ISSN 2278-3091
Volume 13, No.2, March - April 2024
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCS1E/static/pdf/file/ijatcse061322024.pdf
https://doi.org/10.30534/ijatcse/2024/061322024

Sentiment Analysis - An optimized Weighted Horizontal Ensemble approach

George E. Okereke[1], Collin N. Udanor[2], Stephenson C. Echezona[3], Caroline N. Asogwa[4], Izuchukwu Uzoh[5]
[1]Computer Science Department, University of Nigeria, Nsukka, Nigeria, george.okereke@unn.edu.ng
[2]Computer Science Department, University of Nigeria, Nsukka, Nigeria, callings.udanor@unn.edu.ng
[3]Computer Science Department, University of Nigeria, Nsukka, Nigeria, stephenson.echezona@unn.edu.ng
[4]Computer Science Department, University of Nigeria, Nsukka, Nigeria, caroline.asogwa@unn.edu.ng
[5]Computer Science Department, University of Nigeria, Nsukka, Nigeria, izuchukwu.uzoh@unn.edu.ng
Corresponding author: george.okereke@unn.edu.ng

Received Date : February 29, 2024 Accepted Date: March 23, 2024 Published Date: April 06, 2024

ABSTRACT

Sentiment Analysis has gained authority as one of the primary means of analyzing feedbacks and opinion by individuals, organizations and governments. The result of sentiment analysis informs an organization on areas to improve and how best to manage customers. While sentiment analysis may be misleading as no algorithm has been considered 100% efficient, the choice of algorithms can optimize the result based on the dataset in question. This paper aims at studying various algorithms and implementing a weighted horizontal ensemble algorithm as a panacea to low confidence level in the results of sentiment analysis. We designed a system that implements the original Naive Bayes algorithm, Multinomial Naïve Bayes algorithm, Bernoulli Native Bayes algorithm, Logistic Regression algorithm, Linear Support Vector Classifier algorithm and the Stochastic Gradient Descent algorithm. Our dataset was sourced from the Stanford University. It contains fifty thousand (50,000) movie reviews. Dataset from the Nigerian movie review was used to test the models. The reviews were encoded as a sequence of word indices. An accuracy of over 91% was achieved. The Ensemble technique delivered an F1-measure of 90%. Ensemble technique provides a more reliable confidence level on sentiment analysis. The researchers also discovered that change in writing style can affect the performance of sentiment analysis.

Key words: Ensemble Algorithm, Machine Learning, Opinion Mining, Sentiment Analysis

1. INTRODUCTION

Sentiment analysis is of importance and finds active usage in so many fields such as consumer information application, marketing, websites, books, and social media. Sentiment analysis is essential in decision-making for these organizations [1]. Firms that produce for the common person are continually tweaking their products based on reviews from previous products feedbacks. The same is true for the creative arts industry; what kind of movie will trend, what genre of music? In politics, governments use sentiment analysis to gauge the state of a nation and predict election outcomes [2]. Increase in social media penetration creates a potential for private and public organizations to make more accurate predictions based on collected data. Businesses thrive when a good feedback mechanism is in place, this is true given that the probability that an online customer relies on review for a purchase is 0.9; this is because people prefer to take advice from others to invest sensibly [3].

Sentiment analysis is regarded as Opinion mining [4]. In opinion mining, opinion can be represented with a quadruple (s,g,h,t) such that s=> sentiment orientation; g=> sentiment target; h=>opinion holder; t=>time. Sentiment refers to the underlying feeling, emotion and evaluation given to an opinion. The orientation of a sentiment can be positive, neutral or negative. In the study of sentiment analysis, opinion mining is considered an interchangeable synonym [5].

Sentiment analysis is the computational study of emotions and opinions of people as embedded in text. The growth of social media has made available large volume of data for sentiment analysis. Researchers have studied sentiment analysis at three main levels of granularities: document, sentence, and aspect levels. Detecting sentiments in figurative expressions is a challenge [6]. There are mainly three approaches towards sentiment analysis; machine learning based methods, lexicon-based methods and linguistic analysis [7].

Researchers and those implementing sentiment analysis are faced with challenges of choice of algorithm, optimization of chosen algorithm and the possibility of a better algorithm for a given dataset. We encapsulate these in the problem statements below:

1. Few data on the comparative performance of different algorithms
2. Variation of results on the same dataset using different algorithms.
3. Low performance of existing algorithms

The aim of this paper is to ascertain the impact of a multi-algorithm approach on sentiment analysis. We therefore incline to study the performance of individual algorithms in sentiment analysis and analyze the performance of the algorithm on various datasets.

Current technologies rely on amassing as much information as possible given that capable analytical machines now exist. The drive for data has led to unauthorized/stealth access to information using innocuous applications. These data can be used to alter the natural flow of events such as political landscape as seen in [2], purchasing preferences [8], relationship and traveling plans [8], terrorism and stock markets ([9], [10]).

A challenge for most organizations is how to process the large volume of information in the internet space. Organizations edge on the capabilities of sentiment analysis to sieve valuable information from the chunk of data as a means to have better understanding of consumers' opinion and as such, take efficacious actions. Numerous organizations use sentiment analysis in modern day business space to perform department-specific actions ranging from marketing to customer service. Beyond the analysis of product reviews, sentiment analysis is widely used in stock market prediction, disaster management and more**.** Microblogs, blogs in general and other social media platforms such as WhatsApp, have been used to predict elections. Researchers have successfully been able to predict hate speech in sentiments analysis from tweets [11]. Mined data, generate information which is used by organizations to locate new opportunities and enhance marketing. A good example is the United States' use of sentiment analysis to gauge citizens' perception of public policy [12].

1.1. Types of Sentiment Analysis

There are different types of sentiment analysis. Two major types are subjectivity or objectivity identification and feature or aspect-based sentiment analysis. Each one uses different strategies and techniques to pick out the sentiments embodied in a given text [13]. **Subjectivity-Objectivity Identification** is the classification of a sentence of fragments of a piece of text as either subjective or objective. A major challenge with this type of classification analysis is the variation in meaning of a word or phrase due to contextual usage. **Feature-Aspect Identification** on the other hand, permits the determination of varying sentiments or opinion (features) with respect to the different aspects of an entity. Differing from subjectivity-objectivity identification, the feature-aspect based identification gives room for a more nuanced analysis of opinions and feelings [13] [14].

Levels of Sentiment Analysis

Sentiment analysis may be performed at three distinct levels, these are document level sentiment analysis, Sentence level sentiment analysis and Aspect level sentiment analysis [15][16][17][18].

Document level sentiment analysis extracts the sentiment from an entire document and determines its polarity as positive, negative or neutral. In some works, a new line is used as the delimiter for a document level analysis.

Sentence/Phrase level sentiment analysis uses a period punctuation as the delimiter to extract the sentiment in a given data. That is, each sentence has its own defined polarity.

Aspect/Aspect-level sentiment analysis classifies data based on features. This is applicable when the focus is on the subset of the given data. Consider the text "The food is good but the service is bad". If we are studying the quality of food, at sentence level we may give the wrong sentiment.

1.2. Sentiment Classification Techniques

Sentiment Analysis is achievable using Statistics, Natural Language Processing (NLP) and Machine Learning techniques [19][20]. Sentiment classification can be achieved using either Machine learning approach, lexicon-based approach or a hybrid [21][22][23][24]. In the machine learning approach, machine learning algorithms and linguistics features are used to achieve sentiment analysis classification. The lexicon-based approach uses the sentiment lexicon, which is a collection of precompiled sentiment terms. The lexicon-based approach is further divided into dictionary-based approach and corpus-based approach. The corpus-based approach uses statistical or semantic technique to spot the polarity of sentiment. The hybrid of Machine learning and lexicon-based approach is very common. In the hybrid approach, sentiment lexicons play important roles in many of the methods employed. Machine learning approach to test classification can be supervised or unsupervised [23]. The supervised methods make use of labeled training dataset. The unsupervised method is used when training dataset is unavailable.

The lexicon-based approach relies on identifying the opinion lexicon that was used to analyze the text. Two methods available for this approach are dictionary-based approach and corpus-based approach. The dictionary-based approach finds the opinion seed words, then searches a dictionary of synonyms and antonyms. On the other hand, the corpus-based approach is initiated with a seed list of opinion words, then other opinion words in a large corpus are found to help in spotting opinion words with specific context orientations; this is possible using statistical or semantic methods [25].

In this paper, we used a hybrid approach, that is, machine learning and lexicon-based approach to sentiment analysis. Both sentence and document-level sentiment analyses were performed.

Research Gap

Having studied some of the available literature we observed that efforts are continually being made to optimize sentiment analysis performances. Authors in [22] assert that the use of hybrid approach (Machine learning and lexicon-based approach) is more likely to produce better results in sentiment analysis. Ensemble technique has been applied using various algorithms, however, given the literatures assessed, none of them has been implemented using up to five algorithms with a weighted majority voting scheme.

The rest of the paper is organized thus: Section 2 explains how the algorithms were built with different movie review datasets, while section 3 presents the results of the different algorithms from the training and testing datasets, section 4 discusses the results of the experiments when compared with other existing NLP applications like Stanza and other algorithms, and section 5 presents the conclusion and recommendation for further works.

2. METHOD

In this section we present various algorithms by implementing a weighted horizontal ensemble algorithm. We designed a system that implements the original Naive Bayes algorithm, Multinomial Naïve Bayes algorithm, Bernoulli Native Bayes algorithm, Logistic Regression algorithm, Linear Support Vector Classifier algorithm and the Stochastic Gradient Descent algorithm. Our dataset was sourced from the Stanford University. It contains fifty thousand (50,000) movie reviews. Dataset from the Nigerian movie review was used to test the models. The reviews were encoded as a sequence of word indices.

2.1. Design of the Experiment

The system is designed using a simple weighing average to calculate the majority voting. In order to avoid error in computation, none of the established algorithms is tweaked. Six algorithms are trained using the training dataset. The accuracy of each of the algorithm is used as a weighing factor in the ensemble voting.

2.2 Algorithm of the Proposed System

1. Receive text
2. For each algorithm
   2.1 Classify text
3. Perform accuracy-weighed-majority voting
4. Return Classification and Confidence level

Implementation Processes

1. Data Acquisition: The data used in this paper is gotten from Stanford University website, 2011. The choice of data is arising from availability of literature to compare results. Also, the language used in Movie Review is not fast evolving as diction is fairly constant. There is no legal requirement for the data beyond a reference to the source of data. The source of data is anonymised. The data is converted to CSV and cleaned using Spacy and regular expression from Python. We removed emails, html tags, set text to lower case, removed special characters, and removed urls, among others.

2. Term Frequency- Inverse Document Frequency (TFIDF): we utilized the SCIKIT TFIDF to extract features then vectorized our feature set. We set a maximum feature of top 4000 features with a minimum frequency of 100 and a maximum occurrence in 60% of the entire training set. This helps us to eliminate words that have too high frequency but less sentiment.

3. Data Split: The dataset is split using the SCIKIT test_train_split() function at test size of 20% of the dataset.

4. Training: The Model is then trained using Scikit version of the algorithms used in this work. In the case of Linear Support Vector, the model uses Calibrated Classifier in order to utilize the predict_proba() function.

2.3 Sentiment Analysis

Sentiment analysis applications are designed using machine learning algorithms and statistical approaches. Researchers and programmers use one out of the numerous algorithms to implement sentiment analysis solutions. Prior to the use of computers, humans have always used word of mouth, opinion boxes and questionnaires to get feedback. These feedbacks are analyzed by manually reading each one of them and assigning sentiment based on the perception of the reader. The use of computerized approach is subjective and high accuracy is still being sought in the research world.

The Challenge for a new system stems from overt dependence on the output of a single algorithm. It is important that a classified text should have to be indeed certified by majority classifiers. Given that there are multitudes of algorithms, it is important that the output of an algorithm be confirmed by one or more algorithms before a conclusive output is given.

2.4 Analysis of the Proposed Ensemble

The proposed system implements a hybrid ensemble as proposed by [26]. It aims to compare and synergize the output of multiple algorithms in the classification of a given text by utilizing the Gaussian Naïve Bayes Algorithm, the Multinomial Naïve Bayes, the Bernoulli-Naïve Bayes, Logistic regression, Support Vector Machine and Stochastic Gradient Descent algorithms. Majority voting was done to determine the consensual sentiment of the given text. The accuracy of each algorithm is used as a weighing factor in the ratio of Majority Voting. Our dataset was sourced from Stanford University [27]. It contains fifty thousand (50,000) movie reviews. The reviews were encoded as a sequence of word indices. This dataset was created in 2011 by researchers at Stanford University

2.5 Implementation

The algorithms listed in section 2.2 were implemented in this paper using Python 3.8. The algorithms were trained using the movie review dataset [27]. This dataset has been widely used across various research works and therefore, it is easy to compare our results with that of others.

Implementation Architecture

Results from each of the six algorithms were fed to the ensemble. The ensemble calculates the polarity using defined parameters:

Given algorithms A_1 to A_6, with corresponding accuracy C_1 to C_6, it follows that

$$\text{Ensemble} = \frac{A1C1 + A2C2 + \dots + AnCn}{n} \quad (1)$$

$$\text{Ensemble} = \frac{1}{n}\sum_{i=1}^{n} AiCi \quad (2)$$

The ensemble equations are shown in eq 2.1 and 2.2 in which the average of the summation of the ensemble algorithms' accuracies is computed.

Field Use Test:

We tested the algorithm using the partitioned (20%) of the dataset from the IMDB database from Stanford. We also downloaded the latest movie reviews of Squid Game (released September, 2021) and Return of the King of Boys (released August, 2021) from the Nigerian movie review database.

3. RESULTS

The following section shows the results of the ensemble algorithms.

3.1 Multinomial Naïve Bayes

The Multinomial Naïve Bayes report shows an improvement over Gaussian with an accuracy of 0.85. During test, the algorithm correctly polarized a test document where other algorithms failed. We performed secondary test on the algorithm using movie reviews from Return of King of Boys which returned an accuracy of 0.31 and Squid Game with an accuracy of 0.81.

3.2 Logistic Regression

Logistic regression is another algorithm that showed improvement with increase in dataset. When exposed to more dataset, accuracy increased from 88% to 91%, an appreciable increase of 3%. At 91%, we consider Logistic regression to be one of our top three algorithms which is used for stricter sentiment Analysis. When tested with live-data, we got the following result: 0.83 and 0.88 for Return of the King of Boys and Squid Game, respectively.

3.3 Stochastic Gradient Descent

The stochastic gradient descent offers a more promising result with accuracy at 90%, recall at 90% and F1 score at 90%. We however subjected it to some live-data, and the result is as follows: 0.82 and 0.91 for Return of the King of Boys and Squid Game, respectively.

3.4 Support Vector Machine

The Support Vector Machine showed the highest improvement with increase in dataset. From 87%, the accuracy increased to 91.2% when exposed to more dataset. It is the most promising of all the algorithms used in this research. When subjected to live data, we achieved the following results: 0.85 and 0.88 for Return of the King of Boys and Squid Game, respectively.

3.5 Bernoulli Naïve Bayes (BNB)

The Bernoulli report is the most promising of the naïve bayes. It reached an accuracy of 86%, F1 score of 87%, Precision at 87% and recall at 86%. When tested with live-data, we got the following report 0.88 and 0.82 for Return of the King of Boys and Squid Game, respectively.

3.6 Summary of Testing Results

Table 1 shows the summary of how the six algorithms performed on the test sets.

Table 1: Summary of testing results

SN	Algorithm	Precision	Recall	F1-Measure	Accuracy
1	Linear Support Vector	0.91	0.91	0.91	0.91
2	Stochastic Gradient Descent	0.90	0.90	0.90	0.90
3	Logistic Regression	0.91	0.91	0.91	0.91
4	Bernoulli Naïve Bayes	0.87	0.86	0.87	0.86
5	Multinomial Naïve Bayes	0.86	0.85	0.84	0.85
6	Gaussian Naïve Bayes	0.84	0.83	0.83	0.83

Table 1 shows the performance of all the algorithms. The Gaussian Naïve Bayes with an accuracy of 83% did not show appreciable improvement even when the dataset was increased. The TF-IDF vectorizer had the following configuration.

Figure 1 shows a bar chart plot of the performances of all the algorithms on test data. It shows that the logistic regression, linear support vector and the stochastic gradient descent had the best results in that order.

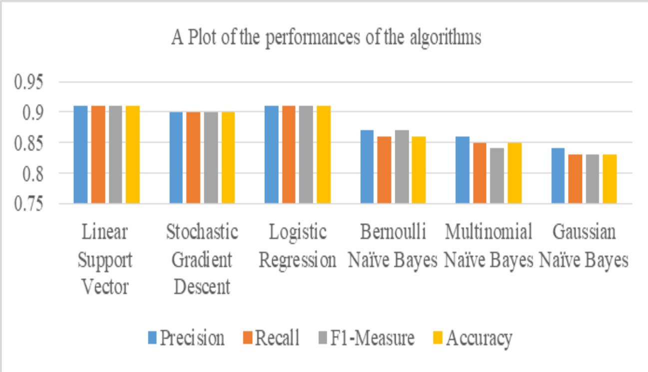

Figure 1: A plot of the performances of all the algorithms on test data

4. DISCUSSION

We have discovered that using ensemble technique increases the accuracy of the classifiers. The ensemble classifier showed an F1 measure of 90% as against 91% being the highest achieved by a single algorithm. The models have been tested with live data from current movie reviews. Stanza, an NLP library released by researchers from Stanford University was used to compare with the system among others. The results are as shown in section 2.

4.1 Comparing Results with Stanza NLP

On Squid Game movie review, we downloaded 70 positive reviews and 70 negative reviews. We used 5-star and 4-star rating for positive reviews, and 1,2-star rating for negative reviews. On Return of the King of Boys, we were able to access 90 reviews from Nigerian audience. A combined confusion matrix is show in table 2a.

Table 2a: Combined Confusion Matrix

	Negative	Positive
Negative	66	10
Positive	21	133

Calculating from table 2a, we get Recall of 0.8636, Precision, **0.9432** and F1 Score of **0.9016**.

Comparing our result with the performance of Stanza sentiment analysis module built with Convoluted Neural Network from Stanford University, we have the following result shown in table 2b

Table 2b: Combined Confusion Matrix

	Negative	Positive
Negative	73	3
Positive	41	113

Calculating from table 2b we get Recall of 0.7334, Precision, **0.9741** and F1 Score of **0.837**

4.2 Summary of Algorithm Performances

1 **Linear Support Vector Classifier (SVM)** appreciates better with increase in size of dataset when compared to other algorithms. It achieved an accuracy rating of 91.2% from 87% when exposure to increased dataset. When exposed to live data, the accuracy fluctuated between 86% and 88%.
Logistic Regression (LOG) came second in accuracy with 90.96%; however, when subjected to live-data, the accuracy of logistic regression fluctuated between 83% and 88% in accuracy.
Stochastic Gradient Descent (SGD) with accuracy of 90.38% gives an impressive performance with 82% as lowest accuracy and 90% as highest accuracy when exposed to live data.
Bernoulli Naïve Bayes (BNB) Model was developed with an accuracy of 86.4% when exposed to live data, the performance fluctuated between 82% for Stanford data and 89% for Nigerian data. Bernoulli is the only top four algorithm to give King of Boys review a higher accuracy than Squid Game.
Multinomial Naïve Bayes (MNB) was implemented at 85% accuracy. When exposed to live data, the accuracy fluctuated between 38% for King of Boys and 81% for Squid Game.
Gaussian Naïve Bayes (GNB) performed least in training at 82% accuracy when implemented. When exposed to live data, the accuracy fluctuated between 75% and 80%.
Given the nature of the live dataset, in particular the review on King of Boys, it is pertinent to infer that MNB does not perform well on sentiment analysis at sentence level, whereas BNB performs best at both sentence and document level classification.
2. We successfully developed an ensemble classifier at two levels, one that considers all the six algorithms and another that considers only the top three algorithms. The ensemble uses the accuracy of each algorithm as its weight in computation as thus:

Given Algorithms A_1 to A_6, with corresponding accuracy C_1 to C_6, it follows that

$$\text{Ensemble} = \frac{A1C1+A2C2+\ldots+AnCn}{n} \quad (3)$$

$$\text{Ensemble} = \frac{1}{n}\sum_{i=1}^{n} AiCi \quad (4)$$

3. We subjected the ensemble to test with three datasets and it showed appreciable performance, reaching an F measure of 90%.

4.3 Comparing with Previous Works

Table 3 shows the results of other research works reviewed.

Table 3: Comparative analysis of reviewed works

Research Works	[28]	[29]	[30]	[31]	[32]	[33]
Techniques/Accuracy	SVM (84%) Naïve Bayes (82%)	Logistic Regression (89%)	Naïve Bayes (81.4 %)	KNN (53%) Random Forest (78%) Naïve Bayes (89%) SVM (4%)	82 %	SVM (86.62%)

Using the F1 measure, accuracy, recall and precision, we can see that our ensemble outperformed most of the previously implemented algorithms.

5. CONCLUSION

This paper investigated how much works had been done in the area of sentiment analysis with particular focus on the machine learning methods used. After reviewing some available literature we observed that though, efforts are continually being made to optimize sentiment analysis, it appears that the use of hybrid approach, that is, machine learning and lexicon-based approach is more likely to produce better results in sentiment analysis. Though ensemble technique had been applied using various algorithms, our investigation showed that none of them has been implemented using up to five algorithms with a weighted majority voting scheme. This motivated our investigation in which we made use of six algorithms on movie review datasets, both from the Stanford IMDB and the Nigerian movie review datasets.

At the end of the investigation, we observed appreciable level of optimization. The demo system developed to showcase the system can be implemented for any organization where sentiment analysis is required. Politicians and business competitors can leverage on this system to assess their performance on the internet scene. Wherever branding is of importance, this system offers a seamless and very accurate means of assessing branding performance. The ensemble technique is best when the algorithms are carefully selected. It is suggested that researchers should test individual algorithms before selecting a pool for ensemble. On the research scale, it is yet to be determined reasons why Nigerian reviews performed poorly under the algorithm.

5.1 Recommendation for Further Works

It has been established that increase in training dataset, increases the performance of Support Vector Machines as observed in this paper. We therefore recommend that further optimization should be carried out, and more dataset provided until the accuracy, precision and recall are well above 90% for all algorithms. We do not recommend the use of Gaussian Naïve Bayes for classification in future works. Bernoulli Naïve Bayes, Support Vector Classifier, Logistic Regression and Stochastic Gradient Descent should be further studied to increase Optimization. Researchers should use an unrelated source of dataset to test algorithms to ascertain true level of optimization achieved as the Nigerian movie review was used to test the Stanford movie review.

REFERENCES

[1] Mohammad EhsanBasiri, ArmanKabiri Words Are Important: Improving Sentiment Analysis in thePersian Language by Lexicon Refining. *ACM Transactions on Asian and Low-Resource Language Information Processing*.2018 doi:https://doi.org/10.1145/3195633

[2] Collins Udanor, Blessing Ogechi Ogbuokiri, Stephen Aneke. Determining social media impact on the politics of developing countries using social network analytics. *Program Electronic Library and Information Systems,* vol.50 no 4, 2016 doi:10.1108/PROG-02-2016-0011

[3] Khan, Muhammad Taimoor; Durrani, Mehr; Armughan Ali, IrumInayat; ShehzadKhalid; Khan, Kamran Habib. Sentiment analysis and the complex natural language.*Complex Adaptive System Modelling,* vol.4 no 2, 2016 doi; 10.1186/s40294-016-0016-9.

[4] Zhang Lei, Liu Bing. Sentiment Analysis and Opinion Mining. 2017 Pp. 31-39

[5] Bo Pang, Lillian Lee. *Opinion mining and sentiment analysis.*2018 url: https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf: Cornell Uviversity. Retrieved from https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf

[6] Mohammed, Saif. Challenges in Sentiment Analysis. In C. Eric, D. Dipankar, B. Sivaji, & F. Antonio (Eds.), *A practical guide to sentiment analysis* 2017 Vol. 5, pp. 61-83; url: https://link.springer.com/chapter/10.1007/978-3-319-55394-8_4). Cham: Springer. Retrieved 10 8, 2019, from https://link.springer.com/chapter/10.1007/978-3-319-55394-8_4

[7] Thakker, Harsh; Dhiren, Patel. *Approaches for sentiment analysis on twitter: A state of art study.*2009 url : https://arxiv.org/ftp/arxiv/papers/1512/1512.01043.pdf. Retrieved October 5, 2019, from https://arxiv.org/ftp/arxiv/papers/1512/1512.01043.pdf

[8] B.H.B.Honan, D.Richer. The Human Advantage: Leveraging the Power of Predictive Analytics to Strategically Optimize Social Campaigns. In *Sentiment Analysis in Social Networks* 2017 doi: https://doi.org/10.1016/B978-0-12-804412-4.00014-0.

[9] Costa, Tammy Da. *Using Sentiment Analysis to Examine Stocks*.2020 url: https://www.dailyfx.com/education/understanding-the-stock-market/stock-market-sentiment-analysis.html: DailyFx. Retrieved February 9, 2021, from https://www.dailyfx.com/education/understanding-the-stock-market/stock-market-sentiment-analysis.html

[10] Geetha, ShriBharathi. Angelina. Sentiment Analysis for Effective Stock Market Prediction.*International Journal of Intelligent Engineering and Systems*, vol.10 no 3, 146-154, 2017 doi: 10.22266/ijies2017.0630.16.

[11] Udanor Collins, and Chinatu Anyanwu**.** Combating the challenges of Social Media Hate Speechin a Polarized Society: A Twitter Ego Lexalytics Approach Emerald Data Technologies and Applications Vol. 53, no. 4, pp. 501-527. 2019

[12] Andrea Ceron, FedraNegri. Public Policy and Social Media: How Sentiment Analysis Can Support Policy-Makers Across the Policy Cycle. *RivistaItaliana di PolitichePubbliche*, pp. 309-338, Dec. 2015 doi: 10.1483/81600.

[13] Sims, Scott. *Sentiment Analysis 101*. 2015 Retrieved February 15, 2020, from https://www.kdnuggets.com/2015/12/sentiment-analysis-101.html

[14] Pavlopoulos, Loannis. *Aspect Based Sentiment Analysis.* Athens: Athens University of Economics and Business 2014

[15] Priyanka Patil, PratibhaYalagi. Sentiment Analysis Levels and Techniques: A Survey. *International Journal of Innovations in Engineering and Technology*, 2016 url: http://ijiet.com/wp-content/uploads/2016/05/72.pdf. Retrieved March 7, 2021, from http://ijiet.com/wp-content/uploads/2016/05/72.pdf

[16] PenubakaBalaji, o. Nagaraju, D Haritha. Levels of Sentiment Analysis and its Challenges: A literature review. *International Conference on Big Data Analytics and Computational Intelligence (ICBDAC).* 2017 doi: 10.1109/ICBDACI.2017.8070879.

[17] Bo Wang, Min Liu. *Deep Learning for Aspect-Based Sentiment Analysis.*StandFord, 2015 url: https://cs224d.stanford.edu/reports/WangBo.pdf: StandFord University. Retrieved January 7, 2021, from https://cs224d.stanford.edu/reports/WangBo.pdf

[18] Kumar Ravi, Vadlamani Ravi. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems, 89*, June 2015 doi: http://dx.doi.org/10.1016/j.knosys.2015.06.015.

[19] Sangharshjit Kamble, A. R. Itkikar.Study of supervised machine learning approachesfor sentiment analysis. Vol. 5 no 4, April 2018 url: https://www.irjet.net/archives/V5/i4/IRJET-V5I4673.pdf. Retrieved February 3, 2020, from https://www.irjet.net/archives/V5/i4/IRJET-V5I4673.pdf

[20] Yulan Hea, DeyuZhoub. Self-Training from Labeled Features for Sentiment Analysis. 2011 url: http://publications.aston.ac.uk/id/eprint/18291/1/Self_training_from_labeled_features_for_sentiment_analysis_2011.pdf. Retrieved February 18, 2011, from http://publications.aston.ac.uk/id/eprint/18291/1/Self_training_from_labeled_features_for_sentiment_analysis_2011.pdf

[21] Rana Alnashwan, Adrian O'Riordan, Humphrey Sorensen, Cathal Hoare. Improving Sentiment Analysis Through Ensemble Learning of Meta-level Features. *2nd International Workshop on Knowledge Discovery on the WEB.* Cagliari, 2016 url: https://cora.ucc.ie/handle/10468/3972. Retrieved from https://cora.ucc.ie/handle/10468/3972

[22] Mohammad Ehsan Basiri, ArmanKabiri. Words Are Important: Improving Sentiment Analysis in thePersian Language by Lexicon Refining. *ACM Transactions on Asian and Low-Resource Language Information Processing*. 2018 doi:https://doi.org/10.1145/3195633

[23] Othman, Rania; Abdelsadek, Youcef; Chelghoum, Kamel; Kacem, Imed; Faiz, Rim. (2019). Improving Sentiment Analysis in Twitter Using Sentiment Specific Word Embeddings. *10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS).* 2019 doi: 10.1109/IDAACS.2019.8924403.

[24] Jeremy Barnes, Erik Velldal, LiljaØvrelid. Improving Sentiment Analysis with Multi-task Learning of Negation.*Journal of Natural Language Engineering special issue on Negation*, 2020 doi: 10.1017/S1351324920000510.

[25] Rupali P. Jondhale, Manisha P. Mali. Study on Distinct Approaches for Sentiment Analysis. *International Journal of Computer Applications*, vol.111 no 17, Feb. 2015 url: https://research.ijcaonline.org/volume111/number17/pxc3901487.pdf. Retrieved February 7, 2020, from https://research.ijcaonline.org/volume111/number17/pxc3901487.pdf

[26] Satuluri Vanaja, Meena Belwal. Aspect-Level Sentiment Analysis on E-Commerce Data.*International Conference on Inventive Research in Computing Applications*, 2018.

[27] StanFord. *aclimdb repository*.2021 Retrieved from Stanford University: http://ai.stanford.edu/~amaas/data/sentiment/aclImdb_v1.tar.gz

[28] Ketan Sarvakar, Urvashi K Kuchara. Sentiment Analysis of movie reviews: A new feature-based sentiment classification. *International Journal of Scientific Research in Computer Science and Engineering*, vol. 6 no3, pp. 8-12. June 2018 Retrieved October 4, 2021, from https://www.isroset.org/pub_paper/IJSRCSE/2-IJSRCSE-0938.pdf

[29] Ankit Goyal, Amey Parulekar. *Sentiment Analysis for Movie Reviews.* Retrieved October 04, 2021,

from https://cseweb.ucsd.edu/classes/wi15/cse255-a/reports/fa15/003.pdf

[30] Palak Baid, ApoorvaGupta ,NeelamChaplot. Sentiment Analysis of Movie Reviews using Machine Learning Techniques.*International Journal of Computer Applications, 179*(7). 2017 Retrieved October 4, 2021, from https://www.researchgate.net/profile/Neelam-Chaplot/publication/321843804_Sentiment_Analysis_of_Movie_Reviews_using_Machine_Learning_Techniques/links/5b6cc3b6299bf14c6d97d82a/Sentiment-Analysis-of-Movie-Reviews-using-Machine-Learning-Techniques.pdf?_sg%5B

[31] Abinash Tripathya, Ankit Agrawalb, Santanu Kumar Rath. Classification of Sentimental Reviews Using Machine Learning Techniques.*3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)*. Procedia Computer Science , Elsevier B.V. pp. 821-829 2015 Retrieved October 4, 2021, from https://pdf.sciencedirectassets.com/280203/1-s2.0-S1877050915X00184/1-s2.0-S1877050915020529/main.pdf?X-Amz-Security-Token=IQoJb3JpZ2luX2VjEMz%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJHMEUCIQDa4H3dWANpLGoLJRyWZuLbOEqJCBsD6dfLw8qHfnev1QIgeDKaphYxMoM0

[32] Gurshobit Singh Brar, Ankit Sharma. Sentiment Analysis of Movie Review Using Supervised Machine Learning Techniques. *International Journal of Applied Engineering Research*, vol. 13 no 16, pp. 12788-12791. 2018 Retrieved from https://www.ripublication.com/ijaer18/ijaerv13n16_53.pdf

[33] Reza Maulana, PannyAgustiaRahayuningsih, WindiIrmayani, DediSaputra, WantyEkaJayanti. Improved Accuracy of Sentiment Analysis Movie Review Using Support Vector Machine Based Information Gain. *ICAISD 2020 Journal of Physics: Conference Series*. 2020 doi:10.1088/1742-6596/1641/1/012060