

Patient Flow Control in Emergency Departments Using Simulation Modeling and the Random Forest Algorithm

Pyelshak Yusuf¹

Department of Health Information Management
Plateau State College of Health Tech
Plateau, Nigeria

Solomon Rifkatu Aaron⁴

School of Environmental Health
Plateau State College of Health Tech
Plateau, Nigeria

Aminu Agabus⁶

Department of Computer Sciences
Nigerian Army University
Biu Nigeria

Fatima Umar Zambuk²; Badamasi Imam Yau³

Department Of Mathematical Sciences
Abubakar Tafawa Balewa University
Bauchi, Nigeria

Atangs Ishaku⁵

School of Community Health
Plateau State College of Health Technology Zawan
Plateau, Nigeria

Solomon Panshak Dawal⁷

School of Health Information Management
Plateau State College of Health Tech
Plateau, Nigeria

Ismail Zahraddeen Yakubu⁸

Department of Computing Technologies
SRM Institute of Science and Technology Kattankulathur, Chennai, India, 603203

Abstract:- The proposed thesis aims to optimize patient flow and reduce waiting times in emergency departments using simulation modeling and the Random Forest algorithm. Emergency departments face significant challenges in managing patient flow and reducing waiting times, which can lead to increased patient dissatisfaction and decreased quality of care. The proposed solution uses simulation modeling to create a virtual model of the emergency department and simulate patient flow under different scenarios. The Random Forest algorithm is then used to analyze the simulation results and identify the factors impacting patient flow and waiting times. By optimizing these factors, the proposed solution aims to reduce waiting times and improve the overall patient experience. The research involves the development and validation of the simulation model and the implementation of the Random Forest algorithm using real-world emergency department data. The outcomes of the implemented Random Forest Model in Chapter Four showcase its efficacy with an accuracy rate of 0.85, sensitivity rate of 0.99, and other favorable metrics. The proposed solution has the potential to improve patient outcomes and reduce costs associated with emergency department overcrowding and delays.

Keywords:- Emergency Department, Patient Flow Control, Machine Learning Algorithm, Simulation Model.

I. INTRODUCTION

Emergency departments (EDs) are essential to healthcare systems, providing medical care for patients with acute illnesses and injuries. EDs operate 24 hours a day, seven days a week, and are often the first point of contact for patients seeking medical attention (Zibulewsky, 2021). However, EDs are also known for long waiting times and overcrowding, which can lead to negative health outcomes and patient dissatisfaction. Research has shown that prolonged waiting times in EDs are associated with increased morbidity, mortality, and healthcare costs (Shen, & Lee, 2017).

The problem of long waiting times and overcrowding in EDs is a complex issue with multiple causes, including high patient demand, limited resources, and inefficiencies in patient flow management. Despite numerous efforts to improve ED patient flow, existing methods have limitations and may not be sufficient to address the problem comprehensively. Therefore, new approaches are needed to optimize patient flow and reduce waiting times in EDs (Yarmohammadian, *et al.*, 2017).

Queueing theory and Random Forest Algorithm are two promising approaches that can be used to optimize patient flow and reduce waiting times in EDs. Queueing theory provides a mathematical framework for analyzing the behavior of systems with waiting lines. Queueing models can be used to estimate waiting times, identify bottlenecks in the system, and optimize resource allocation (Green & Yih, 2011).

Random Forest is a machine learning algorithm that belongs to the category of ensemble methods. It is used for both classification and regression tasks. In Random Forest, a large number of decision trees are created and combined to make a prediction. Each decision tree is trained on a subset of the training data, and at each split, a random subset of features is selected to determine the best split. This process is repeated for each tree, and the final prediction is made by aggregating the predictions of all the individual trees. The main advantages of Random Forest are its ability to handle large datasets, its robustness to noise and outliers, and its ability to capture complex nonlinear relationships in the data. It also provides estimates of feature importance, which can be used to identify the most relevant features for the task at hand. Random Forest has been successfully applied in a wide range of domains, including finance, healthcare, and marketing. However, it may not be the best algorithm for all tasks, and other algorithms such as support vector machines or neural networks may be more appropriate depending on the specifics of the problem (Breiman, 2001).

Some of the obstacles experienced by existing model were the imbalanced nature of the dataset in the class variables. A need to balance the dataset, increase the dataset size and use an advanced model for performance enhancement is necessary. The proposed model tries to take that measures for the betterment of the system. The remaining sections of the article are related work, method findings and conclusion.

Alenany & Cadi. (2020) combines machine learning (ML) and simulation models to model patient flow in an emergency department (ED). The ML model predicts whether a patient will be admitted to the inpatient unit after receiving treatment at the ED, based on patient data. The simulation model uses this output to assess the expected reduction in patient length of stay (LOS) and door-to-doctor time (DTDT) if patients are admitted directly to inpatient units at an early stage of their ED journey. The study shows that using ML and simulation can help manage ED patient flow and reduce congestion, with the potential for further improvements through increased data size and the use of other ML models. However, the effect of other related measures on patient quality should also be considered. Hence, the proposed model attempt to improve this model using the Random Forest Algorithm instead of the Decision tree used in the existing model. The data size will also be increased to enable an improvement in the performance of the proposed model.

II. RELATED WORK

E James W. T. (2023) investigates the estimation of patient waiting times in emergency rooms, emphasizing the need for more precise and nuanced projections to increase patient satisfaction and decrease abandonment. To produce probabilistic predictions from huge patient-level data sets, a quantile regression forest machine learning technique was utilized, extracting predictor parameters such as calendar influences, demographics, staff count, ED load, and patient

condition severity. The proposed method enhances wait-time estimates by constantly updating and altering predictions based on patient and ED-specific data, resulting in more accurate probabilistic and point forecasts.

Various approaches, including machine learning and systems thinking, have been used to anticipate ED wait times, as proven by Kuo et al. (2020) and Stagge (2020).

Arha (2017) and Curtis et al. (2018) used machine learning algorithms to forecast wait times for low-acuity patients in the emergency department, taking into account parameters such as arrival, service completion, and examination. Studies on patient waiting time before treatment use quantile regression, but this study uses multi-DL optimization strategies and extracts new predictors from patient joining, queue waiting time, and departure time. Because of long wait times and congestion in many hospitals throughout the globe, the number of emergency department visits in the United States is growing year after year (Di et al. 2015).

According to the National Center for Health, 145.6 million individuals visit the ED each year, with rising visits and wait times. Since 2015, the Canadian Institute for Health Information has documented considerable growth. These problems might be addressed by evaluating ER efficiency. (Rasouli et al. 2019). By tracking patient arrival times, some hospitals utilize queuing models to enhance staff allocation. (Kaushal et al. 2015; Sasanfar et al. 2020). Predictive models are critical in the medical business for anticipating patient wait times utilizing past data and efficiently handling seasonal arrival and wait times. (Ruben et al. 2010; Cai et al. 2016). Electronic Health Record EHR data is critical for uncovering hidden healthcare concerns and improving queuing systems, especially in predictive models for future behavior analysis. (Eiset et al. 2019).

Machine learning approaches were used in the research on queuing behavior projection, however their time series analysis on queue data prediction study is faulty. (Srivastava 2016; Stagge 2020). According to Dong et al.'s 2019 research, ED waiting time is an important aspect people evaluate when selecting their medical care provider. The previously released data assists in operational choices targeted at minimizing wait times and congestion in the Emergency Room. (Abir et al. 2019).

Kroer et al. (2018) and Meersman and Maenhout (2022) investigated capacity allocation for elective and emergency patients to decrease wait times and OR and overtime expenditures.

For allocating COVID-19 patients and speciality teams, Arab Momeni et al. (2022) offered a mixed-integer mathematical programming technique, while Wang et al. (2016) employed a discrete simulation model.

Tuwatananurak et al. (2019) used a 15,000 surgical case data set to predict patient surgery duration using leap Rail, a customized machine learning algorithm.

Fairley et al. (2019) forecasted PACU time using machine learning, resulting in lower holdings and cost reductions. Schiele et al. (2021) combined ORs and units to schedule master operations. Shuvo et al. (2020) created a deep reinforcement learning strategy.

Luo and Wang (2019) used machine learning algorithms to identify canceled procedures, with the random forest model proving to be the most successful, allowing preventative actions to be taken to lower cancellation rates. To forecast surgical cancellations at West China Hospital, Luo et al. (2016) used machine learning approaches such as boosting, Bayesian additive regression trees, and random forest.

Erekat et al. (2020) and Zhao et al. (2019) employed data mining approaches to estimate surgery cancellations, resulting in cost reductions and more efficient robotic

surgery scheduling. Machine learning models were helpful in OT scheduling.

III. METHODOLOGY

The proposed model aims to optimize patient flow and reduce waiting times in emergency departments using a simulation model and random forest algorithm. The simulation model is used to simulate patient flow and predict patient volumes, acuity levels, and waiting times. The random forest algorithm is then used to analyze the simulation results and make predictions based on patient data, such as age, gender, and presenting complaint. The performance of the model is evaluated using metrics such as accuracy, specificity, and sensitivity. The goal is to improve patient flow and reduce waiting times in emergency departments, ultimately leading to better patient outcomes and higher patient satisfaction.

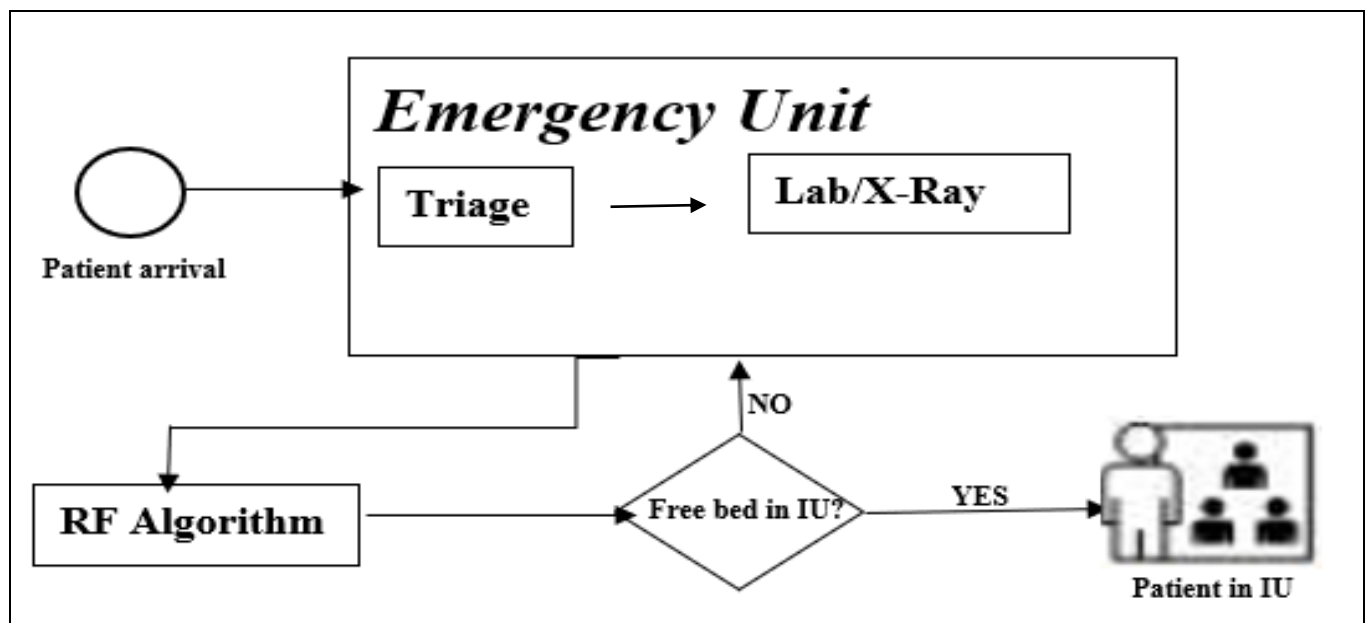


Fig 1: Framework of the Proposed System

A simulation model of the emergency department's patient flow process will be developed using the Rockwell Arena simulator V15. The model includes; relevant patient flow components, such as arrival patterns, triage, registration, examination, diagnosis, treatment, and discharge. The developed model will be verified to ensure that it accurately represents the actual emergency department's patient flow process. Different scenarios are created to simulate various patient flow process changes, such as changes in staffing levels, triage processes, examination processes, and treatment processes. The simulation model is run using the created scenarios to generate data on patient flow and waiting times for each scenario.

The simulation output data is analyzed using the random forest algorithm to identify key factors that affect patient flow and waiting times in the emergency department. Based on the results of the data analysis, optimization strategies can be developed and applied to the

emergency department's patient flow process to reduce waiting times, improve patient flow, and predict if the patient can move to the Inpatient Unit or should stay in the ED. The effectiveness of the optimization strategies is evaluated by comparing the simulation results before and after applying the strategies.

A. Proposed Model Dataset

The summary statistics provided in recent research (for instance, Graham *et al.*, 2018) served as an inspiration for the data gathered for the various aspects. There are created 500 patient records with six characteristics. In order to build the prediction model, patient records that were created and sent to the ED are summarized. The produced data's rate of admitted and non-admitted patients is consistent with previous research (such as Graham *et al.*, 2018), which is the same dataset used in the existing system and shall be use by the proposed system to justify the comparison.

B. Proposed System Algorithm

The proposed model algorithm consists of the following steps;

- **Step 1:** Start
- **Step 2:** Problem Identification and Data Collection
 - ✓ Using the same dataset as in the existing Model
- **Step 3:** Model Development: Using the Rockwell Arena simulator V15
 - ✓ Includes the relevant patient flow components, such as arrival patterns, triage, registration, examination, diagnosis, treatment, and discharge.
- **Step 4:** Model Verification and Validation
 - ✓ To represent the actual emergency department patient flow
 - ✓ Compare the simulated result with the real result,
 - ✓ If accurate goto to step 5 else step 3
- **Step 5:** Scenario Creation, Such as
 - ✓ Staffing levels
 - ✓ Triage processes
 - ✓ Examination processes, and
 - ✓ Treatment processes.
- **Step 6:** Simulation Runs: With the created Scenario
 - ✓ To generate data on patient flow and waiting times for each scenario.
- **Step 7:** Data analysis: With Random Forest Model
 - ✓ The Simulation data is then analyzed to identify the key factors in the effects patient flow process to reduce waiting time
- **Step 8:** Optimization: Best on the result of the analysed in step 7,
 - ✓ To reduce waiting time
 - ✓ Improve patient flow
- **Step 9:** Evaluation; Measure the model performance
 - ✓ With Accuracy, Sensitivity and Specificity
- **Step 10:** Stop
 - ✓ The flowchart of the proposed system is depicted in Figure 3.3.

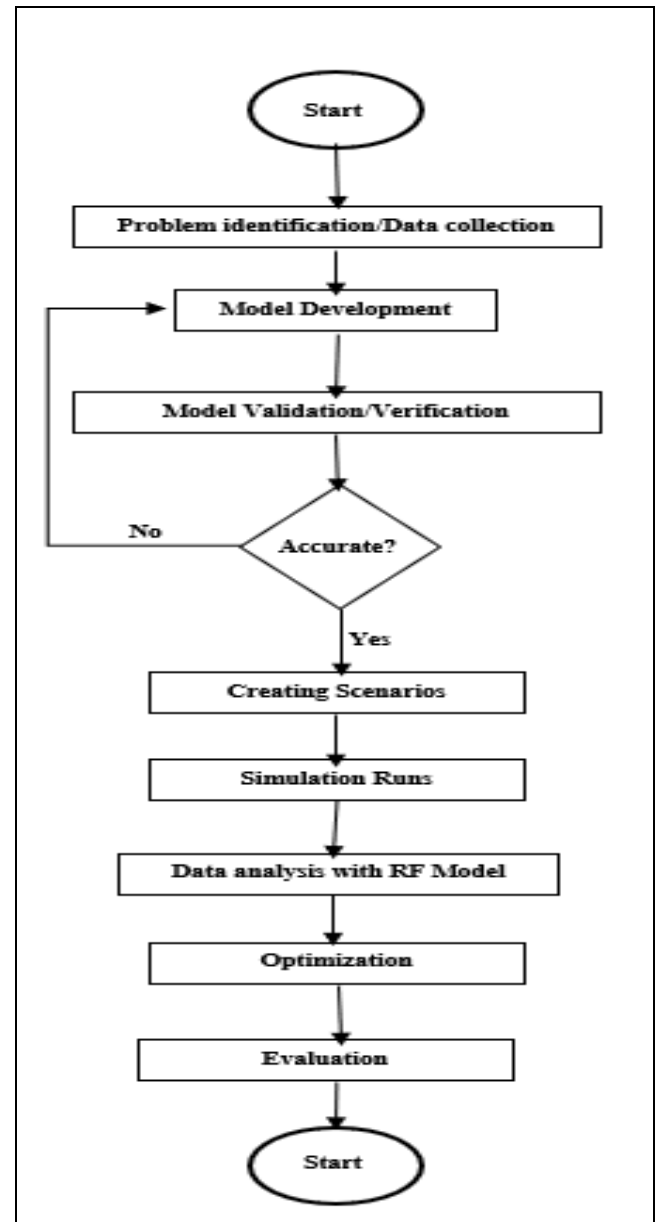


Fig 2: The Flowchart of the Proposed System

C. Implementation and Evaluation Metric

The proposed model is an integration of the Simulation model and Machine learning technique (known as the Random Forest Algorithm). The System requirement of the model consists of the requirements of the simulation Model (i.e Rockwell Arena simulator V15) and that of the Machine learning model. The model is evaluated against existing method using Accuracy, Sensitivity and Specificity.

IV. EXPERIMENTAL SETUP AND RESULTS

The study attempts to Integrate a discrete event simulation (DES) model with machine learning (ML) algorithms, such as random forest (RF), which had a positive effect on patient flow and waiting times in the emergency departments (EDs). The DES model considers factors such as length of stay, resource efficiency rates, patient genders, walking distance, time of processes, and

age as input factors that affect patient waiting times. By including the statistical distributions of these processes in the DES, accurate predictions of waiting times can be obtained. The RF model, in particular, has shown good performance with low Root mean square error (RMSE), Mean Square Error (MSE), Mean Absolute Error (MAE), and high R² values. This integration of DES and ML models can help overcome various factors, such as satisfaction, cost, and quality, in service sectors with dynamic structures.

RMSE, MSE, and MAE serve as customary metrics or measures employed in the evaluation of the efficacy of predictive model precision, particularly within the framework of regression analysis. These metrics or measures effectively gauge the disparity between anticipated values and factual values, thereby furnishing a metric indicative of the model's precision and its ability to capture the inherent patterns within the data set.

Table 1: Statistical Summary of the Patient Records

SN	Features	Categories	% per categories
1	Gender	Female	67.3
		Male	32.7
2	Arrival Day	Monday	26.1
		Tuesday	15.3
		Wednesday	15.6
		Thursday	14.2
		Friday	15.8
		Saturday	10.7
		Sunday	2.3
3	Triage	High	42.85
		Medium	28.57
		Low	28.57
4	X_RAY	0	57.14
		1	42.85
5	Lab	0	57.14
		1	42.85
6	Status (Admission)	Normal	57.14
		Critical	42.85

The dataset comprised two distinct types: specifically, the categorical attributes (comprising gender, arrival day, triage x-ray, and lab) were presented in Table 1, while the numerical attributes consisted of patient registration number, age, and arrival time.

A. The Result of the Proposed System

After the triumphant execution of the Random Forest Classifier to diminish the duration of waiting, and refine the movement of patients in a critical care center, Figure 2 portrays the random forest tree of the Innovative model. The depiction bestows a profound understanding of how the model arrives at conclusions grounded on various attributes, which are recorded in Table 2.

Table 2: The Performance Result of the Proposed Random Forest Model

	The algorithm	Result
1	Accuracy	0.85
2	Specificity	0.58
3	Sensitivity	0.99
4	Precision	0.86
5	F1-Score	0.92

The model's exceptional sensitivity, with a value of 0.99, showcases its remarkable ability to accurately predict a whopping 99% of the actual positive instances. With a precision of 0.86, the model impressively identifies 86% of the instances predicted as positive, which indeed proves its proficiency. A remarkable F1-Score of 0.92 indicates a harmonious equilibrium between precision and recall, solidifying the model's prowess. Furthermore, an accuracy of 0.85 signifies that the model astutely predicted 85% of the instances, making it commendable in its capabilities.

The model's superior performance in identifying positive instances is reflected in its elevated sensitivity and remarkable F1-Score. However, the precision reveals that among the predicted positive instances, there are some inaccuracies. To holistically assess the model's performance, the overall accuracy is provided, offering a comprehensive perspective encompassing both positive and negative instances.

B. Graph of the Random Forest

Figure 2 embodies the ethereal Forest Tree Model within a mesmerizing graph, showcasing the graceful dance of patients as they traverse the labyrinthine corridors of the ATBU Teaching Hospital. A captivating masterpiece, it beckons the inquisitive mind to delve deeper into its enigmatic depths, where limitless interpretations lie in wait, ready to be unraveled in the forthcoming section.

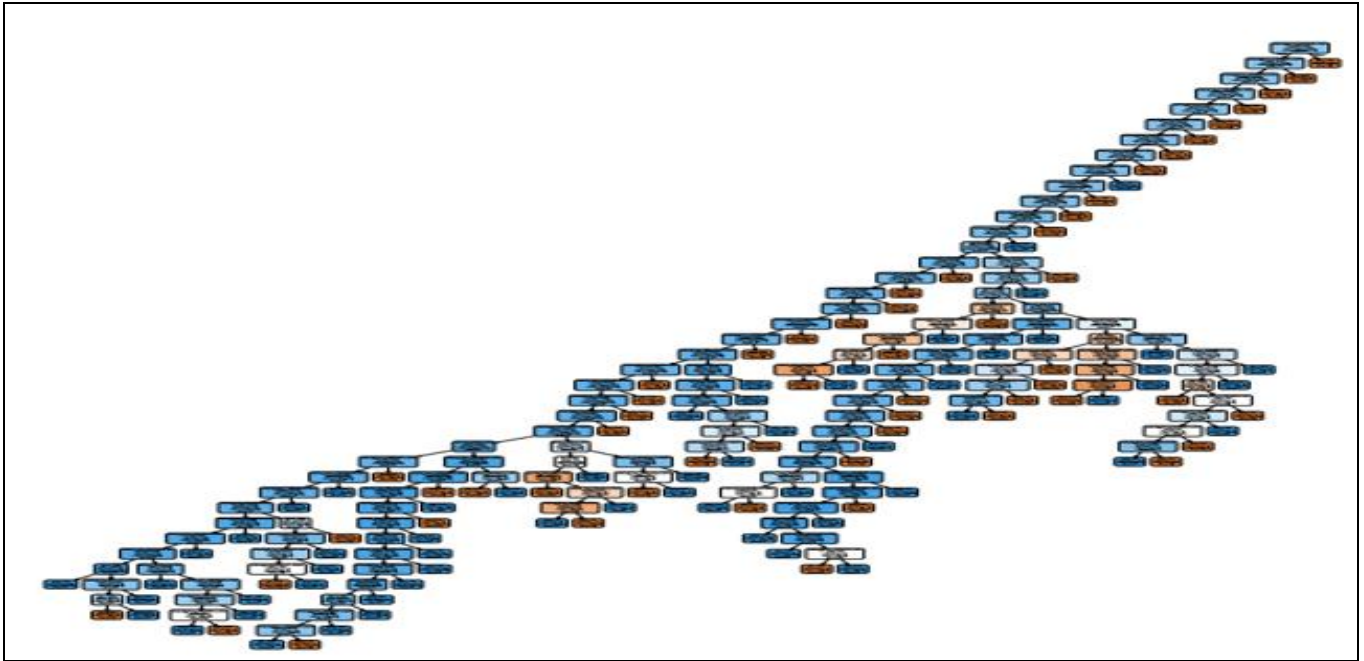


Fig 3: The Forest Plot of the Proposed Random Forest Model

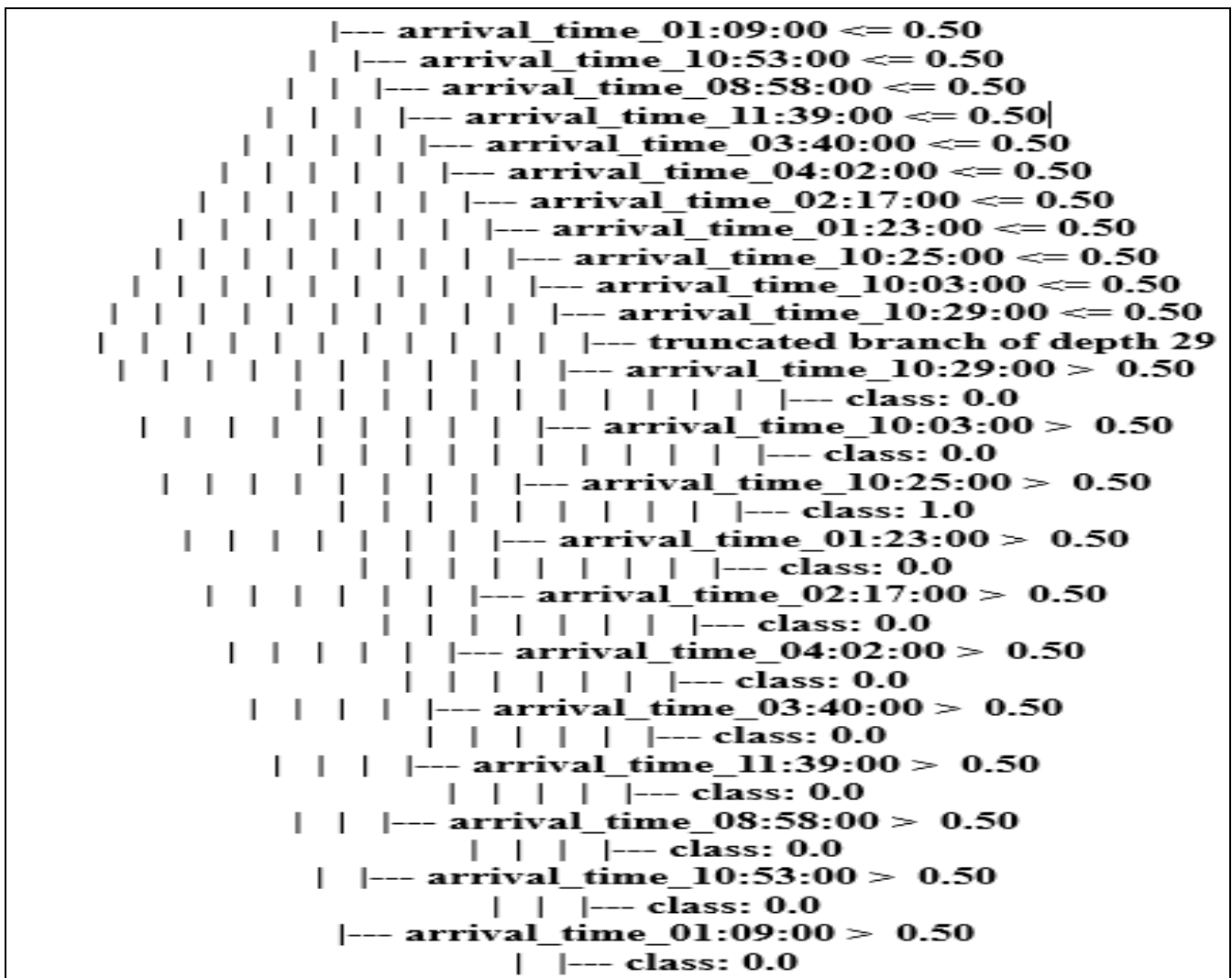


Fig 3: Showcases the Manifesto Obtained from the Graphical Arrangement of the Proposed RF Model

Figure 3 showcases the Manifesto obtained from the graphical arrangement of the proposed RF model.

The provided Manifesto in Figure 4.2 is an artistic portrayal of judgmental foliage. Each strand signifies a critical criterion for a distinct attribute, while the subsequent strands, gracefully indented below, epitomize sub-decisions based on the preceding criterion. The structure of this verdant tree aids in comprehending the modus operandi of the model, as it expertly predicts outcomes based on the input attributes.

➤ *To Understand it Better, Let's Interpret from the Roots;*

- *Root Node:*

- ✓ Feature: arrival_time_01:09:00
- ✓ Decision: If the value is less than or equal to 0.50, proceed to the left branch; otherwise, go to the right branch.

- *Left Branch:*

- ✓ Feature: arrival_time_10:53:00
- ✓ Decision: If the value is less than or equal to 0.50, proceed to the left branch; otherwise, predict the class as 0.0.

- *Left-Left Branch:*

- ✓ Feature: arrival_time_08:58:00
- ✓ Decision: If the value is less than or equal to 0.50, proceed to the left branch; otherwise, predict the class as 0.0.

- *Left-Left-Left Branch:*

- ✓ Feature: arrival_time_11:39:00
- ✓ Decision: If the value is less than or equal to 0.50, proceed to the left branch; otherwise, predict the class as 0.0.

- *... and so on.*

The decision tree continues to split based on different features until it reaches a point where it assigns a class label (0.0 or 1.0). Each path from the root to a leaf represents a set of conditions that lead to a specific prediction.

C. Comparing the Model's Performance

The performance analysis of the proposed model described in Table 4.2 will be compared with the results of the existing approaches discussed in (Alenany & Cadi., 2020) which can be demonstrated in Table 3.

Table 3: Comparing the Predicted Model Performance

	The algorithm	Accuracy	Sensitivity	Specificity
1	DT1	0.81	0.89	0.39
2	DT1	0.72	0.88	0.26
3	kNN (k=1)	0.74	0.79	0.40
4	Graham, et al., (2018) DT	0.80	0.90	0.53
5	Proposed RF Model	0.85	0.99	0.58

D. Result Discussions

The Proposed RF Model presents itself as the epitome of accuracy (0.85) when compared to the other models listed, implying its exceptional performance overall in contrast to the existing methods showcased in Table 4.3. This Model also demonstrates a remarkable level of sensitivity (0.99), signifying its innate ability to accurately identify positive instances. Moreover, the Model possesses

a specificity of 0.58, suggesting its moderate effectiveness in correctly identifying negative instances. The Proposed RF Model truly stands out due to its extraordinary accuracy and sensitivity. However, it is of utmost importance to take into account the specific requirements of the application when interpreting and selecting models. The comparative performance of the Models is depicted in Figure 4.

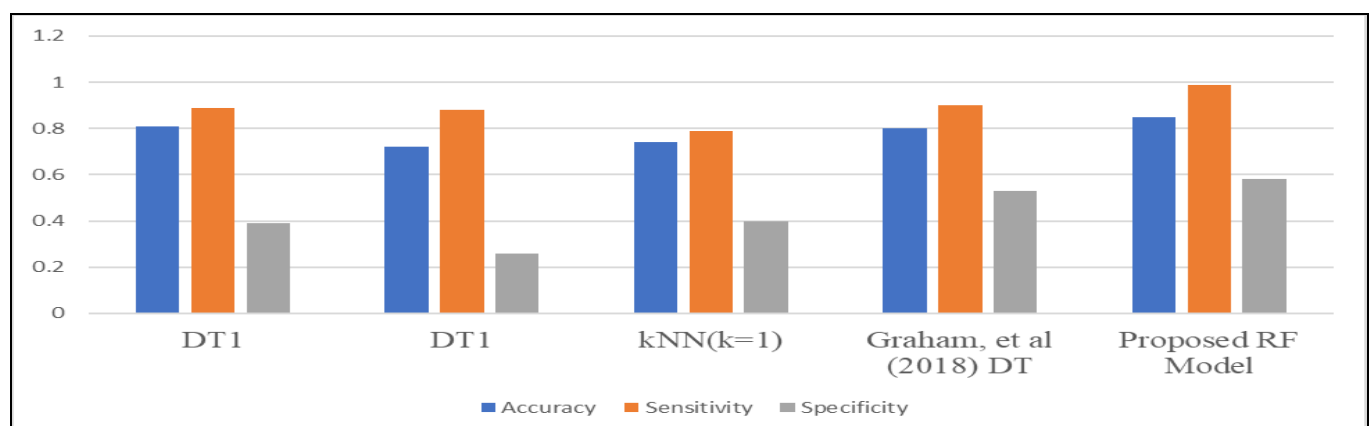


Fig 4: The Comparative Analysis of the Proposed System with the Existing Systems

E. Model Development

A modeling software was utilized to craft an ingenious model of the hospital emergency department. Meticulous attention was paid to the precise layout of the hospital, which proved instrumental in the development of the said model. A thorough analysis of the patients' flow in the section as mentioned earlier served as a valuable

foundation for accurately modeling the movement of patients across hospital.

Figure 5, is a visual representation showcasing the hospital's layout and the seamless transition of patients between different departments. The blue line elegantly illustrates the enchanting split flow pattern that guides patients throughout the entirety of the hospital.

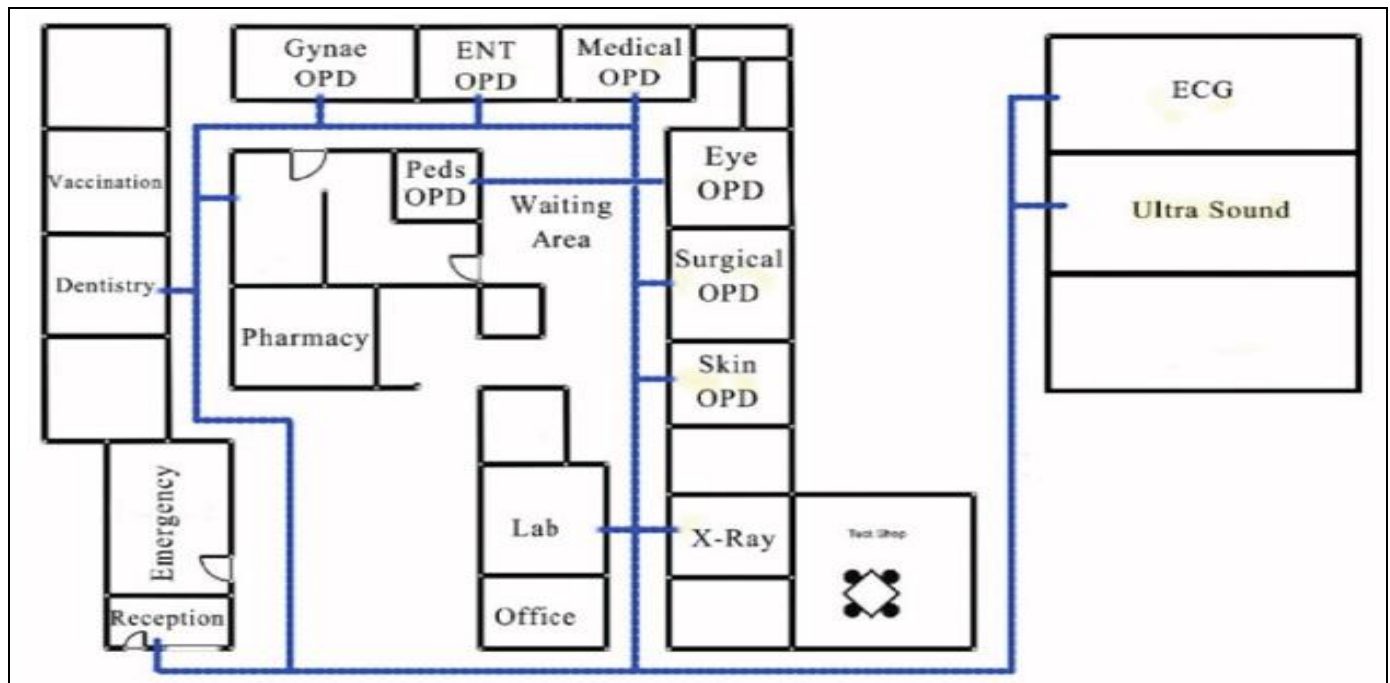


Fig 5: Visual Representation of the Hospital's Layout

V. CONCLUSION

In conclusion, this research endeavors to tackle the obstacles faced by emergency departments, particularly the issue of lengthy waiting times and overcrowding, by employing innovative methodologies such as queueing theory and the Random Forest Algorithm. The significance of this problem is well-established, highlighting its global impact on healthcare systems. The extensive review of existing literature identifies gaps and informs the development of the proposed system, which incorporates a simulation model with the Random Forest Algorithm to enhance the optimization of patient flow. Chapter Three outlines the methodology, showcasing the step-by-step process and emphasizing improvements over the current model. The outcomes of the implemented Random Forest Model in Chapter Four showcase its efficacy with an accuracy rate of 0.85, sensitivity rate of 0.99, and other favorable metrics. While the proposed model surpasses existing approaches, it is crucial to emphasize the importance of considering specific application requirements. This research not only contributes to the advancement of patient flow optimization but also provides a framework for future evaluations and comparisons between systems, ensuring ongoing refinement and adaptability in addressing emerging healthcare challenges.

Based on the discoveries and accomplishments delineated in the investigation, it is highly recommended that healthcare establishments, particularly emergency departments, contemplate the adoption and implementation of the proposed system that integrates the principles of queueing theory with the Random Forest Algorithm. The triumphant application of this model in the optimization of patient flow and reduction of waiting times, as evidenced by an accuracy of 0.85, sensitivity of 0.99, and other favorable metrics, indicates its potential to augment operational efficiency. Institutions should prioritize the fusion of both simulation modeling and machine learning algorithms, considering the specific attributes outlined in the study, such as patient registration, age, gender, arrival particulars, triage, x-ray, lab, and admission status. However, it is imperative for decision-makers to judiciously assess and customize the model to the distinctive characteristics and requisites of their respective healthcare environments. Regular evaluations and comparisons with existing models should be conducted to ensure continued improvement and pertinence in addressing the ever-evolving challenges in the management of emergency healthcare.

ACKNOWLEDGMENT

We wish to appreciate the expert support of our supervisors Dr. F. U. Zambuk and Dr. B. I Ya'u towards the success of this research.

REFERENCES

- [1]. Alenany, E., & Cadi, A. A. E. (2020). Modeling patient flow in the emergency department using machine learning and simulation. arXiv preprint arXiv:2012.01192.
- [2]. Elalouf, A., & Wachtel, G. (2021, December). Queueing problems in emergency departments: a review of practical approaches and research methodologies. In *Operations Research Forum* (Vol. 3, No. 1, p. 2). Cham: Springer International Publishing.
- [3]. Finkelstein, S. M., Wong, E., Lin, H., & Li, Y. (2021). Reducing Patient Waiting Time in Emergency Departments Using Reinforcement Learning and Queueing Theory. *INFORMS Journal on Applied Analytics*, 51(3), 162–179. <https://doi.org/10.1287/inte.2020.1064>
- [4]. Haripriya, G., Abinaya, K., Aarthi, N., & Kumar, P. P. (2021). Random Forest Algorithms in Health Care Sectors: A Review of Applications.
- [5]. Jarvis, P. R. E. (2016). Improving emergency department patient flow. *Clinical and experimental emergency medicine*, 3(2), 63.
- [6]. Khalilnejad Tabari, M., Aghajari, S., & Salahshour, S. (2022). Reinforcement Learning for Scheduling in Emergency Departments. In 2022 IEEE International Conference on Healthcare Informatics (ICHI) (pp. 10-16). IEEE.
- [7]. Lee, H., Lee, J., Lee, H., Lee, J., Lee, Y., & Kim, Y. (2021). Reducing Patient Waiting Time in Emergency Departments Using Reinforcement Learning and Multi-Agent Systems. *Applied Sciences*, 11(8), 3429. <https://doi.org/10.3390/app11083429>
- [8]. Li, J., Cao, W., Yang, Y., Hu, S., & Wei, S. (2021). Data-driven optimization of emergency department operations using machine learning algorithms. *Health Care Management Science*, 24(1), 32-46.
- [9]. Liu, J., Xie, J., Cheng, X., Zhang, Z., Li, X., & Li, Z. (2022). A Novel Reinforcement Learning Algorithm for Emergency Department Patient Flow Optimization. *Journal of Healthcare Engineering*, 2022, 1-11.
- [10]. Liu, Y., Liu, X., Li, Q., Zhu, Y., Ma, J., & Wang, Y. (2021). An intelligent triage system for an emergency department based on machine learning and queueing theory. *Journal of Medical Systems*, 45(9), 1-10.
- [11]. Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9, 381-386.
- [12]. Oh, S. H., Park, J., Lee, S. J., Kang, S., & Mo, J. (2022). Reinforcement learning-based expanded personalized diabetes treatment recommendation using South Korean electronic health records. *Expert Systems with Applications*, 206, 117932.
- [13]. Razavi, S., de la Hoz, E., & Akhavan-Tabatabaei, R. (2021). Optimizing Patient Flow in Emergency Departments Using Discrete Event Simulation and Deep Reinforcement Learning. *International Journal of Environmental Research and Public Health*, 18(9), 4722. <https://doi.org/10.3390/ijerph18094722>