Ensemble Learning in Predicting Medical Appointment No-shows

Harsh Yadav (2020IMG-026)

Supervisor: Dr. Narinder Singh Punn

ABV-Indian Institute of Information Technology and Management Gwalior

Outline

Motivation

- We can help patients stay engaged by sending personal reminders and reaching out to people who might be at higher risk of missing appointments. This can help patients feel more connected to their healthcare providers and make it more likely that they'll come to their appointments.
- By figuring out when appointments are likely to be missed, hospitals can be smarter about how they use their staff, space, and time. This means less wasted resources and a smoother experience for everyone.
- If we can reduce the number of missed appointments, patients can get diagnosed and treated faster. This helps them get better more quickly and improves their overall health and well-being.

Research Gap

- Existing studies have mainly emphasized achieving high accuracy rates[1]. However, accuracy alone overlooks the varying impact of false positives and false negatives on healthcare operations and patient management.
- Prior studies have predominantly focused on individual predictive models[2]. Our research addresses a gap by utilizing the ensemble learning technique which leads to a more robust model.
- Less Emphasis was given on cross-validation and hyper-parameter tuning to increase the precision, recall, and roc_auc scores[3].

Research Objective

- 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 such that there is not much decrease in accuracy.
- Evaluate their performance using the metrics; precision, recall, accuracy, roc_auc.

Proposed Methodology

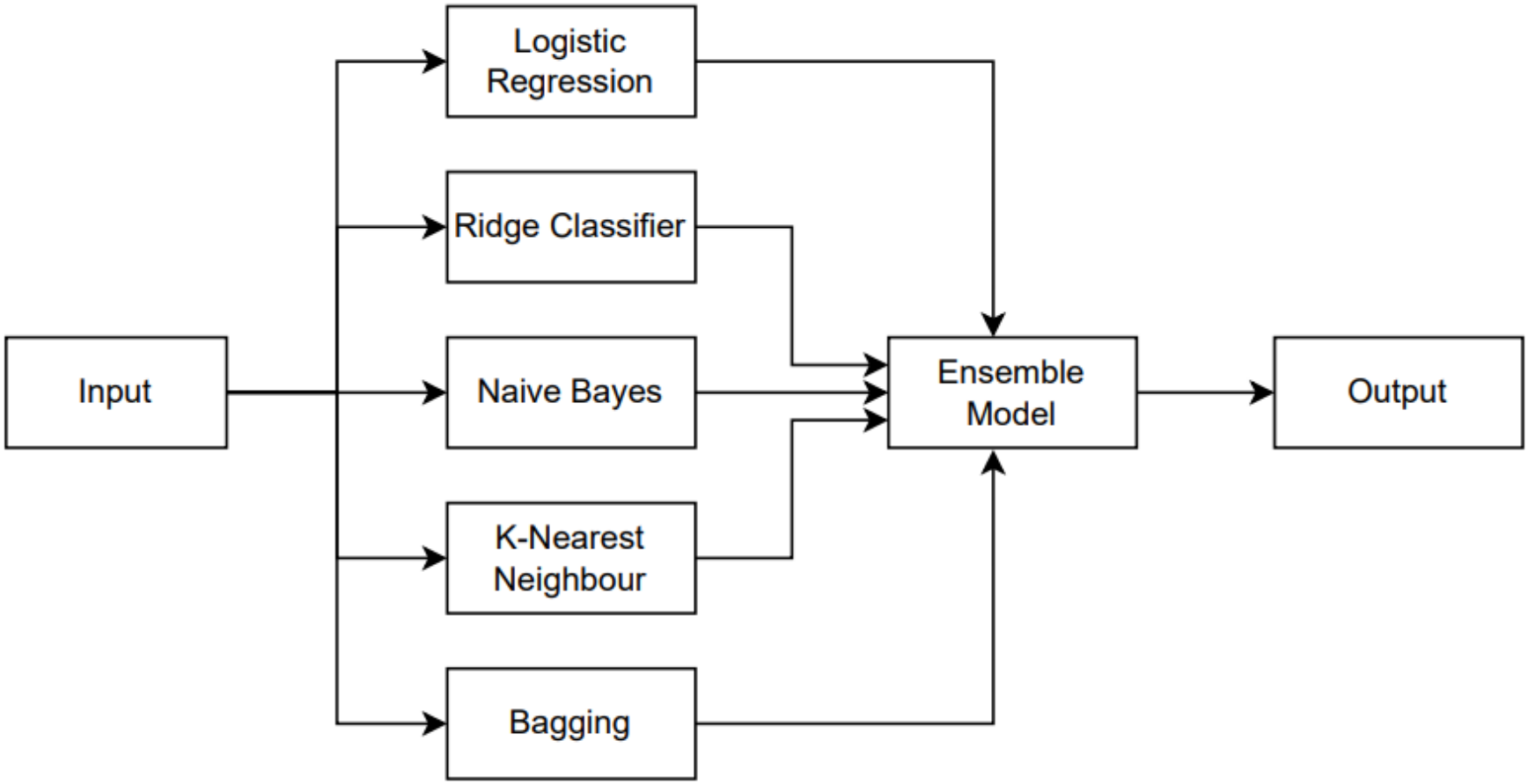

Figure 1: Proposed Methodology for predicting no-shows

Dataset

- We were given 14 features like Age, Gender, Alcoholism, Handicap, SMS received, no-show, etc.
- We added one more feature which is waiting days by taking the difference between the Appointment day and the scheduled day.
- No-show is our target. No-show '0' represents the patient showing up for the appointment and no-show '1' represents the patient didn't show up for the appointment.

Results

Table 1: 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
Naive 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

Confusion Matrix

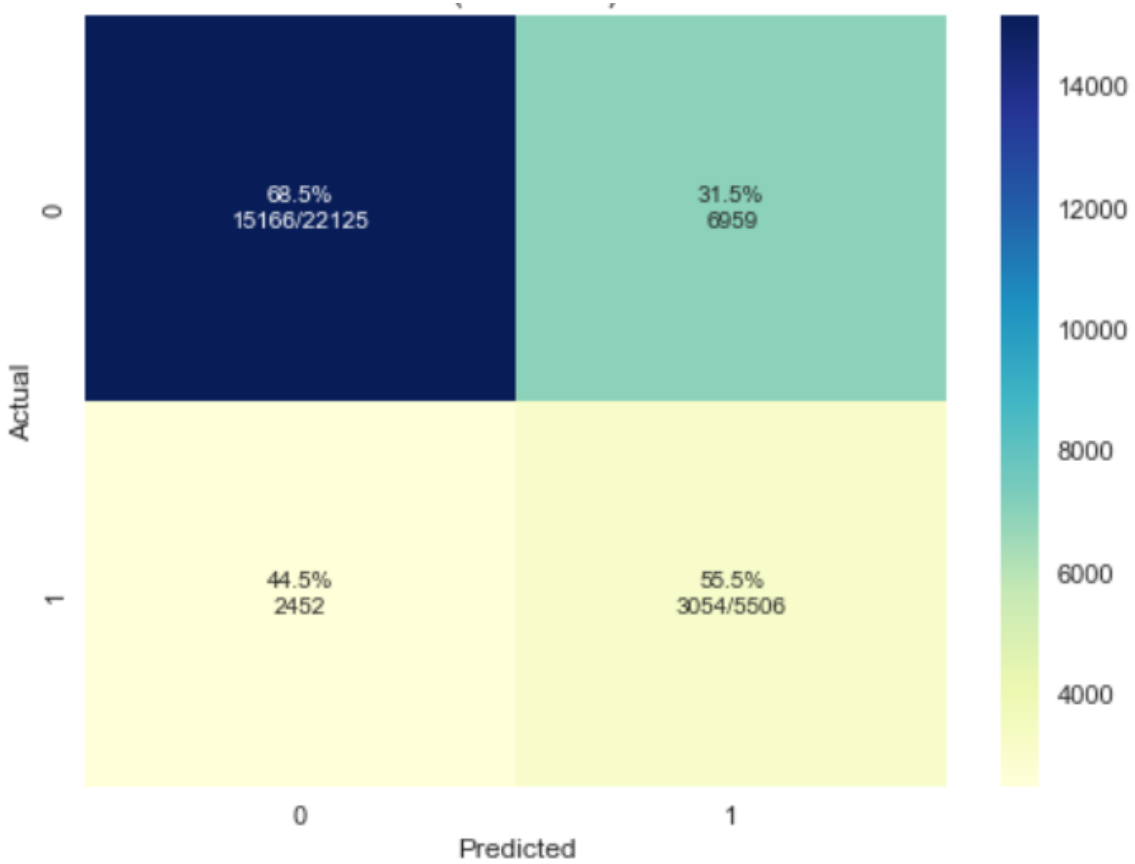

Figure 2: Ridge classifier confusion matrix

Exploratory Data Analysis

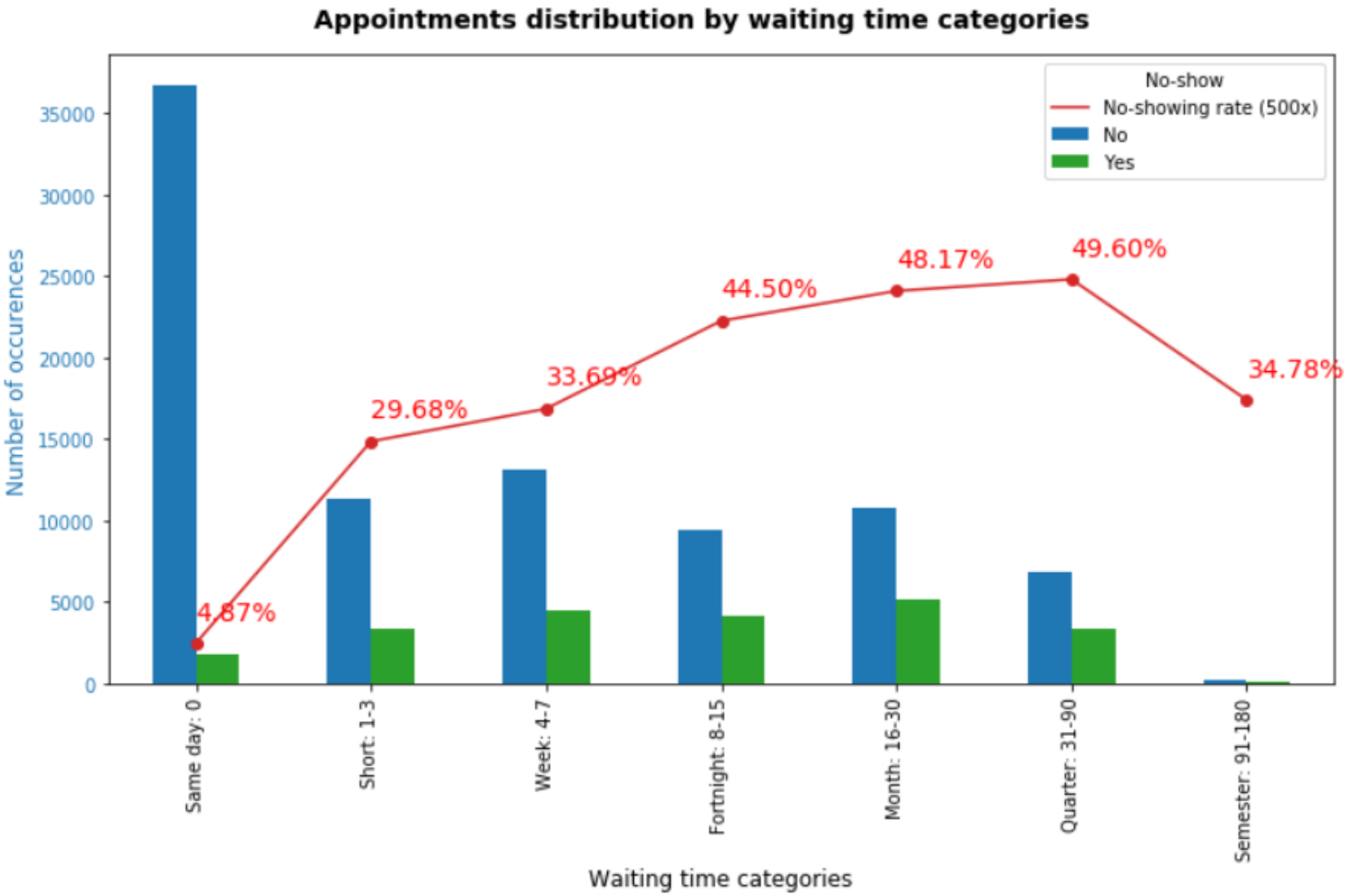

Figure 3: No-show rate Vs Waiting days

Conclusion

- We can conclude that waiting day is one of the most important feature.
- We were able to tune the hyper-parameters such that it increased the precision, recall, and roc_auc scores.
- Ridge Classifier is the model that has highest recall without much decrease in precision and accuracy.
- Overall, the research presents a novel solution to the problem of predicting no-shows.

Future Scope

- Investigate methods to optimize and fine-tune the Hyper-parameters to further improve recall and accuracy.
- Explore the techniques that can handle data imbalance in Medical Appointment no-shows[4].
- Explore techniques to make even more robust model which can help in making even more efficient scheduling and reminder systems[5].

References I

[1] A. Alshammari, "Developing a predictive model of predicting appointment no-show by using machine learning algorithms,"

[2] A. Moharram, "Data analytics and predictive modeling for appointments no-show at a tertiary care hospital,"

[3] J. S. R. Ramson, "Developing prediction model for hospital appointment no-shows using logistic regression,"

[4] C. R. . S. Yu, "A probabilistic model for predicting the probability of no-show in hospital appointments,"

[5] A. R. Teo, "No-show rates when phone appointment reminders are not directly delivered,"

Thank you!

E-mail: img_2020026@iiitm.ac.in