

Article

Estimation of Tool Life in the Milling Process—Testing Regression Models

Andrzej Paszkiewicz ^{1,*}, Grzegorz Piecuch ², Tomasz Żabiński ², Marek Bolanowski ¹, Mateusz Salach ¹
and Dariusz Rączka ³

- ¹ Department of Complex Systems, Faculty of Electrical and Computer Engineering, Rzeszow University of Technology, al. Powstancow Warszawy 12, 35-959 Rzeszow, Poland; marekb@prz.edu.pl (M.B.); m.salach@prz.edu.pl (M.S.)
- ² Department of Computer and Control Engineering, Faculty of Electrical and Computer Engineering, Rzeszow University of Technology, al. Powstancow Warszawy 12, 35-959 Rzeszow, Poland; gpiecuch@prz.edu.pl (G.P.); tomz@prz.edu.pl (T.Ż.)
- ³ Faculty of Electrical and Computer Engineering, Rzeszow University of Technology, al. Powstancow Warszawy 12, 35-959 Rzeszow, Poland; draczka@prz.edu.pl
- * Correspondence: andrzejp@prz.edu.pl

Abstract: The article presents an attempt to identify an appropriate regression model for the estimation of cutting tool lifespan in the milling process based on the analysis of the R^2 parameters of these models. The work is based on our own experiments and the accumulated database (which we make available for further use). The study uses a Haas VF-1 milling machine equipped with vibration sensors and based on a Beckhoff PLC data collector. As the acquired sensor data are continuous, and in order to account for dependencies between them, regression models were used. Support Vector Regression (SVR), decision trees and neural networks were tested during the work. The results obtained show that the best prediction results with the lowest error values were obtained for two-dimensional neural networks using the LBFGS solver (93.9%). Very similar results were also obtained for SVR (93.4%). The research carried out is related to the realisation of intelligent manufacturing dedicated to Industry 4.0 in the field of monitoring production processes, planning service downtime and reducing the level of losses resulting from damage to materials, semi-finished products and tools.

Keywords: IIoT; CNC machine; machine learning; smart manufacturing; tool condition monitoring



Citation: Paszkiewicz, A.; Piecuch, G.; Żabiński, T.; Bolanowski, M.; Salach, M.; Rączka, D. Estimation of Tool Life in the Milling Process—Testing Regression Models. *Sensors* **2023**, *23*, 9346. <https://doi.org/10.3390/s23239346>

Academic Editors: Nancy Alonistioti, Spyros Panagiotakis and Evangelos K. Markakis

Received: 18 October 2023
Revised: 17 November 2023
Accepted: 22 November 2023
Published: 23 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of modern industry, especially with regard to the Industry 4.0 concept, requires the interaction of manufacturing machines, sensors and specialised software to ensure greater efficiency, productivity and reliability of production processes. Ensuring the continuity of these processes is a significant challenge for industry [1–3]. The most typical example is bearing diagnostics. This subject has been continuously developed for decades [4–7], but there are fewer solutions in the context of real-time process monitoring. New ideas and solutions must be sought and implemented in order to achieve the desired goals in this area. One important aspect of ensuring the continuity of production processes is the prediction of wear and tear on the tools used. The comparisons included in [8–12] demonstrate how important and current this topic is. Proper estimation makes it possible to reduce costs related to possible unplanned production downtime as well as costs related to wasted material/semi-finished products due to damage to individual tools or machines [13].

The current and well-established realm of research pertains to the investigation of forecasting the extent of tool wear during the execution of technological operations, including drilling and milling. A literature analysis indicates that experiments are predominantly

conducted using numerically controlled or conventional machine tools with various alloys of steel, aluminium, or Inconel as the processed materials. Numerous authors endeavour to devise methodologies for diagnosing cutting processes and predicting tool deterioration through the utilisation of machine learning [14,15].

Anticipating the wear process of a tool until it becomes unsuitable for use, known as Remaining Useful Life (RUL), permits the estimation of potential failure or damage time, thus minimising unforeseen downtime. This is particularly significant, considering that tool damage accounts for 7% to 20% of machine downtime and tool replacement contributes to 3% to 12% of production costs [16,17]. Employing an IT system built upon the Tool Condition Monitoring (TCM) approach proves instrumental in curtailing costs, reducing downtime, and optimising tool usage. Such activities are particularly important in the aviation and automotive industries, which are characterised by multi-series production focused on efficiency while maintaining a very high quality of production. This relates to reducing unplanned downtime and the waste of usually expensive materials and semi-finished products. Developing a solution based on the integration of manufacturing devices, sensors, and a data collection and processing system that will allow the continuous supervision of the manufacturing process is important. Moreover, the work carried out allowed for the preparation of a large standardised set of data that can be shared, which can be used to conduct further research work.

By analysing sensory data and the operating parameters of a device, artificial intelligence algorithms can contribute to the earlier detection of signs of future problems or failures [18–21]. This makes it possible to take preventive action, which in turn increases system reliability and saves costs. To classify tool conditions or predict failure instances, various signals from sensors, process variables, and images are harnessed. Accelerometers, force sensors, spindle current, work table drive axes, and acoustic emission sensors are the most commonly utilised. These measurement signals are assessed individually or combined, considering the intricacy of the problem. In particular, regression methods are often used for equipment damage prediction because of their ability to model relationships between different variables [22]. In the case of equipment damage prediction, we often deal with continuous data, such as equipment operating parameters, operating time or sensory values, and regression methods allow us to model these relationships and predict continuous values. Also of great importance is the fact that regression algorithms can be trained on historical data mapping specific operating conditions as well as patterns that are specific to the device. Of course, the models can be updated as a result of new data being provided. It is also not insignificant that regression models can be effectively trained and implemented in environments where a lot of data are received, providing scalability, including in continuous production environments. Therefore, by analysing the correlation between the signal and tool wear across time, frequency, and time–frequency domains, conventional metrics like RMS are determined and employed in shallow (ANN, SVM, HMM, SVR, LR) and deep (RNN, LSTM, CNN) classification and regression models.

In the manufacturing industry, there is a significant demand for intelligent production systems tasked with monitoring, for example, the wear progression of tools. These systems are expected to alert operators in real time based on the gathered data about potential failure or quality discrepancies in the product [23]. During the implementation of such systems, decisions are made regarding which signals to measure and how they should be collected and processed. The authors of [24] indicate that vibrations directly reflect the operational state of devices, considering the mechanical behaviour of devices, such as motor rotation, flow rates in pipes, etc. They consider vibration to be independent of other external factors such as temperature and humidity. They also propose the hypothesis that machine vibration effectively reflects the ageing process of devices, making it an excellent indicator of early failure symptoms. In [25], it is added that machine vibration analysis brings many benefits, including increased revenue, production reliability, shorter downtime, and lower repair costs.

Vibration levels generated by individual machine components are most commonly recorded, but signals such as acoustic emissions, power consumption, forces, temperatures, etc., are also often collected. With the increasing number of recorded signals, which can be several dozen on a single machine, there is a growing demand for efficient information systems capable of handling large amounts of streaming data (big data). Researchers and practitioners refer to such a model as 4V—volume, velocity, variety, and veracity [26,27]. A large dataset size also means that transferring data for processing is uneconomical; instead, calculations should be performed on the same machine where the data are stored. They also highlight the need to build analytical platforms for diagnostics based on additional sensors, given that factory-installed sensors typically provide data only for internal control systems. Moreover, in critical machine parts, pre-installed sensors are often lacking [26]. As the number of data increases exponentially with the development of IoT technology [28], research is being conducted to develop methods that maintain good classification quality while reducing the amount of data used in the learning process [29,30].

The methods of computational intelligence are utilised both for diagnosing and predicting the wear of CNC machine components, engines, gear assemblies, wind turbines, transformers, and many other devices. In practice, however, it is challenging to encounter a functioning real-world diagnostic or predictive system, which is largely due to the uncertainty in assessing and predicting the equipment's condition and/or high initial investment costs. Much of the ongoing research in this field remains theoretical. Most often, these studies are conducted based on laboratory objects or publicly available datasets. There is a lack of research on the production reality and evidence of efficiency improvement based on the solutions described in the literature. Thus, the aim of this study is to verify the R^2 parameter of regression models for estimating tool life in the milling process based on real data from the production environment. This article presents sample results of the cutting tool life estimation for a real CNC machine and a dataset collected in house, which we share more widely.

The article consists of several sections. Section 2 describes the laboratory bench built and used during the research, consisting of a three-axis Haas VF-1 milling machine and expanded to include additional vibration sensors, specifically accelerometers and current transducers. This stand enabled the collection of data necessary for the machine learning algorithms. Section 3 contains the characteristics of the regression models used, together with the parameters considered, as well as the results obtained from the tests carried out.

2. Testbed and Data Description

The research stand involves working with a three-axis Haas VF-1 milling machine. This specific machine is equipped with a proprietary measurement system and is augmented by additional vibration sensors, specifically accelerometers and current transducers (Figure 1).



Figure 1. Haas VF-1 milling machine and mounting location of 3 sensors on the spindle.

The VF-1 milling machine integrates a meticulously designed measurement system based on industrial automation components: Industrial PC (IPC) Beckhoff C6920, along with a combination of I/O modules, specifically EL1008, EL2008, EL3632 (4 units), EL3702, and EL3413 (4 units). The software driving this system is TwinCAT 3.1.

The measurement system includes strategically positioned sensors:

- A total of eight accelerometers (sensitivity 100 mV/g)—Dytran 3055A2 and one Dytran 3056D2T. Three of them on the spindle (Figure 1—Ch1: +Y axis, Ch2: −Z axis, Ch3: −X axis) and five sensors on the drives of the work table were mounted (for X direction Ch4: +Z axis, Ch5: −X axis and for Y direction Ch6: +Z axis, Ch7: +Y axis, Ch8: −X axis).
- A total of twelve current transducers—Wago 855-4001/0100-0001, accompanied by Wago 2007-8875 terminal blocks, are also integral to the system and were mounted in the electrical cabinet for each phase (L1, L2, L3) for the spindle and the X, Y, and Z axes.

When designing the experiments, consideration was given to other studies [31–38] to ensure the comparability of results. Experiments similar to previous studies but with a larger number of workpieces were proposed. The choice of material was influenced by its widespread use in the aviation industry [39] and the automotive industry [40–42] as well as the fact that steel workpieces, unlike aluminium, accelerate tool wear. Material 42 CrMo4 with manufacturing tolerance IT12 and hardness 38 ± 2 HRC (thermally hardened) was selected, although not all workpieces met the specified hardness parameter. Hardness measurements ranged from 35.33 to 41.67 HRC. Cuboid samples were utilised to facilitate clamping in the machine vice with dimensions of 80 mm × 80 mm × 150 mm chosen to ensure secure fastening to a depth of approximately 20 mm (Figure 2).



Figure 2. Samples before and after milling (on the left) and during the milling (on the right).

The key to conducting the experiments properly was continuous milling without retracting the tool from the material. To reflect real-life conditions and adapt to industrial machining, the milling path followed a clockwise movement around the workpiece contour. Due to the specificity of the milling process, the vibrations, power consumption, and roughness obtained differ depending on whether the milling is co-rotating or counter-rotating. In this case (omnidirectional movement), both types of milling occurred, so the direction of machining CW or CCW does not matter; only the order is different.

Some experiments used cutting parameters recommended by technologists, while others used unfavourable parameters. This was intended to accelerate the degradation of cutters as well as to examine the negative impact of changing milling parameters on their service life. The side recess of the cutter in the material was a crucial parameter with the manufacturer's recommendation limiting it to 45% of the tool's diameter. This allowed for 4 cycles per layer with 80% cutter insertion (non-optimal) and 7 cycles per layer with 45% cutter insertion (optimal) for a sample of this size (RDOC—Radial Depth of Cut values were 4.5 and 8 mm). The number of milling cycles also depended on the Axial Depth of Cut (ADOC), resulting in 48 to 84 cycles at ADOC = 5 mm and 24 to 42 cycles at ADOC = 10 mm. Two types of 10 mm diameter cutters were used: high quality (Van Hoorn VHVTR 4 100 070 10 03 050) and medium quality (PARA Tooling RS4 10.0x70). The parameter values for both cutters are: $d = 10$ mm (diameter), $L = 70$ mm (total length),

$L_1 = 22$ mm (cutting length), and $z = 4$ (number of flutes). Additionally, the impact of the tool holder length on system vibration was compared in the tests (two different lengths).

The collected data were labelled by sample number (P), layer number (F) and cycle number (C). Each sample was a single cuboid block of the material being processed, which was divided into layers corresponding to the thickness of material removed at each stage of the experiments. Cycles, each consisting of a single machining program run clockwise, were distinguished within the layers. The process of milling is illustrated in Figure 3. The number of layers and cycles for a given sample depended on the adopted ADOC and RDOC values. The basic process parameters were as follows: spindle speed: 3200 rpm, feed rate 640 mm/min, and tool holder length: 80 and 160 mm. Other process data and details are described on <https://datasets.kia.prz.edu.pl/> (accessed on 17 October 2023).

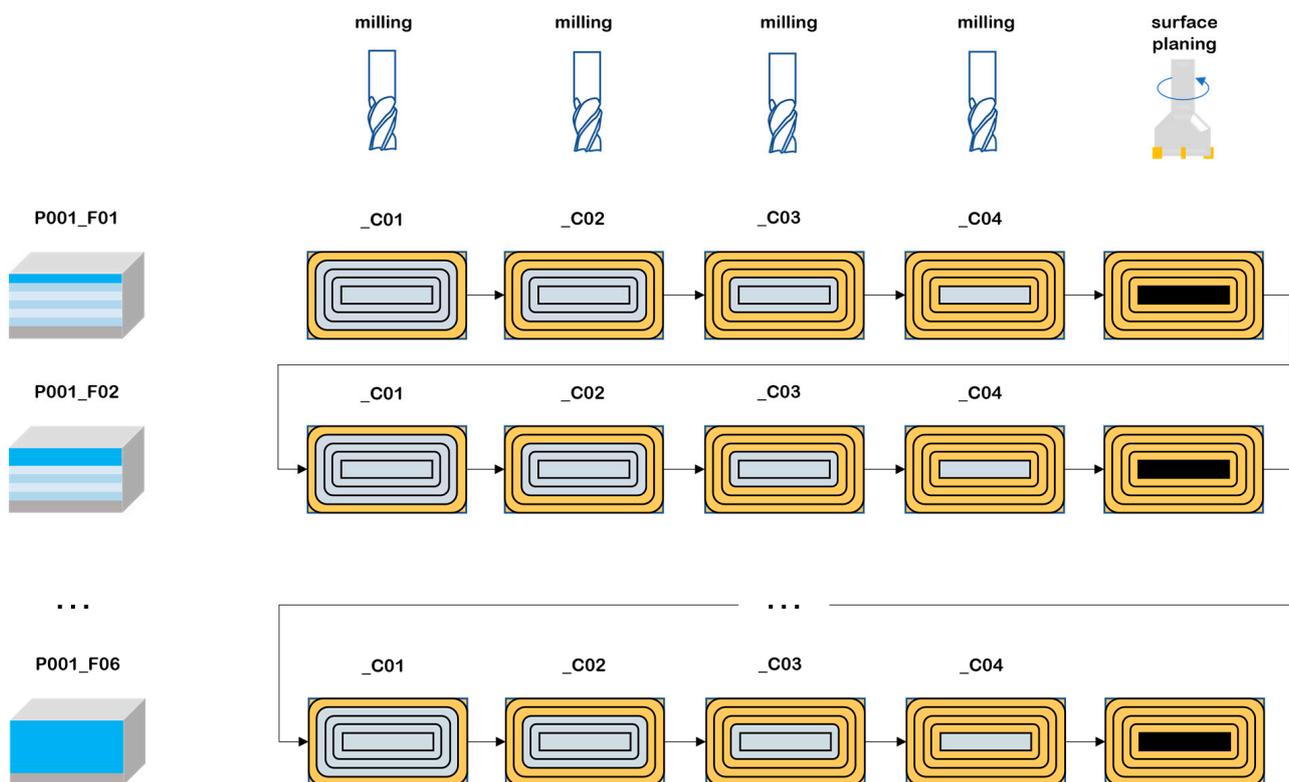


Figure 3. Process of conducting experiments.

Table 1 contains all statistical data regarding working time and the amount of material collected by individual cutters. Tool ID numbers from 0 to 100 apply to Van Hoorn cutters and over 100 apply to PARA Tooling cutters. Tools with ID 10 and 11 in the 160 mm tool holder length were mounted (for tool no. 11, the breakage was immediate). The collected data (Table 1) show that the RDOC parameter significantly affects the tool life, which directly affects the number of cycles performed. Based on the data, it was also observed that the ADOC parameter has no noticeable impact on the cutter lifespan. Therefore, assuming the same number of cycles possible for ADOCs of 5 and 10 mm, the volume of material collected from the sample will be twice as large with an ADOC of 10 mm. The conclusion is consistent with the HEM strategy (High-Efficiency Milling) [43], which consists of using the entire tool length (larger ADOC value) at the expense of reducing the radial depth of cut. The advantage of this milling strategy is the possibility of greater heat distribution in the tool, which increases its service life. For three PARA Tooling cutters, a decrease in the number of cycles performed was visible while maintaining optimal process parameters.

Table 1. Statistical data about tool lifetime—approximate values.

Tool ID	ADOC (mm)	RDOC (mm)	Number of Cycles Performed by the Tool	Total Tool Operating Time (s)	Total Length of the Tool Path (mm)	Total Volume of Milled Material (mm ³)
2	5	8	49	1526	17,932	717,280
3	10	8	49	1526	17,932	1,434,560
4	5	8	40	1240	14,560	582,400
5	10	8	53	1605	19,388	1,551,040
6	5	8	29	906	10,652	426,080
7	10	4.5	126	3852	44,352	1,995,840
8	5	4.5	115	3531	40,696	915,660
9	5	4.5	149	4569	52,628	1,184,130
10	10	4.5	14	528	4928	221,760
11	10	4.5	0	0	0	0
101	10	4.5	83	2528	29,108	1,309,860
102	5	4.5	84	2568	29,568	665,280
103	10	4.5	59	1809	20,876	939,420
105	5	4.5	124	3812	43,892	987,570

3. Preparation and Testing of Regression Models

For each of the tested learning algorithms, multiple models were created with different parameters specific to each algorithm. To achieve this, the GridSearchCV function from the sklearn library was employed as it is one of the default methods to optimise hyperparameters in the sklearn library. To prevent overfitting during training, a 10-fold cross-validation was used. The training data subset was randomly divided into 10 equal parts, and the model with the specified parameters was trained 10 times, each time using a different subset as the validation set.

The evaluation metric used to assess the quality of the models was the R^2 score, which takes a value of 1.0 in the case of a perfect result. The closer the quality metric value is to 0.0, the lower the accuracy of the created model. Finally, the best model was selected for each algorithm (the model with the best parameter set) and tested on a separate 20% test dataset.

3.1. SVR

The research was conducted using the SVR (Support Vector Regression) function from the sklearn library. It is a regression technique based on support vector machines (SVMs) that allows the modelling and prediction of continuous values using C and epsilon-free parameters. As a result of the initial analysis in machine learning, the utilisation of a kernel based on the radial basis function proved to be the most effective. This kernel allows for the transformation of data into a space of infinite dimensionality. The radial basis function (RBF) is the default kernel method in the sklearn library's Support Vector Regression (SVR). Since the initial learning results for other kernel functions were unsatisfactory, the analysis focused solely on the radial basis function.

To find the optimal model, 271,441 different combinations of the following parameters were examined [44]:

- Parameter C—This is a regularisation parameter that controls the importance of allowable errors. A higher value of this parameter gives more weight to errors (loose values falling outside the decision boundaries of the model), leading to more flexible and complex models (the model will try harder to fit the data). However, an excessively high value of this parameter can lead to overfitting. The default value of parameter C is 1.0. In the conducted experiments, 521 values were tested in the range from 10^{-9} to 10^4 , which were evenly spaced on a logarithmic scale.
- Parameter gamma—This is a kernel parameter that determines how far the influence of a single data point reaches on the model's prediction. A high value of the gamma

parameter makes points close to each other have a greater influence on the prediction, while a low value of this parameter emphasises points that are far apart. In the sklearn library, the default value of the gamma parameter is calculated according to the formula: $1/(\text{number of features} \times \text{variance})$. In the conducted experiments, 521 values were tested in the range from 10^{-4} to 10^9 , which were evenly spaced on a logarithmic scale.

Figure 4 presents the results of the obtained learning for the SVR for both the full range of tested parameters and the range of best results obtained.

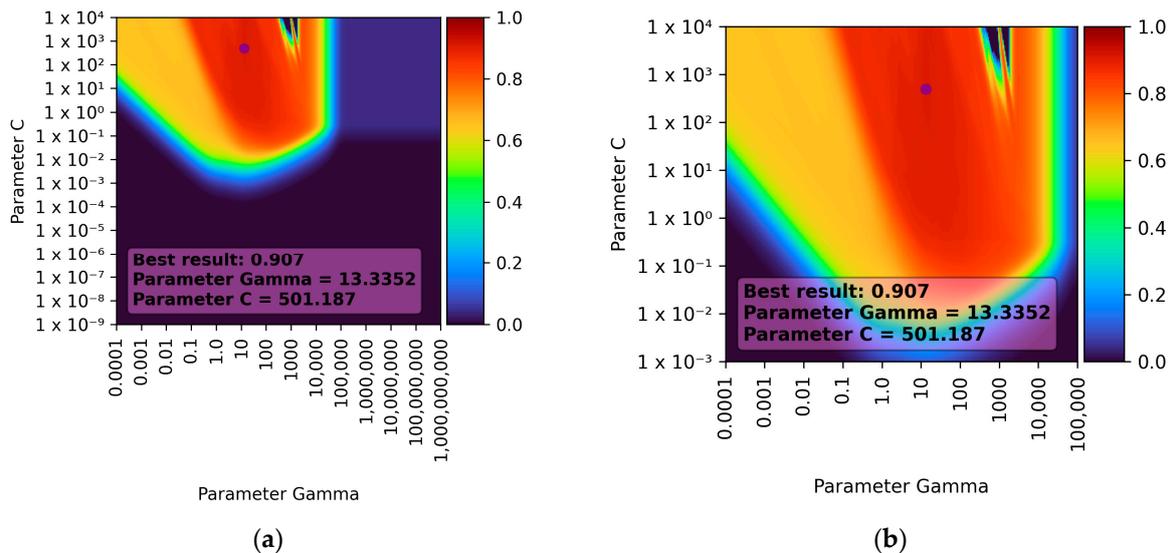


Figure 4. Visualisation of teaching results for the SVR model: (a) full range of tested parameters; (b) range of best results.

It can be observed that the best result was achieved with a gamma parameter of approximately 13.3352 and a C parameter of approximately 501.187. For this parameter combination, the R^2 score was 0.907 ± 0.119 (double standard deviation), which translates to approximately 90% prediction accuracy. When the C parameter was smaller than approximately 10^{-3} and the gamma parameter was larger than 10^5 , the learning results were consistently low. It is also noticeable that the gamma parameter has a greater influence on the learning results (for the C parameter, the learning results remained nearly constant after reaching a certain threshold). These observations and the learning results depending on individual parameters (for the best combination) are presented in Figure 5.

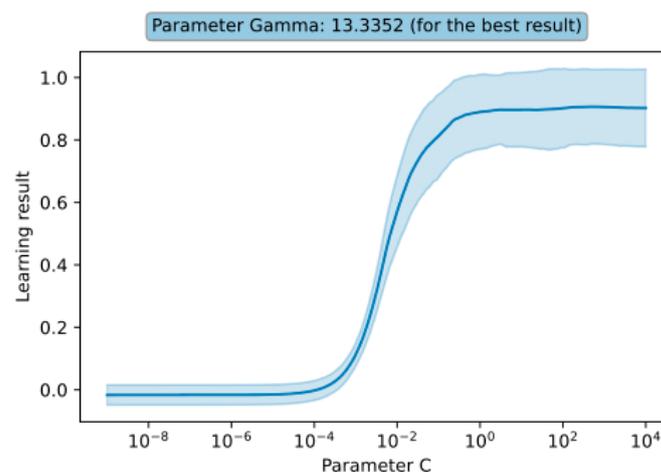


Figure 5. Results of teaching the SVR model depending on the C parameter.

3.2. Regression Decision Tree

The research was conducted using the DecisionTreeRegressor function from the sklearn library. It is the default method for fitting regression models using decision trees, which are flexible tools for modelling nonlinear relationships and are widely used in regression tasks. To find the optimal model, 77,616 different combinations of the following attributes and parameters [45] were tested:

- Splitting criterion (criterion parameter)—describing the function used to measure the quality of a split. The following functions were tested:
 - a. Mean Squared Error (“squared_error”)—minimises the Mean Squared Error between predicted and actual values. It is calculated as the arithmetic mean of the squared prediction errors for each sample in the dataset.
 - b. Friedman Mean Squared Error (“friedman_mse”)—operates similarly to Mean Squared Error but incorporates a Friedman correction to minimise the impact of noise on prediction results.
 - c. Mean Absolute Error (“absolute_error”)—minimises the Mean Absolute Error between predicted and actual values.
 - d. Poisson Deviance (“poisson”)—uses Poisson deviance reduction to find appropriate splits.
- Splitting strategy at each node (“splitter”)—two options were available within the algorithm, both of which were tested:
 - a. Best (“best”)—selects the best possible split.
 - b. Random (“random”)—selects the best random split.
- Minimum number of samples required to split a node (“min_samples_split”)—values were tested in the range from 2 to 99.
- Minimum number of samples required to create a leaf node (“min_samples_leaf”)—values were tested in the range from 1 to 99.

In the case of decision tree utilisation, the best learning results were achieved using the criterion based on Poisson deviance. The trained model exhibited approximately 92% accuracy with an R^2 parameter of 0.915 ± 0.103 (double standard deviation). Similar learning results were obtained with criteria based on the Mean Squared Error; almost identical values were achieved in both cases. The criterion based on mean absolute error performed significantly worse in learning.

The choice of different splitting strategies did not affect the best result achieved by individual criteria. However, it is noticeable that the best split strategy resulted in better prediction results across the entire range of tested parameters. The experiments showed that in each of the tested cases, the best learning results were achieved with a relatively small number of samples required to split a node or create a decision tree leaf node.

3.3. Regression by One-, Two- and Three-Layer Neural Networks

The research was conducted using the MLPRegressor function from the sklearn library. It is the default method for fitting a regression model using a multi-layer perceptron, which allows the modelling of complex nonlinear relationships between features and the explanatory variable. To find the optimal artificial neural network model, a total of 195,088 different combinations of learning parameters were tested. The parameters tested included the following [46]:

1. Activation function parameter—defining the activation function used in hidden layers. All available activation functions in the sklearn library were tested (identity, logistic, tanh, ReLU).
2. Solver parameter—determining the optimisation algorithm used to minimise the cost function and find the appropriate set of weights in the neural network. The Adam and LBFGS solvers were tested.
3. Number of neurons in hidden layers:

- a. For single-layer networks—the number of neurons in the hidden layer was tested from 1 to 496 with a step of 5 (a total of 100 possibilities).
 - b. For two-layer networks—the number of neurons in individual layers was tested from 2 to 299 with a step of 3 (a total of 10,000 combinations).
 - c. For three-layer networks—the number of neurons in individual layers was tested from 5 to 205 with a step of 10 (a total of 9261 combinations).
4. For single-layer networks, the influence of the alpha parameter on network learning was also tested. This parameter determines the regularisation penalty added to the cost function and directly affects the regularisation effect. The default value of the alpha parameter in the sklearn library is 10^{-4} . In the conducted experiments, 51 values were tested in the range from 10^{-5} to 10^5 , which were evenly spaced on a logarithmic scale.

3.3.1. Single-Layer Network

In the case of single-layer networks, the best learning results were achieved using the ReLU activation function and the LBFGS solver. The best-trained model exhibited approximately 96% accuracy with an R^2 parameter of 0.962 ± 0.085 (double standard deviation). When using the Adam solver, the best result was also achieved with the ReLU activation function, but the learning accuracy was much lower, around 73%, with an R^2 parameter of 0.728 ± 0.214 .

Satisfactory learning results were also achieved with the LBFGS solver when using the hyperbolic tangent function (for fewer than 100 neurons) and the logistic function (for fewer than 20 neurons). In other cases, the prediction accuracy did not exceed 70%. For all learning combinations, it was observed that the alpha parameter had a nearly uniform impact. Results rapidly decreased to almost zero for values from 1 to 10 and remained at that level until the maximum tested value. The best results were observed for small values of the alpha parameter (values less than 1). The aggregate learning results for different activation functions and the LBFGS solver are presented in Figure 6.

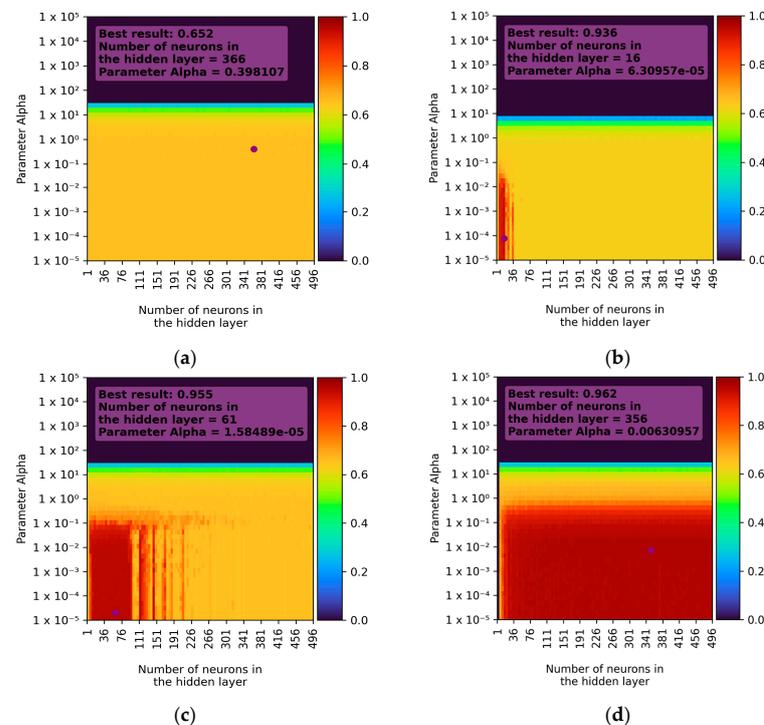


Figure 6. Example of learning results for the LBFGS solver depending on activation function (single-layer network): (a) activation function—identity; (b) activation function—logistic; (c) activation function—Tanh; (d) activation function—ReLU.

3.3.2. Two-Layer Network

In the case of two-layer networks, similar to single-layer networks, the best learning results were achieved using the ReLU and Tanh activation functions with the LBFGS solver. The best models exhibited approximately 96% accuracy with an R^2 parameter of 0.959 ± 0.065 for ReLU and 0.959 ± 0.034 for Tanh (Figure 7). When using the Adam solver, the best result was also achieved with the ReLU activation function, but the learning accuracy was lower, around 88%, with an R^2 parameter of 0.877 ± 0.104 .

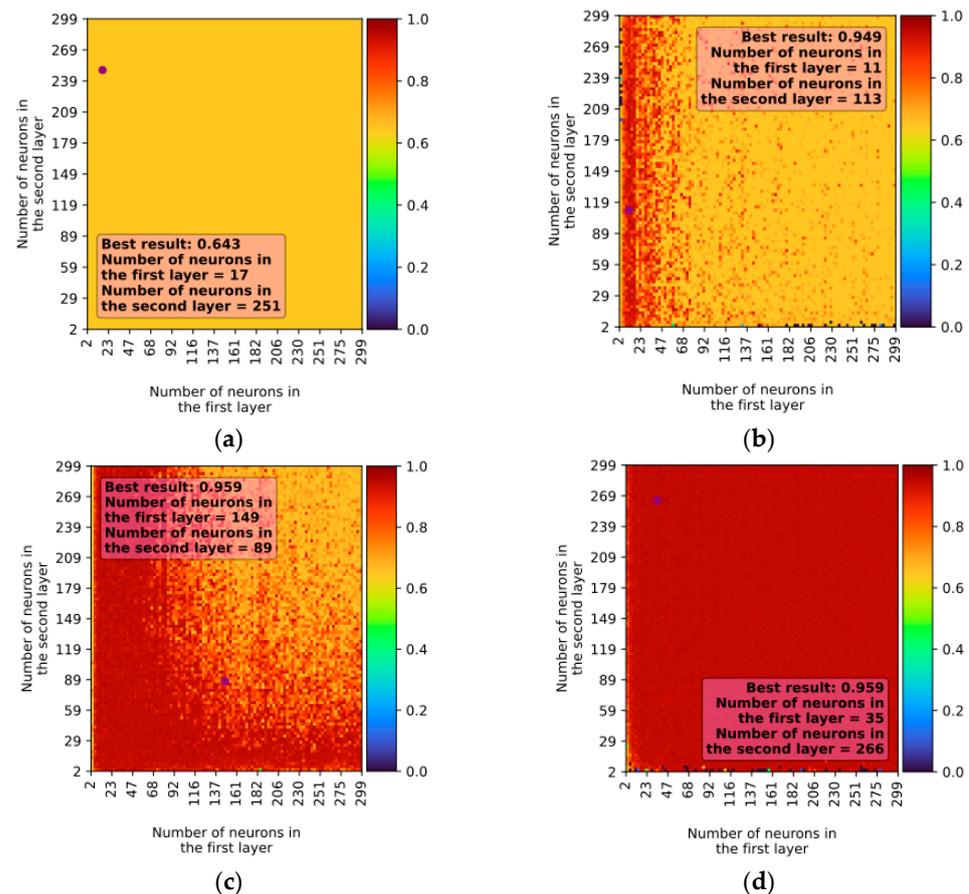


Figure 7. Example of learning results for the LBFGS solver depending on activation function (two-layer network): (a) activation function—identity; (b) activation function—logistic; (c) Activation function—Tanh; (d) activation function—ReLU.

3.3.3. Three-Layer Network

For both solvers, the identity activation function resulted in only approximately 65% learning accuracy across the entire range of tested parameters. Using the LBFGS solver, the results obtained by using a linear activation function differed only within 0.001 (Figure 8a,b), which corresponds to the results obtained for one-layer and two-layer networks. From these results, we can conclude that there are no linear relationships between the input data.

In the tested cases, the logistic activation function performed poorly with exceptionally low learning results (Figure 8c). The vast majority of results did not exceed an accuracy of more than 1%. There were a few combinations of parameters when using the LBFGS solver that achieved higher learning results (up to around 90%—Figure 8d). Higher regression results were achieved for neural networks with fewer neurons in each layer (as in the cases of single-layer and two-layer networks); however, this may have been achieved by randomly fitting the test data to the model (bias). The results obtained confirm the information found in the sources that the logistic function is definitely better suited in the context of classification problems than for regression problems.

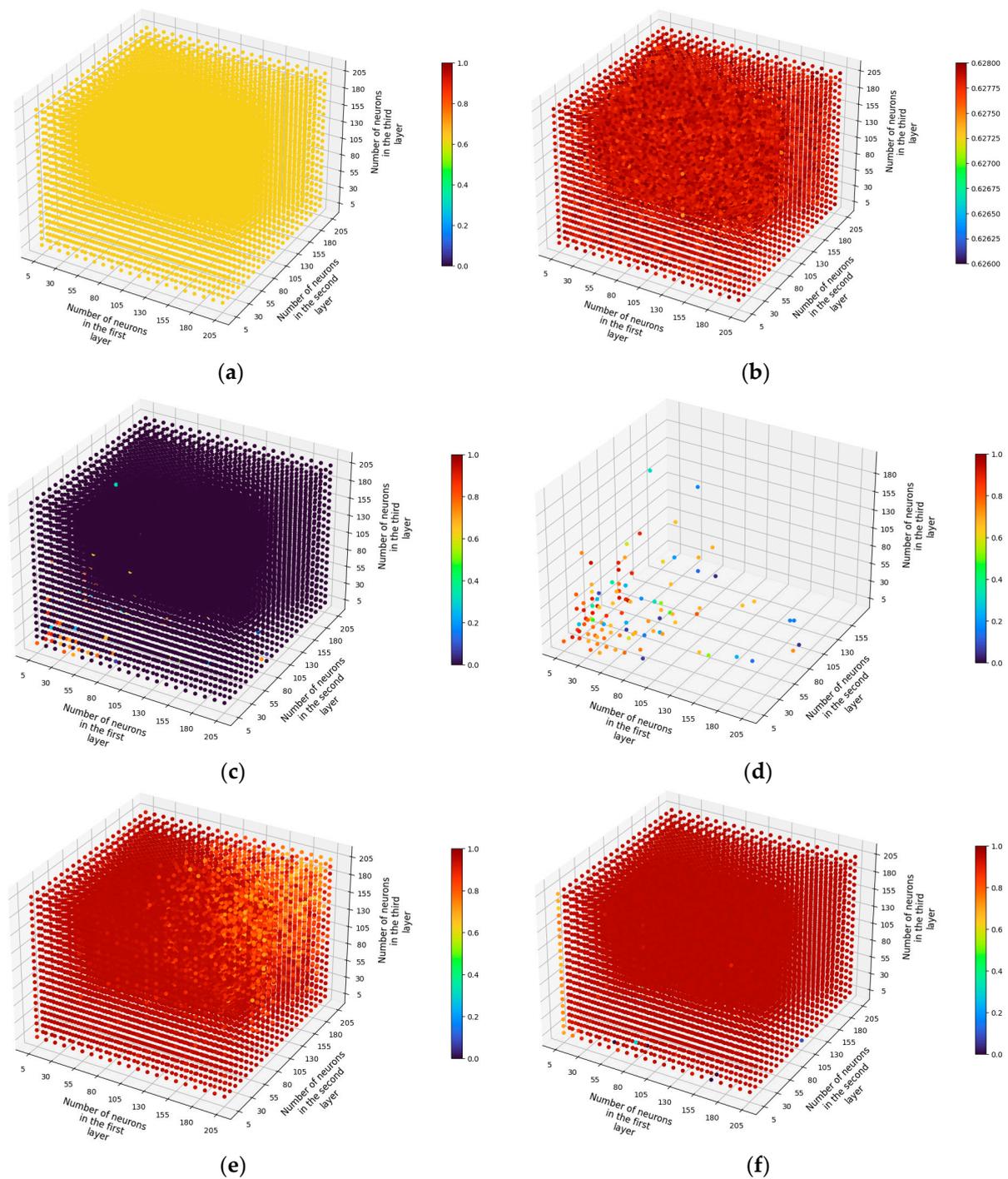


Figure 8. Example of learning results for the LBFGS solver depending on activation function (three-layer network): (a) activation function—identity; (b) activation function—identity, scale adjusted to results; (c) activation function—logistic; (d) activation function—logistic, accuracy greater than 1%; (e) activation function—Tanh; (f) activation function—ReLU.

Using the Adam solver, the hyperbolic tangent function returned results similar to those obtained with the identity function. The LBFGS solver performed significantly better, allowing for a maximum learning accuracy of approximately 96% with an R^2 parameter of 0.96 ± 0.052 being the best-achieved value for the tested three-layer networks. The ReLU activation function consistently allowed for high learning results. For the LBFGS solver, the maximum accuracy achieved with this function was approximately 96% with an R^2

parameter of 0.958 ± 0.052 . For the Adam solver, the highest accuracy was around 93% with an R^2 parameter of 0.93 ± 0.073 . The aggregate learning results for different activation functions and the LBFGS solver are presented in Figure 8.

4. Discussion

The problem lies in selecting an appropriate classification or regression model for the dataset under consideration. If correct assumptions about the dataset are not made, no computational intelligence algorithm can be deemed superior to others as a consequence of the No Free Lunch Theorem. It often happens that data or their features exhibit a linear nature, making the application of linear models most appropriate. On the other hand, for more complex data, the use of models built on neural networks, for example, might be more suitable. Therefore, for each dataset, several to several dozen algorithms should be appropriately selected and tested.

The best results obtained on the learning dataset were achieved for one-, two- and three-layer neural networks using the LBFGS solver. These results were characterised by the lowest error values. In the case of the Adam solver, it was observed that increasing the number of layers positively impacted learning accuracy. For support vector machines and decision trees, similar results were achieved with accuracy levels exceeding 90%. A summary of the results achieved for each of the machine learning methods is shown in Figure 9.

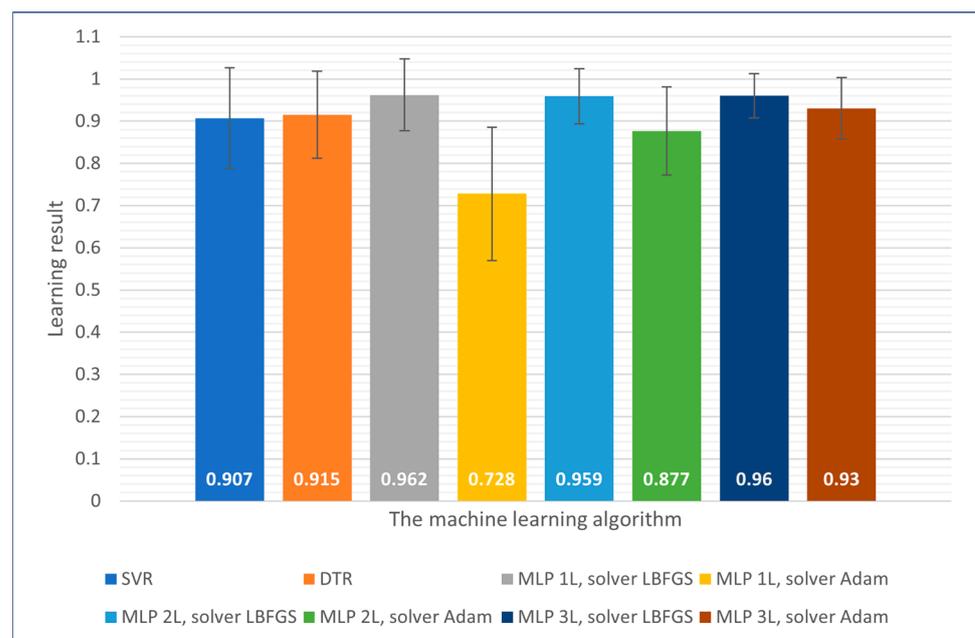


Figure 9. The results achieved for each of the machine learning methods—training dataset.

The machine learning models obtained were then tested on a 20% randomly selected test dataset, which was separated before the learning process. This step was aimed at checking that the models were not overfitting to the training data. The obtained results are presented in Figure 10.

The obtained prediction results have similar levels to the predictions achieved on the training datasets and fall within the range of the error obtained through 10-fold cross-validation. Considering that the created models have not had prior access to the input data on which the final validation was carried out, we can exclude the occurrence of overfitting.

With the exception of single-layer neural networks using the Adam solver, all machine learning algorithms enabled the achievement of satisfactory and high-quality prediction models.

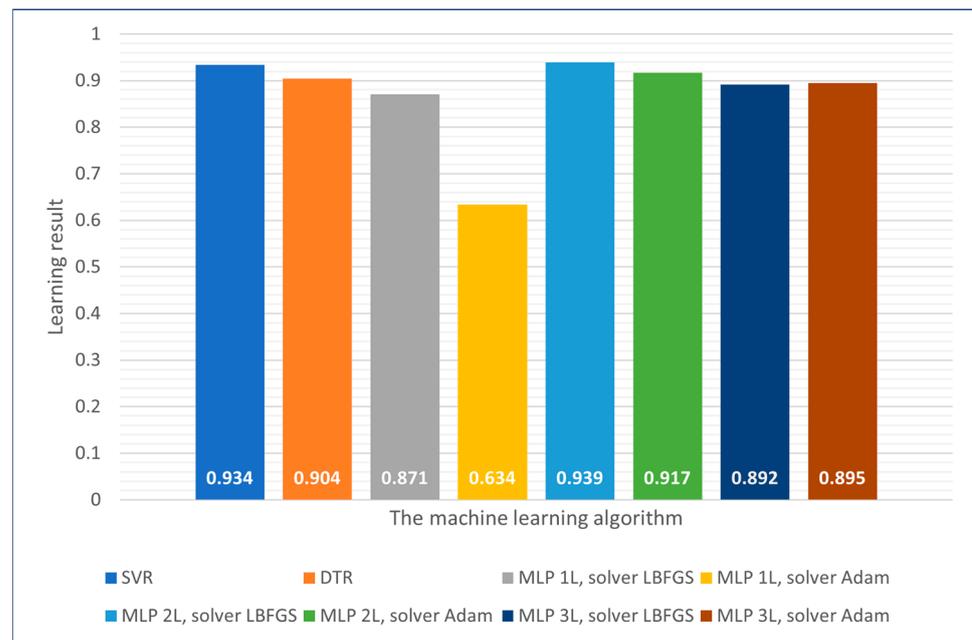


Figure 10. Results of the best models of individual machine learning algorithms—test dataset.

The fastest learning algorithms were decision trees and the support vector machine algorithm. In the case of artificial neural networks, the learning speed was inversely proportional to the number of network layers.

A comparative analysis of the various machine learning regression models made it possible to show how they behave in a specific context, in this case, in predicting the failure of a milling machine. This allows the selection of the best method or approach to a given problem, which can significantly contribute to improving the efficiency and quality of industrial processes.

The models created would achieve very high learning efficiency. Referring to the literature and other studies that assume the prediction of the life cycle of a tool, it can be considered that the results achieving an accuracy level of more than 90% are very satisfactory. However, comparisons of this type are not conclusive, since it is impossible to objectively compare models based on different learning data; therefore, the effectiveness of each model should be evaluated individually. Sometimes, due to the high complexity of the data, it may not be possible to achieve such high accuracy as in the case we studied, but this does not mean that such results are not equally valuable.

The use of artificial intelligence in industry is of great importance in the context of improving efficiency, optimising costs and improving the safety of production processes. The proper selection and optimisation of machine learning methods is one of the key aspects of creating effective solutions.

The conducted research allows us to understand what factors affect the effectiveness of individual methods. By performing a series of tests, it was possible to optimise learning parameters, which can be a valuable indication for engineers and specialists in the field of industrial process automation. Improperly selected learning parameters can prolong the learning process even several times without providing any benefit in the achieved results.

In addition, by using machine learning models to monitor and predict failures, companies can introduce a proactive approach to equipment maintenance. This makes it possible to avoid prolonged production downtime and costly repairs, which significantly contributes to saving time and resources.

This research is also crucial for the continuous improvement of industrial processes. Based on the results of benchmarking, machine learning models can be adjusted and improved, leading to more accurate predictions and the more efficient use of production resources.

It is worth noting that the artificial intelligence-based solutions developed can be adapted and scaled to different types of industrial machinery and equipment, making them versatile tools in the process of automation and optimisation in industry.

5. Conclusions

The use of artificial intelligence in industry is becoming increasingly widespread. As the performance of the prediction mechanisms is strictly dependent on the adopted training dataset, a significant aspect of the conducted work involved constructing an appropriate dataset under industrial conditions. These data formed the basis for further research regarding the application of regression methods for estimating tool life in the milling process. As a result of the conducted experiments, the following conclusions were drawn:

- The best results for both the training and testing datasets were achieved using two-layer neural networks with the LBFGS solver (95.9% and 93.9%, respectively).
- The worst results were obtained for both datasets in the case of one-layer neural networks using the Adam solver (72.8% for the training dataset and 63.4% for the testing dataset).
- Other employed models yielded results around 90% for both datasets.

Taking into account the obtained results, it can be assumed that regression models allow the estimation of cutting tool wear in the milling process. The information obtained in this way can be used in the IT system to plan production downtime, but above all, it can lead to counteracting uncontrolled failures and damage to manufactured components and semi-finished products. Such activities are particularly important in the aviation and automotive industries, where materials and manufactured elements are often expensive and, moreover, must meet high-quality standards. Therefore, elements manufactured with possible defects resulting from faulty tools are not allowed. The use of tested models, especially neural networks, may contribute to the elimination or significant reduction in these undesirable situations.

Author Contributions: Conceptualisation, A.P., G.P. and D.R.; methodology, A.P., G.P., T.Ż., M.B., M.S. and D.R.; software, A.P., G.P. and D.R.; validation, A.P., G.P., T.Ż., M.B., M.S. and D.R.; formal analysis, A.P., G.P., T.Ż., M.B., M.S. and D.R.; investigation, A.P., G.P., T.Ż., M.B., M.S. and D.R.; resources, A.P., G.P. and D.R.; data curation, A.P. and G.P.; writing—original draft preparation, A.P., G.P., T.Ż., M.B., M.S. and D.R.; writing—review and editing, A.P., G.P., T.Ż., M.B., M.S. and D.R.; visualisation, A.P., G.P., T.Ż., M.B., M.S. and D.R.; supervision, A.P. and G.P.; project administration, A.P.; funding acquisition, A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This project is financed by the Minister of Education and Science of the Republic of Poland within the “Regional Initiative of Excellence” program for the years 2019–2023. Project number 027/RID/2018/19, amount granted 11 999 900 PLN.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and photos are available publicly and free of charge on the website: <https://datasets.kia.prz.edu.pl/>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kalsoom, T.; Ramzan, N.; Ahmed, S.; Ur-Rehman, M. Advances in Sensor Technologies in the Era of Smart Factory and Industry 4.0. *Sensors* **2020**, *20*, 6783. [[CrossRef](#)]
2. Jamwal, A.; Agrawal, R.; Sharma, M.; Giallanza, A. Industry 4.0 Technologies for Manufacturing Sustainability: A Systematic Review and Future Research Directions. *Appl. Sci.* **2021**, *11*, 5725. [[CrossRef](#)]
3. Kosieradzka, A.; Smagowicz, J.; Szwed, C. Ensuring the Business Continuity of Production Companies in Conditions of COVID-19 Pandemic in Poland—Applied Measures Analysis. *Int. J. Disaster Risk Reduct.* **2022**, *72*, 102863. [[CrossRef](#)] [[PubMed](#)]
4. Mao, W.; Ding, L.; Tian, S.; Liang, X. Online detection for bearing incipient fault based on deep transfer learning. *Measurement* **2020**, *152*, 107278. [[CrossRef](#)]

5. Minervini, M.; Mognaschi, M.E.; Di Barba, P.; Frosini, L. Convolutional Neural Networks for Automated Rolling Bearing Diagnostics in Induction Motors Based on Electromagnetic Signals. *Appl. Sci.* **2021**, *11*, 7878. [[CrossRef](#)]
6. Borghesani, P.; Smith, W.A.; Randall, R.B.; Antoni, J.; El Badaoui, M.; Peng, Z. Bearing Signal Models and Their Effect on Bearing Diagnostics. *Mech. Syst. Signal Process.* **2022**, *174*, 109077. [[CrossRef](#)]
7. Peng, B.; Bi, Y.; Xue, B.; Zhang, M.; Wan, S. A Survey on Fault Diagnosis of Rolling Bearings. *Algorithms* **2022**, *15*, 347. [[CrossRef](#)]
8. Çınar, Z.M.; Nuhu, A.A.; Zeeshan, Q.; Korhan, O.; Asmael, M.; Safaei, B. Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0. *Sustainability* **2020**, *12*, 8211. [[CrossRef](#)]
9. Achouch, M.; Dimitrova, M.; Ziane, K.; Sattarpanah Karganroudi, S.; Dhoub, R.; Ibrahim, H.; Adda, M. On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges. *Appl. Sci.* **2022**, *12*, 8081. [[CrossRef](#)]
10. Pilar Lambán, M.; Morella, P.; Royo, J.; Carlos Sánchez, J. Using Industry 4.0 to Face the Challenges of Predictive Maintenance: A Key Performance Indicators Development in a Cyber Physical System. *Comput. Ind. Eng.* **2022**, *171*, 108400. [[CrossRef](#)]
11. Jimenez, J.J.M.; Schwartz, S.; Vingerhoeds, R.; Grabot, B.; Salaün, M. Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics. *J. Manuf. Syst.* **2020**, *56*, 539–557. [[CrossRef](#)]
12. Wang, J.; Gao, R.X. Innovative Smart Scheduling and Predictive Maintenance Techniques. In *Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 181–207, ISBN 978-0-12-823657-4.
13. Rojek, I.; Jasiulewicz-Kaczmarek, M.; Piechowski, M.; Mikołajewski, D. An Artificial Intelligence Approach for Improving Maintenance to Supervise Machine Failures and Support Their Repair. *Appl. Sci.* **2023**, *13*, 4971. [[CrossRef](#)]
14. Kounta, C.A.K.A.; Kamsu-Foguem, B.; Nouredine, F.; Tangara, F. Multimodal Deep Learning for Predicting the Choice of Cut Parameters in the Milling Process. *Intell. Syst. Appl.* **2022**, *16*, 200112. [[CrossRef](#)]
15. Nasir, V.; Sassani, F. A Review on Deep Learning in Machining and Tool Monitoring: Methods, Opportunities, and Challenges. *Int. J. Adv. Manuf. Technol.* **2021**, *115*, 2683–2709. [[CrossRef](#)]
16. Sayyad, S.; Kumar, S.; Bongale, A.; Kamat, P.; Patil, S.; Kotecha, K. Data-Driven Remaining Useful Life Estimation for Milling Process: Sensors, Algorithms, Datasets, and Future Directions. *IEEE Access* **2021**, *9*, 110255–110286. [[CrossRef](#)]
17. Zhou, Y.; Xue, W. Review of Tool Condition Monitoring Methods in Milling Processes. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 2509–2523. [[CrossRef](#)]
18. Javaid, M.; Haleem, A.; Singh, R.P.; Suman, R. Artificial Intelligence Applications for Industry 4.0: A Literature-Based Study. *J. Ind. Intg. Mgmt.* **2022**, *07*, 83–111. [[CrossRef](#)]
19. Rajesh, A.S.; Prabhuswamy, M.S.; Krishnasamy, S. Smart Manufacturing through Machine Learning: A Review, Perspective, and Future Directions to the Machining Industry. *J. Eng.* **2022**, *2022*, 9735862. [[CrossRef](#)]
20. Zhang, M.; Matta, A. Models and Algorithms for Throughput Improvement Problem of Serial Production Lines via Downtime Reduction. *IIEE Trans.* **2020**, *52*, 1189–1203. [[CrossRef](#)]
21. Serin, G.; Sener, B.; Ozbayoglu, A.M.; Unver, H.O. Review of Tool Condition Monitoring in Machining and Opportunities for Deep Learning. *Int. J. Adv. Manuf. Technol.* **2020**, *109*, 953–974. [[CrossRef](#)]
22. Soo Lon Wah, W.; Chen, Y.-T.; Owen, J.S. A Regression-Based Damage Detection Method for Structures Subjected to Changing Environmental and Operational Conditions. *Eng. Struct.* **2021**, *228*, 111462. [[CrossRef](#)]
23. Corne, R.; Nath, C.; El Mansori, M.; Kurfess, T. Study of Spindle Power Data with Neural Network for Predicting Real-Time Tool Wear/Breakage during Inconel Drilling. *J. Manuf. Syst.* **2017**, *43*, 287–295. [[CrossRef](#)]
24. Jung, D.; Zhang, Z.; Winslett, M. Vibration Analysis for IoT Enabled Predictive Maintenance. In Proceedings of the 2017 IEEE 33rd International Conference on Data Engineering (ICDE), San Diego, CA, USA, 19–22 April 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1271–1282.
25. Patil, S.S.; Gaikwad, J.A. Vibration Analysis of Electrical Rotating Machines Using FFT: A Method of Predictive Maintenance. In Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1–6.
26. Patwardhan, A.; Verma, A.K.; Kumar, U. A Survey on Predictive Maintenance Through Big Data. In *Current Trends in Reliability, Availability, Maintainability and Safety*; Lecture Notes in Mechanical Engineering; Kumar, U., Ahmadi, A., Verma, A.K., Varde, P., Eds.; Springer: Cham, Switzerland, 2016; pp. 437–445, ISBN 978-3-319-23596-7.
27. Zhang, Q.; Yang, L.T.; Chen, Z.; Li, P. A Survey on Deep Learning for Big Data. *Inf. Fusion* **2018**, *42*, 146–157. [[CrossRef](#)]
28. Lei, Y.; Jia, F.; Lin, J.; Xing, S.; Ding, S.X. An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning Towards Mechanical Big Data. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3137–3147. [[CrossRef](#)]
29. Sinha, J.K.; Elbhah, K. A Future Possibility of Vibration Based Condition Monitoring of Rotating Machines. *Mech. Syst. Signal Process.* **2013**, *34*, 231–240. [[CrossRef](#)]
30. Khaleghi, B.; Khamis, A.; Karray, F.O.; Razavi, S.N. Multisensor Data Fusion: A Review of the State-of-the-Art. *Inf. Fusion* **2013**, *14*, 28–44. [[CrossRef](#)]
31. Al-Habaibeh, A.; Gindy, N. Self-Learning Algorithm for Automated Design of Condition Monitoring Systems for Milling Operations. *Int. J. Adv. Manuf. Technol.* **2001**, *18*, 448–459. [[CrossRef](#)]
32. Zeng, H.; Thoe, T.; Li, X.; Zhou, J. Multi-Modal Sensing for Machine Health Monitoring in High Speed Machining. In Proceedings of the 2006 IEEE International Conference on Industrial Informatics, Singapore, 16–18 August 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 1217–1222.

33. Hesser, D.F.; Markert, B. Tool Wear Monitoring of a Retrofitted CNC Milling Machine Using Artificial Neural Networks. *Manuf. Lett.* **2019**, *19*, 1–4. [[CrossRef](#)]
34. Lamraoui, M.; El Badaoui, M.; Guillet, F. Chatter Detection in CNC Milling Processes Based on Wiener-SVM Approach and Using Only Motor Current Signals. In *Vibration Engineering and Technology of Machinery; Mechanisms and Machine Science*; Sinha, J.K., Ed.; Springer: Cham, Switzerland, 2015; Volume 23, pp. 567–578, ISBN 978-3-319-09917-0.
35. Chen, G.S.; Zheng, Q.Z. Online Chatter Detection of the End Milling Based on Wavelet Packet Transform and Support Vector Machine Recursive Feature Elimination. *Int. J. Adv. Manuf. Technol.* **2018**, *95*, 775–784. [[CrossRef](#)]
36. Eski, İ. Vibration Analysis of Drilling Machine Using Proposed Artificial Neural Network Predictors. *J. Mech. Sci. Technol.* **2012**, *26*, 3037–3046. [[CrossRef](#)]
37. He, Z.; Shi, T.; Xuan, J. Milling tool wear prediction using multi-sensor feature fusion based on stacked sparse autoencoders. *Measurement* **2022**, *190*, 110719. [[CrossRef](#)]
38. Corne, R.; Nath, C.; Mansori, M.E.; Kurfess, T. Enhancing Spindle Power Data Application with Neural Network for Real-Time Tool Wear/Breakage Prediction During Inconel Drilling. *Procedia Manuf.* **2016**, *5*, 1–14. [[CrossRef](#)]
39. Qureshi, W.; Cura, F.; Mura, A. Prediction of Fretting Wear in Aero-Engine Spline Couplings Made of 42CrMo4. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2017**, *231*, 4684–4692. [[CrossRef](#)]
40. Bechouel, R.; Ben Moussa, N.; Terres, M.A. Improvement of the Predictive Ability of Polycyclic Fatigue Criteria for 42CrMo4 Nitrided Steels. In *Advances in Mechanical Engineering and Mechanics; Lecture Notes in Mechanical Engineering*; Benamara, A., Haddar, M., Tarek, B., Salah, M., Fakher, C., Eds.; Springer: Cham, Switzerland, 2019; pp. 212–220. ISBN 978-3-030-19780-3.
41. Vazdirvanidis, A.; Pantazopoulos, G.; Louvaris, A. Failure Analysis of a Hardened and Tempered Structural Steel (42CrMo₄) Bar for Automotive Applications. *Eng. Fail. Anal.* **2009**, *16*, 1033–1038. [[CrossRef](#)]
42. Hausnerova, B. Powder Injection Moulding—An Alternative Processing Method for Automotive Items. In *New Trends and Developments in Automotive System Engineering*; Chiaberge, M., Ed.; InTech: London, UK, 2011; ISBN 978-953-307-517-4.
43. Gómez, G.; De Lucio, P.F.; Del Olmo, A.; De Pissón, G.M.; Jimeno, A.; González, H.; De Lacalle, L.N.L. Comparison between milling roughing operations in full slotting manufacturing: Trochoidal, plunge and conventional milling. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Philadelphia, PA, USA, 2021; Volume 1193, No. 1.
44. Sckit Learn, “sklearn.svm.SVR”. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html> (accessed on 3 October 2023).
45. Sckit Learn, “sklearn.tree.DecisionTreeRegressor”. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html> (accessed on 3 October 2023).
46. Sckit Learn, “sklearn.neural_network.MLPRegressor”. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html#sklearn.neural_network.MLPRegressor (accessed on 3 October 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.