Sentiment Analysis for IMDb Movie Review Using Support Vector Machine (SVM) Method

Fidya Farasalsabila[1], Verra Budhi Lestari[2], D. Diffran Nur Cahyo[3], Hanafi[4], Tutik Lestari[5], Fahmi Rusdi Al Islami[6], M. Akbar Maulana[7]

[1,2,3]*Magister Teknik Informatika, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia*
[4]*Informatics Department, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia*
[5,6,7]*Informatics Department, Institut Teknologi Tangerang Selatan, Tangerang Selatan, Indonesia*
[1,2,3][fidya, verrabl, diffran]@students.amikom.ac.id
[4]hanafi@amikom.ac.id(*)
[5,6,7][tutik.lestari, fahmi.rusdi, maulana]@itts.ac.id

Received: 2022-12-29; Accepted: 2023-01-26; Published: 2023-03-18

Abstract**—** Many researchers currently employ supervised, machine learning methods to study sentiment analysis. Analysis can be done on movie reviews, Twitter reviews, online product reviews, blogs, discussion forums, Myspace comments, and social networks. Support Vector Machines (SVM) classifiers are used to analyze the Twitter data set using different parameters. The analysis and discussion were undertaken to allow for the conclusion that SVM has been successfully implemented utilizing the IMDb data for this study (Support Vector Machine). To complete this study, the preprocessing phase, which consisted of filtering and classifying data using SVM with a total of 50.000 data points, was completed after collecting up to 40.000 reviews to use as training data and 10.000 reviews to use as testing data. 25.000 positive and 25.000 negative points make up the view. In this study, we adopted an evaluation matrix including accurate, precision, recall, and F1-score. According to the experiment report, our model achieved SVM with Bags of Word (BoW) used to get results for the highest accuracy test, which was 88,59% accurate. Then, using grid-search, optimize against the SVM parameters to find the best parameters that SVM models can use. Our model achieved Term Frequency–inverse Document Frequency (TF-IDF) was used to get results for the highest accuracy test, which was 91,27% accurate.

Keywords— Sentiment Analysis, IMDb, Movie Review, TF-IDF, SVM

I. INTRODUCTION

A movie is a leisurely activity. Many movies are now available to watch on the Internet or in theatres. IMDb is a popular website for viewing movie reviews today. There are several different movie comments on the IMDb website. Based on the star rating feature, the comments are visible. As a result, users find it challenging to interpret other users' comments. As a result, this research was conducted to discover remarks based on the IMDb site's star rating feature [1].

Several previous studies have been conducted on sentiment analysis. The majority of these studies treat Sentiment Analysis as a classification task (for example, Support Vector Machine (SVM [1], Naive Bayes (NB) [2], the impact of bias on ML [3], and so on). In this regard, recent work has shown promising results by improving the performance of this algorithm. The first approach proposes a new model based on BERT and deep learning for text sentiment analysis [4].

The paper presents a better way to introduce Internet content to customers using a referral system. The recommendation system calculates product recommendations by detecting past user behavior. Past user behavior may alter to determine the degree to which numerous consumers are similar. Document terminology is one of the main user behaviors. Most document interpretation models in the recommender system use traditional NLP models like TF-IDF and LDA models. From the perspective of NLP, contextual knowledge is a drawback of classical NLP. The researcher has created a novel model for producing contextual awareness to address the issue mentioned above by incorporating two crucial factors: a delicate word and a word sequence. The researcher used RNN-LSTM to implement GLOVE-based word integration and sequential word detection. The contextual preview of the IMDb movie document was successfully captured by our model, according to the qualitative review report [5].

The paper presents this study using Structural Equation Modeling (SEM) with a 5-item Likert rating. It includes chi-squared, probability level, CMIN/df, CFI, RMSEA, TLI, GFI, and AGFI as the model's qualifying index. The researcher found that complete information and information-rich websites lead to customer satisfaction, which reflects the quality of e-services on e-commerce websites. Better online service quality will impact customer satisfaction when frequently using e-commerce sites of their choice. Also, through customer perceived value, the caliber of online services impacted e-commerce customer satisfaction and loyalty. It increased knowledge for those working in e-Commerce to implement high-quality e-Services to win over customers [6].

Recent work has shown promising results in this regard by improving the performance of this algorithm. The predictive performance of five models was compared to the SVM, RNN, LSTM, BERT, and KoBERT [7]. The findings confirmed that KoBERT outperformed all predictive performance indicators (71%) (Accuracy, precision, and F1 score). It is suggested that a feature selection mechanism that gives more context to smaller features than a frequency can outperform some traditional selection methods (such as term-frequency, Chi-squared, etc.). This study uses two feature extraction methods, TF-IDF and Bags of Word (BOW) Vectorized, to increase the performance of SVM [8].

II. RESEARCH METHODOLOGY

DOI : https://doi.org/10.25139/inform.v8i2.5700

The model proposed in this study is Support Vector Machine. Several steps need to be taken before implementing this model into research. The review text is then prepared for the trained model.

Check the working model of this research's outcome last. The technique used for this research project will be applied to the discussion in this section.

A. Proposed Model

The model proposed in this study is Support Vector Machine. Several steps need to be taken before implementing this model into research.

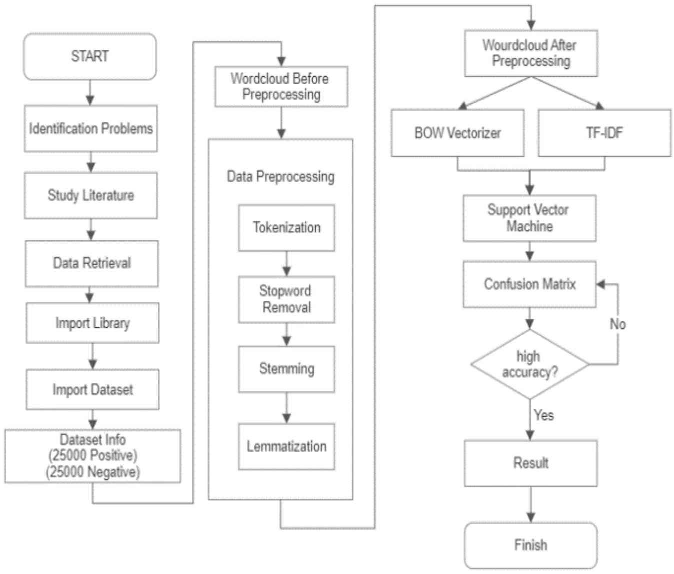

Figure 1. Research Flow

B. Sentiment Analysis

Sentiment analysis, often known as opinion mining, examines how individuals feel about a particular thing or quality as they express themselves in written text. The questioned entity could be goods, services, businesses, people, occasions, problems, or subjects. Numerous terms that refer to the same concept but slightly distinct tasks include sentiment analysis, subjectivity analysis, affect analysis, emotion analysis, and review mining. All of these terms fall under the general heading of sentiment analysis [9]

Sentiment analysis has been one of the most active study fields in natural language processing since the early 2000s. The goal of sentiment analysis is to define an automated tool capable of extracting sentiments and other subjective information from texts and natural language, thereby producing organized and useful knowledge that either decision support systems or decision-making can use. Researchers disagree on whether the discipline should be dubbed sentiment analysis or opinion mining due to uncertainty about the distinction between sentiment and opinion. In contrast to opinion, defined as a view, judgment, or judgment made in mind on a specific issue, sentiment is described in the Merriam-Webster Collegiate Dictionary as an attitude, thinking, or judgment motivated by feelings. The differences are quite slight, and both share several characteristics. The definition demonstrates that while sentiment is more of a feeling than an opinion, opinion is more of a person's concrete perception of something [10].

C. IMDb

IMDb (Internet Movie Database) [11] is an online database of information about movies, TV shows, home videos, video games, and shows on the Internet, including cast lists and biographies of the production team. The Output and staff, story summaries, quizzes, reviews, and ratings by fans. Another fan feature, the message board, was discontinued in February 2017. Fans initially ran the site, but the database was later owned and operated. by imdb.com Inc., a subsidiary of imdb.com Inc. Amazon.

The majority of the database's information comes from volunteers. Registered users of this website can edit and add new documents. Those with a track record of providing accurate data will be given fast clearance for additions or adjustments to staff demographics, actors, awards, and other media productions. However, changes to images, names, character names, plot synopsis, and titles are meant to be read before publication and typically take 24-72 hours to appear [12].

D. Dataset

The material used in this research is the IMDb (Internet Movie Database) dataset obtained from kaggle.com. One thousand data were used in this study. The dataset used will be divided into two types of data for this study's purposes: training data and test data. Training data is used as much as 80% to train the algorithm in finding the appropriate model. While the Test Data used as much as 20% of new data that does not yet have a class, a classification process is needed to determine the appropriate class. The data obtained includes the class (label), namely positive and negative [11].

E. Preprocessing

Tokenization, Stop Word Removal, Stemming, and Lemmatization are examples of fundamental text processing procedures [13].

1) Tokenization: Tokenization is recognizing words in the character input sequence, primarily by separating punctuation marks and by recognition, abbreviations, etc. In some cases, if the content was taken from a web page, the tokenization process additionally includes procedures for normalizing the text, such as removing HTML tags, true casing, or lowercase lettering.

2) Stop Word Removal: Stop words, which also go by the names of function words or closed-class words, are high-frequency words like subjects, affixes, pronouns, determinants, prepositions, conjunctions, and others. Words that are not crucial to the text document will be eliminated throughout this process.

3) Stemming: Although many words in natural language share similarities, their various forms render recognition

DOI : https://doi.org/10.25139/inform.v8i2.5700

useless. In order to retrieve the parent (root) of a phrase, which all associated words will share, stemming is the act of eliminating suffixes and prefixes from an input word. For instance, computations and computations will all be derived from the same root. When the "consumer" of a root word is a system rather than a human person, stemming frequently produces root words that are invalid or irrelevant.

4) Lemmatization: Lemmatization, an alternative to stemming, lowers a word's inflectional form to its root form. In contrast to stemming, lemmatization produces legitimate word forms, the earliest forms of words found in dictionaries. Lemmatization, therefore, has the advantage of making the output human-readable. Still, it necessitates a more computationally demanding procedure because it needs a list of grammatical forms to handle regular inflection and a lengthy list of irregular words.

F. Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF [14] is a combination of Term Frequency (TF) and Inverse Document Frequency (IDF). Because it directly uses the original word frequency values from the document, the TF representation is one of the most straightforward TWSs. The TF is predicated on the idea that terms with higher phrase frequency values are regarded as having greater importance than those with lower term frequency values. It depends on how frequently particular terms appear in the local document. Due to its ignorance of collection frequency, TF's ability to separate all pertinent documents from other irrelevant ones is very poor. Inverse document frequency (IDF) in terms of coverage frequency has been presented as a solution to this problem [15]. This enhances the terms' capacity to be distinguished for text classification. IDF goes beyond document frequency (DF) to include the number of documents that contain the phrase. The idea behind this is that terms that appear in fewer papers are thought to be more significant than terms that feature in more texts [16]. The IDF value for a particular term can be determined using Equations (1), (2), and (3).

$$IDF\ (t,d,D)\ = log\frac{|D|}{DF(t,D)} \tag{1}$$

$$IDF\ (t,d,D)\ = log\frac{|D|+1}{DF(t,D)+1} \tag{2}$$

$$\text{TF-}IDF\ (t,d,D)\ = TF(t,d) * IDF(t,d,D) \tag{3}$$

G. Bag of Words (BoW)

The bag of words (BoW) model, commonly referred to as a vector space model, is a straightforward representation used in natural language processing (NLP) and "IR" information retrieval [17]. This approach views a sentence or text as a collection of several sets of words, irrespective of word order and grammar, but keeping polysemy. A model that learns vocabulary from all papers and then models each document by counting the number of times each word appears is another way to define BoW [18]

H. Support Vector Machine

The Support Vector Machine (SVM) algorithm [19] is a model of binary classification. It is the best-fit segmentation between two data classes since it is a straight line in two dimensions. An ideal decision plan must be established for high-dimensional data sets as the segmentation reference type. According to SVM's fundamental principle, when a classification problem is solved, the distance between the nearest sample point and the decision surface must be at its maximum, i.e., the distance between two classes of sample points to effectively separate the samples [20].

The classification function f(x) = wTx + b represents it. The support vector on the dotted line r is defined as the geometric distance, and it is equal to the distance between the two dotted lines r, or the interval of two dotted lines to 2r using Equation (4).

$$r\ =\ yr\ =\ \frac{\check{r}}{w} \tag{4}$$

Where $\hat{r}$ = *y(wTx + b) = yf(x)*, the $\hat{r}$ variable is a function interval.

I. Hyperparameter Tuning

A hyperparameter is a parameter in a machine-learning context set before the learning process begins. Hyperparameter tuning is an architecture from deep learning to improve the performance of predictive models. Training the data yields the values of model parameters. The model parameters are the weights and coefficients the algorithm derives from the data. Every algorithm has a set of hyperparameters, such as a decision tree depth parameter [21]. Randomly selecting hyperparameters can never ensure a consistent and widely acceptable result. As a result, in addition to manual tuning methods, automated tuning methods have grown in popularity in recent years [22].

J. Evaluation Matrix

A confusion Matrix is a matrix that displays how accurately a model predicts a given situation, as follows in Table I. One method for evaluating a classification model's performance is the confusion matrix. The confusion matrix has a 2x2 table for categorizing models with data as A or B in its most basic form [4].

TABLE I
CONFUSION MATRIX

Actual Class	**Predict Class**	
	Positive	**Negative**
Positive	True Positive	False Negative
Negative	False Positive	True Negative

The effectiveness of the classification of test data that hasn't been seen is assessed using a variety of evaluation methodologies. Precision, recall, F-measure, and accuracy are the most frequently applied to text classification. A common objective is maximizing all metrics with 0 and 1. Higher values, therefore, indicate improved categorization performance [23].

Precision and recall are two metrics frequently combined to assess the effectiveness of information retrieval and are widely used statistics in text categorization. More specifically, recall counts how many relevant documents were successfully

DOI : https://doi.org/10.25139/inform.v8i2.5700

retrieved, whereas accuracy counts how many relevant documents were retrieved. The formula is found in Equations (5) and can be used to determine both metrics using Equation (6).

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Rarely are precision and recall considered in isolation. The F-measure, which offers a single weighted statistic for assessing overall performance, frequently combines these two measurements. Using the formula in the Equation, the F-measure may be computed using Equation (7).

$$F1 - measure = 2\,x\,\frac{Precision\ x\ Recall}{Precision + Recall} \tag{7}$$

Accuracy is a different metric used to gauge categorization performance. The quantity of correctly identified samples is how accuracy is calculated. If the classification consistently predicts one class, which can be determined using a formula such as equation (8), accuracy is used [7].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}\,x\,100\% \tag{8}$$

III. Result and Discussion

The method used for sentiment analysis in this study is the SVM method with Feature Extraction Bow and TF-IDF.

A. Data Collecting

At the data collection stage, the author explores data from various datasets, one of which is through the public dataset site, Kaggle. The author obtained a dataset about IMDb review movies, where the total dataset is 50,000 datasets. Furthermore, the dataset is divided into 50% positive and 50% negative, as shown in Figure 2.

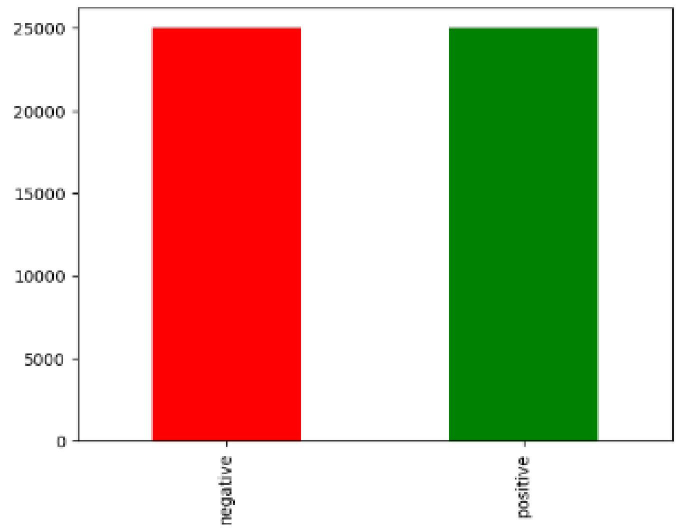

Figure 2. Split Dataset

B. Implementation

In this study, the authors carried out several implementations with several parameters to get maximum results. The parameters tested were Wordcloud before and after preprocessing and then proceeded to the feature extraction stage, namely Bags of Word and TF-IDF.

For the first, the parameter being tested is Wordcloud before preprocessing. In this wordcloud, it processes all the data totaling 65.521.550 words. In Figure 3, these are the words contained in the wordcloud before preprocessing.

Figure 3. Wordcloud Before Preprocessing

After processing all the data, totaling 65.521.550 words, the writer paraphrases and removes the subject words, conjunctions, and typos so that the preprocessing system can read the processed words. Furthermore, the author uses the useful tqdm library to display a progress bar with a simple loop. Figure 4 shows the results of using tqdm.

```
'one reviewers mentioned watching oz episode hooked rig
ht exactly happened first thing struck oz brutality unf
linching scenes violence set right word go trust not sh
ow faint hearted timid show pulls no punches regards dr
ugs sex violence hardcore classic use word called oz ni
ckname given oswald maximum security state penitentary
focuses mainly emerald city experimental section prison
cells glass fronts face inwards privacy not high agenda
em city home many aryans muslims gangstas latinos chris
tians italians irish scuffles death stares dodgy dealin
gs shady agreements never far away would say main appea
l show due fact goes shows would not dare forget pretty
pictures painted mainstream audiences forget charm forg
et romance oz not mess around first episode ever saw st
ruck nasty surreal could not say ready watched develope
d taste oz got accustomed high levels graphic violence
not violence injustice crooked guards sold nickel inmat
es kill order get away well mannered middle class inmat
es turned prison bitches due lack street skills prison
experience watching oz may become comfortable uncomfort
able viewing thats get touch darker side'
```

Figure 4. Result After Stopwords

DOI : https://doi.org/10.25139/inform.v8i2.5700

Furthermore, the second parameter tested is the wordcloud after preprocessing. In this preprocessing, perform data processing that was previously processed before preprocessing. In this preprocessing, there are 41,346,691 words. In Figure 5. these are the words from wordcloud after preprocessing. After that, it is continued to the feature extraction process and evaluation matrix.

Figure 5. Wordcloud After Preprocessing

Following the conversion of the data into a matrix in the feature extraction stage, the matrix and labels obtained in the following stage will be learning material for SVM, allowing SVM to predict the label of a matrix with the same feature as the previously studied matrix. The SVM's method of operation is to create a hyperplane that divides data into several classes. Then, using grid-search, optimize against the SVM parameters to find the best parameters that SVM models can use. The Grid-search will attempt to find the best hyperparameter by using SVM to evaluate each combination of hyperparameters. And then, choose the best combination of hyperparameter values from the options provided.

C. *Evaluation Model*

It may be concluded from the analysis and discussion conducted that SVM has been successfully implemented using the IMDb data for this study (Support Vector Machine). In order to perform this study, the preprocessing process for filtering and classifying data using SVM was completed with 50.000 data, aggregating up to 40.000 reviews for training data and 10.000 reviews for test data. The opinion is divided into 25.000 positive data and 25.000 negative data. 88,59% of the highest accuracy test results were achieved using SVM with Bags of Word (Bow) in Figure 6.

The test results in this study are presented in Table II. The results are 90% precision, 89% recall, and 89% f1-score.

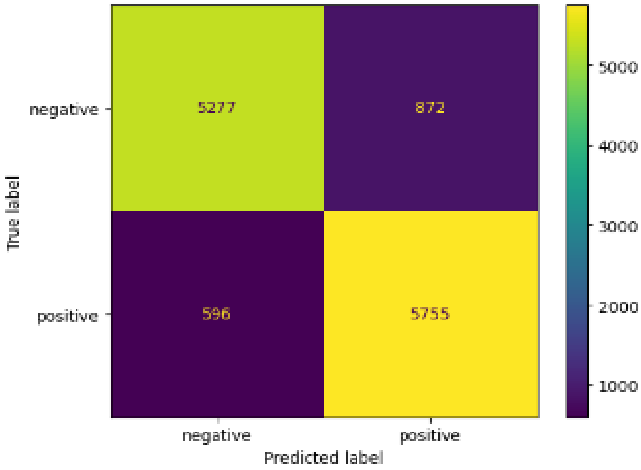

Figure 6. Result Bag of Words (BoW)

TABLE II
RESULT BOW

	Precision	Recall	F1-Score
Negative	0.87	0.90	0.88
Positive	0.90	0.88	0.89
Accuracy	0.90	0.89	**0.89**

Then, 91,27% of the highest accuracy test results were achieved by using SVM with TF-IDF in Figure 7. This result is the best result when compared with BoW.

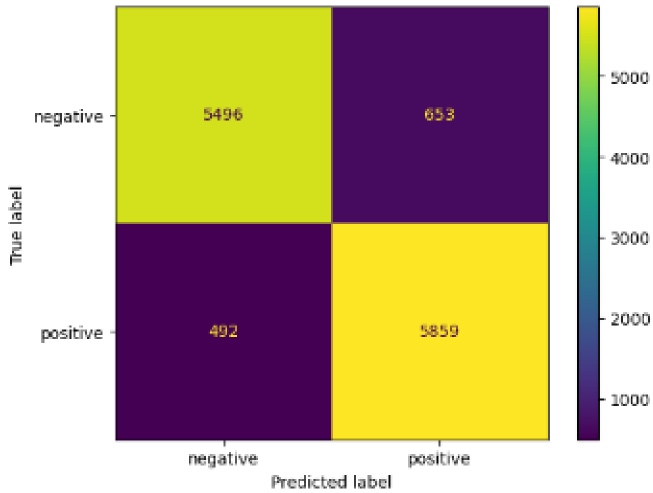

Figure 7. The Result TF-IDF

The test results in this study are presented in Table II. The results are 91% precision, 92% recall, and 91% f1-score.

TABEL III
RESULT TF-IDF

	Precision	Recall	F1-Score
Negative	0.90	0.92	0.91
Positive	0.92	0.91	0.91
Accuracy	0.91	0.92	**0.91**

DOI : https://doi.org/10.25139/inform.v8i2.5700

IV. CONCLUSION

Taking into account the findings of the analysis and discussion that have been conducted, it can be interpreted that the data for this study succeeded in implementing the Support Vector Machines (SVM) algorithm for IMDb Movie Review. The preprocessing process for filtering data and calculating word weights using TF-IDF to classify data using Support Vector Machines (SVM). Using grid-search, optimize against the SVM parameters to find the best parameters that SVM models can use. It was successfully carried out with 50,000 data by grouping as many as 40,000 tweet data for training data and 10,000 tweets for test data to conduct this research. The results of the accuracy test showed our model achieves Support Vector Machines (SVM) with Bags of words (BoW) used to obtain the highest accuracy test results with an accuracy of 88,59%. Then, our model achieves the Term Frequency–Inverse Document Frequency (TF-IDF) used to obtain the highest accuracy test results with an accuracy of 91.27%.

A complete word dictionary (known as a "library stop word" in English) and linguistic specialists are anticipated to be used in future research to increase the results' accuracy. To analyze models and algorithms that have high accuracy for handling specific sentiment analysis cases following the data held for results with high accuracy. Case studies with various social media platforms and comparisons with other models and algorithms are also advised for further research.

ACKNOWLEDGMENT

We thank NLP, Support Vector Machine, and Machine Learning for cardinal support and for helping us complete the research work. Thanks also to Kaggle for the support in the form of datasets. Amikom University supported this work.

REFERENCES

[1] D. Can, A. Kazemzadeh, F. Bar, H. Wang, and S. Narayanan, "A System for Real-time Twitter Sentiment Analysis of 2012 US Presidential Election Cycle Singing View project Emotion Prediction From Movies View project A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle," no. July, pp. 8–14, 2012, [Online]. Available: https://www.researchgate.net/publication/262326668.

[2] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 4, pp. 1–25, 2018, doi: 10.1002/widm.1253.

[3] E. Kouloumpis, T. Wilson, and J. Moore, "Twitter Sentiment Analysis: The Good the Bad and the OMG!," *Proc. Int. AAAI Conf. Web Soc. Media*, vol. 5, no. 1, pp. 538–541, 2021, doi: 10.1609/icwsm.v5i1.14185.

[4] H. Li, Y. Ma, Z. Ma, and H. Zhu, "Weibo text sentiment analysis based on bert and deep learning," *Appl. Sci.*, vol. 11, no. 22, 2021, doi: 10.3390/app112210774.

[5] Hanafi, N. Suryana, and A. S. H. Basari, "Generate Contextual Insight of Product Review Using Deep LSTM and Word Embedding," *J. Phys. Conf. Ser.*, vol. 1577, no. 1, 2020, doi: 10.1088/1742-6596/1577/1/012006.

[6] Hanafi, N. Suryana, and A. S. Bashari, "Evaluation of e-Service Quality, Perceived Value on Customer Satisfaction and Customer Loyalty: a Study in Indonesia," *International Business Management*, vol. 11, no. 11. pp. 1892–1900, 2017, [Online]. Available: https://medwelljournals.com/abstract/?doi=ibm.2017.1892.1900.

[7] G. Eom, S. Yun, and H. Byeon, "Predicting the sentiment of South Korean Twitter users toward vaccination after the emergence of COVID-19 Omicron variant using deep learning-based natural language processing," *Front. Med.*, vol. 9, no. September, pp. 1–13, 2022, doi: 10.3389/fmed.2022.948917.

[8] P. Ashokkumar, G. Siva Shankar, G. Srivastava, P. K. R. Maddikunta, and T. R. Gadekallu, "A Two-stage Text Feature Selection Algorithm for Improving Text Classification," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 3, 2021, doi: 10.1145/3425781.

[9] S. Bhatia, M. Sharma, and K. K. Bhatia, "Sentiment Analysis and Mining of Opinions," *Stud. Big Data*, vol. 30, no. May, pp. 503–523, 2018, doi: 10.1007/978-3-319-60435-0_20.

[10] F. A. Pozzi, E. Fersini, E. Messina, and B. Liu, *Sentiment Analysis in Social Networks*. 2016.

[11] A. Hakim Dalimunthe, R. Aditiya, and R. Watrianthos, "Implementation Naïve Bayes Classification for Sentiment Analysis on Internet Movie Database," *Technol. Sci.*, vol. 4, no. 1, pp. 4–9, 2022, doi: 10.47065/bits.v4i1.1468.

[12] V. Korsunova and O. Volchenko, "Cultural Modernisation And Film Industry: Naked Facts From IMDB," *SSRN Electron. J.*, 2021, doi: 10.2139/ssrn.3915414.

[13] G. Ignatow and R. Mihalcea, *An Introduction to text Mining*. 2017.

[14] Z. Jiang, B. Gao, Y. He, Y. Han, P. Doyle, and Q. Zhu, "Text Classification Using Novel Term Weighting Scheme-Based Improved TF-IDF for Internet Media Reports," *Math. Probl. Eng.*, vol. 2021, no. ii, 2021, doi: 10.1155/2021/6619088.

[15] M. Lan, C. L. Tan, J. Su, and Y. Lu, "Supervised and traditional term weighting methods for automatic text categorization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 721–735, 2009, doi: 10.1109/TPAMI.2008.110.

[16] T. Sabbah *et al.*, "Modified frequency-based term weighting schemes for text classification," *Appl. Soft Comput.*, vol. 58, pp. 193–206, 2017, doi: 10.1016/j.asoc.2017.04.069.

[17] M. McTear, Z. Callejas, and D. Griol, "The conversational interface: Talking to smart devices," *Conversational Interface Talk. to Smart Devices*, no. 2009, pp. 1–422, 2016, doi: 10.1007/978-3-319-32967-3.

[18] S. Raj, Pethuru; Deepu, "A Framework for Text Analytics using the Bag of Words (BoW) Model for Prediction," *Int. J. Adv. Netw. Appl.*, pp. 975–0282, 2016, [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Bag+of+Words.

[19] W. S. Noble, "What is a support vector machine?," *Nat. Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, 2006, doi: 10.1038/nbt1206-1565.

[20] R. Freund, F. Girosi, and E. Osuna, "An Improved Training Algorithm for Support Vector Machines 1 Introduction 2 Support Vector Machines," *Neural Networks Signal Process. [1997] VII. Proc. 1997 IEEE Work.*, pp. 276–285, 1997, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.37.7400%5Cnhttp://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=622408&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D622408.

[21] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis," *Informatics*, vol. 8, no. 4, pp. 1–21, 2021, doi: 10.3390/informatics8040079.

[22] R. Hossain and D. D. Timmer, "Machine learning model optimization with hyper parameter tuning approach," *Glob. J. Comput. Sci. Technol.*, vol. 21, no. 2, pp. 7–13, 2021.

[23] E. Beauxis-Aussalet and L. Hardman, "Simplifying the Visualization of Confusion Matrix," pp. 1–2.

DOI : https://doi.org/10.25139/inform.v8i2.5700