Estimating Remaining Useful Life of Cutting Tools

ENME691

Team Members:

Arka Bera, Het Mevada, Dongyang Zhen, John Pope

12/08/2023

Abstract

The cutting tool of a CNC machine is subject to degradation over time, eventually leading to failure and affecting the productivity of industry. The purpose of our project was to reduce downtime and maximize production by predicting the remaining useful life “RUL” of a CNC machining cutting tool. In this study, we have done prediction of RUL of tool by applying machine learning models. We have implemented feature extraction and feature engineering on the dataset collected from different sensors, and also carried out principal component analysis. The SVR model with PCA had the best prediction values in terms of root mean square error (0.074) and mean absolute error (0.058) compared to linear regression and neural network. Our study also extended to model the dataset without force components as features, and had the same RMSE and MAE as before with neural network as the best model. With this non-linear model such as neural network or SVR we can predict the RUL with very high accuracy and avoid productivity loss.

Introduction:

CNC machines or "computer numerical controlled" machines are controlled by computers and have transformed the manufacturing processes.CNC machines introduced a level of precision, efficiency, consistency and automation that was previously unattainable with manual processes. The cutting tool is a critical component of a CNC machining tool, and is responsible for machining accuracy and precision. Cutting tool degradation can be a limiting factor in performance and should be properly accounted for. Machining failure results in downtime to repair or replace parts. downtime in a manufacturing plant is multifaceted, ranging from lost production and delays to increased costs, reduced inventory, and compromised quality. Such disruptions also exert strain on the supply chain, highlighting the importance of proactive maintenance practices in order to maximize manufacturing operations.

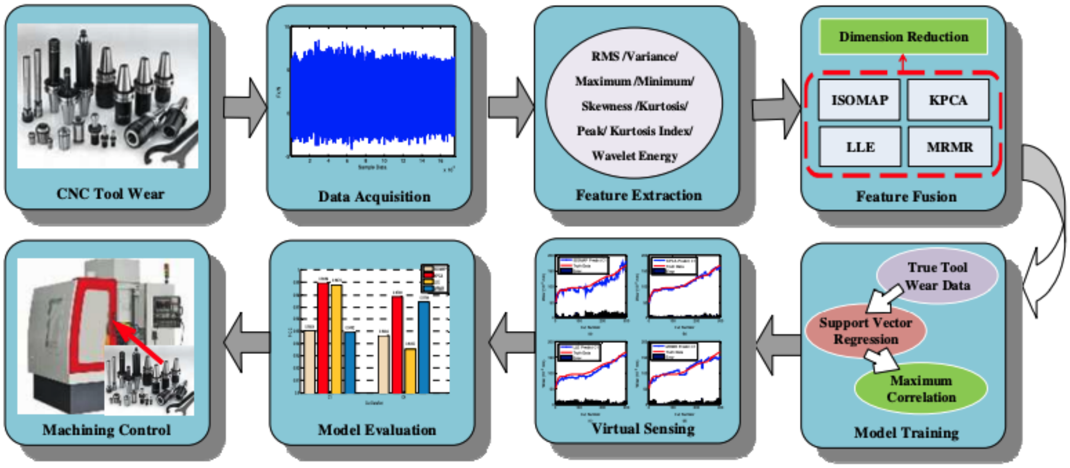
The integration of predictive maintenance strategies such as remaining useful life (RUL) have emerged as an effective technique for condition-based maintenance (predicting failure), minimizing downtime, ensuring product quality, and averting potential accidents. A model that can predict failure would enable projected and actionable maintenance strategies, potentially maximizing long term machine performance and reduced downtime.

Previous model evaluation techniques have used a wear sensing model that was evaluated using a leave-one-out cross-validation approach, which involves using one dataset for testing and the remaining datasets for training, repeated across different datasets. Time-Frequency Domain (TFD) features with different deep learning models [1] and feature network dictionaries to focus on enlarging features with limited sensors [2].

Various dimension reduction methods, such as Kernel Principal Component Analysis (KPCA), Locally Linear Embedding (LLE), Isometric Feature Mapping (ISOMAP), and Minimum Redundancy Maximum Relevance (mRMR), were investigated for multisensory feature selection and fusion. Dimension Reduction Techniques have utilized degradation models for the wear process with Expectation-Minimization for online RUL [3]. Support vector regression (SVR) has been used to reflect the relationship between monitoring data and tool life for online RUL [4] and Capsule BiLSTM has been used to predict RUL even when tool wear mechanism is not understood with prediction accuracy up to 94% [5]

- Data Collection and Processing: Around 300 data files were collected during the tool life test, which included measurements of force and vibration under different tool wear conditions. These measurements indicated that amplitudes of vibration and force increased with progressive tool wear.
- Validation of Virtual Sensing Model: The effectiveness of the virtual tool wear sensing technique was validated experimentally with machining tool run-to-failure tests on a CNC milling machine, demonstrating that virtual sensing estimates of tool wear width were comparable to those measured offline with a microscope.

Figure 1: Model Comparison



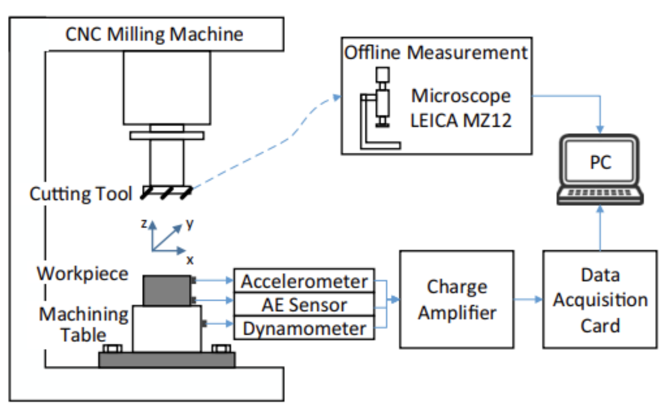
Problem Statement: Industrial productivity is dependent on the reliability and accuracy of the machine cutting tool and machine downtime results in lost production. The **purpose** of our project was to reduce downtime and maximize production by predicting the remaining useful life "RUL" of a CNC machining cutting tool. Computers are already used to monitor performance, we aim to predict future performance and system/component longevity using a RUL prediction model. Our Primary Aim was to build a RUL model that can predict failure in order to maximize long term machine performance. A more **specific Aim** was to identify the appropriate machine learning algorithm to predict remaining useful life of a machining cutting tool based on mean absolute error (MAE) or mean squared error (MSE) between the predicted model values and the target data set values provided.

Methods:

To accomplish this task a set of experimental data measured from a high speed CNC machine under dry milling operations was used. The milling machine consisted of a three-flute ball nose tungsten carbide cutter (stainless steel, HRC52) and was tested in a down milling operation.

For the experimental setup a sensors sweet (Figure 1) was applied to the CNC milling machine to collect operational function: i) Three Kistler piezo accelerometers were mounted on the workpiece to measure the machine tool vibrations of the cutting process in X, Y and Z directions, respectively ii) A Kistler acoustic emission sensor was mounted on the workpiece to monitor the high frequency stress wave generated by the cutting process iii) A Kistler quartz 3-component platform dynamometer was mounted between the workpiece and

Figure 2



machining table to measure the cutting forces.

The operation parameters for sensor data collection were set as follows: i) the running speed of the spindle (10,400 rpm) with the feed rate set at 1555 mm/min in the x direction; ii) the depth of cut (radial) was set to 0.125 mm in y direction; iii) the depth of cut (axial) was 0.2 mm in z direction.

For the data acquisition workflow, A NI DAQ PCI 1200 board was used to capture the voltage signals after the charge amplifiers. A LEICA MZ 12 microscope was used to measure the flank wear of each individual flute after finishing each surface. Finally, seven channels of signals (force_x, force_y, force_z, vibration_x, vibration_y, vibration_z, AE-rms) were captured, and the flank wear was set to be the target value.

Feature Extraction / Algorithms / Models: We implement multiple types of machine learning prediction models to assess reliability and accuracy. Feature extraction in the time and frequency domain was used to identify distinctive characteristics or patterns important for solving the prediction problem (figure 2). Three health assessment algorithms were chosen to perform diagnostic analysis and two Prediction Model Algorithms were chosen as the RUL solver (table 1).

Table 1. Health assessment and model prediction algorithms

Health Assessment Algorithm	
Linear Regression	used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data.
Support Vector Regression	identifies hyperplane that best fits the data while minimizing the margin of error
Multi-layer neural network	computational model of interconnected layers of artificial neurons designed to process and learn complex patterns
Prediction Model Algorithms	
Autoregressive moving average model (ARMA)	time series forecasting model that combines autoregressive (AR) and moving average (MA) components to capture and predict patterns in sequential data.
Exponential curve fitting :	best-fitting exponential function to a set of data points, allowing for the modeling and prediction of exponential growth or decay patterns

Principal Component Analysis (PCA): Principal Component Analysis (PCA) is a statistical technique for data analysis for interpretation of complex datasets. PCA aims to reduce the dimensionality of a dataset by transforming the original variables into a set of uncorrelated variables known as principal components. Once data is standardized to ensure consistent scaling, PCA then calculates the covariance matrix, to identify relationships among different variables. PCA identifies principal components that capture the maximum variance in the data, by computing the eigenvalues and corresponding eigenvectors from the covariance matrix. The selection of the principal components is based on eigenvalues, with the first principal component representing the most variance. PCA simplifies data representation by projecting the original data onto a reduced-dimensional space while retaining crucial information. PCA's applications range from dimensionality reduction and noise reduction to aiding in data visualization and feature extraction across diverse fields such as machine learning, image processing, and statistics.

Machine Learning and Statistical Models:

In our project, we utilize a diverse array of models: Linear Regression, Support Vector Regression, Neural Networks, ARIMA, and Exponential Curve Fitting, each chosen for their unique ability to capture specific types of data relationships and patterns. Linear Regression is used for its straightforward handling of linear relationships, while Support Vector Regression is ideal for capturing more complex relationships in higher-dimensional spaces. Neural Networks are included for their strong capacity to model nonlinear patterns. ARIMA is particularly adept at analyzing time series data, capturing underlying trends. Lastly, Exponential Curve Fitting is employed to effectively model scenarios where data exhibits exponential growth or decay. This selection of models allows us to comprehensively address varied data characteristics, providing a robust framework for our analysis.

Linear Regression: Linear Regression is a fundamental statistical and machine learning technique used to model the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship between the input variables (X) and the single output variable (Y), where Y can be calculated from a linear combination of the input variables. Linear regression is simple yet powerful, and it's used in various fields for predictive modeling and trend analysis. It's particularly useful when the data shows a linear trend and you need a straightforward solution for forecasting or understanding relationships between variables.

Support Vector Machine: Support Vector Machine is a versatile machine learning algorithm primarily used for classification tasks but adaptable for regression. It works by finding the hyperplane that best separates the classes in the feature space. The SVM algorithm attempts to maximize the margin between the data points of different classes, which is defined by the support vectors, or the data points closest to the hyperplane. SVMs are highly effective in high-dimensional spaces and are versatile as they can be used with various kernel functions to adapt to different types of data.

Neural Network: Neural Networks are a cornerstone of modern machine learning, inspired by the structure and function of the human brain. They consist of layers of interconnected nodes, or neurons, which process information using their internal state and a set of learned weights. Neural Networks are highly flexible and can learn complex patterns from data, making them suitable for a wide range of tasks like image recognition, natural language processing, and time series forecasting. Their ability to learn from raw data and improve with experience makes them powerful tools for predictive modeling.

AutoRegressive Integrated Moving Average (ARIMA): ARIMA is a popular statistical method for time series forecasting. It combines autoregressive (AR) and moving average (MA) models and integrates differencing to make the time series stationary, thus addressing trends and seasonality. ARIMA models are characterized by three parameters: p (autoregressive), d (differencing), and q (moving average). They are widely used in finance, economics, and business for forecasting future trends based on past data. ARIMA is best suited for time series data that show evidence of trends or seasonal patterns.

Exponential Curve fitting: Exponential curve fitting involves modeling data using an exponential function. It's particularly useful when the rate of change of a dataset increases or decreases at a constant rate. This method is common in fields like biology, chemistry, and physics, where processes often exhibit exponential growth or decay (like population growth, radioactive decay). In exponential curve fitting, the goal is to find the exponential function that closely approximates the data points. This method provides a simple yet powerful way to forecast and understand the behavior of systems that follow an exponential trend.

Table 2: Model Comparison

Feature	Linear Regression	Support Vector Regression (SVR)	Neural Network	ARIMA
Type of Model	Statistical regression model	Machine learning regression model	A computational model that mimics the human brain	Time series forecasting model
Main Use	Predicting a dependent variable from independent variables	Predicting a dependent variable using a non-linear approach	classification, regression, and diverse domain	Forecasting future values in a time series
Algorithm Approach	Assumes a linear relationship between input and output	Uses kernel functions to transform data and find an optimal boundary	Uses interconnected layers of nodes or neurons connect all the layers	Uses lags of the variable and the error terms for modeling
Data Requirements	Assumes linear relationship, independence, and homoscedasticity of residuals	Works well with both linear and non-linear data, requires scaling of features	Requires large datasets for effective learning and generalization; benefits from scaled and normalized data	Requires stationary time series data

Interpretability	High (coefficients can be easily interpreted)	Low (complex transformations and boundary)	Low (black Box)	Moderate (understanding of time series components needed)

Data Description:

The dataset being analyzed is the result of a thorough experimental investigation carried out on a high-speed CNC (Computer Numerical Control) milling machine. The investigation focused on the performance of a three-flute ball nose tungsten carbide cutter designed for stainless steel with an HRC52 hardness rating. The feed rate was fixed at 1555 mm/min along the x-direction, and the spindle speed was adjusted to 10,400 rpm. The axial depth of cut in the z-direction was kept constant at 0.2 mm, whereas the radial depth of cut in the y-direction was 0.125 mm.

A complex instrumentation setup was used to fully capture the dynamics of the milling operation. To measure cutting forces, a Kistler quartz 3-component platform dynamometer was mounted between the workpiece and the machining table. The workpiece was equipped with three Kistler piezo accelerometers to track vibrations of the machine tool in the X, Y, and Z directions. A Kistler acoustic emission sensor was installed on the workpiece to provide a more detailed understanding of high-frequency stress waves during cutting. A National Instruments (NI) DAQ (Data Acquisition) PCI 1200 board was used to ensure high-quality data acquisition. This board made it easier to record voltage signals after they went through charge amplifiers, which guaranteed that the different stresses and vibrations that occurred during the milling process would be accurately recorded.

In order to evaluate flank wear precisely on each flute of the tungsten carbide cutter, a LEICA MZ 12 microscope was used. When each milling operation was finished, this microscopic evaluation acted as an important value in determining tool wear and performance deterioration. Seven signal channels have been incorporated into the dataset: force_x, force_y, and force_z; vibration_x, vibration_y, and vibration_z; and the root mean square of acoustic emission (AE-rms). Flank wear was selected as the target variable, providing information about tool wear and tear under the given machining conditions. The following figures are the box-plot and the distribution of the variables.

Figure 3. Box-plot

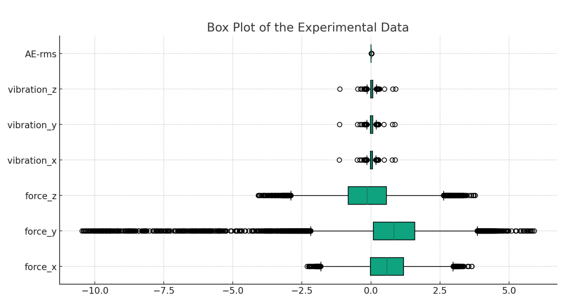
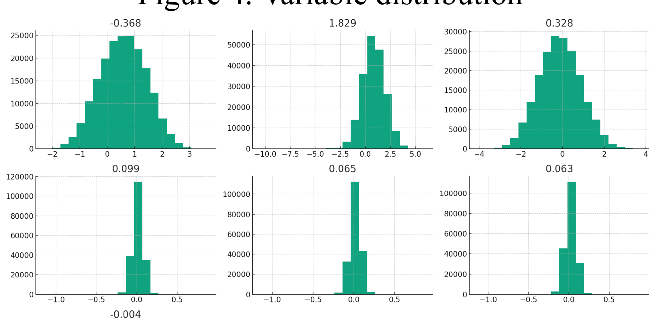
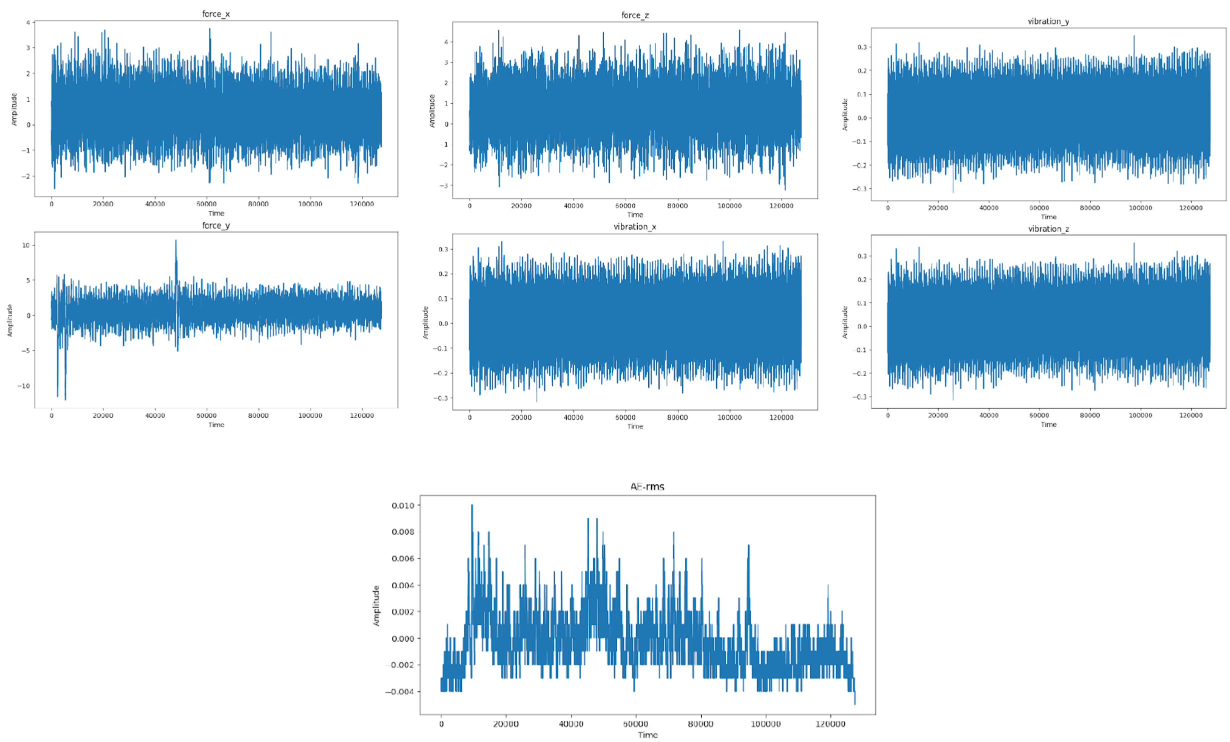


Figure 4. Variable distribution



Three machining conditions were provided for this project (although the PHM Challenge has more data available on their website: (https://phmsociety.org/phm_competition/2010-phm-society-conference-data-challenge/ [6]). The three cutting tools were labeled as C1, C4, and C6 data. For our project, we used C1 and C4 datasets to train the model, and C6 data was used to test them. These datasets are time-domain data, recorded for 315 time steps.

Figure 5. Data Visualization



All three datasets have wear data of three flutes. These wear values formed the ‘target’ for our supervised machine learning algorithms. For this project, we have taken the L2 norm of the three flute wear data to obtain one target value (other approaches have been used in the literature, such as ‘max’ value at each time step, or simply taking all three wear values as output). This was further normalized before feeding to the ML models.

The data was loaded using csv parsers and several pre-processing tasks were performed to obtain the feature space for the Machine Learning models implemented in this project. The large size of the dataset put a lot of load on the RAM and on the system, so the models were initially built and tested for the first 50 data points of each dataset, and then scaled up using more powerful resources to analyze the full dataset.

The time domain data was analyzed and the following features were extracted: i) RMS, ii) Variance, iii) Maximum, iv) Skewness, v) Kurtosis, vi) Peak-to-peak. The time domain data was then converted to the frequency domain to extract features: i) Spectral skewness ii) Spectral kurtosis. In addition, wavelet energy was calculated from the time-frequency domain of the dataset. All features obtained were scaled appropriately. Thus, we got 9 metrics for each of the 6 channels of the input dataset (force_x, force_y, and force_z; vibration_x, vibration_y, and vibration_z), resulting in a feature space of 54 times 315 time steps. We also performed Principal Component Analysis (PCA) to reduce the dimensionality of the feature space.

NOTE: Dynamometers are expensive equipment and not always available in academic laboratories. To see if we can get accurate results even in the absence of Force data, we performed additional testing using only Vibration components of the dataset.

Domain	Features	Expression
Statistical	RMS	$z_{RMS} = \sqrt{\frac{1}{n}(z_1^2 + z_2^2 + \cdots + z_n^2)}$
	Variance	$z_{\text{var}} = {}^{1}\!/_{N} \sum_{N} (z_i - \bar{z})^2$
	Maximum	$z_{max} = \max(z)$
	Skewness	$z_{skew} = E\left[\left(\frac{z-\mu}{\sigma}\right)^3\right]$
	Kurtosis	$z_{kurt} = \frac{1}{n}\sum_{n}\left(\frac{z_i-\mu}{\sigma}\right)^4$
	Peak-to-Peak	$z_{p\text{-}p} = \max(z)\text{-}\min(z)$
Frequency	Spectral skewness	$f_{skew} = \sum_{i=1}^{k}(\frac{f_i - \bar{f}}{\sigma})^3 S(f_i)$
	Spectral Kurtosis	$f_{kurt} = \sum_{i=1}^{k}(\frac{f_i - \bar{f}}{\sigma})^4 S(f_i)$
Time-Frequency	Wavelet energy	$E_{WT} = \sum_{i=1}^{N} wt_{\varphi}^2(i) / N$

Table 3: Calculations for Feature Extraction

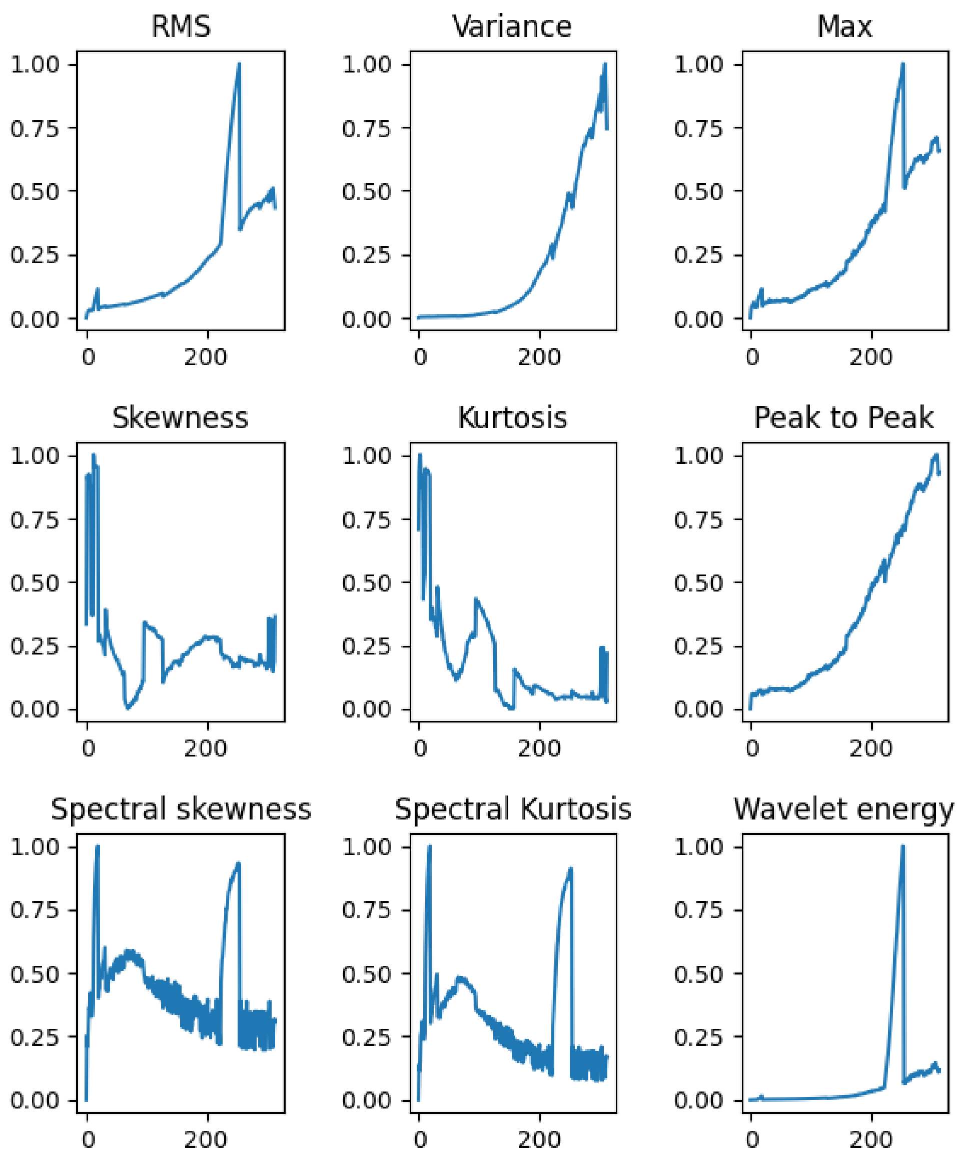


Figure 6. Feature Extraction

Results:

Firstly, we started with feature extraction and engineering to make better feature space to comprehensively train the model. The different space and frequency domain based features are presented in Figure 6 for c1 force x-component.

Principal component analysis (PCA) :

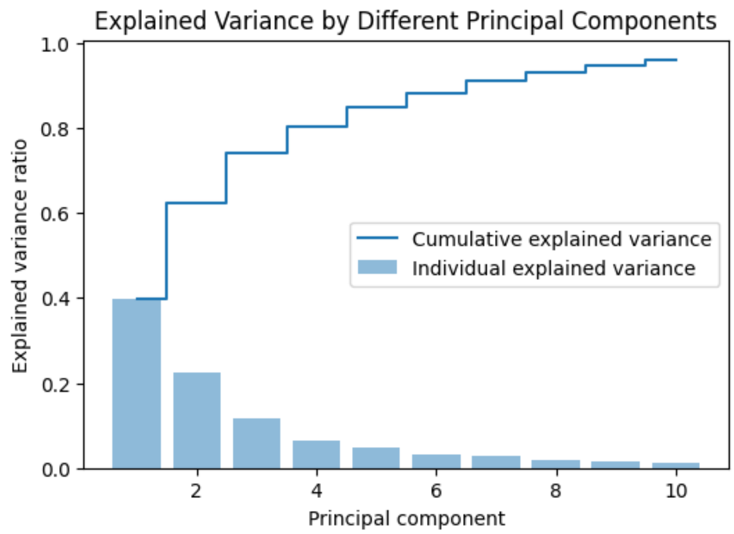
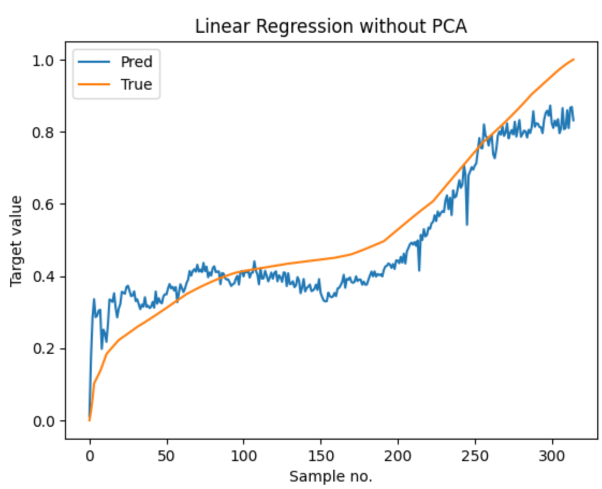
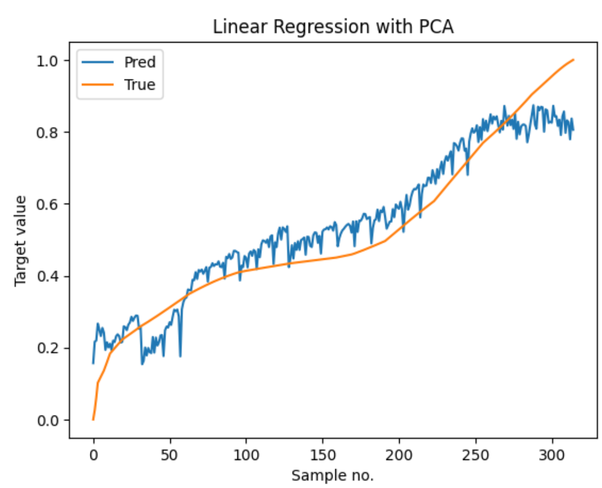


Figure 7. PCA visualization

Logistic Regression: We have tried implementing different models to better predict the actual condition of the tool's remaining useful life. In our first approach, we tried implementing the linear model since it is a very basic model to realize the actual trend in target value. In feature engineering, we also carried out PCA to reduce the number of features in model training without losing the essence of the data. Through PCA we found out that out of the total 54 features, only 9 features are required to explain 95% variance in the dataset. The important features in respective components through PCA were evaluated and they are tabulated in Table 3. Hence, the linear regression model was implemented for the complete feature dataset and another with only 9 components, and the output of the model is compared in Figure 4. From the figure the fit of mode with PCA is definitely better than model without PCA.



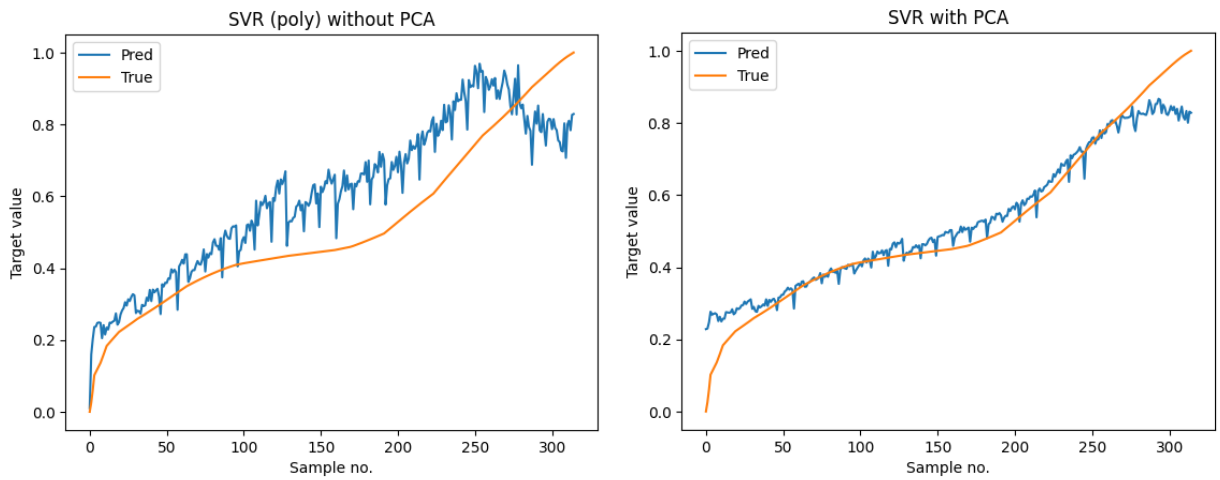
(a)



(b)

Figure 8: Predicted value vs true target value for different samples predicted using Linear Regression (a) with and (b) without PCA

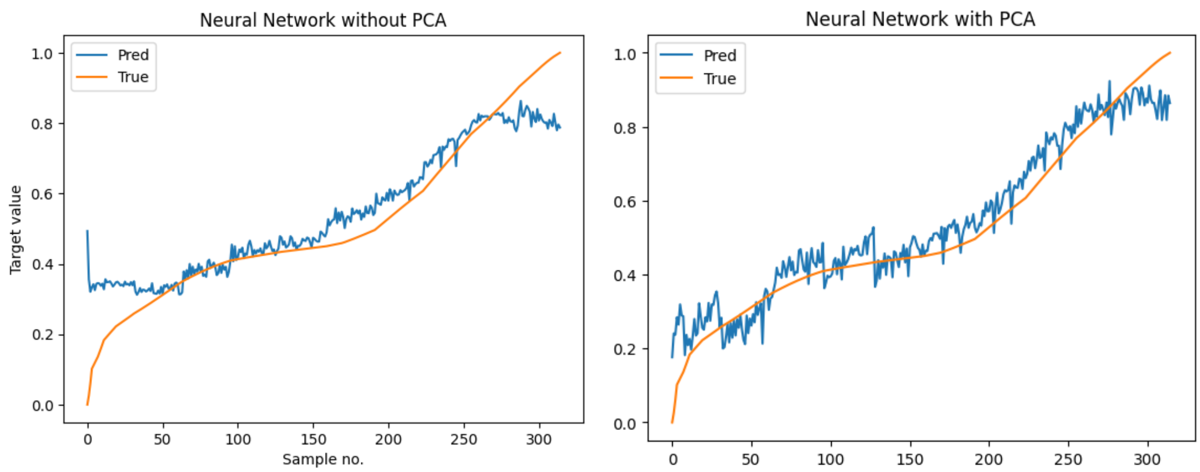
Support Vector Regression (SVR): The linear model fit was not the best we obtained to predict the behavior of our dataset, and so we tried with few non-linear models. Support vector regression is a good start to implement different kernels to fit out the dataset. Our approach also included tuning the hyperparameters for the SVR model. The hyperparameter tuning of SVR model with PCA showed that a model with rbf kernel and C and gamma value of 10 and 0.001 respectively were the best predictors. Hence, Figure 5 shows the prediction of SVR model with and without PCA, and has significantly better results than the linear model.



(a) (b)

Figure 9: Predicted value vs true target value for different samples predicted using Support Vector Regression (a) with and (b) without PCA

Neural Network: With a more comprehensive modeling approach we wanted to try neural networks to predict our target values. In our neural network model approach, we implemented relu and sigmoid activation functions in the hidden layers. Additionally, we also tried different epochs and batch size and determined the optimum value of epochs from the output loss function to minimize the RSME. Figure 6 shows that a neural network with PCA has better performance than without PCA.



(a) (b)

Figure 10: Predicted value vs true target value for different samples predicted using Neural Network (a) with and (b) without PCA

Table 3: Performance of different models in fitting the target values

Model	Without PCA		With PCA	
	RMSE	MAE	RMSE	MAE
Linear Regression	0.079	0.066	0.073	0.062
SVR	0.136	0.121	0.057	0.040
Neural network	0.083	0.058	0.061	0.049

ARMA/ARIMA and Exponential Curve Fitting: For these statistical models, the best prediction values were observed when exogenous regressors were not considered for training. A stepwise prediction algorithm was implemented for the ARMA/ARIMA model. 66% of the dataset was used initially to obtain the next time step, which was then added to the training set to obtain further time steps. This ensures that the model learns any changes in trend while training itself.

Both ARIMA and ARMA models were implemented and compared. ARIMA performed better than ARMA, with RMSE values at 0.079 and 0.109 respectively, and MAE values at 0.023 and 0.082 respectively. A grid search algorithm was implemented to obtain the best values of p, q, and d.

For exponential curve fitting, the function $y = Ae^{x}$ was fitted to 66% of the dataset, and then predictions were obtained from the weights obtained. This shows a significant deviation from the testing dataset, as the trends are not learned once training is stopped (RMSE:36.705, MAE: 31.320).

For statistical analysis, ARMA/ARIMA models are recommended over curve fitting algorithms.

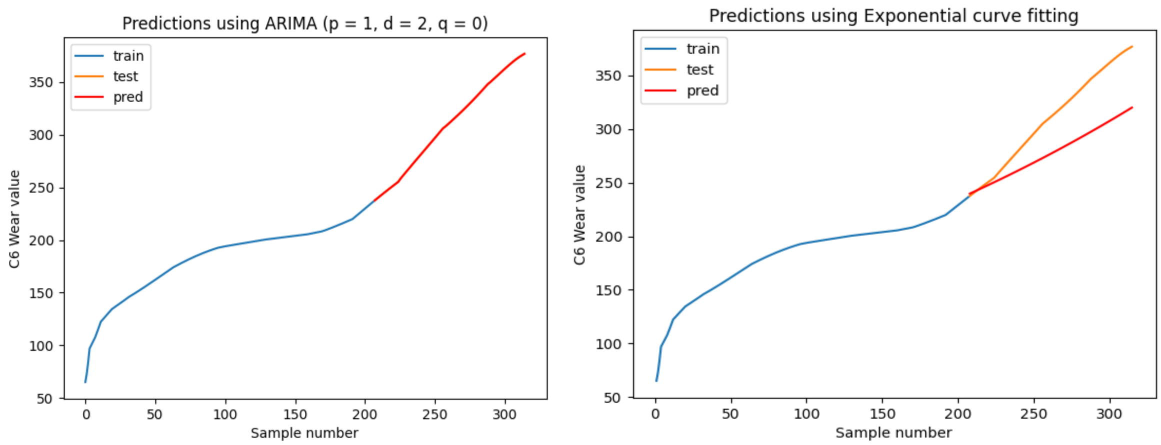


Figure 11. ARIMA and Exponential Curve Fitting

Comparison with PHM10 Data Challenge: When comparing the results with the PHM10 Data Challenge scoring method, their absolute error goes up from 3 to 70 along with sample number, our results are more stable when the sample number increases. Thus, the difference in our approach compared to the PHM10 Data Challenge are significant. Normalizing the data by column changes the scale and distribution of your variables, potentially affecting the models' performances. Furthermore, selecting the max value among the three flutes, instead of treating them as separate targets, alters the prediction task. In the PHM10 challenge, the focus is on estimating the maximum safe cuts for each individual flute, whereas our approach aggregates this information, which could lead to different error metrics.

Accuracy Improvement by Parameter Tuning: Parameter tuning for a SVM is a crucial step to improve the performance. It involves finding the optimal values for hyperparameters, which include regulation parameter "c", the kernel coefficient "gamma", and the type of kernel function (rbf or poly). The process uses grid search, along with cross-validation to prevent overfitting and ensure robustness. After trying different combinations of hyperparameters, the set that yields the best cross-validated performance is chosen. The optimal results are C: 10, indicating a moderate regularization; gamma: 0.0001, suggesting a large similarity radius; and an 'rbf' kernel, suitable for non-linear problems. The chosen parameters are then used to train the final SVM model. The following graph is the flowchart of grid search and the SVR performance after parameter tuning.

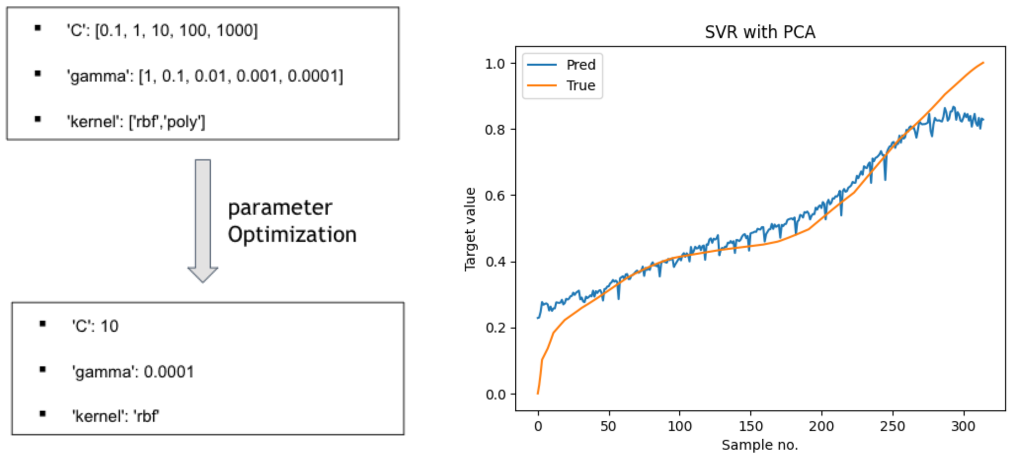


Figure 12. Hyperparameter Tuning

Discussion:

The purpose of our project was to reduce downtime and maximize production by predicting the remaining useful life “RUL” of a CNC machining cutting tool. The dataset analyzed was collected during experimental investigation carried out on a high-speed CNC (Computer Numerical Control) milling machine. A sensor suite (Figure 1) was applied to the CNC milling machine to collect operational function: i) accelerometers to measure machine tool vibrations ii) acoustic emission sensor to monitor the high frequency stress waves iii) a platform dynamometer to measure the cutting forces.

In practical scenarios, the cost of a dynamometer is really high and may not be a feasible sensor for many stakeholders. So, we explored excluding the force components from our feature space and to determine the accuracy of our model’s prediction. In our approach, we implemented PCA and hyperparameter tuning for SVR and neural networks. The PCA finding suggested that out of a total 27 feature space, we only require 6 features to determine 95% variance of the dataset. Hence, linear regression, SVR and neural network models were implemented with and without PCA to study which approach is significantly better. From Figure 13, we can see that the neural network has the best prediction accuracy compared to SVR and linear regression. The RMSE and MAE of a neural network is 0.074 and 0.058 respectively, compared with SVR and linear regression’s value of 0.101 and 0.092, and 0.095 and 0.087 respectively. Additionally, the model with PCA feature set was better than without PCA.

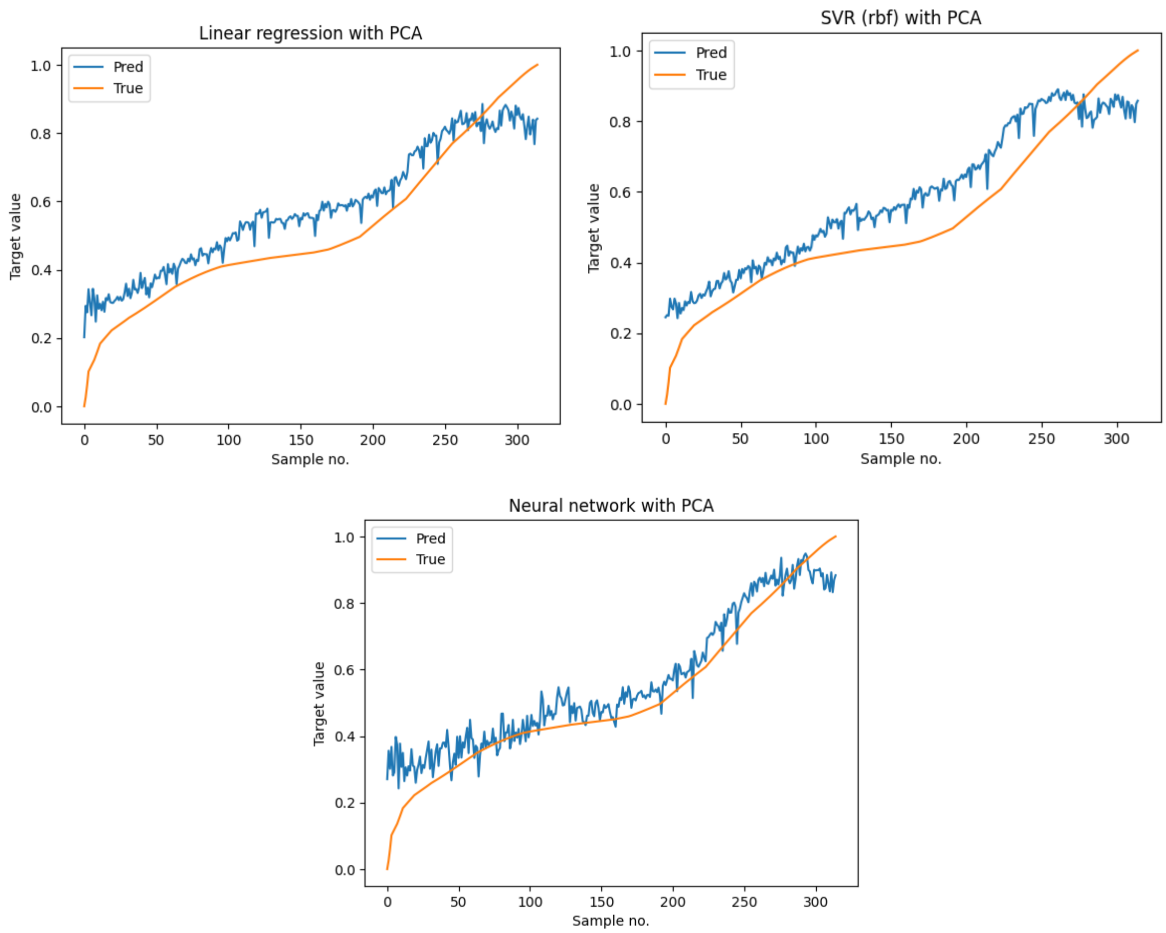


Figure 13: Prediction of different models with PCA

Conclusion:

In conclusion, our exploration into prediction of RUL in a machining cutting tool has yielded valuable insights. Recognizing the impracticality of a high-cost dynamometer we explored alternative approaches for predicting outcomes in the absence of expensive sensors and investigated the impact of excluding force components from our feature space. Furthermore, the inclusion of PCA-enhanced features demonstrated improved model performance compared to the model without PCA.

Our methodology involved employing PCA and fine-tuning hyperparameters for Support Vector Regression (SVR) and neural networks. The results of the PCA study revealed that, out of the initial 54 features, only 9 were necessary to capture 95% of the dataset's variance. Neural network model consistently outperformed SVR and linear regression in terms of prediction accuracy. Notably, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for the neural network were 0.074 and 0.058, respectively, surpassing SVR and linear regression.

This comprehensive analysis underscores the efficacy of our chosen approach for RUL prediction and provides a promising alternative for stakeholders facing cost constraints in utilizing expensive sensors for predictive modeling.Our approach can be used to direct the development and advancement of prediction models in industries where the cost of sensors/data acquisition is a limiting factor. The performance of neural networks, as highlighted in our results, may encourage industry professionals to explore and integrate these models into their predictive analytics workflows. Additionally, the study sets the stage for investigations into the broader application of PCA-enhanced features in various industry settings.

Future research directions could focus on further optimizing and refining the proposed methodologies and investigate the scalability of the approach to different contexts and domains. Furthermore, exploring additional feature engineering techniques and refining hyperparameter tuning, would contribute to the ongoing development of practical, cost-effective predictive modeling solutions. Ultimately, our research opens avenues for innovation, making accurate predictive modeling more accessible and cost-efficient.

References:

[1] Sayyad S, Kumar S, Bongale A, Kotecha K, Abraham A. Remaining Useful-Life Prediction of the Milling Cutting Tool Using Time-Frequency-Based Features and Deep Learning Models. Sensors (Basel). 2023 Jun 17;23(12):5659. doi: 10.3390/s23125659. PMID: 37420825; PMCID: PMC10301425.

[2] Qin X, Huang W, Wang X, Tang Z and Liu Z 2022 Real-Time Remaining Useful Life Prediction of Cutting Tools Using Sparse Augmented Lagrangian Analysis and Gaussian Process Regression Sensors 23 413 Online: http://dx.doi.org/10.3390/s23010413

[3] Huang Y, Lu Z, Dai W, Zhang W and Wang B 2021 Remaining Useful Life Prediction of Cutting Tools Using an Inverse Gaussian Process Model Applied Sciences 11 5011 Online: http://dx.doi.org/10.3390/app11115011

[4] Y C Liu et al 2019 IOP Conf. Ser.: Mater. Sci. Eng. 576 012021

[5] Sun L, Zhao C, Huang X, Ding P, Li Y. Cutting tool remaining useful life prediction based on robust empirical mode decomposition and Capsule-BiLSTM network. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2023;237(14):3308-3323. doi:10.1177/09544062221142197

[6] “PHM10 Data Challenge - Scoring.” PHM Society, phmsociety.org/phm10-data-challenge-scoring/.

<u>Contribution:</u>

Arka Bera: Leader, ML algorithm development, Data processing, Slides, Report
Het Mevada: PCA, Hyperparameter tuning, Slides, Report.
Dongyang Zhen: Results Visualization, Comparison with benchmark, Sides, Report
John Pope: Statistical Model Development, ML model comparison, Slides, Report