Aspect-Based Sentiment Analysis from User-Generated Content in Shopee Marketplace Platform

Andharini Dwi Cahyani
Informatics Engineering, Faculty of Engineering, Universitas Trunojoyo Madura, Indonesia

ARTICLE INFO

Article history:

Received May 06, 2023
Revised June 14, 2023
Published June 17, 2023

Keywords:

Aspect-Based Sentiment Analysis;
Customer review;
Marketplace;
User-generated content;
Fashion merchant

ABSTRACT

A number of businesses, such as TripAdvisor, Open Table, and Yelp, have successfully utilized aspect-based sentiment analysis in order to gain insights from reviews provided by customers and enhance the quality of their goods or services. Businesses are able to swiftly discover any unfavorable sentiment or possible harm to their brand when they analyze client input across numerous aspects from social media, online reviews, and conversations with customer care representatives. This study aims to explain how aspect-based semantic analysis of market-collected user-generated data through performance comparisons of Doc2vec and TF-IDF vectorization. Both Doc2Vec and TF-IDF have their own distinctive qualities, which might vary according on the nature of the job, the dataset, and the volume of the available training data. For the objectives of this research, the data was obtained from several of fashion merchants that run their companies by means of the Shopee platform, which is a well-known online marketplace platform in Indonesia. In this research, the accuracy and F1 Score achieved by Doc2Vec vectorization was superior to those achieved by TF-IDF vectorization. Our findings shows that Doc2Vec vectorization is better for classifying customer ratings because it can pull out the semantic meaning of words in a document. The findings also shows that the score of c and gamma parameter have significant impact to the score of Accuracy and F1 Score of the classifier.By precisely categorizing client sentiment, this study enables businesses to improve their services, respond to customers' problems, and increase their customer satisfaction.

Corresponding Author:

Andharini Dwi Cahyani, Informatics Engineering, University of Trunojoyo Madura, Indomesia.
Email: andharini.cahyani@trunojoyo.ac.id

1. INTRODUCTION

User-generated content often includes feedback and suggestions for product improvements or new features [1]. By monitoring and analyzing customer reviews, businesses can identify areas where they are excelling and areas where improvements are needed. Addressing customer concerns [2] and improving their satisfaction levels can enhance customer loyalty, encourage repeat purchases, and attract new customers through positive word-of-mouth. Furthermore, businesses may get significant information from customer feedback by gathering sentiments about certain aspects. They are able to understand which aspects of a product or service a customer appreciates and which ones they find challenging, as well as identify specific areas of dissatisfactionand elements that contribute to overall satisfaction [3]. This data is utilized to improve the products, enhance the experiences of customers, and more effectively address concerns raised by customers [4].

Aspect-based sentiment analysis (ABSA) examines opinions at the aspect or feature level to provide a more nuanced understanding of sentiment [5]. Instead of classifying the entire text as positive, negative, or neutral, it identifies emotions directed at specific features of the target item [6]. Li *et al.* [7] reported that ABSA may assists businesses in identifying areas of strength and weakness in their products or services. They employed ABSA to predict the restaurant survival possibilities from online customer-reviews. Aspect-based sentiment analysis provides insight into consumer preferences and market trends [7], [8]. By analyzing sentiments across multiple dimensions [9], businesses can recognize emerging patterns, comprehend customer requirements [10], and gain a competitive edge [11]. Businesses are able to manage their reputations [12] through the monitoring of sentiments conveyed about various aspects. They are able to identify negative sentiments associated with particular features [13], [14] or aspects and to proactively address consumer concerns [15]. This contributes to the development and maintenance of a positive brand image, the enhancement of consumer trust, and the prevention of reputational harm [16].

Businesses can increase overall customer satisfaction, strengthen customer loyalty, and decrease customer attrition by addressing customer concerns and enhancing aspects that elicit negative sentiment [13]. Aspect-based sentiment analysis provides decision-making and strategy formulation with data-driven insights [7]. It assists organizations in making informed decisions regarding product positioning, marketing campaigns, consumer support, and resource allocation. Businesses can optimize business outcomes by aligning their strategies with consumer preferences using aspect-based sentiment analysis [4].

Aspect-based sentiment analysis is used by lodging and hotel companies [17], transportation companies, and tourism organizations [9], [18], [19] to examine social media and review sites like TripAdvisor for consumer reviews [20], [21]. They prioritize factors such as service quality, hygiene, location, and amenities. In addition, e-commerce and retail companies frequently employ aspect-based sentiment analysis to analyze consumer reviews and feedback [21]. Components may refer to a variety of features of a product (such as its performance, design, or pricing) or characteristics of a service (such as its customer support or delivery speed). The companies gain insight into consumer preferences, identify areas for refinement, and direct product development by extracting sentiments associated with various product features [22], [23]. This allows them to increase customer satisfaction, make data-driven decisions, and maintain market competitiveness [24].

The present research contributes to the understanding of vectorization comparisons carried out in aspect-based semantic analysis of user-generated information collected from marketplace. For the purpose of this research, the data was obtained from several fashion merchants on the Shopee platform, one of a popular marketplace in Indonesia.

1.1. User-Generated Content

Being aware of user-generated content and reviews enables businesses to gain insights into customer preferences [25], manage their reputation, improve their products and services [26], [27], increase customer satisfaction [28], remain competitive, and utilize customer feedback effectively in their marketing and sales efforts [29], [30]. Customer evaluations and other user-generated content provide valuable insight into the experiences, opinions, and emotions of actual users. Patel and Agrawal [31] suggested that by studying this material, organizations might acquire a more in-depth grasp of the requirements, preferences, and pain points of their respective consumers. This data can inform product development, marketing strategies, and consumer support initiatives [32]. Furthermore, user-generated content has a substantial effect on the reputation of a business or product. According to Giachanaou *et al.* [33] and Wu *et al.* [34], positive reviews can enhance a company's reputation and credibility, whereas negative reviews can have the opposite effect. Being aware of user-generated content enables businesses to monitor their online reputation, identify and promptly resolve any negative feedback or issues, and take the steps necessary to maintain a positive brand image [35].

According to Song *et al.* [36], user-generated content reflects consumer satisfaction levels. By monitoring and analyzing consumer evaluations, businesses can determine where they excel and where they need to make improvements. Addressing customer concerns and increasing their levels of satisfaction can increase customer loyalty, encourage repeat purchases, and attract new customers via positive word of mouth. Moreover, according to Arzaghi *et al.* [37], user-generated content frequently includes suggestions for product enhancements or new features. By actively listening to consumers through their evaluations, businesses can collect innovative ideas and prioritize product improvements based on customer requirements and expectations. This can result in improved products and higher consumer satisfaction.

Additionally, user-generated content provides insights into the products and services of competitors. By assessing consumer evaluations of competing products, businesses can determine where rivals excel and where they fall short [38]. This information is useful for benchmarking performance, identifying market opportunities, and effectively positioning their own products. User-generated content can be utilized in marketing and sales campaigns. Potential consumers can be won over by showcasing positive reviews and testimonials from existing customers [39]. User-generated content can also be a valuable source of user-generated marketing materials such as social media posts, videos, and testimonials that can be shared to increase brand visibility and engagement.

1.2. Aspect-based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) is a natural language processing (NLP) method that attempts to find and study written thoughts or feelings about certain parts or features of a target object [40], [41]. The task entails dividing the text into a number of smaller components, known as aspects or features, and determining the emotion about each one individually. ABSA's target object might be a product, service, event, or other interest. Aspects are the various components, qualities, or characteristics of the target object. For example, a review of a restaurant could talk about how good the food is, how friendly the staff is, how fun the setting is, how much it costs, and how clean the place is [7], [42].

Aspect-based sentiment analysis for marketplace products entails analyzing consumer evaluations or feedback pertaining to particular aspects or characteristics of products sold on an online marketplace. It seeks to gather insights about customer sentiment regarding various aspects of the product, thereby enabling vendors and marketplace administrators to comprehend customer opinions and make informed decisions [43].
The following illustrates how aspect-based sentiment analysis can be applied to market products [4]:

1. Aspect Recognition
   The first step is to find and record the features and traits that the customer stated. For example, if the product is a smartphone, important features might include battery life, photo quality, speed, style, user interface, etc. After figuring out what the parts are, the sorting of emotion is done to find out how people feel about each one. For each part, the customer reviews are looked at to see if they are good, bad, or neutral. This can be done with the help of machine learning techniques and ways based on dictionaries.
2. Aspect-based Analysis
   This step is the process of figuring out the sentiment scores for each aspect. These scores show how customers feel about each part of the product. This analysis helps businesses figure out what customers like and dislike about their goods or services.
3. Aspect-based Opinion Aggregation
   A product's overall sentiment is determined by summing the sentiment evaluations for its multiple features. This can be done by thinking about how important or relevant each factor is. For instance, if the quality of the camera is an important part of the product, its sentiment may be given more weight in the total product sentiment.
4. Insights and Determination
   Vendors and marketplace administrators can gain valuable insight from the outcomes of aspect-based sentiment analysis. They can figure out what needs to be improved, highlight the best things about a product, and solve specific problems that customers have with it. These findings can be used to help make business choices about product development, marketing, customer service, and other business decisions.

Aspect-based sentiment analysis for marketplace products enables vendors and marketplace administrators to comprehend the precise determinants of customer satisfaction and dissatisfaction. It enables them to enhance their products, provide superior consumer experiences, and make data-driven decisions to increase their market competitiveness.

2. METHODS

2.1. Data Collection

The data for this study will be derived from consumer reviews obtained from the Shopee marketplace for several fashion-related vendors in January 2023. The data was gathered using Instant Data Scraping techniques, yielding 855 records in total.
This study's selection of data was based on the following criteria:
1. Processed data consists of user evaluations as opposed to simple ratings. This investigation contained 126 rating data without reviews. Consequently, the data cannot be further processed for sentiment analysis.
2. The evaluated information is written in impeccable Indonesian. This pertains to the library used during the data pre-processing. The total number of data received after applying the aforementioned selection criteria was 729.

2.2. Supervised Vector Machine

SVM is a frequently used and successful method, according to Borg and Bolt [44], especially for situations with small to moderate-sized datasets and where interpretability of findings is critical. Support Vector Machines (SVMs) are a flexible and frequently used supervised learning method that handles classification and regression problems successfully by locating optimum decision boundaries in the feature space [45], [46]. SVM provides a number of benefits, including the capacity to handle high-dimensional data [47], resilience to outliers, and generalization ability. In addition, it is less prone to overfitting than certain other classification methods [48]. SVM, on the other hand, may be computationally demanding, particularly when working with huge datasets. SVM is highly successful for managing complicated datasets, according to Liu *et al.* [49] and Gao *et al.* [50], and it is extensively employed in a variety of disciplines, including text classification, picture recognition, and bioinformatics.

Here are some of the most important benefits of SVM, according to the literature [49] – [52]:

1. SVM works well even in places with a lot of dimensions, which makes it a good choice for datasets with a lot of features. It can deal with complex interactions and define complex choice limits. SVM does not overfit because it uses the idea of "maximum margin," which tries to find a hyperplane that divides classes by the biggest difference possible. Compared to algorithms like decision trees and neural networks, this method improves generalization and makes it less likely that the model will become too perfect.
2. When the number of training data is less than the number of variables, SVM works well. SVM looks at the support vectors, which are the data points that are closest to the decision limits. It finds the best way to separate these important points, so that unimportant or duplicate traits have less of an effect on SVM. SVM is not too affected by errors in the training data. The maximum margin idea tries to find a decision level that works well with unseen data and reduces the effect that individual errors have on the performance of the classification as a whole.
3. SVM has clear lines for making decisions, which makes its results easy to understand and explain. As support vectors are a key part of setting decision boundaries, they can show what factors affect classification choices.
4. SVM can control how complicated a model is by changing things like the regularization parameter (C) and kernel parameters This allows users to find a good mix between how simple the model is and how well it classifies.

SVM is built on solid mathematics, particularly convex optimization and statistical learning theory. This theoretical foundation ensures that SVM responses are well-defined and can be theoretically examined. SVM may not be the ideal solution for every problem, and huge datasets, processing demands, and random data might impair its performance. However, fine-tuning elements like the kernel function and regularization parameter may improve outcomes [53]. Therefore, SVM is still a popular classification and regression method due to its merits.

The following are the stages in the Support Vector Machine classification technique process:

1. Set the starting values of, c, epsilon, lambda, and gamma.
2. Enter training data that is based on the presence of keywords in a single phrase.
3. Use the [K]kernel function to get the dot product for each data set. The formula for the linear kernel function is as follows:

$$K(x.x') = \exp(-\gamma \mid |x - x'| \mid^2) \tag{1}$$

As kernel functions, linear functions are employed. To perform the AxA^T matrix multiplication, the data must first be transposed.

4. Use the following formula to compute the matrix:

$$D_{ij} = Y_i Y_j\left(K\left(X_i X_j\right) + \lambda^2\right) \tag{2}$$

Where D_{ij} is Element matrix ij, Y_i is Data class to i, Y_j is Data class to j, λ^2 is Theoretical bounds.

5. Use the following formula to get the error value:

$$Ei = \sum_{j=1}^{i} \alpha_j D_{ij} \tag{3}$$

6. Using the following formula, calculate the value of delta alpha:

$$\delta\alpha_i = min\,\{[\gamma(1 - Ei) - \alpha_i]\; C - \alpha_i\} \tag{4}$$

7. Use the following formula to get a new alpha value:

$$\alpha_i = \alpha_i + \delta\alpha_i \tag{5}$$

8. Use the following formula to get the bias value:

$$b = -\frac{1}{2}(w.x^+ + w.x^-) \tag{6}$$

9. The testing phase starts after the values of, w, and b are determined. To run the first test, use the kernel function to compute the dot product of all training data between data testing:

$$K(x.xi) = x.xi \tag{7}$$

Following that, the decision function is tested as follows:

$$f(x) = w.x + b \; atau \; f(x) = \sum_{i=1}^{m} \alpha_i x_i K(x, x_i) + b \tag{8}$$

2.3. Confusion Matrix

A confusion matrix, also known as an error matrix, is a table that displays the counts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions to offer a complete perspective of the performance of a classification model. It is often used to assess the performance of machine learning algorithms, especially in binary classification problems [54].

The confusion matrix is organized in a tabular format, with two rows and two columns containing the expected and actual class labels. Typically, it is structured as Table 1.

Table 1. Confusion Matrix

Actual Value	**Prediction Value**	
	Positive	**Negatif**
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

Each cell of the confusion matrix corresponds to a distinct classification outcome.

- True Positive (TP): When the actual class was positive, the model correctly predicted the positive class.
- True Negative (TN): When the actual class was negative, the model correctly predicted the negative class.
- False Positive (FP): The model inaccurately predicted a positive class when the class in question was negative.
- False Negative (FN): When the actual class was positive, the model inaccurately predicted a negative class.

Various performance metrics can be derived from the values in the confusion matrix to assess the model's accuracy, precision, recall (also known as sensitivity or true positive rate), specificity (true negative rate), F1 score, and other evaluation measures.

The F1 score is a popular way to measure how well a classification model works [55]. It takes the accuracy of a model's precision and memory and adds them together into a single number. The F1 score is the harmonic mean of the model's accuracy and memory, and it runs from 0 to 1 [56]. A higher number means that the model is doing better.

The following method is used to figure out the F1 score:

$$F1\ Score \;=\; 2(Precision \;\times\; Recall) \,/\, (Precision \;+\; Recall) \tag{9}$$

Precision is the number of correctly expected positive events out of all the positive events that were predicted. It checks how well the model can avoid making fake positive mistakes. Precision is found by dividing the number of true positives by the number of true positives plus the number of fake positives.

Recall, which is also called sensitivity or true positive rate, is the number of times a positive event was correctly forecast out of all the real positive events. It checks how well the model can find all positive cases while avoiding fake negatives. Recall is determined by dividing True Positives by True Positives plus False Negatives.

The F1 score is crucial for a variety of factors, for some reasons below:

1. The F1 score achieves a balance between recall and accuracy [56]. Recall is how well the model can identify all positive cases while precision measures how well positive predictions can be produced. The F1 score considers both variables, making it applicable in circumstances where false positives and false negatives have differing implications.

2.Accuracy might not be the best metric to use to assess how well the model performs if the dataset is unbalanced [55], which occurs when one class has much more samples than the other. The F1 score is more accurate in these situations since it considers both accuracy and memory.

3.When comparing models, the F1 score offers a foundation for analysing the possibilities in an unbiased manner. Models with a higher F1 score are typically considered to have superior performance when the dataset is unbalanced [57] or the cost of false positives and false negatives vary.

Although useful, the F1 score may not be the greatest predictor in all situations. Domain-specific rating measures, for example, may be superior in cases when the cost of false positives and false negatives is highly different. However, taking into consideration both accuracy and recall, the F1 score is an effective and often used method of gauging how well a classification model performs [58].

3. RESULTS AND DISCUSSION

The vectorization techniques Doc2Vec and TF-IDF were examined in this study. It is common practise to represent and analyse text using the natural language processing (NLP) techniques Doc2Vec and TF-IDF (Term Frequency-Inverse Document Frequency). The performance of the model was evaluated using five-fold cross validation after the SVM classifier had been deployed. In Table 2 and Table 3, the accuracy and F1 Score of the SVM classification using TF-IDF and Doc2Vec vectorization are displayed.

Table 2. TF IDF vectorization and SVM for ABSA

Parameter	Size Aspect		Material Aspect		Delivery Aspect		Overall	
	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score
gamma= 0.01; c=1	0.65	0.47	0.60	0.41	0.68	0.44	0.64	0.44
gamma= 0.01; c=10	0.68	0.52	0.63	0.49	0.71	0.49	0.67	0.50
gamma= 0.01; c=100	0.72	0.61	0.67	0.55	0.75	0.58	0.71	0.58
gamma= 0.1; c=1	0.60	0.42	0.55	0.36	0.63	0.39	0.59	0.39
gamma= 0.1; c=10	0.63	0.47	0.58	0.41	0.66	0.44	0.62	0.44
gamma= 0.1; c=100	0.69	0.56	0.64	0.50	0.72	0.53	0.68	0.53
gamma= 1; c=1	0.60	0.42	0.55	0.40	0.63	0.39	0.59	0.41
gamma= 1; c=10	0.63	0.47	0.58	0.41	0.66	0.43	0.62	0.44
gamma= 1; c=100	0.69	0.58	0.64	0.52	0.70	0.55	0.68	0.55

The TF-IDF gives higher weights to terms that appear frequently in a single document but less frequently over the entire corpus. It prioritises term frequency and document frequency. It doesn't record word or document context or semantic relationships. A statistical measure called TF-IDF evaluates a term's importance in a document in relation to a broader corpus of documents. Based on the frequency of terms (TF) within each document and their inverse document frequency (IDF) over the entire corpus, it represents each document as a vector [59]. Table 2 shows that the average TF-IDF F1 Score is 47.63 % and the average TF-IDF Accuracy is 64.63 %.

Table 3. Doc2Vec vectorization and SVM for ABSA

Parameter	Size Aspect		Material Aspect		Delivery Aspect		Overall	
	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score
gamma= 0.01; c=1	0.80	0.78	0.75	0.72	0.83	0.75	0.79	0.75
gamma= 0.01; c=10	0.84	0.82	0.79	0.76	0.87	0.79	0.83	0.79
gamma= 0.01; c=100	0.88	0.85	0.83	0.79	0.91	0.82	0.87	0.82
gamma= 0.1; c=1	0.75	0.73	0.70	0.67	0.78	0.70	0.74	0.70
gamma= 0.1; c=10	0.79	0.77	0.74	0.71	0.82	0.74	0.78	0.74
gamma= 0.1; c=100	0.85	0.80	0.80	0.74	0.88	0.77	0.84	0.77
gamma= 1; c=1	0.75	0.75	0.70	0.69	0.78	0.72	0.74	0.72
gamma= 1; c=10	0.79	0.77	0.74	0.71	0.82	0.74	0.78	0.74
gamma= 1; c=100	0.85	0.82	0.80	0.76	0.88	0.79	0.84	0.79

Doc2Vec, also known Paragraph Vector, is an unsupervised learning method that learns embeddings of documents or paragraphs as fixed-length vectors. It considers the local word order within a document to determine the semantic meaning and contextual information of words and documents. Doc2Vec incorporates documents into continuous vector representations that convey the semantic meaning and context of a document's words. To generate document vectors, it considers the adjacent words and their order within a document [60]. These embeddings can be used for various NLP tasks, such as document similarity, clustering, and classification, to capture relationships between documents. From Fig. 1, the average of Doc2Vec Accuracy is 80.27 %, while the average of Doc2Vec F1 Score is 75.89 %.

From Table 2 and Table 3, it is clear that gamma and c value contributes to the value of accuracy and F1 score both in TF-IDF and Doc2Vec vectorization. According to the results of our experiment, the accuracy and F1 score decrease as gamma parameter value decreases. However, or result shows that the higher c parameter correlates to higher accuracy and F1 scores.

Gamma is a parameter that regulates the impact of a single training example on the decision threshold. A lower gamma value suggests a larger effect, with points farther out from the decision border having a substantial impact. The reduced gamma values make the decision boundary smoother and, if set too low, can result in underfitting [61]. Moreover, the training instances that are closer to the decision border are given greater weight when the gamma value is increased. However, higher gamma values make the decision boundary more complex and, if not appropriately tailored, can lead to overfitting. Therefore, finding the optimal gamma value is crucial for attaining a proper harmony between model complexity and generalization.

Moreover, C is the regularization parameter in SVM, which represents the penalty for misclassifications. A smaller C value permits more misclassifications, resulting in a larger margin and a more straightforward decision boundary [61]. The greater degree of misclassification of training examples may result in underfitting. Furthermore, greater value of C penalizes misclassifications more severely, resulting in a narrower margin and a more complicated decision boundary. Although higher C values give prominence to correctly classifying training examples, it may result in overfitting if the data contains noise or outliers. Depending on the dataset, cross-validation or grid search techniques are typically used to determine the optimal C value.

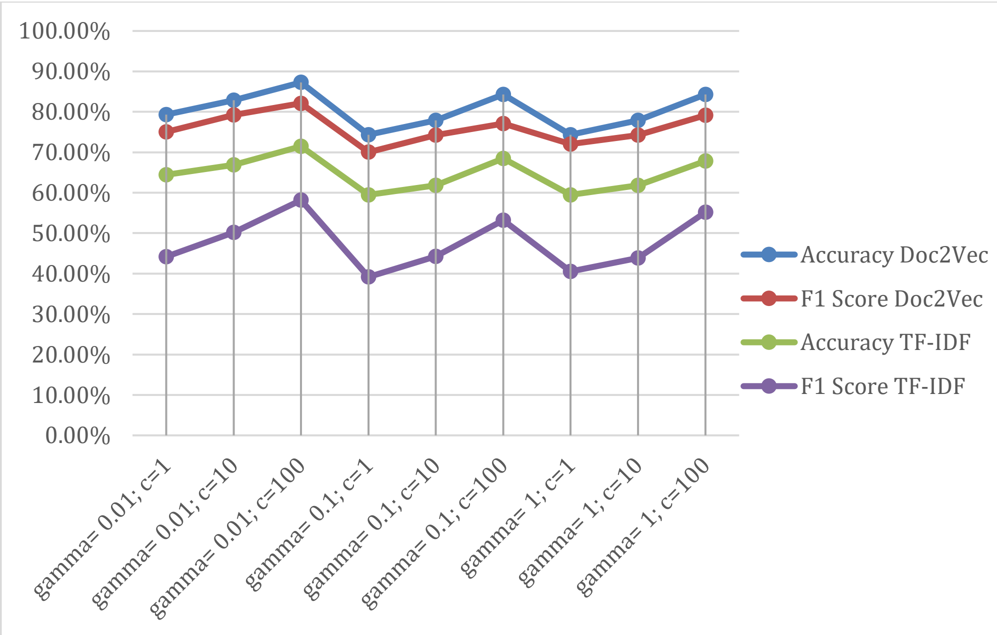


Fig. 1. Comparison result of Doc2Vec and TF-IDF vectorization

Doc2Vec and TF-IDF vectorization are two different techniques used in natural language processing, and their effectiveness depends on the specific task and data at hand. From our results as shown in Fig. 1, it is clear that the performance of SVM algorithm using Doc2Vec vectorization outperformed the SVM algorithm using TF-IDF vectorization. While both techniques have their strengths, Doc2Vec can provide certain advantages over TF-IDF vectorization in certain scenarios

Doc2Vec is a neural network-based approach that learns distributed representations (vectors) for documents [62]. It captures not only the frequency of words but also their semantic meaning and contextual information. This allows Doc2Vec to represent documents in a more comprehensive and nuanced way compared to TF-IDF, which solely relies on term frequencies and inverse document frequencies. According to Jeon *et al.* [63], Doc2Vec can handle out-of-vocabulary words effectively. During training, it learns word embeddings that can generalize to similar words it has encountered, even if they were not present in the training data. This allows it to represent words that may not be present in the vocabulary explicitly, improving its ability to capture the meaning of unseen words during inference. In contrast, TF-IDF vectorization treats out-of-vocabulary words as unseen and may struggle to represent their meaning accurately.

Moreover, Doc2Vec embeddings can be useful for measuring document similarity and performing clustering tasks. By representing documents in a continuous vector space, Doc2Vec allows for more meaningful comparisons between documents. Similar documents tend to have similar vectors, facilitating tasks like finding similar documents, clustering related documents, or identifying document similarity thresholds. TF-IDF, on the other hand, represents documents as sparse vectors based on term frequencies, making direct similarity comparisons less straightforward. According to Hidayat *et al.* [64], and Chen and Sokolova [65], Doc2Vec can be trained in an unsupervised manner, allowing it to learn document representations without the need for labeled data. This makes it suitable for scenarios where labeled data may be scarce or unavailable. TF-IDF, on the other hand, does not inherently capture semantic meaning and relies on term frequencies calculated based on labeled data.

4. CONCLUSION

This study makes use of text data in the form of customer reviews collected from various merchants on the Shopee platform. Out of the 855 data points that were collected, 729 are ready to be further handled for sentiment analysis. The SVM model built in this paper was tested using two distinct vectorization approaches, Doc2Vec and TF-IDF. The performance evaluation model for each classifier and vectorization technique employs accuracy and F1 score. Different parameters are employed systematically for each trial scenario to identify the effect of each parameter. The values of c investigated in this research are 1; 10; and 100. The gamma values utilized are: 0.01; 0.1; and 1. These experiments revealed that both gamma and c value contributes to the accuracy and F1 score. This study also find that Doc2Vec vectorization outperformed TF-IDF vectorization. These findings suggest that the Doc2Vec approach is more suited for customer review categorization because it is capable of recording semantic information in the generated document vector. Future research can investigate the optimization of kernel functions and kernel-specific parameter using Doc2Vec embeddings. This can involve techniques like grid search, Bayesian optimization, or more advanced optimization algorithms to find the optimal hyperparameter settings. Additionally, future research can explore techniques to address class imbalance in sentiment analysis using Doc2Vec SVM, such as oversampling, undersampling, or more advanced methods like SMOTE (Synthetic Minority Over-sampling Technique) or ADASYN (Adaptive Synthetic Sampling). Moreover, cost-sensitive learning techniques can be explored to assign different misclassification costs for different sentiment classes.

REFERENCES

[1] A. Kumar S. Chakraborty and P. K. Bala, ”Text mining approach to explore determinants of grocery mobile app satisfaction using online customer reviews,” *Journal of Retailing and Consumer Services*, vol. 73, p. 103363, 2023, https://doi.org/10.1016/j.jretconser.2023.103363.

[2] M. Nilashi, R. A Abumalloh, S. Samad, M. Alrizq, S. Alyami, and A. Alghamdi, ”Analysis of customers' satisfaction with baby products: The moderating role of brand image,” *Journal of Retailing and Consumer Services*, vol. 73, p. 103334, 2023, https://doi.org/10.1016/j.jretconser.2023.103334.

[3] R. Li, Y. Q. Li, W. Q. Ruan, S. N. Zhang, and M. Y.Wang, “Sentiment mining of online reviews of peer-to-peer accommodations: Customer emotional heterogeneity and its influencing factors,” *Tourism Management,* vol. 96, p. 104704, 2023, https://doi.org/10.1016/j.tourman.2022.104704.

[4] Y. Zheng, X. Li, and J. Y. Nie, Store, share and transfer: Learning and updating sentiment knowledge for aspect-based sentiment analysis,” *Information Sciences*, vol. 635, pp. 151-68, 2023, https://doi.org/10.1016/j.ins.2023.03.102.

[5] L. Jiang, Y. Li, J. Liao, Z. Zou, and C. Jiang, “Research on non-dependent aspect-level sentiment analysis, *Knowledge-Based Systems*, vol. 266, p. 110419, 2023, https://doi.org/10.1016/j.knosys.2023.110419.

[6] L. Xu, and W. Wang, “Improving aspect-based sentiment analysis with contrastive learning,” *Natural Language Processing Journal*, vol. 3, p. 100009, 2023, https://doi.org/10.1016/j.nlp.2023.100009.

[7] H. Li, X. B. Bruce, G. Li, and H. Gao, “Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews,” *Tourism Management*, vol. 96, p. 104707, 2023, https://doi.org/10.1016/j.tourman.2022.104707.

[8] J. Zhang, X. Lu, and D. Liu, “Deriving customer preferences for hotels based on aspect-level sentiment analysis of online reviews,” *Electronic Commerce Research and Applications*, vol. 49, p. 101094, 2021, https://doi.org/10.1016/j.elerap.2021.101094.

[9] Y. Bian, R. Ye, J. Zhang, and X. Yan, “Customer preference identification from hotel online reviews: A neural network based fine-grained sentiment analysis,” *Computers & Industrial Engineering*, vol. 172, p. 108648, 2022, https://doi.org/10.1016/j.cie.2022.108648.

[10] M. Zhang, L. Sun, Y. Li, G.A. Wang, and Z. He, "Using supplementary reviews to improve customer requirement identification and product design development," *Journal of Management Science and Engineering,* 2023, https://doi.org/10.1016/j.jmse.2023.03.001.

[11] P. Mehra, "Unexpected Surprise: Emotion analysis and aspect based sentiment analysis (ABSA) of user generated comments to study behavioral intentions of tourists," *Tourism Management Perspectives*, vol. 45, p. 101063, 2023, https://doi.org/10.1016/j.tmp.2022.101063.

[12] S. Farzadnia, and I. R. Vanani, "Identification of opinion trends using sentiment analysis of airlines passengers' reviews," *Journal of Air Transport Management*, vol. 103, p. 102232, 2022, https://doi.org/10.1016/j.jairtraman.2022.102232.

[13] T. Zhang, Z. He, and A. Mukherjee, "Monitoring negative sentiment scores and time between customer complaints via one-sided distribution-free EWMA schemes," *Computers & Industrial Engineering*, vol 180, p. 109247, 2023, https://doi.org/10.1016/j.cie.2023.109247.

[14] J. Kim, and C. Lim, "Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses," *Advanced Engineering Informatics*, vol. 49, p. 101304, 2021, https://doi.org/10.1016/j.aei.2021.101304.

[15] G. Roy, "Travelers' online review on hotel performance–Analyzing facts with the Theory of Lodging and sentiment analysis," *International Journal of Hospitality Management*, vol. 111, p. 103459, 2023, https://doi.org/10.1016/j.ijhm.2023.103459.

[16] Y. Xiao, C. Li, M. Thürer, Y. Liu, and T. Qu, "Towards Lean Automation: Fine-Grained sentiment analysis for customer value identification," *Computers & Industrial Engineering*, vol. 169, p. 108186, 2022, https://doi.org/10.1016/j.cie.2022.108186.

[17] J. M. Luo, H. Q. Vu, G. Li, and R. Law, "Understanding service attributes of robot hotels: A sentiment analysis of customer online reviews," International Journal of Hospitality Management, vol. 98, p. 103032, 2021, https://doi.org/10.1016/j.ijhm.2021.103032.

[18] R. Sann, and P. C. Lai, "Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry," *International Journal of Hospitality Management*, vol. 91, p. 102678, 2020, https://doi.org/10.1016/j.ijhm.2020.102678.

[19] Y. Tao, W. Liu, Z. Huang, and C. Shi, "Thematic analysis of reviews on the air quality of tourist destinations from a sentiment analysis perspective," *Tourism Management Perspectives*, vol. 42, p. 100969, 2022, https://doi.org/10.1016/j.tmp.2022.100969.

[20] R. Khorsand, M. Rafiee, and V. Kayvanfar, "Insights into TripAdvisor's online reviews: The case of Tehran's hotels," *Tourism Management Perspectives,* vol. 34, p. 100673, 2020, https://doi.org/10.1016/j.tmp.2020.100673.

[21] P. Rita, R. Ramos, M. T. Borges-Tiago, and D. Rodrigues, "Impact of the rating system on sentiment and tone of voice: A Booking. com and TripAdvisor comparison study," *International Journal of Hospitality Management*, vol. 104, p. 103245, 2022, https://doi.org/10.1016/j.ijhm.2022.103245.

[22] Z. Zhan, and B. Xu, "Analyzing review sentiments and product images by parallel deep nets for personalized recommendation," *Information Processing & Management*, vol. 60, no. 1, p. 103166, 2023, https://doi.org/10.1016/j.ipm.2022.103166.

[23] Z. Ahmed, and J. Wang, "A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments," *Alexandria Engineering Journal*, vol. 65, pp. 731-747, 2023, https://doi.org/10.1016/j.aej.2022.10.037.

[24] N. Punetha and G. Jain, "Bayesian game model based unsupervised sentiment analysis of product reviews," *Expert Systems with Applications*, vol. 214, p. 119128, 2023, https://doi.org/10.1016/j.eswa.2022.119128.

[25] J. Zhang, A. Zhang, D. Liu, and Y. Bian, "Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews," *Knowledge-Based Systems,* vol. 228, p. 107259, 2021, https://doi.org/10.1016/j.knosys.2021.107259.

[26] S. Jardim, and C. Mora, "Customer reviews sentiment-based analysis and clustering for market-oriented tourism services and products development or positioning," *Procedia Computer Science,* vol. 196, pp. 199-206, 2022, https://doi.org/10.1016/j.procs.2021.12.006.

[27] M. Zhang, L. Sun, G. A. Wang, Y. Li, and S. He, "Using neutral sentiment reviews to improve customer requirement identification and product design strategies," *International Journal of Production Economics,* vol. 254, p. 108641, 2022, https://doi.org/10.1016/j.ijpe.2022.108641.

[28] P. Eachempati, P. R. Srivastava, A. Kumar, J. M. de Prat, and D. Delen, "Can customer sentiment impact firm value? An integrated text mining approach," *Technological Forecasting and Social Change*, vol. 174, p. 121265, 2022, https://doi.org/10.1016/j.techfore.2021.121265.

[29] R. V. Karthik, and S. Ganapathy, "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce," *Applied Soft Computing*, vol. 108, p. 107396, 2021, https://doi.org/10.1016/j.asoc.2021.107396.

[30] J. Joung, and H. Kim, "Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews," *International Journal of Information Management*, vol. 70, p. 102641, 2023, https://doi.org/10.1016/j.ijinfomgt.2023.102641.

[31] A. Patel, P. Oza, and S. Agrawal, "Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model," *Procedia Computer Science,* vol. 218, pp. 2459-2467, 2023, https://doi.org/10.1016/j.procs.2023.01.221.

[32] S. Park, J. Cho, K. Park, and H. Shin, "Customer sentiment analysis with more sensibility," *Engineering Applications of Artificial Intelligence*, vol. 104, p. 104356, 2021, https://doi.org/10.1016/j.engappai.2021.104356.

[33] A. Giachanou, J. Gonzalo, and F. Crestani, "Propagating sentiment signals for estimating reputation polarity," *Information Processing & Management* vol. 56, no. 6, p. 102079, 2019, https://doi.org/10.1016/j.ipm.2019.102079.

[34] Z. Wu, Q. He, J. Li, G. Bi, and M. F. Antwi-Afari, "Public attitudes and sentiments towards new energy vehicles in China: A text mining approach," *Renewable and Sustainable Energy Reviews* vol. 178, p. 113242, 2023, https://doi.org/10.1016/j.rser.2023.113242.

[35] T. Ali, B. Marc, B. Omar, K. Soulaimane, and S. Larbi, "Exploring destination's negative e-reputation using aspect based sentiment analysis approach: case of Marrakech destination on TripAdvisor," *Tourism Management Perspectives*, vol. 40, p. 100892, 2021, https://doi.org/10.1016/j.tmp.2021.100892.

[36] Y. Song, K. Liu, L. Guo, Z. Yang, and M. Jin, "Does hotel customer satisfaction change during the COVID-19? A perspective from online reviews," *Journal of Hospitality and Tourism Management*, vol 51, pp. 132-138, 2022, https://doi.org/10.1016/j.jhtm.2022.02.027.

[37] M. Arzaghi, I. H. Genc, and S. Naik, "Rating vs. Reviews: Does Official Rating Capture What is Important to Customers?," *Heliyon*, vol. 9, no. 5, 2022, https://doi.org/10.1016/j.heliyon.2023.e16337.

[38] D. Perez, I. Stockheim, and G. Baratz, "Complimentary competition: the impact of positive competitor reviews on review credibility and consumer purchase intentions," *Journal of Retailing and Consumer Services*, vol. 69, p. 103093, 2022, https://doi.org/10.1016/j.jretconser.2022.103093.

[39] S. Kurtmollaiev, L. Lervik-Olsen, and T. W. Andreassen, "Competing through innovation: Let the customer judge!," *Journal of Business Research*, vol. 153, pp. 87-101, 2022, https://doi.org/10.1016/j.jbusres.2022.08.002.

[40] H. Huang, B. Zhang, L. Jing, X. Fu, X. Chen, and J. Shi, "Logic tensor network with massive learned knowledge for aspect-based sentiment analysis," *Knowledge-Based Systems*, vol. 257, p. 109943, 2022, https://doi.org/10.1016/j.knosys.2022.109943.

[41] Z. A. Diekson, M. R. Prakoso, M. S. Putra, M. S. Syaputra, S. Achmad, and R. Sutoyo, "Sentiment analysis for customer review: Case study of Traveloka," *Procedia Computer Science*, vol. 216, pp. 682-690, 2023, https://doi.org/10.1016/j.procs.2022.12.184.

[42] C. F. Tsai, K. Chen, Y. H. Hu, and W. K. Chen, "Improving text summarization of online hotel reviews with review helpfulness and sentiment," *Tourism Management*, vol. 80, p. 104122, 2020, https://doi.org/10.1016/j.tourman.2020.104122.

[43] Z. Wang , P. Gao, and X. Chu, "Sentiment analysis from Customer-generated online videos on product review using topic modeling and Multi-attention BLSTM," *Advanced Engineering Informatics,* vol. 52, p. 101588, 2022, https://doi.org/10.1016/j.aei.2022.101588.

[44] A. Borg, and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert Systems with Applications*, vol. 162, p. 113746, 2020, https://doi.org/10.1016/j.eswa.2020.113746.

[45] H. A. Fayed, and A. F. Atiya, "Decision boundary clustering for efficient local SVM," *Applied Soft Computing*, vol. 110, p. 107628, 2021, https://doi.org/10.1016/j.asoc.2021.107628.

[46] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Computers & Operations Research*, vol. 152, p. 106131, 2023, https://doi.org/10.1016/j.cor.2022.106131.

[47] H. Wang, G. Li, and Z. Wang, "Fast SVM classifier for large-scale classification problems," *Information Sciences*, p. 119136, 2023, https://doi.org/10.1016/j.ins.2023.119136.

[48] K. R. Singh, K. P. Neethu, K. Madhurekaa, A. Harita, and P. Mohan, "Parallel SVM model for forest fire prediction," *Soft Computing Letters*, vol. 3, p. 100014, 2021, https://doi.org/10.1016/j.socl.2021.100014.

[49] X. Liu, X. Cheng, and K. Lee, "GA-SVM-based facial emotion recognition using facial geometric features, *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11532-11542, 2020, https://doi.org/10.1109/JSEN.2020.3028075.

[50] Z. Gao, S. C. Fang, X. Gao, J. Luo, and N. Medhin, "A novel kernel-free least squares twin support vector machine for fast and accurate multi-class classification," *Knowledge-Based Systems*, vol. 226, p. 107123, 2021, https://doi.org/10.1016/j.knosys.2021.107123.

[51] M. N. Sha'Abani, N. Fuad, N. Jamal, and M. F. Ismail, "kNN and SVM classification for EEG: a review," In *ECCE2019: Proceedings of the 5th International Conference on Electrical, Control & Computer Engineering - Malaysia,* pp. 555-565, 2019, https://doi.org/10.1007/978-981-15-2317-5_47.

[52] A. Alsarhan, M. Alauthman, E. A. Alshdaifat, A. R, Al-Ghuwairi, and A. Al-Dubai, "Machine Learning-driven optimization for SVM-based intrusion detection system in vehicular ad hoc networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 6113–6122, 2023, https://doi.org/10.1007/s12652-021-02963-x.

[53] Q. Gu, Y. Chang, X. Li, Z. Chang, and Z. Feng, "A novel F-SVM based on FOA for improving SVM performance," *Expert Systems with Applications,* vol. 165, p. 113713, 2021, https://doi.org/10.1016/j.eswa.2020.113713.

[54] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating trust prediction and confusion matrix measures for web services ranking," *IEEE Access*, vol. 8, pp. 90847-90861, 2020, https://doi.org/10.1109/ACCESS.2020.2994222.

[55] D. Chicco, and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics,* vol. 21, pp. 1-13, 2020, https://doi.org/10.1186/s12864-019-6413-7

[56] T. B. Alakus, and I. Turkoglu, "Comparison of deep learning approaches to predict COVID-19 infection," *Chaos, Solitons & Fractals*, vol. 140, p. 110120, 2020, https://doi.org/10.1016/j.chaos.2020.110120.

[57] G. A. Sambasivam, and G. D. Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egyptian informatics journal*, vol. 22, no. 1, pp. 27-34, 2021, https://doi.org/10.1016/j.eij.2020.02.007.

[58] N. W. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, and P. Lestantyo, "Cross-validation metrics for evaluating classification performance on imbalanced data," *In 2019 International conference on computer, control, informatics and its applications (IC3INA),* pp. 14-18, 2019, IEEE, https://doi.org/10.1109/IC3INA48034.2019.8949568.

[59] S. Wehnert, V. Sudhi, S. Dureja, L. Kutty, S. Shahania, and E. W De Luca, "Legal norm retrieval with variations of the bert model combined with tf-idf vectorization," *In Proceedings of the eighteenth international conference on artificial intelligence and law,* pp. 285-294, 2021, https://doi.org/10.1145/3462757.3466104.

[60] S. Amasaki and P. Leelaprute, "The Effects of Vectorization Methods on Non-Functional Requirements Classification," *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Prague, Czech Republic*, pp. 175-182, 2018, https://doi.org/10.1109/SEAA.2018.00036.

[61] H. A. Fayed and A. F. Atiya, "Speed up grid-search for parameter selection of support vector machines," *Applied Soft Computing*, vol. 80, pp. 202-10, 2019, https://doi.org/10.1016/j.asoc.2019.03.037.

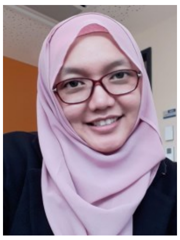
[62] M. Hanifi, H. Chibane, R. Houssin, and D. Cavallucci, "Problem formulation in inventive design using Doc2vec and Cosine Similarity as Artificial Intelligence methods and Scientific Papers," *Engineering Applications of Artificial Intelligence,* vol. 109, p. 104661, 2022, https://doi.org/10.1016/j.engappai.2022.104661.

[63] D. Jeon, J. M. Ahn, J. Kim, and C. Lee, "A doc2vec and local outlier factor approach to measuring the novelty of patents," *Technological Forecasting and Social Change*, vol. 174, p. 121294, 2022, https://doi.org/10.1016/j.techfore.2021.121294.

[64] T. H. Hidayat, Y. Ruldeviyani, A. R. Aditama, G. R. Madya, A. W. Nugraha, M. W. Adisaputra, "Sentiment analysis of twitter data related to Rinca Island development using Doc2Vec and SVM and logistic regression as classifier," *Procedia Computer Science*, vol. 197, pp. 660-667, 2022, https://doi.org/10.1016/j.procs.2021.12.187.

[65] Q. Chen, and M. Sokolova, "Unsupervised sentiment analysis of objective texts," In *Advances in Artificial Intelligence: 32nd Canadian Conference on Artificial Intelligence Proceedings,* vol. 32, pp. 460-465, May 2019, Springer International Publishing, https://doi.org/10.1007/978-3-030-18305-9_45.

BIOGRAPHY OF AUTHOR



Andharini Dwi Cahyani, an associate professor at Universitas Trunojoyo Madura, where she is on the faculty of the Department of Informatics Engineering. The researcher is particularly interested in the fields of information system management, natural language processing, multi-criteria decision making, learning analytics, and machine learning. Email: andharini.cahyani@trunojoyo.ac.id. Orcid: https://orcid.org/0000-0003-3560-7796 .