SVM-Algorithm for Supervision, Monitoring and Detection Vibration in Wind Turbines

Javier Vives1*, Juan Palací2, Janverly Heart2

1Department of Systems Engineering and Automation, University Polytechnic of Valencia, Valencia, Spain
2Red Engineering Technology Limited, London, UK
Email: javifus@posgrado.upv.es

Abstract

With the implementation of supervised machine learning techniques, wind turbine maintenance has been transformed. A wind turbine’s electrical and mechanical components can be automatically identified, monitored, and detected to predict, detect, and anticipate their degeneration using this method of automatic and autonomous learning. Two different failure states are simulated due to bearing vibrations and compared with machine learning classifier and frequency analysis. A wind turbine can be monitored, monitored, and faulted efficiently by implementing SVM. With these technologies, downtime can be reduced, breakdowns can be anticipated, and aspects can be imported if they are offshore.

Keywords

Vibrations, Wind Turbines, Fault Diagnosis, Machine Learning, Condition Monitoring, Deep Learning

1. Introduction

The use of renewable sources of electricity is becoming more popular due to global warming and increased energy consumption. Wind production has increased by about 60% in the past few years, and researchers have developed new techniques to maintain wind power infrastructure. The use of advanced monitoring and fault diagnosis can enhance the reliability, safety, and profitability of wind turbines. Historically, wind turbines have been maintained by performing spectral analysis and fault tree analyses [1].

With the advancements in digital and mobile technology as well as smart and data-driven technologies, artificial intelligence (AI) is becoming a more popular option. A growing amount of data is available to the industry right now, which
has many implications, including scheduling [2], maintenance management [3], and quality improvement [4].

As a result of new hardware and cloud-based solutions, machine learning has had an even greater impact in these areas [5]. The cause of vibrations is usually either a mechanical or electrical failure. Vibrations can also indicate gear and bearing failures. Since the bearing surface position adjusts continuously with respect to the load, due to rotation speed, bearings are subject to wear largely due to their rolling elements. As a result of new hardware and cloud-based solutions, machine learning has had an even greater impact in these areas [6]. The cause of vibrations is usually either a mechanical or electrical failure. Vibrations can also indicate gear and bearing failures. Since the bearing surface position adjusts continuously with respect to the load, due to rotation speed, bearings are subject to wear largely due to their rolling elements. In addition to geometric imperfections, vibrations may be caused by component failing, cage failure, and imbalance and misalignments. In several studies [7], spectral analysis has been used to detect structural failures and rotating machine failures by detecting bearing failures caused by mechanical failures. Wind turbine generators and their structures have been studied using various diagnostic techniques in the past [8]. Machine Learning was found to work perfectly and to continue to work perfectly according to AI [9]. However, there are some limitations and drawbacks to this kind of methodology. Through a series of maintenance methodologies, the function of a malfunctioning component can be detected, detected, and classified automatically. Machine learning reduces response times and virtually eliminates errors, according to [10], while data management and analysis allow for flexible offshore implementation and feedback learning, as per [11].

To successfully implement AI methods on a real system without causing costly errors, they must be validated. By analysing and preventing failures, AI methodologies protect you from all kinds of failures that you desire to monitor. Developing new techniques, performing studies, etc., using prototypes or test benches is useful when validating fault diagnosis techniques, as well as understanding how these systems work. A broken wind turbine can cause considerable losses due to two reasons: first, the cost of replacing them, and second, the loss of energy as they cannot be produced during peak energy times. In offshore wind farms, especially those that are subject to high repair and maintenance costs, the use of fault detection and diagnosis techniques is essential if the machine is to be stopped early if it has a fault. In addition, as the costs of downtime and defective products decline, managing maintenance activities efficiently becomes an increasingly important task. In comparison to existing systems, a prototype is developed to detect, supervise, and anticipate failures through the application of algorithms designed to anticipate and prevent problems. In this article, an algorithm for monitoring and diagnosing faults in a prototype wind turbine using vibration analysis is presented. The algorithm for detecting different bearing failures autonomously. This study begins with a review of the literature, followed by an analysis of the data collection and data set, followed by an
assessment of the classification results. The study concludes with some meaningful conclusions.

2. Research Methodology

A vibration in bearings in a wind turbine can be diagnosed and monitored using different methods, and each bearing can have a variety of characteristics. As a result, the general characteristics of any bearing may not necessarily correspond to the fault characteristics. Based on the vibration measurements of another bearing, this study demonstrates how machine learning can improve accuracy and predict possible failures.

2.1. Machine Learning

Machine learning techniques for wind turbine fault detection mainly address two tasks: detecting anomalous behaviour and classifying faults. In addition to improving the system’s performance and security, this technique allows corrective measures to be taken very quickly in the event of a failure or an anticipated issue. The most common machine learning method, by far, is supervised machine learning [12], and there is an unsupervised method. When you use supervised learning, you already know the output. With unsupervised learning, you do not know the outcome. The process is the same on the way in and out. Unlike supervised learning, unsupervised learning relies only on the input data and binary logic that all systems use. No references are used at all.

The first step in applying any type of learning is to classify the data. This problem can be solved by applying different classification algorithms, which identify an object by fusing the functionality of the object with a number of categories or classes from the input information it provides.

This leads to a two-phase classification process:

● Training means supplying a large amount of sample data and correctly classifying it, and then adjusting the parameters to reach the optimal performance.
● Input data is used to provide an output once the algorithm has been trained.

2.2. Support Vector Machines

An SVM is a machine learning algorithm based on statistical learning theory. This method works well for classification and regression, such as in fault diagnosis, when it uses small samples. It is shown that a linear classifier can separate two simple classes. These two types of samples are represented by triangles and squares in Figure 1. Two classes can be separated by a hyperplane $H$. In these two classes, the planes $H_1$ and $H_2$ (shown in dashed lines) are parallel to $H$ and pass-through samples that are closest to $H$. Margins are calculated by taking the distance between $H_1$ and $H_2$. In the SVM, linear boundaries are placed between two distinct classes $H_1$ and $H_2$. The margin is maximized, so the generalization error is smallest. Support vectors are often used to measure margins,
and they include the closest points to the margin. A quadratic function is minimized under linear inequality constraints by reducing it to convex optimization \[13\]. Assume we have a training set of samples \([(x_i), (y_i)]\), where \(i = 1 \text{ to } N\), and \(N\) represents the total number of samples. To find the separation plane with the least generalization error out of each linear separation plane, it needs to determine how to divide the input samples into two classes. It is possible to divide the samples into two classes: triangular and square. A triangle class has a \(y_i = -1\) label. A square class has a \(y_i = +1\) label. For non-separable data, slack variables are not considered (nor \(P_0\)). Using the following optimization problem, you can obtain the hyperplane for \(f(x) = 0\) from the given data.

Minimize \[
\frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i
\]

Subject to \[
y_i \left( w^T x_i + b \right) \geq 1 - \xi_i
\]
\[
\xi_i \geq 0, \quad i = 1, 2, \ldots, N
\]

where \(C\) is a constant representing the error penalty. Introducing Lagrange multipliers to the optimization problem above leads to the following result:

Minimize \[
W(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=0}^{N} y_i y_j \lambda_i \lambda_j \langle x_i, x_j \rangle
\]

Subject to \[
0 \leq \lambda_i \leq C
\]
\[
\sum_{i=1}^{N} \lambda_i y_i = 0, \quad i = 1, 2, \ldots, N
\]

Using the sequential minimum optimization (SMO) algorithm, the dual problem that results from SVM derivation can be efficiently solved. SMO breaks down the general QP problem into QP subproblems.

### 2.3. Binary Classification for Linearly Non-Separable Cases

Previously, it has been explained how separation hyperplanes are good classifiers when the classes are perfectly separable or quasi-perfectly separable. However,
the application of a support vector classifier for cases that are clearly not linearly separable (most real problems) lacks practical interest.

One possibility to deal with nonlinear boundaries between classes is to manually increase the space of predictors by means of polynomial functions or with interaction terms. Still, run the risk of ending up with too many predictors. As an alternative, there is the use of support vector machines or support vector machines (SVM), which are an extension of support vector classifiers that increase dimensionality in a specific way, using kernels, a more computationally efficient approach [14]. Kernels are functions that transform a low-dimensional space into a higher-dimensional space through complex data transformations. It can also be defined as a function that quantifies the similarity between two observations in a new dimensional space.

2.3.1. Linear Kernel

\[ K(x_i, x_j) = \sum_{j=1}^{p} x_{ij} x_{ij} \]  

(3)

The linear kernel quantifies the similarity of a pair of observations using Pearson’s correlation. With a linear kernel, the obtained classifier is equivalent to a support vector classifier.

2.3.2. Polynomial Kernel

\[ K(x_i, x_j) = \left(1 + \sum_{j=1}^{p} x_{ij} x_{ij}\right)^d \]  

(4)

A polynomial kernel of degree \( d \) (where \( d > 1 \)) allows for a much more flexible decision boundary. When a support vector classifier is combined with a nonlinear kernel, a support vector machine is obtained.

2.3.3. Radial Kernel

\[ K(x_i, x_j) = \exp\left(-\gamma \sum_{j=1}^{p} (x_{ij} x_{ij})^2\right) \]  

(5)

where \( \gamma \) is a positive constant such that the larger it is, the greater the flexibility of the SVM. Assuming that the observation \( x' = x'1...x'p \) is far from a training observation \( x_i \) in terms of Euclidean distance, then \( K = x', x_i \) will be very small, which means that \( x_i \) will not influence \( K(x') \). The radial kernel has a very local behaviour, in the sense that only training observations close to a test observation will affect its classification [15].

It is important to note that more flexibility does not necessarily improve predictions, as a very flexible model may fit the training data too much.

2.4. Classification with More Than Two Classes

There are several extensions to SVMs for classification problems with more than two classes \( (K > 2) \), two of the most popular being:

- One versus one.
- One-versus-all.
2.4.1. One-vs-One Ranking
This method builds $(k^2)$ SVMs, corresponding to $K(K-1)/2$, comparing a pair of classes. A test observation is classified using each of the SVMs, by counting the number of times this observation is assigned to each of the $K$ classes [16]. The final predicted class will be the one to which the observation has been assigned in most SVMs.

2.4.2. One-vs-All Ranking
$K$ SVMs are fitted, each time comparing one of the $K$ classes (coded as +1) with the rest of the $K-1$ classes (coded as −1). Being $\beta_{0k}$, $\beta_{1k}$, $\beta_{pk}$ are the parameters resulting from the fit of an SVM and $x^*$ observation, the observation will be assigned to the class for which
$$
\beta_{0k} + \beta_{1k}x_{i}^* + \cdots + \beta_{pk}x_{p}^* 
$$
is greater. That is, the magnitude of $f(x^*)$ indicates how far $x^*$ is from the separation hyperplane, and the further it is, the higher the level of confidence that the observation $x^*$ has been correctly classified [17].

3. Case Study
It contains a description of the industrial environment and the components within which the system will operate, as well as a description of how the sensors are distributed. It also explores the characteristics and connections of a data acquisition card.

3.1. Prototype and Sensor Distribution
It is very useful to diagnose problems with components like the small wind turbine prototype shown in Figure 2 as it can detect deterioration and wear on the parts and what its effects are [18]. The purpose of this system is to allow easy exchange of parts without waiting for deterioration to occur and, therefore, to test diagnostic techniques before deterioration occurs. In measuring the vibrations of generator, gearbox, and bearings, the vibration sensors are positioned near the fast shaft coupling.

![Figure 2. Component distribution in the prototype.](image)
Sensors should be placed in the input bearing of a generator to monitor the vibrations caused by the fast shaft coupling to the generator. In the multiplier, the sensors should be placed in each stage in accordance with the state monitoring techniques and the design of the machine. This will allow the propagation of the signal between the stages and how various failures affect the vibrations. Another interesting element for measuring the prototype is the bearing located on the slow axis. This element can be changed on some of the deteriorated bearings to examine the behaviour of the signal following a failure as well as studying the signal in normal operation and the deterioration of the component itself over time.

Thus, following the previous considerations, it was decided to place a total of 10 accelerometers, distributed as follows:
- Generator: 2 accelerometers.
- Multiplier: 7 accelerometers.
- Bearing: 1 accelerometer.

To measure vibrations, accelerometers are selected. These are general-purpose accelerometers with a 2-pin MIL-C-5015 connector [19].

### 3.2. Data Collection and Description

Vibrations are measured using accelerometers. The accelerometers have a two-pin MIL-C-5015 NI connector. Using the PCI-4472B acquisition card for measured vibration using eight channels of dynamic signal acquisition. The eight input channels of accelerometers and microphones are integrated simultaneously using IEPE. An eight-channel input system covers a spectrum ranging from DC to 45 kHz. The PCI 4472B performs with a cut-off frequency of only 0.5Hz when AC coupled with very low frequency AC vibration measurements. Audio measurements, fractional octave analysis, frequency analysis, transitory analysis, and order tracking are all performed using the NI Sound and Vibration Measurement Suite and the NI Sound and Vibration Toolkit [20]. Each PCI-4472B card provides 8 inputs, which two PCI-4472B cards. As you can see in Figure 3. In addition to eight accelerometer inputs, the NI PCI-4472B also comes with eight PCI cards that can be connected just like regular PCI cards (Table 1).

### 4. Results and Discussion

Both traditional and artificial methods are successfully used in this section to run the simulation. There are five speeds available for rotating the prototype, from 0 rpm to 1500 rpm. The medium speed was chosen to be 300 rpm in this instance. Using automated learning systems, it is possible to track, diagnose, and prevent wind turbine failures using traditional vibration analysis methods. A graphical presentation at 1 KHz was generated from an average of 5000 samples generated from the selected sensors. Automated learning systems can predict failures as well as track, prevent, and diagnose them. It is important to train the algorithm to get feedback before analysing and classifying the data independently.
Figure 3. Connection of accelerometers to the data acquisition system.

Table 1. The data description is shown in table.

<table>
<thead>
<tr>
<th>Data Attributes</th>
<th>Data Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>First PCI-4472B</td>
<td></td>
</tr>
<tr>
<td>Channel 0</td>
<td>Accelerometer LA.</td>
</tr>
<tr>
<td>Channel 1</td>
<td>Accelerometer LOA.</td>
</tr>
<tr>
<td>Channel 2</td>
<td>Accelerometer E2V.</td>
</tr>
<tr>
<td>Channel 3</td>
<td>Accelerometer E2H.</td>
</tr>
<tr>
<td>Channel 4</td>
<td>Accelerometer E3V.</td>
</tr>
<tr>
<td>Channel 5</td>
<td>Accelerometer E3H.</td>
</tr>
<tr>
<td>Channel 6</td>
<td>Accelerometer 3EA.</td>
</tr>
<tr>
<td>Second PCI-4472B</td>
<td></td>
</tr>
<tr>
<td>Channel 0</td>
<td>Accelerometer EV1.</td>
</tr>
<tr>
<td>Channel 1</td>
<td>Accelerometer E1H.</td>
</tr>
<tr>
<td>Channel 2</td>
<td>Accelerometer ROD.</td>
</tr>
</tbody>
</table>

for it to make a correct prediction. The purpose of this section is to describe how to train, teach, and get accurate results from the algorithm. For accurate prediction, two states are simulated: imbalance and breakage, which provided good feedback. The training algorithm was run about six times.

Using SVM, the two states in the final analysis are compared. A four-phase analysis was conducted, beginning with the acquisition of data using the PCI-4472B acquisition card, then filtering and processing. It is necessary to transform the signal into something that is not random to stabilize the analysis. It is crucial to apply appropriate conditioning and efficient processing to extract patterns from signals of this type when using machine learning algorithms. It is difficult to process and learn from the signal due to its time variation. For the algorithm to work correctly, the first stage of filtering and conditioning is essential. Invariant characteristics are read in time by signal processing algorithms. It
is necessary to extract features to figure out if a fault or condition exists. Arithmetic means are determined by adding each example (for each predetermined condition) and dividing it by the number of tests considered. The data set is then reduced to the minimum number of variables necessary to represent the original variables using principal component analysis.

Additionally, can make future decisions can be made based on the understanding of the current state, as well as what is happening, in addition to determining the standard deviation for each failure condition. According to the data presented, there most points are close to the average, so it should work since there is variation across many states. A few training sessions follow the entire process so that the algorithm becomes self-operating in the future. To make the algorithm work, it only needs a few training sessions and some new data.

Following, each simulated state is analyzed in more detail. Due to the failure due to an imbalance, the SVM algorithm is a little out of date regarding feedback (Figure 4(a)). The failure (Figure 4(b)) can be attributed to bearing race breakage based on the data analysis. Due to different ways of classifying and analyzing the data, the data are not well grouped, since they do not follow a specific pattern. No matter which failure condition is considered, the algorithm predicts the actual output with high accuracy and similarity.

An example of a confusion matrix can be found in Figure 5. In 91% of the

![Figure 4](image_url)
cases, the algorithm correctly classified the imbalance variable, while in 9% of the cases, the algorithm misclassified it as a bearing break. Bearing breaks are classified correctly 96% of the time, while imbalances are classified incorrectly 4% of the time. Our wind turbine prototype is assumed to have a lot of similarities with SVM based learning methodology that can assist in predicting the failure of the prototype with a high degree of accuracy.

5. Conclusion

Acquisition and classification of data are critical to AI’s success and proper functioning. The use of machine learning systems is improving the ability to detect, monitor, and diagnose wind turbine faults. This document explores several techniques for analysing vibrations to diagnose and prevent wind turbine bearing failures using artificial intelligence. Bearing faults have been diagnosed theoretically and practically using SVM models. They are very suitable for this type of study because they are robust, highly accurate, and very fast. Due to its ease of classification and prediction, spectral analysis is displacing traditional methods like spectral analysis. To identify or prevent possible breakdowns, this methodology can be applied to other mechanical components of wind turbine prototypes, thereby providing good predictions for the stipulated failure conditions. As a result of this prototype, fault diagnosis and supervision techniques can be studied, developed, and validated through the possibility of replacing defective or worn parts with other components. Before their installation in high-power wind turbines, prototype wind turbines can be tested to verify, adjust, and correct the diagnostic algorithms, reducing costs and time.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.
References


