Industrial Internship Report on

"Quality Prediction in a Mining Process"

Prepared by

Mahesh Ketam

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 on predictive modeling using various machine learning algorithms. The primary goal was to develop regression models, including Linear Regression, Lasso, Ridge, and Decision Tree Regressor, to predict outcomes based on a dataset. Exploratory Data Analysis (EDA) was also conducted to understand the data and its features. Through the process, various evaluation metrics were computed to assess model performance, such as Mean Squared Error (MSE). The final model performance was visualized using scatterplots and other visual aids.

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

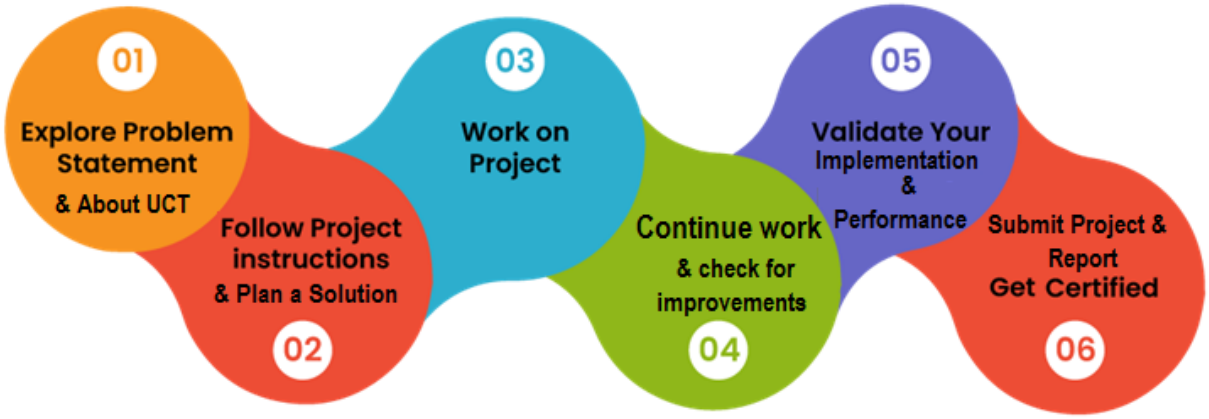
Summary of the whole 6 weeks' work.

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 , who have helped you directly or indirectly.

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.

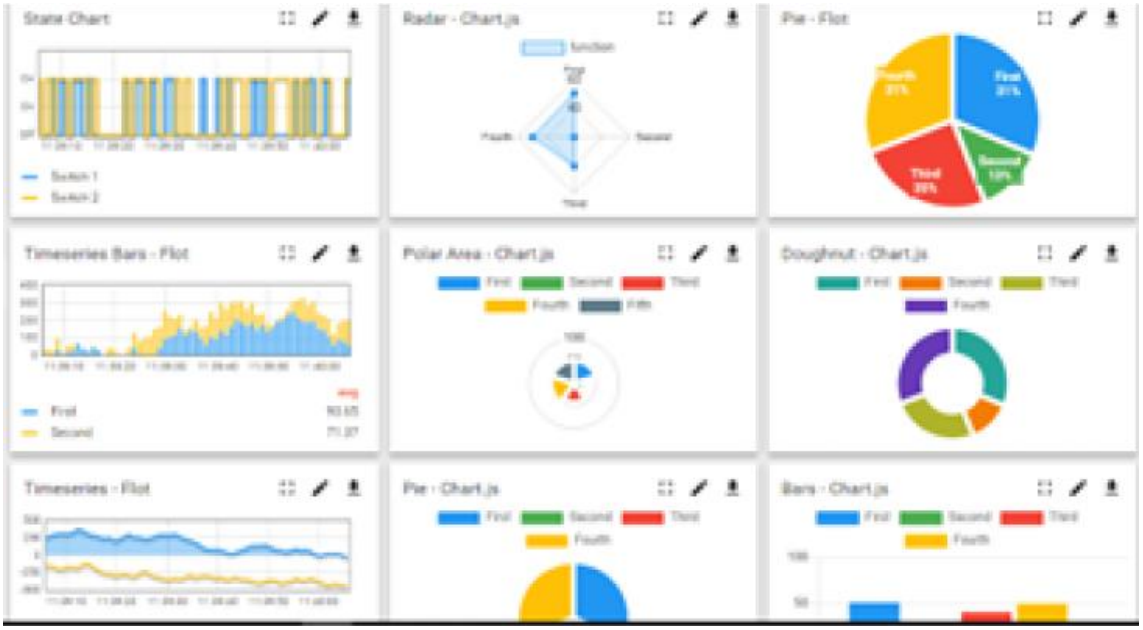
i. UCT IoT Platform (UCT Insight)

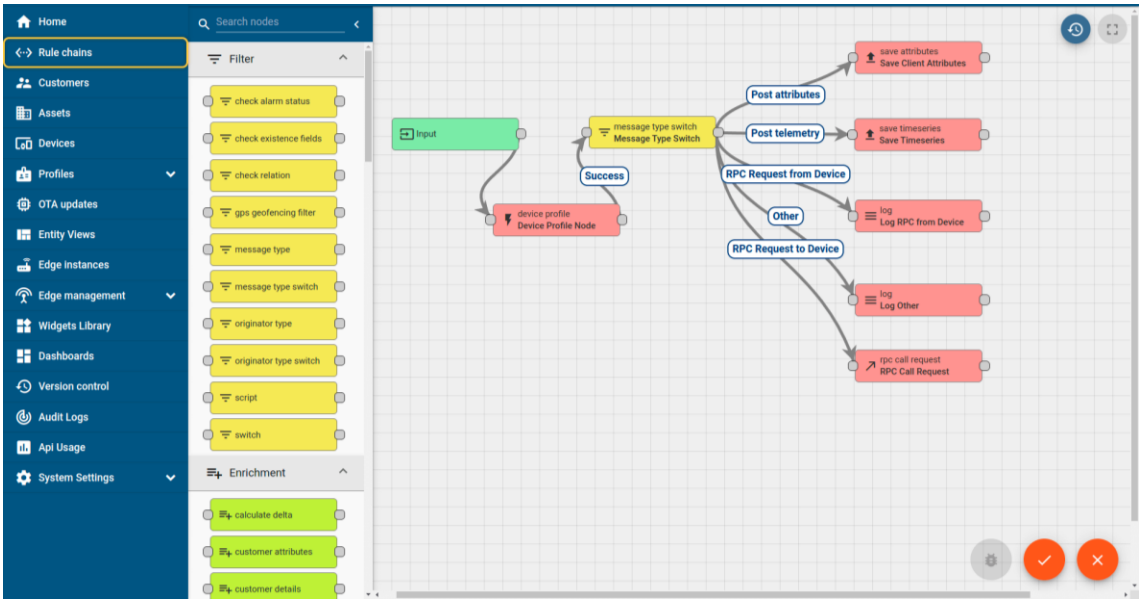
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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.

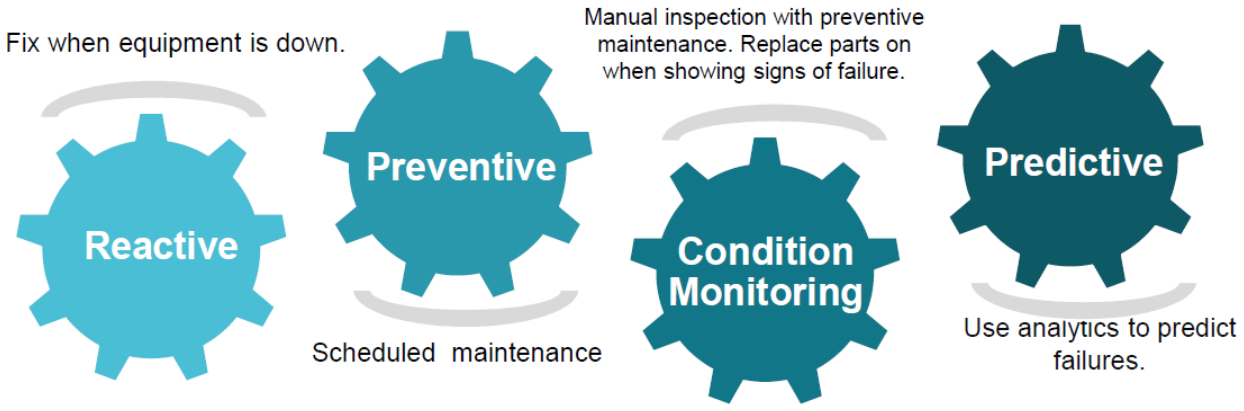
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	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 teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

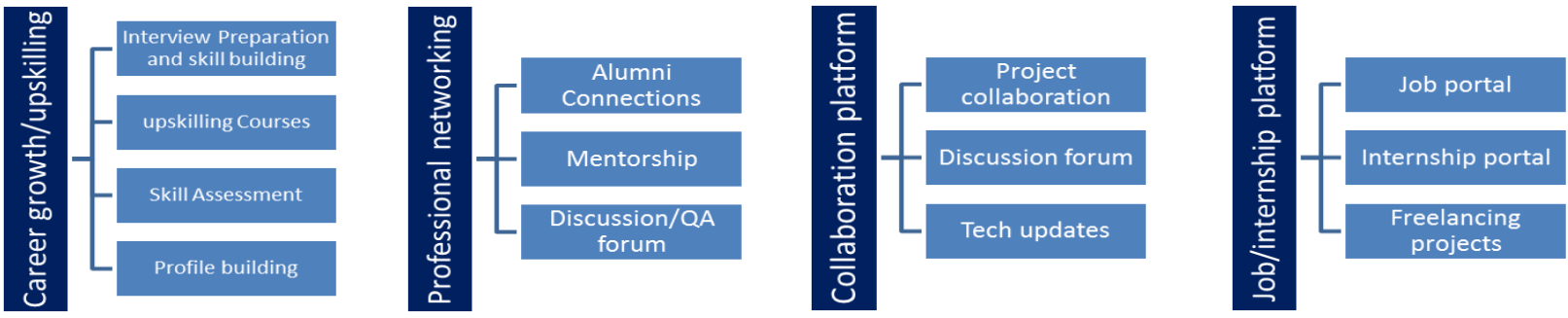
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.



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.

3 Problem Statement

In the assigned problem statement

The project aimed to predict certain outcomes (dependent variable) using various independent variables in the dataset. Different regression algorithms were tested to find the model with the best performance in terms of prediction accuracy and efficiency.

3.1 Code submission (Nirikshan95/Quality-Prediction-in-a-Mining-Process: This project predicts silica levels in iron ore concentrate using real-world data from a mining plant . The goal is to help engineers optimize plant operations and reduce environmental impact. (github.com))

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

4 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 final outcome.

- **Data Collection:** Data was collected and preprocessed by handling missing values, normalizing features, and splitting the dataset into training and testing sets.
- **Algorithms Used:**
    - **Linear Regression:** A basic model that assumes a linear relationship between the input variables and the output.
    - **Lasso Regression:** A regularized regression model that adds a penalty term to reduce overfitting.
    - **Ridge Regression:** Another regularized version that helps prevent overfitting while keeping all features.
    - **Decision Tree Regressor:** A non-linear algorithm used for capturing complex relationships in the data.

5 Exploratory Data Analysis (EDA)

EDA was performed to understand the distribution of the data, identify outliers, and discover relationships between variables. Various plots such as histograms, scatterplots, and correlation matrices were used for data visualization.

6 Performance Test

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

- **Evaluation Metrics:**
  The models were evaluated using **Mean Squared Error (MSE)**, **R-squared**, and other relevant metrics. The MSE for each model was calculated and compared.
    - Linear Regression MSE: 0.404
    - Lasso Regression MSE: 0.417
    - Ridge Regression MSE: 0.242
    - Decision Tree Regressor MSE: 0.007

The Decision Tree Regressor achieved an MSE of 1.21, which indicates the error in the predictions. This value could be considered reasonable depending on the scale of the data.

7 My learnings

During this internship, I gained experience with multiple machine learning algorithms, data preprocessing, and model evaluation. I also developed a better understanding of how to use Python libraries like scikit-learn for implementing machine learning models and matplotlib/seaborn for visualization and more .

8 Future work scope

The project could be extended by exploring more advanced algorithms, such as Random Forest or Gradient Boosting, and implementing hyperparameter tuning to improve model performance. Additionally, expanding the dataset and including more features could lead to better results.