House Price Prediction with Machine Learning Model

Submitted by : Md. Biplob Mina

Supervised by : Dr. Md Forhad Rabbi
Associate professor
Shahjalal University of Science and Techonology, Sylhet

Project Github Link: https://github.com/Biplob68/House-Price-Prediction-with-Machine-Learning-Model

Agenda

- Introduction
- Datasets
- Model Flow Diagram
- Data cleaning, pre-processing
- Regression models used
- Evaluation
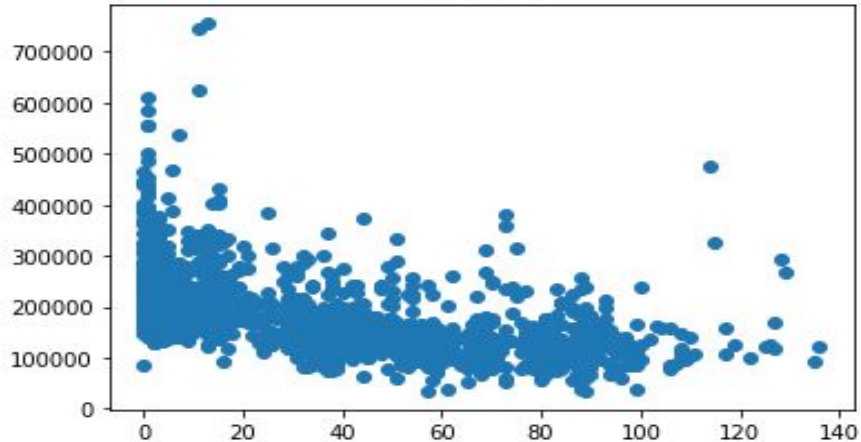- Result
- Conclusion

Introduction

The aim of this project is to identify the suitable model to make the prediction for the house price with significant predictor variable and used a supervised machine learning technique.

Two dataset, train and test are provided and the price of the test data is to be estimated. Here, In this model i have used XGBoost for prediction.
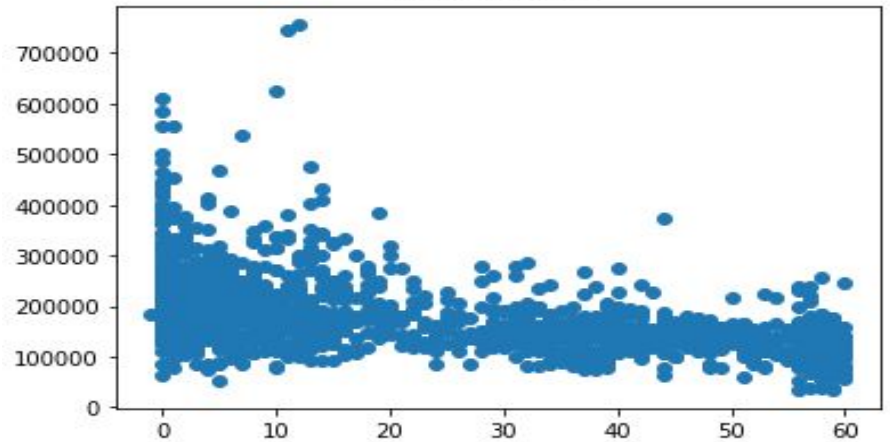
Datasets

We see that the house price is more for the house which is recently built and sold.
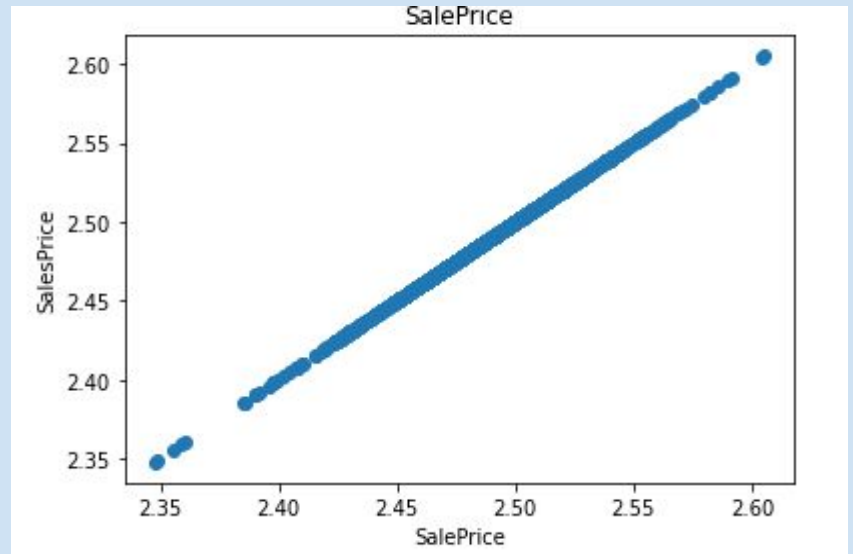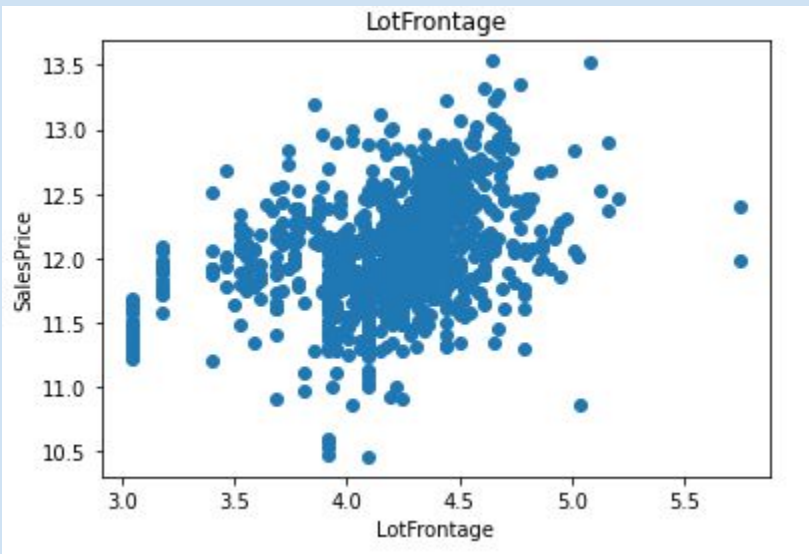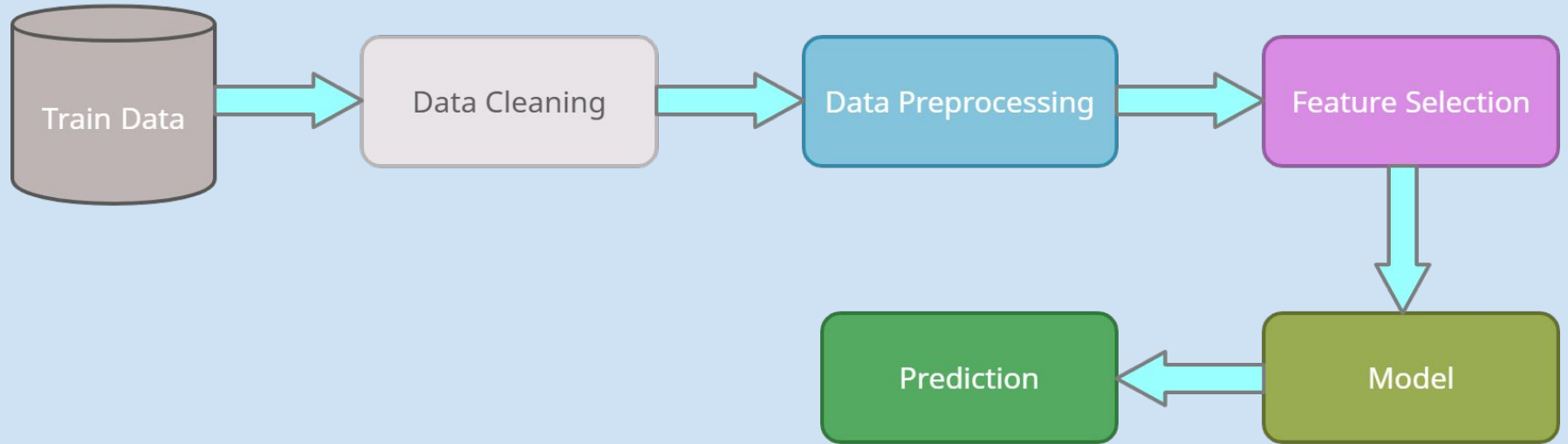
Datasets

Model Flow Diagram

Data Cleaning, Pre-Processing

- Filling the missing value we need to see the test and train data simultaneously.
- We will be replacing the null values with mode for categorical values, discrete numerical values and year variables.
- We will be replacing the null values with the mean for continuous numerical values.
- We will delete columns with more than 50% null values as the available information add no value for our model.

Data Cleaning, Pre-Processing...
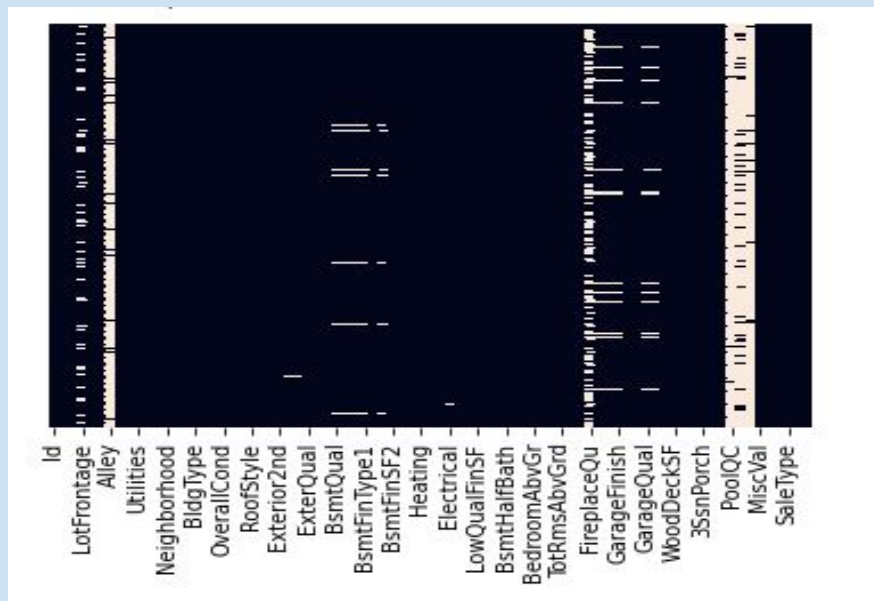
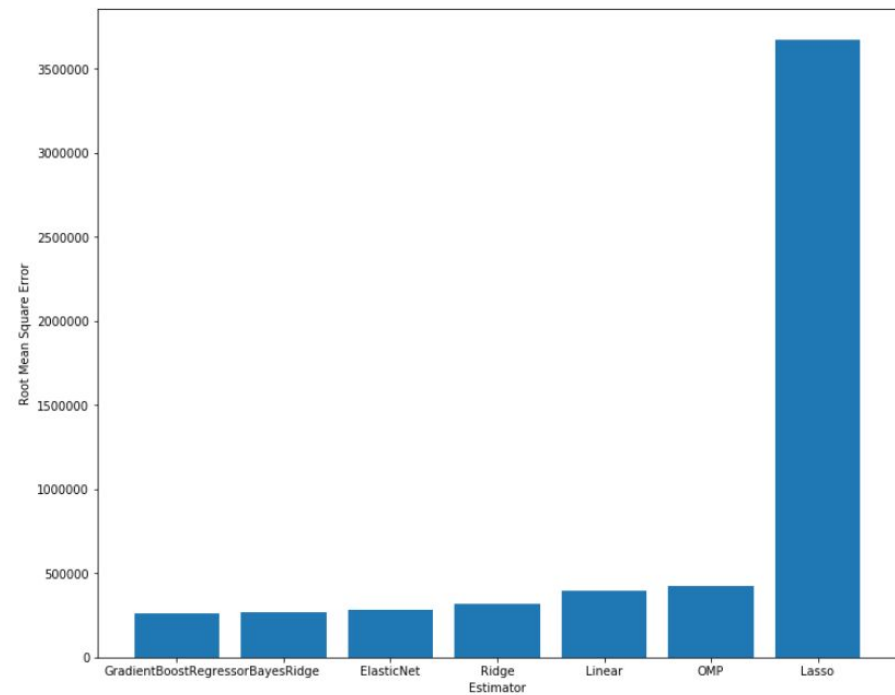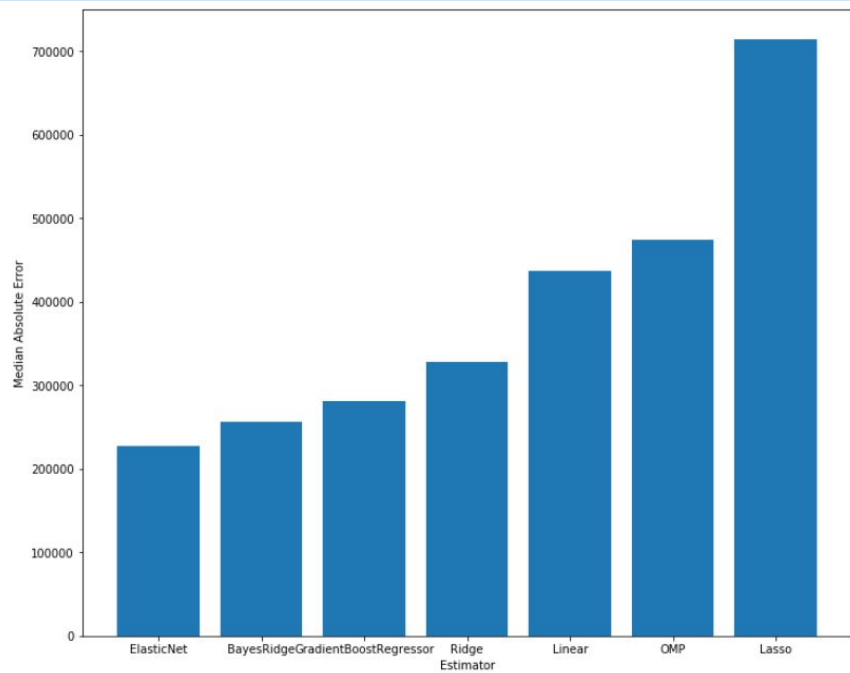Heatmap for visualizing the null values

Regression Models Used

- The Extreme Gradient Boost Regressor is chosen to build the model.

```
# implementing XGBoost regressor
import xgboost
classifier=xgboost.XGBRegressor()
classifier.fit(X_train,y_train)
```

Evaluation

Result

Using XGBoost Regression RMSE value is 0.13439.

Conclusion

This project is a predictive model to solve a regression problem of predicting house sales value.

We performed data cleaning, data preprocessing tasks before applying the model. We trained and evaluated a handful of regression models to choose the best applicable algorithm.