Machine learning framework for predictive maintenance in milling

Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnLine 52-13 (2019) 177–182

Machine Learning Framework for Predictive Maintenance in Milling

Emiliano Traini * **Giulia Bruno** * **Gianluca D'Antonio** * **Franco Lombardi** *

* *Politecnico di Torino, Torino, Italy (email: {emiliano.traini, giulia.bruno, gianluca.dantonio, franco.lombardi}@polito.it)*

Abstract: In the Industry 4.0 era, artificial intelligence is transforming the manufacturing industry. With the advent of Internet of Things (IoT) and machine learning methods, manufacturing systems are able to monitor physical processes and make smart decisions through real-time communication and cooperation with humans, machines, sensors, and so forth. Artificial intelligence enables manufacturers to reduce equipment downtime, spot production defects, improve the supply chain, and shorten design times by using machine learning technologies which learn from experiences. One of the last application of these technologies is the development of Predictive Maintenance systems. Predictive maintenance combines Industrial IoT technologies with machine learning to forecast the exact time in which manufacturing equipment will need maintenance, allowing problems to be solved and adaptive decisions to be made in a timely fashion. This study will discuss the implementation of a milling Cutting-tool Predictive Maintenance solution (including Wear Monitoring), applied to a real milling data set as validation of the framework. More generally, this work provides a basic framework for creating a tool to monitor the wear level, preventing the breakdown, of a generic manufacturing tool, in order to improve human-machine interaction and optimize the production process.

Keywords: Machine Learning - Industry 4.0 - Artificial Intelligence - Predictive Maintenance Condition Monitoring - Milling - Vibration Analysis

1. INTRODUCTION

Predictive Maintenance (PM) is a method to monitor the status of machinery in order to prevent expensive failures from occurring and to perform maintenance when it is really required. It has a long history. From visual inspection, which is the oldest method, PM has evolved to automated methods that use advanced signal processing techniques. Traditionally maintenance creates a trade-off situation in which one must choose between maximizing the useful life of a part at the risk of machine downtime (run-to-failure) and maximizing up-time through early replacement of potentially good parts (time-based PM), which has been demonstrated to be ineffective for most equipment components, considered as flawed and unreliable in recent years, in (Mobley, 2002). PM aims to break these trade offs by empowering companies to minimize maintenance by forecasting it ahead of time. Adoption of PM allows to maximize useful life of assets by reducing the frequency of maintenance activities, avoiding unplanned breakdowns and eliminating unnecessary preventive maintenance. This results in substantial time and cost savings and higher system reliability.

To implement a PM approach, a Condition Monitoring (CM) system is necessary. Using the words of Hassan et al. (2018), CM is "the process of monitoring one or more parameters of machine to predict its potential faults early". Examples of measures useful for this purpose are temperature and acoustic emission in (Ravindra et al., 1997), or vibration. A specific CM is the Tool Condition Monitoring (TCM) that investigates how these parameters affect tool wear. A TCM could prevent these problems, allow optimum utilization of the tool life and improve the efficiency of the machine. TCM has a foremost importance in metal cutting owning to its direct impact on the quality of the machined surface, its dimensional accuracy and, consequently, the economics of machining. In fact even if the tool is not broken yet, its degradation reduces the work surface quality and leads to a significant loss of dimensional accuracy. On the other hand, an excessive preventive replacement of tools involves higher costs and production time: it will require additional tools to be purchased, which are generally expensive, as well as considerable time to change the tool. The IoT enabled presence of abundant sensors that real time collect big data, composed of time domain features. The current technologies are so developed that the scientific community is no longer studying how to detect manufacturing data, but which method is the most economical one. So in this scenario PM is performed with Machine Learning (ML) methods that are much more accurate and can take into account all the factors provided by the sensors. The goal of this study is to create a ML model to predict when the cutting tool used in milling operations must be replaced in order to minimize the effects of disastrous failures and manage the planning of activities.

The rest of the paper is organized as follows. Section 2 describes the main factors that play a fundamental role

10.1016/j.ifacol.2019.11.172

in the topic. Section 3 describes the proposed framework, from data management to the improvement of ML algorithms. Section 4 presents the application of the framework in the use case of a milling data set and the relative results. Finally, Section 5 draws conclusions and states future work perspectives.

2. WEAR MONITORING FOR MILLING

Generally, the TCM is based on factors that can be grouped in two categories: a-priori and a-posteriori. The first ones are factors known at the beginning of the milling activity, such as machine parameters. Conversely, the second type of factors are the output of monitoring activities of the machine and in particular of the tool, through the use of sensors. For the ML tool the a-priori variables are all associated with the factor variable, as they identify a specific case due to the choice of certain conditions. The a-posteriori factors are time series analyzed by ML algorithms to capture the trend of the wear level over time.

2.1 A-priori factors: cutting parameters and materials

Milling consists in using a rotating tool with multiple cutting edges, called cutter, to remove material from the surface of a work-piece. To do this, it requires the definition of several cutting parameters, which influence job duration, quality and accuracy of the finished surface, the life of the tool and the cost of production. These parameters are chosen during the manufacturing activity and for example they could be the **cutting speed**, related to the speed of rotation of the tool, the **spindle speed**, i.e. the rotational frequency of the spindle, the **feed rate**, that refers to the velocity at which the cutting tool is advanced through the work-piece to remove material, and, as the last example, the **depth of cut**, that is the penetration of the cutting in the work-piece.

Another factor that influences the wear trend is the use of cutting tools of different materials. A cutting tool must present three important characteristics: the **hardness**, i.e. the ability of the material to maintain its strength at elevated temperatures, the **toughness**, i.e. the capacity of the material to absorb energy without failing, usually deriving from a combination of strength and ductility, and the **wear resistance**, i.e. the attainment of acceptable tool life before cutting tool need to be replaced. Hardness is the most important property needed to resist abrasive wear.

2.2 A-posteriori factors: cutting forces and emissions

The interaction of milling cutter and work-piece is a complex process and generates different physical effects. During the engagement of the tool in the work part, energy is released which results in cutting forces, vibration and acoustic emission. **Cutting forces** appear in shear zones where plastic deformation takes place. The predominant cutting action in milling involves deformation of the work-piece to form a chip in the primary shear zone. **Vibration emission** is a low frequency oscillation due to the acceleration of the object because of the dynamic changes of cutting forces resulting from periodical changes in tool geometry, chip formation, and built up edges. **Acoustic emission** is a high frequency oscillation which occurs spontaneously within metals when they are deformed or fractured. It is caused by the release of strain energy as the micro structure of the material is rearranged. Acoustic emission is generated in the shear zones, the primary as well as the secondary along the chip-tool interface through bulk deformation and sliding, respectively, and, lastly, at the tool flank work-piece interface due to friction, (R., 1999).

2.3 Wear measurement and life criteria for the tool

As milling proceeds, the various deterioration mechanisms result in increasing levels of wear on the cutting tool. Generally, wear of cutters in milling processes depends on tool material and geometry, work-piece materials, cutting parameters (cutting speed, feed rate and depth of cut), cutting fields and machine-tool characteristics. As explicated by Sunday Abolarin et al. (2015), speed of cutting is the most influential parameter for the rate of wear; depth of cut and feed rate also affect the tool life.

There are two basics zones of wear in cutting tools: flank wear and crater wear. Crater wear consists of a cavity in the rake face of the tool that forms and grows from the action of the chip sliding against the surface. This is somewhat normal for tool wear and does not seriously degrade the use of a tool until it becomes serious enough to cause a cutting edge failure. Conversely, flank wear occurs on the flank, or relief face, of the tool. It results from rubbing between the newly generated work surface and the flank face adjacent to the cutting edge. Flank wear is measured by the width of the wear band, VB.

In milling, operating the tool until final catastrophic failure is one way of defining tool life. However, in production, it is often a disadvantage to use the tool until this failure occurs because of difficulties in re-sharpening the tool and problems with work surface quality. As an alternative, a level of tool wear can be selected as a criterion of tool life, and the tool is replaced when wear reaches that level. A conventional tool life criterion is to define a threshold width of the flank wear, because of its influence on work-piece surface roughness and accuracy. A critical VB_{max} in milling is always recommended by the tool manufacturers guided by industrial applicability and ISO standards: Tool life testing in milling part 1, face milling, and part 2, end milling. The level depends on several parameters, such as the type of milling, the cutting speed and the material of the cutter. For instance, the standard ISO 8688-1:1989 defines a maximal flank wear $VB_{max} = 0.6mm$ as effective tool life for cemented carbides, high-speed steels (HSS) and ceramics tools applied in face milling operations.

3. FRAMEWORK

The proposed framework is a particular application of a most general one described for example in the book (P. N. Tan and Kumar, 2005). This approach can be used when ML algorithms are used to predict the wear level of production tool, based on data coming from (analogical) sensors. Like for all application, after the collection of data, the framework contains a first part of pre-processing, in

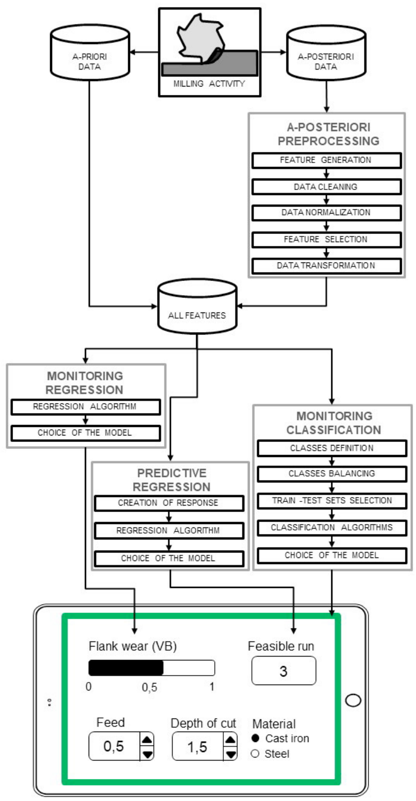

Fig. 1. Structure of the framework. This figure shows a possible application of the ML results.

which data are analyzed and manipulated to prepare them for the application of ML algorithms.

Fig. 1 is a scheme that resumes this framework. At the beginning there is the management of physical data: it is necessary transforming a single multi-dimensional observation in a set of mono-dimensional ones. A second general data management is done to optimize the training of algorithms. Further there are the feature selection and the preparation of data for classification and regression technique. The last part is about the training, valuation and choice of the model.

The first regression model and the classification one have the issue of determining if in the actual moment the cutting tool should be replaced. To be able to exactly specify the Remaining Useful Life, RUL (expected time period during which the tool is likely to operate before it requires repair or replacement), of the tool, another regression method must be implemented. The RUL is the remaining runs that can be performed with the same cutter before it wears. The criteria used for the wear is the critical flank wear VB_{max}: since the wear was measured after the relative run, it is assumed to last practicable run for each case is the first one corresponding to a VB equal or greater than VB_{max}. Therefore, the estimation of the remaining runs for each observation is the difference between this maximum number of feasible runs, different for each case, and the current run. In this way, only those cases where the maximum VB is greater than VB_{max} are considered and the RUL feature is created as previous explicated. The algorithm chosen are the same of the regression regarding VB but the response variable is RUL.

3.1 Feature generation

Feature generation is the process of transforming raw data into meaningful features more suitable for a machine data mining task, which act as input for ML algorithms and help in improving the overall predictive model performance. Generally, feature generation starts from an initial set of measured data and lead to derived values (features) intended to be explanatory and essential, simplifying the subsequent learning and modeling phases. According to the works of Zhang et al. (2016) and Caesarendra and Tjahjowidodo (2017), the time domain features to be considered are maximum, mean, root mean square, standard deviation, Skewness, Kurtois, peak-to-peak (maximum less minimum) and crest factor. A frequency-domain analysis is a representation of how much of the signal lies within each given frequency band over a range of frequencies. It is an important tool in signal processing and is also referred to as power spectrum analysis. The power spectrum describes the signal's power distribution over a range of frequencies. Spectral analysis considers the problem of determining the spectral content (i.e., the distribution of power over frequency) of a time series from a finite set of measurements. By looking at the spectrum, one can find how much power is contained in the frequency components of the signal. A signal can be converted between the time and frequency domains by using the Fourier transform, which converts a time function into a sum or integral of sine waves of different frequencies, each of which represents a frequency component. The features considered are maximum, sum, mean, standard deviation, Skewness, Kurtosis and peak-to-peak of band power spectrum.

3.2 Feature cleaning and normalization

Data cleaning is the process of identifying and correcting (or removing) incomplete, improper and inaccurate data. The aim is to address what are referred to as data quality issues, which negatively affect the quality of the model and compromise the analysis process and results. There are several types of data quality issues, including missing values, duplicate data, outliers, inconsistent or invalid data. Some techniques to clean data involve removing data records with missing values, merging duplicate records, generating a best, or at most reasonable, estimate for invalid values. In the case in which the response variable is missing, it is necessary to remove the entire row. Another important check is about the presence of outliers: it is necessary remove some rows considered outliers to improve the quality of the ML training.

Data normalization, also known as feature scaling or data standardization, is an important step in the data preprocessing phase. It is a technique used to standardize the range of features in the data set, which means to adjust values of numeric columns measured on different scales to a notionally common scale, without altering differences in the values' ranges or losing information. The goal of normalization is to improve the overall quality of a data set by re-scaling the dimension of the data and avoiding situations in which some values over-weighting others. A general choice for the normalization is a Min-Max normalization, according to the work of Al Shalabi et al. (2006).

3.3 Feature selection

The feature generation allows to transform data into syntetic information. However, some of these features could be either irrelevant or redundant and could negatively influence the performance of the trading activity. For this purpose, the feature selection is adopted by this protocol and it is done with the application of two cutting criteria: correlation coefficient and multicollinearity, as in the work of Senawi et al. (2017). The cutting activity is performed by evaluating a correlative matrix: all the features with a correlation factor higher than 0.9 are cut, as they are not adding information to the model. The issue of this approach is that it depends on the order of the features. To solve it, before applying the method, the features are ordered using as rank the correlation between each single feature and the response variable VB: in this way the more a feature is related to the response variable, the more likely it remains in the data set.

3.4 Data transformation

The resulting data set contains only numeric values, though some of them do not represent numbers but categories. So, it is important to specify that some variables, such as material, are categorical variables rather than numerical.

3.5 Regression

Regression analysis is a technique used to analyze a series of data consisting of a dependent variable, VB and RUL, and one or more independent variables, that are the founded features. The aim is to estimate a possible functional relationship existing between the dependent variable and the independent variables. The dependent variable in the regression equation is a real number with a particular domain.

As first step, the instances are randomly split into training and testing set. The proportion chosen are 70% for training and 30% for testing. All ML algorithms are trained and tested with the same subsets.

For the regression task, the following ML algorithms are trained, tested and evaluated:

- Linear Regression (LR);
- Decision Forest regression (DF);
- Bayesian Linear Regression (BLR);
- Boosted Decision Tree regression (BDT);
- Neural Network regression (NN);

where the first ones are used for their simplicity, while the last one, NN regression, because it is the most used acording to J. V. Abellan-Nebot (2010).

The evaluation metrics used for regression models are Mean Root Mean Squared Error (RMSE), Relative Squared Error (RelSE) and R^2. The error metrics measure the quality, i.e. predictive performance, of a regression model in terms of the mean deviation of its estimates from the real values. The lower the error values, the more the model is considered accurate in making predictions.

3.6 Classification

A way to predict when the cutting-tool must be replaced is to implement a binary classification where instances are classified according the value of flank wear: two classes are created using the critical value VB_{max} as partition between the classes *safe*, less than VB_{max}, and *worn*. This method is implemented to have a ML method focused only on alerting the operator who is operating in non-safety conditions (the margins of the tablet in Figure 1 began red). Another important step of the data manipulation is the control about the balancing of the classes: an imbalanced dataset can cause problems with how the model will classify instances since ML algorithms tend to produce unsatisfactory results faced when the class of primary interest is under-represented. The Synthetic Minority Oversampling Technique (SMOTE), presented in (Chawla et al., 2010) is used to treat imbalanced data: it increases the number of cases in the data set in a balanced way, by generating new instances of the minority class. As for regression, the data set is randomly divided into training and testing set, with the proportion of 70% and 30%: this choice, the simplest one, ensures the maintaining of class balancing even for training and testing sets. For the classification task, five ML algorithms are trained, tested and compared against each other. Since the classes of the label are two, binary classification algorithms are chosen to train the model. The choices fall for:

- Logistic Regression (LogR);
- Decision Forest (DF);
- Decision Jungle (DJ);
- Boosted Decision Tree (BDT);
- Neural Network (NN).

The parameter of evaluation for classification is based on the confusion matrix values, which are information about the actual and the predicted class. The metrics to evaluate the performance of the model are the percentage of correct safe classifications, the percentage of correct worn classifications and the accuracy (correct classifications on the total): this metrics are respectively indicated with WP, SP and ACC.

3.7 Models improvement

Model improvement has the objective of boosting the performance of ML algorithms. This is made by Hyperparameter Tuning method and K-Fold Cross Validation (KF-CV) technique. The first one, the Hyperparameter Tuning method, according to the work (Bergstra and Bengio, 2012), is used to find the optimal parameters value for each ML algorithm, both for the regression and for the classification. Usually, one does not know in advance which is the optimal configuration for a given model and therefore should evaluate all different combinations. When a ML algorithm is tuned for a given problem, the goal is to figure out the parameters resulting in the most accurate predictions.

Another technique to improve the models is the KF-CV. The main goal of any ML model is to learn from samples in such a way that the model can generalize the learning to unseen (real) data, that is to accurately perform in practice when making predictions on new data from a real problem.

When evaluating the performance of a model, it is always necessary to assess how well the learner will generalize to an unseen data set, this is called validation. KF-CV is an important technique often used in ML to assess both the variability of a data set and the reliability of any model trained using that data. In KF-CV, the entire data set is randomly partitioned into k groups of approximately equal size. One at a time, a single partition is treated as validation set for testing the model and the remaining $k-1$ sub-samples are put together and used as training data. For instance, considering 5 folds, the module would generate five models during CV, each model trained using 45 of the data, and tested on the remaining 15. Hence, the process is repeated k times and the evaluation results are computed for each validation. The k results can be averaged to obtain a single performance metric of the total effectiveness of the model. As a general rule and empirical evidence, 5 or 10 as value of k is generally preferred, as these values have been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance, but nothing's fixed and it can take any value.

4. USE CASE AND RESULTS

The chosen use case is the Milling data set collected by Agogino and Goebel (2007) in the Prognostic Center of Excellence (NASA – PCoE) which provides a collection of prognostic data sets from universities, agencies and companies representing experiments from runs on a milling machine under various operating conditions, such as different feeds, depth of cut and work-piece material. In particular, tool wear was investigated in a regular cut as well as entry cut and exit cut and the flank wear of milling insert was recorded.

The data set includes information about 16 cases with varying number of runs. The experiment has been carried out by employing the Matsuura machining center MC-510V, a CNC vertical milling machine equipped with a face milling cutter with six inserts. The cutting parameters for the different cases were guided by industrial applicability and recommended manufacturer's settings. Therefore, the cutting speed was set to $200mmin$. Two different depths of cut were chosen, $1.5mm$ and $0.75mm$. Also, two feeds were taken, $0.5mms$ and $0.25mms$ two types of material, cast iron and stainless steel J45 were used for the work-pieces, which size was $0.483m \times 0.178m \times 0.51m$. The milling inserts were of type KC710, a carbide with a PVD TiN coating over a general-purpose alloyed substrate. These choices equal 8 different settings. All experiments were done a second time with the same parameters with a second set of inserts, for a total of 16 different cases. The number of runs for each case was dependent on the degree of flank wear (VB) that was measured between runs at irregular intervals up to a wear limit (and sometimes beyond). Flank wear was not always measured and at times when no measurements were taken, no entry was made. The acquired data is organized in a 167×13 structure with fields divided as shown in table 1. It shows the values of first rows of the features (and the response variable). The case column is related to the choice of the fixed parameters for the runs in the same case (DOC, material and feed), the run is the unit of measurement

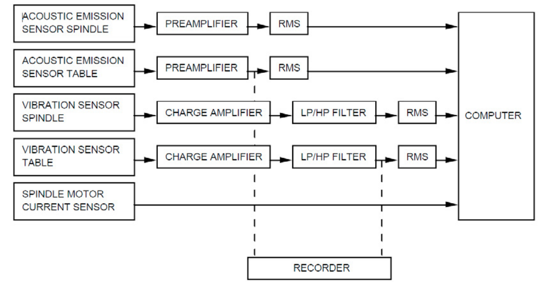

Fig. 2. Experimental setup (from Agogino and Goebel (2007)).

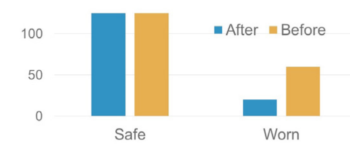

Fig. 3. Classes distribution across the dataset.

used (a run is equivalent to a single working) and each run takes place in a particular time. The values of the last six features are not shown because in each cell there is a multidimensional array corresponding to the analog measurement of the sensors: it is necessary apply a feature generation to transform these multidimensional arrays in a set of mono-dimensional data. The setup of the experiment is as depicted in Fig. 2. The basic setup encompasses the spindle and the table of the Matsuura maching center MC-510V. An acoustic emission sensor and a vibration sensor are each mounted to the table and the spindle of the machining center. The signals from all sensors are amplified and filtered, the fed through two RMS before they enter the computer for data acquisition. The signal from a spindle motor current sensor is fed into the computer without further processing. So, in this way, the experiment design provides 5 sensors output.

The data set presents some empty values in the VB measurement, so it is necessary to remove the entire row. The outliers' management is done by looking the data and the result is that only one is removed because it has signal features whose values greatly differ from the ones of all other observations. As shown by Fig. 3, after the creation of them, the classes result not balanced since the number of observation belonging the class *worn* (20) is significantly lower than that of the *safe* class (125): this is an expected behavior as stopping the measurements according to critical values of VB.

The results of the application of the proposal framework are summarized in Table 2 for the first regression task and the one about the RUL variable. The algorithm for the regression on VB with the best performance is the NN regression. This model returns the minimum error values, as well as the greatest R^2, meaning that it can explain a high percentage of the variance of VB. In the same way, also for the second one regression the best algorithm is the NN regression with a high value for R^2.

Table 3 shows the results for the classification task. In this case the best algorithm is the Two-Class Boosted Decision Tree. It shows the highest evaluation metrics, with an

Table 1. First 4 rows related to the features (and response variable) of milling dataset. The symbol [· · ·] means that in the cell there is a muldimensional array.

case	run	VB	time	DOC	feed	material	smcAC	smcDC	vib_tab	vib_sp	AE_tab	AE_sp
1	1	0	2	1.5	0.5	1	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]
1	2	-	4	1.5	0.5	1	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]
1	3	-	6	1.5	0.5	1	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]
1	4	0.11	7	1.5	0.5	1	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]	[· · ·]

accuracy of almost 96% (the precision is not shown but it is of 92%).

Another important result in this use case is the list of the selected features, that includes 23 out of the 78 (obviously adding the untouchable features such as material and depth of cut). For each type of signal from the sensors, 4 features have been selected amoung the generated ones, except for the *smcDC* from which 3 features have been selected. The situation is practically complete balance between the types of analog signals and therefore presupposes the use of all types of signals.

5. CONCLUSIONS AND FUTURE WORKS

The main aim of this work is to give a general framework that is applicable to cases of predictive maintenance of generic manufacturing tools. Particularly, this method is applicable, as a support to PM, to all tools which activity is managed by parameters provided by the operators and monitored through the application of analog sensors. Using sensors of this type and applying the algorithms and methodologies shown, it is possible creating a prototype to improve the man-machine collaboration in production.

Two activities will follow this work. The first consists in considering more types of sensors, such as thermal ones, and finding the optimal set-up as a compromise between prediction performance and technology costs (find the minimum set of sensors needed and for each of them the minimum sensitivity. The second activity will provide an analysis of the consequences that the use of this technology brings to warehouse management, and consequently to the economy of a company.

ACKNOWLEDGMENTS

This work has been partially supported by "Ministero dell'Istruzione, dell'Università e della Ricerca" Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6".

Table 2. Results of regression algorithms on VB an RUL.

	VB Regression			RUL regression		
Algo.	RMSE	RelSE	R^2	RMSE	RelSE	R^2
LR	0.110	0.182	0.817	1.671	0.178	0.822
DF	0.123	0.225	0.781	1.517	0.134	0.866
BLR	0.116	0.194	0.813	1.640	0.174	0.826
BDT	0.122	0.218	0.794	1.615	0.156	0.844
NN	0.110	0.179	0.821	0.581	0.022	0.979

Table 3. Results of classification algorithms.

Algo.	SP	WP	ACC
LogR	0.960	0.900	0.941
DF	0.936	0.917	0.930
BLR	0.952	0.917	0.941
BDT	0.960	0.950	0.957
NN	0.936	0.900	0.924

REFERENCES

Agogino, A. and Goebel, K. (2007). Milling dataset. URL `http://ti.arc.nasa.gov/project/prognostic-data-repository`.

Al Shalabi, L., Shaaban, Z., and Kasasbeh, B. (2006). Data Mining: A Preprocessing Engine. *J. Comput. Sci.*, 2(9), 735–739.

Bergstra, J. and Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.*, 13, 281–305.

Caesarendra, W. and Tjahjowidodo, T. (2017). A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing. *Machines*, 5(4), 21.

Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. (2010). SMOTE: Synthetic Minority Oversampling Technique. 16, 15–30.

Hassan, S., Tahir, M.M., Badshah, S., Hussain, A., and Anjum, N.A. (2018). Classification of Rigid Rotor Faults Using Time Domain Features Extracted from Multiple Vibration Sensors. 23(3), 43–52.

J. V. Abellan-Nebot, F.R.S. (2010). A review of machining monitoring systems based on artificial intelligence process models. *Int J Adv Manuf Technol*, 47, 237–257.

Mobley, R.K. (2002). *An introduction to Predictive Maintenance*. IEEE Transactions on Instrumentation and Measurement, Butterworth – Heinemann, St. Louis.

P. N. Tan, M. Steinbach, A.K. and Kumar, V. (2005). *Introduction to Data Mining*. Pearson.

R., C. (1999). *Introduction to Manufacturing Processes and Materials*. CRC Press, Boca Raton.

Ravindra, H.V., Srinivasa, Y.G., and Krishnamurthy, R. (1997). Acoustic emission for tool condition monitoring in metal cutting. 212(1), 78–84.

Senawi, A., Wei, H.L., and Billings, S.A. (2017). *A new maximum relevance-minimum multicollinearity (MR-mMC) method for feature selection and ranking*, volume 67.

Sunday Abolarin, M., Oluwafemi Ayodeji, O., Jonathan, J., Solomon Olaoluwa, P., and Clement Kehinde, A. (2015). Effect of cutting speed and feed rate on tool wear rate and surface roughness in lathe turning process. *International Journal of Engineering Trends and Technology*, 22, 171–173. doi:10.14445/22315381/IJETT-V22P236.

Zhang, C., Yao, X., Zhang, J., and Jin, H. (2016). Tool condition monitoring and remaining useful life prognostic based on awireless sensor in dry milling operations. *Sensors (Switzerland)*, 16(6).