Satellite Imagery-Based Property Valuation

Overview

Traditional real-estate valuation models rely heavily on structured attributes such as square footage, number of bedrooms, and location coordinates. While effective, such models fail to capture visual and environmental context-for example, neighborhood greenery, surrounding infrastructure, or urban density-which significantly influence perceived and market value.

In this project, we build a multimodal regression pipeline that integrates:

Tabular housing data (structural and locational features), and

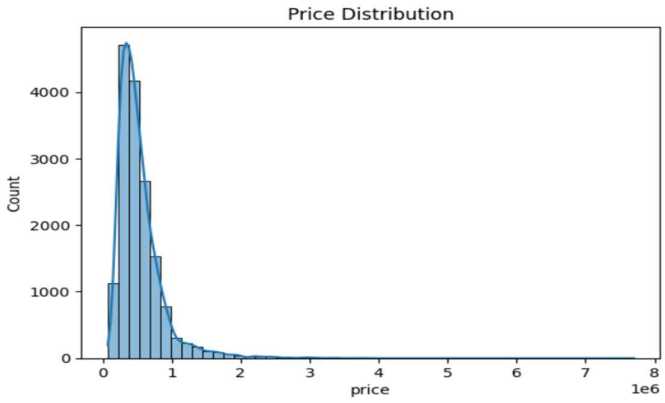
Satellite imagery (capturing neighborhood context)

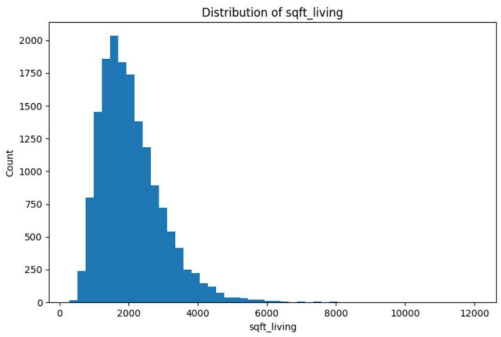
The core idea is to extract high-level spatial features from satellite images using a Convolutional Neural Network (CNN) and combine them with engineered tabular features in a unified model. This approach aims to improve predictive performance while also enabling interpretability of visual factors affecting property prices.

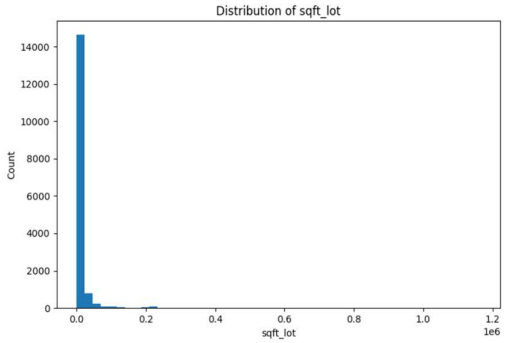
Exploratory Data Analysis (EDA)

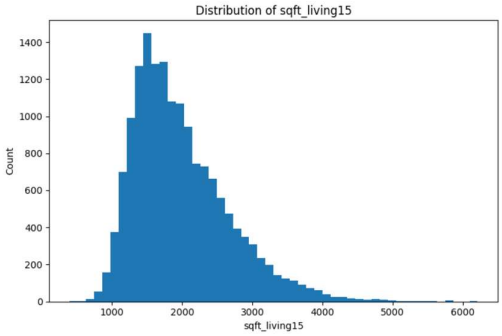
Price Distribution

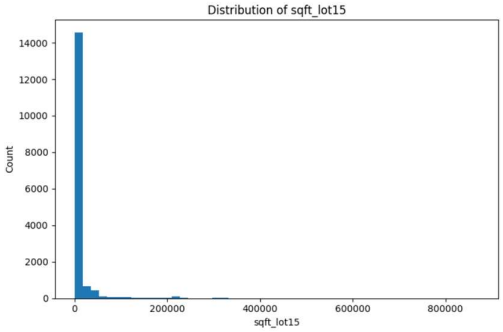
The target variable (property price) exhibits a right-skewed distribution, with a majority of properties concentrated in the lower to mid price ranges and a smaller number of high-value outliers. This behavior is typical of real-estate markets and motivates the use of robust metrics such as RMSE and R2 rather than raw error alone.



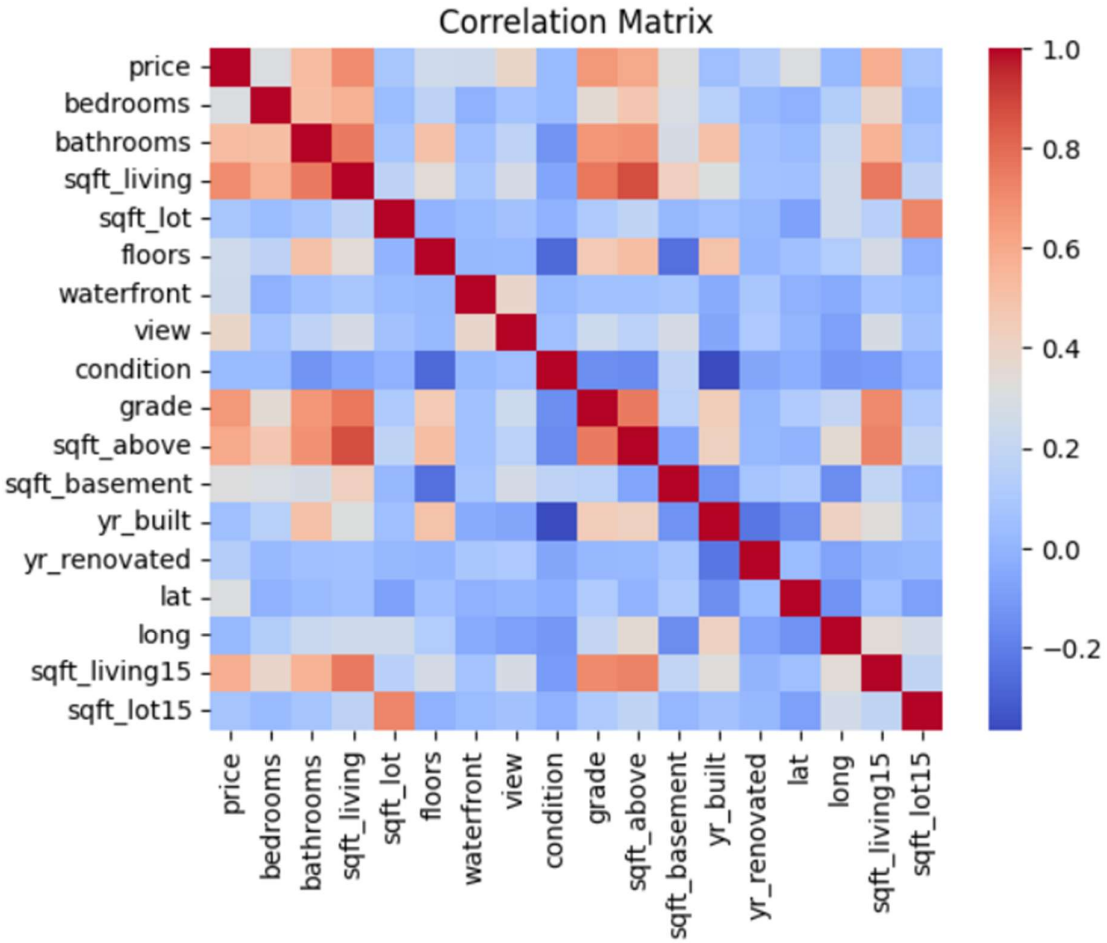








Tabular Feature Observations

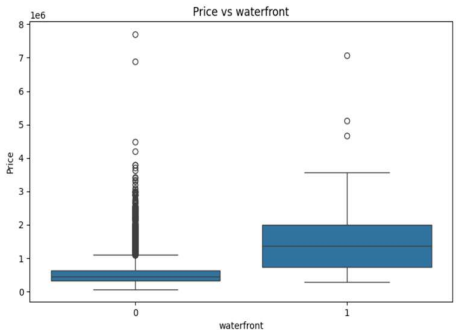


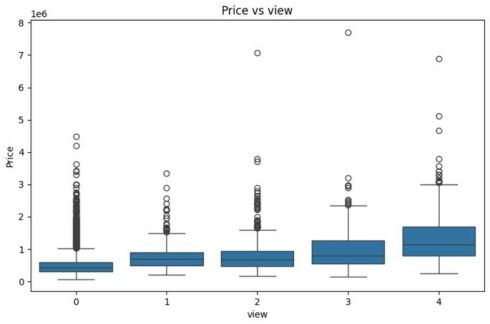
Living area-related features (e.g., square footage, number of bathrooms) show strong positive correlation with price.

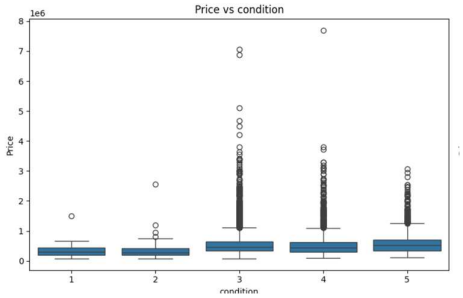
Location-derived features (latitude, longitude, distance to city center) reveal spatial clustering of higher prices.

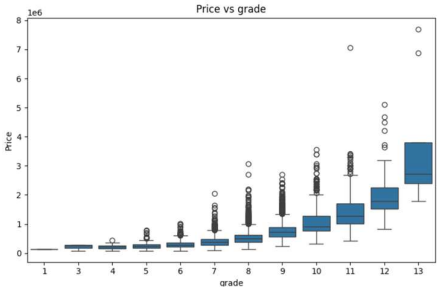
Discrete attributes such as waterfront presence and view quality show noticeable price premiums.

These findings justify the baseline tabular regression model and also indicate that location alone does not fully explain valuation differences.









Financial and Visual Insights

Satellite imagery provides indirect signals of neighborhood quality and long-term desirability:

Green cover (trees, lawns, parks) often correlates with higher property prices, reflecting better environmental quality and lifestyle appeal.

Organized infrastructure, such as structured road networks and planned housing layouts, tends to associate with higher valuations.

High concrete density with minimal open space frequently corresponds to lower prices, especially in congested urban zones.

From a financial perspective, these visual indicators act as proxies for:

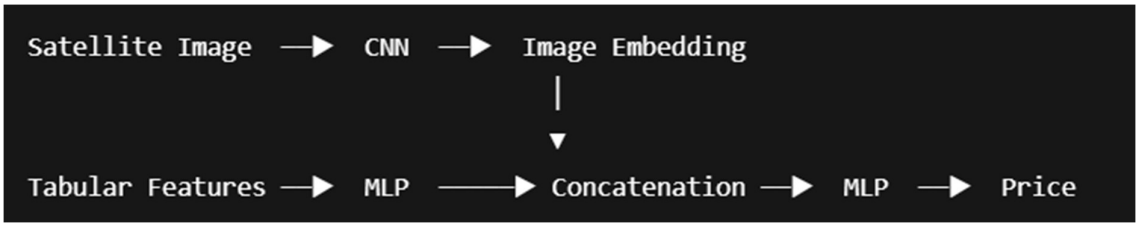
Quality of life

Future appreciation potential

Neighborhood socioeconomic status

The multimodal model is therefore not merely improving prediction accuracy but also aligning with real-world valuation logic.

4. Model Architecture



5. Results and Evaluation

Baseline: Tabular Data Only

The tabular-only model performs reasonably well, capturing most structural and locational effects. However, residual errors remain high for properties whose value is strongly influenced by neighborhood context rather than internal attributes alone.

Multimodal Model: Tabular + Satellite Images

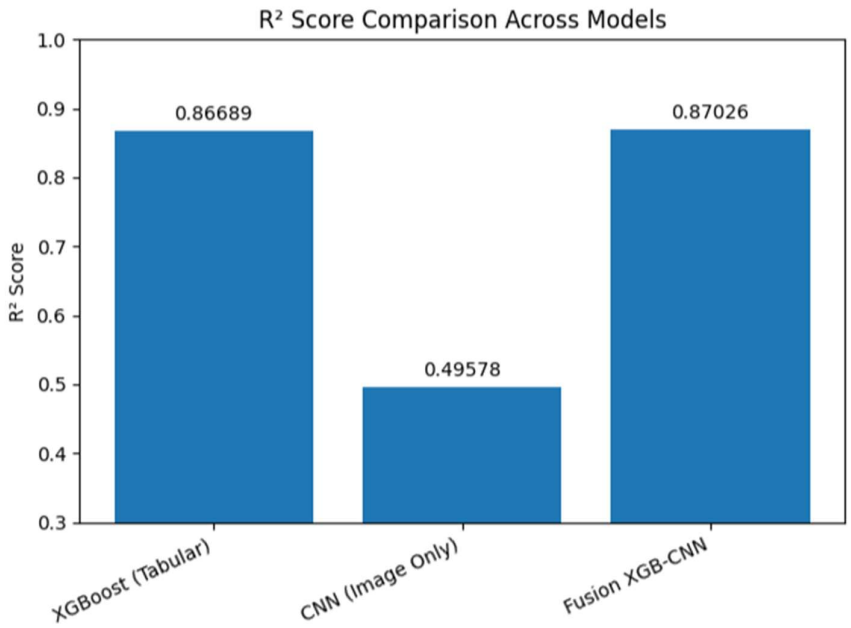
Adding satellite imagery leads to:

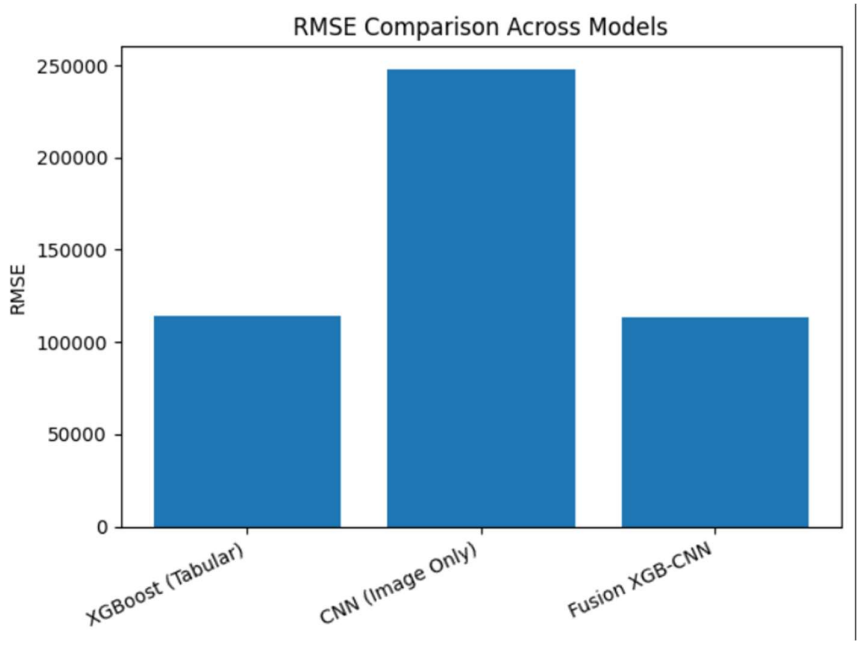
Lower RMSE, indicating improved absolute prediction accuracy.

Higher R2, showing better explanation of variance in prices.

The improvement confirms that visual neighborhood context provides non-redundant information beyond tabular features.

Comparative Summary





Conclusion

This project demonstrates that integrating satellite imagery with traditional housing data leads to more accurate and financially meaningful property valuations. Visual features such as greenery, infrastructure density, and neighborhood layout significantly influence market prices and can be effectively captured using CNNs.

The multimodal framework not only improves prediction metrics but also aligns with real-world real-estate valuation principles, making it a strong candidate for practical deployment in analytics-driven property assessment systems.