JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EMERGENCY DEPARTMENT WAITING TIME ANALYSIS USING MACHINE LEARNING AND SIMULATION

[1]Randa Abdulhadi Al Mansour, [2]Ghala Mohammed Al Mansour, [3] Ahmad M. Alkheder Almasabi
[1]B.Sc Student, [2]B.Sc Student, [3]Assistant Professor
[1]443300687@nu.edu.sa, [2]443303614@nu.edu.sa, [3]amalkheder@nu.edu.sa
Department of Computer Science,
Najran University, Najran, Saudi Arabia

Abstract : This project presents a combined machine learning and simulation approach to predicting and analyzing patient waiting times in the Emergency Department (ED). A Random Forest classifier was trained using a synthetic dataset of 500 patient encounters to classify waiting time into short-wait and long-wait categories. The model achieved high performance, with an accuracy of 95.33% and an F1-score of 96.26%. In the second phase, a discrete-event simulation was implemented using the SimPy environment to model patient arrivals, queuing behavior, and doctor resource utilization in the ED. The simulation produced operational indicators such as average waiting time and total number of patients served. Together, the machine learning model and the simulation provide a comprehensive understanding of ED performance and support decisions related to patient flow and resource management.

***IndexTerms* - Emergency Department, Waiting Time Prediction, Machine Learning, Random Forest, Simulation, SimPy, Healthcare Analytics**

I. INTRODUCTION

Emergency Departments (EDs) face significant challenges related to overcrowding, unpredictable patient arrivals, and limited medical staff. These factors contribute to increased waiting times, which can affect patient satisfaction and clinical outcomes. Predicting waiting time categories before the patient enters the service pipeline can help improve resource allocation and identify potential bottlenecks.

Machine learning (ML) provides a powerful method for predicting waiting time categories based on patient characteristics and ED conditions at arrival. In addition, discrete-event simulation offers a dynamic way to model patient flow and assess operational performance under varying resource configurations.

II. RESEARCH METHODOLOGY

This section describes the methodology used to develop and evaluate the machine learning model and the simulation environment. It includes dataset preparation, model selection, training process, evaluation metrics, and the SimPy simulation model design.

2.1 Dataset Description

A synthetic dataset of 500 Emergency Department (ED) patient records was generated to train the machine learning model. The dataset included the following features:

- Arrival Hour
- Triage Level
- Patient Age
- Number of Doctors on Shift
- Number of Nurses
- Crowd Level
- Severity Score

The target variable was a waiting time category, classified into:

- 0 = Short Wait
- 1 = Long Wait

The dataset was split into:

- 70% for training
- 30% for testing

2.2 Machine Learning Method

A Random Forest Classifier was selected due to its high stability, ability to handle mixed data, and strong performance with non-linear patterns.

Steps followed:

1. Data preprocessing (encoding + splitting).
2. Training the Random Forest model with default parameters.
3. Model evaluation using:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

The model achieved:

- Accuracy: 95.33%
- F1-score: 96.26%

2.3 Simulation Method (SimPy)

A discrete-event simulation was developed using SimPy to model the operational behavior of the ED, including:

- Random patient arrivals
- Queue formation
- Assignment to available doctors
- Service time per patient

Simulation Parameters:

- Number of doctors: 3
- Average inter-arrival time: Exponential distribution
- Service duration: Uniform distribution
- Total simulation time: 120 minutes

The simulation produced:

- Total number of patients served
- Average waiting time
- Maximum waiting time

This simulation complements the machine learning model by showing real system performance under dynamic conditions.

III. RESULTS AND DISCUSSION

3.1 Machine Learning Results

The machine learning model demonstrated strong and reliable performance in classifying Emergency Department (ED) waiting time into short-wait and long-wait categories. The confusion matrix obtained from the test dataset is presented in Figure 1.

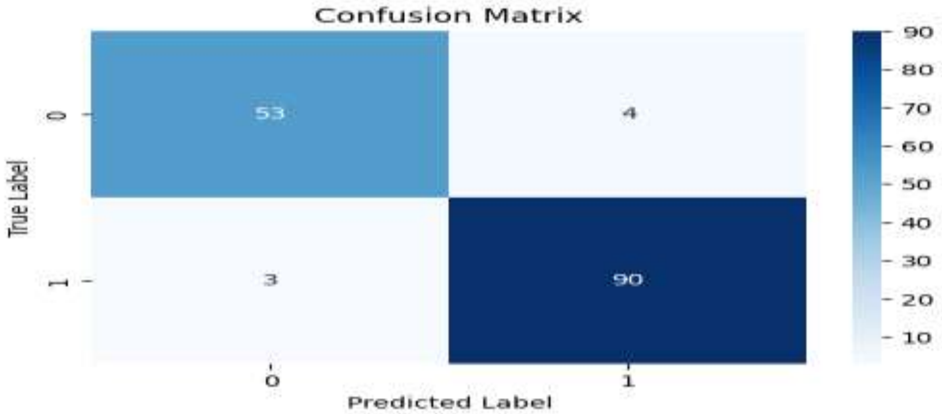

Figure 1: Confusion Matrix

The confusion matrix shows the following numerical outcomes:

- 53 short-wait cases were correctly classified.
- 90 long-wait cases were correctly classified.
- 4 short-wait cases were misclassified as long-wait.
- 3 long-wait cases were misclassified as short-wait.

The classification performance metrics are shown in Figure 2, with the exact computed values listed below:

- Accuracy: 95.33%
- Precision: 95.74%
- Recall: 96.77%
- F1-score: 96.26%

These results indicate that the Random Forest model achieved high accuracy and consistently strong performance in both classes, demonstrating its suitability for ED waiting time prediction.

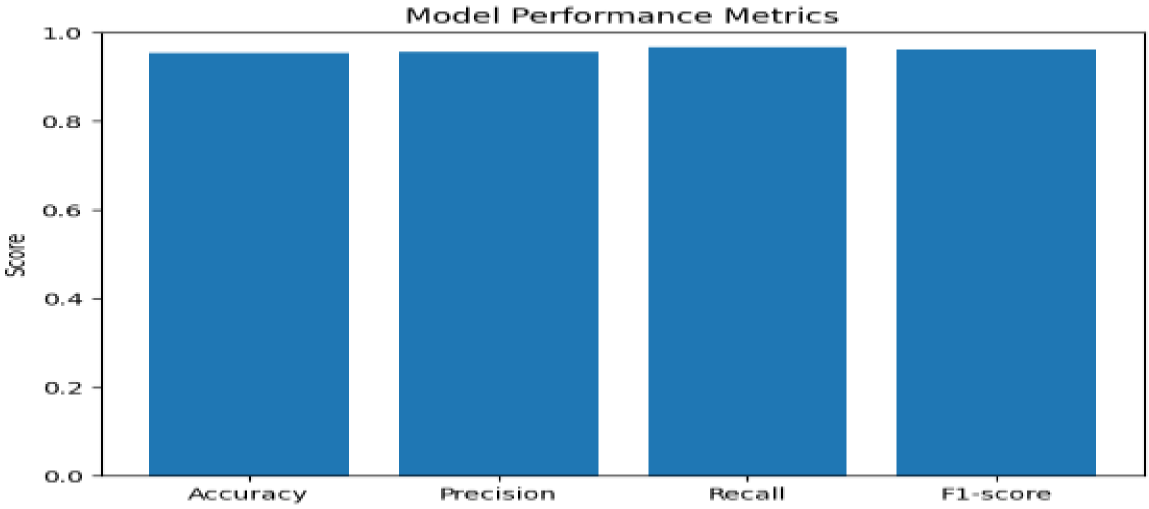

Figure 2: Performance Metrics (Accuracy, Precision, Recall, F1)

3.2 Simulation Results

A SimPy-based discrete event simulation was used to represent patient arrivals, waiting times, and resource utilization in the Emergency Department. The simulation results are summarized in Figure 3, and the numerical outcomes are listed below.

- Total patients simulated: 44
- Average waiting time: 1.35 minutes
- Maximum waiting time: 14.36 minutes

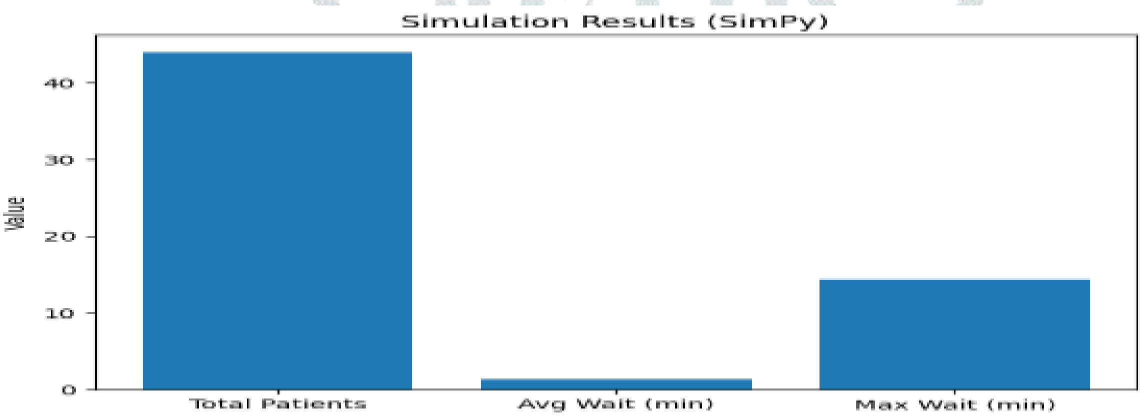

Figure 3: Simulation Output Summary (Total Patients, Avg Wait, Max Wait)

IV. CONCLUSION

This study presented a combined machine learning and discrete-event simulation approach to analyzing and predicting waiting times in the Emergency Department (ED). The Random Forest model demonstrated strong predictive performance, achieving 95.33% accuracy, 95.74% precision, 96.77% recall, and an F1-score of 96.26%, indicating its reliability in classifying short-wait and long-wait scenarios. In addition, the SimPy simulation model provided valuable insights into ED operational behavior, showing a total of 44 patients, with an average waiting time of 1.35 minutes and a maximum waiting time of 14.36 minutes under the given resource configuration.

Together, the machine learning predictions and simulation outputs offer a comprehensive framework that can support ED decision-making, improve resource allocation, and enhance patient flow management. Future work may include using real hospital data, expanding feature sets, or testing alternative machine learning models to further improve performance and practical applicability..

V. ACKNOWLEDGMENT

The authors would like to express their appreciation to Najran University for providing the academic environment and support that contributed to the completion of this research work.

REFERENCES

[1] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
[2] Banks, J., Carson, J., Nelson, B., & Nicol, D. (2010). *Discrete-Event System Simulation*. Prentice Hall.
[3] Rossetti, M. (2016). *Simulation Modeling and Arena*. Wiley.
[4] SimPy Documentation: https://simpy.readthedocs.io/