USING MACHINE LEARNING TO REDUCE HOSPITAL READMISSIONS

By Aaron Bastian

OUTLINE

1. Introduction
2. Project Goals
3. Data
4. Methods and Models
5. Results
6. Value and Application
7. Conclusion

THE PROBLEM OF HOSPITAL READMISSIONS FOR DIABETES PATIENTS

- Hospital readmissions **within 30 days of discharge** are a common problem, particularly for patients with chronic conditions such as diabetes

- Hospital readmissions…
  - Negatively impact patient outcomes
  - **Increase healthcare costs**
  - Strain healthcare system capacity

- In response to this problem, Medicare established the Hospital Readmissions Reduction Program (HRRP), which financially penalizes hospitals with higher-than-expected rates of readmissions

A SOLUTION USING MACHINE LEARNING

- Machine learning models can **identify patients who are at risk** for 30-day readmissions
- **Key features and patterns** used by the model can inform treatment plans, reducing the likelihood of readmissions
- **Real-time** use of the model can **predict the probability** of a patient's readmission and enable providers to modify treatment plans dynamically

PROJECT GOALS

1. **Develop** a machine learning model to predict <30-day readmission based on patient treatment and discharge data
2. **Identify** key patient and treatment features that are most predictive of readmissions
3. **Provide recommendations** to help reduce the number of preventable readmissions

DATA

OVERVIEW

- Source
  - **University of California - Irvine, Machine Learning Repository**
    - "Diabetes 130-US hospitals for years 1999-2008"
- Represents **10 years** of clinical care at **130 US hospitals** and integrated delivery networks
- It includes **101,766 instances** with over **50 features** such as...
  - **Demographic** Information
  - **Health History** Information
  - **Admission** Information
  - **Treatment** Information
  - **Discharge** Information

DEFINING OUR TARGET VARIABLE

- The dataset contains three classes for readmission:
  - NO: Not readmitted
  - >30: Readmitted more than 30 days later
  - <30: Readmitted within 30 days

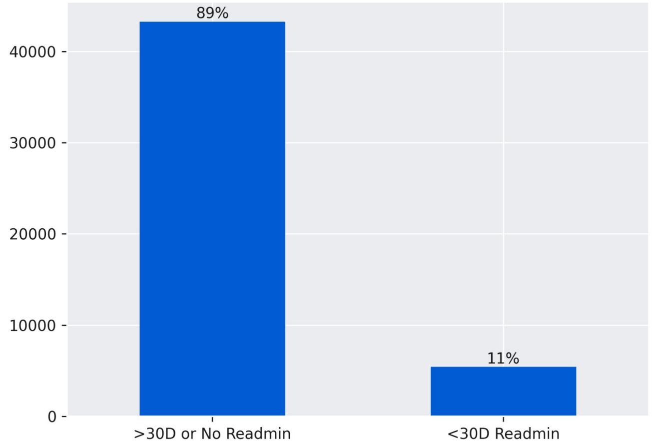
- To simplify the analysis, the target was transformed into a **binary classification**:
  - **Class 0**: Includes NO and >30 readmissions
  - **Class 1**: Includes <30 readmissions

PREPARING DATA FOR MODELING

The data was preprocessed by:

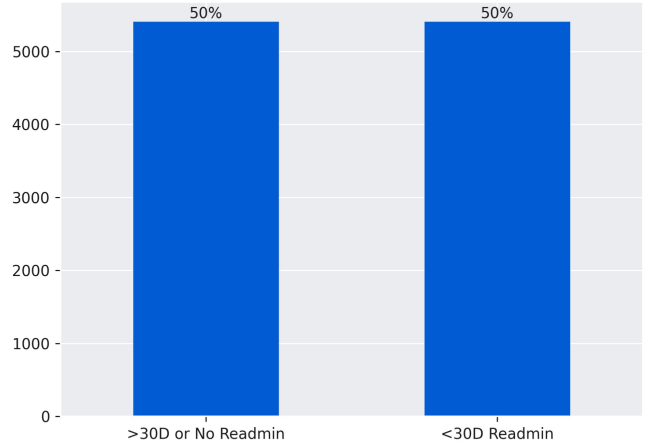
- De-noising
- Feature Engineering
- Undersampling to Address Class Imbalance

48,657 Instances



UNDERSAMPLING

10,816 Instances



METHODS AND MODELS

MODELS USED

SIMPLE

Random Forest

SVC (Support Vector Classifier)

LightGBM & **CatBoost** (Gradient Boosting)

Recurrent Neural Network (RNN)

COMPLEX

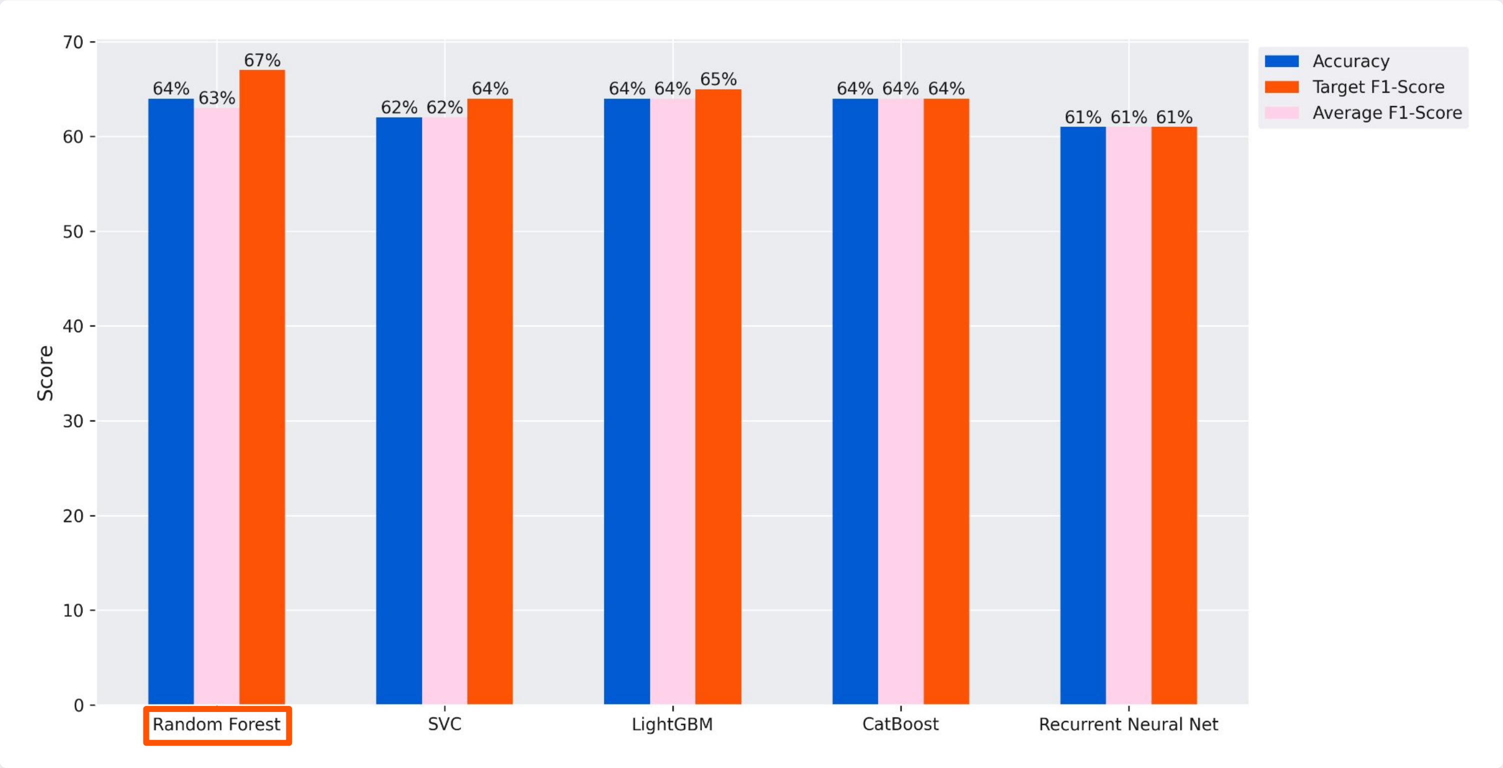
EVALUATION METHODS

- The data was split into **Train (80%)** and **Validation (20%)** sets
  - Models learn the train set
  - Models evaluated on validation set performance

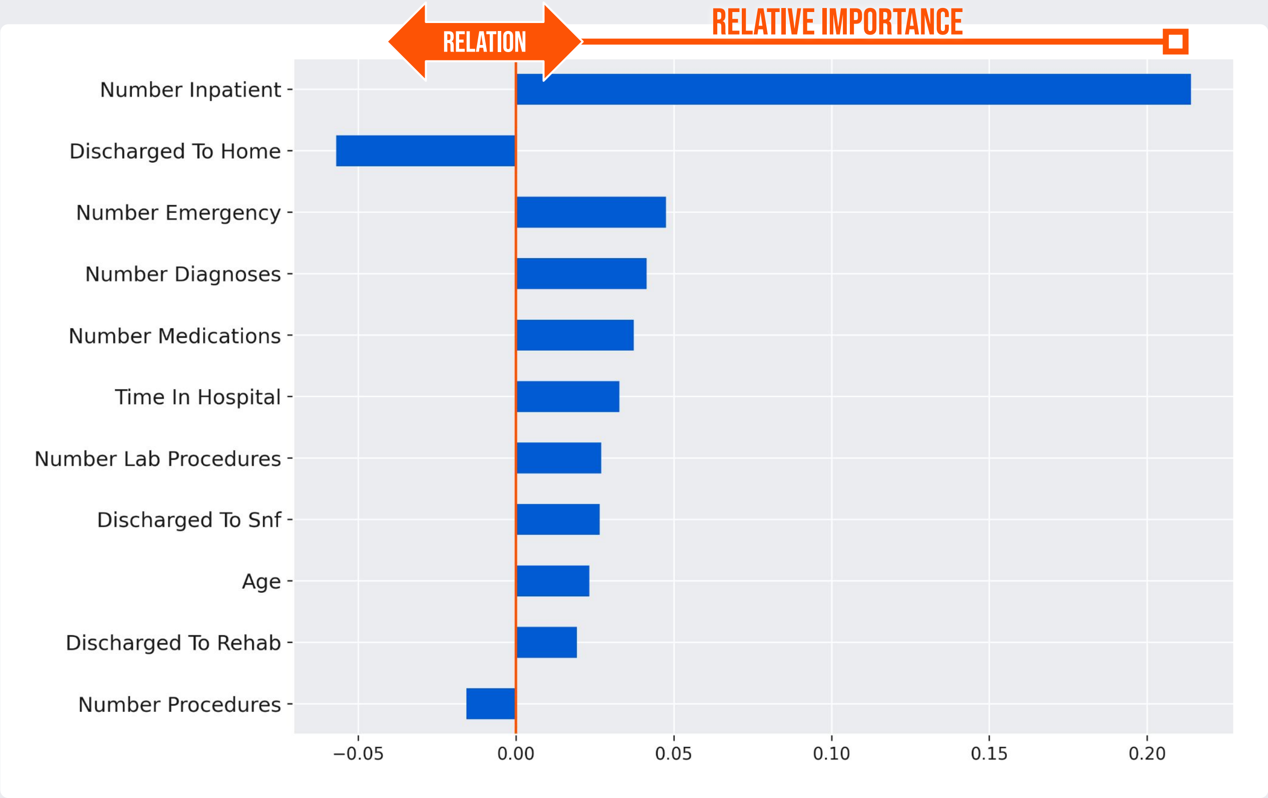
- Models were evaluated on their…
  - **Accuracy** - percent correct
  - **F-1 Score** - how accurate a machine learning model is at correctly identifying positive cases, while also minimizing errors

RESULTS

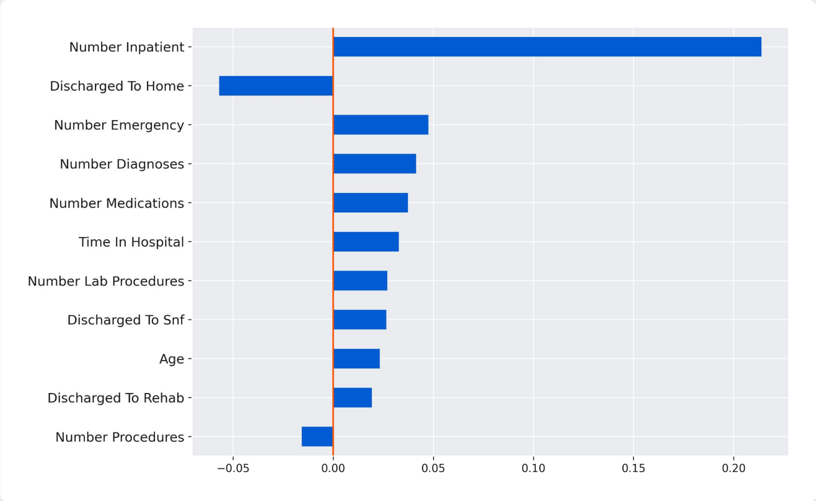
MODEL PERFORMANCE



IMPORTANT FEATURES



IMPORTANT FEATURES CONTINUED



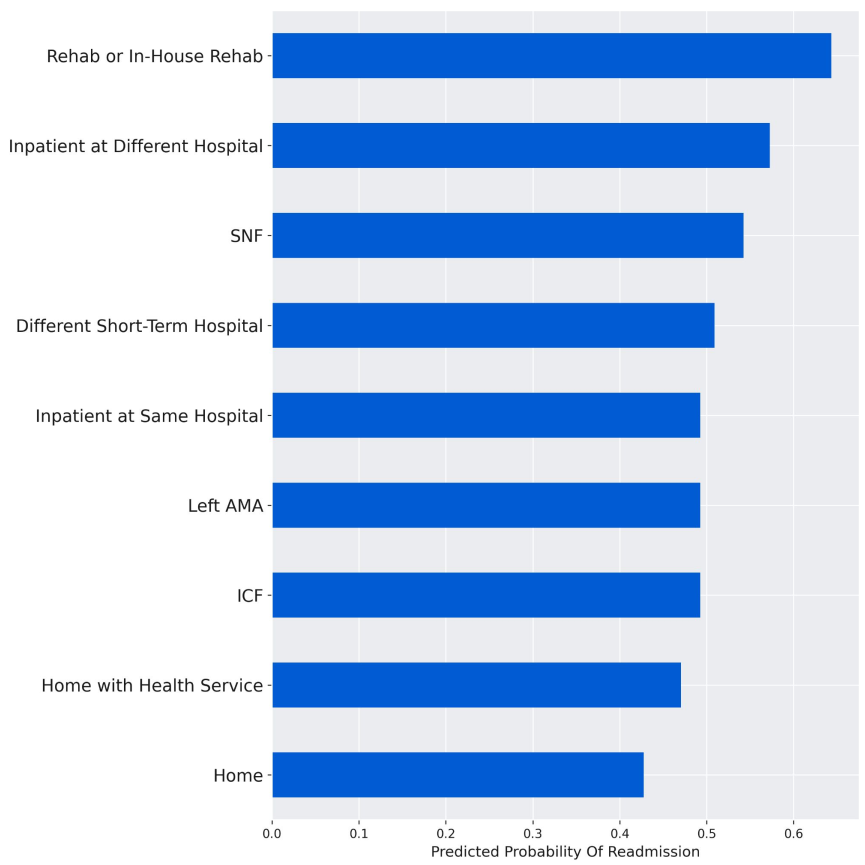
- **Correlation, not causation**

- Many of these are often **secondary to a patient's health status**, which may be the primary predictor of readmission risk
  - Hospital visits
  - Number of diagnoses
  - Number of medications
  - Discharge facility

- Interestingly, more procedures are associated with a lower risk of readmission

DISCHARGE FACILITIES BY READMISSION PROBABILITY

- **Rehab facilities** had the **highest** predicted readmission probability
- Discharge **home** had the **lowest** predicted readmission probability
- Transferring to inpatient care at the **same hospital** was associated with a **lower readmission probability** than transferring to a **different hospital**
- Again, these findings demonstrate **correlation, not causation**



VALUE AND APPLICATION

DEPLOYMENT AND APPLICATION

- The model can be integrated with **electronic health record (EHR)**
- Assess patients' risk of readmission in **real-time**
- Allow healthcare providers to intervene early and provide **targeted interventions** and **informed care plans**

BENEFITS OF A 30-DAY READMISSION PREDICTION MODEL

1. **Real-time** risk assessment
2. **Informed** care and discharge plans for high-risk patients
3. Optimized resource allocation
4. **Quality improvement (QI)** initiatives

CONCLUSION

- Hospital Readmissions…
    - Negatively impact patient outcomes
    - Decrease profits via Medicare's HRRP
    - Increase healthcare costs
    - Strain healthcare system capacity

- Our model can help providers…
    - **identify** high-risk patients
    - **improve** patient outcomes
    - **reduce** healthcare costs
    - **optimize** resource allocation

THANK YOU!

QUESTIONS?

Presenter Information:
Aaron Bastian, B.S.
P: (818) 282-3868
E: aaronbastian31@gmail.com