Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

Comparison of TF-IDF and GloVe Word Embedding for Sentiment Analysis of 2024 Presidential Candidates

Tiara Firdausa Abdillah*, Hasmawati, Bunyamin

Informatics, Faculty of Informatics, Telkom University, Bandung, Indonesia
Email: tiarafa@student.telkomuniversity.ac.id, hasmawati@telkomuniversity.ac.id, bunbunyamin@telkomuniversity.ac.id
Correspondence Author Email: tiarafa@student.telkomuniversity.ac.id
Submitted: **31/07/2024**; Accepted: **10/09/2024**; Published: **11/09/2024**

Abstract−In the ongoing digital era, social media, particularly the social media X, formerly known as Twitter, has become one of the main platforms for sharing public opinions. On the social media, users have the opportunity to express their sentiments or views, including those regarding the presidential election in Indonesia. The main problem in this study is the extent to which public opinion on presidential candidates is reflected in conversations on the social media X. This study involves the combination of Support Vector Machine (SVM) and GloVe Word Embedding algorithms to improve the accuracy of sentiment analysis towards presidential candidates. The performance of the method will be evaluated using a confusion matrix. The results of the study show that while GloVe has the ability to capture global semantic relationships, TF-IDF is more effective in identifying variations and nuances in diverse sentiment data. Therefore, TF-IDF can be a more effective choice for political sentiment analysis in Indonesia, providing more consistent and accurate results. It is seen on the Anies dataset, TF-IDF achieved an accuracy of 0.84 compared to GloVe's 0.82. For the Ganjar dataset, TF-IDF performed better in terms of F1-Score and precision. For the Prabowo dataset, TF-IDF slightly outperformed GloVe in recall, although both techniques had nearly equal high accuracy around 0.93.

Keywords: Presidential Candidates; 2024 Elections; SVM; GloVe; Social media X

1. INTRODUCTION

In the ongoing digital era, social media, particularly the social media x, formerly known as Twitter, has become one of the main platforms for sharing public opinions [1]. Twitter is a relevant source for disseminating information and allows for heterogeneous development [2]. On social media x, users have the opportunity to express their sentiments or views, including those regarding the presidential election in Indonesia. General Elections (Pemilu) are mechanisms used to realize the sovereignty of the people and produce a democratic government in accordance with Pancasila and the 1945 Constitution of the Republic of Indonesia. This election aims to choose the President and Vice President, members of the DPR, DPD, DPRD, as well as regional heads and their deputies who are capable of reflecting democratic values and represent the aspirations of the people in alignment with national development [3].

In its implementation, issues often arise related to fraud, political polarization, and diverse public opinions regarding certain candidates or political parties. Their views on the presidential candidates can be analyzed using sentiment analysis methods, where it can be determined whether the public holds positive, negative, or neutral sentiments. Sentiment analysis is a discipline that specifically examines opinions, evaluations, assessments, attitudes, and emotions expressed in text regarding various subjects, including products, services, organizations, individuals, and other entities [4], [5]. The process of sentiment analysis involves detecting the tendencies of individuals through their writings or texts, aiming to extract information related to sentiments commonly expressed on social media [6]. The method involves identifying the positive, negative, or neutral meanings of the text.

Sentiment analysis on this platform has become a very relevant and interesting focus of research, as evidenced by several recent studies. Some recent studies have achieved good results by adopting innovative methods. For instance, research by Nardilasari et al. addressed the low accuracy of the Naïve Bayes algorithm by replacing it with Support Vector Machine (SVM) in sentiment analysis of the 2024 presidential candidates on Twitter [1]. SVM is a machine learning method that can be used for classification and regression. Operates by identifying the optimal hyperplane that can best separate two classes within the feature space [7]. Similarly, another study exploring various word embedding methods such as GloVe, Word2Vec, BERT, and FastText, demonstrated that the choice of embedding method affects the accuracy of sentiment analysis, with GloVe embedding achieving the highest accuracy of 87.94% in sentiment analysis on the IMDB movie review dataset [8]. GloVe, short for Global Vectors, is a word representation method in vector form used in natural language processing. GloVe is designed to capture the meaning of words based on the statistics of their occurrence in a text corpus [9]. Unlike other methods such as Word2Vec, GloVe integrates global statistical information from the entire corpus, not just relying on local context information of the words [8], [10].

Several related studies have also garnered attention, including research on Covid-19 vaccination sentiment on Twitter using the SVM method with Word2Vec feature extraction. This study improved performance by about 4% compared to using TF-IDF features [11]. There is also sentiment analysis on the projection of the 2024 presidential election using SVM, which achieved high accuracy and provided an accurate picture of public support and views towards the 2024 presidential candidates [12]. Research by Xiaoyan et al. explored the GloVe-CNN-BiLSTM model

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

for sentiment analysis on diverse Twitter texts, demonstrating high accuracy in classifying sentiments across complete, long, and short texts [9].

Given the results achieved by previous studies, this research aims to implement the SVM algorithm to analyze public sentiment towards presidential candidates in the 2024 election on social media X. The objectives are to evaluate the performance of the SVM algorithm in classifying positive and negative sentiments towards the candidates, and to compare the effectiveness of GloVe word embedding and TF-IDF methods in improving sentiment analysis accuracy. This research will focus on evaluating and comparing the performance of SVM with GloVe word embedding and TF-IDF methods in sentiment analysis. The comparison is expected to reveal which method provides more accurate results in depicting public sentiment towards presidential candidates in the 2024 election in Indonesia. The limitations of this research include focusing on sentiment analysis in the Indonesian language, using data from social media X, and tweets posted from January 2023 to December 2023, leading up to the 2024 presidential election. Thus, this research will explore in-depth the potential of GloVe and TF-IDF in enhancing the performance of sentiment analysis in this specific political context.

2. RESEARCH METHODOLOGY

2.1 Research Stages

This study will proceed by following these steps: crawling, labeling, preprocessing, feature extraction using TF-IDF and GloVe, splitting data into training and testing sets, model training with SVM, hyperparameter tuning, and evaluation. The research scheme can be seen in Figure 1.

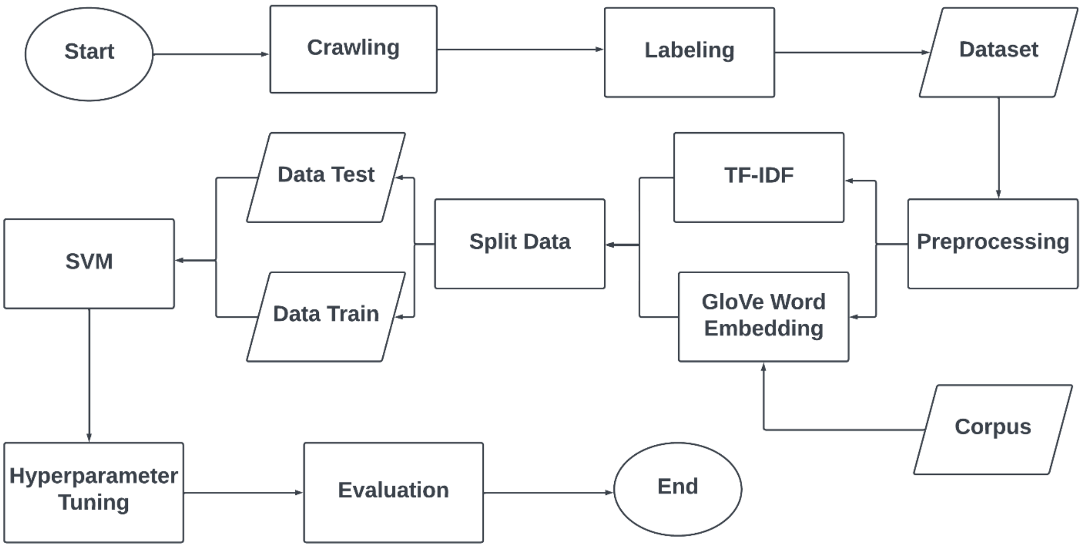

Figure 1. Research Stages

2.2 Data Collection

Data collection for this research was conducted through a crawling process utilizing the Application Program Interface (API) X provided by social media X. The collected data revealed that the keyword Ganjar Pranowo had 5,863 tweets, Prabowo Subianto had 5,988 tweets, and Anies Baswedan had 5,978 tweets. Data collection was carried out in stages due to the limitations imposed by social media X, which allows each account to retrieve no more than 1,000 tweets per day.

2.3 Data Labeling

Sentiment labeling of the collected data was performed manually and divided into two main classes: positive and negative. Labeling was conducted by 3 annotators, and the most frequent label among them was chosen to reduce subjectivity in the labeling process.To assign a positive sentiment label, annotators were guided to label texts that clearly expressed support, satisfaction, or a positive view of the presidential candidate. Conversely, negative sentiment was labeled for texts expressing disagreement, dissatisfaction, or a negative view of the presidential candidate.

2.4 Pre-processing

Data pre-processing was conducted to transform unstructured text data into a format suitable for analysis. The pre-processing stages included several key steps:

a. Text Cleaning: Texts were cleaned of irrelevant elements such as URLs, emojis, and other symbols not necessary for text classification. This step aimed to focus the text on its main content.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

b. Case Folding: This process converted all letters in the text to lowercase to standardize letter forms, ensuring no differentiation between uppercase and lowercase letters, which could hinder sentiment analysis.
c. Tokenization: Texts were divided into small units called tokens, typically referring to individual words. Tokenization is crucial for text analysis as it allows the calculation of word frequency and contextual analysis of words.
d. Stop Word Removal: Unimportant words, such as conjunctions, were removed using the Sastrawi library as a guide for Indonesian stop words.
e. Stemming: This step simplified words to their base forms by removing affixes. The Sastrawi library was also used as a reference for base words in the Indonesian language.

2.5 TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF is a statistical method used to assess the significance of a word within a document compared to a collection of documents (corpus) [10] . In this research, TF-IDF was applied to convert the pre-processed text data into numerical representations. This method involves two main components: Term Frequency (TF) reflects the frequency of a word within a document, while Inverse Document Frequency (IDF) assesses how rare that word is throughout the entire corpus [7], [10]. The TF-IDF score for each word is computed by multiplying TF and IDF, emphasizing words that are common in a particular document but rare across the whole corpus. This transformation enables the subsequent machine learning models to effectively differentiate between texts with positive and negative sentiments based on the weighted importance of words.

The formula for TF-IDF is captures in the formula (1):

$$TF - IDF(t,d) = TF(t,d) \times IDF(t,d) \tag{1}$$

$TF(t,d)$ is the is the term frequency of term t in document d, calculated as formula (2):

$$TF(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \tag{2}$$

$IDF(t,d)$ is the is the inverse document of term t, calculated as formula (3):

$$IDF(t) = \log(\frac{N}{|\{d \in D : t \in d\}|}) \tag{3}$$

Here, $f_{t,d}$ represents the frequency of term t in document d, N represents the total number of documents in the corpus, and $|\{d \in D : t \in d\}|$ indicates the count of documents that include the term t.

2.6 GloVe Word Embedding

The next step was to apply the GloVe Word Embedding technique. This process began with building a vocabulary from the collected text corpus, gathering unique words, and counting their frequencies to determine vector weights [9]. Each sentence from social media X was then converted into vectors by considering vector dimensions and algorithm speed. The co-occurrence probability of two words in context was calculated, considering global statistical information [8], [10].

The core idea of GloVe is captured in the formula (4):

$$J = \sum_{i,j=1}^{V} f\left(X_{ij}\right)(w_i^T \tilde{w}_j + b_i + \tilde{b}_j - \log(X_{ik}))^2 \tag{4}$$

The GloVe model's core components work together to capture word relationships within a text. X_{ij} represents the co-occurrence frequency of word j in the context of word i, providing a measure of how often these words appear together. The vectors w_i and $\widetilde{w}_j$ represent the semantic meanings of word i and its context word j, enabling the model to understand their relationships in a vector space. Bias terms b_i and $\tilde{b}_j$ account for any systematic biases in the words and their contexts, refining the model's predictions. Finally, the logarithm of the co-occurrence frequency, $\log(X_{ik})$, normalizes the data, making the model's learning process more stable and effective. These components collectively enable GloVe to create meaningful vector representations of words based on their contextual usage.

Optimization methods such as gradient descent were used to adjust the word weight matrix, improving the quality of vector representation. The final output was a vector representation for each word in the vocabulary, which was then used for sentiment classification using SVM (Table 1).

Table 1. Sample of GloVe Representative Vector

Word	GloVe Representative Vector
prabowo	[0.23, 0.45, -0.12, ...]
subianto	[-0.56, 0.78, 0.34, ...]
hubung	[0.09, -0.67, 0.21, ...]
khusus	[0.65, 0.43, -0.98, ...]
…	

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

2.7 Support Vector Machine

Following the application of GloVe Word Embedding, the next step was to use SVM for sentiment classification. SVM is a method utilized for both classification and regression tasks, designed to identify the best hyperplane that can divide two classes within the feature space [7], [12]. The SVM framework includes two classes, +1 and -1, separated by this hyperplane. The margin, which is the distance between the hyperplane and the support vectors (the nearest data points to the hyperplane), defines how close the data points are to the hyperplane. The solid line in the middle is the optimal hyperplane that maximizes this margin, as shown in Figure 2.

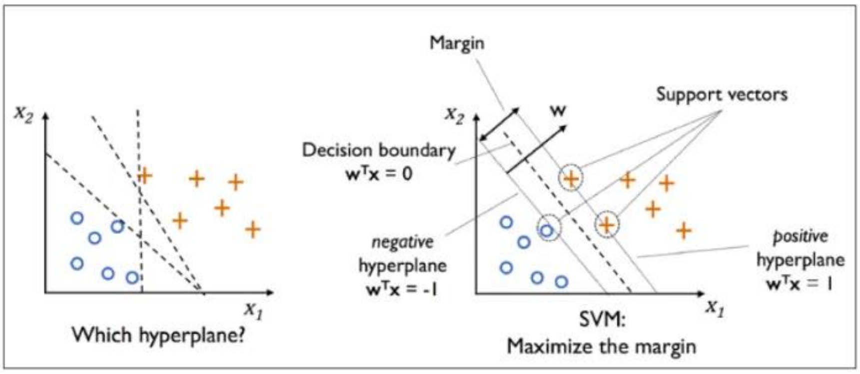

Figure 2. SVM Structure

The dashed lines indicate the best hyperplane that lies in the middle of the two classes. The mathematical equation to find the hyperplane is shown in (5).

$$w \cdot x_i + b = 0 \tag{5}$$

where (w) is the weight vector, (x) is the input variable, and (b) is the bias term. For class +1, the equation is shown in (6).

$$w \cdot x_i + b \geq 1 \tag{6}$$

and for class -1, the equation is shown in (7).

$$w \cdot x_i + b \leq -1 \tag{7}$$

The advantage of SVM lies in its ability to use kernels to separate non-linear inputs in high dimensions. SVM provides several types of kernels such as Radial Basis Function (RBF), polynomial, and linear, allowing the model to handle complex patterns and non-linear relationships in the data [1]. The training data, which had been converted into vectors through GloVe, was used to train the SVM model, with parameter adjustments such as C and kernel for optimization.

Specifically, hyperparameter tuning was conducted using a grid search to explore different values for the penalty parameter (C), the kernel coefficient (gamma), and the type of kernel (linear or RBF). After training, the SVM model could predict the sentiment of new text data by inputting word vector representations from GloVe, resulting in sentiment classification such as positive and negative based on the training data.

2.8 Evaluation with Confusion Matrix

The classification results from the SVM model were evaluated using a Confusion Matrix, which shows the extent to which the model correctly classified sentiments. To evaluate the model's performance in sentiment analysis, metrics such as accuracy, precision, recall, and F1-score were computed [1]. This evaluation provided insights into the model's effectiveness in classifying text data.

3. RESULT AND DISCUSSION

3.1 Data Preparation

The first step is to split the dataset into training and testing subsets, using the 'train_test_split' function with 'test_size' set to 0.2 and 'random_state' set to 42. This approach allocates 20% of the data for testing the model's performance, while the remaining 80% is used for training. This partitioning is essential for evaluating how well the model generalizes to new, unseen data, helping to prevent overfitting.

3.2 Text Representation Techniques

In the analysis, two primary text representation techniques are utilized: TF-IDF Vectorization and GloVe Embeddings. TF-IDF Vectorization is implemented using the TfidfVectorizer, which transforms text data into TF-IDF vectors by

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

calculating word weights based on their frequency in individual documents relative to their frequency across the entire dataset. This method highlights words that are more significant for distinguishing sentiment, thereby enhancing the model's ability to capture relevant information. Conversely, GloVe Embeddings involve using pre-trained embeddings loaded through the load_glove_embeddings function. Sentences are represented by averaging the GloVe vectors of their constituent words, a process facilitated by the apply_glove_embeddings function. This approach results in fixed-size vector representations for each sentence, capturing global semantic relationships and providing a different perspective on text data.

3.3 Model Training and Optimization

SVM classifier is employed for sentiment classification, utilizing Grid Search Cross-Validation (GridSearchCV) to optimize its hyperparameters and enhance model performance. Key hyperparameters tuned include the Penalty Parameter (C), which regulates the balance between minimizing errors on the training data and controlling model complexity, with values tested ranging from 0.1 to 100. The Kernel Coefficient (gamma) is another critical parameter, defining the extent of influence a single training example has, with values tested including 1, 0.1, 0.01, and 0.001. Additionally, the choice of Kernel Type is essential for determining the form of the decision boundary, with options including linear for data that is linearly separable and rbf for handling non-linear data relationships. Following the hyperparameter tuning, the model's performance is evaluated using various metrics such as accuracy, F1 score, precision, and recall, providing a comprehensive assessment of its effectiveness in sentiment classification.

3.4 Result

Sentiment analysis of three Indonesian political figures, namely Anies, Ganjar, and Prabowo, reveals a diverse range of public perceptions. This data provides insights into how the public evaluates the performance, policies, and image of each figure based on the sentiments they express as shown in Figure 3. The analysis results show that Anies receives a predominantly positive sentiment, with around 5,000 positive data points compared to only about 1,000 negative data points. This dominance of positive sentiment indicates that Anies's policies and actions are generally well-received by the public. Factors such as the success of his work programs and a positive public image may contribute to the high level of positive sentiment.

Ganjar shows a sentiment pattern similar to Anies, with approximately 4,800 positive sentiment data points and 1,000 negative sentiment data points. Although slightly lower than Anies, the high positive sentiment towards Ganjar indicates that the public has a positive perception of his performance and personality. This can be attributed to the success of public policies and an effective communicative approach.

In contrast, Prabowo has a more balanced distribution of positive and negative sentiments. Positive sentiment towards Prabowo amounts to around 3,500 data points, while negative sentiment reaches around 2,500 data points. This indicates polarization in public perception of Prabowo. The high level of negative sentiment may be due to political controversies or unpopular policies among certain segments of the public. This polarization reflects the challenges Prabowo faces in building a consistent and widely accepted image among the public.

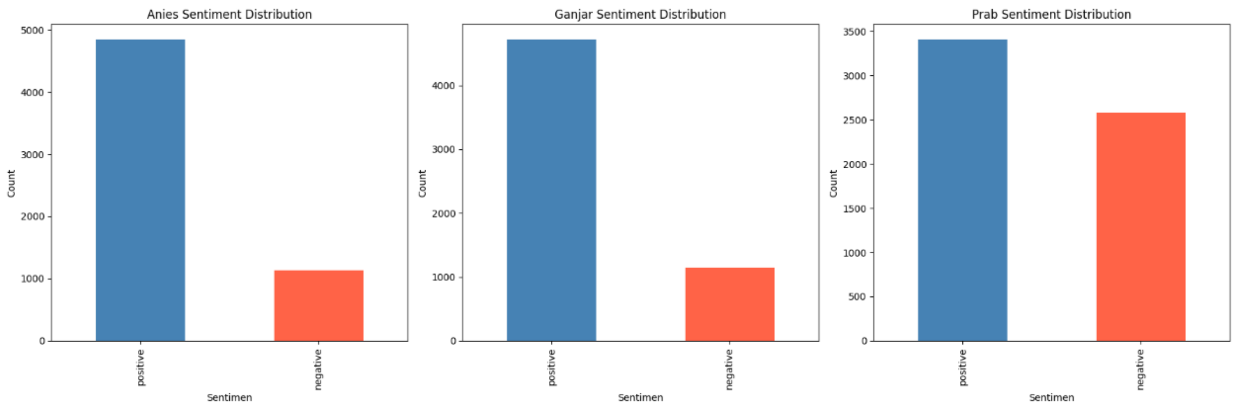

Figure 3. Comparison of Sentiment Distribution

This evaluation provided insights into the effectiveness of each method in handling sentiment data for the three political figures: Anies, Ganjar, and Prabowo (Figure 4). For Anies's dataset, the model using TF-IDF showed an accuracy of 0.84 with an F1 Score of 0.80, precision of 0.85, and recall of 0.84. In contrast, using GloVe, the accuracy was slightly lower at 0.82 with an F1 Score of 0.73, precision of 0.67, and recall of 0.82. These results indicate that TF-IDF is more effective in capturing relevant information for sentiment analysis in the case of Anies. The higher precision of TF-IDF suggests that it is better at correctly identifying positive sentiments, which is crucial for understanding the nuances in public perception. Additionally, the balanced recall indicates TF-IDF's ability to capture a comprehensive range of sentiments, contributing to a more holistic analysis (Table 2).

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

Table 2. Performance Results for Anies

Method	Accuracy	F1 Score	Precision	Recall
TF_IDF	0.84	0.80	0.85	0.84
GloVe	0.82	0.73	0.67	0.82

For Ganjar's dataset, the TF-IDF model also performed better, achieving an accuracy of 0.80, an F1-Score of 0.73, precision of 0.79, and recall of 0.80. Meanwhile, the model using GloVe showed the same accuracy but with a slightly lower F1-Score of 0.71, precision of 0.64, and recall of 0.80. This demonstrates that TF-IDF provides better results in terms of F1-Score and precision for Ganjar's dataset. The superior precision of TF-IDF implies a higher rate of true positive sentiment detection, which is particularly important in sentiment analysis for political figures where the precision of sentiment identification can impact strategic decisions. The consistency in recall between TF-IDF and GloVe indicates that both methods are equally effective in detecting the range of sentiments but differ in the accuracy of those detections (Table 3).

Table 3. Performance Results for Ganjar

Method	Accuracy	F1-Score	Precision	Recall
TF_IDF	0.80	0.73	0.79	0.80
GloVe	0.80	0.71	0.64	0.80

For Prabowo's dataset, both techniques showed excellent performance. The model using TF-IDF had an accuracy of 0.93, an F1-Score of 0.93, precision of 0.93, and recall of 0.94. Similarly, the model using GloVe also showed an accuracy of 0.93, an F1-Score of 0.92, precision of 0.93, and recall of 0.93. Although the performance of both models was nearly equal, TF-IDF had a slight advantage in recall. This slight edge in recall for TF-IDF indicates its slightly better ability to identify all relevant instances of sentiment, ensuring that fewer sentiments are missed. This is particularly valuable in political sentiment analysis, where capturing the full spectrum of public opinion is critical for accurate representation and strategic planning (Table 4).

Table 4. Performance Results for Prabowo

Method	Accuracy	F1-Score	Precision	Recall
TF_IDF	0.93	0.93	0.93	0.94
GloVe	0.92	0.92	0.93	0.93

Additionally, an evaluation was conducted using TF-IDF and presented as graphic shown in Figure 5. For Anies's dataset, the model using TF-IDF showed an accuracy of 0.84, higher compared to the model using GloVe, which had an accuracy of 0.82. This difference indicates that TF-IDF is more effective in analyzing sentiment in Anies's case. TF-IDF can capture more relevant and specific information, which is crucial in a political context, resulting in better performance than GloVe.

For Ganjar's dataset, both techniques showed the same accuracy, which was 0.80. However, when looking at other metrics such as F1-Score and precision, TF-IDF still outperformed GloVe. This shows that even though the overall accuracy is the same, TF-IDF is better at balancing precision and recall, resulting in a higher F1-Score. This highlights TF-IDF's superiority in capturing finer and more complex sentiment nuances in Ganjar's sentiment data.

For Prabowo's dataset, both techniques demonstrated similarly high accuracy, with the TF-IDF model achieving an accuracy of 0.93 and GloVe slightly lower but still at 0.93. Nevertheless, TF-IDF showed a slight advantage in recall, meaning that TF-IDF is better at identifying all instances of a particular class in the dataset. This shows that TF-IDF not only provides accurate results but also more consistently identifies sentiment comprehensively.

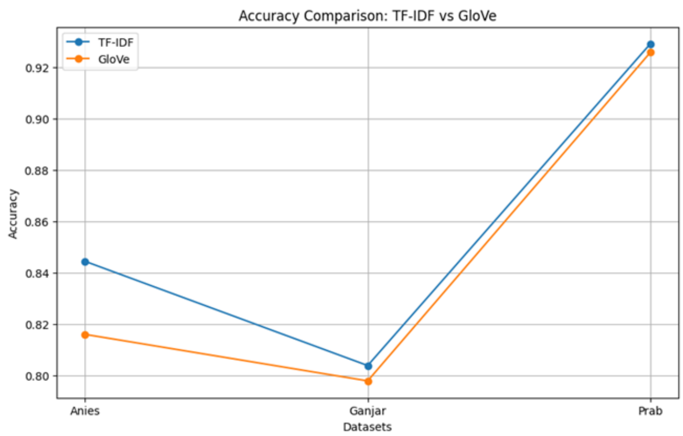

Figure 5. Result of Accuracy Comparison between TF-IDF and GloVe

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

Based on these results, it is evident that TF-IDF generally outperforms GloVe in terms of accuracy. TF-IDF is better at capturing specific contexts, which is crucial in political sentiment analysis, while GloVe focuses more on global semantic meaning.

3.5 Discussion

Fundamentally,sentiment analysis techniques are crucial for categorizing and interpreting user's emotions regarding various topics.. However, Sentiment analysis methods[13], [14] fundamentally assist in classifying and understanding user feelings about topics of interest. However, analyzing social media events is challenging [8], [9] due to the intricate nature of socio-technical systems and the large-scale data characteristics of complex networks [15], [16]. This complexity is highlighted in real-world political campaign case studies, which are particularly valuable for understanding human behavior, detecting patterns, and identifying common strategies for analyzing user interactions in online social networks. This is also reflected in the sentiment analysis results of three Indonesian political figures, namely Anies, Ganjar, and Prabowo, which show significant variations in public perception. From the sentiment distribution results, it is evident that Anies and Ganjar receive the majority of positive sentiments, while Prabowo faces more polarization with sentiments balanced between positive and negative. These differences in sentiment distribution reflect how the policies, actions, and image of each political figure influence public views.

Additionally, sentiment analysis is recognized for its ability to uncover public opinion by utilizing representations, models, and algorithms that can convert "simple unstructured text" to "detailed insights" [17], [18]. This is reflected in the analysis conducted on approximately 17,800 unstructured text data obtained through the Social media X and processed with sentiment analysis, which provided the insight that Anies and Ganjar tend to receive greater support from Twitter users, while Prabowo faces more critical and polarized public opinions. This sentiment analysis is essential for grasping the dynamics of public opinion and can be used to formulate more effective communication strategies for these political figures. However, it should be understood that conducting sentiment analysis for the political domain is a challenging task. It is said that some background information about the authors is required to determine the true meaning of the comments, such as political party affiliation, political standpoint (according to general taxonomy, such as right vs. left), and so forth [19].

In the context of evaluating sentiment analysis models, the researchers used two different text representation techniques, namely TF-IDF and GloVe, providing important insights into the effectiveness of each method. In the Anies dataset, the model using TF-IDF showed an accuracy of 0.84, higher than the GloVe model which achieved an accuracy of 0.82. This indicates that TF-IDF is more effective in capturing relevant information for sentiment analysis in Anies's case. The ability of TF-IDF to focus on the specific context of words contributes to its better performance. Similarly, in the Ganjar dataset, the TF-IDF model showed an accuracy of 0.80. Although the GloVe model also achieved the same accuracy, other metrics such as F1-Score and precision were slightly lower in GloVe. This indicates that TF-IDF not only matches GloVe in overall accuracy but also excels in balancing precision and recall, resulting in a higher F1-Score. This confirms the superiority of TF-IDF in capturing more subtle and complex sentiment nuances in Ganjar's sentiment data.

Meanwhile, in the Prabowo dataset, both techniques showed very high and nearly equal performance, with an accuracy of about 0.93. However, TF-IDF showed a slight advantage in recall, meaning that TF-IDF was better at identifying all instances of a particular class in the dataset. This indicates that TF-IDF not only provides accurate results but is also more consistent in identifying sentiments comprehensively. These evaluation results show that the choice of text representation techniques greatly influences the performance of sentiment analysis models. In general, TF-IDF is superior or equal to GloVe in terms of accuracy, F1-Score, and precision. The superiority of TF-IDF lies in its ability to capture the specificity of words in a particular context, which is crucial in political sentiment analysis. Conversely, GloVe, which focuses more on global semantic representation, shows suboptimal performance in some metrics.

These evaluation results demonstrate that the choice of text representation techniques significantly affects the performance of sentiment analysis models. Overall, TF-IDF is superior or equal to GloVe in terms of accuracy, F1-Score, and precision. According to research [20], the advantage of TF-IDF lies in its ability to capture the specificity of words in a particular context, which is crucial in political sentiment analysis. This can also be attributed to TF-IDF's ability to assign more weight to unique and important words in a document, which is highly relevant in political sentiment analysis. Conversely, GloVe, which focuses more on global semantic representation, shows suboptimal performance in some metrics [20]. This weakness may be due to its inability to capture the specific contexts often required in detailed and nuanced sentiment analysis. Ultimately, these findings indicate that while GloVe has the ability to capture global semantic relationships, TF-IDF is more effective in identifying variations and nuances in diverse sentiment data.

Furthermore, these findings highlight the importance of choosing the right text representation technique based on the specific requirements of sentiment analysis tasks. TF-IDF's ability to assign higher weights to unique and relevant terms makes it particularly effective in contexts where specific word choices are significant, such as political discourse. On the other hand, GloVe's focus on capturing global semantic relationships may not be as advantageous in scenarios where the specificity of language plays a crucial role. The comparative analysis also suggests that while GloVe's semantic embeddings provide a strong baseline, TF-IDF's contextual weighting can lead to more accurate

ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

and nuanced sentiment classifications. This is evident from the higher F1-Scores and precision rates achieved by TF-IDF across multiple datasets. Consequently, for applications in political sentiment analysis, where the precision and recall of sentiment detection are paramount, TF-IDF appears to offer a more reliable approach.

Overall, these findings indicate that while GloVe has the ability to capture global semantic relationships, TF-IDF is more effective in identifying variations and nuances in diverse sentiment data. TF-IDF can be a more effective choice for political sentiment analysis in Indonesia, providing more consistent and accurate results. The advantage of TF-IDF in capturing specific contextual nuances makes it a more reliable tool for understanding public perception of political figures. This insight is crucial for developing more targeted and responsive communication strategies to public opinion. In a political context where public opinion can be very diverse and dynamic, the ability to capture sentiment nuances with high accuracy is invaluable. Therefore, using TF-IDF in political sentiment analysis in Indonesia is recommended to obtain a more accurate and reliable picture of public perception.

4. CONCLUSION

This study aims to implement the SVM algorithm and evaluate the combination of SVM and GloVe Word Embedding methods in sentiment analysis of the 2024 presidential candidates on social media X. The sentiment analysis results show that Anies receives predominantly positive sentiment with around 5,000 positive data points and 1,000 negative data points. Ganjar also shows predominantly positive sentiment with approximately 4,800 positive data points and 1,000 negative data points. In contrast, Prabowo has a more balanced sentiment distribution with about 3,500 positive data points and 2,500 negative data points. The evaluation of models using TF-IDF and GloVe shows that TF-IDF generally outperforms in terms of accuracy, F1-Score, and precision. For Anies and Ganjar's datasets, TF-IDF shows better results, while for Prabowo's dataset, both techniques perform almost equally, though TF-IDF has a slight edge in recall. Overall, TF-IDF is more effective in capturing the specific context, which is crucial in political sentiment analysis, making it a more reliable choice for understanding public perceptions of political figures. This insight is important for developing more targeted and responsive communication strategies to public opinion. Therefore, using TF-IDF in political sentiment analysis in Indonesia is recommended to obtain a more accurate and reliable picture of public perception.

REFERENCES

[1] S. S. H. T. A. Putri Nardilasari A. Lia Hananto and B. Priyatna, “Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter,” *Journal of Information Technology and Computer Science*, vol. 8, no. 1, pp. 11–18, 2023, doi: 10.31328/jointecs.v8i1.4265.

[2] E. Cano-Marin, S. Sánchez-Alonso, and M. Mora-Cantallops, “Unleashing Competitive Intelligence: News Mining Analysis on Technology Trends and Digital Health Driving Healthcare Innovation,” *IEEE Trans Eng Manag*, pp. 1–15, Nov. 2023, doi: 10.1109/TEM.2023.3326233.

[3] A. S. Y. I. Salim Puad Garno, “Analisis Sentimen Masyarakat Pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naïve Bayes,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, Jun. 2023.

[4] W. Kurnia, “Sentimen Analisis Aplikasi E-Commerce Berdasarkan Ulasan Pengguna Menggunakan Algoritma Stochastic Gradient Descent,” vol. 4, no. 1, pp. 138–143, 2023, doi: 10.33365/jtsi.v4i2.2561.

[5] C. F. Hasri and D. Alita, “Penerapan Metode Naïve Bayes Classifier dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona di Twitter,” *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, vol. 3, no. 2, pp. 145–160, 2022, doi: 10.33365/jatika.v3i2.2026.

[6] Herwinsyah and A. Witanti, “Analisis Sentimen Masyarakat Terhadap Vaksinasi COVID-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM),” *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 5, pp. 2622–6901, 2022, doi: 10.47080/simika.v5i1.1411.

[7] S. Khairunnisa A. Adiwijaya and S. Al Faraby, “Pengaruh Text Preprocessing Terhadap Analisis Sentimen Komentar Masyarakat Pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.

[8] P. Shah S. Shah and S. Joshi, “A Study of Various Word Embeddings in Deep Learning,” in *2022 3rd International Conference for Emerging Technology, INCET 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/INCET54531.2022.9824963.

[9] L. Xiaoyan R. C. Raga and S. Xuemei, “GloVe-CNN-BiLSTM Model for Sentiment Analysis on Text Reviews,” *J Sens*, 2022, doi: 10.1155/2022/7212366.

[10] J. Ravi and S. Kulkarni, “Text Embedding Techniques for Efficient Clustering of Twitter Data,” *Evol Intell*, vol. 16, no. 5, pp. 1667–1677, Oct. 2023, doi: 10.1007/s12065-023-00825-3.

[11] M. Sahbuddin and S. Agustian, “Support Vector Machine Method with Word2vec for Covid-19 Vaccine Sentiment Classification on Twitter,” *Journal of Informatics and Telecommunication Engineering*, vol. 6, no. 1, pp. 288–297, Jul. 2022, doi: 10.31289/jite.v6i1.7534.

[12] A. Azzawagama Firdaus A. Yudhana and I. Riadi, “Analisis Sentimen Pada Proyeksi Pemilihan Presiden 2024 Menggunakan Metode Support Vector Machine,” *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 236–245, Jun. 2023, doi: 10.51454/decode.v3i2.172.

Volume 6, No 2, September 2024, Page: 961−969
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5668

[13] F. Aftab *et al.*, "A Comprehensive Survey on Sentiment Analysis Techniques," *International Journal of Technology*, vol. 14, no. 6, pp. 1288–1298, 2023, doi: 10.14716/ijtech.v14i6.6632.
[14] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.
[15] G. Lombardo, M. Mordonini, and M. Tomaiuolo, "Adoption of social media in socio-technical systems: A survey," *Information (Switzerland)*, vol. 12, no. 3, Mar. 2021, doi: 10.3390/info12030132.
[16] Z. Stojanov, J. Stojanov, G. Jotanovic, and D. Dobrilovic, "Weighted networks in socio-technical systems: Concepts and challenges," 2020.
[17] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi, "Text mining in big data analytics," *Big Data and Cognitive Computing*, vol. 4, no. 1, pp. 1–34, Mar. 2020, doi: 10.3390/bdcc4010001.
[18] J. Li, A. Media, S. Thabit Albarhami, and K. Kyritsis, *Advances in Sentiment Analysis - Techniques, Applications, and Challenges*, vol. 22. in Artificial Intelligence, vol. 22. IntechOpen, 2024. doi: 10.5772/intechopen.111293.
[19] E. Kubin, C. Puryear, C. Schein, and K. Gray, "Personal experiences bridge moral and political divides better than facts," *PNAS*, vol. 118, no. 6, 2021, doi: 10.1073/pnas.2008389118/-/DCSupplemental.
[20] Y. Didi A. Walha and A. Wali, "COVID-19 Tweets Classification Based on a Hybrid Word Embedding Method," *Big Data and Cognitive Computing*, vol. 6, no. 2, p. 58, 2022, doi: 10.3390/bdcc6020058.