JJMIE

Volume 16, Number 5, December. 2022
ISSN 1995-6665
Pages 821– 833

Jordan Journal of Mechanical and Industrial Engineering

A Conceptual Framework for Cyber-Physical Quality Monitoring System using Machine Learning

Mathew Chacko[1,*] , Atul[2] , Satish Babu Boddapati[3]

[1,2] Mechanical Department, Alliance University, India

[3] Department of Petroleum Engineering, RGIPT, Amethi, India

Received 1 Aug 2022 Accepted 5 Nov 2022

Abstract

In this paper, we propose a Cyber-Physical Quality System (CPQS) integrated framework that can predict, analyze, and validate the quality monitoring system in manufacturing with 95% accuracy in real-time using machine learning techniques. CPQS framework analyses real-time sensor networks and configures the importance of artificial intelligence-driven big data analytics for predicting the quality of cyber-physical production networks. Cyber-physical data like speed, feed, depth of cut, coolant temperature, vibrations, tangential cutting forces, and tool life for 400 parts were collected from the various sensors placed on Computerized Numerical Control (CNC) machines after doing modal analysis. Various machine learning techniques were used to predict the quality of the part wherein the inputs affecting it were predominately dominated by vibration and temperature.

Extreme Gradient Boosting (XGB) machine learning techniques out of many could predict the quality of the part with 96.2%. accuracy. The caveat for the present results is that it has been tried out only for Titanium Alloy parts and the tool wear has been approximated using Taylor's equation which can be enhanced by using image processing. The model deployed in real-time could produce defect-free parts quickly. This could reduce the cost of quality by 80%, thereby increasing the production line's productivity, quality, and efficiency.

Keywords: Cyber-Physical Systems, Machine Learning, Modal Analysis, Tangential Cutting Forces.

1. Introduction

The prediction of the quality of a machined part, while it is being manufactured, saves unnecessary time and money spent on inspection in the Cyber-Physical Quality System (CPQS) [1-2]. The accurate prediction of the quality of a machined part, while it is being manufactured, prevents quality failures in real-time and maintains compliance. It optimizes material usage, quickly isolates defects, and increases contribution margins [3-4]. This will help manufacturing companies stay competitive because the part can be manufactured quickly, at low cost, and with high quality. This study deals with the implementation of CPQS in a real factory setting to produce maximum parts and predict the quality of manufacturing of a flange in a CNC milling machine using different machine learning techniques. [5]. Industry 4.0 envisions new technologies [6], like the Internet of Things (IoT), Cyber-Physical-Systems (CPS), Big Data, High-Performance Computing (HPC), Edge Computing, and Cloud Computing for setting up a Digital Twin Shop-Floor that encompass an effective way to create the physical-virtual convergence of the real, virtual world and their connections. With the digital twin [7] serving as a digital controller of the real-world manufacturing system we can use applications powered by Artificial Intelligence (AI) [9] to understand the manufacturing parameters that affect quality. With AI technology becoming more mature and affordable, new applications have been introduced into production systems to support manufacturers in complex decision-making and business processes [10]. Machine Learning (ML), a subset of AI, focuses on extracting useful knowledge [11] into the bottleneck of the problem through the learning and training process with a large volume of both structured and unstructured data [12-14]. Some of the machine learning applications used in manufacturing process diagnosis involve Logistic Regression, K Nearest Neighbor's (KNN), Support vector machine classifier (SVC), Gaussian Naïve Bayes (GNB), Decision Tree, Random Forest, Extreme Gradient Boosting (XGB), and Multi-Layer Perceptron (MLP) techniques [15-16]. These help in optimizing the manufacturing process for better quality and cheaper costs. Cyber-physical systems are also widely used in the manufacturing and processing industry to monitor product quality in real time. Computer vision systems are used to control robots, CNC machines, conveyors, and other equipment in the autonomous production line to automatically detect anomalies during machining [17]. The various modeling techniques that

* Corresponding author e-mail: chacko1959@yahoo.com.

have been used include Petri Nets (PNs) which simulate various types of asynchronous and concurrent processes in an industrial production system [18]. Smart Partial Least Squares/ Statistical Package for Social Sciences (PLS/SPSS) techniques are used to identify the correlation functions between the input and output parameters and machine-learning techniques are used to identify the anomalies.

1.1. Literature Survey

Manufacturing quality control is the ability to measure parts and provide assurance that parts have been produced according to their specifications. This is achieved by inspecting all incoming raw materials and establishing inspection points along the manufacturing process, to ensure part quality and spot any quality variations before they have an impact on the dimensional accuracy and surface finish. Broadly all research papers written so far can be classified as either Theoretical, Practical, or Simulation type based on the parameters of dimensional quality, surface finish, or both.

Table 1. Classification of papers

Focus Area / Type of Research	Dimensional Quality	Surface finish
Theoretical	[19], [29], [23], [24], [25],[27],[28]	[21], [22], [26]
Experimental	[31],[33], [34], [35], [36]	[29] [32], [35]
Simulation	[37], [38], [39], [40],[41],[43],[43],[44],	

[ ] Reference papers

1.1.1. Theoretical:

Some of the theoretical studies done on this topic include

1. Developed a methodology for the development of an intelligent quality function deployment (IQFD)application for the Manufacturing Process. [19]
2. Review the challenges and limitations of the optimization techniques used in optimizing machining parameters in milling operations. [20]
3. Studied the correlation between primary waviness and roughness during hard turning through mathematical modeling. [21]
4. A paper that talks of various surface quality improvement techniques, including how to reduce surface defects, surface roughness, and dimensional inaccuracy [22]
5. b) Study on dimensional quality and distortion analysis of thin-walled alloy parts [23]
6. c) Study on combining online testing, sensor, network, and database technologies and quality control methods to realize the online process quality control system. [24]
7. Quality control methods for product reliability and safety using optimization techniques [25]
8. In this study, mathematical models were developed that established the correlation between input variables and quality characteristics in the plasma Computer Aided Manufacturing (CAM) process using Response Surface Methodology (RSM). [26]
9. Review paper on studies conducted on the interoperability between Internet of Things-based real-time production logistics and cyber-physical process monitoring systems. [27]
10. This paper reviews the current research on the Internet of Things-based real-time production logistics, sustainable industrial value creation, and artificial intelligence-driven big data analytics in cyber-physical smart manufacturing systems. [28]

1.1.2. Experimental:

Some of the experimental studies conducted include

1. A detailed study of the effects of machining parameters on the surface roughness in the end-milling process. [29]
2. Optimization of surface roughness in end milling using the Response Surface Method and Radian Basis Function Network. [30]
3. Setting up a web-based automated inspection of manufactured parts wherein they developed a platform to study the quality parameters. [31]
4. Select process parameters based on the Taguchi orthogonal array technique and use the analysis of variance (ANOVA) to establish a relationship between input parameters and surface roughness as output characteristics. [32]
5. Tool wears monitoring using in-process machine vision for Cyber-Physical Production Systems (CPPS): The author of this study [33] proposes a four-phased approach based on the CPS for in-process tool wear monitoring using machine vision.
6. Using a machine vision system to measure
7. tool wear parameters: In this study, they captured tool wear photographs with digital cameras and used image processing techniques to determine the tool wear zone to take necessary action. [34].
8. Data from a spindle probe, a coordinate measuring machine, and surface roughness data are used to characterize machine quality features, namely dimensional accuracy, and surface roughness. [35].
9. The impact of variables cutting speed, feed rate, depth of cut cooling method, blank size, and work material on the dimensional accuracy and surface quality of turned parts were investigated in this study. [36].

1.1.3. Simulation:

Some of the studies done about simulation are to do with understanding the relationship between deep learning-assisted smart process planning and Internet of Things-based real-time production logistics as regards cyber-physical smart manufacturing systems. This include

1. Design of a model for turning involving a neural network controller to track the desired vibration level of the turning machine. [37]
2. The Service-Oriented Cross-layer infRAstructure for Distributed smart Embedded devices (SOCRADES) is an initiative to achieve predetermined automation goals in which networked systems made up of smart embedded devices, collected data from a service-oriented ecosystem. [38]
3. Integration of process and quality control using multi-agent technology (MAT) (Ref: GRACE European Project). Using multi-agent system (MAS) principles,

performed real-time data analysis to dynamically modify production factors wherein concepts like Product Type Agents (PTA), Product Agents (PA), Resource Agents (RA), Independent Meta Agents (IMA), and other dynamic self-adaptation techniques along with feedback control loops were implemented. [39]

4. Adaptive Production Management (see: ARUM European Project). The project developed production planning, scheduling, and optimization strategies using agent technology to respond to anomalies according to Service-Oriented Architecture (SOA) principles. [40]
5. Deterministic models also have traditionally tested extraordinarily beneficial during the current industrial revolution involving digital data. Key deterministic models for distributed cyber-bodily systems have sensible faithful realizations through Cyber-physical structures. [41].
6. 5C architecture for the implementation of a CPS involves (i) Smart Connection (ii) Data-to-information conversion (iii) Cyber level (iv) Cognition (v) Configuration [42].
7. Empirical studies on the IoT- based real-time production logistics, cyber-physical process monitoring systems, and industrial artificial intelligence in sustainable smart manufacturing. [43]
8. Simulation studies and analyses on how data-driven supervision, predictive analytics, and optimization systems integrate product traceability, maintenance, and process performance in smart manufacturing. [44]

As per the literature survey done above, there is a need to analyze real-time sensor networks and configure the importance of artificial intelligence-driven big data analytics for use in cyber-physical production networks which is what this paper intends to do.

The rest of the paper is organized as follows. In section 2, the machining process parameters and their responses are discussed. Section 3 describes the proposed approach to developing a cyber-physical quality system. Section 4 gives a sneak preview of the experimental setup. Section 5 is all about data extraction and analysis. Results are explained in section 6. Conclusions are shown in section 7.

2. Machining process parameters and responses

The behavior of the product is depicted in Figure 1. below

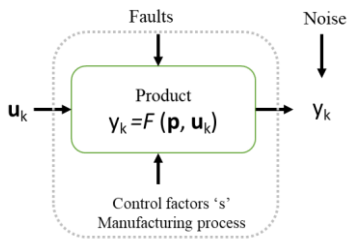


Figure 1. The general scheme of the product

can be described by the following relation

$$yk = F(p, uk) + \varepsilon k \quad (1)$$

where uk and yk are product inputs and outputs, respectively. p are parameters representing the physical characteristics of product components. Control factor s influence these values, in the production process. F(.)is the relation between inputs, output, and parameters describing the behavior or properties of the product and εk represents the noise

The control factors 's' involved in CNC manufacturing are:

- Feed rate: The feed rate CNC parameter is the speed at which the cutter moves across the face of the material. It is measured in distance units per minute (e.g. millimeters per minute, or inches per minute).
- Plunge rate: Plunge rate is the speed at which the bit enters the material, meaning that this CNC parameter affects only pure vertical movement. It is measured in distance units per minute (e.g. millimeters per minute, or inches per minute).
- Depth per pass: Usually, a non-industrial CNC machine doesn't have enough power to cut through all the material thickness in a single attempt, unless you are cutting a soft and thin piece of material with a large bit. That's why your project will likely require multiple passes to get the desired depth. The CNC parameter – depth per pass dictates how deeply your machine carves down into your material on each pass.
- Spindle speed: Spindle speed is the speed at which your cutting tool rotates. It is measured in RPM (revolutions per minute).

2.1. Surface roughness

The theoretical one-dimensional expression of surface roughness R for a surface of profile length d is

$$R = \frac{1}{d}\int_0^d |f(x)|dx \quad (2)$$

where f(x) is the difference between the local surface height at position x and the mean height over the profile based on the assumption that the overall profile is even. If the height f_n is measured at N locations along with the profile length d, the expression of the roughness is:

$$R \approx \frac{1}{N}\sum_{i=1}^{n} |f_n| \quad (3)$$

Converting the expression of surface roughness to a two-dimensional surface profile area A. The surface roughness of area A with N × M tested differences f_{ij} can be approximated as:

$$R \approx \frac{1}{NM}\sum_{i=1}^{N}\sum_{i=1}^{M} |f_{ij}| \quad (4)$$

2.2. Dimensional accuracy

Though CNC machining has the ability of manufacturing parts with complicated shapes, the dimensional accuracy is a limitation of the machining parameters like speed, feed, depth of cut, coolant temperature, material properties, and tool condition.

Dimensional accuracy is the deviation between the nominal size and the measured size of an as-built part. In this work, the lengths (L), widths (W), and heights (H) of the Computer-Aided Design (CAD) model of the part and the built part is used to define the dimensional accuracy [18]. The expressions of dimensional accuracy are

$$DL = |Ld - Le| \quad (4)$$

$$DW = |Wd - We| \quad (5)$$

$$DH = |Hd - He| \quad (6)$$

where DL, DW, and DH denote the deviations, and the subscripts ‘d’ and ‘e’ denote the nominal size from the CAD model and the measured size from the as-built part, respectively.

3. Proposed Approach

A Cyber-Physical Quality System (CPQS) is an integration of computation, networking, and physical processes for measuring the quality of manufactured parts. It involves embedded computers and sensory networks that monitor and control the physical processes, with feedback loops, wherein physical processes affect computations and vice versa. The first step involves designing a CPQS architecture framework based on the collaboration among, manufacturing execution systems, the internet of things (IoT), simulations using artificial intelligence, advanced need quality systems for quality prediction, and operation control as in Figure 2.

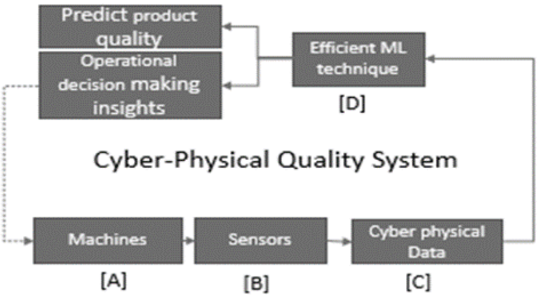


Figure. 2 The cyber-physical quality system framework

Based on this framework we can create a reconfigurable quality system as shown below in Figure.3.

which can predict the quality of the component being manufactured and appropriately take actions to ensure a defect-free production. This reconfigurable quality management system involves collecting the data from the CNC controllers, multi-sensory systems, and the local terminal which monitors the machining processes and transmits it to the database server. The features from these data and the signals are its time /frequency are then extracted and using different machine learning techniques are used to predict the tool wear, and surface finish and thereby accurately predicting the quality of the part being manufactured. This leads to producing quality parts and in case of anomalies taking actions to change the tool or stop the process thereby ensuring quality output at all times. The various environments involved in such a setup are described in detail.

3.1. Machining Environment

The manufacturing environment consists of the process, machine, tools, coolant, the part to be machined, and inspection gadgets.

3.1.1. Part

The part to be manufactured is the Hinge as shown in Figure 4

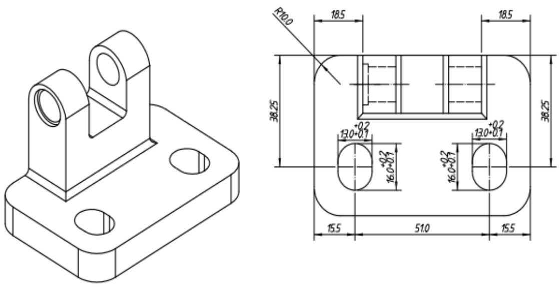


Figure 4. Hinge Type A

This is used to support the doors of any vehicle. It consists of two numbers per door and is different for the left vs right side of the vehicle. Its dimensions are crucial and they have to be within the tolerance limits to ensure that the doors close properly. 25 crucial parameters need to be checked for dimensions and the surface quality has to be within the permissible limits for the part to be adjudged as having passed the quality test

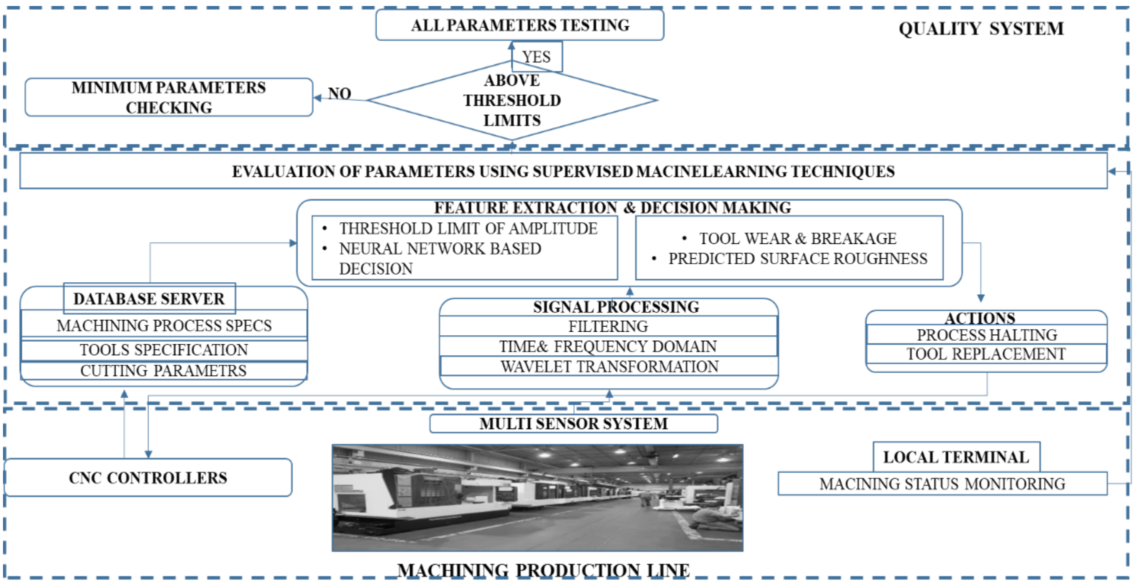


Figure 3. Reconfigurable quality management system

3.1.2. Machine

The machine used for manufacturing this part is the most versatile OKADA VM500 milling machine fitted with all the sensors and cameras as shown in Figure 5.

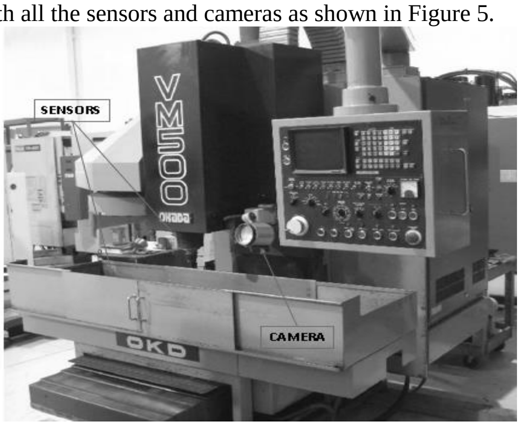


Figure 5. OKADA VM500 machine with sensors and cameras

3.1.3. Tools

To manufacture the part, we use mainly 3 tools. We use the milling cutter having carbide inserts (T1) for all the milling operations. This tool has Taylor's tool life exponent n = 0.143 and the constant C=48.1 for a cutting speed of 30m/sec. and is changed typically, after producing 50 components. The second tool (T2) is HSS used for all drilling operations. This tool has Taylor’s tool life exponent n = 0.2 and the constant C=63.53 for a cutting speed of 30m/sec. This tool is changed after every 40 components are produced. The last tool (T3) is a forming/chamfering tool with Taylor's tool life exponent n = 0.143 and the constant C=48.1 for a cutting speed of 30m/second is used for making 80 components before it is re-sharpened or a new tool is replaced.

3.2. Sensor Environment

The various sensors used are for temperature, displacements, and vibration. Three types of vibrations are generated throughout the turning process: free, forced, and self-excited. These vibrations are caused by the machine tool system's lack of dynamic stiffness/rigidity, work material, machine, tool, and holder. Free vibrations are caused by shock, while forced vibrations are caused by machine tool imbalance, misalignment, mechanical rigidity, and gear faults. Frictional chatter is caused by rubbing on the clearance face, which causes vibration in the cutting force (Fc) and thrust force (Ft) directions. Temperature and strain rate in the plastic cause thermo-mechanical chatter. The presence of chatter has the following negative consequences: Poor surface quality, tool wear and damage, lower rate of material removal, waste of time, effort, energy, and higher costs in terms of production time. Hence the vibration sensors must be placed rightly in the right places for the experiments to be meaningful accordingly we need to do the vibration analysis of the spindle shown in Figure 6.

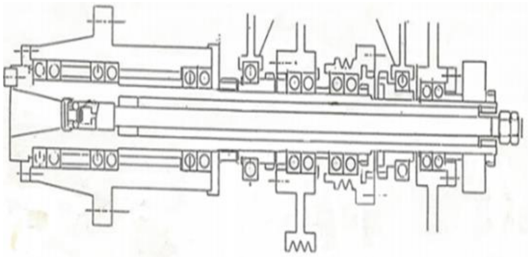


Figure 6. Schematic diagram of the spindle

3.2.1. 3D Model of the spindle

To carry out the modal analysis we need to first model the spindle, bearing, tool holder, and tool using the Computer-Aided Three dimensional Interactive Application (CATIA) software. A model of the same is shown in Figure 7.

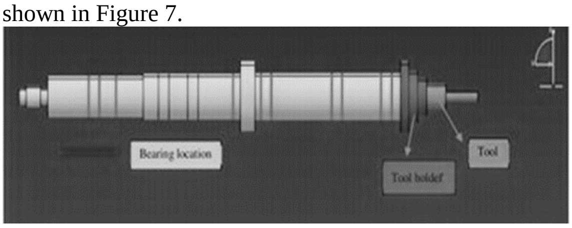
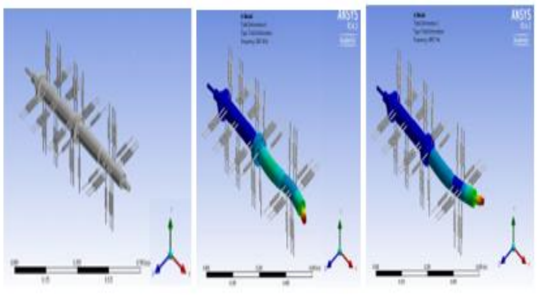


Figure 7. Model in CATIA

3.2.2. Modal analysis of the spindle

The spindle system model was imported into ANSYS® software and we used BEAM188 and SOLID187 elements. COMBIN14 elements for the bearings. The material property was that of tool steel with young’s modulus, E = 210 GPa, and density = 7850 Kg/m3 applied. four spring-damper elements replaced each bearing location as shown in Figure 4. At k = 2.25 x 108 N/m is the natural frequency. The first natural frequency was zero as the spindle was not constrained along the longitudinal translational direction, the second/third, fourth/fifth, and sixth/seventh natural frequencies represent the bending of the spindle in Z/Y directions respectively, and the eighth natural frequency represents the torsional vibration of the spindle, as shown in Figure 8.



(a) **(b)** **(c)**

Figure 8. a) Solid with spring-damper b)8th c) 2nd -5th natural frequencies

Based on the above analysis the location to fix the vibration sensor was finalized which would give us the best output.

3.3. Digital Environment

The digital environment consists of two phases namely the development of a prediction model and quality performance prediction using machine learning techniques as shown in Figure 9

The steps involved in the prediction model development involve determining the Critical To Quality (CTQ) of the component to be manufactured and identifying the significant parameters that impact CTQ. The next step would be to determine the location of the sensors and build the model with the appropriate transfer functions determined through machine learning techniques

The performance metrics in the classification problem are different from regression problems. In classification, accuracy is a basic metric and more robust metrics are F1 score, Precision, and Recall [45]. The proper metric is selected based on the problem that is dealt with. In this scenario, the target is to predict the occurrence of output 0, so that we can make the required changes in the input features so that the overall cost of production is lowest. Thus, Recall and Precision of classification 0 are more important than accuracy. Recall gives the fraction of 0's that the model can identify. Precision gives the fraction of relevant instances to retrieved instances [19]. We need high recall to identify all the Not ok quality manufactured pieces and the max possible precision.

Models with accuracy > 79% for Dimension prediction and models with accuracy > 86.25% for surface finish are reported in the Results section because any lesser accuracy means that any model giving 1's to all input can achieve the cut-off accuracy levels.

4. Experimental Set-Up

The overall experimental setup consists of the physical resources namely machines, tools, and sensors. Along with that, we have the local and database servers where the data of the processes and pre-processing of the sensor signals are done. We use cloud services for advanced signal processing and cognitive decisions making as shown in Figure 10.

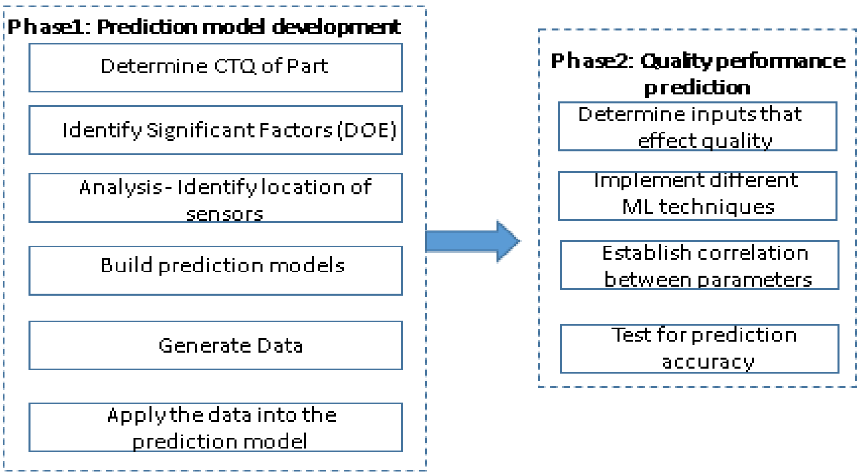


Figure 9. Stages of prediction model development and quality

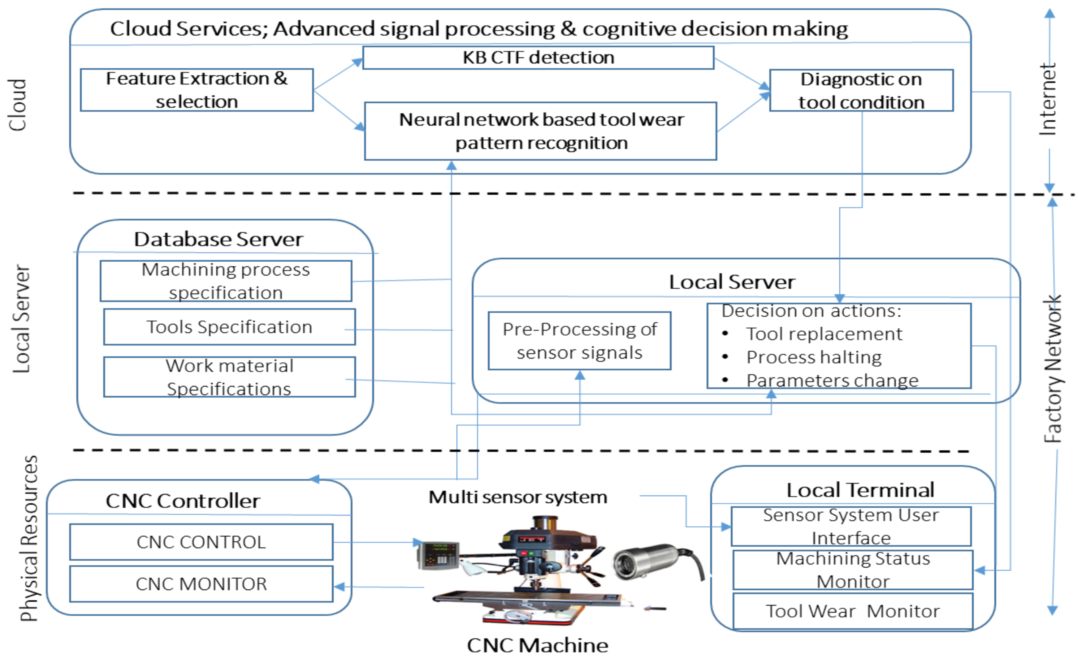


Figure 10. Experimental set-up

5. Feature Extraction and Analysis

5.1. Understanding Data

The data is collected for 400 flanges manufactured by varying different input parameters and values of target variables are decided after physical observation. Data is of dimension 400*11. 400 rows with 11 columns that contain 8 input features and 3 target variables.

Input features are speed (m/min), Feed (mm), depth (mm), coolant temperature (Celsius), Vibration (mm/sec), and the tool life of T1, T2, and T3 tools measured in min. Target variables include dimension quality, quality of surface finish, and final quality all taking the values 1 or 0, interpreted as Quality is ok, and Quality is not Ok respectively. The final quality is 1 if both the dimensional quality and quality of the surface finish are 1, else 0. In this paper, we try to predict the independent target variables, since the final quality can be calculated based on these 2 target variables. Statistical analysis of the results was done, to ensure that the differences in performance are statistically significant or not by using Friedman Aligned Rank Test using IBM SPSS gave the following results as in Table 2.

Table 2. Friedman Test statistics

Test Statistics[a]

N	400
Chi-Square	9851.329
df	27
Asymp. Sig.	<.001

a. Friedman Test

wherein it was observed that the p-value < 0.05.

5.2. Data Preparation

The data is clean devoid of any null values and there are no irregular data types. The next step is to check for outliers. Box plot is constructed for the input parameters and observed for any outlier values based on [46]. Only the Coolant temperature is found to have 16 points as outliers where the temperature is less than 28.5 C as shown in Figure 11.

5.2.1. Coolant temperature distribution

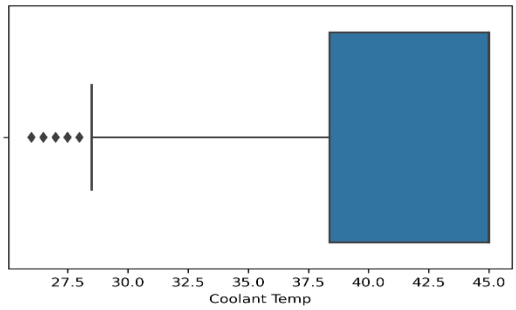


Figure 11. Box plot for Outlier detection

These 16 points correspond to 4% of the data and their effect on the result is negligible. Hence no manipulation is made of the data. The reason is that this temperature reflects the starting point of the CNC machine and as the machine's working time progresses, the temperature increases [47].

5.3. Data Visualization

Pair plots and correlation heat maps [48] are made for the data to identify dependencies and pattern recognition. The correlation heat map is shown in Figure 12.

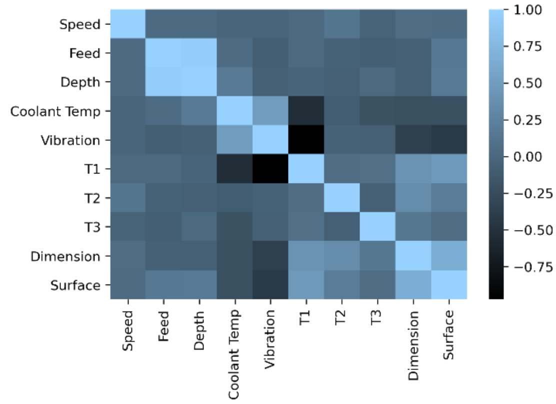


Figure 12. Correlation heat map

From the heat map, it is observed that dimensional quality is highly correlated to T1, and T2 tool life & vibration, and weakly correlated to coolant temperature. Surface finish quality is highly correlated to T1 & vibration and weakly correlated to coolant temperature and T2. Distribution plots for the input features are prepared. To understand the data better and check for the usefulness of features in predicting the target variables, kde (kernel density estimation) of the parameters T1, T2, and Vibration are shown in Figure.13 and Figure.14 using target variables as hue. As predicted high density of target variable 0 is observed at a high level of vibration and low tool life.

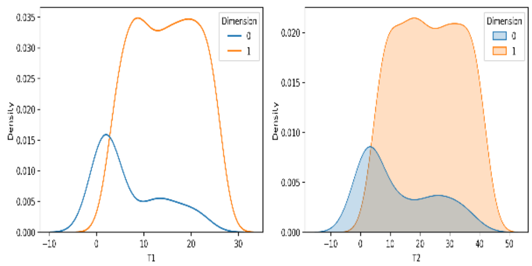


Figure 13. Kde Plots of T1 and T2 using dimension as hue.

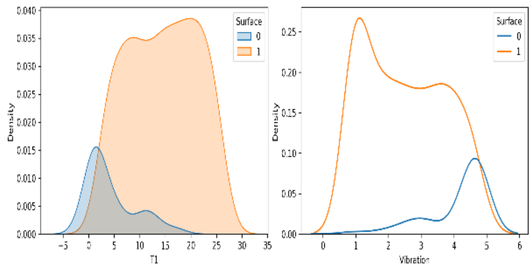


Figure 14. Kde Plots of T1 and Vibration using surface quality as hue

5.4. Data Modeling

This is a case of multi-label classification supervised learning. For optimal performance and high control of the process, the target variables are modeled independently and in two ways in each case. The results are compared to finalize the best model and the final quality is predicted by combining the predicted targets from the selected best models. The summary is shown in Table 3.

Table 3. Input features and target predicted for each method

Method	Input features used	Target predicted
1	T1, T2, and Vibration	Dimension quality
2	All 8 parameters	Dimension quality
3	T1 and vibration	Surface finish
4	All 8 parameters	Surface finish

The first step of modeling is to check if the data set is balanced or not. The count plot for Dimension and surface finish is shown in Figure 15. Category 0 implies the quality is not ok, while 1 implies ok. There is an imbalance in the data set but the modeling performed with stratified training and test data gave good results. Thus, no special techniques are used for further data processing.

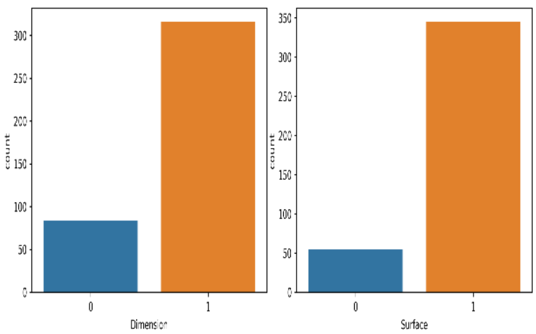


Figure 15. Count plot for Dimension and Surface finish

Dimension has 84-0's and 316-1's resulting in 21% minority class, while the Surface finish has 55-0's and 345-1's giving 13.75% as a minority class.

5.5. Data Analysis

The data is split using sklearn [49]test train split, test size =0.3, random state=42, and stratified using output variable considered in each method and then different models are applied to training data and the results are validated on the test set. The machine learning techniques evaluated are Logistic Regression, K Nearest Neighbors (KNN), Support Vector Machine Classifier (SVC), Gaussian Naïve Bayes (GNB), Decision Tree, Random Forest, Extreme Gradient Boosting (XGB), Multi-Layer Perceptron (MLP) classifiers and Artificial Neural Networks (ANN).

5.5.1. Dimensional accuracy

Dimensional accuracy is measured using method 1 and method 2 mentioned above with a cut-off of 79%.

1. *Method-1*: The performance of these models for Recall, F1 score, and Accuracy are shown in Figure 16. concerning input parameters T1, T2, and Vibration. To see that the models are not overfitted, Grid search CV is used to find optimized hyperparameters of the model and are validated on test data. In some cases, there are already available techniques in the model used instead of Grid search CV.

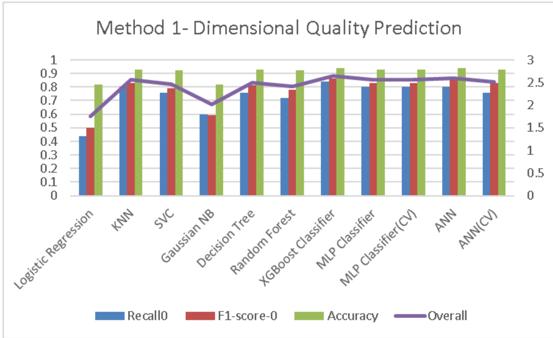


Figure 16. Plot for Dimensional accuracy using different ML techniques

It can be seen that Logistic regression and Gaussian Naïve Bayes have accuracies higher than cut-off accuracies, however, the Recall and F1 scores are poor. K Nearest Neighbours (KNN),

SVC, Decision tree, Random Forest, MLP, and ANN produced similar results but the best was given by XG Boost with values of 0.84for Recall,0.86 for F1 score and 0.96 for accuracy.

2. *Method- 2: The performance of these models concerning Recall, F1 score, and Accuracy are shown in Figure 17. for all 8 input parameters.*

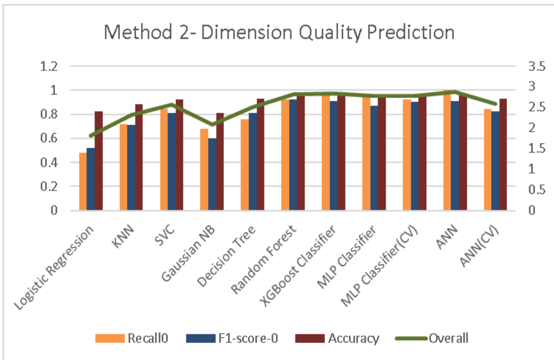


Figure 17. Plot for Dimensional accuracy using different ML techniques- method2

Here also we see that Logistic regression and Gaussian Naïve Bayes have accuracies higher than cut-off accuracies, however, the Recall and F1 scores are poor. K Nearest Neighbours (KNN), SVC, and Decision tree produced similar results, and Random Forest and MLP gave better results but again XG Boost and ANN gave marginally better results of 0.96for Recall,0.91 for F1 score, and 0.96 for accuracy.

5.5.2. Surface Finish

Surface Finish is measured using method 3 and method 4 mentioned above with a cut-off at 86.25%.

1. *Method-3*: The performance of these models for Recall, F1 score, and Accuracy are shown in Figure 18. for input parameters T1and Vibration. To see that the models are not overfitted, Grid search CV is used to find optimized hyperparameters of the model and are

validated on test data. In some cases, there are already available techniques in the model used instead of Grid search CV

It can be seen that Logistic regression. K Nearest Neighbours (KNN), Decision tree and Random Forest have accuracies higher than cut-off accuracies, however, the Recall and F1 scores are poor, SVC, MLP, XG Boost, and ANN produced similar results of 0.59for Recall,0.71 for F1 score and 0.93 for accuracy. All of them are equally good at predicting.

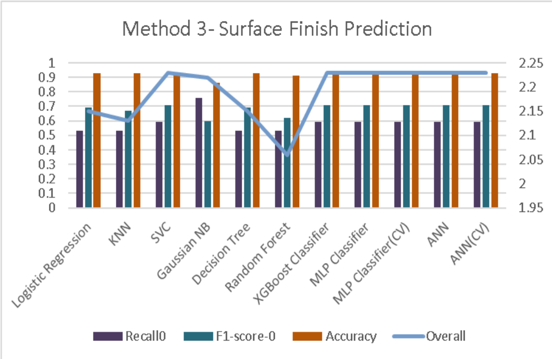


Figure 18. Plot for Surface finish prediction using different ML techniques- method 3

2. *Method- 4: The performance of ML models for Recall, F1 score, and Accuracy are shown in Figure 19. for all 8 input parameters.*

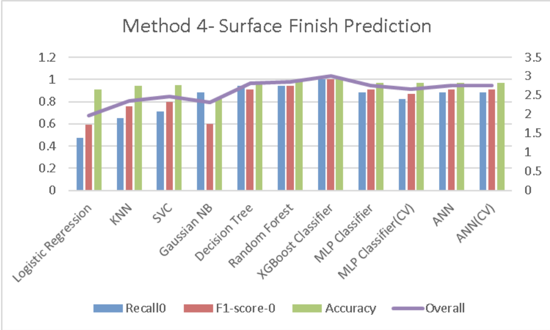


Figure 19. Plot for Surface finish prediction using different ML techniques- method 4

It can be seen that Logistic regression. K Nearest Neighbours (KNN), SVC, and Gaussian NB have accuracies higher than cut-off accuracies, however, the Recall and F1 scores are poor, Decision tree, Random Forest, MLP, and ANN produced similar results, but the best was given by XG Boost with values of 1.0for Recall, 1.0 for F1 score and 1.0 for accuracy.

It can be concluded from the above that Extreme Gradient Boosting as shown in Figure 20. is a supervised learning algorithm that is similar to RF, that tries to accurately predict a target variable by combining the estimates of a set of simpler, weaker models.is the best for predicting both the dimensional accuracy and surface finish.

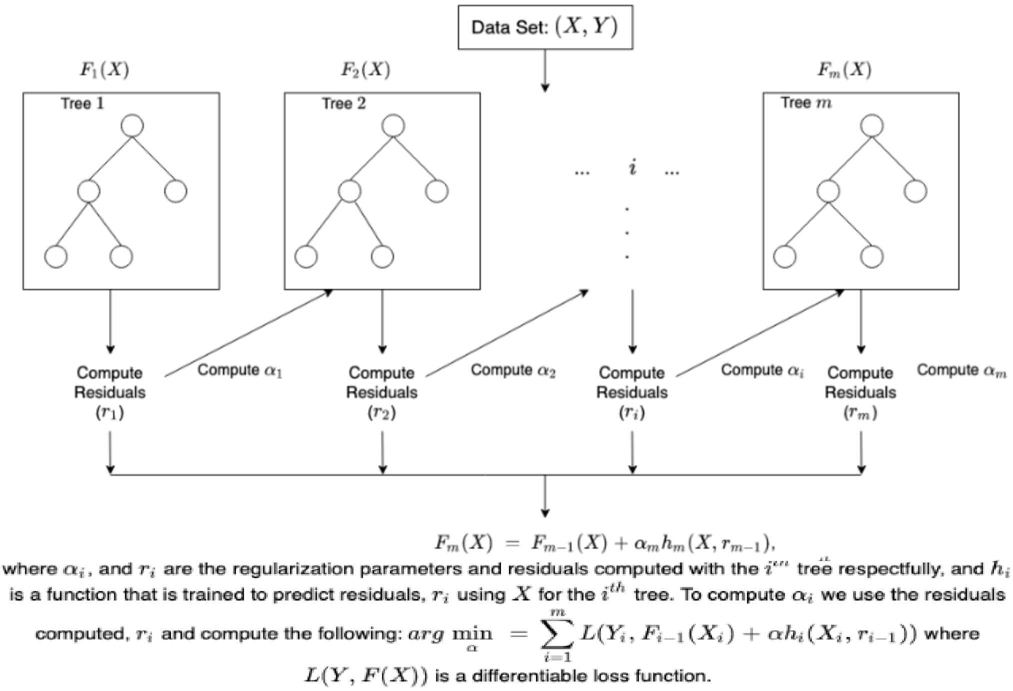


Figure 20. XG Boost illustration

To look at the overfitting issue, the classification error on both training data and test data is observed and early stopping is used. Training loss and testing loss of different methods are shown in Figure 21.

6. Results and Discussions

6.1. Quality Performance Prediction

This integrated framework wherein all the input parameters are captured in real-time with the sensors placed at the appropriate positions based on modal analysis and using the appropriate models can predict up to 25 output parameters with 96.2% accuracy. The results of all the models are shown in Table 4 below of which the top 4 models that are tuned with the best hyperparameters decided using Grid search CV are SVC, Random Forest, XG Boost Classifier, and MLP Classifier (CV).

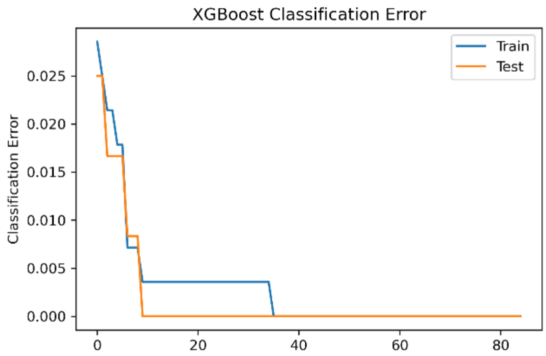


Figure 21. Classification error

The parameters used in modeling the above are given in Table 5 below:

Table 4. Recall-0, F1 score -0, and accuracy of all models

Model	cutoff= 79%						cutoff= 86.25%					
	Dimension Quality prediction						Quality of Surface finish prediction					
	Method1			Method2			Method3			Method4		
	Recall0	F1-score-0	Accuracy	Recall0	F1-score-0	Accuracy	Recall0	F1-score-0	Accuracy	Recall0	F1-score-0	Accuracy
Logistic Regression	0.44	0.5	0.82	0.48	0.52	0.82	0.53	0.69	0.93	0.47	0.59	0.91
KNN	0.8	0.83	0.93	0.72	0.71	0.88	0.53	0.67	0.93	0.65	0.76	0.94
SVC	0.76	0.79	0.92	0.84	0.81	0.92	0.59	0.71	0.93	0.71	0.8	0.95
Gaussian NB	0.6	0.59	0.82	0.68	0.6	0.81	0.76	0.6	0.86	0.88	0.6	0.83
Decision Tree	0.76	0.81	0.93	0.76	0.81	0.93	0.53	0.69	0.93	0.94	0.91	0.97
Random Forest	0.72	0.78	0.92	0.92	0.92	0.97	0.53	0.62	0.91	0.94	0.94	0.98
XGBoost Classifier	0.84	0.86	0.94	0.96	0.91	0.96	0.59	0.71	0.93	1	1	1
MLP Classifier	0.8	0.83	0.93	0.96	0.87	0.94	0.59	0.71	0.93	0.88	0.91	0.97
MLP Classifier(CV)	0.8	0.83	0.93	0.92	0.9	0.96	0.59	0.71	0.93	0.82	0.87	0.97
ANN	0.8	0.85	0.94	1	0.91	0.96	0.59	0.71	0.93	0.88	0.91	0.97
ANN(CV)	0.76	0.83	0.93	0.84	0.82	0.93	0.59	0.71	0.93	0.88	0.91	0.97

Table 5. Parameters used in modeling

Parameters used in modelling					
Model	Data	Method1	Method2	Method3	Method4
Logistic Regression	Actual	Default Parameters	Default Parameters	Default Parameters	Default Parameters
KNN	normalize	n_neighbors=3	n_neighbors=4	n_neighbors=6	n_neighbors=4
SVC	normalize	C=100,gamma=1,Kernel=rbf	C=10,gamma=1,Kernel=poly	C=100,gamma=1,Kernel=rbf	C=100,gamma=0.1,Kernel=rbf
Gaussian NB	Actual	Default Parameters	Default Parameters	Default Parameters	Default Parameters
Decision Tree	Actual	max_depth =3	max_depth =5	max_depth =1	max_depth =5
Random Forest	Actual	max_depth =3, estimators=9	max_depth =5, estimators=9	max_depth =1,estimators=15	max_depth =5,estimators=6
XGBoost Classifier	Actual	early stopping	early stopping	early stopping	early stopping
MLP Classifier	Actual	Default Parameters	Default Parameters	Default Parameters	Default Parameters
MLP Classifier(CV)	Actual	hidden_layer_size =40	hidden_layer_size =50	hidden_layer_size =25	hidden_layer_size =100
ANN	Actual	neurons=100, ann_default	neurons=100, ann_default	neurons=100, ann_default	neurons=100, ann_default
ANN(CV)	Actual	neurons=45, ann_default	neurons=50, ann_default	neurons=25, ann_default	neurons=60, ann_default
ann_default = Dense, 1 hidden layer, Relu activation, drop out = 0.3, adam optimizer					

All the models improved when the number of input features increased. Method 2 and method 4 have higher performance metrics than method 1 and method 3 respectively. It is observed that tree models have performed better than other models. XGB being a robust algorithm was able to predict the surface finish quality very perfectly using all features. RF and MLP gave very close results in method 2. Overall XGB is the best method for predicting both dimension quality and surface quality. RF and MLP are based on random generators and hence the results are not the same for each run, but similar results are achieved.

6.2. Decision-Making Insights:

The insights that we get into the live conditions of the manufacturing parameters and their effect on the dimensional and surface quality assist the operator in taking decisions like changing the tools or machine parameters resulting in error-free manufacture of components. It also helps in extracting maximum usage of the tool before it is worn out and brings in a paradigm shift in the process of changing tools after a specific number of operations. It would reduce the time spend on quality inspections as we are able to predict the quality accurately by up to 96.2%. This can be further increased if we can bring in image processing to determine tool wear and train the system over a larger sample size. In the long run, these could be automated and would help in reducing the cost of operation and make defect free parts

7. Conclusion

Contrary to individual post facto studies on quality this integrated real-time CPQS framework analyses real-time sensor networks together with artificial intelligence-driven big data analytics for predicting the quality of cyber-physical production networks can be a game changer. A detailed characterization of the manufacturing parameters responsible for the quality of a component manufactured on a 5-axis milling CNC machine showed the importance of vibration and the coolant temperature. Predicting the quality (dimensional and surface finish) of the part manufactured by using various machine learning techniques showed that the XG Boost

algorithm with 96.2% accuracy was the best. The major observations can be summarized as follows:

1. Placing the sensors at the right position using modal analysis to capture vibration is important.
2. The correlation heat map showed that dimensional quality is highly correlated to T1, T2 tool life & vibration, and weakly correlated to coolant temperature while the surface finish quality is highly correlated to T1 & vibration, and weakly correlated to coolant temperature and T2.
3. This is a case of multi-label classification supervised learning hence for optimal performance and high control of the process, the target variables are modeled independently and in two ways in each case
4. The performance metrics giving the best results were accuracy, F1 score, Precision, and Recall
5. It is observed that tree models have performed better than other models. XGB being a robust algorithm was able to predict the surface finish quality very perfectly using all features. RF and MLP gave very close results.

In continuation to the present work, the author believes that the framework can be further enhanced by introducing image processing techniques to determine the tool life rather than using Taylor's equation and increasing the sample size to improve the accuracy further and automate the production to produce defect-free parts under the auspicious Industry 4.0 paradigm.

Acknowledgments

Authors acknowledge the use of various open-source software [50-52] and thank the developer communities, Alliance university faculties, the CEO, and staff of FerAux Engineering, Bangalore India. Raw data that was generated at FerAux Engineering is available at DOI: https://doi.org/10.6084/m9.figshare.19367234

References

[1] M. Chacko & Atul. "Cyber-Physical Quality Systems in Manufacturing." Turkish Journal of Computer and Mathematics Education (TURCOMAT). Vol. 12 No. 2 (2021).

[2] T. Wang, X. Wang, R. Ma, X. Li, X. Hu, F.T.S Chan & J. Ruan. Random Forest-Bayesian Optimization for Product Quality Prediction with Large-Scale Dimensions in Process Industrial Cyber-Physical Systems. IEEE Internet of Things Journal, 7(9), 8641–8653.

[3] Y. Chen. Integrated and Intelligent Manufacturing: Perspectives and Enablers. *Engineering*, 3(5), 588–595.

[4] T.B. Tseng, Y. J. Kwon. "Characterization of machining quality attributes based on spindle probe, coordinate measuring machine, and surface roughness data", Journal of Computational Design and Engineering, *Vol. 1, No. 2* (2014) pp 128-135

[5] T.T. Ademujimi, M.P. Brundage & V.V. Prabhu, "A Review of Current Machine Learning Techniques Used in Manufacturing Diagnosis." Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing, 2017, Volume 513

[6] J. Lee, B. Bagheri and H. Kao." A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems" Science Direct Manufacturing *Letters* 3 (2015) 18–23.

[7] K. Alexopoulos, N. Nikolakis, & G. Chryssolouris. "Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing." International Journal of Computer Integrated Manufacturing, 33(5), 429–439.

[8] K.S. Aggour, et al (2019). "Artificial intelligence/machine learning in manufacturing and inspection." A GE perspective. MRS Bulletin, 44(7), 545–558.

[9] L. Monostori. "AI and machine learning techniques for managing complexity, changes, and uncertainties in manufacturing." Engineering Applications of Artificial Intelligence, 16(4), 277–291.

[10] U.M.R. Paturi & S. Cheruku. "Application and performance of machine learning techniques in the manufacturing sector from the past two decades: A review." Materials Today: Proceedings, 38, 2392–2401.

[11] M. Sharp, R. Ak & T. Hedberg. "A survey of the advancing use and development of machine learning in smart manufacturing." Journal of Manufacturing Systems, 48, 170–179.

[12] T. Wuest, D. Weimer, C. Irgens & K.D. Thoben. "Machine learning in manufacturing: advantages, challenges, and

applications." Production & Manufacturing Research, 4(1), 23–45.

[13] A. Dogan & D. Birant. "Machine learning and data mining in manufacturing." Expert Systems with Applications, 166, 114060.

[14] M. Hossin, & M.N Sulaiman. "A Review of Evaluation Metrics for Data Classification Evaluations." International Journal of Data Mining & Knowledge Management Process, 5(2), 01–11.

[15] T. El-Midany, M. El-Baz & M. Abd-ElWahed." A Proposed Performance Prediction Approach for Manufacturing Process Using Artificial Neural Networks." (Dept.M). MEJ. Mansoura Engineering Journal, 36(4), 39–49.

[16] A. Essien & C. Giannetti. "A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders." IEEE Transactions on Industrial Informatics, 16(9), 6069–6078.

.[17]H. Yetiş and M. Karaköse, "Adaptive vision-based condition monitoring and fault detection method for multi robots at production lines in industrial systems". International Journal of Applied Mathematics, Electronics and Computers 4 (Special Issue-1) pp.271–276

[18] T. Murata. "Petrinets: properties, analysis, and applications", Proc. IEEE77 (4) (1989) pp.541–580

[19] Z. Ahmed and A. Moosa. "An Intelligent Quality Function Deployment (IQFD) for Manufacturing Process Environment." Jordan Journal of Mechanical and Industrial Engineering (3) (2009) pp 23-30.

[20] S. Mohsen and A. Mohammed. "A Review of the Recent Development in Machining Parameter Optimization." Jordan Journal of Mechanical and Industrial Engineering (16)(2) (2022) pp 205-223.

[21] T. Mite, B. Prangoski and K. Paweł. "Mathematical Modelling and Correlation Between the Primary Waviness and Roughness Profiles During Hard Turning." Jordan Journal of Mechanical and Industrial Engineering. 15. 243-249.

[22] A.W. Hashmi, H.S. Mali, and A.Meena. "A comprehensive review on surface quality improvement methods for additively manufactured parts", Rapid Prototyping Journal, Vol. ahead-of-print No. ahead-of-print.

[23] A. Ahmed, A. Majeed, Z. Atta, and G. Jia. "Dimensional Quality and Distortion Analysis of Thin-Walled Alloy Parts of AlSi10Mg Manufactured by Selective Laser Melting." Journal of Manufacturing and Material. Processing. Vol 3, Issue 2, 2019, 51

[24] T. Yu and G. Wang, "Research of online process quality control system," IEEE International Conference on Automation and Logistics, 2008, pp. 351-356,

[25] B. Mrugalska and E. Tytyk," Quality Control Methods for Product Reliability and Safety", Procedia Manufacturing, Volume 3,2015, Pages 2730-2737, ISSN 2351-9789.

[26] A. Bahram, V. Durga, N. Clara, J. Jesus and R. Fischer. "Optimizing the quality of parts manufactured by the automated plasma cutting process using response surface methodology." 20th Annual International Solid Freeform Fabrication Symposium, SFF 2009.

[27] M. Andronie, G. Lăzăroiu, M. Iatagan, C. Uță, R. Ștefănescu and M. Cocoșatu. "Artificial Intelligence-Based Decision-Making Algorithms, Internet of Things Sensing Networks, and Deep Learning Assisted Smart Process Management in Cyber-Physical Production Systems," Electronics 10,(20), 2497.

[28] E. Nica, C.I. Stan, P.A.G Luțan and R.S. Oașa (Geambazi). "Internet of Things-based Real Time Production Logistics, Sustainable Industrial Value Creation, and Artificial Intelligence-driven Big Data Analytics in Cyber-Physical Smart Manufacturing Systems," Economics, Management, and Financial Markets 16(1): 52–62.

[29] M.T. Hayajneh, M.S. Tahat, and J. Bluhm. "A study of the Effects of Machining Parameters on the Surface Roughness in the End-Milling Process". Jordan Journal of Mechanical and Industrial Engineering (1) (2007) pp1-5

[30] K. Kadirgama, M.M. Noor, N.M.Zuki.N.M, M.M. Rahman, M.R.M. Rejaba, R.Dauda, K. A. Abou-El-Hossein. "Optimization of Surface Roughness in End Milling on Mould Aluminium Alloys (AA6061-T6) Using Response Surface Method and Radian Basis Function Network." Jordan Journal of Mechanical and Industrial Engineering. Vol 2, No 4, 2008 pp 209- 214.

[31] K. Soundararajan, S.M. Ramkumar, and I. Edinborough. "A Web-Based Approach To Automated Inspection And Quality Control Of Manufactured Parts Paper" American Society for Engineering Education Annual Conference & Exposition (ASEE), Montreal, CA, 2002

[32] M.N. Islam."Effect of additional factors on dimensional accuracy and surface finish of turned parts", Journal Machining Science and Technology, Volume 17, (2013) - Issue 1 pp145-162

[33] A.A. Thakre, V.L. Aniruddha and K. Mala," Measurements of Tool Wear Parameters Using Machine Vision System". Hindawi Modelling and Simulation in Engineering, Volume 1 (2019), Article id 1876489.

[34] R.G. Linsa, P.R.M. Araujoa, M, Corazzimb "In-process machine vision monitoring of tool wear for Cyber-Physical Production Systems". Robotics and Computer Integrated Manufacturing 61 (2020) pp.85-101

[35] T.B. Tseng, Y. J. Kwon. "Characterization of machining quality attributes based on spindle probe, coordinate measuring machine, and surface roughness data", Journal of Computational Design and Engineering, Vol. 1, No. 2 (2014) pp 128-135

[36] M.N. Islam. "Effect of additional factors on dimensional accuracy and surface finish of turned parts", Journal Machining Science and Technology, Volume 17, (2013) - Issue 1 pp145-162

[37] B. I. Kazem and N. F. H. Zangana. "A Neural Network-Based Real-Time Controller for Turning Process." Jordan Journal of Mechanical and Industrial Engineering, Vol 1, No 1, 2007 pp 43-55

[38] J.L. Berg et al. "Self-Optimization and self-ConfiguRATion in wirelEss networkS (SOCRATES)." COST 2100TD (08)422, Wroclaw, Poland, Feb 2008.

[39] Universita Politecnica Delle Marche "Integration of process and quality control using multi-agent technology (GRACE)" report Cordis. European Union project id:246203

[40] A. Schirrmann." Adaptive Production Planning and Scheduling: The ARUM approach based on MAS and SOA technologies" Springer International Publishing, 2019 - Technology & Engineering

[41] E.A. Lee." The Past, Present, and Future of Cyber-Physical Systems: A Focus on Models". Sensors 2015, 15(3), 4837-4869.

[42] J. Lee, B. Bagheri and H. Kao." A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems" Science Direct Manufacturing Letters 3 (2015) 18–23.

[43] U.M.R. Paturi & S. Cheruku. "Application and performance of machine learning techniques in the manufacturing sector from the past two decades: A review." Materials Today: Proceedings, 38, 2392–2401.

[44] L. Monostori, B.Kadar, Z.J. Viharosy, L. Mezgar, & P.Stefan. "Combined Use of Simulation and AI/Machine Learning Techniques in Designing and Manufacturing Processes and Systems." Journal for Manufacturing Science and Production, 3(2–4),.2000, pp111–118.

[45]T. Zeugmann, et. al. “Precision and Recall.” Encyclopedia of Machine Learning, 781.

[46] Y. Fathy, M. Jaber, and A. Brintrup. “Learning with Imbalanced Data in Smart Manufacturing: A Comparative Analysis.” IEEE Access, 9, 2734–2757.

[47] Z. Ge, Z. Song, S.X. Ding, and B. Huang. “Data Mining and Analytics in the Process Industry: The Role of Machine Learning.” IEEE Access, 5, 20590–20616.

[48] A. Kusiak. “Smart manufacturing must embrace big data.” Nature, 544(7648), 23–25.

[49] Pedregosa, et al., Scikit-learn: Machine Learning in Python, *JMLR* 12, pp. 2825-2830, 2011.

[50] G. van Rossum. “Python tutorial, Technical Report CS-R952.” Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.

[51] J. D. Hunter. “Matplotlib: A 2D Graphics Environment.” *Computing in Science &* Engineering, vol. 9, no. 3, pp. 90-95, 2007.

[52] M. Waskom et al. mwaskom/seaborn: *Zenodo* v0.8.1 (September 2017),