

1.INTRODUCTION

1.1 Overview

As internet is growing bigger, its horizons are becoming wider. Social Media and Micro blogging platforms like Facebook, Twitter, Tumblr dominate in spreading encapsulated news and trending topics across the globe at a rapid pace. A topic becomes trending if more and more users are contributing their opinion and judgements, thereby making it a valuable source of online perception [3]. These topics generally intended to spread awareness or to promote public figures, political campaigns during elections, product endorsements and entertainment like movies, award shows. Large organizations and firms take advantage of people's feedback to improve their products and services which further help in enhancing marketing strategies. One such example can be leaking the pictures of upcoming iPhone to create a hype to extract people's emotions and market the product before its release. Thus, there is a huge potential of discovering and analysing interesting patterns from the infinite social media data for business-driven applications. Sentiment analysis is the prediction of emotions in a word, sentence or corpus of documents. It is intended to serve as an application to understand the attitudes, opinions and emotions expressed within an online mention. The intention is to gain an overview of the wider public opinion behind certain topics. Precisely, it is a paradigm of categorizing conversations into positive, negative or neutral labels.

1.2 Purpose

The purpose of this project is to build an algorithm that can accurately classify Twitter messages as positive or negative, with respect to a query term. Our hypothesis is that we can obtain high accuracy on classifying sentiment in Twitter messages using machine learning techniques. Generally, this type of sentiment analysis is useful for consumers who are trying to research a product or service, or marketers researching public opinion of their company

2. LITERATURE SURVEY

2.1 Existing Problem

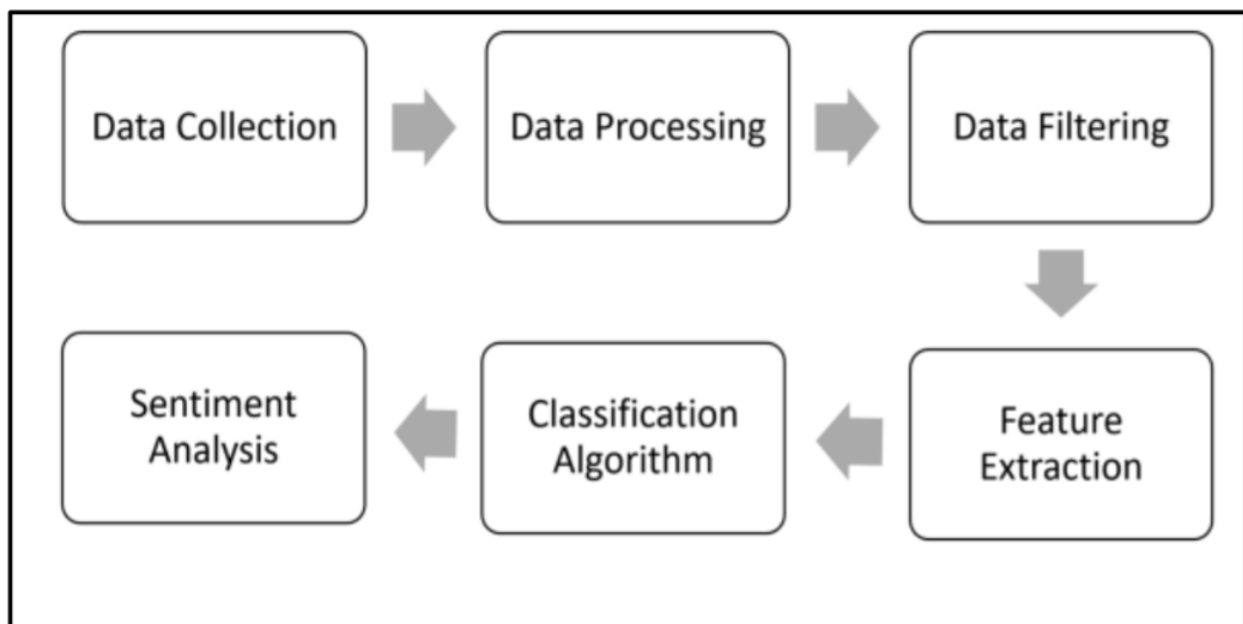
Every day massive amount of data is generated by social media users which can be used to analyze their opinion about any event, movie, product or politics. It is still difficult for a vast majority of tools to precisely evaluate what truly is a negative, neutral, and a positive statement. While it can be extremely helpful, there's still a long way to go before sentiment analysis becomes more accurate. Even though it's an automated process, marketers still need to evaluate themselves whether a negative mention is truly negative, for instance.

2.2 Proposed Solution

Taking necessarily, the prerequisite inputs from which sentiment analysis can be predicted beforehand helps the user to learn a trend of tweets which people tweet on this vast platform of twitter. This sentiment analysis is done with the concept of natural language processing that has highest accuracy infers the finest algorithm.

3. THEORITICAL ANALYSIS

3.1 Block Diagram



3.2 Hardware/Software designing

- DATA COLLECTION

Collection of the data set related to the sentiment analysis of twitter

- TEXT PRE-PROCESSING

—

1. Importing the dataset
2. Removing the punctuations, numbers and unwanted stopwords using the appropriate libraries.
3. Coverting the each sentence of the given data set into lower case
4. Stemming
5. Splitting the dataset into two groups- train and test sets. Train sets will be used to train the model and after training the model the test sets will be used to test the model using random values.

- THE MODEL BUILDING

6. Importing the model building libraries
7. Initializing the Model
8. Adding the input layer
9. Adding the Hidden layer
10. Adding the output layer
11. Configure the learning process
12. Train and test the model
13. Optimize the model
14. Save the model(.h5 file)
15. Predictions- using random values as inputs predict the output using the saved model.

- APPLICATION BUILDING

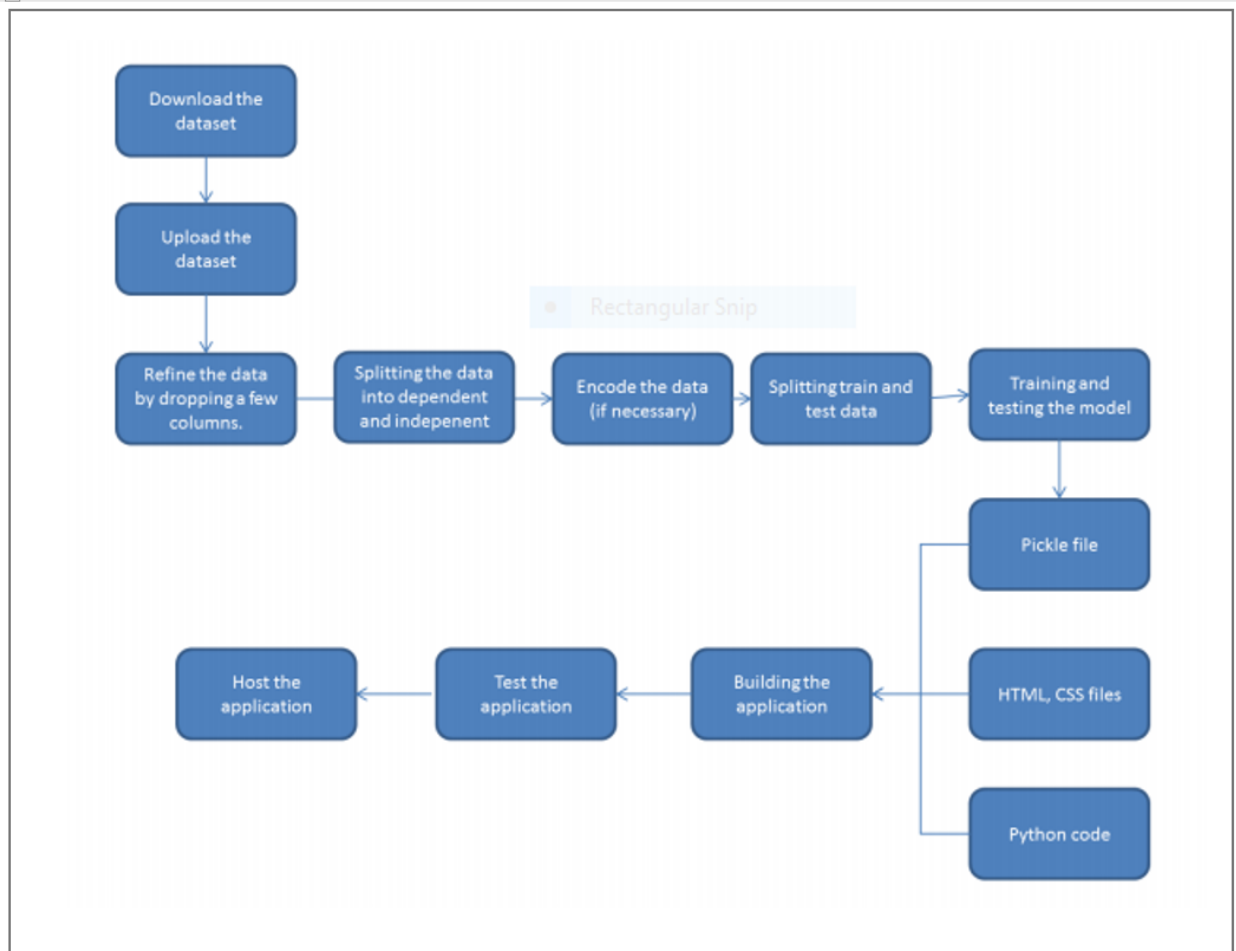
16. Creating an HTML file
17. Creating a python file

4. EXPERIMENTAL INVESTIGATION

In our prediction we collect information i.e. inputs required for the Twitter Sentiment Analysis. In our task we collect parameter which is the collection of tweets which people have tweeted expressing their views, thoughts, ideas, expressions or simply how they

are doing today.

5. FLOWCHART



6. RESULT

The natural processing language model can predict the output (0 for positive and 1 for negative) based on certain groups of words present in the dataset.

7. ADVANTAGES AND DISADVANTAGES

Advantages

- Accurate predictions
- Easy usable interface
- Results shown are without any errors
- It gives information about the sentiment of the people whether is it negative or positive.

Disadvantages

- For predictions we might need more data to train the set. Sometimes if a particular data has both negative and positive words it will be difficult to predict.

8. APPLICATIONS

Useful in predicting whether a particular word or sentence indicates positivity or negativity. It can sort the given data set with the help of prediction. It doesn't require any human interference in predicting.

9. CONCLUSIONS

Hence this model will help us to find out the correct analysis of a given tweet. Based on the criteria we can predict the sentiment (whether positive or negative) precisely.

10. FUTURE SCOPES

This model can form a base to predict a larger group of data. Not only positive or negative it can predict the exact emotion or the sentiment based on certain parameters of data. Hence, it can be a powerful tool in analysing data and sorting it based on the

positive or negative effects.

11. BIBLIOGRAPHY

Wikipedia, ResearchGate, TensorFlow