Satellite Imagery-Based Property Valuation

Dakshit Tanay (24114030)

2nd year, B. Tech Cse

1. Overview (Approach & Modelling Strategy)

Problem Statement

Property valuation depends not only on a house's internal characteristics but also on its **surrounding environment**. While traditional regression models use structured attributes such as size and quality, they often ignore **neighbourhood-level visual context**, which can influence perceived value.

Proposed Approach

This project develops a **multimodal regression pipeline** that integrates:

- **Tabular housing attributes** (e.g., bedrooms, bathrooms, square footage, grade)
- **Satellite imagery** capturing neighbourhood context using geographic coordinates

Satellite images are programmatically fetched using the **ESRI World Imagery service** via the ArcGIS REST API:

ESRI World Imagery API
https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer/export

Modelling Strategy

The modelling pipeline consists of:

1. **Baseline Model (Tabular Only)**
   - XGBoost regressor trained on structured housing features
   - Serves as a strong reference model
2. **Visual Feature Engineering**
   - Satellite images extracted using latitude and longitude
   - Images processed through a pretrained **ResNet-18 CNN**
   - CNN used strictly as a **feature extractor**, producing fixed-length embeddings
3. **Multimodal Learning**
   - CNN-derived embeddings combined with tabular features
   - XGBoost used as the final regression model due to its robustness on structured data.

2. EDA and Financial/Visual Insight

(i) Price Distribution

The raw house price distribution is highly **right-skewed**, with a small number of very expensive properties. To stabilize variance and reduce the influence of outliers, a **log transformation** is applied:

$$\text{log\_price} = \log(\text{price} + 1)$$

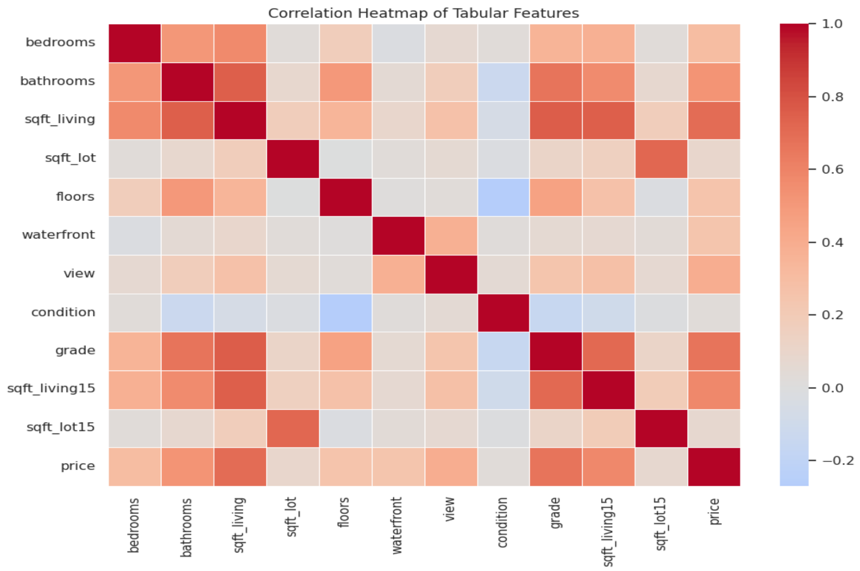
This transformation produces a more symmetric distribution suitable for regression modelling.

(ii) Correlation & Heatmap Analysis

The correlation heatmap of tabular features reveals that **house price is most strongly influenced by:**

- **grade** (construction quality and architectural design)
- **sqft_living** (interior living area)

These two variables exhibit the highest positive correlation with price, indicating that **structural quality and size dominate valuation**.



(iii) Key Visual Insights

- Property prices exhibit clear spatial clustering across geographic coordinates.(as we see gradual changes in the price as we move)
- CNN attention focuses on neighbourhood layout rather than individual houses.
- Lower surrounding house density and open spaces are associated with higher influence whereas middle of a fields have very less affect on the prediction. In a nutshell, neighbourhood to spaces such as parks, fields affect the price most as compared to the field itself.
- Satellite imagery contributes contextual environmental cues to the multimodal model

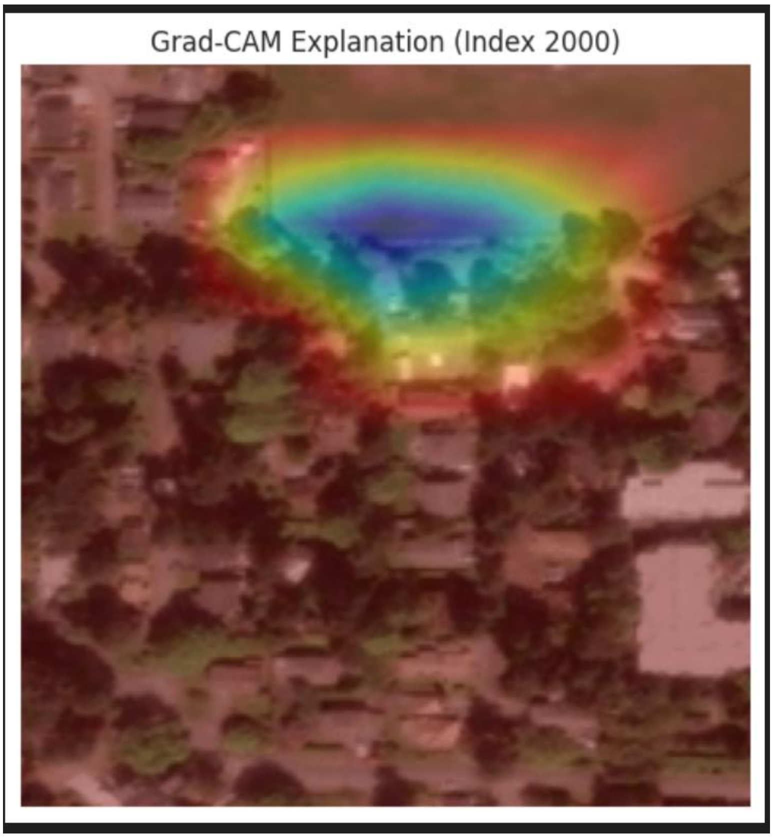


Fig : Grade-Cam Explainability (Red > Yellow>Green > Blue(order of importance))

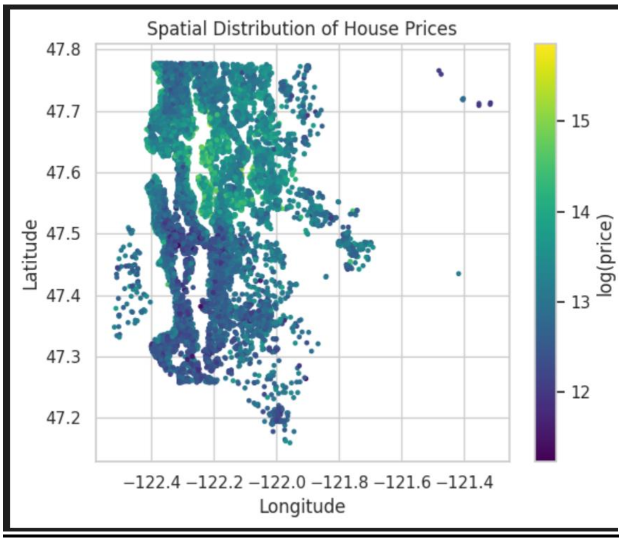


Fig : Spatial Clustering of prices

3. Tried Architecture: Random Forest Regression

- A **Random Forest regressor** was evaluated as an alternative multimodal model using concatenated tabular features and CNN-derived satellite embeddings.
- The model achieved **an R^2 score of 0.7273 (log scale)**, which is lower than the XGBoost-based multimodal model.
- Random Forest constructs **trees independently using bagging**, whereas XGBoost **uses sequential boosting**, enabling more effective error correction across trees.
- The presence of **high-dimensional and noisy CNN embeddings** negatively impacted Random Forest performance, as it lacks gradient-based optimization to selectively down-weight weak visual features.
- While Random Forest performs well on purely tabular data, it is less effective in multimodal settings where feature importance varies significantly across modalities.
- Based on comparative performance and stability, X**GBoost was selected as the final regression model** for multimodal price prediction.

4. Final Architecture Diagram (Multimodal Pipeline)

```
Tabular Features

(bedrooms, sqft_living, grade, ...)

     |

     |---------+-------Images (ESRI World Imagery API) ---------|

     |                                                        |

     v                                                        v

XGBoost (Baseline Model) |                    Pretrained CNN (ResNet-18)

                                                              |

                                                              v

                                    Image Embeddings (CNN) + Tabular Features

                                                              |

                                                              v

                                           XGBoost (Multimodal Regression)

                                                              |

                                                              v

                                                 Final Price Prediction
```

Explanation:

- CNN extracts visual representations from satellite imagery.
- Embeddings are fused with tabular features.
- XGBoost performs final regression using both modalities.

5. Results & Performance Comparison

Evaluation Metrics

- R^2 Score
- RMSE (Root Mean Squared Error)

All models are evaluated on log-transformed prices, with RMSE additionally reported on the original price scale.

Model Performance

Model	**Modalities Used**	**R^2 (log scale)**	**RMSE (price scale)**
XGBoost Baseline	Tabular only	**0.6993**	**`187,610.99**
Multimodal XGBoost	Tabular + Satellite Images	**0.7663**	**169,231.18**