Manarat International University

Department of Computer Science and Engineering
Course: Artificial Intelligence

Report
On
House Price Prediction

Problem Title: House Prices: Advanced Regression Techniques.

Team Name: Techtrio
Team Member:

01. Name: Aysha Rahman
    ID: 1640CSE00472
02. Name: Sumaiya Raha
    ID: 1640CSE00483
03. Name: Fatima Tuj Jahra
    ID: 1640CSE00495

Kaggle Account: https://www.kaggle.com/techtrio
Git Repository: https://github.com/Techtrio2019

Project Goal:

The goal of this competition is to predict the sales prices of houses based on these variables. We collaborated on certain parts of the project and completed other parts individually as if it were a research project. The goal of this project, as aspiring data scientists, was to utilize our arsenal of machine learning knowledge to predict housing prices.

Data Description:

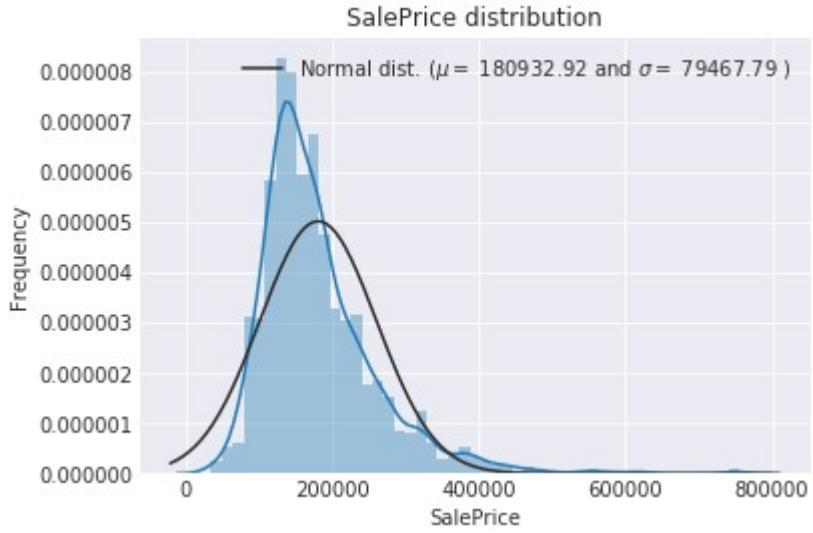
Training dataset contains 1,460 price training samples and 1459 testing samples.

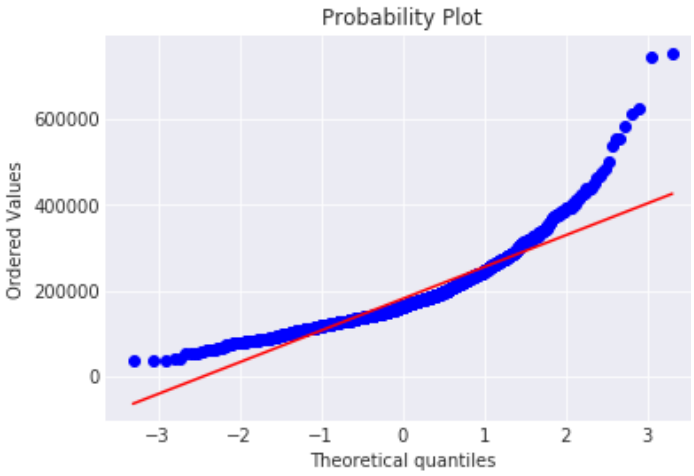
Data Preprocessing:

.

Target Variable:

SalePrice is the variable we need to predict. So let's do some analysis on this variable first.





The target variable is right skewed. As (linear) models love normally distributed data, we need to transform this variable and make it more normally distributed.

Data Correction:

Correlation map to see how features are correlated with SalePrice.