

EVALUATION OF CNC MACHINE OPERATION  
THROUGH MACHINE LEARNING EFFECTIVE MODEL



Author:

Muhammad Hussain

Regn Number

320084

Supervisor

Dr. Abbas Hussain

Manufacturing Engineering and Management  
DEPARTMENT OF INDUSTRIAL AND MANUFACTURING ENGINEERING  
PAKISTAN NAVY ENGINEERING COLLEGE, KARACHI  
NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

ISLAMABAD

AUGUST 2023

Evaluation of CNC Machine Operation through Machine Learning  
Effective Model

Author:

Muhammad Hussain

Regn Number:

320084

A thesis submitted in partial fulfillment of the requirements for the  
degree of MS Mechanical Engineering

Thesis Supervisor:

Dr. Abbas Hussain

Thesis Supervisor Signature:



Manufacturing Engineering and Management  
DEPARTMENT OF INDUSTRIAL AND MANUFACTURING ENGINEERING  
PAKISTAN NAVY ENGINEERING COLLEGE, KARACHI  
NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY  
ISLAMABAD  
AUGUST 2023

# National University of Sciences & Technology

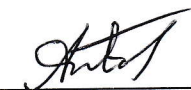
## MASTER THESIS WORK

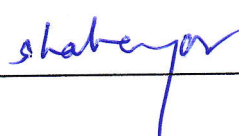
We hereby recommend that the dissertation prepared under our supervision by:


Muhammad Hussain & Reg No: 320084

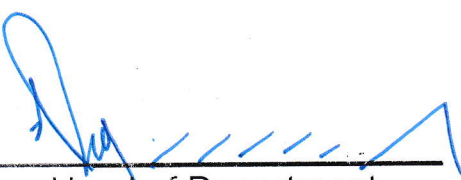
Titled: "Evaluation of CNC Machine Operations through machine learning effective model" be accepted in partial fulfillment of the requirements for the award of MS(ME&M) degree and awarded grade 'B+' (Initial) AE

### Examination Committee Members

1. Name: Dr. Antash Najib Signature: 

2. Name: Dr. Sheharyar Atta Khan Signature: 

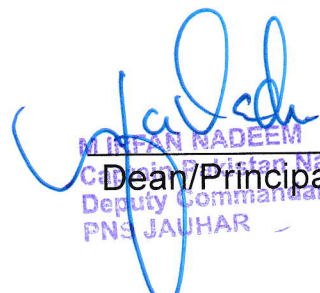
Supervisor's name: Dr. Abbas Hussain Signature:   
Date: 28/08/2023

  
Head of Department  
DR. ZEESHAN RIAZ  
Captain Pakistan Navy  
Dean IME

29/08/2023  
Date

### COUNTERSIGNED

Date: 31/08/2023

  
M. IRFAN NADEEM  
Captain Pakistan Navy  
Deputy Commandant  
PNS JAUHAR


## THESIS ACCEPTANCE CERTIFICATE

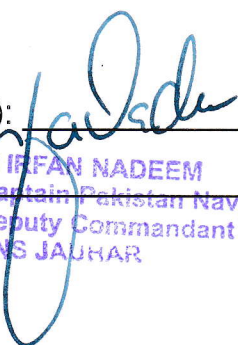
Certified that final copy of MS thesis written by Muhammad Hussain Registration No. 00000320084, of PNEC College has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS Degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature:  \_\_\_\_\_

Name of Supervisor: Dr. Abbas Hussain

Date: 22/08/2023

Signature (HOD):  \_\_\_\_\_  
DR. ZEE SHAN RIAZ  
Captain Pakistan Navy  
Date: 29/08/2023  
Dean IME

Signature (Dean/Principal):  \_\_\_\_\_  
Date: 31/08/2023  
M IFFAN NADEEM  
Captain Pakistan Navy  
Deputy Commandant  
PNS JAUHAR

## Declaration

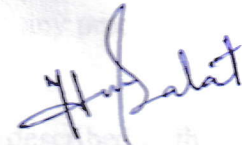
I certify that this research work titled as "*Evaluation of CNC Machine Operation through Machine Learning Effective Model*" is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledge / referred.



Muhammad Hussain  
2023-NUST-MS-IME-320084

## Plagiarism Certificate (Turnitin Report)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.



Muhammad Hussain  
320084



Signature of Supervisor

## **Copyright Statement**

- Copyright in the text of this thesis rests with the student author. Copies (by any process) either in full or of extracts, may be made only by instructions given by the author and lodged in the Library of PNEC, NUST. Details may be obtained by the Librarian. This page must form part of any such copies made. Further copies (by any process) may not be made without the author's permission (in writing).
- The ownership of any intellectual property rights which may be described in this thesis is vested in PNEC, NUST, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the PNEC-NUST, which will prescribe the terms and conditions of any such agreement.
- Further information on the conditions under which disclosures and exploitation may take place is available from the Library of PNEC – NUST, Karachi.

## **Acknowledgments**

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought which You set up in my mind to improve it. Indeed I could have done nothing without Your priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was Your will, so indeed none be worthy of praise but You.

I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout every department of my life.

I would also like to express special thanks to my supervisor Dr. Abbas Hussain for his help throughout my thesis.

I would also like to pay special thanks to Dr. Shehryar Atta for his tremendous support and cooperation. Each time I got stuck in something, he came up with a solution.

I would also like to thank Dr. Antash Najib, for being on my thesis guidance and evaluation committee and express my special thanks for his help.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study.



*Dedicated to my exceptional parents and adored siblings whose  
tremendous support and cooperation led me to this wonderful  
accomplishment.*

## Abstract

All machining processes involve vibrations that are generated either through the interaction of cutting tools and workpieces or by machine structure. These relative vibrations between the cutting tool and the workpiece are considered important for the final machining output. These vibrations lead to a machining process failure if such generating vibration don't address on time. Besides this, the occurrence of machine failure in the machining process can also cause due to other factors like misaligning of the tool, chip in the chuck, chip clamping, tool breakage, etc. In recent decades, the machining process has made significant technological advancements. To follow the agile philosophy in the industry, it is necessary to use algorithms based on data to tackle the increasing complexity. Specifically in Computer Numerical Control machining (CNC Machining). To improve machining, a simple and time-effective method has been used in this study. An auto-machine learning tool applied to data extracted from real-world production plants over two years. In this study, an auto machine learning tool was used for training data, training done by numerous models, and then the most effective and accurate model was used for predicting new data.

**Key Words:** *Machine Learning; Computer Numerical Control machining; Milling; Quality Prediction*

# Table of Contents

## Contents

<b>Declaration</b> .....	<b>v</b>
<b>Plagiarism Certificate (Turnitin Report)</b> .....	<b>vi</b>
<b>Copyright Statement</b> .....	<b>vii</b>
<b>Acknowledgments</b> .....	<b>viii</b>
<b>Abstract</b> .....	<b>x</b>
<b>Table of Contents</b> .....	<b>xi</b>
<b>List of Tables</b> .....	<b>xii</b>
<b>List of Figures</b> .....	<b>xii</b>
<b>INTRODUCTION</b> .....	<b>1</b>
1.1 Background and Motivation.....	1
1.1.1 Process Monitoring.....	2
1.1.2 Machine Learning.....	3
1.2 Research Objectives .....	4
<b>CHAPTER 2: WORKING METHODOLOGY</b> .....	<b>5</b>
2.1 Experimental Setup / Data Acquisition System .....	5
2.1.1 Data Description.....	5
2.1.2 Challenges in Data.....	6
2.2 Data Pre-processing .....	7
2.3 Features Extraction and Model Selection.....	8
2.3.1 Time Domain Features Extraction.....	8
2.3.2 Auto Machine Learning in MATLAB .....	9
<b>CHAPTER 3: RESULTS AND CONCLUSION</b> .....	<b>12</b>
3.1 RESULTS: .....	12
3.1.1 Confusion Matrix .....	12
3.1.1 Accuracy of models on new data:.....	14
3.2 Conclusion .....	16
<b>REFERENCES</b> .....	<b>17</b>

## List of Tables

Table 1 <i>Trained Model and their corresponding accuracies</i> .....	12
Table 2 <i>Accuracy Result of Selected Model on New Data</i> .....	15

## List of Figures

<i>Figure 1 Process Monitoring process through Chatter Diagnosis Process [12]</i> .....	3
<i>Figure 2 Structure of data, stored in the directory</i> .....	5
<i>Figure 3 Class Distribution, process health concerning Operation</i> .....	6
<i>Figure 4 Process Health distribution w.r.t Operation after balancing data of 5 Operations</i> .....	8
<i>Figure 5 Machine Learning Workflow</i> .....	10
<i>Figure 6 Auto ML Workflow</i> .....	10
<i>Figure 7 Confusion matrix of two trained models with higher accuracy</i> .....	14
<i>Figure 8 Confusion of selected model on new data set</i> .....	16

# INTRODUCTION

## 1.1 Background and Motivation

In the past 20 years, machining technologies used in manufacturing have greatly improved due to a competitive and innovative market. As a result, computer numerical control (CNC) machines have advanced from basic punch-driven devices to complex computer-operated tools. With the continuous progress in software and information technologies, along with advancements in hardware and control technologies, machining accuracy and efficiency have been consistently improved. However, despite these advancements, imprecision sources in machine tools remain a significant limitation for machining accuracy and productivity in metal cutting processes. [1]

Computer Numerical Control (CNC) machines are one of the most durable and enduring components of the production process. Machining centers that are highly automated are known for their fast production, but they can also be quite complex. The intense environmental conditions and speedy processing can lead to problems like tool breakage, improper tool clamping, and chip jamming. With a wide range of tool types and operations that vary in shape, coatings, material, surface finishing changes over time, traditional analytics face significant challenges in terms of robustness and generalization [2]. When there are variations in machining constraints (parameters) and maintenance methods, such as lubricating mechanisms, the complexity of the processes can increase.

To address these challenges, extensive research has been conducted on improving the quality of processes. [3, 4] and few in tool health monitoring [5, 6] has been performed. In order to advance research in the field, various machining datasets have been made available, including the SMART LAB Milling Dataset. [7], The University of Michigan compiled this dataset from 18 experiments to investigate tool wear and inadequate clamping detection through direct measurements. The NASA Milling Dataset [8] includes data on tool wear measured by three types of sensors: vibration, current and acoustic emission. It is important to note that both experiments were performed in laboratory within a limited time frame. A dataset from a production plant was collected to solve a real-world problem. [9]. The dataset was collected over two years from three

brownfield CNC milling machines at different time intervals. It was designed to address feature drifts, tool operation variation, and class imbalance. [9].

This study advances the work on a data set collected through real-life production plant [9], work consists of sorting the dataset as per requirement, balancing it concerning Operation, and then applying the auto machine learning tool of MATLAB which provides a robust method to find an accurate model. This might help the researcher to understand the robustness of the Auto Machine learning tool and make use of it to avoid time-consuming iterations as well as model can be used for process monitoring.

### **1.1.1 Process Monitoring**

The study of process monitoring is a recent field in process condition research that has emerged alongside the progress of Industry 4.0 and soft computing. In the machining industry, various researchers have suggested and implemented the self-diagnostic machining system. Consequently, manufacturing sectors are adopting automation to meet global demand and remain competitive. [10]. In recent years, researchers have studied process monitoring extensively to reduce the need for skilled operators. Figure 1 shows the flow chart for the sensor-based online chatter identification process used in process monitoring. Firstly, the signal acquisition system collects the raw signals. Following this, three operations are carried out: signal processing, signal transformation, and feature extraction. Finally, a classifier is obtained using the soft computing method to predict chatter.

Users of machine tools can now shift from routine maintenance to condition-based maintenance strategies by utilizing process monitoring. This involves monitoring process variables using various sensors in machining-based operations [11].

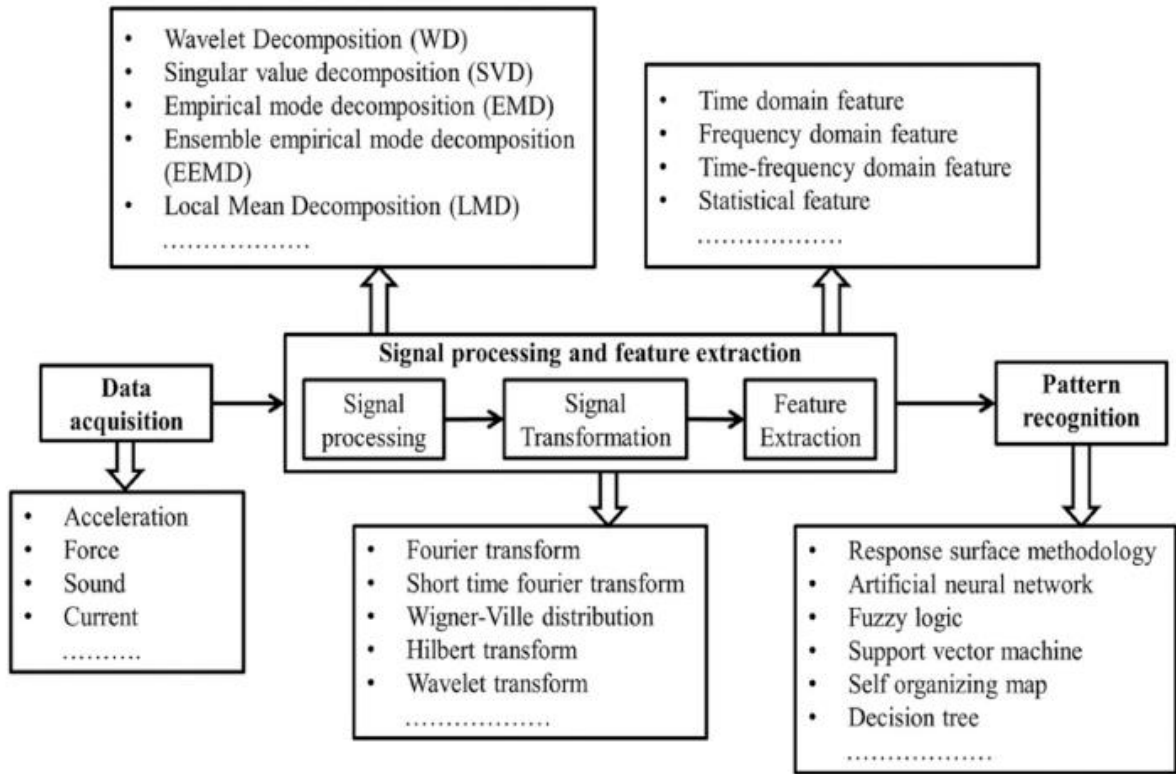


Figure 1 Process Monitoring process through Chatter Diagnosis Process [12]

### 1.1.2 Machine Learning

Machine learning (ML) involves creating algorithms that enable computers to learn from input data and make predictions or decisions. It is a subfield of artificial intelligence (AI). AI and ML are important evolutions in computer science and can enhance almost every technology-enabled service, product, and industrial application. Machine learning enables computers to learn and improve from experience, without explicit programming. ML is a subset of artificial intelligence that has applications in various fields of study [13].

Manufacturing companies are concerned about the future impact of CNC machining due to machine learning and artificial intelligence.. [14]. The learning process of a machine relies on real-time data, analytics, and deep learning. Operators require datasets to comprehend individual machine operations and machine group interactions. [ 15 ]. Thanks to the advancements in sensor technology and communication systems, machine learning can now be applied to tool condition monitoring in new and innovative ways. [16] . Artificial intelligence can improve accuracy in CNC

machining by enhancing productivity and efficiency. [17]. Machines can analyze and generate real-time production data, enhancing productivity in part manufacturing. Machine users can easily modify machine operations using advanced machine learning algorithms [18]. Incorporating machine learning and artificial intelligence into part of manufacturing operations can enhance productivity and minimize work floor downtime throughout the process. [19].

In this study, a Machine learning tool will be used for the purpose described above, CNC data collected over two years period were trained and a model is developed. The developed model will implement on other data to analyze the robustness of the model.

## **1.2 Research Objectives**

Developing a machine learning model for a dataset collected through a CNC Milling machine leads to numerous objectives. However, the main objectives are as follows:

1. Anomaly Detection and Quality Control
2. Predictive Maintenance of Equipment and tools
3. Tool Wear Prediction
4. Process parameter Monitoring
5. Process Monitoring and Control



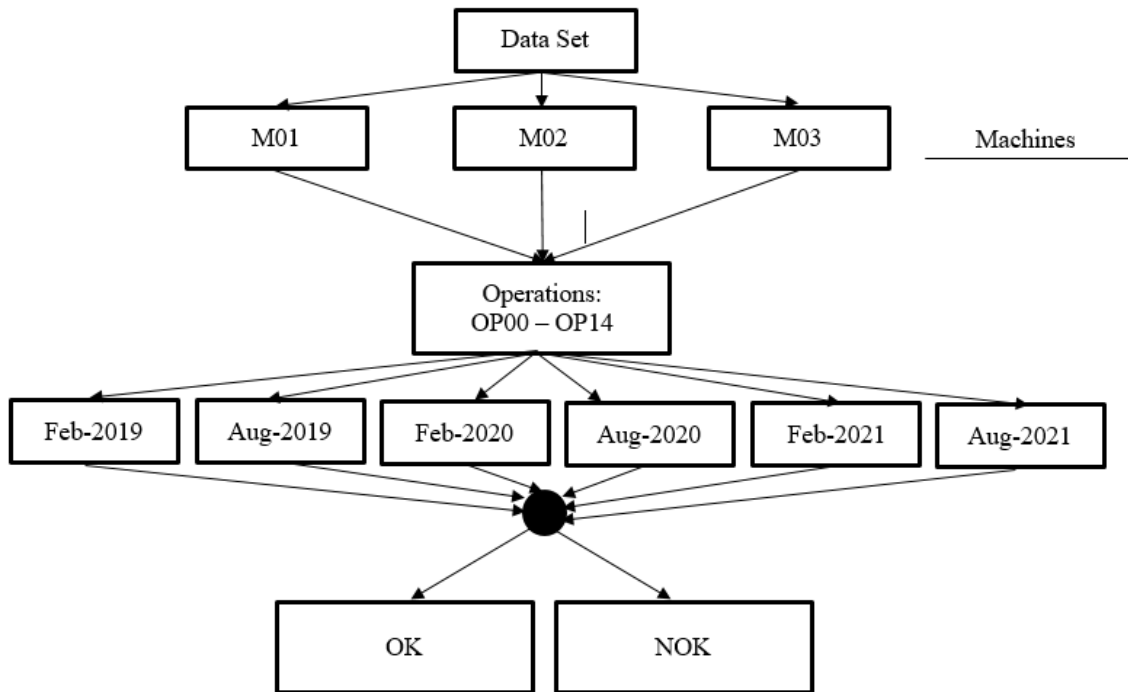
## CHAPTER 2: WORKING METHODOLOGY

### 2.1 Experimental Setup / Data Acquisition System

A dataset used in this study was collected through the “Smart Data Collection System for Brownfield CNC Milling Machines”[9] and the data set is available at: [https://github.com/boschresearch/CNC\\_Machining](https://github.com/boschresearch/CNC_Machining).

#### 2.1.1 Data Description

Data is being collected from three different CNC machines within a production plant, Figure 5 depicts the structure of the data.



*Figure 2 Structure of data, stored in the directory*

Data is collected from October 2018 to August 2021 at regular intervals. The dataset includes 15 tool operations performed on 3 machines at various times to analyze drift over time and between machines. Different parts are produced by the machines, and the process flow varies. The time frame is identified by "Month Year" and represents the 6-month interval before the label, e.g. "Aug 2019" refers to the period between February 2019 and August 2019.

### 2.1.2 Challenges in Data

During machining, high-speed operations require frequent tool mounting and unmounting on the spindle chuck. These factors occasionally cause process failures due to tool misalignment, chip-in chuck, chip clamping, and tool breakage. To ensure optimum product quality, a professional on the shop floor manually controls the workpiece in a gauging station and annotates the process's condition after each batch. However, labeling during production remains challenging due to manual gauging, leading to wrongly labeled processes and missing precise annotations. The dataset that has been published focuses specifically on failures in the quality process, with the OK class indicating an error-free process and NOK referring to a faulty process.

A major trial in industrial datasets is the significant imbalance between OK and NOK samples, as depicted in Figure 6. In our dataset, there are 816 OK samples for every 35 NOK samples. However, in our actual production, the number of OK samples is much higher. To build a model, we selected a limited dataset with an equal proportion of OK and NOK samples, as mentioned in section 2.3

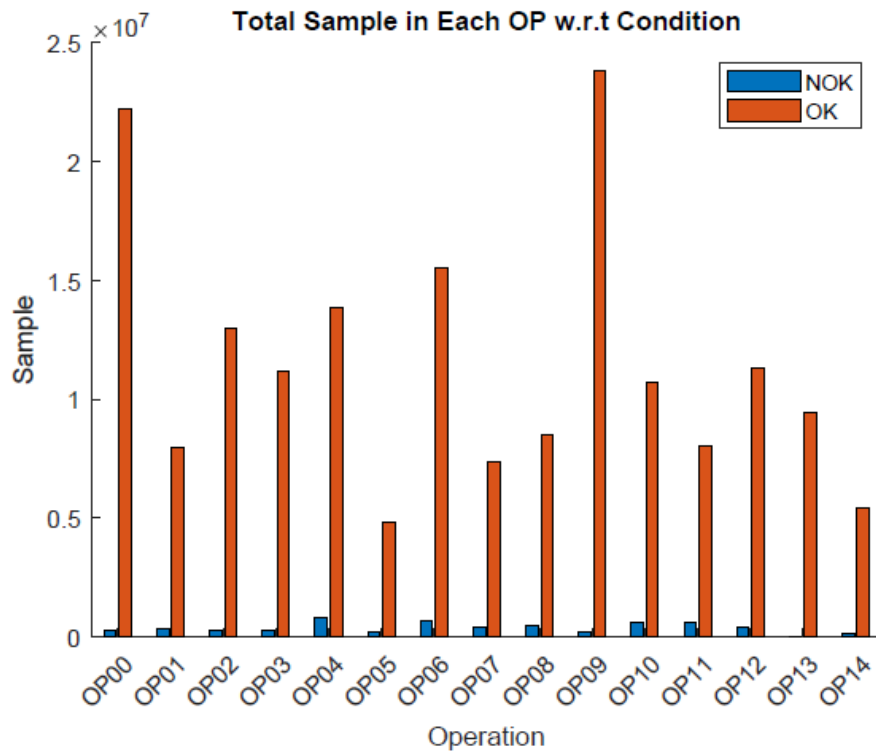


Figure 3 Class Distribution, process health concerning Operation

## **2.2 Data Pre-processing**

Before proceeding to develop an effective model, data pre-processing was done to create biases between OK and NOK data. Initially, to reduce OK samples without avoiding the cutting region, each OK sample is reduced while maintaining sample values that range in its cutting region. Second, each sample is windowed per 2000 sample rate and an overlapping windowing technique per 1000 samples adopt. After this, step-by-step data were analyzed to balance the data w.r.t OK and NOK samples.

5 operations, OP01, OP02, OP04, OP07 and OP10 out of 15 were selected for developing an effective model and the remaining operations data will be used for prediction. Figure 7 shows the Process health distribution w.r.t to Operation after pre-processing of the initial data set.

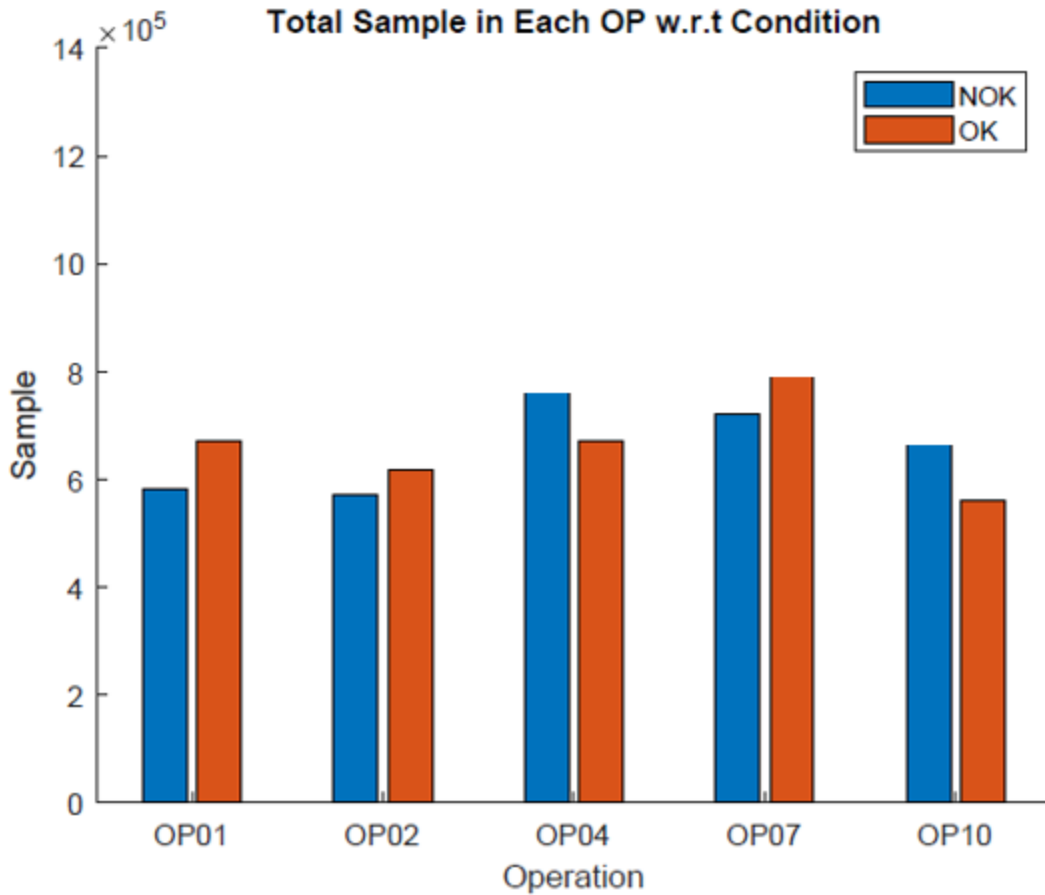


Figure 4 *Process Health distribution w.r.t Operation after balancing data of 5 Operations*

## 2.3 Features Extraction and Model Selection

### 2.3.1 Time Domain Features Extraction

Data acquired is accelerometer data and accelerometer data is typically represented in the time domain. An accelerometer measures acceleration, which is a physical quantity that changes over time. The data captured by an accelerometer sensor provides information about the acceleration experienced by an object or device in the form of a time series signal. Each sample in the signal corresponds to a specific point in time and contains information about the acceleration at that moment. While dealing with accelerometer data, time-domain features give meaningful insights. These feature helps in characterizing the acceleration signals and enable further analysis and model training for machine learning.

Following Time domain features extracted from the signal for further analysis and machine learning model:

- **Mean:** In the context of time-domain signal analysis, the mean represents the central tendency of the signal's amplitude values over a specific time interval.
- **Standard Deviation:** It indicated how much the values deviate from the mean. A high standard deviation implies greater variability in the signal.
- **Root Mean Square (RMS):** In signal processing, RMS provides a measure of the effective amplitude of the signal and is commonly used to quantify signal power or energy.
- **Skewness:** A measure of the asymmetry of a probability distribution.
- **Kurtosis:** A measure of the "tailedness" or shape of a distribution. High kurtosis indicates a distribution with heavier tails and a more pronounced peak, while low kurtosis indicates lighter tails and a flatter peak.
- **Zero Crossing Rate:** The rate at which the signal changes polarity or crosses the zero amplitude level. It provides information about the frequency of the changes in the signal's direction.
- **Shannon Entropy:** In the context of signal, it quantifies the amount of randomness or unpredictability in its amplitude values.
- **Energy:** In signal processing, energy refers to the total magnitude of the signal's amplitude values over a specific time interval. It provides a measure of the overall strength or intensity of the signal.

### 2.3.2 Auto Machine Learning in MATLAB

Before proceeding to auto Machine learning, it is better to know the workflow of the machine learning. Figure 8 shows the workflow of Machine learning.

Machine learning starts with accessing data either by accelerometer data stored offline or any particular live machining data. Once the data is gathered, steps move forward for preprocessing which is cleaning, labeling, and data reduction. In some cases feature extract for a more reliable machine learning model.

After preprocessing the data, develop predictive models like deep learning, statistics, etc. As there are a number of learning models in machine learning, so there are hyperparameters for

each machine learning model for a more robust and accurate model and then proceed to validate the result on the predictive model. Once the end user satisfied with the predictive model, proceeds to deploy algorithm of a selective model for predicting future data.

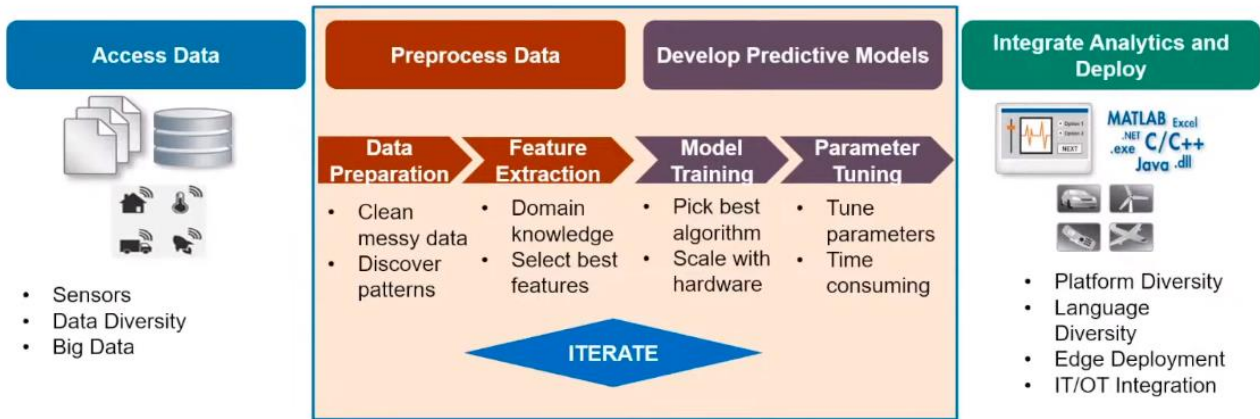


Figure 5 Machine Learning Workflow

As Machine learning gives challenges in developing attractive models because it contains the nth number of iterations before finalizing any best performance model. So Auto Machine learning assists to go through these iterative steps automatically. Figure 9 presents the Auto ML workflow.

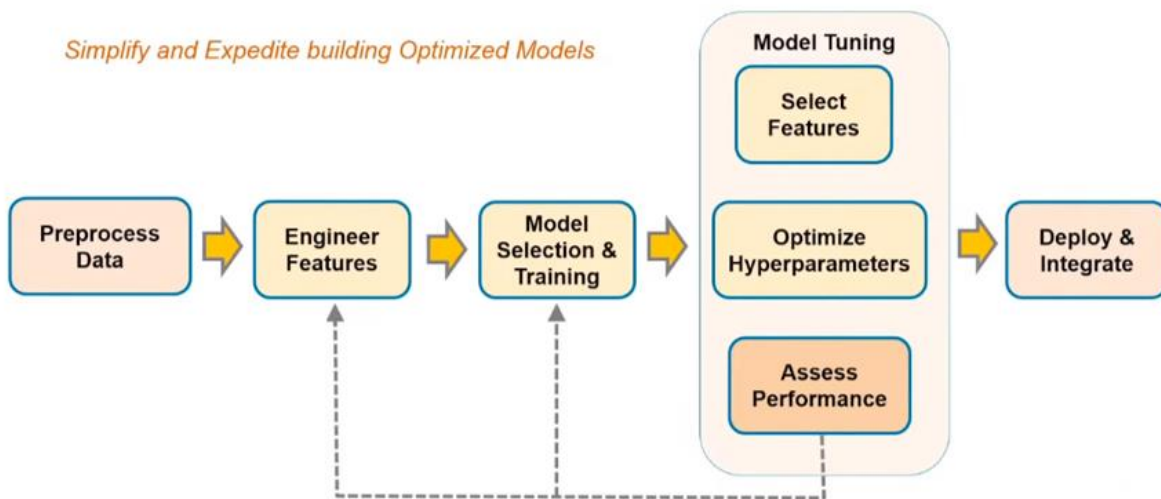


Figure 6 Auto ML Workflow

In Auto ML, after preprocess data iterative selection of a robust model is done automatically, and generalize model is ready to predict future data.

A list of learning Models along with accuracy is mention in the tree of Classification Learner app used for Auto ML, it allows user to select the most accurate model and even allow user to train model on optimizable model type of any model. Optimizable model type feature allows you to perform hyper parameter tuning for various machine learning models.

**Hyper parameters** are parameters that are set before the learning process begins, and they control the behaviors and performance of the machine learning algorithm. Optimizing these hyper parameters can lead to improved model performance.

Optimizable model feature work in the following way in classification learner app :

**Hyperparameter tuning:** It involves finding the best combination of hyperparameter values that result in the optimal model performance.

**Automated Search:** App performs an automated search to find the combination of hyperparameters that leads to the best model performance.

**Grid Search and Bayesian Optimization:** The app uses techniques such as grid search or Bayesian optimization to explore the hyperparameters space efficiently. Bayesian optimization is a more advanced technique that adapts the search based on the results of previous evaluations.

**Cross Validation:** The "Optimizable model" feature often uses cross-validation to estimate the model's performance for different hyperparameter configurations. Cross-validation involves splitting the dataset into training and validation subsets multiple times and averaging the results to get a robust estimate of model performance.

**Best Model Selection:** Once the hyperparameter tuning process is complete, the app identifies the best-performing model based on the specified evaluation metric. This model can then be used for making predictions on new, unseen data.

Using the "Optimizable model" feature can help you find the best set of hyperparameters for a given machine learning algorithm without manual trial and error. This can lead to improved model accuracy and generalization to new data. It's a valuable tool for practitioners who want to achieve the best possible performance from their chosen classification algorithm.

## CHAPTER 3: RESULTS AND CONCLUSION

### 3.1 RESULTS:

Initially, 30 features were selected as predictors and 10 unique operation names with their respective conditions were chosen as the response variable for model learning in the classification learner app for the Auto ML model.

All models in Table 1 were trained to select the best model based on accuracy. The “Cubic SVM”, 'Quadratic SVM', and “Ensemble Bagged Trees” models had the highest accuracies, but none exceeded 90.4%. Optimizable models were trained as a result and their results are also shown in Table 1. An Optimizable Ensemble Machine learning model was ultimately chosen based on its highest accuracy of 90.8% for predicting new data.

*Table 1 Trained Model and their corresponding accuracies*

<u>Trained Model</u>	<u>ACCURACY %</u>	<u>Trained Model</u>	<u>ACCURACY %</u>
<b>Decision Tree</b>		<b>Nearest Neighbor Classifier</b>	
Fine Tree	65.0%	Fine KNN	85.0%
Medium Tree	50.9%	Medium KNN	76.7%
Coarse Tree	36.4%	Coarse KNN	62.7%
<b>Discriminant Analysis</b>		Cosine KNN	77.6%
Linear Discriminant	62.6%	Cubic KNN	74.0%
Quadratic Discriminant	78.3%	Weighted KNN	83.4%
<b>Naïve bayes classifier</b>		<b>Ensemble classifiers</b>	
Gaussian Naïve Bayes	52.4%	Boosted Trees	57.4%
Kernel Naïve Bayes	68.4%	Bagged Trees	86.0%
<b>Support Vector Machine</b>		Subspace Discriminant	57.7%
Liner SVM	67.7%	Subspace KNN	52.1%
Quadratic SVM	87.6%	RUSBoosted Trees	55.2%
Cubic SVM	90.4%	<b>Optimizable Model</b>	
Fine Gaussian SVM	57.8%	Optimizable SVM	88.1%
Medium Gaussian SVM	85.8%	Optimizable Ensemble	90.8%
Coarse Gaussian SVM	63.0%		

#### 3.1.1 Confusion Matrix

The confusion matrix is a valuable tool for assessing classification model performance in both binary and multi-class classification problems. It categorizes classification results into four groups.

- **True Positive (TP):** These are instances that are actually positive and also predicted correctly.
- **True Negative (TN):** These are instances that are truly negative and have been accurately identified as negative by the model.



- **False Positive (FP):** When the model incorrectly predicts a negative instance as positive (Type I error).
- **False Negative (FN):** Instances that are truly positive but mistakenly classified as negative by the model (Type II error).

Confusion Matrix for the trained models corresponding to higher accuracies mentioned in Fig 10.

		Model 1.10									
True Class	OP01_bad	66	1			3	1	1	1		
	OP01_good	1	81				2				
	OP02_bad			70			1	1			
	OP02_good				71				5		1
	OP04_bad	2				156	2	3		2	
	OP04_good	2	4		1	2	69		6		
	OP07_bad		1	2		3	5	75	2	2	
	OP07_good				6		1		104		1
	OP10_bad	3	2			2		2	2	113	1
	OP10_good		1		1	1	2		4	5	56
		OP01_bad	OP01_good	OP02_bad	OP02_good	OP04_bad	OP04_good	OP07_bad	OP07_good	OP10_bad	OP10_good
		Predicted Class									

Figure 7a. Confusion Matrix for Cubic SVM Model, Accuracy 90.4%

		Model 4									
True Class	OP01_bad	67				3		1	2		
	OP01_good	2	78			4					
	OP02_bad	1		70				1			
	OP02_good	1			74				2		
	OP04_bad	3				153	4	3			2
	OP04_good	1	9		2	4	62		6		
	OP07_bad	1				7	4	75	2	1	
	OP07_good				3		1		108		
	OP10_bad	1				3	1	3	1	114	2
	OP10_good		1		1	1	1			3	63
		OP01_bad	OP01_good	OP02_bad	OP02_good	OP04_bad	OP04_good	OP07_bad	OP07_good	OP10_bad	OP10_good
		Predicted Class									

Figure 7b. Confusion Matrix for Optimizable Ensemble Model, Accuracy 90.8%

Figure 7 Confusion matrix of two trained models with higher accuracy

### 3.1.1 Accuracy of models on new data:

These models were then applied to new data to confirm their accuracy. The performance was evaluated using the MacroF1 score. Table 2 Compares the F1 score and accuracy of these two models. Figure 11 shows the confusion matrix. F1 score is a metric commonly used to access the performance of a binary classification model. It takes into account both the precision and recall of the model to provide a single measure of its accuracy in making correct positive predictions while minimizing false positives and false negatives. To evaluate the result of the multi-classification model, macro-averaging or weighted average approach for F1 scores will be evaluated. Macro averaging F1 score, compute the average F1 score across all classes. This gives an overall performance measure that treats each class equally, regardless of class size.

Table 2 Accuracy Result of Selected Model on New Data

Model Name	Accuracy %	F1 Score
Cubic SVM	91.4	0.919
Optimizable Ensemble	90.6	0.918

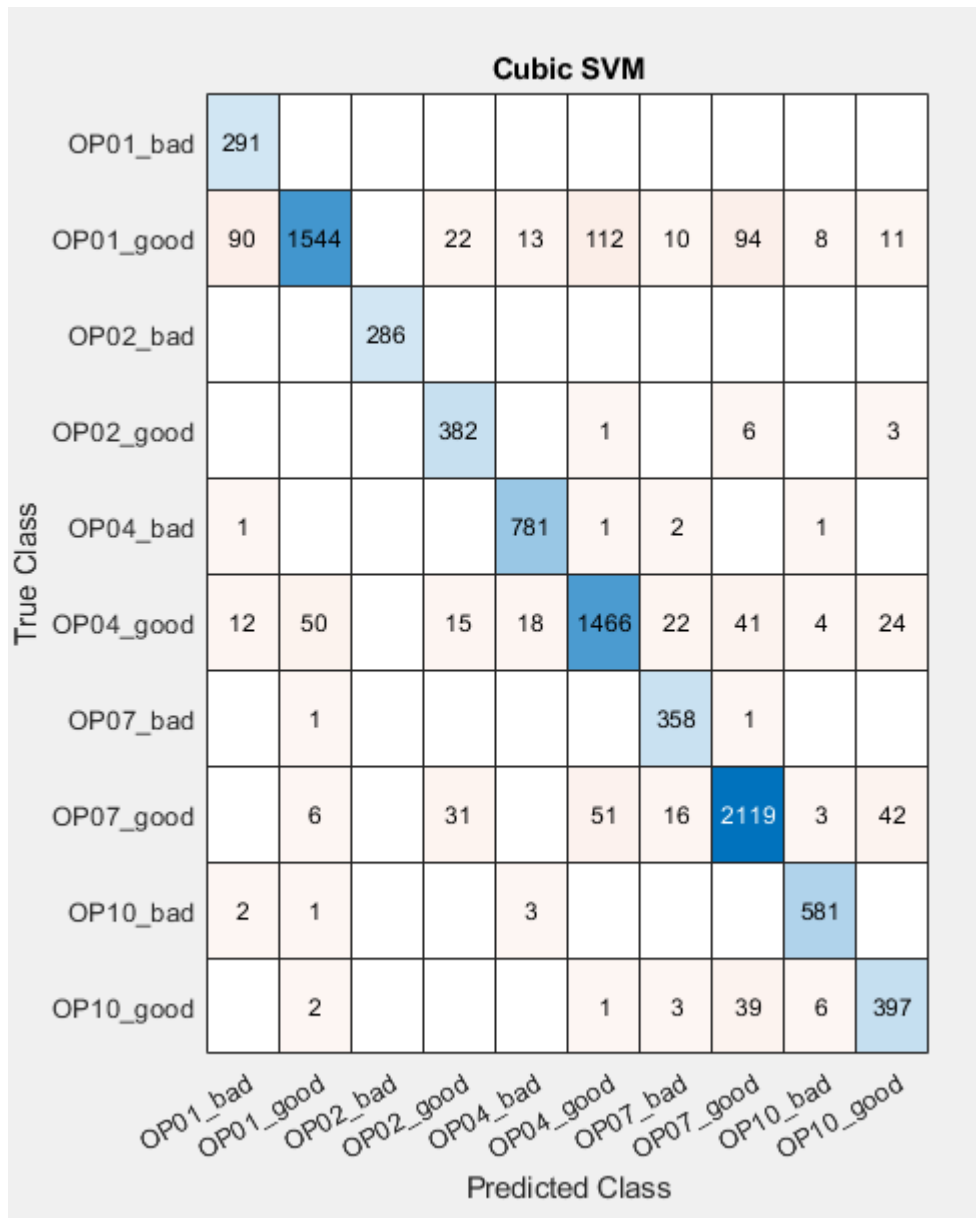


Figure 8a. Confusion Matrix Cubic SVM Model on new data set

		Optimizable Ensemble									
True Class		OP01_bad	OP01_good	OP02_bad	OP02_good	OP04_bad	OP04_good	OP07_bad	OP07_good	OP10_bad	OP10_good
	OP01_bad	291									
OP01_good	52	1517		30	41	113	6	120	1	24	
OP02_bad			286								
OP02_good				390		1					1
OP04_bad					780				1	5	
OP04_good	28	65		30	40	1360	19	79	4	27	
OP07_bad								360			
OP07_good		4		24		43	3	2178			16
OP10_bad	1	1			1					584	
OP10_good					5	1	4	1	49		388
		OP01_bad	OP01_good	OP02_bad	OP02_good	OP04_bad	OP04_good	OP07_bad	OP07_good	OP10_bad	OP10_good
		Predicted Class									

Figure 8b. Confusion Matrix Optimizable Ensemble Model on new data set

Figure 8 Confusion of selected model on new data set

### 3.2 Conclusion

It has been concluded that the accuracy of a trained model can only be confirmed on a new dataset. The accuracies of each model may vary slightly depending on the conditions and dataset. This paper proposes an effective model for predicting class based on operations. The dataset used in this paper [9] can be utilized for future work to predict results based on other variables.

## REFERENCES

- [1] Zhao Zhang, Feng JIANG, Ming LUO, Baohai WU, Dinghua ZHANG, Kai TANG, 2022, “ Geometric error measuring, modeling, and compensation for CNC Machine tools: A review”
- [2] Chandra Nath. Integrated tool condition monitoring systems and their applications: a comprehensive review. *Procedia Manufacturing*, 48:852–863, 2020.
- [3] Zhiyuan Lu, Meiqing Wang, and Wei Dai. Machined surface quality monitoring using a wireless sensory tool holder in the machining process. *Sensors*, 19(8):1847, 2019.
- [4] Vinh Nguyen and Shreyes N Melkote. Manufacturing process monitoring and control in Industry 4.0. In *Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing*. Springer, 144– 155, 2020.
- [5] Daniel Frank Hesser and Bernd Markert. Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. *Manufacturing letters*, 19:1–4, 2019.
- [6] T Mohanraj, S Shankar, R Rajasekar, NR Sakthivel, and Alokesh Pramanik. Tool condition monitoring techniques in milling process review. *Journal of Materials Research and Technology*, 9(1):1032–1042, 2020.
- [7] System-level Manufacturing and Automation Research Testbed (SMART) at the University of Michigan. CNC milling dataset. <https://www.kaggle.com/shasun/tool-wear-detection-in-cnc-mill>, 2018. (Date accessed: 14.12.2021).
- [8] A. Agogino and K. Goebel. Best lab, UC Berkeley. milling data set, nasa ames prognostics data repository. <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>, 2007. (Date accessed: 14.12.2021).
- [9] Mohamed Ali Tnani, Michael Feil, Klaus Diepold, 2022, “Smart Data Collection System for Brownfield CNC Milling Machines: A New Benchmark Dataset for Data-Driven Machine Monitoring”.
- [10] C. Yue, H. Gao, X. Liu, et al., A review of chatter vibration research in milling (Chinese Society of Aeronautics and Astronautics:), *Chin. J. Aeronaut.* 32 (2) (2019) 215–242,
- [11] A.R. AL-OBAIDI, Experimental comparative investigations to evaluate cavitation conditions within a centrifugal pump based on vibration and acoustic analyses techniques (Polska Akademia Nauk:), *Arch. Acoust.* 45 (3) (2020) 541–556,

- [12] Rohit Mishra, Matta S.N.S. Kiran, Manikantadhar Maheswaram, Akshat Upadhyay, Bhagat Singh, “Investigation of optimal feature for milling chatter identification using supervised machine learning techniques”
- [13] A.V. Joshi, Machine learning and artificial intelligence, (2020).
- [14] L.C. Moreira, W. Li, X. Lu, M.E. Fitzpatrick, Supervision controller for real-time surface quality assurance in CNC machining using artificial intelligence, *Comput. Ind. Eng.* 127 (2019) 158–168
- [15] Z.M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, B. Safaei, Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0, *Sustainability* 12 (2020) 8211.
- [16] P. Krishnakumar, K. Rameshkumar, K. Ramachandran, Acoustic emission-based tool condition classification in a precision high-speed machining of titanium alloy: a machine learning approach, *Int. J. Comput. Intell. Appl.* 17 (2018) 1850017.
- [17] L. Li, C. Li, Y. Tang, Q. Yi, Influence factors and operational strategies for energy efficiency improvement of CNC machining, *J. Clean. Prod.* 161 (2017) 220–238.
- [18] S. Nallusamy, Enhancement of productivity and efficiency of CNC machines in a small scale industry using total productive maintenance, *Int. J. Eng. Res. Africa, Trans. Tech. Publ.* (2016) 119–126.
- [19] Y. Yang, T. Hu, Y. Ye, W. Gao, C. Zhang, A knowledge generation mechanism of machining process planning using cloud technology, *J. Ambient Intell. Humans. Comput.* 10 (2019) 1081–1092.
- [20] Bosch Connected Devices and Solutions GmbH. Connected industrial sensor solution. [https://www.bosch-connectivity.com/media/downloads/ciss/ciss\\_datasheet.pdf](https://www.bosch-connectivity.com/media/downloads/ciss/ciss_datasheet.pdf), 2020. (Date accessed: 14.12.2021).
- [21] Yang Hui, Xuesong Mei, Gedong Jiang, Tao Tao, Changyu Pei, and Ziwei Ma. Milling tool wear state recognition by vibration signal using a stacked generalization ensemble model. *Shock and Vibration*, 2019.
- [22] Grzegorz Wszolek, Piotr Czop, Jakub Słoniewski, and Halit Dogrusoz. Vibration monitoring of cnc machinery using mems sensors. *Journal of Vibration Engineering*, 22(3):735–750, 2020.

- [23] Claude Elwood Shannon. Communication in the presence of noise. Proceedings of the IRE, 37(1):10–21, Jan 1949.