Analysis Of Behavior Extraction On Social Life Issues Using Sentiment Analysis: A Review

Obaidullah, Dr. Faiyaz Ahmad

Abstract— In modern days there is a different type of services which are popular in the social site and use for different purpose like FB, Twitter, etc. For the forecast the sentiment analysis we utilized information store on the social site stockpiling. Our goal is to retrieve data from a social site and our machine will decide what is the emotion of a particular person related to any topics. Twitter is an example of a social site where lots of data are stored in a structured and unstructured format. Since many useful links with a various field like politics, education, health, agriculture, religious, marketing post on twitter so we fetch the data from Twitter for sentiment analysis. When we take a shot at unstructured information then numerous issues emerge in text mining. There is a different idea like artificial intelligence, machine learning, deep learning which are utilized for handling the unstructured information. Be that as it may, we are applying machine learning to deal with the unstructured information.

Keywords — Machine Learning, Opinion Mining, NLP, Hybrid, automated classification.

I. INTRODUCTION

Present time preserve microblog websites play an important role in our life, social relationship as well as act as a crucial information source. Every day lots of users utilizing microblog sites for expressing their views, opinions, experiences, and feelings about specific topics. There is various kind of social sites by which we can speak with one another, similar to Twitter, Facebook, and another person to person communication sites.

nowadays we have lots of data related to different department like health, education, sport, politics, religion. Social media containing twitter creates a space for explanation and circulation of thoughts and opinions on distinct topics and disparate events. Twitter with 600

a million users and 250 million messages per day have become a beneficial resource for organization s to monitor their own character. Organizations invoke posted emotions in tweets about their achievement and thus can make marketing activity again most competitors. These tweets show the different behavior of a person that the person belongs to which field. Here we will analyze the sentiment of a person and different areas people mindset which type of thinking and behavior. We can find the interest of the people they are religious, political, sportsman, etc. In this system, we monitor the tweets of the peoples what they want. which type of mode of people what they are demanding and what they are declining. and we also try to know what are the thinking and what are the views of the hot topic. For example, we monitor the people's views of the election. Then we collect the twitter data and try to identify what are the opinions of the public on the parties. Also, which party they give a positive feeling and which party they give a negative sentiment. We can also identify according to the tweets which leader they are liked and which type of leader they want. Here we are enhancing a system that takes data on Twitter and tries to identify the thinking of people on the basis of the tweets. through this mechanism, the system can automatically identify people thinking according to the tweets.

A. Sentiment Analysis Scope

Document-level sentiment analysis gets the sentiment of a total report or passage.
Sentence level sentiment analysis gets the sentiment of a single sentence.
Sub-sentence level sentiment analysis acquires the sentiment of sub-expressions inside a sentence.

B. Sentiment Analysis Classification

There are numerous techniques and calculations to execute sentiment analysis systems, which can be named
.
1. Machine learning:- It is the use of AI. It gives the capacity to the system to learn things naturally and enables

International Journal of Research in Advent Technology, Vol.7, No.6, June 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

programming applications to turn out to be progressively exact without being explicitly programmed. They can recognize the text which are express the sentiments. Automatic systems that depend on machine learning techniques to learn from the information.

- **Supervised Learning Techniques:-** Support Vector Machines(SVM), Bayesian Network, Naïve Bayes Classification, Decision Tree are most generally utilized methods.
- **Unsupervised Learning Techniques:-** Clustering Algorithm, Matrix factorizations are unsupervised learning procedures.

2. Rule-based systems that perform sentiment analysis dependent on a lot of manually crafted rules.

3. Hybrid:- The combination of machine learning and dictionary-based ways to deal with location Sentiment Analysis is called Hybrid. Despite the fact that not regularly utilized, this strategy generally delivers more encouraging outcomes than the methodologies. Hybrid systems that join both rule-based and automatic approaches.

II. LITERATURE SURVEY

The following papers are surveyed in this section:-

In this paper [1] The author considered a neural network-based approach to analyzing a sentiment based on political tweets firstly the author represent the text by dens vector which consists of subword information that helps in detecting word similarities more accurately by utilizing both morphology and semantic/. After that, a CNN was trained to learn so as to differentiate the tweets depending on sentiment, on the basis of the available labeled data set. Then the model is applied to implement the sentiment analysis of a cluster of tweets collected during the days before UK general election the result shows that 29% positive tweets, 30% negative tweets and 93% neutral are done.

In this paper [2] The author proposed an automated topic modeling technique LDA which help in identifying the attracted matter of arguments from large scale tweets co-related to any two leading political leaders of India. In this paper, the author takes a spark R framework on which topics modeling techniques are performed to upgrade to speed and performance for large scale social data processing and its analysis in real time. At that point sentiment analysis, those tweets are finished by utilizing dictionary based ways to deal with arrange the general population's sentiment against the two leaders. The result shows that the proposed method has improved the speed compared to the normal R tool.

[3] Nowadays various existing studies offense use emotions as noise emotion levels or identical emotion indicator to effectively trained the classifier but ignore their emotional capacity. In this paper, the author takes this issue by creating an emotional space as a feature representation matrix and projecting emoticons and words into the emotional space on the basis of semantic composition.so to improve the perform of analyzing the sentiment, the author proposes a new emotion semantic enhanced convolutional neural network (ECNN) model result shows that the proposed model is better than the other model on the data set of several sentiment tasks.

In this paper [4] word embedding was introduced by the author which is obtained by unsupervised learning on huge twitter's text collection that uses latent contextual semantic relationship and their co-occurrence statistical characteristics between the word in tweets this approach connects the pre-trained word embedding feature which was generated by using gloVe word sentiment voting features based sentiment glossary and N-grams features as the sentiment feature vector of the tweets and also tacking in the feature set to a deep CNN. This model implements better state-of-the-art methodologies and baseline model.

In this paper [5] The author is a machine learning for handling a huge amount of data which is stored on social network service such as Twitter that also has huge data of users posts. So the author tried to do research on text mining of these post. The author uses 1000 data set of each positive and negative for training or testing this experiment is done by using tenser flow for creating the network based on the experiment the accuracy for Twitter sentiment analysis is 77.45% for train and test set for (MLP) multilayer perceptron the accuracy is 67.45% and 52.60% for train and test set.

In this paper [6] An analysis for sentiment behavior of Twitter data is shown which was applied on large data set in terms of positive, negative and neutral. The author indicates how to collect original collection for sentiment and refinement that is essential with such type of data and also the author applied a hybrid of naïve Bayes and fuzzy classifier this data set to execute sentiment analysis. The outcome demonstrates that the proposed algorithms give better execution.

In this paper [7] The author rewired various technologies and approaches that help in the gathering of information to the system. This selection of classification models is done on the basis of training time available ,resources accuracy requirement etc.naive Bayes classifier is chosen when processing power and memory is an issue because of its low memory processing power requirements whereas Max-entropy is taken when less training time available but powerful processing system and memory is provided but Max-entropy is a bit costly.

In this paper [8] The author tracks the power of deep CNN to illustrate sentences and implement sentiment analysis this approach helps in processing the whole sentence

International Journal of Research in Advent Technology, Vol.7, No.6, June 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

automatically otherwise it would be done by advanced NLP technique. this approach helps in analyzing acceptable content based relation from word level to phrase level to sentence level and then document level by using this approach an accuracy of 80.69% is reached.

In this paper [9] The author proposed a structure for the sentiment analysis of geospatial of the disaster-related internet-based life information object. The author fundamentally addresses three kinds of difficulties, the inaccuracy, and variety identified with different content and pictures sentiment classifier, the geo-sentiment varieties among Different information object in the nearby land territory and watched particular sentiment from an interactive media information article to defeat the difficulties the author proposed system. which comprise of sentiment analysis, Special worldly dividing and geo-sentiment molding.

In this paper [10] The author discussed the motivation behind the use of sentiment analysis and opining mining In terms of companies and the government bodies, to improve their product and service. The author provides some techniques which were mainly based on specially teller deep learning architecture or machine learning method this research is based on fault. Firstly the author analyzes the prompt feature for machine learning approaches with the aim to integrate external knowledge into machine learning techniques.secondly for the better representation of text the author experiments with different Deep learning architecture.

In this paper [11] The author proposed a language-agnostic translation free method for Twitter sentiment analysis, which helps in making use of deep CNN with character level embedding for showing the proper polarity of tweets that were written in different languages result shows that the proposed method was more accurate than other deep neural architecture while involving considerably less parameter. The knowledge learned by this proposed method helps in qualitatively verifying its effectiveness for the classification of tweets sentiment.

In this paper [12] The author main object is to provide an interactive automatic system which figures out the emotions of tweets of the people posted in social media by applying Hadoop, that can process a large amount of data there are few challenges arise, namely, sentiment classification, feature-based classification handling negations. The main aim of the proposed method is to execute real-time sentiment analysis on the tweets which were extracted from Twitter. The training module of the processed system is done with the help of Hadoop and map reduced were as classification based on naïve Bayes, time variant analytics and the continuous learning system. The advantage is proposed system is it's cost-effective and the analysis is done at real time. The disadvantage of the proposed system is its is being unigram naïve.

In this paper [13] has introduced a new data set of an Arabic language which was mainly an opinion on various health services and these data set were collected from Twitter. In this paper, the author first collected the data from twitter and then the processor of filtering, preprocessing and defining the Arabic content in order to build a big data set. In this paper, the author uses various machine learning algorithms such as naïveBayes, SVM and logistic regression side by side deep CNN was utilized for sentiment analysis an health data set. The result a having accuracy in between 85% and 91% and the best classifier was SVM using leaner support vector classification and stochastic gradient descent.

In this paper [15] The author proposed an innovative system to select the behavioral information from tweets for context awareness services. In this paper, the author takes real Japanese language and extracts only those words that indicate behavior the experiments is done with precision, recall, and f-mission. The result shows that the proposed method has higher precision and recall then the existing method also the proposed method do not require to prepare a dictionary during its use or in advance which helps in reducing much load for humans. The proposed method is able to handle both common words as well as twitter specific words as effective features.when the system extracts behavioral information by using character N-gram tokenization.

III. CONCLUSION

Nowadays Sentiment Analysis is a very popular area of research to analyze the human thinking pattern. This area is taking more growth in many fields like business, politics, education, health, etc, In this Survey paper, we study the different algorithm and method used by the various researchers in these algorithms have many pros and cons.many algorithm some advantage specific task but it has some disadvantage of tasks. In my survey, many hybrid approaches are better than to the traditional algorithms. hybrid approaches give better results and accuracy.

IV. FUTURE WORK

Future work includes a broad examination of various text mining and sentiment analysis approaches on various data sets collections gained from multiple resources and in different languages. We will also work towards finding most computationally inexpensive algorithms for various tasks and sub-tasks. Today people are sharing their views on a different topic just like Politics, Education, Religion, Dearness, and corruption. Some people give positive and some people give negative views on these topics. In this method, we take Twitter data to analyze to know People's behavior. And give the overall positive and negative views on the different topic what exactly they want. In the future, it is easy to understand

International Journal of Research in Advent Technology, Vol.7, No.6, June 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

the people problem. Various prediction applications will also be studied.

REFERENCES

[1] Pota, M., Esposito, M., Palomino, M. A., & Masala, G. L. (2018, May). A Subword-based Deep Learning Approach for Sentiment Analysis of Political Tweets. In *2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA)* (pp. 651-656). IEEE.

[2] Monish, P., Kumari, S., & Babu, C. N. (2018, July). Automated Topic Modeling and Sentiment Analysis of Tweets on SparkR. In *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

[3] Yang, G., He, H., & Chen, Q. (2018). *Emotion-Semantic Enhanced Neural Network. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 1–1.*

[4] Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep convolution neural networks for Twitter sentiment analysis. *IEEE Access*, *6*, 23253-23260.

[5] Ramadhani, A. M., & Goo, H. S. (2017, August). Twitter sentiment analysis using deep learning methods. In *2017 7th International Annual Engineering Seminar (InAES)* (pp. 1-4). IEEE.

[6] Mehra, R., Bedi, M. K., Singh, G., Arora, R., Bala, T., & Saxena, S. (2017, July). Sentimental analysis using fuzzy and naive Bayes. In *2017 International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 945-950). IEEE.

[7] Hazra, T. K., Bihani, Y., & Mishra, S. (2017). *Sentiment learning using Twitter ideograms. 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON).*

[8] Chachra, A., Mehndiratta, P., & Gupta, M. (2017, August). Sentiment analysis of text using deep convolution neural networks. In *2017 Tenth International Conference on Contemporary Computing (IC3)* (pp. 1-6). IEEE.

[9] Alfarrarjeh, A., Agrawal, S., Kim, S. H., & Shahabi, C. (2017, October). Geo-spatial multimedia sentiment analysis in disasters. In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 193-202). IEEE.

[10] Alexander Gelbukh. (2017), Sentiment Analysis and Opinion Mining. *6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*.

[11] Wehrmann, J., Becker, W., Cagnini, H. E., & Barros, R. C. (2017, May). A character-based convolutional neural network for language-agnostic Twitter sentiment analysis. In *2017 International Joint Conference on Neural Networks (IJCNN)*(pp. 2384-2391). IEEE.

[12] Trupthi, M., Pabboju, S., & Narasimha, G. (2017, January). Sentiment analysis on twitter using streaming API. In *2017 IEEE 7th International Advance Computing Conference (IACC)* (pp. 915-919). IEEE.

[13] Alayba, A. M., Palade, V., England, M., & Iqbal, R. (2017, April). Arabic language sentiment analysis of health services. In *2017 1st International Workshop on Arabic Script Analysis and Recognition (ASAR)* (pp. 114-118). IEEE.

[14] Minab, S. S., Jalali, M., & Moattar, M. H. (2015). *A new sentiment classification method based on the hybrid classification on Twitter. 2015 International Congress on Technology, Communication, and Knowledge (ICTCK).*

[15] Yano, Y., Hashiyama, T., Ichino, J., & Tano, S. I. (2014, July). Behavior extraction from tweets using character N-gram models. In *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1273-1280). IEEE.

AUTHORS PROFILE

Obaidullah is a student of the integral university. He achieved his b.tech degree from the integral university and currently, he has pursuing M.tech in Computer Science and Engineering from the integral university.his Research work in "Analysis of behavior extraction on social life issues using sentiment analysis ".

Faiyaz Ahmad has currently working in the department of computer science and engineering as an Assistant Professor in Integral University. He has completed his Ph.D. in "clinical Decision support system using soft computing technique" in 2018 .he has published one book chapter in Springer and one book title name is "cyber low and information security" he has published 18 papers in a reputed journal.