Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, No. 4, July : 2023

SENTIMENTAL ANALYSIS BASED ON SOCIAL NETWORKING

Miss. **Monica S. Todkar,** M.Tech Student**,**
Dr. Jaydeep B. Patil*, Assistant Professor (M.Tech Guide),
Mr. **Rajwardhan S. Todkar,** M.Tech Student,
Dept. of Computer Science and Engineering (AI &ML), D. Y. Patil Agriculture and Technical University, Talsande, 416 112, India. **Email address:** jaydeep.patil@dyp-atu.org

Abstract
Web discussions are mostly used for information sharing, peer review, and publicity distribution. The majority of recent research on the preparation of content data focuses on the real world rather than the evaluation space, and content mining is seen as being essential to online data gathering and sentiment analysis. Social media is now used by users of all types to express their thoughts and experiences. Some people are neutral in their opinions, some have good aspects, and some have negative aspects. In a subsequent stage, the system does pre-process to remove noise from the collected tweets. then group data into good, negative, or neutral categories. The System uses visualization techniques to display sentiment scores. The separated hyperplane that maximizes margin across the various classes can be found using SVM. The practice of removing subjective information from text data to ascertain the sentiment or emotion represented by the author is known as sentiment analysis, also known as opinion mining. Sentiment analysis in social media has received a lot of attention as a result of the social networks' quick expansion.

Keywords: Sentimental analysis, opinion mining, SVMs, tokenization, stemming, and trained datasets are some of the keywords.

I. Introduction

Social media is now a major hub for information exchange. Everyday interactions on social media involve people starting conversations and others sharing their ideas. Social media websites are therefore a source of information for opinion mining. Organizations feel compelled to explore public opinion before making choices as a result of the extensive usage of the internet, the World Wide Web, and social media. However, to examine the polarity of opinions, accurate, intelligent information needs to be filtered. Thus, automated techniques for sentiment analysis and opinion mining are required. Sentiment analysis is viewed as a classification task that divides a text's orientation into positive, negative, and neutral categories. Sentiment analysis can be done at any one of three levels: document, phrase, or phrase level. To determine whether a document's attitude is good, negative, or neutral, a synopsis of the full document is taken first.

Support Vector Machines (SVM), a supervised machine learning technique, are used by this system to categorize tweets. The separated hyperplane that maximizes the margin for the various classes can be found using SVMs. In other words, the method generates an optimum hyperplane that classifies fresh testing data given labeled training data (supervised learning). SVM is efficient, precise, and does well with little training data. The third section of this paper provides an overview of the system design, and the fourth section explains the pre-processing stage. The proposed feature extraction and classification methods are discussed in the fifth section, which is followed by the experimental sentiment analysis results on crime and education in the sixth section.

Essentially, sentiment analysis on social media means keeping an eye on brand mentions and going over the text information acquired. Companies can use this information to gain analytical insight into the attitudes, routines, and feelings of their audience. What factors—and why—do they consider when mentioning your brand in a tweet? Additionally, Python is used in this project to write soft code. Why to utilize python is described in paper later still several of crucial points surrounding python surely mentioned i.e. Because of a variety of helpful features it offers, Python stands out from

other programming languages and is popular. Procedural programming is allowed, object-oriented programming is supported, and dynamic memory allocation is offered.

II. Literature

In the past, a variety of methodologies have been employed for opinion mining activities. Pang et al. [5], Mukras R.J [6], and others use data from product, movie, and consumer feedback reviews. They employ several statistical feature selection strategies and immediately deploy machine learning technologies. These observations demonstrate that the machine learning method alone does not perform well in the sentiment categorization task. They demonstrate that word presence or absence seems to be a better indicator of content than word frequency.

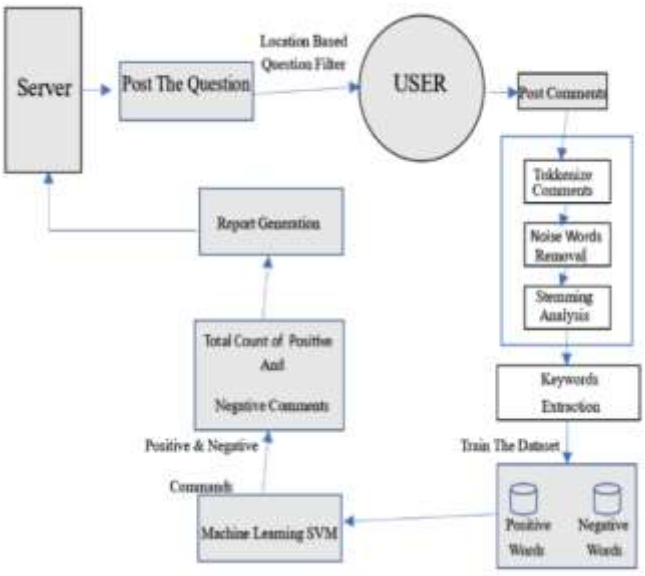
With the addition of various social site-specific cues such as retweets, hashtags, links, uppercase words, emotions, and exclamation and question marks, [7] classifies the subjectivity of social media messages based on traditional criteria. Further, a Part-Of-Speech (POS) specific prior polarity feature and a tree kernel are included in [8] to avoid the requirement for laborious feature engineering. Researchers have partially focused on this issue.

In [9] this study, the authors analyze Twitter, a major microblogging site, and create models for two classification tasks. These tasks can be divided into binary classifications of positive and negative emotions as well as three-way classifications of positive, negative, and neutral emotions. Additionally, they experimented with the unigram, feature, and tree kernel-based models.

III. Discussion

Sentimental Analysis

The term "feeling examination" refers to the methodical recognition, extrication, measurement, and concentration of numerous feeling states and emotional data using characteristic language preparation, content research, and computational semantics. The watchwords can be removed by the administrator after pre-processing. Apply support vector machine calculations to the watchword order as well. Straight SVM requires some direct polynomial math, which is outside the scope of this introduction to SVM, to complete the learning of the hyperplane.



System architecture

SVM Algorithm:

Support Vector Machine, sometimes known as SVM, is a linear model for classification and regression issues. It works well for many real-world issues and can solve both linear and non-linear problems.

The fundamental idea behind the SVM method is to partition the data into classes by drawing a line or a hyperplane. Components of SVM. SVM is capable of both linear and nonlinear behavior. Hard margin and soft margin are the two varieties of linear SVM. In hard margin SVM but not in soft margin SVM, a hyperplane can completely linearly segregate the data. It is simplest to understand the SVM approach by focusing on its primary type, the SVM classifier. In an N-dimensional space, the SVM classifier aims to produce a hyper-lane that splits the data points into several classes. However, as it provides the biggest margin out of the two, this hyper-plane is chosen.

Stepwise SVM Algorithm functioning:

First, the classes are predicted by the SVM method. The classes are given a number 1 for one and a number -1 for the other.

The second step is to transform the business issue into a mathematical equation with unknown variables, as is done by all machine learning algorithms. By turning the issue into an optimization issue, these unknowns are then discovered. In the case of the SVM classifier, a loss function called the hinge loss function is utilized and adjusted to obtain the maximum margin because optimization issues always aim at maximizing or minimizing something while searching for and adjusting for the unknowns.

Step 3: To make this loss function easier to grasp, it may also be thought of as a cost function, with a cost of 0 in the absence of any inaccurate predictions for any class. In contrast, error/loss is assessed if this is not the case. The existing system has a trade-off between margin maximization and the loss that results from highly aggressive margin maximization. An additional regularisation parameter is provided to theoretically ground these ideas.

Step 4: As with the majority of optimization issues, weights are optimized by computing gradients using advanced calculus ideas, namely partial derivatives.

Step 5: Only when there is no classification error using the regularisation parameter and when misclassification is present using the loss function are the gradients updated. Step 6: When there is no classification error, the gradients are just updated using the regularisation parameter. The loss function is also used when a classification error occurs.

- **Why Python:**

  Python offers a variety of helpful features that set it apart from other programming languages and make it popular. It offers dynamic memory allocation, allows procedural programming, and supports object-oriented programming. A few key characteristics are described below.

  1.Simple to Use and Understand

  Comparatively speaking to other programming languages, Python is simple to learn. Simple and largely identical to English, its syntax is easy to understand. No semicolon or curly braces are used; the indentation identifies the code block instead. Beginner programmers are advised to start with this language.

  2. Expression of Ideas

  Python only requires a few lines of code to complete complicated tasks. Just write print ("Hello World") to run the hello world program as an example. While Java or C require numerous lines to run, it simply needs one.

  3. Interpretation of Language

  Python is an interpreted language, which implies that each line of the Python program is run individually. The benefit of being an interpreted language is that debugging is quick and portable.

4. Language that is Cross-Platform
Python may function equally well on a variety of operating systems, including Windows, Linux, UNIX, Macintosh, etc. We can therefore claim that Python is a portable language. It enables programmers to create software for a number of competing platforms with only one piece of code.

5. Free and Open Source
Everyone can access Python for free. On the organization's main website, www.python.org, it is publicly accessible. A sizable global community is devotedly striving to create new Python modules and functions. The Python community is open to contributions from anyone. "Anyone can download its source code without paying any money," is what open-source software means.

6. Object-Oriented Language
Classes and objects are concepts introduced by Python's support for object-oriented languages. It supports polymorphism, encapsulation, inheritance, and other concepts. The object-oriented approach enables programmers to construct reusable code and create apps with fewer code.

7. Extendable
It implies that the code can be translated into other languages, such as C or C++, and that this will enable future use of the code in our Python programs. The program is transformed into byte code, which can be used on any platform.

8. Huge Standard Library
It offers a wide variety of libraries for many different professions, including web development, machine learning, and scripting. There are many machine learning libraries, including Tensor Flow, Pandas, Numpy, Keras, and Pytorch, among others. The most well-liked Python web development frameworks are Django, Flask, and Pyramids.
9. GUI Programming Assistance
The developing desktop application uses a graphical user interface. The libraries that are utilized for creating web applications are PyQT5, Tkinter, and Kivy.

10. Incorporated
Languages like C, C++, JAVA, and others are simple to integrate with it. Like Java, C, and C++, Python executes code line by line. It makes code debugging simple.

11. Embedded
The Python source code is compatible with other programming languages' code. Another programming language can also use Python source code. It can incorporate foreign languages into our programming.

12. Adaptive Memory Allocation
The variable's data type does not need to be specified in Python. It automatically allots RAM to the variable at run time when we assign it a value. We don't need to write int x = 15 if the integer value allocated to x is 15, for example. Put simply, x = 15.

IV. Objective:
1. To analyze the positive and negative responses on a social network by SVM Algorithm.
2. To enhance the sentiments of End users by networking.

3. To minimize the time consumption of the user at the end of a task.

V. Methodology

1. Support vector machine (SVM) algorithm is used to solve classification and regression issues.
2. Python code implementation for persuasion.
3. Use of the stemming and tokenization analysis procedure.

VI. Results:

Based on factors like the clients' location, age, gender, and employment, the administrator can manipulate their emotions. the summary of every observation. enhanced sensation mining's accuracy in the Python structure. channels the feelings based on client regions.

VII. Conclusion

That content conclusion review is also known as the passionate extremity analysis. Has grown into a prosperous outpost in the network of content mining. This study looks at how to find and measure internet-gathering hotspots using angle analysis and content-mining approaches. The material obtained through information collecting creates an enticing source of knowledge for speculation and concept analysis. A client can receive various examples of information on the most popular trends thanks to an area-based hypothesis analysis of a large number of client records.

VIII. References

1. Sathish Kumar T Assistant Professor, Department of CSE Saranathan College of Engineering Trichy, Tamil Nadu, India. Fellow “Sentimental Analysis (Opinion Mining) in Social Network by Using SVM Algorithm.”
2. A machine learning-based classification for social media messages R. Nivedha and N. Sariam.
3. Sentiment analysis of smartphone product review using SVM algorithm-based particle swarm optimization
4. docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm /introduction_to_svm.html
5. Pang B et al. Thumbs up? Sentiment Classification Using Machine Learning Techniques, Procs. Of the Conference on Empirical Methods in Natural Language Processing (EMNLP), ACL Press. 2002, pp 79-86.
6. Mukras R, Carroll J. A comparison of machine learning techniques applied to sentiment classification. 2004, pp 200-204.
7. Barbosa L, Feng J. Robust sentiment detection on Twitter from biased and noisy data. In: Proceedings of COLING. 2010, pp. 3644.
8. Salloum, Said A., et al. "A survey of text mining in social media: Facebook and Twitter perspectives." Adv. Sci. Technol. Eng. Syst. J 2.1 (2017): 127-133.
9. Balazs, Jorge A., and Juan D. Velásquez. "Opinion mining and information fusion: a survey." Information Fusion 27 (2016): 95-110.
10. Megha Rathi, Aditya Malik, Daksh Varshney, Rachita Sharma, Sarthak Mendiratta. “Sentiment Analysis of Tweets using Machine Learning Approach”2018.