Mid-semester Project: Customer Satisfaction Prediction for Invistico Airlines

Overview

In this assignment, you will work with a dataset provided by Invistico Airlines (anonymized) – Airline_data.csv - that contains information about customer feedback on various aspects of their flights. Your task is to predict whether a future customer would be satisfied with the airline's service based on the parameters in the dataset. Additionally, you will analyze which service aspects have the most significant impact on customer satisfaction, helping the airline focus on areas that can improve overall satisfaction.

Instructions

1. **Dataset Overview**: The dataset contains detailed feedback from customers who have flown with Invistico Airlines. Each row in the dataset represents a customer and includes the following types of information:
    - **Customer demographics**: Age, gender, etc.
    - **Flight details**: Type of travel, class of travel, etc.
    - **Service feedback**: Ratings on different aspects of the service (e.g., cleanliness, baggage handling, leg room service, etc.).
    - **Satisfaction status**: The target label indicating whether the customer was satisfied or not.
2. **Objective**:
    - Your primary objective is to build a **classification model** to predict whether a customer will be satisfied with their experience based on the features provided in the dataset.
    - Additionally, based on analysis of the data, make suggestions to the company to improve customer satisfaction.

Tasks

1. **Data Preprocessing**:
    - **Load and explore the dataset**. Ensure that all necessary libraries are installed and data is imported correctly.
    - Handle any missing values in the dataset by appropriate imputation techniques.
    - **Encode categorical variables** (if any) using techniques like one-hot encoding or label encoding.

- **Normalize or standardize** numerical features to ensure that features are on a comparable scale.

2. **Model Building**:
   - Split the data into **training** and **testing sets**.
   - Build a **classification model** for the data.
   - Evaluate the performance of each model using appropriate metrics.
3. **Report and Presentation**:
   - Prepare a brief **report** that includes:
     - Data preprocessing steps and challenges.
     - Insights from Exploratory Data Analysis.
     - Model Evaluation and Results.
     - Recommendations for the airline.
   - Ensure that your code is well-documented and clear.

Deliverables

1. A **Jupyter notebook (.ipynb)** with all the code used to complete the assignment.
2. A **report** (PDF or Word) summarizing your analysis, model results, and recommendations.

Submission Guidelines:

- The **notebook/script** should be well-commented and include all the steps mentioned above.
- The **report** should be concise and explain your approach, results, and key insights for the airline.
- Submit your files through the course submission portal by **1st November 2024**.

Grading Criteria:

Note:

- The purpose of this assignment is to assess your creativity and analytical skills. The specific details of the implementation are at your discretion; however,

ensure that your approach is logical and that you are prepared to justify your decisions

- The use of AI is strictly to generate code is strictly prohibited!

- **Data Preprocessing**: 30%
- **Model Building and Evaluation**: 60%
- **Report Clarity and Code Quality**: 10%

Good luck, and remember to think about how your model can provide actionable insights for Invistico Airlines to improve customer satisfaction!