**Title Page 4**

Predicting Accuracy for Automated Aero Assist Recommendation using Random Forest and Compared with Naive Bayes with Improved Accuracy

K.Varshini 1,K.Sashi Rekha2

K. Varshini1

Research Scholar,

Department of Computer Science and Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, Pincode: 602105.

kethuvarshini1556.sse@saveetha.com

K.Sashi rekha 2

Research Guide, Corresponding Author,

Department of Computer Science and Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, Pincode: 602105.

sashirekhak.sse@saveetha.com

**Keywords:** Novel Random Forest, Naive Bayes, Airline Reservation System, Airline, Travel, Flight, Ticket, Reservation.

**ABSTRACT**

**Aim:** The importance of flight information systems for airlines is the target of this effort. By optimizing flight ticketing, sales, and air travel operations with the help of the Airline Reservation System, we want to strengthen the bonds between airlines and their customers. **Materials and Methods:**The research project comprises two segments. Group 1 involves the innovative Random Forest, while Group 2 focuses on Naive Bayes. Employing an 80% G-power and conducting 10 iterations, the evaluation was conducted on 10,000 sample sets. With a 95% confidence level and using a specific programming language, the identification utilized a Kaggle dataset containing twenty-five thousand field files.**Results:** The most recent enhancement to the Random Forest algorithm achieved a precision of 96.57%, but the newly developed Naive Bayes approach yields a precision of 86.52%. The observed data has substantial statistical ramifications. The p-value is 0.016.The statistical significance level is less than 0.05. **Conclusion:** The Novel Random Forest approach demonstrates superior performance in number identification compared to the Naive Bayes strategy.

**Keywords:** Novel Random Forest, Naive Bayes, Airline Reservation System, Airline, Travel, Flight, Ticket, Reservation.

**INTRODUCTION**

The Airline Flight Information System was once autonomous and functioned as distinct systems. Every airline had its own autonomous system that was not interconnected with that of other airlines or ticket agencies. Access to this system was restricted to a select group of approved airline staff. During the 1970s, travel agents campaigned for access to the airlines' systems. The current method of overseeing air travel data entails the interlinking, storage, and retrieval of information via a network of Computer Reservations Systems. [(Wittmer, Bieger, and Müller 2011)](https://paperpile.com/c/MQ35Et/XHDX) Direct consumer access to these technologies is made possible by the worldwide distribution system, which also allows for the integration of ticket sales and purchase across several airlines. [(Kasabov and Warlow 2012)](https://paperpile.com/c/MQ35Et/64z5) Users are able to purchase tickets, choose seats, and reserve hotels and rental cars directly via web-based gateways and portals that are part of the worldwide distribution system. [(Mckinsey Chief Marketing & Sales Officer Forum 2014)](https://paperpile.com/c/MQ35Et/0PKp) Several advantages come with crossing international boundaries.Reviews have an immediate effect on interactions with consumers. Either a single review (positive or negative) or several reviews (conflicting) could express an opinion. [(Smit et al. 2023)](https://paperpile.com/c/MQ35Et/soR5) A less-studied subject, conflicting online reviews have lately acquired prominence. Airline companies may use these ratings to find and correct service issues, and customers can use them to pick an airline. [(Rana et al. 2019)](https://paperpile.com/c/MQ35Et/sdZX) The problem of conflicting airline reviews is being tackled by a new research paradigm that seeks for passenger impressions that cause uncertainty or ambivalence while making a choice [(National Research Council et al. 2011)](https://paperpile.com/c/MQ35Et/vKdP) Airline applications include mobile booking, in-flight entertainment, and crew scheduling apps for passengers, while operational systems manage reservations, revenue, and maintenance to optimize overall airline efficiency.

This paper reviews the existing literature on the subject and explores the possible policy and social effects of autonomous driving, while also proposing avenues for further research. Using the concept of ripple effects, we may classify the impacts of autonomous vehicles into three levels: first-order (traffic, travel cost, travel alternatives), second-order (vehicle ownership, location, land use, and transport infrastructure), and third-order (all three). Google Scholar and Science Direct, two databases of highly-cited research papers, had 772 articles related to this area. To automate ticketing and booking and to provide all of its offices with real-time flight data, American Airlines needed a system in the late 1950s, and so the first airline reservation system was born. The Semi-Automated Business Research Environment was launched in 1964 as a result of this. [(Cook and Billig 2017)](https://paperpile.com/c/MQ35Et/cpLD) Agents all across the globe now have access to Sabre's innovative real-time inventory accuracy. Traditional, paper-based systems need central reservation hubs that use real cards to represent inventory, such as airplane seats. [(International Civil Aviation Organization 2004)](https://paperpile.com/c/MQ35Et/JfC2). Due to the Civil Aeronautics Board's stringent controls on airline routes and price, early American commercial aviation carried few passengers. [(Cusick, Cortes, and Rodrigues 2017)](https://paperpile.com/c/MQ35Et/QsyL) With the electromechanical Reservisor, American Airlines was the first to provide automated booking in 1946. The next machine in line was the innovative Magnetronic Reservisor, which used a magnetic drum for temporary storage. Many companies, including Goodyear, airlines, and Sheraton Hotels, have found success using this kind of inventory management[(Cheverton 2008)](https://paperpile.com/c/MQ35Et/Yp4J) In order to read the results, ticketing agents had to call a booking office, where operators gave a small team the instructions to run the Reservisor. The agents could not immediately contact the system for any inquiries. [(International Association for Food Protection 2016)](https://paperpile.com/c/MQ35Et/hKMe) An Airline Reservation System is part of Passenger Service Systems, which enables direct passenger engagement. [(Krajangta, Mahidon, and Witthayāsāt 1997)](https://paperpile.com/c/MQ35Et/SzWq). The Computer Reservations System was formerly known as the Airline Reservation System. via the use of a Global Distribution System, a Computer Reservation System facilitates the booking of major airlines via various channels, including travel agencies. [(Ren 2000)](https://paperpile.com/c/MQ35Et/ayDV)

The research gap identified Online platforms differentiating features are used to describe online platforms. Review sites, online stores, social media, blogs, and forums are all examples of online communities that get reviews. Customers who have recently used a product or service (e.g., text-based opinion) are the most likely to provide ratings and reviews online. According to studies, customer feedback has a major influence on new product development. The research evaluates the advice of airline critics by analyzing online reviews. Service providers may use the results of this research to better understand the impact of customer evaluations on their marketing efforts. Previous findings are expanded upon in this study. Examine the correlation between online reviews and suggestions for services. [(Joseph Pine, Pine, and Gilmore 1999)](https://paperpile.com/c/MQ35Et/2YSb) In order to help both customers and airlines, this article offers suggestions on how to handle ambivalence. We used natural language processing techniques to preprocess traveler reviews in the recommender system, which was the first of two contributions provided by this work. Secondly, data was gathered from different social networks by building the Convolutional Neural Network model. [(Jayakumari et al. 2020)](https://paperpile.com/c/MQ35Et/yV6y)

**MATERIALS AND METHODS**

The training and testing of airline reservation systems are performed in the open source Laboratory,Saveetha School of Engineering and Saveetha Institute of Medical and Technical Sciences. This group used in this study are two , a Random Forest, then a new Logistic Regression. Clinal determined sample size. The sample size was calculated using a 0.05 threshold, 80% G power, and 95% confidence range from previous studies.

Windows (version 11), Jupyter Notebooks, Google Collab, and 512 MB of Intel Core i5 RAM are some of the other features. The statistical assessment for the precision was conducted using IBM SPSS. The data used to train the algorithms has 25,000 rows and 24 columns, with a batch size of 10,000. Data from the Kaggle website was extracted in the form of CSV files. These files include various pieces of information, such as customer sex, age, travel type, flight class, distance, wifi while flying, good departure/arrival time, online booking, and more. To compare the two approaches for dependability, a second T-test research is conducted using SPSS.

**Random Forest:**

Random Forest machine learning can do regression and classification. A decision tree "forest" is created by creating several decision trees. The Random Forest algorithm aggregates all tree projections to get the final forecast. The Random Forest approach is strong and frequently used because it can handle complicated non-linear relationships, missing data, and outliers and produce a probabilistic result estimate.

**Random forest Algorithm:**

Step 1: Start by loading data.

Step 2: Divide the data into training and testing.

Step 3: Build a Random Forest classifier using hyperparameters like n\_estimators, maximum depth, etc.

Step 4: Train Random Forest classifier using data.

Step 5: Forecast using test results.

Step 6: Assess model accuracy, precision, and recall.

Step 7: Print or save model performance evaluation findings.

**Naive bayes:**

Classification using Naive Bayes is straightforward and effective. This formula employs Bayes' theorem to estimate the probability of an occurrence given a prior one.Naive Bayes classifiers presume class label-independent data point attributes. Though silly, this assumption usually works. A Naive Bayes classifier calculates a data point's class probability before categorizing it. Multiply each class's data point feature chances. Data points go to the most likely class.

**Naive bayes Algorithm:**

Step 1:Set up your picture dataset by loading and labeling photos.

Step 2: Split the dataset into training and testing.

Step 3: Determine model architecture.

Step 4: Compile the Naive bayes model with these settings:

Step 5: Train the Naive Bayes model using training data using 'fit'.

Step 6: Evaluate the model on test data using suitable measurements.

Step 7: Adjust hyperparameters and regularization to improve results.

Step 8 saves the learned model for subsequent use or deployment.

**Statistical Analysis**

Statistical analysis performed using the SPSS tool. The dependent variables in this study file include information such as Identification, Sex, Customer Category, Age, Travel Type, Class Flight, Distance, Wireless internet in flight, Good departure/arrival time, and online booking with improved accuracy, which are influenced by the independent factors.

**RESULT**

One new Random Forest classifier that works well with Naive Bayes is shown in Table 1. You can see how well Random Forest and Naive Bayes do in terms of accuracy in the table below. Airline flight information systems are vital, and this study is going to highlight that. involves taking a large dataset and breaking it down into smaller pieces.

One fresh method is shown in Table 2, which is a hybrid of the Random Forest and Naive Bayes algorithms for classifiers. Independent sample T-tests In terms of significance, mean, and standard error, the table below compares the Novel Random Forest method to Naive Bayes. The SPSS analysis was used to identify these differences, and the results showed statistical significance (p<0.05) with a p-value of 0.016.

The statistical statistics about the performance of the Random Forest and Naive Bayes algorithms are shown in Table 3. Both the Naive Bayes and Random Forest models' means, standard deviations, and standard errors were computed by entering the findings from both approaches into SPSS. Naive Bayes achieves an accuracy of 96.57%, whereas Novel Random Forest achieves an accuracy of 86.52%.

The results show that compared to Naive Bayes' 86.52% success rate, the proposed method has a 96.57% success rate. An accompanying 95% confidence range and standard deviation of +/- 2 SD are provided with these values.

**DISCUSSION**

In comparison to the Naive Bayes classifier, the Random Forest method outperformed it with higher accuracy (96.57%) and precision (86.52%). Examining the visitors' recommendation system via various evaluations was the purpose of the research. There is a statistical significance (p-value = 0.016, < 0.05). Based on the results of these tests, Naive Bayes is not very good at detecting sentiment. A Random Forest is proposed as a possible solution in this research.

When it comes to automating the organization's activities, the Airline management system is crucial. It helps the company's dairy operations run smoothly. The company's bottom line has benefited from the automation of its processes. [(Wittmer, Bieger, and Müller 2011)](https://paperpile.com/c/MQ35Et/XHDX) The findings will help management improve income generation and streamline day-to-day operations. Anyone involved with the operation of flights, air travel, travel agencies, flight operators, or airline reservation systems will find this presentation useful. [(Whiteley 2017)](https://paperpile.com/c/MQ35Et/1ivq) Consumers vent their feelings in online forums. These feelings may be positive or negative; sometimes, they might be a mix of the two, leading to conflicting assessments. Twitter, Instagram, and e-commerce sites are just a few examples of the many online venues where customers may share their opinions and ideas on businesses, products, and services. [(Davison et al. 2023)](https://paperpile.com/c/MQ35Et/pN7F) There is a higher chance that customers may incorrectly perceive information when airline firms and consumers engage in contradictory, disputed, or bivalent behaviors, according to the research. By their very nature, they make one feel conflicted and uncertain. [(Christou 2006)](https://paperpile.com/c/MQ35Et/7ai0) Therefore, we argue that consumers experience moral ambiguity when confronted with objects that include paradoxes, like online reviews. The object has made them feel conflicted and unsure, which is the root cause of their ambivalence. Customers end up confused and unable to make a decision. [(U.S. Department of Health and Human Services 2019)](https://paperpile.com/c/MQ35Et/HRmR)

In an online environment, there is a correlation between e-loyalty, perceived value, and the quality of e-services. Further, I looked at how a tourist's experience relates to the existence of a perceived value-satisfaction-loyalty relationship. Consistent with previous research, the results show that both perceived value and satisfaction significantly impact happiness and loyalty. Nevertheless, further research is necessary. [(Gale 2009)](https://paperpile.com/c/MQ35Et/UTNf) Furthermore, it will benefit those in computer-related fields who want to have a comprehensive understanding of the system, as well as those doing research on related subjects. [(Marwedel 2021)](https://paperpile.com/c/MQ35Et/Drjr)

**CONCLUSION**

This study aimed to analyze tourist reviews from various sources to assess the effectiveness of the recommendation system, employing the Random Forest model for outcome calculation. Evaluation metrics, including F-measure, Precision, and Recall, were considered for result quality improvement. Through N=10 accuracy iterations, the investigation explored recommendation system functioning, feature extraction, and the impact of feature count. Despite Naive Bayes achieving an accuracy of 86.52%, Random Forest outperformed with a classification accuracy of 96.57%, offering more relevant and insightful results. The comparison yields statistical significance at p=0.016 (p<0.05)."

**DECLARATIONS**

**Conflicts of interest**

There is no reported conflict of interest in this study.

**Author's Contribution**

Author KV was responsible for collecting data, analyzing the data, and composing the manuscript of Natural Language Processing. Author KSR provided conceptual direction, supervised the work, and thoroughly reviewed the manuscript.

**Acknowledgements**

The authors extend their appreciation to the Saveetha School of Engineering and Saveetha Institute of Medical and Technical Science (formerly Saveetha University) for offering the necessary support to successfully complete this project.

**Funding**

The authors express their gratitude to the following organizations for their financial support, which made it possible to carry out the study.

1. CK technology Pvt. Ltd Chennai.
2. Saveetha School of Engineering.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha University.

**REFERENCES**

[Cheverton, Peter. 2008. *Global Account Management: A Complete Action Kit of Tools and Techniques for Managing Key Global Customers*. Kogan Page Publishers.](http://paperpile.com/b/MQ35Et/Yp4J)

[Christou, Anastasia. 2006. *Narratives of Place, Culture and Identity: Second-Generation Greek-Americans Return “Home.”* Amsterdam University Press.](http://paperpile.com/b/MQ35Et/7ai0)

[Cook, Gerald N., and Bruce G. Billig. 2017. *Airline Operations and Management: A Management Textbook*. Routledge.](http://paperpile.com/b/MQ35Et/cpLD)

[Cusick, Stephen K., Antonio I. Cortes, and Clarence C. Rodrigues. 2017. *Commercial Aviation Safety, Sixth Edition*. McGraw Hill Professional.](http://paperpile.com/b/MQ35Et/QsyL)

[Davison, Michelle, Jason Chan, Meg Clarke, Caroline Mitchell, Alan Yan, and Elwyn Henaway. 2023. “Yarning to Reduce Take Own Leave Events in First Nations Patients Presenting to the Emergency Department-Presenting the Qualitative Themes and Co-Design of the Deadly RED Project.” *Health Promotion Journal of Australia: Official Journal of Australian Association of Health Promotion Professionals*, December. https://doi.org/](http://paperpile.com/b/MQ35Et/pN7F)[10.1002/hpja.835](http://dx.doi.org/10.1002/hpja.835)[.](http://paperpile.com/b/MQ35Et/pN7F)

[Gale, Bradley. 2009. *Managing Customer Value: Creating Quality and Service That Customers Can See*. Simon and Schuster.](http://paperpile.com/b/MQ35Et/UTNf)

[International Association for Food Protection. 2016. *Procedures to Investigate Waterborne Illness*. Springer.](http://paperpile.com/b/MQ35Et/hKMe)

[International Civil Aviation Organization. 2004. *Manual on the Regulation of International Air Transport*.](http://paperpile.com/b/MQ35Et/JfC2)

[Jayakumari, J., George K. Karagiannidis, Maode Ma, and Syed Akhter Hossain. 2020. *Advances in Communication Systems and Networks: Select Proceedings of ComNet 2019*. Springer Nature.](http://paperpile.com/b/MQ35Et/yV6y)

[Joseph Pine, B., Joseph Pine, and James H. Gilmore. 1999. *The Experience Economy: Work Is Theatre & Every Business a Stage*. Harvard Business Press.](http://paperpile.com/b/MQ35Et/2YSb)

[Kasabov, E., and A. Warlow. 2012. *The Compliance Business and Its Customers: Gaining Competitive Advantage by Controlling Your Customers*. Springer.](http://paperpile.com/b/MQ35Et/64z5)

[Krajangta, Jurairat, Mahāwitthayālai Mahidon, and Mahāwitthayālai Mahidon Khana Witthayāsāt. 1997. *Self Service Airline Reservation System VIA Internet*.](http://paperpile.com/b/MQ35Et/SzWq)

[Marwedel, Peter. 2021. *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things*. Springer Nature.](http://paperpile.com/b/MQ35Et/Drjr)

[Mckinsey Chief Marketing & Sales Officer Forum. 2014. *Big Data, Analytics, and the Future of Marketing and Sales*. Createspace Independent Pub.](http://paperpile.com/b/MQ35Et/0PKp)

[National Research Council, Division of Behavioral and Social Sciences and Education, Board on Behavioral, Cognitive, and Sensory Sciences, and Committee on Behavioral and Social Science Research to Improve Intelligence Analysis for National Security. 2011. *Intelligence Analysis: Behavioral and Social Scientific Foundations*. National Academies Press.](http://paperpile.com/b/MQ35Et/vKdP)

[Rana, Nripendra P., Emma L. Slade, Ganesh P. Sahu, Hatice Kizgin, Nitish Singh, Bidit Dey, Anabel Gutierrez, and Yogesh K. Dwivedi. 2019. *Digital and Social Media Marketing: Emerging Applications and Theoretical Development*. Springer Nature.](http://paperpile.com/b/MQ35Et/sdZX)

[Ren, Dazhou. 2000. *An Online Airline Reservation System with XML*.](http://paperpile.com/b/MQ35Et/ayDV)

[Smit, Chloé C. H., Maarten Lambert, Kris Rogers, Steven P. Djordjevic, Antoine M. Van Oijen, Caitlin Keighley, Katja Taxis, Hamish Robertson, and Lisa G. Pont. 2023. “One Health Determinants of Antimicrobial Resistance in Humans in the Community: An Umbrella Review.” *International Journal of Molecular Sciences* 24 (24). https://doi.org/](http://paperpile.com/b/MQ35Et/soR5)[10.3390/ijms242417204](http://dx.doi.org/10.3390/ijms242417204)[.](http://paperpile.com/b/MQ35Et/soR5)

[U.S. Department of Health and Human Services. 2019. *TIP 35: Enhancing Motivation for Change in Substance Use Disorder Treatment (Updated 2019)*. Lulu.com.](http://paperpile.com/b/MQ35Et/HRmR)

[Whiteley, David. 2017. *An Introduction to Information Systems*. Bloomsbury Publishing.](http://paperpile.com/b/MQ35Et/1ivq)

[Wittmer, Andreas, Thomas Bieger, and Roland Müller. 2011. *Aviation Systems: Management of the Integrated Aviation Value Chain*. Springer Science & Business Media.](http://paperpile.com/b/MQ35Et/XHDX)

**TABLES AND FIGURES**

**Table 1.** Accuracy comparison between Naive Bayes and Random Forest Classifier. The following table displays the precision of Random Forest and Naive Bayes.

|  |  |  |
| --- | --- | --- |
| **S. No** | **Random Forest** | **Naive Bayes** |
| 1 | 96.16 | 86.94 |
| 2 | 96.61 | 86.50 |
| 3 | 96.02 | 86.03 |
| 4 | 96.99 | 86.52 |
| 5 | 96.34 | 86.75 |
| 6 | 96.76 | 86.95 |
| 7 | 96.93 | 86.30 |
| `8 | 96.97 | 86.81 |
| 9 | 96.83 | 86.01 |
| 10 | 96.66 | 86.34 |

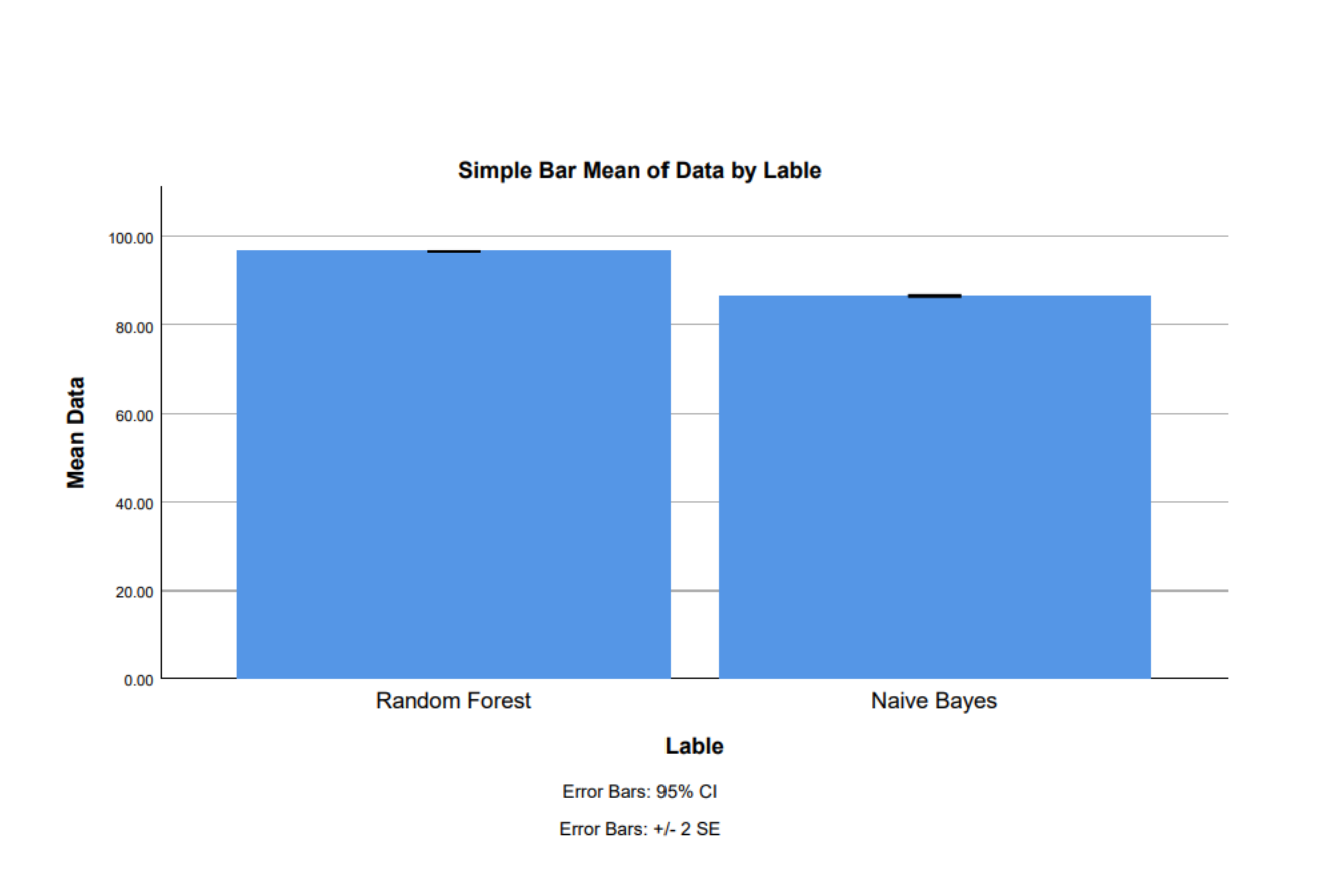
|  |
| --- |
|  |

**Table 2**. Random Forest using Naive Bayes classifier and independent sample T-tests. The table below presents a comparison of the significance, mean, and standard error differences between the Novel Random Forest and Naive Bayes approaches. These comparisons were made using SPSS analysis, with a significance level of p = 0.016 (p<0.05).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Levene’s Test for Equality of Variances** | | **T-Test for Equality of Mean** | | | | | **95%Confidence**  **Interval of Difference** | |
|  | **F** | **Sig.** | **t** | **df** | **Sig. (2-tailed)** | **Mean Difference** | **Std. Error Difference** | **Lower** | **Upper** |
| Equal variances assumed | 7.10 | 0.016 | 84.65 | 18 | 0.000 | 10.05 | 0.118 | 9.803 | 10.302 |
| Equal variances assumed |  |  | 84.65 | 12.08 | 0.000 | 10.05 | 0.118 | 9.794 | 10.311 |

**Table 3.** The method outputs for both Random Forest and Naive Bayes were used in SPSS to get the mean, standard deviation, and standard error. The Random Forest model achieves an accuracy of 96.57%, whereas the Naive Bayes model achieves an accuracy of 86.52%.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Algorithm** | **N** | **Mean** | **Std.Deviation** | **Std.Error Mean** |
| **Accuracy** | Novel Random Forest | 10 | 96.57 | 0.145 | 0.046 |
| Naive Bayes | 10 | 86.52 | 0.346 | 0.109 |

****

**Fig. 1.** Statistical analysis was conducted using the SPSS tool to determine the accuracy of the data. The findings of the suggested method yielded an accuracy of 96.57% using Random Forest. This accuracy was compared to a baseline accuracy of 86.52% which is Naive Bayes, with an error bar of 95% confidence interval and a standard deviation of +/- 2.