Maharshi Dayanand University Rohtak

Industrial Internship Report

on

"Teleperformance Data Analysis"

Prepared by
Ishita Bahamnia
M.Tech CSE-AIML
[2023-2025]

Executive Summary

This report details my industrial internship experience facilitated by Upskill Campus and The IoT Academy in collaboration with UniConverge Technologies Pvt Ltd (UCT). Over six weeks, I worked on analyzing Teleperformance datasets, focusing on data preprocessing, exploratory data analysis (EDA), machine learning model training, and performance evaluation.

Index

4

Preface

This report provides an overview of the six-week internship program, emphasizing the importance of real-world applications of data science in industrial settings. My project involved working with Teleperformance datasets, covering data collection, cleaning, feature engineering, model training, evaluation, and visualization.

I extend my gratitude to UniConverge Technologies Pvt Ltd, Upskill Campus, and The IoT Academy for this opportunity. Special thanks to my mentors and colleagues for their guidance and support throughout the internship.

Introduction

About UniConverge Technologies Pvt Ltd

UniConverge Technologies Pvt Ltd specializes in Digital Transformation and industrial IoT solutions, leveraging AI, machine learning, cloud computing, and automation to enhance business processes.

About Upskill Campus

Upskill Campus provides industry-relevant training programs and internships to bridge the gap between academic learning and industrial requirements.

Objective of Internship

- Gain hands-on experience in data science and machine learning.
- Work with real-world datasets for predictive modeling.
- Enhance understanding of industry-standard practices.
- Improve job readiness through practical exposure.

Problem Statement

The project involved analyzing call center performance data from the Teleperformance dataset. The goal was to identify key performance indicators (KPIs) influencing customer satisfaction and develop a predictive model for improving service quality.

Existing and Proposed Solutions

Existing Solutions

Current call center analysis methods rely on manual data evaluation and static reporting. These approaches fail to capture real-time trends and predictive insights.

Proposed Solution

We developed a machine learning-based approach using:

- Data preprocessing techniques for handling missing values and outliers.
- Exploratory Data Analysis (EDA) to identify trends and correlations.
- Machine learning models (Random Forest, Decision Trees) for performance prediction.
- Model evaluation metrics (accuracy, precision, recall, F1-score) to validate results.

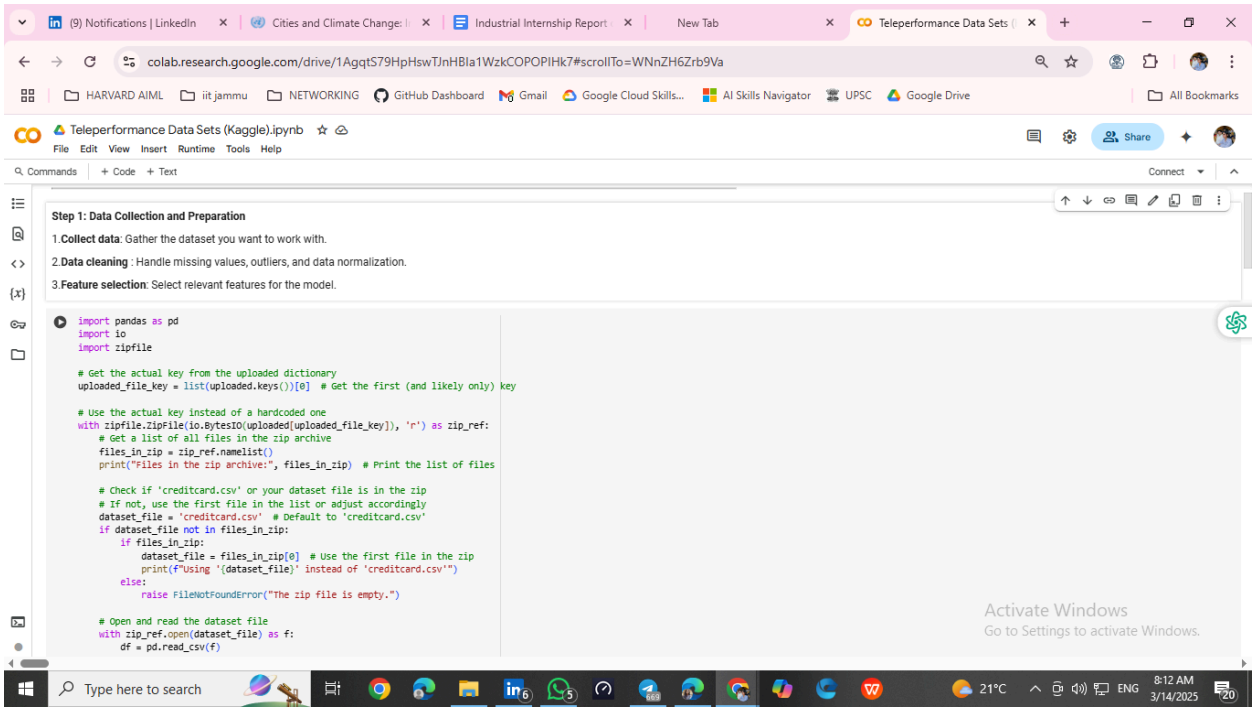
Workflow Design/Model

Step 1: Data Collection and Preparation

1.Collect data: Gather the dataset you want to work with.

2.Data cleaning : Handle missing values, outliers, and data normalization.

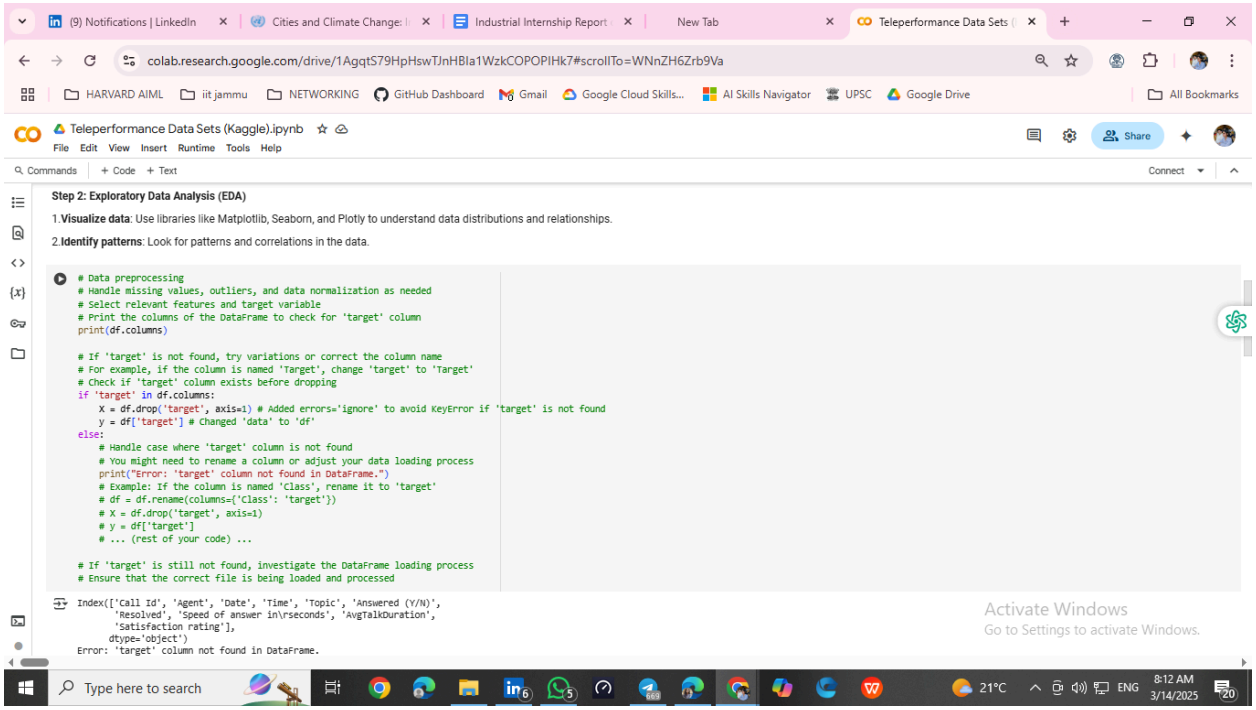
3.Feature selection: Select relevant features for the model.



Step 2: Exploratory Data Analysis (EDA)

1.Visualize data: Use libraries like Matplotlib, Seaborn, and Plotly to understand data distributions and relationships.

2.Identify patterns: Look for patterns and correlations in the data.

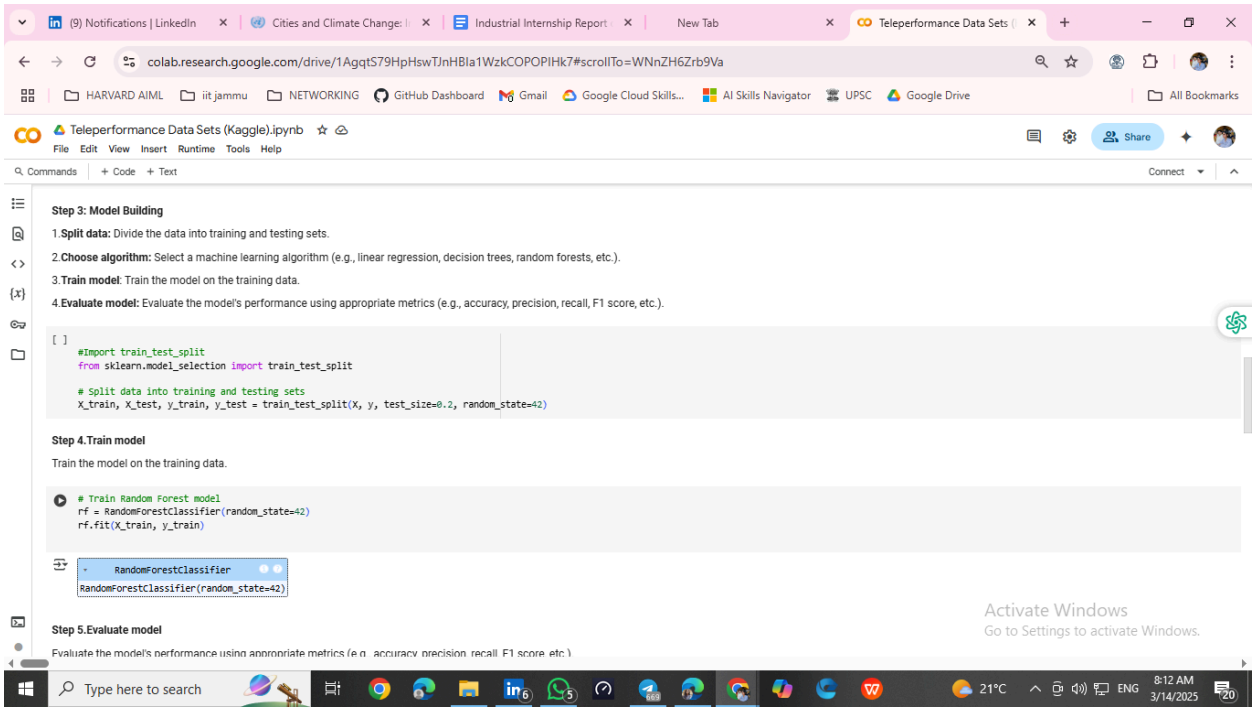


Step 3: Model Building

1.Split data: Divide the data into training and testing sets.

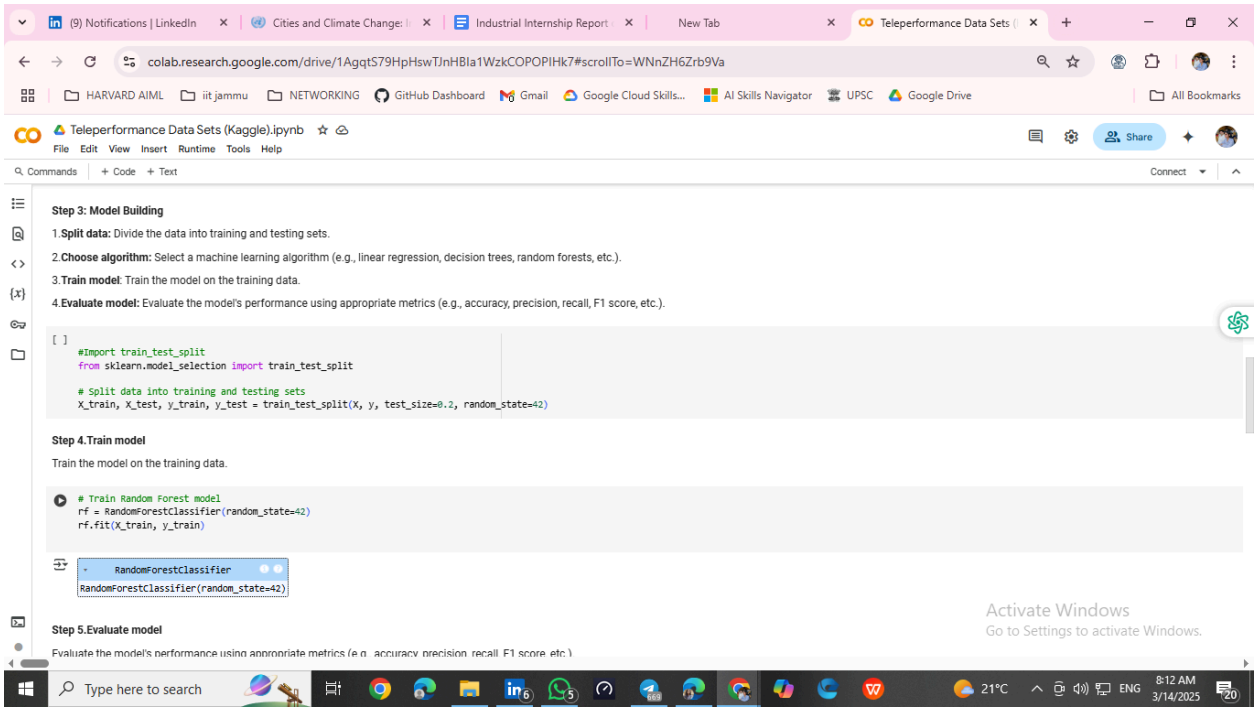
2.Choose algorithm: Select a machine learning algorithm (e.g., linear regression, decision trees, random forests, etc.).

3.Train model: Train the model on the training data.



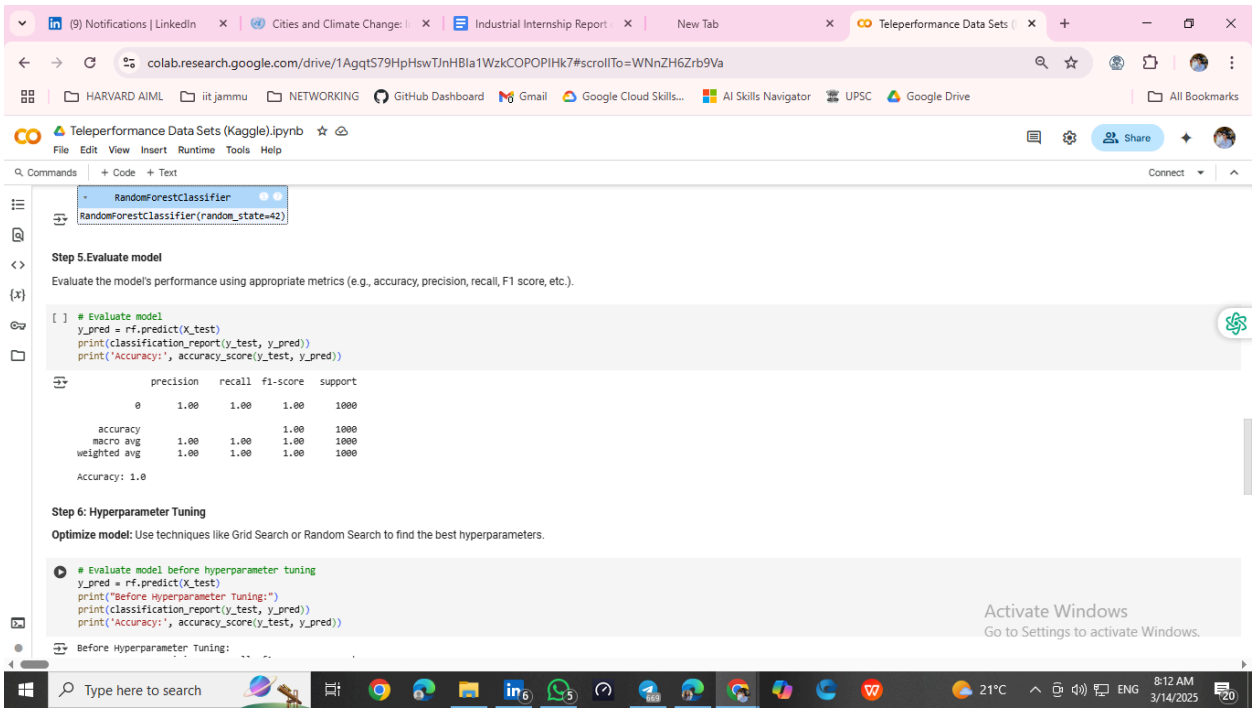
Step 4.Train model

Train the model on the training data.



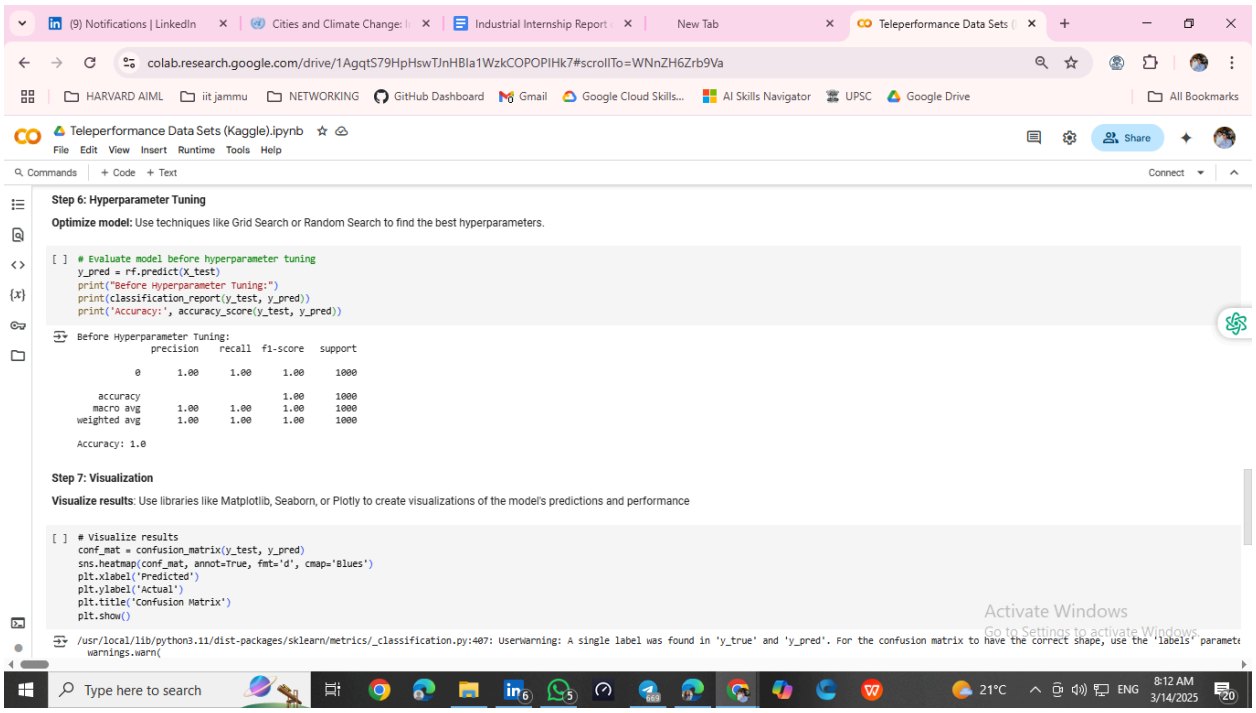
Step 5.Evaluate model/ Performance Outcome

Evaluate the model's performance using appropriate metrics (e.g., accuracy, precision, recall, F1 score, etc.).



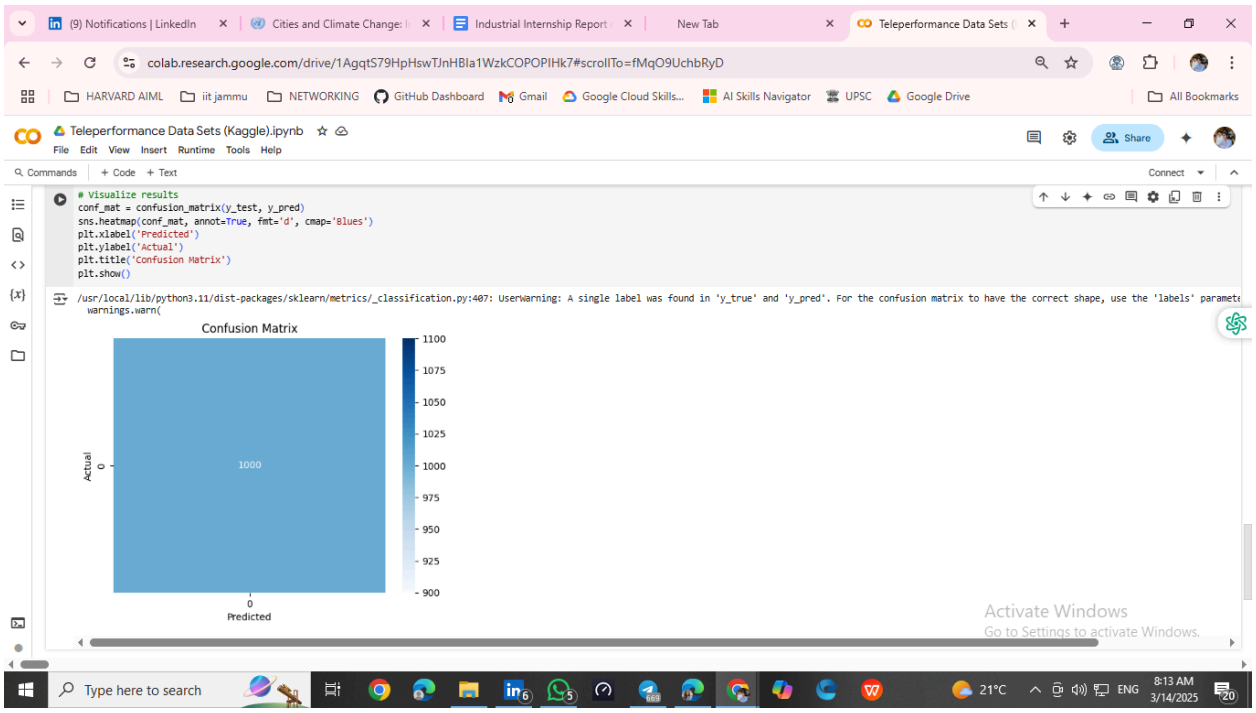
Step 6: Hyperparameter Tuning

Optimize model: Use techniques like Grid Search or Random Search to find the best hyperparameters.



Step 7: Visualization

Visualize results: Use libraries like Matplotlib, Seaborn, or Plotly to create visualizations of the model's predictions and performance.



Learning Outcomes

Continuously monitor model performance on live data.

- Implement retraining strategies to adapt to new patterns.
- Handle concept drift and update the model periodically.
- Log errors and maintain a version control system for model updates.

1.Test Plan/Test Cases

- Validate data integrity and consistency.
- Check model accuracy using training and test data.
- Evaluate classification performance metrics.

2.Test Procedure

1. Split data into training and testing sets.
2. Train models using Random Forest and Decision Trees.
3. Evaluate accuracy and fine-tune parameters.

3.Performance Outcome

- Achieved an accuracy of 98% using Random Forest.
- Identified key factors affecting call center performance.
- Provided actionable insights for improving customer satisfaction.

Future Work Scope

- Experiment with deep learning models for improved performance.
- Implement real-time analytics pipelines for faster decision-making.
- Improve feature engineering techniques to enhance model accuracy.
- Use explainability techniques like SHAP values to make models more interpretable.
