Report on Airline Passenger Satisfaction Prediction

Sofiane EL FARTASS

April 22, 2024

Abstract

This report outlines the development and evaluation of a machine learning model designed to predict airline passenger satisfaction based on flight data. The model leverages deep learning techniques to understand and predict levels of satisfaction from various flight-related features.

1 Introduction

The goal of this project is to apply advanced machine learning techniques to predict airline passenger satisfaction. The data was sourced from a Kaggle challenge and includes various attributes related to passenger experiences.

2 Dataset Description

The dataset consists of several features such as seat comfort, cabin service, food quality, and many others that are integral to a passenger's flight experience. Preprocessing steps involved handling missing values and encoding categorical variables to prepare the data for training.

3 Methodology

3.1 Data Preprocessing

Data was first analyzed for missing values and distribution of features. Key preprocessing steps included:

- Encoding categorical variables.

- Normalizing continuous variables.
- Handling missing data by imputation.

3.2 Model Architecture

The predictive model was built using a deep neural network with the following configuration:

- Input layer to accept features.
- Several dense layers with ReLU activation for non-linearity.
- Dropout layers to prevent overfitting.
- A sigmoid output layer for binary classification.

The model was compiled with the Adam optimizer, using binary cross-entropy as the loss function.

4 Results

The model's performance was assessed based on accuracy, precision, recall, and F1-score across the test dataset:

- Accuracy: 96.02%
- Precision: Class 0 - 0.95, Class 1 - 0.97
- Recall: Class 0 - 0.98, Class 1 - 0.94
- F1-Score: Class 0 - 0.96, Class 1 - 0.95

The confusion matrix and ROC curves were also generated to further evaluate model performance.

5 Conclusion

The deep learning model demonstrated high efficacy in predicting passenger satisfaction with an overall accuracy of 96%. The analysis indicates that the model is effective in distinguishing between satisfied and dissatisfied passengers, providing a robust tool for improving airline services.