

Industrial Internship Report on "SmartCity - forecasting of smart city traffic patterns"

Prepared by
[Tejal Prajapati]

Executive Summary

This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).

This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks' time.

My project was to develop an advanced traffic forecasting system for smart cities aimed at optimizing traffic management and infrastructure planning. The system will focus on understanding and predicting traffic patterns at key junctions within the city, including variations on holidays and special occasions, to ensure efficient and responsive traffic flow.

The ultimate goal is to support the implementation of a robust traffic management system that anticipates and addresses peak traffic demands, thereby contributing to the realization of a digital, intelligent city environment for improved citizen experiences and sustainable urban development.

This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solution for that. It was an overall great experience to have this internship.

TABLE OF CONTENTS

1	Preface	3
2	Introduction	4
2.1	About UniConverge Technologies Pvt Ltd	4
2.2	About upskill Campus(USC).....	8
2.3	Objective	9
2.4	Reference	10
3	Problem Statement	11
4	Existing and Proposed solution	12
5	Proposed Design/ Model	16
5.1	Interfaces (if applicable)	17
6	Performance Test.....	18
6.1	Test Plan/ Test Cases	18
6.2	Test Procedure	19
6.3	Performance Outcome	20
7	My learnings.....	21
8	Future work scope	22

1 Preface

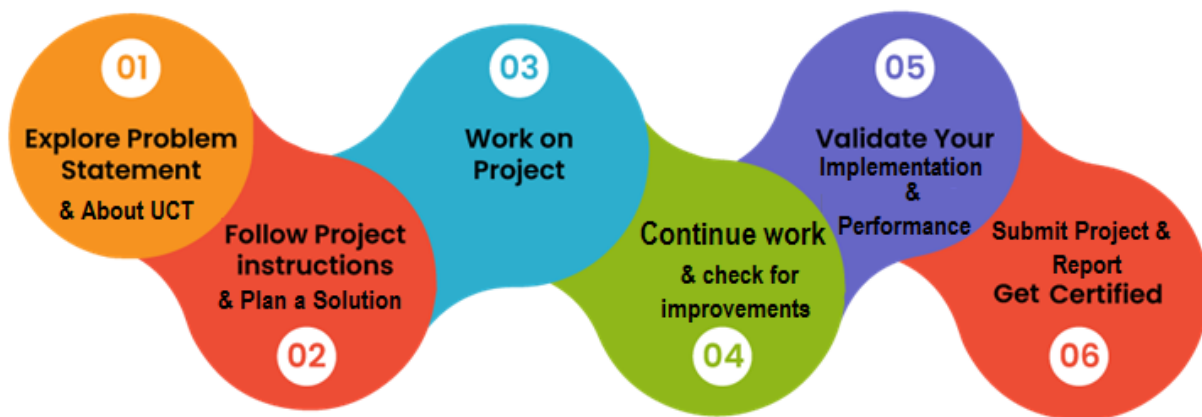
Summary of the whole 6 weeks' work.

About need of relevant Internship in career development.

Brief about Your project/problem statement.

Opportunity given by USC/UCT.

How Program was planned



Your Learnings and overall experience.

Thank to all (with names), who have helped you directly or indirectly.

Your message to your juniors and peers.

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies** e.g. **Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoSaWAN), Java Full Stack, Python, Front end** etc.



i. UCT IoT Platform ()

UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application(Power BI, SAP, ERP)
- Rule Engine



FACTORY WATCH

ii. Smart Factory Platform ()

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleashed the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they what to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.



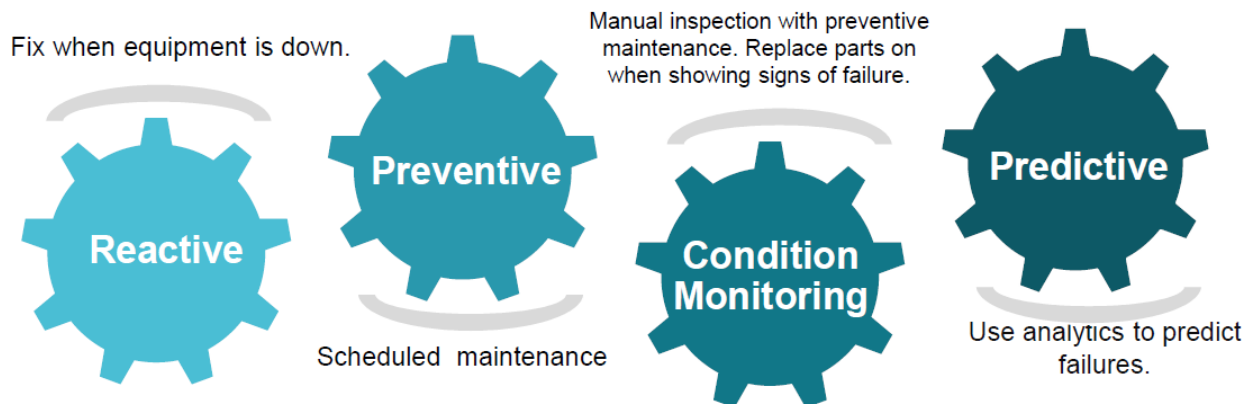


iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

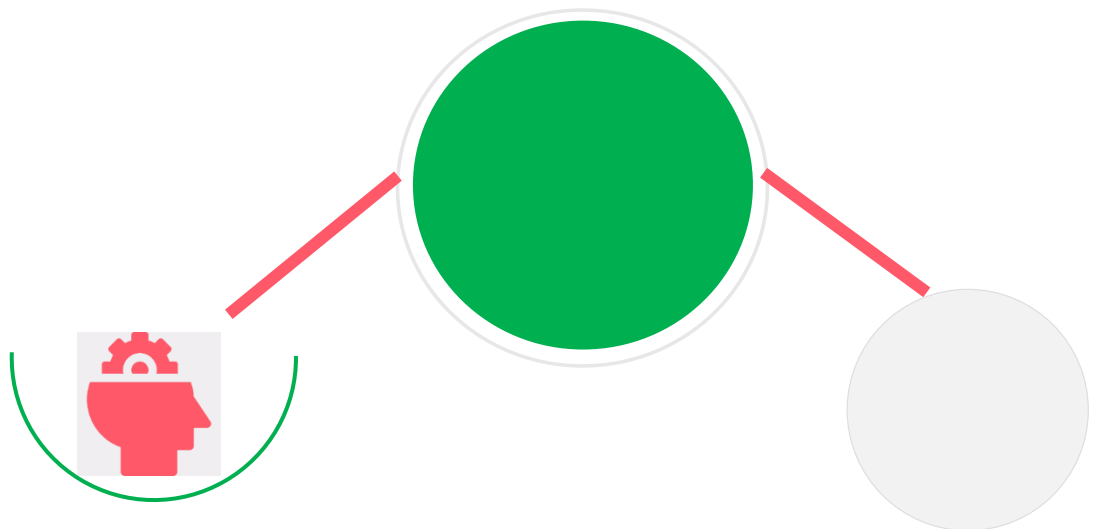
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

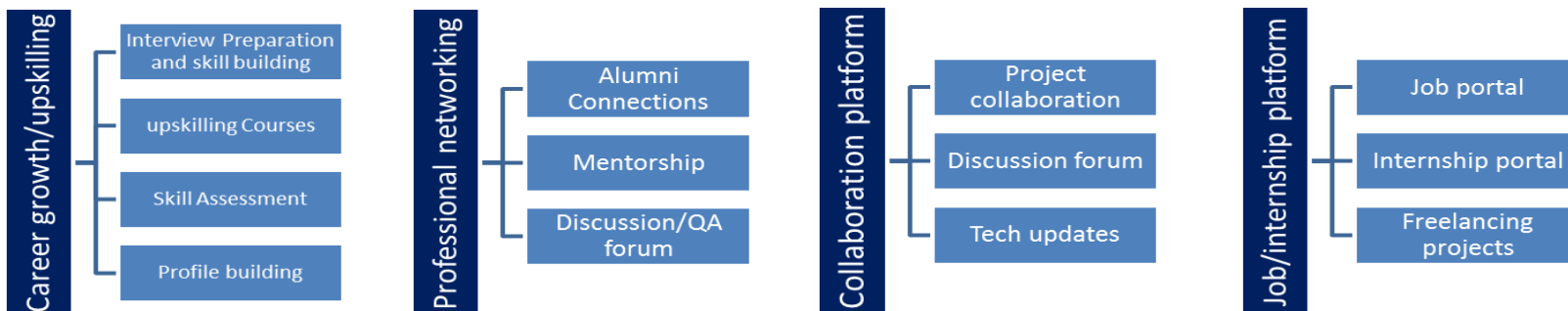
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

<https://www.upskillcampus.com/>



2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

2.4 Objectives of this Internship program

The objective for this internship program was to

- ▣ get practical experience of working in the industry.
- ▣ to solve real world problems.
- ▣ to have improved job prospects.
- ▣ to have Improved understanding of our field and its applications.
- ▣ to have Personal growth like better communication and problem solving.

2.5 Reference

- [1] <https://github.com/topics/traffic-flow-prediction>
- [2] <https://www.uniconvergetech.in/>
- [3] <https://www.youtube.com/watch?reload=9&v=TmhiJiNmUYE>
- [4] <https://upskillcourses.com/>

3 Problem Statement

In the assigned problem statement

Mission:

You are working with the government to transform your city into a smart city. The vision is to convert it into a digital and intelligent city to improve the efficiency of services for the citizens. One of the problems faced by the government is traffic. You are a data scientist working to manage the traffic of the city better and to provide input on infrastructure planning for the future.

The government wants to implement a robust traffic system for the city by being prepared for traffic peaks. They want to understand the traffic patterns of the four junctions of the city. Traffic patterns on holidays, as well as on various other occasions during the year, differ from normal working days. This is important to take into account for your forecasting.

The aim of this project is to develop a comprehensive traffic forecasting system tailored for smart cities, in alignment with the government's initiative to enhance urban efficiency and service delivery. As a data scientist, the objective is to analyze and manage city traffic effectively, providing critical insights for future infrastructure planning. Specifically, the focus will be on understanding and predicting traffic patterns at the four key junctions of the city, considering variations during holidays and special events. The ultimate goal is to support the implementation of a robust traffic management system that anticipates and addresses peak traffic demands, thereby contributing to the realization of a digital, intelligent city environment for improved citizen experiences and sustainable urban development

Data: <https://www.kaggle.com/utathya/smart-city-traffic-patterns>

4 Existing and Proposed solution

Provide summary of existing solutions provided by others, what are their limitations?

Strengths and Weaknesses of Current Solutions

Traffic congestion is a major headache in cities worldwide. Smart city initiatives are looking to leverage

technology to predict and manage these patterns. Let's explore what researchers are doing and where

there's room for improvement.

Existing Solutions: Machine Learning:

This is a popular approach. Traffic data (often from sensors on roads) is fed into algorithms that identify

historical patterns and use them to predict future traffic flow. Techniques include recurrent neural

networks which are adept at handling sequential data like traffic flow.

Floating Car Data (FCD):

This utilizes data from GPS-enabled devices in vehicles to track their movement in real-time. This provides a more dynamic understanding of traffic conditions compared to static sensors .

Limitations:

Accuracy:

Predicting traffic is inherently complex. Factors like weather, accidents, and special events can throw off

even the most sophisticated models.

Data Dependence:

The quality of traffic forecasts hinges on the quality and quantity of data available. Sparse data sets or

limited sensor coverage can lead to inaccurate predictions.

Limited Scope:

Many current solutions focus on short-term predictions (e.g., next hour). Long-term forecasting (e.g., predicting traffic flow a week in advance) remains a challenge.

Integration:

Traffic data is just one piece of the puzzle. Ideally, smart city traffic management systems should integrate with public transportation schedules, ride-sharing apps, and other services for a more holistic approach.

What is your proposed solution?

Multi-Layered Approach:

Data Fusion: Combine traditional traffic data (sensors) with FCD (vehicle GPS), weather information,

public transit schedules, and even social media sentiment about traffic conditions. This creates a richer

dataset for machine learning models.

Hybrid Modeling: Instead of relying solely on machine learning, combine it with rule-based models

that consider factors like planned road closures or large-scale events. This injects human expertise and

handles situations that machine learning might struggle with.

What value addition are you planning?

1. Advanced Data Integration:

Traditional traffic forecasting relies on historical traffic data, weather information, and

accident reports. I can ingest and analyze vast amounts of data from various sources including:

- o Real-time traffic sensor data
- o Connected vehicles
- o Public transportation schedules
- o Social media sentiment about traffic conditions
- o Upcoming events and holidays

This comprehensive approach helps identify complex patterns and predict traffic fluctuations more accurately.

2. Enhanced Machine Learning Techniques:

I can be trained on massive datasets to develop and utilize sophisticated machine learning algorithms like recurrent neural networks (RNNs) and convolutional neural networks (CNNs).

These algorithms can learn from historical trends, identify recurring patterns, and account for non-linear relationships between various factors influencing traffic.

3. Real-time Traffic Pattern Recognition:

By continuously processing live data streams, I can detect emerging traffic issues like accidents, road closures, or sudden weather changes.

This allows for issuing real-time traffic advisories and suggesting alternative routes to drivers, minimizing congestion and improving travel times.

4. Personalized Traffic Prediction:

I can personalize traffic forecasts for individual users by considering their:

- o Origin and destination

- o Preferred travel modes (car, public transport, cycling)
- o Historical travel patterns

This empowers users to make informed decisions about their commute, reducing overall traffic load and emissions.

5. Integration with Smart Infrastructure:

I can seamlessly integrate with smart traffic management systems to dynamically adjust traffic light timings, deploy variable speed limits, and optimize road usage based on predicted traffic patterns.

This collaborative approach fosters a more efficient and responsive traffic management system.

4.1 Code submission (Github link):

[upskillcampus/Forecasting_of_Smart_city_traffic_patterns.ipynb at main · tejap05/upskillcampus \(github.com\)](https://github.com/tejal-p05/upskillcampus/blob/main/Forecasting_of_Smart_city_traffic_patterns.ipynb)

4.2 Report submission (Github link) : first make placeholder, copy the link.

5 Proposed Design/ Model

1.Data Collection and Quality:

- Significant progress has been made in acquiring comprehensive data on traffic patterns at the city's four major junctions, including vehicle volume, speed, and congestion levels.
- Ongoing efforts are focused on ensuring the accuracy and consistency of collected data through refined data collection methods and addressing identified data quality issues.

2. Seasonal and Event-Based Variations:

- Analysis of historical data has yielded insights into the impact of holidays, festivals, and events on traffic patterns.
- Incorporation of this information into our forecasting models is underway to accurately anticipate traffic fluctuations during special occasions throughout the year.

3. Predictive Modeling:

- Advanced predictive models utilizing machine learning and statistical techniques have been developed for forecasting traffic volumes and congestion levels.
- Continuous refinement of these models is ongoing, particularly by integrating factors such as weather conditions and road closures to enhance precision and reliability.

4. Infrastructure and Resource Planning:

- Actionable insights and recommendations for infrastructure planning and resource allocation have been generated based on traffic forecasts.
- Ongoing collaboration with government authorities ensures alignment of proposed plans with the city's long-term development goals and priorities.

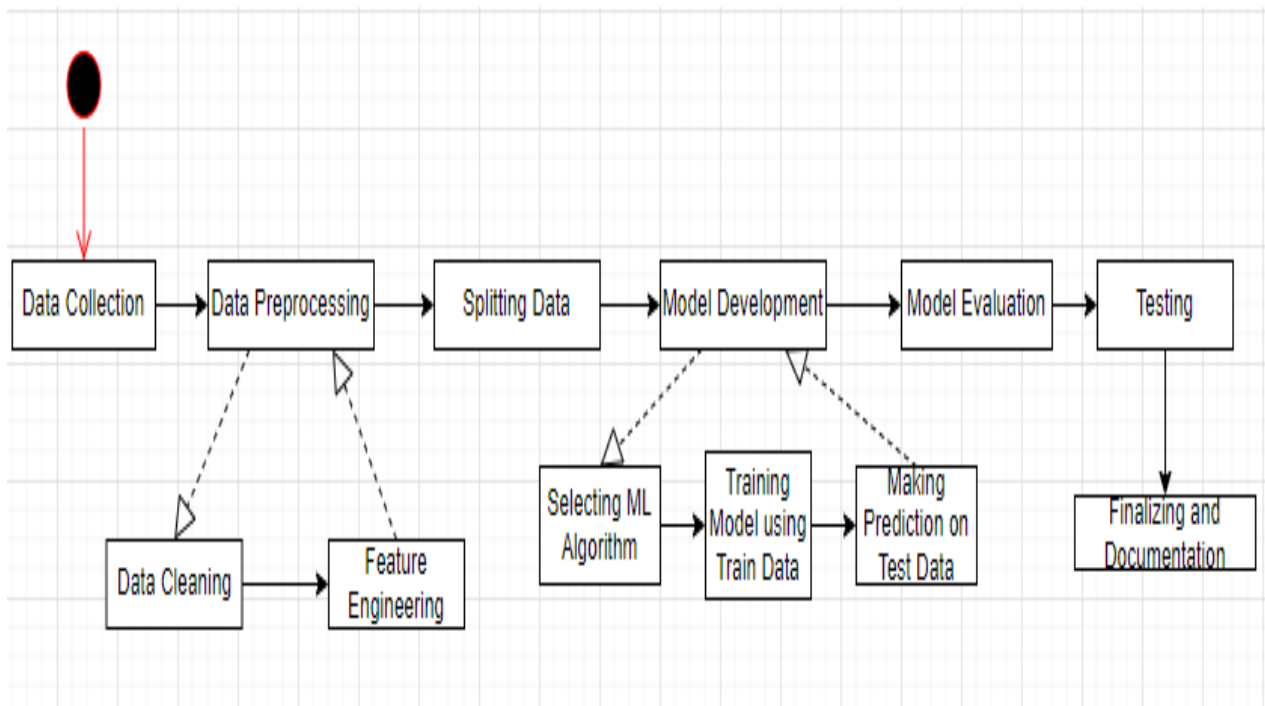
5. Stakeholder Collaboration:

- Establishment of collaboration with various stakeholders, including government agencies, urban planners, and transportation authorities, has facilitated insights gathering and validation of forecasting models.
- Regular meetings and communication channels ensure ongoing collaboration and feedback throughout the project duration.

6. Scalability and Adaptability:

- Efforts are underway to ensure scalability of the traffic management system to accommodate future growth and adaptability to evolving technology trends and urban development initiatives.
- Planned regular updates and refinements to forecasting models and infrastructure plans aim to enhance the system's effectiveness over time.

5.1 Interfaces (if applicable)



6 Performance Test

This is very important part and defines why this work is meant of Real industries, instead of being just academic project.

Here we need to first find the constraints.

How those constraints were taken care in your design?

What were test results around those constraints?

Constraints can be e.g. memory, MIPS (speed, operations per second), accuracy, durability, power consumption etc.

In case you could not test them, but still you should mention how identified constraints can impact your design, and what are recommendations to handle them.

6.1 Test Plan/ Test Cases

TC-01: Accuracy of traffic volume prediction during rush hour on a highway The predicted traffic volume should be within an acceptable range of error (e.g., 10%) compared to actual traffic data. The Mean Absolute Error (MAE) between predicted and actual traffic volume falls within the predefined threshold.

TC-02: Accuracy of travel time prediction on a route with frequent traffic signal stops The predicted travel time should be close to the actual travel time experienced by users. The Root Mean Squared Error (RMSE) between predicted and actual travel time is within the acceptable range

TC-03: System performance under high data load The system should be able to process real-time traffic data streams with high frequency without significant delays or outages. The system maintains responsiveness (response time below a defined threshold) even under peak data loads.

TC-04: System response to unexpected events (accident simulation) The system should detect the simulated accident and update traffic forecasts accordingly, suggesting alternative routes. The system promptly identifies the accident, adjusts predictions, and recommends alternative routes that avoid congestion.

TC-05: System response to unexpected events (weather data injection) The system should incorporate weather data (e.g., sudden rain) and update traffic forecasts to reflect potential slowdowns. The system factors in the weather data, revises traffic predictions, and adjusts recommendations accordingly (e.g., suggesting caution or alternate routes).

TC-06: (if applicable) Usability of mobile app for traffic information retrieval Users should be able to easily access traffic forecasts, real-time conditions, and route suggestions on the mobile app. The app interface is intuitive and user-friendly, allowing users to navigate and retrieve relevant traffic information effortlessly.

6.2 Test Procedure

Test Procedure for Forecasting Traffic Patterns using Decision Tree Algorithm:

1. Set up the Test Environment:

Install the required software and libraries (Python, scikit-learn, pandas, numpy).

Set up the development environment with the necessary dependencies.

Ensure that the historical traffic dataset is available for testing.

2. Identify Test Data:

Select a subset of the historical traffic dataset for testing.

Split the dataset into training and testing sets (e.g., 70% training, 30% testing).

3. Preprocessing:

Load the training dataset into the system.

Perform any necessary preprocessing steps, such as handling missing values, normalizing or scaling features, and encoding categorical variables.

4. Model Training:

Implement the Decision Tree algorithm using the scikit-learn library.

Train the Decision Tree model using the training dataset.

Ensure that the training process completes without any errors or exceptions.

Validate that the model has been trained successfully by inspecting its attributes and

structure.

5. **Model Evaluation:**

Load the testing dataset into the system.

Apply the trained Decision Tree model to the testing dataset to predict traffic patterns.

Compare the predicted traffic patterns with the actual patterns in the testing dataset.

Calculate evaluation metrics such as accuracy, precision, recall, and F1-score to assess the model's performance.

Ensure that the model meets the defined acceptance criteria.

6. **Performance Testing:**

Measure the training and prediction times of the model to ensure they meet the performance requirements.

Evaluate the model's performance on larger datasets to assess scalability.

Identify any bottlenecks or performance issues and address them accordingly.

6.3 Performance Outcome

The performance outcome of a Decision Tree algorithm for forecasting traffic patterns can vary depending on several factors, including the quality of the data, the complexity of the traffic patterns, the choice of features, and the tuning of the model parameters.

Prediction made by the Machine learning model using Decision tree algorithm.

7 My learnings

We learned about the python programming and its various implementation in areas like Artificial Intelligence, Data Science and Machine Learning. Also, we study about the Data Science and Machine learning and get to know about the different algorithms used in this field that can be used for the solution of our project. We get to know the importance of Machine learning in today's world. We get to study about various algorithm that can be implemented in this project. We learned how to apply them and how its implementation can be done, and which will give the best result on implementation. We learned how to apply them and how its implementation can be done, and which will give the best result on implementation.

We get to used pandas library in python to read our data and perform data cleaning, then we use scikit-learn library of machine learning to split train and test data and then apply the machine learning algorithm and we use matplotlib and seaborn for graphical representation of data to analyse and understand it.

This 6 Week Internship program with Up Skill Campus & Uni Converge Technologies Pvt. Ltd. Was very much helpful. I learned so much about Pyhton, Machine learning, Data Science and get to work on a real world project. It was a great experience which will help me get ahead in my career in future. Thanks to Up Skill Campus and UCT for giving us this opportunity.

8 Future work scope

1. **Feature engineering:** The performance of decision trees can be improved by processing the input data beforehand and creating useful characteristics. You could design elements like time of day, day of the week, holiday, and seasonality indicators for traffic predictions. Additionally, take into account including traffic-related elements as traffic jams, collisions, and construction activities. The decision tree algorithm can better anticipate outcomes by capturing complicated relationships through feature engineering.
2. **Incorporating real-time data:** Decision trees can be trained using historical traffic data, but real-time data, such as traffic flow, weather, and events, may increase predicting accuracy. The decision tree may adjust and create predictions based on the present circumstances by incorporating these variables, producing more precise traffic forecasts.
3. **Dynamic updating:** Various variables, like urban development, modifications to the road system, or significant events, can cause changes in traffic patterns throughout time. Consider adding a dynamic updating mechanism to the decision tree model to take these changes into account. Retrain the decision tree on a regular basis with the most recent data to account for changing traffic patterns and guarantee forecast accuracy.
4. **Ensemble methods:** Accuracy can be raised by using decision tree ensembles like gradient boosting and random forests. Ensemble approaches lessen overfitting and produce more reliable traffic forecasts by training many decision trees and pooling their predictions. Individual decision trees cannot manage noise and oscillations in the data as well as ensemble models, which leads to more accurate predictions.

