Predicting no-shows for dental appointments

Yazeed Alabdulkarim[1], Mohammed Almukaynizi[1], Abdulmajeed Alameer[2], Bassil Makanati[1], Riyadh Althumairy[3] and Abdulaziz Almaslukh[1]

[1] Information Systems Department, King Saud University, Riyadh, Saudi Arabia
[2] Computer Science Department, King Saud University, Riyadh, Saudi Arabia
[3] Department of Restorative Dental Sciences, King Saud University, Riyadh, Saudi Arabia

ABSTRACT

Patient no-shows is a significant problem in healthcare, reaching up to 80% of booked appointments and costing billions of dollars. Predicting no-shows for individual patients empowers clinics to implement better mitigation strategies. Patients' no-show behavior varies across health clinics and the types of appointments, calling for fine-grained studies to uncover these variations in no-show patterns. This article focuses on dental appointments because they are notably longer than regular medical appointments due to the complexity of dental procedures. We leverage machine learning techniques to develop predictive models for dental no-shows, with the best model achieving an Area Under the Curve (AUC) of 0.718 and an F1 score of 66.5%. Additionally, we propose and evaluate a novel method to represent no-show history as a binary sequence of events, enabling the predictive models to learn the associated future no-show behavior with these patterns. We discuss the utility of no-show predictions to improve the scheduling of dental appointments, such as reallocating appointments and reducing their duration.

INTRODUCTION

Patient no-show, which refers to patients missing their booked appointments, is a major problem in healthcare. Its rates may reach up to 80% (*Kheirkhah et al., 2015*; *Rust et al., 1995*). Reports estimated that the no-show of patients costs the United States' healthcare system an average of $150 billion a year (*Gier, 2017*). Patients may miss their appointments for various reasons, such as scheduling factors, forgetness, transportation, and work-related issues (*Marbouh et al., 2020*). The consequences of this problem are wide, affecting patients, doctors, and the overall healthcare system. Patients are perhaps affected the most by no-shows, whether it is the patient who has missed the appointment or the others who could not book appointments due to unavailabilities. When patients miss their appointments, important treatments get delayed, leading to complications that require costly and complex medical procedures.

Outpatient clinics implement several mitigation strategies to tackle the no-show problem, yet, this brings limited success. For example, sending reminders to booked patients may not reduce the no-show rates (*Li et al., 2019b*). Blind overbooking is ineffective because the number of extra appointments is often estimated based on factors

that do not consider the historical no-show patterns of the scheduled patients. If the overbooking strategy is implemented aggressively, it would crowd the clinic, over-utilize its resources and increase patient waiting time. On the other hand, if overbooking is conservative, resources would be wasted, and the clinic is yet again underbooked (*Zeng, Zhao & Lawley, 2013*).

Several papers in the literature have utilized machine learning techniques to develop predictive models for patient no-shows. See "Related Work" for details. These models use patient no-show history to predict no-shows on a per-patient basis, which empowers targeted and more effective mitigation strategies. No-show behavior varies across health clinics and specialties (*AlMuhaideb et al., 2019*; *Ding et al., 2018*; *Mohammadi et al., 2018*). For instance, reported percentages of no-shows for Urology and Oncology were 6.5% and 6.9%, respectively. In contrast, no-shows for Surgery, Optometry, and Dietary reached 16.6%, 23.8% and 27.23%, respectively (*AlMuhaideb et al., 2019*). This calls for fine-grained studies covering specific specialties to discover useful patterns that may help to predict no-shows.

This problem could get even worse for some geographical locations. For example, previous research reported that 58.1% of patients missed their dental appointments in Saudi Arabia, impacting their oral health (*Baskaradoss, 2016*; *Shabbir, Alzahrani & Khalid, 2018*). Furthermore, dental appointments are notably longer than regular medical appointments due to the complexity of dental procedures. The average duration of primary care appointments is 17.4 min, while the average for dental appointments is 48.7 min, as reported by the American Dental Association (*American Dental Association, 2021*). Dental appointments are, on average, almost thrice as much as the duration of primary care appointments. Consequently, dental no-shows are costlier and blind overbookings are impractical. In this article, we leverage machine learning techniques on a newly proposed feature creation approach to develop models to precisely predict no-shows to dental appointments.

Prior no-show history is one of the most significant predictors for future no-shows (*AlMuhaideb et al., 2019*; *Goffman et al., 2017*; *Harvey et al., 2017*). This article proposes a novel method to represent no-show history as a binary sequence of events. Each previous appointment is represented as 1 if the patient has attended and as 0 if the patient has missed the appointment, resulting in a representation of a sequence of binary numbers for each of the appointment records. We evaluate this approach for capturing no-show history considering the last three, five, seven, and ten appointments. The intuition behind this approach is that the model should learn the behavior associated with each sequence of no-show history in the training phase rather than reducing no-show history to a number based on fixed calculations.

Additionally, we discuss the utility of no-show predictions to improve the scheduling of dental appointments. Several studies (*Chiara & Domenic, 2018*; *Li et al., 2019b*; *Srinivas & Ravindran, 2018*) suggested utilizing the predictions of no-shows to implement more effective overbooking strategies. These strategies may be impractical for dental appointments because of their long duration, see "Discussion" for details. For this reason, we propose to reduce the time of appointments with high no-show probabilities to

minimize the impact of no-shows. Last, the dataset is publicly available for the research community to facilitate further research and comparative studies. The contribution of this article is as follows:

- Leveraging machine learning models to predict no-shows for dental appointments.
- Proposing a novel method to capture prior no-show history, one of the strongest predictors of no-shows, as a binary sequence.
- Reviewing and discussing various techniques to mitigate no-shows in the field of dentistry.
- Making the dataset publicly available for further research (see Supplemental File).

The organization of this article is as follows. "Problem Definition" details our approach to capturing no-show history. The collected data and our research methodology are described in "Data Analysis" and "Experimental Setup", respectively. Experimental results and their discussion are reported in "Results" and "Discussion". Last, "Related Work" provides a literature review and "Conclusion" concludes our work.

PROBLEM DEFINITION

In this section, we first formally define the problem of predicting no-shows to dental appointments. Then we describe our approach to capturing the no-show history patterns from our dataset.

Predicting dental appointment no-shows

We formulate the problem of predicting no-shows as a binary classification problem, where the goal is to assign a class label belonging to {no-show, show} for each booked dental appointment. In our settings, we consider the no-show as the *positive class* for which our models are optimized to predict. For a given booked dental appointment n_i, such that $n_i \in N$, where N is the set of all booked appointments, we seek to define a prediction algorithm that assigns a score $S(n_i) \in [0, 1]$. The score models the predicted likelihood that the dental appointment n_i will be unattended, *i.e.*, the higher the prediction score, the more likely that patient will not attend the appointment. A decision boundary $\gamma \in [0, 1]$ can be used to attain class label assignments in accordance to the prediction function $F: N \rightarrow$ [no-show, show], such that:

$$F(N) = \begin{cases} no-show, & if \quad S(N) \geq \gamma, \\ show, & otherwise. \end{cases}$$

Capturing no-show history

Prior no-show history is one of the significant predictors for future no-shows (*AlMuhaideb et al., 2019*; *Chiara & Domenic, 2018*; *Goffman et al., 2017*; *Harvey et al., 2017*). Most papers in the literature capture it as an empirical probability, a likelihood of an event occurring based on historical data. Instead of computing the no-show history for all previous appointments, the no-show percentage is usually calculated based on a moving

window, such as the latest 10 appointments. This is because older appointments have less impact on predicting future no-shows than recent ones.

This article captures no-show history as a binary sequence of events, representing each previous appointment as 1 if the patient attended and 0 if the patient missed the appointment. For instance, a patient with a previous no-show history sequence as (show, no-show, no-show, show, no-show) is represented as a (1, 0, 0, 1, 0). This approach is applied to sequences of n last appointments, *i.e.*, $(a_i)_{i=1}^{n}, a_i \in \{0, 1\}$. We perform an evaluation for these sequences where $n \in \{3, 5, 7, 10\}$; see "Results" for details. The intuition behind this method is that the model should learn the behavior associated with each sequence of no-show history in the training phase rather than reducing no-show history to a single number, *e.g.*, moving average. For example, consider a patient with a no-show history sequence as (show, show, no-show, no-show). The empirical probability of this no-show history is 50%. This approach treats all previous appointments equally, computing the same percentage (50%) for other different sequences, such as (no-show, no-show, show, show). It fails to capture possibly different future no-show behavior associated with these sequences.

DATA ANALYSIS

In order to evaluate our approach, we collected and analyzed a dataset of dental appointments obtained from a dental clinic. In "Data Collection and Preprocessing ", we describe our approach for collecting and preprocessing the dataset, and in "Feature Analysis ", we describe our feature analysis process.

Data collection and preprocessing

For this article, we collected a dataset of dental appointments from a local dental clinic in Riyadh, Saudi Arabia, for one year between January 1st, 2019, and December 31st, 2019. Our dataset consists of 262,140 appointments. Among these appointments, we delete 66,122 records that are missing critical data elements, such as appointment date/time, or containing erroneous values (see "Data Preprocessing" for details). The resulting data consists of 196,018 appointments. Among these appointments, 42.68% (83,663) belong to the positive class, *i.e., no-show*.

Each record in our dataset consists of 10 predictive attributes (features), in addition to the target class (*show* or *no-show*). Among the predictive attributes are data specific to each patient, such as date of birth, marital status, gender, and nationality. The records also contain information about the appointments, such as appointment date, time, *show* or *no-show*, booking date and time, appointment duration, and a variable specifying whether a confirmation SMS is sent for that appointment. Description of each variable is shown in Table 1.

Feature analysis

We explore the dataset to determine the distribution of values for different variables. In our dataset, 75.4% (133,772) of the appointments are for male patients, while 24.6% are for female patients; 89.3% are for single patients, and 10.7% are for married patients. The

Table 1 Description of the variables in the dataset.

Category	Variable	Description
Patient	Date of birth	Patient's date of birth
	Marital status	Specifies whether the patient is single or married
	Gender	Patient sex (male or female)
	Nationality	The nationality of the patient
	Appointment date	Date of the booked appointment
	Appointment time	Time of the booked appointment
	Show or No-show	Specifies whether the appointment is attended or missed
	Booking date and time	Date and time in which the appointment was booked
	Appointment duration	Length of the appointment in minutes
	SMS	Specifies whether a confirmation SMS is sent or not sent
Appointment	Doctor ID	Dentist identification number

average age of patients at appointment time is 29.77 years, as illustrated in Fig. 1. The average appointment duration is 35.99 min, as shown in Fig. 2.

We utilize the variables, shown in Table 1, to compute features, such as the appointment lead time (*i.e.*, days between appointment booking date and actual appointment date), appointment month, and day. Furthermore, we obtain historical weather data for the appointment date, including temperature and condition. The complete list of computed features and their description is shown in Table 2. For instance, (*In Holiday*) is a feature to capture whether the appointment date is on a public holiday in Saudi Arabia, and (*In Ramadan*) is a feature to capture whether the appointment date is in Ramadan, which is the holy month of fasting for Muslims. We also compute features to capture the previous no-show history in various ways. First, we calculate the percentage of no-shows based on the last 10 appointments, namely 'No-show10 %', and based on all previous appointments, namely 'No-show All %'. Second, we represent previous no-show history as a binary sequence of events. We provide a detailed description of the no-show binary sequence representation in "Problem Definition".

We investigate the correlation between the target variable (*show* or *no-show*) and each feature in the dataset for preliminary analysis. We observe multiple features that correlate with the target class variable. These features are Age at Appointment Date, Lead Time, Hour, and the variables associated with previous no-show history.

By looking at the age variable (Fig. 3), the records in the dataset indicate that patients who need a companion (*i.e.*, kids and elderly patients) are more likely to show up for their appointments. By investigating the Lead Time variable (Fig. 4), it can be clearly seen that appointments that are booked closer to the actual appointment date have a higher rate of showing up. When looking at previous no-show history (Fig. 5), it is evident that patients with a high rate of prior no-shows are more likely to not show up for their future appointments.

To further investigate the relationship between appointment hours and the likelihood of no-shows, we bin the 'Hour' variable into different shifts to identify shifts that are likely to
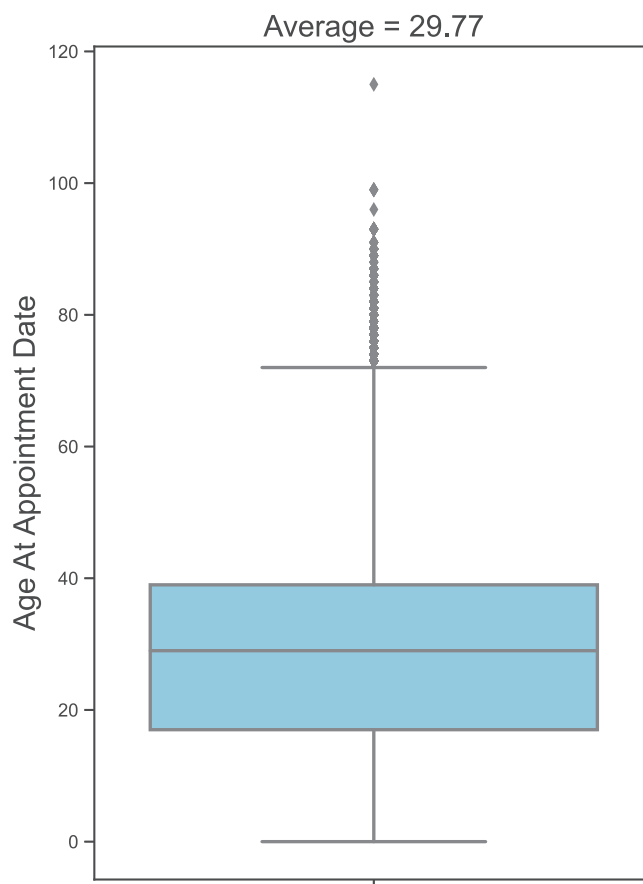
Figure 1 Patients' age distribution at appointment date.

Full-size 🖼 DOI: 10.7717/peerj-cs.1147/fig-1

exhibit more no-shows. As seen in Fig. 6, appointments scheduled early in the morning (*i.e.*, before 10 AM) have the highest likelihood of being unattended.

Overall, this preliminary analysis shows clear correlations between some features and the target class label. This implies that machine learning methods can be effective in building a well-performing model for predicting no-shows for dental appointments.

EXPERIMENTAL SETUP

Several supervised learning algorithms are examined to predict appointment no-shows with the best performance possible. This section explains the steps followed to preprocess and clean our data. It details our approach to partitioning data into training and testing sets. Moreover, the model selection process and hyperparameter optimization applied to our datasets are explained in this section. Additionally, we present the evaluation metrics used to evaluate the performance of the machine learning models. We do so by using the scikit-learn Python machine learning library.

Data preprocessing

Our data is produced by an appointment scheduling system that enables the dental clinic to manage patient appointments. The data is pulled out from a relational database. Some
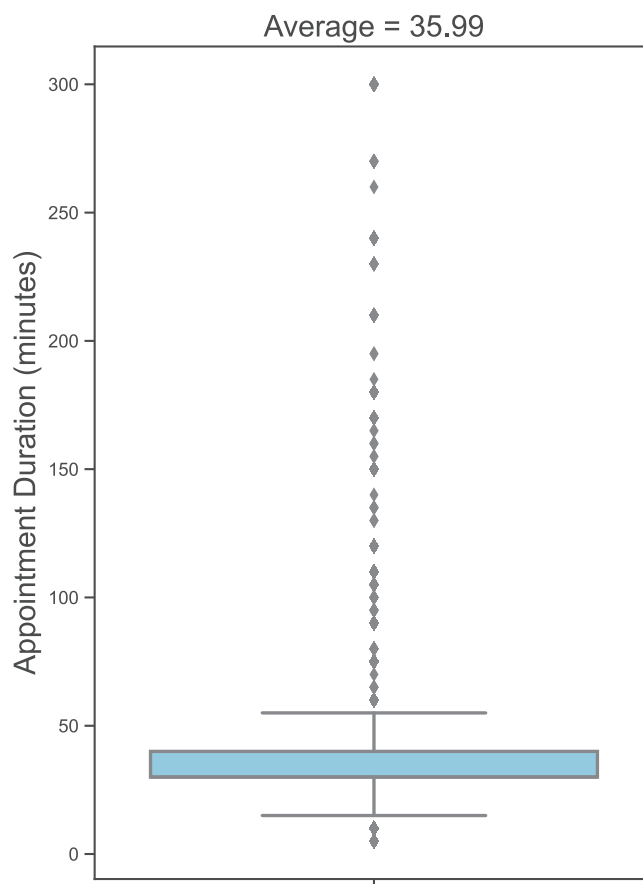
**Figure 2 Distribution of appointment duration in minutes.**

attributes in the database may have been manually altered by users having certain privileges without validating the inputs. Manual alteration has resulted in issues with values in some database attributes. For example, the attribute 'duration' represents the estimated time needed to complete the dental appointment's work. Hence, it may only take positive values. The dataset has zero- and negative-duration appointments. All records with any such invalid values have been removed from the dataset. Invalid values are found in four attributes, explained in Table 1. That includes records where: (1) 'Appointment Duration' is less than 10 min, (2) 'Date of Birth' is a future date, (3) 'Appointment Date' is missing, or (4) 'Appointment Time' is missing. Records of patients without appointments (*i.e.*, walk-ins) are removed because they are only admitted if physicians are available. Therefore, they are not part of the prediction task.

Training and testing data split

Machine learning models are generally trained by fitting a training dataset. They are then evaluated on a different testing dataset. The train and test datasets are often assumed to be drawn from the same underlying data distribution. However, this may not always be the case with time-series data where the data distribution changes over time, aka., concept drift

Table 2 Description of the computed variables in the dataset.

Variable	Description
Multi appointments	More than one appointment on the same day
Age at appointment date	Patient age at the appointment date
Lead time	Days between booking and appointment dates
Days since last appointment	Days between the last and upcoming appointments
Appointment day	Day of the week that the appointment is booked in
Is weekend	If the appointment is on the weekend
Month	The month that the appointment was booked in
In holiday	If the appointment is scheduled to be during a national holiday
In Ramadan	If the appointment is in Ramadan
Temperature	The temperature on the appointment day
Weather	The weather on the appointment day
Hour	The appointment time in hours
No-show All #	Number of all no-shows
No-show All %	Percentage of all no-shows of a particular patient
No-show10 #	Number of no-shows based on the last ten appointments
No-show10 %	Percentage of no-shows based on the last ten appointments
No-show Seq3	Binary sequence representation of last three appointments
No-show Seq5	Binary sequence representation of last five appointments
No-show Seq7	Binary sequence representation of last seven appointments
No-show Seq10	Binary sequence representation of last ten appointments
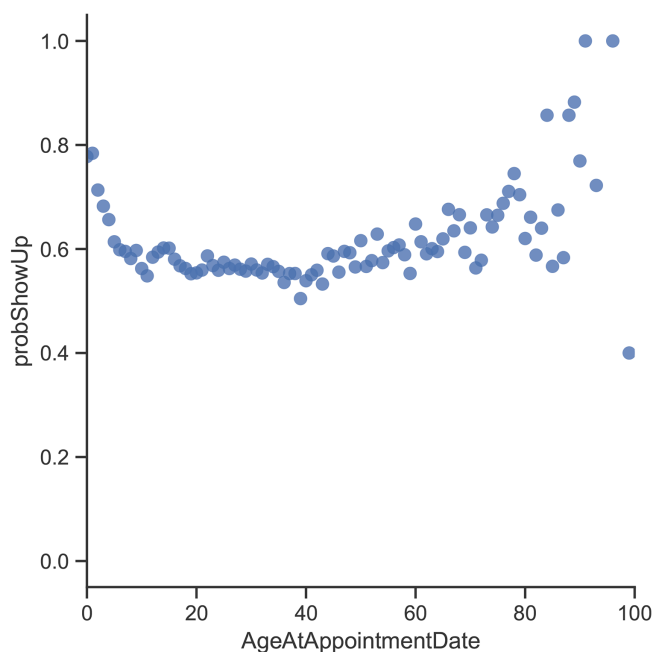
Figure 3 Probability of showing up with respect to the patient's age at appointment date.
Full-size ☑ DOI: 10.7717/peerj-cs.1147/fig-3
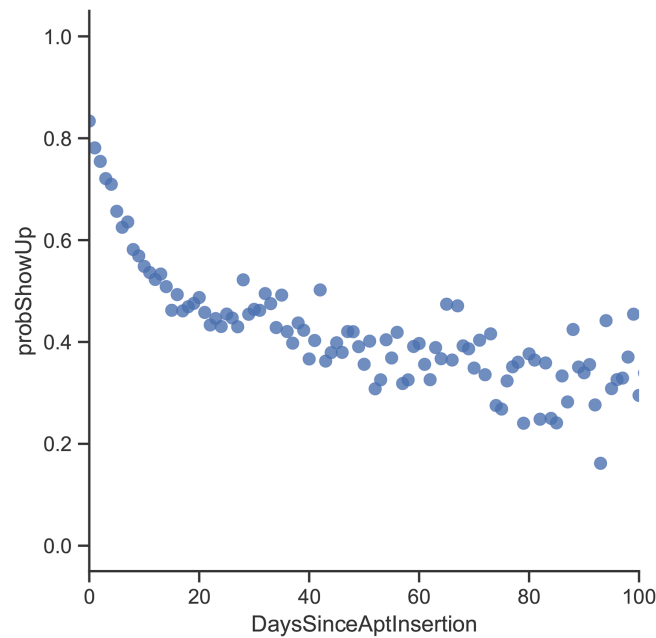
Figure 4 Probability of showing up with respect to lead time.
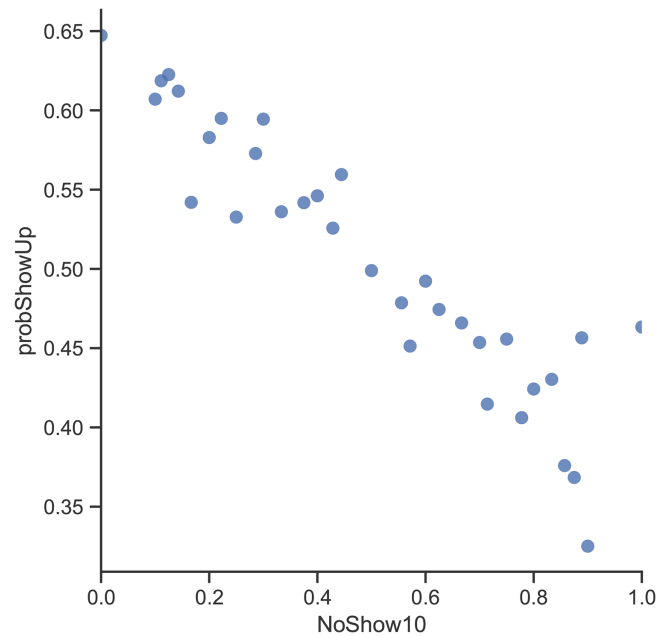Full-size 🖾 DOI: 10.7717/peerj-cs.1147/fig-4

Figure 5 Probability of showing up with respect to no-show rate in the last 10 appointments.
Full-size 🖾 DOI: 10.7717/peerj-cs.1147/fig-5

(see (*Gama et al., 2014*) for a survey). Data may be partitioned into training and testing data in different ways. For example, in k-fold cross-validation, performance is reported on all the data after k runs, *i.e.*, data is partitioned into k folds. For each fold/run, models are
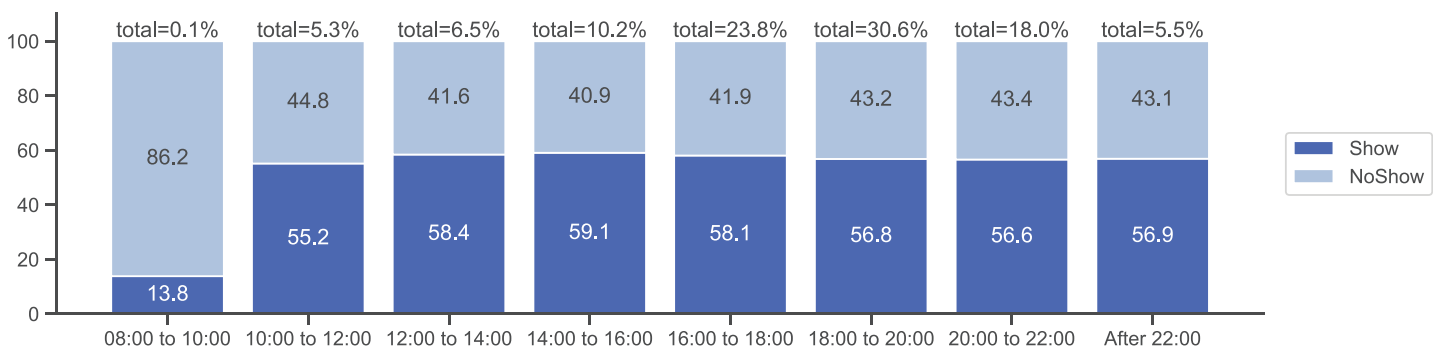
Figure 6 Probability of showing up with respect to the hour shift of the day.   Full-size ⬚ DOI: 10.7717/peerj-cs.1147/fig-6

fit to the remaining k−1 folds, and the fold is used to test the models. Such a setting violates the temporal aspect of time-series data where concept drift may exist.

Our task of predicting appointment no-show has the temporal dimension, *i.e.*, time-dependent. Therefore, we follow experimental settings that best resemble real-world, deployed predictive models on time-series data where concept drift may exist. To do so, appointments are sorted by the appointment date for all experiments. After sorting, 90% of the data points are used to train models, and 10% of the data points are reserved for testing the models. We do so to avoid timely-intermixing the appointments. This way guarantees that none of the appointments in the testing set preceded any appointment in the training set. Following these settings, the first set of experiments is done to predict the no-show on all appointments, regardless of the appointment duration. In this set of experiments, models are trained on several feature sets, as explained in "Results".

In another set of experiments, the models are fit to predict longer appointments, *e.g.*, appointments for more lengthy procedures such as crown mounting and root canal. This category of appointments costs the clinic significantly more than shorter appointments as they rely on the high involvement of skilled physicians and at least two assistants.

Optimizing hyperparameters

Most supervised machine learning algorithms can be controlled to learn representations from data through a set of parameters, often known as hyperparameters, *e.g.*, learning rates, number of predictors in ensemble machine learning algorithms, depth of the decision trees, types of kernels, *etc*. The optimal set of hyperparameters should be empirically identified from the training dataset. The goal of such optimization is to produce the best-performing models (in terms of the efficiency of the learning process and the accuracy of the models). Grid search is perhaps the most commonly used hyperparameter tuning algorithm because of its implementation simplicity (*Bergstra & Bengio, 2012*). Grid search attempts to exhaustively search through a pre-defined set of hyperparameter space of a given algorithm.

In this article, we run a grid search for all experiments and models. Table 3 summarizes the hyperparameter search space and the optimal set of parameters that resulted in the

Table 3 Description of the hyperparameter search space and the optimal set of parameters that resulted in the highest F1 score.

ML model	Hyperparameter	Description	Search values	Optimal values by grid search
LR	'penalty'	Specifies the norm of penalty term	L1, L2	L2
	'C'	A constant controlling the regularization strength	0.001, 0.01, 0.1, 1, 10	0.001
	'max_iter'	Maximum number of iterations for the solvers to converge	100, 250, 500	500
RF	'n_estimators'	Number of trees in the forest	100, 250, 500	100
	'max_depth'	Maximum depth of trees in the forest	5, 8, 10, 30, 40, 50	10
	'min_samples_split'	Minimum number of samples required to be at a leaf node	2, 5, 10, 20, 50, 70, 80, 90, 100	70
GB	'learning_rate'	A constant controlling the contribution of each tree	0.001, 0.01, 0.1, 1	1
	'n_estimators'	Number of boosting stages performed	50, 100, 200, 300, 400, 500, 550, 600	500
	'max_depth'	Number of nodes in each individual tree	2, 3, 5, 8	2
	'min_samples_leaf'	Minimum number of samples required to be at a leaf node	1, 5, 10	1

highest F1 score. Grid search runs are parallelized on a CPU computer equipped with an 8-core Intel Core i7 Processor.

Supervised learning

After reviewing the related literature and empirically testing over eight different machine learning models, including k-nearest neighbors, naive Bayes, and decision tree, among others, we use grid search to find the optimal hyperparameter values on the three best-performing supervised learning models: (1) logistic regression (LR), (2) random forests (RF), and (3) gradient boosting (GB).

Logistic regression (LR): is a statistical model that is widely used for classification applications in various scientific fields. In logistic regression, a logistic function is fit to data to model the uncertainty that a dependant variable belongs to a certain binary class label, a setting that is similar to the problem addressed in this study. However, there are various extensions to logistic regression for multinomial classification problems (*El-Habil, 2012*).

Random forests (RF): is an ensemble learning model that is commonly used for classification (*Breiman, 2001*). A random forest works by constructing a set of classification tree predictors; each tree is induced from a set of data points sampled independently and trained on a sampled set of features, depending on the selected method. A large body of work in the literature has proposed different approaches to training individual trees in the forest for varying predictive tasks. In general, a significant improvement in predictive accuracy is achieved when the ensemble combines votes from the individual tree predictions.

Gradient boosting (GB): is similar to random forests, and it is an ensemble supervised machine learning approach that uses a set of decision trees. However, gradient boosting is an iterative model that builds the trees in a stage-wise manner by adding a weak learner at each learning stage to optimize the current ensemble of learners, *i.e.*, boosting (*Friedman, 2002*).

Table 4 Evaluation metrics. TP, true positives; FP, false positives; FN, false negatives; TN, true negative.

Metric	Formula
Precision	$TP/TP + FP$
Recall	$TP/TP + FN$
F1	$2 * precision * recall/precision + recall$

Performance evaluation

To report the prediction performance, we compare the labels predicted by our models with the actual labels from the ground truth testing dataset. We use different evaluation metrics to quantify the prediction performance. The first set of metrics is the commonly used precision, recall, and F1 score. In the problem of predicting dental appointment no-shows, precision is the fraction of appointments that are correctly predicted to be unattended from the total number of predicted appointments, while recall is the fraction of appointments that are correctly predicted to be unattended from the total unattended appointments. F1 score is the harmonic mean of precision and recall. Table 4 summarizes how each metric is computed.

Another metric we use is the Receiver Operating Characteristic (ROC) curve, a graphical chart that illustrates the performance of a binary predictor. At each confidence score threshold, ROC plots the true positive rate against the false positive rate. True positive rate (aka., sensitivity or recall) in this study is the fraction of the correctly predicted no-show appointments from all no-show appointments. Moreover, the false positive rate (aka., false alarm ratio) is the fraction of the appointments that are wrongly predicted as no-shows from all attended appointments. ROC only shows the performance graphically. Therefore, we use the area under the ROC curve (AUC) to quantify the performance of the predictors. A perfect predictor would have an AUC of 1, while a random predictor with no discrimination power would be expected to have an AUC that is close to 0.5. AUC is reported for all results in "Results".

RESULTS

This section shows the performance results for the three selected models explained in "Supervised Learning". We choose 'No-show Seq10' of the no-show features, shown in Table 2, because it captures more history, see "Comparing No-Show Features" for details. Fig. 7 presents the ROC curve for the three no-show predictor models on the testing dataset. It is clearly shown that the three models are performing almost the same in terms of AUC. Yet, they are offering different tradeoffs between true positive and false positive rates. Such tradeoffs are presented in the precision-recall curve as in Fig. 8.

Next, we show the F1 scores and compare the three models across various threshold values in Fig. 9. The models have similar F1 scores, but it is noticeable that GB tolerates threshold changes better than LR and RF. Given that the prior no-show probability is 42.68% in our testing dataset, our highest model with an F1 of 66.5% outperforms the baseline with an F1 increase of over 55%.
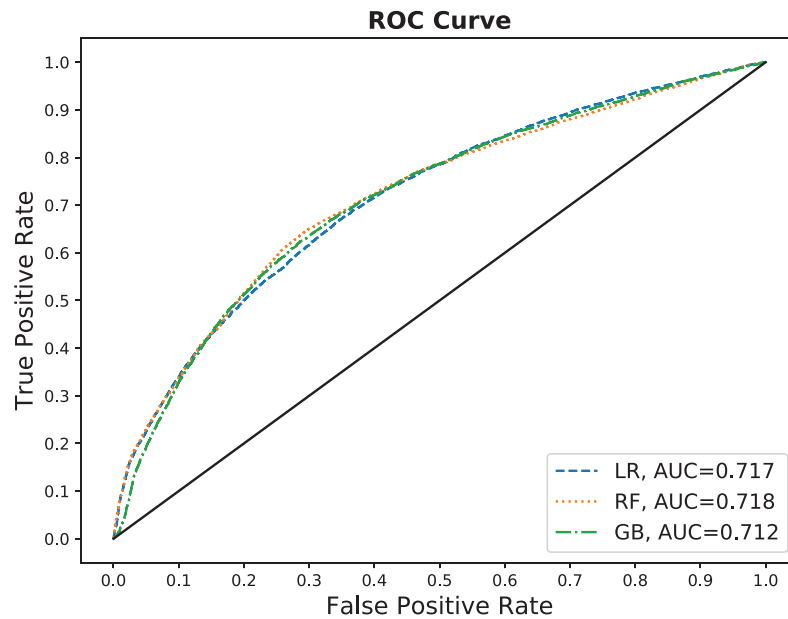
ROC Curve

Precision-Recall Curve

In the following subsections, we show results for four sets of experiments: (1) experiments providing a comparison between the different no-show features, (2) experiments evaluating the performance of our no-show prediction models on appointments with long duration, (3) experiments reporting the results of an ablation test
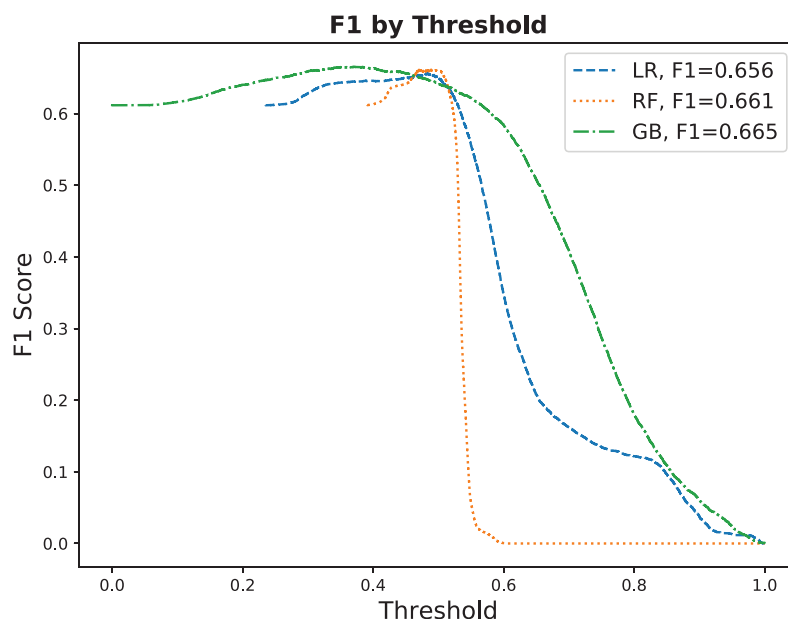
F1 by Threshold

Table 5 Reported AUC results for the three models on the testing dataset using various representations of no-show history.

Model	Baseline	No-show All #	No-show All %	No-show10 #	No-show10 %	No-show Seq3	No-show Seq5	No-show Seq7	No-show Seq10	Markov model
LR	0.676	0.674	0.695	0.677	0.695	0.718	0.719	0.718	0.717	0.712
RF	0.726	0.723	0.725	0.726	0.726	0.728	0.731	0.727	0.718	0.716
GB	0.724	0.726	0.724	0.732	0.727	0.728	0.722	0.722	0.712	0.711

to identify important features that are most indicative of class label, and (4) experiments evaluating the generalization of our models on a different dataset.

Comparing no-show features

This article proposes a novel method to represent no-show history as a binary sequence of events, as discussed in "Comparing No-Show Features". To show the utility of the no-show history representation approach, we evaluate this approach for the last n previous appointments $(a_i)_{i=1}^{n}, a_i \in \{0, 1\}$. This sequence representation is compared to other methods that represent previous no-show history as a numerical feature, namely 'No-show All #', 'No-show All %', 'No-show10 #', and 'No-show10 %', as described in Table 2, to evaluate the value of learned no-show sequences.

Table 5 provides a comparison between these no-show features. Each row shows the reported AUC for one of the prediction models. The second column, 'Baseline', represents the baseline result for each prediction model where all no-show history features are removed. We keep all other features that are not related to no-show history in all experiments. Each of the remaining columns represents adding its corresponding no-show history feature.

Table 6 No-show percentage and the number of appointments in the training and testing data sets for each grouping of duration.

Stats	Appointment length					
	≥ 15	≥ 30	≥ 40	≥ 45	≥ 50	≥ 60
No-show %	42%	43%	44%	43%	43%	43%
Number of training appointments	173,537	143,899	61,155	42,522	36,152	31,036
Number of testing appointments	8,599	7,225	3,234	2,249	1,869	1,631

Table 7 Reported AUC results for the three models showing the most significant features.

Model	All features	Lead time	No-show history
LR	0.717	0.642	0.676
RF	0.718	0.640	0.726
GB	0.712	0.633	0.724

Considering the logistic regression model, the binary sequence representations of no-show history outperform the numerical features by 2–5%. There is no significant difference between the examined lengths of no-show sequences. For the numerical no-show history features, percentage values outperform counter values by about 2%. For the remaining RF and GB models, there are insignificant differences between the various no-show history features. The reported AUC baseline numbers (second column) for the RF and GB models are very close to the AUC numbers when adding the no-show history features.

We compare our binary sequence representation of no-show history with the Markov model approach to compute no-show probabilities based on the last 10 appointments, proposed in (*Goffman et al., 2017*). The AUC values for the Markov model, see the last column of Table 5, are very close to our binary sequence representation for the three machine learning models.

Long appointments

Appointments that have longer duration are more costly as they occupy more time slots and probably include advanced medical procedures. Consequently, these appointments are more important for predicting no-shows accurately. Furthermore, we hypothesize that they have lower no-show rates as patients should be more considerate not to miss them. For this set of experiments, we look at appointments with duration $\geq d$ where $d \in \{15, 30, 40, 45, 50, 60\}$. We report the no-show rate and the number of appointments in training and testing data sets for each group d in Table 6. It turns out that the no-show rates are very similar for the various duration groups. We do not report the AUC for the various duration groups because they are comparable, indicating an insignificant impact.

Important features

We perform an ablation test to determine the important features. First, we add all features and measure AUC as a baseline, see second column of Table 7. Then, we remove each feature independently and report AUC to determine the importance of the feature. In

Table 8 Evaluating the generalization of our models that are trained on the first dataset on a new dataset.

(a) AUC

Model	First dataset	New dataset
LR	0.708	0.706
RF	0.731	0.748
GB	0.727	0.710

(b) F1 scores

Model	First dataset	New dataset
LR	0.652	0.657
RF	0.673	0.665
GB	0.669	0.653

other words, each experiment includes all features except for the tested feature.[1] We show the results for two features only, see Table 7, as other features are insignificant with less than 2% difference. Lead time is the most significant feature causing the AUC to drop 7.5%, 7.8%, and 7.9% with LR, RF, and GB, respectively, an observation that is confirmed by the existing literature (*AlMuhaideb et al., 2019*; *Chiara & Domenic, 2018*; *Goffman et al., 2017*; *Harvey et al., 2017*) that appointment lead time is one of the strong predictors of future no-shows. Eliminating no-show history features causes the AUC to drop by 4.1% with the LR model and is insignificant with the other models.

Model generalization

It is critical for machine learning models not to be limited to their training and testing datasets. This ensures they are widely applicable and adapt to data changes without significant performance degradation. To evaluate the generalization of the trained models, we obtain a new dataset from a small dental clinic in Riyadh, Saudi Arabia. The dataset consists of 49,007 appointments with 41% no-shows.

We use this new dataset to test the three machine models trained on the first dataset, which is from a different clinic. We slightly modify the training dataset by removing the 'SMS' feature, which specifies whether an SMS confirmation message has been received because the new dataset does not have it. We also use the 'No-show Seq3' feature instead of 'No-show Seq10' due to the small size of the new dataset that generated insufficient no-show history.

Table 8 shows the AUC and F1 scores comparing the performance of the three models that are trained on 90% of the first dataset. The second column of Tables 8a and 8b shows the results for the 10% testing of the first dataset, while the third column shows the results for the whole new dataset. The results for both AUC and F1 scores for the new dataset are very similar to scores on the first dataset, showing that the three models generalize very well. Furthermore, given a 41% no-show for the new dataset as a baseline, our models outperform it by a percentage increase of over 62% in the F1 score.

[1] Since we are r es may minimize the impact of each other.

DISCUSSION

The patient no-show problem has a tremendous negative impact on dental clinics financially and operationally. Classical strategies to mitigate this problem include overbooking and sending a reminder to the patients. However, these strategies have limited success as they may not consider patients individually. In the literature, several works have adopted machine learning algorithms to develop a model to predict patients' no-shows based on various features. This work utilizes machine learning algorithms to build a very effective predictive model of the no-show patient and improves the implemented classical strategies.

Previous studies considered the patient history of no-shows and the lead time as the essential features in building the no-show model. Other features such as age, appointment time, duration, distance from the clinic, whether a reminder has been sent, etc., are also considered. However, the importance of these features is relatively less than the history of no-shows and the lead time as they do not provide important information to the model. Similarly, this work focuses on these two features as the main features to develop the model. However, the history of no-shows is further categorized into a sequence of binary (0 for missed appointments and 1 for attended appointments). This representation allows the model to learn the behavior rather than depending on statistical calculations.

Implementing traditional strategies, including overbooking or sending reminders blindly to the patients, could unexpectedly increase the overall cost or overwhelm the dental clinic resources, eventually diminishing the quality of the provided services. The proposed no-show model may help implement a careful overbooking strategy by scheduling a reduced appointment time slot for patients with a high no-show prediction in contrast to the traditional scheduling where all patients have the same appointment time length. Thus, it would help to balance the cost of no-shows and the use of the clinic resources effectively.

Our results demonstrate the importance of considering the right representation and weighting the importance of one feature over others. More specifically, our results of the developed models utilizing the sequence of binary representation show that the accuracy score has improved noticeably as the best model achieves AUC of 0.718 beating the baseline approach with a percentage increase of over 62%. Additionally, the results of evaluation of the generalization of the model, using the develop models to a new dataset, show that the developed models are not limited to a specific dataset confirming the generality and ensuring the applicability with negligible performance degradation.

Several potential applications could utilize our model. For instance, deploying a recommender system on top of the predictive no-show model could even boost certainty and improve costs, resources, and patient satisfaction. For a given patient, the recommender system, with the aid of the no-show model, suggests the optimal time slot at which she is more likely to attend. For example, if a patient had missed some appointments in the past in the early morning, the recommender system may suggest an afternoon time slot. Additionally, the system could suggest a reduced time slot appointment for a patient with a high no-show risk; even if the patient attended, the inconvenience for the following

appointments would be minimal. On the other hand, her absence would not considerably affect the utilization of the available resources.

Our predictive model has some limitations. First, some variables in the datasets are unavailable, such as home address and distance to the clinic, which could even help the model accuracy overall and reveal more meaningful insights. Second, external factors such as the daily weather or some significant events have not been considered. Including such factors could improve the model accuracy.

RELATED WORK

There is an abundance of papers that studied the behavior of no-shows in various domains (*Lawrence, Hong & Cherrier, 2003*; *Li et al., 2019a*; *Phumchusri & Maneesophon, 2014*), such as healthcare, airlines, hotel, and restaurant reservations. In the healthcare domain, earlier work in the literature (*Dove & Schneider, 1981*; *Oppenheim, Bergman & English, 1979*; *Shonick, 1977*) studied the factors of no-show behavior and evaluated the effectiveness of various interventions, such as sending reminders and charging service fees, to mitigate no-shows (*Garuda, Javalgi & Talluri, 1998*). For predicting no-shows and the number of no-show patients, early work used statistics and data mining techniques, such as decision trees (*Dove & Schneider, 1981*) and association rules (*Glowacka, Henry & May, 2009*).

These techniques are used to discover patterns and properties of the data. They are unable to effectively capture and predict complex behavior, such as no-shows, that involve several factors. Furthermore, these techniques are not able to adapt to changing behavior and new data effectively. They also failed to implement an appropriately targeted intervention such as using text messages or reminder calls 24 h in advance to reduce no-shows.

The early 2010s have seen a proliferation of studies to predict no-shows for individual patients, using various machine learning and probabilistic modeling techniques. This includes employing one or more techniques, such as logistic regression (*Alaeddini et al., 2011*; *Alshaya, McCarren & Al-Rasheed, 2019*; *Chiara & Domenic, 2018*; *Ding et al., 2018*; *Goffman et al., 2017*; *Harvey et al., 2017*; *Lenzi, Ben & Stein, 2019*; *Mohammadi et al., 2018*; *Nelson et al., 2019*; *Salazar et al., 2020*; *Srinivas & Ravindran, 2018*), decision trees (*Salazar et al., 2020*), random forests (*Alshaya, McCarren & Al-Rasheed, 2019*; *Ferro et al., 2020*; *Nelson et al., 2019*; *Salazar et al., 2020*; *Srinivas & Ravindran, 2018*), neural networks (*Ferro et al., 2020*; *Mohammadi et al., 2018*; *Srinivas & Ravindran, 2018*), gradient boosting (*Alshaya, McCarren & Al-Rasheed, 2019*; *Denney, Coyne & Rafiqi, 2019*; *Nelson et al., 2019*; *Srinivas & Ravindran, 2018*), NaïveBayes (*Mohammadi et al., 2018*), Bayesian Belief Networks (*Li et al., 2019b*; *Simsek et al., 2021*; *Topuz et al., 2018*), Support Vector Machine (*Alshaya, McCarren & Al-Rasheed, 2019*; *Nelson et al., 2019*), Hoeffding trees (*AlMuhaideb et al., 2019*) and others (*Alaeddini et al., 2011*; *Denney, Coyne & Rafiqi, 2019*). *Carreras-García et al. (2020)* conduct a comprehensive systematic literature review analyzing and highlighting the limitations of the proposed predictive patient no-show models.

Some papers used different techniques to capture features (*Goffman et al., 2017*), select variables (*Simsek et al., 2021*; *Topuz et al., 2018*), and balance the classes of data (*Alshaya,*

McCarren & Al-Rasheed, 2019; Nelson et al., 2019; Simsek et al., 2021). Techniques include class weights (*Nelson et al., 2019*), random undersampling (*Nelson et al., 2019; Simsek et al., 2021*), Synthetic Minority Oversampling Technique (SMOTE) (*Nelson et al., 2019; Simsek et al., 2021*), and elastic net (*Topuz et al., 2018*). To effectively capture no-show history, *Goffman et al. (2017)* capture patients' no-show history as binary sequences and uses a Markov model to compute a probability for each observed sequence of no-shows in the dataset. First, the training dataset is processed to compute the no-show probability for each observed sequence of no-show history. Each probability is computed by considering possible next future outcomes for the corresponding no-show history sequence in the dataset. Our work is similar in that it represents no-show history as binary sequences. However, our approach maintains the no-show history sequences and tries to learn the behavior associated with different sequences from data training, including all features. Their work requires processing the dataset to compute probability tables for different lengths of sequences and reducing them to numbers based solely on observed no-show history in the dataset.

Similar to our findings, several papers (*AlMuhaideb et al., 2019; Chiara & Domenic, 2018; Goffman et al., 2017; Harvey et al., 2017*) reported that the features of previous no-show history and appointment lead time were strong predictors of future no-shows. *Dantas et al. (2018)* summaries in a condensed systematic review of 105 studies focused on no-shows in appointment scheduling revealing that socioeconomic status and history of no-shows have a huge impact on the patients to miss their appointments. The importance of other features, such as insurance (*Harvey et al., 2017*), multiple same-day appointment (*Goffman et al., 2017*), gender (*Chiara & Domenic, 2018; Simsek et al., 2021*), age (*Chiara & Domenic, 2018; Simsek et al., 2021*), care cost (*Nelson et al., 2019*), season (*Li et al., 2019b*), count of previous appointments (*Nelson et al., 2019*), online doctor rating (*Fan et al., 2021*), and utilizing a local weather data (*Liu et al., 2022*), varies across studies. This is because the no-show behavior varies by demographics, clinic specialties, and other factors. This calls for fine-grained studies covering specific specialties, clinics (*Ding et al., 2018; Mohammadi et al., 2018*) and demographics to discover useful patterns that may help to predict no-shows. For example, (*Mohammadi et al., 2018*) collected data from 10 facilities and observed differences in the strong predictors of no-shows across clinics and facilities. *Ding et al. (2018)* examined patient appointments across 14 specialties, not including dentistry, and 55 clinics. Their results showed that fine-grained data, which captures clinic and specialty details, is more likely to perform better.

Several research papers have contributed by studying data from different populations and specialties, see Table 9. These papers collected data from the United States (*Alaeddini et al., 2011; Denney, Coyne & Rafiqi, 2019; Ding et al., 2018; Goffman et al., 2017; Harvey et al., 2017; Li et al., 2019b; Mohammadi et al., 2018; Srinivas & Ravindran, 2018; Topuz et al., 2018*), United Kingdom (*Nelson et al., 2019*), Brazil (*Alshaya, McCarren & Al-Rasheed, 2019; Lenzi, Ben & Stein, 2019; Salazar et al., 2020; Simsek et al., 2021*), Italy (*Chiara & Domenic, 2018*), Columbia (*Ferro et al., 2020*), and Saudi Arabia (*AlMuhaideb et al., 2019*). Their data covered various clinics and specialties, such as primary care (*Denney, Coyne & Rafiqi, 2019; Ferro et al., 2020; Lenzi, Ben & Stein, 2019; Simsek et al.,*

Table 9 Studied specialties and population in the literature.

Study	Speciality	Location
Salazar et al. (2020)	Unknown	Brazil
Ferro et al. (2020)	Primary care	Columbia
AlMuhaideb et al. (2019)	Various	Saudi Arabia
Alshaya, McCarren & Al-Rasheed (2019)	Unknown	Brazil
Denney, Coyne & Rafiqi (2019)	Primary care	United States
Nelson et al. (2019)	Radiology	United Kingdom
Lenzi, Ben & Stein (2019)	Primary care	Brazil
Li et al. (2019b)	Various	United States
Simsek et al. (2021)	Primary care	Brazil
Chiara & Domenic (2018)	Radiology	Italy
Srinivas & Ravindran (2018)	Primary care	United States
Ding et al. (2018)	Various	United States
Mohammadi et al. (2018)	Primary care (community health centers)	United States
Topuz et al. (2018)	Pediatric	United States
Goffman et al. (2017)	Primary care (veterans health centers)	United States
Harvey et al. (2017)	Radiology	United States
Alaeddini et al. (2011)	Primary care (veterans health centers)	United States

2021; *Srinivas & Ravindran, 2018*), community health centers (*Mohammadi et al., 2018*), veteran facilities (*Alaeddini et al., 2011*; *Goffman et al., 2017*), Radiology (*Chiara & Domenic, 2018*; *Harvey et al., 2017*; *Nelson et al., 2019*), Pediatrics (*Topuz et al., 2018*), and a mix of specialities (*AlMuhaideb et al., 2019*; *Ding et al., 2018*; *Li et al., 2019b*). Compared to previous studies, this article focuses on predicting no-shows for dental appointments in Saudi Arabia.

No-show predictions are utilized to improve the scheduling of appointments, resulting in better utilization of resources and satisfaction of patients. No-show predictions may be used to guide overbooking decisions (*Chiara & Domenic, 2018*; *Li et al., 2019b*; *Srinivas & Ravindran, 2018*), facilitate targeted interventions (*Li et al., 2019b*; *Simsek et al., 2021*) and reduce the cost of appointment reminders (*Ferro et al., 2020*; *Nelson et al., 2019*). For example, *Srinivas & Ravindran (2018)* proposed multiple overbooking and sequencing strategies for multi-stage appointments. Similarly, *Chiara & Domenic (2018)* presented an algorithm to compute the minutes eligible for overbooking.

Li et al. (2019b) developed two models: one to predict whether a patient would confirm her appointment or not and the second to compute the no-show probability. The first model clusters patients into three groups, which may be used for targeted interventions, and its output is fed to the second model, along with other features. The output of the second model is used to improve overbooking strategies with respect to several parameters, such as patient waiting cost, clinic overtime cost, and idle time cost. *Nelson et al. (2019)* proposed an equation to calculate the benefit of sending reminders to a particular set of patients compared to all patients.

CONCLUSION

We study the patient no-show problem in the context of the dental appointment where more than 40% of the patients miss their appointments for various reasons. Consequently, it has a significant negative impact on the clinic resources and the patient's oral health. Traditional solutions such as overbooking and sending a reminder have limited success and failed to address the problem effectively. In this work, we utilize state-of-the-art machine learning algorithms to develop predictive models to anticipate no-shows. We use locally collected dataset to develop the predictive models. The key feature utilized in the developing the predictive models is the patient prior no-show history. Thus, we propose a novel representation as a binary sequence of events that the model can effectively learn the behavior of the attendance. The results show that our approach is both viable and generalizable. The experiments show the best performing model achieves AUC of 0.718, and the developed models outperformed the baseline approach with a percentage increase of over 62%. Finally, we discuss various applications, such as a personalized recommender system for appointment time, that could be built on the deployed models to boost the utilization of the dental clinic resources and increase overall patient satisfaction. Our code and the dataset are freely available for the community to reproduce and investigate other ideas.

In future work, we plan to collect the dataset from the same dental clinics to study the patient no-show problem during and post the outbreak of the COVID-19 pandemic and how the results are different before the pandemic. Additionally, combining predictive no show model and cancellation model is another promising direction that may reduce the cost and utilize the healthcare clinic resources effectively.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Yazeed Alabdulkarim conceived and designed the experiments, performed the experiments, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.
- Mohammed Almukaynizi conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Abdulmajeed Alameer analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Bassil Makanati conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.
- Riyadh Althumairy analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Abdulaziz Almaslukh analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw data is available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.1147#supplemental-information.

REFERENCES

Alaeddini A, Yang K, Reddy C, Yu S. 2011. A probabilistic model for predicting the probability of no-show in hospital appointments. *Health Care Management Science* **14(2)**:146–157 DOI 10.1007/s10729-011-9148-9.

AlMuhaideb S, Alswailem O, Alsubaie N, Ferwana I, Alnajem A. 2019. Prediction of hospital no-show appointments through artificial intelligence algorithms. *Annals of Saudi Medicine* **39(6)**:373–381 DOI 10.5144/0256-4947.2019.373.

Alshaya S, McCarren A, Al-Rasheed A. 2019. Predicting no-show medical appointments using machine learning. In: *International Conference on Computing*. Springer, 211–223.

American Dental Association. 2021. Survey reports. *Available at https://www.ada.org/resources/research/health-policy-institute/dental-care-market*.

Baskaradoss JK. 2016. The association between oral health literacy and missed dental appointments. *The Journal of the American Dental Association* **147(11)**:867–874 DOI 10.1016/j.adaj.2016.05.011.

Bergstra J, Bengio Y. 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13(2)**:281–305 DOI 10.5555/2188385.2188395.

Breiman L. 2001. Random forests. *Machine learning* **45(1)**:5–32 DOI 10.1023/A:1010933404324.

Carreras-García D, Delgado-Gómez D, Llorente-Fernández F, Arribas-Gil A. 2020. Patient no-show prediction: a systematic literature review. *Entropy* **22(6)**:675 DOI 10.3390/e22060675.

Chiara A, Domenic S. 2018. Using overbooking to manage no-shows in an italian healthcare center. *BMC Health Services Research* 180–185 DOI 10.1186/s12913-018-2979-z.

Dantas LF, Fleck JL, Oliveira FLC, Hamacher S. 2018. No-shows in appointment scheduling–a systematic literature review. *Health Policy* 122(4):412–421 DOI 10.1016/j.healthpol.2018.02.002.

Denney J, Coyne S, Rafiqi S. 2019. Machine learning predictions of no-show appointments in a primary care setting. *SMU Data Science Review* 2(1):2.

Ding X, Gellad ZF, Mather C III, Barth P, Poon EG, Newman M, Goldstein BA. 2018. Designing risk prediction models for ambulatory no-shows across different specialties and clinics. *Journal of the American Medical Informatics Association* 25(8):924–930 DOI 10.1093/jamia/ocy002.

Dove HG, Schneider KC. 1981. The usefulness of patients' individual characteristics in predicting no-shows in outpatient clinics. *Medical Care* 19(7):734–740 DOI 10.1097/00005650-198107000-00004.

El-Habil AM. 2012. An application on multinomial logistic regression model. *Pakistan Journal of Statistics and Operation Research* 8(2):271–291 DOI 10.18187/pjsor.v8i2.234.

Fan G, Deng Z, Ye Q, Wang B. 2021. Machine learning-based prediction models for patients no-show in online outpatient appointments. *Data Science and Management* 2(1):45–52 DOI 10.1016/j.dsm.2021.06.002.

Ferro DB, Brailsford S, Bravo C, Smith H. 2020. Improving healthcare access management by predicting patient no-show behaviour. *Decision Support Systems* 138:113398 DOI 10.1016/j.dss.2020.113398.

Friedman JH. 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38(4):367–378 DOI 10.1016/S0167-9473(01)00065-2.

Gama J, Žliobaitė I, Bifet A, Pechenizkiy M, Bouchachia A. 2014. A survey on concept drift adaptation. *ACM Computing Surveys (CSUR)* 46(4):1–37 DOI 10.1145/2523813.

Garuda SR, Javalgi RG, Talluri VS. 1998. Tackling no-show behavior: a market-driven approach. *Health Marketing Quarterly* 15(4):25–44 DOI 10.1300/J026v15n04_02.

Gier J. 2017. Missed appointments cost the us healthcare system $150 b each year. *Health Management Technology* 2:.

Glowacka KJ, Henry RM, May JH. 2009. A hybrid data mining/simulation approach for modelling outpatient no-shows in clinic scheduling. *Journal of the Operational Research Society* 60(8):1056–1068 DOI 10.1057/jors.2008.177.

Goffman RM, Harris SL, May JH, Milicevic AS, Monte RJ, Myaskovsky L, Rodriguez KL, Tjader YC, Vargas DL. 2017. Modeling patient no-show history and predicting future outpatient appointment behavior in the veterans health administration. *Military Medicine* 182(5–6):e1708–e1714 DOI 10.7205/MILMED-D-16-00345.

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. *Journal of the American College of Radiology* 14(10):1303–1309 DOI 10.1016/j.jacr.2017.05.007.

Kheirkhah P, Feng Q, Travis LM, Tavakoli-Tabasi S, Sharafkhaneh A. 2015. Prevalence, predictors and economic consequences of no-shows. *BMC Health Services Research* 16(1):1–6 DOI 10.1186/s12913-015-1243-z.

Lawrence RD, Hong SJ, Cherrier J. 2003. Passenger-based predictive modeling of airline no-show rates. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 397–406.

Lenzi H, Ben ÂJ, Stein AT. 2019. Development and validation of a patient no-show predictive model at a primary care setting in Southern Brazil. *PLOS ONE* **14(4)**:e0214869 DOI 10.1371/journal.pone.0214869.

Li K-P, Chen S-L, Chiu W-H, Lu W-C. 2019a. Strategies of reduce customerâ€™ s no-show probability at restaurants. *The Journal of Social Sciences Research* **5(1)**:145–152.

Li Y, Tang SY, Johnson J, Lubarsky DA. 2019b. Individualized no-show predictions: effect on clinic overbooking and appointment reminders. *Production and Operations Management* **28(8)**:2068 DOI 10.1111/poms.13033.

Liu D, Shin W-Y, Sprecher E, Conroy K, Santiago O, Wachtel G, Santillana M. 2022. Machine learning approaches to predicting no-shows in pediatric medical appointment. *NPJ Digital Medicine* **5(1)**:1–11 DOI 10.1038/s41746-022-00594-w.

Marbouh D, Khaleel I, Al Shanqiti K, Al Tamimi M, Simsekler MCE, Ellahham S, Alibazoglu D, Alibazoglu H. 2020. Evaluating the impact of patient no-shows on service quality. *Risk Management and Healthcare Policy* **13**:509–517 DOI 10.2147/RMHP.S232114.

Mohammadi I, Wu H, Turkcan A, Toscos T, Doebbeling BN. 2018. Data analytics and modeling for appointment no-show in community health centers. *Journal of Primary Care & Community Health* **9**:2150132718811692 DOI 10.1177/2150132718811692.

Nelson A, Herron D, Rees G, Nachev P. 2019. Predicting scheduled hospital attendance with artificial intelligence. *NPJ Digital Medicine* **2(1)**:1–7 DOI 10.1038/s41746-019-0103-3.

Oppenheim GL, Bergman J, English E. 1979. Failed appointments: a review. *Journal of Family Practice* **8(4)**:789–796.

Phumchusri N, Maneesophon P. 2014. Optimal overbooking decision for hotel rooms revenue management. *Journal of Hospitality and Tourism Technology* **5(3)**:261–277 DOI 10.1108/JHTT-03-2014-0006.

Rust CT, Gallups NH, Clark WS, Jones DS, Wilcox WD. 1995. Patient appointment failures in pediatric resident continuity clinics. *Archives of Pediatrics & Adolescent Medicine* **149(6)**:693–695 DOI 10.1001/archpedi.1995.02170190103018.

Salazar LH, Fernandes A, Dazzi R, Garcia N, Leithardt VR. 2020. Using different models of machine learning to predict attendance at medical appointments. *Journal of Information Systems Engineering and Management* **5(4)**:em0122 DOI 10.29333/jisem/8430.

Shabbir A, Alzahrani M, Khalid AA. 2018. Why do patients miss dental appointments in eastern province military hospitals, Kingdom of Saudi Arabia? *Cureus* **10(3)**:e2355 DOI 10.7759/cureus.2355.

Shonick W. 1977. An approach to reducing the adverse effects of broken appointments in primary care systems. *Medical Care* **15(5)**:419–429 DOI 10.1097/00005650-197705000-00008.

Simsek S, Dag A, Tiahrt T, Oztekin A. 2021. A bayesian belief network-based probabilistic mechanism to determine patient no-show risk categories. *Omega* **100**:102296 DOI 10.1016/j.omega.2020.102296.

Srinivas S, Ravindran AR. 2018. Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: a prescriptive analytics framework. *Expert Systems with Applications* **102(4)**:245–261 DOI 10.1016/j.eswa.2018.02.022.

Topuz K, Uner H, Oztekin A, Yildirim MB. 2018. Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and bayesian belief network. *Annals of Operations Research* **263(1)**:479–499 DOI 10.1007/s10479-017-2489-0.

Zeng B, Zhao H, Lawley M. 2013. The impact of overbooking on primary care patient no-show. *IIE Transactions on Healthcare Systems Engineering* **3(3)**:147–170 DOI 10.1080/19488300.2013.820239.