Classification of Comments on the basis of Toxicity

by

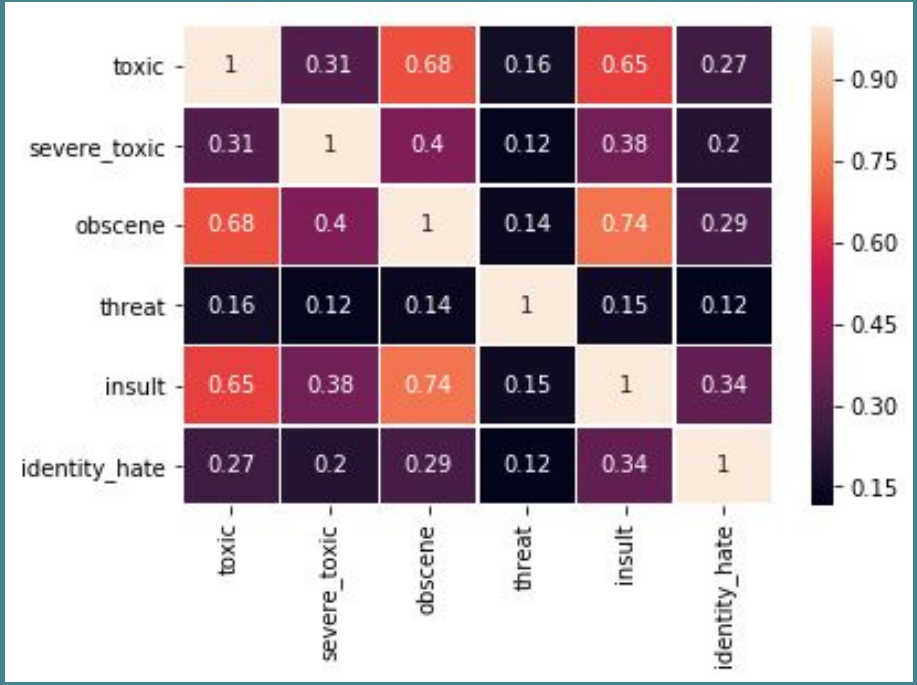
Ashish Jain (MT18052)
Shubham Gupta (MT18055)
Sarosh Hasan (MT18084)

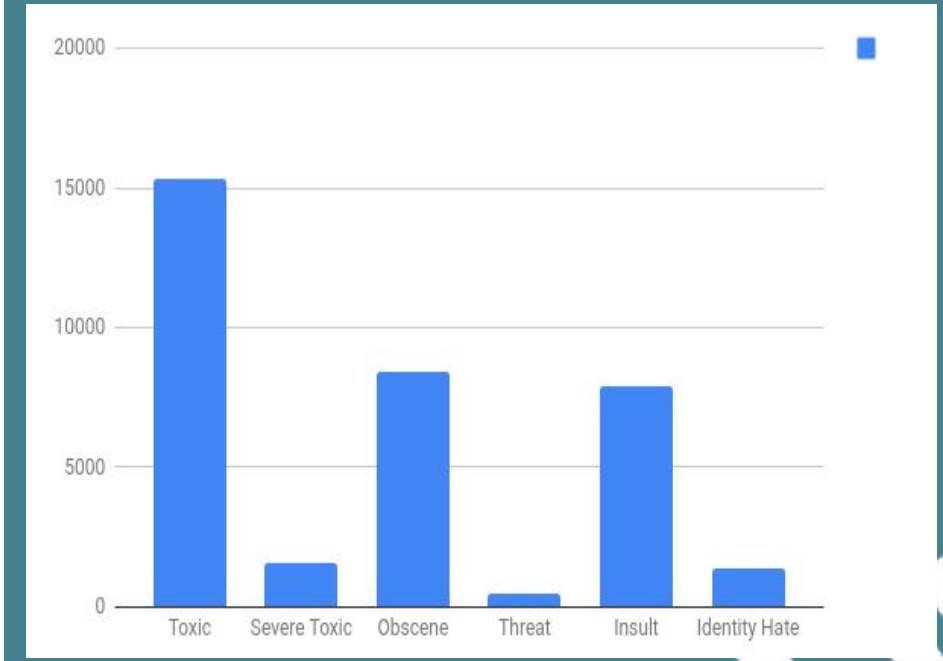
INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Motivation

To make discussion/review/comment section more approachable and readable by the users, As many users often do not prefer to contribute to such sections because of certain comments from other users that can hurt or offend one person.

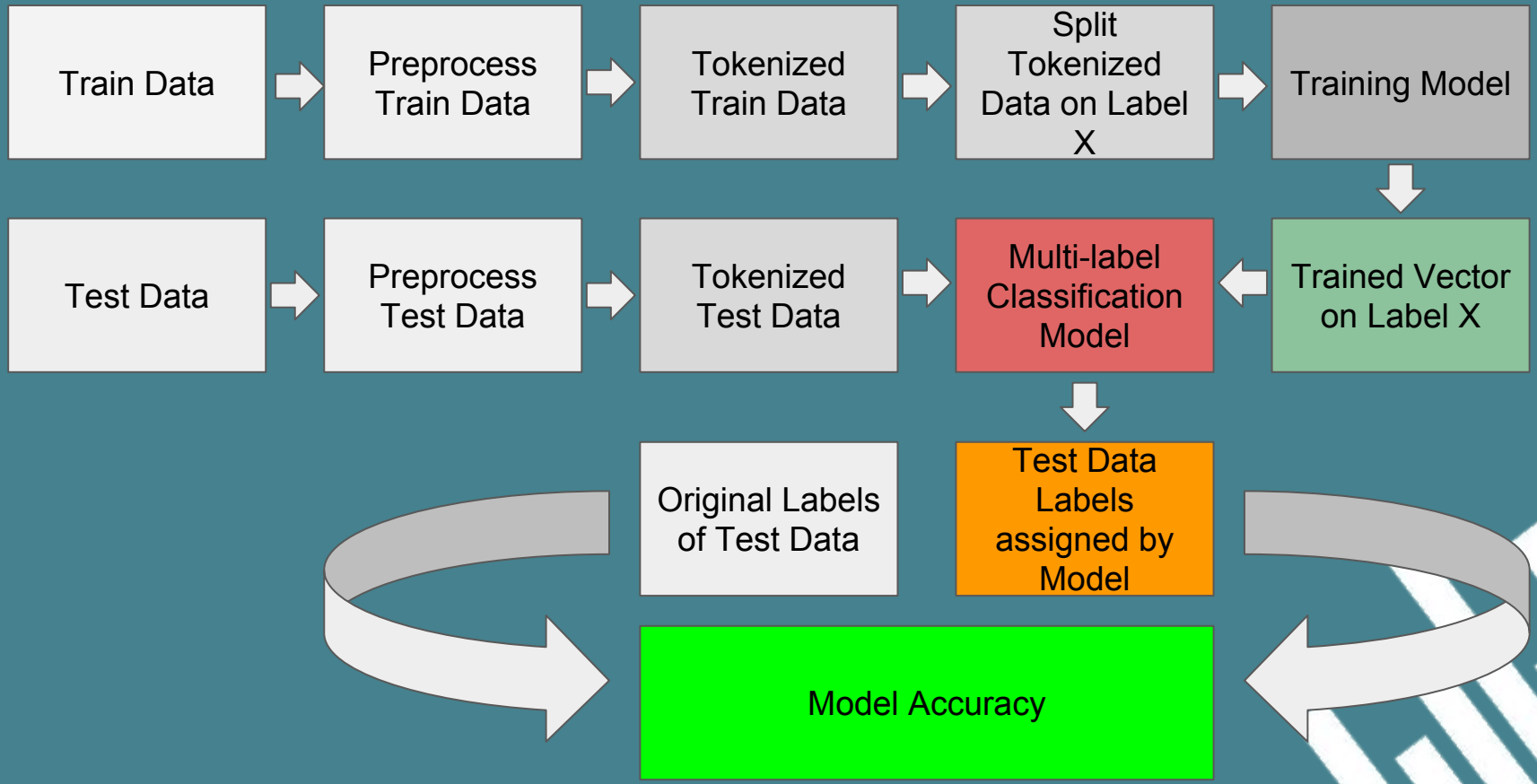
Visualisation of Data





Comments data set is taken from Kaggle.The data set contains 6 classes. Number of a comments in each class is shown in figure.

Proposed Model



Approach Used

Firstly we approach our problem using Weka :-

1.Used NaiveBayesMultinomialText -

->Used on Text data only.

->Give poor result with accuracy 17.25%

-->Maybe due to model trained bad on text data.

2. Used Multinomial Naive Bayes -

->on Numeric data generated from StringToWordVec.

->Give poor result with accuracy 38.9%

->Problem is that one comment can be the part of atmost six classes.So the model trains that comment in all classes (in which it belongs) and hence model is not getting trained properly.

Approach Used

From our previous Approach Problems we came to know that our data is a MultiLabel data and hence we apply MultiLabel classification Approach on dataset .
We use Problem Transformation method in our project. This method can be carried out in different ways as:

1. Binary Relevance
2. Classifier Chains.

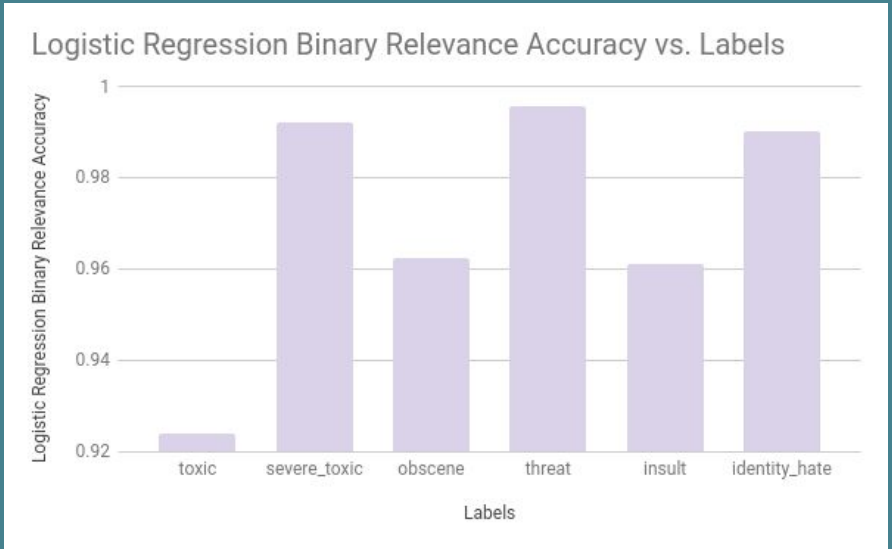
Approach Used

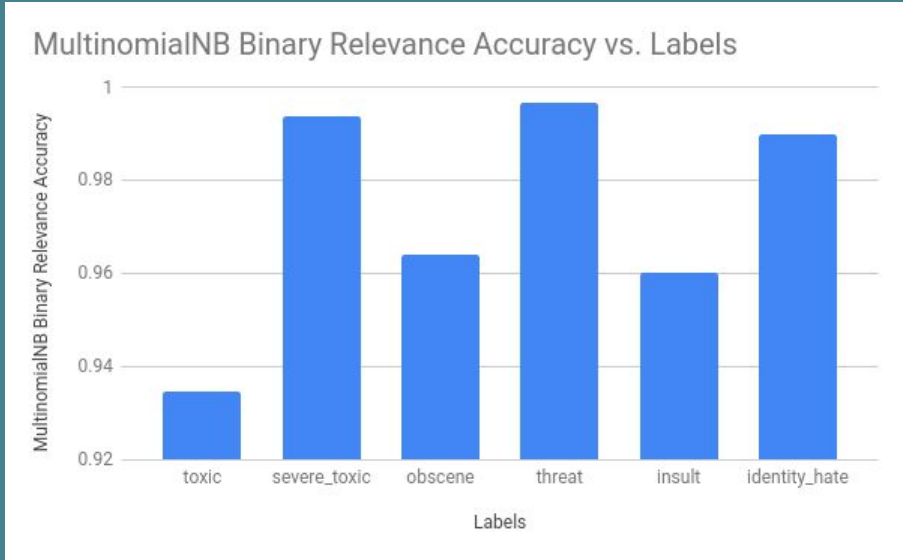
Models trained using the Problem Transformation techniques are-

1. Logistic Regression with Binary Relevance
2. Multinomial Naive Bayes with Binary Relevance
3. Logistic Regression with Classifier Chains
4. Multinomial Naive Bayes with Classifier Chains

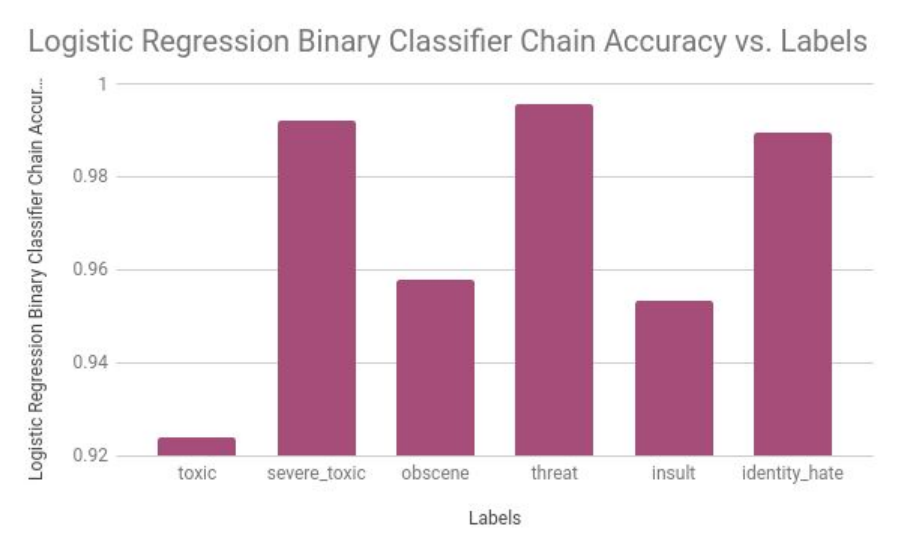
The accuracy of the model trained using this approach increases significantly as the choice of MultiLabel Classification works perfect.

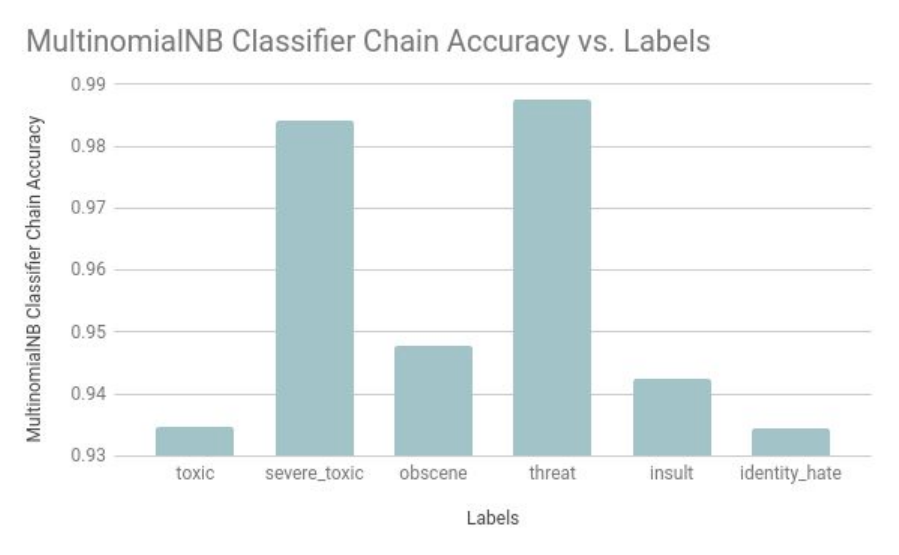
Accuracy of Binary Relevance Models





Accuracy of Classifier Chain Models





Conclusion

We used the techniques(Naive Bayes,logistic Regression) and tool(Weka) that we learned in Data mining course. We learned how to use these technique and tool in approaching a problem.

We came across different approaches and methodology that can be used to achieve our goal using data.

Our trained Model is able to classify a comment into a correct labeled class .

References

- https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/data
- http://blohttp://blog.datumbox.com/machine-learning-tutorial-the-naive-bayes-text-classifier/g.datumbox.com/machine-learning-tutorial-the-naive-bayes-text-classifier/
- https://www.datacamp.com/community/tutorials/text-analytics-beginners-nltk
- https://www.analyticsvidhya.com/blog/2017/08/introduction-to-multi-label-classification/
- https://www.kaggle.com/rhodiumbeng/classifying-multi-label-comments-0-9741-lb
- https://towardsdatascience.com/journey-to-the-center-of-multi-label-classification-384c40229bff