Identifying Airline Passenger Satisfaction Using Machine Learning

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

Activity 1: Define Problem Statement

Problem Statement: Using machine learning, this project aims to predict airline passenger satisfaction based on various factors. For instance, by analyzing data on flight punctuality, in-flight services quality, and customer service interactions, the model will classify passengers as satisfied or dissatisfied. This classification will help airlines proactively address service deficiencies and enhance customer experience.

Airline Problem Statement Report: Click here

Activity 2: Project Proposal (Proposed Solution)

Problem Statement: Propose a machine learning model trained on data encompassing flight punctuality, service quality, and customer feedback to predict passenger satisfaction levels. Use algorithms like logistic regression or random forests for classification, enabling airlines to prioritize improvements that enhance overall customer experience and loyalty.

Airline Project Proposal Report: Click here

Activity 3: Initial Project Planning

Problem Statement: Define objectives and metrics for passenger satisfaction. Collect and preprocess relevant data on flight details, services, and feedback. Select and train machine learning models (e.g., logistic regression, random forests).Evaluate model performance using cross-validation. Deploy and monitor the model for real-time satisfaction prediction and continuous improvement.

Airline Project Planning Report: Click here

Milestone 2: Data Collection and Preprocessing Phase

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

Problem Statement: Gather flight details, customer service logs, and passenger surveys. Ensure data includes variables like flight punctuality, service ratings, and passenger demographics. Verify data completeness and accuracy. Address missing values and inconsistencies. Ensure privacy and security compliance.

Airline Data Collection Report: Click here

Activity 2: Data Quality Report

Problem Statement:

Completeness: All essential fields (flight details, service ratings) are present.
Accuracy: Data entries match actual flight and service records.
Consistency: Uniform format across datasets.
Validity: Data conforms to predefined formats and ranges.
Integrity: No duplicate or erroneous entries.

Airline Data Quality Report: Click here

Activity 3: Data Exploration and Preprocessing

Problem Statement: In exploring airline passenger satisfaction with machine learning, begin with data cleaning, feature selection, and normalization. Address missing values and outliers, ensuring data readiness for modeling to derive insights into factors influencing satisfaction.

Airline Data exploration and preprocessing report: Click here

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for loan approval. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, Logistic Regression), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

Activity 1: Feature Selection Report

Problem Statement: The Feature Selection Report outlines the rationale behind choosing specific features (e.g., Gender, Travel Type, Cleanliness) for the loan approval model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to discern credible loan applicants.

Airline Feature selection report: Click here

Activity 2: Model Selection Report

Problem Statement: The Model Selection Report details the rationale behind choosing Random Forest, Decision Tree, KNN, and Logistic regression models for passenger satisfaction prediction. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Airline Model selection report: Click here

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

Problem Statement: The Initial Model Training Code employs selected algorithms on the Airline passenger satisfaction dataset, setting the foundation for predictive modelling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in predicting airline passenger outcomes.

Airline Model development phase report: Click here

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

The Random Forest model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Random Forest model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Random Forest as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal loan approval predictions.

Airline model optimisation and tunning phase: [Click here](#)

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, For the documentation, Kindly refer to the link. [Click here](#)

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.