IMPLEMENTATION OF A DECISION SUPPORT SYSTEM AND BUSINESS INTELLIGENCE ALGORITHMS FOR THE AUTOMATED MANAGEMENT OF INSURANCE AGENTS ACTIVITIES

Alessandro Massaro, Antonio Panarese, Michele Gargaro, Costantino Vitale and Angelo Maurizio Galiano

Dyrecta Lab, IT Research Laboratory, via Vescovo Simplicio, 45, 70014 Conversano (BA), Italy

ABSTRACT

Data processing is crucial in the insurance industry, due to the important information that is contained in the data. Business Intelligence (BI) allows to better manage the various activities as for companies working in the insurance sector. Business Intelligence based on the Decision Support System (DSS), makes it possible to improve the efficiency of decisions and processes, by improving them to the individual characteristics of the agents. In this direction, Key Performance Indicators (KPIs) are valid tools that help insurance companies to understand the current market and to anticipate future trends. The purpose of the present paper is to discuss a case study, which was developed within the research project "DSS / BI HUMAN RESOURCES", related to the implementation of an intelligent platform for the automated management of agents' activities. The platform includes BI, DSS, and KPIs. Specifically, the platform integrates Data Mining (DM) algorithms for agent scoring, K-means algorithms for customer clustering, and a Long Short-Term Memory (LSTM) artificial neural network for the prediction of agents KPIs. The LSTM model is validated by the Artificial Records (AR) approach, which allows to feed the training dataset in data-poor situations as in many practical cases using Artificial Intelligence (AI) algorithms. Using the LSTM-AR method, an analysis of the performance of the artificial neural network is carried out by changing the number of records in the dataset. More precisely, as the number of records increases, the accuracy increases up to a value equal to 0.9987.

KEYWORDS

Decision Support System, Artificial Intelligence, Long Short Term Memory, Key Performance Indicator, Data Mining.

1. INTRODUCTION

In recent years, the number of applications of Data Mining (DM) to the Human Resources (HR) domain has increased, because it provides valuable tools for talent management [1]. DM algorithms allow to extract knowledge, which is necessary for the Decision Support System (DSS) in a decision-making perspective [2]-[9]. The aim of this work is to discuss a case study related to the implementation of an intelligent platform based on DSS applied in the insurance sector by adopting Artificial Intelligence (AI) facilities. Due to Machine Learning (ML), AI is a powerful tool that can be applied in various fields [8]-[16]. AI can learn from data and adapt to a
The platform integrates different AI algorithms such as k-means, scoring algorithms and Long Short-Term Memory (LSTM) neural network, into the data mining engine to process the gathered data. In particular, the LSTM algorithms predict the Key Performance Indicators (KPIs) of the agents. As for companies working in other sectors [2], [3], [20], insurance companies use KPI or metric to measure its performance and efficiency. In fact, the use of insurance metrics permits the identification of areas that are not functional. In this sector there are many variables that make analysis complicated, therefore it is important to have KPIs that make it possible to focus on essential and information-rich parameters. The prediction of these parameters allows the company to not only monitor the efficiency and performance of agents, but also to increase their motivations.

The most significant KPIs, which make it possible to trace the motivations of insurance agents and improve the performance of the company, are:

- **Profits**, critical indicator to companies working in any sector, because it motivates strongly and directly an agent;
- **Sales**, the backbone of the insurance industry, which allow the company and the agent to make a profit;
- **Undertaken Initiatives**, the number of initiatives taken by an agent is very important to test his motivation to increase the sold products;
- **Quote Rate**, the ratio between the number of quotes that a staff member was able to provide and the number of its contacted leads;
- **Problem Resolution Rate**, it measures the ability to resolve the customer problems quickly and effectively; agents who have a high value of this KPI are efficient and save time; this leads to generate great profit in a short space of time, therefore, it motivates the agent;
- **Retention**, the number of customers who renews the insurance coverage after its expiration. It is an index of the agent's ability to retain a customer; retaining a customer is very convenient for insurance companies, because it is less expensive than finding a new customer;
- **New Policies per Agent**, the number of new insurance policies sold by an agent in a specific period of time; this metric can be monitored by product line;
- **Bind Rate**, the percentage number of quotes that turn into policies. This metric measures the agent’s ability to close a deal;
- **Sales Growth Rate**, the measure of sales tendency of an agent over a certain period; this KPI can be it can be referred to the number of new policies and/or the number of policy renewals, and can be calculated by considering new policies and / or renewals.

The monitoring of these KPIs allows an insurance company to analyse in depth the operational performance of its employees, providing a useful tool for both evaluating agents and developing corporate strategies. In fact, the metrics make it possible to identify company flows and business areas on which it is necessary to intervene with improvements of greater impact. The goal of the paper is to provide an innovative tool useful for the improvement of service business by managing efficiently insurance agents’ activities. The paper is structured as follows:

- is discussed the platform design defining the requirements, and using the Unified Modeling Language (UML) model approach;
are provided the results about KPI and the applied LSYM model (the LSTM model processes the information contained in the input dataset that is organized according to the following features: initiative type, post code of the client, agent name and event date); are shown the layouts of the prototype implemented platform.

2. DESIGN OF THE INTELLIGENT PLATFORM

The platform combines DM algorithms with an LSTM neural network in order to provide a DSS to the agents, supporting the actions to perform to increase performances, such as the best policy to propose to the customer. Furthermore, the DSS system, on the basis of the outputs provided by the DM algorithms and the predictions obtained from LSTM, offers E-Learning support to the agent. Figure 1 shows the Use Case Diagram of the project platform, where are indicated the actors of the system. Agents use a mobile app to access the platform and to take E-learning courses that have the dual function of training and motivating them. The data collected by the mobile app and by the company information system, are analyzed by the DM algorithms. Then is applied the K-means algorithm performing client clustering. This is one of the inputs that will allow the DSS to suggest the most profitable policies to the agent. The DM algorithms are also the basis of the scoring mechanism that assigns a score to agents based on their activities, such as number of initiatives undertaken, number of policies sold, frequency of use of the application, e-learning courses followed and etc. Scoring is also based on the use of KPIs that allow the insurance company to track the performance of agents.

In addition to the functions previously analyzed, the platform integrates an LSTM neural network into the data mining engine implemented to process the data collected by means of the mobile app in order to predict the KPIs of individual operators. The prediction of these parameters will allow the company not only to monitor the efficiency and performance of agents, but also to increase their motivations.

![Use Case Diagram](image-url)
Figure 2 shows the architecture of the LSTM network on which the predictive model is based.

![Architecture of the LSTM neural network for KPI prediction](image)

An LSTM layer and a dense layer provide the output data. Adam was used as an adaptive learning rate optimization algorithm. This adaptive learning rate method is very useful for training deep neural networks, because it identifies the weights that minimize the difference between predicted and actual values. Adam calculates separate learning rates for different parameters and adapt the learning rate for every single weight of the neural network.

Ultimately, the algorithms implemented in the platform allow the scheduling of activities by managing them through an automated decision support calendar, according to the following rules:

- important activities due
- important activities not due
- unimportant activities due
- unimportant activities not due.

Using the BI tools via a mobile responsive app, the system administrator is able to consult the result of data processing by the tools implemented to support the improvement of business performance. Among these results, are include the prediction of KPI values of strategic importance to track the behavior of users of the mobile application, namely the agents.

The platform integrates the following modules for analyzing and managing the performance of agents (Figure 3):

- A mobile app through which the agents can follow online training courses and view the activity to be carried out; the functions of Google Maps are also integrated to facilitate and speed up logistical operations; the application is also useful for traceability of
customer management and for storing customer data, necessary for customer clustering, in a local database.

- A Data Mining Engine based on:
  - Scoring algorithms for profiling the agents based on their experience and aptitude;
  - Clustering algorithms for customer segmentation (K-means);
  - LSTM to predict the KPIs of the agents and their activities associated with the geolocation of customers.

- A DSS-based mobile responsive application, able to:
  - to plan agents’ activities by managing priorities (scheduling);
  - to monitor the daily progressions on objectives monthly of the single operator (real time monitoring);
  - to execute the data mining algorithms (data mining engine) whose outputs provide graphic dashboards / indicators to support decisions (more reliable operators for specific types of activities, risk indicators, economic indicators based on customer segmentation).

- An E-learning platform suitable for the training of operators; the contents can also be uploaded through the management application.

![Diagram](image.png)

Figure 3. Functional diagram of the intelligent platform.

The designed applications exploit a single Database (DB) significant to receive the required information, and to deposit the data necessary for the correct development of the system. Therefore, the DB allows to define the activities to be assigned to each individual agent and to keep track of the history of activities, earnings and meetings.
3. Results and Discussion

The LSTM neural network processes the data collected through the prototype mobile application and predicts the motivational KPI values of agents. In addition, the LSTM also predicts the activities of agents in relation to the geolocation of potential customers based on the zip code. The prediction by geographical unit allows to optimize the logistics and the movements of the agents because it supports the decisions of the operators in the planning of the movements to be carried out in the areas where a higher level of activity is expected.

Managers use performance measurement systems with the aim of influencing the behavior of the employees, increasing their motivations and finalizing them to achieve company objectives. For this reason, it is necessary to clearly know the reasons that stimulate agents to work effectively. The KPIs allow to track the performance of agents and, therefore, to know the parameters that most positively influence their performance.

The implemented LSTM network has been tested on the prediction of the KPI "number of initiatives undertaken". This KPI is very important because it measures the activities undertaken by an agent and aimed at increasing the number of sales. A high number of initiatives undertaken will lead to a high number of sales and therefore to a greater profit, giving continuity to the agent's performance. If a high number of initiatives is predicted for an agent, a certain reward can be anticipated, in such a way as to incentivize him to follow the predicted positive trend. Figure 4 shows the architecture of the LSTM used for the prediction of this KPI.

The input data consist of files in the csv format, with the following variables on the columns: initiative type, post code of the client, agent name and event date. Each line of the file constitutes a record. The LSTM input layer reshapes the input data by re-writing them in a suitable form for processing by the LSTM. Finally, the dense output layer gives the predictive output. A preliminary test made it possible to establish that the best LSTM configuration is based on 256 neurons and that 100 epochs are sufficient to train this network.

![Figure 4. Schematic diagram of the LSTM network through its layers](image-url)
A dashboard has been created to display the results of the AI algorithms and to test the prototype system. New data can be uploaded through the dashboard in order to improve the accuracy of the KPI prediction model. In fact, as shown in the following Figure 5, the accuracy of the model progressively increases as the available data increases. Performance measurement is very important for both basic and industrial research [1], [21]-[23]. In the present case study, the number of records that make up the dataset is crucial to improve network performance. For this reason, has been adopted the LSTM-Artificial Records (LSTM-AR) method, to test and validate the implemented network. The LSTM-AR approach is used to enrich data-poor datasets [24]. In this way the starting dataset has been enriched with records generated in a random way starting from those in possession. Figure 5 summarizes the results of the validation made starting from a dataset of 100 records and progressively enriched: 100 records, 200 records, 300 records, 400 records, 500 records, 1000 records and finally 10000 records. Accuracy grows as the number of epochs increases, but reaches a plateau only with datasets with a number of records at least equal to 300. For datasets consisting of 500, 1000 and 10000 records the accuracy reaches the excellent value of about 0.9987. The same behavior is assumed for the loss function during training and validation, in fact the best results are those relating to 500, 1000 and 10000 records.

During the testing of the LSTM, the cross-validation technique was used, i.e., the training set was divided into two distinct sets:

- training sub-set, which is actually used for training the model;
- validation sub-set, which is used to measure the generalization capacity of the model and to validate it.

During the training test the optimal value of the parameters is found, while during the validation test the system calibrates the hyperparameters. To optimize testing, the training sub-set was constructed by randomly extracting 80% of the dataset records, the remaining 20% constituting the validation sub-set [24]. Figure 6 shows the accuracy during the training and the loss function during the validation of the model, in the chosen case of a total dataset of 10,000 records.
An agent, by accessing the platform, can use a series of features, including 'Initiative Prediction', which is the output of the LSTM. To obtain the prediction of the tested KPI, number of initiatives, it is necessary to set the following parameters:

- Type of initiative (Healthcare, Assistance, Fire, …);
- Customer's postcode (70100, 70116, ...);
- agent code (015, 021, ...);
- LSTM batch (2, 4, 8, ..., 512, 1024);
- Days LSTM (2, 4, 8, ..., 512, 1024);
- Batch Prediction (2, 4, 8, ..., 512, 1024);
- Days Prediction (2, 4, 8, ..., 512, 1024).

Figure 7 shows two different examples of predicting the number of initiatives, corresponding to different choices of the parameters indicated above.

To further motivate agents, an ‘initiative coefficient’, which indicates the performance of agents with respect to the number of initiatives on a scale ranging from 0 to 10, is also indicated for each of the 4 clusters into which the clientele has been divided (Figure 8). For this important KPI a minimum threshold to be reached for each cluster was set and equal to 7 in order to motivate the agent to undertake a high number of initiatives. Figure 8 shows the values reached by an agent.
The data collected from the use of the mobile app are processed by k-means algorithms that perform the segmentation of customers into 4 clusters. For the development of the k-means algorithm, the python programming language was used, which provides the Scikit-learn library. In detail, the sklearn.cluster class was used to group customers into sets. For this reason, the following features relating to customers have been considered: product, type of product expiration, product branch, number of policies in the name of the customer, years of indemnity, customer income, property owned by the customer, pets of the customer, customer hobby. The clustering analysis carried out by the k-means algorithms allows to suggest to the agent the products to be proposed to the customer based on his characteristics.

It is observed that the LSTM approach is suitable for augmented and artificial data [25]. This validates the choice to use the LSTM tools for the specific case study. Other approaches which can adopted are Convolutional Neural Network (CNN) [26], Recurrent Neural Network (RNN) [27], Gated Recurrent Units (GRU) [27], and Artificial Neural Network (ANN) [28],[29].

4. CONCLUSIONS

The implementation of DM algorithms to classify corporate customers and score the agents, in combination with the introduction of insurance KPI measuring and analyzing performance, has led to a remarkable understanding of business dynamics. However, the crucial aspect of the developed platform is the LSTM artificial neural network, as it allows KPIs to be predicted with excellent accuracy. In fact, the training and validation of the model made possible an accuracy equal to 0.9987. This precision has been achieved through the use of the LSTM-AR method, which allows to enhance the data in cases where a large amount of data is not available. The integration of the tools described in the platform will allow the company to better manage business flows and achieve the set objectives. The use of the different technologies increases the business efficiency because during the time are stored more data in the backend system. Moreover, the business dynamics is ensured by means the mobile app which is useful also for data entry by furthermore enriching the training dataset. The proposed methodology is extensible to other use cases by formulating new KPI according to the service to improve, and by designing a new mobile app facilitating human resource activities. The business approach can be associated to business risk approach following the models found in literature [30]. Other tools to facilitate the business performance can be the E-Learning and the chatbot integrating AI [31]. The analyzed KPI and algorithms can be implemented into an information system integrating ESB and big data technologies.

ACKNOWLEDGEMENTS

The work has been developed within the framework of the Italian project: “Applicativi per l’automazione intelligente dello scheduling profilato delle attività, e per lo scoring dei venditori, orientato al trasferimento tecnologico: DSS/BI HUMAN RESOURCES” [Applications for intelligent automation of profiled scheduling of activities and for vendor scoring, oriented towards technology transfer].
REFERENCES


AUTHORS

Alessandro Massaro (corresponding author): Professor Alessandro Massaro (ING/INF/01, FIS/01, FIS/03) carried out scientific research at the Polytechnic University of Marche, at CNR, and at Italian Institute of Technology (IIT) as Team Leader by activating laboratories for nanocomposite sensors for industrial robotics. He is in MIUR register as scientific expert in competitive Industrial Research and social development, and he is currently head of the Research and Development section and scientific director of MIUR Research Institute Dyrecta Lab Srl. Member of the International Scientific Committee of Measurers IMEKO and IEEE Senior member, recently received an award from the National Council of Engineers as Best Engineer of Italy 2018 (Top Young Engineer 2018).

Antonio Panarese received his PhD degree in Physics from University of Bari (Italy) in 2011. From 2011 to 2016 he held various positions as a post-DOC researcher at the CNR and at the University of Bari, where his research covered numerical modeling and parallel computing in Plasma Physics and Gas Dynamics. From 2016 to 2017 he worked at the CMCC (Euro - Mediterranean Center for Climate Change) as member of ASC (Advanced Scientific Computing) Division and HEC (High End Computing) research group, Lecce (Italy), where his research tasks were the analysis of alternative parallel approaches for performance optimization and parallelization of NEMO (Nucleus for European Modelling of the Ocean) oceanic model. He is currently a researcher of MIUR Research Institute Dyrecta Lab. His research focuses on Artificial Intelligence with particular emphasis on Artificial Neural Networks, Predictive Machine Learning and Data Science.

Michele Gargaro. Michele Gargaro received the Master's Degrees in Information technology and Cybersecurity at University of Bari (Italy) in 2019. Actually, he is a researcher of the Dyrecta Lab Institute studying topics about artificial intelligence applied on cybersecurity in industrial applications.

Costantino Vitale. Costantino Vitale graduated in Information technology and Software Development at University of Bari (Italy) in 2020. Actually, he is a researcher at Dyrecta Lab Institute studying topics about artificial intelligence applied to telemedicine and industrial applications.

Angelo Maurizio Galiano. Angelo Maurizio Galiano is CEO at Dyrecta Lab Srl - research institute accredited by the Italian Ministry of University and Scientific Research. He has more than 20 years of experience in the field of Information Technologies. He received M.S. degree in Education Science in 2009. His current research interests include neural networks, smart health, predictive analytics.