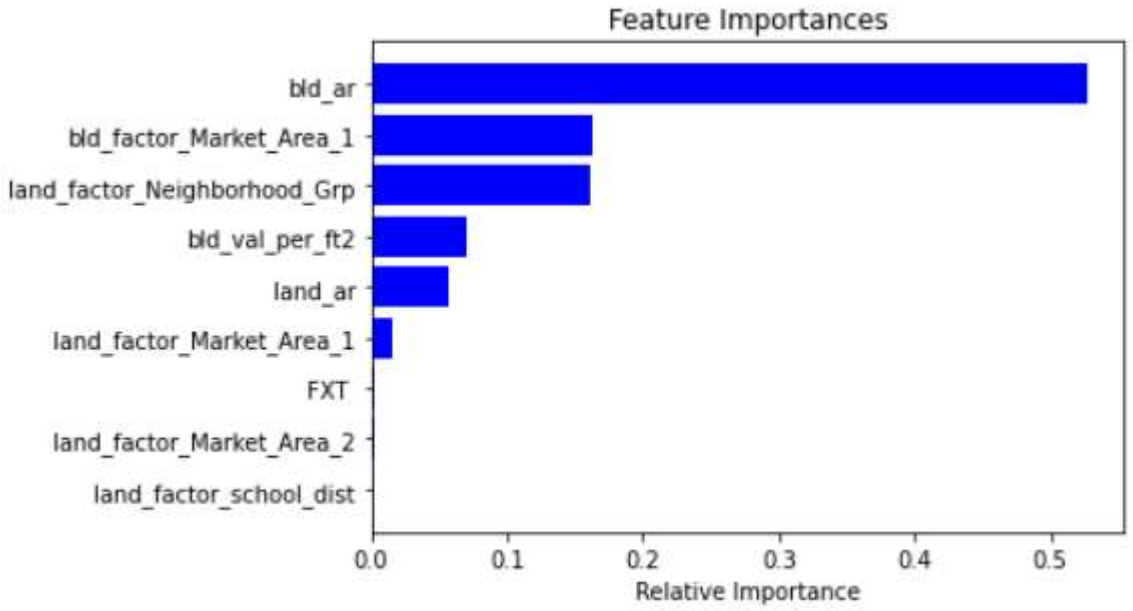
Model metrics

This document is intended to document the selected models' features, parameters, hyperparameters, and performance metrics

Tot_appr_val:

The best performing model evaluated in this project was the Random Forest Regression model from the sklearn library.

The following chart summarizes the features and their importance in the final model.



Parameters and hyperparameters:

Random_state = 47

N_estimators = 100

Criterion = mse

Min_samples_split = 2

min_samples_leaf = 1

Min_weight_fraction_leaf = 0.0

Max_features = auto

Min_impurity_decrease = 0.0

Bootstrap = True

Oob_score = False

Verbose = 0

Ccp_alpha = 0

Metrics:

R2 score training = 0.9964663497295445

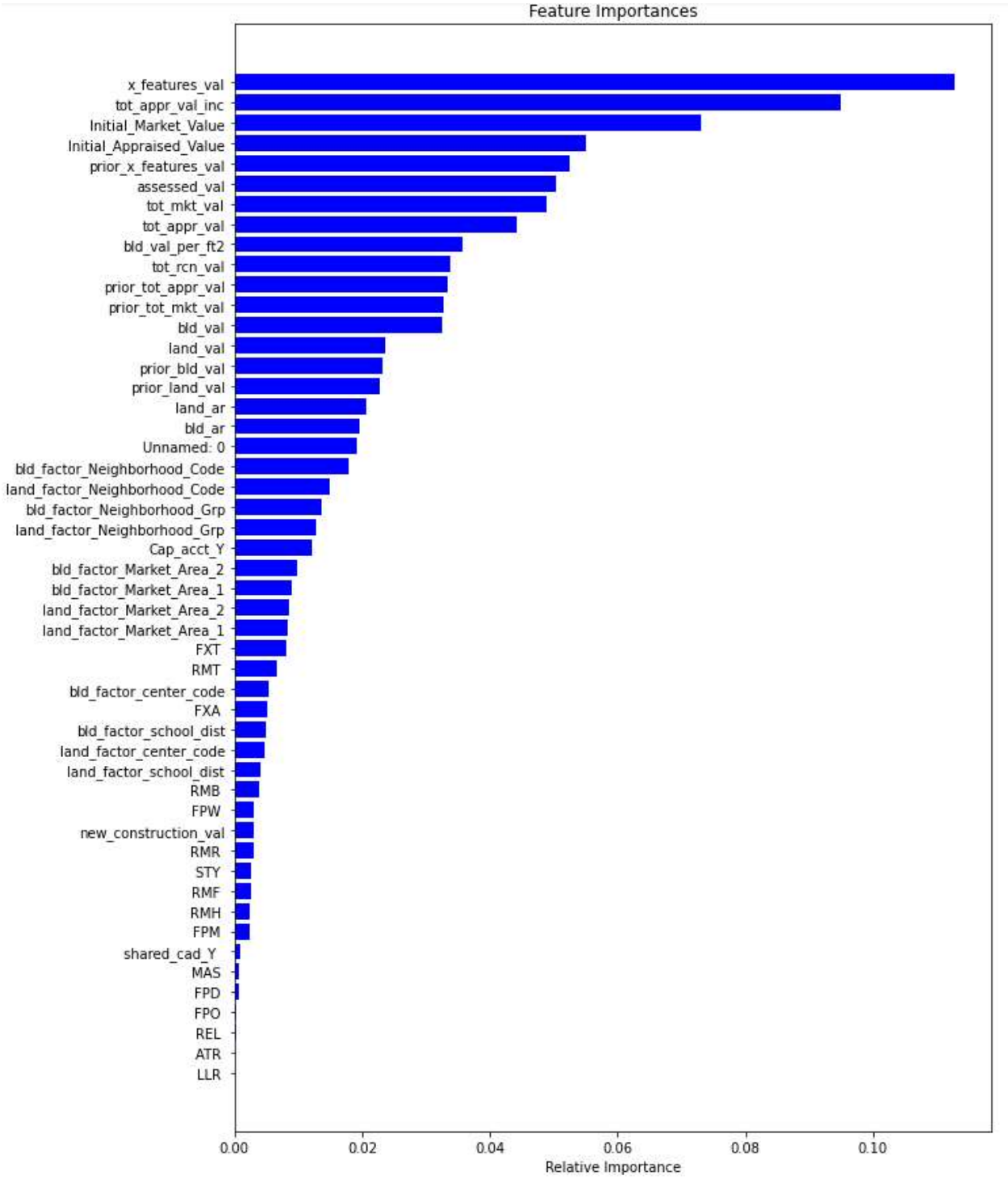
R2 score testing = 0.974500867057765

RMSE = 50404.648048785755

Protested:

The best performing model evaluated in this project was the Random Forest Classification model from the sklearn library.

The following chart summarizes the features and their importance in the final model.



Parameters and hyperparameters:

Random_state = 47

N_estimators = 100

Criterion = gini

Min_samples_split = 2

min_samples_leaf = 1

Min_weight_fraction_leaf = 0.0

Max_features = auto

Min_impurity_decrease = 0.0

Bootstrap = True

Oob_score = False

Verbose = 0

Warm_start = False

Ccp_alpha = 0

Metrics:

Accuracy: 0.9167983194239333

Balanced accuracy: 0.8702956693259484

Precision score for "Yes" 0.9163689772345378

Precision score for "No" 0.9169265190680249

Recall score for "Yes" 0.7671035905557755

Recall score for "No" 0.9734877480961213

Predicted	0	1
Actual		
0	246711	6719
1	22352	73622