e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

AUTOMATE SENTIMENTAL ANALYIS OF TEXTUAL COMMENTS AND FEEDBACK

A. Ramesh*1, CH. Vachan sai*2, G. Balasai Sesha Simha*3, N. Bala Sainath*4

*1Associate Professor, CSE Department, Ace Engineering College, Hyderabad, Telangana, India

*2Student,CSE Department, Ace Engineering College, Hyderabad, Telangana, India

*3Student,CSE Department, Ace Engineering College, Hyderabad, Telangana, India

*4Student,CSE Department, Ace Engineering College, Hyderabad, Telangana, India

DOI: https://www.doi.org/10.56726/IRJMETS-NCASCTE202212

ABSTRACT

Now a days everything is Automated. We watch movies and most of us give the feedback about the movie and also, many people believe in the reviews given in the review websites like Twitter and so on. The reviews may be going towards good things or bad things, some people say good some say, bad but knowing overall true review is difficult. Some people may use irrelevant words, unwanted data also. This should not happen. Reading all the reviews is difficult, and finding the relevant words about the movie is also difficult. So, if we can make this easy to the audience (People) it would be very nice and it would be great thing. So, we collect the reviews preprocess it so that redundancies are removed and data becomes consistent. Then, vectorizing the content takes place so that model can easily process the data. Now, we start building the model and we split the dataset to train and test the model accuracy. To analyze the model results we will have the accuracy plot, we also try to make a function which takes text as an input and gives the sentiment of the review (positive or negative). At last, we can see the reviews sentiment using Data Visualization by Using Naïve Baye's classifier we have achieved 94% accuracy.

Keywords: SentimentalAnalysis,NaïveBaye'sClassifier,Machinelearning,ConfusionMatrix,Vectori

I. INTRODUCTION

Automate sentiment analysis of textual comments and feedback, is a project where I need to do the analysis and predict the sentiment of the sentence.

Here the dataset used is Twitter tweets reviews and the method using Naïve Bayes classifier.

The prediction should be done for the given review and should name it as positive or negative.

The opinions of others have a significant influence in our daily decision-making process. These decisions range from buying a product such as a smart phone to making investments to choosing a school—all decisions that affect various aspects of our daily life. Before the Internet, people would seek opinions on products and services from sources such as friends, relatives, or consumer reports.

The Internet and the Web have now (among other things) made it possible to find out about the opinions and experiences of those in the vast pool of people that are neither our personal acquaintances nor well-known professional critics — that is, people we have never heard of. And conversely, more and more people are making their opinions available to strangers via the Internet.

The interest that individual users show in online opinions about products and services, and the potential influence such opinions wield, is something that is driving force for this area of interest. And there are many challenges involved in this process which needs to be walked all-over in-order to attain proper outcomes out of them.

II. METHODOLOGY

Importing Libraries

We import required libraries like nltk (natural language tool kit), pandas ,numpy re (regular expression) , seaborn, matplotlib.

Import Data

We import data as a csv file from the social media platforms like twitter, facebook,imdb website etc.,

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

Here, we taken the tweets from the twitter with racist and non-racist tweets.

Labelling the data

We label the data as 0 and 1 to differentiate between the data referring the required output like positive, negative.

Here label 0 for racist tweets and label 1 racist tweets

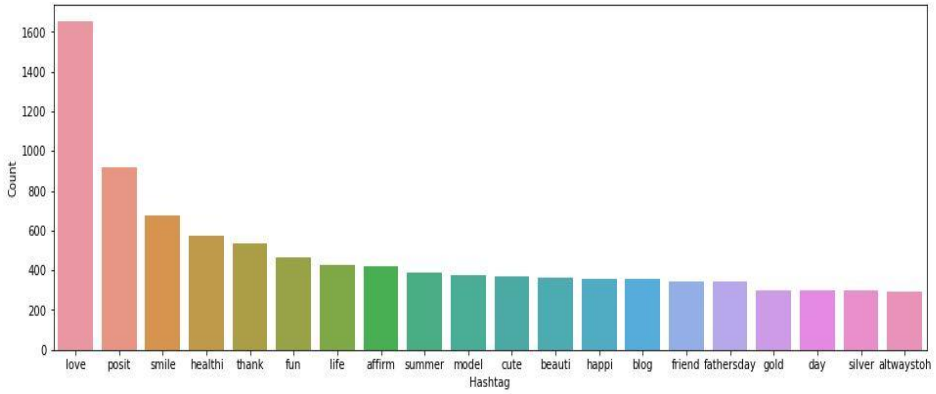
Training and Testing

We train the data collected from the outside to csv fail and we test the data we follow 0.8 : 0.2 format of train and test data

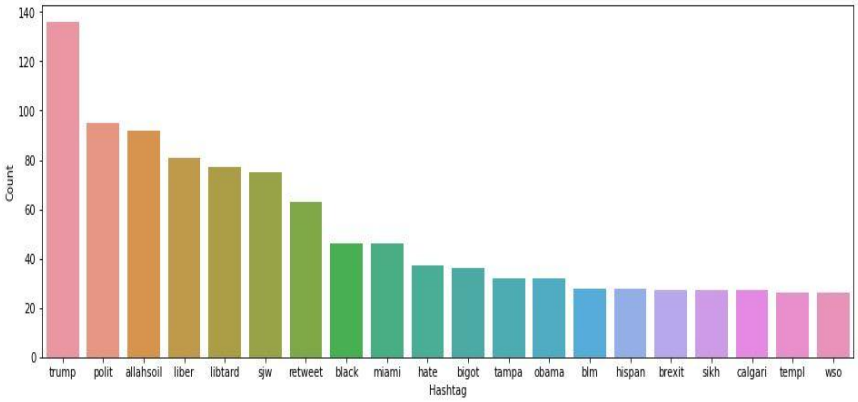
We get the label 0 and label 1 values

For our data we got 28000 label 0 values, 2200 label values

We calculate length of the train, test data sets and check in a histogram like this:



Racist Tweets



Non-Racist Tweets

Removing twitter handle, emojis and punctions

By using string functions and vectorization techniques we remove twitter handle, emojis and punctions

Then we remove small words to get accuracy (words with length less than 4)

Tokenization

We split the words with white spaces and store them in the form of lists.

Then we create word clouds to check which words are highly used and we get the hash tags.

Stemitization

We convert all the words into their root words so that we can easily get the frequency of the words

Then we split the train data and test data

Applying Multinomial Naïve Bayes Classifier

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

Now we apply the multinomial naïve bayes classifier to fit the train data and test the data by giving the train and test data and get the out put with the help of confusion matrix

Confusion matrix

It has actual values and predicted values which are as follows:

True actual value and correctly Predicted

True actual value and Wrongly Predicted

False actual value and correctly Predicted

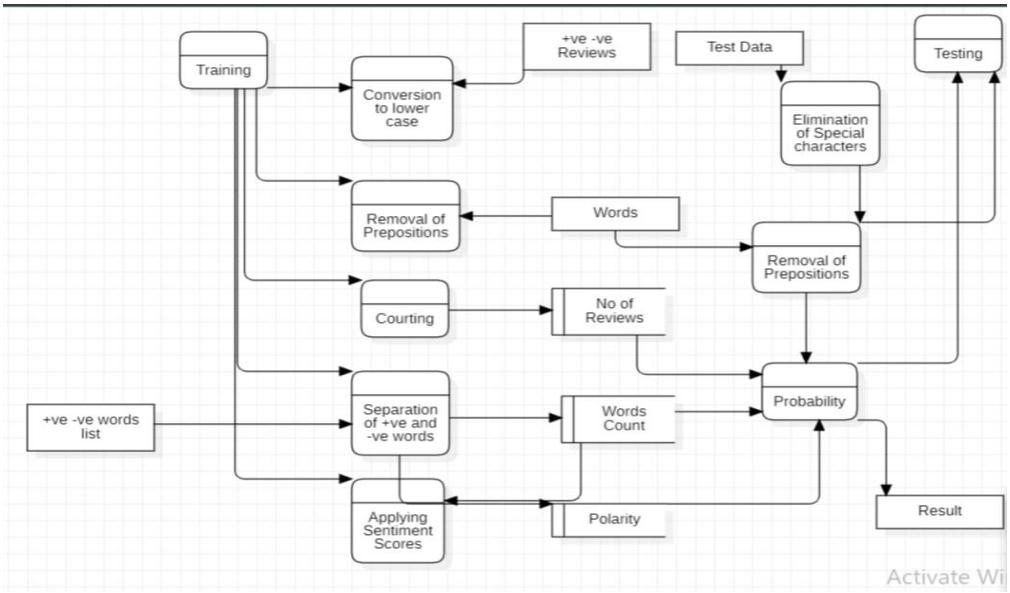
False actual value and correctly Predicted

Then we can check the accuracy of our model

III. MODELING AND ANALYSIS

We Can Depict the model in the form of some data flow diagrams and uml diagrams

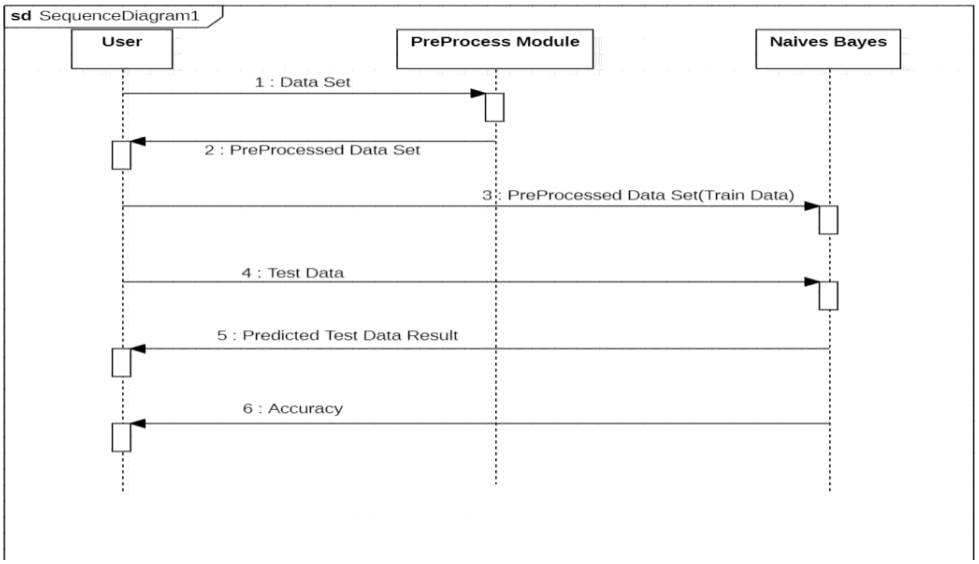
DFDs



UML Diagrams

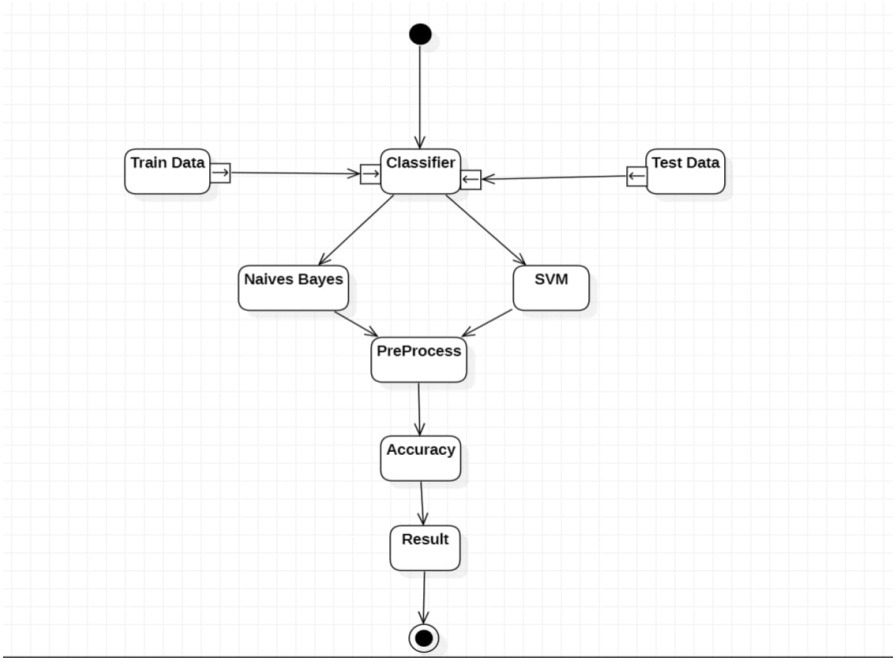
We show some the diagrams

Sequence Diagram



e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

Activity Diagram



IV. RESULTS AND DISCUSSION

Collection Of Twitter Raw Data And Coverting Into Csv File.

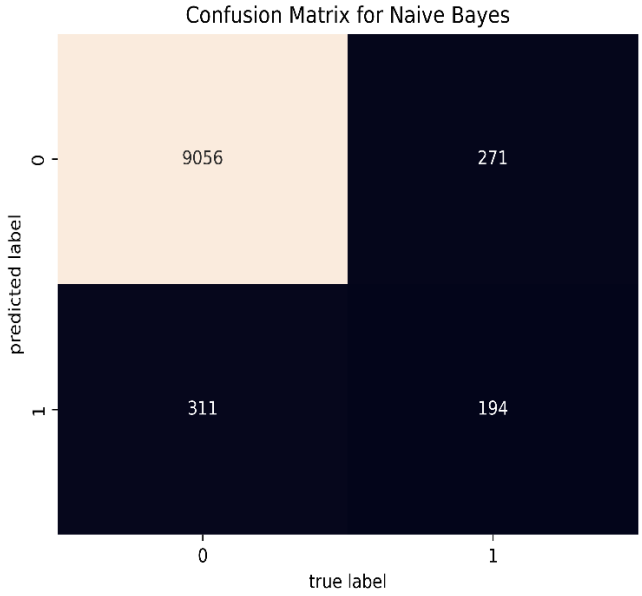
```
id,tweet
31963,#studiolife #aislife #requires #passion #dedication #willpower  to find #newmaterialsâ•¦
31964, @user #white #supremacists want everyone to see the new â••  #birdsâ•• #movie â•• and hereâ••s why
31965,safe ways to heal your #acne!!    #altwaystoheal #healthy  #healing!!
31966,"is the hp and the cursed child book up for reservations already? if yes, where? ð•••ð•••ð•••  #harrypotter
#pottermore #favorite"
31967,"  3rd #bihday to my amazing, hilarious #nephew eli ahmir! uncle dave loves you and missesâ•¦ "
31968,choose to be   :) #momtips
31969,something inside me dies ð••¦ð••¿â•"  eyes ness #smokeyeyes #tired  #lonely #sof #grungeâ•¦
31970,#finished#tattoo#inked#ink#loveitâ•¤ï¸.• #â•¤ï¸.•â•¤ï¸.•â•¤ï¸.•â•¤ï¸.• #thanks#aleeee !!!
31971, @user @user @user i will never understand why my dad left me when i was so young.... :/ #deep #inthefeels
31972,#delicious   #food #lovelife #capetown mannaepicure #resturantâ•¦
31973,1000dayswasted - narcosis infinite ep.. make me aware.. grinding neuro bass #lifestyle
31974,one of the world's greatest spoing events   #lemans24 #teamaudi
31975,half way through the website now and #allgoingwell very
31976,"good food, good life , #enjoy and  ð•••ð•••ð•••ð•••ð•••ð••• this is called ~garlic bread~ ... #iloveitâ•¦ "
31977,i'll stand behind this #guncontrolplease  #senselessshootings #taketheguns #comicrelief #stillsad
31978,"i ate,i ate and i ate...ð•••ð•••  #jamaisasthi #fish #curry #prawn #hilsa #foodfestival #foodies "
31979, @user got my @user limited edition rain or shine set today!!  ! @user @user @user @user
31980,&amp; #love &amp; #hugs &amp; #kisses too! how to keep your #baby     #parenting #healthcare
31981,"ð•••ð•••ð••• #girls   #sun #fave @ london, united kingdom "
31982,thought factory: bbc neutrality on right wing fascism  #politics #media #blm #brexit #trump #leadership &gt;3
31983,hey guys tommorow is the last day of my exams i'm so happy yay
31984, @user @user  @user  #levyrroni #recuerdos memoriesð•••â•¤Ïð•••ð•••«ð••• #recuerdos    #friends #life #triunfodelamor
31985,my mind is like ð•••ð•••ð••%ð••• but my body like ð•••ð••´ð••µð••¾....   #sleepy #stillallinð•••
31986,never been this down on myself in my entire life.
31987,"check  twitterww - trends: ""trending worldwide 11:14 am bst""1. #oscarpistorius2. #diplomalÄtliselilerayakta3.  â•¦"
31988,i thought i saw a mermaid!!! #ceegee  #smcr   #inshot #girls #cute #summer #blur #sun  â•¦
31989, chick gets fucked hottest naked lady
31990,happy bday lucyâ•¨¨â•¨¨ð••• xoxo #love #beautiful #pizza   #instagood #mileycyrus #demilovatoâ•¦
31991, haroldfriday have a weekend filled with sunbeams everyone!   #healthy #weekend
31992,"@user @user tried that! but nothing - will try again! know you loved #2, but the 3rd light my fave &amp;deep rivers  "
31993,i'll #never be #120 #again i'm   #i'm a #thick #women #blacktina
31994, @user new @user episode tk! and a really cool 2-year anniversary ep in the works for august. :o)
31995,#orangechicken   attack bull game 3d: do you really think that his head was empty around the city. each side
31996,suppo the #taiji fisherman! no bullying! no racism! #tweet4taiji #thecove #seashepherd
31997,i say we because i'm speaking collectively. i've always known. 2016 showed a lot.  andâ•¦
```

The above data represents the tweets of racist and non-racist comments'

Running program code into jupyter notebook in order to generate the output of an analyzed data into graphical representation.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

Using Confusion Matrix In Order To Predict Exact Data Acuuracy




```
In [50]:
from sklearn.metrics import accuracy_score

score_naive = accuracy_score(predicted_naive, y_test)
print("Accuracy with Naive-bayes: ",score_naive)
```

```
Accuracy with Naive-bayes:  0.9407038242473555
```

We have achieved

Accuracy 94% with the help of Naïve-Bayes Classifier.

V. CONCLUSION

We have used Naïve bayes Classifier and Confusion Matrix to predict the Both Positive and Negative reviews of racist comments as an Outcome by analyzing tweets from the raw random collected Data from Twitter In order to achieve the at most Accuracy.

Where we have gained the accuracy of 94% predicting the review of both, Positive and Negative Comments tweets.

VI. REFERENCES

[1] Shiv Dhar, SuyogPednekar, KishanBorad- Methods for Sentiment Analysis Computer Engineering, VIVA Institute of Technology, University of Mumbai, India.

[2] RavinderAhujaa, AakarshaChuga, Shruti Kohlia, Shaurya Guptaa, Pratyush Ahujaa- The Impact of Features Extraction on the Sentiment Analysis Noida, India.

[3] Richa Mathur, Devesh Bandil, Vibhakar Pathak-Analyzing Sentiment of Twitter Data using Machine Learning Algorithm.

[4] Alec Glassford and Berk Coker Multiclass Emotion Analysis of Social Media Posts Stanford University.

[5] Kavya Suppala, Narasinga Rao "Sentiment Analysis Using Naïve Bayes Classifier" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8 June, 2019.64

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Conference:01/December-2022 Impact Factor- 6.752 www.irjmets.com
1st National Conference on Applications of soft Computing Techniques in Engineering NCASCTE-2022
Organized by Department of Electrical & Electronics Engineering, ACE Engineering College, Hyderabad

[6] Bhagyashri Shinde , N.R.Wankhade" Sentiment Analysis on Twitter Data using Naive Bayes" International Journal of Advanced Research in Computer and Communication Engineering ISO 3297: 2007 Certified

[7] Introduction to Machine Learning Alex Smola and S.V.N. Vishwanathan Yahoo! Labs Santa Clara –and– Departments of Statistics and Computer Science Purdue University –and– College of Engineering and Computer Science Australian National University