Applied and Computational Mechanics 20 (2026) XXX–YYY

Comparative analysis of support vector machine and logistic regression for gear fault detection and classification

A. Lakikza[a], T. Khoualdia[a], A. Lakehal[a,*], F. Mahfoudi[a], N. Guerti[a], S. Aknouch[a]

[a] *Laboratory of Research on Electromechanical and Dependability, University of Souk Ahras, Souk Ahras, Algeria*

Received 25 July 2025; accepted 31 January 2026

Abstract

Bevel gears play a crucial role in mechanical systems, particularly in power transmission applications where changes in shaft orientation are required. Early fault detection and diagnosis in gearboxes are essential for ensuring operational efficiency and preventing costly failures. This study evaluates and compares the effectiveness of logistic regression (LR) and support vector machine (SVM) classification methods in identifying gear faults, focusing on data-driven condition monitoring rather than predictive maintenance. To enhance the performance of these models, hyperparameter tuning was performed using grid search cross-validation. These techniques are essential for improving model performance, reducing overfitting and increasing classification accuracy. Gear fault classification was carried out using vibration data from a test bench under different speeds, loads, and measurement directions. While LR initially achieved higher accuracy (64.10 %) compared to SVM (38.46 %), hyperparameter tuning significantly improved SVM performance, allowing it to reach an accuracy of 92.31 % compared to 82.05 % in the case of LR. These findings underscore the capability of the optimized SVM model to provide more sensitive and precise fault diagnosis, highlighting its suitability for robust data-driven diagnostics of gear conditions.

Keywords: bevel gears fault diagnosis, support vector machine, logistic regression, grid search cross-validation, data-driven diagnostics

1. Introduction

Gearbox fault diagnosis is essential for ensuring the reliability and safety of mechanical transmission systems. Early detection of gear defects helps to prevent unexpected failures, reduce downtime and maintain operational performance. With the growing availability of vibration data and advancements in machine learning, data-driven methods have become powerful tools for identification and classification of gear faults [3]. Today, pattern recognition, object detection, language interpretation and other academic subjects use machine learning. Artificial intelligence (AI) creates data-trend-based algorithms using machine learning [6]. In the recent decade, various machine learning methods have been used for gearbox fault diagnosis. We can mention artificial neural networks (ANN) [18], logistic regression (LR) [10], support vector machine (SVM) [7], decision tree-based ensemble learning [2], random forest [4] and gradient boosting decision tree [24]. Data models are analyzed using classification. It uses future data forecasts to identify sample trends. Classification is important and widely used and uses machine learning to group and predict data. The elements are then categorized to increase prediction accuracy. After refining training data, classification techniques focus on dataset paths to conclusion.

*Corresponding author. Tel.: +213 675 789 468, e-mail: a.lakehal@univ-soukahras.dz.
https://doi.org/10.24132/acm.2026.1024

In this contribution, we consider the following classification methods: LR and SVM. The first one is used to characterize data and explain the relationship between a binary response variable (indicating two possible outcomes) and a binary predictor variable (consisting of two categories) with one or more ordinal (ranked categorical variables), nominal (variables with multiple categories), interval (where measurement differences are significant), or ratio-level independent variables. On the other hand, LR works for dichotomous dependent variables but can be used for polytomous or multinomial dependent variables. Based on predictor variable values, the LR algorithms determine the chance of a categorical response variable being classified [17]. A popular method for optimizing predicted results is SVM. The classification problem uses a function based on examples to distinguish two categories, ensuring that the classifier works effectively on fresh instances and has high generalization. It should be recalled that multiple linear classifiers can separate data, but only one optimizes margin. A linear classifier is called the optimal separating hyperplane [21].

Gears are used in cars, machines, and wind turbine equipment. Unexpected equipment faults can lead to functional failures and result in significant operational or financial losses. Early gear fault diagnosis prevents major incidents, maintains industrial machinery performance, and protects operators. Condition-based maintenance of gearbox systems is mainly based on lubricant analysis and sound and vibration monitoring. Vibration analysis is rigorously investigated for early concerns and vibration signals are essential gearbox indicators. Alongside these methods, AI has given more robustness and reliability to diagnosis and decision-making. In [1], Abdul and Al-Talabani proposed an SVM model based on feature concatenation. Despite the significant contributions to existing gear diagnostic models in the literature, the authors demonstrated that their model allows high effectiveness of gear fault detection. A global representation by concatenating frame-based features outperforms global statistical and time-series feature representations. In other works, SVM is combined with other AI techniques to perform more reliable and precise fault diagnosis in gearbox systems. For example, for online fault diagnosis and monitoring of a critical gearbox, a nonparametric filter technique gives a good solution for an optimized usage of the internet of things (IoT) server [14]. The proposed solution technique integrates the energy operator, genetic algorithm, and SVM. Fault diagnosis can be further improved by achieving high classification accuracy, which is also attainable using an SVM classifier optimized with the whale optimization algorithm [16]. Although SVM classifiers are widely applied for gearbox fault diagnoses, researchers have continued to enhance their performance through various optimization techniques. As a result, several hybrid approaches have been developed, including SVM classifiers optimized using the grey wolf optimizer [26], particle swarm optimization [27], and the genetic algorithm [15]. In this paper, a hybrid classifier based on SVM optimized through grid search cross-validation is proposed.

The LR algorithm is widely used and has shown promising results in fault diagnosis in geared power transmission systems and mechanical systems in general. To choose the best AI techniques, several scientists have compared these techniques (for example: decision tree, LR, SVM, and ANN), but results have shown that each technique has advantages compared to the others and also depends on the system which we try to diagnose [20, 23]. Fault diagnosis by the LR algorithm can be improved by using other AI techniques. A combination of LR and the stochastic gradient descent (LR-SGD) gives a strong accuracy when used to gearbox fault diagnosis. In this case, the accuracy can be doubled or tripled compared to using only LR [8]. The same as SVM goes for this technique; it will be combined with grid search cross-validation (GSCV-LR) to explore results and advantages.

Although numerous machine learning methods have been applied to gear fault diagnosis in the past, most of the existing studies have focused on complex models or hybrid optimization techniques, while the comparative performance of simpler and widely used classifiers, such as LR and SVM, remains insufficiently explored for bevel gears. This knowledge gap is important because bevel gear transmissions are critical components in many industrial applications and reliable fault classification requires models that are both accurate and computationally efficient. Furthermore, the influence of hyperparameter tuning on the performance of these baseline classifiers has not been systematically evaluated under varying operating conditions.

To address this gap, the present study conducts a structured comparison between LR and SVM models for multiclass bevel gear fault classification using vibration features. The analysis evaluates accuracy, precision, recall, and F1-Score both before and after applying grid search cross-validation (GSCV), allowing a clear assessment of how tuning affects each classifier. This contribution provides practical insights into the suitability of these two commonly used algorithms for vibration-based condition monitoring and establishes a foundation for future work using larger datasets, different feature types and more advanced diagnostic models.

2. Methodology

2.1. Data description

The time-domain analysis method employs several traditional time-domain signal features for vibration assessment, such as peak, peak-to-peak, mean, root mean square (RMS), crest factor, skewness, kurtosis, clearance factor, impulse factor, and form factor [9]. Table 1 presents the most commonly used temporal vibration features, combining statistical and physically interpretable measures [13]. These features are essential for analyzing vibration signals and identifying potential faults in rotating machinery, particularly in bevel gears [12]. Consequently, time-domain signal features effectively capture both the operating condition and the characteristics of the pinion within the gearbox. In this study, classical and widely adopted vibration features were selected and computed from the time-domain signal [11, 25].

After computing the vibratory features listed in Table 1, a data file was generated to form the input matrix $\mathbf{X}$ used for algorithms training. This file contains the simulated features corresponding to the different operational states of the pinion and the results were compiled into a CSV dataset. The data are organized according to four classes: No Fault, Fault 1, Fault 2, and Fault 3.

The dataset consists of j rows, each representing a complete feature set, including the maximum value, RMS, mean square value, variance, kurtosis and others. Each row corresponds to k extracted vibration features obtained under different combinations of rotational speed, load, direction, and pinion condition. In addition to these continuous vibration features, the last columns of the dataset include categorical variables describing the operating conditions, namely: rotational speed, load level, and measurement direction. These categorical variables are later encoded (via one-hot encoding) before model training. Consequently, the input matrix $\mathbf{X}$ is expressed as

$$\mathbf{X} = \begin{bmatrix} X_{1,1} & \dots & X_{1,k} & c_{1,1} & \dots & c_{1,m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ X_{j,1} & \dots & X_{j,k} & c_{j,1} & \dots & c_{j,m} \end{bmatrix}, \tag{1}$$

where $\mathbf{X}$ is the vibration feature matrix, j denotes the number of rows, each representing a specific pinion state obtained under different speed, load, and measurement direction conditions,

Table 1. Statistical and physical time-domain features, where x_i represents individual data points, N is the number of data points, and σ is the standard deviation

Feature	Formula	Description
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	Represents the central tendency of the data by calculating the average of all values in the dataset.
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2$	Measures the spread of data points around the mean, indicating how dispersed the values are.
Skewness	$S = \frac{1}{N}\sum_{i=1}^{N} \left(\frac{x_i - \mu}{\sigma}\right)^3$	Describes the asymmetry of the data distribution. Positive skewness suggests a longer right tail, while negative skewness indicates a longer left tail.
Kurtosis	$K = \frac{1}{N}\sum_{i=1}^{N} \left(\frac{x_i - \mu}{\sigma}\right)^4$	Measures the distribution peaks. Higher kurtosis means more extreme values (sharp peak), while lower kurtosis indicates fewer extreme values (flatter peak).
Absolute mean	$\mu_a = \frac{1}{N}\sum_{i=1}^{N} \lvert x_i \rvert$	Measures the average magnitude of a signal, ignoring sign variations. It is useful in vibration analysis to assess both positive and negative fluctuations.
Peak-to-peak	$P_p = x_{\max} - x_{\min}$	Determines the total range of variation between the maximum and minimum values.
Root mean square	$x_{\text{RMS}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$	Computes the square root of the mean of the squared data points, providing an indicator of signal magnitude for both positive and negative values.
Impulse factor	$I_f = \frac{x_{\max}}{\mu_a}$	The ratio of the maximum value to the absolute mean, useful for detecting transient spikes in the data.
Shape factor	$S_f = \frac{x_{\text{RMS}}}{\mu_a}$	The ratio of RMS to the absolute mean, offering insight into the waveform shape and how energy is distributed.

k is the number of extracted continuous vibration features, and m is the number of categorical variables after encoding (e.g., one-hot encoded speed, load, and direction). This structure ensures that each sample contains both the relevant vibrational characteristics and the operational context required for accurate gear fault classification.

2.2. Preprocessing and standardization

This study characterizes each sample using a collection of statistical time-domain features derived from vibration signals (e.g., mean, variance, root-mean-square value, skewness, kurtosis, peak-to-peak value), in conjunction with three categorical operating conditions: rotational

speed, load, and measurement direction. The objective variable is the gear condition (fault kind or healthy state).

Let $\boldsymbol{s} \in \mathbb{R}$ denote the vector of statistical vibration features, and let v, ℓ, and d represent the speed, load, and direction, respectively. Each unprocessed input sample can thereafter be expressed as $(\boldsymbol{s}, v, \ell, d)$. The complete dataset is initially divided into a training set comprising 80 % of the samples and a testing set consisting of 20 %. The division is executed randomly with a predetermined random seed, ensuring that the partition is precisely reproducible. The testing set is not utilized during training or hyperparameter optimization. It is solely designated for the final performance assessment.

Standardization is exclusively applied to the continuous statistical features $\boldsymbol{s}$, but the one-hot encoded categorical variables (speed, load, and direction) retain their binary format. For each continuous variable s_j, the mean $\mu_{j,\text{train}}$ and standard deviation $\sigma_{j,\text{train}}$ are computed on the training set only. The standardized feature is given by

$$s_{\{n,j\}}^{\{*\}} = \frac{s_{\{n,j\}} - \mu_{j,\text{train}}}{\sigma_{j,\text{train}}}. \tag{2}$$

This transformation is executed on the training samples (fit + transform) and then on the testing samples (transform only, utilizing the same $\mu_{j,\text{train}}$ and $\sigma_{j,\text{train}}$).

Standardization of the vibration characteristics guarantees that all continuous variables contribute on a comparable scale, enhances the convergence of the optimization process and results in more stable and trustworthy classification outcomes. Continuous variables contribute on a similar scale, enhance the convergence of the optimization process and result in more stable and reliable classification outcomes.

2.3. Encoding methods for categorical variables

Each sample's input vector is comprised of continuous statistical features and three categorical operational conditions: rotational speed, load level, and measurement direction. The three variables are classified as nominal categories and encoded by one-hot encoding. The target variable (gear condition) is a category label denoting the sort of fault or the healthy state.

(A) One-Hot

Let v, ℓ, and d denote the categorical variables corresponding to speed, load, and direction, respectively. Each of these variables takes values in a finite set of categories, for example,

$$v \in \left\{c_1^{\{(v)\}}, \ldots, c_{\{K_v\}}^{\{(v)\}}\right\}, \quad \ell \in \left\{c_1^{\{(\ell)\}}, \ldots, c_{\{K_\ell\}}^{\{(\ell)\}}\right\}, \quad d \in \left\{c_1^{\{(d)\}}, \ldots, c_{\{K_d\}}^{\{(d)\}}\right\}. \tag{3}$$

For a generic categorical variable $z \in \{c_1, \ldots, c_K\}$, the one-hot encoded representation is a binary vector

$$e_{(z)} = [e_1, \ldots, e_k], \qquad e_k = \begin{cases} 1 & \text{if } z = c_k, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In the proposed pre-processing pipeline, speed, load, and direction are each transformed into a corresponding one-hot vector and these vectors are concatenated with the continuous statistical features [5].

(B) Ordinal or label encoding

The output variable y denotes the gear condition, indicating either a fault type or a healthy state. It can assume a limited number of values (the various fault categories). For compatibility with the learning method, these classes are substituted with integers (0, 1, 2, ...). This encoding is strictly technical. It solely serves to represent the classes without establishing any hierarchy among them.

3. Experimental setup

3.1. Data collection

Fig. 1 shows the schematic drawing of the test rig used in our experiments for vibration analysis and fault diagnosis in rotating machinery. The setup consists of an electric motor, which drives a shaft supported by bearings and connected to a belt and pulley system for power transmission. An adjustable load is applied to simulate varying operational conditions. A gearbox, enclosed within a dashed box, contains gears that modify speed and torque, making it a critical component for studying gear faults. As shown in Fig. 1, a triaxial piezoelectric accelerometer is mounted directly on the gearbox housing, above the pinion-side rolling bearing, as illustrated in the enlarged view in Fig. 1. The figure also identifies the three measurement directions, horizontal (H), vertical (V), and axial (A), corresponding to the sensor's orthogonal axes. These measurements are essential for detecting potential faults in the system. Fig. 1 illustrates the sensor location and orientation on the gearbox. This test rig is designed to analyze the dynamic behavior of rotating components and diagnose issues such as misalignment, imbalance, bearing defects, and gear failures through vibration signals.

The test rig used for fault diagnosis in bevel gears consists of a gear transmission system mounted on a rigid steel frame to minimize external vibrations. The system includes a three-phase induction motor with a variable speed range of 0–3 000 RPM, controlled via a frequency inverter. The bevel gear assembly consists of an 18-tooth pinion and a 27-tooth gear. Faults were

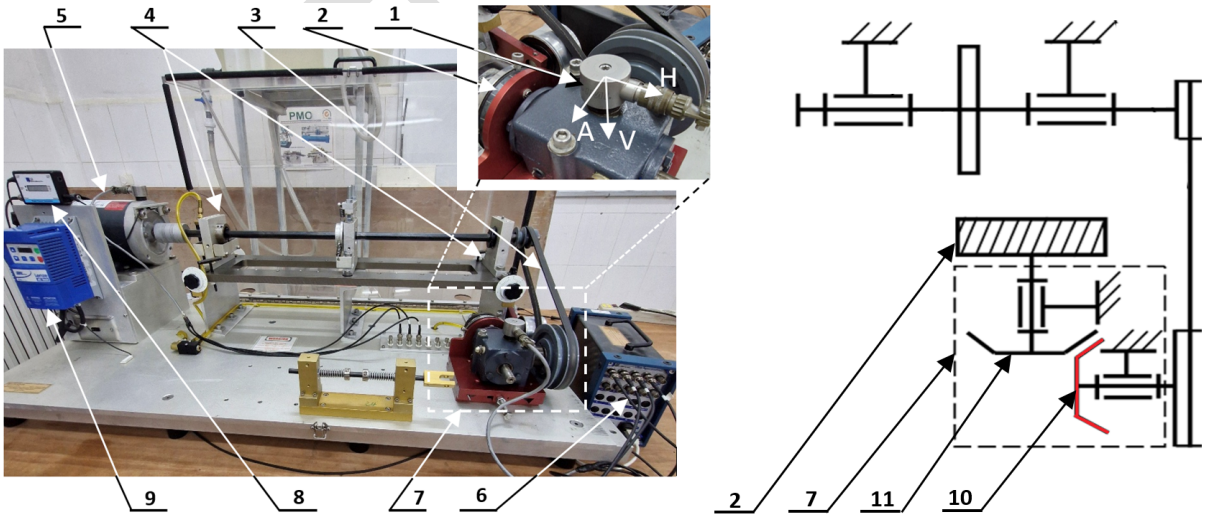
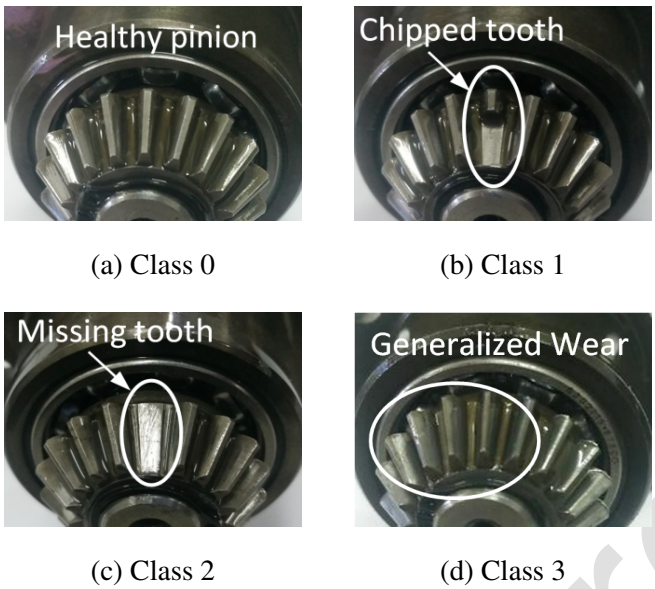


Fig. 1. Vibration test rig and gearbox schematic drawing: (1) piezoelectric accelerometer (triaxial), (2) adjustable load, (3) belt and pulley, (4) rolling bearings, (5) induction motor, (6) data acquisition system, (7) gearbox and gearbox sketch, (8) digital tachometer display, (9) motor speed controller, (10) location of the bevel gears used in the experiments, (11) bevel gear



(a) Class 0 (b) Class 1

(c) Class 2 (d) Class 3

Fig. 2. Different states of the bevel gear

Table 2. Bevel gear parameters

Gear ratio	Pitch gear diameter	Pitch pinion diameter	Pressure angle	Number of teeth (gear)	Number of teeth (pinion)
1.5:1	42.862 5 mm	28.575 mm	20°	27	18

present exclusively on the pinion supplied with the test rig by the manufacturer. The considered fault conditions include a chipped tooth, a missing tooth, and generalized tooth surface wear. The chipped-tooth and missing-tooth faults were manufactured by the test-rig manufacturer through controlled material removal from the pinion tooth, while generalized surface wear was produced by uniformly removing material from the tooth flanks to simulate wear progression. Fig. 2 illustrates the different fault states of the pinion.

The relevant bevel gear parameters are listed in Table 2. A magnetic particle brake provides a controlled load variation between 0–1.13 N m to simulate real working conditions. As indicated in Fig. 1, the triaxial accelerometer is fixed on the pinion-side bearing seat using a stud-mount arrangement to ensure high-frequency signal transmission. The sensor placement was selected to maximize sensitivity to fault-induced vibration components transmitted through the rolling bearings. For vibration measurement, three directions piezoelectric accelerometer with a sensitivity of 100 mV g^{-1} and a frequency range of 0.5–50 kHz is mounted on the pinion bearing and gearbox housing. The signals are acquired using a VibraQuest software data acquisition system. Experiments are conducted at speeds of 600, 900, 1 200, and 1 500 RPM under varying loads (0, 0.37, 0.75, and 1.13 N m).

Table 3 summarizes the experimental design used to investigate the dynamic behavior of the bevel gear under these operating conditions for both healthy and faulty states. Four gear conditions are considered (see Fig. 2): a healthy pinion, a chipped tooth, a missing tooth, and generalized surface wear. Vibration measurements were recorded in three orthogonal directions using a triaxial accelerometer. The collected vibration signals are analyzed to identify fault-specific patterns and to support the development of AI-based condition monitoring strategies for bevel gears.

Based on the three factors listed in Table 3, experiments for each gear condition (healthy, chipped tooth, missing tooth, and generalized wear) were conducted under 16 distinct operating conditions, obtained from the combination of four rotational speeds and four load levels. For each operating condition, vibration data were acquired using a triaxial accelerometer, which simultaneously recorded the horizontal, vertical, and axial components. Each operating condition was repeated 10 times to ensure system stabilization and repeatability, with 8 seconds of vibration data recorded per repetition. As a result, each gear condition yielded 480 vibration signals (4 speeds $\times$ 4 loads $\times$ 3 axes $\times$ 10 repetitions). These signals were processed independently for feature extraction and classification.

This structured methodology enables a systematic evaluation of the gear's dynamic response and facilitates the identification of significant interactions between operational factors and fault conditions. By employing this factorial design, the study improves the robustness of fault characterization and supports the development of AI-based diagnostic strategies for condition monitoring of bevel gears.

3.2. Vibration signal and frequency spectrum characteristics

Fig. 3 illustrates the time-domain vibration signals and the corresponding frequency spectra for the four pinion conditions used in this study. The healthy pinion (Fig. 3a) exhibits a low-amplitude signal resembling random noise, indicating stable operation with minimal mechanical disturbance. In the frequency domain, only small peaks appear at the shaft rotational speed f_n, the gear mesh frequency f_m, and their harmonics, confirming normal behavior.

For the chipped-tooth condition (Fig. 3b), the time-domain signal shows a slightly higher amplitude due to localized impacts, occurring each time the damaged tooth engages. The spectrum reveals amplified components at frequencies related to the fault, including sidebands around the gear mesh frequency, reflecting the irregularity caused by the tooth damage.

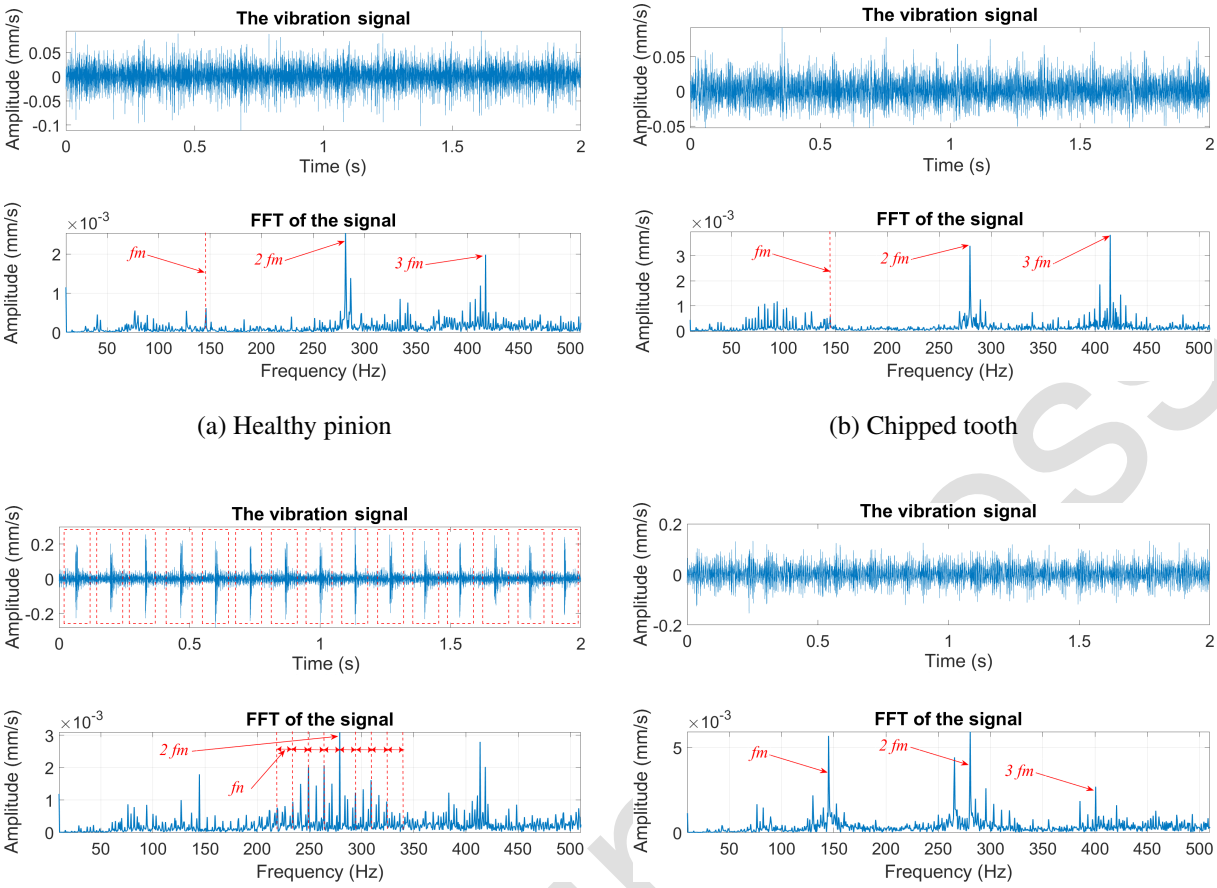
The missing-tooth fault (Fig. 3c) produces the most distinguishable signal pattern, characterized by strong periodic impacts and pronounced amplitude modulation. These impacts arise whenever the void in the gear mesh passes the contact point. Accordingly, the frequency spectrum shows dominant peaks at the gear mesh frequency f_n, its harmonics and modulation sidebands, with significantly higher overall amplitude compared to other conditions.

In contrast, generalized wear (Fig. 3d) leads to increased overall vibration amplitude without clear periodicity. This condition causes distributed surface degradation across multiple teeth. The spectrum displays elevated broadband noise and numerous small peaks, indicating a widespread and nondiscrete fault mechanism rather than localized impacts.

Together, all the figures demonstrate that each fault type produces a distinct vibration signature. Clear periodic impacts characterize the missing-tooth condition, moderate localized disturbances identify the chipped tooth, and broadband spectral changes reflect generalized wear, whereas the healthy gear remains dominated by low-amplitude components at f_n and f_m. These

Table 3. Measuring conditions and their levels

Factor No.	Parameters	Level 1	Level 2	Level 3	Level 4
1	speed [RPM]	600	900	1 200	1 500
2	load [N m]	0	0.37	0.75	1.13
3	direction	horizontal	vertical	axial	–



(a) Healthy pinion

(b) Chipped tooth

(c) Missing tooth

(d) Generalized wear

Fig. 3. Vibration analysis of different pinion conditions

differences form the basis for extracting discriminative time-domain features used in subsequent machine learning classification.

4. Fault diagnosis using classification models

4.1. SVM-based fault classification for bevel gears

SVM is a powerful machine learning technique used for classifying faults in bevel gears based on vibration analysis [1]. In this study, SVM is applied to classify four conditions: healthy pinion, chipped tooth, missing tooth, and generalized wear. The classification process follows three main steps: (A) vibration signals are acquired from the MFS Magnum simulator, (B) time-domain statistical features are extracted from each signal, and (C) these features are used as input to machine learning models for fault classification.

The classification models use a set of statistical time-domain features calculated from each vibration signal. These include maximum value, RMS, variance, mean square value, kurtosis factor, and others, as listed in Table 1. Each sample is therefore represented by a multi-dimensional feature vector, not a single encoded value.

Given a dataset $D = \{x_i, y_i\}_{i=1}^{N}$, where $x_i \in \mathbb{R}^d$ represents the feature vector extracted from vibration signals, and $y_i \in \{0, 1, 2, 3\}$ denotes the fault class, the objective of SVM is to find an optimal hyperplane that separates these classes while maximizing the margin. Here, the following four physical gear conditions are considered (Fig. 2): Class 0 – healthy gear, Class 1

– chipped tooth fault, Class 2 – missing tooth fault, and Class 3 – generalized wear fault. For a binary classification problem, the decision function is

$$f(x) = \mathrm{sgn}(wx + b), \tag{5}$$

where w is the weight vector, and b is the bias term. However, since this is a multi-class classification problem, the one-vs-one (OvO) or one-vs-all (OvA) strategy is employed, where multiple binary classifiers are trained and combined.

To handle nonlinear relationships in vibration features, the radial basis function (RBF) kernel is commonly used

$$K\left(x_i, x_j\right) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right), \tag{6}$$

where σ is a kernel parameter controlling the influence of training samples.

4.2. Linear logistic regression

LR estimates the probability that a given sample x belongs to a specific class. It predicts the class of a binary classification sample based on the computed probability of the positive class [22]

$$P\left(Y = 1 \middle| x\right). \tag{7}$$

One can specify a threshold to obtain a discrete value instead of probability. Let us consider categorizing samples with $P\left(Y = 1 \middle| x\right)$ as positive using a threshold of 0.5. When misidentifying a sample as positive is costly, a conservative method may utilize 0.8.

To demonstrate the link between input x and probability $P(Y = 1 - x)$, a standard approach is to employ an affine translation of the input data, followed by the logistic or sigmoid function [22]

$$a = w^t x, \tag{8}$$

where w represents the weight vector (model parameters), x is the input feature vector, and $w^t x$ is the linear combination of input features. The transformed value a is then passed through the logistic (sigmoid) function to map it to a probability value as

$$P\left(Y = 1 \middle| x\right) = \frac{e^a}{1 + e^a}. \tag{9}$$

Since the sigmoid function returns values in the range (0,1), it is ideal for representing probabilities. A simplified notation for the sigmoid function is

$$P\left(Y = 1 \middle| x\right) = \mathrm{sigmoid}(a) = \widehat{y}, \tag{10}$$

where $\widehat{y}$ represents the predicted probability. This probability can then be compared against a chosen threshold to make a final classification decision.

In this study, LR is implemented using the multinomial (softmax) formulation, which allows the model to assign probabilities to more than two classes. The model uses a one-vs-rest (OvR) or softmax multiclass strategy depending on the solver. This enables LR to classify samples into the four fault categories.

4.3. Grid search cross-validation

Grid search cross-validation (GSCV) is a systematic approach for improving the performance of machine learning models by fine-tuning their hyperparameters. The hyperparameters are predefined settings that influence the learning process, such as the regularization strength in LR or the penalty and kernel parameters in SVM. Selecting appropriate hyperparameter values is essential for improving model accuracy, stability and generalization.

In GSCV, model performance is optimized by exhaustively searching over a predefined grid of hyperparameter values. For each hyperparameter combination, the model is trained and evaluated using k-fold cross-validation. The dataset D of size N is divided into k equal-sized subsets $\{D_1, D_2, \ldots, D_k\}$. For each fold i, the model is trained on D/D_i and validated on D_i. The validation performance metric M_i is computed for each fold and averaged to obtain the final cross-validation score

$$M_{cv} = \frac{1}{k} \sum_{i=1}^{k} M_i. \tag{11}$$

In this study, GSCV was applied to both SVM and LR classifiers. For SVM, the regularization parameter C, kernel type (linear and radial basis functions), and kernel coefficient γ (for the RBF kernel) were tuned. The parameter C was varied over a logarithmic range, while γ was tested over multiple orders of magnitude to capture both smooth and highly nonlinear decision boundaries. For LR, the regularization strength C and the solver algorithm were tuned. Different values of C were evaluated to control the degree of regularization, and commonly used solvers were considered to ensure convergence and numerical stability.

The hyperparameter combination yielding the highest average cross-validation performance was selected as the optimal configuration. Although GSCV is computationally demanding, it provides a robust and reproducible framework for hyperparameter optimization and was therefore adopted in this work to enhance the performance of the SVM and LR models.

5. Results and analysis

5.1. Impact of GSCV on SVM and LR performance

Figs. 4 and 5 present the confusion matrices of the SVM and LR classifiers before and after GSCV, expressed in absolute sample counts. Each matrix is based on 4 320 feature samples per gear condition, obtained by extracting nine temporal features from 480 vibration signals per class.

Prior to hyperparameter optimization, both classifiers achieve perfect recognition of Class 1, with all 4 320 samples correctly classified. However, notable misclassification is observed for the remaining classes. In the SVM model, Class 3 is strongly confused with Class 0, with 2 880 samples incorrectly predicted as Class 0 and only 720 correctly identified. Similarly, Class 2 exhibits substantial overlap with Class 0, with 2 400 misclassified samples.

The LR classifier shows improved discrimination for Class 3, correctly classifying 3 240 samples, but suffers from increased confusion in Class 2, where 1 440 samples are incorrectly assigned to Class 0. These results indicate that, before optimization, LR performs better for Class 3, while SVM shows relatively stronger performance for Class 2, albeit with considerable inter-class confusion.

After applying GSCV, both models SVM-GSCV and LR-GSCV show noticeable improvements in classification performance. The optimized SVM achieves perfect classification for

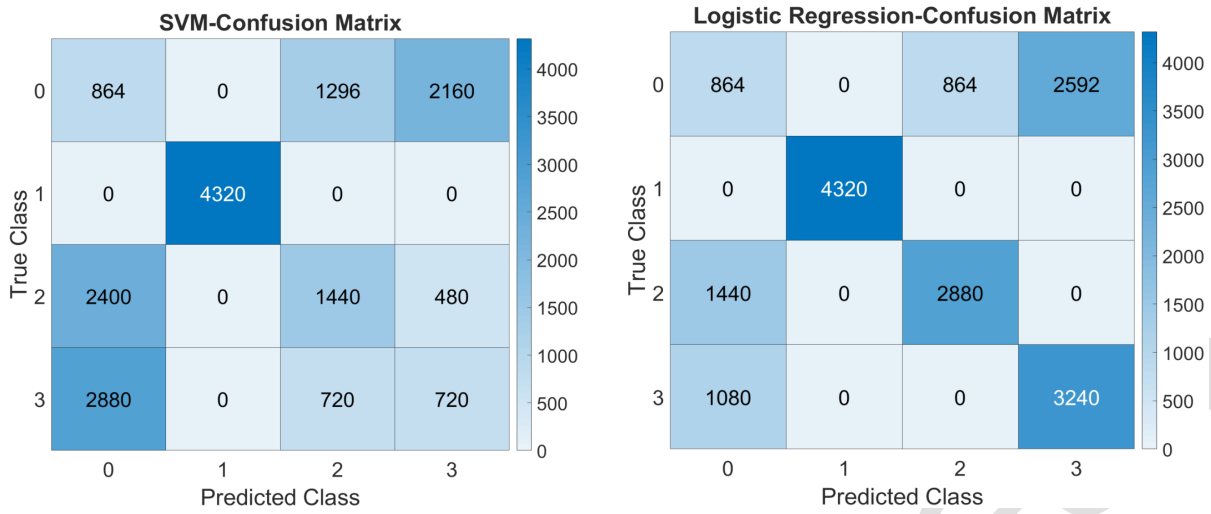


Fig. 4. Performance of SVM and LR before tuning

Class 1 and Class 2, while substantially improving the recognition of Class 0, with 3 888 correctly classified samples. Although minor confusion remains for Class 3, 3 600 samples are correctly identified, representing a marked improvement compared to the non-optimized model.

Similarly, the optimized LR model maintains perfect classification of Class 1 and achieves near-perfect performance for Class 2, with 3 840 correctly classified samples. While Class 0 remains the most challenging for LR, with some misclassification toward Classes 2 and 3, the overall error rate is considerably reduced compared to the pre-GSCV configuration.

5.2. Predicted and true labels across all models

Fig. 6 illustrates the correspondence between predicted and true class labels at the individual sample level across the test set. This visualization highlights the dispersion and recurrence of misclassifications along the sample index, thereby complementing the aggregate performance information provided by the confusion matrices.

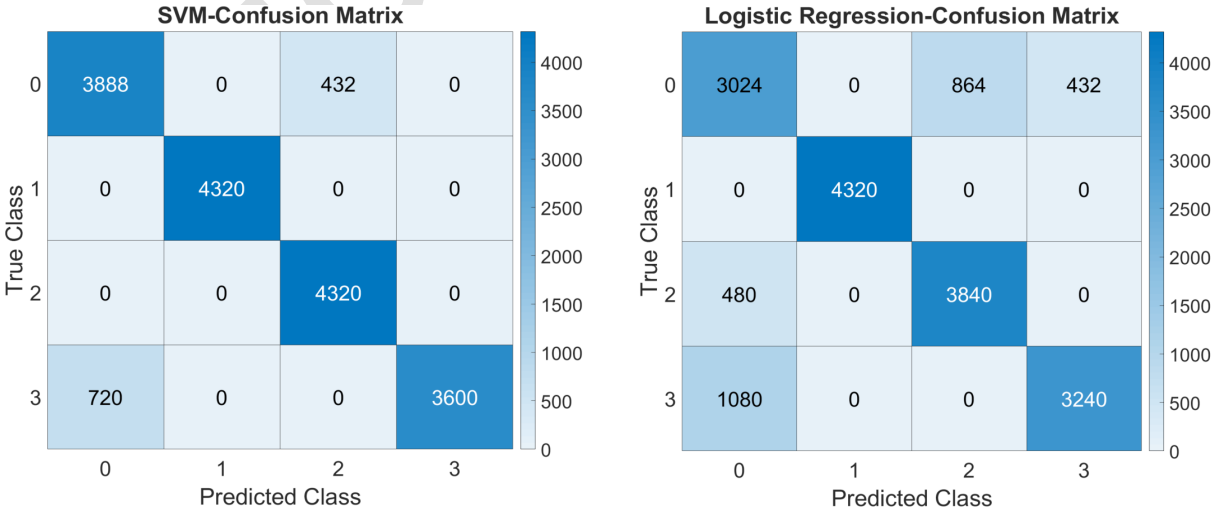


Fig. 5. Performance of SVM-GSCV and LR-GSCV

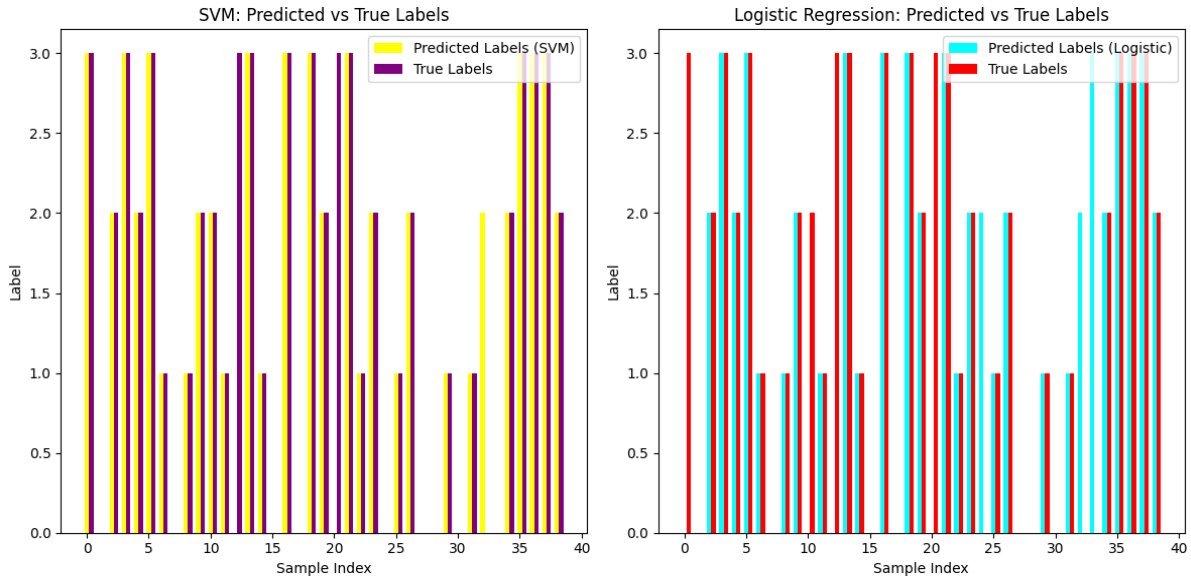


Fig. 6. Predicted vs true labels for the SVM and LR models before and after GSCV tuning

In Fig. 6, the comparison of predicted vs actual labels for SVM-GSCV and LR-GSCV models reveals differences in their classification performance. In the SVM-GSCV plot on the left, the predicted labels (in yellow) align more closely with the true labels (in purple), suggesting better accuracy. In contrast, the LR-GSCV plot on the right shows more noticeable deviations, with the predicted labels (in cyan) frequently misaligned with the true labels (in red). Misclassifications are distributed across the sample index range in both models, but they appear more frequent in LR-GSCV. This indicates that SVM-GSCV may provide more reliable predictions compared to LR-GSCV for this classification task.

5.3. Quantitative metrics

Performance metric values and model parameters play a crucial role in evaluating and optimizing the effectiveness of a classification model, ensuring its reliability, accuracy, and generalization to real-world scenarios.

Accuracy is calculated as the sum of true positives and true negatives divided by the total number of predictions. Precision measures the proportion of correctly predicted positive instances out of all predicted positives, computed as true positives divided by the sum of true positives and false positives. Recall, also known as sensitivity, represents the ability of the model to correctly identify positive instances and is given by true positives divided by the sum of true positives and false negatives. The F1-Score provides a balanced measure of precision and recall, calculated as twice the product of precision and recall divided by their sum [19].

However, as shown in Fig. 7, LR achieves a global accuracy of 64.10 %, compared to SVM's 38.46 %. Overall, both models demonstrate relatively low accuracy, with LR performing slightly better than SVM. Both models have difficulty distinguishing between certain classes, suggesting that additional feature engineering, hyperparameter tuning, or hybrid classification techniques may improve performances.

The quantitative comparison in Fig. 8 evaluates the performance of SVM-GSCV and LR-GSCV across four key metrics (accuracy, precision, recall, and F1-Score) for different classes. The accuracy results indicate that SVM consistently outperforms LR, achieving higher correct

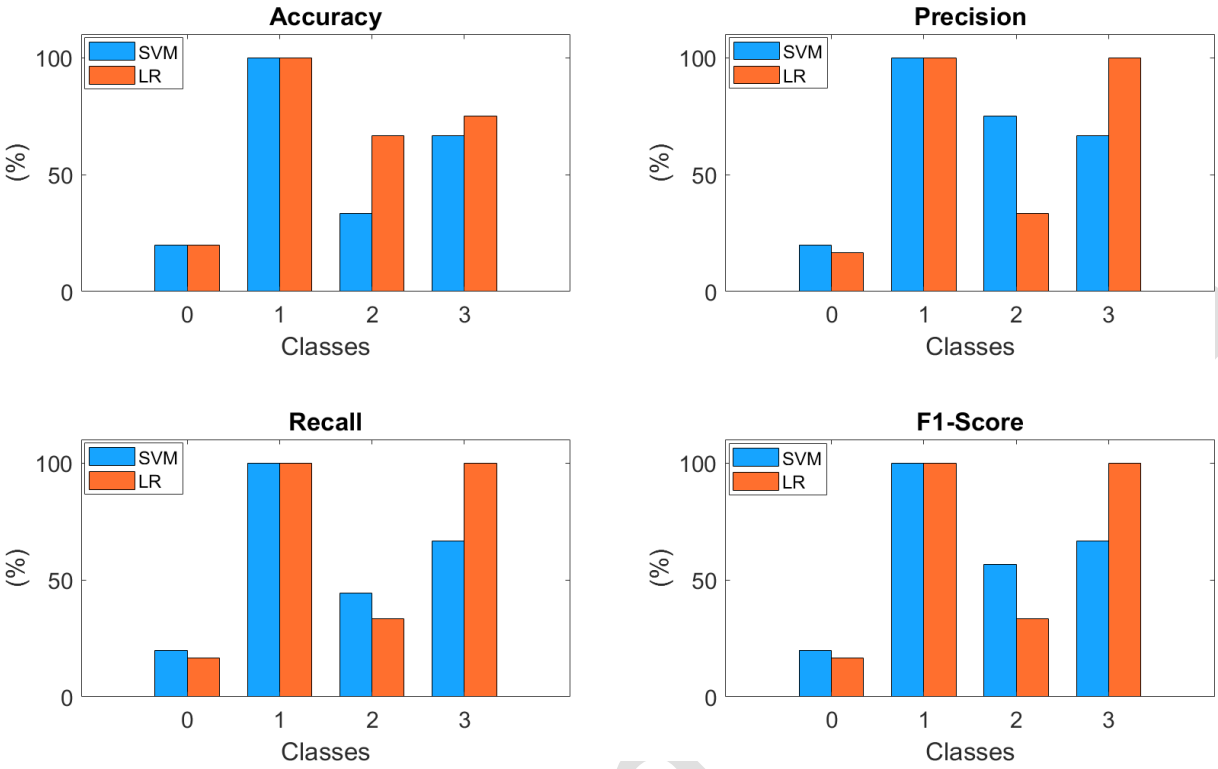


Fig. 7. Comparison of SVM and LR metrics by classes before tuning

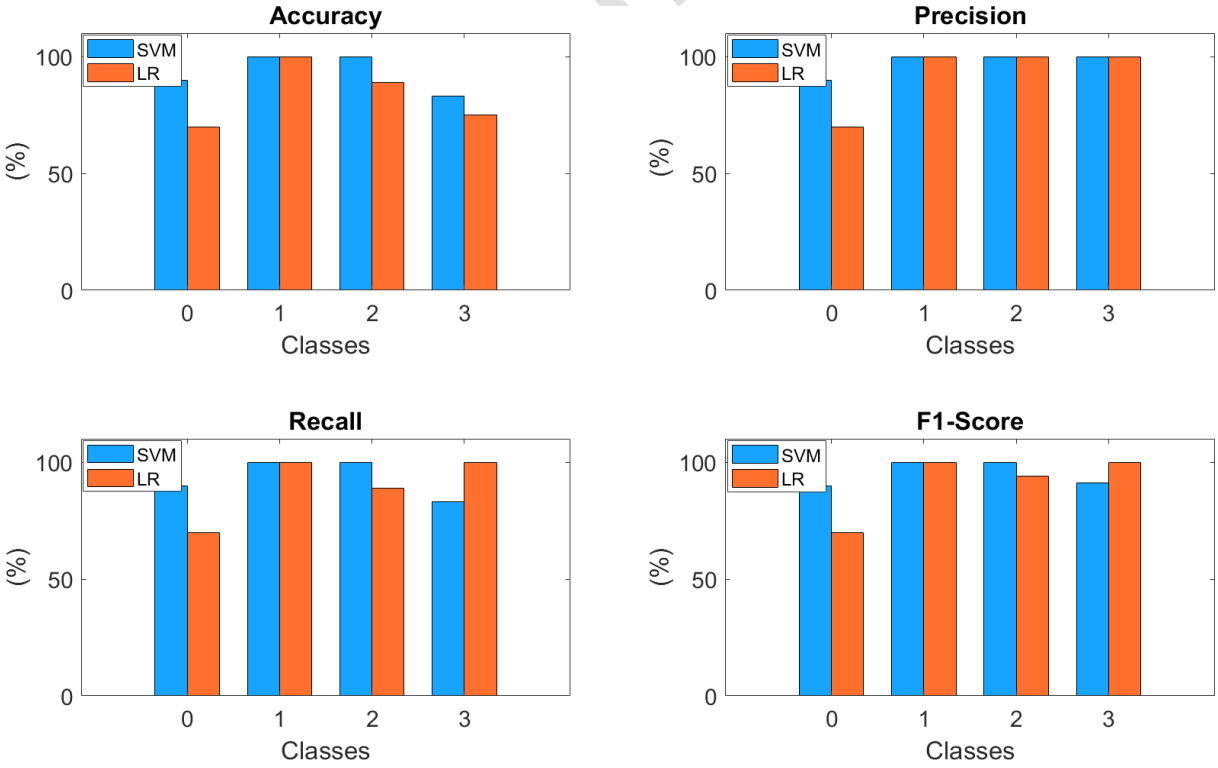


Fig. 8. Comparison of SVM-GSCV and LR-GSCV metrics by classes

classifications across most classes. Precision values for both models are relatively similar, suggesting that false positive rates are comparable. However, SVM demonstrates better recall, meaning it captures more actual positive instances, whereas LR tends to miss more true positives. As a result, the F1-Score, which balances precision and recall, also favors SVM in most cases, reinforcing its overall robustness. These results confirm that SVM-GSCV provides better classification performance than LR-GSCV, particularly in terms of accuracy and recall.

Following improvements, the SVM-GSCV model exhibited a significant enhancement in performance, achieving a global accuracy of 92.31 %, exceeding the 82.05% accuracy of the LR-GSCV model. Overall, the tuning process has significantly enhanced both models, making SVM-GSCV the preferable choice due to its more balanced and higher accuracy across all classes.

6. Discussion

The results show that before hyperparameter tuning, both SVM and LR struggle to reliably separate some fault classes. This difficulty arises from overlapping patterns in the extracted time-domain features, which limits the discriminative power of the initial models. LR performs better overall at this stage, particularly for Class 3, while SVM shows a relative advantage in identifying Class 2. These differences indicate that each model emphasizes distinct properties of the feature space.

After applying GSCV, SVM exhibits substantial improvements across nearly all fault categories. This confirms that SVM is highly sensitive to hyperparameter selection and benefits strongly from optimized kernel and regularization settings. The increases observed in accuracy, precision, recall, and F1-Score demonstrate that the tuned SVM model captures fault-related patterns more effectively than the LR model. Although LR also improves after tuning, its performance remains less stable, especially for Class 0.

These observations are consistent with findings in earlier studies, which reported that the SVM performance significantly increases when appropriate kernels and parameters are selected, whereas LR, being a linear classifier, faces inherent limitations in capturing nonlinear class boundaries.

7. Conclusions

In this study, a comparison of the effectiveness of SVM and LR in the detection and classification of defects in bevel gears was conducted. The results showed that LR outperforms SVM in terms of accuracy, although it remains relatively low. However, after applying hyperparameter tuning with GSCV, both models demonstrate a significant improvement, with SVM achieving the highest overall performance. In terms of class-wise performance, SVM generally achieves higher precision and F1-Scores across all bevel gear faults, while LR demonstrates strong recall for the missing-tooth failure mode following improvement.

Overall, SVM proves to be the superior model across most evaluation metrics, particularly after tuning. Thus, the SVM-GSCV model demonstrates a notable improvement in performance, attaining a global accuracy of 92.31 %, surpassing the 82.05% accuracy of the LR-GSCV model. The tuning process substantially improves both models, making SVM-GSCV the more robust choice due to its better balance and higher accuracy across all classes.

The study presented some promising results; however, the experiments were performed on a controlled laboratory test bench utilizing a single bevel gear configuration, and the overall

dataset size is still constrained. Additionally, the analysis was based only on time-domain statistical features, which might not be able to fully capture the vibration patterns that are linked to different fault mechanisms. Future research may mitigate these limitations by augmenting the dataset to encompass a wider array of operating conditions and gear types, in addition to integrating frequency- and time-frequency-domain features. Furthermore, the implementation of real-time monitoring and the evaluation of model robustness across various operational conditions would augment the applicability of the proposed methodology in industrial contexts.

Acknowledgement

The authors would like to thank the Directorate General for Scientific Research and Technological Development (DGRSDT) for financial support (PRFU code: A01L09UN410120230002).

References

[1] Abdul, Z. K., Al-Talabani, A. K., Highly accurate gear fault diagnosis based on support vector machine, Journal of Vibration Engineering & Technologies 11 (7) (2023) 3 565–3 577. https://doi.org/10.1007/s42417-022-00768-6

[2] Bhandare, R. V., Phalle, V. M., Handikherkar, V. C., Patil, S. S., Weighted aggregation approach and tree-based ensemble technique for anomaly detection and fault severity analysis in gear fault diagnosis, Journal of Failure Analysis and Prevention 25 (2025) 864–876. https://doi.org/10.1007/s11668-025-02148-0

[3] Cen, J., Yang, Z., Liu, X., Xiong, J., Chen, H., A review of data-driven machinery fault diagnosis using machine learning algorithms, Journal of Vibration Engineering & Technologies 10 (7) (2022) 2 481–2 507. https://doi.org/10.1007/s42417-022-00498-9

[4] Chen, S., Yang, R., Zhong, M., Graph-based semi-supervised random forest for rotating machinery gearbox fault diagnosis, Control Engineering Practice 117 (2021) No. 104952. https://doi.org/10.1016/j.conengprac.2021.104952

[5] Dahouda, M. K., Joe, I., A deep-learned embedding technique for categorical features encoding, IEEE Access 9 (2021) 114 381–114 391. https://doi.org/10.1109/ACCESS.2021.3104357

[6] Dhall, D., Kaur, R., Juneja, M., Machine learning: A review of the algorithms and its applications, Proceedings of ICRIC 2019: Recent Innovations in Computing, 2020, pp. 47–63. https://doi.org/10.1007/978-3-030-29407-6_5

[7] Dubaish, A. A., Jaber, A. A., Comparative analysis of SVM and ANN for machine condition monitoring and fault diagnosis in gearboxes, Mathematical Modelling of Engineering Problems 11 (4) (2024) 976–986. https://doi.org/10.18280/mmep.110414

[8] Hammood, A. S., Taki, A. G., Ibrahim, N. S., Mohammed, J. G., Jasim, R. K., Jasim, O. M., Optimizing failure diagnosis in helical gear transmissions with stochastic gradient descent logistic regression using vibration signal analysis for timely detection, Journal of Failure Analysis and Prevention 24 (1) (2024) 71–82. https://doi.org/10.1007/s11668-023-01814-5

[9] Jain, P. H., Bhosle, S. P., Analysis of vibration signals caused by ball bearing defects using time-domain statistical indicators, International Journal of Advanced Technology and Engineering Exploration 9 (90) (2022) 700–715. http://doi.org/10.19101/IJATEE.2021.875416

[10] Jing, Y., Su, H., Wang, S., Gui, W., Guo, Q., Fault diagnosis of electric impact drills based on timevarying loudness and logistic regression, Shock and Vibration 2021 (1) (2021) No. 6655090. https://doi.org/10.1155/2021/6655090

[11] Kafeel, A., Aziz, S., Awais, M., Khan, M. A., Afaq, K., Idris, S. A., Alshazly H., Mostafa, S. M., An expert system for rotating machine fault detection using vibration signal analysis, Sensors 21 (22) (2021) No. 7587. https://doi.org/10.3390/s21227587

[12] Khoualdia, T., Lakehal, A., Chelli, Z., Practical investigation on bearing fault diagnosis using massive vibration data and artificial neural network, Proceedings of the 3rd International Conference on Big Data and Networks Technologies (BDNT 2019), Springer, 2019, pp. 110–116. https://doi.org/10.1007/978-3-030-23672-4_9

[13] Khoualdia, T., Lakehal, A., Chelli, Z., Khoualdia, K., Nessaib, K., Optimized multi-layer perceptron artificial neural network based fault diagnosis of induction motor using vibration signals, Diagnostyka 22 (1) (2021) 65–74. https://doi.org/10.29354/diag/133091

[14] Kumar, V., Mukherjee, S., Verma, A. K., Sarangi, S., An AI-based nonparametric filter approach for gearbox fault diagnosis, IEEE Transactions on Instrumentation and Measurement 71 (2022) 1–11. https://doi.org/10.1109/TIM.2022.3186700

[15] Lu, D., Qiao, W., A GA-SVM hybrid classifier for multiclass fault identification of drivetrain gearboxes, Proceedings of 2014 IEEE Energy Conversion Congress and Exposition (ECCE), 2014, pp. 3 894–3 900. https://doi.org/10.1109/ECCE.2014.6953930

[16] Men, Z., Hu, C., Li, Y.-H., Bai, X., A hybrid intelligent gearbox fault diagnosis method based on EWCEEMD and whale optimization algorithm-optimized SVM, International Journal of Structural Integrity 14 (2) (2023) 322–336. https://doi.org/10.1108/IJSI-12-2022-0145

[17] Miller, A., Panneerselvam, J., Liu, L., A review of regression and classification techniques for analysis of common and rare variants and gene-environmental factors, Neurocomputing 489 (2022) 466–485. https://doi.org/10.1016/j.neucom.2021.08.150

[18] Nguyen, V.-T., Diep, Q. B., Vibration-based gearbox fault diagnosis using a multi-scale convolutional neural network with depth-wise features concatenation, PLOS One 20 (7) (2025) No. e0324905. https://doi.org/10.1371/journal.pone.0324905

[19] Priyadharshini, N., Selvanathan, N., Hemalatha, B., Sureshkumar, C., A novel hybrid extreme learning machine and teaching-learning-based optimization algorithm for skin cancer detection, Healthcare Analytics 3 (2023) No. 100161. https://doi.org/10.1016/j.health.2023.100161

[20] Rayjade, G., Bhagure, A., Kushare, P. B., Bhandare, R., Matsagar, V., Chaudhari, A., Performance evaluation of machine learning algorithms and impact of activation functions in artificial neural network classifier for bearing fault diagnosis, Journal of Vibration and Control 31 (9–10) (2025) 1 859–1 873. https://doi.org/10.1177/10775463241235778

[21] Roy, A., Chakraborty, S., Support vector machine in structural reliability analysis: A review, Reliability Engineering & System Safety 233 (2023) No. 109126. https://doi.org/10.1016/j.ress.2023.109126.

[22] Seger, C., An investigation of categorical variable encoding techniques in machine learning: Binary versus one-hot and feature hashing, Bachelor thesis, KTH Institute of Technology, Stockholm, Sweden, 2018.

[23] Sikora, M. V., Comparison of machine learning binary classifiers for detection of gear defects, Master thesis, Federal University of Santa Catarina, Florianópolis, 2024.

[24] Su, Y., Meng, L., Kong, X., Xu, T., Lan, X., Li, Y., Small sample fault diagnosis method for wind turbine gearbox based on optimized generative adversarial networks, Engineering Failure Analysis 140 (2022) No. 106573. https://doi.org/10.1016/j.engfailanal.2022.106573

[25] Xiao, X., Li, C., He, H., Huang, J., Yu, T., Rotating machinery fault diagnosis method based on multi-level fusion framework of multi-sensor information, Information Fusion 113 (2025) No. 102621. https://doi.org/10.1016/j.inffus.2024.102621

[26] Xuan, H. U., Chun, L. I., Kehua, Y. E., Application of GWO-SVM in wind turbine gearbox fault diagnosis, Journal of Mechanical Strength 43 (5) (2022) 1 026–1 034. https://doi.org/10.16579/j.issn.1001.9669.2021.05.002

[27] Zhang, H., Guo, X., Zhang, P., Improved PSO-SVM-based fault diagnosis algorithm for wind power converter, IEEE Transactions on Industry Applications 60 (2) (2023) 3 492–3 501. https://doi.org/10.1109/TIA.2023.3341059