Industrial Internship Report on

”Quality Prediction in Mining Process”

Prepared by

Aman Kumar Emaley

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 (Tell about ur Project) 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

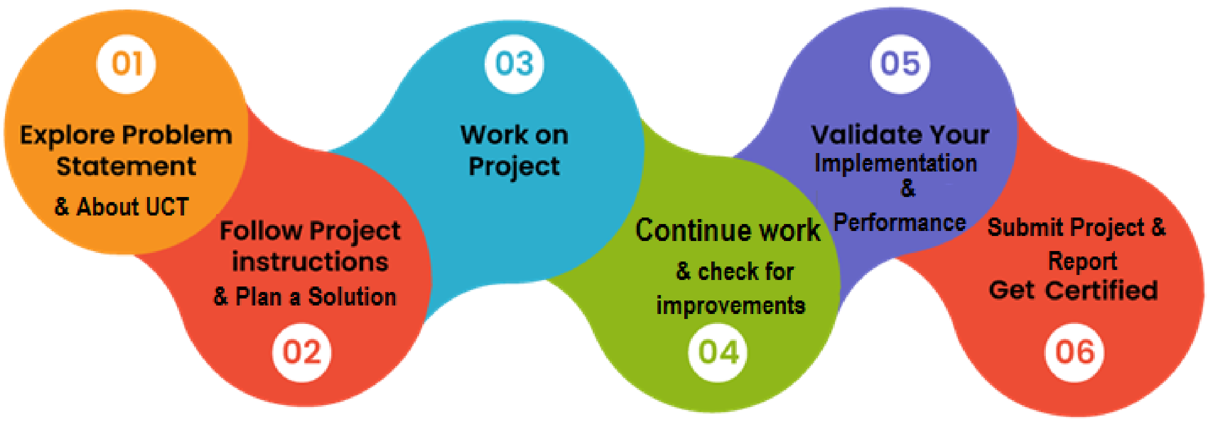
Summary of the whole 6 weeks' work.

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)

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

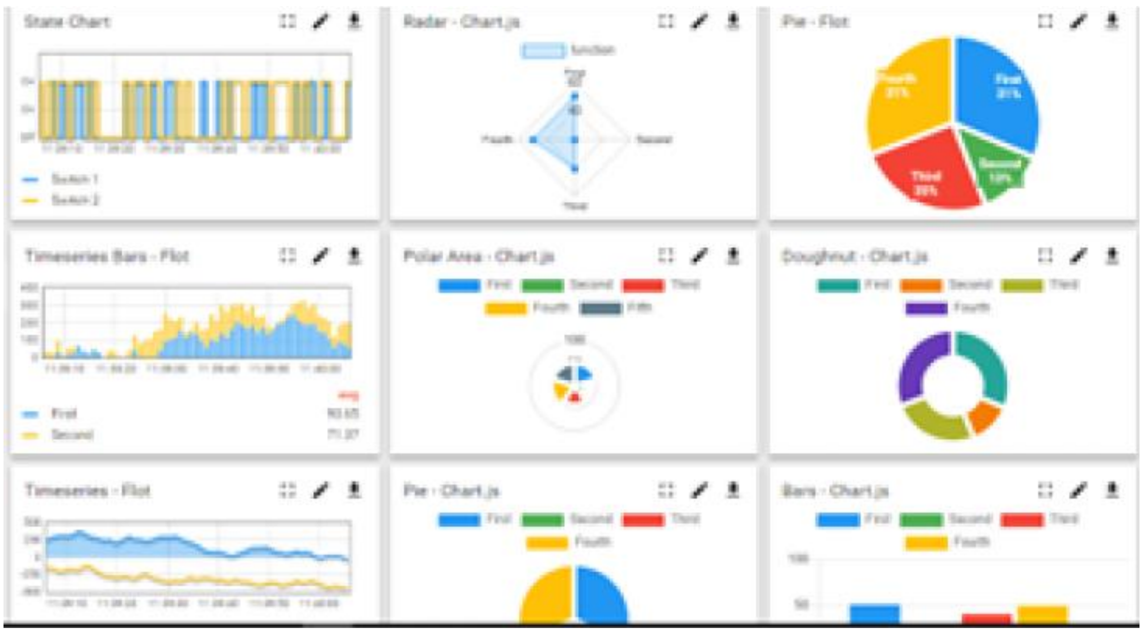
• Build Your own dashboard

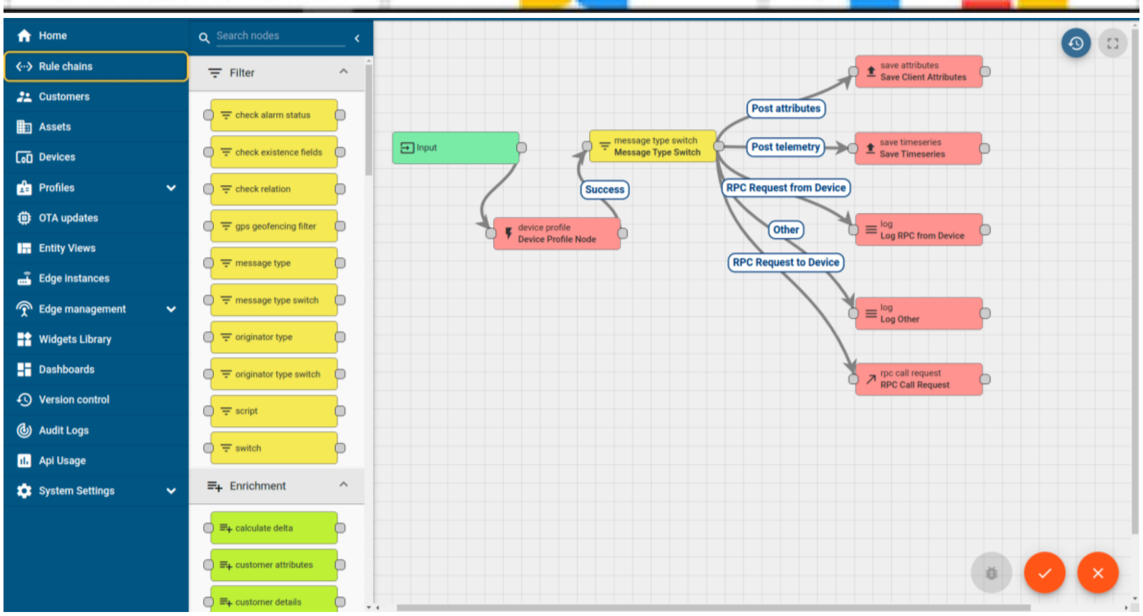
• Analytics and Reporting

• 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.

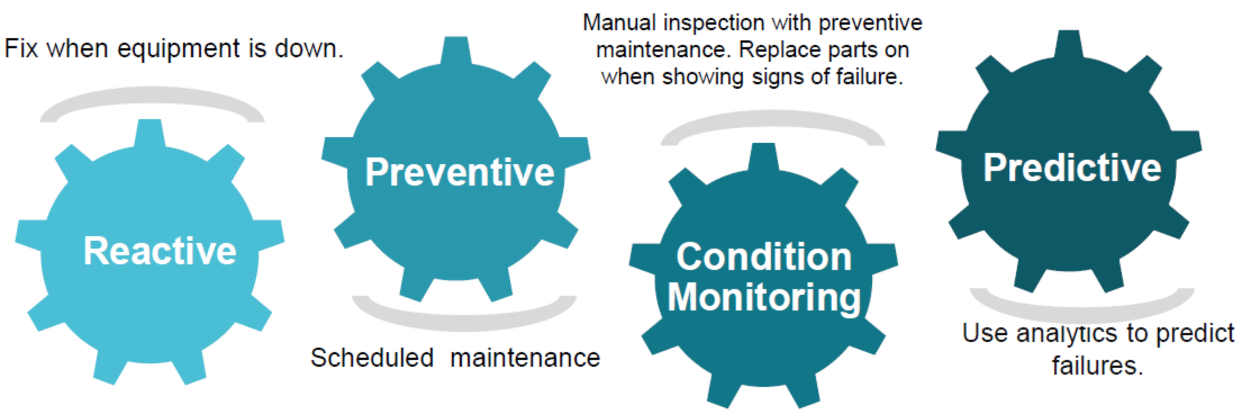
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

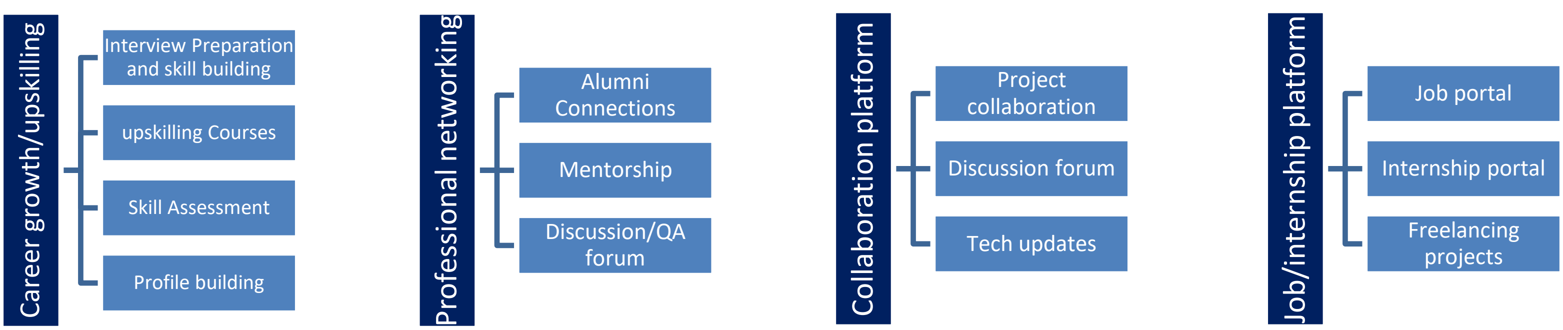
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

2.5 Reference

[1] https://www.researchgate.net/publication/237086918_A_Data_Mining_Approach_for_Developing_Quality_Prediction_Model_in_Multi-Stage_Manufacturing.

[2] https://www.sciencedirect.com/science/article/pii/S1474034620300707.

[3] https://www.kaggle.com/datasets/edumagalhaes/quality-prediction-in-a-mining-process

3 Problem Statement

Quality Prediction In Mining Process

Explore real industrial data and help manufacturing plants to be more efficient

About Dataset

3.1.1 Context

It is not always easy to find databases from **real world** manufacturing plants, specially mining plants. So, I would like to share this database with the community, which comes from one of the most important parts of a mining process: a flotation plant!.

The main goal is to use this data to predict how much impurity is in the ore concentrate. As this impurity is measured every hour, if we can predict how much silica (impurity) is in the ore concentrate, we can help the engineers, giving them early information to take actions (empowering!). Hence, they will be able to take corrective actions in advance (reduce impurity, if it is the case) and also help the environment (reducing the amount of ore that goes to tailings as you reduce silica in the ore concentrate).

3.1.2 Content

The first column shows time and date range (from march of 2017 until september of 2017). Some columns were sampled every 20 second. Others were sampled on a hourly base.

The second and third columns are quality measures of the iron ore pulp right before it is fed into the flotation plant. Column 4 until column 8 are the most important variables that impact in the ore quality in the end of the process. From column 9 until column 22, we can see process data (level and air flow inside the flotation columns, which also impact in ore quality. The last two columns are the final iron ore pulp quality measurement from the lab.Target is to predict the last column, which is the % of silica in the iron ore concentrate.

4 Existing and Proposed solution

Existing solutions for quality prediction in mining processes often involve the application of machine learning and statistical modeling techniques. Some common methods used include:

Regression Models: Linear regression and its variants are commonly employed to predict the quality of the mining process based on input features such as temperature, ore composition, and other process parameters.

Classification Models: Classification algorithms, such as decision trees, random forests, and support vector machines, can be used to classify the quality of the mining process into discrete categories (e.g., high, medium, low) based on input features.

Time Series Analysis: If the dataset contains temporal information, time series analysis techniques like ARIMA or LSTM (Long Short-Term Memory) networks can be used to capture temporal dependencies and predict the quality of future mining processes.

Limitations

- **Limitations of existing solutions** can vary depending on the specific methods employed and the characteristics of the dataset. Some common limitations include:
- **Limited Feature Set:** If the dataset lacks important features or if relevant information is missing, it can impact the accuracy and reliability of the predictions.
- **Data Quality**: Poor data quality, such as missing values, outliers, or inconsistencies, can affect the performance of prediction models.
- **Lack of Domain Knowledge:** Without a deep understanding of the mining process, it can be challenging to select appropriate features, interpret the results, or address specific domain-related challenges.
- **Scalability:** Some models may not scale well to large datasets or real-time prediction requirements, which can limit their practical application in mining operations.

Proposed solution

The proposed solution for the "Quality Prediction in a Mining Process" involves several steps. First, we process dataset needs to be explored and preprocessed, addressing any missing values, outliers, or data quality issues. Feature engineering techniques can be applied to create new features based on domain knowledge and the mining process. Next, suitable machine learning models are selected, such as regression or classification algorithms, and trained on the dataset. Next, hyperparameter tuning was performed using GridsearchCV and lastly the models were finalized and validated with the best estimator previously discovered during the tuning phase. The results of the models are interpreted and analyzed to gain insights into the factors influencing the quality prediction.

Value addition

In the future, the "Quality Prediction in a Mining Process" project can be enhanced in several technical aspects. the project can be extended to incorporate deep learning approaches, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), to capture intricate patterns in the data. Real-time prediction capabilities can be developed by establishing streaming data pipelines, integrating with real-time data sources, and deploying scalable models. Anomaly detection techniques can be integrated to identify deviations or abnormalities in the mining process. Predictive maintenance methods can be explored to predict equipment failures or maintenance requirements. Additionally, interactive visualizations and reporting dashboards can be created to provide stakeholders with intuitive insights, while deployment and integration considerations can ensure the models are efficiently integrated into existing mining operations. These technical advancements will contribute to more accurate quality predictions, operational efficiency, and informed decision-making in the mining process.

4.1 Code submission (Github link)

https://github.com/Akcount007/projects.git

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

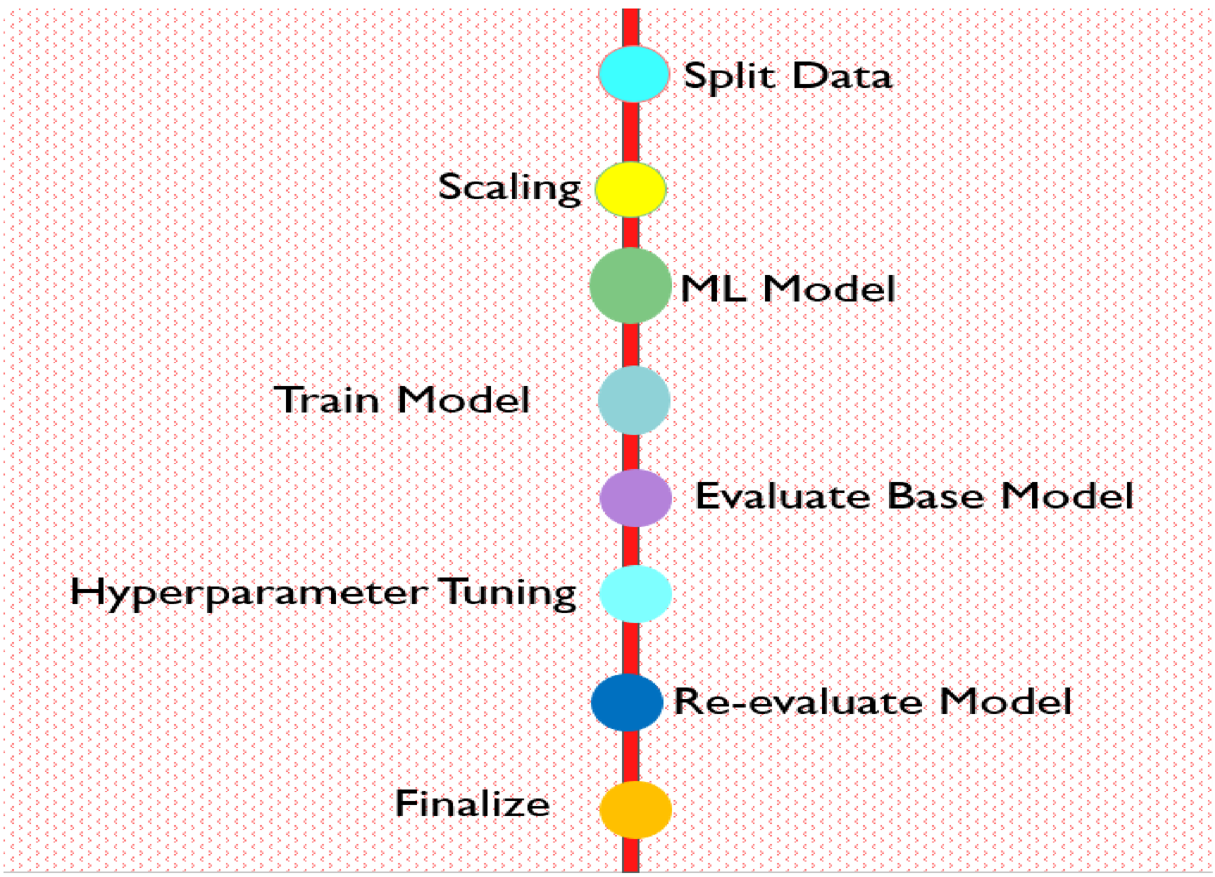
https://github.com/Akcount007/projects.git

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 final outcome.

5.1 Interfaces

FLOW Chart



6 Performance Test

We started off by splitting the data into training and testing sets. We then built a pipeline that consisted of a transformer and an estimator. Next, we used cross validation on the training data in order to determine which regression algorithm performs the best on out of sample data. The best respective model for each approach was then chosen to build the regression model. We then evaluated the models with the testing data in order to get a baseline model. Next, hyperparameter tuning was performed using GridsearchCV and lastly the models were finalized and validated with the best estimator previously discovered during the tuning phase.

6.1 Test Plan/ Test Cases

we evaluated the models with the testing data in order to get a baseline model. Then hyperparameter tuning was performed using GridsearchCV and lastly the models were finalized and validated with the best estimator previously discovered during the tuning phase. After using various algorithm's, we documented their performance using the RMSE score, R^2, and accuracy.

6.2 Performance Outcome

After using various algorithm's, we documented their performance using the RMSE score, R^2, and accuracy. We found out that the Gradient Boosting Regressor performed better on the resampled dataset that aggregated the values every hour from the original dataset. We were able to predict the % Silica Concentrate with an accuracy of 93.80% on the large original dataset and an accuracy of 62.89% on the small resampled dataset.

7 My learnings

While working on this project we learned several valuable lessons. We learn more about different regression model while using in this project. Effective communication was key to ensure that everyone understood the project goals and tasks, enabling us to work cohesively towards a common objective. We realized the importance of dividing tasks based on individual strengths and expertise, allowing team members to contribute their unique skills to different aspects of the project. Collaboration and knowledge sharing fostered a creative and innovative environment, where diverse perspectives and ideas were encouraged. Time management was crucial to meet project milestones and deadlines, and regular progress updates ensured that everyone stayed on track. Working as a team enhanced our problem-solving abilities,taught us the value of cooperation, and reinforced the significance of effective teamwork in achieving project success.

8 Future work scope

From this Project We able to Predict the how much impurity is in the ore concentrate. As this impurity is measured every hour, if we can predict how much silica (impurity) is in the ore concentrate, we can help the engineers, giving them early information to take actions (empowering!). Hence, they will be able to take corrective actions in advance (reduce impurity, ifit is the case) and also help the environment (reducing the amount of ore that goes to tailings as you reduce silica in the ore concentrate).