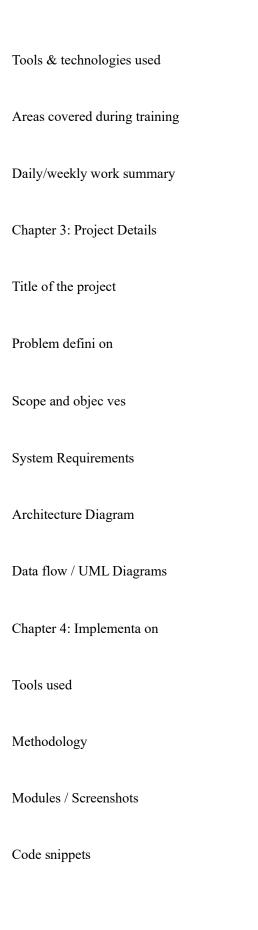
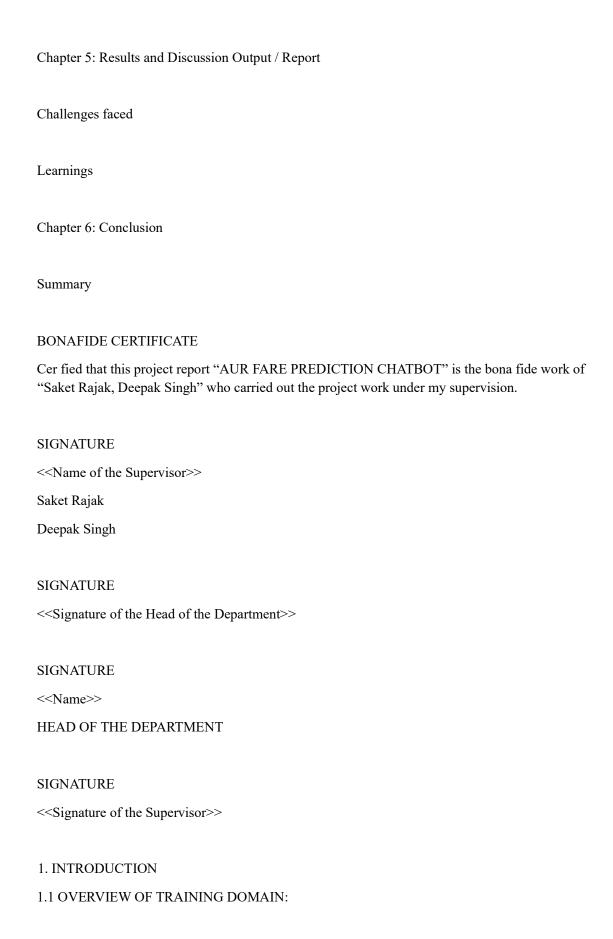
SUMMER TRAINING PROJECT REPORT (Term June–July 2025)

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 Cover page Cer ficate Table of Contents Chapter-wise Report Chapter 1: Introduc on Overview of training domain Objec ve of the project

Chapter 2: Training Overview





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:

TC1 / '	1 1	, 1	1 .	1		1 .	• ,	. 1	1
I ha too	เดอกส	tachno	100100	11000	tor t	hic	nrolact	1110	11/40.
The too	is anu	. icciiio	เบยเบล	uscu	IOI L	ans.	DIOICCL	11101	uuc.
							,		

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

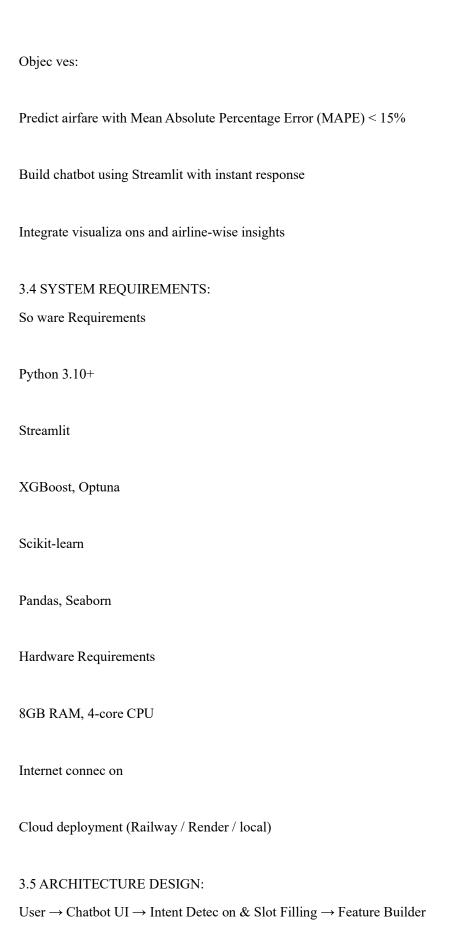
Hyperparameter tuning with Optuna

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

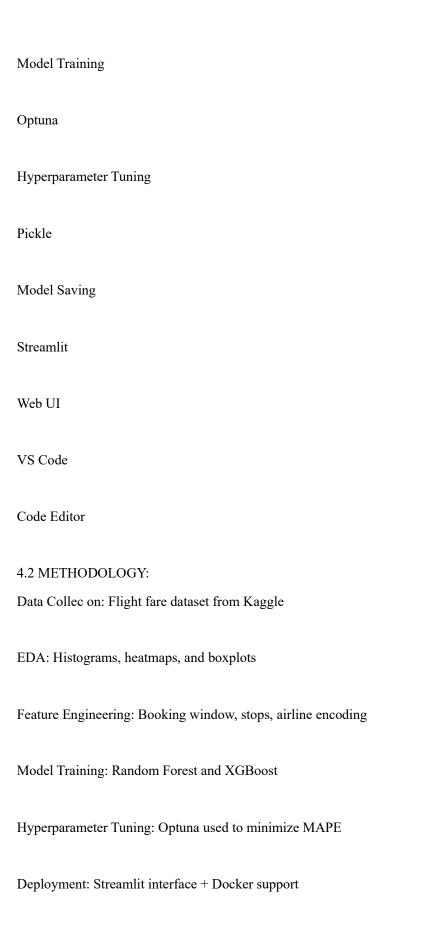
3. PROJECT DETAILS
3.1 TITLE OF THE PROJECT:
"Aur Fare Predic on Chatbot"
3.2 PROJECT DETAILS:
Flight prices fluctuate constantly due to demand, season, availability, and routes. Passengers o en struggle to find the best me to book a cket. Our solu on leverages Machine Learning to build a chatbot that 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

Tes ng, documenta on, deployment, final report



\rightarrow ML Model (XGBoost) \rightarrow Fare Predic on \rightarrow 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
Response sent back to chat window
Historical fare trend visualized
4. IMPLEMENTATION
4.1 TOOLS USED:
Tool/Library
D.
Purpose
Pandas
Data Handling
~ .
Seaborn
Visualiza on
XGBoost



```
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%
```

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

Streamlit UI predicts fare in under 1 second

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

TT 7 1		• .	
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.