Journal of Computer and Communications, 2022, 10, 1-12
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Monitoring and Detection of Wind Turbine Vibration with KNN-Algorithm

Javier Vives

Institute of Automatic and Industrial Informatics, Universitat Politècnica de València, Valencia, Spain
Email: javifus@posgrado.upv.es

How to cite this paper: Vives, J. (2022) Monitoring and Detection of Wind Turbine Vibration with KNN-Algorithm. *Journal of Computer and Communications*, **10**, 1-12.
https://doi.org/10.4236/jcc.2022.107001

Received: June 1, 2022
Accepted: July 5, 2022
Published: July 8, 2022

Abstract

Maintenance for wind turbines has been transformed using supervised machine learning techniques. This method of automatic and autonomous learning can identify, monitor, and detect electrical and mechanical components of wind turbines and predict, detect, and anticipate their degeneration. Using a machine learning classifier and frequency analysis, we simulate two failure states caused by bearing vibrations. Implementing KNN facilitates efficient monitoring, monitoring, and fault-finding for wind turbines. It is possible to reduce downtime, anticipate breakdowns, and import offshore aspects through these technologies.

Keywords

Wind Turbines, Vibrations, Fault Diagnosis, Machine Learning, Condition Monitoring, Internet of Things

1. Introduction

Global warming and increased energy consumption are driving the use of renewable sources of electricity. Research has developed new techniques for maintaining wind power infrastructure, and wind power production has increased by about 40% in the past few years. By using advanced monitoring and fault diagnosis, wind turbines can be made more reliable, safer, and more profitable. A spectral analysis and fault tree analysis have traditionally [1] been used to maintain wind turbines.

A growing number of businesses are turning to artificial intelligence (AI) thanks to the growth of digital technology, mobile technology, and smart and data-driven technologies. The industry is currently able to access a growing amount of data, which can have many implications, including scheduling [2], maintenance management [3], and quality improvement [4].

Machine learning has become even more relevant in these areas with the advent of cloud-based solutions and new hardware [5]. Failure to replace a motor or switch usually causes vibrations. Vibrations can also indicate gear and bearing failures. The rolling elements in bearings wear primarily because the bearing surface position continuously changes with respect to the load, because of the rotation speed. The advent of new hardware and cloud-based solutions has made machine learning even more impactful in these fields [6]. The cause of vibration is usually either mechanical or electrical failure. Vibrations can also indicate gear and bearing failures. Due to their rolling elements, bearings are subject to wear in large part due to the way their surface position is continuously adjusted with respect to the load. A vibration can be caused by geometric imperfections, cage failure, as well as imbalance and misalignment. The use of spectral analysis to detect bearing failures caused by mechanical failures has been used in several studies [7]. In the past, various diagnostic techniques have been used to study wind turbine generators and their structures [8]. Artificial Intelligence [9] found that Machine Learning worked perfectly and continues to work perfectly. However, there are some limitations and drawbacks to this kind of methodology. Maintenance methodologies can automatically diagnose and classify a malfunctioning component's function. Data management and analysis allow for flexible offshore implementation and feedback learning, according to [10], while machine learning reduces response times and virtually eliminates errors, as per [11].

Validating AI methods is essential for implementing them successfully on real systems without costly errors. With AI methodology, you are protected against all types of failures by analysing and preventing them. Validating fault diagnosis techniques and understanding how these systems work is accomplished by developing new techniques, performing studies, etc., using prototypes or test benches. Wind turbines broken during peak energy times can cause considerable losses for two reasons: first, they cost a great deal to replace, and second, their inability to generate energy contributes to the loss. The use of fault detection and diagnosis techniques is essential to avoid high repair and maintenance costs in offshore wind farms, especially those subject to high repair and maintenance costs. Aside from reducing downtime costs, it becomes increasingly important to manage maintenance activities efficiently. By applying algorithms designed to anticipate and prevent problems, we developed a prototype that detects, supervises, and anticipates failures in comparison to existing systems. The purpose of our article is to present a method in which vibration analysis can be used to monitor and diagnose faults in a prototype wind turbine. Automatically detecting bearing failures is presented in this paper. After reviewing the literature, data collection and analysis, the classification results are assessed before concluding with a review of the literature. As a result of the study, some significant conclusions can be drawn.

2. Research Methodology

There are many ways to diagnose and monitor the vibration of bearings in a wind turbine, and each bearing has different characteristics. Thus, bearing characteristics may be different from fault characteristics in general. In this study, machine learning is used to improve accuracy and predict potential failures based on vibration measurements from another bearing.

2.1. Machine Learning

Machine learning techniques for wind turbine fault detection mainly address two tasks: detecting anomalous behavior and classifying faults. The system is made more reliable and secure by using this technique, which also makes it possible to take corrective measures very quickly if the system fails.

To classify the different bearing data and determine which ones are in good condition or which ones have a fault and, in this case, what type of fault it is, Machine learning algorithms have been used [12].

Machine learning is a subfield of computer science which belongs to the field of Artificial Intelligence and whose objective is to create systems that learn automatically. Machine learning algorithms use computational methods to extract information from the data that is entered into the algorithm without relying on a specific equation as a model, and by increasing the number of data that we enter into the algorithm, it adapts and improves its results.

Nowadays, with the increase in the amount of data that computers can store and manage, Machine learning has become a very useful tool to solve problems or make better decisions. Several of the sectors where algorithms of this type are used facial recognition, motion detection, manufacturing, predictive maintenance, etc.

There are two types of techniques when obtaining machine learning algorithms, unsupervised learning and supervised learning.

Unsupervised learning consists of obtaining patterns not visible to the naked eye in the data that we introduce to the algorithm, while supervised learning tries to train a model with known input and output values, to predict the output values when introducing certain values, input data. Of these two types of machine learning, the one used in this work has been supervised learning.

Within supervised learning, two types of techniques are used to obtain models: regression and classification techniques.

Regression techniques are used to predict the value of continuous variables over time; however, in this work classification techniques have been used since they are used to predict discrete variables.

2.2. Decision Trees

The algorithm that is based on some input data (inputs) provides a solution (output) following a set of conditions or decisions which go from a root node, which is the beginning of the tree, to a terminal node [13]. The conditions at each node and the number of nodes are determined when training the algorithm. Some of the advantages of this algorithm are: that is easy to interpret, it uses little memory space, and it can be adjusted to nonlinear patterns. As for disadvantages: it is sensitive to anomalous data and does not usually offer very high precision.

2.3. Discriminant Analysis

This type of algorithm classifies data by looking for linear combinations of features or predictors. They assume that the data belonging to each class follows a Gaussian distribution [14]. Gaussian distributions are characterized by certain parameters which are calculated for each class when training this type of algorithm. From these parameters, the limits that separate the data regions of different classes are obtained and thanks to these limits, new input data are classified.

2.4. Naïve Bayes Algorithms

They are quite simple algorithms and easy to interpret. They use the Bayes' theorem to decide to which class the input data belongs. When using this theorem, it is assumed that the predictor variables are independent of each other, that is, that the presence of a variable or predictor in the data is not correlated with any other predictor [15]. When introducing new data to the algorithm, it will classify them in the class to which they are most likely to belong, calculating this probability using Bayes' theorem.

2.5. Support Vector Machines

These algorithms try to obtain the "border" that best separates the different classes. These borders between classes are called hyperplanes; the objective is to obtain a hyperplane that separates the data of different classes, leaving the greatest possible margin between them. The points from which the distance or margin is measured are called support vectors, which are the points or points that are closest to the hyperplane [16].

This algorithm is usually used when there are two possible classes to classify the input data (although it can also be used for problems with more than two classes, using the same theoretical concept that is used when classifying between two classes, the problem becomes more complex).

2.6. K-Nearest Neighbout (KNN)

It is a learning algorithm based on the principle that instances within a data set will generally exist close to other instances with similar properties [17]. This methodology does not generate a model resulting from learning with training data, but rather learning happens at the same time that the test data is tested. This type of algorithm is also known as the lazy learning method [18].

Its operation is very simple, for a given training group of classified instances $T = \left[(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\right]$, where x_i is the vector of characteristics of the unlabeled instance, y_i is the label $y_1 = c_1, c_2, \cdots, c_K$, $i = 1, 2, \cdots, N$. For a training sample (*x, y*), the k-NN algorithm finds the *k* closest instances to *x* based on a given distance metric. The area that contains these k instances is represented by $N_k(x)$. Therefore, the test sample label x can be calculated based on the decision rules:

$$y = \arg\max_{c_j} \sum_{x_i \in N_k(x)} I(y_i = c_j), i = 1, 2, \cdots, N; j = 1, 2, \cdots, K \quad (1)$$

where I is the indicator function.

If the instances are tagged using a tagger, then the tag of an unclassified instance can be obtained by analysing its closest neighbours, as shown in Figure 1.

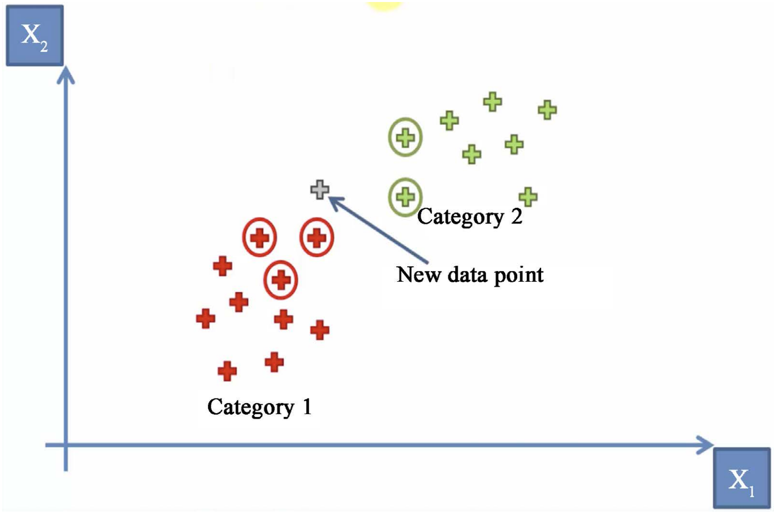


Figure 1. K-NN diagram with different samples.

3. Case Study

The document describes the industrial environment and how the sensors will be distributed, as well as the components within which the system will operate. In addition, a data acquisition card's characteristics and connections are discussed.

3.1. Sensor Distribution and Prototype Development

A small wind turbine prototype like the one in Figure 2 can be useful for

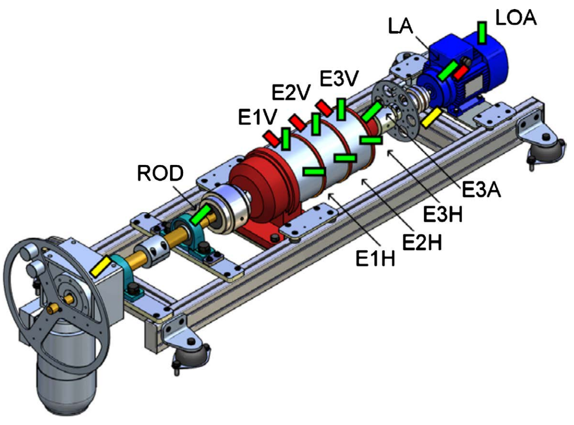


Figure 2. Component distribution in the prototype.

diagnosing problems with components since it can detect damage and wear on the parts and how these affect them [19].

As a result of this system, parts can be exchanged without waiting for deterioration to occur, allowing for testing of diagnostic techniques before the parts deteriorate. Vibration sensors are installed close to the fast shaft coupling to measure vibrations from generators, gearboxes, and bearings.

To monitor the vibrations generated by the fast shaft coupling to the generator, sensors should be installed in the input bearing of the generator. Depending on the state monitoring techniques and design of the machine, sensors should be placed in each process stage of the multiplier. In doing so, you can see how different failures affect vibrations and see how they propagate between stages. A bearing located on the slow axis is another interesting metric for measuring the prototype. We may change this element on some of the deteriorated bearings to observe the signal's behavior following a failure as well as to see how the signal behaves in normal operation and how the component deteriorates over time.

After considering the previous points, it was decided to place 10 accelerometers in total, distributed as follows:

- Bearing: 1 accelerometer;
- Multiplier: 7 accelerometers;
- Generator: 2 accelerometers.

Vibrations are measured with accelerometers. They have a 2-pin MIL-C-5015 connector [19], so they are generally useful accelerometers. Some characteristics are presented in Table 1.

Table 1. Accelerometer characteristics.

Measured range	Sensitivity (±10%)	Frequency range (±3 dB)
±50 g (±490 m/s^2)	100 mV/g (10.2 $mV/(m/s^2)$)	30 a 600,000 cpm (0.5 a 10,000 Hz)

3.2. Data Collection and Description

For vibration measurement, accelerometers are used. The accelerometers have a two-pin MIL-C-5015 NI connector. To measure vibration, we used the PCI-4472B acquisition card, which offers eight-channel dynamic signal acquisition. For accelerometers and microphones, IEPE is used to integrate the signals of the eight input channels simultaneously. The eight input channels cover a bandwidth from DC to 45 kHz. When the card is AC coupled with very low-frequency AC vibration measurements, the PCI 4472B performs with a cut-off frequency of only 0.5 Hz. NI sound and vibration analysis software, including the NI Sound and Vibration Measurement Suite and the NI Sound and Vibration Toolkit, provides signal processing functionality to perform audio measurements, fractional octave analysis, frequency analysis, transitory analysis and order tracking [20]. We use two PCI-4472B cards since the prototype has 10 accelerometers, and each card only provides 8 inputs. In Table 2, a short description (label) of accelerometers for the data acquisition system is presented. As you can see in

Table 2. Description accelerometers to the data acquisition system.

Data Attributes	Data Description
First PCI-4472B	
Channel 0	Accelerometer LA.
Channel 1	Accelerometer LOA.
Channel 2	Accelerometer E2V.
Channel 3	Accelerometer E2H.
Channel 4	Accelerometer E3V.
Channel 5	Accelerometer E3H.
Channel 6	Accelerometer 3EA.
Second PCI-4472B	
Channel 0	Accelerometer EV1.
Channel 1	Accelerometer E1H.
Channel 2	Accelerometer ROD.

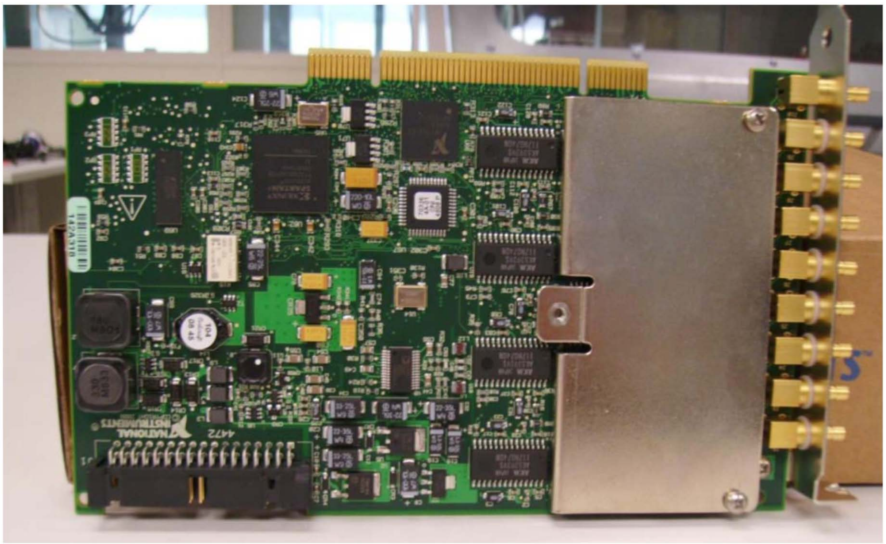


Figure 3. Connection of accelerometers to the data acquisition system.

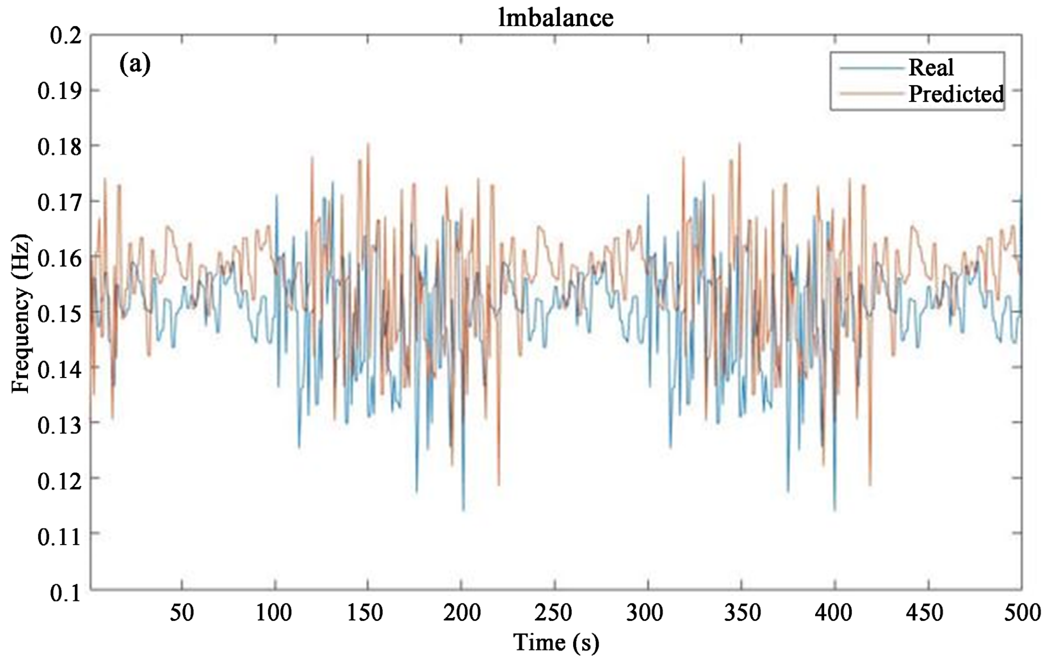
Figure 3, the NI PCI-4472B also has eight accelerometer inputs, which can be connected as you would any other PCI card.

4. Results and Discussion

The simulation is successfully run using traditional methods versus the artificial methods described in this section. Historically, vibrational motion is measured with spectral analysis. With traditional methods compared to the artificial methods described here, the simulation is successfully run. Spectral analysis has historically been used to measure vibrational motion. An accelerometer array around the wind turbine prototype is shown in figure X. 0 rpm to 1500 rpm is the speed range for the prototype. A medium speed of 300 rpm was selected for this particular case. Wind turbines can be tracked, diagnosed, and prevented using traditional vibration analysis methods coupled with automated learning systems. Based on an average of 5000 samples collected by the selected sensors, a graphical presentation is generated at 1 KHz. Automation can be used to prevent, diagnose, and track wind turbine failures. The algorithm needs to be trained so it can analyse and classify the data independently, so it can make a correct prediction. The following section explains how to train and teach the algorithm.

Based on the training, we were able to simulate 2 states of analysis: imbalance and breakage. We believe that the feedback we received is valuable for accurate prediction. About eight times, the algorithm was trained. As a final analysis, we used KNN to compare the two states. A total of four phases of analysis were conducted, starting with the acquisition of the data using a PCI-4472B card, followed by filtration and processing. Stabilizing the analysis involves transforming the signal into something that is not random. Machine learning algorithms must be properly conditioned and processed efficiently to extract patterns from signals of this type. Signals of this type are time-varying and therefore hard to process. To ensure the algorithm works correctly, the filtering and conditioning phase must be completed. Signal processing algorithms can be used to read the invariant characteristics of signals in real-time. Identifying faults or conditions requires the extraction of features. The arithmetic mean is computed by adding the examples of each problem (based on the predetermined issue condition) and dividing by the number of tests considered. The principal component analysis is then used to reduce the number of variables in the data set to a minimum number that produces the same outcome as the original variables. Furthermore, by understanding the current state as well as what is happening, we can make future decisions based on their standard deviation. It shows that there are differences or dispersions between many of the states, which implies that most of the points are around the average, which is why the study should work. Training sessions follow the entire process so that the algorithm eventually becomes self-operating. The algorithm can be made to work with a few training sessions; it just needs new data.

Each state will be explored separately. It is important for KNN to update the limits used in their feedback algorithm due to the fact that the limits used in this case are out of date with respect to the problem of imbalances (**Figure 4(a)**). The failure is caused by bearing race breakage, as shown in **Figure 4(b)**. Here's a breakdown of each state. Firstly, the classifier follows the same pattern. Since KNN does not follow a set pattern for sorting and classifying data, it produces more grouped data, as a result of the techniques used. Additionally, the algorithms in both cases are highly accurate and are similar to the actual and predicted outputs in both cases. **Figure 5** shows the confusion matrix. Let us examine the imbalance variable as an example. 95% percent of the time, this variable



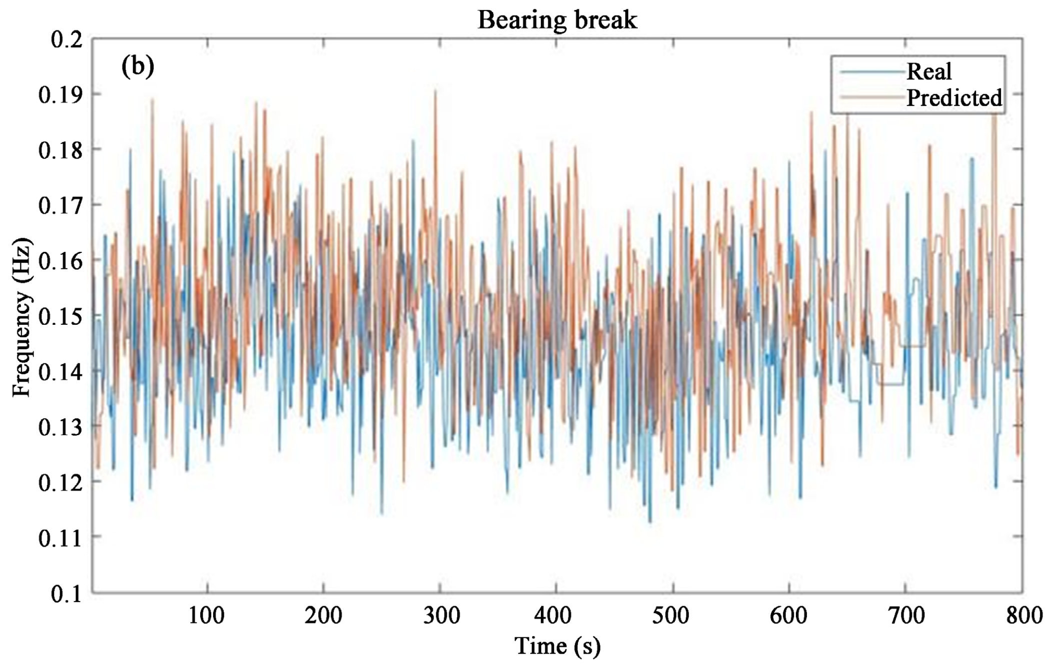


Figure 4. (a) Imbalance. Real output vs. predicted output algorithm; (b) Bearing break. Real output vs predicted output algorithm.

is correctly classified, while 5% percent of the time, it is incorrectly classified. A total of 95% of true positives and just 5% of false positives were detected. In this way, KNN is able to accurately predict the wind turbine's failure thanks to its many similarities with the prototype. As a result of this simulation, we have observed that both the KNN and the BNN classifications were quite robust to the interference (noise). The KNN algorithm also allows for regression classification to be implemented.

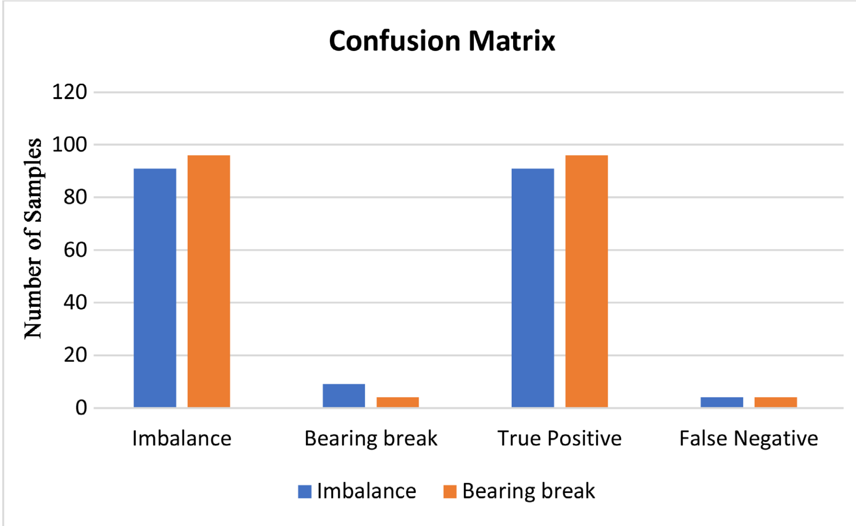


Figure 5. Confusion matrix (KNN).

5. Conclusion

A large part of the success and proper functioning of AI depends on the acquisition and classification of data. Machine learning systems are changing how wind turbine faults can be detected, monitored, and diagnosed, making them better accessible. The purpose of this document is to explore several different AI techniques for analysing vibrations to diagnose and prevent failures in wind turbine bearings. KNN models have been used to diagnose bearing faults from a theoretical and practical perspective. This model has a lot of advantages, including its robustness, high accuracy, and high processing speeds, which make them very suitable for this type of study. Traditional methods such as spectral analysis are being displaced due to some of their advantages, such as their ease of classification and prediction. As a result, the methodology provided good predictions for the stipulated failure conditions, allowing this methodology to be used for other mechanical components of wind turbine prototypes, with the goal of identifying or preventing possible breakdowns. This prototype facilitates the study, development, and validation of fault diagnosis and supervision techniques by providing the possibility of replacing defective or worn parts with other components. In the phase leading up to their installation in high-power wind turbines, prototype wind turbines allow testing of the designed diagnostic algorithms, reducing costs and time, and allowing them to be verified, adjusted, and corrected.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Yang, B., Liu, R. and Chen, X. (2018) Sparse Time-Frequency Representation for Incipient Fault Diagnosis of Wind Turbine Drive Train. *IEEE Transactions on In-*

strumentation and Measurement, **67**, 2616-2627. https://doi.org/10.1109/TIM.2018.2828739

[2] Sánchez, P.J.B., Ramirez, I.S. and Márquez, F.P.G. (2021) Wind Turbines Acoustic Inspections Performed with UAV and Sound Frequency Domain Analysis. 2021 *7th International Conference on Control, Instrumentation and Automation* (*ICCIA*), Tabriz, 23-24 February 2021, 1-5.

[3] Ramirez, I.S., Urioso, E.C., Peco, A.M., Kotorov, R., Chi, L., Padhye, R.G., García Márquez, F.P., *et al.* (2021) Motif Analysis in Internet of the Things Platform for Wind Turbine Maintenance Management. In: Xu, J.P., *et al.*, Eds., *International Conference on Management Science and Engineering Management*, Springer, Cham, 74-86. https://doi.org/10.1007/978-3-030-79203-9_7

[4] Zhang, R.G., Wang, B.J., Wang, J.B., Lei, H. and Xia, S.C. (2021) Wind Turbine Blade Failure Case Analysis and Control Measures. *Journal of Physics*: *Conference Series*, **2047**, Article No. 012014. https://doi.org/10.1088/1742-6596/2047/1/012014

[5] Lee, G.S., Choi, J., Kang, M., Park, S. and Lee, J. (2020) Study on Enhancement of Data Processing Algorithm in SaaS Cloud Infrastructure to Monitor Wind Turbine Condition. *New & Renewable Energy*, **16**, 25-30. https://doi.org/10.7849/ksnre.2020.2047

[6] Roy, S., Kundu, B. and Chatterjee, D. (2021) Cloud Based Real-Time Vibration and Temperature Monitoring System for Wind Turbine. In: Shaw, R.N., Mendis, N., Mekhilef, S. and Ghosh, A., Eds., *AI and IOT in Renewable Energy*, Springer, Singapore, 89-99. https://doi.org/10.1007/978-981-16-1011-0_9

[7] Kong, Y., Wang, T. and Chu, F. (2019) Meshing Frequency Modulation Assisted Empirical Wavelet Transform for Fault Diagnosis of Wind Turbine Planetary Ring Gear. *Renewable Energy*, **132**, 1373-1388. https://doi.org/10.1016/j.renene.2018.09.027

[8] Karaagac, U., Mahseredjian, J., Gagnon, R., Gras, H., Saad, H., Cai, L., Wang, L., *et al.* (2019) A Generic EMT-Type Model for Wind Parks with Permanent Magnet Synchronous Generator Full Size Converter Wind Turbines. *IEEE Power and Energy Technology Systems Journal*, **6**, 131-141. https://doi.org/10.1109/JPETS.2019.2928013

[9] Ti, Z., Deng, X.W. and Yang, H. (2020) Wake Modeling of Wind Turbines Using Machine Learning. *Applied Energy*, **257**, Article ID: 114025. https://doi.org/10.1016/j.apenergy.2019.114025

[10] Kreutz, M., Ait-Alla, A., Varasteh, K., Oelker, S., Greulich, A., Freitag, M. and Thoben, K.D. (2019) Machine Learning-Based Icing Prediction on Wind Turbines. *Procedia CIRP*, **81**, 423-428. https://doi.org/10.1016/j.procir.2019.03.073

[11] Liu, Z., Xiao, C., Zhang, T. and Zhang, X. (2020) Research on Fault Detection for Three Types of Wind Turbine Subsystems Using Machine Learning. *Energies*, **13**, 460. https://doi.org/10.3390/en13020460

[12] Gu, G.X., Chen, C.T. and Buehler, M.J. (2018) *De Novo* Composite Design Based on Machine Learning Algorithm. *Extreme Mechanics Letters*, **18**, 19-28. https://doi.org/10.1016/j.eml.2017.10.001

[13] Chacón, A.M.P., Ramirez, I.S. and Márquez, F.P.G. (2021) Support Vector Machine for False Alarm Detection in Wind Turbine Management. 2021 *7th International Conference on Control, Instrumentation and Automation* (*ICCIA*), Tabriz, 23-24 February 2021, 1-5.

[14] Deng, Z., Kammoun, A. and Thrampoulidis, C. (2019) A Model of Double Descent for High-Dimensional Binary Linear Classification. *Information and Inference: A*

Journal of the IMA, **11**, 435-495. https://doi.org/10.1093/imaiai/iaab002

[15] Ye, J.C. (2022) Linear and Kernel Classifiers. In: Ye, J.C., Ed., *Geometry of Deep Learning*, Springer, Singapore, 29-44. https://doi.org/10.1007/978-981-16-6046-7_2

[16] Zhang, Z.L., Luo, X.G., Yu, Y., Yuan, B.W. and Tang, J.F. (2018) Integration of an Improved Dynamic Ensemble Selection Approach to Enhance One-vs-One Scheme. *Engineering Applications of Artificial Intelligence*, **74**, 43-53. https://doi.org/10.1016/j.engappai.2018.06.002

[17] Utami, E., Iskandar, A.F. and Raharjo, S. (2021) Multi-Label Classification of Indonesian Hate Speech Detection Using One-vs-All Method. 2021 *IEEE* 5*th International Conference on Information Technology*, *Information Systems and Electrical Engineering* (*ICITISEE*), Yogyakarta, 9-10 November 2021, 78-82. https://doi.org/10.1109/ICITISEE53823.2021.9655883

[18] Dzulfikri, Z., Nuryanti, N. and Erdani, Y. (2020) Design and Implementation of Artificial Neural Networks to Predict Wind Directions on Controlling Yaw of Wind Turbine Prototype. *Journal of Robotics and Control* (*JRC*), **1**, 20-26. https://doi.org/10.18196/jrc.1105

[19] Diaferio, M., Foti, D., La Scala, A. and Sabbà, M.F. (2021) Design of the Set-Up for Ambient Vibration and Dynamic forced Tests on a Cable-Stayed Bridge. 2021 *AEIT International Annual Conference* (*AEIT*), Milan, 4-8 October 2021, 1-5. https://doi.org/10.23919/AEIT53387.2021.9626962

[20] Quiles, E., Garciia, E., Cervera, J. and Vives, J. (2019) Development of a Test Bench for Wind Turbine Condition Monitoring and Fault Diagnosis. *IEEE Latin America Transactions*, **17**, 907-913. https://doi.org/10.1109/TLA.2019.8896812