Airline Passenger Satisfaction Prediction

Objective:

Build a machine learning system to predict airline passenger satisfaction based on service ratings, travel details, and demographics.

Dataset Overview:

Includes features like:

- **Demographics**: Gender, Age, Customer Type
- **Travel Info**: Class, Flight Distance, Type of Travel
- **Service Ratings**: Inflight WiFi, Food, Boarding, Check-in, Cleanliness
- **Time Delays**: Departure & Arrival Delay

Target: **Satisfaction** (Satisfied, Neutral or Dissatisfied)

Requirements:

1. **Data Preprocessing**
    - Handle missing values, encode categories, and normalize features
    - **Bonus** for using **advanced** or **non-trivial techniques** not covered in labs
        - You must understand the technique
2. **Exploratory Data Analysis (EDA)**
    - Visualize data distribution and correlation
    - Identify key satisfaction factors

3. **Model Development**

    - Train and compare **at least 4 ML algorithms** (e.g., Logistic Regression, Random Forest, SVM, KNN, Gradient Boosting, Neural Networks)
        - Using [**train.csv]**
        - **Bonus** for using **advanced** or **non-trivial techniques** not covered in labs (You must understand the model)
    - **Evaluate** Using [**test.csv]**
        - **Use proper metrics:** accuracy, precision, recall, F1-score
        - Use a confusion matrix

4. **GUI Application (Bonus)**

    - Desktop app using **Tkinter or PyQt**
    - Allow user input to predict satisfaction
    - Display prediction and confidence
    - Bonus: Charts or model comparison inside GUI

5. **Reporting (PDF) (Soft Copy)**

    - Clearly explain all preprocessing steps (e.g., missing values, encoding, scaling, feature selection).
    - Data Analysis & Visualization
        - Present and interpret relevant plots.
        - Explain the relationships between key features and satisfaction.
    - List all models. Include hyperparameters and tuning approach.
    - Describe any additional techniques or optimizations applied (e.g., feature engineering, ensemble methods).
    - Evaluation
        - Compare models using metrics such as accuracy, precision, recall, F1-score, and confusion matrix.
        - Discuss the strengths and weaknesses of each model.
    - Summarize findings, model performance, and insights gained.

Mentor

T.A. Mohamed Mosa

Email: mohamed_mosa@cis.asu.edu.eg