate links and the organization's own efforts to spread word regarding their fundraising. However, these websites require donors to go out of their way to find the website and then proceed to donate, which limits the scale of their reach. Another issue with this form of donation is that the proceeds are extremely hard to track. Many organizations can create a fundraising event and use the donations towards something else, with little to no transparency on where the money is actually going.

Our tool is a web3 token-based gaming donation platform. Because it is a dApp, it runs on a blockchain of computers and is free from the influence of a single authority [7]. In comparison to others. In comparison to other donation platforms who operate under one site owner and spread through word of mouth, Donahub's API is integrated into distributed video games and takes donations from a wide variety of players. A strong feature of using an online wallet for the organization to receive funds is that it offers transparency in where the money is going. By tracking the organization's wallet ID and its transactions, it guarantees that the organization is actually receiving the money and also spending it on the right things. Another feature that we have is the ability for an admin to determine how the donations will be spread among various organizations, allowing a single donation to benefit more than one organization.

In practice, Donahub demonstrates how the web3 token-based structured donation platform increases the effectiveness of donations. First, we have complete transparency, showing users the recent transactions list including addresses for easier tracking. Donahub also utilizes HTTP requests in order to make new donations, making it highly compatible with most games and coding languages [8]. Donations can be made from multiple sources as long as HTTP requests are being sent, letting you make donations from web browsers, games, and any other source. The app is also secure, with all the data being stored in a Moralis database using private keys to both encrypt and decrypt the information. Donahub has its own distribution algorithms, permitting users to donate equally among organizations or give priority to specific organizations.

The rest of the paper will follow this outline: Section 2 describes challenges we encountered during the process of creating and carrying out this experiment, Section 3 explains the methodology we used and the solution we came up with, Section 4 evaluates the experiment and provides extra details, Section 5 presents related works, and Section 6 gives concluding remarks.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Web3-Based Portal

The development of a web3-based portal [9]. As the idea of Web3, a decentralized version of the world wide web, is still extremely new, there is little to no help regarding the development of a platform based on it. Moreover, it is necessary to connect every little bit of data collected from the platform to a database as well as connect a wallet to its user. Because of the novelty of dApps, it is a challenge to figure out the connection process and make sure that data being collected is being stored in the right places and is also able to be called whenever needed. It is also difficult to develop both the backend and frontend of such a platform, figuring out how to handle the money as well as create a pleasing user interface that provides all the information necessary.

2.2. Integration Compatibility

The compatibility of the platform to integrate with different game applications. Due to the use of our platform depending on a game's code and integrating our SDK/API, it is a challenge to ensure that our platform is compatible with the plethora of games that exist. Because different developers use different tokens, different engines, different code, our platform might not be compatible with all kinds of games. Moreover, developers might want to integrate our SDK/API in their own way.

2.3. Fund Distribution Algorithm

The diversity of the fund distribution algorithm. Our platform is going to work by distributing donations given to us to a number of different charities using a distribution algorithm. However as a platform, we most likely will be wanting to use a wide variety of algorithms, encompassing individual developers' priorities. The algorithm also needs to be flexible, able to adapt to the needs of different donors.

3. SOLUTION

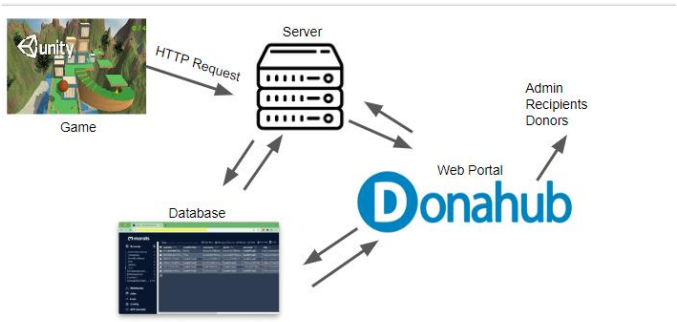


Figure 1. Overview of the solution

Donahub is a web3 decentralized blockchain-based donation platform that integrates into games in order to gain donations from players [10]. Users have two options on how to donate; either from a game or directly from our web platform. Our platform has 4 main components, the games, our server, a database, and our web portal. Players of integrated games will have the option to send an HTTP request containing their account name and donation amount to our server, which communicates with our database and our web platform. People donating from the web platform don't need to send their account name because they will have to be signed in to donate, instead only needing to send their donation amount to our server and database. Receiving an HTTP request from anywhere will create and save a new donation object in our database, which will then be displayed on our web portal for a final confirmation. The web portal houses important information such as past and pending transactions and a list of all recipients. The website is also where donors will be able to choose how their funds get distributed amongst the organizations, with an equal distribution function and a priority distribution one.

Moralis is a platform that provides the full-stack workflow for people looking to develop their own dApps. It provides you with a database that will automatically store all of the data that you have configured and is easily accessible from the frontend. All of the features that Moralis offers are cross-chain, so the data can flow easily between different blockchains. We utilized Moralis's functions in order to create our signup and signin functions as well as easily keep track of transactions and display them.

```
async function connectWithWallet() {
                let user = Moralis.User.current();
                if (!user) {
                  user = await Moralis.authenticate();
                }
                console.log("logged in user:", user);
                console.log("logged in user:", user.attributes.ethAddress);
                document.getElementById("input_walletid").value =
user.attributes.ethAddress;
                const DonaUser = Moralis.Object.extend("DonaUser");
                const query = new Moralis.Query(DonaUser);

                query.get(currentUser.objectId)
                    .then((donaUser) => {
                      // The object was retrieved successfully.
                      donaUser.set("ethAddress", user.attributes.ethAddress);
                      donaUser.set("ethUserObjectId", user.id);

                      donaUser.save().then((donaUser) => {
                          console.log("Update the object successfully.");
                          console.log(donaUser);
                      });
                    }, (error) => {
                      alert("Failed to update the user.");
                    });
                }
```

Figure 2. Screenshot of code 1

Portal

- Design
- For each screen/page
- User Homepage

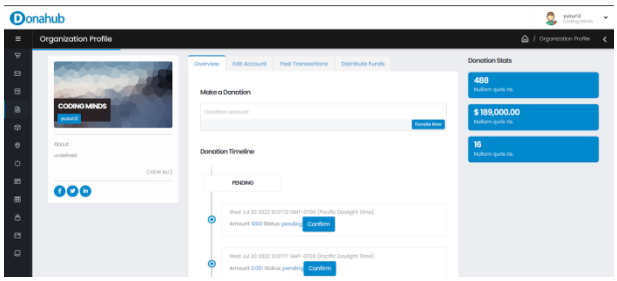


Figure 3. Screenshot of making donation

Our user profile page is where everyone is taken after they sign into their account. On the left hand side, we can see our organization logo along with its name, the user's username, and a short description of the organization. At the top middle section, we have 4 tabs: Overview, Edit Account, Past Transactions, and Distribute Funds. In Overview, we can either make a fast donation using the Make a Donation input box or confirm any pending donations that we have. Everything that we see on the left hand side can be edited using the edit account tab.

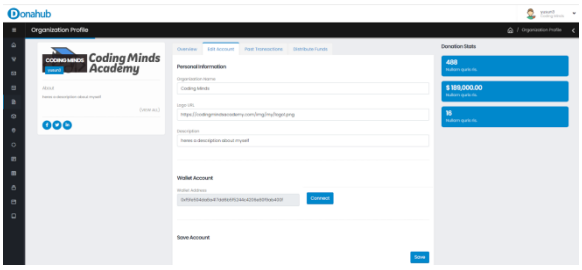


Figure 4. Screenshot of filling information

The Past Transactions tab is where you can see a timeline of every single past donation/transaction that has been made.

```
async function getUserTransactions() {
  let user = Moralis.User.current();
    if (!user) {
      user = await Moralis.authenticate();
    }
  // create query
  const query = new Moralis.Query("EthTransactions");
  console.log(user.get("ethAddress"));
  query.equalTo("to_address", user.get("ethAddress"));

  // run query
  const results = await query.find();
  console.log("user transactions:", results);
  for (var i = 0; i < results.length; i++) {
    var liItemHTML = '<li><div class="tm-box"><p class="text-muted mb-0">'
+ results[i].createdAt + '.</p>'
                + '<p>From <span class="text-primary">' +
results[i].attributes.from_address+ '</span> to <span class="text-primary">' +
results[i].attributes.to_address + '</span></p></div></li>';
      document.getElementById("transaction_list").innerHTML += liItemHTML;
    }
  }
```

Figure 5. Screenshot of code 2

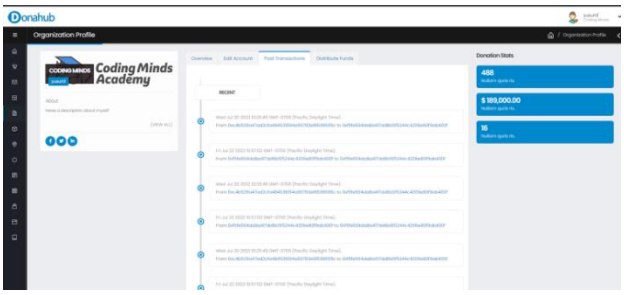


Figure 6. Screenshot of post transaction

The Distribute Funds tab is where you can decide on how you are going to donate. An equal distribution button allows for you to enter an amount. That amount will then be donated to every organization that has signed up with DonaHub.

```
async function equalDonation() {
  var amount = document.getElementById("equal_distribution_amount").value;
  const DonaUser = Moralis.Object.extend("DonaUser");
  const query = new Moralis.Query(DonaUser);
  const results = await query.find();
  console.log("Successfully retrieved " + results.length + " users.");
  if (results.length > 0) {
    for(var i = 0; i < results.length; i++) {
      donationAmountForWallet(results[i].attributes.ethAddress, amount);
      console.log("donated " + amount + " to " +
results[i].attributes.ethAddress);
    }
  }
}
```

Figure 7. Screenshot of code 3

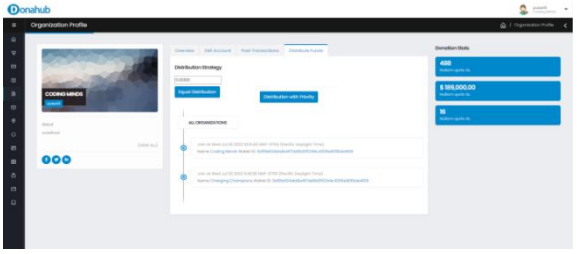


Figure 8. Screenshot of Funds

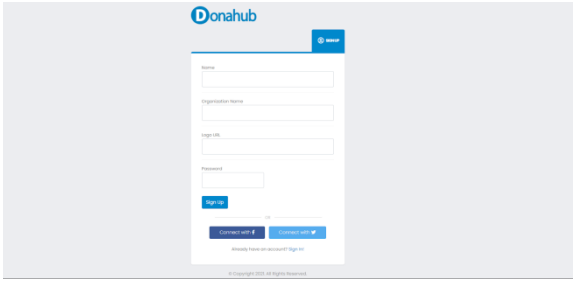


Figure 9. Sign up page

The signup page consists of 4 fields. Submitting your information here and clicking sign up will save all 4 fields into a single object (Donauser) stored in our database. There is also a button at the bottom that will allow you to sign in if you already have an account.

```
function signup() {
    // 1. collect all the values from the input boxes
    console.log("sign up clicked");
    var username = document.getElementById("input_username").value;
    var password = document.getElementById("input_password").value;
    var organization = document.getElementById("input_orgname").value;
    var logourl = document.getElementById("input_logourl").value;

    // 2. send it to server to save the user info
    const DonaUser = Moralis.Object.extend("DonaUser");
    const donaUser = new DonaUser();

    donaUser.set("username", username);
    donaUser.set("password", password);
    donaUser.set("organization_name", organization);
    donaUser.set("logourl", logourl);

    donaUser.save().then((donaUser) => {
        console.log("Save the object successfully.");
        console.log(donaUser);
        // 3. show a message "Signup request submitted"
        window.location.href = "pages-signin.html";
    });
}
```

Figure 10. Screenshot of code 4

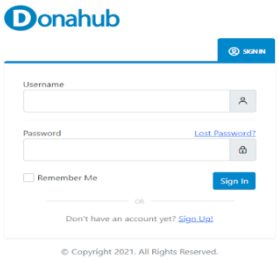


Figure 11. Sign in page

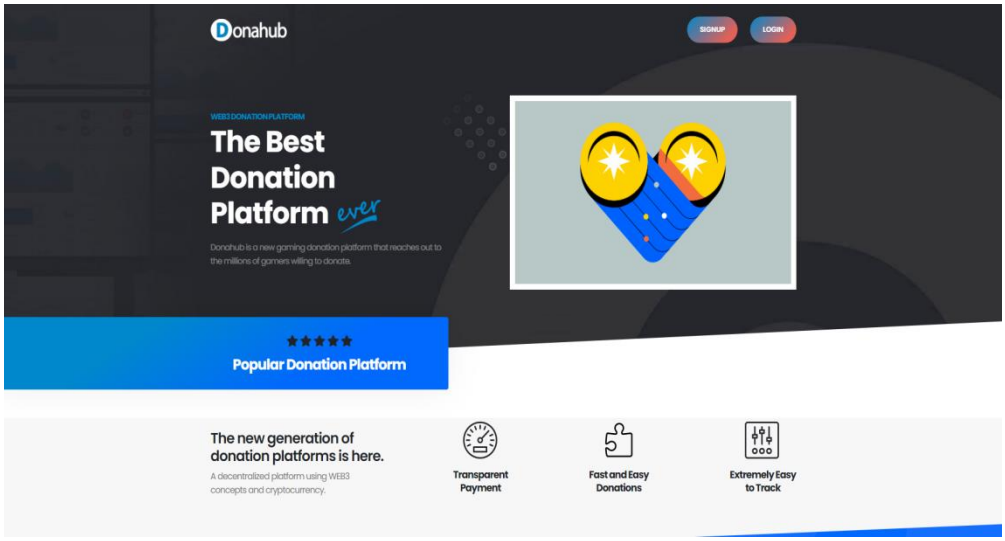
The sign-in page is where users will input their username and password in order to access their organizations profile page. The input forms are collected and compared to the data in the database in order to check for a match.

```
async function login() {
        console.log("signup clicked");
        var username = document.getElementById("input_username").value;
        var password = document.getElementById("input_password").value;

        console.log(username);
        console.log(password);

        const DonaUser = Moralis.Object.extend("DonaUser");
        const query = new Moralis.Query(DonaUser);
        query.equalTo("username", username);
        const results = await query.find();
        console.log("Successfully retrieved " + results.length + " users.");
        // Do something with the returned Moralis.Object values
        if (results.length > 0) {
          const object = results[0];
          console.log(object.id + ' - ' + object.get('username'));
          if (object.get("password") == password) {
            console.log(JSON.stringify(object));
            localStorage.setItem("currentUser", JSON.stringify(object));
            alert("sign in successful");
            window.location.href = "pages-user-profile.html";
          } else {
              alert("Incorrect or invalid password");
            }
        } else {
            alert("Invalid Username");
          }
    }
```

Figure 12. Screenshot of code 5

Homepage

- The homepage contains much needed information about our donation platform as well as provides two buttons in the top right; sign-in and sign-up. It is written using Javascript and HTML and CSS.



Server

- This server allows us to externally call upon our donation functions, enabling people to make donations from outside of our portal. This is how we will be integrating our donation platform with other games, by having them send an HTTP request to our server.
- The server uses Node.js, a back-end JavaScript runtime environment that works to execute JavaScript code externally; outside a web browser. This allows us to use JavaScript for server-side scripting and also allows us to use Express.js, which is a back-end web app framework for Node.js. Express.js is what actually allows for us to send external HTTP requests and is what we use to create all of our APIs.

- We are able to send requests through a website url, a command-line shell, and through unity. In all of these methods, the account name and amount trying to be donated is required in order to successfully make the donation.
- Code for Server:

```
app.get('/makedonation', (req, res) => {
  const Donation = Moralis.Object.extend("Donation");
                                const donation = new Donation();

                                donation.set("account", req.query.account);
                                donation.set("amount", parseInt(req.query.amount, 10));
                                donation.set("date", (date.getMonth()+1) + "/" + date.getDate() +
"/" + date.getFullYear());
                                donation.set("status", "pending");

                                donation.save().then((donation) => {
                                        console.log("Save the object successfully.");
                                        console.log(donation);
                                });
    res.send('Donation received')
})

app.get('/showall', (req, res) => {
  const donation = Moralis.Object.extend("donation");
        const query = new Moralis.Query(donation);
        const results = query.find();
        console.log("Successfully retrieved " + results.length + " users.");
    res.send('Here are all the donations')
})
```

Figure 13. Screenshot of code 6

Database

- We created a database using Moralis.
- When signing up, each user is stored in the database as an object called a donaUser. The donaUser object stores the organization's name, the user's username, logo url, and password.
- Whenever someone creates a donation, it creates a new donation object within the database.
- Game Integration with the Server API
- Within Unity, we can send an HTTP request by utilizing a coroutine that uses Unity's web request functions.
- public class NewBehaviourScript : MonoBehaviour

```
{
    // Start is called before the first frame update
    void Start()
    {
        Debug.Log("Started");
        StartCoroutine(MakeDonation());
        Debug.Log("Ended");
    }

    IEnumerator MakeDonation() {
            Debug.Log("Sending ...");
            UnityWebRequest www =
UnityWebRequest.Get("https://EricProject-1.sunyu912.repl.co/makedonation?account=yusun3&amount
=0.5");
        yield return www.SendWebRequest();
    }
```

Figure 14. Screenshot of code 7

Inside of the Unity game, we can utilize the provided TextMeshPro packages to create input fields and a button. All we need to do is have the button send an HTTP request using the code above except replacing the account name and donation amount with the values inside of the text input fields. We can do this easily by simply creating two TMP input field variables and then adding their text value to the UnityWebRequest.

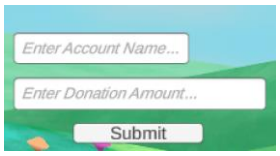


Figure 15. Put donation information page

```
public class SendDonation : MonoBehaviour
{

    public TMP_InputField account_inputField;
    public TMP_InputField donation_inputField;

    public void sendDonation()
    {
        Debug.Log("Started");
        StartCoroutine(MakeDonation());
        Debug.Log("Ended");
    }

    IEnumerator MakeDonation()
    {
        Debug.Log("Sending ...");
        UnityWebRequest www =
UnityWebRequest.Get("https://EricProject-1.sunyu912.repl.co/makedonation?account=" +
account_inputField.text + "&amount=" + donation_inputField.text);
        yield return www.SendWebRequest();

        if (www.result != UnityWebRequest.Result.Success)
        {
            Debug.Log(www.error);
        }
        else
        {
            Debug.Log("Request completed!");
        }
    }

}
```

Figure 16. Screenshot of code 8

4. EXPERIMENT

4.1. Experiment 1

On account of DonaHub being built on a blockchain network, it automatically becomes more secure than regular donations. This is because, while it is still susceptible to cyber threats, all transactions are encrypted and recorded on a distributed ledger, any hackers would need to change the entire chain in order to alter the records.

HTTPS is the protocol that we use to send information between the website and the web browser. HTTPS is a secure protocol due to its use of an asymmetric public key infrastructure. It contains a public and a private key. The public key is used by everyone who wants to access the website, encrypting information sent. The private key is used to decrypt information encrypted on the public key. The data encrypted on the public key can only be decrypted by the private key, making the website more secure.

HTTPS has been used in all of the communications in our Web3 app:

- Calling our nodejs server
- Moralis communication
- Moralis database

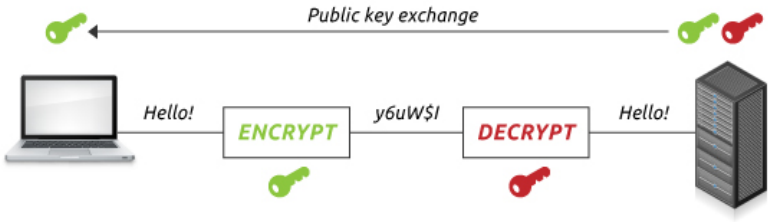


Figure 17. HTTPS security

Data encryption is important simply due to the extra security it provides. Especially because our platform handles matters regarding money, if unencrypted data was gained by a third party, they would be able to easily read the data and use it. All of our data is stored in a Moralis database which is securely encrypted using private keys. Private keys are used to both encrypt and decrypt the data in this case.

All of the data is stored in the Moralis database, which is securely encrypted using private keys.

4.2. Experiment 2

Our donation platform is shown to greatly increase transparency in transactions and in tracking what donations are used for. Due to the usage of blockchain technology as well as our web portal's display functionalities, all transactions of the digital wallet are displayed in the user homepage, allowing for easier tracking and much more transparency compared to other donation platforms, which will accept donations and never track where the money actually goes.

Donations can be made whenever someone sends an HTTPS request to our server. This creates a new donation object inside of our database. When the object is created, the donation is automatically marked as pending. Within the admin account, users will be able to click the confirm button, which will finalize the transaction and actually send the donation. Once the donation has been made, it will be publicly available to see through different ETH explorers such as: https://etherscan.io/.

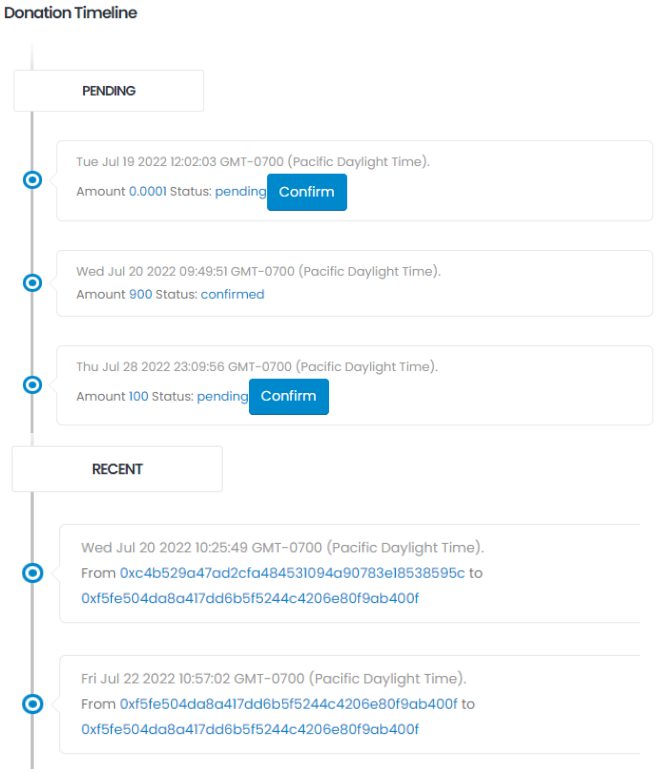


Figure 18. Donation timeline

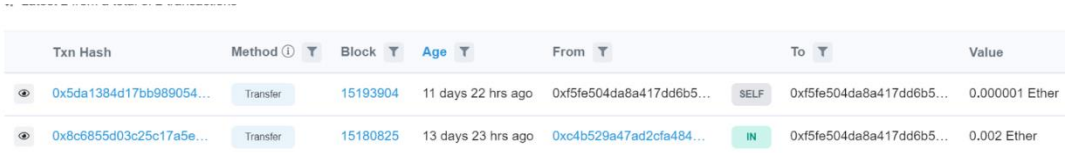


Figure 19. Files (https://etherscan.io/address/0xf5fe504da8a417dd6b5f5244c4206e80f9ab400f)

It is also possible to have the donations distributed among the different organizations that have signed up with DonaHub. Inside of the admin account user homepage, there are two options: Equal distribution or priority distribution. Equal distribution takes the amount you want to donate, splits it evenly among the number of organizations, and then sends the donations. Priority distribution allows for you to choose which organizations will receive the funds, either allowing some to get either most or all of the donated money.

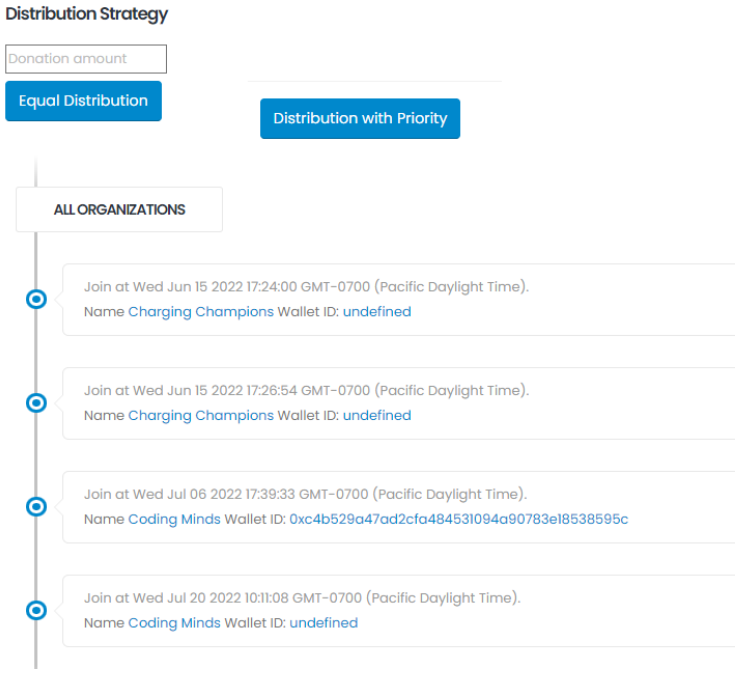


Figure 20. Distribution strategy

Our platform is extremely compatible with most programs, regardless of different game engines, coding language, etc. The only necessary component that an application needs to connect to our services is the ability to send an HTTP request. Because nearly all coding languages have incorporated HTTP requests, it becomes especially simple to send a donation request to our server.

Game Type	Library
Unity	UnityWebRequest https://docs.unity3d.com/Manual/UnityWebRequest.html
ReactJS	AJAX https://reactjs.org/docs/faq-ajax.html
GameMaker	HTTP Request Function https://manual.yoyogames.com/GameMaker_Language/GML_Reference/Asynchronous_Functions/HTTP/http_request.htm
Other Game Engine …	
Other Game Engine …	
Web - HTML/Javascript	
iOS - Swift	
Android - Java	
Flutter - Dart	Dart HTTP Library https://pub.dev/packages/http

Figure 21. Table of game type and library

5. RELATED WORK

This work focuses on the decreasing trust and transparency within the field of donations as well as offers some solutions to aforementioned issues [11]. The work offers two use cases, with the first one solving transparency issues utilizing blockchains and declaring an organization's status. The second use case focuses more on a general approach towards donations and presents a rough draft for a donation model with a DonationToken that has a special ID and features. This work was much more broad in terms of execution and was not as focused on actually creating a functioning web3 donation platform.

This work discusses the distrust regarding the usage of donated money [12]. This work mainly focuses on blockchains and their ability to create transparency and discusses the implementation of a platform that would track donations based on blockchain technology. The work states that this would offer transparent accounting and provide a transparent donation route. This work is much more broad in terms of its research, also mainly focusing on discussing the implementation as compared to actually executing.

This work analyzes the feasibility of a blockchain based financial product that would solve the many challenges as well as optimize the process of donation [13]. The authors created a dApp and tested it using Ethereum and fully evaluated the throughput and feasibility of such a blockchain based product. The work is similar in its creation of a financial product, however it focuses more on testing the feasibility rather than creating a practical product.

6. CONCLUSIONS

In summary, DonaHub is a Web3 token based donation platform that seeks to provide transparency and security in the field of fundraising [14]. Utilizing blockchain technology, DonaHub allows for easy and traceable transactions that tell donors and organizations exactly where or to where the money is going. DonaHub also provides its own distribution system, allowing users to either equally distribute funds among organizations, or give priority to specific ones. The application is also easily integrated into most games, tapping into the wide variety of gamers willing to donate money. The only factor necessary in integrating DonaHub is the ability to send an HTTPS request to our server.

The application used requires game developers to send HTTPS requests to our server in order for donations to work. There are limitations to this because, while rare, some engines do not support easily sending such requests. Moreover, there is also an issue of simply getting organizations to sign up for our portal. Another limitation is that crypto currency and digital wallets are still relatively new and not as many people have them.

The first issue can be solved by simply allowing our server to take more different kinds of requests other than just HTTPS ones, such as JSON-RPC. Other limitations can be fixed by just expanding our reach and through reach out to companies. Lastly, the number of people owning a digital wallet will gradually increase over time as cryptocurrency grows in popularity [15].

REFERENCES

[1] Almeida, Fernando, and Alexandre Cunha. "Digital Donation Platform for Nonprofit and Charity Organizations." International Journal of Information Communication Technologies and Human Development (IJICTHD) 10.3 (2018): 14-27.

[2] Rao, Thammavarapu R. N., and K-H. Nam. "Private-key algebraic-code encryptions." IEEE Transactions on Information Theory 35.4 (1989): 829-833.

[3] Mukhopadhyay, Ujan, et al. "A brief survey of cryptocurrency systems." 2016 14th annual conference on privacy, security and trust (PST). IEEE, 2016.
[4] Liu, Yukun, and Aleh Tsyvinski. "Risks and returns of cryptocurrency." The Review of Financial Studies 34.6 (2021): 2689-2727.
[5] Funk, Jeanne B. "Reevaluating the impact of video games." Clinical pediatrics 32.2 (1993): 86-90.
[6] Piper, Holly. "Case study: Charities Aid Foundation and St Mungo's." Demystifying Social Finance and Social Investment. Routledge, 2020. 140-145.
[7] Kovacova, Maria, Jakub Horak, and Michael Higgins. "Behavioral Analytics, Immersive Technologies, and Machine Vision Algorithms in the Web3-powered Metaverse World." Linguistic and Philosophical Investigations 21 (2022): 57-72.
[8] Nielsen, Henrik Frystyk, et al. "Network performance effects of HTTP/1.1, CSS1, and PNG." Proceedings of the ACM SIGCOMM'97 conference on Applications, technologies, architectures, and protocols for computer communication. 1997.
[9] Wang, Qin, et al. "Exploring Web3 From the View of Blockchain." arXiv preprint arXiv:2206.08821 (2022).
[10] Zheng, Zibin, et al. "Blockchain challenges and opportunities: A survey." International journal of web and grid services 14.4 (2018): 352-375.
[11] Enria, Luisa, et al. "Trust and transparency in times of crisis: Results from an online survey during the first wave (April 2020) of the COVID-19 epidemic in the UK." PloS one 16.2 (2021): e0239247.
[12] Callen, Jeffrey L. "Money donations, volunteering and organizational efficiency." Journal of Productivity Analysis 5.3 (1994): 215-228.
[13] Zhang, Zhonghua, et al. "Recent advances in blockchain and artificial intelligence integration: feasibility analysis, research issues, applications, challenges, and future work." Security and Communication Networks 2021 (2021).
[14] Vesterlund, Lise. "The informational value of sequential fundraising." Journal of public Economics 87.3-4 (2003): 627-657.
[15] Rathore, Hem Shweta. "Adoption of digital wallet by consumers." BVIMSR's journal of management research 8.1 (2016): 69.

AN EFFICIENT AI MUSIC GENERATION MOBILE PLATFORM BASED ON MACHINE LEARNING AND ANN NETWORK

Jiacheng Dai[1] and Yu Sun[2]

[1]International Department of the Affiliated High School of SCNU,
Zhongshan Road West No.1, Guangzhou, China
[2]California State Polytechnic University, Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

The aim of this paper is to provide a solution to the growing need for fresh music to use in media, as adding music can greatly enhance the media's atmosphere and the viewers' experience [6]. Our solution to this issue was the creation of a mobile application named MFly that can output music using the sentiment from an inputted message. To test the effectiveness of this new music-generating method, an experiment was conducted in which twenty-three participants inputted a message with a positive and negative sentiment each and recorded whether each outputted musical piece accurately represented the sentiment from the message [7]. A post-experiment survey was also provided to each of the participants to gauge the convenience and practicality of the application. The results indicated that MFly was largely successful at conveying messages into appropriately fitting music. However, the practicality of the application could use some work, as generating music based on the sentiment does not always seem to match up with the original inputted message's sentiment, especially with messages that have a negative sentiment. Furthermore, feedback from participants indicated that the application could still improve with the addition of more features, such as the ability to save the generated music for later use.

KEYWORDS

Machine Learning, AI, Mobile application.

1. INTRODUCTION

Music has played a very prevalent part in human culture for centuries. From films and television shows to festivals and funerals, they are present in many aspects of life. Music has been shown in studies to provide benefits to those who listen to it. One of the benefits of music is improved memory because the brain can link sounds and music with certain experiences and places. Music can provide other various benefits related to improving the mental well-being of those who listen to it, such as reducing stress levels while boosting mood and improving performance in cognitive tasks [1]. In real life, music can also provide people with a path for them to express their creativity and channel their thoughts and emotions, help connect people with different cultural backgrounds with the power of music, and help create certain atmospheres like the ones required in night clubs.

Music is significant because it is capable of enhancing an event or a form of media [8]. By adding music to a film or a video game, the viewer or player can experience a greater level of immersion by having the music reflect and intensify the atmosphere [9]. Another quality that can

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 43-55, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121705

makes music an important part of society is that it can help bring people together [10]. People can bond with each other over their passion for music, whether it be by listening to music, playing music on an instrument, singing, or creating their own music. This sense of bonding is not limited to a personal level; entire communities and cultures can blend together with the help of music. By having two cultures share their music with one another, both cultures can gain a better understanding of each other, which can improve cultural diversity globally. These many common uses of music call for a quick and convenient method of generating new music.

There have been attempts to generate music using artificial intelligence by taking in an existing song or melody as input [15]. Some of these programs have options in order to provide more flexibility to those who want a certain quality or element of music in particular. One way this has been done is by producing multiple different music pieces based on the one melody or audio file that was inputted. The individual parts that make up the music, such as the melody, the bass, and the drums, could also be mixed and matched in some cases. This gave the program users much control over the generation of the final output music. Another way to bring the program users more flexibility is to provide them with values to adjust. Using sliders, they can adjust qualities such as whether the song is in a major key or minor key and how conventional the song is.

However, the ability of these programs to generate music is limited by whatever is being inputted. Because audio usually has to be inputted in order for the artificial intelligence to generate music, it will usually be very similar to whatever the original audio is. Despite the flexibility and increased options of having multiple possible musical elements to choose from and adjust, the generated music still bears a striking resemblance to what was inputted in most scenarios. There are relatively few existing programs that can competently output music based on user-inputted values without any inputted reference music. Another general issue with AI-generated music is that certain “personal touches” that human musicians with superior knowledge and understanding of music and specific music styles are generally difficult to emulate or reproduce with artificial intelligence. Therefore, artificial intelligence that specializes in generating music may need to be further refined before it can see widespread usage.

The method presented in the research is a mobile application named MFly that generates music using artificial intelligence. The two main features of the application are generating music based on inputted reference audio and generating music based on the sentiment of inputted messages. With the reference audio, the artificial intelligence will attempt to generate a musical piece that sounds similar to the reference audio. Using messages, the application will attempt to derive the sentiment based on the connotation of the words used. Sentiments can either be positive or negative. Then, the application will output music that fits as closely to the selected sentiment as possible.

Other applications have a similar approach in terms of using reference audio to generate music, almost as if the application was trying to predict what an extension of the original audio would sound like. What sets MFly apart from many other applications that were developed for the same purpose is that this application aims to convert messages into music. While it is not advanced enough to detect specific keywords and make intricate changes like adjusting the specific style of music, this project aims to be a stepping stone towards an application that can eventually do so.

The remainder of the paper is structured into sections labeled 2 through 6. Section 2 provides insight into the challenges that were faced when creating the application. Section 3 dives into the overall development of the application as well as some specific key parts of the application, while Section 4 explains the experiments that took place and how the experiments’ results proved the effectiveness of this application at properly providing music based on sentiment. In Section 5, related works involving music generated by artificial intelligence are explored and

compared to this paper. Lastly, Section 6 ends the paper with concluding remarks and how the limitations of the application can be resolved in the future.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Brainstorming the Music-Generating Methods

One major hurdle that had to be overcome when developing the MFly mobile application was figuring out what options the users should be provided with when generating the music in the MFly application. To make the application more welcoming to those who are familiar with applications of the same purpose, the traditional implementation of inputting a reference song was added as a choice for the users to generate music. However, the application would preferably have a relatively unconventional method of generating music compared to most other AI-generated music applications. When brainstorming ideas for what such a method could be, the concept of being able to return musical pieces using written messages sounded unique and intriguing. Our vision behind this second music-generating method was that depending on what the message said and the connotation that the words in the message had, the application would output a musical piece to match the tone of the message.

2.2. Implementing AI-Generation of Music using Messages

The next challenge, however, came with the method of generating music using messages. Such a task sounded incredibly daunting at first, as there does not seem to be one simple step to convert a written message into a piece of music. Rather, the message must instead undergo multiple steps to reach the final desired output. The message would first be converted into two possible sentiments: positive or negative. The plan would be to use the keywords and the connotation of certain words to determine the sentiment. After retrieving the sentiment, the application can select a song based on a list of songs, and the one that is selected depends on whether the mood of the message is determined to be positive or negative. By breaking the process down into two main parts, the issue of outputting a song to the user with artificial intelligence using a written message was successfully overcome.

2.3. Testing the AI-Generation of Music

Lastly, one of the largest challenges was gauging the accuracy and quality of the AI-generated music through experiments or surveys. As people have varying tastes in music and what sounds like good music to some might sound like terrible music to others, whether the music generated by the AI was high-quality or not is very subjective. While such a question could still be included in a post-experiment survey, a more concrete value to test for was whether the sentiment in the written message was reflected accurately in the music. For example, a positive message should result in a positive and happy-sounding piece of music, and a sad and negative message should result in a sad-sounding piece of music. As there are still differences in opinion as to what music accurately reflects the sentiment of a message and what does not among people, as many participants as possible from multiple diverse communities would be needed to reflect the opinion of the general public as accurately as possible.

3. SOLUTION

MFly was created using Flutter and Python; Flutter primarily makes up the front-end of the code, while Python acts as the back-end of the code. Using multiple Dart files, the application was able to cycle through various screens, including the splash screen when first opening the application and the home screen that appears right after [11]. There are two possible options on the home screen labeled by two buttons. The first button leads to a screen for inputting audio, and the second button leads to a screen for inputting a message. On both screens, after the data has been inputted, the application transfers over to the screen that plays the generated music. The Python file was responsible for handling the artificial intelligence aspect of the application. Artificial intelligence is primarily made using Keras, which is an open-source software library. In particular, the Long Short-Term Memory neural network was used, which is a kind of recurrent neural network that utilizes sequential information and learns through gradient descent. Because the Long Short-Term Memory neural network has the capability to recognize and encode patterns over an extended period of time, it excels at tasks such as music generation. When generating music, the application starts with a MIDI file [12]. First, the sequences used by the neural network are prepared. Then, the model is created and is used to generate a prediction of how the music should continue using the patterns it recognizes. Lastly, the MIDI file is converted back into an audio file for the audio player screen.

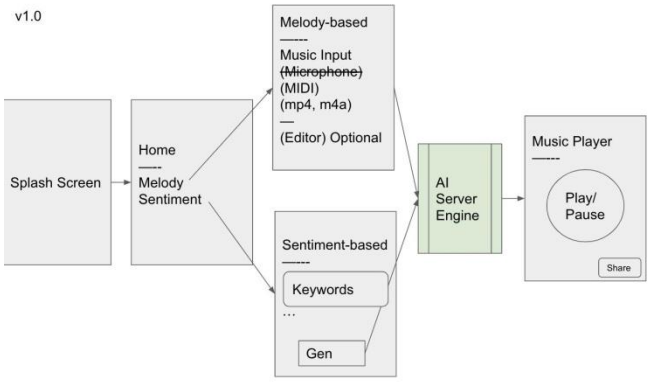


Figure 1. Overview of the solution

When the application is first opened, the splash screen appears. This splash screen is handled with a single Dart file in which its sole purpose is to flash the application logos on the screen for a few seconds. This was performed by creating a Future that would switch to the main menu screen after a 3-second delay. From the main menu screen, elevated buttons allow the user to navigate around the different screens in the application with Flutter's Navigator. The two main options in the main menu are Melody AI and Sentiment AI; Melody AI uses reference audio as input, while Sentiment AI uses a message as input.

To implement the "Melody AI" option of the main menu screen, the audio file is converted into a MIDI file. The audio is retrieved using the path where the file is located, then the audio is loaded with Librosa (a Python package for music and audio analysis). In the MIDI file, only the piano was used in order to make the training of the model as effective as possible. Therefore, the generated music's only instrument is the piano. A pre-trained model was fetched for the purpose of this application [5]. The notes of the MIDI file are extracted, then the notes used to train the pre-trained model are extracted and sorted so that all tested pitch names could be saved. The extracted notes from the input audio and the pitch names from the train data are used to prepare the neural network sequences, then a Long Short-Term Memory neural network is created.

Exactly 250 notes for the song are created using the LSTM model, and the generated notes are compiled into a MIDI file [13]. The MIDI file is converted into a WAV file, which appears on the music player screen and allows the users to listen to the video. During this process of conversion, a Transcription method based on Magenta Onsets and Frames is used, which is a newly improved transcription method with the help of CNN and LSTM models. The method hugely improves the capability of music transcription by separating note detections into two stacks of neural networks instead of one. After installing Magenda, importing libraries, and fetching Magenda, the MIDI file can finally be generated.

To implement the "Sentiment AI" option of the main menu screen, the sentiment analysis was performed by taking the words of the message as a String and running the words through a method to generate the sentiment. The Natural Language Toolkit, or NLTK for short, is imported into Python and the VADER lexicon is downloaded. VADER stands for Valence Aware Dictionary and sEntiment Reasoner, which is a rule-based sentiment analysis tool that specializes in the ability to detect the sentiment of texts. A sentiment intensity analyzer that was imported from the Natural Language Toolkit is created. Then, a method to find the polarity scores using the aforementioned sentiment intensity analyzer is called with the String as a parameter and its value is returned. The polarity score can be anywhere between -1 and 1, in which a negative score means that the message has a negative sentiment and a positive score means that the message has a positive sentiment. Depending on what the score is, a variable representing mood will be assigned a certain value to determine which song will be chosen to be played.

```
void uploadFileToServer(var filePath) async {
  print("uploading ... ");
  setState(() {
    loading=true;
  });
  var url = Config.serverUrl + '/melody';
  print(url);
  http.MultipartRequest request = http.MultipartRequest('POST', Uri.parse('$url'));

  request.files.add(
    await http.MultipartFile.fromPath(
      'song',
      filePath,
      //contentType: MediaType('audio', 'midi'),
    ),
  );
  // The following line will enable the Android and iOS wakelock.
  Wakelock.enable();
  request.send().then((r) async {
    print(r.statusCode);
    // print(json.decode(await r.stream.transform(utf8.decoder).join()));
    if (r.statusCode == 200) {
      Directory tempDir = await getTemporaryDirectory();
      String tempPath = tempDir.path + "/test.wav";

      File f = new File(tempPath);
      await f.writeAsBytes(await r.stream.toBytes());
      print(f.path);
      print(f.length());
      setState(() {
        loading=false;
      });
      // The next line disables the wakelock again.
      Wakelock.disable();
      Navigator.of(context).push(MaterialPageRoute(builder: (context) => PlayerPage(title: tempPath)));
```

```
StreamBuilder<PlayerState>(
  stream: player.playerStateStream,
  builder: (context, snapshot) {
    final playerState = snapshot.data;
    final processingState = playerState?.processingState;
    final playing = playerState?.playing;
    if (processingState == ProcessingState.loading ||
        processingState == ProcessingState.buffering) {
      return Container(
        margin: EdgeInsets.all(8.0),
        width: 64.0,
        height: 64.0,
        child: CircularProgressIndicator(),
      );
    } else if (playing != true) {
      return IconButton(
        icon: Icon(Icons.play_arrow),
        iconSize: 64.0,
        onPressed: player.play,
      );
    } else if (processingState != ProcessingState.completed) {
      return IconButton(
        icon: Icon(Icons.pause),
        iconSize: 64.0,
        onPressed: player.pause,
      );
    } else {
      return IconButton(
        icon: Icon(Icons.replay),
        iconSize: 64.0,
        onPressed: () => player.seek(Duration.zero),
      );
    }
  },
),
```

```
def generate(song_path, output_filename):
  song_notes = extract_notes(song_path)
  """ Generate a piano midi file """
  #load the notes used to train the model
  with open('/content/Classical-Piano-Composer/data/notes', 'rb') as filepath:
      notes = pickle.load(filepath)

  # Get all pitch names
  pitchnames = sorted(set(item for item in notes))
  # Get all pitch names
  n_vocab = len(set(notes))

  network_input, normalized_input = prepare_sequences(song_notes, pitchnames, n_vocab)
  model = create_network(normalized_input, n_vocab)
  prediction_output = generate_notes(model, network_input, pitchnames, n_vocab)
  create_midi(prediction_output,output_filename)
```

```
def create_network(network_input, n_vocab):
  """ create the structure of the neural network """
  model = Sequential()
  model.add(LSTM(
      512,
      input_shape=(network_input.shape[1], network_input.shape[2]),
      recurrent_dropout=0.3,
      return_sequences=True
  ))
  model.add(LSTM(512, return_sequences=True, recurrent_dropout=0.3,))
  model.add(LSTM(512))
  model.add(BatchNorm())
  model.add(Dropout(0.3))
  model.add(Dense(256))
  model.add(Activation('relu'))
  model.add(BatchNorm())
  model.add(Dropout(0.3))
  model.add(Dense(n_vocab))
  model.add(Activation('softmax'))
  model.compile(loss='categorical_crossentropy', optimizer='rmsprop')

  # Load the weights to each node
  model.load_weights('/content/Classical-Piano-Composer/weights.hdf5')

  return model
```

```python
def generate_notes(model, network_input, pitchnames, n_vocab):
  """ Generate notes from the neural network based on a sequence of notes """
  # pick a random sequence from the input as a starting point for the prediction
  start = numpy.random.randint(0, len(network_input)-1)


  int_to_note = dict((number, note) for number, note in enumerate(pitchnames))

  pattern = network_input[start]
  # print(type(pattern))

  prediction_output = []

  # generate 250 notes. Generate a 30 sec songs
  for note_index in range(250):
      prediction_input = numpy.reshape(pattern, (1, len(pattern), 1))
      prediction_input = prediction_input / float(n_vocab)

      prediction = model.predict(prediction_input, verbose=0)

      index = numpy.argmax(prediction)
      result = int_to_note[index]
      prediction_output.append(result)

      pattern.append(index)
      pattern = pattern[1:len(pattern)]

  return prediction_output
```

```python
import nltk
from nltk.sentiment import SentimentIntensityAnalyzer
import ast
nltk.download('vader_lexicon')

def nltkSentiment(sentence):
  #NLTK anaylzer provides 4 datapoints in a dictionary --> compound datapoint is used;
  anaylzer = SentimentIntensityAnalyzer()
  return anaylzer.polarity_scores(sentence)['compound']

print(nltkSentiment("Sentiment analysis has not been good."))
```

Figure 2. Code of MFly

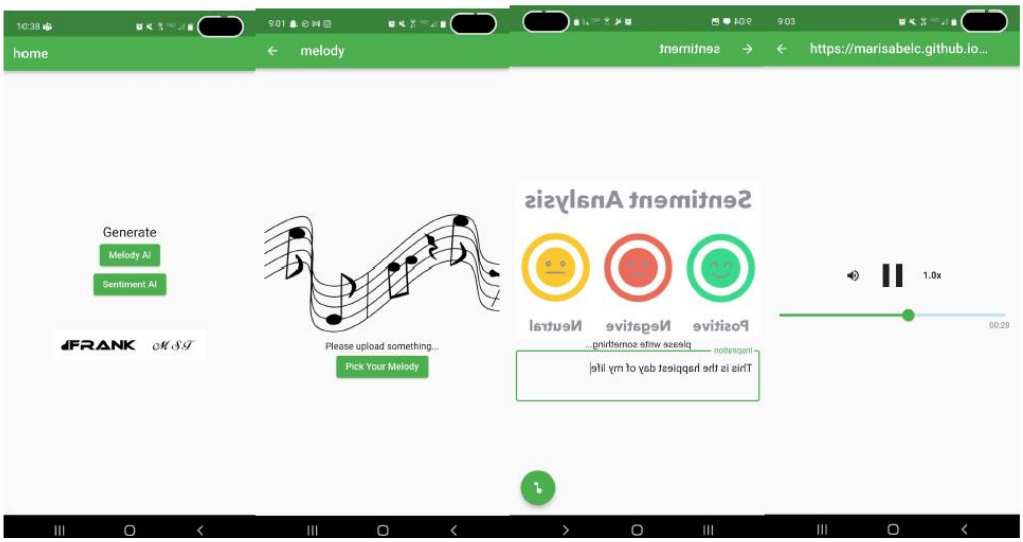


Figure 3. Screenshots of MFly

4. EXPERIMENT

4.1. Experiment 1

MFly resolves the issues of struggling to obtain music that suits people's needs in terms of atmosphere and mood. In order to test how effective the system of generating music is at

accounting for the intent of the user when using a message from the user as input, twenty-three participants downloaded the mobile application on their mobile phones, then tested this implementation by going into the sentiment page of the application, typing in a message, and recording whether the music that was outputted matches the sentiment of the message. The participants would perform this process two times, one for each possible sentiment (positive and negative).

Participant #	Positive Sentiment Accurately Portrayed	Negative Sentiment Accurately Portrayed
1	Yes	Yes
2	Yes	Yes
3	No	No
4	Yes	No
5	Yes	Yes
6	No	Yes
7	Yes	Yes
8	Yes	Yes
9	Yes	Yes
10	Yes	No
11	Yes	No
12	Yes	Yes
13	Yes	Yes
14	Yes	Yes
15	Yes	Yes
16	Yes	Yes
17	Yes	No
18	Yes	No
19	Yes	Yes
20	Yes	Yes
21	Yes	Yes
22	Yes	Yes
23	Yes	No
Accuracy	21/23	16/23

Figure 4. Table of possible sentiment

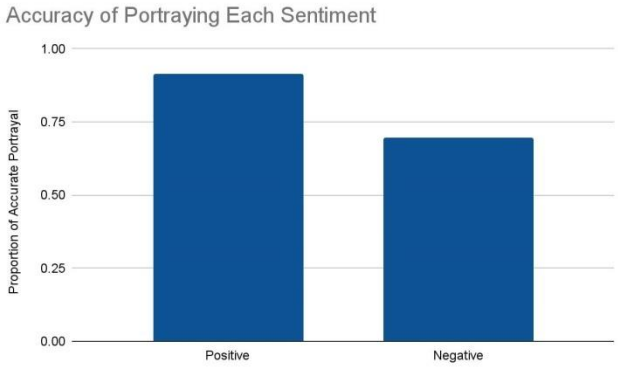


Figure 5. Accuracy of portraying each sentiment

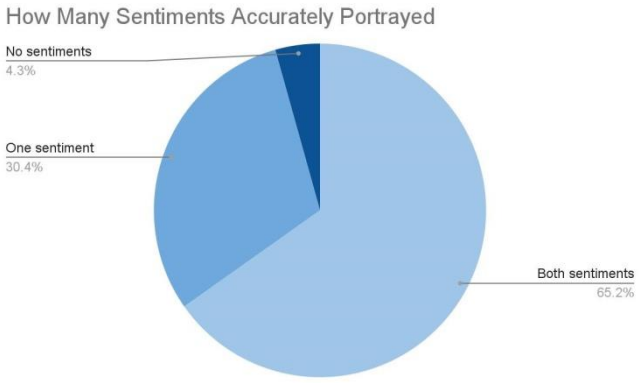


Figure 6. Chart of sentiments accurately portrayed

According to the charts and graphs, it appears that the accuracy of properly portraying the sentiment of a message is highest in messages that have a positive sentiment, as 21 out of the 23 participants recorded that their inputted message with a positive sentiment resulted in a musical piece that they believed to fit a positive atmosphere. On the other hand, the lowest accuracy comes from messages with a negative sentiment, with only 16 out of 23 participants indicating that their negative message returned appropriately fitting music. A possible explanation for these results is that the AI that is being used to select the music is not well-trained enough in sentiment analysis to choose the correct song. As for each participant, approximately 65.2% of the participants believed that their positive and negative messages both returned appropriately fitting music. Therefore, 34.8% of the participants found that at least one of the music outputs did not match the corresponding original message's sentiment.

4.2. Experiment 2

This application also seeks to provide convenience for those who wish to quickly have access to freshly created music. After the previous experiment is finished, the participants will take a postexperiment survey that gauges how convenient using the application was and whether they would use it again in the future. These values will be asked in two separate questions, "How convenient is the application to use?" and "How practical is using this application in daily life?", and the participants will have the option to answer them using a scale from one to ten. An optional free-response question for any additional feedback will also be provided to the users at

the end of the survey. As the survey features the same twenty-three participants from the other experiment, there is more than enough of a sample size to account for variability.

Participant #	Convenience of the Application	Practicality of the Application
1	9	6
2	8	7
3	10	5
4	8	5
5	7	9
6	9	6
7	8	8
8	10	8
9	7	6
10	8	5
11	6	4
12	9	7
13	7	7
14	8	6
15	7	8
16	8	7
17	8	6
18	9	5
19	6	5
20	7	7
21	8	7
22	9	8
23	9	7
Average	8.0435	6.4783

Figure 7. Table of convenience and practicality of the application

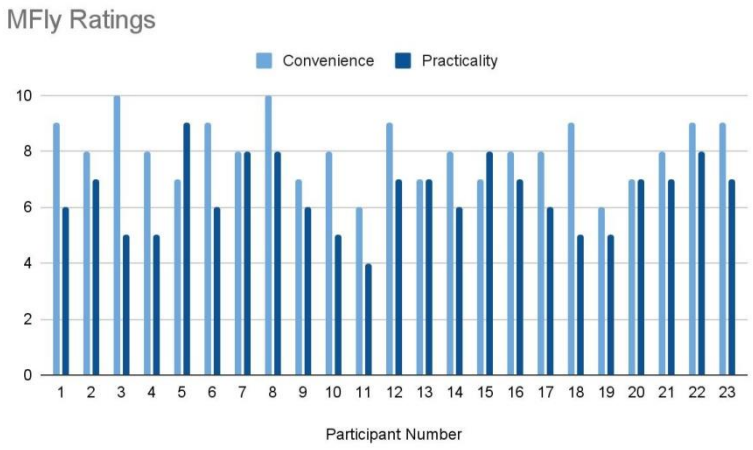


Figure 8. Graph of MFly ratings

Generally, both the convenience and the practicality of the application were rated somewhat highly. A pattern that can be seen among the participants is that they generally seem to rate the convenience of using the application higher than the practicality of the application. The difference between the two values being rated can be emphasized by comparing the average ratings, in which the average convenience rating of approximately 8.0435 is significantly higher

than the average of approximately 6.4783. The primary criticism from the participants 'feedback was that there was no way within the application to easily share the AI-generated music to other platforms or social media. Another recurring point of feedback was that the application could include more features to make quality-of-life improvements for the users by adding features. However, participants also noted that the simple look and the clearly labeled buttons made navigation of the application very easy.

According to the experiment that involves testing the success of the application at accurately reflecting the sentiment of the inputted message in the outputted music, a decent portion of the participants believed that there was at least one instance in which the sentiment from the message did not align with the tone of the music and the feelings and emotions that the music gave off. These results indicate that although the application can generate music with a specific sentiment to some degree, there is still work to be done when it comes to improving the accuracy of portraying the sentiment of a message in the outputted music properly.

From the post-experiment survey, it appears that while the application is convenient and simple to use for the general public, the application can still use some room for improvement when it comes to its practicality. The primary complaints based on the participants' feedback are that there is no way to share the audio outside the application and there is not enough features in the application, which indicates that the most urgent changes to make to the application should attempt to create more features to improve user experience, especially for exporting the audio from the application.

5. RELATED WORK

Ramon Lopez de Mantaras and Josep Lluis Arcos from the Spanish National Research Council surveyed computer music systems that used artificial intelligence techniques focusing on compositional systems, improvisation systems, and performance systems. Case-based reasoning has been shown to be an incredibly useful technique for composing music with artificial intelligence because it has the capability to directly utilize extracted knowledge from recorded human performance examples in an implicit manner [2]. While de Mantaras and Arcos thoroughly describe and analyze the AI techniques to generate music and explore the challenge behind giving these generated pieces of music the expressiveness of human-created music, this work aims to develop a mobile application that generates music based on inputted audio or message and evaluate its effectiveness at doing so.

Ryan Louie from Northwestern University teamed up with researchers from Google Brain to produce a study on developing tools to steer artificial intelligence so that music composers would be able to work alongside the AI in a clear and coherent manner. Some of these tools included directing the voices towards a specific range or adjusting certain values based on sliders. Results indicated that these tools were effective at improving the trust, understanding, and sense of control and productiveness of users [3]. Both the work of Louie et al. and this work focus on developing an interface for people to conveniently generate music to their liking; however, MFly focuses much more on using sentiment as a whole for the input for music generation than the interface created by Louie et al. does.

Emma Frid from the KTH Royal Institute of Technology collaborated with Celso Gomes and Zeyu Jin from Adobe Research to conduct studies on generating music with artificial intelligence by inputting an example reference of a song and mixing AI-generated parts of the song. In recent times, creating videos for social media platforms and finding suitable, copyright-free music to accompany them can be a difficult challenge, and online content creators for these platforms have indicated that composing music in conjunction with the AI is preferable [4]. The work of

Frid et al. is similar to this work in that an AI and a person both have a part in the creation of the final musical piece. Frid et al. focused on more complex ways to customize the music, such as mixing from a selection of generated melody and bass tracks, while this work simply used the sentiment of an inputted message to create the music.

6. CONCLUSIONS

The mobile application MFly was designed to generate new musical pieces using artificial intelligence [14]. The two implemented methods of music generation that the user can choose from are inputting reference audio or inputting a message. By taking in reference audio as input, the artificial intelligence would attempt to emulate the audio's musical style while generating the music. By taking in a message as input, the application would extract the sentiment of the message and use the sentiment to influence the generation of the music. The goal of the application is to become a convenient and flexible source of music for its users.

In order to test the effectiveness of the application, an experiment was conducted with the method of music generation that involves an inputted message. Twenty-three participants were asked to input one message with a positive sentiment and one message with a negative sentiment. Then, they would record whether they believed each generated piece of music accurately portrayed the sentiment of the corresponding message. The results of this experiment were that the application was the most accurate when generating music with a positive sentiment and the least accurate when generating music with a negative sentiment. A post-experiment survey was also conducted using the same participants, in which each participant would rate the convenience and practicality of the application from 1 to 10. Survey results indicated that the application was very simple to use and was fairly practical in daily usage. However, the participants also mentioned that the addition of a feature to share or export the generated music in the application would greatly enhance the user experience.

One of the most limiting features of the MFly mobile application is the ability to save the music that is created. Once the music is generated by artificial intelligence, the application moves to a screen in which the generated audio is played. However, there is no current way to save the without the use of external tools, such as a mobile screen recorder that captures the audio playing on the device. Without such a feature, the application would be much less practical for the general public, as the audio would not be easily transferable outside of the application for other means.

To tackle this limitation in the future, something that could be done is updating the interface to create a download button. Once pressed, this download button will save the AI-generated music onto the device, and the user will be free to share it online in various social media platforms. This will greatly improve the quality of life for MFly's users.

REFERENCES

[1] How does music affect the mind and body? Medford Leas. (2022, July 19). Retrieved August 30, 2022, from https://medfordleas.org/how-does-music-affect-the-mind-and-body/

[2] Frid, E., Gomes, C., & Jin, Z. (2020). Music Creation by Example. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1 – 13. https://doi.org/10.1145/3313831.3376514

[3] Louie, R., Coenen, A., Huang, C. Z., Terry, M., & Cai, C. J. (2020). Novice-ai music co-creation via AI-steering tools for deep generative models. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1 – 13. https://doi.org/10.1145/3313831.3376739

[4] de Mantaras, R. L., & Arcos, J. L. (2002). AI and Music: From Composition to Expressive Performance.AI Magazine, 23(3). https://doi.org/https://doi.org/10.1609/aimag.v23i3.1656
[5] Skuldur. (2019). Skuldur/classical-piano-composer. GitHub. Retrieved August 30, 2022, from https://github.com/Skuldur/Classical-Piano-Composer
[6] Collins, Karen, ed. From Pac-Man to pop music: interactive audio in games and new media. Routledge, 2017.
[7] Bown, Oliver Roland, and Aengus Martin. "Autonomy in music-generating systems." Eighth Artificial Intelligence and Interactive Digital Entertainment Conference. 2012.
[8] Juslin, Patrik N., László Harmat, and Tuomas Eerola. "What makes music emotionally significant? Exploring the underlying mechanisms." Psychology of Music 42.4 (2014): 599-623.
[9] Donnelly, Kevin J., William Gibbons, and Neil Lerner. "Music in Video Games." All Your Bass are Belong to Us (= Routledge Music and Screen Media Series), New York (2014).
[10] North, Adrian C., David J. Hargreaves, and Jon J. Hargreaves. "Uses of music in everyday life." Music perception 22.1 (2004): 41-77.
[11] Weisser, Martin. "DART – The dialogue annotation and research tool." Corpus Linguistics and Linguistic Theory 12.2 (2016): 355-388.
[12] Cambouropoulos, Emilios. "From MIDI to traditional musical notation." Proceedings of the AAAI Workshop on Artificial Intelligence and Music: Towards Formal Models for Composition, Performance and Analysis. Vol. 30. 2000.
[13] Xue, Hao, Du Q. Huynh, and Mark Reynolds. "SS-LSTM: A hierarchical LSTM model for pedestrian trajectory prediction." 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.
[14] Dubnov, Shlomo, et al. "Using machine-learning methods for musical style modeling." Computer 36.10 (2003): 73-80.
[15] Fetzer, James H. "What is Artificial Intelligence?." Artificial Intelligence: Its Scope and Limits. Springer, Dordrecht, 1990. 3-27.

RESEARCH ON CREATIVE THINKING MODE BASED ON CATEGORY THEORY

Tong Wang

Guangzhou College of Technology and Business, China

ABSTRACT

The research on the brain mechanism of creativity mainly has two aspects, one is the creative thinking process, and the other is the brain structure and functional connection characteristics of highly creative people. The hundreds of millions of nerve cells in the brain connect and interact with each other. The human brain has a high degree of complexity at the biological level, especially the rational thinking ability of the human brain. Starting from the connection of molecules, cells, neural networks and the neural function structure of the brain, it may be fundamentally impossible to study the rational thinking mode of human beings. Human's rational thinking mode has a high degree of freedom and transcendence, and such problems cannot be expected to be studied by elaborating the realization of the nervous system. The rational thinking of the brain is mainly based on the structured thinking mode, and the structured thinking mode shows the great scientific power. This paper studies the theoretical model of innovative thinking based on category theory, and analyzes the creation process of two scientific theories which are landmarks in the history of science, and provides an intuitive, clear interpretation model and rigorous mathematical argument for the creative thinking. The structured thinking way have great revelation and help to create new scientific theories.

KEYWORDS

Category theory, creative thinking mode, structured thinking way.

1. INTRODUCTION

cogito, ergo sum (I think, therefore I am)
---René Descartes

From ancient times to the present, the discussion about human rational thinking has always been one of the philosophy topics. Aristotle put forward formal logic in "On Instruments" and studied the syllogistic rules of propositional reasoning, which is the generalization and summary of human rational thinking by ancient philosophers. Similarly, the discussion of human rational thinking is also at the core of modern philosophy. Philosophers have established various theories of mind in the name of "anthropology" or "human knowledge research", forming two camps of "rationalism" and "empiricism". In "Critique of Pure Reason", Kant proposed the cognitive model of rationalism and empiricism---transcendental idealism and rational intuition. Russell was extremely critical of Kant's philosophy, believing that human rational thinking lies in free creation and does not recognize the presupposition of transcendental idealism. Scientists in modern philosophy and cognitive research believe or assume that human beings have a certain innate cognitive structure of the brain. By studying and understanding this cognitive structure, we can solve the problem that how rational human cognition is possible.

The development of modern technical means has made great progress in the study of the brain's thinking mode. For example, functional magnetic resonance imaging technology provides

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 57-65, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121706

technical support for revealing the creative brain structure and functional network connections [1], providing people to study the brain structure and connection with function. Brain imaging technology is used to study the unique phenomena such as epiphany and the internal connection of the brain [2]. Based on electroencephalogram (EEG) and brain functional imaging research technology, Psychologists directly observe the activity of the brain when processing complex information, so as to explore the brain mechanism of creative thinking (Bowden, Jung-Beeman, & Zf, 2007; Fink et al., 2007; Luo & Knoblich, 2007; Srinivasan, 2007). So far, there are two main aspects of brain mechanism research on creativity, one is the creative thinking process, and the other is the brain structure and functional connection characteristics in highly creative people [3] [4] [5].

Creativity is the advanced manifestation of human rational thinking, the glory of human mind, and the source of the progress of human civilization. However, creative thinking itself is complex. As is seen in above literature, with different experimental conditions and technical means, the research results of the brain mechanism of creative thinking are very different. It is therefore difficult to have a clear and deterministic conclusion about the brain mechanisms of creative thinking (one-to-one explanatory model). No studies have shown that creative brain functional connectivity can be altered by training and then have practical meaning.

Hundreds of millions of nerve cells in the brain are interconnected and interact with each other, producing people's superb motor movements, and also emerging people's complex advanced cognitive activities, such as perception, language, thinking, etc. The human brain has a high level of biological complexity, especially the rational thinking ability of the human brain. Research in brain science involves neuroanatomy, neurophysiology, neural networks, neurolinguistics, and more. Starting from the connection of molecules, cells, neural networks and the neural function structure of the brain, it may be fundamentally infeasible to study the rational thinking mode of human beings. Human's rational thinking mode has a high degree of freedom and transcendence, and such problems cannot be expected to be studied by elaborating the realization of the nervous system.

The rational thinking of the brain is mainly based on the structured thinking mode. The structured thinking mode of the brain shows the great scientific power, and this thinking mode is the main mode which had the major scientific discoveries in the whole mankind history. This paper studies the structural characteristics of the brain's thinking mode based on the category theory, and the brain's structural thinking mode is clearly explained and understood. Understanding and mastering the structured way of thinking has great enlightenment and help to create new scientific theories.

2. CATEGORY THEORY

2.1. History and Definition of Category Theory

As early as in ancient Greece, Aristotle wrote his famous "Categories", which explored the classification of objects that can be recognized by humans. In the context of modern mathematics, categories have other meanings and precise mathematical definitions. But, the basic intuition remains unchanged: a category is formed by grouping together a class of similar objects. Category theory was founded in MacLane and Eilenberg's 1945 paper "General theory of natural equivalences" [6]. Category theory has expanded into most fields of modern mathematics in a very short period of time, and category theory is the language and thinking way of describing abstract mathematical structures.

Definition 1: A category C includes:

(1) A collection of a set of objects O; as a mathematical object, it has internal operation rules;

(2) A collection of morphisms, in which morphisms (or arrows, functors f) f: X → Y ; X is the domain of f, and Y is the codomain, denoted as $\mathrm{dom}(f) = X$, $\mathrm{cod}(f) = Y$. Hom(X,Y) to denote the entirety of arrows from object X to Y. This is a collection, as shown in Figure 1(a) below.

For any object X, there is an arrow, as shown in Figure 1(b) below, $1X : X \to X$, 1X is named the identity of X, which is the unit law. For any f, $f \circ 1X = f = 1Y \circ f$. There are two arrows f and g, such that $\mathrm{cod}(f) = \mathrm{dom}(g)$ as shown in Figure 1(c) below. g∘f, denote the composition of f and g. A graph is said to be "commutative" when any arrow composite has the same range and equal domains. As shown in Figure 1(d) below.

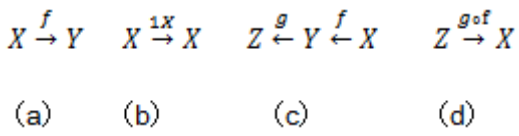


Figure 1. Definition of category

With functors as morphisms, all categories form a larger Cat. So naturally, we have the notion of isomorphism between categories: as isomorphisms within other categories, two categories X, Y are isomorphic if and only if there are two functors F: C→D, G:D→C, so that both sides are identity functors after they are combined, as shown in Figure 2 below:

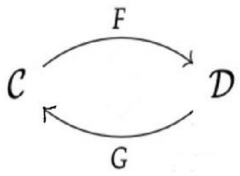


Figure 2. Category isomorphism

(3) $G \circ F = \mathrm{id}C,\ F \circ G = \mathrm{id}D$

The concept of "isomorphism" is defined by a compositional functor, a thinking way that is consistent with the philosophy of category theory, which defines mathematical structures in terms of morphisms. In other words, the basic ideas of category theory reflect the way we organize the structure of information. The idea that morphism represent category has been widely used in the foundations of mathematics. For example, a subgroup can be equivalently regarded as a single group homomorphism, and the quotient group (or normal subgroup) of a group can be equivalently regarded as a full group homomorphism; thus, a subset can be equivalently defined in terms of injective, a quotient can be equivalently expressed in terms of surjective. Developing this idea, the concept of homology which measuring the difference between the exact sequence and the surjective of the category can be defined.

From the above definition, the basic idea of category theory is to study the properties of an object through the relationship between objects. This relationship is usually defined by morphism, so that morphism reflects the structure of mathematical objects. "Structure" refers to a class of mathematical objects described by axioms, and mathematics is concerned with "structure" rather than a specific "object". Bourbaki believe that the structure in mathematics is divided into three categories: algebraic structure (group, ring, field), order structure (partial order, total order) and topological structure (limit, continuity, connectivity, neighborhood). With the improvement of

the abstraction degree and the expansion of the field of mathematics, the structure itself is the "category theory" as a study object.

Category theory is the study of mathematical structures in a concise, general and abstract way. For example, in topology, a doughnut is the same as a coffee cup. According to category theory, they are isomorphic in a category Top, the object is a topological space, and the arrow is a continuous map (i.e., continuous transformation), so a doughnut and a coffee cup are homeomorphic.

The advantages of category theory are as follows:

A very obvious trend in the development of modern science is the division between disciplines; knowledge in different fields seems to be divided more and more finely, and it is more and more difficult for us to have an overall understanding of it. At the conceptual level, category theory unifies definitions and concepts from different branches of mathematics. Category theory has partially unified the division of mathematics at the conceptual level, and it has found the same conceptual basis among different branches of mathematics.

It’s promising to study difficult interdisciplinary scientific problems applying the structural ideas of category theory. For example, physicists apply category theory to solve the problem of topological condensed matter. A good study of problems in mathematical logic. Applying the structured thinking mode of category theory, the structured thinking mode of the brain is researched, and through the Bi-interpretation perspective of the category theory, the structural characteristics of the brain's thinking is interpreted.

3. CATEGORY THEORY MODEL OF STRUCTURED THINKING

3.1. Transcendence of Creative Thinking

As early as in ancient Greece, the study of human rational thinking has already begun and established. Plato vividly expresses his philosophical "idea" view through the cave metaphor, in which the "idea" is seen as a being that illuminates everything. And to reach knowledge of this idea, mind must go through a series of turns, from shadows to things, to firelight, to things outside the cave, and finally to the sun. The path of individual liberation from opinions and phenomena is the same path [7].

Parmenides was the first to put forward the schema of "thought" (rational thinking) which is opposite to feeling. This is the first real thinking model in the study of human reason. Later philosophers including Kant modified and developed this model.

As shown in Figure 3, Parmenides classified objective matter and subjective spirit, and proposed two philosophical categories of "existence" and "non-existence", "feeling" and "thought". In it, what can be "thought is the same as that which can be", and feeling and non-being are the same. Parmenides gave a negative answer to the question of whether there is a connection between existence and non-existence, feeling and thought. Rational cognition leads to the "path of truth", while perceptual cognition leads to It is the "opinion road", the two roads are diametrically opposed [8].

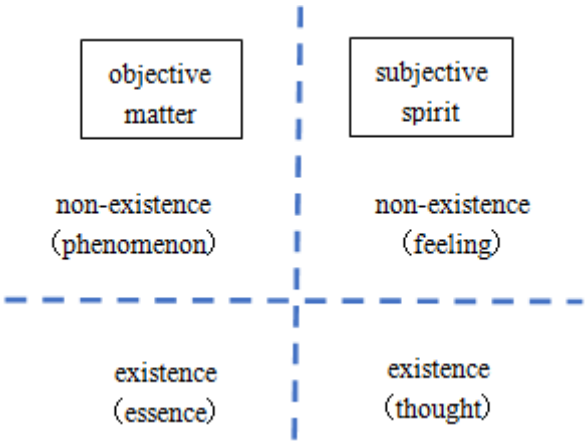


Figure 3. Parmenides' model of rational thinking

After the Renaissance, modern philosophy, physics, mathematics, and neuroscience have developed by leaps. The cognition of rational thinking has undergone a huge change, and Descartes put forward the famous slogan "cogito, ergo sum (I think, therefore I am)", which is logically impossible to refute, that is, all the world of experience, all senses and all idea are some kinds of illusion. But the existence of this activity itself indicates a being, which is "I", and rational cognition of the world can only proceed from doubt. From the perspective of human philosophy, Cassirer, modern philosopher and scientist, regards man as a symbolic animal, and only in man, the question of possibility and reality can be realized. Only in man, mythical, religious, linguistic, artistic, historical and scientific symbols can be created and applied. Man use symbols to create culture, and the creation and using of symbols is the difference between humans and animals [9].

In the 20th century, with the establishment of the two major scientific theories of relativity and quantum mechanics, people have a deeper understanding of the characteristics and structure of scientific theories. Now generally, it is believed that the structure of a scientific theory consists of three elements: concepts; basic principles, linking to these concepts; the logical conclusions, i.e., various specific special laws and predictions derived from these concepts and principles. Among them, basic concepts, basic principles and propositions constitute the core elements of theory.

Einstein said, "in the structure of my thinking, written or spoken words do not seem to have any role... The various concepts that appear in our thinking and our verbal expressions are logically free creations of the mind, they cannot be derived inductively from sensory experience. The reason this is not so easy to notice is only because we are so accustomed to associating certain concepts and their relations (propositions) with certain sensory experiences in such certainty that we are unaware that such a logical an insurmountable chasm that separates the world of sensory experience from the world of concepts and propositions" [10].

According to Einstein, "There is nothing that can be said a priori about the formation of concepts and how they are related to one another, and how we oppose these concepts to sensory experience. Comparing these with the rules of the game, in the game, the rules themselves are arbitrary, but the game is possible only if they are strictly followed. But such rules are never final, and they will only be effective if they are obeyed in a specific field"[11].

Reviewing the historical development of rational thinking, it’s obviously that the transcendence of brain thinking, beyond the world of experience, and even beyond human intuition, which guided scientists to establish various scientific theories.

3.2. Category Theory and Natural Transformations

In order to study the structural characteristics of brain thinking, it is necessary to further study the structural characteristics of scientific theories, that is, to categorize and axiomize scientific theories. In this way, from the perspective of categorization, the isomorphic characteristics of theoretical models of different disciplines and the categorization characteristics of brain thinking can be shown.

Definition 2:

A scientific theory category C includes:
(1) A collection T of a set of objects; the set contains an operator O and its operation rules;
(2) A collection of morphisms, in which morphisms (or arrows, functors f) f: T→T ； f is said to be an automorphic functor induced by object collection. Let Hom(T,T) denote the entirety of arrows from object T to T.

Definition 3:

A scientific theory is a Hom(T,T).
The conception of scientific theory is defined through the predicate "is a Hom(T,T)". As far as scientific theory is concerned, the predicate Hom(T,T) represents a mathematical structure, which is the basic structure of scientific theory. If only the mathematical structure of this theory is considered, this basic structure is sufficient to explain the characteristics of the theory itself.

3.3. The Category Theory Model of Structured Thinking

Definition 4:

F,G: A→B is a functor, a natural transformation, or a morphism t:F→G, is collection of t_x:x∈Ob(A) such that:
(1) t_x:F(x)→G(x) is a morphism in category B, and
(2) For each morphism in A (internal operation in A) ϕ:x→y , follow the commutative diagram is shown in Figure 4:

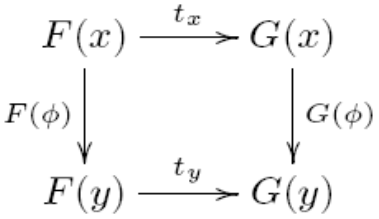


Figure 4. Natural transformation

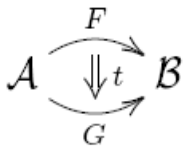


Figure 5. simplified commutative diagram

Sometimes, commutative diagram is simplified as shown in Figure 5, indicating that t:F→G morphism.

The morphism between functors is named the natural transformation of the functor. There is a mathematical structure of natural transformations between functors, that is, a set of objects, morphisms, and morphisms between morphisms, which is called a 2-category. A general category is called a first-order category. For any category K of order 2, there is also a natural equivalent notion that is weaker than isomorphism. For two objects X, Y in any K, the two objects are

equivalent in K, if the composite map of two morphisms f: X→Y, g: Y→X and the identical map meet the demand of definition of higher-order isomorphism.

By investigating the process of scientific theory discovery, the category theory model of structured thinking of scientific theory is to be established. The father of modern science, Galileo, faced a lot of challenging problems when he created new physics, such as the following:

1. How to find simple laws from the complex appearances, and what principles and tools can be based on? 2. Compared with the old theories, how can the proposed laws make breakthroughs in the interpretation of experience? 3. What kind of mathematical properties (structural) should the found laws have?

The first question, Galileo, combined with his own study of motion, believed that the natural world itself is simple and harmonious. The language of mathematics is simple and clear, and scientific research should follow this language to explore the secrets of nature. This grand book (universe) is written in the mathematical language, and its characters are triangles, circles or other geometric figures, without which it is impossible to humanly understand a word; without these one is wandering in a dark labyrinth [7].

The second question, Aristotle's theory is closely integrated with people's daily experience. Compared with Galileo's new theory, we even can say that Aristotle's theory has more support from empirical evidence of various motions on the earth's surface. Galileo's argument revealed that existing experience must be reinterpreted critically, thus allowing his theory to be reborn from the new interpretation. Galileo's interpretation also takes man out of the Aristotle's cave of sensory experience.

The third question is very crucial, even determining whether the new theory can become a scientific theory. To answer this question, Galileo proposed the principle of inertia (relativity) principle and Galileo transformation. The principle of inertia became the foundation of physics, and this principle evolved into the basic idea of modern physics: Symmetry determines the equations of physics.

The Ox'y'z' coordinate system moves at a constant speed u relative to the Oxyz system. When t=t'=0, O coincides with O'. The space-time correspondence between the two systems is

$$\begin{cases} t = t' \\ r = r' + ut' \\ v = v' + u \end{cases} \quad (1)$$

This is the Galileo transformation. The transformation is written in matrix T, T connects the two systems and keeps the corresponding physical phenomena consistent in the two systems, which is the isomorphism of the two systems, see Definition 4 and Figure 4.

Similarly, if we propose a different transformation relation (non-isomorphism functors), then a different scientific theory is proposed. For example, the transformation is Lorentz transformation, then the special relativity theory is proposed. These theories are isomorphic (physically equivalent theories), if the proposed functors have natural transformation which have properties of composite mappings and identity mappings. In Fig. 5, the A coordinate system and the B coordinate system are isomorphic if the transformation is Lorentz transformation, and they are physically equivalent special relativity.

Another example of the pinnacle of human creative thinking is the introduction of general relativity. According to the special theory of relativity, all inertial reference frames that move at a constant speed relative to each other are of equal weight, and the laws of physics have the same form in any inertial reference frame. Should the laws of physics in an accelerating frame of reference also be in the same form as those in an inertial frame of reference? Einstein also often thought about the problem of "gravity ": how to include gravitation in the framework of relativity? [12]

Einstein devised a thought experiment in which a person in an elevator could not tell whether he was ascending at a uniform acceleration or in the gravitational field of the ground. The reason for this phenomenon is that we cannot distinguish gravitational fields from inertial fields by any physical experiment. Gravity is very different from other forces, such as electricity. Because we can't cancel out electricity with something like acceleration! But, why can gravity be eliminated? Perhaps gravity should not be regarded as a force at all, but rather as a property of curved space-time. From the idea that gravity is equivalent to acceleration, another surprising conclusion can be deduced: gravity can be eliminated by choosing an appropriate acceleration reference frame. The idea of gravity as a property of space-time led Einstein to the general theory of relativity [11]. With the equivalence principle, any non-inertial frame can be converted into an inertial frame, as long as the influence of a gravitational field is properly processed.

This inference is crucial in the development of general relativity, and its significance is equivalent to the Lorentz transformation in special relativity. From the viewpoint of equivalence of the second-order category, if the transformation F is decomposed and expressed as F=L+A, where L is the Lorentz transformation and A is the additional acceleration vector (or gravitational field acceleration), then F=L+A still holds in Figure 5. The corresponding physical phenomena in the two systems A and B are consistent, that is to say, the two systems are isomorphisms of order 2. in Definition 4. Physically, it means that the laws of physics in the local inertial frame can be extended to the laws of physics in any reference frame through generalized space-time coordinate transformation. Einstein took the equivalence principle as a starting point, and finally established the general relativity following the Poisson equation of the gravitational field and Riemann geometry.

Functors and their natural transformations fully demonstrate that the brain has the transcendence of creative thinking, and such transcendence is manifested everywhere in the process of major scientific discoveries. A scientific theory has its specific mathematical structure, and this mathematical structure originates from the free creation of scientists. In the process of free creation, scientists consciously or unconsciously apply the ideas of category theory. The structured view of category theory simplifies the complex and tentative thinking process, making the thinking process clear and structured. Therefore, category theory is very important in knowledge expression and reasoning. The thinking way based on category theory has gradually become a typical paradigm in the innovative thinking.

4. CONCLUSION

Most of the research on innovative thinking is based on neuroscience. This paper analyzes the limitations of existing research methods from the perspective of category theory, and proposes research model on innovative thinking based on category theory. This paper applies the thinking method of categorical equivalence to analyze the creation process of two landmark scientific theories in the history of science, the case is real, and a clear mathematical argument is provided.
Category theory studies structure which is formed by the relation between objects, so the focus of research lies in the relation between objects. Concepts in category theory such as morphism, functor, natural transformation, and isomorphism are representation of relation. In a category, the

object only needs to satisfy the morphism relation in it. Category theorists do not need to point out or care about what constitutes an object and a single object has no significance in a category. The relation between objects is the focus of category theory.

Category theory deals with mathematical concepts in an abstract way, turning these concepts into sets of morphisms. If we have two theories each represented as a category whose object is the theoretical model, the morphism preserves the structure of the model in a suitable sense. If there is an equivalence relation for these categories, we can say that the two theories are categorically equivalent.

Many scientific theories can actually be described as a collection of certain mathematical structures. There are two main reasons for applying the equivalence of category theory in the study of scientific theories. First, the inherent properties of scientific theories requirement, for example, the theory has property of coordinate transformation invariance (isomorphism). Second, the equivalence idea of category theory can simplify theoretical models, so that the model can be greatly simplified and the research work can be achieved easily. The structured thinking way has great revelation and help to master and create new scientific theories.

REFERENCES

[1] Shen W B, Yuan Y, Luo J, et al. Neural basis underlying creative essentials of wisdom (in Chinese). Chin Sci Bull, 2015, 60: 2726–2738, doi:10.1360/N972014-01155.
[2] Carlsson, I., Wendt, P. E., & Risberg, J. (2000). On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects. Neuropsychologia,38(6), 873–885.
[3] Chavez, R. A., Graff-Guerrero, A., Garcia-Reyna, J. C.,Vaugier, V., & Cruz-Fuentes, C. (2004). Neurobiology of creativity: prelimary results from a brain activation study. Salud Mental, 27(3), 38–46.
[4] Chavez-Eakle, R. A., Graff-Guerrero, A., Garcia-Reyna, J. C.,Vaugier, V., & Cruz-Fuentes, C. (2007). Cerebral blood flow associated with creative performance: A comparative study. Neuroimage, 38(3), 519–528.
[5] Colombo, J. A., Reisin, H. D., Miguel-Hidalgo, J. J. & Rajkowska, G. (2006). Cerebral cortex astroglia and the brain of a genius: A propos of A. Einstein's. Brain Research Review, 52, 257–263.
[6] Tom Leinster. Basic Category Theory. Cambridge University Press［M］2014. 9-22
[7] James Adam. The Republic of Plato. Cambridge University Press［M］ 1902.
[8] Bertrand Russell. A History of Western Philosophy. Simon & Schuster［M］1945.
[9] Ernst Cassirer. An Essay on Man. Yale University Press［M］2021.
[10] Liangying Xu. "The Collected Works of Einstein" Vol. 1. Commercial Press [M] 1976.
[11] https://einsteinpapers.press.princeton.edu/
[12] Tianrong Zhang. General Relativity and Riemann Geometry Series [J]. Physics. 2015.(44)-2016(45)

AUTHOR

Tong Wang, was born in January 1,1974, associate professor, philosophy doctor, graduated from Tianjin university, China. From 2013 to 2014, researched in Australian national university as visiting scholar.

AN INTELLIGENT SYSTEM TO AUTOMATE THE DETECTION OF ONLINE CHEATING ACTIVITIES USING AI AND CONTEXT AWARE TECHNIQUES

Qinyuhan Zhao[1], Mingze Gao[2] and Yu Sun[3]

[1]Pacific Academy Irvine, 4947 Alton Pkwy, Irvine, CA 92604
[2]University of California, Irvine, 204 Aldrich Hall. Irvine, CA 92697
[3]California State Polytechnic University, Pomona,
CA, 91768, Irvine, CA 92620

ABSTRACT

In the environment of online courses and online exams, cheating in online courses is prevalent [1]. To better ensure fairness in exams, schools and educational institutions need to use technology to detect and deter cheating [2]. Starting from practical application, this paper discusses 3 different methods to detect cheating behavior, and proposes a new way. for online exam supervision.

KEYWORDS

Machine Learning, Audio Detection, Context-Aware.

1. INTRODUCTION

At the end of 2019 a disease called covid-19 is sweeping across the world, and it has soon become a global disaster [3]. After that people on earth entered the age of pandemic, where people stay at their houses to prevent the spread of the disease. Shortly after people entered the pandemic remote learning came into people's sight and it has changed people's understanding of online learning drastically [4]. We are amazed at how advanced our technologies have become so that it allows all of us to meet online to study even though at our own house. Huge progress has been made by humans, but still, there are a lot of flaws in this system.

As soon as the pandemic entered the end of the first year I noticed a serious problem, a problem that's gonna put the whole human race on a string between life and death. I found out that many of my friends and classmates that I know are all cheating on their quizzes and tests to achieve better grades, and this perfectly demonstrates the flaws of online learning [5]. Different from in-person teachers were able to check on the environment around students in real-time, but when it comes to online quizzes and tests teachers are not able to see the environment around the students to help them demonstrate academic honesty. After I realized this has become a serious issue out of curiosity I did a lot of research online and the result shows the rate of academic dishonesty that occurred during the pandemic has increased significantly in comparison to in-person studies. This "blind spot" from the teachers gives a lot of students chances to cheat during their exams, and to put this in perspective soon there will be surgeons that cheat on their biology class in high school performing surgeries for us. I realize this going to become a national crisis that must be stopped so I decided to do something about it. I started to look into ways to prevent. online cheating.

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 67-74, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121707

There are a lot of apps or websites that offer ways to prevent cheating but none of them was ideal enough to make sure the student doesn't get a chance to cheat [6]. Ever since that, I made up my mind to create a program of my own to prevent the issue of online academic dishonesty.

There are actually a lot of methods and applications on the internet to prevent academic dishonesty/ online cheating. However, many of the anti-cheating techniques and systems that have been developed are not working effectively. An obvious example is College board uses a lockdown browser as their method to make sure the student can't cheat on their test, but in reality, the students that are taking the test might have their book opened and possibly another device on the side helping them search for formulas they need to use for that certain question. By just using a lockdown browser students might also just ask for their friend's or tutor's help during the exam. The lockdown browser sounds prominent with accountability but in reality, there's no guarantee that the student doesn't cheat because there are too many flaws in this method that allows the student to have the chance to cheat on their exam. Another method that I usually see online is that during the exam I will take a picture of the test taker to see if they are the actual person that's taking the exam. I think this is a smart idea but it's not ideal. The benefit is that the exam will know whether the student is actually taking the exam or other people that are helping the student taking the exam, but on the other hand, the student might communicate with other people during the exam through another device to establish communication with their friends or tutor to help them during the exam. The intention of this system and methods are good, but they are not working ideally as they are expected to be because there are flaws and imperfections within them.

My method is to utilize a sound monitor and mouse movement tracking technology in our system to ensure cheating is impossible during the exam [7]. The current existing method is to lockdown browsers but students might cheat by getting outside help at their house. My sound monitoring system ensured that students cannot talk to other people during the exam. If they get any outside help the system will capture the voice and the student will be determined as cheating.

We set up three experimental groups and the first one is the control group by not using any anti-cheating application during the exam, the second group uses a lockdown browser during the exam, and the third group uses the application I developed. We first placed the three candidates in three separate empty rooms, and then we gave them three exact same exams and another computer. The first candidate was able to cheat easily because it doesn't have any application or any restriction on him. The second candidate had a locked-down browser on the test he had so he used another computer to call his friend for help and was able to easily pass the exam too. The third candidate used my application and was trying to do the same as what the second candidate was trying to do with the other laptop but he was determined to be cheating because the sound system recognized the candidate was talking to someone else. This shows my application does work.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the experiment and designing the sample; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the related work in Section5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follow.

My largest challenge is to learn the programming language python because without it I can't even start my project [8]. At first, I didn't know what to do, so I asked to do some research online. I found out that there are many education centers that can teach python. After I checked most of them I decided to go to coding minds and study python there. At first, I was struggling with the coding process and I almost gave up, but when I think about how my program might benefit the whole society. I kept learning python until I was capable of doing the project, and eventually, I accomplished my goal.

3. SOLUTION

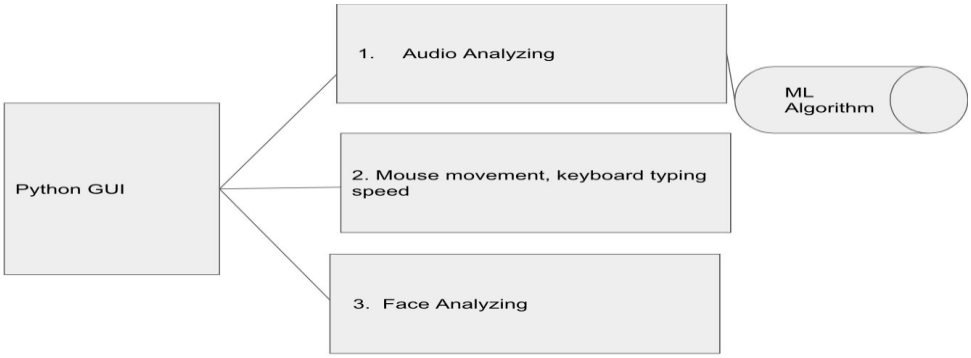


Figure 1. Overview of the solution

The cheating detection system is accomplished by using the language of Python that the system supports based on three main features which are audio analyzing, mouse movement with keyboard typing speed and face analyzing (Figure 1). The current audio analyzing system takes steps to calculate the user's sound wave amplitude that compares with the default value detecting any conversation or discussion behavior during the test. In the future, the system is planning to collect and train the data from users at the same time, where the machine learning algorithms will be implemented in order to predict and analyze the sound from users [9]. By doing this, machine learning will capture and learn the relationships between data providing the result of the detection. The process of checking the mouse movement with keyboard typing speed also plays an important role to catch any abnormal activities during the exams or quizzes. For example, if the mouse moves too often or the keyboard types too fast, those frequent signs will cause an alert and warning for cheating detection. In the meantime, the system will activate the face capture that monitors all the sensitive facial movements such as mouth movement, eye movement, head movement, etc. By combining all three results from the detection features, the system will make a decent result based on the user's activities under the detection. The three detection features have minimized the deviation of the result. The system instructions are easy to follow, and users will download the software on their laptops or desktops. Once the software is successfully downloaded and opened, tests will be provided by the instructor. The detector will be activated automatically once the exam is started. When users are finished taking the exams, students are required to click the "Submit Exam" button, then the system will provide the relevant results to the instructor.

Two algorithms have been implemented into the feature of mouse detection where the algorithm will check two factors, boundaries and rapid movements. When the mouse has either vertically or horizontally moved outside of the testing window, warnings will appear. In addition, if the mouse has moved too fast that exceeds a certain range of speed, this behavior will be detected as a rapid movement which causes warnings as well. The cheating will be detected based on 2 or more combinations of warnings. For example, if a student tries to open a browser during the exam, he

or she quickly moves the mouse outside of the testing window. This kind of behavior will cause at least two warnings that the mouse goes over the boundary and the moving speed is exceeded. The warnings will cause the cheating detection to be activated, and the system will immediately stop the exam, generating the report to the instructor at the same time [10]. As shown in Figure 2, warning detection are implemented based on different factors. The algorithm uses a number of four warnings to determine both mouse out-of-bound and rapid movements. If the warning has been detected, it will be set to TRUE. The 'warning1' will be TRUE if the mouse moves too fast to the left or right, and 'warning2' will focus on the speed of the mouse moving up and down. In the meantime, the 'warning3' will be set to TRUE if the mouse moved out of the bond horizontally, while 'warning4' checks whether the mouse moved out of the bond vertically. If 2 or more warnings are being set to TRUE, the cheating will be detected. For example, once warning1 and warning3 are set to TRUE, the algorithm will shut down the exam, and the student will be prohibited from taking the test. In addition, if warning1, warning2, and warning3 have been detected, the exam will be ended as well, as long as there are at least two warnings set to TRUE.

```
def detect_warnings(self, x, y):
  warning1 = bool
  warning2 = bool
  warning3 = bool
  warning4 = bool

  if self.x != 0 and self.y != 0:
    # Rapid Movements
    if ((self.x - x) > 225) or ((x - self.x) > 225):
      self.mouse_movements.append("Warning-1")
      warning1 = True
    if ((self.y - y) > 225) or ((y - self.y) > 225):
      warning2 = True
      self.mouse_movements.append("Warning-2")
    # Out-of-bounds
    if x >= 448 or x <= 4:
      warning3 = True
      self.mouse_movements.append("Warning-3")
    if y >= 411 or y <= 4:
      warning4 = True
      self.mouse_movements.append("Warning-4")
  self.x, self.y = (x, y)
```

Figure 2. Mouse Movement Detection Code

The sound detection feature will make sure that the environment around the student is quiet enough, so that it detects whether the student is talking to someone else during the exam or not. The sound detection will check the amplitude of the sound waves instead of reading the words or sentences from the user. The system checks sound waves every 100ms. If the current sound wave. is greater than the initialized sound amplitude, the algorithm will identify that the student has cheated during the exam because of talking. Based on Figure 3, the proper measurements and global requirements for audio recording have been initialized for sound identification. For example, if the collected sound amplitude from the student is greater than the initial value of 'VOICE_THRESHOLD' which means the noise or sound is louder than normal. Since the value of 'VOICE_THRESHOLD' has been carefully calculated, it avoids the small noise or sounds during the exam such as typing, sketching, clicking, etc.

```
VOICE_THRESHOLD = 0.28
FORMAT = pyaudio.paInt16
SHORT_NORMALIZE = (1.0/32768.0)
CHANNELS = 1
RATE = 44100
INPUT_BLOCK_TIME = 0.05
INPUT_FRAMES_PER_BLOCK = int(RATE*INPUT_BLOCK_TIME)
```

Figure 3. Initialized Measurements for Sound Detection

By looking at Figure 4, it shows the main process of sound detection where the algorithm will first check to select the appropriate input device [11]. If there is no preferred input, the system will set the input device as default. Once the input has been set, the function of 'open_mic_stream' will set all the needed requirements such as format, channels, rate, etc. When the input device is ready and the microphone stream is opened, the sound detector will check the amplitude every round to compare with the 'VOICE_THRESHOLD' that has been initialized. The algorithm will always return TRUE if the sound exceeds the maximum amplitude. If the current amplitude does not encounter any problem, the next sound amplitude will be checked after 100ms.

```
def find_input_device(self):
    device_index = None
    for i in range( self.pa.get_device_count() ):
        devinfo = self.pa.get_device_info_by_index(i)
        print( "Device %d: %s"%(i,devinfo["name"]) )

        for keyword in ["mic","input"]:
            if keyword in devinfo["name"].lower():
                print( "Found an input: device %d - %s"%(i,devinfo["name"]) )
                device_index = i
                return device_index
    if device_index == None:
        print( "No preferred input found; using default input device." )
    return device_index

def open_mic_stream(self):
    device_index = self.find_input_device()
    print(device_index)
    stream = self.pa.open(  format = FORMAT,
                            channels = CHANNELS,
                            rate = RATE,
                            input = True,
                            input_device_index = device_index,
                            frames_per_buffer = INPUT_FRAMES_PER_BLOCK)
    return stream

def detector(self):
    block = self.stream.read(INPUT_FRAMES_PER_BLOCK, exception_on_overflow = False)
    amplitude = get_rms(block)
    if amplitude > VOICE_THRESHOLD:
        return True
    else:
        return False
```

Figure 4. Sound Amplitude Detection Code

The third detection feature is face movement analyzing which is still in progress [12]. The plan is broken into two sub-parts. The first part will be setting the dimensions of the camera to monitor the student's face within specific boundaries by considering enough space for normal movements. The second part is designing a face movement capture system that catches all the sensitive facial movements. For example, if the eyes are looking outside of the screen, the system will identify it as a warning. If the eyes are looking outside of the screen for longer than 5 seconds, then the algorithm will detect this behavior as cheating. Similar algorithms will apply to head movements and mouse movements as well, which checks the degree of head rotation and frequency of mouth movement during the exam.

4. EXPERIMENT

4.1. Experiment 1

In order to evaluate the accuracy of the mouse movement detection. The first experiment contains three main parts which test various speeds of mouse movement: slow speed, normal speed, and

fast speed. Each test will take different combinations such as slow speed with exceeded boundaries or fast speed with exceeded boundaries. By taking a close look at Table 1, the results supported that the system was sensitive to any moving speed of the mouse that once the mouse has moved outside of any two boundaries, the system will easily detect that the user has moved the mouse out of the testing window. However, the experiment also pointed out an issue of the algorithm that when a mouse was slowly moved out of one edge, the system did not detect this behavior. Recall back to the mouse detection algorithm, the cheating behavior will only detect when the user has touched at least two warnings. Since the user slowly moves the mouse out to a specific boundary, the rapid movement detection did not activate, the only warning was the mouse exceeded one of the boundaries, which caused the system not to be satisfied with the base case resulting in a not detected result. Based on this issue during the experiment, it also brought up a great idea for the improvement of the system, in which the improved algorithm could limit the moving area of the mouse clicker. For example, once the exam has started, the mouse can only move within the testing window, once the mouse clicker touches any boundaries, it will be relocated to the center of the testing window. By doing this, the system easily handles the issue of the boundaries that not only locked the mouse within a certain area but also limit students to only interact with the testing information without checking anything outside of the testing screen.

Table 1. Mouse Movement Detection Experiment Result

Mouse Moving Speed	Left Boundary	Right Boundary	Up Boundary	Down Boundary	Result
Slow	Exceeded	Not Exceeded	Exceeded	Not Exceeded	Detected
Slow	Not Exceeded	Exceeded	Not Exceeded	Not Exceeded	Not Detected
Slow	Not Exceeded	Exceeded	Exceeded	Not Exceeded	Detected
Slow	Not Exceeded	Not Exceeded	Not Exceeded	Not Exceeded	Not Detected
Normal	Not Exceeded	Exceeded	Exceeded	Not Exceeded	Detected
Normal	Exceeded	Not Exceeded	Not Exceeded	Exceeded	Detected
Normal	Not Exceeded	Not Exceeded	Not Exceeded	Not Exceeded	Not Detected
Normal	Not Exceeded	Exceeded	Exceeded	Not Exceeded	Detected
Fast	Exceeded	Not Exceeded	Not Exceeded	Not Exceeded	Detected
Fast	Exceeded	Not Exceeded	Exceeded	Not Exceeded	Detected
Fast	Not Exceeded	Not Exceeded	Not Exceeded	Not Exceeded	Detected
Fast	Not Exceeded	Not Exceeded	Not Exceeded	Exceeded	Detected

By considering different cheating methods and applications, it is necessary to consider various anti-cheating strategies. The system contains a sound detection system in order to prevent students asking help from someone else such as calling a friend or discussing with a tutor. The second experiment created a realistic testing environment to test how the sound detection system works. Three separate empty rooms were prepared for three different students. One laptop was placed on the desk, and students were allowed to take their phones. All three students were required to take the same test. There was no anti-cheating application used for student 1. Student 2 was required to use a lockdown browser system, and student 3 was using our system. As shown in Table 2, it recorded students' behaviors, test results, and cheating detection results during the test. Student 1 had no problem passing the test because no detectors were placed on the laptop. Student 2 faced a little problem searching for resources online, but student 2 decided to ask help from his friend by calling a voice chat using the phone, but still passed the test at the end. Student 3 did the same thing as student 2, but he had no chance to cheat on the test because the sound detection system had received a high amplitude of the sound wave and ended the test immediately. Although student 3 did not use the browser searching for the help, the test would still be ended based on the mouse movement detection. By doing this experiment, the cheating detection system had been working in a good way to prevent cheating behaviors. The results from experiments have reached my first round of development and testing expectations where the detection algorithms could deal with the common cheating methods while students are taking the tests.

Table 2. Sound Detection Experiment Result

Candidates	**Laptop Status**	**Phone Status**	**Test Condition**	**Cheating Detection**
Student 1	Easily searched all the answers on websites	Not Used	Passed	Not Detected
Student 2	Not able to open the browser	Used to call his friend	Passed	Not Detected
Student 3	Taking the test	Used to call his friend	Not Passed	Detected

5. RELATED WORK

Razan B. et al mentioned a computer-vision algorithm to prevent cheating behaviors in online exams [13]. The system takes action to detect the eyes movements of the users by taking images.

This idea is similar to our current face capture system that is still being designed. The system will capture any sensitive movement from the user's face but the algorithms have to be implemented carefully where different factors need to be considered into the algorithms such as eyes focusing time on a specific position or the user moving out of the testing position.

Mohammad M. et al identified a great idea on recording pieces of video during the online exam [14]. Once the students have any prohibited behavior, the system will record that piece of video for future checking. By doing this, it saves a big amount of time from the instructor who only checks small pieces of videos instead of watching the whole recording of how students are doing in tests. Related back to our sound detection system, it focuses on the amplitude of sound waves. Based on the recording idea, it could be helped for the improvement of the sound detection system that the system would record a piece of time for the unusual sounds so that the instructor could double check the detection result based on those short recordings.

Ali M. et al discussed an approach of unsupervised learning that iterates and learns from group data in order to predict cheating behavior while taking the tests [15]. Machine learning plays an important role in handling and predicting results. It would be good to keep implementing machine learning into our mouse movement detection system where the algorithm will be more advanced to catch any students who have the same or similar mouse movements as other detected students. By using machine learning, it not only improves the flexibility of the algorithm but also increases the accuracy of each detection in the future.

6. CONCLUSIONS

Not only has the pandemic of Covid-19 caused many of the classes to switch to the online format, there are still other programs or online courses that have been set as remote in order to provide a more convenient learning experience for other countries' students. However, one of the biggest issues during online education is how to effectively prevent the issue of online academic dishonesty. By making this exam cheating detection system, it helps instructors alleviate the issues of online exams. The system contains various aspects to detect any cheating behaviors happening during the exams. Mouse movement detection restricts students from looking up any prohibited resources through the browser or any other information outside of the testing window. Sound detection supports that the exam is taken by students themselves without having a conversation or discussion with someone else. As the face capture system will be finished later, the system will be more comprehensive and checks any sensitive facial movement to avoid students from looking at resources by hand such as textbooks or notes, and it ensures that the student is not moving on the seat during the exam. According to the experiment, it proves that the

system can effectively prevent the issue of online academic dishonesty that catches any students who are having prohibited behaviors while taking the exam. However, the system does have a deviation in some cases. For example, if the student accidentally touched the mouse causing the mouse to move outside of the boundary with a rapid movement, the system would lock the exam right after it, but the student did not mean to cheat. Although this is a small probability, it is still important to consider how to avoid this kind of mistake. The current idea is that the system needs to be improved to count how long the mouse clicker has been staying outside of the boundaries. By taking this improvement, if the student has accidentally touched the mouse, they could have a few seconds to bring the mouse clicker back to the testing window. The cheating detection will be activated only if the mouse has stayed outside of the testing window for certain seconds. Another solution has been discussed in the earlier section that limits the mouse within a moving area. Once the mouse touches any boundaries, the mouse clicker will be pulled back to the center of the window. In addition, an external feature is also considered to be implemented into the system in the future, where the instructor can decide to set the exam to be the open book format so that the mouse movement detection will be closed because students will be allowed to check any resources. By doing this, the improvement improves the flexibility of the system that provides different formats and restrictions for the exams.

REFERENCES

[1] Benigno, Vincenza, and Guglielmo Trentin. "The evaluation of online courses." Journal of computer assisted learning 16.3 (2000): 259-270.
[2] Ball, Nicholas L., Robert E. Wood, and Gove Allen. "USING TECHNOLOGY TO DETECT AND DETER ELECTRONIC CHEATING AT A LARGE PUBLIC UNIVERSITY." Issues in Information Systems 17.4 (2016).
[3] Pokhrel, Sumitra, and Roshan Chhetri. "A literature review on impact of COVID-19 pandemic on teaching and learning." Higher Education for the Future 8.1 (2021): 133-141.
[4] Ali, Wahab. "Online and remote learning in higher education institutes: A necessity in light of COVID-19 pandemic." Higher education studies 10.3 (2020): 16-25.
[5] Raines, Deborah A., et al. "Cheating in online courses: The student definition." Journal of Effective Teaching 11.1 (2011): 80-89.
[6] Mubin, Omar, et al. "Can a robot invigilator prevent cheating?." AI & SOCIETY 35.4 (2020): 981-989.
[7] Kieslich, Pascal J., et al. "Mouse-tracking: A practical guide to implementation and analysis 1." A handbook of process tracing methods. Routledge, 2019. 111-130.
[8] De Smedt, Tom, and Walter Daelemans. "Pattern for python." The Journal of Machine Learning Research 13.1 (2012): 2063-2067.
[9] Mahesh, Batta. "Machine learning algorithms-a review." International Journal of Science and Research (IJSR).[Internet] 9 (2020): 381-386.
[10] Bawarith, Razan Hamza. "Student Cheating Detection System in E-exams." (2017).
[11] Dufaux, Alain, et al. "Automatic sound detection and recognition for noisy environment." 2000 10th European Signal Processing Conference. IEEE, 2000.
[12] Zarkasi, Ahmad, et al. "Face movement detection using template matching." 2018 International Conference on Electrical Engineering and Computer Science (ICECOS). IEEE, 2018.
[13] Bawarith, Razan, et al. "E-exam cheating detection system." International Journal of Advanced Computer Science and Applications 8.4 (2017).
[14] Masud, Mohammad M., et al. "Smart Online Exam Proctoring Assist for Cheating Detection." International Conference on Advanced Data Mining and Applications. Springer, Cham, 2022.
[15] Duhaim, Ali M., Safaa O. Al-Mamory, and Mohammed Salih Mahdi. "Cheating detection in online exams during COVID-19 pandemic using data mining techniques." (2021).

A GRADIENT DESCENT INSPIRED APPROACH TO OPTIMIZATION OF PHYSICS QUESTION

Feihong Liu[1] and Yu Sun[2]

1Crean Lutheran High School, 12500 Sand Canyon Ave, Irvine, CA 92618
2California State Polytechnic University,
Pomona, CA, 91768, Irvine, CA 92620

ABSTRACT

Many people believe that the crouch start was the best way to start a sprint [1]. While it seems intuitive, when the process of running is dissected using specific physical and mathematical representations, the question of "what is the best starting position" becomes harder to answer [2]. This paper aims to examine this phenomenon through a computer science approach inspired by gradient descent. Specifically, this paper aims to maximise the distance covered by a runner in ten steps. Assuming that runners do their best on every step and that their motion is not slowed by friction or air resistance, we will generate a hypothetical environment to study what the best strategy is for reaching the furthest distance within ten steps.

KEYWORDS

Gradient Descent.

1. INTRODUCTION

Imagine a runner sprinting ten steps [10]. On each step, they are applying some force on the ground at some angle [3]. Assuming that air resistance and internal forces are negligible, their motion before they hit the ground can be modelled roughly by a projectile motion [4]. Considering if the runner is trying to maximise their distance, they will not hold back on their force and will try their best on every single step. These reasoning give us our first two conditions: a constant max force and an angle in which this force will be applied for every step.

Realistically, a runner will be affected by various forces such as friction, air resistance, and internal forces. Because these forces can vary significantly (they can vary due to weather, heat, among many other factors), those forces will be ignored in our study. We aim to only explore how the different combinations of angles in which forces are applied will result in varying total distances. Thus, we will assume that horizontal velocity is not lost between every step (the velocity from the previous step will carry forward in the next step) and that the vertical velocity will be reset between every step (the runner will not bounce between each step).

With these conditions in mind, the rest of the paper is organised as follows: Section 2 introduces our initial experimentation; Section 3 analyses the results of Section 2 and proposes potential solutions; Section 4 explores the methods proposed in Section 3; Section 5 discusses the

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 75-85, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121708

implementation of these methods. Finally, Section 6 gives concluding remarks, as well as pointing out the future work of this project.

2. INITIAL EXPERIMENTATION

To get a better sense of how to approach this problem, our situation is simplified to using only 5 angles. To understand a potential approach to this problem, different sequences of angles are generated with their output distance. For example, the list [5, 5, 5, 5, 5] will represent the motion of a runner who applies their force at a degree of 5 at every step. Using a series of algorithms, the distance resulting from this sequence of angles will be generated. Likewise, the list [5, 5, 5, 5, 10] will represent the motion of a runner who applies their force at a degree of 5 for every step other than the last one, which will be an angle of 10. Through an iterative process, all possible combinations of acute angles that are divisible by 5 are generated, and written in a data file. This means that the data point with the angle list [5, 5, 5, 5, 5] is first generated, followed by [5, 5, 5, 5, 10], [5, 5, 5, 5, 15], and so on. To do this, two helper functions are defined to find the distance with a particular list of angles.

```
def Calculate(current_angle, current_x_velocity, current_distance):
    # Physics here
    velocity = 5
    x_velocity, y_velocity = velocity * math.cos(current_angle) + current_x_velocity, velocity * math.sin(current_angle)
    time = 2 * y_velocity / 9.8
    jump_distance = x_velocity * time
    return x_velocity, jump_distance + current_distance


def Final_Calculation(angle_list):
    # Return angle list and the distance
    current_vx = 0
    current_d = 0
    for angles in angle_list:
        angles = angles * math.pi / 180
        current_vx, current_d = Calculate(angles, current_vx, current_d)
    return [angle_list, current_d]
```

Figure 2.1. Screenshot of Helper methods

The calculate function returns the horizontal velocity and distance with some specifically inputted angle, velocity and distance. This function is called repeatedly in a for loop which will iterate through a list of angles. The final distance will be outputted along with the angle list that outputted this distance. The Final_Calculation function is called repeatedly using different angle lists starting from [5, 5, 5, 5, 5] up to [90, 90, 90, 90, 90] in increments of five. It is assumed that the runner is able to accelerate his velocity by 5 m/s. To save space, the best distances with each unique combination of the four five angles are saved in a data file. For example, in the simplest case, the best angle for the following sequence of angles [90, 90, 90, 90]

Since we have a four-dimensional input and one-dimensional output for every data point, we cannot visualise our data in a two-dimensional plane. Thus, a separate function is run to assign each angle list an x value. The list [5, 5, 5, 5] is given the x-value of 0, [5, 5, 5, 10] is given an x-value of 2, [5, 5, 10, 5] is given 18 ([5, 5, 5, 90] is given 17) and so on. All multidimensional data will be represented in this way throughout this paper. After plotting the data points using matplotlib, the following graph is generated [5]:

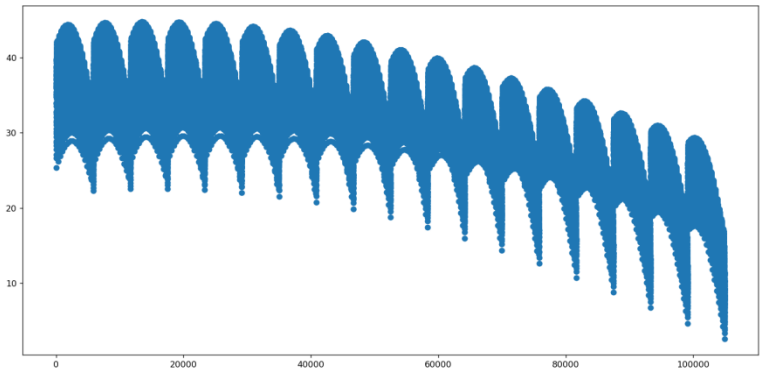


Figure 2.2. Initial Experiment Data

This graph looks roughly like a parabolic curve, composed of smaller curves. Interestingly though, when one zooms in on on curve, we can see that it is made up of smaller parabolic curves similar to this one:

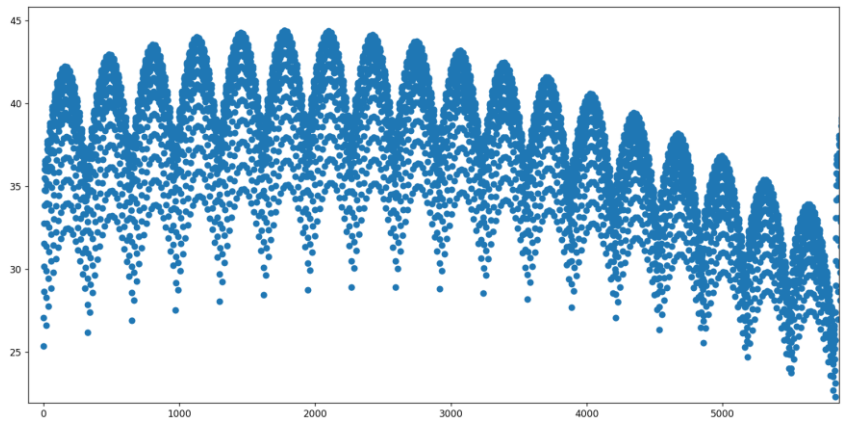


Figure 2.3. Initial Experiment Data (1st Zoom)

If we continue zooming in, we will see the following graphs:

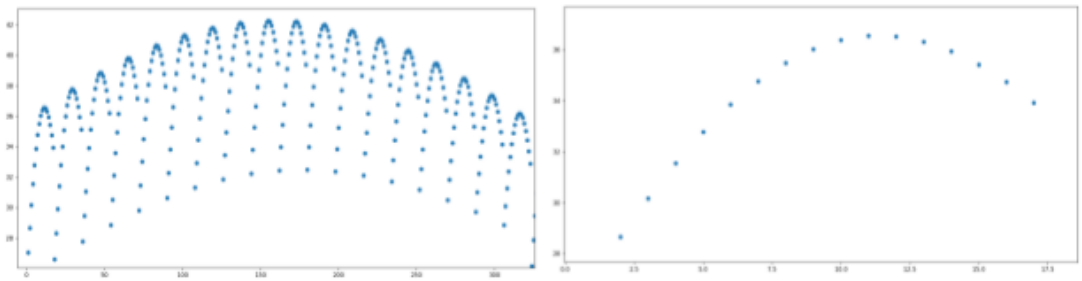


Figure 2.4. Initial Experiment Data (2nd and 3rd Zoom)

Each graph shown above is a zoomed in version of the graph above it. This shows an interesting relationship between each angle. At the smallest level (last image shown above), within each parabola, the first four angles of every point are the same. The last angle of each term is increased by increments of 5, with the first having 5 as its last angle and the last having 90 as its last angle.

Going back up and zooming out from the last graph, we see that each angle corresponds to a specific parabola. The last angle corresponds to the smallest parabola, as shown in the last figure. Similarly, the angle before that corresponds to the second to last figure, which is composed of smaller parabolas. The other angles have similar relationships with the other larger parabolas. It is also notable that there are only four of these parabola patterns because the fifth angle is already taken care of by the way we saved our data, since we are selecting only the optimal fifth angle.

3. POTENTIAL SOLUTIONS

Considering the example above, we can make several observations. First, we see that there are several smaller parabolas that make up the larger parabola in Figure 2.2 [6]. All data in each one of these smaller parabolas all have the same leading angle. This is represented by the following figure:

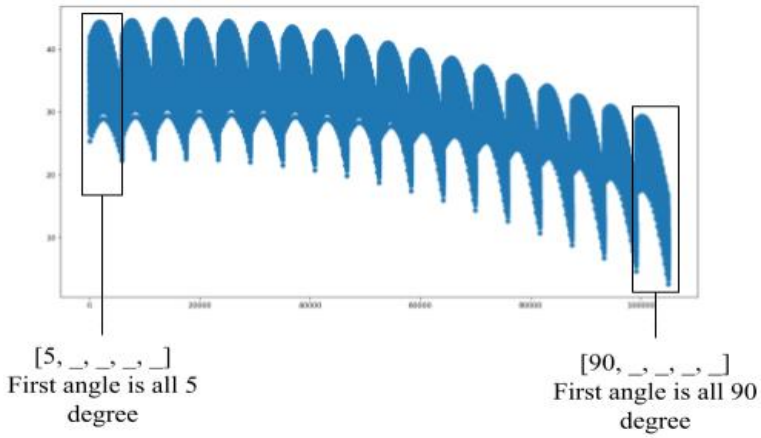


Figure 3.1. Grouping of Overall Data

Considering that these groups come to form a parabolic shape, a good way to represent each large group is to consider the average of its elements. For example, in the five-angle example shown in the previous portion, each group is very complex and composed of many points. By creating a graph of the average value of each group, we have a simpler way of finding the optimal value for this roughly parabolic shape.

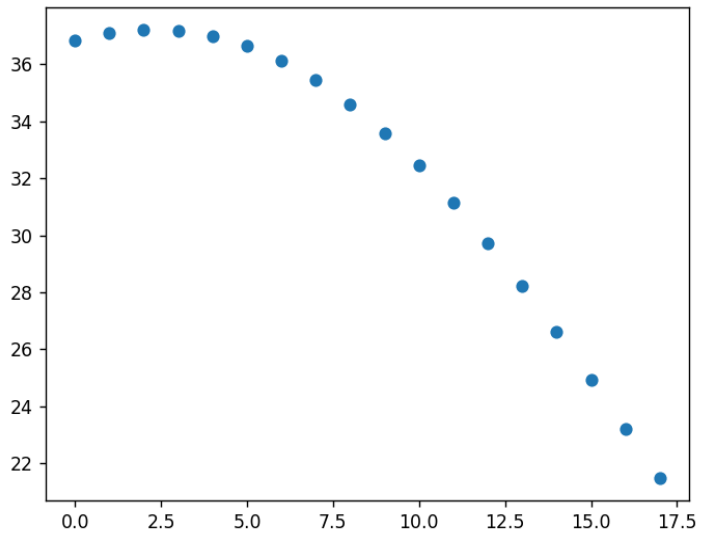


Figure 3.2. Each Group Represented by a Point With a its Average Distance

Thus, utilising this fact, we can generate data with larger increments (i.e. generating data with increments of 30 instead of 5), and find new data points based on calculated data points.

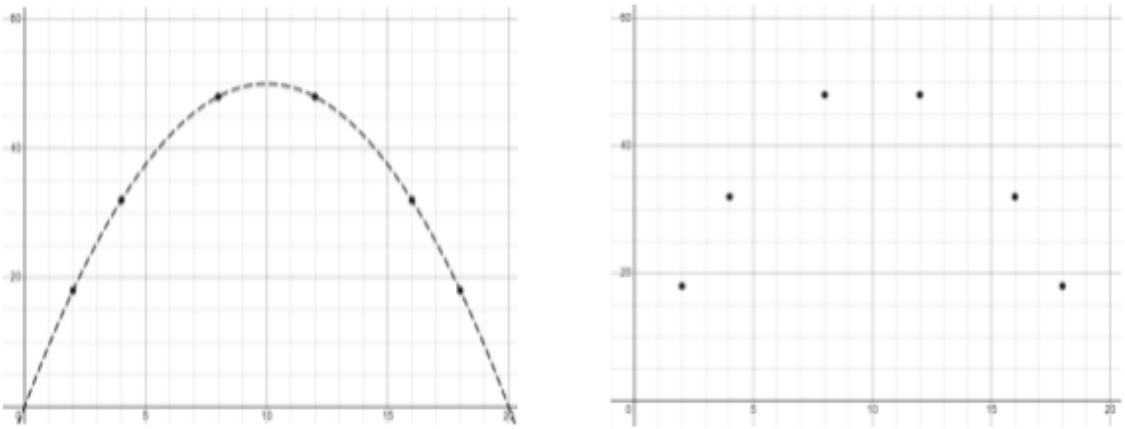


Figure 3.3. Generating Groups With Larger Increments

In the illustration shown above, we have six data points calculated but we are unsure what the optimal value is [7]. This presents us two ways to potentially approximate the optimal solution. We can use quadratic regression to find a quadratic equation that fits the given points or deduce the x value of the optimal by repeatedly comparing the x values of the three data points with the largest y value (this will be explained later on).

Repeating this process for every one of the ten angles. We are able to pinpoint, with relative accuracy and efficiency, the location of the optimal set of angles.

4. PROPOSED METHODS

In this section, we will explore the two methods mentioned in the previous section.

4.1. Regression

Because the general shape for each angle can be roughly modeled by a quadratic curve, recursion seems like the most direct route to finding an optimal solution [8]. However, after experimentation, quadratic regression is shown to have two main issues: relatively large margin or error, and relatively inaccurate models in select situations.

Firstly, when quadratic regression is applied to the average value plot shown in the previous section, we result with the following plot.

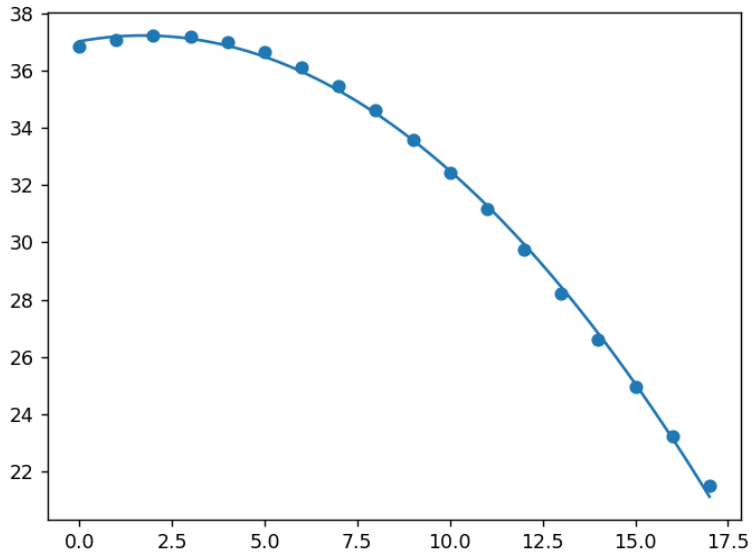


Figure 4.1.1 Regression applied to Figure 3.2

As shown, the plot shown above is slightly inaccurate, predicting that a value around 2.5 (or a sequence of angles beginning with 25 degrees) will produce the optimal. However, the optimal solution produced from the data points appears to be shifted to the right compared to the optimal produced by the regression line. This difference, while seemingly small, has a relatively. important impact in practice. A difference by ten degrees alone could drastically impact the path of the runner. Thus, the strategy of regression is considered to be relatively inaccurate.

4.2. Case Scenarios

Instead of approaching the problem directly with regression, we can take a similar approach by using some observations of several scenarios. For the following case scenarios, consider only the three points with the highest distance value. We are able to ignore the rest because they are unnecessary to find an optimal in these cases. Considering the three angles, we outline two

scenarios: when the three are all ascending/descending or when the angle with the largest distance lies between the other two.

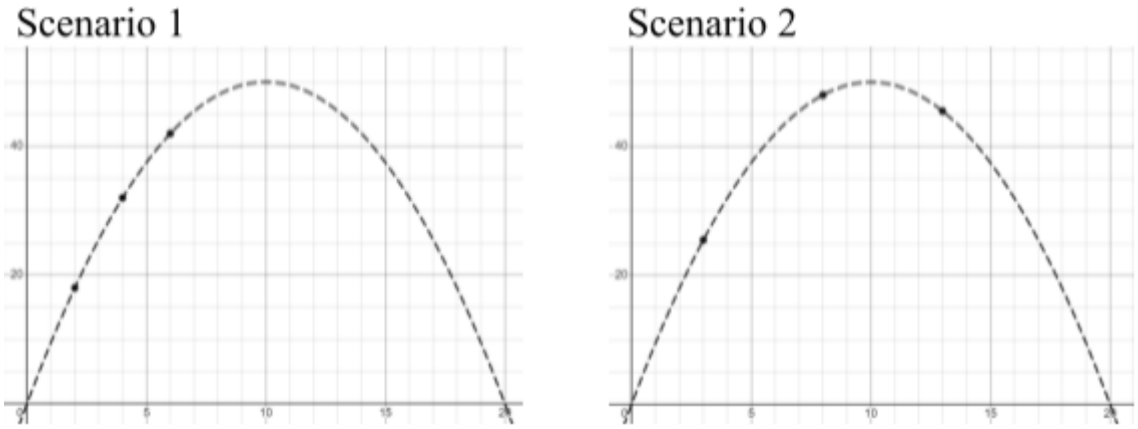


Figure 4.2.1. Two case scenarios

Our goal is to have angles of scenario 2, since we will be able to find another point with the angle equal to the average between the top two. This new data point will, theoretically, have an average value greater than the lesser of the two points. This process of averaging the top two angle measures can be repeated until the margin of error is small enough. However, if it happens that we come across scenario 1, we can simply replace the point with the smallest distance with a point to the right of the point with the largest distance. Through a combination of this process, we will, theoretically, be able to reach an approximate optimal solution with a margin of error of our choosing. The implementation of this method will be discussed below.

5. IMPLEMENTATION

Firstly, a program is run to generate an appropriate amount of data and save it into a data file. This will be important later on to more accurately localize the optimal sequence of angles. Because of the runtime for generating data, increments of 15 are used between each angle. Furthermore, for simplification, only the best cases for unique combinations of the first three angles will be saved. The rest will be optimized through the Group class shown below.

Figure 5.1. Screenshot of code 1

Simply, when a group object is initiated, it contains a list of angles that begin with some values. Other values in the list will be defined as zero, which will be important for the it's two functions. The zero values of this array will be values in which the functions will be replaced with other values. For example, if [15, 15, 15, 0, 0, 0, 0, 0] is passed into the function, it will alter the fourth to eighth angles. This is used for the main function, Custom_Return_Value. This function will iterate through all angles incrementing by a value determined by the input n. For example, if a Group object is defined with [15, 15, 15, 0, 0, 0, 0, 0], the function will look through angles [15, 15, 15, 0, 0, 0, 0, 15], [15, 15, 15, 0, 0, 0, 0, 30], etc and return the angles with the furthest distance. However, an additional statement is implemented into this function to save run time. Consider that a ball is climbing up a parabola, a decrease in height will only happen if the ball has passed the vertex. This is essentially the logic of the statement, which will exit the for loop immediately when a decrease in "height" (distance in our case) is detected. The following is the generated data.

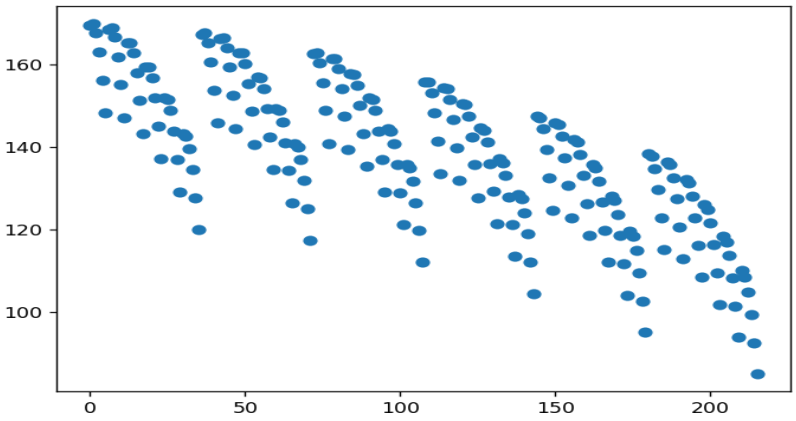
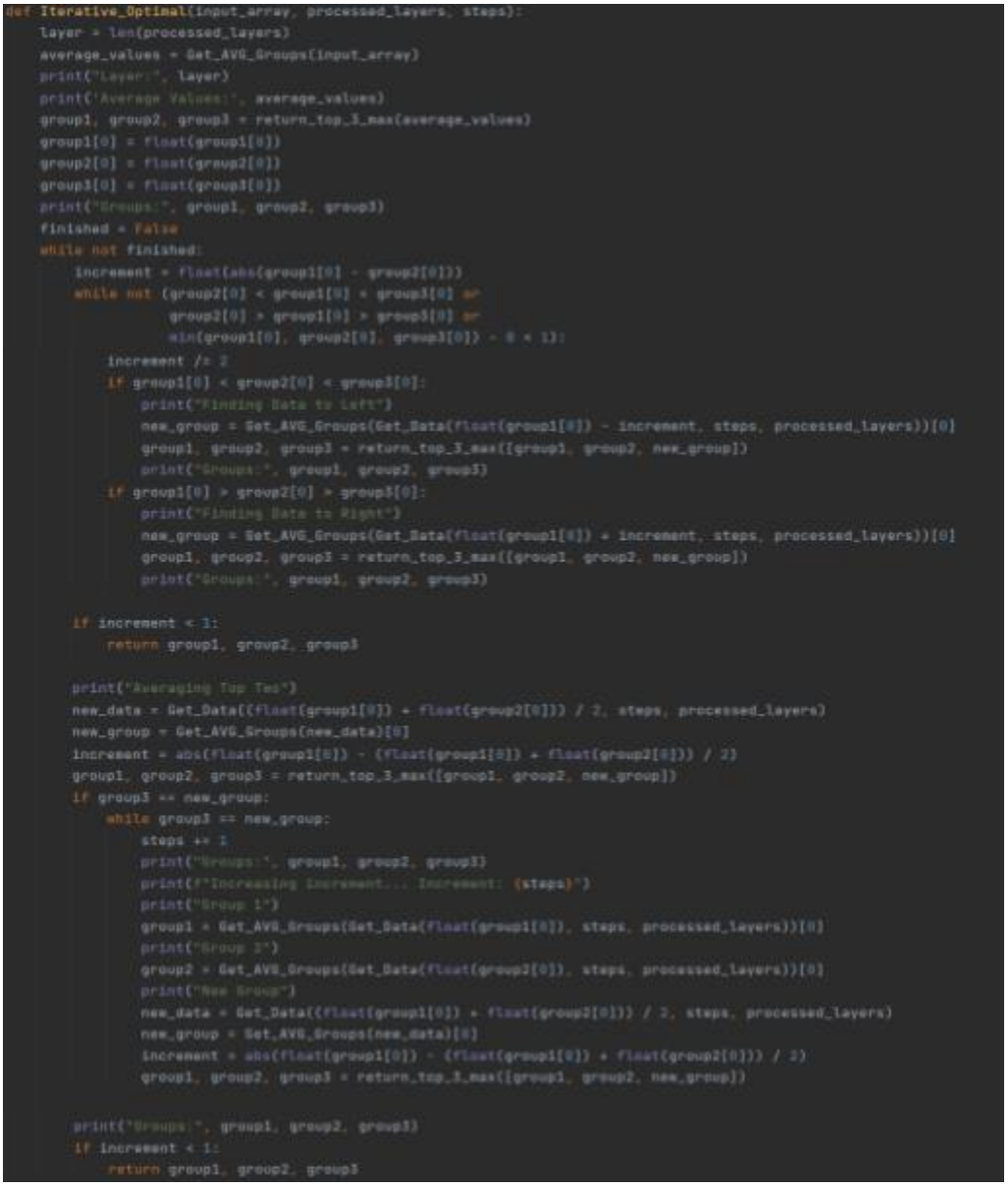


Figure 5.2. Generated data (not averaged by groups)

After this data is generated, an iterative processing function is called. This function will first go through the angles, and select the one being altered, from the first to the last angle. With each

angle, the top three distances are found and the corresponding angle in which this occurred is stored into three groups. Afterwards, our second method from the last section will begin processing these angles (see Method 2 of Experimentation). These processes will keep iterating until the margin of error is less than 1. The following code (and helper function) demonstrates this process.



```
def Iterative_Optimal(input_array, processed_layers, steps):
    layer = len(processed_layers)
    average_values = Get_AVG_Groups(input_array)
    print("Layer:", layer)
    print('Average Values:', average_values)
    group1, group2, group3 = return_top_3_max(average_values)
    group1[0] = float(group1[0])
    group2[0] = float(group2[0])
    group3[0] = float(group3[0])
    print("Groups:", group1, group2, group3)
    finished = False
    while not finished:
        increment = float(abs(group1[0] - group2[0]))
        while not (group2[0] < group1[0] < group3[0] or
                   group2[0] > group1[0] > group3[0] or
                   min(group1[0], group2[0], group3[0]) - 0 < 1):
            increment /= 2
            if group1[0] < group2[0] < group3[0]:
                print("Finding Data to Left")
                new_group = Get_AVG_Groups(Get_Data(float(group1[0]) - increment, steps, processed_layers))[0]
                group1, group2, group3 = return_top_3_max([group1, group2, new_group])
                print("Groups:", group1, group2, group3)
            if group1[0] > group2[0] > group3[0]:
                print("Finding Data to Right")
                new_group = Get_AVG_Groups(Get_Data(float(group1[0]) + increment, steps, processed_layers))[0]
                group1, group2, group3 = return_top_3_max([group1, group2, new_group])
                print("Groups:", group1, group2, group3)

        if increment < 1:
            return group1, group2, group3

        print("Averaging Top Two")
        new_data = Get_Data([float(group1[0]) + float(group2[0])] / 2, steps, processed_layers)
        new_group = Get_AVG_Groups(new_data)[0]
        increment = abs(float(group1[0]) - (float(group1[0]) + float(group2[0])) / 2)
        group1, group2, group3 = return_top_3_max([group1, group2, new_group])
        if group3 == new_group:
            while group3 == new_group:
                steps += 1
                print("Groups:", group1, group2, group3)
                print(f"Increasing Increment... Increment: {steps}")
                print("Group 1")
                group1 = Get_AVG_Groups(Get_Data(float(group1[0]), steps, processed_layers))[0]
                print("Group 2")
                group2 = Get_AVG_Groups(Get_Data(float(group2[0]), steps, processed_layers))[0]
                print("New Group")
                new_data = Get_Data([float(group1[0]) + float(group2[0])] / 2, steps, processed_layers)
                new_group = Get_AVG_Groups(new_data)[0]
                increment = abs(float(group1[0]) - (float(group1[0]) + float(group2[0])) / 2)
                group1, group2, group3 = return_top_3_max([group1, group2, new_group])

        print("Groups:", group1, group2, group3)
        if increment < 1:
            return group1, group2, group3
```

```
def Get_Data(fixed_value, n, processed_layer):
    final_array = []
    for i in range(0, 90, int(90 / n)):
        print(f"Getting Data...{int(i * n / 90 + 1)}/{n}")
        for j in range(0, 90, int(90 / n)):
            temp = Group(processed_layer + [fixed_value, i + 90 / n, j + 90 / n] +
                         [0] * (5 - len(processed_layer))).Custom_Return_Value(n)
            final_array.append([temp[0][len(processed_layer):], temp[1]])
            with open("Calculated_Data.txt", "a") as f:
                f.write(str(temp))
                f.write("\n")
    return final_array
```

Figure 5.3. Screenshot of code 2

This process is iterated through all the angles, finding the approximate optimal solution (with a 1 degree margin of error). However, because of the nature of how the Group class is set up, the last four angles are not returned. Thus, a final function is used to iterate through all combinations for the last four angles with an increment of 1. All the angles are rounded and returned. This means that, overall, this approach will return an angle list with the first six angles as approximations of an optimal solution with a margin of error of 1, and the last four angles are optimal integer solutions. This finally produces the following angles list. [8, 16, 25, 34, 42, 52, 57, 65, 73, 81].

6. Conclusions And Future Works

The approach outlined in the paper demonstrated a more runtime-efficient algorithm that approximates an optimal solution with a margin of 1 degree for the first six angles [9]. This is an acceptable margin of error mainly due to two reasons. Firstly, people in reality are not able to have pinpoint accuracy in the direction they apply their force, meaning that this 1 degree of error will have little impact in a real life situation. Second, it is clearly shown in numerous figures in this paper that the overall shape of this scenario resembles a parabola, meaning that little margins of error towards the optimal solution will result in little to no change in the total distance travelled (change in vertical distance increases the
further away a point is horizontally from the vertex).

Although this situation assumes many conditions that are too good to be true for real life, the general result found in these projects add reason to many strategies in running. For example, Olympic runners start off their run with pedals that allow them to exert a smaller angle of force. Observers often see them straightening out within a few steps. Our results demonstrate that increasing angles of force does indeed allow runners to travel an optimal distance (although our program does not account for time efficiency).

Among many issues that the current algorithm has, most notable to me is the issue of time efficiency. This algorithm is not very time-efficient, and will not work nearly as well if the number of angles were increased. Thus, modifications and better approaches can be considered to significantly improve this aspect of our project. Additionally, adding conditions that are closer to

reality could also be beneficial for this program, as it can give us a strategy that we can use in real life.

REFERENCES

[1] Salo, Aki, and Ian Bezodis. "Athletics: Which starting style is faster in sprint running standing or crouch start?." Sports Biomechanics 3.1 (2004): 43-54.
[2] Graber, Kenneth C. The effect of starting body position on velocity in running. Diss. 2013.
[3] Frishberg, Barry A. "An analysis of overground and treadmill sprinting." Medicine and science in sports and exercise 15.6 (1983): 478-485.
[4] Parker, G. W. "Projectile motion with air resistance quadratic in the speed." American Journal of Physics 45.7 (1977): 606-610.
[5] Barrett, Paul, et al. "matplotlib--A Portable Python Plotting Package." Astronomical data analysis software and systems XIV. Vol. 347. 2005.
[6] Richmond, Bettina, and Tom Richmond. "How to recognize a parabola." The American Mathematical Monthly 116.10 (2009): 910-922.
[7] Zhang, Xinpeng. "Reversible data hiding with optimal value transfer." IEEE Transactions on Multimedia 15.2 (2012): 316-325.
[8] DeVORE, Ronald A., and Zheng Yan. "Error analysis for piecewise quadratic curve fitting algorithms." Computer aided geometric design 3.3 (1986): 205-215.
[9] Huang, Xiaoqiu, and Webb Miller. "A time-efficient, linear-space local similarity algorithm." Advances in applied mathematics 12.3 (1991): 337-357.
[10] Young, W. A. R. R. E. N., B. R. I. A. N. Mc Lean, and J. A. M. E. S. Ardagna. "Relationship between strength qualities and sprinting performance." Journal of sports medicine and physical fitness 35.1 (1995): 13-19.

MUSICAPP, A MUSIC SHEET TRANSCRIBING MOBLIE PLATFORM USING MACHINE LEARNING AND NATURE LANGUAGE PROCESSING

Daniel Wang[1] and Yu Sun[2]

[1]Troy High School, 2200 Dorothy Ln, Fullerton, CA 92831
[2]California State Polytechnic University, Pomona,
CA, 91768, Irvine, CA 92620

ABSTRACT

As technology advances, we have found more practical uses for it. This ranges from such things as cleaning the house using machines to serving restaurants with robots. Using technology, what if we can use machines to automatically write sheet music for us, transcribing it from audio [1]. This paper designs an application to do exactly that. We used Java to write a program and app that would be able to transcribe audio into sheet music and store it on an app. We applied our application to multiple cases and conducted a qualitative evaluation of the approach. The results show that it is possible with some fine tuning and may be usable in the near future.

KEYWORDS

Music, Sheet Music, Audio to Sheet Music, Flutter.

1. INTRODUCTION

The program in mind is being made into an app. We believe the usage of it will be able to transcribe audio into sheet music that will be able to be saved in an app [2]. I think it will be popular with a certain niche but will be helpful in that certain niche. The benefits of the app will be the ability to easily make audio into sheet music. This will possibly benefit the people that want to write music using a computer and allow them to easily edit the music with different recordings and cut time by needing to manually write the music. The only consequence that I foresee happening is the possible mistranslation of music that can cause the sheet music to mess up and transcribe the wrong notes. I think this topic is very important as music is a massive part of culture and ways to improve the accessibility of music is extremely important.

Some of the existing techniques and systems that have been proposed to that were the same as mine are the numerous programs and apps. They are about the same as mine, being able to upload music using a WAV or MP3 file [3]. They also have the feature to record music and have it uploaded to the site. However these programs are all tied to the internet and computer which might make it difficult to record. Their programs might lead to some issues that are caused by bad recording methods because you have to record it on the computer. This is inferior to the computer as you would need your instrument near your computer to record the music, while the app being on a phone will allow you to record it wherever you want.

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 87-95, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121709

My tool is an app that is able to record and transform WAV files into sheet music [4]. Our method was inspired by other tools but there are some other good features. First, our music is on the phone, allowing for easier recording and uploading. Secondly, our app allows you to record the music on the device, which allows for a quicker and easier way to get your sheet music as opposed to recording and uploading a WAV file every time. This provides an easier method than other methods because we allow our app on mobile devices and allow them to record them as well, allowing for easier recording.

We took 3 sample cases, Twinkle Little Star, Final Duet, and Electronic Music (Midi) and passed them through the app which converted them into sheet music and compared these sheet music with the actual sheet music from the songs to see how accurate our conversion was [5]. All three test cases are vastly different sounding. Twinkle Twinkle little star is a simple piano rift that everyone has heard. Final duet is a duet between a piano and a violin and electronic music uses digital synthesizers with a midi interface to create piano music.

The rest of the paper is organized as follows: Section 2 gives the details on the challenges that we met during the creation of the app and experiment; Section 3 focuses on the details of our solutions corresponding to the challenges that we mentioned in Section 2; Section 4 presents the relevant details about the experiment we did, following by presenting the related work in Section 5. Finally, Section 6 gives the conclusion remarks, as well as pointing out the future work of this project.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Converting Audio to a Midi File

The first challenge was turning audio into computer midi files. We need it to input it into our program so our program can actually read the audio data [6]. A midi file is a sheet music for the computer. A midi file pretty much holds musical information such as note numbers, pitch, duration, and other musical information for a piece. A midi file is organized by event messages that carry information for each note. Using a sequencer, we can view and edit the audio data as we would like. One way we can think of it is by using a computer to record a pitch and use a system to convert the pitch and test it to see if the computer recognizes it. There are points in the audio file where there are recorded key-frames that are called samples. These samples hold a pitch that has been recorded by our computer. Training a program to read in audio data becomes hard because in a single song, there could be a million samples.

2.2. Converting a Midi File into Sheet Music

While midi holds important note information pertaining to the particular piece you created, parsing through that data in attempts to make it into sheet music poses a problem. Firstly with the midi file format, it also comes with someone figuring out the notation, what is quarter note, ace note, half note. This happens because midi doesn't use the standardized format for time signatures. A time signature is the traditional way we signify how many beats are in each measure. A measure is a section of a song. It records it in a different way. It records audio duration in beats. The issue is that midi uses a different time measurement than normal. It uses ticks instead of beats and is recorded in delta time. It doesn't use a tick of normal time, it uses a tick of delta time, a measure of time from now to the last time a midi event happened. The main

challenge that comes with converting midi to sheet music is the conversion of time between midi standard time and regular time signature.

2.3. Connecting Our Powerful Musical Backend to a Comfortable Front End Available for Other Users

The third challenge is the creation of the app ui, we used flutter to create an android app using an android emulator [7]. Flutter is constantly used for creation of android and IOS app and doesn't supply much back end resources for audio or musical uses [8].We had to create the back end that wasn't linked with flutter and is on the google collab that is able to pull online python packages capable of overcoming the challenges needed [9]. Since flutter and google collab are not linked by any direct path or package, we had to create our own way to send audio information from the app into the converter and from the converter back into the app. One connection could be firebase but we needed another key to help post the image and communicate from flutter to the google collab server.

3. SOLUTION

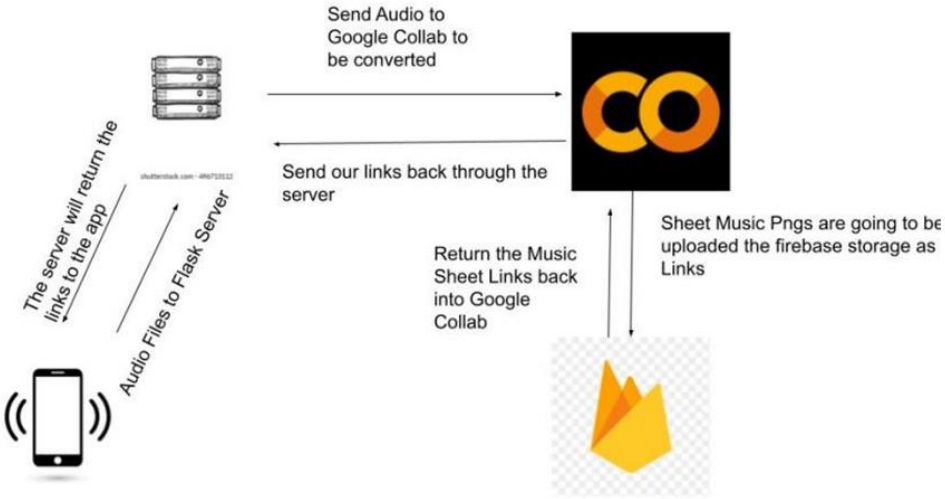
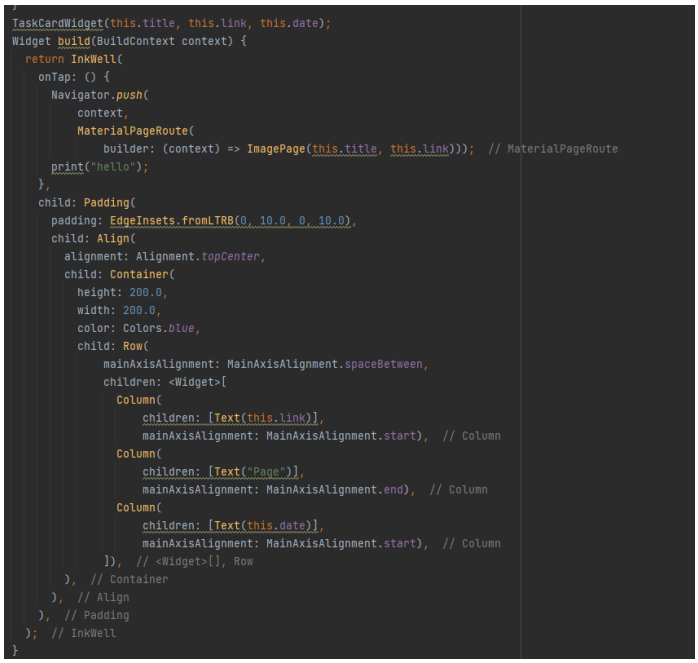


Figure 1. Overview of the solution

This is by running the audio file through the magenta package which uses Onsets and Frames to convert the audio to a midi file. Once we have a midi file our computer can start calculating the correct musical notation. For this we can use the music 21 library to convert the midi note number and tick format to western musical notation. Lastly, in the google collab, we use firebase service storage to upload our image to firebase [10]. We access the storage server service by using the firebase python package. In firebase, the image is hosted to the firebase server, meaning we have a web link to our image. That web link is sent back to the google collab which is then returned to the server and finally returned to our flutter app [11]. This image is displayed in the user on a new screen for our user using an image widget.

The first component we use to create our app is flutter. Flutter is a plug in for android studio code that easily allows people to make UIs for android and IOS applications. For this app, we made different screens, using dart files in flutter, each having its different use. We had a record page, a save notes page, and a sheet music viewing page. For our UI, we used the flutter built in widgets

and for app functionality such as connecting to the internet and sending files to the firebase. We used dart packages.



```
TaskCardWidget(this.title, this.link, this.date);
Widget build(BuildContext context) {
  return InkWell(
    onTap: () {
      Navigator.push(
          context,
          MaterialPageRoute(
              builder: (context) => ImagePage(this.title, this.link)));  // MaterialPageRoute
      print("hello");
    },
    child: Padding(
      padding: EdgeInsets.fromLTRB(0, 10.0, 0, 10.0),
      child: Align(
        alignment: Alignment.topCenter,
        child: Container(
          height: 200.0,
          width: 200.0,
          color: Colors.blue,
          child: Row(
              mainAxisAlignment: MainAxisAlignment.spaceBetween,
              children: <Widget>[
                Column(
                    children: [Text(this.link)],
                    mainAxisAlignment: MainAxisAlignment.start),  // Column
                Column(
                    children: [Text("Page")],
                    mainAxisAlignment: MainAxisAlignment.end),  // Column
                Column(
                    children: [Text(this.date)],
                    mainAxisAlignment: MainAxisAlignment.start),  // Column
              ]),  // <Widget>[], Row
        ),  // Container
      ),  // Align
    ),  // Padding
  );  // InkWell
}
```

Figure 2. Screenshot of code 1

(THIS IS UI CODE) This is the code to create a new custom task card widget. It stores all the info for our converted sheet music. We can see a couple of UI flutter mechanics in here such as container, row, and text.

```
void pickFiles() async {

  _resetState();
  try {
    _paths = (await FilePicker.platform.pickFiles(
      type: _pickingType,
      onFileLoading: (FilePickerStatus status) => print(status),
      allowedExtensions: ['mp3', 'wav'],
    ))
        ?.files;
    _uploadFiletoServer();
  } catch (e) {
    _logException(e.toString());
  }
}
```

Figure 3. Screenshot of code 2

This is code that examples dart packages and in this function, we use file picker package to open the native file picker on either android or IOS devices. We can see that we only allow mp3 and wav files into our program. After we pick a file, our flutter code calls a very important function called uploadFiletoServer, this allows us to send our files to the ngrok server. The ngrok server is the second major component in our app, it provides a tunnel to our web-page that holds our music files in its directories. We use the flask python package to capture the music files from our server into our google collab with python code.

Our third major collaboration is the ability to turn audio files into sheet music. This is by calling the function, generate sheet music. This takes in a file path that comes from flask server and first what generate sheet music does The first thing that generate sheet music does is convert our audio file to midi. We do this by using the function generatemidi, which is going to use the magenta package to convert a solo piano recording into a midi sequence. The way magenta processes this is by using a new model they created uses onset and frames. This model tracks notes across two stacks of programs. One stack is detecting when a note begins and if a note is currently active. Using this output, we can restrict some notes that are picked up by the detector that might not be there.

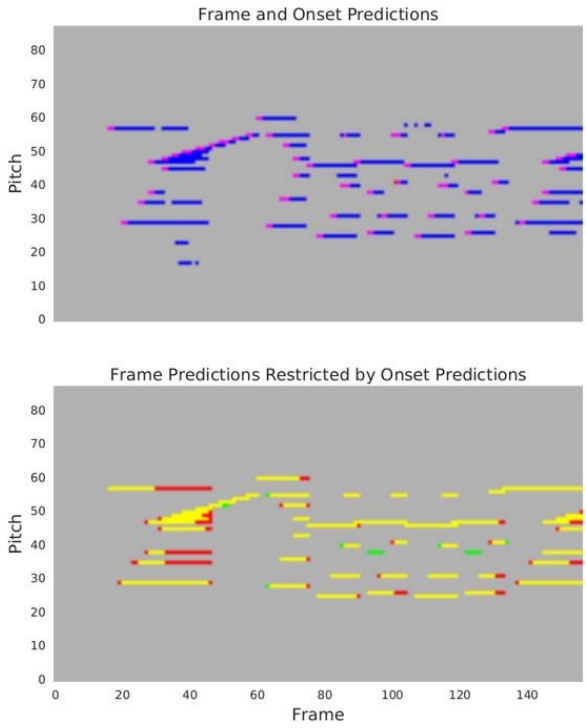


Figure 4. Graph of predictions

The first image in blue and magenta is the prediction of notes without restricting it by onset predictions every time a note is not actually playing, the magenta block is not at the front of the note. The image below shows the results of restricting our prediction by our onset prediction which in turns yields more accurate midi data. After we return the midi data to generate sheet music, we then pass it to the function midi2sheet_music This is going to take in a midi path and use the music 21 package to convert the midi information into a png image of the sheet music. The music 21 library was created in the university of MIT and is used as a tool kit for computer aided musicology. We can use the python library to make sheet music directly from python using the note class but for us we wanted to convert a whole midi sequence so we have to use the converter class which is a part of the music 21 library. This class is a midi converted into a music 21 object. Once it's in the music 21 library, we can write it to a png file which is returned to generate sheet music. Then generate sheet music and upload that png to firebase storage. Then it will take the image links and turn it back into ngrok server, then using the flutter which will be displayed using the image screen.

4. EXPERIMENT

4.1. Experiment 1

The first experiment that I had in mind was that we were going to take an audio file and convert it into sheet music using our app, then compare it against the written sheet music piece. To analyze this, we will be listening to both of the pieces played to see if there are any audio differences or if there is any notation difference. The first piece we are going to examine is a simple melody on a solo piano. This piece is going to be a twinkle twinkle little star. This is a base case because twinkle twinkle little star is a simple song that the AI should be able to recognize.

This is the result of the generated sheet music using our app, we can tell that this has visible different notation than our original notation. Although listening to both of the pieces side by side, we can see.

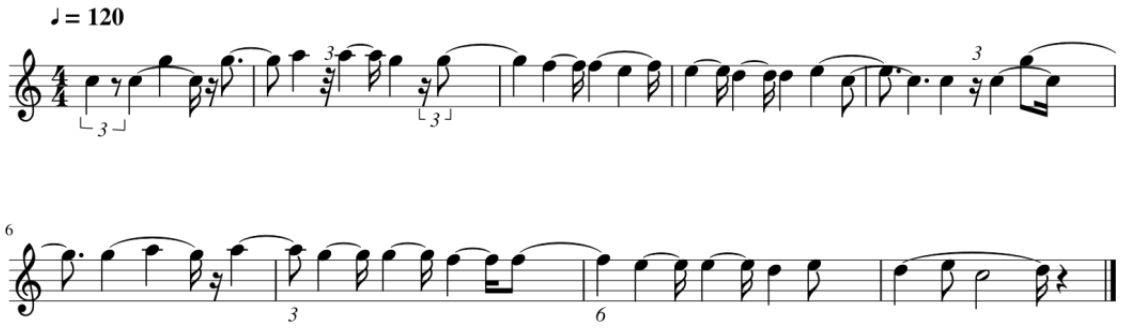


Figure 5. Music pieces 1

4.2. Experiment 2

This experiment is going to use our app to generate sheet music of the same twinkle twinkle little star melody. This melody is written in 4/4 time meter and one of the easiest melodies to play on piano. The variable that we are changing is going to piano timbre. Timbre is an umbrella term that usually stands for the sound of the instrument. For instance, a grand piano has a distinct sound, so does an electric guitar. We can use the computer to create synthetic timbres and use a simple midi sequencer, so we created an audio file that uses the same melody but instead of a grand piano playing, its a bright synth lead.

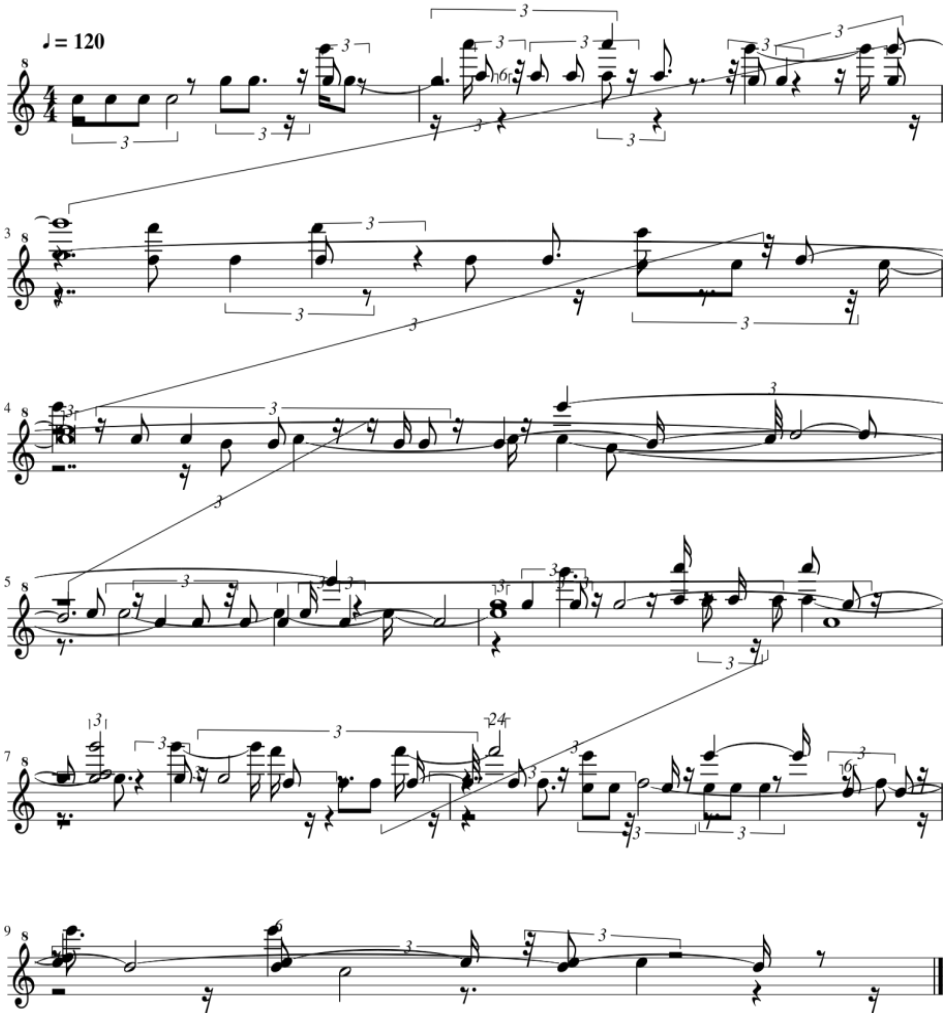


Figure 6. Music pieces 2

This is the generated sheet music from our app from the timbre shifted audio file. We can already see a similarity from both our results that seems to overcomplicate the notation of the sheet music. We can also see there are more lines there as more measures have more notes in it than the last result. Playing this sheet music alongside the bright synth lead.

Analyzing the results we saw from our experiments, we can determine that challenge one was successfully met. We were able to convert audio into a midi file. We were able to see this by checking the output and comparing the midi file to the original midi file of twinkle twinkle little stars. In both cases, timbre is shifted, the audio was actually able to recognize the midi information. Converting that midi to sheet music also worked but didn't meet expectations [15]. There are a couple issues converting a full midi file into music 21 library. One is that it overcomplicates the notation and also that there is no bass clef included. Although for simple piano tracks such as twinkle twinkle little star this played in solo piano. We were able to receive sheet music that actually sounds like the original sheet music so overall, this app could be used as a sort of practice for beginners that are unaware how to notate their piano melodies.

5. RELATED WORK

In this paper, the four scientists Fremery, Clauson, Ewert, and Muller researched two different approaches to automate a sheet music to audio identification [12]. Their pieces used various instruments with different complexities. The two approaches they used are first in OMR software, which is an optical music software that extracts information from the musical symbols as the program scans the music sheet. The second approach is a music to audio matching search, which compares musical content to a sequence of bars that we're currently looking at. In the end, the scientists proved that both approaches seem to be more useful in certain situations than others. This research paper took a much larger experiment pool of various tracts, pages of music, with varying different complexities. Our work focused more on mainly simple melodies. In our work, the strength is that we actually showed the results of our generated sheet music, but since they're experimental pool was so large, they were unable to show the sheet music in the research paper.

6. CONCLUSIONS

In this paper, we discover more automated use of machine learning to generate the music sheet base [13]. And we used advanced dev technology in dart based mobile dev framework names flutter. The way that flutter is connected to the server to send audio files is to use the HTTP package. We use the HTTP package for the flutter app to upload our audio file into the flask server. We use ngrok to host our server which is currently running on our google collab. After we upload our audio file from our app to our flask server, our google collab takes the audio files from the ngrok server. Then we use python code to take this audio file and convert it into an image of sheet music. The application is designed to perform all the functions mentioned before. We made this music app and published it on Google Play and did a lot of analysis with the app. Next step we are ready to publish the App to the Ios platform [14].

REFERENCES

[1] Saling, Lauren L., and James G. Phillips. "Automatic behaviour: efficient not mindless." Brain research bulletin 73.1-3 (2007): 1-20.
[2] Dorfer, Matthias, et al. "Learning audio–sheet music correspondences for cross-modal retrieval and piece identification." Transactions of the International Society for Music Information Retrieval 1.1 (2018).
[3] Liu, Qingzhong, Andrew H. Sung, and Mengyu Qiao. "Detecting information-hiding in WAV audios." 2008 19th International Conference on Pattern Recognition. IEEE, 2008.
[4] Islam, Rashedul, Rofiqul Islam, and Tohidul Mazumder. "Mobile application and its global impact." International Journal of Engineering & Technology 10.6 (2010): 72-78.
[5] Loy, Gareth. "Musicians make a standard: the MIDI phenomenon." Computer Music Journal 9.4 (1985): 8-26.
[6] Li, Dongge, et al. "Classification of general audio data for content-based retrieval." Pattern recognition letters 22.5 (2001): 533-544.
[7] Chen, Qiuyuan, et al. "How should I improve the UI of my app? a study of user reviews of popular apps in the Google Play." ACM Transactions on Software Engineering and Methodology (TOSEM) 30.3 (2021): 1-38.
[8] Garg, Shivi, and Niyati Baliyan. "Comparative analysis of Android and iOS from security viewpoint." Computer Science Review 40 (2021): 100372.
[9] Gunawan, Teddy Surya, et al. "Development of video-based emotion recognition using deep learning with Google Colab." TELKOMNIKA (Telecommunication Computing Electronics and Control) 18.5 (2020): 2463-2471.
[10] Khawas, Chunnu, and Pritam Shah. "Application of firebase in android app development-a study." International Journal of Computer Applications 179.46 (2018): 49-53.

[11] Payne, Rap. "Developing in Flutter." Beginning App Development with Flutter. Apress, Berkeley, CA, 2019. 9-27.
[12] Ewert, Sebastian. Signal Processing Methods for Music Synchronization, Audio Matching, and Source Separation. Diss. Universitäts-und Landesbibliothek Bonn, 2012.
[13] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.
[14] Wetchakorn, Thara, and Nakornthip Prompoon. "Method for mobile user interface design patterns creation for iOS platform." 2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE, 2015.
[15] Tanprasert, Thitaree, et al. "Midi-sheet music alignment using bootleg score synthesis." arXiv preprint arXiv:2004.10345 (2020).

PREDICTION AND KEY CHARACTERISTICS OF ALL-CAUSE MORTALITY IN MAINTENANCE HEMODIALYSIS PATIENTS

Mu Xiangwei[1], Zhu Mingjie[2], Liu Shuxin[2], Li Kequan[1],
You Lianlian[2] and Che Shuang[2]

[1]School of Maritime Economics and Management,
Dalian Maritime University, Dalian 116026, Liaoning, China
[2]Dalian Key Laboratory of Intelligent Blood Purification, Dalian Municipal
Central Hospital affiliated with Dalian Medical University,
Dalian 116033, Liaoning, China

ABSTRACT

Predict and analyze key features of all-cause death in maintenance hemodialysis patients to provide guidance for later diagnosis and treatment. Four machine learning methods were used to establish an all-cause death prediction model for maintenance hemodialysis patients and compare their performance. Analyze the key characteristics that have an important impact on all-cause death, and conduct user portraits for patients of different ages and genders. After comparison, the random forest algorithm works best, and an important factor affecting the all-cause death of patients is obtained. Among them, the all-cause death of all patients is related to factors such as albumin, blood potassium, blood magnesium, and urea; With age, the importance of factors such as blood sodium and phosphorus increases, and the importance of factors such as cardiac ultrasound ejection fraction decreases. Finally, there were also differences in the importance of analyzing patients of different ages and different sexes affecting their all-cause death. It is useful for residents to adjust their dialysis index timely.

KEYWORDS

Maintenance Hemodialysis, All-cause Mortality, Randomized Forest, Feature Importance, Prognosis.

1. INTRODUCTION

The growing number of patients with end-stage renal disease plagues global public health, and hemodialysis is one of the more important treatments, the rationalization of which contributes to the prognosis of patients. According to statistics, the number of patients with end-stage renal disease requiring hemodialysis treatment is estimated to be between 4.9 million and 7.08 million worldwide [1].

Although China has invested heavily in the treatment of end-stage renal disease and the reimbursement rate for dialysis treatment has received a large increase from 50% to 80%, the mortality rate of maintenance dialysis patients is still at a high level and their all-cause mortality rate is much higher than that of the normal population. Studies in Europe and the United States and other countries have shown that the mortality rate within 1 year of dialysis is about 20% [2]. Meanwhile, the mortality rate of maintenance hemodialysis patients in the United States has risen

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 97-107, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121710

since the outbreak of the new crown epidemic in 2020, with a mortality rate of 37% for patients receiving dialysis treatment in 2020 [3]. There are many factors influencing the mortality rate of dialysis patients, and dialysis equipment cannot analyze the past of dialysis patients to adjust the dialysis dose according to past data, and machine learning can precisely address this deficiency. The aim of this study was to investigate all-cause mortality in maintenance hemodialysis patients and its associated influencing factors. The analysis yielded important influencing factors that could provide guidance for patient prognosis.

2. MATERIALS AND METHODS

2.1. Data Source and Processing

2.1.1. Data Source

Data for this study were obtained from a large hemodialysis center in northeastern China, and data were recorded for 5 years for 1052 patients, including basic patient information such as gender, age, primary disease, and age on dialysis, baseline data such as cardiac ultrasound ejection fraction, reduced diastolic function, and heart failure, and blood chemistries including creatinine, albumin, potassium, sodium, and calcium, which were recorded quarterly. Due to the presence of survival, death, and departure data in the dataset, where the departure data are the result of patients being transferred to a hospital or undergoing renal transplantation, etc., it is not instructive for studying all-cause mortality. Therefore, this part of the data was chosen to be excluded from this paper. The study was approved by the hospital ethics committee, and all patients gave informed consent and signed an extensive informed consent form.

2.1.2. Data Processing

In this paper, we introduce a padding method using missing values of adjacent data points - K-Nearest Neighbor algorithm [4-6], which identifies adjacent points by measuring the distance between points, estimates and fills missing values, and is robust to noise.

The complete samples in dataset D are denoted as D_i and the samples containing missing values are denoted as D_a. The complete samples are clustered and the Euclidean distance is used to calculate the clustering similarity between dataset D_i and D_a. The distance between samples $X_i(X_{i1}, X_{i2}, \cdots, X_{in})$ and $X_j(X_{j1}, X_{j2}, \cdots, X_{jn})$ is denoted as:

$$dist = \sqrt{\sum_{k=1}^{n} (X_{ik} - X_{jK})^2} \qquad (1)$$

In the end, there were 9,992 processed data. Forty-six features were incorporated, as shown in Figure 1.

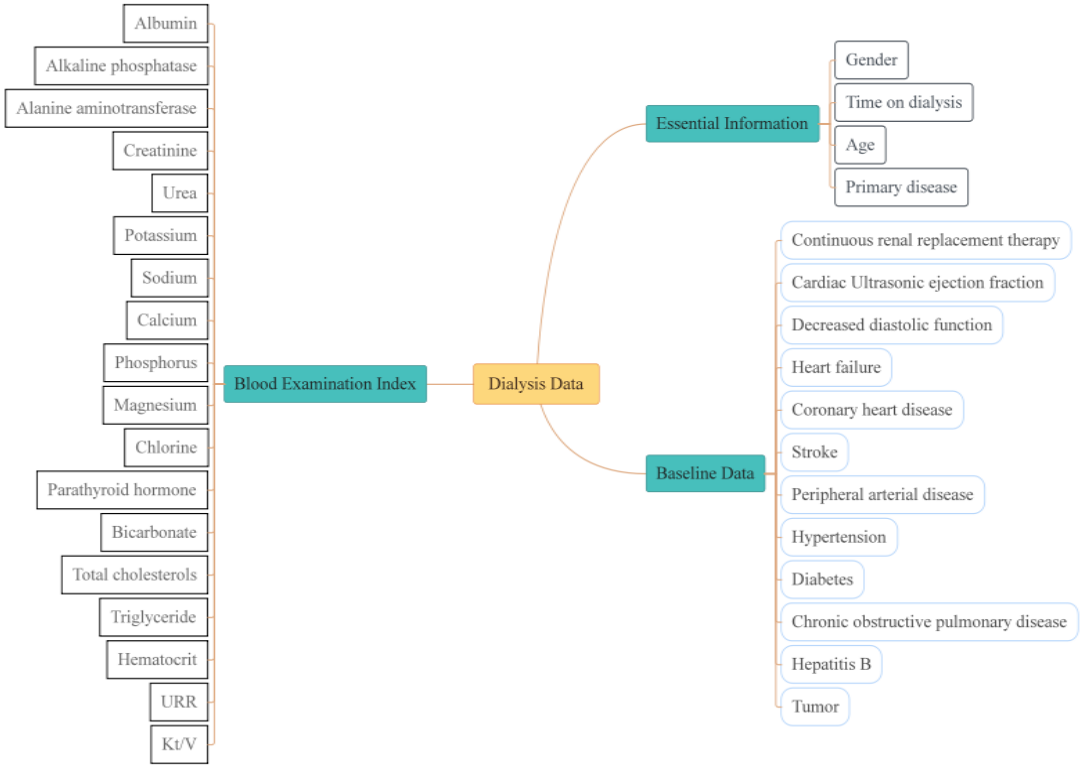


Figure 1. Dialysis data characteristics

Among them, the primary diseases include hypertensive nephropathy, diabetic nephropathy, chronic interstitial nephritis, hypertensive benign renal arteriosclerosis and polycystic kidney. Sodium, magnesium, and potassium plasma were included in the pre-dialysis and post-dialysis data.

2.2. Methods

In this paper, four machine learning algorithms-random forest algorithm, support vector machine algorithm, Adaboost algorithm, and logistic regression algorithm are used to construct a prediction model for all-cause mortality in maintenance hemodialysis patients. Due to the complexity of dialysis treatment, the algorithms with specific formulas such as K-NN and naïve bayes cannot describe them in detail, so the above algorithm was selected to establish an all-cause death model for maintenance hemodialysis patients, At the same time, the effect of different age groups and patient gender on mortality, and the user portrait.

2.2.1. Methods

(1) Random forest algorithm

Random forest is a supervised learning algorithm that consists of a series of independent decision trees trained by multiple Bagging integrated learning techniques for regression and classification. The variability between models is increased by constructing different training sets; the implementation of this algorithm obtains a sequence of classification models by training k rounds of initial data and uses this sequence of classification models to form a multi-classification system,

and finally, the final classification results of the system use simple majority voting method with the following classification decisions [7]:

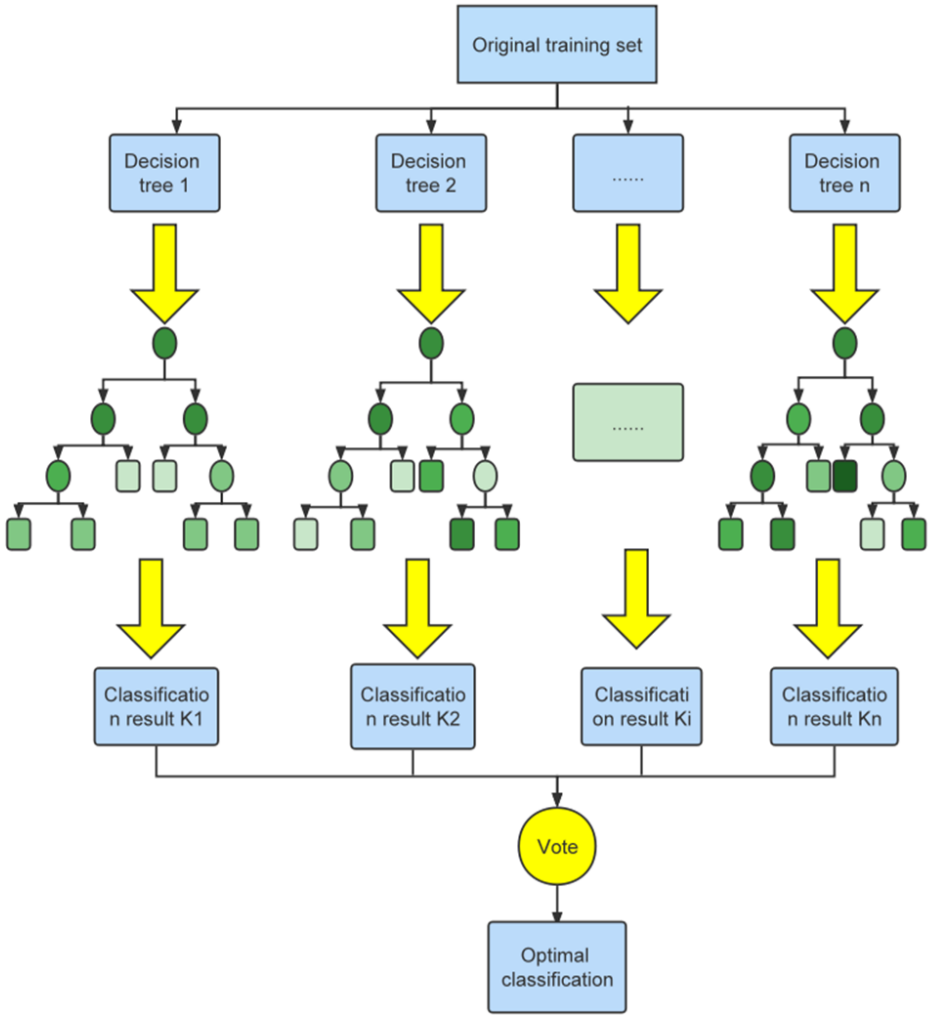


Figure 2. Schematic diagram of random forest

(2) Support vector machine algorithm

Support vector machine [8] is a classification model and the algorithm solves the binary classification problem by finding the optimal hyperplane. The optimal hyperplane is a multidimensional plane which finds the maximum interval of the binary classification problem and the solution of the algorithm can be expressed as the following problem:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \tag{2}$$

$$y_i[(w \cdot x_i) + b] - 1 \geq 0, i = 1, \cdots, N \tag{3}$$

(3) Adaboost algorithm

Similar to the random forest algorithm, the principle of the Adaboost algorithm [9] is to combine multiple if classifiers by certain methods to form a strong classifier. The algorithm is implemented by an iterative method. In the implementation of the algorithm, one weak classifier is trained at a time, and in the next training, the last trained weak classifier is used for iteration, and the implementation process of the algorithm is shown in Figure 3.

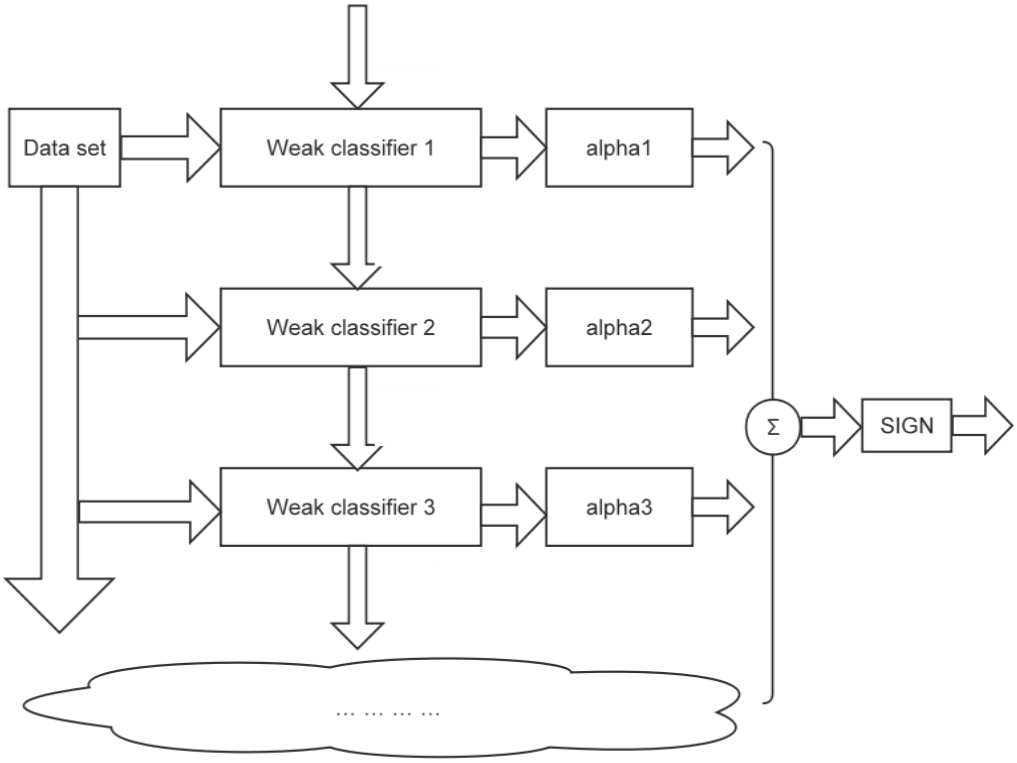


Figure 3. Schematic diagram of Adaboost algorithm

(4) Logistic regression algorithm

Similar to linear regression, the logistic regression algorithm [10] is a classical binary classification algorithm, and the algorithm is implemented by adding a Sigmoid function to the results of linear regression, transforming the linear regression results into results between 0 and 1.

2.2.2. Construction of Machine Learning Models

The 9,992 cases of data are divided into training set and test set according to the ratio of 7:3, and 70% of the data are used for training to build the best model, and the remaining 30% of the data are used as test set for testing for model evaluation and result output.
For model building, this paper introduces the 10-fold cross-validation method. The 10-fold cross-validation method is to divide the data set into ten parts, and select nine data in turn for training and one data for testing, and finally take the mean value of 10 test results to evaluate the performance of each model, and the specific results are shown in Table 2.

2.2.3. Evaluation of the Model

In this paper, the confusion matrix is used to evaluate the trained models, as shown in Table 1.

Table 1. Confusion matrix.

Confusion matrix		**True Value**	
		Positive	Negative
Predicted value	Positive	TP	FP
	Negative	FN	TN

Based on the confusion matrix, the model evaluation metrics: accuracy, precision, recall, F1-score and ROC curve can be calculated.

Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

Recall:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

F1-score:

$$\text{F1} - \text{score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

3. RESULTS

3.1. Comparison of Model Performance

The four models were trained to obtain the confusion matrix. As shown in Table 2, the random forest algorithm is greater than the remaining three algorithms in terms of accuracy, precision, recall, F1-score and AUC score, and all four metrics are above 90%. Collectively, the random forest algorithm has the best performance among the four algorithms.

Table 2. Results of model predictions.

model	**accuracy**	**precision**	**recall**	**F1-score**	**AUC**
Random forest algorithm	90%	90%	99%	93.50%	99%
Adaboost algorithm	84%	83%	84%	83%	87.9%
Support vector machine algorithm	79%	83%	79%	70%	75.8%
Logistic regression algorithm	79%	75%	79%	74%	77.3%

3.2. Feature Importance of Random Forest and Adaboost Models

In the training of Random Forest with Adaboost model, the importance of each influencing factor (feature) was obtained, and it was ranked to extract the features that have a greater impact on all-cause mortality, and the results are shown in Figure 4.

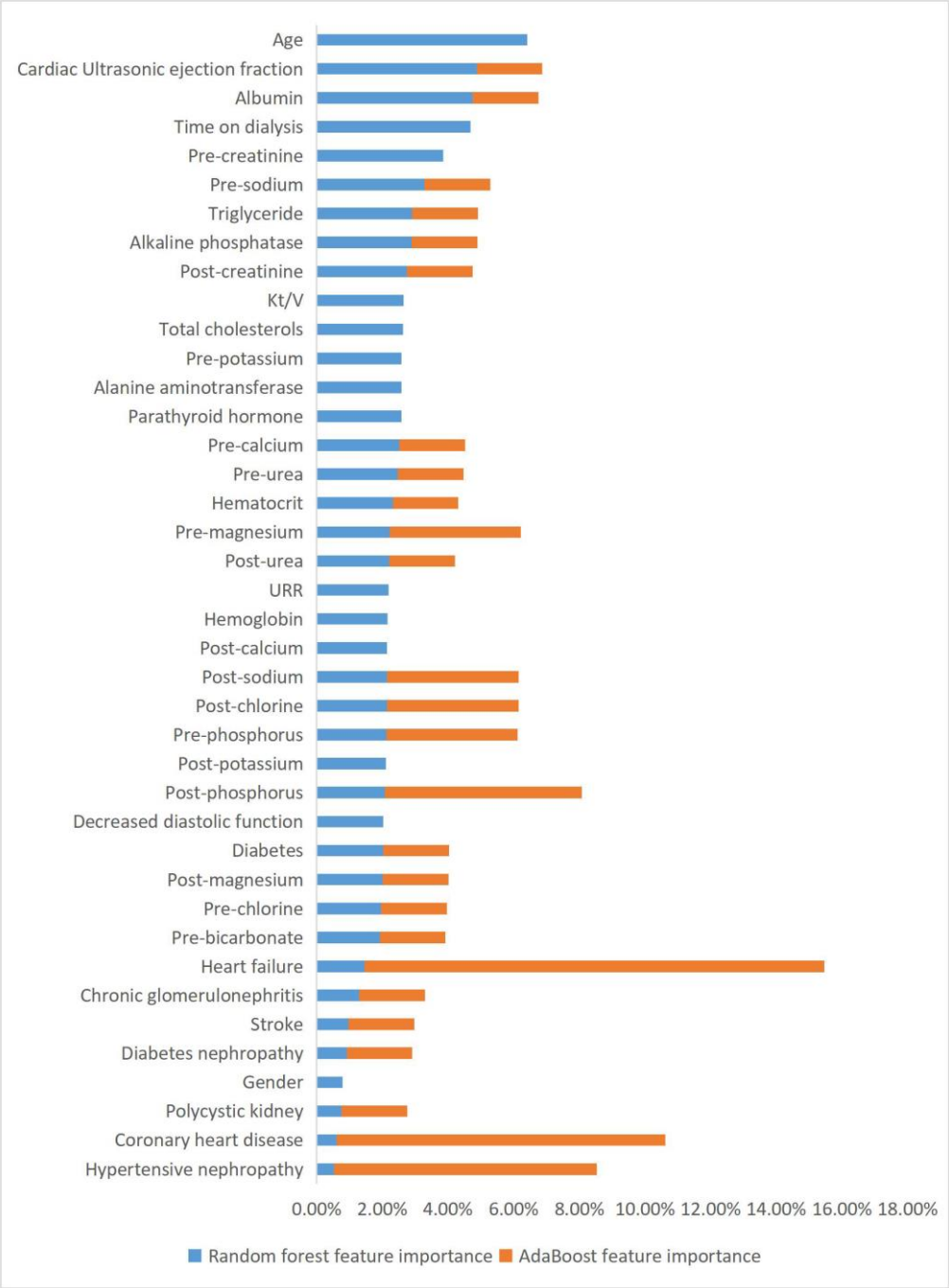


Figure 4. Ranking of feature importance

3.3. Further Analysis of the Random Forest Algorithm

The patient groups were divided according to age groups [11] into group A (<40), group B (40-59), group C (60-79), and group D (>=80). A random forest model was developed for the data of each group of patients separately, and a large difference in the importance of 16 characteristics, including dialysis age, cardiac ultrasound ejection fraction, and sodium, was found in the different age groups of patients, as shown in Figure 5.

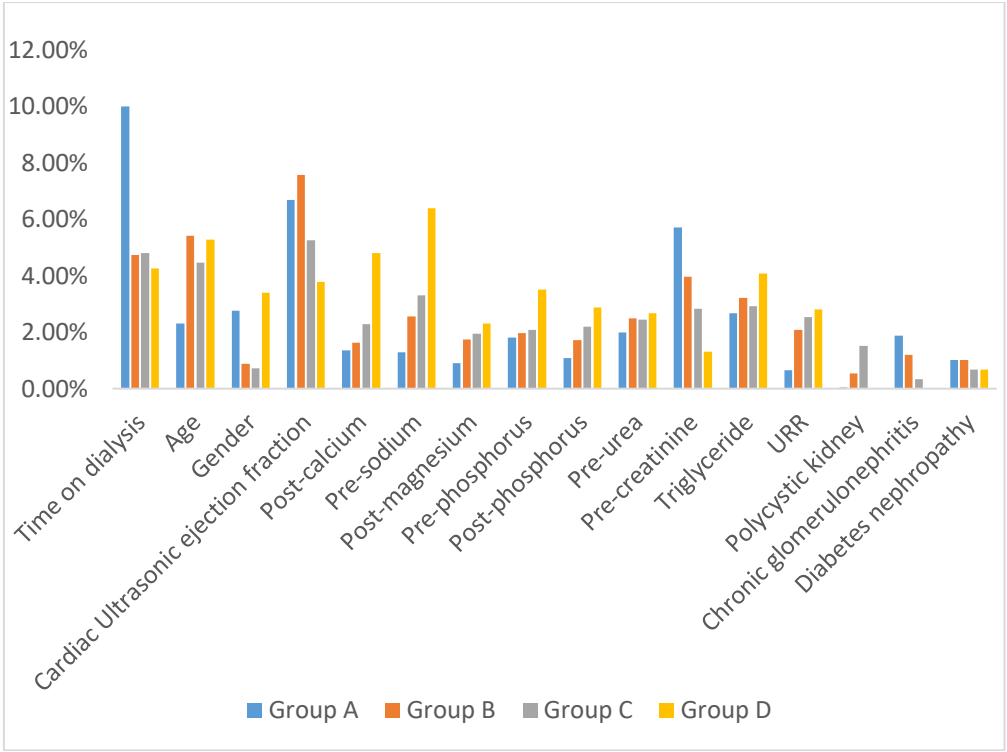


Figure 5. Fluctuations in the importance of age-segmented characteristics

For all patients, the weight of the effect of gender on all-cause mortality was relatively small, but after grouping by age group, the effect of gender characteristics on all-cause mortality was found to be higher in groups A and D. Therefore, the gender of patients in groups A and D was divided to explore the differences between factors influencing all-cause mortality in patients of different genders under this age group, and certain characteristics with greater variability were derived, as shown in Table 3.

Table 3. Gender classification assessment weights/% for groups A and D.

Features	**Group A**		**Group D**	
	Male	Female	Male	Female
Time on dialysis	16.61%	5.82%	17.00%	5.41%
Age	3.83%	2.76%	7.06%	4.74%
Cardiac Ultrasonic ejection fraction	2.31%	5.89%	0.64%	3.22%
Pre-calcium	1.23%	2.67%	1.69%	1.83%
Post-calcium	1.33%	1.58%	2.33%	2.48%
Pre-sodium	1.68%	1.90%	2.28%	3.52%
Post-sodium	1.31%	0.49%	1.33%	1.24%
Pre-magnesium	1.56%	0.51%	5.70%	1.03%
Total cholesterols	3.50%	1.46%	2.97%	2.19%
Parathyroid hormone	4.04%	0.94%	2.31%	1.31%
Alanine aminotransferase	1.95%	0.85%	5.19%	2.20%
Kt/v	1.49%	0.95%	1.25%	0.79%

4. CONCLUSIONS

Using four algorithms, all-cause mortality in maintenance hemodialysis patients was focused on, and the Random Forest algorithm was found to have the best performance among the four algorithms, while for all patients, the features with higher impact on all-cause mortality were derived using the Random Forest algorithm and Adaboost algorithm, as shown in Figure 5; finally, a gender- and age-specific user portrait of patients was conducted using the Random Forest algorithm for further analysis, with the following results:

Findings from a study of all patients: Random forest and Adaboost algorithms were used for the analysis of the importance of characteristics of all patients, and both reached consistent conclusions in the analysis of the importance of most characteristics, and the experimental results showed that in the patient population, the characteristics that had a greater impact on all-cause mortality were age on dialysis, age, albumin, potassium, cardiac ultrasound ejection fraction, calcium, sodium, potassium, urea, and alkaline phosphatase (alp). For the above findings, it has been found that age, diabetes, calcium, albumin, sodium, potassium, and alkaline phosphatase are associated with all-cause mortality in patients. According to a related study, malnutrition is one of the complications of maintenance hemodialysis patients and affects their prognosis, and serum albumin level is an important factor in determining the nutritional status of patients. Therefore, the study of serum albumin levels has significance for the prognosis of patients. It has been found,that alkaline phosphatase has a correlation with cardiac lesions and is its independent risk factor, which affects the survival rate of patients. The important features derived from the training can be focused on for further studies.

Conclusions from patients by age group: When the patient population was divided into four groups according to age, the importance of certain features fluctuated widely across age groups. Features that showed an overall increasing trend in importance with increasing age were: post-calcium, pre-sodium, post-magnesium, pre-phosphorus, post-phosphorus, and URR. The features that showed an overall decreasing trend in importance with decreasing age were: cardiac ultrasound ejection fraction, pre-creatinine, and chronic glomerulonephritis in primary disease.

Conclusions from grouping patients in groups A and D according to gender: Certain characteristics were analyzed in the two groups of patients with a large difference in their importance. Among them, in groups A and D, the importance of characteristics such as cardiac ultrasound ejection fraction, pre-calcium, pre-sodium, and so on were significantly higher in female patients than in male patients under this age group; while for characteristics such as dialysis age, age, pre-magnesium, total cholesterol, parathyroid hormone, alanine transferase, and Kt/v were higher in males than in females, as shown in Table 4. Therefore, in patients under this age group, dialysis should be performed with differentiation according to gender Treating, for different index values with different characteristics in different genders, dialysis should be reasonably arranged and timely interventions should be made to achieve the goal of reducing all-cause mortality in patients.

In this paper, the importance ranking and analysis of the characteristics related to all-cause mortality of all patients as well as maintenance hemodialysis patients of different ages and genders were performed, but the classification of patients was not detailed enough. Therefore, a more in-depth analysis will be conducted in the future to provide a more detailed user profile of patients, to identify patients accurately and efficiently, and to improve diagnostic efficiency.

ACKNOWLEDGEMENTS

This work was supported by Dalian Maritime University Undergraduate Education and Teaching Reform Research Project 2022-63, Humanities and Social Sciences Research Project of the Ministry of Education 18YJC630124, Liaoning Provincial Department of Education Science and Technology Research Project L2014203, Liaoning Provincial Social Science Planning Fund Project L14BGL012. We thank Dalian Municipal Central Hospital for the help with the clinical knowledge related to the extraction of our dataset and for providing clinical insights on the case study.

REFERENCES

[1] Thomas M. Cover, and Peter E. Hart."Nearest neighbor pattern classification.." IEEE Trans. Information Theory 13.1(1967).
[2] Tadashi Yamamoto, et al."Predialysis and Postdialysis pH and Bicarbonate and Risk of All-C ause and Cardiovascular Mortality in Long-term Hemodialysis Patients." American Journal of Kidney Diseases 66.3(2015).
[3] Lu Jiayue, et al."The relationship between survival rate and intradialytic blood pressure changes in maintenance hemodialysis patients.." Renal failure 39.1(2017).
[4] Viknesh Selvarajah, et al."Pre-dialysis systolic blood pressure-variability is independently associated with all-cause mortality in incident haemodialysis patients.." PLoS ONE 9.1(2017).
[5] Lv Ji-Cheng and Zhang Lu-Xia."Prevalence and Disease Burden of Chronic Kidney Disease.." Advances in experimental medicine and biology 1165.(2019).
[6] Mei Wang, et al."SKNN Algorithm for Filling Missing Oil Data Based on KNN." IOP Conference Series: Materials Science and Engineering 612.3(2019).
[7] Wenbo Wu, et al."A Cross-Sectional Machine Learning Approach for Hedge Fund Return Prediction and Selection." Management Science 67.7(2020). doi:10.1287/mnsc.2020.3696.
[8] "Cancer; Researchers from Department of Computer Sciences and Engineering Provide Details of New Studies and Findings in the Area of Cancer (Detecting biomarkers from microarray data using distributed correlation based gene selection)." Computers, Networks & Communications. (2020).
[9] Chen Y W, et al."[Early mortality and risk analysis in adult patients with maintenance hemodialysis].." Zhonghua nei ke za zhi 60.1(2021).
[10] Battisti Rodrigo, et al."Machine learning modeling and genetic algorithm-based optimization of a novel pilot-scale thermosyphon-assisted falling film distillation unit." Separation and Purification Technology 259.(2021). doi:10.1016/J.SEPPUR.2020.118122.
[11] Johansen Kirsten L., et al."US Renal Data System 2020 Annual Data Report: Epidemiology of Kidney Disease in the United States." American Journal of Kidney Diseases 77.4S1(2021).
[12] Theodorakopoulou Marieta, et al."SEX DIFFERENCES IN AMBULATORY BLOOD PRESSURE TRAJECTORIES AND BLOOD PRESSURE VARIABILITY IN HEMODIALYSIS PATIENTS." Journal of Hypertension 40.Suppl 1(2022).
[13] Bridle Tristen G, et al."Physiologically relevant hCys concentrations mobilize MeHg from rabbit serum albumin to form MeHg-hCys complexes.." Metallomics : integrated biometal science 14.3(2022).
[14] Patel Harsh, et al."Oropharyngeal cancer patient stratification using random forest based-learning over high-dimensional radiomic features.." Scientific reports 11.1(2021).
[15] Musbah Hmeda, Aly Hamed H.,and Little Timothy A.."Energy management of hybrid energy system sources based on machine learning classification algorithms." Electric Power Systems Research 199.(2021). doi:10.1016/J.EPSR.2021.107436.
[16] Ren Jiaolong, et al."Design optimization of cement grouting material based on adaptive boosting algorithm and simplicial homology global optimization." Journal of Building Engineering 49.(2022). [4]Wang Xingcheng, Meng Cheng,and Wang Yuanyu."Insight for the construction of R-T phase boundary in KNN piezoceramics from the view of energy band structure and electron density." Ceramics International 47.20(2021).
[17] Kathuria, Charu, Mehrotra, Deepti,and Misra, Navnit Kumar."A novel random forest approach to predict phase transition." International Journal of System Assurance Engineering and Management .prepublish(2021).

[18] Jiang Chengcheng, et al."Spatial modeling of gully head erosion on the Loess Plateau using a certainty factor and random forest model." Science of the Total Environment 783.(2021).

[19] Ali Ahmad Hassan, et al."A Model Incorporating Serum Alkaline Phosphatase for Prediction of Liver Fibrosis in Adults with Obesity and Nonalcoholic Fatty Liver Disease.." Journal of clinical medicine 10.15(2021).

AUTHORS

Mu Xiangwei, Associate Professor at Dalian Maritime University; His research focuses on machine learning and medical big data.

ARTIFICIAL INTELLIGENCE DESIGNED FOR ATTENDANCE

Jasmin Liao and Yu Sun

Trabuco Hills High School, 27501 Mission Viejo, CA
Cal Poly Pomona, 3801 W Temple Ave, Pomona, CA

ABSTRACT

Engaging online students is a challenge for many teachers. While I was a student, I saw teachers struggling to take attendance due to the number of students leaving their classes after attendance. Students would be held responsible for their work using facial recognition technology. To simplify the process of applying absences to students in each class, this paper proposes an application that would allow teachers to stay on top of their work. We applied our software to test "students" in the classroom and used various libraries/CSC styles to create a classroom that is easy for both the student and the teacher to read. Our designs are built upon OpenCV and PIL which are used as geometric classifiers to determine if the student is present. We tested several faces to see if the algorithm was suitable for the program. After conducting a qualitative evaluation of the approach, we've begun to implement registration, create new classrooms with different databases, and apply verification. With the addition of HTML code, we wereable to create a classroom that is safe, engaging, and easy to use.

KEYWORDS

Machine Learning, Artificial Intelligence, Python, JavaScript.

1. INTRODUCTION

As a student, I had a difficult time attending classes during quarantine. Ever since covid cases have risen, I realized the difficulty of managing the classrooms as I watched my teachers take ten minutes trying to take attendance. Covid precautions added more stress to teachers which severely affected the relationships between a student and their teachers by focusing on adjustment rather than focusing on the students themselves. It's not like traditional campus-based classes which require students to have a strong ability to create and stick to a schedule. Between class times, club meetings, homework, and jobs or work-study; online classes can be viewed as redundant and a hassle. Using our program would entice students to prepare for classes while also keeping them accountable for their schedules. This allows people and instructors to easily manage classrooms, meetings, and the number of people who attended the class. However, some consequence of this app is the flexibility. Ai- Attendance is fairly new and only has the basic foundations of the app which mainly concerns the website and its functionality. We are planning to develop feedback into the program so educators or mentors can have a mailbox dedicated to engagement in the classroom. It's a new look for educators and digital learners and it allows teachers to have more efficiency in classroom management.

a. What are existing related methods/tools on this topic?

To correlate the faces to the students in the database we use a library called OpenCV. OpenCV is a library used for image processing and plays a large role in programs with machine learning

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 109-113, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121711

algorithms. Companies like Google, Yahoo, and Microsoft use OpenCV to create amazing applications. However, performing face detections with Cascade Classifiers could not reliably detect humans in certain images. We've tested several different images to find that in some of them they would detect random spots in the corners with no faces at all. When tested with the video we found that the program was not compatible with moving pictures. We resolved to call a PIL module that stores pictures of individuals that would allow the computer to train and recognize the faces through a function called finding_closest_match. Here, teachers can use screenshots of the classroom to determine who is present in the classroom. This function uses a folder of people's pictures and then maps the image for similar geometric details from the database. In this paper, we follow the same line of research by Google Classroom. Our goal is to create a classroom that allows teachers to manage students entirely. Our method is inspired by Google Classroom's database; their learning platform is what inspires us to make an app that would allow teachers to navigate who is attending or leaving the classroom. There are some good features of our app. We have two roles within our app, the administrator and the student. The teacher position contains special admin roles and is able to create classrooms and send invitations to the students to join theirs. The process is similar to Google Classrooms except the teachers are able to create passwords for each classroom. While in the homepage, teachers have the ability to create new classrooms and add in new students. Students are able to access the classroom once given the passcode for the classroom. From then, teachers are able to see when the student enters the classroom. Our app also ensures that no stranger comes into the classroom with our verification app. To prevent the possibility of people trying to gain access to an online account that isn't theirs, we use two authentication factors to protect the user's credentials. Teachers and students will be able to log in by using a TOTP app that gives them the code.

Before creating Ai-Attendance, I took some courses in Java and HTML to understand the basics of web design. I've been inspired by the decorated websites and the ease of access that they contained. Searching through several different templates for HTML and CSS websites, we decided to settle on a template that was user-friendly. I tested several different students inside the classroom to see whether the program can recognize the students. Using a process-based evaluation, we checked to see whether each function in the website would work as intended and look aesthetically pleasing. We found a couple of kinks concerning the app route but were able to implement a solution each time. For the testing of the student's faces, we used a variety of faces to find the closest match in the classroom. Some libraries were unable to work as it was incompatible with some photos that used jpeg formats. Instead of using OpenCv, we translated from CV2 to PIL. From PIL we are able to use images from Google Drive to then match the individual's face to the ones in the drive. We also tested the different types of forms on the website by modifying the address bar. We tried to ensure that no one would be able to access the accounts through the website by changing the address or being able to inspect the password.

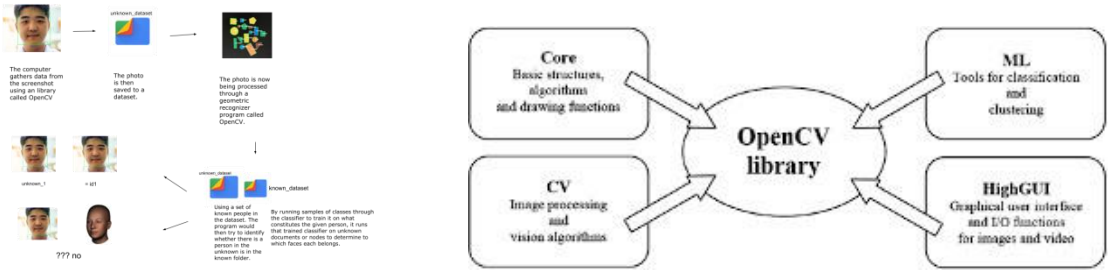
CHALLENGES

Actively reading the face from multiple different angles proved to be a challengebecause of the measurements in the facial recognition functions. Due to the inability ofthe program to recognize faces in live video, the program would often result in errors.

METHODOLOGY/SOLUTION

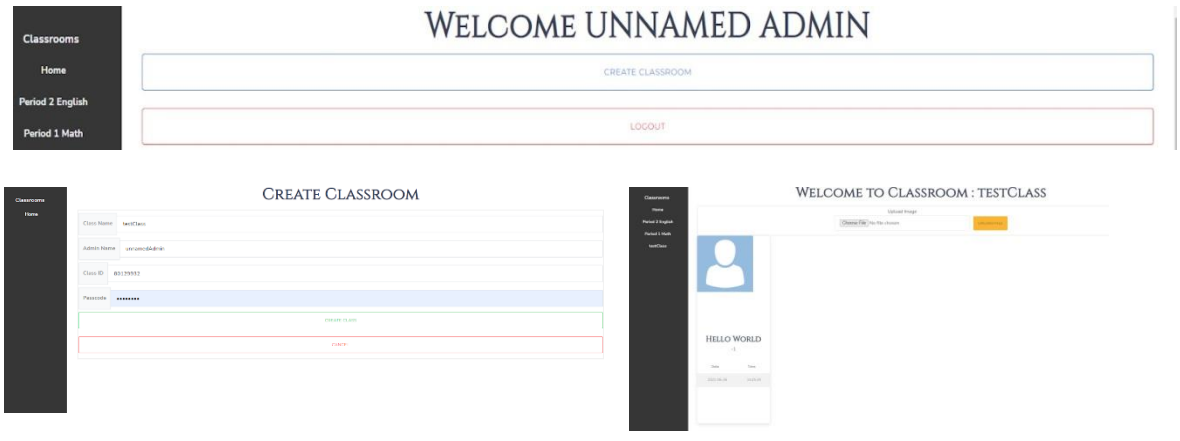
Ai-Attendance relies on facial recognition to identify students. First, we create two folders to help us distinguish the faces of the students. Our certificated classroom is created using a firebase folder linked to the firebase server. There would be a server that holds all of the student information, teacher information, class information, and second verification information. The student should be provided with a name, id number, picture, and a verification code. A fiducial

point recognizer program called OpenCVwould then be used to process the system. We imported cv2_imshow from OpenCV and imported ImageDraw for mapping out the face encodings since we were using Google's web IDE. By breaking down the image, they train the computer's deep neural network to better understand the dataset so it can recognize the student once they log on. Our program would run the function facial_recognition from the firebase console so it would only find the location of the known faces. It associates facial encoding with their identification by comparing the folders and known persons. The tolerance determines the distance between faces to consider the faces a match and returns a true or false if the tolerance is lower. In the event that they recognize the face as belonging to the student, the computer records its attendance. An overview of the process is shown in Figure 1.



DESIGN EXPERIMENT RESULTS

1. The reason we created Ai-Attendance was mainly for the convenience of the teachers during the lockdown and to expand my knowledge of machine learning and web development. I conducted research with my professor on how OpenCV works. Several tests were made on the project to ensure that facial recognition would work. Matching celebrities' faces with various photos were one of the experiments we practiced. We however were unable to create a live video as the database only took photos. My hope is that I will be able to further develop this project in the future. Our website tested the teacher's abilities as well as the students' perspectives.



How does your solution solve the problems?

My solution solves the problem of addressing classroom management by allowing teachers to upload a picture of the classroom to determine who is currently inside the classroom. We sampled different accounts to check whether the students have been recognized in the program and whether they have difficulty logging into a newclassroom. The results show that students were able to access and be added to theclassroom. Based on the feedback I received, the

classroom should have limits on the student's facial tests, which I agree with and hope to address later.

RELATEDWORK

1. Kaur, Paramjit, et's "Facial-recognition algorithms: A literature review." describe a method that discusses the broad range of methods used for face recognition and attempts to discuss their advantages and disadvantages. Initially, we present the basics of face-recognition technology, its standard workflow, background and problems, and the potential applications. Their job is very useful.
2. In Richards, Gregor, et al. "An analysis of the dynamic behavior of JavaScript programs." Proceedings of the 31st ACM SIGPLAN Conference on Programming Language Design and Implementation. They perform an empirical study of the dynamic behavior of a corpus of widely-used JavaScript programsand analyze how and why the dynamic features are used. We report on the degree ofdynamism that is exhibited by these JavaScript programs andcompares that with assumptions commonly made in the literature and accepted industry benchmark suites.
3. Hashemipour, Sadegh, and Maaruf Ali. "Amazon web services (AWS)–an overview of the on-demand cloud computing platform." They discuss how AWS contributes to a range of industries making them dependent on AWS's agility, scalability, and deliverability of services. Implementation using information available on AWS leading portals as well as the practicalexperiments involving the discussed features are given. As more and more industries become dependent on Big data and AI, one can state that AWS will play a key role in future developments across myriads of industries.

CONCLUSION

As a result of the Pandemic, I was inspired to develop an application assisting teachers with their attendance. It was through this project that I learned how to incorporate machine learning with web development. Platforms like Repl.it, Firebase, and Colab made it possible for me to create an environment for my application. I solved challenges such as implementing facial recognition with my new experience in machine learning using OpenCv, Tensor Flow, and GitHub. After finishing our software for facial recognition, we focused on safety and the ability to create new classrooms in the app.The TOTP generator and secrets allowed us to make a second authentication as well as reset passwords for users. The project turned out to be a success, however, we are still developing the program so we could expand on machine learning and the limits to classroom safety. Previous experiments from the algorithm proved to be effective as it identified the person from 10 tests with up to 15 people in a picture. The effectiveness of the project is mostly limited to the classroom page and the TOTP generator.Currently, we are at a disadvantage in terms of managing our classrooms. I want to be able to create an environment that allows teachers to easily add students. In spite of this, they are only able to add them one at a time, which makes it inconvenient for them. By making the buttons more visible and navigable with CSS and HTML, I will be able to improve accessibility through web routes and page functionality. In the future, I'm going to be improving upon the designs, data management, and sourcing of the project.

REFERENCES

https://codepen.io/pen/

https://www.w3schools.com/w3css/w3css_templates.asphttps://pug2html.com

https://pugjs.org/api/getting-started.htmlhttp://www.html2jade.org/

https://www.tutorialspoint.com/html/html_quick_guide.htm

Documentation: https://cirrus-ui.netlify.app/getting-started/developing

Github:https://github.com/Spiderpig86/Cirrus

CDN: https://unpkg.com/cirrus-uihttps://github.com/ageitgey/face_recognition

AUTHORS

Jasmin Grace Liao, I am currently a senior at Trabuco Hills High School! I spend most of my time involved with coding competitions such as Gamegalas and like to involve myself in coding projects in and outside of school!

Dr. Yu Sun Professor at Calpoly

A DATA-DRIVEN AND COLLABORATIVE MOBILE APPLICATION TO ASSIST SENSORS USING ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Aaron Tse[1] and Yu Sun[2]

[1]Arcadia High School, 180 Campus Dr, Arcadia, CA 91006
[2]California State Polytechnic University, Pomona,
CA, 91768, Irvine, CA 92620

ABSTRACT

There are numerous arguments as to what the best GPS software is [1]. However, there are no definitive answers as to which is the best. In this paper, we use multiple GPS applications tracking user locations to determine what GPS is best in terms of tracking users through a mobile app [3]. The app utilizes a GPS as well as a Google Firebase Realtime Database to manage, pinpoint, and track users' locations [2]. The application is explicitly applied to track locations of people that need care, such as the elderly. This will allow concerned caretakers to help keep track and take care of people in need.

KEYWORDS

Android, GPS, Flutter, Firebase, Google Maps.

1. INTRODUCTION

A GPS (Global Positioning System) has many uses in the modern age [4]. When it was started by the US Department of Defense in 1973, it was originally designed to assist soldiers and military vehicles in accurately determining their locations world-wide. In the modern day, however, it has taken on a whole new use. Whether it is helping people navigate to a new place or finding your missing phone, the GPS has become a central part of our lives by helping us get around. With 2.5 billion registered Android users and 1.8 billion Apple devices worldwide (all of them having GPS on them), it is imperative that this system is a well functioning part of every device. One of the many useful uses for this system is to locate people in need of help. This includes missing people, kids that aren't coming home on time, etc. One group that would greatly benefit from this app would be the elderly. With approximately 125,000 search and rescue missions where volunteer teams are deployed to find missing Alzheimer's patients every year, people with mental disabilities and their caretakers would greatly benefit from this app as well. Despite the potential dangers that this app poses, it is a necessary system to implement. With over 90,000 people missing in the USA at any given time on average, a GPS locator app will be imperative in aiding tracking and the rescue of people. Whether it be concerned parents or medical personnel utilizing the app, it will be used to save lives. There are many GPS apps on the market, with the two biggest ones being Google Maps and Apple Maps. According to BatuhanKilic and Faith Gulgen, Google Maps is accurate because the positional accuracy of the Google Map is mathematically more significant [5]. However, how does this mathematical accuracy of Google Maps measure up to that of Apple Maps?

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 115-124, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121712

Some of the existing technologies that have been proposed to enable or expedite location tracking, such as LifeAlert, have been marketed as being useful to people in need, especially seniors [14]. The primary function of LifeAlert is that it is a necklace with a button on it. Once the button is pressed, a user would be automatically connected with a 911 operator, to whom the person in need is able to speak to. Although the necklace itself doesn't have a GPS system, the 911 operator is able to determine the victim's location through tracing of the call. This slows down response times, taking away precious seconds of life. Although sounding promising, these implementations have largely faded into obscurity for multiple reasons. GPS tracking systems that aren't connected to mobile devices are plagued with issues. The first being battery life. Since they cannot be charged while still in use, this leaves people in need vulnerable during the charging duration. Secondly, most of these devices charge subscription fees in order to function to their full effect, such as LifeAlert with their subscription service price being up to $89.95 a month. Compounded by the fact that most services similar to services like LifeAlert bind people in multi-year contracts, these devices are nonoptimal. Another problem that is prevalent with medical alert devices is that they do not function in low-reception areas. This is not an issue for mobile devices, as they can utilize wifi as well to make calls. Although these medical alert devices have gained some traction, for the large part, they have not broken through the market. An ideal solution would be some device that would be near effortless to operate and not be forgotten when being brought out. Since people always have their mobile device on hand, it would be natural to have some location software on a mobile device [6].

In this paper, we aim to solve the aforementioned problems regarding the Google Maps GPS and how accurate it is [7]. My goal is to pinpoint the accuracy of the GPS by going to multiple locations and tracking the accuracy by using the Google provided position and measuring it against the actual position the person is at. There are various different studies that have been conducted regarding this topic. One such study studies the accuracy of the GPS relative to multiple cellular data options, such as 2G, 3G, and 4G. Another study focuses on the potential of Google Maps location history data to characterize individual air pollution studies. My experiment is to solely test the accuracy of the Google Maps GPS relative to your actual location in the real world.

In order to make informed and reasonable predictions on the accuracy of the Google Maps GPS, I have developed an app that tracks users' locations. By making an account and pressing the button to get your location, the app shows your location with a red pin on a map. Additionally, by "friending" other people by inputting their email, you are able to see their locations as well. By comparing where the red pin is relative to my relative location in the real world, I am able to judge the accuracy of the Google Maps GPS. The accuracy will be tested in multiple locations, ranging from places with a high density of people with places with a low density of people to assure that there are no other outside factors that may adversely affect the study. Additionally, I will utilize multiple apps, such as Apple Maps to further verify the legitimacy of the location listed.

The rest of the paper is organized as follows: Section 2 details the multiple challenges that I ran across developing the location tracking app and performing the experiments; Section 3 focuses on the details of my solutions corresponding to the challenges that were encountered in Section 2; Section 4 presents the relevant details obtained as a result of the experiments, with Section 5 presenting related work to this experiment. Finally, Section 6 will give the conclusion on the study, as well as state any future work that may be done.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. App Requires Real-Life Unpredictable Variables (Internet Speed)

On top of the accuracy of the map itself, there are a multitude of unpredictable variables in real life that may affect the accuracy of the map even more. The first variable, internet speed, can play a huge factor. If the user isn't connected to the internet either through wi-fi or a cellular data connection, they are unable to update their location in the database, not being able to track their location as well as others'. Another unpredictable variable is the phone itself. As every phone is a little bit different, one wrong setting or outdated operating system can throw off the accuracy of the GPS pinpoint.

2.2. Google Services May Not Yield Proper Results For Various Reasons

The app relies on various google services, such as Google Firebase and Google Maps, to function properly. It relies on Google Firebase to read and display the information of various users, such as their user information and their location. It relies on Google Maps to show the locations of "friended" users. If any of these services are rendered inoperable for even a moment, the app will cease to function properly. Another disadvantage to using google services has to do with the Google Play store. Compared to the iOS App Store, the Google Play Store is difficult for new users to find apps. Google Play Store has limited options as to how developers can promote their apps. Additionally, Google is also unable to produce optimal results based on user ratings, popularity, etc.

3. SOLUTION

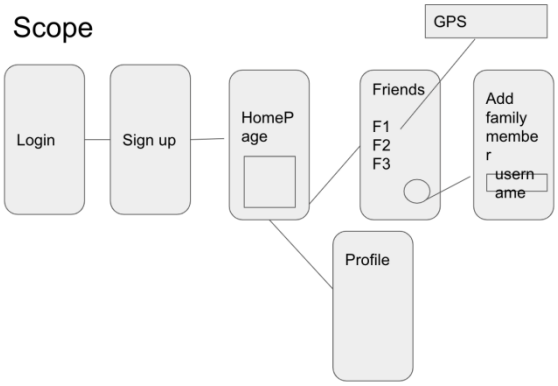


Figure 1. Overview of the solution

Figure 1 displays user navigation paths for the application. This application allows users to track each other's locations through a "friend" system. Once you have "friended" someone else, you are able to see your friend's location on a map, along with their name and their latitude/longitude coordinates. This application utilizes flutter and dart as its coding languages. Additionally, it uses Google Firebase as a database to store and encrypt user information [8]. First, the user will navigate to the "Get Location" page. Once at that page, you can press the "Get Location" button to use Google Map Services to pinpoint your location. Once your location has been pinpointed, you can press the "Push" button to push your latitude and longitude coordinates to the database.

Once your information has been pushed, your coordinates will be shown on screen for you and your friends to see.

The frontend of the app was built using Android Studio, using Google's Open Source UI Software Development Kit, Flutter. Flutter uses the object-oriented programming language, Dart.

The application, targeted towards any person who wants to track the location of others, consists of 3 parts. The first part is the login/sign up page. In order to sign up, you have to provide various pieces of information about yourself, such as your username and email. Following that, you are now able to either change the credentials of your newly created account, view the map where the locations of yourself or your friends are located, update your location, or "friend" people.

All the functions of this app are made possible with the communication of the front end of the app, as well as the back end, which is a cloud database, made possible by Google Firebase. Google Firebase provides many advantageous features such as user authentication by email, efficient communication between front and back end, and, most importantly, is supported by Flutter.

Get Location:

```
Future<void> _getLocation() async {
  setState(() {
    _error = null;
    _loading = true;
  });
  try {
    final _locationResult = await getLocation(
      settings: LocationSettings(ignoreLastKnownPosition: true),
    );

    setState(() {
      _location = _locationResult;
      _loading = false;
    });
```

Figure 2. Screenshot of code 1

```
onPressed: () {
  FirebaseDatabase.instance.ref().child("users/${FirebaseAuth.instance.currentUser!.uid}").update({
    "latitude": _location?.latitude?.toDouble(),
    "longitude": _location?.longitude?.toDouble()
  });
},
```

Figure 3. Screenshot of code 2

Figure 2 and Figure 3 use the Flutter Package named Location to obtain the most recent location of the user. It grabs the latitude and longitude coordinates from the device, pushes it to the account's Google Firebase database location. In other words, these pieces of code work in tandem with the database to display your latitude and longitude coordinates.

```dart
void getKeys() async {
  ref.onValue.listen((event) {
    final love = Map<String,dynamic>.from(event.snapshot.value as Map);
    keys = love.keys.toList();
    });
}

void getYourMarker() async {
  ref1.child("${FirebaseAuth.instance.currentUser?.uid}").onValue.listen((event) {
    final yourInfo = Map<String,dynamic>.from(event.snapshot.value as Map);
    print(yourInfo['uid']);
    markers.add(
        Marker(
          markerId: MarkerId(yourInfo['uid'] as String),//document['placeID'] as String
          icon: BitmapDescriptor.defaultMarkerWithHue(_pinkHue),
          position: LatLng(
            yourInfo['latitude'],
            yourInfo['longitude'],
          ),  // LatLng
          infoWindow: InfoWindow(
            title: yourInfo['username'] as String, //document['name'] as String
            snippet: yourInfo['description'] as String, //document['address'] as String
          ),  // InfoWindow
        ),  // Marker
    );
  });
}
```

Figure 4. Screenshot of code 3

The code in Figure 4 grabs important user information from the database, including their uid, latitude and longitude coordinates, username, and description. It continuously updates the information using a StreamBuilder.

```dart
@override
Widget build(BuildContext context) {
  getKeys();
  getYourMarker();
  print(markers);
  //getFriendsInfo();
  return Scaffold(
    appBar: AppBar(
      title: Text(""),
      leading: IconButton(
        icon: const Icon(Icons.arrow_back, color: Colors.black),
        onPressed: () => Navigator.of(context).pop(),
      ),  // IconButton
      backgroundColor: Colors.transparent,
      elevation: 0.0,
    ),  // AppBar
    extendBodyBehindAppBar: true,
    body: StreamBuilder(
        stream: ref1.orderByKey().onValue,
        builder: (context, snapshot) {
         if(snapshot.hasData) {
           final tilesList = <ListTile> [];
           final allUsers = Map<String,dynamic>.from(
               (snapshot.data! as DatabaseEvent).snapshot.value as Map);  // Map.from
           for (int i = 0; i < keys.length; i++) {
             if (allUsers.keys.contains(keys[i])) {
                  final nextMarker = Marker(
                    markerId: MarkerId(allUsers[keys[i]]['uid'] as String),//document['placeID'] as String
                    icon: BitmapDescriptor.defaultMarkerWithHue(_pinkHue),
                    position: LatLng(
                      allUsers[keys[i]]['latitude'],
                      allUsers[keys[i]]['longitude'],
                    ),  // LatLng
                    infoWindow: InfoWindow(
                      title: allUsers[keys[i]]['username'] as String, //document['name'] as String
                      snippet: allUsers[keys[i]]['description'] as String, //document['address'] as String
                    ),  // InfoWindow
                  );  // Marker
                  final nextTile = ListTile(
```

Figure 5. Screenshot of code 4

Shown in Figure 5, the StreamBuilder starting from Line 74 streams all the information of the people on the app. After, an “if “statement is made to filter out all the people that aren’t your friends. The StreamBuilder provides a constant stream of constantly updating friend information.

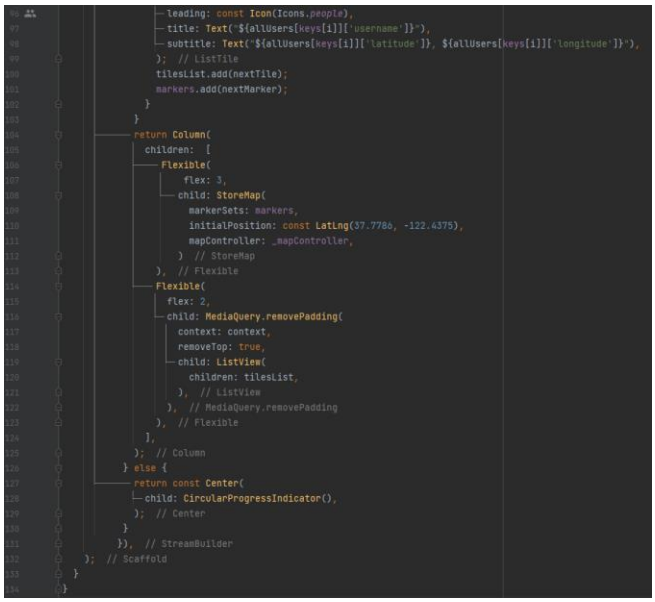


Figure 6. Screenshot of code 5

Figure 6 displays the StreamBuilder's constant stream of updating information displayed on the bottom of the Google Map. This information includes information about their name and their latitude and longitude as shown in Figure 7.

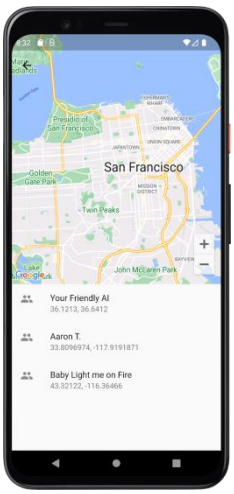


Figure 7. Screenshot of name, latitude, and longitude

```dart
class StoreMap extends StatelessWidget {

  const StoreMap({
    Key? key,
    required this.markerSets,
    required this.initialPosition,
    required this.mapController,
  }) : super(key: key);

  final LatLng initialPosition;
  final Completer<GoogleMapController> mapController;
  final Set<Marker> markerSets;

  @override
  Widget build(BuildContext context) {
    return GoogleMap(
        initialCameraPosition: CameraPosition(
          target: initialPosition,
          zoom: 12,
        ),  // CameraPosition
        markers: markerSets,

        onMapCreated: (mapController) {
          this.mapController.complete(mapController);
        }
    );  // GoogleMap

  }

}
```

Figure 8. Screenshot of code 6

Figure 8 displays the code that starts by setting up the map with its initial position (where the map initially shows when first launched). Then, it places markers on the map with the information from the StreamBuilder.

In conclusion, the location is grabbed by the device sending its own location to the Google Firebase database [10]. This database then constantly updates and sends information to the app, which displays this information on the bottom of the map.

Friend Users:

```dart
onPressed: () {
  ref1.orderByChild("email").equalTo(emailController.text).onValue.listen((event) {
    friendsProfile = event.snapshot.value;
    final cardList = Map<String,dynamic>.from(friendsProfile);
    cardList.forEach((key, value) {
      final nextCard = Map<String,dynamic>.from(value);
      var friendsUID = nextCard['uid'];
      //print(friendsUID);
      friendsUID1 = friendsUID ;
    });
    DatabaseReference listPush= FirebaseDatabase.instance.ref().child("friends/$uid");
    listPush.update({
      friendsUID1.toString(): "true",
    });
  });
  const snackBar = SnackBar(
    content: Text('You have successfully add this person!'),
  );  // SnackBar
```

Figure 9. Screenshot of code 7

The code displayed in Figure 9 allows users to type the information of their friend's name and email to make them their friend. After the request is sent, it adds that friend in the user's Google Firebase database.

4. EXPERIMENT

4.1. Experiment 1

This solution solves the problem of locating your loved ones by accurately pinpointing their position. This experiment aims to determine how accurate this position is displayed compared to other mapping programs such as Apple Maps and Waze [9].

This experiment was carried across 3 different devices: An LG G5, a Google Pixel 4a, and an Iphone XS Max as shown in Figure 10. Through the process of testing the app on multiple phones, many points of feedback were received, all of which are to be implemented in the future.

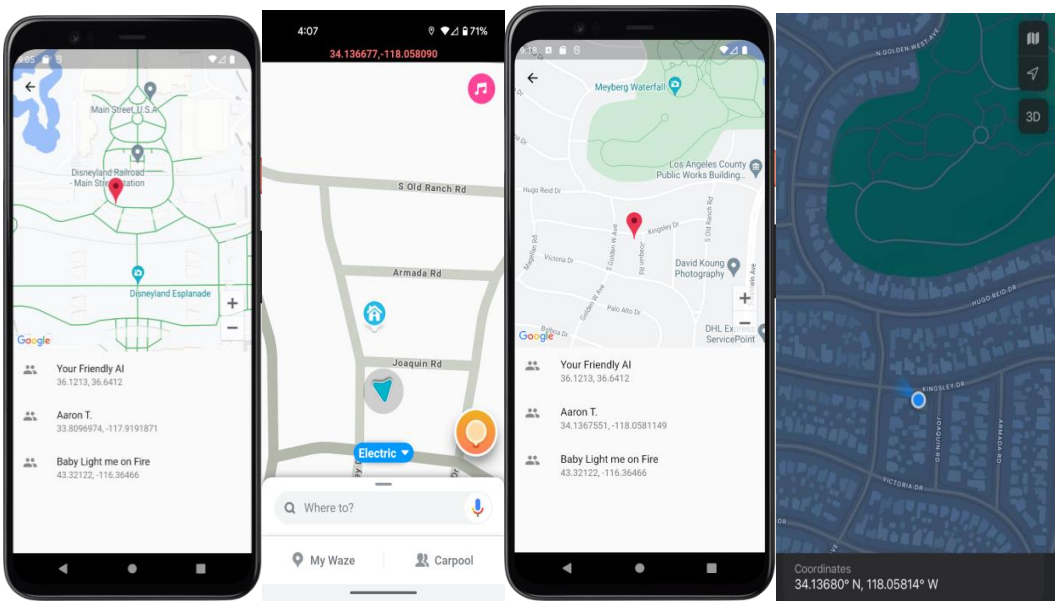


Figure 10. Screenshots of of locations

Google Maps	Apple Maps	Difference
34.1367551, -118.0581149	34.13680, -118.05814	-0.0000449, 0.0000251
33.8096974, -117.9191871	33.812511, -117.918976	-0.0028136, -0.0002111

Figure 11. Table of different location between Google Maps and Apple Maps

From Figure 11, it is clear that the latitudes and the longitude measurements of Google and Apple Maps do not differ by much.

In the beginning of the experiment, it was hypothesized that the coordinates of the Google Maps and Apple Maps services would be almost identical, and that hypothesis was somewhat true. With the deviations between the latitudes and longitudes measured never going above 0.003 (due to an outlier), it is safe to say that the two map services are accurate when compared to one another. The results are proven to be legitimate by the corroboration of the multiple sources, as well as evidence from other papers backing up this claim as well.

5. RELATED WORK

In the paper Zandbergen, Paul A., and Sean J. Barbeau. "Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones." presents an empirical analysis of the positional accuracy of location data gathered using a high-sensitivity GPS-enabled mobile phone [11]. The performance of the mobile phone is compared to that of regular recreational grade GPS receivers.

In the paper Payne, Rap. "Developing in Flutter." Beginning App Development with Flutter [12]. Apress, Berkeley. The paper discusses how flutter enables us to create apps that run on the Web, on desktop computers, and on mobile devices (which seems to be the main draw). This paper is very useful for the developers who are new to the flutter framework

In the paper Khawas, Chunnu, and Pritam Shah. "Application of firebase in android app development-a study" [13]. They show the application of Firebase with Android and aims at familiarizing its concepts, related terminologies, advantages and limitations. tries to demonstrate some of the features of Firebase by developing an Android app.

6. CONCLUSIONS

The safety of our loved ones, especially our elderly, is never a given. Loved Ones Locator provides a way to know that they are safe, or in the event that they are not, be able to locate them. The experiment tested whether the location shown on the device is accurate or not. This accuracy was measured by comparing the latitude and longitude of the user measured on the Google Map to those measured on other apps, including Apple Maps and Waze. Google Maps, Apple Maps, and Waze were used in the experiment because they are the most-used GPS apps [15].

Since the app shares very personal information such as name, email, and location information, it is of the utmost importance to ensure that this information cannot be obtained easily by the wrong people. Additionally, the way the information is transferred between the frontend and the database can be improved, making it more optimized. This will result in faster communication between the frontend of the app and the backend Firebase Database. As well as the above changes, there are some Quality of Life changes to be made as well. This will make the user experience more enjoyable, as well as simpler to navigate the menus.

To ensure the safety of the users' information, the method of data transfer and storage between the app and the database needs to be encrypted. To optimize the data transfer between the frontend and backend, redundancies can be eliminated from the code. For the quality of life improvements, a function could be added that lets the users edit their account data, such as their username and a confirmation pop-up confirming that users have created an account.

REFERENCES

[1] Tolman, Brian W., et al. "The GPS toolkit… open source GPS software." Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2004). 2004.

[2] Moroney, Laurence. "The firebase realtime database." The Definitive Guide to Firebase. Apress, Berkeley, CA, 2017. 51-71.

[3] Rajendran, G., M. Arthanari, and M. Sivakumar. "GPS Tracking Simulation by Path Replaying." International journal of innovative technology & creative engineering 1.1 (2011): 20-26.

[4] McNeff, Jules G. "The global positioning system." IEEE Transactions on Microwave theory and techniques 50.3 (2002): 645-652.

[5] Zamir, Amir Roshan, and Mubarak Shah. “Accurate image localization based on google maps street view.” European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010.
[6] Punja, Shafik G., and Richard P. Mislan. "Mobile device analysis." Small scale digital device forensics journal 2.1 (2008): 1-16.
[7] Li, He, and Lai Zhijian. “The study and implementation of mobile GPS navigation system based on Google Maps.” 2010 International Conference on Computer and Information Application. IEEE, 2010.
[8] Chatterjee, Nilanjan, et al. "Real-time communication application based on android using Google firebase." Int. J. Adv. Res. Comput. Sci. Manag. Stud 6.4 (2018).
[9] Sathiyanarayanan, Mithileysh, and Tobias Mulling. “Map navigation using hand gesture recognition: A case study using myo connector on apple maps.” Procedia computer science 58 (2015): 50-57.
[10] Moroney, Laurence. "An Introduction to Firebase." The Definitive Guide to Firebase. Apress, Berkeley, CA, 2017. 1-24.
[11] Zandbergen, Paul A., and Sean J. Barbeau. “Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones.” The Journal of Navigation 64.3 (2011): 381-399.
[12] Singh, Mohit, and G. Shobha. "Comparative Analysis of Hybrid Mobile App Development Frameworks."
[13] Khawas, Chunnu, and Pritam Shah. “Application of firebase in android app development-a study.” International Journal of Computer Applications 179.46 (2018): 49-53.
[14] Ahn, BoHyun. "Life Alert System for vulnerable adults who live alone." (2016).
[15] Pack, Michael, and Nikola Ivanov. “Are you gonna go my Waze?.” Institute of Transportation Engineers. ITE Journal 87.2 (2017): 28.

LISTING2SPEAK: A DATA-DRIVEN ANALYTICAL SYSTEM TO EVALUATE THE E-COMMERCE PRODUCT LISTING USING ARTIFICIAL INTELLIGENCE AND BIG DATA ANALYSIS

Zihan Xu[1] and Yu Sun[2]

[1]Crean Lutheran High School, 12500 Sand Canyon Ave, Irvine, CA 92618
[2]California State Polytechnic University, Pomona, CA, 91768,
Irvine, CA 92620

ABSTRACT

In e-commerce, product pages are important to the success of a website [4]. The ultimate goal of an e-commerce site is sales and it takes a great product page to achieve that. However, today, e-commerce entrepreneurs are confused about how to make their product page more attractive [5]. This paper designs a data-driven analytical system to analyze the relationship between different web page features with sales, in order to give users feedback on how to improve their product web pages [6].

KEYWORDS

Data Analysis, E-commerce, Artificial Intelligence.

1. INTRODUCTION

As the E-commerce system develops rapidly and becomes mature, e-commerce platforms have become an inseparable part of people's lives [7]. There is always an interesting phenomenon that two Amazon stores sell the exact same thing, but one sells much more than the other. In fact, for e-commerce websites, besides advertisement, the design of product pages is extremely important to the user's experience, and can even affect their purchasing decisions. First of all, the product details page is a product display area that can further stimulate customers' desire to buy. The biggest difference between online shopping and brick-and-mortar shopping is that customers cannot really touch the product, and can only rely on the product pictures displayed by the seller and the reviews of buyers who have already purchased to guess the quality of the product, and then decide whether to buy it or not [8]. In order to let products make better sales, the product details page of each product is our top priority. As a result, creating an E-commerce product page analysis project is necessary to help E-commerce sellers better improve their product web page. It brings users convenience and supports them to get better sales.

Some of amazon product listing optimization techniques and systems that have been proposed to identify the various factors that influence the success of your conversion and product listing sales, which allows the user‘s product to have a better ranking in its specific product category .These already existing tools claim to provide users with some of the best possible product

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 125-133, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121713

choices and ensure that their products are highly relevant when customers search for them by reading and analyzing all the data available on the market. However, these proposals assume that getting a higher ranking when customers search on Amazon is an absolute determinant of their sales, which is rarely the case in practice. Their implementations are also limited in scale because their system requires them to fill in product descriptions, keywords, etc. and upload images and videos on their own which will waste users a lot of time. Other techniques, such as "amify". They offer an expert team to optimize customers' Amazon Listing. Because people's preference may change over time, the method of using people to optimize product listing often results in inaccuracy.

In this paper, we follow the same line of research by … Our goal is to create an E-Commerce Product Listing reviewer based on a data analytical System [9]. Our method is inspired by one of machine learning ' s functions: it can use algorithms to parse data, learn from it, and make decisions or predictions about something in the world which are compliant with our website's goal. There are some good features both in our website page and inside our website's backend. First, we divide different kinds of products into different categories, which can produce more accurate forecasts. Secondly, we use the data collected from Amazon to do Sklearn in order to analyze and find the relationship between aspects of product listings and sales [10]. Sklearn provides us the ability to transform these huge amounts of data to predictions and analysis results. Thirdly, we gathered a huge amount of data by using selenium to do web scraping. Although our computers do not allow us to collect all the sales data, top 100 products in each Amazon category can still enable us to make precise analyses.

The following of this paper is organized as follows: Section 2 details the challenges we encountered while writing the backend of the program and during our experiments; Section 3 highlights the details of our solutions to the challenges mentioned in Section 2 and the main METHOD used in our project; Section 4 presents the details related to the evaluation of our results by the experiments we performed, followed by the related work of our project in Section 5. Finally, Section 6 gives the conclusion and points out the existing limitations of this project and the aspects and work that can be improved in the future.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Finding All the Features that can be Included in Our Website Page

The first challenge is finding all the features that can be included in our website page. In an Amazon product page, there is plenty of information such as description, pictures, videos and ratings [11]. Definitely I can not include all these features in my evaluation and feedback website because this would take our advice out of focus and some of the information can not convert to numbers in order to find a data relationship. To solve this challenge, my professor and I brainstormed on what users really want to make changes on their products. First, we eliminated ratings because ratings are something that cannot be changed by the design of the product page, as customers will only rate the product once they get the actual product, so the rating depends more on the quality of the product itself. Then we decided to include the number of keywords, description words, pictures and videos in our website which allows us to build AI models to analyze data relationships.

2.2. Difficulty Collecting Data on Amazon Sales

Amazon's whole product sales data is so huge that we can't collect all the sales data and do data analysis and future prediction with just the arithmetic power of our computers. We need a method that does not need to collect so much data but can provide a database for my ai model. We then found that collecting the top 100 lists for each product type on Amazon would provide both an adequate data set and not require excessive computer computing power.

3. SOLUTION

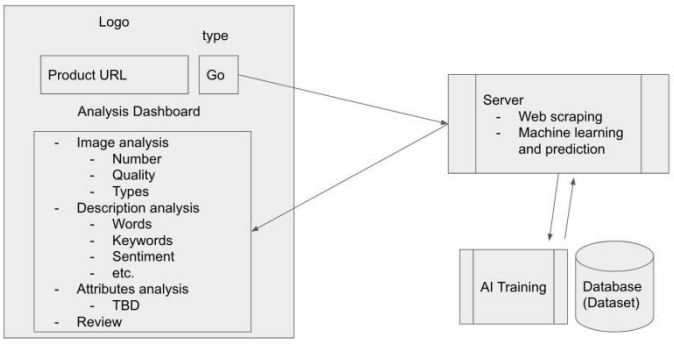
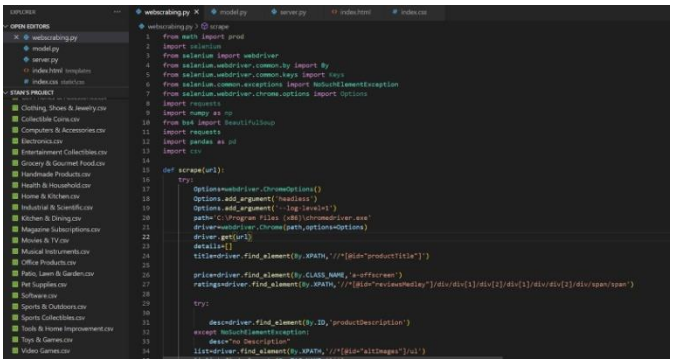


Figure 1. Overview of the solution

Listing2Speak is a Data-Driven Analytical System aimed to Evaluate the E-Commerce Product Listing using Artificial Intelligence and Big Data Analysis [12]. Listing2Speak is a powerful system for users to consider the way up their products' sales. It is able to analyze the input data and give users reviews immediately. There are four main parts in Listing2Speak, which are the website itself, server, AI model and database [13]. The website page of Listing2speak is the platform for users to input their product listing and get reviews. Our system's reviews on aspects of product listing include description, number of description words, number of images, number of videos, number of keywords and sales. After a user enters their product listing page into our web page and chooses their product type, our system will analyze different aspects of their product page and give them feedback on where they should make improvements. Our algorithm adapts to the buying patterns of its customers since the database contains all top 100 products in each amazon product category. Using the datasets collected by web scraping codes, Listing2Speak created a vector regression model in SKlearn machine learning in order to predict and analyze input data. This AI based method provides our system ability to give the most appropriate improvement suggestion for users' product listing in different product categories. Due to the regression model, our system has identified various factors that play into the success of conversions and sales of product listings. Therefore, based on the data relationship of the listing aspects of the product page constructed by the AI model, the server is able to make predictions with input data.

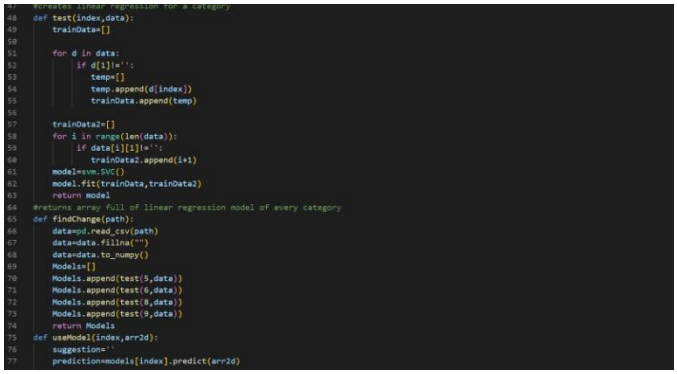


```
from math import prod
import selenium
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.common.keys import Keys
from selenium.common.exceptions import NoSuchElementException
from selenium.webdriver.chrome.options import Options
import requests
import numpy as np
from bs4 import BeautifulSoup
import requests
import pandas as pd
import csv

def scrape(url):
    try:
        Options=webdriver.ChromeOptions()
        Options.add_argument('headless')
        Options.add_argument('--log-level=1')
        path='C:\Program Files (x86)\chromedriver.exe'
        driver=webdriver.Chrome(path,options=Options)
        driver.get(url)
        details=[]
        title=driver.find_element(By.XPATH,'//*[@id="productTitle"]')

        price=driver.find_element(By.CLASS_NAME,'a-offscreen')
        ratings=driver.find_element(By.XPATH, '//*[@id="reviewsMedley"]/div/div[1]/div[2]/div[1]/div/div[2]/div/span/span')

        try:

            desc=driver.find_element(By.ID,'productDescription')
        except NoSuchElementException:
            desc="no Description"
        list=driver.find_element(By.XPATH,'//*[@id="altImages"]/ul')
```

Figure 2. Screenshot of code 1

3.1. Web Scraper

This is a screenshot of my web scaper's codes. By using selenium, a comprehensive project with a set of tools and libraries to support web browser automation, my web scraper is able to scrape amazon products' sales, number of description words, number of keywords, and number of videos from each amazon product category. When my scraper runs, the program automatically opens the chrome browser and scrapes the data I need from those specific rows of the inspect page I assigned in the scraper. The data collected by web scraper is stored in csv files for Listing2Speak's database. With the database, we then developed an AI model in order to build up the main algorithms of Listing2Speak. As a result, it leads to my AI model codes.



```
#creates linear regression for a category
def test(index,data):
    trainData=[]

    for d in data:
        if d[1]!='':
            temp=[]
            temp.append(d[index])
            trainData.append(temp)

    trainData2=[]
    for i in range(len(data)):
        if data[i][1]!='':
            trainData2.append(i+1)
    model=svm.SVC()
    model.fit(trainData,trainData2)
    return model
#returns array full of linear regression model of every category
def findChange(path):
    data=pd.read_csv(path)
    data=data.fillna("")
    data=data.to_numpy()
    Models=[]
    Models.append(test(5,data))
    Models.append(test(6,data))
    Models.append(test(8,data))
    Models.append(test(9,data))
    return Models
def useModel(index,arr2d):
    suggestions=''
    prediction=models[index].predict(arr2d)
```

Figure 3. Screenshot of code 2

3.2. Training and Predicting

The second main part I implemented in Listing2Speak is the AI model. We used machine learning libraries Sklearn for training and predicting datasets in the model. Sklearn is a very powerful machine learning library provided by a third party in Python that covers everything from data pre-processing to training models. I mainly utilize sklearn's data prediction and regression features. The above code shows how I create a linear regression model for each Amazon product category and how it returns an array full of linear regression models of every category. Based on the linear relationship of each product listing aspect we found from the linear regression model, once users input their own product listing page, the model can further make predictions on the users' listing pages' percentile among all Amazon e-commerce stores.

3.3. Website's UI

Figure 4. Screenshot of website

The frontend website is developed using css and HTML [15]. CSS is a program language that defines style structures such as fonts, colors, positions, etc. and is used to describe the way information is formatted and displayed on a web page. CSS styles can be stored directly on HTML pages or in separate style sheet files. This screenshot shows the main UI of Listing2Speak's website. The two red fields below the title are for entering your product listing URL and selecting your product type. After you click on the go button, our backend will start computing and display the content in the bottom half of the image within a few seconds. On the left side of the analysis area below, it shows different aspects that Listing2Speak analyze on of users' input's product listing page. In contrast, in the section on the right, it tells users if they need to improve their product pages and the percentile of their product listing pages in our dataset.

4. EXPERIMENT

In order to verify that our solution can effectively solve problems at different levels and have good user feedback, we decided to select multiple experimental groups and comparison groups for several experiments. For the first experiment, we want to prove that our solution works stable and continuously, so we choose a group size of 100 different products in 4 categories from different retailers and record the revenue changes of them. The goal of the first experiment is to verify if our AI could analyze all products and give the effectively improve suggestions if the AI works good for all categories of the products. Through sampling 10 groups of different targets. Result is collected by statistics if the AI could find and analyze correctly. Experiments have shown that almost all targets in different types tested with the good result which could appearly improve the sell amount of the retailers. Retailer sells Toothpaste has the most improve rates, which means our AI are works more better in analyze pictures. This experiment could explain that the data sets do have a obvious impact on the finding the elements in the sells' website and how to make it better to sell more products. The average increasing sells rate of 4 different categories of products shows below:

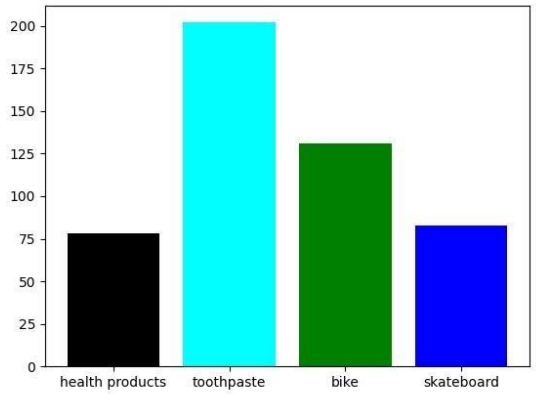


Figure 5. Graph result of experiment

A good user experience is as important as a good product. So a perfect solution should have excellent user experience feedback. In order to prove that our solution has the best user feedback, we specially designed a user experience questionnaire base on the US system usability questionnaire rules. We statistics the feedback result from 100 users, Track the user's data for 10 days using, let them explore freely on the functionality of the web. we divide those users into Four different groups sells health products, toothpaste, bike, skateboard. The goal of the first experiment is to verify if our AI could analyze all products and give the effectively improve suggestions high feedback scores shows high performance We collect the feedback scores form these 4 different group of users and analyze it. Experiments have shown that users who ages who sells toothpaste give the highest result feedback to our web. Which may because of they increasing their revenue a lot by using the web. The experiment graph shows below:

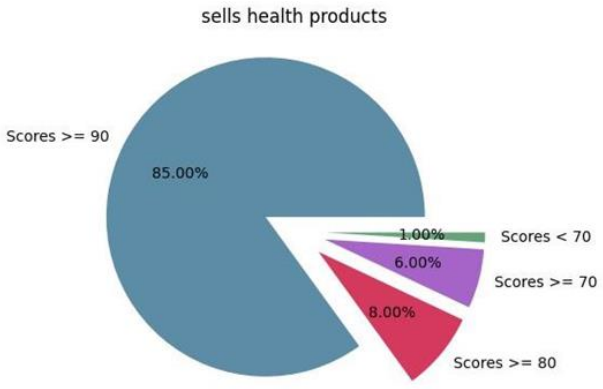


Figure 6. Chart of sells health products

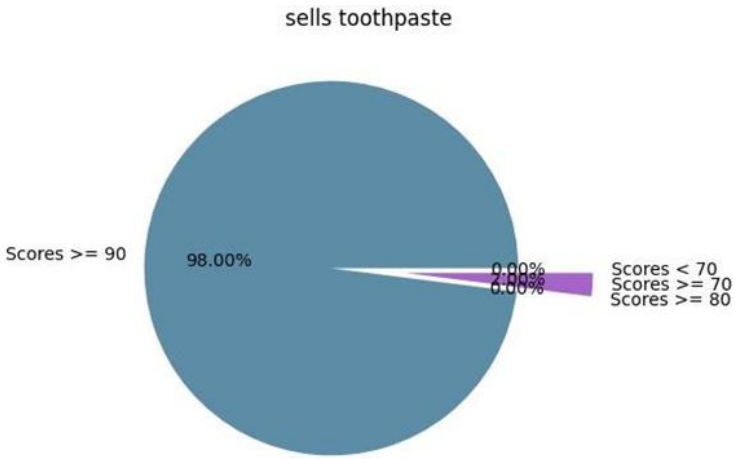


Figure 7. Chart of sells toothpaste

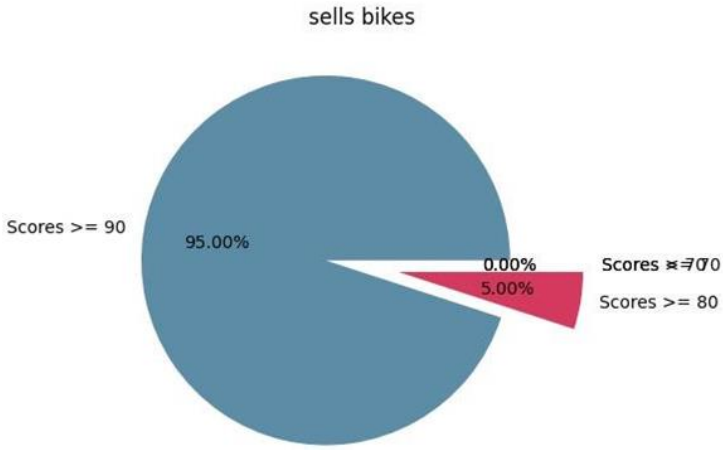


Figure 8. Chart of sells bikes

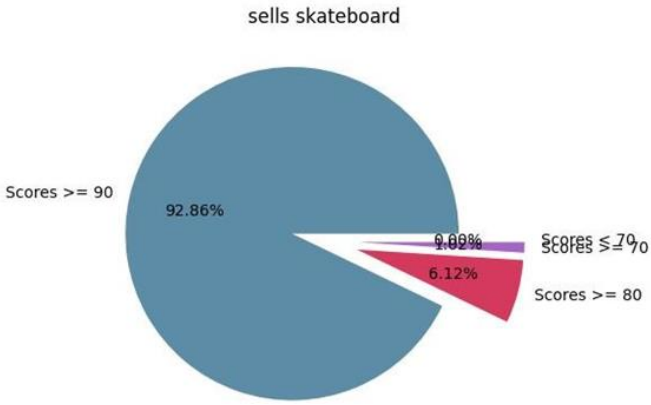


Figure 9. Chart of sells skateboard

5. RELATED WORK

Yan-Kwang C. et al. synthesized the visual attention literature and shelf space allocation theory to build a mathematical planning model with genetic algorithms to find the optimal solution to the focal problem by investigating the effects of display mode (image text vs. plain text) and display format (list vs. array) on users' online shopping performance and attitudes [1]. Their method combined the results of experiments and surveys to provide data support and eventually constructed a mathematical model using functional equations and genetic algorithms. We are using existing data from the Amazon website and linear regression method to create a machine learning model. Yan-Kwang C. et al gives readers a concept how to increase their product's profit, by contrast, Listing2Speak is used to analyze specific user's input and provide ways to improve their sales.

Weiyin Hong et.al's work focuses on two e-commerce page design features: presentation mode and information format [2]. Through their experimental design and observations, they show that the graphic display mode and the list information format are higher than the text-only display mode and the array information format in terms of user performance and attitude, respectively. Their passage is dedicated to proving that combining graphs is better than simple text information. In contrast, Listing2Speak is not concerned with the overall page design of the product page, but rather with the some specific design aspects of the product page that are linked to sales.

Using advanced quantum knowledge and mathematical algorithms, Naoki Nishimura et.al proposed a method to optimize item listings on e-commerce websites, excluding the effect of product popularity and diversity on sales [3]. Their work decomposes the problem by transforming and decomposing it to exclude confounding variables, while Listing2Speak divides the machine learning database into different categories according to Amazon product types to analyze their linear regression separately. The work of Naoki Nishimura et al. has strengths in the mathematics and scientific theory behind their work that support their findings. Listing2Speak uses a machine learning Ai model that focuses more on real-time analytics and provides an improvement solution for individual product pages.

6. CONCLUSIONS

Currently I don't have many project analysis categories, and the database of keyword analysis is not large enough, which leads to most of the analyzed ratings being lower than they would have been. And another limitation about our project is that it can only work when you input an Amazon link which means that you need to distribute your product on Amazon first, only then you can use our e-commerce reviewer. In addition, I want my project to evaluate all the ecommerce products pages, but now I only gathered the database from only the Amazon website. As a result, our project is sometimes not practicable if users don't want to sell their product on Amazon or just test their design demo.

In the future, we are going to gather data from different kinds of e-commerce platforms to make users' choice more diverse. For example, if users want to sell their products through Shopify, they can choose the Shopify category and enter in their Shopify product page's link [14]. We are going to not only expand data for different platforms, but also the keywords dataset which can make our results more accurate.

REFERENCES

[1] Yan-Kwang Chen, Fei-Rung Chiu, Ciao-Jyun Yang, "An Optimization Model for Product Placement on Product Listing Pages", Advances in Operations Research, vol. 2014, Article ID 357693, 9 pages, 2014. https://doi.org/10.1155/2014/357693

[2] Weiyin Hong, James Y.L. Thong, Kar Yan Tam, Designing product listing pages on e-commerce websites: an examination of presentation mode and information format, International Journal of Human-Computer Studies, https://doi.org/10.1016/j.ijhcs.2004.01.006.

[3] Front. Comput. Sci., 16 July 2019, Sec. Quantum Engineering and Technology, https://doi.org/10.3389/fcomp.2019.00002

[4] Tian, Yan, and Concetta Stewart. "History of e-commerce." Encyclopedia of e-commerce, egovernment, and mobile commerce. IGI Global, 2006. 559-564.

[5] Burt, Steve, and Leigh Sparks. "E-commerce and the retail process: a review." Journal of Retailing and Consumer Services 10.5 (2003): 275-286.

[6] Markovsky, Ivan, and Florian Dörfler. "Behavioral systems theory in data-driven analysis, signal processing, and control." Annual Reviews in Control 52 (2021): 42-64.

[7] Tran, Lobel Trong Thuy. "Managing the effectiveness of e-commerce platforms in a pandemic." Journal of Retailing and Consumer Services 58 (2021): 102287.

[8] Rajamma, Rajasree K., Audhesh K. Paswan, and Gopala Ganesh. "Services purchased at brick and mortar versus online stores, and shopping motivation." Journal of Services Marketing (2007).

[9] Calixto, Iacer, et al. "Using images to improve machine-translating e-commerce product listings." Association for Computational Linguistics (ACL), 2017.

[10] Komer, Brent, James Bergstra, and Chris Eliasmith. "Hyperopt-sklearn." Automated Machine Learning. Springer, Cham, 2019. 97-111.

[11] Haque, Tanjim Ul, Nudrat Nawal Saber, and Faisal Muhammad Shah. "Sentiment analysis on large scale Amazon product reviews." 2018 IEEE international conference on innovative research and development (ICIRD). IEEE, 2018.

[12] Arena, Fabio, and Giovanni Pau. "An overview of big data analysis." Bulletin of Electrical Engineering and Informatics 9.4 (2020): 1646-1653.

[13] McCarthy, John. "What is artificial intelligence." URL: http://www-formal. stanford. edu/jmc/whatisai. html (2004).

[14] Alfida, P. P., et al. "SHOPIFY – The Smart Shopping Cart." International Research Journal of Engineering and Technology 6.6 (2019): 616-620.

[15] Alawar, Mariam W., and Samy S. Abu-Naser. "CSS-Tutor: An intelligent tutoring system for CSS and HTML." (2017).

LABELLE: A DEEP LEARNING APP THAT HELPS YOU LEARN BALLET

Sarah Fan[1], Kevin Guo[2] and Yu Sun[3]

[1]Sage Hill School, 20402 Newport Coast Dr, Newport Beach, CA 92657
[2]University of Southern California, Los Angeles, CA 90007, Irvine, CA 92606
[3]California State Polytechnic University, Pomona,
CA, 91768, Irvine, CA 92620

ABSTRACT

Human Pose Estimation has proven versatility in improving real-world applications in healthcare, sports, etc. [1]. Proper stance, form and movement is instrumental to succeeding in these activities. This paper will explain the research process behind the deep learning mobile ballet app, LaBelle [2]. LaBelle takes in two short videos: one of a teacher, and one of a student. Utilizing MediaPipe Pose to identify, analyze, and store data about the poses and movements of both dancers, the app calculates the angles created between different joints and major body parts. The app's AI Model uses a K-means clustering algorithm to create a group of clusters for both the student dataset and the teacher dataset [3]. Using the two sets of clusters, LaBelle identifies the key frames in the student-video and searches the teacher cluster set for a matching set of properties and frames. It evaluates the differences between the paired frames and produces a final score as well as feedback on the poses that need improving. We propose an unsupervised guided-learning approach with improved efficiency in video comparison, which is usually both time and resource consuming. This efficient model can be used not just in dance, but athletics and medicine (physical therapy like activities) as well, where stance, form, and movements are often hard to track with the naked eye.

KEYWORDS

Artificial Intelligence, Machine Learning, Deep Learning, K-means Clustering, Computer Vision, Ballet.

1. INTRODUCTION

Deep learning methods, especially in computer vision, are beginning to gain more and more popularity as technology advances to produce accurate models. Human Pose Estimation (HPE) addresses this area of deep learning research, computer vision [4]. It has versatile real-world applications in healthcare, sports, etc. [5]. We will be discussing dance, specifically ballet, which is the foundation for most forms of dance. Ballet serves as a steppingstone to most other forms of dance, so it is universally appreciated as one of the best ways to improve one's dance skills.

There are many ways to learn ballet [6]. With today's technology, you can watch a professional or teacher recording and learn through imitation. Alternatively, you can record yourself dancing, and compare it to the teacher's dance, critiquing your own movements to improve. Manually comparing two videos, however, is tedious, inefficient, and resource-consuming. We want to develop a deep learning-based App, *LaBelle*, that helps users learn ballet. To achieve this, we first needed to develop an accurate deep learning model that can classify, detect, track, and segment videos for useful data and information. For these tasks, deep convolutional neural

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 135-146, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121714

networks (CNN) are the most efficient, given their ability to extract patterns from target video clips (images) with high precision and accuracy [7]. Next, we need a good algorithm to efficiently compare two videos. This algorithm must process and store data for hundreds to thousands of video frames both accurately and quickly, to ensure a smooth and consistent user experience. Lastly, we need to design a clean UI layout in a practical app to display our score and feedback.

The rest of the paper is organized as follows: Section 2 describes our research background, direction, and crucial research elements in detail; Section 3 presents our approach to resolving any challenges we faced and relevant details about the experiments we carried out; Section 4 presents the results and analysis; Section 5 gives a brief summary of other works that tackle a similar problem; Section 6 gives closing remarks and discusses the future of this project.

2. CHALLENGES

Figure 1 shows the problem our research aims to solve. The process can be summarized as the following: (1) user inputs teacher's dance video (maximum length two minutes) and student's dance video (maximum length 40 seconds); (2) app processes the videos and applies our deep learning algorithm (using HPE) to the frames to analyze the data; (3) user receives a standardized score and is shown video frames with several poses to improve on.

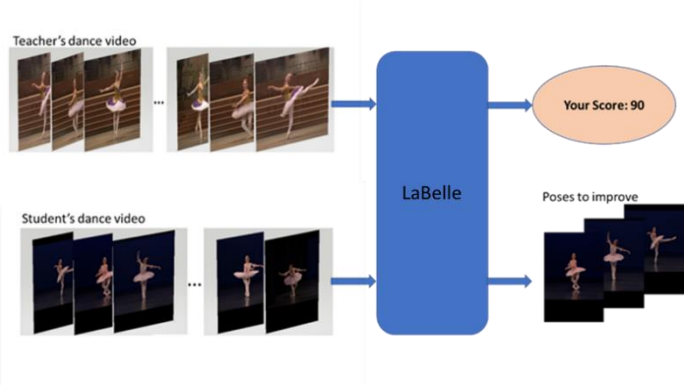


Figure 1. Problem definition

2.1. Human Pose Estimation

Human Pose Estimation (HPE) is a computer vision task that includes detecting, associating, and tracking semantic key points. On each video frame of target videos, the key points on a body need to be identified and tracked to describe the dancer's pose, referred to as landmarks. The connection drawn between key points and their movement is used to model the dancer's movement. There are three models we follow to process the human body: skeleton-based model, contour-based model, and volume-base model (shown in Figure 2).

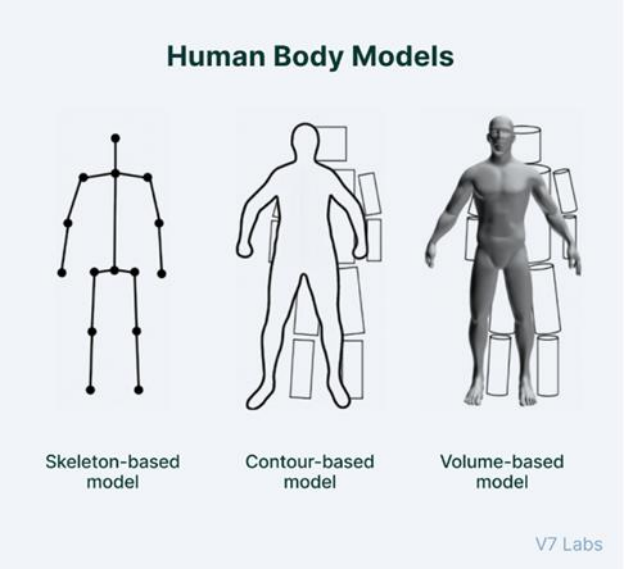


Figure 2. Human body models

The performance of semantic key-point tracking in live video requires high computational resources. The teacher's video clip may be up to 3 minutes in length, with each second consisting of 24, or 60 frames. That's 4000-10,000 frames in total. For the student's video, it can vary from 30 seconds to one minute. That's up to 1800 frames. This is an extremely large set of data to be collecting and processing.

2.2. Key Frame Detection

Within a dance video clip, the key frame is the frame/shot defining the starting or ending point of a transition (by the dancer). When we compare two video clips, one of the teacher and one of the student, both will be dancing to the same piece of music, but the video lengths will differ. Not only that, the starting and ending points of the video clips will be different. Additionally, the performers' height and weight are rarely the same, with arm, leg, and torso/body proportions making contour and volume vary greatly despite performers completing the same dance. Finally, the angle at which the videos are taken may differ.

2.3. Clustering Algorithms

It's almost impossible to compare two videos frame by frame in real time, even if they were the same length. Even if the student video is short, it could still consist of up to 900 frames. To compare the two videos efficiently, we need to group the frames together in a process called clustering, where frames with similar properties are group around a centroid set of desired properties.

2.4. Datasets

Strong datasets are a fundamental factor in a deep learning system. An extensive and diverse dataset is a crucial requirement for the successful training of a deep neural network. For our research purposes, we decided to create our own dataset, deriving data from sample videos of professionals, paired with several videos of students performing to the same piece. These videos are cut into multiple clips: 10 professionally performed classical dances are downloaded, and correspondingly, 10X6 student video clips are generated.

3. SOLUTION -- LABELLE

Our solution to this problem is deep learning-based App: LaBelle. LaBelle is structured into seven layers of processing. The first layer (L1) consists of two MediaPipe Post models (M1 and M2) to process the input video clips, one for the teacher video clip and one for the student video clip. The second layer (L2) consists of two blocks (B1 and B2) that calculate the major angles for important joints in the skeleton-based dancer model. The third layer(L3) consists of two blocks (C1 and C2) that group the angle-based properties into clusters for both teacher and student. The fourth layer (L4) consists of one block (D11) to detect the key frames for student video clip. The fifth layer (L5) consists of one block (D12) to find the centroids' feature property for each cluster of student video clip based on the starting and ending frame. The sixth layer (L6) has one block (F2) to find the matching property in the teacher video clip for each centroid's feature property generated from the student video. The last layer (L7) has one block (S12) to compare matching properties of teacher and student, generating a score to send to the user along with the frames corresponding to the centroids bearing lowest score to improve on.

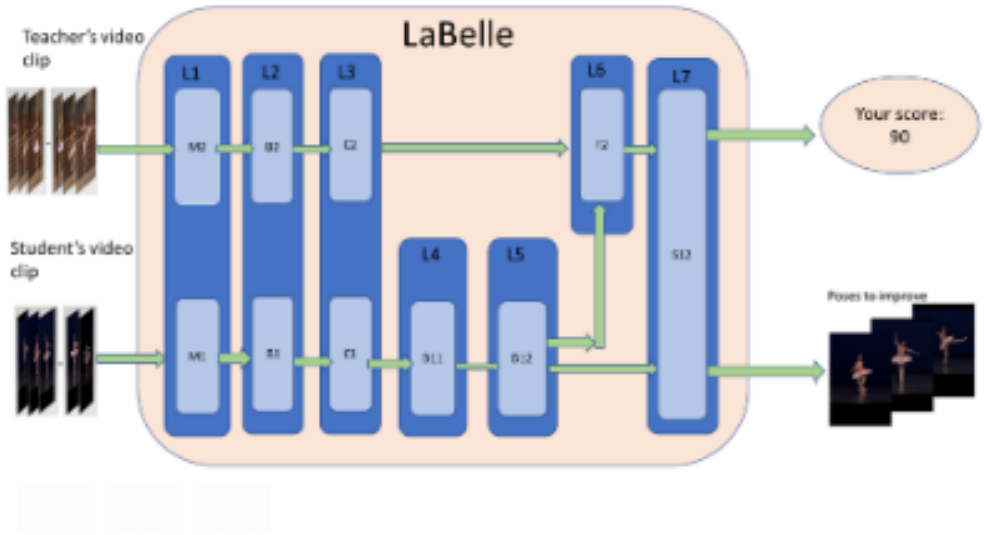


Figure 3. Our solution

3.1. MediaPipe Pose (L1 Layer: M1 and M2)

MediaPipe is a cross-platform library developed by Google. It provides high quality ready-to-use Machine Learning solutions for computer vision tasks. MediaPipe Pose is one of the solutions we use for high-accuracy body pose detection and tracking [9]. The Pose Landmark Model (BlazePose GHUM 3D) allows us to predict the location of 33 pose landmarks.

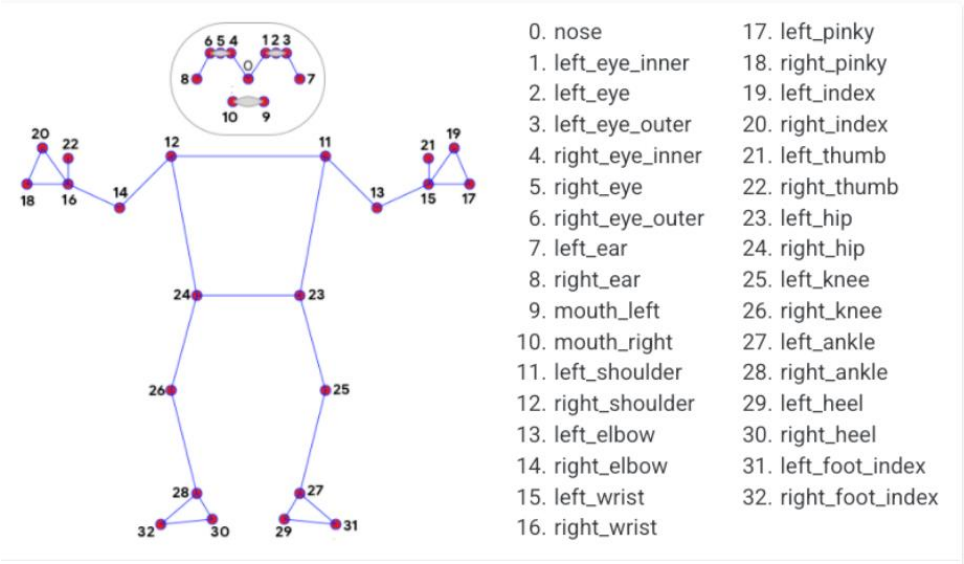


Figure 4. MediaPipe Pose landmarks [5]

3.2. Property Extraction: Angle Calculation (L2 Layer: B2 and B1)

MediaPipe Pose detects/tracks 33 landmarks, outputs 33 landmark connection. However, to improve the performance, we define a set of key angles between two landmark connections. For example, we use (∠16, 14 ,12) for the right elbow. We use (∠14, 12, 24) for the right shoulder. We use (∠24, 26, 28) for the right knee. This is done for the left side of the body as well. As shown in Figure 5 and Figure 6, we use MediaPipe's built in libraries to plot the human's pose on a 3D graph. The coordinates of 3 adjacent joints are called into the plot angle function of our algorithm, which calculates and returns the angle between the three landmarks. This process is repeated and carried out for each video frame until all desired joint angles are returned, resulting in a final tuple array consisting of all major body angles for individual video frames.

The pre-defined key angles can be altered. As dancers' skill level progresses, more angles in the hand and feet can be applied to provide an even more rigorous level of tracking and critique.

```
angle, image = plot_angle(mp_pose.PoseLandmark.LEFT_SHOULDER.value,
                          mp_pose.PoseLandmark.LEFT_ELBOW.value,
                          mp_pose.PoseLandmark.LEFT_WRIST.value, landmarks, image, h, w + val)
```

Figure 5. MediaPipe code implementing key joints into angle calculation

```
def plot_angle(p1, p2, p3, landmarks, image, h, w):
    # Get coordinates
    a = [landmarks[p1].x,
         landmarks[p1].y]
    b = [landmarks[p2].x, landmarks[p2].y]
    c = [landmarks[p3].x, landmarks[p3].y]

    # Calculate angle
    angle = calculate_angle(a, b, c)
    # print(angle)
    draw_angle(tuple(np.multiply(b, [w, h]).astype(int)), image, round(angle))
    return angle, image
```

Figure 6. MediaPipe code to generate angles given key joints

Table 1. Example of key properties (angles) for random frames

	∠16, 14 ,12	∠14, 12, 24	∠24, 26, 28	∠11,13,15	∠13,11,23	∠23, 25, 27
Frame 0	50	80	180	50	80	120
Frame 1	180	120	100	45	180	180
Frame 2	70	45	180	180	45	70
Frame 3	90	35	130	120	90	180
Frame 4	45	75	120	180	180	120
Frame 5	120	140	180	110	55	60
Frame n	30	160	150	180	45	180

3.3. K-means Clustering (L3 Layer: C2 and C1)

We limit our teacher's video length to about 3 minutes. For each second, we can generalize about 30 frames. That's approximately 5400 frames in total. For the students' video, typically 30 seconds long, there's approximately 900 frames. We want to cluster our large number of frames into similar groups based on their properties, using our key angles as properties. The videos uploaded by the user will change each time, so it's impossible to preset cluster/centroid properties. To solve this problem, we need an unsupervised and autonomous clustering algorithm.

K-means (Macqueen, 1967) is widely used in unsupervised learning algorithms that tackle clustering tasks. It uses "centroids," different randomly initiated points in the data, and assigns every data point to the nearest centroid (distance is dependent on similarity between data point and centroid). Once every data point has been assigned, the centroid is adjusted to the average of all the points assigned to it. Sklearn.cluster. KMeans works with KMeans using Lloyd's or Elkan's algorithm.

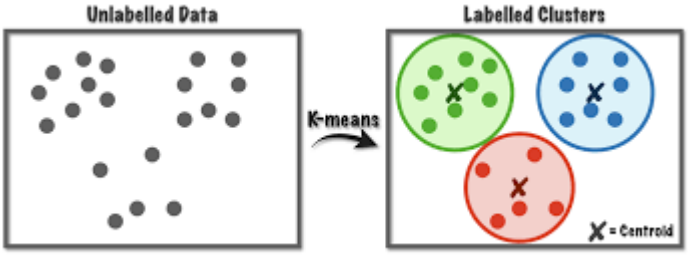


Figure 7. K-means clustering: unsupervised clustering algorithm

3.4. K-mean Hyper Parameter Tuning using Silhouette Score

Both video clips need to be processed using clustering. To determine the optimal number of clusters for both video clips, we use silhouette scoring to tune the hyper parameters. One way we can do this is by assigning a random number of clusters ranging from 3 to 30 with random initial centroid values. Alternatively, we can assign meaningful initial centroid values: by studying Basic Ballet Positions and Movements shown in Figure 8, we can summarize ballet positions into 14 initial centroid frames [15]. We summarize the properties for those unique positions shown in Table 2. With this guide, we use random selection, random range (14, 30), to determine the number of clusters, and initial centroid properties will be used to initialize the sklean.cluster. KMeans algorithm.

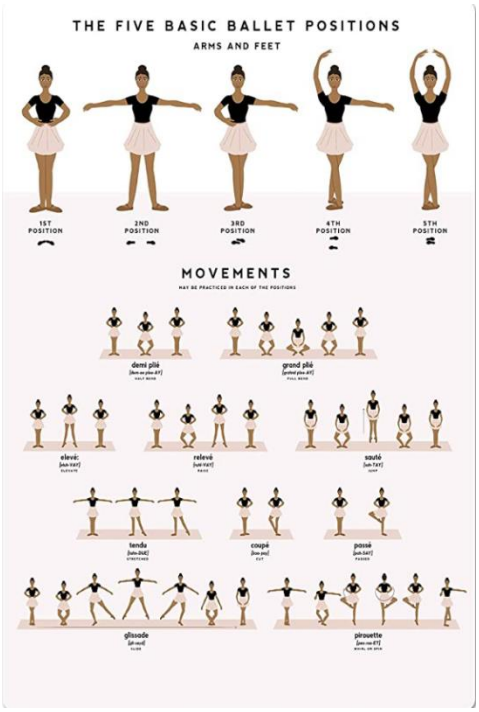


Figure 8. Basic ballet positions and movements

Table 2. Initial centroids based on basic ballet positions and movements

	∠16, 14 ,12	∠14,12,24	∠24, 26, 28	∠11,13,15	∠13,11,23	∠23, 25, 27	∠12, 24, 26	∠11, 23, 25
1st Position	120	30	180	120	30	180	180	180
2nd Position	180	90	180	180	90	180	160	160
3rd Position	120	30	180	180	90	180	180	180
4th Position	180	160	180	180	90	180	180	180
5th Position	160	160	180	160	160	180	160	160
Demi Plie	120	30	120	120	30	120	160	160
grand Plie	120	15	60	120	15	60	120	120
Coupe	120	30	90	120	30	180	120	180
passe	120	30	60	120	30	180	90	180
eleve	120	30	180	120	30	180	160	160
tendu	180	90	180	180	90	180	180	160
Glissade1	180	60	160	180	60	180	160	135
Glissade2	180	90	180	180	90	180	135	135
Glissade3	180	60	180	180	60	160	135	160

3.5. Key Frame Identification (L4 and L5 Layer: D11 & D12)

Based on the clusters generated from the student's video clip, each cluster will determine its own key frame, then the middle frame is identified and used as key frame.

3.6. Compare and Generate Final Score (L6 and L7 Layer: F2 & C12)

To compare the two videos, we go through labels of all the centroids in the student video clip and find the modified labels by fitting the centroids within the teacher's clusters [10]. Then we find the most similar property from the teacher's video clip, producing the most similar poses for both student and teacher, allowing us to calculate the score. Show in Figure 9 is part of the algorithm, given a predetermined video frame. We use matrix norms from linear algebra to help us find the smallest distance (least difference between frames). We conclude the final score and output the pose that generates the lowest score, which means least matching, for user to improve in the future practice.

```
def get_nearest_neighbor(image, indexes, frames):
    a = np.array(image)
    min_dist = sys.maxsize
    nearest = indexes[0]
    for idx in indexes:
        b = np.array(frames[idx])
        dist = np.linalg.norm(a - b)
        if min_dist > dist:
            nearest = idx
            min_dist = dist
            # print(min_dist, nearest)
    return nearest
```

Figure 9. Portion of algorithm used to find the two most similar frames

4. EXPERIMENT

In this section we give some examples with real data sets to demonstrate the performance of the proposed k-means algorithm. We show these unsupervised learning behaviors to get the best number of clusters for our algorithm to match the video between teacher and student. The dataset we are using is from the data we collected through the training. To find the best K to using in the production level we try to train the dataset for both 40 k-means random states and 80 random states, run several k-means, increment k with each iteration, and record the SSE.

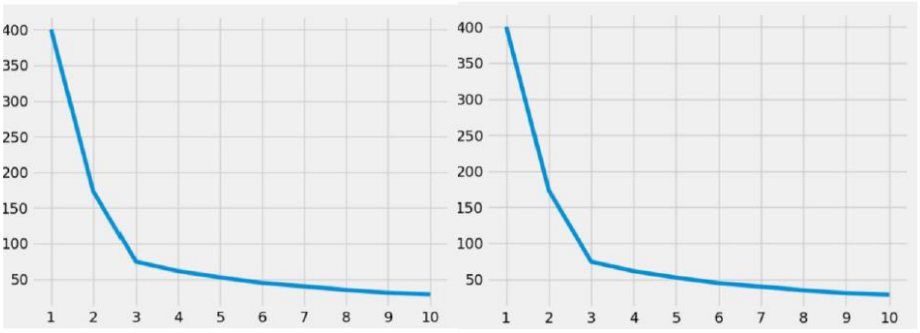


Figure 10. SSE curve with N cluster changes for tested with 40 and 80 random states

The diagram shows K bigger than 3 has the acceptable SSE control.

The silhouette coefficient is a measure of cluster cohesion and separation. It quantifies how well a data point fits into its assigned cluster. instead of computing SSE, compute the silhouette coefficient, to determine the best k for our algorithm:

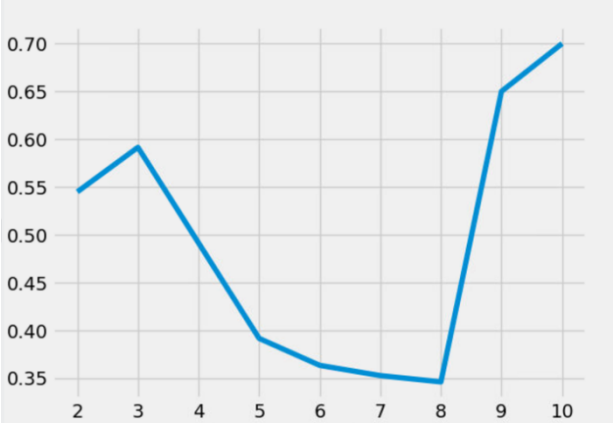


Figure 11. Silhouette scores for k shows that the best k is 10 since it has the maximum score

Ground truth labels categorize data points into group. These types of metrics do their best to suggest the correct number of clusters. To determine the best number of clusters, we fitted our dataset and get the plot.

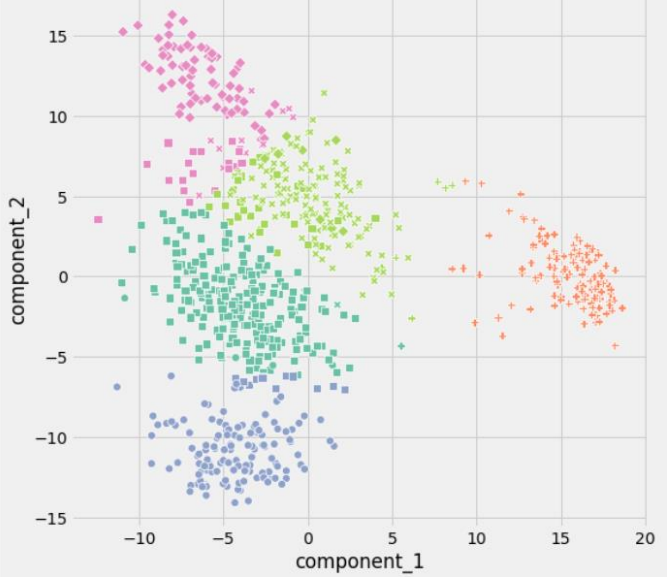


Figure 12. Ground truth labels for clusters 2, 4, 6, 8, 10

Ground truth labels also shows that cluster with 10 (orange points) has the best error value.
The comparison shows that clusters with 8 has the best result to matching the student video and coach's video.

To test the effectiveness of the application, we also create a survey in Google Forms [13]. 100 participants were gathered to take the survey, which is a large enough sample size to mitigate the effects of any variability.

Scores	Numbers
Under 5	2
6	7
7	11
8	10
9	28
10	42

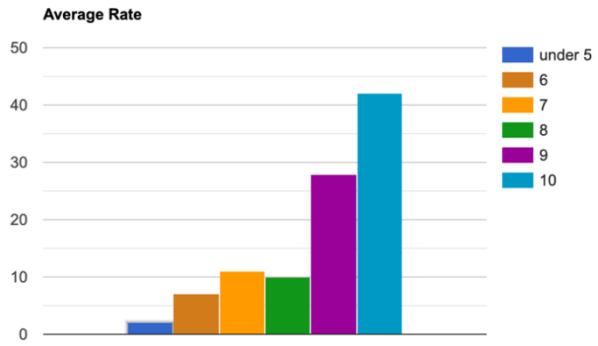


Figure 13. Survey scores for using the app

According to the results of the first experiment, we choose 8 as the best performance K values for matching the video between student and coach. The second experiment shows the app can be a great help to dancers. The participants are regarded as a small sample of the public, which means that the general public can make use of the application.

5. RELATED WORK

Several researchers have worked on the K-means algorithm. In the paper “Unsupervised K-means clustering algorithm.", Sinaga, Kristina P., and Miin-Shen Yang construct an unsupervised learning schema for the k-means algorithm so that it is free of initialization without parameter selection [11]. They also simultaneously find an optimal number of clusters. In the paper, the k-means Algorithm: A Comprehensive Survey and Performance Evaluation, Syed Islam explains the advance of k-means algorithm and discuss the using of comparison among different k-means clustering algorithms.

In the paper "Real-time comparison of movement of Bulgarian folk dances.", Uzunova, Zlatka. explain a method recording the dancing with two Kinetic sensors and analyze the data with machine learning algorithm. To understand how to apply and pick the data, we also check the research paper related to the machine learning filed, In the paper " Machine learning algorithms-a review” Mahesh, Batta published, he explained how different machine learning algorithm applying in data analyzing.

6. CONCLUSION

In this paper, we explain a method for analyzing student ballet movement by Comparing coach's movements and the student movements. We cut the coach's 3 minutes video length into approximately 5400 frames and student’s 30 second video into approximately 900 frames to compare with a K-means algorithm [14]. Our results show that to make the k-means algorithm work best, we need a good dataset and use different methods to find the best K number that fits the algorithm dataset. Both SSE curve and silhouette scores show that the algorithm is stable.

Further improvement: In the future we could collect more data from the user to improve the model performance and processing.

REFERENCES

[1] Andriluka, Mykhaylo, et al. "2d human pose estimation: New benchmark and state of the art analysis." Proceedings of the IEEE Conference on computer Vision and Pattern Recognition. 2014.
[2] Joorabchi, Mona Erfani, Ali Mesbah, and Philippe Kruchten. "Real challenges in mobile app development." 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. IEEE, 2013.
[3] Likas, Aristidis, Nikos Vlassis, and Jakob J. Verbeek. "The global k-means clustering algorithm." Pattern recognition 36.2 (2003): 451-461.
[4] Deng, Li, and Dong Yu. "Deep learning: methods and applications." Foundations and trends® in signal processing 7.3–4 (2014): 197-387.
[5] Sarker, Iqbal H. "Machine learning: Algorithms, real-world applications and research directions." SN Computer Science 2.3 (2021): 1-21.
[6] Novack, Cynthia J. "Ballet, gender and cultural power." Dance, gender and culture. Palgrave Macmillan, London, 1993. 34-48.
[7] O'Shea, Keiron, and Ryan Nash. "An introduction to convolutional neural networks." arXiv preprint arXiv: 1511.08458 (2015).
[8] I. H. Witten, "Data Mining: Practical Machine Learning Tools and Techniques, Third Edition," Morgan Kaufmann Series in Data Management Systems, 2011.
[9] Zhang, Fan, et al. "Mediapipe hands: On-device real-time hand tracking." arXiv preprint arXiv: 2006.10214 (2020).
[10] Yuan, Fang, et al. "A new algorithm to get the initial centroids." Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826). Vol. 2. IEEE, 2004.
[11] Sinaga, Kristina P., and Miin-Shen Yang. "Unsupervised K-means clustering algorithm." IEEE access 8 (2020): 80716-80727.
[12] Uzunova, Zlatka. "Real-time comparision of movement of Bulgarian folk dances." Computer Science and Education in Computer Science 14.1 (2018): 93-108.
[13] Vasantha Raju, N., and N. S. Harinarayana. "Online survey tools: A case study of Google Forms." National Conference on Scientific, Computational & Information Research Trends in Engineering, GSSS-IETW, Mysore. 2016.
[14] Likas, Aristidis, Nikos Vlassis, and Jakob J. Verbeek. "The global k-means clustering algorithm." Pattern recognition 36.2 (2003): 451-461.
[15] Yuan, Fang, et al. "A new algorithm to get the initial centroids." Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826). Vol. 2. IEEE, 2004.

ON THE ENERGY EFFICIENCY OF IEEE 802.11AX HIGH DENSITY WLANS

Zineb Machrouh, Abdellah Najid and Iyad Lahcen-Cherif

Department of Communication Systems, INPT, Rabat, Morocco

ABSTRACT

Wireless communications evolved in a remarkable way during the last decade and is well on its way to surpass wired internet. The demand shifted towards higher transmission speed for more users and heavier traffics. In this paper, we present an IEEE 802.11ax scenario, in which we study the energy efficiency for the key metrics of the MAC layer. In this latest edition, called High Efficiency WLAN (HEW) energy is a main concern in order to satisfy scenarios of internet of things and wireless sensor networks. we prove that some of the new features such as the higher order modulation and coding schemes enhance remarkably the energy efficiency. We also show the impact of an increase in the number of users on the system and prove the payoff of using IEEE 802.11ax. We evaluate the contention window size performance as one of the most important metrics, on which throughput highly depends.

KEYWORDS

802.11ax, Contention Window, Energy Efficiency, Throughput, Wireless Networks.

1. INTRODUCTION

Since appearing in the late 90's, IEEE 802.11 standard has had subsequent amendments. The latest version, IEEE 802.11ax offers a new approach with the focus shifting towards enhancing the overall efficiency instead of the throughput. The aim is to accommodate real time applications, 4K/8K videos, virtual reality, as well as coexist with 4G, 5G, and LTE. Internet of Things (IoT) can also benefit from this amendment and its new features [1], with efficient allocation of low data rate connections over an extended range, which will remarkably improve battery life for IoT sensors.

For these reasons, energy consumption in Wireless Local Area Networks (WLANs) is a key consideration in the design and the implementation of physical layer (PHY) or medium access layer (MAC). In every scenario, it depends highly on the system throughput, the reliability and the achievable lifetime. The stations consume energy during transmission, reception, and in idle states due to collisions, transmission errors, channel sensing, listening, and the use of wider bandwidths. Remedying these energy-consuming problems has been the subject of several studies.

In [2], the authors show the importance of the contention window size. They investigate the energy efficiency in an IEEE 802.11 network with an analytical model and prove that the maximum energy efficiency is a function of optimal packet size and contention window size. The work in [3] uses a configuration for IEEE 802.11, which aims to optimize the energy efficiency. The authors show that although we can have an energy-optimized configuration, it does not necessarily mean that throughput and energy efficiency can have a joint maximization. The

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 147-153, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121715

backoff window effect on the throughput is a very important aspect in saturation conditions. Paper [4] presents an energy consumption model for different numbers of users and its impact on the system throughput.

The number of users affects resource allocation, which in turn influences the throughput and energy efficiency. In [5], a bi-directional distributed coordination protocol improves the energy efficiency remarkably. The authors use an optimal configuration to reduce overhead and channel contention and balance throughput and energy efficiency. A closely related matter; power efficiency is examined in [6]. The authors present a tradeoff between power efficiency and bandwidth efficiency. They prove that a joint maximization is not possible and that only suboptimal maximization is possible. [7], presents a throughput performance for IEEE 802.11ax which is crucial in determining the energy efficiency.

The rest of our paper presents an overview of the IEEE 802.11ax amendments in section 2, then we layout the system model in section 3, we present our results and their interpretations in section 4, and section 5 concludes the paper with summary and perspectives.

This document describes, and is written to conform to, author guidelines for the journals of AIRCC series. It is prepared in Microsoft Word as a .doc document. Although other means of preparation are acceptable, final, camera-ready versions must conform to this layout. Microsoft Word terminology is used where appropriate in this document. Although formatting instructions may often appear daunting, the simplest approach is to use this template and insert headings and text into it as appropriate.

2. OVERVIEW OF IEEE 802.11AX

While IEEE 802.11ax conserves most of the features provided by previous amendments and enhances them, it also adds features specific for highly dense networks. Among the new additions, we mention the following. Enhancements for MU-MIMO include both the Downlink and the Uplink with eight possible streams, while the SU transmission benefit from up to four streams. The 1024 QAM with 3/4 and 5/6 rates, joins the Modulation and Coding Schemes (MCS) as a new addition. The increase in the number of subcarriers to 256 enables longer Guard Interval (GI) and a longer symbol duration. While the GI helps decrease the Inter Symbol Interference (ISI), longer symbol duration allows connectivity in larger coverage areas and makes the system a robust one against propagation delays.

802.11ax introduces the concept of Resource Units (RU), which allows several users the use of the same frequency domain. Thus, providing multiplexing in the frequency domain. A modification in the frame format includes trigger frame, which contains information about the stations involved in a transmission. This information contains a list of users, their RU, resource allocation, bandwidth and MCS. The stations start to transmit only after receiving this frame from the Access Point (AP). A variant of trigger frame is MU Request-To-Send (MU-RTS) frame; it includes information about the stations and the channel widths. Each station replies with a Clear-To-Send (CTS) frame and the AP transmits only to the station that have replied.

For more details about the technical features of IEEE 802.11ax, refer to [8] and [9].

3. SYSTEM MODEL

The Distributed Coordination Function (DCF) The Distributed Coordination Function (DCF) is the default access mechanism for MAC layer in IEEE 802.11 standards. It is similar in principle

to the Carrier Sense Multiple Access mechanism with Collision Avoidance (CSMA/CA). In order to transmit, a station must listen to the medium until it encounters an idle state. When a period equal to Distributed Inter-Frame Space (DIFS) passes, the station can transmit. To avoid collisions, the station waits a random backoff time, which is also used between two consecutive transmissions. If the channel is busy, the station continues listening for a DIFS in order to transmit.

DCF employs either the basic access scheme or the four-way handshaking RTS/CTS. In the latter, the station sends a RTS frame and waits for the CTS before transmitting the packet. RTS/CTS is very useful against the hidden nodes problem. When the transmission is successful, the receiver sends the ACK. During this period, the other stations can update their Network Allocation Vector (NAV). For more details about DCF see [10]. Since IEEE 802.11ax offers MU transmissions in both the Downlink and the Uplink directions, the same can be said about both directions. When the AP is transmitting to several stations or when it is receiving from several stations.

In our paper, we use the well-known Markov Chain Model of [11] to represent the system throughput and the expression of the energy efficiency. The Markov Chain analytical model is very accurate in evaluating thethroughput of both IEEE 802.11ac such in [12] and IEEE 802.11ax.

We note the number of stations as n, in saturation conditions, every station has a packet ready for transmission. The Markov Chain model allows us to deduce the transmission probability (1) that a station transmits in a randomly chosen slot time τ and the collision probability p in (2).

$$\tau = \frac{2(1-2p)}{(1-2p)(CW_{min}+1) + pCW_{min}(1-(2p)^m)} \tag{1}$$

$$p = 1-(1-\tau)^{n-1} \tag{2}$$

To compute the system throughput S we define P_{tr} as the probability that there is a transmission, and P_S the probability that said transmission is successful. Therefore, we express the system throughput as the ratio of successfully transmitted bits in a slot time $E[P]$ against said slot time (5).

$$P_{tr} = 1-(1-\tau)^n \tag{3}$$

$$P_S = \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n} \tag{4}$$

$$S = \frac{P_{tr}P_S E[P]}{(1-P_{tr})\sigma + P_{tr}P_S T_S + (1-P_S)P_{tr}T_C} \tag{5}$$

The expressions σ, T_S, and T_C represent respectively an empty slot time, the time spent for a successful transmission, and the time spent during a collision. The expressions of both T_S and T_C are detailed respectively in (6) and (7). Certain expressions extensions encompass the new modifications added by 802.11ax and allow calculating the throughput [7].

$$T_S = T_{MU-RTS} + 3 * SIFS + CTS + T_{DATA} + BA + AIFS + \sigma \tag{6}$$

$$T_C = T_{MU-RTS} + SIFS + CTS + DIFS + \sigma \tag{7}$$

$$T_{MU-RTS} = PHY + [\frac{L_{SF} + L_{MU-RTS} + L_{TB}}{R}]\sigma \quad (8)$$

$$T_{DATA} = PHY + [\frac{L_{SF} + N_a(L_{MD} + L_{MH} + L_{DATA}) + L_{TB}}{R}]\sigma \quad (9)$$

With the transmission rate R expressed as:

$$R = \frac{N_{DBPS}}{T_{SYM}} \quad (10)$$

N_{DBPS} is the number of bits per OFDM symbol. It can be calculated as in [13] using the number of data subcarriers, the number of information bits in a modulation, and the coding rate. The symbol duration T_{SYM} is equal to the number of subcarriers divided by the used bandwidth plus the Guard Interval (GI).

$$P_{success} = \tau(1-\tau)^{n-1} \quad (11)$$

$$E = (1-\tau)^n p_i + \tau p_t T_S + (1-\tau)(1-(1-\tau)^{n-1}) p_r T_S \quad (12)$$

We can finally express the energy efficiency EE, which is the ratio between transmitted bits and the energy consumed in a slot time as:

$$EE = \frac{P_S E[P]}{E} \quad (13)$$

E is the energy per slot. In Table 1, details the parameters with constant numerical values (see more in [14]).

Table 1. Simulation Values

Parameter	**Value**
L_{SF}	16 bits
L_{TB}	18 bits
L_{MD}	32 bits
L_{MH}	360 bits
L_{DATA}	1500 Bytes
N_a	64
CW_{min}	32
CW_{max}	1024
L_{RTS}	160 bits
L_{CTS}	128 bits
$SIFS$	16 µs
$DIFS$	34 µs
$AIFS$	34 µs
m	5
σ	16 µs
PHY	164 µs
p_e	1.15 W
p_t	1.65 W
p_r	1.5 W

4. RESULTS AND DISCUSSION

We consider a Basic Service Set (BSS) with IEEE 802.11ax capable devices, in which we have a single AP and n stations. All stations are within hearing distance from the AP, as well as from each other. These stations possess the same capabilities and are able to operate in MU-MIMO for both transmission and reception (Uplink and Downlink). In addition, we assume that every station has a packet ready for transmission every time. We consider all the bandwidths available for IEEE 802.11ax amendment, and the full capabilities of spatial streams and diversity. N_a packets are transmitted using Aggregated MAC Protocol Data Unit (A-MPDU) packet aggregation scheme. We present the energy efficiency for different number of users, varying contention window sizes and different MCS.

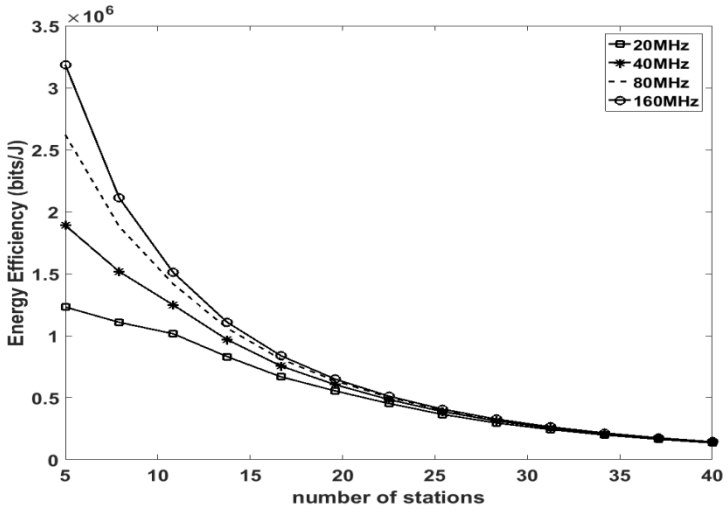


Figure 1. Energy Efficiency against the number of STAa

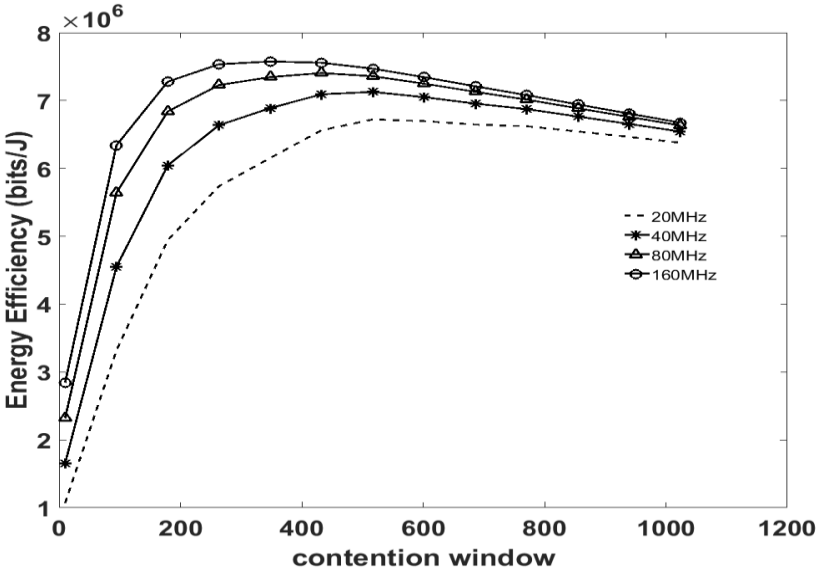


Figure 2. Energy Efficiency against Contention Window

We start with the energy efficiency against the number of stations depicted in figure 1, the energy efficiency obviously decreases as the number of users increase. This is due to the energy

consumed during retransmissions and overhearing besides the collisions, which present the most wasteful condition for energy.

For another important metric, the contention window size in figure 2, for small values of CW a lot of energy is wasted in collisions. Choosing an optimal minimal CW can enhance the energy efficiency regardless of the channel bandwidth. This is mainly because CW affects the collisions in the system. Finding the best value for every case scenario can remarkably minimize contentions and provide a better energy efficiency. There is a CW value, which maximizes EE devoted to successful transmission.

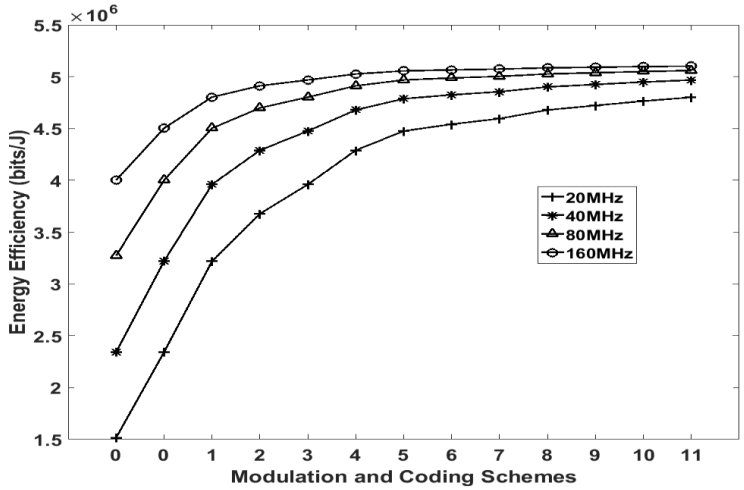


Figure 3. Energy Efficiency for MCS

Finally, as to a very important metric in the IEEE 802.11ax, the modulation and coding schemes. With the newly added 1024QAM or MCS11, the energy efficiency provides remarkable results. Despite reaching saturation faster the higher the MCS gets, we found that low MCS offer low energy efficiency, this is due to the fact that the throughput is lower and the transmission is longer (see figure 3). We obtain the maximum EE, by combining both the optimal packet size and CW optimal value.

5. SUMMARY

In this paper, we presented an overview of the new amendment IEEE 802.11ax; we showed in our analysis the performance of the energy efficiency of the norm as a function of several key parameters in every WLAN with a PHY and MAC layer based on IEEE 802.11ax. We compared the results for various bandwidths and proved the impact of the CW size on the optimization of the energy efficiency. We also proved that the newly added features of aggregation and throughput remarkably improve the performance for higher MCS schemes. Finally, we show the impact of a high number of users (i.e. dense networks) on the energy efficiency.

REFERENCES

[1] Boris Bellalta (2016) "IEEE 802.11ax: High-efficiency WLANs", *IEEE Wireless Communications*, Vol. 23, No. 1, pp38-46.

[2] Xiaodong Wand & Jun Yin & Dharma Agrawal (2006) “Analysis and Optimization of the Energy Efficiency in the 802.11 DCF”, *Mobile Networks and Applications* Vol. 11, No. 2, pp279-286.

[3] Pablo Serrano & Matthias Hollick & Albert Banchs (2010) "On the Trade Off between Throughput Maximization and Energy Consumption Minimization in IEEE 802.11 WLANs", *Journal of Communications and Networks*, Vol. 12, No. 2, pp150-157.
[4] Eleni-Constantina Davri & Emmanouil Kafetzakis & Kimon Kontovasilis & Charlabos Skianis (2014) "An Accurate Model for Energy Efficiency in IEEE 802.11 WLANs", *Proceedings CAMAD, Athens, Greece*, pp385-389.
[5] Palacios Raúl et al. (2013) "Energy efficiency of an enhanced DCF access method using bidirectional communications for infrastructure-based IEEE 802.11 WLANs." *IEEE 18th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks*, pp38-42.
[6] Zhao Liqiang & Yang Kuanli & Zhao Guogang. (2016) "Energy Efficient Fundamental Theory and Technical Approach", *Energy Management in Wireless Cellular and ad-hoc Networks*, Springer Publishing, pp189-213.
[7] Zineb Machrouh & Abdellah Najid (2018) "High Efficiency WLANs IEEE 802.11ax Performance Evaluation", *Proceedings ICCAD, Marrakech, Morocco.*
[8] Cisco (2018) "802.11ax: The Sixth Generation of Wi-Fi", *Technical White Paper.*
[9] Aruba (2018) "802.11ax", *Technical White Paper.*
[10] Ilenia Tinnirello & Giuseppe Bianchi & Yang Xiao (2010) "Refinements on IEEE 802.11 Distributed Coordination Function Modeling Approaches", *IEEE Transactions on Vehicular Technology*, Vol.59, No.3, pp1055-1067.
[11] Giuseppe Bianchi (2000) "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", *IEEE journal on selected areas in communications*, Vol. 18, No. 3, pp535-547.
[12] Zineb Machrouh & Abdellah Najid (2016) "Performance Analysis of 802.11ac DCF Enhancement for VHT with Frame Aggregation", *iJES*, Vol. 4, No. 3, pp17-21.
[13] Ruizhi Liao & al. (2013) "Performance analysis of IEEE 802.11ac wireless backhaul networks in saturated conditions", *Journal of Wireless Communications and Networking*, Vol. 1, No. 1
[14] IEEE (2016) Proposed TGax Draft Specification

AUTHORS

Zineb Machrouh has Master degree in Electronics and Telecommunications and is currently preparing her Ph.D. thesis about Wireless Networks at the National Institute of Posts and Telecommunications in Morocco.

Abdellah Najid receivedthe M.Sc. degree in networking and communication systems and the Ph.D. degree in electronic engineering from ENSEEIHT, Toulouse, France. He has several years of research experience with ENSEEIHT, ENSTA, INRIA, and ALTEN. He joined the National Institute of Posts and Telecommunications as a full professor. He is the chief of Communications Systems department and Cybersecurity Master's Coordinator.

Iyad Lahcen Cherif received his Ph.D. in computer sciences from the university of Paris-Saclay and dedicated his research to improving wireless communication systems with artificial intelligence, he joined the National Institute of Posts and Communication in Rabat as an assistant professor.

RECONFIGURABLE GAMIFICATION PLATFORM FOR THE AUTONOMOUS LEARNING OF LOW VALUE MEDICAL PRACTICES

César Fernández[1], María Asunción Vicente[1], Susana Lorenzo[2], Irene Carrillo[1] and Mercedes Guilabert[1]

[1]Department of Health Psychology, Miguel Hernández University, Elche, Spain
[2]Hospital Universitario Fundación Alcorcón, Madrid, Spain

ABSTRACT

Failure to follow do-not-do recommendations (also known as low-value practices) is one of the causes of the lack of quality care in all health systems in all countries. Healthcare professionals must be provided with information about these low-value practices that are still frequently performed and their implications for patients and the healthcare system.

Continuous education is a key factor in this scenario, so that health students, health professionals, and even patients are kept updated with the main do-not-do recommendations. Gamified platforms are one of the most valuable options for continuous education, as they combine learning efficiency with a high level of engagement for the students. Besides, the effectiveness of gamification platforms can be improved by adding artificial intelligence techniques.

In this paper, a novel gamified platform focused on improving knowledge about low-value practices is proposed. AI techniques, as well as NLP tools are used to optimize the effectiveness of learning by adapting the platform to each user, at an individual level. Besides, the engagement of students is encouraged by their participation in a common project, namely the creation of a specialized dictionary for do-not-do terms.

Hardware development is currently in progress. A basic gamification platform has already been developed for the two main mobile operating systems. Developing IA and NLP techniques to analyse the training outputs and make the platform adaptable to each student is progressing.

The proposed learning tool can significantly improve healthcare quality and be applied to many other learning fields, particularly when continuous training is required.

KEYWORDS

Low-value practices, Do-not-do, Gamification, Artificial Intelligence, Natural Language Processing, Transformers.

1. INTRODUCTION

Failure to follow do-not-do recommendations (also known as low-value practices, overuse, overdiagnosis, overtreatment, non-recommended practices) is one of the causes of the lack of quality care in all health systems in all countries. It concerns the provision of health services in

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 155-163, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121716

circumstances where the potential risk of harming the patient exceeds the potential benefits [1], representing a risk both for patient safety [2] and for the sustainability of health systems [3].

The volume of patients subjected to low-value practices varies, depending on the type of practice and the country, between 1 and 80% [4]. The use of potentially unsuitable medicines might come as high as 57.6% [5] in some studies. In terms of costs, in the USA, where 18% of gross domestic product is spent on the health sector, overuse represents an additional annual cost which, according to the most optimistic estimates [6], ranges from $75.7 to $101.2 billion, whereas other studies put the figure at $158-226 billion [7]. In terms of damages to patients, the study presented in [8] estimates that between 0.2% and 15.0% of hospitalized patients admitted suffering adverse events because of low-value practices. In primary care, a study in Spain [9] estimated that ignoring the do-not-do recommendations resulted in 5.1% of adult patients suffering adverse events

Healthcare professionals must be provided with information about these low-value practices that are still frequently performed and their implications for patients and the healthcare system.

Continuous education is a key factor in this scenario, so that health students, health professionals, and even patients are kept updated with the main do-not-do recommendations. Previous research [10] concluded that gamified platforms are one of the most valuable options for continuous education, as they combine learning efficiency with a high level of engagement for the students. Besides, the effectiveness of gamification platforms can be improved by adding artificial intelligence techniques, as it is also stated in previous research [11].

The term gamification refers to the use of game attributes (mainly challenges and rewards) in a non-gaming context, particularly in education (Game-Based Learning or GBL). When focused on education, GBL techniques make use of games as assistant tools for learning, concept assimilation or knowledge evaluation. One of the main advantages of GBL is its ability to capture the user's attention, since it provides them with a fun and motivating environment. The game presents the users with situations in which they must reflect and make the right decisions, solve errors, and recognize mistakes. Attributes commonly found in gamification include scoring systems, win/lose challenges, rewards, leader boards, and social elements.

Different gamification strategies have been previously used in healthcare education. The effectiveness of these techniques has been widely analysed in multiple studies and reviews. Among the more recent reviews, those presented in [12] and [13] agree on the main conclusion: although rigorous studies are scarce, some of them report improvements, with respect to control groups, in different aspects like knowledge, skills, satisfaction, changing learning behaviours and improving attitudes towards learning.

There are multiple theories that can explain the reasons behind the success of gamification platforms. Among them, the called Social Comparison Theory, which is related to the introduction of scoring systems and leader boards [14]; the Experiential Learning Theory, which states that being part of a game increases the feeling of experiencing new situations and improves learning [15]; the Reinforcement Learning Theory, which explains that rewards and punishments (common in computer Reinforcement Learning algorithms and also in games) allow students to obtain more skills [16]; or the Self-Directed Learning, particularly for gamification platforms where students can choose among different games or tests [17].

In all cases, part the benefits of gamification are also related to the additional development of a social network for learners [18], where leader boards and scoring systems are complemented with different ways of interaction between participants.

An additional tool commonly used in leaning environments is Natural Language Processing (NLP), which includes multiples techniques for the automatic processing of huge amounts of textual information. Among the different applications of NLP, we will focus on information extraction, sentiment analysis, and automatic translation. NLP has evolved considerably in the last few years, mostly thanks to the development of Deep Learning [19]. More recently, the concept of "attention" was introduced [20] and a new architecture emerged: the transformer. Transformers are particularly well suited for translation tasks, where they clearly outperform all previous approaches [21]. Among the main developments in this field, we can cite different open-sourced NLP platforms, like BERT [22], ROBERTa [23] and BETO [24]. The availability of open-sourced code for these platforms allows for an easy integration and extension of capabilities.

It must also be considered that the combination of gamification and NLP is extremely beneficial [25]. Basically, when NLP is complemented with gamification techniques, data gathering becomes easier and all kinds of information extraction are improved. In fact, some gamification platforms are designed with the only purpose of acquiring data for further NLP processing [26].

In this paper, a novel gamified platform focused on improving knowledge about low value practices is proposed. Artificial Intelligence (AI) techniques, as well as NLP tools are used to optimize the effectiveness of learning by adapting the platform to each user, at an individual level. Besides, the engagement of students is encouraged by their participation in a common project, namely the creation of a specialized dictionary for do-not-do terms.
The paper is structured as follows: First, the methodology used for content generation is described in section 2. Sections 3 and 4 present the gamified learning platform developed and a proposal for improving student engagement, respectively. Software development is detailed in section 5; and section 6 analyses the data that can be gathered during platform usage. Finally, some conclusions are summarized in section 7.

2. CONTENT GENERATION

Content selection was performed through searches for do-not-do recommendations in different scientific societies and initiatives such as "choosing wisely". The following sources of information were used:

- American Academy of Orthopaedic Surgeons
- American Academy of Ophthalmology
- American College of Surgeons
- American Society for Transplantation and Cellular Therapy
- Cell Therapy Transplant Canada
- American Society of Haematology
- American Society of Paediatric Haematology Oncology
- American Society of Haematology
- American Society of Nuclear Cardiology
- Appropriateness Evaluation Protocol
- Choosing Wisely initiative

Besides, at a national level, 24 additional Spanish scientific associations were also searched for extra recommendations.

Next, recommendations were converted to teaching material for the gamification platform. The material as questions of different complexity: from short sentences that must be marked as true or

false to practical cases which will need a reasoned answer. In all cases, a reward for the correct answers is established. A total of 186 questions or challenges were created.

To improve the learning efficiency, each question was complemented with a detailed explanation for the students, which was shown after they answered, either correctly or incorrectly.

3. RECONFIGURABLE GAMIFIED PLATFORM

The basic behaviour of a gamified learning system is depicted in figure 1, where the three main components (learning system, gamification elements or rewards, and social network of students) are represented. Basically, by carrying out the tasks, assignments, tests, or games, the students are rewarded with different rewards.

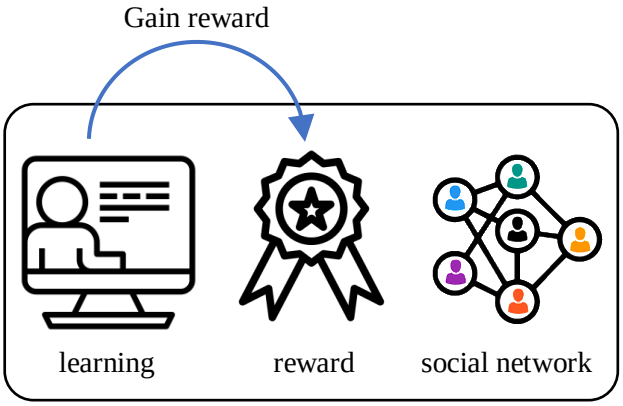


Figure 1. Basic behaviour of a gamified learning platform

We propose to add a continuous evaluation of 1) effectiveness and 2) engagement for each student. This evaluation will be based on all information gathered during training: first, learning results but, also, information extracted via natural language processing (NLP) form the participation of the student in the ad-hoc created social network, which may even include sentiment analysis. All data will be combined by means of artificial intelligence techniques (AI) to infer the best adaptation of the gamification scheme to offer optimized learning performance at the individual level. The proposal is represented in figure 2.

The items that can be adapted to each student are:

- Question level adjustment: should be adapted depending on the previous results of each user.
- Question types: some users may prefer challenges, other may prefer tests, etc. The same concept can be acquired using different question types.
- Level of gamification: some of the game-related elements (medals, trophies, sounds, animations) may be excessive for certain users.

Basically, two types of information are fed to the A.I. algorithms: evolution of the platform evaluation by users (which directly correlates to the suitability of the adjustments for those users); and textual information extracted from the interactions between users, which is analysed through NLP to indirectly detect levels of satisfaction. These NLP techniques include sentiment analysis, among other searches for keywords representing the acceptability of the platform.

A closed-loop control system continuously monitors the information received and adjusts the platform. This process is transparent for the users.

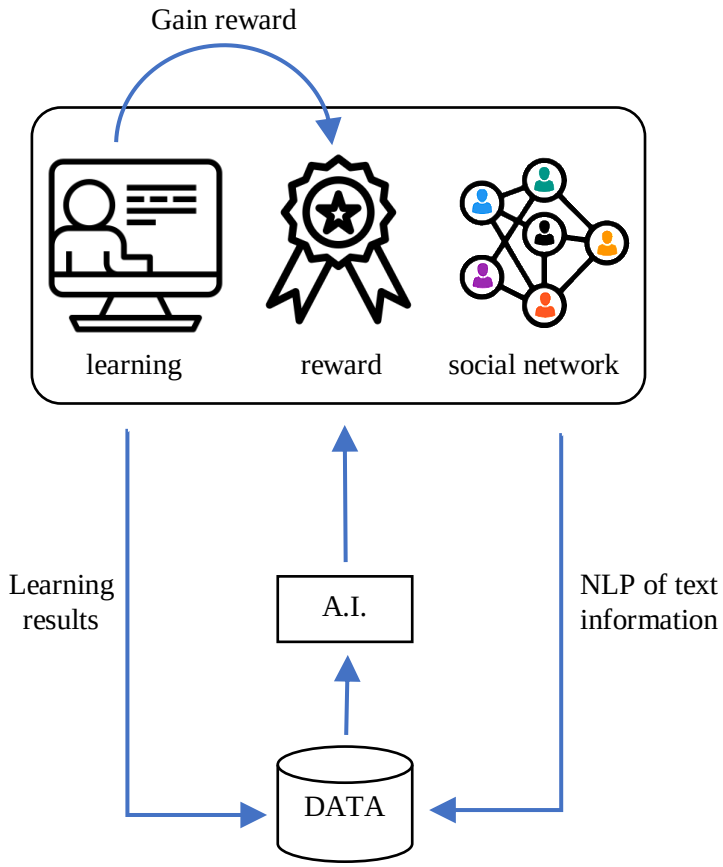


Figure 2. Our proposal of reconfigurable learning platform

4. IMPROVING ENGAGEMENT

To obtain more value from the system, and to increase the engagement of the students, all participants become part of a collaborative project whose objective is to create a dictionary specialized in "do-not-do" terms. Assuming that students from different countries may use the learning tool, each participant will be encouraged to add new items (new translations) to the dictionary (from English to their mother tongues), as well as to evaluate the other participants' translations. Both actions will be rewarded in the gamified platform in order to engage students. The dictionary will be used to train an automatic translation system via NLP and transformers. Besides being a means to engage students, this automatic translation system will finally become an extra source of information for all students, thus creating a learning system that evolves and is partially self-maintained. A schematic representation can be found in figure 4.

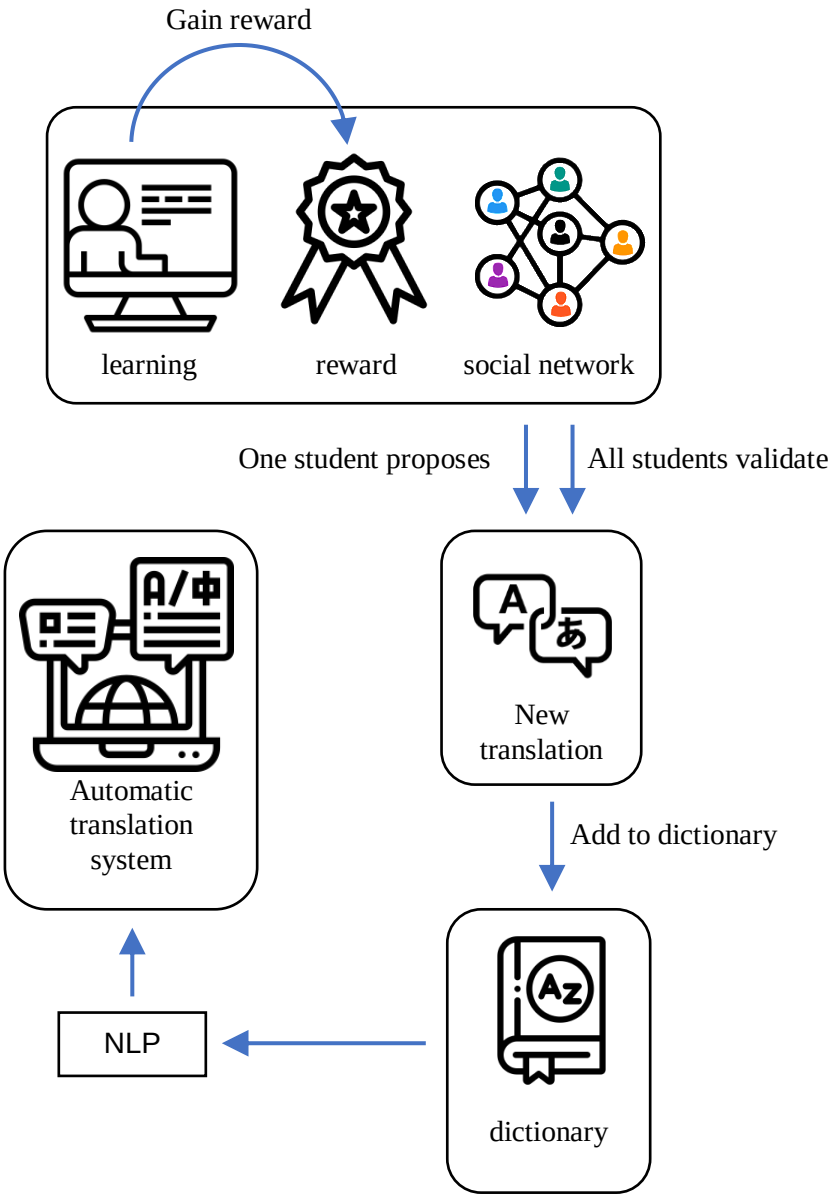


Figure 3. Collaborative creation of a specialized automatic translation system via NLP and transformers

5. SOFTWARE DEVELOPMENT

Software development is currently in progress. A basic gamification platform has already been developed for the two main mobile operating systems and can be downloaded from the store of Apple (for iOS devices) and Google (for Android devices). So far, they have been developed as native applications, which implies a specific development for each operating system. Future work includes a reengineering of the platform in order to develop a hybrid application that can run on both operating systems (Android and iOS), as well as in web browsers, in order to allow students to use it from their mobile devices and also from their computers. The IONIC platform [27] can be used for this purpose.

The development of IA and NLP techniques to analyse the training outputs and make the platform adaptable to each student are still under development. The automatic translation system (fed from the collaborative-created dictionaries) is based on open-sourced code available as a Python API [28].

6. DATA TO GATHER

The data to be collected during the use of the platform includes:

- Demographic, but not identification-capable, data: gender, age, medical specialty, and years of professional experience (the latter two, only for professionals).
- Game usage: questions answered, results, awards.
- Social network interactions, in order to apply NLP techniques to extract information, particularly sentiment analysis.
- Contributions to the collaborative dictionary.

Provided that the registered data will not include any identification-capable information, all these data available for further research, following the recommendations of the MareData network [29] and fulfil the FAIR Data Action Plan, as proposed by the European Commission [30] In brief, the data will be made available to researchers fulfilling Findability, Accessibility, Interoperability and Reusability.

7. CONCLUSIONS

The training of medical students and professionals in the field of low-value practices can have an important impact on the improvement of healthcare quality, as well as in the reduction of adverse effects and the reduction of health care costs, by avoiding overuse of medicines and medical tests. Engagement of students is crucial for increasing the effectiveness of training, particularly when a continuous training is required. The proposed learning platform is focused on reaching high engagement levels, since 1) it is based on a gamified environment, which per-se enhances engagement; 2) includes an adaptative behaviour to adapt to each student at the individual level; and 3) adds a collaborative project where all students are encouraged to participate to obtain additional rewards.

The proposed learning tool, which is centred on training on low-value medical practices, can be applied to many other learning fields, particularly when continuous training is required.

REFERENCES

[1] Chassin MR, Galvin RW. The Urgent Need to Improve Health Care Quality. JAMA.1998;280:1000-5.
[2] Lipitz-Snyderman A, Korenstein D. Reducing Overuse-Is Patient Safety the Answer? JAMA. 2017;317:810-1.
[3] Korenstein D, Chimonas S, Barrow B, et al. Development of a Conceptual Map of Negative Consequences for Patients of Overuse of Medical Tests and Treatments. JAMA Intern Med.2018;178:1401-07.
[4] Brownlee S, Chalkidou K, Doust J, et al. Evidence for Overuse of Medical Services Around the World. Lancet. 2017;390:157-68.
[5] Alhawassi TM, Alatawi W, Alwhaibi M. Prevalence of potentially inappropriate medications use among older adults and risk factors. BMC Geriats. 2019; 19:154.
[6] Shrank WH, Rogstad TL, Parekh N. Waste in the US Health Care System: Estimated Costs and Potential for Savings. JAMA. 2019 [Online ahead of print] doi: 10.1001/jama.2019.13978

[7] Berwick D, Hackbarth A. Eliminating Waste in US Health Care. JAMA. 2012;307:1513-6.
[8] Badgery-Parker T, Pearson SA, Dunn S, et al. Measuring Hospital-Acquired Complications Associated With Low-Value Care. JAMA Intern Med. 2019;179:499-505.
[9] Mira JJ, Carrillo I, Silvestre C, Pérez-Pérez P, Nebot C, Olivera G, González de Dios J, Aranaz Andrés JM. Drivers and strategies for avoiding overuse. A cross-sectional study to explore the experience of Spanish primary care providers handling uncertainty and patients' requests. BMJ Open. 2018;8(6):e021339. doi: 10.1136/bmjopen-2017-021339. FI (JCR, 2018): 2,376 - Q2 Medicine, General & Internal - SCIE.
[10] Fernández C, Lorenzo S, Vicente MA, Checa ML. Estudio comparativo de metodologías de enseñanza online para la reducción de las prácticas de poco valor. Actas del VII Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud. Volumen I, 2021.
[11] Fernández C. Lorenzo S, Carrillo I, Oveja R. Utilización de técnicas de inteligencia artificial en una plataforma digital de enseñanza de las recomendaciones nohacer. Actas del VII Congreso Internacional en Contextos Psicológicos, Educativos y de la Salud. Volumen I, 2021.
[12] van Gaalen, A.E.J., Brouwer, J., Schönrock-Adema, J. et al. Gamification of health professions education: a systematic review. Adv in Health Sci Educ 26, 683–711 (2021). https://doi.org/10.1007/s10459-020-10000-3
[13] Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, Dauwels-Okutsu S, Nikolaou CK, Zary N, Campbell J, Car J. Serious Gaming and Gamification Education in Health Professions: Systematic Review. J Med Internet Res 2019;21(3):e12994
[14] Huschens, Martin & Rothlauf, Franz & Rothe, Ricarda. On the Role of Social Comparison Processes in Gamified Work Situations. Hawaii International Conference on System Sciences, 2019. 10.24251/HICSS.2019.176.
[15] Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. Experiential learning theory: Previous research and new directions. Perspectives on Cognitive, Learning, and Thinking Styles, 2001, 1, 227–247.
[16] Chen, P. H., Roth, H., Galperin-Aizenberg, M., Ruutiainen, A. T., Gefter, W., & Cook, T. S. Improving abnormality detection on chest radiography using game-like reinforcement mechanics. Academic Radiology, 2017, 24(11), 1428–1435.
[17] Salinas Navarro, David Ernesto, Pacheco-Velazquez, Ernesto, Arana-Solares Iván, Mendoza, Jaime. (2020). Gamification and Self-Directed Learning. Proceedings of the European Conference on Games-based Learning, 2020. 10.34190/GBL.20.159.
[18] Amy K Saenger, Michael Berkwits, Simon Carley, Shannon Haymond, Marie Ennis-O'Connor, Jonathan Sherbino, Stephen W Smith, The Power of Social Media in Medicine and Medical Education: Opportunities, Risks, and Rewards, Clinical Chemistry, Volume 64, Issue 9, 1, 2018, Pages 1284–1290, https://doi.org/10.1373/clinchem.2018.288225
[19] Y. Bengio, A. Courville and P. Vincent, Representation Learning: A Review and New Perspectives, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798-1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.
[20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is All You Need, 2017, NIPS, arxiv.org/pdf/1706.03762.pdf
[21] S. Karita et al., A Comparative Study on Transformer vs RNN in Speech Applications, 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2019, pp. 449-456, doi: 10.1109/ASRU46091.2019.9003750.
[22] J. Devlin, M.W. Chang, Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing. Google AI Blog, November 2, 2018
[23] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, RoBERTa: A Robustly Optimized BERT Pretraining Approach. Preprint, 2019.
[24] J. Cañete, G. Chaperon, R. Fuentes, J.H. Ho, H. Kang, J. Pérez, Spanish pre-trained BERT model and evaluation data, International Conference on Learning Representations, 2020.
[25] J. Chamberlain, R. Bartle, U. Kruschwitz, C. Madge, M. Poesio, Metrics of games-with-a-purpose for NLP applications, 15th European Chapter of the Association for Computational Linguistics (EACL), 2017.
[26] A. Scharl, M. Fols, uComp Language Quiz - A Game with a Purpose for Multilingual Language Resource Acquisition, 15th European Chapter of the Association for Computational Linguistics (EACL), 2017.
[27] IONIC, cross-platform app development tool. https://ionicframework.com/
[28] Hugging Face Transformer API. https://huggingface.co/docs/transformers

[29] Maredata, thematic network of open research data. https://maredata.net/
[30] Final Report and Action Plan from the European Commission Expert Group on FAIR Data. https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf

AUTHORS

César Fernández is Associate Professor at Miguel Hernandez University, Telematics Engineering area, since 1997. He obtained his Industrial Engineering degree from Polytechnical University of Madrid (Spain) and his PhD degree from Miguel Hernandez University in Elche. His main research topics are machine learning, computer vision and telemedicine. https://orcid.org/0000-0002-9391-9192

María Asunción Vicente received her MS degree in Electronics Engineering from University of Valencia in 1999. In 2007, she obtained her PhD degree (extraordinary doctorate award) from Miguel Hernandez University. Associate Professor at Miguel Hernandez University, Telematics Engineering area. She has carried research on innovation in engineering education, computational neuroscience, and computer vision. Currently, her research is focused on mobile health. https://orcid.org/0000-0002-8630-7251

Susana Lorenzo is PhD in Medicine by the Complutense University of Madrid and has a master's degree in public health policy administration by the University of Michigan. She is head of the Quality and Patient Management Area of the Fundación Alcorcón University Hospital. Director of the Journal of Healthcare Quality Research and member of editorial committees of other national and international journals. She has been president of the Spanish Society for Healthcare Quality. She is author of more than a hundred publications in books and national and international scientific journals. https://orcid.org/0000-0003-0954-7830

Irene Carrillo is Assistant Professor of the Department of Health Psychology at the Miguel Hernández University of Elche (UMH). Degree in Psychology (2013), Master's Degree in Health Psychology (2014), Master's Degree in General Health Psychology (2016) and PhD (2018) by the UMH. Main lines of research: quality of care, patient safety, second victim phenomenon, technological solutions in health and recommendations not to do. https://orcid.org/0000-0002-6981-7284

Mercedes Guilabert is Professor in the Department of Health Psychology of Miguel Hernández University (UMH). Degree in Psychology from Miguel Hernández University (2005). PhD in Psychology from Miguel Hernández University PhD program (2011). Master's degree in Patient Safety and Healthcare Quality (2014). Researcher in several projects, earning competitive research contracts financed by national and international agencies. Specialist in qualitative research techniques. Co-author of national and international publications. Invited speaker in national and international conferences and writer of technical reports for both public and private institutions. https://orcid.org/0000-0002-0706-9911

USING SINGULAR VALUE DECOMPOSITION IN A CONVOLUTIONAL NEURAL NETWORK TO IMPROVE BRAIN TUMOR SEGMENTATION ACCURACY

Pegah Ahadian, Maryam Babaei and Kourosh Parand

Department of Data and Computer Sciences,
Shahid Beheshti University, Tehran, Iran

ABSTRACT

A brain tumor consists of cells showing abnormal brain growth. The area of the brain tumor significantly affects choosing the type of treatment and following the course of the disease during the treatment. At the same time, pictures of Brain MRIs are accompanied by noise. Eliminating existing noises can significantly impact the better segmentation and diagnosis of brain tumors. In this work, we have tried using the analysis of eigenvalues. We have used the MSVD algorithm, reducing the image noise and then using the deep neural network to segment the tumor in the images. The proposed method's accuracy was increased by 2.4% compared to using the original images. With Using the MSVD method, convergence speed has also increased, showing the proposed method's effectiveness.

KEYWORDS

Artificial Neural Network, SV, Brain tumor, Brain MR, Segmentation.

1. INTRODUCTION

A brain tumor is an abnormal mass in the brain that can be cancerous (malignant) or non-cancerous (benign). Such as the type, location, size, and manner of tumor expansion that affect the degree of the threat of a tumor. Diagnosis of brain tumors due to size, form, and appearance in the brain is a complex task. Today, with the advancement of image segmentation algorithms, it is possible to diagnose a brain tumor. It is also provided automatically. These methods make it possible to store the state of tumor growth during treatment and reduce errors in human actions. However, due to the diversity of tumors, achieving high accuracy in tumor diagnosis is challenging. From diagnosis methods, a brain tumor can be detected by color histogram method, MRI image segmentation methods, and adaptive neuro-fuzzy inference system method. The artificial neural network was one of the available methods for detecting brain tumors in MRI images. The convolutional neural network is one of the deep methods of machine learning which is used due to its high efficiency. It has become prevalent to extract features and categorize images, which has a considerable image clarification impact. This way, the algorithm performance is improved by targeting the noises in the image [1]. The most critical removal methods are the noise of the images can be referred to as average filters, Wiener filters, non-local averages, and singular values decomposition. MSVD, which is obtained using the singular value analysis method, in addition to removing the noises of images, also preserves the quality of the images. This paper has tried firstly by choosing the appropriate number of individual values in the analysis of individual values of brain MRI images, to reduce the volume of existing noises in

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 165-171, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121717

the images, and then by using a convolutional neural network method to detect brain tumors in the images.

2. METHOD

2.1. Singular Value Decomposition

The decomposition of singular values plays a crucial role in various fields of numerical linear algebra. Due to the properties and efficiency of this decomposition, many modern algorithms and methods are based on singular value decomposition. Singular Value Decomposition of an [M x N] matrix A is obtained from the matrix equation:

$$A = \mathrm{USV}^{\mathrm{T}}$$

Where U is an orthonormal matrix of size [M x M] and V is an orthonormal matrix of size [N x N], and V^{T} is transpose of V. S is a diagonal matrix (of singular values) of size [M x N] with positive values on the [M x M] in the following form:

$$\mathrm{S} = \begin{bmatrix} \partial_1 & & \\ & \ddots & \\ & & \partial_n \end{bmatrix}$$

and

$$\partial_1 \geq \partial_2 \geq \partial_3 \geq \cdots \geq \partial_n$$

The elements in S are square roots of eigenvalues of $\mathrm{A}^{\mathrm{T}}\mathrm{A}$. Moreover, the diagonal elements in S are sorted in descending order [2]. The SVD bases have frequency interpretations for images. As S(1,1) has the greatest value, its significance in describing image features is the highest. Values of the diagonal elements of S decrease and so also their significance in describing the image features.

One of the applications of this analysis is to use this method in image processing for Image compression is the removal of noise and scratches in a photo, etc. [3]. Usually, matrices corresponding to images have a large number of values that are small singulars. Suppose there is (n − k) a small singular value of matrix A that can be neglected. Then the matrix $\partial_1 u_1 \vartheta^T{}_1 + \partial_2 u_2 \vartheta^T{}_2 + \ldots + \partial_k u_k \vartheta^T{}_k$ is an excellent approximation of A, and such an approximation can be sufficient in applications [4]. Fig. 1 shows the K-SVD method's effect on MRI images by choosing different K values.

As part of this study, we used a multiresolution SVD for image denoising. Multiresolution singular value decomposition (MSVD) method is similar to the multiresolution analysis (MRA) employed in the wavelet domain [5]. The noisy image is decomposed into four sub-bands of varying frequency in the first level of MSVD decomposition. The low-frequency sub-band represents most of the signal. As noise is generally of high frequency, it is present primarily in the detailed highfrequency sub-bands. The present noise in detailed sub-bands is removed using suitable thresholding techniques. Synthesis of the approximation sub band and threshold detailed sub-band is done using inverse multiresolution SVD (IMSVD) to get the denoised image. MSVD algorithm is given in the following semi-code:

Input Image A

$$A1 = Reshape(A.4.16384)$$
$$[U\ S\ V] = svd(A1)$$
$$T = U^T\ A1$$
$$Y.LL = Reshape(T(1.:).128.128)$$
$$Y.LH = Reshape(T(2.:).128.128)$$
$$Y.HL = Reshape(T(3.:).128.128)$$
$$Y.HH = Reshape(T(4.:).128.128)$$
$$T1(1.:) = Reshape(\ (Y.LL).1.16384)$$
$$T1(2.:) = Reshape\ ((Y.LH).1.16384)$$
$$T1(3.:) = Reshape(\ (Y.HL).1.16384)$$
$$T1(4.:) = Reshape\ ((Y.HH).1.16384)$$
$$A2 = UT1\ A2\ is\ (4.16384)$$
$$A = Reshape(A2\ .256.256)$$
Output A

Algorithm 1. MSVD algorithm

The Y.LL band contains low-frequency details. The other three contain high-frequency (horizontal, vertical, and diagonal) details.

2.2. Convolutional Neural Network

Among machine learning methods, an artificial neural network is a popular method that simulates the neural network mechanisms of biological organisms. It consists of a network of nodes called neurons and weighted edges. In this system, the inputs flow in the network, and a series of outputs are produced. Then the outputs are compared with valid data, and the paths that lead to the detection of correct cases will be strengthened and the network paths with the wrong detection corrected. The network can finally achieve perfect accuracy and the desired task by repeating this process. One of the essential steps in the development of artificial neural networks is the design of the network architecture, which greatly affects network performance. In designing the architecture of artificial neural networks, factors such as the number of hidden layers, the number of neurons in each layer, transformation functions, and training algorithms must be determined.

One of the common architectures is called a convolutional neural network (CNN). Convolutional neural networks are a class of deep neural networks that are usually used to perform image or speech analysis in machine learning. The fundamental motivation for these networks is that the recognition function of the visual part of the cat's brain was found, which seemed to stimulate specific neurons in certain parts of the visual field. This rule was more widely used to design the architecture of these networks. In convolutional neural network architecture, each layer of the network is 3-dimensional, with a spatial range and depth corresponding to the number of features. Historically convolutional neural networks are the most successful type of neural network. These networks are widely used for image recognition, object localization, and even text processing, and the performance of these networks has recently surpassed humans in the problem of image classification.

Among machine learning methods, an artificial neural network is a popular method that simulates the neural network mechanisms of biological organisms. It consists of a network of nodes called neurons and weighted edges. In this system, the inputs flow in the network, and a series of outputs are produced. Then the outputs are compared with valid data, and the paths that lead to

the detection of correct cases will be strengthened and the network paths with the wrong detection corrected. The network can finally achieve perfect accuracy and the desired task by repeating this process. One of the essential steps in the development of artificial neural networks is the design of the network architecture, which greatly affects network performance. In designing the architecture of artificial neural networks, factors such as the number of hidden layers, the number of neurons in each layer, transformation functions, and training algorithms must be determined.

One of the common architectures is called a convolutional neural network (CNN). Convolutional neural networks are a class of deep neural networks that are usually used to perform image or speech analysis in machine learning. The fundamental motivation for these networks is that the recognition function of the visual part of the cat's brain was found, which seemed to stimulate specific neurons in certain parts of the visual field. This rule was more widely used to design the architecture of these networks. In convolutional neural network architecture, each layer of the network is 3-dimensional, with a spatial range and depth corresponding to the number of features. Historically convolutional neural networks are the most successful type of neural network. These networks are widely used for image recognition, object localization, and even text processing, and the performance of these networks has recently surpassed humans in the problem of image classification.

2.3. U-Net in Tumor Segmentation

In this work, we first tried to reduce the noises in the images by using the MSVD method and detecting brain tumors with image segmentation. Image segmentation involves partitioning an input image into different segments with a strong correlation with the region of interest in the given image. The field of medical image analysis is growing, and the segmentation of the organs, diseases, or abnormalities in medical images has become demanding [6]. Medical image segmentation aims to represent a given input image in a meaningful form to study the volume of tissue in order to measure the tumor size and help in deciding the dose of medicine, planning of treatment prior to applying radiation therapy, or calculating the radiation dose. Recently, deep neural models have shown application in various image segmentation tasks. We used the U-Net, a convolutional network for Biomedical Image Segmentation. The U-NET was developed by Olaf Ranneberger et al. for Biomedical Image Segmentation [7]. The architecture contains two paths. The encoder path is used to capture the context of the image. The encoder is just a traditional stack of convolutional and max pooling layers. The decoder is the symmetric expanding path that is used to enable precise localization using transposed convolutions. This network only contains convolutional layers and does not contain any dense layer because of which it can accept images of any size. The neural network architecture of the U-Net that we use in this paper is given in figure 2.

In general, the proposed structure includes a part to reduce image noise, and another part is a UNet neural network for brain tumor segmentation.

2.4. Experimental Results

To check the effectiveness of the proposed method, we have considered a brain MRI dataset, including 3900 images with a size of 256,256 pixels. The images were obtained from The Cancer Imaging Archive (TCIA). After that, the images were improved using the MSVD method. Then, diagnosed images are used as input to the U-Net neural network. With the Adam optimizer, the learning rate is set to 1e-4. We have also employed data augmentation methods to improve the robustness of the model, including scaling, shifting, flipping, and horizontal flipping. The model is trained using 70 epochs. The U-Net architecture of the proposed method is illustrated in Fig. 2.

In Fig. 3, we present the network's outputs and an MRI image with the prediction mask. We have considered three evaluation metrics to compare the results: binary accuracy, Intersection-OverUnion (IOU), and Dice Coefficient, which are defined as follows:

$$\text{Binary accy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$IOU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

$$Dice = \frac{2 \times Area\ of\ Overlap}{Total\ number\ of\ pixels}$$

A comparison of final evaluation metrics is displayed in table 1. Furthermore, we can see that using denoised images causes faster network convergence in Fig. 4.

Table 1. Comparison of final evaluation metrics

data	**Binary accuracy**	**IOU**	**Dice-coef**
Orginal image	0.9980	0.7984	0.8865
Denoised image with MSVD	0.9982	0.8116	0.8946

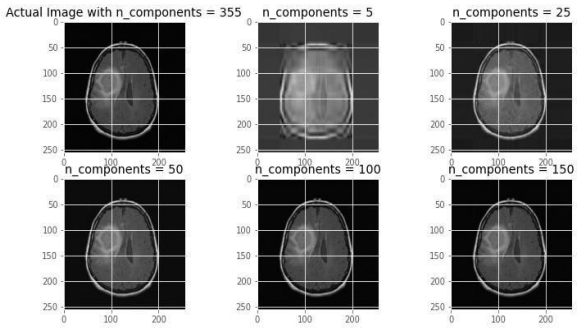


Figure 1. The effect of using the K-SVD method on MRI images by choosing different K values

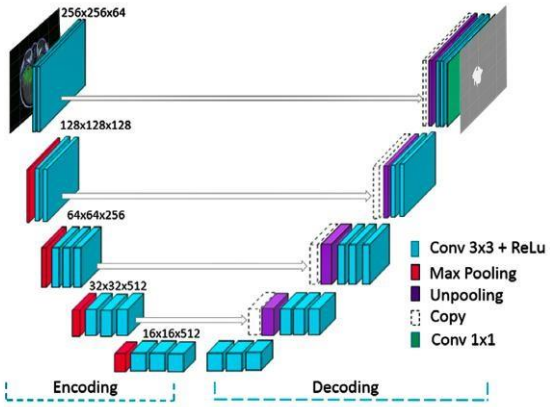


Figure 2. U-Net architecture related to mentioned method

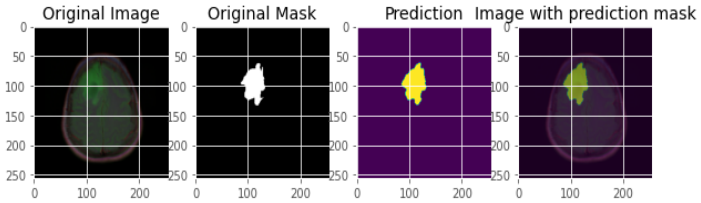


Figure 3. comparing the network output with the original mask

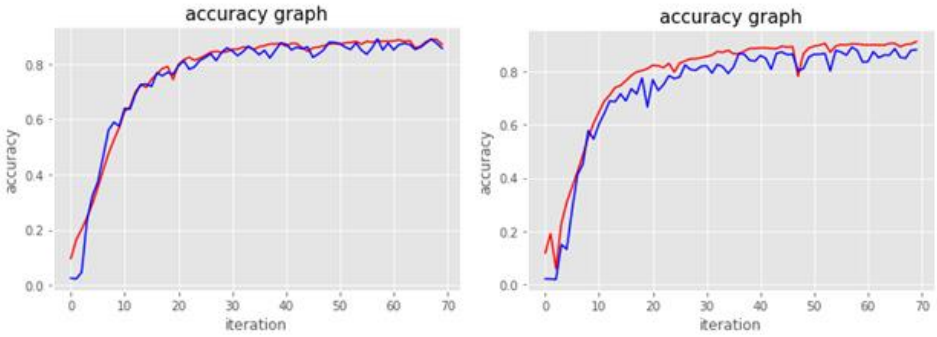


Figure 4. Comparing the accuracy of U-Net with (a) denoised data and (b) initial data

3. CONCLUSIONS

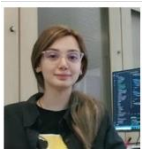
One of the challenges in processing and diagnosing diseases according to MRI and CT scan images is reducing the image noise, which can be achieved by analyzing individual values. This research investigates the effect of using the MSVD method on speed and accuracy detection in a convolutional neural network. The results show that using a singular value experiment to destroy the noise and denoise the images has reduced the execution time of each epoch of neural network learning. Besides, the speed of convergence of the network has increased significantly.

REFERENCES

[1] Lotfi Y, Parand K. Efficient image denoising technique using the meshless method: Investigation of operator splitting RBF collocation method for two anisotropic diffusion-based PDEs. Computers & Mathematics with Applications. 2022 May 1;113:315-31.
[2] Workalemahu T. Singular value decomposition in image noise filtering and reconstruction.
[3] Knapp AW. Advanced algebra. Springer Science & Business Media; 2007 Oct 11.
[4] Ghaderi-Kangavari A, Rad JA, Parand K, Nunez MD. Neuro-cognitive models of single-trial EEG measures describe latent effects of spatial attention during perceptual decision-making. bioRxiv. 2022 Jan 1.
[5] Malini S, Moni RS. Image denoising using multiresolution singular value decomposition transform. Procedia Computer Science. 2015 Jan 1;46:1708-15.
[6] Sharma N, Aggarwal LM. Automated medical image segmentation techniques. Journal of medical physics/Association of Medical Physicists of India. 2010 Jan;35(1):3.
[7] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234-241). Springer, Cham.

AUTHORS

Pegah Ahadian



ENSEMBLES FOR CLASS IMBALANCE PROBLEMS IN VARIOUS DOMAINS

Deepakindresh N, Gauthum J, Jeffrin Harris,
Harshavardhan J and Shivaditya Shivganesh

School of Computer Science Engineering,
Vellore Institute of Technology, Chennai, India

ABSTRACT

The paper is an analysis of class imbalance problems from various domains such as the medical field, sentiment analysis, software de-fects, water portability, and relationship status of students and summarizes the performance of data resampling techniques such as random undersampling and oversampling. Synthetic minority oversampling techniques combined with the power of ensemble methods such as bagging, boosting, and hybrid techniques are generally used to solve the class imbalance problem.

KEYWORDS

Machine Learning, Class Imbalance, Multi Domain Analysis.

1. INTRODUCTION

Data being the most important aspect of machine learning and data science, without which no matter how strong or complex models are used, it cannot solve a problem without having an abundant size of data. Although in many cases this is not true, for instance, real-time data does not always promise equally skewed data for all classes since one class might dominate over the other class and hence lead to class imbalance wherein the smaller class is misconstrued as noise in the dataset, thus misleading machine learning algorithms. This is a major problem as the models would generally underfit the data and if tweaked with a lot of hyperparameter tuning, it again quickly escalates to overfitting, which is why data resampling techniques are used where the data corresponding to the lower class are either randomly resampled or oversampled using nearest neighbors and the other method is under-sampling the higher classes. Both have their advantages and they are analyzed in this paper. Apart from data resampling techniques, another great way to counteract the class imbalance problem is using ensemble models which combine the advantages of many models to give priority and weights to the models that perform best in one particular subset of the data than others.

Ensembles for class-imbalance problems (ECIP)[7] are a subset of ensembles that use data resampling to pre-process imbal-anced data before learning. The performance ofvarious ECIPs is experimentally compared in this work utilizing the performance metrics F1 scoreand g-Mean over five datasets that predict the rate of heart failure, customer satisfaction, software defect, student relationship and water potability. Each of these datasets from different domains has a class imbalance problem due to which nor-mal stand-alone models suffer from true positive identification which is identifying the lower class.

David C. Wyld et al. (Eds): SIP, CRBL, WiMNeT, BIOM, EDUPT, NLDM, SOFTFM, COMIT - 2022
pp. 173-180, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121718

2. DATA ANALYSIS

Data was collected from kaggle and it can be clearly seen from the graphs how big of a class imbalance problem is present in the dataset. For preprocessing columns with no correlation to the dataset were removed with a threshold set to less than 0.2 and ows with NA values were removed. For software defect, dataset with features such as lines of code etc was used to predict if the data has software defects or not, from the [Fig 1] class imbalance problem is visible where the system with no defects is the higher class and with defects is the subordinate class. The data in both the classes are approximately in an 11:89 ratio which could cause tremendous bias in the classification.

For students, dataset [2] the response variable is the relationship status of the studentwhich is predicted using features such as age, gen-der, marks, parents' occupation etc and as seen from the [Fig 5] the ratio of single students to com-mitted is 6:19.

For the Heart failure dataset, various features were used to determine if the patient suffers from heart disease or not and the class imbalance ratio is close to 14:86. The data split can be seen in [Fig 2].

For water, potability [4] dataset the features used where chemical quanti-ties such as ph values etc, and class imbalance are visible from [Fig 4] where the imbalance ratio is 4:21%.

For the airline, [3] dataset the features were ratings on each aspect such as reviews on maintenance, safety, communication, food quality, etc were used to predict the total re-view whether the travel experience was good or bad.The class imbalance ratio is 19:81 and is visualized in [Fig 3]

3. CLASSIC ENSEMBLES

There are two sections namely the Classic ensembles and Ensembles for the class imbalance problem. The Classic ensembles employed in this research are briefly described in the first section, and the ECIP is described in the next section.

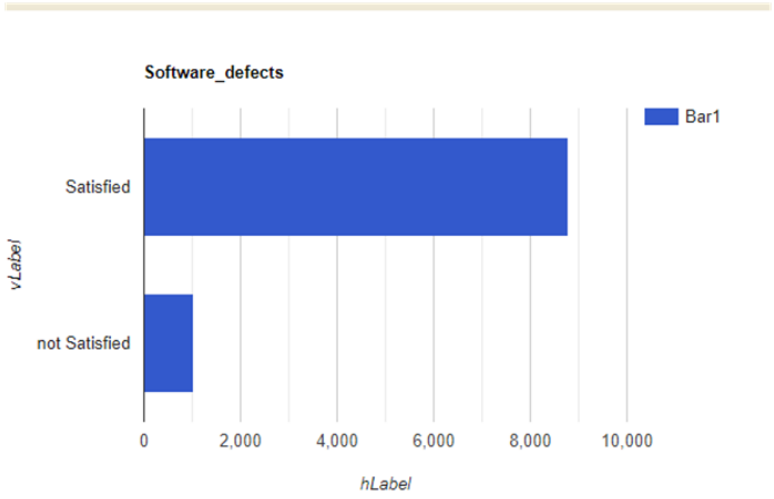


Figure 1. Distribution of defect and no defect class labels

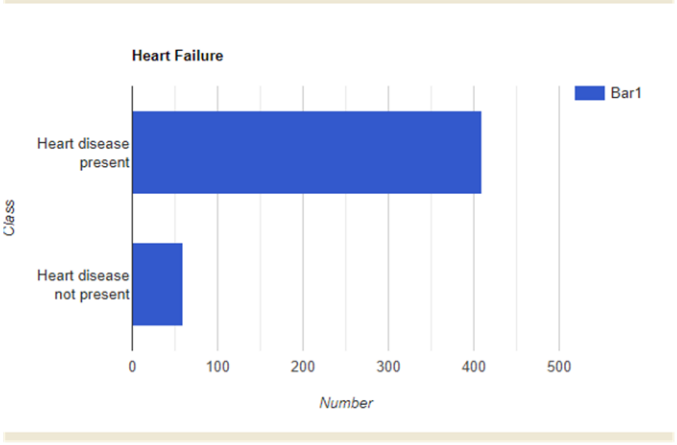


Figure 2. Distribution of disease and no disease class labels

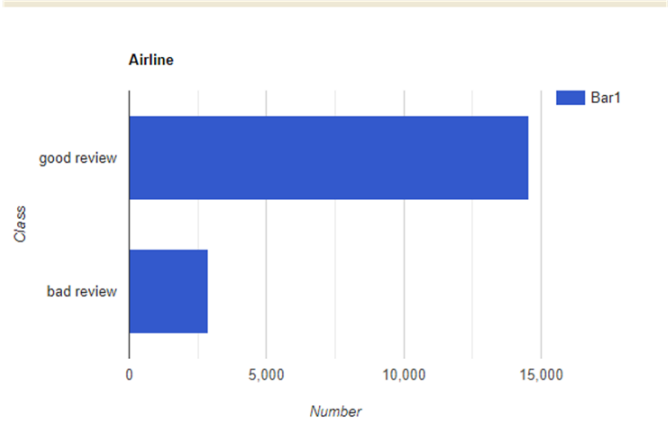


Figure 3. Distribution of good and bad review class labels

3.1. Bagging

Bootstrap Aggregation (aka Bagging) is a basic yet effective ensemble approach. The Bootstrap approach is applied to a high-variance machine learning system, such as decision trees [8], in the process of bagging. Consider the case when there are N observations and M features. A sample from observation is selected randomly with replacement(Bootstrapping). A subset of features is selected to create a model with a sample of observations and a subset of features. The feature from the subset that yields the best split on the training data is chosen. This process is repeated to produce a large number of models, each of which is trained in parallel. A forecast is made by aggregating all of the models' predictions. Bagging is the ideal approach if the single model's issue is overfitting. Boosting, on the other hand, does not prevent over-fitting; in fact, this strategy is plagued by this issue. As a result, Bagging is more successful than Boosting in this situation.

3.2. Boosting

Boosting is a set of algorithms that apply weighted averages to turn poor learners become good learners. In contrast, bagging had each model run in dependently before aggregating the outputs

in the end without giving any model preference. Boosting is all about ”collaboration.” Each model that runs determines which features will be prioritized in the following model. If the issue is that a single model performs poorly, Bagging is unlikely to produce a better bias. Boosting, on the other hand, may result in a combined model with reduced errors by maximizing the advantages and minimizing the drawbacks of a single model.

4. ECIP MODEL

The hybridization of data resampling techniques and classic ensembles is the ensemble for the class-imbalance problem. For the learning process, data resampling algorithms give balanced data distribution to ensembles. The baseline model that was used for these problems is the Decision tree classifier which runs on non-resampled data to differentiate the performance between normal models and ECIP models. The ECIP is categorized into three parts: boosting-based ensembles, bagging-based ensembles, and hybrid ensembles.

4.1. Boosting based ECIP Models

4.1.1. Adaboost

AdaBoost is a boosting technique. The technique’s main focus is on difficult-to-understand cases. The complete dataset is presented to each classifier sequentially at the start of the learning phase, with identical weights assigned to all of the instances. Following iteration, the algorithm’s main attention is on the cases that are difficult to classify, which are correctly categorized in succeeding iterations.

4.1.2. SMOTE Boost

SMOTEBoost is a blending of SMOTE and AdaBoost into a single method. With the aid of SMOTE, synthetic instances are produced in SMOTEBoost to oversample the minority class with each round of boosting. Each poor learner receives balanced material in this manner, allowing them to learn more effectively.

4.1.3. RUS Boost

RUSBoost is a combination of random undersampling and AdaBoost. It works similarly to SMOTE-Boost and MSMOTEBoost, but it ensures that each round of boosting has a balanced data distribution by removing the majority of class instances at random. RUSBoost is a simpler and quicker ensemble to deal with class-imbalance problems than SMOTEBoost and MSMOTEBoost since a lower amount of data points are provided to the classifiers during each round of boosting

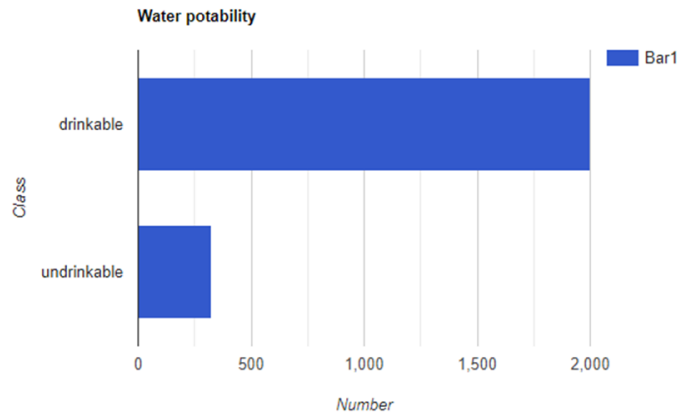


Figure 4. Distribution of drinkable and not drinkable class labels

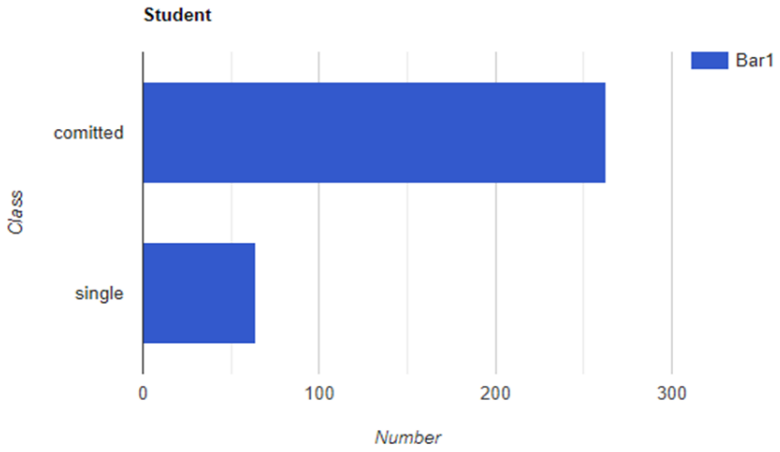


Figure 5. Distribution of in-relationship and not in relationship review class labels

4.2. Bagging based ECIP Models

4.2.1. Over Bagging

This technique is an amalgam of oversampling sampling with Bagging. In this approach, each subset of training data is generated by oversampling the instances of the minority class. The subsets of the balanced training dataset are then used to train the individual classifiers in forming the ensemble.

4.2.2. Under Bagging

This technique is a combination of undersampling and bagging. Each classifier in the ensemble is built iteratively from a subset of a balanced training dataset obtained via undersampling, which contains the majority of class instances. When an unseen instance is provided to the ensemble after the individual classifiers have been constructed, the majority vote on the individual classifiers' predictions is used to output the class label of the unseen instance.

5. RESULTS

The results of the experiments with all the models mentioned above for the software defects, airline, heart disease, water potability, and student relation-ship are provided in the tables below.

From Table 1 It can be inferred that RusBoost works with the best efficiency for the Software Defects dataset followed by SmoteBoost and EasyEnsemble.

Table 1. Software Defects

S.No	Model	g-mean	F1-score
1	Decision Tree	0.3963	0.158590
2	Bagging	0.192285	0.069665
3	AdaBoost	0.096198	0.018182
4	Easy Ensemble	0.615819	0.202381
5	RusBoost	0.618666	0.227378
6	SmoteBoost	0.615819	0.202381
7	underBagging	0.556696	0.184332
8	Over Bagging	0.213254	0.069444

From the Table 2 it is prominent that under bagging works with best efficiency for the Airline dataset followed by Over Bagging and Bagging.

Table 2. Airline

S.No	Model	g-mean	F1-score
1	Decision Tree	0.915663	0.851518
2	Bagging	0.918186	0.890930
3	AdaBoost	0.865817	0.820467
4	Easy Ensemble	0.891907	0.740586
5	RusBoost	0.906828	0.828364
6	SmoteBoost	0.891907	0.740586
7	underBagging	0.928717	0.870704
8	Over Bagging	0.921145	0.885077

From the Table 3 It can be clearly deduced that SmoteBoost is the most efficient for the Heart Disease dataset followed by Easy Ensemble and Over Bagging.

Table 3. Heart Disease

S.No	Model	g-mean	F1-score
1	Decision Tree	0.600481	0.480000
2	Bagging	0.693375	0.592593
3	AdaBoost	0.725563	0.500000
4	Easy Ensemble	0.764811	0.645161
5	RusBoost	0.762713	0.564103
6	SmoteBoost	0.764811	0.645161
7	underBagging	0.740322	0.666667
8	Over Bagging	0.745177	0.692308

From the Table 4 It can be understood that under Bagging works with best efficiency for the Water potability dataset followed by Over Bagging and RusBoost.

Table 4. Water Potability

S.No	Model	g-mean	F1-score
1	Decision Tree	0.915663	0.851518
2	Bagging	0.918186	0.890830
3	AdaBoost	0.865817	0.820467
4	Easy Ensemble	0.891907	0.740586
5	RusBoost	0.906828	0.828364
6	SmoteBoost	0.891907	0.740586
7	underBagging	0.928717	0.870704
8	Over Bagging	0.921145	0.885077

From the Table 5 It can be seen that RusBoost works best for the Student Relationship dataset followed by Under Bagging and Decision tree.

Table 5. Student Relationship

S.No	Model	g-mean	F1-score
1	Decision Tree	0.647382	0.357143
2	Bagging	0.551447	0.352941
3	AdaBoost	0.445904	0.234294
4	Easy Ensemble	0.572263	0.285714
5	RusBoost	0.709171	0.413793
6	SmoteBoost	0.572263	0.285714
7	underBagging	0.683986	0.434783
8	Over Bagging	0.450254	0.250000

5.1. Metrics and Evaluation

5.1.1. F1 Score

F1 Score [6] is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. The equation1 is provided below.

$$F1_{Score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

5.1.2. G-Mean

G-Mean [5] represents the geometric mean of true positive rate and true negative rate. A good prediction model should have high g-Mean. The equation 2 is provided below

$$gMean = \sqrt{Recall * Specificity}$$

6. CONCLUSION

Class imbalance problems are a widespread concern in the current era of Big Data Collection. Implementation of efficient methods to counteract class imbalance problems is vital for progression in the machine learning domain. From our experiments it is evident that the best models for every domain consistently are RusBoost, UnderBagging followed by a little lesser performance from Overbagging, Easy ensemble, and Smote Boost. As this experiment was

conducted on datasets from multiple domains this could be a proof of concept that ECIP models will produce a better result for datasets having class imbalance problems as our best models generalize well for all types of data with various sizes.

ACKNOWLEDGEMENTS

We wish to express our sincere thanks and deep sense of gratitude to our project guide, Dr. Shivani Gupta, Associate Professor, School of Computer Science and Engineering (SCOPE), for her consistent encouragement and valuable guidance pleasantly offered to us throughout the project work.We are extremely grateful to Dr. Neela Narayanan, Dean of the School of Computer Science and Engineering, VIT Chennai, for extending the facilities of the School towards our project and for his unstinting support. We also take this opportunity to thank all the faculty of the School for their support and the wisdom imparted to us throughout the course. We thank our parents, family, and friends for bearing with us throughout our project and for the opportunity they provided us in undergoing this course at such a prestigious institution.

REFERENCES

[1] S. Mohan, C. Thirumalai, και G. Srivastava, 'Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques', IEEE Access, τ. 7, σσ. 81542–81554, 2019.

[2] H. P. Libbey, 'Measuring student relationships to school: Attachment, bonding, connectedness, and engagement', The Journal of school health, τ. 74, τχ. 7, σ. 274, 2004.

[3] M. An και Y. Noh, 'Airline customer satisfaction and loyalty: impact of in-flight service quality', Service Business, τ. 3, τχ. 3, σσ. 293–307, 2009.

[4] A. Y. Itah και C. E. Akpan, 'Potability of drinking water in an oil impacted community in southern Nigeria', 2005.

[5] R. P. Espíndola και N. F. F. Ebecken, 'On extending f-measure and g-mean metrics to multi-class problems', WIT Transactions on Information and Communication Technologies, τ. 35, 2005.

[6] H. Huang, H. Xu, X. Wang, και W. Silamu, 'Maximum F1-score discriminative training criterion for automatic mispronunciation detection', IEEE/ACM Transactions on Audio, Speech, and Language Processing, τ. 23, τχ. 4, σσ. 787–797, 2015.

[7] R. Malhotra και K. Lata, 'Using Ensembles for Class-Imbalance Problem to Predict Maintainability of Open Source Software', International Journal of Reliability, Quality and Safety Engineering, τ. 27, σ. 2040011, 2020.

[8] F. Pedregosa κ.ά., 'Scikit-learn: Machine Learning in Python', Journal of Machine Learning Research, τ. 12, σσ. 2825–2830, 2011.

AUTHOR INDEX