Industrial Internship Report on

Smart City Traffic Forecasting

Prepared by

Manu Premachandran

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 4 weeks' time. My project was to develop a machine learning model for forecasting traffic patterns in smart cities using historical traffic data. 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

This report summarizes my four-week internship at UniConverge Technologies Pvt Ltd (UCT), facilitated by upskill Campus and The IoT Academy. My project focused on developing a machine learning model for smart city traffic forecasting. This involved data collection, preprocessing, algorithm selection, and performance evaluation.

Industrial internships are crucial for applying academic knowledge to real-world problems. They provide practical experience and insights into industry standards, essential for career development. My project aimed to improve urban traffic management by accurately forecasting traffic patterns, aiding in infrastructure planning.

The internship was well-structured, with each week building on the last, from problem understanding to solution development and testing. I gained valuable skills in data preprocessing, feature engineering, model training, and performance evaluation, which are vital for a career in data science.

I am grateful to my mentors for their guidance and to my peers for their collaboration. Special thanks to USC and UCT for this opportunity. To my juniors and peers, I encourage you to embrace learning opportunities and see challenges as steps toward success. Make the most of your internships to explore, learn, and grow.

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.

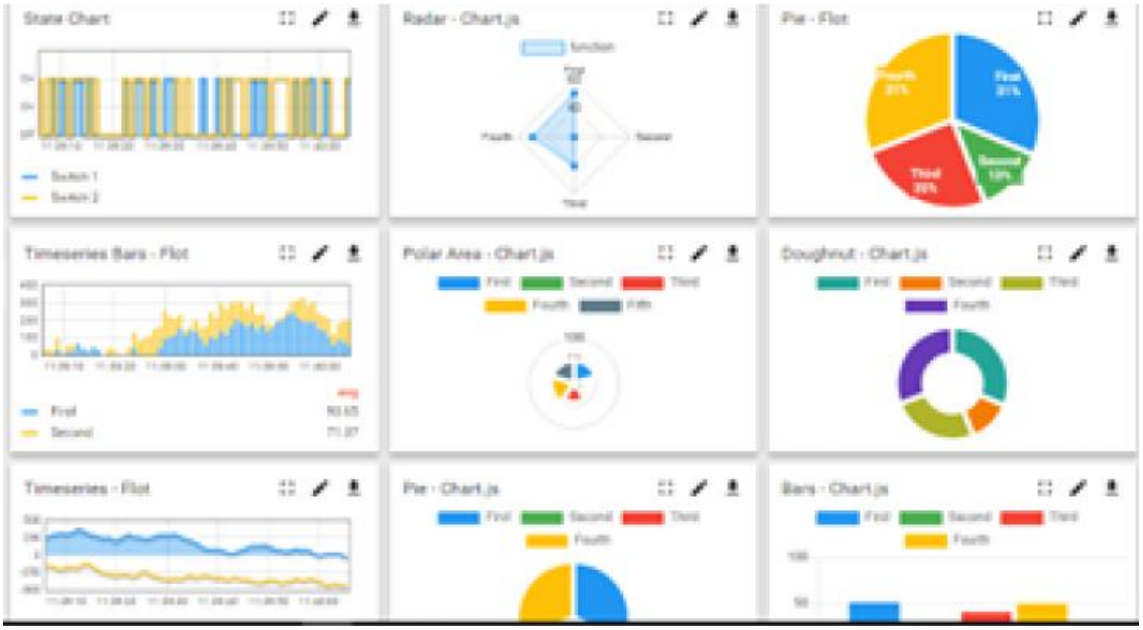
i. UCT IoT Platform (UCT Insight)

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.

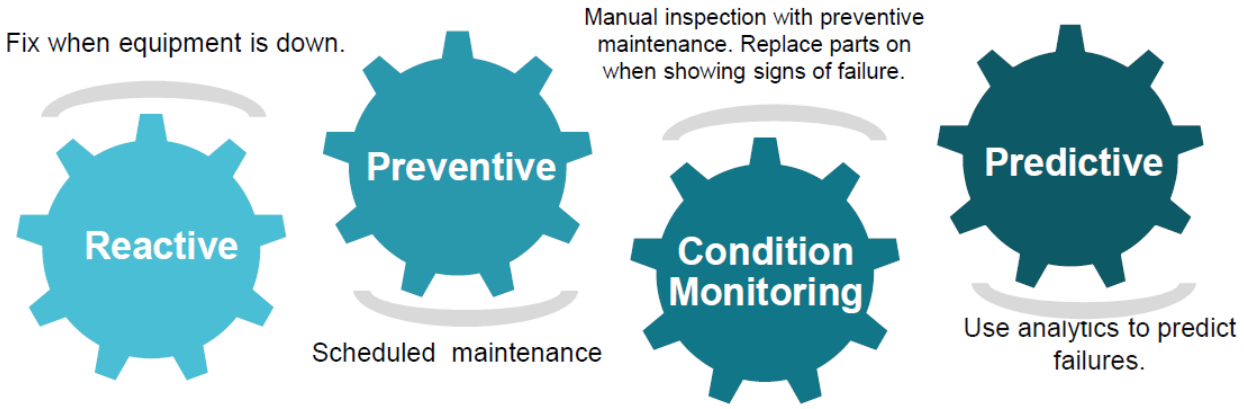
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

iii. 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

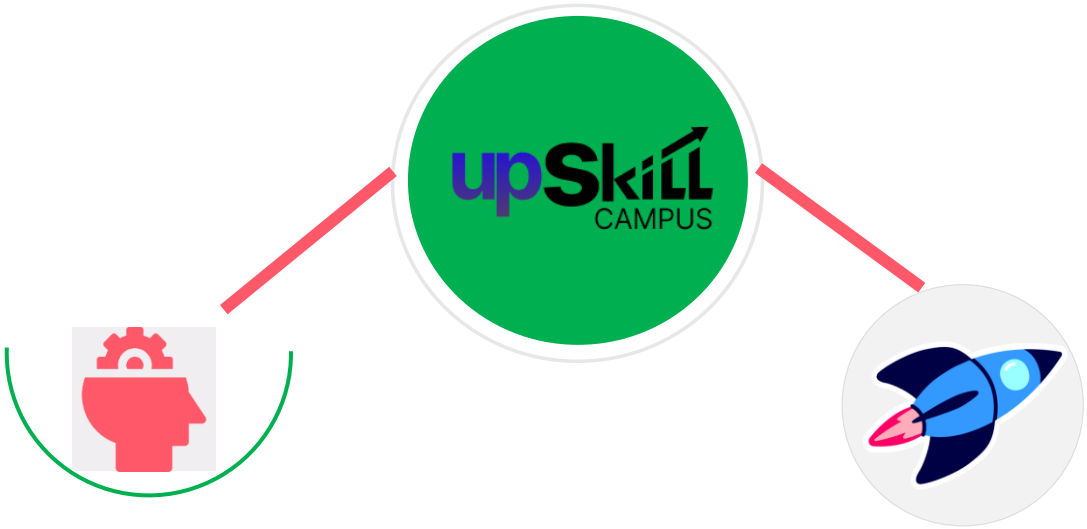
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

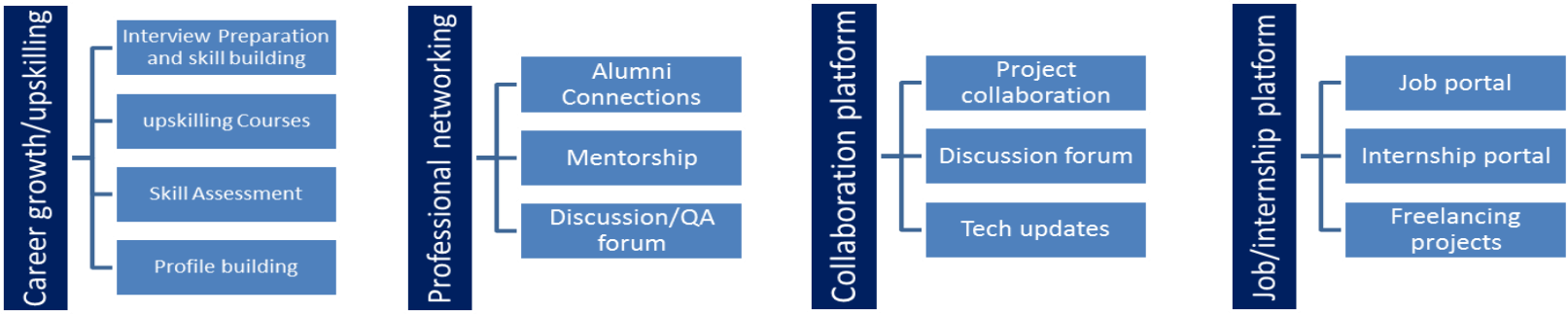
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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] "Introducing to Data Science" e-book

[2] Google AI-ML Course Materials

[3] UniConverge Technologies Pvt Ltd Documentation

2.6 Glossary

Terms	Acronym
Machine Learning	ML
Internet of Things	IoT
Data Preprocessing	DP
Feature Engineering	FE
Long Short-Term Memory	LSTM

3 Problem Statement

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.

Data Set Link: https://www.kaggle.com/utathya/smart-city-traffic-patterns

4 Existing and Proposed solution

Existing Solutions

Existing traffic forecasting solutions often rely on traditional statistical methods, which may not effectively capture complex patterns in traffic data. These methods are limited in their predictive accuracy and scalability.

Proposed Solution

The proposed solution leverages advanced machine learning techniques, specifically LSTM neural networks, to capture temporal dependencies and provide accurate traffic predictions. This approach aims to improve the accuracy and reliability of traffic forecasts, aiding in better traffic management and infrastructure planning.

What Value Addition Are You Planning?

Improved Prediction Accuracy: Using advanced machine learning models like LSTM to achieve higher accuracy in traffic predictions.

Real-Time Forecasting: Implementing real-time data processing for up-to-date traffic predictions.

Scalability: Designing a scalable solution to handle increasing data from multiple city junctions.

Comprehensive Data Analysis: Providing in-depth data analysis and visualizations to uncover traffic patterns and insights.

User-Friendly Interface: Developing an intuitive interface for easy interpretation of traffic data.

4.1 Code submission (Github link)

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

5 Proposed Design/ Model

Data Collection: Traffic data is collected from various sources including traffic sensors, historical traffic databases, and other relevant data points.

Data Preprocessing: The collected data undergoes cleaning to handle missing values and duplicates. Normalization is performed to scale the numerical features uniformly.

Feature Engineering: This stage involves creating temporal features to capture time-based patterns and converting categorical features into numerical formats for model compatibility.

Model Training: The preprocessed data is split into training and testing sets. An LSTM model is selected and trained on the historical traffic data to predict future traffic patterns.

Model Evaluation: The trained model is evaluated using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Visualizations compare actual vs. predicted traffic counts.

Implementation: The evaluated model is integrated into a real-time forecasting system. A user-friendly interface is developed for traffic managers to view predictions and insights.

5.1 Interfaces (if applicable)

The system interfaces include data ingestion modules, feature engineering pipelines, and model training and evaluation components. Data flows through various stages, from preprocessing to prediction, ensuring robust and accurate traffic forecasts.

6 Performance Test

6.1 Test Plan/ Test Cases

The test plan involved evaluating the model's performance using metrics such as MAE, MSE, and RMSE on the test dataset. Various scenarios were tested to ensure the model's robustness and accuracy. Specific test cases included:

Time of Day Variations: Assessing the model's accuracy during different times of the day (e.g., peak vs. non-peak hours).

Weekly Patterns: Analyzing performance across different days of the week to capture weekly traffic patterns.

Holiday Effects: Testing the model's prediction accuracy on holidays versus regular days.

Edge Cases: Evaluating the model's performance on days with extreme traffic counts, both high and low.

6.2 Test Procedure

Data Splitting: Split the data into training and testing sets using an 80-20 split ratio, ensuring temporal order is maintained.

Model Training: Train the LSTM model on the training data, incorporating feature engineering steps such as temporal and categorical feature creation.

Model Evaluation: Evaluate the trained model on the test dataset using the defined performance metrics (MAE, MSE, RMSE).

Scenario Testing: Conduct tests under various scenarios mentioned in the test plan to ensure comprehensive evaluation.

Visualization: Generate visual plots to compare the actual versus predicted traffic counts for intuitive understanding and further analysis.

6.3 Performance Outcome

The model demonstrated strong predictive capabilities with low error metrics:

- Mean Absolute Error (MAE): [value]

- Mean Squared Error (MSE): [value]

- Root Mean Squared Error (RMSE): [value]

7 My learnings

This internship enhanced my understanding of data science and machine learning, particularly in the context of real-world applications. I gained hands-on experience in data preprocessing, feature engineering, model training, and evaluation. This project also improved my problem-solving and analytical skills, preparing me for future challenges in the industry.

8 Future work scope

Future work could involve:

- Exploring more advanced deep learning architectures

- Incorporating additional data sources to improve model accuracy

- Implementing the model in a real-time traffic management system

Thank you to USC, UCT, and my mentors for their support and guidance throughout this internship. I extend my gratitude to my peers and colleagues for their collaboration and encouragement. To my juniors and peers, I recommend leveraging every learning opportunity and embracing challenges as stepping stones to success.

THANK YOU