International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 4, May-June 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Sentimental Emotion Analysis using Python and Machine Learning

Mohit Chaudhari

Computer Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT

Sentiment analysis is used in opinion mining. It helps businesses understand the customers' reviews with a particular product by analyzing their emotional from the product reviews they post, the online recommendations they make, their survey responses and other forms of social media text. Businesses can get feedback on how happy or sad the customer is, and use this insight to gain a competitive edge. In this article, we explore how to conduct sentiment analysis on a piece of text using some machine learning techniques. Python happens to be one of the best programming language, when it comes to machine learning as it is easy to learn, is open source, and is effective in catering to machine learning requirements like processing big datasets and performing mathematical computations. Natural Language ToolKit (NLTK) is one of the popular packages in Python that can use for in sentiment analysis.

KEYWORDS: *Python, Sql, Haddop, Database, Machine Learning, Natural Language*

How to cite this paper: Mohit Chaudhari "Sentimental Emotion Analysis using Python and Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-4, June 2021, pp.147-149, URL: www.ijtsrd.com/papers/ijtsrd41198.pdf

IJTSRD41198

1. INTRODUCTION

With the Large amount of increase in the web technologies, the no of people expressing their views and the opinion via web. This information is useful for everyone like businesses, governments and individuals with 500+ million reviews per day, twitter is becoming a major source of information. Input to our model is the raw data extracted from reviews. For the same, we automate the process of text extraction and categorizing it into two categories i.e. positive or negative. The content in twitter generated by the user is about different kinds of products, event, people and political affair.

Performing sentiment analysis on text is considered best due to the following reasons:

1. Text are abstract in nature.
2. Analysis in real time can be done.
3. A vast variety of text for performing the analysis.

Our Proposal

The main reason of model of twitter data analysis will be implemented using Anaconda python. Anaconda is open source distribution of the Python and R programming languages for data science and machine learning related applications. It can also install on Windows, Linux, and MacOS. Conda is an free source, cross-platform, package management system.

The texts can be analysed and characterized based on the emotions used by the social users. We attempt to classify the polarity of the text where it is either positive or negative. If the text has both positive and negative elements, the more dominant sentiment should be take as the final label. We use the dataset from Kaggle which was crawled and labelled positive/negative. The data provided comes with emotions, usernames and hash tags which are required to be processed and converted into a standard form. It also needs to extract useful features from the text such as unigrams and bigrams which is a form of representation of the "text".

2. SOCIAL NETWORK ANALYSIS

Social network analysis is the study of people's Interactions and communications on different topics and nowadays it has received more attraction. Millions of people give their opinion of different topics on a daily basis on social medias like Facebook and Twitter. It has many applications in different areas of research from social science to business. Twitter nowadays is one of the popular social media which according to the statistic currently has over 300 million accounts. Twitter is the rich source to learn about people's opinion and sentimental analysis. For each text it is important to determine the sentiment of the text whether is it positive, negative, or neutral. Another challenge with twitter is only 140 characters is the limitation of each tweet which cause people to use phrases and works which are not in language processing. Recently twitter has extended.

3. HYPHOTHETICAL SUGGESTED APPROCH

An assumed number of reviews depend on whether or not it is ironic. Using a learning algorithm to classify after a tweet abstract, a set of features is assigned to fitting. The extraction of the feature is done to detect the sarcasm in reviews.

Hypothetical Data:

An assumed number of reviews depend on whether or not it is ironic. Using a learning algorithm to classify after a tweet

abstract, a set of features is assigned to fitting. The extraction of the feature is done to detect the sarcasm in reviews.

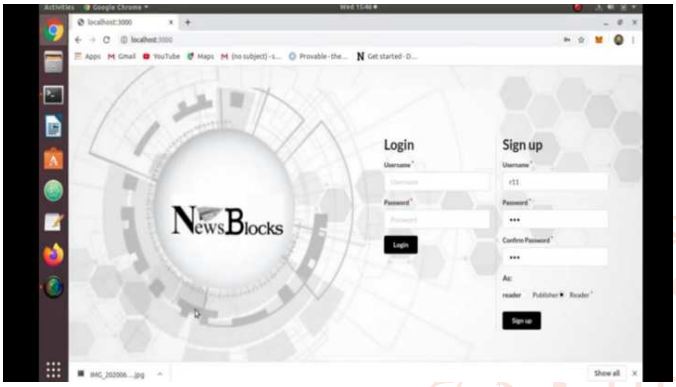
Mortgage data:

The Twitter's allow us to collect the reviews. To collect satirical tweet with the #sarcasm hash tag. Although the writer says # tag. However, works in the additional stressed that this # tag can be used for only a few purposes. However, the hash tag is not robust but can mainly be used for this purpose:

- Use as an anchor for research
- To find an irony marker in a truly sensitive sarcasm wherever it is extremely difficult to achieve

Induce a clear marker to ylack them, as "It was fun today. For the first time in weeks! #sarcasm

RESULT AND DISCUSSION



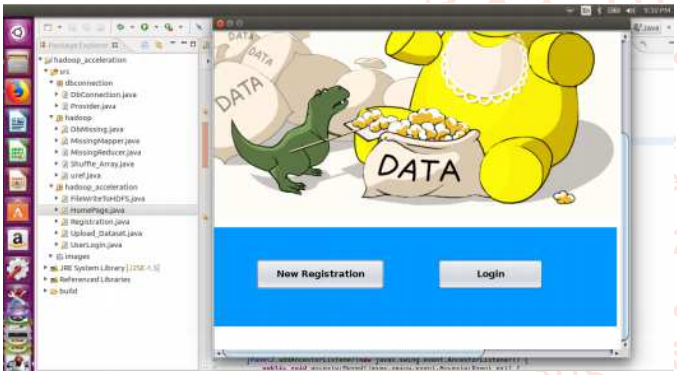


Fig Login Page

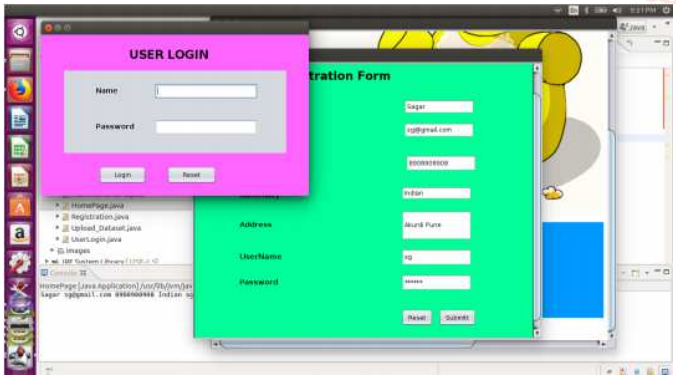


Fig User Registration

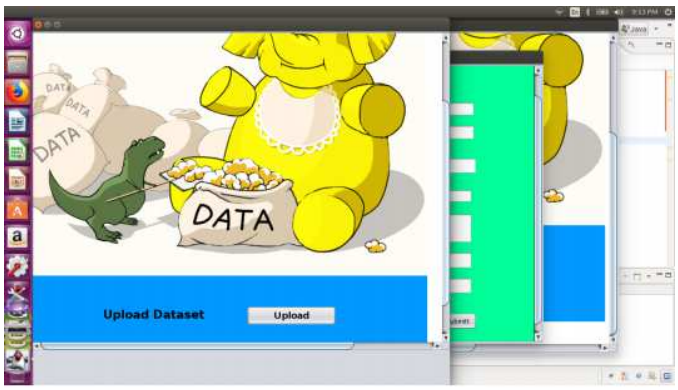


Fig Upload Dataset

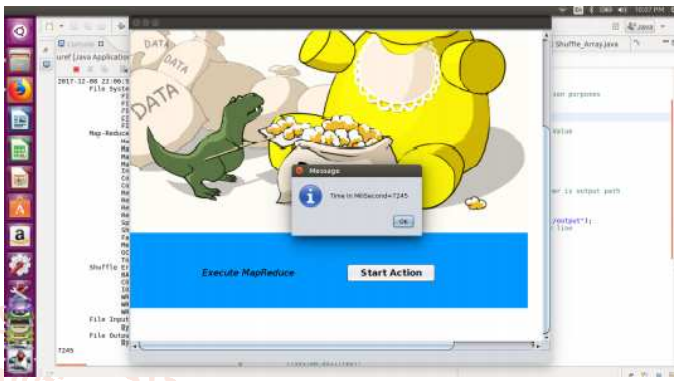


Fig Time Complexity

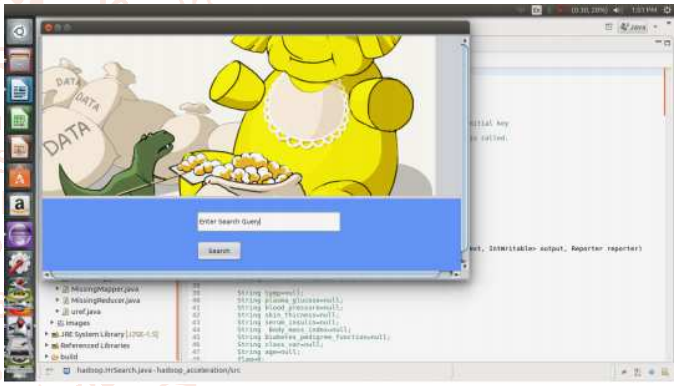


Fig Search Query

CONCLUSION

The analysis of Text format is being done in different points of view to mine the opinion or sentiment. Our proposed approach classify the texts as Positive and Negative texts which further helps in sentiment analysis and uses that sentiment analysis for further decision making. For our prototype, Twitter API is used to gather data in real-time. The prototype back-end tests on retrieving and processing the API data indicate that it is successful in gathering huge amounts of data from popular search terms in real-time. We will use various machine learning algorithms to conduct sentiment analysis using the extracted features. However, just relying on individual models did not give a high accuracy so we pick the top few models to generate a model.

REFERENCES

[1] Harshali P. Patil, Dr. Mohammad Atique, "Sentiment Analysis for Social Media: A Survey", Department of Computer Engineering Thakur College of Engineering & Technology Mumbai, India.

[2] Megha Rathi, Aditya Malik, Daksh Varshney, Rachita Sharma, Sarthak Mendiratta, "Sentiment Analysis of Reviews using Machine Learning Approach", Jaypee Institute of Information Technology JIIT Sec-62 Noida, India.

[3] Pitiphat Santidhanyaroj, Talha Ahmad Khan, "A SENTIMENT ANALYSIS PROTOTYPE SYSTEM FOR SOCIAL NETWORK DATA," Faculty of Engineering and Applied Science University of Regina Regina, Saskatchewan, Canada.

[4] Subramaniam.G, Ranjitha.M," User Emotion Analysis using Twitter Data", IFET COLLEGE OF ENGINEERING Cuddalore, India.

[5] M.Trupthi, Suresh Pabboju," SENTIMENT ANALYSIS ON TWITTER USING STREAMING API" 2017 IEEE 7th International Advance Computing Conference

[6] Rasika Wagh, Payal Punde," Sentiment Analysis using Twitter Dataset", Department of Computer Science and IT, Dr. BAMU Aurangabad, India.

[7] https://en.wikipedia.org/wiki/Twitter

[8] Po-Wei Liang, Bi-Ru Dai, "Opinion Mining on Social MediaData", IEEE 14th International Conference on Mobile Data Management, Milan, Italy, June 3- 6, 2013, pp 91-96, ISBN: 978-1494673-6068-5

[9] P. D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews," in Proceedings of the 40th annual meeting on association for computational linguistics, pp. 417–424, Association for Computational Linguistics, 2002.