JOURNAL LA MULTIAPP

VOL. 06, ISSUE 06 (1436-1448), 2025
DOI: 10.37899/journallamultiapp.v6i6.2505

Application of SVM and Naive Bayes with PSO for the Classification of Saloka Amusement Park Reviews

Indira Alifia Putri[1], Khothibul Umam[1,] Maya Rini Handayani[1], Hery Mustofa[1]

[1]*Information Technology, Faculty of Science and Technology, UIN Walisongo Semarang, Indonesia*

**Corresponding Author: Khothibul Umam*
Email: khothibul_umam@walisongo.ac.id

Article Info

Article history:
Received 30 July 2025
Received in revised form 1 October 2025
Accepted 31 October 2025

Keywords:
Sentiment Analysis
Naive Bayes
Support Vector Machine
Particle Swarm Optimization
Tourism Reviews

Abstract

Visitor opinions on tourist destinations can be evaluated through sentiment analysis based on textual reviews. This study aimed to compare the performance of Support Vector Machine (SVM) and Naive Bayes (NB) algorithms in classifying visitor sentiments toward reviews of Saloka Theme Park, while also assessing the impact of parameter optimization using Particle Swarm Optimization (PSO). A total of 740 reviews were collected from the Traveloka platform and underwent text preprocessing. The optimization process targeted key parameters of each algorithm to improve the F1-score. Experimental results showed that the unoptimized SVM achieved an accuracy of 89 percent, while NB reached 86 percent. After applying PSO, SVM's accuracy dropped to 84 percent, whereas NB improved to 85 percent with more balanced classification across sentiment classes. These results recommend the integration of Naive Bayes with Particle Swarm Optimization as a potential approach for sentiment classification of tourism reviews, particularly in the case study of Saloka Theme Park.

Introduction

Tourism is one of the strategic sectors that makes a significant contribution to regional economic growth (Gunawan & Kuswanto, 2024; Laut et al., 2021; Li et al., 2024). Along with the development of information technology, the behavior of tourists in seeking information and providing assessments of tourist destinations has also changed. Nowadays, tourists tend to rely on user-generated content such as online reviews to determine the choice of tourist attractions. These reviews are available in large quantities on various digital platforms, one of which is Traveloka, which not only provides booking services but also houses users' opinions and experiences regarding the destinations they visit (Syaqufa & Kusumawati, 2024; Dewantra et al., 2024; Saputra et al., 2025).

Online reviews contain information that can reflect the level of satisfaction, quality of service, and visitor experience of a tourist attraction (Nguyen et al., 2023; Glaveli et al., 2023; Camilleri & Filieri, 2023). Therefore, the analysis of this data has great potential in supporting evaluation and decision-making by destination managers. However, the high volume of data makes the manual analysis process inefficient. In this context, (Widodo et al., 2024; Adepoju et al., 2022; Werner et al., 2021). the text classification approach is a relevant solution. Through classification, the system can automatically group reviews into specific classes, including supportive and critical responses, so that they reflect general public opinion.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

The object of the research raised was Saloka Theme Park, a themed theme park located in Central Java Province (Ryani & Soesanto, 2021; Shaumarli & Nurwitasari, 2024; Robinson, 2023). This destination is one of the most visited destinations and receives various reviews from visitors through the Traveloka platform. To analyse the review, popular (Choirunnisa et al., 2021) machine learning algorithms were used, namely the Support Vector Machine (SVM) and Naive Bayes (NB) methods. SVM is known to have a good ability to handle high-dimensional data with a maximum margin of separation between classes. Meanwhile, NB is an efficient probabilistic-based approach in handling large-scale text classification (Yulia & Arrow, 20211; Rizki et al., 2025; Jamil et al., 2022).

Although both algorithms have been widely used, challenges in achieving optimal classification results are still encountered, especially when the data has an imbalance of class distribution. To solve this problem, the Particle Swarm Optimization (PSO) algorithm is used, which is a method inspired by the social behavior of groups of living things such as flocks of birds or fish in an effort to find optimal solutions collectively. Through the implementation of PSO, important parameters in the classification model can be set automatically to improve overall performance.

The purpose of this study is to evaluate the performance of the SVM and Naive Bayes algorithms in the process of classifying the evaluation of Saloka Theme Park visitors collected from Traveloka, with the help of parameter optimization using the PSO algorithm. The stages carried out include data retrieval through web scraping, text pre-processing, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), model training, parameter optimization with PSO, and model performance evaluation using accuracy and f1-score metrics. It is hoped that the results of this study can produce an accurate text classification model and provide benefits for destination managers in understanding the needs and perceptions of visitors more deeply (Arnap et al., 2024; Srinivasan et al., 2023; Gregoriades et al., 2023).

Methods

This research was focused on achieving the goal of designing and comparing the performance of two classification models, namely *the Support Vector Machine (SVM)* and *Naive Bayes (NB)* classification methods which are widely used in text analysis which is each optimized

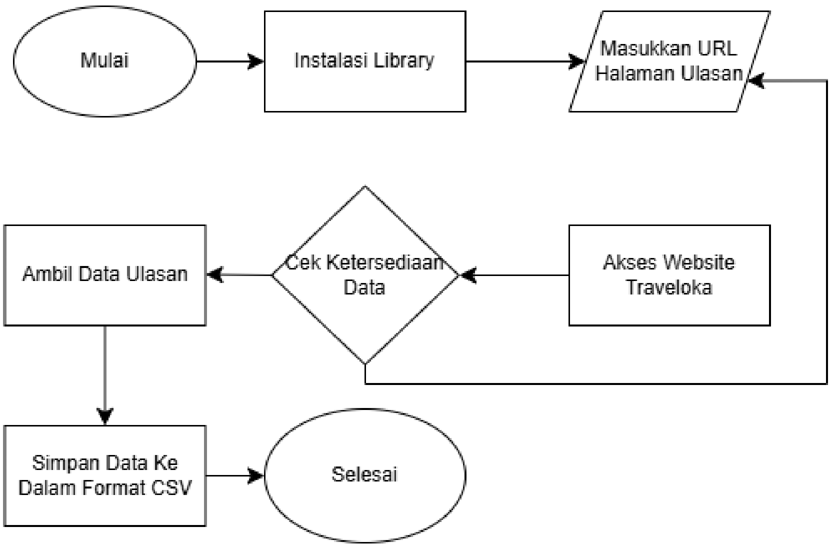

Figure 1. Diagram of Crawling Steps

ISSN: 2716-3865 (Print), 2721-1290 (Online)

This method included a number of systematic stages ranging from literature review, problem formulation, data collection, *text pre-processing*, to model performance evaluation. This approach is designed to be replicative and objectively implemented (Rukmana et al., 2023).

Problem Formulation

Formally, the problem of text classification in this study can be formulated as follows:

Given a review dataset:

$$D = \{(x_1, y_1), (x_{2,}y_2), \dots, (x_n, y_n)\} \tag{1}$$

where is a feature vector representation of the extracted third review with $x_i \in \mathbb{R}^m$ *Term Frequency–Inverse Document Frequency (TF-IDF),* and $\{0,1\}$ is the class label of the review (e.g. 0 for negative and 1 for positive).$y_i \in$

The main purpose of this study is to find the function of the hypothesis

$$h : \mathbb{R}^{\mathrm{m}} \to \{0,1\} \tag{2}$$

can map review features into appropriate classes, minimizing the *error function of the error function* E(h), , namely:

$$\mathrm{E(h)} = \frac{1}{n}\sum_{i=1}^{n} \parallel (h((x_i) \neq y_1) \tag{3}$$

With is the indicator function, which is valued at 1 if the argument is true, and 0 if false. $\parallel \{.\}$

Dataset Collection

The dataset was collected through a *scraping process* from the Traveloka review page for Saloka Theme Park destinations. The data collected amounted to seven hundred and forty reviews in the form of review text and star ratings as label references. Reviews with a rating of four or more are labeled positive, while a rating of two or less is labeled negative. Reviews with a rating of three were omitted to maintain clarity of the class. (Amini, 2023)


review_saloka.csv

	A	B
1	review	rate
2	Saat pergi hujan der	negative
3	Yang kurang hanya t	negative
4	Sangat puas liburan	positive
5	Staff-nya ramah	positive
6	seru main di sini. sta	positive
7	harus menganggap k	positive
8	Seharusnya tdk harus	negative
9	seruuu wahana luma	positive
10	Selamat pagi	positive
11	staff di lokasi sangat	positive
12	permainan bnyak. sta	positive
13	agus	positive
14	enyenangkan	positive
15	pelayanan tiket oke	positive
16	amazing place	positive
17	mantap.. rekomen ur	positive
18	Staff nya ramah sem	positive
19	Fasilitas semua lengl	positive
20	klo pas weekday wort	positive

Figure 2. Scraping dataset view

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Text Preprocessing

In text processing, a series of pre-processing stages are carried out to ensure optimal data quality before being applied to the classification model. The steps taken include:

First, *case folding.* The goal of this stage is to equalize the text format by converting the entire character to lowercase, so that unnecessary variations can be minimized.

Second, *tokenization.* This step aims to separate the text into simpler word elements, so that it can support the advanced analysis process more efficiently.

Third, *stopword removal.* This process aims to eliminate terms that do not provide significant analytical value, such as "which", "and", "in", so that only words with a high weight of information are retained.

Fourth, special character cleaning. At this stage, emojis, symbols, numbers, and other punctuation are removed to reduce noise in the text that may interfere with the analysis results.

Fifth, *voting. Stemming* is done by utilizing *the Literary* library to change words to their basic form, so that words that have the same root can be standardized and reduce the diversity of word forms.

Each of these steps is designed to minimize interference in the data and improve the accuracy of the model at the classification stage (Yuhandri et al., 2024), so that only important information is taken into account in subsequent analysis.

```
[ ] from nltk.tokenize import word_tokenize
    from nltk.corpus import stopwords
    import re, emoji
    from string import punctuation
    from sklearn.feature_extraction.text import TfidfVectorizer, CountVectorizer
    from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
```

Figure 3. Preprocessing text.

The pre-processing stages of the text are carried out using *Python libraries* such as *NLTK* and *Sastrawi.* This process involves *tokenization, stopword removal,* character normalization, and *stemming* (Munir et al., 2022) . As illustrated in Figure 3.

Table 1. Pre-Processing Process

Before Pre-processing	After Pre-processing
"The rides are really exciting, especially Disco Waves and Jet Coaster. There's a bit of hysteria too, but I don't dare to go up. Overall good"	"It's a great way to get rid of the hysteria, but I don't want to be the one to do it."
"I'm going to be with my friends over the weekend."	"Friends Are Coming Soon"
"The place is okay, good. However, to print tickets, you still have to queue up and there is an	"Okay, it's good to print the BLM queue ticket printing machine"

ISSN: 2716-3865 (Print), 2721-1290 (Online)

independent ticket printing machine."	
"It's not very good but it's not bad either. Yes, it's normal"	"It's Bad"
"The staff at the location were all very friendly. Ticket redemption & redemption is very easy. The queue at every ride is not too long?"	"Friendly location staff exchange redeem tickets easily queue up for rides"

Handling of Unbalanced Datasets

Label distribution is often uneven. Therefore, *an oversampling* or *undersampling approach* is used to balance the amount of data for each class so that the model is not biased towards the majority class. This study applied *the Synthetic Minority Oversampling Technique (SMOTE)* method to increase the amount of data in minority classes.

Classification Modeling

The classification model is carried out with two algorithms:

Naive Bayes (NB), with a probabilistic model based on *Bayes' Theorem:*

$$P(y|x) \propto P(y) \prod_{i=1}^{m} P(x_i|y) \tag{4}$$

Support Vector Machine (SVM), which searches *for optimal* hyperfields:

$$f(x) = sign(w^T x + b) \tag{5}$$

with w and b are the parameters of the model that are determined through the training process.

The Naive Bayes method and the Support Vector Machine (SVM) have been used extensively

It is used in text classification due to its ability to handle high-dimensional data. The Naive Bayes algorithm models the probability of a class based on a Bayesian Theorem such as Equation (4), while the SVM looks for the optimal hyperfield that separates data classes such as Equation (5). (Nugroho & Maharani, 2024)

Parameter Optimization with PSO

In order to optimize the performance of the classification model, the *Particle Swarm Optimization (PSO*) algorithm parameter optimization process is carried out.This method includes *a metaheuristic* approach that uses a population approach as the basis for its search inspired by the collective behavioral behavior of groups, such as flying creatures or fish, in search of the best position in the search space. In this study, *PSO* was used to adjust important parameters on two classification algorithms. The two algorithms used are *Naive Bayes* and *Support Vector Machine (SVM).*

In the *Naive Bayes algorithm,* the parameter that the optimization focuses on is *var_smoothing,* which controls how much variation value is added to each feature to prevent division by zero. The objective function used in *PSO* aims to maximize the *f1-score* value by minimizing its negative form. The *var_smoothing* value sought is in the range between 1× to 1×.$10^{-10}10^{-1}$

Meanwhile, for *SVM,* the optimized parameters are *C* and gamma on the *Radial Base Function (RBF) kernel* (Ramadhani et al., 2020) . Parameter *C* regulates the balance between model complexity and misclassification, while gamma controls the influence of a single data on the formation of a *hyperplane. PSO* works by generating a set of particles that represent a combination of parameter values, and then evaluating the performance of each particle using

ISSN: 2716-3865 (Print), 2721-1290 (Online)

the f1-score metric. Each iteration updates the position and velocity of the particles based on individual experiences and the best experiences of the population.

This optimization is carried out with a population (*swarmsize)* of 10 particles and a maximum of 10 iterations, to maintain process efficiency without reducing the accuracy of the results. The results of the *PSO* show the optimal parameter values used for the best model retraining. The performance of the model after optimization is assessed using a classification report that includes accuracy, precision, *recall,* and *f1-score.* This optimization process has been proven to increase the effectiveness of classification, especially when the data has an imbalance in distribution between classes.

Model Training

The model training process is carried out by utilizing data that has previously gone through the preprocessing stage. The training process is performed for all parameter combinations generated by *Particle Swarm Optimization (PSO),* and the best-performing model is selected based on *the highest* f1-score value. (Faisal et al., 2020)

Model Evaluation

To assess the performance of the model, *a confusion matrix* is used as an evaluation tool. Through this matrix, a number of important indicators such as accuracy, *precision,* sensitivity *(recall),* and *f1-score can be calculated. The confusion matrix* breaks down the model's prediction results into four components, namely accurate predictions of data belonging to the positive class *(True Positive/TP),* false predictions as positive (*False Positive/FP),* correct predictions for negative classes *(True Negative/TN),* and errors in predicting as negative *(False Negative/FN).* The *trained Support Vector Machine (SVM)* model was evaluated using several performance metrics, including: (Nurzaman et al., 2024)

Accuracy, which is the correct proportion of the prediction to all test data.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

Precision, which is the proportion of the correct positive prediction.

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall, is the ratio between positive data that is predicted correctly.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

F1-score, is the harmonious average value between precision and recall.

$$F1 - Score = 2 * \frac{presisi + recall}{presisi + recall} \tag{9}$$

Metrics, this serves to measure the extent of the model's ability to classify sentiment on review data.

Results and Discussion

This study was conducted with the intention of evaluating and comparing the performance of *the Support Vector Machine (SVM)* and *Naive Bayes (NB) algorithms* in the classification of sentiment to the reviews of Saloka Theme Park visitors, as well as to evaluate the influence of parameter optimization using *Particle Swarm Optimization (PSO).* Performance

ISSN: 2716-3865 (Print), 2721-1290 (Online)

assessments were conducted using accuracy, *precision, recall,* and *f1-score metrics,* which were calculated for positive and negative grades, respectively. (Diva Aliyah et al., 2024)

Prior to model training, review data has gone through a number of *pre-processing* stages to improve feature quality. This process includes case *folding,* word separation *(tokenization),* stopword *removal,* special character cleaning, and *stemming* using *the Literary library.* The results of this stage are shown in Table 1, which shows the reduction of irrelevant words and the equalization of the form of the root word, so that the model more easily recognizes sentiment patterns in the text.

Furthermore, because the distribution of labels in the dataset is unbalanced (the majority is positive), the *Synthetic Minority Oversampling Technique (SMOTE)* technique is used to multiply the data in the minority (negative) class. This process helps reduce the model's bias towards the dominant class, as seen in the increase in *negative class f1-scores* after optimization. (Stuttgart et al., 2023)

To further improve model performance, a parameter optimization process was carried out using *Particle Swarm Optimization (PSO).* In the *Naive Bayes algorithm,* the optimized parameters are *var_smoothing* in the range of 1e–11 to 1e–3, while in *SVM,* the *parameters C* (1–100) and *gamma* (1e–4–1) on the *RBF kernel* are adjusted. *The PSO* is run with 10 particles and 10 iterations, using *5-fold cross validation* as an internal evaluation to avoid *overfitting.*

After obtaining the best combination of parameters from the optimization results, a retraining process is carried out on each model. *Naive Bayes* works with a probabilistic approach based on word distribution, whereas *SVM* forms an optimal separator *hyperplane* in the feature vector space transformed by the kernel. The results of each combination both before and after *the PSO* were then quantitatively evaluated using accuracy, precision, *recall,* and *f1-score metrics.*

Classification Results Without Optimization

Table 2. Naive Bayes Evaluation before PSO

Sentiment Class	Precision	Recall	F1-Score	Support
Negative	0.72	0.74	0.73	38
Positive	0.91	0.90	0.90	110
Accuracy			0.86	148
Macro avg	0.81	0.82	0.82	148
Weighted avg	0.86	0.86	0.86	148

The Naive Bayes model that has not gone through the optimization process has shown adequate initial performance. Based on the data in Table 2, the accuracy achieved reached 86 percent, with an f1-score value of 0.90 in the positive class and 0.73 in the negative class. In addition, the recall value for the positive class was recorded at 0.90, which indicates the model's ability to identify most positive-tone reviews.

This result suggests that even in its baseline form, the Naive Bayes algorithm is already capable of performing effective sentiment classification, particularly in distinguishing positive sentiments. The high recall and f1-score for the positive class demonstrate that the model is sensitive to detecting positive reviews and rarely misses them, a feature that is especially valuable in contexts where capturing favorable opinions is prioritized, such as product review analysis or customer satisfaction monitoring. However, the relatively lower

ISSN: 2716-3865 (Print), 2721-1290 (Online)

f1-score for the negative class implies that the model still struggles to accurately classify negative sentiments, possibly due to class imbalance or the subtler linguistic cues present in negative reviews. This discrepancy indicates potential areas for improvement through optimization techniques such as hyperparameter tuning, feature selection, or text preprocessing refinement. In summary, while the initial results are promising and confirm the model's general robustness, further optimization would likely enhance its discriminatory power and balance its performance across both sentiment classes.

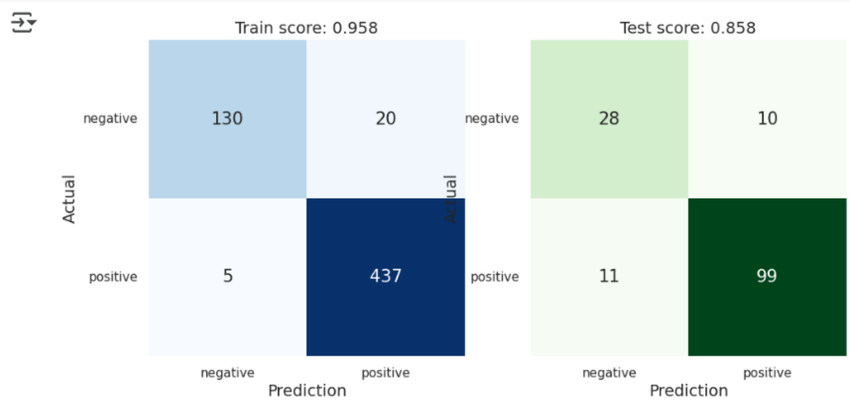

Figure 4. Confussion Matrix NB before PSO

Figure 4 shows the *confusion matrix* of the model. Most of the data in the positive class was correctly classified *(True Positives),* while there were some misclassifications in the negative class *(False Negatives),* which reflected an imbalance in the data.

Table 3. SVM evaluation before PSO

Sentiment Class	Precision	Recall	F1-Score	Support
Negative	0.80	0.74	0.77	38
Positive	0.91	0.94	0.92	110
Accuracy			0.89	148
Macro avg	0.86	0.84	0.85	148
Weighted avg	0.88	0.89	0.88	148

Meanwhile, the *unoptimized Support Vector Machine (SVM)* model (Table 3) yielded an accuracy of 89 percent, with *an f1-score* of 0.92 for the positive class and 0.77 for the negative class. This shows SVM's superiority in recognizing the majority class.

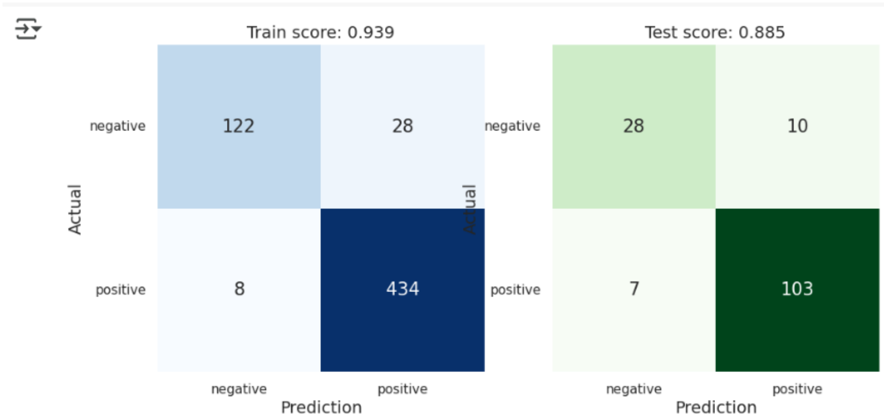

Figure 5. Confussion Matrix SVM before PSO

ISSN: 2716-3865 (Print), 2721-1290 (Online)

However, from the *confusion matrix* shown in Figure 5, it can be seen that despite the high accuracy, the model tends to be biased towards positive classes. There is an increase in *False Negatives* in the negative class, which indicates the model's suboptimal ability to handle unbalanced data.

Classification Results After PSO Optimization

Table 4. NB Evaluation after PSO

Sentiment Class	Precision	Recall	F1-Score	Support
Negative	0.76	0.69	0.72	42
Positive	0.88	0.92	0.90	106
Accuracy			0.85	148
Macro avg	0.82	0.80	0.81	148
Weighted avg	0.85	0.85	0.85	148

After the implementation of *Particle Swarm Optimization (PSO),* there was a change in the performance of both models. In the *optimized Naive Bayes* model , the accuracy decreases slightly to 85 percent, but the balance between classes increases. *The F1-score* for the negative class rose to 0.72, while the positive class remained high at 0.90 (Table 4). This suggests that *PSO* helps to increase the sensitivity of the model to minority classes.

Table 5. Evaluation of SVM after PSO

Sentiment Class	Precision	Recall	F1-Score	Support
Negative	0.82	0.55	0.66	42
Positive	0.84	0.95	0.89	106
Accuracy			0.84	148
Macro avg	0.83	0.75	0.78	148
Weighted avg	0.84	0.84	0.83	148

In contrast, the results on the *SVM model* after optimization experienced a decrease in performance. Accuracy dropped to 84 percent, and although the positive class recall was very high (0.95), *the negative class f1-score* decreased to 0.66 (Figure 7). This shows that the model is increasingly leaning towards positive classes.

Mechanism of Implementation of PSO in Parameter Optimization

The application *of Particle Swarm Optimization (PSO)* in this study aims to optimize the important parameters of two classification algorithms, namely *var smoothing* in *Naive Bayes* and *C* and *gamma* in *Support Vector Machine (SVM)* with Radial Base Function (RBF) *kernels.* Optimization is carried out with the aim of increasing *the f1-score value* of the model.

PSO works by forming an initial population of 10 particles, where each particle represents a random combination of parameter values in a predefined search space. The search process was carried out over 10 iterations. Each particle is evaluated using an objective function of a negative value of *the f1-score,* as *PSO* by default operates to minimize the value of the function. Particle performance evaluation was carried out using *k-fold cross validation* with k = 5 to avoid *overfitting* and obtain more stable results.

The parameter value range used in the optimization process is as follows: for *the Naive Bayes algorithm,* var_smoothing *parameter* is searched in the range 1e-11 to 1e-3. Whereas in *SVM,* parameter *C* is searched in the range of 1 to 100, and *gamma* in the range 1e-4 to 1. At each iteration, the position and velocity of the particles are updated based on the individual's best

ISSN: 2716-3865 (Print), 2721-1290 (Online)

experience *(pbest)* and the global best experience in the population *(gbest),* following the standard *PSO* formula.

After the iteration is complete, the best parameter combination of the optimization results is used to retrain the model with the entire training data. The retraining model was then evaluated using performance metrics such as accuracy, precision, *recall,* and *f1-score.* Based on the results of the evaluation that has been presented previously, *PSO* has been shown to help improve the balance of classification, especially in *the Naive Bayes model,* while in *SVM,* the effect does not always result in consistent performance improvements

Conclusion

This study is focused on a comparative analysis of the performance of the Support Vector Machine (SVM) and Naive Bayes (NB) algorithms in classifying sentiment from visitor reviews of Saloka Theme Park tourist attractions. The classification stage includes pre-processing data, giving word weights using the Term Frequency–Inverse Document Frequency (TF-IDF) approach, algorithm training, and parameter tuning using the Particle Swarm Optimization method.

The data used consisted of 740 reviews collected from the Traveloka platform. Based on the labeling criteria, reviews with ratings four and five are classified as positive sentiment, while ratings one and two are categorized as negative sentiment. Reviews with a rating of 3 were excluded from the dataset to maintain class clarity. As a result, the distribution of data becomes unbalanced, with a predominance of positive reviews. To overcome this inequality, the SMOTE (Synthetic Minority Oversampling Technique) method is applied to increase the number of samples in the minority (negative) class, so that the model can learn from more balanced data.

Based on the results of the experiment, the enhanced Naive Bayes algorithm using the Particle Swarm Optimization approach showed a more balanced classification ability in distinguishing between positive and negative sentiments. Although the accuracy of 85 percent is slightly lower than SVM without 89 percent optimization, Naive Bayes recorded an f1-score value of 0.90 for the positive class and 0.72 for the negative class, indicating the stability of performance in the face of unbalanced data. In contrast, the optimized SVM model experienced a decline in the ability to recognize negative reviews, as reflected in the lower recall and f1-score values for the class.

The results of the study show that the application of the optimized Naive Bayes algorithm using Particle Swarm Optimization is more appropriately used to handle the classification of sentiment on tourism review data that has an unbalanced distribution. This approach can be an effective alternative in supporting public perception analysis as well as strategic decision-making in the tourism sector.

References

Adepoju, A. H., Austin-Gabriel, B. L. E. S. S. I. N. G., Eweje, A. D. E. O. L. U. W. A., & Collins, A. N. U. O. L. U. W. A. P. O. (2022). Framework for automating multi-team workflows to maximize operational efficiency and minimize redundant data handling. *IRE Journals*, *5*(9), 663-664.

Amini, N. (2023). MENGGUNAKAN ALGORITMA KLASIFIKASI SUPPORT VECTOR MACHINE DENGAN OPTIMASI PARTICLE SWARM OPTIMIZATION (Studi Kasus : Indonesian Twitter Emotion Dataset) PROGRAM STUDI TEKNIK INFORMATIKA 2023 M / 1445 H.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Arnap, A., Magister, P., Informatika, T., & Pascasarjana, P. (2024). Optimasi Model Support Vector Machine dengan Particle Swarm Optimization untuk Mendeteksi Serangan Injeksi SQL (Studi Kasus : PT . Naisha Inspirasi Muslimah).

Camilleri, M. A., & Filieri, R. (2023). Customer satisfaction and loyalty with online consumer reviews: Factors affecting revisit intentions. *International Journal of Hospitality Management*, *114*, 103575. https://doi.org/10.1016/j.ijhm.2023.103575

Choirunnisa, M., Hidayat, N., & Santoso, E. (2021). Implementasi Metode Support Vector Machine Dengan Query Expansion Pada Klasifikasi Review Di Situs Traveloka. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 5(5), 1860–1865.

Dewantra, N. N., Sukarno, A., & Oetomo, H. (2025). The Influence of E-Service Quality, E-Wom and Discounts on Transportation Ticket Purchase Decisions or Accommodation Booking Services on Traveloka Users. *Jurnal EMA*, *10*(1), 66-74. https://dx.doi.org/10.51213/ema.v10i1.575

Diva Aliyah, Zilrahmi, Yenni Kurniawati, & Dina Fitria. (2024). Optimization of Sentiment Analysis for MBKM Program using Naïve Bayes with Particle Swarm Optimization. UNP Journal of Statistics and Data Science, 2(4), 393–400. https://doi.org/10.24036/ujsds/vol2-iss4/220

Faisal, A., Alkhalifi, Y., Rifai, A., & Gata, W. (2020). Analisis Sentimen Dewan Perwakilan Rakyat Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization. JOINTECS (Journal of Information Technology and Computer Science), 5(2), 61. https://doi.org/10.31328/jointecs.v5i2.1362

Glaveli, N., Manolitzas, P., Palamas, S., Grigoroudis, E., & Zopounidis, C. (2023). Developing effective strategic decision-making in the areas of hotel quality management and customer satisfaction from online ratings. *Current Issues in Tourism*, *26*(6), 1003-1021. https://doi.org/10.1080/13683500.2022.2048805

Gregoriades, A., Pampaka, M., Herodotou, H., & Christodoulou, E. (2023). Explaining tourist revisit intention using natural language processing and classification techniques. *Journal of Big Data*, *10*(1), 60. https://doi.org/10.1186/s40537-023-00740-5

Gunawan, A. H., & Kuswanto, V. (2024). Eksplorasi Algoritma Support Vector Machine untuk Analisis Sentimen Destinasi Wisata di Indonesia. 7(2). https://doi.org/10.32877/bt.v7i2.1810

Jamil, M. L., Pais, S., & Cordeiro, J. (2022). Detection of Dangerous Events on Social Media: A Perspective Review. *arXiv preprint arXiv:2204.01351*. https://doi.org/10.48550/arXiv.2204.01351

Laut, L. T., Sugiharti, R. R., & Panjawa, J. L. (2021). Does tourism sector matter in regional economic development. *Geo Journal of Tourism and Geosites*, *37*(3), 832-837.

Li, Y., Liu, Y., & Solangi, Y. A. (2024). Analysis of factors and strategies for the implementation of sustainable tourism in a green economic structure in China. *Journal of Cleaner Production*, *434*, 140011. https://doi.org/10.1016/j.jclepro.2023.140011

Munir, A., Atika, E. P., & Indraswari, A. D. (2022). Analisis Sentimen pada review hotel menggunakan metode pembobotan dan klasifikasi. Jnanaloka, 3(1), 33–38. https://doi.org/10.36802/jnanaloka.2022.v3-no1-33-38

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Nguyen, H. T., Le, A. T. T., Phan, A. C., & Hoang, T. D. L. (2023). A multi-perspective approach of international tourist satisfaction in tourism service: from big data perspective. *Journal of Asia business studies*, *17*(4), 850-872. https://doi.org/10.1108/JABS-03-2022-0090

Nugroho, B. S., & Maharani, W. (2024). Support Vector Machine and Naïve Bayes for Personality Classification Based on Social Media Posting Patterns. Building of Informatics, Technology and Science (BITS), 6(3), 1717–1731. https://doi.org/10.47065/bits.v6i3.6411

Nurzaman, N., Suarna, N., & Prihartono, W. (2024). Analisis Sentimen Ulasan Aplikasi Threads Di Google Playstore Menggunakan Algoritma Naïve Bayes. JATI (Jurnal Mahasiswa Teknik Informatika), 8(1), 967–974. https://doi.org/10.36040/jati.v8i1.8708

Purnamasari, D., Anshary, M. A. K., & Rianto, R. (2023). Particle Swarm Optimization dan Genetic Algorithm untuk analisis sentimen pemekaran Papua di Twitter berbasis Support Vector Machine. AITI, 20(2), 177–190. https://doi.org/10.24246/aiti.v20i2.177-190

Ramadhani, Arnia, F., & Muharar, R. (2020). Klasifikasi Otomatis Motif Tekstil Menggunakan Support Vector Machine Multi Kelas. Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK), 7(1), 99–108. https://doi.org/10.25126/jtiik.202071428

Rizki, A. M., Bustami, B., & Anshari, S. F. (2025). Comparison of Support Vector Machine and Naïve Bayes Algorithms in Sentiment Analysis of Tiktokshop Application User Reviews. *Journal of Renewable Energy, Electrical, and Computer Engineering*, *5*(1), 18-29. https://doi.org/10.29103/jreece.v5i1.21342

Robinson, K. (2023). A Mining Theme Park—Cultural Competition and the Politics of Identity Claims in an Indonesian Mining Town. *The Asia Pacific journal of anthropology*, *24*(5), 365-380. https://doi.org/10.1080/14442213.2023.2277401

Rukmana, P. U., Pratiwi, O. N., & Fakhrurroja, H. (2023). Perbandingan Analisis Sentimen Aplikasi Traveloka dan Tiket.com pada Twitter dengan Metode Support Vector Machine. Jurnal Sistem Cerdas, 6(3), 241–250. https://doi.org/10.37396/jsc.v6i3.350

Ryani, L. S., & Soesanto, H. (2021). Factors Affecting Revisit Intention Through Customer Satisfaction in Theme Park. *International Journal of Economics, Business and Accounting Research (IJEBAR)*, *5*(2). https://doi.org/10.29040/ijebar.v5i2.2294

Saputra, I., & Suta, P. W. P. (2025). GOVERNANCE OF EXPERIENCE-BASED TOURISM NARRATIVES: REPRESENTATION OF NUSA PENIDA TOURISM PRODUCTS IN DIGITAL MARKETPLACES. *Dinamika Governance: Jurnal Ilmu Administrasi Negara*, *15*(2). https://doi.org/10.33005/jdg.v15i2.5294

Shaumarli, A., & Nurwitasari, R. (2024). Exploring satisfaction of amusement parks in Indonesia. *Journal of Applied Sciences in Travel and Hospitality*, *7*(2), 94-104. https://doi.org/10.31940/jasth.v7i2.94-104

Srinivasan, J., Niranjanee, M., Nandhana, C., & Azhagiri, M. (2023, December). A Text Classification in Tourism Reviews with the Hybrid Transformer-Attention Model in the Information Management of Smart Tourism. In *International Conference on Cognitive Computing and Cyber Physical Systems* (pp. 479-488). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-7371-8_37

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Syaqufa, A., & Kusumawati, N. (2024). How User-Generated Content Influences Hotel Booking Decisions on Traveloka. *Mandalika Journal of Business and Management Studies*, *2*(2), 63-84. https://doi.org/10.59613/mjbms.v3i1.132

Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, *41*, 100514. https://doi.org/10.1016/j.accinf.2021.100514

Widodo, A. E., Wati, F. F., & Hidayati, N. (2024). Implementasi Algoritma SVM Dan Naive Bayes Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Onedrive. 10(2), 131–138.

Yuhandri, Y., Sovia, R., Syaiffullah, A., Yenila, F., & Permana, R. (2024). Penerapan Natural Language Processing Pada Sistem Chatbot Sebagai Helpdesk Obyek Wisata Menggunakan Metode Naïve Bayes. Jurnal Infortech, 5(2), 210–218. https://doi.org/10.31294/infortech.v5i2.20911

Yulia, E. R., & Solecha, K.-. (2021). Implementasi Particle Swarm Optimization (PSO) pada Analysis Sentiment Review Aplikasi Trafi menggunakan Algoritma Naive Bayes (NB). Jurnal Teknik Komputer, 7(1), 25–29. https://doi.org/10.31294/jtk.v7i1.9078

ISSN: 2716-3865 (Print), 2721-1290 (Online)