ANALYZING CLINICAL CHARACTERISTICS AND PREDICTING HOSPITALIZATION OF OLDER EMERGENCY PATIENTS

Ala' Karajeh [1] and Rasit Eskicioglu[2]

[1] Independent Researcher, Winnipeg, Canada
[2] Department of Data Science and Analytics, Atlas University, Turkey

ABSTRACT

Older patients often present with multiple comorbidities and distinct clinical patterns, making their severity assessment more challenging in emergency settings. This study analyzes two clinical databases from Beth Israel Deaconess Medical Center to examine the triage characteristics and disposition outcomes of older emergency patients. A framework of four machine-learning models was developed and compared to a baseline logistic regression model to predict whether a patient is likely to be hospitalized or discharged based on triage information. The models demonstrated reasonable predictive performance and highlight the potential of using machine learning-based triage tools to support early risk identification and improve decision-making for this patient group.

KEYWORDS

Emergency Medicine, Hospitalization Prediction by Machine Learning, Emergency Older Patients Classification, and Emergency Older Patients Data Analytics.

1. INTRODUCTION

Hospitals and medical institutions take in patients who seek timely care in their emergency departments. The received cases comprise severe conditions such as difficulty breathing and suicidal ideations and range to minor ones including dental pain and suture removal. Identifying the urgency of received cases is usually carried out by a specific nurse, trained on such assessment with the assistance of an acuity score guideline such as Emergency Severity Index (ESI). The ESI triage system recommends considering the patient's clinical readings, realized pain, and reason for visiting the emergency department. The triage nurse reckons the available resources during assessment and assigns the case to a score on a five-level scale. Level one represents immediate urgency while level five indicatesa non-urgent situation [1][2].

Various populations head to the emergency in order to obtain timely care including older adults and children. Patients seek emergency services due to many reasons such as their subjective judgement of their health issue, advice from a relative or a friend, or being referred by a healthcare provider [3]. Older adults represent a critical population considering the physiological changes and cognitive deterioration human bodies exhibit at this stage, leading to less receptiveness to pain [4]. Furthermore, aging is often accompanied by multiple conditions such as diabetes and dementia [5]. Consequently, the aging population is more prone to revisit the emergency due to health deterioration following their initial discharge [6].

David C. Wyld et al. (Eds): DPPR, CNDC, AIAA, DSA, WeST, ACITY, IoTE – 2025
pp. 97-105, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.152408

Although machine learning (ML) applications in emergency medicine are increasing, research focused specifically on predicting outcomes for older adults remains limited. Mowbray et al. examined hospitalization prediction among older ED patients using several ML models, including support vector machines, but did not evaluate neural networks [7]. Tan et al. explored ML-based forecasting of multiple outcomes, including hospitalization, among older adults with influenza; however, their analysis was restricted to a disease-specific cohort [8].

Given the complexity of clinical presentations in older adults and the limited evidence on ML-based admission prediction in this population, this study develops a comparative framework evaluating neural networks, ensemble learning methods, and logistic regression models to predict hospitalization outcomes among older emergency patients—a comparison not previously examined to our knowledge. Additionally, it applies health data analytics to describe the clinical characteristics of this population and proposes a two-group disposition stratification based on predicted admission probability.

2. METHODOLOGY

2.1. Database Access and Pre-processing

Two clinical databases were utilized and accessed through the Physionet.org repository. These databases were granted after meeting the requirements specified by the hosting website. Patient information was collected and de-identified by the Beth Israel Deaconess Medical Center in Boston, USA. The data encompassed emergency and demographic data, where emergency data comprised triage specifics such as vital signs and acuity scores besides case disposition. The emergency database covered emergency cases received by the center from 2011 to 2019.

Dataset analysis and model development were carried out using RStudio (v2023.12.1 - Build 402) and the R language (v4.2.1). Three different datasets were extracted to retrieve the patient data required for this research. These datasets were then merged based on the identifiers of a single emergency visit and an individual patient. Demographic data were obtained from the MIMIC-IV (v2.0) database, while emergency data were retrieved from the MIMIC-IV-ED (v2.0) database.

2.2. Dataset Processing

Prior to performing the data analysis, the dataset was manipulated to facilitate the next steps. The year of the emergency visit was derived from the registration time variable. This derived feature was then used to calculate the age of the emergency patients. Cases without pre-assigned acuity scores were eliminated from the dataset. Additionally, non-relevant variables were excluded. Missing values (NAs) were handled by being replaced with the values from the closest possible case using the K-Nearest Neighbor method, which is usually more accurate than conventional imputation methods.

2.3. Data Analysis

The Corrgram and Psych packages were used to compute the correlation coefficients between features and acuity outcomes and to determine their statistical significance. Following this, individual relationships between predictors and the outcome variable were examined to understand how these features vary among the five acuity score groups. The data analysis included not only the acuity outcome but also extended to another outcome variable – disposition. Disposition was analyzed by stratifying its various groups based on the age variable.

2.4. Model Development

Following the analysis phase, four machine learning-based classifier models were developed using the Caret package, which are SGB (Stochastic Gradient Boosting), MLP (Multi-Layer Perceptron), RF (Random Forests) and CART (Classification and Regression Trees), to distinguish between admitted and discharged cases. Additionally, a logistic regression model was established as a reference model. For reproducibility, the seed was set to 12345.A subset consisting of 30,559 records was utilized and divided into 80% for training and 20% for testing. Five-fold cross-validation was performed during the training phase to enhance the generalizability of the models. Another seed, set to 1234, was used before executing the training. The subset incorporated 12 predictors: temperature, heart rate, oxygen saturation, respiratory rate, diastolic blood pressure, systolic blood pressure, age, reported pain, triage acuity score, chief complaint, arrival method, and gender.

2.5. Model Evaluation

The developed models were assessed by predicting the hospitalization outcomes for the test set. These predicted results were then compared with the actual outcomes from the datasets. Performance measures were enumerated using a confusion matrix, considering sensitivity, specificity, and balanced accuracy. Sensitivity represents a model's ability to identify positive outcomes among all actual positive outcomes. Specificity indicates a model's ability to detect negative outcomes among all actual negative outcomes. Balanced accuracy is the average of sensitivity and specificity.

3. RESULTS

Heart rate readings were considerably higher for the immediate patients group, scoring a mean of 103.61 with a standard deviation 35.69. On the contrary, the four remaining groups of acuity were relatively convergent, where the lowest mean heart rate was attained in acuity group four, with an average of 78.23 and a standard deviation of 14.24. Also, the patients of these four groups exhibited outlier readings compared to the immediate cases. Further information is available in the following figure.

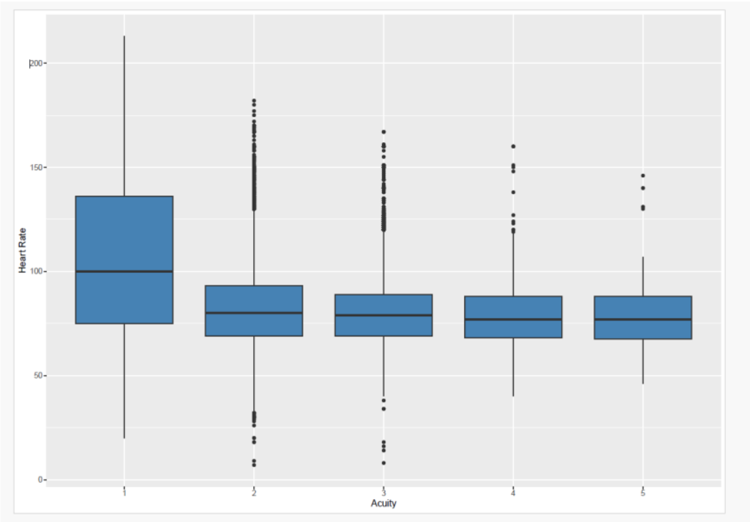

Figure 1: Acuity vs Heart Rate for the recorded older patients visits in the emergency department.

Pain feature was analyzed too, where semi-urgent cases came first in this regard with a mean of 4.48, followed by urgent cases whose perceived pain average was 4.02. Nonetheless, pain means for the three remaining groups did not exceed three as level five cases came last, scoring a mean of 2.06±3.68. Further details on the analysis of this feature are illustrated in the below figure.

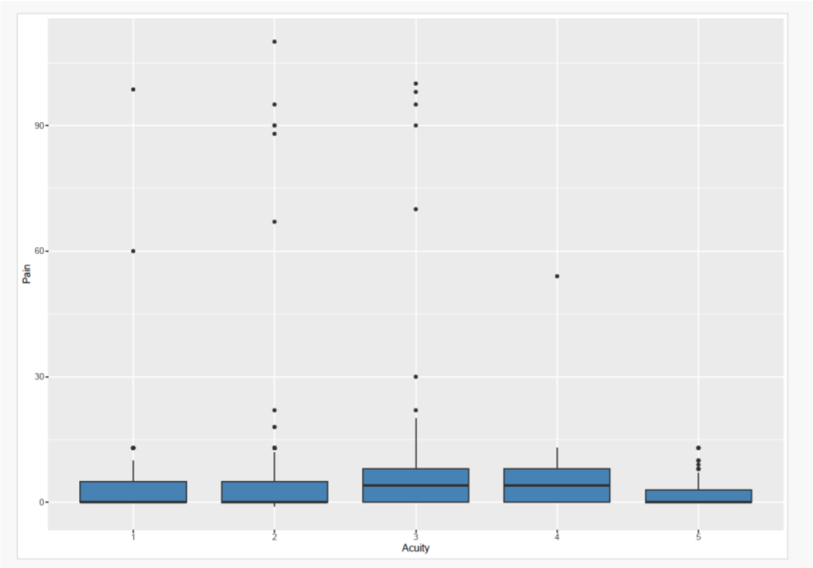

Figure 2: Acuity vs Pain for the recorded older patients visits in the emergency department.

Respiratory rates for the most severe patients were significantly higher than the four lower groups. Level one cases scored a mean of 22.88±7.62, while the averages of the lower acuity groups did not exceed 18.2. The lowest recorded values were for level four cases, with an average of 17.01 and a standard deviation of 2.09. Notably, this feature's most outstanding data dispersion was revealed in the level one group among the five acuity groups. Additional details on this are provided in the following figure.

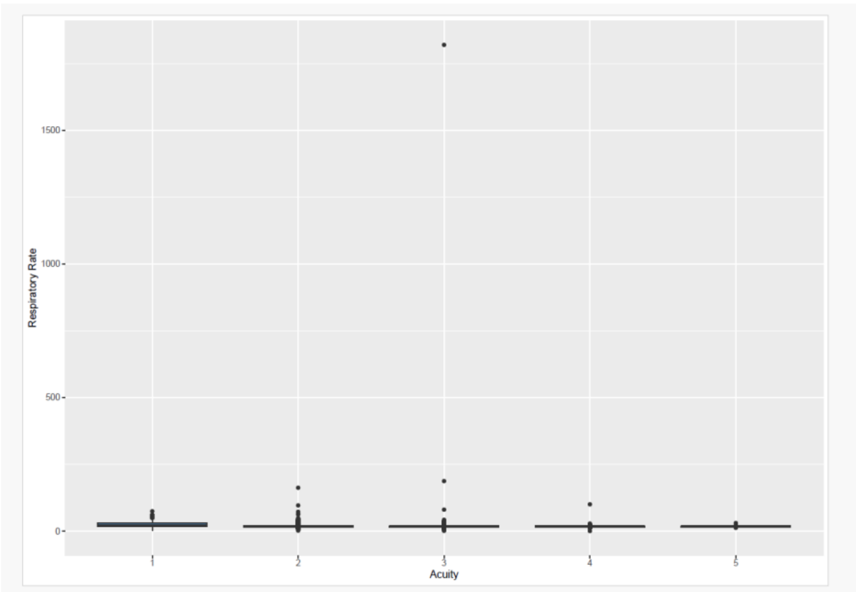

Figure 3:Acuity vs Respiratory Rate for the recorded older patients visits in the emergency department.

Concerning the means of arrival to the respective emergency department, the majority of older adult patients were transported to the venue via ambulance. 60.18% of the ambulance cases were identified as unstable and assigned to levels one and two accordingly. Walk-in arrival mean came second in this regard, demonstrating a lower portion of unstable cases in favour of a more significant percentage of level three cases. Level three cases among this group comprised 54.36% of the overall count. Regarding the helicopter group, unstable cases were considerably high, forming 98.62% of the cases. Further information on the five groups is revealed in the below figure.

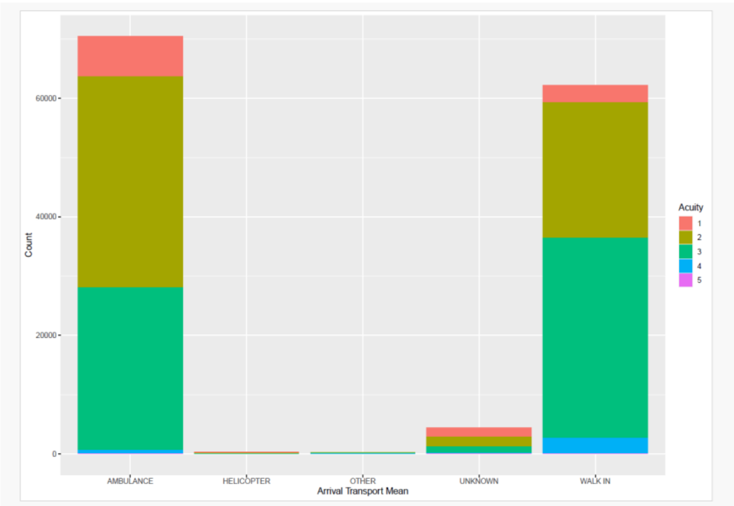

Figure 4:Received cases stratified by arrival method and severity score.

Analyzing the disposition groups of the examined emergency department shed light on variations of age specifics among these groups. Cases transferred from the emergency department exhibited significantly higher ages, with a mean of 80.23. However, cases that ended up being discharged or admitted to the hospital had age averages of 75.94±8.24 and 77.87±8.56, respectively. An additional slump in age particulars was shown in the groups that represented anomalous end-points comprising eloped patients, left with being seen and left against medical advice. Disposition groups stratified by age are explicated in the following figure.

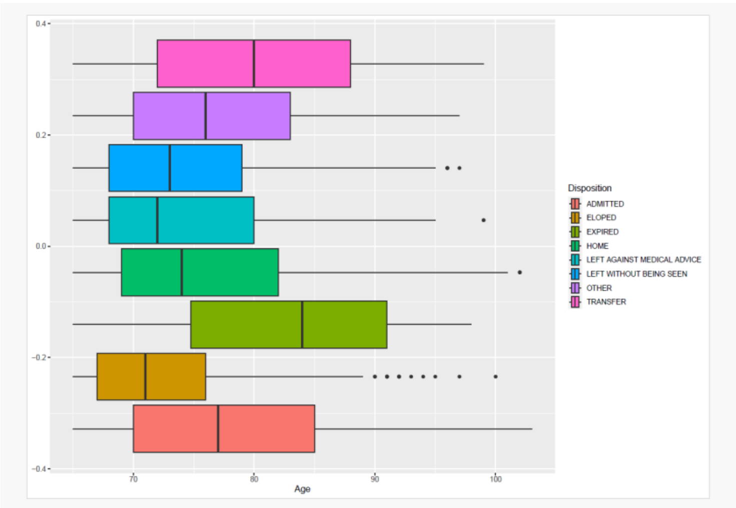

Figure 5: Age medians and ranges for the disposition groups of older patients in the emergency department.

Unveiling the most frequently reported complaints yielded varying results concerning the acuity scores assigned to them. Considerable amounts of cases that reported a clinical manifestation of dyspnea, chest pain, bright red blood per rectum and altered mental state were triaged to unstable scores of levels one and two. A decline in unstable cases was observed in groups of fall incidents and weakness. Moreover, the lowest proportion of these cases was found in patients who gave an account of wounds as a reason for their emergency visit, which made up 14.83% of this complaint group. The following plot reveals more information on the top complaints and acuity scores.

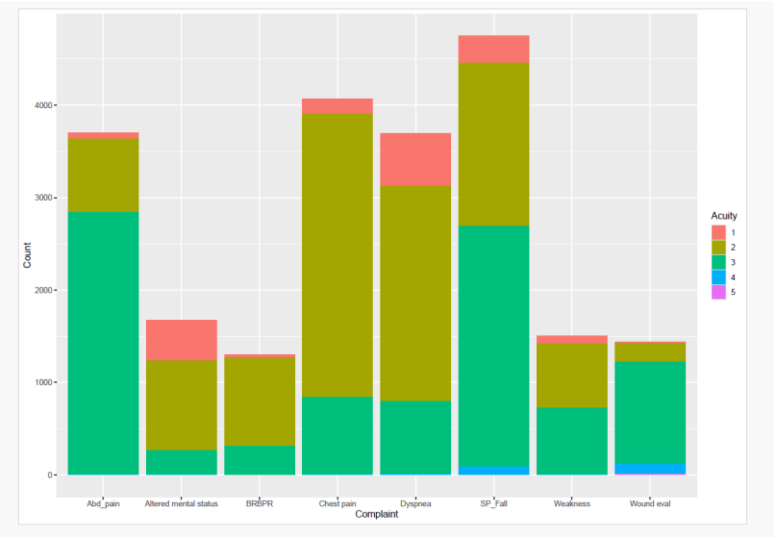

Figure 6:The most frequent complaints among the received senior patients, categorized by their acuity score.

The developed classifying models showed varying performance levels when tested on the separate test set. The highest-performing model was the SGB, with an overall accuracy of 71.81%. The MLP model achieved a balanced accuracy of 69.17%, making it the most balanced among the five models in terms of sensitivity and specificity. Only one model surpassed the reference logistic regression model in overall accuracy, which was the SGB. In contrast, the CART model had the weakest performance, with an accuracy of 69.23%. Further details on the models' performance are provided in the table below.

Table 1: Testing results of the developed models, including the reference model.

Model	Overall Accuracy	Sensitivity	Specificity	Balanced Accuracy
LR	71.55%	83.03%	51.84%	67.44%
SGB	71.81%	85%	48.60%	66.97%
MLP	71.24%	77%	61.34%	69.17%
CART	69.23%	89.39%	34.62%	62.01%
RF	71.32%	84.09%	49.40%	66.75%

4. DISCUSSION

This research aims to predict the likelihood of hospitalization or discharge from the emergency department for older patients based on vital signs, pain levels, age, gender, arrival means, reported complaints, and acuity scores. Moreover, It examines data collected during triage assessments to explore correlations with both acuity scores and patient disposition. This study provides insights into this vulnerable population and identifies patterns within different acuity groups and disposition outcomes.

The descriptive analysis helps clarify the models' performance by showing the clinical patterns present in the data. Several of the most common complaints, including dyspnea, chest pain, altered mental status, and gastrointestinal bleeding, were mainly triaged to the higher-acuity levels, reflecting the complexity of this population. Vital signs of heart rate and respiratory rate also showed noticeable separation across severity groups, which may expound the stronger sensitivity achieved by models such as SGB and RF. The MLP model's more balanced performance may be due to its ability to capture the subtler differences in factors including pain, and moderate-acuity complaints that were not as distinctly separated. On the other hand, the weaker results of the CART model likely stem from its limited capacity to handle the variability seen in higher-acuity groups and the wide range of complaints reported by older adults. Overall, these clinical patterns help explain why the ensemble models and the neural network were better suited to this dataset than the simpler decision tree approach.

Conventional triage scales indicate that high-severity cases identified at triage, such as patients at levels one and two on the ESI scale, are more likely to be hospitalized [9]. However, the benefit of these indicators is limited due to inadequate triage accuracy, which does not exceed 60% according to previous studies [10]. Additionally, the issue of under-triaging, where patients are incorrectly identified as level three instead of level two, can exacerbate conditions for older patients who usuallyhead to emergency for more pressing conditions compared to younger emergency patients [11][12]. There has also been a spike in the older emergency patient's turnout in the recent years [13]. A more reliable approach to anticipate cases likely to require admission following an emergency visit would aid in dealing with these circumstances, especially for such a vulnerable population. Data-driven prediction of hospitalization for the elderly would enable a more consistent and precise approach, significantly improving the care provided to this vulnerable segment.

This study has a few limitations worth noting. The models were built using two datasets from a single medical center (2011–2019), which means the findings may not fully generalize to other settings. The chief-complaint variable was also reduced to fifteen categories to fit the scope of this work. Therefore, the developed models did not capture the entire range of reported cases. In addition, the algorithms relied on the automatic parameter tuning of the ML library, which may have limited the model performance.

5. CONCLUSIONS AND FUTURE WORK

This study provides comprehensive insights into the clinical characteristics of older emergency patients based on triage information and disposition outcomes. It also introduces a comparative framework of machine-learning classification models designed to predict hospitalization likelihood among this population. The findings establish a foundation for future exploration of additional algorithms and model enhancements. Integrating such predictive frameworks into clinical decision-support tools could assist emergency clinicians in managing older patients more effectively and ultimately improve their health outcomes.

Future work could focus on improving the models through more advanced hyperparameter tuning to achieve stronger performance as well as further statistical analysis of models' performance. Expanding the set of clinical variables, especially by adding more detailed chief complaints or additional patient information, may also assist in capturing a wider range of cases and improving model training. These steps could support moving toward clinical validation, where healthcare providers can assess how reliable and beneficial the model would be in real practice.

The code used for data processing, analysis, and model development is available at [14].

REFERENCES

[1] J. M. Zachariasse et al., "Performance of Triage Systems in emergency care: A systematic review and meta-analysis," BMJ Open, vol. 9, no. 5, May 2019. doi:10.1136/bmjopen-2018-026471

[2] J. S. Hinson et al., "Triage performance in emergency medicine: A systematic review," Annals of Emergency Medicine, vol. 74, no. 1, pp. 140–152, Jul. 2019. doi:10.1016/j.annemergmed.2018.09.022

[3] L. Lobachova et al., "Patient and provider perceptions of why patients seek care in emergency departments," The Journal of Emergency Medicine, vol. 46, no. 1, pp. 104–112, Jan. 2014. doi:10.1016/j.jemermed.2013.04.063

[4] R. Haque, M. Bezerko, L. Tibbits, and K. Tate, "Pain assessment and documentation for older adults presenting with non-surgical conditions in emergency room," Innovation in Aging, vol. 5, no. Supplement_1, pp. 594–595, Dec. 2021. doi:10.1093/geroni/igab046.2284

[5] K. Gadó, A. Szabo, D. Markovics, and A. Virág, "Most common cardiovascular diseases of the elderly – a review article," Developments in Health Sciences, vol. 4, no. 2, pp. 27–32, Mar. 2022. doi:10.1556/2066.2021.00048

[6] M. van Loon-van Gaalen, I. E. Voshol, M. C. van der Linden, J. Gussekloo, and R. C. van der Mast, "Frequencies and reasons for unplanned emergency department return visits by older adults: A cohort study," BMC Geriatrics, vol. 23, no. 1, May 2023. doi:10.1186/s12877-023-04021-x

[7] F. Mowbray et al., "Predicting hospital admission for older adults in the emergency department: A machine learning approach," *Acad. Emerg. Med.*, vol. 27, no. 8, pp. 786–795, 2020. doi:10.1111/acem.13983

[8] T. H. Tan et al., "Machine learning prediction of clinical outcomes for older adults with influenza: A multicenter retrospective study," *BMC Geriatrics*, vol. 21, no. 1, p. 210, Apr. 2021. doi:10.1186/s12877-021-02229-3

[9] N. Gilboy, P. Tanabe, D. A. Travers, A. M. Rosenau, D. R. Eitel, and Others, 'Emergency severity index, version 4: implementation handbook', *Rockville, MD: Agency for Healthcare Research and Quality*, vol. 1, 2005.

[10] H. L. Tam, S. F. Chung, and C. K. Lou, "A review of triage accuracy and future direction," *BMC Emergency Medicine*, vol. 18, no. 1, Dec. 2018. doi:10.1186/s12873-018-0215-0

[11] O. Ivanov *et al.*, "Improving ed emergency severity index acuity assignment using machine learning and Clinical Natural Language Processing," *Journal of Emergency Nursing*, vol. 47, no. 2, Mar. 2021. doi:10.1016/j.jen.2020.11.001

[12] C. O'Keeffe, S. Mason, R. Jacques, and J. Nicholl, "Characterising non-urgent users of the Emergency Department (ED): A retrospective analysis of routine Ed Data," *PLOS ONE*, vol. 13, no. 2, Feb. 2018. doi:10.1371/journal.pone.0192855

[13] K. B. Ahsan, M. R. Alam, D. G. Morel, and M. A. Karim, "Emergency department resource optimisation for improved performance: A Review," *Journal of Industrial Engineering International*, vol. 15, no. S1, pp. 253–266, Nov. 2019. doi:10.1007/s40092-019-00335-x

[14] https://github.com/karajeh-ala/ED-Elderly-Patients-Analysis-and-Hospitalization-Prediction

AUTHORS

Ala' Karajeh holds an MSc in Biomedical Engineering and works as a Data Labeling Analyst. He graduated from the University of Manitoba in 2024 and conducted his graduate research under the supervision of Dr. Rasit Eskicioglu at the Internet of Things Laboratory, where he developed expertise in biomedical analytics and machine learning. His research interests include health monitoring, health data analytics, and the use of AI to enhance clinical care. He is also interested in applying computer vision and natural language processing algorithms in healthcare settings.

Dr. Rasit Eskicioglu is a Professor at Atlas University in Turkey. His research interests are experimental systems, particularly computer systems, systems software, operating systems, distributed, cluster, and grid computing, high-speed network interconnects, mobile networks, AI, and pervasive computing. Most recently, he is looking at wireless sensor networks (WSNs) and their applications for solving real-world problems, such as indoor localization, monitoring, and tracking, and the Internet of Things (IoT).