

# Predictive Maintenance using Machine Learning on Industrial Water Pumps

Dr. Sharda Chhabria<sup>1</sup>, Rahul Ghata<sup>2</sup>, Varun Mehta<sup>3</sup>, Ayushi Ghosekar<sup>4</sup>, Manasi Araspure<sup>5</sup>,  
Nandita Pakhide<sup>6</sup>

<sup>1</sup>Associate Professor, <sup>2</sup>UG Student

<sup>1</sup>Dept. of Artificial Intelligence

<sup>2</sup>Dept. of Computer Science and Engineering

G H Rasoni Institute of Engineering and Technology Nagpur, India- 440028

*rahulghata456@gmail.com*

**Received on:** 11 June ,2022

**Revised on:** 07 August ,2022,

**Published on:** 09 August,2022

**Abstract** – To help the industry, detect critical problems in production or maintenance devices, Predictive Maintenance or Predictive Preservation are combined with the Internet of Things (IoT). We propose in this research paper a system architecture model for detecting early water pump system failures by utilizing prevailing measurements captured by the monitored controlled devices. In the tentative portion, we have worked on real measured datasets, events, and failures from the water pumps sector. There are different preservation approaches or tactics that are being used to keep the effectiveness of the industries. For any specific industry, preservation disturbs the value of product works. To avoid shocking or later breakdown, the preservation methods ought to be deliberate in such a mode that the preservation tasked to decrease the conservation costs and the time required to implement them properly. This research study describes the deployment of a keen and Machine Learning architecture system for Predictive Maintenance or Preventive preservation, based on the Random Forest method in an Industry sector, that considers IoT (Internet of Things) and Machine Learning(ML) technologies to support the real-time statistics, an online collection of data, and analysis for detecting machine breakdowns sooner, allowing the real-time monitoring on the data, visualization of the statistics and schedule of preservation interventions to

mitigate the existence of such failures in the water pumps. The deployed system architecture also integrates Machine Learning techniques to provision the technicians during the execution of preservation interventions..

**Keywords**— *Machine Learning, Predictive Maintenance (PdM), Classifiers, Regression, Internet of Things, Water Pumps, Sensors..*

## I - INTRODUCTION

**P**redictive Maintenance or generally known as “online monitoring”, “condition-based maintenance” or “risk-based maintenance”, maybe the content of the many recent journals with a history behind it. It states the equipment of intelligent monitoring to avoid future failures or future consequences. Predictive Preservation has progressed from the first technique which is a visual inspection of the data to automated approaches using advanced signal processing techniques or methods based on Pattern Recognition methods and Fuzzy Logic, Machine Learning (ML), etc. The automated approaches provide a viable resolution to many industries sectors for detecting and gathering sensitive statistics from the types of equipment which are mainly machines, where human sight or ears can cease to do so. Integrated Sensors and

predictive preservation can evade unnecessary equipment part replacements, reduce machine downtime or the failure, discover the root cause of the fault, and this way improves efficiency and save costs. Predictive preservation overlays with the scope of preventive conservation in terms of preparing the maintenance commotion in advance to evade machine failures and machine downtime. In contrast to conservative preventive preservation, predictive preservation schedule activities are based on collecting datasets from different sensors and doing analysis algorithms (examination).

The Machine Learning technique is also known as the data-driven method, which uses a historical dataset to train a model of system behavior. The model-based practice has the flexibility to embrace a physical understanding of the target product, counting on the analytical approach model to represent the behavior of the industry's system. Machine Learning approaches are commonly used in areas where the obtain ability of statistics is increasing such as the preservation in industry sectors. Progressively, it provides operative solutions, cloud-based resolutions, and newly introduced procedures. Machine Learning created predictive preservation may be separated into the subsequent two main approaches: Supervised, wherever data on the occurrence of downtime are present within the modeling knowledge set, and unattended - where supplying method information is offered however no preservation information exists. The obtain ability of preservation information is highly dependent on the nature of the preservation management policy. The most preferable resolution is supervised maintenance. From the perspective of Machine Learning methods, two different approaches for supervised problems are probable, depending on the endpoint of the datasets: classification problems (presumptuous categorical values) and regression problems (presumptuous continuous values).

It is possible to move from preventive to predictive preservation with the use of equipment monitoring and fault detection. Preventive preservation necessitates extensive data on the devices' current state of health. To obtain the statistics, further metering, physical progression modeling methods, or data-driven procedures can be applied. In part, leveraging advanced analytics and integrating devices with existing data allowed for the rapid creation of data-driven modeling systems that delivered perceptions at a low cost. Preventive preservation entails reducing strategic and unplanned disruptions, as well as better-planned preservation strategies.

The production environment in the analytical work is complicated by decisive the validity of the information. In the circumstance of physically entered information, such as in failure logs, the information rarely includes the specific period of the event, and the content must be interpreted by a professional. This makes it difficult to correctly categorize measures and to correctly time their existence in the streams of data originating from the sensors. The sensor statistics are frequently subject to meddling from other systems devices in accumulation to measurement error. For example, vibrations may be recorded from the linked module; temperature measurement outcomes are affected by climate situations. Occasionally, measurement values are non taken from sensors but are replicated. This exercise helps to avoid incorrect alarms in the security system. The overhead factors make the manufacturing environment significantly different from the laboratory situation, so it is important to treat information with a high degree of vagueness.

**II- LITERATURE SURVEY**

Table 1 compares the many projects available in the company like Vending machines, Cutting Machine IoT, Industrial Pumps, Packaging Robots, and the project we describe here which is Predictive Maintenance (PdM) using Machine Learning (ML) in Industrial Water Pumps system. All the key capabilities such as Temperatures, Vibrations, Sounds, and Voltages all these features give remarkable precise predictions using the Machine Learning (ML) Algorithm.

*Table 1 - Comparison Between Different Projects Available For Company*

Sr No.	System	ML Algorithm & Model Performances	Limitations	Ref
1	Industrial Water Pump	RF: <ul style="list-style-type: none"> <li>• Accuracy : 97%</li> <li>• Recall: 96.5%</li> <li>• Precision : 98%</li> </ul>	<ul style="list-style-type: none"> <li>• More time to prepare data.</li> <li>• If sensors are replaced, maintaining data consistency becomes challenging.</li> </ul>	[2]
2	Cutting Machine IIOT	RF: <ul style="list-style-type: none"> <li>• Accuracy : 90%</li> <li>• Recall: 91.5%</li> <li>• Precision : 90%</li> </ul>	<ul style="list-style-type: none"> <li>• Different classifiers lead to increasing complexity.</li> </ul>	[1]
3	Vendor	SVM, RF	<ul style="list-style-type: none"> <li>• Precision and</li> </ul>	[4]

	Machine	& GBM: <ul style="list-style-type: none"> <li>• Accuracy : 96%</li> <li>• Recall: 96.5%</li> <li>• Precision : 97%</li> </ul>	accuracy are limited to 80% on average.	
4	Induction motor, 2.2 KW	RF: <ul style="list-style-type: none"> <li>• Accuracy : 95%</li> <li>• Recall: 96.5%</li> <li>• Precision : 93%</li> </ul>	<ul style="list-style-type: none"> <li>• The cost is higher since two train models were used.</li> <li>• The effects of undervoltage, overvoltage, and unbalanced voltage are not taken into account.</li> </ul>	[7]
5	Packaging Robot	ANN: <ul style="list-style-type: none"> <li>• Accuracy : 97%</li> <li>• Recall: 96.5%</li> <li>• Precision : 98%</li> </ul>	<ul style="list-style-type: none"> <li>• Not having IoT technology.</li> <li>• Offline Data is collected manually.</li> </ul>	[8]
6	SEW Euro drive 1.1 kW motor	RF: <ul style="list-style-type: none"> <li>• Accuracy : 95%</li> <li>• Recall: 96.5%</li> <li>• Precision : 98%</li> </ul>	<ul style="list-style-type: none"> <li>• Imbalance cannot be accurately detected.</li> <li>• Multi-sensor uses increase cost.</li> </ul>	[9]
7	1200 RPM Wind Turbine	ANN: <ul style="list-style-type: none"> <li>• Accuracy : 97%</li> <li>• Recall: 96.5%</li> <li>• Precision : 98.1%</li> </ul>	<ul style="list-style-type: none"> <li>• The feature required is a more complex computation.</li> </ul>	[10]
8	Printing Machine	LR, XG Boost & RF: <ul style="list-style-type: none"> <li>• Accuracy : 95%</li> <li>• Recall: 96.2%</li> <li>• Precision : 96%</li> </ul>	<ul style="list-style-type: none"> <li>• Data that is missing owing to errors in the data collection process.</li> <li>• Data processing complexity lets to cost.</li> </ul>	[11]
9	Woodworking Industrial Machines	GBM, RF, XG Boost & NNC: <ul style="list-style-type: none"> <li>• Accuracy : 98.9%</li> <li>• Recall: 99.6%</li> <li>• Precision : 99.1%</li> </ul>	<ul style="list-style-type: none"> <li>• Obtaining data from a defective woodworking machine.</li> <li>• Uncertainty Propagation Using Statistical Models</li> </ul>	[12]

In this field, several more research, and methods have been undertaken and developed. A technique for early or later identification of defects in water pump systems based on signals acquired with the use of several sensors mounted on a water pump system is presented in this

research article. We employ a Random Forest (RF) classifier model created for the specific water pump system being observed in real-time for this resolution. The classifier model allows deviations from the operating state to be detected without the need to label data. As a result, our method reduces the need for expert knowledge, particularly knowledge regarding the classification of errors and the precise timing of their occurrence, which might be difficult to determine.

### III- SYSTEM ARCHITECTURE & METHODOLOGY

A proper Predictive Preservation system design should include specific modules as well as overall module combinations that can lead to proper Predictive Preservation system implementation.

The system's functioning starts with information procurement, which records and stores information acquired from multiple sensor modules in a real-time way. This information will be used to analyze data by grouping multiple parameters or analyzing individual parameters, the information provided for data analysis is sent to the state detection block at the same time. This enables real-time monitoring of system status and can be used for early fault detection or maintenance. The results of the data analysis are utilized to determine the current health and prognosis of the water pumping system. Finally, after assessing the health state and

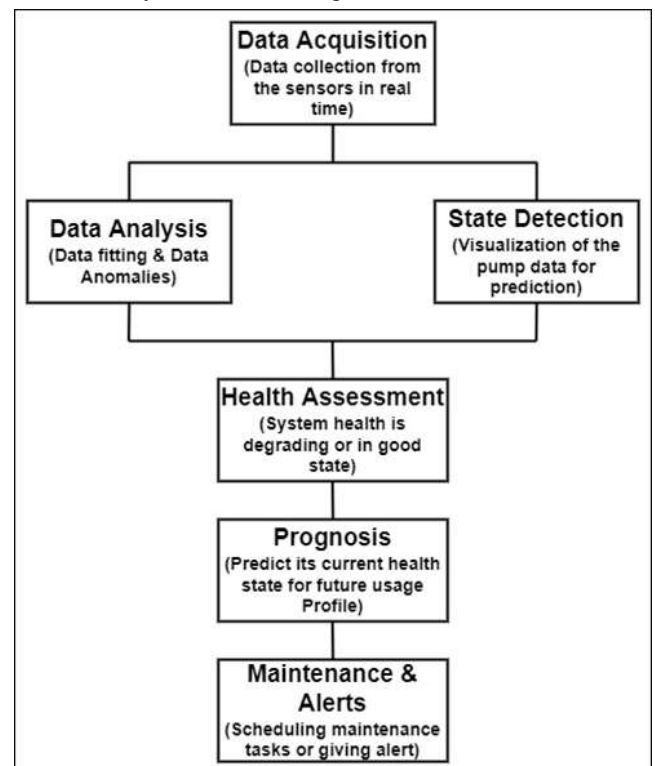


Fig 1- System Architecture

prognosis, relevant maintenance actions can be selected and planned using the maintenance action block.

The functionality or operation of this system, as well as each component of its architecture as shown in Fig 1, can be described beginning with the collection of sensor data.

Finally, after assessing the health state and prognosis, relevant maintenance activities can be selected and planned using the maintenance action block. Various parameters must be recorded for those. This can be accomplished by placing many sensors on the water pump system, which will collect data, and then using the MCU (Micro Controller Unit) module to send and store the data in the database.

The information gathered from sensors and MCU modules are used for data analysis, model training, and system state detection. Data from the data acquisition system is a historical record of the system. Various strategies can be employed in data analysis and model training to identify defects, discover errors, and so on. While state detection is used for visualizing the present condition of the water pump system and early detection of faults by just observing the system's behavior.

The data analysis results are used for a health evaluation, which includes urgent or planned water pump maintenance. Based on its behavioral history, a health assessment informs if the system's health is deteriorating or improving, and prognosis will help estimate its current health condition in the future by considering future usage. Once the prognosis results are known, it will be possible to determine the proper maintenance chores to be performed based on the parameters affected by the water pump system and, as a result, schedule the jobs or send alerts to avoid a breakdown.

#### IV- IMPLEMENTATION

This proposed system architecture's functionality or operation is shown in Fig 2. It consists of multiple sensor nodes (Temperature Sensor, Voltage Sensor, Vibration Sensor, and Sound Sensor), a Wi-Fi MCU (Micro Controller Unit) module, cloud services, and graphical user interfaces make up the system (GUI). The methods proposed can be used in any sector. Temperatures, vibrations, voltages, currents, and other typical parameters that may be required in any industry for any system are some of the most prevalent. We need historical machine data collected over time to train and create the machine learning (ML) model. To gather and

convey the data, a sensor node was created containing temperature, sound, current, and voltage sensors. The machine learning model is essential for real-time machine data analysis. It must store the model to reuse it; otherwise, it will have to be trained repeatedly during the prediction phase. The data from the Running Machine is then collected in real-time. After the data has been collected, it is saved for later analysis. Following that, the data is pre-processed, and the model is trained. A cloud database will be developed for this purpose. Because the prediction algorithms or preservation were intended to be implemented in the cloud, the database was developed there. The benefit is that the same predictive model may be utilized by various devices, which reduces the expense of establishing a dedicated predictive model.

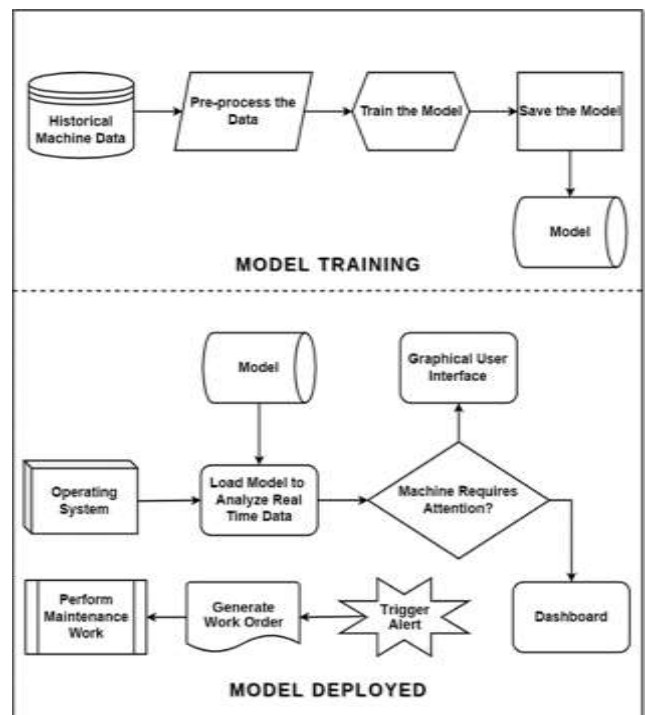


Fig 2 -Block Diagram

The framework unit for sensor nodes was created as shown in Figure 3, the WI-FI module ESP8266 was employed as the microcontroller unit for sensor node creation. The temperature sensor that was used was the LM35 temperature sensor. A three-phase power meter B25 was used to monitor voltage and current. The sound sensor used to detect sound frequency was the Sound Sensor LM393. A vibration sensor SW-420 was utilized to measure vibration. All of these sensors' data was gathered. The gateway received all of the data read by the microcontroller via wireless transmission. As a gateway, an ESP8266 MCU WIFI module was employed. The data was formatted as needed once it

arrived at the gateway. To upload it to the cloud, use the JSON format. The cloud platform used was Amazon Web Services (AWS). The database was built utilizing the AWS platform's Elastic Search Offering service. Under that service, a MySQL database was constructed, and all sensor data was stored in the cloud. The database's information was saved. Dashboards for visualizing the present state of the water pump system were created using Flask and Figma (UI/UX). The dashboards are accessible from any device because they are built on a cloud platform.

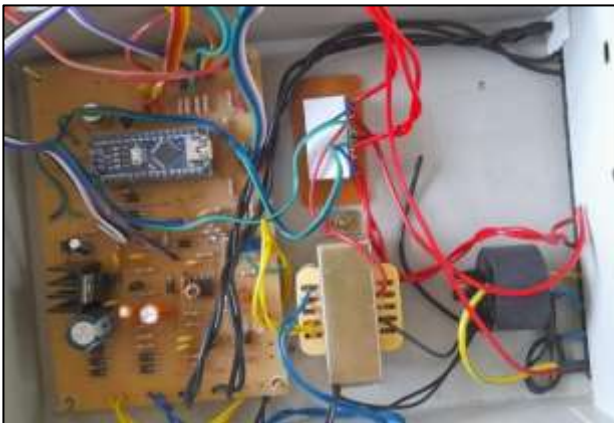


Figure 3 Sensor Nodes

From the dashboard, as shown in **Error! Reference source not found.**, in real-time, the present state of the water pumps system and the predictions from the machine learning (ML) model may be monitored. If an abnormal machine condition is detected, the Alarm is raised and the notification is sent to the maintenance engineer for maintenance work. The work order for the maintenance works is automatically generated as well.



Fig 4- Dashboard

## V - CONCLUSION

This paper concludes that predictive maintenance (PdM) or predictive preservation is revolutionizing the industry

because due to machine learning (ML) the maintenance revenue will be cut to 15% from 30% which was used in preventive maintenance. Predictive Maintenance (PdM) is an efficient method to improve the remaining useful life of a machine as well as enhance performance. It's possible to improve the detection of machine status using more trained models and can give better results in the future. It uses two models at the same time to collect and predict genuine positive values, making the system more trustworthy than single model prediction machine learning techniques and the type of data used. Predictability retention is still an efficient technique to boost performance in all types of engine and breakdown equipment situations. Temperature analysis, vibration analysis, voltage analysis, and acoustic analysis can all help you attain a more stable condition of motion. Furthermore, combining approaches can strengthen the evidence for detecting errors or inefficiencies.

Given that machine learning is a new approach to the use of predictive care, the preservation of speculation is considered as having significant growth potential. According to the literature, machine learning is frequently employed. The Random Forest algorithm, which is employed in most motors and systems, including rotation machinery, industrial engines, cutting machines, wind turbines, sales equipment, and many others, is used in the Predictive Maintenance approach. The approach was also extremely sensitive to any changes in the system under examination, particularly when one of the sensors failed. The algorithm's benefit, on the other hand, was the ability to forecast large failures long before they occurred, as well as signal detection that suggested a likely source of the failure event. The advantage of the method described above in previous methods presented in related research was its success in the case of weak descriptive data events anticipated solely based on process data. The method operated quickly and did not require a lot of hardware because of its simplicity.

## VI- FUTURE SCOPE

Internet of Things (IoT) based forecasting model applications can be researched and used to make this more commendable in the future. A combination of more than one error or defect recognition machine learning (ML) approach and models can produce the most advanced forecast when compared to the use of individual models.

## REFERENCES

- [1] Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E. and Loncarski, J., 2018, July. *Machine learning approach for predictive maintenance in industry 4.0*. In *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)* (pp. 1-6). IEEE.
- [2] Amihai, I., Gitzel, R., Kotriwala, A.M., Pareschi, D., Subbiah, S. and Sosale, G., 2018, July. *An industrial case study using vibration data and machine learning to predict asset health*. In *2018 IEEE 20th Conference on Business Informatics (CBI)* (Vol. 1, pp. 178-185). IEEE.
- [3] Alves, F., Badikyan, H., Moreira, H.A., Azevedo, J., Moreira, P.M., Romero, L. and Leitão, P., 2020, June. *Deployment of a smart and predictive maintenance system in an industrial case study*. In *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)* (pp. 493-498). IEEE.
- [4] Xiang, S., Huang, D. and Li, X., 2018, October. *A generalized predictive framework for data driven prognostics and diagnostics using machine logs*. In *TENCON 2018-2018 IEEE Region 10 Conference* (pp. 0695-0700). IEEE.
- [5] Manjare, A.A. and Patil, B.G., 2021, March. *A Review: Condition Based Techniques and Predictive Maintenance for Motor*. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 807-813). IEEE.
- [6] Butte, S., Prashanth, A.R. and Patil, S., 2018, April. *Machine learning based predictive maintenance strategy: a super learning approach with deep neural networks*. In *2018 IEEE Workshop on Microelectronics and Electron Devices (WMED)* (pp. 1-5). IEEE.
- [7] Dos Santos, T., Ferreira, F.J., Pires, J.M. and Damásio, C., 2017, May. *Stator winding short-circuit fault diagnosis in induction motors using random forest*. In *2017 IEEE International Electric Machines and Drives Conference (IEMDC)* (pp. 1-8). IEEE.
- [8] Onur, K.O.C.A., Kaymakci, O.T. and Mercimek, M., 2020, May. *Advanced Predictive Maintenance with Machine Learning Failure Estimation in Industrial Packaging Robots*. In *2020 International Conference on Development and Application Systems (DAS)* (pp. 1-6). IEEE.
- [9] Janssens, O., Loccufer, M. and Van Hoecke, S., 2018. *Thermal imaging and vibration-based multisensor fault detection for rotating machinery*. *IEEE transactions on industrial informatics*, 15(1), pp.434-444.
- [10] Biswal, S. and Sabareesh, G.R., 2015, May. *Design and development of a wind turbine test rig for condition monitoring studies*. In *2015 international conference on industrial instrumentation and control (icic)* (pp. 891-896). IEEE.
- [11] Binding, A., Dykeman, N. and Pang, S., 2019, April. *Machine Learning Predictive Maintenance on Data in the Wild*. In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)* (pp. 507-512). IEEE.
- [12] Calabrese, M., Cimmino, M., Fiume, F., Manfrin, M., Romeo, L., Ceccacci, S., Paolanti, M., Toscano, G., Ciandrini, G., Carrotta, A. and Mengoni, M., 2020. *SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0*. *Information*, 11(4), p.202.
- [13] Moleda, M., Momot, A. and Mrozek, D., 2020. *Predictive maintenance of boiler feed water pumps using SCADA data*. *Sensors*, 20(2), p.571.
- [14] Manjare, A.A. and Patil, B.G., 2021, March. *A Review: Condition Based Techniques and Predictive Maintenance for Motor*. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 807-813). IEEE.
- [15] Motaghare, O., Pillai, A.S. and Ramachandran, K.I., 2018, December. *Predictive maintenance architecture*. In *2018 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)* (pp. 1-4). IEEE.
- [16] Poór, P., Basl, J. and Zenisek, D., 2019, March. *Predictive Maintenance 4.0 as next evolution step in industrial maintenance development*. In *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)* (pp. 245-253). IEEE.
- [17] Susto, G.A., Schirru, A., Pampuri, S., McLoone, S. and Beghi, A., 2014. *Machine learning for predictive maintenance: A multiple classifier approach*. *IEEE transactions on industrial informatics*, 11(3), pp.812-820.
- [18] Cachada, A., Barbosa, J., Leitão, P., Grcaldcs, C.A., Deusdado, L., Costa, J., Teixeira, C., Teixeira, J., Moreira, A.H., Moreira, P.M. and Romero, L., 2018, September. *Maintenance 4.0: Intelligent and predictive maintenance system architecture*. In *2018 IEEE 23rd international conference on emerging technologies and factory automation (ETFA)* (Vol. 1, pp. 139-146). IEEE.
- [19] Heim, S., Clemens, J., Steck, J.E., Basic, C., Timmons, D. and Zwiener, K., 2020, December. *Predictive maintenance on aircraft and applications with digital twin*. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 4122-4127). IEEE.
- [20] Azamfar, M., Singh, J., Bravo-Imaz, I. and Lee, J., 2020. *Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis*. *Mechanical Systems and Signal Processing*, 144, p.106861.