Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

Comparison of Naive Bayes and SVM Methods for Identifying Anxiety Based on Social Media

Endri Rizki Nugraha*, Warih Maharani

School of Computing, Informatics, Telkom University, Bandung, Indonesia
Email: [1,*]exdxix@student.telkomuniversity.ac.id, [2]wmaharani@telkomuniversity.ac.id
Correspondence Author Email: exdxix@student.telkomuniversity.ac.id
Submitted: **23/12/2024**; Accepted: **26/02/2025**; Published**: 01/03/2025**

Abstract−This research aims to detect anxiety patterns from social media posts using Naive Bayes (NB) and Support Vector Machine (SVM) algorithms. Tweets are extracted using Data Crawling techniques, then continued their way into labeling using Depression Anxiety Stress Scale (DASS-42) questionnaire along with Random Oversampler to balance out the unbalanced dataset and NB and SVM were chosen for their effectiveness in text sentiment classification. This study integrates textual features obtained from the Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words (BoW) methods. The study compares the performance of these algorithms in detecting anxiety using datasets from the X platform. The comparison aims to identify the advantages and limitations of each method in handling textual sentiment data. This research aims to analyze sentiment data by calculating accuracy, recall, and F1-score to determine the most optimal performance outcome. The results indicate that the SVM with TF-IDF feature extraction achieved the highest accuracy of 72% and an average F1-Score of 61%, while the NB with BoW achieved 56% accuracy and an average F1-Score of 49%. These findings highlight the effectiveness of combining SVM and TF-IDF features which improve model effectiveness with SVM producing the best overall result in identifying anxiety from social media data.

Keywords: Anxiety; Naive Bayes; SVM; DASS-42; Sentiment Analysis; Social Media

1. INTRODUCTION

In this era of globalization, technology has advanced significantly, and the development of communication technology and the internet has made everyday life easier, more effective, and efficient. The internet has both positive and negative impacts on human life [1]. According to Aldiansyah Putra et al. [2], the number of social media users increased by 13.2% in January 2021, reaching 4.2 billion users compared to the same period the previous year.

Anxiety is a state experienced in response to threats that are either distal or uncertain, and involves changes in an individual's subjective state, behavior and physiology that facilitate detection of a potential threat within the environment [3]. A study by Mukhlasin shows that anxiety can be triggered by genetic factors, past experiences that affect psychological well-being, clinging too much on some substances, or too much time spent in social media. Lately, social media use is often linked to mental health problems [4]. Excessive use of social media has become a concern for researchers, parents, and society [4]. The American Psychological Association (APA) explains that anxiety among teenagers is specified by stress, constant worry, physical reaction such as increased heart rate, and concerns about uncertain things. Anxiety mainly appear from a fear of unknown objects or situations without a clear reason [5]. In this era of digitalization, social media has become a platform to exchange information and to acknowledge the condition of others that makes it one of the most relevant data source for analyzing a person's mental health [6]. X (formerly Twitter) is commonly used by society to express opinions on a range of issues, making it one of the most reliable data source for sentiment analysis to this day [6].

Two commonly used methods in sentiment classification are Naive Bayes (NB) and Support Vector Machine (SVM). Research performed by Ridho Fazal et al. [6] compares NB and SVM algorithms for sentiment analysis on X, indicated higher performance for SVM, with an accuracy of 88.52%, while NB achieved an accuracy of 82.51%. Another study by Widia Ningsih et al. [7] NB and SVM in sentiment analysis on X in the case of electric vehicle usage in Indonesia, resulting in an accuracy of 70.83% for SVM and 63.02% for NB. Another study by Moh. Aminullah Al Fachri and Ummi Athiyah [8] compared the NB and SVM algorithms to analyze the issue of expensive cooking oil using a dataset of 9,194, which showed an accuracy of 81.6% for both SVM and NB. In research by Rani Yunita et al. [9], SVM and NB algorithms were applied to sentiment analysis of ChatGPT on X, using a dataset of 1,000, with results showing 59% accuracy for SVM and 47% for NB. A study by Dewi Setiyawati and Nuri Cahyono [10] analyzed sentiments about smoking in Indonesia on X, using a dataset of 2,772, resulting in an accuracy of 60.8% for SVM and 62.1% for NB. However, these studies often encountered challenges with imbalanced datasets or inadequate feature extraction methods for this kind of task, leading to inconsistent accuracy and F1-scores. Therefore, this study is necessary to improve the performance of NB and SVM by addressing these limitations through the use of advanced feature extraction methods like TF-IDF and BoW and balancing techniques like Random Oversampler. Based on the above explanation, this study focuses on comparing the prediction algorithms for anxiety using NB and SVM to evaluate the performance of both. The aim of comparing these two methods is to determine which algorithm performs the best.

This study's contributions are twofold. First, the work presents a comparative analysis of the two popular algorithms in the context of mental health analytics. Then, it also points out the role of feature engineering in improving model performance and provides practical insight into future research and applications. Based on data

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

from X, this study further develops anxiety detection and emphasizes the use of social media for early intervention in mental health.

2. RESEARCH METHODOLOGY

2.1 Research Stages

In this research, our system is developed to classify tweets into two kinds of sentiment classes. To implement it, a structured methodology with numerous steps is applied. Under here is the detailed process illustrated in Figure 1 Flowchart System:

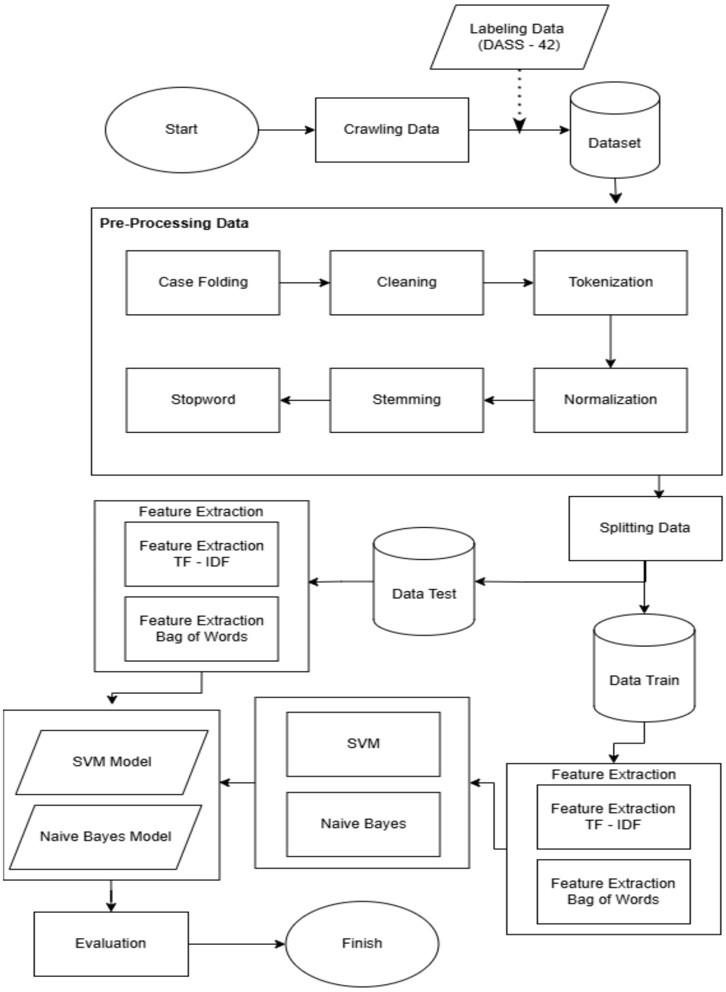

Figure 1. Flowchart System

Based on Figure 1 It is explained that the process of collecting data for this research involves a few steps, starting with the link distribution to the DASS-42 questionnaire for a wider audience. The questionnaire results are used to collect anxiety labels. Furthermore, the participants that have agreed to share their data for this study are crawled to collect their overall tweets. The collected data is then put into a single .csv dataset that stores the usernames, labels, and tweets of the participant. After the data collection process comes the preprocessing stage. This makes it an important step to guarantee consistency and improve the quality of the data and to make the preprocessed dataset ready for the next stages. Preprocessing includes converting text to lowercase, removing emojis and symbols, tokenization, normalization, stemming, and eliminating stopwords, as can be seen in Table 6.

The next stage is to split the dataset into several training and testing sets for the different testing scenarios to be tested. The first testing scenario is to change the proportion of data used for training. Then, feature extraction techniques such as Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) are applied to each testing scenario. BoW creates numerical vectors based on word frequency in the text, while TF-IDF assigns weights to words by considering their frequency within the document relative to their presence across the entire dataset. These numerical representations are used for training and evaluating the performance of the models.

The performance of the Naive Bayes and Support Vector Machine algorithms is compared in this next stage. While the NB algorithm predicts the class with the highest probability based on Bayes theorem, the SVM algorithm identifies the optimal hyperplane to separate data into distinct groups. Kernel

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

functions are used by SVM to handle such a high-dimensional space of the dataset. The obtained output of both the models would give important insights on various strengths and limitations that each of their applications might have with text classification tasks. The confusion matrix is finally used, showing several performance metrics that involve accuracy, precision, recall, and F1-score for each model.

2.2 Dataset

The dataset used in this research includes a total of 156 X account that gave permission for the data to be used and is willing to be a part of this research. The dataset value variates from 300 tweets to 50.000+ tweets per user that makes the 156 data entries enough for this research, which were subsequently tested and labeled based on the results of the DASS-42 questionnaire as can be seen in Table 3,4 and 5. The labeling categorizes the data into two classes: 0 (not diagnosed with anxiety) and 1 (diagnosed with anxiety). Table 1 provides an overview of the data distribution.

Table 1. Imbalanced Label

Label	Total Data
0	48
1	108

The raw dataset used in this research is imbalanced, with fewer entries labeled as class label 0 (not diagnosed with anxiety) compared to class label 1 (diagnosed with anxiety). This imbalanceness can cause bias in the classification model, making it difficult to accurately predict the minority class. To handle this imbalance dataset, the random oversampler technique is then used in this research to reduce bias. Random oversampler works by duplicating samples from the minority class (label 0) and deleting some samples from the majority class (label 1) to create a balanced dataset. By increasing the minority class and reducing the majority class ensures that the model does not become biased towards the majority class (label 1) [11]. Repetition can be done either randomly or by intentionally selecting samples that lie on the boundaries between the majority and minority classes. The use of oversampling in this study is to readjust the proportions of minority and majority samples within a dataset, without introducing any new information [11]. Resulting in a more evenly distributed and balanced dataset that can improve the model's ability to learn and accurately predict both classes. Table 2 provides an overview of the data distribution after balancing process.

Table 2. Balanced Label

Label	Total Data
0	64
1	64

2.2.1 X

X is one of the most used media platforms that is known as a place to communicate and share your thoughts among the public. One of the services provided by X is the creation of status messages called tweets or tweets that can be read by other X users [12]. Typically, these tweets express user's opinions on various topics, with a character limit of 140 [12]. Apart from that, X is also known for the use of hashtags (#) which allows users to mark relevant keywords or phrases in their tweets. This helps categorize tweets and make it easier to search and participate in trending topics. The use of hashtags also makes it easier to classify text because it can show the emotions or opinions of X users. Thus, X has become one of the platforms that collects opinion data from people around the world [12].

2.2.2 Anxiety Identification

Anxiety is a condition where a person feels worried or afraid that something negative will happen in the future. This feeling causes uneasy feeling, anxiety or tension, a sense of uncertainty that causes unease, worry, and even fear [13]. Identification is the process of determining or predicting future events using relevant historical information through scientific methods. The goal is to gain insight into what may happen in the future with a high probability of occurrence [13]. Anxiety identification can help in recognizing early signs of anxiety, enabling well timed interventions and personalized strategies to manage symptoms before they escalate into a bigger mental health issues. This approach can lead to better outcomes by addressing the root causes and providing support when needed.

2.2.3 DASS-42

The DASS-42, or Depression, Anxiety, and Stress Scales – 42 items, is a psychological instrument designed to assess the levels of depression, anxiety, and stress in individuals. This tool consists of 42 questions, each scored using a 4-point Likert scale, ranging from "0 (Not at all)" to "3 (Very Much)." Based on the total score from the Likert scale, individuals can be classified as experiencing depression, anxiety, or stress within categories of mild, moderate, severe, or extremely severe [13]. Tabel 3 provides an overview of DASS-42.

Copyright © 2025 **Author**, Page 2186
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

Table 3. Measured Variable

Measured Variable	Question number in questionnaire
Stress	1, 6, 8, 11, 12, 14, 18, 22, 27, 29, 32, 33, 35, 39.
Anxiety	2, 4, 7, 9, 15, 19, 20, 23, 25, 28, 30, 36, 40, 41.
Depression	3, 5, 10, 13, 16, 17, 21, 24, 26, 31, 34, 37, 38, 42.

Table 4 under provides the DASS-42 measurement score. This table outlines the scoring mechanism for the DASS-42 questionnaire, where each score corresponds to how frequently an individual identifies with a specific statement. The scores range from 0 to 3, representing levels of agreement from "never" to "very often". After the measurement results containing the total scores for each variable (stress, anxiety, and depression) are obtained, these scores will be categorized into several categories: normal, mild, moderate, severe, and extremely severe. This categorization allows the DASS-42 questionnaire to classify the severity of stress, anxiety, and depression for individuals [13].

Table 4. Measurement Score

Score	Description
0	Does not match yourself at all (never)
1	Match yourself to a certain degree (sometimes)
2	Match yourself to a considerable extent (often)
3	Match yourself (very often)

The scoring mechanism in Table 4 above forms the basis for establishing the severity of mental health conditions. The total score, after summing up the individual scores for each variable, would help in identifying whether the person falls into a normal or heightened category of stress, anxiety, or depression. This step is quite important to ensure that the data is well labeled for analysis in this study.

Table 5. Questionnaire Category

Category	Stress	Anxiety	Depression
Normal	0-14	0-7	0-9
Mild	15-18	8-9	10-13
Moderate	19-25	10-14	14-20
Severe	26-33	15-19	21-27
Very Severe	34+	20+	28+

Table 5 above classifies each of the variables into different scores representing a particular range of severity; in doing so, enabling the researcher to obtain precisely the exact intensity. For instance, regarding the case of stress, any score ranging between 0 and 14 defines the normal score, and at or over 34 it was considered very severe. Therefore, such classification can yield accurate categorization regarding mental disorder based on the DASS-42 questionnaire.

2.3 Data Preprocessing

This process concentrates on the preprocessing of data collected in this process. These involve various key steps, including converting text to lowercase, emoji and symbols removal, tokenization, normalization, stemming, and removing stopwords. These steps are are important for putting the text into a standardized form and preparing it for the following analysis. The preprocessing workflow can be adjusted or expanded depending on the characteristics and needs of the dataset. A complete illustration of this process is provided below at Table 6 below.

Table 6. Preprocessing Steps

Steps	Preprocessed Text
Full Text	OMG I'm so tired today! Who wants to join me for a quick coffee?
Case Folding	omg i'm so tired today! who wants to join me for a quick coffee?
Cleansing	omg im so tired today who wants to join me for a quick coffee
Tokenization	['omg', 'im', 'so', 'tired', 'today', 'who', 'wants', 'to', 'join', 'me', 'for', 'a', 'quick', 'coffee']
Normalization	['oh', 'my', 'god', 'im', 'very', 'tired', 'today', 'who', 'wants', 'join', 'me', 'quick', 'coffee']
Stemming	['oh', 'my', 'god', 'im', 'veri', 'tire', 'today', 'who', 'want', 'join', 'me', 'quick', 'coffe']
Stopword Removal	['oh', 'god', 'tire', 'today', 'want', 'join', 'quick', 'coffe']

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

With these refinements, the dataset was prepared for feature extraction and training the machine learning models. The final dataset was saved for future use, with the preprocessed text organized into a cleaned format suitable for classification tasks.

2.4 Feature Extraction

Feature extraction involves converting a text document from its original format into a set of features that can be easily handled by text classification methods that makes it a crucial preprocessing step in data mining and text classification, where feature values within documents are calculated. Therefore, techniques such as TF-IDF and BoW are very much used because they are effective in weighting terms [14]. The dataset was split into training and testing sets prior to feature extraction, which will later be used for evaluating the performances of the SVM and NB models. Feature extraction techniques like BoW and TF-IDF are applied to convert these splits of data into numerical formats. In the BoW technique, the process involved is to create vectors depending on the frequency of occurrence of words. Each of these vector values reflects the frequency of occurrence for a given word in any document. On the other hand, the TF-IDF method performs weighting regarding how often a word appears in just one document versus how it is generally spread across all other documents. By this stage, data preprocessing will have been concluded in preparation for the training and subsequent testing of the models.

2.4.1 TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) is a method of weighing the words in a document to indicate their importance regarding the representation of its content. It takes the frequency of a word inside the document, represented as TF, and measures the inverse of how many documents contain the same word, which is represented by IDF. The equation of TF-IDF is given in equation (1) [15].

$$IDF_t = log\frac{N}{DF_t} \tag{1}$$

The last step in TF-IDF weighting involves multiplying the TF and IDF values, as seen in equation (2) below [15].

$$w_t = TF_t \times IDF_t \tag{2}$$

2.4.2 Bag of Words

Bag of Words (BoW) is a basic feature extraction that is commonly used in text classification. It converts textual data into a numerical vector, where each element represents the frequency of a particular word in the document [16]. BoW refers to the analysis of text by ignoring the order, grammar, and context of words, focusing only on their existence. This method is highly efficient and widely used for machine learning models. However, it does not capture the relationships between each words, which can limit its ability to fully understand the context. But, the simplicity makes it a useful starting point for many text classification tasks. The BoW formula is shown in equation (3) below [17].

$$BoW(w, d) = Frequency(w\ in\ document\ d) \tag{3}$$

2.5 Model Classification

This next step performs model classification to evaluate and compare the performance of the NB model against the SVM model. Naive Bayes works in such a way that it predicts the class with the highest probability, using the calculation of prior and posterior probabilities for each class based on Bayes theorem. While Support Vector Machine looks for the best decision boundary or hyperplane that can differentiate data points into distinct groups. SVM uses kernel functions to handle high-dimensional feature datasets effectively or even infinite dimensional spaces effectively. This helps in comparing how these two algorithms work in several different situations and gives insights into their strengths and limitations when applied to text classification tasks.

2.5.1 Naive Bayes

The Naive Bayes algorithm uses the principles of probability, that is used in classification for sentiment analysis. This algorithm is known for the simplicity of it, as well as its ability to estimate events effectively based on classification results [18]. NB is a basic Bayesian statistical technique, this method is given the nickname Naïve because it uses the assumption that each variable has a role in the classification process and that they are all independent of each other, a circumstance known as class conditional independence [19]. In NB algorithm, each opinion is represented by a pair of attributes "x1, x2, x3,...xn" where x1 is the first word, x2 is the second word and so on [20]. This is based on Bayes principle about conditional probability which can be seen in formula (4).

$$P(Ci|X|) = \frac{P(X|Ci|)\,P(Ci)}{P(X)} \tag{4}$$

In the Naive Bayes algorithm, **P(Ci|X)** represents the posterior probability, which is the probability of a class **Ci** given the data **X**. **P(Ci)** denotes the prior probability, reflecting the initial probability of a class before observing the data. **P(X|Ci)** represents the likelihood, which is the probability of observing the data **X** given that it belongs to

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

class **Ci**. Lastly, **P(X)** acts as the evidence or normalization factor, ensuring that the sum of probabilities across all classes equals 1.

2.5.2 Support Vector Machine

SVM is a method for finding the optimal hyperplane by maximizing the distance between data classes. SVM is a model used to carry out classification and predictions, where no lines or areas can be created to separate data classes. In its concept, SVM tries to find the best hyperplane among unlimited functions. The equation of SVM [7] can be seen in equation (5).

$$f(Xd) = \sum_{i=1}^{ns} aiyi\vec{xi}\,\vec{xd} + b \quad (5)$$

In the Support Vector Machine (SVM) equation, $\boldsymbol{n_s}$ indicates the number of support vectors that determine the hyperplane. The term $\boldsymbol{a_i}$ refers to the weight value of each data point, while $\boldsymbol{y_i}$ represents the class label of the corresponding data point. $\boldsymbol{x_i}$ refers to the support vector variables, and $\boldsymbol{x_d}$ is the data point being classified. Finally, $\boldsymbol{b}$ signifies the error or bias term that adjusts the hyperplane's position to maximize the margin between different classes.

2.6 Evaluation

After the dataset went through preproccessing and modelling, the next step is evaluation. The performance of the sentiment analysis model is evaluated using a Confusion Matrix along with several metrics such as accuracy, precision, recall, and F1-score. The formula for calculating this evaluation is as follows:
Accuracy: This metric assesses the overall predictive capability of the model.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision: This measures the proportion of correct positive predictions among all positive predictions made.

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

Recall: This calculates the ratio of correct positive predictions to the total actual positive instances.

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

F1-score: This is the harmonic means of precision and recall, providing a balance between the two.

$$F1\ Score = \frac{2(Recall.Precision)}{Precision+Recall} \quad (9)$$

3. RESULT AND DISCUSSION

This section presents the results of the experiments conducted in this study, accompanied by a detailed discussion. The focus is on evaluating model performance, comparing different approaches, and analyzing the outcomes using various evaluation metrics.

3.1 Data Gathering and Labeling

This research uses data crawled from social media, specifically focusing on tweets from the X platform. Tweets that fit keywords related to anxiety and mental health were gathered so that this dataset would match the purposes of this paper. The Depression Anxiety Stress Scale, commonly known as DASS-42, questionnaire was used as a guide to allow the labeling of data. Each of the tweets was labeled in categories corresponding to the DASS-42 criteria of anxiety scoring: normal, mild, moderate, severe, and very severe. This labeling methodology ensures that a structured dataset is maintained. Due to the nature of social media data, which is always imbalanced, the number of tweets in the "normal" category is very large, the Random Oversampler technique was used to balance the dataset. This ensured that the models trained on a balanced representation of each class to reduce potential biases.

3.2 Data Crawling and Preprocessing

These tweets are gathered using the Twitter API itself, in search for anxiety related keywords. After collecting the data in raw form, it had gone through some preprocessing steps as shown in Table 6 such as case folding, text cleaning, tokenization, normalization, stemming, and stopword removal. The preprocessed steps of normalizing text and eliminating noise were much needed for this dataset. therefore, more meaningful and feature extractive data resulted.

3.3 Feature Extraction

The preprocessed data was transformed into numerical form for machine learning, two techniques were used in this study: TF-IDF and BoW. TF-IDF was performed to give significant weight to the words over their importance in the general dataset, hence reducing common terms that may be overwhelmingly present. Meanwhile, the BoW

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

representation simply translates text into a word occurrence matrix. Both methodologies were chosen for their excellent performance in the capture of patterns in texts. These feature extraction techniques were then compared to decide which one works best with the Naive Bayes and Support Vector Machine algorithms.

3.4 Training Scenario Result

Before presenting the training scenario results that can be seen in Table 7, the testing scenario were choosed to evaluate the performance of the classification models under different conditions. These scenarios involve varying the split ratios of the dataset into training and testing sets. Commonly used split ratios, such as 40/60, 30/70, 20/80, and 10/90, are used to understand how the proportion of data used for training impacts the model's predictive capability. The two feature extraction is used to convert the textual data into numerical format for the machine learning model. This study tries to explore different settings of split ratios, and feature extraction methods to get the optimal combination. In this study, we evaluated the performance of Multinomial Naive Bayes and Support Vector Machine using TF-IDF and BoW for the feature extraction. In Table 7 below, a series of test scenarios are presented, comparing the various splits of test and training data. These different splits were used to assess the model's accuracy under different conditions. To ensure consistency across tests, the max features on BoW is set to 150 and keeping the default random parameter value state at 42. The results allow for the comparison of the highest accuracy values achieved across different data splits and feature extraction methods.

Table 7. Splitting Comparison

Split Ratio	Model	TF-IDF	BoW
40/60	NB	0.5625	0.5714
40/60	SVM	0.6984	0.5714
30/70	NB	0.5319	0.5106
30/70	SVM	0.6170	0.5532
20/80	NB	0.5625	0.5625
20/80	SVM	0.7188	0.5625
10/80	NB	0.4375	0.5625
10/80	SVM	0.6875	0.6875

From the data split experiments, it was found that the 20/80 test-train split yields the best overall accuracy that has an average accuracy of 0.6015. The SVM model achieved the highest accuracy of 0.7188 when using TF-IDF as the feature extraction method. Meanwhile, the Naive Bayes model reached its best performance with BoW feature extraction, achieving an accuracy of 0.5714.

3.5 Final Result

In Table 8 below is the SVM evaluation matrix with data split at 20/80 and TF-IDF feature extraction.

Table 8. SVM Evaluation

Class	Precision	Recall	F1-Score	Support
0	0.60	0.30	0.40	10
1	0.74	0.91	0.82	22
Accuracy		0.72		
Macro Avg	0.67	0.60	0.61	32
Weighted Avg	0.70	0.72	0.69	32

The SVM model performs effectively with an overall accuracy of 0.72, which means it is able to correctly predict 72% of the total data. Class 1 (diagnosed with anxiety) performed best with an F1-Score of 0.82. This model shows a limitation in classifying class 0 (not diagnosed with anxiety), as it achieves an F1-Score of only 0.40 for this class. This indicates that the model frequently misclassifies class 0.

Table 9. NB Evaluation

Class	Precision	Recall	F1-Score	Support
0	0.30	0.30	0.30	10
1	0.68	0.68	0.68	22
Accuracy		0.56		
Macro Avg	0.49	0.49	0.40	32
Weighted Avg	0.56	0.56	0.56	32

Table 9 above shows the result of the NB model on split data in 20/80 and using BoW feature extraction. As shown, the model achieves an overall accuracy of 0.56, indicating that it can correctly predict 56% of the total data. Class 1 (diagnosed with anxiety) performed best with an F1-Score of 0.68. This model shows a limitation in classifying class 0 (not diagnosed with anxiety), as it achieves an F1-Score of only 0.30 for this class, a bit worse than the SVM

This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

model does. This indicates that the model frequently misclassifies class 0. The following Table 10 presents the evaluation from each model with 20/80 split ratio.

Table 10. Final Comparison

Model	Accuracy	Macro Avg Precision	Macro Avg Recall	Macro Avg F1-Score
SVM (TF-IDF)	0.72	0.67	0.60	0.61
NB (BoW)	0.56	0.49	0.49	0.49

The results indicate that the SVM model using TF-IDF feature extraction outperforms the NB model with BoW feature extraction in all evaluated aspects. The SVM model demonstrated superior performance, achieving an accuracy of 0.72, a Macro Average Precision of 0.67, a Macro Average Recall of 0.60, and a Macro Average F1-Score of 0.61. These results highlight that the SVM model, when paired with TF-IDF feature extraction, excels in analyzing data patterns and making more accurate predictions compared to the top-performing Naive Bayes model.

3.6 Suggestion

a. *Dataset contains too many Retweets*
Having too many retweets in a dataset can greatly disrupt the accuracy of sentiment analysis. Retweets often lack the original context or sentiment of the content they share and typically represent the act of sharing rather than the retweeter's own feelings or opinions. Retweets often contain incomplete phrases, emojis, or bits out of context that don't provide anything meaningful for analysis. That can also lead to a huge amount of repetitive content able to distort the balance of sentiment. This might be the starting point for further research: during data collection, the filtering mechanism could be set in such a way as to exclude retweets, thus assuring a richer and more varied dataset with original textual content.

b. *Limited Dataset*
The dataset used in this research is gathered from the DASS-42 questionnaire that was distributed to a wide audience. Though many responded to the questionnaire, few agreed to the usage of their information in this research. Because of this constraint, it resulted in a small dataset, which might affect the model's generalization capability. Other future research on data collection development could extend to other sources like online forums or public social media posts that would hugely help in adding size to the dataset with diversity to improve the validity of the analysis.

4. CONCLUSION

This study demonstrates that the SVM algorithm combined with TF-IDF feature extraction outperforms Naive Bayes (NB) paired with Bag of Words (BoW) in detecting anxiety from social media posts. The SVM model achieved an accuracy of 72% and a macro-average F1-score of 0.61, indicating a better balance of precision and recall compared to NB, which reached 56% accuracy and a macro-average F1-score of 0.49. The superior performance of SVM is consistent with previous studies, such as those by Ridho Fazal et al. and Widia Ningsih et al., which also highlighted the robustness of SVM in text classification tasks. However, the relatively lower performance in this study compared to others may be attributed to the smaller dataset size. The random oversampler technique performed very well in this study, mostly impacting the recall and F1-Score of the overall performance results. In general, this study has proved that the TF-IDF feature extraction, especially in the SVM model with linear kernel has consistently provided the best result, showing the importance of pairing the right feature extraction method with the right algorithm to enhance the accuracy and effectiveness of the models. These findings emphasize the importance of selecting a suitable algorithms, balanced dataset, and feature extraction methods to optimize model performance. Future work should focus on expanding the dataset from gaining extra participants that are willing to fill the DASS-42 questionnaire and agreeing on having their data studied and used as a dataset in this research. This research highlights the potential of using social media data for early detection of mental health issues. By analyzing publicly available social media data, such models could be integrated into social media platforms to provide users with personalized feedback or suggestions, this could help adressing mental health issues before they escalate, especially in populations that may not actively seek professional help due to lack of access to resources. However, deploying these models in real-world scenarios also raises ethical considerations, such as ensuring user privacy, avoiding misclassification, and addressing potential biases in the data. Addressing these challenges is essential to building trust and maximizing the positive impact of these technologies on mental health care.

REFERENCES

[1] A. Trikandini and L. Kurniasari "Hubungan intensitas penggunaan media sosial dengan tingkat kecemasan pada mahasiswa UMKT." *Borneo Studies and Research*, vol. 3, no. 1, 2021.
[2] A. Putra, D. Haeirudin, H. Khairunnisa, and R. Latifah, "Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Pada Media Sosial Twitter Menggunakan Algoritma Svm," *Seminar Nasional Sains dan Teknologi*, p. 1. 2021.

Copyright © 2025 **Author**, Page 2191
This Journal is licensed under a Creative Commons Attribution 4.0 International License

Building of Informatics, Technology and Science (BITS)
Volume 6, No 4, March 2025 Page: 2184−2192
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i4.6506

[3] M. M. Kenwood, N. H. Kalin, and H. Barbas, “The prefrontal cortex, pathological anxiety, and anxiety disorders,” *Neuropsychopharmacology*, vol. 47, no. 1, 2021, doi: 10.1038/s41386-021-01109-z.

[4] R. Mukhlasin, “Dampak Negatif Media Sosial Terhadap Kesehatan Mental dalam Fotografi Ekspresi” *Digilib Institut Seni Indonesia Yogyakarta*, 2022. Available: https://digilib.isi.ac.id/13317/

[5] N. Oktavia Hidayati *et al.*, “Kecemasan Remaja Selama Pandemi Covid-19.” *Jurnal Ilmu Keperawatan Jiwa,* vol. 5, no. 2, 2022.

[6] R. Fazal, L. Andraini, and T. Komputer, “Membandingkan Support Vector Machines Dan Naive Bayes Pada Analisis Sentimen Data Twitter.” *Jurnal Portal Data* 2, vol. 2, no. 10, 2022.

[7] W. Ningsih, B. Alfianda, R. Rahmaddeni, and D. Wulandari, “Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, 2024, doi: 10.57152/malcom.v4i2.1253.

[8] Moh. A. Al Fachri and U. Athiyah, “Komparasi Model Analisis Sentimen Pada Twitter Terhadap Kemahalan Minyak Goreng dengan Metode Naive Bayes dan Support Vector Machine,” *Infotekmesin,* vol. 14, no. 2, 2023, doi: 10.35970/infotekmesin.v14i2.1759.

[9] Rani Yunita and Kamayani Mia, “Perbandingan Algoritma SVM Dan Naïve Bayes Pada Analisis Sentimen Kebijakan Penghapusan Kewajiban Skripsi,” *Indonesian Journal of Computer Science,* vol. 12, no. 5, 2023.

[10] J. Khatib Sulaiman, D. Setiyawati, N. Cahyono, and U. Amikom Yogyakarta, “Analisa Sentimen Pengguna Sosial Media Twitter Terhadap Perokok di Indonesia,” *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 1, 2023, doi: 10.33022/ijcs.v12i1.3154.

[11] R. Asokan, D. P. Ruiz, and S. Piramuthu, “Smart data intelligence : proceedings of ICSMDI 2024” *Springer*, 2024. doi: 10.1007/978-981-97-3191-6.

[12] T. Agustiranti, A. Khalfani Izzati Kurdiana, B. Al Ghiffari, E. Dwi Juniar, and D. Gita Purnama, “Penerapan Naive Bayes Terhadap Sentimen Analisis Media Sosial Twitter Pengguna Kereta Cepat Jakarta-Bandung (Whoosh),” *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI),* vol. 7, no. 1, 2024.

[13] S. R. Marsidi, “identification of Stress, Anxiety, and Depression Levels of Students in Preparation for the Exit Exam Competency Test,” *Journal of Vocational Health Studies*, vol. 5, no. 2, 2021, doi: 10.20473/jvhs.v5.i2.2021.87-93.

[14] A. I. Kadhim, “Term Weighting for Feature Extraction on Twitter: A Comparison Between BM25 and TF-IDF,” in *2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019, doi: 10.1109/ICOASE.2019.8723825.

[15] A. Putri Riani, N. Sulistyowati, T. Ridwan, and A. Voutama, "Klasifikasi Emosi Publik Terhadap Larangan Penggunaan Obat Sirup Menggunakan Algoritma Naive Bayes." *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, vol. 7, no. 2, 2023, doi: 10.46880/jmika.Vol7No2.pp325-339.

[16] L. Galke and A. Scherp, “Bag-of-Words vs. Graph vs. Sequence in Text Classification: Questioning the Necessity of Text-Graphs and the Surprising Strength of a Wide MLP,” *arXiv* 2021, Available: https://doi.org/10.48550/arXiv.2109.03777.

[17] Y. HaCohen-Kerner, D. Miller, and Y. Yigal, “The influence of preprocessing on text classification using a bag-of-words representation,” *PLoS One*, vol. 15, no. 5, 2020, doi: 10.1371/journal.pone.0232525.

[18] F. Amaliah, I. Kadek, and D. Nuryana, “Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter,” *Journal of Informatics and Computer Science (JINACS)*, vol. 3, No. 3 2022, https://doi.org/10.26740/jinacs.v3n03.p384-393.

[19] D. Marutho, “Perbandingan Metode Naive Bayes, KNN, Decision Tree pada Laporan Water Level Jakarta” *Jurnal Ilmiah Infokam* vol. 15, no. 2, 2019, doi: 10.53845/infokam.v15i2.175

[20] M. Imam Syafii, “Sentimen analisis Pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier (NBC).” *Jurnal Teknologi Pintar* Vol. 2, no. 7, 2022.

Copyright © 2025 **Author**, Page 2192
This Journal is licensed under a Creative Commons Attribution 4.0 International License