Ensemble Learning in Predicting Medical Appointment No-shows

A

report submitted in partial fulfillment for the award of the degree of

Bachelors of Technology

in

Information Technology

By

Harsh Yadav : 2020IMG-026

Under the Supervision of

Dr.Narinder Singh Punn

Department of Computer Science

ABV-INDIAN INSTITUTE OF INFORMATION TECHNOLOGY
AND MANAGEMENT GWALIOR

GWALIOR, INDIA

DECLARATION

I hereby certify that the work, which is being presented in the report/thesis, entitled Ensemble Learning in Predicting Medical Appointment No-shows, in fulfillment of the requirement for the award of the degree of Bachelor of Technology and submitted to the institution is an authentic record of my/our own work carried out during the period May-2023 to August-2023 under the supervision of Dr. Narinder Singh Punn. I also cited the reference about the text(s)/figure(s)/table(s) from where they have been taken.

Dated: **Signature of the candidate**

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dated: **Signature of supervisor**

Acknowledgements

I am highly indebted to Dr. Narinder Singh Punn, for his esteemed mentorship, and for allowing me to freely explore and experiment with various ideas in the course of making this project a reality. The leeway I was given went a long way towards helping cultivate a genuine hunger for knowledge and keeping up the motivation to achieve the best possible outcome. I can genuinely say that this Bachelor's Thesis Project (BTP) made me explore many areas of machine learning that are new to me, and kindled an interest to further follow up on some of those areas. Moreover, the semi-successful completion of this project has brought with it great satisfaction and more importantly, confidence in my ability to produce more high-quality non-trivial artificial intelligence systems that can make a difference in the real-world. I would like to sincerely express my gratitude to this prestigious institution for providing me and my colleagues with the opportunity to pursue this BTP. It is an honor to be able to work on such an important academic project under the guidance and support I am provided with. I am grateful for the resources and facilities provided by this institution, which have been instrumental in enabling me to conduct my research and complete this project. Moreover, I deeply appreciate the efforts of my professors in mentoring and fairly evaluating our works.

Harsh Yadav

Abstract

The purpose of this project is to explore a dataset of medical appointments and investigate the factors that are associated with patients not showing up for their appointments. The dataset comprises over 100,000 medical appointments to investigate the factors correlated with patients not showing up for their scheduled appointments. The dataset includes essential variables such as patient age, gender, medical history, and whether they received reminder SMS notifications. The primary objectives of the study are twofold: first, to identify the key factors associated with missed appointments, and second, to accurately predict the no-shows so that we can improve the efficiency of the health care system. The project includes the application of various machine learning models, such as Logistic Regression, Decision Trees, Random Forest, Gradient Boost, Extra tree, XGBoost, AdaBoost, Bagging, KNN, Ridge Classifier, Naive Bayes and a Voting Classifier ensemble, to accurately predict appointment no-shows. Through data-driven insights, this project aspires to propose personalized patient reminder strategies, targeted communication approaches, and optimized appointment management to cater to individual patient needs and preferences. The ultimate goal is to enhance patient adherence, optimize resource allocation, and improve overall healthcare efficiency, leading to better patient outcomes and satisfaction. Missed appointments can have serious consequences for both patients and healthcare providers, such as delayed diagnosis and treatment, wasted resources, and decreased patient satisfaction. Therefore, understanding the factors that contribute to missed appointments is important for improving the quality of care and reducing costs. By addressing the pressing issue of medical appointment no-shows, this project seeks to contribute to the advancement of healthcare delivery, ensuring that healthcare facilities can effectively meet patient needs, reduce operational costs, and provide high-quality, patient-centered care.

Keywords: *Medical Appointments, No-Show, Decision Tree, Random Forest, Ensemble Learning, Patient-Centered Care, Healthcare Optimization*

Contents

List of Figures

List of Tables

List of Acronyms

KNN	K-Nearest Neighbours
Ada-Boost	Adaptive Boost
ETC	Extra Tree Classifier
EDA	Exploratory Data Analysis

1

Introduction

This chapter offers an overview of the subject matter by presenting background information on Medical Appointment No-shows and emphasizes the potential of deep ensemble learning. Moreover, this chapter entails the motivating factors that instigated the investigation of this topic.

1.1 Introduction

In modern healthcare systems, missed medical appointments pose a significant challenge to both patients and healthcare providers, leading to delayed diagnoses, suboptimal treatment, wasted resources, and reduced patient satisfaction. These no-show occurrences are not only detrimental to individual patients but also have far-reaching implications for the overall efficiency and cost-effectiveness of healthcare facilities. Consequently, understanding the factors associated with patients not attending their scheduled appointments is of paramount importance to improve the quality of care and optimize resource utilization. This project aims to explore a vast dataset containing information on over 100,000 medical appointments to identify the primary determinants of appointment non-attendance. The dataset includes crucial variables, such as patient age, gender, medical history, and whether they received reminder SMS notifications.

Various machine learning models, such as Logistic Regression, Decision Tree, Random Forest, XGBoost, AdaBoost, etc [1] have been used to accurately predict appointment no-shows. These models aid in uncovering underlying patterns and correlations within the dataset, contributing to evidence-based recommendations for the improvement of the scheduling process [2]. The motivation behind this project lies in its potential to revolutionize healthcare delivery. By understanding the factors influencing appointment no-shows, we can offer evidence-based recommendations to healthcare providers, leading to personalized patient reminders, targeted communication strategies, and optimized appointment management. These interventions, tailored to individual patient needs, aim to enhance patient adherence and ultimately improve healthcare system efficiency. The significance of this project lies in its potential to empower healthcare providers with actionable insights to address the complex challenge of missed appointments. Through a data-driven approach, we aspire to contribute to the larger goal of optimizing healthcare systems and delivering high-quality care that is accessible, efficient, and patient-centered.

By uncovering the factors contributing to no-show appointments and constructing a predictive model, this project endeavors to empower healthcare providers with a strategic tool for resource allocation, appointment scheduling optimization [3], and patient engagement strategies. Ultimately, the insights gained from this research have the potential to reduce the number of no-show appointments, enhance the efficiency of healthcare delivery, and improve patient outcomes. As we delve into the realm of medical appointments and no-show predictions, the following sections of this report will delve deeper into the dataset, analysis methods, findings, and implications, leading to the development of a robust predictive model.

1.2 Motivation

In the realm of healthcare, the harmonious synchronization between patients and their medical appointments is essential to ensure the delivery of timely and effective medical care. However, the persistence of "no-show appointments" - instances where patients fail to attend their scheduled medical visits without prior notice - presents a pressing challenge to healthcare providers worldwide. This phenomenon not only disrupts the healthcare workflow but also jeopardizes patient well-being and strains valuable healthcare resources. Therefore, delving into the intricacies of no-show appointments and understanding the underlying factors becomes a crucial endeavor with far-reaching implications.

The motivation for this project stems from a deep-seated desire to bridge the gap between patient expectations and healthcare provider schedules. No-show appointments, while seemingly inconspicuous, are indicative of a complex interplay of individual, social, and logistical factors. They often carry implications for both patients and healthcare systems, leading to inefficient resource utilization, increased wait times for other patients, delayed diagnoses, and overall compromised healthcare quality. By exploring a dataset encompassing over 100,000 medical appointments, this project aspires to shed light on the key determinants that contribute to no-show appointments. Unraveling these determinants holds the promise of transforming the healthcare landscape in profound ways:

(i) Improved Patient Engagement and Care: Understanding why patients miss appointments can open avenues for healthcare providers to enhance patient engagement and communication. Tailored strategies, such as personalized reminders or proactive outreach for high-risk individuals, could foster a stronger patient-provider relationship and increase appointment attendance rates.

(ii) Resource Optimization: Predicting no-show appointments empowers healthcare facilities to allocate resources more judiciously. With a clearer understanding of the variables that correlate with no-shows, hospitals and clinics can manage staff, facilities, and time slots more efficiently, reducing both financial strain and patient frustration.

(iii) Informed Decision-Making: Armed with data-driven insights, healthcare administrators and policy-makers can make more informed decisions regarding appointment scheduling, patient education initiatives, and resource allocation. This data-driven approach can lead to more patient-centered healthcare practices.

(iv) Better Health Outcomes: Minimizing no-show appointments has the potential to expedite diagnoses and treatments, thus leading to improved patient health outcomes. Timely medical care is often crucial, and reducing no-shows can contribute to faster interventions and improved overall well-being.

In essence, the motivation behind this project is not merely statistical analysis, but a profound drive to create a tangible impact on the healthcare experience for patients, providers, and the healthcare system as a whole. By delving into the realm of no-show appointments and their multifaceted determinants, this project aspires to contribute to a more efficient, patient-centric, and equitable healthcare ecosystem.

2

A Review on Medical Appointment No-shows

This chapter responds to the significant amount of research assigned to understanding the efficient methods of Medical Appointment no-shows. We examine some of the literature and briefly review the development of former proposed methods and the research gaps.

2.1 Review on Existing Methods

In the realm of healthcare operations and patient management, accurately predicting missed medical appointments has become an essential focus. Earlier models, encompassing a range of techniques including logistic regression, random forest, and decision trees, have sought to enhance the accuracy of predicting no-show rates. This literature review delves into the landscape of research endeavors that have aimed to improve the understanding and prediction of patient behavior with regards to appointment attendance. Earlier Studies explored missed medical appointments date back to before the 1980s when they were initially referred to as "broken appointments". Overtime, the term "no-shows" gained prevalence, marking a shift in research nomenclature. Todd Molfenter (2013) [4] noted that the prevalence of no-shows varies across countries, healthcare systems, and clinics, with rates typically falling in the range of 20-30%. Researchers have sought to identify factors contributing to no-shows, such as forgetfulness, health conditions, socioeconomic status, lack of transportation, and patients' appointment history.

In recent years, machine learning techniques have gained traction for predicting missed appointments. Machine learning models offer the advantage of identifying both linear and non-linear patterns within the data. Iman Mohammadi (2018) [5] employed Logistic Regression, Decision Trees, and Random Forest to predict no-shows with accuracy rates ranging from 71.8% to 72.9%. Similarly, Andrew McCarren (2019) [6] developed a predictive model using empirical Markov models and logistic regression to identify patients at risk of missing appointments. Efforts to optimize model thresholds have emerged as a technique to improve predictive accuracy. Studies like that of have introduced cost-sensitive approaches to determine optimal thresholds for logistic regression models. By assigning varying costs to show and no-show misclassifications, thresholds were tailored to minimize errors while considering the differential impact of these errors.

2.2 Research Analysis of Existing Methods

The research landscape surrounding accurate prediction of missed medical appointments has seen significant evolution, with earlier models and techniques gradually giving way to more sophisticated approaches. This literature review explores the historical context of research endeavors aimed at understanding and predicting patient behavior related to appointment attendance, tracing the progression from the term "broken appointments" to the more contemporary concept of "no-shows." The concept of missed medical appointments dates back to the 1960s, where early investigations shed light on the phenomenon. These foundational studies laid the groundwork for subsequent research, leading to the emergence of the term "no-shows" to describe appointment non-attendance. Over time, the research focus shifted from mere descriptive observations to a more comprehensive understanding [7] of the factors contributing to missed appointments.

Alan C. Wilson, Karrie Fields (2010) [8] contributed valuable insights by highlighting the variability in no-show rates across different countries, healthcare systems, and clinics, typically falling within the range of 20-30%. Researchers began to identify various factors [9] influencing appointment attendance, such as forgetfulness, health conditions, socioeconomic status, lack of transportation, and patients' appointment histories. This contextual understanding of the phenomenon paved the way for more targeted predictive modeling approaches. In recent years, the application of machine learning techniques has gained traction in the field. Machine learning models offer a distinct advantage by uncovering both linear and non-linear patterns within the complex patient data. Similarly, Dianbo Liu (2022) [10] leveraged machine learning algorithms including Logistic Regression, Decision Trees, and Random Forest to predict no-shows. The accuracy rates achieved, ranging from 71.8% to 72.9%, indicated the potential of these techniques in enhancing prediction capabilities.

Similarly, Eduard Incze (2022) [11] embraced a predictive modeling approach using empirical Markov models and logistic regression. By identifying patients at risk of missing appointments, they contributed to the ongoing efforts to reduce no-show rates. This shift towards more advanced models demonstrated the growing recognition of the significance of accurate predictions in healthcare operations. Efforts to optimize model thresholds have emerged as a valuable strategy to further enhance predictive accuracy. Studies like the one referenced as have introduced cost-sensitive approaches to determine optimal thresholds for logistic regression models. By assigning varying costs to show and no-show misclassifications, these approaches aimed to fine-tune the thresholds to minimize errors while considering the varying impact of these errors. This refinement of model thresholds represents an iterative process aimed at achieving more precise predictions and aligning the model's outcomes with the real-world implications.

The research journey in predicting missed medical appointments has evolved from historical observations to data-driven insights and machine learning-driven predictions. The adoption of machine learning techniques and the exploration of optimization strategies for model thresholds underscore the commitment of researchers to enhancing patient management and healthcare operations through accurate prediction of appointment attendance behavior. As technology and methodologies continue to advance, this area of research holds promise for further improving patient outcomes and resource utilization.

2.3 Research Gaps

While the existing research landscape surrounding the prediction of missed medical appointments has witnessed significant progress, there exists a notable research gap that our project can effectively address. The historical studies and recent advancements have predominantly emphasized accuracy as a primary performance metric [12] for predictive models. However, our project's focus on precision and recall brings attention to an important dimension that has been relatively underexplored in this domain.

The existing literature, as highlighted in the provided information, has primarily concentrated on achieving higher accuracy rates in predicting no-shows. While accuracy is undoubtedly a crucial metric, it can be insufficient in scenarios where the consequences of false positives (predicting a patient will not attend an appointment when they actually do) and false negatives (predicting a patient will attend an appointment when they do not) have varying degrees of impact on healthcare operations and patient management. Accuracy does not explicitly account for the trade-off between correctly identifying potential no-shows and minimizing the occurrence of missed appointment.

By shifting the emphasis from accuracy to precision and recall, our project contributes a nuanced perspective to the prediction of missed appointments. It acknowledges that the consequences of incorrect predictions are not uniform and aims to strike a balance that aligns with the priorities of healthcare providers. This research gap is significant as it recognizes that reducing both false positives and false negatives can lead to improved resource allocation, patient engagement, and overall operational efficiency.

3

Problem Statement based on Identified Research Gaps

This chapter explains the formulation of the problem that this thesis addresses, as well as it outlines the thesis objectives.

3.1 Problem Formulation

In the realm of healthcare operations and patient management, accurately predicting missed medical appointments has gained prominence due to its potential to optimize resource allocation, enhance patient engagement, and improve overall healthcare system efficiency [13]. However, the existing focus on accuracy as a primary performance metric for predictive models falls short when dealing with imbalanced datasets, such as those characterized by a no-show rate of approximately 20%. This research gap underscores the critical need to shift the evaluation paradigm towards precision and recall, which are more relevant in scenarios where the consequences of false predictions have varying degrees of impact.

The problem at hand revolves around the challenge of developing a predictive model that is specifically tailored to address the unique demands of healthcare operations and patient behavior within the context of missed medical appointments. Traditional accuracy-oriented models [14] fail to capture the nuances of an imbalanced dataset, potentially leading to an overemphasis on the majority class while overlooking the minority class of potential no-show patients. As a result, valuable opportunities to engage with patients who are likely to miss appointments may be missed, adversely affecting patient outcomes and healthcare resource utilization. In healthcare operations, accurately predicting missed appointments is crucial for efficient resource allocation and patient engagement. However, the focus on accuracy for predictive models is problematic with imbalanced datasets, like a 20% no-show rate. Shifting to precision and recall as primary metrics is essential due to varying impact of false predictions. The challenge is crafting a predictive model tailored to healthcare nuances, as traditional accuracy-based models overlook potential no-shows. This project aims to optimize models for precision and recall, prioritizing patient outcomes and resource utilization. The central question this project aims to address is:
How can predictive models be refined and optimized to prioritize precision and recall over accuracy in predicting potential no-show patients within an imbalanced dataset context?

3.2 Thesis Objective

- To explore and implement the literature work on various Medical Appointment no-shows.
- To do Exploratory Data Analysis(EDA) to identify the factors that contribute to patients not showing up for their scheduled appointments.
- To optimize precision and recall metrics in predictive models. Fine-tune models to strike a balance between minimizing false positives (precision) and false negatives (recall), considering the costs associated with each.
- Evaluate their performance using the metrics; precision, recall, accuracy, roc_auc.

4

Proposed Methodology

This chapter provides a comprehensive discussion of the methodology employed in the project, including the various models that were used in the ensemble learning.

4.1 Data Preprocessing, Feature Engineering, Feature Scaling, Cross-Validation

We have a dataset of over 100,000 medical appointments, including patient demographics, medical history, and appointment details. First we performed data preprocessing tasks like handling missing values, outlier detection, and addressing data inconsistencies. Then we did feature engineering in which we added a feature called waiting day which is the number of days a patient has to wait for the appointment. We calculated waiting day by taking the difference between scheduled day and appointment day. Then we have done feature scaling in which we normalize or standardize the features to ensure that all variables have similar scales, preventing certain algorithms from dominating due to their magnitude.

Hyperparameter tuning plays a pivotal role [15] in enhancing the performance of machine learning models by finding the optimal configuration of parameters that yield the best results. In the context of predicting medical appointment no-shows, achieving higher precision and recall scores is of paramount importance to ensure effective allocation of healthcare resources and minimize missed appointments. To rigorously evaluate the models and ensure their generalization capability, we employed a shuffle split cross-validation strategy. This approach involves randomly shuffling the dataset and splitting it into training and validation sets for multiple iterations. Each iteration involves a different random split, allowing us to mitigate any potential biases that could arise from a fixed training-validation split.

4.2 Ensemble Learning

Ensemble learning is a powerful methodology that harnesses the strength of multiple individual models to create a more robust and accurate predictive system. Given the inherent complexities and challenges posed by imbalanced datasets, ensemble learning offers a compelling solution to improve the precision and recall of predictions, thus aligning with the patient-centric goals of healthcare operations. Ensemble learning operates on the principle that combining diverse models can overcome the limitations of individual algorithms by leveraging their complementary strengths. The different models used are as follows:

4.2.1 Logistic Regression

It is one of the important classification techniques that uses supervised learning [16]. It forecasts the likelihood of an event occurring for given values of the predictor variable. It creates a Decision boundary that classifies the different classes into different Segments.

4.2.2 Decision Tree Classifier

It is used for both regression and classification tasks that utilises a supervised learning [17]. This algorithm divide the population into 2 or more sub-samples that are homogeneous in nature and have less impurity. The tree grows in top-down approach and finds the Best split on the basis of different algorithm like gini index, entropy, Chi-square and Reduction in Variance.

4.2.3 Random Forest Classifier

It is a bagging technique that is used for both regression and classification problems. Random forest trains many decision trees by selecting some of the samples and features randomly [18] from the dataset and then averaging the Predictions.

4.2.4 Ridge Classifier

Ridge Classifier is a classifier version of the Ridge regressor. After converting binary targets to -1, 1, this classifier considers the problem as a regression job. The projected class is determined by the regressor's prediction sign.

4.2.5 Gradient Boost Classifier

It is also a Boosting techniques. So, it also combines multiple Weak learner to create strong learners. It minimize the loss using gradient boosting algorithm.

4.2.6 Extra Tree Classifier

A random sample of features is utilised in the same way as in random forests, but in Extremely Randomized Classifiers, random thresholds are generated for each input feature, and the best of these randomly generated thresholds is chosen as the splitting criterion.

4.2.7 Bagging Classifier

It is an ensemble technique that chooses random samples of the original dataset to fit on the Base classifier and then find the final result with max voting for classification problem or by averaging the prediction of each models for Regression problem.

4.2.8 AdaBoost Classifier

AdaBoost is a Boosting technique which generate iterative ensembles [19]. By combining several weak classifiers, this method creates a powerful classifier, and that strong classifier will help us to achieve better accuracy of the models.

4.2.9 K-Nearest Neighbors (KNN)

It is a simple supervised learning classification algorithm [20] that classifies the point to a particular class by looking at its most K nearest neighbors which measured using the distance metric and then chooses the most common KNN class and assigns it to the provided data point.

4.2.10 XGBoost Classifier

This algorithm is the Extreme version of Gradient Boost. It is a great mix of hardware and software optimization approaches that produce superior outcomes in the shortest period of time with the least amount of computer resources. This algorithm has built-in methods for Regularization, Missing values Imputation, Cross-Validation, Hyperparameter tuning and many others to improve the computation time and performance of the Model.

4.2.11 Naive Bayes

It is also a supervised learning classification method that based on Bayes' theorem with the "Naive" assumption [21]. This is a very fast algorithm which may be used for both Multiclass as well as Binary classification.

4.2.12 Voting Classifier

A Voting Classifier is an algorithm that takes input as an Ensemble of Machine Learning Models and based on the Prediction of these Models, it computes the Max Voting for a specific class. The Voting Classifier considers the class that receives the most votes to be the Predicted class. It computes the Max voting either using Soft or Hard voting.

5

Experiment and Results

The contents of this chapter encompass the description of the experiment set-up that was employed in the project, as well as a detailed account of the outcomes and results that were obtained.

5.1 Experiment setup

After doing Data preprocessing, Feature Engineering and Feature Scaling we do the experimental setup. First we setup the cross validation technique. Cross-validation, a crucial step in model evaluation, was carried out using the shuffle split technique to enhance the reliability of the results. In this approach, the dataset was randomly partitioned into training and validation sets for each iteration. Unlike fixed splits, shuffle split mitigates potential biases, providing a more accurate assessment of the model's generalization capabilities. By repeatedly shuffling and dividing the data, the model is exposed to diverse subsets, thus obtaining a robust understanding of its performance across different scenarios. This technique ensures that the model's effectiveness isn't contingent on a specific data split, contributing to a more comprehensive evaluation process. A variety of classification models were employed to predict medical appointment no-shows, including Logistic Regression, Decision Tree, Random Forest, Gradient Boost, Extra Tree, Bagging, AdaBoost, Ridge Classifier, K-Nearest Neighbors (KNN), XGBoost, and Naïve Bayes. The choice of these models was motivated by their suitability for classification tasks and their ability to capture different types of patterns in the data.

In this project, hyperparameter tuning was executed using techniques like GridSearchCV and RandomizedSearchCV. These methods systematically explore a range of parameter values to identify the combination that maximizes the model's performance metrics. GridSearchCV exhaustively searches through predefined parameter grids, testing all possible combinations. This method is particularly useful when the search space is relatively small and manageable. On the other hand, RandomizedSearchCV randomly samples a specified number of parameter combinations from the defined search space. This approach is beneficial for larger search spaces and can be more efficient in terms of computational resources. This project can be used to conduct various experiments by tweaking many parameters according to the users experimental setup and requirement. Table 5.1 shows the parameter settings for the experiment and simulations:

Table 5.1: Hyperparameters for Each Model

Model	**Hyperparameters**
Logistic Regression	{'C': 0.2}
Decision Tree	{'min_samples_leaf': 34}
Random Forest	{'n_estimators': 100, 'min_samples_split': 0.004, 'min_samples_leaf': 0.004, 'max_features': 'sqrt', 'criterion': 'gini', 'class_weight': 'balanced', 'bootstrap': False}
Gradient Boost	{'random_state': 42, 'n_estimators': 140, 'min_samples_leaf': 5, 'max_depth': 7, 'learning_rate': 0.05}
Extra Tree	{'class_weight': 'balanced', 'min_samples_leaf': 0.1, 'min_samples_split': 0.1, 'n_estimators': 120}
Bagging	{'max_features': 0.85, 'max_samples': 0.6, 'n_estimators': 10, 'random_state': 42, 'warm_start': False}
AdaBoost	{'n_estimators': 120, 'learning_rate': 0.12}
Ridge Classifier	{'alpha': 0.1, 'class_weight': 'balanced'}
KNN	{'n_neighbors': 8}
XGBoost	{'subsample': 0.8, 'scale_pos_weight': 3.9315801999048072, 'objective': 'binary:logistic', 'n_estimators': 80, 'min_child_weight': 2, 'max_depth': 8, 'learning_rate': 0.1, 'lambda': 4, 'gamma': 0, 'eval_metric': 'auc', 'alpha': 5}
Naïve Bayes	{'var_smoothing': 1e-06}

5.2 Results and discussion

5.2.1 Evaluation Metrics

Each model's performance is measured in terms of precision, recall, AUC Score and Accuracy. True Positive (TP) denotes that the patient is doing no-show and Predicted to be no-show, whereas False Positive (FP) denotes that the patient shows up but Predicted to no-show. The terms True Negative (TN) denotes that the patient no-shows and predicted to be no-show and False Negative (FN) denotes that patient no-show but predicted to show up.

Precision and recall are fundamental metrics for evaluating classification models, particularly in scenarios where class imbalance exists, as is often the case in predicting medical appointment no-shows. Precision measures the proportion of correctly predicted positive cases among all predicted positive cases. A high precision indicates that when the model predicts a positive outcome (no-show), it is likely to be correct. Recall, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive cases among all actual positive cases. A high recall indicates that the model is effectively capturing a significant portion of the actual positive cases.

$$\text{Accuracy (A)} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

5.2.2 Results

From our EDA we have found the following results:

- Younger patients tend to be associated with a higher likelihood of no-shows could be attributed to various factors. Younger individuals might perceive their health needs differently, underestimating the importance of medical appointments.

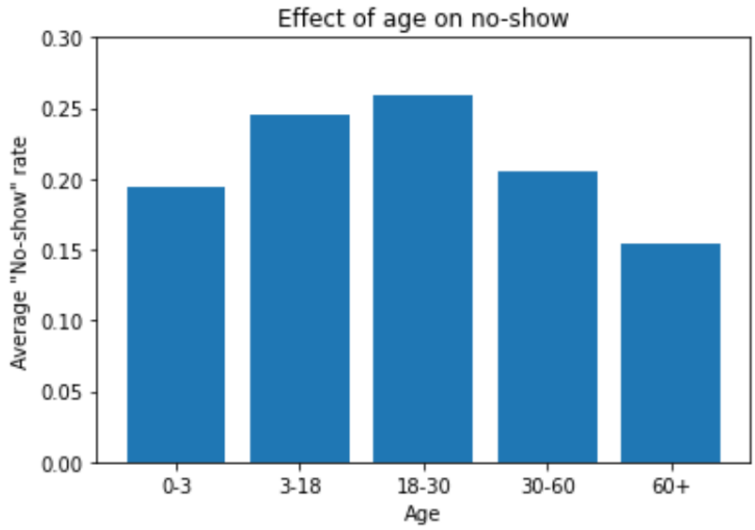

Figure 5.1: No-shows rate vs Age

- The observation that patients with a waiting time of 0 days tend to show up for their appointments suggests that same-day or immediate appointments are associated with higher attendance rates.

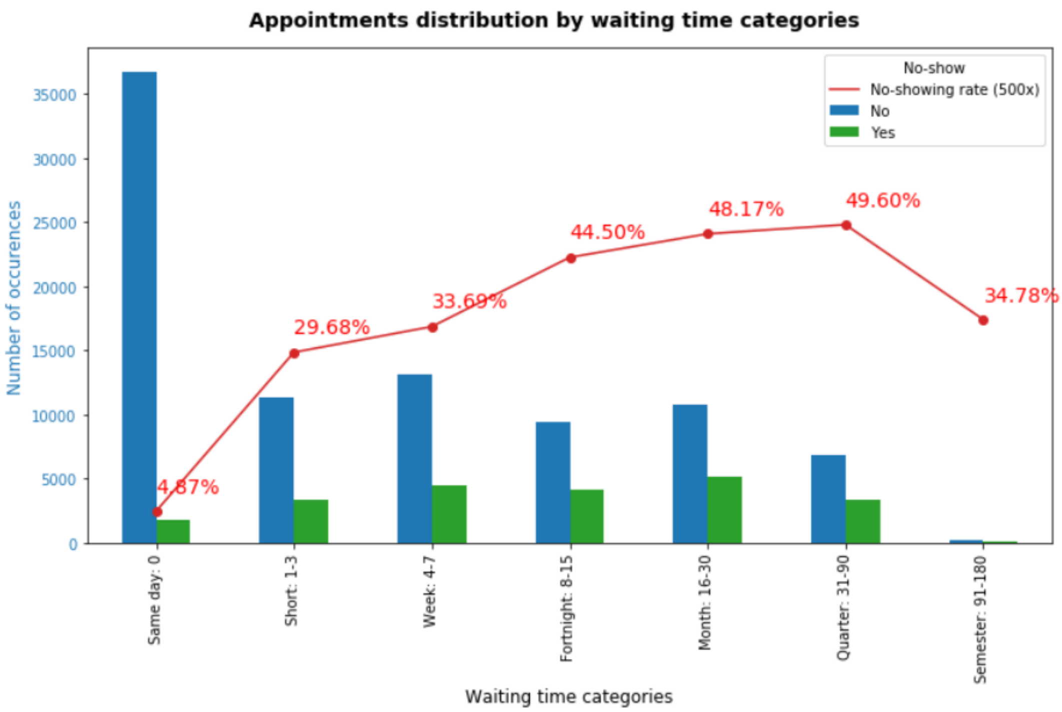

Figure 5.2: No-shows rate vs Waiting time

- Surprisingly, the finding that patients who receive SMS reminders tend to no-show suggests that the effectiveness of SMS reminders might not be as straightforward as assumed. Several reasons could contribute to this observation.

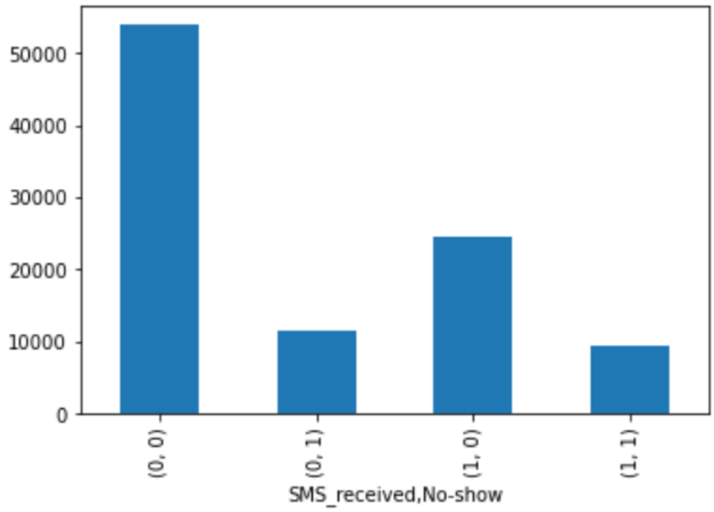

Figure 5.3: No. of patients vs SMS received, No-show

Now coming onto the results of the models used. Comparing the models based on their precision and recall, we observe the following trends:

- XGBoost, AdaBoost, Gradient Boost have high precision but low recall.
- Ridge Classifier, Logistic Regression, Voting Classifier have the highest recall and also the precision is also not affected much.

For all the models we can see that the accuracy has reduced but the precision and recall has increased which is an important evaluation metric in our imbalanced dataset. Among the various models evaluated for predicting missed medical appointments, the Ridge Classifier stands out with the highest recall score. A high recall indicates the Ridge Classifier's proficiency in correctly identifying potential no-show patients, which is a crucial aspect of healthcare operations.

The confusion matrix is a valuable tool to gain a deeper understanding of a model's performance, particularly in binary classification tasks such as predicting missed medical

Table 5.2: Results for each models

Model	Precision	Recall	Roc_Auc	Accuracy
Ridge Classifier	0.86	0.69	0.62	0.66
Logistic Regression	0.86	0.67	0.65	0.65
Voting Classifier	0.87	0.62	0.65	0.62
Naïve Bayes	0.86	0.61	0.63	0.61
KNN	0.88	0.6	0.68	0.61
Bagging	0.9	0.53	0.7	0.58
Random Forest	0.92	0.52	0.73	0.58
Decision Tree	0.91	0.52	0.7	0.58
Extra Tree	0.9	0.52	0.66	0.57
Gradient Boost	0.93	0.51	0.74	0.58
AdaBoost	0.93	0.5	0.73	0.57
XGBoost	0.94	0.49	0.74	0.56

appointments. Let's analyze the confusion matrix for the Ridge Classifier to glean insights into its strengths and limitations:

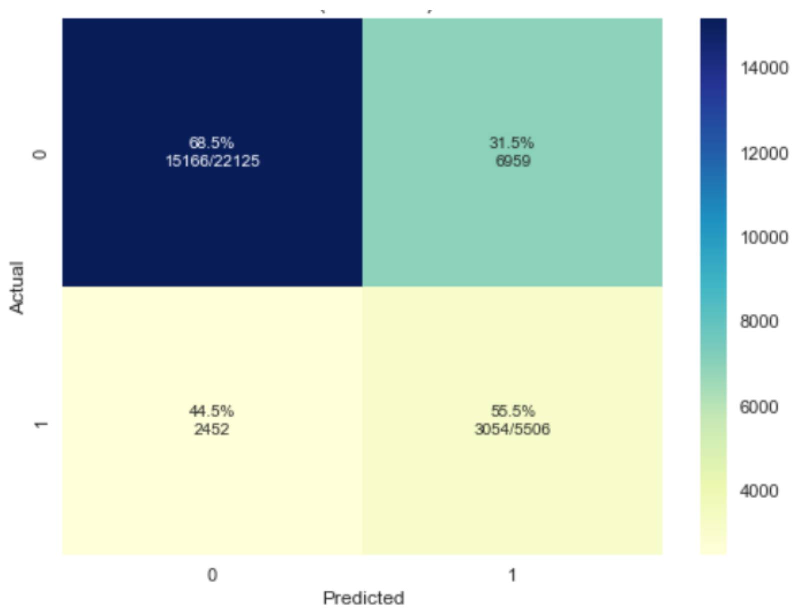

Figure 5.4: Ridge Classifier Confusion Matrix

Here 0 represents the show and 1 represents the no-show. For healthcare operations, where identifying potential no-show patients is crucial for efficient resource allocation and patient-centric care, a high recall score is of paramount importance. The Ridge Classifier proficiency in capturing true no-show instances signifies its effectiveness in addressing the challenges posed by imbalanced datasets and the consequences of false negatives. We have ensembled the top five models that is Logistic Regression, Ridge Classifier, Naive Bayes, KNN, Bagging by taking there votes using a Voting Classifier. So from the above results we can select the models. The selected models is Ridge Classifier because it has the highest recall with very less decrease in precision and also it has good Roc_Auc and accuracy socres.

6

Conclusions and Future Scope

This chapter of the report is dedicated to the conclusion and future scope. This section presents a thorough summary of the principal discoveries and understandings attained through the research conducted in the preceding sections.

6.1 Conclusions

Various past research or studies have various characteristics, advantages, disadvantages, and, most importantly, future scope. We have proposed using Ensemble Learning the medical appointment no-shows. The key contributions and objectives that have been achieved are: Tuning the hyperparameters of the models such that we can get more precision and recall, also we have done exploratory data analysis. In conclusion, all thesis objectives were achieved. According to the data attained from this project thus far, these graphs and results allow us to evaluate the performance of the Ensemble Learning in a Medical Appointment no-shows. The performance is shown that Ridge Classifier outperformed the other models, followed by Logistic Regression.

6.2 Future Scope

In light of this potential, a future research direction for our proposed work could be pursued is modification of hyperparameters to get good accuracy as well with increase in precision and recall. Furthermore, in recent years, the demand for Health care is going to increase. Thus, more efficient models needs to be trained. The landscape of predicting medical appointment no-shows holds immense potential for growth and innovation. Advancements in technology, data availability, and the growing recognition of patient-centered care will continue to shape the direction of research in this field. The future lies in developing sophisticated, adaptable, and ethically sound predictive models that seamlessly integrate into healthcare operations, ultimately improving patient outcomes and healthcare system efficiency.

Bibliography

[1] F. L. C. O. S. H. Leila F. Dantas, Julia L. Fleck, “No-shows in appointment scheduling – a systematic literature review,” 2018. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0168851018300459

[2] H. R. Tugba Cayirli, Emre Veral, “Designing appointment scheduling systems for ambulatory care services,” 2006. [Online]. Available: https://link.springer.com/article/10.1007/s10729-006-6279-5

[3] M. Samorani and L. G. Blount, “Machine learning and medical appointment scheduling: Creating and perpetuating inequalities in access to health care,” 2020. [Online]. Available: https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2020.305570

[4] T. Molfenter, “Reducing appointment no-shows: Going from theory to practice,” 2013. [Online]. Available: https://www.nature.com/articles/s41746-022-00594-w#Sec10

[5] H. W. Iman Mohammadi and B. N. Doebbeling, “Data analytics and modeling for appointment no-show in community health centers,” 2018. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/2150132718811692

[6] A. A.-R. . Sara Alshaya, Andrew McCarren, “Predicting no-show medical appointments using machine learning,” 2019. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-36365-9_18#author-information

[7] M. A. Christos Zacharias, “Joint panel sizing and appointment scheduling in outpatient care,” 2016. [Online]. Available: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2016.2532

[8] A. C. W. K. F. N. M. C. J. B. K. Amay Parikh, Kunal Gupta, “The effectiveness of outpatient appointment reminder systems in reducing no-show rates,” 2010. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0002934310001087

[9] L. G. B. H. L. M. A. S. Michele Samorani, Shannon L. Harris, “Overbooked and overlooked: Machine learning and racial bias in medical appointment scheduling,” 2021. [Online]. Available: https://pubsonline.informs.org/doi/full/10.1287/msom.2021.0999

[10] E. S. K. C. O. S. G. W. . M. S. Dianbo Liu, Won-Yong Shin, “Machine learning approaches to predicting no-shows in pediatric medical appointment,” 2022. [Online]. Available: https://www.nature.com/articles/s41746-022-00594-w#Sec10

[11] G. H. A. W. Eduard Incze, Penny Holborn, “Using machine learning tools to investigate factors associated with trends in ‘no-shows’ in outpatient appointments,” 2022. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1353829220318906

[12] H. T. H. V.-A. A. A.-H. S. E. Mohammad Reza Mazaheri Habibi, Fahimeh Mohammadabadi, "Effect of an online appointment scheduling system on evaluation metrics of outpatient scheduling system: a before-after multicenterstudy," 2019. [Online]. Available: https://link.springer.com/article/10.1007/s10916-019-1383-5

[13] J. L. B. J. L. E. S. Peng Zhao, Illhoi Yoo, "Web-based medical appointment systems: A systematic review," 2017. [Online]. Available: https://www.jmir.org/2017/4/e134/

[14] A. M. R. Q. L. Luiz Henrique Américo Salazar, Wemerson Delcio Parreira, "No-show in medical appointments with machine learning techniques: A systematic literature review," 2012. [Online]. Available: https://www.mdpi.com/2078-2489/13/11/507

[15] V. R. . H. R. J. Dunstan, F. Villena, "Predicting no-show appointments in a pediatric hospital in chile using machine learning," 2023. [Online]. Available: https://link.springer.com/article/10.1007/s10729-022-09626-z

[16] J.-Y. Huang, "An outpatient appointment scheduling system based on the logistic regression method," 2015. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/02522667.2014.1001607

[17] S. S. M.D.Anto Praveena, J.Sai Krupa, "Statistical analysis of medical appointments using decision tree," 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8918766

[18] L. H. S. A. M. R. F. R. D. J. Raduenz, "Prediction of attendance at medical appointments based on machine learning," 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9140973?casa_token=bvHa7IaHDhQAAAAA:v58MhUMubu1XeU6AfknriY_Dn4i1lmwTH4-55UDXmxYdMDNGCrN0fYOml397o4eN960hOF_WhLs

[19] R. A. Abdulwahhab Alshammari and R. Alshammari, "Developing a predictive model of predicting appointment no-show by using machine learning algorithms," 2021. [Online]. Available: http://www.jait.us/uploadfile/2021/0719/20210719053619657.pdf

[20] Z. Zhang, "Introduction to machine learning: k-nearest neighbors," 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4916348/

[21] M. A. David L. Olson, "Naïve bayes models in healthcare," 2023. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-28113-6_12