SvedbergOpen
DISSEMINATIONOFKNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Extensive Feature Extraction Approach For Fault Diagnosis In Predictive Maintenance Enabled By Machine Learning

Sreyas Jose[1]*, Dr. Salaja Silas[2], Dr. Elijah Blessing Rajsingh[3]

[1] Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore – 641 114, Tamil Nadu, India
ORCID: https://orcid.org/0000-0001-9851-6205[2], https://orcid.org/0000-0003-3424-812X[3]
*Corresponding author: Sreyas Jose, Email: sreyasjose@karunya.edu.in, ORCID: https://orcid.org/0000-0002-0511-0164

Abstract
With the advent of autonomous vehicles, there is a growing need for reliable and efficient classifiers in complex vehicular environments to detect faults. The diagnosis of faults is a key step to detect operating conditions that are abnormal early and in a way that will help in effective predictive maintenance. Machine-learning-based approaches offer potential solutions for fault diagnosis of vehicles based on the data collected from the vehicles, and with the development of wireless and intelligent vehicular technologies, it is now possible to use this observed data to complete this diagnosis. Four feature extraction and dimensionality-reduction methods, Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), Non-Negative Matrix Factorization (NNMF), and Linear Discriminant Analysis (LDA), are tested and compared with a Support Vector Machine (SVM) for vehicle fault diagnosis. The extracted feature representations are fed to the SVM classifier, which classifies the vehicle conditions being observed. The results of the experiments reveal that the accuracy of PCA-SVM, t-SNE-SVM and NNMF-SVM is 99.14%, while that of LDA-SVM is 88.03%. In the set of the evaluated methods, t-SNE-SVM has the highest recall on 79.05% and F-measure of 79.02%, keeping an accuracy of 80% as PCA-SVM and NNMF-SVM methods. The results show the effect of feature representation on the detection of faults and that the representation t-SNE gives a relatively good representation of the data under evaluation. The proposed method can be used for automating vehicle fault diagnosis and be used as a framework for a predictive-maintenance vehicle application for the observed operating conditions of the vehicle.

Keywords: Machine Learning Classifiers, Principle Component Analysis, t-distributed Stochastic Neighbor Embedding (t-SNE), Linear Discriminant Analysis (LDA), Non-Negative Matrix Factorization (NNMF), SVM Classifier.

1. Introduction

The advancements and growing amount of sensors and the rudimentary data fusion techniques are considered significant factors in order to analyse software faults and hardware risk increases in terms of processing failures, actuators malfunctions, and sensor failures. Therefore, these issues are overcome by detecting various faults using control strategy methods to ensure certain factors, such as reliable performance results and autonomous systems (Mohamed, A et al., 2018). Most of the existing research on Machine Learning (ML) and statistical models mostly depends on the data that are accumulated from multiple sensors. The safety of autonomous vehicles is basically dependent on the reliance, sensor faults, and the resulting error propagation in order to ensure their safety and reliability of it (Zholghadri et al., 2014). A fault is considered one of the significant terms in autonomous vehicles that classify and detects the abnormal state of the vehicles. It is investigated by the properties of the vehicles or the parameters of the systems. The fault system can be divided into sensor faults, actuator faults or any process faults. Sensor faults occur in the input modules and the actuator faults occur in the output modules (Wang et al., 2019). In order to remove the dangers and risks, the heavy vehicles are manufactured with the fault diagnosis units. The intention of the fault diagnosis systems is to detect, isolate and recognize the faulty sensors with the prediction of enhancing the reliability of sensors and also for making wiser decisions based on the prediction. In order to remove the potential threats, the vehicles are constructed with the electronic architecture that intelligently predicts the fault present in the system, fault isolation that is used to find which sensor is faulty, fault identification that determines how the sensor has failed (Chen et al., 2018).

The rest of the sections are as follows: Section 2 provides the existing techniques of the research that include various ML classifiers and other feature extraction techniques in the autonomous vehicle system. Section 3 provides a novel technique that discusses the design and functioning of the new proposed classifiers. Section 4 presents the Experimental Results and Discussion that portrays the achieved results of the proposed system. Section 5 discusses the implications of the study for research, predictive maintenance, and vehicle fault diagnosis. Finally, section 6 represents the proposed research's conclusion that indicates the study's findings.

2. Related Work

In this related work section, we discuss the various Machine Learning techniques used in autonomous vehicle systems, especially in fault diagnosis.

Faults that are identified in-vehicle sensors system have the capability to propagate the control system based on unchecked throughout. In case of undetected faults, automatic systems and health monitoring techniques are considered a vital part of automotive applications. Therefore, traditional fault diagnosis systems exhibit an effectual output, especially in complex systems that include autonomous vehicles. To overcome this kind of issue, a fault detection method has been used (LiamBiddle et al., 2021) with appropriate identification and isolation for multi-sensor systems with real-time experimentation. Additionally, SVM methods are utilized to identify the faults found in sensors. Further, DL-based methods are utilized by (Naeem Ayoub et al., 2020) for identifying faults, especially in power-line components. Therefore, a YOLOv4-tiny based DNN - Deep Neural Networks for identifying high accuracy components. To execute these DL-based techniques in a real-time scenario, various single-board devices, including Nvidia, TX2, Raspberry Pi4, and much more.

Moreover, functional safety plays an essential role in the automotive industry with minimum costs has turned into a significant drawback. A PdM - predictive maintenance has been utilized in (Andreas Theissler et al., 2021) with operating data. Consequently, the existing research was analyzed from the ML perspective. The open challenges are identified and recognized based on the research activities. The available data which was publicly available would enhance the research activities. Further, most of the research concentrates only on supervised techniques needed for labeled data. Next, integrating multiple-data sources have the possibility to enhance the preciseness. And the utilization of DL-based methods will enhance the outcome; however, improvement was needed with the interpretable techniques with the availability of labeled data. The autonomous robotic manipulation systems are implemented in (Eduardo Godinho Ribeiro et al., 2021) with the capability to generalize to various objects with the high-processing speed that enables its application in real-time robotic systems scenarios.

Furthermore, in the current environment, ML has been utilized in the system based on several conditions, including human performance. Most of the systems need safety assurance in the automotive system by developing appropriate ML-classifiers to diagnose and rectify faults. Hence, AMLAS - Assurance of ML for utilization in Autonomous Systems comprises a certain case pattern exhibited by associating safety assurance while developing the ML components (N. Adbennour et al., 2020). Most of the implementations resulted in certain limitations on diagnosing the fault in a real-time environment. To increase the accuracy in identifying the fault in the autonomous vehicle system, we have utilized ML classifiers in the research.

3. Proposed Methodology

The proposed methodology aims to diagnose vehicle faults through a machine-learning-based framework, which consists of data preprocessing, feature extraction/reduction and supervised classification. Five key stages are used in the overall framework: (i) data collection, (ii) data pre-processing and sampling, (iii) feature extraction and dimensionality reduction, (iv) the implementation of a SVM-based fault classification system, and (v) performance evaluation. The aim of the proposed framework is to find the most informative display of vehicle fault data and then to classify the operating state of the vehicle into the corresponding fault categories. The overall architecture of the proposed model is shown in Figure 1.

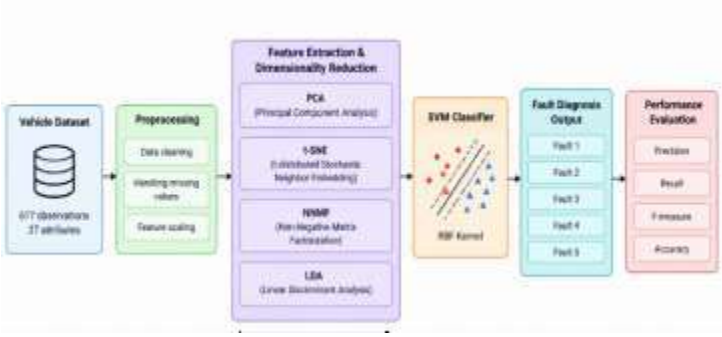

Fig 1. Overall Architectural Diagram

3.1 Dataset Collection and Description

The vehicle fault data for this study was downloaded from the open-source repository. The dataset contains data collected from various operating and fault conditions of the vehicles. The data for each observation is composed of a number of attributes that describe the vehicle condition in the observation and the target variable defines the fault class.

The dataset was first checked for missing data, duplicate observations, inconsistent records, and class imbalance before applying the machine-learning algorithms. Noise and redundancy in the attributes are significant in the pre-processing stage as they may affect the learning process adversely and may lead to poor classification results.

The data was split into attributes and classes. Have the original data set represented as in Eq. (1).

$$D = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

Here, the term N denotes the number of observations, $x_i \in \mathbb{R}^d$ denotes the d-dimensional feature vector of the ith observation, and y_i denotes the fault class of the observation.

3.2 Data Pre-processing and Sampling

Data pre-processing was done prior to feature extraction and classification. The pre-processing stage was developed to enhance the quality of the data, minimize redundancy, and ensure that the data fed into the machine-learning algorithms is presented in a consistent format.

3.2.1 Data Cleaning

The data collected was first checked for missing, duplicate, irrelevant and inconsistent data. Handling of missing or invalid observations was done based on the characteristics of the data sets. In cases where it was necessary to eliminate observations due to duplication of effort, they were removed to avoid giving too much weight to repeated observations.

3.2.2 Class Distribution and Sampling

Fault data sets might have unequal numbers of observations in the different fault categories. Such class imbalances may lead to a bias of the classifier towards the majority class. A sampling procedure was thus adopted to get a representative training dataset therefore.

Suppose that the dataset$D$ is given as$N$ to denote the observations. A sampleD_s of a dataset is collected as in Eq. (2).

$$D_s \subseteq D, \quad | D_s | < | D | \quad (2)$$

Here, $D_s$is the sample that is obtained which is used for the next processing. The sampling process helps to decrease the number of observations to be processed when training the model, while keeping the representation of the relevant fault categories. The sampling method was chosen based on the distribution of the data set in the classes.

3.2.3 Feature Scaling

The features were scaled prior to classification, due to different ranges of numbers in the extracted features. Assume the original value of the feature is x, with minimum and maximum being$x_{\min}$ and $x_{\max}$. Min-max normalization is given as in Eq. (3).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

Here, the term x' is the normalized value of the feature. For SVM, the scale of the individual attribute will be important, so feature scaling is especially crucial.

3.3 Feature Extraction and Dimensionality Reduction

The feature extraction and dimensionality reduction process is an important part of the proposed fault diagnosis process. The number of attributes present in the collected vehicle fault data set can be quite high, there may be redundant information and correlated features, which can lead to high computational complexity in the classification process and may have a negative effect on the ability of the classifier to learn the classification. To this end, suitable feature transformation methods are used to get a compact and informative representation of the original vehicle data. In the present study, four methods, namely, the Principal Component Analysis (PCA), the Linear Discriminant Analysis (LDA), the Non-Negative Matrix Factorization (NNMF) and the t-distributed Stochastic Neighbor Embedding (t-SNE) are studied. These techniques are chosen because they offer various methods for representation of high dimensional data. PCA determines the directions of maximum variation, LDA maximizes the separation between the predefined fault classes, NNMF latent representation of the input data is kept non-negative, and t-SNE captures the nonlinear local relationships between the data observations. These techniques are then applied to transform the representations, which are fed into the SVM classifier. Experiments are conducted in similar conditions, and the correctness of the classification is compared for the four approaches to find the most appropriate feature representation for vehicle fault diagnosis.

3.3.1 Principal Component Analysis (PCA)

PCA is an unsupervised dimensionality reduction method which maps a high dimensional data set to a lower-dimensional one of uncorrelated variables called Principal Components (PCs). In a vehicle fault diagnosis, it is possible to find highly correlated or redundant information in different attributes, because it is possible for several vehicle parameters to have similar variations within the same operating or fault condition. If such attributes are given directly to the classifier, the input space can become quite high dimensional and may lead to expensive computations. To solve this, PCA computes new directions which represent as much variance as possible from the original directions.

Let$X$ represent the preprocessed feature matrix. PCA finds a transformation matrix$P$ based on the input data's covariance structure, with the reduced representation being the result as in Eq. (4).

$$Z = XP \tag{4}$$

Where$Z$ is the transformed feature matrix and P are the selected principal components. The components are ordered by the eigenvalues corresponding to them, such that the components with larger eigenvalues contain a larger percentage of the total variation of the original data. Only the necessary number of components are kept to achieve a lower dimensional representation of the data but still including the necessary information in the original components. PCA is hence used in the proposed fault diagnosis framework to reduce the dimensions of features, reduce the effect of correlation between different attributes, and to get a compact representation of SVM classification. The PCA-based representation is then fed into the SVM classifier and the diagnostic performance is compared with the other feature extraction methods.

3.3.2 Linear Discriminant Analysis (LDA)

LDA is a dimensionality reduction method that uses class information to transform the original feature space and is supervised. Whereas PCA is used to retain the overall variance of the data, LDA tries to find a projection in which observations in different classes are well separated while observations in the same class are closely grouped. This property makes LDA very applicable to fault diagnosis problems, where the focus is on determining the vehicle fault categories. The LDA transformation is determined by considering the within-class and between-class variations of the data. The goal is to achieve the highest between-class vs within-class value as in Eq. (5).

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|} \tag{5}$$

Here, the between-class scatter matrix is denoted by S_B, the within-class scatter matrix is indicated by S_W and, W represents the transformation matrix. The resulting projection allows a lower dimensional feature space suited with more discriminative information with respect to the pre-defined fault classes. The class labels are used during transformation in LDA, so it can highlight the features which discriminate among the fault conditions. In the proposed methodology, the preprocessed vehicle dataset is used to obtain the discriminative feature representations using LDA and fed to SVM to classify the vehicles. This LDA-SVM is evaluated for its performance to see if the class-oriented representation improves fault classification.

3.3.3 Non-Negative Matrix Factorization (NNMF)

NNMF is an unsupervised matrix-factorization technique that decomposes the original non-negative matrix into two smaller non-negative matrices, to provide a lower-dimensional representation of the data. When the attributes of the input have non-negative values, the technique is useful because the resulting components can be thought of as additive combinations of the original values. NNMF can be used to discover latent patterns in the dataset and to reduce the dimensionality of the dataset by representing the original observations in a lower-dimensional space, which is important in the field of vehicle fault diagnosis.

Let V be a non-negative input matrix. NNMF is an approximation of the original matrix as the product of two lower dimensional non-negative matrices W and H as shown in Eq. (6).

$$V \approx WH \quad (6)$$

Here, the term H is a matrix of the factors and W shows the lower-dimensional bias, where the number of latent components that is less than the number of features. Using this factorization, a few latent components are used to represent the original high-dimensional data. In doing this, we are reducing the dimensionality of our data, while preserving the important structures. The NNMF is applied to the preprocessed vehicle fault dataset in the proposed methodology to get a reduced non-negative representation. This representation is then fed into the SVM classifier for fault identification. Based on the NNMF representation, the classification performance is evaluated and compared to that of other representations, such as PCA- and LDA-based and tSNE-based representations, to see how well the NNMF-based representation can be used for the problem of fault diagnosis under investigation.

3.3.4 t-Distributed Stochastic Neighbor Embedding (t-SNE)

t-SNE is a dimensionality reduction method that is a non-linear version of Stochastic Neighbor Embedding (SNE). t-SNE is able to capture nonlinear aspects of relationships between observations, which is not possible with linear dimensionality-reduction methods like PCA and LDA. This property is important in vehicle fault diagnosis, when considering that the appearance of different fault modes may not necessarily be separable linearly in the original feature space.

The t-SNE algorithm first measures the similarity between observations in the original high dimensional space and then tries to maintain that similarity in the lower dimensional space. In the lower-dimensional space the similarity between observations is modeled by a student's t-distribution. Optimization is provided by minimizing similarity distances between the original space and the lower dimensional space to transform the space. Kullback–Leibler divergence is used as the optimization objective as shown in Eq. (7).

$$C = \sum_i \sum_j p_{ij} \log \left(\frac{p_{ij}}{q_{ij}}\right) \quad (7)$$

where p_{ij} is the similarity in the original space and q_{ij} is the similarity in the lower dimensional space. This study uses t-SNE to explore if there are any nonlinear local relationships in the vehicle fault data that can give a better representation to be classified. The output of the t-SNE is fed to the SVM classifier and its performance is compared with the other feature extraction methods. The classification result shows that the nonlinear local structure discovered by t-SNE is especially helpful to classify the fault categories in the investigated dataset, if the model based on t-SNE has the highest classification performance.

3.4 Construction of the SVM Classifier

After feature extraction the reduced feature representations produced by PCA, LDA, NNMF, and t-SNE are fed into a Support Vector Machine (SVM) classifier independently of each other. SVM is a supervised machine-learning method that can be used for classification, and it is a widely-adopted method due to its capacity to build efficient decision boundaries in high dimensional feature spaces. The SVM is utilized for classification in the proposed methodology to classify the vehicle observations into their fault category based on transformed feature representation. In a binary classification problem, the separating hyperplane of SVM can be represented in Eq. (8).

$$w^T x + b = 0 \quad (8)$$

Here, the term w is the weight vector, x is the input feature vector and b is the bias term. The classifier is going to find an optimum separating hyperplane which has to maximize the margin between the hyperplane and the nearest observations in the training set. The closest observations in these are called support vectors. The maximization of the classification margin is important to have good generalization capability when unseen vehicle observations are fed to the classifier.

The RBF kernel is used in the proposed SVM-based fault Classifier, because the relationship between the vehicle attributes and fault conditions may be nonlinear. The RBF kernel allows to map the input observations into a higher dimensional feature space and it allows to build nonlinear decision boundaries. The RBF kernel is given by Eq. (9).

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2) \quad (9)$$

Here, the termγ controls the influence of individual training observations. The regularization parameter C determines the balance between maximizing the separation between the different classes and minimizing the misclassification rates. The SVM parameter's values are determined based on the experimental setup adopted in this study. Four feature representations are used to compare the SVM configuration used to all of them is the same.

3.5 Integration of Feature Extraction Techniques with SVM

The proposed methodology is based on comparing the effectiveness of the four feature extraction techniques by using the same SVM classification model for each one of them. Pre-processing and feature scaling are then performed on the original vehicle data set, and each feature extraction method is used individually to reduce the dimensionality of the data. The resulting representation is then fed into RBF-based SVM classifier. The four configurations of the classifiers, i.e., PCA-SVM, LDA-SVM, NNMF-SVM and t-SNE-SVM are studied. The common use of the same classifier for all feature representations is used to ensure that the comparison is mainly between the influence of the feature extraction technique and not the difference in the classification methods used. The training and testing process, the configuration of the classifiers, and the performance measures are the same for each model. This experimental setup serves as a systematic basis to find out which feature representation is best suited for the detection of fault conditions in a vehicle.

3.6 Training and Testing of the Classification Models

The data is split into training and testing set after preprocessing and feature transformation. The former training subset is used to build the classification model, while the latter testing subset is kept to test the model on a set of previously unseen observations. If the different categories of faults are expected to be distributed unevenly, stratified partitioning can be used to ensure that the relative proportions of the various classes are the same in each of the two subsets. The feature transformation applied to the testing observations is done with the feature transformation that was learnt from the training data. This procedure is crucial as the information from the testing data shouldn't affect the training process. This separation helps minimize data leakage and is more beneficial for estimating the generalization capability of the proposed fault diagnosis model. This training and testing process is kept the same across PCA-SVM, LDA-SVM, NNMF-SVM and t-SNE-SVM to reasonably compare the results obtained.

3.7 SVM Parameter Selection

The RBF based SVM performance is sensitive to its hyperparameters including the regularization parameter C and the kernel parameterγ. The parameter C sets the trade-off between attaining a large classification margin and tolerating a high number of classification errors, whereas the parameter γcontrols the effect of each observation in the nonlinear feature space. It is therefore important to have a proper choice of the parameters to get a robust classification model. The optimal SVM parameters are determined depending on the experimental method used in the research. If the model has been optimized using cross-validation or grid search, then this should be explained here. For instance, candidate values of C and γ can be tested with cross validation on the training set, and the one that achieves the best validation performance could be chosen for the final model. To ensure an equal comparison to all four feature representations, the same procedure for selecting the parameters should be used.

3.8 Vehicle Fault Diagnosis

New vehicle observations are matched with trained SVM models to identify the fault condition associated with the new observation. In the process of diagnosis, the first step is to preprocess the input observation and transform it into a feature vector using the same preprocessors and transformations that were used during the training process. The resulting feature vector is then fed into the trained SVM classifier, which classifies the observation to its respective fault category. During diagnosis, the system can determine if any faults are likely to be present in the vehicle, according to the characteristics it has learned from the data. The results obtained from the PCA-SVM, LDA-SVM, NNMF-SVM and t-SNE-SVM classifiers can be compared to each other to determine which one is the most reliable to classify the investigated vehicle fault conditions because the proposed framework evaluates multiple feature representations.

4. Experimental Results and Discussion

The performance of the proposed machine-learning-based fault diagnosis approach was tested by varying the type of feature extraction and dimensionality reduction methods used in the system, such as PCA, t-SNE, NNMF, and LDA, and applying them with an SVM classifier. The aim of the experiments was to investigate the effect of the various feature representations on the classification of vehicle fault conditions. The data set has 617 observations and 27 original attributes, of which 20 attributes were reduced to a smaller size feature space for

classification. The performance of each feature extraction method is measured through the precision, recall, F-measure and accuracy. The results acquired by each of the four approaches are analyzed separately and then compared against to determine which representation of features is best for the fault-diagnosis problem under consideration.

4.1 Performance Evaluation

The effectiveness of the proposed fault diagnosis models is tested by the standard classification measures. Accuracy is the term used to determine the overall percentage of correctly classified observations and is calculated as in Eq. (10).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

where P , TN , FP, and FNdenote true positive, true negative, false positive, and false negative predictions, respectively. While accuracy is generally considered overall effectiveness of classification, it is not enough when the distribution of the fault classes is imbalanced. Hence, precision, recall and F1-score are also taken into consideration as it gives a wider analysis of the classification models. Precision is the percentage of correctly identified observations of all observations predicted to be of a particular class, whereas recall is the percent of observations of that class identified correctly.

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

The F1 score is both a precision and recall score. Further, a confusion matrix is used to analyse the classification behaviour for each individual fault category, as well as to determine which fault classes are most often misclassified by the classifier.

4.2 PCA-SVM Performance

To reduce the dimensionality of the set of vehicle attribute data to a lower dimensional space that would preserve the dominant information, PCA was employed. The generated features were fed to SVM classifier to classify the fault. The achieved objective function during the process is shown in Fig. 2, which is based on PCA. The minimum value of the objective function rapidly decreases in the early function evaluations, and then stabilizes, proving convergence of the optimization process. The PCA representation obtained was used as an input to SVM which showed a higher classification performance than the other methods taken into consideration in the study. The PCA results show that linear dimensionality reduction is capable of capturing sufficient information from the original vehicle attributes for good fault classification.

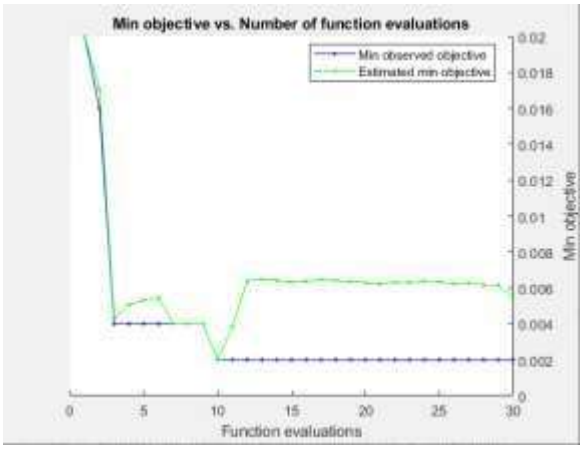

Fig 2. Objective-function convergence for the PCA-based feature extraction process

4.3 t-SNE-SVM Performance

The vehicle data was represented in a lower dimensional space in a nonlinear way using t-SNE before applying the SVM classifier. The process is presented with the objective-function behaviour in Fig. 3. During the first few function evaluations, the minimum observed objective drops significantly after which it tends to a stable area, which means that the optimization process is converging. By using the nonlinear representation created by t-SNE, the SVM classifier can be used to classify data in a different feature space, and the classification model can

be based on the relationships between the transformed data. The effectiveness of the t-SNE based approach in terms of fault detection performance is clearly illustrated in the values of recall and F-measure obtained in the overall comparison given later in this section.

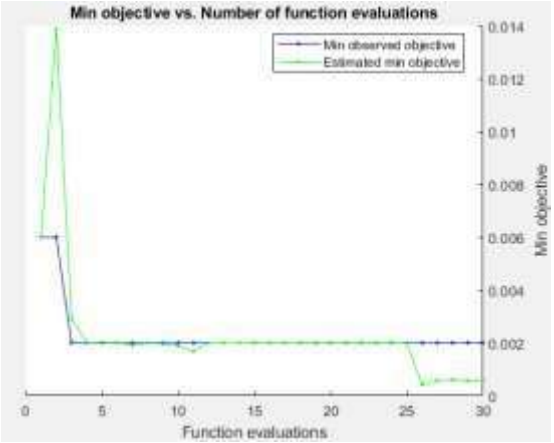

Fig 3. Objective-function convergence for the t-SNE-based feature extraction process with SVM

4.4 NNMF-SVM Performance

The original vehicle data was represented in a lower dimensional non-negative space using NNMF. The process of factorization breaks down the data into non-negative factors that can be used to interpret patterns in the data. The convergence of the objective function achieved in the NNMF process is presented in Fig. 4. The target value rapidly drops in the first function evaluations and then moves towards a stable value, which implies that the iterative process is converging towards a target value. NNMF representation was fed to the SVM vehicle fault classification. The results showed that NNMF can be used as an effective reduced representation of the classification task, and yielded comparable performance with PCA- and t-SNE-based approaches.

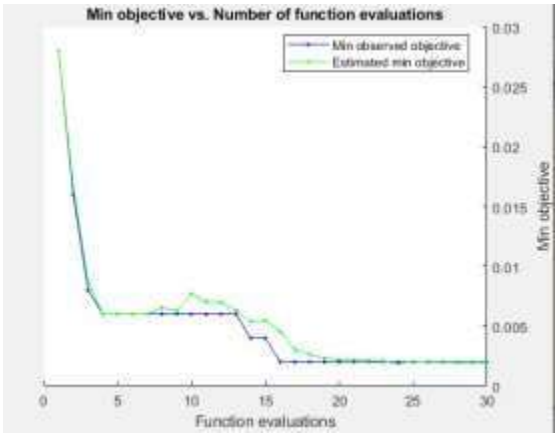

Fig 4. Objective-function convergence for the NNMF-based feature extraction process with SVM

4.5 LDA-SVM Performance

LDA was used as a supervised dimensionality reduction method to get the feature representation with class-related information. The LDA-based process objective-function behaviour is shown in Fig. 5. The minimum observed objective decreases at the first evaluations of the function and is then relatively constant meaning that the optimization process is converging. This representation was then fed to an SVM classifier, which gave an LDA representation. LDA-SVM outperformed the other feature extraction techniques based on the experimental settings used in this study. This means that the representation obtained by LDA was not as well suited for separation of the vehicle fault conditions in the test set.

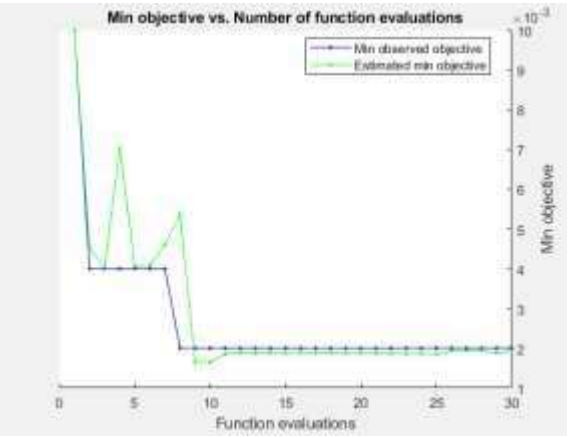

Fig 5. Objective-function convergence for the LDA-based feature extraction process with SVM

4.6 Comparative Performance Analysis

The overall classification results of PCA-SVM, t-SNE-SVM, NNMF-SVM and LDA-SVM are shown in Table 1. PCA-SVM, t-SNE-SVM, NNMF-SVM achieved same accuracy of 99.14% while LDA-SVM accuracy is 88.03%. Even though all the first 3 approaches resulted to the same accuracy values for 99.14%, only the t-SNE-SVM approach achieves the recall for 79.05% and F-measure for 79.02%; PCA-SVM achieved highest recall value for 77.60% and F-measure value for 78.82%. For the NNMF-SVM approach recall value achieved is 77.77% and F-measure value for 78.80%. The result generated by LDA-SVM are substantially lesser having a value for precision of 40%, for recall 37.58 and F-measure of 38.71. The performance results show that the t-SNE-SVM could perform well on recall and F-measure without obtaining the highest accuracy as compared to PCA-SVM and NNMF-SVM.

Table 1. Comparative performance of feature extraction techniques with the SVM classifier

SVM classifier	PCA	t-SNE	NNMF	LDA
Precision	80%	80%	80%	40%
Recall	77.6%	79.05%	77.77 %	37.58%
F- Measure	78.82%	79.02%	78.8%	38.71%
Accuracy	99.14%	99.14%	99.14%	88.03%

All four approaches are graphed and compared in Fig. 6. As can be seen from the figure, the performance of PCA-SVM is almost comparable to that of t-SNE-SVM and NNMF-SVM while LDA-SVM gives poorer performance. Especially, the precision values of the first three methods are quite similar and t-SNE-SVM method has minor benefit in the recall and F-measure value. However, the comparative results show that, even when the overall accuracy is the same, the choice of feature extraction method has an impact on the results of fault detection.

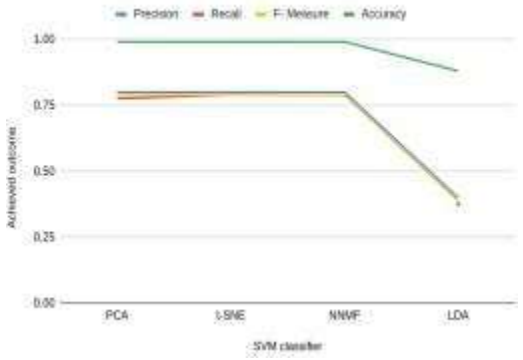

Fig 6. Comparative performance of PCA-SVM, t-SNE-SVM, NNMF-SVM, and LDA-SVM

4.7 Discussion of Results

The experimental results show that representation of features is also significant in the vehicle fault diagnosis system using SVM. The accuracy of PCA-SVM, t-SNE-SVM and NNMF-SVM have been significantly improved as compared to that of LDA-SVM, and all three gave an accuracy of 99.14%. But accuracy alone would not distinguish the three approaches since their values are the same. The accuracy achieved by the t-SNE-SVM classification algorithm is slightly better on recall and F-measure, suggesting a slight benefit in the identification of actual fault conditions as well as to keep the balance between precision and recall. The result indicates that the use of nonlinear feature representation might be useful for the analyzed dataset. However, the difference between the t-SNE-SVM and the PCA-SVM and NNMF-SVM is not significant and should be considered as a comparative advantage in this experimental set-up, not as t-SNE being always superior. The results also indicate that LDA-SVM achieved significantly less accuracy under the same experimental conditions showing the significance of selecting a feature representation that is relevant to the nature of the data and fault-classification problem.

5. Implications of the Study

5.1 Research Implications

Based on this study, some research implications can be drawn from the comparison between PCA, t-SNE, NNMF and LDA along with a common SVM classifier. The results show that although the accuracy is about the same, the two sets of features can lead to different fault-classification behavior. In particular, the representation of the input data can affect the potential of a classifier to recognize fault conditions, as t-SNE-SVM showed the best recall and F-measure values among the different approaches tested. This result could be a starting point to explore other nonlinear feature representation in machine-learning-based fault diagnosis. Further studies can be conducted by comparing more dimensionality reduction methods, feature selection algorithms, classifiers, and deep learning models. Validation with other larger datasets, other independent datasets, cross validation, and real-time data from vehicle operations would also yield more evidence of generalizability of the results.

5.2 Implications for Predictive Maintenance and Fault Diagnosis

The results are useful for predictive-maintenance and vehicle fault-diagnosis systems. Feature extraction prior to SVM classification can help to reduce the dimensionality of the vehicle data and give a compact representation for automated fault classification. In the current experiment, the t-SNE-SVM approach resulted in the highest recall and F-measure values, which is applicable for predictive maintenance because of increased recall, a larger proportion of actual fault conditions can be identified. This model may help the maintenance staff to monitor the condition, screen the faults, prioritize maintenance, and plan maintenance. The proposed approach is however intended as a decision support approach, not a substitute for engineering inspection or physical diagnosis. Further validation with real-world and continuously collected vehicle data is needed before it can be put into use in safety-critical predictive-maintenance systems.

5.3 Limitations and Future Research

There are several limitations to the findings like it was evaluated using 617 observations and 27 original attributes, and the data set does not necessarily represent all the variability of real world operating conditions of a vehicle. Furthermore, four feature extraction methods were used with the SVM classifier, and other feature representations and classifiers can yield different results. The high accuracy values must be considered in conjunction with precision and recall and F-measure values, because accuracy alone does not necessarily reflect fault-detection capability, especially in the case of an unbalanced class distribution. Larger and diverse datasets, repeated cross validation, independent external validation, class-wise performance analysis and real-time experiments should, therefore, be considered in future work. The additional extensions would give the proposed fault-diagnosis framework strength and usefulness.

6. Conclusion

In this study, PCA, t-SNE, NNMF, and LDA were used to represent features of vehicles and an SVM was used to classify the features. The results of the experiments are presented and show that the accuracy of PCA-SVM, t-SNE-SVM and NNMF-SVM are 99.14% and LDA-SVM are 88.03%. The best performance in terms of recall (79.05%) and F-measure (79.02%) was obtained with the t-SNE-SVM approach among the evaluated approaches. Thus, t-SNE-SVM gave the best results for fault detection recall and F-measure and a comparable result for accuracy, when compared with PCA-SVM and NNMF-SVM. The results show that the representation of the features can affect the performance of machine learning-based fault diagnosis. However, it also gives a comparison to choose among the feature extraction methods for predictive-maintenance applications and also emphasizes the significance of precision, recall and F-measure apart from accuracy. The proposed approach

needs to be further validated based on larger, more diverse and real-time vehicle datasets before it can be suggested for practical or safety critical use.

Statements and Declarations

Competing Interests

The authors declare that they have no competing financial or non-financial interests that could have influenced the work reported in this manuscript.

References

1. Abdennour, N., Ouni, T., & Amor, N. B. (2021). Driver identification using only the CAN-Bus vehicle data through an RCN deep learning approach. *Robotics and Autonomous Systems, 136*, 103707. https://doi.org/10.1016/j.robot.2020.103707
2. Ayoub, N., & Schneider-Kamp, P. (2021). Real-time on-board deep learning fault detection for autonomous UAV inspections. *Electronics, 10*(9), 1091. https://doi.org/10.3390/electronics10091091
3. Biddle, L., & Fallah, S. (2021). A novel fault detection, identification and prediction approach for autonomous vehicle controllers using SVM. *Automotive Innovation, 4*(3), 301–314. https://doi.org/10.1007/s42154-021-00138-0
4. Chen, F., Niu, J., & Jiang, G. (2018). Nonlinear fault-tolerant control for hypersonic flight vehicle with multi-sensor faults. *IEEE Access, 6*, 25427–25436. https://doi.org/10.1109/ACCESS.2018.2820008
5. Chen, Z., Xiong, R., Tian, J., Shang, X., & Lu, J. (2016). Model-based fault diagnosis approach on external short circuit of lithium-ion battery used in electric vehicles. *Applied Energy, 184*, 365–374. https://doi.org/10.1016/j.apenergy.2016.10.026
6. Hawkins, R., Paterson, C., Picardi, C., Jia, Y., Calinescu, R., & Habli, I. (2021). Guidance on the assurance of machine learning in autonomous systems (AMLAS). *arXiv*. https://doi.org/10.48550/arXiv.2102.01564
7. Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (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.ymssp.2019.106587
8. Liu, Z., & He, H. (2015). Model-based sensor fault diagnosis of a lithium-ion battery in electric vehicles. *Energies, 8*(7), 6509–6527. https://doi.org/10.3390/en8076509
9. Meckel, S., Schuessler, T., Jaisawal, P. K., Yang, J. U., & Obermaisser, R. (2020). Generation of a diagnosis model for hybrid-electric vehicles using machine learning. *Microprocessors and Microsystems, 75*, 103071. https://doi.org/10.1016/j.micpro.2020.103071
10. Mohamed, A., Ren, J., El-Gindy, M., Lang, H., & Ouda, A. N. (2018). Literature survey for autonomous vehicles: Sensor fusion, computer vision, system identification and fault tolerance. *International Journal of Automation and Control, 12*(4), 555–581. https://doi.org/10.1504/IJAAC.2018.095104
11. Namigtle-Jiménez, A., Escobar-Jiménez, R. F., Gómez-Aguilar, J. F., García-Beltrán, C. D., & Téllez-Anguiano, A. C. (2020). Online ANN-based fault diagnosis implementation using an FPGA: Application in the EFI system of a vehicle. *ISA Transactions, 100*, 358–372. https://doi.org/10.1016/j.isatra.2019.11.003
12. Qi, X., Qi, J., Theilliol, D., Zhang, Y., Han, J., Song, D., & Hua, C. (2014). A review on fault diagnosis and fault tolerant control methods for single-rotor aerial vehicles. *Journal of Intelligent & Robotic Systems, 73*(1), 535–555. https://doi.org/10.1007/s10846-013-9954-z
13. Ribeiro, E. G., de Queiroz Mendes, R., & Grassi Jr., V. (2021). Real-time deep learning approach to visual servo control and grasp detection for autonomous robotic manipulation. *Robotics and Autonomous Systems, 139*, 103757. https://doi.org/10.1016/j.robot.2021.103757
14. Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., & Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability Engineering & System Safety, 215*, 107864. https://doi.org/10.1016/j.ress.2021.107864
15. Tian, G. Y., Sophian, A., Taylor, D., & Rudlin, J. (2005). Wavelet-based PCA defect classification and quantification for pulsed eddy current NDT. *IEE Proceedings—Science, Measurement and Technology, 152*(4), 141–148. https://doi.org/10.1049/ip-smt:20045011
16. Wang, Y., Tian, J., Chen, Z., & Liu, X. (2019). Model based insulation fault diagnosis for lithium-ion battery pack in electric vehicles. *Measurement, 131*, 443–451. https://doi.org/10.1016/j.measurement.2018.09.007
17. Xiong, R., Sun, W., Yu, Q., & Sun, F. (2020). Research progress, challenges and prospects of fault diagnosis on battery system of electric vehicles. *Applied Energy, 279*, 115855. https://doi.org/10.1016/j.apenergy.2020.115855

18. Xiong, R., Yu, Q., & Shen, W. (2018). Review on sensors fault diagnosis and fault-tolerant techniques for lithium ion batteries in electric vehicles. In *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)* (pp. 406–410). IEEE. https://doi.org/10.1109/ICIEA.2018.8397751
19. Yang, R., Xiong, R., He, H., & Chen, Z. (2018). A fractional-order model-based battery external short circuit fault diagnosis approach for all-climate electric vehicles application. *Journal of Cleaner Production, 187*, 950–959. https://doi.org/10.1016/j.jclepro.2018.03.259
20. Yao, L., Xiao, Y., Gong, X., Hou, J., & Chen, X. (2020). A novel intelligent method for fault diagnosis of electric vehicle battery system based on wavelet neural network. *Journal of Power Sources, 453*, 227870. https://doi.org/10.1016/j.jpowsour.2020.227870
21. Zolghadri, A., Henry, D., Cieslak, J., Efimov, D., & Goupil, P. (2014). *Fault diagnosis and fault-tolerant control and guidance for aerospace vehicles*. Springer. https://doi.org/10.1007/978-1-4471-5313-9