MEDICAL APPOINTMENT NO SHOWS

Presented by:
Alanoud Alosaimi
Shaikha Bin Ateeq

INTRODUCTION

One of the problems that medical clinics face is when a person books an appointment but does not show up. Every appointment that is not attended deprives another patient of the opportunity to benefit from it.

GOAL

Predict whether the patient will show or not show.

DATASET

Data Source

Kaggle

Which Contains

Data about clinics in Brazil and their patients

Rows & Columns

Rows: 110527
Columns: 14

Features

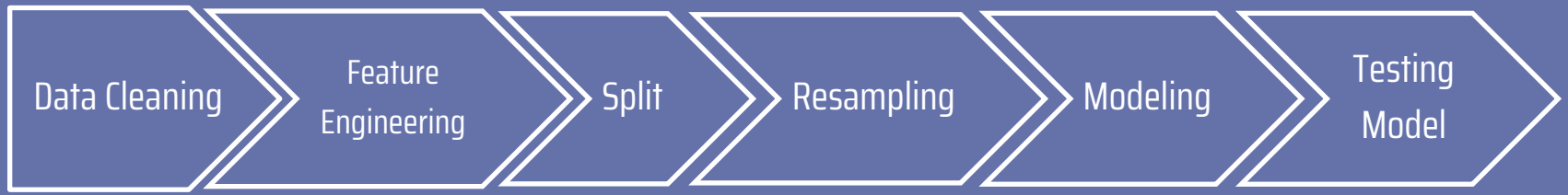
Patient &
AppointmentID
Scheduled &
AppointmentDay
Age
Gender
Neighborhood
Scholarship
Hipertension
Diabetes
Alcoholism
Handcap
SMS_received

Target

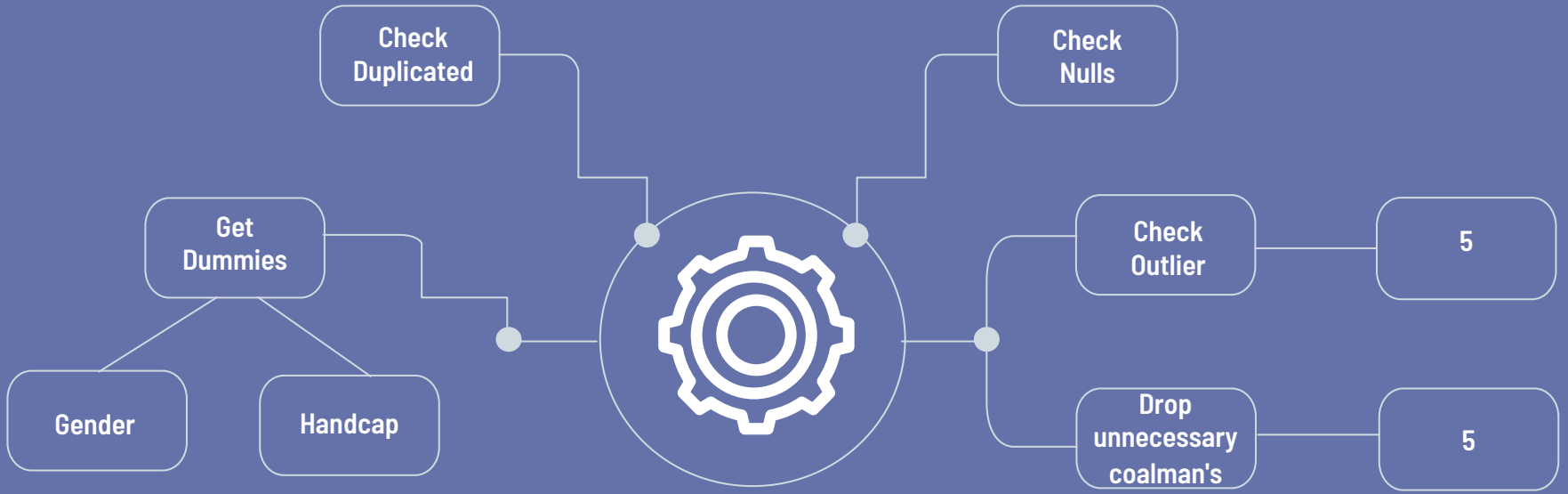
No Show

TOOLS

PROJECT WORKFLOW



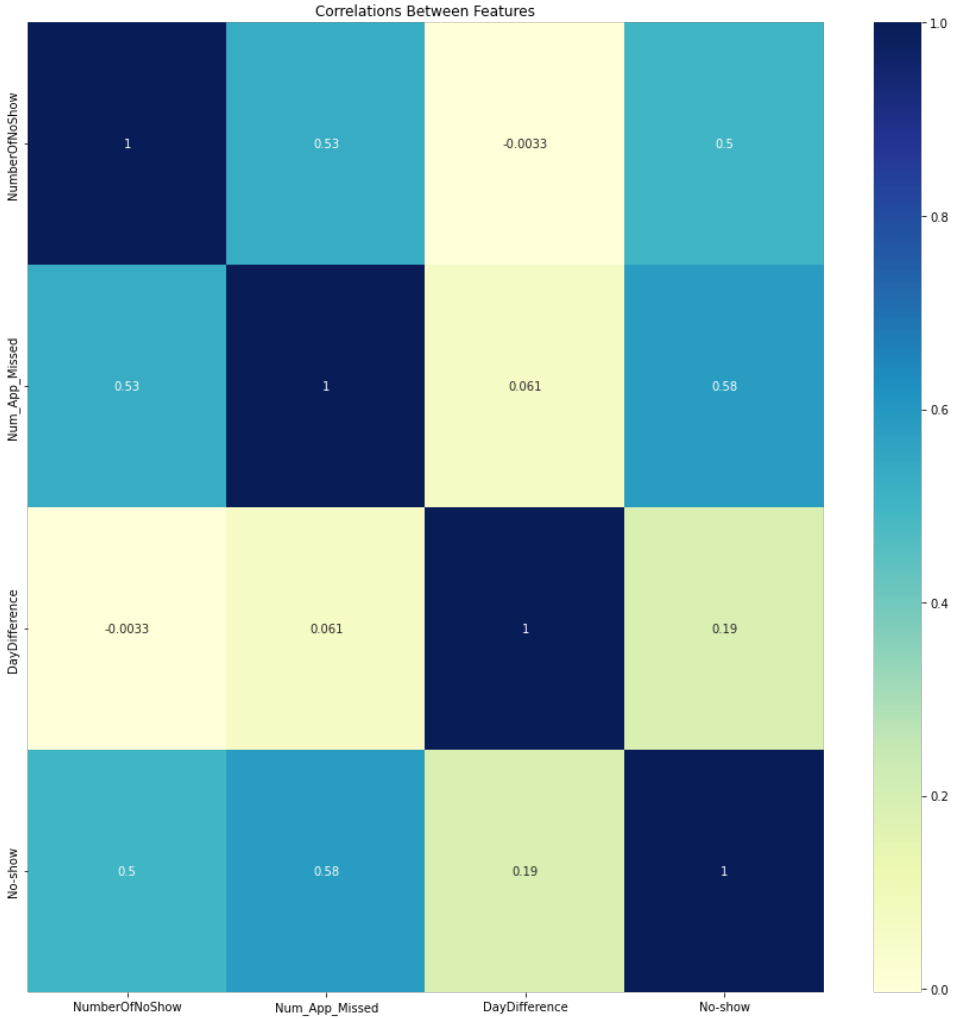
DATA PREPROCESSING



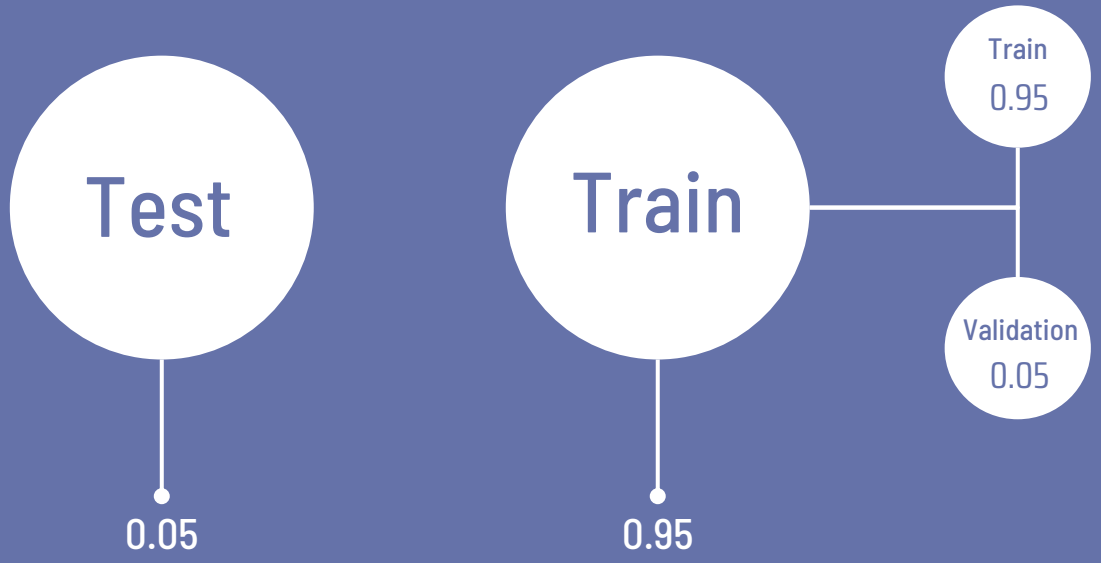
Feature Engineering

- We have 6 new Features.
- We have a good correlations with the target.

BEFORE	AFTER
Rows: 110527	Rows: 110522
Columns: 14	Columns: 21

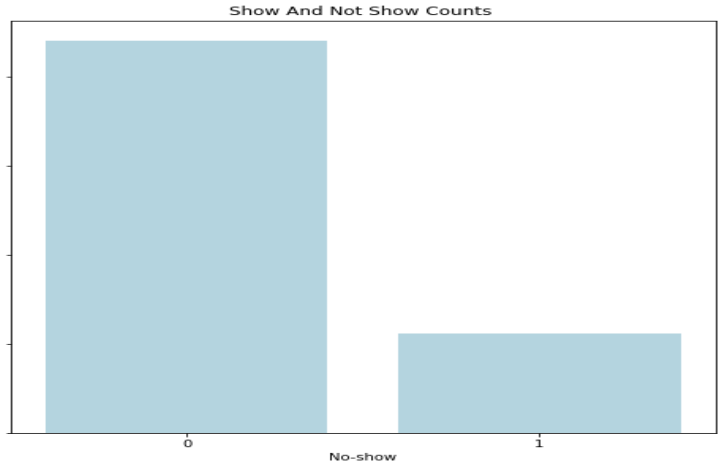


SPLIT

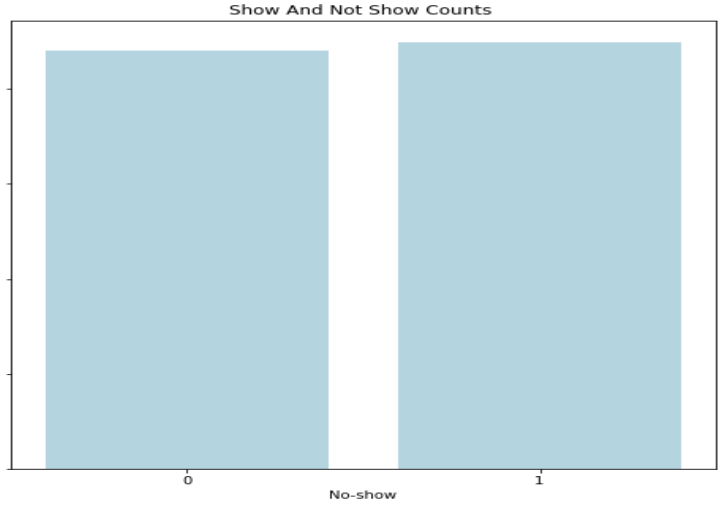


RESAMPLING

Data Before Oversampling



Data After Oversampling

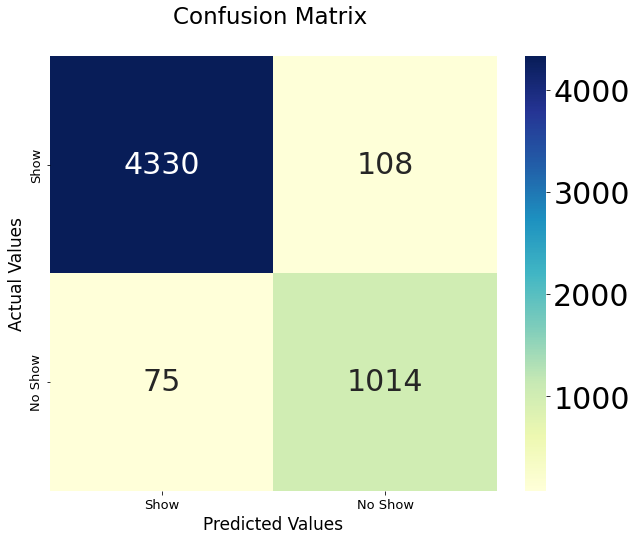


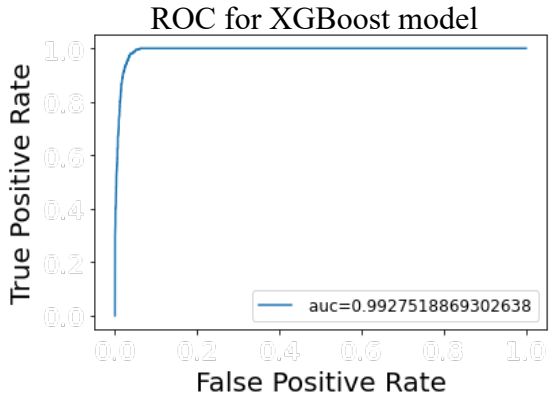
MODELING

Model	Train	Validation
Baseline (LR)	0.8	0.8
Logistic Regression	0.94	0.92
K-Nearest Neighbors	0.97	0.93
Decision Tree	0.99	0.95
Extra Trees	0.99	0.95
Random Forest	0.99	0.96
Support Vector Machine	0.96	0.94
Naive Bayes (Gaussian)	0.8	0.86
Naive Bayes (MultinomialNB)	0.93	0.89
XGBoost	**0.97**	**0.95**

CONCLUSION

- The best model in the term of performance is XGBoost.
- The train accuracy was 0.99 which the test accuracy was 0.97





Thank You