iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43661

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Twitter Data Visualization using SVM (Machine Learning Approach)

Sanket Gaikwad[1], Vishal Ingle[2], Razique Raza[3], Prof. Pooja Waware[4]
[1, 2, 3]*Department of Computer Engineering, ISBM College Of Engineering, SPPU, India.*
[4]*Guide*

Abstract: ***Twitter produces a massive amount of data due to its popularity that is one of the reasons underlying big data problems. One of those problems is the classification of tweets due to use of sophisticated and complex language, which makes the current tools in- sufficient. We present our framework HTwitt, built on top of the Hadoop ecosystem, which consists of a MapReduce algorithm and a set of machine learning techniques embedded within a big data analytics platform to efficiently address the following problems.***

Keywords: Support Vector Machine (SVM), Machine learning, Big Data, MapReduce.

I. INTRODUCTION

Nowadays, people from all around the world use social media sites to share information. Twitter for example is a platform in which users send, read posts known as 'tweets' and interact with different communities. Users share their daily lives, post their opinions on everything such as brands and places. Companies can benefit from this massive platform by collecting data related to opinions on them. The aim of this system is to present a model that can perform sentiment analysis of real data collected from Twitter. Data in Twitter is highly unstructured which makes it difficult to analyse. Online social media such as Twitter, Facebook, and Instagram allow users to communicate with the whole world. Write their own opinions about products or share their moments, even influence politics and companies. Twitter for example, al- most every huge company has an account on Twitter to know about their customers' feedback about their services or products. Sentiment analysis, known as opinion mining, for classifying specific words.

II. LITERATURE SURVEY

Sr. no.	Publication Details	Author	Abstract	Research Gap Identified
1	Machine Learning based Sarcasm Detection on Twitter Data Year-2020	Neha Pawar; Sukhada Bhingarkar	Sarcasm is a subtle type of irony, which can be widely used in social networks. It is usually used to transmit hidden information to criticize and ridicule a person and to recognize. Sentiment analysis refers to internet users of a particular community, expressed attitudes and opinions of identification and aggregation. In this paper, to detect sarcasm, a pattern-based approach is proposed using Twitter data.	Accuracy Very less
2	Opinion Mining Using Live Twitter Data Year-2019	Andleeb Aslam 1, Usman Qamar 2, Reda Ayesha Khan 3, Pakizah Saqib 4, Aleena Ahmad 5, Aiman Qadeer	Opinion mining and extracting Sentiments of people is the need of today. This is the era of big data. Because of social networking sites, it's easy to analyze sentiments of people. Sentiment analysis is a technique to extract opinions of people regarding any product, issue or personality. This paper is about extracting live twitter data regarding any topic and converting it into structured form from unstructured one. Opinions are extracted from the text data and polarity is assigned against each tweet.	Network issue

3	Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model Year- 2021	CHRISTOPHER IFEANYI EKE 1,2, AZAH ANIR NORMAN 1, AND LIYANA SHUIB 1	Nowadays mobile devices such as smartphones had widely been used. People use smartphones not limited for phone calling or sending messages but also for web browsing, social networking and online banking transaction. To certain extend, all confidential information is kept in their smartphone. As a result, smartphones became as one of the cyber-criminal main targets especially through an installation of mobile botnet. Eurograbber is an example of mobile botnet that being installed via infected mobile application without victim knowledge. It will pretence as mobile banking application software and steal financial transaction information from victim's smartphone. In 2012, Eurograbber had caused a total loss of USD 47 Million accumulatively all over the world. Based on the implications posed by this botnet, this is the urge where this research comes in.	Less Accuracy
4	Sentiment Analysis of Tweets from Airlines in the Gulf Region Using Machine Learning Year-2021	Mazen M. Hrazi , Abdulrahman M. Althagafi , Abdullah T. Aljuhani , Jenifar Rahman, Md. Mahfuzur Rahman, Mohammad Shorfuzzaman	Over the past few years' people have been using Twitter messages on a daily basis to express their views and share their feelings. As a result, the size of data is increasing dramatically creating opportunity for researchers to use these tweets as sources for data mining and extract valuable information. Being popular in Saudi Arabia, we believe that twitter messages (tweets) are a good source to capture the sentiment of people.	Not Efficient
5	Discovering Public Opinions by Performing Sentimental Analysis on Real Time Twitter Data Year-2018	G. Kavitha; B. Saveen; Nomaan Imtiaz	Data streaming is an evolutionary concept in big data where the size of data increases a lot from social media, trending websites, and mobile applications. Nowadays, Streaming data tends to collect data from live streaming to run analysis and generate reports for data prediction. This process requires skilled professional for acquiring data from live stream using complex coding and queries. The above drawback is overcome in this research work by implementing streaming algorithm to fetch data from twitter using a keyword search. The Twitter data visualization application is designed for data visualization, report generation and its analysis. The live twitter data is fetched by configuring the system with Hadoop, Hive warehouse and Apache Flume.	Less Efficient

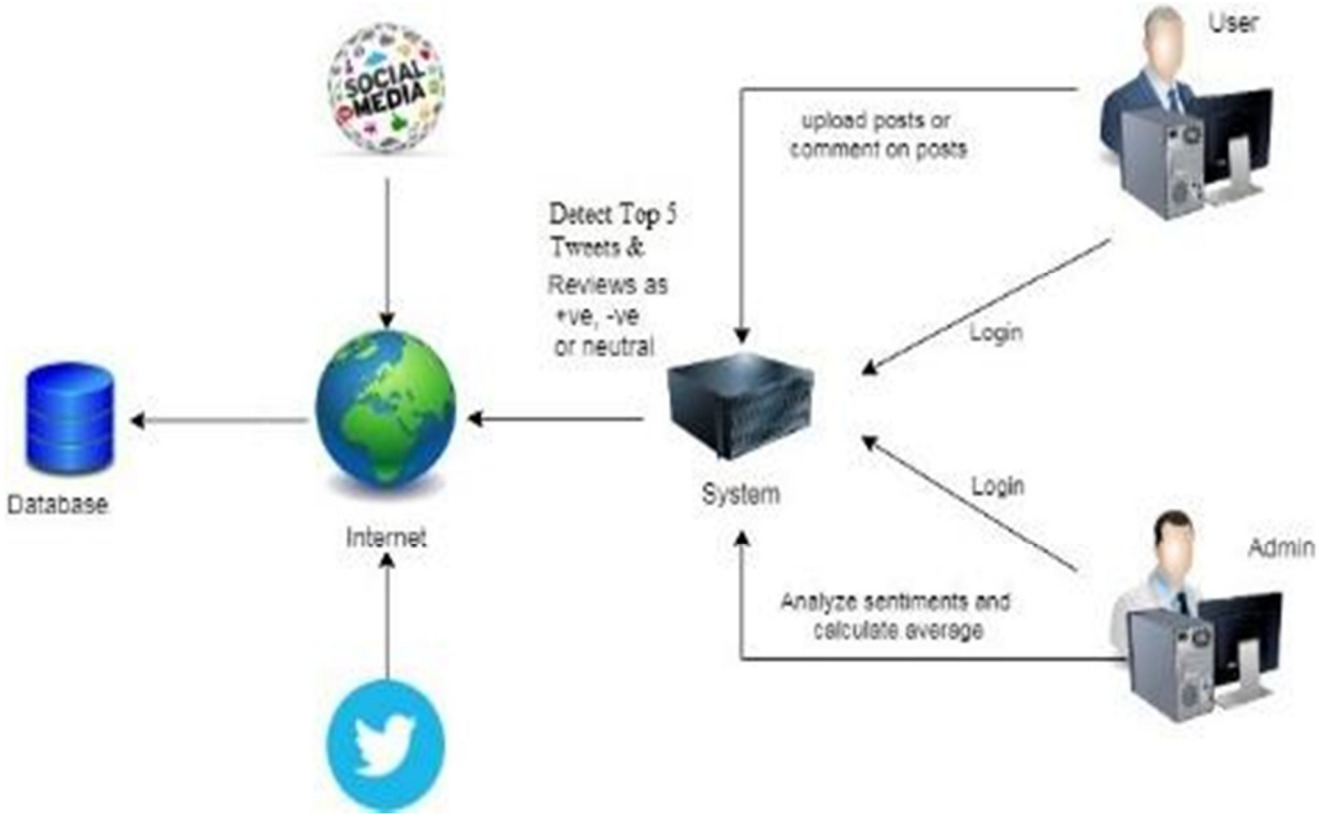
III. RELATED WORK

Analysis is a technique widely used in text mining. Twitter Sentiment Analysis, therefore means, using advanced text mining techniques to analyse the sentiment of the text (here, tweet) in the form of positive, negative and neutral.

The framework will accept both structure and unstructured data. And so, with machine learning, this method can improve the accuracy of prediction. The limitations faced are the lack of information and research about the symptoms of disease.

Exporting the dataset into the desirable format is one more setback we have to face. The proper user input query must be checked before dealing with processing.

IV. ARCHITECTURAL DIAGRAM



V. CONCLUSION

"TWITTER DATA VISUALIZATION " is presented that is implemented using simple programming concepts of Python and JavaScript. Twitter Analyzer is capable of finding out the top ten trending hashtags and users at any given point in time and plotting them against their frequency using a bar graph. The model explained here can be extended to improve user experience, provide additional functionalities and optimize processing power. Machine Learning techniques are simpler and efficient than Symbolic techniques. These techniques can be applied for twitter sentiment analysis. Classification accuracy of the feature vector is tested using different classifiers like Naive Bayes, SVM, Maximum Entropy and Ensemble classifiers. All these classifiers have almost similar accuracy for the new feature vector.

REFERENCES

[1] Anukarsh G Prasad, Sanjana S, Skanda M Bhat, B S Harish "Sentiment Analysis for Sarcasm Detection on Streaming Short Text Data", 2nd International Conference on Knowledge Engineering and Applications, IEEE, 2018.
[2] Savan K. Patel,Jigna B. Prajapati "A Study on Developing Effective Option Trading Strategy On Nifty Index in National Stock Exchange using Data Mining " International Research Journal of Engineering and Technology (IRJET), volume 04, issue 10, pages 201-204.
[3] ParasDharwal, Tanupriya Choudhury, Rajat Mittal, Praveen Kumar, "Automatic Sarcasm Detection using Feature Selection", International Conference on Applied and Theoretical Computing and Communication Technology, IEEE.
[4] Sindhu. C, G. Vaidhu, Mandala Vishal Rao, "A Comprehensive Study on Sarcasm Detection Techniques in Sentiment Analysis", International Journal of Pure and Applied Mathematics, volume 118, pages 433-442, 2018.
[5] Tanya Jain, NileshAgrawal, GarimaGoyal, Niyati Aggrawal, "Sarcasm Detection of Tweets: A Comparative Study", Tenth International Conference on Contemporary Computing (IC3), IEEE, August 2017.
[6] Levy, M. (2016). Playing with Twitter Data. [Blog] R-bloggers. Available at: https://www.r-bloggers.com/playing-with-twitter-data/ [Accessed 7 Feb. 2018].
[7] Popularity Analysis for Saudi Telecom Companies Based on Twitter Data. (2013). Research Journal of Applied Sciences, Engineering and Technology. [online] Available at: http://maxwellsci.com/print/rjaset/v6-4676-4680.pdf [Accessed 1 Feb. 2018].
[8] Zhao, Y. (2016). Twitter Data Analysis with R – Text Mining and Social Net- work Analysis. [online] University of Canberra, p.40. Available at: https://paulvanderlaken.files.wordpress.com/2017/08/rdataminingslides-twitter- analysis.pdf [Accessed 7 Feb. 2018].
[9] Alrubaiee, H., Qiu, R., Alomar, K. and Li, D . Sentiment Analysis of Arabic Tweets in e-Learning. Journal of Computer Science. [online] Available at: http://thescipub.com/PDF/jcssp.2016.553.563.pdf [Accessed 7 Feb. 2018].

[10] Qamar, A., Alsuhibany, S. and Ahmed, S. (2018). Sentiment Classification of Twitter Data Belonging to Saudi Arabia Telecommunication Companies. (IJACSA) International Journal of Advanced Computer Science and Applications, [online] 8. Available https://thesai.org/Downloads/Volume8No1/Paper 50-Sentiment Classification of Twitter Data Belonging.pdf [Accessed 1 Feb. 2018].

[11] R. M. Duwairi and I.Qarqaz, "A framework for Arabic sentiment analysis using supervised classification" , Int. J. Data Mining, Modeling and Management, Vol. 8, No. 4, pp.369-381.

[12] Hossam S. Ibrahim, Sherif M. Abdou, Mervat Gheith, "Sentiment Analysis for Modern Standard Arabic and Colloquial", International Journal on Natural Language Computing (IJNLC), Vol. 4, No.2, pp. 95-109.

[13] Anukarsh G Prasad, Sanjana S, Skanda M Bhat, B S Harish "Sentiment Analysis for Sarcasm Detection on Streaming Short Text Data", 2nd International Conference on Knowledge Engineering and Applications, IEEE.

[14] Sana Parveen, Sachin N. Deshmukh, "Opinion Mining in Twitter – Sarcasm Detection" International Research Journal of Engineering and Technology (IRJET), volume 04, issue 10, pages 201-204.

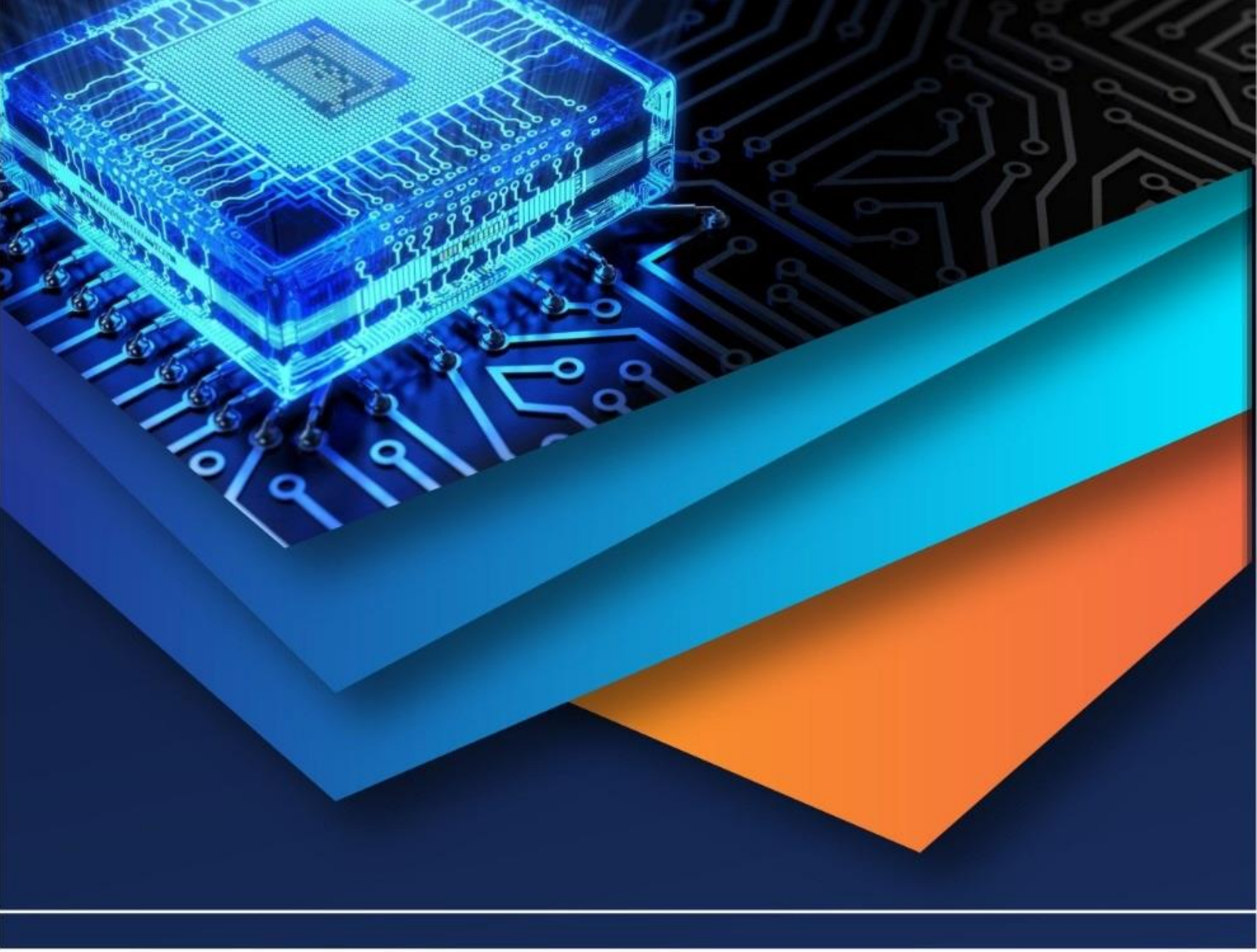
[15] Paras Dharwal, Tanupriya Choudary, Rajat Mittal, Praveen Kumar, "Automatic Sarcasm Detection using Feature Selection", International Conference on Applied and Theoretical Computing and Communication Technology, IEEE.

[16] Sindhu. C, G. Vaidhu, Mandala Vishal Rao, "A Comprehensive Study on Sarcasm Detection Techniques in Sentiment Analysis", International Journal of Pure and Applied Mathematics, volume 118, pages 433-442, 2018.

[17] Anusha, K. S., and A. D. Radhika. "A Survey on Analysis of Twitter Opinion Mining Using Sentiment Analysis.".

[18] K. Abdalgader and A. Al Shibli, ''Context expansion approach for graph-based word sense disambiguation,'' *Expert Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114313.

[19] A. Onan and M. A. Tocoglu, ''A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification,'' *IEEE Access*, vol. 9, pp. 7701–7722, 2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY