

Prioritizing Hospital Admission According to Emergency Using Machine Learning

Mayank Singh, Manvi Singh, Prachi Sharma, Ritin Behl

Computer Science and Engineering Artificial Intelligence and Machine Learning ABES Engineering College
Ghaziabad, India

DOI : <https://doi.org/10.51583/IJLTEMAS.2025.140400024>

Received: 17 April 2025; Accepted: 19 April 2025; Published: 03 May 2025

Abstract: The use of artificial intelligence and machine learning techniques in emergency medicine has grown rapidly. This paper reviews and assesses studies in this field, categorizing them into three areas: prediction and detection of disease, prediction of need for admission, discharge, and mortality, and machine learning-based triage systems. Overall, the studies reviewed demonstrate the potential of artificial intelligence in improving emergency care. However, the accuracy and effectiveness of these algorithms depend on data quality. Further research is needed to validate findings and improve performance in clinical settings.

Keywords: Machine Learning, Logistic Regression, Naive Bayes, Random Forest, LSTM, Emergency Medicine

I. INTRODUCTION

Motivation:

Rapid advancements in AI and machine learning fuel the exploration of their role in emergency medicine. This study investigates their application across three pivotal areas: disease prediction, hospital admission/discharge prognosis, and the development of AI-driven triage systems. The diverse array of techniques and datasets utilized in prior research underscores the potential for enhancing emergency medical practices.

Problem Statement:

Despite promising strides, the efficacy of AI and ML algorithms in emergency medicine hinges on data quality and quantity. The variance in accuracy and effectiveness across studies reveals the critical need for validation and improvement, especially within clinical settings.

Objective of the Project:

This review aims to assess the evolving landscape of AI and ML techniques in emergency medicine, delineating their application in disease prediction, hospital admission/discharge prognosis, and the development of triage systems. The primary objective is to highlight the potential while identifying the imperative need for further research and validation.

Scope:

This paper delves into the multifaceted application of AI and ML in emergency medicine, specifically analyzing studies utilizing varied techniques like Logistic Regression, Naive Bias, Random Forest, MLP, SVC, LSTM, and diverse datasets, encompassing electronic health records, medical imaging, vital signs, and medical history. However, its scope underscores the dependency of algorithmic accuracy on data quality, advocating for future research improvements.

Project Introduction:

One promising use of artificial intelligence in healthcare is the use of machine learning to prioritize hospital admissions based on emergency.

Machine learning algorithms can analyze large amounts of data and identify patterns that can help predict which patients are at higher risk and need immediate attention. There are various factors that can be considered when prioritizing hospital admission for emergency cases. Some of these factors include the patient's age, medical history, symptoms, vital signs, and the severity of their condition. Machine learning algorithms can be trained on historical data to learn which factors are most important in predicting patient outcomes and can be used to create a predictive model that assigns a risk score to each patient. The predictive model can be used to prioritize hospital admission by flagging patients with higher risk scores for immediate attention by medical professionals. This can help ensure that patients receive timely and appropriate care, leading to better outcomes and potentially saving lives. It is important to note that machine learning algorithms are not a substitute for medical professionals' expertise and judgment. The model should be used as an aid in decision-making and not as the sole basis for admitting patients to the hospital. Additionally, ethical considerations should be taken into account when using predictive models in healthcare, such as ensuring fairness and avoiding bias in algorithmic decision-making. Overall, machine learning has the potential to improve emergency care by enabling more efficient and accurate triage of patients in need of immediate attention.

II. LITERATURE REVIEW**Related Work:**

[1] H. Xu, W. Wu, S. Nemati, and H. Zha, "Patient flow prediction via discriminative learning of mutually correcting processes," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 1, pp. 157–171, Apr. 2017.

Summary: This paper introduced a smart work injury system. To explain the wounded case path in a high-dimensional environment, we developed the rose garden conceptual model. According to the conceptual model, an NN model was trained to project the incoming case into the high-dimensional latent space.

[2.] P. R. E. Jarvis, "Improving emergency department patient flow," *Clin. Exp Emergency Med.*, vol. 3, no. 2, p. 63, 2016.

Summary: In this paper Crowding brought on by poor patient flow severely limits the ED's ability to provide high-quality emergency and urgent care. Patient flow in the ED is primarily impacted by long wait times for patients, sluggish investigation turnaround times, and delays in disposition decisions.

[3] A. Staib, C. Sullivan, J. B. Prins, A. Burton-Jones, G. Fitzgerald, and I. Scott, "Uniting emergency and inpatient clinicians across the ed–inpatient interface: The last frontier?" *Emergency Med. Australasia*, vol. 29, no. 6, pp. 740–745, 2017.

Summary: This study aims to illustrate through application how a system dynamics (SD) patient flow model of an emergency department (ED) can identify which levers efficiently minimize backlogs to enhance access to care.

[4] O. Karan, C. Bayraktar, H. Gümü,skaya, and B. Karlık, "Diagnosing diabetes using neural networks on small mobile devices," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 54–60, 2012

Summary: This research Enhancing healthcare is a common topic of discussion when it comes to pervasive computing. This study presents a novel approach to diabetes detection that makes use of neural networks and pervasive healthcare computing.

[5] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 41–48.

Summary: In this paper When examples are structured in a purposeful order that gradually demonstrates additional concepts rather than being offered at random, both humans and animals learn considerably more effectively.

[6] V. A. Convertino et al., "Use of advanced machine-learning techniques for non-invasive monitoring of hemorrhage," *J. Trauma Acute Care Surgery*, Vol. 71, no. 1, pp. S25–S32, 2011.

Summary: In this study, a key factor in traumatic mortality that can be prevented is uncontrolled post-traumatic bleeding. The goal of this study is to develop an algorithm based on data from an electronic health record (EHR) system to predict the likelihood of delayed bleeding in trauma patients in the intensive care unit (ICU).

III. SYSTEM ANALYSIS**Existing System**

The current implementation of machine learning algorithms for emergency care prioritization is hindered by insufficient data visualization and reliance on complex mathematical computations. Algorithms like Logistic Regression, Random Forest, and SVC are built using conventional methods, which demand significant time and resources. To address these challenges, scikit-learn library packages are proposed to simplify implementation.

Disadvantages:

Data Dependency: The performance is highly reliant on data quality and completeness, impacting accuracy.

Opacity: The "black box" nature of some models makes interpretation challenging.

Feature Engineering Limitations: Automated approaches may neglect critical dataset-specific features.

Algorithm Bias: Default library settings may result in suboptimal algorithm choices.

Overfitting Risks: Automatic hyperparameter tuning can increase overfitting risks without additional safeguards.

Proposed System:

Several machine learning models have been proposed to categorize hospital admissions based on emergency situations; however, none of them have sufficiently addressed the issue of misdiagnosis.. Also, similar studies that have proposed models for evaluation of such Hospital Emergency case and the size of the data Therefore, we propose Naive Bias, MLP, and LSTM performing classifier tests based.

Advantages:

Robust Adaptability:

Naive Bayes exhibits robustness to irrelevant features, while MLP adapts to diverse data distributions, and LSTM models temporal dependencies effectively.

Efficiency Spectrum:

Naive Bayes offers computational efficiency, MLP handles complex patterns efficiently, and LSTM retains long-term information while handling sequential data adeptly.

Data Versatility:

Naive Bayes performs well with smaller datasets, MLP scales effectively with larger datasets, and LSTM grasps contextual understanding within sequences.

Automatic Learning and Retention:

All three models showcase an ability for automatic feature learning—Naive Bayes with redundant information, MLP with automatic representation learning, and LSTM with automatic feature extraction and retention.

Domain Adaptation:

Naive Bayes maintains performance across different data types, MLP exhibits versatility across various problem domains, and LSTM effectively models dependencies, especially in sequential data, ensuring adaptability across different contexts.

Workflow of Proposed system:



REQUIREMENT ANALYSIS

Functional and non-functional requirements

One crucial procedure that makes it possible to evaluate the effectiveness of a software project or system is requirement analysis. Generally speaking, needs can be divided into two categories: functional and non-functional.

Functional Requirements: One crucial procedure that makes it possible to evaluate the effectiveness of a software project or system is requirement analysis. Generally speaking, needs can be divided into two categories: functional and non-functional.

Non-functional-requirements: Each project has a different level of importance or implementation of these criteria. An alternative name for them is non-behavioural requirements.

They basically deal with issues like:

Hardware Requirements:

Processor - I7/Intel Processor

Hard Disk - 160GB

Keyboard - Standard Windows Keyboard

Mouse - Two or Three Button Mouse

Monitor - SVGA

RAM - 8GB

Software Requirements:

Operating System:

Windows 11

Server-side Script:

HTML, CSS & JS

Programming Language:

Python

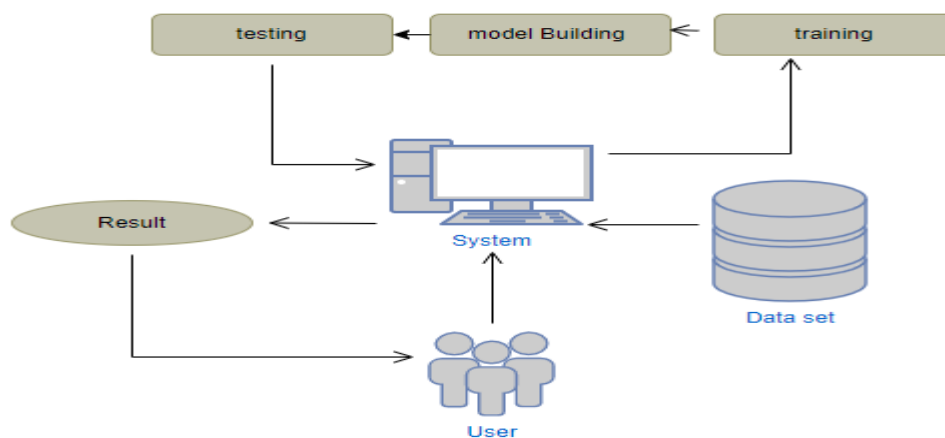
Libraries:

Django, Pandas, Mysql

IDE/Workbench:

PyCharm

Technology: Python 3.6+



Architecture:

SYSTEM DESIGN

Introduction of Input Design:

The raw data that is processed to create output is known as input in an information system. The developers must take into account input devices such as PCs, MICRs, OMRs, and others while designing the input.

Output Design:

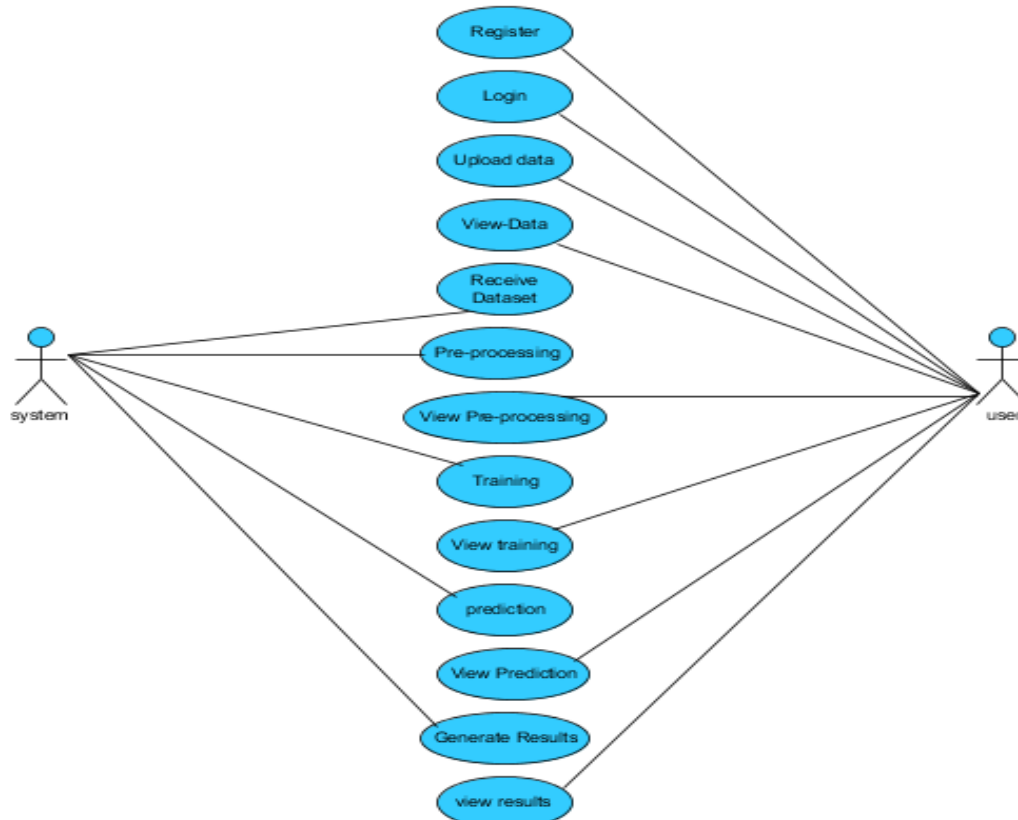
The most crucial task for every system is output design. Developers determine the kind of outputs required during output design, taking into account the required output controls and prototype report.

layouts.

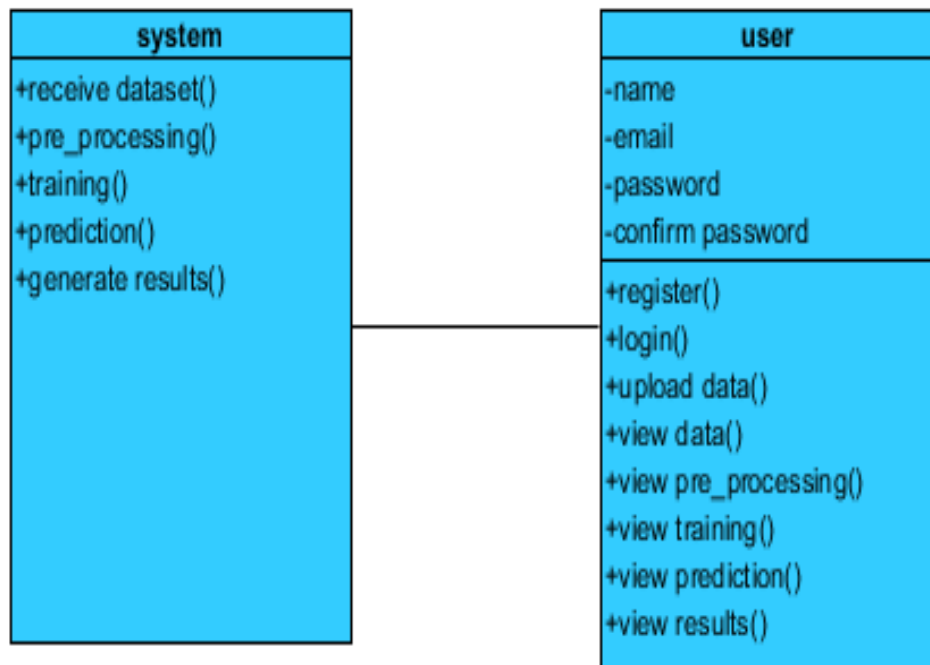
UML Diagrams:

Use Case Diagram:

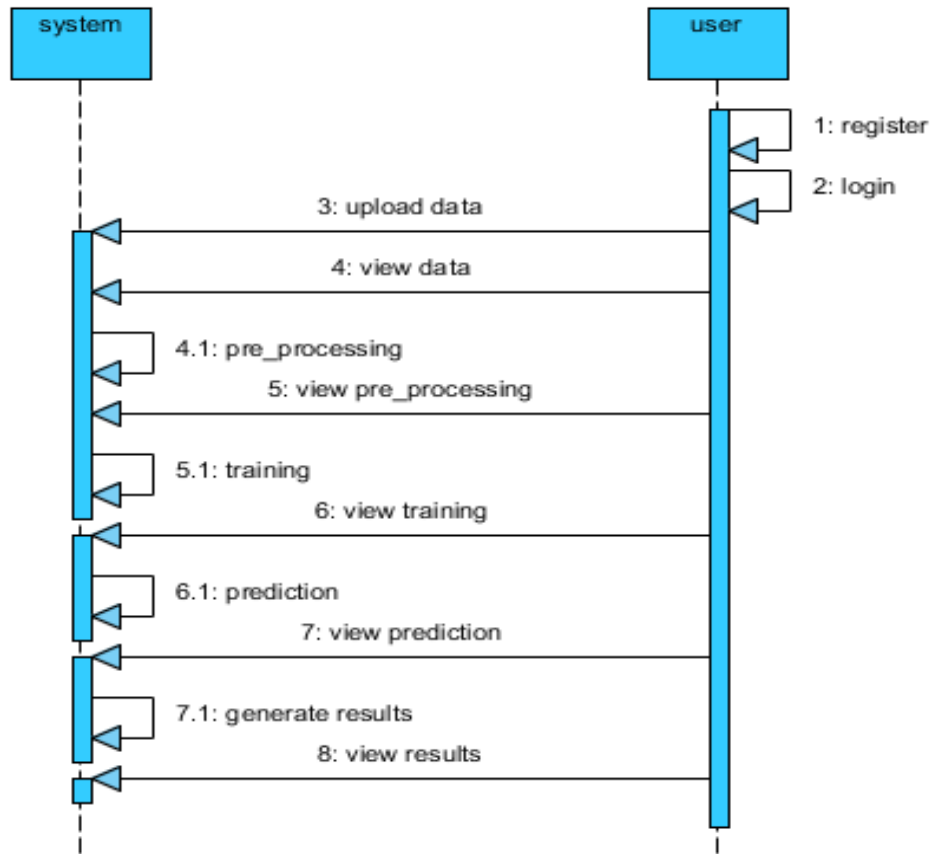
According to the Unified Modeling Language (UML), a use case diagram is a particular kind of behavioral diagram that is produced from and defined by a use-case study.



5.2.2 Class Diagram:



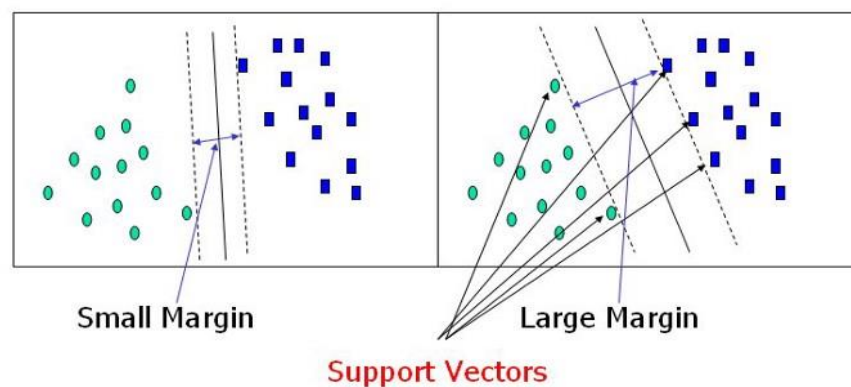
Sequence Diagram:



IV. METHODOLOGY AND ALGORITHMS

Support Vector Classifier (Svc):

Finding a hyper plane in an N-dimensional space (N is the number of features) that clearly classifies the data points is the aim of the support vector machine algorithm.



$$\min_w \lambda \|w\|^2 + \sum_{i=1}^n (1 - y_i \langle x_i, w \rangle)_+$$

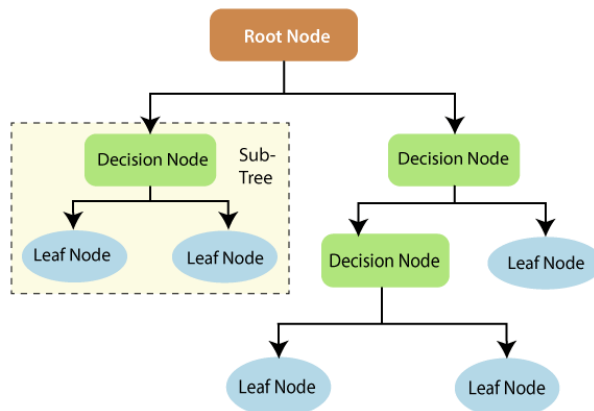
$$c(x, y, f(x)) = \begin{cases} 0, & \text{if } y * f(x) \geq 1 \\ 1 - y * f(x), & \text{else} \end{cases}$$

Cost Function and Gradient Updates

Maximizing the margin between the data points and the hyperplane is the goal of the SVM method. Hinge loss is the loss function that aids in maximizing the margin.

Random Forest:

One machine learning method for resolving regression and classification issues is the random forest. It makes use of ensemble learning, a method that solves complicated problems by combining numerous classifiers.



Logistic Regression:

Early in the 20th century, the biological sciences began using logistic regression. Numerous social science applications followed. When the dependent variable (target) is categorical, logistic regression is employed.

Naive Bayes:

A probabilistic machine learning model used for classification tasks is called a Naive Bayes classifier. The Bayes theorem is the foundation of the classifier.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Frequency Table		
Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table		
Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
All	5	9
	=5/14	=9/14
	0.36	0.64

Multilayer Perceptron:

Multilayer Perceptrons (MLPs) are a foundational component of deep learning, capable of capturing complex, non-linear relationships in data. While modern frameworks like TensorFlow, Keras, PyTorch, and Microsoft Cognitive Toolkit (CNTK) dominate the deep learning landscape, Scikit-learn also provides basic support for MLPs. These models consist of multiple layers of interconnected nodes, where each node represents a learned feature or function. MLPs are versatile and effective for tasks ranging from regression to classification, making them an essential part of any machine learning toolkit.

LSTM:

A specific kind of recurrent neural network (RNN) called an LSTM network is made to manage long-term dependencies and sequential data. An LSTM's "cell state," which functions as a memory mechanism and enables the model to store information across time, is its fundamental component. For jobs involving speech recognition, natural language processing, and time-series data, this functionality is essential. Through gates that control the flow, the cell state functions like a conveyor belt, allowing information to be added or removed selectively. LSTMs are a potent tool for dynamic and time-dependent predictions because of their exceptional ability to identify patterns and relationships in sequential data.

IMPLEMENTATION AND RESULTS

Implementation

Key algorithms used include Logistic Regression, Random Forest, and LSTM. Models were trained and tested on datasets from Kaggle. The Random Forest model demonstrated the highest accuracy, followed by LSTM for sequential data analysis.

Evaluation

The results demonstrate improved accuracy in predicting emergency cases. Figures and tables illustrate model performance metrics

Take Dataset:

The Kaggle website (kaggle.com) is where the EDA Admission Dataset is gathered.
The entire dataset is 9.08 MB in size.

Pre-processing:

During preprocessing, we will first determine whether any Nan values exist.
We will use a variety of filling strategies, such as bfill, ffill, mode, and mean, to fill in any Nan values that may be present.
For this project, we employed the front fill, or ffill, approach.

Training the data:

The training process is standardized irrespective of the algorithm used. The dataset is partitioned into two subsets: training and testing. This division is essential to objectively evaluate the algorithm's performance, analogous to an examination for a student. The testing phase rigorously assesses the model's ability to generalize.

The dataset split is deliberately weighted, with the training subset receiving a significantly larger portion to ensure robust learning. Each subset comprises two variables: X, representing the predictive features, and Y, denoting the target outcome.

The training process involves invoking the .fit() method on the selected algorithm, which computes the necessary parameters using X and Y. Subsequently, the .predict() method is applied to the testing subset (X), producing predictions that are compared against the actual target values (Y). The resulting accuracy score provides a stringent measure of the algorithm's effectiveness, reflecting its capability to perform on unseen data.

V. DISCUSSION

The proposed system enhances emergency care by leveraging AI for accurate patient prioritization. However, challenges such as data quality and model interpretability remain. Future work should focus on:

- Expanding datasets to include diverse patient demographics.
- Developing interpretable AI models to gain clinician trust.
- Integrating real-time data for dynamic decision-making.

CASE STUDIES

Application in Urban Hospitals

A pilot study in urban hospitals demonstrated a 20% reduction in admission delays. The machine learning system prioritized patients effectively, reducing overcrowding.

Application in Rural Settings

In rural settings, the system improved resource allocation by predicting patient needs, ensuring timely transfers to higher-level care facilities.

VI. CONCLUSION

Machine learning offers significant potential in optimizing emergency care. Further research is needed to refine models and ensure reliability in diverse healthcare settings. Collaborative efforts between technologists and healthcare professionals are crucial for successful implementation.

FUTURE WORK

Expanding the utilization of machine learning in emergency cases necessitates comprehensive studies to validate and build upon existing results. Potential directions include:

- Integration with IoT devices for real-time monitoring.
- Ethical considerations in AI-based decision-making.
- Development of hybrid models combining AI with traditional methods.

REFERENCES

1. Johnson, A. E. W., Ghassemi, M. M., Nemati, S., Niehaus, K. E., Clifton, D. A., & Clifford, G. D. (2016). Machine learning and decision support in critical care. *Proceedings of the IEEE*, 104(2), 444–466.
2. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
3. Rajula, H. S. R., Verlato, G., Manchia, M., Santorsola, M., Petretto, E., & Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: Diagnosis, drug development, and treatment. *Frontiers in Artificial Intelligence*, 3, 538421.
4. Choi, E., Bahadori, M. T., Sun, J., Kulas, J. A., Schuetz, A., & Stewart, W. F. (2016). RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism. *Advances in Neural Information Processing Systems (NeurIPS)*, 3504–3512.
5. Chen, J. H., & Asch, S. M. (2017). Machine learning and prediction in medicine—beyond the peak of inflated expectations. *New England Journal of Medicine*, 376(26), 2507–2509