Predicting Airline Passenger Satisfaction Using Different Machine Learning Models

Arun Reddy Yeduguru Madhulika Pimple Rajul Jain

University of Michigan-Dearborn

IMSE 586

Background

In the present day, airlines are competing more fiercely with one another in their quest to attract new consumers and retain consumers who previously used their services. Over the years, it has become imperative to investigate passenger satisfaction on airlines as customer's not only demand exceptional services but also low price for tickets [1]. This research aims to analyze important service features on flights, so airlines can re-strategise their business model and put majority focus on the services that play a crucial role in passenger satisfaction to gain more market share. This research will not only benefit the incumbent airlines but will also be useful for new prospects who plan to enter this industry.

Dataset

- We have used data from the US Airline Passenger Satisfaction survey available on Kaggle.
- The data set consists of 129880 rows and 24 columns of customer data from US airlines.
- The output variable is 'satisfied' and 'dissatisfied or neutral'.
- The data set includes categorical data as well as numerical data consisting of flight information and customer rating of the services on the scale of 0-5.

Below tables show a brief description of the data set:

Categorical Variable	Description	Categorical Variable	Description
Satisfaction	**Satisfied, neutral or dissatisfied**	Gender	Passenger gender (Female, Male)
Type of Travel	(Personal Travel, Business Travel)	Class	Travel class (Business, Eco, Eco Plus)
Type	Customer type (Loyal, disloyal)	Flight cancelled	Flight cancelled (Yes, No)

Numerical Variable	Description	Numerical Variable	Description
id	Passenger ID number	Online boarding	On 0-5 Scale
Age	Passenger age	Inflight entertainment	On 0-5 Scale
Flight distance	Flight distance of this journey	Food and drink	On 0-5 Scale
Departure Delay in Minutes	Minutes delayed in departure	Seat comfort	On 0-5 Scale
Arrival Delay in Minutes	Minutes delayed in Arrival	On-board service	On 0-5 Scale
Inflight wifi service	On 0-5 Scale	Leg room service	On 0-5 Scale
Ease of Online booking	On 0-5 Scale	Departure/Arrival time convenient	On 0-5 Scale
Baggage handling	On 0-5 Scale	Gate location	On 0-5 Scale
Cleanliness	On 0-5 Scale	Check-in service	On 0-5 Scale

Research questions

In our project, we aim to delve into the following research questions.

→ Can the passenger satisfaction level be predicted as a function of service features?

→ Are certain features more significant than the others in the success of an airline business?

→ How do we rank the significant features in order of highest priority for a successful airline business?

→ Which is the best classification model among Logistic Regression, Random Forest, KNN, SVM, Naive Bayes and Neural Networks in predicting the customer satisfaction based on accuracy and F1 score?

→ If the labels are ignored, how do the results of K-Means clustering match with the best classification model?

Methodology I

In order to solve the research questions we followed the following methodology:

1. Data checking & Preprocessing - The data was cleaned by removing features with missing values, standardized and split into 80:20 ratio for training and testing. Some features have high correlation with each other and so we have removed one of the features to improve the performance.

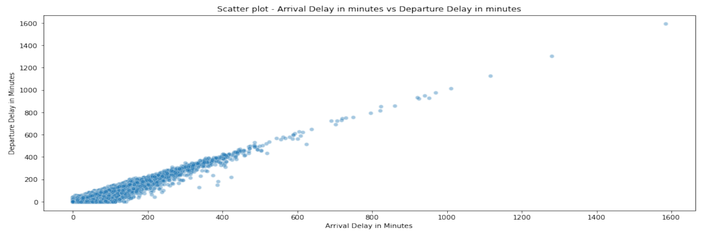

Figure 1. Scatter plot of features with high correlation

2. Feature selection - Initially we had 23 features. After forward selection are left with 18 features which are most critical for predicting passenger satisfaction.
3. Training the model - The data was trained using the following machine learning algorithms: Logistic Regression, Random Forest, K-Nearest Neighbour, Naive Bayes Support Vector Machine (SVM) and Neural Networks.
4. Hyperparameter tuning - Halving Grid Search was used to tune the hyperparameters and get the best model.
5. Feature Ranking - We rank the 18 service features from highest to lowest in predicting passenger satisfaction based on their Shapley values.

Methodology II

- Forward Selection: It begins with a null model and in each iteration, one feature is added at a time till we get the best performance [2].
- Halving Grid Search: We have provided a set of values and parameters to the halving grid search algorithm. After each iteration, the combinations of parameters that did not produce sufficient results are removed to produce the best combination of parameters [3].
- Shapley values: In Shapley values, each feature is given a significance value for a particular prediction by analyzing its effect on the output by removing that feature. The result ranks each feature by the order of importance [4] [5].

Model Performance & Results

Algorithm	Accuracy	Precision	Recall	f1 score
Logistic Regression	0.83	0.85	0.85	0.85
Random Forest Classifier	0.93	0.93	0.94	0.94
K-Nearest Neighbour	0.93	0.94	0.93	0.94
Support Vector Machine	0.94	0.95	0.95	0.95
K-means clustering	0.76	0.76	0.82	0.78
Naive Bayes Classifier	0.83	0.84	0.84	0.84
Neural Networks	0.95	0.94	0.96	0.95

Table 1. Performance metrics of machine learning models

Feature Ranking

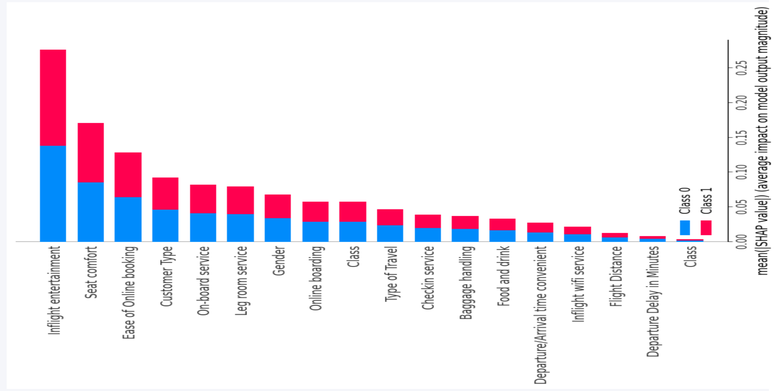

Figure 2. Ranking different features based on their Shapley values

In the above graph, we can see that the features are ranked in the order of significance when predicting the customer satisfaction. Services such as 'Inflight entertainment', 'Seat comfort', 'Ease of online booking', 'Customer type' and 'Leg room space' are the top five most important features to the passengers on flights. These services should be given a high priority to build a successful airline business.

Conclusion & Limitations

Among all the machine learning models, Neural Networks had the best accuracy of 95.01% and F-1 score of 95%. The airlines can use our best performing machine learning model-Neural Networks for future predictions to check the satisfaction level of customers. In order to improve the customer satisfaction and generate more revenues, the airlines can focus on improving the top-5 or top-10 highest ranked features, based on their budget. This research can be extremely useful for new companies who are planning to enter the airline business and have limited budget, so they can only focus on top ranked features.

In another approach, we classified the data into two clusters using unsupervised learning algorithm K-means clustering which gave an accuracy of 76%. This shows the importance of having data with output classes.

One of the limitations we faced during the project was the computational bottlenecks because of halving grid search and Shapley values processing times. The best model was also selected based on limited parameter changes and thus full capability of the model could not explored.

References

[1] R. Jou, S. H. Lam, D. Hensher, C.-C. Chen, and C.-W. Kuo, "The effect of service quality and price on international airline competition," *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, pp. 580–592, 07 2008.

[2] H. Singh, "Forward feature selection and its implementation," 2021.

[3] S. Babu, "Different hyperparameter optimization techniques,"

[4] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," pp. 4765–4774, 2017.

[5] A. Messalas, Y. Kanellopoulos, and C. Makris, "Model-agnostic interpretability with shapley values," in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pp. 1–7, 2019.