Machine learning

## PROJECT SUBMISSION 3

1. Did you do the same that was done in the references?

As per the previous submission, papers were presented which acted as a roadmap or a pipeline for this machine learning project.

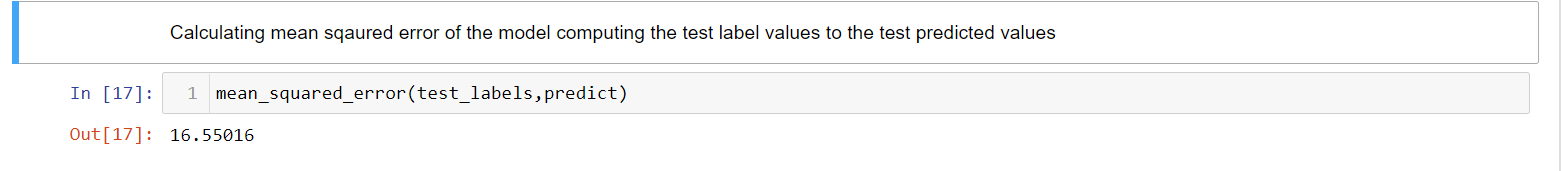
According to the papers submitted, the Model in it’s preprocessing and model training phases used Pre-processing methods such as Tokenization and Removal of stop words. There were various Feature extractions such as Positive sentiment words extraction and Negative words extraction. Feature Reduction methods such as Entropy was also proposed in the form Information Gain, where Information Gain was based on the measure of entropy.

1. Any enhancement to the work described in the references?

To further increase the efficiency of the model, steps were taken during the preprocessing and training stages.

Apart from the normal preprocessing methods applied various other libraries and functions were used to remove HTML Tags, Removal of punctuations, Removal of numbers and apart from just Positive sentiment words extraction and Negative words extraction Neutral words and compound values for the reviews have also been used. The Compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1(most extreme negative) and +1 (most extreme positive).

Mean Squared Error



Predictions of model

