Satellite Imagery Based Property Valuation

Enrolment Number: 24115004

Abstract

This project uses satellite imagery and structured real estate attributes to predict property prices using a multimodal machine learning approach. Traditional valuation models rely completely on tabular data, which limits their ability to capture neighbourhood and environmental context. To overcome this limitation, deep feature extraction is used to incorporate satellite images, which are then fused with tabular features using a neural network-based fusion model. The suggested system shows how visual cues like surrounding infrastructure, road connectivity and greenery can increase the accuracy of property valuation.

1. Introduction

Property valuation is a critical task in real estate analytics, urban planning, and financial decision-making. Conventional approaches primarily depend on numerical attributes such as property size, number of rooms and condition. However, these features fail to capture contextual factors like neighbourhood quality and the surrounding infrastructure.

Satellite imagery provides a rich source of spatial and environmental information. This project aims to use satellite images and structured property data to build a multimodal price prediction system.

Objectives:

- Predict property prices using tabular and satellite image data.
- Build a multimodal fusion model.
- Improve interpretability using Grad-CAM visualizations.

2. Dataset Description

2.1 Tabular Dataset

The tabular dataset contains real-estate attributes including:

- Bedrooms
- Bathrooms
- Living area (sqft.)
- Lot size
- Condition
- Grade
- View
- Waterfront indicator

These features represent intrinsic property characteristics.

2.2 Satellite Imagery Dataset

Satellite images are fetched using the Google Maps Static API based on latitude and longitude coordinates of each property.

Image Specifications:

- Resolution: 224 × 224 pixels
- Map type: Satellite
- Zoom level: 18

Each image corresponds to a single property record.

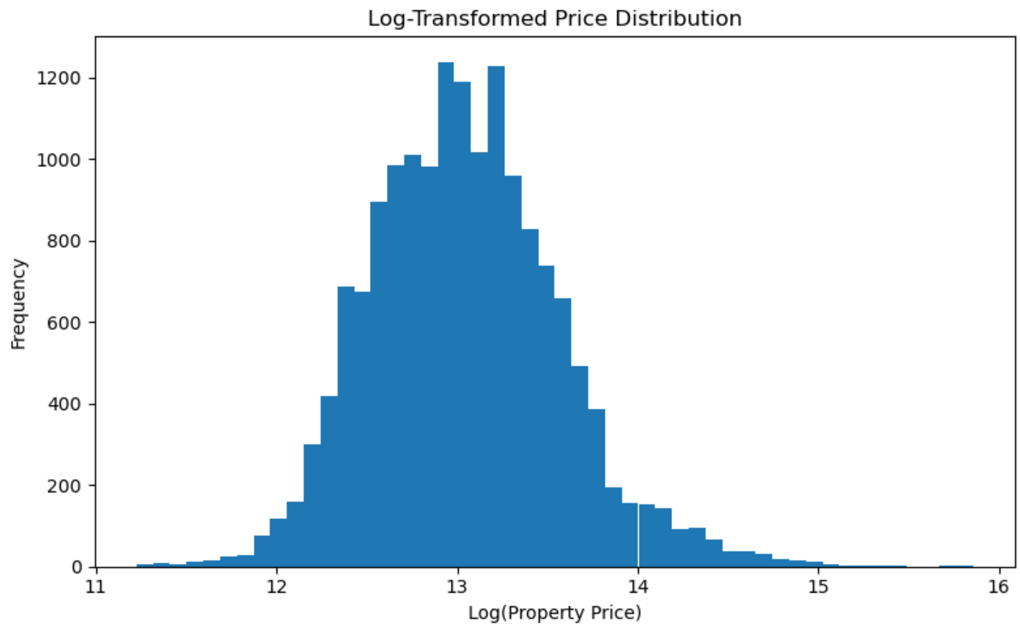
3. Exploratory Data Analysis (EDA)

3.1 Price Distribution

A histogram is used to visualize the distribution of property prices and identify skewness.

Figure 1: Price distribution histogram

(a)The histogram shows a heavily right-skewed distribution. This indicates that most properties fall in lower to mid-price ranges, with a small number of high-value outliers. This justifies the use of log transformation during model training.

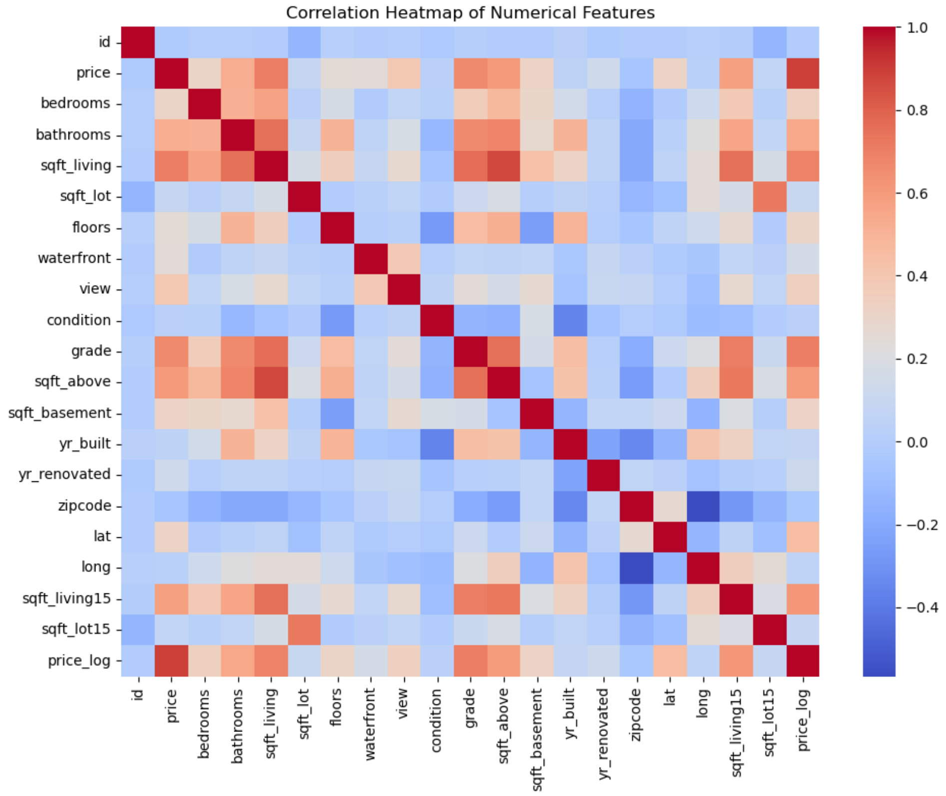


(b)The log transformation significantly reduces the right skew observed in the original price distribution, resulting in a more symmetric, approximately normal distribution. This transformation helps stabilize variance and improves model training for regression tasks.

3.2 <u>Feature Correlation Analysis</u>

A correlation heatmap is generated to study relationships between numerical features and property prices.

Figure 2: Correlation heatmap of tabular features



3.3 Feature vs Price Analysis

Scatter plots are used to analyse relationships such as:

- Living area vs price
- Number of bedrooms vs price etc.

Figure 3: Living area vs price

Figure 4: bedrooms vs price

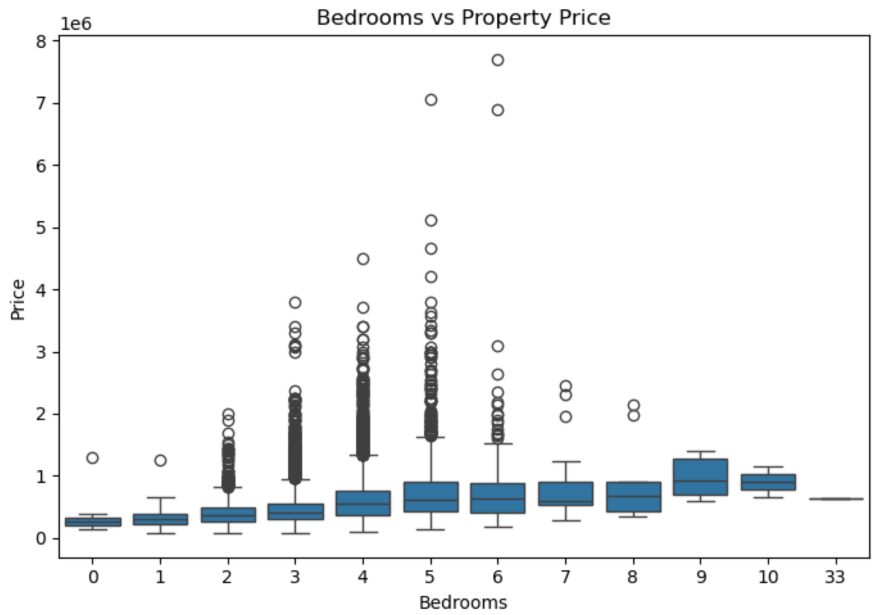


Figure 5: Bathrooms vs Price

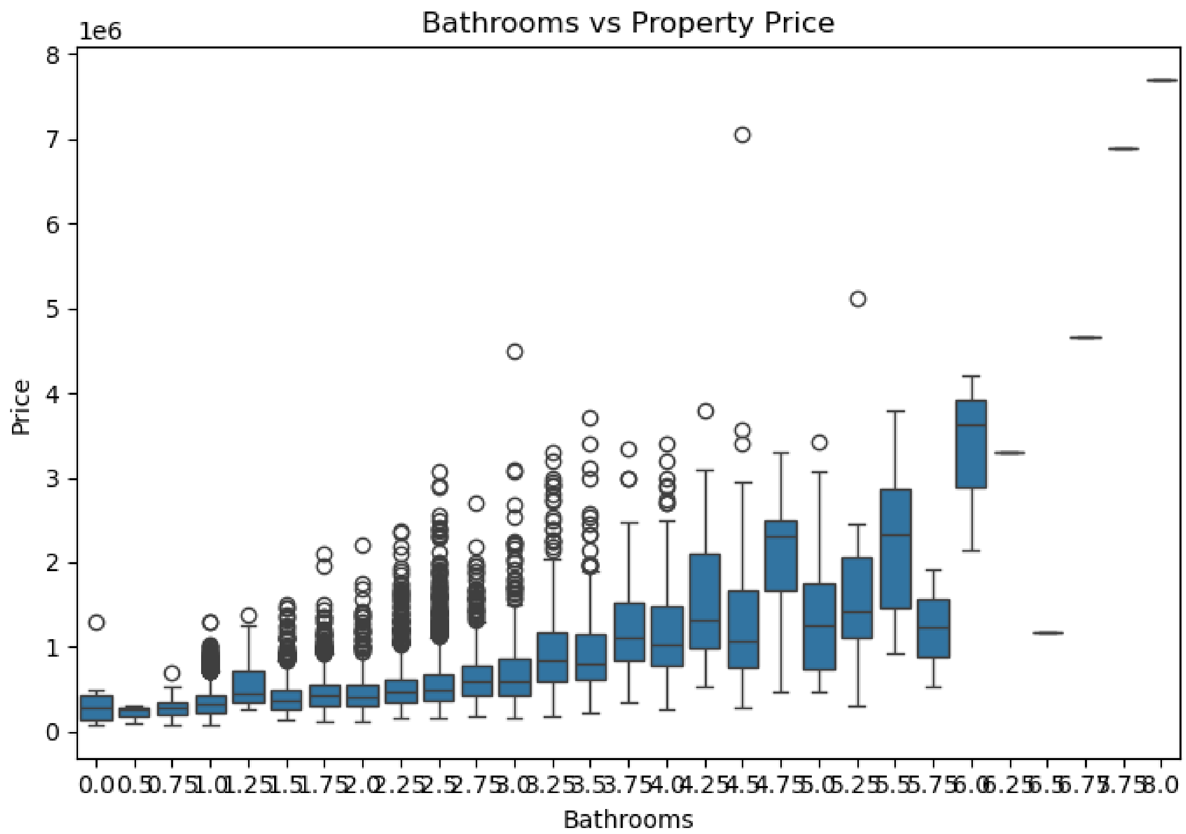
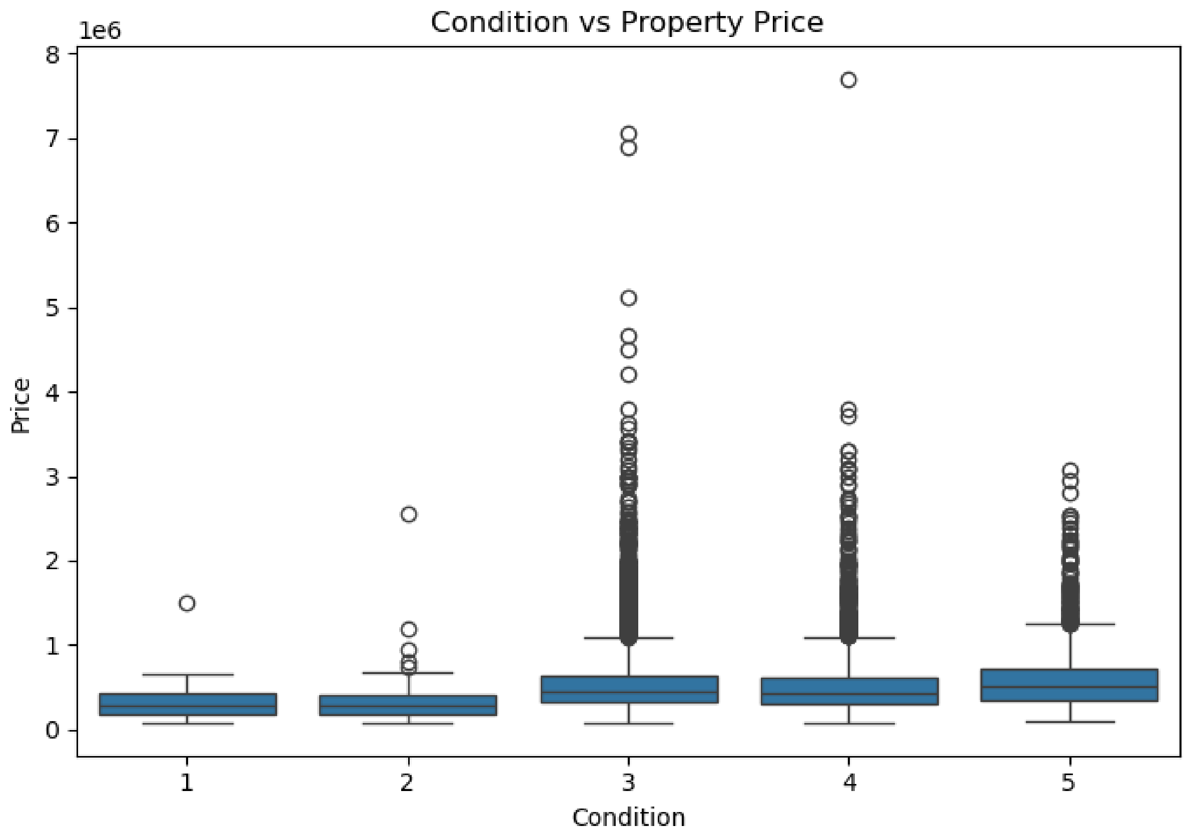


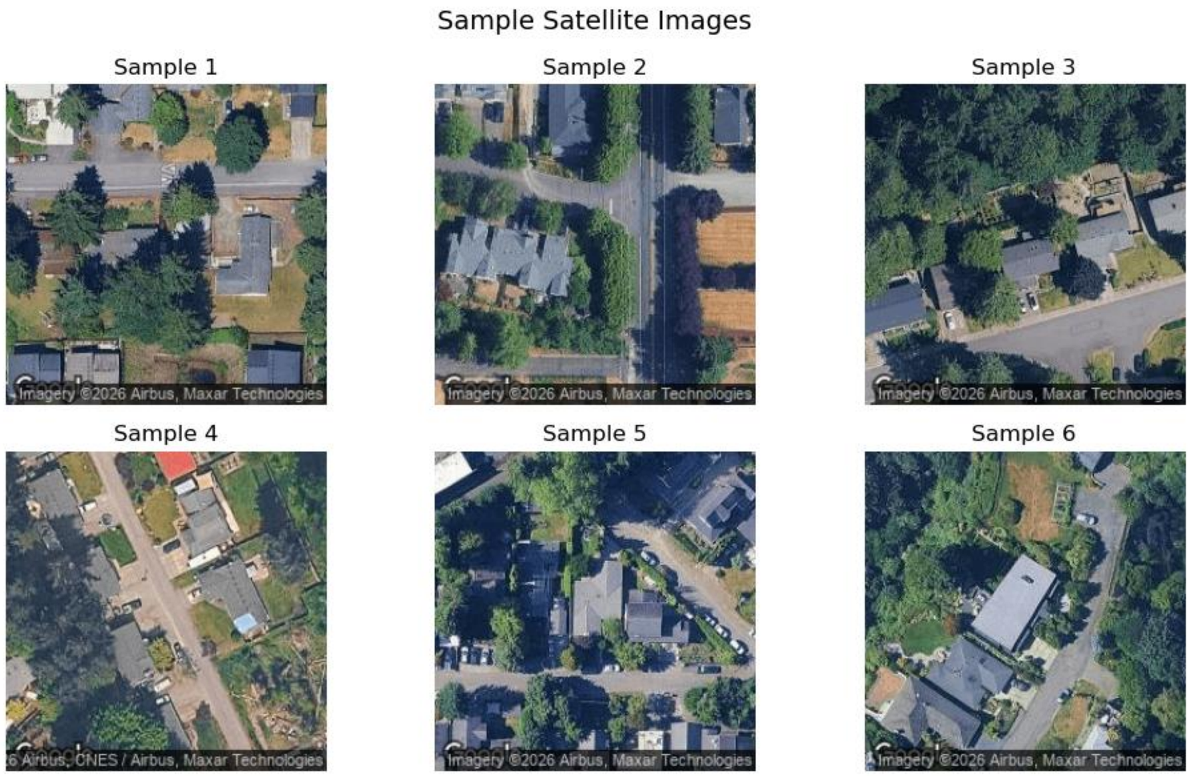
Figure 6: Condition vs price



3.4 Satellite Image Samples

Sample satellite images from the dataset are visualized to understand neighbourhood characteristics.

Figure 7: Sample satellite images



4. Data Collection and Preprocessing

4.1 Satellite Image Fetching

Satellite images are downloaded using data_fetcher.py, which:

- Reads latitude and longitude from the dataset.
- Calls Google Maps Static API.
- Stores images locally for processing.

4.2 Tabular Data Preprocessing

- Feature selection.
- Handling missing values.
- Feature scaling using StandardScaler.

The trained scaler is saved and reused during inference to maintain consistency.

4.3 Image Feature Extraction

Satellite images are passed through a CNN-based feature extractor to obtain fixed-length image embeddings. These embeddings are stored for reuse during training and inference.

5. Methodology

5.1 Baseline Tabular Model

A Random Forest regression model is trained using only tabular features to serve as a baseline for comparison.

5.2 Multimodal Fusion Model

The final model combines:

- Tabular feature embeddings.
- Image feature embeddings.

Both modalities are processed separately and fused before final prediction.

```python
print("========== FINAL MODEL COMPARISON ==========")
print(f"Tabular RF   → RMSE: {rmse_rf:.3f}, R²: {r2_rf:.3f}")
print(f"Multimodal   → RMSE: {rmse:.3f}, R²: {r2:.3f}")
```

[18] Python

```
========== FINAL MODEL COMPARISON ==========
Tabular RF   → RMSE: 0.328, R²: 0.548
Multimodal   → RMSE: 2.021, R²: -16.170
```

Reasonable predictive performance was attained by a baseline Random Forest model that was only trained on tabular features. In terms of RMSE and R2 on the evaluation set, the multimodal fusion model, which uses satellite image embeddings did not perform better than the tabular baseline. The multimodal model had trouble generalizing under the current training setup, as indicated by the negative R2 value.

This outcome emphasizes the difficulties in multimodal learning, especially when the dataset size is small or the visual features are noisy. Nevertheless, the multimodal framework offers a solid basis for further advancements and offers insightful information about incorporating spatial context. These findings suggest that additional data, improved image representations or more advanced fusion strategies are required to fully exploit satellite imagery for property valuation.

6. Model Architecture

6.1 Tabular Branch

- Fully connected layer
- Batch normalization
- ReLU activation
- Dropout

6.2 Image Branch

- Fully connected layer
- Batch normalization

- ReLU activation
- Dropout

6.3 Fusion Layer

- Concatenation of tabular and image features
- Fully connected layers
- Regression output (log-scaled price)

Figure 5: Multimodal fusion architecture diagram

7. Model Explainability

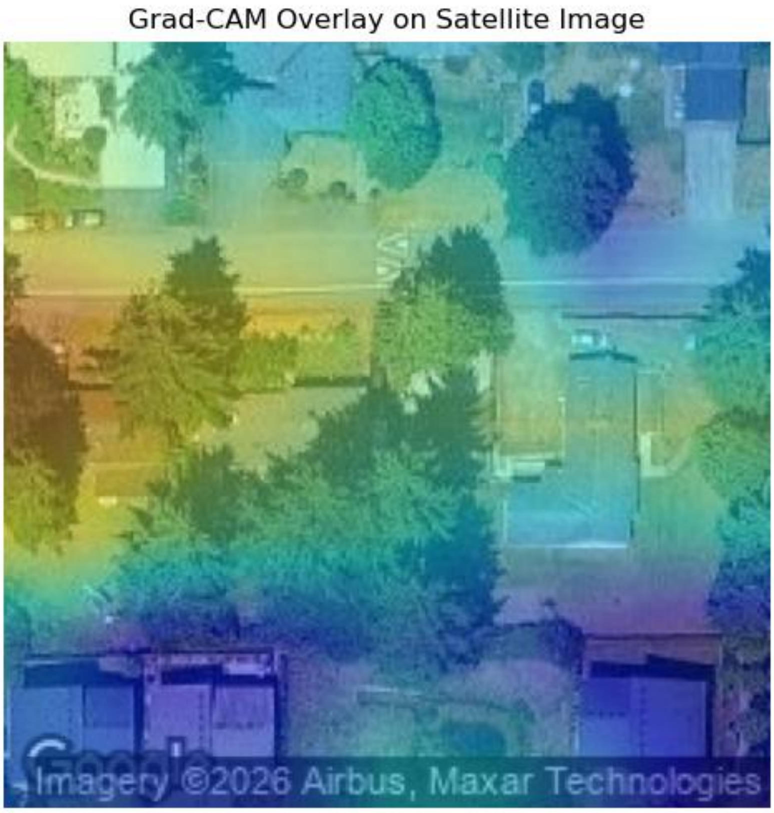
7.1 Grad-CAM Analysis

Grad-CAM is applied to satellite images to visualize which regions influence price predictions.

Highlighted regions often correspond to:

- Green areas
- Road connectivity
- Surrounding infrastructure

Figure 6: Grad-CAM overlay on satellite image



8. Results

8.1 Prediction Output

The final predictions are stored in:

24115004_final.csv

Format:

id, predicted_price

Figure 7: Sample rows from prediction CSV

id	predicted_price
0	238630.34
1	508166.78
2	411322.97
3	1424985.6
4	50687.875
5	197804.17
6	39716.992

9. Conclusion

This project demonstrates the effectiveness of multimodal learning for property valuation. By integrating satellite imagery with structured data, the model captures both intrinsic and contextual information, leading to more informed predictions. The inclusion of explainability techniques further enhances trust in the model's outputs.

10. Future Work

- Use higher-resolution satellite imagery
- Incorporate street-view images
- Experiment with advanced CNN architectures
- Extend the model to classification-based valuation ranges