Universidade Estadual Paulista - UNESP
Faculdade de Engenharia de Ilha Solteira - FEIS
Departamento de Engenharia Mecânica - DEM
Programa de Pós Graduação em Engenharia Mecânica - PPGEM

Information Gain XGBoost for Feature Selection in structural health monitoring of bridges

Research Local: Departamento de Engenharia Mecânica, Faculdade de Engenharia de Ilha Solteira (FEIS), Universidade Estadual Paulista (UNESP).

Candidate: Heitor Nunes Rosa, h.rosa@unesp.br
Advisor: Elói João Faria Figueiredo, eloi.figueiredo@ulusofona.pt

Ilha Solteira, August 2024.

FICHA CATALOGRÁFICA
Desenvolvido pela Diretoria Técnica de Biblioteca e Documentação

R788i

Rosa, Heitor Nunes.
Information gain XGBoost for feature selection in structural health monitoring of bridges / Heitor Nunes Rosa. -- Ilha Solteira: [s.n.], 2024
38 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Mecânica dos Sólidos, 2024

Orientador: Eloi Joao Faria Figueiredo
Inclui bibliografia

1. Structural health monitoring. 2. Decision tree learning. 3. Feature importance. 4. Information gain. 5. XGBoost.

Elaborado por Raiane da Silva Santos - CRB-9/9999

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Ilha Solteira

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Information Gain XGBoost for Feature Selection in structural health monitoring of bridges

AUTOR: HEITOR NUNES ROSA
ORIENTADOR: ELÓI JOÃO FARIA FIGUEIREDO

Aprovado como parte das exigências para obtenção do Título de Mestre em Engenharia Mecânica, área: Mecânica dos Sólidos pela Comissão Examinadora:

Prof. Dr. ELÓI JOÃO FARIA FIGUEIREDO (Participaçao Virtual)
Engenharia Mecanica / Universidade Lusofona de Humanidades e Tecnologias

Prof. Dr. MARCIO ANTONIO BAZANI (Participaçao Presencial)
Departamento de Engenharia Mecânica / Faculdade de Engenharia de Ilha Solteira - UNESP

Prof. Dr. ADAM DREYTON FERREIRA DOS SANTOS (Participaçao Virtual)
Faculdade de Sistemas de Informação / Universidade Federal do Sul e Sudeste do Pará - UNIFESSPA

Ilha Solteira, 29 de agosto de 2024

Abstract

Recent efforts in structural health monitoring (SHM) have focused on strategies utilizing statistical pattern recognition, where extracting damage-sensitive features from raw data is critical. The large volume of features generated from multiple sensors and signals often includes irrelevant data and noise, challenging accurate damage detection. Selecting the most informative features is paramount to transit bridge SHM from research to practice. This paper contributes to the bridge SHM field by introducing a decision tree-based information gain approach to automate feature selection, enhancing sensitivity to damage. The study uses data from the Z24 Bridge, analyzing its undamaged state and eight different damaged state conditions. Features were extracted using frequency- and time-domain techniques. The proposed approach was compared with two other approaches, based on permutation importance and F-statistics, demonstrating a superior trade-off in classification performance, identifying a smaller and more effective subset of features. An XGBoost classifier was used to overcome the challenges of missing data inherent to field monitoring. The time-domain features were enhanced for real-world bridge applications and proposed as an alternative to traditional frequency-domain features. Additionally, a sensor-level sensitivity analysis revealed new possibilities for applying tree-based algorithms in SHM sensor analysis, underscoring the practical impact of this method even for damage localization.

Keywords: Structural Health Monitoring, Decision Tree Learning, XGBoost, Feature Importance, Information Gain.

List of Figures

List of Tables

Summary

1 Introduction

1.1 Motivation

Bridges are essential for ensuring connectivity within large road and rail networks, and their structural deterioration has increased the need for effective monitoring strategies to optimize maintenance and prevent failures (WENZEL, 2008). Over the past three decades, structural health monitoring (SHM) has emerged as a key strategy for damage identification, including detection, localization, type, severity, and prognosis. When the goal is to detect and localize damage, it normally relies on the statistical pattern recognition (SPR) paradigm to differentiate between undamaged and damaged structural conditions under varying operational and environmental factors (FIGUEIREDO; BROWNJOHN, 2022). As illustrated in Figure 1, data-based models within this paradigm have been proposed to address the limitations of physics-based models and the lack of long-term experimental data, using statistical techniques directly on response data to assess the structural condition of bridges (FARRAR; WORDEN, 2013).

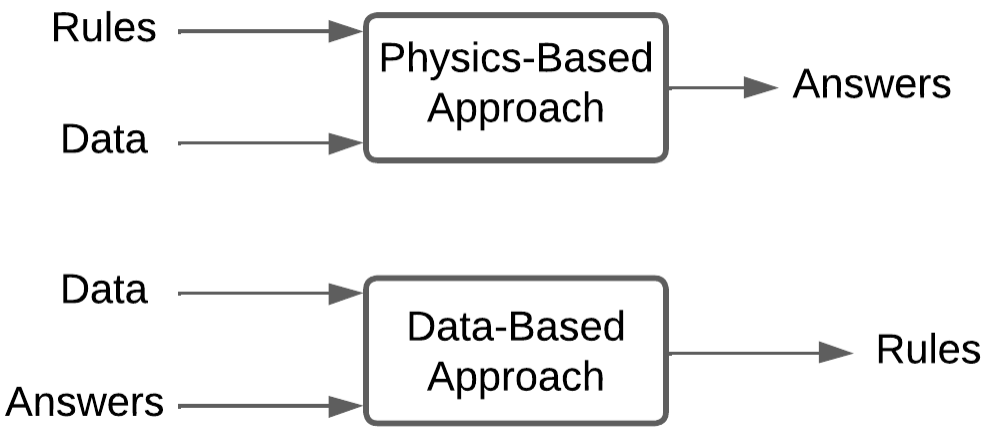


Fig. 1 – Conceptual comparison between physics- and data-based models.

Source: Adapted from Chollet (2021).

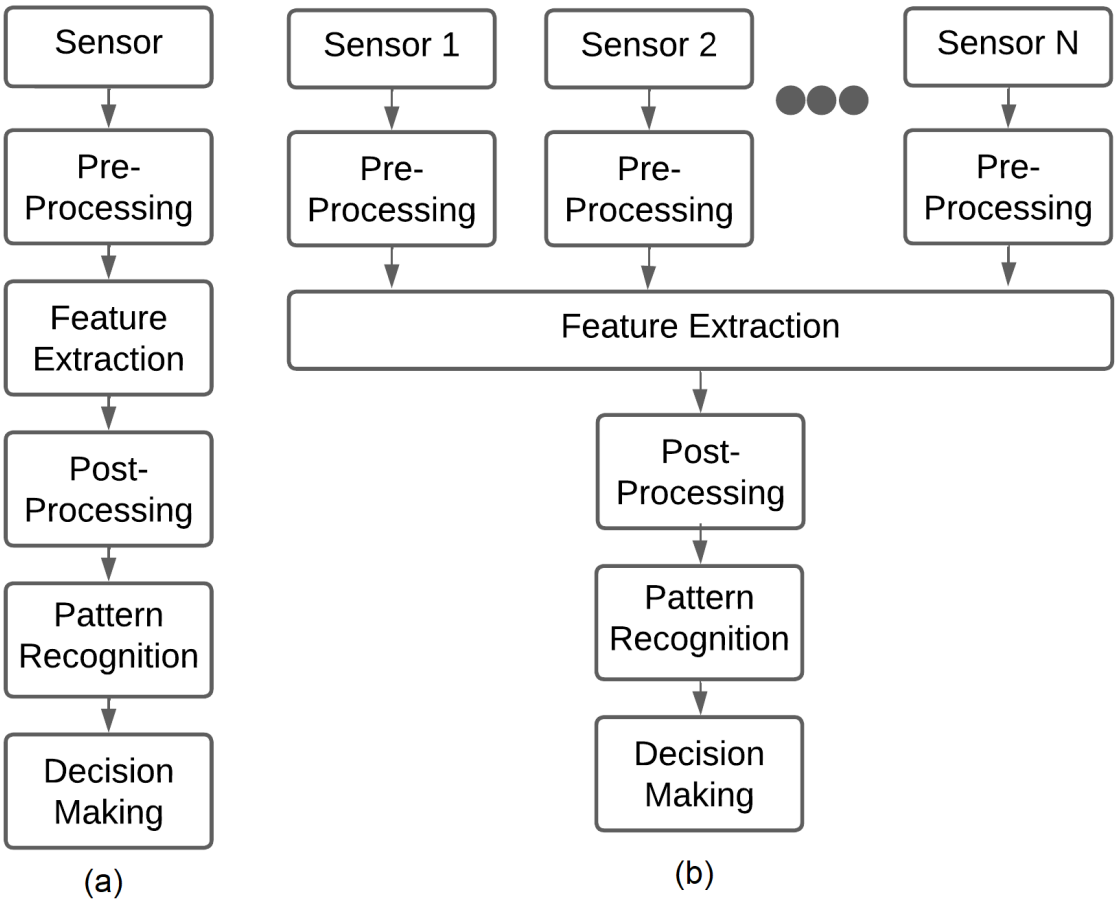
In the SPR paradigm, feature extraction is an essential stage. This process generally applies to signals transforming raw data into damage-sensitive features that distinguish between different conditions. It can be done using multivariate and univariate methods, as schematized in Figure 2. Multivariate methods use signals from multiple sensors mounted in a structure to capture its dynamic characteristics. Frequent-domain features, such as modal properties (natural frequencies, damping ratios, and mode shapes), are identified, giving a more global

sensing perspective. Many authors have applied multivariate methods to detect damage in bridges, particularly by using natural frequencies estimated from multiple sensors distributed across the structures (BUD et al., 2024; NOVAIS; SILVA; FIGUEIREDO, 2024). Even though modal properties have been used for damage detection for several decades, they still strive to be sensitive to identify damage globally. There are examples where the sensitivity arises only when damage is relatively severe (FARRAR et al., 1994). They generally fail information regarding damage localization or sensor sensitivity for a given damage, especially under operational and environmental conditions (PEETERS; ROECK, 2001; KIM et al., 2003; SOYOZ; FENG, 2009). Besides, the transformation of signals from time to frequency domain can result in information losses, primarily the non-stationary parts of signals that vary with time (A. . . , 2023). On the other hand, univariate methods can independently extract multiple time-domain features from sensors' signals, which can potentially be more sensitive to local information from the whole structure. A review of time-domain techniques found that these techniques are mostly used in linear systems while they are highly sensitive to noise (VIBRATION. . . , 2022). Time-domain signals can effectively display the temporal characteristics of vibration as they change over time. They may also be related to damage, particularly in the case of minor or local damage.

Numerous statistical features can be extracted from a single signal. Given that some structures have multiple sensors, the number of extracted features increases immensely, most irrelevant, highly correlated, or noisy. Those features can be removed without a significant loss of information. This issue is represented in Figure 3. The pair of features on the left is less useful than the one on the right for a classification algorithm to distinguish between red and black classes. Therefore, the set of features from the left can be eliminated from the classification.

Feature selection is fundamental to finding informative features. This process can be done through unsupervised or supervised learning strategies. In the unsupervised strategy, the selection discards features that could be less informative for classification, such as those with high correlation, low variance, missing data, and spurious values. Such features do not necessarily promote improving the classifier's performance, but their removal helps reduce dimensionality. In the supervised strategy, the feature selection is done when knowledge of the varying damage states or classes is available to identify the subset of informative features that best discriminate different classes. As long as an optimal feature subset is identified, it could be used as input for unsupervised novelty detection algorithms.

Feature selection is paramount to transition SHM from research to practice. Recently, Buckley et al. (BUCKLEY; GHOSH; PAKRASHI, 2023) showed that features extracted from



Source: Adapted from Farrar and Worden (2013).

Fig. 2 – Univariate (a) and multivariate (b) methods for feature extraction process schematization.

univariate, single-sensor acceleration signals could accurately distinguish between multiple damage states in full-scale structures, such as Z24 and S101 Bridges. Using various feature selection methods (permutation importance, F-statistics, correlation clustering), subsets of these features were obtained, which achieved similar classification performance of damage states to the entire feature set. The classification algorithms were based on supported vector machines, random forest, k-nearest neighbor, and Naive Bayes. The authors also showed that hierarchical clustering effectively reduces the feature set without much loss in classification performance by removing correlated features.

However, there is still a need for more studies to validate the applicability of time-domain features to detect damage as an alternative to traditional natural frequencies or to improve the damage detection along with them. Therefore, as a major contribution, this paper proposes a decision tree-based information gain approach for feature selection, considering the sensitivity of features to damage. It compares the approach with consolidated ones, such as permutation

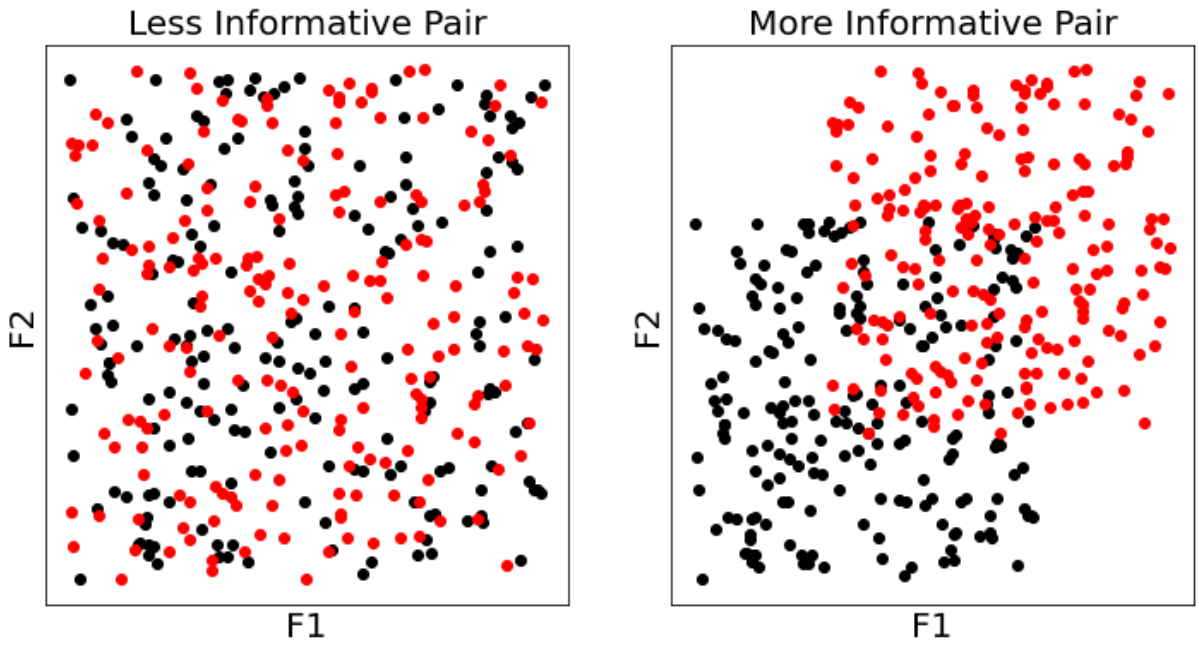


Fig. 3 – Representation of pairs of features that are more or less informative for the classification task.

importance and F-statistics. The classification is conducted through XGBoost, as it handles the missing data challenge that is present most of the time in real-world bridge applications, like bridges, without the need for imputation methods. As a result, time-domain features for bridges are unveiled as alternative features to the frequency-domain ones (e.g., natural frequencies). After the feature selection, a sensor-level sensitivity analysis is conducted using decision tree-based structures, revealing new opportunities for applying it to sensor sensitivity analysis in SHM. Unlike Buckley et al. (BUCKLEY; GHOSH; PAKRASHI, 2023), this research highlights time-domain features as an alternative to frequency-domain features, showing their effectiveness for real-world bridge monitoring applications.

1.2 Objectives

The main objective of this dissertation consists of investigating the use of Information Gain from Tree-Based Models in the context of Structural Health Monitoring, Using the Z-24 Bridge experiment database.

As for specific objectives, there are:

- Provide a concise and effective approach to feature selection.
- Compare Information Gain approach with other established Feature Selection approaches.

- Evaluate how each approach deals with features missing data.
- Evaluate the most informative statistical features by summarizing the amount of times they were selected by the approaches.
- Propose a data-based approach to sensor sensitivity analysis with Information Gain.

1.3 Outline

This work is organized as follows. Besides this section, section 2 presents the methodology, starting with some background knowledge on the decision tree-based algorithms and then on the three approaches for feature selection. It also summarizes the frequency- and time-domain features to be tested in this paper. Then, in section 3, the Z24 Bridge benchmark is presented as a case study, and the quality of the raw data is discussed. In Section 4, results from feature selection (frequency and time domains) and classification are summarized and discussed. The missing data analysis and sensitivity of sensors to damage are also discussed. Finally, section 5 presents the conclusions and future works.

2 Methodology

A graphical representation of the proposed methodology is presented in Figure 4, as a flowchart. First, raw acceleration data is obtained as response signals and segmented through windowing. Next, feature extraction is performed in both time and frequency domains. Feature selection is then implemented using three approaches to define an optimal set of features for damage classification.

Nevertheless, this section begins with background knowledge on tree-based algorithms to support the three approaches to feature selection. It also summarizes the frequency—and time-domain features that will be used in this paper.

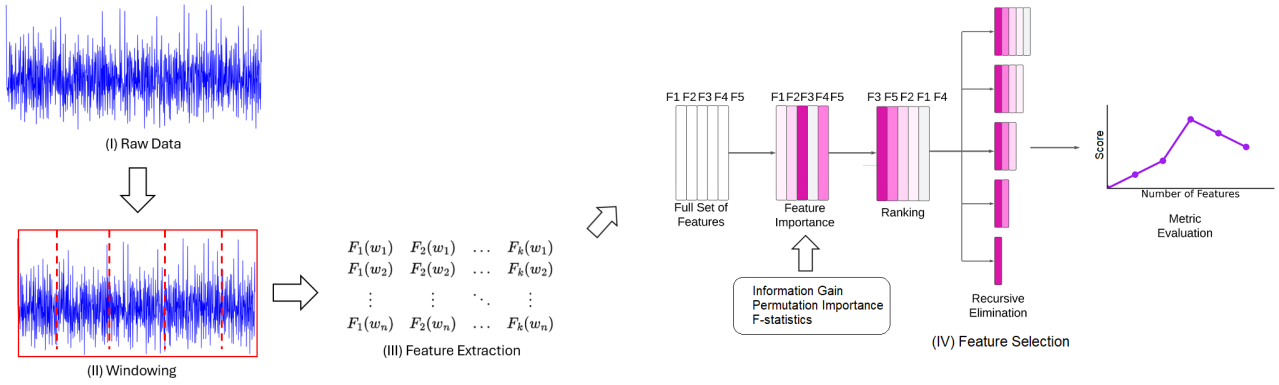


Fig. 4 – Graphical methodology representation, including feature extraction, feature selection, and recursive feature elimination.

2.1 Tree-based algorithms

A decision tree is a supervised nonparametric algorithm for regression and classification tasks. Its objective is to create a model to predict an outcome (terminal or leaf node) based on subsequential logical tests (internal nodes) resembling a tree, which gives rise to the algorithm's name (BREIMAN et al., 2017).

A decision tree is trained rather than explicitly programmed as a machine learning algorithm. It presents many examples relevant to a task. It finds a statistical structure in these examples that eventually allows a model to devise rules for automating the task (CHOLLET, 2021). Therefore, the order and thresholds are not decided by a programmer but by the algorithm itself.

A decision tree grows greedily by selecting the best split in an internal node for all data

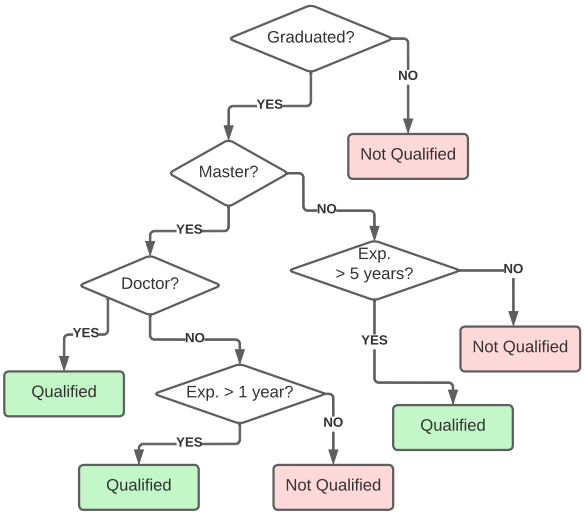


Fig. 5 – Decision tree scheme evaluating a candidate qualification.

set features. The best split is based on the Gini impurity criterion for classification tasks and mean squared error for regression tasks.

Firstly, the algorithm evaluates all the thresholds of all features and selects the one that reduces the loss function the most, splitting the set into two. Then, the procedure continues to the new internal nodes until the loss function is zero, i.e., until it has terminal nodes with only one class. The algorithm is greedy because it makes a locally optimal split, not ensuring a globally optimal decision tree.

Gini impurity is a criterion for classification tasks. It can be defined as a measure that quantifies the variability of a set. It is computed by summing the probability of an item with the label being chosen times the probability of this item not being chosen. It reaches its minimum (zero) when all cases in the node fall into a single target category (ZIEGEL, 2003). Eq. 1 expresses this relation, where m is the number of classes, and p_i is the probability of selecting an item, calculated by its frequency in the set.

$$GI = \sum_{i=1}^{m} p_i(1 - p_i) = \sum_{i=1}^{m} p_i - \sum_{i=1}^{m} p_i^2 = 1 - \sum_{i=1}^{m} p_i^2 \tag{1}$$

To evaluate the Gini impurity of a given split, the algorithm uses a loss function (L), a weighted average of the resulting left and right sets expressed in the Eq. 2, where n, n_L, n_R, GI_L, and GI_R are, respectively, the total number of observations, the number of observations in the left and right sets, and the Gini impurity in the left and right sets.

$$L = \frac{n_L}{n} GI_L + \frac{n_R}{n} GI_R \tag{2}$$

Ideally, a decision tree should have leaf nodes with one class ($L = 0$), resulting in a perfect classification.

Although a decision tree creates complex structures, it might lead to overfitting if not appropriately trained, making inaccurate predictions outside the training data. This is due to the objective of reducing the loss function to zero, which makes the tree overly specific and does not generalize well. To avoid this problem, limits must be put before its training through hyperparameters.

Hyperparameters can define certain aspects of the tree, such as how many splits can occur, the minimum number of observations to perform a split, the maximum number of leaves or terminal nodes, etc. The *scikit-learn* (PEDREGOSA et al., 2011) decision tree has plenty implemented, being the most important:

- **max_depth:** The maximum depth of the tree.
- **min_samples_split:** The minimum number of observations required to split an internal node.
- **min_samples_leaf:** The minimum number of observations required at a leaf node.
- **max_leaf_nodes:** The maximum number of leaf nodes.
- **min_impurity_decrease:** Defines a threshold in which if the gini impurity of the node is smaller or equal to its value, the node won't split.

The hyperparameter tuning improves the algorithm's results, but only to a certain degree. To mitigate its limitations, the decision tree must be implemented with an ensemble, such as a Random Forest, Extremely Randomized Trees, or Gratient Boosting.

2.2 Gradient Boosting

Gradient Boosting is an ensemble algorithm inspired by the gradient descent optimization method, which directs the search for the minimum of the problem's loss function to the largest

possible increment in the attribute vector and multiplies it by a negative learning rate (FRIEDMAN, 2001). Given a training data set $D = \{x_i, y_i\}_1^N$, the goal of gradient boosting is to find an approximation, $F(\mathbf{x})$, which maps observations $\mathbf{x}$ to their output values y, by minimizing the expected value of a given loss function, $L(y, F(\mathbf{x}))$. Gradient boosting builds an additive approximation of $F(\mathbf{x})$ as a weighted sum of functions (or trees). The approximation is constructed iteratively, as illustrated by the Eq. 3, where ρ_m is the weight of the m^{th} function $h_m(\mathbf{x})$.

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \rho_m h_m(\mathbf{x}) \tag{3}$$

First, a constant approximation of $F(\mathbf{x})$ is obtained as shown in Eq. 4.

$$F_0(\mathbf{x}) = \underset{\alpha}{argmin} \sum_{i}^{N} L(y_i, \alpha) \tag{4}$$

The subsequent models are minimized, given the following Eq. 5.

$$(\rho_m, h_m(\mathbf{x})) = \underset{\rho,h}{argmin} \sum_{i}^{N} L(y_i, F_{m-1}(\mathbf{x_i}) + \rho h(\mathbf{x_i})) \tag{5}$$

This algorithm can suffer from overfitting if the iterative process is not adequately regularized (FRIEDMAN, 2001). For some loss functions, if the model h_m fits the pseudo-residuals perfectly, the pseudo-residuals become zero in the next iteration, and the process terminates prematurely. To control the additive process of gradient boosting, a shrinkage hyperparameter is considered to reduce each gradient descent step, resulting in the following Eq. 6.

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \eta\rho_m h_m(\mathbf{x}) \tag{6}$$

Where η is the learning rate. Conventional implementations of Gradient Boosting need to scan all the data instances for every feature to estimate the information gain of all possible split points. Therefore, their computational complexities will be proportional to the number of features and instances. These makes these implementations very time-consuming when handling big data.

XGBoost, or Extreme Gradient Boosting, comes as an improvement over the Gradient

Boosting algorithm, as described by the developers, XGBoost refers to the engineering goal to push the limit of computational resources for boosted tree algorithms (CHEN, 2015). The major changes in the algorithm are introducing a regularization term and expanding the gradient to the second order (Hessian), increasing the control of model complexity and its convergence to an optimal result. In addition to these changes, using sparse matrices and improved data structures was added for better training and result prediction data processing, drastically reducing the computational cost (CHEN; GUESTRIN, 2016).

2.3 Feature Extraction

Feature extraction refers to transforming the measured data into a quantitative form where the correlation with the damage is more readily observed. In vibration-based SHM, the feature extraction process is often based on fitting some model, either physics-based or data-based, to the measured structure response data. The parameters of these models, quantities derived from the parameters, or the predictive errors associated with these models then become the damage-sensitive features (FARRAR; WORDEN, 2013).

In this present work, two categories of signal processing methods related to the domains of extracted features are considered, namely (i) frequency domain and (ii) time domain. Frequency-based signal processing techniques can be categorized into parametric and nonparametric ones. Parametric signal processing techniques are based on extracting modal parameters such as natural frequencies, mode shapes, and damping ratios of the analyzed signals (GUL; CATBAS, 2009). These techniques are commonly applied to signals with known sources. However, nonparametric techniques are employed when there is no prior knowledge about signal sources, such as signal responses measured under moving traffic loads, speed-rated signals, or occupancy signals (BURSI* et al., 2014). The frequency-domain signal processing techniques can also be classified into four types in terms of their damage detection performance based on the changes of modal parameters, including damage detection based on the changes of natural frequencies, mode shapes, damping ratios, and updating the structural model parameter (HAIJEMA; HENDRIX, 2014). Among these techniques, frequency-based methods are more commonly used for structural damage detection. However, frequency-based techniques have disadvantages for damage localization and severity. The time-domain signal processing techniques mainly refer to statistical time series techniques for SHM. The statistical time series techniques are based on the random extraction procedure of vibration signals, developing the statistical model of the

analyzed signals, and the statistical estimation of the analyzed structure's state (undamaged or damaged) (NATKE; CEMPEL, 2012).

As summarized in Figure 4, the feature extraction starts with the raw data referring to the accelerometers. Then, a window size is defined so that the signal is divided into several segments for processing. The size definition varies for each experiment, depending on the sampling rate. High sampling rates lead to greater coverage of signal characteristics; however, it reduces the number of observations for analyzing a structure, which might lead to model overfitting (CASUSOL et al., 2021). It also increases the processing time, which makes the system unsuitable for real-time applications (TAEE et al., 2022).

After defining the size of each window, features will be extracted and transformed into a feature vector, which is stored in a matrix. The lines represent an observation, and the columns represent a related feature. After this process, a machine learning algorithm can classify the structure condition.

2.3.1 Frequendy-domain features

Natural frequencies are the most common features in the frequency domain of bridge SHM. Damage detection methods employing measurement of changes in frequencies are based on the fact that damage reduces the local stiffness of the structure and thereby affects its natural frequencies. While in frequency-based methods, a reduction in natural frequencies indicates damage, the real inverse problem consists of identifying that damage exists and determining its location and size. The frequency-domain features struggle to tackle this issue since the information extracted from it refers to the global structure condition, not directly tied to the location of each sensor. To take advantage of the information acquired by all sensors, the Power spectrum density averaging normalization process is applied, in which the Power spectrum density of each sensor is subtracted by its average and then added, guaranteeing the extraction of more consistent natural frequencies. In case of failure to transform a signal, such as constant or noisy values, the Power spectral density of that signal for that specific window is discarded.

2.3.2 Time-domain Features

Statistical methods are applied directly to signals without any domain transformation. In recent years, the application of these techniques in the context of SHM has been expanded (OSTACHOWICZ et al., 2011). For this present work, nine different time domain features were used, namely: absolute energy (Eq. 7), area (Eq. 8), signal distance (Eq. 9), variance (Eq. 10), mean absolute value (Eq. 11), wavelength (Eq. 12), skewness (Eq. 13), kurtosis (Eq. 14), amplitude (Eq. 15).

$$Energy = \sum_{n=1}^{N-1} \Delta y^2 \tag{7}$$

$$Area = \sum_{i=1}^{n} \frac{1}{2}(y_{1,i} + y_{2,i})\Delta t \tag{8}$$

$$SigDist = \sum_{n=1}^{N-1} \sqrt{1 + \Delta y^2} \tag{9}$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2 \tag{10}$$

$$\text{MAV} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \bar{y}| \tag{11}$$

$$WaveLength = \sum_{n=1}^{N-1} \sqrt{\Delta y^2} \tag{12}$$

$$\gamma_1 = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^3}{\left(\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2\right)^{3/2}} \tag{13}$$

$$\gamma_2 = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2\right)^{2}} - 3 \tag{14}$$

$$Amplitude = \frac{(max(y) - min(y))}{2} \tag{15}$$

Two other time-domain features were considered: mean crossing, which counts the number of times the signal crosses its mean, and the Anderson normality test (D'AGOSTINO, 2017). Note that the mean crossing is an adaptation of the zero crossing feature, which counts the number of times the signal crosses zero. This adaptation reduced drift effects in the signal (KAHRIZI; KABUDIAN, 2023).

2.4 Feature Selection

In the previous feature extraction process, several features may not bring new information regarding the classification of an undamaged condition structure. To achieve this, some techniques were developed to select the most important features in carrying out such tasks. It can be divided into two groups: (i) filter and (ii) wrapper methods. The filter method incorporates an independent measure for evaluating feature subsets without involving a learning algorithm (KUMAR; MINZ, 2014). This method is efficient and fast to compute (computationally efficient). However, filter methods can miss features that are not useful but can be very useful when combined with others. Wrapper methods, on the other hand, integrate the algorithm with feature subset search. Included in this way, the model finds, generates, and evaluates different subsets of resources. The suitability of a subset of features is evaluated through training and testing on the algorithm. So this means the best optimal subset of the features is essentially "wrapped" around the algorithm (GHOJOGH et al., 2019).

This work focuses on wrapped methods, specifically the Recursive Feature Elimination (RFE), as shown in Figure 4. Initially, an importance value is given to each feature, expressed by color saturation, by approaches that will be further discussed, and ranked by their importance. The RFE process then begins: The subset of features is trained and evaluated, the value of the evaluation is stored, The subset is trained and evaluated again, but with the less important feature eliminated, and the process repeats after all but one of the features is eliminated. The registered results can be analyzed to decide which subset of features is the most important. In the example presented, the subset of three features is optimal.

2.4.1 Information Gain

Decision tree-based algorithms have a particular method for evaluating the importance of a feature, i.e., information gain. Given a trained decision tree, the information gain of a feature in one node is the difference between the node's Gini impurity and the Gini impurity weighted

sum of the resulting nodes after the split or child nodes (Eq. 16). Then, for each feature, the information gain is summed in all nodes considering the number of observations in each node (Eq. 17). The results are then normalized so that the sum of the feature's importance is one.

$$IG(f_i, node) = GI(f_i, node) - \sum_{v=1}^{2} \frac{n_v}{n_{node}} GI(f_i, child_v) \tag{16}$$

$$IG(f_i) = \sum_{b}^{nodes} \frac{n_b}{n} IG(f_i, b) \tag{17}$$

2.4.2 Permutation Importance

An alternative approach to determining the importance of the features for the classification model is to calculate the importance of permutation, which is the decrease in a model's score when the value of a single feature is randomly shuffled. This is repeated for each feature in the data set. So that the entire process can be repeated multiple times. The result is an average value of importance for each input feature (and distribution of points considering repetitions). A decline in the model score indicates how strongly the model depends on a specific feature. This procedure breaks the relationship between the functionality and the target variable. Permutation based on feature importance avoids the problem with the model importance as it is calculated on unseen data. The function swap is too sensitive to multicollinearity. If two entities are correlated and one of the functions is replaced, the model will still have access to the functionality via their correlated functionality. This could result in a lower importance ranking for two features, which are very important for predictions.

2.4.3 F-statistics

The ANOVA (analysis of variance) test is a parametric statistical hypothesis test for determining whether the means of multiple samples are from the same distribution. The F-statistic is then used to calculate the ratio of the explained and unexplained variance from the ANOVA test and rank the features based on their level of variance across the damage states. The higher the F-statistic, the more the mean values of the feature differ between the classes (MARKOWSKI; MARKOWSKI, 1990). ANOVA test is generally used as a filter method by removing any feature in which the p-value calculated from the F-statistics is less than a threshold (ex. $p \leq 0.05$). While filter methods for feature selection are computationally efficient, they are not based on

the selection criteria of the features related to the effectiveness of the classification model. But, by ranking the F-statistics of each feature, a wrapper method can be applied.

2.5 Sensor Sensibility Analysis

Sensor sensitivity analysis is gaining attention in the field of SHM. Not only is it necessary to develop systems capable of capturing damage, but also to improve the sensitivity and positioning of sensors. A robust system is costly, so reducing the number of sensors needed to monitor any structure is essential.

Lin (LIN; XU; LAW, 2018) proposed a covariance-based multi-objective multi-type sensor placement method. The method is evaluated in a three-dimensional finite element frame model. Slimani (SLIMANI et al., 2022) suggested a vibration sensor placement method in Carbon Fibre fiber-reinforced polymer composite structures in small structure applications where the measuring instrument weight can affect the vibrational characteristics. Li (LI; YAM, 2001) addresses both sensor sensitivity and damage detection of thin-plate systems by parameter variation. Hou (HOU et al., 2019) used the Sparse Recovery Theory for optimal sensor positioning. The theory requires that the columns of the sensing matrix inform certain independence criteria (minimal mutual coherence). Since sensor location is a combinatorial problem, genetic optimization can be applied to reduce the sensitivity matrix's mutual coherence, thus providing the best sensor system configuration. Loi (LOI; AYMERICH; PORCU, 2022) explores how the position of the sensing transducer, concerning the modal shape of the pump excitation, may influence the sensitivity of the VAM technique for impact damage detection in composite laminates.

The proposed sensor sensitivity analysis exploits the cumulative characteristic of information gain. As previously stated, information gain adds the importance of a feature to all nodes of all trees in an ensemble model. It is proposed that the importance of each feature be added and extracted from each sensor, thus resulting in information gained from that specific sensor. This way, the sensor's importance for detecting damage can be classified as more or less sensible to damage.

3 Case Study: Description and Data Sets

Z24 Bridge was a post-tensioned concrete box girder bridge composed of a main span of 30 m and two side spans of 14 m. The abutment columns were partially hinged at both ends, and the main piers were built as diaphragms, fully connected to the superstructure (PEETERS; ROECK, 2001). A view of the bridge before the demolition is presented in Figure 6.

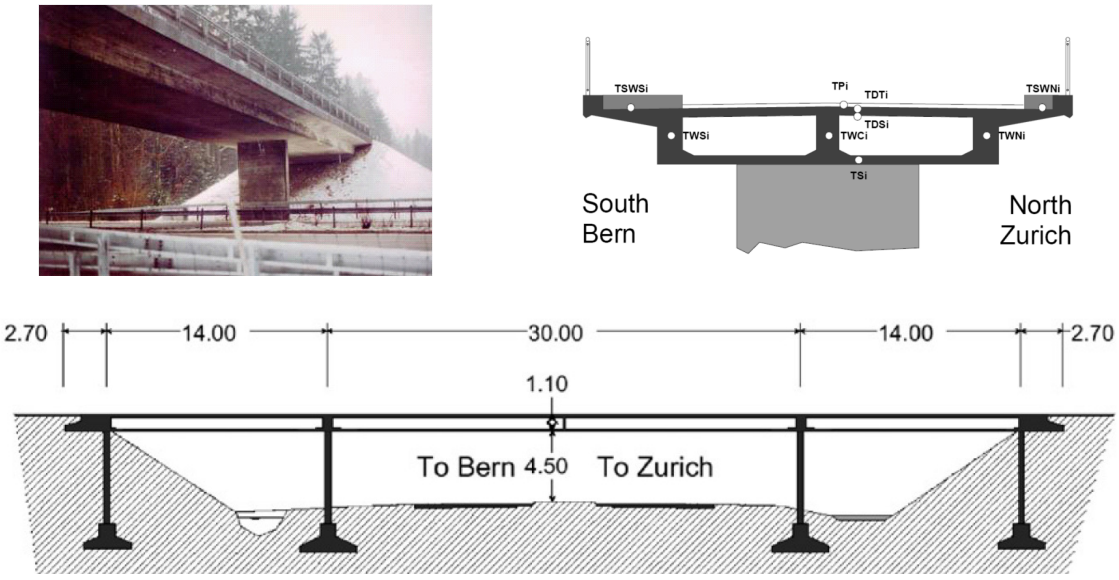


Fig. 6 – Z24 Bridge. Source: https://bwk.kuleuven.be/bwm/z24.

Before demolition, the bridge was instrumented and tested to provide a feasibility benchmark for vibration-based SHM. A long-term monitoring program was carried out from November 11, 1997, until September 10, 1998, to monitor changes in the natural modes of vibration caused by environmental variability and controlled damage. The monitoring program consisted of two phases. From November 11, 1997, to August 5, 1998, the bridge was monitored under environmental variability only, using a network of sensors. Eight accelerometers captured the bridge's vibrations every hour, while other sensors measured environmental parameters, including temperature, at several locations. Originally, the operation had sixteen accelerometers, but only eight lasted the complete program (PEETERS; ROECK, 2001). A total of 4475 observations were obtained this way (for technical reasons, the monitoring system was sometimes down). The sensors marked failed during the operation period, and the valid ones begin with numbers 03, 05, 06, 07, 10, 12, 14, and 16.

From August 5, 1998, a controlled damage campaign was conducted on the bridge. It involved the installation of hydraulic jacks into one of the piers, through which pier settlements and foundation tilts were enforced, followed by simulations of concrete spalling, landslide at

abutment, concrete hinge failure, failure of anchor heads, and rupture of tendons. The measurements were conducted hourly but with a fine network of accelerometers (PEETERS; ROECK, 2001). The test procedures fell within the sensor network paradigms (FARRAR et al., 2006). The various scenarios of the monitoring program are listed in Table 1. A database of 5201 observations was assembled. Observations 1–4475 correspond to the undamaged condition, observations 4476–4605 are considered uncertain, and observations 4606–5201 correspond to the damaged condition according to various damage scenarios. It is noted that observations 4476–4605, corresponding to the installation of hydraulic jacks in one of the piers, are listed as "undamaged" in Peeters (PEETERS; ROECK, 2001). However, the dynamic properties of the bridge were changed during this operation, as the installation of the jacks involved cutting through the respective pier. Therefore, the measurements performed between August 5, 1998, and August 12, 1998, are considered damaged in this study.

Tab. 1 – Description of the monitoring observations

Obs.	**Condition**	**Scenario Desc.**
4474	Undamaged	Env. variability
129	Damaged	Found. settlement of 2 cm
155	Damaged	Found. settlement (4, 8, 9.5 cm)
104	Damaged	Settlement Recovered
45	Damaged	Spalling of concrete
91	Damaged	Landslide
46	Damaged	Failure of concrete hinges
42	Damaged	Failure of anchor heads
106	Damaged	Failure of tendon wires

3.1 Data Wrangling and Feature Extraction

Initially, the acceleration time histories were stored in files with the extension ".AAA", with its name structured by a sequence **WWDHHSS**, where:

WW : A two-digit number indicating how many weeks passed after the experiment began.

D : A letter from A to G indicates the data acquisition's weekday.

HH : A two-digit number indicating the data acquisition time of day.

SS : A two-digit number indicating the sensor.

To facilitate the extraction of features, organization, and data labeling, the raw data was manipulated so that the data collected by sensors simultaneously for the same day composed a

single file in the form of a matrix, with each column representing a sensor. The files were saved with the name **YYYY-MM-DD-HH**, where:

YYYY : A four-digit number indicating the year of the data acquisition.

MM : a two-digit number indicating the month of the data acquisition.

DD : A two-digit number indicating the day of the month of the data acquisition.

HH : A two-digit number indicating the data acquisition time of day.

To reduce data storage and processing time without losing information, the data was saved in 32 bits in the ".npy" file format, a Python programming language-specific file format, improving the optimization even further.

4 Results and Discussion

In this chapter, the results obtained in the selection of the subset of features for classifying structure condition based on the three proposed approaches will be presented, as well as a discussion about their performance, how the missing data features were treated for each of the approaches and finally a evaluation of the proposed method for verifying sensor sensitivity.

5 Results and Discussion

This section will present the results obtained from selecting a subset of features for structural classification based on the three proposed approaches. These results will also be discussed regarding classification performance, the treatment of missing data for each approach, and an evaluation of the proposed approach for verifying sensor sensitivity.

5.1 Feature Extraction

Feature extraction was performed using a sample size of 65,536 acceleration data points per hour in the time domain. Roughly, this resulted in 5,652 observations over a year, with 4,736 from undamaged and 925 from damaged conditions.

The features were extracted in the frequency and time domains. The natural frequencies were extracted from average normalized power spectral densities in the frequency domain. Some signals were not considered, as certain sensors did not work properly. In the time domain, 11 statistical features were considered for each of the eight sensors, so 88 features were obtained. Some samples had invalid data points, resulting in missing data. Figure 7 shows the arrangement of missing values, where each column represents a feature. A dark tone represents the null values, and a light tone represents valid values. Most of the missing values are associated with sensor 10 (features 44 to 55), which shows high constant values in its measured time series.

Observations with missing values must be discarded or handled using imputation methods for some machine learning algorithms. However, a tree-based algorithm can inherently manage missing values, making it ideal for applications like SHM, where some data loss is expected.

After the feature extraction, the ranking process was applied to the three approaches: information gain, permutation importance, and F-statistics. The processing time of the three

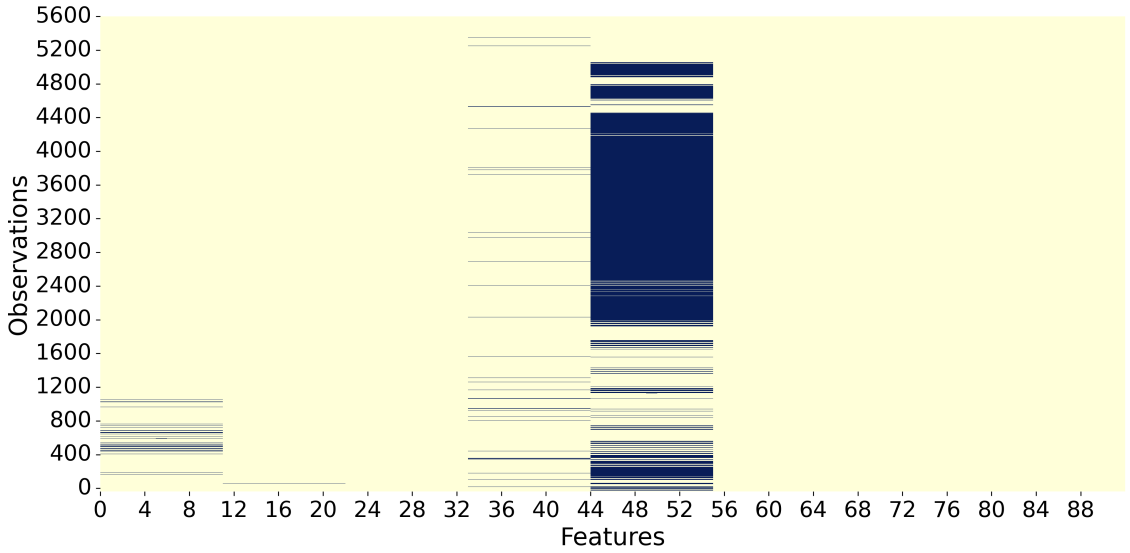


Fig. 7 – Missing values for all 92 features.

approaches for ranking features is summarized in Table 2. The shortest ranking time for the information gain approach is obtained, followed by F-statistics and permutation importance. Information gain is faster, using the model's already trained structure to extract the Gini impurity. Regarding F-statistics, it is worth mentioning that it was necessary to treat missing values for the hypothesis test, requiring longer processing time, using K-nearest neighbors to carry out the imputation method (TROYANSKAYA et al., 2001). On the other hand, F-statistics takes less time than permutation importance because a hypothesis test takes less processing time than training a machine learning algorithm.

Tab. 2 – Processing time comparison for ranking features.

Approach	**Time [s]**
Information gain	$2.242.10^{-3}$
Permutation importance	$1.448.10^{1}$
F-statistics	$1.591.10^{-2}$

5.2 Feature Selection

The XGBoost is used for classification, assuming 65% (3674) of the observations for training, and 30% (1978) for test. The ranking process for each approach was carried out as described in the methodology section. The importance of each feature for the information gain, permutation importance, and F-statistics approaches can be seen in Figures 8, 9 and 10 respectively. The number after the feature's name refers to the sensor where that feature was extracted. The natural frequencies are noted for the three approaches, particularly the second one, ranked

as the most valuable feature. This is most likely due to the damage applied to the bridge, which significantly reduced its stiffness, reflecting changes in its mechanical properties. We can also highlight that the frequencies were obtained considering all sensors' information, ensuring robust features and noise reduction. In Figure 9, it is difficult to visualize the importance of most features, largely because the first feature causes a huge impact on the model's performance when replaced by random features, undermining other features' importance.

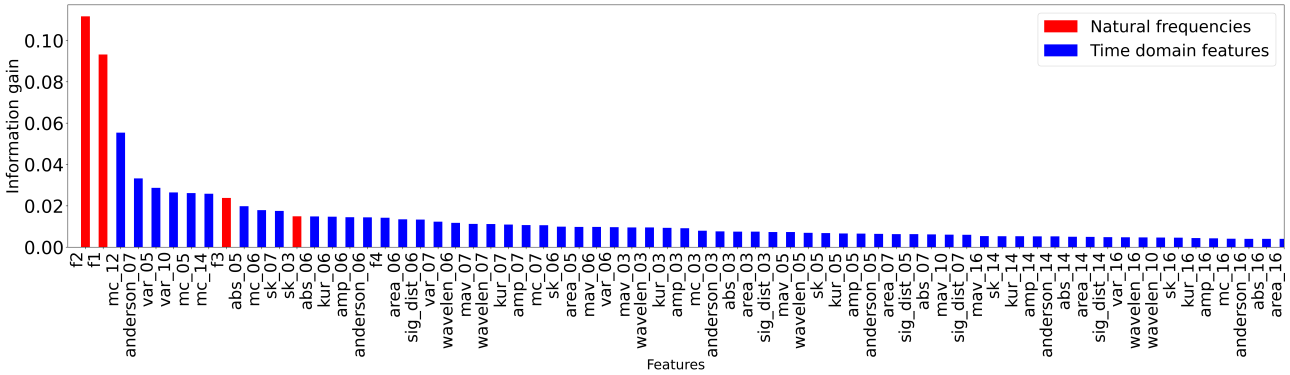


Fig. 8 – Features ranked by their information gain.

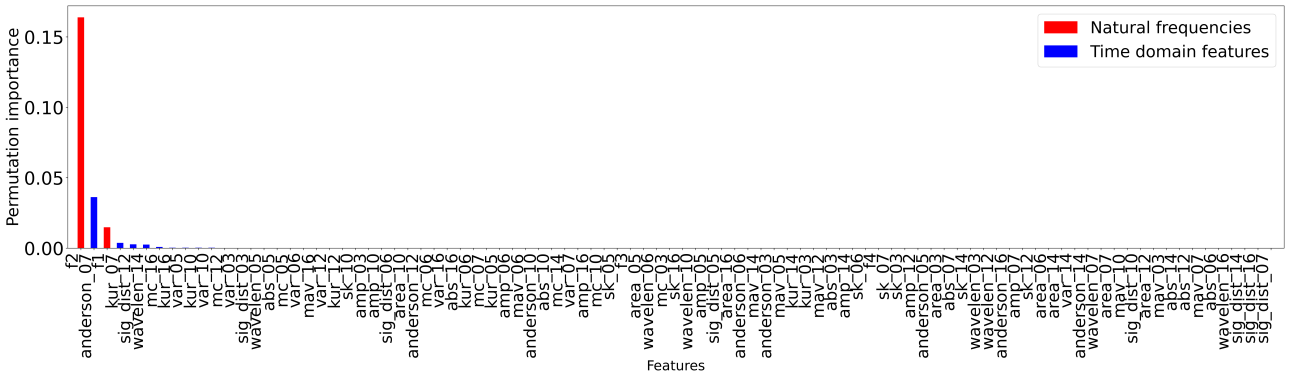


Fig. 9 – Features ranked by their permutation importance.

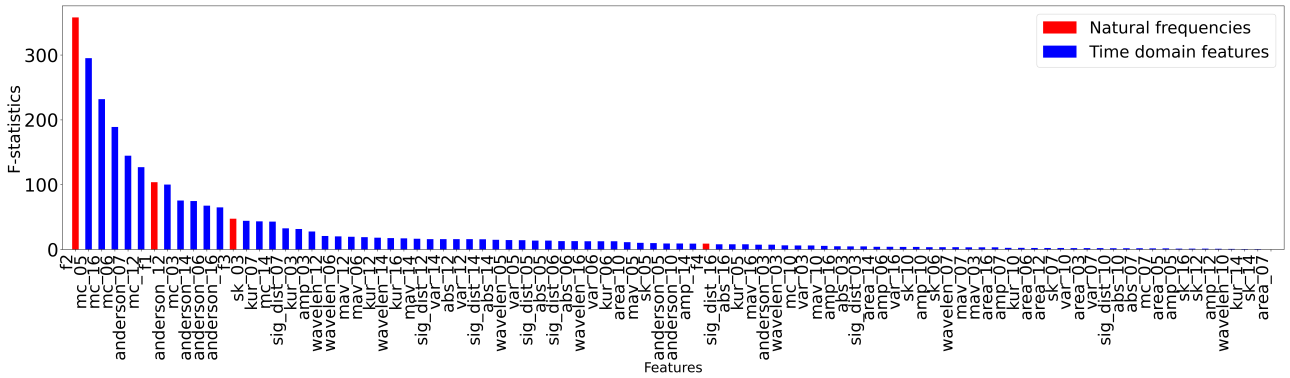


Fig. 10 – Features ranked by their F-statistics.

The RFE process is applied to each approach. A summary of the number of features selected after obtaining an optimal subset and three classification scores (weighted F1 score,

precision, and recall) are presented in Table 3. For a deeper understanding of the importance of several features used for classification, Figure 11 shows the classification performance for each approach in terms of F1 score. The information gain approach has the smallest set of selected features (13), followed by permutation importance (28) and F-statistics (43). However, note that the approach with the best performance regarding all three scores (0.948, 0.947, and 0.948) is F-statistics, with the largest number of selected features. Therefore, a trade-off exists when selecting the feature subset: whether the performance improvement justifies the added complexity of the final model. A larger number of features increases computational time for feature extraction in real-world applications.

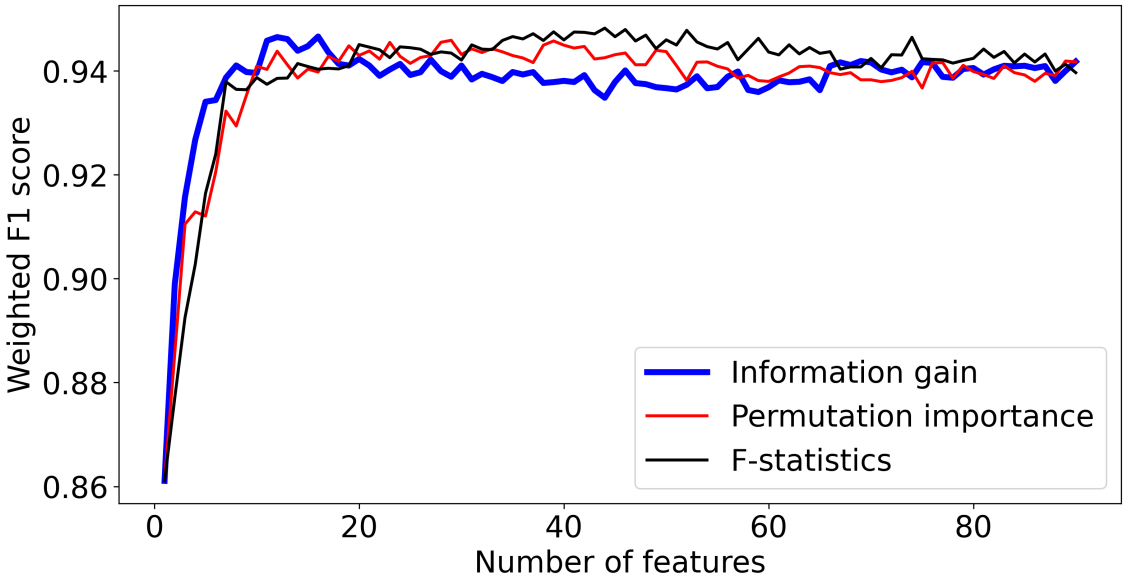


Fig. 11 – Classification performance evolution as a function of several features based on an RFE for each approach.

Tab. 3 – Weighted scores for different feature selection approaches.

Approach	**Features**	**F1 score**	**Precision**	**Recall**
Baseline	92	0.943	0.938	0.941
Information gain	13	0.946	0.944	0.947
Permutation importance	28	0.945	0.946	0.945
F-statistics	43	0.948	0.947	0.948

Table 4 contains the features selected by each approach and Table 5 summarizes the number and type of features each approach selects. As expected, the natural frequencies were selected by all approaches, with the information gain one selecting all four. All three approaches selected mean crossing, variance, Anderson, absolute energy, and skewness. Note that mean crossing as an adaptation of zero crossing is the one with better performance in the time domain. It is also important to point out that absolute energy performs better than signal distance, even though

they have similar formulations. The features area, amplitude, and mean absolute value proved little important to the present case study.

Tab. 4 – Ordered selected Features in each approach.

Information Gain	**Permutation Importance**	**F-Statistics**
2nd Frequency	2nd Frequency	2nd Frequency
1st Frequency	Anderson 7	Mean Crossing 5
Mean Crossing 12	1st Frequency	Mean Crossing 16
Variance 5	Kurtosis 7	Mean Crossing 6
Variance 10	Signal Distance 12	Anderson 7
Anderson 7	Wave Length 14	Mean Crossing 12
Mean Crossing 5	Mean Crossing 16	1st Frequency
Mean Crossing 14	Kurtosis 16	Anderson 12
3rd Frequency	Variance 5	Mean Crossing 3
Mean Crossing 6	Kurtosis 10	Anderson 14
Absolute Energy 5	Variance 10	Anderson 6
Skewness 7	Mean Crossing 12	Anderson 16
4th Frequency	Variance 3	3rd Frequency
	Signal Distance 3	Skewness 3
	Wave Length 5	Kurtosis 7
	Absolute Energy 5	Mean Crossing 14
	Mean Crossing 5	Signal Distance 7
	Variance 6	Kurtosis 3
	Mean Absolute Value 16	Amplitude 3
	Variance 12	Wave Length 12
	Kurtosis 12	Wave Length 6
	Skewness 10	Mean Absolute Value 12
	Amplitude 3	Mean Absolute Value 6
	Signal Distance 6	Kurtosis 12
	Area 10	Wave Length 14
	Anderson 12	Kurtosis 16
	Mean Crossing 6	Mean Absolute Value 14
		Signal Distance 12
		Variance 14
		Absolute Energy 12
		Variance 12
		Signal Distance 14
		Absolute Energy 14
		Wave Length 5
		Variance 5
		Signal Distance 5
		Absolute Energy 5
		Signal Distance 6
		Absolute Energy 6
		Wave Length 16
		Variance 6
		Kurtosis 6
		Area 10

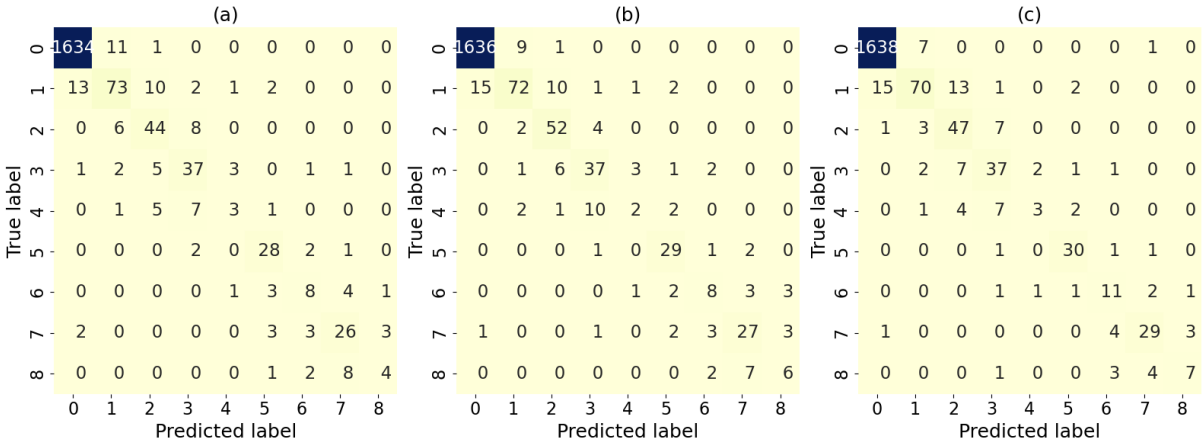


Fig. 12 – Confusion matrices for each feature selection approach: (a) information gain; (b) permutation importance; and (c) F-statistics.

Tab. 5 – Number and type of features selected by each approach.

Feature	Information gain	Permutation importance	F-statistics
Frequencies	4	2	3
Mean crossing	4	4	6
Variance	2	5	4
Anderson	1	2	5
Absolute energy	1	1	4
Skewness	1	1	1
Kurtosis	0	4	5
Signal distance	0	3	4
Wavelength	0	1	5
Mean absolute value	0	1	3
Amplitude	0	1	1
Area	0	1	1

Figure 12 shows the confusion matrices. The undamaged condition is labeled state 0, and labels 1 to 8 are associated with different damaged conditions. The confusion matrices summarize the classification of the 1978 observations in terms of predicted labels with their respective true labels for each approach. The main diagonal represents the true classification. As shown in Figure 12, all approaches have similar confusion matrices. Based on the F1 score (Table 6), overall, the F-statistics performs better. However, as noted before, the information gain has a better trade-off. Note that F-statistics use a larger set of features for classification.

5.3 Missing Data Analysis

When selecting features, it is intuitive to assume that features with missing data are eliminated in favor of complete features since these have more consistency in the structure's con-

Tab. 6 – F1 score for each state condition and approach.

State condition	Information gain	Permutation importance	F-statistics
0	0.99	0.99	0.99
1	0.77	0.75	0.76
2	0.81	0.72	0.73
3	0.71	0.70	0.74
4	0.24	0.17	0.26
5	0.79	0.82	0.87
6	0.48	0.48	0.59
7	0.71	0.68	0.78
8	0.44	0.35	0.54

dition. Figure 13 shows missing features for the selected subsets. It is noted that the fifth most relevant feature of the information gain approach is the one with more missing values. The same occurs for the permutation importance approach, which notices more features with missing values. Therefore, for classification, those features were important, even if they did not have complete information; one of the explanations could be the correlation of missing data with the definition of a class.

An important point to consider is the subset of features from F-statistics. Because an imputation was applied to conduct the variance analysis, the features were contaminated with imputed data, impacting the hypothesis test. Therefore, such features were considered to be of no significance and were therefore taken to the end of the ranking, thus being initially eliminated. This may explain the need for only one feature with a significant amount of missing data at the end of the set, maintaining complete features that may not be relevant to the classification at the beginning of the ranking, occurring in the largest subset of the three approaches.

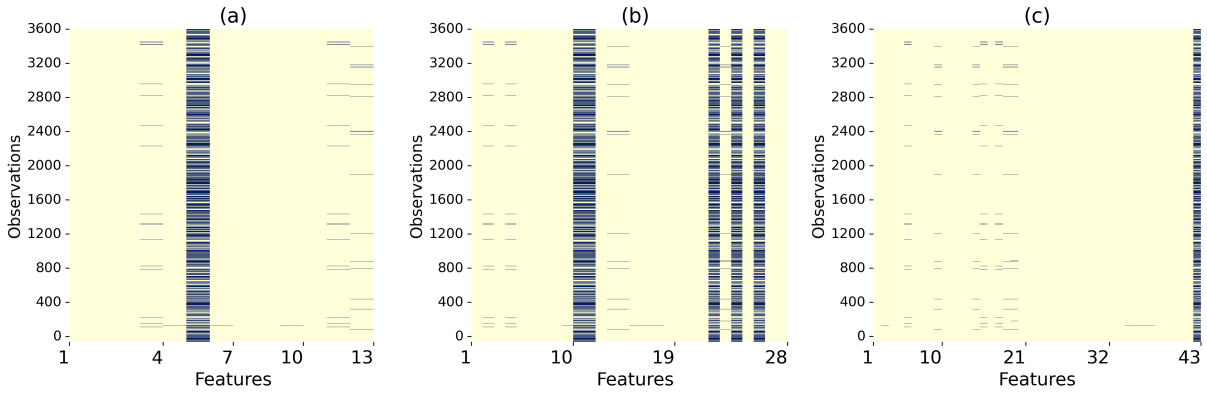


Fig. 13 – Missing features for the subset of features selected for each approach: (a) information gain; (b) permutation importance; and (c) F-statistics.

5.4 Feature Domain Analysis

As discussed previously, the natural frequencies of the Z24 Bridge have played a major role in classifying the undamaged condition. However, the time domain features are usually easier to directly estimate from raw time series, as there is a direct interpretation and no transformation is required (e.g., Fourier transform). In order to show the potential of time-domain features and assuming the subset of features selected from the information gain approach, a classifier is trained for three scenarios: (i) time-domain features only, (ii) frequency-domain features only, and (iii) all features of the subset. Table 7 summarizes the results obtained for those scenarios. As expected, the model using four natural frequencies (frequency domain) provides better coverage than relying solely on nine-time domain features (e.g., F1 score of 0.900 against 0.884, respectively). Nevertheless, it is fair to state that the classifier's performance with time-domain features for practical applications is very close to that of the one based on natural frequencies. Notably, integrating both feature types, comprising a subset of 13 features, results in an improved classification (e.g., F1 score equal to 0.945).

Tab. 7 – Weighted scores for different training features: frequency and time domains.

Training features	**No. features**	**F1 scores**	**Precision**	**Recall**
Frequency frequencies	4	0.900	0.898	0.903
Time domain features	9	0.884	0.880	0.891
All subset features	13	0.945	0.944	0.947

5.5 Sensor Sensitivity Analysis

As mentioned in the methodology section, the proposed sensor sensitivity analysis exploits the accumulated information gain of features per sensor. Therefore, the sum of the information gained from the 11 time-domain features shows the sensitivity of each sensor to damage. As shown in Figure 14, sensor 5 is considered the most damage-sensitive due to the highest information (0.168) shown. As shown in Figure 14b, sensor 5 has the highest weighted F1 score (0.857).

Therefore, when chosen individually, the sensor with the highest accumulated information gain (sensor 5) presents the highest weight F1 score value. However, there is no positive correlation between accumulated information gain and weighted F1 score. For instance, it turns out that sensor 10, which has the second-highest accumulated information gain (0.166),

has the lowest weighted F1 score (0.794). The scatter plot in Figure 15 shows the correlation between accumulated information gain and weighted F1 score to make it more comprehensive.

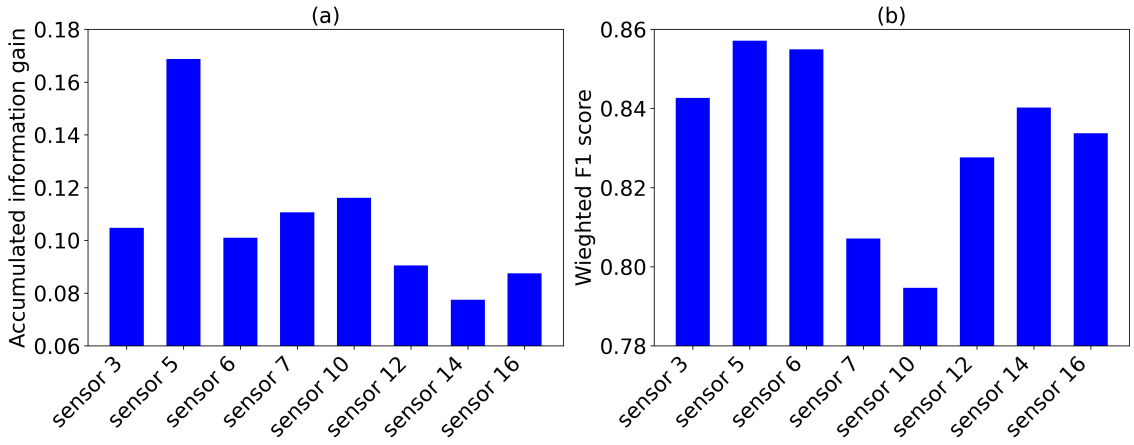


Fig. 14 – Sensor-level sensitivity: (a) accumulated information gain and (b) weighted F1 score.

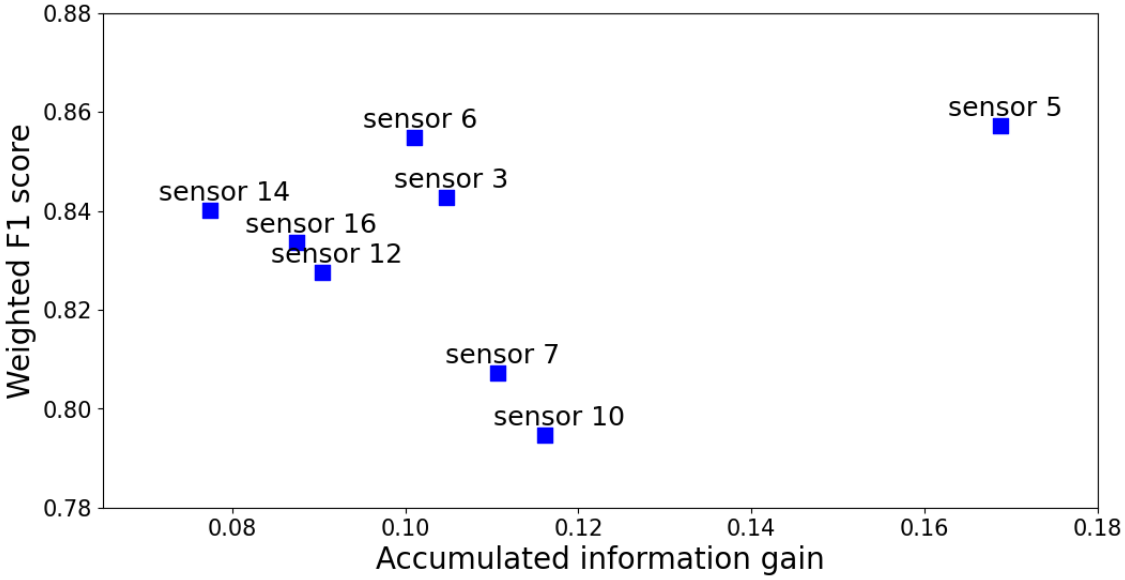


Fig. 15 – Relationship between accumulated information gain and weighted F1 score per sensor.

6 Conclusion

Concise and effective feature selection approaches offered flexibility in ranking damage-sensitive features. The proposed decision tree-based information gain approach was compared with two established approaches: permutation importance and F-statistics. To evaluate the classification performance, an XGBoost classifier was employed for all three approaches. The information gain approach demonstrated the best overall classification performance, achieving a weighted F1 score of 0.946 with a reduced set of just 13 frequency—and time-domain features.

This study confirmed on data sets from the Z24 Bridge that frequency-domain features (e.g., natural frequencies) outperform time-domain features in classification tasks. However, the difference was considered small, which makes time-domain features an alternative for real-world applications. One should note that time-domain features are generally easier to estimate directly from raw time series, as they offer straightforward interpretation without requiring transformations like the Fourier transform. Notably, integrating both frequency- and time-domain features improved classification performance. Among the time-domain features, mean crossing, variance, and the Anderson test consistently ranked as the most important for damage classification across all three approaches.

This study also allowed for an assessment of the robustness of the approaches to missing data. It was observed that features with missing values can still provide relevant information for classification, as two of the approaches selected features with missing values in their subsets. In contrast, based on hypothesis testing rather than performance, the F-statistics approach excluded all features with significant missing values.

Finally, from the sensor-level analysis, the sensor with the highest accumulated information gain also has better classification performance. However, no positive correlation between those variables was found. Therefore, an in-depth study on the applicability of tree-based information gain for sensor sensitivity analysis in SHM is recommended for future work.

7 References

A Time-Domain Signal Processing Algorithm for Data-Driven Drive-by Inspection Methods: An Experimental Study. *Materials (Basel)*, v. 16, n. 7, 2023.

BREIMAN, L. et al. **Classification and regression trees**. [S.l.]: Routledge, 2017.

BUCKLEY, T.; GHOSH, B.; PAKRASHI, V. A feature extraction & selection benchmark for structural health monitoring. **Structural Health Monitoring**, v. 22, n. 3, p. 2082–2127, 2023. Disponível em: <https://doi.org/10.1177/14759217221111141>.

BUD, M. A. et al. Hybrid approach for supervised machine learning algorithms to identify damage in bridges. **Journal of Bridge Engineering**, v. 29, n. 8, 2024.

BURSI*, O. S. et al. Identification, model updating, and validation of a steel twin deck curved cable-stayed footbridge. **Computer-Aided Civil and Infrastructure Engineering**, Wiley Online Library, v. 29, n. 9, p. 703–722, 2014.

CASUSOL, A. J. et al. Optimal window size for the extraction of features for tool wear estimation. In: **2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)**. [S.l.: s.n.], 2021. p. 1–4.

CHEN. **What is the difference between the R gbm (gradient boosting machine) and xgboost (extreme gradient boosting)?** 2015. Online, Data de Acesso: 5 de novembro de 2021. Disponível em: <https://bit.ly/3q9LA4f>.

CHEN, T.; GUESTRIN, C. XGBoost: A scalable tree boosting system. In: **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. New York, NY, USA: ACM, 2016. (KDD '16), p. 785–794. ISBN 978-1-4503-4232-2. Disponível em: <http://doi.acm.org/10.1145/2939672-.2939785>http://doi.acm.org/10.1145/2939672.2939785.

CHOLLET, F. **Deep learning with Python**. [S.l.]: Simon and Schuster, 2021.

D'AGOSTINO, R. B. Tests for the normal distribution. In: **Goodness-of-fit-techniques**. [S.l.]: Routledge, 2017. p. 367–420.

FARRAR, C. et al. **Dynamic Characterization and Damage Detection in the I-40 Bridge over the Rio Grande**. [S.l.]: Los Alamos National Laboratory, LA-12767-MS, 1994.

FARRAR, C. R. et al. Sensor network paradigms for structural health monitoring. **Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures**, Wiley Online Library, v. 13, n. 1, p. 210–225, 2006.

FARRAR, C. R.; WORDEN, K. **Structural Health Monitoring: A Machine Learning Perspective**. [S.l.]: John Wiley & Sons, Ltd, 2013. ISBN 9781119994336.

FIGUEIREDO, E.; BROWNJOHN, J. Three decades of statistical pattern recognition paradigm for shm of bridges. **Structural Health Monitoring**, 2022.

FRIEDMAN, J. H. Greedy function approximation: a gradient boosting machine. **Annals of statistics**, JSTOR, p. 1189–1232, 2001.

GHOJOGH, B. et al. Feature selection and feature extraction in pattern analysis: A literature review. **arXiv preprint arXiv:1905.02845**, 2019.

GUL, M.; CATBAS, F. N. Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications. **Mechanical Systems and Signal Processing**, Elsevier, v. 23, n. 7, p. 2192–2204, 2009.

HAIJEMA, R.; HENDRIX, E. M. Traffic responsive control of intersections with predicted arrival times: a markovian approach. **Computer-Aided Civil and Infrastructure Engineering**, Wiley Online Library, v. 29, n. 2, p. 123–139, 2014.

HOU, R. et al. Genetic algorithm based optimal sensor placement for l1-regularized damage detection. **Structural Control and Health Monitoring**, Wiley Online Library, v. 26, n. 1, p. e2274, 2019.

KAHRIZI, M. R.; KABUDIAN, S. J. Long-term multi-band frequency-domain mean-crossing rate (fdmcr): A novel feature extraction algorithm for speech/music discrimination. **Circuits System Signal Process**, v. 42, p. 6929–6950, 2023. Disponível em: <https://doi.org/10-.1007/s00034-023-02440-0>.

KIM, C.-Y. et al. ffect of vehicle weight on natural frequencies of bridges measured from traffic-induced vibration. **Earthquake Engineering and Engineering Vibration**, v. 2, p. 109–115, 2003.

KUMAR, V.; MINZ, S. Feature selection. **SmartCR**, v. 4, n. 3, p. 211–229, 2014.

LI, Y.; YAM, L. Sensitivity analyses of sensor locations for vibration control and damage detection of thin-plate systems. **Journal of Sound and Vibration**, Elsevier, v. 240, n. 4, p. 623–636, 2001.

LIN, J.-F.; XU, Y.-L.; LAW, S.-S. Structural damage detection-oriented multi-type sensor placement with multi-objective optimization. **Journal of Sound and Vibration**, Elsevier, v. 422, p. 568–589, 2018.

LOI, G.; AYMERICH, F.; PORCU, M. C. Influence of sensor position and low-frequency modal shape on the sensitivity of vibro-acoustic modulation for impact damage detection in composite materials. **Journal of Composites Science**, MDPI, v. 6, n. 7, p. 190, 2022.

MARKOWSKI, C. A.; MARKOWSKI, E. P. Conditions for the effectiveness of a preliminary test of variance. **The American Statistician**, Taylor & Francis, v. 44, n. 4, p. 322–326, 1990.

NATKE, H. G.; CEMPEL, C. **Model-aided diagnosis of mechanical systems: Fundamentals, detection, localization, assessment**. [S.l.]: Springer Science & Business Media, 2012.

NOVAIS, H. C.; SILVA, S. da; FIGUEIREDO, E. Co-kriging strategy for structural health monitoring of bridges. **Structural Health Monitoring**, 2024.

OSTACHOWICZ, W. et al. **Guided waves in structures for SHM: the time-domain spectral element method**. [S.l.]: John Wiley & Sons, 2011.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEETERS, B.; ROECK, G. D. One-year monitoring of the z24-bridge: environmental effects versus damage events. **Earthquake engineering & structural dynamics**, Wiley Online Library, v. 30, n. 2, p. 149–171, 2001.

SLIMANI, M. et al. Experimental sensitivity analysis of sensor placement based on virtual springs and damage quantification in cfrp composite. **Journal of Materials and Engineering Structures «JMES»**, v. 9, n. 2, p. 207–220, 2022.

SOYOZ, S.; FENG, M. Q. Long-term monitoring and identification of bridge structural parameters. **Computer-Aided Civil and Infrastructure Engineering**, v. 24, n. 2, p. 82–92, 2009.

TAEE, A. A. A. et al. The effectiveness of narrowing the window size for ld hd emg channels based on novel deep learning wavelet scattering transform feature extraction approach. In: **2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)**. [S.l.: s.n.], 2022. p. 3698–3701.

TROYANSKAYA, O. et al. Missing value estimation methods for dna microarrays. **Bioinformatics**, Oxford University Press, v. 17, n. 6, p. 520–525, 2001.

VIBRATION feature extraction using signal processing techniques for structural health monitoring: A review. **Mechanical Systems and Signal Processing**, v. 177, p. 109175, 2022. ISSN 0888-3270.

WENZEL, H. **Health Monitoring of Bridges**. [S.l.]: Wiley, 2008. ISBN 9780470740187.

ZIEGEL, E. R. **The elements of statistical learning**. [S.l.]: Taylor & Francis, 2003.