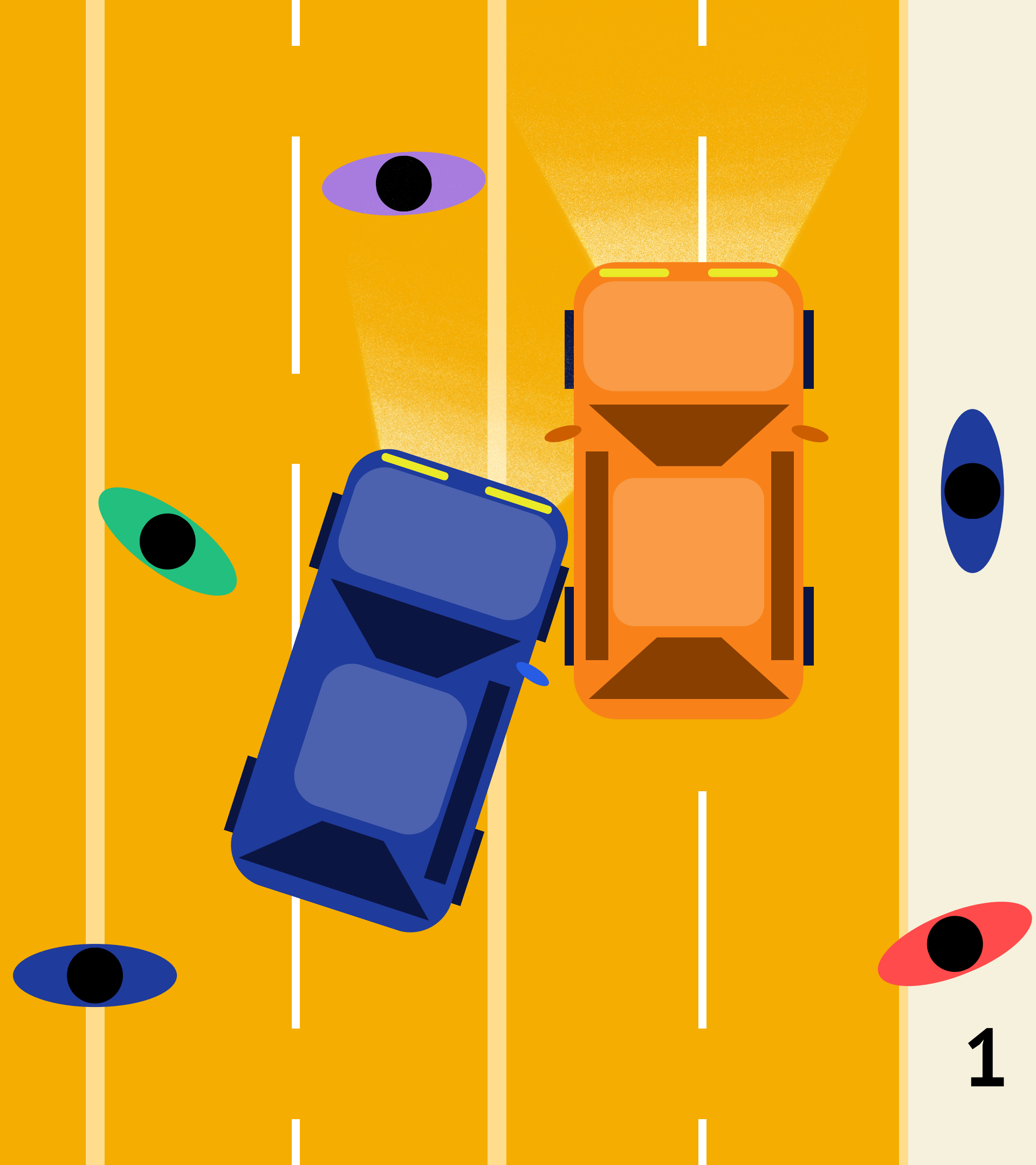# PARKING TICKETS

CITY OF TORONTO

Alla B.

OVERVIEW

- Generate multimillion revenue for the City of Toronto
- Declining revenue
- Reverse this decline
- Efficient Administrative Penalty System (APS)
- Machine learning: predicting fine amounts
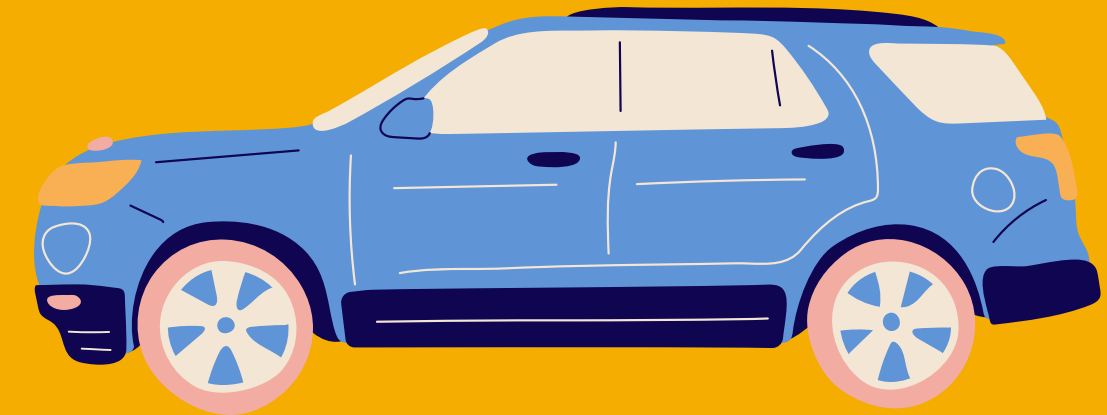- Impact: safer city, more attractive to residents and tourists

1

DATASET AND PREPROCESSING

- time of infraction -> hour, minutes

- infraction description

- date of infraction -> day, month, year

- location -> lat, lon

- DATASET: https://open.toronto.ca
- 2020, 2021, 2022
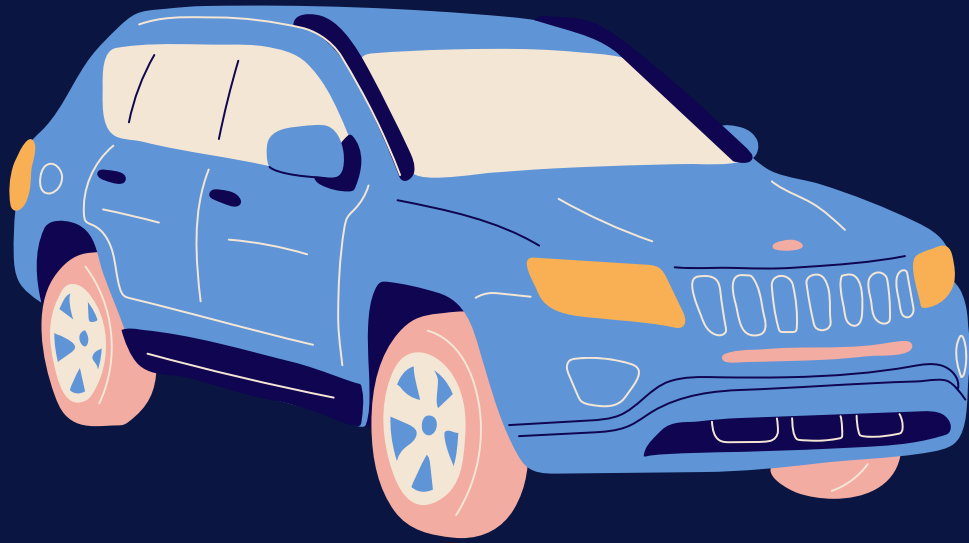- 4707259 rows
- 11 columns

FEATURE ENGINEERING

- one hot encoding:
  - infraction description

- geocoding:
  - Nominatim by OpenStreetMap
  - 2500 unique values

2

FINDINGS FROM EDA AND MODELLING

- regression -> classification

- out of 17 classes picked most frequent 9

- LogisticRegression

  Experimenting with other algorithms
  no hyperparameter tuning

- RandomForest
- XGBoost
- AdaBoost

- accuracy score 0.99
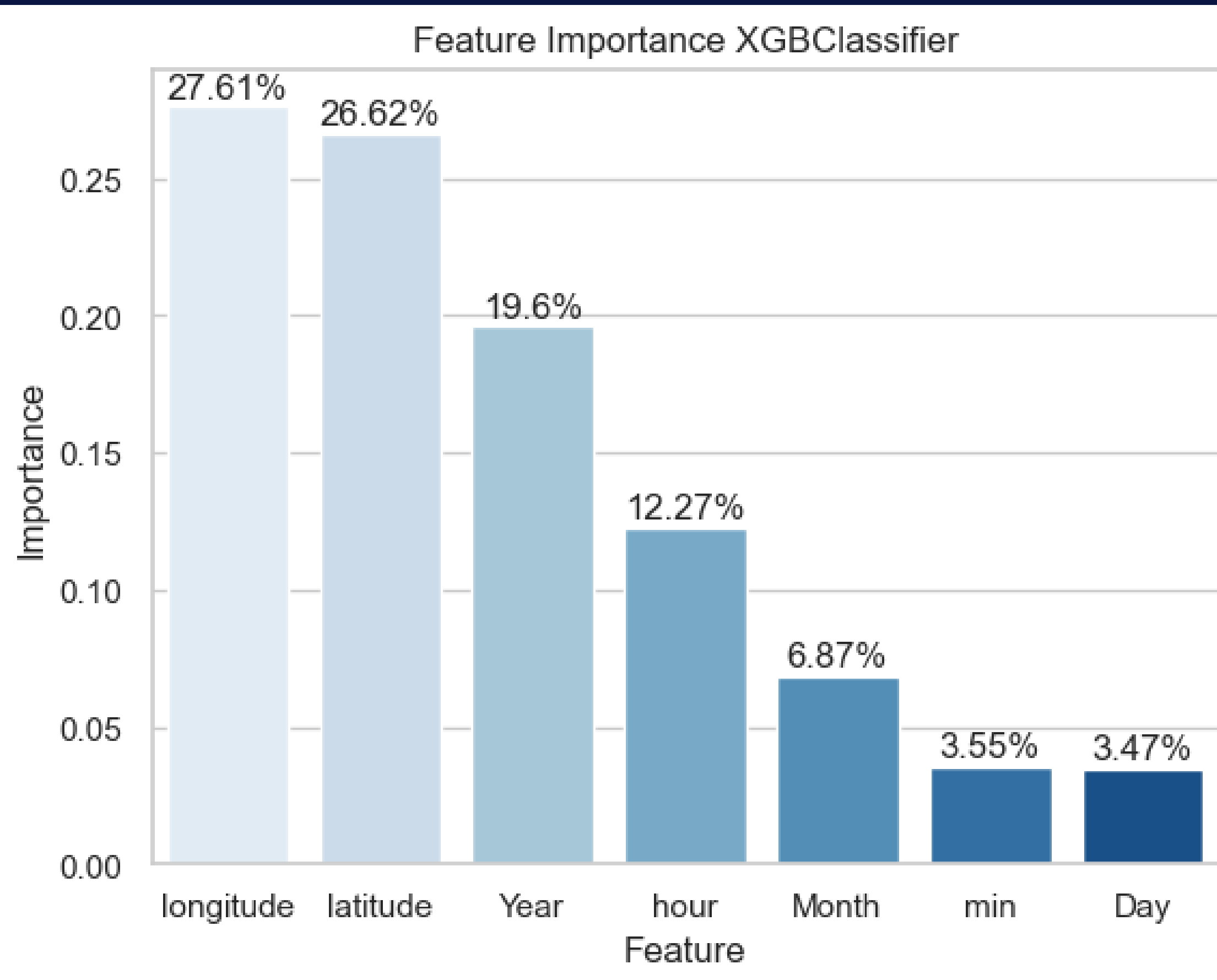
MODELLING

- RandomForest: 0.66 -> 0.68
- param_grid = { 'max_depth': [5, 10, 15, 20],
                   'min_samples_split': [1,2,3],
                   'n_estimators': [64, 128, 256, 512]}

- XGBoost: 0.75 -> 0.73
- param_grid = {'learning_rate': [0.01, 0.1, 0.2],
                 'max_depth': [3, 4, 5],
                 'colsample_bytree': [0.8, 0.9, 1.0],
                 'n_estimators': [32, 128, 256]}

- AdaBoost: 0.34 -> 0.34
- param_grid = {'n_estimators': [50, 100, 200],
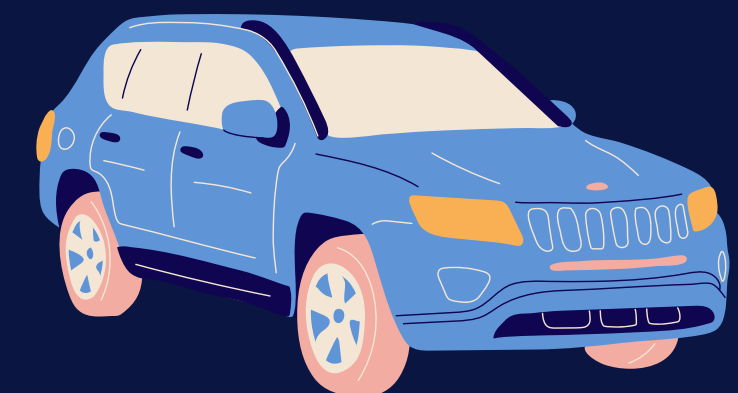   'learning_rate': [0.01, 0.1, 1.0]}

4

Feature Importance XGBClassifier

latitude and longitude are the most important features

NEXT STEPS

- Convert all addresses to Latitude and Longitude

- More hyperparameter tuning

- Confusion matrix

- ROC curve

- Streamlit app

THANK YOU!