# Airline Passenger Satisfaction Analysis Using Machine Learning and Data Analytics

**Abstract**

In today's competitive aviation industry, understanding and improving passenger satisfaction is more important than ever. Airlines constantly strive to enhance the travel experience in order to build brand loyalty, optimize operations, and stay ahead in the market. The "Airline Passenger Satisfaction Analysis" project aims to uncover the factors that influence passenger satisfaction by applying modern data analytics and machine learning techniques. Using real-world passenger feedback data, the project provides insights that can help airlines improve services, increase customer retention, and make informed business decisions.

**Problem Statement and Overview**

Airlines receive a large volume of passenger feedback through surveys and digital channels. While this data is rich and valuable, it can be difficult to analyze due to its complexity and size. The main challenge is identifying which service elements—such as seat comfort, in-flight entertainment, or check-in experience—most significantly affect passenger satisfaction.

This project addresses that problem by analysing a structured dataset containing various attributes like travel type, class, flight distance, and satisfaction scores. The goal is to classify passengers as either satisfied or neutral/dissatisfied and understand what influences these outcomes the most. By doing so, airlines can focus their efforts on the areas that truly matter to customers.

**Tools and Technologies Used**

To analyze the dataset and build accurate predictive models, the following tools and technologies were used:

- Python: The core programming language used for data manipulation and model building.
- Pandas and NumPy: For efficient data processing and numerical operations.
- Matplotlib and Seaborn: For creating insightful visualizations and understanding data trends.

- Scikit-learn: A machine learning library used to implement and evaluate classification models.
- Jupyter Notebook: Used for developing, documenting, and presenting the code and results interactively.

**Submodules of the Project**

1. Data Preprocessing
    - Cleaned the dataset to remove inconsistencies
    - Handled missing values and encoded categorical features
    - Normalized data for consistent scaling
2. Exploratory Data Analysis (EDA)
    - Visualized distributions of variables
    - Analyzed correlations between features and satisfaction scores
    - Identified outliers and patterns using graphs
3. Model Building
    - Trained classification models including Logistic Regression, Decision Tree, and Random Forest
    - Evaluated performance using metrics like Accuracy, Precision, Recall, F1-score
4. Feature Importance Analysis
    - Identified which features had the most impact on predicting satisfaction (e.g., in-flight Wi-Fi, cleanliness, and seating comfort)
5. Result Interpretation and Visualization
    - Presented model predictions through summary tables and visual charts
    - Translated technical results into actionable business insights.

**Design and Flow of the Project**

The project follows a structured data science workflow:

1. Data Collection and Understanding
2. Data Cleaning and Transformation

3. Exploratory Analysis
4. Model Training and Testing
5. Feature Importance Evaluation
6. Final Insights and Interpretation

This linear and modular approach ensures that each stage of the analysis builds upon the previous one, providing clarity, consistency, and scalability for future enhancements or similar problems.

**Conclusion / Expected Output**

The expected result of this project is a reliable machine learning model that can predict whether a passenger is satisfied based on their travel experience and preferences. More importantly, the project identifies the most significant factors that influence satisfaction, such as:

- Quality of inflight amenities
- Efficiency of check-in and boarding
- Cleanliness and seat comfort

These findings provide airlines with a roadmap for enhancing customer service. By targeting these key areas, they can personalize services, streamline operations, and improve overall satisfaction scores.

In conclusion, this project bridges the gap between raw passenger feedback and strategic service improvements, showcasing how data analytics and machine learning can transform customer experience in the aviation sector.