Jurnal Informatika Ekonomi Bisnis

http://www.infeb.org

2023 Vol. 5 Iss. 4 Hal: 1493-1500 e-ISSN: 2714-8491

Analysis of Airline Passenger Satisfaction Using Decision Tree and Naïve Bayes Algorithms

Devita Sertivia Suprapto[1], Raymond Sunardi Oetama[2]✉

[1,2]Information Systems Department, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara

raymond@umn.ac.id

Abstract

In the dynamic landscape of airline services, comprehending the intricacies that mold customer satisfaction is paramount to elevating overall service quality. This study aspires to dissect these pivotal elements, contributing nuanced insights that can propel the enhancement of customer satisfaction within the industry. A multifaceted investigation encompasses analyzing demographic data, exploring underlying factors significantly shaping passenger satisfaction, and identifying the most adept model for forecasting imminent passenger satisfaction outcomes. A model was meticulously crafted by leveraging a decision tree algorithm to discern the substantial variables influencing passenger satisfaction. Simultaneously, the Naïve Bayes algorithm was harnessed to prognosticate forthcoming passenger satisfaction. The findings underscore the diverse facets of the flying experience impacting satisfaction, with both ctree and rpart decision tree algorithms spotlighting critical factors, such as online boarding, inflight entertainment, WiFi service, class, and travel type. The Naïve Bayes algorithm demonstrates around 87% accuracy in predicting passenger satisfaction, underscoring its efficacy in discerning patterns within this complex realm.

Keywords: Airline Passenger Satisfaction, Decision Tree, Naïve Bayes, In-Flight Services, Customer Satisfaction.

1. Introduction

In a rapidly evolving world, human mobility is on the rise. People are moving from one region to another within a single country and covering tens of kilometers to reach other countries in the shortest time possible. Therefore, air travel has emerged as the fastest mode of human transportation. This mode of travel has significantly expedited movement, with aircraft technology continuously advancing to ensure enhanced safety. Given the sense of security and efficiency, air transport has garnered increasing popularity among the public. Number of Flight in 2019 and 2020 on Figure 1.

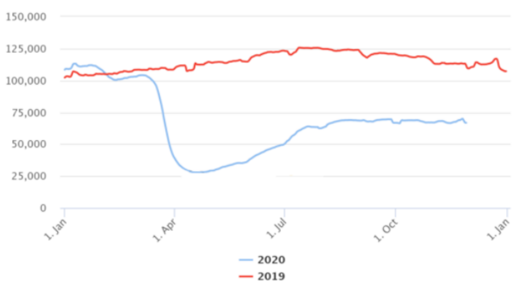

Figure 1. Number of Flight in 2019 and 2020

Consequently, airlines strive to provide the best facilities and services to secure passenger loyalty. When passengers are content with the services they receive, the likelihood of them choosing the same airline for their future flights increases significantly. This contentment can also influence their friends and acquaintances through word-of-mouth recommendations. Such dynamics have a substantial impact on both the productivity and reputation of the airline industry. It becomes evident that passenger satisfaction is a paramount aspect of management within the transportation sector, particularly in the aviation industry.

Figure 1 illustrates the number of flights in the years 2019 and 2020. Despite experiencing a decline during the Covid-19 pandemic, the aviation industry is showing signs of recovery. The aviation industry is undeniably substantial, as evidenced by nearly every country's airlines, hundreds of thousands of flights daily, and airports rarely devoid of visitors. This industry significantly contributes to a country's economic growth. Consequently, various institutions routinely release annual rankings of the world's best airlines, with customer satisfaction being one of the critical determinants of an airline's quality. Skytrax World Airline Award 2019 on Table 1.

Table 1. Skytrax World Airline Award 2019

Rank	Airlines
1	Qatar Airways
2	Singapore Airline
3	ANA All Nippon Airways
4	Cathay Pacific Airways
5	Emirates
6	Eva Air
7	Hainan Airline
8	Qantas Airways
9	Lufthansa
10	Thai Airways
11	Japan Airlines
12	Garuda Indonesia

Accepted: 11-11-2023 | Revision: 30-12-2023 | Publication: 31-12-2023 | doi: 10.37034/infeb.v5i4.728

One of the most prestigious assessment organizations in the aviation industry is Skytrax. Skytrax is an accredited organization researching air transportation (Sezgin & Yuncu, 2016). International awards from Skytrax are highly coveted by airlines, as they aim to encourage air transport service providers to improve the quality and performance of their services consistently. As depicted in Table 1, in 2019, Garuda Indonesia, a source of national pride, slipped out of the top 10 airlines in the world, according to Skytrax, and found itself in 12th place. This event prompted the researchers to identify the factors influencing passenger satisfaction with an airline. These factors can be valuable to Garuda Indonesia and other airlines, helping them maximize their performance and secure top positions among the best airlines in Asia and worldwide. While Skytrax assesses various aspects of airline performance, including customer satisfaction, cabin service, crew performance, profitability, fleet age, and more, the researchers are specifically interested in examining one of the criteria used by Skytrax: customer satisfaction. One of the methods to measure customer satisfaction is by conducting surveys, as customer feedback serves as a gauge for determining the success of a business.

Several previous studies have examined customer satisfaction in the airline service sector. They study customer satisfaction, customer complaints, and customer sentiment. When they study customer satisfaction with airline services, they learn about customer feedback, customer loyalty, and factors that influence customer satisfaction. In terms of algorithms, some employed classification techniques such as Random Forest, KNearest Neighbour, and Natural language processing. Many applied the Decision Tree algorithm and Naïve Bayes, among the most commonly utilized methods. Additionally, certain studies compare Decision Tree and Naïve Bayes.

In this study, we've acquired comprehensive and transparent customer satisfaction survey data from US Airlines that closely aligns with our research objectives. Additionally, the dataset is sufficiently extensive to facilitate the application of Decision Trees and Naïve Bayes algorithms. Our research addresses several crucial questions. First, we aim to investigate whether demographic data can effectively discern passenger satisfaction, focusing on identifying trends and correlations among specific passenger groups and their satisfaction levels. Second, we delve into the myriad factors influencing passenger satisfaction to uncover the underlying determinants of satisfaction levels. Lastly, we explore various modeling and forecasting methodologies to determine which model offers the most efficient means for predicting future passenger satisfaction outcomes.

2. Research Method

The subject of this study encompasses passenger data, flight types, delay information, and passenger evaluations of various airline services. Out of the 24 variables obtained from the dataset, only 19 relevant variables are utilized in this research. The data for this study was acquired from passengers of US Airlines through the Kaggle.com website. The choice to use data from US Airlines is justified by the fact that airlines, in principle, serve a diverse clientele, including domestic and international passengers. The United States is a global economic hub, attracting people from various parts of the world who have likely used American Airlines. By utilizing this data, the research aims for greater accuracy, and the conclusions drawn can apply to airlines worldwide, not limited to Indonesian carriers alone. Research Flow on Figure 2.

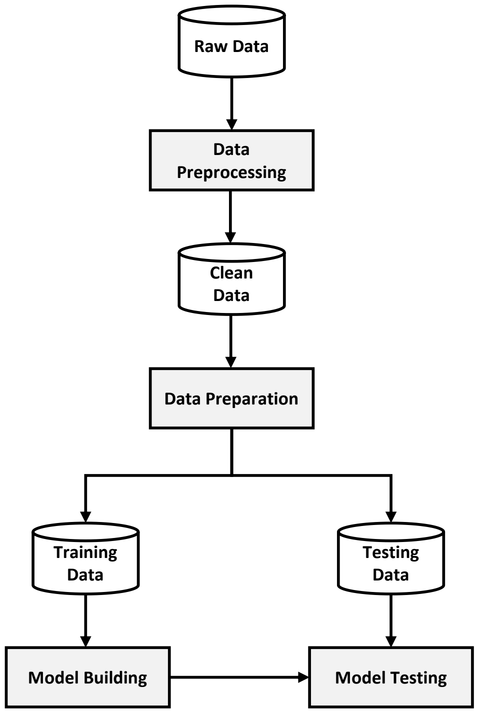

Figure 2. Research Flow

The process begins with Raw Data, which undergoes Data Preprocessing to ensure cleanliness and validity. After this step, the clean data is further divided into Data Preparation, which includes Data Training and Data Testing. The data is split into 70% for training and 30% for testing. Data Training is utilized for constructing the model, while Testing Data is employed to evaluate the model's performance. This approach ensures the model is developed with high-quality, reliable data and rigorously tested to assess its accuracy and effectiveness. All processes are performed by using R programming.

The Decision Tree algorithm is a method used to learn and predict patterns within data and represent the relationships between attribute variables in the form of a tree structure. A decision tree consists of the root node (the top node), decision nodes (branches generated by the parent node), and leaf nodes (the final nodes in the tree, which represent conclusions). The advantages of the decision tree algorithm include its easy-to-understand visualization, the ability to determine the best and worst values for various scenarios, simplification of complex and global decision areas, no need to calculate linear relationships between parameters, and the capability to process both numerical and categorical data.

On the other hand, the Naïve Bayes algorithm is a method that seeks the highest probability to classify data into the most appropriate category. It is done based on Bayes' theorem, which assumes that all variables are independent and have no correlation that can affect the results, hence the term naïve. The first step involves calculating the total probability of each event class, which is done by dividing the number of data cases for the event class by the total number of data in the table. Second, it calculates the detailed probabilities of variables within each category. It is achieved by dividing the number of variable cases matching the event class by the number of data cases for the event class. It is done for all the variables under investigation. Third, it multiplies all the variable results for each event class, repeating this process for all event class conditions. Finally, it compares the final results for each category, and the decision is made based on the highest impact. The advantages of the Naïve Bayes algorithm include its ability to work with quantitative and qualitative data, not requiring extensive data, and being accessible to create, understand, and effectively. Naïve Baye can also predict the probability for each class member.

3. Result and Discussion

Can passenger satisfaction be discerned through the analysis of demographic data?. Upon thorough scrutiny of the gender data, a notable equilibrium is apparent, indicating a balanced representation of both male and female passengers. Yet, a nuanced observation reveals a subtle disparity, with a slightly higher number of female passengers than their male counterparts. This differentiation is conspicuously manifest in the graphical illustration delineated in Figure 3.

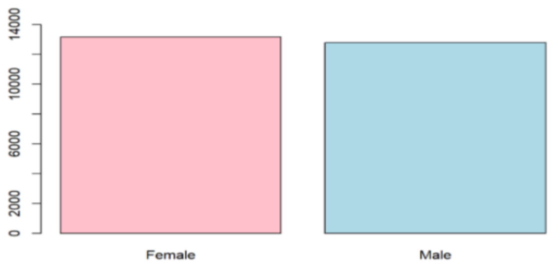

Figure 3. Demographic Data – Customer Gender

On the customer type, it is noteworthy that Loyal Customers significantly outnumber Disloyal Customers within the context of US Airlines, as illustrated in Figure 4. This observation highlights a prevailing trend where a substantial portion of passengers demonstrate loyalty towards the airline, implying a positive and enduring relationship between these passengers and the airline. Demographic Data – Customer Type on Figure 4.

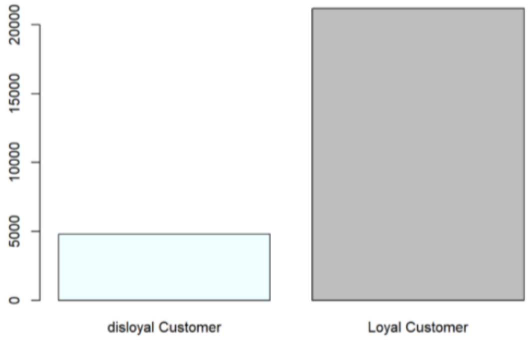

Figure 4. Demographic Data – Customer Type

In Figure 5, the depiction of various types of travel provides a noteworthy insight. It's evident that passengers traveling for business purposes significantly outnumber those who are traveling for personal reasons. This observation underlines a prevalent trend in the dataset where many passengers are categorized as business travelers. This dominance of business travelers has several implications. First, it suggests a substantial demand for air travel related to business activities, such as corporate meetings, conferences, or work-related trips. It also points to airlines' importance in catering to business travelers' specific needs and preferences, including schedule flexibility, onboard amenities, and booking options.

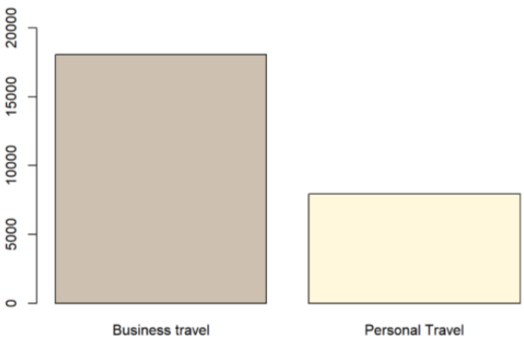

Figure 5. Demographic Data – Types of Travel

In Figure 6, we are presented with an illustration of passenger class preferences, encompassing Business, Economy, and Economy Plus. Business Class takes the lead, with a remarkable count exceeding 12,000 passengers. Subsequently, Economy Class occupies the second position regarding passenger favorability, while Economy Plus Class follows in the third position. This distribution of passenger class preferences offers valuable insights into passenger choices and priorities. The substantial preference for Business Class indicates a significant segment of travelers who highly value the premium services, enhanced comfort, and amenities associated with this class. Simultaneously, the popularity of Economy Class underscores the significance of cost-effective travel options and prudent budgetary considerations among passengers. Comprehending the hierarchy of class preferences is paramount for airlines as it can influence decisions regarding seating configurations, pricing strategies, and service enhancements. Additionally, this knowledge can aid in customizing marketing strategies to target specific passenger segments, ultimately contributing to enhanced overall customer satisfaction and the airline's financial performance.

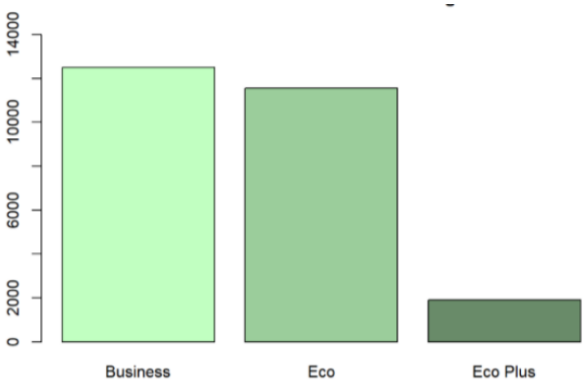

Figure 6. Demographic Data – Class

A decision tree algorithm, specifically the ctree model, is integral to this analysis. This model has generated a decision tree, represented in Figure 7, which provides valuable insights into the factors influencing customer satisfaction. The decision tree highlights the substantial impact of passenger demographic data on customer satisfaction, with particular emphasis on two key variables: flight class and travel type. It becomes evident that the most satisfied passengers are those with a business travel type who choose the business class. This observation suggests that individuals traveling for business purposes place a premium on service quality and comfort. Business travelers often have unique needs and expectations, given the demanding nature of their trips. The decision tree's conclusion aligns with this understanding, as the business class typically offers top-tier service, catering to the requirements of individuals who prioritize enhanced comfort, efficiency, and a premium travel experience. On the other hand, the economy or economy plus class appears to be more suitable for personal travel. Individuals traveling for personal reasons often prioritize affordability, especially on shorter flight routes where the economy class can meet their needs. This analysis underscores the importance of tailoring services to specific passenger segments. Airlines can focus on enhancing service quality and comfort in business class, recognizing that it is often the choice of travelers with higher service expectations. Conversely, affordability and cost-effectiveness are critical considerations for passengers traveling for personal reasons, and airlines can concentrate on delivering value in economy class. In summary, the decision tree model's insights emphasize the significance of aligning service offerings with passenger demographics' distinct preferences and needs, ultimately contributing to higher customer

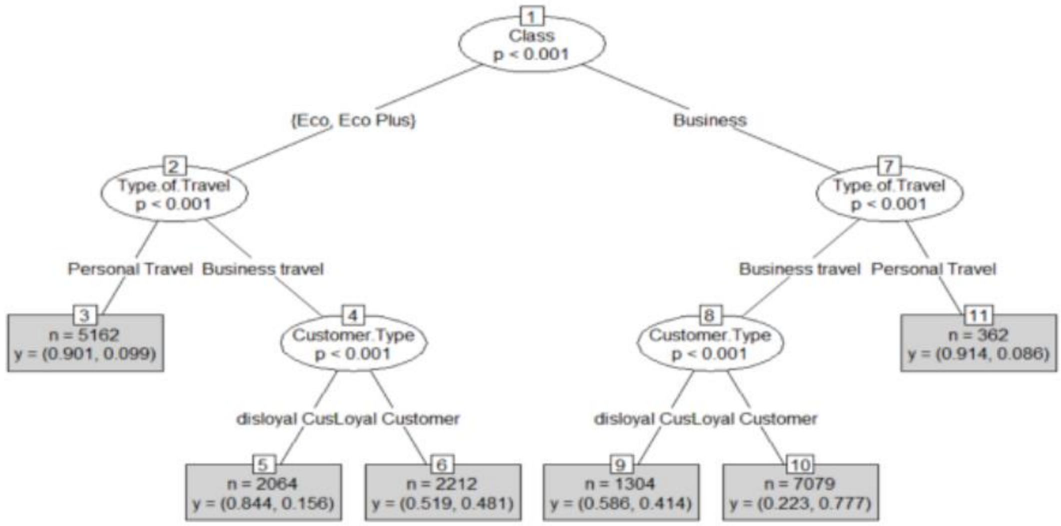

Figure 7. The Decision Tree for Passenger Satisfaction Based on Passenger Demographic Data.

satisfaction and loyalty in the competitive airline industry.

```
Confusion Matrix and Statistics

predict_party            neutral or dissatisfied satisfied
  neutral or dissatisfied                     3718      1048
  satisfied                                    646      2381

               Accuracy : 0.7826
                 95% CI : (0.7733, 0.7917)
    No Information Rate : 0.56
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5533

 Mcnemar's Test P-Value : < 2.2e-16

            Sensitivity : 0.8520
            Specificity : 0.6944
         Pos Pred Value : 0.7801
         Neg Pred Value : 0.7866
             Prevalence : 0.5600
         Detection Rate : 0.4771
   Detection Prevalence : 0.6116
      Balanced Accuracy : 0.7732

       'Positive' Class : neutral or dissatisfied
```

Figure 8. The Decision Tree Confusion Matrix and Statistics

The model mentioned above was then tested for accuracy, and as indicated by the confusion matrix in Figure 8, it achieved a reasonable accuracy rate of 78.26%. For all algorithms, the data was divided into 70% training data and 30% testing data.

What are the underlying factors that exert influence on passenger satisfaction levels?. Furthermore, we aim to determine the factors that affect passenger satisfaction following their journey. During their travels, passengers are exposed to various services provided by the airline, such as in-flight WiFi, entertainment (movies, magazines, music), cleanliness, legroom, punctuality, food, drinks, and many more. Online booking, food and beverages,

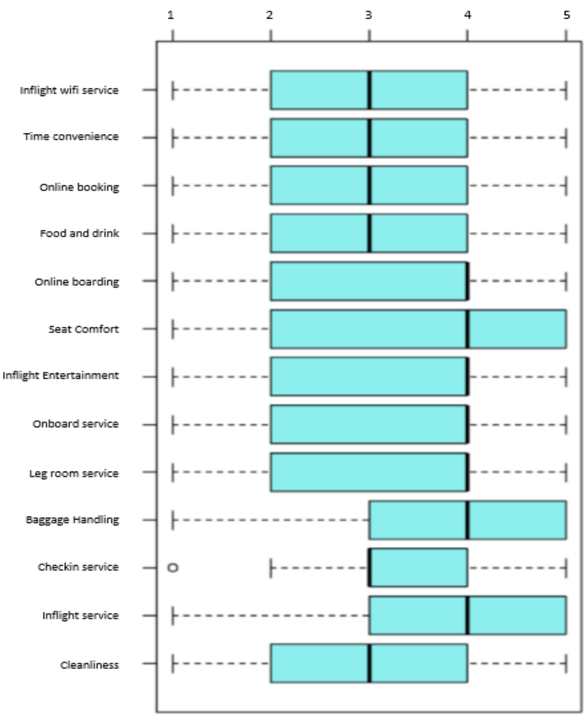

Figure 9. Boxplot of US Airline Services

Figure 9 displays box plots for each of the services provided by US Airlines. As indicated by their median values, several services with relatively high satisfaction levels include online boarding, seat comfort, inflight entertainment, onboard service, legroom services, baggage handling, and inflight service. Online Boarding: High satisfaction levels regarding the online check-in process suggest that passengers value the convenience and efficiency of pre-departure check-in procedures.

Seat Comfort: Passenger satisfaction with seat comfort is crucial to the overall travel experience. High satisfaction indicates that passengers find the seating comfortable during their flights. Inflight Entertainment: The quality of inflight entertainment, including movies, music, or games, can significantly enhance the passenger experience. High satisfaction levels here indicate the presence of adequate entertainment options. Onboard Service: The service provided by cabin crew, including cabin cleanliness and the responsiveness of the flight attendants to passenger needs, plays a vital role in passenger satisfaction. High satisfaction in this category reflects positively on the airline's service quality.

Legroom Services: Adequate legroom is essential for passenger comfort during the flight. High satisfaction levels suggest that passengers find the legroom accommodating and satisfactory. Baggage Handling: The efficiency and accuracy of baggage handling, from check-in to retrieval at the destination, are critical aspects of the passenger journey. High satisfaction levels indicate the airline's competence in handling passengers' luggage. Inflight Service: The quality of inflight service, including meals and beverages, also contributes to passenger satisfaction. High satisfaction levels in this category reflect positively on the in-flight dining and overall service quality.

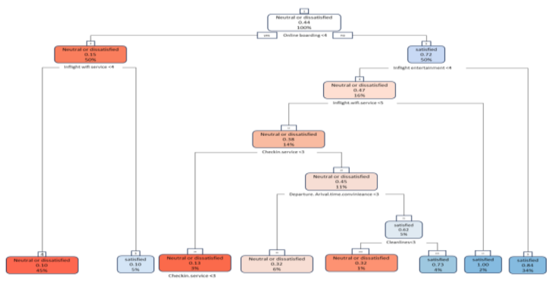

Figure 9. The Decision Tree for Customer Satisfaction Based on Satisfaction with Each Service.

Several services still require further enhancement as their medians fall below satisfactory levels. These include inflight WiFi service, time convenience, check-in service, and cleanliness. First, The availability and quality of inflight WiFi service have become increasingly important to passengers, particularly in this era of growing digital connectivity. If the median value remains below expectations, passengers may be dissatisfied with the internet options, connection speed, or service stability. Secondly, time convenience involves the airline's adherence to flight schedules. If the median in this aspect is unsatisfactory, it may suggest passenger dissatisfaction with delays or schedule changes that can disrupt travel plans. Third, The check-in process marks the initial stage of air travel and can significantly influence the overall passenger experience. If median satisfaction with the check-in service is low, it may reflect issues such as long queues, difficulties with self-service check-in options, or unresponsive staff. Finally, the cleanliness of aircraft and airport facilities is a critical factor in passengers' perception of the airline. If median satisfaction with cleanliness is still lacking, it may indicate a lack of attention to cleanliness, poor cabin hygiene, or inadequately maintained airport facilities.

Furthermore, we employed the decision tree algorithm with the rpart model, as shown in Figure 9, to analyze these factors. From the decision tree, we can observe that online booking or check-in services, in-flight entertainment, and WiFi availability during the flight play pivotal roles in overall customer satisfaction. Additionally, offline check-in services, punctuality, and cleanliness significantly improve passenger satisfaction. The higher the score, typically three or more on a 1-5 point scale, the greater the likelihood of passengers being satisfied with the airline. Rpart Model Performance on Figure 10.

```
predict_rpart            neutral or dissatisfied satisfied
  neutral or dissatisfied                    3680       605
  satisfied                                   684      2824

               Accuracy : 0.8346
                 95% CI : (0.8262, 0.8428)
    No Information Rate : 0.56
    P-Value [Acc > NIR] : < 2e-16

                  Kappa : 0.6652

 Mcnemar's Test P-Value : 0.02981

            Sensitivity : 0.8433
            Specificity : 0.8236
         Pos Pred Value : 0.8588
         Neg Pred Value : 0.8050
             Prevalence : 0.5600
         Detection Rate : 0.4722
   Detection Prevalence : 0.5499
      Balanced Accuracy : 0.8334

       'Positive' Class : neutral or dissatisfied
```

Figure 10. Rpart Model Performance

The testing of the Rpart model for accuracy by predicting the testing data and creating a confusion matrix with an accuracy rate of 83.46%, as displayed in Figure 10, holds significant implications. This high accuracy rate suggests that customer satisfaction with individual services after a flight can be a robust indicator of overall passenger satisfaction. This finding implies that airlines can gain valuable insights into overall passenger satisfaction by evaluating aspects of the passenger experience, such as in-flight services, check-in processes, baggage handling, or any other post-flight services. Which model offers the most effective means for forecasting forthcoming passenger satisfaction outcomes?. Naïve Bayes Model on Figure 11.

```
[1] 700
[1] 300
$apriori
grouping
neutral or dissatisfied               satisfied
              0.5457143               0.4542857

$tables
$tables$Class
                         var
grouping                    Business        Eco   Eco Plus
  neutral or dissatisfied 0.27748691 0.59162304 0.13089005
  satisfied               0.72641509 0.21698113 0.05660377

$tables$Type.of.Travel
                         var
grouping                  Business travel Personal Travel
  neutral or dissatisfied       0.5418848       0.4581152
  satisfied                     0.9245283       0.0754717

$tables$Online.boarding
                            [,1]     [,2]
neutral or dissatisfied 2.732984 1.099707
satisfied               4.053459 1.156645

$tables$Inflight.entertainment
                            [,1]    [,2]
neutral or dissatisfied 2.950262 1.32417
satisfied               3.965409 1.02434

$tables$Inflight.wifi.service
                            [,1]      [,2]
neutral or dissatisfied 2.384817 0.9424823
satisfied               3.267296 1.5425407


$levels
[1] "neutral or dissatisfied" "satisfied"

$call
NaiveBayes.default(x = X, grouping = Y)
```

Figure 11. Naïve Bayes Model

Even though both decision trees above have identified the factors affecting passenger satisfaction, the researcher aims to construct an even more accurate model for predicting passenger satisfaction. The researcher uses the two most significant variables from the above models and calculates their probabilities through the Naïve Bayes algorithm. While the decision tree algorithm employed all the available data, the researcher randomly selected 1000 data samples for testing. This sample selection will provide more transparent and more understandable data visualization for the Naïve Bayes algorithm. The results are depicted in Figure 11.

The data in Figure 12 shows that the model has an accuracy of 86.67%. Thus, this model is the best predictor of future passenger satisfaction. Furthermore, to simplify the analysis, the researcher also provides a visualization for the confusion matrix of the Naïve Bayes algorithm. It can be observed that the number of true positives and true negatives is higher compared to false positives and false negatives. True positives and true negatives represent accurate predictions, while false positives and false negatives denote prediction errors. In total, there are 260 correct predictions and 40 incorrect predictions. Naïve Bayes Model on Figure 12.

```
                         neutral or dissatisfied satisfied
  neutral or dissatisfied                     157        17
  satisfied                                    23       103

               Accuracy : 0.8667
                 95% CI : (0.8229, 0.903)
    No Information Rate : 0.6
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.7245

 Mcnemar's Test P-Value : 0.4292

            Sensitivity : 0.8722
            Specificity : 0.8583
         Pos Pred Value : 0.9023
         Neg Pred Value : 0.8175
             Prevalence : 0.6000
         Detection Rate : 0.5233
   Detection Prevalence : 0.5800
      Balanced Accuracy : 0.8653

       'Positive' Class : neutral or dissatisfied
```

Figure 12. Naïve Bayes Model Performance

Business travelers might place a higher value on efficiency and premium services. Moreover, the satisfaction uptick for online boarding and check-in when ratings reach four or higher is significant. It suggests that passengers who experience a smoother and more efficient online boarding and check-in process are considerably more satisfied. It emphasizes the importance of user-friendly online check-in systems for improving overall satisfaction. Finally, the upward trajectory in satisfaction from a rating of 3 and beyond for in-flight entertainment suggests that passengers appreciate having various entertainment options during their flight. It can significantly enhance their travel experience. Density Plot Per variables on Figure 13.

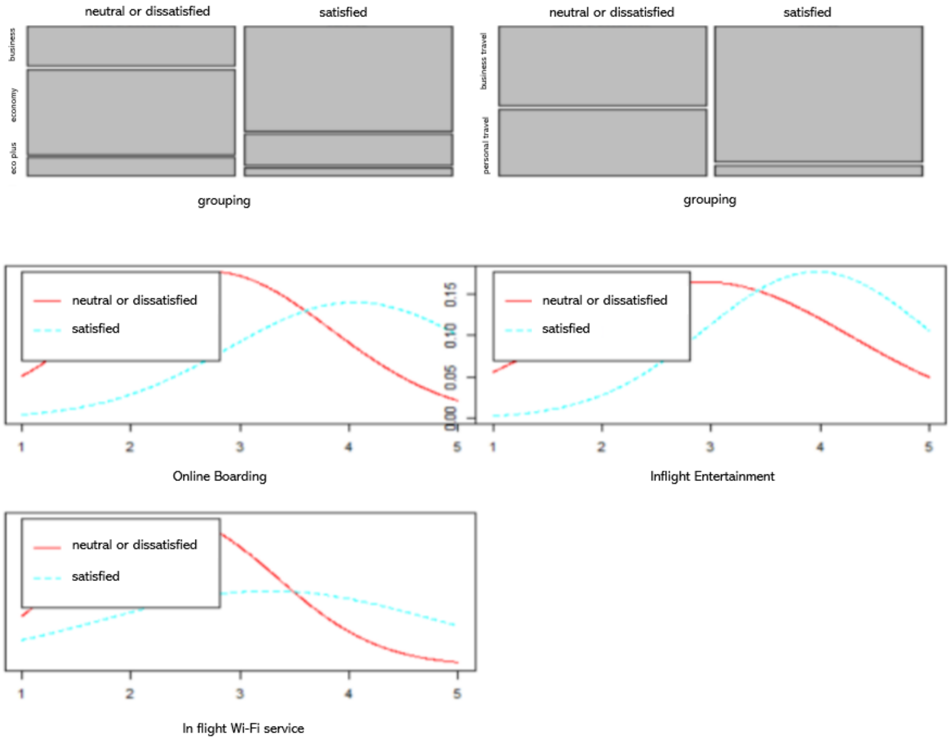

Figure 13. Density Plot Per variables

The density plot in Figure 13 offers valuable information about the factors influencing passenger satisfaction. The plot clearly shows that the highest satisfaction levels are found in the business class. It suggests that passengers flying in business class are generally more satisfied with their experience. Business class passengers often expect and receive premium services and comfort, contributing to their higher satisfaction levels. Afterwards, the plot related to travel types also indicates that the most excellent satisfaction is associated with business trips. It is an interesting finding as it highlights that passengers traveling for business tend to have higher satisfaction levels than those traveling for personal reasons.

4. Conclusion

From the data analysis above, we can conclude that all components in the flight influence passenger satisfaction. The decision tree algorithm can identify the variables with the top nodes (the most significant) as online boarding, inflight entertainment, WiFi service, class, and type of travel. Both decision tree models, ctree, and rpart, are equally suitable for creating a decision tree. As for the Naïve Bayes algorithm, creating a predictive model with significant variables can achieve a high accuracy of around 87%, as seen in the data distribution of satisfaction in the density plot that has been created.

Based on the research findings and discussions, several recommendations can be offered. Airlines should focus on developing and maintaining online boarding and check-in services, enhancing in-flight entertainment options, and adding WiFi services for long-distance flights to improve passenger satisfaction. Passengers should select their flight class according to their specific needs, with business class being a favorable choice for long-distance business travel. Regarding the algorithms used in this study, decision tree algorithms are valuable for uncovering patterns and extracting insights from raw data. In contrast, the Naïve Bayes algorithm effectively builds predictive models, mainly when dealing with limited data resources.

Acknowledgment

We extend our heartfelt appreciation to Multimedia Nusantara University for their unwavering support, provision of facilities, and financial assistance that were instrumental in the successful execution and completion of our research.

References

[1] Ashwika, Dishali G K, & Hemalatha N. (2020). Airline Passenger Satisfaction Prediction Using Machine Learning Algorithms. *Redshine Archive*, *1*, 8–24. DOI: https://doi.org/10.25215/8119070682.24 .

[2] Ban, H. J., & Kim, H. S. (2019). Understanding Customer Experience and Satisfaction Through Airline Passengers' Online Reviews. *Sustainability (Switzerland)*, *11*(15), 4066. DOI: https://doi.org/10.3390/su11154066 .

[3] Hulliyah, K. (2021). Predicting Airline Passenger Satisfaction with Classification Algorithms. *IJIIS: International Journal of Informatics and Information Systems*, *4*(1), 82–94. DOI: https://doi.org/10.47738/ijiis.v4i1.80 .

[4] Ludwig, S. A., Picek, S., & Jakobovic, D. (2018). Classification of Cancer Data: Analyzing Gene Expression Data Using A Fuzzy Decision Tree Algorithm. *International Series in Operations Research and Management Science*, *262*, 327–347. DOI: https://doi.org/10.1007/978-3-319-65455-3_13 .

[5] Mustopa, A., Wildah, S. K., Wijaya, G., Gata, W., & Agustiani, S. (2020). Pengaruh Media terhadap Pengambilan Keputusan dalam Menjalankan Program Keluarga Berencana dengan Algoritma Decision Tree. *Paradigma - Jurnal Komputer dan Informatika*, *22*(2), 145–152. DOI: https://doi.org/10.31294/p.v22i2.8141 .

[6] Nurdina, A., & Puspita, A. B. I. (2023). Naive Bayes and KNN for Airline Passenger Satisfaction Classification: Comparative Analysis. *Journal of Information System Exploration and Research*, *1*(2). DOI: https://doi.org/10.52465/joiser.v1i2.167 .

[7] Ranggadara, I., Wang, G., & Kaburuan, E. R. (2019). Applying Customer Loyalty Classification with RFM and Naïve Bayes for Better Decision Making. *Proceedings - 2019 International Seminar on Application for Technology of Information and Communication: Industry 4.0: Retrospect, Prospect, and Challenges, ISemantic* 2019, 564–568. DOI: https://doi.org/10.1109/ISEMANTIC.2019.8884262 .

[8] Sezgin, E., & Yuncu, D. (2016). The SWOT Analysis of Turkish Airlines Through Skytrax Quality Evaluations in the Global Brand Process. *Development of Tourism and the Hospitality Industry in Southeast Asia*, 65–81. DOI: https://doi.org/10.1007/978-981-287-606-5_5 .

[9] Yunus, W., Desanti, R. I., & Wella, W. (2020). Data Visualization And Sales Prediction of PD. Asia Agung (Ajinomoto) Pontianak in 2019. *IJNMT (International Journal of New Media Technology)*, *7*(2), 51–57. DOI: https://doi.org/10.31937/ijnmt.v7i2.1697 .

[10] Zhang, H., Jiang, L., & Yu, L. (2020). Class-specific attribute value weighting for Naive Bayes. *Information Sciences*, *508*, 260–274. DOI: https://doi.org/10.1016/j.ins.2019.08.071 .

[11] Aileen Chun Yueng Hong, Khaw, K. W., Xinying Chew, & Wai Chung Yeong. (2023). Prediction of US airline passenger satisfaction using machine learning algorithms. *Data Analytics and Applied Mathematics (DAAM)*, 8–24. DOI: https://doi.org/10.15282/daam.v4i1.9071 .

[12] Botchey, F. E., Qin, Z., & Hughes-Lartey, K. (2020). Mobile Money Fraud Prediction-A Cross-Case Analysis on The Efficiency of Support Vector Machines, Gradient Boosted Decision Trees, and Naïve Bayes Algorithms. *Information (Switzerland)*, *11*(8). DOI: https://doi.org/10.3390/INFO11080383 .

[13] Dinesh, T. (2021). Higher Classification of Fake Political News Using Decision Tree Algorithm Over Naive Bayes Algorithm. *Revista Gestão Inovação e Tecnologias*, *11*(2), 1084–1096. DOI: https://doi.org/10.47059/revistageintec.v11i2.1738 .

[14] Noviriandini, A., & Nurajijah, N. (2019). Analisis Kinerja Algoritma C4.5 Dan Naïve Bayes Untuk Memprediksi Prestasi Siswa Sekolah Menengah Kejuruan. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, *5*(1), 23–28. DOI: https://doi.org/10.33480/jitk.v5i1.607 .

[15] Maheswari, S., & Pitchai, R. (2018). Heart Disease Prediction System Using Decision Tree and Naive Bayes Algorithm. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *15*(8), 712–717. DOI: https://doi.org/10.2174/1573405614666180322141259 .

[16] Yogesh, L., Arunadevi, M., & Prakash, C. P. S. (2021). Predicton of MRR & surface roughness in wire EDM machining using decision tree and naive bayes algorithm. *In 2021 International Conference on Emerging Smart Computing and Informatics, ESCI 2021* (pp. 527–532). Institute of Electrical and Electronics Engineers Inc. DOI: https://doi.org/10.1109/ESCI50559.2021.9396857 .

[17] Mabe-Madisa, G. V. (2018). A Decision Tree and Naïve Bayes Algorithm for Income Tax Prediction. *African Journal of Science, Technology, Innovation and Development*, *10*(4), 401–409. DOI: https://doi.org/10.1080/20421338.2018.1466440 .

[18] Sumpena, S., Akbar, Y., Nirat, N., & Hengky, M. (2019). ICU Patient Prediction for Moving with Decision Tree C4.5 and Naïve Bayes Algorithm. *SinkrOn*, *4*(1), 88. DOI: https://doi.org/10.33395/sinkron.v4i1.10150 .

[19] Permana, A. P., Ainiyah, K., & Holle, K. F. H. (2021). Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up. *JISKA (Jurnal Informatika Sunan Kalijaga)*, *6*(3), 178–188. DOI: https://doi.org/10.14421/jiska.2021.6.3.178-188 .

[20] Putri, T. A. Q., Triayudi, A., & Aldisa, R. T. (2023). Implementasi Algoritma Decision Tree dan Naïve Bayes Untuk Klasifikasi Sentimen terhadap Kepuasan Pelanggan Starbucks. *Journal of Information System Research (JOSH)*, *4*(2), 641–649. DOI: https://doi.org/10.47065/josh.v4i2.2949 .