Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 07, July -2018

TWITTER CLASSIFICATION OF THE COMMENTS USING ARTIFICIAL NEURAL IN SOCIAL MEDIA

Miss. Pooja Giri

DY Patil College of Engineering, Ambi, Pune

Abstract: *With the advancement in net technology, there's vista amount of data gift within the web world. These info or information is provided by the web the web the net} users by accessing the information on the market on internet and every now and then giving necessary feedback, so adding a lot of helpful info. Thanks to the presence of enormous quantity of data like user views, feedback and suggestions on the market on the online, it's become extremely essential to explore and analyze their views for higher understanding and deciding. Opinion mining conjointly referred to as sentiment analysis could be a technique victimisation which of them views or sentiments are often classier into positive or negative comments. The most standard communication tool currently days is small blogging. Conjointly great deal of knowledge is shared on the social networking sites. so social networking sites will function an excellent platform for sentiment analysis. We have done state of the art rigorous literature survey in sentiment analysis. We've got projected algorithms that perform higher and square measure a lot of efficient in evaluating the opinions.*

Keywords *- Sentiment Analysis, Opinion Mining, Stream Data Analysis, Polarity Detection, Sentiment Classification, Keyword Correlation, Natural Language Processing, Sentiment Mining, Twitter*

I. Introduction:

With the advancement in internet technology, there's large quantity of data gift within the web world. These data or knowledge is provided by the web the web the net users by accessing the information obtainable on internet and sometimes giving necessary feedback, therefore adding additional helpful data. Because of the presence of enormous quantity of data like user views, feedback and suggestions obtainable on the net, it's become extremely essential to explore and analyze their views for higher understanding and deciding. Opinion mining additionally known as sentiment analysis may be a technique exploitation which of them views or sentiments may be classified into positive or negative comments. With the explosion of internet a pair of.0, varied styles of social media reminiscent of blogs, discussion forums and peer-to-peer networks gift a wealth of knowledge which will be terribly useful in assessing the overall public's sentiment and opinions towards product and services. Recent surveys have disclosed that opinion-rich resources like on-line reviews are having larger economic impact on each shoppers and firms compared to the standard media. Driven by the demand of gleaning insights into such nice amounts user-generated knowledge, work on new methodologies for machine-controlled sentiment analysis and discovering the hidden information from unstructured text knowledge has bloomed splendidly. Among varied sentiment analysis tasks, one amongst them is sentiment classification, i.e., characteristic whether or not the linguistics orientation of the given text is positive, negative or neutral. though abundant work has been exhausted this line, most of the present approaches have faith in supervised learning models trained from labelled corpora wherever every document has been labelled as positive or negative before coaching. However, such labelled corpora don't seem to be perpetually simply obtained in sensible applications.

II. Literature Survey

[1]. Twitter, as a social media could be a very fashionable method of expressing opinions and interacting with others within the online world. once taken in aggregation tweets will give a reflection of public sentiment towards events. In this paper, we offer a positive or negative sentiment on Twitter posts employing a well-known machine learning methodology for text categorization. additionally, we tend to use manually tagged (positive/negative) tweets to make a trained methodology to accomplish a task. The task is searching for a correlation between twitter sentiment and events that have occurred. The trained model is predicated on the Bayesian supplying Regression (BLR) classification methodology. we tend to used external exicons to notice subjective or objective tweets, added Unigram and written word options and used TF-IDF (Term Frequency-Inverse Document Frequency) to separate the features. victimization the FIFA World Cup 2014 as our case study, we used Twitter Streaming API and a few of the official world cup hashtags to mine, filter and method tweets, in order to investigate the reflection of public sentiment towards unexpected events. constant approach, may be used as a basis for predicting future events.

[2]. Sentiment analysis or opinion mining aims to use automatic tools to sight subjective info resembling opinions, attitudes, and feelings expressed in text. This paper proposes a completely unique probabilistic modeling framework referred to as joint sentiment-topic (JST) model supported latent Dirichlet allocation (LDA), that detects sentiment and

topic at the same time from text. A reparameterized version of the JST model referred to as Reverse-JST, by reversing the sequence of sentiment and topic generation within the modelling method, is also studied. Though JST is like Reverse-JST while not hierarchal previous, in depth experiments show that once sentiment priors square measure other, JST performs systematically higher than Reverse-JST. Besides, not like supervised approaches to sentiment classification which often fail to provide satisfactory performance once shifting to different domains, the weakly-supervised nature of JST makes it highly moveable to different domains. this is often verified by the experimental results on datasets from 5 totally different domains wherever the JST model even outperforms existing semi-supervised approaches in a number of the datasets despite victimisation no labeled documents. Moreover, the topics and topic sentiment detected by JST square measure so coherent and informative. we have a tendency to theorize that the JST model will pronto meet the demand of large-scale sentiment Analysis from the net in an open-ended fashion.

[3]. Description: Sentiment analysis seeks to spot the viewpoint(s) underlying a text span; associate degree example application is classifying a motion-picture show review as "thumbs up" or "thumbs down". to work out this sentiment polarity, we propose a unique machine-learning methodology that applies text-categorization techniques to only the subjective parts of the document. Extracting these parts is enforced victimisation economical techniques for locating minimum cuts in graphs; this greatly facilitates incorporation of cross-sentence contextual constraints.

[4]. Merchants marketing merchandise on the online typically raise their customers to share their opinions and active experiences on products they need purchased. Sadly, reading through all client reviews is tough, particularly for widespread things, the number of reviews are often up to a whole bunch or maybe thousands. This makes it tough for a possible client to browse them to make AN au fait call. The Opinion Miner system designed in this work aims to mine client reviews of a product and extract high elaborated product entities on that reviewers express their opinions. Opinion expressions area unit known and opinion orientations for every recognized product entity area unit classified as positive or negative. Totally different from previous approaches that used rule-based or applied math techniques, we propose a unique machine learning approach designed below the framework of linguistic process HMMs. The approach naturally integrates multiple necessary linguistic options into automatic learning. During this paper, we tend to describe the design and main components of the system. The analysis of the projected method is bestowed supported process the net product reviews from Amazon and different publically offered datasets.

[5]. In this paper, we have a tendency to specialise in object feature one based review report. Completely different from most of previous work with linguistic rules or statistical ways, we have a tendency to formulate the review mining task as a joint structure tagging downside. We propose a replacement machine learning framework supported Conditional Random Fields (CRFs). It will use made options to jointly extract positive opinions, negative opinions and object options for review sentences. The linguistic structure are often naturally integrated into model illustration. Besides linear chain structure, we have a tendency to conjointly investigate conjunction structure and syntactical tree structure in this framework. Through in depth experiments on motion-picture show review and merchandise review data sets, we have a tendency to show that structure-aware models outperform several progressive approaches to review mining.

[6]. In this paper, we have a tendency to target the opinion target extraction as a part of the opinion mining task. we have a tendency to model the matter as associate degree data extraction task, that we have a tendency to address based on Conditional Random Fields (CRF). As a baseline we have a tendency to use the supervised formula by Tai et al. (2006), that represents the progressive on the utilized data. we have a tendency to assess the algorithms comprehensively on datasets from four totally different domains annotated with individual opinion target instances on a sentence level. Moreover, we investigate the performance of our CRF-based approach and therefore the baseline during a single- and cross-domain opinion target extraction setting. Our CRF-based approach improves the performance by 0.077, 0.126, 0.071 and 0.178 relating to F-Measure within the single-domain extraction in the four domains. Within the cross domain setting our approach improves the performance by 0.409, 0.242, 0.294 and 0.343 relating to F-Measure over the baseline.

[7]. —Design and development of energy efficient routing protocols for Wireless Sensor Network (WSN) is one of the active research fields. Cluster based routing protocols have proven to be energy efficient and LEACH is one of most popular cluster based routing protocol for WSN. But, LEACH suffers from several drawbacks such as possibility of choosing a low energy node as Cluster Head (CH), non-uniform distribution of CHs, etc. In this paper EiP-LEACH (Energy influenced Probability based LEACH) protocol is proposed which is an enhanced version of LEACH protocol that is influenced by the energy parameter for CH selection. EiP-LEACH helps in deciding the better CH nodes and thereby contributes towards network life prolongation. EiPLEACH is compared with basic LEACH in terms of number of alive nodes, average energy depletion, First Node Dead (FND) and Last Node Dead (LND) and found that EiP-LEACH is farbetter.

III. Scope of the project:

Sentiment analysis is in itself an innovative technique using which the vast amount of information present in the internet world can be analyzed. These information or data may include internet users feedback or reviews on various sites.

Opinion mining also called as sentiment analysis helps in extracting, exploring and analyzing the data which helps in tapping into the minds of internet users for knowing their opinion. Sentimental analysis of social networking sites can be useful in various domains .As social networking sites are the most visited sites by today's internet users. Therefore any opinion of the user can be analyzed for feedback purpose, product and movie reviews, or just to get an insight to the users mind. The system will be able to identify and classify sentiments represented in digital text. This will enable the people to analyze and organize the views and therefore make better decisions.

IV. Objectives of the project:

The main objectives of the project are to analyze tweet using classification & sentiment mining. In This project we have used Naive bayes as a classification algorithm which takes feature of tweets an train using history data. We have also used Negation algorithm & it applied on text. The main motivation behind of building this project to analyzed on which location there are maximum crime tweets comes.

V. Mathematical model:

Let S be the system where,
S= {S, e, I, O,F, I}
Let S: Start state
E: end state
Let I be the set of Inputs Where
I = { I1, I2,…..In},{referred as set of comments}
O be the set of Output where,
O = {O1, O2…..On} { Positive/Negative comments}
F be the set of function … where
F= {F1…… Fn}
F1=download .comments ()
F2= Parse comments ()
F3= Remove stop words ()
F4= Apply Negation Algorithm ()
F5= Extract Feature ()

VI. Algorithms

1. **Negation algorithm**

```
Check if ➔ Negation;
{
Check for next word
{
    If(next word is positive)
    {Score  positive}
                else
                   {Score Negative}
    }
          Else
    {
    If(positive word)
    {
    Positive
    }
    else
    {
    Negative;
    }
      }
       If(pos>neg)==. Positive
      Else negative
       }
```

2. Navie Bayesian

- It has Three Steps:

1. Initial Problem
2. Conditional Probability

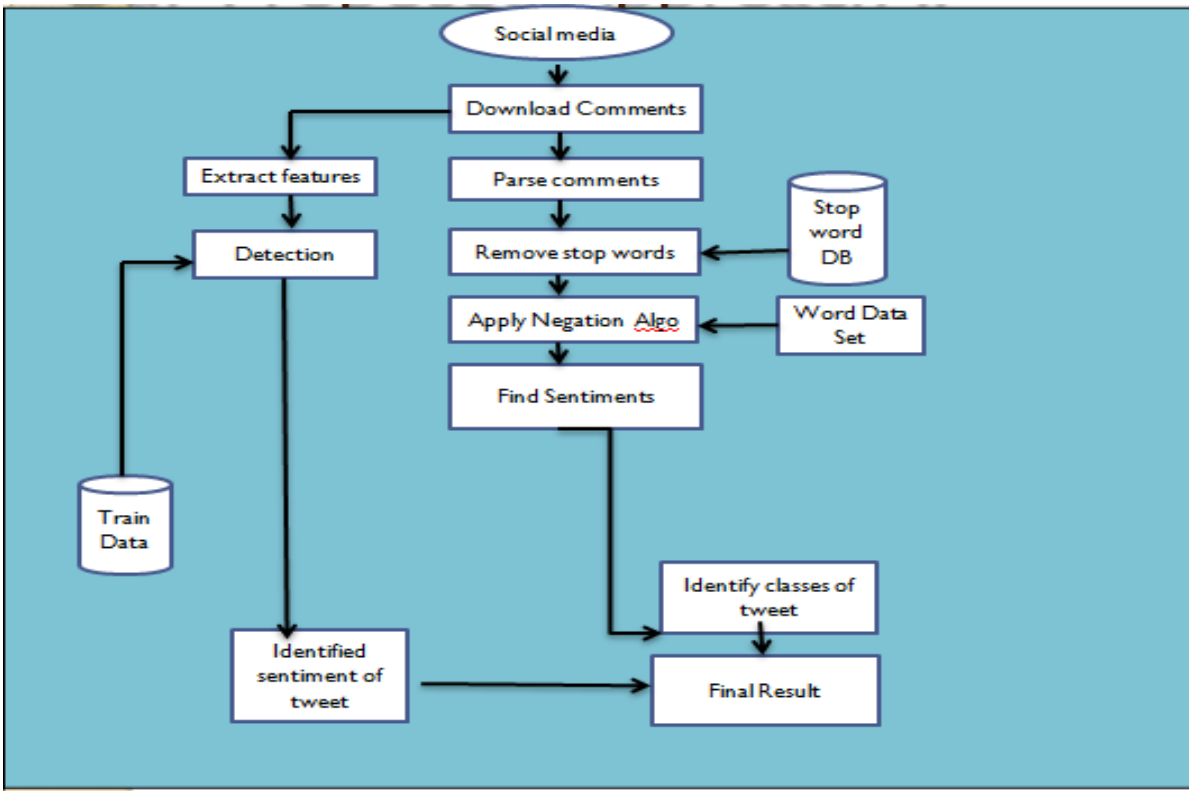
3. Final Probability

$$P(ai|vj) = \frac{nc + mp}{n + m}$$

where:

- n = the number of training examples for which v = vj
- nc = number of examples for which v = vj and a = ai
- p = a priori estimate for P(aijvj)
- m = the equivalent sample size

VII. System Architecture:



VIII. Results

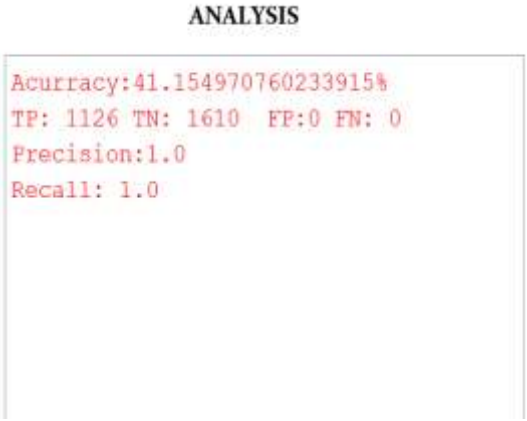
1. Proposed Accuracy

ANALYSIS

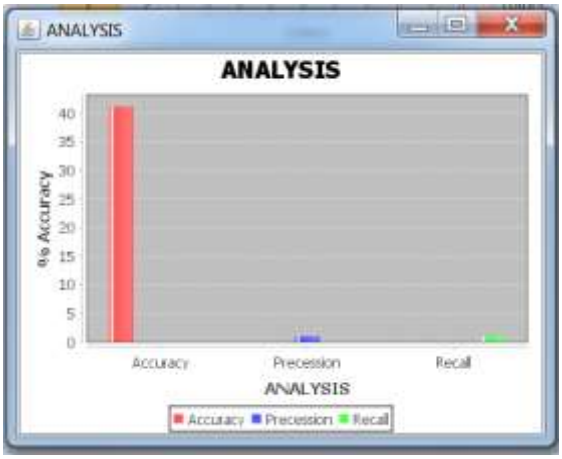
```
Acurracy:58.845029239766085%
TP: 1610 TN: 0  FP:0 FN: 1126
Precision:1.0
Recall: 0.5884502923976608
```


2. Proposed Graph

3. Existing Accuracy only Navie Bayesian



4. Existing using only Navie Bayesian



IX. Conclusion:

In this system having sentiment analysis technique using Negation algorithm and Navie Bayesian gives a advantage as It will be accurate, identification of the sentiments from the data will be faster, efficient classification of the identified sentiments will be done.

Sentiment analysis is in itself an innovative technique using which the vast amount of information present in the internet world can be analyzed. Data Analysis can be made possible for all the social networking sites the system will be able to identify and classify sentiments represented in digital text. This will enable the people to analyze and organize the views and therefore make better decisions.

VIII. References:

[1]. Peiman Barnaghi and John G. Breslin, Parsa Ghaffari "Opinion Mining and Sentiment Polarity on Twitter and Correlation Between Events and Sentiment" 978-1-5090-2251-9/16 $31.00 © 2016 IEEE DOI 10.1109/BigDataService.2016.36
[2]. Chenghua Lin, Yulan He, Richard Everson" Weakly-supervised Joint Sentiment-Topic Detection from Text"2017 IEE E/ACM 39th IEEE International conference
[3]. Xing Fang and Justin Zhan," Sentiment analysis using product review data" Fang and Zhan Journal of Big Data (2015) 2:5DOI 10.1186/s40537-015-0015-2
[4]. Peiman Barnaghi Parsa Ghaffari, " Opinion Mining and Sentiment Polarity on Twitter and Correlation Between Events and Sentiment."Big data Computing service and application Year: 2016 IEEE
[5] Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu " Structure-Aware Review Mining and Summarization" Machine learning, vol. 39, pp. 135-168, 2000.
[6]. Niklas Jakob, Iryna Gurevych," Extracting Opinion Targets in a Single- and Cross-Domain Setting with Conditional Random Fields", Korea under the ITRC support program supervised by NIPA
[7]. Bongale, Anupkumar M., Anand Swarup, and Shashank Shivam. "EiP-LEACH: Energy influenced probability based LEACH protocol for Wireless Sensor Network." IEEE International Conference on Emerging Trends & Innovation in ICT (ICEI), 2017
[8]. Bongale, Anupkumar M., and C. R. Nirmala. "EOICHD: A Routing Scheme for Wireless Sensor Network Based on Energy and Optimal Inter Cluster Head Distance." International Journal of Applied Engineering Research, Vol. 11, Issue 11, 7256-7266, 2016.
[9]. Malmaz Roshanaei and Shivakant Mishra," An Analysis of Positivity and Negativity Attributes of Users in Twitter" 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)
[10] V. N. Vapnik and V. Vapnik, Statistical learning theory vol. 1: Wiley New York, 1998.
[11] G. Salton and M. J. McGill, "Introduction to modern information retrieval," 1986.
[12] S. Dumais, J. Platt, D. Heckerman, and M. Sahami, "Inductive learning algorithms and representations for text categorization," in Proceedings of the seventh international conference on Information and knowledge management, 1998, pp. 148-155.
[13] S. M. Weiss, C. Apte, F. J. Damerau, D. E. Johnson, F. J. Oles, T. Goetz, et al., "Maximizing text-mining performance," IEEE Intelligent systems, pp. 63-69, 1999.
[14] R. Feldman, "Techniques and applications for sentiment analysis," Communications of the ACM, vol. 56, pp. 82-89, 2013.
[15] S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," in LREC, 2010, pp. 2200-2204.
[16] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in LREC, 2010, pp. 1320-1326.
[17] T. Blog, "Insights into the #WorldCup conversation on Twitter," in Twitter Blog, ed, 2014.

