

## Article

# Exponential Local Fisher Discriminant Analysis with Sparse Variables Selection: A Novel Fault Diagnosis Scheme for Industry Application

Zhengping Ding<sup>1</sup>, Yingcheng Xu<sup>1</sup> and Kai Zhong<sup>1,2,\*</sup> 

<sup>1</sup> Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China; q21301154@stu.ahu.edu.cn (Z.D.); q22301198@stu.ahu.edu.cn (Y.X.)

<sup>2</sup> Key Laboratory of Intelligent Computing and Signal Processing of the Ministry of Education, Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China

\* Correspondence: kaizhong0402@ahu.edu.cn

**Abstract:** Local Fisher discriminant analysis (LFDA) has been widely applied to dimensionality reduction and fault classification fields. However, it often suffers from small sample size (SSS) problem and incorporates all process variables without emphasizing the key faulty ones, thus leading to degraded fault diagnosis performance and poor model interpretability. To this end, this paper develops the sparse variables selection based exponential local Fisher discriminant analysis (SELFDA) model, which can overcome the two limitations of basic LFDA concurrently. First, the responsible faulty variables are identified automatically through the least absolute shrinkage and selection operator, and the current optimization problem are subsequently recast as an iterative convex optimization problem and solved by the minimization-maximization method. After that, the matrix exponential strategy is implemented on LFDA, it can essentially overcome the SSS problem by ensuring that the within-class scatter matrix is always full-rank, thus more practical in real industrial practices, and the margin between different categories is enlarged due to the distance diffusion mapping, which is benefit for the enhancement of classification accuracy. Finally, the Tennessee Eastman process and a real-world diesel working process are employed to validate the proposed SELFDA method, experimental results prove that the SELFDA framework is more excellent than the other approaches.

**Keywords:** matrix exponential; SSS problem; sparse variables selection; fault diagnosis



**Citation:** Ding, Z.; Xu, Y.; Zhong, K. Exponential Local Fisher Discriminant Analysis with Sparse Variables Selection: A Novel Fault Diagnosis Scheme for Industry Application. *Machines* **2023**, *11*, 1066. <https://doi.org/10.3390/machines11121066>

Academic Editor: Ahmed Abu-Siada

Received: 7 November 2023

Revised: 28 November 2023

Accepted: 29 November 2023

Published: 1 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Due to the gradually increasing requirements on system property, product quality as well as economic benefits, modern industrial processes have become even more complicated [1]. Hence, modern industrial processes urgently need advanced fault detection and isolation (FDI) techniques. Reviewing the existing FDI methods, there are two main subclasses: the data-based ones and the model-based ones. Specially, the data-based methods have been ever-accelerated recently with the continuously development of data collection and storage technologies, and received broad attention in both academy and industry domains [2–5]. Recently, lots of mature data-based approaches have been found applications in FDI field with great successes, which is also the focus of this paper.

Among all the data-based methods mentioned above, it is noteworthy that Fisher discriminant analysis (FDA) [6,7] and principal component analysis (PCA) [8,9] are the two most popular ones. In general, PCA and its extensions are an unsupervised feature extraction method and ignore the correlations among different faults. Hence, it is more suitable for fault detection rather than fault classification. In contrast, FDA belongs to supervised dimension reduction and feature extraction method, it selects a set of vectors

that maximize the distance among classes while minimize the distance between classes simultaneously, have also becomes one of the research hotspots in classifying the detected process abnormalities nowadays [10].

Despite the fact that FDA is a popular fault classification model, it tends to show degraded performance if the samples in a class show multimodality. In response to this problem, a method named local Fisher discriminant analysis (LFDA) was originally presented in [11] for dimensionality reduction. LFDA aims to overcome the limitations of local preserving projection (LPP), which cannot select the most discriminative basis vectors to construct the subspace. Simultaneously, LFDA addresses the issue of unsatisfactory classification performance of traditional FDA when dealing with multi-modal problems. By combining locality-preserving projection and Fisher Discriminant Analysis, LFDA enhances the discriminative power of features. It achieves this by preserving similarity within the local structure of the data and optimizing global discriminability through the Fisher criterion. LFDA is particularly suitable for handling multi-modal data. Then, Yu applied this method to complex chemical process monitoring [12] and shown high sensitivity in diagnosing multiple faults. After that, the authors of [13] extended the LFDA to its multiway variation and performed much better in classifying faults as well as detecting abnormal operating conditions in fed-batch operation. Furthermore, the authors of [14] developed JFDA method to describe the process dataset from both the global and local view in a high-dimensional space and received satisfactory diagnosis performance in Tennessee Eastman (TE) process. Recently, Zhong et al. [15] proposed the sparse LFDA model, which can exploit the local data information from multiple dimensions and ease the problem of multimodality and nonlinearity. Ma et al. [16] presented the hierarchical strategy based on LFDA and canonical variable analysis (CVA) for hot strip mill process monitoring.

Nevertheless, the above LFDA-based methods typically assume that the relationships between samples have been correctly described. However, this assumption can be easily violated since the relations among the operating variables are also intricate. In such context, the variable selection strategy is necessary for feature filtering and better model interpretability, drawing great attention in both industry and academia. What is more, when the dimension of dataset is bigger than the number of samples (i.e., the small-sample-size (SSS) problem). Then the LFDA-based methods will confront with singular problem due to the irreversibility of the within-class scatter matrix. Given the reliability of LFDA in the field of fault diagnosis and classification, the following two problems maybe formidable challenges: One major challenge is how to implement proper improvements on LFDA to release the hypothesis that every sample of training data should be labeled accurately beforehand. Another challenge lies in the adverse effects on model classification performance that caused by singular data matrix.

To deal with the first predicament, an alternative way could be utilizing the useful information from labeled and unlabeled samples simultaneously [17,18], thus both the supervised information and intrinsic global structure are considered. Such methods are called the semi-supervised ones and they also have been applied to fault classification field. For example, ref. [18] proposed the semi-supervised form of the FDA (SFDA), which incorporated the supererogatory unlabeled samples when conducting the fault diagnosis model, showing better fault classification over FDA and PCA. Yan et al. presented the semi-supervised mixture discriminant analysis (SMDA) [19] to monitor batch processes and both the known and unknown faults have been diagnosed correctly. Recently, researchers combined the active learning with the semi-supervised EDA model and applied to process industries successfully [20], thus improving the applicability of the traditional EDA model in real industrial processes. However, the semi-supervised LFDA for fault classification has not been studied yet in existing literatures.

When encountering the likely and common SSS problem in practical scenarios, none of these approaches can be applied to that case directly since the within-class scatter matrix in question is singular, if an appropriate technique is utilized to solve up the SSS problem,

then the model could be more broadly applicable. Fortunately, several modified methods have been provided in [21,22] to relieve the SSS problem to some extent. Nevertheless, the intrinsic limitation of FDA has not eliminated and the SFDA model could completely solve the SSS problem neither. Afterwards, Zhang et al. [23] proposed a novel advantageous FDA model from the matrix exponential perspective, which has settled the SSS problem thoroughly and show superior performance on various experiments. Soon after that, Adil et al. [24] applied the EDA model to fault diagnosis in industrial process with improved classification accuracy. Note that the authors of [25] also incorporated the exponential technique into LDE and showed better performance than the current discriminant analysis techniques in face recognition. More recently, Yu et al. [26] developed the exponential slow feature analysis (SFA) model for adaptive monitoring, which can correctly identify various operation statuses in different simulation processes. Nevertheless, the intentions of these advanced approaches are not overcome or provide practical solutions to solve the SSS problem in LFDA.

Based on the above discussions and current research status. We tend to present the sparse variables selection based exponential local Fisher discriminant analysis, referred to as SELFDA, which has not been reported in fault diagnosis field before. The salient contributions of this paper lies in the following aspects:

- The SELFDA can maximize the between-class separability and reserve the within-class local structure simultaneously through the localization factor. That is means, the multimodality of operating data has been preserved from sample dimension.
- The least absolute shrinkage and selection operator (LASSO) is used to select the responsible variables for SELFDA model effectively. Then the sparse discriminant optimization problem is formulated and solve by minimization-maximization method. Thus, the data characteristics can be well exploited from the variable dimension.
- Besides, the matrix exponential strategy is integrated into the framework of LFDA. As a consequence, the SELFDA method can function well when encountering the common SSS problem in despite of the dimensions of the input samples.
- Although SELFDA is an LFDA-based method, it is able to jointly overcome the two limitations of conventional LFDA. Thus, SELFDA is more feasible and universal in engineering practices. To our best knowledge, this paper is also the first time to leverage the SELFDA for fault classification of real-world diesel engine.

The outline of this work is structured as below. In Section 2, the classic LFDA model is reviewed. The motivation of this study as well as the specific description of the proposed SELFDA algorithm is presented in detail in Section 3. Simulation experiments are conducted on a simulation process and a real-world diesel working process in Section 4. Finally, conclusions are made in Section 5.

## 2. Revisit of LFDA

Assume the  $X_L = \{x_1, x_2, \dots, x_l\} \in R^{m \times l}$  is the labeled dataset matrix, where the vector  $x_i$  is from  $m$ -dimensional space  $R^m$ . And we assume that there are  $n_k$  labeled samples in the  $k$ th ( $1 \leq k \leq K$ ) class  $C_k$ .

Let  $S_b$  and  $S_w$  be the between-class scatter matrix and within-class scatter matrix. In order to better understand LFDA, the pairwise manner of FDA [18] is the necessary prerequisite knowledge.

$$S_b = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^b (x_i - x_j)(x_i - x_j)^T \quad (1)$$

$$S_w = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^w (x_i - x_j)(x_i - x_j)^T. \quad (2)$$

where

$$W_{i,j}^b = \begin{cases} \frac{1}{n} - \frac{1}{n_k} & \text{if } x_i \in C_k, x_j \in C_k \\ \frac{1}{n} & \text{otherwise} \end{cases} \quad (3)$$

$$W_{i,j}^w = \begin{cases} \frac{1}{n_k} & \text{if } x_i \in C_k, x_j \in C_k \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

Based on the pairwise forms of FDA, the pairwise forms of LFDA can be rewritten as below [11]:

$$S_{lb} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^{lb} (x_i - x_j)(x_i - x_j)^T$$

$$S_{lw} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^{lw} (x_i - x_j)(x_i - x_j)^T. \quad (5)$$

The weighting matrices  $W_{i,j}^{lb}$  and  $W_{i,j}^{lw}$  are given as

$$W_{i,j}^{lb} = \begin{cases} A_{i,j} \left( \frac{1}{n} - \frac{1}{n_k} \right) & \text{if } x_i \in C_k, x_j \in C_k \\ \frac{1}{n} & \text{otherwise} \end{cases} \quad (6)$$

$$W_{i,j}^{lw} = \begin{cases} \frac{A_{i,j}}{n_k} & \text{if } x_i \in C_k, x_j \in C_k \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

where  $A_{i,j}$  is the  $(i, j)$ th element of affinity matrix  $A$ , being the affinity between the  $i$ th sample and  $j$ th sample. Then the projection vectors of LFDA are obtained by the objective function described as below.

$$J_{\text{LFDA}} = \arg \max_{v_{l,i} \neq 0} \left\{ \frac{v_{l,i}^T S_{lb} v_{l,i}}{v_{l,i}^T S_{lw} v_{l,i}} \right\}. \quad (8)$$

where  $v_{l,i}$  represents the  $i$ -th discriminant vector in LFDA. It has already been proved that the solution of the above optimization problem is given by:

$$S_{lw}^{-1} S_{lb} w_{l,i} = \lambda_{l,i} w_{l,i}. \quad (9)$$

where  $\lambda_{l,i}, w_{l,i}, i = 1, 2, \dots, m$  are the generalized eigenvalues and corresponding eigenvectors, respectively.

### 3. Methodology

#### 3.1. Problem Statement and Motivation

*Statement and Motivation 1:* Generally speaking, there are always plentiful variables in real cases, especially in the complex industrial process, often resulting in inaccurate classification models. To improve the suboptimal classification performance caused by imbalanced number of normal and faulty samples, the model should explore the useful discriminant information and manifold structures from both the labeled and unlabeled samples, which is beneficial for fault classification.

However, most of the exciting LFDA models are unable to carry out variable selection in fault diagnosis area, which incorporates all process variables without emphasizing the key faulty ones. Therefore, the LASSO method is used to select the responsible variables for SELFDA model effectively. Then the sparse discriminant optimization problem is formulated and solve by feasible gradient direction method. Thus, the data characteristics can be well exploited from the variable dimension.

*Statement and Motivation 2:* Since the within class matrix  $S_{lw}$  of LFDA is non-invertible when encounter the frequent SSS problem, thus the discriminant information corresponding to the eigenvalues of  $S_{lw}$  that are equal to 0 has been ignored by LFDA, hence the LFDA cannot be applied to the SSS case. As a consequence, the application scope of fault classification methods that developed upon LFDA is largely narrowed, which has brought bottlenecks to the popularizations and implement of these methods in actual industrial process. Therefore, an efficient LFDA-based classification model able to cope with the SSS problem is badly needed.

As described in former sections, if the dimension of dataset exceeds the number of samples, then the  $S_{lw}$  in (9) is non-invertible. Thus, the optimization problem of (8) is unsolvable and the fault cannot be accurately classified through the LFDA-based methods. Actually, the SSS problem is quite common in the complex industrial process since the faulty samples are insufficient and hard to obtain in most cases. Therefore, to alleviate the SSS problem is imperative for fault classification models. In the proposed method, the matrix exponential strategy is carried out to develop the favorable model, which can completely solve the SSS issue without reducing the data dimensionality compulsively, and more practical and robust in practical applications. What is more, it also inherits the discriminant nature from LFDA and allows for the enhanced classification performance by the distance diffusion mapping.

### 3.2. SELFDA

(1) *Sparse Local Fisher Discriminant Analysis (SLFDA):* Based on LFDA, SLFDA with LASSO sparsity makes the model more concise and interpretable, the object function (9) can be reformulated as below:

$$\begin{aligned} \max_{\hat{w}_k} \quad & \hat{w}_k^T \hat{S}^b \hat{w}_k \\ \text{s.t.} \quad & \hat{w}_k^T \hat{S}^w \hat{w}_k \leq 1 \\ & \hat{w}_k^T \hat{S}^w \hat{w}_i = 0 \quad (\forall i < k). \end{aligned} \quad (10)$$

where  $\hat{w}_k$  is the  $k$ th discriminant direction of SLFDA.

The LFDA model can realize variable selection through adding L0 penalty term, which is a NP-hard problem. So the LFDA in (10) can be reformulated by adding LASSO penalty as follows:

$$\begin{aligned} \max_{\hat{w}_k} \quad & \hat{w}_k^T \hat{S}^b \hat{w}_k - \lambda \|\hat{w}_k\|_1 \\ \text{s.t.} \quad & \hat{w}_k^T \hat{S}^w \hat{w}_k \leq 1 \\ & \hat{w}_k^T \hat{S}^w \hat{w}_i = 0 \quad (\forall i < k). \end{aligned} \quad (11)$$

where  $\lambda$  is the LASSO penalty factor. In general, the interpretability and discriminant performance of SLFDA model increase with the increase of  $\lambda$  within a certain range, and then decrease.

(2) *Solution of SLFDA:* As orthogonal constraint in (11) is difficult to satisfy directly. Aiming at this problem, a new between class scatter matrix  $\hat{S}_k^b$  is designed to replace  $\hat{S}^b$ . So that the  $k$ th discriminant direction can be calculated as  $\hat{S}_k^b = \Psi^T P_k^\perp \Psi$ . If  $k = 1$ , let  $P_1^\perp = I$ , then  $\hat{S}_1^b = \hat{S}^b$ . Or else, the  $\hat{S}^b$  can be expressed as  $\hat{S}^b = \Psi^T \Psi$  through Cholesky decomposition. Then, a new orthogonal projection matrix  $M_k^\perp$  projects onto the space that orthogonal to  $\Psi w_i$  ( $i = 1, 2, \dots, k - 1$ ), which can be written as:

$$M_k^\perp = I - \Psi W (\Psi W)^+ \quad (12)$$

where  $W = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_{k-1}]$ ,  $(\Psi W)^+$  is the Moore Penrose pseudoinverse of  $\Psi W$ .

Generally speaking, problem (11) can be worked out by Lagrangian multiplier regardless of the LASSO penalty as

$$\widehat{S}_k^b \widehat{w}_k = \lambda_k \widehat{S}^w \widehat{w}_k. \tag{13}$$

And Equation (13) can be turn into the following form by multiplying  $\widehat{w}_i^T$

$$\widehat{w}_i^T \widehat{S}_k^b \widehat{w}_k = \lambda_k \widehat{w}_i^T \widehat{S}^w \widehat{w}_k. \tag{14}$$

Then, the left part of (14) can be changed into

$$\widehat{w}_i^T \widehat{S}_k^b \widehat{w}_k = (\Psi \widehat{w}_i)^T \{I - \Psi W [(\Psi W)^T \Psi W]^{-1} (\Psi W)^T\} \Psi \widehat{w}_k. \tag{15}$$

Set  $m_i = \Psi \widehat{w}_i$ , so  $m_i^T m_i = \widehat{w}_i^T \widehat{S}^b \widehat{w}_i = \lambda_i \widehat{w}_i^T \widehat{S}^w \widehat{w}_i = \lambda_i$ , and  $m_i^T m_j = 0$  when  $i \neq j$ . Set  $M = \Psi W$ , then  $M = [m_1, m_2, \dots, m_{k-1}]$ . Therefore, Equation (15) can be changed into (16)

$$\lambda_k \widehat{w}_i^T \widehat{S}^w \widehat{w}_k = [m_i^T - m_i^T M (M^T M)^{-1} M^T] m_k. \tag{16}$$

Since  $(M^T M)^{-1}$  in (16) is diagonal matrix that consist of  $\frac{1}{\lambda_i}$ , so  $m_i^T M (M^T M)^{-1} M^T$  can be expressed as:

$$[0, \dots, \lambda_i, \dots, 0] \begin{pmatrix} 1/\lambda_1 & & \\ & \ddots & \\ & & 1/\lambda_{k-1} \end{pmatrix} \begin{pmatrix} m_1^T \\ m_2^T \\ \vdots \\ m_{k-1}^T \end{pmatrix} = m_i^T. \tag{17}$$

Then we have  $\lambda_k \widehat{w}_i^T \widehat{S}^w \widehat{w}_k = 0$ , that is to say  $\widehat{w}_i^T \widehat{S}^w \widehat{w}_k = 0$ . After that, Equation (11) can be calculate as

$$\begin{aligned} \max_{\widehat{w}_k} \quad & \widehat{w}_k^T \widehat{S}_k^b \widehat{w}_k - \lambda \|\widehat{w}_k\|_1 \\ \text{s.t.} \quad & \widehat{w}_k^T \widehat{S}^w \widehat{w}_k \leq 1. \end{aligned} \tag{18}$$

since the problem in (18) is non-convex. And the minimization-maximization (MM) method that is a common choice for figuring out nonconvex functions. Thus, (18) is finally transformed into the following iterative optimization problem by MM algorithm:

$$\begin{aligned} \max_{\theta} \quad & \theta^T \widehat{S}_k^b \widehat{w}_k^i - \lambda \|\theta\|_1 \\ \text{s.t.} \quad & \theta^T \widehat{S}^w \theta \leq 1. \end{aligned} \tag{19}$$

where  $\theta$  is the parameter vector used to maximize the objective function.  $\widehat{w}_k^i$  is the optimal solution of the last iteration. In this way, the  $k$ th discriminant direction can be approximated by iterated operation, which can be solved by the feasible gradient direction method [15].

Then, the regularized forms of scatter matrixes  $S_{rlb}$  and  $S_{rlw}$  are defined in following form:

$$\begin{aligned} S_{rlb} &= (1 - \beta) \widehat{S}_k^b + \beta S_t. \\ S_{rlw} &= (1 - \beta) \widehat{S}^w + \beta I_m. \end{aligned} \tag{20}$$

where  $S_t$  is the total scatter matrix of the whole dataset [18],  $I_m$  is an identity matrix, and  $\beta \in [0, 1]$  denotes the weighting factor. Under most condition, one may choose different  $\beta$  value to increase the flexibility of the model.

(3) *Derivation Procedure of SELFDA*: In order to extract the discriminant information contained in the null space of within-class scatter matrix, the matrix exponential strategy is carried out here. Analogous to the kernel methods, in SELFDA model, suppose there is

a non-linear distance diffusion mapping  $\varphi$ , then the scatter matrices  $S_{rlb}$  and  $S_{rlw}$  can be mapped into a new high-dimensional space.

$$\varphi : \mathbb{R}^n \rightarrow F \quad (21)$$

$$S_{rlb} \rightarrow \varphi(S_{rlb}) = \exp(S_{rlb}). \quad (22)$$

$$S_{rlw} \rightarrow \varphi(S_{rlw}) = \exp(S_{rlw}). \quad (23)$$

Specifically, taking the covariance matrix  $S_{rlb}$  as an example, its exponential form can be calculated as below:

$$\exp(S_{rlb}) = I + S_{rlb} + \frac{S_{rlb}^2}{2!} + \dots = Qe^{\wedge}Q. \quad (24)$$

where  $S_{rlb} = Q \wedge Q^T$ ,  $Q$  is an orthogonal matrix and  $\wedge$  is a diagonal matrix.

Similar to the LFDA introduced in Section 2, the projection directions of SELFDA can be obtained by solving the exponential of  $S_{rlb}$  and  $S_{rlw}$  from the following optimization problem:

$$J_{\text{SELFDA}} = \arg \max_{v_{ce,i} \neq 0} \left\{ \frac{v_{ce,i}^T \exp(S_{rlb}) v_{ce,i}}{v_{ce,i}^T \exp(S_{rlw}) v_{ce,i}} \right\}. \quad (25)$$

And the matrices  $S_{rlb}$  and  $S_{rlw}$  should be normalized beforehand to prevent the appearance of large values of the numbers originating from  $\exp(S_{rlb})$  and  $\exp(S_{rlw})$ . Since the  $\exp(S_{rlb})$  and  $\exp(S_{rlw})$  are full-rank matrices according to the Theorem 3 in [23]. Thus, the SSS problem of LFDA is solved because there is no need for (25) to consider the singularity of within-class scatter matrix.

Similarly, the solution of (25) is acquired through the following Lagrangian multiplier method:

$$\exp(S_{rlb})w_{ce,i} = \lambda_{ce,i} \exp(S_{rlw})w_{ce,i}. \quad (26)$$

where  $\lambda_{ce,i}$  and  $w_{ce,i}$  are the eigenvalues and the corresponding eigenvectors, respectively. The first  $d$  eigenvectors  $W_d = [w_{ce,1}, \dots, w_{ce,d}]$  are used to span the subspace.

### 3.3. Discriminant Power of SELFDA

It is noted that SELFDA is equivalent to transforming the original data into a new space by exponential mapping, after that, the LFDA model with comprehensive sample information is carried out in such a new space, which might shown some analogous characteristics of kernel mapping. The only difference between them is that the latter maps the feature vectors while the SELFDA maps the scatter matrices. After that, the SELFDA may show enhanced performance over LFDA when involving nonlinear circumstance.

The LFDA tends to search the optimal discriminate direction, which can minimize the within-class distance and maximize the between-class distance simultaneously. In mathematics, the *trace* of the scatter matrices can be given:

$$\begin{aligned} \text{trace}(S_{rlb}) &= \sum_{i=1}^K n_k \|\mu_k - \mu\|_2^2 \\ &= \lambda_{rlb,1} + \lambda_{rlb,2} + \dots + \lambda_{rlb,m}. \end{aligned} \quad (27)$$

$$\begin{aligned} \text{trace}(S_{rlw}) &= \sum_{k=1}^K \sum_{x_i \in C_k} \|(x_i - \mu_k)\|_2^2 \\ &= \lambda_{rlw,1} + \lambda_{rlw,2} + \dots + \lambda_{rlw,m}. \end{aligned} \quad (28)$$

Since the eigenvalues of  $\text{trace}(S_{rlb})$  are often used to describe the separation between classes, while the eigenvalues of  $\text{trace}(S_{rlw})$  are often used to describe the closeness of the samples within classes. Hence, the discriminant vector that corresponds to the bigger ratio of  $\lambda_{rlb,i} / \lambda_{rlw,i}$  owns stronger discriminant power.

And the eigenvalue of the exponential matrix can be obtained by the following equation:

$$\exp(S_{rlb})w_{rlb} = Iw_{ce} + S_{rlb}w_{rlb} + \dots = e^{\lambda_{rlb}}w_{rlb}. \quad (29)$$

where  $w_{rlb}$  is the eigenvector of  $S_{rlb}$ , and  $\lambda_{rlb}$  is its corresponding eigenvalue.

Therefore, the *trace* of the SELFDA can be calculated as following:

$$\begin{aligned} \text{trace}(\exp(S_{rlb})) &= e^{\lambda_{rlb,1}} + \dots + e^{\lambda_{rlb,m}} \\ \text{trace}(\exp(S_{rlw})) &= e^{\lambda_{rlw,1}} + \dots + e^{\lambda_{rlw,m}}. \end{aligned} \quad (30)$$

It is noticeable that for most of the eigenvalues in (29) and (30), we have the inequality  $\lambda_{rlb,i} > \lambda_{rlw,i}$  and  $e^{\lambda_{rlb,i}} > e^{\lambda_{rlw,i}}$ .

Since  $\lambda_{rlb,i} > \lambda_{rlw,i} > 0$ , then  $\lambda_{rlb,i} - \lambda_{rlw,i} > 0$  and consequently since  $\forall a > 0, e^a > 1 + a$ , we have

$$e^{\lambda_{rlb,i} - \lambda_{rlw,i}} > 1 + \lambda_{rlb,i} - \lambda_{rlw,i}. \quad (31)$$

Also we have  $1 + \lambda_{rlb,i} - \lambda_{rlw,i} > 1 + \frac{\lambda_{rlb,i} - \lambda_{rlw,i}}{\lambda_{rlw,i}}$ , since  $\lambda_{rlw,i} > 1$  for large eigenvalues. We obtain the following equation through transitivity

$$e^{\lambda_{rlb,i} - \lambda_{rlw,i}} > 1 + \lambda_{rlb,i} - \lambda_{rlw,i} > \frac{\lambda_{rlb,i}}{\lambda_{rlw,i}}. \quad (32)$$

That gives us

$$\frac{e^{\lambda_{rlb,i}}}{e^{\lambda_{rlw,i}}} = e^{\lambda_{rlb,i} - \lambda_{rlw,i}} > \frac{\lambda_{rlb,i}}{\lambda_{rlw,i}}. \quad (33)$$

From the above analysis, we know that diffusion scale to between class distance is bigger than that of within class. That is means, SELFDA can enlarge the margin between different categories compared with LFDA, which is desirable for fault classification.

In order to better understand the working process of SELFDA algorithm, the pseudo-codes are given in Algorithm 1.

---

#### Algorithm 1 SELFDA

---

**Input:** Training data  $X \in R^{m \times n}$

**Output:** The data matrix projection

- 1: Establish the of SELFDA model according to (10)–(20)
  - 2: **for**  $i, j = 1$  to  $n$ ,  $\beta \in [0, 1]$  **do**
  - 3:  $S_{lb} \leftarrow \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^{lb} (x_i - x_j)(x_i - x_j)^T$
  - 4:  $S_{lw} \leftarrow \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n W_{i,j}^{lw} (x_i - x_j)(x_i - x_j)^T$
  - 5:  $S_{rlb} \leftarrow (1 - \beta)\hat{S}_k^b + \beta S_t$
  - 6:  $S_{rlw} \leftarrow (1 - \beta)\hat{S}^w + \beta I_m$
  - 7: Implement the matrix exponential on  $S_{rlb}$  and  $S_{rlw}$  to construct the optimization problem of (25)
  - 8:  $\{\lambda_{ce,i}, w_{ce,i}\} \leftarrow$  Solve the optimization problem in (25) by  $\exp(S_{rlb})w_{ce,i} = \lambda_{ce,i} \exp(S_{rlw})w_{ce,i}$
  - 9: Rank the eigenvectors  $w_{ce,i}$  according to the eigenvalues  $\lambda_{ce,i}$  in descending order
  - 10: **end for**
  - 11: **for**  $d < n$  **do**
  - 12: Choose the first  $d$  eigenvectors associated with the first  $d$  eigenvalues defined by  $W_d = [w_{ce,1}, \dots, w_{ce,d}]$
  - 13: **end for**
  - 14: The projection of  $X$  into the discriminant subspace is given by  $W_d^T X$
-

### 3.4. SELFDA-Based Fault Diagnosis Scheme

After introducing the theoretical framework of SELFDA, the Bayesian inference is borrowed to realize fault classification. Suppose all the samples follow the Gaussian distributions and thereby the priori probability of each category is  $P(x \in C_k) = \frac{1}{K}$ . Hence, the conditional probability density function (PDF) of the new testing sample  $x$  is given as:

$$P(x|x \in C_k) = \frac{\exp\left[-\frac{(x-\bar{x}_k)^T W_d \Sigma_k^{-1} W_d^T (x-\bar{x}_k)}{2}\right]}{(2\pi)^{d/2} \det[(\Sigma_k)]^{1/2}} \tag{34}$$

$$\Sigma_k = \frac{1}{n_k - 1} W_d^T \exp(S_k) W_d.$$

where  $\bar{x}_k$  and  $S_k$  are the mean vector and within-class scatter matrix of the labeled training dataset in  $C_k$ .  $W_d$  is a matrix used for dimensionality reduction of the data. In the light of Bayes rule, the posterior probability of the  $x$  belonging to  $i$ th fault category is expressed as:

$$P(x \in C_k|x) = \frac{P(x|x \in C_k)P(x \in C_k)}{\sum_{i=1}^K P(x|x \in C_k)P(x \in C_k)}. \tag{35}$$

As a result, the testing samples are classified into the certain type by the classification criterion defined in (36).

$$C(x) = \arg \max_{1 \leq k \leq K} \{P(x \in C_k|x)\}. \tag{36}$$

To simplify the discriminant task in practice, a discriminant function can be redefined as follows

$$g_k(x) = -\frac{(x - \bar{x}_k)^T W_d \Sigma_k^{-1} W_d^T (x - \bar{x}_k)}{2} - \frac{\ln[\det(\Sigma_k)]}{2}. \tag{37}$$

In conclusion, the flowchart of the favorable SELFDA method is briefly showed in Figure 1.

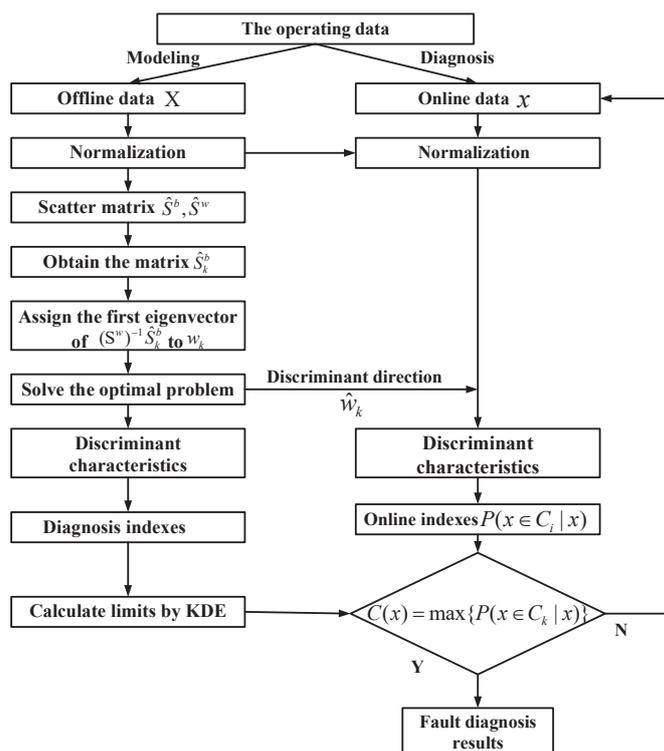


Figure 1. Flowchart of the SELFDA based fault diagnosis.

### 4. Experimental Results and Discussion

#### 4.1. TE Process

The TE process is a benchmark simulation platform of the real chemical process, which is firstly proposed by Downs and Vogel [27]. It mainly composed of five parts: a reactor, a condenser, a recycle compressor, a separator and a stripper. In this process, each sample contains 52 variables, including 12 manipulated variables and 41 measured variables. There are also 21 types of programmed fault conditions. More information about the TE process can refer to [26,28]. The process mainly produces two products with four raw materials. And Figure 2 shows the schematic diagram of the TE process.

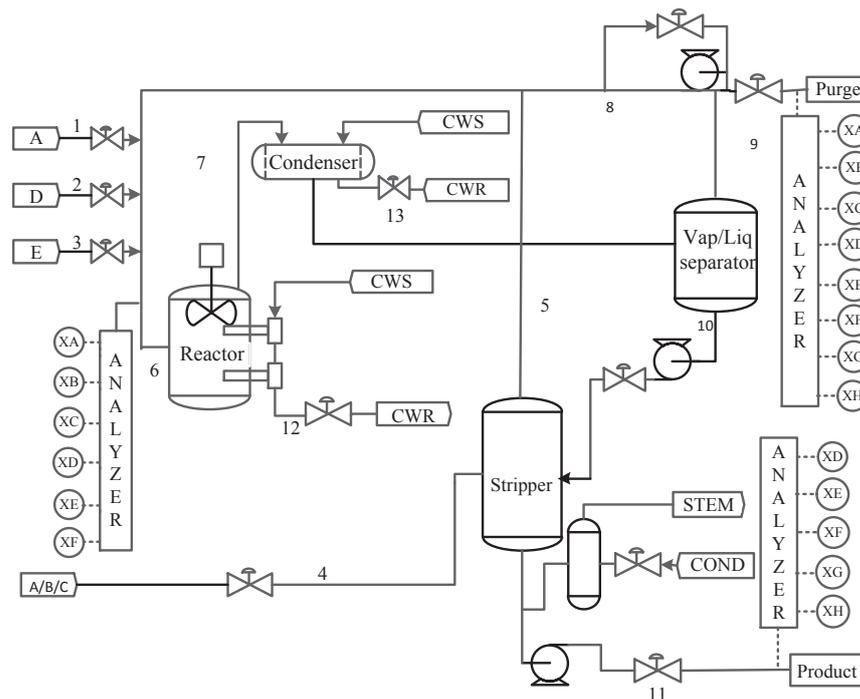


Figure 2. The schematic diagram of benchmark TE process.

In order to carry out the fault classification experiment, all the 41 measurement variables are used in this work, and Faults 1, 2 and 5 are taken as an example to carry out detailed comparative studies between FDA and LFDA, SELF and SELFDA, since they are all step faults involving process variables [18], hence it is suitable to verify the SELFDA in this paper. The description and types of the three faults are showed in Table 1. Every fault contains 200 samples for training and the labeled samples account for 5% of each class, which are selected randomly. The testing dataset include 1200 faulty samples (each failure contains 400 samples). For convenience, the value of weighting factor parameter  $\beta$  is chosen as 0.4.

Table 1. Three Faults in The TE Process.

No	Description	Type
Fault 1	A/C feed ratio B composition constant	Step
Fault 2	B composition, A/C ration constant	Step
Fault 5	Condenser cooling water inlet temperature	Step

(1) Analysis of the Visual Performance: To be noted, the classification results obtained by different methods are shown in Figure 3a–d. Specifically, there is plenty of overlaps between Fault 1 and Fault 5 in Figure 3a, that is means, the FDA cannot separate the three faults well. And the similar situation happens in Figure 3b,c. In contrast, from the visual

results in Figure 3d, we know that SELFDA method shows no overlaps in separating the three faults, which means the SELFDA achieves best classification performance with least misclassification points. To demonstrate the superiority of SELFDA better, the Euclidean distance between the centers of the three faults in the low-dimensional space are calculated and shown in Figure 4. The bar charts in different colors denote the distances obtained by different faults, and normalized them into [0, 1]. From the figure, it is easy to conclude that the three bar charts achieved by SELFDA is higher than the corresponding ones of the competitors, that is to say, the three faults are farther parted by SELFDA, which verify the classification capacity of SELFDA from another point of view.

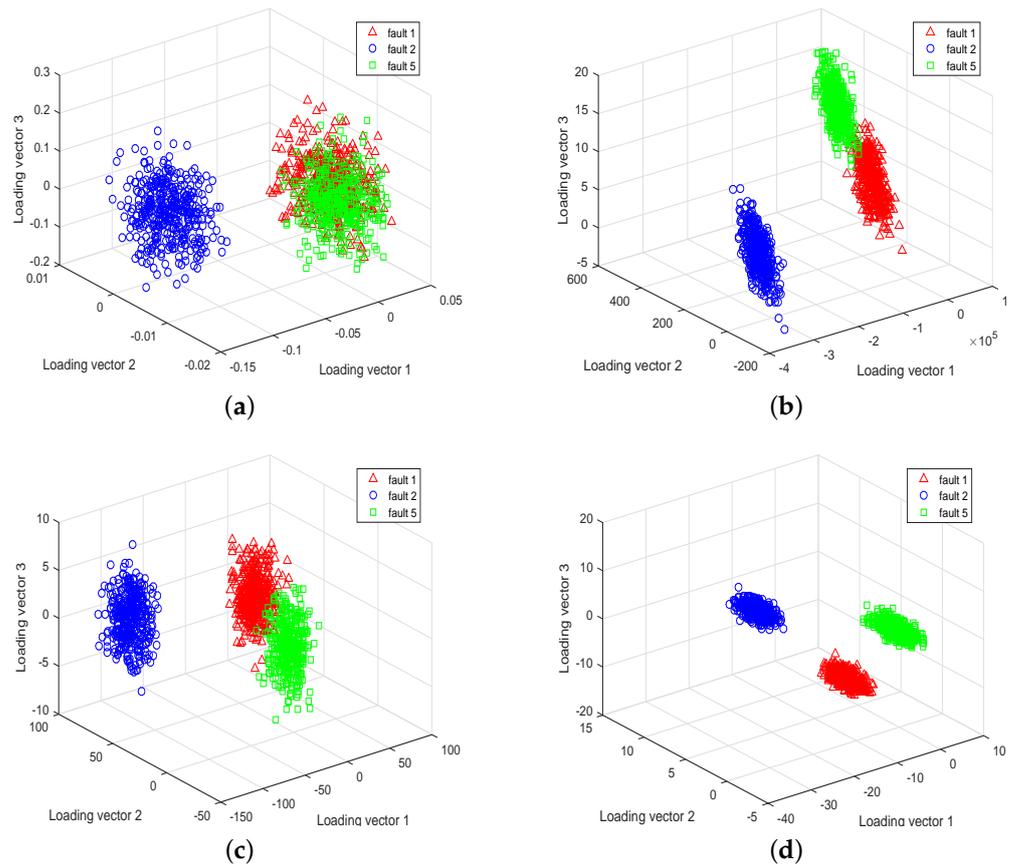


Figure 3. Classification results by (a) FDA, (b) LFDA, (c) SELF, and (d) SELFDA.

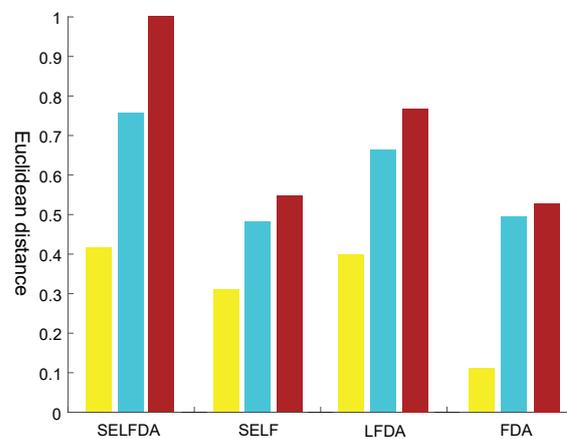


Figure 4. Between-class distances produced by different methods.

(2) *Discussions of the Classification Accuracy*: Here, we choose accuracy as the performance metric for diagnosis. Accuracy refers to the proportion of correctly classified samples by a classification model among the total number of samples. Furthermore, the quantified classification results of different algorithms are given in Table 2. Both the FDA and LFDA show undesirable performance since none of them can explore the comprehensive discriminant information. The classification performance of SELF is improved to some degree since it involves the whole dataset, however, the SSS problem is not included. By contrast, SELFDA acquires the best classification accuracy (highlighted in bold) regarding Fault 1 (100%), Fault 2 (100%) and Fault 5 (96.25%), respectively. And the average classification accuracy is also the highest (98.75%), which further explained the effectiveness of the SELFDA method. The enhanced discrimination performance of SELFDA over the comparing algorithms can be explained by two reasons. First, SELFDA makes full use of the discrimination information that contained in both the labeled and unlabeled samples. Second, it can enlarge the margins distance between different fault categories after the introduction of matrix exponential strategy.

**Table 2.** Classification Accuracy of The Faults In TE Process.

	FDA	LFDA	SELF	SELFDA
Fault 1	83.25%	83.25%	64%	<b>100%</b>
Fault 2	39.5%	0.95%	100%	<b>100%</b>
Fault 5	63.75%	69.5%	66.5%	<b>96.25%</b>
Average	62.17%	54.08%	76.83%	<b>98.75%</b>

#### 4.2. Real-World Diesel Working Process

Due to the thermal efficiency, long useful life and high reliability of the diesel engine, it has been widely utilized in ships. Nevertheless, since the aging process of the machine is irreversible, the working environment is harsh, and various faults occur frequently, thus resulting in economic losses and even casualties. To this end, ever-increasing attention should be paid to accurate and timely fault diagnosis for diesel engine. And abundant scientific payoffs have been proposed recently from both academia and industry. For example, the authors of [29] analysed the vibration data by artificial neural networks to realize fault classification in a four-stroke gasoline engine. More recently, ref. [30] presented LAMDA algorithm to classify the faults in an automotive diesel engine operating under some smooth operating conditions.

However, in-depth analysis of the above researches, none of them can handle the SSS problem or large number of unlabeled fault samples, which are quite common in diesel working process. In this paper, two universal process failures are used to confirm the superiority of SELFDA in fault classification. One of them is the exhaust pipe blockage (i.e., fault 1), the other is insufficient cooling of the air cooler (i.e., fault 2), all the operating data are collected during the working process of the 6S35ME-B9 diesel engine produced by the MAN company, which is demonstrated in Figure 5, the main system parameters and monitored sensor variables of 6S35ME-B9 diesel engine are given in Tables 3 and 4, respectively. And the detail introduction of the system parameters and data acquisition process can refer to [15].

In this experiment, each type of the faults contains 200 samples and 900 normal samples for training and the labeled samples account for 10% of each class, which are selected randomly. The testing dataset includes 100 normal samples (Fault 0) and 200 faulty samples (each type of fault contains 100 samples). For simplicity, weighting factor parameter  $\beta$  is set as 0.5 here. The detailed classification performance of different methods is shown in Figure 6. The horizontal axis denotes the position of the sample points, the first 100 samples are Fault 0, the 101–200th ones are Fault 1, and the last 100 ones are Fault 2, the vertical axis presents the label of the corresponding sample. From Figure 6, all four methods show excellent performance in classifying Fault 1. However, the classification performance of

SELFDA is obviously better than the other three methods regarding Fault 0 and Fault 2. This improvement is mainly ascribed to the advantages of SELFDA discussed earlier.

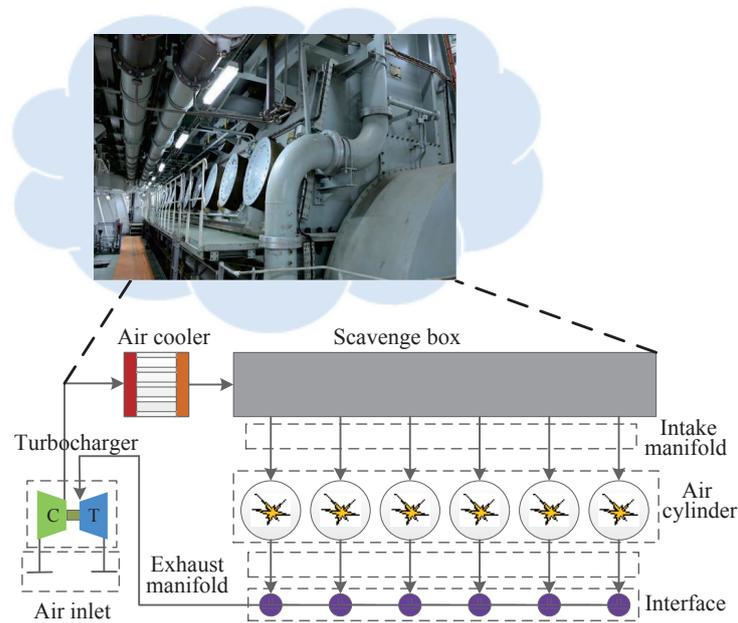


Figure 5. Layout and schematic diagram of 6S35ME-B9 diesel engine.

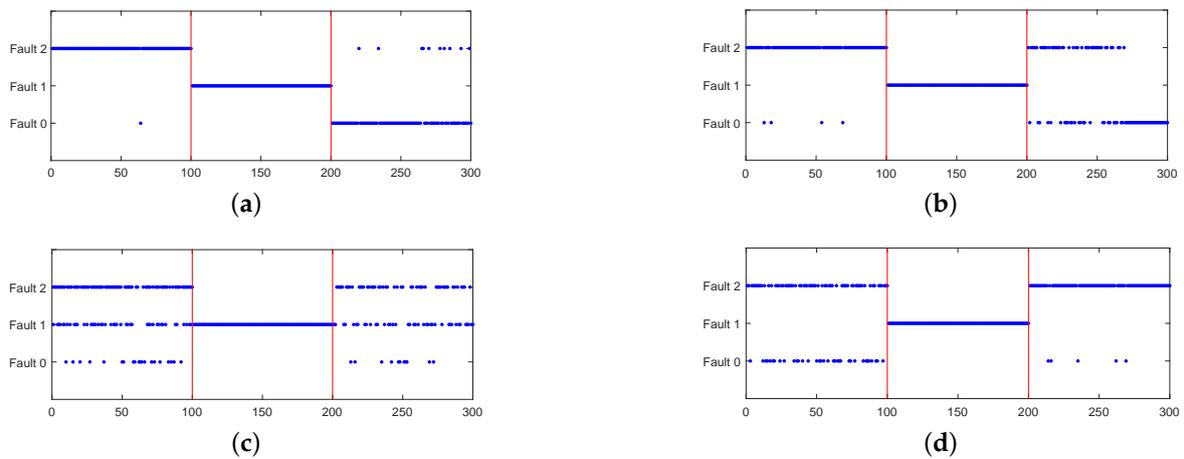


Figure 6. Fault classification performance of testing samples by (a) FDA, (b) LFDA, (c) SELF, and (d) SELFDA.

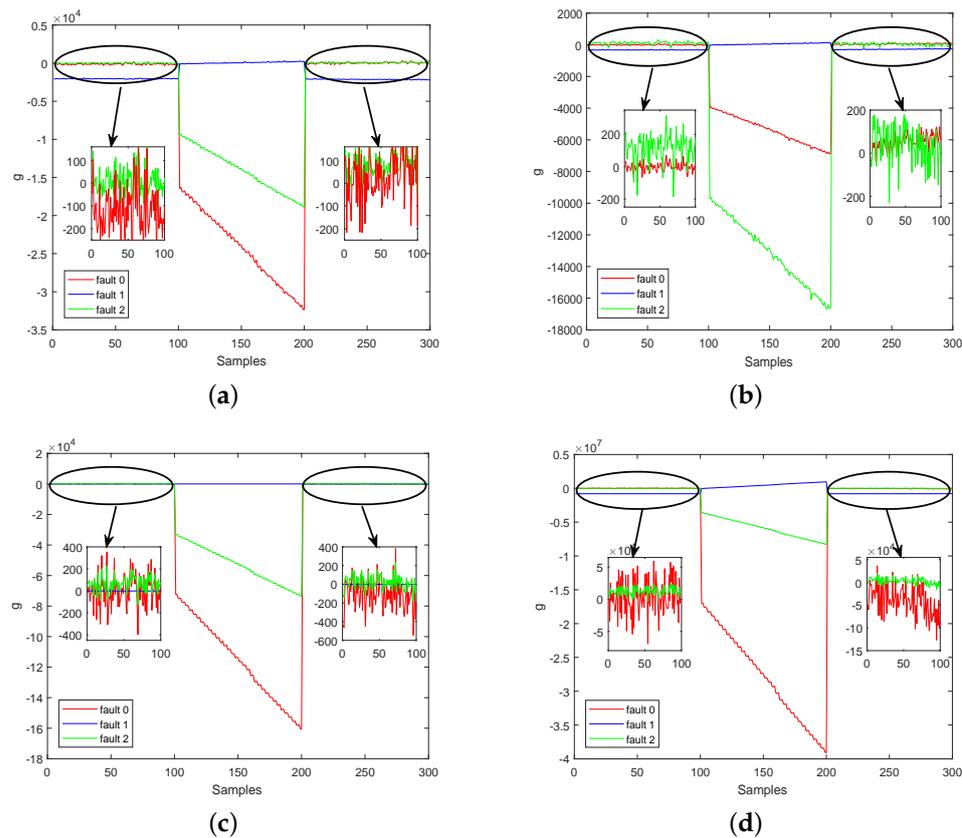
Table 3. Main System Parameters of 6S35ME-B9 Diesel Engine.

Parameter	Value	Unit
Rated power	3570	Kw
Rated speed	142	r/min
Cylinders	6	N
Fuel consumption	174.36	g/kw·h
Stroke	2	t
Oil	MGO	-
Viscosity	3–5 at 100 °C	cSt
Density	≤0.887 at 15 °C	g/cm <sup>3</sup>

**Table 4.** Monitored Variables in Diesel Engine Working Process.

No.	Variable Description	Units	No.	Variable Description	Units
1	Diesel power	kW	9	Scavenge air pressure	Bar
2	Exhaust manifold pressure	Bar	10	Scavenge air temp	°C
3	Press flow	kg/c	11	Pressure difference	Bar
4	Outlet temp of press	°C	12	Exhaust gas tempe	°C
5	Outlet pressure of press	Bar	13	Exhaust pipe pressure	Bar
6	Intercooler post temp	°C	14	Turbocharger inlet tempe	°C
7	Fuel consumption	g/kw·h	15	Turbocharger outlet tempe	°C
8	Intercooler post pressure	Bar			

Also, the discriminant functions of different methods are detailedly demonstrated in Figure 7a–d, it is easy to see that the discriminant functions of Fault 1 produced by the four methods own the maximum output in the corresponding interval (100–200th samples), which means all of them can separate Fault 1 well. However, there are a lot of intersections and overlaps between the discriminant function curves of Fault 0 and Fault 2, with a mass of false and missing classified points in all the three contrast algorithms. However, this situation is greatly relieved by SELFDA in Figure 7d, which further exhibit the superiorities of SELFDA in complex industrial processes.



**Figure 7.** Discriminant function curve of three faults achieved by (a) FDA, (b) LFDA, (c) SELF, and (d) SELFDA.

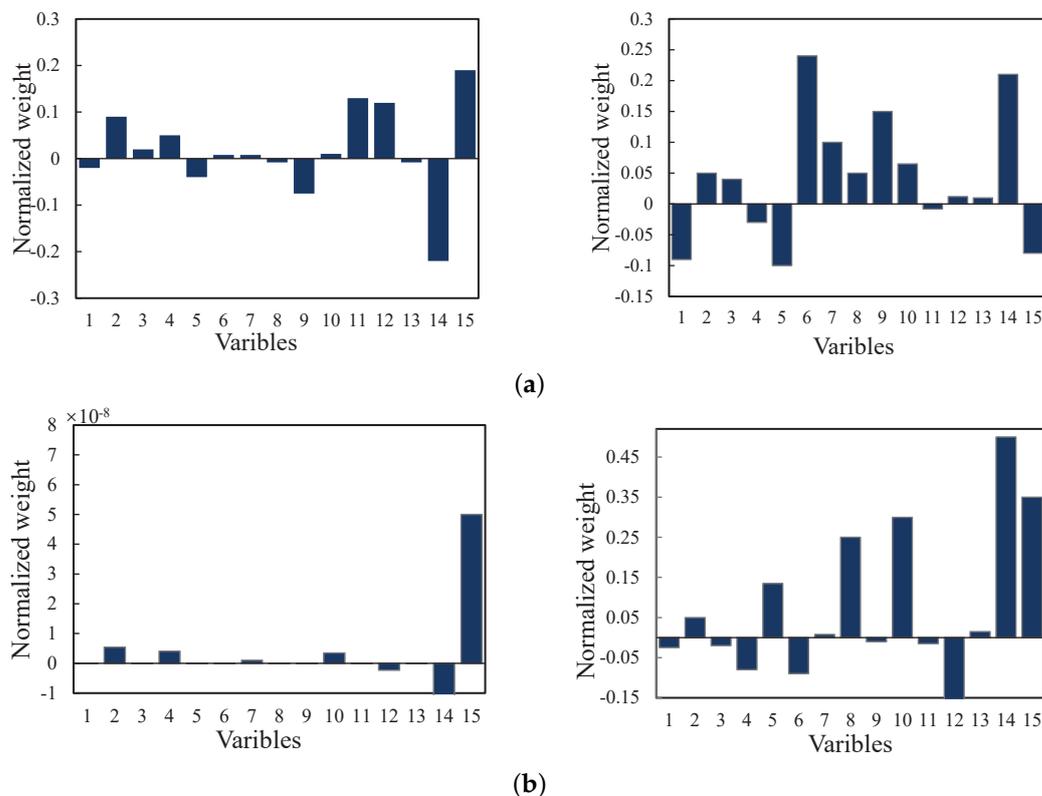
Besides, the quantitative classification results are tabulated in Table 5. And the highest classification accuracy (in bold) are all achieved by SELFDA (35% for Fault 0, 100% for Fault 1, 95% for Fault 2 and 76.67% for average). More specifically, the classification accuracy produced by SELFDA is 15% and 46% higher than the suboptimal SELF model

for Fault 0 and Fault 2, respectively. The average classification accuracy of SELFDA is also significantly enhanced compared to FDA (76.67% versus 37.00%), LFDA (76.67% versus 49%) and SELF (76.67% versus 56.33%).

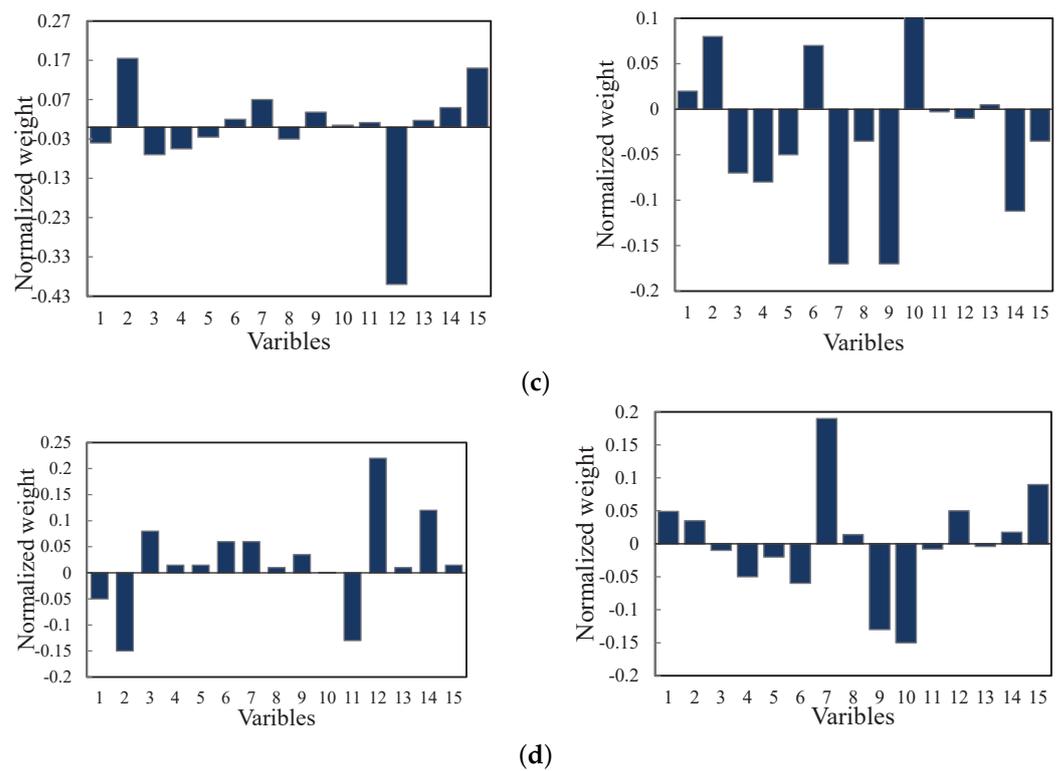
**Table 5.** Classification Accuracy of The Faults In Real-Word Diesel Working Process.

	FDA	LFDA	SELF	SELFDA
Fault 0	1%	4%	20%	35%
Fault 1	100%	100%	100%	100%
Fault 2	10%	43%	49%	95%
Average	37%	49%	56.33%	76.67%

In addition, normalized weight of each variable regarding fault 1 and fault 2 of the real-word diesel working process obtained by the four methods are given in Figure 8a–d. Specifically, from Figure 8a, we know that the first three key variables of fault 1 are misidentified as variables 11, 14, 15 by FDA. Analogously, the key variables fault 2 are also misidentified as variables 6, 9, 14. And the similar situation happens in Figure 8b since the LFDA approach is unable to handle the variable selection issue. By contrast, SLFDA can recognize the real responsible variables (variables 2, variable 12 for fault 1, variable 7, variable 9 and variable 10 for fault 2) in Figure 8c successfully, which is in consistence with actual condition. Besides, the proposed SELFDA method can also automatically select the key variables for different faults, and the details are demonstrated in Figure 8d.



**Figure 8.** Cont.



**Figure 8.** Normalized weight of each variable in Fault 1 (left) and Fault 2 (right) of diesel engine working process. (a) FDA, (b) LFDA, (c) SLFDA, (d) SELFDA.

## 5. Conclusions

This paper presents a practical SELFDA model, which improves the discrimination performance of conventional LFDA and provides a promising way to handle the challenges of multimodality, model interpretability. Besides, the novel method is insensitive to SSS problem that can enhance the performance of traditional LFDA in real industrial cases. And the simulation results on both the simulated and real industrial process demonstrated that this new fault classification framework outperforms the FDA and LFDA methods. In the future work, embedding the classification method into software systems deserves further explorations.

**Author Contributions:** Conceptualization, K.Z. and Z.D.; software, Y.X. and Z.D.; validation, Y.X. and Z.D.; writing—review and editing, Y.X. and Z.D.; supervision, K.Z.; project administration, K.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Anhui Provincial Natural Science Foundation (2208085QF205), Key Projects of Natural Science Research of Universities in Anhui Province under Grant KJ2021A0071.

**Data Availability Statement:** The data that support the findings of this study are available on request from the corresponding author, [initials] upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

The following abbreviations are used in this manuscript:

FDA	Fisher discriminant analysis
LFDA	Local Fisher discriminant analysis
SSS	Small sample size
FDI	Fault detection and isolation
PCA	Principal component analysis

LPP	Local preserving projection
TE	Tennessee Eastman process
CVA	Canonical variable analysis
SMDA	Semi-supervised mixture discriminant analysis
SFA	Slow feature analysis
LASSO	Least absolute shrinkage and selection operator
SLFDA	Sparse local Fisher discriminant analysis
PDF	Probability density function
SELFDA	Sparse variables selection based exponential local Fisher discriminant analysis

## References

1. Yin, S.; Li, X.; Gao, H.; Kaynak, O. Data-Based Techniques Focused on Modern Industry: An Overview. *IEEE Trans. Ind. Electron.* **2015**, *62*, 657–667. [\[CrossRef\]](#)
2. Subrahmanya, N.; Shin, Y.C. A data-based framework for fault detection and diagnostics of non-linear systems with partial state measurement. *Eng. Appl. Artif. Intell.* **2013**, *26*, 446–455. [\[CrossRef\]](#)
3. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer Science & Business Media: Berlin, Germany, 2000.
4. Wang, H.; Cai, Y.; Fu, G.; Wu, M.; Wei, Z. Data-driven fault prediction and anomaly measurement for complex systems using support vector probability density estimation. *Eng. Appl. Artif. Intell.* **2018**, *67*, 1–13. [\[CrossRef\]](#)
5. Tidriri, K.; Tiplica, T.; Chatti, N.; Verron, S. A generic framework for decision fusion in Fault Detection and Diagnosis. *Eng. Appl. Artif. Intell.* **2018**, *71*, 73–86. [\[CrossRef\]](#)
6. Jiang, Y.; Yin, S. Recent Advances in Key-Performance-Indicator Oriented Prognosis and Diagnosis with a MATLAB Toolbox: DB-KIT. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2849–2858. [\[CrossRef\]](#)
7. Nor, N.M.; Hassan, C.R.C.; Hussain, M.A. A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. *Rev. Chem. Eng.* **2019**, *36*, 513–553. [\[CrossRef\]](#)
8. Jiang, Q.; Yan, X.; Huang, B. Performance-Driven Distributed PCA Process Monitoring Based on Fault-Relevant Variable Selection and Bayesian Inference. *IEEE Trans. Ind. Electron.* **2016**, *63*, 377–386. [\[CrossRef\]](#)
9. Tong, C.; Lan, T.; Zhu, Y.; Shi, X.; Chen, Y. A missing variable approach for decentralized statistical process monitoring. *ISA Trans.* **2018**, *81*, 8–17. [\[CrossRef\]](#)
10. Zhao, C.; Gao, F. Critical-to-fault-degradation variable analysis and direction extraction for online fault prognostic. *IEEE Trans. Control Syst. Technol.* **2017**, *25*, 842–854. [\[CrossRef\]](#)
11. Sugiyama, M. Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis. *J. Mach. Learn. Res.* **2007**, *8*, 1027–1061.
12. Yu, J. Localized Fisher discriminant analysis based complex chemical process monitoring. *AIChE J.* **2011**, *57*, 1817–1828. [\[CrossRef\]](#)
13. Yu, J. Nonlinear bioprocess monitoring using multiway kernel localized Fisher discriminant analysis. *Ind. Eng. Chem. Res.* **2011**, *50*, 3390–3402. [\[CrossRef\]](#)
14. Feng, J.; Wang, J.; Zhang, H.; Han, Z. Fault Diagnosis Method of Joint Fisher Discriminant Analysis Based on the Local and Global Manifold Learning and Its Kernel Version. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 122–133. [\[CrossRef\]](#)
15. Zhong, K.; Han, M.; Qiu, T.; Han, B. Fault Diagnosis of Complex Processes Using Sparse Kernel Local Fisher Discriminant Analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 1581–1591. [\[CrossRef\]](#)
16. Liang, M.; Jie, D.; Kaixiang, P.; Chuanfang, Z. Hierarchical Monitoring and Root Cause Diagnosis Framework for Key Performance Indicator-Related Multiple Faults in Process Industries. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2091–2100.
17. Adeli, E.; Thung, K.H.; An, L.; Wu, G.; Shi, F.; Wang, T.; Shen, D. Semi-supervised discriminative classification robust to sample-outliers and feature-noises. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 515–522. [\[CrossRef\]](#)
18. Zhong, S.; Wen, Q.; Ge, Z. Semi-supervised Fisher discriminant analysis model for fault classification in industrial processes. *Chemom. Intell. Lab. Syst.* **2014**, *138*, 203–211. [\[CrossRef\]](#)
19. Yan, Z.; Huang, C.C.; Yao, Y. Semi-supervised mixture discriminant monitoring for chemical batch processes. *Chemom. Intell. Lab. Syst.* **2014**, *134*, 10–22. [\[CrossRef\]](#)
20. Liu, J.; Song, C.; Zhao, J. Active learning based semi-supervised exponential discriminant analysis and its application for fault classification in industrial processes. *Chemom. Intell. Lab. Syst.* **2018**, *180*, 42–53. [\[CrossRef\]](#)
21. Chen, L.F.; Liao, H.Y.M.; Ko, M.T.; Lin, J.C.; Yu, G.J. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognit.* **2000**, *33*, 1713–1726. [\[CrossRef\]](#)
22. Ye, J.; Li, Q. A two-stage linear discriminant analysis via QR-decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 929–941.
23. Zhang, T.; Fang, B.; Tang, Y.Y.; Shang, Z.; Xu, B. Generalized discriminant analysis: A matrix exponential approach. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2009**, *40*, 186–197. [\[CrossRef\]](#)
24. Adil, M.; Abid, M.; Khan, A.Q.; Mustafa, G.; Ahmed, N. Exponential discriminant analysis for fault diagnosis. *Neurocomputing* **2016**, *171*, 1344–1353. [\[CrossRef\]](#)

25. Dornaika, F.; Bosaghzadeh, A. Exponential local discriminant embedding and its application to face recognition. *IEEE Trans. Cybern.* **2013**, *43*, 921–934. [[CrossRef](#)]
26. Yu, W.; Zhao, C. Recursive Exponential Slow Feature Analysis for Fine-Scale Adaptive Processes Monitoring With Comprehensive Operation Status Identification. *IEEE Trans. Ind. Inform.* **2019**, *15*, 3311–3323. [[CrossRef](#)]
27. Downs, J.J.; Vogel, E.F. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17*, 245–255. [[CrossRef](#)]
28. Tong, C.; Lan, T.; Shi, X. Double-layer ensemble monitoring of non-gaussian processes using modified independent component analysis. *ISA Trans.* **2017**, *68*, 181–188. [[CrossRef](#)]
29. Ahmed, R.; Sayed, M.E.; Gadsden, S.A.; Tjong, J.; Habibi, S. Automotive Internal-Combustion-Engine Fault Detection and Classification Using Artificial Neural Network Techniques. *IEEE Trans. Veh. Technol.* **2015**, *64*, 21–33. [[CrossRef](#)]
30. Ruiz, F.A.; Isaza, C.V.; Agudelo, A.F.; Agudelo, J.R. A new criterion to validate and improve the classification process of LAMDA algorithm applied to diesel engines. *Eng. Appl. Artif. Intell.* **2017**, *60*, 117–127. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.