Application of Bayesian Family Classifiers for Cutting Tool Inserts Health Monitoring on CNC Milling

Abhishek D. Patange[1*], Jegadeeshwaran R[2]

[1,2]*School of Mechanical Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India*
abhipatange93@gmail.com
jegadeeshwaran.r@vit.ac.in

[1]*Department of Mechanical Engineering, Govt. College of Engineering, Pune, 411005, India*

ABSTRACT

The customized usage of tool inserts plays an imperative role in the economics of machining operations. Eventually, any in-process defects in the cutting tool lead to deterioration of complete machining activity. Such defects are untraceable by the conventional practices of condition monitoring. The characterization of such in-process tool defects needs to be addressed smartly. This would also assist the requirement of 'self-monitoring' in Industry 4.0. In this context, induction of supervised Machine Learning (ML) classifiers to design empirical classification models for tool condition monitoring is presented herein. The variation in faulty and fault-free tool condition is collected in terms of vibrations during the face milling process on CNC (Computer Numerically Controlled) machine tool. The statistical approach is incorporated to extract attributes and the dimensionality of the attributes is reduced using the J48 decision tree algorithm. The various conditions of tool inserts are then classified using two supervised algorithms viz. Bayes Net and Naïve Bayes from the Bayesian family.

1. INTRODUCTION

A cutting tool is presumed to commit high persistence, strength, and most important repeatability in the machining operation (Grzesik 2017). The systematic usage of cutting tool executes an imperative role in the economics of machining operation. Thereby, any defect in the cutting tool leads to deterioration of complete machining activity. The consequences such as poorer surface finish, a discrepancy in the dimension of workpiece, substantial power consumption of drive, discontinuance of machining process, etc. are dominating and unendurable (Roth, Djurdjanovic, Yang, Mears, and Kurfess 2010). The choice of conservative input factors (Depth of cutting, Speed of machining and table feed) which satisfies the utmost cutting conditions is the primary and irreplaceable step to initiate the machining. However, this may not address the dynamics in the machining and corresponding effects such as excessive vibrations, temperature, noise, and power consumption leading to in-process failure of tools (Engin and Altintas 2001). Also, as far as the current cutting tool market scenarios are concerned, highly advanced metallurgical solutions and several analytical and numerical studies merely endure the dynamics of the machining process (Drori 2015). Thus, the experimental approach aimed particularly for monitoring tool condition and random dynamics serves to be the most realistic (Daneshmand and Pak 1986, Dan and Mathew 1990, Maj, Modica and Bianchi 2006).

The process of experimental condition monitoring commences with the acquisition of signals which describes the change in behavior of machine component. The vibration developed during the machining is considered as the commendable directive of association between the health of a tool, deviations in accuracy and surface finish of job. The machining vibrations are deliberated as most instructive, reactive, intent, self-explanatory, and trustworthy as compared to other signals such as temperature, noise, spindle load, and power consumption (Mohanraj, Shankar, Rajasekar, Sakthivel and Pramanik 2020). Also, the instrumentation involved in vibration acquisition is considerably less; an accelerometer as a sensor and a data acquisition system (Dimla and Lister 2000, Dimla 2000, Zhou and Xue 2018). Thus, the study dealing with the assessment of vibrations evolved during milling in order to report resultant cutter health is presented herein. Hakan, Ali, Sadettin, and Ersan (2016) used statistical parameters to examine the relation of machining parameters, respective tool conditions, and associated vibration experimentally. The estimate of the tool wear and corresponding work-piece roughness was presented with the intention of predictions. Cuka and Kim (2017) investigated the correlation between machining speed and tool life using frequency response functions. It is proved that by monitoring machining speed, the tool life is extended and interruptions due to failure are avoided. The study presented

by Zhou and Xue (2018) & Zhang, Lu, Wang, Li (2018) also includes the multi-sensor approach for milling tool condition monitoring but acts unsuitable for on-board fault identification as the complexity of instrumentation and signal processing is a concern. Ong, Lee, Jit, and Lau (2019) estimated wear rate and respective roughness of work in CNC end milling. The image processing was used to highlight the descriptor of wear zone and hence further used for prediction. Torabi, Li, Lim, and Peen (2016) presented a fault diagnosis of the tool by examining time-frequency changes in end milling operation. The learning model was developed to provide a prediction on the status of the process by adapting to variations. The study presented by Dilma (2002) analyzed the tool wear and showed that catastrophic failure occurs if the magnitude of nose wear becomes greater than or equal to 200 microns. Conversely, notch wear grows linearly when the machining starts and retained unaffected for rest of the tool life. Shiba (2003) presented the use of piezoelectric vibrators to identify the smallest variation in vibration response when the cutting tool is damaged. The time-domain plot of vibration response is an appropriate tool to inspect errors rising while machining to determine the deterioration level of a cutting tool (Siddhpura & Bhave 2008). In the experiment carried by Balla, Sarcar, and Satish Ben (2010) a vibrometer was used to capture the signal and observed an increase in the amplitude of displacement with the progression of wear for finding the association between flank wear, surface roughness with cutting tool vibrations. Narayanan and Namboothiri (2010), performed nonlinear time series analysis for the vibration involved during machining and found that the phase space trend along with the correlation dimension compute got increased as flank wear increased. Ming, Jiawei, and Dinghua (2016) examined the relation between cutting forces and the deviation in thickness of the chip removed. It has projected that dramatic change in the direction of cutting forces causes excessive vibrations at the exit of the cutter.

The traditional condition monitoring implicates the study and analysis of tool condition with the aid of manual supervision which has left out of favor in the boom of artificial intelligence (AI). A study presented by Surendar and Elangovan (2017) used the M5P tree, regression tree (RT), and multiple linear regression (MLR) algorithms for predicting the roughness of the workpiece. The vibration response in terms of statistical features was extracted with the help of Principle Component Analysis (PCA). In the investigation of classification accuracy, when histogram and statistical attributes were served to classifiers, it was found that statistical attributes yield more accuracy i.e. 86.34% than histogram attributes as 73.61% (Elangovan, Ramachandran, and Sugumaran 2010). Sambayi (2012) showed four features i.e. kurtosis, RMS (Root mean square), crest factor and peak are most appropriate to classify vibration signals of different drill conditions. But later, it was noted that as the peak is a non-averaging statistical

attribute hence not appropriate in concern of stochastic response of the tool. At last, the superior results were obtained considering kurtosis and crest factor. Jegadeeshwaran and Sugumaran (2015) used the best first search method in attribute evaluator for examining the influence of several attributes on classifier efficiency. It is based on Greedy hill climb augmentation to discover the attribute subsets space using the principle of back-tracing. In this study, the supervised subset evaluator (CfsSubsetEval) has been used for the feature selection study. It is observed that when nine features were selected, the classification algorithm provides the maximum classification accuracy. Hence using supervised subset evaluator (CfsSubsetEval), the top nine features were selected for classification. CfsSubsetEval not only considers prognostic aptitude of the individual attribute but also redundancy degree between them to evaluate the significance of attributes subset. In the final step, the features exhibiting less inter-correlation were selected. Madhusudana, Kumar, and Narendranath (2016) presented an experimental investigation for failure prediction of milling cutter based on J48 tree and K star classifier with the histogram features approach. A classifier 'K star' from the family of Lazy algorithms was utilized for categorization of the tool conditions. The study presented by Bohara, Jegadeeshwaran, and Sakthivel (2017) exhibits the classification of tool condition in turning performed on a traditional lathe using J48 and random tree classifiers. Since single carbide coated insert was employed, a study involving multi-insert tool and possible failure conditions can't be correlated. Also, tool failure conditions such as nose and notch wear were not considered. Recently, another research presented by Patange and Jegadeeshwaran (2019) incorporates machine learning scheme for monitoring health of milling tool considering 5 different tree-based algorithms separately and comparative study is presented amongst them. The highest accuracy achieved was 90% and the time consumed for constructing the model was 8.6 seconds which is considerably higher.

The designing and training of Bayesian classifiers for multi-point tool inserts differ from other machining operations. The machining by milling is different from shaping and turning, in the sense that material in milling is removed by a rotary cutting tool. It consists of multiple cutting edges revolving around a fixed axis hence known as 'multi-point tool'. The milling process provides intermittent cutting, with each tooth producing chips of variable thickness. This is an inherent difference between turning, shaping, and milling because, in turning and shaping, there is a continuous chip production. Moreover, the milling tool bears vertical as well as sideways thrust produced during machining. Thus as far as training of vibration signature evolved due to the engagement of multi-point tool inserts is considered, the application of the Bayesian family has not been reported yet. The current research proposes a vibration-based tool condition monitoring framework capable of fault classification considering the Supervised Machine Learning

approach. It is also learned that the studies on health monitoring of cutting tools have been reported for a conventional lathe (Elangovan et al., 2010, Painuli, Elangovan and Sugumaran 2014, Elangovan 2011, Elangovan, Sugumaran and Ramachandran 2011, Shewale, Mulik, Deshmukh 2018, Nalavade, Patange, Prabhune 2018) and conventional milling machines (Cuka et al., 2017, Zhou et al., 2018, Zhang et al., 2018, Ong et al., 2019, Torabi et al., 2016). The literature showcasing the health monitoring in the intention for finding faults in multi-point tool insert employed in computer numerical controlled milling machines in small/medium scale industry need to be explored to meet the requirements of Industry 4.0 (Fink, et al. 2020).

The research gap is stated herein:

- The traditional approach of the selection of various levels and combinations of input parameters is process characterization and arbitrarily comments about tool health are derived from the same study. However beyond that there exist unknown moments resulting in in-process faults and need to be explored. The characterization of such faults and corresponding vibrations are not mapped with the tool condition.
- The study reported by M. Elangovan et al. (2010) demonstrated the use of Bayesian classifiers for monitoring the health of cutting tools in turning operations where the tool consisted of the only single insert. The study of multi-point tool inserts needs to be undertaken to examine vibration signatures. Another study presented by Karandikar et. al. (2015) demonstrated the use of cutting force to train Bayesian classifiers. As far as training of vibration signature evolved due to the engagement of multi-point tool inserts is considered, the application of the Bayesian family has not been reported yet.
- The previous study published is only confined to the application of classifiers for the study of data distributions to propose the classification. However, the use of the learner's output to classify the blind datasets has not been previously addressed.

The unique contributions and innovations of the current study are:

- Demonstrates a real-time multi-point tool insert fault classification scheme
- Characterization of data sets generated for unknown moments of vibration evolved from in-process tool faults (i.e. wear at nose & flank and notch wear)
- Design and training the Bayesian family algorithms for classification of tool inserts fault and its validation considering test data set
- Classification of predefined classes for blind data sets based on learner's output

2. Detailed Experimentation

2.1. Experimentation arrangement

The experimentation arrangement (Figure 1) comprises of a Computer Numerically Controlled Milling machine, an accelerometer (Model PCB 352c33), and a data acquisition system (National Instruments Inc. Model 9234).
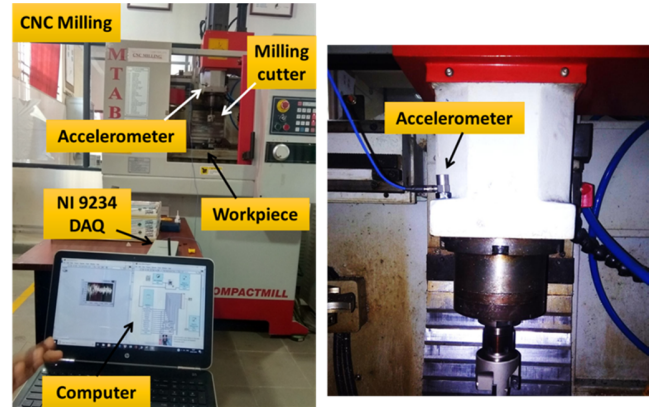

Figure 1. The schematic of the experimentation arrangement

The Mild Steel workpiece is machined using a face milling cutter of four inserts (coat of carbide) where the single level of the input parameter is selected as:

1. Spindle speed = 1200 revolutions per minute,
2. Feed = 50 mm/min,
3. Cut Depth = 0.35 mm.

The parameters were estimated using the formula as per the technical tool selection manual provided by Komet. The KOMET GROUP is a global technology leader for metal cutting tools and effective, high-precision solutions in the areas of drilling, bore machining, reaming, threading, and milling (Kom Guide 2016). The technical tool selection manual provided by Komet is thus commonly referred to while selecting machining parameters.

The typical parameters in milling operations are depth of cutting, machining speed and table feed. Its choice based on the type of operation to be performed, the material of the workpiece, the diameter of the cutter, and standard feed per tooth.

The cutting speed can be calculated by,

$$V_c = (\pi \times D_m \times N)/1000 \qquad (1)$$

where D_m = Diameter of cutter = 40 millimeters, N = Spindle speed = 1200 revolutions/minute

By substituting these values in equation 1, we get, machining speed = V_c = 150.72 meters/minute

Now, table feed is calculated by,

$$V_f = F_z \times N \times z \qquad (2)$$

where F_z= Standard feed per tooth for cast-iron workpiece = 0.01 millimeters/tooth, Z = Number of inserts

By substituting these values in equation 2, approximately we get, feed = V_f = 50 millimeters/minute

Lastly, cut depth was selected as 0.35 millimeters.

The procedure stated above is adopted for selecting the best choice of depth of cutting, machining speed and table feed according to workpiece material and standard catalog. For cast iron workpiece and face milling operation the possible ranges of these parameters are (Kom Guide 2016):

Machining speed = 110 – 180 meters/minute

Feed = 30 – 80 millimeters/minute

Depth of cut = 0.2 – 0.5 millimeters

Here a total of 5 machining operations were performed with the same level of machining parameters and with different tool conditions specified in Tables 1 and 2. The three most commonly occurring wear conditions namely flank, nose, and notch are considered in the study.

Insert/Operation No.	1	2	3	4	5
Insert i	N	N	N	N	N
Insert ii	N	N	N	N	FW
Insert iii	N	N	N	N	NSW
Insert iv	N	FW	NSW	NTW	NTW

Table 1. Insert conditions for machining

Operation No.	Category label
1	4N
2	3N_1D_1
3	3N_1D_2
4	3N_1D_3
5	1N_3D_1_2_3

Table 2. Insert conditions & operation category label

where N: Normal, FW: Flank wear, NSW: Nose wear, NTW: notch wear

2.2. Tool faults considered

The tool wear is nothing but gradual progressive removal of material with molecular or atomic subtractive mechanisms. Most commonly three kinds of deterioration occur namely wear at flank & nose and notch wear (Figure 2). The elementary theoretical mechanism for these tool inserts deterioration is discussed here:

Flank wear is a prevalent style of tool wear due to its constant progression and is easy to identify. As the name implies, it is wearing off the flank face due to harsh abrasion through the hard workpiece. It is usually observed in the removal of steels, cast iron with the abrasives such as iron carbide and non-metallic additives. Flank wear can largely be attributed to a mechanical process initiated when primary particles are released from the tool face by a temperature-dependent process. As these primary particles are released and travel down the flank, they scour its surface and in the process release secondary particles by abrasion. These secondary particles also act similarly to further abrade the rest of the flank. Progressively, this process results in a wear land being formed on the flank. Flank wear grows faster for higher depth of cut, feed, and machining speeds. Higher cutting forces resulting from a 0.5-0.6 mm dimension of flank wear causes severe tool damage. Abrasion by fragments of the built-up edge is one of the reasons behind flank wear caused due to excessive friction in the tool-workpiece interface (Bermingham, Palanisamy, and Dargusch 2012).

Notch wear is observed because of the oxidation of primary and sub-tool-tips, otherwise harsh abrasion. It is favorable for work materials that possess high work hardening and creates higher cutting edges heat, i.e., nickel-base super-alloys, titanium, and non-magnetic stainless steel (gamma-iron). It causes the insert to wear quickly in the cutting zone. The local stress concentration may also cause notching. As a result of the compressive stress along the cutting edge – and lack of the same behind the cutting edge – the insert is particularly stressed at the depth of the cut line. This region develops a thin work-hardened layer that can originate a groove. A groove or notch develops in a region that undergoes work-hardening. Oxide layers on a workpiece also contribute to notch wear because these are hard and abrasive (Bermingham, Palanisamy and Dargusch 2012).

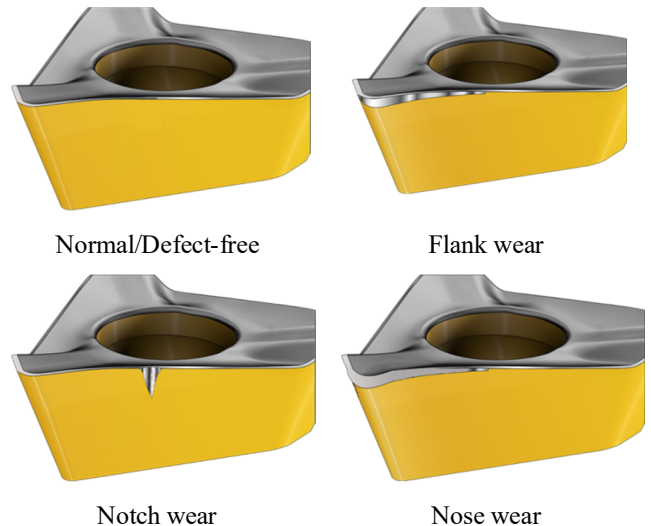Normal/Defect-free   Flank wear

Notch wear   Nose wear

Figure 2. Faults considered

Nose wear is nothing but dulling of cutting edge observed due to the weak deformation resistance of an insert with excessive heat generation during machining. By providing this nose radius the carbide inserts are strengthened to withstand the high cutting force created during the cutting operation. If not provided, edge frittering takes place at the

nose causing scratch marks on the component. Nose wear is rounding off a sharp radius due to mechanical and thermal effects. The excessive chip population causes rubbing of the tool over the workpiece increasing the temperature and leads to wear of the nose radius (Bermingham, Palanisamy and Dargusch 2012).

2.3. Data acquisition

The accelerometer used in the study is single-axial. The vibration evolved during machining transferred on either side, i.e., on a tool spindle or a workpiece side. However intending to record tool fault signatures, any change in spindle motion needs to be detected; hence locating the accelerometer near to a tool spindle is necessary. It is learned that the vibration with vertical components (Z-direction) is best responsive to tool wear than the other two axes (Dimla and Lister 2000). Hence the piezoelectric accelerometer was directly attached to the spindle holder (frame) vertically with the help of sticky material. This location of the accelerometer will ensure that most of the vibration signals collected by the accelerometer come from the tool and will be least affected by other components of the CNC Milling machine. The data acquisition system (DAQ) is employed to acquire a time-domain response. According to the Nyquist proposition, the signal sampling was carried out at frequency of 20 kHz. In this study, 250 samples describing all tool conditions are selected. This means each tool condition is described by 50 samples. Each sample constitutes of 12000 data points of the acquired vibration signal.

The vibration signal depicted in figure 3(a)–(e) respectively plotted considering a single sample of each configuration. Each sample consists of 12000 data points and is collected for 0.6 seconds. For machining activity 1-5, the corresponding time dependent graphs exhibit characteristic variation in the acceleration. The time-domain graph depicts the periodic and cyclic nature of vibration signals corresponding to the instantaneous influence of fault on the process. This typical behavior of the vibration signature is necessary to be understood by a learner while studying the influence of faults on the process involving multi-point cutting tools and is evident when one observes the time-domain graph. The time-based features such as skewness, standard deviation, standard error, etc. can sufficiently classify the tool condition as evident from the decision tree algorithm. Hence the initial study is carried out using time-domain signals which require no mathematical computation additionally and are easy for real-time monitoring. The fault diagnosis studies using the time-domain analysis reported that the features extracted from the time-dependent signal alone are capable of mapping insert failures to the vibrations evolved during the tool-workpiece interface (Rubeo and Schmitz 2016, Liu, Kothuru and Zhang 2020, Aralikatti et al. 2020, Alamelu and Jegadeeshwaran 2020).
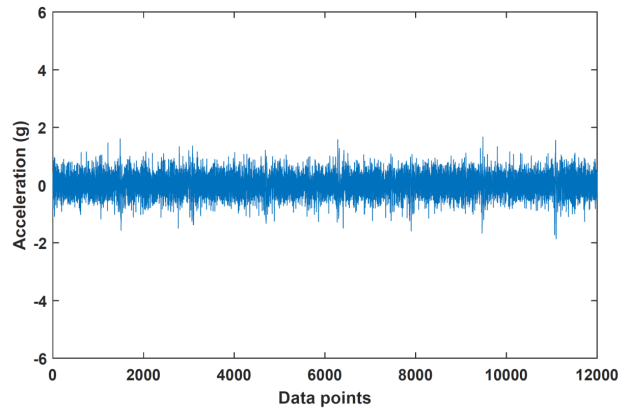

Figure 3 (a). Time-domain plots for operation 1 with class '4N'

Different tool configurations involved in machining reflect the distinct change in vibration signals. In order to quantify this analog-type signal and show accurate difference between tool conditions, descriptive statistics & ML algorithms are deployed and presented in further sections.
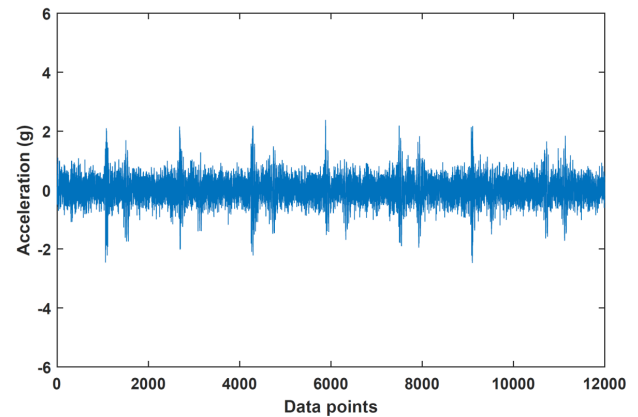

Figure 3 (b). Time-domain plots for operation 2 with class '3N_1D_1'
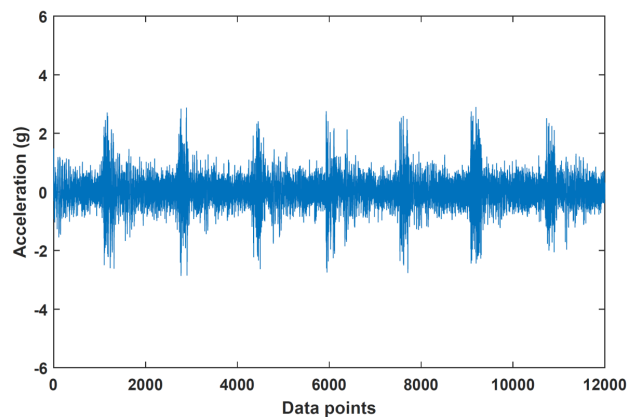

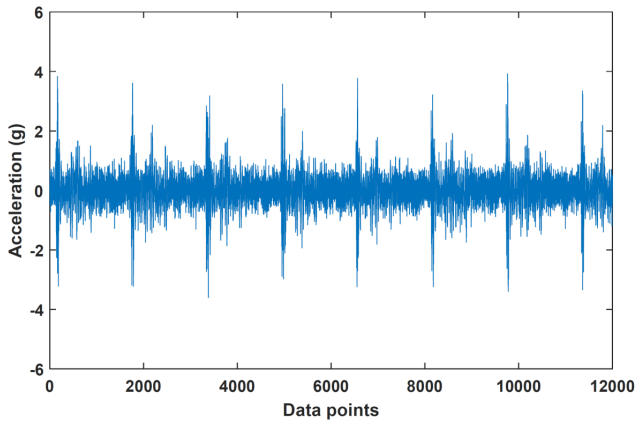Figure 3 (c). Time-domain plots for operation 3 with class '3N_1D_2'

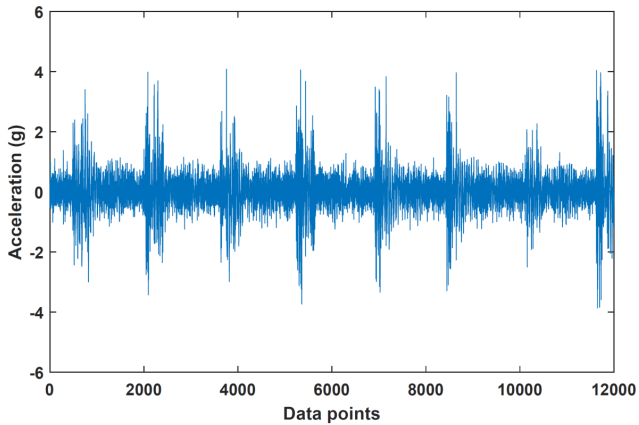Figure 3 (d). Time-domain plots for operation 4 with class '3N_1D_3'

Figure 3 (e). Time-domain plots for operation 5 with class '1N_3D_1_2_3'

3. MACHINE LEARNING FRAMEWORK

3.1. Attributes extraction

The change in vibration signal reflects the significant statistical information about corresponding tool insert conditions. Statistical attributes were estimated for data collected from the time-dependent response of the vibration signal for all 5 machining operations. Here total of 16 attributes such as (1) Standard Error, (2) Minimum value, (3) Range, (4) Maximum value, (5) Shape factor (6) Count, (7) Skewness, (8) Summation, (9) Standard Deviation, (10) Variance, (11) Mode, (12) Median (13) Kurtosis, (14) Impulse factor (15) K-factor (16) Mean were estimated from the coding. The mathematical expressions for statistical attributes are given in Annexure A. The pseudo-code in Python is presented in Annexure B.

3.2. Attributes selection

The dimensionality of the attributes is reduced with the help of a decision tree (J48). This stage is considered to be an important phase as it suggests the best distinct features amongst overall descriptive statistics set from training data and thus decides the accuracy. Feature selection using a decision tree (J48) is easy, computes negligible time, and exhibits appropriate results. The principle of 'information gain' and 'entropy reduction' is used in the decision tree classifier. It discretizes the instances first and then exhibits binary structured framework comprising homogeneous instances. It is constructed in form of primary root, various nodes, leaves, and branches. The top-down ordinance assists the classification of classes. This means the induction starts at primary root to ends at leaves via different divisions. The primary root defines the best significant attribute, every interior node defines a test on a feature, every branch describes a result of the trial and each terminating node (leaf) describes a class. All existing attributes in a tree are deliberated to be key attributes among others (Kingsford and Salzberg 2008, Jankowski and Jackowski 2014).

3.3. Attributes classification

Several other algorithms like Support Vector Machine, Perceptron, Lazy, Trees, Regressions, and Artificial Neural Networks can be utilized; provided that superior accuracy and less computation time are attained. The 'Bayesian family classifiers' performs with superior efficiency and in addition to that, consumes lesser period to construct the model. Thus 'BayesNet and NaiveBayes' classifiers are chosen for this investigation. It is the family of 'supervised-probabilistic based' algorithms are popularly called as Bayesian network classifiers works on the principle of 'Bayes' law' with robust liberal assumptions amongst the attributes. Naive Bayes develops the model which allocates instances category characterized as vectors of attributes mined from a certain predetermined finite set. The assumption made is that the magnitude of a specific attribute is not depending on any other attribute in the same category. Bayes law is a 'condition-based rule' apparently trusting the indication that expected to be classified, and mathematics that illustrates confidence about the indication to be presented with (Friedman and Geiger 1997, Wiggins, Saad and Litt 2008).

Theoretically, naive Bayes is a model where the condition to be categorized is characterized by,

$$vector\ v = (v_1, v_2, v_3 \ldots \ldots, v_z) \tag{3}$$

where 'z's are several attributes, which allocates to output condition possibilities.

$$p(Ct_m\ l\ v_1, v_2, v_3 \ldots \ldots, v_z) \tag{4}$$

for each of m probable output category Ct_m

Eventually, for the larger number of samples and attributes, constructing this model on probability formulation is unattainable hence seems to be the drawback of this approach. Hence this necessitates reconstructing the model to achieve an attainable model.

Using Bayes' law, the conditional probability can be decomposed as

$$p(Ct_m \, l \, v) = [\, p(Ct_m) \, p(v \, l \, Ct_m) \,]/p(v) \qquad (5)$$

Simply speaking the above formulation is regarded as

Prediction = (Past*Probability)/Proof-mark

On careful observation, it is learned that the denominator remains constant as it is independent of Ct_m and attribute magnitude v_i is also already provided. The value of the numerator is considered to be equivalent as of probability model

$$p(Ct_m, v_1, v_2, v_3 \ldots \ldots, v_z) \qquad (6)$$

Reformulating this by utilizing chain law for recurrent presentations of the characterization of uncertain probability,

$$p(Ct_m, v_1, v_2, \ldots \ldots, v_z)$$

$$= p(\, v_1, v_2, \ldots \ldots, v_z, Ct_m) \qquad (7)$$

$$= p(\, v_1 \, l \, v_2, \ldots \ldots, v_z, Ct_m) \, p(\, v_2, \ldots \ldots, v_z, Ct_m) \qquad (8)$$

$$= p(\, v_1 \, l \, v_2, ., v_z, Ct_m) \, p(\, v_2 \, l \, v_3, ., v_z, Ct_m) \ldots \ldots$$

$$\ldots p(\, v_3, ., v_z, Ct_m) \qquad (9)$$

$$= \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots. \qquad (10)$$

$$= p(\, v_1 \, l \, v_2, \ldots \ldots, v_z, Ct_m) \, p(\, v_2 \, l \, v_3, \ldots \ldots, v_z, Ct_m) \ldots$$

$$\ldots p(\, v_{z-1} \, l \, v_z, Ct_m) p(\, v_z \, l \, Ct_m) p(Ct_m) \qquad (11)$$

Here the assumption of naive condition appears as it assumes that all attributes 'V' are communally non-dependent, conditional on the class Ct_m. With this,

$$p(\, v_i \, l \, v_{i+1}, \ldots \ldots, v_z, Ct_m) = p(v_i \, l \, Ct_m) \qquad (12)$$

Therefore, the combined approach gives a model which can be stated as,

$$p(\, Ct_m \, l \, v_1, \ldots \ldots, v_z) \propto p(\, Ct_m, v_1, \ldots \ldots, v_z)$$

$$= p(\, Ct_m) \, p(v_1 \, l \, Ct_m) \, p(v_2 \, l \, Ct_m) \, p(v_3 \, l \, Ct_m) \ldots \ldots$$

$$= p(\, Ct_m) \prod_{i=1}^{Z} p(\, v_i \, l \, Ct_m) \qquad (13)$$

Therefore, with this assumption, the categorization of Ct_m is,

$$= p(\, Ct_m \, l \, v_1, \ldots, v_z) = \frac{1}{E} \, p(\, Ct_m) \prod_{i=1}^{Z} p(\, v_i \, l \, Ct_m) \qquad (14)$$

where, evidence $E = p(v) = \sum p(Ct_m) \, p(v \, l \, Ct_m)$

is a scaling factor dependent only on $v_1 \ldots v_z$ otherwise, remains constant if attribute values are known (Elangovan et al., 2010, Karandikar et. al. 2015).

4. RESULTS AND DISCUSSION

Stepwise result of machine learning scheme i.e. statistical attribute extraction, selection, and classification is discussed here. The time-domain response of acceleration as depicted in Fig. 3(a)–(e) differs for each machining operation as a result of various defective tool configurations. Figure 3 (a) is recorded corresponding to a healthy condition and evident that the amplitude of vibrations is less and uniform in nature. Figure 3(a)–(e) depicts the periodic and cyclic nature of vibration signals corresponding to the instantaneous influence of fault on the process. It is very challenging to quantify variation between the classes due to the analogs nature of the plots; hence descriptive statistics demonstrating the signal are extracted and explained in with the help of illustrative graphs.

It can be witnessed in figure 4 (a) to (e) that the variation of standard error, range, K factor, and skewness exclusively differs for healthy tool conditions over faulty conditions. The variation of the mean for healthy conditions matches other faulty conditions in a few instances. This is because of the configurations of tools considered for each operation and inherent faults. Along with this the variation of standard error, range, K factor, and skewness notably differs for tool condition '1N_3D_1_2_3' as this configuration involves 3 faulty inserts. The visualization of this graph provides a general overview of classification. However, to quantify the classification, the Bayesian family algorithms are designed and trained. The results are discussed herein.
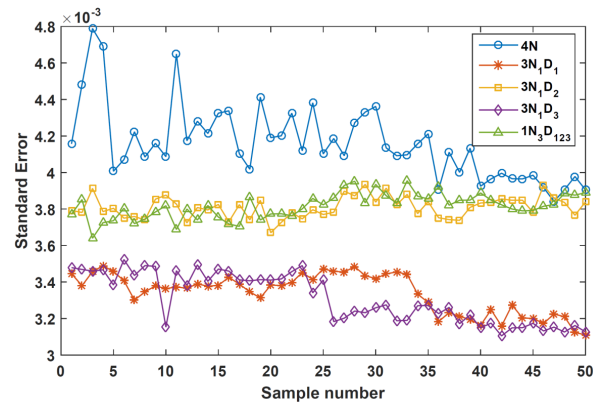

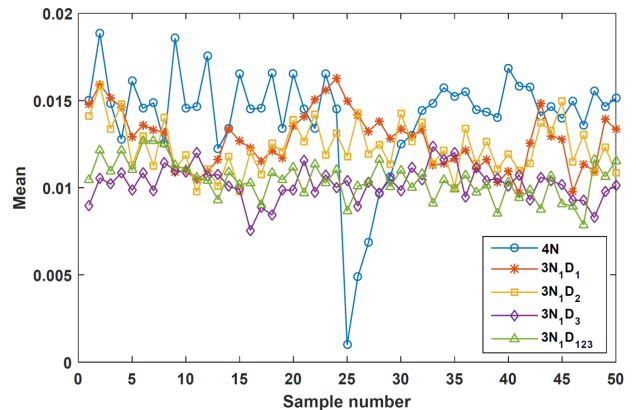Figure 4 (a). Class wise variation of standard error

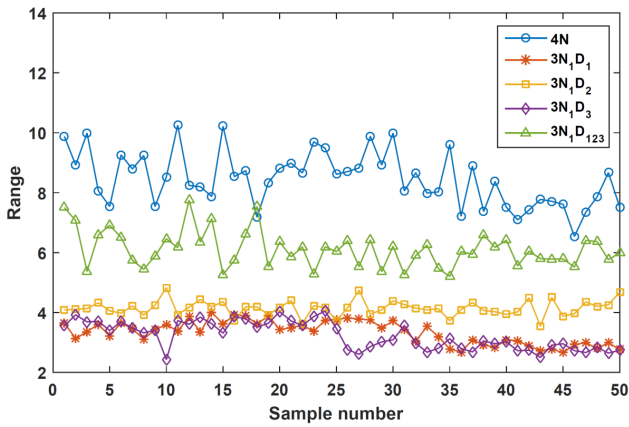Figure 4 (b). Class wise variation of Mean

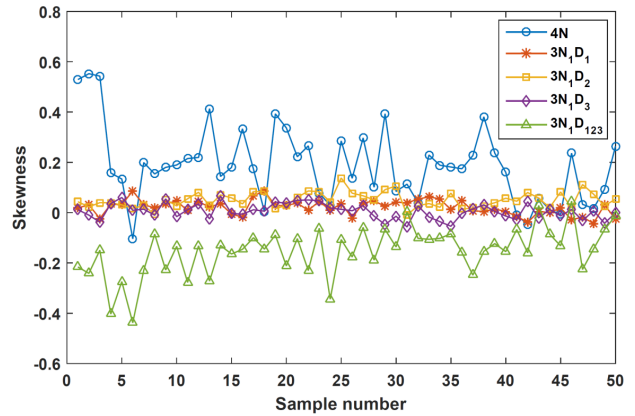Figure 4 (c). Class wise variation of Range

Figure 4 (e). Class wise variation of Skewness
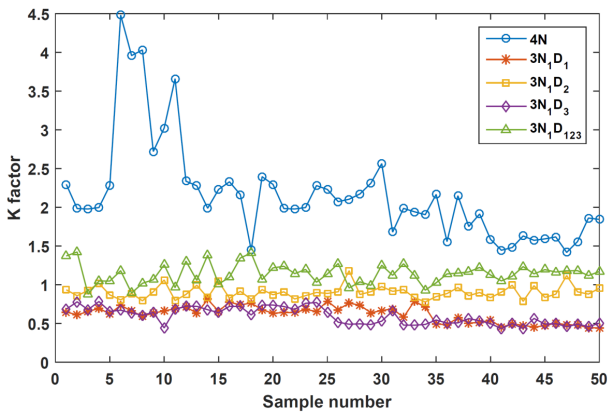
Figure 4 (d). Class wise variation of K factor

The second step is attribute selection. Here, all 16 attributes of 250 samples (50 samples per class) are applied to the decision tree algorithm, and structure is exhibited in figure 5.

The attributes displayed in the framework of decision tree created using the J48 classifier are served to be 'most distinctive' in a set of all extracted features i.e. only 5 attributes i.e. 'Standard error, Mean, Range, K-factor and Skewness' and hence used for classification. The tree consists of 9 leaves and 17 sub-divisions in size. The interpretation of the decision tree is given here. In a case where 'standard error' is larger than 0.003523 and 'range' is smaller than or equivalent to 4.829225, then all 50 samples of tool condition '3N_1D_2' i.e. nose wear are correctly classified.

Figure 5. Dimensionality reduction using decision tree

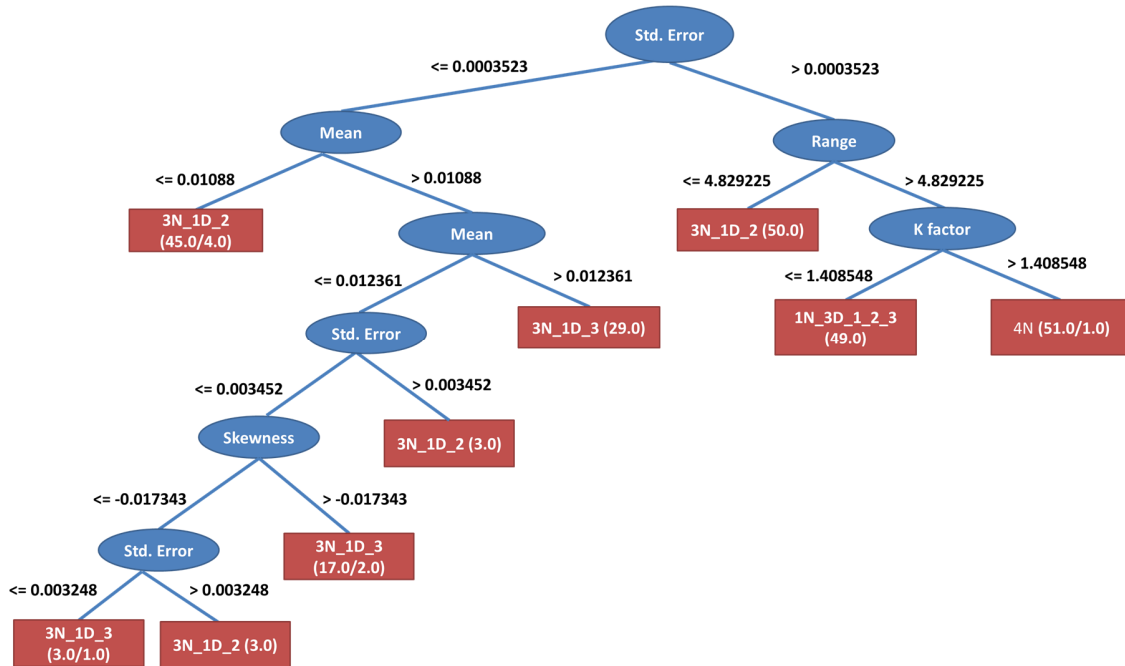But when 'range' is greater than 4.829225 and 'k-factor' is lesser than/equivalent to 1.408548, at that moment 49 samples of tool condition '1N_3D_1_2_3' i.e. 3 defectives are correctly categorized. Likewise, other classes are categorized. The further description is self-explanatory. Now the last step is attribute classification. The result classification of attributes with respect to various conditions/classes /categorize of tool condition is explained here.

The detailed performance of classifiers is elaborated using various terms like Recall, Precision, True Positive (TP) rate, Receiver Operating Characteristics (ROC) region, and False Positive (FP) rate. Simply stated, true positive represents % of truly recognized samples, and false-positive represents the % of wrongly identified samples. A perfect classification model must yield the TP rate near to unity and the FP should be null. The ratio of TP rate to all the samples in the category is represented by 'Precision'. ROC tells about how the model has performed overall. The detailed class-wise performance for BayesNet is shown in Table 3.

True Positive	False Positive	Precision	ROC	Tool Condition
0.98	0	1	1	4N
0.84	0.035	0.857	0.974	3N_1D_1
0.98	0.005	0.98	1	3N_1D_2
0.86	0.04	0.843	0.975	3N_1D_3
1	0.005	0.98	1	1N_3D_1_2_3
0.932	**0.017**	**0.932**	**0.99**	**Weighted Average**

Table 3. Tool condition-wise performance for BayesNet

The accuracy of the BayesNet classifier is found to be 93.2% under 10 fold cross-validations. The misperception matrix for BayesNet is shown in figure 6. The performance of a model is presented by the misperception matrix (Popularly called a 'confusion matrix') in a most significant way. The elements placed in the central diagonal of this matrix exhibit correctly classified elements whereas elements placed in non-diagonal locations exhibit wrongly classified samples. The misperception matrix demonstrates that for tool condition '4N' where four inserts are normal, 49 samples placed properly under 'All defect-free' and 1 sample wrongly placed under condition '3N_1D_2' (i.e. 1 with nose wear). Next to that, for condition '3N_1D_1' (1 is worn flank and 3 are normal), 42 samples placed properly under actual condition '3N_1D_1', whereas 8 samples wrongly placed under condition '3N_1D_3' (notch wear) condition.

This confusion matrix can be explicitly used to understand further categorization w.r.t. the rules stated above. The validation of this experiment can be examined from the Misperception matrix, as it was created based on

the 10 folds cross-validation. The comparative study reveals that the BayesNet algorithm yields more accuracy (93.2%) than the NaiveBayes algorithm (88.8%). Also, the time consumed for training the model for the BayesNet algorithm is 0.7 seconds and the NaiveBayes algorithm is 0.2 seconds.

Predicted class						Actual class	Tool Condition
A	**B**	**C**	**D**	**E**			
49	0	1	0	0	A		4N
0	42	0	8	0	B		3N_1D_1
0	0	49	0	1	C		3N_1D_2
0	7	0	43	0	D		3N_1D_3
0	0	0	0	50	E		1N_3D_1_2_3

Figure 6. Misperception matrix for BayesNet algorithm

The naive Bayes algorithm yields a classification accuracy of 88.8%. The misperception matrix for NaiveBayes is represented in figure 7.

Predicted class						Actual class	Tool Condition
A	**B**	**C**	**D**	**E**			
42	0	0	0	8	A		4N
0	39	0	11	0	B		3N_1D_1
0	0	49	0	1	C		3N_1D_2
0	9	0	41	0	D		3N_1D_3
0	0	0	0	50	E		1N_3D_1_2_3

Figure 7. Misperception matrix for NaiveBayes algorithm

The previous section presented a classification using K-fold cross-validation mode. In this mode, the predictor has an unfair advantage, as they were chosen in 1st based on all of the samples. Leaving samples out after the variables have been selected does not correctly mimic the application of the classifier to a completely independent test set, since these predictors 'have already seen' the left out samples (Hastie T., Tibshirani R. and Friedman J. 2009). To test the robustness of the classifier to a completely independent test set with no labels is considered. Thus the current section presents the application of the learner's output to classify the blind datasets. Figure 8 shows the framework for class prediction used in this study. The pseudo-code for the predictor model is presented in Annexure C.
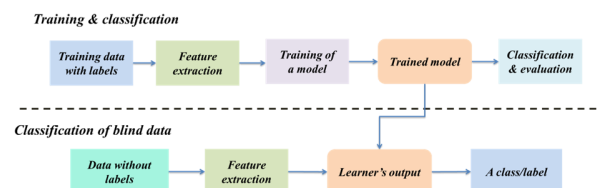
Figure 8. Framework for classification of blind data

To demonstrate the convergence of the trained model for classifying blind datasets adequate trials (500 runs) have been undertaken. In this test, 100 samples per each class were fed to the classifier model as blind data. The misperception matrix depicted in figure 9 presents the result and is self-explanatory. The trained model was deployed towards predicting the class for blind datasets and the results showcased a high convergence rate (i.e. 442/500 samples correctly predicted giving accuracy of 88.4%). Hence the same trained model can facilitate the prediction of a class within explicit domains.

Predicted class							Tool Condition
A	**B**	**C**	**D**	**E**			
95	0	2	0	3	A		**4N**
0	88	0	2	10	B		**3N_1D_1**
0	0	85	6	9	C	Actual class	**3N_1D_2**
0	0	4	88	8	D		**3N_1D_3**
0	4	4	6	86	E		**1N_3D_1_2_3**

Figure 9. Misperception matrix for classification of blind datasets

5. LIMITATIONS & FUTURE SCOPE

The framework presented herein limits itself to the tool condition classification study for vibration signatures collected considering fixed machining factors. Any variations in the depth of cut, feed & speed, will affect the tool condition and thus the trained Bayesian classifier may categorize the tool class equivalent to whichever of the pre-defined categorize. Due to alteration machining parameters, the model may come across different phases of faults within a class and thus results in misclassification. To avoid this, the training consisting of different levels of machining parameters and phases of faults – a generic model can be investigated; that would increase the accuracy of classification. Testing of the algorithm against varying CNC parameters in addition to the varying damage conditions requires further experimentation & data expansion and be pursued in near future. However based on the framework presented herein, one can easily adopt advanced experimentation and training of data.

6. CONCLUSION

A novel machine learning framework for cutting tool inserts monitoring on CNC milling based on vibration analysis was successfully investigated. The statistical approach is incorporated to extract attributes. The dimensionality of the attributes is reduced using the J48 decision tree algorithm. The various conditions of tool inserts are then classified using two supervised algorithms viz. BayesNet and NaiveBayes from the Bayesian family and accuracy of 93.2% and 88.88% were achieved

respectively. This framework illustrates the supervisory methodology to be more competent and can be effectively implemented to assure higher productivity and ultimately to enhance tool-machine life. Also, it serves as the most appropriate scheme on-board fault diagnosis. The training model as a descriptor and a data fusion scheme can be incorporated for future correspondence to make machine tools self-intelligent pursuing generalizability and repeatability.

ANNEXURE A:

The mathematical expressions for descriptive statistical attributes:

Sr. No.	Attribute	Mathematical Expression
1	Kurtosis	It is an estimate of the 'tailedness' of the probability distribution of a real-valued random variable. $$\left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{S_d} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$$
2	Standard Error	Standard Error is a measure of the deviation of the sample means from the population. The standard error of a sample statistic is an estimate of the standard deviation of the sampling distribution of that sample statistic. It helps you to find out confidence intervals for that statistic at different significance levels. $$\sqrt{ \frac{1}{n-2} \left(\sum (y - \bar{y})^2 - \frac{\sum[(x - \bar{x})(y - \bar{y})]^2}{\sum(x - \bar{x})^2} \right) }$$
3	Maximum value	It is the highest data point value.
4	Skewness	It defines the inclination of the spread of data on either side. $$\frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{S_d} \right)^3$$
5	Minimum value	It is the lowest data point value.
6	Range	It is an estimate of subtraction between maximum and minimum values of data points.
7	Count	It is an estimate of the number of data points in each sample.
8	Summation	It is an estimate of the sum of all feature values for each sample.
9	Variance	It is the expectation of the squared deviation of a random variable from its mean. $$\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}$$
10	Standard Deviation	It is a measure of the amount of variation or dispersion of a set of values. $$\sqrt{ \frac{n \sum x^2 - (\sum x)^2}{n(n-1)} }$$
11	Mode	It is an estimate of the number which occurs most frequently in a set of data points.

12	Median	It is an estimate of the middle value segregating the higher and lower splits of a data set.
13	Mean	It is an estimate of the average (arithmetic) of a set of data points. $$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots\ldots + x_n}{n}$$
14	Impulse factor	The impulse factor estimates the impact created by considering the ratio of maximum value and average value. $$\frac{Maximum\ value\ of\ x}{\bar{x}}$$
15	K-factor	The K-factor is a product of root mean square value and maximum value. $$\frac{\sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \ldots\ldots + x_n^2}{n}}}{Maximum\ value\ of\ x}$$
16	Shape factor	The shape factor is the ratio of root mean square value and average value. $$\frac{\sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \ldots\ldots + x_n^2}{n}}}{\bar{x}}$$

ANNEXURE B:

The pseudo-code is for computing statistical attributes in Python:

```
import glob as glob
import pandas as pd
import numpy as np
import scipy.stats as ss
from scipy import stats
path = r'C:\Users\HP\Desktop\Python\4N'
filenames = glob.glob(path + "\*.xlsx")
def stat(DataFrame, name):
    x = df.iloc[:,0]
    a = np.mean(x)
    b = np.median(x)
    c = stats.mode(x).mode[0]
    d = ss.kurtosis(x)
    e = ss.skew(x)
    f = np.std(x)
    g = np.var(x)
    h = ss.sem(x)
    i = np.amax(x)
    j = np.amin(x)
    k = np.sum(x)
    l = np.ptp(x)
    m = np.sf(x)
    o = np.imf(x)
    p = np.kf(x)
    n = np.stde(x,50)
    R1 = pd.DataFrame({'File Name': [name[31:-5]], 'Mean':[a],
'Median':[b], 'Mode':[c], 'Kurtosis':[d], 'Skewness':[e], 'Std
Dev':[f], 'Var':[g], 'STd Err':[h], 'Max':[i], 'Min':[j],'Sum':[k],
'Range':[l], 'Shape factor':[m], 'Impulse factor':[o], 'K factor':[p],
'stde':[n]})
    R1.set_index('File Name')
    return R1
all_data = pd.DataFrame()
for f in filenames:
    df = pd.read_excel(f)
    mdf = stat(df, f)
    all_data = all_data.append(mdf, ignore_index = True)
all_data.to_excel
('C:/Users/HP/Desktop/Python/4NOutput.xlsx',sheet_name   =
"sheet 1")
```

ANNEXURE C:

The predictor model used for classifying blind data:

```
Std error <= 0.003523
|  Mean  <= 0.01088: 3N_1D_3 (45.0/4.0)
|  Mean  > 0.01088
|  |   Mean  <= 0.012361
|  |  |   Std error <= 0.003452
|  |  |  |   Skewness  <= -0.017343
|  |  |  |  |   Std error <= 0.003241: 3N_1D_1 (3.0/1.0)
|  |  |  |  |   Std error > 0.003241: 3N_1D_3 (3.0)
|  |  |  |   Skewness  > -0.017343: 3N_1D_1 (17.0/2.0)
|  |  |   Std error > 0.003452: 3N_1D_3 (3.0)
|  |   Mean  > 0.012361: 3N_1D_1 (29.0)
Std error > 0.003523
|   range <= 4.829225: 3N_1D_2 (50.0)
|   range > 4.829225
|  |   k factor <= 1.408548: 1N_3D_1_2_3 (49.0)
|  |   k factor > 1.408548: 4N (51.0/1.0)
```

REFERENCES

Wit Grzesik. (2017) Machining Economics and Optimization, *Advanced Machining Processes of Metallic Materials: Theory, Modeling, and Applications*, 2, 265–283.

Roth J. T., Djurdjanovic D., Yang, X., Mears, L., Kurfess, T. (2010) Quality and Inspection of Machining Operations: Tool Condition Monitoring, *Journal of Manufacturing Science and Engineering ASME, 132*(4), 1–15.

Engin S. and Altintas Y. (2001) Mechanics and dynamics of general milling cutter, Part II: Inserted cutters, *International Journal of Machine Tools and Manufacture*, *41*(15), 2213–2231.

Gidi Drori. (2015) Proper Cutting Tool Choice is Vital to Productivity. *Production Machining: Gardner Business Media Incorporation*.1–4.

Daneshmand L. K. and Pak H. A. (1986) Adaptive Control System using economic performance index in turning, *Journal of dynamic systems measurement and control*, *108*, 215–222.

Dan Land, Mathew J. (1990) Tool wear and failure monitoring techniques or turning – a review, International *Journal of Machine Tool Manufacture*, *30*, 579–598.

Maj R, Modica F., Bianchi G. (2006) Machine Tools – Mechatronic Analysis, *Proc. Institute of Mechanical Engineers, 220 (B),* 345–353.

Mohanraj T., Shankar S., Rajasekar R., Sakthivel N. R., Pramanik A. (2020) Tool condition monitoring techniques in milling process – A review, *Journal of Materials Research Technology, 9*(1), *1032–1042.*

Dimla D. and Lister P. (2000) On-line metal cutting tool condition monitoring: force and vibration analyses, *International Journal of Machine Tools Manufactures*, *40,* 739–768.

Dimla D. E. (2002) The correlation of vibration signal features to cutting tool wear in a metal turning operation, *The International Journal of Advanced Manufacturing Technology, 19*, 705–713.

Zhou Y., Xue W. (2018) Review of tool condition monitoring methods in milling processes, The International *Journal of Advanced Manufacturing Technology*, *96,* 2509–2523.

Hakan Arslan, Ali Osman Er, Sadettin Orhan, Ersan Aslan. (2016) Tool Condition Monitoring in Turning Using Statistical Parameters of Vibration Signal, International, *Journal of Acoustics and Vibration*, *21*(4), 371–378.

Besmir Cuka and Dong-Won Kim. (2017) Fuzzy logic based tool condition monitoring for end-milling, *Robotics and Computer-Integrated Manufacturing*, *47*(C), 22–36.

Yuqing Zhou, Wei Xue. (2018) A Multi-sensor Fusion Method for Tool Condition Monitoring in Milling, *Sensors*, *18* (11), 38–66.

Zhang X. Y., Lu X., Wang S., Wang W., Li W. D. (2018) A multi-sensor based online tool condition monitoring system for milling process, *Proceedia CIRP, 72,* 1136–1141.

Pauline Ong, Woon Kiow Lee, Raymond Jit, Hoo Lau. (2019) Tool condition monitoring in CNC end milling using wavelet neural network based on machine vision, *The International Journal of Advanced Manufacturing Technology, 104,* 1369–1379.

Torabi A. J., Er M. J., Li X., Lim B. S., Peen G. O. (2016) Application of Clustering Methods for Online Tool Condition Monitoring and Fault Diagnosis in High-Speed Milling Processes, *IEEE Systems, 10* (2), 721–732.

Shiba K. (2003) Development of a miniature abrasion-detecting device for a small precision lathe, *Sensors Actuators: A Physics*, *109,* 137–142.

Siddhpura M., Siddhpura A., Bhave S. (2008). Vibration as a parameter for monitoring the health of precision machine tools, *International conference on frontiers in design and manufacturing engineering.*

Balla S. P., Sarcar M. M. M., Satish Ben B. (2010) Development of a system for monitoring tool condition using acousto-optic emission signal in face turning–an

experimental approach, *International Journal of Advanced Manufacturing Technology, 51,* 57–67.

Narayanan R. V., Namboothiri V. (2010) Flank wear detection of cutting tool inserts in turning operation: application of nonlinear time series analysis, *Soft Computation Fusion Found Methodology Application, 14,* 913–919.

Ming L., Jiawei M., Dinghua Z. (2016) Time-domain modeling of a cutter exiting a workpiece in the slot milling process, *Chinese Journal of Aeronautics, 29,* 1852–1858.

Surendar S., Elangovan M. (2017) Comparison of Surface Roughness Prediction with Regression and Tree Based Regressions during Boring Operation, *Indonesian Journal of Electrical Engineering and Computer Science, 7*(3), 887–892.

Elangovan M., Ramachandran K. I., Sugumaran V. (2010) Studies on Bayes classifier for condition monitoring of single point carbide tipped tool based on statistical and histogram features, *Expert Systems with Applications-Elsevier, 37,* 2059–2065.

Sambayi P. M. K. (2012) Drill wear monitoring using instantaneous angular speed: A comparison with conventional technology used in drill monitoring systems, *Masters Theses-University of Pretoria.*

Jegadeeshwaran R., Sugumaran V. (2015) Brake fault diagnosis using Clonal Selection Classification Algorithm (CSCA): A statistical learning approach, *Engineering Science and Technology, an International Journal, 18,* 14–23.

Madhusudana C. K., Kumar H., Narendranath S. (2016) Condition monitoring of face milling tool using K-star algorithm and histogram features of vibration signal, *Engineering Science and Technology, An International Journal-Elsevier, 19,* 1543–1551.

Navneet Bohara, Jegadeeshwaran R., Sakthivel G., (2017) Carbide Coated Insert Health Monitoring Using Machine Learning Approach through Vibration Analysis, International *Journal of Prognostics and Health Management, 24,* 1–14.

Patange A. D., Jegadeeshwaran R. (2019) Milling cutter condition monitoring using machine learning approach, *IOP Conference Series: Material Science and Engineering, 624:012030*

Painuli S., Elangovan M. and Sugumaran V. (2014) Tool condition monitoring using K-star algorithm, *Expert Systems with Applications, 41,* 2638–2643.

Elangovan M. (2011). Evaluation of expert system for condition monitoring of a single point cutting tool using principle component analysis and decision tree algorithm, *Expert Systems with Applications, 38,* 4450–4459.

Elangovan M., Sugumaran V., Ramachandran K. I. (2011) Effect of SVM kernel functions on classification of vibration signals of a single point cutting tool, *Expert Systems with Applications, 38* (12), 15202–15207.

Shewale M. S., Mulik S. S., Deshmukh S. P., Patange A. D. (2018), A novel health monitoring system, *Advances in Intelligent Systems and Computing Springer Singapore Proc. of 2nd Inter. Conf. on Data Engineering and Communication Technology, 828,* 461–468.

Nalavade S. P., Patange A. D., Prabhune C. L. and Mulik S. S., (2018) Development of 12 Channel Temperature Acquisition System for Heat Exchanger Using MAX6675 and Arduino Interface, *Lecture Notes in Mechanical Engineering, Springer, 1,* 119-125

Olga Fink, Qin Wang, Markus Svensen, Pierre Dersin, Wan-Jui Lee, Melanie Ducoffe (2020) Potential challenges and future directions for deep learning in prognostics and health management applications, *Engineering Applications of Artificial Intelligence, 92,* 103678

Karandikar J., Tom M., Sam T., Tony S. (2015) Tools wear monitoring using naive bayes classifiers. *International Journal of Advanced Manufacturing Technology, 77,* 1613–1626.

Kom Guide (2016) Technical Manual Drilling, Threading, Reaming, Milling, *Komet Group*, 1–325.

Bermingham M. J., Palanisamy S., Dargusch M. S. (2012) Understanding the Tool Wear Mechanism during Thermally Assisted Machining Ti–6Al–4V, *International Journal of Machine Tools and Manufacture*; *62,* 76–87.

Rubeo M. A., Schmitz T. L. (2016) Global stability predictions for flexible workpiece milling using time-domain simulation, *Journal of Manufacturing Systems, 40,* 8–14.

Liu R, Kothuru A, Zhang S. (2020) Calibration-based tool condition monitoring for repetitive machining operations, *Journal of Manufacturing Systems, 54,* 285–293.

Aralikatti S. S., Ravikumar K. N., et al. (2020) Comparative Study on Tool Fault Diagnosis Methods Using Vibration Signals and Cutting Force Signals by Machine Learning Technique, *Structural Durability & Health Monitoring, 14*(2), 128–145.

Alamelu M. T. M. and Jegadeeshwaran R. (2019) Vibration based brake health monitoring using wavelet features: A machine learning approach, *Journal of Vibration and Control,* 1–17.

Kingsford C., Salzberg S. L. (2008) What are decision trees? *Nature Biotechnology, 26*, 1011–1013

Jankowski D., Jackowski K., (2014) Evolutionary Algorithm for Decision Tree Induction, *13th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM) - Ho Chi Minh City, Vietnam,* 23–32.

Friedman N., Geiger D. (1997) Bayesian network classifier, *Machine Learning, 29,* 131–163.

Wiggins M., Saad A., Litt B. (2008) Evolving a Bayesian classifier for ECG-based age classification in medical applications, *Applied Soft Computing, 8* (1), 599–608.

Hastie T., Tibshirani R. and Friedman J. (2017) High-dimensional problems, *in The Elements of Statistical Learning, Springer-Verlag,* 649-694

BIOGRAPHIES

Abhishek D. Patange, is currently working as an Assistant Professor, in Mechanical engineering, Govt. College of Engineering, Pune (COEP). He holds bachelor's degree in Mechanical from COEP, a master's degree in Mechanical Design Engineering. He is pursuing Ph.D. from VIT Vellore in field of condition monitoring using machine learning approach. He works in domain of application of machine learning for machine fault diagnosis, Machine health monitoring using vibration analysis, Mechatronics and Controller integration with mechanical systems, Designing GUI based control algorithms using MATLAB & Python.

Dr. R. Jegadeeshwaran is currently working as an Associate Professor, in School of Mechanical Engineering, Vellore Institute of Technology, Chennai. He holds B.E. in Mechanical, M.Tech in Mechatronics and PhD in Fault Diagnosis. With fifteen + years of academic & research experience, he has more than 40 research papers at national and international conferences and published papers in national and international journals and filed 3 patents. He serves as a principal investigator for 'Experimental Investigation for Developing an on-board diagnostic tool for monitoring the health condition of a brake system of an automobile' under DST SERB and received funding of 25 Lakhs. Currently, he is guiding 3 Ph.D. scholars. He also serves as a member of various professional bodies.