1. INTRODUCTION

The condition of the tool plays a significant role in producing the required surface finish, which determines the quality of the product. The failure of the cutting tool happens over a period due to deterioration [1]. Further, the replacement of defective tools demands unplanned downtime and leads to production loss. Hence, monitoring the state of the tool is essential to schedule preventive maintenance [2]. Essentially, monitoring the tool condition prevents unexpected rejection of workpiece, minimizes downtime, achieves targeted workpiece dimensions, and, as a result, lowers production costs and manpower [3]. This monitoring can be of two methods, direct method, in which the state of the tool is measured directly through optical sensors, or wear width can be measured using tool maker’s microscope. This method gives more accurate results but is economically not viable [4]. The main drawback of the direct method is the noise that exists in the signal due to coolant and metal chips. On the other method, the tool conditions can be predicted indirectly by capturing the signals such as Current [5], Power [6], Force [7], Acoustic Emission (AE) [8], Sound [9], and Vibration [10]. These indirect methods are more feasible and the tool states are correlated well with the signals acquired [11].

In general, Tool Condition Monitoring (TCM) in indirect methods involves two fundamental processes, they are extraction of features from the sensor signal pool and the next step is to process the signal to diagnose the tool conditions. Extraction of useful features such as statistical feature [12], histogram feature [4], wavelet feature [6,10] is in current research.

The features extracted are processed using Machine Learning Algorithms such as Artificial Neural Network (ANN) [13], Naive Bayes classifier [9], Decision Tree [12], SVM [14], Random Forest [15], Fuzzy classifier [16], K-star algorithm [4].

This Machine Learning (ML) is gaining momentum in data analysis and pattern classification problems of various domains. Syed Shaul Hameed et al. (2021) compared the performance of the fuzzy classifier with ANN in planetary gearbox condition monitoring. The authors used histogram features and concluded that the results of both the classifiers showed encouraging results [17]. Fault diagnosis of Self-Aligning Troughing Roller (SATR) in belt conveyor systems was reported in [18]. The authors compared the performance of ANN and the decision tree algorithm using vibration signals. An image based surface texture classification of the machined surface using ANN and Random Forest algorithm was carried out by [19]. Susai Mary et al. (2019) predicted the surface roughness in the drilling process. ANN approach was used for tuning the drilling parameters to yield the required surface finish [20]. Altobi et al. (2019) presented work on centrifugal pump fault diagnosis. Authors used ML algorithms such as backpropagation multilayer perceptron with genetic algorithm and SVM. Continuous wavelet transform (CWT) features under different conditions of the pump were considered and concluded that the multilayer perceptron approach achieved an accuracy of 99.5% [21]. Zhong et al. (2019) investigated the fault diagnosis of a gas turbine by transfer learning method based on Convolution Neural Network (CNN) and SVM. Authors redesign the CNN to make it capable to diagnose the fault with limited data. They concluded that the proposed methods require less expert knowledge and hence, reduce the preprocessing work [22]. Joshua and Sugumaran (2018) reported the condition monitoring of wind turbine blades. The authors have used the Auto-regressive Moving Average (ARMA) feature set to classify the faults in the turbine blade [23]. Rahul Kumar Sharma et al. (2017) did work on condition monitoring of
roller bearing. The classification accuracies of the K-Star algorithm and K-Nearest Neighbor (KNN) algorithm were compared by acquiring sound signals [24].

From the literature, one can understand the remarkable role of ML in the pattern classification of machine tools and machine elements. It is evident from the literature that the vibration signal has a strong correlation with the tool state [25,26]. In the material removal process, face milling is a primary process that produces flat surfaces using a multi-point cutting tool. SVM is one of the popular ML algorithms that find limited application in milling TCM. The modest way to analyze data is through statistical features. Therefore, in the present work, the performance of SVM is studied in classifying the different conditions of the milling tool. Further, the work is compared with K-Star, another prominent ML with limited implication in the field of fault diagnosis.

2. METHODOLOGY

In the present work, the statistical parameters of the acquired vibration signal were extracted during the face milling of mild steel with different conditions of the tool. The salient features are selected from the decision tree and given as input to the SVM algorithm. The work was extended with the K-Star algorithm and the performance of algorithms compared. The flow of the work is depicted in figure 1.

![Diagram of methodology](image1)

**Figure 1. Flow chart for the fault diagnosis of milling tool**

3. EXPERIMENTAL SETUP AND PROCEDURE

The experiments are conducted in a vertical milling machine that has eight cutting inserts. The cutting inserts were mounted in the milling cutter. Using a tri-axial piezoelectric accelerometer (Model No: 8763B500BB Kistler make) the vibration signals during the milling process were obtained. The accelerometer was mounted on the spindle head as shown in figure 3. An analog-to-digital converter is used to convert the vibration signal into a digital form. Then in computer memory, the digital signals were stored. Figure 3 shows the actual representation of the experimental setup. Initially, the milling cutter was fitted with defect-free inserts in the vertical milling machine. The machine was allowed to run for about three minutes to stabilize the vibrations and rough machining was carried on the workpiece.

After the machine gets stabilized, the vibration signals corresponding to the new unused tool were acquired by the piezoelectric triaxial accelerometer. Then the experiments were conducted with new inserts and one insert with flank wear and the corresponding vibration signals were acquired. In the same manner, the insert with the other faults taken up in the study such as tool with wear on rake face and tool with broken tip was introduced and the respective vibration signals were recorded. From the experiments, a total of 1000 signals were collected out of which 250 signals belonged to each condition. A sample vibration signal recorded is presented in figure 4. The acceleration amplitude varies with the condition of the tool. The cutting parameters were kept constant for all the above experiments and the mild steel was used for ease of machining. The cutting parameters are given in table 1.

<table>
<thead>
<tr>
<th>Cutting Parameters</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>250 RPM</td>
</tr>
<tr>
<td>Feed rate</td>
<td>14 mm/minute</td>
</tr>
<tr>
<td>Depth of cut</td>
<td>0.5 mm</td>
</tr>
</tbody>
</table>

Table 1. Details of Machining Parameters

The photographic image of the tool shown in figure 2 gives a better understanding of the various faults of the tool.

![Photographic images](image2)

**Figure 2. Inserts with various faults and good condition**

a. Fault-free (G)  b. Flank wear (FW)  c. Wear on rake face (C)  d. Broken Tip (B)
4. FEATURE EXTRACTION

In their raw form, data are rarely valuable. A sampling rate of 2048 samples per second was used to capture the corresponding vibrations for the different conditions of the tool for 10 seconds. Therefore, each signal has 20480 data points. A total of 250 instances for each condition are recorded.

Managing and interpreting such a massive amount of data becomes tedious, requires large storage, and is time-consuming. Therefore, it is vital to extract and pick the simplest and most valuable features that contain meaningful information. This is possible with the use of modern signal processing techniques and is the most important phase in tool condition monitoring because they form the input to the machine learning algorithms.

Deriving insights manually from the raw signals requires domain knowledge and is a time-consuming process. Hence, to draw some meaningful conclusions the key statistical features are extracted as discussed in [12], [30].

Figure 3. Experimental Setup

Figure 4. Representative Time Domain signal for different conditions of the tool
5. FEATURE SELECTION

The process of selecting useful features from the extracted features is known as feature selection. The selected features have a high potential to classify the good and fault conditions of the tool. This needs to be done to increase the classification accuracy and to reduce the data size. The decision tree algorithm is popularly used in time series data to select the prominent features. The statistical features are given as input to the decision tree algorithm and decision rules are obtained as output by computing the entropy and information gain from the input statistical features.

In the decision tree, the information gain is calculated for all the features and the one having high information gain is placed at the top of the tree which is called the root node. The features with subsequent information gain are placed based on their order of importance. These nodes are called branch nodes. The branching continues until the decision is reached, the target class. The entropy and the information gain are calculated using equations (1) and (2).

\[
\text{Entropy}(S) = \sum_{i=1}^{c} -P_i \log P_i \tag{1}
\]

where \(P_i\) is the probability of a particular class and \(c\) is the total number of classes.

\[
\text{Gain}(S, A) = H(s) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} H(S_v) \tag{2}
\]

where \(H(S)\) - Entropy before the split, \(S_v\): No of instances of a particular class variable in one of the subsets of the split tree, \(S\): Total no of class variables in that particular subset of the split tree, \(H(S_v)\): Entropy of the particular subset of the split tree.

6. CLASSIFIERS

6.1 SVM

In classification problems, SVM is widely applied. It works based on supervised learning [27]. In SVM, a plane called a hyperplane is created to classify the data. An optimal separable hyperplane that separates the two classes is shown in figure 5. The bounding plane is called a margin, represented by a discontinuous line. SVM tries to maximize the margin to minimize the error. Support Vectors are the data that is near to the bounding planes which help in the determination of margin. The creation of a hyperplane is critical in SVM. In figure 5, two classes are considered (‘+’ and ‘-’). The hyperplane is obtained by solving the following equations (3) and (4).

\[
\text{minimize} \quad \frac{1}{2} ||w||^2 + c \sum_{i=1}^{L} \xi_i \tag{3}
\]

\[
\text{Subject to} \quad \begin{align*}
 y_i (w^T x_i + b) & \geq 1 - \xi_i \\
 \xi_i & \geq 0; i = 1 to L \tag{4}
\end{align*}
\]

Here the input vector is represented by \(x_i\) and the indicator vector is represented by \(y_i\) and \(L\) represents the total data points and \(w\) is the weight vector. The distance between the margin is measured by the slack variable \(\xi\), \(b\) gives the bias, and the penalty hyperparameter is represented by positive constant \(c\).

The prediction of belongingness of a particular class by the data points is given by equation 5.

\[
f(x) = (w^T x - y) \tag{5}
\]

The hyperparameter \(\gamma\) determines the curvature of the decision boundary. If the decision function \(f(x)\) is positive, data points are classified as ‘+’ otherwise ‘-’. [28].

![Figure 5. Working principle of SVM](image)

6.2 K-STAR

K-star is a lazy family classifier that works well with both instance-based and rule-based learning. It solves the smoothness problem by adding up all the probabilities in each conceivable path. This technique makes a significant contribution to its effectiveness. It includes a mechanism for dealing with missing values as well as the merging of real and symbolic valued features[4].

The entropic distance metric is used by this instance-based classifier. Parameters \(x_0\) and \(s\) for real and symbolic features should be set to a value for every other dimension. If the symbolic feature \(s\) has a value of about 1, data points with symbols other than one will have a very low transformation probability, while data points with the same symbol will have a much greater transformation probability.

As a result, the distance function will outperform its nearest neighbors. The probability distribution of symbols will be reflected by transformation probability when the value of symbolic features approaches zero. As a result, favorable symbols appear frequently [24].
7. RESULTS AND DISCUSSIONS

In this work, experiments were carried out with fault-free and different fault conditions of the milling tool, and its corresponding vibration signals were acquired. The time-domain plot shown in figure 4, depicts the variation in vibration amplitude between the different conditions of the tool.

Eleven statistical features were extracted from the acquired vibration signal. The representative statistical feature extracted from the vibration signals for four conditions of the tool is given in table 2. All the statistical features are not useful in classifying the conditions of the tool. A decision tree algorithm was used to select the most significant features. The information gain is computed by the decision tree using equation (2) for all the statistical parameters after computing the entropy from equation (1). A logic test is performed at each node to make decisions, thereby then rules are generated to discriminate the different tool conditions. The most significant features are identified as standard deviation, mean, median, sample variance, and kurtosis by the decision tree algorithm which contains the maximum information to discriminate the tool conditions and are utilized further for classification.

Figure 6 represents the part of THE DECISION TREE OBTAINED from the features. The ellipse represents the statistical feature and the rectangular box represents the conditions of the tool. These boxes form the leaves in the decision tree. From the tree, it is observed standard deviation forms the root node which contains the maximum information GAINED to classify the condition of the tool. The majority of the flank wear (FW) tool condition is classified by following the decision rule from the tree: if standard deviation and median are greater than 0.004472 and 0.07201 respectively, and sample variance is greater than -0.040545 and kurtosis is less than 4.637346, then the data belongs to FW. Similarly, the If-Then decision rules are generated by the decision tree to classify the various tool conditions taken in this study.

The discriminating capability of the tool condition by the features selected from the decision tree algorithm can be visualized through box plots. The box plots of standard deviation and kurtosis are represented in figures 7 and 8 respectively. In figure 7, the horizontal line present inside the blue colour box refers to the median of the standard deviation of the data corresponding to THE Good condition of the tool (G). The box covers 50% of the data referred TO AS THE INTER QUARTILE RANGE (IQR) [29], and the median line divides the data in equal proportions. The vertical lines are drawn on either side of the box which is referred to as whiskers. The length of the whiskers is drawn, 1.5 times of IQR or drawn up to the last data point. These whiskers cover most of the remaining data. However, the data which lies outside the whiskers are called outliers and are represented by dots. In figure 7, it is inferred that the standard deviation feature has different range distribution for various tool conditions which are represented as boxes.

Similarly, in figure 8, it can be inferred that the kurtosis feature can classify FW, C, and B tool conditions effectively as there is no overlap between the boxes. Whereas, there is an overlap between the G and FW boxes, which indicates that kurtosis is least prominent in classifying these tool conditions within the selected features.

The selected significant features were given as input to the SVM algorithm. The polynomial kernel function yielded higher classification accuracy.
Table 2. Details of Statistical Features

<table>
<thead>
<tr>
<th>Statistical Features</th>
<th>G</th>
<th>FW</th>
<th>C</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.039219794</td>
<td>-0.037041507</td>
<td>-0.0284527</td>
<td>-0.036961073</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.000491885</td>
<td>0.000515288</td>
<td>0.00044815</td>
<td>0.000414131</td>
</tr>
<tr>
<td>Median</td>
<td>-0.039263818</td>
<td>-0.037060256</td>
<td>-0.028647174</td>
<td>-0.037189302</td>
</tr>
<tr>
<td>Mode</td>
<td>-0.033467374</td>
<td>-0.056226044</td>
<td>-0.041704426</td>
<td>-0.052931223</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.070392882</td>
<td>0.073741983</td>
<td>0.064134076</td>
<td>0.059265674</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>0.004955158</td>
<td>0.00543788</td>
<td>0.00411318</td>
<td>0.00351242</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.3507962</td>
<td>3.99663516</td>
<td>1.210265481</td>
<td>2.247830533</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.0117191796</td>
<td>-0.059697216</td>
<td>0.005714433</td>
<td>-0.038439466</td>
</tr>
<tr>
<td>Range</td>
<td>1.04030916</td>
<td>0.989300453</td>
<td>0.69252252</td>
<td>0.67128923</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.549533936</td>
<td>-0.560211596</td>
<td>-0.399314514</td>
<td>-0.392358781</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.490775224</td>
<td>0.42908885</td>
<td>0.29320806</td>
<td>0.27893045</td>
</tr>
</tbody>
</table>

Figure 7. Sample Deviation Box Plot

Figure 8. Kurtosis Box Plot

The performance of SVM is shown in the form of a confusion matrix given in figure 9. In this work, 250 instances for each tool condition were considered. The correctly classified instances are shown in the diagonal of the confusion matrix. In the first row, 180 instances of the fault-free condition are correctly classified, whereas, 42 instances are misclassified as FW, 5 instances are misclassified as C and 23 instances are misclassified as B. For the tool condition FW, 229 instances are correctly classified, 7 instances are misclassified as G and 14 instances are misclassified as C.

Figure 9. Confusion Matrix of SVM

In the third row, all the 250 instances are correctly classified as C. For the other tool condition B, 248 instances are correctly classified and 2 misclassified as G. The importance in determining the algorithm’s quality is indicated by True Positive Rate (TPR) and False Positive Rate (FPR). TPR with 1 is considered highly accurate. For the acquired signal, TPR and FPR for the good condition of the tool are found to be 0.72 and 0.012. For the other conditions, TPR is more than 0.9. The overall classification accuracy of the SVM algorithm is found to be 90.7%. The class-wise detailed accuracy can be seen in table 3.

In figure 10, the performance of the K-Star algorithm is shown in confusion matrix form. It is understood that 213 instances are correctly classified as fault-free condition (G) of the tool from the 250 instances. The other elements in the first row represent the misclassifications namely, 33 instances of G are misclassified as FW and 4 instances are misclassified as B respectively. Likewise, in FW 242 instances are classified correctly and 5 misclassified as G, and 3 misclassified as C. However, in C 249 instances are correctly classified and 1 got misclassified as FW. In the fourth row, all the instances are correctly classified in B. The correctly classified instances from each condition occupied the diagonal of the matrix.
Figure 10. Confusion Matrix of K-Star

The detailed accuracy is presented in table 4. For good condition, the TPR and FPR are found to be 0.855 and 0.009 respectively. From table 4, it is observed that there is an increase in TPR for all the conditions of the tool considered in the study when compared with the detailed accuracy of the SVM algorithm given in table 3.

Table 3. Detailed accuracy - SVM

<table>
<thead>
<tr>
<th>Class</th>
<th>TPR</th>
<th>FPR</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0.720</td>
<td>0.012</td>
<td>0.952</td>
<td>0.720</td>
<td>0.820</td>
</tr>
<tr>
<td>FW</td>
<td>0.916</td>
<td>0.056</td>
<td>0.845</td>
<td>0.916</td>
<td>0.879</td>
</tr>
<tr>
<td>C</td>
<td>1.000</td>
<td>0.025</td>
<td>0.929</td>
<td>1.000</td>
<td>0.952</td>
</tr>
<tr>
<td>B</td>
<td>0.992</td>
<td>0.031</td>
<td>0.915</td>
<td>0.992</td>
<td>0.937</td>
</tr>
</tbody>
</table>

Table 4. Detailed accuracy – K-Star

<table>
<thead>
<tr>
<th>Class</th>
<th>TPR</th>
<th>FPR</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0.852</td>
<td>0.007</td>
<td>0.977</td>
<td>0.852</td>
<td>0.910</td>
</tr>
<tr>
<td>FW</td>
<td>0.968</td>
<td>0.045</td>
<td>0.877</td>
<td>0.968</td>
<td>0.920</td>
</tr>
<tr>
<td>C</td>
<td>0.996</td>
<td>0.004</td>
<td>0.988</td>
<td>0.996</td>
<td>0.992</td>
</tr>
<tr>
<td>B</td>
<td>1.000</td>
<td>0.005</td>
<td>0.984</td>
<td>1.000</td>
<td>0.992</td>
</tr>
</tbody>
</table>

The classification accuracy of good condition in the K-Star algorithm is 85%. The other conditions are classified with more than 96% classification accuracy. The overall classification accuracy of the K-Star algorithm is found to be 95.4%. The mean absolute error in SVM is 0.2581 whereas in K-Star it is 0.0355 and the root means squared error of SVM and K-Star are 0.324 and 0.131 respectively. This shows that the K-Star algorithm classifies the tool condition with less error. Since misclassification between the good and the faulty conditions is reduced in the K-Star algorithm and the measure of error is also less in K-Star, the reliability of the algorithm is relatively high when compared with SVM.

8. CONCLUSION

In this study, the condition monitoring of cutting tool has been carried out while machining mild steel workpiece with carbide inserts under different conditions of the tool such as fault-free tool(G), flank wear (FW), wear on rake face (C), tool with broken tip(B) by utilizing the vibration signal. Experiments were carried out under specific machining conditions. For each condition, 250 instances were obtained. Statistical features are extracted from the acquired vibration signal. From the decision tree, significant features are selected and given as input to SVM and K-Star algorithms. The classification accuracy of SVM and K-Star are obtained as 90.7% and 95.4% respectively. The computation time of SVM is found to be 0.09 seconds whereas K-Star clocked in 0.03 seconds. Hence, it is concluded that the K-Star algorithm with statistical features can be recommended for face milling tool condition monitoring.

REFERENCES


Резултати се пореде са перформансама К-Стар алгоритма. Добијена тачност класификације је охрабрујућа, стога се студија препоручује за примену у реалном времену.