

Industrial Internship Report on Multi-Stage Continuous Flow Manufacturing Process

Prepared by

Mitali Saini

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 Multi-Stage Continuous Flow Manufacturing Process, which predicts the factory output.

This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solutions for that. It was an overall great experience to have this internship.

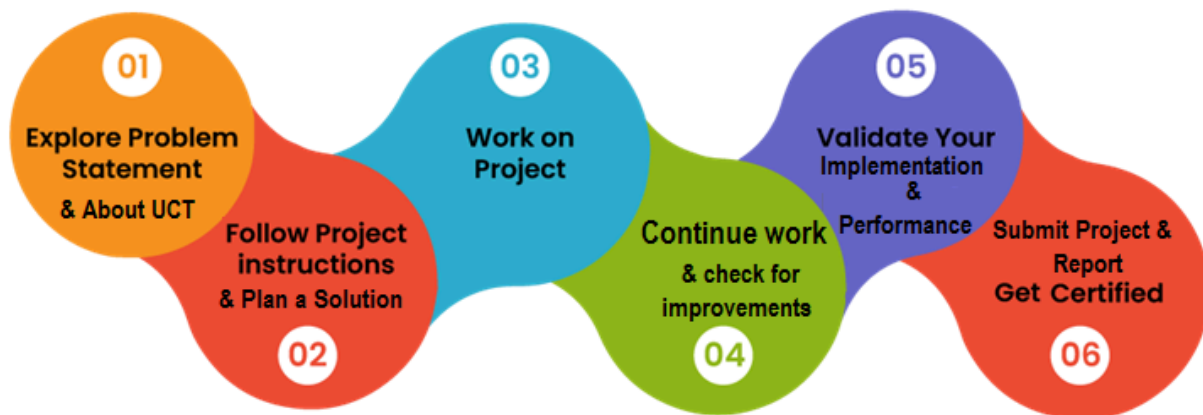
TABLE OF CONTENTS

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
3	Problem Statement.....	11
4	Existing and Proposed solution.....	12
5	Proposed Design/ Model.....	13
5.1	High Level Diagram (if applicable).....	14
5.2	Boxplot.....	15
6	Performance Test.....	16
6.1	Test Plan/ Test Cases.....	16
6.2	Test Procedure.....	17
6.3	Performance Outcome.....	17
7	My learnings.....	18
8	Future work scope.....	19

1 Preface

This project report encapsulates my enriching experience of a 6-week internship, a milestone in my academic and professional journey. I firmly believe that relevant internships are indispensable in career development, offering a platform to apply theoretical knowledge in real-world settings, cultivate industry specific skills and other networking opportunities.

The project undertaken at upskills campus, aimed to predict the factory output with better accuracy. This opportunity not only allowed me to work on a challenging problem but also provided me with invaluable hands-on experience.



Throughout this internship, I acquired a profound understanding of Data Science, Machine Learning, Predictive Modeling, Probability, Big Data Analytics, etc. The experience was truly enriching and I am grateful for the opportunity to work with the upskills campus and its team who helped me get the idea about the project through the video contents and ebooks.

I extend my sincere appreciation to all those who have supported me directly or indirectly throughout this journey.

As I reflect on my experience, I would like to advise my juniors and peers that internships offer a unique chance to explore your interests, develop your skills, and build your professional network. I encourage you to be proactive, seek out opportunities and make the most of your experiences.

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in the Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and Role.

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.



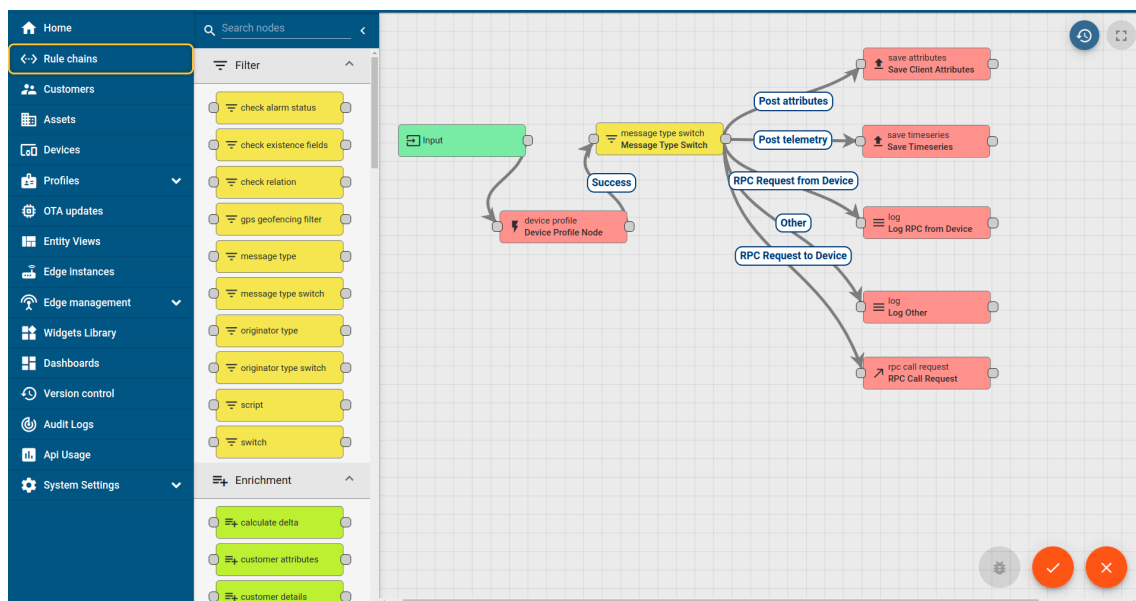
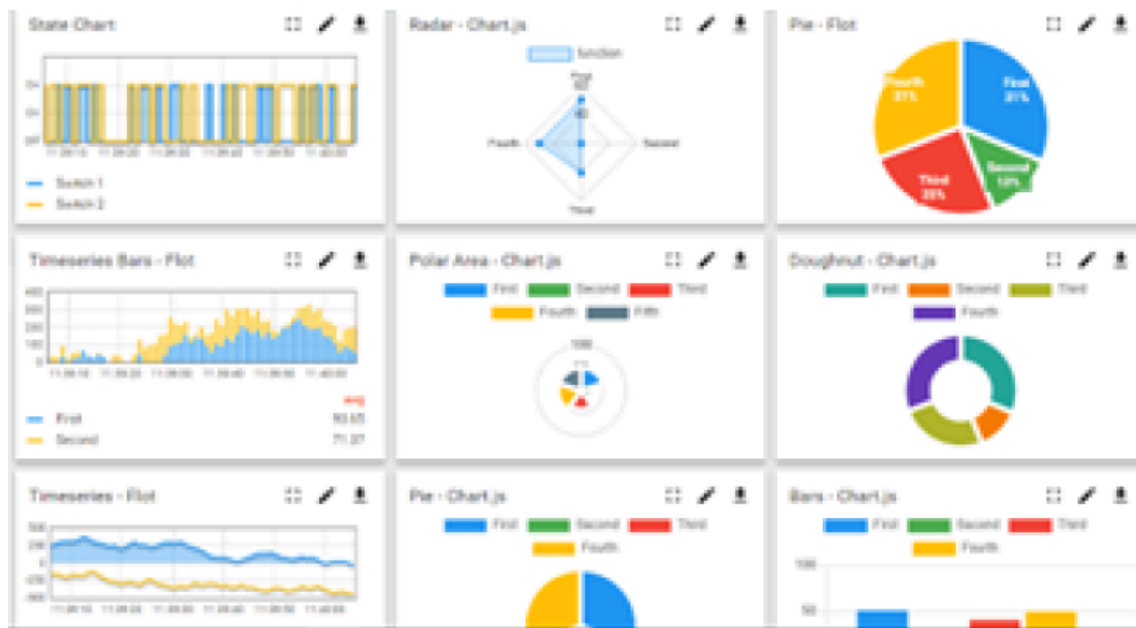
i. UCT IoT Platform ()

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.



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



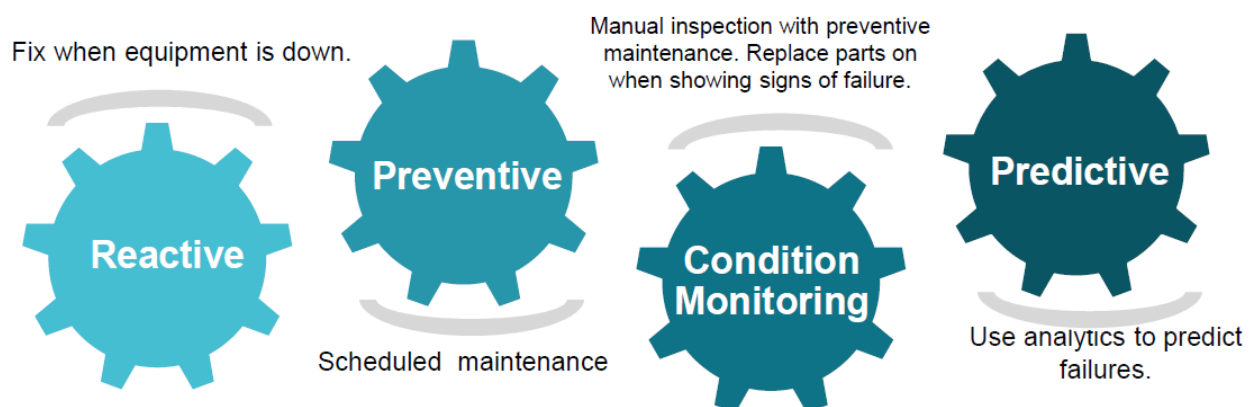


iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN technology and provides solutions 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.



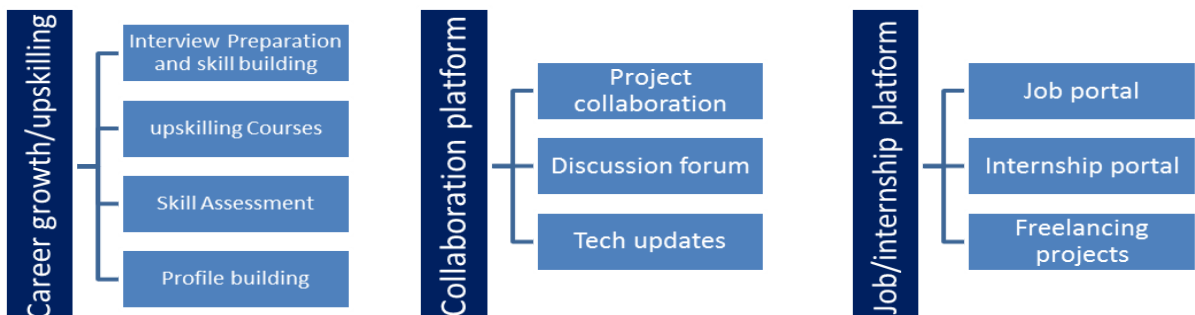
Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services



upSkill Campus aiming to upskill 1 million learners in next 5 year

<https://www.upskillcampus.com>

7



2.3 The IoT Academy

The IoT academy is the 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] www.techsaksham.org

3 Problem Statement

This data was taken from an actual production run spanning several hours. The goal is to predict certain properties of the line's output from the various input data, to optimize the performance of a multi-stage continuous flow manufacturing process to improve efficiency, quality and sustainability. The line is a high-speed, continuous manufacturing process with parallel and series stages.

The data comes from a continuous flow manufacturing process with multiple stages. Sample rates are 1 Hz.

- In the first stage, Machines 1, 2, and 3 operate in parallel, and feed their outputs into a step that combines the flows.
- Output from the combiner is measured in 15 locations. These measurements are the primary measurements to predict.
- Next, the output flows into a second stage, where Machines 4 and 5 process in series.
- Measurements are made again in the same 15 locations. These are the secondary measurements to predict.

4 Existing and Proposed solution

Maintaining consistent product quality and process stability across multiple stages especially in the presence of disturbances. Ensuring the reliability and availability of equipment to maintain continuous flow and minimize downtime, could be the existing solutions provided by others.

I analysed some of the previous problems like this and proposed a solution by utilizing some of the machine learning and ensemble learning techniques to enhance the accuracy score and develop a model.

I plan to add value to the project by leveraging my expertise in data analysis to improve efficiency. Specifically, I aim to develop and implement a model to track key performance indicators within the first 6 weeks. I took a target variable with size 80% and got 100% accuracy by utilizing Decision Tree Regressor.

4.1 Code submission (Github link):

<https://github.com/mitalis-03/upskillcampus/blob/main/MultiStageContinuousFlowProcess.ipynb>

4.2 Report submission (Github link) :

5 Proposed Design/ Model

The proposed design for this project:

Step1: Data Collection:

Data Collected by utilizing the dataset in the given link.

Step2: Data Preprocessing:

Check for the missing values.

Analyze the data and make graphs, heatmap, etc.

Divided the dataset into 2 sub-datasets, one for machines 1, 2 and 3, and the other for machines 4

and 5.

Step3: Machine Learning Model

(i) Model Selection:

Supervised learning (e.g., logistic regression, decision trees, random forests)

Ensemble methods (e.g., random forest, adaboost, bagging)

(ii) Model Training:

Split data into training (80%) and testing sets (20%)

Train model using training data

(iii) Model Evaluation:

Evaluate model performance using testing data

Metrics: accuracy, r2 score, mean squared error, mean absolute error, etc.

5.1 High Level Diagram (if applicable)

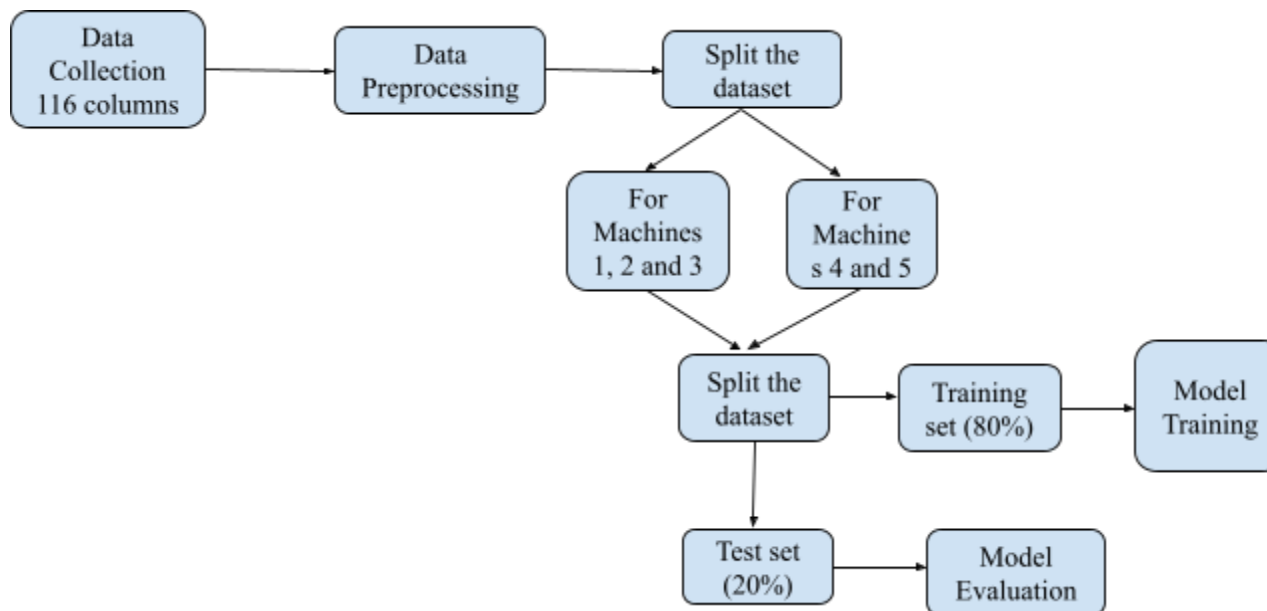
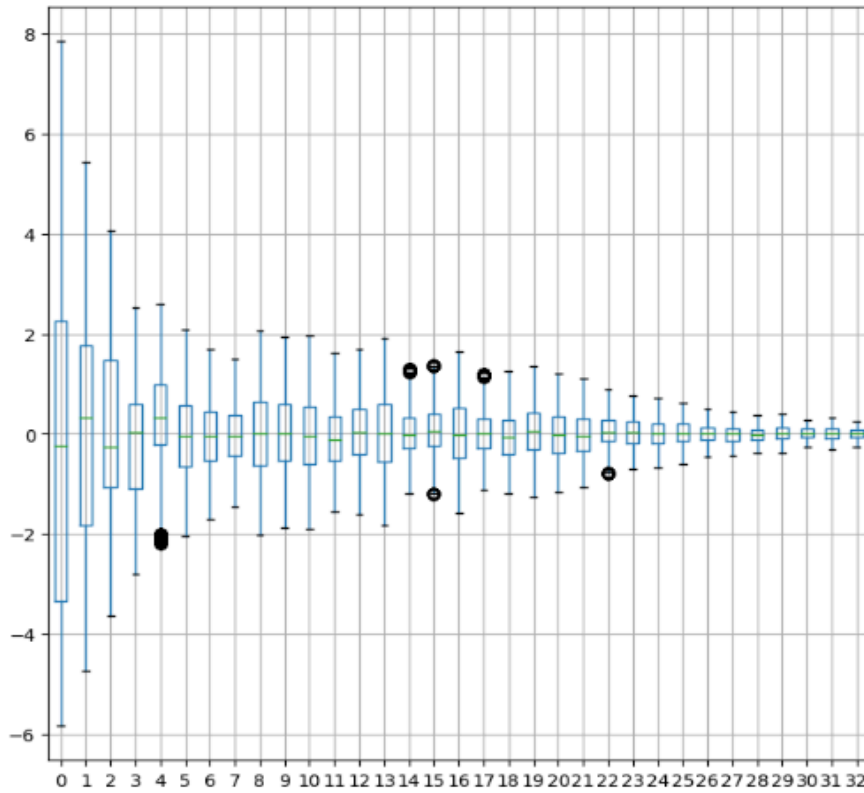


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

5.2 Boxplot



without any outlier

Figure 2: BOXPLOT OF THE FINAL DATASET

6 Performance Test

This project is not only useful for academics, but also for real industries as it helps in enhancing the accuracy so that we can develop a model that gives the most accurate results. Initially, the collected dataset contained 116 attributes, by applying PCA feature selection, the attributes reduced to 33. It includes 14088 rows. In this project, machine learning techniques such as K Nearest Neighbors, DecisionTree Regressor and ensemble learning techniques such as GradientBoostingRegressor, RandomForestRegressor and AdaBoostRegressor and feature selection technique such as Principal Component Analysis(PCA) have been utilized.

The constraints involve raw materials, zone temperature, material pressure, material temperature, motor amperage, motor rpm, ambient conditions such as temperature and humidity, etc.

These constraints play a major role in designing our model as with greater accuracy, they can predict the factory output of what equipment should be chosen based on its capacity of temperature, pressure, humidity and some other factors.

6.1 Test Plan/ Test Cases

- **Functional Test Cases**

1. Data Ingestion: Test that the system can ingest data from various sources such as CSV.
2. Data Preprocessing: Test that the system can preprocess data such as handle missing values, normalize data. Verify that the system can handle different preprocessing techniques and parameters.
3. Model Training: Test that the system can train ML models (linear regression, decision trees, etc).
4. Model Evaluation: Test that the system can evaluate ML models (calculate accuracy, r2 score, etc) and verify that the system can handle different evaluation metrics and techniques.

- **Performance Test Cases**

1. Training Time: Test that the system can train ML models within a reasonable time frame and verify that the system can handle large datasets and complex models.

2. Inference Time: Test that the system can make predictions using trained ML models within a reasonable time frame and verify that the system can handle high-volume prediction requests.

3. Memory Usage: Test that the system can handle large datasets and complex models without excessive memory usage and verify that the system can handle memory-intensive tasks (e.g., data preprocessing, model training).

6.2 Test Procedure

1. Test Objective: The objective of this test procedure is to verify that the ML model meets the required specifications and functions as expected.
2. Test Scope: The test scope includes the following:
 - Data ingestion and preprocessing
 - Model training and evaluation
 - Model deployment and prediction
 - Error handling and logging
3. Test Environment: The test environment consists of:
 - Hardware: 16GB RAM, Windows 11 OS, etc.
 - Software: jupyter notebook, python libraries such as pandas, numpy, scikit-learn, etc.
4. Test Cases: The test cases used are the following mentioned above.

6.3 Performance Outcome

Decision Tree proved to be the best model as it gives 100% accuracy.

Decision Tree Regressor:

Accuracy: 100%,

rmse: 7.113699,

mae: 3.701448,

r2 score: 1.0

7 My learnings

We explored the multi-stage continuous flow manufacturing process which includes data collection, data preprocessing, model selection, model training and evaluation. We developed a test procedure to ensure the ML model meets the required specifications and functions as expected. We discussed the importance of specifying hardware configurations, including processors, memory, storage, etc. We understood the importance of accuracy in the model. With this project, we learnt to use several algorithms and to detect the errors with patience.

By working on a complex ML project, we'll develop strong problem-solving skills, which are highly valued in the industry. We'll gain expertise in the field of data science and machine learning and develop an understanding of the challenges and opportunities in that space. We'll enhance our technical skills in ML, including data preprocessing, model training, and model evaluation. By working on a comprehensive test procedure, we'll develop strong collaboration and communication skills, essential for working with cross-functional teams. We'll learn to adapt to new technologies, tools, and methodologies, making us a more flexible and versatile professional.

8 Future work scope

Feature Engineering: Some of the feature engineering techniques could not be applied such as mutual information, Boruta algorithm, etc. that would be applied in the future projects.

Explainable AI techniques: Some of the XAI techniques could not be applied such as SHAP, LIME to show the predicted results using some plots or graphs that would be applied in the future projects.

Model Deployment: Model could not be deployed. We can do it by utilizing Python's Flask or web development tools that would be done in future projects.

Future Career Scopes:

ML Engineer: With expertise in ML and domain knowledge in manufacturing, we'll be well-positioned for a career as an ML engineer.

Data Scientist: Our skills in data preprocessing, model training, and deployment will serve us well in a data scientist role.

Research and Development: You'll be well-equipped to work in research and development, exploring new applications of ML in manufacturing and other industries.