Industrial Internship Report on

Project Name: - Forecasting of Smart city traffic patterns

Student name: - Kanan Patel

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 Smart City Traffic patterns forecasting. 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

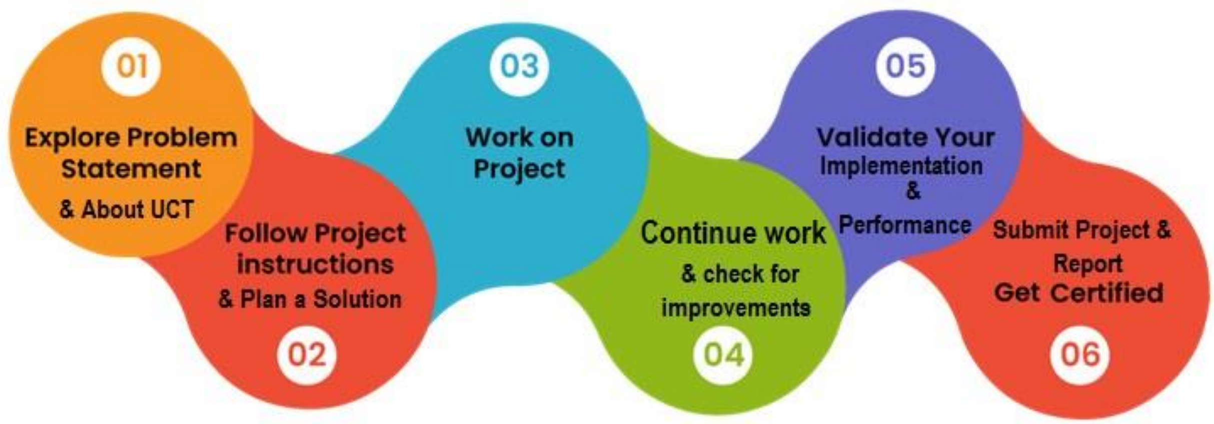
Summary of the whole 6 weeks' work.

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/LoRaWAN), Java Full Stack, Python, Front end** etc.

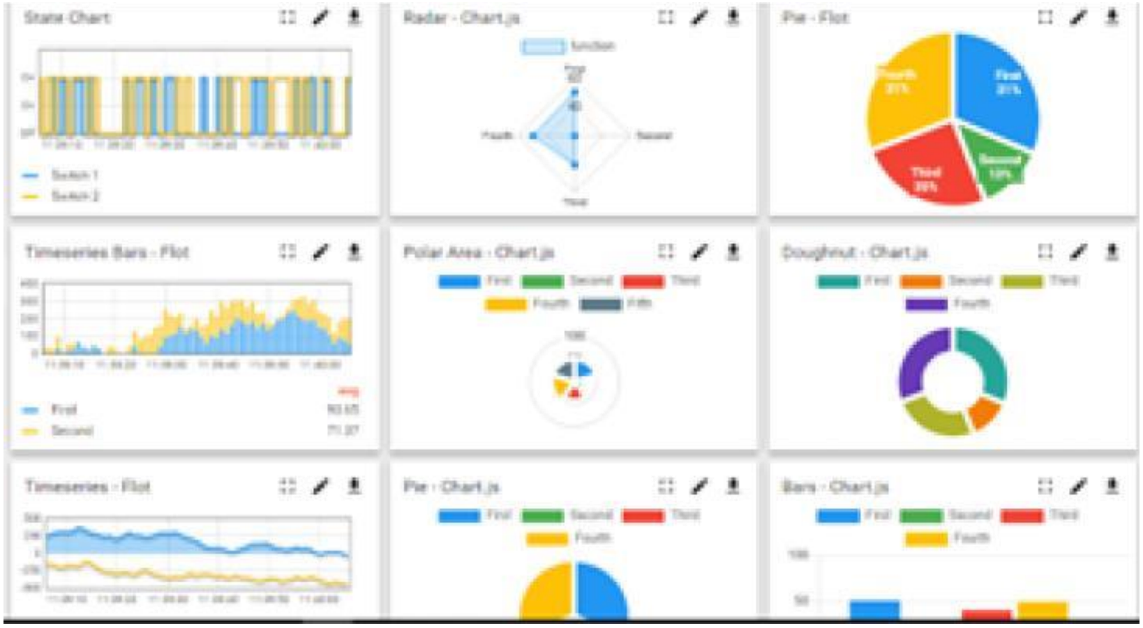
i. UCT IoT Platform ( )

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.

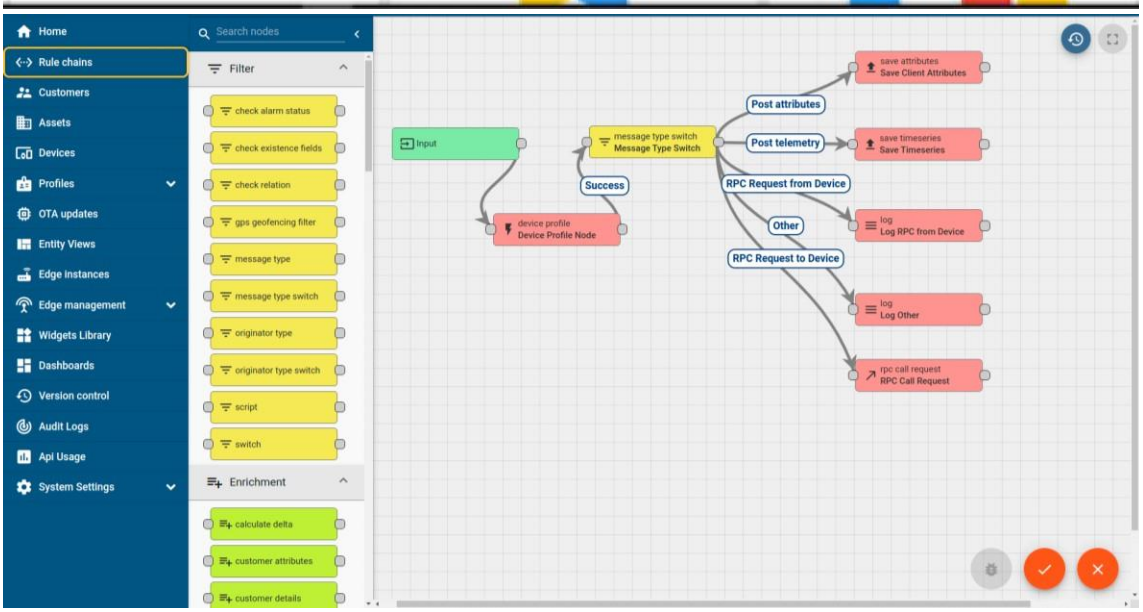
- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

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



ii.



Smart Factory Platform (FACTORY WATCH)

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 unleased 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.

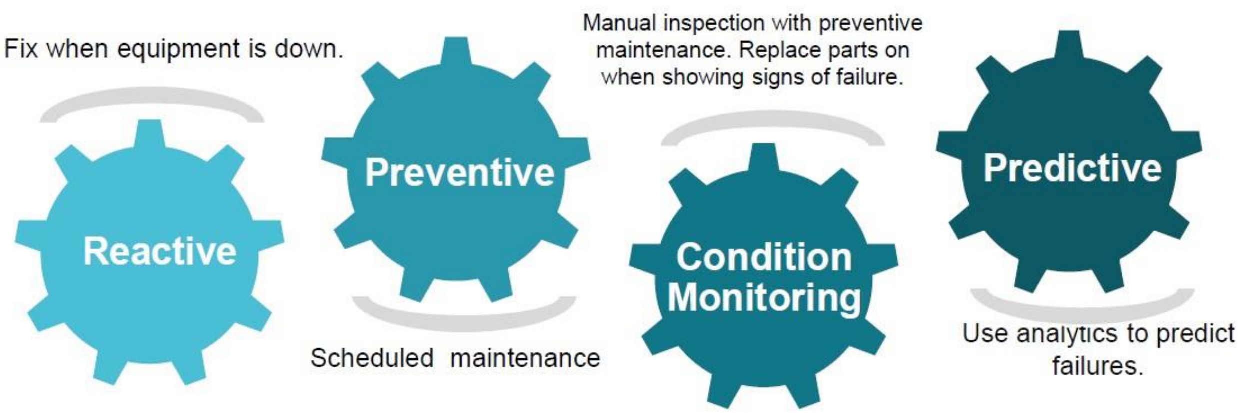
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i

iii. based Solution

UCT is one of the early adopters of LoRAWAN technology 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 measurab

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 ☛

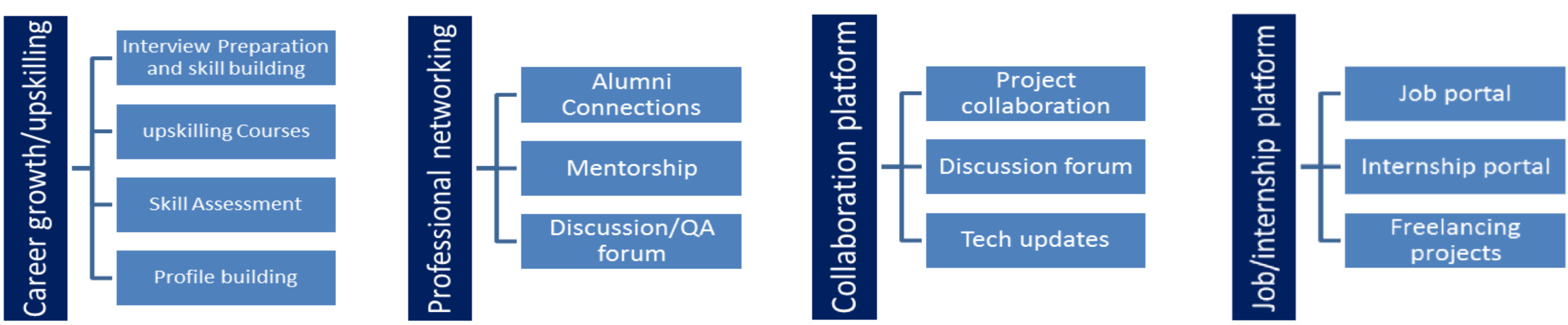
get practical experience of working in the industry.

☛ 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/mratsim/McKinsey-SmartCities-Traffic-Prediction

[2] https://github.com/topics/traffic-flow-prediction

[3] https://www.youtube.com/watch?reload=9&v=TmhiJiNmUYE

2.6 Glossary

Terms	Acronym
Reduce congestion: pen_spark	By predicting traffic flow, cities can implement measures to prevent gridlock, like dynamic traffic light adjustments or rerouting suggestions for drivers.
Improve travel times:	Accurate traffic forecasts help commuters choose optimal routes and departure times, leading to faster journeys.
Lower emissions:	Reduced congestion means less idling and smoother traffic flow, which can significantly decrease pollution levels.
Enhanceurban planning	Traffic forecasting informs infrastructure development, allowing cities to build roads and public transport systems that meet future needs.

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.

Your task: To predict traffic patterns in each of these four junctions for the next 4 months.

Data: The sensors on each of these junctions were collecting data at different times, hence you will see traffic data from different time periods. To add to the complexity, some of the junctions have provided limited or sparse data requiring thoughtfulness when creating future projections. Depending upon the historical data of 20 months, the government is looking to you to deliver accurate traffic projections for the coming four months. Your algorithm will become the foundation of a larger transformation to make your city smart and intelligent.

The evaluation metric for the competition is RMSE. Public-Private split for the competition is 25:75.

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 ○ Real-time traffic sensor data ○ Connected vehicles ○ public transportation schedules
    - social media sentiment about traffic conditions ○ 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:

- Origin and destination
- 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)

https://github.com/TechInnovatorKP/Upskill-Campus.git

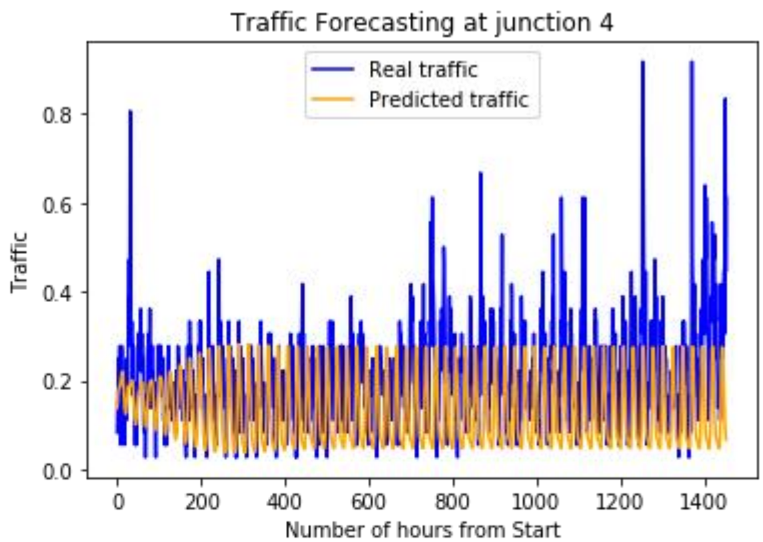
4.2 **Report submission (Github link) :** first make placeholder, copy the link.

https://github.com/mratsim/McKinsey-SmartCities-Traffic-Prediction

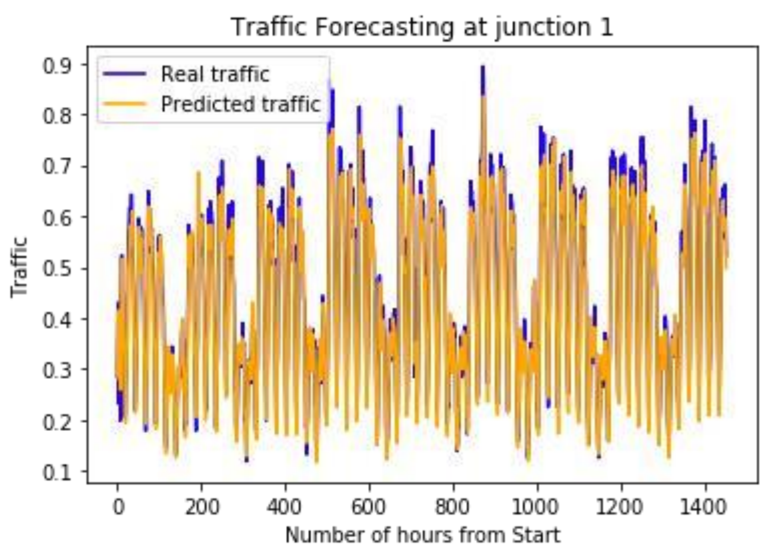
5 Proposed Design/ Model

Given more details about design flow of your solution. This is applicable for all domains. DS/ML Students can cover it after they have their algorithm implementation. There is always a start, intermediate stages and then outcome.

5.1 High Level Diagram (if applicable)

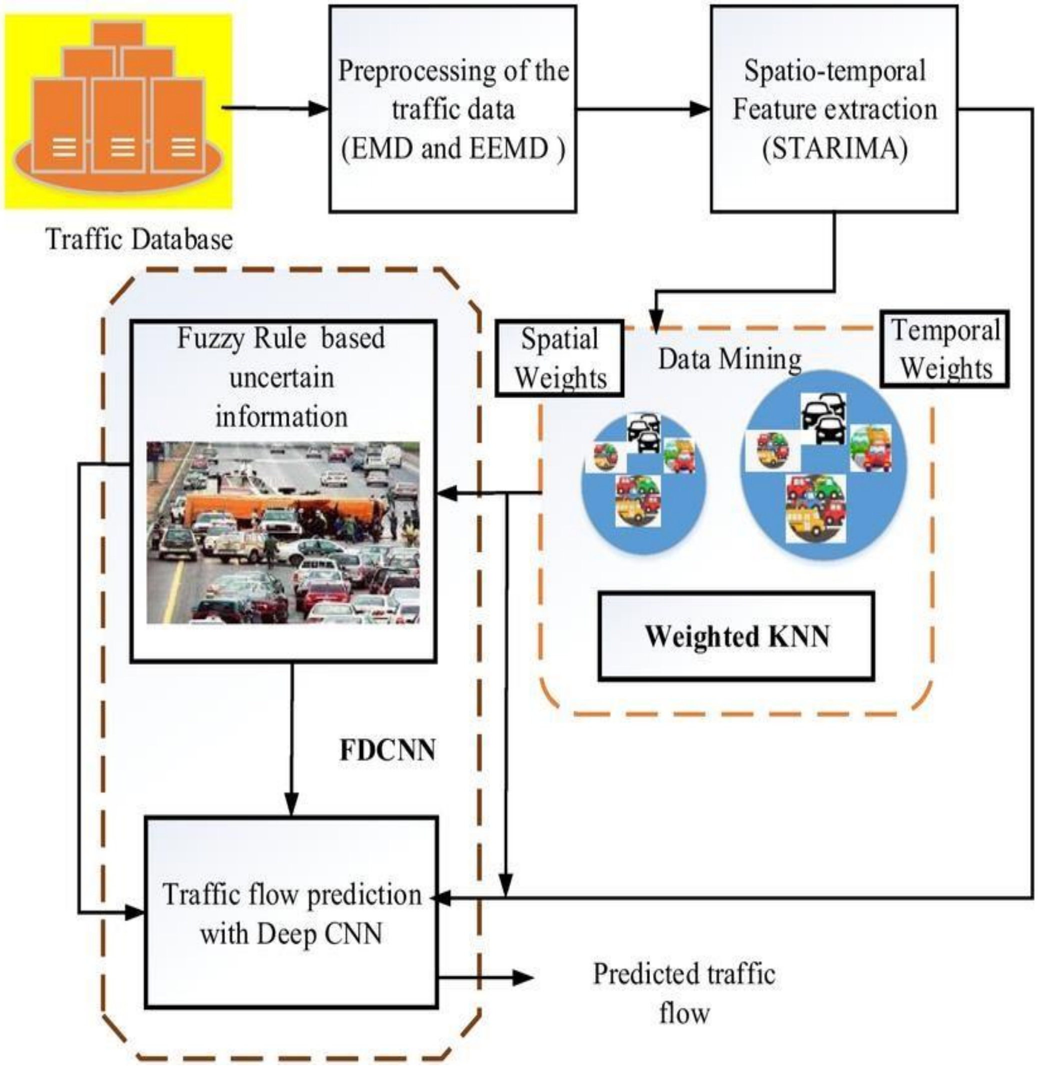


5.2 Low Level Diagram (if applicable)

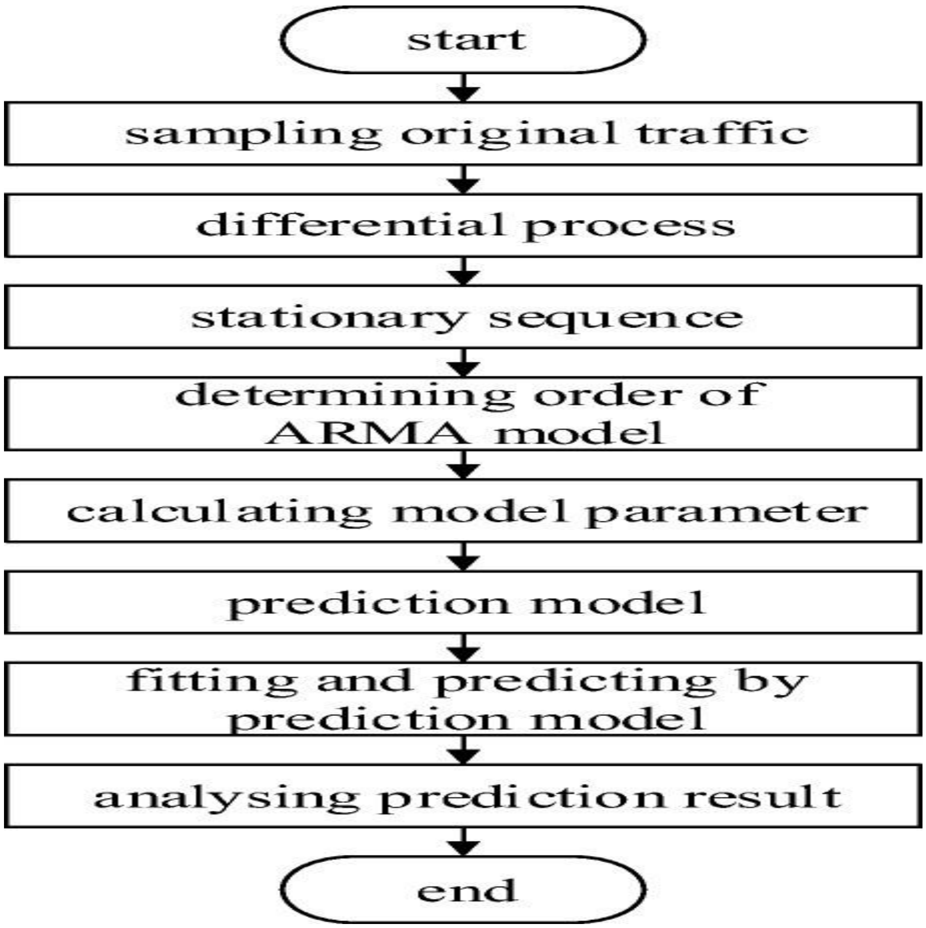


5.3 Interfaces (if applicable).

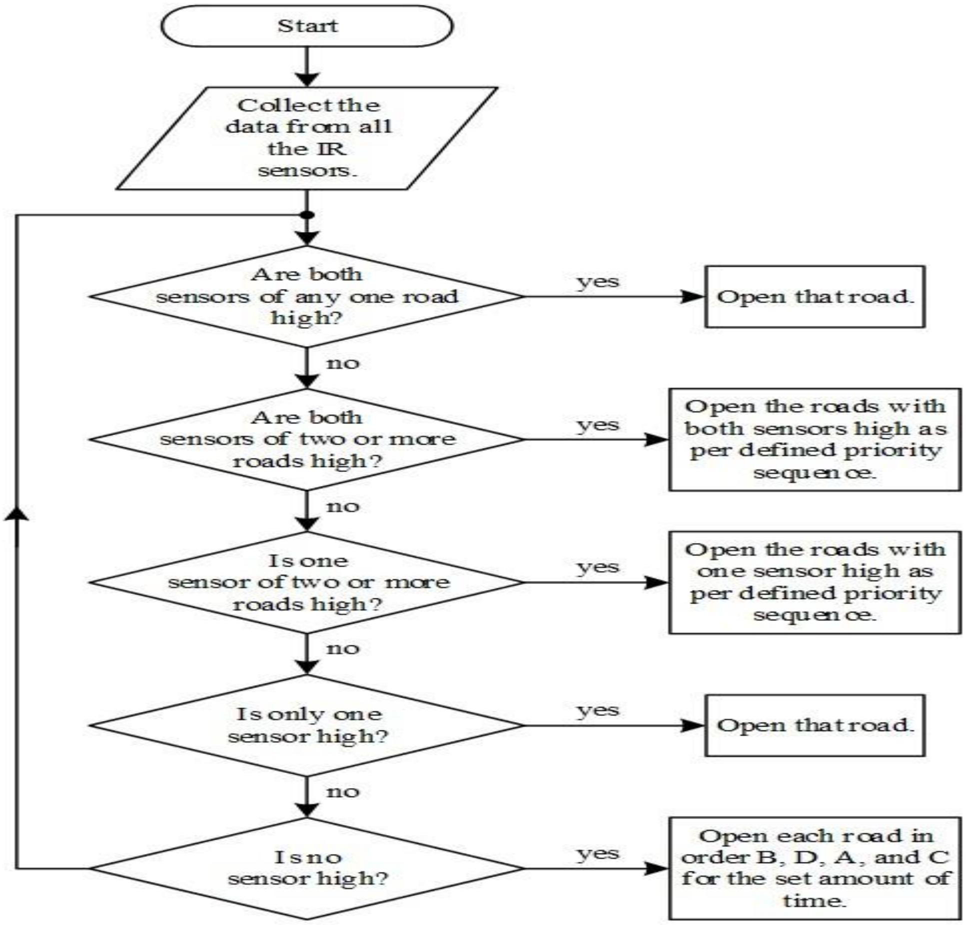
- **Block Diagrams: -**


Traffic Database
Preprocessing of the traffic data (EMD and EEMD)
Spatio-temporal Feature extraction (STARIMA)
Fuzzy Rule based uncertain information
Spatial Weights
Data Mining
Temporal Weights
Weighted KNN
FDCNN
Traffic flow prediction with Deep CNN
Predicted traffic flow

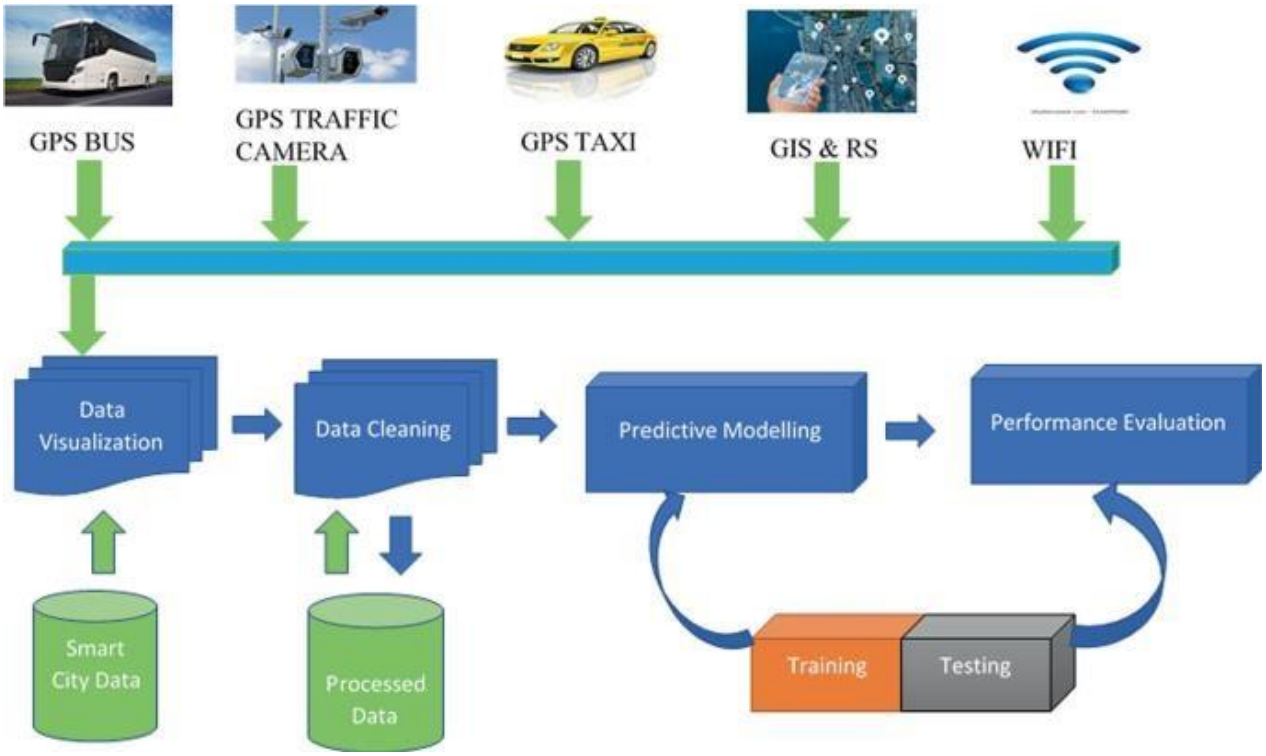
Data flow Diagram:-



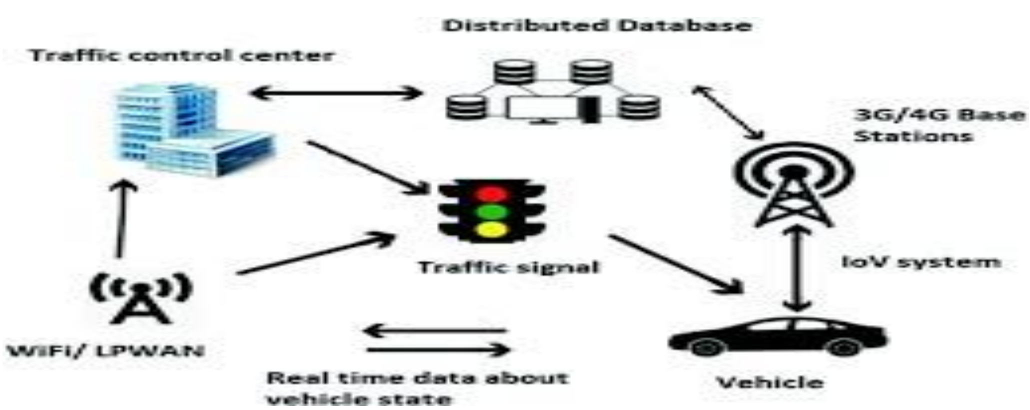
FLOW Charts: -



State Machines: -



Memory Buffer Management:-



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 were those constraints 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

This test plan outlines the strategy to validate the functionalities of the smart traffic pattern forecasting system. The test cases will assess the system's accuracy, performance, and robustness under various traffic conditions and data scenarios.

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

This document outlines the step-by-step procedure for testing the functionalities of the smart traffic pattern forecasting system. It complements the test plan by detailing the actions involved in executing the test cases.

Pre-requisites:

The smart traffic pattern forecasting system is fully functional and deployed in a test environment.

All necessary testing tools and data sets are prepared (e.g., traffic simulators, historical traffic data, weather data).

Test cases are clearly defined and documented (refer to section 1.1.2 of the Test Plan).

Test Procedure:

1.System Configuration:

Configure the system with appropriate parameters like data sources, prediction horizons, and machine learning models (if applicable).

Ensure data pipelines are functional, and real-time data feeds (if used) are established.

2.Test Case Execution:

For each test case:

- Set up the testing environment based on the specific scenario (e.g., rush hour traffic simulation, weather data injection).
- Trigger the system to generate traffic forecasts for the defined timeframe and geographical area.
- Collect the predicted traffic data (e.g., traffic volume, travel time) from the system.

3.Data Collection and Analysis:

Gather actual traffic data (e.g., from historical records, traffic sensor readings) corresponding to the test scenario.

Employ statistical methods to compare the predicted and actual traffic data.

Calculate metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to assess prediction accuracy.

Evaluate response times for processing real-time data loads.

4.User Interface Testing (if applicable):

If the system includes user interfaces (mobile app, dashboard), conduct usability testing.

Recruit participants to navigate the interface and retrieve traffic information.

Observe user interactions and gather feedback on ease of use and functionality.

5.Result Recording and Evaluation:

Document all test results, including collected data, calculated metrics, and observations from user interface testing.

Analyze the results against the pass/fail criteria defined in the test plan (refer to section 1.1.2).

Identify any discrepancies or unexpected behavior in the system's performance.

6.Defect Reporting and Re-testing:

Report any identified issues or defects in the system's functionality or performance.

Address the defects through bug fixes or system modifications.

Re-execute relevant test cases to verify the effectiveness of the implemented solutions.

6.3 Performance Outcome

The performance outcome of a smart traffic pattern forecasting system can be evaluated using various metrics depending on the specific goals and functionalities. Here are some key performance indicators (KPIs) to consider:

Accuracy:

Mean Absolute Error (MAE): -

Measures the average difference between predicted and actual traffic volume/travel time. Lower MAE indicates higher accuracy.

Root Mean Squared Error (RMSE):-

Captures the spread of errors between predicted and actual values. Lower RMSE signifies better accuracy.

Timeliness:

Prediction Latency:-

Measures the time it takes for the system to generate forecasts after receiving new data. Lower latency is desirable for real-time applications.

Performance:

Throughput: The rate at which the system can process data and generate forecasts, especially under high data loads. Higher throughput ensures responsiveness during peak traffic periods.

User Experience (if applicable):-

User Satisfaction Surveys: -

Gather feedback from users of mobile apps or dashboards to assess ease of use, clarity of information, and overall satisfaction.

App Engagement Metrics:-

Track metrics like app downloads, daily active users, and time spent using the app to gauge user interest and effectiveness.

Real-World Impact:-

Reduction in Traffic Congestion: -

Measured through metrics like average travel time or speed during peak hours. Lower travel times indicate a positive impact on traffic flow.

Improved Travel Efficiency:-

Evaluated by user surveys or tracking changes in route choices suggested by the system. User feedback and a reduction in unnecessary travel time can indicate improved efficiency.

It's important to establish performance benchmarks for these KPIs before deploying the system. This allows for a clear comparison between predicted and actual performance after implementation.

Here's an example of how performance outcomes can be reported:

The smart traffic forecasting system achieved an average MAE of 5% in predicting traffic volume during rush hour.