VISHNU UNIVERSAL LEARNING | BVRIT H

BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

TITLE

Real Estate Price Prediction

Team BrainStormers:

R.Sai Srinidhi- 6631

K.Mahima- 6616

V.Chandu Priya- 6618

U.Tejaswini- 6633

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

INDEX

VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

PROBLEM STATEMENT

The real estate market is highly dynamic, with property prices fluctuatin.g based on various factors such as location, amenities, demand, and economic conditions. Traditional property valuation methods rely on manual assessments, which can be time-consuming, inconsistent, and error-prone. This project aims to develop a machine learning-based price prediction system that provides accurate property valuations based on real-time data

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

ABSTRACT

The Real Estate Price Prediction System is designed to provide accurate property valuations using machine learning and real-time data integration. The project addresses the challenge of inconsistent and manual pricing in the real estate market by leveraging data-driven predictive analytics. The solution involves web scraping real estate listings, extracting relevant features, and training ML models to predict property prices. The system will be deployed as an API for seamless integration with real estate platforms and mobile applications. By automating price estimation, this project aims to enhance market transparency, reduce pricing discrepancies, and empower buyers, sellers, and investors with actionable insights.

VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Partial Implementation:

Tools Used:

- Selenium for Web Scraping
- Selenium Webdriver
- CSV to save files in csv

Extracted the data and saved into csv file

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Literature Survey

.

1.Khatri et al. (2021) Used Selenium to scrape property data from MagicBricks for structured dataset creation and model training. IJERT

2.Sharma & Verma (2020) Combined Selenium and Google Maps API to enrich real estate datasets with location-based features. IJCSIT

3.Roy et al. (2022) Applied ML models on Selenium-scraped data; Random Forest and XGBoost showed high accuracy. IJEAT

4. Dutta et al. (2023) Built a full MLOps pipeline using Selenium, Docker, and MLflow for automated price prediction. IJAICT

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Methodologies

1. Web Scraping with Selenium : Automated the extraction of land listings, including price, plot size, and location, from property website 99acres.

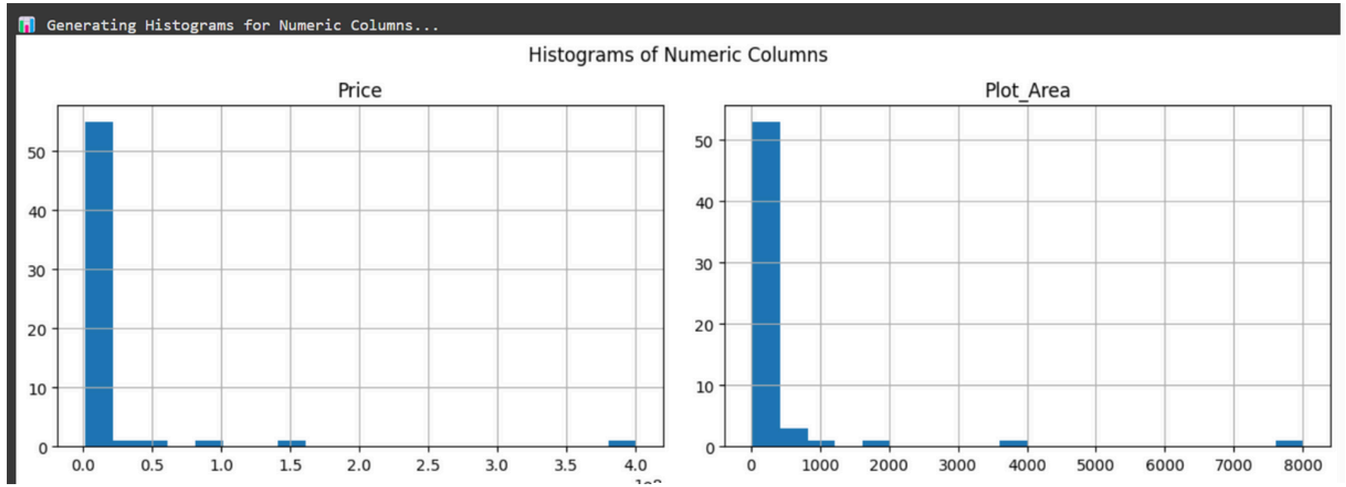
2. Data Cleaning & Preprocessing: Filtered out incomplete listings, removed banners/ads, and standardized price and plot area for consistency.

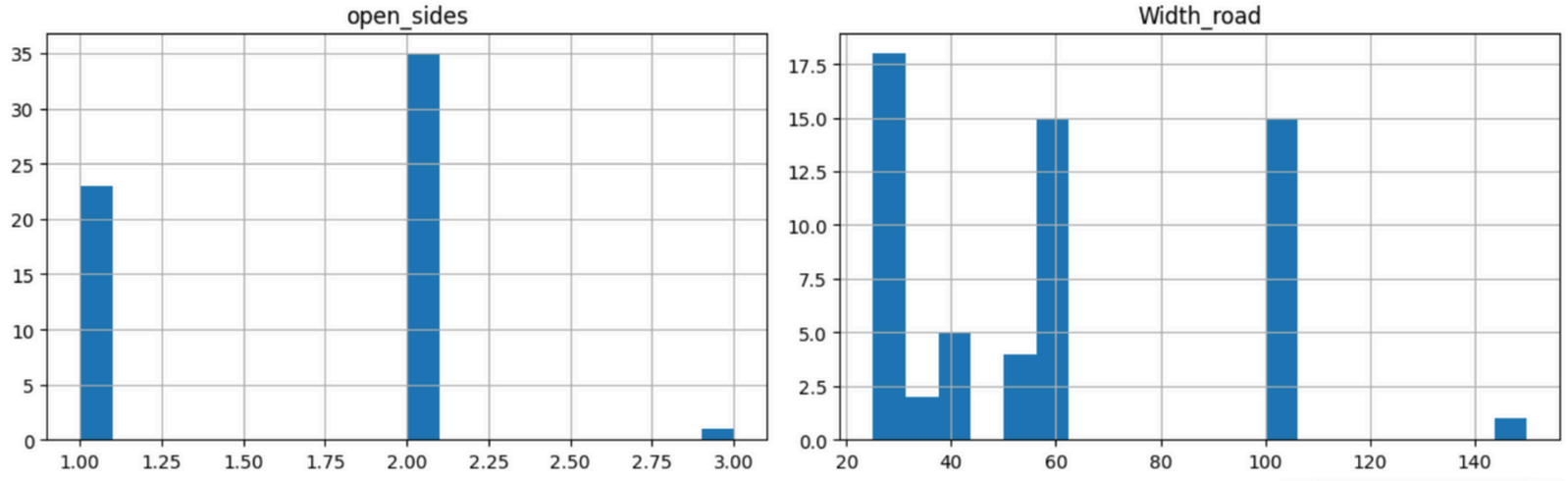
3. Feature Selection & Engineering : Focused on key factors like plot size, locality, facing direction, and road width; applied label encoding for categorical fields.

4. Model Training & Validation: Used Linear regression model to predict land prices, validated with metrics like RMSE and MAE.

VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

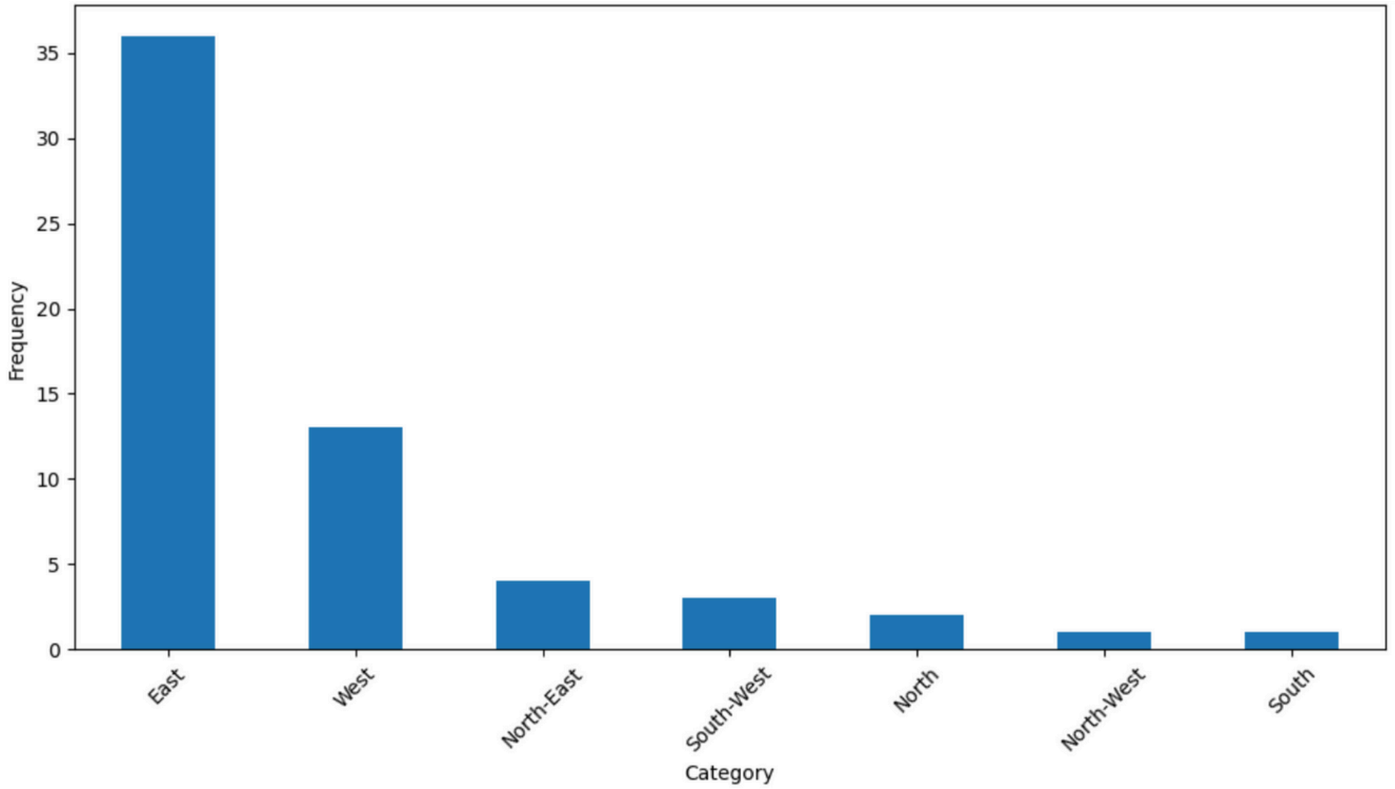
Results for EDA:





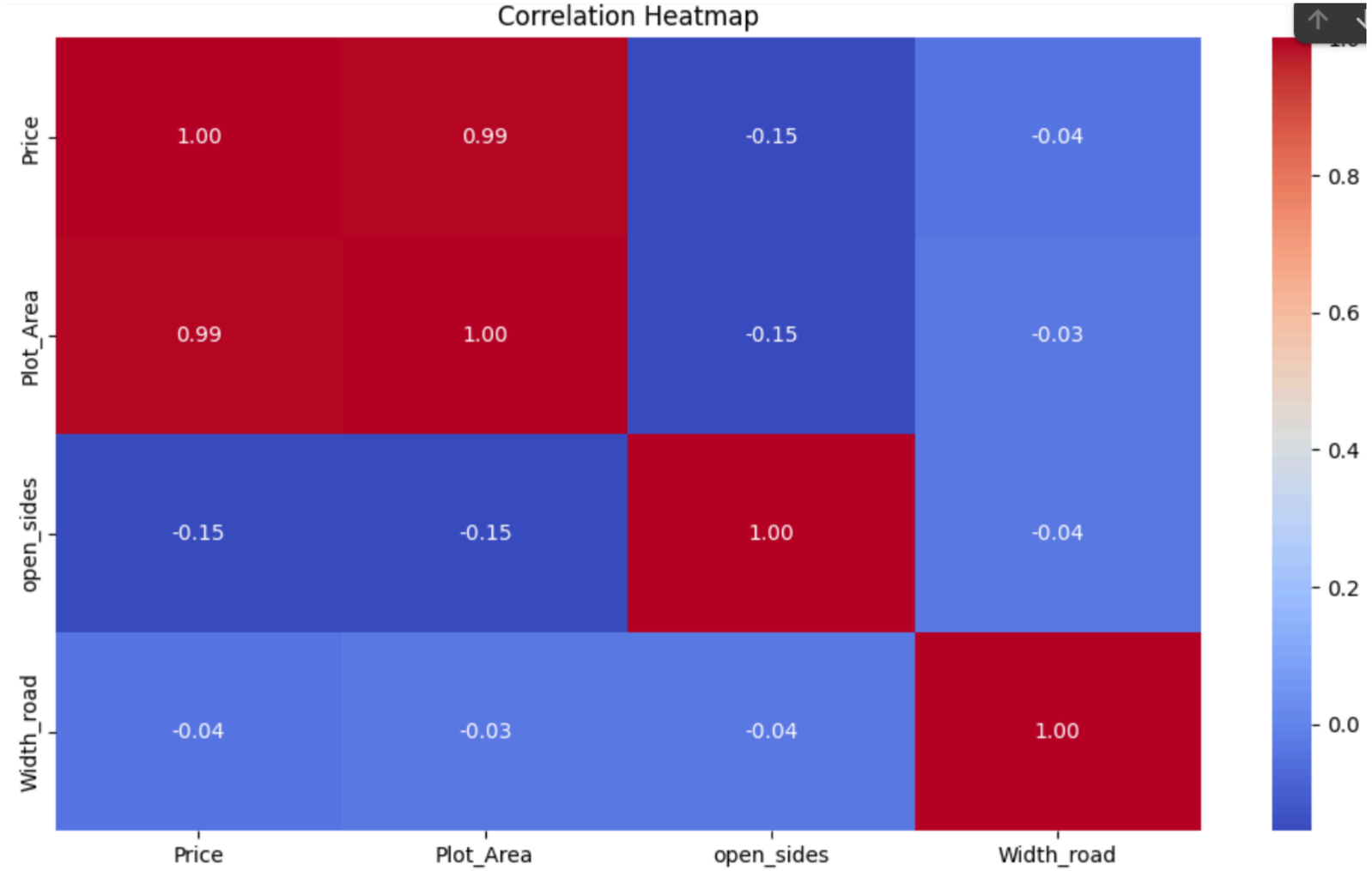
VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Results for EDA:



VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Results for EDA:



VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Model Evaluation:

```
🔍 Linear Regression Model Evaluation:
R² Score          : 0.9334
MAE               : 6851691.4701
MAPE (%)          : 67.08
Train Time (s)    : 0.024
```

VISHNU UNIVERSAL LEARNING | BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Price prediction

```
200
East
3
60
Predicted Price: ₹9575523.94
```

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Conclusion

This project successfully demonstrates land price prediction using machine learning models trained on data scraped with Selenium. By automating data collection from real estate websites, we built a clean and structured dataset. The trained models achieved reliable performance, with evaluation metrics like R^2 indicating strong correlation between predicted and actual prices, and low MSE reflecting accurate estimations. This approach can support better decision-making for buyers, sellers, and investors in the real estate market.

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

Future Enhancements

1. Include Satellite Imagery & Maps :Integrate visual data (e.g., surroundings, terrain) from sources like Google Earth to improve location-based insights.
2. Add Time-Based Trends: Incorporate historical price data to capture market fluctuations and predict future land value trends.
3. Enhance Location Precision: Use geo-tagging or pin-code level data to improve accuracy in predicting prices across micro-locations.
4. Real-Time Model Updating : Automate periodic scraping and model retraining using MLOps tools to keep predictions up-to-date.

VISHNU UNIVERSAL LEARNING | BVRIT H
BVRIT HYDERABAD College of Engineering for Women (UGC-Autonomous)

References

1. 99acres – Property Listing Website
https://www.99acres.com
Used as the primary data source for extracting property listings in Patancheru.

2. Selenium – Python Web Automation Library
https://www.selenium.dev/
Employed for automated web scraping and browser interaction to dynamically load and extract data