RAJALAKSHMI
ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

25 years
OF ACADEMIC
EXCELLENCE

Department of Computer Science and Engineering

Flight Fare Prediction Using Machine Learning Techniques

Mrs. M. Divya M.E.
Asst.professor ,
dept of cse,
Rajalakshmi
Engineering College

220701503
Deebak N

Problem Statement and Motivation

Problem Statement:

Accurately predicting flight fares is a complex challenge due to multiple influencing factors such as brand, time, aircraft specs, and demand variability

Motivation:

With rising airline usage, customers and airlines alike can benefit from an accurate, intelligent pricing system for better decision-making and competitive advantage.

Existing System

- Traditional systems rely on linear regression or hedonic pricing models.
- Limited ability to handle non-linearity and high-cardinality categorical data.
- Prone to overfitting and poor generalization

Objectives

- Predict flight fare resale prices using machine learning.
- Improve accuracy and generalization using ensemble models.
- Identify key features influencing price.
- Deploy a scalable, real-world model.

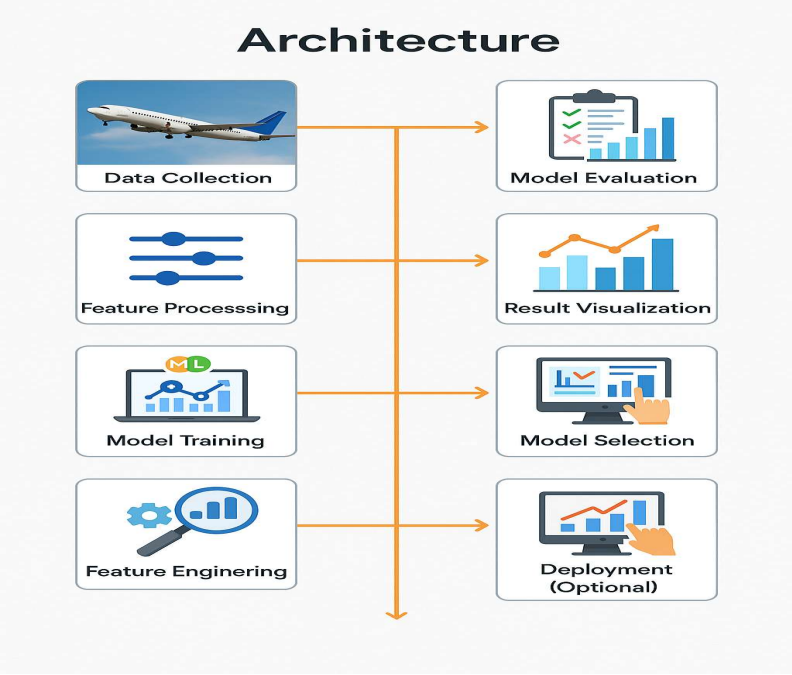
Abstract

This project uses historical flight fare data to build a machine learning-based predictive model. A variety of algorithms are employed—including ensemble methods like XGBoost—to capture complex relationships in pricing. Results show that these models outperform traditional methods in terms of accuracy and robustness.

Proposed System

- Use a cleaned and preprocessed dataset of historical flight fares.
- Apply multiple ML algorithms including regression and ensemble methods.
- Evaluate model performance with metrics like MAE, MSE, and R^2 Score.
- Select and visualize the best-performing model

System Architecture



List of Modules

- Data Collection and Preprocessing
- Feature Engineering
- Model Training (ML Algorithms)
- Performance Evaluation
- Result Visualization
- Deployment Readiness (Optional)

Functional Description for each modules with DFD and Activity Diagram

- **Data Collection:** Fetch historical flight fare data.
- **Preprocessing:** Handle missing values, encode categoricals, scale numericals
.
- **Modeling:** Train models (e.g., Linear, RF, XGBoost).
- **Evaluation:** MAE, MSE, R^2 Score.
- Include a basic **DFD and activity diagram** showing flow

Implementation & Results of Module

Conclusion:

XGBoost outperformed other models, effectively predicting flight fares by learning complex patterns. Feature engineering and tuning enhanced model accuracy.

Future Work:

- Add real-time API for dynamic fare prediction.
- Use deep learning for image/text-based ticket data.
- Apply explainable AI for transparency.
- Explore time series modeling

Conclusion & Future Work

Conclusion:

XGBoost outperformed other models, effectively predicting flight fares by learning complex patterns. Feature engineering and tuning enhanced model accuracy.

Future Work:

- Add real-time API for dynamic fare prediction.
- Use deep learning for image/text-based ticket data.
- Apply explainable AI for transparency.
- Explore time series modeling

References

Saini, R., & Tiwari, A. (2021)
Flight Fare Prediction System Using Random Forest and XGBoost.
Journal of Emerging Technologies and Innovative Research (JETIR), 8(4), 101–107.

Rajesh, A., & Arun, V. (2020)
Airfare Price Prediction Using Supervised Learning Techniques.
International Journal of Scientific & Technology Research, 9(1), 282–285.

Kumar, V., & Goel, D. (2019)
Flight Ticket Price Prediction Using Machine Learning.
International Journal of Computer Sciences and Engineering, 7(6), 622–628.

Thank You