Institut Riset dan Publikasi Indonesia (IRPI)
IJATIS: Indonesian Journal of Applied Technology and Innovation Science
Journal Homepage: https://journal.irpi.or.id/index.php/ijatis
Vol. 2 Iss. 1 February 2025, pp: 42-52
ISSN(P): 3032-7466 | ISSN(E): 3032-7474

Comparison of Supervised Learning Algorithms for Predicting Airline Passenger Satisfaction

Agil Irman Fadri[1*], Abid Zahfran[2], Taylan Irak[3], Naufal Helga Firjatullah[4], Jelita Ekaraya Herianto[5]

[1,2]Department of Information Systems, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[3]Faculty of Engineering and Natural Sciences, Sabanci University, Turkey
[4]Bachelor Media and Communication, Asia Pacific University, Malaysia
[5]Bachelor of Mass Communication, Taylors University, Malaysia

E-Mail: [1]12250314181@students.uin-suska.ac.id, [2]12250314503@students.uin-suska.ac.id , [3]taylan.irak@sabanciuniv.edu, [4]naufalf1005@gmail.com, [5]ekarayaj@gmail.com

Received Dec 29th 2024; Revised Feb 25th 2025; Accepted Feb 27th 2025; Available Online Feb 28th 2025, Published Feb 28th 2025
Corresponding Author: Agil Irman Fadri

Abstract

Service quality and airline passenger satisfaction are the main factors in business success in the modern aviation industry. This research compares the performance of supervised learning algorithms, namely K-Nearest Neighbor (K-NN), Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine (SVM), to predict passenger satisfaction. The k-fold cross-validation method with k=20 was applied to ensure comprehensive model evaluation by dividing the data proportionally. Using a high value of k was chosen to optimize the stability of the model estimates, reduce the risk of overfitting, and produce more accurate evaluation metrics. The research results show that the Random Forest algorithm provides the highest accuracy of 95.78%, followed by Decision Tree (93.82%) and K-NN (91.85%). These results indicate that the Random Forest algorithm better classifies passenger satisfaction than other algorithms. This research confirms the potential of machine learning algorithms as a practical solution in data analysis to support strategic decision-making, especially for airlines that want to improve customer experience.

Keyword: Decision Tree, K-Nearest Neighbors, Naïve Bayes, Random Forest, Support Vector Machine

1. INTRODUCTION

Flight experience is now the main aspect that differentiates one airline from another. According to Baker (2013), airlines allocate resources to meet high-quality standards, maintain superior service, and ensure passenger satisfaction [1]. Since the 2000s, technological advances have provided access to a wealth of information, making airline customers more aware of the market and competitors. With increased bargaining power, customers now have a dominant role in company-consumer relationships [2]. In the aviation industry, service quality and passenger satisfaction are increasingly considered crucial elements in business performance and key strategies for achieving a competitive advantage [3]. The services most directly felt by airline customers are the in-flight services provided by flight attendants and the facilities on board the aircraft. Passengers usually rate airlines based on their satisfaction with the service experience [4]. Evaluation of airline service quality can be done by measuring consumer satisfaction using data mining techniques. One approach used is a classification model to predict passenger satisfaction [5].

Various data mining algorithms have been applied to predict airline passenger satisfaction levels with varying results. In this research, five machine learning algorithms were selected: Random Forest, K-Nearest Neighbors (K-NN), Decision Tree, Naïve Bayes, and Support Vector Machine (SVM). The selection of these algorithms is based on their unique advantages. Random Forest is known for its ability to visualize feature importance, ease of interpretation, and resistance to overfitting [6]. K-NN is a simple algorithm that does not require initial training and performs well without assuming data distribution [7]. Decision Tree is efficient for large datasets, easy to visualize, and quickly generates decision rules [8]. Naïve Bayes is effective in handling both quantitative and qualitative data, requires minimal data, and is simple to implement [9]. Meanwhile, SVM offers clear visualization for data separation, making it suitable for pattern recognition tasks [10], though it poses challenges when applied to large-scale problems [11]. Several previous studies have explored the effectiveness of these algorithms in predicting airline passenger satisfaction. A study by A.C.Y. Hong et al. (2023) utilized Random Forest, K-NN, Decision Tree, and Naïve Bayes, reporting that Random Forest achieved the highest accuracy of 89.29% with a 75:25 train-test split, followed by K-NN (87.20%), Decision Tree (82%), and Naïve Bayes (76.80%) [12]. Another study by B. Herawan Hayadi et al. (2021) demonstrated similar results, with Random Forest achieving 99.1% accuracy when using a max depth of 17 [4]. Despite these findings, prior studies have not incorporated the SVM algorithm, which has been widely recognized for its robustness in classification tasks such as pattern recognition and satellite image classification. This study differs from previous research by incorporating SVM alongside Random Forest, K-NN, Decision Tree, and Naïve Bayes in airline passenger satisfaction prediction. While prior research has largely focused on Random Forest as the most effective algorithm, this study explores whether SVM can offer competitive performance. Additionally, by comparing all five algorithms within the same dataset, this research aims to provide a more comprehensive evaluation of their predictive capabilities.

This research aims to predict future passenger satisfaction using machine learning algorithms. This research will use data taken from Kaggle to compare classification algorithms such as K-NN, Naive Bayes, Decision Tree, Random Forest, and SVM. The main goal is to determine the algorithm that provides the highest accuracy for classifying passenger satisfaction levels. Apart from that, this research also tries to identify features that have a strong correlation with the level of passenger satisfaction. The results of this study can serve as a guide for airlines to improve their services to match customer expectations while strengthening their competitiveness in the market.

2. MATERIAL AND METHOD

This research methodology was designed to compare the performance of classification algorithms (Decision Tree, K-NN, Naive Bayes, Random Forest, and SVM) in predicting airline passenger satisfaction. The stages of the research methodology are explained as figure 1.

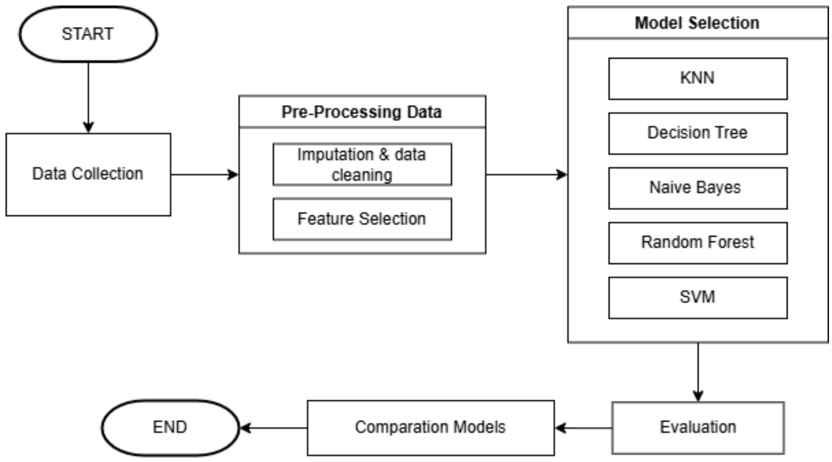


Figure 1. Research Methodology

2.1. Data Collection

The dataset used in this research is publicly available and has been utilized by other researchers, as it is sourced from Kaggle. The dataset, titled "Airline Passenger Satisfaction", is a classification dataset containing approximately 25,000 records with 25 attributes that capture various aspects of customer profiles, travel experiences, and satisfaction levels. Several studies have previously used this dataset to analyze and predict airline passenger satisfaction. For instance, A.C.Y. Hong et al. (2023) applied multiple machine learning algorithms, including Random Forest, K-NN, Decision Tree, and Naïve Bayes, and reported that Random Forest achieved the highest accuracy of 89.29% [12]. Similarly, B. Herawan Hayadi et al. (2021) utilized this dataset and demonstrated that Random Forest with a max depth of 17 achieved an accuracy of 99.1% [4]. These prior studies indicate that this dataset is well-suited for machine learning-based prediction tasks due to its comprehensive features and balanced class distribution. Given the proven effectiveness of this dataset in previous research, this study aims to further explore its potential by evaluating additional algorithms, including SVM, and comparing their predictive performance in determining airline passenger satisfaction.

2.2. Preprocessing Data

After the data is collected, the next step is to prepare the data systematically before carrying out analysis using a machine learning model. This process is known as Data Preprocessing, which aims to clean and modify the data to make it more suitable for model training. This stage includes handling missing data, deleting irrelevant data, and data normalization [9]. Data preprocessing aims to prepare the dataset so that it can be processed using data mining algorithms. This process includes handling missing or incomplete data [10].

2.3. Classification

Classification is a grouping process, namely collecting similar objects or entities and separating different objects or entities [15]. Data grouping is the process of organizing data based on the similarity of features resulting from the extraction of certain parameters. Classification also involves grouping data with other data that has similar characteristics [16].

2.4. K-Fold Cross-Validation

K-Fold Cross-Validation is a method used to divide a dataset into training and testing data. This method divides the data into k proportional subsets. This process is carried out repeatedly k times, with each subset taking turns being the test data, while the other subset is used for training [17]. This technique divides the data into 'k' folds to define the data used in training and testing [18].

2.5. K-Nearest Neighbor (K-NN)

K-NN is an algorithm that is included in the supervised learning technique and is known as one of the easiest algorithms to implement in Machine Learning. Although it can be used for both classification and regression, K-NN is generally used more often for classification purposes [19]. The working principle of K-NN is to find the shortest distance between the data to be evaluated and the K nearest neighbors of the training data [20]. K-NN determines the class of an instance based on the majority of the classes of its k nearest neighbors. The selection of k should be considered with certain techniques and should preferably be odd in binary classification to avoid draws. The performance of K-NN also depends on the distance measurement method used [21].

The most common distance calculation used in calculations on the K-NN algorithm is using Euclidean distance calculations. The formula is like equation (1).

$$\text{euc} = \sqrt{(\sum (\text{pi} - \text{qi})\ 2\ )\ \text{n i} = 1} \quad (1)$$

Here, p_i and q_i are the values of each variable in the ith dimension, and n is the number of dimensions. This formula measures geometric distance by adding the squares of the differences between each pair of values and then taking the square root.

2.6. Decision Tree

Decision trees are hierarchical trees built by partitioning data into several sets based on input variables, as a data mining method for item classification. This decision tree has several types, such as ID3, CART, C4.5, and so on [22]. This technique aims to identify optimal decision rules to more accurately predict individual classes based on certain variables [23].

To build a decision tree, the first step is to determine the attribute that will be used as the root[17]. This selection is based on the lowest entropy value and the highest gain value. Entropy is calculated using equation (2).

$$\text{Entropy}(y) = \sum_{i=1}^{n} - p_i \log_2 p_i = 1 \quad (2)$$

Where p_i is the probability of each class in the dataset. After the entropy is calculated, the next step is to calculate the gain value using formula (3).

$$\text{Gain (y,A)} = \text{Entropy}(y, A) - \sum c \in \text{nilai (A)} \frac{yc}{y} \text{ entropy } (y_c) \quad (3)$$

Where y_c is a subset of data based on the value of attribute A. The gain value indicates how much the attribute reduces uncertainty. Next, a branch is created for each attribute value, and the data on each branch is divided based on the attribute value. This process is repeated on each branch until all data has the same class or meets certain criteria.

2.6. Naïve Bayes

Naïve Bayes is a simple probabilistic prediction method, which relies on the application of Bayes' theorem or rules with strong independence assumptions (naïve). The Bayes method itself is a statistical approach used to draw inductive conclusions in classification problems [25]. The steps in the Naive Bayes process are calculating the number of classes or labels, calculating the number of possibilities per class, multiplying all class variables, and comparing the results of the multiplication per class [26]. Bayes' theory is expressed mathematically as equation (4).

$$P(B) = \frac{P(A)P(A)}{P(B)} \tag{4}$$

Where $P(B \mid A)$ is the conditional probability of event B occurring given that condition A has occurred. The value $P(A \mid B)$ is the conditional probability of event A if B has occurred, $P(B)$ is the probability of event B occurring, and $P(A)$ is the probability of event A occurring. This equation is used to update confidence in a hypothesis (event B) based on emerging new evidence (event A).

2.7. Random Forest

RF is an ensemble learning method introduced by Breiman to handle classification and regression problems [27]. RF is a popular machine learning algorithm for classification research of various types of data, which improves accuracy by building multiple decision trees [28]. The Random Forests (RF) formula uses entropy and information gain equations to determine the best division in building a decision tree. Entropy, expressed by equations (5) and (6).

$$\text{Entropy } (Y) = -\sum i\ p\ (c \mid Y)\ \log^2\ p\ (c \mid Y) \tag{5}$$

Measures the degree of irregularity or uncertainty in the set of cases Y. Here, Y is the set of cases, while $p(c \mid Y)$ is the proportion of cases in Y that belong to class c. The lower the entropy value, the greater the homogeneity of the data in a class. Then, information gain is used to evaluate how effectively certain attributes separate data into classes. The information gain equation is expressed as:

$$\text{Information gains (Yes)} = \text{Entropy } (Y) = -\sum_v \varepsilon\text{values } (a)\ \frac{Yv}{Ya}\ \text{Entropy } (Y_v) \tag{6}$$

Where values(a) is the set of values for attribute a, Y_v is a subset of data Y for value v, and Y_a is the entire data relevant to attribute a. This formula measures the difference between the initial entropy and the average entropy after division by a particular attribute, giving an idea of how much uncertainty that attribute has reduced.

2.8. Support Vactor Machine (SVM)

The SVM algorithm is a method used to analyze data and recognize patterns in the classification process [16]. The main goal of SVM is to find the hyperplane with the maximum margin that can separate the classes linearly [29]. SVM is known to have superior performance compared to other classifiers, especially for problems that are not linearly separable. In such cases, SVM with a non-linear kernel such as RBF is the right choice. SVM is also effective in high-dimensional spaces, such as text classification, although its training time is relatively long [30]. The Support Vector Machine (SVM) equation is found in equation (7).

$$f(x) = w^t\ \phi(x) + b \tag{7}$$

Where f(x) is the decision function used to separate data into certain classes. The parameter w represents the weight or vector that indicates the direction and scale of the separating hyperplane. is a collection of variables or features of the data. The feature transformation function $\phi(x)$ is used to map data from the original space to a higher feature space so that non-linear data becomes linearly separable. Meanwhile, b is the bias, which functions to adjust the hyperplane position in feature space. This equation is the essence of SVM in determining the optimal separator that maximizes the margin between data classes.

2.9. Confusion Matrix

Confusion Matrix is a classification performance evaluation method that compares true and false data. This matrix calculates accuracy, precision, recall, and error rate to assess model performance based on the level of correctness and error of the classification results[20]. Confusion Matrix is a table that displays the number of test data that are correctly classified and the number of test data that are incorrectly classified [17]. The Confusion Matrix diagram can be seen in Figure 2.

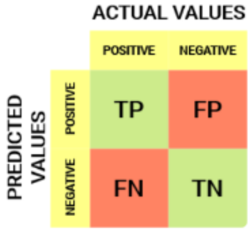


Figure 2. Confusion Matrix Diagram

3. RESULTS AND DISCUSSION

3.1. Initial Data

The first step in this research is collecting data. The dataset used in this research contains 25,975 records about airline passenger satisfaction, obtained from Kaggle (https://www.kaggle.com/datasets/teejmah al20/airline-passenger-satisfaction). This dataset includes 25 attributes. Before entering the classification stage, data analysis is carried out first to understand each variable. This process involves data pre-processing to ensure the data is ready to be processed thoroughly.

Table 1. Airline Passenger Satisfaction Dataset

Unnamed: 0	id	Gender	Customer Type	Age	...	Cleanliness	Departure Delay in Minutes	Arrival Delay in Minutes	Satisfaction
0	19556	Female	Loyal Customer	52	…	5	50	44.0	Satisfied
1	90035	Female	Loyal Customer	36	…	5	0	0.0	Satisfied
2	12360	Male	Disloyal Customer	20	…	2	0	0.0	Neutral or dissatisfied
3	77959	Male	Loyal Customer	44	…	4	0	6.0	Satisfied
4	36875	Female	Loyal Customer	49	…	4	0	20.0	Satisfied
…	…	…	…	…	…	…	…	…	…
25973	37675	Female	Loyal Customer	17	…	2	0	0.0	Neutral or dissatisfied
25974	90086	Male	Loyal Customer	14	…	4	0	0.0	Satisfied
25975	34799	Female	Loyal Customer	42	…	1	0	0.0	Neutral or dissatisfied

3.2. Preprocessing Data

Pre-processing is a stage carried out by eliminating inappropriate data or changing data into a form that is easier to process by the system. The processed data will be cleaned first to remove unwanted data such as unnamed, ID, etc. so that the results will be as shown in Table 2.

Table 2. Preprocessing Dataset

Gender	Customer Type	Age	Type of Travel	Class	...	Checkin service	Inflight service	Cleanliness	Satisfaction
Female	Loyal Customer	52	Business Travel	Eco	…	2	5	5	Satisfied
Female	Loyal Customer	36	Business Travel	Business	…	3	4	5	Satisfied
Male	Disloyal Customer	20	Business Travel	Eco	…	2	2	2	Neutral or dissatisfied
Male	Loyal Customer	44	Business Travel	Business	…	3	1	4	Satisfied
Female	Loyal Customer	49	Business Travel	Eco	…	4	2	4	Satisfied
…	…	…	…	…	…	…	…	…	…

Gender	Customer Type	Age	Type of Travel	Class	...	Checkin service	Inflight service	Cleanliness	Satisfaction
Female	Loyal Customer	17	Loyal Customer	Eco	…	5	4	2	Neutral or dissatisfied
Male	Loyal Customer	14	Loyal Customer	Business	…	4	5	4	Satisfied
Female	Loyal Customer	42	Loyal Customer	Eco	…	1	1	1	Neutral or dissatisfied

In the data pre-processing stage, this research utilized the Python programming language and the Google Colab platform. The initial data preparation process includes steps such as cleaning, filtering, and trimming. Some features that are not used in this research are Unnamed: 0, id, Departure Delay, Arrival Delay, Departure/Arrival time convenient, and Flight Distance. Unnamed Feature: 0 and id are the line numbers in the CSV file and unique identification of each record, so they are not relevant for analysis. Meanwhile, Departure Delay, Arrival Delay, Departure/Arrival time convenience, and Flight Distance are not used because they have a low negative correlation with the passenger satisfaction level classification.

3.3. K-Nearest Neighbor

In this research, the classification process was carried out using the Google Colab platform and the first algorithm model used was K-NN. Model validation was carried out using the k-fold cross-validation method with k=20. The value k=20 was chosen because the dataset used is quite large (27,975 records), so dividing the data into 20 subsets still produces representative training and testing data groups. This method allows each subset to be used alternately as training and testing data. This approach is designed to provide a more comprehensive evaluation of model performance, reduce the risk of overfitting, and ensure more accurate and reliable results. The results of the K-NN algorithm can be seen in Figure 3.

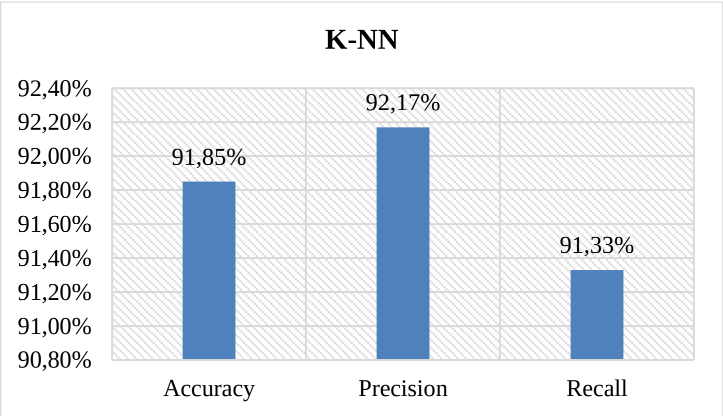


Figure 3. K-NN Classification Results

Based on Figure 3, the analysis results show that the classification model with the K-Nearest Neighbors (K-NN) algorithm produces an accuracy value of 91.85%, precision of 92.17%, and recall of 91.33%. These results reflect that the model has quite good performance in classifying data, with high accuracy and precision. A precision value that is higher than recall indicates that most of the positive predictions made by the model are correct (low false positives), although there are still some positive cases that are not detected (false negatives). This shows that the model tends to be more selective in predicting the positive class, resulting in more precise predictions, but potentially missing some important data. To improve model performance, further optimization is needed, especially to increase the recall value. This step will help the model be more sensitive in detecting all positive cases consistently. Optimization can be done by adjusting the parameter k, choosing a more appropriate distance metric, or performing better data preprocessing, such as normalization or feature selection.

3.4. Decision Tree

The next classification process is carried out using a decision tree algorithm model, the process is also carried out using the same platform, namely, Google Colab, and the same validation, namely 20-Fold. The results of the DT algorithm can be seen in Figure 4.

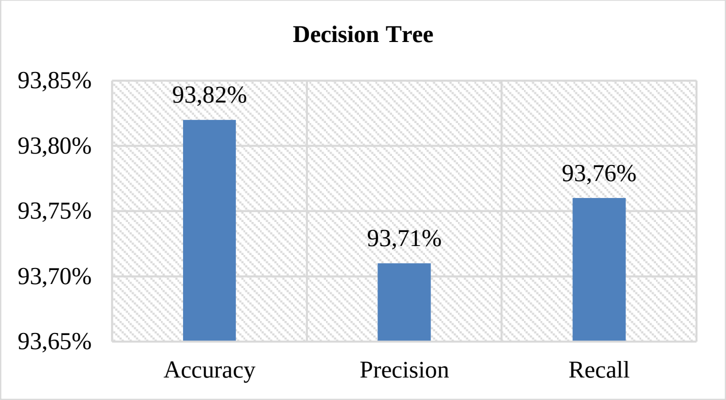


Figure 4. Decision Tree Classification Results

Based on Figure 4, the use of the decision tree algorithm in classifying airline passenger satisfaction produces an accuracy value of 93.82%, which reflects a very good level of accuracy. The precision value obtained was 93.71%, slightly lower than the accuracy, while the recall value reached 93.76%, which shows good performance in detecting positive cases. The classification results using decision trees show that this model has very good performance in classifying data.

High accuracy and recall values indicate that the model is able to recognize patterns well while detecting the majority of positive cases. In addition, the precision value which is close to the recall shows that the positive predictions produced are mostly correct, reflecting the suitability of the model to the data used.

3.5. Naïve Bayes

The next classification process is carried out using the Naïve Bayes algorithm and also uses the same validation. The results of this algorithm can be seen in Figure 5.

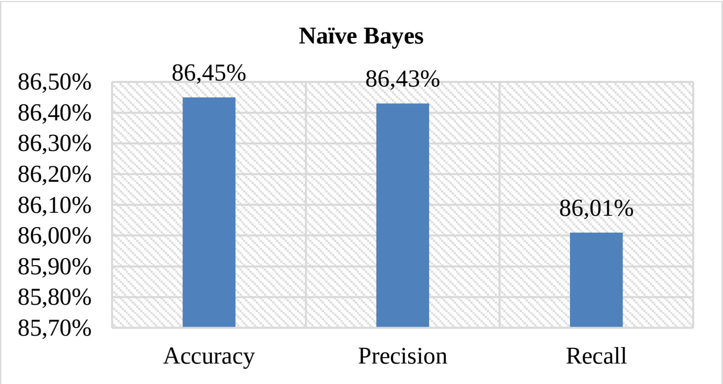


Figure 5. Naïve Bayes Classification Results

Based on Figure 5, the use of the Naïve Bayes algorithm in classifying airline passenger satisfaction produces an accuracy value of 86.45%, precision of 86.43%, and recall of 86.01%. Overall, the Naïve Bayes algorithm is quite good at classifying data, especially because the precision value is close to accuracy indicating that most of the positive predictions are correct. However, more attention needs to be paid to recall, because a low recall value indicates that several positive cases were not detected by the model. This is important to improve, especially in a context where detecting all positive cases is a priority.

3.6. Random Forest

The fourth classification process will use the Random Forest algorithm model. The results can be seen in Figure 6.

Based on Figure 6, the use of the Random Forest algorithm in classifying airline passenger satisfaction produces an accuracy value of 95.78%, which reflects a very good level of accuracy. The resulting precision value was 95.83%, slightly higher than accuracy, while the recall value reached 95.9%, which is the highest value among other metrics. This shows that the model has a very good performance in detecting positive cases. The classification results using the Random Forest algorithm show very good performance in classifying data. A high recall value indicates that the model can detect the majority of positive cases well, reflecting its ability to recognize relevant patterns in the data. In addition, a precision value that is close to the recall value indicates that most of the positive predictions produced by the model are correct, so the error rate in producing false positive predictions is relatively low.

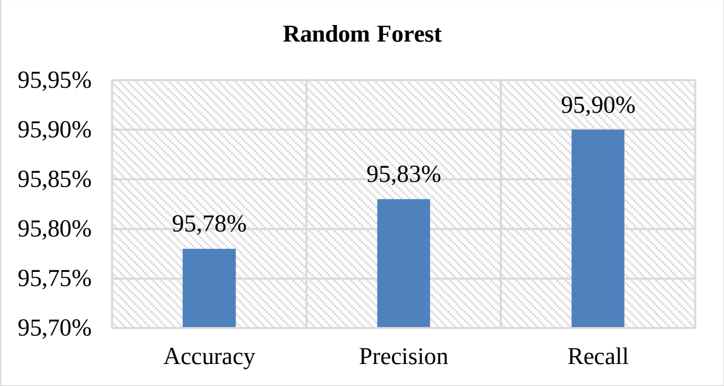


Figure 6. Random Forest Classification Results

However, there is still an opportunity to improve precision and recall simultaneously, especially if certain applications require higher positive prediction accuracy. This improvement can be achieved through adjusting model parameters or more optimal data processing.

3.7. Support Vector Machine (SVM)

The final classification process will use the SVM algorithm model, the process and validation also use the same techniques as the previous algorithm models. The results can be seen in Figure 7.

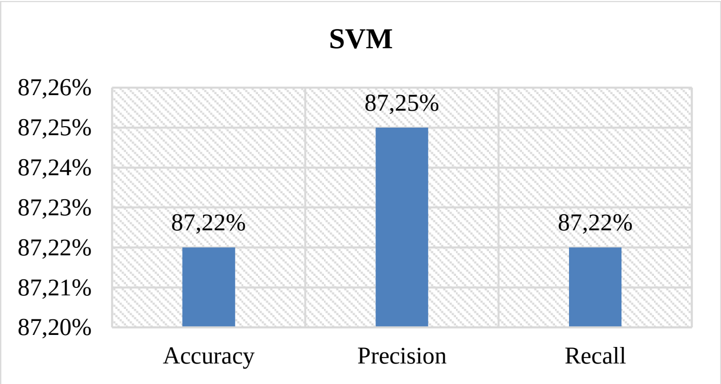


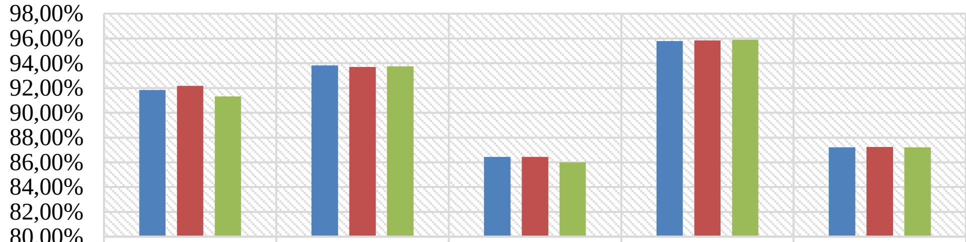
Figure 7. SVM Classification Results

Based on Figure 7, the Support Vector Machine (SVM) algorithm used for airline passenger satisfaction classification produces the same accuracy and recall values of 87.22%, while the precision value is slightly higher, namely 87.25%. These results show that the SVM model has quite good performance, even though it is not optimal. The accuracy and recall values of 87.22% indicate that the model can recognize patterns and detect the majority of positive cases well. However, this performance can still be improved further to produce a more optimal classification.

3.8. Comparison of Results

This analysis section discusses the comparison of classification results using the K-NN, Decision Tree, Naïve Bayes, Random Forest, and SVM algorithms. This research uses the k-fold cross-validation method with a value of k=20, which ensures that the model evaluation is carried out thoroughly. This approach is applied to classify airline passenger satisfaction, resulting in a more accurate and reliable assessment of model performance. From a series of experiments carried out with these five algorithms, researchers have found the best algorithm that produces the highest accuracy for classifying passenger satisfaction data. A comparison of the results of the K-NN, Decision Tree, Naïve Bayes, Random Forest, and SVM algorithms in classifying passenger satisfaction data can be seen in Figure 8.

Comparison of Results



	K-NN	Decision Tree	Naïve Bayes	Random Forest	SVM
Accuracy	91,85%	93,82%	86,45%	95,78%	87,22%
Precision	92,17%	93,71%	86,43%	95,83%	87,25%
Recall	91,33%	93,76%	86,01%	95,90%	87,22%

Accuracy Precision Recall

Figure 1. Algorithm Comparison Results

Experimental results show that the Random Forest algorithm has the best performance in predicting airline passenger satisfaction, with the highest accuracy value of 95.78%, precision of 95.83%, and recall of 95.90%. This algorithm is very effective in classifying data, both in detecting correct satisfies and minimizing prediction errors. The Decision Tree algorithm ranked second with 93.82% accuracy, 93.71% precision, and 93.76% recall, showing accurate and consistent performance. Followed by K-NN which has 91.85% accuracy, 92.17% precision, and 91.33% recall, although it requires additional optimization. The Naïve Bayes algorithm shows the lowest performance with 86.45% accuracy, 86.43% precision, and 86.01% recall, while SVM has 87.22% accuracy, 87.25% precision, and 87.22% recall, but still inferior to Random Forest, Decision Tree, and K-NN.

The superiority of Random Forest in this study is closely related to the nature of the dataset used. With the large number of features, the possibility of correlation between features, and the combination of numerical and categorical data, Random Forest is the best choice because it is able to handle complex data, reduce overfitting, and provide more stable prediction results than other algorithms. Random Forest is the best algorithm for passenger satisfaction classification in this case, outperforming the other four algorithms in all evaluation metrics.

3.9. Discussion

The results of the algorithm performance comparison in this study show that the use of 19 features correlated with passenger satisfaction produces the best performance in the Random Forest algorithm, achieving the highest accuracy of 95.78% based on k-fold validation (k=20). These findings confirm that Random Forest is highly effective in classifying airline passenger satisfaction.

To validate these findings, this study refers to two key previous studies: A.C.Y. Hong et al. (2023) and B. Herawan Hayadi et al. (2021). The selection of these studies as main references is based on their relevance in methodology, dataset usage, and focus on algorithm performance evaluation. The study by Hong et al. (2023) analyzed the prediction of airline passenger satisfaction using machine learning algorithms, including Random Forest, K-Nearest Neighbors (K-NN), Decision Tree, Naïve Bayes, Logistic Regression, and AdaBoost [12]. Their results demonstrated that Random Forest achieved the highest accuracy of 89.20%, outperforming the other algorithms. However, their study was limited to using only five selected features, which might restrict the generalizability of their findings. Despite this limitation, Hong et al. (2023) remains an important reference as it confirms the superiority of Random Forest over other classification algorithms in a similar context. Meanwhile, the study by B. Herawan Hayadi et al. (2021) serves as a more comprehensive benchmark, as it compared the performance of six algorithms Decision Tree, Logistic Regression, K-NN, Random Forest, Gaussian Naïve Bayes, and an Ensemble method—on an airline passenger satisfaction dataset[4]. Their research reported that Random Forest achieved the highest accuracy of 99.1% with an AUC (Area Under Curve) of 0.993, surpassing all other algorithms tested. Hayadi et al.'s study is particularly relevant because it used a more extensive feature set, making it methodologically closer to this research.

By integrating insights from Hong et al. (2023) and Hayadi et al. (2021), this study strengthens the argument that Random Forest consistently performs better in passenger satisfaction classification. The findings from Hayadi et al. particularly validate the impact of using a larger feature set, which aligns with the approach taken in this study. Additionally, both references confirm that Random Forest effectively handles complex patterns in the dataset, reinforcing its suitability for predictive modeling in the airline industry.

Nevertheless, it is important to acknowledge that model performance depends on various factors, including dataset quality, preprocessing techniques, and hyperparameter optimization. Despite these variations, the consistency of Random Forest's superior performance across multiple studies further supports its effectiveness for airline passenger satisfaction classification.

4. CONCLUSION

This research aims to evaluate the effectiveness of various classification algorithms, namely Decision Tree, Naïve Bayes Classifier, K-NN, Random Forest, and SVM. These algorithms are applied to predict airline passenger satisfaction obtained by Kaggle. The k-fold cross-validation technique with a value of k=20 was applied to ensure the model evaluation was carried out thoroughly by dividing the data into 20 subsets for training and testing. This approach helps reduce the risk of overfitting while providing a more accurate and reliable assessment of model performance. The research results show that the Random Forest algorithm provides the best performance with the highest accuracy, namely 95.78%, precision 95.83%, and recall 95.90%. The Decision Tree algorithm is in second place with an accuracy of 93.82%, followed by K-NN with an accuracy of 91.85%. The SVM algorithm achieved an accuracy of 87.22%, while Naïve Bayes had the lowest performance with an accuracy of 86.45%.

The findings of this study confirm that Random Forest is the most reliable algorithm for airline passenger satisfaction classification, as it effectively handles complex relationships between features, reduces overfitting, and delivers consistently high accuracy. These results are consistent with previous studies, particularly those conducted by Hong et al. (2023) and Hayadi et al. (2021), which also demonstrated the superiority of Random Forest in similar classification tasks. The ability of this algorithm to process large datasets while maintaining high predictive accuracy makes it an ideal choice for airline companies seeking data-driven insights into customer satisfaction. The practical implications of this research extend beyond algorithm comparison. Machine learning models, such as the one developed in this study, can assist airlines in identifying key factors that influence passenger satisfaction and optimizing their services accordingly. By leveraging predictive analytics, airlines can proactively enhance customer experience, allocate resources more efficiently, and refine service strategies to meet passenger expectations. The ability to analyze passenger satisfaction patterns also allows airlines to continuously improve service quality based on data-driven evaluations.

Future research could explore hybrid models or deep learning approaches to further improve predictive accuracy. Additionally, applying this model to real-time passenger feedback, such as social media sentiment analysis or in-flight service ratings, could provide more dynamic insights that contribute to continuous improvements in airline service quality.

REFERENCES

[1] N. Vojtek and B. Smudja, "Improving The Passenger Feedback Process In Airline Industry," vol. 9, no. 2, pp. 255–269, 2019, [Online]. Available: doi: http://dx.doi.org/10.7708/ijtte.2019.9(2).10

[2] S. Soklaridis, A. M. Geske, and S. Kummer, "Journal of the Air Transport Research Society Key characteristics of perceived customer centricity in the passenger airline industry : A systematic literature review," J. Air Transp. Res. Soc., vol. 3, no. July, p. 100031, 2024, doi: 10.1016/j.jatrs.2024.100031.

[3] M. B. Gorzałczany and F. Rudzi, "applied sciences Business Intelligence in Airline Passenger Satisfaction Study — A Fuzzy-Genetic Approach with Optimized," 2021, [Online]. Available: https://doi.org/10.3390/app11115098

[4] B. H. Hayadi, J. Kim, K. Hulliyah, and H. T. Sukmana, "Predicting Airline Passenger Satisfaction with Classification Algorithms," vol. 4, no. 1, pp. 82–94, 2021, [Online]. Available: https://doi.org/10.47738/ijiis.v4i1.80

[5] M. H. Setiono, "Komprasi Algoritma Decision Tree, Random Forest, SVM dan K-NN Dalam Klasifikasi Kepuasan Penumpang Maskapai Penerbangan," vol. 17, no. 1, pp. 32–39, 2022.

[6] P. Sun et al., "IBRO Neuroscience Reports Bipolar disorder : Construction and analysis of a joint diagnostic model using random forest and feedforward neural networks," vol. 17, no. December 2023, pp. 145–153, 2024.

[7] S. Kang, "k -Nearest Neighbor Learning with Graph Neural Networks," 2021, [Online]. Available: https://doi.org/10.3390/math9080830

[8] S. Lee, C. Lee, K. G. I. Mun, and D. Kim, "Decision Tree Algorithm Considering Distances Between Classes," vol. 10, no. July, 2022, [Online]. Available: 10.1109/ACCESS.2022.3187172

[9] D. S. Suprapto and R. S. Oetama, "Jurnal Informatika Ekonomi Bisnis Analysis of Airline Passenger Satisfaction Using Decision Tree and Naïve Bayes Algorithms," vol. 5, pp. 1493–1500, 2023, doi: 10.37034/infeb.v5i4.728.

[10] I. T. Akinola, Y. Sun, I. G. Adebayo, and Z. Wang, "Daily peak demand forecasting using Pelican

Algorithm optimised Support Vector Machine (POA-SVM),” Energy Reports, vol. 12, no. October, pp. 4438–4448, 2024, doi: 10.1016/j.egyr.2024.10.017.
[11] A. S. Nugroho, A. B. Witarto, and D. Handoko, “Support Vector Machine,” 2003, [Online]. Available: https://asnugroho.net/papers/ikcsvm.pdf
[12] A. C. Y. Hong, K. W. Khaw, X. Y. Chew, and W. C. Yeong, “Prediction of US airline passenger satisfaction using machine learning algorithms,” vol. 4, no. 1, pp. 7–22, 2023, [Online]. Available: https://doi.org/10.15282/daam.v4i1.9071
[13] A. Handayanto, K. Latifa, N. D. Saputro, and R. R. Waliyansyah, “Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi (Analysis and Application of Algorithm Support Vector Machine (SVM) in Data Mining to Support Promotional Strategies),” vol. 7, no. November, pp. 71–79, 2019.
[14] P. F. Pratama, D. Rahmadani, and R. S. Nahampun, “Random Forest Optimization Using Particle Swarm Optimization for Diabetes Classification,” vol. 1, no. July, pp. 41–46, 2023.
[15] B. P. Christie, D. M. Tat, and T. Irwin, “Comparation of K-Nearest Neighboor (K-NN) and Naive Bayes Algorithm for the Classification of the Poor in Recipients of Social Assistance Comparation of K-Nearest Neighboor (K-NN) and Naive Bayes Algorithm for the Classification of the Poor in Recipients of Social Assistance”, doi: 10.1088/1742-6596/1641/1/012077.
[16] A. Rahmah, N. Sepriyanti, and M. H. Zikri, “Implementation of Support Vector Machine and Random Forest for Heart Failure Disease Classification,” vol. 1, no. July, pp. 34–40, 2023.
[17] Z. C. Dwynne, “Comparison Of Machine Learning Algorithms On Sentiment Analysis Of Elsagate Content,” 2024 Int. Conf. Smart Comput. IoT Mach. Learn., pp. 239–243, 2024, doi: 10.1109/SIML61815.2024.10578186.
[18] S. M. Ganie and M. B. Malik, “Healthcare Analytics An ensemble Machine Learning approach for predicting Type-II diabetes mellitus based on lifestyle indicators,” Healthc. Anal., vol. 2, no. August, p. 100092, 2022, doi: 10.1016/j.health.2022.100092.
[19] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor , Genetic , Support Vector Machine , Decision Tree , and Long Short Term Memory algorithms in machine learning,” Decis. Anal. J., vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.
[20] Z. C. Dwinnie et al., “Application of the Supervised Learning Algorithm for Classification of Pregnancy Risk Levels,” vol. 1, no. July, pp. 26–33, 2023.
[21] W. Cherif, “ScienceDirect ScienceDirect Optimization of K-NN algorithm by clustering and reliability coefficients : application to breast-cancer diagnosis,” Procedia Comput. Sci., vol. 127, pp. 293–299, 2018, doi: 10.1016/j.procs.2018.01.125.
[22] A. I. Putri, N. A. Husna, N. M. Cia, and M. A. Arba, “Implementation of K-Nearest Neighbors , Naïve Bayes Classifier , Support Vector Machine and Decision Tree Algorithms for Obesity Risk Prediction,” vol. 2, no. July, pp. 26–33, 2024.
[23] G. Pappalardo, S. Cafiso, A. Di Graziano, and A. Severino, “Decision Tree Method to Analyze the Performance of Lane Support Systems,” 2021, [Online]. Available: https://doi.org/10.3390/su13020846
[24] A. F. Lubis, H. Z. Haq, I. Lestari, and M. Iltizam, “Classification of Diabetes Mellitus Sufferers Eating Patterns Using K-Nearest Neighbors , Naïve Bayes and Decission Tree,” vol. 2, no. July, pp. 44–51, 2024.
[25] D. R. Prehanto, A. D. Indriyanti, and K. D. Nuryana, “Preprocessing Using Correlation Based Features Selection on Naive Bayes Classification Preprocessing Using Correlation Based Features Selection on Naive Bayes Classification”, doi: 10.1088/1757-899X/982/1/012012.
[26] M. R. Romadhon, “A Comparison of Naive Bayes Methods , Logistic Regression and K-NN for Predicting Healing of Covid-19 Patients in Indonesia,” pp. 41–44, 2021.
[27] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Member, “Support Vector Machine Versus Random Forest for Remote Sensing Image Classification : A Meta-Analysis and Systematic Review,” vol. 13, pp. 6308–6325, 2020.
[28] C. Avci, M. Budak, N. Yagmur, and F. B. Balcik, “Comparison between random forest and support vector machine algorithms for LULC classification,” vol. 8, no. 1, pp. 1–10, 2023, doi: 10.26833/ijeg.987605.
[29] D. M. Abdullah and A. M. Abdulazeez, “Machine Learning Applications based on SVM Classification : A Review,” pp. 81–90, doi: 10.48161/Issn.2709-8206.
[30] E. Y. Boateng, J. Otoo, and D. A. Abaye, “Basic Tenets of Classification Algorithms K -Nearest-Neighbor , Support Vector Machine , Random Forest and Neural Network : A Review,” pp. 341–357, 2020, doi: 10.4236/jdaip.2020.84020.