Readmission Predictions for Diabetes Patients

ELLIOT MACY, DAVIDA ROSENSTRAUCH

JANUARY 8, 2021

Overview, Business Problem

- Model that classifies whether a diabetes patient in a given hospital encounter is likely to be readmitted to the hospital within 30 days of discharge.

- This information will help insurance companies, hospitals, and medical professionals determine individual patients' risk of returning in the hospital in this time period.

Data

- 101,000 hospital encounters of approximately 71,000 individual patients, across 130 US hospitals, and recorded over the 10-year period between 1999 and 2008.

- Includes over 50 features representing patient and hospital outcomes.

- Dataset suffers from lack of class separation, presenting a challenge for generating predictions.

Methods

▶ Due to a class imbalance of approximately 90%-to-10% between target variables, undersampling was performed and models were conducted on a dataset of 22,714 hospital encounters.

▶ Three classification models were performed to determine best fit: a decision tree, random forest, and XGBoost. Accuracy, precision, recall, and F1 scores were analyzed to determine the best-fit model, with an emphasis on accuracy and recall in an effort to minimize false predictions of no readmittance.
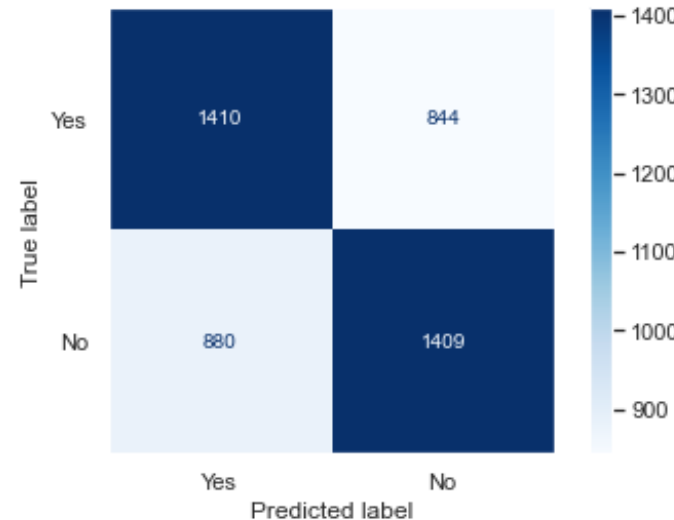
Results

▶ Of the three models, the XGBoost model most accurately predicted readmission, as determined via accuracy (62.36%) and recall (62.12%).
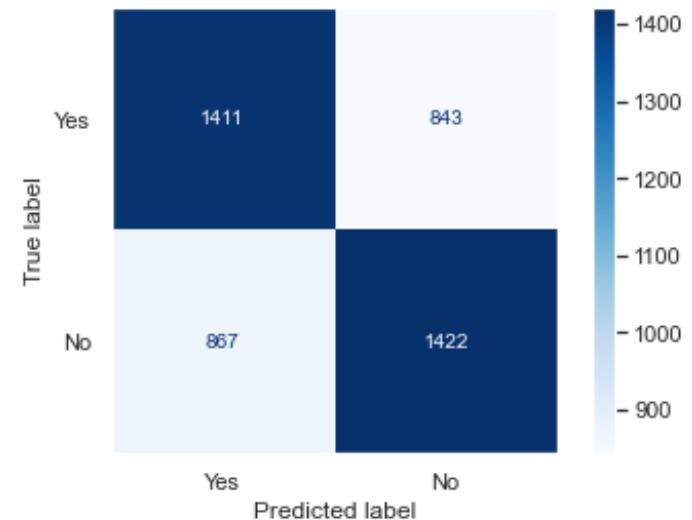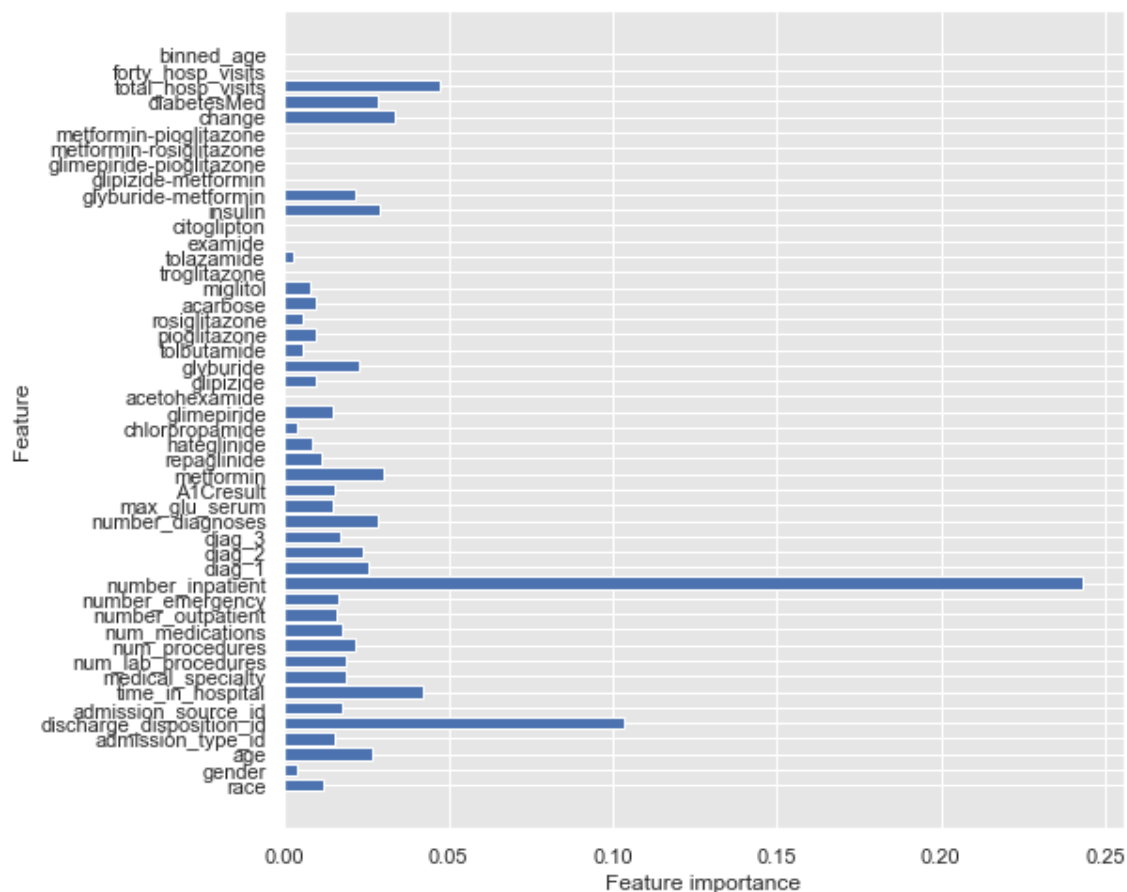
Decision Tree:
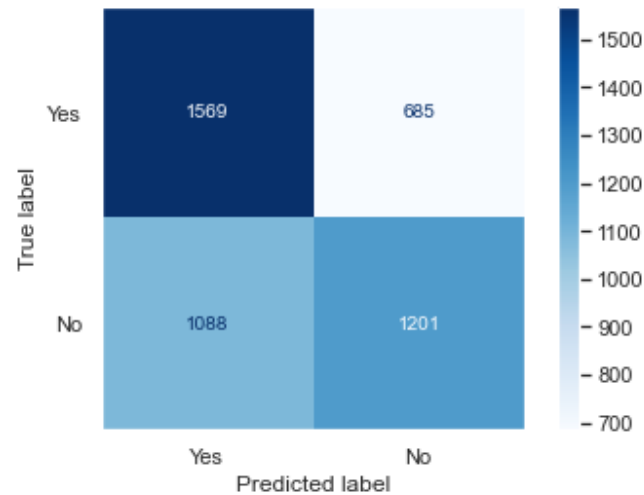
Random Forest:

XGBoost:

Results

▶ Of the features included in this most successful model, number_inpatient, discharge_disposition_id, total_hosp_visits, and time_in_hospital held the most predictive power.

Recommendations, Next Steps

▶ Based on these findings, we recommend insurance companies, hospitals, and medical professionals look to diabetes patients' number of inpatient admissions over the previous year, discharge disposition id, total number of hospital visits in the previous year, and number of days spent in the hospital in the observed encounter to predict likelihood of hospital readmission within 30 days.

▶ Next steps:

  ▶ Create a multi-class classification model to predict likelihood of hospital readmission at different intervals above 30 days.

  ▶ Incorporate patient weight data, which was omitted here due to lack of data.

  ▶ Conduct more recent research as modern medical innovation and technology may yield different results.

  ▶ Use neural networks to improve model accuracy and recall.

Summary

▶ An XGBoost model successfully predicted hospital readmission for diabetes patients with 62.36% accuracy and 62.12% recall for 22,714 hospital encounters recorded over 10 years.

▶ Of the features used to build this model, the following held the most predictive power:

  ▶ Number of inpatient admissions over the previous year

  ▶ Discharge disposition id

  ▶ Total number of hospital visits in the previous year

  ▶ Number of days spent in the hospital in the observed encounter

▶ We recommend insurance companies, hospitals, and medical professionals look to these features when predicting likelihood of readmission for diabetes patients.

Thank you!

- Any questions?