

Manarat International University
Department of Computer Science and Engineering
Artificial Intelligence (CSE – 411)
Assignment Milestone report

Problem Title: House Prices: Advanced Regression Techniques
Name of the Team: **rdnasim**
Contestants Name: Riadul Islam Nasim
Mahfuzur Rahman
Student ID: 1640CSE00467
1640CSE00519
Kaggle Account: <https://www.kaggle.com/rdnasim/>
Git Repository link: <https://github.com/rdnasim/House-Prices-Advanced-Regression-Techniques>

Problem statement:

The problem will be addressed if we build predictive models using Advanced Regression Techniques and train the model and pick the best model using validators so that it can accurately predict the value of House price.

The data set is obtained from Kaggle and it is a Competition Data set which contains 79 features which influences the price of Homes at Ames, Iowa. The competition organizer also boasts about the advantages of not just estimating house price using number of bedrooms or the fence around the house usually done by the brokers. Yeah, he is right, when you can accurately predict the house price using Advanced Regression techniques then why bother about a House Broker's price estimate?

We need a variety of per-processing steps before dealing with the Big Picture (i.e) Finding the Best Model. The below steps make sure that we are on track to achieve the prime target i.e Predicting the Sale Price

1. Data Cleaning, Outlier Analysis and Exploratory Data Analysis
2. Find the Correlation among features
3. Analyze all the features and perform log transform for the features
4. Data Visualization
5. Finding Most Important Features
6. Model building
7. Model Validation
8. Conclusion by submitting the code, findings and result

Technical Approach:

1. LASSO Regression
2. Elastic Net Regression
3. Kernel Ridge Regression
4. RandomForestRegression
5. Gradient Boosting Regression
6. XGBoost
7. LightGBM
8. RidgeRegression

9. SVR
10. StackingRegressor
11. LGBM Regressor

Expected Submission:

- Times : I will submit **13times**.
- Score : **0.10449** and that will be **99%** Accuracy
- Position : **100 – 200**