



A Study on Various Approaches for Sentiment Analysis

Priyanka Tyagi*

Assistant Professor, Department of Management, LLDIMS, Affiliated to GGSIPU, New Delhi, India

***Corresponding Author:** Priyanka Tyagi, Assistant Professor, Department of Management, LLDIMS, Affiliated to GGSIPU, New Delhi, India.

Received: November 20, 2020

Published: January 07, 2021

© All rights are reserved by **Priyanka Tyagi**.

Abstract

This article helps you the study of various approaches used in the sentiment analysis on the based of Hybrid approaches, lexicon based approaches and machine learning based approaches..

Keywords: Hybrid; Lexicon; Machine Learning

Hybrid for Sentiment Analysis

M Anagha, *et al.* (2015) introduced a hybrid technique for extracting sentiments from the reviews of Malayalam film [1]. The recommended scheme combined Machine Learning and Fuzzy Logic for sentiment analysis. The recommended approach was used for the tagging of input file. The tags were given by TnT tagger. The recommended approach used some specific rules for handling the exemptions. The sentiments were extracted using Fuzzy membership functions. The tested outcomes demonstrated the efficiency of the recommended approach in decision making regarding a review. It was required to insert sentiment terms for novel fields. The future work would be based on the automatic production of sentiment lexicons. This phenomenon would decrease manual involvement in dictionary preservation.

Amira Shoukry, *et al.* (2015) proposed a hybrid technique for sentiment analysis of Egyptian language Tweets [2]. The recommended scheme combined machine learning and SO (semantic orientation) approaches for this purpose. The achieved tested results showed noteworthy enhancement in different performance parameters. Thus, these results confirmed the efficacy of the recommended scheme in the sentence level sentiment analysis. Some

more efforts would be made in the nearby future for creating a more wide-ranging list of positive and negative sentiment terms. This work would improve the performance of semantic orientation scheme for sentiment analysis.

Ping Han, *et al.* (2016) recommended a hybrid approach for extracting sentiment words. The recommended approach was the combination of lexicon-based and similarity-based technique [3]. The sentiment polarity of small message was obtained using semantic rules and emoticons. Semantic rules based on Chinese language were used in this work. The sentiment based and similarity based techniques were combined together for the extraction of sentiment terms. Also, Emoticons revised the achieved results one more time. The tested outcomes depicted that the recommended scheme efficiently classified dual sentiments. The words from micro blogs were used in this work for validating the accuracy of the recommended scheme. The recommended approach could be generalized and implemented easily on other small texts based on Chinese language.

Farkhund Iqbal, *et al.* (2019) recommended a hybrid system for sentiment analysis using genetic algorithm based feature reduction [4]. The recommended approach integrated lexicon-based

methods and machine learning methods for improving accuracy and scalability in sentiment analysis. The size of feature set could be reduced up to 42% without affecting the accuracy by the recommended hybrid system. The recommended approach was compared with other popular feature reduction approaches such as PCA (Principal Component Analysis) and LSA (Latent Semantic Analysis) in this work for performance analysis. The obtained results depicted that the recommended approach achieved enhanced accuracy rate of 15.4% and 40.2% over PCA and LSA approaches respectively. The future work would be focused on the expansion of this system for cyber-intelligence. This system would assist to generate suggestions for regulatory bodies on the basis of users' views.

Mariam Biltawi, *et al.* (2017) recommended a hybrid approach for classifying the sentiments of Arabic dialect [5]. The recommended approach was the combination of lexicon based and corpus based methods. The main objective of this approach was to present review for CBS (Corpus Based Approach) just like it represented in LBM (Lexicon Based Approach). In order to achieve this aim, the polarity words were replaced with their equivalent tag 'POS' (Positive) or 'NEG' (Negative) in the lexicon. Several classification models were compared in this work for performance evaluation. The tested outcomes depicted that the recommended scheme performed better than CBA. The recommended scheme achieved maximum accuracy rate of 96.34%. However, the research work is still going on additional datasets for expanding lexicons, and improving some classification model using just general words.

A. HariPriya, *et al.* (2018) recommended a hybrid technique to carry out real-time sentiment analysis of highly trending incidents on the basis of location [6]. Several sentiment lexicons, unigram and bi-gram language paradigm with different machine learning algorithms (e.g. Naïve Bayes and Support Vector Machine) were combined together for study in this work. The evaluation of the achieved outcomes was carried out for verifying the performance of the recommended system on the basis of several metrics. The recommended system outperformed the other location based existing systems in terms of accuracy. The future work would be focused on extending the recommended system to multiple dialects with emoticons. Also, more lexicons could be discovered for classifying polarity.

Orestes Appel, *et al.* (2016) recommended a hybrid technique for analyzing sentiments. The recommended approach combined semantic rules, fuzzy sets, unsupervised machine learning algorithms and a sentiment lexicon modified by the Senti-WordNet for sentiment analysis [7]. At first, a hybrid typical classification was performed. This approach was converted into a Hybrid Advanced approach through modification. This approach integrated linguistic classification of semantic polarity. Fuzzy sets were used here to model the polarity of sentiments. The recommended approach was used for computing the polarity of a particular text and a standard openly existing dataset based on film review.

Vani Kansal, *et al.* (2019) recommended a hybrid Approach for financial sentiment analysis using Artificial Intelligence and Cuckoo Search [8]. Initially, the training of ANN (Artificial Neural Network) based sentiment analysis model was done to get positive, negative and neutral sentiments. However, cuckoo search algorithm was recommended in this work due to some limitations of this classifier. Optimizing the fetched attributes with a new fit function was the main aim of this algorithm. The recommended algorithm returned suitable feature sets on the basis of positive and negative sentiments. Every sentiment analysis model required an optimization method due to the occurrence of rough data. However, the recommended system used different pre-processing techniques for improving the data. The future work would involve different types of optimization techniques. In addition, more efforts would be made to deal with sarcastic tweets.

Lexicon based for sentiment analysis

Hunaida Awwad, *et al.* (2017) proposed a hybrid stemming scheme for improving lexicon oriented sentiment analysis (SA) in Arabian language [9]. In order to improve the chances of sentiment text's discovery, the recommended scheme used root-based and light stemmers collectively. The implementation of recommended scheme was carried out on two datasets and four lexicons. The recommended scheme was compared with root-based stemmer and light stemmers independently. The achieved outcomes depicted that the recommended scheme outperformed the other approaches in terms of accuracy when these approaches were used individually.

Annet John., *et al.* (2019) recommended context based sentiment analysis using hybrid lexicon. This work combined a wide-ranging lexicon and domain based lexicon along with other contextual sentiment adjustment rules for improving the sentiment based tweets' classification [10]. The main aim here was to handle the context based polarity of tweet. In such a case, the previous polarity of the word articulated in the lexicon could be differed from the polarity articulated in the tweet. The tested outcomes demonstrated that the recommended approach outperformed the other existing approaches in terms of different performance metrics. This work would be extended in nearby future by integrating more flexible sentiment rules.

Rezvaneh Rezapour., *et al.* (2017) recommended and assessed an improved model for SA (sentiment analysis) [11]. In order to enhance the accurateness of SA, the recommended model integrated informative hashtags into a lexicon. The tweets were collected regarding every contender of president post prior to year 2016 in general election held in USA. In order to detect the polarity of every post, the symbol '# informed' LBA (Lexicon Based Approach) was used. On the basis of these tweets, the ranking was given to the contenders as per their reputation. The tested outcomes depicted that the intuitive hints about a contender's fame particularly on social media could be provided by assigning ranks to contenders on the basis of sentiments expressed in tweets. However, this work had several limitations. Firstly, sarcastic and symbolic texts were very frequent on Twitter. The approaches employed in this work did not explain these effects in specific manner.

Endang Wahyu Pamungkas., *et al.* (2016) performed lexicon-based sentiment analysis of Indonesian language [12]. The classification of sentiment data was carried out in three sentiments such as positive, negative, and neutral. The tested outcomes demonstrated that the recommended approach achieved classification accuracy of 0.68. In general, the recommended approach achieved good results in sentiment classification. This work detected some problems as well. It was possible to use these issues as a base to carry out more research work. Initially, a non-standard language was detected in the dataset. The next issue was related to the dissatisfaction of expectation event. The very last issue was related to vagueness.

Machine learning for sentiment analysis

Mohammed H. Abd El-Jawad., *et al.* (2018) compared several machine learning and deep learning algorithms for performance analysis [13]. Moreover, a novel hybrid model was recommended in this work. This model used text mining and neural networks to classify sentiments. A dataset containing over 1 million tweets gathered from five fields was used in this work. The training and testing of the recommended model was carried out using 75% and 25% part of the dataset. The recommended model outperformed other existing approaches in terms of accuracy (83.7%). The future work would be focused on combining sentiments and text for analyzing sentiments. In addition, a hybrid classification technique would be implemented in nearby future for sentiment analysis of Arabic tweets.

Megha Rathi., *et al.* (2018) classified emotions on the basis of data collected from twitter. The ensemble machine learning algorithms were used in this work for improving classification outcomes in the area of sentiment analysis [14]. In this work, a hybrid approach was recommended for sentiment analysis. This approach combined two classification algorithms called support vector machine and decision tree together. The ensemble algorithms improved the efficacy and trustworthiness of the recommended approach. The tested results depicted that the recommended scheme provided better classification outcomes in terms of different performance parameters.

Pragya Juneja., *et al.* (2017) used several supervised machine learning algorithms for classifying twitter's data into sentiments (positive or negative) [15]. The main aim of this work was to make predictions about the results of Delhi Corporation Elections. In this work, the most efficient machine learning classification model was also identified. This work highlighted the behavior of several classification models on the basis of twitter dataset of different political parties. The tested outcomes revealed that the Multinomial Naïve Bayes classification model outperformed the other classifiers and achieved classification accuracy of 78%. The recommended model could be modified in the nearby future to deal with the texts having different meanings.

Conclusion and Recommendation

The main aim of this work was to make predictions about the results of Delhi Corporation Elections. In this work, the most efficient machine learning classification model was also identified. This work highlighted the behavior of several classification models on the basis of twitter dataset of different political parties. The tested outcomes revealed that the Multinomial Naïve Bayes classification model outperformed the other classifiers and achieved classification accuracy of 78%. The recommended model could be modified in the nearby future to deal with the texts having different meanings.

Bibliography

1. M Anagha., *et al.* "Fuzzy logic based hybrid approach for sentiment analysis of Malayalam movie reviews". IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES) (2015).
2. Amira Shoukry and Ahmed Rafea. "A Hybrid Approach for Sentiment Classification of Egyptian Dialect Tweets". First International Conference on Arabic Computational Linguistics (ACLing) (2015).
3. Ping Han., *et al.* "A Topic-Independent Hybrid Approach for Sentiment Analysis of Chinese Microblog". IEEE 17th International Conference on Information Reuse and Integration (IRI) (2016).
4. Farkhund Iqbal., *et al.* "A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction". IEEE Access 7 (2019).
5. Mariam Biltawi., *et al.* "Arabic Sentiment Classification: A Hybrid Approach". International Conference on New Trends in Computing Sciences (ICTCS) (2017).
6. A Haripriya., *et al.* "Location Based Real-time Sentiment Analysis of Top Trending Event Using Hybrid Approach". International Conference on Advances in Computing, Communications and Informatics (ICACCI) (2018).
7. Orestes Appel., *et al.* "A hybrid approach to sentiment analysis". IEEE Congress on Evolutionary Computation (CEC) (2016).
8. Vani Kansal and Rakesh Kumar. "A Hybrid Approach for Financial Sentiment Analysis Using Artificial Intelligence and Cuckoo Search". 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (2019).
9. Hunaida Awwad and Adil Alpkocak. "Using Hybrid-Stemming Approach to Enhance Lexicon-Based Sentiment Analysis in Arabic". International Conference on New Trends in Computing Sciences (ICTCS) (2017).
10. Annet John., *et al.* "Context Deployed Sentiment Analysis Using Hybrid Lexicon". 1st International Conference on Innovations in Information and Communication Technology (2019).
11. Rezvaneh Rezapour., *et al.* "Identifying the Overlap between Election Result and Candidates' Ranking Based on Hashtag-Enhanced, Lexicon-Based Sentiment Analysis". IEEE 11th International Conference on Semantic Computing (ICSC) (2017).
12. Endang Wahyu Pamungkas., *et al.* "An experimental study of lexicon-based sentiment analysis on Bahasa Indonesia". 6th International Annual Engineering Seminar (InAES) (2016).
13. Mohammed H Abd El-Jawad., *et al.* "Sentiment Analysis of Social Media Networks Using Machine Learning". 14th International Computer Engineering Conference (ICENCO) (2018).
14. Megha Rathi., *et al.* "Sentiment Analysis of Tweets Using Machine Learning Approach". Eleventh International Conference on Contemporary Computing (IC3) (2018).
15. Pragya Juneja and Uma Ojha. "Casting online votes: To predict offline results using sentiment analysis by machine learning classifiers". 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (2017).

Assets from publication with us

- Prompt Acknowledgement after receiving the article
- Thorough Double blinded peer review
- Rapid Publication
- Issue of Publication Certificate
- High visibility of your Published work

Website: www.actascientific.com/

Submit Article: www.actascientific.com/submission.php

Email us: editor@actascientific.com

Contact us: +91 9182824667