Hospital Length of Stay: A Predictive Tool

Ironhack, Data Analytics 2023
Sandra Cunha

TABLE OF CONTENTS

Hospital Bed Shortage Problem

- **Impact on Patients and Healthcare Professionals**

  Patient safety compromise;
  Increase stress in Healthcare Professionals;

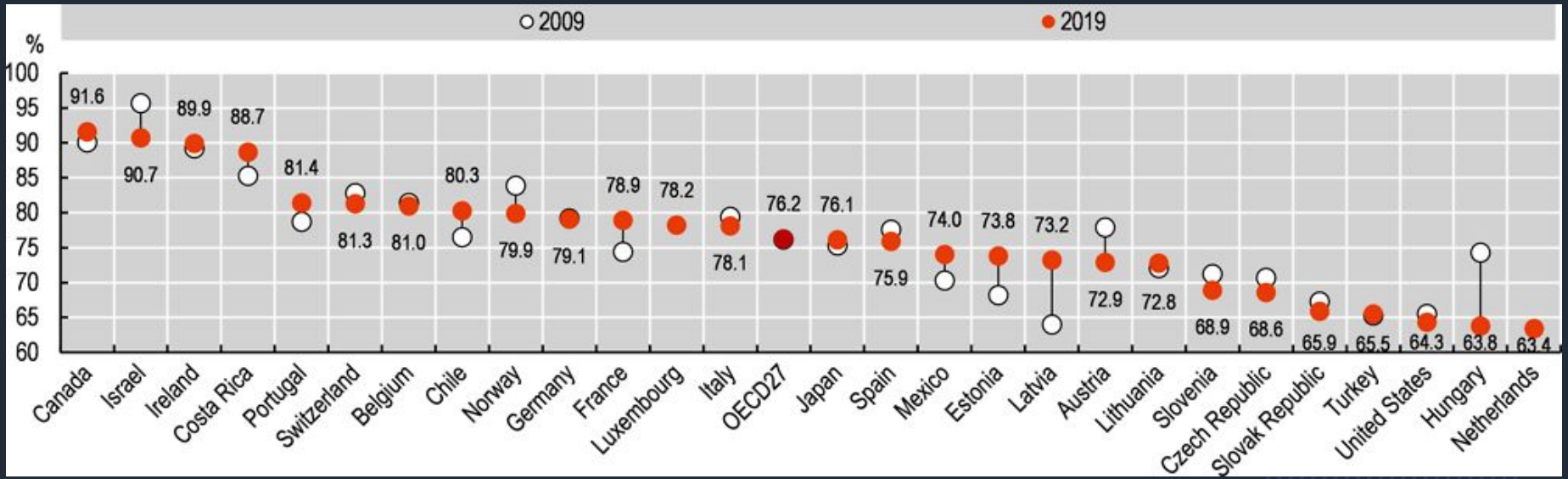
- **Inefficient Bed Management**

  Increase in hospital costs;
  Inadequate hospital resources utilization;

- **Prolonged Hospitalizations**

  High impact on overall bed occupancy;
  Often unnecessary (leading to an increase in complications on patients);

Occupancy Rate of Acute Care Beds, 2009 and 2019



Source: OECD Health Statistics 2021

Prediction of Hospital Length of Stay in General Surgical Patients

Capacity Planning

Anticipates bed requirements;
Improves patient flow;

Early Intervention

Target patients with high risk of prolonged stay;
Optimize treatment plan and discharge planning;

Financial Efficiency

Prevents surgery cancellations;
Better management of staff levels and medical equipment;

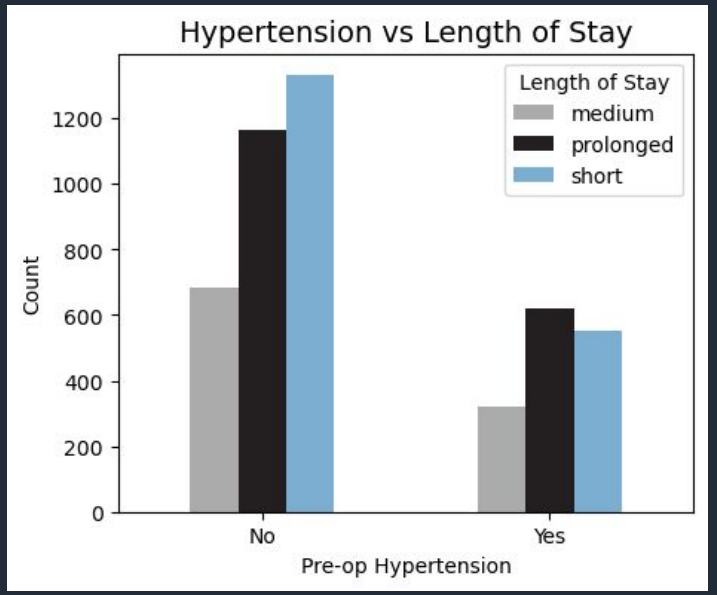
Medical efficiency

Patient safety improves;
Healthcare professionals more satisfied;

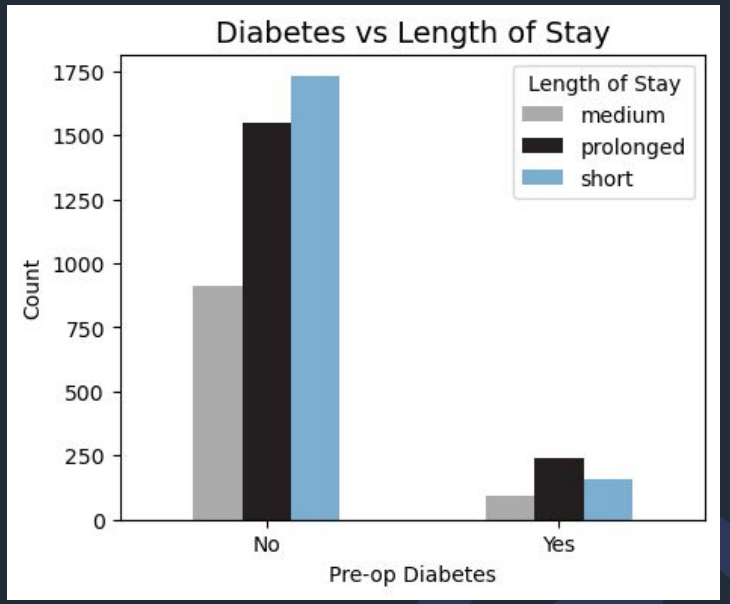
VitalDB

- Open Dataset containing surgical information of General, Thoracic, Urology and Gynaecologic patients;
- Seoul National University Hospital, Republic of Korea, 2022;
- 6,388 patients;
- 73 Clinical parameters and 34 Laboratory results parameters;

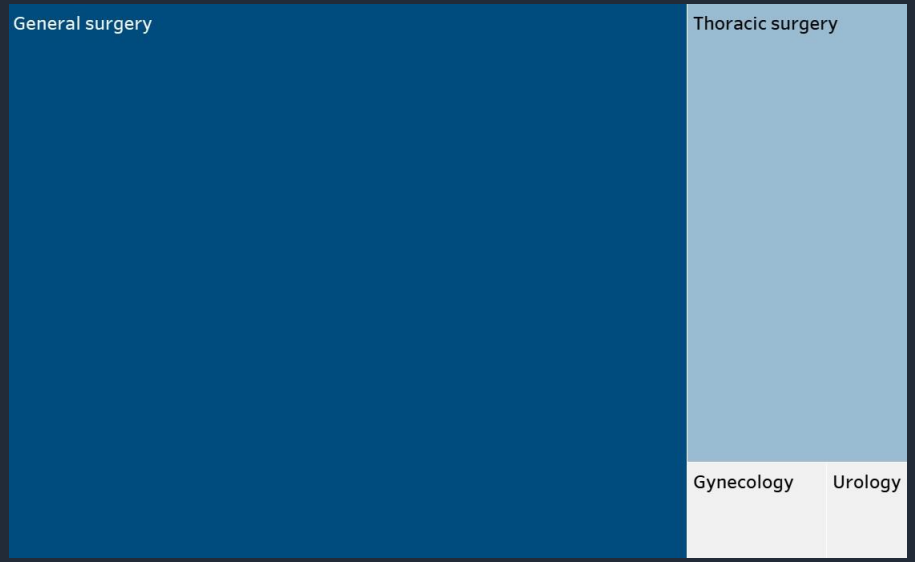
Factors Associated with a Prolonged Length of Stay in Surgical Patients

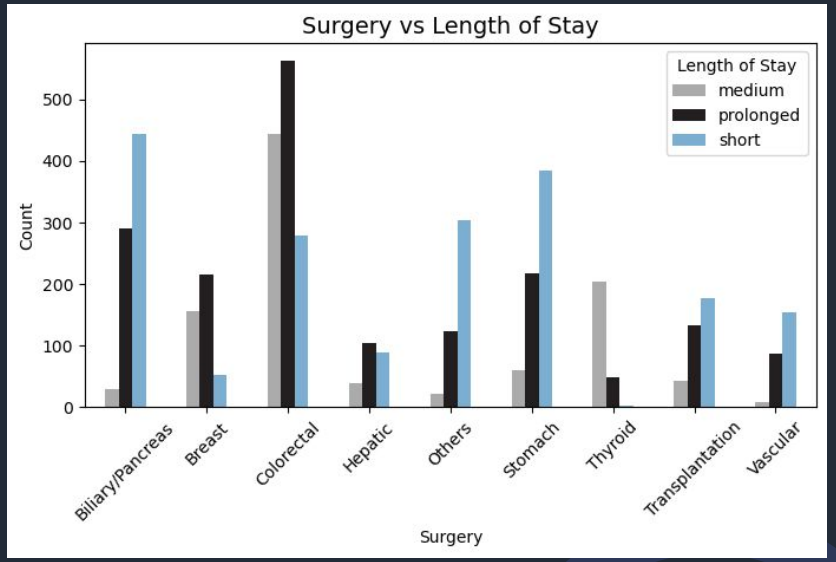


Pre-existing Medical Conditions



Factors Associated with a Prolonged Length of Stay in Surgical Patients

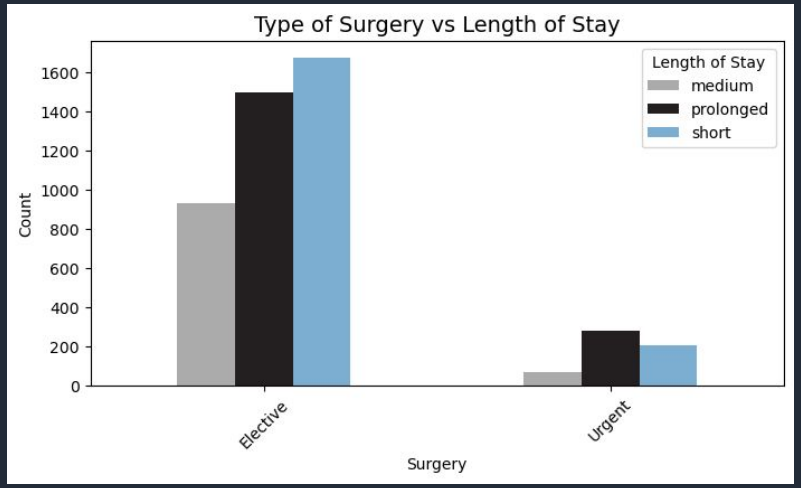


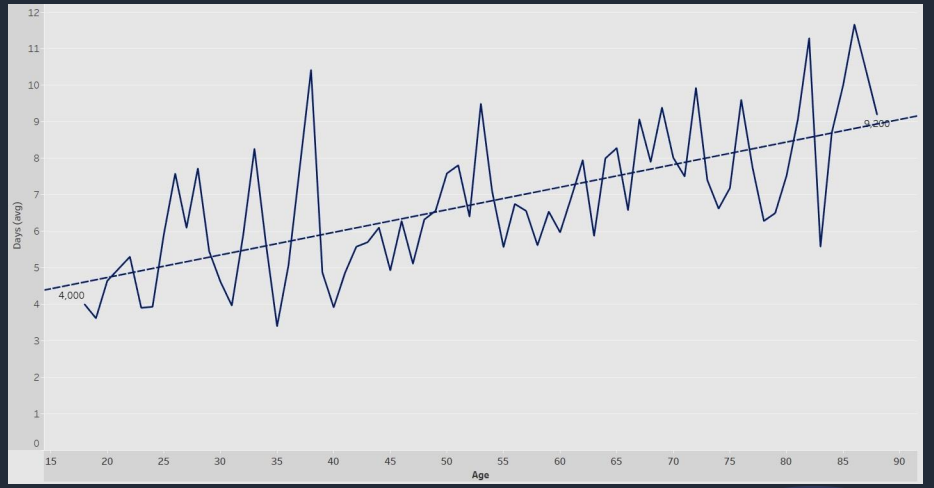


Surgical Department

Surgical Complexity

Factors Associated with a Prolonged Length of Stay in Surgical Patients





Type of Surgery

Age

Supervised Machine Learning Classification Problem

01

Load Data / Preprocessing

Data cleaning, EDA
Multiclass Classification Target
(Short, Medium, Prolonged)

02

ML Baseline Metrics

SVM Classifier
KNeighbors Classifier
Logistic Regression
Extreme Gradient Boosting
Random Forest Classifier

03

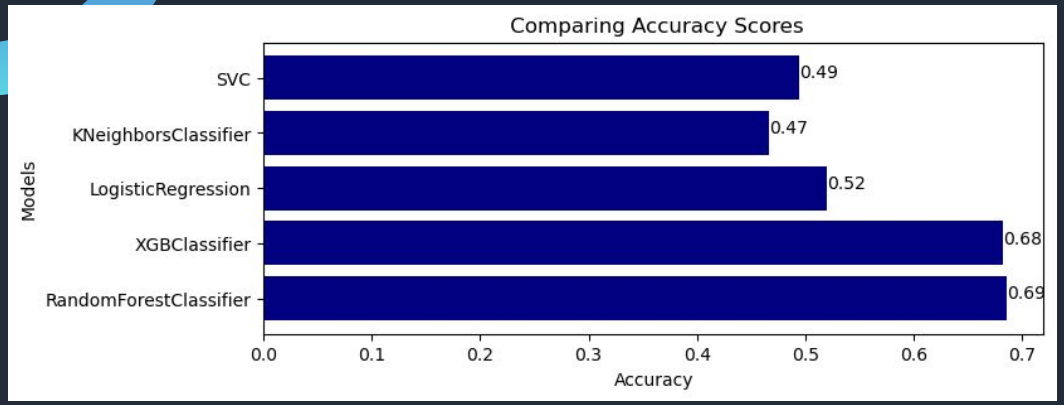
Development / Refinement

Multicollinearity
Outliers Handling
Data Transformation
Class Imbalance
Feature Selection

04

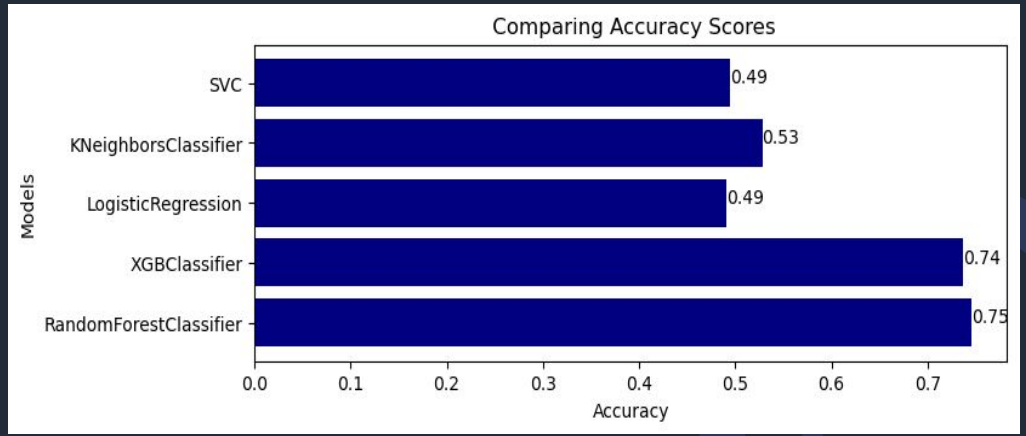
RF Optimization and Model Evaluation

Hyperparameter Tuning
K-fold Cross Validation
Accuracy
Confusion Matrix
ROC-AUC Curve



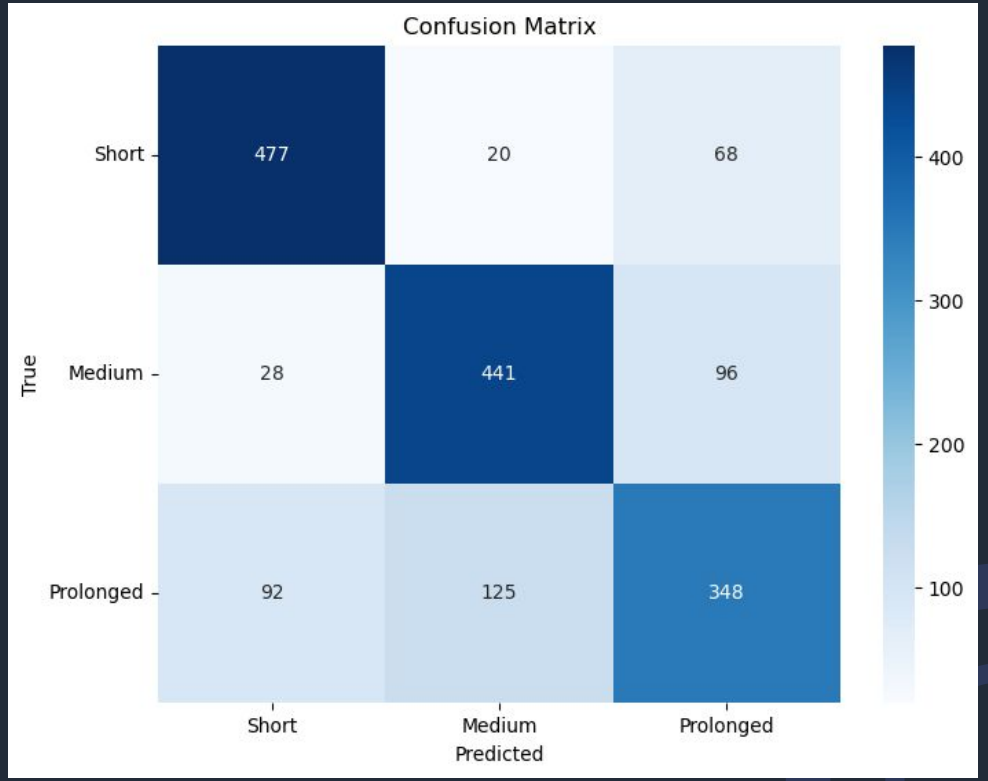
Baseline Metrics

Upsampling Minority Classes (Medium and Prolonged Stay)



Performance Metrics of Random Forest Classifier

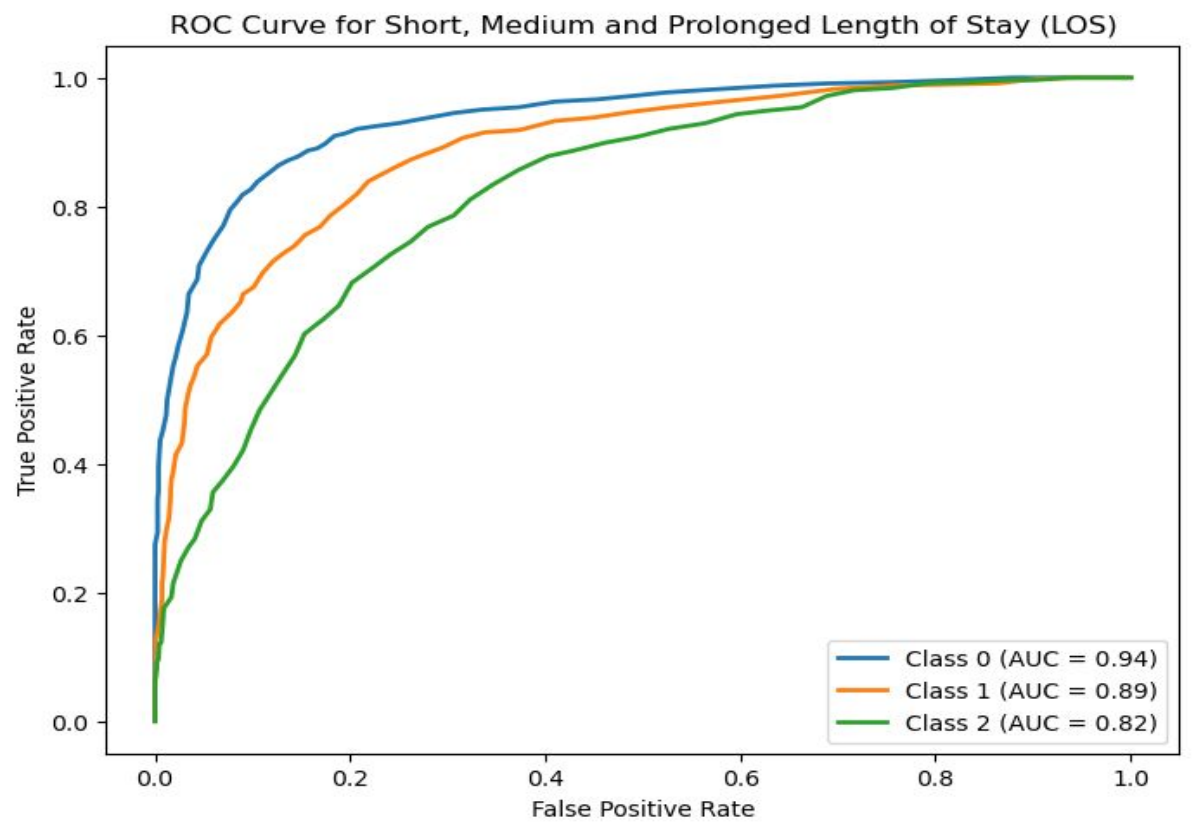
Overall Accuracy
74%



Performance Metrics in Predicting Different Hospital Length of Stay Classes

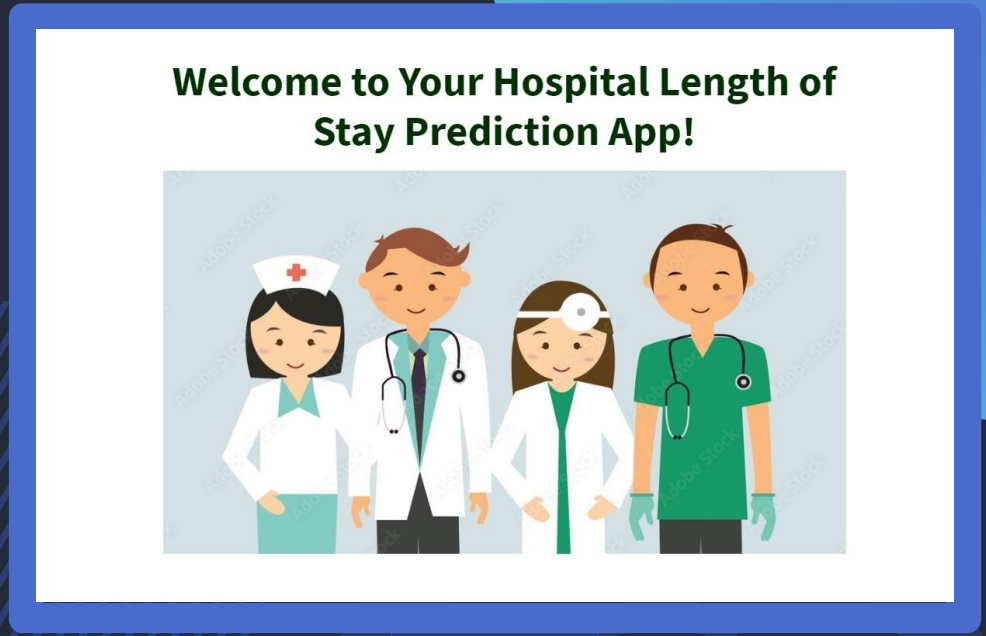
	Short	Medium	Prolonged
Precision	**80%**	75%	68%
Recall	84%	78%	62%
f1-score	82%	77%	65%

ROC Curve - AUC Score



Model Deployment

Surgical_Patient Application Demo



Conclusion

- This model is particularly effective in correctly predicting Short Length of Stay class, as evident from the higher precision;
- Applying the model to the real-world unseen patient data should be evaluated carefully, especially when predicting on "Medium" and "Prolonged" classes, since this model was trained on balanced target classes;
- Limited dataset from a single Hospital with records from a single ethnicity (Asian Patients) and lack of important features such as vital signs on patient admission, which could be important in predicting Length of Stay.
- If improved, it can be a reliable predictive tool that can assist hospital administrators, bed managers and healthcare staff in managing bed capacity effectively enhancing financial and medical efficiency;

Thank You!

Do you have any questions?

sandra_cris_cunha@hotmail.com

www.linkedin.com/in/sandra-cunha-a553a41b6