Prediction of Outpatient No-Show Appointments Using Machine Learning Algorithms for Pediatric Patients in Saudi Arabia

Abdulwahhab Alshammari[1], Fahad Alotaibi[2], Sana Alnafrani[3]
Health Informatics Department, College of Public Health and Health Informatics[1]
Information Technology Department[2], Pediatric Department[3]
King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia[1, 2, 3]
King Abdullah International Medical Research Center, Riyadh, Saudi Arabia[1, 2, 3]
Ministry of the National Guard-Health Affairs, Riyadh, Saudi Arabia[1, 2, 3]

Abstract*—Patient no-shows are prevalent in pediatric outpatient visits, leading to underutilized medical resources, increased healthcare costs, reduced clinic efficiency, and decreased access to care. The use of machine learning techniques provides insights to mitigate this problem. This study aimed to develop a predictive model for patient no-shows at the Ministry of National Guard Health-Affairs, Saudi Arabia, and evaluate the results of various machine learning algorithms in predicting these events. Four machine learning algorithms - Gradient Boosting, AdaBoost, Random Forest, and Naive Bayes - were used to create predictive models for patient no-shows. Each model underwent extensive parameter tuning and reliability assessment to ensure robust performance, including sensitivity analysis and cross-validation. Gradient Boosting achieved the highest area under the receiver operating curve (AUC) of 0.902 and Classification Accuracy (CA) of 0.944, while the AdaBoost model achieved an AUC of 0.812 and CA of 0.927. The Naive Bayes and Random Forest models achieved AUCs of 0.677 and 0.889 and CAs of 0.915 and 0.937, respectively. The confusion matrix demonstrated high true-positive rates for no-shows for the Gradient Boosting and Random Forest models, while Naive Bayes had the lowest values. The Gradient Boosting and Random Forest models were most effective in predicting patient no-shows. These models could enhance outpatient clinic efficiency by predicting no-shows. Future research can further refine these models and investigate practical strategies for their implementation.

Keywords*—*No-show; pediatric; machine learning; algorithms; prediction; outpatients

I. Introduction

Patient no-shows are one of the main challenges in the healthcare sector, disturbing the workflow or affecting cost load, reflecting on the quality and performance [1]. Reducing the number of no-shows significantly impacts healthcare institutions' services, reducing financial costs and effectively utilizing resources to improve patient service [2]. The issue of no-shows is a recurring problem that hinders the efficient utilization of human resources [2,3]. In addition, it increases patient waiting time and negatively impacts the workflow by wasting time for healthcare providers [3].

No-shows are a significant concern for healthcare institutions and may be expensive and inconvenient [3]. Capacity is underutilized, and costly assets are underused [4]. Researchers have found that eliminating non-cancelled no-show appointments may considerably influence productivity, profitability, and clinical outcomes [4]. Machine learning techniques would provide a solution [2]. Thus, finding a way to predict no-shows would facilitate the effective utilization of hospital resources and enhance the satisfaction of both providers and patients, ultimately improving healthcare quality [3–5]. There are some practices used to overcome no-shows, such as overbooking and walk-ins, but these methods are still not the ideal solutions to address the issue of no-shows; there is no effective tool in the electric healthcare system to detect patients at a higher risk of not showing up [6].

Prediction is the most challenging part of human behavior; presuming and predicting this pattern or behavior of a no-show [7]. Finding associations with variables and attributes would facilitate the prediction of no-show appointments [8]. Enabling a prediction model to predict no-shows would help to effectively utilize human resources, reduce financial losses, and increase patient satisfaction [9]. It would also help to improve appointment scheduling, reduce waiting time, and increase the number of patients seen daily [10]. Therefore, a no-show prediction tool is an added value for any organization [11].

A study by Alshammari aims to predict no-shows through machine learning [12]. The dataset includes more than thirty-three million outpatient appointments [12]. The dataset was extracted for a period of nearly three years (January 2016 - July 2019) from the Health Information System (HIS) at all facilities in the central region of the Ministry of National Guard Health Affairs (MNGHA), Saudi Arabia. The authors state that nearly 77,000 outpatient appointments were scheduled monthly at the MNGHA in the Riyadh Region. The patients' ages ranged from 5 to 69 years old. The highest no-show rate was observed among patients over 45 years old. Almost 85% of the no-shows were a national citizen. The study utilized three machine-learning algorithms: Deep Neural Network, AdaBoost, and Naive Bayes. The results of this study were promising, showing that it achieved a 98% precision rate using the deep learning model [12].

The authors of Alshammari's other research paper attempt to develop a prediction model based on a machine learning algorithm for cases where patients do not attend their scheduled appointments [13]. The dataset was obtained from the Kaggle database of hospital appointments booked between April 29, 2016, and June 8, 2016. The dataset included (110,528) medical appointments. Recursive Feature Elimination (RFE) was implemented using Python to exclude unrelated items from the dataset. After running the RFE to remove the unrelated attributes, as including all variables can lead to highly complex modeling, nearly 83,000 appointments were included. The no-show rate was approximately 20%. The dataset has been divided into 70% for the training dataset and 30% for the test dataset. The machine learning algorithms used in this study are Decision Trees and AdaBoost. Multiple variables were used to determine the optimal model for predicting no-shows. The results showed high precision and recall, indicating that the Decision Tree outperformed the AdaBoost results [13].

A study by AlMuhaideb used machine learning to create a model for predicting no-shows in outpatients [14]. The research data were extracted from the health information system, which captures records of patient visits for outpatients. The dataset contains almost more than 1 million outpatient records. The period of this dataset was between January and December 2014, and the no-show rate was 11.3%. The machine prediction models used are JRip and Hoeffding tree algorithms. The machine learning software used was Weka. The dataset was cleansed and preprocessed to conduct the modeling analysis. Both the JRip and Hoeffding algorithms provided rational degrees of accuracy levels of almost 77%. The study showed that the no-shows could be predicated using a machine-learning model [14].

Hamdan, A., and Abu Bakar, A. published a study in 2023 on outpatient no-show appointments in a Malaysian tertiary hospital [15]. The study aimed to develop a model for predicting patient no-show appointments using machine learning. The data were collected through 2019 and included 246,943 appointment records with 14 attributes, including demographics and appointment data. The result shows that 69,173 patients did not attend their appointment, which accounts for about 28% of the dataset. The machine learning model used seven algorithms: logistic regression (LR), decision tree (DT), k-nearest neighbors (k-NN), Naïve Bayes (NB), random forest (RF), gradient boosting (GB), and multilayer perceptron (MLP). Three different train and testing splits were applied at 60:40, 70:30, and 80:20, and ten folding validations were performed on each split using Python. The evaluation metrics included accuracy, AUC value, and F1 score. The GB scored the highest accuracy of 78%.

Therefore, finding a way to predict the no-show or high no-show candidates will help healthcare organizations overcome this issue [16]. Developing a prediction model will help stakeholders mitigate the anticipated effects of no-shows and enhance healthcare efficiency by optimizing resource utilization [17]. The prediction model can help to identify patients at high risk of no-shows based on factors such as age, gender, appointment type, past behavior, and geographic location [18]. A machine learning model that predicts patient no-shows can enhance clinical efficiency by optimizing resource allocation, reducing wasted time through overbooking appointments without compromising patient care, and allocating efficient slot allocation by understanding no-show patient patterns.

The main objective of this study is to develop a machine-learning model capable of accurately predicting the likelihood of a pediatric patient missing a scheduled appointment. Other objectives include identifying key factors influencing pediatric no-show rates to inform targeted interventions and optimize appointment scheduling and resource allocation based on no-show predictions. This model can improve patient satisfaction by reducing wait times and increasing appointment availability. From a broader perspective, this model can help better understand pediatric patient behavior and healthcare utilization.

In this study, we aim to develop a predictive model and evaluate its performance using machine learning algorithms to predict pediatric patient no-shows in pediatric outpatient visits at the Ministry of National Guard - Health Affairs (MNGHA) using machine learning techniques. This study differs from the mentioned studies [12–15]. This study targets a more specific population of pediatric patients. It uses primary data extracted from a tertiary hospital and applies multiple or different machine learning algorithms in a single study.

II. Method

This study is a retrospective exploratory/predictive study. It aims to predict the no-show based on machine learning techniques and Machine Learning (ML) algorithms such as Gradient Boosting (GB), AdaBoost, Naive Bayes (NB), and Random Forest, which are all supervised Machine Learning algorithms.

This study was conducted ethically, following established guidelines and protocols to ensure patient privacy and data confidentiality. This study has been approved by an Institutional Review Board (IRB) committee from the King Abdullah International Medical Research Center (KAIMRC).

A. Study Area, Settings, and Subjects

The study was conducted on pediatric patients at the Ministry of National Guard Health Affairs (MNGHA) in Saudi Arabia. The data were extracted from the BESTCare health information system used in the MNGHA. They encompass appointments scheduled throughout the day from January 1, 2021, to May 5, 2022.

The patient records eligible for the study must fulfill the inclusion and exclusion criteria. All patients under the age of 14 who had a scheduled visit to a pediatric outpatient clinic (pediatric patients) were included. A patient no-show is a visit in which the patient fails to attend a scheduled appointment without providing prior notice. Canceled appointments before the clinic were not counted as no-shows to ensure all missed appointments were not intervenable. Emergency visits, unscheduled visits, such as walk-ins, and patients older than 14 years were excluded.

B. Data Collection, Management, and Analysis Plan

The dataset used in this study comprises 358,759 outpatient appointment visits, with a mix of nominal, ordinal, and numeric attributes related to patient demographics, appointment details, and medical history. The dataset includes data on patients' age groups, gender, nationality, appointment types, region, and appointment times to ensure a representative sample of the study population across different age groups, genders, and appointment types.

This study utilizes historical data on patient visits to predict the likelihood of no-shows. The analysis plan encompasses the standard stages of the data mining process, including data collection and understanding, data preparation, model selection, model building, and model evaluation.

The study collected various attributes from the dataset to analyze and predict patient no-show appointments. Table I summarizes these attributes, including their descriptions and types. The attributes capture information such as visit ID, region, facility, department, clinic, patient demographics (gender, nationality, age), appointment details (date, time), diagnosis information, appointment message status, patient's address, sponsor eligibility, and more. One notable attribute is the lead time, which represents the difference between the booking and appointment dates. This table references the attributes used in analyzing and predicting no-show appointments.

The data was cleaned and preprocessed to ensure the quality of the dataset. This included handling missing values, removing irrelevant attributes, and transforming variables required for model development. For example, the lead time variable was derived by calculating the difference between the booking and appointment dates. In addition, the appointment time was categorized into AM and PM.

Relevant features were selected based on their potential impact on predicting no-show appointments. Factors such as age group, gender, nationality, appointment type, region, and appointment time were included in the analysis, as these were expected to contribute to the prediction of no-show appointments.

TABLE I. DESCRIPTION OF THE COLLECTED AND DERIVED ATTRIBUTES

Data Attributes			
No.	**Attribute Name**	**Description**	**Type**
1	Visit_ID	Visit ID	Numeric
2	Region	Region Name	Nominal
3	Facility	Facility Name	Nominal
4	HSP_TP_CD	Hospital or facility type code	Numeric
5	HSPI_TP_CD	Internal hospital or facility code	Numeric
6	Department	Medical Department	Nominal
7	Department_CD	Medical Department Code	Numeric
8	Clinic	Clinic name	Nominal
9	Clinic_CD	Clinic Code	Numeric
10	MRN	Patient's Medical record number	Numeric
11	Appointment_DT	Appointment Date	Ordinal
12	Appointment_TIME	Appointment Time	Ordinal
13	Appointment_DTM	Appointment Date & Time	Ordinal
14	Visit_Type	Patient Visit type	Nominal
15	Appointment_Booking_DTM	Appointment Booking Date & Time	Ordinal
16	Appointment_Booking_TIME	Appointment Booking Time	Ordinal
17	ICD10_CD	ICD10 code for the diagnosis	Nominal
18	Diagnosis	Diagnosis	Nominal
19	Flag	Show/no-show	Nominal
20	MSG_SENT_YN	Appointment message sends the status	Nominal
21	MSG_Status	Appointment message status	Nominal
22	Gender	Patient Gender	Nominal
23	Nationality	Patient Nationality	Nominal
24	Age	Patient Age	Ordinal
25	Address1	Patient Region or an area name or code of the patient's residence	Nominal
26	Address2	Patient district name or code of patient's residence	Nominal
27	Sponsor_Eligibility	Patient's Sponsor_Eligibility	Nominal
28	ETPR_PT_NO	Patient Enterprise record number	Numeric
29	Cachement_Area_CD	Area name of the patient's residence	Numeric
30	Cachement_Area_NAME	Area Code of the patient's residence	Nominal
31	Cachement_FCLT_NO	Facility Code of the patient's residence	Numeric
32	Cachement_FCLT_NAME	Facility name of the patient's residence	Nominal
33	Lead time	Difference between Appointment Booking and Appointment dates	Numeric
34	Appointment time AM/PM	AM/PM	Ordinal

C. Model Selection, Building, and Evaluation

We applied four machine learning algorithms to the preprocessed data: Gradient Boosting, AdaBoost, Naive Bayes, and Random Forest. Gradient Boosting, AdaBoost, Naive Bayes, and Random Forest are flexible, well-suited algorithms for handling complex relationships in large datasets. These algorithms are also well-suited for handling categorical features and numerical data [19,20]. We used 10-fold cross-validation to assess model performance and avoid overfitting. Each model evaluation was based on various metrics, including Area Under the Receiver Operating Characteristic Curve (AUC), Classification Accuracy (CA), F1 score, Precision, and Recall. Furthermore, the preprocessing steps and hyperparameters for each model were recorded, ensuring the integrity and consistency of the input data.

Gradient Boosting (XGBoost): The model was trained and constructed by combining 100 individual decision trees and a learning rate of 0.3, which determines the weight given to each tree's prediction when they are combined. In this case, a learning rate of 0.3 balances responsiveness and stability in the model's predictions. The maximum depth of individual trees was set at 20, providing a good balance between the model's complexity and its ability to learn the underlying patterns in the data.

Regularization was applied with a lambda value of 7 to prevent overfitting. Regularization helps prevent overfitting when a model becomes too complex and starts to memorize the training data instead of learning the underlying patterns.

The lambda value of 7 represents the strength of the regularization. A higher lambda value increases the penalty for complex models, encouraging the model to simplify its predictions and avoid overfitting. By applying regularization with a lambda value of 7, the model aims to balance capturing essential patterns in the data and avoiding excessive complexity.

We also experimented with different fractions of training instances and features for each tree, level, and split. The fraction was set to 1.0 in all cases to use all available data and features. We fixed the random seed for replicable training to ensure that specific conditions did not affect our results.

The preprocessing steps for Gradient Boosting were removing instances with unknown target values, customizing categorical variables using one-hot-encoding, removing empty columns, and imputing missing values with mean values.

AdaBoost: The model was built using a base estimator (a decision tree) and 100 additional estimators. A high learning rate of 0.999 was set to give more weight to the most recent data. The classification algorithm SAMME was used to boost the model. As in the case of XGBoost, we ensured the replicability of results by fixing the random seed.

AdaBoost's preprocessing steps included removing instances with unknown target values, customizing categorical variables using one-hot encoding, removing empty columns, and imputing missing values with mean values.

Naive Bayes: The Naive Bayes algorithm does not have specific hyperparameters like other algorithms, but preparing the data well for this model is essential. For Naive Bayes, the preprocessing step was removing empty columns.

Random Forest: The model was trained with 100 trees, and the number of attributes considered at each split was set to 5. This allowed the model to consider a balanced number of attributes at each node to achieve a good compromise between bias and variance. Growth control measures were applied to avoid creating complex models that could lead to overfitting. We ensured that subsets smaller than a certain threshold were not split.

Preprocessing of the Random Forest included removing instances with unknown target values, customizing categorical variables using one-hot encoding, removing empty columns, and imputing missing values with mean values.

III. Results and Analysis

This section presents the research project's findings based on the statistical analysis and the development of the machine learning model described in the previous section. The results are presented as descriptive statistics, model performance, and critical findings from the analysis.

A. Description Analysis

This section provides an overview of the dataset, showing patterns and trends in no-show appointments among different patient demographics and appointment attributes.

1) Age group: In Table II, the data show different age groups, including infants (0-12 months), toddlers (1-3 years), preschoolers (3-6 years), school-age children (6-12 years), and adolescents (12-14 years). The table displays the number of patients who attended their appointments and those who did not (no-shows) for each age group. The "Show" column represents the number of patients who attended their appointments, while the "No-Show" column represents the number of patients who did not show up. The "Total" column indicates the total number of patients in each age group. The school-age children (6-12 years) accounted for the highest proportion of no-show appointments, followed by infants (0-12 months) and preschoolers, see Table II.

TABLE II. Description and Distribution of Age Groups with No-Show Percentages

Age Groups Categories			
Age Group	**Show (%)**	**No-show (%)**	**Total**
Infant (0-12 Months)	79,948 (92.7%)	6,281 (7.3%)	86,229
Toddler (1-3 Years)	53,150 (92.1%)	4,460 (7.9%)	57,610
Preschool (3-6 Years	59,248 (91.1%)	5,740 (8.9%)	64,988
School-age (6-12 Years)	101,212 (91.1%)	9,831 (8.9%)	111,043
Adolescent (12-14 Years)	35,264 (90.6%)	3,624 (9.4%)	38,888

2) Gender: Table III presents the data for gender distribution. The table displays the number of patients who attended their appointments and those who did not (no-shows) for each gender. The dataset contained a higher proportion of shows with male patients than with female patients.

TABLE III. DISTRIBUTION OF AND PERCENTAGES OF GENDER

Gender and Distribution			
Gender	**Show (%)**	**No-show (%)**	**Total**
Male	171,469 (91.9%)	15,382 (8.2%)	186,851
Female	157,353 (91.5%)	14,554 (8.5%)	171,907

3) Nationality: Table IV presents the data for individuals of Saudi and non-Saudi nationality. The majority of patients were of Saudi Nationality, with a small percentage of non-Saudi patients. The rate of no-shows is higher for non-Saudi patients.

TABLE IV. DISTRIBUTION OF AND PERCENTAGES OF NATIONALITY

Nationality and Distribution			
Nationality	**Show (%)**	**No-show (%)**	**Total**
Saudi	325,944 (91.7%)	29,409 (8.3%)	355,353
Non-Saudi	2,878 (84.6%)	527 (15.4%)	3,405

4) Appointment types: Table V presents data on different appointment types, including New Patient (NP), First Visit (FV), and Follow-up (FU). Follow-up appointments had the highest no-show rates, followed by first visits and new patient appointments.

TABLE V. DISTRIBUTION AND PERCENTAGES OF APPOINTMENT TYPES

Appointment Types and Distribution			
Appointment Type	**Show (%)**	**No-show (%)**	**Total**
New Patient (NP)	10,182 (95.9%)	446 (4.1%)	10,628
First visit (FV)	171,415 (92.6%)	13,681 (7.4%)	185,096
Follow-up (FU)	147,225 (90.3%)	15,809 (9.7%)	163,034

Table VI shows the data for the Central, Eastern, and Western regions, showing attendance and no-show numbers. Geographically, the proportion of no-shows was highest in the Central region, followed by the Western and Eastern regions.

TABLE VI. DISTRIBUTION OF AND PERCENTAGES OF INCLUDED REGIONS

Regions and Distribution			
Region Name	**Show (%)**	**No-show (%)**	**Total**
Central	218,151 (93.9%)	13,994 (6.1%)	232,145
Eastern	32,206 (93.1%)	2,388 (6.9%)	34,594
Western	78,465 (85.2%)	13,554 (14.8%)	92,019

5) Appointment time and hours: Table VII provides information on the appointment times of the day, namely AM and PM, indicating the number of patients who showed up and those who did not. The data showed that appointments in the morning (AM) had slightly higher no-show rates than afternoon (PM) appointments.

TABLE VII. DISTRIBUTION OF AND PERCENTAGES OF APPOINTMENT TIME OF THE DAY

Time of the Day and Distribution			
Gender	**Show (%)**	**No-show (%)**	**Total**
AM	166,642 (90.8%)	16,737 (9.2%)	183,379
PM	162,180 (92.5%)	13,199 (7.5%)	175,379

Table VIII provides information on appointment hours, including the number of no-shows and appointments attended for each period. The table displays data for time slots 7-9, 9-12, 12-15, 15-17, and beyond working hours. Table IX presents data on appointment days of the week, showing the number of no-shows and shows for each day.

TABLE VIII. DISTRIBUTION OF APPOINTMENT HOURS CATEGORIES

Appointment Hours Categories of the Day			
Appointment Hour	**Show (%)**	**No-show (%)**	**Total**
7-9	3,855 (8.7%)	40,597 (91.3%)	44,452
9-12	12,867 (9.6%)	121,583 (90.4%)	134,450
12-15	9,221 (7.9%)	106,923 (92.1%)	116,144
15-17	3,659 (9.8%)	33,723 (90.2%)	37,382
After working hours	334 (1.3%)	25,996 (98.7%)	26,330

TABLE IX. DISTRIBUTION OF APPOINTMENT DAY

Appointment Day			
Appointment Day	**Show (%)**	**No-show (%)**	**Total**
Sunday	6,040 (8.4%)	66,002 (91.6%)	72,042
Monday	6,495 (8.2%)	72,645 (91.8%)	79,140
Tuesday	7,423 (9.9%)	67,213 (90.1%)	74,636
Wednesday	5,662 (7.8%)	66,647 (92.2%)	72,309
Thursday	4,168 (7.7%)	49,965 (92.3%)	54,133
Friday	85 (2.8%)	2,958 (97.2%)	3,043
Saturday	63 (1.8%)	3,392 (98.2%)	3,455

B. Model Performance

The results suggest that all four models (Gradient Boosting (GB), AdaBoost, Naive Bayes (NB), and Random Forest) have performed reasonably well, but Gradient Boosting stood out as the most robust model in our study. Our sensitivity analysis, which involved varying the training and testing data splits and the machine learning algorithms' hyperparameters enhanced the reliability of our results. Comprehensive evaluation ensures that the models are reliable and not overly sensitive to the specific selection of parameters, confirming the robustness of the findings. The consistency in preprocessing across models further strengthens the credibility of the outcomes. The performance of the four machine learning models was evaluated using cross-validation with ten subsets and a separate testing set.

Table X summarizes the model performance, with Gradient Boosting demonstrating the best performance by achieving the highest values for AUC, CA, F1 score, precision, and recall. This indicates that Gradient Boosting outperformed the other models in accurately predicting the likelihood of no-show appointments.

TABLE X. MODEL'S ALGORITHMS AND EVALUATION METRICS PERFORMANCE

Algorithms and Evaluation Metrics					
Algorithm	**AUC**	**CA**	**FI**	**Precision**	**Recall**
Gradient Boosting	0.902	0.944	0.937	0.939	0.944
AdaBoost	0.812	0.927	0.926	0.924	0.927
Naive Bayes	0.677	0.915	0.877	0.861	0.915
Random Forest	0.889	0.937	0.925	0.931	0.937

The Receiver Operating Characteristic (ROC) analysis is a graphical representation that illustrates the performance of a binary classification model. In this study, the ROC plot (Fig. 1) demonstrates the performance of the machine learning models in predicting no-show appointments.

The x-axis represents the False Positive Rate (FPR), which measures the proportion of false positives (show appointments incorrectly classified as no-shows) to all actual negatives (show appointments). The y-axis represents the True Positive Rate (TPR), which measures the proportion of true positives (no-show appointments correctly classified as a no-show) to all actual positives (no-show appointments).

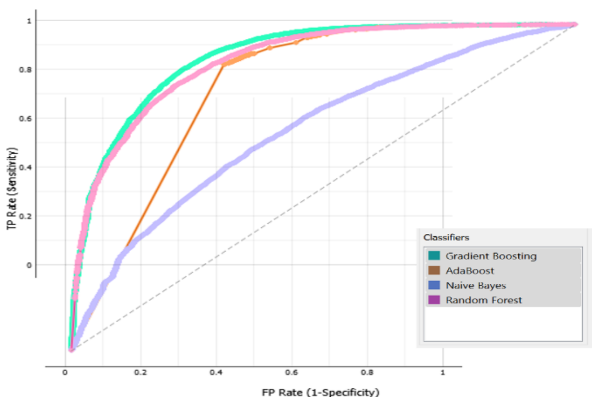

Fig. 1. ROC analysis diagram with classifier.

A curve on the ROC plot represents each model. The closer the curve is to the top-left corner, the better the model's performance. The ideal scenario is a model with a curve that reaches the top-left corner, indicating a high TPR and a low FPR.

By examining the ROC plot, we can observe that the Gradient Boosting model exhibits the highest performance among the four models. It shows the highest TPR for a given FPR threshold, indicating its ability to identify and classify no-show appointments accurately. The other models, including AdaBoost, Naive Bayes, and Random Forest, also demonstrate varying performance levels, with their respective curves positioned below that of Gradient Boosting.

The ROC Analysis Diagram visually represents the models' performance, distinguishing between show and no-show appointments. It helps evaluate and compare the predictive capabilities of models and select the most suitable one for accurately predicting no-shows in future scenarios.

Table XI shows the confusion matrix for the four represented models and the predicted versus actual outcomes for no-show and show appointments. The matrices provide insight into each model's performance by showing the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) rates.

TABLE XI. DISTRIBUTION OF APPOINTMENT DAY (CONFUSION MATRIX FOR THE FOUR REPRESENTED MODELS)

Confusion Matrix			
Gradient Boosting (GB) Algorithm			
		Actual Values	
		Positive	**Negative**
Predicted Values	Positive	TP (80.4 %)	FP (4.9 %)
	Negative	FN (19.6 %)	TN (95.1 %)
AdaBoost Algorithm			
		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP (56.9 %)	FP (4.3 %)
	Negative	FN (43.1 %)	TN (95.7 %)
Naïve Bayes			
		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP (24.8 %)	FP (8.3 %)
	Negative	FN (75.2 %	TN (91.7 %)
Random Forest			
		Actual Values	
		Positive	**Negative**
Predicted Values	Positive	TP (81.3 %)	FP (5.8 %)
	Negative	FN (18.7 %)	TN (94.2 %)

1) Gradient Boosting (GB):

- The Gradient Boosting model delivered the highest performance among all the evaluated models with an AUC of 0.902, CA of 0.944, F1 score of 0.937, precision of 0.939, and recall of 0.944.
- The model accurately predicted no-show appointments at 80.6% and showed appointments at 95.1%.
- The model misclassified 4.9% of actual no-show appointments and 19.6% of existing show appointments as no-show appointments.
- Gradient Boosting (GB) demonstrated high accuracy in predicting no-show appointments, with 80.4% True Positives (correctly predicted no-shows) and 95.1% True Negatives (correctly predicted shows). However, the model also exhibited a 19.6% False Positive rate (incorrectly predicted no-shows) and a 4.9% False Negative rate (incorrectly predicted shows).

2) AdaBoost:

- This model achieved an AUC of 0.812, CA of 0.927, F1 score of 0.926, precision of 0.924, and recall of 0.927.

- The model accurately predicted no-show appointments at 56.8% and showed appointments at 95.7%.
- The model misclassified 4.3% of actual no-show appointments and 43.1% of existing show appointments as no-show appointments.
- The AdaBoost model yielded 56.9% True Positives and 95.7% True Negatives, exhibiting 43.1% False Positive and 4.3% False Negative rates. This performance suggests a moderate ability to predict no-show appointments correctly.

3) Naive Bayes (NB):

- This model had an AUC of 0.677, CA of 0.915, F1 score of 0.877, precision of 0.861, and recall of 0.915.
- The model accurately predicted no-show appointments at 24.8% and showed appointments at 91.7%.
- The model misclassified 8.3% of actual no-show appointments and 75.2% of existing show appointments as no-show appointments.
- The Naive Bayes (NB) model achieved a lower accuracy in predicting no-show appointments, with a 24.8% True Positive rate and a 91.7% True Negative rate. The model had a high False Positive rate of 75.2% and a False Negative rate of 8.3%.

4) Random Forest (RF):

- This model reported an AUC of 0.889, CA of 0.937, F1 score of 0.925, precision of 0.931, and recall of 0.937.
- The model accurately predicted no-show appointments at 81.3% and showed appointments at 94.2%.
- The model misclassified 5.8% of actual no-show appointments and 18.4% of existing show appointments as no-show appointments.
- The RF model strongly predicted no-show appointments, with an 81.3% True Positive rate and a 94.2% True Negative rate. The model had an 18.4% False Positive rate and a 5.8% False Negative rate.

The Gradient Boosting model exhibits effective performance with a high true positive rate for no-show and show appointments and relatively low misclassification rates. Specifically, it correctly identified 80.4% of no-show appointments and 95.1% of show appointments. This level of performance indicates a strong ability of this model to distinguish between the two classes accurately.

On the contrary, the Naive Bayes model demonstrates the poorest performance, with the lowest true positive rate for no-show appointments (24.8%) and a relatively lower success rate for show appointments (91.7%). It had the highest misclassification rate for show appointments, signaling potential weaknesses in the model's ability to identify true positives in a balanced manner correctly.

The Random Forest model showed strong performance with a high true positive rate for no-show appointments (81.3%) and a high success rate for show appointments (94.2%). This reflects a balanced performance for both classes, making it a reliable model for this prediction task.

Finally, while not as proficient as Gradient Boosting or Random Forest, the AdaBoost model still showed a reasonable true positive rate for no-show appointments (56.9%) and a high success rate for show appointments (95.7%).

In conclusion, based on these results, the Gradient Boosting and Random Forest models demonstrate superior performance in predicting no-show appointments compared to the AdaBoost and Naive Bayes models. This comprehensive evaluation gives insights into each model's strengths and weaknesses. It provides valuable information for selecting the most suitable model for predicting no-show appointments in future studies.

IV. Discussion

The primary aim of this study was to evaluate and predict no-show appointments at MNGHA pediatric outpatient visits using machine learning models. Our research builds upon the existing literature by developing a predictive model specifically for pediatric outpatient settings, focusing on a large dataset that includes demographic, appointment-related, and geographic factors.

Our findings contribute to the existing body of knowledge by identifying patterns and trends in no-show appointments across various patient demographics and appointment attributes. This information can help healthcare providers better understand the factors contributing to no-show appointments and develop targeted strategies for reducing no-show rates [21,22].

The results of this study indicate that the GB and RF models outperformed other models in predicting no-show appointments. This superior performance can be attributed to the model's ability to capture complex relationships between various features in the dataset, making it particularly suitable for our research objective.

The strengths of this study include the large and diverse dataset, which allowed us to develop a robust and reliable predictive model. Moreover, the use of multiple machine learning models and the implementation of cross-validation for model evaluation ensure the validity of our findings [23].

A. Limitation

First, the data used in this study were limited to a single healthcare organization, "MNGHA," which may not be representative of other pediatric outpatient settings [24]. Future research could consider including data from multiple healthcare systems to further validate the predictive model's generalizability. Second, the dataset should have included certain factors such as socioeconomic status, transportation availability, weather conditions, and patient preference; including these factors might enhance the model's predictive capabilities [25].

Future research directions could involve the following:

1) Expanding the dataset to include additional pediatric outpatient settings to validate the predictive model's performance across different healthcare organizations [24].

2) Incorporating other relevant factors, such as socioeconomic status, transportation availability, and weather conditions, further enhances the model's predictive capabilities [25,26].

3) Investigating the potential impact of targeted interventions, such as appointment reminders or personalized follow-up, on reducing no-show rates based on the predictive model's output [25–27].

In conclusion, this study's results provide valuable insights into the factors associated with no-show appointments in pediatric outpatient settings. The machine learning model developed can aid healthcare providers in predicting no-show appointments, optimizing resource management, and improving patient care.

V. Conclusion

The primary contribution of this research project is developing a machine learning model to predict no-show appointments in pediatric outpatient settings. Our study has identified patterns and trends in no-show appointments by analyzing a large and diverse dataset. This analysis can assist healthcare providers in optimizing resource management and enhancing patient care. The GB and RF models emerged as the best performers in predicting no-show appointments, demonstrating their potential utility in pediatric outpatient settings.

Our findings build upon existing literature, highlighting the importance of understanding factors contributing to no-show appointments in pediatric populations. These insights can guide healthcare providers in developing targeted strategies for reducing no-show rates and enhancing overall healthcare delivery.

While our study has some limitations, such as focusing on a single healthcare system and excluding certain factors, it provides a solid foundation for future research. Expanding the dataset to include additional pediatric outpatient settings, incorporating other relevant factors, and investigating the impact of targeted interventions based on the predictive model's results may further deepen the understanding of no-show appointments and help improve healthcare management.

In conclusion, this study provides valuable insight into the factors associated with no-show appointments in outpatient pediatrics. It gives healthcare providers a powerful tool for effectively predicting and managing missed appointments. Through continued research and model improvement, we can further enhance our understanding of no-show appointments and optimize resource allocation in outpatient pediatric care.

Acknowledgment

The authors would like to acknowledge Reem Alamr, editor at the Office of Research, King Saud Bin Abdulaziz University for Health Sciences, for her contribution to editing and proofreading the manuscript.

References

[1] Huang Y, Hanauer DA. Patient no-show predictive model development using multiple data sources for an effective overbooking approach. Appl Clin Inform. 2014;5(3):836–60.

[2] Denney J, Coyne S, Rafiqi S. Machine Learning Predictions of No-Show Appointments in a Primary Care Setting. SMU Data Sci Rev. 2019;2(1):1–32.

[3] Norris JB, Kumar C, Chand S, Moskowitz H, Shade SA, Willis DR. An empirical investigation into factors affecting patient cancellations and no-shows at outpatient clinics. Decis Support Syst. 2014 Jan;57:428–43.

[4] Huang Z, Ashraf M, Gordish-Dressman H, Mudd P. The financial impact of clinic no-show rates in an academic pediatric otolaryngology practice. Am J Otolaryngol. 2017;38(2):127–9.

[5] Mochón F, Elvira C, Ochoa A, Gonzalvez JC. Machine-Learning-Based No Show Prediction in Outpatient Visits. Int J Interact Multimed Artif Intell. 2018;4(Special Issue on Big Data and e-Health):29–34.

[6] Devasahay SR, Karpagam S, Ma NL. Predicting appointment misses in hospitals using data analytics. mHealth. 2017 Apr;3(12):1–9.

[7] Goffman RM, Harris SL, May JH, Milicevic AS, Monte RJ, Myaskovsky L, et al. Modeling Patient No-Show History and Predicting Future Outpatient Appointment Behavior in the Veterans Health Administration. Mil Med. 2017 May;182(5):e1708–14.

[8] Dantas LF, Hamacher S, Cyrino Oliveira FL, Barbosa SDJ, Viegas F. Predicting Patient No-show Behavior: a Study in a Bariatric Clinic. Obes Surg. 2019 Jan;29(1):40–7.

[9] Nelson A, Herron D, Rees G, Nachev P. Predicting scheduled hospital attendance with artificial intelligence. NPJ Digit Med. 2019 Apr;2(26):1–7.

[10] Roy S, Gupta A, Datta S, Das A, Pradhan S, Das T, et al. Prediction of heart diseases using machine learning. AIP Conf Proc. 2023 Nov;2851(1):020001.

[11] Alhamad Z. Reasons for missing appointments in general clinics of primary health care center in Riyadh Military Hospital, Saudi Arabia. Int J Med Sci Public Health. 2013;2(2):258–67.

[12] Alshammari R, Daghistani T, Alshammari A. The Prediction of Outpatient No-Show Visits by Using Deep Neural Network from Large Data. Int J Adv Comput Sci Appl. 2020;11(10):533–9.

[13] Alshammari A, Almalki R, Alshammari R. Developing a Predictive Model of Predicting Appointment No-Show by Using Machine Learning Algorithms. J Adv Inf Technol. 2021;12(3):234–9.

[14] AlMuhaideb S, Alswailem O, Alsubaie N, Ferwana I, Alnajem A. Prediction of hospital no-show appointments through artificial intelligence algorithms. Ann Saudi Med. 2019 Dec;39(6):373–81.

[15] Ahmad Hamdan AF, Abu Bakar A. Machine Learning Predictions on Outpatient No-Show Appointments in a Malaysia Major Tertiary Hospital. Malays J Med Sci MJMS. 2023 Oct;30(5):169–80.

[16] Dashtban M, Li W. Deep Learning for Predicting Non-attendance in Hospital Outpatient Appointments. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019. p. 3731–40.

[17] Harvey HB, Liu C, Ai J, Jaworsky C, Guerrier CE, Flores E, et al. Predicting No-Shows in Radiology Using Regression Modeling of Data Available in the Electronic Medical Record. J Am Coll Radiol JACR. 2017 Oct;14(10):1303–9.

[18] Carreras-García D, Delgado-Gómez D, Llorente-Fernández F, Arribas-Gil A. Patient No-Show Prediction: A Systematic Literature Review. Entropy. 2020 Jun 17;22(6):675.

[19] Tanha J, Abdi Y, Samadi N, Razzaghi N, Asadpour M. Boosting methods for multi-class imbalanced data classification: an experimental review. J Big Data. 2020 Sep 1;7(1):70.

[20] Wickramasinghe I, Kalutarage H. Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation. Soft Comput. 2021 Feb 1;25(3):2277–93.

[21] Mohammadi I, Wu H, Turkcan A, Toscos T, Doebbeling BN. Data Analytics and Modeling for Appointment No-show in Community Health Centers. J Prim Care Community Health. 2018 Nov 17;9:2150132718811692.

[22] Daggy J, Lawley M, Willis D, Thayer D, Suelzer C, DeLaurentis PC, et al. Using no-show modeling to improve clinic performance. Health Informatics J. 2010 Dec;16(4):246–59.

[23] Lenzi H, Ben ÂJ, Stein AT. Development and validation of a patient no-show predictive model at a primary care setting in Southern Brazil. PLoS ONE. 2019 Apr;14(4):e0214869.

[24] Joseph J, Senith S, Kirubaraj AA, Ramson SRJ. Machine Learning for Prediction of Clinical Appointment No-Shows. Int J Math Eng Manag Sci. 2022 Jul;7(4):558–74.

[25] Daghistani T, AlGhamdi H, Alshammari R, AlHazme RH. Predictors of outpatients' no-show: big data analytics using apache spark. J Big Data. 2020 Dec;7(108):1–15.

[26] Lee G, Wang S, Dipuro F, Hou J, Grover P, Low LL, et al. Leveraging on Predictive Analytics to Manage Clinic No Show and Improve Accessibility of Care. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA) [Internet]. IEEE; 2017 [cited 2024 Apr 29]. p. 429–38. Available from: https://ieeexplore.ieee.org/document/8259804.

[27] Alaeddini A, Yang K, Reeves P, Reddy CK. A hybrid prediction model for no-shows and cancellations of outpatient appointments. IIE Trans Healthc Syst Eng. 2015;5(1):14–32.