**Major Assignment 1**

**Twitter Sentiment Analysis Report**

**Roopesh Bharatwaj K R**

**A00279933**

**1. Problem Statement**

Perform Twitter Sentiment Analysis , the target of the problem is classified into two phases they are listed below.

## **Phase 1**

Using Tweepy, gain a critical mass of tweets and then annotate the sentiment of each tweet: Positive, negative, neutral.

Clean and organise the data, remove duplicates and use techniques from text classification exercise such as stopping and stemming to prepare the dataset.

## **Phase 2**

Create a model which can accurately assign a polarity to a group of new tweets within the chosen topic. Test your model against a different domain e.g., politics model on sports or tech.

**Deliverables**

1. A dataset with annotated training data for sentiment polarity

2. A model which predicts polarity of given tweet

3. A report which includes a comparison with VADER(an off the shelf sentiment analysis tool)

# **2. Introduction**

Twitter Sentiment Analysis which tells about the mood of the user towards the (Hashtag ) or #@ user, and Twitter is one of the Largest Social Media Platform with Millions of users in them and Millions of tweets each day and more. In this Assessment we going to Choose the Political Model - 4 different Twitter Handlers and the people sentiment towards them “Obama”, “Hilary Clinton”, “Donald Trump”, “Bernie Sanders”), they are famous politicians in the United States of America . Let us Find the Collective Sentiment toward the Tweets by users towards the Handlers.

**2.1 . Process (What I, did)**

* Fetched Data using Tweepy on Multiple days from March till April 2021.
* Hand Annotated sentiments Complete 750+ Tweets after fetching of data,
* Removal of Un wanted Columns and The Re-Tweets in Tweets
* Removal of Patterns in the tweet
* Using VADER to find the polarity of the Tweets
* Virtualization of the Tweets using the Vader vs Hand Annotated to find comparison.
* Creating the Final CSV for the Tweets with the sentiment Polarity as part of Deliverables.
* NLP -Stop words, Stemmer, Lemmatizes, TFIDF
* Preparing the Model (Naïve Bayes, Multilayer Perceptron, Perceptron)
* Overall Accuracy for the Hand Annotated Tweets
* Comparing Vader with each Models for Accuracy
* Overall Accuracy for Vader tweets with Model result
* Testing of Airlines Dataset
* Airline Dataset Cleaning and Performing NLP functions(Stop words, Stemmer, Lemmatizes, TFIDF)
* Testing tweets with each Models using given sentiment in dataset( Naive Bayes/SVM/MLP/Perceptron)
* Creating Vader Sentiment for Airline Dataset Tweets
* Testing the Models for the Airline Vader Sentiment (Naive Bayes/SVM/MLP/Perceptron)
* Overall Tested Results.

**3. Hand Annotating :**

This dataset has the total of 5000 + Rows in it. I used the Programming to filter the retweets and marked them True or False based on the Retweets and I Hand Annotated complete 750+ Tweets by entering SENTIMENTS (Positive/Negative/Neutral) by reading each tweets. Which is though time consuming but still will tells us the difference between the efficiency of Vader and our hand annotated sentiment.

Graphical user interface, application, table

Description automatically generated

Next , using the programming, dropped the Nan Values which automatically drops the True section, which are retweets. And I started with the removal of patterns (like Text: @, # , Symbols , Http/Https, Most Used Emoji. So that this analysis will be perfect one.

Graphical user interface, text, application, email

Description automatically generated

**3.1. Vader Annotation**

VADER ( Valence Aware Dictionary for Sentiment Reasoning) in the NLTK package and can be applied directly to unlabelled text data. A model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion. It is available.

VADER sentimental analysis relies on a dictionary that maps lexical features to emotion intensities known as sentiment scores. The sentiment score of a text can be obtained by summing up the intensity of each word in the text.

dictionary of scores in each of four categories as Negative, Neutral, Positive, and compound Normalized scores of Positive, Negative, Neutral)

Graphical user interface, text, application, email

Description automatically generated

**Vader Annotated Tweets Count**

Chart, bar chart

Description automatically generatedText

Description automatically generated

From the Vader result of tweets, we can understand the counts of the Positive, Negative, Neutral. On further the compound is considered for the(Positive , Negative, neutral classification)

**NLP**

After the Vader now the input Tweets are once again fed into the further purification process functions like Stop-words, remove Punctuation, Lemmatizes, Tokenizer, Tfidf .

**4. Model Building to predict Polarity of Tweets**

For modelling, I used mainly like (SVM, Naïve bayes) on further testing I used Multilayer Perceptron Classifier and Perceptron. For the Sentiment analysis

**4.1 Naïve Bayes**

Why I choose Naïve Bayes because, it does not Require as much as training data. It handles both Continuous and discrete data.it is Highly Scalable with the number of predictors and data points. It is Fast and cab be used to make real-time predictions.

Naïve Byes gives the Best Accuracy for the dataset when compared with the VADER and the predicted polarity using the Multinomial NB

Table

Description automatically generated

Figure 3, Naive Bayes report

Comparision of the PREDICTED Polarity in the Naïve Bayes Vs the Hand Annotated Sentiment Vs the Vader Sentiment.

Table

Description automatically generated

From this we can understand that the Naïve bayes performance are similar and accurate prediction and it tops in the prediction and the accuracy rate for the text classification for this dataset, Tops in Accuracy with 58%.

**4.2 Support Vector Machine**

SVM is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. But Mostly used for the classification problems, the idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes.

Table

Description automatically generated

SVM is the Next Highest accuracy for this dataset with 52% accuracy.

Table

Description automatically generated

**4.3 Multilayer Perceptron Classifier.**

MLP classifier is a Neural Network , on comparing with our dataset and the predicted model seems to be low in performance and the results show less accuracy compared with SVM and Naïve Bayes.

**Below Shows the Comparison of Actual Vs Predicted MLP**

**Table

Description automatically generated**

Table

Description automatically generated

Visual Representation of the Polarity in MLP classifier with accuracy

**4.4 Perceptron**

Perceptron is a Neural Network , on comparing with our dataset and the predicted model seems to be low in performance and the results with accuracy are compared to be very less with SVM and Naïve bayes.

**Table

Description automatically generated**

**Below Shows the Comparison of Actual Hand Annotated Vs Predicted Perceptron Vs Vader Annotated**

Table

Description automatically generated

Visual Representation of the Polarity in Perceptron.

**5. Results**

From the Overall Results we can tell that the Naïve Bayes is the best Model on comparing with the Vader Analysis Toll which has the 100% Match Accuracy than other Models.Table

Description automatically generated

**5.1 Comparison with Vader:**

Here the same models as (SVM, Naïve bayes, MLP, Perceptron) is made using of the Vader sentiment and the Tweets for a verification purpose to find how its better from the Hand annotated one. And how will be the accuracy, when its automatically done instead of the Hand Annotated is studied,

This study tells that they are in a very slight variation (Accuracy performance with Vader Vs Hand Annotated) and only SVM-Rbf kernel outperforms and give us the 54.9% on compared with the SVM of Hand Annotated rest other models are close to each other on comparison.

Chart, bar chart

Description automatically generated

|  |  |  |
| --- | --- | --- |
| **Model** | **Hand Annotation\_ Sentiment** | **VADER Sentiment** |
| Naïve Bayes | 57% accuracy | 56% accuracy |
| SVM-rbf Kernel | 51% accuracy | 54% accuracy |
| MLP-Classifier | 35% Accuracy | 36%Accuracy |
| Single Layer Perceptron | 35% accuracy | 35% accuracy |

On Manually looking the comparison with the Hand Annotated and the Vader we can tell that Vader is much more effective

**5.2 Testing Model with Different Dataset**

**5.2.1 Testing with Given Annotation of the Airline Dataset.**

Dataset Description: US airlines Dataset which contains the tweets of most of the airlines in United States like(Virgin, Atlantic, Jet blue, United Airlines) these are the tweet posted and its available in the Kaggle.

For testing, Airline dataset has been used and which run over all similar parameters functions like filtering and removing patterns like( retweets, Http/Https) and then into the NLP Function which has (Stop-words, remove Punctuation, Lemmatizes, Tokenizer, Tfidf) and further tweets fed into models and we can see that that the perceptron predicts the sentiment accurately followed by the Naïve Baye, SVM, MLP-classifier, Perceptron.

Chart, bar chart

Description automatically generated

**5.2.1 Testing with Vader sentiment of Airline Dataset (Additionally for self-testing purpose)**

On a further study how the testing models works, I have done a complete Vader analysis for the testing Airline Dataset, why this was carried out because to understand more about the Vader and to have a sure shot to compare our own dataset (US Political Persons). So that this study is used and essential to complete.

Model has been built (Using Vader Sentiment) and performed the Naïve Baye, SVM, MLP-classifier, Perceptron Algorithm.

After doing this its very clearly understood that the Accuracy was very less compared with the Testing model of the Airline Dataset(With Given Annotation Sentiment). Moreover, which is in the similar case of the Us Personal Dataset, why because : there also Vader Accuracy was less from hand annotated sentiment accuracy. From this we can tell that the Vader sentiment is not that much logical and perfect.

Chart, bar chart

Description automatically generated

|  |  |  |
| --- | --- | --- |
| **Testing- Model Airline** | **Given Sentiment** | **Vader Sentiment** |
| Naïve Bayes | 73% accuracy | 64.5% |
| SVM-rbf Kernel | 72% accuracy | 56.8% |
| MLP-Classifier | 69% Accuracy | 65.2% |
| Single Layer Perceptron | 65% accuracy | 66.6% |

**6. Conclusion:**

I tested the Both Vader Model with the hand Annotated Model where we can tell that Hand Annotated is good why because its more logical and sensible and it has good accuracy over Vader annotated.

Vader completely uses the lexical and performs the polarity. Vader also has drawbacks as it does not understand the full logic of the sentences say example (if there is positive in the sentence ) it considers positive. In my case I checked what, Vader does- for example, a tweet( TRUE,TRUE,TRUE was mentioned. In hand annotation I, gave its as Neutral but in Vader it took it as a Positive sentiment.) when running the Model, it predicted as the Neutral. This reconfirms our predicted Hand Annotated Model is more Accurate and correct.

Next, on testing with another dataset we can reconfirm the claim that Vader is less effective than hand Annotated. Why because I performed both Vader sentiment and given sentiment in the dataset. Results showed the Vader is poor in Accuracy and the Given Annotations for the Airline Dataset is better and good.