The Prediction of Pediatric Outpatient No-Show Visits by Employing Machine Learning Framework

Abdulwahhab Alshammari[1], Hanoof Alaboodi[2], Riyad Alshammari[3]
College of Public Health and Health Informatics,
King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia[1, 2]
King Abdullah International Medical Research Center, Riyadh, Saudi Arabia[1, 2]
Ministry of the National Guard-Health Affairs, Riyadh, Saudi Arabia[1, 2]
Pediatric Department, Ministry of the National Guard-Health Affairs, Riyadh, Saudi Arabia[2]
National Centre for Artificial Intelligence (NCAI),
Saudi Data and Artificial Intelligence Authority (SDAIA), Riyadh, Saudi Arabia[3]

***Abstract*—Patient no-show for a booked medical appointment is a significant problem that negatively impacts healthcare resource utilization, cost, efficiency, quality, and patient outcomes. This paper developed a machine learning framework to predict pediatric patients' no-shows to medical appointments accurately. Thirty months of outpatient visits data were extracted from data warehouse from January 2017 to July 2019 of the Ministry of National Guard Health Affairs (MNGHA), Saudi Arabia. The researchers retrieved the data from all healthcare facilities in the central region, and more than 100 attributes were generated. The data includes over 100,000 pediatric patients and more than 3.7 million visits. Five machine learning algorithms were deployed, where Gradient Boosting (GB) algorithm outperformed the other four machine learning algorithms: decision tree, random forest, logistic regression, and neural network. The study evaluated and compared the performance of the five models based on five evaluations criteria. GB achieved a Receiver Operating Characteristic (ROC) score of 97.1%. Furthermore, this research paper identified the factors that have massive potential for effecting patients' adherence to scheduled appointments.**

***Keywords*—No-show; machine learning; healthcare medical appointments; predictive analytics**

I. Introduction

Outpatient no-show in a healthcare setting is the non-attendance of confirmed medical appointment by the patient [1-2]. The outpatient medical appointment no-show is a significant issue for healthcare facilities since utilizing resources is ineffective and widens the appointment waiting time. Thus, accurately predicting the patient no-show will reduce the financial cost, increase productivity, enhance the quality of care noticeably, and increase the patient's satisfaction [2-8] considerably.

The demand for outpatient medical services in the Kingdom of Saudi Arabia (KSA) is increasing substantially [5]. There is an average of 70K medical appointments at the Ministry of National Guard Health Affairs (MNGHA) central region medical, of which around 10% are related to pediatric cases. Presently, there is no effectible deployed digital tool in the Electronic Health Records (EHR) that can accurately predict the patient of high risk of no-show [5-6, 9]. As a result, the healthcare in KSA could utilize the advancements in machine learning techniques to construct a digital solution that can identify outpatients with a high possibility of not attending their medical reservation.

Machine learning can be used to build an intelligent digital solution that can provide healthcare with a data-driven system to enhance the patient's adherence to the medical appointment and predict patient behavior toward non-attendance of a scheduled medical appointment. Historical collected data in the EHR system can be utilized to forecast future patient visits. Therefore, an intelligent digital health solution can allow healthcare facilities to strategically create and manage a long-term projection plan for their medical resources [10-16].

This research study aims to develop an intelligent data-driven approach based on a machine learning technique to learn from more than three million extracted pediatric medical records from MNGH database systems to predict outpatient no-show medical appointments smartly. Other objectives are to predict patient behavior and identify the aspects that can be useful to forecast no-shows.

The remaining research paper is organized as follows. Section II covers the related work, while Section III details the methods. The results and discussion are presented in Section IV. Finally, Section V gives the conclusions and future work.

II. Related Work

This section summarized machine learning solutions to predict no-shows in healthcare facilities. Most of the articles in the literature focused on adult cases. Harvey et al. [17] used a logistic regression algorithm to predict the non-appearance of patients from booked radiology examinations. They achieved 75.3% for the Area Under the Receiver (AUC) operator. Chua and Chow [18] extracted a no-show administrative dataset and employed Multiple Logistic Regression (MLR). They scored 72% for AUC. Dantas et al. [19] collected more than 13K records related to bariatric surgery no-show appointments. They achieved an accuracy of 71% using the Logistic Regression algorithm. Likewise, Kurasawa et al. [20] used a Logistic Regression algorithm to forecast the no-show of diabetic patients. Their best model scored 75.7% and 65.9% for precision and recall, respectively.

Furthermore, Mohammadi et al. [21] achieved the best result of 86% of AUC using Naive Bayes on a dataset of 73,811 records, while AlMuhaideb et al. [22] achieved 86.1% of AUC using Hoeffding algorithms on outpatient no-show appointments. Goffman et al. [23] used Logistic Regression on a dataset based on demographic and appointment descriptions and the past patient's activities. They achieved an AUC of 71%. Moreover, Nelson et al. [24] created a predictive model for imaging no-show appointments. They got the best result with the Gradient Boosting algorithm, with 85% and 51.1% for AUC and precision, respectively. Lee *et al.* [25] collected two years of data on no-shows, where they applied three machine learning algorithms: Logistic Regression, Decision Tree, and Random Forest. Random Forest achieved the best results (accuracy of 72.9%).

III. Methodology

A. Dataset and Attributes

Thirty months of outpatient visits were extracted from the Ministry of National Guard Health Affairs (MNGHA) data warehouse from January 2017 to July 2019. The data were retrieved from all the medical facilities in the central region. The central region has the largest medical facilities of the MNGHA, where more than 70K appointments are booked monthly. The retrieved dataset consists of more than 3,733,580 million pediatric patient visits related to 104,640 pediatric patients. The patient's arrival time at the clinic is used to label the record as show (timestamp) or no-show (no timestamp).

The distribution of genders is almost equal, with 49.3% as female patients and 50.7% as male. In comparison, most patients were Saudi nationals (90%) since the hospital mainly serves employees of national guards and their families, Table I. However, the Saudis have the highest no-show rate of 99% (miss at least one medical appointment), Table I.

TABLE I. Statistical Description of Patients (N= 104640)

Attributes	*No-Show N%*	*Show N%*	*Total %*
Age: Infant (0-12 Months) Toddler (1-3 Years) Preschool (3-6 Years) School-age (6-12 Years) Adolescent (12-14 Years)	6651 (90.5%) 19187 (64.1%) 7986 (63.9%) 16733 (56.8%) 12422 (48.8%)	701 (9.5) 10729 (35.9%) 4503 (36.1%) 12719 (43.2%) 13009 (51.2%)	7352 (7%) 29916 (28.6%) 12489 (12%) 29452 (28.1%) 25431 (24.3)
Gender: Female Male	22202 (47.5%) 24427 (52.5%)	29373 (50.6%) 28638 (49.4%)	51575 (49.3%) 53065 (50.7)
Nationality: Non-Saudi Saudi	362 (0.7%) 46267 (99.3%)	696 (1.1%) 57315 (98.9%)	1058 (1%) 103582 (99%)

The data warehouse team filtered out unnecessary records to reduce the noise in the extracted data. Categorical attributes such as gender and age were converted into an integer. The age attribute was calculated based on the difference between the birthday and appointment dates. The age attribute was grouped into five categories: i) infant (0-12 months), ii) Toddler (1-3 years,) iii) Preschool (3-6 years), iv) School-age (6-12 years), and v) Adolescent (12-14 years). A new attribute, lead days, was derived from the dataset. The lead days attribute is a derived attribute calculated as the number of days between the booking of the appointment and the scheduling day. The historical behavior of the patients was included since it could contain helpful information to predict future outcomes. Therefore, the number of walk-ins, scheduled appointments, emergency visits, and canceled medical appointments were included for all patients.

TABLE II. Description of the Attributes

Attribute	*Description*
1) Gender	Male, Female
2) Age Category	Five age categories
3) Nationality	The nationality of the patient
4) Medical Department code	The code of medical/clinic
5) Hospital code	The code of referral hospital /clinic
6) Patient Services Department Type: a) Patient service b) Business center	Type of patient insurance
7) Address code	The primary health care clinic location (used for the patient's address)
8)Appointment type code: New Patient (NP) First visit (FV) Follow-up (FU)	Type of Appointment New Patient First Visit Follow-up Visit
9) Patient Services Department Type Cod: First visit (FV) Follow up (FU)	Type of patient insurance First Visit Follow-up Visit
10) Lead Days	Derived: the time difference in days between the booking of the appointments and the day of the appointments
11) Cancellation Flag: a) Yes= appointment cancelled b) No = not cancelled	Derived: count how many Check-ins Cancellation was Yes as total Derived: count how many Check-ins Cancellation was No as total
12) Medical Treatment Reservation Type Code: a) 1= Schedule b) 2= Walk-in	Derived: count how many Scheduled reservations as the total Derived: count how many Walk-in reservations as total
13) Medical Treatment a) Yes= patient treated b) N= not treated	Derived: count how many patients treated as total Derived: count how many patients not treated as total
14) On Foot visit a) 1= Yes b) 0 = No	Derived: count how many On Foot visit was Yes as total Derived: count how many On Foot visit was No as total
15) Emergency visit a) 1= Yes b) 0 = No	Derived: count how many Emergency visits were Yes as total Derived: count how many Emergency visits were No as total
16) Flu Season	Flu season from Oct-Nov and March-April
17) Distance (km)	The distance between the referring clinic and the Central Medical center of MNGHA at Riyadh City
18) Class No-Show	class show=0 or No-show=1

The historical visits of the patients are grouped together, and nine statistical values were calculated for 11 attributes (attributes 7 to 17 from Table II) which were minimum, maximum, standard deviation, variance, mean, median, skewness, and kurtosis. After the preprocessing phase, the dataset consists of 558,721 records with 106 attributes, including the class; Table II shows the description of the attributes.

B. Machine Learning Algorithms and Evaluation Criteria

This study involves five machine learning algorithms for classification data tasks. That includes Random Forest (RF), Gradient Boosting (GB), Logistic Regression (LR), Neural Network (NN), and Decision Tree (C4.5). RF and C5.4 are machine-learning algorithms from the family of decision-tree class algorithms. Gini impurity, information gain, and other techniques are used to build the tree structure; further algorithms description can be found in [26-29].

On the other hand, the Gradient Boosting algorithm is a boosting algorithm. It utilizes a number of weak classifiers through the ensemble method to build strong learners. The algorithm used gradient descent with other techniques to complete building the model; further algorithm descriptions can be found in [26,30. Logistic Regression (LR) can be described as a sample Neural Network (NN). LR can be viewed as a one-layer NN. LR and NN utilized many techniques to build the classification model of a binary class; further algorithm descriptions can be found in [26].

We evaluated and compared the five model's performance based on four evolution criteria: sensitivity (recall), precision, F-score, and accuracy. We considered True Positive (TP) Rate, False Positive (FP) Rate, True Negative (TN), False Negative (FN) Rate, and Receiver Operating Characteristic (ROC). These are defined as:

- True Positive (TP) Rate represents the number of no-show patient events classified as No-show, calculated based on equation 1:

$$TPR=TP / (TP + FN) \tag{1}$$

- False Positive (FP) Rate represents the number of show patient events classified as a no-show, calculated based on equation 2:

$$FPR=FP / (FP + TN) \tag{2}$$

- True Negative (TN) Rate represents the number of show patient events classified as a show, calculated based on equation 3:

$$TPR=TN / (TN + FP) \tag{3}$$

- False Negative (FN) Rate represents the number of no-show patient events classified as a show, calculated based On equation 4:

$$FPR=FN / (FN + TP) \tag{4}$$

- Receiver Operating Characteristic (ROC) score is a classification performance to show a false positive rate versus a true positive rate across a series of cut-off points and selecting the optimal cut-off point.

IV. Results and Discussion

The dataset was divided into 80% training and 20% testing. Table III shows the performance of the machine learning algorithms on an unseen testing dataset. Random Forest (RF) and Gradient Boosting (GB) outperformed the other classifiers on all five evaluation criteria. GB achieved 97.1% for ROC, which was 0.5% higher than RF. On the other hand, the score is higher than GB in terms of false positive rate (5.8% for RF and 5.9% for GB). In general, GB can be considered the champion model with an AUC of 97.1%.

Accurately predicting the intention of patients showing or no-showing to their medical appointment is considered an interesting and challenging goal for healthcare providers. Machine learning has become popular in healthcare research because of machine learning algorithms' ability to discover hidden patterns in the datasets, predict future outcomes and recognize the most relative attributes.

The machine learning algorithms were applied to more than half a million aggregated records with complex relations between attributes (predictors) and class labels with more than 100 attributes. The results of Gradient Boosting in predicting outpatient no-show outperforms traditional algorithms applied by previous research studies in Section II.

Analysis of the most used relevant attributes by machine learning shows that specific attributes significantly impact patient appointment adherence. The attributes are the traveling distance between the patient's residence and the medical center, the increased number of days between booking an appointment and the scheduled day (lead days), clinical medical services, and appointment time and day.

The rate of patient adherence to scheduled appointments can be increased by utilizing other resources. The medical facilities can use Short Message Service (SMS), an automated phone calls reminders 2-3 days before the medical appointment, mainly if the lead time is so long (more than four months). Furthermore, changing the lead days to be no more than 2-3 months [31-32].

Regarding model deployment, the machine learning framework used to build the model is a standard approach that can be integrated effortlessly within the digital healthcare system [33-34]. The research study has two limitations. Firstly, the dataset was retrieved from one region in KSA; therefore, more data from other regions can improve the robustness of the model if the model is deployed on different regions' digital systems. Secondly, the attributes are mainly extracted from an EHR in which other factors that influence the no-show rate are not included. These factors, such as personal reasons, availability of transportation, cultural background, educational level of the parents, and environmental conditions (e.g., weather circumstances), are not available in the digital healthcare system.

TABLE III. The Performance of Classifiers on Test Data

Evaluation Criteria	**DT**	**RF**	**LR**	**NN**	**GB**
ROC	94.8%	**96.4%**	95.0%	95.2%	**97.1%**
TPR	89.7%	**90.0%**	89.5%	89.7%	**90.0%**
FNR	10.3%	**10.0%**	10.5%	10.3%	**10.0%**
TNR	93.1%	**94.2%**	92.1%	91.9%	**94.1**
FPR	6.9%	**5.8%**	7.9%	8.1%	**5.9%**

V. Conclusion

In summary, machine learning's ability to acceptably predict no-shows provides a new potential digital solution for healthcare facilities to take advantage of intelligent solutions. This paper used five machine learning algorithms, namely, Decision Tree, Random Forest, Gradient Boosting, Logistic Regression, and Neural Network, applied to MNGH's extracted dataset from the data warehouse system.

The model was trained and tested using a dataset with different metrics to compare the results. Results show that Gradient Boosting achieves high performance for decision-making by the healthcare appointment management team. Furthermore, the study identified factors that increase the no-show rate, such as location, lead days between the reservation day and appointment day, appointment time of the day, and the type of medical services. Recommendations to increase the adherence of the no-show by patients have also been provided in Section IV.

Future work is to obtain more datasets from healthcare systems (for example, laboratory and medication data) and integrate weather data and patient background data such as educational level, transportation availability, and cultural background.

Disclosures

None of the authors have any competing interests.

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(03):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):2.

[3] Chariatte V, Berchtold A, Akre C, Pierre-André M (2008) Missed Appointments in an Outpatient Clinic for Adolescents, an Approach to Predict the Risk of Missing. J Adolesc Heal 43:38–45.

[4] Carlsen KH, Carlsen KM, Serup J (2016) Non-attendance, predictors and interventions. Adherence in Dermatology 29–35.

[5] Davies M, Goffman R, May J, Monte R, Rodriguez K, Tjader Y, Vargas D (2016) Large-Scale No-Show Patterns and Distributions for Clinic Operational Research. Healthcare 4:15.

[6] O. AlRowaili M, Ahmed AE, Areabi HA (2016) Factors associated with No-Shows and rescheduling MRI appointments. BMC Health Serv Res 16:1–7.

[7] 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;1(57):428–43.

[8] Samorani M, Harris S, Blount LG, Lu H, Santoro MA. Overbooked and Overlooked: Machine Learning and Racial Bias in Medical Appointment Scheduling. Available at SSRN 3467047. 2019 Oct 9.

[9] Srinivas S, Ravindran AR (2018) Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework. Expert Syst Appl 102:245–261.

[10] J. Daggy, *et al*., "Using no-show modeling to improve clinic performance," *Health Informatics J.*, vol. 16 pp. 246-259, 2010.

[11] D. Bentayeb, N. Lahrichi, and L. M. Rousseau, "Patient scheduling based on a service-time prediction model: A datadriven study for a radiotherapy center," *Health Care Manag. Sci.*, vol. 22, pp. 768-782, 2019.

[12] Z. Huang, M. Ashraf, H. Gordish-Dressman, and P. Mudd, "The financial impact of clinic no-show rates in an academic pediatric otolaryngology practice," Am. J. Otolaryngol., vol. 38, pp. 127- 129, 2017.

[13] Joseph, J., Senith, S., Kirubaraj, A. A., & Ramson, S. R. J. (2022). Machine Learning for Prediction of Clinical Appointment No-Shows. *International Journal of Mathematical, Engineering and Management Sciences*, *7*(4), 558-574.

[14] Chen J, Goldstein IH, Lin WC, Chiang MF, Hribar MR. Application of Machine Learning to Predict Patient No-Shows in an Academic Pediatric Ophthalmology Clinic. AMIA Annu Symp Proc. 2021 Jan 25;2020:293-302. PMID: 33936401; PMCID: PMC8075453.

[15] Daghistani, T., AlGhamdi, H., Alshammari, R. *et al.* Predictors of outpatients' no-show: big data analytics using apache spark. *J Big Data* **7**, 108 (2020).

[16] Alshammari, Riyad, Tahani Daghistani, and Abdulwahhab Alshammari. "The Prediction of Outpatient No-Show Visits by using Deep Neural Network from Large Data." *International Journal of Advanced Computer Science and Applications* 11.10 (2020).

[17] Harvey HB, Liu C, Ai J, Jaworsky C, Guerrier CE, Flores E, Pianykh O (2017) Predicting No-Shows in Radiology Using Regression Modeling of Data Available in the Electronic Medical Record. J Am Coll Radiol 14:1303–1309.

[18] Chua SL, Chow WL (2019) Development of predictive scoring model for risk stratification of no-show at a public hospital specialist outpatient clinic. Proc Singapore Healthc 28:96–104.

[19] Dantas LF, Hamacher S, Cyrino Oliveira FL, Barbosa SDJ, Viegas F (2019) Predicting Patient No-show Behavior: a Study in a Bariatric Clinic. Obes Surg 29:40–47.

[20] Kurasawa H, Hayashi K, Fujino A, Takasugi K, Haga T, Waki K, Noguchi T, Ohe K (2016) Machine-Learning-Based Prediction of a Missed Scheduled Clinical Appointment by Patients with Diabetes. J Diabetes Sci Technol 10:730–736.

[21] Mohammadi I, Wu H, Turkcan A, Toscos T, Doebbeling BN (2018) Data Analytics and Modeling for Appointment No-show in Community Health Centers. J Prim Care Community Heal 9:1–11.

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

[23] Goffman RM, Harris SL, May JH, Milicevic AS, Monte RJ, Myaskovsky L, Rodriguez KL, Tjader YC, Vargas DL. Modeling patient no-show history and predicting future outpatient appointment behavior in the Veterans Health Administration. Mil Med. 2017;182(5–6):e1708–14. 2017;14(10):1303–9.

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

[25] G. Lee, *et al.*, "Leveraging on predictive analytics to manage clinic no-show and improve accessibility of care," in *Proc. Int.* Conf. Data Sci. Adv. Anal., 2017, pp. 429-438.

[26] Alpaydin, Ethem. *Introduction to machine learning*. MIT press, 2020.

[27] Fawagreh K, Gaber MM, Elyan E. Random forests: from early developments to recent advancements. Systems Sci Control Engineering Open Access J. 2014;2(1):602–9.

[28] Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogrammetry Remote Sensing. 2012;1(67):93–104.

[29] J R Quinlan, "C4.5: Programs for Machine Learning",Morgan Kaufmann Publishers,isbn=1-55860-238-0, 1993.

[30] Friedman, Jerome H. "Greedy function approximation: a gradient boosting machine." *Annals of statistics* (2001): 1189-12.

[31] Agarwal P, Nathan AS, Jaleel Z, Levi JR. Factors Contributing to Missed Appointments in a Pediatric Otolaryngology Clinic. Laryngoscope. 2022 Apr;132(4):895-900. doi: 10.1002/lary.29841. Epub 2021 Aug 24. PMID: 34427327.

[32] Shahab I, Meili R. Examining non-attendance of doctor's appointments at a community clinic in Saskatoon. Can Fam Physician. 2019 Jun;65(6):e264-e268. PMID: 31189640; PMCID: PMC6738377.

[33] Hofer IS, Burns M, Kendale S, Wanderer JP. Realistically Integrating Machine Learning Into Clinical Practice: A Road Map of Opportunities, Challenges, and a Potential Future. Anesth Analg. 2020 May;130(5):1115-1118. doi: 10.1213/ANE.0000000000004575. PMID: 32287118; PMCID: PMC7584400.

[34] Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a Machine Learning Model Applied to Harmonized Electronic Health Record Data for the Prediction of Incident Atrial Fibrillation. JAMA Netw Open. 2020 Jan 3;3(1):e1919396. doi: 10.1001/jamanetworkopen.2019.19396.