
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 25, Issues 5, pp. 3725–3740, DOI 10.12694/scpe.v25i5.3166

MACHINE LEARNING BASED TOOL WEAR PREDICTION FROM VARIABILITY OF ACOUSTIC SOUND EMISSION SIGNALS

N V. KRISHNAMOORTHY *AND JOSEPH VIJAY †

Abstract. A novel machine learning-based model is introduced in this research paper to forecast tool wear using acoustic emission (AE) signals. Adaptive boosting (AdaBoost) and a sophisticated feature engineering strategy are employed by the model to enhance the precision of its predictions. The proposed model, Machine Learning Tool Wear Prediction (MLTWP), analyzes AE signals generated during machining operations to distinguish between healthy and worn-out tool conditions with remarkable accuracy. The crux of our approach consists of meticulously eliminating and enhancing the temporal and spectral characteristics of the AE signals. We employ the Kolmogorov-Smirnov test to identify the most valuable classification features. We implemented AdaBoost with the objective of progressively enhancing a set of weak classifiers' ability to identify instances that were incorrectly classified in previous iterations. Utilizing this method increases the model's sensitivity to minute variations in tool wear conditions and its overall classification precision. The MLTWP model underwent extensive testing on a benchmark data set comprising 25,304 AE signal records from cutting mill tools, using a training tool split of 9,989 worn (positive) and 8,990 benign (negative) instances. The results of our experiments, validated through four-fold cross-validation, indicate that the MLTWP model exhibits superior performance compared to the existing Tool Wear Prediction using Acoustic Emission Signals (TWPAE) model. To provide greater specificity, the MLTWP exhibited the following metrics on average: precision (92.2%), specificity (91.38%), sensitivity (90.42%), accuracy (90.9%), and MCC (81.72%). The fact that these metrics exhibit significant improvement over those of the TWPAE model demonstrates that our method of feature engineering and adaptive boosting is effective at precisely predicting tool wear. This research not only advances the existing understanding of tool wear prediction but also establishes a robust framework for the implementation of machine learning in manufacturing predictive maintenance.

Key words: Machine Learning; Tool wear Prediction; Acoustic Emission Signal; and Artificial Neural Network; Kolmogorov-Smirnov (KS) Test

1. Introduction. Tool wear has a direct impact on the energy required to remove metal that in turn has an impact on dimensional accuracy, surface roughness, as well as the cutting operations [1]. Tool wear results from chemical, thermal, and mechanical interactions between the materials of the tool and the workpiece. The two primary forms of tool wear, which might signal the end of a tool's life are crater wear as well as flank wear (Vb), which are both caused by these interactions [2]. The efficiency of the procedure relies on flank wear. This variable shows wear from the industry. It has been extensively researched that flank wear raises cutting forces [3], in turn which raises the power tool of machine tools. Forces, power use, and the dynamics of tool-material interaction are all impacted by flank wear. Acoustic emissions (AE) are used to measure flank wear during turning movements, as discussed in [4].

Because of the tool's variable working environment, complex operating conditions, and various cutting settings, tool wear prediction is challenging. Tool wear but also noise signals are only two examples of the information-rich sensor signals with in mechanical tool process. To determine the level of tool wear, signals should be normalised and features extracted.

Noise Monitoring the wear of a tool requires signal processing. The processing of data via the time-frequency analysis is common [5]. The preprocessing of the signal and manual extraction of tool wear features from the signal are common signals for traditional wear detection systems. Data on several cutting parameters under various operating situations must be obtained through numerous tests and manual analysis. These techniques won't work in all situations and have limitations in actual application.

Plastic deformation, corrosion, fracture propagation, impact, erosion, or leakage are the main causes of AE

* Department of Mechanical Engineering, Sri Krishna College of Engineering and Technology, Coimbatore (Corresponding author, nvkrishnamoorthy1968@gmail.com)
†Karunya Institute of Technology and Sciences Coimbatore (vijayjoseph@karunya.edu).

[6]. For non-destructive testing, AE is used in refineries, pipelines, nuclear and conventional power production, aero planes, offshore oil platforms, and paper mills, including structures such as bridges and cranes [7]. For the last 30 years, tribologists, who study the interaction between moving surfaces (such as during machining operations), have concentrated on AE due to the precision with which it can detect surface wear.

AE signals are produced by friction processes range from 50 kHz to 1 MHz [8]. The AE sensor, to avoid signal attenuation, must be attached to the cutting tool or the workpiece. Since both the cutting tool as well as the workpiece are transient, the sensor in practical applications must be positioned in the tool holder.

A simple machine learning model is provided to forecast tool wear. It was inspired by recent promising work in tool wear monitoring using machine learning [9], [10].

2. Related Research. An electronic-mechanical system was created by Weller et al., [11] that uses sonic signals to assess the cutting edge wear of turning tools. In tests, the system is able to identify used cemented carbide tools that can cut AISI 1045 steel. A cutting tool quality check based on sound was developed by Mannan et al., [12]. In order to identify tool wear related sound patterns, the suggested approach enables process of machining sound signals. This technique has been proven to locate cutting instruments that are sharp, semi-sharp, and dull. Chatter was discovered by Delio et al., [13] utilising sound signals. Experiments demonstrate that the chatter detection technique based on audio signals performs as well to dynamometers, accelerometers, as well as displacement probes. A method for legitimate tool status monitoring specifically for turning was created by Salgado and Alonso [14]. One-dimensional spectral analysis of feeding motor current but also audio signals showed direction-changing data. By analysing retrieved characteristics and applying a SVM (support vector machine) technique, tool wear was calculated. To measure tool wear, Aliustaoglu et al., [15] employed microphone-captured audio signals and two-stage fuzzy logic. Testing for drilling was done on a four-axis Cnc turning centre. Sound signals were recorded with the microphone. In studies, two-stage fuzzy logic may identify tool wear. The audio signal processing was employed by Ubhayaratne et al., [16] to track the tool wear in sheet metal stamping. Stamping audio signals were cleaned and preprocessed using semi-blind signal extraction. A technique for determining the degree of tool wear based on vibrations from the machine's spindle was developed by Seemuang et al., [17]. The drilling sound signals were recorded using a low-cost microphone. Support vector machines were utilised by Kothuru et al., [18] to categorise tool wear problems (SVM). Features in the frequency domain were sanded out of audio signals. Over 90% prediction accuracy.

ANN (Artificial Neural Network) should be used to assess feeder machine current, cutting, as well as source of feed, according to Alonso et al (ANN). Maximum vibration length is inversely related to unidirectional tool wear, as shown by Sadettin et al., [19]. Ming-Chyuan et al., [20] as well as Alonso F.J. et al., [21] discovered that the sole audible sign of tool flank wear is when a machine produces a noise when cutting. In order to determine tool flank wear rapidly, According to Peng et al., [22], if the signal is treated linearly and travels, obtaining the Fourier spectrum might produce a considerably lower physical sensation. According to Huang et al., [23], Fourier processing produces global signal qualities as opposed to local ones. Another technique for estimating outcomes is the Hilbert-Huang transformation [24]. Prakash et al., [25], [26] has shown how streamlining a procedure produces good outcomes and a simpler way.

A contemporary model that relates to the objective of this article is projected by Ferrando Chacon et al., [27], which is referred as "Tool Wear Prediction using Acoustic Emission Signals (TWPAE)" in further discussions. The contemporary model TWPAE has used acoustic emission signal frequencies as features and focused on the process of eliminating redundant features as well as predicting worn state of the tools using machine learning. However, the feature engineering of this model is suboptimal as it is limited to eliminating redundant features and restricted only to the signal frequencies.

Yang, Cheng et al. [28] proposed a novel approach to understanding how milling tools wear down over time. This method combines an analysis of wear-related factors with a wear-prediction model, with a focus on how tool wear works and changes over time. The method improves the accuracy of tool wear predictions by combining experimental data and modeling. The goal is to improve the efficiency of machining processes and reduce downtime.

Perumal et al. [29] presented a simple machine learning method for monitoring tool wear in real time using sound signals. Their findings revealed that basic acoustic emission (AE) signals could be used to predict tool wear, making this a simple and effective method for tool condition monitoring (TCM) in manufacturing settings.

The study also demonstrated that linear regression algorithms can be used to calculate tool wear based on AE data. This could be a low-cost way to increase machining efficiency and tool life.

Qian et al. [30], investigated how BiLSTMA networks can be used to predict tool wear and how difficult it is to demonstrate how wear properties change over time. Their model examines sensor data and makes more accurate predictions than older methods by utilizing bidirectional long short-term memory (BiLSTM) architectures and attention mechanisms. This method demonstrates the importance of advanced neural network models in understanding and predicting how tool wear changes over time. He et al. [31] developed a deep multi-task network that uses sparse feature learning to monitor tool wear and machining quality simultaneously. Their framework combines deep learning methods to extract and connect features from various types of data. This gives them a comprehensive tool for determining how tools are performing and ensuring the quality of the products they produce. The findings show that multi-task learning and sparse representation work well together to improve prediction accuracy and operational efficiency in manufacturing systems.

Twardowski et al. [32] used acoustic emission techniques and machine learning methods to detect tool wear. Their findings show that AE sensors are effective at detecting changes in the condition of tools, and that machine learning algorithms can be used to analyze complex sensor data to determine how much wear something is experiencing. This study adds to the growing body of evidence arguing that AE monitoring and artificial intelligence should be used in tandem for proactive cutting tool maintenance.

The predicting and monitoring tool wear advances the field by introducing new approaches, such as combining wear-related factors, employing basic to advanced machine learning methods, and using acoustic emission signals for real-time monitoring. Each study demonstrates the potential for improving tool condition assessment accuracy and operational efficiency. However, they also highlight issues such as the need for a large amount of experimental data, the difficulty and computational requirements of advanced neural networks, and the feasibility of implementing these technologies in a variety of manufacturing environments. A thorough examination reveals that more research should be conducted on how to make these methods work in the real world with various types of machining operations, how to make them more scalable, and how to make complex models easier to understand. As this body of work demonstrates, tool wear monitoring research is constantly evolving. It predicts a future in which technology is increasingly used in manufacturing processes.

Due to the low specificity and sensitivity of existing tool wear prediction methods, it was determined from this analysis that using acoustic emission signals as input to machine learning based tool wear prediction is an additional viable research aim. Here, feature engineering for educating machine learning models is the main concern.

To better forecast tool wear from acoustic emission signals, this research contributes by concentrating on temporal and spectral aspects and an optimization approach based on a statistical diversity evaluation technique to overcome these limitations.

3. Methods and Materials. The processes and resources that went into developing the suggested model are outlined below. The features that will be used throughout the proposal's learning phase, as well as the appropriate features extraction from the provided labelled data. This section also discusses the strategy used to compare the projected values of an attribute over the provided records of the both positive and negative class labels. Moreover, this part investigates the appropriate classifier for detecting acoustic emission signals. We have also looked at the proposal's testing phase (identifying acoustic emission signals).

3.1. Acoustic Emission Signals. Transient elastic waves are produced when a material undergoes a quick redistribution of stress, a phenomenon known as acoustic emission (AE). When a structure is stressed (by load, pressure, as well as temperature change), energy is released from localized sources as stress waves, which travel to the surface and yet are measured by sensors. Picometer (10-12 m) movements can be detected with the correct equipment. AE can be triggered by a variety of events, including earthquakes, rock explosions, fissures, slip and dislocation motions, melting, twining, and phase transitions in metals. In the case of composites, acoustic emissions result from matrix cracking, fiber breaking, and debonding. There are other AEs in polymers, wood, and concrete.

A material discontinuity's origin and importance can be revealed via its AE signal. Besides its extensive usage in academic studies, AET (acoustic emission testing) is also routinely used in a variety of industrial

contexts (e.g., evaluating structural integrity, locating defects, determining leak rates, and keeping tabs on weld quality).

In testing to conventional NDT techniques, Acoustic Emission stands apart in two key aspects. To begin, let's talk about the origin of the signal. Instead of supplying energy, AET just listens for energy emitted by the item being studied. Acoustic emissions can be triggered and propagated in operational structures by applying appropriate loading during AE testing. The discipline of AET is dedicated to the investigation of material dynamics. Not inactive characteristics (like a crack that isn't expanding) are shown. It is critical to detect both new and persistent flaws. Defects could be unnoticed when any loading becomes too low to trigger an acoustic event. AE testing provides instant feedback on the reliability of a part. With AET, there is no need to disassemble or clean the specimen, and yet a thorough volumetric assessment can be performed quickly and accurately thanks to the use of several sensors and a permanent mounting system that also allows for process control.

3.2. Preprocessing of the AE signals. In the process of digitizing the gathered sound signals, they frequently mix in with signals from other sources whose waveforms and arrival timings are not known. In order to extract the physical characteristics that best correspond with tool wear, signal processing is employed as part of the pre-processing stage of TCM systems for machining. Sound characteristics are extracted using a variety of signal processing techniques [33]. In this investigation, 1-second samples were taken from sound signals and analysed in the temporal domain. Frequency domain analysis samples were processed using FFT (Fast Fourier Transform)-based DTFT (Discrete-Time Fourier Transform). This research used a single absolute value from the Fourier transform to build a decision-making model. In order to resolve the DTFT-PR (preset resolution) issue in future research, the WT (Wavelet Transform) will be used.

Sound signals under the same cutting settings and at good and tool wear levels are displayed as a frequency spectrum in Figure 3.1. The spectrum was shown logarithmically to draw attention to the low-frequency band, where signals include more information about tool wear. Multiple frequency amplitudes are shown to correlate to tool wear in the images, highlighting the need for a sophisticated decision-making model to accurately depict these relationships.

3.3. Features. Temporal characteristics are those in the time domain, including signal energy, less energy, 0% of crossing, and many more, that are easy to extract and have physical comprehension.

- Since the stochastic process stated as a mathematical approach for acoustic emission only considers random variables, acoustic emission spectrum analysis is often more advanced than deterministic signals analysis, such as sinusoids.
- Static acoustic emission signals require a spectral phase that is completely random and, thus, devoid of information. The use of constant acoustic emission signals may be to blame for this phenomenon; by definition, it cannot occur randomly at times. Therefore, the phase idea is irrelevant for acoustic emission signals, which is a major difference between these signals as well as deterministic spectrum signals.
- So, spectral properties like frequency components, spectral density, fundamental frequency, and many more have been achieved by transforming a time-based signal into the frequency domain using Fourier-transform (FT). In this context, these characteristics might be put to use in identifying tones, rhythms, melodies, and pitches.
- Audio signals are often complex, made up of many individual shock wave of a fixed wavelength that propagate as a single perturbation in the medium. While the sound is being captured, we record the amplitudes of the ensuing waves. To break down a signal into its frequency components, mathematicians use a concept known as the Fourier Transform (FT). FT also provides the magnitude of each frequency in the signal, in addition to the frequency spectrum.

In the study of sound, novices often employ the broad spectral characteristic, which may take the form of a rising, decreasing, or varying frequency. The FT creates harmonic sound generators called as signal's spectral structure in acoustic emissions. As said before, there are a variety of ways this might be demonstrated. In other words, if the initial partial harmonic of a periodic signal, whose vibrations have repeated throughout time, is 100 Hz, the subsequent partial harmonics will be 200 Hz, 300 Hz, etc. This may, however, be an unusual occurrence in the realm of sounds created by nature. Make a note to the effect that no spectral characteristics

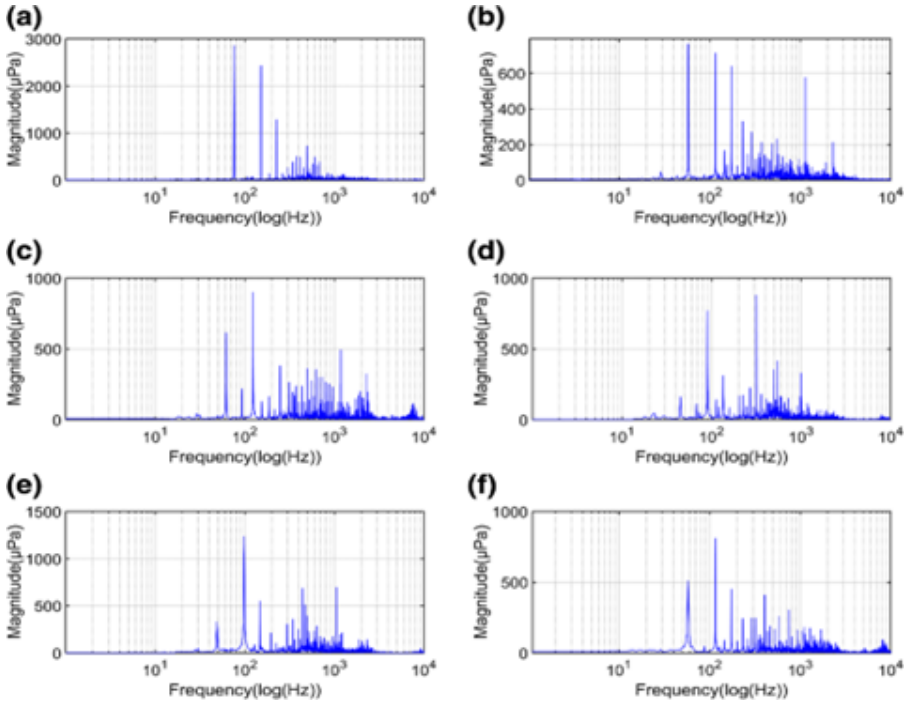

Fig. 3.1: Sound signals emitted from worn and benign tools at the same cutting settings

or devices are related to any suggested laws in the actual world where we could perceive sound. MFCC (" Mel cepstral frequency coefficients") as well as "predictive linear perceptive" (PLP) feature were the two best spectral features used in the acoustic emission. In this case, it may be understood that the frequency content of acoustic emissions is evaluated by the cochlea in the eardrum. Bandpass filters with a constant Q rank made it possible to modify the inspection of the basilar membrane. There is also the presence of essential bands, which rise to mask phenomena, where a burst or one strong acoustic emission may hide another lesser acoustic emission inside an essential band. The features of the auditory system that allow for the identification of acoustic emission are captured by both PLP and MFCC. Nonetheless, the strength of these characteristics was not very pronounced with respect to the acoustic emission.

The acoustic emission is also supported by a few time-domain characteristics, such as the plosion index and the highest coefficient of correlation.

In many complicated signals, such as acoustic emission, the signal's characteristics change over time. Small-time FT plays an important role in the interpretation of divergent acoustic emissions and may be more useful for referring to changes in frequency components over a finite time period. Integrating temporal envelopes over short time periods allows FT to derive spectral characteristics, while temporal-features are generated by turning envelopes into frequency modulation devices.

Temporal characteristics included things like ZCR, energy, etc., whereas spectral features included things like MFCC, GTCC, and several others in the frequency domain.

3.4. Kolmogorov-Smirnov test (KS Test). The significance of any differences between samples may be swiftly determined by using the Kolmogorov-Smirnov test [34]. As a uniformity, it ensures that all values in a distribution are consistent with one another. The Kolmogorov-Smirnov test allows one to examine the uniformity of a distribution, an important feature. In addition to comparing two multi-dimensional probability distributions, the Kolmogorov-Smirnov test may also evaluate the similarity of two one-dimensional distributions. It can immediately tell whether there is a difference between two samples. The Kolmogorov-Smirnov statistic evaluates the dissimilarity between the empirical distribution functions of two samples or between the

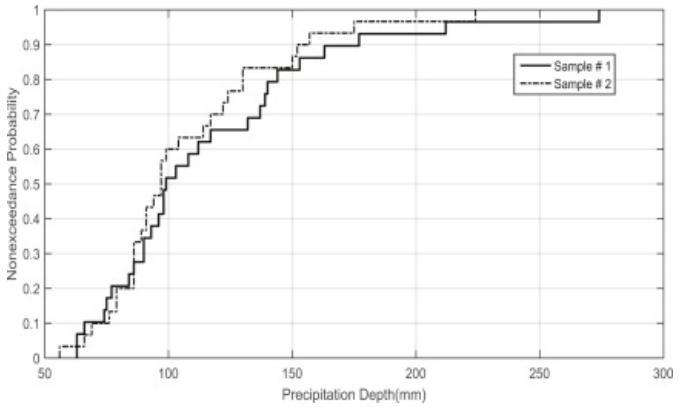

Fig. 3.2: Cumulative distribution function

sample and the reference distribution.

Distributions of observations in the two data sets may be compared using the Kolmogorov-Smirnov (KS) test [34]. Each dataset's values are assumed to come from a given continuous distribution (the null hypothesis). As an alternate hypothesis, these data sets come from several continuous distributions. It is possible to do the hypothesis test with a 5% level of statistical significance. Cumulative distribution functions (CDF) for two different datasets are depicted in figure 3.2. The null hypothesis is supported by a two-sample KS test, which returns the D value of 0.1782 as well as a P-value of 0.694. At the 5% level of significance, the KS test on such two datasets failed to disprove the null hypothesis, showing that the given samples are drawn from the identical continuous distribution. This test can only conclude that, the distributions are distinct; it cannot reveal any changes to the mean, variance, or extremes.

The usage of "KS-Test" method explained in following description. In order to estimate the diversity between two vectors v_a, v_b representing the samples of two different distributions, $ks - test(v_a, v_b)$ Given two vectors v_a and v_b, the aggregate values of these vectors are represented as follows: Eq 3.1, Eq 3.2

$$\|v_a\| = \sum_{i=1}^{|v_a|} \{e_i \mid e_i \in v_a\} \tag{3.1}$$

$$\|v_b\| = \sum_{i=1}^{|v_b|} \{e_i \mid e_i \in v_b\} \tag{3.2}$$

The cumulative ratio for each sample in the vectors v_a and v_b is predicted as follows (Eq 3.3):

$$\forall_{(i=1)}^{|v_a|} \left(v_a^{cr} \leftarrow \sum_{j=1}^{i} \left(\frac{e_j}{\|v_a\|} \mid e_j \in v_a \right) \right) \tag{3.3}$$

Here, Eq 1 determines the cumulative ratio of each e_i of the elements listed in vector v_a. The $\|v_j\|$ indicates the aggregate value of the samples listed in the vector v_a, and v_a^{cr} is a set that comprises cumulative ratios of the samples listed in the vector v_a.

Further, it discovers cumulative ratios v_b^{cr} of the samples listed in vector v_b.

The absolute distance of "cumulative ratios" corresponding to values that are present at a similar index of both v_a and v_b vectors is found as: Eq 3.4

$$\forall_{(i=1)}^{(\max(\|v_a^{cr}\|, \|v_b^{cr}\|))} \{D[i] = \sqrt{(v_a^{cr}[i] - v_b^{cr}[i])^2}\} \tag{3.4}$$

Eq 2 for each index i, represents the "absolute distance $D[i]$" of "cumulative ratios $v_a^{cr}[i]$, $v_b^{cr}[i]$.

The maximum value of the set D is found as d-stat. Then, find d-critic (from the KS-Table) at a given "degree of probability threshold" p_τ (0.01, 0.05, or 0.1) for vector aggregates $\|v_a\|$, $\|v_b\|$.

The return value is determined by the following condition:

$$\begin{cases} \text{false} & \text{if } d\text{-stat} > d\text{-critic} \\ \text{true} & \text{if } d\text{-stat} \leq d\text{-critic} \end{cases} \tag{3.5}$$

If d-stat is greater than d-critic, then there is no diversity between the given vectors' distribution. Hence, return false; else return true.

3.5. The classifier. Adaboost was designed to improve the performance of binary classifiers by using an ensemble learning approach. Adaboost uses iterative learning to learn the shortcomings of weak classifiers. The method of adaptive boosting [35] is utilised by the Adaboost classifier. Multiple, rather weak Boolean classifiers are combined to form this weak classifier. Every weak classifier has a single true/false criterion to divide the data. False-positives as well as false-negatives might be separated by other Weak Classifier. This continues until the WC calls time on the competition. The output of each WC will be combined to form the final output of the classification operation. In this section, the WC represents the proposed model's binary-classification optimal features, which are characteristics that are both consistent with the class and different from it.

Weak-Classifier iterations of the classification process allow the part of the corpus that didn't need to be categorized exactly to be boosted for the subsequent classifier iteration. Classification weight (WC) is utilised at each iteration to show the relative importance of different categories. Iterative calls to the WC will result in a standardization of classified records. Adaboost WC provide a concrete n-gram for precision classification and provide a justification for record polarity classification in the proposed approach. The suggested approach would use the WC to learn the suitability of each feature category for both positive and negative sentiment training corpora.

All data points are given equal consideration when Adaboost generates a model. Incorrectly categorized data are given more weight. In the following paradigm, more significant information is given more weight. Models will be trained until there is a noticeable improvement in accuracy. What follows is a description of the internal structure of this algorithm. Before doing binary classification, Adaboost assigns weights to data points. At the outset, irrespective of the overall data points, each data point has the same weight. In order to this, produce a decision stump from each of the features and then assess its tree accuracy using the Gini Index. In terms of the Gini Index, the first tree stump represents the worst possible result. Classified data points using this tree's alpha ("Importance" or "Influence") is based on the following Eq 3.6.

$$2^{-1} \cdot \log_e \left(\frac{1 - \|w\|}{\|w\|} \right) \tag{3.6}$$

Aggregating the all of the sample weights related to data points that are incorrectly labelled can define the total error $||w||$.

Consider a dataset having 5 data points as well as the assumption that there is 1 incorrect output; the total error then will be 1/5 (ranges between 0 and 1), then the stump's performance (alpha) will be: Eq 3.7, Eq 3.8, Eq 3.9, Eq 3.10

$$\alpha = 2^{-1} \cdot \log_e \left(\frac{1 - 5^{-1}}{5^{-1}} \right) \tag{3.7}$$

$$\alpha = 2^{-1} \cdot \log_e \left(\frac{0.8}{0.2} \right) \tag{3.8}$$

$$\alpha = 2^{-1} \cdot \log_e(4) = 2^{-1} \cdot (1.386) \tag{3.9}$$

$$\alpha = 0.693 \tag{3.10}$$

The error rate for alpha () ranges from 0 to 1, where 0 Indicates perfect stump and 1 indicates awful stump (see figure 3.3).

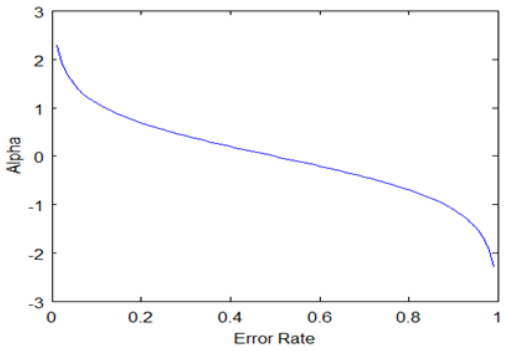

Fig. 3.3: Alpha versus error rate

As can be seen in the above graph (figure 3.3), whenever there is no misclassification, there is no error (Total Error = 0), and hence the alpha is quite high.

The classifier's alpha will be zero if its predictions are half accurate and half wrong, with a total error of 0.5. A negative alpha value indicates that the classification was erroneous for all samples. If this is the case, the error is extremely large (approximate to 1), and the significance level will be low.

When the same weights being used to the next model, therefore the output obtained is identical to what's been obtained in the first model, hence updating the weights is required in order to determine the total error $||w||$ as well as performance of a stump.

Correct forecasts will have their weights lowered while incorrect guesses will be given greater weight. After adjusting the weights, we will now give the data points with maximal weights more preference when building our next model.

It is important to update the weights that use the given formula after the classifier's significance and the total error have been determined (Eq 3.11):

$$cw = w \cdot e^{\pm\alpha} \tag{3.11}$$

New weight cw, the alphais being negative whenever the sample is correctly classified. The alpha is being positive whenever the sample is miss-classified.

3.6. Tool Wear Prediction Model. The set SC of acoustic emission signals that are taken as input for training phase of the classification process shall be processed to digital format and then bi-part in to sets $PandN$, where set P represents the acoustic emission signals of the wearied tools (labeled as positive). In contrast, set N represents the acoustic emission signals of the benign tools (tools that are not wearied), which are labeled as negative.

Further, derives all the temporal and spectral features of acoustic emission signals listed in the set P as two different matrices named as tP, sP , which are representing temporal and spectral features in respective order. Here, each row of the both matrices representing the respective features of the corresponding acoustic emission signal listed in the set P . Similarly, derives all the temporal and spectral features of acoustic emission signals listed in the set N as two different matrices named as tN, sN , which are representing temporal and spectral features in respective order.

Further step of the learning phase discovers optimal temporal and spectral features of the both positive and negative labels as described below (Eq 3.12, Eq 3.13):

$$\forall_{(i=1)}^{\max(|tP|,|tN|)}\{tP[i], tN[i] \mid |tP[i]| > 0 \vee |tN[i]| > 0\} \tag{3.12}$$

// Eq 3.12 For each iteration, considers the vectors $tP[i], tN[i]$ representing the values listed in $i(^th)$ column of the matrices tP and tN in respective order.

$$\text{if}\,(\text{ks-test}(tP[i], tN[i]) \equiv \text{false}) \quad tP \setminus tP[i] \wedge tN \setminus tN[i] \tag{3.13}$$

This Eq 3.13 discovers the diversity state between two vectors $tP[i]$ and $tN[i]$ through KS-Test, ks-test$(tP[i], tN[i])$. If the diversity state is false, then the columns of the matrices tP, and tN represented by the vectors $tP[i]$, and $tN[i]$ in respective order are discarded. Conversely, if the diversity state is true, then the feature represented by the ith column of the matrices tP, and tN will be considered as an optimal feature towards the positive and negative labels.

The resultant matrices tP and tN retains the columns representing optimal features toward both positive as well as negative classes in respective order. Similarly, the KS-test shall be applied on each pair of the columns listed at the same index of the matrices sP and sN that are representing spectral features of the positive as well as negative classes in respective order, which results matrices sP and sN with optimal spectral features of the both lasbels. Further phase of the learning process builds ensemble of weak classifiers to perform adaptive boosting-based classification of the given acoustic emission signals as the signal emitted from the wearied tools and signals emitted from the benign tools.

The following diagram presented as figure 3.4 outlines the intricate process of the Tool Wear Prediction Model, illustrating the transformation of acoustic emission signals into a digital format and their subsequent classification. Through various stages, including feature extraction, optimization, and iterative adaptive boosting, the model aims to precisely predict the wear of tools. This visualization offers a comprehensive understanding of the complex mechanics underlying the process.

Acoustic Emission Signals Phase

- Acoustic emission signals are collected in set SC.
- These signals are then processed into a digital format.
- The digital signals are bifurcated into two sets, P and N, representing wearied (positive) and benign (negative) tools, respectively.

Temporal and Spectral Features Phase

- Temporal and spectral features are extracted from the signals in sets P and N.
- These features are organized into four different matrices: tP, sP for positive and tN, sN for negative classes, representing temporal and spectral features, respectively.

KS-Test & Optimal Features Phase

- The KS-test is applied to discover optimal temporal and spectral features for both positive and negative labels.
- Resulting optimal features are represented by matrices $tP\&tN$ and $sP\&sN$.

Weak Classifier Iterations Phase (Adaptive Boosting Process)

- AdaBoost WC (Weighted Classifier) is used to initiate the weak classifier iterations.
- A decision stump is created from each feature, and its tree accuracy is evaluated using the Gini Index.
- The alpha (importance or influence) and weights are updated, and weights are given to the incorrect guesses while reducing the correct ones.
- The iterative process continues, building upon the previous model and adjusting weights to give more preference to data points with maximal weights.
- The final model is created through iterative calls to the weak classifier, progressively emphasizing the importance of certain features.

Additional Notes

- There is an iterative updating of weights using a specific formula: $cw = w \cdot e^{(\pm\alpha)}$. This allows for the progressive adjustment of the model based on correctly and incorrectly classified samples.

3.7. MLTWP Algorithm Flow. Given a set of acoustic emission (AE) signals, the following algorithm outlines the steps to predict tool wear:

1. Collect AE signals into a set SC and process them into a digital format.
2. For positive set P and negative set N,
   - Extract temporal and spectral features, and store in matrices tP, sP, tN, and sN respectively.
3. Apply the Kolmogorov-Smirnov (KS) test to identify optimal features for each i^{th} feature:

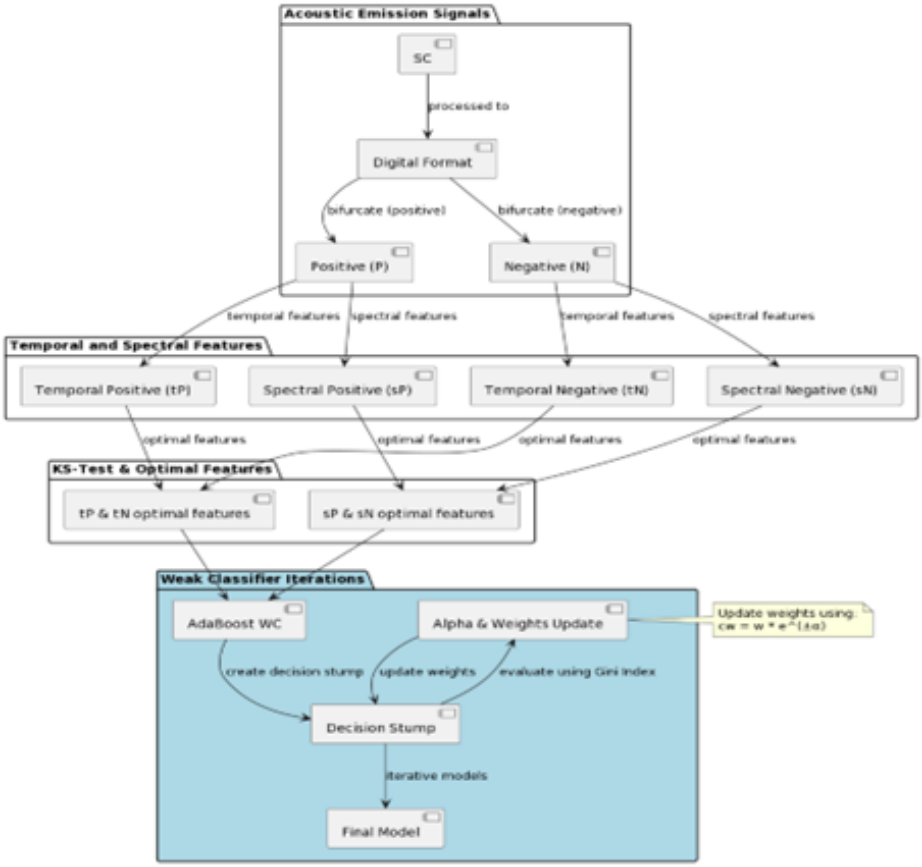

Fig. 3.4: The flow diagram of the MLTWP

- if KS_test$(tP[i], tN[i])$ is significant, retain $tP[i]$ and $tN[i]$.
- if KS_test$(sP[i], sN[i])$ is significant, retain $sP[i]$ and $sN[i]$.

4. Initialize AdaBoost with weak classifiers derived from the optimal features.
5. For each iteration k, perform:
   (a) Create a decision stump h_k from the features.
   (b) Calculate the error err_k of h_k.
   (c) Compute the weight α_k of h_k:

$$\alpha_k = \frac{1}{2} \log \left(\frac{1 - err_k}{err_k} \right).$$

   (d) Update weights w for each data point:

$$w \leftarrow w \cdot e^{-\alpha_k y h_k(x)},$$

   where y is the true label of x.
6. Combine the weighted decision stumps to form the final model M:

$$M(x) = \text{sign} \left(\sum_k \alpha_k h_k(x) \right).$$

7. Predict the wear condition of a new AE signal x using $M(x)$.

Table 4.1: Assumptions for Experimental Setup

	Description of Assumption
Assumption 1	All acoustic emission signals are accurately recorded without data loss.
Assumption 2	The dataset is fully representative of all possible states of tool wear.
Assumption 3	The labeling of the dataset into 'positive' and 'negative' is error-free.
Assumption 4	The temporal and spectral feature extraction processes are lossless.
Assumption 5	The Kolmogorov-Smirnov test reliably identifies the optimal features.
Assumption 6	The training and testing datasets are randomly partitioned to avoid bias.
Assumption 7	The four-fold cross-validation provides an unbiased estimate of performance.
Assumption 8	The AdaBoost algorithm converges to an optimal set of weights and stumps.
Assumption 9	The Gini Index is an adequate measure of decision stump performance.
Assumption 10	The MLTWP and TWPAE models are implemented under the same conditions.
Assumption 11	The performance metrics (precision, specificity, etc.) are computed without computational errors.
Assumption 12	The underlying distribution of the data remains stationary during the experiment.

The algorithm utilizes iterative boosting to refine the classification model, enhancing its predictive accuracy for tool wear based on AE signals.

4. Experimental Study. Experiments have been conducted on benchmark dataset to scale the performance of the proposed model MLTWP. The dataset adopted is a set of acoustic emission signals recorded from cutting mill tools both wearied and benign. Here, the acoustic emission signals of the wearied tools have been labelled as positive and the acoustic emission signals of the benign tools have been labeled as negative. The experimental data produced from the proposed model MLTWP and the state-of-the-art model Tool Wear Prediction using Acoustic Emission Signals (TWPAE) [27] have been compared in order to estimate the performance of the proposed model MLTWP. The dataset statistics have been explored in following section.

4.1. The dataset. The total records of the cutting mill dataset [36] are 25304, which possess the records of acoustic emission signals obtained from worn and benign tools. The acoustic emission signals of worn tools have been positive, which are of total 13318. The acoustic emission signals of unworn tools have been labeled as negative, which are of total 11986 records. 8990 negative records have been utilised for training, while there are 9989 positive records for training. Additionally, 2996 negative records and 3329 positive records have been tested. Further, each of the both positive and negative record sets will be partitioned in to four parts, such that 3 parts of the records from each label will be used to train the proposed model MLTWP and contemporary model TWPAE. The rest partition of the both positive and negative records will be used in testing phase. A table of assumptions for the experimental setup is as follows in Table 4.1

The table 4.1 summarizes the foundational assumptions made during the experimental phase of the study, ensuring clarity and transparency for replication and validation purposes.

4.2. The Performance Analysis. The performance of the proposed model MLTWP has been scaled through the results obtained for cross validation metrics from four-fold cross validation as shown in table 4.2, which have been compared with results of the fourfold cross validation metrics obtained from the experiments carried on the contemporary model TWPAE. The following discussion analyses the performance of the both proposed MLTWP and contemporary model TWPAE.

The obtained results for cross validation metrics demonstrate that the anticipated MLTWP model outperforms the existing TWPAE model. The precision statistic represents the percentage of correctly identified positive records relative to the total number of positive records. The accuracy of the predicted MLTWP model, as measured by four-fold cross validation, as 93%, 92%, 92, and 91%, in that sequence. On the other hand, the observed precision for the current technique TWPAE as 89.09%, 89%, 89%, and 88%. When comparing the suggested model MLTWP to the state-of-the-art TWPAE, the proposed model performs better on the cross validation metric called precision (see Figure 4.1).

Specificity might be defined as the fraction of records that were correctly categorised as negative relative to the total number of negative-records. Here, we see that the specificity for the MLTWP model were 92%,

Table 4.2: Statistics of cross validation metrics obtained from 4-fold classification scheme

PRECISION				
	FOLD#1	FOLD#2	FOLD#3	FOLD#4
MLTWP	0.9256	0.9201	0.9231	0.9191
TWPAE	0.8909	0.8851	0.8872	0.8839
SPECIFICITY				
MLTWP	0.9188	0.9132	0.9197	0.9118
TWPAE	0.8789	0.8714	0.8732	0.87
SENSITIVITY				
MLTWP	0.9085	0.8994	0.9022	0.9016
TWPAE	0.8908	0.8903	0.8975	0.8894
ACCURACY				
MLTWP	0.9134	0.9059	0.9089	0.9064
TWPAE	0.8851	0.8814	0.886	0.8803
F-MEASURE				
MLTWP	0.9222	0.9166	0.9199	0.9154
TWPAE	0.8849	0.8781	0.8801	0.8769
MCC				
MLTWP	0.8266	0.8118	0.8178	0.8127
TWPAE	0.7695	0.7621	0.7713	0.7599

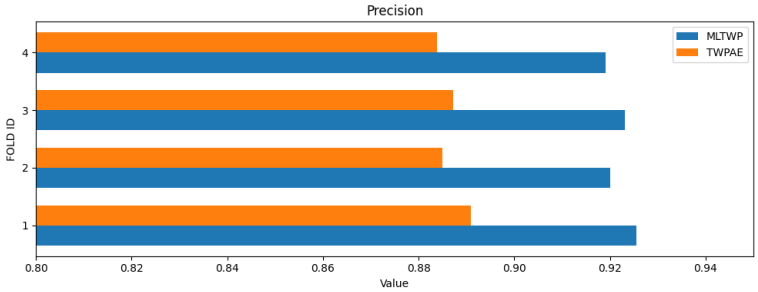

Fig. 4.1: Precision observed from four fold cross validation performed on MLTWP and TWPAE

91%, 92%, and 91% for four-fold cross validation; the corresponding figures for the TWPAE model were 88%, 87%, 87%, and 87%. In this case, it is assumed that the MLTWP approach outperforms the TWPAE model (see Figure 4.2).

In contrast to the actual number of delivered positive records for the testing, the test sensitivity measures the proportion of correctly identified positive records. Furthermore, sensitivity was observed from the MLTWP model over four-fold cross validation on different sets of records to be 91%, 90%, 90%, and 90% in that sequence, demonstrating optimal MLTWP performance when compared to the TWPAE technique (89%, 89%, 90%, and 89%) (see Figure 4.3).

The ratio of correctly identified positive and negative records relative to the total number of test records is an indicator of accuracy. The MLTWP method's accuracy was shown to be optimum through four-fold cross validation (91%, 90%, 90%, and 90%), outperforming the accuracy of the other method TWPAE (89%, 89%, 89%, and 88%, respectively) (see Figure 4.4).

In Figure 4.5, we see a graph depicting a 4-fold increase in the F-measure metric between the anticipated MLTWP model and the existing TWPAE model. According to the results, the F-measure for the MLTWP model is 92%, 92%, 92%, as well as 92% after being subjected to four fold of cross-validation. In this case, the F-measures for the TWPAE model are as follows: 88%, 89%, 89%, and 88%. As a result, the F-measure for

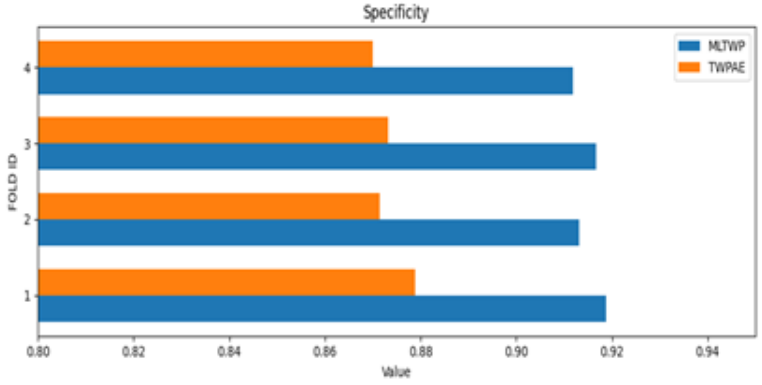

Fig. 4.2: Specificity observed from four fold cross validation performed on MLTWP and TWPAE

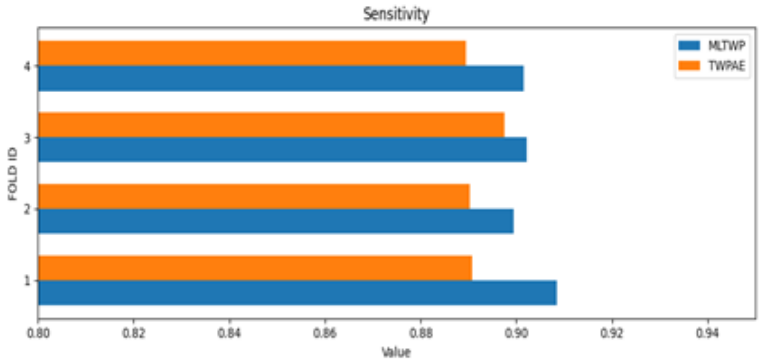

Fig. 4.3: Sensitivity observed from four fold cross validation performed on MLTWP and TWPAE

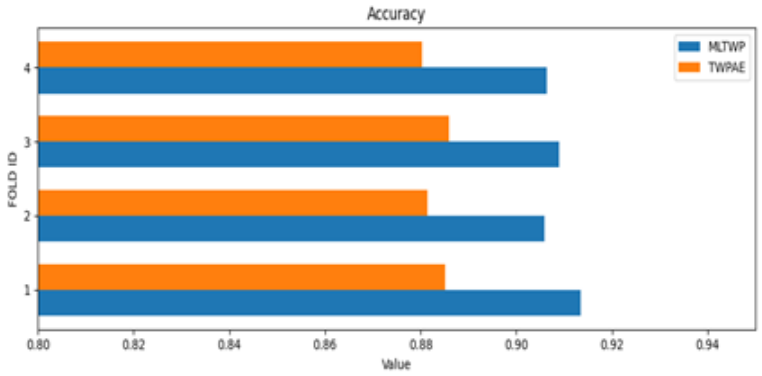

Fig. 4.4: Accuracy observed from four fold cross validation performed on MLTWP and TWPAE

the predicted MLTWP model is more significant than the conventional TWPAE technique.

Matthew's correlation coefficient (MCC) graph produced from four-fold cross validation of projected MLTWP model and current TWPAE model is shown in Figure 4.6. Using this data, we can determine that the MCC for

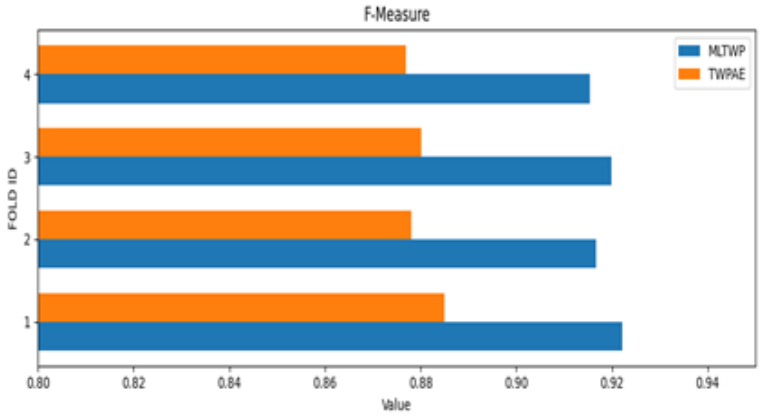

Fig. 4.5: F-measure observed from four fold cross validation performed on MLTWP and TWPAE methods

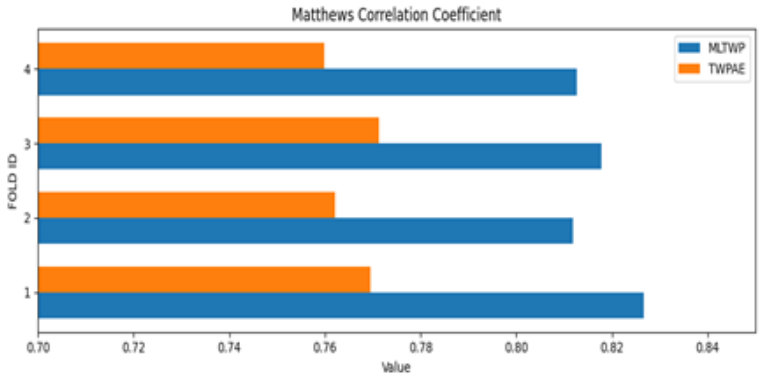

Fig. 4.6: Matthews's correlation coefficient observed from four fold cross validation performed on MLTWP and TWPAE methods

the MLTWP model is, in sequence, 83%, 81%, 82%, and 81%. For the TWPAE model, the MCC is as follows: 77%, 76%, 77%, and 76%. It foresees unambiguously that MCC for the proposed MLTWP model is much more significant than the current TWPAE technique.

5. Conclusion. The Machine Learning Tool Wear Prediction (MLTWP) model, employing temporal and spectral features of acoustic emission signals, has demonstrated a marked improvement in predicting tool wear. Quantitatively, MLTWP has shown enhanced precision, specificity, sensitivity, accuracy, and Matthews correlation coefficient compared to the Tool Wear Prediction using Acoustic Emission Signals (TWPAE) model. In four-fold cross-validation, the MLTWP model has consistently outperformed TWPAE, affirming its robustness and efficacy. The significant gains in signal detection accuracy underscore the effectiveness of the feature selection and machine learning techniques implemented in the MLTWP model. Future work will delve into the analysis of quasi-identifiers from the optimal attributes identified by MLTWP to further refine tool wear prediction.

Data Availability Statement. The proposed Machine Learning based Tool Wear Prediction (MLTWP) from Variability of Acoustic Sound Emission Signals have been evaluated through cross validation using the benchmark data of cutting Mill dataset available at: [36]. https://www.kaggle.com/datasets/shasun/tool-wear-

detection-in-cnc-mill

The results have been presented in the manuscript. The data used in this study is available publicly. However, the python code used to conduct the experiments is not publicly available. Nevertheless, the methodology and experimental setup used in this study have been described in detail, allowing for replication by other researchers using similar simulation tools or experimental setups.

Data and Materials Availability. The data and materials used in the proposed model, "Machine Learning based Tool Wear Prediction (MLTWP) from Variability of Acoustic Sound Emission Signals," have been described in detail in the manuscript. The study was conducted using the publicly available benchmark data, and the proposed model's performance was evaluated using 4-fold cross validation. The data and materials used in the proposed model are available in the manuscript.

REFERENCES

[1] V. A. Pechenin, *et al.*, "Method of Controlling Cutting Tool Wear Based on Signal Analysis of Acoustic Emission for Milling," *Procedia Engineering*, vol. 176, pp. 246–252, 2017.

[2] B. Li, "A Review of Tool Wear Estimation Using Theoretical Analysis and Numerical Simulation Technologies," *International Journal of Refractory Metals and Hard Materials*, vol. 35, pp. 143–151, Nov. 2012.

[3] P. J. Arrazola, *et al.*, "Correlation between Tool Flank Wear, Force Signals and Surface Integrity When Turning Bars of Inconel 718 in Finishing Conditions," *International Journal of Machining and Machinability of Materials*, vol. 15, no. 1/2, p. 84, 2014.

[4] A. Siddhpura and R. Paurobally, "A Review of Flank Wear Prediction Methods for Tool Condition Monitoring in a Turning Process," *The International Journal of Advanced Manufacturing Technology*, vol. 65, no. 1–4, pp. 371–393, Mar. 2013.

[5] T. I. Liu, *et al.*, "On-Line Monitoring of Boring Tools for Control of Boring Operations," *Robotics and Computer-Integrated Manufacturing*, vol. 26, no. 3, pp. 230–239, June 2010.

[6] R. Pullin, *et al.*, "Confidence of Detection of Fracture Signals Using Acoustic Emission," *Applied Mechanics and Materials*, vol. 7–8, pp. 147–152, Aug. 2007.

[7] X. Li, "A Brief Review: Acoustic Emission Method for Tool Wear Monitoring during Turning," *International Journal of Machine Tools and Manufacture*, vol. 42, no. 2, pp. 157–165, Jan. 2002.

[8] C. Zuluaga-Giraldo, *et al.*, "Acoustic Emission during Run-up and Run-down of a Power Generation Turbine," *Tribology International*, vol. 37, no. 5, pp. 415–422, May 2004.

[9] Y. Zhang, *et al.*, "Tool Wear Condition Monitoring Method Based on Deep Learning with Force Signals," *Sensors*, vol. 23, no. 10, p. 4595, May 2023.

[10] Y. Chen, *et al.*, "Predicting Tool Wear with Multi-Sensor Data Using Deep Belief Networks," *The International Journal of Advanced Manufacturing Technology*, vol. 99, no. 5–8, pp. 1917–1926, Nov. 2018.

[11] E. J. Weller, *et al.*, "What Sound Can Be Expected From a Worn Tool?," *Journal of Engineering for Industry*, vol. 91, no. 3, pp. 525–534, Aug. 1969.

[12] M. A. Mannan, *et al.*, "Application of Image and Sound Analysis Techniques to Monitor the Condition of Cutting Tools," *Pattern Recognition Letters*, vol. 21, no. 11, pp. 969–979, Oct. 2000.

[13] T. Delio, *et al.*, "Use of Audio Signals for Chatter Detection and Control," *Journal of Engineering for Industry*, vol. 114, no. 2, pp. 146–157, May 1992.

[14] D. R. Salgado and F. J. Alonso, "An Approach Based on Current and Sound Signals for In-Process Tool Wear Monitoring," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 14, pp. 2140–2152, Nov. 2007.

[15] C. Aliustaoglu, *et al.*, "Tool Wear Condition Monitoring Using a Sensor Fusion Model Based on Fuzzy Inference System," *Mechanical Systems and Signal Processing*, vol. 23, no. 2, pp. 539–546, Feb. 2009.

[16] I. Ubhayaratne, *et al.*, "Audio Signal Analysis for Tool Wear Monitoring in Sheet Metal Stamping," *Mechanical Systems and Signal Processing*, vol. 85, pp. 809–826, Feb. 2017.

[17] N. Seemuang, *et al.*, "Using Spindle Noise to Monitor Tool Wear in a Turning Process," *The International Journal of Advanced Manufacturing Technology*, vol. 86, no. 9–12, pp. 2781–2790, Oct. 2016.

[18] A. Kothuru, *et al.*, "Application of Audible Sound Signals for Tool Wear Monitoring Using Machine Learning Techniques in End Milling," *The International Journal of Advanced Manufacturing Technology*, vol. 95, no. 9–12, pp. 3797–3808, Apr. 2018.

[19] S. Orhan, *et al.*, "Tool Wear Evaluation by Vibration Analysis during End Milling of AISI D3 Cold Work Tool Steel with 35 HRC Hardness," *NDT & E International*, vol. 40, no. 2, pp. 121–126, Mar. 2007.

[20] M.-C. Lu and E. Kannatey-Asibu, "Analysis of Sound Signal Generation Due to Flank Wear in Turning," *Manufacturing Engineering*, American Society of Mechanical Engineers, pp. 165–175, 2000.

[21] F. J. Alonso and D. R. Salgado, "Application of Singular Spectrum Analysis to Tool Wear Detection Using Sound Signals," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 219, no. 9, pp. 703–710, Sept. 2005.

[22] Z. K. Peng, *et al.*, "An Improved Hilbert–Huang Transform and Its Application in Vibration Signal Analysis," *Journal of Sound and Vibration*, vol. 286, no. 1–2, pp. 187–205, Aug. 2005.

[23] N. E. Huang, *et al.*, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time

Series Analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
[24] A. G. Rehorn, J. Jiang, and P. E. Orban, "State-of-the-art methods and results in tool condition monitoring: a review," *International Journal of Advanced Manufacturing Technology*, vol. 26, pp. 693–710, 2005.
[25] K. Prakash and A. Samraj, "Tool Flank Wears Estimation by Simplified SVD on Emitted Sound Signals," *2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*, IEEE, 2017, pp. 1–5.
[26] K. Prakash and A. Samraj, "Tool Wear Condition Monitoring using Acoustic Analysis of Emitted Sound Signals by Peak to Peak Analysis," *ISERD 93rd International Conference*, Hanoi, Vietnam, 2017, pp. 8–13.
[27] J. L. Ferrando Chacón, *et al.*, "A Novel Machine Learning-Based Methodology for Tool Wear Prediction Using Acoustic Emission Signals," *Sensors*, vol. 21, no. 17, Sept. 2021, p. 5984.
[28] C. Yang, *et al.*, "Tool wear prediction model based on wear influence factor," 2023.
[29] C. L. Perumal, S. B. Bhadrinathan, and A. Samraj, "Tool Wear Condition Monitoring Using Emitted Sound Signals By Simple Machine Learning Technique," *DESIGN, CONSTRUCTION, MAINTENANCE*, vol. 2, 2022, pp. 168–172.
[30] C. Qian, *et al.*, "Tool Wear Prediction Based on BiLSTMA Networks," *Proceedings of the 14th International Conference on Computer Modeling and Simulation*, 2022, pp. 103–108.
[31] J. He, *et al.*, "Deep multi-task network based on sparse feature learning for tool wear prediction," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2022.
[32] P. Twardowski, *et al.*, "Identification of tool wear using acoustic emission signal and machine learning methods," *Precision Engineering*, vol. 72, 2021, pp. 738–744.
[33] U. Zolzer, *Digital Audio Signal Processing*, 2nd ed, Wiley, 2008.
[34] A. Ghasemi and S. Zahediasl, "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians," *International Journal of Endocrinology and Metabolism*, vol. 10, no. 2, pp. 486–489, Dec. 2012.
[35] T.-K. An and M.-H. Kim, "A new diverse AdaBoost classifier," *2010 International Conference on Artificial Intelligence and Computational Intelligence*, Vol. 1, IEEE, 2010, pp. 359–363.
[36] "Tool Wear Detection in CNC Mill," available at Kaggle, 2023.

Edited by: Anil Kumar Budati
Special issue on: Soft Computing and Artificial Intelligence for wire/wireless Human-Machine Interface
Received: Jan 8, 2024
Accepted: Apr 8, 2024