Department of Information Technology
NBA Accredited

A.P. Shah Institute of Technology

G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2020-2021

A Project Report on

SENTIMENT ANALYSIS

Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-8)

in
INFORMATION TECHNOLOGY
By

Pratham Dhanesha (19104021)
Pavan Chopra (19104069)
Shreya Desai (16104018)

Under the Guidance of
Name of Guide
Prof. Manasi Choche

1.Project Conception and Initiation

1.1 Objectives

- To identify as well as classify the sentiments that are expressed in the text source.
- To generate a vast amount of sentiment data upon analysis
- To understand the opinion of the people about a variety of topics.
- To use natural language processing, text analysis, and statistics to analyse customer sentiment.
- To detect positive or negative sentiment in text.

1.2 Literature Review

In this survey, knowledge base approach is used to come to classify the opinions and sentiments in a particular text. In the survey, Naive Bayes along with the Semantic Orientation based WordNet which extracts synonyms and similarity for the content feature.
The key idea of the paper is to increase the accuracy of classification by including Natural Language Processing Techniques (NLP).

1.3 Problem Definition

- Twitter sentiment analysis is a model that helps to overcome the challenges of identifying the sentiments of the tweets.
- The problem in sentiment analysis is classifying the polarity of a given text whether the expressed opinion in a document, a sentence or an entity feature/aspect is positive, negative, or neutral .

1.4 Scope

- Twitter sentiment analysis allows you to keep track of what's being said about your product or service on social media.
- Anyone can use sentiment analysis to compile and analyse large amounts of text data, such as news, social media, opinions, and suggestions.
- Sentiment analysis is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs.

1.5 Technology stack

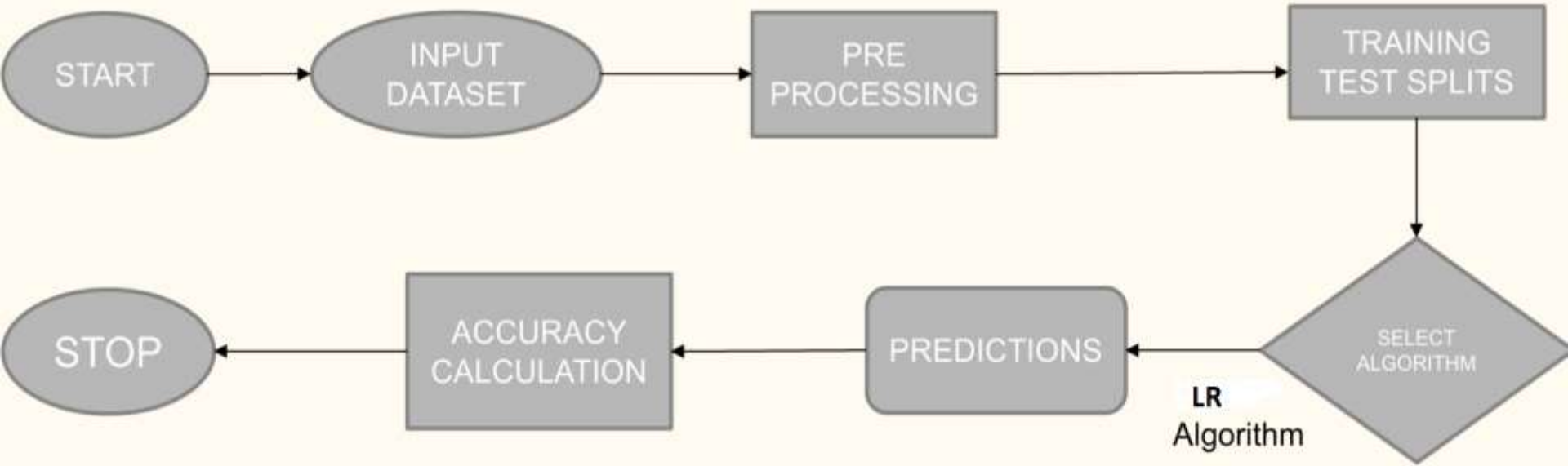
- Jupyter Notebook
- Anaconda Navigator
- Python 3.9.12
- HTML
- CSS
- Streamlit

2. Project Design

2.1 Proposed System

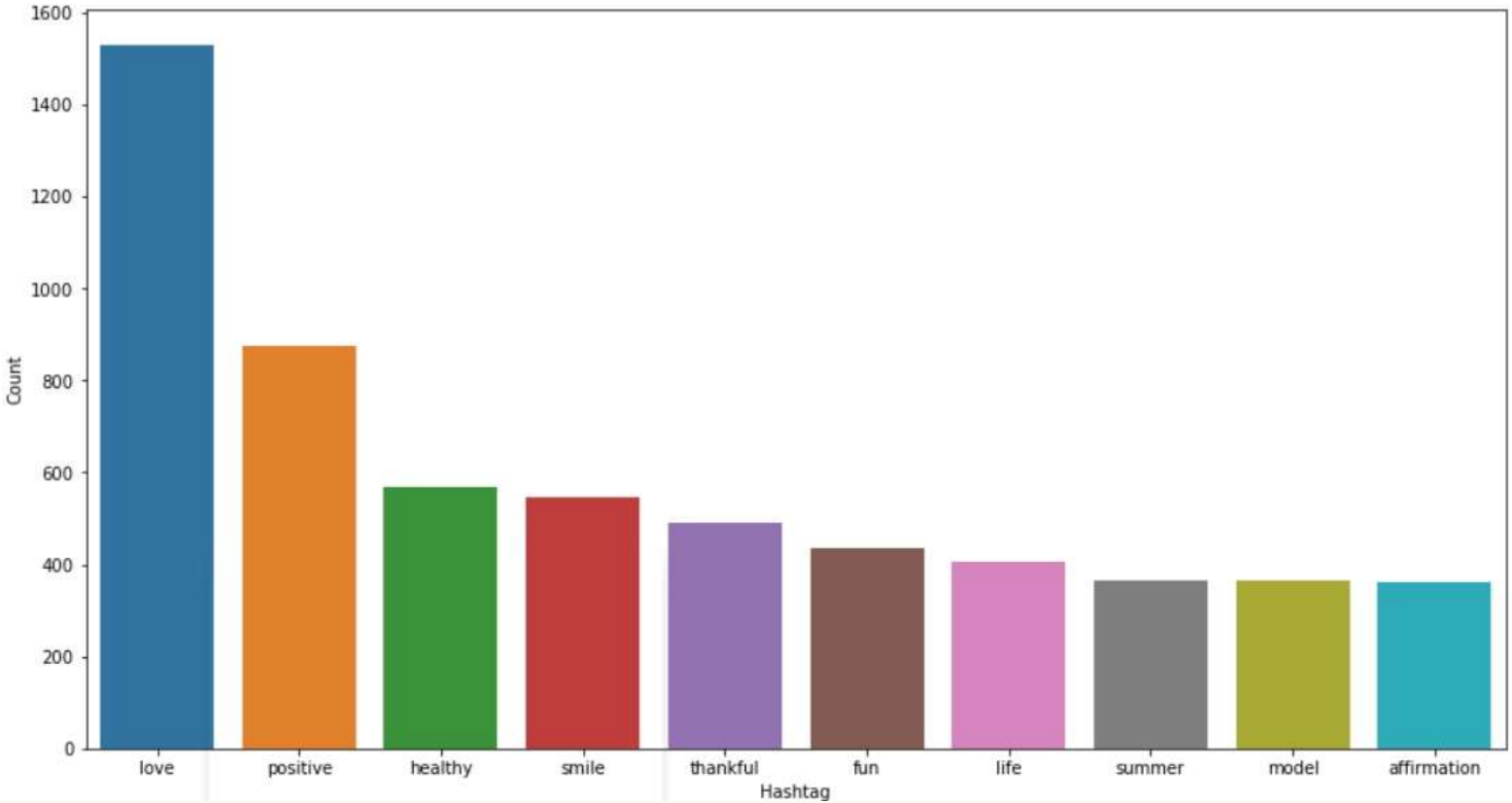
- Sentiments are classified with the help of machine learning algorithms. We have used the datasets from Kaggle which was crawled from the internet and labeled positive/negative.
- The data provided comes with emoticons (emoji), usernames and hashtags which are required to be processed (so as to be readable) and converted into a standard form.
- We use various machine learning algorithms based on Logistic Regression to conduct sentiment analysis using the extracted features.

2.2 Design(Flow Of Modules)



3. Implementation

In [22]:
```
pos = pos.nlargest(columns='Count', n=10)
plt.figure(figsize=(15,8))
sns.barplot(data=pos, x='Hashtag', y='Count')
plt.show()
```



5. Result

In this project, we were able to display graph of sentiment text analyzing the dataset provided using Logistic Regression Algorithm.

6. Conclusion and Future Scope

- The future of sentiment analysis is going to continue to dig deeper, far past the surface of the number of likes, comments and shares, and aim to reach, and truly understand, the significance of social media interactions and what they tell us about the consumers behind the screens.
- It enables you to analyze large amounts of market research data in order to spot emerging trends and better understand consumer buying habits.

References

- M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013, pp. 1-5, doi: 10.1109/ICCCNT.2013.6726818.
- G. Gautam and D. Yadav, "Sentiment analysis of twitter data using machine learning approaches and semantic analysis," 2014 Seventh International Conference on Contemporary Computing (IC3), 2014, pp. 437-442, doi: 10.1109/IC3.2014.6897213.
- M. Kanakaraj and R. M. R. Guddeti, "Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques," Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015), 2015, pp. 169-170, doi: 10.1109/ICOSC.2015.7050801

Thank You