DINAME 2023 - Proceedings of the XIX International Symposium on Dynamic Problems of Mechanics,
A.S. De Paula, A.T. Fabro and M.R. Machado (Editors), ABCM.
Pirenópolis, Go, Brazil, February 26th to March 3rd, 2023

Bolt-loosening Detection Based in Data-Driven of Bolted-Beam Connections by Support Vector Machine Method

Jefferson da Silva Coelho[1], **Amanda Aryda Silva Rodrigues de Sousa**[2], **Marcela Rodrigues Machado**[1] **and Maciej Dutkiewicz**[4]

[1] Department of Mechanical Engineering, University of Brasília. Campus Universitário Darcy Ribeiro, Faculdade de Tecnologia - Asa Norte, Brasília - DF, 70910-900, Brazil

[2] Post-Graduate Program - Integrity of Engineering Materials, University of Brasília. Campus Universitário Darcy Ribeiro, Faculdade de Tecnologia - Asa Norte, Brasília - DF, 70910-900, Brazil

[3] Faculty of Civil, Environmental Engineering and Architecture, University of Science and Technology, Bydgoszcz, Poland

Abstract: Structures are commonly jointed using fasteners such as rivets or bolts arranged in various configurations depending on the required performance. Bolts are widely employed because of the numerous advantages they possess, for instance, avoiding any possible movement and ensuring the stability and security of the bolted joints. However, one of the main disadvantages of fasteners is the loosening that occurs under various causes, such as shock, vibration, and others that can cause serious damage and lead to structural failure. Recent studies have shown that the ML application into bolt joints is still small. This work investigates the application of SVM machine learning to bolt loosening detection based on a data-driven bolt-joint structure. A damage index was calculated using the frequency response to classifying the state of the bolted connection in binary form. It will show some advantages of SVM over other machine learning techniques since it is a technique that allows good results e can be applied to nonlinear classification problems using kernel functions. This work conducts a study of the use of SVM and discusses the challenges of usage, performance and implementation of the technique.

Keywords: Bolt-loosening, SVM Machine Learning, Damage diagnosis, Data-Driven

INTRODUCTION

Structures are commonly jointed using fasteners such as rivets or bolts arranged in various configurations depending on the required performance. Bolts are widely employed because of their numerous advantages, for instance, avoiding any possible movement and ensuring the stability and security of the bolted joints. They secure a preload level, maintaining the bolted system rigidity, making them an important part of the system and influencing its dynamic responses (Ziaja and Nazarko, 2021). Torque loosening is one of the main disadvantages of fasteners and bolts, that occurs under various causes, such as shock, vibration, tightening process inadequate, and fracture, among others that can cause serious damage and lead to structural failure (Wang *et al.*, 2020). Therefore, consistent and systematic maintenance is demanded throughout the structure's lifespan, which is a costly and often dangerous operation requiring an individual to check each bolted joint at regular intervals. Detecting and identifying loosening before failure is still a challenge in several engineering areas (Tran *et al.*, 2020).

Numerous researchers have made significant contributions in exploring methods for detecting bolt loosening (Miao *et al.*, 2020). Nevertheless, a small amount of research has been done on the applications of Machine Learning algorithms (ML) in this area. ML is a technique that uses learning from available data to obtain a model that can make accurate predictions (Avci *et al.*, 2021). Recent studies have been conducted using ML techniques to detect looseness in bolts. Eraliev et al. Eraliev *et al.* (2022) detected and identified bolts loosening using seven ML algorithms comprising a lapped bolt structure. The Random Forest classifier was indicated for a future study of online monitoring. Miguel et al. Miguel *et al.* (2022) used modal parameters to observe the tightening torque loss in bolted joints. They applied two probabilistic machine learning methods, one for damage detection and another for tightening torque quantification. Zakir et al. Zakir Sarothi *et al.* (2022) applied ten ML techniques to identify the failure mode of double-shear bolted connections on a comprehensive database of 455 samples of experimental data. Detailed ML-based failure identification showed that Random Forest best classified the failure modes among the proposed ML methods.

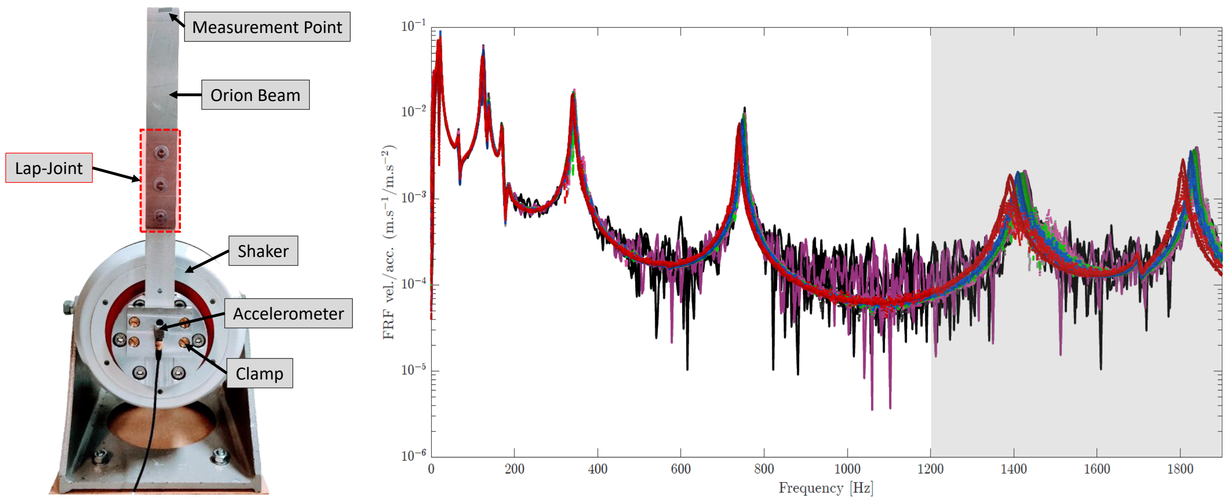
Other works using ML applied to bolted joints were investigated, as presented in Zhou *et al.* (2021) and Cha *et al.* (2016). Furthermore, several different machine learning algorithms can apply. This work proposes using the Support Vector Machine (SVM) for bolt loosening detection in an experimental dataset of a bolted joint. SVM is one of the most widely utilised machine learning techniques for applications such as pattern classification, forecasting and decision-making tasks (Kurian and Liyanapathirana, 2020). The results show the SVM algorithm's satisfactory performance for the bolt-loosening classification in the bolted joint.

DAMAGE DETECTION METHODOLOGY WITH SUPPORT VECTOR MACHINE

Bolts have the function of connecting and maintaining stability between two parts that need to be joined. However, this fixation is not always guaranteed in the long term, a problem that engineering has faced. It is a common structure joined by the bolt to loosen over time due to external vibrations, dynamic loading, or thermal variations. Predicting torque losses is essential and helps engineers develop control strategies for torque tightening. Based on actual torque data, a binary classifier model can be trained to predict whether torque loss occurs for a bolted joint.

Data-Driven

Detecting bolt loosening from vibration data is challenging due to variability and nonlinear effects that occur from the contact interface in bolted joints (Teloli *et al.*, 2022). Therefore, the proposal is to conduct a data-driven strategy to detect loosening bolts from vibration signals. The dataset disposable by Teloli *et al.* (2021) was extracted from an academic scale structure consisting of two assembled beams in a cantilever configuration and connected by three bolts that provide controlled tightening torque conditions on the bolts in a joint lap configuration(Fig. 1). The Orion beam dataset is available at the Mendeley repository (Teloli *et al.*, 2021). To exploit data for diagnostics and prediction, we used the acceleration and velocity spectrum measurements available in Teloli *et al.* (2021) to estimate the transmissibility spectrum.



(a) Experimental setup

(b) Dataset FRF

Figure 1: Data-Driven (Teloli *et al.*, 2021).

Damage Index

The damage index evaluates the structure's condition as damaged or pristine, calculated using the Orion beam's transmissibility considering different torque levels. By decreasing the torque values, changes in the resonance frequency are seen. The torque conditions at 80 cNm and 60 cNm were assumed to be undamaged conditions, and below e.g., 30 cNm, 20 cNm, and 10 cNm were assumed to be damaged conditions due to the loss of connecting properties. The damage index relates the dynamic response of the beam in damaged and health assumption and indicates the torque loss. The damage index used in this work is the frequency response assurance criterion (FRAC) Heylen and Lammens(Heylen and Lammens, 1996). It provides a representative metric in the frequency domain and is expressed such as

$$FRAC_{ij}(f) = \frac{\left| H_{ij}^{damaged}(f)\left(H_{ij}^{undamaged}(f)\right)^{*}\right|^{2}}{H_{ij}^{undamaged}(f)\left(H_{ij}^{undamaged}(f)\right)^{*} H_{ij}^{damaged}(f)\left(H_{ij}^{damaged}(f)\right)^{*}} \tag{1}$$

where "*" defines the complex conjugate operator, "$H_{ij}^{damgaed}(f)$" is the FRF vector on "j" for the damaged excited on "i" and "$H_{ij}^{undamged}(f)$" is the FRF vector for the undamaged, on the same aforementioned coordinates. The values of $FRAC_{ij}$ indicate a scale between zero to unity. In case the $FRAC_{ij}$ is equal to a unit, no damage is found, and the closer to zero is detected an increase in damage severity.

The dataset generates 180 samples damage index. The entire transmissibility signal contains noise and small sensibility to the torque loss. Hence, we calculated the damage index using the truncated signal in a frequency range between 1200 $\sim$ 1940 HZ, as shown in Fig. 1b, comprising the 5th and 6th mode shape. Aside from torque change, the damage indexes were also estimated for different excitation levels and repeat measurements considering the assembly/disassembly of the structure.

Support Vector Machine

Support Vector Machines (SVM) are supervised machine learning techniques developed from the Statistical Learning Theory that can be used for the classification and regression of structured data. In the case of linear classification, with two classes, let $\{(x_i, y_i), ..., (x_n, y_n)\}$, a training dataset with n observations, where x_i represents the set of input vectors and $y_i(+1, -1)$ is the class label of x_i, the hyperplane is a straight line that separates the two classes with a marginal distance (as seen in Fig. 2). The purpose of an SVM is to construct a hyperplane using a margin, defined as the distance between the hyperplane and the nearest points that lie along the marginal line termed as support vectors.

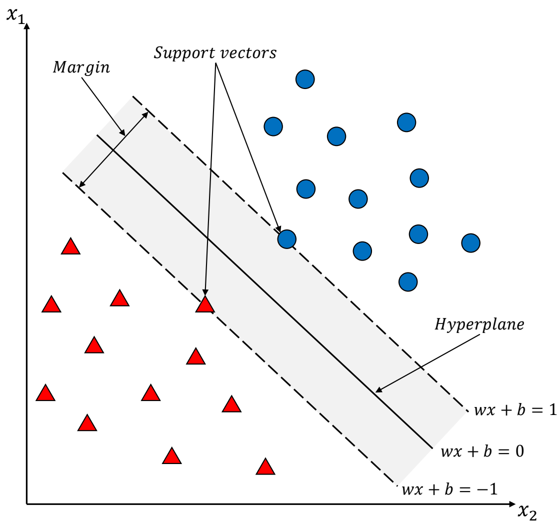


Figure 2: Illustration of SVM.

We can define the hyperplane by Eq. (2), where we have the dot product between x and w added to the term b:

$$D(x) = w^T.x + b = c \quad for \quad -1 < c < 1 \tag{2}$$

where x represents the points within the hyperplane, w are the weights that determine the orientation of the hyperplane, and b is the bias or displacement of the hyperplane. When $c = 0$, the separating hyperplane is in the middle of the two hyperplanes with $c = 1$ and -1. The goal of an SVM is to maximise the data separation margin from the minimisation of $||w||$. This optimization problem can be obtained as the quadratic programming problem given by Eq. (3).

$$min \quad \frac{||w||^2}{2} \quad s.t \quad y_i(w^T.x_i + b) \geq 1 \quad for \quad i = 1, 2, ..., n \tag{3}$$

where $||w||$ is the Euclidean norm.

In real situations, applications are usually not linearly separable, this is due to the presence of noise in the data or the problem may be non-linear. The adoption of soft margins relaxes the restrictions imposed on the optimisation problem with the introduction of slack variables, allowing some data to have one or few misclassified instances with fewer margin violations. In SVM to control error minimization and classification margin maximization is using the C parameter, called the penalty or regularization. The value of C controls the margin, the number of support vectors, and the training and testing errors, having a great influence on the classification performance of the SVM. Small values of C maximize the margin that can lead to underfitting and a large value of C minimizes the margin that can lead to overfitting (Tharwat, 2019).

SVM can be applied to nonlinear classification problems using kernel functions, where it projects the sample space over a higher dimensional space where the data are linearly separable. The most used kernels are Polynomial (Eq.4) and Gaussian radial basis function (RBF) (Eq.5) (Tharwat, 2019).

$$K(x, y) = ((x^T y) + 1)^d \tag{4}$$

where x and y are vectors in input space, and d is the degree of the polynomial kernel.

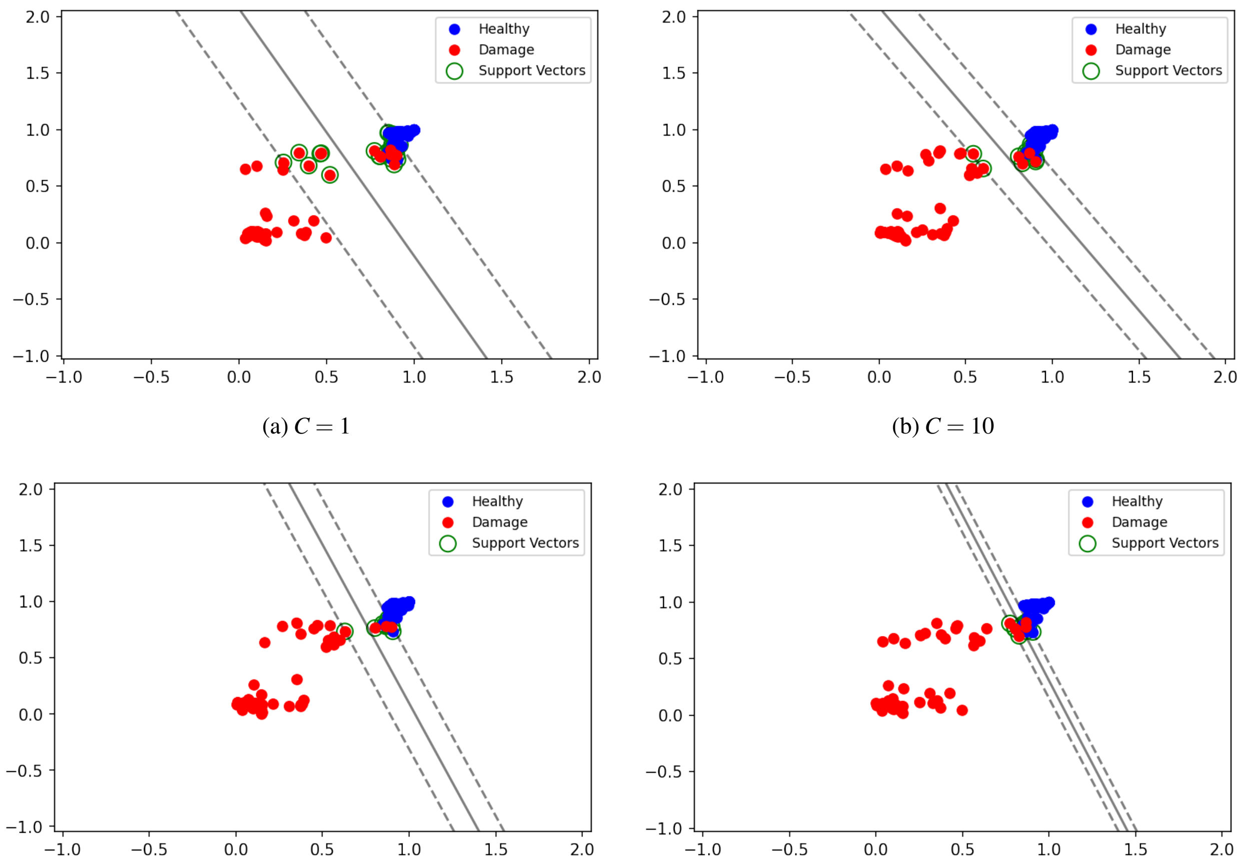
$$K(x,y) = e^{-\gamma \|x-y\|^2} \tag{5}$$

where γ is a positive parameter for controlling the influence of new features on the decision boundary, and $\|x-y\|^2$ is the euclidean distance between x and y.

RESULTS AND DISCUSSION

The data were divided into training and testing sets before training the SVM classifier. Out of the total data samples, 75% of the data was used for training purposes and 25% for testing the model. By providing a training dataset, the algorithm builds a model, which assigns the new test datasets to one of the two categories (Health or Damage). The new test datasets are then mapped to the same space and predicted in their category based on which side of the space they are placed. SVM takes a set of data as the input and predicts the two possible classes of each input.

The data-driven used in this study conting noise and randomness that do not follow a pattern (outliers), adding a non-linearly separable, so the data will be in the middle of the road or on the wrong classification side, causing margin violations. In this way, the most flexible linear SVM model is used to avoid violations by controlling a parameter C. Figure 3 shows the decision limits and margins of the smooth margin SVM classifiers with different C parameters. The idea is to find the best parameter to maintain the balance between the widest path and the least possible violation. For a low value of $C = 1$ (Fig. 3a), the margin was maximized, increasing the number of support vectors and misclassified samples, with a higher value of $C = 100$ (Fig. 3d) the margin is minimized, causing noise or outliers to determine the decision limit, making the classifier sensitive to noise in the data. The parameter $C = 10$ (Fig. 3b) was the best result, as it produces fewer prediction errors and most violations are on the right side of the decision limit. The parameter $C = 50$ (Fig. 3c) could be chosen. However, with a smaller margin, although few samples are on the way, the samples on the right side are closer to the decision limit.



(a) $C = 1$ (b) $C = 10$

(c) $C = 50$ (d) $C = 100$

Figure 3: The impact of increasing C on the performance of linear SVM. Decision boundaries (solid line), two planes (dashed lines), support vectors samples are marked by surrounding it by round shapes (green).

As a result of SVM implementation to classify the bolted connection state in binary form, trained Linear SVM can classify the test datasets for damaged and undamaged with s correct classification accuracy of 97.77%, which is a good result for damage classification. The accuracy used only evaluates the model's overall performance, and the proportion of undamaged damage was correctly classified. The confusion matrix was used to analyze the data classification, providing an overview of configurations identifying damages and their correct classification. The confusion matrix also indicates errors and successes of the model in comparing actual and predicted values. Figure 4 shows the SVM confusion matrices (number and percentage of predictions) in which they allow the correct classification of the state of the bolted joint for the loss of torque to identify damage. The model failed to identify the damage, with an erroneous prediction of 2.22% classified as undamaged (false positive).

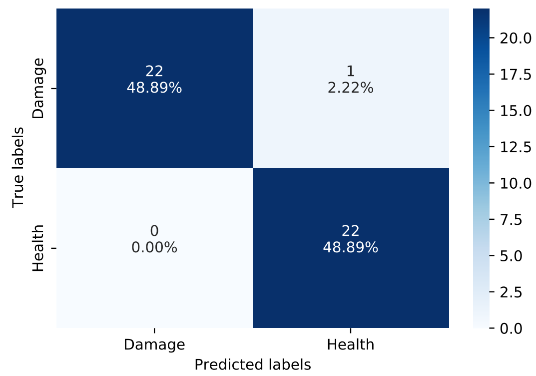
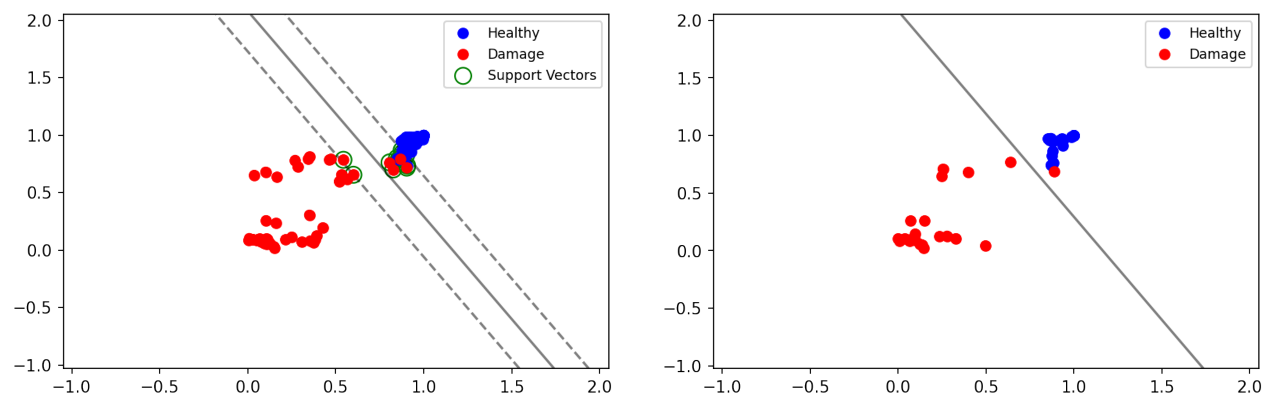


Figure 4: Confusion Matrix SVM Linear.

Figure 5a shows the SVM results calculated considering the training dataset. It can be seen that the calculated SVM algorithm can separate the training datasets into undamaged (blue) and damaged (red), which have a certain violation margin even though the instances are on the correct side of the decision limit. In this case, the algorithm did not lead to prediction errors. Thereby the structure's integrity prediction involves checking the side of the decision boundary from the test point onwards.

Figure 5 shows the SVM result considering the training and testing data. It can be seen from the figure (5a) that the SVM algorithm can separate the training datasets into undamaged (blue) and damaged (red), and there are margin violations on the right side (blue). The samples are on the correct side of the decision limit, and there are outliers on the right side of the margin. After training, predicting the structure's status only involves figuring out which side of the decision boundary will be classified. In this way, the test dataset was used. Fig. 5b shows the result with a good classification of the states of the bolted joint, and it is still possible to see that there are outliers, but that despite the problem, the model did not lead to large prediction errors, as seen previously in the accuracy with 97.77%.

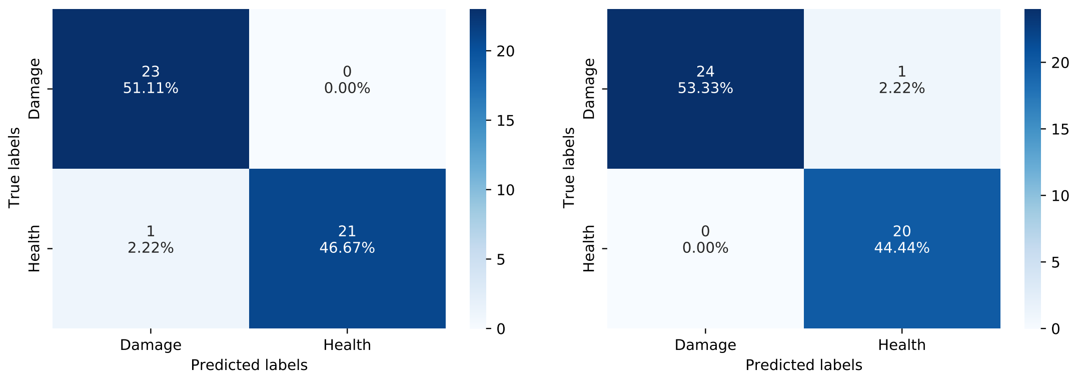


(a) Training (b) Test

Figure 5: Dataset SVM Linear.

The Linear SVM classifier was efficient and worked well in this case, even in the presence of noise and outliers in the dataset. However, there are cases where it is impossible to split the training data by a hyperplane. For this purpose, we use the resource of taking the training data to a higher dimensional space using Kernel functions. The choice of Kernel and the parameters considered affects the performance of the classifier obtained, as they define the induced decision frontier. At this stage of the work, we consider that the data are not linearly separable. Therefore, we observed the behaviour of two SVM Kernels functions, polynomial and RBF, and assessed whether there would be any improvement in SVM performance. Based on the performance results of C, the value of the parameter $C = 10$ was kept ($C = 10$), and the parameter of the polynomial kernel function was of $d = 3$ and the RBF $\gamma = 1$.

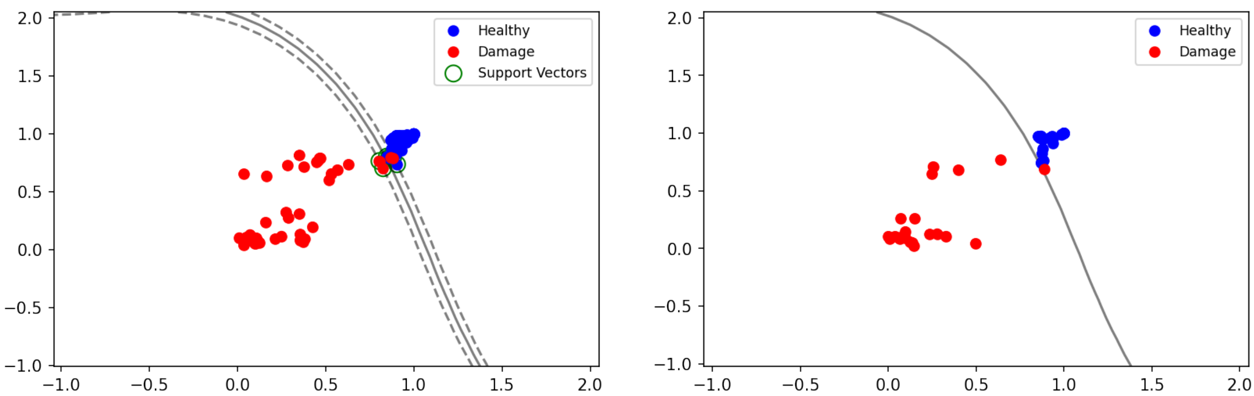
As a result of implementing SVM through the Kernel function, the test dataset regarding Health and Damage got the correct classification accuracy percentage of 97.77% in both Kernel's polynomial and RBF. This result is similar to that found with Linear SVM. Fig 6 shows the confusion matrix where it was possible to observe where each model failed in its prediction. For the nonlinear SVM polynomial kernel, the model failed with an erroneous prediction of 2.22% classified as damaged (Fig 6a), and for the RBF kernel, it also failed with a wrong prediction of 2.22% classified as undamaged (false positive), as shown in Fig 6b.



(a) Polynomial (b) RBF

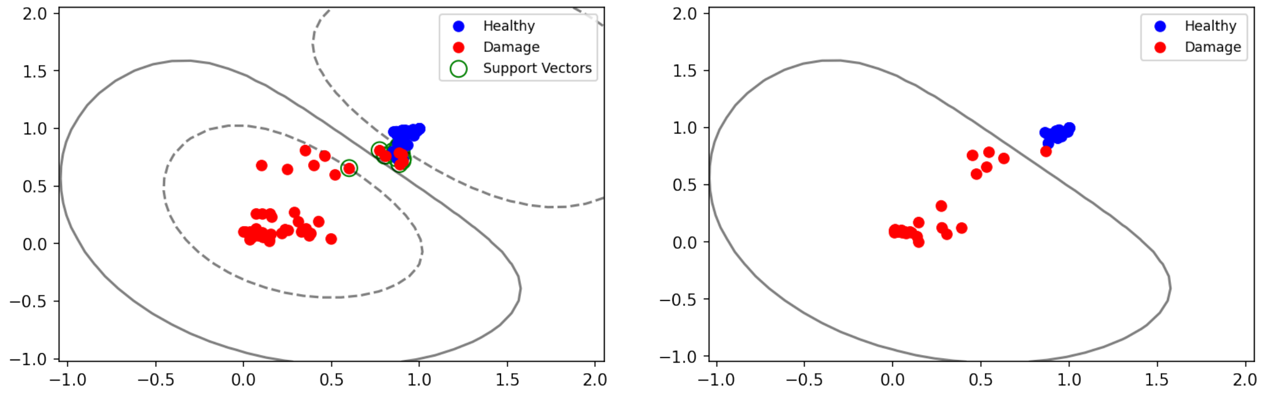
Figure 6: Confusion Matrix SVM Nonlinear Kernel.

Figures 7-8 show the nonlinear SVM results of the kernel functions considering the training and test dataset. Figure 7a-7b shows that the decision limit of the polynomial Kernel was determined by the outlier, affecting its orientation and leading to overfitting. The RBF kernel function (Fig.8a-8b) shows that the decision limits classify the points correctly, covering the data dispersion well, there are still margin violations and outliers on the right side of the decision limit on both training and test data but do not affect model performance.



(a) Training (b) Test

Figure 7: SVM Nonlinear Kernel Polynomial.



(a) Training

(b) Test

Figure 8: SVM Nonlinear Kernel RBF.

CONCLUSIONS

SVM is one of the most popular supervised machine learning algorithms and has been used in many applications to perform regression, linear and nonlinear classification using a kernel trick. As we saw in the results, the SVM classification model's performance depends on the choice of SVM parameters (penalties and kernel) that, if poorly selected, can lead the model to underfit or overfit.

This work investigated the application of SVM machine learning tools to detect bolt loosening in a bolted joint structure. A damage index was calculated using the frequency response to classify the state of the bolted connection in binary form (Health and Damage), and the behaviour of the SVM using kernel functions were also analyzed. The SVM analyzes include illustrative examples to show the test error rate through the confusion matrix and decision limits.

It was possible to show in the results the ability of the Linear SVM classifier to detect the loosening of bolts satisfactorily using the measured vibration data. The results also showed that using a simple linear SVM with experimental data from the Orion beam obtained good accuracy in the damage classification with 97% of accuracy. This result was similar to the SVM models Nonlinear kernel polynomial and RBF.

ACKNOWLEDGMENTS

This optional section must be placed before the list of references.

REFERENCES

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M. and Inman, D.J., 2021. "A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications". *Mechanical Systems and Signal Processing*, Vol. 147, p. 107077. ISSN 08883270. doi:10.1016/j.ymssp.2020.107077. URL `https://doi.org/10.1016/j.ymssp.2020.107077` `https://linkinghub.elsevier.com/retrieve/pii/S0888327020304635`.

Cha, Y.J., You, K. and Choi, W., 2016. "Vision-based detection of loosened bolts using the Hough transform and support vector machines". *Automation in Construction*, Vol. 71, No. Part 2, pp. 181–188. ISSN 09265805. doi: 10.1016/j.autcon.2016.06.008. URL `http://dx.doi.org/10.1016/j.autcon.2016.06.008`.

Eraliev, O., Lee, K.H. and Lee, C.H., 2022. "Vibration-Based Loosening Detection of a Multi-Bolt Structure Using Machine Learning Algorithms". *Sensors*, Vol. 22, No. 3, p. 1210. ISSN 1424-8220. doi:10.3390/s22031210. URL `https://www.mdpi.com/1424-8220/22/3/1210`.

Heylen, W. and Lammens, S., ???? "FRAC: a consistent way of comparing frequency response functions".

Kurian, B. and Liyanapathirana, R., 2020. "Machine Learning Techniques for Structural Health Monitoring". *Lecture Notes in Mechanical Engineering*, pp. 3–24. ISSN 21954364. doi:10.1007/978-981-13-8331-11.

Miao, R., Shen, R., Zhang, S. and Xue, S., 2020. "A review of bolt tightening force measurement and loosening detection". *Sensors (Switzerland)*, Vol. 20, No. 11. ISSN 14248220. doi:10.3390/s20113165.

Miguel, L.P., Teloli, R.d.O., da Silva, S. and Chevallier, G., 2022. “Probabilistic machine learning for detection of tightening torque in bolted joints”. *Structural Health Monitoring*, Vol. 0, No. 0, pp. 1–16. ISSN 17413168. doi: 10.1177/14759217211054150.

Teloli, R.d.O., Butaud, P., Chevallier, G. and da Silva, S., 2022. “Good practices for designing and experimental testing of dynamically excited jointed structures: The Orion beam”. *Mechanical Systems and Signal Processing*, Vol. 163, No. June 2021, p. 108172. ISSN 08883270. doi:10.1016/j.ymssp.2021.108172. URL `https://doi.org/10.1016/j.ymssp.2021.108172`.

Teloli, R.d.O., Butaud, P., Chevallier, G. and da Silva, S., 2021. “Dataset of experimental measurements for the Orion beam structure”. *Data in Brief*, Vol. 39, p. 107627. ISSN 23523409. doi:10.1016/j.dib.2021.107627. URL `https://doi.org/10.1016/j.dib.2021.107627`.

Tharwat, A., 2019. “Parameter investigation of support vector machine classifier with kernel functions”. *Knowledge and Information Systems*, Vol. 61, No. 3, pp. 1269–1302. ISSN 02193116. doi:10.1007/s10115-019-01335-4. URL `https://doi.org/10.1007/s10115-019-01335-4`.

Tran, D.Q., Kim, J.W., Tola, K.D., Kim, W. and Park, S., 2020. “Artificial intelligence-based bolt loosening diagnosis using deep learning algorithms for laser ultrasonic wave propagation data”. *Sensors (Switzerland)*, Vol. 20, No. 18, pp. 1–25. ISSN 14248220. doi:10.3390/s20185329.

Wang, F., Chen, Z. and Song, G., 2020. “Monitoring of multi-bolt connection looseness using entropy-based active sensing and genetic algorithm-based least square support vector machine”. *Mechanical Systems and Signal Processing*, Vol. 136, p. 106507. ISSN 10961216. doi:10.1016/j.ymssp.2019.106507. URL `https://doi.org/10.1016/j.ymssp.2019.106507`.

Zakir Sarothi, S., Sakil Ahmed, K., Imtiaz Khan, N., Ahmed, A. and Nehdi, M.L., 2022. “Machine learning-based failure mode identification of double shear bolted connections in structural steel”. *Engineering Failure Analysis*, Vol. 139, p. 106471. ISSN 13506307. doi:10.1016/j.engfailanal.2022.106471. URL `https://doi.org/10.1016/j.engfailanal.2022.106471`.

Zhou, L., Chen, S.X., Ni, Y.Q. and Choy, A.W.H., 2021. “EMI-GCN: A hybrid model for real-time monitoring of multiple bolt looseness using electromechanical impedance and graph convolutional networks”. *Smart Materials and Structures*, Vol. 30, No. 3. ISSN 1361665X. doi:10.1088/1361-665X/abe292.

Ziaja, D. and Nazarko, P., 2021. “SHM system for anomaly detection of bolted joints in engineering structures”. *Structures*, Vol. 33, No. May, pp. 3877–3884. ISSN 23520124. doi: 10.1016/j.istruc.2021.06.086. URL `https://doi.org/10.1016/j.istruc.2021.06.086` `https://linkinghub.elsevier.com/retrieve/pii/S2352012421005919`.

RESPONSIBILITY NOTICE

The author(s) is (are) the only responsible for the printed material included in this paper.