Methodology

In this approach, the main focus is on cleaning the text data. Removing the special characters, emojis, stop words, punctuationsetc which may not have any impact on the predcitions and stemming the words is done on the provided train data. Count vectorizer is used to create feature vectors from the text data and to train the model Logistic Regression is used.

Result Analysis

After predictions manually analysing the validation data that are predicted as toxic, shows model does a good work predicting the toxic comments. However the confusion matrix shows the bias towards predicting comments as non-toxic as we see good number of predictions are false negatives. This is the reason behind the **higher accuracy** of the model which is **94%** on the validation set comapared to the **much lower recall (45%) and precision (72%)** for toxic class.

Conclusion

This work is in progress.