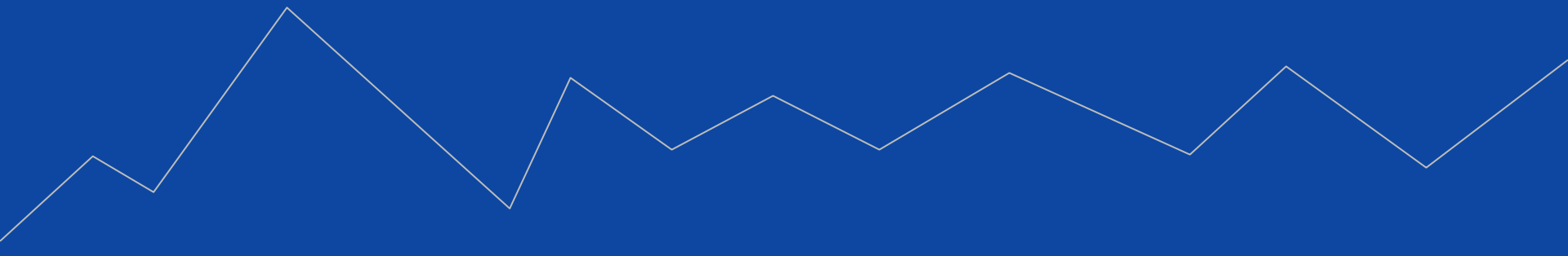
A COMPARATIVE STUDY TO DETECT ABUSIVE TWEETS



DETECTION OF OFFENSIVE CONTENT.. A NECESSITY?

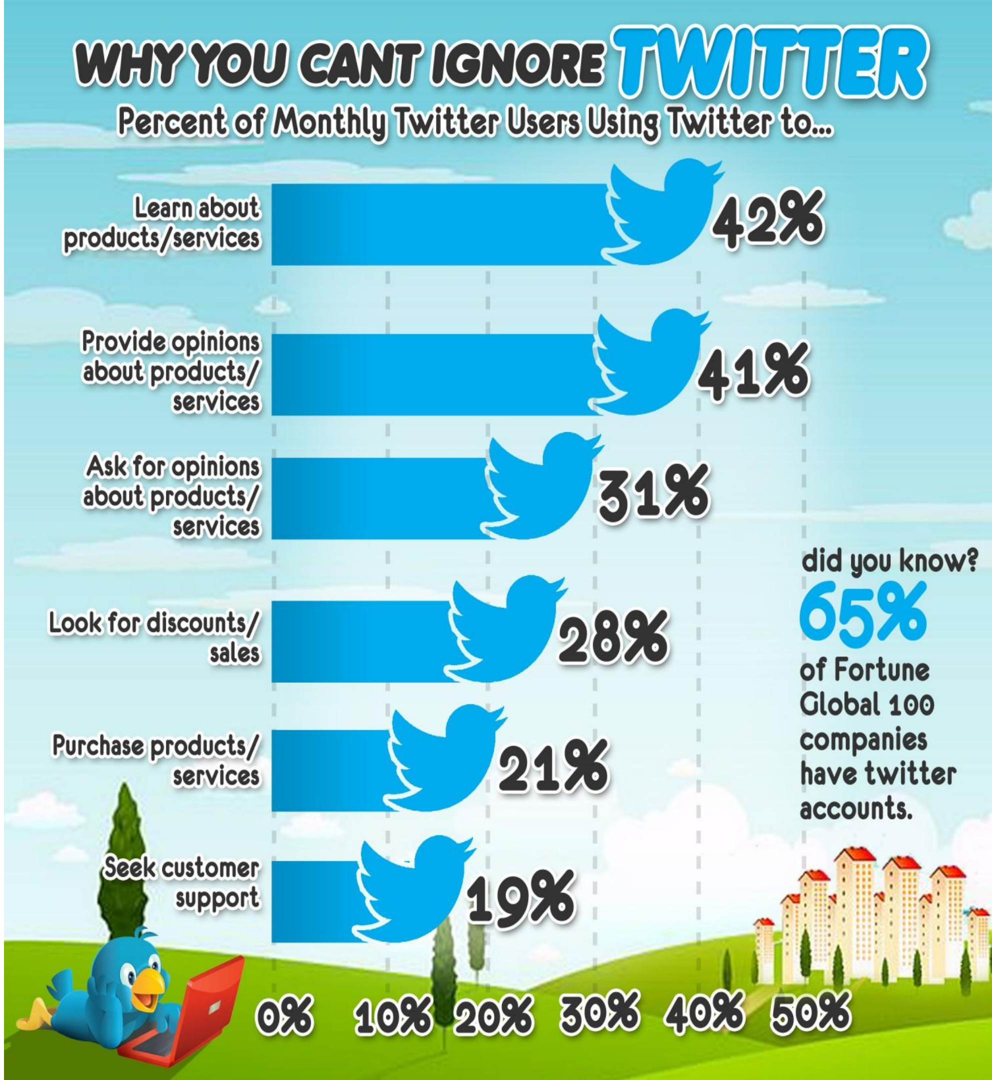
In Today's world, Social Network has become the main medium of communication and expressing views. Hence, to maintain the reliability and integrity of these sites it is necessary to detect and filter out offensive content.

WHY TWITTER ?

Since twitter was found, it had become an indispensable part of our lives. It is used for:

- Expressing views and opinions on major happenings.
- It acts as a medium of communication where people around the globe can reach each other.
- It is widely used for Marketing of products, movies, events etc.

SHOULD YOU USE TWITTER?
START HERE
You live in the 21st century
YES
You have access to the internet
YES
You like to learn & share
NO
NO
NO
YES
No, don't use Twitter
Yes, you should use Twitter!
NO
You have a growth mindset
NO
YES
NO
NO
Use Twitter with caution
YES
You live in a communist state
YES
You like to make international connections
@sylviaduckworth



WHY YOU CANT IGNORE TWITTER
Percent of Monthly Twitter Users Using Twitter to...
Learn about products/services
42%
Provide opinions about products/ services
41%
Ask for opinions about products/ services
31%
Look for discounts/ sales
28%
Purchase products/ services
21%
Seek customer support
19%
0% 10% 20% 30% 40% 50%
did you know?
65%
of Fortune Global 100 companies have twitter accounts.

WHAT DO WE MEAN BY OFFENSIVE CONTENT ?

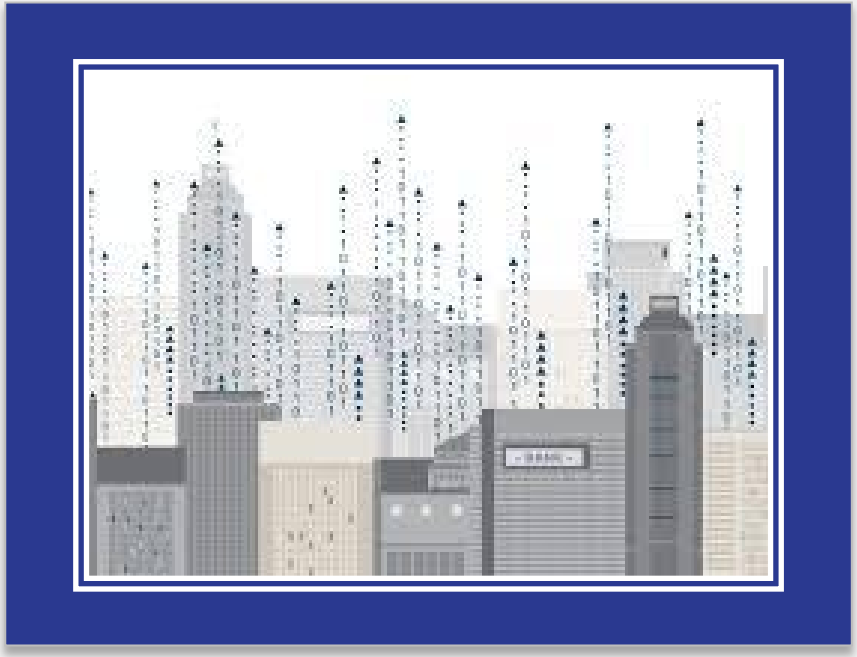
- For the purpose of this project, we employ Jay and Janschewitz (2008) definition of offensive language as vulgar, pornographic, and hateful language.
- Vulgar language refers to coarse and rude expressions, which includes explicit and offensive reference to sex or bodily functions; hateful language includes any communication outside the law that disparages a person or a group on the basis of some characteristic such as race, color, ethnicity, gender, sexual orientation, nationality, and religion.

EXISTING SOLUTIONS

- Implicit Content Filtering tools provided by various websites like Facebook, Twitter etc.
- Some parental control softwares are also present but they are not efficient and are paid like Appen.
- NO blocking of offensive content by twitter rather just blocking of users for you if someone finds content abusive.

CURRENT PROBLEMS

- VOLUME AND DIVERSITY OF CONTENT
- REAL TIME ANALYSIS
- DATA ACCESSIBILITY
- SHORT - LENGTH TEXT(MAX 140 CHARACTERS)
- SPELLING MISTAKES
- SLANG WORDS
- NO PREDEFINED CONTEXTUAL CLASSIFIER



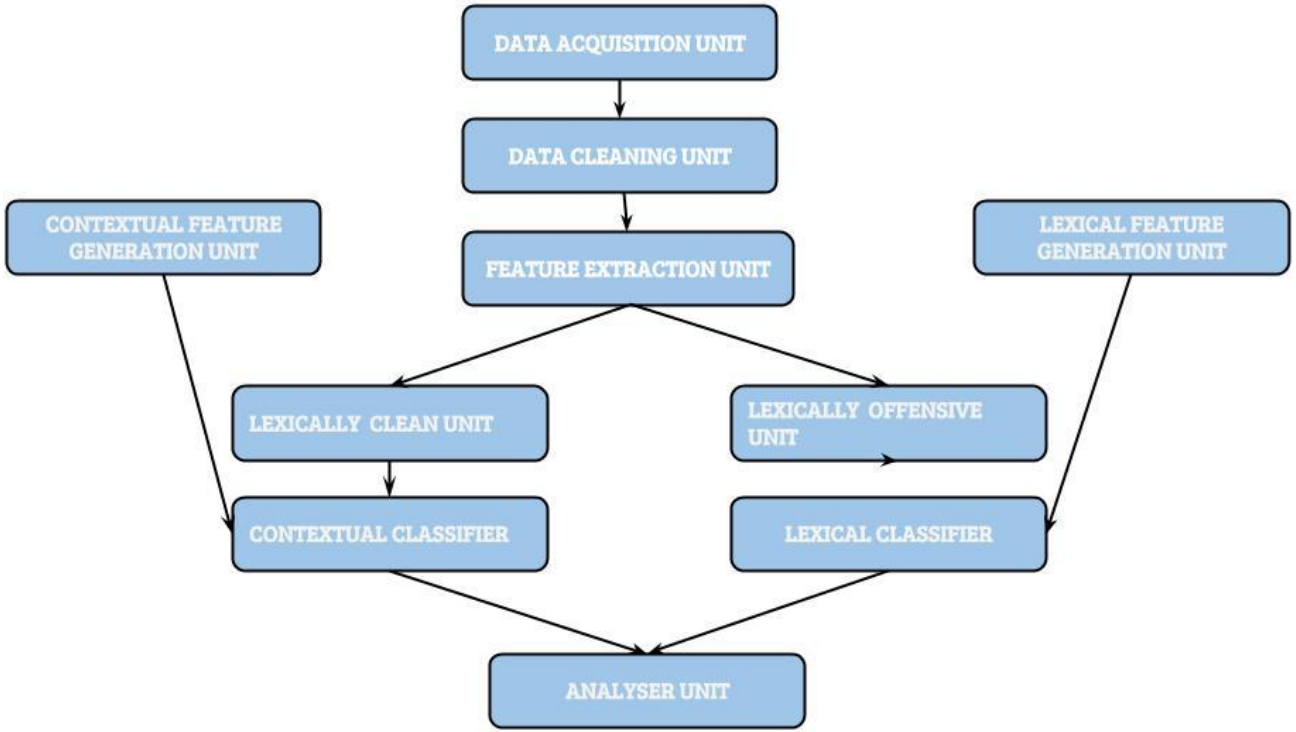
PROPOSED SOLUTION

- 3 Level Architecture.
- Take in consideration lexical as well as contextual as well as lexical aspect of tweet
- Extensive cleaning of raw tweets to reduce processing time.
- New Hybrid Model of LDA and Naive Bayes is created.
- New Rule Based Approach for Naive Bayes.
- Development of new approach for context based classifier -ADJECTIVE BASED APPROACH
- Comparative analysis of newly devised approach with existing COSINE_SIMILARITY.

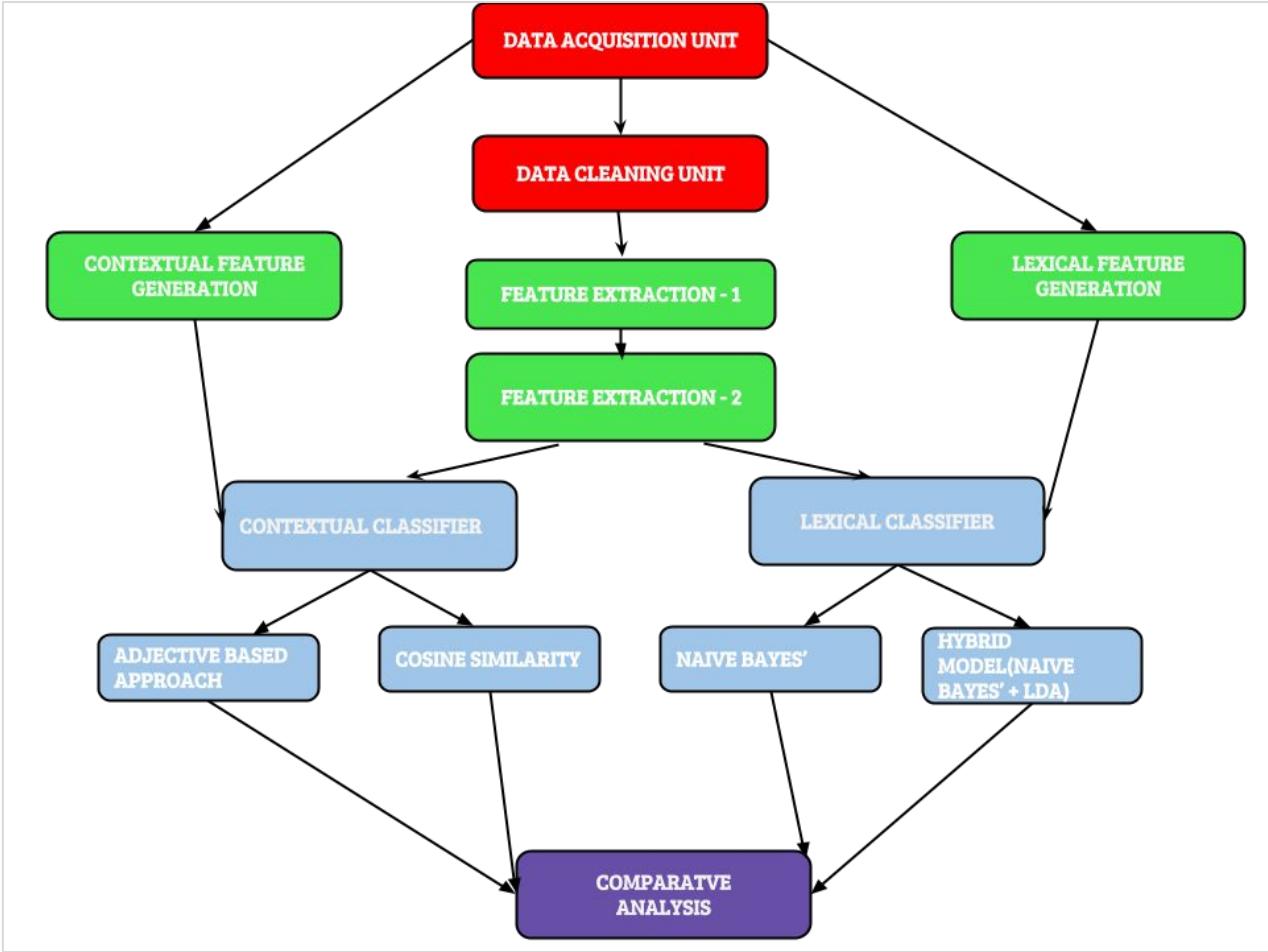
BENEFITS OF PROPOSED SOLUTION

- Extensive Preprocessing by removing Slang words and spelling correction to standardize data.
- Focuses on both Contextual and Lexical Aspects of data
- Comparative analysis is done for both the aspects and thus provides the best technique.
- Techniques intended to provide High Precision and Recall.
- Also, because of extensive preprocessing, the processing time has reduced many folds.
- Tools like apache spark used to reduce processing speed.

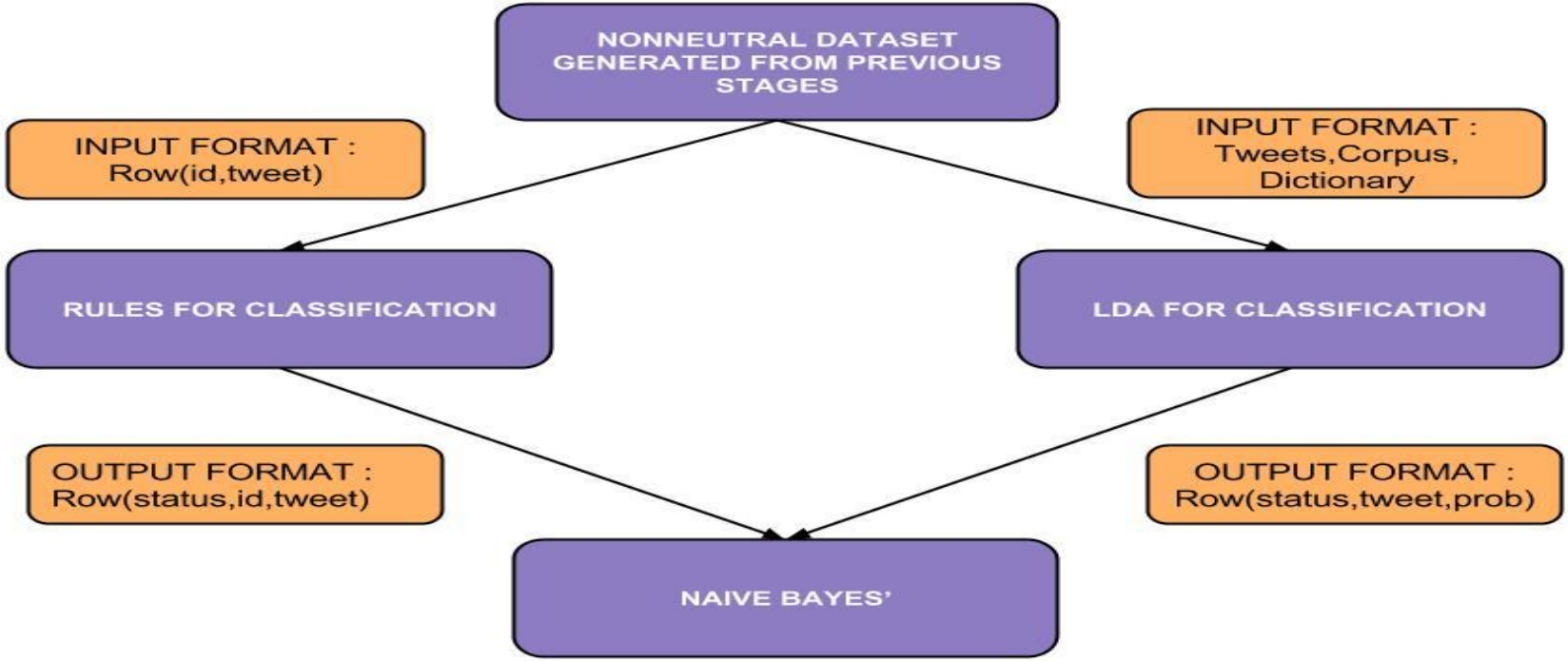
ARCHITECTURE



ALGORITHM

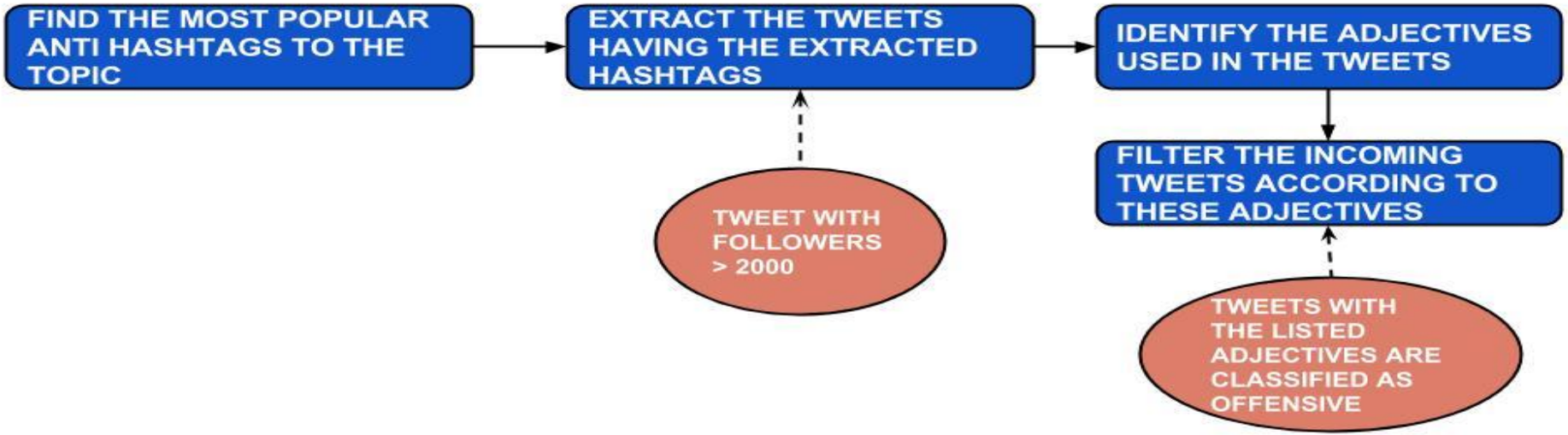

DATA ACQUISITION UNIT
DATA CLEANING UNIT
CONTEXTUAL FEATURE GENERATION
FEATURE EXTRACTION - 1
LEXICAL FEATURE GENERATION
FEATURE EXTRACTION - 2
CONTEXTUAL CLASSIFIER
LEXICAL CLASSIFIER
ADJECTIVE BASED APPROACH
COSINE SIMILARITY
NAIVE BAYES'
HYBRID MODEL(NAIVE BAYES' + LDA)
COMPARATVE ANALYSIS

HYBRID MODEL OF LDA AND NAIVE BAYES'



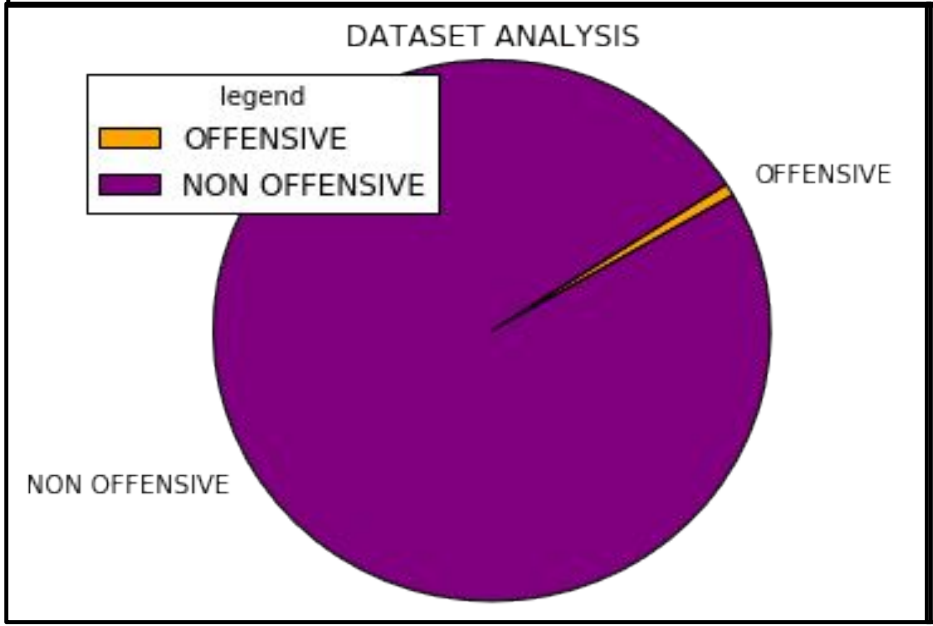
Where Status indicated the 2 clusters formed and Corpus represents Doc-TermId matrix

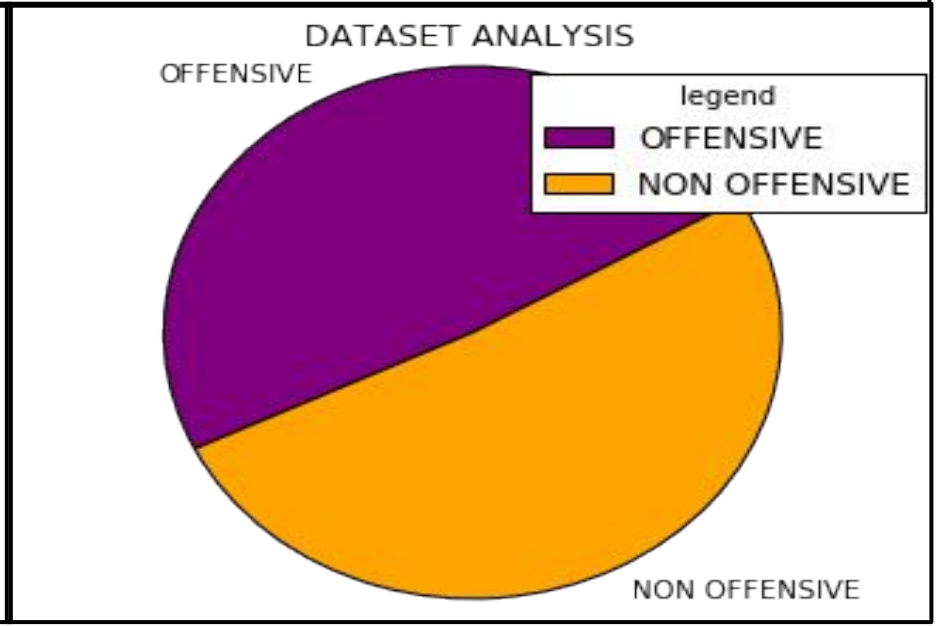
ADJECTIVE BASED APPROACH



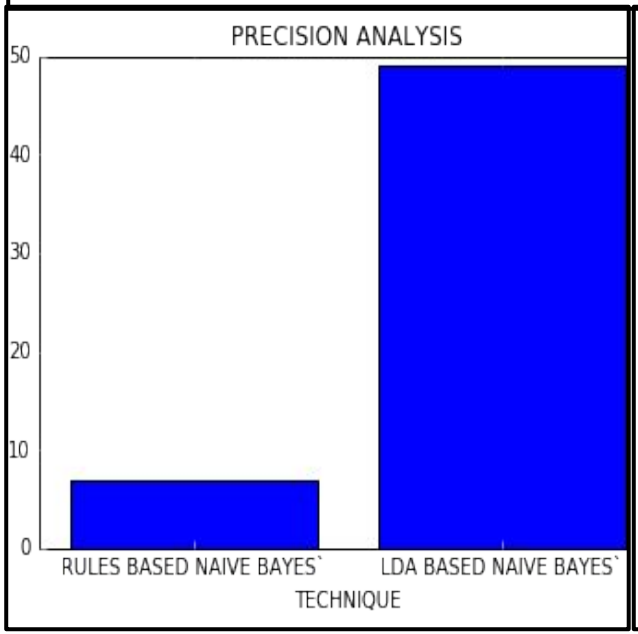
COMPARATIVE ANALYSIS OF LEXICAL CLASSIFIERS

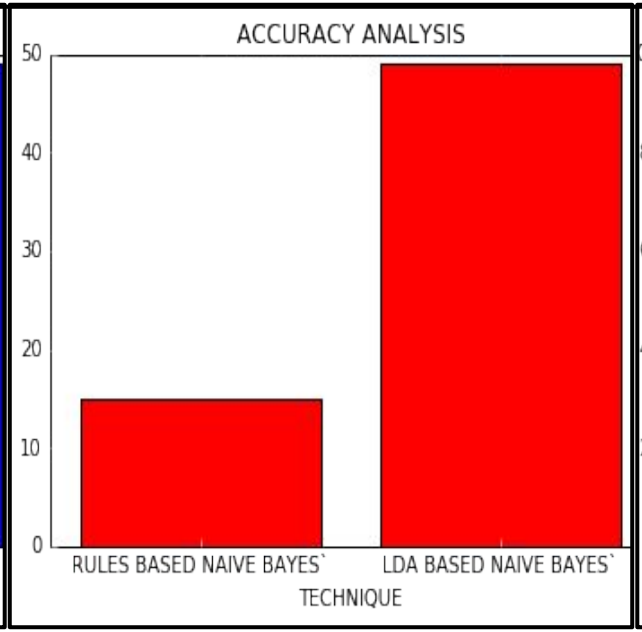
DATASET ANALYSIS

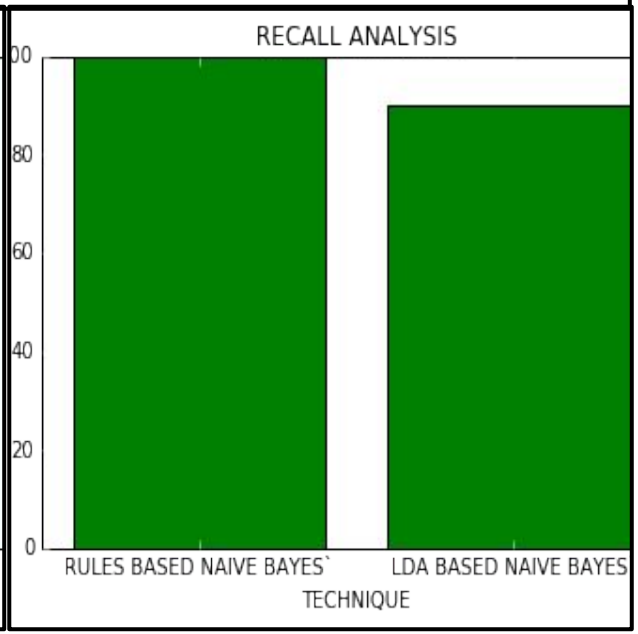




PERFORMANCE ANALYSIS

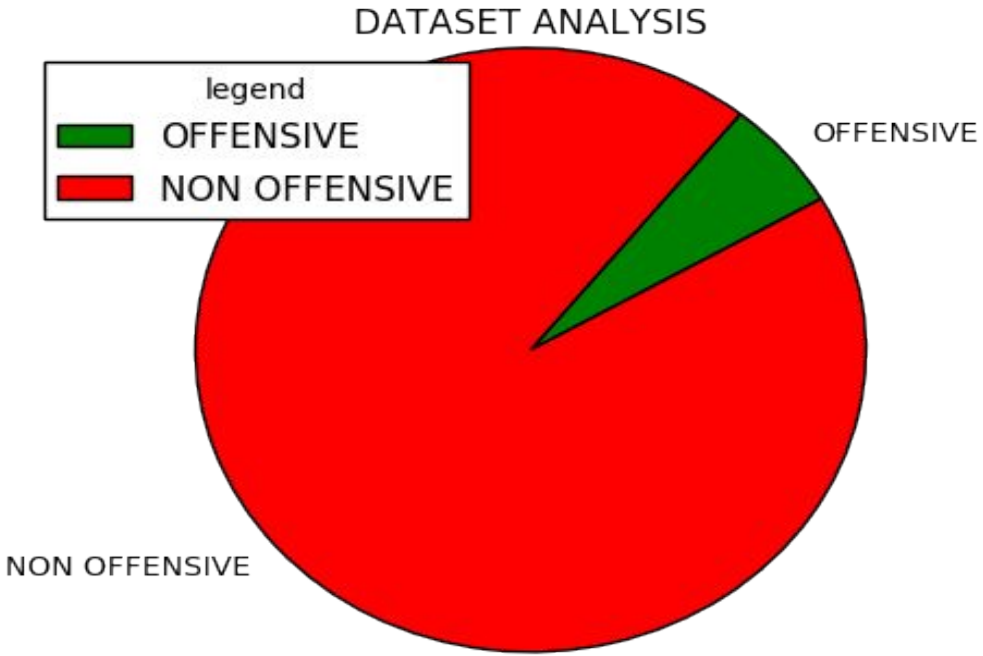




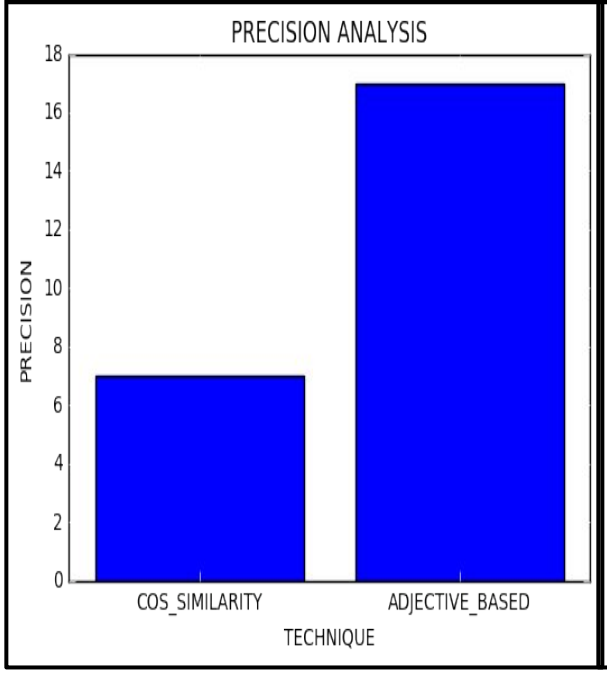


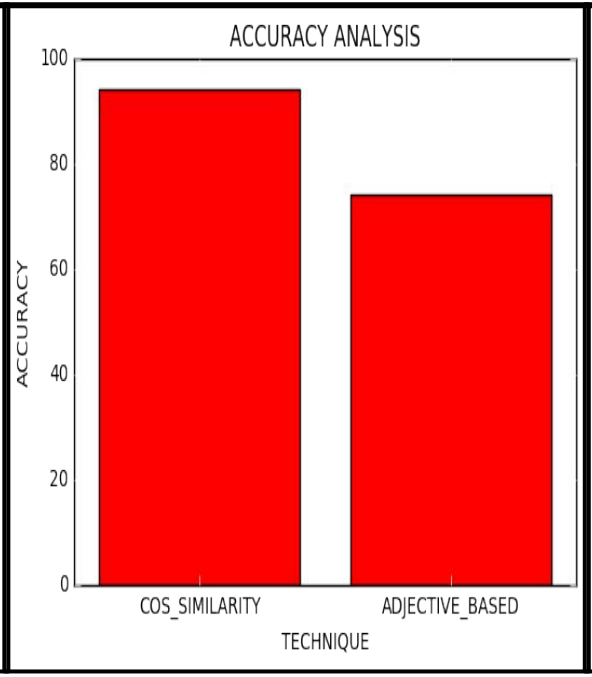
COMPARATIVE ANALYSIS OF CONTEXT BASED CLASSIFIER

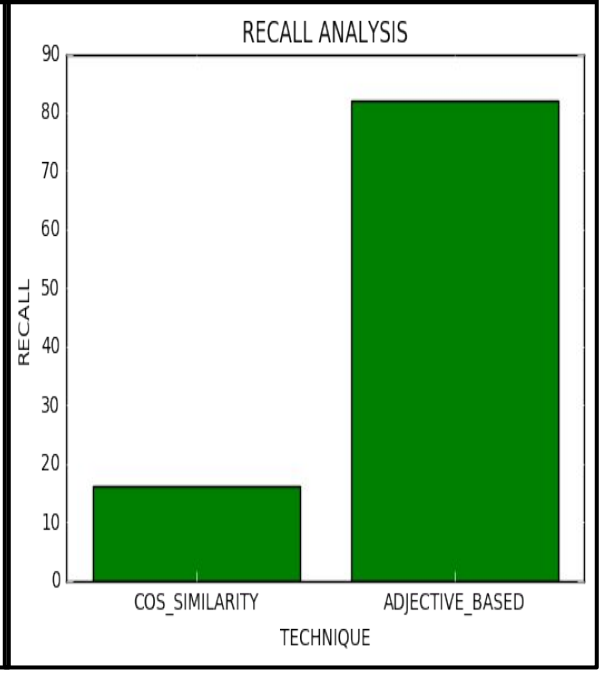
DATASET ANALYSIS



PERFORMANCE ANALYSIS







CONCLUSION

LEXICAL CLASSIFIERS

- it was found that newly created Hybrid Model of LDA and Naive Bayes is better than the newly devised Rule Based Naive Bayes approach in terms of precision and accuracy.
- Main reason for the failure of Rule based Naive Bayes approach is that it only takes into consideration adjectives and proper nouns.
- Hybrid Model of LDA and Naive Bayes approach was the showstopper because the clusters formed by LDA are more detailed and systematic and also takes into consideration the frequency of words in the document.

CONTEXT BASED CLASSIFIER

- It can be concluded that the adjective based approach is better for context based analysis of tweets when the offensive content volume is low in the test dataset.
- The main reason for this is as the adjective based approach only takes in consideration the most important part of speech of text for classification i.e adjectives, it will always perform optimally irrespective of the volume of the offensive content in the test data.