IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 8, Issue 12, December 2020

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 8, Issue 12, December 2020 ||

| DOI: 10.15680/IJIRCCE.2020.0812034 |

Twitter Tweet Analyzer Model Using Natural Language Processing

Charvin Zaveri[1], Divjyot Singh Saluja[1]

U.G. Student, Department of Electronics and Telecommunication Engineering, Dwarkadas .J. Sanghvi College of Engineering, Vile Parle, Mumbai, Maharashtra, India [1]

ABSTRACT: Social media has garnered a lot of attention these days. Public and private perspectives on a variety of topics are voiced and are being continuously disseminated through a wide range of social media. Twitter is one of the most popular social media platforms. Twitter provides organizations with a quick and effective way to analyze customer perceptions of market success. Developing an emotional analysis system is a method that will be used to measure customer comprehension. This paper reports on the construction of emotional analysis, excluding a large number of tweets. Prototyping is used in this development. The results separate customer perceptions with tweets of good and bad, represented on a pie chart. However, this program is planning to upgrade to a web application system, but due to the Django limit that can be used on a Linux or LAMP server, to continue this process should be done.

KEYWORDS: Twitter, customer perceptions, emotional analysis.

I. INTRODUCTION

Over the past few years, there has been a huge increase in the use of microblogging platforms such as Twitter. Encouraged by that growth, companies and media organizations are increasingly looking for ways to present Twitter information about how people think and feel about their products and services. Companies like Twitratr, tweetfeel , and Social Men-tion are just a few of the people who advertise Twitter emotional analysis as one of their services. be a fair amount of research on how emotions are expressed in genres such as online reviews and news articles, how emotions are expressed when given the language barriers and the limitations of microblogging messages have not been read much. Features such as automatic part-of-speech tags and resources such as sentiment lexicons have proven useful in emotional analysis in other domains, but will they also be helpful in sentimental testing on Twitter? In this paper, we begin to investigate the question: Another challenge to making microblogging is the amazing scope of the cover story. It's no exaggeration that people write a tweet about anything and anything. Therefore, in order to be able to build Twitter empathy systems for any given topic, we need a way to quickly identify data that can be used for training. In this paper, we use one method to build such data: using Twitter hash tags (e.g., #Bestfeeling, #epicfail, #news) to identify positive, negative, and neutral post tweets used to train methods three emotional separators.

II. RELATED WORK

Sentiment analysis is a growing area of Natural Language Processing. Research through research from the level of document classificationto the study of sentence-related words. Given the limitations of the characters in the tweets, the distinction of the sentiments of Twitter messages is very similar to the analysis of sentences ratios ; However, the informal and special language used in tweets, and the microbloggingdomain environment make Twitter emotional analysis a very different task. The open question is how the features and technique used in well-designed data will transfer to the themicroblogging domain.Last year there were a number of papers that looked at Twitter feeling and buzz .Some researchers have begun to experiment with the use of speech therapy, but the repetition is still mixed. Common features in microblogging (e.g., thumbnails) are also common, but there has been little research into the use of existing sensory sources made of non-microblogging data. Researchers have also begun investigating various ways to collect training data automatically. Several researchers use icons to describe their training data use existing Twitter emotional sites to collect training data also use hashtags to create training data, but limit their experimental practices to emotional / non-emotional separation, rather than three-way separation.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 8, Issue 12, December 2020 ||

| DOI: 10.15680/IJIRCCE.2020.0812034 |

III. PROPOSED ALGORITHM

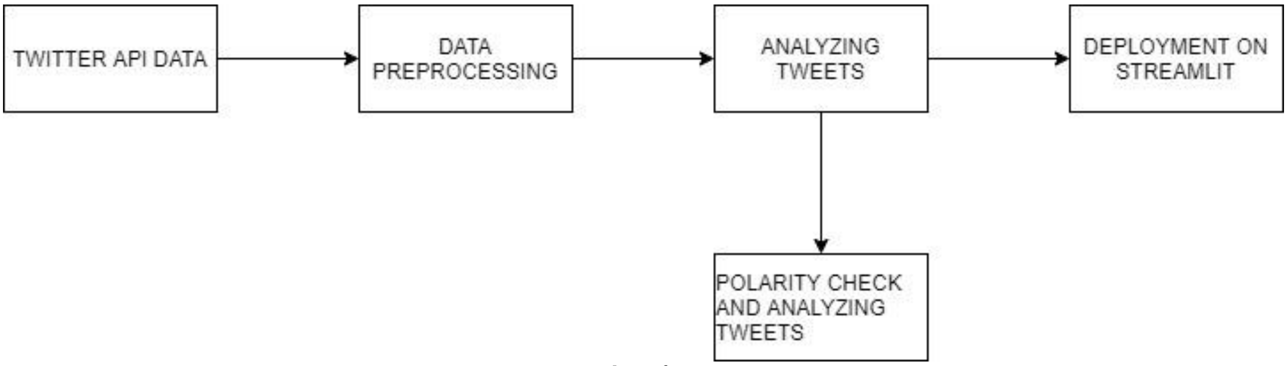

Fig. 1. Flow of Program

A) TWIITER API DATA:

Twitter is what is happening in the world and what people are talking about right now. You can access Twitter via the web or your mobile device. Sharing information on Twitter is as broad as possible, and it provides companies, developers, and users with scheduled access to Twitter data through our APIs (application programming interface). This explains what the Twitter APIs are, what information is made available about them, and other protections Twitter has in place for their use. At the highest level, APIs are the way computers “talk” so that they can request and transmit information. This is done by allowing the software application to call what is known as endpoint: the address corresponding to the type of information we provide (endpoints are usually different as phone numbers). Twitter allows access to parts of our app via APIs to allow people to build software that integrates with Twitter, as a solution that helps a company respond to customer feedback on Twitter. Twitter data is different from data shared by many other social platforms because it shows details of users who choose to share publicly. Our API platform provides extensive access to the Twitter profile of users who have chosen to share the world. We also support APIs that allow users to manage their non-public Twitter information (e.g., Direct Messages) and provide this information to the developers they are authorized to do so.

B) DATA PREPROCESSING:

Data processing has three steps: 1) making tokens, 2) familiarity, and 3) partial marking (POS). Emoticons and abbreviations are identified as part of the token process and treated as individual tokens. In the normal process, the presence of abbreviations within the tweet is noted and the summaries are replaced by their true meaning (e.g., BRB−> go back directly). We also point to informal reinforcement similar to all chemicals (e.g., I LOVE this show !!! and duplicate of letters (e.g. I have a deposit !! happyyyyyy”), note their presence in the tweet. All words are made up of small pages, and duplicate characters are taken place one character. Finally, the presence of any special Twitter tokens is noted (e.g., #hashtags, user tags, and URLs) as well as pronouns indicating the type of token is included.that this modification improves the performance of the POS tagger, which is a final step of preparation.

C) ANALYZING TWEETS:

Sentiment Analysis from a text is a classical problem of NLP. It is used when you try to predict the sentiment of customer feedback in a restaurant, a shopping site, etc. This app is a text analysis technique that detects the sentiment of the user, whether it's positive or negative. This technique allows brands or the companies to listen attentively to their customers by examining their feedback, and tailoring products and services to meet their needs. For instance, using sentiment analysis to automatically analyse 4,000+ reviews about your product can help you discover if customers are happy about your service and pricing plans or not.

D) DEPLOYMENT ON STREAMLIT:

Streamlit is an open-source Python source library it's easy to create and share beautiful, automated web app machine learning and data science. In our project we have customized the entire web app by creating options to show recent tweets, generate a word cloud and performing sentiment analysis and displays it in the form of a bar

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 8, Issue 12, December 2020 ||

| DOI: 10.15680/IJIRCCE.2020.0812034 |

graph. The sentiment can be found in the comments or tweet to provide useful indicators for many different purposes. Also, and stated that a sentiment can be categorized into two groups, which is negative and positive words. Sentiment analysis is a natural language processing technique to quantify an expressed opinion or sentiment within a selection of tweets

IV. RESULTS

The result will be shown in a bar chart which is representing positive, negative and neutral sentiment hash tags. For null hash tag is representing the hash tags that were assigned zero value. However, this program is able to list a top ten positive and negative hash tag.The pie chart is representing of each percentage positive, negative and null sentiment hash tags in different colours.Test results show that deleting URLs barely affects the performance of classifiers on two-dimensional models across all data sets. This indicates that the URLs contain useful information for emotional separation. It shows that there are few effects on the performance of classifiers in N-gram model before and after position stops. One of the reasons could be that words never appear in tweets always. In the Prior polarity model, it removes the stop words leads to a change in the functioning of the partition because a stop words that contain different feelings of polarity.In the future, we will continue our evaluation using different stoplists and acronym dictionaries and will investigate the reasons for the fluctuation of sentiment classification performance using different classifiers on various datasets.

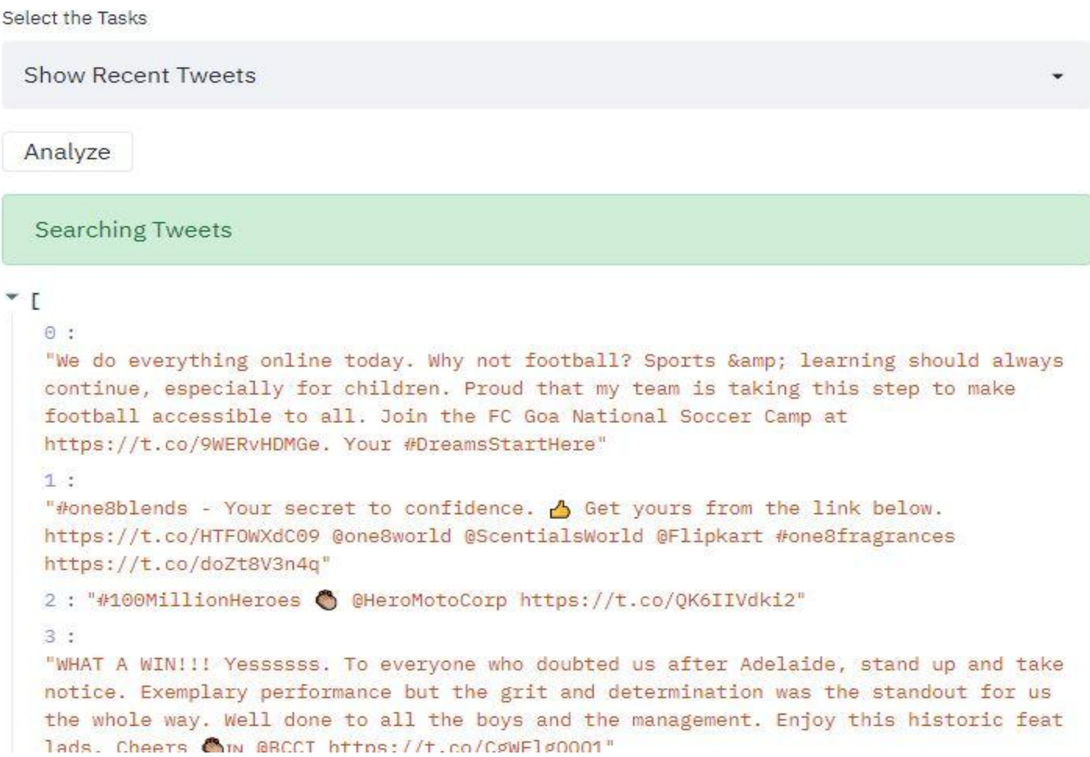

Fig. 2. Sweeping through Tweets

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 8, Issue 12, December 2020 ||

| DOI: 10.15680/IJIRCCE.2020.0812034 |

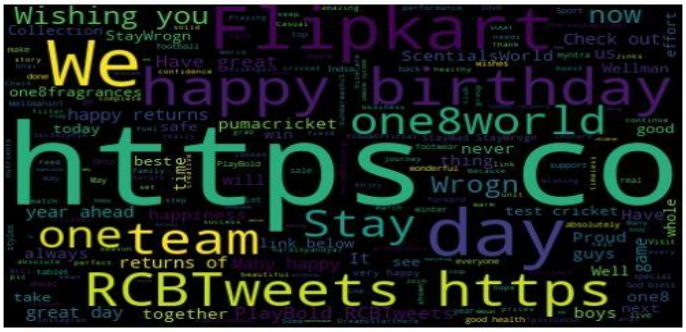

Fig. 3. Word Map

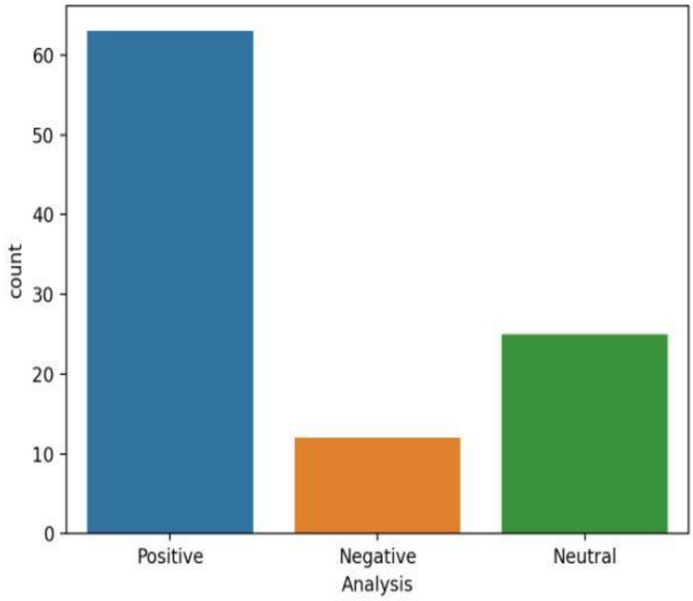

Fig. 4. Comparison of the Responses

V. CONCLUSION AND FUTURE WORK

This paper studies the fact that six different pre-processing methods affect sentiment polarity classification in the Twitter. We conduct a series of experiments using four classifiers to verify the effectiveness of several pre-processing methods on five Twitter datasets. Experimental results indicate that the removal of URLs, the removal of stop words and the removal of numbers minimally affect the performance of classifiers; furthermore, replacing negation and expanding acronyms can improve the classification accuracy. Therefore, removing stop words, numbers, and URLs is appropriate to reduce noise but does not affect performance. Replacing negation is effective for sentiment analysis. We select appropriate pre-processing methods and feature models for different classifiers for the Twitter sentiment classification task.As a result, program will be categorized sentiment into positive and negative, which is represented in a pie chart and html page Although, the program has been planned to be developed as a web application, due to limitation of Django which can only work on Linux server or LAMP. Thus, it cannot be realized. Therefore, further enhancement of this element is recommended in future study.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 8, Issue 12, December 2020 ||

| DOI: 10.15680/IJIRCCE.2020.0812034 |

REFERENCES

[1] M.Rambocas, and J. Gama, "MarketingResearch:TheRoleof SentimentAnalysis". The 5th SNA-KDD Workshop'11. Universityof Porto, 2013.
[2] A. K. Jose, N. Bhatia, and S. Krishna, "TwitterSentimentAnalysis". NationalInstituteof TechnologyCalicut,2010.
[3] P. Lai, "ExtractingStrongSentimentTrendfromTwitter". Stanford University, 2012.
[4] Y. Zhou, and Y. Fan, " A Sociolinguistic Study of American Slang," Theory and Practice in Language Studies, 3(12), 2209–2213, 2013. doi:10.4304/tpls.3.12.2209-2213
[5] M. Comesaña, A. P.Soares, M.Perea, A.P. Piñeiro, I. Fraga, and A. Pinheiro, " Author ' s personal copy Computers in Human Behavior ERP correlates of masked affective priming with emoticons," Computers in Human Behavior, 29, 588–595, 2013.
[6] A.H.Huang, D.C. Yen, & X. Zhang, "Exploring the effects of emoticons," Information & Management, 45(7), 466–473, 2008.
[7] D. Boyd, S. Golder, & G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," System Sciences (HICSS), 2010 …. Retrieved.
[8] A. Pak,and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," Special Issue of International Journal of Computer Application, France:Universitede Paris-Sud, 2010.
[9] S.Lohmann, M. Burch, H. Schmauder and D. Weiskopf, "Visual Analysis of Microblog Content Using Time-Varying Co-occurrence Highlighting in Tag Clouds," Annual conference of VISVISUS. Germany: University of Stuttgart, 2012.
[10] H. Saif, Y.He, and H. Alani, "SemanticSentimentAnalysisof Twitter," Proceeding of the Workshop on Information Extraction and Entity Analytics on Social Media Data. United Kingdom: Knowledge Media Institute, 2011.
[11] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R.Passonneau, "Sentiment Analysis of Twitter Data," Annual International Conferences. New York:Columbia University, 2012.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details