



VitalCare Health System

Reducing Readmission Risk in Diabetic Patients

Bekah McLaughlin





Contents

- ✓ Clinical Problem
- ✓ Dataset
- ✓ Data Analysis and Prep
- ✓ Model Development
- ✓ Results
- ✓ Limitations and Next Steps



Patient Safety Risk

Mitigating Risk for Readmission within 30 Days

- Why 30 Days is an Important Metric
- Risk for Patient
- Cost to Hospital

Diabetic Patient Focus

- Adult population
- Risk to this patient population
- Data driven interventions



The Dataset

Source of Data:
UC Machine Learning
Repository

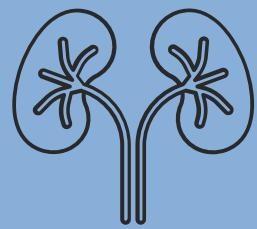
Clinical Data from 1999-
2008
Collected from 130 US
Hospitals
101,766 Rows
50 Features

Features:
Demographic Data
Clinical Categories
Hospitalization Details

Target:
Readmission

After Cleaning and
Prep:
70815 rows
25 Columns



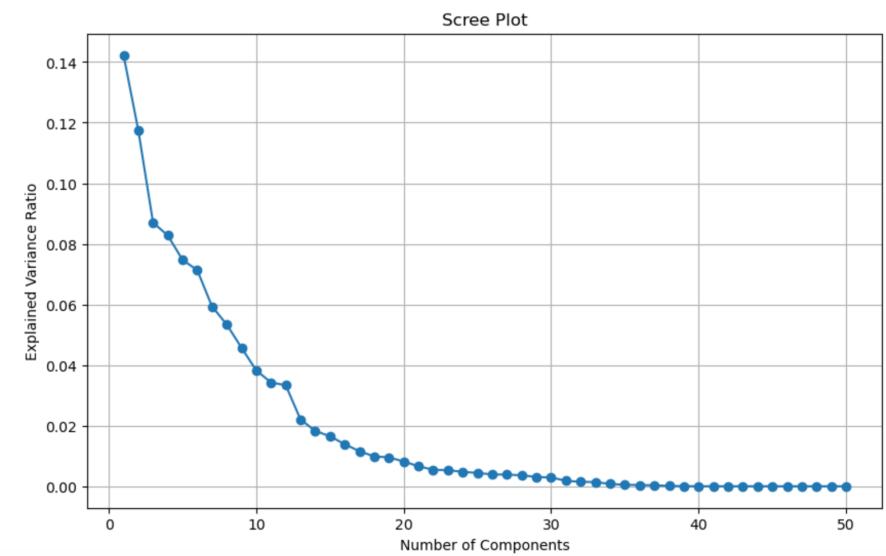
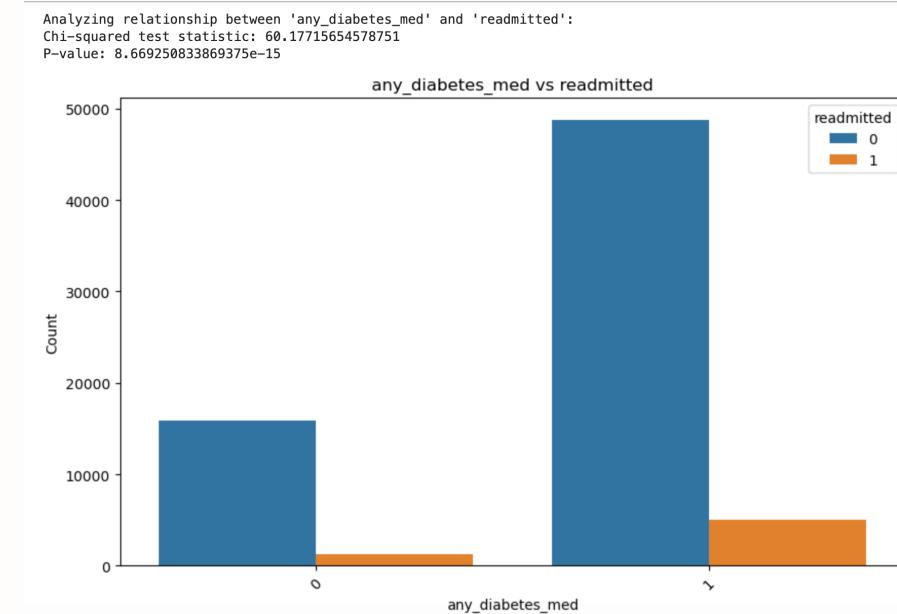


Exploratory Data Analysis and Preparation

Correlation Analysis

Categorical Analysis with
Chi-Squared and Visuals

Principal Component
Analysis



	precision	recall	f1-score	support
0	0.91	0.99	0.95	9684
1	0.19	0.02	0.03	939

Logistic Regression

Final Model Results

- Accuracy = 0.91
- Precision for Class 1 - 0.17
- This model struggled to effectively categorize the target class of interest.
- While the model performs reasonably well overall, it falls short of meeting the specific business requirements for accurately predicting early readmissions.



XG Boost

Final Model Results

- Accuracy = 0.63
- Precision for Class 1 - 0.13
- This model also struggled to effectively categorize the target class of interest.
- However, recall for class 1 is 0.55, indicating that the model captures 55% of all actual instances of class 1 in the dataset.





Random Forest

Final Model Results

- Accuracy = 0.61
- Precision for Class 1 - 0.12
- Random Forest did result in a model that did better than XG Boost, but marginally
- And according to my metric, precision it did worse.



Clinical Case

Results



All Models only about 60% accurate and all very poor at identifying the target



No occurrence of symptom in the past 3 months

Next Steps



Ideally, I would seek input from individuals with more experience and expertise to assess the suitability of this dataset for predictive modeling purposes.



If possible, I'd like to collect more current, balanced data and repeat this project.

Questions?



rebekah.l.mclaughlin@gmail.com



<https://www.linkedin.com/in/bekah-mclaughlin-8a86b02b4/>

