Multimodal House Price Prediction Using Tabular Data and Satellite Imagery

Name: *Vamshi Poojari*

Enrollment Number: 23113157

Approach and Modeling Strategy

This project aims to predict residential property prices using a **multimodal regression framework** that combines traditional tabular housing attributes with **satellite imagery–derived visual context**.

While conventional real estate valuation models rely solely on structured features such as square footage, number of bedrooms, and location coordinates, they fail to capture important neighborhood-level characteristics like green cover, road connectivity, and urban density. To address this limitation, this project integrates **satellite images** fetched using geographic coordinates and extracts visual features using a **Convolutional Neural Network (CNN)**.

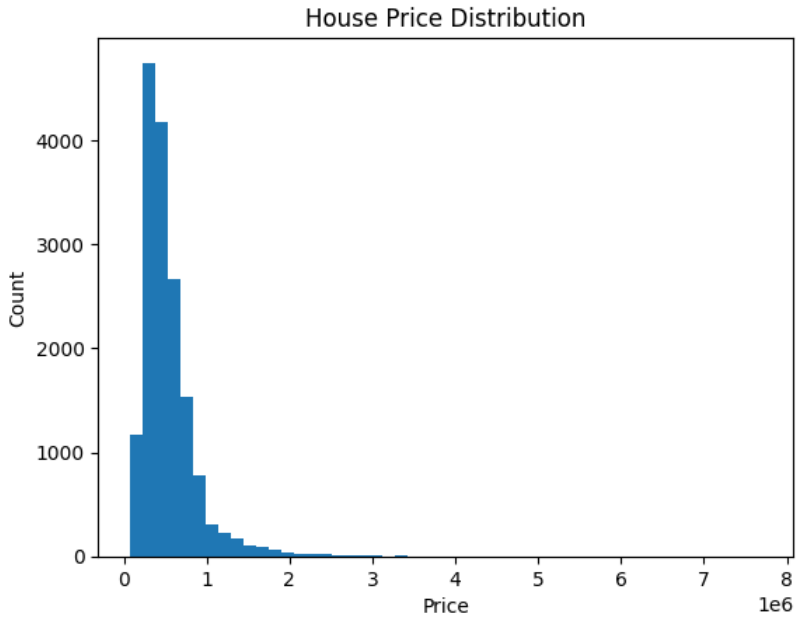
The modeling pipeline follows these steps:

1. Establish a strong tabular baseline using Linear Regression and XGBoost.
2. Programmatically acquire satellite images using latitude and longitude.
3. Extract visual embeddings using a pretrained ResNet18 model.
4. Fuse tabular and image features using late fusion.
5. Apply Grad-CAM to explain the influence of visual regions on predictions.

EXPLORATORY DATA ANALYSIS (EDA)

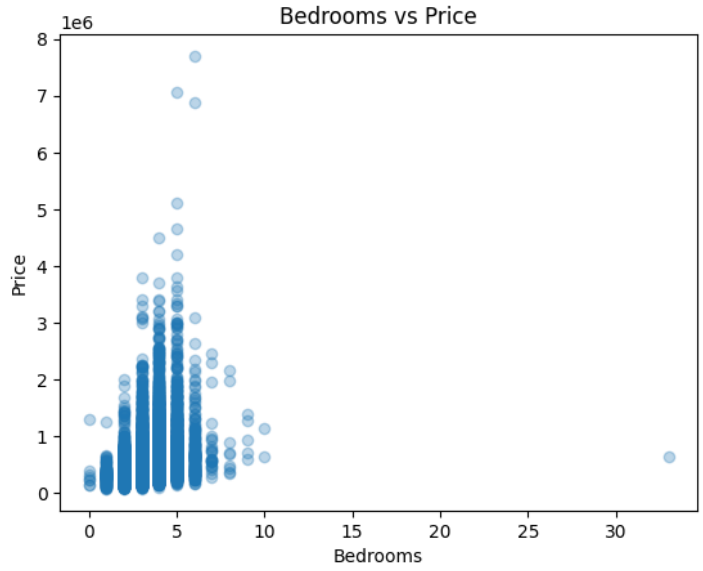
Price Distribution

The distribution of house prices reveals a right-skewed pattern, indicating the presence of high-value properties while most houses fall into a moderate price range.



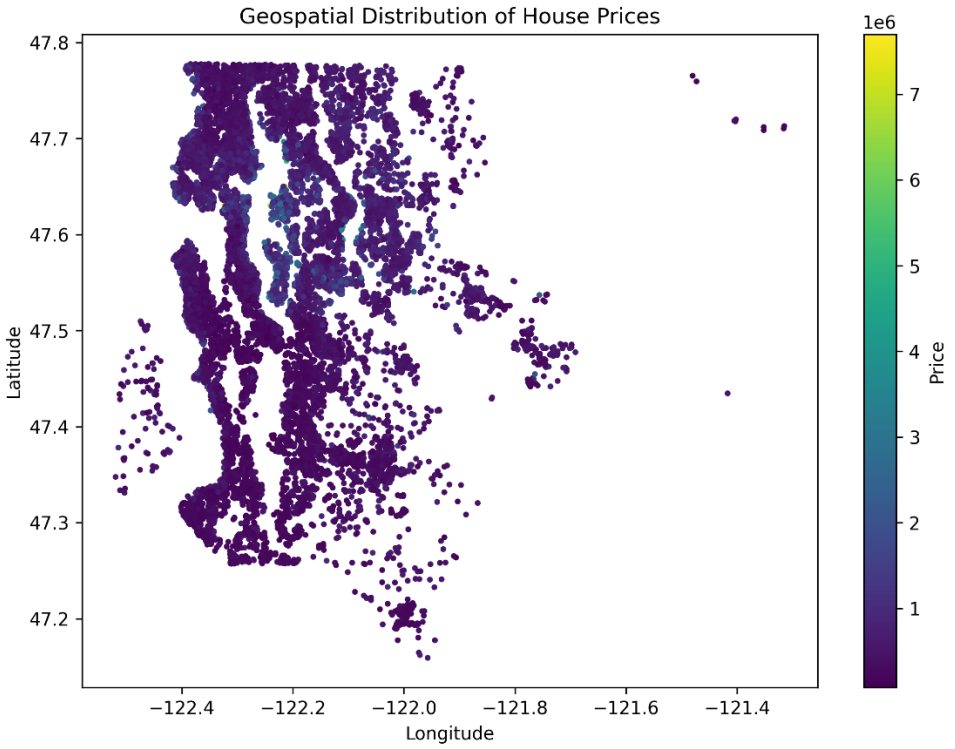
Bedrooms vs House Price

A positive correlation is observed between the number of bedrooms and house price, although variance increases for higher bedroom counts, suggesting that other factors influence valuation.



Geospatial Price Distribution

Mapping house prices using latitude and longitude highlights strong geographic clustering, reinforcing the importance of neighborhood-level information in price prediction.



Sample Satellite Images

Satellite imagery provides visual context about surrounding infrastructure, greenery, and urban density that is not present in tabular data.

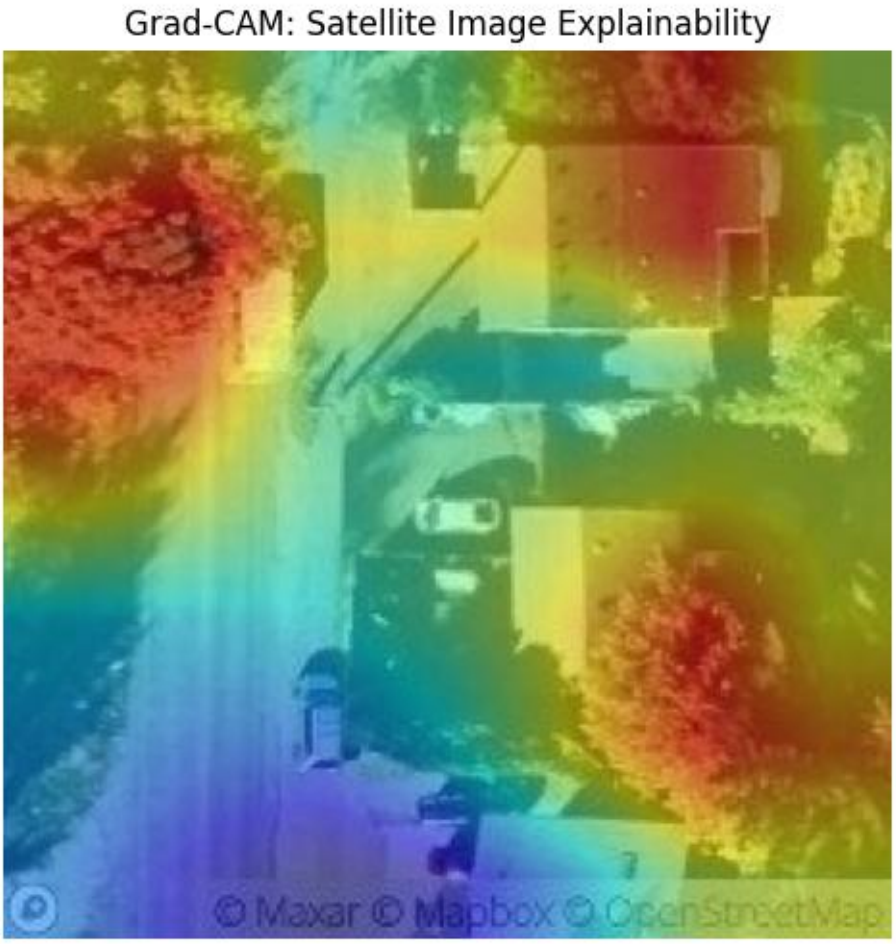
FINANCIAL & VISUAL INSIGHTS

Satellite imagery contributes valuable insights into real estate valuation:

- **Green cover (trees, parks)** is associated with higher property values.
- **Road connectivity** improves accessibility and positively impacts price.

- **Dense built-up regions** indicate urban convenience but may reduce value if overcrowded.

These insights validate the inclusion of visual data alongside traditional features.



RESULTS

Model Performance Comparison

The performance of different models is summarized below:

Model	RMSE	R^2
Linear Regression (Tabular)	~219k	0.62
XGBoost (Tabular)	~139k	0.845
Multimodal (Raw Fusion)	~156k	0.806
Multimodal (PCA-Controlled Fusion)	~143k	0.838

The tabular XGBoost model achieved the best numerical performance. However, the multimodal model provided valuable interpretability and neighborhood-level insights.