International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I2P108

Volume 73 Issue 2, 92-106, February 2025

Original Article

Comparative Analysis of Machine Learning Algorithms for Predicting Consultation Wait Times in Outpatient Clinics

Jeffin Joseph[1,2], S. Senith[1], A. Alfred Kirubaraj[3], S.R Jino Ramson[4]

[1]*Department of Management Sciences, Karunya Institute of Technology and Sciences, Coimbatore, India.*
[2]*Department of Management Studies, St. Berchmans College, Kottayam, India.*
[3]*Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India.*
[4]*Global Foundries, Vermont, USA.*

[2]*Corresponding Author* : *jeffinjoseph@hotmail.com*

Received: 13 August 2024 Revised: 13 January 2025 Accepted: 24 January 2025 Published: 21 February 2025

Abstract - *Accurate prediction of consultation wait times significantly improves patient satisfaction and operational efficiency in outpatient clinics. This paper develops predictive models for consultation wait times using machine learning techniques, which include comprehensive predictors: patient demographics, temporal factors, queueing metrics, and historical wait time data. Fifteen Machine Learning models, including six regression and nine classification algorithms, were trained and evaluated using a dataset of 28,787 patient records from a multispecialty hospital. Results show that the tree-based models have better performance, in which the Decision Tree Regressor has the best performance among regression models with $R^2 = 0.98$, MAE = 0.40, and RMSE = 3.75, while the Random Forest Classifier is among the classification models with Accuracy = 95.65%, ROC-AUC = 98.91%. The analysis of feature importance using linear regression underscores the dominance of temporal factors and queueing metrics over demographic and historical wait time predictors in determining consultation wait times. This paper demonstrates that machine learning algorithms could be a good approach to predicting consultation wait time to help clinic operations. However, these findings need to be interpreted in the context of limitations in the dataset and the exclusion of some potentially important predictors. Future studies should validate the models in different clinical settings and include more variables to increase their generalizability and clinical usefulness.*

Keywords - *Hospital management, Machine learning, Wait time, Outpatient clinic, Random forest, Decision tree.*

1. Introduction

Waiting time in outpatient clinics is one of the important indicators of quality and efficiency in health care, but it is usually not given enough attention. Longer waiting times can have negative impacts on patients' satisfaction and probably lead to worse health outcomes because of increased anxiety and exacerbation of pre-existing medical conditions [1, 2]. Consultation wait time, defined as the time patients spend waiting to meet a physician, is one of the most important factors in shaping the patient experience [3, 4]. This metric is particularly difficult to forecast accurately due to the multi-causal and complex nature of contributing factors. While previous literature has shown promise in the ability of ML algorithms to predict wait times, there are still large gaps. Indeed, ML models have been applied in fields such as radiation oncology [10], phlebotomy [11], and pediatric ophthalmology [12] for predicting wait time, showing their adaptability across healthcare settings. However, most of the studies have limitations, including using a narrow set of predictors and insufficient algorithm performance evaluation. These drawbacks restrain the reliability and generalizability of their findings. The current study will address these gaps with the following primary research questions: (1) How does consultation wait time prediction accuracy change when broader sets of predictors are incorporated into the models? (2) How does the effectiveness of different ML algorithms compare for this application? (3) Which algorithms work best when approached as a regression problem? (4) Which algorithms are best when treated as classification?

Our approach also introduces feature engineering during data preprocessing, which empowers the models to take advantage of important variables such as demographic characteristics, temporal trends, queuing metrics, and historical wait time data. Such a combination of predictors has not been explored well in past research. What's more, this research performs a wide comparison among many ML algorithms: Decision Tree Regressor, Random Forest Regressor, Gradient Boosting Regressor, Bagging Regressor, Neural Network (Multilayer Perceptron), K-Nearest Neighbors Regressor, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, Bagging Classifier, Neural Network (Multilayer Perceptron), K-Nearest Neighbors Classifier, Gaussian Naive Bayes, Linear Discriminant Analysis, and Logistic Regression.

The distinctive contribution of this research lies in its comprehensive variable selection process and its systematic evaluation of algorithmic performance. Metrics such as R^2, MAE, and RMSE are used for regression models, while Accuracy, Sensitivity, Precision, F-measure, and ROC-AUC evaluate classification models. Based on our findings, it seems that among the regression models, the best performing is Decision Tree, while among classification models, Random Forest has the highest accuracy. Compared with previous studies, our approach has improved predictive performance and provided a stable framework for the management of consultation wait time in outpatient settings. The remainder of the paper is organized as follows: Section 1 introduces the study and states the problem statement; Section 2 gives a review of relevant literature; Section 3 focuses on model building and analyzing variables; Section 4 deals with the results, including the evaluation of models and feature importance; Section 5 interprets the findings of the study, including implications and limitations; while Section 6 provides the conclusion.

2. Related Studies

Related studies are analyzed into three categories: the machine learning models used, the variables used, and the performance metrics used to evaluate the models.

2.1. Machine Learning Models for Wait Time Prediction

Efforts to improve the prediction accuracy of clinical wait time for patients in a healthcare setting have been widely investigated through many machine learning algorithms in the literature. The choice of the algorithm is critical to finding a proper trade-off between accuracy and efficiency. Linear Regression has been widely applied in predicting wait times by many studies, which show its applicability in various contexts [14], [10], [15, 16, 17], [6]. The Quartile Regression offers another view by focusing on the distributional aspects of the wait times, hence showing insight into different data segments [18]. Multiple linear regression extends this analysis by simultaneously accommodating multiple predictors, capturing more complex relationships within the data [19], [12]. Elastic Net Regression, combining the strengths of both ridge and lasso regression, has been successfully applied in the literature on wait time prediction and has shown its versatility and robustness [12], [15, 16], [20]. Moreover, the Multivariate Adaptive Regression Splines (MARS) show the elasticity of regression analysis in capturing complex non-linear relationships, which implies its potential usefulness in wait time prediction [16]. Of note, several studies that employ regression analysis coupled with ensemble machine learning approaches have all reported results showing that the ensemble approaches, particularly Random Forest, surpass the conventional regression methods in terms of accuracy as well as reliability [14], [10], [15].

Decision Trees, including the foundational Decision Tree approach and the more specific Classification and Regression Trees (CART), are utilized to segment datasets into more manageable subsets, facilitating a structured approach to decision-making. These models will be highly rated for their simplicity, interpretability, and ability to handle different variable types, thus making them able to provide effective predictions related to clinical wait times. Basic Decision Trees have been used in research works [10], [14], with a more detailed description of Classification and Regression Trees (CART) given in [16]. Ensemble models distinguish themselves by integrating the capabilities of various predictive models, thereby enhancing both precision and resilience. Thus, they are more likely to break through limitations in individual models towards a stronger model. In this category come the Random Forest, Gradient Boosting Machine, and Bagging; examples show their extensive uses within studies for clinical wait time predictions [21], [17], [22], [15], [23, 24].

In this respect, deep learning algorithms lead the way in advanced data modeling and are capable of handling big and complex datasets. Neural Networks have been in use for a long time in wait time estimation and offer strength in dealing with big data [26], [11], [14]. Also, more advanced techniques, such as the Support Vector Machine, K-nearest neighbours, and their variants, have been applied in the analysis and prediction of the wait time to offer different ways of solving the problem [13], [6], [17], [25], [22]. While many studies have reported the performance of different machine learning models, there is some inconsistency regarding the best-performing algorithms. This gap is therefore filled to systematically compare a variety of algorithms based on decision trees, ensemble methods like Random Forest and Gradient Boosting, and Multilayer perceptron—neural networks belonging to the paradigm of deep learning. The tasks of regression and classification are realized to determine the best models to be used in predicting the consultation wait times. This paper also uses LDA, Logistic Regression, and Gaussian Naïve Bayes for the classification tasks so that the performance of the models can be fairly assessed.

2.2. Essential Predictive Variables

A vast literature review of the subject matter identifies a broad array of predictive variables for patient waiting times, which may be grouped into six categories: Demographic information, Temporal factors, Examination-specific characteristics, Queuing theory parameters, Historical waiting time information, and Miscellaneous predictors. Of the demographic information, Age and Gender are the most prominent. These are independent health service needs and, therefore, relate to the design and delivery of health services [25], [15], [10], [17], [26]. Other demographic data that further specify predictive models include patient financial class, new versus returning patient status, country of birth, Indigenous status, and preferred language [12], [15]. Those are important variables that explain variability in patient behaviour and healthcare needs, which may have a very large impact on wait times.

Temporal elements: capturing the changing nature of patient wait times, including "temporal snapshots" to accommodate short-term variations in queue lengths. Studies by [27], [12], [15], [26], [17], [11], and [16] emphasize the importance of the day of the week, specific times within the day, and public holidays in modulating seasonal and weekly trends. Including temporal factors addresses the dynamic nature of patient arrivals and services, which is very important for the accuracy of real-time predictions. Past patterns and trends can be observed from historical wait time data, enhancing the accuracy of predictions, such as those studies by [10], [16], and [12]. Historical data will help establish baseline patterns and variations on which future wait time forecasting depends. Queuing theory metrics, such as those measuring arrival and service rates, increase the accuracy of predictive models by several folds. Applications of the Pollaczek-Khinchine formula, prioritization in patient queues, and "measures of chaos" to quantify disruptions work quite effectively in unpacking patient flow dynamics [9], [16]. Attributes for each examination, such as the type of clinical examination, the preparations required, and the time allotted, all have an impact on determining wait times. Integrating variables such as the total number of patients, historical treatment durations, and International Classification of Diseases (ICD) codes by which patients have been treated makes the predictive models much richer [16], [19], [12], [25].

Furthermore, the inclusion of Miscellaneous predictors, such as weather conditions [27], [21] and modes of arrival by patients [25], [15], and novel variables, including compensable status [15], has brought out the flexibility in applying machine learning models toward wait time prediction. The present study incorporates various predictors, including demographic information, temporal factors, queuing theory metrics, and historical wait time data, in developing comprehensive models for predicting consultation wait time. This rigorous selection is based on vast literature, bringing forth the importance of each category.

2.3. Performance Measures of Previous Studies

When predictive modeling is applied in clinical settings to predict wait times, there is a need to select an appropriate method for measuring model accuracy. Mean Absolute Error (MAE) is the most used due to its simplicity. For example, Hijry and Olawoyin proved that the Deep Learning with Stochastic Gradient Descent (SGD) algorithm shows a lower MAE of 10.80 minutes compared to other methods like RMSprop, Adam, and AdaGrad with MAEs around 12 minutes [26].

Another study using random forest regression reported an MAE of 4.6 minutes and explained 47% of the variation [10]. In other studies devoted to predicting wait times in pediatric outpatient services, Gradient Boosting Decision Tree (GBDT) and Random Forest (RF) performed well with MAEs equal to 5.28 and 5.03 minutes, respectively [6]. Both had an R^2 of 0.97.

However, as MAE does not heavily penalize larger errors, using Root Mean Square Error (RMSE) and Mean Squared Error (MSE) provides a stricter evaluation of errors. For instance, in another study predicting wait time, RMSE values of 33.45 were found using a Random Forest in a Women's Clinic and 16.49 using a Support Vector Machine (SVM) in a Prompt Care Clinic [14].

Table 1. Overview of machine learning models and performance of best models in recent wait time prediction studies

Study	ML Model(s) Used	Performance of Best Models in Recent Studies
[11]	Artificial Neural Networks (ANN)	ANN: 88% Accuracy
[15]	Linear Regression, Random Forest, Elastic Net Regression	Random Forest: MAE = 22.6-44.0 minutes
[26]	Deep Learning (DL) with ANN and optimization algorithms	DL with SGD: MAE = 10.80 minutes
[25]	Logistic Regression	~52% Accuracy
[27]	OLS, Quantile, Ridge, LASSO Regression, Random Forest	LASSO reduces MSPE by 21%
[13]	LDA, SVM, Logistic Regression, KNN, Decision Tree, Naive Bayes, Gradient Boosting, Elastic Net	LDA: 99.02% Accuracy, SVM: 95.84% Accuracy
[6]	Linear Regression, Random Forest, Gradient Boosting, Decision Tree, KNN	GBDT MAE = 5.28, RF MAE = 5.03, R^2 = 0.97
[23]	Random Forest, Neural Network	Random Forest: RMSE = 2.81, MSE = 1.67, MAE = 0.88, R^2 = 0.996
[20]	Ridge, LASSO Regression, Decision Tree, Random Forest, ANN/DNN, Elastic Net Regression	Random Forest: Lowest RMSE = 6.69 minutes
[12]	Random Forest, Elastic Net, Gradient Boosting Machine, SVM, Multiple Linear Regressions	Random Forest: AUC = 81.55%, R^2 = 0.38, RMSE = 24.22

Another study compared the results from the Random Forest model to be excellent with RMSE of 2.81, MSE of 1.67, MAE of 0.88, and R^2 of 0.996 [23]. Comparatively, another study showed the Random Forest model yielding an RMSE of 6.69 minutes [20]. For classification problems, the Area Under the Curve (AUC) is relevant, with studies showing AUCs such as 81.55% for Random Forest [12] and accuracy rates of 99.02% and 95.84% for LDA and SVM, respectively [13]. In the current work, the selected metrics for regression – R^2, MAE, RMSE, MSE; for classification – Accuracy, Precision, Recall, F-measure, ROC-AUC have been selected since they are successfully used in similar studies [26], [14], [12, 13]. Table 1 gives the overview of Machine Learning Models and the performance of the best models from previous studies. Given potential data imbalances, precision, recall, and F-measure also prove essential in performance evaluation.

3. Model Development

3.1. The Framework of the Study

Data for the study is based on 28,787 patient records that the researchers had collected from the electronic health records of a multispecialty hospital in Kerala, India, for four months from October 2022 to January 2023. The data comprised the records from all seventeen outpatient departments of the hospital. Figure 1 illustrates the Patient Consultation Process in the outpatient departments. Patients come and register themselves at the counter shown, after which they sit in the waiting area until their name is called for consultation with the doctor. From this point, the patient may be referred to other departments, like the laboratory or pharmacy, for associated services.

Meanwhile, the next patient moves forward for their consultation with the doctor, hence giving a smooth flow of the process in the system. The original data set had variables like patient ID, 'Age', 'Gender', 'Doctor', 'Department', date and time of registration and the time spent consulting with the doctor. The dependent variable, 'Consultation Wait Time', is determined as the difference between the registration time of the patient and the time of consultation by the doctor. In this step, during data cleaning and preprocessing, outliers were identified and removed to give a refined dataset of 27,151 patient records. Specifically, those patient records where the wait times were more than 120 minutes were excluded based on the assumption that such longer wait times could be due to doctors attending to emergency cases rather than routine consultations. Some feature engineering techniques were used to improve the prediction accuracy. Temporal variables: 'Visit Day', 'Consulting Session', 'Time of Visit', and 'Weekday vs. Weekend' were extracted from the registration date and time data. The 'Consultation Start Time' variable, representing the time a consultation begins on a given day, was extracted from the consultation time data. 'Repeat Arrival' was found by analyzing patient IDs to track patients who visited the facility more than once.

Added to that are variables like 'Avg Wait Time per Doctor', 'Avg Wait Time per Department', 'Median Wait Time per Doctor', 'Median Wait Time per Department', 'Avg Wait Time per Session', 'Median Wait Time per Session', 'Historic Avg Wait Time (3 days)', and 'Historic Median Wait Time (3 days)' to describe the average and median wait time of different categories based on the consultation wait time. Also, queueing theory metrics were derived from the time of registration and consultation time, respectively: 'Arrival Interval per 30 minutes' and 'Service Interval per 30 minutes'. These derived features are included in our model concerning their established importance in the literature. See how the variables were analyzed in further detail in Section 3.2.

The preprocessing also included one-hot encoding to refine the data for further analysis. The model was trained on 75% of the dataset and tested on the remaining 25%. The 75%-25% split ensures that there is sufficient data for training and testing, striking a practical balance between letting models learn enough and rigorously evaluating their performance. This approach avoids overfitting and ensures that the metrics obtained, RMSE and MAE, reflect the predictive power and generalizability of the models, given their complexity [28].

The study employed a chain of regressors and classifiers and carried out fifteen models to predict the consultation waiting time. Namely, the models are Decision Tree Regressor, Random Forest Regressor, Gradient Boosting Regressor, Bagging Regressor, Neural Network using Multilayer Perceptron MLP Regressor, K-Nearest Neighbors Regressor, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, Bagging Classifier, Neural Network using MLP Classifier, K-Nearest Neighbors Classifier, Gaussian Naive Bayes, Linear Discriminant Analysis and Logistic Regression. The models have been selected based on an in-depth analysis of the previous works.

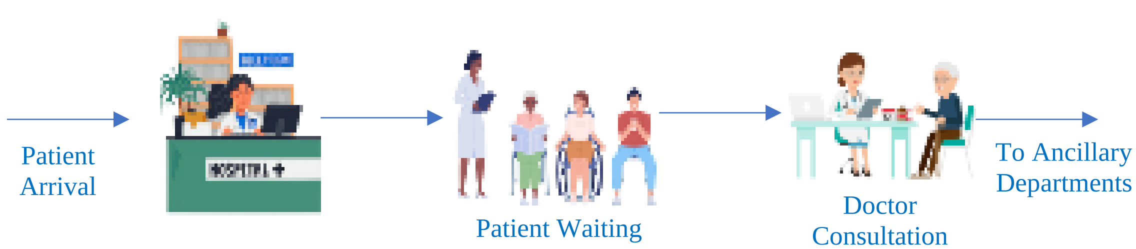

Fig. 1 Patient consultation process in the outpatient department

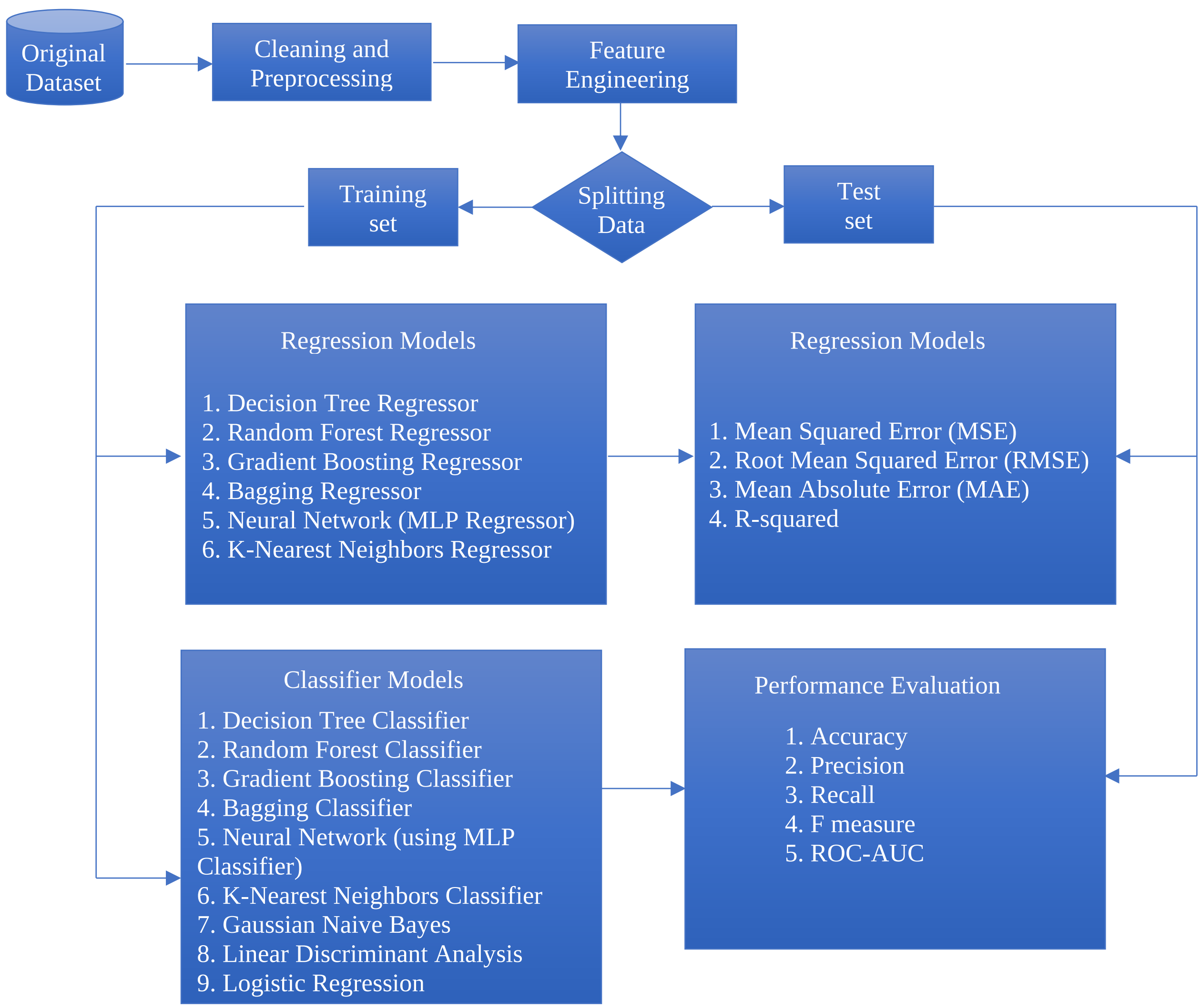

Fig. 2 Framework of predicting consultation wait time

This research uses the Python Scikit-learn library to implement models for predicting wait times. Regression models are considered for predicting the continuous variable "Consultation Wait Time Minutes." Those are evaluated based on some metrics: Mean Squared Error—MSE, Root Mean Squared Error—RMSE, Mean Absolute Error—MAE, and the Coefficient of Determination—R^2. On the other side, classification models were used to predict the categorical outcome "Consultation Wait Time Multinomial." The following metrics were utilized in evaluating their performance: Accuracy, Precision, Recall, F-measure, and the Receiver Operating Characteristic Curve (ROC-AUC). Figure 2 presents the comprehensive framework of the study.

3.2. Analysis of Variables

Categorical variables used for the prediction of the Consultation Wait Time are 'Gender', 'Department', 'Doctor', 'Visit Day', 'Consulting Session', 'Repeat Arrival', and 'Weekday vs weekend'. Continuous variables are 'Age', 'Consulting Start Time', 'Avg Wait Time per Doctor', 'Avg Wait Time per Department', 'Median Wait Time per Doctor', 'Median Wait Time per Department', 'Avg Wait Time per Session', 'Median Wait Time per Session', 'Historic Avg Wait Time 3 days', 'Historic Median Wait Time 3 days', 'Arrival Interval per 30 minutes' and 'Service Interval per 30 minutes'. Thus, the research predicts the continuous variable: 'Consultation Wait Time Minutes' and the categorical variable: 'Consultation Wait Time Multinomial' using the same set of predictor variables.

The 'Consultation Wait Time Multinomial' variable is categorized into three classes: Low, Medium, and High, based on the waiting time. More precisely, a waiting time of fewer than 30 minutes is defined as "Low," between 30 and 60 minutes as "Medium," and longer than 60 minutes as "High." The classification thresholds are based on some of the literature found in the references [29, 30, 31]. The paper carries out an in-depth descriptive analysis of the variables using different visualization techniques, such as bar charts and histograms.

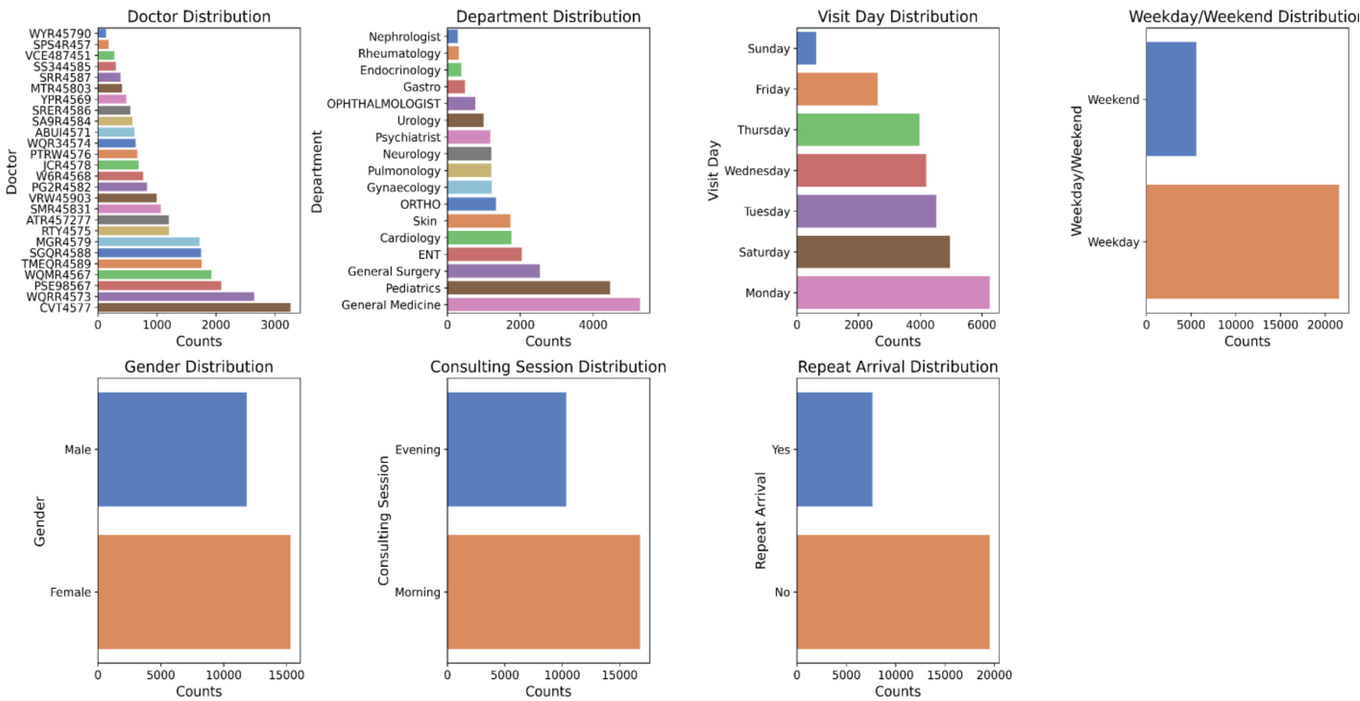

Fig. 3 Bar charts for each categorical variable

3.2.1. Analyzing the Categorical Variables

Figure 3: Bar Charts for different categorical variables like 'Gender', 'Department', 'Doctor', 'Visit Day', 'Consulting Session', 'Repeat Arrival', and 'Weekday vs Weekend'. 'Department' has 17 unique categories, which are General Surgery, Ophthalmology, Neurology, Pediatrics, ENT, Gynecology, General Medicine, Dermatology, Endocrinology, Cardiology, Psychiatry, Nephrology, Orthopedics, Pulmonology, Urology, Gastroenterology, and Rheumatology. The departments of General Medicine, Pediatrics, and General Surgery attract the greatest number of patients. The chart below, 'Doctor' distribution, shows that the doctor with ID CVT4577 gets the most visits from the patients. 'Gender' is divided into Male and Female, showing the distribution across sexes. The 'Visit Day' variable ranges from Monday to Sunday, with Monday having the most patients and Sunday the least. Consulting sessions are grouped into Morning and Evening, showing the spread of visits over these times. 'Repeat Arrival' is broken down into Yes and No, showing how often patients pay repeat visits. Finally, in the comparison between 'Weekday vs Weekends', one can see the trend of patient visits over these times.

3.2.2. Analyzing Continuous Variables

Figure 4: The histograms of the continuous variables: the histogram for 'Age' suggests a nearly normal distribution of ages with, however, a slight skewing to the right—that means there is a greater number of younger patients. One peak lies around the age range of 20–30 years, and the second one is less distinct, around 60–70 years of age. This bimodal distribution may indicate the presence of two large groups of patients visiting the health facility: younger adults and older people. The trend line smooths out the fluctuations and accentuates the normal distribution shape. These trend lines, or kernel density estimates (KDE), present a smoothed picture of the distribution of data, which can be very useful in the identification of patterns that may underlie the data [32].

The following histogram has been prepared for 'Consultation Start Time' minutes past midnight. Such conversion makes the time data numerical and thus amenable to analysis and visualization of the distribution of consultation start times throughout the day. This histogram shows the frequency of consultations starting at different times and helps understand the pattern and the peak consultation hours. This histogram has one cluster of consultations, beginning early in the morning, peaking at around 10 am; then, it just slightly decreases. The trend line shows that most consultations are during the morning hours. The 'Average Wait Time per Doctor' data points are very spread out, with a few peaks that show some doctors have an average much longer wait time.

The trend line would indicate that most of the doctors have relatively short wait times, with some outliers causing spikes in the graph. Similarly to the 'Average Wait Time per Doctor', the 'Average Wait Time per Department' is not distributed equally; there are departments with much longer waits. The trend line shows that more departments have shorter wait times than those with long wait times. The histogram for 'Median Wait Time per Doctor' is much less variable than the average, which may be an indication that the average is more sensitive to outliers than the median.

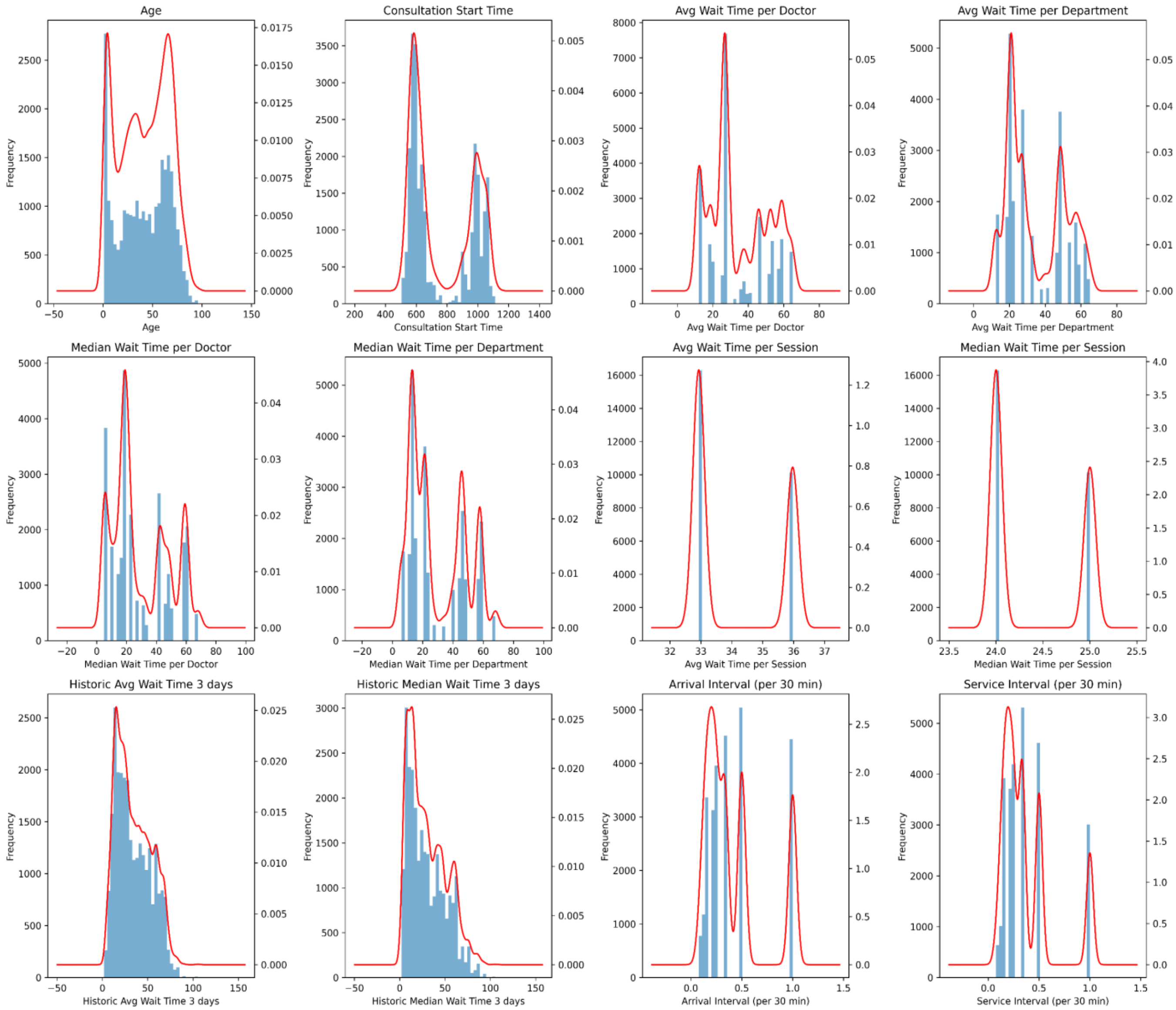

Fig. 4 Histograms for each continuous variable

Also, the trend line is fairly flat, which could indicate that the median wait times are spread out somewhat equally across all doctors. The distribution of 'Median Wait Time per Department' is less variable than that of 'Average Wait Time per Department'. The trend line is flatter, which may suggest that median wait times across different departments are more homogeneous. The 'Average Wait Time per Session' metric is as follows: morning and evening sessions have an average wait time of about 33 minutes and 36 minutes, respectively. On the other hand, the histogram of 'Median Wait Time per Session' has a very narrow distribution, with most of the sessions having wait times of around 24 minutes and 25 minutes for morning and evening sessions, respectively. This could also mean that the wait times for the sessions remain the same for the most part. For the variables 'Historic Avg Wait Time 3 days' and 'Historic Median Wait Time 3 days', the calculation is defined to look back over the last three days. The values for the earlier days are calculated using a smaller time window. Here's how it's done: For Day 1, the historical values are calculated using the values of Day 1 itself, as there are no previous days to use. Day 2 the historic values just have data from Day 1 and Day 2. For Day 3 The historic values now include data from Day 1, Day 2, and Day 3. From Day 4 onward, the average is carried out as desired over the complete 3-day window. This means that for the first three days, the historical values are effectively the same as the actual average and median for those days, limited to the available data. The 'Historic Average Wait Time 3 days' variable has a right-skewed distribution with a peak at the lower end. This indicates that typically, the average wait time is low, but there are periods with longer wait times. 'Historic Median Wait Time 3 days' distribution is also right-skewed.

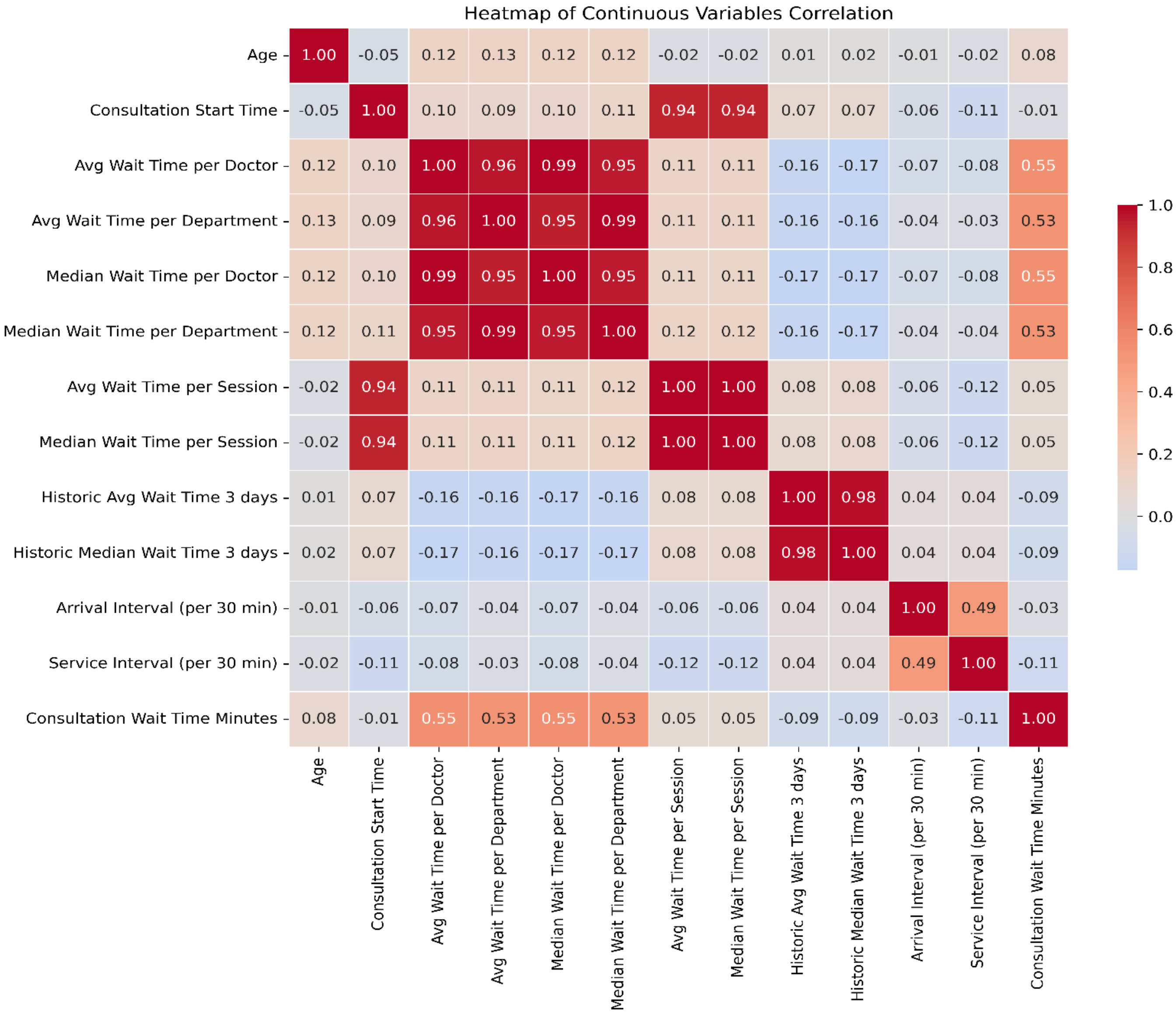

Fig. 5 Heatmap showing the correlation between all continuous variables

The variable 'Arrival Interval (per 30 min)' indicates the time per arrival for each 30-minute session interval. It is calculated as the inverse of the number of arrivals in that interval. Similarly, the 'Service Interval (per 30 min)' indicates the time per service for each 30-minute interval, calculated as the inverse of the number of services completed in that interval. The histogram of 'Arrival Interval (per 30 min)' is skewed to the right with the mass of data concentrated at the lower end, which might hint at the fact that the time per arrival is typically short, suggesting a high arrival rate, but there are periods with significantly longer intervals between arrivals. Similarly, 'Service Interval (per 30 min)' is also skewed to the right with a peak at the lower end. This implies that the time per service is generally short, indicating a high service rate, but there are times when the intervals between services increase significantly.

3.3. Correlation among Continuous Variables

The heatmap in Figure 5 shows the correlation coefficients for a pair of continuous variables. The variables included are 'Age', 'Consultation Start Time', 'Avg Wait Time per Doctor', 'Avg Wait Time per Department', 'Median Wait Time per Doctor', 'Median Wait Time per Department', 'Avg Wait Time per Session', 'Median Wait Time per Session', 'Historic Avg Wait Time 3 days', 'Historic Median Wait Time 3 days', 'Arrival Interval (per 30 min)', 'Service Interval (per 30 min)', and 'Consultation Wait Time Minutes'. The lighter shades of blue in the heatmap represent a weak negative correlation; the transition through light orange into dark red shows an increasingly positive correlation. Light blue would suggest that there may be some sort of faint inverse relationship between the variables: as one goes up, the other tends to go down, but not strongly. The orange color in the heatmap represents moderate positive correlations, which lie between the strong correlations represented by dark red and the negligible correlations represented by white. Numbers in Each Cell are actual values of correlation coefficients. From the heatmap, the following interpretations were drawn: Age is very weakly correlated with all other variables, as shown by values near 0.

This would further suggest that age is not a strong predictor of wait times or other variables. Thus, other variables such as 'Average Wait Time per Doctor' or 'Average Wait Time per Department' might be much more critical in the prediction of wait times. Variables from 'Consultation Start Time' to 'Median Wait Time per Session' all show very high positive correlations with each other—very close to 1 and in red—meaning these metrics do tend to move together: when one goes up, so does the other. The 'Arrival Interval (per 30 min)' and 'Service Interval (per 30 min)' show a moderate positive correlation with each other but a very low correlation with wait times, indicating relatedness without a strong influence on wait times.

This may suggest that there is a moderate level of correlation between arrival and service intervals, which could be an indication that the scheduling and handling of patient arrivals and services are somewhat coordinated. 'Consultation Wait Time Minutes' moderately positively correlates with the 'Avg Wait Time per Doctor', 'Median Wait Time per Doctor', 'Avg Wait Time per Department', and 'Median Wait Time per Department' variables, meaning longer average and median wait times correlate with longer consultation wait times. The study used two of the most popular visualization Python libraries—Seaborn and Matplotlib—to generate various types of visualizations, including heat maps, histograms, and bar charts.

4. Analyzing Performance

4.1. Analysing Performance of Regression Models

The performance of the regression models was evaluated using key metrics: the coefficient of determination, R^2; Mean Absolute Error—MAE; Root Mean Square Error—RMSE; and Mean Squared Error—MSE. The metrics are very relevant to establishing the accuracy and effectiveness of the regression models in predicting wait time for outpatient consultation. They were selected after an extensive review of the literature to ensure their appropriateness and common acceptance in evaluating predictive performance in the setting of wait time prediction. A comparative analysis of regression models shows that the Decision Tree Regressor does best with an R^2 of 0.98, MAE of 0.40, RMSE of 3.75, and MSE of 14.06 (refer to Table 2). The results for the Random Forest Regressor and the Bagging Regressor also look very good. On the other hand, Gradient Boosting Regressor and Neural Network—MLPRegressor perform quite poorly, which could be a sign that there is an overfitting problem, or maybe it just needs more tuning with hyperparameters and feature engineering.

Table 2. Performance measures of regression models for predicting consultation wait time

Model	R^2	MAE	RMSE	MSE
Decision Tree Regressor	0.98	0.4	3.75	14.06
Random Forest Regressor	0.97	2.57	5.08	25.82
Gradient Boosting Regressor	0.4	17.57	23.03	530.24
Bagging Regressor	0.96	2.84	5.97	35.59
Neural Network (MLPRegressor)	0.31	18.08	24.74	612.14
K-Nearest Neighbors Regressor	0.65	12.11	17.56	308.22

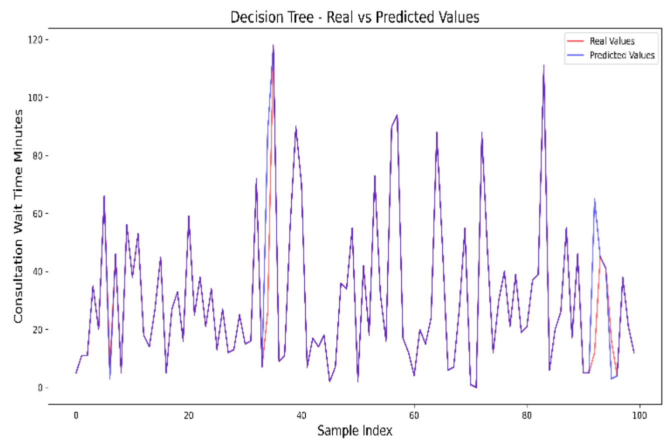

Fig. 6 Real vs Predicted wait time for the decision tree regressor

The K-Nearest Neighbor Regressor does moderately well but can surely be improved with some feature scaling and hyperparameter optimization. The Decision Tree Regressor is the best model because of its superior accuracy and lower error rates.

Figure 6: Line graph showing real vs. predicted consultation wait times for a randomly selected sample of 100 records from the test set using the decision tree model. The red line represents actual wait times, and the blue line represents predicted values.

The difference between the two lines can serve as a gauge of how correct the model is in its prediction of the wait time for consultation. As one may see, the predicted values follow the real values, creating a purple color where the blue and red lines overlap, which means the decision tree model performed well.

4.2. Analysing Performance of Classification Models

The paper developed the classification models with the purpose of classifying waiting times into three classes: Class 1, representing 'Low Waiting Time'; Class 2, representing 'Medium Waiting Time'; and Class 3, representing 'High Waiting Time'. The methodology used here was to decompose the multiclass classification problem into several binary classification tasks by using the One-vs-Rest (OvR) strategy.

Table 3. Performance measures of classification models for predicting consultation wait time

Model	Accuracy	Precision	Recall	F-measure	ROC-AUC
Decision Tree Classifier	94.5	94.5	94.5	94.5	95.53
Random Forest Classifier	95.7	95.64	95.65	95.64	98.91
Gradient Boosting Classifier	67.9	65.75	67.88	65.47	82.23
Bagging Classifier	95.3	95.26	95.27	95.26	98.75
Neural Network MLP Classifier	56.9	56.17	56.87	53.49	76.52
K-Nearest Neighbor Classifier	71	70.4	71	70.65	88.73
Gaussian Naive Bayes	61.2	59.72	61.17	58.74	75.08
Linear Discriminant Analysis	63	59.24	63.03	59.91	76.99
Logistic Regression	62.7	57.94	62.68	57.74	76.03

Results from the individual binary classifiers were combined using a weighted average method to decide the final class [33]. The study assessed the performance of the classifier models on a panel of evaluation metrics: Accuracy, Precision, Sensitivity, the F measure, and the Area under the ROC (Receiver Operating Characteristic) Curve. To evaluate the multiclass problem in detail, the study averaged these key metrics from each binary classifier by using macro-averaging, hence providing a detailed assessment of the model's predictive accuracy. Among the classifiers, the Random Forest performed best with an accuracy of 95.65%, precision of 95.64%, recall of 95.65%, and a very strong ROC-AUC of 98.91% (see Table 3). Similarly, the Bagging Classifier showed strong results in handling complex and high-dimensional data and capturing feature interactions, which proved ideal in the prediction of wait times. The Decision Tree Classifier also performed well, which indicates that individual decision trees can model the underlying patterns in wait times, but using ensemble methods makes the models even more robust and stable. On the other hand, the Gradient Boosting Classifier resulted in lower metrics, which could be indicative of overfitting and sensitivity to noise, hence requiring careful tuning in the context of wait time prediction.

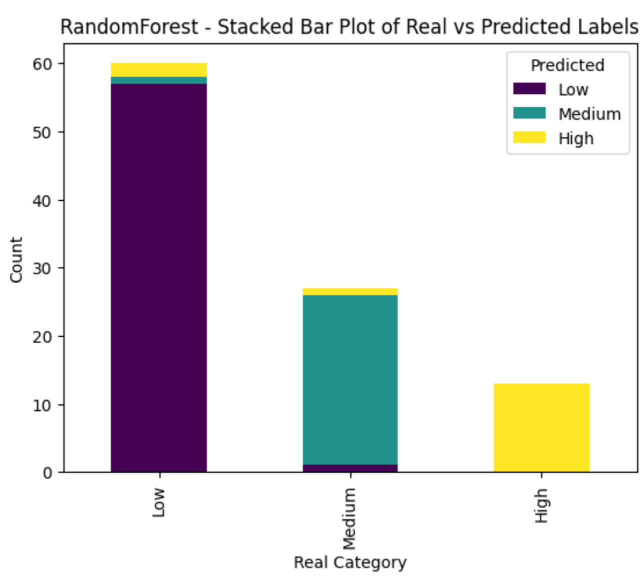

Fig. 7 Real vs Predicted wait time for the random forest classifier

The Neural Network MLP Classifier did not generalize well to the complexities in the wait time data, probably due to its nature, which usually requires large datasets and extended tuning. The K-Nearest Neighbors Classifier showed moderate performance, which can be explained by its limitations regarding dealing with complex relationships in wait time prediction because of high-dimensional data challenges. Gaussian Naive Bayes, Linear Discriminant Analysis, and Logistic Regression were poorer in accuracy since the mentioned models assume linearity and independence among the features—simplistic assumptions that do not explain complex, nonlinear interrelations found in wait-time data. This makes the Random Forest model selected as the best since it is relatively better than the other models. Figure 7: This is a stacked bar plot comparing real vs. predicted consultation waiting time over a sample of 100 records using a Random Forest model. On the x-axis are real wait time categories: Low, Medium, and High; the y-axis shows the count of records. Each of the bars is, in turn, made up of segments showing predictions: Low in purple, Medium in teal, and High in yellow, thus showing insights into the accuracy of the model and misclassification rates across the categories. Most of the Low wait times are predicted correctly, while there is some misclassification in the Medium and High categories.

4.3. Analysis of Feature Importance Using Linear Regression

It is found from Table 4 that the feature importance analysis with linear regression revealed Average Wait Time per Doctor to be the most influential predictor of consultation wait times, with 8.46. This finding underlines the fundamental role that doctor-level efficiency plays in reducing delays. Other significant operational predictors were Average Wait Time per Department (5.99) and Median Wait Time per Doctor (5.87), which suggests that both departmental and individual performance measurements need to be considered for the better allocation of resources. Other temporal factors were also found to be of considerable importance, including the Consultation Start Time at 3.20, queuing-related metrics of Service Interval at 2.95, and Arrival Interval at 1.76. The results underline the predictive power of time-based variables, particularly in modeling dynamic patterns of patient arrivals and service bottlenecks.

Table 4. Feature importance based on linear regression

Feature	Importance
Avg Wait Time per Doctor	8.46
Avg Wait Time per Department	5.99
Median Wait Time per Doctor	5.87
Median Wait Time per Department	4.05
Consultation Start Time	3.2
Service Interval (per 30 min)	2.95
Arrival Interval (per 30 min)	1.76
Avg Wait Time per Session	1.19
Consulting Session	1.19
Median Wait Time per Session	1.19
Repeat Arrival	1.19
Doctor	0.43
Visit Day	0.38
Age	0.24
Gender	0.13
Historic Avg Wait Time (3 days)	0.12
Weekday/Weekend	0.09
Department	0.03
Historic Median Wait Time (3 days)	0.03

In contrast, demographic and categorical variables, including Gender (0.13) and Department (0.03), showed minimal influence, suggesting that operational and temporal factors far outweigh static characteristics in predicting wait times. Historical predictors, such as Historic Average Wait Time (3 days) (0.12) and Historic Median Wait Time (3 days) (0.03), also contributed marginally, indicating limited value in this context. The reason linear regression was chosen for feature importance analysis is that it is interpretable, providing clear coefficients that quantify the impact of each predictor on consultation wait times. These results emphasize the importance of healthcare systems focusing on time-sensitive and operational metrics in predictive modeling.

5. Discussion

This study compares the relative performance of predictive models for consultation wait times, covering both regression and classification techniques. It was determined that the Decision Tree Regressor and Random Forest classifier were the best models for the prediction of consultation wait times. This is following previous studies, which show the strong performance of Random Forest models in clinical applications [10], [12], 15]. However, there are large variations in model performance and metrics used across the studies. The conclusions of the current study are compared with those of earlier studies in the following sections, along with the study's limitations, practical implications, and suggestions for further research.

5.1. Evaluation of Performance

In the study, different regression techniques are considered, and the decision tree regressors, performance is excellent for R^2 to 0.98 and error metrics close to the minimum: MAE to 0.40, RMSE to 3.75, MSE to 14.06. Similarly, the Random Forest Regressor had a strong predicting power with R^2 equal to 0.97, with strong error metrics (MAE of 2.57, RMSE of 5.08, MSE of 25.82). In terms of classification accuracy, Random Forest Classifier and Bagging Classifier have done very well with 95.65% and 95.27%, respectively.

The Random Forest Classifier also scored highly, with ROC-AUC at 98.91%. On the other side, models like the Gradient Boosting Classifier and MLP Classifier performed worse, as shown by the lower accuracy and ROC-AUC scores. Compared with the previous studies on healthcare wait time using predictive modeling, the current study shows a substantial performance improvement. For example, in the study by Hijry and Olawoyin [26], the authors reported an MAE of 10.80 minutes for the SGD model, which is far outperformed by the 0.40 minutes MAE obtained from the Decision Tree Regressor in the present study. Similarly, Joseph et al. [10] had an MAE of 4.6 minutes with a Random Forest model, while Li et al. [6] stated that the GBDT reached an MAE of 5.28 minutes and the RF had an MAE of 5.03 minutes. All are far above the Decision Tree Regressor in the present study, which had an MAE of 0.40 minutes and a Random Forest Regressor with an MAE of 2.57 minutes. In another study, Rastpour and McGregor [14] focused on predicting wait time at mental healthcare clinics. In that, his Random Forest model scored Root Mean Square Error (RMSE) values of 33.45 in the Women's Clinic and an RMSE score of 16.49 using an SVM model in the Prompt Care Clinic. The reported RMSE scores in the said study were higher compared to the RMSE recorded in this present study of 3.75 from the Decision Tree Regressor.

However, in a study conducted in an Emergency Department, Atalan [23] utilized the Discrete Event Simulation and reported smaller values for the RMSE—2.81. The Mean Square Error (MSE) is 1.67 with respect to the Random Forest Model compared to the best-performing one within the current research. These discrepancies show the power and impact of different clinical settings and variables in wait times, which flags the need for tailoring predictive models to the unique characteristics of each healthcare environment.

Also, in the study by Amjed Al-Mousa et al. [20], an RMSE of 6.69 minutes was obtained for their Random Forest model; however, in the present study, the results are shown to be better by the Random Forest Regressor with an RMSE of 5.08 and with a lower RMSE of 3.75 for the Decision Tree Regressor. Lin et al. [12] reported an AUC value of 81.55% for a Random Forest classifier; this is much below the ROC-AUC score of 98.91% observed in this study for the Random Forest Classifier. Moreover, Ataman and Sarıyer [25] discussed an accuracy of 52.247% for the ordinal logistic regression model in predicting wait times, which is way below the 95.65% accuracy of the Random Forest Classifier of this study. Overall, these comparisons underline the great improvements made by the machine learning models of the current study, especially the Decision Tree Regressor and Random Forest Classifier, in predicting wait times as accurately as possible.

5.2. Practical Implications

This paper highlights some of the practical implications of the findings for healthcare organizations seeking to improve wait time prediction and its management. Implementation of machine learning models in the prediction of consultation wait time could help healthcare organizations achieve better utilization of resources, greater patient satisfaction, and a general improvement in the level of healthcare delivery. However, there is a quite high number of barriers to adoption. The first major one is the requirement for specialized expertise in ML and data science; healthcare organizations may be interested in investing in training or hiring personnel with the needed skills to correctly implement and maintain the models.

Another possible barrier is related to concerns about data privacy and security when using ML models. While these barriers exist, some facilitators can enable an organization to adopt these models. For example, the potential to cut costs and improve patient satisfaction can motivate an organization to surmount the challenges in implementation. It can also be facilitated by collaboration with ML and data science experts and involvement in the implementation process stakeholders such as clinicians and administrators. An implementation framework is proposed (Figure 8) that is easily integrated with already existing healthcare information systems, considering critical aspects, such as compatibility with EHR systems, secure data-sharing protocols, and user-friendly interfaces for both clinicians and administrators.

The framework starts with collecting data from EHR and then subjects it to stringent preprocessing for quality assurance. It applies state-of-the-art machine learning models to generate actionable insights, which are integrated through an intermediary layer in a secured manner. Such knowledge is brought to the end user by accessible dashboards, while a feedback loop ensures continuity in updating and improvement. Testing this framework in real clinical settings will further refine its practical utility and guide enhancements in integrating ML technologies into healthcare.

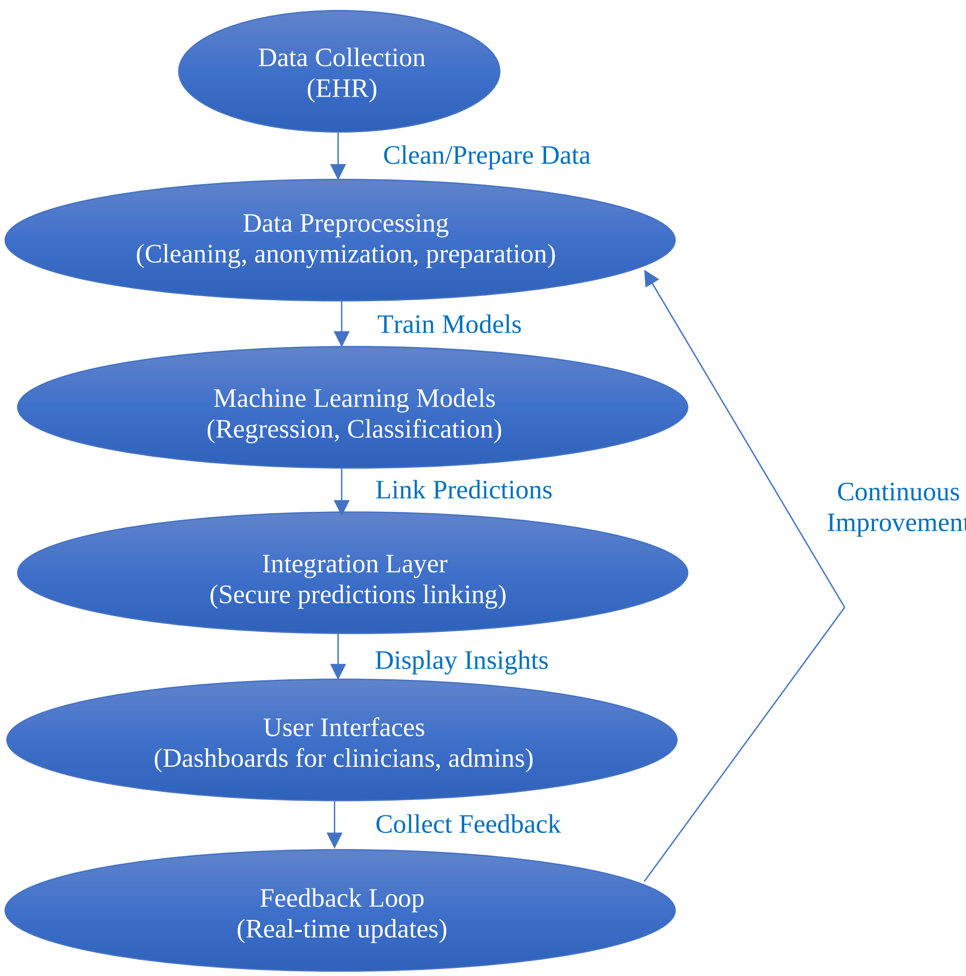

Fig. 8 Implementation framework

5.3. Limitations and Future Research Recommendations

The study's results must be interpreted cautiously, given the limitations associated with the size and source of the dataset. Data were exclusively retrieved from outpatient departments of a multispecialty hospital and are moderate in scale. The limitation was caused mainly by difficulties related to accessing the larger and more diverse dataset from the electronic health record system of the hospital. Secondly, a dataset emanating from a single multispecialty hospital may limit the generalizability of the findings to other clinical settings.

Its unique characteristics—be it the patient population, clinical practices, or operational processes—may not be representative of other hospitals or healthcare facilities. Consequently, the predictive performance and generalizability of the machine learning models in the present study may not be the same when applied to different healthcare settings. Future research should, therefore, focus on the validation of the findings with data from several hospitals that are diverse in terms of patient population, clinical practice, and operational environment.

Also, the variable scope was restricted to those available in the hospital's database. Such a limitation may have excluded potentially important variables to be included in the study. A major limitation is that no study-specific clinical variables are included in the dataset. Such variables, like the type of medical examination, duration of treatments, and International Classification of Diseases—ICD codes—are important for a deeper understanding of the dynamics within medical examinations [19], [25]. Including these variables might go a long way in increasing the model's accuracy since it provides more details on patient consultations.

The class imbalance issue, which frequently impairs machine learning models' predicted performance in classification tasks, was another issue the study had to address. Future studies could also seek to balance this class imbalance problem, perhaps by using either oversampling, under-sampling techniques, or synthetic data generation in general (like SMOTE). Furthermore, feature-selection methods like Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) can reduce the number of features, making prediction easier, models easier to understand, and computations more efficient.

Future research should be carried out in the following directions:

- Expansion of the dataset to include multiple hospital settings to enhance generalizability.
- Including examination-specific variables and other clinical details into the models for better prediction accuracy.
- Exploration of the inclusion of real-time data streams emanating from healthcare information systems for a necessary dynamic prediction.
- Handling class imbalance typical in classification problems by advanced resampling or algorithmic techniques to strengthen the models.
- Feature selection methods can be applied to decide the most pertinent predictors for the reduction of complexity and optimization of performance.

From these areas, future studies will further build on the present research and hence bring innovation in predicting and managing healthcare wait times.

6. Conclusion

This will generally show the potential of machine learning techniques for accurately predicting consultation wait time in outpatient clinics. This large-scale evaluation compares fifteen models, regression, and classification using a diverse set of predictors, thereby giving insight into the relative performance of different algorithms for this task. The Decision Tree Regressor and Random Forest Classifier have higher accuracy, which implies that tree-based models are practically viable for implementation in the operation of a clinic.

However, given the study's limitations, the results should be viewed cautiously. Some of the limitations include dependence on data from a single hospital and not including some potentially important predictors, such as examination-specific variables. Thus, future research will be important to validate the models proposed here in other clinical settings and consider additional relevant variables to increase their robustness and practical utility.

While this study is not without limitations, it adds to the fast-growing body of research in machine learning applications in healthcare. These results underline the importance of using data-driven approaches when seeking to improve patient flow and resource allocation in outpatient clinics. Accurate prediction of consultation wait times allows health organizations to better manage patient expectations, decreasing frustration and optimizing operational efficiency.

For ML-based wait time prediction to reach its full potential, a future focus of work needs to concentrate on developing user-friendly interfaces and integrating models into the existing clinical workflows. Collaboration among data scientists, healthcare providers, and clinic administrators will be important for successfully implementing and adopting the approaches. The approaches developed in this study have the potential to substantially impact the level of patient satisfaction and outpatient quality of care with further refinement and validation.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] John Launer, "Waiting Rooms and the Unconscious," *Postgraduate Medical Journal*, vol. 88, no. 1040, pp. 361-362, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[2] Luigi Siciliani, and Rossella Verzulli, "Waiting Times and Socioeconomic Status among Elderly Europeans: Evidence from SHARE," *Health Economics*, vol. 18, no. 11, pp. 1295-1306, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[3] Shane Mesko et al., "Using Patient Flow Analysis with Real-Time Patient Tracking to Optimize Radiation Oncology Consultation Visits," *BMC Health Services Research*, vol. 22, no. 1, pp. 1-7, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Christoper A. Alarcon-Ruiz, Paula Heredia, and Alvaro Taype-Rondan, "Association of Waiting and Consultation Time with Patient Satisfaction: Secondary-Data Analysis of a National Survey in Peruvian Ambulatory Care Facilities," *BMC Health Services Research*, vol. 19, no. 1, pp. 1-9, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hui Zhang et al., "How to Adjust the Expected Waiting Time to Improve Patient's Satisfaction?," *BMC Health Services Research*, vol. 23, no. 1, pp. 1-8, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Xiaoqing Li et al., "Prediction of Outpatient Waiting Time: Using Machine Learning in a Tertiary Children's Hospital," *Translational Pediatrics*, vol. 12, no. 11, pp. 2030-2043, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Holly Chu et al., "The Psychology of the Wait Time Experience - What Clinics Can Do to Manage the Waiting Experience for Patients: A Longitudinal, Qualitative Study," *BMC Health Services Research*, vol. 19, no. 1, pp. 1-10, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] Diego A. Martinez et al., "Prolonged Wait Time is Associated with Increased Mortality for Chilean Waiting List Patients with Non-Prioritized Conditions," *BMC Public Health*, vol. 19, no. 1, pp. 1-11, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Oleg S. Pianykh et al., "Improving Healthcare Operations Management with Machine Learning," *Nature Machine Intelligence*, vol. 2, no. 5, pp. 266-273, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] A. Joseph et al., "Predicting Waiting Times in Radiation Oncology Using Machine Learning," *16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Cancun, Mexico, pp. 1024-1029, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Hamed Javadifard et al., "Predicting Patient Waiting Time in Phlebotomy Units using a Deep Learning Method," *Innovations in Intelligent Systems and Applications Conference (ASYU)*, Izmir, Turkey, pp. 1-4, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Wei-Chun Lin et al., "Predicting Wait Times in Pediatric Ophthalmology Outpatient Clinic using Machine Learning," *AMIA Annual Symposium Proceedings*, vol. 2019, pp. 1121-1128, 2020. [Google Scholar] [Publisher Link]

[13] Suhaila Zainudin, and Dzulhusni Anjang Ab. Rahman, "Analysis of Outpatient Visit Pattern and Waiting Time Prediction for Selected Public Health Clinics in Selangor," *Proceedings of International Conference on Emerging Technologies and Intelligent Systems*, pp. 603-612, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Amir Rastpour, and Carolyn McGregor, "Predicting Patient Wait Times by Using Highly De-Identified Data in Mental Health Care: Enhanced Machine Learning Approach," *JMIR Mental Health*, vol. 9, no. 8, pp. 1-11, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Katie Walker et al., "Emergency Medicine Patient Wait Time Multivariable Prediction Models: A Multicentre Derivation and Validation Study," *Emergency Medicine Journal*, vol. 39, no. 5, pp. 386-393, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Catherine Curtis MS et al., "Machine Learning for Predicting Patient Wait Times and Appointment Delays," *Journal of the American College of Radiology*, vol. 15, no. 9, pp. 1310-1316, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[17] Lin Lin Guo et al., "Characteristics and Admission Preferences of Pediatric Emergency Patients and Their Waiting Time Prediction Using Electronic Medical Record Data: Retrospective Comparative Analysis," *Journal of Medical Internet Research*, vol. 25, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Yan Sun et al., "Real-Time Prediction of Waiting Time in the Emergency Department, using Quantile Regression," *Annals of Emergency Medicine*, vol. 60, no. 3, pp. 299-308, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[19] Wei-Chun Lin et al., "Secondary Use of Electronic Health Record Data for Prediction of Outpatient Visit Length in Ophthalmology Clinics," *AMIA Annual Symposium Proceedings*, vol. 2018, pp. 1387-1394, 2018. [Google Scholar] [Publisher Link]

[20] Amjed Al-Mousa, Hamza Al-Zubaidi, and Mohammad Al-Dweik, "A Machine Learning-based Approach for Wait-Time Estimation in Healthcare Facilities with Multi-Stage Queues," *IET Smart Cities*, vol. 6, no. 4, pp. 333-350, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Filipe Gonçalves et al., *Predictive Analysis in Healthcare: Emergency Wait Time Prediction*, Ambient Intelligence-Software and Applications-, 9th International Symposium on Ambient Intelligence, Advances in Intelligent Systems and Computing, Springer, Cham, pp. 138-145, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yong-Hong Kuo et al., "An Integrated Approach of Machine Learning and Systems Thinking for Waiting Time Prediction in an Emergency Department," *International Journal of Medical Informatics*, vol. 139, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Abdulkadir Atalan, "Neural Network and Random Forest Algorithms for Estimation of the Waiting Times based on the DES in ED," *1st International Conference on Contemporary Academic Research ICCAR*, Turkey, vol. 1, pp. 14-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Elisabetta Benevento, Davide Aloini, and Nunzia Squicciarini, "Towards a Real-Time Prediction of Waiting Times in Emergency Departments: A Comparative Analysis of Machine Learning Techniques," *International Journal of Forecasting*, vol. 39, no. 1, pp. 192-208, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Mustafa Gökalp Ataman, and Görkem Sarıyer, "Predicting Waiting and Treatment Times in Emergency Departments using Ordinal Logistic Regression Models," *The American Journal of Emergency Medicine*, vol. 46, pp. 45-50, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] Hassan Hijry, and Richard Olawoyin, "Predicting Patient Waiting Time in the Queue System Using Deep Learning Algorithms in the Emergency Room," *International Journal of Industrial Engineering and Operations Management*, vol. 3, no. 1, pp. 33-45, 2021. [Google Scholar] [Publisher Link]

[27] Anton Pak, Brenda Gannon, and Andrew Staib, "Predicting Waiting Time to Treatment for Emergency Department Patients," *International Journal of Medical Informatics*, vol. 145, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[28] Kevin K. Dobbin, and Richard M. Simon, "Optimally Splitting Cases for Training and Testing High Dimensional Classifiers," *BMC Medical Genomics*, vol. 4, no. 1, pp. 1-8, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[29] Mensur Biya et al., "Waiting Time and its Associated Factors in Patients Presenting to Outpatient Departments at Public Hospitals of Jimma Zone, Southwest Ethiopia," *BMC Health Services Research*, vol. 22, no. 1, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[30] Shyamkumar Sriram, and Rakchanok Noochpoung, "Determinants of Hospital Waiting Time for Outpatient Care in India: How Demographic Characteristics, Hospital Ownership, and Ambulance Arrival Affect Waiting Time," *International Journal of Community Medicine and Public Health*, vol. 5, no. 7, pp. 2692-2698, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[31] Deepak Yaduvanshi, Ashu Sharma, and Praful Vijay More, "Application of Queuing Theory to Optimize Waiting-Time in Hospital Operations," *Operations and Supply Chain Management: An International Journal*, vol. 12, no. 3, pp. 165-174, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[32] David W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[33] G.O. Odu, "Weighting Methods for Multi-Criteria Decision-Making Technique," *Journal of Applied Sciences and Environmental Management*, vol. 23, no. 8, pp. 1449-1547, 2019. [CrossRef] [Google Scholar] [Publisher Link]