Predicting Airline Passenger Satisfaction

By Piyush Ramteke

Presentation Agenda

Here's what we'll cover in today's deep dive into predicting airline passenger satisfaction:

01

Project Overview

Understand the scope and objectives of our analysis.

02

Problem Statement

Define the core challenge we aim to address with data.

03

Dataset Description & Preprocessing

Examine the data and steps taken to prepare it for modeling.

04

Exploratory Data Analysis (EDA)

Discover patterns and insights hidden within the dataset.

05

Model Building

Details on the predictive models developed.

06

Model Evaluation

Assessment of model performance and accuracy.

07

Results & Key Insights

Summary of findings and actionable conclusions.

08

Limitations & Future Work

Acknowledging constraints and outlining next steps.

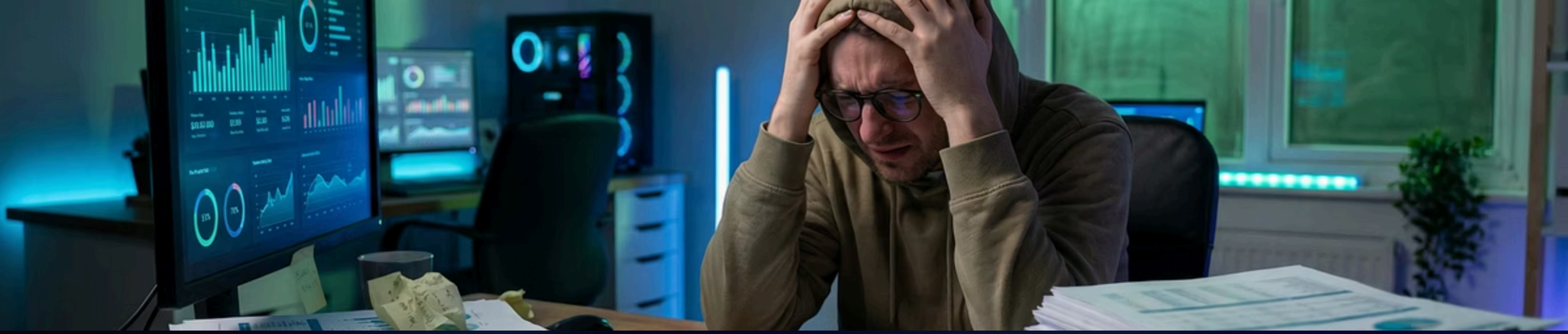
09

Conclusion

Recap of our journey and final thoughts.

Project Overview

This project uses machine learning to predict airline passenger satisfaction, identifying key factors influencing customer experience. The model classifies passengers as satisfied or dissatisfied, helping airlines make data-driven improvements.



Problem Statement

Inefficient Manual Analysis

Manually analyzing vast amounts of passenger feedback is inefficient and often inaccurate.

Challenging Factor Identification

Identifying key satisfaction factors is difficult due to data complexity and volume.

Need for Automation

An automated ML model is needed to accurately predict satisfaction and improve service quality.

Dataset Description & Preprocessing

Dataset Description

Sourced from Kaggle, the dataset contains ~129,880 records with passenger demographics, travel type, service ratings, and flight delays.

Data Preprocessing

Cleaned data by handling missing values, removing inconsistencies, encoding categorical features, and normalizing numerical variables for model training.

Exploratory Data Analysis (EDA)

EDA revealed feature distributions, correlations, and patterns affecting satisfaction. Visual analyses highlighted service ratings, travel type, and delays as key influencers, guiding model selection.

Service Ratings

Travel Type

Flight Delays

Model Building

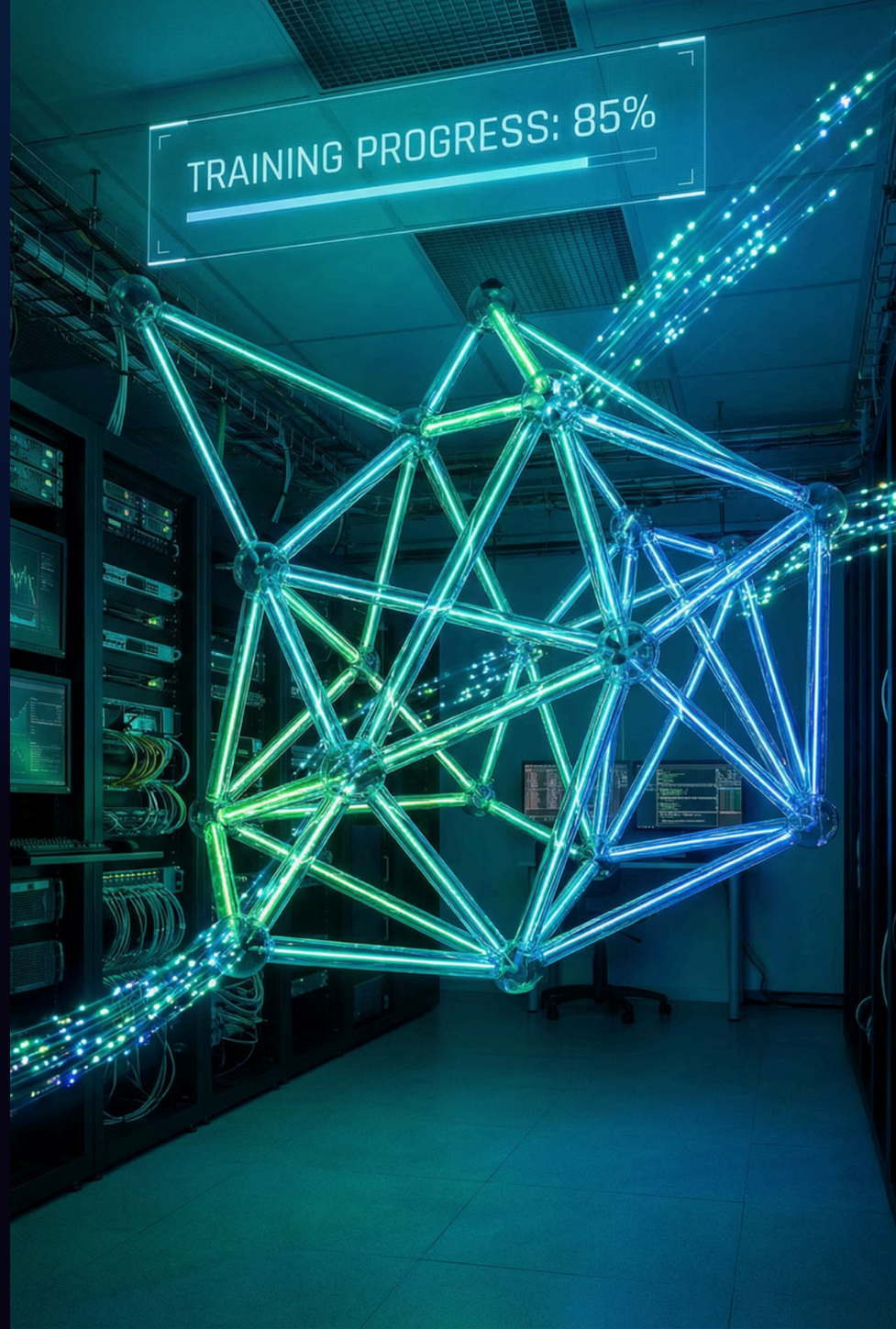
The dataset was split into training and testing sets. Logistic Regression, Random Forest, and XGBoost were trained. Parameters were tuned, and feature importance analyzed to understand satisfaction drivers.

Data Split

Algorithm Training

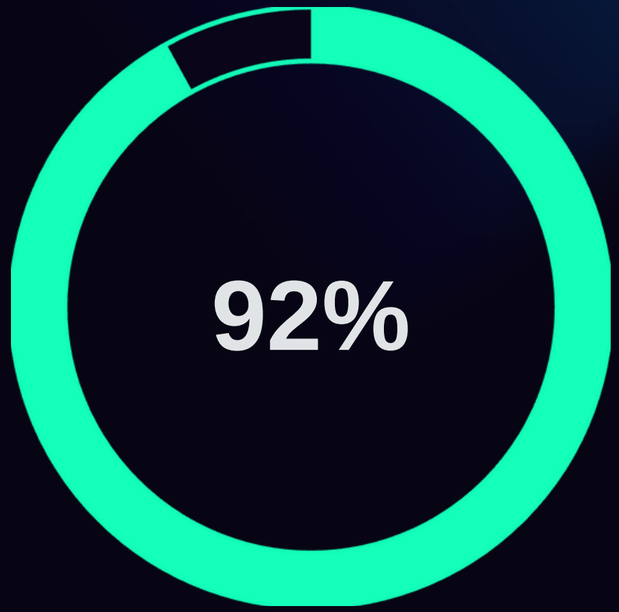
Parameter Tuning

Feature Analysis



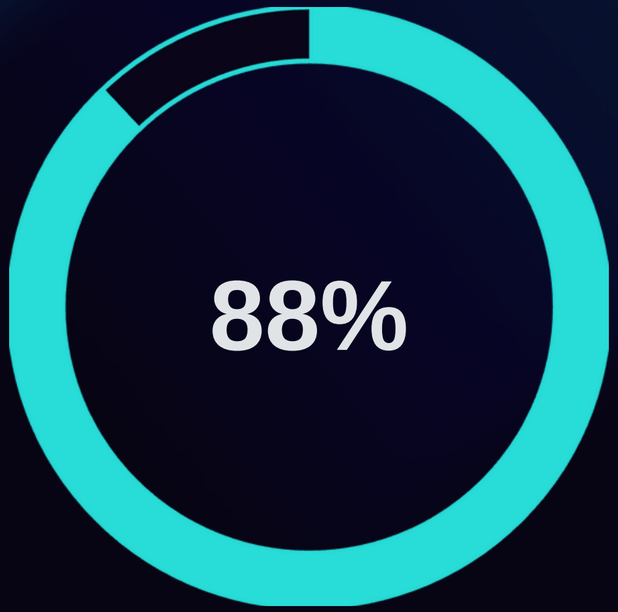
Model Evaluation

Model performance was assessed using accuracy, precision, recall, and F1-score. A confusion matrix analyzed prediction errors, leading to the selection of the best-performing model.



Accuracy

High accuracy in predicting satisfaction.



Precision

Reliable positive predictions.

Recall

Effective at identifying satisfied passengers.

Results & Key Insights

The best model achieved strong accuracy. Key insights highlight seat comfort, onboard service, check-in experience, and flight delays as significant factors influencing passenger satisfaction.

Seat Comfort

Onboard Service

Check-in Experience

Flight Delays

Limitations & Future Work

Limitations

- Dataset from a single airline limits general applicability.
- Subjective passenger ratings may introduce bias.
- Model performance depends on data quality and completeness.

Future Work

- Integrate real-time prediction and explore deep learning for enhanced accuracy.
- Expand with larger, diverse datasets to improve generalization and reduce bias.

Conclusion

This project successfully built an ML model to predict airline passenger satisfaction with strong accuracy. The analysis provides actionable insights for airlines to improve service quality and enhance customer experience through data-driven decisions.