*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

# Sentiment Analysis and Classification on Amazon Products using Improved Support Vector Machine for Multiclass Classification

**D. Geethanjali[1], Dr. P. Suresh[2]**
[1]Research Scholar, Periyar University, Salem-11
e-mail: ageethaa7@gmail.com
[2]HOD Computer Science, Salem Sowdeswari College, Salem-10
e-mail: sur_bhoo71@rediffmail.com

**Abstract**—There is a huge increase in number of peoples who have been accessing many social networking sites especially user post or reviews for a specific product, company, brand, individual, forums and movies etc. These reviews are helpful in judging customer perception on certain thing. The development of algorithms that could automate the categorization of distinct comments based on feedback from consumers became an analyst project, and this automated classification process is known as sentiment analysis. This research main goal is to analyze Amazon product reviews using an approach to Machine Learning (ML) built around TF-IDF and then employ the Support Vector Machine (SVM) algorithm to categorize the sentiment scores and sentences. SVM can handle binomial classification but the customer reviews is mostly classified into positive, negative and neutral and in some applications, it is fine grained into star ratings such 1-5 or sometimes 1-10. Also, in some applications features or attributes are high in number in which some are irrelevant. Hence, this work applies feature subset algorithm and improves the existing SVM to handle multiclass classification. The Sentiment analysis, Rapidminer tool is considered for classification and the results are visualized, assessed with suitable classification metrics.

**Keywords**- Data mining, Sentiment analysis, TF-IDF, SVM, Wrapper feature subset, Exhaustive code.

## I. INTRODUCTION

The perspectives, ideas, beliefs, and decisions of people are communicated in real time on many platforms in this dynamic environment [1]. Gaining insights from these textual data for efficient decision-making due to the enormous number of knowledge being shared from a broad population may be a herculean effort in the case of sentiment analysis due to the intricacy connected with evaluating the textual information for modelling purposes. The analysis of sentiment, commonly referred to as sentiment mining, examines people's attitudes regarding various objects, including goods, services, groups, people, and problems. There are three main ways to sentiment analysis: supervised and unsupervised machine learning approach; lexicon-based approach [2] [3] with dictionary and corpus-based approach; and hybrid approach using both supervised and unsupervised machine learning approach..

### *A. Levels in Sentiment analysis*

Three levels exist: the document level, the phrase level, and the aspect level.

1) The document level: At this stage, the entire text is classified as good, negative, or impartial depending on its attitude [4]. When the user gives their view on a single entity, document level classification is quite helpful.

Example: "I recently purchased an iPhone. This phone is really amazing. I adore it because of the nice touchscreen and the crystal-clear voice.

2) Sentence level - This sort of review is utilized when an individual only writes a single sentence. It entails two steps: categorizing subjective statements into one of two classes, positive or negative, and categorizing subjective phrases into one of two classes, either positive or negative.

As an illustration, "IPhone revenues have been rising regardless of this bad recession."

The third level is the aspect level, which identifies aspects and estimates the sentiment polarity for each one. Aspect sentiment categorization and extraction are two processes involved in this.

Example: "The meal was excellent, but the staff member who served it was slow." Service and food are aspects.

### *B. Types of Sentiment analysis*

Sentiment analysis [5] concentrates on a text's polarity (positive, negative, or neutral), but it also goes further to identify particular emotions and sentiments, ensuring that you can build and customize your classifications based on the way you wish to comprehend client comments and inquiries.

1) Graded Sentiment Analysis - The evaluation is enlarged to include extremely positive, positive, neutral, negative, and

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

very negative sentiments if polarity precision is necessary. This is known as graded or fine-grained and might be interpreted with star rankings; for instance, very positive would receive 5 stars and very negative would receive 1 star.

2) Sentiment evaluation based on emotion - Identifies sentiments that are accompanied by emotions like fulfillment, frustration, rage, sadness, etc. Numerous sentiment recognition systems make use of sophisticated machine learning techniques or lexicons with lists of phrases and the emotions they denote.

3) Aspect-based sentiment examination - This method is used to determine whether individuals are mentioning specific elements or features in a positive, negative, or neutral way.

### *C. Importance of Sentiment analysis*

- Humans convey their thoughts and sensations, hence emotion evaluation is a crucial technique for comprehending the viewpoints of others.
- Analyzing customers' feedback, to learn what makes customers happy or frustrated can aid in tailoring the products and services to meet their needs.
- Tracking and comparing from one quarter to the next will helpful in taking needed action or verifying the reason of sentiment falling or rising.

### *D. Benefits of sentiment analysis*

- Sentiment analysis can reveal crucial concerns in real-time, allowing one to take the appropriate action at the appropriate moment.
- Sentiment analysis helps organizations process enormous amounts of data in a cost-effective and effective way.
- Centralized sentiment analysis system, companies apply criteria to all of the data, helping them improving accuracy and gain better insights.

## II. LITERATURE SURVEY

Esha Tyagi et al [7] implemented a ML algorithm named SVM to categorize the words and attitudes of product reviews using various datasets. In this study, a variety of datasets were used for training and testing to simulate the SVM method and determine the polarity of positive and negative attitudes. Text preprocessing with tokenization, stop word removal is done in prior and TF-IDF method was implemented to get the score for each token. Finally, with the SVM algorithm higher accuracy 89.98% was obtained.

Faithful Chiagoziem Onwuegbuche et al [8] applied SVM method for sentiment analysis of Twitter data from Nigerian banks over a two-year period. Based on training information for positive and negative tweets, WEKA's LibSVM technique was utilized to create a sentiment categorization model. The preprocessing includes removing symbols, URL, hash tags, whitespaces, punctuations, stop words, stemming. Finally, from the implementation of SVM classification, 71.84% accuracy was obtained.

Imamah, Husni et al [9] proposed Text Mining and SVM for the evaluation of sentiments such as positive and negative reviews provided for the Tourist rankings in Bangkalan Regency. In preprocessing, tokenization, filtering, stemming, tagging were carried out. This work used lexicon to calculate score of the sentiment and labeled as zero and one. Word cloud for positive tweets was displayed and the classification of sentiment is done with linear SVM, the overall results showed the accuracy as 70.22%.

Sachin Chawla et al [10] put forth sentiments analysis using SVM with Maximum entropy on National Education Policy tweets. Initially, feature selection with maximum entropy was applied to get most relevant attribute and then with the subset SVM algorithm was implemented to classify positive and negative tweets. 300 tweets with 150 positive and 150 negative tweets were taken from analysis. The parameters lamda variable, gamma constant, epsilon, maximum iteration, and complexity (C) was fixed and 77.47 % accuracy was obtained in the results.

Suman Rani et al [11] dealt with mining sentiments of tweets about Indian politicians using SVM. Unigram and TF-IDF are used as feature extractors for this model and the performance are measured. Data was collected from twitter using twitter API. Removal of URL, stop words, case transformation was done in preprocessing and the sentence is broken into tokens. Unigram with TF-IDF was used to extract features. Two types of SVM was implemented namely linear SVM and Kernel SVM to classify the positive and negative tweets. Linear SVM gave high accuracy as 93 %.

Munir Ahmad et al [12] used SVM classifier for evaluating the sentiments. Two pre classified datasets of tweets are used and for comparative analysis. The performance was analyzed by applying SVM and the polarity detection with positive, negative and neutral of sentiments are classified. First tweet dataset named self-driving cars has 5 class and the second dataset named Apple has 3 classes. Pre-processing includes TF-IDF, Stemmer, stopwords Manager and tokenizer. From the implemented results of SVM on both datasets, the accuracy obtained was 59.91 % for first one and 71.2 % for the second one.

Syahputra et al [13] analyzed public opinion sentiment towards online stores in Indonesia via Twitter using the SVM algorithm. Twitter is classified into three classes such as positive, negative and neutral. Pre-processing of text data was carried out with the case folding, the data cleaning, normalization, removal of stop-words, the stemming, and the data tokenization. The best accuracy was obtained by sigmoid kernel SVM with 82% accuracy on the Shopee online store dataset, 94.7% on the Tokopedia dataset and 75.3% for the Bukalapak dataset.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

*Limitations in existing literatures*

- SVM is used for binomial classifications for positive and negative reviews as SVM cannot be applied for multiclass classification. Some work applied One Vs Rest (1 vs Rest) for trinomial class but the work obtained variant accuracy for the applied datasets.
- There is inadequate accuracy in all the discussed existing literatures.
- Processing time is not given for the implemented methods.

## III. PROPOSED METHODOLOGY

The TF-IDF method is proposed to create and develop the word vector. Analysis of sentiment is done using N-grams along with Tri-Gram. For each feeling, the tool Bag of Words (BoW) creates a word cloud to represent the frequency of words used. Initially, wrapper based feature subset is applied along with SVM and Exhaustive code correction in Error Correction Output Codes (ECOC) is combined to handle multiclass classification.

### *A. Term Frequency- Inverse Document Frequency (TF-IDF)*

The significance of a word is to a document in a corpus is indicated by statistical data. A common weighting method is the TF-IDF in Eq. (1), which grows in values of a direct relationship to the frequency that a word occurs across the entire text.

The formula for TF-IDF is given in equation 1,

$$TF - IDF(t,d,D) = tf(t,d) \,.\, idf\,(t,D) \qquad (1)$$

The ratio of the frequency of phrase t within record d is called term frequency, or tf (t, d).

The tf (t, d) in Eq.(1) is calculated and given in equation 2,

$$tf\,(t,d) = \frac{f_{t,d}}{\Sigma'_{t' \in d} f_{t',d}} \qquad (2)$$

Where, ft, d – count of a term in a document.

Equation 1's provides Inverse Document Frequency - idf (t, D) is a measurement of the quantity of data a single phrase contains. The idf (t, D) is calculated and given in equation 3,

$$idf\,(t,D) = log\ log\ \frac{N}{|\{d \epsilon D : t \epsilon d\}|} \qquad (3)$$

Where, N represented for the total number of documents in a corpus (N=|D|) and {d_D:t_d}| - number of documents in which the term t appears..

### *B. Text Pre-processing*

Tokenization is the process of breaking up a document into sentences or words by removing blank spaces, spelling and grammar, question marks, full stops, and other symbols.

Stopword removal - Pronouns and articles are typically categorized as stop words.

Reduce a term to its root or basis by stemming.

A procedure known as "n-Grams" breaks down a sentence into its component words. If n is 1, only one word is retained, and n is 2, a word with its recurrent word succession is removed. Any number may be analyzed using trigrams, and n is relevant to all numbers.

Tokens can be filtered based on length, which is fixed at 4 to 25 letters in order to prevent the use of very short words that don't adequately convey information.

### *C. SVM-WFSEC*

- Wrapper Based Feature Set

The selection of features is an essential step in any machine learning process since some parameters may be irrelevant or of less importance than others when trying to extract insights from high dimensional data. A specific ML method serves as the foundation for the wrapper-based decision-making procedure.

From Figure 1, the method starts with the whole feature/variables initially. Then the wrapper based feature subset is applied which is combined with machine learning algorithm. In each iteration, performance is assessed with a subset of features and finally the subset is selected with high performance.

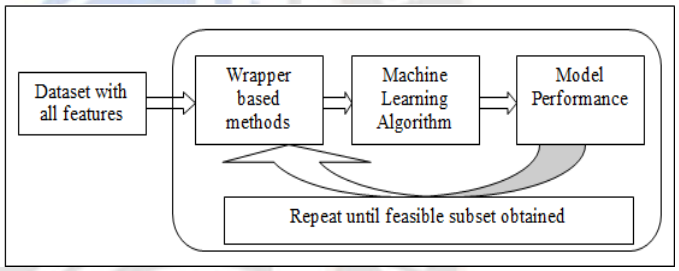


Figure 1. Flow of Wrapper based feature subset method

- Wrapper based feature subset method- Backward elimination

If the dataset has several features, [14] it is possible that not all features are equally important hence, to improve performance ignore some few features using feature selection methods without losing the accuracy. The entire collection of attributes is used when the backward reduction operator first begins. It eliminates each remaining attribute in each round, then estimates performance using the inner operators, such as a cross-validation with classifier. Finally, just the attribute with the smallest effectiveness degradation has been eliminated from the evaluation. After that, the changed selection is used to begin a new round.

- Support Vector machine with Exhaustive code

SVM training algorithm builds a model with the training examples, each categorized into one of two groups (-1/+1 or yes/no). New examples will be assigned to one of the category thus making it a non-probabilistic binary linear classifier. It maximizes the width of the gap between the two categories and the new examples are then mapped into that same space, predicted based on which side of the gap the example will

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

belong to. SVMs may effectively execute non-linear categorization using a kernel-based method in addition to this linear categorization by mapping their inputs into high-dimensional feature spaces [15].

For a given training dataset of n points in the form (x1, y1),...(xn, yn), any hyper plane can be written as,

$$w^T x - b = 0 \quad (4)$$

Where, w is the normal vector to the hyperplane, the algorithm works with the intention of the distance between the hyperplane and the nearest point from either category is maximized.

To avoid the data being falling into the margin the following constraint is followed,

$$w^T x_i - b \geq 1, if\ y_i = 1$$
$$w^T x_i - b \leq 1, if\ y_i = -1 \quad (5)$$

Kernel function - Radial basis kernel (RBF Kernel)

Kernel function is used to maximize hyperplane margin by replacing the dot product in normal SVM with kernel function that allows the classifier to fit the maximum-margin hyperplane in a transformed feature space. Some kernel functions include Gaussian, sigmoid radial basis, dot, polynomial, neural,[14] anova, multiquadratic. This work applies radial basis kernel function. The kernel function for two points $Y_1$ and $Y_2$ computes the similarity. This kernel can be mathematically represented as follows,

$$K(Y_1, Y_2) = exp\ exp\ (- \frac{||Y_1 - Y_2||^2}{2\sigma^2} \quad (6)$$

where,

'σ' is the variance and our hyper-parameter and $||Y_1 - Y_2||$ is the Euclidean Distance between two points $Y_1$ and $Y_2$.

When the points are the same, there is no distance between them.

The kernel value is less than 1 and near to 0, indicating that the points are different, whenever the points are far apart.

Error Correcting Output Codes (ECOC) - Exhaustive code

An ensemble method designed for multi-class classification problem, the task is to decide one label from C > 2 potential options. Exhaustive codes are used when there are fewer than seven classes C (3 C). The length of each code is 2C-1-1. Only one is present in row 1.

TABLE I. Exhaustive code for four-class problem

| Row | Column | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| 3 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |

Row 2 is made up of 2C-2 zeros and 2C-2-1 ones. 2C-3 zeros are found in row 3, which are followed by 2C-3 ones, 2C-3 zeros, and 2C-3-1 ones. If 10 number of classes is very large, random codes can be used instead of exhaustive codes. Thus, the multiclass classification is transformed to multi binary class problem that aids in high accuracy.

Example: Exhaustive code for four-class problem

In Table 1, one, zero are the binary class values for the classes 4 hence the multi class is transformed to multi binary class and the classifier is executed for each class and each inter row has Hamming distance value 4.

*Algorithm of proposed method SVM-WFSEC*

Input Amazon Dataset with 21 features

Output Classified data with sentiments, Word cloud for each sentiments

Step 1 Stratified sampling with the range 500,1000, 1500 2000, 2500

Step 2 Change five classes in the dataset to three classes positive, negative, and neutral using if else expression.

Step 3 Process Documents-Text Preprocessing

Step 3.1 Tokenization with non-letters.

Step 3.2 Filter tokens by length

Step 3.3 Filter Stopwords

Step 3.4 Stemming

Step 3.5 Transform cases

Step 3.6 Generate n-grams (trigram-n=3)

Step 4 Wrapper based feature subset with backward elimination and classifier

Step 4.1 Multi class classification with binomial procedure using exhaustive code

Step 4.2 SVM classifier with RBF kernel

Step 4.3 Performance assessment

Step 4.4 Return the optimal attributes with high accuracy.

- *Advantages of the proposed method SVM-WFSEC*

The text preprocessing combination methods aid in getting informative tokens.

Wrapper based feature selection method gives relevant features for the analysis which then aids in high accuracy.

SVM with exhaustive code transforms the multiclass problem to multi binary-class problem make SVM to handle multiclass problem[16].

## IV. RESULTS ANALYSIS

The Amazon evaluation dataset, which has twenty-one features, was collected through the Kaggle website [15]. The performance is assessed with range of data (500, 1000, 1500, 2000, 2500) using stratified sampling. After the evaluation of wrapper based feature subset with SVM classifier 18 relevant attributes are retained.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

### *A. Word Cloud for Different Polarities*

A graphic illustration of the word occurrence of the more prevalent keyword in the text under analysis is considered. Generally the most frequent term is generated in image form


Figure 2. Word Cloud for positive polarity


Figure 3. Word Cloud for negative polarity

Figure 2, shows the top most terms used by the user for positive polarity, Figure 3, represent the negative polarity and Figure 4, presents the terms regarding neutral polarity. It provides provide quick insights at a glance for three polarities separately.


Figure 4. Word Cloud for neutral polarity

### *B. Performance Analysis*

Three evaluation measures namely accuracy, RMSE, processing time are used for this analysis.

1) Accuracy

$$Accuracy = \frac{Correctly\ classified\ instances}{N} \quad (7)$$

2) Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Predicted_i - Actual_i)^2}{N}} \quad (8)$$

3) Time taken for Processing – The duration required to develop and verify the model.

Table 2, list out the performance values in terms of accuracy, RMSE for the existing SVM - 1 vs Rest and proposed SVM-WFSEC for three class sentiments such as positive, negative and neutral. Five data range are taken for granularity-based Trigram (n=3) analysis.

TABLE II. Accuracy, RMSE analysis of SVM - 1 vs Rest, SVM-WFSEC

| Accuracy in percentage | | | RMSE in range (0-1) | |
|---|---|---|---|---|
| Range | SVM -1 vs Rest | SVM-WFSEC | SVM -1 vs Rest | SVM-WFSEC |
| 500 | 84.15 | 89.37 | 0.125 | 0.096 |
| 1000 | 86.23 | 91.02 | 0.096 | 0.073 |
| 1500 | 87.13 | 92.01 | 0.109 | 0.068 |
| 2000 | 88.72 | 93.14 | 0.092 | 0.054 |
| 2500 | 90.82 | 93.15 | 0.082 | 0.052 |

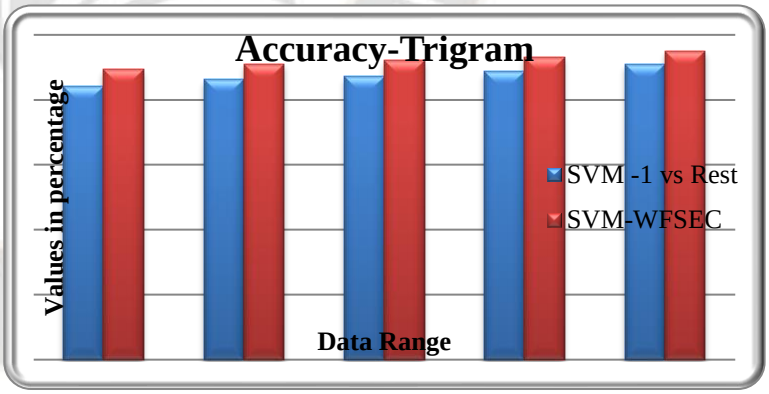


Figure 5. Accuracy analysis of SVM - 1 vs Rest, SVM-WFSEC

From Figure 5, it is proved that for a perfect analysis high range of data is needed. The highest accuracy is around 93.15 % which is obtained for the range 2000 and 2500. Also, the proposed method scores highest accuracy for all range of data.

Table 3, list out the performance values in terms of processing time for the existing SVM - 1 vs Rest and proposed SVM-WFSEC for three class sentiments such as positive, negative and neutral. Five data range are taken for granularity-based Trigram (n=3) analysis.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

TABLE III. Processing Time analysis of SVM - 1 vs Rest, SVM-WFSEC.

| Processing time in minutes | | |
|---|---|---|
| Range | SVM -1 vs Rest | SVM-WFSEC |
| 500 | 0.79 | 0.68 |
| 1000 | 1.98 | 1.10 |
| 1500 | 2.92 | 2.19 |
| 2000 | 5.23 | 4.41 |
| 2500 | 11.01 | 10.12 |

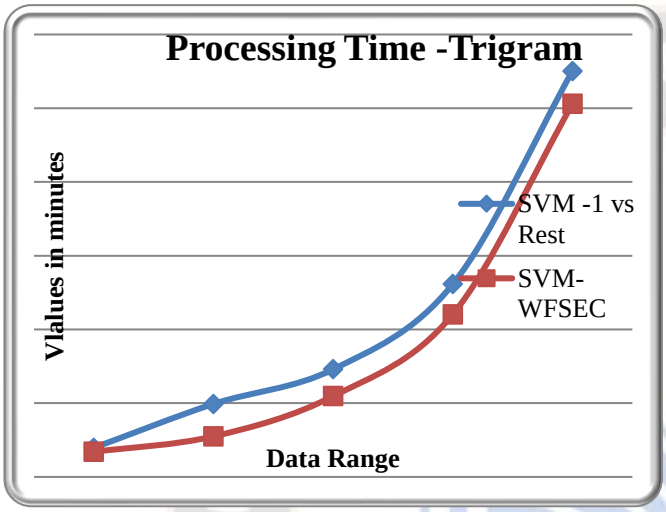


Figure 7. Processing Time analysis of SVM - 1 vs Rest, SVM-WFSEC

From Figure 7, it is shown the proposed method SVM-WFSEC takes less time than SVM-1 vs Rest since the proposed method has feature subset evaluation with Error correcting output codes.

## V. CONCLUSION AND FUTURE WORK

Sentiment analysis helps businesses to monitor brand and product sentiment from customers' feedback, and understand customer needs to improve their sales and services. This work concentrates on Amazon product reviews which has five star rating. The proposed work initially changes the five stars rating into three sentiment polarities as the analysis part is accurately developed with such polarities. SVM-WFSEC incorporates wrapper based feature selection, text processing, trigram analysis, exhaustive correction code for multiclass classification and so the method outperforms the existing SVM-1 vs Rest in terms of accuracy with 93.15 for the data range 2500, RMSE with 0.52 for the same range and less processing time.

In future, the work can be extended using Part-of-speech (POS) tagging instead of n-gram analysis. Variant application may be tried using this proposed method. Other multi-class classification methods one vs one, random access code can be implemented.

## REFERENCES

[1] K. H. Manguri, R. N. Ramadhan, R.N. and P.R.M Amin, "Twitter sentiment analysis on worldwide COVID-19 outbreaks," Kurdistan Journal of Applied Research, 2020, pp.54-65.

[2] K. Ravi, and V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," Knowledge-based systems, vol. 89, 2015, pp.14-46.

[3] M. Al-Ayyoub, S. B. Essa, and I. Alsmadi, "Lexicon-based sentiment analysis of Arabic tweets," International Journal of Social Network Mining, vol. 2, no. 2, pp.101-114, 2015.

[4] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams engineering journal, vol. 5, 2014, pp.1093-1113.

[5] B. M. Jadav, and V. B. Vaghela, "Sentiment analysis using support vector machine based on feature selection and semantic analysis," International Journal of Computer Applications, vol. 146, 2016.

[6] https://monkeylearn.com/sentiment-analysis/

[7] E. Tyagi, and A. K. Sharma, "Sentiment analysis of product reviews using support vector machine learning algorithm," Indian Journal of Science and Technology, vol. 10, 2017, pp.1-9.

[8] F. C. Onwuegbuche, J. M. Wafula, and J. K. Mung'atu, "Support vector machine for sentiment analysis of Nigerian banks financial tweets," Journal of Data Analysis and Information Processing, vol. 7, 2019, p.153.

[9] Assef Raad Hmeed, Jamal A. Hammad, Ahmed J. Obaid. (2023). Enhanced Quality of Service (QoS) for MANET Routing Protocol Using a Distance Routing Effect Algorithm for Mobility (DREAM). International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 409–417. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2687

[10] E. M. Rachman, I. O. Suzanti, I.O. F. A. Mufarroha, "Text Mining and Support Vector Machine for Sentiment Analysis of Tourist Reviews in Bangkalan Regency," In Journal of Physics: Conference Series, vol. 1477, no. 2, pp. 022-023, March 2020, IOP Publishing.

[11] M. Cindo, D. P. Rini, and E. Ermatita, "Sentiment analysis on Twitter by using Maximum Entropy and Support Vector Machine method," Sinergi, vol. 24, pp.87-94, 2020.

[12] S. Rani, and J. Singh, "Sentiment analysis of Tweets using support vector machine," Int. J. Comput. Sci. Mob. Appl, vol. 5, 2017, pp.83-91.

[13] AGYEI , I. T. . (2021). Simulating HRM Technology Operations in Contemporary Retailing . International Journal of New Practices in Management and Engineering, 10(02), 10–14. https://doi.org/10.17762/ijnpme.v10i02.132

[14] M. Ahmad, S. Aftab, and I. Ali, "Sentiment analysis of tweets using svm," Int. J. Comput. Appl, vol. 177, 2017, pp.25-29.

[15] H. Syahputra, "Sentiment analysis of community opinion on online store in Indonesia on twitter using support vector machine algorithm (SVM)," In Journal of Physics: Conference Series vol. 1819, no. 1, pp. 012-030, Mar 2021, IOP Publishing..

[16] K. Gurumoorthy, and P. Suresh, "Supervised Machine Learning algorithm using Sentiment Analysis based on Customer Feedback for Smartphone Product," International Journal of Emerging Trends in Engineering Research, vol. 8, no. 8, pp.1-9, Aug 2020.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 10s*
*DOI: https://doi.org/10.17762/ijritcc.v11i10s.7619*
***Article Received**: 01 June 2023 **Revised**: 25 July 2023 **Accepted**: 09 August 2023*

[17] https://www.analyticsvidhya.com/blog/2020/10/a-comprehensive-guide-to-feature-selection-using-wrapper-methods-in-python/.

[18] Christopher Davies, Matthew Martinez, Catalina Fernández, Ana Flores, Anders Pedersen. Using Machine Learning for Early Detection of Learning Disabilities. Kuwait Journal of Machine Learning, 2(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/172

[19] S. Muthukumaran, and P. Suresh, "A Unified Framework of Sentimental Analysis for Online Product Reviews Using Enhanced Ant Colony Optimization Algorithm," International Journal of Pure and Applied Mathematics (IJPAM), vol. 119, no. 14, pp 489-496, 2018.

[20] https://www.kaggle.com/datasets/yasserh/amazon-product-reviews-dataset.