SUMMER TRAINING PROJECT REPORT (Term June–July 2025)

TITLE : AUR FARE PREDICTION CHATBOT

Submitted by

Name: Saket Rajak

Registra on Number: 12306986

Name: Deepak Singh

Registra on Number: 12316331

Course Code: PETV79

Under the Guidance of

Mahipal Singh Papola

School of Computer Science and Engineering

CONTENTS OF THE REPORT

Tools & technologies used

Areas covered during training

Daily/weekly work summary

Chapter 3: Project Details

Title of the project

Problem defini on

Scope and objec ves

System Requirements

Architecture Diagram

Data flow / UML Diagrams

Chapter 4: Implementa on

Tools used

Methodology

Modules / Screenshots

Code snippets

Chapter 5: Results and Discussion Output / Report

Challenges faced

Learnings

Chapter 6: Conclusion

Summary

BONAFIDE CERTIFICATE

Cer fied that this project report “AUR FARE PREDICTION CHATBOT” is the bona fide work of “Saket Rajak, Deepak Singh” who carried out the project work under my supervision.

SIGNATURE

<<Name of the Supervisor>>

Saket Rajak

Deepak Singh

SIGNATURE

<<Signature of the Head of the Department>>

SIGNATURE

<<Name>>

HEAD OF THE DEPARTMENT

SIGNATURE

<<Signature of the Supervisor>>

1. INTRODUCTION

1.1 OVERVIEW OF TRAINING DOMAIN:

The training focused on the integra on of Machine Learning and Web deployment to build a datadriven chatbot that can predict airfare. The session included working with ML libraries like scikitlearn and XGBoost, designing visualisa ons using Seaborn, deploying interac ve dashboards using Streamlit, and understanding hyperparameter op miza on through tools like Optuna.

Par cipants applied theore cal concepts in a real-world avia on pricing context, with exposure to:

Exploratory Data Analysis (EDA)

Feature engineering and selec on

Regression modelling techniques

Hyperparameter tuning

REST API integra on

Frontend design using Streamlit

1.2 OBJECTIVE OF THE PROJECT:

The primary objec ve of this project is to build a real- me airfare predic on chatbot that allows users to query flight prices using natural language and receive accurate fare forecasts based on historical data. Key objec ves include:

Build ML models to predict airfare with high accuracy

Integrate the model into a chatbot interface for user queries

Deploy the chatbot using Streamlit to make it user-accessible

Include visualiza on and insights to help users make informed decisions

2. TRAINING OVERVIEW

2.1 TOOLS AND TECHNOLOGIES USED:

The tools and technologies used for this project include:

Python: Core programming language

Pandas, NumPy: Data manipula on

Seaborn, Matplotlib: Data visualiza on

Scikit-learn, XGBoost: ML model building

Optuna: Hyperparameter tuning

Streamlit: UI deployment

Pickle: Model saving/loading

Jupyter Notebook/VS Code: IDEs for code development

2.2 AREAS COVERED DURING TRAINING:

The training was structured around the following core areas:

Introduc on to Regression Models

Data preprocessing and feature engineering

Model training with XGBoost and RandomForest

Hyperparameter tuning with Optuna

Streamlit-based chatbot development

Model deployment and web UI

CI/CD and dockeriza on basics

2.3 DAILY/WEEKLY WORK SUMMARY:

Week

Summary

1

Understanding airline pricing and exploring datasets

2

EDA, feature correla on, ini al model training

3

Building XGBoost and Random Forest models

4

Hyperparameter tuning using Optuna

5

Chatbot interface development using Streamlit

6

Tes ng, documenta on, deployment, final report

3. PROJECT DETAILS

3.1 TITLE OF THE PROJECT:

“Aur Fare Predic on Chatbot”

3.2 PROJECT DETAILS:

Flight prices fluctuate constantly due to demand, season, availability, and routes. Passengers o en struggle to find the best me to book a cket. Our solu on leverages Machine Learning to build a chatbot that predicts airfare based on:

Departure & arrival airports

Booking window

Airline

Number of stops

A user-friendly interface is developed for users to type queries like “What’s the cheapest fare from Mumbai to Delhi next Saturday?” and receive a predic on.

3.3 SCOPE AND OBJECTIVES:

Scope:

Provide price predic ons for Indian domes c routes

Serve as a travel assistant through chatbot interface

Aid budget travelers in iden fying op mal booking mes

Objec ves:

Predict airfare with Mean Absolute Percentage Error (MAPE) < 15%

Build chatbot using Streamlit with instant response

Integrate visualiza ons and airline-wise insights

3.4 SYSTEM REQUIREMENTS:

So ware Requirements

Python 3.10+

Streamlit

XGBoost, Optuna

Scikit-learn

Pandas, Seaborn

Hardware Requirements

8GB RAM, 4-core CPU

Internet connec on

Cloud deployment (Railway / Render / local)

3.5 ARCHITECTURE DESIGN:

User → Chatbot UI → Intent Detec on & Slot Filling → Feature Builder

→ ML Model (XGBoost) → Fare Predic on → Response

3.6 DATA FLOW/UML DIAGRAM:

User types flight query

Query parsed using keyword matching

Feature vector created

Model predicts fare

Response sent back to chat window

Historical fare trend visualized

4. IMPLEMENTATION

4.1 TOOLS USED:

Tool/Library

Purpose

Pandas

Data Handling

Seaborn

Visualiza on

XGBoost

Model Training

Optuna

Hyperparameter Tuning

Pickle

Model Saving

Streamlit

Web UI

VS Code

Code Editor

4.2 METHODOLOGY:

Data Collec on: Flight fare dataset from Kaggle

EDA: Histograms, heatmaps, and boxplots

Feature Engineering: Booking window, stops, airline encoding

Model Training: Random Forest and XGBoost

Hyperparameter Tuning: Optuna used to minimize MAPE

Deployment: Streamlit interface + Docker support

4.3 MODULES / SCREENSHOTS: Module 1:

Fare Predic on Model model = XGBRegressor(...)

model.fit(X_train, y_train)

Module 2: Chatbot UI (Streamlit)

st.text_input("Enter flight query:")

st.bu on("Predict Fare")

Module 3: Visualiza on sns.barplot(data=airline_df, x='Airline', y='AvgFare')

Module 4: Fare Forecast JSON Output

```
{
  "fare": 7150,
  "confidence": "±8%"
}
```

4.4 CODE SNIPPETS:

predic on = model.predict([[origin, des na on, date, stops, airline]]) st.success(f"Es mated Fare: ₹{predic on}")

5. RESULTS AND DISCUSSION

5.1 OUTPUT:

Predic on accuracy (MAPE): 12.9%

Streamlit UI predicts fare in under 1 second

Addi onal visualiza ons: average fares per airline, booking window impact Graphical UI supports responsive layout

5.2 CHALLENGES FACED:

Sparse Routes: Low data for rare routes

Realis c Output: Avoiding outlier fares

Intent Recogni on: Parsing user queries in natural language

Deployment Constraints: Streamlit container op miza on

5.3 LEARNINGS:

Experience in ML model lifecycle

UI design without HTML/CSS

API endpoint management in Streamlit

Understanding airline pricing pa erns

Collabora ve project development

6. CONCLUSION

The Aur Fare Predic on Chatbot is a successful blend of Machine Learning and web technologies to address real-world pricing challenges in air travel. With high accuracy, a simple interface, and ac onable insights, the system can help users plan their trips smartly.

The project covered the complete pipeline:

Data wrangling

Model training and tuning

Web app integra on

Visualisa on and deployment

This solu on lays the groundwork for future enhancements like dynamic fare alerts and mul lingual support, showing the role AI can play in the travel industry.