Application of machine learning in monitoring systems of civil structures

João Zeferino[1], Eduardo Gonçalves[1], Paulo Carapito[1], Filipe Santos[1]
[1] VESAM Engenharia S.A., Zona Industrial de Cantanhede, Lote 69, 3060-197 Cantanhede, Portugal
email: joao.zeferino@vesamgroup.com, eduardo.goncalves@vesamgroup.com, paulo.carapito@vesamgroup.com, filipe.santos@vesamgroup.com

ABSTRACT: Structural Health Monitoring (SHM) aims primarily to accurately identify the current state of a structure, assessing damage levels and eventually allowing to predict its future performance. In civil engineering structures, SHM is often responsible for ensuring public safety while handling with large and complex data. Such monitoring problems can be difficult to solve by conventional computing techniques alone, as they require the acquisition of large data sets that need to be thoroughly and carefully analyzed. This yields big data opportunities to use artificial intelligence methodologies. This paper presents the integration of machine learning (ML) techniques for pattern recognition in SHM systems of civil engineering structures. The developed SHM consists of data acquisition both from time series of values observed at regular intervals and from structurally relevant measured values, called events, where specific data are collected. ML is used in the development of statistical models for feature discrimination. Events are classified into different clusters in a semi-supervised learning procedure, which is an extension of an unsupervised learning to allow their identification. A real-world SHM implementation is presented as a case study of the ML application. It consists of an industrial steel tower structure, containing several mechanical equipment with different loads, which operate at various frequencies. A sensor network is installed, acquiring data on strains, accelerations, and weather conditions. A visualization user interface is provided to access all data through a user friendly and accessible tool. The paper presents the main results obtained and illustrates the potentialities of the applied ML methodology.

KEY WORDS: SHM, Machine Learning, Semi-supervised, Real world application.

1 INTRODUCTION

Assessing and ensuring the safety of structures and people is the main goal of Structural Health Monitoring (SHM) systems applied to civil engineering structures. Economic benefits are also expected if a predictive model can successfully forecast the remaining service life and need for maintenance interventions. Structural damages, as well as their location and severity, need to be accurately detected before too long. To this end, different types of SHM systems have been proposed and implemented in recent decades, such as vibration based [1], guided wave based [2], and computer vision-based [3].

Vibration based techniques are a commonly researched area for global damage diagnosis. In the damage detection process, it is first necessary to determine the occurrence of damage according to the global change in structural integrity, by identifying differences in the vibration data before and after the damage occurs. The measured time domain can be converted into frequency domain or modal domain by transforming techniques [1]. Time domain approaches can be more straightforward and might eliminate the need for domain transformations. Data-driven time series methods use well-established statistical concepts. Along with data fusion, data-driven approaches are used to transform massive data obtained from monitoring into meaningful information [4].

The great amount of data provided by SHM yields big data opportunities to use Artificial Intelligence methodologies that allow automatic monitoring, which ideally should be able to assess structural integrity. Machine Learning (ML) is a subset of Artificial Intelligence that incorporates mathematics and statistics in such a way that allows machines to learn hidden rules and patterns in data. In SHM applications, ML has proven to be a powerful technique that can also contribute to determining the location and severity of damage [5]. ML algorithms are generally catalogued in the literature in two broad categories, according to the nature of learning: supervised learning; unsupervised learning [6, 7]. Supervised learning makes use of "labeled data", encompassing an attempt to classify new data sets using known pairs of input data and output data. This type of ML includes regression methods, aimed at predicting quantities, and classification methods, aimed at predicting labels. In classification methods, training datasets are usually required, with many samples of each class label. Examples of classification methods are k-nearest neighbors, support vector machines, and decision trees. Unsupervised learning provides a learning scheme with "unlabeled data", fitting data sets with unspecified outputs. Examples are clustering methods, such as k-means and spectral clustering, which are able to split datasets into groups base on their similarities. Worden et al. [8] suggested the use of unsupervised learning identify the existence and location of damage, while identifying the type and severity of damage can generally only be done through supervised learning. More recently, a third category of ML is catalogued as semi-supervised learning. This represents a combination of the two learning schemes mentioned above, typically with the aim of obtaining a classification of data using both "labeled data" and "unlabeled data" [9].

This paper presents an SHM methodology applied to civil engineering structures. The methodology includes an ML approach to discriminate influences on signals driven by operational conditions, as well as changes in environmental conditions. The case study of an industrial steel tower structure, with a large amount of mechanical equipment, is used to validate the methodology, using vibration measurements to assess the integrity of the structure over time.

2 METHODOLOGY

2.1 Installation and Monitoring

The methodology applied in the present work to monitor civil engineering structures follows the flowchart presented in Figure 1. The first steps relate to the installation procedure of the SHM system. Then, the system is implemented for long-term continuous monitoring.

In the installation process, it is first conducted a survey of structural and equipment elements. This is based on the project documentation and outcomes of software for production control and management of structures, from which it is possible to obtain the real mechanical properties from laboratory test certificates of the raw material. Then, a computational modeling of the structural behavior is performed. After some in-situ vibration measurements, the computer model is recalibrated. Finally, with complete knowledge of the structure, the sensor network is designed and installed.

The monitoring stage begins with the collection of data, made from two main sources: momentary long-term monitoring; and vibration events. The system is configured to perform measurements automatically on a fixed schedule, and according to data triggers such as threshold exceedances. The sensors used are mainly composed of tri-axial accelerometers, inclinometers, strain gauges, as well as temperature, humidity, and other weather-related sensors. When applying the SHM systems in bridges, a set of sensors for a weigh-in-motion measurement systems might also be used. After the raw data is collected, it is processed, with data cleansing and computations. In addition, there is a signal processing procedure, where transformation of vibration signal to its equivalent frequency domain occurs, using the fast Fourier transform. This includes evaluating the frequency response, with natural frequencies determined through a peak picking procedure [10]. The subsequent step is to make a characterization of the structure behavior, using all the pertinent data. This behavior characterization is done through an ML algorithm as described in the following subsection. Then, the study is directed to a time-series analysis of the structural behavior in each type of event. In particular, the variability of eigenfrequencies when the structure is classified as stationary, without vibrations induced by operations, is compared with historical baseline values. When potential damage is identified, structural computer modeling from the installation stage can be used and updated to assist in the evaluation of the location and severity of the damage. Afterwards, repair and maintenance procedures can be planned and performed with confidence.

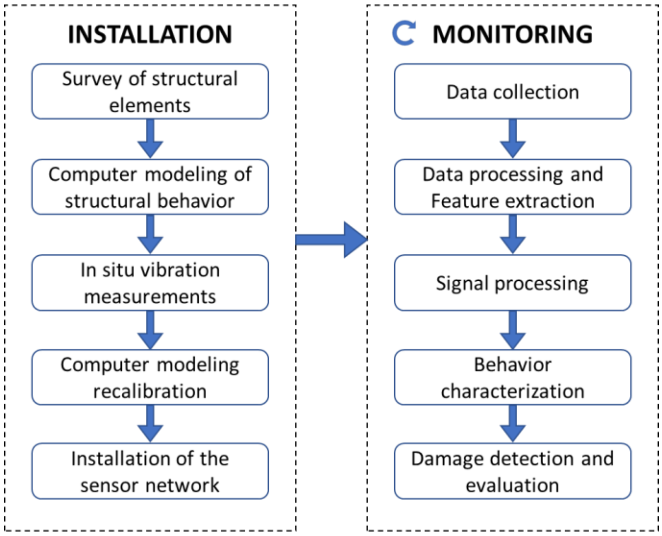

Figure 1. Flowchart of the proposed methodology for SHM

All the steps of the monitoring stage represented in Figure 1 are made automatically. Data is transferred and contained in a secure cloud storage, and can be accessed by system administrators, while visualized in a web-based platform. To this end, a dashboard app is made available to the owners and defined users of the structure. This web-app can be accessed through any browser, showing near real-time data about the structure, as well as its behavior characterization. In addition, an alert system sends direct warnings via SMS and email when severe actions on the structure are identified.

2.2 Behavior characterization algorithm

The algorithm developed in this paper for the characterization of structural behavior is inspired by semi-supervised learning techniques, aiming to obtain a classification of data both from “unlabeled data” and from a small set of “labeled data”. Therefore, an unsupervised k-means clustering technique is used, followed by an assignment technique to appropriately allocate a classification for unlabeled data.

Figure 2 shows the flowchart of the algorithm developed for behavior characterization. Once all the data is collected and processed, the data from all vibration events are used as an input for classification. Some specific events are easily identified and can be classified manually by the structure owners or defined users, or alternatively by the monitoring administrators. In parallel, the k-means clustering algorithm is processed assuming all data as unlabeled. Before applying the k-means clustering, data is normalized for all variables used as input.

The k-means clustering algorithm groups the observations available in k clusters. The cluster centers, called centroids, are randomly distributed at first, and each data point is assigned to the cluster of its nearest centroid in terms of Euclidean distance. Each centroid is then updated to be the average of the points assigned to it. The data points are reassigned to the new cluster and the cluster centroids are recalculated. These operations are repeated until a stop criterion is met, such as the centroids of the clusters do not change or a maximum number of iterations is reached.

The number of clusters, *k* value, can be set by users, which means that it should match the number of labels predefined by them. However, the value of *k* can also be optimized [11]. Indeed, there may be other events that have not yet been labeled during manual description. Therefore, in order to identify the appropriate number of clusters, the compactness of the clusters can be evaluated, for example, comparing the total between-cluster sum of squares with to the total sum of square. A higher value for this ratio value suggests a better compactness of individuals within cluster. Since this ratio increases with *k*, the 'elbow' method can be used to select the best *k*, which should occur at the location where the curve flattens markedly.

The assignment problem technique is one of the earliest applications of linear integer programming [12]. Its typical goal is to assign tasks to assignees, which can be people, machines, plants, etc. Different algorithms are available to solve assignment problems extremely efficiently as long some assumptions are satisfied: the number of assignees and the number of tasks must be the same; each assignee must be assigned to exactly one task; each task must be performed by exactly one assignee; and there must be a cost associated with an assignee performing a certain task. The objective is to determine how all assignments should be made to minimize the total cost. The Hungarian method is a well-known example of an extremely efficiently algorithm to solve the assignment problem [13].

The assignment problem can be used to assign the label of the manual description of events to the output of the k-means clustering, which are unlabeled clusters. Instead of a cost minimization problem, this assignment problem consists of maximizing the number of times that each label appears in a cluster. In summary, the label is assigned to the cluster where it was most named during the manual description. It should be expected that all the same labels belong to the same cluster. If the number of clusters, *k*, is larger than the number of unique labels, cluster that are not assigned to any label are identified as "unknown_1", "unknown_2", etc., and are marked to be manually labeled in the future.

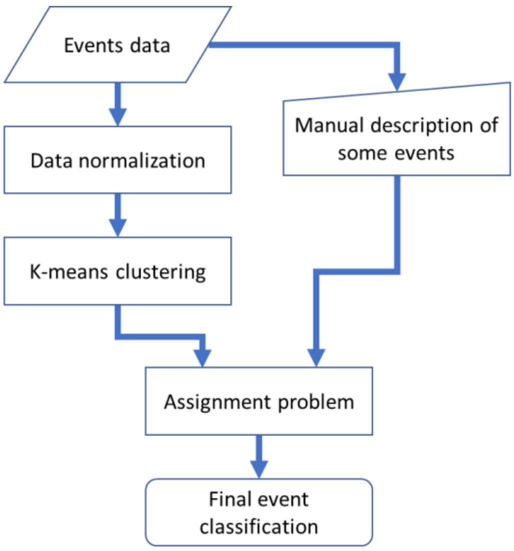

Figure 2. Flowchart of the algorithm for behavior characterization

3 CASE STUDY

The case study refers to a steel tower structure, 28 m high, located in the District of Coimbra, in Portugal (see Figure 3). The steel structure is used for industrial purposes and includes several mechanical equipment operating at varying schedules. The vibration induced during its activity influences the structural behavior and might contribute to issues such as fatigue and loss of connection stiffness.

3.1 Installation

3.1.1 Structural analysis

The preliminary stage was the analysis of the structure. After the first study and analysis of the structural documentation, and the consequent development of a computer modeling of the structure and its structural behavior, a set of in-situ vibrations measurements was carried out.

The measurements were performed with accelerometers, aimed at using the vibration signal obtained to determine the natural frequencies of the structure. To this end, tri-axial accelerometers contained in the sensors to be used during the monitoring were used, but also piezoelectric accelerometers to validate the quality of the solution. With the obtained data was also possible to analyze natural frequencies and the structural modes of vibration.

These in-situ measurements lead to a recalibration of the computer modeling of the structure. During this step, an additional external sound barrier wall made of sandwich panels was included, which had been inserted in one of the facades of the structure since the first computer modeling. The results showed similar vibration modes, with the natural frequencies differing less than 0.2 Hz. The computer modeling also allowed the knowledge of the structure and the locations with higher stress, assisting in the decision of the sensor network design.

3.1.2 Installation of the sensor network

The implemented SHM was optimized for the structure in order to provide meaningful and insightful data for the analysis of structural behavior.

The system consists of a wired network of 9 main locations, all including tri-axial accelerometers, temperature, and relative humidity sensors. The bottom 2 locations also include strain gauges. In addition, there is a weather station for environmental data, and a gateway to manage system data and send it through an LTE network. Figure 3 shows the industrial steel tower structure with the approximate location of the different sensors, gateway, and weather station.

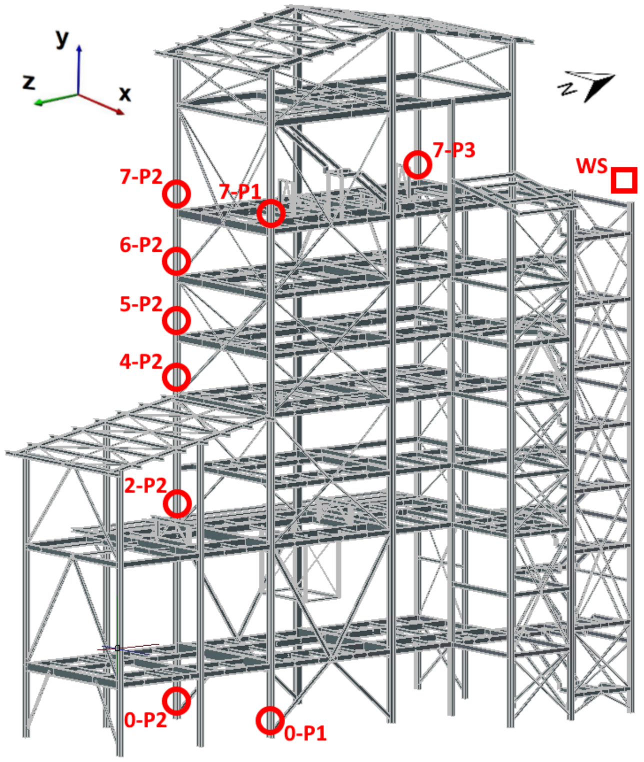

Figure 3. Industrial steel tower structure of the case study.

3.2 Main monitoring

Data collection is processed continuously on the administration servers. The transformed data, including some statistical parameters and signal processing results, is allocated on a specialized cloud server.

The output data can be visualized in a web-app. A concise report is delivered monthly to the client. When some thresholds exceed certain limits or the data is identified as dangerous events, the client receives alerts by SMS and email. Following are shown some examples of the outputs of this installation.

3.2.1 Long-term monitoring

This type of data is collected momentarily, at regular intervals of 5 minutes, with the aim of evaluating both actions and structural responses over time. In the steel structure of this case study, the data collected consists of weather data and strains. Figure 4 shows different long-term monitoring data at a seasonal scale. The data can also be visualized in more detail, such as in daily scale.

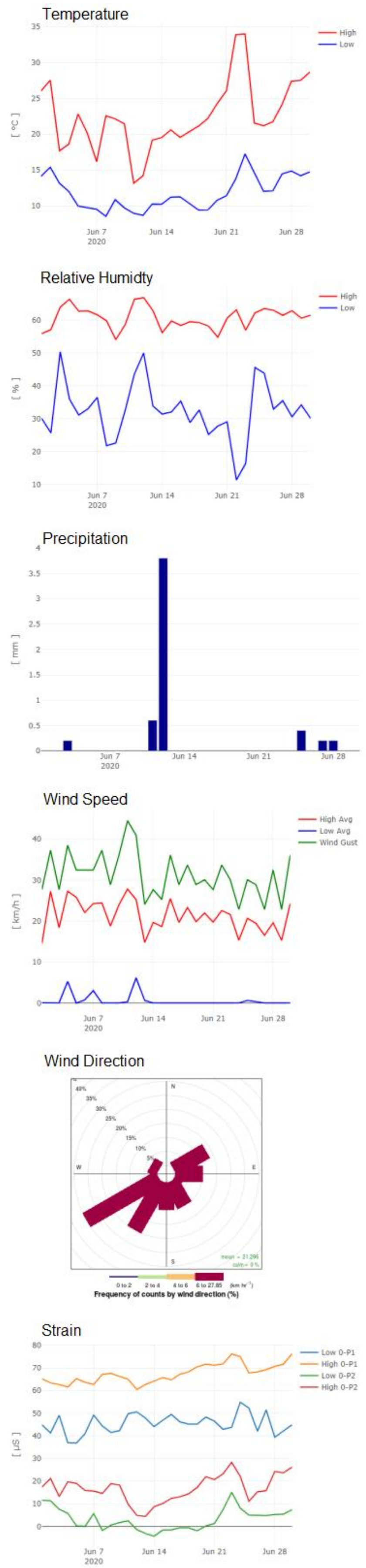

Figure 4. Example of long-term monitoring data

3.2.2 Events

Events data mainly consists of vibration data collected at regular intervals of 30 minutes. In threshold exceedances, vibration data is also collected. The thresholds are updated dynamically, either increasing or decreasing to preset values automatically by the system, avoiding onerous and redundant measurements.

During the collection of vibration data, weather data and strain data are also collected, to evaluate actions on the structure and its structural response.

Figure 5 shows an example of an event, which occurred after a threshold of strong wind was verified. During this event, the wind speed was around 60 km/h, with a NE direction.

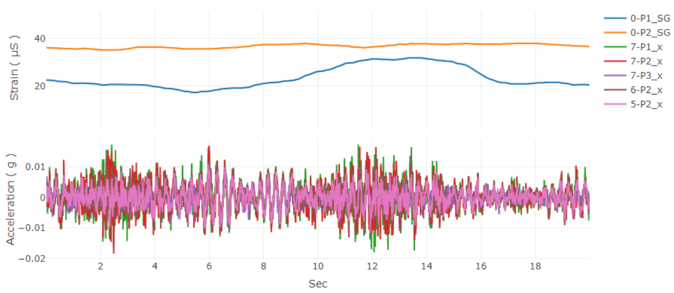

Figure 5. Accelerations and strains during a vibration event of strong wind

3.2.3 Frequencies analysis

The spectral analysis of the vibration signals is made for all vibration events through a peak picking procedure in the Fourier transform of the recorded signals. Together with the classification made through the ML algorithm, it is possible to analyze the natural frequencies of the structure while it is not affected by the operation of the industrial mechanical equipment. Figure 6 shows the Fourier transform of the x-axis direction in the sensors at the top of the structure.

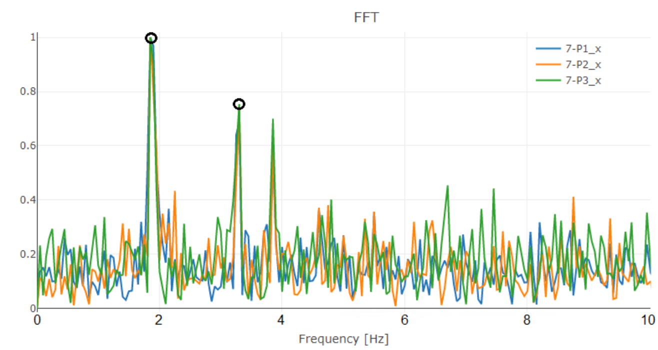

Figure 6. Frequency spectrum during an event

3.3 *Behavior characterization*

The Behavior characterization algorithm is executed regularly to classify the data and allow the update of the structural integrity evaluation. The variables used as input to the algorithm refer to all accelerations, strains and wind, for different sensor channels and for all vibration events.

In this case study, four main type of vibration events were labeled in the manual description. These events are: in operation, when all mechanical equipment is in normal operation; in partial operation, when only some of the mechanical equipment is operating; stationary, when there is no operation; and exceptional, when unusual conditions are observed. Figure 7 provides the plot of the explained variation as a function of the number of clusters. When applying the elbow method to the curve, it was defined the consideration of four clusters, coinciding with the unique clusters defined in the manual description. This elbow method should be updated occasionally, in particular when new manual descriptions are inserted.

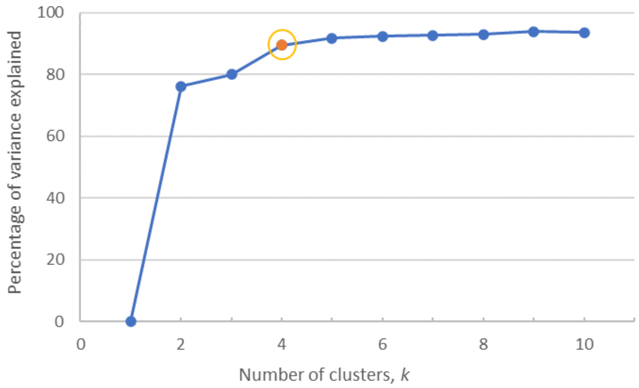

Figure 7. Elbow method applied for k-means clustering.

As the data used has several variables, it is not possible to plot the global clustering obtained. However, it is possible to draw a two-dimensional clustering plot showing the two components that most explain data variability, as presented in Figure 8. After the labels are assigned to the clusters, it is possible to identify, for example, the exceptional cluster events, located at the upper right red area of Figure 8.

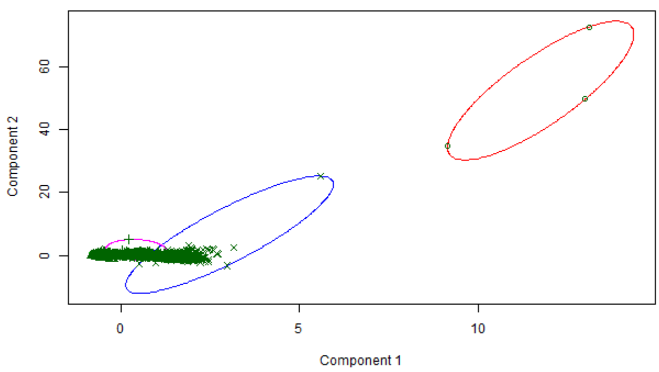

Figure 8. Clustering plot for the two principal components

3.4 *Structural integrity evaluation*

To assess the structural integrity, the evolution of the eigenfrequencies of the structure is analyzed. Given the classification provided by the behavior characterization, the analysis is made focusing on vibration events where the structure is classified as stationary, allowing a better assessment of natural frequencies. Figure 9 shows a time series of the evolution of the main peaks of natural frequencies measured in the sensor at 7-P2, observed since the beginning of structural monitoring. For instance, the first mode of vibration of the structure on this axis, around 1,90 Hz, has not shown any significant variation. The existing variability can be correlated with changes in environmental conditions.

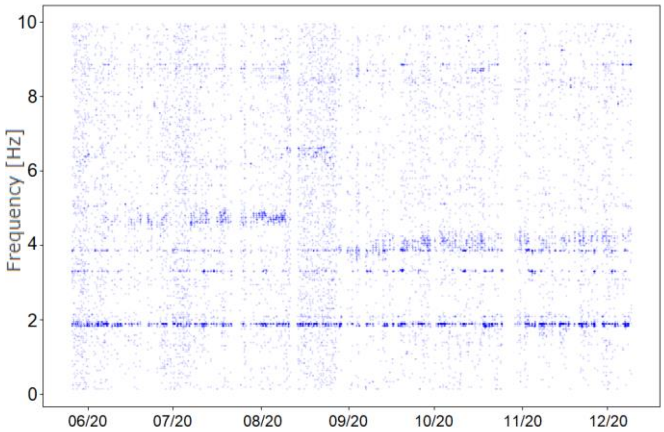

Figure 9. Second sample figure.

4 CONCLUSION

This paper presents a methodology for structural health monitoring applied to civil engineering structures. The methodology includes an ML approach, and was successfully implemented in the installation and monitoring system of an industrial steel tower structure, with a large amount of mechanical equipment. The approach was able to classify the different events occurring in the structure. A large set of vibration measurements over time allowed the visualization of the structural behavior in different situations

The proposed approach is also able to provide a real-time feedback on structural integrity. As the ML algorithm used is applied for every new event and more historical data and information is known about the classification of events, potential damage to the structure is more securely identified. To improve the ML accuracy, including the ability to identify the type and severity of damage, further experiments will be conducted, including laboratorial simulations of damage in similar structures. Future works will involve the development of more sophisticated analysis using power spectral functions matrices and frequency domain decomposition, aiming at improving and automatizing predictive modelling and damage detection.

ACKNOWLEDGMENTS

The authors are grateful for the support received from the Portugal 2020, through the program “Projetos de I&D - Projetos Individuais” 4/SI/2018, co-financed by the European Fund FEDER.

REFERENCES

[1] Kong, X., Cai, C.S., and Hu, J., *The state-of-the-art on framework of vibration-based structural damage identification for decision making*. Applied Sciences, 7(5), 497, 2017.

[2] Mitra, M., and Gopalakrishnan, S., *Guided wave based structural health monitoring: A review*. Smart Materials and Structures, 25(5), 053001, 2016.

[3] Dong, C.Z., and Catbas, F.N., *A review of computer vision–based structural health monitoring at local and global levels*. Structural Health Monitoring, in press, 2020.

[4] Khoa, N.L.D., Alamdari, M.M., Rakotoarivelo, T., Anaissi, A., and Wang, Y., *Structural health monitoring using machine learning techniques and domain knowledge based features*. In: Human and Machine Learning (pp. 409-435). Springer, Cham, 2018.

[5] Kurian, B., and Liyanapathirana, R., *Machine Learning Techniques for Structural Health Monitoring*. In: Proceedings of the 13th International Conference on Damage Assessment of Structures (pp. 3-24). Springer, Singapore, 2020.

[6] Kotsiantis, S.B., Zaharakis, I., and Pintelas, P., *Supervised machine learning: A review of classification techniques*. Emerging artificial intelligence applications in computer engineering, 160(1), 3-24, 2007.

[7] Gentleman, R., and Carey, V.J., *Unsupervised machine learning*. In Bioconductor case studies (pp. 137-157). Springer, New York, NY, 2008.

[8] Worden, K., Farrar, C.R., Manson, G., and Park, G., *The fundamental axioms of structural health monitoring*. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 463(2082), 1639-1664, 2007.

[9] Zhu, X., and Goldberg, A.B., *Introduction to semi-supervised learning*. Synthesis lectures on artificial intelligence and machine learning, 3(1), 1-130, 2009.

[10] Bendat, J.S., and Piersol, A.G., *Random data: analysis and measurement procedures* (Vol. 729). John Wiley & Sons, 2011.

[11] Tibshirani, R., Walther, G., and Hastie, T., *Estimating the number of clusters in a data set via the gap statistic*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2), 411-423, 2001.

[12] Eiselt, H.A., and Sandblom, C.L., *Integer programming and network models*. Springer Science & Business Media, 2013.

[13] Kuhn, H.W., *The Hungarian method for the assignment problem*. Naval research logistics quarterly, 2(1-2), 83-97, 1955.