Predicti g Ba galore Hou e Rate

This project aims to help people interested in buying a home in Bangalore, India by developing a model to accurately predict house rates across the city. Using data collected from Kaggle, we'll explore the factors that influence Bangalore's dynamic real estate market.

I troductio to t e Project

Objective

The primary goal is to create a predictive model that can estimate the price of a house in Bangalore based on various features.

Sig ifica ce

This will empower prospective home buyers with valuable insights, allowing them to make informed decisions and negotiate effectively.

Approac

The project will involve data collection, preprocessing, exploratory analysis, feature engineering, model selection, and deployment using Flask.

Data Collectio a d Preproce i g

1

Data Source

The dataset was obtained from Kaggle, containing 13,320 rows and 9 columns of information on Bangalore houses.

2

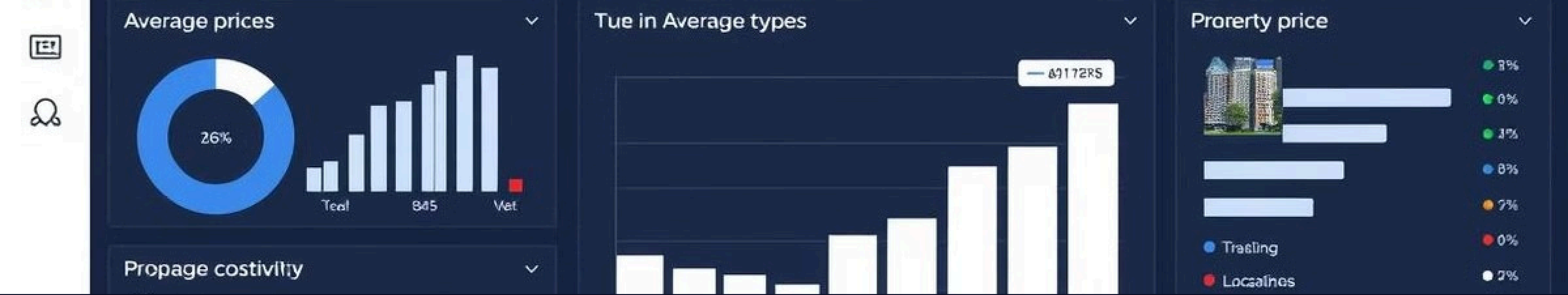
Preproce i g

The 'total_sqft' column, which provided the area in a range format, was converted to a float by taking the average.

3

Clea i g

The data was checked for missing values, outliers, and inconsistencies, which were then addressed to ensure data integrity.



Exploratory Data A aly i (EDA)

1 2Price Di tributio

The distribution of house prices in Bangalore was examined, revealing a wide range from affordable to high-end properties.

I flue tial Factor

The relationship between features like location, area, and number of bedrooms with house prices was explored.

3 Market I ig t

The EDA provided valuable insights into the trends and patterns in Bangalore's dynamic real estate market.

Feature E gi eeri g a d Selectio

Feature E gi eeri g

New features were created by combining and transforming existing variables to improve the model's predictive power.

Feature Selectio

Techniques like correlation analysis and recursive feature elimination were used to identify the most important predictors.

Di e io ality Reductio

Principal Component Analysis (PCA) was applied to reduce the number of features while preserving the maximum variance.

Ha dli g Multicol li earity

Multicollinear features were identified and addressed to ensure the model's stability and reliability.

Model Selectio a d Trai i g

Regre io Model

Multiple regression algorithms, such as Linear Regression, Random Forest, and XGBoost, were evaluated.

Hyperpara eter Tu i g

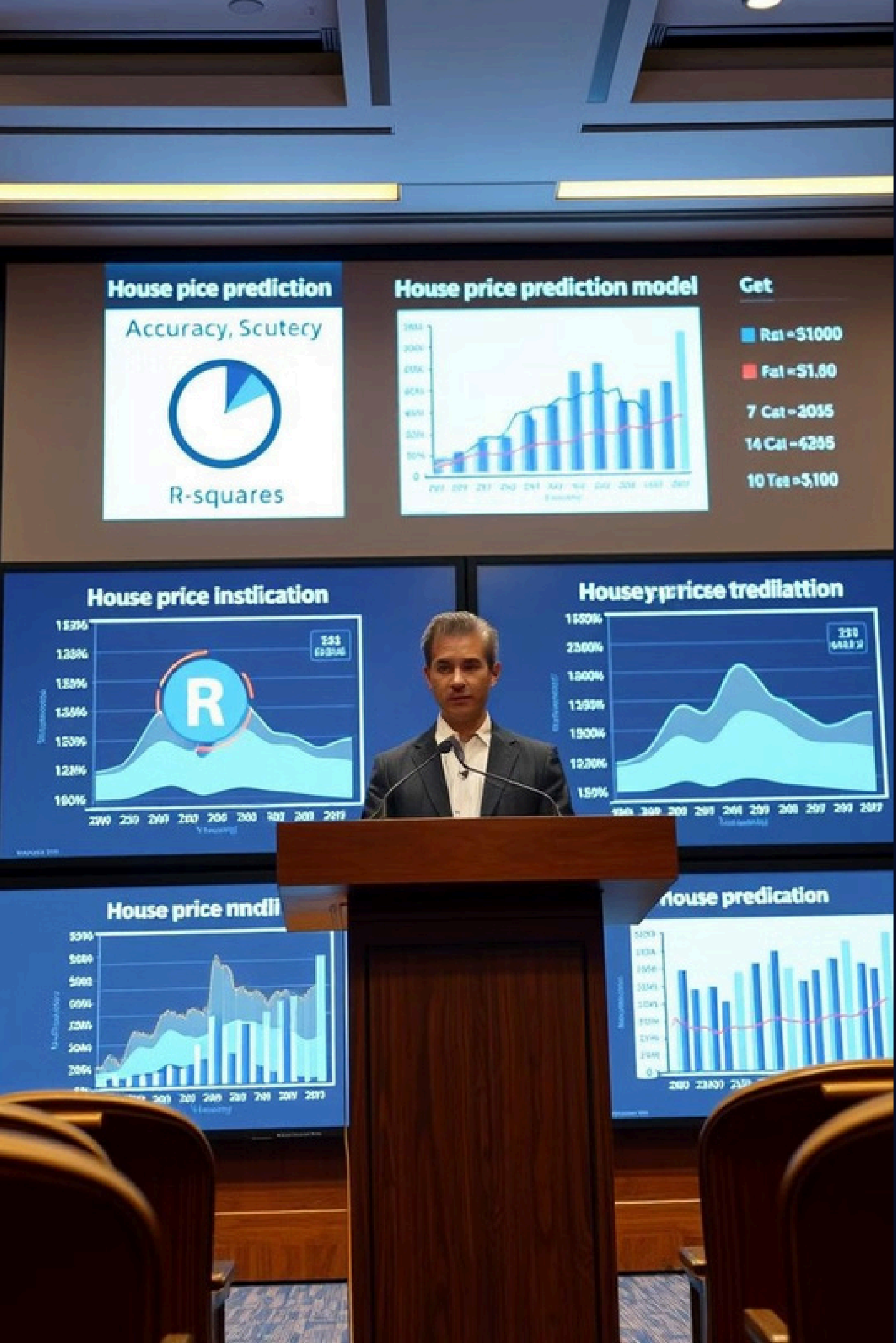
The models' hyperparameters were optimized using techniques like grid search and cross-validation.

Model Evaluatio

The models were assessed using various metrics, including R-squared, Mean Squared Error, and RMSE.

Fi al Model Selectio

The best-performing model was chosen as the final solution for predicting Bangalore house rates.



Model Evaluatio a d Validatio

1

Trai -Te t Split

The dataset was divided into training and testing sets to evaluate the model's generalization performance.

2

Cro -Validatio

Multiple rounds of cross-validation were performed to ensure the model's robustness and stability.

3

Perfor a ce Metric

The model's predictive accuracy was assessed using metrics like R-squared, RMSE, and Mean Absolute Error.

Deploy e t u i g Fla k

Requirement	Description
User Interface	A web-based interface built using Flask, allowing users to input property details and view the predicted price.
Model Integration	The trained predictive model was integrated into the Flask application to provide real-time house price estimates.
Scalability	The deployment setup ensures the solution can handle increasing user traffic and data volume.

Co clu io a d Key Takeaway

1. ### Accurate Predictio
   The predictive model developed in this project can reliably estimate Bangalore house prices based on various factors.

2. ### E poweri g Ho e Buyer
   The web-based application provides home buyers with valuable insights, enabling them to make informed decisions.

3. ### Future E a ce e t
   Ongoing data collection and model refinement can further improve the accuracy and usability of the solution.