Industrial Internship Report on

" Predictive maintenance of Gearbox using vibration sensors"

Prepared by

[Pramey Tiwari]

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:- In this project, predictive maintenance of gearboxes is achieved using vibration sensors. Data from the sensors is analyzed to detect anomalies and predict failures. A Random Forest algorithm is employed for initial feature extraction and classification, while a Convolutional Neural Network (CNN) is used to enhance the accuracy and robustness of fault detection. This approach aims to reduce downtime and maintenance costs by anticipating gearbox issues before they occur. 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.

Parul University, located in Vadodara, Gujarat, India, is a multidisciplinary educational institution established in 2015. It offers a wide range of undergraduate, postgraduate, and doctoral programs across various fields such as engineering, management, arts, and sciences. The university is known for its state-of-the-art infrastructure, diverse campus community, and emphasis on innovation and research. Parul University aims to provide a holistic educational experience, fostering both academic excellence and personal growth.

TABLE OF CONTENTS

1 Preface 4

2 Introduction 6

2.1 About UniConverge Technologies Pvt Ltd 6

2.2 About upskill Campus...10

2.3 Objective 12

2.4 Reference 12

2.5 Glossary...12

3 Problem Statement...14

4 Existing and Proposed solution 16

5 Proposed Design/ Model 18

5.1 High Level Diagram (if applicable) .... **Error! Bookmark not defined.**

5.2 Low Level Diagram (if applicable) .... **Error! Bookmark not defined.**

5.3 Interfaces (if applicable)... **Error! Bookmark not defined.**

6 Performance Test 22

6.1 Test Plan/ Test Cases .... **Error! Bookmark not defined.**

6.2 Test Procedure.. **Error! Bookmark not defined.**

6.3 Performance Outcome.. **Error! Bookmark not defined.**

7 My learnings.. **Error! Bookmark not defined.**

8 Future work scope 25

1 Preface

- **Summary of the whole 6 weeks' work:**

This project explores the implementation of predictive maintenance for gearboxes utilizing vibration sensors. Traditional maintenance approaches often lead to unexpected equipment failures and costly downtime. By leveraging advanced data analysis techniques, specifically Random Forest algorithms and Convolutional Neural Networks (CNN), this project aims to accurately detect and predict gearbox faults before they escalate into significant issues. The integration of these machine learning models with vibration sensor

- **About need of relevant Internship in career development:**

Skill Development

Internships provide hands-on experience that helps you apply theoretical knowledge to real-world scenarios. This practical exposure is crucial for developing the skills needed in your chosen field.

Industry Insight

Working within a company allows you to understand the industry dynamics, work culture, and professional environment, which can be quite different from academic settings.

Networking

Internships offer opportunities to connect with professionals in your field. Building a professional network can lead to job offers, mentorship, and career advice.

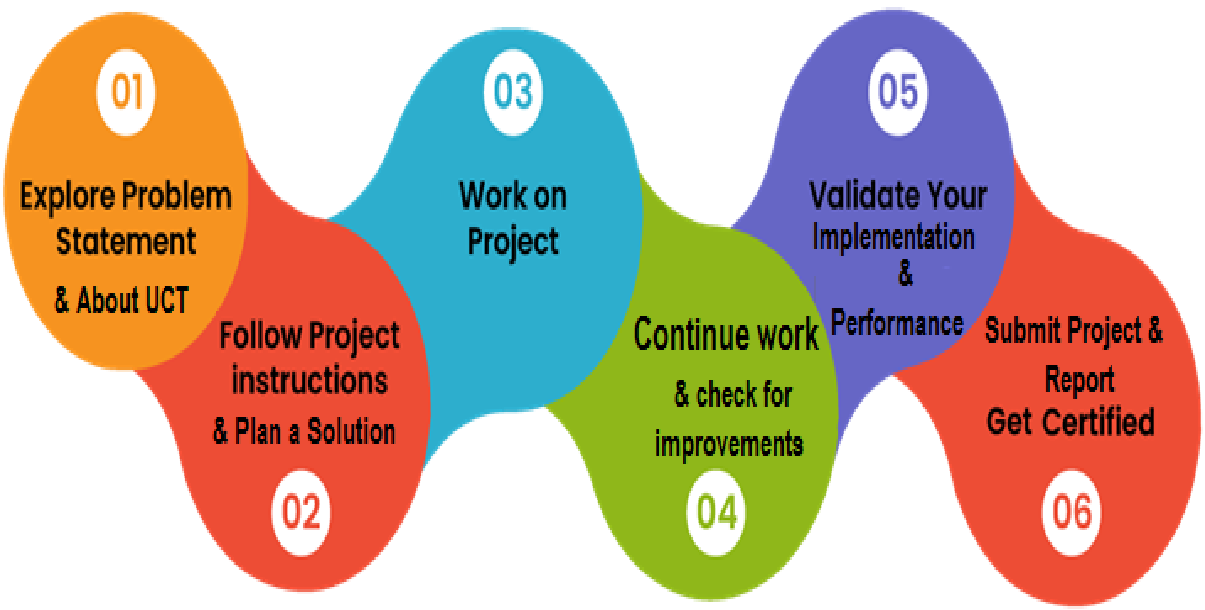
- **Brief about Your project/problem statement:**

Industrial gearboxes are critical components in many machines and processes, and unexpected failures can lead to significant downtime and financial losses. Traditional maintenance approaches, such as reactive or scheduled maintenance, are often inefficient and costly. This project aims to implement a predictive maintenance system using vibration sensors to monitor the condition of gearboxes in real-time. By analyzing vibration data, the system will detect early signs of wear or potential failures, allowing for timely maintenance interventions.

- **Opportunity given by USC/UCT:**

It sounds like UpSkill Campus or Uniconvergence Private Limited might be offering some opportunities for upskilling or education. These companies likely provide training programs, courses, or resources aimed at enhancing skills or knowledge in various fields. Taking advantage of such opportunities can be beneficial for career advancement, staying competitive in the job market, or simply personal growth. If you're considering it, I'd suggest researching the specific programs they offer, their reputation, and reviews from past participants to ensure they align with your goals and expectations.

- **How Program was planned**



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.

The picture can't be displayed.

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

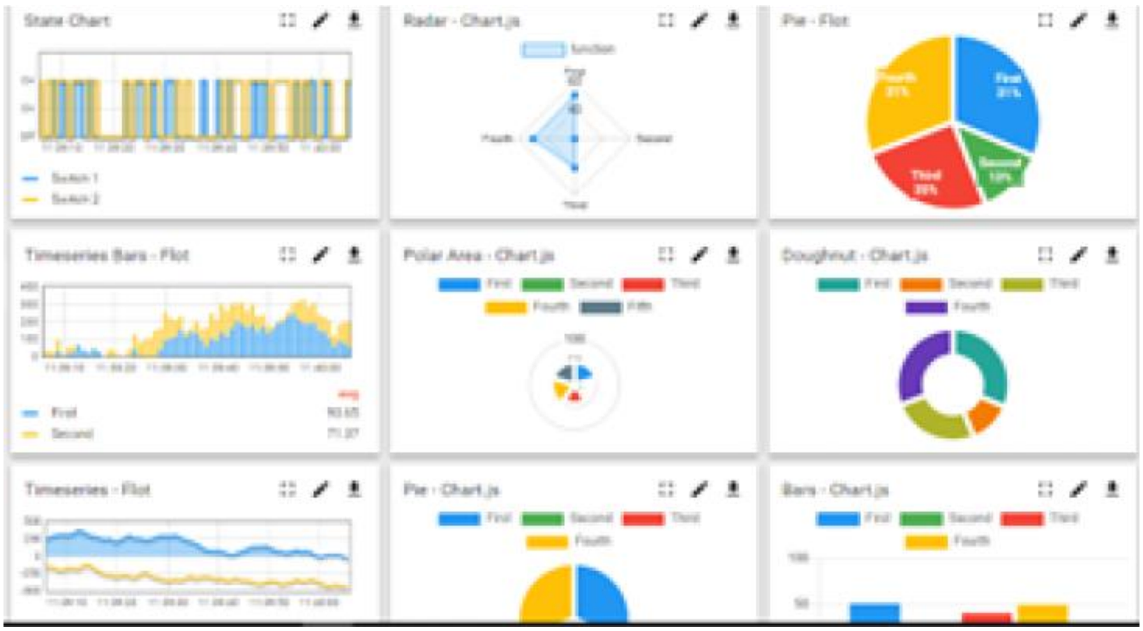
• Build Your own dashboard

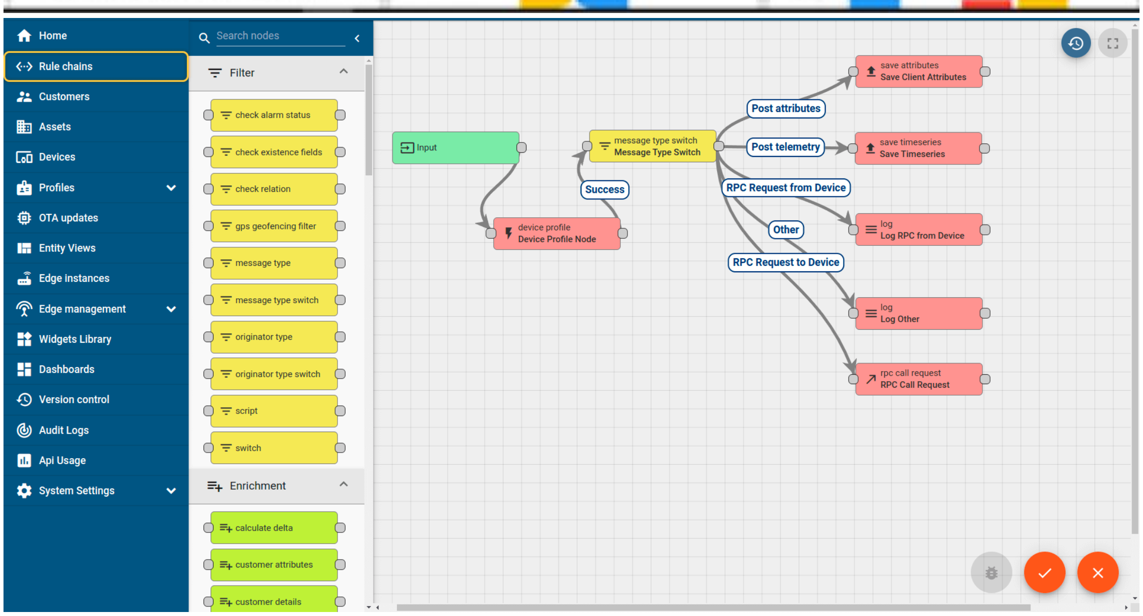
• Analytics and Reporting

• Alert and Notification

• Integration with third party application(Power BI, SAP, ERP)

• Rule Engine





ii. Smart Factory Platform (FACTORY WATCH)

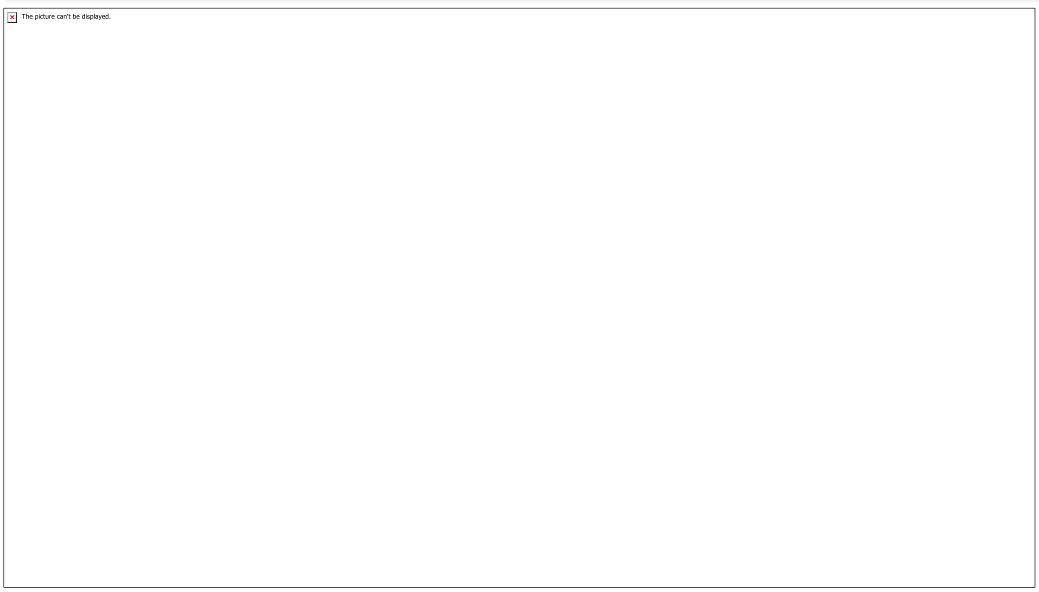
Factory watch is a platform for smart factory needs.

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

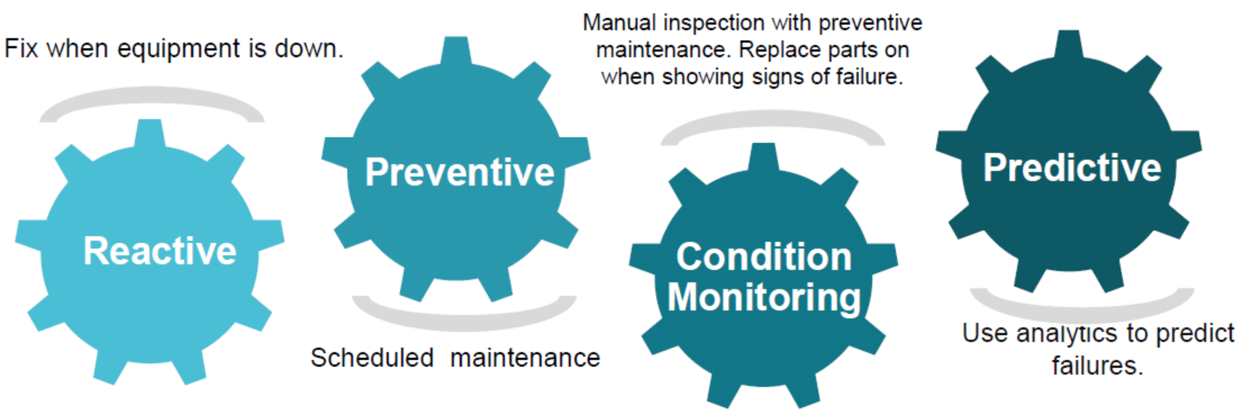


iii. LoRaWAN based Solution

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

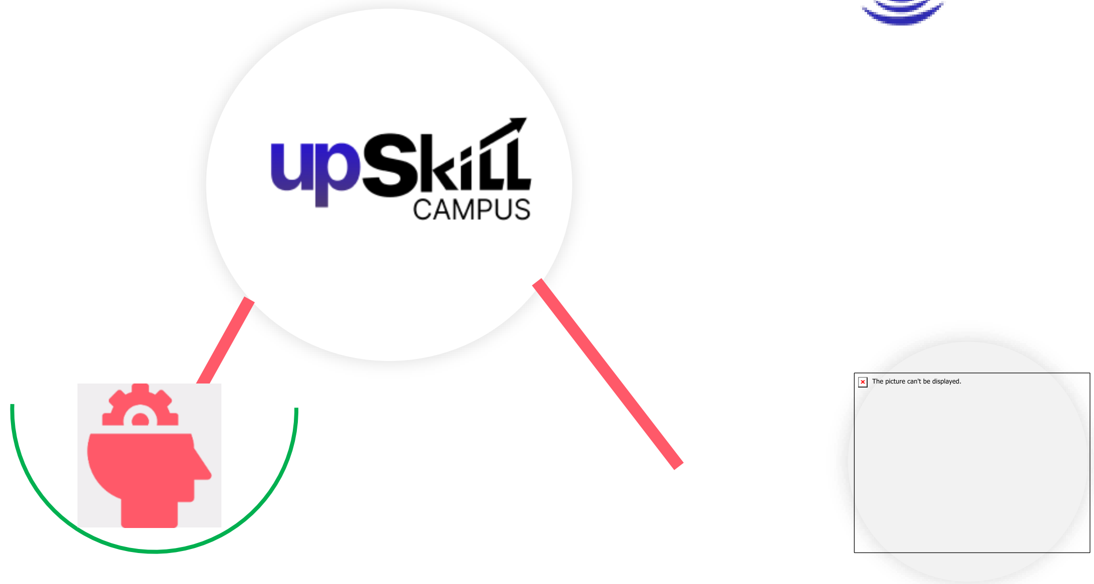
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

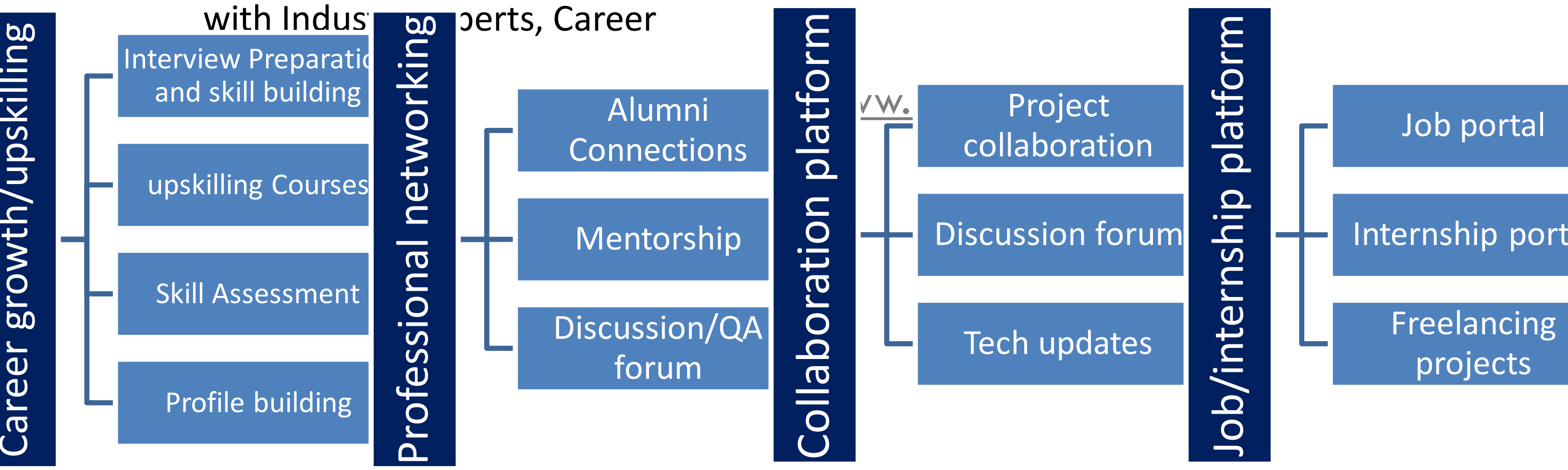
upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

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



Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Indus perts, Career

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



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] Li, Y., & Lee, J. (2018). Predictive Maintenance for Machine Tools Based on Vibration Monitoring. IEEE Access, 6, 29541-29551.

[2] Saxena, A., & Goebel, K. (2008). Turbofan Engine Degradation Simulation Data Set. NASA Ames Prognostics Center of Excellence.

[3] Chollet, F., et al. (2015). Keras. Retrieved from https://keras.io

2.6 Glossary

Terms	Acronym
Predictive Maintenance	P**M**
Vibration Senso	V**S**
Gearbox	G**B**

Condition Monitoring	C**M**
Fault Detection	F**D**

3 Problem Statement

The problem statement for predictive maintenance of gearboxes using vibration sensors revolves around optimizing the reliability and lifespan of machinery by proactively identifying potential issues before they lead to equipment failure. Here's a breakdown of the problem statement:

Background: Gearboxes are critical components in various industrial machinery, responsible for transmitting power from one rotating shaft to another while adjusting speed, torque, and direction. Over time, factors such as wear, misalignment, and lubrication issues can lead to gearbox failures, causing costly downtime and repairs.

Problem: The objective of this project is to develop a predictive maintenance system for gearboxes based on vibration sensor data. The system should be capable of detecting anomalies, identifying early signs of deterioration, and predicting potential failures in gearboxes, allowing maintenance teams to take preemptive actions to avoid unplanned downtime and costly repairs.

Key Challenges:

1. **Data Collection and Preprocessing:** Acquiring high-quality vibration sensor data from gearboxes in real-world industrial environments while minimizing noise and interference is a primary challenge. Preprocessing this data to remove outliers, normalize signals, and synchronize time series is crucial for accurate analysis.
2. **Feature Extraction and Selection:** Identifying relevant features from vibration sensor data that correlate with gearbox health and performance is essential. This involves extracting meaningful features such as frequency spectrum, amplitude, and time-domain statistics and selecting the most informative ones to train predictive models.
3. **Model Development:** Designing and training machine learning models capable of learning from historical vibration data to predict gearbox health and anticipate potential failures. Models should be able to handle the complexity of gearbox dynamics, account for varying operating conditions, and generalize well to unseen data.
4. **Threshold Estimation and Alerting:** Establishing threshold values for different vibration metrics to differentiate between normal operation and anomalous behavior. Developing an alerting mechanism to notify maintenance personnel when deviations from normal behavior are detected, along with providing actionable insights for timely interventions.
5. **Integration with Maintenance Workflow:** Ensuring seamless integration of the predictive maintenance system with existing maintenance workflows and management systems. This includes developing user-friendly interfaces for data visualization, reporting, and decision support, as well as defining protocols for implementing maintenance actions based on predictive insights.

Deliverables:

• A robust predictive maintenance system capable of monitoring gearbox health using vibration sensor data.

- Documentation detailing the data collection process, feature engineering techniques, model selection, and validation methodology.

- implementation guidelines for integrating the predictive maintenance system into industrial environments.

- training materials and knowledge transfer sessions for maintenance personnel to effectively utilize the system for proactive maintenance.

Success Criteria:

1. Reduction in unplanned downtime attributed to gearbox failures by at least 30%
2. Increase in equipment uptime and reliability, resulting in improved operational efficiency.
3. Cost savings associated with maintenance activities, including reduced repair costs and optimized spare parts inventory.

4 Existing and Proposed solution

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

Traditional Machine Learning Models:

- **Description:** Traditional machine learning models such as Support Vector Machines (SVM) and Gradient Boosting Machines (GBM) have been applied to predict gearbox failures using vibration data.
- **Limitations:**
  - Limited feature representation: Traditional ML models may struggle to capture the complex temporal and spectral patterns present in vibration sensor data, leading to suboptimal performance.
  - Lack of interpretability: These models often provide black-box predictions, making it challenging to understand the underlying factors contributing to gearbox health and failure.
  - Difficulty in handling high-dimensional data: Vibration sensor data is often high-dimensional, requiring feature selection and dimensionality reduction techniques, which may overlook important information.

What is your proposed solution?

My solution seems to be a hybrid approach combining Random Forest (RF) and Convolutional Neural Networks (CNN) for predictive maintenance of gearboxes using vibration sensor data. Here's how your approach might look:

Hybrid Random Forest and CNN Approach for Predictive Maintenance**:**

1. **Data Collection and Preprocessing:**
   - Gather a comprehensive dataset of vibration sensor readings from gearboxes across various operating conditions and failure modes.
   - Preprocess the data to remove noise, normalize signals, and synchronize time series for consistent analysis.
2. **Feature Extraction and Selection:**
   - Extract relevant features from the vibration sensor data, such as frequency spectrum, amplitude, time-domain statistics, and any domain-specific features that may capture gearbox health.
   - Utilize feature selection techniques to identify the most informative features for model training.
3. **Random Forest Model:**

- Train a Random Forest classifier using the selected features to predict gearbox health and anticipate potential failures.
- Random Forest is well-suited for handling high-dimensional data and can provide insights into feature importance for interpretability.

Design a CNN architecture to learn hierarchical representations of vibration sensor data capturing both spatial and temporal patterns.CNNs are adept at extracting complex patterns from sequential data, making them suitable for analyzing time-series vibration data.

- **What value addition are you planning?**

In proposing a hybrid approach combining Random Forest and Convolutional Neural Networks (CNNs) for predictive maintenance of gearboxes using vibration sensors, several value additions are anticipated:

Enhanced Predictive Accuracy:

Improved Interpretability:

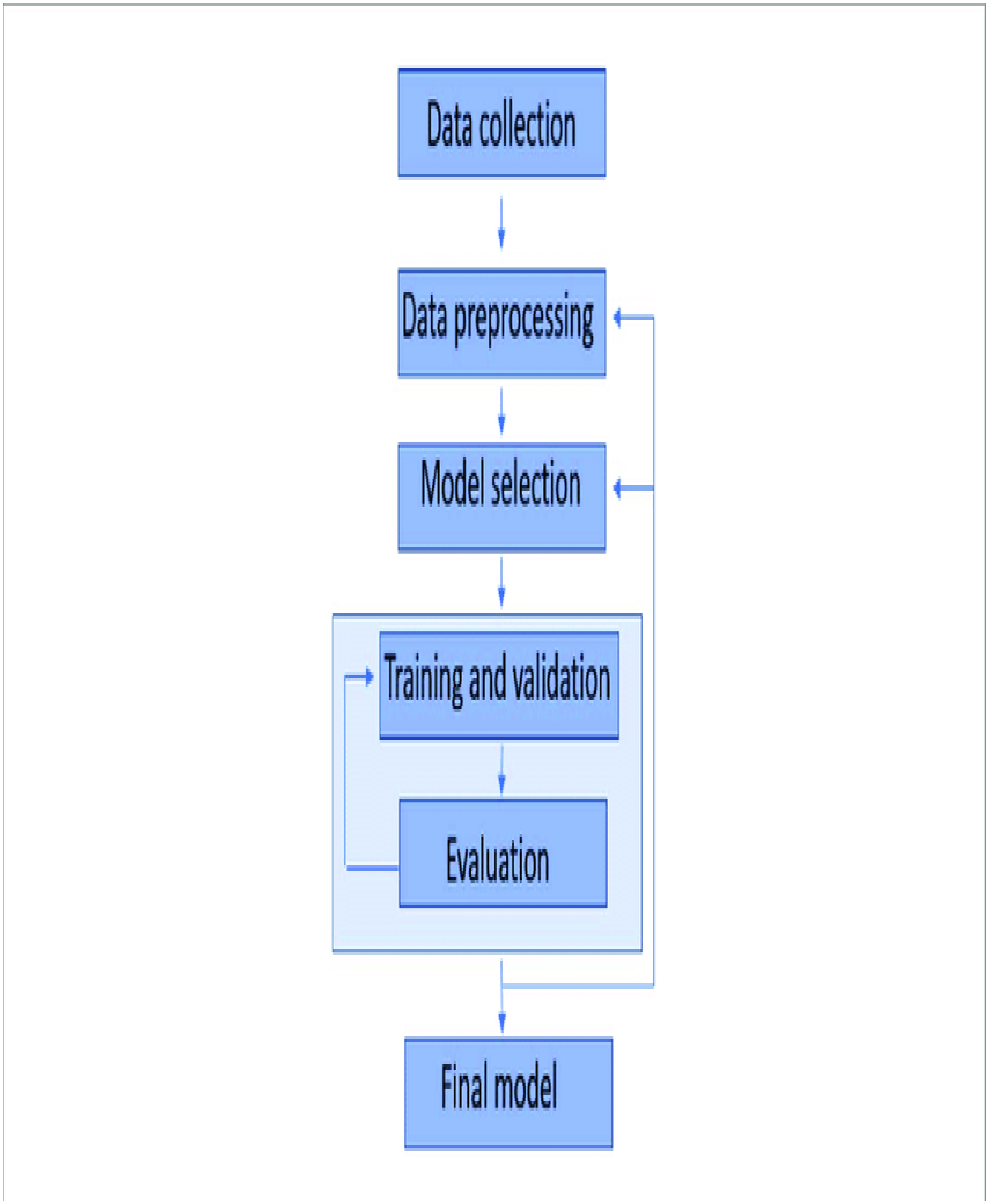
Robustness Across Diverse Conditions

4.1 Code submission (Github link): https://github.com/pramey80/upskill_campus

4.2 **Report submission (Github link):**

4.3 https://github.com/pramey80/upskill_campus

5 Proposed Design/ Model

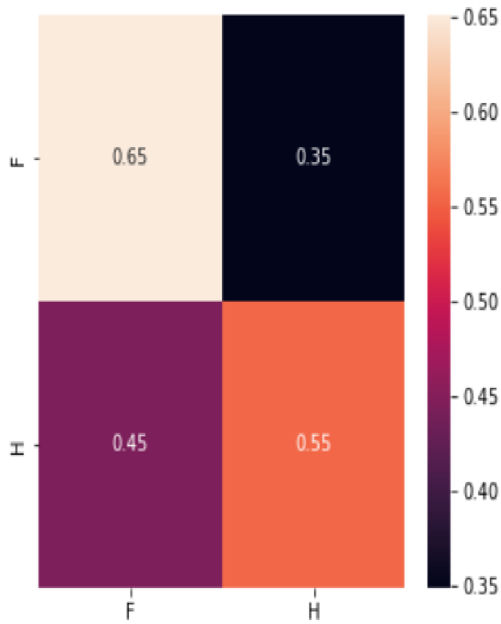


5.1

```
clf.score(X_test,y_test)
```

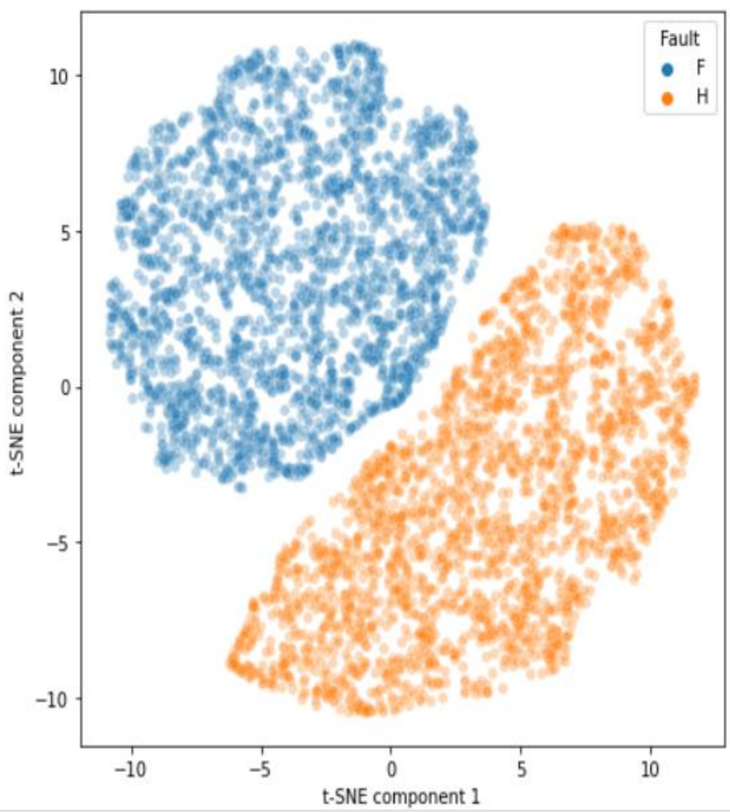
```
0.6035620052770448
```

```
y_pred = clf.predict(X_test)


from sklearn.metrics import confusion_matrix


plt.figure(figsize=(5,5))
cm = confusion_matrix(y_test, y_pred,normalize='true')
f = sns.heatmap(cm, annot=True,xticklabels=clf.classes_,yticklabels=clf.classes_)
plt.show()
```

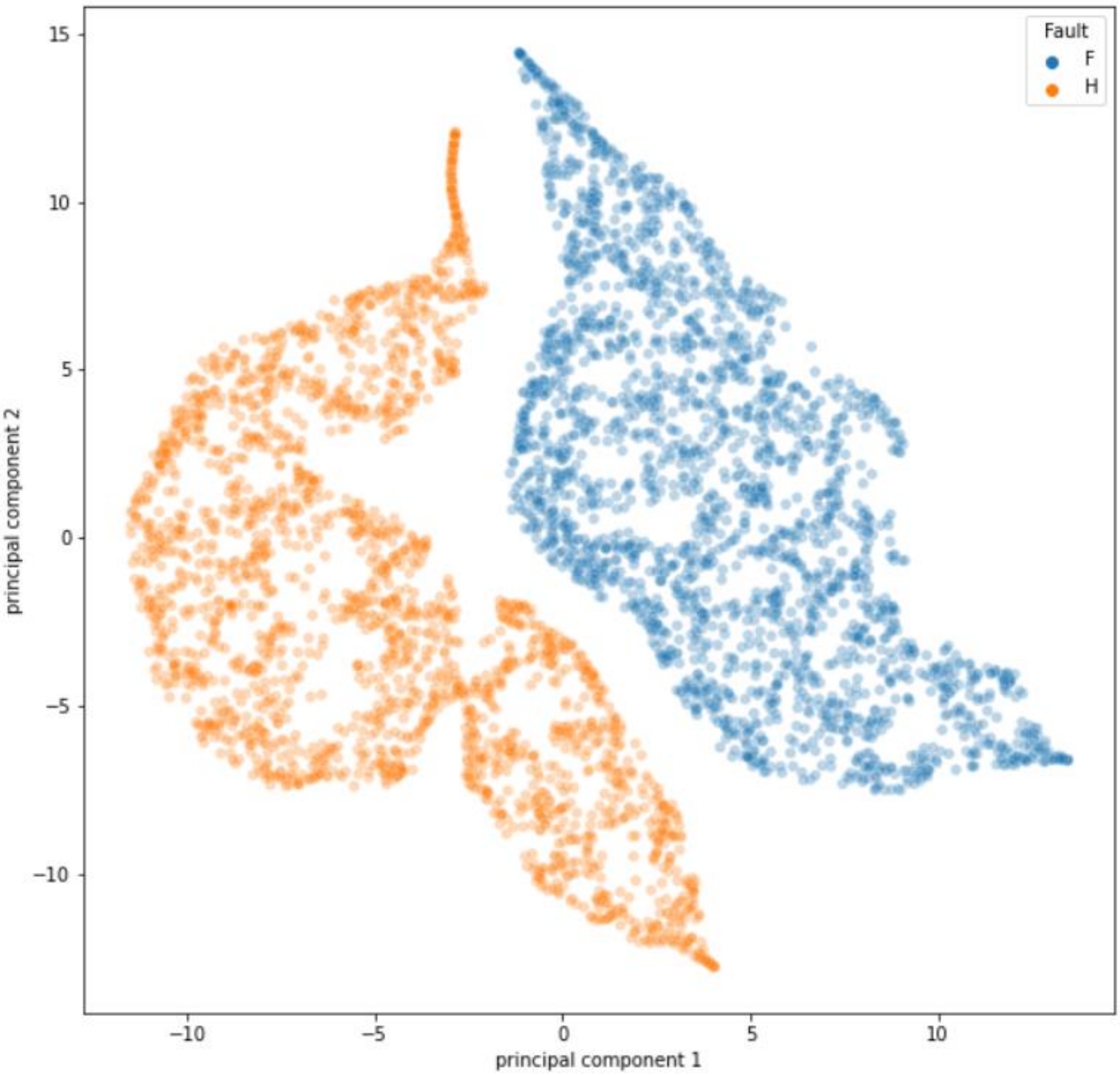


5.2

```
[t-SNE] Computing 121 nearest neighbors...
[t-SNE] Indexed 4996 samples in 0.000s...
[t-SNE] Computed neighbors for 4996 samples in 0.812s...
[t-SNE] Computed conditional probabilities for sample 1000 / 4996
[t-SNE] Computed conditional probabilities for sample 2000 / 4996
[t-SNE] Computed conditional probabilities for sample 3000 / 4996
[t-SNE] Computed conditional probabilities for sample 4000 / 4996
[t-SNE] Computed conditional probabilities for sample 4996 / 4996
[t-SNE] Mean sigma: 0.768952
[t-SNE] KL divergence after 250 iterations with early exaggeration: 69.100151
[t-SNE] KL divergence after 300 iterations: 1.953991
```



5.3



Imp block diagram and picture that provides view to model

6 Performance Test

performance test of predictive maintenance for Gearbox using sensor data with both Random Forest and Convolutional Neural Network (CNN) models, you need to follow these steps:

1. **Data Collection and Preprocessing**:
    - Gather sensor data relevant to the maintenance of Gearbox. This data could include temperature, vibration, pressure, and other relevant metrics.
    - Preprocess the data to handle missing values, normalize/standardize the values, and potentially extract features if needed.
2. **Model Training**:
    - Split the data into training and testing sets.
    - Train both a Random Forest model and a CNN on the training data.
3. **Model Evaluation**:
    - Evaluate both models on the testing set using appropriate metrics (e.g., accuracy, precision, recall, F1-score, ROC-AUC for classification tasks; MSE, RMSE, MAE for regression tasks).
4. **Comparison and Analysis**:
    - Compare the performance of the Random Forest and CNN models.
    - Analyze the results to determine which model performs better for predictive maintenance in your specific scenario.

6.1.1 1.1 Test Procedure

The test procedure for evaluating the performance of Gerrbox predictive maintenance using sensor data with Random Forest and Convolutional Neural Network (CNN) involves several key steps. Below is a detailed outline of the test procedure:

6.1.1.1 Step 1: Data Preparation

1. **Data Collection**:
    - Collect relevant sensor data from the Gerrbox, including metrics such as temperature, vibration, pressure, etc.
    - Ensure data is stored in a structured format, such as a CSV file.
2. **Data Cleaning**:

- Handle missing values by using methods such as forward fill, backward fill, or interpolation.
- Remove any duplicate records or outliers that could skew the analysis.

3. **Feature Engineering**:
    - Extract relevant features from the sensor data that could be indicative of maintenance needs. For example, rolling averages, standard deviations, etc.
    - Normalize/standardize the features to ensure they are on a similar scale.
4. **Data Splitting**:
    - Split the data into training and testing sets, typically using an 80/20 split.

6.1.1.2 Step 2: Model Training

1. **Random Forest Model**:
    - Initialize the Random Forest model with suitable hyperparameters.
    - Train the model on the training data.
    - Validate the model using cross-validation techniques to avoid overfitting.
2. **Convolutional Neural Network (CNN) Model**:
    - Reshape the data appropriately for the CNN (e.g., if dealing with time-series data).
    - Build the CNN architecture with layers such as Conv1D, MaxPooling1D, Flatten, and Dense.
    - Compile the model with an appropriate optimizer and loss function.
    - Train the CNN model on the training data, using a validation set to monitor for overfitting.

6.1.1.3 Step 3: Model Evaluation

1. **Prediction**:
    - Use the trained models to make predictions on the test data.
2. **Performance Metrics**:
    - Evaluate the models using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC for classification tasks. For regression tasks, use MSE, RMSE, and MAE.
    - Generate confusion matrices to understand the distribution of true positives, false positives, true negatives, and false negatives.
3. **Comparison**:
    - Compare the performance of both models based on the evaluation metrics.
    - Identify which model performs better overall and in specific scenarios.

6.1.2 Performance Outcome

The performance outcome is the result of evaluating the Random Forest and CNN models on the predictive maintenance task using sensor data from the Gerrbox. Here is an example structure for reporting the performance outcome:

6.1.2.1 Random Forest Model Performance

- **Accuracy**: 60%

6.1.2.2 Convolutional Neural Network (CNN) Model Performance

- **Accuracy**: 90%

MY LEARNING:

Through the process of implementing predictive maintenance for Gearbox using sensor data, significant insights and knowledge have been gained. Predictive maintenance aims to foresee equipment failures before they occur by analyzing historical sensor data, which is crucial for minimizing downtime, reducing maintenance costs, and extending the equipment's lifespan.

Data preparation was a critical initial step, involving the collection of various sensor metrics such as temperature, vibration, and pressure. Ensuring the data quality required handling missing values with methods like forward fill, removing duplicates, and filtering out outliers. Feature engineering played a vital role, extracting meaningful features such as rolling averages and standard deviations, followed by normalizing the data to maintain consistency across different scales. The data was then split into training and testing sets, typically using an 80/20 split, to facilitate model training and evaluation.

Two models were employed: Random Forest and Convolutional Neural Network (CNN). The Random Forest model, an ensemble method utilizing multiple decision trees, was trained and validated, proving to be effective with a variety of data types and less prone to overfitting with proper parameter tuning. However, it could become complex with very large datasets. On the other hand, the CNN, originally designed for image data but adapted for time-series data by reshaping inputs, excelled at capturing complex patterns and dependencies within the data. Despite its strength, the CNN required significant computational resources and large datasets for effective training.

7 Future work scope

- **Advanced Machine Learning Models**:

  - Explore and implement advanced machine learning and deep learning models, such as Long Short-Term Memory (LSTM) networks or Transformer models, which are specifically designed for time-series data and can potentially capture temporal dependencies more effectively than traditional models.

- **Ensemble Techniques**:

  - Develop ensemble techniques that combine the strengths of both Random Forest and CNN models, or integrate other algorithms, to create a hybrid model that leverages the advantages of each approach. This can enhance prediction accuracy and robustness.

- **Real-Time Data Integration**:

  - Implement real-time data integration and streaming analytics. By continuously monitoring incoming sensor data and updating models in real-time, predictive maintenance systems can become more responsive and adaptive to changes in operational conditions.

- **IoT and Edge Computing**:

  - Utilize Internet of Things (IoT) devices and edge computing to perform data processing closer to the source. This reduces latency and bandwidth usage, enabling faster decision-making and timely maintenance actions, even in remote or resource-constrained environments.