Predicting Readmission of Diabetes Patients

Kyle Hayes, Kaleb Nyquist

Background

- 30.3 million people (9.4% of US population) diabetic
- 4 in 10 diabetic adults describe health as “poor to fair”- 3 times the rate of non-diabetic adults.
- 25% of hospital admissions are diabetic
- Hospitalization costs in 2012: $124 billion

Can we predict repeat hospitalization?

Image Source: http://www.starmulticare.com/home-health-care-services/post-hospitalization/

Data

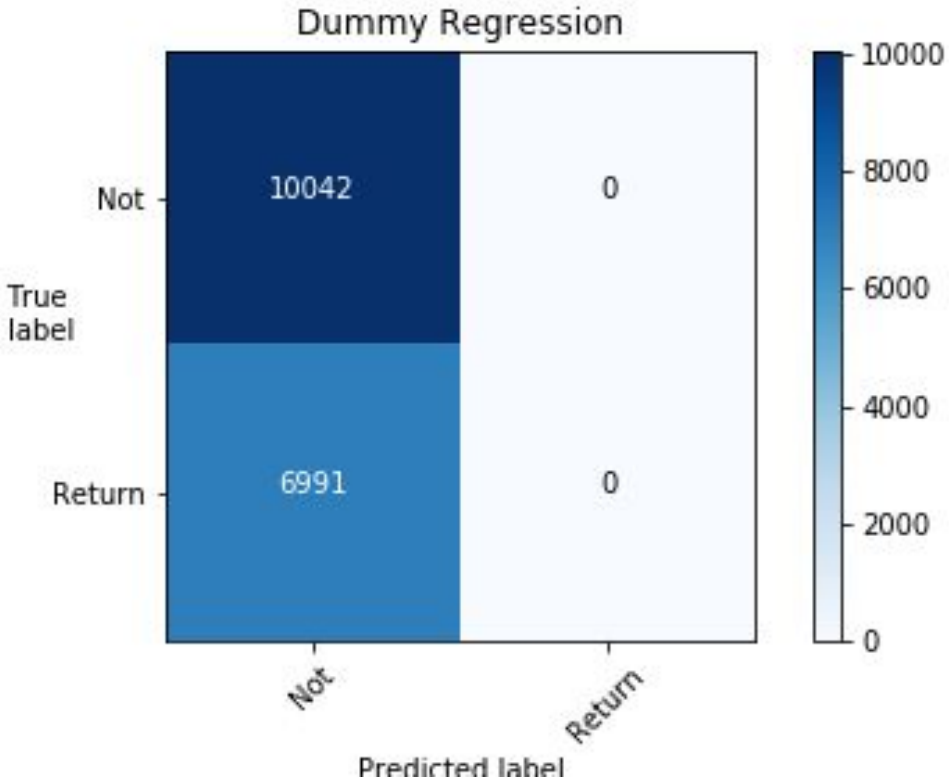
- Diabetes 130-US hospitals for years 1999-2008 Data Set
  (*https://archive.ics.uci.edu/ml/datasets/diabetes+130-us+hospitals+for+years+1999-2008)*
- Inpatient encounters, diabetic patients
- Length of stay: one to 14 days

Model Measurement:

- Measure the number of correct “readmission” predictions divided by total actual readmissions (Recall)

- False “no readmission” predictions are the most harmful

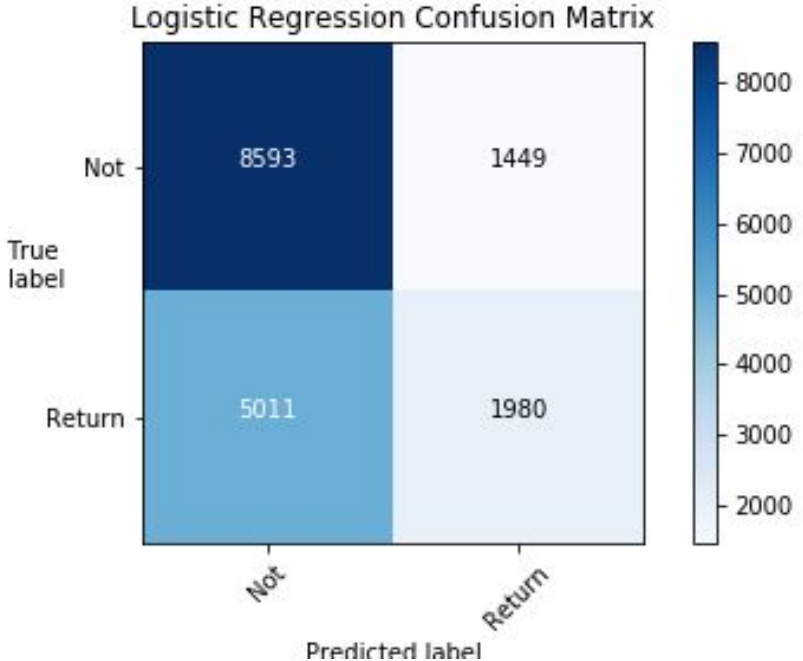
Dummy Model:



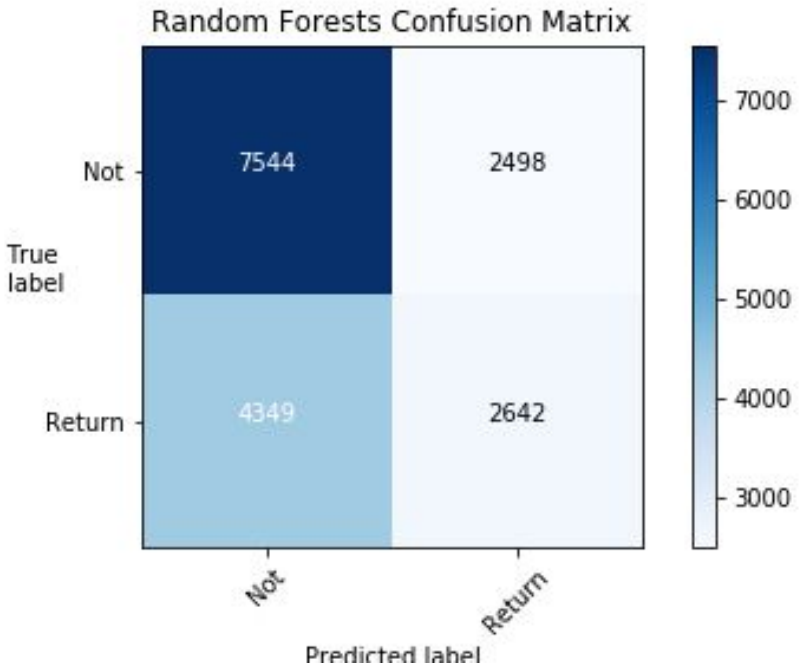
Recall: 0

Non-boosted models

Logistic Regression: 0.283

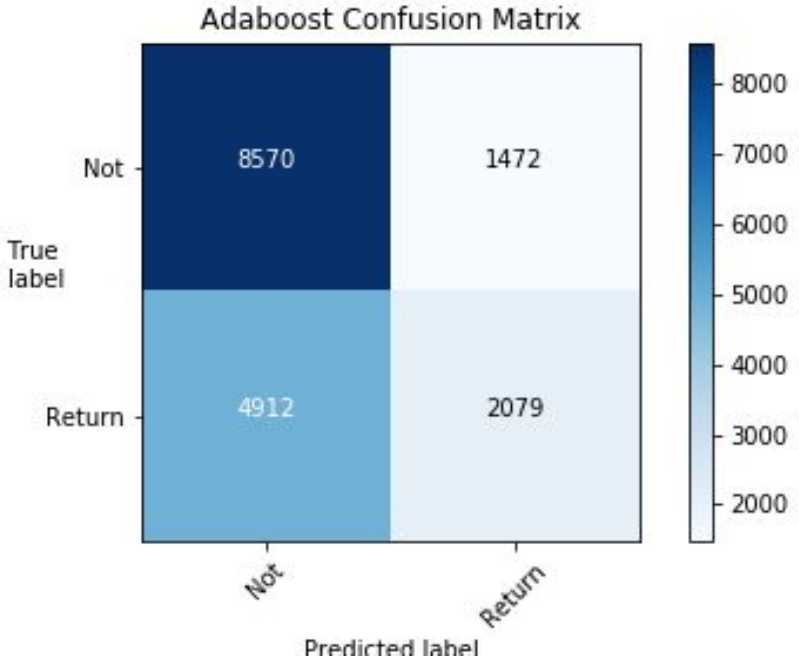


Random Forests: 0.378

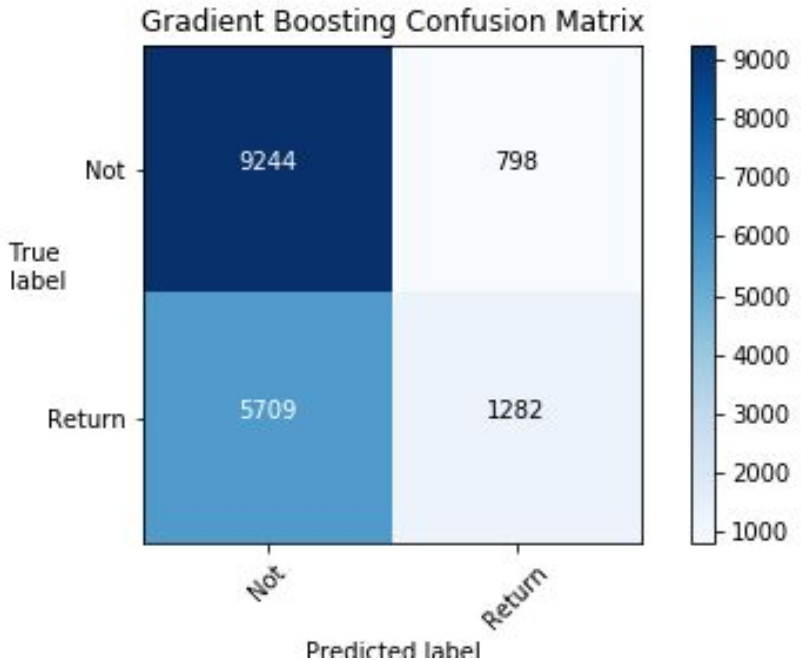


Boosted Models

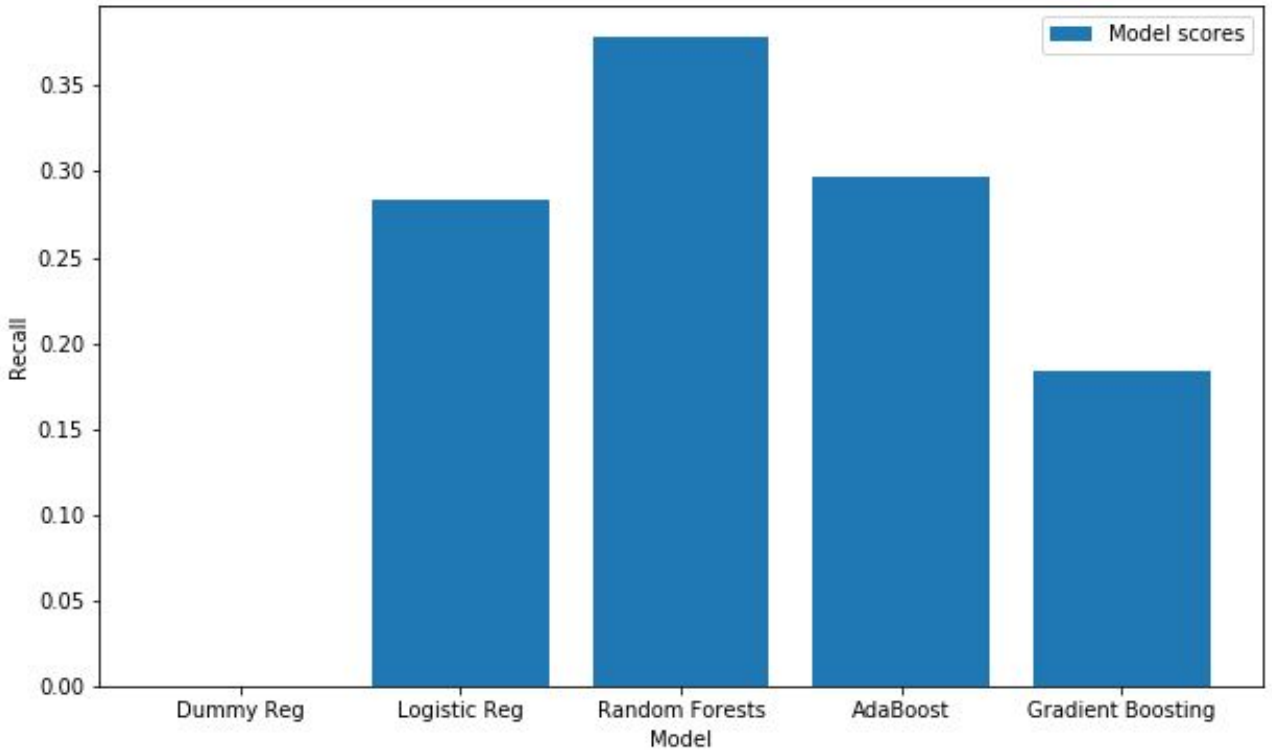
AdaBoost: 0.297



Gradient Boosting: 0.183



Model Comparison



Conclusion

- Random forests is best at predicting readmissions using predicted readmissions / total readmissions as measurement
- Prediction capabilities are low, model will need improvement

Next Steps and Recommendations

- Determine if we can study only subsamples with HbA1c and/or glucose tests
- Look at whether readmission is within 30 days
- Remember that a lack of readmission does not mean a lack of need.

Sources

- Diabetes Research Institute (https://www.diabetesresearch.org/diabetes-statistics)
- Hospital Readmission of Patients With Diabetes (https://link.springer.com/article/10.1007/s11892-018-0989-1)