London House Price Prediction

Sanjay Raju

Problem Space

Dilemma
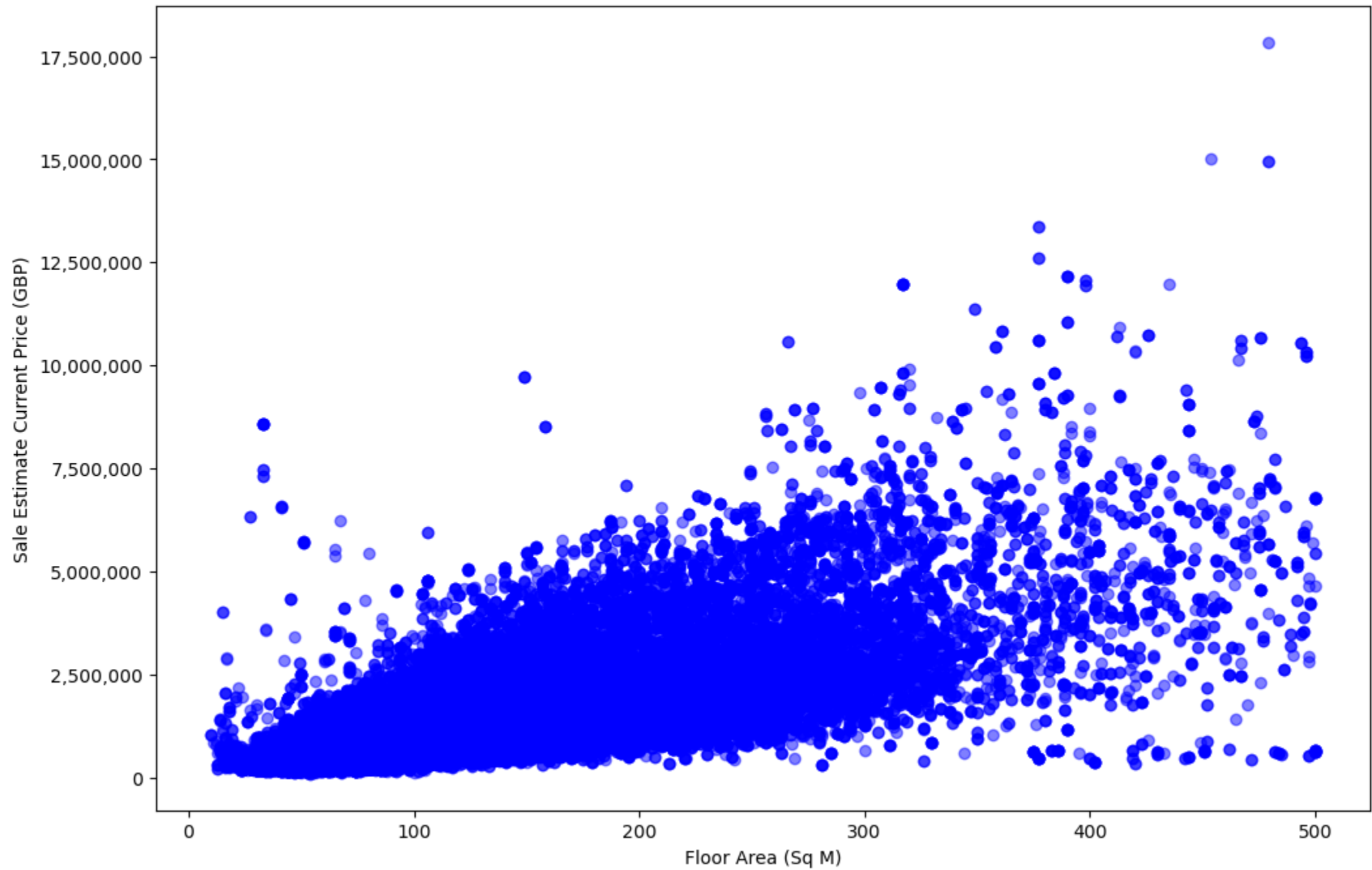Various attributes influence values of property

Objective
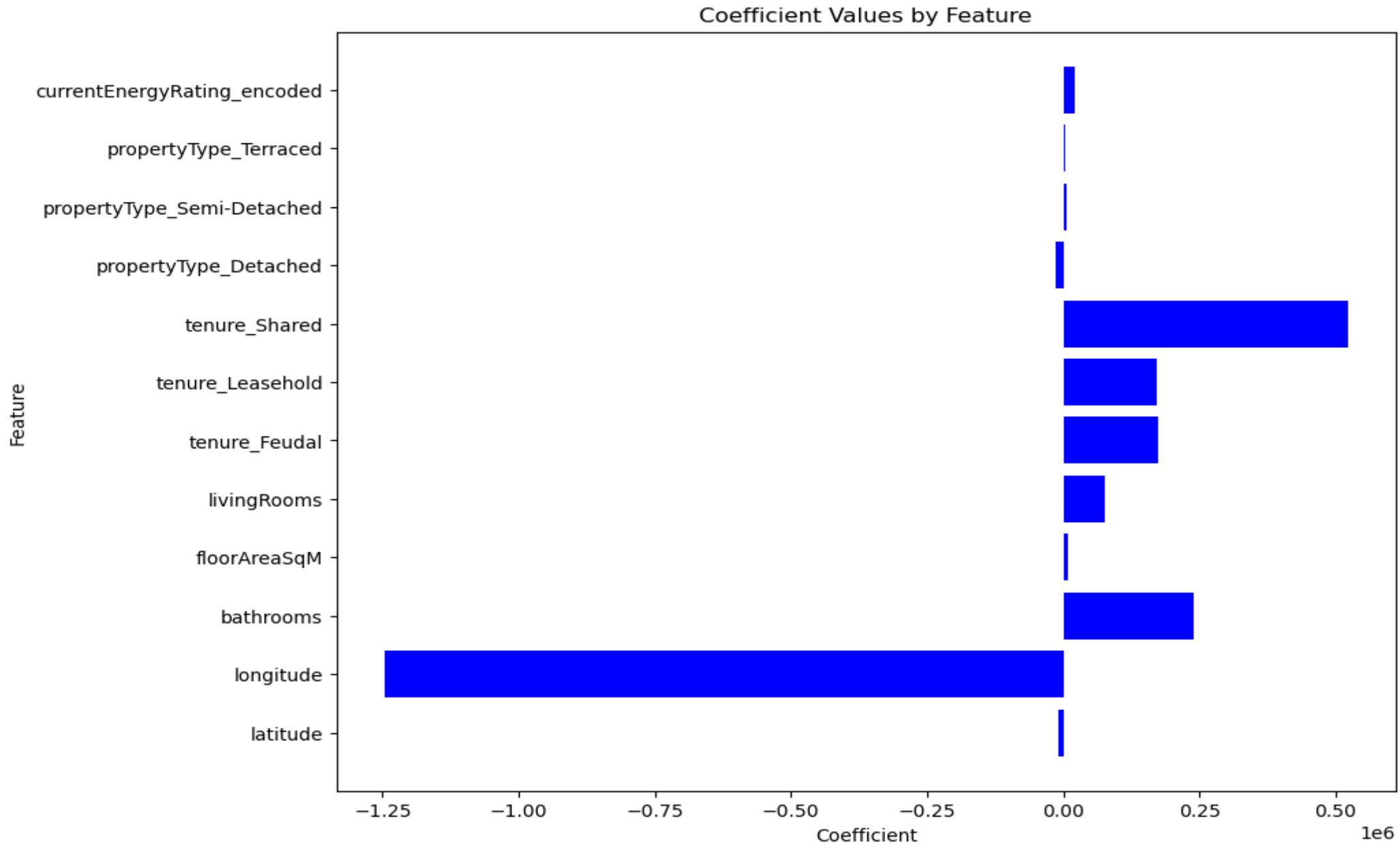Develop machine learning models to predict house prices

Use case
Buyers, sellers and estate agents

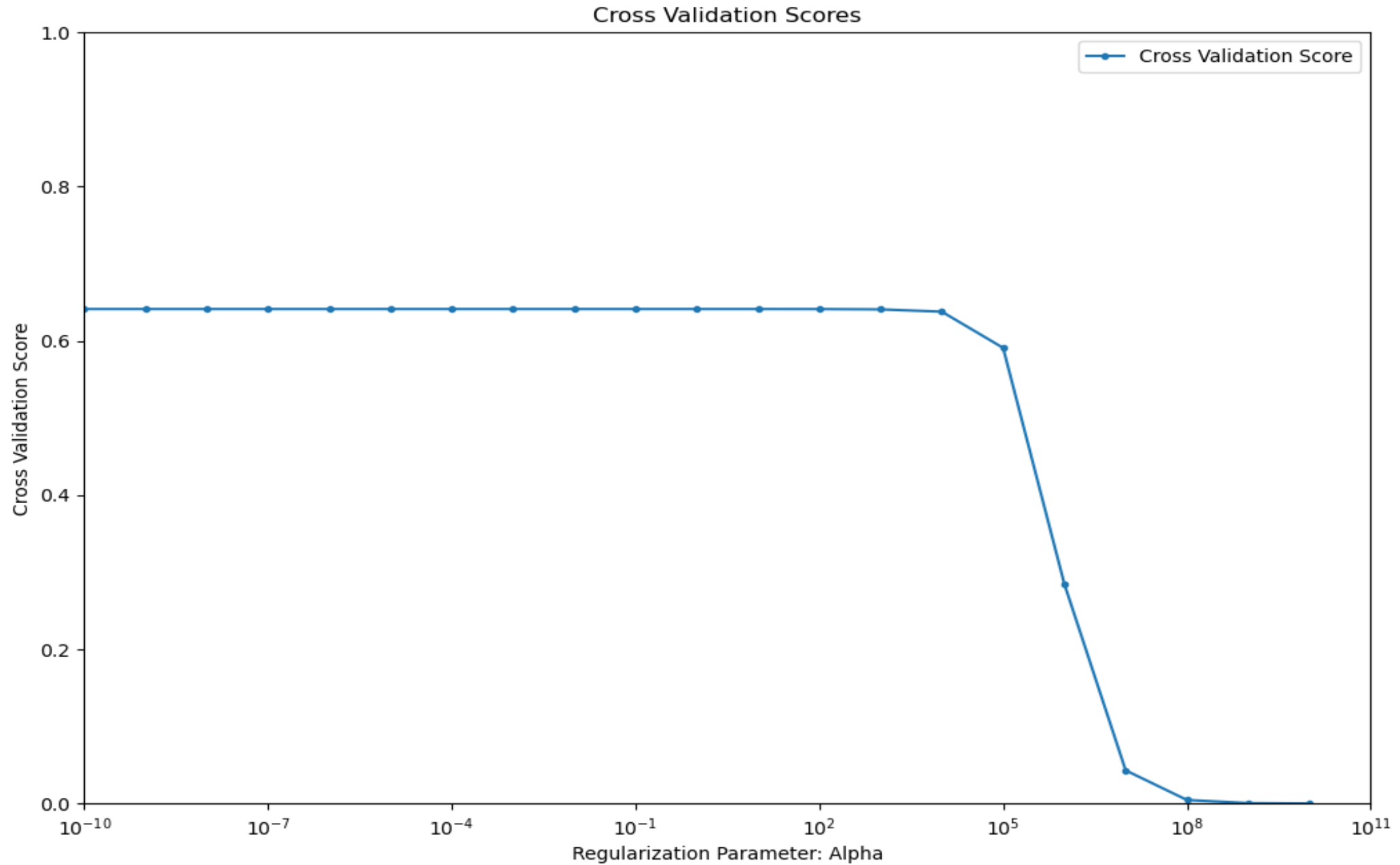Sale Estimate Current Price vs. Floor Area

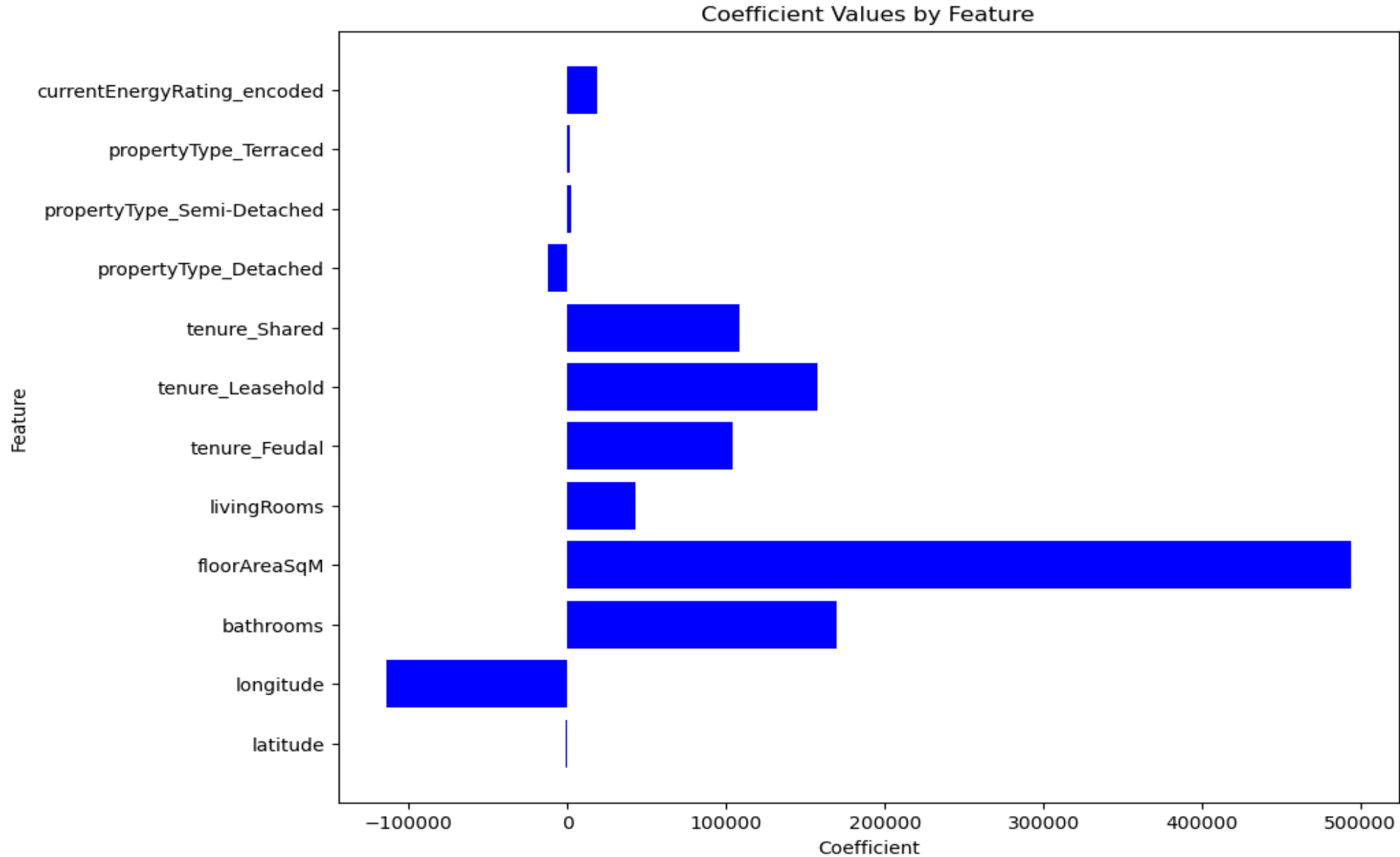# Linear Regression

Coefficient Values by Feature

- Chosen based on feature relationships
- R^2 = 0.64
- Coefficients directions accurate
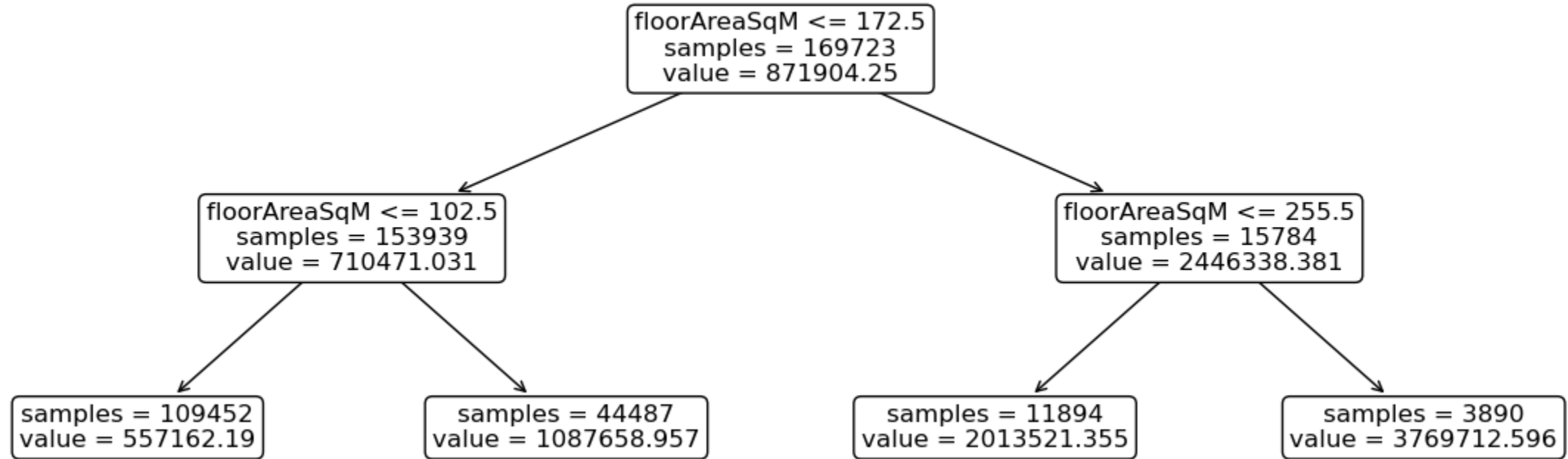
# Ridge Regression – Hyper parameter tuning

Ridge Regression

Coefficient Values by Feature

- R^2 = 0.64
- Coefficients magnitude more representative of actual

Decision Tree

```
                          ┌──────────────────────────┐
                          │ floorAreaSqM <= 172.5    │
                          │ samples = 169723         │
                          │ value = 871904.25        │
                          └──────────────────────────┘
                           /                        \
                          /                          \
        ┌──────────────────────────┐      ┌──────────────────────────┐
        │ floorAreaSqM <= 102.5    │      │ floorAreaSqM <= 255.5    │
        │ samples = 153939         │      │ samples = 15784          │
        │ value = 710471.031       │      │ value = 2446338.381      │
        └──────────────────────────┘      └──────────────────────────┘
         /                  \               /                    \
        /                    \             /                      \
┌────────────────────┐ ┌────────────────────┐ ┌────────────────────┐ ┌────────────────────┐
│ samples = 109452   │ │ samples = 44487    │ │ samples = 11894    │ │ samples = 3890     │
│ value = 557162.19  │ │ value = 1087658.957│ │ value = 2013521.355│ │ value = 3769712.596│
└────────────────────┘ └────────────────────┘ └────────────────────┘ └────────────────────┘
```

- Capture non – linear relationships

- Max depth = 2
- R^2 = 0.54

Decision Tree
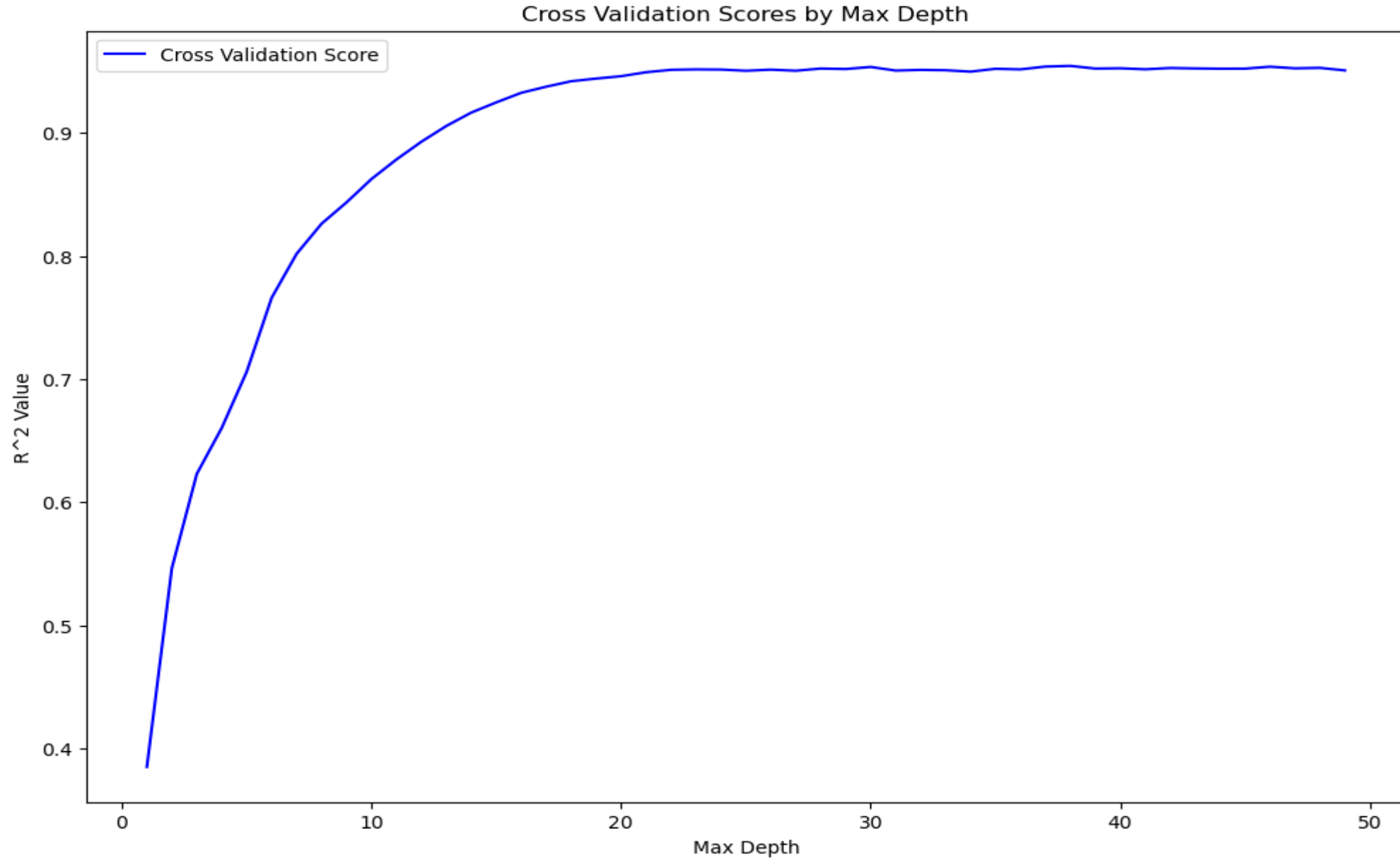
Random Forest Regressor

Cross Validation Scores by Estimators
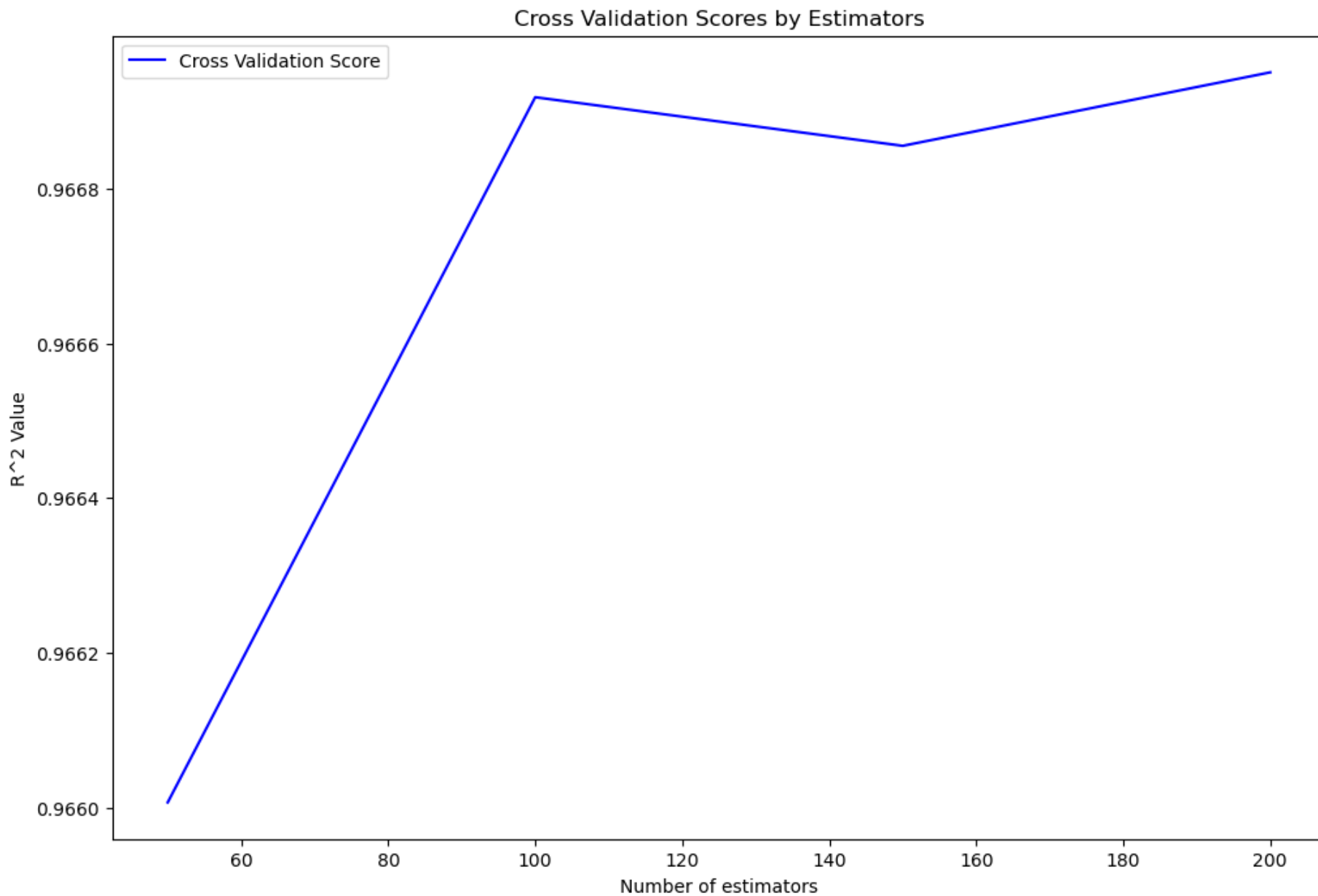
Feature importance
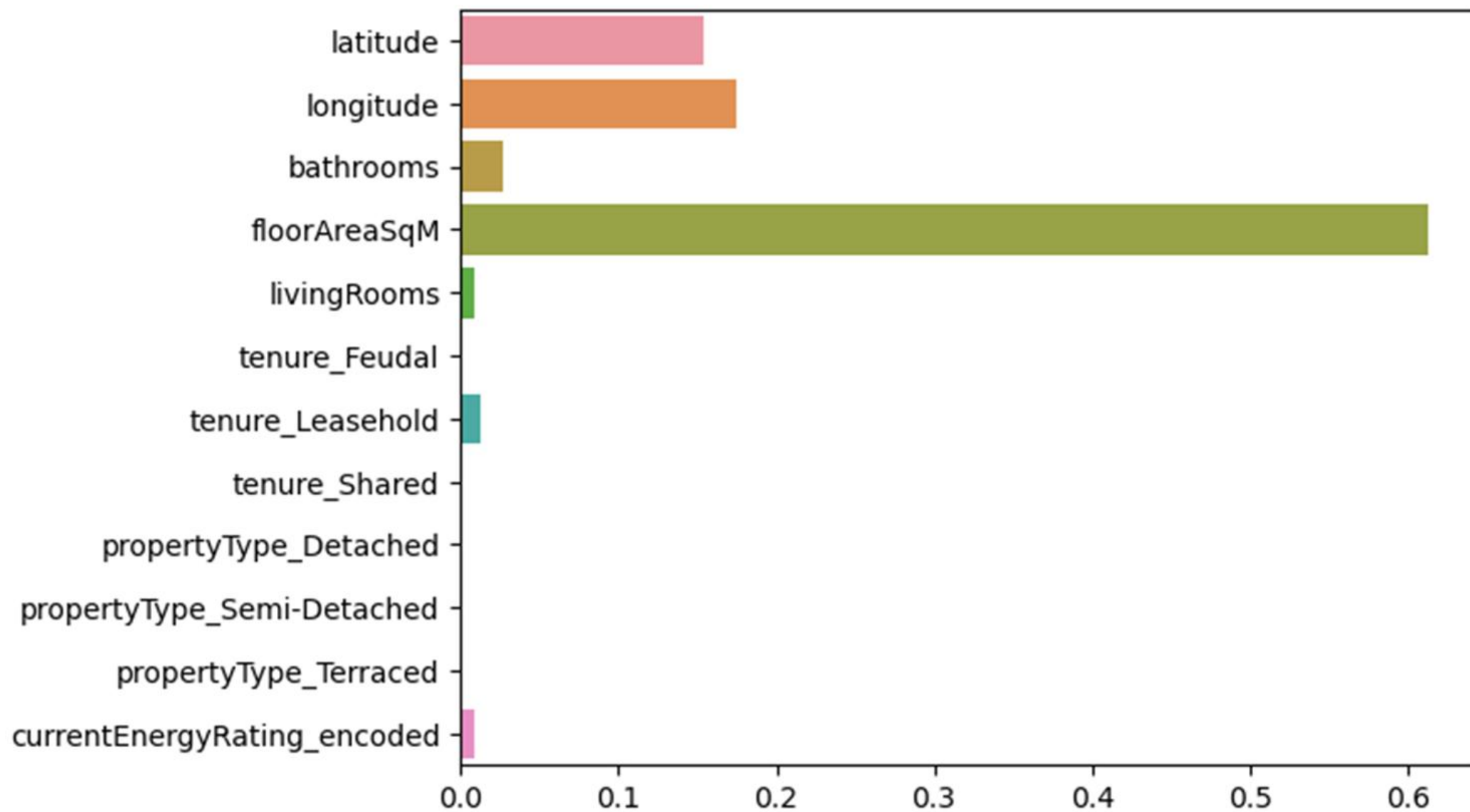
Next Steps

- Evaluate best performing models
- Uncover the 'accuracy' of each
- Final conclusions