SATELLITE IMAGERY BASED PROPERTY VALUATION

By:
Vaibhavi Shinde
(23119042)

1.Problem Statement

I aimed to improve house price prediction by incorporating **visual neighborhood context** along with traditional tabular housing attributes. While classical models rely heavily on structured features (area, rooms, location codes), they fail to capture **surrounding environment signals** such as greenery, road density, and urban layout.

To address this, I designed a **multimodal learning pipeline** that fuses:

- Tabular housing attributes
- Satellite imagery embeddings
- Transportation accessibility features

The objective was to predict the prices using images and tabular data.

2. Dataset Overview

The dataset consists of residential properties with:

- **Structural attributes:** bedrooms, bathrooms, square footage, floors
- **Quality indicators:** condition, grade, view, waterfront
- **Location attributes:** latitude, longitude, zipcode
- **Temporal information:** Date
- **Target variable:** house price

Each house is uniquely identified using an id.

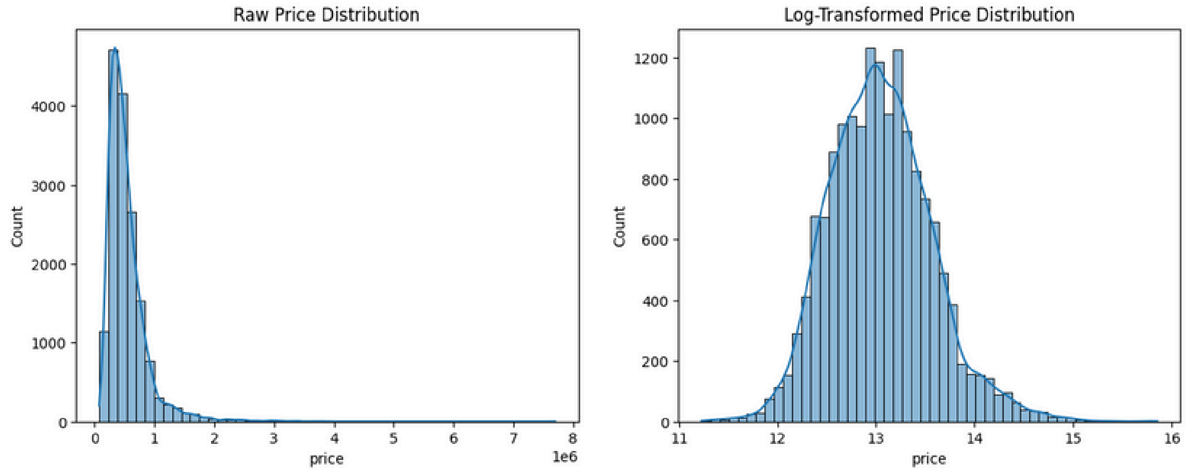
3. Data Cleaning & Preprocessing

I performed the following preprocessing steps:

- Eliminated duplicate property IDs
- Enforced consistent data types (id as string)

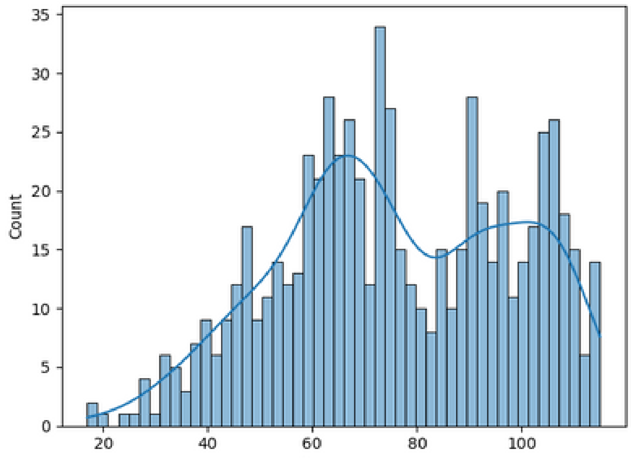
This ensured data integrity across satellite fetching, OSM feature extraction, and model training.

4.EDA

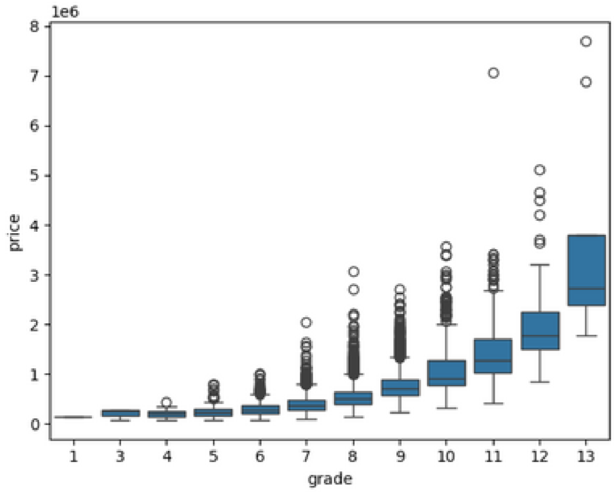


Raw Price Distribution & Log-Transformed Price Distribution

Log transformation reduces skewness and stabilizes variance, making regression more reliable.

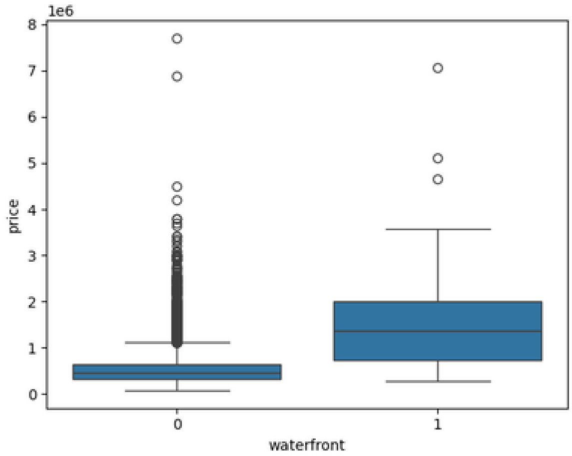


This represents the age distribution of **Renovated** Houses



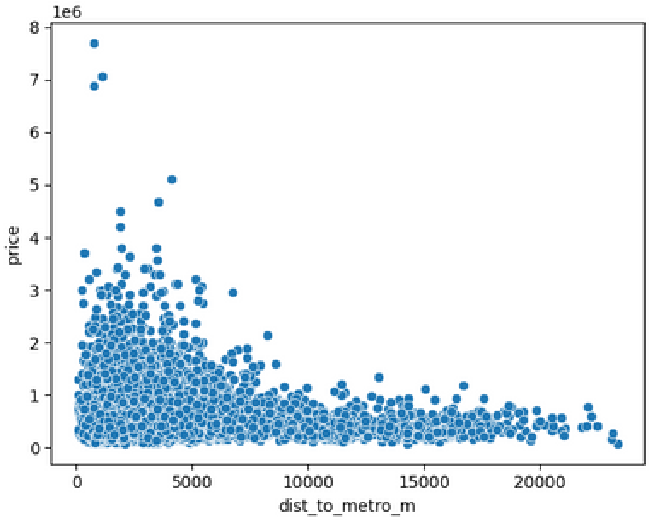
House Price vs. Grade

A strong monotonic increase in median price is observed with higher grades, confirming grade as a highly influential feature in determining property value.

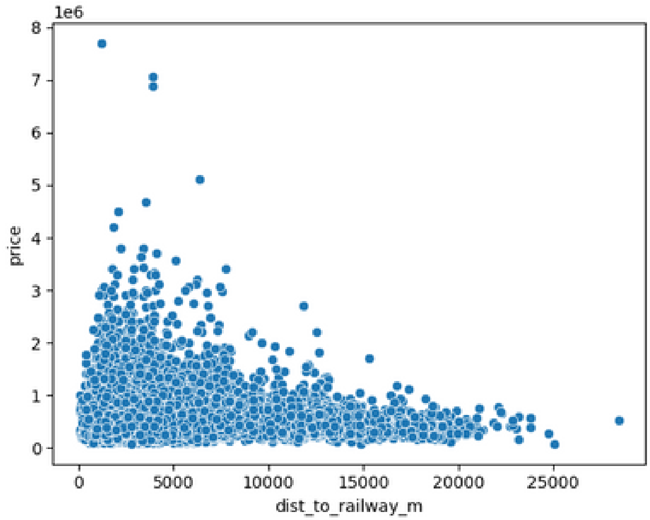


House Price vs. Waterfront

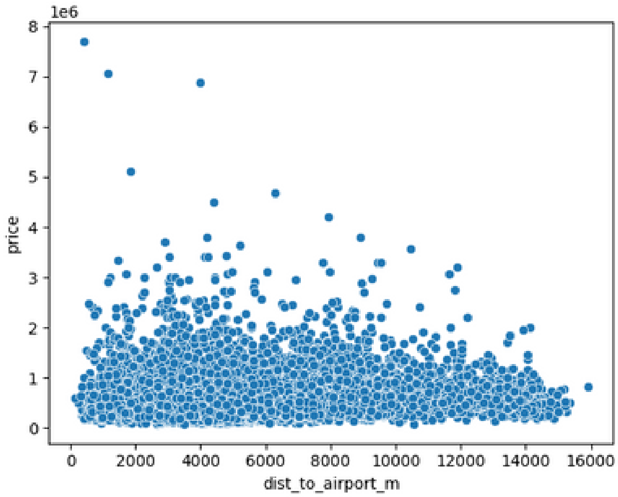
Waterfront homes exhibit significantly higher median prices and greater variance, highlighting the premium associated with waterfront access.



House Price vs. Distance to Metro Station



House Price vs. Distance to Railway Station



House Price vs. Distance to Airport

House prices decrease with increasing distance from major transportation infrastructure (metro, railway, and airport), indicating a clear accessibility premium, with greater variability observed closer to transit hu

5.Satellite Image Collection

For each property, I downloaded high-resolution satellite images using the **Mapbox Static API**.

Configuration:

- Zoom level: **18** (captures neighborhood context)
- Image size: **512×512 (@2x)**
- Map style: **Satellite**
- Rate limiting, retries, and failure logging implemented

Each image was saved using the property ID, ensuring perfect alignment with tabular data.

6. Spatial Infrastructure Feature Engineering

To model accessibility, I extracted **transportation infrastructure** using OpenStreetMap:

- Metro stations
- Railway stations
- Airports

Using OSMnx, I fetched all POIs once and computed **nearest distances** for each house using a **BallTree with Haversine distance.**

Final features included:

- Distance to nearest metro
- Distance to nearest railway station
- Distance to nearest airport
- Log-transformed distances (log1p) to handle skewness

These features capture **connectivity and accessibility effects on pricing.**

7. Spatial Infrastructure Feature Engineering

From the transaction date, I derived:

- Year, month, day
- Day of week
- Weekend indicator

This allows the model to learn **seasonal and temporal pricing patterns**.

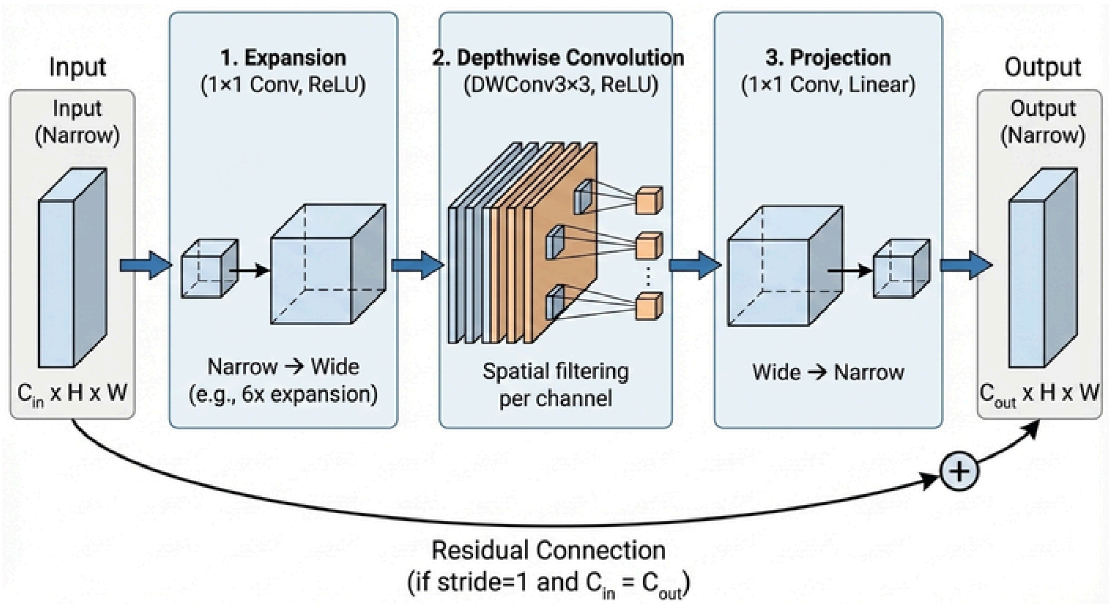
8. Advanced Tabular Feature Engineering

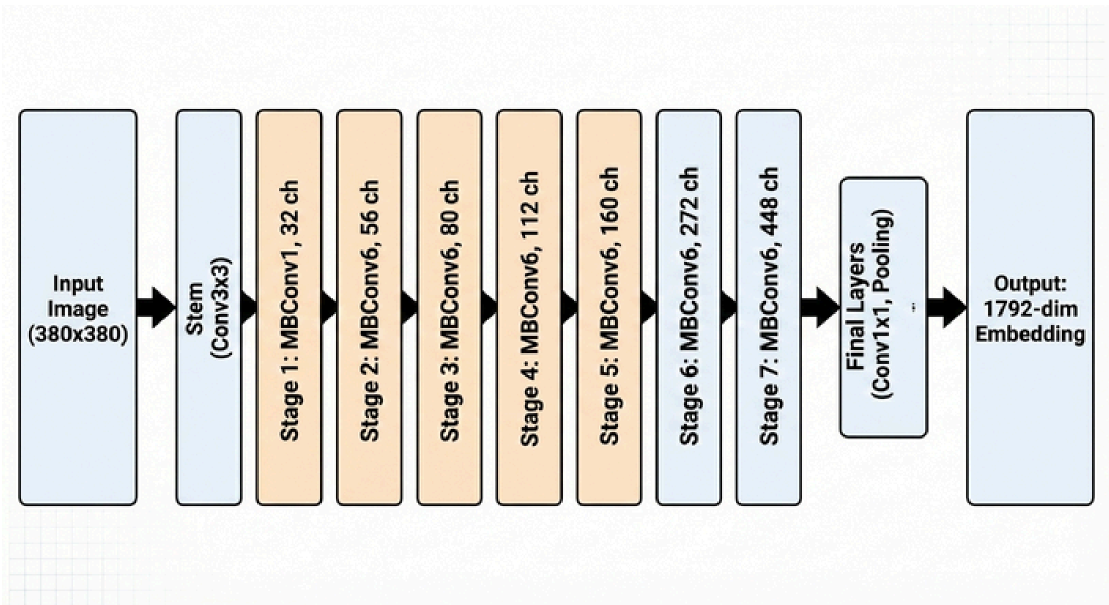
I engineered additional high-signal features:
- House age
- Renovation indicator and years since renovation
- Quality-weighted living area
- Condition-weighted living area
- Relative living and lot size (compared to neighborhood averages)

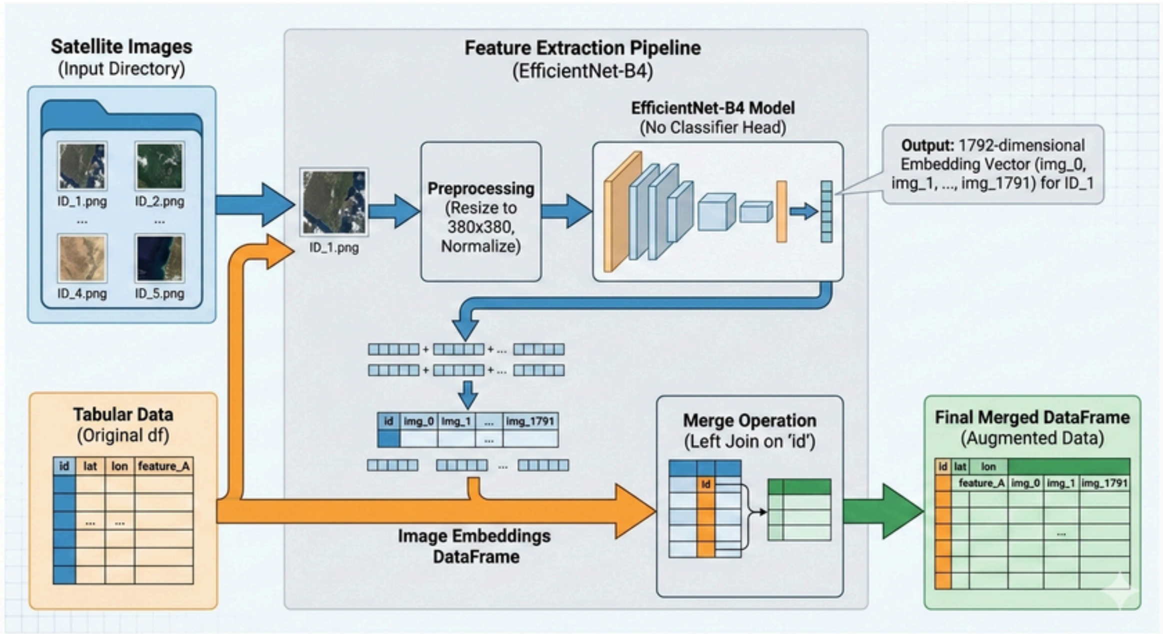
These features improve **price sensitivity modeling.**

9. Satellite Image Feature Extraction (EfficientNet-B4)

To extract visual features, I used **EfficientNet-B4 pretrained** on **ImageNet** as a **fixed feature extractor.**







Why EfficientNet-B4?

- Strong performance with efficient computation
- Compound scaling of depth, width, and resolution
- Well-suited for satellite imagery

Extraction Details:

- Input resolution: **380×380**
- Classification head removed
- Output: **1792-dimensional embedding per image**
- Model run in evaluation mode on GPU

Each house image was converted into a **dense semantic representation of its surroundings**.

10. Fusion Strategy Used

I used **late fusion**, where visual and tabular features are combined **after independent preprocessing.**

Fusion pipeline:

1. Extract 1792-D image embeddings
2. Apply PCA (1792 → 60) to reduce noise and dimensionality
3. Concatenate PCA features with tabular features
4. Train a single regression model on the combined feature space

Why Late Fusion?

- Works well with tree-based models
- Avoids alignment issues between modalities
- Computationally efficient
- Allows independent feature engineering

11. Zipcode Encoding (Leak-Free)

Zipcode was encoded using **smoothed target encoding:**

- Computed strictly on training data
- Leave-one-out strategy for training rows
- Smoothed toward global mean
- Validation zipcodes mapped safely

This avoids **data leakage and location bias.**

12. Dimensionality Reduction (PCA)

High-dimensional image embeddings can cause overfitting.

To address this:

- PCA fitted only on training embeddings
- Retained **60 principal components**
- Preserved most of the visual variance
- Reduced computational and statistical noise

13. Modeling Strategy

Target Transformation

- Model trained on log(price + 1)
- Predictions converted back to real price scale

This stabilizes variance and improves learning.

14. Models Trained

14.1 Tabular-Only XGBoost (Baseline)

- Uses only tabular and spatial features
- No hyperparameter tuning
- Serves as a strong classical benchmark

14.2 Multimodal XGBoost (Tabular + Image)

- Tabular features
- Zipcode target encoding
- PCA-compressed image embeddings
- Randomized hyperparameter tuning
- 3-fold cross-validation

15. Evaluation Metrics

I evaluated models using:

- **RMSE** – penalizes large pricing errors
- **R^2 score** – measures explained variance

Evaluation was performed on **real price values.**

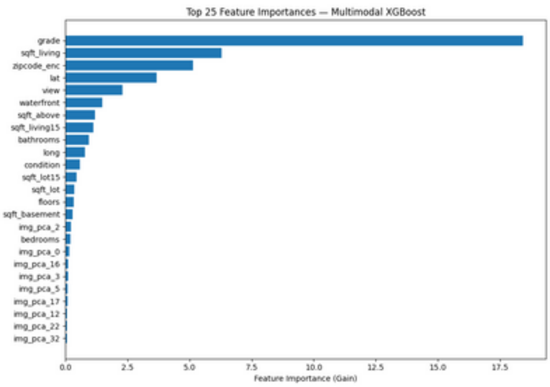
16. Results & Observations

Model	RMSE	R^2
Tabular Only	118597.31	0.8879
Tabular + Image	115809.26	**0.8931**

Key Insight:

Satellite imagery captures **environmental and neighborhood context** that is not present in structured data, leading to **lower error and higher explanatory power.**

17.Feature Importance



Top-25 feature importances (gain) from the multimodal XGBoost model, showing dominant influence of structural and location features with limited but present contribution from image-derived PCA components.

18. Interpretability Insights

- Image embeddings encode:
  - Green cover
  - Road and building density
  - Urban vs suburban patterns
- PCA ensures only dominant visual signals are retained
- Tree-based models allow feature importance analysis across modalities

19. Key Contributions

- Designed a full multimodal ML pipeline from scratch
- Integrated satellite imagery into price prediction
- Implemented leak-free encoding and spatial features
- Demonstrated measurable performance gains

20. Conclusion

Through this project, I demonstrated that **satellite imagery is a powerful complementary signal** for real estate valuation.

By combining deep visual embeddings with structured housing data, the model becomes **more context-aware and accurate**, outperforming traditional tabular-only approaches.

21.GRAD-CAM

Key Observations

- **High-importance** regions **(red/yellow)** consistently correspond to:
  - Dense green cover (trees, parks)
  - Low road density and quiet residential layouts
  - Organized housing patterns with larger plots
- **Lower-importance** regions **(blue)** are often associated with:
  - High concrete density
  - Road intersections and crowded built-up areas

Insights from Visualizations

- Properties surrounded by vegetation and open space receive stronger positive activation, indicating higher perceived value.
- Areas with clear separation between houses show more activation than tightly packed neighborhoods.
- The model focuses on neighborhood-level context, not just the individual building footprint.