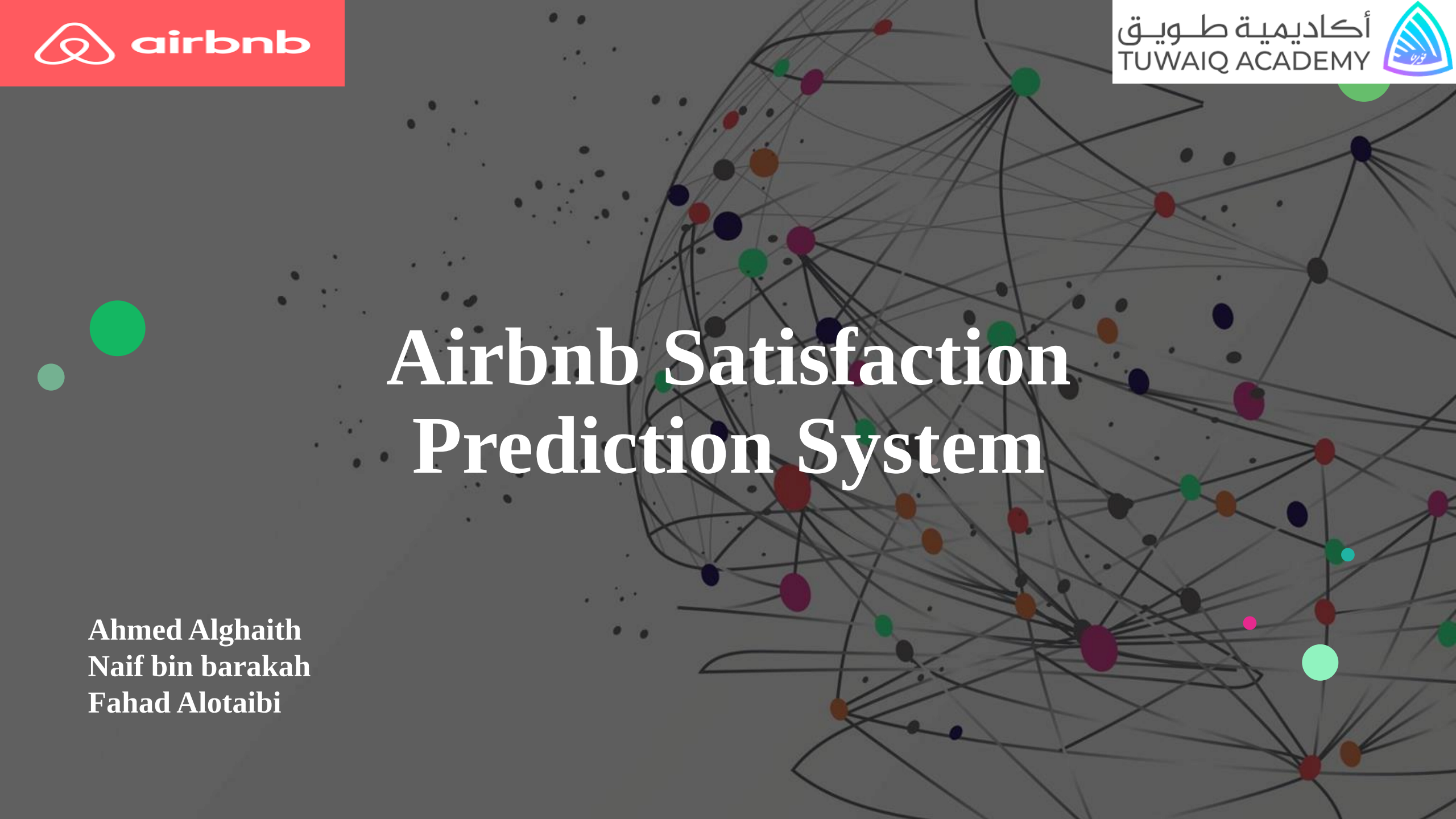


Airbnb Satisfaction Prediction System

Ahmed Alghaith
Naif bin barakah
Fahad Alotaibi

Introduction

- Airbnb
- Why Guest satisfaction
- Potential solution

أكاديمية طويق
TUWAIQ ACADEMY

Problem definition

The process of building an AI model :

- Choosing the dataset
- Pre-processing
- Selecting an appropriate machine learning algorithm
- Training and testing the AI model
- Evaluating and fine-tuning the model

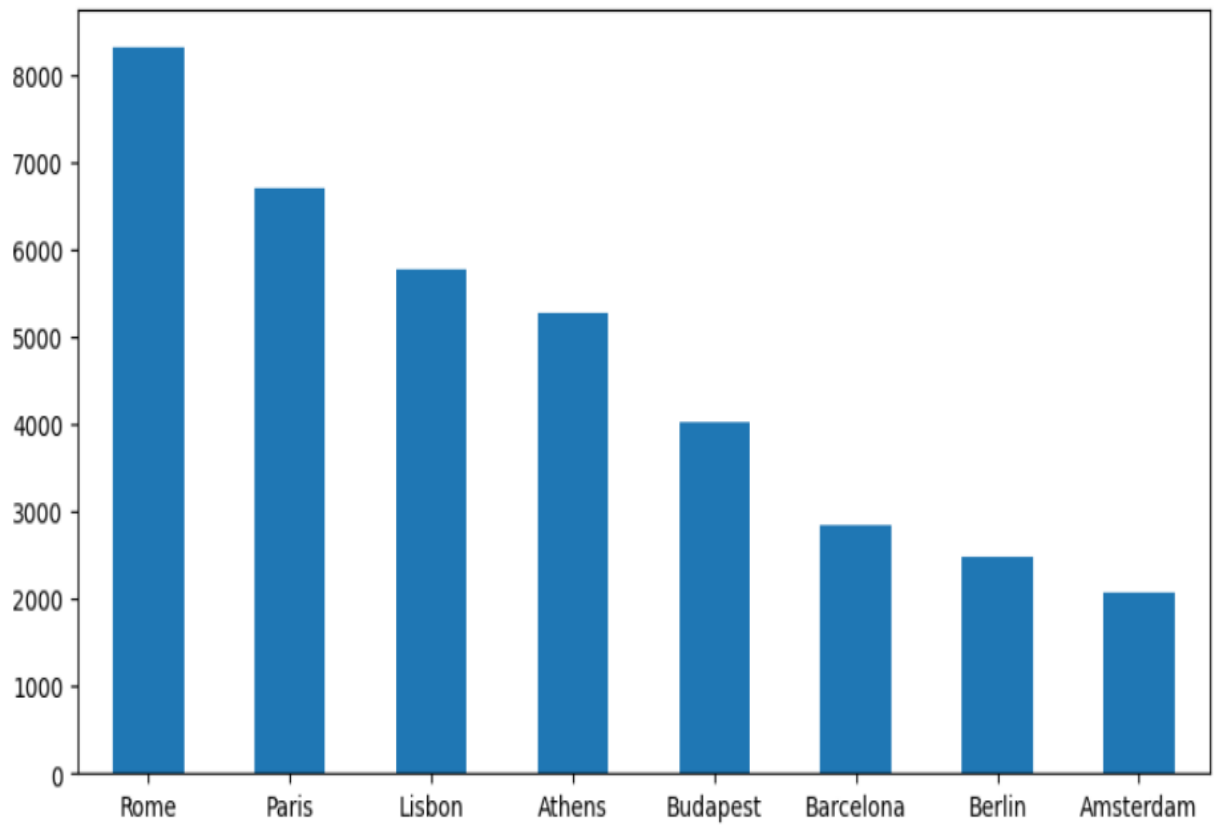
kaggle

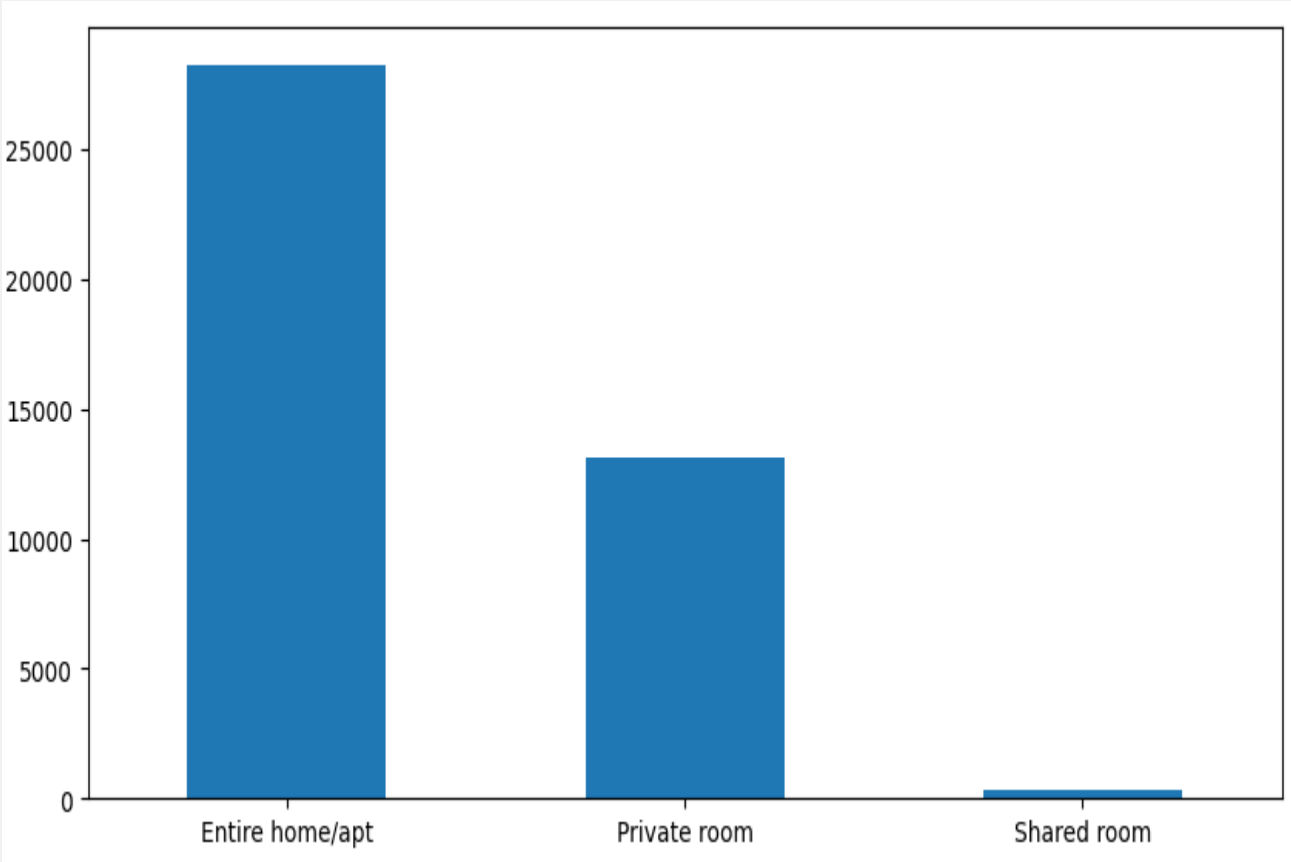
أكاديمية طويق
TUWAIQ ACADEMY

Tools

أكاديمية طويق
TUWAIQ ACADEMY

Data Distribution





- 40.000 Records
- 19 Columns

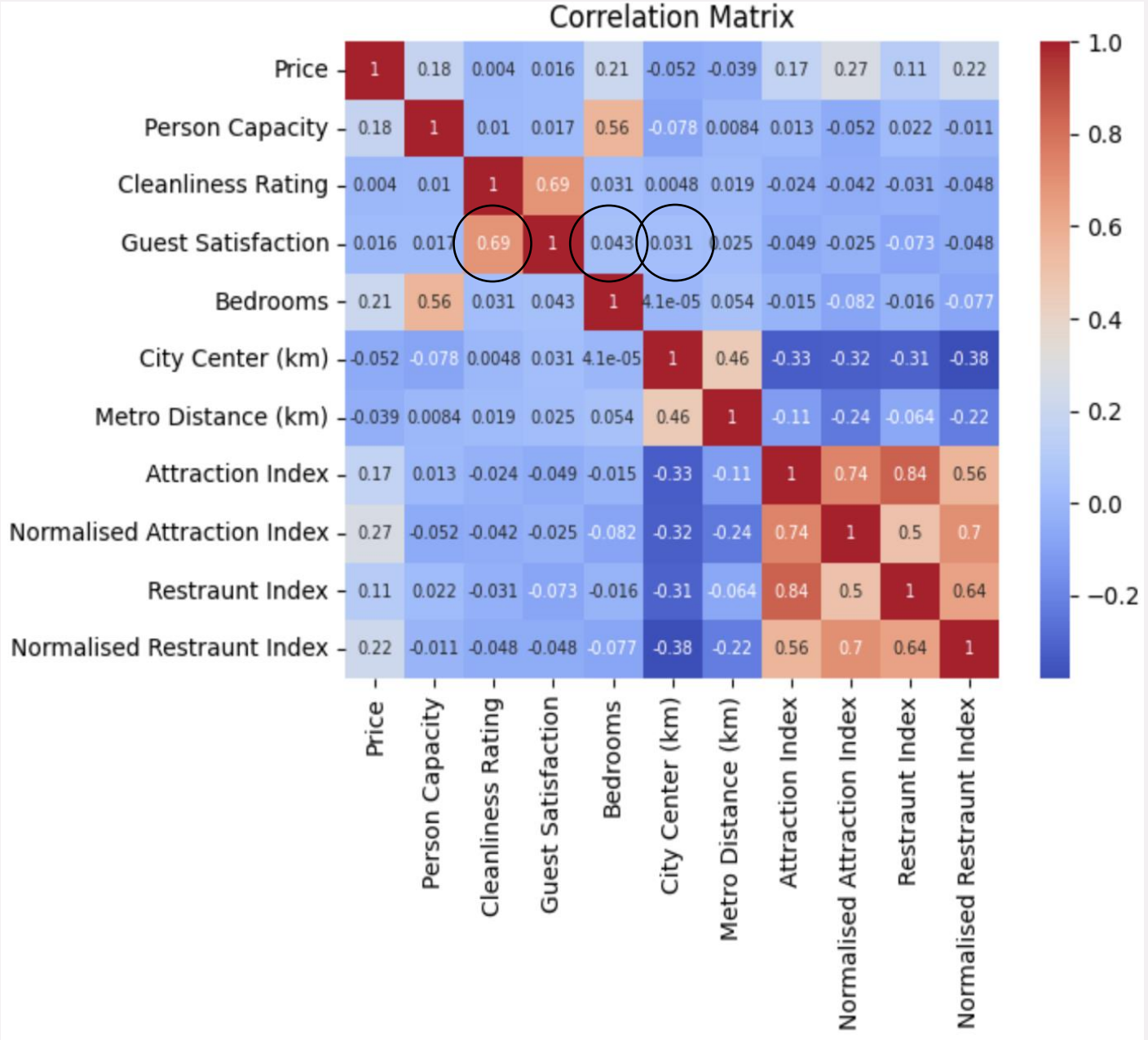
Data Pre-Processing

- **Handling Boolean data**
  - Convert true and false columns to 0 and 1.
- **Handling categorical data**
  - Convert categorical data (e.g., City) into numerical data using one-hot encoding.
- **Feature scaling**
  - Scale numerical features to have similar ranges using techniques such as standardization.
- **Feature engineering**
  - Create new features from existing ones that might be more useful for the machine learning algorithm.

أكاديمية طويق
TUWAIQ ACADEMY

Data Analysis

- **High correlated features**
  1. Cleanliness Rating
  2. Bedrooms
  3. City Center (km)



Neural Network architectures

	Hidden dense layers number	Neurons number with layers	Batch Normalization layers number	Training accuracy	Validation accuracy	Difference
Architecture1	3	All 16	0	92.80%	92.85%	-0.05
Architecture2	6	All 16	0	93.03%	92.81%	0.22
Architecture 3	3	32,16,8	0	92.92%	93.08%	-0.16
Architecture 4	3	All 16	3	93.03%	92.99%	0.04

Models cross validation scores

Gradient Boosting Classifier: 92.76%

Neural Network: 92.72%

Logistic regression: 92.45%

XGBClassifier: 94.07%

Support Vector machine: 92.54%

DecisionTreeClassifier: 92.81%

RandomForestClassifier: 94.15%

Model Hyperparameter tuning

Grid search with cross validation scores

RandomForestClassifier: 94.20%(+0.05%)

XGBClassifier: 95.34%(+1.27%)

DecisionTreeClassifier: 93.32%(+0.51%)

VotingClassifier(Hard): 94.44%

VotingClassifier(Soft): 94.31%

StackingClassifier: 94.17%

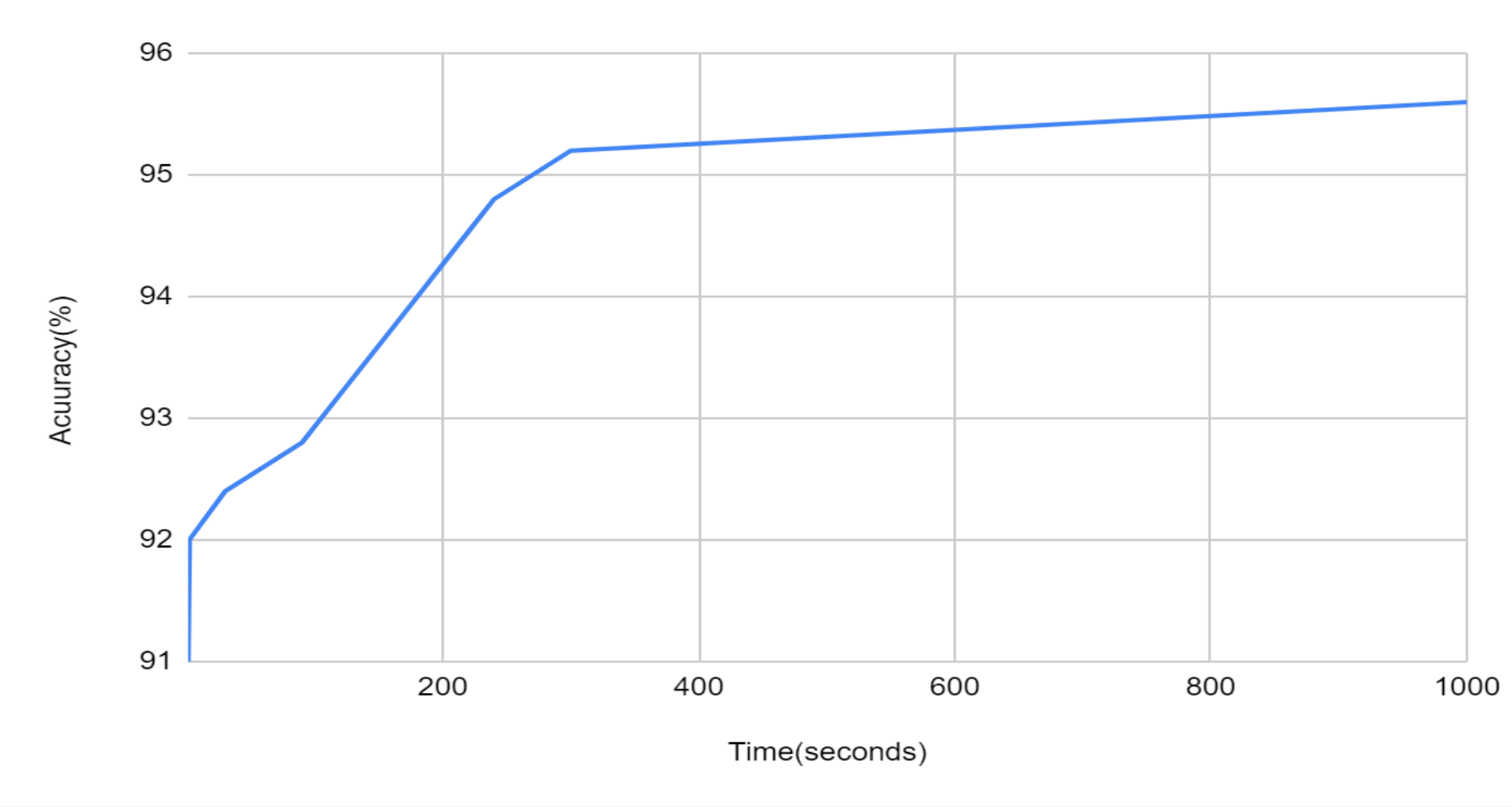
Report the results- XGBClassifier

Test accuracy: 96.27%

cross validation scores: 95.34%

أكاديمية طويق
TUWAIQ ACADEMY

Trade off Time Vs Accuracy



- # Conclusion

  Project opportunities in the future

أكاديمية طويق
TUWAIQ ACADEMY

- Thank you!

GitHub

Project repository

Ahmed

Fahad

Naif