Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5183-4_S02-06-619-cd

Wrapper Selection of Features for Fault Diagnostics of Truss Structures

Mohammad N. Juybari

Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran.
Energy Department, Politecnico di Milano,, Milan, Italy. E-mail: mnajafijuy@in.iut.ac.ir

Piero Baraldi*

Energy Department, Politecnico di Milano, 20156, Milan, Italy. E-mail: piero.baraldi@polimi.it

Antonio Palermo

Alma Mater Studiorum Università di Bologna, Bologna, Italy. E-mails: antonio.palermo6@unibo.it

Ali Eftekhari Milani

Energy Department, Politecnico di Milano, Milan, Italy.
Wind Energy Department, TU Delft, Netherlands. E-mail: A.EftekhariMilani@tudelft.nl

Alessandro Marzani

Alma Mater Studiorum Università di Bologna, Bologna, Italy. E-mail: alessandro.marzani@unibo.it

Enrico Zio

MINES Paris-PSL, CRC, Sophia Antipolis,
Energy Department, Politecnico di Milano, Milan, Italy. E-mail: enrico.zio@polimi.it

Truss structures are used in power systems to support pipelines and auxiliary equipment like pumps, utility stations, manifolds, firefighting equipment, and first-aid stations. The collapse of truss structures supporting pipelines carrying dangerous liquids or gases, such as those used in the petrochemical and chemical industries, can trigger accident chains. The diagnostics of damages in truss structures are, then, important to avoid catastrophic events that can cause severe consequences. In this context, we develop a method for fault diagnostics of truss structures. The method, which exploits the power spectral densities (PSD) derived from measured structural accelerations, is based on the two steps of feature selection and data classification. The feature selection task, which aims at identifying the set of features to be used as input of the diagnostic system, is here performed by a wrapper approach based on Multi-Objective Genetic Algorithms (MOGAs). The selected features are fed to a *k*-nearest neighbor (KNN) classifier for the identification of the damaged scenario of the truss structure. The developed fault diagnostic method is validated on several damage scenarios numerically simulated for an aluminum tower structure. The results show that the proposed approach is able to correctly recognize the damaged scenario with a limited number of misclassifications.

Keywords: Fault diagnostics, Feature selection, Wrapper approach, Multi-objective Genetic algorithm, k-nearest neighbours, Truss structure.

1. Introduction

Truss structures are well-accepted and cost-effective options for supporting structural loads in petrochemical and chemical plants, where they are used to support pipelines carrying dangerous liquids or gases. Truss structures are made by assemblies of beams connected at nodes. During their life they can be exposed to severe loads and degradation phenomena, which can cause their collapse and, therefore, trigger dangerous accident chains that can potentially cause catastrophic consequences. Hence, it is fundamental to detect the onset of abnormal conditions and diagnose the responsible beams (Fang, Luo, and Tang 2005).

The problem of anomaly detection in truss structures has been recently addressed in (Milani et al. 2021) by developing a method based on Principal Component Analysis (PCA) for signal reconstruction. In the present work, we consider the fault diagnostic problem of identifying which beam has caused the anomaly. The problem is addressed by performing the two sequential steps of feature selection and classification.

The objective of feature selection is to identify the quantities to be used as input to the model the identification of the defective beam. To this aim, acceleration signals measured by a set of sensors installed on the structure are processed to extract, in the frequency domain, the corresponding power spectral densities (PSDs) and the related features, namely frequency and intensity of peaks, corresponding to the estimated natural frequencies of the truss structure. Given the large number of features extracted in this way, a phase of selecting the subset of features which allows developing the most accurate classification model is required (Guha et al. 2020).

Feature selection methods are typically classified as wrapper, filter and embedded (Chandrashekar and Sahin 2014). Wrapper methods select an optimal subset of features using the classification model itself. Embedded methods perform feature selection as part of the training of the classification model. Filter methods rank the features according to their statistical association with the response.

This work develops a wrapper method given its superior performance in terms of achieved classification accuracy with respect to filter and embedded methods, despite the large computational effort that it typically requires.

Assuming the availability of n features, an exhaustive search among all possible 2^n combinations requires to train and test a dedicated classifier for each feature subset for performance evaluation, which is unfeasible (Amaldi and Kann 1998). Commonly used suboptimal searching strategies include forward selection, which starts with a small number of features and adds features until the classification model performance is decreased, and backward selection, which starts with all features and removes features until the performance is decreased (Kohavi and John 1997). Since the order of parameter entry (or deletion) may affect the final selection in sequential search strategies, in this work we develop an approach based on Genetic Algorithms (GAs) whose effectiveness for feature selection was demonstrated in (Zio, Baraldi, and Pedroni 2006). GAs main advantages are: i) relatively fast convergence to near-global optimum, ii) superior global searching capability in complicated search spaces, and iii) applicability when gradient information is not achievable, as in wrapper feature selection problems.

With respect to the second step, *i.e.* the identification of the beam responsible of the anomaly, we resort to the *k*-nearest neighbor (KNN) classification algorithm. The choice is justified by the fact that the feature selection algorithm requires the use of a classification algorithm characterized by few parameters to be tuned and fast computational times. In this context, possible choices are Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees (DT) and *k*-Nearest Neighbour (KNN) classifiers. SVMs have been shown to produce satisfactory performance when applied to datasets characterized by many classes (Gryllias and Antoniadis 2012), even when few labelled examples are available for training. On the other side, SVMs require the tuning of the kernel, which can be computationally intensive. ANNs dealing with nonlinear and multi-class classification problems, but they typically require a large amount of data for the ANN training (Li and Ma 1997) and also requires settings the ANN architecture and hyperparameters. Decision Trees (DTs) are simple to understand and to interpret and can handle multi-class classification problems (Parvin, MirnabiBaboli, and Alinejad-Rokny 2015). However, DT learners can create over-complex trees that do not generalize the data well, *i.e.*, they suffer from overfitting and can be unstable with small variations in the data resulting in the generalization of completely different trees. The KNN algorithm is intrinsically non-linear and only requires the setting of: 1) a single scalar parameter, *i.e.* the number k of near neighbours to be considered for the classification, and 2) the metric used to quantify the distance between patterns in the feature space. KNN classifiers have been employed in this work for their simplicity and low computational requirements (Baraldi et al. 2016).

The proposed method is validated on a dataset containing the features extracted from structures with damages in different beams. At present stage, the data have been simulated by using a Finite Element (FE) code properly developed to this aim.

The rest of the paper is organized as follows. Section 2 describes the considered case study. The applied methodology is detailed in Section 3, followed by the analysis of the obtained results in Section 4. Finally, Section 5 contains the conclusions and some suggestions for future works.

2. Case study

We consider an aluminum truss structure assembled in the LISG Laboratory of the University of Bologna. It is made of B=70 beams positioned in 5 cubic blocks of 1×1×1 m^3. A schematic view of the structure is shown in Fig. 1. The structure is part of the portfolio of test cases of the “Manutenzione intelligente (smart maintenance) di impianti industriali e opere civili mediante tecnologie di monitoraggio 4.0 e approcci prognostici” (MAC4PRO) project supported by “Instituto nazionale Assicurazione Infortuni sul Lavoro” (INAIL) within BRIC2018. The beams of the structure are characterized by a hollow circular cross-section with an internal diameter of $d_i = 0.036\,m$ and an external diameter of $d_e = 0.042\,m$, for an overall area of $A = 3.67 \times 10^{-4}\,m^2$. Quasi-spherical nodes are used at the end sections of the beams to provide moment-free connections. The truss structure is constrained at the four nodes placed at its base by means of elastic pads (elastic supports).

A FE-based truss code has been developed within the MAC4PRO project to simulate the dynamic behavior of the structure. The i-th beam, $i = 1,..,B$, of the structure is modeled as a linear member with an equivalent Young’s Modulus E_i, length L_i, cross-section $A_i = A$, and density $\rho_i = \rho$. All diagonal beams are $L_d = 1.42\,m$ long, while the length of horizontal and vertical beams is $L = 1\,m$.

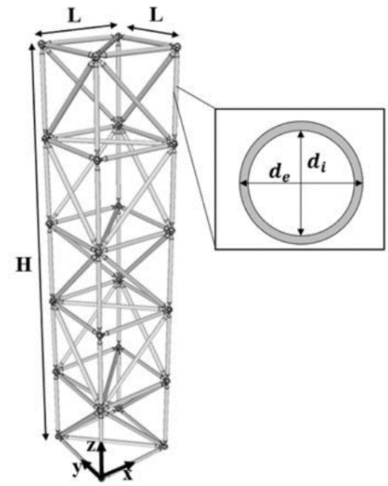

Fig. 1. Schematic view of the truss structure

The Young’s Modulus has been estimated by laboratory tests on selected truss elements while considering the stiffness contribution of the nodes. Accordingly, the Young’s Modulus is set equal to $E_d = 51.94\,GPa$, for the diagonal beams, and to $E_{v,h} = 48.7\,GPa$ for the vertical and horizontal beams. Each beam element is modeled via stiffness matrix with no contributions for ends rotation (bar elements). The material density is equal to $\rho = 3130\ kg/m^3$, where the mass contribution of the spherical nodes is also considered in the density calculation. A lumped mass matrix is used for modelling the mass contribution of each beam.

Time-transient analysis are performed to simulate the response of the truss structure under ambient vibrations. To this aim, a gaussian white noise input is applied at the base of the truss. The transient simulations are performed for a duration of t=300 s, with a sampling frequency fs=1024 Hz using the Newmark time integration scheme.

Acceleration time-histories are computed at each free node of the truss along the three spatial directions (*i.e.*, x, y and z).

The recorded accelerations are then processed to compute the PSD. To this aim, the data are down-sampled at a sampling frequency fs=256 Hz and, then, the Welch method with a window length of 20 s and an overlap of 50% is adopted to estimate the PSDs with reduced noise.

The developed procedure allows computing the PSDs of the acceleration signals for each considered cases, either defective or healthy.

In particular, a dataset accounting for variations in the structural mechanical properties due to environmental and operational conditions (e.g., ambient temperature), is simulated by sampling the following parameters:

1) The elastic supports stiffness along the z direction is taken as $\tilde{k}_g$, which is assumed to be uniformly distributed in $[k_g - 0.05k_g; k_g + 0.05k_g]$.
2) The Young modulus $\tilde{E}_i$ of each beam, which is assumed to be normally distributed $\tilde{E}_i \sim N(E_i, 0.05E_i)$.

For the simulation of damaged structures, it is assumed that the damage will lead to a reduction in the stiffness parameter E_iA_i of the defective beam. In this work, we consider scenarios with a single damaged beam at a time in each truss structure. The damage index of beam i is:

$$DI_i = 1 - \frac{E_i(d)A_i(d)}{E_i(h)A_i(h)}, \qquad i = 1, \dots, B \quad (1)$$

where h and d represent a healthy and damaged structure, respectively. In this structures, damages to 10 different defective beams and with $DI_i = 40\%$ are simulated. The simulation process is repeated 100 times to consider different operational conditions for each simulated damaged structure. The overall dataset $\boldsymbol{X}_D = \{\boldsymbol{x}_D^{(d)}, d = 1, .., D\}$ contains $D = 1000$ simulations of damaged structures of 10 different classes, where each class corresponds to a different defective beam. For each simulation, the PSDs of the acceleration signals in directions x, y and z in correspondence of 24 nodes are obtained. Fig. 2 shows, as example, the PSD of the simulated acceleration signal along direction x on a node of the first floor of the structure.

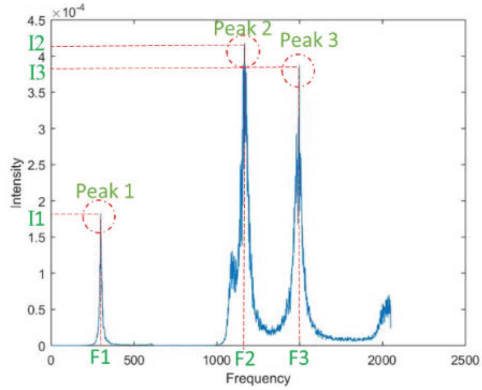

Fig. 2. Example of PSD of the simulated acceleration signal in direction $\boldsymbol{x}$ on a node at the first floor

3. Problem statement

The objective of the present work is to develop a method to identify which is the defective beam in a truss structure by using accelerometers' records.

The problem is addressed by performing the two sequential steps of feature selection and classification. In this work, feature selection considers the possible locations of the accelerometers and the possible features that can be extracted from the PSDs. Regarding the location, we assume that accelerometers measuring vibrations along the x, y, and z axes can be installed in correspondence of the nodes. Regarding the feature extraction, the idea is based on the fact that frequency and intensity of the highest peaks of the PSD contain information on structural response status, healthy or damaged, since they can evidence variation in the natural frequencies of the structure. Considering the simulated data, 174 significant peaks have been identified in the 64 PSDs collected at the 24 nodes along the x, y and z axes (note that the nodes at the base are restrained along the x and y directions but can move along the z direction, being supported by the elastic spring of stiffness k_g). Therefore, since the intensity and frequency features are extracted from each peak, the total number of features is $174 \times 2 = 348$.

4. Method

The method developed in this work to address the problem formulated in Section 3 is based on the two sequential steps of feature selection and classification.

Section 4.1 will describe the developed wrapper feature selection method, whereas Section 4.2 will illustrate the k-Nearest Neighbor algorithm used for the classification.

4.1. Wrapper feature selection

A wrapper feature selector is based on a search engine, which builds candidate groups of features, whose performances are evaluated by using of properly defined fitness functions. A solution is represented by the Boolean vector $\boldsymbol{x} \in [0, 1]^{348}$ whose i-th component, x_i, is equal to 1 if the i-th feature belongs to the solution and to 0 if it does not belong to the solution. Therefore, the lengths of vector $\boldsymbol{x}$ is equal to 348, i.e. the total number of features extracted from the PSDs.

In this work, we consider as objectives the minimization of the following three fitness functions:

$$\boldsymbol{F_1}(\boldsymbol{x}) = \textbf{Misclassification error}$$
$$\boldsymbol{F_2}(\boldsymbol{x}) = \boldsymbol{N}\textbf{umber of selected sensors}$$
$$\boldsymbol{F_3}(\boldsymbol{x}) = \textbf{Computational time}$$

Notice that the computation of the fitness functions F_1 and F_3 requires the development of the classifier fed by the features selected by $\boldsymbol{x}$ and its application to a set of test data, which can be time and resources consuming.
Since we deal with a Multi-Objective Optimization (MOO) problem, the final objective is the identification of the Pareto Optimal Set, $\mathcal{P}^* = \{\boldsymbol{x} \in \boldsymbol{F}(\boldsymbol{x})\ is\ Pareto - optimal\}$, i.e., the set of optimal solutions. A vector of decision variable $\boldsymbol{x}_{opt} \in \mathcal{F}$ is Pareto optimal if it is non-dominated with respect to $\mathcal{F}$, i.e., it does not exist another solution $\boldsymbol{x}' \in \mathcal{F}$ such that $\boldsymbol{F}(\boldsymbol{x}')$ dominates $F(\boldsymbol{x}^*)$:

$$\forall \alpha \in \{1,2\},\ F_\alpha(\boldsymbol{x}') \leq\ F_\alpha(\boldsymbol{x}^*), and$$

$$\exists \tilde{\alpha} \in \{1,2\}, such\ that\ F_{\tilde{\alpha}}(\boldsymbol{x}') \leq\ F_{\tilde{\alpha}}(\boldsymbol{x}^*) \quad (2)$$

The wrapper approach developed in this work uses a NSGA-II Multi Objective Genetic Algorithm (MOGA), which is based on the main steps of population initiation, crossover, mutation, and selection. The reader interested in more details on NSGA-II can refer to (Deb et al. (2002); Singh et al. (2017)). The chromosomes are encoded by the binary vectors, $\boldsymbol{x}$, of the possible solutions. Table 1 reports the parameters used for the NSGA-II search.

Table 1. Parameters of the NSGA-II

Parameter	Value
Max generation	100
Initial population	100
Mutation rate	0.7
Crossover rate	0.4

4.2. Classification

The classification algorithm used in this work is the k-Nearest Neighbor. According to this algorithm, the classification of a test pattern is based on the computation of its distance with all the patterns of the training set and the identification of the k closest patterns. Then, the class assigned to the test pattern is the class with the largest number of representatives among those of its k neighbours (Fix and Hodges 1989). The only parameter of the algorithm is the number, k, of nearest neighbours to be considered. In (Hellman 1970) the empirical rule of setting $k = \sqrt{N}$, where N is the number of training patterns, is suggested. Since the number of patterns available in this work is 1000, k has been set equal to 32. With respect to the metric for computing the distance among the patterns, the Euclidean distance has been used.

5. Results

The data of the simulated $D = 1000$ structures $\boldsymbol{X_D} = \{\boldsymbol{x_D^{(d)}}, d = 1,..,D\}$ are randomly divided into a training and a validation set, made by 75% and 25% of the patterns, respectively. The training set is used for training the KNN classifiers and for feature selection, whereas the validation set is used for the evaluation of the performance of the developed method on those data not used for model development. Regarding the feature selection, which requires itself the training and test of classifiers, the training set is split into 5 folds, where 4 of them are used for training and one for testing the selected features. Note that, the process is repeated 5 times using different test sets for a robust evaluation of the MOGA fitness function.

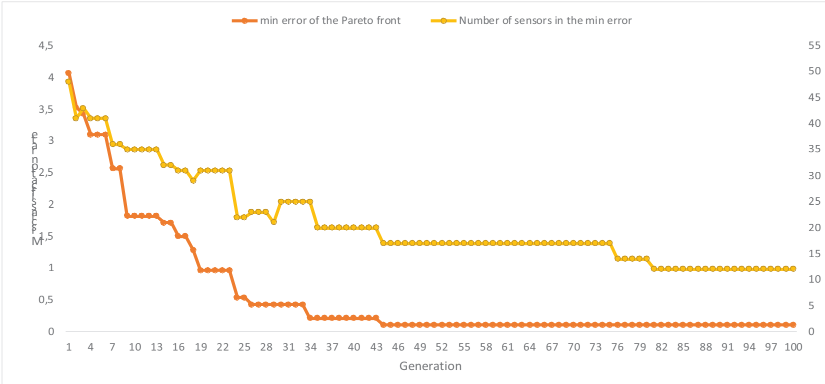

Fig. 3. Evolution during the MOGA search of the misclassification rate and of the number of sensors of the Pareto optimal set solution which has the smallest misclassified rate

5.1. Feature selection results

Fig. 3 shows the evolution during the MOGA search of the misclassification rate and of the number of sensors of the Pareto optimal set solution which has the smallest misclassified rate. We remark that: 1) the misclassification rate significantly reduces from 4.06% in generation 1 to 0.11% in generation 100. At the same time, the number of sensors needed to measure the features used for the classification is reduced from 48 in generation 1 to 12 in generation 100. 2) The solution is not changing after generation 80, which indicates the algorithm convergence.

Fig. 4 shows the obtained Pareto optimal set. Note that the two objectives related to the computational time and the number of sensors tend to be correlated, since a solution with a large number of sensors tend to have several features and, therefore, requires long computational times. On the contrary in Fig. 5, the objectives of reducing misclassification rate and number of sensors tend to be conflictual. Since, as expected, solutions based on less features tend to provide less information for the classification. Finally, the range of values of the computational time in the Pareto optimal set is relatively small and very limited, *i.e.*, between 0.009 s and 0.0105 s, which indicates that the corresponding objective is satisfactorily met by all solutions.

Once the Pareto optimal set is obtained, the solution to be used by the diagnostic system should be selected according to the decision maker desired trade-off among the objectives. In this regards, Table 2 reports:

1. the solution of the Pareto optimal set providing the smallest misclassification rate;

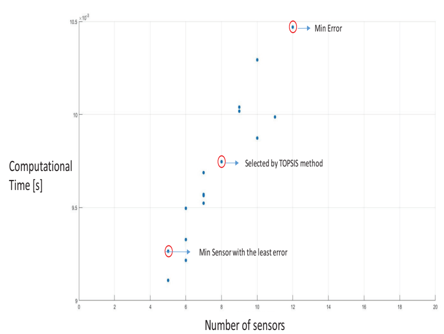

Fig. 4. Projection of the final Pareto optimal set on to the subspaces defined by the fitness function "Number of sensors" and "Computational time"

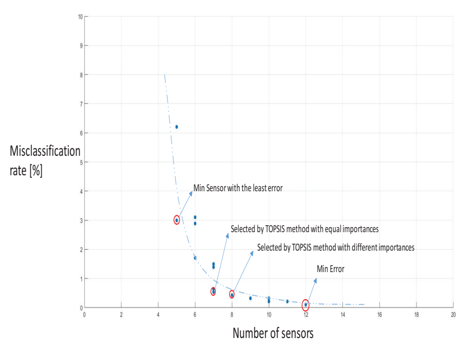

Fig. 5. Projection of the final Pareto optimal set on to the subspaces defined by the fitness function "Number of sensors" and "Misclassification rate"

2. the solution of the Pareto optimal set which requires to install the smallest number of sensors;
3. the solution of the Pareto optimal set selected by the TOPSIS method (Hwang and Yoon 1981) using as importance weights of the three objectives $w_{F_1} = 0.6, w_{F_2} = 0.3$, and $w_{F_3} = 0.1$. This choice allows prioritizing classification accuracy and number of sensors to install and reducing the importance of the computational time that is acceptable for all solutions.

Table 2. Solutions of the Pareto optimal set

Method	# Sensors	Computational time s	Misclassification error %
Min sensor	5	0.0093	5.04
TOPSIS	8	0.0097	0.8
Min error	12	0.0105	0.24

Table 3. Details of the TOPSIS solution

Sensor	Node	Intensity or/and Frequency	Direction
1	1	Intensity	z
5	5	Intensity	x
6	5	Intensity	y
11	7	Intensity	x
13	7	Intensity and Frequency	z
16	8	Intensity	z
42	17	Frequency	y
45	18	Intensity	y

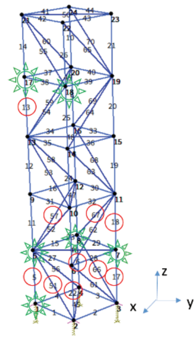

Fig. 6. Location of the sensors selected by the TOPSIS solution (green circles) and beams that are damaged in the 10 classes of faults (red circles)

Table 3 reports the features selected by the TOPSIS solution and Fig. 6 shows the locations of the selected sensors. Notice that 4 out of 6 of the selected sensors are in the first two blocks of the structure, close to 8 of the 10 beams that are damaged by the considered fault classes.

5.2. Classification results

The KNN classifier developed using the features of the TOPSIS solution is used to classify the 250 structures of the validation set. Fig. 7 shows the confusion matrix and reports the obtained misclassification rates for each class of the faults. Note that all structures with defective beams A5, A6, A13, A17, A22 and A51 are correctly classified.

The few misclassifications include: 1) damages of beams A57 and A67, which are both diagonal beams in the second block; and 2) damages of beams A67 and A6. Since A67 is commonly mislabeled in both pairs, the use of a further sensor located near to beam A67 is expected to reduce the misclassifications.

True Classes: Direction	Block	Beam	A5	A6	A13	A17	A18	A22	A51	A57	A66	A67	Total number of samples	Accuracy(%)
Vertical	1	A5	20										20	100
Vertical	1	A6		23									23	100
Vertical	4	A13			19.0								19	100
Vertical	1	A17				19							19	100
Vertical	2	A18					29.8					0.2	30	99
Vertical	1	A22						36					36	100
Diagonal	1	A51							28				28	100
Diagonal	2	A57								20.4		0.6	21	97
Diagonal	1	A66					0.2				23.8		24	99
Diagonal	2	A67		1								29.0	30	97
			Predicted Classes											

Fig. 7. Confusion matrix of the TOPSIS solution

6. Conclusion

We have developed a fault diagnostic method for the identification of the damaged beams in truss structures. The number of sensors to be installed and their locations have been optimized by using a wrapper feature selection approach, where a MOGA has been used as search engine and the *k*-Nearest Neighbor algorithm as classifier. The method has been applied to data generated using a FE code properly developed to this aim. The results show that a satisfactory classification accuracy is obtained by a diagnostic system which requires the installation of a limited number of sensors. Specifically, a misclassification rate of 0.8% is obtained by a *k*-Nearest Neighbor classifier fed by 9 features of intensity and frequency of peaks extracted from the PSDs of acceleration signals measured by 8 sensors.

Acknowledgment

The participations of Alessandro Marzani, Ali Eftekhari Milani, Antonio Palermo, Enrico Zio, Mohammad N. Juybari and Piero Baraldi have been funded by "Smart maintenance of industrial plants and civil structures by 4.0 monitoring technologies and prognostic approaches – MAC4PRO", sponsored by the call BRIC-2018 of the National Institute for Insurance against Accidents at Work – INAIL.

References

Amaldi, E, and V Kann. 1998. "On the Approximability of Minimizing Nonzero Variables or Unsatisfied Relations in Linear Systems." *Theoretical Computer Science* 209 (1–2): 237–60. https://doi.org/10.1016/S0304-3975(97)00115-1.

Baraldi, Piero, Francesco Cannarile, Francesco Di Maio, and Enrico Zio. 2016. "Hierarchical K-Nearest Neighbours Classification and Binary Differential Evolution for Fault Diagnostics of Automotive Bearings Operating under Variable Conditions." *Engineering Applications of Artificial Intelligence* 56: 1–13. https://doi.org/https://doi.org/10.1016/j.engappai.2016.08.011.

Chandrashekar, Girish, and Ferat Sahin. 2014. "A Survey on Feature Selection Methods." *Computers & Electrical Engineering* 40 (1): 16–28. https://doi.org/https://doi.org/10.1016/j.compeleceng.2013.11.024.

Fang, X, H Luo, and J Tang. 2005. "Structural Damage Detection Using Neural Network with Learning Rate Improvement." *Computers & Structures* 83 (25): 2150–61. https://doi.org/https://doi.org/10.1016/j.compstruc.2005.02.029.

Fix, Evelyn, and Joseph Lawson Hodges. 1989. "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties." *International Statistical Review/Revue Internationale de Statistique* 57 (3): 238–47.

Gryllias, Konstantinos C, and Ioannis A Antoniadis. 2012. "A Support Vector Machine Approach Based on Physical Model Training for Rolling Element Bearing Fault Detection in Industrial Environments." *Engineering Applications of Artificial Intelligence* 25 (2): 326–44.

Guha, Ritam, Manosij Ghosh, Shyok Mutsuddi, Ram Sarkar, and Seyedali Mirjalili. 2020. "Embedded Chaotic Whale Survival Algorithm for Filter–Wrapper Feature Selection." *Soft Computing* 24 (17): 12821–43. https://doi.org/10.1007/s00500-020-05183-1.

Hellman, Martin E. 1970. "The Nearest Neighbor Classification Rule with a Reject Option." *IEEE Transactions on Systems Science and Cybernetics* 6 (3): 179–85.

Hwang, Ching-Lai, and Kwangsun Yoon. 1981. "Methods for Multiple Attribute Decision Making." In *Multiple Attribute Decision Making*, 58–191. Springer.

Kohavi, R, and G H John. 1997. "Wrappers for Feature Subset Selection." *Artificial Intelligence* 97 (1–2): 273–324. https://doi.org/10.1016/s0004-3702(97)00043-x.

Li, C James, and Jun Ma. 1997. "Wavelet Decomposition of Vibrations for Detection of Bearing-Localized Defects." *Ndt & E International* 30 (3): 143–49.

Milani, Ali Eftekhari, Piero Baraldi, Antonio Palermo, Alessandro Marzani, and Enrico Zio. 2021. "Damage Detection in Truss Structures Supporting Pipelines and Auxiliary Equipment in Power Plants." In *2021 5th International Conference on System Reliability and Safety (ICSRS)*, 115–19. IEEE.

Parvin, Hamid, Miresmaeil MirnabiBaboli, and Hamid Alinejad-Rokny. 2015. "Proposing a Classifier Ensemble Framework Based on Classifier Selection and Decision Tree." *Engineering Applications of Artificial Intelligence* 37: 34–42.

Zio, E, P Baraldi, and N Pedroni. 2006. "Selecting Features for Nuclear Transients Classification by Means of Genetic Algorithms." *IEEE Transactions on Nuclear Science* 53 (3): 1479–93. https://doi.org/10.1109/TNS.2006.873868.