

Project Design - Phase 1

Proposed solution

The proposed solution aims to predict airline passenger satisfaction using machine learning techniques. Data will be collected from various sources and preprocessed to ensure quality. Models like logistic regression, decision trees, random forests, and neural networks will be built and evaluated. The best performing model will be used to predict passenger satisfaction. The project's output will be a web-based application providing real-time insights for airlines to improve their services and enhance passenger experience.

Solution architecture

The solution architecture for the project can be divided into several components, as outlined below:

1. Data Collection:

- Online Surveys: Design and implement a survey platform to collect feedback from airline passengers, capturing their satisfaction levels and related features.
- Social Media Monitoring: Utilize APIs or scraping techniques to gather publicly available data from social media platforms, extracting relevant posts or comments related to passenger satisfaction.
- Other Data Sources: Explore additional publicly available data sources that provide insights into passenger experiences, such as customer review websites or airline-specific databases.

2. Data Preprocessing and Cleaning:

- Remove duplicates, irrelevant entries, and noisy data.
- Handle missing values through techniques like imputation or deletion.
- Standardize and normalize the data to ensure consistency across features.

3. Feature Engineering:

- Analyze the collected data and identify relevant features that contribute to passenger satisfaction.
- Transform and engineer features if needed, such as converting categorical variables into numerical representations or creating new derived features.

4. Model Development and Evaluation:

- Split the preprocessed data into training and testing sets.
- Explore various machine learning models, such as logistic regression, decision trees, random forests, and neural networks.
- Train and evaluate each model using appropriate evaluation metrics (e.g., accuracy, precision, recall, F1 score) to identify the best performing model.

5. Model Deployment:

- Implement the chosen model into a web-based application using frameworks like Flask or Django.
- Develop a user-friendly interface for airlines to input relevant data for passenger satisfaction prediction.
- Incorporate real-time data processing capabilities to allow for continuous updates and predictions.

6. Application Output and Insights:

- Provide airlines with real-time predictions of passenger satisfaction based on the input data.
- Display the results in a visually appealing and interpretable manner, such as through charts, graphs, or summary statistics.
- Enable airlines to identify specific areas of improvement based on the predicted satisfaction levels and feature importance.

7. Continuous Improvement:

- Monitor the performance of the deployed model and collect feedback from airline users.
- Regularly update and retrain the model using new data to ensure its accuracy and relevance.
- Consider incorporating additional features or data sources, such as sentiment analysis of social media data, to enhance the prediction capabilities of the system.