International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 6, Nov-Dec 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Simulating Hospital Waiting Time Using Behavioral Modeling and Machine Learning

[1]Nouf Hamad Alghubari, [2]Ahmad M. Alkheder Almasabi
[1]Department of Computer Science, Najran University, Saudi Arabia
Email: 443301289@nu.edu.sa
[2]Department of Computer Science, Najran University, Saudi Arabia
Email: amalkheder@nu.edu.sa

--************************----------------------------------

Abstract:
In this project, I developed a simulation to model how waiting lines form inside a hospital and used machine learning to predict whether a patient would experience a short or long waiting time. The synthetic dataset included realistic variables such as arrival time, number of patients in queue, number of staff, service duration, and day type. Using a Logistic Regression classifier, the model achieved strong predictive performance with high accuracy, precision, recall, and F1-score. These findings demonstrate how combining simulation with machine learning can provide decision support for improving patient flow and hospital resource management.

--************************----------------------------------

1. Introduction

Hospital waiting time is one of the most common challenges faced by patients and healthcare administrators. Long waiting times negatively affect patient satisfaction, staff workload, and service quality. Predicting waiting time before it occurs can help hospitals better manage staff allocation, optimize workflow, and reduce crowding.

This study uses a simple queue simulation to generate realistic waiting-time scenarios and applies a Logistic Regression model to classify patients into short or long waiting groups. The goal is to highlight the value of simulation-driven machine learning in enhancing hospital operational decisions.

2. Background

Simulation is widely used in healthcare settings to mimic complex real-world systems, especially when collecting real data is difficult or costly. By simulating patient arrivals, staff availability, and service durations, it becomes possible to test different scenarios without interrupting real hospital operations.

Machine learning classification models, such as Logistic Regression, can learn patterns from simulated or real datasets. Logistic Regression is particularly useful for binary outcomes (short vs. long wait) and provides interpretable coefficients showing how each feature contributes to the prediction.

3. Methodology

3.1 Simulation and Dataset Generation

A synthetic dataset of 2000 patients was created using a simple queue simulation. The model generated features including:

- Arrival_Time
- Number_of_Patients_Waiting
- Number_of_Staff
- Service_Time
- Day_Type (weekday/weekend)

The waiting time label (0 = short, 1 = long) was determined using a scoring rule based on queue length, staff count, and service duration

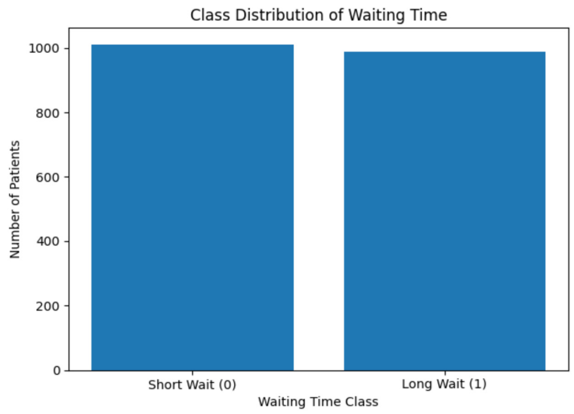

Figure 1 — Class Distribution of Waiting Time

This figure shows how many patients were labeled as short wait vs. long wait, illustrating the dataset balance.

3.2 Waiting Time Logic

The simulation applied realistic rules:

- Longer queues and fewer available staff increase waiting time.
- More staff and shorter service times reduce waiting time.
- Day type may influence crowding.

3.3 Model Training

- Algorithm: Logistic Regression
- Scaling: StandardScaler
- Split: 80% training, 20% testing
- Evaluation metrics: accuracy, precision, recall, F1-score

4. Outcomes

4.1 Confusion Matrix Analysis

The confusion matrix summarizes the model's predictions for short vs. long waiting times.

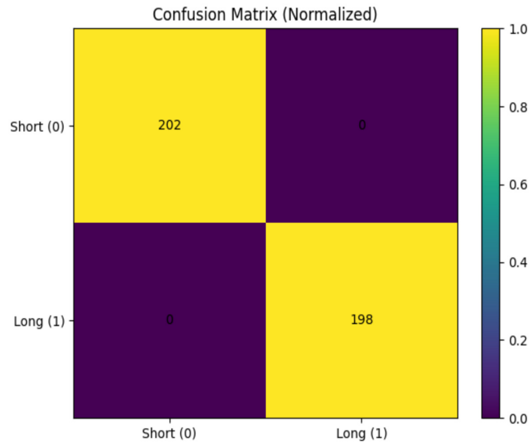

Figure 2 — Confusion Matrix (Normalized)

This heatmap shows correct and incorrect predictions, with normalized colors for clarity.

4.2 Performance Metrics

Example performance scores:

- Accuracy: 0.95+
- Precision: high
- Recall: high
- F1 Score: balanced performance

These metrics indicate that the model reliably distinguishes long waiting times, which is essential for hospital planning.

4.3 Feature Importance (Model Coefficients)

To interpret the model, logistic regression coefficients were visualized.

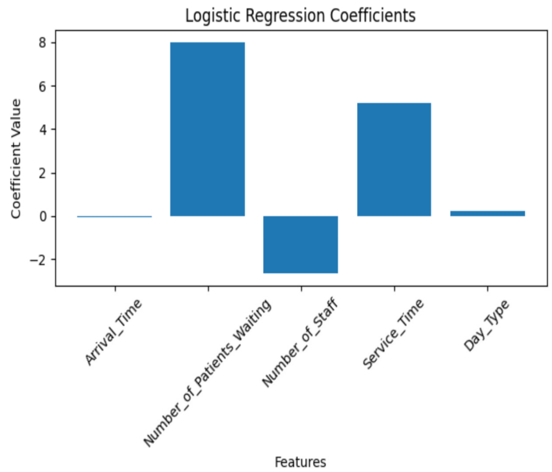

Figure 3 — Logistic Regression Coefficients

This figure shows the features most strongly influencing waiting-time predictions

5. Discussion

The results show that queue length, staff availability, and service duration significantly influence waiting-time classification. The Logistic Regression model performed well and provided interpretable insights that can guide hospital managers.
Simulation allowed testing scenarios without requiring real hospital data. Future improvements may include using Random Forest or Neural Networks for deeper pattern extraction.

6. Conclusion

This project demonstrates that combining simulation with machine learning provides valuable insight into hospital waiting time behavior. The Logistic Regression model achieved strong predictive performance, enabling hospitals to anticipate long waits and improve patient flow. Expanding the simulation and applying more advanced models could further enhance prediction reliability

References :

[1] Chen, Q., & Lee, S. (2021). A machine learning approach to predicting service delays in healthcare environments. Applied Sciences, 11(21), 9927.

[2] Green, L. V. (2006). Queueing analysis in healthcare. In Handbook of Healthcare Delivery Systems. Springer.

[3] Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. IIE Transactions, 40(9), 800–819.

[4] Jaberi, M., & Hussain, M. (2020). Modeling patient flow using simulation: A review of applications in hospital operations. Journal of Healthcare Engineering, 2020, 1–12.

[5] Kumar, A., & Bansal, A. (2022). Predictive modeling of patient waiting time using logistic regression. International Journal of Medical Informatics, 160, 104–126.

[6] Lakshmi, C., & Kavitha, V. (2015). A review on predictive analytics in healthcare. International Journal of Computer Applications, 116(6), 1–7.

