



**Infosys Springboard Vitual Internship**  
**Batch -8**

**PropertyA**

**Mentor : Durga Prasad**  
**Coordinator : Belovedinla**

# Problem Statement : Global Valuation Complexity

Real estate price estimation is often inconsistent and unreliable due to fragmented data, market variability, and geographic differences across regions.

## Global Valuation Complexity



### Heterogeneous Property Data



→ Fragmented & Non-Standardized Property Data



### Cross-Country Market Variability



→ High Cross-Market Price Volatility



### Location-Specific Price Drivers



→ Strong Location-Dependent Price Sensitivity



### Manual & Subjective Evaluation



→ Manual, Time-Intensive Valuation Processes



### Lack of Scalable Valuation Tools



→ Limited Scalable & Automated Valuation Systems

# Key Features: The PropertyAI



## Instant Valuations

Instant property valuations with precision confidence scoring



## Analytics Dashboard

Comprehensive market analytics with Chart.js-powered visual intelligence



## Global Heatmap

Leaflet.js price density mapping (50+ cities), across worldwide markets.

**AI Chatbot Assistant** - Gemini-powered natural language queries.

**Blog & Market Insights** - Real estate articles + AI-generated trends.

**Prediction History & Audit Trail** - Complete valuation archive with timestamps

**Role-Based Access Control** - Secure user/admin authentication system

# Model training



## Model

- Stacking Ensemble Regressor
- Tuning: RandomizedSearchCV
- Meta-Learner: Ridge Regression
- Gradient Boosting Regressor



## Accuracy

- Test R<sup>2</sup> Score: 0.737
- Mean Absolute Error: ~\$46K
- Variance Explained: ~74%
- RMSE: ~0.41



## Tech Stack

- Programming: Python
- ML Libraries: scikit-learn
- Boosting: XGBoost, LightGBM
- Data Tools: pandas, numpy

# Challenges Faced



## Data Challenges

- Large multi-country property datasets
- Handling missing values and outliers
- Feature standardization across regions



## Model Challenges

- Feature engineering from limited attributes
- Selecting optimal models for ensemble learning
- Preventing overfitting and underfitting



## System Challenges

- ML model & Flask API integration
- Performance and response-time optimization
- Secure authentication and session handling

Addressing these challenges improved **model reliability** and **system scalability**

# FUTURE SCOPE



## Architectural & Model Enhancements

### 1. Real-time Data Scraping

Enables continuous model updates to reflect current market conditions.

### 2. Deep Learning Integration

Helps capture temporal trends and spatial relationships between neighborhoods.



## Feature Expansion (User Experience)

### 1. Multi-Lingual & Voice-Enabled Interaction

Regional language support with voice-based property queries powered by NLP.

### 2. Adaptive Model Retraining

Ensures long-term accuracy for consistent and reliable user results over time.



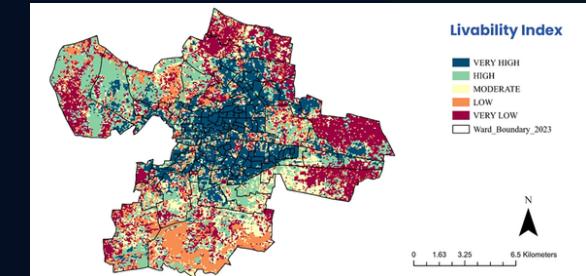
## Ecology & Biodiversity Mapping

### 1. Green Canopy Cover

Use satellite imagery to calculate the percentage of tree cover within a 1km radius.

### 2. Air Quality Index (AQI) History

Show a 12-month average of air quality, not just today's data, to help buyers with respiratory concerns.



# Conclusion

This AI-powered web application uses Random Forest and XGBoost to deliver accurate real estate valuations across Indian and global markets. By blending sophisticated machine learning with localized trends and multi-currency support, the system reduces information gaps and provides a scalable, data-driven solution for modern property appraisal.

