

Industrial Internship Report on "Forecasting Of Smart City Traffic Pattern"

**Prepared by
[N Kalyani Pushpa]**

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 is Forecasting smart city traffic patterns which involves usage of data analytics and technology to predict how traffic will behave in urban areas

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

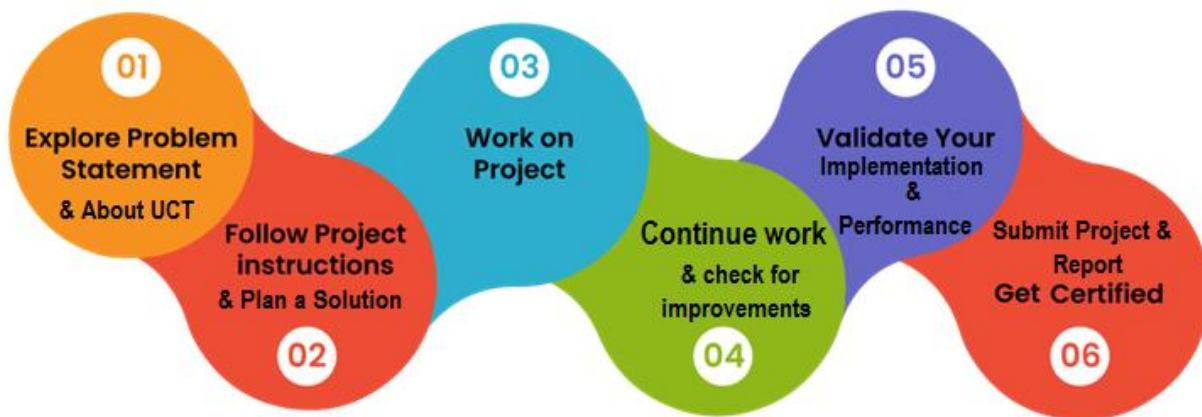
1	Preface	3
2	Introduction	4
2.1	About UniConverge Technologies Pvt Ltd	4
2.2	About upskill Campus	8
2.3	Objective	10
2.4	Reference	10
2.5	Glossary	10
3	Problem Statement	11
4	Existing and Proposed solution	11
5	Proposed Design/ Model	13
5.1	Regression model	13
6	Performance Test	14
6.1	Performance Outcome	16
7	My learnings	18
8	Future work scope	18

1. Preface

In a 6-week project on forecasting smart city traffic patterns, I developed predictive models using historical and real-time data to anticipate traffic congestion and road conditions. This work contributes to improving traffic management and transportation efficiency in smart cities. About need of relevant Internship in career development.

- **Problem Statement :** Develop a smart city traffic management solution to improve citizen services by understanding and forecasting traffic patterns at four city junctions, including variations on holidays and special occasions, to enable efficient infrastructure planning and prepare for traffic peaks.
- **Solution:** Implement data-driven traffic forecasting models and intelligent traffic management systems for the city's four key junctions, incorporating holiday and special occasion patterns, to optimize traffic flow and enhance future infrastructure planning."

How Program was planned



- **My Learnings and overall experience:** Through my upskill internship in data science and machine learning, I gained valuable hands-on experience, improving my skills in data analysis, modeling, and programming. I worked on real-world projects, collaborated with a diverse team, and enhanced my problem-solving abilities. This internship has been a pivotal learning experience, equipping me with practical knowledge and a deeper understanding of the field.

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.



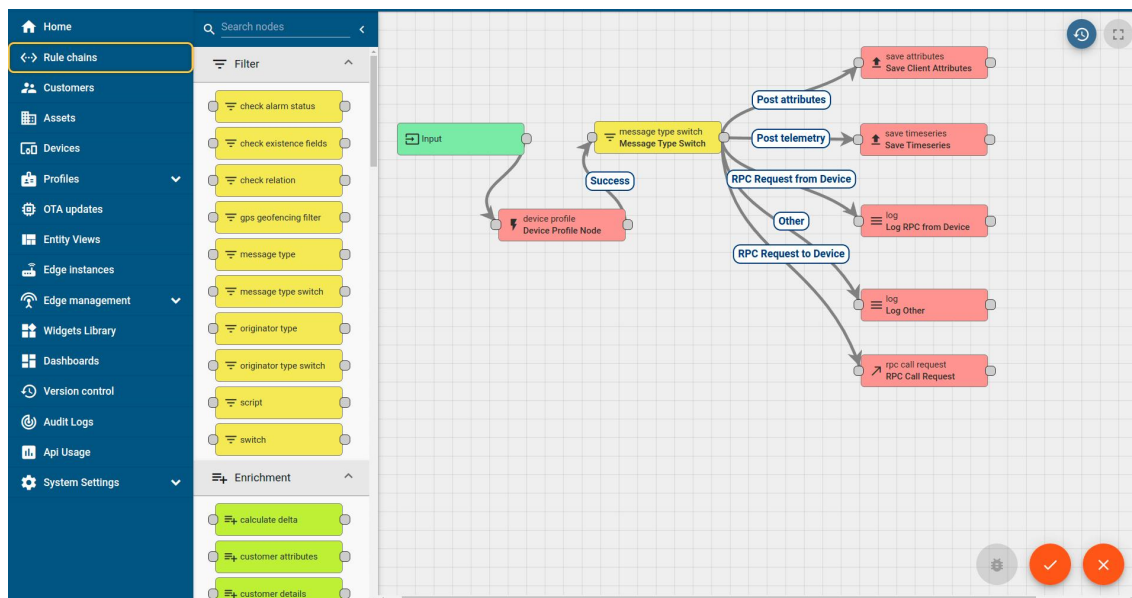
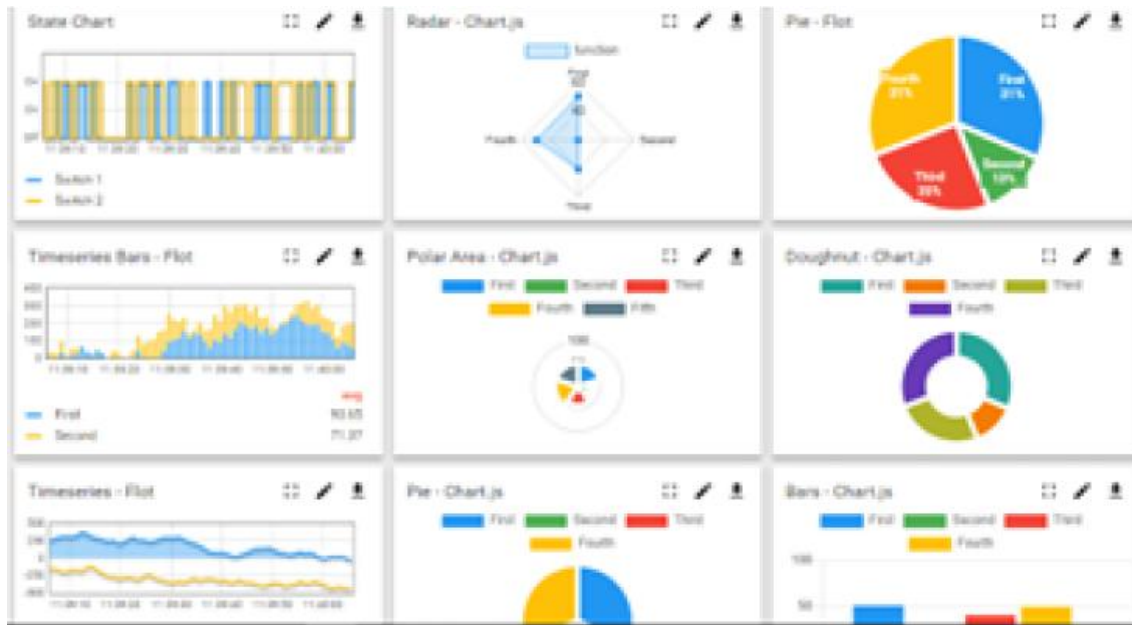
i. UCT IoT Platform ()

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



FACTORY **WATCH**

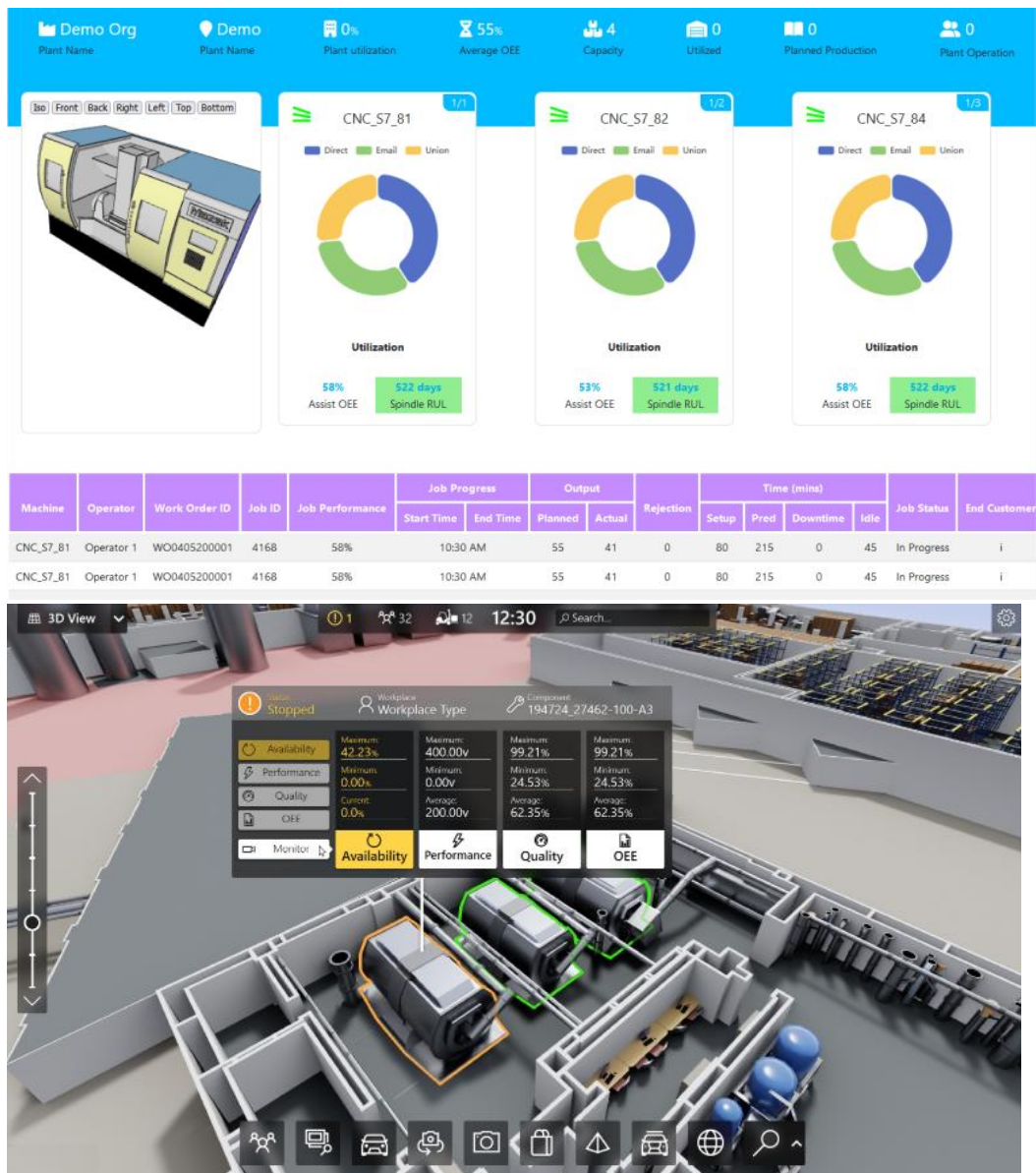
ii. Smart Factory Platform ()

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 unleash 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 want 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.



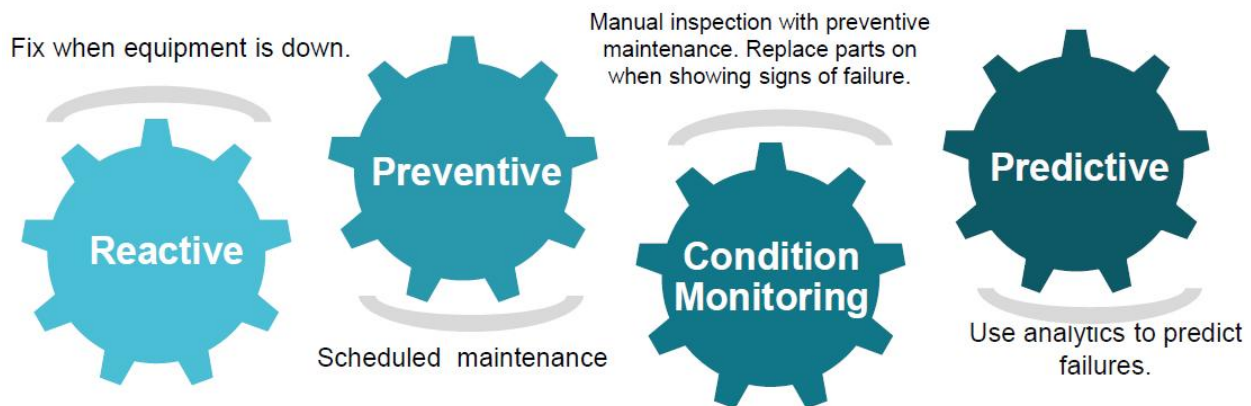


iii. LoRaWAN based Solution

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

iv. Predictive Maintenance

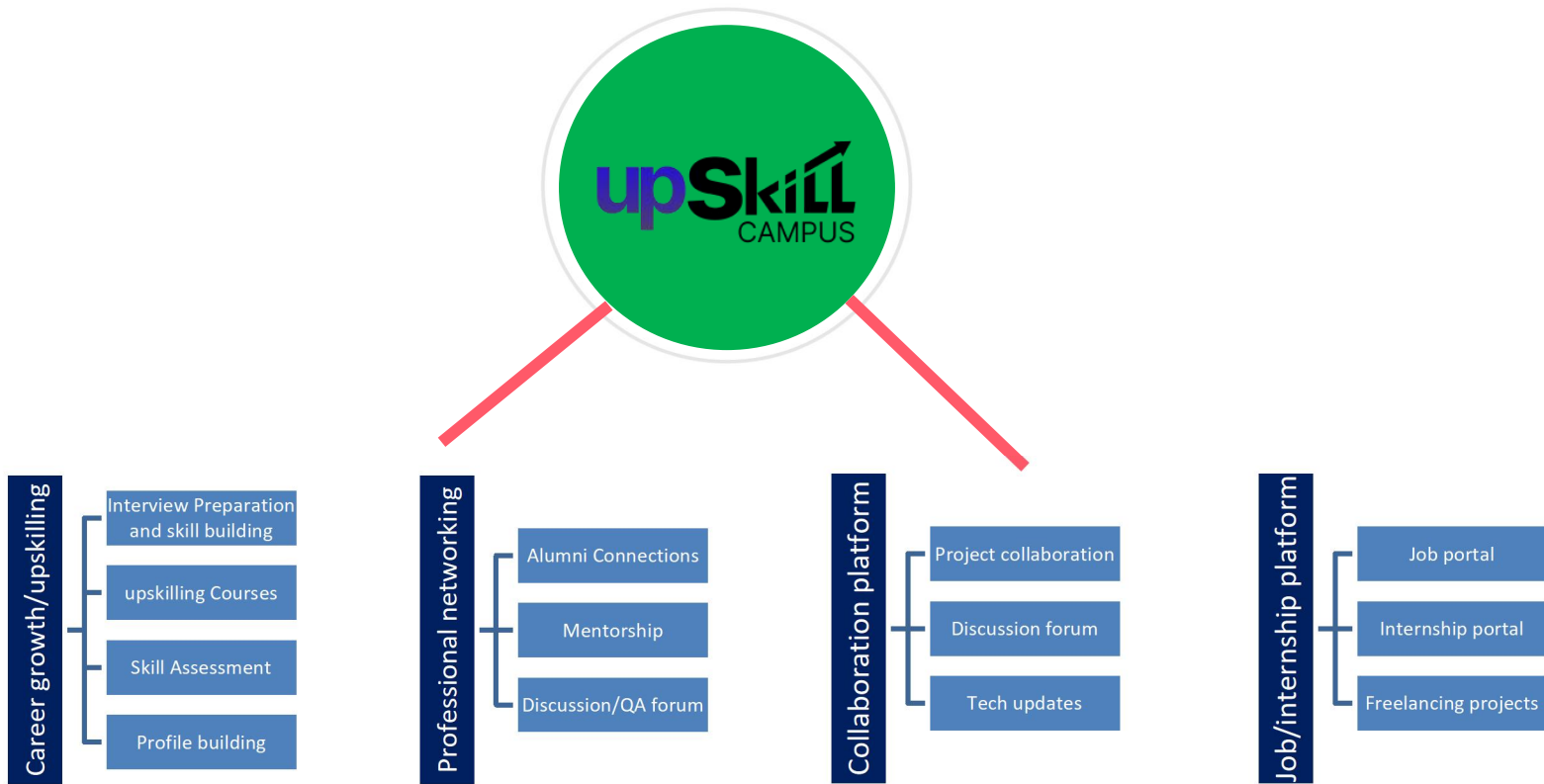
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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 measurable way.



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.

2.3 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.

2.4 Reference

- [1] Introducing Data Science by : Davy Cielen, Arno D.B. Maysman and Mohamed Ali
- [2] An Introduction To Probability And Statistics by: Vijay K. Rohatgi and A.K.Md Ehsanes Saleh
- [3] Introduction To Machine Learning by: Alex Smola and S.V.N. Vishwanathan

2.5 Glossary

Terms	Acronym
Big Data	Extremely large and complex datasets that cannot be effectively managed with traditional data processing tools.
Clustering	A machine learning technique that groups similar data points together based on certain features or characteristics.
Cross-validation	A method to assess the performance of a machine learning model by dividing the dataset into subsets for training and testing.
Hyperparameter	Parameters in a machine learning algorithm that are not learned from the data but set prior to training.
Deep Learning	A subset of machine learning that focuses on artificial neural networks with multiple layers, used for complex tasks like image and speech recognition.

3 Problem Statement

We are working with the government to transform various cities 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.

4 Existing and Proposed solution

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

- **Machine Learning Models:** Machine learning algorithms, such as neural networks and decision trees, are used to analyze historical data and predict traffic patterns. They can provide accurate forecasts but may struggle with real-time data integration and require substantial computational resources.
- **Data Fusion:** Some solutions integrate data from multiple sources, combining traffic cameras, mobile apps, and other sensors. This approach enhances accuracy but can be complex to implement and costly.
- **Traffic Prediction Platforms:** Companies and startups provide traffic prediction platforms that leverage various data sources and machine learning. These platforms offer comprehensive solutions but may require ongoing subscription fees and vendor lock-in.

➤ Limitations

- **Privacy Concerns:** Collecting and sharing real-time data may raise privacy issues, particularly when using data from GPS and mobile apps.
- **Data Accuracy:** The accuracy of predictions depends on the quality and coverage of data sources. Incomplete or inaccurate data can lead to less reliable forecasts.
- **Integration Challenges:** Integrating data from various sources and different vendors can be complex and may require significant technical expertise.

ii. What is your proposed solution?

Many reports of the traffic data are of actual time but it is not favorable and accessible to many users as we need to have prior decision in which route we need to travel. For example, During working days, we need to have daily traffic information or at times we need hourly traffic information but then the traffic congestion occurs; for solving this issue the user need to have actual time traffic prediction. Many factors are responsible for the traffic congestion. This can be predicted by taking two datasets; one with the past year and one with the recent year's data set. If traffic is so heavy then the traffic can be predicted by referring the same time in the past year's data set and analyzing how congested the traffic would be. With the increasing cost of the fuel, the traffic congestion changes drastically. The goal of this prediction is to provide real-time gridlock and snarl up information. The traffic on the city becomes complex and are out of control these days, so such kind of systems are not sufficient for prediction. Therefore, research on traffic flow prediction plays a major role in ITS

1.1 Code submission (Github link): <https://github.com/Nkpsuga/project>

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

5 Proposed Design/ Model

In traffic congestion forecasting there are data collection and prediction model.

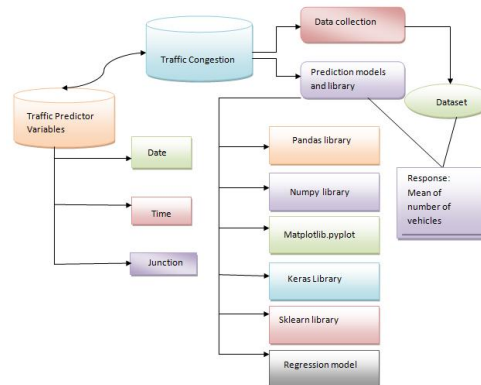


Figure 1. Outline of the traffic prediction in this paper

5.1 Regression model

Regressor model analysis could even be a mathematical technique for resolving the connection in the middle of one dependent (criterion) variable and one or more independent (predictor) variables.

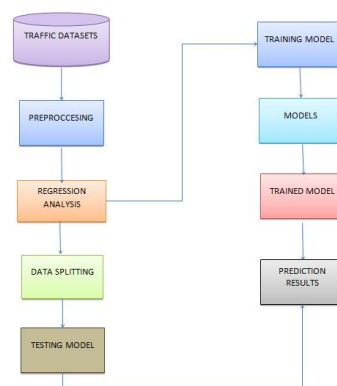
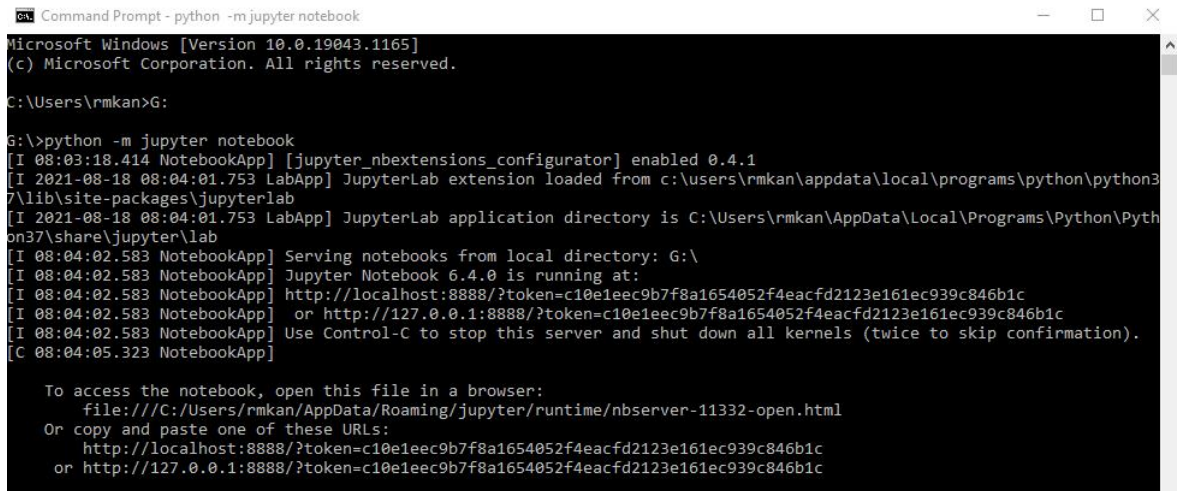


Figure 2. Regression model of traffic prediction in this paper

6 Performance Test

The command prompt is the local host in this paper to initialize the jupyter notebook.



```
Command Prompt - python -m jupyter notebook
Microsoft Windows [Version 10.0.19043.1165]
(c) Microsoft Corporation. All rights reserved.

C:\Users\rmkam>G:

G:\>python -m jupyter notebook
[I 08:03:18.414 NotebookApp] [jupyter_nbextensions_configurator] enabled 0.4.1
[I 2021-08-18 08:04:01.753 LabApp] JupyterLab extension loaded from c:\users\rmkam\appdata\local\programs\python\python37\lib\site-packages\jupyterlab
[I 2021-08-18 08:04:01.753 LabApp] JupyterLab application directory is C:\Users\rmkam\AppData\Local\Programs\Python\Python37\share\jupyter\lab
[I 08:04:02.583 NotebookApp] Serving notebooks from local directory: G:\
[I 08:04:02.583 NotebookApp] Jupyter Notebook 6.4.0 is running at:
[I 08:04:02.583 NotebookApp] http://localhost:8888/?token=c10e1eec9b7f8a1654052f4eacfd2123e161ec939c846b1c
[I 08:04:02.583 NotebookApp] or http://127.0.0.1:8888/?token=c10e1eec9b7f8a1654052f4eacfd2123e161ec939c846b1c
[I 08:04:02.583 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).

[C 08:04:05.323 NotebookApp]

To access the notebook, open this file in a browser:
file:///C:/Users/rmkam/AppData/Roaming/jupyter/runtime/nbserver-11332-open.html
Or copy and paste one of these URLs:
http://localhost:8888/?token=c10e1eec9b7f8a1654052f4eacfd2123e161ec939c846b1c
or http://127.0.0.1:8888/?token=c10e1eec9b7f8a1654052f4eacfd2123e161ec939c846b1c
```

Figure 3. The figure signifies the initializing the jupyter notebook through command prompt

The local host contains the nbextensions which we modify to our convenience

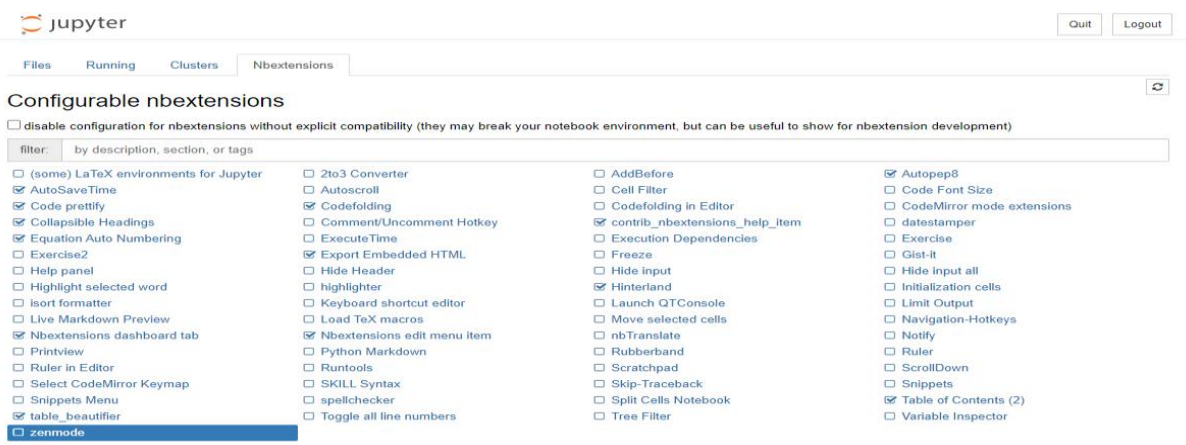


Figure 4. The figure denotes the necessary nbextensions that is needed for the prediction to take place

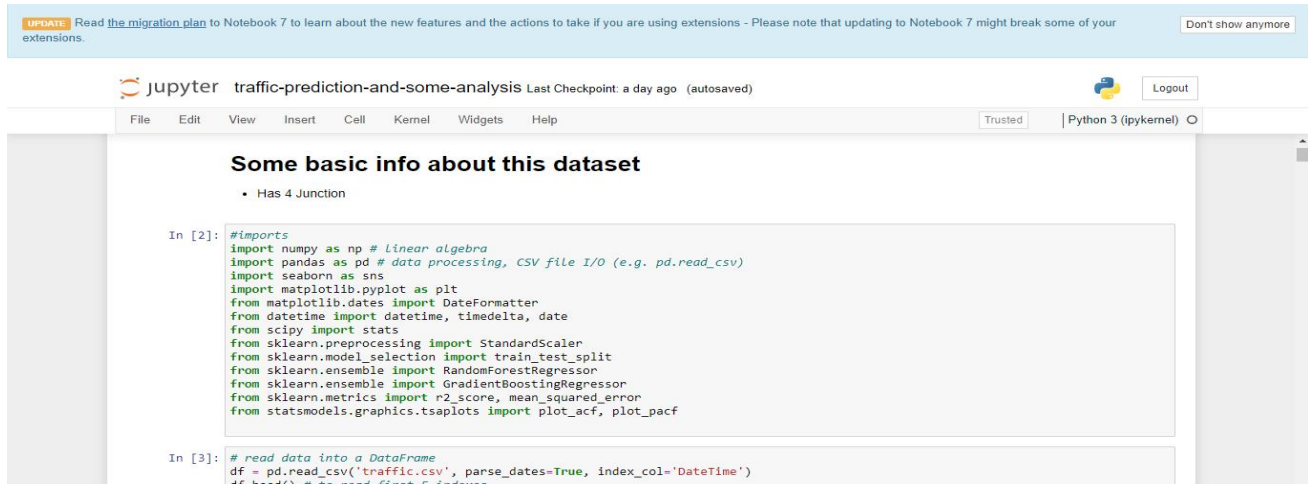


Figure 5. The figure implies the page of the jupyter notebook

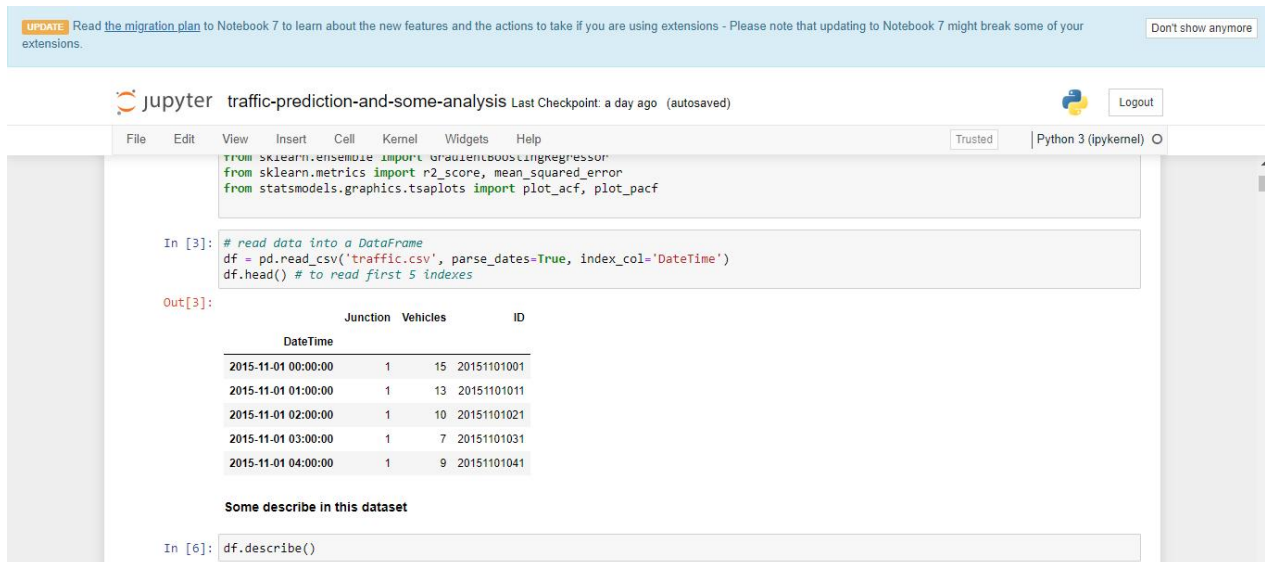


Figure 6. The figure above denotes the result after running the particular command in the prediction process will be displayed under out command

The below mentioned are the libraries used for the prediction of traffic

```
import pandas as pd    # library imported
import numpy as np     # library imported
```

6.1 Performance Outcome

The results of the traffic are as follows which by the matplotlib library

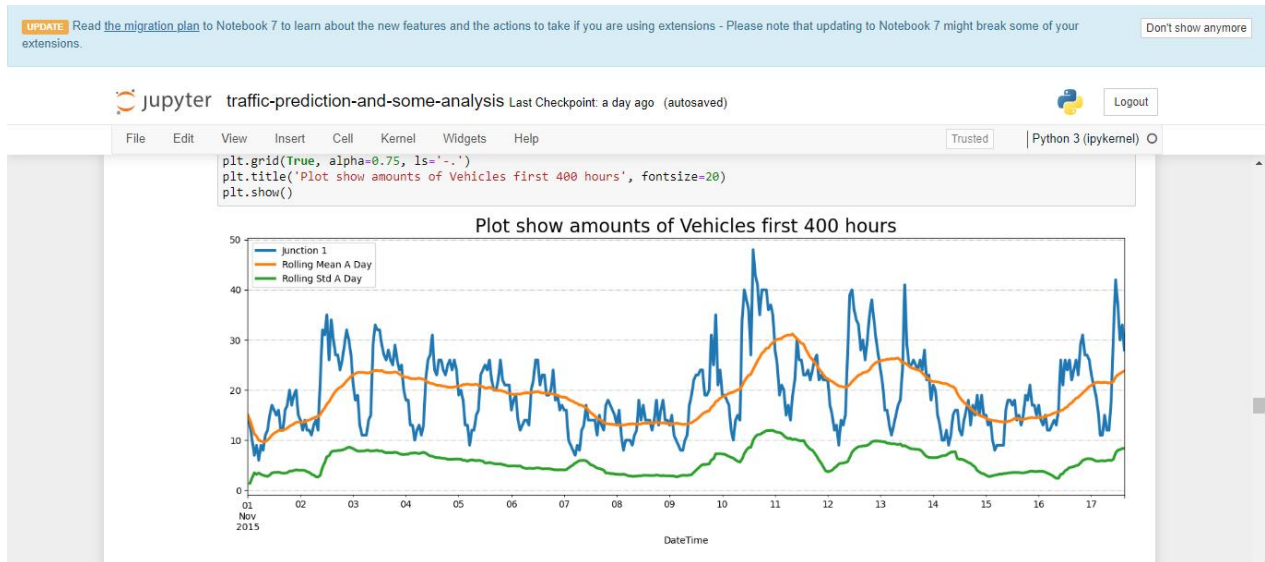


Figure 7. The figure signifies the Traffic prediction of Junction 1 from the datasets

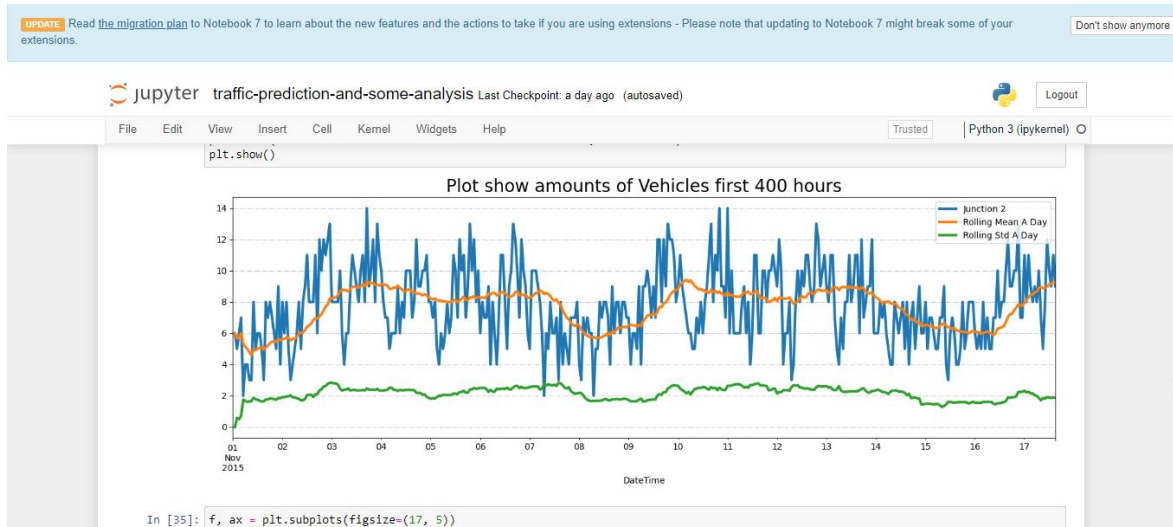


Figure 8. The figure signifies the Traffic prediction of Junction 2 from the datasets

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

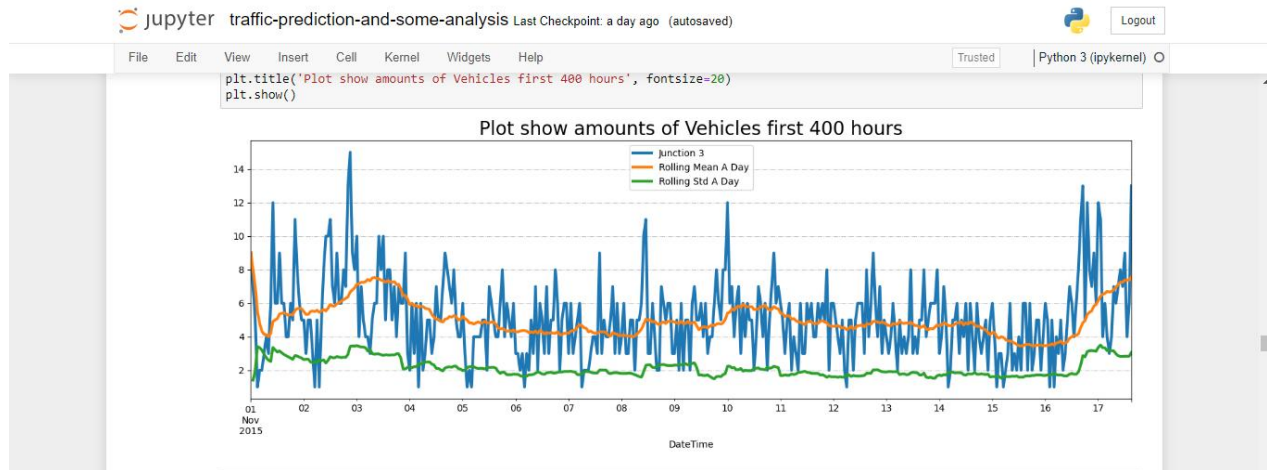


Figure 9. The figure signifies the Traffic prediction of Junction 3 from the datasets

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

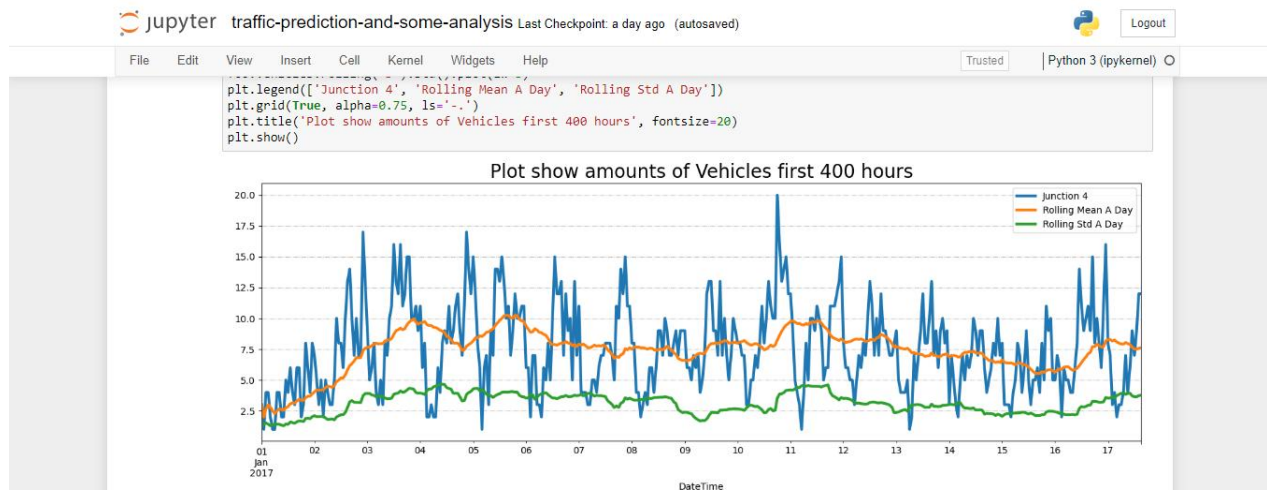


Figure 10. The figure signifies the Traffic prediction of Junction 4 from the datasets

7 My learnings

My learning about forecasting smart city traffic patterns has equipped me with a valuable skill set that holds great potential for my career growth. Understanding the intricacies of traffic data analysis, predictive modeling, and the application of advanced technologies in smart city planning has not only broadened my knowledge base but also opened doors to various career opportunities. With this expertise, I can contribute significantly to urban planning and transportation management initiatives, aiding in the development of more efficient and sustainable smart cities. This knowledge is not only relevant to roles in urban planning and transportation management but also in data science, AI development, and consulting, where I can provide valuable insights and solutions for traffic-related challenges. Furthermore, it positions me as a valuable asset in the ever-evolving field of smart city development, offering the potential for career advancement and a meaningful impact on improving urban living conditions.

8 Future work scope

In the future, the system are often further improved using more factors that affect traffic management using other methods like deep learning, artificial neural network, and even big data. The users can then use this technique to seek out which route would be easiest to achieve on destination. The system can help in suggesting the users with their choice of search and also it can help to find the simplest choice where traffic isn't in any crowded environment. Many forecasting methods have already been applied in road traffic jam forecasting. While there's more scope to create the congestion prediction more precise, there are more methods that give precise and accurate results from the prediction. Also, during this period, the employment of the increased available traffic data by applying the newly developed forecasting models can improve the prediction accuracy. These days, traffic prediction is extremely necessary for pretty much every a part of the state and also worldwide. So, this method of prediction would be helpful in predicting the traffic before and beforehand. For better congestion prediction, the grade and accuracy are prominent in traffic prediction. within the future, the expectation are going to be the estimation of established order accuracy prediction with much easier and user-friendly methods so people would find the prediction model useful and that they won't be wasting their time and energy to predict the information. There will be some more accessibility like weather outlook, GPS that's the road and accident-prone areas will be highlighted in order that people wouldn't prefer using the paths which aren't safe and simultaneously they'll predict the traffic. This will be done by deep learning, big data, and artificial neural networks