Lecture Notes in Mechanical Engineering

José Machado · Filomena Soares · Justyna Trojanowska · Sahin Yildirim · Jiří Vojtěšek · Pierluigi Rea · Bogdan Gramescu · Olena O. Hrybiuk *Editors*

Innovations in Mechatronics Engineering II

Springer

Lecture Notes in Mechanical Engineering

Lecture Notes in Mechanical Engineering (LNME) publishes the latest developments in Mechanical Engineering—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNME. Volumes published in LNME embrace all aspects, subfields and new challenges of mechanical engineering. Topics in the series include:

- Engineering Design
- Machinery and Machine Elements
- Mechanical Structures and Stress Analysis
- Automotive Engineering
- Engine Technology
- Aerospace Technology and Astronautics
- Nanotechnology and Microengineering
- Control, Robotics, Mechatronics
- MEMS
- Theoretical and Applied Mechanics
- Dynamical Systems, Control
- Fluid Mechanics
- Engineering Thermodynamics, Heat and Mass Transfer
- Manufacturing
- Precision Engineering, Instrumentation, Measurement
- Materials Engineering
- Tribology and Surface Technology

To submit a proposal or request further information, please contact the Springer Editor of your location:

China: Ms. Ella Zhang at ella.zhang@springer.com
India: Priya Vyas at priya.vyas@springer.com
Rest of Asia, Australia, New Zealand: Swati Meherishi at swati.meherishi@springer.com
All other countries: Dr. Leontina Di Cecco at Leontina.dicecco@springer.com

To submit a proposal for a monograph, please check our Springer Tracts in Mechanical Engineering at https://link.springer.com/bookseries/11693 or contact Leontina.dicecco@springer.com

Indexed by SCOPUS. All books published in the series are submitted for consideration in Web of Science.

More information about this series at https://link.springer.com/bookseries/11236

José Machado · Filomena Soares ·
Justyna Trojanowska · Sahin Yildirim ·
Jiří Vojtěšek · Pierluigi Rea ·
Bogdan Gramescu · Olena O. Hrybiuk
Editors

Innovations in Mechatronics Engineering II

Springer

Editors
José Machado
Department of Mechanical Engineering
University of Minho
Guimraes, Portugal

Filomena Soares
Department of Industrial Electronics
University of Minho
Guimraes, Portugal

Justyna Trojanowska
Poznan University of Technology
Poznan, Wielkopolskie, Poland

Sahin Yildirim
Department of Mechatronics Engineering,
Faculty of Engineering
Erciyes University
Kayseri, Turkey

Jiří Vojtěšek
Faculty of Applied Informatics
Tomas Bata University in Zlín
Zlín, Czech Republic

Pierluigi Rea
Department of Mechanical, Chemical
and Materials Engineering
University of Cagliari
Cagliari, Italy

Bogdan Gramescu
Department of Mechatronics
and Precision Mechanics
Polytechnic University of Bucharest
Bucharest, Romania

Olena O. Hrybiuk
Institute of Information Technology
and Learning Tools
National Academy of Educational Sciences
Kyiv, Ukraine

ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-3-031-09384-5 ISBN 978-3-031-09385-2 (eBook)
https://doi.org/10.1007/978-3-031-09385-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume of Lecture Notes in Mechanical Engineering gathers selected papers presented at the Second International Scientific Conference (ICIE'2022), held in Guimarães, Portugal, on 28–30 June 2022. The conference was organized by the School of Engineering of University of Minho, throughout MEtRICs and Algoritmi Research Centres.

The aim of the conference was to present the latest engineering achievements and innovations and to provide a chance for exchanging views and opinions concerning the creation of added value for the industry and for the society. The main conference topics include (but are not limited to):

- Innovation
- Industrial engineering
- Mechanical engineering
- Mechatronics engineering
- Systems and applications
- Societal challenges
- Industrial property

The organizers received 139 contributions from 16 countries around the world.

After a thorough peer review process, the committee accepted 81 papers written by 335 authors from 15 countries for the conference proceedings (acceptance rate of 58%), which were organized in three volumes of Springer Lecture Notes in Mechanical Engineering.

This volume, with the title "Innovations in Mechatronics Engineering II", specifically focuses on cutting-edge control algorithms for mobile robots, automatic monitoring systems and intelligent predictive maintenance techniques. They cover advanced scheduling, risk-assessment and decision-making strategies, and their applications in industrial production, training and education, and service organizations. Last but not least, it analyses important issues proposing a good balance of theoretical and practical aspects. This book consists of 26 chapters, prepared by 93 authors from 8 countries.

Extended versions of selected best papers from the conference will be published in the following journals: Sensors, Applied Sciences, Machines, Management and Production Engineering Review, International Journal of Mechatronics and Applied Mechanics, SN Applied Sciences, Dirección y Organización, Smart Science, Business Systems Research and International Journal of E-Services and Mobile Applications.

A special thanks to the members of the International Scientific Committee for their hard work during the review process.

We acknowledge all that contributed to the staging of ICIE’2022: authors, committees and sponsors. Their involvement and hard work were crucial to the success of ICIE’2022.

June 2022

José Machado
Filomena Soares
Justyna Trojanowska
Sahin Yildirim
Jiří Vojtěšek
Pierluigi Rea
Bogdan Gramescu
Olena O. Hrybiuk

Contents

Machine Multi-sensor System and Signal Processing for Determining Cutting Tools Service Life

Edward Kozłowski[1], Katarzyna Antosz[2](✉), Dariusz Mazurkiewicz[3], Jarosław Sęp[2], and Tomasz Żabiński[4]

[1] Faculty of Management, Lublin University of Technology, Nadbystrzycka 38, 20–618 Lublin, Poland
e.kozlovski@pollub.pl
[2] Faculty of Mechanical Engineering and Aeronautics, Rzeszow University of Technology, Powstańców Warszawy 8, 35-959 Rzeszów, Poland
{katarzyna.antosz,jsztmiop}@prz.edu.pl
[3] Mechanical Engineering Faculty, Lublin University of Technology, Nadbystrzycka 36, 20–618 Lublin, Poland
d.mazurkiewicz@pollub.pl
[4] Faculty of Electrical and Computer Engineering, Rzeszow University of Technology, W. Pola 2, 35–959 Rzeszów, Poland
tomz@prz.edu.pl

Abstract. The constantly growing data resources are a challenge for manufacturing companies in every industry. This is due, inter alia, to a significant increase in the number of devices that generate data. Currently, the concept of Indus-try 4.0 and related technologies facilitate the collection, processing, and use of large amounts of data. For this purpose, the possibility use of the data mining method (wavelet analysis and logistic regression) to develop the model for supporting the decision-making process in determining the service life of the cutting tool was discussed in this article. The developed model will support to identify the parameters influencing the condition of the cutter. The predictive ability of the obtained model was assessed with the use of indicators to assess the quality of the classification.

Keywords: Cutter condition · Wavelet analysis · Logistic regression quality prediction

1 Introduction

Cutting tool is one of the core elements of the machine tool, which condition is significant bot for the production process efficiency and reliability of the technological machine. Machine tool is also the element which is most often relaced due to its wear, which makes it an important cost-creating element of production. On the other hand, machine tools together with other technological machines used typically in production systems of high technology industry form complex systems functioning as Industry 4.0 elements.

J. Machado et al. (Eds.): icieng 2022, LNME, pp. 1–12, 2022.
https://doi.org/10.1007/978-3-031-09385-2_1

All this creates a lot of several technological and research challenges, for example [1, 2], related to their service life, health monitoring and reliability, especially with respect to predicting future states to enable inference and implementation of executive activities in terms of failure-preventing servicing. Current development of condition monitoring technology and machine learning, makes available more and more fault diagnosis methods, mostly related to signal processing and deep learning technology [3]. For example, for the purpose of cutter reliability decision-making process, it was applied [1, 2] a multi sensor measurement system integrated with signal processing. Signal processing-based fault diagnosis or remaining useful life prediction is based on feature extraction and pattern recognition. Feature extraction based on wavelet transform makes possible signal decomposition, while advanced statistical analysis results in fault features. On the other hand, Support Vector Machine (SVM) can be successfully applied as a classifier in pattern recognition [4]. All these methods usually work well when analysed theoretically, with the use of simulation data or with limited access to data [5–8]. There is a very limited feedback on their results when applied to real data obtained from industrial measurement systems. New researches with real-world data for verification of advanced data processing methods are necessary, especially due to the Industry 4.0 vision. In this vision data digitalization and data processing is expected to bring major changes in manufacturing. Novel technologies will enable a stepwise increase of productivity in manufacturing companies. From this perspective, described in this article milling machine multi-sensor system with proposed and verified data processing model may be considered as a decision-making process tool in determining cutting tools service life, extending the time of their effective use in the production process. Effective prediction of cutting tool state and its remaining useful life makes its replacement time as optimal as possible. The structure of the paper is as follows. Sections 2 presents the research problem and methodology. The structure of the paper is as follows. Section 3 presents a short literature review on materials and methods used in the research. In Sect. 4 the results and discussion are presented. Finally (Sect. 5), the conclusion are presented.

2 Problem Formulation and Research Methodology

The main goal of the presented research was to develop a predictive model that would allow to identify the state of the cutting tool (milling cutter) edge. To develop the model, the data obtained from the sensor installed on the technological machine (force and torque signal reading) were used. The collected sensor signals were preprocessed by application Discrete Wavelet Transformation (DWT). In the analyzed case, the identification consisted in determining whether the cutter is sharp or blunt. The obtained value of the variable is qualitative, which is related to the classification task. Many classifiers can be used to analyze a classification problem. In this study, logistic regression were used.

3 Materials and Methods

3.1 Data Collection- Test Stand

The tests were carried out on a specially organized test stand. Detailed elements of the test stand, their characteristics and data collection methods are presented in Table 1.

Table 1. The test stand – elements and parameters

No	Element	Parameters
1	Machine	Haas VM-3 CNC machine with 12.000 RPM direct drive spindle
2	Milling cutter	Four-blade solid carbide with TiAln (aluminum titanium nitride) coating. The rotational speed of the spindle during experiments – 860 rpm
3	Material	Inconel 718
4	Multi – component sensor	Type - CL16 ZEPWN. Parameters: force measurement – max 10 kN, torque measurement – max 1 kNm, the accuracy - 0.5, the sensitivity – 1 mV/V Collected signals: force sensor (P1x, P2y, P3z) and moments (M1x, M2y, M3z)
5	Data collecting system	Industrial computer (Beckhoff C6920), system: EtherCAT-based distributed I/O, real-time task developed using ST (Structured Text) language - environment: the TwinCAT 3 Beckhoff and custom Matlab/Simulink projects

Data was collected in real time with a sampling interval of 2 ms. The sampling interval of the real-time data collection job was 2 ms. The duration of the signal buffer stored in one file was 640 m/s. During the tests, data was collected from various tasks in the milling process on the machine.

3.2 Discrete Wavelet Transformation (DWT)

To design a classifier it was analyzed the signals obtained from the force sensors (P1x, P2y, P3z) and moment sensors (M1x, M2y, M3z), which were preprocessed. To extract the necessary information from these signal it was used the wavelet transformation. In comparison to harmonic analysis in wavelet analysis we utilize the wavelet functions, which usually are irregular, asymmetric, and not periodic. Based on wavelet transformation the signal is decomposed into components, which created as a wavelets with a different scale (scale/frequency) and position (time/space) [9, 10]. For a certain scale and position the wavelet coefficients describe the extent of similarity of wavelet function to the fragment of analyzed signal. By applying the wavelet transform we estimate the coefficients for wavelets of various scales and positions. These coefficients provide us the essential information about signal.

Let $\mathbb{N}$ denotes the set of natural numbers, $\mathbb{R}$ - set of real numbers, $\mathbb{Z}$ - set of integer numbers. Niech $\{x_t\}_{t\in\mathbb{Z}}$ will be a time series, $\Psi(t)$ - the orthogonal wavelet basis (mother wavelet) and $\phi(t)$ - the scaling function (father wavelet) corresponding to walvet $\Psi(t)$. For decomposition level $j \in \mathbb{N}$ we define the sequence of mother wavelets $\left\{\Psi_{jk}\right\}_{k\in\mathbb{Z}}$ as follows:

$$\Psi_{jk}(t) = \frac{1}{2^{j-1}}\Psi\left(\frac{t}{2^j} - k\right) \tag{1}$$

and the sequence of father wavelets $\left\{\phi_{jk}\right\}_{k\in\mathbb{Z}}$ as follows:

$$\phi_{jk}(t) = \frac{1}{2^{j-1}}\phi\left(\frac{t}{2^j} - k\right) \tag{2}$$

Thus the time series $\{x_t\}_{t\in\mathbb{Z}}$ can be expressed as:

$$x_t = \sum_{k=-\infty}^{\infty} c_{jk}\phi_{jk}(t) + \sum_{i=-\infty}^{j}\sum_{k=-\infty}^{\infty} d_{ik}\Psi_{ik}(t) \tag{3}$$

where c_{jk} is scaling coefficient, d_{ik} is detailed coefficient.

For Haar decomposition [11, 12] the mother wavelet has non zero values on interval $[0,\ 1)$. Because the observed signal (time series)$\{x_t\}_{1\le t\le n}$ has a finite number of observations then the level j should meet the condition $1 \le j \le m = \max\{s \in \mathbb{N} : 2^s \le n\}$. For simplicity, we assume that $n = 2^s$.

From above the time series $\{x_t\}_{1\le t\le n}$ we can present as follows:

$$x_t = \sum_{k=0}^{\frac{n}{2^j}-1} c_{jk}\phi_{jk}(t) + \sum_{i=0}^{j}\sum_{k=0}^{\frac{n}{2^i}-1} d_{ik}\Psi_{ik}(t) \tag{4}$$

for $1 \le j \le m$. From (4) for the sake of level j, $1 \le j \le m$ the time series $\{x_t\}_{t\in\mathbb{Z}}$ can be presented in different forms. According to [8], we define the time series projection operator $\{x_t\}_{1\le t\le n}$ for level j in the base $\left\{\phi_{jk}(t)\right\}_{0\le k\le\frac{n}{2^j}-1}$

$$P^j x_t = \sum_{k=0}^{\frac{n}{2^j}-1} c_{jk}\phi_{jk}(t) \tag{5}$$

In further part of paper the sequence $\left\{c_{jk}\right\}_k$ will be depict the information contained in observed signal $\{x_t\}_{1\le t\le n}$. More about DWT can be found in [11, 12].

3.3 Logistic Regression

Our aim is to recognize a cutter state based on strict or preprocessed measurements obtained from transducers obtained from transducers, sensors. To classify the cutter state we apply the logistic regression. Let $D = \{\mathrm{Y}, \mathrm{X}\}$ be a data set, where vector $\mathrm{Y} \in R^n$ contains the realizations of response variable, but matrix $\mathrm{X} \in R^{n\times m}$ contains the realizations of predictors (e.g. series of input variables) and:

$$\mathrm{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathrm{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} = \begin{bmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(n)} \end{bmatrix} \tag{6}$$

For $i-$ th case when the cutter is blunt then we accept $y_i = 1$ otherwise when the cutter is sharp then we take $y_i = 0$. Thus by observing the values of input variables

$x_{(i)} = (x_{i1}, x_{i2}, \ldots, x_{im}) \in R^m$ it is necessary to classify the cutter state. The task consists in definition a classifier $f : R^m \to \{0, 1\}$, which would allow to assessment the cutter based on observation of predictors.

Let (Ω, F, P) be a probability space and $Y : \Omega \to \{0, 1\}$ be a random variable with binomial distribution (see e.g. [12, 13]). Logistic regression models are applied usually to estimate the probability of realization of response variable (see e.g. [13, 15, 16]). $P(Y = 1|x)$ value denotes the success probability based on observation of predictors $x \in R^m$, $P(Y = 0|x)$ value denotes the defeat probability, but the ratio of success to defeat:

$$\Theta(x) = \frac{P(Y = 1|x)}{P(Y = 0|x)} = \frac{P(Y = 1|x)}{1 - P(Y = 1|x)} \tag{7}$$

means the odds. The success probability $p(x) \in (0, 1)$, thus from (7) we have $\Theta(x) \in (0, \infty)$ but $ln(\Theta(x)) \in (-\infty, \infty)$. The logarithm of odds is called log-odds or logit. Logistic regression describes the dependencies between logit and predictors as follows:

$$ln\Theta(x) = ln\left(\frac{p(\beta, x)}{1 - p(\beta, x)}\right) = \beta_0 + x^T\beta \tag{8}$$

where $\beta_0 \in R$, $\beta = (\beta_1, \ldots, \beta_m) \in R^m$. The aim of logistic regression consist in determining the probability of success $P(Y = 1|x) = p(\beta_0, \beta, x)$ based on observation of predictors $x \in R^m$. From Eqs. (7)–(8) we can determine the success probability as follows:

$$p(\beta_0, \beta, x) = \frac{e^{\beta_0 + x^T\beta}}{1 + e^{\beta_0 + x^T\beta}} \tag{9}$$

To estimate the unknown parameters $(\beta_0, \beta) \in R^{m+1}$ we apply the maximum likelihood method. We define the likelihood function as a product of probabilities of successes and from (9) we have:

$$L(\beta_0, \beta, \mathrm{Y}, \mathrm{X}) = \prod_{i=1}^{n} p\big(\beta_0, \beta, x_{(i)}\big)^{y_i}\big(1 - p\big(\beta_0, \beta, x_{(i)}\big)\big)^{1-y_i} \tag{10}$$

Maximum likelihood method consists in maximizing the objective function (10) and by solving the task

$$\max_{\beta_0, \beta} L(\beta_0, \beta, \mathrm{Y}, \mathrm{X}) \tag{11}$$

we obtain the estimators of unknown parameters β_0 and β. Usually we solve the auxiliary task

$$\max_{\beta_0, \beta} \ln L(\beta_0, \beta, \mathrm{Y}, \mathrm{X}) \tag{12}$$

instead the task (11), where the logarithm of likelihood function is equal:

$$\ln L(\beta_0, \beta, \mathrm{Y}, \mathrm{X}) = \sum_{i=1}^{n}\Big(y_i\Big(\beta_0 + x_{(i)}^T\beta\Big) - ln\Big(1 + e^{\beta_0 + x_{(i)}^T\beta}\Big)\Big). \tag{13}$$

To solve the auxiliary task (12) Newton-Raphson algorithm was applied.

3.4 Elasticnet

Sometimes the predictors used in linear system (8) are correlated. Then the direct application Newton-Raphson algorithm to solve the task (12) does not provide the expected effect. Due to multicollinearity problem of predictors we obtain the large absolute values of estimators of unknown parameters β_0 andβ. This implies the instability of forecasts. Thus the problem depends on the selection of appropriate predictors or transformation of predictors, which should be included to logistic regression (8). On the one hand, the input variables should influence the value of response variable, on the other they should not generate multicollinearity.

To overcome the multicollinearity problem usually we apply such techniques as singular value decomposition, regularization, least angle regression.

In our task from one hand to reduce the absolute values of estimator of unknown parameters and from the other to select the importance predictors to model (8) we apply the elasticnet method (see e.g. [14, 16, 17]), which consists in the addition of penalty depended on values of estimators of unknown parameters into objective function. Such techniques implies a shrinkage of values of estimators. From above we solve the task:

$$\max_{\beta_0,\beta} \ln L(\beta_0, \beta, \mathrm{Y}, \mathrm{X}) - \lambda P_\alpha(\beta) \tag{14}$$

where $\lambda > 0$ and the penalty $P_\alpha(\beta)$ is defined as linear combination of norms in spaces L_1 and L_2 of β values. Thus the penalty can be presented as follows:

$$P_\alpha(\beta) = \frac{1-\alpha}{2}\|\beta\|_{L_2} + \alpha\|\beta\|_{L_1} \tag{15}$$

For $\alpha = 0$ we have a classical Tikhonov regularization (ridge regression), but for $= 1$ the Least Absolute Shrinkage and Selection Operator (LASSO). The elastic net is α a connection between ridge regression and LASSO. The application of elastic net method allowed to receive classifier based on logistic regression (8) with more accurate and stable detection of cutter state.

3.5 The Quality Assessment of Prediction Models - Receiver Operating Characteristics (ROC) Analysis

For the quality assessment of the developed prediction models Receiver Operating Characteristics (ROC) analysis were used. ROC analysis is one of the most important methods used in the evaluation of the models in machine learning. This analysis uses the confusion matrix (YP – True positive, TN – true negative, FP - false positive, FN – false negative) assessing the accuracy of the developed classifier. The errors of the developed classifier are defined as: FP and FN. The classifier quality assessment is carried out by assessing whether the objects have been properly classified from positive to negative class and vice versa [18–22]. In the paper for evaluation the performance of the developed models the indicators presented in the Table 2 were used.

To calculate the indicators for all predicted models the confusion matrices were generated. The following assumptions were made: the sharp cutter as a negative case (N) and the blunt cutter as a positive case (P). The values specified the confusion matrix

Table 2. The quality assessment of prediction model - indicators

Indicator	Description	Equation
Acc - Accuracy	Properly classified cases (positive and negative) to the total number of predictions (ability of prediction model)	$Acc = \frac{TP+TN}{TP+TN+FP+FN}$
TPR - True Positives Rate	True positive cases classified to all true positive and false negative cases classified	$TPR = \frac{TP}{TP+FN}$
TNR - True Negatives Rate	True negative cases classified to all true negative and false positive cases classified	$TNR = \frac{TN}{TN+FP}$
PPV - Positive Predictive Value	True positive cases classified to all true and false positive cases classified	$PPV = \frac{TP}{TP+FP}$
NPV - Negative Predictive Value	True negative cases classified to all cases classified as true and false negative	$NPV = \frac{TN}{TN+FN}$
PV - Prevalence	True positive and false negative cases classified to the total number of predictions	$PV = \frac{TP+FN}{TP+TN+FP+FN}$
DR – Detection Rate	True positive cases classified to the total number of predictions	$DR = \frac{TP}{TP+TN+FP+FN}$
DPV – Detection Prevalence	All positive classified (true and false) cases to the total number of predictions	$DPV = \frac{TP+FP}{TP+TN+FP+FN}$

are as follows: TP (True Positive), which means a number of cases for which the cutter state was properly recognized (as blunt), TN (True Negative) - a number of cases for which the cutter state was properly recognized (as sharp), FP (False Positive) - a number of cases for which the cutter state was not properly recognized (sharp instead of blunt), FN (False Negative) - a number of cases for which the cutter condition was not properly recognized (blunt instead of sharp).

4 Results and Discussion

The main aim of the research was to develop a model that would allow to recognize the cutter state (sharp or blunt). During the research, carried out on the test stand (Table 1), 2172 observations of the following signals: P1x, P2y, P3z, M1x, M2y and M3z were obtained. Then, the obtained data with using of wavelet analysis were pre-processed. The wavelets of various types and levels for preprocessing were used. In the Table 3 the used types and levels of wavelets are presented.

Table 3. The used types and levels of wavelets.

Type	Level		
	l = 3	l = 4	l = 5
Daubechies (d)	–	d12, d14, d16, d18, d20	d2, d4, d6, d8, d10
Least asymmetric (la)	–	la12, la14, la16, la18, la20	la8, la10
Best localized (bl)	–	bl14, bl18, bl20	–
Coiflet (c)	c24, c30	c12, c18	c6

Table 4. The quality of prediction models - the indicators values.

Wavelet	Level	Acc	TPR	TNR	PPV	NPV	PV	DR	DPV
d2	5	0.8964	0.7599	**1.0000**	**1.0000**	0.8459	0.4314	0.3278	0.3278
d10	5	0.8973	0.7631	0.9992	0.9986	0.8475	0.4314	0.3292	0.3297
d18	4	0.8959	0.7599	0.9992	0.9986	0.8458	0.4314	0.3278	0.3283
d20	4	0.8969	0.7620	0.9992	0.9986	0.8469	0.4314	0.3287	0.3292
la8	5	0.8978	0.7641	0.9992	0.9986	0.8481	0.4314	0.3297	0.3301
la10	5	0.8969	0.7620	0.9992	0.9986	0.8469	0.4314	0.3287	0.3292
la12	4	0.8964	0.7609	0.9992	0.9986	0.8464	0.4314	0.3283	0.3287
bl14	4	0.8973	0.7631	0.9992	0.9986	0.8475	0.4314	0.3292	0.3297
bl18	4	0.8950	0.7577	0.9992	0.9986	0.8446	0.4314	0.3269	0.3273
bl20	4	0.8955	0.7588	0.9992	0.9986	0.8452	0.4314	0.3273	0.3278
c6	5	**0.8983**	0.7652	0.9992	0.9986	0.8487	0.4314	0.3301	0.3306
c12	4	0.8950	0.7577	0.9992	0.9986	0.8446	0.4314	0.3269	0.3273
c18	4	0.8969	0.7620	0.9992	0.9986	0.8469	0.4314	0.3287	0.3292
c24	3	0.8955	0.7577	**1.0000**	**1.0000**	0.8447	0.4314	0.3269	0.3269
c30	3	0.8955	0.7577	**1.0000**	**1.0000**	0.8447	0.4314	0.3269	0.3269

Legend: Prediction model with the highest ACC value; Prediction models with the highest TNR and PPV value

To develop the predictive models for all wavelets logistic regression (8) and elastic net were used. The use of elastnic net allowed to receive classifier based on logistic regression with more accurate and stable detection of cutter state. For the prepared base, the learning time did not exceed 0.05 s and the tool condition detection time does not exceed 0.00003s.. The indicators presented in Table 2 were used to assess the quality of the developed models. In the Table 4 the obtained values of indicators for chosen developed prediction models are presented.

The results presented in Table 4 show that the prediction ability of the developed models for individual wavelets is comparable. The value of the Accuracy indicator is in the range from 0.8983 to 0.8950. The highest value of Acc was obtained for wavelets c6 at the level l = 5, and the lowest value for wavelets la16, bl18, c12 (all at the same level l = 4). Furthermore, the highest recognition of negative class (N - blunt) (TNR = 1) is obtained by the prediction models for the wavelets d2 level l = 5, c24 and c30 level l = 3. For the same models positive predictive value is also the highest (PPV = 1). Additionally, the detection rate for true positive class (P - sharp) (DR) and predicted positive cases (DPV) were the highest for the wavelet c6 level l = 3, the values 0.3301 and 0.3306 respectively. The value of the PV indicator for all wavelets was equal 0.4314.

Table 5 presents the confusion matrix for the model with the highest Acc value (wavelet c6 at the level l = 5). Analyzing the results presented in the Table 5, it should be noted that the highest impact on the obtained value of the Acc indicator have the number of cases for which the cutter state was not properly recognized (FN (False Negative) - sharp instead of blunt).

Table 5. Confusion matrix for prediction model with the highest Acc value.

	Reference		
	State	Blunt	Sharp
Prediction	Blunt	717	1
	Sharp	220	1234

Moreover, 221 cases from 2172 were not properly recognized, which means that the prediction error in analyzing model is ca. 10% (1- Acc*100%).

Furthermore, the value of the FN affects the True Positive Rate and Negative Predictive Value indicators too. Therefore, for the developed predictive models, the value of these indicators were also comparable. TPR indicator was in the range from 0.7577 to 0.7652 and NPV indicator was in the range 0.4787 to 0.8446 (Table 4). In the case of high ability of the model these values should be close to the 1.00.

An additional element of the analysis of the developed predictive models was the influence of the regularization parameters (α and λ) on the Acc indicator value. On the Fig. 1 the influence of the regularization parameters (α and λ) for the model with highest Acc value (wavelet c6 l = 5) is presented. On the other hand, on the Fig. 2 the influence of the regularization parameters (α and λ) for the d2 level l = 5 (Acc = 0.8964) is presented. This prediction model is the one of the models with the highest TNR and PPV value.

When analyzing the graphs presented in Figs. 1 and 2, it should be noted that for the given α and λ parameters, the distribution of the Acc values obtained for the models differ significantly. In the model developed for the c6 wavelet, the Acc values close to the maximum Acc value are obtained for a wider range of parameters α and λ. This model is more stable (Fig. 1). It is different in the case of the db2 wavelet model. Only for selected values of α and λ parameter the Acc values are close to the max Acc value.

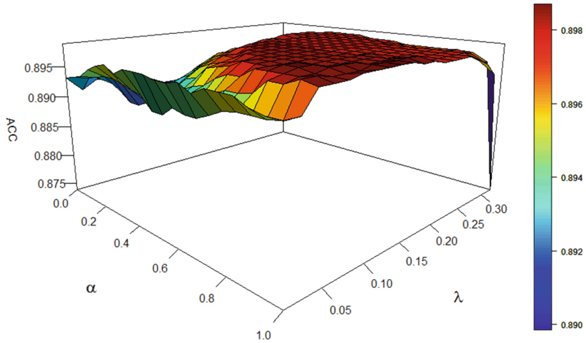

Fig. 1. The influence of the regularization parameters (α and λ) for the model with highest Acc (wavelet c6 level l = 5).

The obtained Acc values for this model are changing abruptly - the model is less stable (Fig. 2).

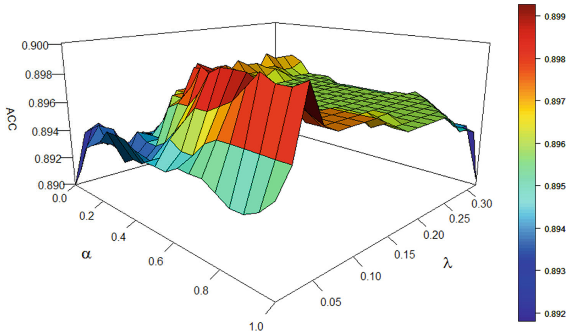

Fig. 2. The influence of the regularization parameters (α and λ) for the chosen model with highest TNR and PPV value (d2 level l = 5) is presented.

For the analyzed models presented in Figs. 1 and 2, the optimal values of the parameters α and λ were determined. For the wavelet c6 level $l = 5$ the values $\alpha = 0.7895$ and $\lambda = 0.0753$ were estimated. For a wavelet d2 level $l = 5$ they were $\alpha = 0.8421$ and $\lambda = 0.0426$. With these values, the analyzed models obtained the highest Acc value, and thus the lowest prediction error.

5 Conclusions

Currently, many manufacturing companies face the problem of huge amounts of data. These data are generated, inter alia, by various types of devices that carry out operations in the technological process or also by production process monitoring systems. Properly conducted analysis of production data can reveal important information useful for predicting models in manufacturing. The main goal of the presented research was to develop a predictive models that would allow to identify the state of the cutter. To develop the model, the data obtained from the sensors installed on the technological machine were used. During the research the model with different wavelets have been developed. For developed models the Accuracy indicator is in the range from 0.8983 to 0.8950 was obtained. The highest value of Acc was obtained for wavelets c6 and the lowest value for wave-lets la16, bl18, c12. The presented models allow to identify in the easy way the condition of the cutter and thus will reduce the number of their use in the manufacturing process. The developed model can be used on the other machines equipped with the same type of sensors. For the other machines with other different types of sensors the model can be developed according to proposed methodology.

References

1. Kozłowski, E., Antosz, K., Mazurkiewicz, D., Sęp, J., Żabiński, T.: Integrating advanced measurement and signal processing for reliability decision-making. Eksploatacja i Niezawodnosc – Mainten. Reliab. **23**(4) 777–787 (2021)
2. Antosz, K., Mazurkiewicz, D., Kozłowski, E., Sęp, J., Żabiński, T.: Machining process time series data analysis with a decision support tool. In: Machado, J., Soares, F., Trojanowska, J., Ottaviano, E. (eds.) icieng 2021. LNME, pp. 14–27. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-79165-0_2
3. Yao, D.C., Liu, H.C., Yang, J.W., Li, X.: A lightweight neural network with strong robustness for bearing fault diagnosis. Measurement **159**, 1–11 (2020)
4. Kozłowski, E., Mazurkiewicz, D., Żabiński, T., Prucnal, S., Sęp, J.: Machining sensor data management for operation-level predictive model. Expert Syst. Appl. **159**, 1–22 (2020)
5. Li, H., Wang, W., Li, Z., Dong, L., Li, Q.: A novel approach for predicting tool remaining useful life using limited data. Mech Syst Sig Process **143**, 106832 (2020). https://doi.org/10.1016/j.ymssp.2020.106832
6. Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., Gao, R.X.: Deep learning and its applications to machine health monitoring. Mech. Syst. Signal Process. **115**, 213–237 (2019)
7. Arrazola, P., Özel, T., Umbrello, D., Davies, M., Jawahir, I.: Recent advances in modelling of metal machining processes. CIRP Ann. **62**, 695–718 (2013)
8. Arrais-Castro, A., Varela, M.L.R., Putnik, G.D., Ribeiro, R.A., Machado, J., Ferreira, L.: Collaborative framework for virtual organisation synthesis based on a dynamic multi-criteria decision model. Int. J. Comput. Integr. Manuf. **31**(9), 857–868 (2018)
9. Daubechies, I.: Orthonormal bases of compactly supported wavelets. Commun. Pure Appl. Math. **41**, 909–996 (1988)
10. Edwards, T.: Discrete Wavelets Transform: Theory and Implementation, Stanford University (1991)
11. Daubechies, I.: Ten lectures on wavelets. Society for Industrial and Applied Mathematics (1992). https://doi.org/10.1137/1.9781611970104
12. Walnut, D.F.: An Introduction to Wavelet Analysis. Springer, Boston (2004)
13. Freedman, D.A.: Statistical Models: Theory And Practice. Cambridge University Press (2009)
14. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, New York Inc. (2009)
15. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning. Springer, New York (2013)
16. Fox, J., Weisberg, S.: An R Companion to Applied Regression. SAGE Publications, Inc. (2019)

17. Rymarczyk, T., Kozłowski, E., Kłosowski, G., Niderla, K., Logistic regression for machine learning in process tomography. Sensors **19**(15), 1–19 (2019)
18. Fawcett, T.: An introduction to ROC analysis. Pattern Recogn. Lett. **27**, 861–874 (2006). https://doi.org/10.1016/j.patrec.2005.10.010
19. Fawcett, T.: Using rule sets to maximize ROC performance. In: Proceedings of the IEEE International Conference on Data Mining (ICDM-2001), pp. 131–138 IEEE (2001)
20. Provost, F., Fawcett, T., Kohavi, R.: The case against accuracy estimation for comparing classifiers. In: Proceedings of the ICML-98. pp.445–453. Morgan Kaufmann, San Francisco (1998)
21. Powers, D.: Evaluation: from precision, recall and F-score to ROC, unforcedness, nakedness & correlation. J. Mach. Learn. Technol. **2**, 37–63 (2011)
22. Antosz, K., Jasiulewicz-Kaczmarek, M., Paśko, Ł., Zhang, C., Wang, S.: Application of machine learning and rough set theory in lean maintenance decision support system development. Eksploatacja i Niezawodnosc – Maint. Reliab. **23**(4), 695–708 (2021)

Automatic Anomaly Detection in Vibration Analysis Based on Machine Learning Algorithms

Pedro Torres[1,2](✉), Armando Ramalho[1,3], and Luis Correia[1]

[1] Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal
{pedrotorres,aramalho,lcorreia}@ipcb.pt
[2] SYSTEC - Research Center for Systems and Technologies, Porto, Portugal
[3] CEMMPRE - Centre for Mechanical Engineering, Materials and Processes, Coimbra, Portugal

Abstract. This paper presents an approach for automatic anomaly detection through vibration analysis based on machine learning algorithms. The study focuses on induction motors in a predictive maintenance context, but can be applied to other domains. Vibration analysis is an important diagnostic tool in industrial data analysis to predict anomalies caused by equipment defects or in its use, allowing to increase its lifetime. It is not a new technique and is widely used in the industry, however with the Industry 4.0 paradigm and the need to digitize any process, it gains relevance to automatic fault detection. The *Isolation Forest* algorithm is implemented to detect anomalies in vibration datasets measured in an experimental apparatus composed of an induction motor and a coupling system with shaft alignment/misalignment capabilities. The results show that it is possible to detect anomalies automatically with a high level of precision and accuracy.

Keywords: Industry 4.0 · Anomaly detection · Isolation forest · Vibration analysis · BigML

1 Introduction

According to the Europe Predictive Maintenance Market - Industry Trends and Forecast to 2027 report [1], the predictive maintenance market is expected to growing market with a CAGR (Compound Annual Growth Rate) of 39.6% in the forecast period 2020 to 2027. The increased use of new and emerging technologies to gain valuable insights into decision making has contributed to the growth of the predictive maintenance market. Several vertical end users are increasingly in need of cost reduction and downtime, which has spurred the growth of the predictive maintenance market.

Industry 4.0 marks the revolution of digitization of traditional manufacturing industries supported by modern technologies such as automation, interconnectivity, real-time data processing and intelligence based on machine learning

J. Machado et al. (Eds.): icieng 2022, LNME, pp. 13–23, 2022.
https://doi.org/10.1007/978-3-031-09385-2_2

techniques. With the growth of automation technologies, the sensing component is the basis of perception, fundamental to the smart factory concept. Condition monitoring techniques are maintenance methodologies to monitor the operating conditions of an equipment in real time, through measurement and extraction of information that allow understanding its health, wear, degradation or significant changes in operation. The collected data is used to find trends, predict failures and estimate the remaining lifetime of an asset.

Through vibration analysis, by analysing the frequency, amplitude, phase, position and direction of vibrations in machinery, it is possible to identify many common faults. For example it is possible differentiate between wear on a specific gear or bearing, a lack of lubrication, an imbalance, a misalignment, a loose mounting or an electrical fault. Fault detection can be carried out before a machine is stopped, reducing downtime to its bare minimum. Early detection and predictive maintenance can also prevent more serious faults from developing.

Machine learning algorithms have the ability to analyze large amounts of data, and automatically perform diagnoses, without human intervention, based on historic and correlations with failure situations, but also by self-learning. Increasingly, they are the appropriate tool for decision making with high levels of accuracy.

The Fig. 1 shows a generalist architecture supported in the context of Industry 4.0, which illustrates the overview of fault detection systems in electrical machines based on vibration analysis and on which we base this work. The Operational Technology (OT) and Information Technology (IT) parts were aligned to design an Industrial Control System (ICS) in the laboratory, for acquiring, controlling and monitoring the operating status of rotating machines, producing reports, automatic alerts and recommending actions to take as a prescriptive maintenance system.

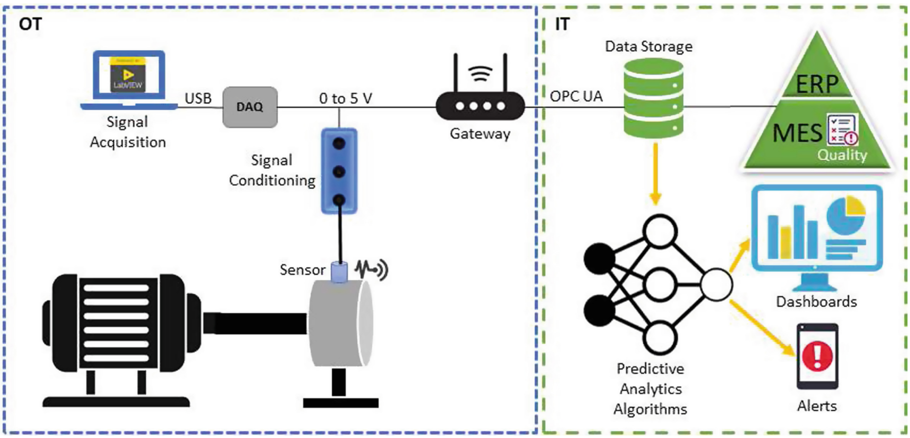

Fig. 1. Architecture overview for data acquisition and its interconnection with high-level industry management systems.

The remainder of this paper is organized as follows: Sect. 2 describes the state of the art and related work. Section 3 describes the materials and methods for the automatic anomaly detection. Section 4 presents the experimental results achieved and, finally, Sect. 5 presents the conclusions and future work.

2 Related Work

This section presents some scientific works related to vibration analysis in rotating machines using different machine learning techniques that help to support the validity of the work of this paper and its importance in the current context of the industry. Any vibration measurement experiment for diagnosing machine operation must be in line with ISO 22096:2007.

Vibration signals carry very important information for predictive maintenance applications, which is why it is widely used. The paper [2] presents and describes some condition maintenance techniques in a predictive maintenance context, in particular it shows an overview of some vibration analysis tools, such as ICA (independent component analysis), TFA (time-frequency analysis), ED (energy distribution) and CD (change detection).

In [3], anomaly detection techniques using machine learning models such as K- Nearest Neighbour (KNN), Support Vector Regression (SVR) and Random Forest (RF) have been applied to vibration data for early fault detection of industrial electric motors. According to the authors the Random Forest presented the best performance compared to SVR and KNN, based on less number of false positives and the detection time.

In [4], is proposed a method for providing a visual explanation of the predictions of a convolutional neural network (CNN)- based anomaly detection system. The CNN takes the monitoring target machine's vibrational data as input and predicts whether the target's state is healthy or anomalous.

For a relation between the motor speed and vibration signals, [5] proposes a CNN based deep learning approach for automatic motor fault diagnosis. In the same research line [6] establishes a comparison of fault motor diagnosis using RNN (Recurrent Neural Networks) and k-means in vibration analysis.

For accurate detection, it is important that the acquisitions and the entire experimental acquisition chain are also calibrated and that the best sensors are used for the best application. Accelerometers are widely used in this type of applications, however, other types of sensors such as magnetoresistive sensors have shown their potential due to their high sensitivity. In [7] was studied the comparison between magnetoresistive sensors and accelerometers in the acquisition of vibration signals to validate the accuracy in vibration analysis condition monitoring systems.

Most of the research work presented in this section involves extensive mathematical formulations in the development and implementation of the algorithms, as well as their adaptation to fault detection scenarios. In real implementations, it is sometimes not feasible to implement these solutions in industrial controllers capable of running in real-time. On the other hand, the solution proposed in this work uses *BigML* [8] as a machine learning tool, as it is a freely available and web-based tool, accessible even to developers with few mathematical bases on artificial intelligence algorithms. With this approach, anomaly detection processes can be more quickly and efficiently applied in real-world vibration analysis scenarios.

3 Materials and Methods

According to [9], vibration analysis methodology in rotating machines help to determine, unbalanced, misalignment, looseness, bearing faults, gear defects, belt wear an tear, pum cavitation and others. Analysis is usually performed by measuring mechanical movements with accelerometers or magnetic sensors. Typically the signal amplitude is analyzed on the time and frequency through a FFT computation. Based on the ISO-10816 Vibration Severity Chart standard, it is possible to categorize the severity of the problem into 4 classes depending on the power and size of the rotary machine. According to the same standard, typical faults produce unusual low-frequency vibrations, so the analysis range should be from 10 to 1000 Hz. To be more precise in analysing the data, it is known that imbalances, misalignments and looseness are recorded at frequencies up to 300 Hz. The relationship between the failures that occur in rotating machines and the frequency spectrum is illustrated on Fig. 2.

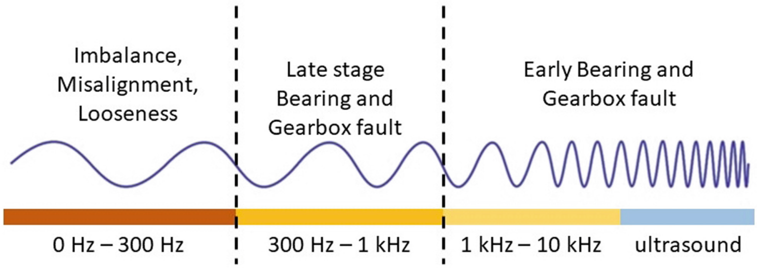

Fig. 2. Machines typical faults distributed in the frequency spectrum.

In the state of the art there are some algorithms that are currently widely used to implement anomaly detection solutions. *Robust Covariance* [10] is an algorithm that detecting anomalies and outliers by means of the *Mahalanobis* distance. *One-class SVM* [11] is an unsupervised algorithm that learns a decision function for novelty detection: classifying new data as similar or different to the training set. The *Local Outlier Factor (LOF)* [12] algorithm is an unsupervised anomaly detection method which computes the local density deviation of a given data point with respect to its neighbours. It considers as outliers the samples that have a substantially lower density than their neighbours. For this work, the implementation of the *Isolation Forest* through framework *BigML* was considered.

Isolation Forest (IF) [13,14] is an unsupervised model, without need a predefined labels, based on decision trees, extensively used for outlier detection. In an *Isolation Forest*, randomly sub-sampled data is processed in a tree structure based on randomly selected features. The samples that travel deeper into the tree are less likely to be anomalies as they required more cuts to isolate them. Similarly, the samples which end up in shorter branches indicate anomalies as it was easier for the tree to separate them from other observations.

As illustrated on Fig. 3, the algorithm try "isolate" outliers from the normal points. In order to isolate a data point, the algorithm recursively generates partitions on the sample by randomly selecting an attribute and then randomly selecting a split value for the attribute, between the minimum and maximum values allowed for that attribute.

According to the original propose of isolation forest [13], the anomaly score (s) in a instance (x) is calculating by:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{1}$$

where $h(x)$ is the path length of a point x and $E(h(x))$ is the average of $h(x)$ from a collection of isolation trees. $c(n)$ is the constant value to normalize the average path length for n trees.

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n} \tag{2}$$

where H(i) can be estimated by $ln(i) + 0.5772156649$ (Euler's constant) as the harmonic number.

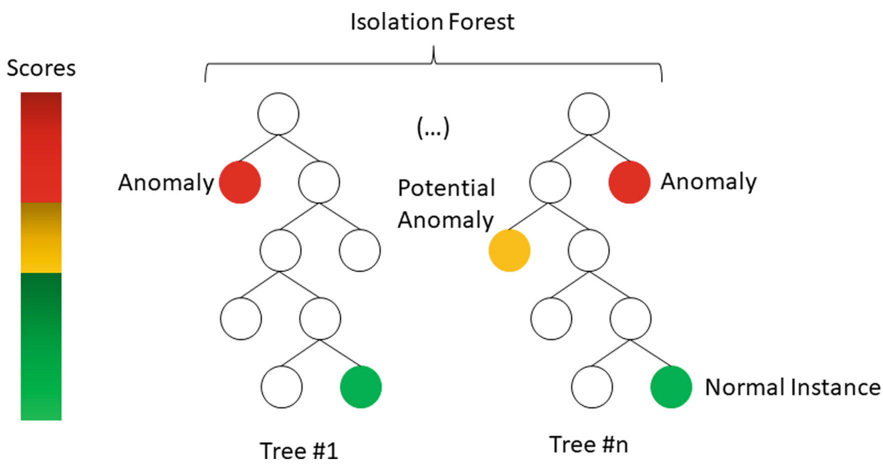

Fig. 3. Overview of the isolation forest method.

4 Experimental Results

This section presents the experimental results achieved in the laboratory through different tests in an experimental apparatus consisting of a single-phase 0.2 kW/3000 rpm motor and a shaft alignment/misalignment system, as illustrated on Fig. 4. Vibration acquisitions were performed with a DYTRAN model 3134D piezoelectric accelerometer with a sensitivity of 500 mV/g through a PCB PIEZOTRONICS 482A21 ICP signal conditioner, a National Instruments data acquisition board (NI DAQ 6008) and a data acquisition virtual instrument software developed in LabVIEW. The accelerometer and the data acquisition chain was calibrated with a PCB PIEZOTRONICS 394C06 handheld portable shaker that oscillates a 159.2 Hz. The data acquisition setup is aligned with the Operational Technology (OT) part of Fig. 1.

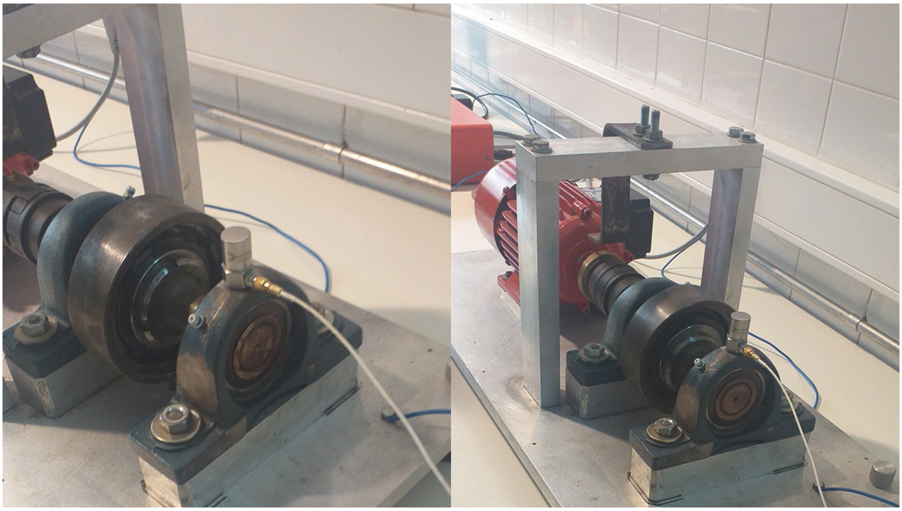

Fig. 4. Experimental apparatus.

According to the ISO-10816 standard, all acquisitions were performed up to 1000 Hz, with a resolution of 0.5 Hz. Different experiments were carried out, namely misalignments and loosening at different rotation speeds, which allowed the creation of a diversity of datasets for analysis. On Fig. 5 is represented the frequency spectrum of one acquisition. As expected, signals with information about the motor status operation are identified at low frequencies, until 300 Hz.

For better interpretation of the signals, a truncated to 300 Hz representation is performed, as shown in Fig. 6. Three groups of signals with greater amplitude are perfectly visible, which means that the detection of any anomaly will have to go through the analysis of these signals. This acquisition was performed with the motor at 570 rpm, measured with a strobe lamp. It means that the first peak

9.5 Hz (570 rpm/60) corresponds to the motor speed. From the spectral analysis it is also possible to pre-determine a threshold for which we can consider that the value is an outlier, a potential anomaly. This threshold corresponds to the value of mean plus *3*$\times$ *standard deviation* as shown in Fig. 6.

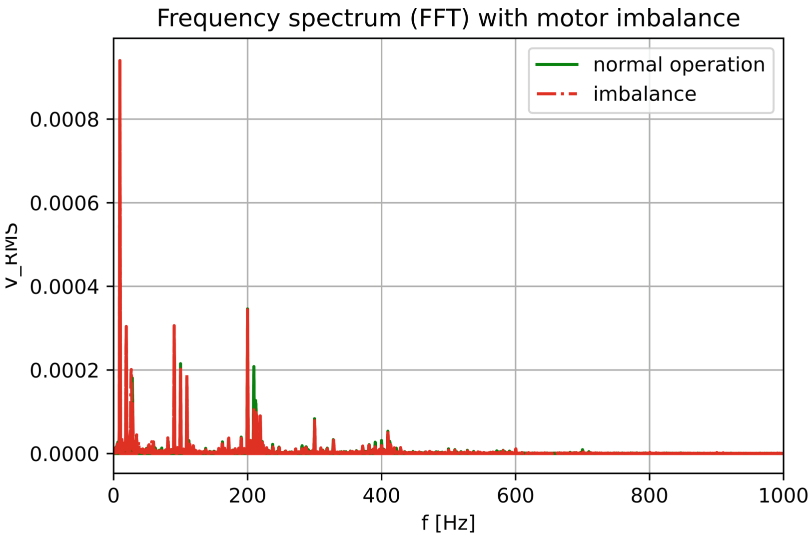

Fig. 5. Frequency spectrum of an acquisition.

In this experiment a misalignment of the motor shaft has been imposed, which implies a significant increase in the amplitude of the frequencies associated to the motor, as better illustrated in Fig. 7. This misalignment caused an imbalance in the motor, which is why all the peaks gained amplitude. At the motor speed frequency, the peak reaches a vibration velocity $V_{RMS} = 0.858$ mm/s, which according to ISO 10816 vibration severity standards, for small machines which is the case (Class I), corresponds to a **satisfactory** severity.

The automatic anomaly detection was implemented in the BigML framework, explorating the potentialities of unsupervised learning isolation forest algorithm. Figure 8 shows a snapshot of the BigML output with the identification of the top 5 anomalies detected. The algorithm is configured to search 20 anomalies along the spectrum and is automatically identified 16 anomalies with an accuracy score of 90.61% as presented on Table 1. This results are very promissory and shows that the isolation forest is an adequate algorithm for this kind of analysis.

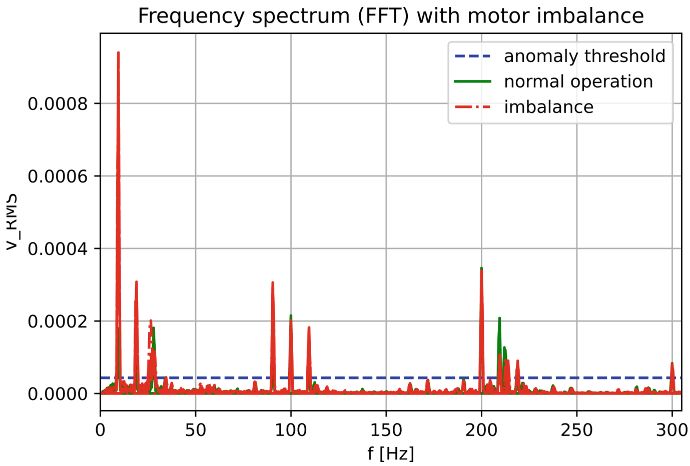

Fig. 6. Frequency spectrum [0–300 Hz].

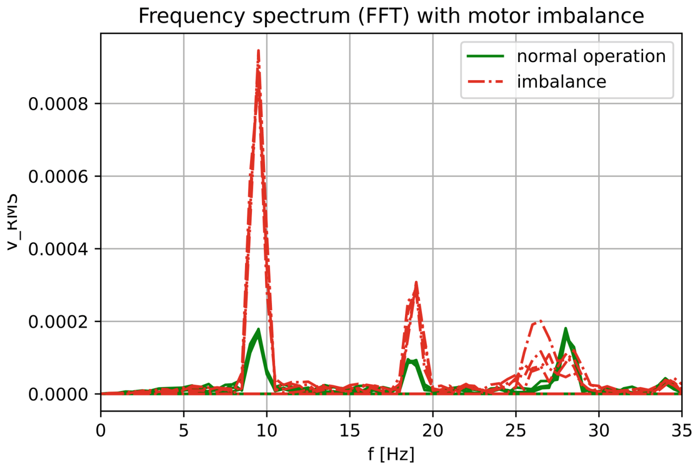

Fig. 7. Frequency spectrum [0–35 Hz].