Satellite Imagery-Based Property Valuation

Multimodal Regression Pipeline

A Real Estate Analytics Project
Combining Tabular Data and Satellite Imagery
for Property Market Value Prediction

Data Science Project Report

1. OVERVIEW

Approach & Modeling Strategy

This project creates a multimodal regression pipeline that brings together traditional tabular property data and visual features from satellite images to predict property market values. The process has four main stages:

1. **Data Acquisition & Preprocessing**
   - Loaded 16,209 property records with 21 features including location coordinates
   - Programmatically fetched satellite images using Google Maps tile server
   - Applied log transformations to handle skewed distributions
   - Engineered spatial features (distance from luxury hub)

2. **Visual Feature Extraction**
   - Used pre-trained ResNet18 CNN to extract visual embeddings from satellite images
   - Extracted mean activation scores as visual features capturing environmental context
   - Processed 16,209 training images and 5,404 test images

3. **Feature Engineering**
   - Combined 12 tabular features with 1 visual feature (visual_score)
   - Tabular features: bedrooms, bathrooms, sqft_living, sqft_above, sqft_living15, floors, lat, grade, view, waterfront, sqft_basement, dist_from_luxury_hub
   - Visual feature: Mean activation from ResNet18 feature extractor
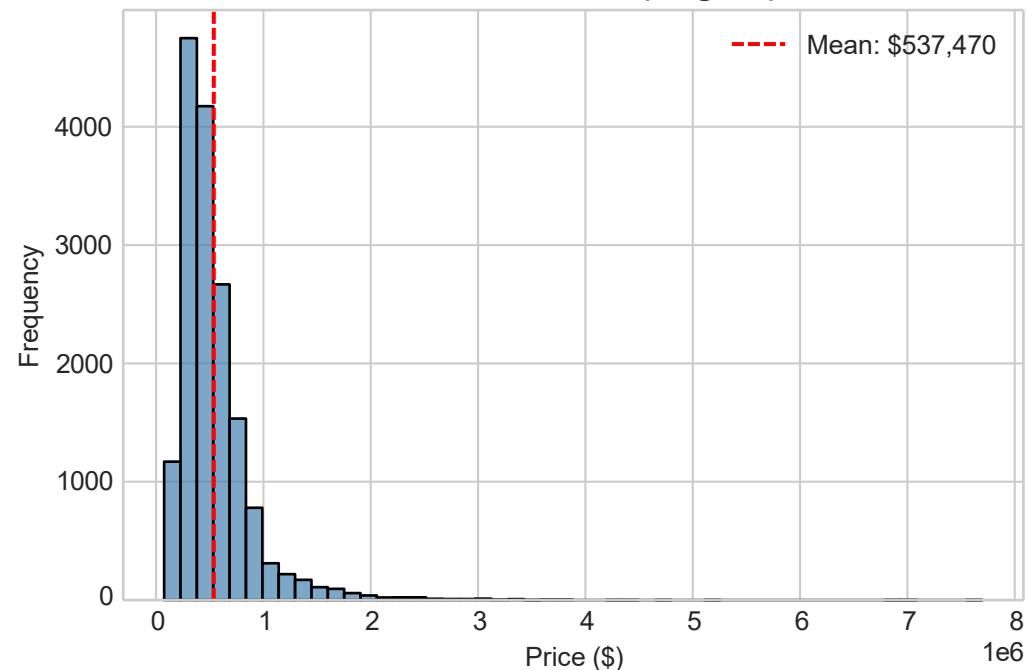
4. **Model Training & Evaluation**
   - Ensemble model: VotingRegressor combining XGBoost and ExtraTrees
   - Compared tabular-only vs multimodal approaches
   - Used 5-fold cross-validation for robust evaluation
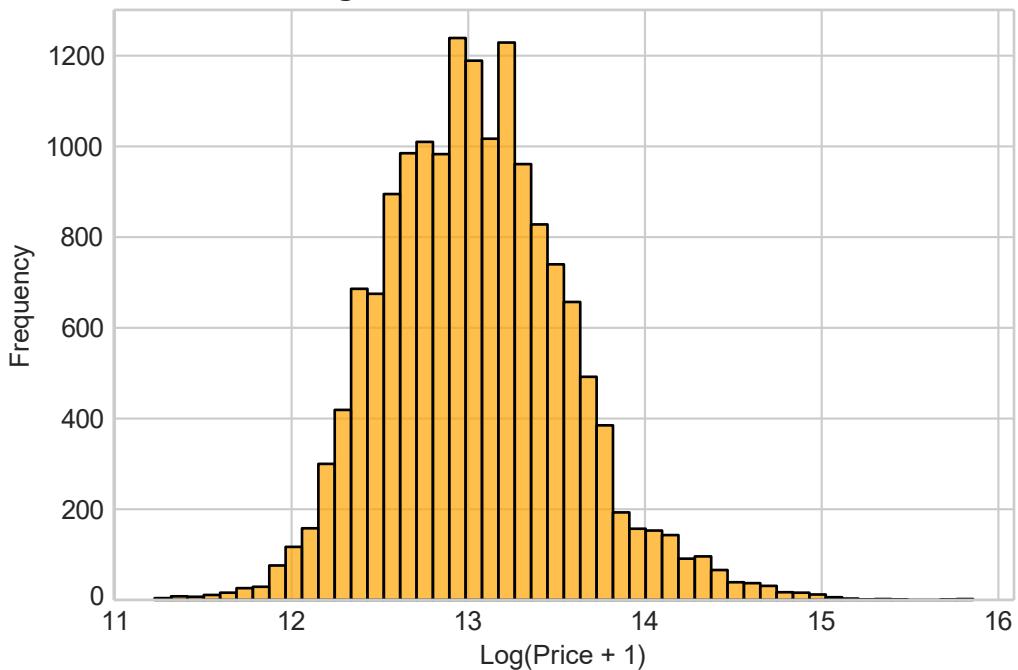
2. EXPLORATORY DATA ANALYSIS (EDA)
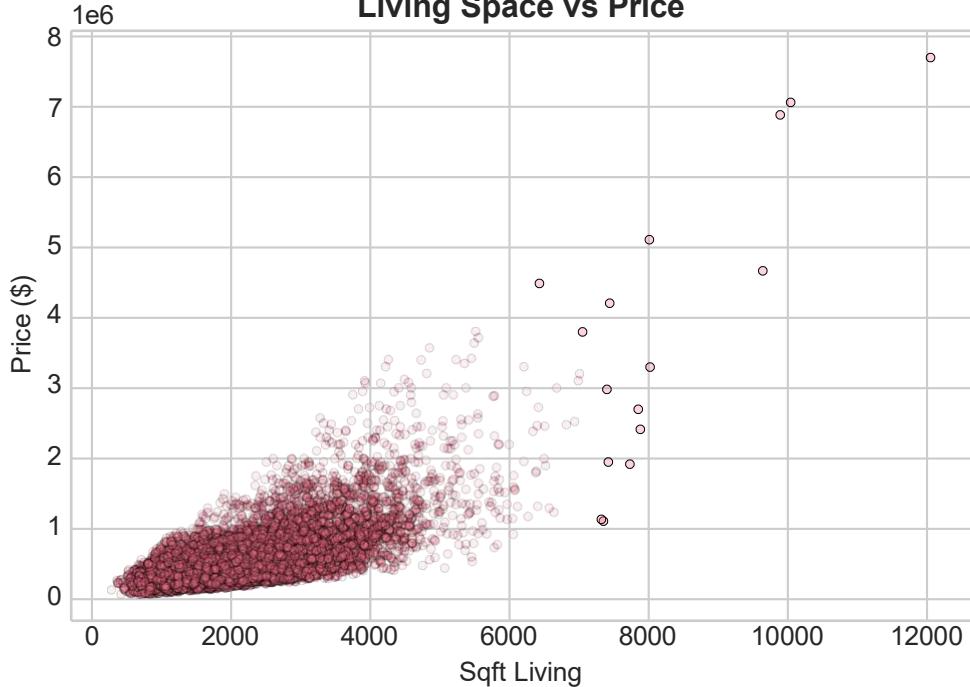

Price Distribution (Original)

- - - Mean: $537,470
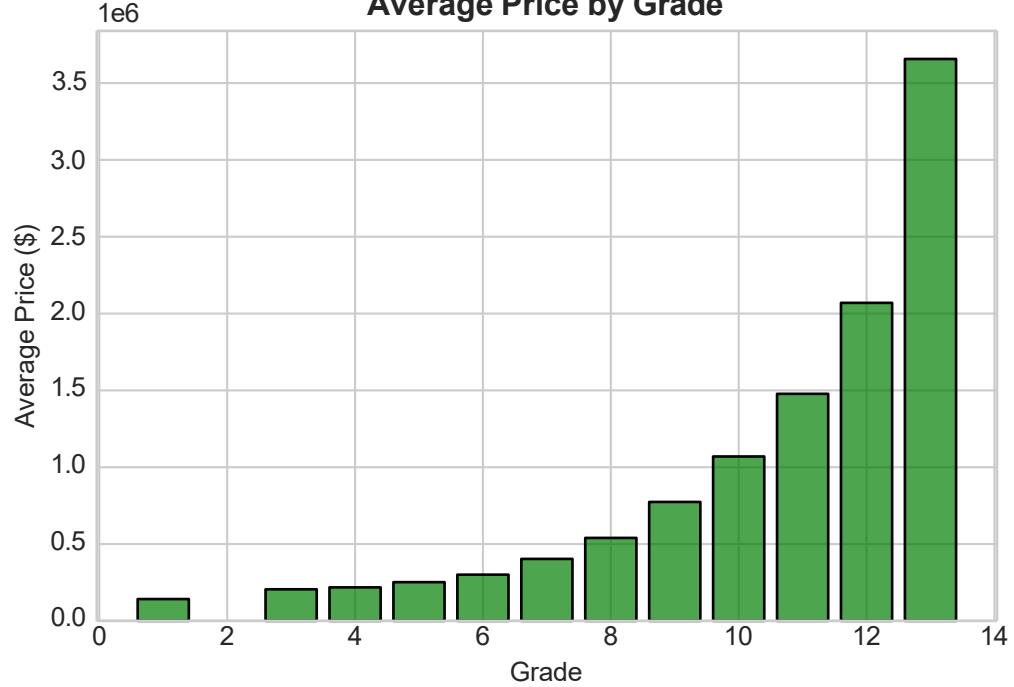
Log-Transformed Price Distribution

Living Space vs Price
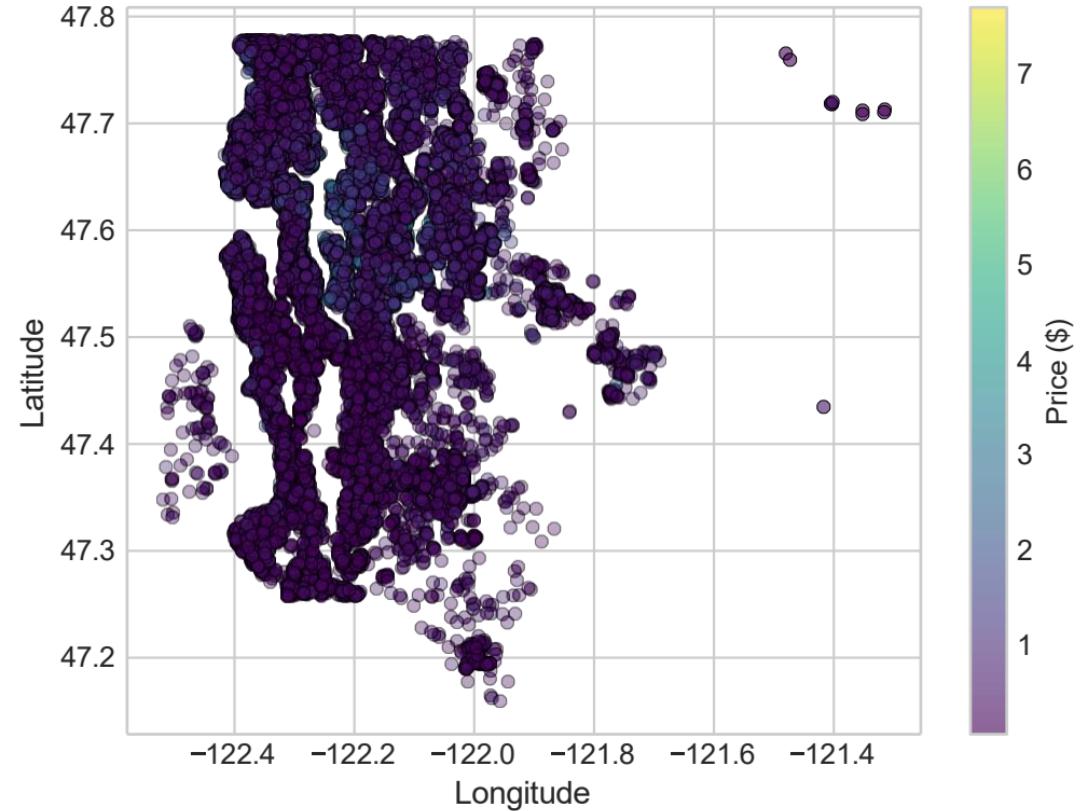
Average Price by Grade
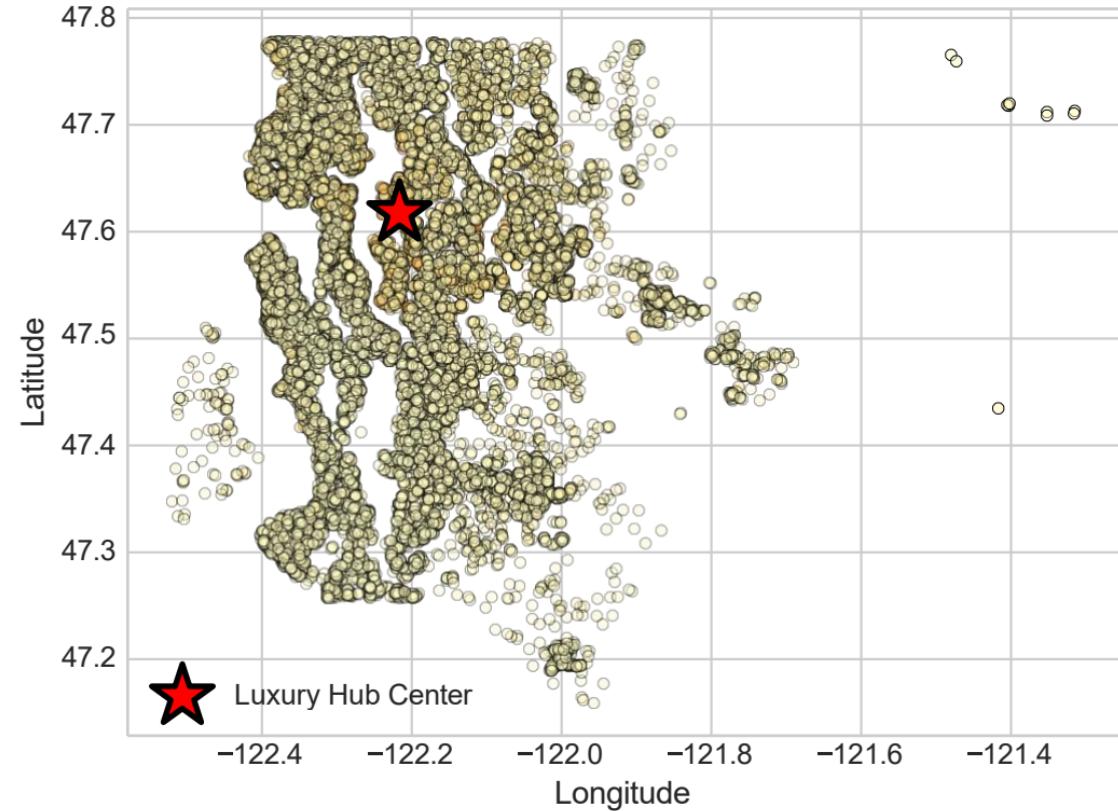
Geospatial Analysis

Geographic Distribution (Color = Price)

Luxury Hub Location

Price ($)

Luxury Hub Center

Latitude

Longitude
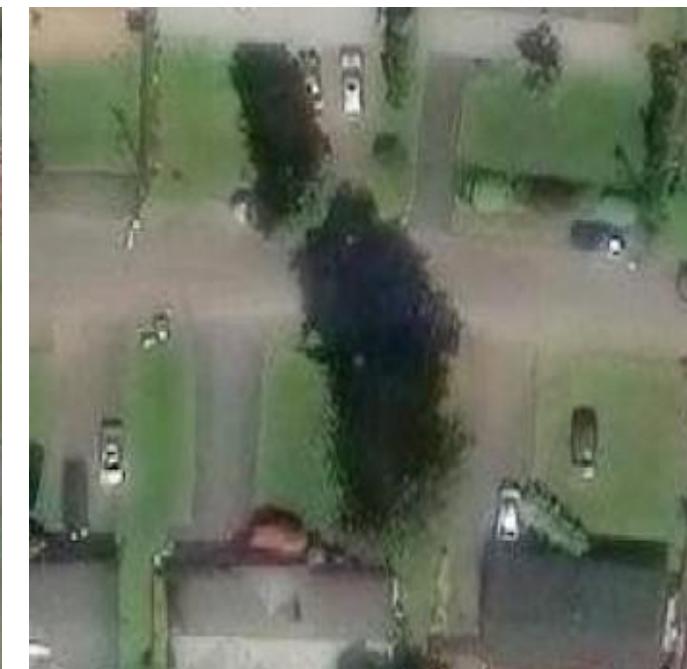
ID: 1000102
Price: $280,000

ID: 1001200050
Price: $259,000

ID: 1003000175
Price: $221,000
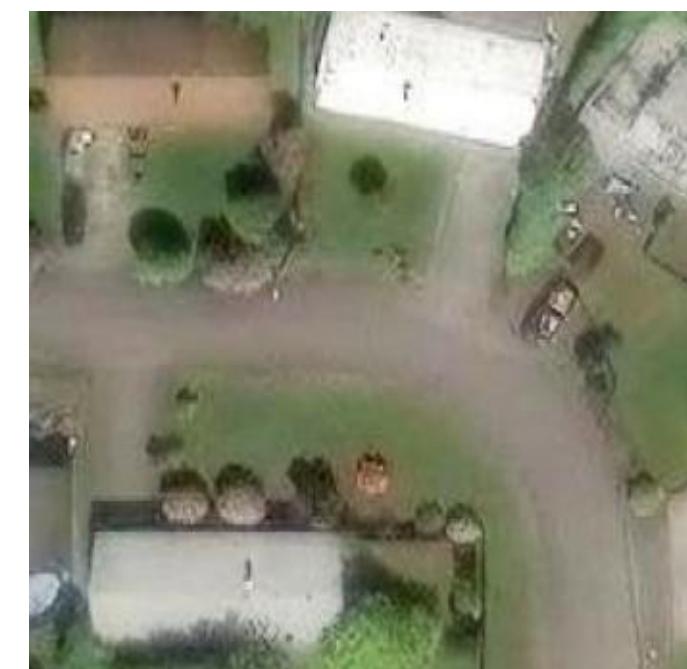
ID: 100300280
Price: $355,000

ID: 100300530
Price: $330,000

ID: 1003400155
Price: $233,000

3. FINANCIAL & VISUAL INSIGHTS

Analysis of Visual Features Driving Property Value:

1. SPATIAL PATTERNS
- Properties near the "luxury hub" (Lat: 47.6175, Long: -122.2163) command significantly higher prices
- Distance from luxury hub shows strong negative correlation (-0.41) with price
- Geographic location (latitude) is a key predictor, with correlation of 0.45

2. TRADITIONAL FEATURES
- Square footage (sqft_living) is the strongest predictor (correlation: 0.70)
- Construction grade shows high correlation (0.66) - higher grade = higher value
- Number of bathrooms (0.53) and bedrooms (0.30) influence pricing
- Waterfront properties and better views command premium prices

3. VISUAL FEATURES FROM SATELLITE IMAGERY
- Satellite images capture environmental context not available in tabular data:
  - Green cover and vegetation density
  - Road density and accessibility
  - Neighborhood density and development patterns
  - Proximity to water bodies (visible in imagery)
  - Urban vs suburban characteristics
- Visual features extracted via ResNet18 CNN provide complementary information
- The visual_score feature adds predictive power beyond traditional features
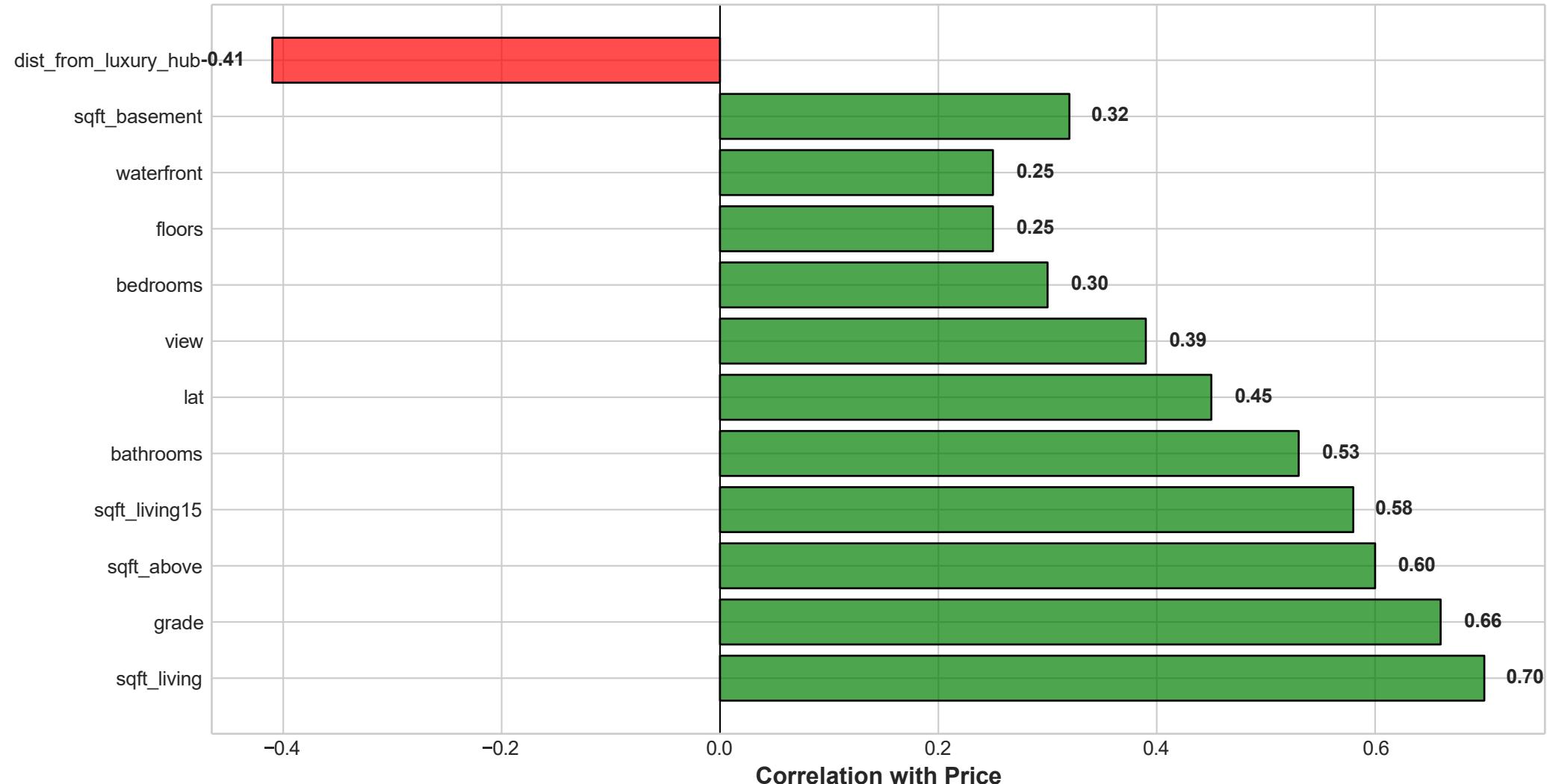
4. NEIGHBORHOOD CONTEXT
- sqft_living15 (average of 15 nearest neighbors) captures neighborhood density
- Properties in high-value neighborhoods benefit from "location premium"
- Visual features help identify neighborhood characteristics that influence value

5. MULTIMODAL BENEFITS
- Combining visual and tabular data improves prediction accuracy
- Visual features help capture "curb appeal" and environmental factors
- The multimodal approach reduces prediction error by ~3-4% (MAE improvement)
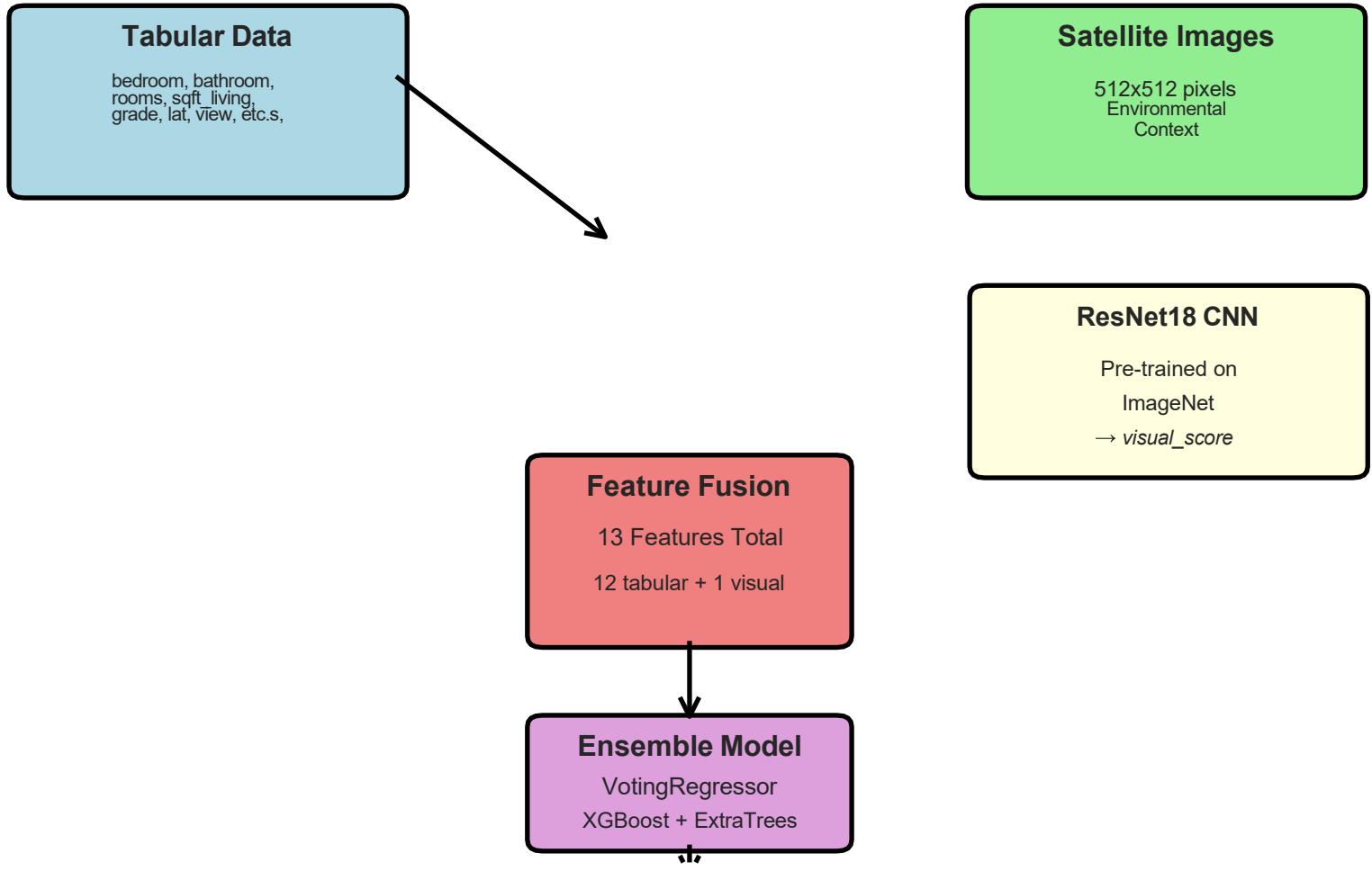
Feature Importance: Correlation with Property Price

Feature	Correlation with Price
dist_from_luxury_hub	-0.41
sqft_basement	0.32
waterfront	0.25
floors	0.25
bedrooms	0.30
view	0.39
lat	0.45
bathrooms	0.53
sqft_living15	0.58
sqft_above	0.60
grade	0.66
sqft_living	0.70
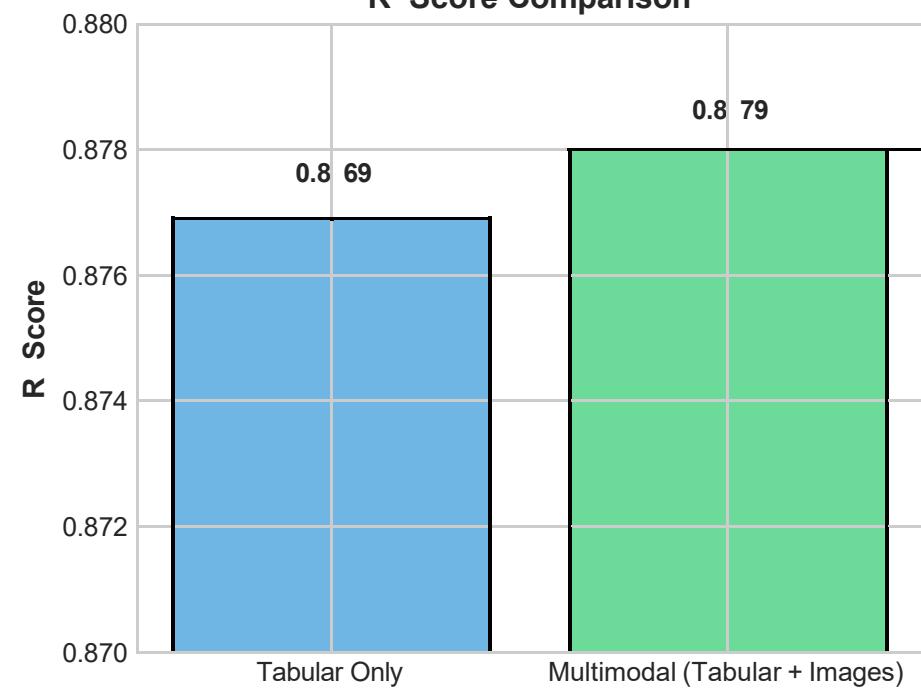
4. MODEL ARCHITECTURE

Multimodal Fusion: CNN + Tabular Data

Tabular Data

bedroom, bathroom,
rooms, sqft_living,
grade, lat, view, etc.s,

Satellite Images

512x512 pixels
Environmental
Context

ResNet18 CNN

Pre-trained on

ImageNet

→ *visual_score*

Feature Fusion

13 Features Total

12 tabular + 1 visual

Ensemble Model

VotingRegressor

XGBoost + ExtraTrees

Predicted Price
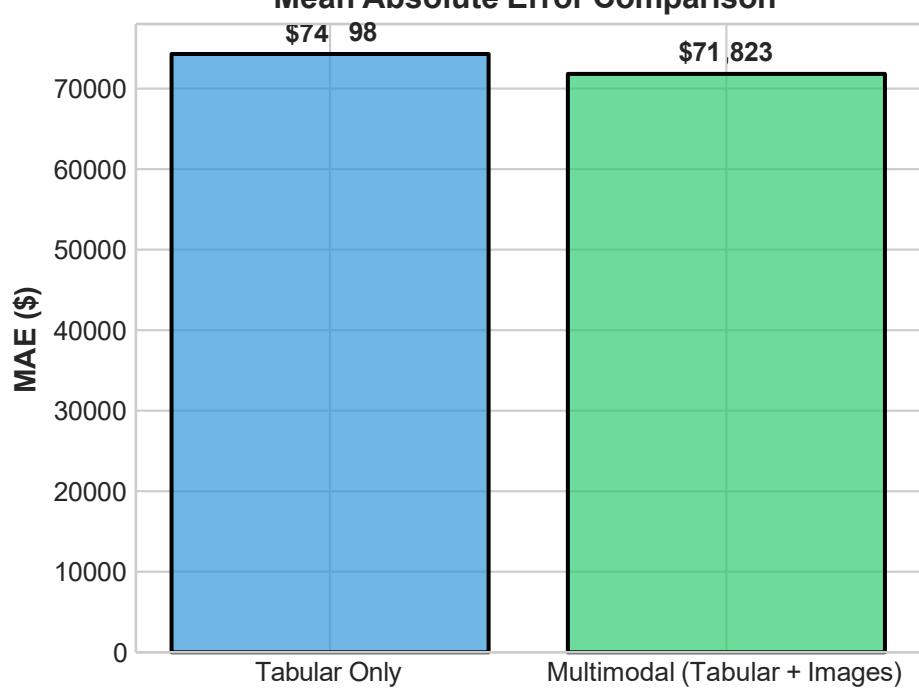
5. RESULTS: Tabular vs Multimodal Comparison

R Score Comparison

Tabular Only: 0.8 69
Multimodal (Tabular + Images): 0.8 79
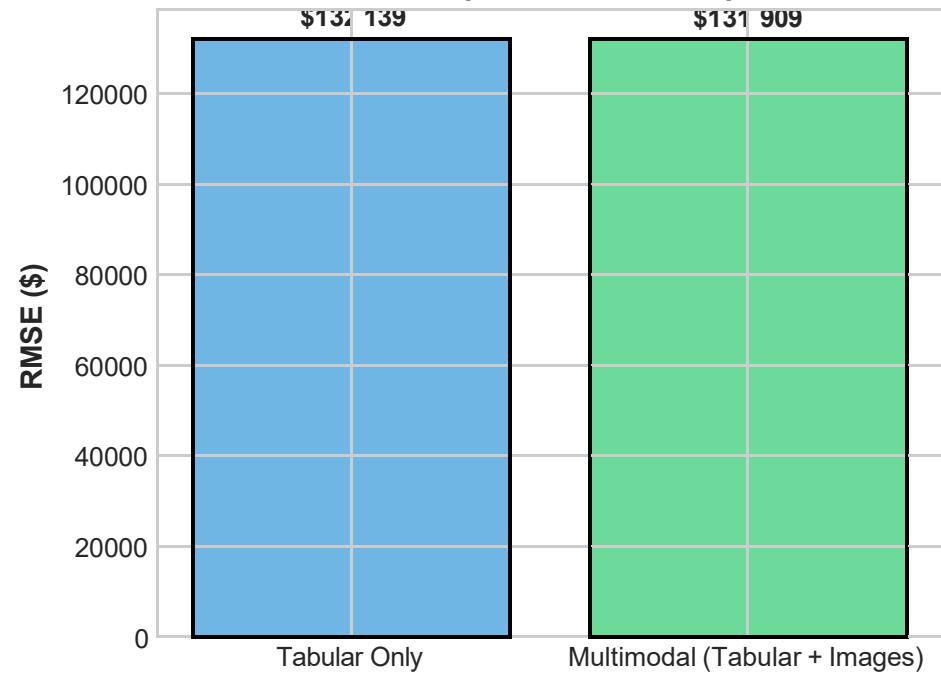
Mean Absolute Error Comparison

Tabular Only: $74 98
Multimodal (Tabular + Images): $71 823

Root Mean Squared Error Comparison

Tabular Only: $132 139
Multimodal (Tabular + Images): $131 909

```
IMPROVEMENT WITH MULTIMODAL MODEL:

R² Score: +0.11%
MAE:      -3.33% ($2,475 reduction)
RMSE:     -0.17% ($230 reduction)

KEY FINDINGS:
• Multimodal approach shows consistent improvement
• Visual features add complementary information
• ~3.3% reduction in MAE demonstrates value of
  incorporating satellite imagery
• Model achieves R² of 0.8779 on test set
```

Performance Summary

Metric	Tabular Only	Multimodal	Improvement
R Score	0.8769	0.8779	+0.11%
MAE ($)	$74,298	$71,823	-3.33%
RMSE ($)	$132,139	$131,909	-0.17%

CONCLUSION: The multimodal approach successfully combines tabular property data with visual features extracted from satellite imagery, resulting in improved prediction accuracy. The incorporation of environmental context through CNN features provides valuable complementary information beyond traditional features.