

Real Estate Price Prediction System

Data-Driven Property Valuation Using Machine Learning

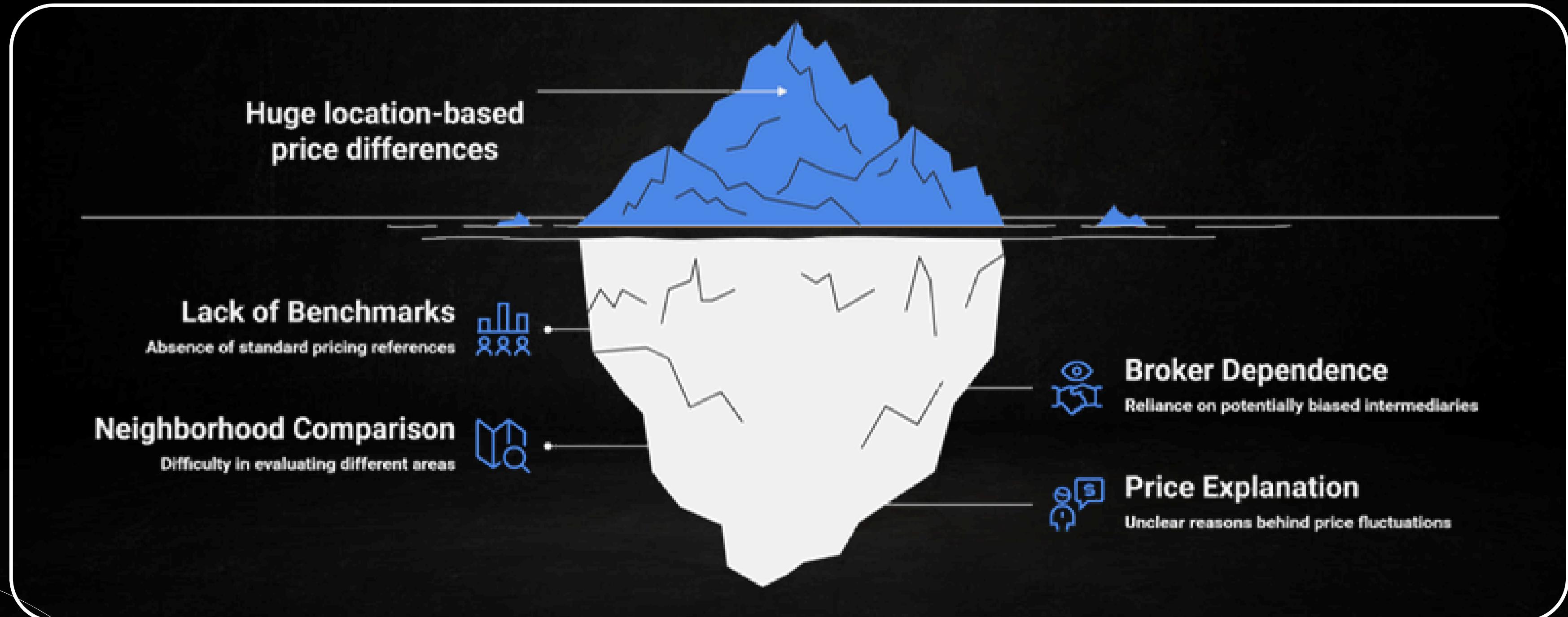
MEMBERS -
UCE2023529 -Kanchan Jadhav
UCE2023530 -Disha Janve
UCE2023546 -Anjali Patel





Problem Statement

Property prices vary widely across Bangalore, making it difficult for buyers to identify a fair price. The real challenges behind property pricing lie hidden beneath surface-level factors.

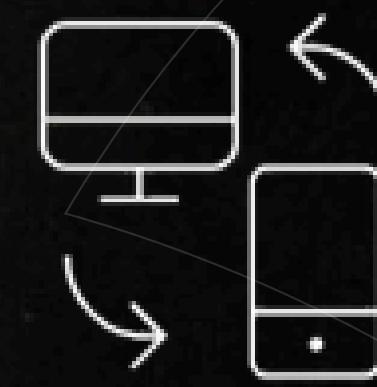
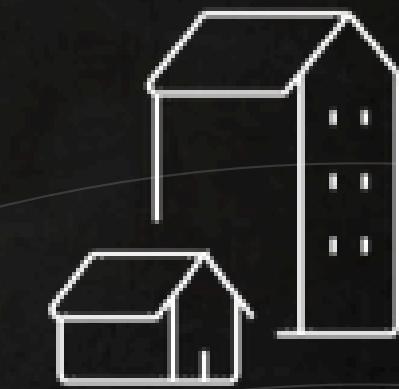
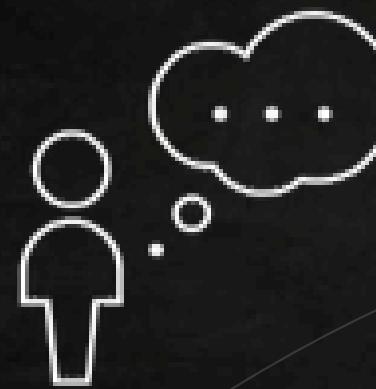
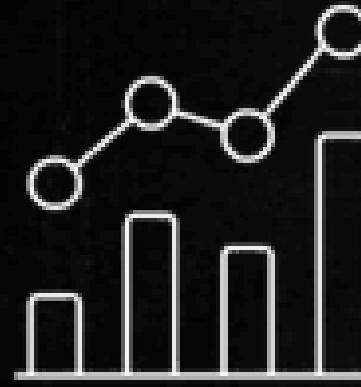




Solution

A fast, transparent and AI-powered platform that predicts accurate property prices and helps users make smarter real-estate decisions.

Property Prediction Features



Accurate Prediction

Predicts property value instantly using sqft, BHK, bath & location.

Explainable Insights

Shows price justification with Fair-Price Meter & feature influence.

Location Comparison

Compares 2–5 areas side-by-side with visual charts.

Real-Time Application

Modern, responsive UI powered by Flask API & dynamic JS.

1

2

3

4



Project Objectives

Develop Accurate ML Model

Create a model for price prediction



Compare ML Algorithms

Evaluate different algorithms for performance



Design Interactive Web UI

Create a user-friendly interface



Clean and Preprocess Data

Prepare real-world data for analysis



Build API Backend

Implement a Flask-based backend



Provide Additional Insights

Offer comparisons, fairness, and influence analysis

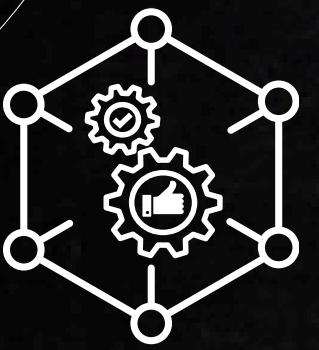


A fast transparent and AI & ML -powered platform that predicts accurate property prices and helps users make smarter real-estate decisions.



Why Property Prices Are Rising In Bangalore?





System Architecture



Frontend (HTML/CSS/JS)



Backend API setup



Model loading and execution



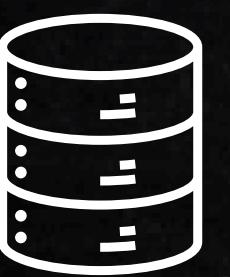
Predictions+Insights



Trained Model
The core component for making predictions

Feature Columns
Provides structure and context for data interpretation

Components Ensuring Model Efficiency and Accuracy



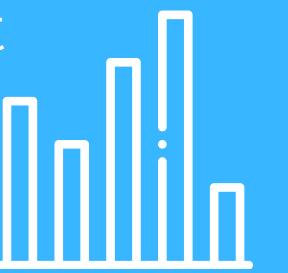
Dataset Overview

Source : Bangalore House Price Dataset (Kaggle)

Records : 13,000+

Features : Sqft, BHK, Bath, Location, Price

After Processing : 140+ columns (One-Hot Encoding)



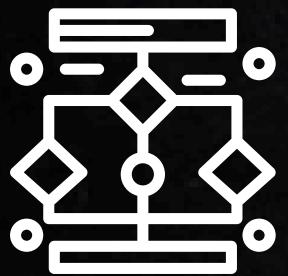
Why This Dataset?

- Rich location diversity
- Good mix of residential data
- Suitable for regression tasks



Data Preprocessing





Algorithms Used

Linear Regression

Good accuracy after
cleaning ($R^2 \approx 0.84\text{--}0.86$)

Sensitive to outliers

Fastest (<1 ms inference)

Highly interpretable
coefficients

Random Forest Regression

Higher accuracy, handles
complexity well

Naturally robust to outliers

Slower, but manageable

Medium (feature importance only)

XGBoost Regression

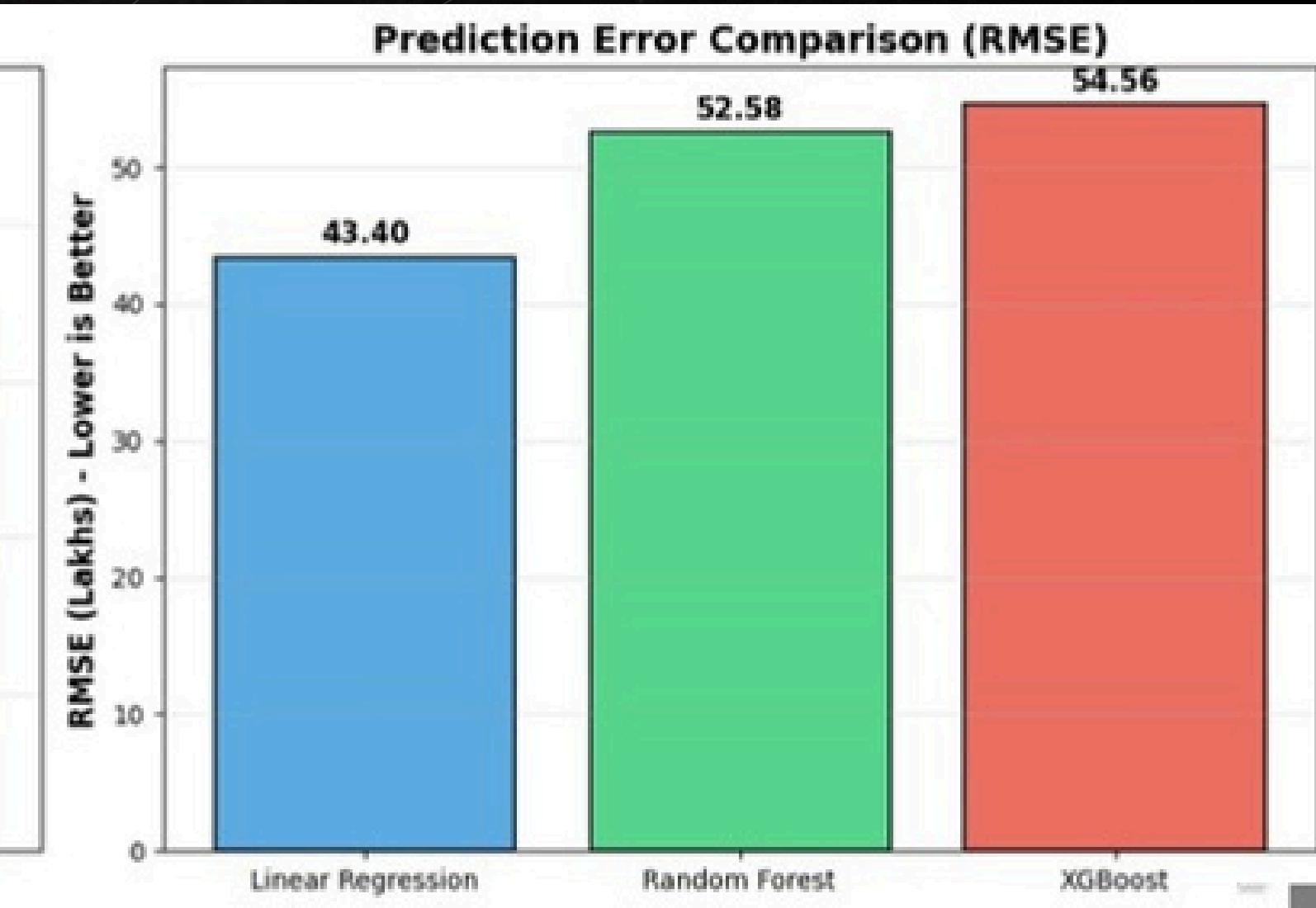
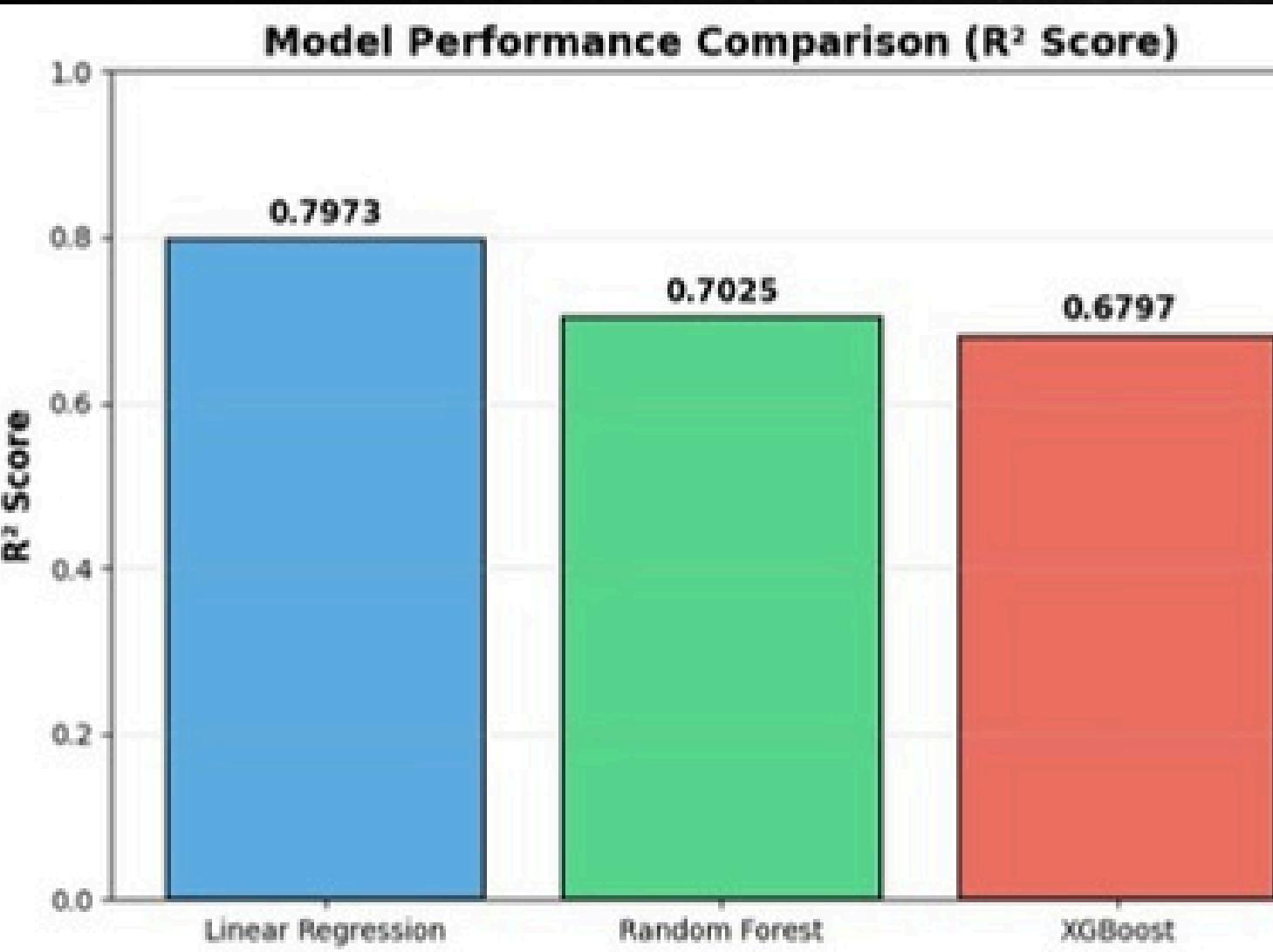
Highest accuracy in most
real-estate datasets

Very robust due to boosting
+ regularization

Slower than RF, high training time

Low interpretability (complex
boosting)

Model Comparison



R^2 Score Comparison (Higher is Better)

Model
Linear Regression
Random Forest
XGBoost

R^2 Score
0.7973
0.7025
0.6797

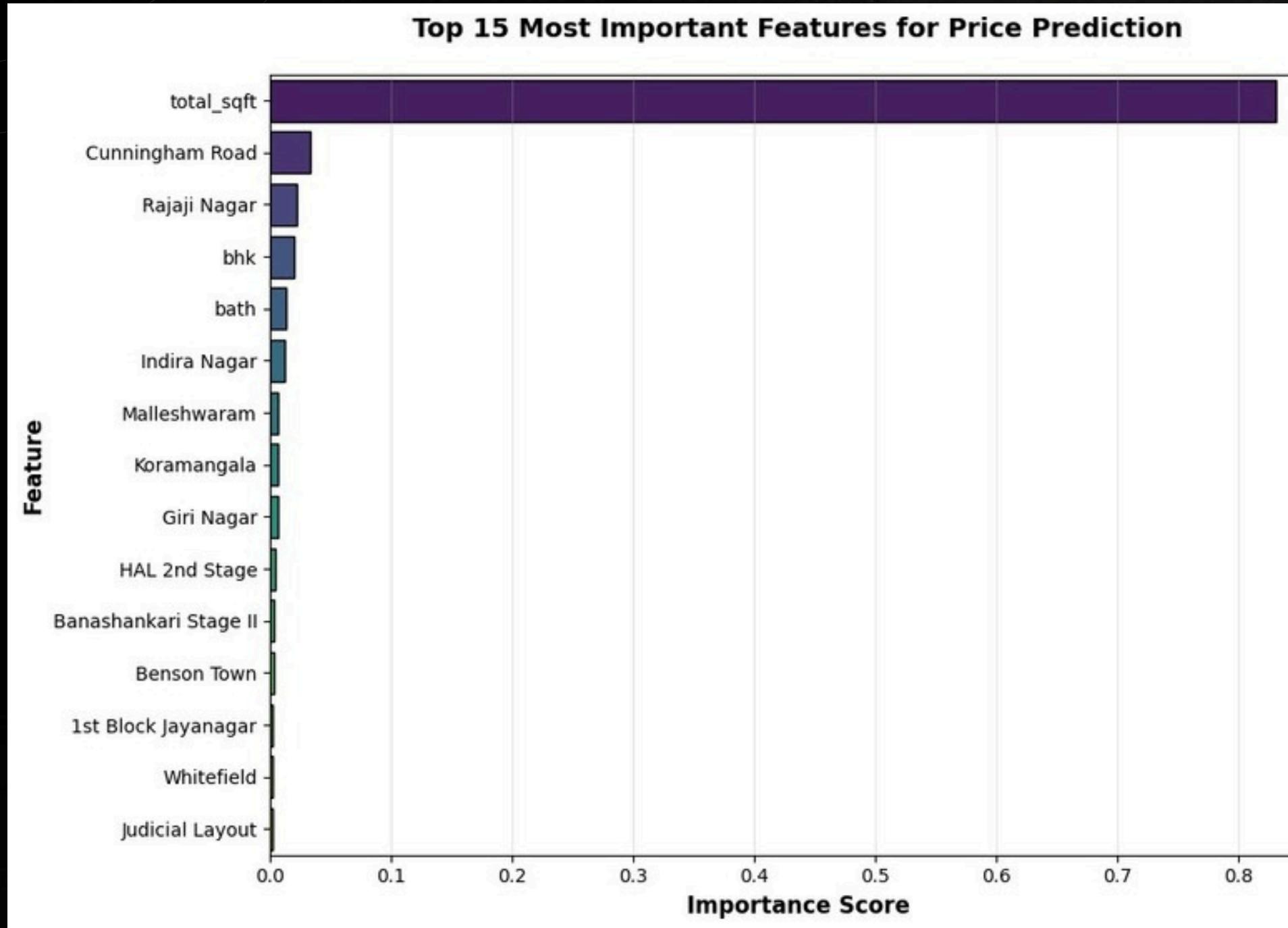
RMSE Comparison (Lower is Better)

Model
Linear Regression
Random Forest
XGBoost

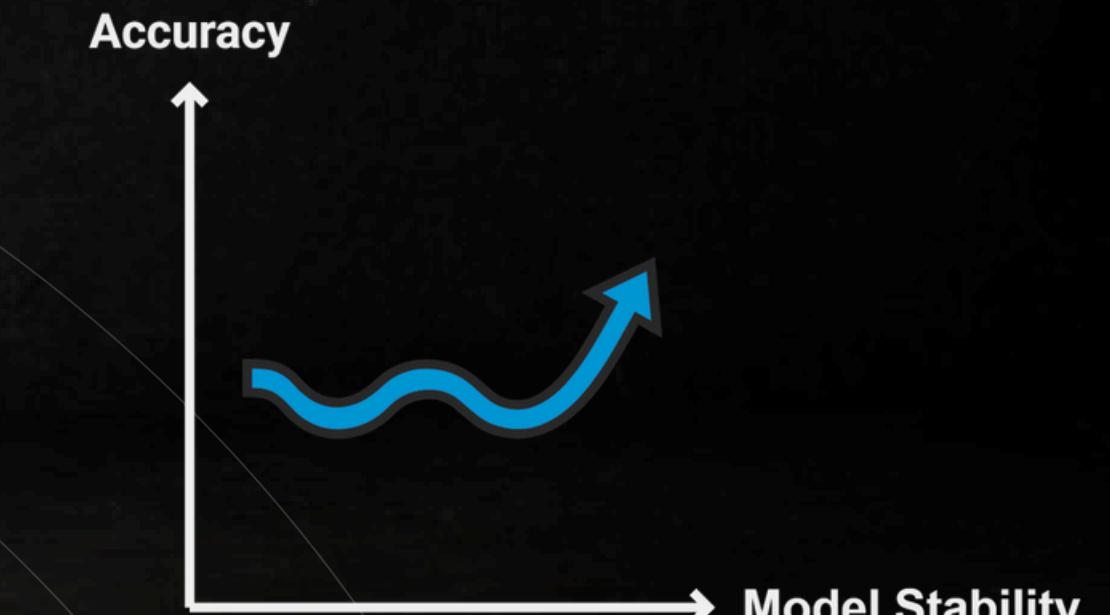
RMSE (Lakhs)
43.40 Lakhs
52.58 Lakhs
54.56 Lakhs



Key Insights

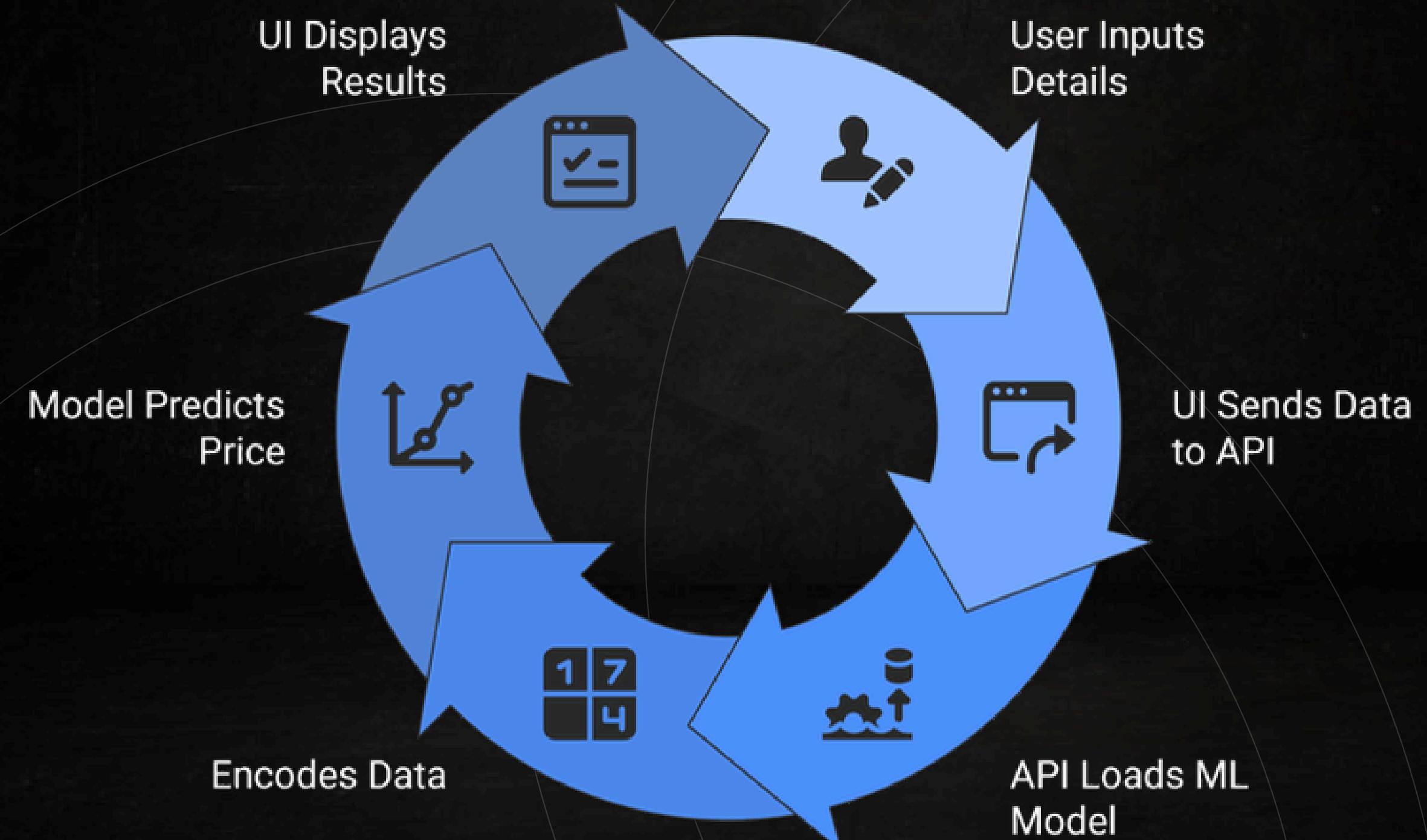


1. Linear Regression is the Most Reliable Model
2. Complex Models Did Not Perform Better
3. Location Plays a Bigger Role Than Model Complexity
4. Linear Regression is ideal for deployment - fast, stable, interpretable, and lightweight.
5. High RMSE Values Reflect Real-World Market Volatility



Accuracy increases with model stability

System Workflow



Unique Features & Limitations

Unique Features

1. AI-Based Price Prediction

- Fast and accurate valuation using ML.



2. Fair-Price Meter

- Shows if a property is underpriced, fair, or overpriced.

3. Market Value Comparison

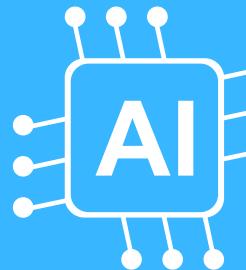
- Uses median, low & high locality prices.

4. Multi-Location Comparison

- Compare 2–5 localities side-by-side.

5. Visual Analytics

- Graphs for price trends across locations.



Limitations

1. Only Bangalore dataset

- Limited location coverage.



2. No Amenity Details

- Age, amenities, and building type not included.

3. Sparse Location Data

- Some localities have too few samples.

4. Complex Models Underperform

- Tree-based models overfit due to data size.

5. No Live Market Integration

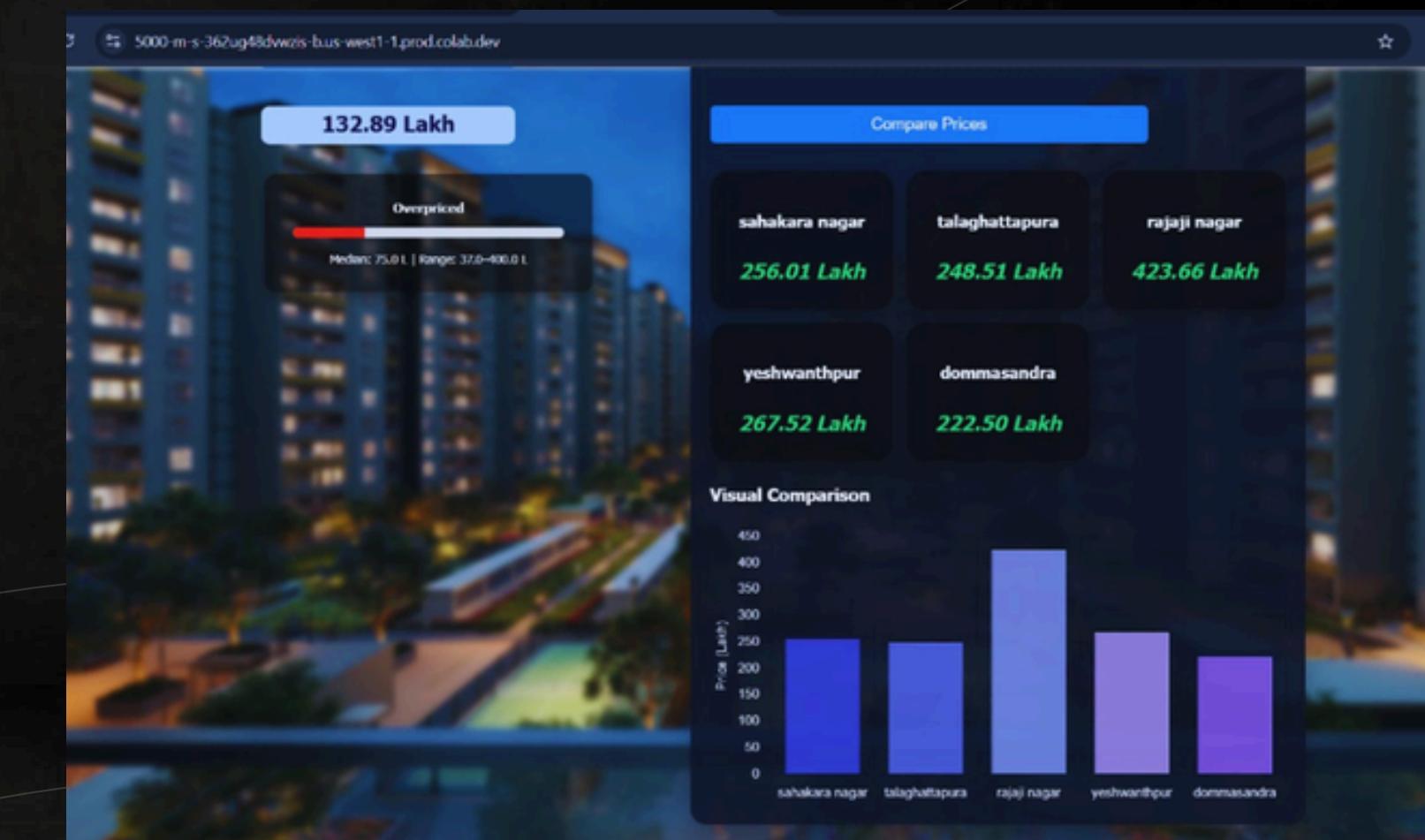
- Prices not updated in real time yet.



Application Demo



The screenshot shows the application's main interface. On the left, there are input fields for "Area (Square Feet)" (2000), "BHK" (4), and "Bath" (3). Below these are dropdown menus for "Location" (whitefield) and "Choose 2 to 5 Locations" (sahakara nagar, talaghattapura, rajaji nagar, yeshwanthpur, dommasandra). A blue button labeled "Estimate Price" is present. The estimated price is displayed as **132.89 Lakh**, with a red bar indicating it is "Overpriced". A note below states "Median: 75.0 L | Range: 37.0-400.0 L". On the right, a "Compare Prices" section shows prices for five locations: sahakara nagar (256.01 Lakh), talaghattapura (248.51 Lakh), rajaji nagar (423.66 Lakh), yeshwanthpur (267.52 Lakh), and dommasandra (222.50 Lakh).



FUTURE SCOPE

Advanced
Machine Learning
Models

Multi-City &
Nationwide
Expansion

Live Market Data
Integration

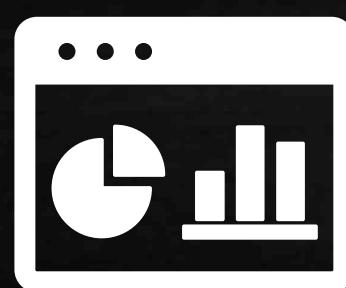
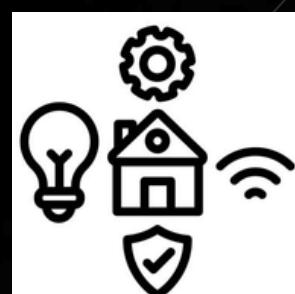
Geospatial
Intelligence
(GIS Integration)

Amenities &
Property
Attributes

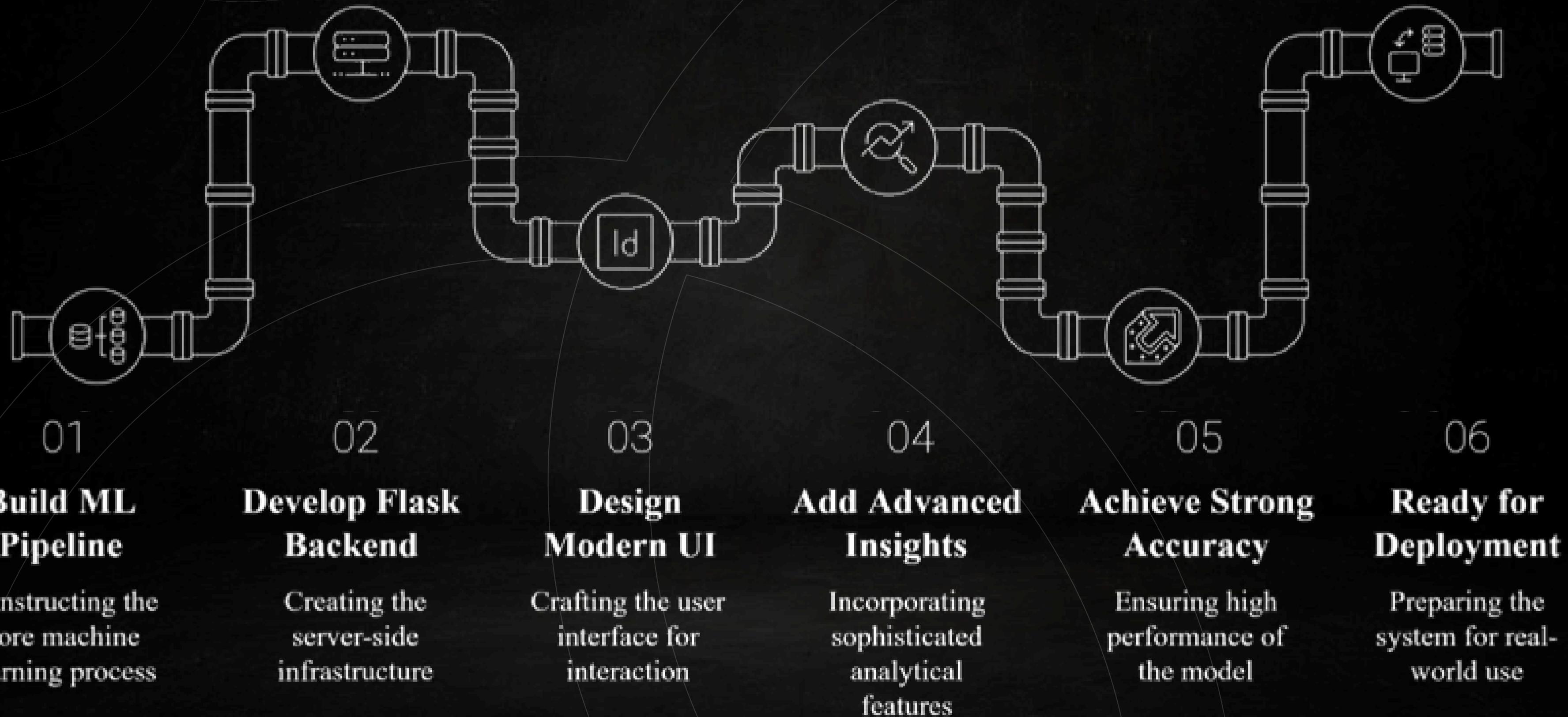
Personalized
User Dashboard

Predictive Market
Analytics

Cloud Deployment
& Scaling



Conclusion



References :

- 1.Kaggle. “Bengaluru House Price Dataset.” Available at: <https://www.kaggle.com/amitabhajoy/bengaluru-house-price-data>
- 2.Kaggle Notebook — Bengaluru House Price Analysis by Amit Abhijoy
- 3.Scikit-Learn Documentation. Available at: <https://scikit-learn.org>
- 4.Flask Web Framework Documentation. Available at: <https://flask.palletsprojects.com>
- 5.Kumar, R., & Singh, A. (2021). “House Price Prediction Using Machine Learning Algorithms.” International Journal of Computer Applications.
- 6.Pagi, V., & Sharma, P. (2020). “Real Estate Price Prediction using AI Techniques.” IJERT.

Thank You