Satellite Based Property Valuation

Ankit Kumar

Enrol. No. : 241170133

B.tech DSAI

IIT ROORKEE

Data Collection and Satellite Image Downloading

To enrich the dataset with visual context, satellite images were programmatically fetched using the latitude and longitude of each property. A satellite imagery API (from maptiler.com) was used to download top-view images centred around the given coordinates. These images capture neighbourhood-level features such as road networks, greenery, and surrounding infrastructure.

For each property:

- The latitude and longitude were passed to the API.
- Different Zoom parameters were tested and zoom=17 was finalised.
- A satellite image was downloaded in .tif format.
- The image was saved locally using the property id as the filename to maintain alignment with the tabular data.

Speeding Up and Progress Recovery for Image download :

- Satellite images were saved immediately after each successful download, ensuring that already downloaded images were not lost if the process was interrupted.
- Before downloading an image, the local image directory was checked. If an image with the corresponding property id already existed, the download step was skipped. This enabled safe resumption across multiple runs.
- To speed up the image download process, the dataset was split into two non-overlapping parts.
- Two separate VS Code windows were opened, each running the image download script on a different subset of the dataset using different API keys.
- Since each process operated on a distinct set of property IDs and saved images using unique filenames, there were no file conflicts between the parallel runs.
- This parallel execution effectively halved the total download time while remaining within API rate limits.

Exploratory Data Analysis (EDA)

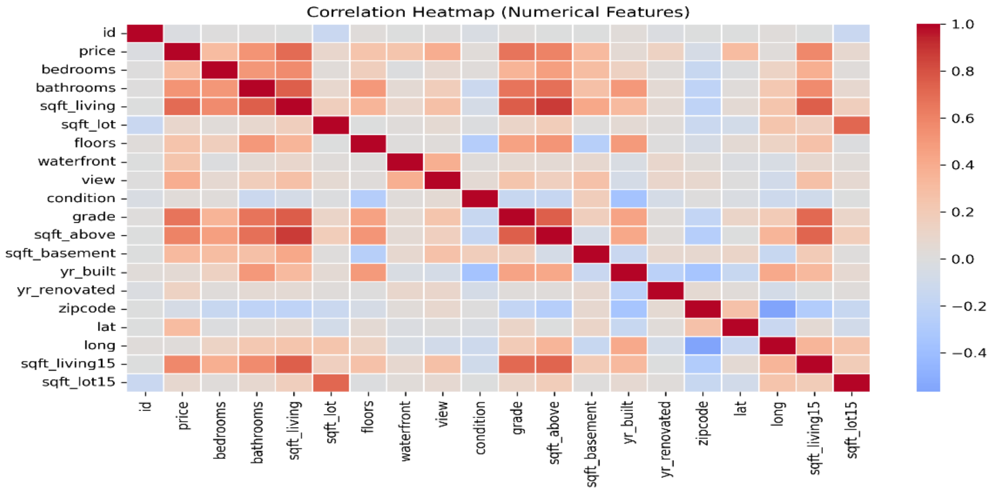
EDA was performed to understand the dataset, analyse feature distributions, and guide preprocessing decisions. First, a general insight of the dataset was taken using some common functions of Pandas library (like: .head(), .info(), .describe(), etc). After this, detailed plots were generated using Matplotlib as shown below:

1. Target Variable Distribution

- The distribution of property price was analysed to study spread, skewness, and outliers.

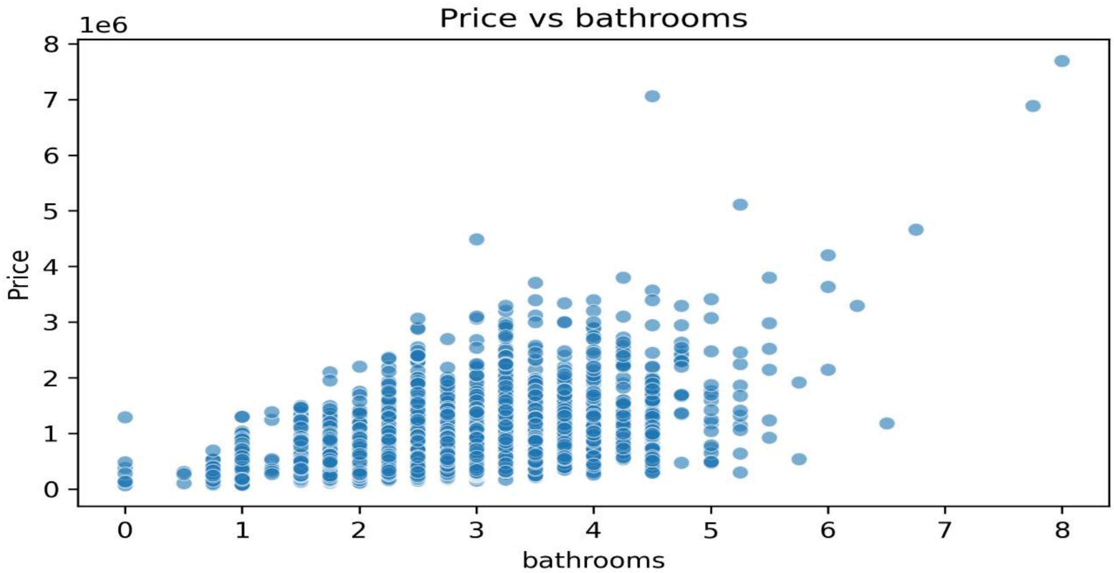
2. Correlation Analysis

- Correlations among numerical features and property price were examined to identify influential variables.



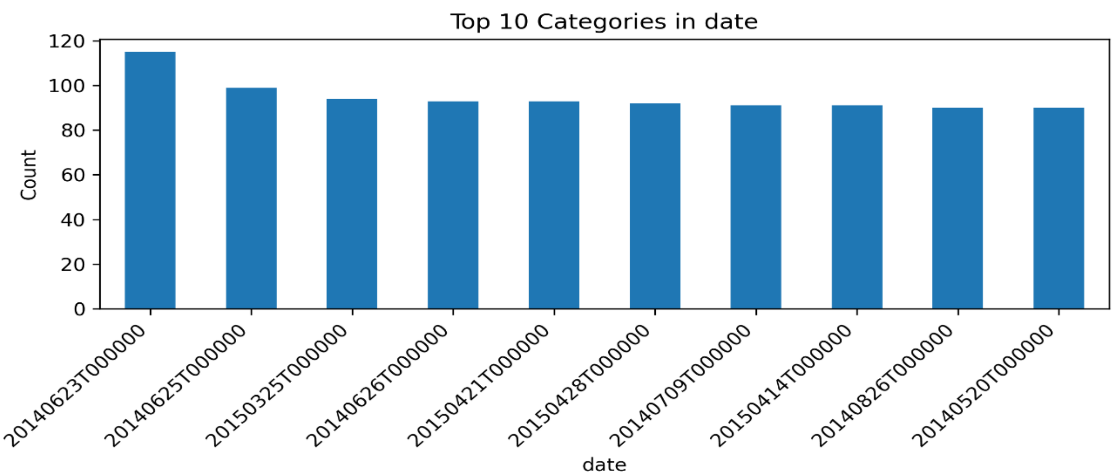
3. Price vs Numerical Feature

- Scatter plots were used to observe the relationship between property price and key numerical features. Here, I have shown only "Bathrooms" feature as there are too many numerical features.



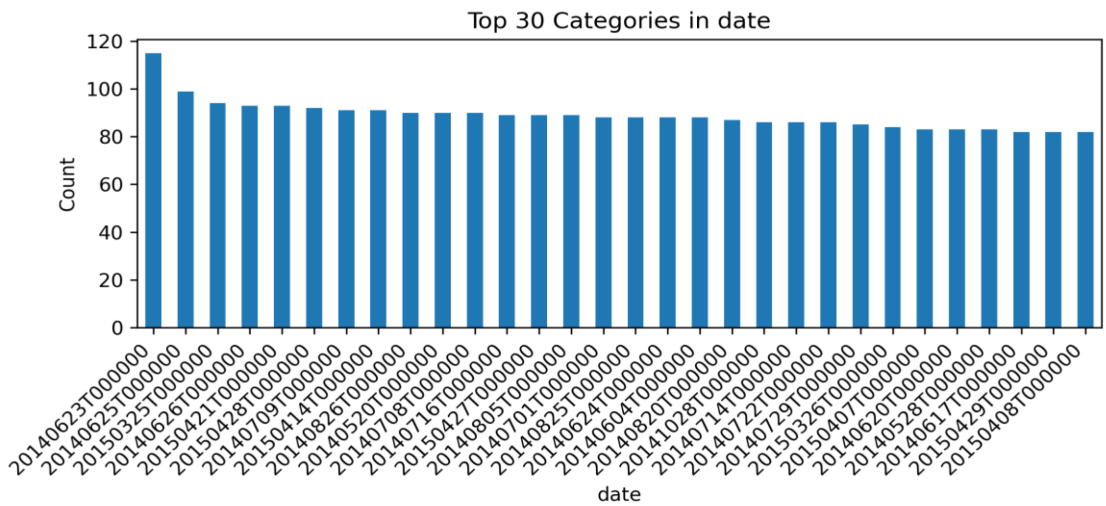
4. Categorical Feature Distribution

- The frequency distribution of categorical variables was visualized to detect imbalance.



5. Price Variation Across Categories

- Average property price was analysed across different categories.



6. Missing and Duplicate Value Analysis

- Missing values were examined to determine suitable imputation strategies.

EDA Summary

- Property prices show skewness and outliers
- Some numerical features correlate with price
- Categorical feature, date, gives little to no information so dropping it may be a good option.
- No missing values are found and there were some duplicate IDs (which were dropped in the pre-processing stage).

Image Preprocessing and Feature Extraction

Satellite images were transformed into numerical features using a **pretrained ResNet-18 Convolutional Neural Network**.

- The final classification layer was removed, and the network was used purely as a feature extractor and all CNN parameters were frozen to prevent training and reduce computational cost.
- Images were resized to 224 × 224 pixels and normalized using ImageNet mean and standard deviation.
- Extracted image features were cached to disk as "X_image.npy" to avoid repeated computation in subsequent runs.

Tabular Data Preprocessing :

The target variable for the regression task was property price, while all remaining columns were treated as input features.

- Numerical features
  - No missing values were found.
  - Duplicate IDs were dropped.
  - Features were standardized using z-score normalization
  - Id (Identifier), zipcode (high-cardinality categorical and yr_renowated (sparse & noisy) columns were dropped.
- Categorical features
  - 'Date' column was dropped and accuracy scores improved so I left it dropped.
  - One-hot encoding was applied with safe handling of unseen categories

A ColumnTransformer pipeline was used to apply these transformations in a structured and reproducible manner.

Train–Test Split

The final dataset was split into training and testing sets using an 80–20 split with a fixed random seed, 42. The same split was applied consistently across both tabular features and image embeddings.

Model Training and Evaluation

Multiple models were trained and evaluated to assess the contribution of satellite imagery to property price prediction. Performance was measured using R^2 and RMSE on the held-out test set (20% of the training dataset).

1. Baseline Model (Tabular Data Only)

- A baseline regression model was trained using only tabular property features.
- An XGBoost regressor was combined with the preprocessing pipeline.
- This model serves as a reference to evaluate the impact of image features.

Evaluation Metrics

- R^2 and RMSE were computed on the test set.

2. Image-Only Model

- A separate XGBoost regressor was trained using only CNN-extracted image embeddings.
- This model evaluates how much predictive information is present in satellite images alone.

Evaluation Metrics

- R^2 and RMSE were computed on the test set.

3. Early Fusion Model

- In early fusion, tabular features and image embeddings were concatenated at the feature level.
- Tabular features were pre-processed using the column transformer.
- Image embeddings were standardized and combined with tabular features.
- A single XGBoost regressor was trained on the fused feature space.

Evaluation Metrics

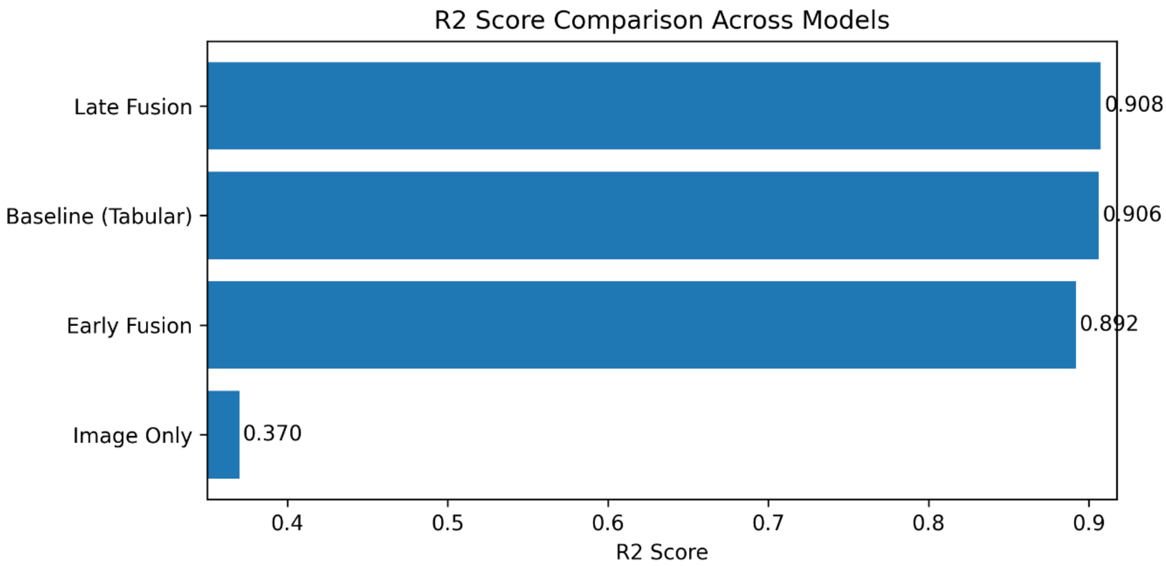
- R^2 and RMSE were computed on the test set.

4. Late Fusion Model

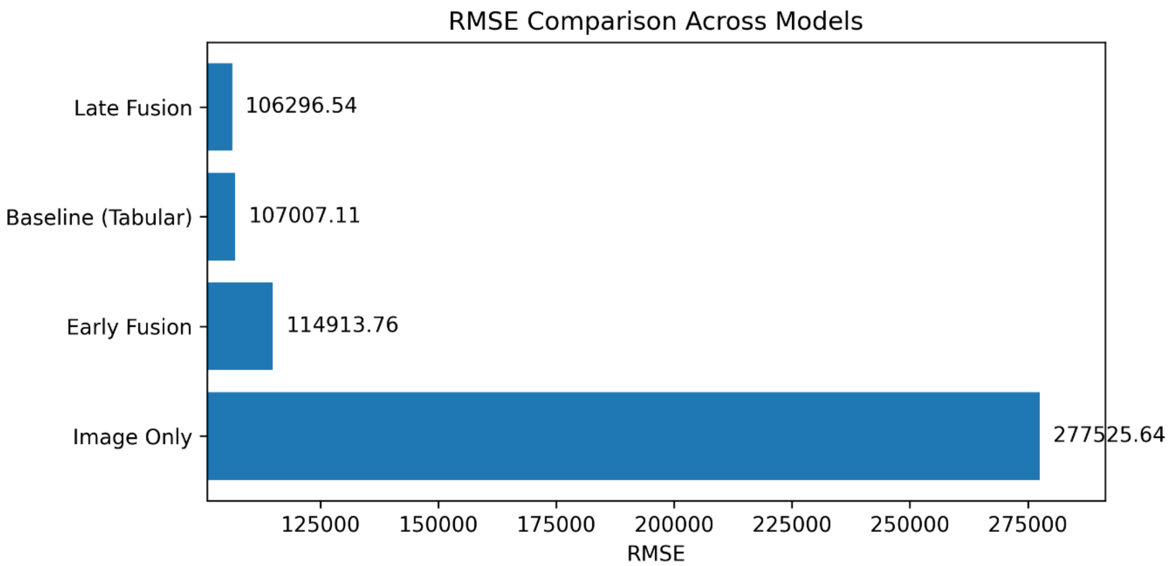
- In late fusion, separate models were trained for tabular and image features.
- Final predictions were obtained using a weighted combination of both model outputs.
- The fusion weight (α) was tuned by minimizing RMSE on the validation set.

5. Model Comparison

To compare all approaches, evaluation metrics from each model were summarized and visualized as shown below:



Comparison of R^2 scores across models



Comparison of RMSE across models

Evaluation Summary

- The baseline model establishes strong tabular performance.
- The image-only model captures useful neighbourhood information.
- Early fusion improves performance by combining structured and visual features.

- Late fusion achieves the best balance by optimally weighting both modalities so it was **Finalised.**

Model Explainability using Grad-CAM

To improve interpretability of the image-based component of the model, Grad-CAM (Gradient-weighted Class Activation Mapping) was used to visualize which regions of satellite images most influenced the model's predictions.

Purpose of Explainability

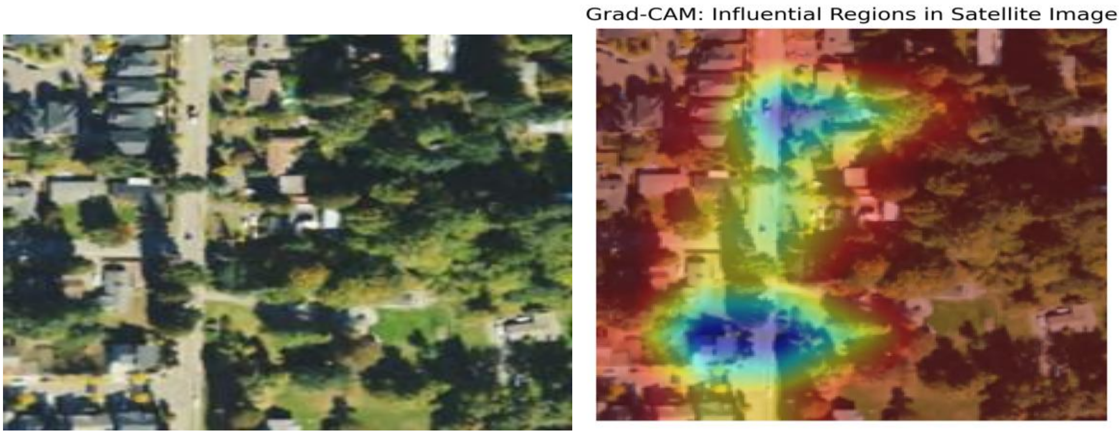
- Deep learning models, especially CNNs, are often treated as black boxes.
- Grad-CAM provides visual explanations by highlighting image regions that contribute most to the model output.
- This helps verify whether the model focuses on meaningful neighbourhood features such as roads, buildings, and open spaces.

Grad-CAM Methodology

- A pretrained ResNet-18 CNN (used earlier for feature extraction) was employed.
- Grad-CAM was applied to the last convolutional layer of the network.
- During backpropagation:
    - Gradients of the output with respect to feature maps were computed.
    - These gradients were spatially averaged to obtain importance weights.
    - A weighted combination of feature maps was used to generate a heatmap.
- The heatmap was normalized and overlaid on the original satellite image.

Visualization Output

The resulting Grad-CAM visualization highlights regions of the satellite image that had the greatest influence on the model's internal representation. Below is one of the satellite images and its highlighted version put side by side for comparison.



Grad-CAM visualization showing influential regions in the satellite image

Observations

- The highlighted regions correspond to meaningful spatial features such as:
    - Road networks
    - Dense built-up areas
    - Surrounding infrastructure
- This indicates that the CNN successfully captures relevant neighbourhood context rather than irrelevant background regions.

Prediction on Unseen Test Data

To evaluate the practical applicability of the proposed multimodal model, predictions were generated for a new unseen dataset provided in test2.xlsx **using the finalised model which s trained on the entire training dataset.**

- Satellite images corresponding to the test properties were downloaded using their geographic coordinates.
- CNN-based image embeddings were extracted using the same preprocessing and feature extraction pipeline used during training.
- Extracted image features were cached to disk as **"X_image_test2.npy"** to avoid repeated computation in subsequent runs.
- The tabular features were transformed using the previously fitted preprocessing pipeline to ensure consistency.
- Final property price predictions were obtained using the **late fusion model**, which combines tabular and image-based predictions using the optimized fusion weight (α).

The predicted prices were saved in a CSV file as **“enrollno_final.csv”** containing the property ID and the corresponding estimated price.