Prediction of Customer Satisfaction in Airline Hospitality services for improved service delivery using Support Vector Machine

[1]Adedayo A. Sobowale, [2]Olukemi O. Osadare[2],Afeez A. Soladoye[3]* and Peace O. Sobowale[4]
[1,3]Department of Computer Engineering, Federal University Oye-Ekiti
[2]Department of Hospitality Technology Management, Federal Polytechnic, Ede
[4]Daboss Polytechnic Idominasi
adedayo.sobowale@fuoye.edu.ng| osadare.olukemi@federalpolyede.edu.ng| afeez.soladoye@fuoye.edu.ng| |nnamanipeace@gmail.com

Received: 26-MAY-2024; Reviewed: 20-JUNE-2024; Accepted: 21-JUNE-2024
https://dx.doi.org/10.4314/fuoyejet.v9i2.13

ORIGINAL RESEARCH

Abstract — The recent entry of Air peace airline service to London and competition resulting from this with other airlines showcased the trending competitiveness in airline industry which might affect the patronage of an airline over others. In line of this, ensuring customer's satisfaction would enable business growth, sustainability and provision of improved service delivery. Gauging customer satisfaction would help in knowing steps to take to ensure better service delivery but employing traditional approach might be subjected to travel experience, time consumption, inaccurate which would hinder its objective. However, application of Machine learning techniques would give a faster, accurate and concise prediction resulting from historical customer's reviews. This study aims to employing Support Vector Machine a traditional machine learning classifier for prediction of airline customer's satisfaction in hospitality services employing an open access dataset comprising of different attributes like age, gender, flight distance among others. This dataset was pre-processed using various pre-processing techniques like normalization using standard scalar, transformation using label encoder, removal of missing values, feature selection using forward-backward feature selection techniques. The well pre-processed dataset was implemented using the SVM classifier with its three kernels for comparative analysis. The study used hold-out evaluation method with 80-20 split and gave average accuracies with Rbf, Linear and poly kernels of 93, 85 and 90% respectively. This study showed that the RBF kernel of SVM gave the best predictive performance with average accuracy, precision and F1-score of 93, 93 and 92% respectively with its good predictive airlines can identify the major factors that contribute to customer's satisfaction to prioritize to improve service delivery and optimization of staffing at during strategic period. This would enhance customer's satisfaction and ensure improved good service delivery.

Keywords Airline services, customer satisfaction, machine learning, support vector machine,
.

1 INTRODUCTION

Ensuring customer's satisfaction is one of the major measures and factor in hospitality industry. Airline service is mainly based on hospitality as this determines the leisure, comfort and hitch free experience customers witness during the course of their journey. The growing competition in the aviation sector and the recent breakthrough achieved by airpeace airlines with their ground-breaking entrance to the Europe and their acceptability and adoption by Nigerians going to United Kingdom showed that their indigenous hospitality techniques like food, movies, ushering among other further triggered their patronage as well as their cheaper ticket. Identifying Customer's satisfaction in due time would help in placing good policy in place to further enhance this satisfactory services or better some unsatisfied ones as it gives business insight on the best action to take to ensure increase patronage from enhanced customer satisfaction (Bockhorst ,*et al.*2017).

Customer's satisfaction is based on their psychological and emotional need which have been backed by many theories and studies to explain customer's perception to services they received in hospitality industry, but this approach might be inaccurate of time consuming as customers might use their present unbalanced emotional state to give a review and this would count during cross examination of their

review. Moreover, ensuring even distribution of different hospitality attributes related to airline services and grading them by the level of their satisfaction and dissatisfaction by customers would really help in covering a large area regarding the prediction of their satisfaction (Oh et al., 2022).

Moreover, ensuring customer's satisfaction in airline service is prominent to ensure patronage and as earlier said, good hospitality in airline service is the best approach to achieve an improved customer's satisfaction by placing in place various hospitality techniques and services that would ensure smooth travelling and pleasant travel experience as their journey through the cloud progresses with ensured services quality and ensuring this is a key advantage over other competitors (Li, et al., 2017). Owing to the growing advancement and adoption of technology major Artificial Intelligence and Data science with respect to Hospitality and Tourism with prediction of customer's satisfaction though customer's hotel review, and information, managers were able to know that placing images and good painting at strategic areas in the hotel would boost customer patronage and satisfaction.

Employing conventional approach like statistical techniques or mixed methods like interview,

*Corresponding Author

Section B- ELECTRICAL/COMPUTER ENGINEERING & COMPUTING SCIENCES
Can be cited as:
Sobowale A.A., Osundare O.O., Soladoye A.A., and Sobowale P.O. (2024). Prediction of Customer Satisfaction in Airline Hospitality services for improved service delivery using Support Vector Machine, FUOYE Journal of Engineering and Technology (FUOYE**JET), 9(2), 241-245.** https://dx.doi.org/10.4314/fuoyejet.v9i2.13

questionnaire or survey to determine customer's satisfaction might be time consuming as statistically analysing such data might take a while and data collection might be prolonged. However, using the technological approach of electronic survey form and employing machine learning approach to analyse the data would be easier, faster and more accurate as it would help in knowing the best attributes with best significance to customer's satisfaction and help the hospitality industry to focus on those attributes for improved service delivery (Lucini *et al.* 2020).

This study aim to employ machine learning approach to prediction customers' satisfaction in airline hospitality services so as to improve delivery of services and enhance customer satisfaction for increased patronage using Support vector machine (SVM) as the classifier. Different kernels of SVM would be compared to check the best performing ones among them and to check the best combination of attributes that would give the best performance and in turn give insight on where and what to give focus.

2 RELATED WORKS

The widely acceptance of AI techniques to improvement of business has increased the focus of researchers on conducting various researcher in the adoption of AI and its techniques in hospitality industry. Some of these studies are reported and reviewed in this section to create a background knowledge regarding studies that have been carried out so far in this domain.

Chan, Kwong and Kremer (2020) developed framework to determine customer satisfaction using online customers review, this study employed empirical models as the classifiers with committee of member selection approach d=employed to motivate the undeterministic nature shown by Genetic programming. This framework focus on predicting customers' satisfaction in the design of electric hair dryer using customers reviews obtained from Amazon.com so as to check the level of their satisfaction using the product and help the manufacturer to improve the design or operation of the product. Oh *et al.* (2022) conducted a study on employing big data analytics to predict customer satisfaction in hospitality services based on customer's hotel review, information about different hotels considering and images of the hotel and services provided, the proposed fused model employed by the authors gave a very good predictive performance and managerial implication of hospitality was well presented as well as academic's.

Lee, Kwon and Back (2021) stated that big data analytics enables researchers and practitioners to tap and discover hidden fact and knowledge from information which would help in decision making process through big data. This made them to analyze 1,483,853 reviews by restaurant users acquired from and online platform, the authors employed four machine learning algorithms and comparing these models Extreme Gradient Boosting gave the best predictive performance for review helpfulness over SVM, RF and multivariate linear regression. This study showed that, predicting review helpfulness would help restaurant owner/managers to attract customers and employing this model would help in enhancing customer's attitude and reduced information overloading leading to enhanced and increase patronage of restaurant and sales. Similarly, predicting customers satisfaction from hostel reviews gathered online would increase the value of the hospitality business as discussed by Zhao, Xu and Wang (2019) in their study, where a sample of 127,629 reviews were analysed using different technical attributes. This study compared to earlier study focused majorly on hotel customers and used lesser instances than the latter. The study found out that long reviews are usually related to lower customer satisfaction and involvement of customer's review positively influence their satisfaction. This study focused majorly on understanding the customer's review and not just classification of this review based on their sentiment. Resulting from this study, the authors were able to show the hotel management various technical attributes that are included in their customer's review like subjectivity, readability, length, customer rating and they implore hotel management for further understand customer review by implementing good electronic review management system to enhance service delivery and increase patronage

3 METHODOLOGY

This study employed machine learning techniques leading to the generic methodology phases of data acquisition, pre-processing, prediction and evaluation. The overview of this methodology is shown in Figure 1 for clearer representation.

3.1 Data Acquisition

Primarily acquiring a raw data from airline hospitality customers would need some level of protocols and request, in order to ensure easier research conduct, this study employed an open access dataset available online from Kaggle on airline customer satisfaction prediction. This dataset consist of a total of 24 attributes namely: Age, Gender, type of travel, lass, Inflight service, inflight entertainment, food and drink, seat comfort, on-board service, baggage handling among other. Careful analysis of the listed attributed would obviously show that this is centered on the hospitality services provided to the customers by the airline service.

The publicly available dataset would have been well curated by professionals and ensures the inclusion of most significant attributes with instances collected over a long period of time from different airlines customers traveling over varying period of time.

https://dx.doi.org/10.4314/fuoyejet.v9i2.13 engineering.fuoye.edu.ng/journal

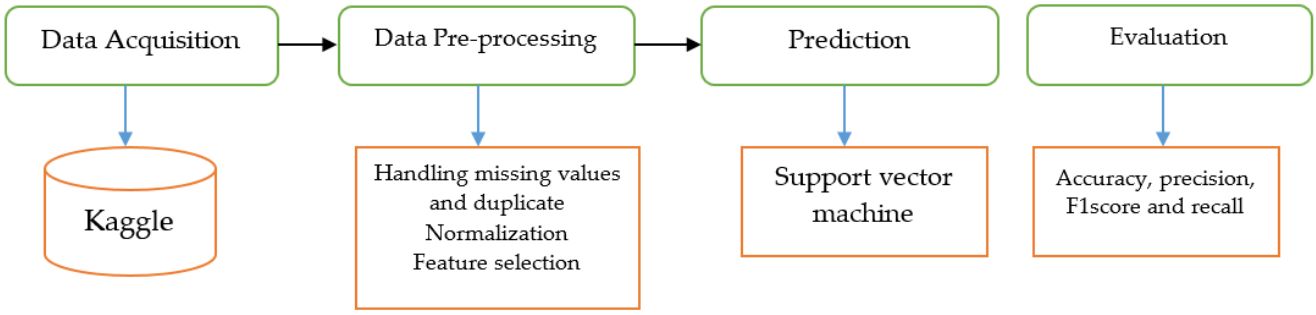

Figure 1: Overview of research methodology

3.2 DATA PRE-PROCESSING

Owing to the application of machine learning techniques with data analytics in this study, the nature of the data have great impact on the performance of the model and the expected outcome of the study. As a result of that, pre-processing the data would help in improving the model's performance and as a result of that, this study employed some pre-processing techniques on the data like Removal of duplicate values as this would reduce overfitting and redundancy. Duplicate instances or values would only increase the volume of the dataset without any significant impact in the performance of the model, therefore its removal is the best approach for handing duplicate values. Moreover, missing values were also handled the same way, however, there are mostly filled with the mean or mode of their columns, but the researchers decided to remove the missing values since the dataset is big already. Furthermore, there is usually variation in the range of the variables values, with big values out shinning and dominating the small values and this would greatly affect the attributes section and significance. In order to ensure all the variables are within the same range of values, standard scalar was used for standardization of the values, this would ensure that the data in the right form for processing and implementation on the desired machine learning algorithm.

3.3 SUPPORT VECTOR MACHINE

SVM is a supervised learning algorithms that is majorly used for classification and regression problem. SVM is based on the derivation of hyper plane and support vector. In a dataset, there are usually many decision boundaries that separate different class labels, these labels are usually related base on their attributes, the best decision boundary that helped in accurate classification of this multiple decision boundaries is what is known as the hyperplane (Qin, Zhou and Jiang, 2021). This algorithm have different kinds of kernels for implementation and three of these kernels namely: Linear, Polynomial and Gaussian Radial basis function were implemented in this study. This simple representation of SVM and it functionality for hyperplane is shown in Figure 2 for easier understanding and representation.

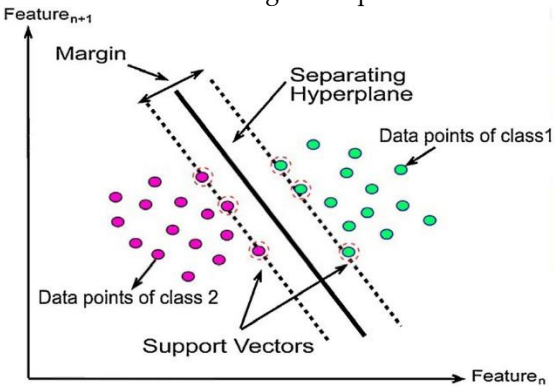

Figure 2: Support vectors machine (Sarp and Ozcelik, 2017)

As earlier discussed, the principle of SVM is based on hyperplane, this airline customer satisfaction dataset consists of varying attributes, these attributes might have instances with close values which might be regarded as the multiple decision boundaries, the hyperplane phenomenon of SVM would help in correct and accurate classification of the instances to the most correct class based on the decision made with the hyperplane.

Different kernels of SVM were also employed for this classification problem after which their experimental results were compared to know the best kernels for this problem domain.

3.4 PERFORMANCE EVALUATION

This study was evaluated using various performance evaluation metrics for its performance evaluation to know its predictive performance. Some of the metrics employed are

https://dx.doi.org/10.4314/fuoyejet.v9i2.13 engineering.fuoye.edu.ng/journal

accuracy, precision, f1score and recall as shown in Equation 1-4

$$\text{Accuracy} = \frac{TP+TN}{(TP+FP+TN+FN)} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{F1score} = \frac{2*precision*recall}{precision+recall} \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (4)$$

4 RESULTS AND DISCUSSION

During the pre-processing state, only 18 out of the 23 attributes were used as they gave the best predictive performance to the determination and prediction of customers satisfaction in airline hospitality service, the insignificant ones are age, gender, type of travel. These attributes were dropped as fundamentally they were known to have no significant effect on the decision of the customers for the satisfactory decision of the hospitality services provided by the airline while the major attributes that contributed to the customer satisfaction are the various hospitality services provided by the airline to its passengers like the type of food, drinks, comfort experienced, handling of their baggage, inflight services among others. This shows that hospitality services have a great impact to the satisfactions airline service customers derive from the industry. Moreover, no balancing technique was employed on the dataset, as the result obtained as shown in Table 1 does not show any effect of imbalanced data, as the macro average of the metrics does not show any effect of the imbalance nature of the dataset on the result obtained.

As earlier discussed, SVM was used as the machine learning classifier for prediction of customer's satisfaction using various hospitality attributes as the input to the classifier. The prediction was performed for the three aforementioned SVM kernels namely: Linear, Polynomial and Gussian RBF, this is done to compare the performance of these kernels and determine the best performing among them during the transformation of the input dataset into the required form needed and appropriate for classification.

The performance of these kernels were measured using the average performance evaluation metrics as discussed in the earlier section. These results is presented in Table 1 for clearer representation. As shown in Table 1, there are two classes present in the dataset namely 0 (unsatisfied) 1 (satisfied). These two classes were obtained from diverse customer satisfactory response as some stated that they were fully satisfied, partially satisfied and satisfied, all this categories were later renamed and classes as satisfied while other categories apart from these were classed as unsatisfied.

Table 1: Experimental Results of prediction of Customer's satisfaction in airline hospitality services

Class	Precision	Recall	F1-score	Support
Gaussian Polynomial Kernel				
0	0.88	0.95	0.92	14565
1	0.94	0.85	0.89	12226
Accuracy			0.90	26791
Macro avg	0..91	0.90	0.90	26791
Weighted avg	0.91	0.90	0.90	26791
Linear Kernel				
0	0.83	0.90	0.87	14565
1	0.87	0.79	0.83	12226
Accuracy			0.85	26791
Macro avg	0.85	0.84	0.85	26791
Weighted avg	0.85	0.85	0.85	26791
Gaussian RBF Kernel				
0	0.90	0.98	0.93	14565
1	0.97	0.87	0.91	12226
Accuracy			0.93	26791
Macro avg	0.93	0.92	0.92	26791
Weighted avg	0.93	0.93	0.93	26791

A total of 26791 customers were used to evaluate the model as the study employed hold-out evaluation method with 20% of the whole data amount to the aforementioned figure was used for testing.

The three aforementioned SVM kernels were used for comparison of their performance as shown in Table 1, Polynomial kernel gave an average accuracy of 90% which is actually a great predictive performance and the best precision was obtained from the satisfactory class with average precision of 94%, and the whole dataset gave an average micro and weighted precision of 91% respectively, this shows that the dataset were well processed and the SVM (polynomial) kernel was able to evenly understand the pattern in the two classes. Similarly, Linear Kernel of the SVM algorithm give a lower predictive performance than what was obtained from the Polynomial Kernel, however, the average predictive accuracy obtained was also encouraging with 85% and the macro and weighted average precision, f1-score and recall of 85%, 85%, 84%, 85%, 85% and 85% respectively. With the satisfactory class having the higher precision of 87% while unsatisfactory class have the best recall and f1-score of 90% and 87% respectively. This result obtained differs from that earlier presented as with linear kernel, unsatisfactory class have an edge over the satisfactory class as regards the recall and f1-score evaluation metrics.

Furthermore, the Gaussian RBF kernel gave the best performance as the SVM model when evaluated with the same metrics and testing size. This kernel gave an average accuracy of 93% which is the highest among them all and the macro and weighted average precision, f1-score and recall of 93%, 93%, 92%, 93%, 92% and 93% respectively. This is such an improved performance compared to the earlier presented results. This implies that for Binary classification problem with the employed dataset Gaussian RBF kernel is the best SVM model for implementation to give the best result.

Owing to this, SVM with Gaussian RBF kernel is a good model that can be implemented in real life to classify customer's satisfaction prediction in airline hospitality service for managers to understand the best attributes to consider in identifying customer's satisfaction and improve on their satisfaction. Additionally, airline services can explore the attributes aforementioned with significant effect to the satisfaction of their customers, so as to ensure these attributes are improved on to have improved satisfaction and patronage.

5 CONCLUSION AND RECOMMENDATION

This study have been able to employ Machine learning techniques to prediction customer's satisfaction ins airline hospitality service and SVM with three of its kernels were compared with Gaussian RBF kernel having the best predictive performance with the highest average accuracy, precision and F1sscore of 93, 93 and 92% respectively . From this result it is known that this kernel could be used when developing such model for real life application for effective and accurate result to be obtained. Moreover, hospitality services provided by the airline to their customers like food, drink, entertainment, inflight services among others are the major attributes that contributes to the prediction of customer's satisfaction and airline industry.

However, this study only employed an open-access dataset available to the researchers during the course of this study, future works might use locally acquired dataset so as to ensure its validation and recency, while different class balancing techniques can also be used so as to prevent bias and ensure even distribution of all the classes in the dataset. The future study should also ensure that the dataset that would be used might be review based so as to capture some recommendations given by the customers and analysis their sentimental response through the use of Natural Language Processing techniques, in order to improve delivery of services by the airline services.

REFERENCES

Bockhorst, J., Yu, S., Polania, L and Fung G. (2017). "Predicting self-reported customer satisfaction of interactions with a corporate call center" *Y.Altun et al. (Eds): ECML PKDD 2017, part III, LNAI 10536, pp.179-190*

Chan, K. T., Kwong, C. K. and Kremer, G. E. (2020). "Predicting customer satisfaction based on online reviews and hybrid ensemble genetic programming algorithms" *Engineering Applications of Artificial Intelligence, 95*

Lee, m., Kwon, W. and Back, K (2021). "Artificial intelligence for hospitality big data analytics: developing prediction model of restaurant review helpfulness for customer decision-making" *International Journal of Contemporary Hospitality, 33(6), pp. 2117-2136*

Lucini. F. R., Tonetto, L. M., Fogliatto, F. S. and Anzanello, M. J. (2020) "Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews" *Journal of Air transport Management, 83*

Oh, S., Ji, H., Kim, J., Park, E. and P.del, A. (2022). "Deep learning model based on expectation-confirmation theory to predict customer satisfaction in hospitality service" *Information Technology and Tourism, 24, pp.109-126*

Qin Q, Zhou X. And Jiang Y. (2021) "Prognosis Prediction of Stroke based on Machine Learning and Explanation Model" International Journal of Computers Communications & Control, 16(2). https://doi.org/10.15837/ijccc.2021.2.4108

Sarp, G and Ozcelik, M. (2017) "Water body extraction and change detection using time series: A case study lake Burdur, Turkey" *Journal of Taibah University of Science. 1,(1) 391-391*

Zhao, Y., Xu, X. and Wang, M. (2019)."Predicting Overall Customer satisfaction: Big data evidence from hotel online textual reviews." *International Journal of Hospitality Management. 76, pp. 111-121*

https://dx.doi.org/10.4314/fuoyejet.v9i2.13 engineering.fuoye.edu.ng/journal