

Mining the public sentiment for *wayang climen* preservation and promotion



Aji Prasetya Wibawa ^{a,1,*}, Adjie Rosyidin ^{a,2}, Fitriana Kurniawati ^{a,3}, Gwinny Tirza Rarastris ^{a,4},
Ilham Ari Elbaith Zaeni ^{a,5}, Suyono ^{b,6}, Agung Bella Putra Utama ^{a,7}, Felix Andika Dwiyanto ^{c,8}

^a Department of Electrical Engineering and Informatics, Faculty of Engineering, Universitas Negeri Malang, Indonesia

^b Department of Indonesian Literature, Faculty of Letters, Universitas Negeri Malang, Indonesia

^c Faculty of Computer Science, Electronics, and Telecommunications, AGH University of Science and Technology, Poland

¹ aji.prasetya.ft@um.ac.id; ² adjie.rosyidin.1705356@students.um.ac.id; ³ fitriana.kurniawati.1705356@students.um.ac.id; ⁴

gwinny.tirza.1705356@students.um.ac.id; ⁵ ilham.ari.ft@um.ac.id; ⁶ suyono.fs@um.ac.id; ⁷ agung.bella.2305349@students.um.ac.id; ⁸ dwiyanto@agh.edu.pl

* corresponding author

ARTICLE INFO

Article history

Received 2023-09-17

Revised 2023-10-09

Accepted 2023-10-18

Keywords

Sentiment Analysis

Wayang Climen

Social Media

SVM

Random Forest

Naïve Bayes

ABSTRACT

Indonesia is a country that has a variety of cultural arts, one of which is shadow puppetry (*Wayang*). *Wayang*, in a staged, simple, and minimalist manner, is called *Wayang Climen*. *Wayang Climen* has been performed since the COVID-19 pandemic as a solution to keep working while still complying with health protocols. Utilization through YouTube social media attracts people to watch and provide opinions through comments. This opinion is beneficial and can be used as a feasibility study through sentiment analysis information classified as positive, negative, and neutral opinions. Sentiment analysis determines a person's opinion and tendency to opinionated sentences. The methods used are Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB). The dataset comes from YouTube comments of *Dalang* Seno and Ki Seno Nugroho. The best accuracy is generated by SVM (70.29%). The positive sentiment shows the public's appreciation for the *Wayang Climen* performance, which ultimately represents the performance even though it is staged densely. This research contributes to effectively utilizing digital platforms for cultural preservation and audience engagement during challenging times, demonstrating the potential for innovative solutions in traditional arts and entertainment.



This is an open-access article under the [CC-BY-SA](#) license.



1. Introduction

Indonesia, a country characterized by a rich diversity of civilizations, possesses a remarkable cultural treasure known as *Wayang*, which holds a prominent position in its artistic heritage [1]. The esteemed cultural legacy, a valuable repository of customs and practices, can be historically traced, with its origins believed to date back to approximately 1500 BC [2]. *Wayang*, an enthralling artistic expression, provides a riveting insight into the human condition [3]. The performing arts play diverse and significant roles within society, extending beyond their entertainment value to serve as a potent means of imparting moral guidance to the community [4]. They also serve as a sincere tribute to ancestral spirits, seeking their blessings. Furthermore, the performing arts function as an educational tool, shaping the minds of young individuals. Additionally, during the era of Sunan Kalijaga, they were utilized as a medium for spreading religious beliefs [5]. The ongoing appeal of *Wayang* lies in its moral tales and timeless storytelling, which have successfully endured the dynamic terrain of entertainment, spanning from the emergence of television to the proliferation of social media [6]. Clearly, *Wayang* has not only persevered but has flourished, successfully assimilating into contemporary culture, including the extensive domain of social media.

In the contemporary era of technological advancements, social media has emerged as a prominent means of connecting individuals across great geographical distances, serving as a unifying force accessible to a wide range of people [7]. The platform functions as an interactive public space where ideas, information, and opinions are exchanged without restrictions [8]. However, the global pandemic has had a significant impact, leading governments across the globe, including Indonesia, to implement tough restrictions such as the Large-Scale Social Restrictions (PSBB). These measures were implemented in order to restrict community movement and meetings, which resulted in a shared need for the vibrant cultural experiences that were formerly integral to communal existence [9]. Amidst the current circumstances characterized by difficulties, a novel and inventive idea emerged within the innermost being of *Dalang Seno*. *Dalang Seno*, driven by an unyielding dedication to the preservation of the *Wayang* tradition, utilized the influential platform of YouTube as a virtual stage that effectively reaches audiences in the digital realm. Through this medium, Seno successfully reignited the fervor for *Wayang* within a receptive and eager viewership. Therefore, the emergence of the *Wayang Climen* performance can be attributed to a deviation from conventional practices, as it embraces a minimalist and streamlined approach [10]. Utilizing the platform YouTube, this digital theatre served as a valuable resource for anyone seeking engagement with the art form, providing a sanctuary of cultural abundance within the digital realm.

However, the narrative does not conclude at this point; rather, it marks the commencement of a new phase. YouTube, as a widely accessible platform, serves as a conduit for the dissemination of public opinions and sentiments pertaining to *Dalang Seno's Wayang Climen* performances. Sentiment analysis is a complex procedure that involves extracting the underlying emotions present in textual data. It serves as a tool for observing and understanding the audience's responses [11]. In the realm of text categorization, approaches such as Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB) are employed to reveal the underlying sentiment that is obscured inside comments. With this valuable understanding, a comprehensive marketing and preservation strategy for *Wayang Climen* performances can be carefully developed, guaranteeing the sustained prosperity of this ancient tradition among future generations. This strategy will encompass both offline and online platforms, effectively bridging the divide between the past and the future while maintaining cultural continuity.

Social media is online content that uses a technology easily accessible to the public [12]. People use social media to exchange information and thoughts. People actively participate in social media to convey opinions and comments [13]. During the pandemic, the government restricted community mobility by implementing the PSBB (Large-Scale Social Restrictions) regulation. People are encouraged to stay at home and not gather in crowds [14]. Due to government regulations, *Dalang Seno* initiated an innovative *Wayang Climen* performance broadcast through YouTube social media to treat people's longing for *Wayang* performances. *Dalang Seno* performs *Wayang Climen* with different characteristics from the usual *Wayang* performances. *Wayang Climen* is made with a minimalist and simple concept [15]. YouTube social media can provide public access to opinions on *Dalang Seno's Wayang Climen* performances. Sentiment analysis of public comments is carried out to determine how the public comments on *Wayang Climen*. Sentiment analysis is the process of processing textual data into important sentiment information contained in a sentence [16]. Sentiment analysis usually uses text classification methods like Random Forest (RF), Support Vector Machine (SVM), and Naïve Bayes (NB). Once it is implemented, the best algorithm is used to develop a marketing and preservation strategy for *Wayang Climen's* performance

2. Method

The foundation of our research is based on a comprehensive and abundant dataset collected from the diverse array of comments on YouTube, which serves as a valuable source of information. This collection of conversations encompasses the remarkable *Wayang Climen* performances of two esteemed artists, namely *Dalang Seno* and *Ki Seno Nugroho*. The virtuosos have graciously disseminated their artistic performances on the online platform YouTube, providing accessibility to anybody interested in exploring the captivating realm of *Wayang Climen*. In order to commence this digital journey, individuals just need to access *Dalang Seno's* YouTube platform, which serves as a virtual domain teeming with diverse forms of artistic expression. The URL of *Dalang Seno's* YouTube can be accessed at <https://www.youtube.com/channel/UCHNQcvuYixP35Cf3tLb83pA>, while *Ki Seno Nugroho's* YouTube can be accessed through <https://www.youtube.com/channel/UCv9wXGBmh9PQcg6S0O4hBXA>. The dataset utilized in this

study comprises a collection of remarks, whispers, and exclamations that were carefully selected from the *Wayang Climen* performances hosted on internet platforms. The temporal fabric depicted in the tapestry encompasses the period from July to September in the year 2020, encapsulating the distinctive characteristics of a singular juncture in history. During the specified time period, our data collection efforts resulted in a substantial collection of 2126 comments originating from the artistic realm of *Dalang Seno*, as well as 794 remarks originating from the digital domain of *Ki Seno Nugroho*. The comments provided by individuals serve as valuable resources for extracting thoughts and ideas that contribute to our study and enhance our comprehension of the cultural significance of *Wayang Climen* in the era of digital technology

2.1. Labeling

Within the complex framework of our study, the pivotal role of assigning labels takes on a prominent position, leading us toward a deeper comprehension of the sentiments expressed within the comments. Similar to skilled craftsmen crafting a work of art, we carefully classify these comments into three distinct categories based on sentiment: positive, negative, and neutral. The aforementioned process, as exemplified in the comprehensive study conducted by Ardani and Sujaini, is executed by a laborious and dedicated manual effort. During the rigorous process of manual labeling, every comment is carefully examined with precise attention to detail. Through the utilization of a practical methodology, we engage in the process of analyzing digital manifestations, thereby distinguishing the nuanced emotions and attitudes that are trapped inside them. In order to guarantee the highest level of precision and thoroughness in our undertaking, our data passed a meticulous validation process. The responsibility for this assignment was given to the highly regarded linguist. The inclusion of the linguist's experience enhanced the academic rigor and linguistic elegance of our labeling procedure, hence reinforcing the credibility and reliability of our findings. The core of our labeling procedure centers around the differentiation of adjectives and verbs, which are fundamental components of language used to express attitudes and emotions. Every word in a comment is thoroughly evaluated, taking into account its connotations, and carefully scrutinized to determine whether it falls into the category of positive, negative, or neutral feeling. A visual representation of the process of labeling is depicted in [Table 1](#), providing valuable insights into the complex relationship between language and sentiment that forms the foundation of our analysis.

Table 1. Labeling process

No.	Comment	Adjectives and Verbs			Sentiment
		Positive	Neutral	Negative	
1	<i>Mbelisa cantik sekali suarane bagus</i>	<i>cantik, bagus</i>	-	-	Positive
2	<i>Terlambat, gak masalah nggih ki Seno... Tetap terhibur...</i>	<i>terhibur</i>	<i>tetap</i>	<i>terlambat</i>	Neutral
3	<i>Mas soundnya pecah mas, ground pecah</i>	-	-	<i>pecah</i>	Negative
4	<i>Latamahusadi itu artinya apa ya lur?</i>	-	-	-	Neutral

2.2. Data Preprocessing

The translated dataset needs to be preprocessed. This process is vital in all relevant data mining applications, especially sentiment analysis [17]. Several preprocessing techniques are performed; the first stage is case folding. The *Wayang Climen* dataset has the use of capital and lowercase letters. So, it is necessary to adjust the letters so that there is no duplication of words with the same meaning. In this research, the capital letters in the comments are changed to lowercase letters. The flow of case-folding can be seen in [Fig. 1](#).

Input	: Translated text
Output	: Text in lowercase
Process	:
Begin	
	1) Input the translated text into the program or system
	2) Change the capital letters of the text to lowercase
	3) Produce output in the form of text in lowercase letters
End	

Fig. 1. Pseudocode of the case folding process

The following preprocessing stage is to remove punctuation. This process removes components that do not affect the classification process [18]. These components include full stops (.), commas (,), question marks (?), and numbers. The flow of removed punctuation can be seen in Fig. 2.

Input	: Text in lowercase
Output	: Text without punctuation and symbols
Process	:
Begin	
	1) Input text in lowercase letters in the program or system
	2) Eliminate punctuation and symbols in the text
	3) Produce output in the form of text without punctuation and symbols
End	

Fig. 2. Pseudocode of the remove punctuation process

Next is the tokenizing process. This stage involves cutting sentences into individual words and organizing them into single chunks [19]. Single pieces are commonly referred to as tokens. This stage is needed for the stopword removal stage. The flow of the tokenizing process can be seen in Fig. 3.

Input	: Text without punctuation and symbols
Output	: Collection of single words
Process	:
Begin	
	1) Input text without punctuation and symbols in the program or system
	2) Separating text in sentence form into single words
	3) Produce output in the form of a collection of single words
End	

Fig. 3. Pseudocode of the tokenizing process

The fourth preprocessing stage is stopword removal. This stage removes personal pronouns or connectives such as I, and, with, which, certainly, if, when, also, because, will, etc. The flow of the stopword removal process can be seen in Fig. 4.

Input	: A collection of single words
Output	: A collection of single words without personal pronouns and connecting words
Process	:
Begin	
	1) Input a collection of words into the program or system
	2) Eliminate the pronouns and connecting words in the group of words
	3) Produce output in the form of a collection of single words without personal pronouns and connecting words
End	

Fig. 4. Pseudocode of stopword removal process

The fifth stage is Part of Speech (POS) Tagging. This stage is to tag each word. Researchers used Fam Rashel's data as training data in the tagging process [20]. The data has two hundred thousand tokens that already have tags. The flow of the post-tagging process can be seen in Fig. 5.

```

Input : A collection of single words without personal pronouns and connecting words
Output : A collection of single words with tags
Process :
Begin
    1) Input a collection of single words without personal pronouns and conjunctions into the program or system
    2) Add tags to each word according to the type of word
    3) Produce output in the form of a collection of single words with tags
End

```

Fig. 5. Pseudocode of the post-tagging process

The last preprocessing stage is resampling. The balance of class distribution needs to be considered in the classification process. The positive, negative, and neutral class distribution of the *Wayang Climen* data is uneven. So, the resampling stage is needed. The resampling method used in this research is the Synthetic Minority Over-Sampling Technique (SMOTE), with a development ratio of 100%. Researchers chose SMOTE because it can prevent the loss of important information in the data [21]. SMOTE can prevent overfitting. The flow of resampling can be seen in Fig. 6.

```

Input : The unequal distribution of each class
Output : Even distribution of each class
Process :
Begin
    1) Determine the K value of the nearest neighbor sample for each sample in the minority class
    2) Randomly select N samples from each nearest neighbor
    3) Compute new samples from the minority class and seitan N is a new synthesis, and the synthesis is added to the minority class
    4) Repeat the process until the minority class meets the class number requirements
End

```

Fig. 6. Pseudocode of resampling process

The distribution of *Wayang Climen* data for each class after resampling can be seen in Table 2 and Table 3.

Table 2. Data distribution of *Dalang Seno*

Class Name	Data distribution of <i>Dalang Seno</i>					
	<i>Before Oversampling</i>			<i>After Oversampling</i>		
	<i>Total</i>	<i>Percentage</i>	<i>Amount</i>	<i>Total</i>	<i>Percentage</i>	<i>Amount</i>
Positive	925	43.50%	2126	1073	33.33%	3219
Negative	128	6.02%		1073	33.33%	
Neutral	1073	50.48%		1073	33.33%	

Table 3. Ki Seno Nugroho data distribution

Class Name	Data distribution of <i>Dalang Seno</i>					
	<i>Before Oversampling</i>			<i>After Oversampling</i>		
	<i>Total</i>	<i>Percentage</i>	<i>Amount</i>	<i>Total</i>	<i>Percentage</i>	<i>Amount</i>
Positive	352	44.33%	794	391	33.33%	1173
Negative	51	6.42%		391	33.33%	
Neutral	391	49.25%		391	33.33%	

When viewed in Table 2, *Dalang Seno*'s data after resampling, the initial data amounted to 2126 to 3219 with an equal distribution for each class, namely 1073. From Table 3, as for Ki Seno Nugroho's data, the initial data amounted to 794 after resampling to 1173.

2.3. Support Vector Machine (SVM)

At this stage, researchers use the SVM algorithm, which is used for classification. SVM is a method for class prediction based on the training process [22]. SVM uses a statistical classification approach to find the most significant margin between instances and the hyperplane [23]. An illustration of the SVM algorithm can be seen in Fig. 7.

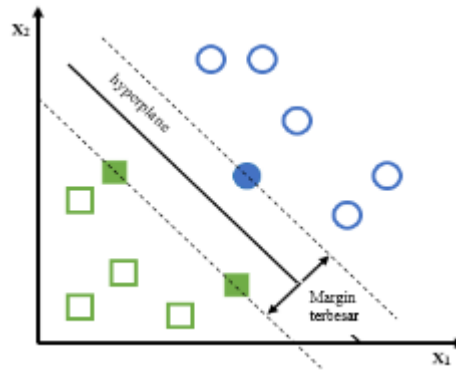


Fig. 7. Pseudocode of resampling process

The distance between neighboring instances and separating classes is called the margin. The algorithm looks for the most significant margin and will form the optimum hyperplane. The instance that is closest to the hyperplane is called the support vector. According to Song, the advantages of the SVM algorithm are that it provides the smallest generalization error of other methods and can solve problems with high dimensions and limited samples [24]. The weakness of SVM is its complicated computation for high-dimensional data [25]. The equation of the SVM algorithm is defined as follows in (1) to (2) [26].

$$f(x) = w \cdot x + b \quad (1)$$

or

$$f(x) = \sum_{i=1}^m a_i y_i K(x, x_i) + b \quad (2)$$

Where w = hyperplane; x = data point; a_i = data point value; y_i = data class; $K(x, x_i)$ = kernel function, and b = bias.

2.4. Naïve Bayes (NB)

In this research, the NB algorithm is used in the classification stage. NB is a supervised learning algorithm requiring training data for model-building classification [27]. NB has the principle that attributes in a data set do not affect each other. This principle is called class conditional confidence. The NB algorithm classification process analyzes document samples and determines the category value [28]. The advantages of this algorithm are that it can handle discrete and continuous data, is easy to implement, can provide good results from a variety of cases, and can automatically classify documents [29]. Moreover, the implementation is simple, and the processing time required is short [30]. As for the weaknesses, it cannot model the relationship between variables because NB has the principle that the variables in the data do not affect each other [31] and requires many records to produce a good accuracy value [32]. Bayes' theorem is a theory based on conditional probability. In general, Bayes' theory can be seen as in (3) [33]. Where A is the label of the data, and B is the feature contained in it.

$$P(A|B) = \frac{P(A|B)P(A)}{P(B)} \quad (3)$$

2.5. Random Forest (RF)

At this classification stage using the RF algorithm, RF is a collection of learning methods using decision trees as a base classifier built and combined [34]. According to [35], the RF method is a development of the decision tree method combined with the bootstrap aggregating (bagging) method and random feature selection. Decision Tree (DT) uses information gain and gini index as the selection of attribute criteria used to determine the root node and rule in building the tree, and this also applies to random forest in building more than one tree.

Information gain is each attribute in the data set is calculated to obtain the information gain value using equation (4). The results of this calculation will select the attribute with the highest information gain value [36].

$$Entropy(S) = \sum_{i=1} -p_i \log_2 p_i \quad (4)$$

$Entropy(S)$ is the set of cases, n is the number of classification classes, and p_i is the probability value in class i .

In the Gini index, each attribute in the data set is calculated to get the Gini index value using equation (5). The results of this calculation will be selected for the attribute with the lowest Gini index value.

$$Gini(A) = 1 - \sum_{k=1}^C P_k^2 \quad (5)$$

Where k is the attribute class, C is the number of Y variable classes, and P_k is the proportion of the number of classes in attribute K to the number of classes C .

In the classification process, each tree votes, and the class with the most votes is selected. This mechanism is called majority voting [28]. The illustration of RF can be seen in Fig. 8.

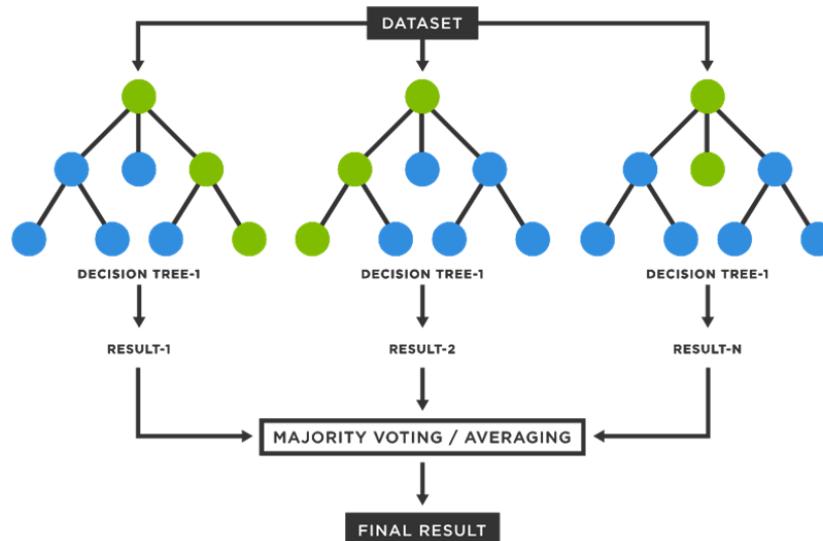


Fig. 8. Random Forest Illustration (Source <https://www.tibco.com/reference-center/what-is-a-random-forest>)

The more trees in the random forest, the higher the accuracy [37]. According to [38] RF method, the advantages of this algorithm are that it can handle multiclass classification, predictor variables can be categorical or continuous, relatively fast to train and predict, depending on one or two parameters, each tree built has an error estimate (OOB), and can overcome overfitting problems. Random Forest also has weaknesses, namely in the stability of data accuracy [39].

3. Results and Discussion

The last stage is obtained after classification, namely conducting an evaluation using the confusion matrix method, which measures the accuracy, precision, recall, and f1 score results from the data obtained. The classification of *Wayang Climen's* comments has three classes: positive, negative, and neutral. The number of classes in the dataset used is a multiclass confusion matrix.

3.1. Sentiment Analysis Performance

Table 4 shows the performance comparison results of the three methods used in this study. Accuracy in sentiment analysis can vary depending on various factors, including data type, training data size, text processing, and the model used. In the case of sentiment *Wayang*, SVM can have better accuracy than NB or RF due to the following factors.

Table 4. Classification model performance results

Scenario	Accuracy	Precision	Recall	F1 Score
Random Forest (RF)	62.72%	52.28%	49.40%	50.08%
Naïve Bayes (NB)	63.22%	52.12%	50.39%	50.84%
Support Vector Machine (SVM)	63.72%	52.92%	49.09%	49.96%

The field of sentiment analysis, which is as complex as human emotions, frequently faces a complex terrain where straightforward connections between data characteristics and sentiment classifications are insufficient. The Support Vector Machine (SVM) is recognized as a prominent and reliable method in this context. In contrast to linear models, Support Vector Machines (SVM) provide the capability to effectively handle the intricacies inherent in human sentiment, which often deviate from conventional patterns. Within the realm of fluctuating emotions, support vector machines (SVMs) emerge as unwavering entities, possessing the ability to navigate a path where alternative approaches may encounter difficulties. The study of sentiment, similar to the intricate tapestry of existence, frequently manifests with varying intensity among different sentiment categories. The environment is often characterized by neutrality, with limited presence of both positive and negative elements. In such situations, Support Vector Machines (SVMs) demonstrate their versatility by effectively managing imbalances in class distribution. By skillfully adjusting the weights assigned to different classes, Support Vector Machines (SVMs) are able to maintain a balanced representation, so preventing any sentiment from being unjustly overwhelmed. However, the capabilities of Support Vector Machines (SVM) go beyond their adaptability, as they also possess the potential to be tuned for optimal performance. The parameters of the system, such as the kernel and the elusive C, function as the fundamental components utilized in its operations. Through the process of expert tuning, Support Vector Machines (SVMs) are able to smoothly integrate with the training data, resulting in the creation of decision boundaries that are beyond the limitations of conventional methods. What is the result? Enhanced precision is a valuable asset within the domain of sentiment analysis. Support Vector Machines (SVMs) demonstrate their effectiveness, particularly in the complex domain of high-dimensional data, such as the sophisticated field of text analysis. Whether it is the mysterious choreography of one-hot encoding representations or the seamless integration of word vectors, support vector machines (SVMs) thrive in intricate scenarios, surpassing models such as Naive Bayes and Random Forest.

Support Vector Machines (SVMs) possess a notable advantage in mitigating the issue of overfitting, a significant challenge frequently encountered with models such as Random Forest. Support Vector Machines (SVMs) exhibit a notable level of restraint and elegance as they construct models that deliberately avoid incorporating noise present in the training data. Consequently, Support Vector Machines (SVMs) have demonstrated superior performance compared to other methods in scenarios including unfamiliar data, thereby exhibiting their efficacy in practical contexts. However, it is important to exercise caution as support vector machines (SVMs) may not possess universal efficacy. The performance of the system is subject to variability, as it is influenced by the data it encounters and the parameters it possesses. In the pursuit of achieving optimal precision in sentiment analysis, it is widely acknowledged that experimentation serves as the fundamental principle. Various datasets can have varying results in terms of identifying champions, and the process of selecting models and parameters is still an untapped resource that needs to be explored for each individual project.

3.2.Strategy for Developing *Wayang Climen* Based on Sentiment Analysis

Within the domain of sentiment and language analysis, the concept of positive sentiment serves as a prominent manifestation of benevolence and optimism. It represents the experience of profound joy, the expression of deep fulfillment, or the enthusiastic recognition of a valued entity, whether it be a subject matter, a product, a service, or a circumstance. When an individual expresses a favorable emotion towards a compelling *Wayang* presentation, it indicates an acknowledgment of the artistic qualities of the performance, the harmonic melodies that accompany it, and the narrative that creates a captivating effect. Within the expansive realm of social media, an abundance of affirmative emotions manifests as virtual reverberations of happiness, satisfaction, affection, and genuine gratitude.

Sentiment analysis is a complex process that involves analyzing emotions expressed in text and material, with the aim of categorizing them into three main categories: positive, negative, or neutral.

Positive sentiment is characterized by a cheerful attitude towards the topic at hand, comprising feelings of enjoyment, approval, affection, or appreciation. On the other hand, negative sentiments serve to obscure, therefore exposing sentiments of disappointment, dissatisfaction, disapproval, rage, or discomfort. Furthermore, one must consider the tranquil domain of neutral feeling, a state of emotional balance in which objectivity holds paramount importance. This state is distinguished by the absence of overt emotional evaluations and frequently encompasses objective facts or descriptions. The potential applications of sentiment analysis are diverse, encompassing a wide range of human emotions. This study establishes a foundation within the complex and intricate realms of business, social media, market research, and brand monitoring, shedding light on the various ways in which individuals perceive and experience different aspects of life. Positive sentiment analysis plays a crucial role for artists, serving as a navigational tool that aids them in navigating the complex landscape of public feedback. This compilation provides a wealth of valuable perspectives, granting a glance into the shared recognition and accomplishments of their offerings, provisions, and corporate identities. Positive evaluations and criticism serve as more than just plaudits; they play a crucial role in assisting artists in honing their skills and adapting their approaches.

Positive feelings have revolutionary power within the digital sphere. They shape the reputation of an artist's brand, conferring upon it a strong and distinct identity. The ability to resonate with the audience results in the establishment of brand recognition, trust, and steadfast devotion. Within the domain of *Wayang*, an analysis of positive sentiment reveals the underlying connections that resonate with the emotional experiences of the audience, hence facilitating enhanced product development and ongoing improvement. When conducting surveys and feedback loops, the collection of positive feelings can be likened to the quantification of audience satisfaction levels. These pleasant sentiments indicate that individuals have not only undergone but also greatly relished their voyage. Furthermore, within the realm of marketing and advertising, favorable emotions function as valuable assets, enhancing the visual appeal and effectiveness of promotional products and campaigns. The positive testimonies and evaluations provided by contented viewers serve as a testament to the establishment of trust. During periods of instability, artists have the ability to utilize good mood as a guiding force to navigate away from pessimistic perspectives. By emphasizing the favorable aspects and highlighting their endeavors towards enhancement, individuals can effectively alter the course of the discourse. Sentiment analysis is a valuable tool for artists to gain insights into audience opinions and identify chances for distinction by monitoring the fluctuations of positive sentiment toward competition. In the realm of social media analytics, the concept of positive sentiment serves as a reflective mechanism that gauges the manner in which viewers perceive and respond to material. Social media analytics functions as a metric to assess the effectiveness of social campaigns and provides significant insights into customer perceptions, pricing strategies, market segmentation, and marketing approaches within the field of market research. The symphony of positive attitude serves as a catalyst for several areas, fostering not just applause but also the virtuosic display of comprehension, adaptability, and achievement. It serves as a reminder that inside the realm of emotions, the vibrant tones of positivity persistently reverberate.

4. Conclusion

In conclusion, the utilization of Support Vector Machine (SVM) analysis has demonstrated its remarkable precision in unraveling the intricate tapestry of sentiment across various domains. Positive sentiment, characterized by its cheerful disposition, has emerged as a pivotal influencer affecting diverse human endeavors, including commerce and the arts. A positive mood is paramount in guiding product development and shaping brand reputations in the business world. Sentiment analysis offers invaluable insights for corporations, allowing them to understand their clientele's emotions and opinions better, leading to enhanced customer satisfaction and more effective marketing strategies. Moreover, in the realm of arts and culture, as exemplified by *Wayang Climen*, sentiment analysis on social media platforms serves as a potent tool for preserving traditional art forms, nurturing cultural history, and inspiring a new generation's appreciation for the arts. Looking ahead, the versatility of sentiment analysis holds promise for application in diverse global artistic expressions, transcending boundaries and enriching the international cultural heritage, ultimately enhancing the overall human experience. Through its discerning methodology, sentiment analysis stands as a steadfast advocate, invigorating art and culture worldwide and fostering admiration for the rich tapestry of human creativity.

Acknowledgment

The author would like to thank the Faculty of Engineering, Faculty of Letters, Universitas Negeri Malang, Indonesia, and Faculty of Computer Science, Electronics, and Telecommunications, AGH University of Science and Technology, Poland, for the granted support.

Declarations

Author contribution	:	All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement	:	None of the authors have received any funding or grants from any institution or funding body for the research
Conflict of interest	:	The authors declare no conflict of interest
Additional information	:	No additional information is available for this paper

References

- [1] N. W. Somawati and I. W. Karja, "Contemporary Wayang Beber: Transformative Narration in the Discourse of Philosophy and Society," *Int. J. Soc. Sci. Hum. Res.*, vol. 6, no. 08, Aug. 2023, doi: [10.47191/ijsshr/v6-i8-52](https://doi.org/10.47191/ijsshr/v6-i8-52).
- [2] A. Arafahan, Saidin, and F. Y. Sitepu, "Copyright Protection Of Wayang Kulit As A Traditional Indonesian Culture," *POLICY, LAW, NOTARY Regul. ISSUES*, vol. 1, no. 4, pp. 1–8, Oct. 2022, doi: [10.55047/polri.v1i4.413](https://doi.org/10.55047/polri.v1i4.413).
- [3] A. Windarsih, "Traditional Art's Survival in the Digital Era: Puppet Performance on YouTube," in *Deconstructing Culture and Communication in the Global South*, 2023, pp. 141–157. doi: [10.4018/978-1-6684-8093-9.ch010](https://doi.org/10.4018/978-1-6684-8093-9.ch010)
- [4] A. Alacovska, P. Booth, and C. Fieseler, "A Pharmacological Perspective on Technology-Induced Organised Immaturity: The Care-giving Role of the Arts," *Bus. Ethics Q.*, vol. 33, no. 3, pp. 565–595, Jul. 2023, doi: [10.1017/beq.2022.39](https://doi.org/10.1017/beq.2022.39).
- [5] M. Muslih, A. Rohman, A. Ahmad, and A. Saifullah, "Sunan Kalijaga's Da'wah Strategy In Suluk Linglung And Its Implication To Indonesian Radicalism Movement," *El-Harakah (Terakreditasi)*, vol. 23, no. 1, pp. 1–19, Jun. 2021, doi: [10.18860/eh.v23i1.11672](https://doi.org/10.18860/eh.v23i1.11672).
- [6] R. U. Dona, "Character Wayang Timplong of Panji Asmarabangun Figure and its Relevance to the Pancasila Students ' Profile," vol. 15, pp. 2846–2856, 2023, doi: [10.35445/alishlah.v15i3.3177](https://doi.org/10.35445/alishlah.v15i3.3177).
- [7] M. Hänska and S. Bauchowitz, "Can social media facilitate a European public sphere? Transnational communication and the Europeanization of Twitter during the Eurozone crisis," *Soc. Media + Soc.*, vol. 5, no. 3, p. 205630511985468, Jul. 2019, doi: [10.1177/2056305119854686](https://doi.org/10.1177/2056305119854686).
- [8] B. Bimber and H. Gil de Zúñiga, "The unedited public sphere," *New Media Soc.*, vol. 22, no. 4, pp. 700–715, Apr. 2020, doi: [10.1177/1461444819893980](https://doi.org/10.1177/1461444819893980).
- [9] S. G. Purnama and D. Susanna, "Attitude to COVID-19 Prevention With Large-Scale Social Restrictions (PSBB) in Indonesia: Partial Least Squares Structural Equation Modeling," *Front. Public Heal.*, vol. 8, Oct. 2020, doi: [10.3389/fpubh.2020.570394](https://doi.org/10.3389/fpubh.2020.570394).
- [10] C. Respa and J. D. Imelda, "Digital Capital As Informal Social Protection In The Pandemic Covid-19's Era (Case Study Of The Wayang Kulit Association In Yogyakarta)," *JHSS (Journal Humanit. Soc. Stud.)*, vol. 5, no. 3, pp. 242–246, Oct. 2021, doi: [10.33751/jhss.v5i3.3994](https://doi.org/10.33751/jhss.v5i3.3994).
- [11] C. Baden, C. Pipal, M. Schoonvelde, and M. A. C. G. van der Velden, "Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda," *Commun. Methods Meas.*, vol. 16, no. 1, pp. 1–18, Jan. 2022, doi: [10.1080/19312458.2021.2015574](https://doi.org/10.1080/19312458.2021.2015574).
- [12] E. S. Rice, E. Haynes, P. Royce, and S. C. Thompson, "Social media and digital technology use among Indigenous young people in Australia: a literature review," *Int. J. Equity Health*, vol. 15, no. 1, p. 81, Dec. 2016, doi: [10.1186/s12939-016-0366-0](https://doi.org/10.1186/s12939-016-0366-0).
- [13] M. L. Khan, "Social media engagement: What motivates user participation and consumption on YouTube?," *Comput. Human Behav.*, vol. 66, pp. 236–247, Jan. 2017, doi: [10.1016/j.chb.2016.09.024](https://doi.org/10.1016/j.chb.2016.09.024).

-
- [14] L. Maunder, "Motivating people to stay at home: using the Health Belief Model to improve the effectiveness of public health messaging during the COVID-19 pandemic," *Transl. Behav. Med.*, vol. 11, no. 10, pp. 1957–1962, Oct. 2021, doi: [10.1093/tbm/ibab080](https://doi.org/10.1093/tbm/ibab080).
 - [15] I. Meloni and E. O. Allasso, "Going 'Viral,'" *J. World Pop. Music*, vol. 10, no. 1, Jul. 2023, doi: [10.1558/jwpm.26376](https://doi.org/10.1558/jwpm.26376).
 - [16] Kauffmann, Peral, Gil, Ferrández, Sellers, and Mora, "Managing Marketing Decision-Making with Sentiment Analysis: An Evaluation of the Main Product Features Using Text Data Mining," *Sustainability*, vol. 11, no. 15, p. 4235, Aug. 2019, doi: [10.3390/su11154235](https://doi.org/10.3390/su11154235).
 - [17] C. Zucco, B. Calabrese, G. Agapito, P. H. Guzzi, and M. Cannataro, "Sentiment analysis for mining texts and social networks data: Methods and tools," *WIREs Data Min. Knowl. Discov.*, vol. 10, no. 1, Jan. 2020, doi: [10.1002/widm.1333](https://doi.org/10.1002/widm.1333).
 - [18] 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, p. e0232525, May 2020, doi: [10.1371/journal.pone.0232525](https://doi.org/10.1371/journal.pone.0232525).
 - [19] C. Daryani, G. S. Chhabra, H. Patel, I. K. Chhabra, and R. Patel, "An Automated Resume Screening System Using Natural Language Processing And Similarity," in *Ethics And Information Technology*, Jan. 2020, pp. 99–103, doi: [10.26480/etit.02.2020.99.103](https://doi.org/10.26480/etit.02.2020.99.103).
 - [20] A. Briliani, B. Irawan, and C. Setianingsih, "Hate Speech Detection in Indonesian Language on Instagram Comment Section Using K-Nearest Neighbor Classification Method," in *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*, Nov. 2019, pp. 98–104, doi: [10.1109/IoTaIS47347.2019.8980398](https://doi.org/10.1109/IoTaIS47347.2019.8980398).
 - [21] A. Ishaq *et al.*, "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques," *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: [10.1109/ACCESS.2021.3064084](https://doi.org/10.1109/ACCESS.2021.3064084).
 - [22] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, "Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting," *Inf. Fusion*, vol. 54, pp. 128–144, Feb. 2020, doi: [10.1016/j.inffus.2019.07.006](https://doi.org/10.1016/j.inffus.2019.07.006).
 - [23] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, Sep. 2020, doi: [10.1016/j.neucom.2019.10.118](https://doi.org/10.1016/j.neucom.2019.10.118).
 - [24] C. Song, A. Shafieezadeh, and R. Xiao, "High-Dimensional Reliability Analysis with Error-Guided Active-Learning Probabilistic Support Vector Machine: Application to Wind-Reliability Analysis of Transmission Towers," *J. Struct. Eng.*, vol. 148, no. 5, May 2022, doi: [10.1061/\(ASCE\)ST.1943-541X.0003332](https://doi.org/10.1061/(ASCE)ST.1943-541X.0003332).
 - [25] X.-F. Song, Y. Zhang, D.-W. Gong, and X.-Z. Gao, "A Fast Hybrid Feature Selection Based on Correlation-Guided Clustering and Particle Swarm Optimization for High-Dimensional Data," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9573–9586, Sep. 2022, doi: [10.1109/TCYB.2021.3061152](https://doi.org/10.1109/TCYB.2021.3061152).
 - [26] S. Dong, "Multi class SVM algorithm with active learning for network traffic classification," *Expert Syst. Appl.*, vol. 176, p. 114885, Aug. 2021, doi: [10.1016/j.eswa.2021.114885](https://doi.org/10.1016/j.eswa.2021.114885).
 - [27] P. Sarang, "Naive Bayes," in *The Springer Series in Applied Machine Learning*, 2023, pp. 143–152. doi: [10.1007/978-3-031-02363-7_7](https://doi.org/10.1007/978-3-031-02363-7_7)
 - [28] R. Ardianto, T. Rivanie, Y. Alkhalifi, F. S. Nugraha, and W. Gata, "Sentiment Analysis On E-Sports For Education Curriculum Using Naive Bayes And Support Vector Machine," *J. Ilmu Komput. dan Inf.*, vol. 13, no. 2, pp. 109–122, Jul. 2020, doi: [10.21609/jiki.v13i2.885](https://doi.org/10.21609/jiki.v13i2.885).
 - [29] I. Ibrahim and A. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 10–19, Mar. 2021, doi: [10.38094/jastt20179](https://doi.org/10.38094/jastt20179).
 - [30] K. Maswadi, N. A. Ghani, S. Hamid, and M. B. Rasheed, "Human activity classification using Decision Tree and Naïve Bayes classifiers," *Multimed. Tools Appl.*, vol. 80, no. 14, pp. 21709–21726, Jun. 2021, doi: [10.1007/s11042-020-10447-x](https://doi.org/10.1007/s11042-020-10447-x).
-

-
- [31] H. Chen, S. Hu, R. Hua, and X. Zhao, "Improved naive Bayes classification algorithm for traffic risk management," *EURASIP J. Adv. Signal Process.*, vol. 2021, no. 1, p. 30, Dec. 2021, doi: [10.1186/s13634-021-00742-6](https://doi.org/10.1186/s13634-021-00742-6).
- [32] P. P. Surya and B. Subbulakshmi, "Sentimental Analysis using Naive Bayes Classifier," in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, Mar. 2019, pp. 1–5, doi: [10.1109/ViTECoN.2019.8899618](https://doi.org/10.1109/ViTECoN.2019.8899618).
- [33] R. van de Schoot *et al.*, "Bayesian statistics and modelling," *Nat. Rev. Methods Prim.*, vol. 1, no. 1, p. 1, Jan. 2021, doi: [10.1038/s43586-020-00001-2](https://doi.org/10.1038/s43586-020-00001-2).
- [34] A. B. Shaik and S. Srinivasan, "A Brief Survey on Random Forest Ensembles in Classification Model," in *Lecture Notes in Networks and Systems*, Singapore: Springer, 2019, pp. 253–260. doi: [10.1007/978-981-13-2354-6_27](https://doi.org/10.1007/978-981-13-2354-6_27)
- [35] T.-H. Lee, A. Ullah, and R. Wang, "Bootstrap Aggregating and Random Forest," in *Advanced Studies in Theoretical and Applied Econometrics*, 2020, pp. 389–429. doi: [10.1007/978-3-030-31150-6_13](https://doi.org/10.1007/978-3-030-31150-6_13)
- [36] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Idris, A. M. Bamhdi, and R. Budiarto, "CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection," *IEEE Access*, vol. 8, pp. 132911–132921, 2020, doi: [10.1109/ACCESS.2020.3009843](https://doi.org/10.1109/ACCESS.2020.3009843).
- [37] X. Zhou, P. Lu, Z. Zheng, D. Tolliver, and A. Keramati, "Accident Prediction Accuracy Assessment for Highway-Rail Grade Crossings Using Random Forest Algorithm Compared with Decision Tree," *Reliab. Eng. Syst. Saf.*, vol. 200, p. 106931, Aug. 2020, doi: [10.1016/j.res.2020.106931](https://doi.org/10.1016/j.res.2020.106931).
- [38] T. Zhu, "Analysis on the Applicability of the Random Forest," *J. Phys. Conf. Ser.*, vol. 1607, no. 1, p. 012123, Aug. 2020, doi: [10.1088/1742-6596/1607/1/012123](https://doi.org/10.1088/1742-6596/1607/1/012123).
- [39] G. Wang, B. Zhao, B. Wu, C. Zhang, and W. Liu, "Intelligent prediction of slope stability based on visual exploratory data analysis of 77 in situ cases," *Int. J. Min. Sci. Technol.*, vol. 33, no. 1, pp. 47–59, Jan. 2023, doi: [10.1016/j.ijmst.2022.07.002](https://doi.org/10.1016/j.ijmst.2022.07.002).
-