

Framework for Single Misfire Identification in a Marine Diesel Engine using Machine Learning

Victor Nicodemos Guerra¹, Brenno Moura Castro¹, Dionísio Henrique Carvalho de Sá Só Martins¹, Ricardo Homero Ramírez Gutiérrez² and Ulisses Admar Barbosa Vicente Monteiro¹

Received: 12 December 2023 / Accepted: 14 September 2024
© Harbin Engineering University and Springer-Verlag GmbH Germany, part of Springer Nature 2025

Abstract

Misfire is a common fault in compression ignition engines, characterized by the absence or flame loss due to insufficient fuel in the cylinders. This fault is difficult to diagnose and resolve due to its multiple potential causes. This study focuses on identifying misfires in a 12-cylinder V-type marine diesel engine by analyzing vibration data collected from 15 accelerometers mounted on the engine block. Three machine learning algorithms—*K*-Nearest Neighbors (*K*-NNs), support vector machines (SVMs), and random forests (RFs)—were employed to classify engine conditions using 18 time-domain features. Results showed that the *K*-NN, SVM and RF algorithms achieved F1 scores of 99.87%, 100%, and 99.87%, respectively, when using 18 time-domain features and all 15 accelerometers mounted on the engine block. Additionally, the study evaluated classification performance while reducing the number of accelerometers and features using two methods: Relief-F and general combinatory analysis (GCA). Although the GCA method yields better results when using only two accelerometers and nine features for misfire classification, its overall process required substantially more computational time compared to Relief-F. The best result obtained with Relief-F was achieved using 3 accelerometers and 18 features. Therefore, Relief-F proved to be more practical and take less overall computational time within the proposed framework.

Keywords Misfire fault; Vibration; Marine diesel engine; *K*-NN; SVM; Random forest

1 Introduction

Misfire is a common failure in compression ignition (CI) engines, occurring when insufficient fuel in the engine's cylinders causes ignition failure or flame extinction. This issue can arise from a variety of causes, including a nota-

bly poor fuel/air mixture within the combustion chamber, clogged injector nozzles, poor fuel quality, deteriorated sealing rings, intake/exhaust valve sealing failure, incorrect injection pressure, or low operating temperatures. Owing to its numerous potential causes, misfire is one of the most difficult problems to diagnose and resolve (Merhige, 2016).

Concerning its severity, Tao et al. (2019) noted that misfire faults in single and double cylinders are the most commonly observed forms. However, the occurrence of faults in more than two cylinders can lead to severe vibrations in the engine block.

In common rail engines, when a cylinder fails to ignite, the consequences may be propagated to other cylinders because the fuel distribution from the common rail may become deregulated. Therefore, one cylinder might receive an excessive amount of fuel.

For partially premixed combustion, Cui et al. (2022b) stated that misfire occurs when the equivalence ratio is low or when the thermodynamic conditions inside the cylinder cannot sustain flame development. Furthermore, they asserted that misfire in PCC results from the synergistic effects of local equivalent ratio and temperature.

Maurya et al. (2019) stated that misfires can lead to an increase in long-term gas emissions, which is a notable

Article Highlights

- Misfire in compression ignition engines, caused by insufficient fuel and resulting in flame loss, is a common but challenging fault to diagnose due to its many potential causes.
- Vibration signals were collected via 15 accelerometers mounted on the engine block and 18 time-domain features were extracted and three machine learning algorithms were used for classification.
- SVM achieved a perfect F1 score (100%), while *K*-NN and RF achieved 99.87%, after applying feature and sensor selection methods.

✉ Victor Nicodemos Guerra
victorguerra@oceanica.ufrj.br

¹ Federal University of Rio de Janeiro (UFRJ), Ocean Engineering Program (PENO), Centro de Tecnologia, Bloco I-108, Cidade Universitária, Ilha do Fundão ZIP 20945-970, Rio de Janeiro/RJ, Brazil

² State University of Amazonas, Naval Engineering Department, Escola Superior de Tecnologia, Parque Dez de Novembro, 69050-020, Manaus/AM, Brazil

concern today. According to Wang et al. (2023), the increase in soot emissions with rising intake pressure is due to a fuel/air equivalence ratio below 0.175 (indicating a lean mixture). Meanwhile, Zhang et al. (2023) stated that NO_x emissions increase with higher intake temperatures, a consequence of an excessively poor fuel/air mixture.

In addition, misfires can cause power loss, leading to reduced ship performance and sudden decreases in engine speed, as well as fluctuations in speed and torque (Maurya et al., 2019) and unbalanced forces. Given these possible consequences, misfire detection is a critical aspect of onboard diagnostics, and the diagnostic system must be capable of identifying the specific cylinder where the failure has occurred.

This study focuses on cases of complete single-cylinder misfires, limiting the possible causes to those involving a total absence of fuel in the cylinder chamber, such as fully clogged injector nozzles. Simulations were conducted by cutting off fuel injection to one cylinder at a time.

Regarding misfire detection methods, Zheng et al. (2019) proposed tracking crankshaft angular speed in a four-cylinder spark-ignition engine using an optimized Luenberger sliding mode observer. Their results demonstrated that this method could effectively diagnose misfires under transient engine conditions.

Charles et al. (2009) used encoder signals to construct instantaneous angular speed waveforms representing torsional vibrations, aiming to predict misfires in the cylinders of medium-speed engines. However, this method requires devices capable of detecting minimal changes in angular speed during misfire and must account for random variations in acceleration during operation (Devasenapati et al., 2010).

Tamura et al. (2011) proposed using exhaust gas temperature analysis at a low sampling rate for misfire detection in internal combustion engines. This approach enabled an understanding of signal behavior, supporting decision-making for preventive maintenance. However, the study was limited to engines operating under stationary conditions and did not account for drastic load changes. Additionally, Zhang et al. (2021) noted that this method is only applicable to a specific range of ignition-induced misfires.

Lujan et al. (2010) proposed using the derivative of the in-cylinder pressure signal for combustion detection in diesel engines because any drastic pressure changes within the cylinder chamber would be amplified by the derivative of the corresponding signals. The results were compared with classic thermodynamic analysis and validated through experiments on diesel engines. However, the authors highlighted the need for further studies considering transient operating conditions and cyclic variations.

Firmino et al. (2021) used vibration and acoustic analysis for misfire detection in a four-stroke spark-ignition engine. After comparing the two methods, they concluded

that vibration analysis offers greater accuracy than acoustic analysis.

Jafarian et al. (2018) used four sensors to monitor vibration signals in an automobile engine to investigate different levels of misfire severity and excessive valve clearances. The proposed method achieved a fault diagnosis accuracy of over 94%, demonstrating its effectiveness in accurately detecting misfires and its wide applicability.

This technique was adopted in the current study due to the proven effectiveness of vibration signal monitoring, along with its ease of installation and nonintrusive nature.

Machine learning (ML) techniques are increasingly being applied to a wide range of predictive tasks. For example, Cui et al. (2022a) developed a method based on a backpropagation neural network (BPNN) model aimed at predicting the parameters of multicomponent fuel surrogates using datasets derived from simpler surrogates. The results showed that BPNNs with two hidden layers outperformed those with a single hidden layer, enabling more accurate parameter estimation for real fuels.

In the literature, several techniques have been proposed for diagnosing failures in diesel engines. Among them, ML algorithms hold considerable importance because they can learn the specific characteristics of different fault types, thereby reducing diagnostic errors.

Devasenapati et al. (2010) employed decision trees using the C4.5 algorithm for misfire detection in a four-stroke, four-cylinder petrol engine. Their approach involved time-domain vibration signal analysis, extraction of eight features, and feature selection based on gain ratio. The method achieved a classification accuracy of 95%. However, decision trees are known to exhibit high variance in predictions (Fratello and Tagliaferri 2018), a limitation that can be mitigated through ensemble methods such as the random forest (RF). In a related study, Sharma et al. (2014) applied five types of tree-based algorithms, including RF, for misfire detection in a CI engine using time-domain vibration signal analysis. They extracted eight features and performed feature selection using information ratio and entropy reduction. Their approach achieved 100% classification accuracy in distinguishing normal operation and misfire conditions.

Zhang et al. (2018) proposed a support vector machine (SVM)-based system for misfire diagnosis in a 16-cylinder V-type marine diesel engine. Features were extracted in the frequency domain using a polar representation method, where quantitative radii were used to indicate the position and severity of the fault. The system achieved high accuracy in identifying normal operating conditions and various levels of misfire.

Jafarian et al. (2018) explored the use of artificial neural networks (ANNs), SVMs, and k -nearest neighbors (K -NNs) for detecting misfires and valve clearance faults in internal combustion engines based on multisensor vibration signal

monitoring. Fast Fourier transform was applied to convert time-domain signals into the frequency domain. All three classification methods achieved accuracy levels above 95%.

Vakharia et al. (2017) decomposed vibration signals using wavelet transform before feature extraction, which was performed in the time and frequency domains. They employed information gain and Relief-F for feature selection and used SVM and RF to identify bearing faults. Their approach achieved a cross-validation efficiency of 98.38% when using Relief-F in combination with RF.

This study aims to identify single-cylinder misfires in a 12-cylinder V-type marine diesel engine using vibration data collected from 15 accelerometers mounted on the engine block. Three ML techniques, K -NN, SVM, and RF, were used for misfire classification. In addition, this study introduces a new contribution by addressing the potential for reducing the number of accelerometers used in data acquisition, an aspect not considered in previous studies. To examine fault diagnosis performance under reduced sensor configurations, the Relief-F method was applied for data dimensionality reduction. This approach investigates whether fewer sensors can be used for signal acquisition without compromising diagnostic accuracy. The performance of the Relief-F method is then compared with that of a general combinatory analysis (GCA) method to assess its effectiveness in identifying the optimal number of sensors and its overall feasibility for sensor reduction in fault diagnosis applications.

This work is organized into six sections. Section 2 provides an overview of the ML methods used in this study. Section 3 details the methodology implemented to prepare the data for ML classification. Section 4 describes the data collection process and feature extraction procedures. Section 5 presents the results, along with a discussion of the findings. Finally, Section 6 offers concluding remarks and suggestions for future work.

2 Theoretical background

ML techniques play a crucial role in fault detection. However, some techniques are more suitable for specific scenarios than others, regardless of their complexity. Therefore, understanding the theoretical background of the methods used is essential to assess their potential applicability to the current study.

2.1 Classification algorithms

The methods used for misfire detection and identification are described in this section.

2.1.1 K -Nearest Neighbors (K -NNs)

K -NN is an instance-based learning method developed to allow the analysis of characteristics in cases where estab-

lishing parametric approximations of probability densities is difficult (Taunk et al., 2019). This method is one of the simplest and most effective approaches for pattern classification (Tamura et al., 2011). K -NN can be applied as a classifier and regression method.

K -NN uses a distance metric to identify the K samples nearest to a given unlabeled sample (Lei et al., 2020). The majority label among these K samples is then assigned to the unlabeled sample as its predicted result (Lei et al., 2020; Abu Alfeilat et al., 2019), as depicted in Figure 1. The distance and similarity between the test and training samples play an important role in the K -NN algorithm, because its classification performance heavily depends on the two mentioned parameters (Abu Alfeilat et al., 2019).

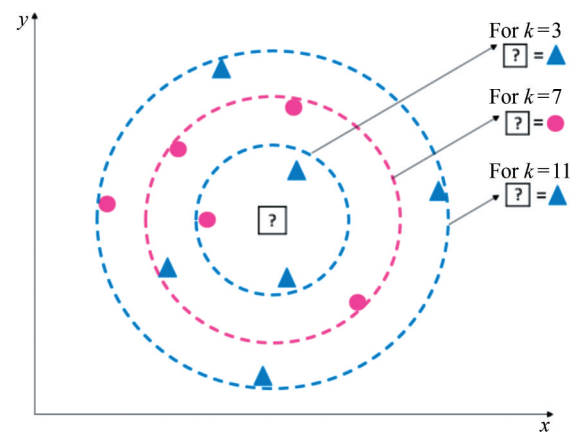


Figure 1 k -nearest neighbors (K -NNs) scheme for output data classification (modified from Lei et al., 2020)

The choice of K -neighbors determines the nature of the approximation used to classify the unlabeled sample. For smaller neighborhoods, the prediction is localized, and K -NN is prone to overfitting. However, for larger K , the algorithm tends to generalize more, potentially overlooking smaller agglomeration patterns (Kramer, 2013).

The metric used in this study is the Euclidean distance, which is defined in Eq. (1) as follows:

$$ED(a, z) = \sqrt{\sum_{i=1}^n |a_i - z_i|^2} \quad (1)$$

where a and z are the coordinates of two different points in the i -th dimension of Euclidean space, and n is a positive number representing the total magnitude of Euclidean space.

2.1.2 SVM

SVM is a supervised learning model used for classification and regression tasks. The main concept behind SVMs lies in the use of boundaries that separate two or more classes. For example, in two dimensions, a straight line is used to divide the classes, while in three-dimensional space,

a plane divides the zones. This approach can be mathematically extended to higher dimensions (Noble, 2006). These boundaries are called hyperplanes. The best hyperplane is the one that separates the classes with the maximal distance, known as the Maximal Margin Classifier or Hard Margin SVM (Mammone et al., 2009), and it is defined by Eq. (2) below (Boser et al., 1992):

$$D(x) = \mathbf{w} \cdot \mathbf{x}_i + b, i = 1, \dots, m \tag{2}$$

where \mathbf{w} is the normal vector to the hyperplane, \mathbf{x} is the input vector of the sample i , b is the bias, and m is the total number of samples. Changing b results in a parallel shift of the hyperplane.

In this algorithm, classes are separated into zones whose boundaries are defined by the nearest samples to the Maximal Margin Classifier. These points are called support vectors, and they determine the boundary hyperplanes. The distance between these hyperplanes is given by Eq. (3):

$$D_m = \frac{2}{\|\mathbf{w}\|} \tag{3}$$

To maximize D_m , i.e., the distance between these boundaries, $\|\mathbf{w}\|$ should be minimized. The SVM classification scheme for two classes is depicted in Figure 2.

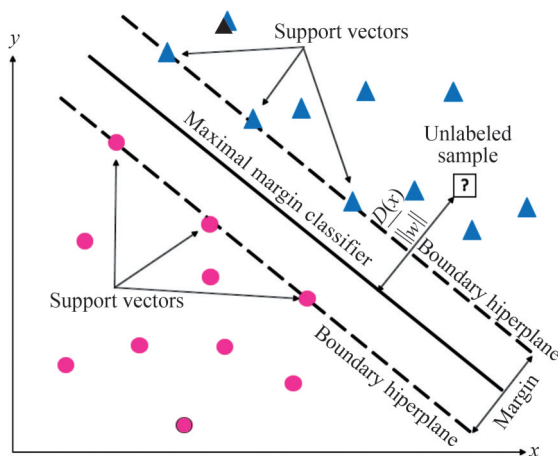


Figure 2 Support vector machine (SVM) classification scheme considering two classes (based on Cortes & Vapnik, 1995)

However, in practice, experimental data often contains noise, which prevents a perfect linear separation between classes. Therefore, the maximal margin hypothesis cannot be applied directly (Mammone et al., 2009). Aiming to address this issue, the algorithm developed by Cortes and Vapnik (1995) introduced the concept of relaxing the boundary conditions in Eq. (2) when necessary, thereby allowing some misclassification of training samples (Mammone et al., 2009). Therefore, the boundary hyperplanes can be defined by Eq. (4):

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \zeta_i, \zeta_i \geq 0 \tag{4}$$

where ζ_i is a positive slack variable that measures how far a sample lies from the boundary of its original class.

The optimal separating hyperplane is obtained by minimizing Eq. (5):

$$\min \left[\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \zeta_i \right] \tag{5}$$

where C is the regularization term that controls the balance between maximizing the margin and minimizing the training error (Ukil, 2007).

For data that cannot be linearly separated, the “Kernel trick” is adopted (Noble, 2006). This technique involves mapping the data into a higher-dimensional feature space where it becomes linearly separable (Figure 3). Common kernel functions include linear, polynomial and Gaussian kernels.

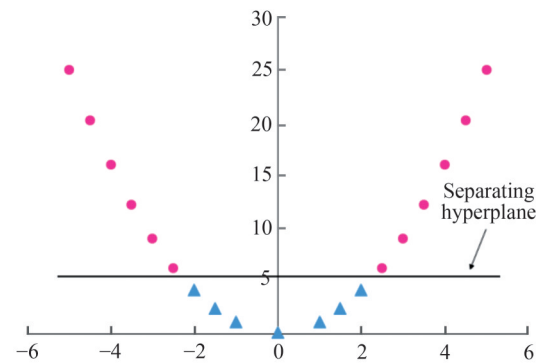


Figure 3 Kernel trick representation in two dimensions (modified from Noble, 2006)

2.1.3 RF

RF is a tree-based ensemble learning algorithm used for regression and classification tasks (Cutler et al., 2012; Breiman, 2001). Each tree casts a vote to classify a given feature vector (Figure 4). The bagging method (Breiman, 1996) then creates a committee of trees by generating subsets of observations through sampling with replacement from the training set. This method enables the creation of independent models by choosing different attributes to split nodes in each tree. The ensemble approach of RF reduces the bias that typically arises from using a single decision tree.

2.1.4 Comparison of ML methods

Table 1 presents a comparison of the methods used in this study, highlighting their respective advantages and disadvantages.

Although K -NN is a time-consuming method, it was chosen for an initial approach due to its user-friendly implementation and minimal hyperparameter tuning requirements. SVM was selected for its capability to handle nonlinear systems, its effectiveness in high-dimensional spaces, and

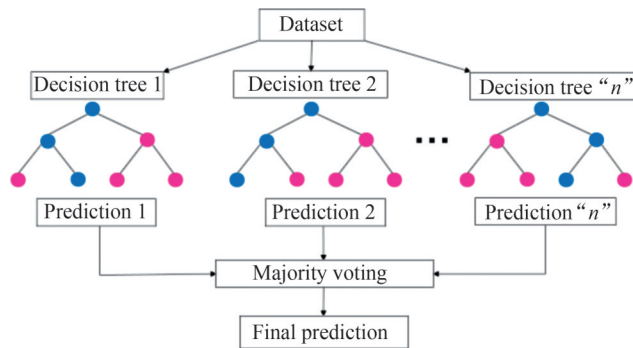


Figure 4 General structure of a random forest (RF) algorithm (Verikas et al., 2011)

its low overfitting tendency. RF was employed for its capability to manage nonlinear parameters while also reducing overfitting and variance. All three classifiers share the advantage of being robust to noise.

2.2 Feature selection method

Dimensionality reduction is applied due to several reasons, such as the decreasing amount of information to be stored for improved computational efficiency (Jović et al., 2015), eliminating irrelevant and redundant features (Zhou et al., 2021), and increasing the speed of algorithm training and prediction (Phyu and Oo, 2016).

The filter-based method Relief-F is an extension of the original Relief algorithm (Kononenko, 1994). In the original version, a sample is randomly chosen from the dataset, and its nearest neighbors from the same and different classes

are identified. The algorithm then compares the attribute values of the sampled instance with those of its nearest neighbors, assigning importance weights. Attributes with higher scores are considered more effective in distinguishing between instances from different classes.

The advantage of Relief-F over the original version lies in its capability to handle multiclass problems, greater robustness, and its capacity to manage noisy and incomplete data (Robnik-Šikonja and Kononenko, 2003). Additionally, this technique has a low bias, supports interactions among features, and can detect local dependencies that other methods may fail to identify (Bolón-Canedo et al., 2013).

3 Proposed methodology

This section presents the proposed methodology designed to address the three main objectives of this study:

- Detect and identify misfires in a 12V4000 marine diesel engine;
- Evaluate the possibility of reducing the number of accelerometers while maintaining strong prediction performance;
- Assess the effectiveness of dimensionality reduction techniques in preserving high prediction accuracy.

The methodology roadmap is illustrated in Figure 5. First, an engine test rig was used to collect vibration signals. The acquired data was then preprocessed and organized into different sets. Afterward, the dataset was divided into training and testing datasets. Time-domain features were extracted, and dimensionality reduction was evaluated using feature selection techniques. Outliers were

Table 1 Comparison of the ML methods used

Method	Advantages	Disadvantages
K-NN	Robust to noise (Boateng et al., 2020); Easy to implement (Boateng et al., 2020); Few hyperparameters to optimize (Boateng et al., 2020);	Might function slowly in the presence of an excessive number of samples in the dataset (Boateng et al., 2020); Absence of generalization of the training dataset (Taunk et al., 2019).
SVM	Effective in high-dimensional spaces (Boisberranger et al., 2023); Effective when the number of features is greater than the number of samples (Boisberranger et al., 2023); Versatile when different kernel functions are adopted (Boisberranger et al., 2023); Capable of solving nonlinear and unknown systems (Ukil, 2007); Robust to noise and no overfitting (Ukil, 2007).	Once SVM is a binary classification, it might not be optimum for multiclass classifications. For these cases, pair-wise classifications may be used (Ukil, 2007); In cases where the number of features is notably greater than the number of samples, choosing kernel functions and regularization terms is crucial (Boisberranger et al., 2023).
RF	Not sensible to attribute value scaling (Nazarenko et al., 2019); Discrete and continuous signals are effectively processed (Nazarenko et al., 2019); Methods for tree construction may be used in datasets with missing value attributes (Nazarenko et al., 2019); Methods for feature individual relevance assessment inside a model (Nazarenko et al., 2019); Data-processing capability with a high quantity of features and classes (Nazarenko et al., 2019); Robust to noise (Nazarenko et al., 2019).	Substantial size of resulting models (Nazarenko et al., 2019); Classification time increases for larger input datasets (Boateng et al., 2020).

subsequently removed, and the data was normalized through feature scaling (except in the case of RF, as described in Table 1).

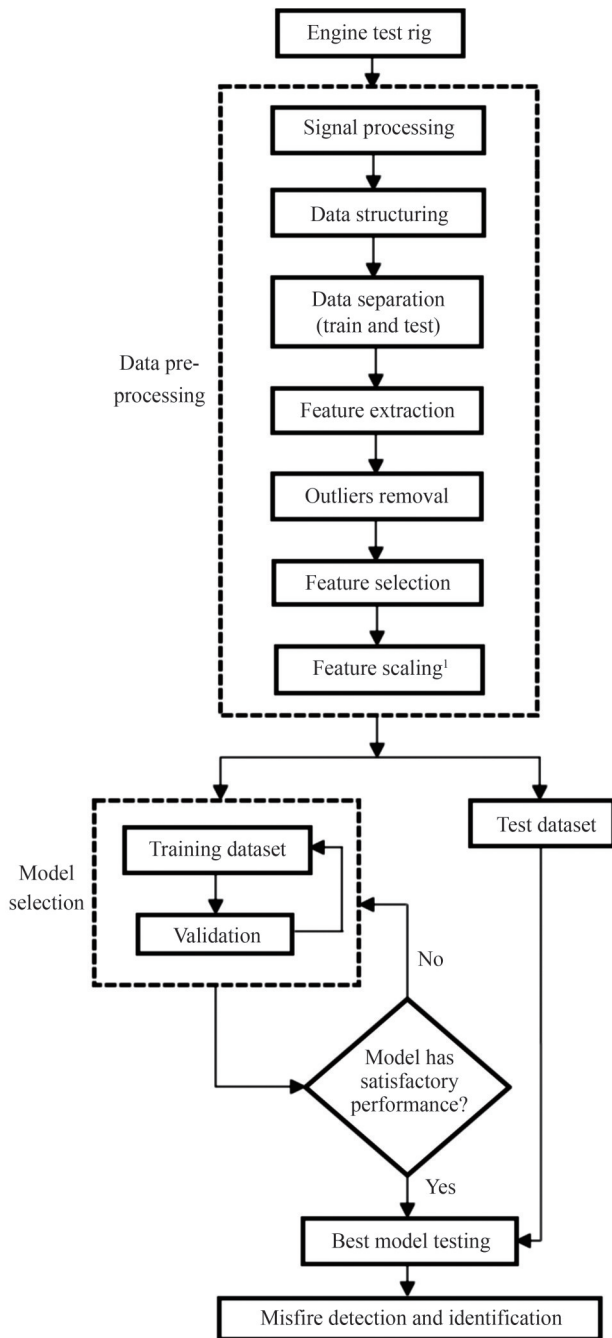


Figure 5 Roadmap of the proposed methodology for misfire detection and identification

After the preprocessing step, the data was prepared for processing by ML algorithms. A portion of the training dataset was separated into k -folds to evaluate the performance of the ML algorithms. Once a model demonstrated satisfactory performance during cross-validation, it was applied to the test dataset for misfire detection and identification.

4 Engine test rig

The experimental procedure was conducted on a 12V4000 diesel marine engine (MTU Detroit Diesel, 2016), with its specifications presented in Table 2.

Table 2 Specifications of the marine diesel engine

Manufacturer/Series	MTU/12V4000
Type	T1237k11
Number of Cylinders/Stroke	12 V-type cylinders/4-stroke
Fuel	Diesel
Rated power (kW)	1 193
Rated speed (RPM)	1 900
Max. Torque (N·m)	7 595
Engine stroke length (mm)	190
Engine bore diameter (mm)	165

The following boundary conditions were considered for the experimental procedure:

- Single-cylinder misfires;
- Misfires induced by cutting off fuel injection, resulting in no combustion within the cylinder chamber. This condition reduces the number of possible misfire causes (i.e., completely clogged injectors).

In the engine test rig, 15 piezoelectric accelerometers were used to record vibration data on the engine block (Figure 6). One accelerometer was mounted on each of the 12-cylinder V-type heads, while the remaining three were positioned on the engine foundation, as shown in the schematic representation in Figure 7.

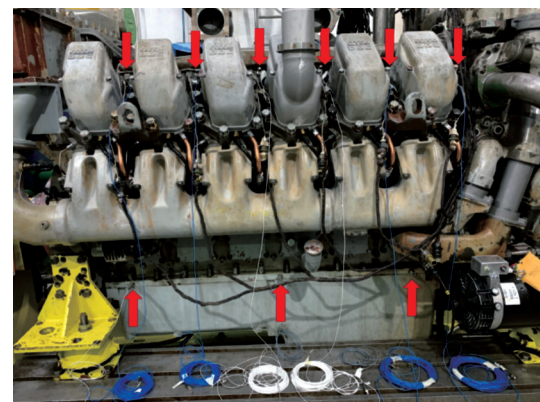


Figure 6 Installation of accelerometers in an MTU 12V4000 engine block

The signals were recorded under 13 different operational conditions: one representing normal operation and one for an induced single-cylinder misfire in each of the 12 cylinders. This distinction is crucial for accurate misfire detec-

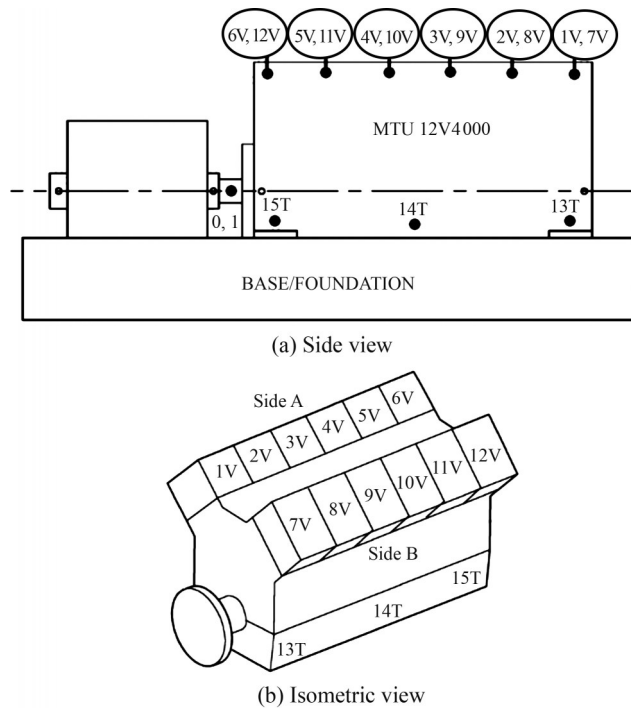


Figure 7 12V4000 IC Engine with the indicative positions of the 15 accelerometers

tion and identification. The obtained signals were passed through an analogical-to-digital converter and stored as a dataset. The sampling rate was 10240 Hz.

4.1 Data acquisition and processing

A “bandpass” filter was applied for noise removal. According to the International Standard ISO 10816-6: 1995, the main excitation frequencies in reciprocating machines typically range from 2 to 1 000 Hz. Additionally, a detrend function was applied to all datasets to eliminate linear trends in the measurements.

Measurements were conducted at constant engine speed while varying torque levels. Torque variation was achieved by injecting different amounts of fuel. The resulting data were grouped into different operational scenarios, as summarized in Table 3.

Table 3 Operation scenarios for data measurement

Scenario	Speed (RPM)	Torque (N·m)	Power (kW)
N°1	1 500	850	127.5
N°2	1 500	2 040	306
N°3	1 500	3 272	409

The dataset was split into two parts before feature extraction: 80% of the data were allocated for training the ML algorithms, and the remaining 20% were reserved for test-

ing (Gholamy et al., 2018). As previously mentioned, measurements were collected for one normal condition and 12 misfire conditions across three different operational scenarios, resulting in a dataset with 13 classes and three observations per class. However, this setting yields a relatively small number of observations, which may lead to a high-dimensional dataset after feature extraction. This scenario, where the number of features notably exceeds the number of observations, can introduce some issues, such as high variance and overfitting (Hastie et al., 2009).

To address this issue, a data augmentation method will be adopted (Martins et al., 2023). Given the large volume of measurement samples, the dataset contains sufficient information to allow multiple subdivisions without losing important data. Therefore, the training and testing datasets will be partitioned (Hassanat et al., 2022). For the training dataset, an optimal partitioning strategy will be explored to achieve the best prediction performance. The number of partitions considered in this study is presented in Table 4. For the test dataset, it was divided into 20 parts to ensure that its size remains less than 30% of the training dataset.

Table 4 Signal Partition

Training dataset		Test dataset	
Partitions	Observations	Partitions	Observations
80	3 120		
90	3 510		
100	3 900	20	780
110	4 290		
120	4 680		

Data was collected for 13 classes across three different operational scenarios. Thus, considering that feature extraction will be applied to each partition, the number of observations in the resulting feature dataset can be represented by Eq. (6):

$$f_v = n_c \cdot n_s \cdot n_p \tag{6}$$

where f_v is the number of observations, n_c is the number of classes, n_s is the number of scenarios, and n_p is the number of partitions.

4.2 Feature extraction

Features were extracted from the collected data using the time-domain analysis. Eighteen (18) statistical features in the time domain were initially selected for feature extraction analysis. The features are described in Table 5, where x_i represents the vibration signal in the time domain, N is the vibration signal length, and $p(c_i)$ denotes the probability that x_i corresponds to a given value in the sequence c_i (Martins et al., 2023).

Table 5 Extracted features (based on Martins et al., 2023)

Feature	Definition
1. Mean	$MEAN = \frac{1}{N} \sum_{i=1}^N x_i$
2. Variance	$VAR = \frac{1}{N} \sum_{i=1}^N (x_i - MEAN)^2$
3. RMS	$RMS = \sqrt{\frac{1}{N} \left[\sum_{i=1}^N (x_i)^2 \right]}$
4. Standard deviation	$STD = \sqrt{VAR}$
5. Kurtosis	$KUR = \frac{1}{N} \sum_{i=1}^N \left(\frac{(x_i - MEAN)}{STD} \right)^4 - 3$
6. Standard error	$STDER = \frac{STD}{\sqrt{N}}$
7. Minimum value	$MIN = \min(x_i)$
8. Peak	$PEAK = \max(abs(x_i))$
9. Peak to peak	$P2P = \max(x_i) - \min(x_i)$
10. Skewness	$SK = \frac{1}{N} \sum_{i=1}^N \left(\frac{(x_i - MEAN)}{STD} \right)^3$
11. Median	$MEDIAN = \text{median}(x_i)$
12. Sum	$SUM = \sum_{i=1}^N x_i$
13. Entropy	$ENTROPY = - \sum_{i=1}^N p(c_i) \text{lb}_p(c_i)$
14. Energy	$ENERGY = \sum_{i=1}^N x_i ^2$
15. Crest factor	$CR = \frac{\max}{RMS}$
16. Clearance factor	$CL = \frac{\max}{\frac{1}{N} \sum_{i=1}^N (x_i)^2}$
17. Shape factor	$SF = \frac{RMS}{\frac{1}{N} \sum_{i=1}^N x_i }$
18. Impulse factor	$IF = \frac{\max}{\frac{1}{N} \sum_{i=1}^N x_i }$

In this study, 18 time-domain features were extracted from each of the 15 accelerometers, resulting in an initial total of 270 attributes.

Outliers were identified using the interquartile range method. Subsequently, they were replaced using the Winsorization method (Blaine, 2018), which substitutes extreme values (those beyond the upper and lower limits) with the respective boundary values. In other words, extremely low

values were replaced by the lower limit, and extremely high values were replaced by the upper limit.

For feature scaling, the Min–Max technique was used because it is a feature normalization method recommended for datasets without outliers (De Amorim et al., 2023). This technique was applied after the removal of outliers from the dataset. This technique scales the values of a feature to a range between a minimum value of zero and a maximum value of 1, with all other values proportionally adjusted within this range. The equation for the Min–Max scaling method is defined in Eq. (7):

$$x_{fs} = \frac{(x_o - x_{min})}{(x_{max} - x_{min})} \tag{7}$$

where x_{fs} is the scaled value of the feature, x_o is the original value of the feature, x_{min} is the minimum value of the feature, and x_{max} is the maximum value of the feature.

Min–Max scaling was applied to only *K*-NN and SVM algorithms because they are sensible to feature scaling. However, this technique was not applied to RF due to its inherent robustness to unscaled data (Singh and Singh, 2022).

5 Results and discussion

5.1 Hyperparameter tuning

For misfire detection and identification, *K*-NN, SVM, and RF were used to classify the 13 classes (normal operation and misfires in each one of the 12 cylinders) based on vibration signals. The performance was evaluated across five different training dataset partitions, as illustrated in Table 4.

The hyperparameters used for each one of the three classification methods are described in Table 6. For cross-validation, a fivefold technique ($k = 5$) was employed. That is, from the 80% of data allocated for training, 20% was reserved for validation in each of the five iterations.

Table 6 Hyperparameters considered for each classification method employed in this study

Method	Hyperparameters	
<i>K</i> -NN	Number of neighbors:	1–10
	Distance method:	Euclidean
	Hyperparameters tuning:	Grid search
SVM	Regularization term C:	From 2^{-5} to 2^{15} , in powers of 4 steps
	Kernel function	Gaussian
	Kernel polynomial order	3
Random forest	Number of trees:	1–30
	Number of leaves:	5
	Hyperparameter tuning:	Grid search

Regarding the evaluation parameter for assessing classification performance, the F1 score was chosen because it balances precision and sensibility metrics, which are critical for this analysis.

Regarding the minimum F1 score required, Chen et al. (2018) stated that a classification performance should be at least 90% of F1-Score to be considered satisfactory. However, aiming for higher standards, this work sets the minimum acceptable F1 score at 95%.

Table 7 shows the F1 score of each of the three ML algorithms, as well as the prediction time per sample (to measure the response time of the algorithm). According to the table, *K*-NN, SVM, and RF achieved F1 scores of 99.87%, 100%, and 99.87%, respectively, using a dataset configuration with 80 partitions and 18 features. These results demonstrate that the employed methods have excellent capability to detect and identify misfires.

For SVM, the F1 score was observed to be 100% across all analyzed partitions. Considering these outstanding results, the risk of overfitting was notably mitigated using data augmentation and cross-validation of the training dataset. Furthermore, previous work by Zhang et al. (2018) achieved 100% classification accuracy for normal operation and misfire detection.

The use of 18 features per sensor, resulting in 270 attributes in total, creates an extensive dataset that leads to high computational cost. To alleviate this issue, dimensionality reduction was applied, enabling the use of fewer features without compromising prediction accuracy and supporting the decision to eliminate less essential sensors.

5.2 Evaluation of accelerometer removal

The main idea behind evaluating sensor removal is feature ranking (Urbanowicz et al., 2018). Feature selection techniques assign scores to each attribute to measure their relevance. Therefore, Relief-F was used to rank the 18 time-domain features across all five training dataset partitions. As shown in Figure 8, each of the five configurations produced a different feature ranking, indicating that feature importance may change depending on the dataset size.

5.3 Evaluation of feature removal

The influence of the number of features on prediction performance was also investigated. In other words, whether using fewer features (and consequently requiring less computational effort) would maintain the capability of ML algorithms to make accurate predictions was verified.

Therefore, simulations were performed using fewer features. Aiming to avoid an exhaustive “step-by-step” search, experiments were conducted with three different feature quantities for performance assessment: 18, 9, and 5 features. The selection criterion was based on the highest-ranked features. This process was repeated for all five training dataset partitions, resulting in 15 configurations.

Then, the accelerometers were ranked in accordance with their importance for classification based on the feature rankings. Each feature is associated with a specific accelerometer; thus, the scores of all features belonging to each sensor were summed. The accelerometers were then sorted in descending order, with the most relevant sensor receiving the highest score. Figure 9 shows the ranking of accelerometers by Relief-F weights for each of the five dataset partitions, considering 5, 9, and 18 features.

To determine the optimal number of sensors, simulations were conducted using the smallest dataset configuration (80 partitions and 5 features). Table 8 shows the F1 score obtained for each classifier as the least important sensors, considering the Relief-F importance criteria, were progressively eliminated. Figure 10 depicts the classification performance trend as sensors are eliminated. As shown in Table 8 and Figure 10, the minimum required F1 score is still achieved after removing 11 accelerometers. In other words, the minimum effective number of sensors required is four.

However, this outcome cannot be considered definitive because simulations must be extended to the remaining 14 dataset configurations to explore the possibility of further reducing the number of sensors. Performing the same exhaustive simulations for each dataset would be highly time-consuming. To avoid this phenomenon, the classification performance analysis began with the last four remaining sensors, with results shown in Table 9. Subsequently, perfor-

Table 7 F1 score for a dataset with 18 features

N° of Partitions	<i>K</i> -NN		SVM		RF	
	F1 Score (%)	Time (s)	F1 Score (%)	Time (s)	F1 Score (%)	Time (s)
80	99.87%	0.0026915	100%	0.000806	99.87%	0.000388
90	99.74%	0.0032461	100%	0.000765	88.58%	7.36×10^{-5}
100	99.87%	0.0031457	100%	0.000735	38.1%	0.000106
110	99.74%	0.0035323	100%	0.001179	NaN	–
120	100%	0.0043258	100%	0.000780	NaN	–

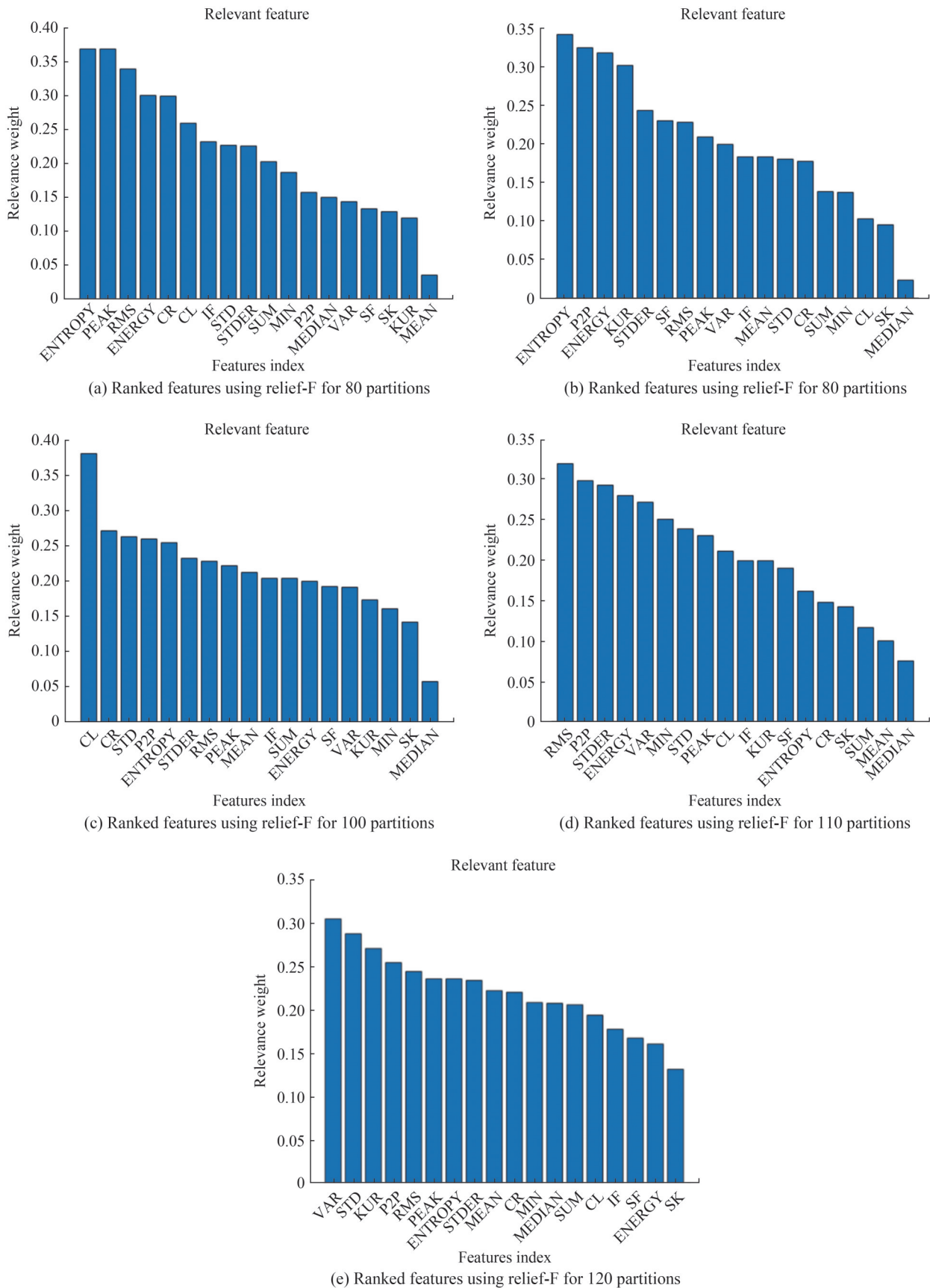


Figure 8 Ranked features using Relief-F

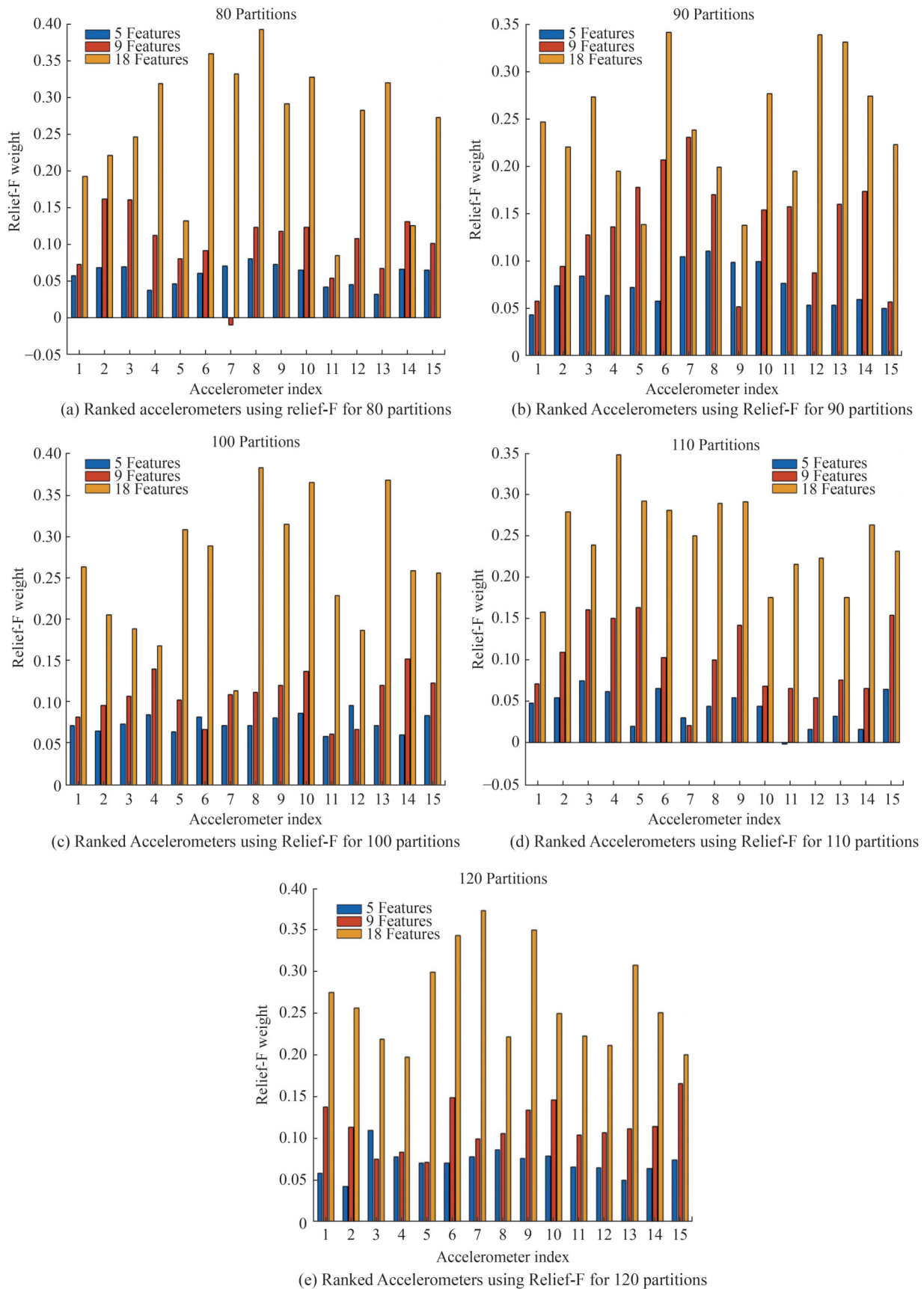


Figure 9 Ranked accelerometers using Relief-F

Table 8 F1 score for a dataset with 80 partitions and 5 features

N° of Eliminated sensors	F1 Score (%)		
	K-NN	SVM	RF
0	99.36	100	99.87
1	99.49	99.87	100
2	99.23	99.87	99.74
3	97.47	99.74	99.36
4	96.96	99.74	99.36
5	97.84	99.49	99.36
6	97.60	99.62	98.97
7	97.72	99.49	99.23
8	96.71	98.98	98.72
9	97.72	98.73	98.46
10	96.71	98.85	98.46
11	95.94	97.61	96.91
12	89.52	92.46	92.30
13	82.20	83.82	88.93
14	47.33	48.33	62.76

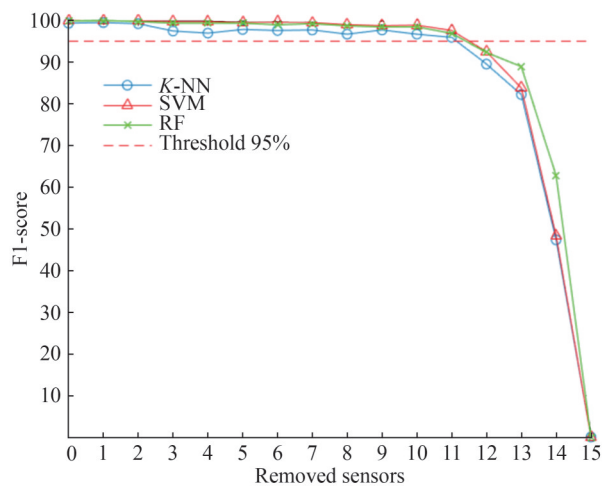


Figure 10 F1 score for a dataset with 80 Partitions and 5 Statistical features

Table 9 F1 score for 11 sensors eliminated

N° of Partitions	F1 Score (%)								
	5 Features			9 Features			18 Features		
	K-NN	SVM	RF	K-NN	SVM	RF	K-NN	SVM	RF
80	95.94	97.61	96.91	88.29	94.75	94.97	95.01	99.84	98.46
90	95.95	96.05	NaN	97.05	98.45	47.16	96.40	99.10	68.71
100	85.48	93.02	NaN	81.51	93.64	NaN	97.95	99.10	30.05
110	87.29	89.08	NaN	80.02	85.62	NaN	95.89	98.58	NaN
120	93.77	91.41	96.15	73.18	85.38	NaN	97.69	98.98	NaN

mance was assessed after eliminating 12 sensors, as shown in Table 10. An additional attempt was made using only two sensors, but no combination achieved the minimum required F1 score. Therefore, the minimum number of accelerometers that can be used is three.

Table 11 shows the three selected accelerometers for each simulation (considering Relief-F ranking in Figure 8), along with the respective classifier and prediction time for a single observation. The criterion used to determine the best sensor combination was the shortest prediction time. This condition was achieved using sensors n°4, n°6, and n°15 with a dataset configuration of 80 partitions and 18 features. A radar plot detailing the classification performance obtained using this optimal sensor combination is depicted in Figure 11.

Subsequently, the prediction capability of the classifier was analyzed using only a single sensor. This process was repeated for each accelerometer across all 15 dataset configurations. SVM outperformed K-NN and RF in misfire detection and identification; thus, it was adopted as the classification method for this phase of the study. The results are depicted in Figure 12. In some cases, an F1 score above 80% was observed using only one sensor.

Next, the possibility of combining two sensors to obtain an F1 score above 95% was evaluated. Two selection criteria were applied to determine the best pair of accelerometers. The first criterion was based on the two highest-ranked sensors according to Relief-F weights for each dataset configuration. The second criterion involved selecting the sensor pair that achieved the best F1 score using the GCA method within the same dataset configuration. For both cases, classification performance was assessed, with results presented in Tables 12–14, including F1 score and prediction time per sample for each analysis.

According to the results presented in the tables, the GCA criterion outperformed Relief-F in almost all cases, not only in terms of classification performance but also in terms of prediction time. Furthermore, the sensor pairs selected using the Relief-F ranking method only achieved a minimum F1 score of 95%. In contrast, the GCA method achieved the target performance in six cases: one corresponding to the dataset with 9 features and 90 partitions, and the remaining

Table 10 F1 score for 12 sensors eliminated

N° of Partitions	F1 Score (%)								
	5 Features			9 Features			18 Features		
	K-NN	SVM	RF	K-NN	SVM	RF	K-NN	SVM	RF
80	89.52	92.46	92.30	82.52	90.46	90.01	91.63	93.95	95.04
90	88.40	91.74	NaN	91.28	92.95	46.98	93.39	97.36	57.81
100	81.74	90.76	NaN	74.28	84.67	NaN	94.98	98.07	NaN
110	76.90	78.12	NaN	73.72	80.29	NaN	93.32	97.20	NaN
120	86.11	85.24	92.34	74.34	81.12	NaN	95.31	96.78	NaN

Table 11 Selection of sensors using the Relief-F method, considering dataset configurations

N° of Partitions	N° of Features	Selected Sensors	F1 Score (%)	Classifier	Time (s)
80	18	n°15, n°6 e n°4	96.36	RF	4.12692×10^{-5}
90	18	n°6, n°13 e n°14	97.36	SVM	8.39359×10^{-5}
100	18	n°8, n°13 e n°10	98.07	SVM	9.10538×10^{-5}
110	18	n°4, n°5 e n°9	97.20	SVM	9.43026×10^{-5}
120	18	n°7, n°9 e n°6	96.78	SVM	0.000108241

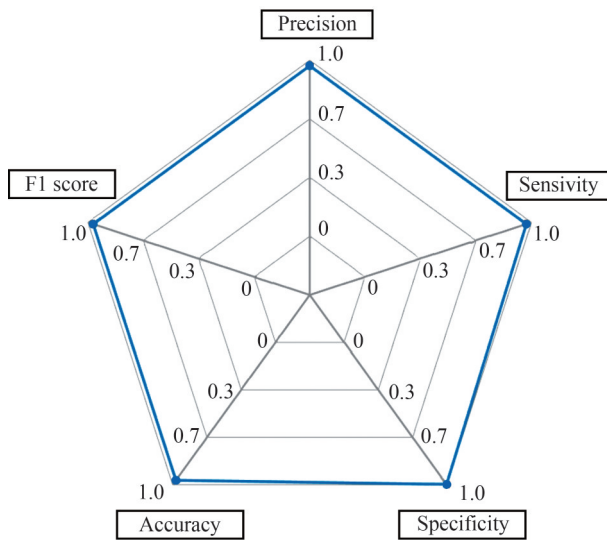


Figure 11 Radar Plot for Relief-F classification when using sensors n°4, n°6, and n°15 in a dataset configuration of 80 partitions and 18 features

five across all dataset configurations with 18 features. These findings highlight the possibility of using only two sensors while still maintaining satisfactory prediction performance.

Finally, the study assessed which sensor pair should be recommended for implementation in this method. In the current analysis, three different pairs of accelerometers reached an F1 score above 95%: sensors n°2 with n°9, n°4

with n°13, and n°9 with n°13. Among these, the selection criterion was the shortest prediction time. Therefore, the best-performing pair was sensors 2 and 9, using a dataset with 90 partitions and 9 features. A radar plot detailing the classification performance with this selected sensor combination is depicted in Figure 13.

Although the GCA method yielded better results, its overall procedure required substantially more computational time compared to the Relief-F process to determine the optimal number of sensors. Considering that the sensor elimination procedure based on the Relief-F criteria is faster and was able to eliminate up to 12 accelerometers while maintaining acceptable classification performance, the Relief-F method is more suitable for sensor quantity optimization than the GCA method.

Although Xi et al. (2018) stated that determining the best sensor position is impossible due to the complexity of engine structures and the unknown location of engine faults, this study was able to identify sensor positions that yielded good prediction performance.

The results showed that other sensor combinations also achieved the required F1 score of 95%, indicating that various sensor placements can lead to satisfactory classification results. Therefore, the optimal sensor positions depend on multiple factors, including the type of failure, the sensor type, and the method employed for misfire detection and identification (i.e., data structuring, feature selection, and classification method).

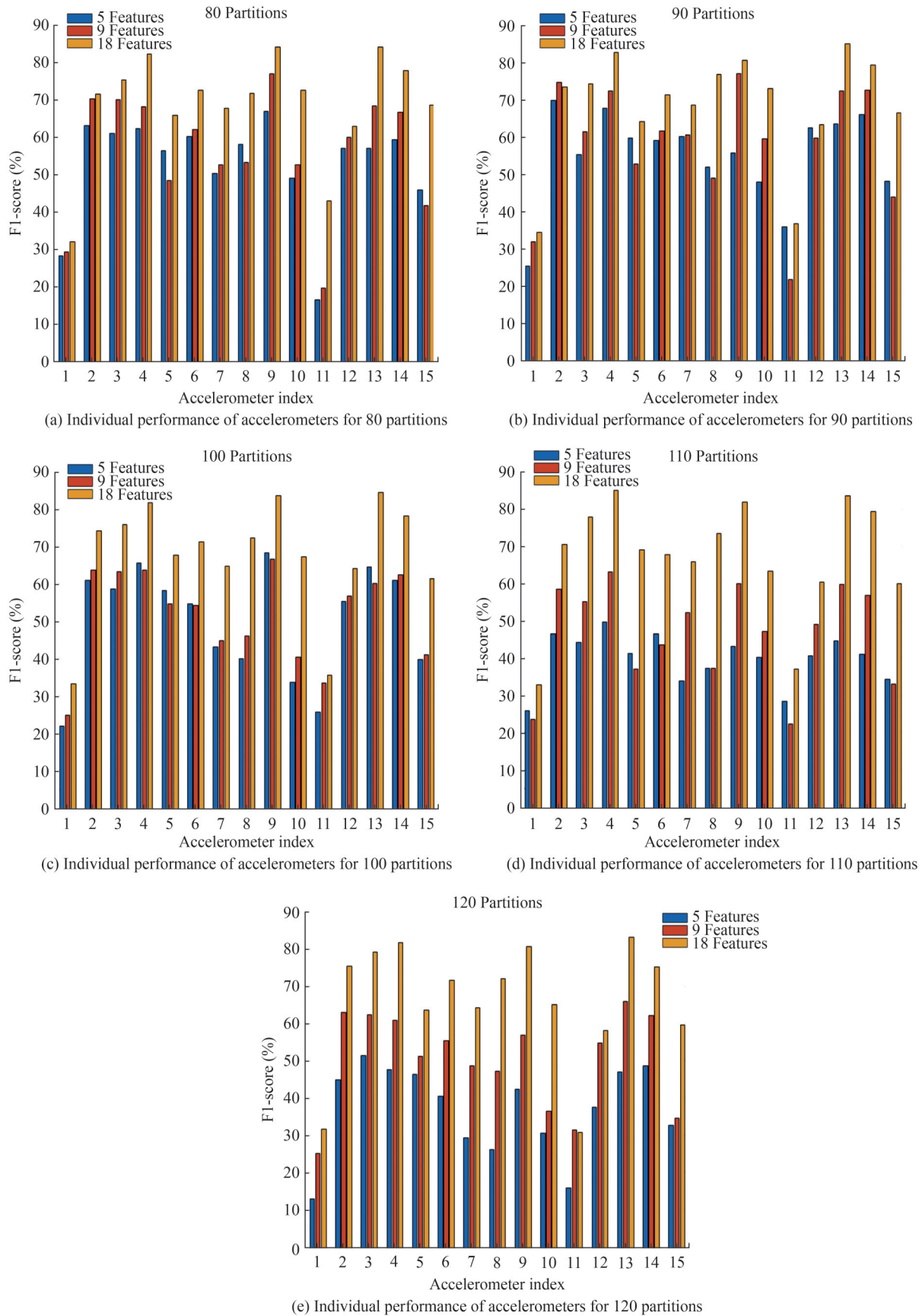


Figure 12 Individual performance of accelerometers

Table 12 Prediction performance for two sensors using SVM in a dataset with five features

N° Partitions	SVM for Five Features					
	Relief-F			GCA		
	Sensors	F1 Score (%)	Time (s)	Sensors	F1 Score (%)	Time (s)
80	n°8; n°9	87.95	6.201×10^{-5}	n°9; n°2	91.51	8.08846×10^{-5}
90	n°8; n°7	77.45	0.0006927	n°2; n°4	90.62	5.81359×10^{-5}
100	n°12; n°10	73.59	9.431×10^{-5}	n°9; n°4	90.80	6.64282×10^{-5}
110	n°4; n°5	66.58	0.0001432	n°4; n°2	75.72	0.000117333
120	n°3; n°8	70.46	0.0001132	n°3; n°14	80.98	0.000275264

Table 13 Prediction performance for two sensors using SVM in a dataset with nine features

N° Partitions	SVM for Nine features					
	Relief-F			GCA		
	Sensors	F1 Score (%)	Time (s)	Sensors	F1 Score (%)	Time (s)
80	n°2; n°3	85.07	6.664×10^{-5}	n°9; n°2	94.65	6.27936×10^{-5}
90	n°7; n°6	86.25	9.343×10^{-5}	n°9; n°2	96.36	5.96551×10^{-5}
100	n°14; n°4	80.45	9.331×10^{-5}	n°9; n°2	89.92	0.0000828
110	n°5; n°3	69.37	8.088×10^{-5}	n°4; n°9	89.03	6.77282×10^{-5}
120	n°15; n°6	68.47	0.0001085	n°13; n°2	89.55	0.000107968

Table 14 Prediction performance for two sensors using SVM in a dataset with 18 features

N° Partitions	SVM For 18 features					
	Relief-F			GCA		
	Sensors	F1 Score (%)	Time (s)	Sensors	F1 Score (%)	Time (s)
80	n°8; n°6	89.20	8.986×10^{-5}	n°9; n°13	96.93	8.12936×10^{-5}
90	n°6; n°12	86.03	9.956×10^{-5}	n°13; n°4	96.68	8.25833×10^{-5}
100	n°8; n°13	94.46	9.234×10^{-5}	n°13; n°9	96.92	8.89231×10^{-5}
110	n°4; n°5	92.70	0.0001054	n°4; n°13	96.67	9.08538×10^{-5}
120	n°7; n°9	94	0.0001052	n°13; n°4	95.64	0.000174871

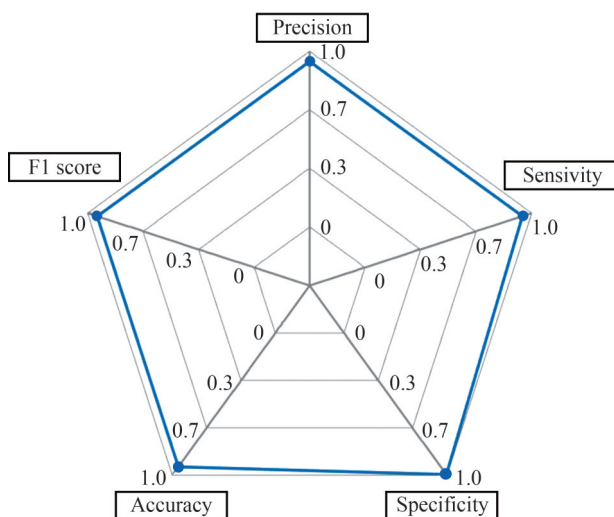


Figure 13 Radar Plot for GCA classification when using sensors n°2 and n°9 in a dataset configuration of 90 partitions and 9 features

6 Conclusions and recommendations

The results showed outstanding performance from all three analyzed ML models in detecting and identifying misfires, with *K*-NN, SVM, and RF achieving F1 scores of 99.87%, 100% and 99.87%, respectively, for a dataset comprising 18 features with 80 partitions. However, the large dataset size resulted in high computational costs.

Aiming to mitigate this issue, Relief-F was applied for dimensionality reduction, generating dataset configurations with nine and five features. Subsequently, accelerometers were ranked by importance based on Relief-F weights, and this process was repeated for each of the 15 dataset configurations.

The verification of the minimum number of sensors required to achieve the desired prediction performance was conducted using the following two methods: the first

relied on selecting the top-ranked accelerometers according to Relief-F weights, and the second employing the GCA method, which involves combining sensors until the required performance is reached.

Results show that the sensors selected by GCA not only allowed the use of only two accelerometers while maintaining good prediction performance but also outperformed the pairs of accelerometers selected by Relief-F.

However, although the GCA method showed better results using only two accelerometers and nine features, its overall procedure requires substantially more computational time compared to Relief-F to reach the optimal number of sensors. Considering the excellent results obtained by Relief-F with three accelerometers and 18 features, Relief-F is more suitable for accelerometer selection than the GCA method.

Regarding the optimal sensor placement on the engine block, the results of this study indicate that multiple combinations of sensor pairs can achieve good prediction performance. In other words, the positions of the accelerometers on the engine may vary depending on the methods used for misfire detection and identification (i.e., data structuring, feature selection, and classification method).

For future work, one suggestion is to analyze the dataset in the frequency domain, with a focus on the phases of frequency peaks. Another recommendation is to explore the time-frequency domain to evaluate the capability of these attributes to characterize faulty signals. Additionally, studying cases with multiple simultaneous misfires in a 12-cylinder V-type diesel engine is suggested, aiming to develop methods for the detection and identification of such faults.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Abu Alfeilat HA, Hassanat AB, Lasassmeh O, Tarawneh AS, Alhasanat MB, Eyal Salman HS, Prasath VS (2019) Effects of distance measure choice on k-nearest neighbor classifier performance: a review. *Big Data* 7(4): 221-248. <https://doi.org/10.1089/big.2018.0175>
- Blaine BE (2018) Winsorizing. *The sage encyclopedia of educational research, measurement, and evaluation*: 1817-1818. <https://doi.org/10.4135/9781506326139.n747>
- Boateng EY, Otoo J, Abase DA (2020) Basic tenets of classification algorithms k-nearest-neighbor, support vector machine, random forest and neural network: a review. *Journal of Data Analysis and Information Processing* 8(4): 341-357. <https://doi.org/10.4236/jdaip.2020.84020>
- Boisberranger J, Estève L, Fan TJ, Gramfort A, Grisel O (2023) Support Vector Machines. In: *Scikit-learn*. Available via dialog. <https://scikit-learn.org/stable/modules/svm.html>. Accessed 3 Aug 2023
- Bolón-Canedo V, Sánchez-Marroño N, Alonso-Betanzos A (2013) A review of feature selection methods on synthetic data. *Knowledge and Information Systems* 34: 483-519. <https://doi.org/10.1007/s10115-012-0487-8>
- Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In *Proceedings of The Fifth Annual Workshop on Computational Learning Theory*: 144-152. <https://doi.org/10.1145/130385.130401>
- Breiman L (2001) Random forests. *Machine learning* 45(1): 5-32. <https://doi.org/10.1023/A:1010933404324>
- Breiman L (1996) Bagging predictors. *Machine learning* 24: 123-140. <https://doi.org/10.1007/BF00058655>
- Charles P, Sinha JK, Gu F, Lidstone L, Ball AD (2009) Detecting the crankshaft torsional vibration of diesel engines for combustion related diagnosis. *Journal of Sound and Vibration* 321(3-5): 1171-1185. <https://doi.org/10.1016/j.jsv.2008.10.024>
- Chen SK, Mandal A, Chien LC, Ortiz-Soto E (2018) Machine learning for misfire detection in a dynamic skip fire engine. *Sae International Journal of Engines* 11(6): 965-976. <https://www.jstor.org/stable/26649141>
- Cortes C, Vapnik V (1995) Support-vector networks. *Machine Learning* 20: 273-297. <https://doi.org/10.1007/BF00994018>
- Cui Y, Liu H, Wang Q, Zheng Z, Wang H, Yue Z, Yao M (2022a) Investigation on the ignition delay prediction model of multi-component surrogates based on back propagation (BP) neural network. *Combustion And Flame* 237: 111852. <https://doi.org/10.1016/j.combustflame.2021.111852>
- Cui Y, Liu H, Wen M, Feng L, Ming Z, Zheng Z, Yao M (2022b) Optical diagnostics of misfire in partially premixed combustion under low load conditions. *Fuel* 329: 125432. <https://doi.org/10.1016/j.fuel.2022.125432>
- Cutler A, Cutler DR, Stevens JR (2012) *Random forests*. Ensemble Machine Learning: 157-175. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-9326-7_5
- De Amorim LB, Cavalcanti GD, Cruz RM (2023) The choice of scaling technique matters for classification performance. *Applied Soft Computing* 133: 109924. <https://doi.org/10.1016/j.asoc.2022.109924>
- Devasenapati SB, Sugumaran V, Ramachandran KI (2010) Misfire identification in a four-stroke four-cylinder petrol engine using decision tree. *Expert Systems with Applications* 37(3): 2150-2160. <https://doi.org/10.1016/j.eswa.2009.07.061>
- Firmino JL, Neto JM, Oliveira AG, Silva JC, Mishina KV, Rodrigues MC (2021) Misfire detection of an internal combustion engine based on vibration and acoustic analysis. *Journal of The Brazilian Society of Mechanical Sciences and Engineering* 43(7): 336. <https://doi.org/10.1007/s40430-021-03052-y>
- Fratello M, Tagliaferri R (2018) Decision trees and random forests. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*: 374-383. <https://dx.doi.org/10.1016/B978-0-12-809633-8.20337-3>
- Gholamy A, Kreinovich V, Kosheleva O (2018) Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. [https://doi.org/10.6148/IJTAS.201806_11\(2\).0003](https://doi.org/10.6148/IJTAS.201806_11(2).0003)
- Hassanat AB, Tarawneh AS, Abed SS, Altarawneh GA, Alrashidi M, Agamid M (2022) Rdpvr: Random data partitioning with voting rule for machine learning from class-imbalanced datasets. *Electronics* 11(2): 228. <https://doi.org/10.3390/electronics11020228>
- Hastie T, Tibshirani R, Friedman JH, Friedman JH (2009) *The elements of statistical learning: data mining, inference, and prediction*, Vol. 2: 1-758. New York: Springer. <https://doi.org/10.1007/978-0-387-21606-5>
- ISO 10816-6 (1995) *Mechanical vibration-Evaluation of machine vibration by measurements on non-rotating parts-Part 6: Reciprocating machines with power ratings above 100 Kw*
- Jafarian K, Mobin M, Jafari-Marandi R, Rabiei E (2018) Misfire and

- valve clearance faults detection in the combustion engines based on a multi-sensor vibration signal monitoring. *Measurement* 128: 527-536. <https://doi.org/10.1016/j.measurement.2018.04.062>
- Jović A, Brkić K, Bogunović N (2015) A review of feature selection methods with applications. In 2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO): 1200-1205. IEEE. <https://doi.org/10.1109/MIPRO.2015.7160458>
- Kononenko I (1994) Estimating attributes: Analysis and extensions of RELIEF. In European conference on machine learning: 171-182. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-57868-4_57
- Kramer O (2013) Dimensionality reduction with unsupervised nearest neighbors Vol. 51: 13-23. Berlin: Springer. <https://doi.org/10.1007/978-3-642-38652-7>
- Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK (2020) Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems And Signal Processing* 138: 106587. <https://doi.org/10.1016/j.ymsp.2019.106587>
- Lujan JM, Bermudez V, Guardiola C, Abbad A (2010) A methodology for combustion detection in diesel engines through in-cylinder pressure derivative signal. *Mechanical Systems and Signal Processing* 24(7): 2261-2275. <https://doi.org/10.1016/j.ymsp.2009.12.012>
- Mammone A, Turchi M, Cristianini N (2009) Support vector machines. *Wiley Interdisciplinary Reviews: Computational Statistics* 1(3): 283-289. <https://doi.org/10.1002/wics.49>
- Martins DHCSS, De Lima AA, Pinto MF, Hemerly DO, Prego TM, Silva FL, Tarrataca L, Monteiro UA, Gutiérrez RHR, Haddad DB (2023) Hybrid data augmentation method for combined failure recognition in rotating machines. *Journal of Intelligent Manufacturing* 34: 1795-1813. <https://doi.org/10.1007/s10845-021-01873-1>
- Maurya RK (2019) Reciprocating engine combustion diagnostics. IN-Cylinder Pressure Measurement and Analysis. Cham, Switzerland: Springer International Publishing. <https://doi.org/10.1007/978-3-030-11954-6>
- Merhige R (2016) Engine misfire largely to blame for vibration onboard. In: Triton. Available via dialog. <https://www.the-triton.com/2016/09/engine-misfire-largely-to-blame-for-vibration-onboard/>. Accessed 20 Dec 2022
- MTU Detroit Diesel (2016) Surface Mining 12V4000-T1237K11. Detroit: MTU Detroit Diesel, 20p
- Nazarenko E, Varkentin V, Polyakova T (2019) Features of application of machine learning methods for classification of network traffic (features, advantages, disadvantages). In 2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon): 1-5. IEEE. <https://doi.org/10.1109/FarEastCon.2019.8934236>
- Noble WS (2006) What is a support vector machine? *Nature biotechnology* 24(12): 1565-1567. <https://doi.org/10.1038/nbt1206-1565>
- Phyu TZ, Oo NN (2016) Performance comparison of feature selection methods. In MATEC web of conferences Vol. 42: 126, 521-535. <https://doi.org/10.1051/mateconf/20164206002>
- Robnik-Šikonja M, Kononenko I (2003) Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning* 53: 23-69. <https://doi.org/10.1023/A:1025667309714>
- Sharma A, Sugumaran V, Devasenapati SB (2014) Misfire detection in an IC engine using vibration signal and decision tree algorithms. *Measurement* 50: 370-380. <https://doi.org/10.1016/J.MEASUREMENT.2014.01.018>
- Singh D, Singh B (2022) Feature wise normalization: An effective way of normalizing data. *Pattern Recognition* 122: 108307. <https://doi.org/10.1016/j.patcog.2021.108307>
- Tamura M, Saito H, Murata Y, Kokubu K, Morimoto S (2011) Misfire detection on internal combustion engines using exhaust gas temperature with low sampling rate. *Applied Thermal Engineering* 31(17-18): 4125-4131. <https://doi.org/10.1016/j.applthermaleng.2011.08.026>
- Tao J, Qin C, Li W, Liu C (2019) Intelligent fault diagnosis of diesel engines via extreme gradient boosting and high-accuracy time-frequency information of vibration signals. *Sensors* 19(15): 3280. <https://doi.org/10.3390/s19153280>
- Taunk K, De S, Verma S, Swetapadma A (2019) A brief review of nearest neighbor algorithm for learning and classification. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS): 1255-1260. IEEE. <https://doi.org/10.1109/ICCS45141.2019.9065747>
- Ukil A (2007) Intelligent systems and signal processing in power engineering. Springer Science & Business Media. <https://doi.org/10.1007/978-3-540-73170-2>
- Urbanowicz RJ, Meeker M, La Cava W, Olson RS, Moore JH (2018) Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics* 85: 189-203. <https://doi.org/10.1016/j.jbi.2018.07.014>
- Vakharia V, Gupta VK, Kankar PK (2017) Efficient fault diagnosis of ball bearing using ReliefF and Random Forest classifier. *Journal of The Brazilian Society of Mechanical Sciences and Engineering* 39(8): 2969-2982. <https://doi.org/10.1007/s40430-017-0717-9>
- Verikas A, Gelzinis A, Bacauskiene M (2011) Mining data with random forests: A survey and results of new tests. *Pattern Recognition* 44(2): 330-349. <https://doi.org/10.1016/j.patcog.2010.08.011>
- Wang C, Yue Z, Zhao Y, Ye Y, Liu X, Liu H (2023). Numerical simulation of the high-boosting influence on mixing, combustion and emissions of high-power-density engine. *Journal of Thermal Science* 32(3): 933-946. <https://doi.org/10.1007/s11630-023-1796-9>
- Xi W, Li Z, Tian Z, Duan Z (2018) A feature extraction and visualization method for fault detection of marine diesel engines. *Measurement* 116: 429-437. <https://doi.org/10.1016/j.measurement.2017.11.035>
- Zhang M, Zi Y, Niu L, Xi S, Li Y (2018) Intelligent diagnosis of V-type marine diesel engines based on multifeatures extracted from instantaneous crankshaft speed. *IEEE Transactions on Instrumentation and Measurement* 68(3): 722-740. <https://doi.org/10.1109/TIM.2018.2857018>
- Zhang P, Gao W, Li Y, Wang Y (2021) Misfire detection of diesel engine based on convolutional neural networks. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 235(8): 2148-2165. <https://doi.org/10.1177/0954407020987077>
- Zhang Z, Liu H, Yue Z, Li Y, Liang H, Kong X, Yao M (2023) Effects of intake high-pressure compressed air on thermal-work conversion in a stationary diesel engine. *International Journal of Green Energy* 20(3): 338-351. <https://doi.org/10.1080/15435075.2022.2040509>
- Zheng T, Zhang Y, Li Y, Shi L (2019) Real-time combustion torque estimation and dynamic misfire fault diagnosis in gasoline engine. *Mechanical Systems and Signal Processing* 126: 521-535. <https://doi.org/10.1016/j.ymsp.2019.02.048>
- Zhou H, Zhang J, Zhou Y, Guo X, Ma Y (2021) A feature selection algorithm of decision tree based on feature weight. *Expert Systems with Applications* 164, 113842. <https://doi.org/10.1016/j.eswa.2020.113842>