­­­­­Week 8

I found that our model has testing accuracy around 46% ~ 48% in predicting the correct rating when the input is the concatenation of positive and negative reviews.

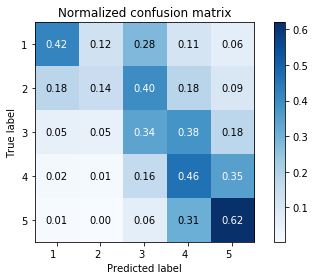
 Figure 1.

Figure 1 is the confusion matrix of testing data fitted into the model. We can see that rating 5 has the highest true positive rate (62%) compared to the true positive rates of rating 1, rating 3 and rating 4, which are below 50%. Rating 2 is the worst (14%), which implies that this model is terribly bad in correctly predicting reviews that are rated as 2. This model is also confused when it comes to predict reviews that are rated as 3 and 4. For instance, we can see that 34% of times our model will correctly predict reviews rated as 3 while 38% of times it will predict reviews with a rating of 3 as 4.