

Language: Python

Data tidying:

1. Removed outliers of column GridLivArea which are greater than 4000 sq ft. They are not needed.
2. Changed values of some particular columns directly in test data so that they don't make problem while using in prediction.
3. Manually converting columns like ExterQual, ExterCond and many more to numerical data.
4. Converted categorical data as textual data into numerical data using LabelEncoder.
5. Changing NA values to numerical values.
6. Converted some columns like 2ndFlrSF, OpenPorchSF and many others to 0 , 1 state before using one hot encoder.
7. Converted the all the values to normal distribution using skewness technique.
8. Then Scale the data using StandardScaler of Sklearn.
9. Converted Categorical numeric data to 0 and 1 forms so that models can understand those values better, using one hot encoder.
10. Dropped columns that are not in test data but exist in train data like _RoofMatl_Metal, _RoofMatl_Roll and many more.

Feature Engineering:

1. Converted multiple new features so that models can recognized these better like OverallQual to SimplOverallQual. It means that converting huge categorical data to more simpler data.
2. LotShape to IsRegularLotShape based on the assumption that if it is regular then it is more meaningful than other shapes. Same pattern is followed by multiple columns like
 - a. LandSlope to IsLandSlope
 - b. Electrical to IsElectricalSBKrEtc. Of course all above fields are to convert it to more binary form.

Models:

1. Used stacked regression technique in which I have used Lasso, Elastic and Kernel Ridge and Gradient Boosting. At the end I got the averages of them.
2. Also used XGBoost and LightGBM but later given very low weightage of them at the end.
3. Weightage technique is based on giving 75% weight to Average Stacked models and 12.5% and 12.5% to both XGBoost and LightGBM.

Training:

1. Training is based on average of 10 folds' cross validation.
2. Selected Lasso, Elastic and Kernel Ridge and Gradient Boosting for K-Folds training.