Enhancing Machine Learning Algorithms Using Particle Swarm Optimization (PSO) to Analyze Coffeeshops in Padang City

Ilmawati [1], Nadya Alinda Rahmi[2], Rahmatia Wulan Dari[3]
[1,2,3,] Faculty Computer Science, Universitas Putra Indonesia YPTK Padang
ilmawati@upiyptk.ac.id

Submitted: November 10, 2025; accepted: December 14, 2025

Abstract

Sentiment analysis of customer reviews is a crucial component in understanding consumer perception of coffeeshop services, especially in Padang City. This study aims to enhance text classification performance by employing two machine learning algorithms, Support Vector Machine (SVM) and Naïve Bayes (NB) and optimizing them using Particle Swarm Optimization (PSO). To address data imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The dataset consists of customer reviews categorized into three sentiment classes: negative, neutral, and positive. Evaluation results indicate that the optimized SVM+PSO model achieved the highest accuracy of 78.57%, while NB+PSO achieved 73.21%. The implementation of PSO successfully improved the accuracy and f1-score compared to baseline models. This study demonstrates that combining SMOTE and PSO can significantly enhance the performance of multi-class sentiment classification in local business domains such as coffeeshops.

Keywords*: sentiment analysis, machine learning, SMOTE, PSO, coffeeshop.*

1. Introduction

Coffee shops have become an important part of urban lifestyle, including in the city of Padang, West Sumatra. Places such as Kubik Coffee are popular destinations for enjoying a relaxing atmosphere while savoring quality coffee. The popularity of these coffee shops has led many customers to leave comments or reviews on platforms such as Google Review. These reviews provide valuable insights into customer experiences regarding product quality, service, price, location, and ambiance. From a business perspective, analyzing customer feedback is highly useful for improving service quality, understanding customer preferences, and supporting better strategic decision-making.

Sentiment analysis is one of the approaches that can be used to automatically identify customer opinions from review data. Machine learning methods have been widely applied in sentiment analysis because they are able to process large amounts of textual data efficiently and classify opinions into specific sentiment categories [1]. Among various machine learning algorithms, Support Vector Machine (SVM) and Naïve Bayes (NB) are commonly used in text classification tasks. SVM is known for its ability to find an optimal hyperplane for separating data classes, especially in high-dimensional feature spaces [2]. Meanwhile, Naïve Bayes is a probabilistic algorithm that is simple, efficient, and suitable for text data with many features [3].

Several previous studies have demonstrated the use of machine learning algorithms for sentiment analysis in various domains. Munawaroh and Ridhoi [4] applied Naïve Bayes using Orange to analyze public sentiment toward the risks of IKN development in East Kalimantan. The study showed that Naïve Bayes could be used effectively for public opinion classification. However, the research did not compare the performance of Naïve Bayes with other algorithms and did not address the problem of class imbalance. Fremmuzar and Baita [5] evaluated several SVM kernels for sentiment analysis of Telkomsel services on Twitter. Their study showed that SVM could achieve strong classification performance, but the research focused mainly on kernel testing and did not apply an optimization technique to improve model parameters. Sitorus and Zufria [6] used Naïve Bayes to classify user reviews of the Shopee application on the Play Store. Although the method produced acceptable results, the study suggested that data balancing techniques and additional algorithms could be used to improve classification performance. Ditami et al. [7] also applied SVM for sentiment analysis related to marketplace promotional events on Twitter and obtained promising results. However, their research did not explore further optimization methods or data balancing strategies.

Based on these previous studies, it can be observed that SVM and Naïve Bayes have been widely used and proven effective in sentiment analysis. However, most prior research still focused on using conventional machine learning models without sufficient attention to two important issues: data imbalance and parameter optimization. In real-world customer review datasets, sentiment classes are often imbalanced because positive reviews usually dominate neutral and negative reviews. This condition may cause the model to be biased toward the majority class and reduce its ability to recognize minority classes. Therefore, the Synthetic Minority Over-sampling Technique (SMOTE) is needed to balance the number of samples in each class [8].

In addition to data imbalance, the performance of machine learning models is also influenced by the selection of optimal parameters. SVM requires appropriate parameter settings such as C and gamma, while Naïve Bayes requires suitable smoothing parameters to improve classification performance. If these parameters are not properly selected, the model may not achieve optimal results. Particle Swarm Optimization (PSO) is one of the metaheuristic optimization algorithms that can be used to search for optimal parameter combinations. PSO is inspired by the social behavior of bird flocks or fish schools in finding the best solution [9]. Previous research has shown that PSO can improve the performance of machine learning models by optimizing their parameters [10], including in sentiment analysis tasks involving SVM [11].

Although previous studies have applied SVM, Naïve Bayes, SMOTE, and PSO in different contexts, there is still limited research that combines SMOTE and PSO to improve multi-class sentiment classification of coffee shop customer reviews, particularly in the local business context of Padang City. This research gap becomes the main basis for conducting the present study. Therefore, this study proposes the integration of SMOTE and PSO to improve the performance of SVM and Naïve Bayes in classifying customer review sentiments into negative, neutral, and positive classes. The main objective of this study is to evaluate whether PSO-based parameter optimization can enhance the performance of SVM and Naïve Bayes after the dataset is balanced using SMOTE. The results of this study are expected to contribute to the development of more accurate sentiment classification models and provide useful insights for coffee shop businesses in understanding customer perceptions.

2. Method

This study was conducted to develop and evaluate sentiment classification models for coffee shop customer reviews in Padang City. The proposed research flow consists of several main stages, namely data collection, sentiment labeling, text preprocessing, feature extraction, data balancing using SMOTE, model development using Support Vector Machine (SVM) and Naïve Bayes (NB), parameter optimization using Particle Swarm Optimization (PSO), and model evaluation. The complete research flow is illustrated in Fig. 1.

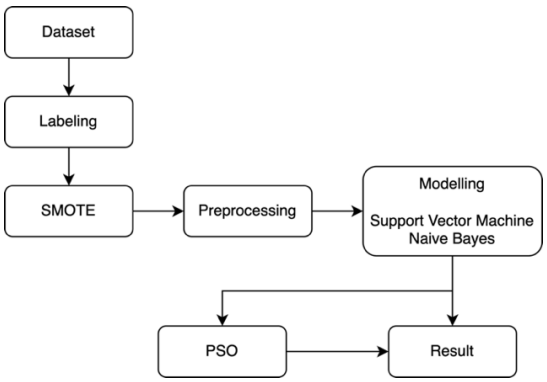


Fig. 1 Research Methodology

The research flow in Fig. 1 shows that the process begins with collecting customer review data and ends with evaluating the classification performance of the optimized models. Each stage is designed to ensure that the textual review data can be transformed into structured numerical features and classified into sentiment categories accurately.

2.1. Data Collection

The first stage of this research is data collection. The dataset used in this study consists of 500 customer review data obtained from Google Review for several coffee shops in Padang City. The collected

data include customer review text and sentiment labels. The review text is used as the main input feature because it contains customer opinions related to service quality, product quality, place atmosphere, price, and overall customer experience.

The dataset was classified into three sentiment classes, namely negative, neutral, and positive. The negative class represents reviews that contain complaints, dissatisfaction, or unfavorable customer experiences. The neutral class represents reviews that do not clearly express satisfaction or dissatisfaction. Meanwhile, the positive class represents reviews that contain satisfaction, appreciation, or favorable customer experiences toward the coffee shop.

Before the data balancing process, the dataset distribution was imbalanced. The negative class consisted of 39 reviews, the neutral class consisted of 161 reviews, and the positive class consisted of 300 reviews. This distribution shows that positive reviews dominated the dataset, while negative reviews were the minority class. Such an imbalance can affect the classification model because the model may tend to learn more from the majority class and have difficulty recognizing minority sentiment patterns.

2.2. Preprocessing

Text preprocessing is performed to clean and standardize the review text before it is processed by machine learning algorithms [12]. Customer reviews often contain unnecessary characters, inconsistent capitalization, punctuation, symbols, numbers, and informal expressions. Therefore, preprocessing is required to reduce noise and produce cleaner text data. The preprocessing stages in this study include:

a. Case folding, which converts all letters into lowercase form to ensure consistency between words [13].
b. Cleaning, which removes punctuation marks, special characters, numbers, URLs, and irrelevant symbols [14].
c. Tokenization, which splits sentences into individual words or tokens [15].
d. Stopword removal, which removes common words that do not have strong meaning for classification [16].
e. Normalization, which converts informal words, abbreviations, or slang into standard forms [17].
f. Stemming, which reduces words to their root forms [18].

2.3. Feature Extraction Using TF-IDF

After preprocessing, the cleaned text data are converted into numerical features using Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF is used to measure the importance of a word in a document based on how frequently the word appears in a review and how rare the word is across all reviews [19].

2.4. Data Blancing using SMOTE

Before applying the classification models, the distribution of sentiment classes was first analyzed to identify whether the dataset had an imbalance problem. Class imbalance is an important issue in sentiment classification because the model may become biased toward the majority class and fail to recognize minority classes properly [20]. Therefore, SMOTE was applied to balance the number of samples in each sentiment class [21]. The comparison of class distribution before and after applying SMOTE is shown in Fig. 2.

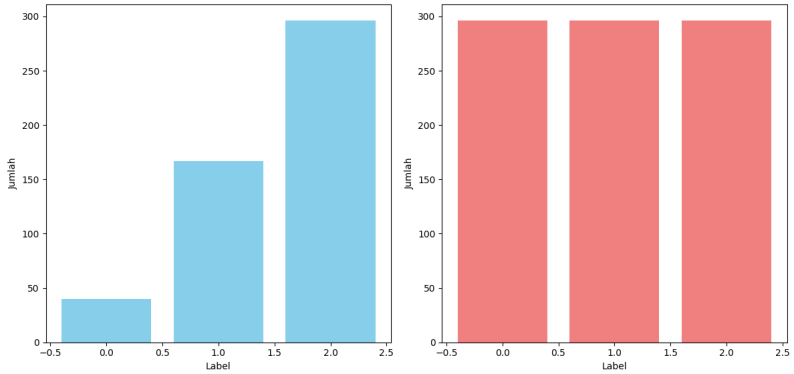


Fig. 2 Label Distribution Before and After SMOTE

.

Fig. 2 shows the distribution of sentiment labels before and after applying the Synthetic Minority Over-sampling Technique (SMOTE). Before SMOTE was applied, the dataset had an imbalanced class distribution. Label 2, which represents positive sentiment, had the highest number of data samples, while label 0, which represents negative sentiment, had the lowest number of samples. Label 1, which represents neutral sentiment, was positioned between the two classes. This condition indicates that the dataset was dominated by positive reviews, which could cause the classification model to learn more from the majority class and reduce its ability to correctly recognize minority classes.

After SMOTE was applied, the number of samples in each class became balanced. The negative, neutral, and positive classes had the same number of data samples. SMOTE works by generating synthetic data for minority classes based on the nearest neighbors of existing samples. This process helps the model receive a more equal learning opportunity for each sentiment class. Therefore, the application of SMOTE is expected to reduce classification bias toward the majority class and improve the model's ability to identify negative and neutral reviews more fairly.

2.5. Classification using Naïve Bayes

Naïve Bayes is used as one of the classification models in this study because it is simple, efficient, and widely used in text classification [22]. Naïve Bayes predicts the class of a review based on the probability of the words appearing in each sentiment class [23]. The classification process in Naïve Bayes can be expressed as:

$$Predicted\ class = argmax\ P(class)\ \times\ P(word \mid class) \quad (1)$$

In this formula, P(class) represents the probability of each sentiment class, namely negative, neutral, and positive. Meanwhile, P(words | class) represents the probability that the words in a review appear in a certain sentiment class. The class with the highest probability value is selected as the final sentiment prediction.

2.6. Classification using Support Vector Machine

Support Vector Machine is also used in this study because it performs well in high-dimensional data such as text classification. SVM works by finding the optimal hyperplane that separates data points into different sentiment classes. For binary classification, the decision function of SVM can be expressed as:

$$f(x) = w.x + b \quad (2)$$

In this formula, x represents the TF-IDF feature vector of a review, w represents the weight vector, and b represents the bias value. The value of f(x) determines the position of a review toward the separating hyperplane. In this study, SVM is used to classify reviews into three sentiment classes: negative, neutral, and positive.

2.7. Parameter Optimization using Particle Swarm Optimization

Particle Swarm Optimization is applied to improve the performance of SVM and Naïve Bayes by searching for the best parameter combinations [21], [24]. In this study, PSO is used to optimize the smoothing parameter (\alpha) in Naïve Bayes and the (C) and gamma parameters in SVM.

PSO works by moving a group of particles in the search space. Each particle represents a possible solution or parameter combination. The position and velocity of each particle are updated using the following equations:

$$v_{new} = w \times v_{old} + c1 \times r1 \times (pbest - x) + c2 \times (gbest - x) \quad (3)$$

In this formula, v_{new} is the updated velocity of a particle, v_{old} is the previous velocity, w is the inertia weight, c1 and c2 are learning factors, r1 and r2 are random values, pbest is the best solution found by an individual particle, gbest is the best solution found by all particles, and x is the current position of the particle. Each particle represents a possible parameter combination. The best parameter combination is selected based on the highest model performance, such as accuracy or F1-score.

2.8. Model Evaluation

The final stage is model evaluation. The performance of each model is measured using accuracy, precision, recall, and F1-score. These metrics are used to evaluate how well the models classify customer reviews into negative, neutral, and positive sentiment classes.

3. Result and Discussion

This section presents the empirical findings and a comprehensive discussion of each result. The goal is to explain the outcomes in sufficient depth to address the problem statements outlined in the Introduction. The study evaluates performance gains for two machine learning algorithms—Naive Bayes (NB) and Support Vector Machine (SVM)—in classifying coffee shop customer reviews from Padang. We compare baseline (non-optimized) models against models optimized with Particle Swarm Optimization (PSO). Evaluation metrics include accuracy, precision, recall, F1-score, supported by confusion matrices and label distribution plots.

3.1. Evaluation of the Naive Bayes Model

The first model evaluated in this study is Naïve Bayes. This model was used as a baseline classifier because it is simple, efficient, and commonly applied in text classification tasks. The evaluation was conducted to observe how well Naïve Bayes classified coffee shop customer reviews into three sentiment classes: negative, neutral, and positive. The confusion matrix of the Naïve Bayes model is shown in Fig. 3.

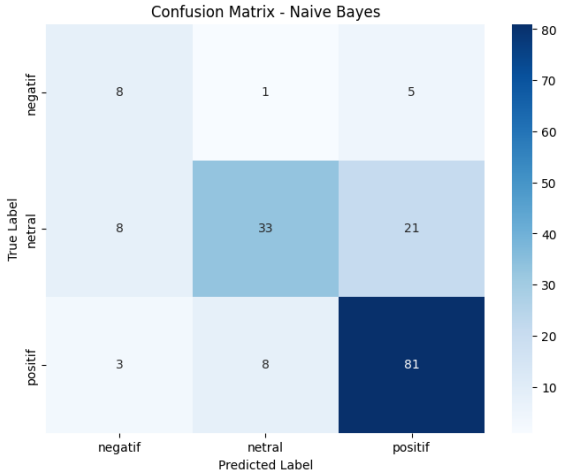


Fig. 3 Confusion Matrix - Naïve Bayes

Fig. 3 shows the classification results of the Naïve Bayes model for each sentiment class. The matrix compares the actual labels and the predicted labels. The diagonal values represent correctly classified data, while the non-diagonal values represent misclassified data.

For the negative class, the model correctly classified 8 reviews as negative. However, 1 negative review was incorrectly predicted as neutral, and 5 negative reviews were incorrectly predicted as positive. This result shows that the model still has difficulty recognizing negative sentiment patterns, especially because the number of negative data samples is relatively small compared to other classes.

For the neutral class, the model correctly classified 33 reviews as neutral. However, 8 neutral reviews were predicted as negative, and 21 neutral reviews were predicted as positive. This indicates that the neutral class is often confused with the positive class. This may occur because neutral reviews sometimes contain words or expressions that are not strongly negative but may still resemble positive opinions.

For the positive class, the model achieved the best performance. A total of 81 positive reviews were correctly classified as positive. Only 3 positive reviews were predicted as negative, and 8 positive reviews were predicted as neutral. This shows that Naïve Bayes is more effective in identifying positive sentiment than negative and neutral sentiment. Overall, the confusion matrix indicates that the model tends to perform better on the positive class, while classification errors are still found in the negative and neutral classes.

To provide a more detailed quantitative evaluation, the classification report of the Naïve Bayes model is presented in Fig. 4. This report includes precision, recall, F1-score, and support for each sentiment class.

```
Accuracy: 0.73
Classification Report:
              precision    recall  f1-score   support

           0       0.42      0.57      0.48        14
           1       0.79      0.53      0.63        62
           2       0.76      0.88      0.81        92

    accuracy                           0.73       168
   macro avg       0.65      0.66      0.64       168
weighted avg       0.74      0.73      0.72       168
```

Fig. 4 Classification Report - Naïve Bayes

Fig. 4 presents the classification performance of the Naïve Bayes model. The model achieved an overall accuracy of 0.73, meaning that approximately 73% of the total test data were classified correctly. From a total of 168 test data, the model correctly classified most of the reviews, especially in the positive class.

For class 0, which represents negative sentiment, the model obtained a precision of 0.42, recall of 0.57, and F1-score of 0.48 with 14 data samples. The low precision value indicates that several reviews predicted as negative were actually from other classes. Although the recall is slightly higher, the F1-score shows that the model's performance on the negative class is still limited.

For class 1, which represents neutral sentiment, the model achieved a precision of 0.79, recall of 0.53, and F1-score of 0.63 with 62 data samples. The high precision indicates that when the model predicts a review as neutral, the prediction is often correct. However, the lower recall shows that many actual neutral reviews were classified into other classes, especially positive sentiment.

For class 2, which represents positive sentiment, the model achieved the best performance with a precision of 0.76, recall of 0.88, and F1-score of 0.81 with 92 data samples. The high recall value indicates that the model was able to identify most positive reviews correctly. This result is consistent with the confusion matrix, where the positive class had the highest number of correct predictions.

The macro average values show a precision of 0.65, recall of 0.66, and F1-score of 0.64, indicating that the model performance across all classes is moderate. Meanwhile, the weighted average values are higher, with precision of 0.74, recall of 0.73, and F1-score of 0.72, because the positive class has a larger number of samples. Overall, the Naïve Bayes model provides acceptable baseline performance, but its ability to classify negative and neutral sentiment still needs improvement. Therefore, optimization using Particle Swarm Optimization is applied in the next stage to improve the classification performance.

3.2. Naive Bayes + PSO Evaluation Model

After evaluating the baseline Naïve Bayes model, Particle Swarm Optimization was applied to improve the model performance by searching for the best smoothing parameter. Based on the optimization process, the best alpha value obtained was 0.9038. The classification report of the Naïve Bayes model after PSO optimization is shown in Fig. 5.

```
Akurasi: 73.21%

Classification Report with optimized alpha (0.9038007922722548):
              precision    recall  f1-score   support

           0       0.42      0.57      0.48        14
           1       0.79      0.55      0.65        62
           2       0.76      0.88      0.82        92

    accuracy                           0.73       168
   macro avg       0.66      0.67      0.65       168
weighted avg       0.75      0.73      0.73       168
```

Fig. 5 Classification Report Naïve Bayes + PSO

Fig. 5 shows the performance of the Naïve Bayes model after optimization using Particle Swarm Optimization. The optimized Naïve Bayes model achieved an accuracy of 73.21%, which is slightly higher than the baseline Naïve Bayes accuracy. This indicates that PSO was able to find a better smoothing parameter for the Naïve Bayes model, although the improvement was relatively small.

For class 0, which represents negative sentiment, the model obtained a precision of 0.42, recall of 0.57, and F1-score of 0.48 with 14 data samples. These values are the same as the baseline Naïve Bayes model, indicating that PSO optimization did not significantly improve the model's ability to classify

negative reviews. This result shows that negative sentiment remains the most difficult class to identify, possibly because the number of negative reviews is smaller and the textual patterns are more varied.

For class 1, which represents neutral sentiment, the model achieved a precision of 0.79, recall of 0.55, and F1-score of 0.65 with 62 data samples. Compared with the baseline Naïve Bayes model, the recall and F1-score of the neutral class increased slightly. This means that after PSO optimization, the model became better at recognizing actual neutral reviews, although some neutral reviews were still classified into other sentiment classes.

For class 2, which represents positive sentiment, the model obtained a precision of 0.76, recall of 0.88, and F1-score of 0.82 with 92 data samples. This is the best performance among all classes. The high recall value indicates that the optimized Naïve Bayes model was able to correctly identify most positive reviews. This result also shows that positive sentiment has clearer textual patterns compared to negative and neutral sentiment.

The macro average values show a precision of 0.66, recall of 0.67, and F1-score of 0.65. These values indicate that the overall performance across the three sentiment classes improved slightly after optimization. Meanwhile, the weighted average values show a precision of 0.75, recall of 0.73, and F1-score of 0.73. The weighted average is higher because the positive class has more data samples than the other classes.

Overall, the Naïve Bayes model optimized using PSO produced a slight improvement compared with the baseline model. The improvement can be seen particularly in the neutral and positive classes. However, the negative class still showed limited performance. Therefore, although PSO contributes to improving the Naïve Bayes model, further improvement is still needed, especially to increase the model's ability to recognize negative sentiment.

3.3. Evaluation of the Support Vector Machine Model

. After evaluating the Naïve Bayes model, the next model tested in this study is Support Vector Machine. SVM was used because it is known to perform well in high-dimensional text classification tasks, especially when review data have been transformed into TF-IDF feature vectors. The evaluation was conducted to observe the ability of SVM to classify coffee shop customer reviews into three sentiment classes: negative, neutral, and positive. The confusion matrix of the SVM model is shown in Fig. 6.

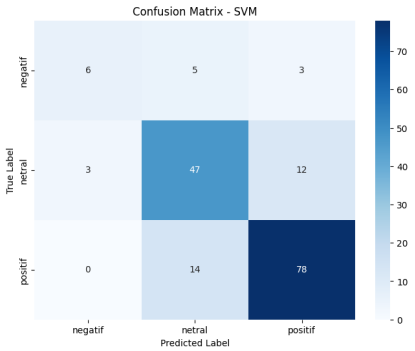


Fig. 6 Confusion Matrix - SVM

Fig. 6 shows the classification results of the SVM model based on the actual and predicted sentiment classes. The diagonal values represent the number of correctly classified reviews, while the non-diagonal values represent classification errors.

For the negative class, the SVM model correctly classified 6 reviews as negative. However, 5 negative reviews were incorrectly classified as neutral, and 3 negative reviews were incorrectly classified as positive. This result shows that SVM still has difficulty identifying negative reviews. The limited number of negative samples may cause the model to have fewer learning patterns for this class.

For the neutral class, the model correctly classified 47 reviews as neutral. However, 3 neutral reviews were predicted as negative, and 12 neutral reviews were predicted as positive. This indicates that the SVM model has a fairly good ability to recognize neutral sentiment, although some neutral reviews are still confused with positive sentiment. This can happen because neutral reviews may contain general expressions that are closer to positive comments than negative comments.

For the positive class, the SVM model achieved the best classification result. A total of 78 positive reviews were correctly classified as positive. However, 14 positive reviews were predicted as neutral, while

no positive review was predicted as negative. This result indicates that SVM can recognize positive sentiment patterns quite well. The absence of positive reviews classified as negative also shows that the model can distinguish positive and negative sentiment more clearly.

Overall, the confusion matrix shows that the SVM model performs better than Naïve Bayes, especially in classifying neutral and positive reviews. However, the negative class remains the most challenging class to classify correctly. This indicates that even though SVM has strong classification ability, the model still needs further optimization to improve its performance across all sentiment classes.

To provide a more detailed quantitative evaluation, the classification report of the SVM model is presented in Fig. 7. This report shows the precision, recall, F1-score, and support values for each sentiment class.

```
Akurasi: 77.98%
Classification Report for SVM:
              precision    recall  f1-score   support

           0       0.67      0.43      0.52        14
           1       0.71      0.76      0.73        62
           2       0.84      0.85      0.84        92

    accuracy                           0.78       168
   macro avg       0.74      0.68      0.70       168
weighted avg       0.78      0.78      0.78       168
```

Fig. 7 Classification Report SVM

Fig. 7 presents the performance evaluation of the SVM model. The model achieved an overall accuracy of 77.98%, or approximately 78%. This means that the SVM model correctly classified most of the test data. Compared with the baseline Naïve Bayes model, SVM produced better overall performance.

For class 0, which represents negative sentiment, the model obtained a precision of 0.67, recall of 0.43, and F1-score of 0.52 with 14 data samples. The precision value indicates that when the model predicts a review as negative, the prediction is relatively reliable. However, the low recall value shows that many actual negative reviews were still classified into other classes, especially neutral. This means that the model is still not sensitive enough in detecting negative sentiment.

For class 1, which represents neutral sentiment, the model achieved a precision of 0.71, recall of 0.76, and F1-score of 0.73 with 62 data samples. These values show that SVM performs reasonably well in classifying neutral reviews. The recall value of 0.76 indicates that most neutral reviews were successfully recognized by the model. However, some neutral reviews were still misclassified as positive, which shows that the boundary between neutral and positive sentiment remains difficult to separate.

For class 2, which represents positive sentiment, the model achieved the best performance with a precision of 0.84, recall of 0.85, and F1-score of 0.84 with 92 data samples. These values indicate that SVM is highly effective in recognizing positive reviews. The high precision and recall values show that the model can classify positive sentiment consistently and accurately.

The macro average values show a precision of 0.74, recall of 0.68, and F1-score of 0.70. These values indicate that the model performance across all classes is good, although the negative class still lowers the overall balance of performance. Meanwhile, the weighted average values show a precision of 0.78, recall of 0.78, and F1-score of 0.78. These weighted values are higher because the positive class has more data samples and is classified more accurately by the model.

Overall, the SVM model achieved better performance than the Naïve Bayes model, particularly in the neutral and positive classes. The accuracy increased to 77.98%, showing that SVM is more suitable for this sentiment classification task. However, the relatively low recall and F1-score for the negative class indicate that further optimization is still needed. Therefore, Particle Swarm Optimization is applied in the next stage to improve the parameter selection of SVM and enhance the overall classification performance.

3.4. Evaluation of the Support Vector Machine + PSO Model

After evaluating the baseline SVM model, Particle Swarm Optimization was applied to improve the model performance by optimizing the main SVM parameters. Based on the optimization process, the best parameter values obtained were C = 355.3433 and gamma = 0.0283. These parameters were then used to build the optimized SVM model. The classification report of the SVM model after PSO optimization is shown in Fig. 8.

```
Akurasi: 78.57%

Classification Report with optimized C (355.343315388919) and gamma (0.028315525604811122):
              precision    recall  f1-score   support

           0       0.62      0.36      0.45        14
           1       0.75      0.74      0.75        62
           2       0.82      0.88      0.85        92

    accuracy                           0.79       168
   macro avg       0.73      0.66      0.68       168
weighted avg       0.78      0.79      0.78       168
```

Fig. 8 Classification Report SVM + PSO

Fig. 8 presents the classification performance of the SVM model after optimization using Particle Swarm Optimization. The optimized SVM model achieved an accuracy of 78.57%, which is slightly higher than the baseline SVM accuracy of 77.98%. This result indicates that PSO was able to improve the overall classification performance by finding a better combination of SVM parameters.

For class 0, which represents negative sentiment, the model obtained a precision of 0.62, recall of 0.36, and F1-score of 0.45 with 14 data samples. The precision value shows that most reviews predicted as negative were relatively correct. However, the low recall value indicates that many actual negative reviews were still classified into other sentiment classes. This means that negative sentiment remains the most difficult class to identify, even after PSO optimization.

For class 1, which represents neutral sentiment, the model achieved a precision of 0.75, recall of 0.74, and F1-score of 0.75 with 62 data samples. These values show that the optimized SVM model performed well in recognizing neutral reviews. Compared with the baseline SVM model, the F1-score for the neutral class increased, indicating that PSO helped improve the balance between precision and recall for this class.

For class 2, which represents positive sentiment, the model achieved the best performance, with a precision of 0.82, recall of 0.88, and F1-score of 0.85 with 92 data samples. The high recall value indicates that most positive reviews were successfully identified by the model. This result shows that the optimized SVM model is highly effective in detecting positive sentiment patterns in coffee shop customer reviews.

The macro average values show a precision of 0.73, recall of 0.66, and F1-score of 0.68. These values indicate that the overall performance across all classes is still influenced by the low recall of the negative class. Meanwhile, the weighted average values show a precision of 0.78, recall of 0.79, and F1-score of 0.78. The weighted average is higher because the positive class has more data samples and is classified more accurately by the model.

Overall, the SVM model optimized with PSO produced the highest accuracy in this study, reaching 78.57%. The improvement was mainly observed in the neutral and positive classes. However, the performance for the negative class decreased compared with the baseline SVM model. This result shows that PSO can improve overall accuracy, but further improvement is still needed to increase the model's ability to recognize negative sentiment more effectively.

3.5. Comparative Analysis

Overall, the experimental results show that Support Vector Machine produced better performance than Naïve Bayes in classifying coffee shop customer reviews into negative, neutral, and positive sentiment classes. The baseline Naïve Bayes model achieved an accuracy of 73%, while Naïve Bayes optimized with PSO slightly improved the accuracy to 73.21%. This indicates that PSO was able to improve the Naïve Bayes model, although the improvement was relatively small.

The SVM model achieved better performance than Naïve Bayes, with a baseline accuracy of 77.98%. After applying PSO optimization, the SVM model obtained the highest accuracy of 78.57%. This result shows that PSO successfully improved the SVM parameter selection, particularly through the optimization of C and gamma values. Among all tested models, SVM + PSO produced the best classification performance.

Based on the classification reports, the positive sentiment class achieved the highest performance in almost all models. This indicates that positive reviews have clearer textual patterns and are easier for the models to recognize. In contrast, the negative sentiment class had the lowest recall and F1-score, showing that negative reviews were more difficult to classify correctly. This may be caused by the smaller number of negative samples and the variation of negative expressions in customer reviews.

.

In conclusion, the combination of SMOTE and PSO can improve machine learning performance in multi-class sentiment classification. SMOTE helps balance the data distribution, while PSO helps find better model parameters. However, further improvement is still needed, especially to increase the model's ability to recognize negative sentiment more accurately. This study also obtained quite good results compared to previous research on sentiment analysis as shown in Table 1.

Table 1. Comparison With Previous Studies

No	Researcher	Model	Accuracy
1	[25]	Support Vector machine	72.12%
2	[26]	Support Vector machine	59%
3	[27]	Support Vector machine	75,5%
		Naïve Bayes	75%
4	[28]	Naïve Bayes	72,14%
5	[29]	Support Vector machine	70,82%
		Naïve Bayes	63,02%
6	**This Research**	**Support Vector machine + PSO**	**78.57%**
		Naïve Bayes + PSO	**73.21%**

Table I presents a comparison of the model performance in this study with a number of previous studies that used Support Vector Machine (SVM) and Naïve Bayes algorithms in text classification, particularly in the context of sentiment analysis or customer reviews.

From the table, it can be seen that most previous studies used SVM as the main model. For example, the study referred to by [25] showed an accuracy of 72.12%, and [26] achieved only 59%, which is the lowest value among all results. Study [27] used two models, namely SVM (75.5%) and Naïve Bayes (75%), which showed competitive results, but still below the model proposed in this study.

Study [28], which used only Naïve Bayes, obtained an accuracy of 72.14%, while [29] applied both models with SVM accuracy of 70.82% and Naïve Bayes of 63.02%. This study recorded the best results of all those compared, namely:

- SVM + PSO with an accuracy of 78.57%
- Naïve Bayes + PSO with an accuracy of 73.21%

This shows that integrating the PSO optimization algorithm into the SVM and Naïve Bayes models provides a significant improvement in classification performance. This improvement is most likely due to PSO's ability to optimally adjust hyperparameters, which was not carried out in most previous studies.

Overall, this table reinforces the finding that using metaheuristic optimization techniques such as PSO can not only improve accuracy, but also provide a competitive advantage over conventional approaches previously used in similar studies.

4. Conclusion

This study successfully demonstrates that applying the Particle Swarm Optimization (PSO) optimization method can improve the performance of two popular machine learning algorithms, namely Support Vector Machine (SVM) and Naïve Bayes (NB), in sentiment classification of coffee shop customer reviews in the city of Padang. With the support of data balancing using SMOTE, the models can avoid bias toward the majority class and produce a fairer classification distribution. The evaluation results show that SVM optimized with PSO achieved the highest accuracy of 78.57%, followed by NB+PSO with an accuracy of 73.21%. These findings reinforce the effectiveness of combining SMOTE and PSO in addressing imbalance issues and automatically tuning model parameters. For future research, it is recommended that this approach be tested in other review domains, such as the culinary or tourism sectors, and that the results be compared with other optimization methods such as Grid Search or Genetic Algorithm to obtain more comprehensive outcomes.

Reference

[1] S. U. Hassan, J. Ahamed, and K. Ahmad, "Analytics of machine learning-based algorithms for text classification," *Sustainable Operations and Computers*, vol. 3, pp. 238–248, Jan. 2022, doi: 10.1016/j.susoc.2022.03.001.

[2] Y. A. Singgalen, "Performance Evaluation of Sentiment Classification Models: A Comparative Study of NBC, SVM, and DT with SMOTE," *Media Online*, vol. 4, no. 5, pp. 2539–2548, Apr. 2024, doi: 10.30865/klik.v4i5.1827.

.

[3] S. V RamaKrishna Reddy, D. V L N Somayajulu, and A. R. Dani, "Classification of Movie Reviews Using Complemented Naive Bayesian Classifier," *International Journal of Intelligent Computing Research (IJICR)*, vol. 2, no. 3, pp. 148–153, 2011, doi: 10.20533/ijicr.2042.4655.2011.0019.

[4] A. Munawaroh and R. Ridhoi, "Sentiment Analysis dengan Naïve Bayes Berbasis Orange Terhadap Resiko Pembangunan IKN," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 1, pp. 587–592, Feb. 2024, doi: 10.36040/jati.v8i1.8454.

[5] P. Fremmuzar and A. Baita, "Uji Kernel SVM dalam Analisis Sentimen Terhadap Layanan Telkomsel di Media Sosial Twitter," *Komputika : Jurnal Sistem Komputer*, vol. 12, no. 2, pp. 57–66, Sep. 2023, doi: 10.34010/komputika.v12i2.9460.

[6] R. A. Sitorus and I. Zufria, "Application of the Naïve Bayes Algorithm in Sentiment Analysis of Using the Shopee Application on the Play Store," *Digital Zone*, vol. 15, no. 1, pp. 53–65, 2024, doi: 10.31849/digitalzone.v15i1.19828.

[7] G. R. Ditami, E. F. Ripanti, and H. Sujaini, "Implementasi Support Vector Machine untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja pada Marketplace," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 3, pp. 508–516, Dec. 2022, doi: 10.26418/jp.v8i3.56478.

[8] M. M. R. Mubarak, Y. H. Chrisnanto, and P. N. Sabrina, "Implementation of Random Forest Using Smote and Smoteenn in Customer Churn Classification in E-Commerce," *Enrichment: Journal of Multidisciplinary Research and Development*, vol. 1, no. 8, pp. 463–477, Nov. 2023, doi: 10.55324/enrichment.v1i8.69.

[9] N. W. Susanto and H. Suparwito, "SVM-PSO Algorithm for Tweet Sentiment Analysis #BesokSenin," *Indonesian Journal of Information Systems (IJIS)*, vol. 6, no. 1, pp. 36–47, 2023, doi: 10.24002/ijis.v6i1.7551.

[10] Y. Zhou, N. Wang, and W. Xiang, "Clustering Hierarchy Protocol in Wireless Sensor Networks Using an Improved PSO Algorithm," *IEEE Access*, vol. 5, pp. 2241–2253, 2017, doi: 10.1109/ACCESS.2016.2633826.

[11] R. Wati, S. Ernawati, and I. Maryani, "Optimasi Parameter Pso Berbasis Svm Untuk Analisis Sentimen Review Jasa Maskapai Penerbangan," *Evolusi: Jurnal Sains dan Manajemen*, vol. 8, no. 2, pp. 64–71, 2020, doi: 10.31294/evolusi.v8i2.9248.

[12] A. N. Ulfah, M. K. Anam, N. Y. S. Munti, S. Yaakub, and M. B. Firdaus, "Sentiment Analysis of the Convict Assimilation Program on Handling Covid-19," *JUITA: Jurnal Informatika*, vol. 10, no. 2, pp. 209–216, 2022, doi: 10.30595/juita.v10i2.12308.

[13] M. K. Anam, M. B. Firdaus, F. Suandi, Lathifah, T. Nasution, and S. Fadly, "Performance Improvement of Machine Learning Algorithm Using Ensemble Method on Text Mining," in *ICFTSS 2024 - International Conference on Future Technologies for Smart Society*, Kuala Lumpur: Institute of Electrical and Electronics Engineers Inc., Sep. 2024, pp. 90–95. doi: 10.1109/ICFTSS61109.2024.10691363.

[14] S. G. Begum and P. K. Sree, "Drug Recommendation Using a 'Reviews and Sentiment Analysis' By a Recurrent Neural Network," *Indonesian Journal of Multidisciplinary Science*, vol. 2, no. 9, pp. 3085–3094, Jun. 2023, doi: 10.55324/ijoms.v2i9.530.

[15] H. Taherdoost and M. Madanchian, "Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research," *Computers*, vol. 12, no. 2, pp. 1–15, Feb. 2023, doi: 10.3390/computers12020037.

[16] S. R. Aisy and B. Prasetiyo, "Sentiment Analysist of the TPKS Law on Twitter Using InSet Lexicon with Multinomial Naïve Bayes and Support Vector Machine Based on Soft Voting," *Recursive Journal of Informatics*, vol. 1, no. 2, pp. 93–101, Sep. 2023, doi: 10.15294/rji.v1i2.68324.

[17] B. Kane *et al.*, "CNN-LSTM-CRF for aspect-based sentiment analysis: A joint method applied to French reviews," in *ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence*, SciTePress, 2021, pp. 498–505. doi: 10.5220/0010382604980505.

[18] H. Basri, M. B. S. Junianto, and I. Kusyadi, "Enhancing Usability Testing Through Sentiment Analysis: A Comparative Study Using SVM, Naive Bayes, Decision Trees and Random Forest," *Jurnal Teknologi Sistem Informasi dan Aplikasi*, vol. 7, no. 4, pp. 1603–1610, Oct. 2024, doi: 10.32493/jtsi.v7i4.45117.

[19] B. P. Zen, I. Susanto, and D. Finaliamartha, "TF-IDF Method and Vector Space Model Regarding the Covid-19 Vaccine on Online News," *SinkrOn*, vol. 6, no. 1, pp. 69–79, Oct. 2021, doi: 10.33395/sinkron.v6i1.11179.

[20] P. P. Putra, M. K. Anam, A. S. Chan, A. Hadi, N. Hendri, and A. Masnur, "Optimizing Sentiment Analysis on Imbalanced Hotel Review Data Using SMOTE and Ensemble Machine Learning

Techniques," *Journal of Applied Data Sciences*, vol. 6, no. 2, pp. 936–951, May 2025, doi: 10.47738/jads.v6i2.618.

[21] H. Hamdani, S. Susanti, L. Lathifah, M. K. Anam, and R. Pradipta, "Optimization of Machine Learning Models in Student Graduation Prediction Systems Using Ensemble Learning with PSO and SMOTE," *sinkron*, vol. 9, no. 4, pp. 3149–3158, Oct. 2025, doi: 10.33395/sinkron.v9i4.15335.

[22] R. Rahmiati, D. Paradila, M. K. Anam, M. Mardainis, and M. Machdalena, "Application of Naïve Bayes Algorithm for Non-Cash Food Assistance Recipients in Kampar Regency," *Sinkron*, vol. 8, no. 1, pp. 433–441, Jan. 2023, doi: 10.33395/sinkron.v8i1.12032.

[23] S. Susanti, A. A. Sari, M. K. Anam, M. Jamaris, and H. Hamdani, "Sistem Prediksi Keuntungan Influencer Pengguna E-Commerce Shopee Affiliates menggunakan Metode Naïve Bayes," *Edumatic: Jurnal Pendidikan Informatika*, vol. 6, no. 2, pp. 394–403, Dec. 2022, doi: 10.29408/edumatic.v6i2.6787.

[24] T. B. Sasongko, "Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Klasifikasi Jalur Minat SMA)," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 2, no. 2, pp. 244–253, 2016, doi: 10.28932/jutisi.v2i2.476.

[25] G. R. Ditami, E. F. Ripanti, and H. Sujaini, "Implementasi Support Vector Machine untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja pada Marketplace," *JEPIN (Jurnal Edukasi & penelitian Informatika)*, vol. 8, no. 3, pp. 508–516, 2022, doi: 10.26418/jp.v8i3.56478.

[26] M. H. Arfian *et al.*, "Analisis Sentimen Pada Media Sosial Menggunakan Metode Support Vector Machine," *Jurnal Ilmu Teknik dan Komputer*, vol. 9, no. 1, pp. 1–6, 2025, doi: 10.22441/jitkom.v9i1.001.

[27] E. Eskiyaturrofikoh and R. R. Suryono, "Analisis Sentimen Aplikasi X Pada Google Play Store Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine (SVM)," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1408–1419, Aug. 2024, doi: 10.29100/jipi.v9i3.5392.

[28] N. A. Maulana and D. Darwis, "Perbandingan Metode SVM dan Naïve Bayes untuk Analisis Sentimen pada Twitter tentang Obesitas di Kalangan Gen Z," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 3, pp. 655–666, Mar. 2025, doi: 10.52436/1.jpti.691.

[29] W. Ningsih, B. Alfianda, R. Rahmaddeni, and D. Wulandari, "Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 556–562, Feb. 2024, doi: 10.57152/malcom.v4i2.1253.