PREDICTION OF DIABETES READMISSIONS

Aisha Kamara, Data Scientist
October 4th, 2021

PRESENTATION OVERVIEW

01 **Introduction**

Brief background on Diabetes and readmission rates in the US.

02 **Problem Statement**

Problem statement

03 **EDA Findings**

EDA findings

04 **Modeling Results**

Models and results

05 **Conclusion**

Findings and recommendations

INTRODUCTION

3.3 Million

Readmitted Patients

$41 Billion

Hospital Costs

DIABETES

“Diabetes is a chronic (long-lasting) health condition that affects how your body turns food into energy.”- CDC, 2021

Dataset Overview

- The data obtained represents 100,000+ unique inpatient diabetes medical visits over 10 years (1999–2008) of clinical care at 130 hospitals and integrated delivery networks in the United States.

- It included 50 features such as a patient's age, gender, duration in hospital, number of lab tests performed during their medical visit, the amount of medication prescribed for the patient, etc.

PROBLEM STATEMENT

PROBLEM STATEMENT

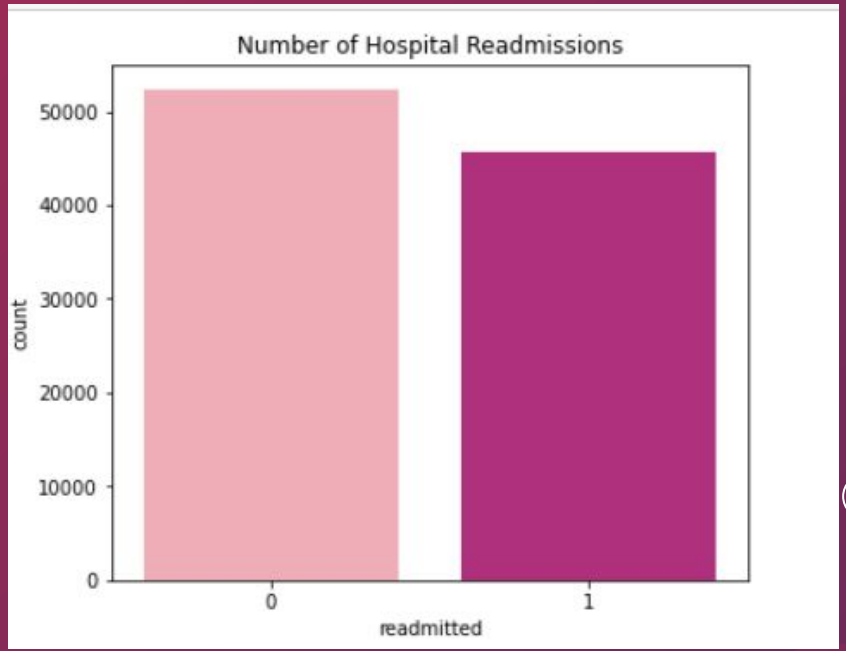
Problem

What factors are the strongest indicators of hospital readmission for a diabetic patient?

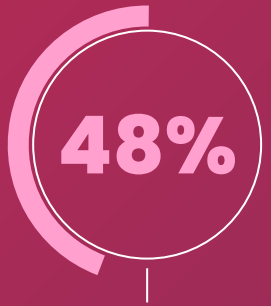
EDA

READMITTANCE RATES

- 98,052 Observations
- 47 Features
- 1 Target Feature
- 54% of patient had no record of readmission
- 46% of patients, were readmitted after discharge



DEMOGRAPHICS



Age

48 % of patients readmitted are between 60-80 years of age.

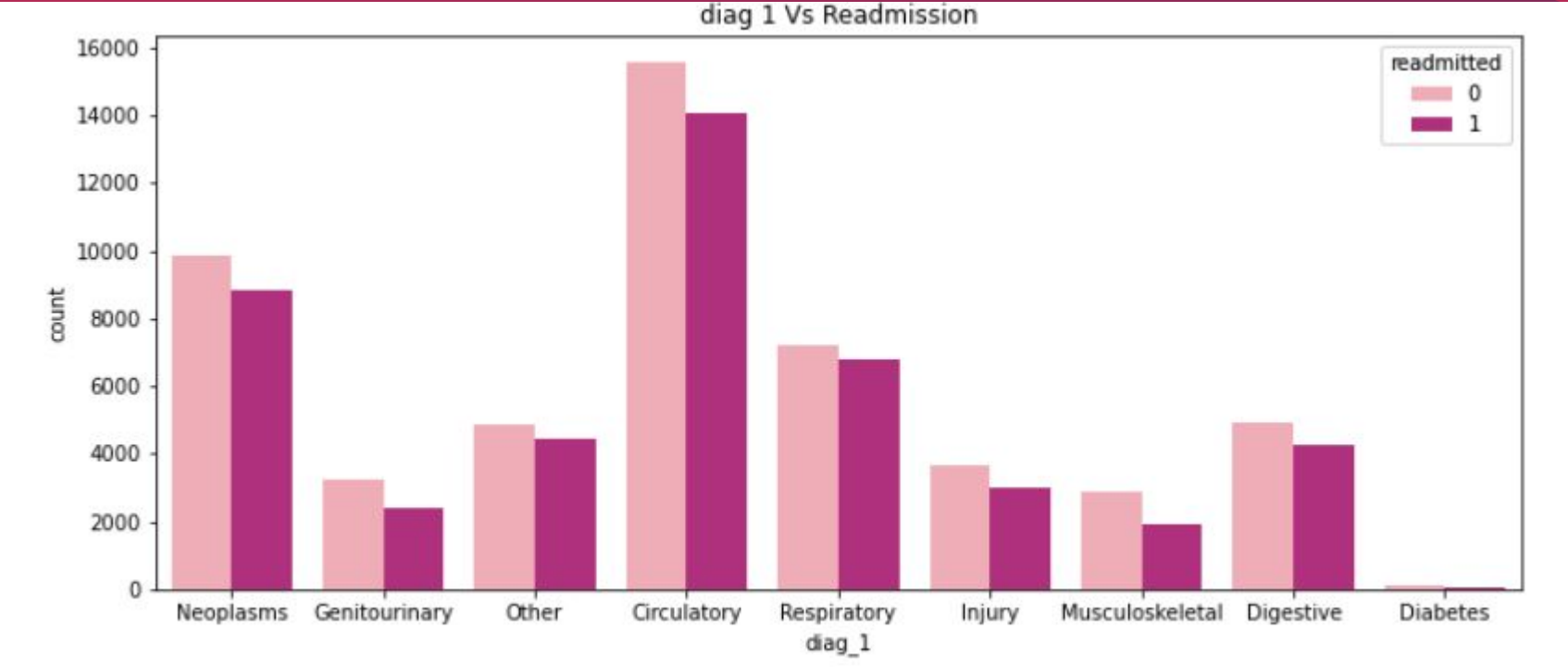
Gender

53 % of readmitted patients are females.

Race

76% of readmitted patients are Caucasian.

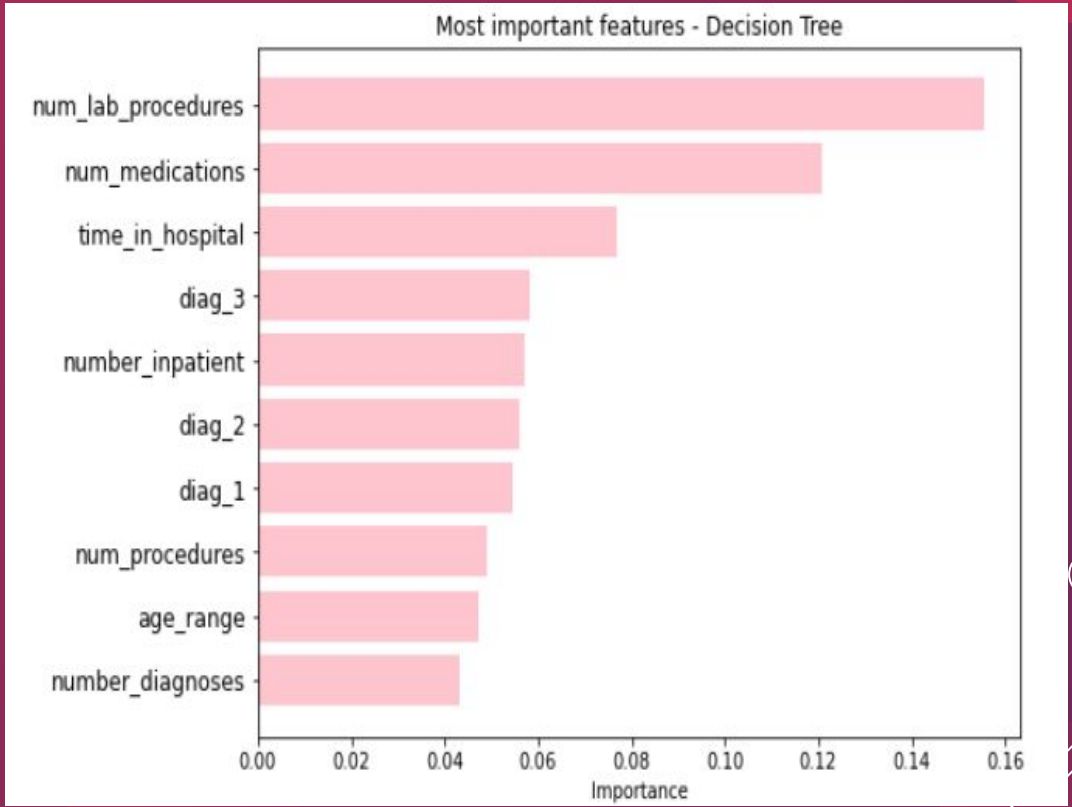
DIAGNOSIS TYPE



MODELING

MOST IMPORTANT FEATURES

- Decision Tree Classifier
- Most important feature option
- Limited to 10 features out of the 50 in the initial dataset
- Limited 3 classification models to just these 10 features.



GRID SEARCH FINE TUNING

- Grid search is a tuning technique that attempts to compute the optimum values of information fed into a model for training.

METRIC GOALS

Accuracy score

How best did we train our model to accurately predict our desired outcome. The higher the score the better.

Sensitivity Score

How "sensitive" is the classifier to detecting positive instances? We want to identify “high risk” patients most likely to be readmitted. The higher the score the better.

RESULTS

Important Features	Important Features	Grid Search	Grid Search
Accuracy Score	Sensitivity Score	Accuracy Score	Sensitivity Score
0.5947	0.4912	0.6180	0.5166

CONCLUSION

CONCLUSION

- Ten major features are found to have high impact on diabetes patient readmission.
- Although not the best scores, still beneficial for medical practitioners to pay attention to these features
- Using Grid search to fine tune for each of our classification models produced the best accuracy and sensitivity score.
- Our best model provided an accuracy score of 0.62 and a sensitivity score of 0.52.
- Attempt to use other models moving forward.
- Including new data such as family history may be helpful in increasing primary diagnosis rates and effectively decrease readmission rates.

THANKS

Do you have any questions?

akamara@xyz.com
+1 240 333 4456
xyzconsulting.com

CREDITS: This presentation template was created by **Slidesgo,** including icons by **Flaticon** and infographics & images by **Freepik**
Please keep this slide for attribution