

Establish Induction Motor Fault Diagnosis System Based on Feature Selection Approaches with MRA

Authors:

Chun-Yao Lee, Meng-Syun Wen

Date Submitted: 2021-02-22

Keywords: artificial neural network, correlation and fitness value-based feature selection, multiresolution analysis, correlation-based feature selection, feature selection, Fault Detection

Abstract:

This paper proposes a feature selection (FS) approach, namely, correlation and fitness value-based feature selection (CFFS). CFFS is an improvement feature selection approach of correlation-based feature selection (CFS) for the common failure cases of the induction motor. CFFS establishes the induction motor fault detection (FD) system with artificial neural network (ANN). This study analyzes the current signal of the induction motor with multiresolution analysis (MRA), extracts the features, and uses feature selection approaches (Relief, CFS, and CFFS) to reduce the number of features and maintain the accuracy of the induction motor fault detection system. Finally, the induction motor fault detection system is trained by the feature selection approaches selected features. The best induction motor fault detection system will be established through the comparison of the efficiency of these FS approaches.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):

LAPSE:2021.0046

Citation (this specific file, latest version):

LAPSE:2021.0046-1

Citation (this specific file, this version):

LAPSE:2021.0046-1v1

DOI of Published Version: <https://doi.org/10.3390/pr8091055>

License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

Establish Induction Motor Fault Diagnosis System Based on Feature Selection Approaches with MRA

Chun-Yao Lee * and Meng-Syun Wen

Department of Electrical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Taoyuan City 320, Taiwan; g10878027@cycu.org.tw

* Correspondence: cyl@cycu.edu.tw; Tel.: +886-3-2654827

Received: 6 August 2020; Accepted: 26 August 2020; Published: 29 August 2020



Abstract: This paper proposes a feature selection (FS) approach, namely, correlation and fitness value-based feature selection (CFFS). CFFS is an improvement feature selection approach of correlation-based feature selection (CFS) for the common failure cases of the induction motor. CFFS establishes the induction motor fault detection (FD) system with artificial neural network (ANN). This study analyzes the current signal of the induction motor with multiresolution analysis (MRA), extracts the features, and uses feature selection approaches (ReliefF, CFS, and CFFS) to reduce the number of features and maintain the accuracy of the induction motor fault detection system. Finally, the induction motor fault detection system is trained by the feature selection approaches selected features. The best induction motor fault detection system will be established through the comparison of the efficiency of these FS approaches.

Keywords: fault detection; feature selection; multiresolution analysis; correlation-based feature selection; correlation and fitness value-based feature selection; artificial neural network

1. Introduction

As the fourth industrial revolution emerges, the way factories work will never be the same. Factory automation has benefits such as economizing manpower, avoiding malfunction, as well as smart machines analyzing and diagnosing issues without human intervention. Motors are widely used in factories and could affect the operation of the system and cause production equipment failure. Therefore, motor maintenance is important for detecting motor failure, and then fixing the problem with online systems, before production equipment gets severely affected [1–3], while also reducing the cost of equipment failure. There are various common causes of motor failure, among which 45% are bearing failure, 35% are stator winding failure, 10% are rotor failure, and another 10% are other damages [4]. This study analyzes motor failure in different conditions, including a functional motor, bearing failure, stator winding failure, and rotor failure. We measure the current signal of the motor and analyze the captured signal feature through signal processing.

Recently, there have been many signal processing methods, such as (1) fast Fourier transform (FFT) [5,6], which accelerates the calculation of discrete Fourier transform (DFT), also known as the basic digital signal process (DSP) approach. FFT analyzes the frequency of the signal but cannot analyze the correlation between frequency and time, according to [7,8]. (2) Gabor transform (GT) analyzes the frequency bandwidth fuzzy. (3) Winger distribution function (WDF) analyzes the frequency clearly, but has cross-term problems, causing quantities of noise at other bandwidths. (4) Gabor-Winger transform (GWT) solves the problems of Gabor transform and Winger distribution function, yet GWT is not used in this study as it cannot analyze other bandwidths in the frequency domain. (5) Wavelet transform (WT) [9,10] affects the analysis result through mother wavelet, scaling function, and displacement function. (6) Multiresolution analysis (MRA) [11] changes the scaling function and

displacement function based on the level of the analysis. MRA has the characteristics, such as high pass filter and low pass filter, to capture the feature, which is not significant for signal but important for the classification system. Additionally, MRA is widely used to analyze the power failure issues concerning power quality and power lines transmission [12,13]. Therefore, this study uses MRA to analyze the current signal to establish the FD system.

According to [14], through feature engineering, design feature sets could improve the effect of recognition systems and reduce operating costs. Common features of engineering include (1) feature construction [15,16], which defines the new features based on the original features according to unobservable information from the data, and uses all the features for the recognition system; (2) feature extraction [17,18], the transfer function based on high-dimensional feature sets reduces feature dimensions, improves the effect of the recognition system, and avoids the Hughes phenomenon, which increases the number of features and would lead to a decrease in the accuracy of the system; (3) feature selection (FS) [19] can be divided into two parts: filter and encapsulation (wrapper). The filter method selects the features based on the correlation between features, and uses correlation analysis algorithms: Relief [20,21] and ReliefF [22], and uses feature selection approaches: ReliefF and correlation-based feature selection (CFS) [23] to select the features based on the correlation, and the calculation cost efficiency is higher. The wrapper method selects the features based on evaluation function, and the calculation cost efficiency is lower.

This study uses artificial neural networks (ANN) to establish recognition systems with features that are selected with feature selection approaches. According to [24], artificial neural networks are supervised machine learning, suitable as a classifier with non-linear input data. The ANN module includes the probabilistic neural network (PNN), radial basis function neural network (RBNN), and back propagation neural network (BPNN), etc. As the induction motor current signal is limited, this study uses ANN as a recognition system and stochastic gradient descent (SGD) [25] to calculate the error and fix the network of each sample. The time of training will be longer, but the effect will be better if the parameters are properly adjusted in neural network.

2. Signal Analysis and Classifier

2.1. Wavelet Transform (WT)

Wavelet transform was developed by A. Grossmann and J. Morlet. Wavelet transform is shown as (1). Mother wavelet $\psi(t)$ decomposed into several daughter wavelets $\psi_{a,b}(t)$ shown as (2). Calculate the wavelet coefficient $CWT(a,b)$ by inner product of original signal $f(t)$ and daughter wavelets, and reconstruct the original signal by wavelet coefficient, shown as (3).

$$CWT(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt \quad (1)$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (2)$$

$$f(t) = \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{b^2} CWT(a,b) \frac{1}{\sqrt{b}} \psi\left(\frac{t-b}{a}\right) da db \quad (3)$$

where, a is scaling function, and b is displacement function.

2.2. Multiresolution Analysis (MRA)

This study uses multiresolution analysis (MRA) for signal processing and selects the features of time domain and frequency domain of the signal that is reconstructed by the wavelet coefficient. Multiresolution analysis was developed by Mallat. Multiresolution analysis is shown as (4).

MRA decomposes the original signal $f(t)$ into detail coefficient d_j and approximation coefficient a_j with scaling function $\varphi(t)$ and wavelet function $\psi(t)$, shown as (5) and (6).

$$f(t) = \sum_k a_{j0,k} \varphi_{j0,k}(t) + \sum_j \sum_k d_{j,k} \psi_{j,k}(t) \tag{4}$$

$$\varphi(t) = \sum_k g_0(k) + \varphi_k(2t - k) \tag{5}$$

$$\psi(t) = \sqrt{2} \sum_k h_0(k) + \varphi_k(2t - k) \tag{6}$$

where, g_0 and h_0 are coefficients of filter. The schematic diagram of MRA shown in Figure 1.

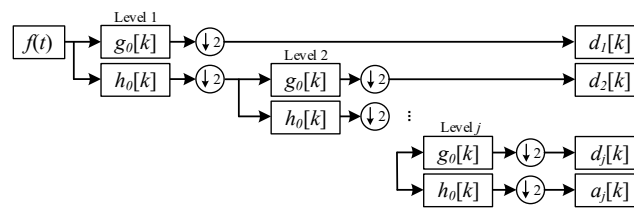


Figure 1. Schematic diagram of multiresolution analysis (MRA).

2.3. Artificial Neural Network Training (ANN)

This study uses artificial neural network training to classify and recognize the four cases of induction motor failure with the features that were selected by feature selection approaches. The construct of the neural network was developed by McCulloch and Pitts, as shown in Figure 2. The network is composed of an input layer (j), hidden layers (k, l), and output layer (z) with neurons. The input layer inputs the data into network, the hidden layers calculate the input with weight, bias, and activation function, and you receive the result from the output layer. Backpropagation [26] is a combination of multi-layer feed-forward neural network [27] and error back propagation [28]. This artificial neural network was designed with the model as nervous system, using stochastic gradient descent to fix the coefficient of the neural network to promote the classification system.

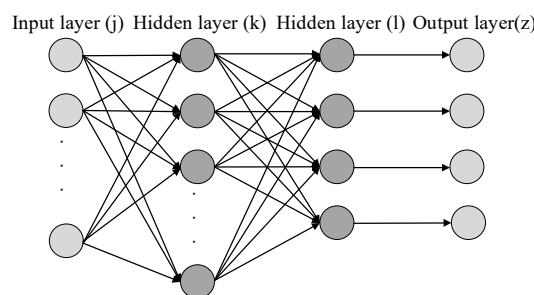


Figure 2. Construct of neural network.

The feedforward propagation is shown in Figure 3a. Calculate the output with input layer, hidden layers, and output layer. Each neuro of the network has one weight w_{jk} . All of the input data that multiplies with the weight will be summated, and one bias b_{k1} of the neuron will be added to get the output z_{k1} transfer with activation function $f(x)$. The back propagation is shown in Figure 3b. Calculate the error between the output of feedforward propagation and the target output and select the learning rate to fix the weight and bias of the network with error stochastic gradient descent. We could get the new weight and bias for neural network.

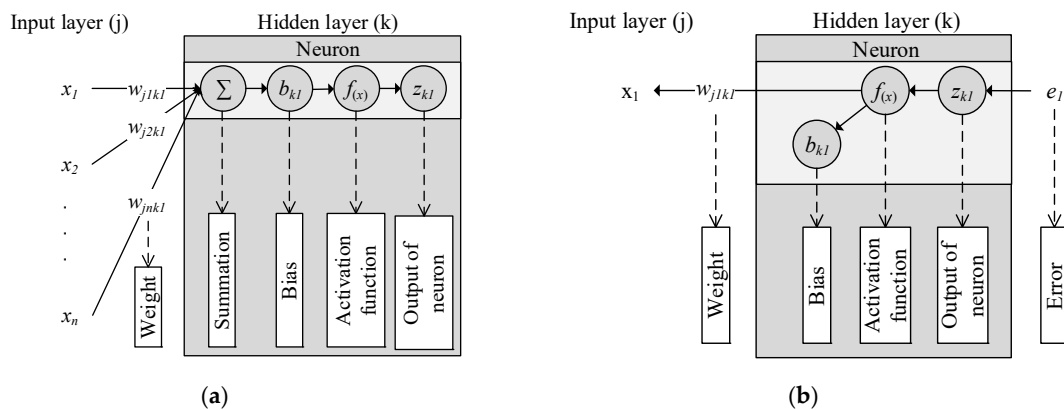


Figure 3. Schematic diagram of artificial neural network (ANN). (a) Feedforward propagation; (b) back propagation.

The construct of ANN of this study is shown as Figure 2. The steps are listed as follows:

Step 1. Set the input data (z_j), target output (t), numbers of feature (f), numbers of neurons (c), weights (w_{jk}) between input layer and hidden layer (k), biases (b_k) of hidden layer (k), activation function ($f_{k(x)}$) of hidden layer (k), weights (w_{kl}) between input layer and hidden layer (l), biases (b_l) of hidden layer (l), activation function ($f_{l(x)}$) of hidden layer (l), learning rate (ln), times of training ($Times$), times of iteration (T). This study sets $c = f + 2$, $ln = 0.007$, $T = 50$, all the weights and biases are random numbers between 0 and 1, activation function ($f_{k(x)}$) is hyperbolic tangent sigmoid transfer function [29], activation function ($f_{l(x)}$) is soft max function [30].

$$w_{jk} = \begin{pmatrix} w_{j1k1} & \cdots & w_{j1kf} \\ \vdots & \ddots & \vdots \\ w_{jck1} & \cdots & w_{jckf} \end{pmatrix}, b_k = \begin{pmatrix} b_{k1} \\ \vdots \\ b_{kc} \end{pmatrix}, w_{kl} = \begin{pmatrix} w_{kl11} & \cdots & w_{kl1c} \\ \vdots & \ddots & \vdots \\ w_{kl41c} & \cdots & w_{kl41c} \end{pmatrix}, b_l = \begin{pmatrix} b_{l1} \\ \vdots \\ b_{l4} \end{pmatrix}$$

where, hyperbolic tangent sigmoid transfer function, shown as (7), limits the output data to -1 and 1 , as shown in Figure 4. Softmax function, shown as (8), calculates the input data and let summation of output data equal 1, shown as Figure 5, and the maximum term is regarded as the classification result.

$$f_{k(x)} = \frac{2}{1 + e^{-2x}} - 1 \quad (7)$$

$$f_{l(x_j)} = \frac{e^{x_j}}{\sum_{j=1}^n e^{x_j}}, x_j = x_1, x_2, \dots, x_n \quad (8)$$

Step 2. Calculate the output (z_l) of feed-forward propagation as (9)–(12).

$$h_k = w_{jk} \times z_j + b_k \quad (9)$$

$$z_k = f_k(h_k) \quad (10)$$

$$h_l = w_{kl} \times z_k + b_l \quad (11)$$

$$z_l = f_l(h_l) \quad (12)$$

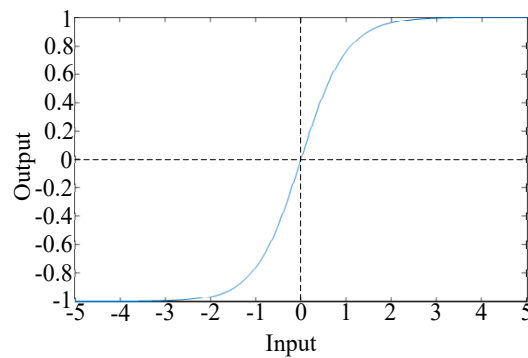


Figure 4. Hyperbolic tangent sigmoid transfer function.

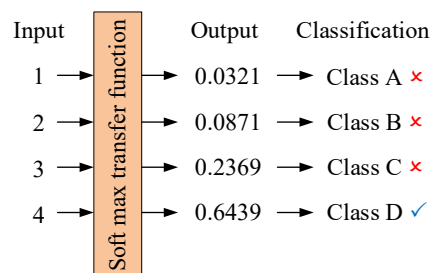


Figure 5. Schematic diagram of softmax.

Step 3. Calculate the error (E) between output (z_l) and target output (t) by cross entropy, shown as (13).

$$E = -\frac{1}{4} \sum_{i=1}^4 t_i \log(z_l) \quad (13)$$

Step 4. Use stochastic gradient descent algorithm (SGD) to calculate the error about weights (w_{kl}) and biases (b_l) from output layer to hidden layer (l) as (14) and (15), fix weights (w_{kl}) and biases (b_l) as (16) and (17).

$$\frac{\partial e}{\partial w_{kl}} = \frac{\partial e}{\partial z_l} \times \frac{\partial z_l}{\partial h_l} \times \frac{\partial h_l}{\partial w_{kl}} \quad (14)$$

$$\frac{\partial e}{\partial b_l} = \frac{\partial e}{\partial z_l} \times \frac{\partial z_l}{\partial h_l} \times \frac{\partial h_l}{\partial b_l} \quad (15)$$

$$w_{kl} = w_{kl} - 0.007 \times \frac{\partial e}{\partial w_{kl}} \quad (16)$$

$$b_l = b_l - 0.007 \times \frac{\partial e}{\partial b_l} \quad (17)$$

Step 5. Use stochastic gradient descent algorithm (SGD) to fix weights (w_{jk}) and biases (b_k), the process is shown as (14)–(17).

Step 6. If $Times$ is not equal T , calculate $Times$ as (18), and back to **Step 2**.

$$Times = Times + 1 \quad (18)$$

Step 7. ANN completely established. The flowchart of ANN is shown in Figure 6.

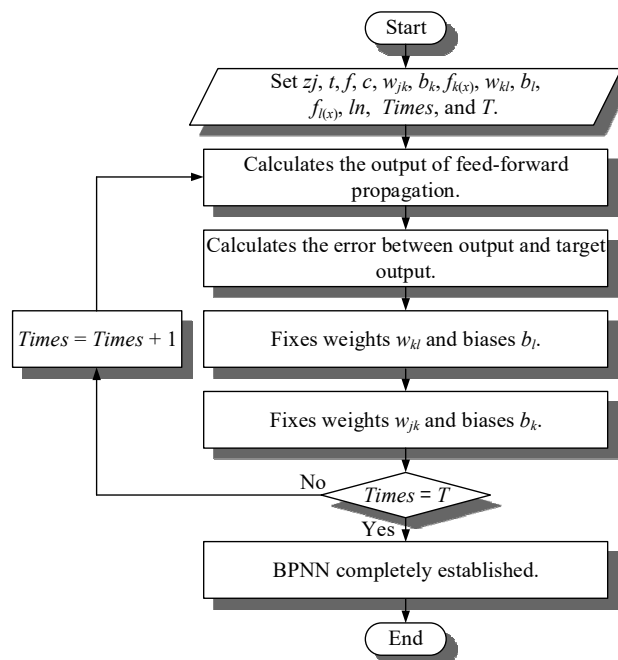


Figure 6. Flowchart of the ANN.

3. Correlation Analysis Algorithm

This chapter introduces the correlation analysis algorithms: Relief and ReliefF algorithms were used in the research. Additionally, this chapter compares three feature selection approaches: ReliefF, CFS, and correlation and fitness value-based feature selection (CFFS). The features extracted from the current signal of induction motor are not all useful for the classifier, some of the features do not affect the recognition result, and some of the features reduce the recognition effect. This study uses feature selection approach to delete the useless features to increase the recognition of the classifier and reduce the cost of calculation.

3.1. Relief

This study uses Relief to calculate the correlation between features. Relief was proposed by K. Kira and L.A. Rendell. This feature selection approach was designed for application to binary classification problems. Relief calculates the correlation between feature (F_h) and target (F_m), selects one stochastic value (f_h) from feature (F_h), chooses one nearest value (f_{nh} , Near-hit) of feature (F_h) and one nearest value (f_{nm} , Near-miss) of target (F_m) with it. The steps of Relief are listed as follows, and the flowchart of Relief is shown in Figure 7.

- Step 1.** Set the set of features (F_i), and maximum times of sampling (k). Initialize the correlation (W_i) to zero, and times of sampling (i) to one. Set of features means all of the features sampling $F_i = \{F_1, F_2, \dots, F_p\}$. F_p is one of the features, value (f_1, f_2, \dots, f_t).
- Step 2.** Choose two features (F_h, F_m) from set of features.
- Step 3.** Select one value (f_h) from one of the features (F_h).
- Step 4.** Select one nearest value (f_{nh}) with f_h from F_h , and a nearest value (f_{nm}) with F_h from F_m .
- Step 5.** Calculate the correlation (R_f) with Formula (19),

$$R_f = W_i - \left(\frac{1}{k}\right) \text{diff}(f_h, f_{nh})^2 + \left(\frac{1}{k}\right) \times \text{diff}(f_h, f_{nm})^2 \quad (19)$$

where $\text{diff}(f_h, f_{nh})$ is the distance between f_h and f_{nh} , $\text{diff}(f_h, f_{nm})$ is the distance between f_h and f_{nm} .

Step 6. If $i \geq k$, select i with Formula (20), and W_i with Formula (21). Go back to **Step 3**.

$$W_i = R_f \quad (20)$$

$$i = i + 1 \quad (21)$$

Step 7. Get the correlation (R_f) between feature F_h and feature F_m .

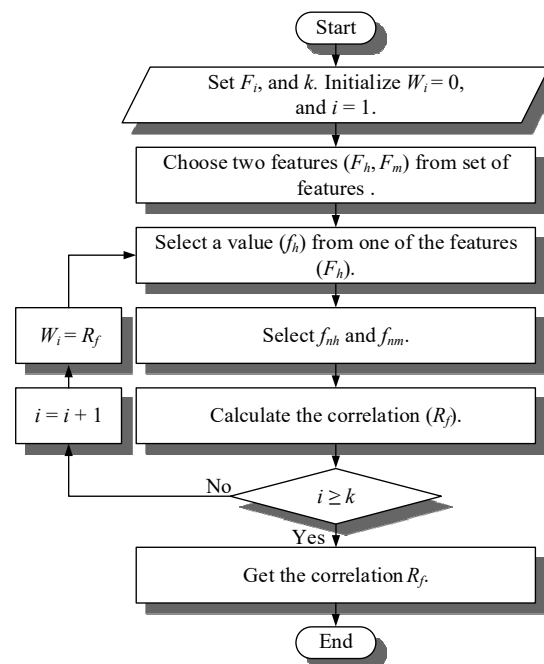


Figure 7. Flowchart of the Relief.

3.2. ReliefF

This study uses ReliefF to calculate the correlation between feature and classification. ReliefF was proposed by I. Kononenko. This feature selection approach is designed for multiclass classification. ReliefF calculates the correlation between feature (F_h) and classification (C), selects one stochastic value (f_h) from feature, chooses one nearest value (f_{sh} , Near-hit) of the same classification with f_h , and one nearest value (f_{sm} , Near-miss) of each different classification with f_h . The steps of ReliefF are listed as follows, and the flowchart of ReliefF is shown in Figure 8.

- Step 1.** Set the set of features (F_i), maximum times of sampling (k), classification of sampling ($C_{a,b}$), and the maximum sample number of the nearest value (s). Initialize the correlation (W_i) to zero, times of sampling (i) to one, and sample number of the nearest value (j). The set of features means all of the feature sampling $F_i = \{F_1, F_2, \dots, F_p\}$. F_p is one of the features, value (f_1, f_2, \dots, f_i) . Classification of sampling means all of the feature sampling is classification to m class $C_{a,b} = (c_{1,1}, c_{2,1}, \dots, c_{i,1}, c_{(i+1),2}, c_{(i+2),2}, \dots, c_{j,2}, c_{(j+1),3}, \dots, c_{t,m})$, where a is the order about value of feature, b is class b .
- Step 2.** Choose one feature (F_h) from set of features and select a value (f_h) from F_h . f_h is belong to n class of all classification (m).
- Step 3.** Select a nearest value (f_{nh}) with f_h from n class of F_h , and the nearest value (f_{nmb}) with f_h from each classification of F_h in addition to n class, where $f_{nmb} = (f_{nm1}, f_{nm2}, \dots, f_{nm(n-1)}, f_{nm(n+1)}, \dots, f_{nmm})$.

Step 4. Calculate the correlation (R_{fF}) with formula is shown as (22),

$$R_{fF} = W_i - \left(\frac{1}{km}\right) \text{diff}(f_h, f_{nh}) + \left(\frac{p(mn)}{1-p(n)}\right) \left(\frac{1}{km}\right) \times \text{diff}(f_h, f_{nmb}) \quad (22)$$

where $\text{diff}(f_h, f_{nh})$ is the distance between f_h and f_{nh} , $\text{diff}(f_h, f_{nmb})$ is the summation of distance between f_h and f_{nmb} .

Step 5. If $j \geq s$, select j with Formula (23), and W_i with Formula (24). Go back to **Step 3**.

$$j = j + 1 \quad (23)$$

$$W_i = R_{fF} \quad (24)$$

Step 6. If $i \geq k$, select i with Formula (25), and W_i with Formula (24). Go back to **Step 2**.

$$i = i + 1 \quad (25)$$

Step 7. Get the correlation (R_{fF}) between feature F_h and classification of sampling ($C_{a,b}$).

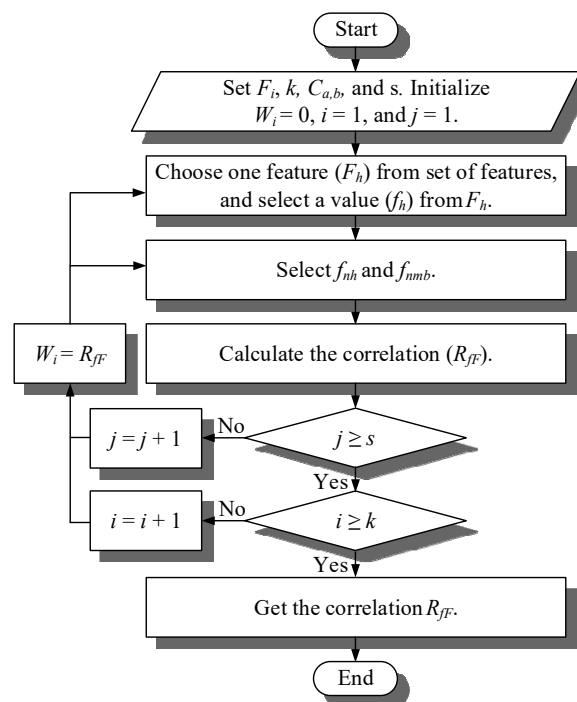


Figure 8. Flowchart of the ReliefF.

3.3. CFS

This study uses CFS to select the feature. CFS was proposed by M. A. Hall. CFS calculates the merit value based on the correlation between features (R_f), and the correlation between feature and classification (R_{fF}). CFS selects the useful features for increased accuracy and reduces the number of features and runtime. In M. A. Hall's paper, CFS was designed for binary classification problems, and used Relief to calculate R_f and R_{fF} . As the study is discussing multiclass classification in this paper, we therefore, use Relief to calculate R_f , and ReliefF to calculate R_{fF} . The steps of CFS are listed as follows, and the flowchart of CFS is shown in Figure 9.

Step 1. Set the target of feature $T = \{ \}$, set of feature $F_i = \{F_1, F_2, \dots, F_n\}$, correlation between features (R_f), and correlation between feature and classification (R_{fF}) where, R_{fij} is the correlation between feature F_i and feature F_j , R_{fFi} is the correlation between feature F_i and classification.

$$R_f = \begin{pmatrix} 1 & R_{f12} & R_{f13} & \dots & \dots & R_{f1n} \\ 0 & 1 & R_{f23} & \dots & \dots & \vdots \\ \vdots & 0 & \ddots & \dots & R_{fij} & \vdots \\ \vdots & 0 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \vdots & 0 & 1 & \vdots \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}, R_{fF} = \begin{pmatrix} R_{fF1} \\ \vdots \\ R_{fFi} \\ \vdots \\ R_{fFn} \end{pmatrix}$$

Step 2. Calculate the merit value (F_{Mi}) based on R_f and R_{fF} with Formula (26) of each feature. $F_{Mi} = \{F_{M1}, F_{M2}, \dots, F_{Mn}\}$.

$$\text{Merit} = \frac{k \times \bar{R}_{fFi}}{\sqrt{k + k(k-1) \times \bar{R}_{fij}}} \tag{26}$$

where, k is number of features (at this step, $k = 1$), \bar{R}_{fFi} is the average of R_{fij} , \bar{R}_{fFi} is the average R_{fFi} (at this step, $\bar{R}_{fFi} = R_{fFi}$).

Step 3. Choose the feature (F_j) that have maximum merit value, set it into T , $T = \{F_j\}$, and $F_i \notin F_j$, $F_i = \{F_1, F_2, \dots, F_{(j-1)}, F_{(j+1)}, \dots, F_n\}$.

Step 4. Calculate the merit value between T and F_i with Formula (26).

Step 5. If $F_i \neq \emptyset$, go back to **Step 3**.

Step 6. Get the merit value of each number of features.

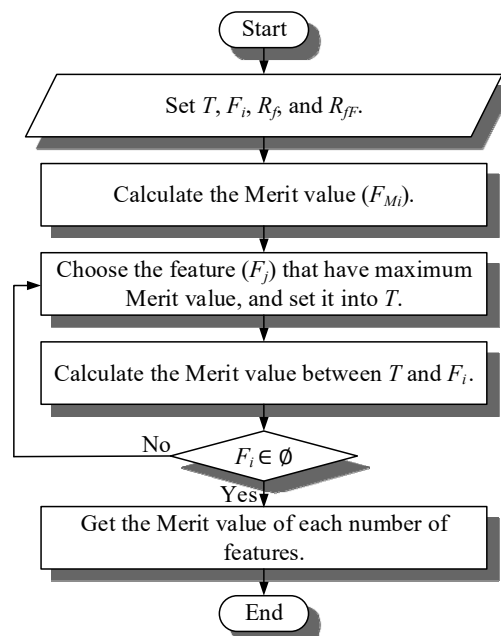


Figure 9. Flowchart of the correlation-based feature selection (CFS).

3.4. CFFS

CFFS is a feature selection approach, improved with CFS. Use this approach to select critical features from all of the features. This study uses CFFS to reduce the number of features of the classification system. Calculate the merit value based on the correlation between features (R_f), the correlation

between feature and classification (R_{fi}), fitness value W_{fi} of each feature, and the merit_new value with Formula (27).

$$\text{Merit_new} = \text{Merit} \times W_{fi} \quad (27)$$

The fitness value was calculated by particle swarm optimization (PSO). PSO is the optimize algorithm that can optimize the coefficients to establish the best neural networks. This study use PSO to optimize the weights of features [31,32], and record the swarm's best-known solution of PSO. We chose the number of features and weights of features after training ANN. The steps of PSO are listed as follows, and the flowchart of PSO is shown in Figure 10a.

Step 1. Set number of particles (p), number of iterations (k), maximum number of iteration (k_{max}), inertia weight (w), acceleration constants (c_1, c_2), particle's position in space (X_p), particle's velocity (V_p), particle's best-known solution (P_{bestp}), and swarm's best-known solution (G_{best}). Particle's position in space is means weights of features $X_p = (X_{p1}, X_{p2}, \dots, X_{pj})$, particle's velocity $V_p = (V_{p1}, V_{p2}, \dots, V_{pj})$, where j is number of features.

Step 2. Calculate the fitness value (W_{fi}) of each feature. Fitness value is particle's accuracy of ANN.

Step 3. Make sure $P_{best_new} > P_{best_old}$, replace P_{best_old} with P_{best_new} as (28), if not, go to **Step 5**. P_{best_old} is particle's best-known solution before fixed, P_{best_new} is particle's best-known solution after fixed, particle's best-known solution $P_{best} = (P_{best1}, P_{best2}, \dots, P_{bestp})$.

$$P_{best_old} = P_{best_new} \quad (28)$$

Step 4. Make sure $G_{best_new} > G_{best_old}$, replace G_{best_old} with G_{best_new} as (29), if not, go to **Step 5**. G_{best_new} is swarm's best-known solution before fixed, G_{best_new} is swarm's best-known solution after fixed, swarm's best-known solution is the best solution of particle's best-known solution.

$$G_{best_old} = G_{best_new} \quad (29)$$

Step 5. Fix the particle's velocity as (30) and (31).

$$V_{p_new} = w \times V_p + c1r1(P_{best} - X_p) + c2r2 \times (G_{best} - X_p) \quad (30)$$

$$V_p = V_{p_new} \quad (31)$$

Step 6. Fix the particle's position in space as (32) and (33).

$$X_{i_new} = X_i + V_p \quad (32)$$

$$X_i = X_{i_new} \quad (33)$$

Step 7. Make sure $k = k_{max}$, if not, calculates k as (34) and back to **Step 2**.

$$k = k + 1 \quad (34)$$

Step 8. Get the swarm's best-known solution for ANN.

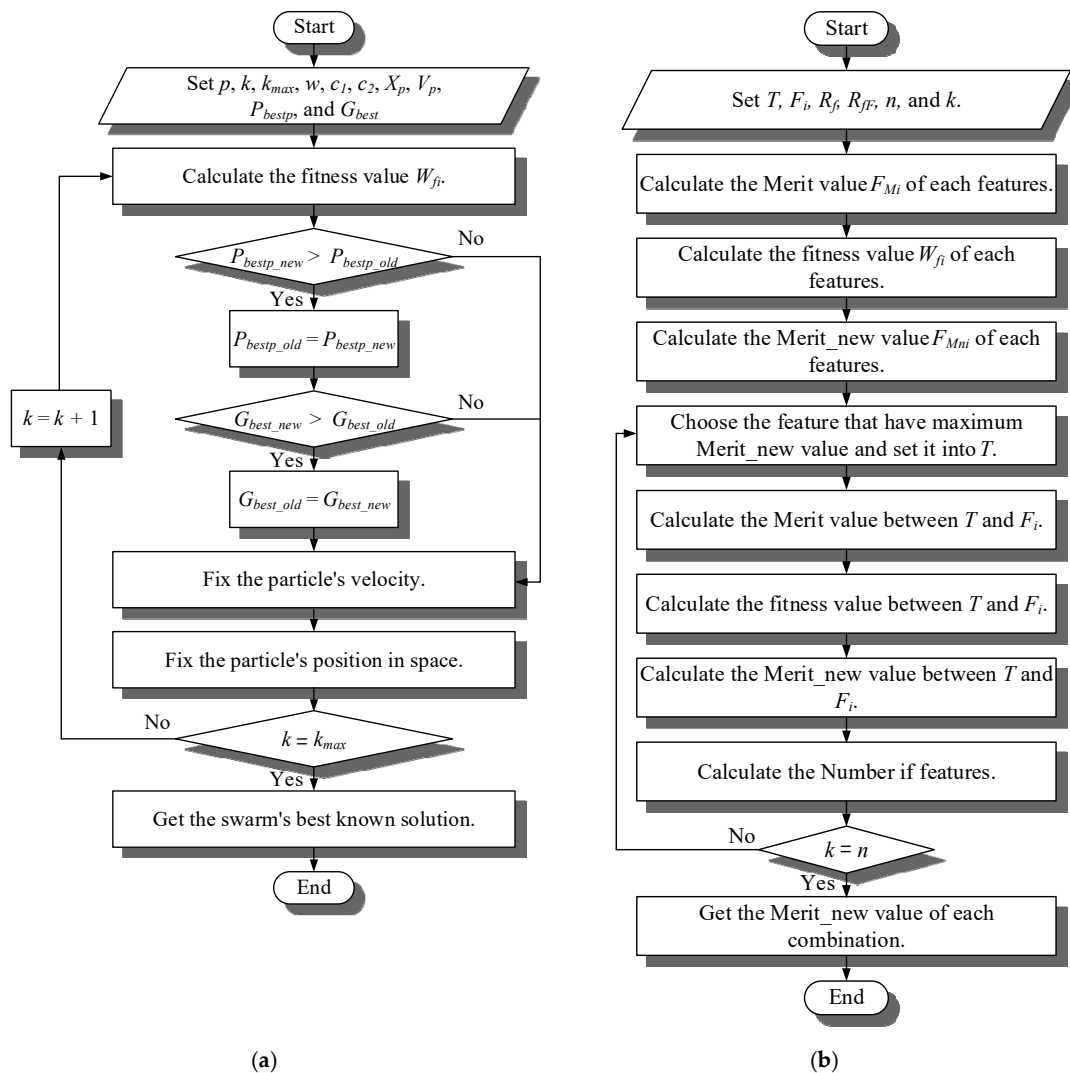


Figure 10. Flowchart of the correlation and fitness value-based feature selection (CFFS). (a) Calculates the fitness value; (b) calculates the merit_new.

Formula (27) calculates the merit_new value with the fitness value from PSO to select the features that are useful to the classification system and get the best accuracy. The steps of CFFS are listed as follows, and the flowchart of CFFS is shown in Figure 10b.

Step 1. Set the target of feature $T = \{ \}$, set of feature $F_i = \{F_1, F_2, \dots, F_n\}$, correlation between features (R_f), correlation between feature and classification (R_{fF}), number of all features (n), set of feature $F_i = \{F_1, F_2, \dots, F_n\}$, and number of features $k = 1$.

$$R_f = \begin{pmatrix} 1 & R_{f12} & R_{f13} & \dots & \dots & R_{f1n} \\ 0 & 1 & R_{f23} & \dots & \dots & \vdots \\ \vdots & 0 & \ddots & \dots & R_{fij} & \vdots \\ \vdots & 0 & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \vdots & 0 & 1 & \vdots \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}, R_{fF} = \begin{pmatrix} R_{fF1} \\ \vdots \\ R_{fFi} \\ \vdots \\ R_{fFn} \end{pmatrix}$$

where, R_{fij} is the correlation between feature F_i and feature F_j , R_{fFi} is the correlation between feature F_i and classification. Calculate the merit value (F_{Mi}) of each features, based on R_f and

R_{fF} with Formulas (3)–(8), $F_{Mi} = \{F_{M1}, F_{M2}, \dots, F_{Mn}\}$, where, k is number of features (at this step, $k = 1$), \bar{R}_{fFi} is the average of R_{fFij} , \bar{R}_{fFi} is the average R_{fFi} (at this step, $\bar{R}_{fFi} = R_{fFi}$).

- Step 2.** Calculate the fitness value (W_{fi}) of each feature. Fitness value is the accuracy that calculated of artificial neural network. $W_{fi} = (W_{f1}, W_{f2}, \dots, W_{fn})$.
- Step 3.** Calculate the merit_new value (F_{Mni}) with Formula (27) of each feature. $F_{Mni} = (F_{Mn1}, F_{Mn2}, \dots, F_{Mnn})$.
- Step 4.** Choose the feature (F_j) that have maximum merit_new value, and set it into T , $T = \{F_j\}$. $F_i \notin F_j$, $F_i = \{F_1, F_2, \dots, F_{(j-1)}, F_{(j+1)}, \dots, F_n\}$.
- Step 5.** Calculate the merit value as Formulas (3)–(8) between T and F_i . $F_{Mi} = F_{Mi} = \{F_{M1}, F_{M2}, \dots, F_{M(j-1)}, F_{M(j+1)}, \dots, F_{Mn}\}$.
- Step 6.** Calculate the fitness value (W_{fi}) with PSO optimizes the weights of features between T and F_i . $W_{fi} = (W_{f1}, W_{f2}, \dots, W_{f(j-1)}, W_{f(j+1)}, \dots, W_{fn})$.
- Step 7.** Calculate the merit_new value as Formula (27) between T and F_i . Choose the feature (F_x) that have maximum merit_new value, and set it into T .
- Step 8.** Calculate the number of features as (35).

$$k = k + 1 \quad (35)$$

Step 9. Make sure $k = n$, if not, back to **Step 6**.

Step 10. Get the merit_new value of each combination of number of features.

4. Measurement and Analysis of the Current Signal of the Induction Motor

The specification of the induction motor is shown in Table 1. The events of motor failure are shown in Figure 11a–c. (1) Damage of shaft output of bearing. Electro-discharge machining at inner bearing of motor and make hole about 1.96 mm in length and 0.53 mm in width. Simulate the damage caused by foreign material. (2) Layer short, make electrical breakdown at five coils. Simulate the damage caused by the insulation problem. (3) Broken rotor bar, made two holes: about 8 mm in diameter and 10 mm deep. Simulate the damage caused by external force or high temperature.

Table 1. Specification of the induction motor.

Induction Motor			
Rated voltage	220 V	Power frequency	60 Hz
Output	2 Hp	Rated speed	1764 rpm
Number of poles	4 poles	Power factor	0.8

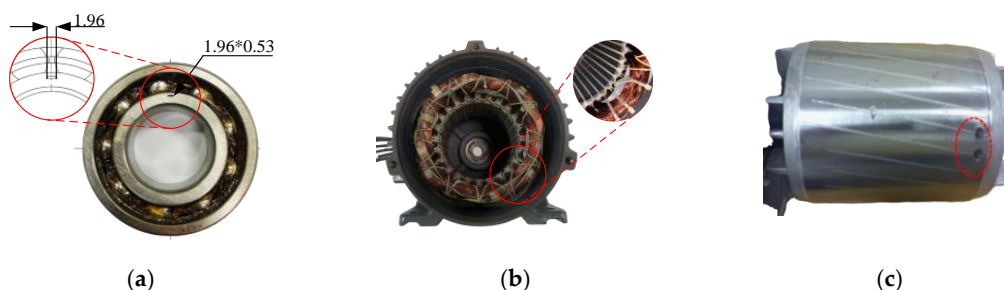


Figure 11. Motor failure sample. (a) Bearing of motor (length 1.96 mm and width 0.53 mm); (b) Coils of stator (five coils); (c) Rotor bar (two holes, diameter: 8 mm and depth: 10 mm).

This study used power supply for the induction motor with three phases, 220 V, and alternating current (AC) power that current would reverse periodically. Control panel sets the AC servo motor to simulate the load to the induction motor. Finally, we used NI PXI-1033 to capture the current signal of induction motor at one of the three phases. Time sampling for 2 s and sampling frequency for 1000 Hz of each observation, so that each observation has 2000 data points. In addition, the observations database is also publicly posted in this link (<https://reurl.cc/gmmWgN>) in the form of a Matlab file. Equipment layout is shown in Figure 12. The steps of experiment are listed as the following:

- Step 1.** Prepare all the samples of the motor. This study prepares four classes of motor: normal, damage of shaft output of bearing, layer short, and broken rotor bar.
- Step 2.** Connect the three phases R, S, and T of motor with AC power supply. R, S, and T are same AC power but have different angle.
- Step 3.** Setup the NI PXI-1033 and computer and choose one of the three phases to measure current signal.
- Step 4.** Use AC power supply to input the power to the motor.
- Step 5.** Use control panel to select torque of AC servo motor to simulate the load of the motor. This study selected the half load for the motor.
- Step 6.** Use Labview to record data from NI PXI-1033.
- Step 7.** Use Matlab to analyze the current signal of each motor.

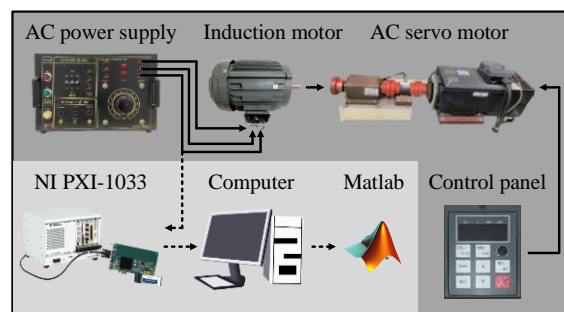


Figure 12. Equipment layout to capture the current signal of induction motor.

Current signal of induction motor processing with MRA and standardization. **d1**, **d2**, **d3**, **d4**, **d5**, and **a5** were decomposed from current signal of normal motor as shown in Figure 13a–d, figures are shown for one of all observations of each motor class: normal, bearing damage, layer short circuit damage, and rotor bar damage. Comparison of the MRA of the normal motor and the other motors: (1) The damage of shaft output of bearing: **d1** and **d4** are smaller than that of the normal motor. **d2**, **d3**, **d5**, and **a5** are similar to the normal motor. (2) The layer short: **d1** is smaller than that of the normal motor. **d2**, **d3**, **d4**, **d5**, and **a5** are similar to the normal motor. (3) The broken rotor bar: **d1** is smaller than that of the normal motor. **d2** is larger than that of the normal motor. **d3**, **d4**, **d5**, and **a5** are similar to the normal motor.

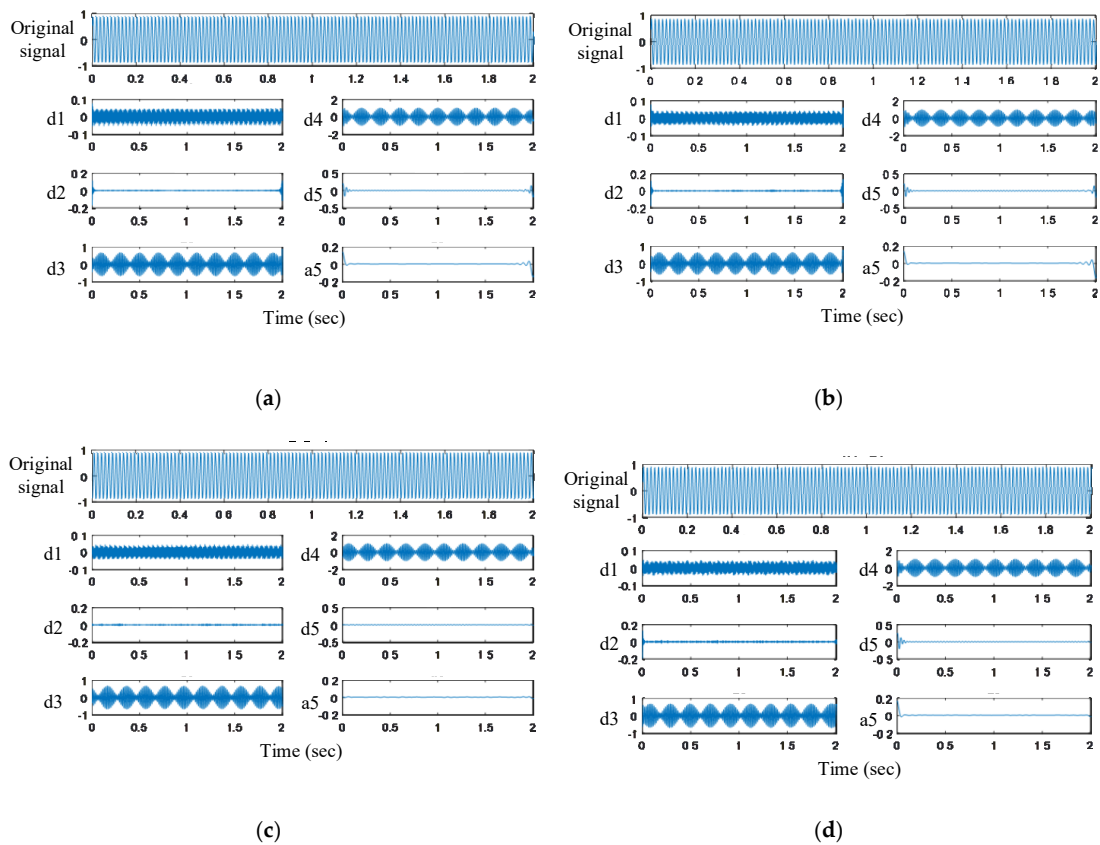


Figure 13. (a) MRA of normal motor; (b) MRA of bearing failure; (c) MRA of stator winding failure; (d) MRA of rotor failure.

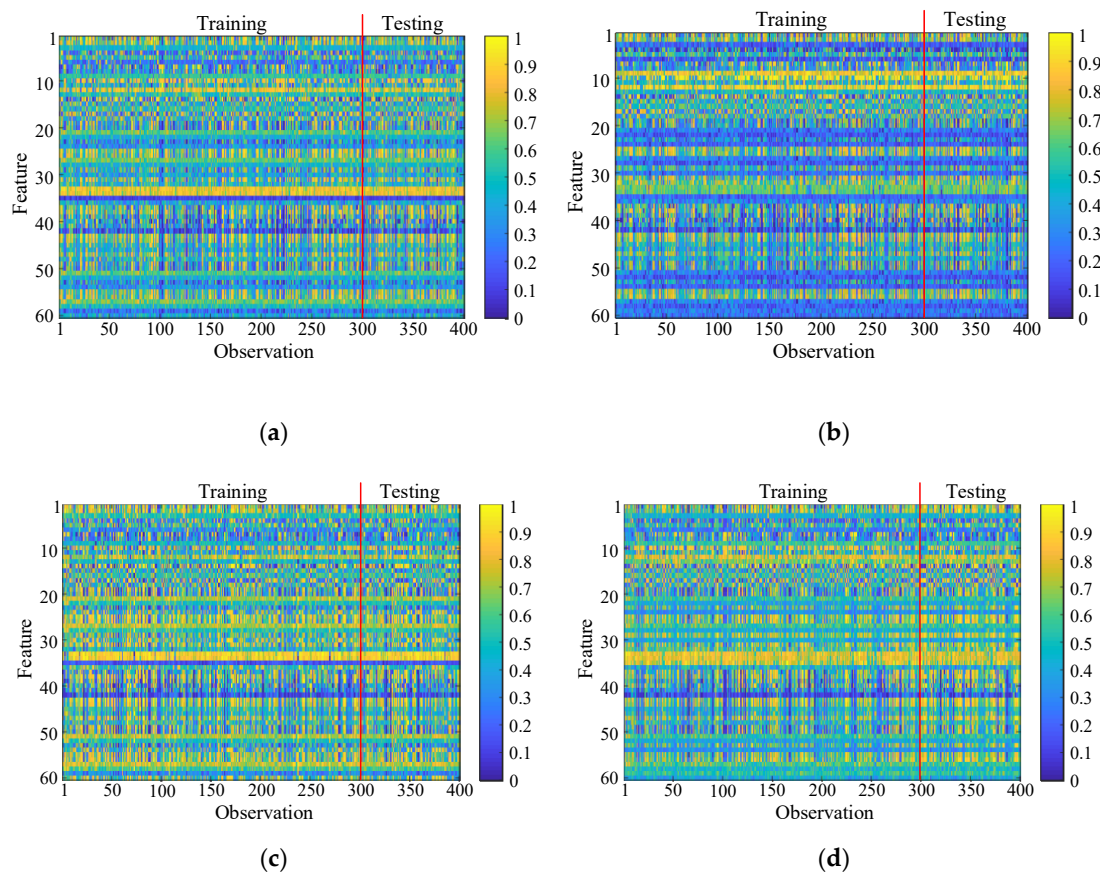
5. Classification of Induction Motor Failure

5.1. Motorcurrent Signal Analysis of Induction Motor Failure

The study used NI PXI-1033 to capture 400 observations of each motor, used normalization and MRA to process the observations and captured the features on Matlab, and selected critical features by ReliefF, CFS, and CFFS. Finally, trained the ANN with 300 observations, and 100 observations for testing, and calculated the accuracy of this neural network. MRA decomposes the signal to level five, using detail coefficients of level one to five (**d1** to **d5**) and approximation coefficients of level five (**a5**) to compose the signal and extract the features, as shown in Table 2, maximum value (T_{max}), minimum value (T_{min}), average value (T_{mean}), root mean square value (T_{mse}), standard value (T_{std}) of time domain and maximum value (F_{max}), minimum value (F_{min}), average value (F_{mean}), root mean square value (F_{mse}), and standard value (F_{std}) of frequency domain. Feature distribution of induction motor, as shown in Figure 14a–d, x -axis is for samples, y -axis is for features. In the figures, some slightly different features and compared bearing damage, layer short circuit damage, and rotor bar damage with normal motor are observed. Compared the feature distribution of MRA of normal motor and the other motors: (1) The damage of shaft output of bearing: F3, F9, F21, F22, F24, F27, F28, F30, F33, F35, F34, F51, F52, F54, F57, F58, F59, and F60. F24, F30, F54, F59, and F60 were extracted from **d1**. F35 was extracted from **d2**. F22, F28, F34, F52, F58, and F59 were extracted from **d3**. F3, F9, F21, F27, F33, F51, and F57 were extracted from **d4**. (2) The layer short: F9, F21, F22, F24, F27, F30, F51, F57, and F60. F24, F30, and F60 were extracted from **d1**. F9, F21, F27, F51, and F57 were extracted from **d4**. (3) The broken rotor bar: F13, F23, F29, F32, F35, F53, F59, and F60. F24, F30, and F60 were extracted from **d1**. F23, F29, F35, F53, and F59 were extracted from **d2**. F32 was extracted from **d5**. F13 was extracted from **a5**.

Table 2. Feature Extraction of MRA.

	a5	d5	d4	d3	d2	d1
<i>Tmax</i>	F1	F2	F3	F4	F5	F6
<i>Tmin</i>	F7	F8	F9	F10	F11	F12
<i>Tmean</i>	F13	F14	F15	F16	F17	F18
<i>Tmse</i>	F19	F20	F21	F22	F23	F24
<i>Tstd</i>	F25	F26	F27	F28	F29	F30
<i>Fmax</i>	F31	F32	F33	F34	F35	F36
<i>Fmin</i>	F37	F38	F39	F40	F41	F42
<i>Fmean</i>	F43	F44	F45	F46	F47	F48
<i>Fmse</i>	F49	F50	F51	F52	F53	F54
<i>Fstd</i>	F55	F56	F57	F58	F59	F60

**Figure 14.** (a) Feature distribution of normal motor; (b) feature distribution of bearing failure; (c) feature distribution of stator winding failure (d) feature distribution of rotor failure.

5.2. Accuracy of Classifier

Feature selection approaches: ReliefF, CFS, and CFFS have different feature selection criterion that result in different consequences for the approaches. ReliefF calculates the correlation between feature and classes and selects features in correlation numerical order. CFS calculates the merit value with Relief and ReliefF, and selects features in merit numerical order. CFFS calculates the merit_new value with merit value and weights, selects features in merit_new numerical order. The critical features and the features combination are selected with the least amount of features, and the accuracy can be improved by using the proposed feature selection method.

(1) ReliefF:

According to Table 3, the results of ReliefF indicate that the critical features at 38, 39, 40, and 45 number of features increase accuracy with a sharp rise: between 37 and 38 number of features, a 9.04% (from 26.81% to 35.85%) increase can be observed at F13; between 38 and 39 number of features, a 8.12% (from 35.85% to 43.97%) increase can be observed at F29; between 39 and 40 number of features, a 19.38% (from 43.97% to 63.35%) increase can be observed at F9; between 42 to 43 number of features, a 6.01% (from 65.44% to 71.45%) increase can be observed at F3.

(2) CFS:

Table 3. Result of ReliefF.

Feature Numbers	Accuracy	The Elements of the Feature Vector
37	26.81%	F1, F2, F4, ..., F8, F10, F11, F12, F14, ..., F20, F25, F26, F31, F32, F37, ..., F50, F55, F56
38	35.85%	F1, F2, F4, ..., F8, F10, F11, F12, F13, F14, ..., F20, F25, F26, F31, F32, F37, ..., F50, F55, F56
39	43.97%	F1, F2, F4, ..., F8, F10, F11, F12, F13, F14, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F55, F56
40	63.35%	F1, F2, F4, ..., F8, F9, F10, F11, F12, F13, F14, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F55, F56
42	65.44%	F1, F2, F4, ..., F20, F23, F25, F26, F29, F31, F32, F37, ..., F50, F53, F55, F56
43	71.45%	F1, F2, F3, F4, ..., F20, F23, F25, F26, F29, F31, F32, F37, ..., F50, F53, F55, F56

According to Table 4, the results of CFS indicate that the critical features at 20, 32, 34, and 35 number of features increase accuracy with a sharp rise: between 19 and 20 number of features, a 7.27% (from 30.58% to 37.85%) increase can be observed at F32; between 31 and 32 number of features, a 14.78% (from 36.78% to 51.56%) increase can be observed at F29; between 33 and 34 number of features, a 6.18% (from 57.7% to 60.58%) increase can be observed at F59; between 34 to 35 number of features, a 14.88% (from 60.58% to 75.46%) increase can be observed at F52.

(3) CFFS:

Table 4. Result of CFS.

Feature Numbers	Accuracy	The Elements of the Feature Vector
19	30.58%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F31, F38, ..., F41, F47, F49, F50
20	37.85%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F31, F32, F38, ..., F41, F47, F49, F50
31	36.78%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F25, F26, F31, F32, F37, ..., F50, F55, F56
32	51.56%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F55, F56
33	54.4%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F53, F55, F56
34	60.58%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F53, F55, F56, F59
35	75.46%	F1, F2, F8, F10, F11, F14, F15, F17, ..., F20, F25, F26, F29, F31, F32, F37, ..., F50, F52, F53, F55, F56, F59

According to Table 5, the results of CFFS indicate that the critical features at 1, 2, 3, and 5 number of features increase accuracy with a sharp rise: F6 has 39.5% accuracy at 1 number of features; between

1 and 2 number of features, a 16% (from 39.5% to 55.5%) increase can be observed at F34; between 2 and 3 number of features, a 28.75% (from 55.5% to 84.25%) increase can be observed at F57; between 4 and 5 number of features, a 5.5% (from 84.25% to 90%) increase can be observed at F35.

Table 5. Result of CFFS.

Feature Numbers	Accuracy	The Elements of the Feature Vector
1	39.5%	F6
2	55.5%	F6, F34
3	84.25%	F6, F34, F57
4	84.5%	F6, F34, F52, F57
5	90%	F6, F34, F35, F52, F57

The accuracy of ReliefF, CFS, and CFFS are shown in Figure 15. ReliefF accuracy curve is black wire (—), the number of features have best effect at the red star mark (★), the number of features have best accuracy at the red square mark (■). CFS accuracy curve is blue wire (—), the number of features have best effect at the purple star mark (★), the number of features have best accuracy at the purple square mark (■). CFFS accuracy curve is green wire (—), the number of features have best effect at the yellow star mark (★), the number of features have best accuracy at the yellow square mark (■).

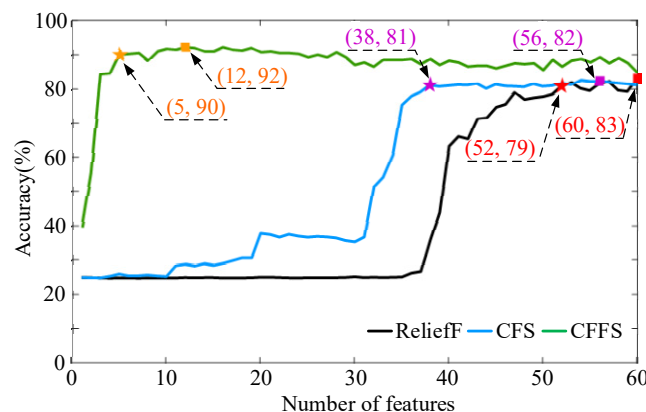


Figure 15. Accuracy curve of ReliefF, CFS, and CFFS.

The ReliefF accuracy curve becomes stable at a point of 52 number of features (accuracy: 79.27%) and reaches its maximum point at 60 number of features (accuracy: 83.03%). The CFS accuracy curve becomes stable at a point of 38 number of features (accuracy: 81.21%) and reaches its maximum point at 56 number of features (accuracy: 82.31%). The CFFS accuracy curve becomes stable at a point of five number of features (accuracy: 90%) and reaches its maximum point at 12 number of features (accuracy: 92.25%).

6. Conclusions

This study uses MRA to analyze current signal, extract the features, and uses feature selection approaches to select critical features for induction motor classification to establish the fault detection system with ANN. ReliefF, CFS, and CFFS are different feature selection approaches, and these approaches could reduce the number of features and the system operation cost while maintaining good accuracy. The results of the study show that ReliefF, CFS, and CFFS have better efficiency than the unused feature selection approach, and in compared to the unused feature selection approach (number of features: 60, accuracy: 83.03%) the accuracy of ReliefF (number of features: 52, accuracy: 79.27%) is lower 0.76%, reduces 13.3% number of features. The accuracy of CFS (number of features:

38, accuracy: 81.21%) is lower 1.82%, reduces 36.7% number of features. The accuracy of CFFS (number of features: 5, accuracy: 90%) is higher 6.97%, reduces 90% number of features. Therefore, the CFFS has the best efficiency of the three feature selection approaches, as it reduces most of the features, and also improves the accuracy of the induction motor fault detection system.

Author Contributions: Conceptualization, C.-Y.L.; Methodology, C.-Y.L. and M.-S.W.; Software, C.-Y.L. and M.-S.W.; Validation, C.-Y.L.; Formal Analysis, C.-Y.L. and M.-S.W.; Investigation, C.-Y.L.; Resources, C.-Y.L.; Data Curation, C.-Y.L. and M.-S.W.; Writing-Original Draft Preparation, C.-Y.L. and M.-S.W.; Writing-Review & Editing, C.-Y.L. and M.-S.W.; Visualization, C.-Y.L. and M.-S.W.; Supervision, C.-Y.L.; Project Administration, C.-Y.L.; Funding Acquisition, C.-Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Zhang, P.Y.; Shu, S.; Zhou, M. An online fault detection model and strategies based on SVM-grid in clouds. *IEEE/CAA J. Autom. Sin.* **2018**, *5*, 445–456. [[CrossRef](#)]
- Wang, H.; Lu, S.; Qian, G.; Ding, J.; Liu, Y.; Wang, Q. A two-step strategy for online fault detection of high-resistance connection in BLDC motor. *IEEE Trans. Power Electron.* **2020**, *35*, 3043–3053. [[CrossRef](#)]
- Mao, W.; Chen, J.; Liang, X.; Zhang, X. A new online detection approach for rolling bearing incipient fault via self-adaptive deep feature matching. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 443–456. [[CrossRef](#)]
- Bazurto, A.J.; Quispe, E.C.; Mendoza, R.C. Causes and failures classification of industrial electric motor. In Proceedings of the 2016 IEEE ANDESCON, Arequipa, Peru, 19–21 October 2016.
- Kral, C.; Habetler, T.G.; Harley, R.G.; Pirker, F.; Pascoli, G.; Oberguggenb, H.; Fenz, C.J.M. A comparison of rotor fault detection techniques with respect to the assessment of fault severity. In Proceedings of the 4th IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives, Atlanta, GA, USA, 24–26 August 2003; pp. 265–269.
- Liu, W.; Liao, Q.; Qiao, F.; Xia, W.; Wang, C.; Lombardi, F. Approximate designs for fast Fourier transform (FFT) with application to speech recognition. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2019**, *66*, 4727–4739. [[CrossRef](#)]
- Pei, S.C.; Ding, J.J. Relations between Gabor transforms and Fractional Fourier Transforms and their applications for signal processing. *IEEE Trans. Signal Process.* **2007**, *55*, 4839–4850. [[CrossRef](#)]
- Cho, S.H.; Jangand, G.; Kwon, S.H. Time-frequency analysis of power-quality disturbances via the Gabor–Wigner transform. *IEEE Trans. Power Deliv.* **2010**, *25*, 494–499.
- Laurence, C.; Pierre-Francois, M.; Lise, D.; Kim, P.G.; Francis, G.; Sá, R.C. A method for the analysis of respiratory sinus arrhythmia using continuous wavelet transforms. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1640–1642.
- Ghunem, R.A.; Jayaram, S.H.; Cherney, E.A. Investigation into the eroding dry-band arcing of filled silicone rubber under DC using wavelet-based multiresolution analysis. *IEEE Trans. Dielectr. Electr. Insul.* **2014**, *21*, 713–720. [[CrossRef](#)]
- Kang, M.; Kim, J.; Wills, L.M.; Kim, J.M. Time-varying and multiresolution envelope analysis and discriminative feature analysis for bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2015**, *62*, 7749–7761. [[CrossRef](#)]
- Ananthan, S.N.; Padmanabhan, R.; Meyur, R.; Mallikarjuna, B.; Reddy, M.J.B.; Mohanta, D.K. Real-time fault analysis of transmission lines using wavelet multi-resolution analysis based frequency-domain approach. *IET Sci. Meas. Technol.* **2016**, *10*, 693–703. [[CrossRef](#)]
- Biscaro, A.A.P.; Pereira, R.A.F.; Kezunovic, M.; Mantovani, J.R.S. Integrated fault location and power-quality analysis in electric power distribution systems. *IEEE Trans. Power Deliv.* **2016**, *31*, 428–436.
- Al-Otaibi, R.; Jin, N.; Wilcox, T.; Flach, P. Feature construction and calibration for clustering daily load curves from smart-meter data. *IEEE Trans. Ind. Inform.* **2016**, *12*, 645–654. [[CrossRef](#)]
- Imani, M.; Ghassemian, H. Band clustering-based feature extraction for classification of hyperspectral images using limited training samples. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1325–1329. [[CrossRef](#)]
- Neshatian, K.; Zhang, M.; Andrae, P. A filter approach to multiple feature construction for symbolic learning classifiers using genetic programming. *IEEE Trans. Evolut. Comput.* **2012**, *16*, 645–661. [[CrossRef](#)]

17. Panigrahy, P.S.; Santra, D.; Chattopadhyay, P. Feature engineering in fault diagnosis of induction motor. In Proceedings of the 2017 3rd International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), Rupnagar, India, 16–18 November 2017.
18. Godse, R.; Bhat, S. Mathematical morphology-based feature-extraction technique for detection and classification of faults on power transmission line. *IEEE Access* **2020**, *8*, 38459–38471. [[CrossRef](#)]
19. Rauber, T.W.; de Assis Boldt, F.; Varejão, F.M. Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2014**, *62*, 637–646. [[CrossRef](#)]
20. Fu, R.; Wang, P.; Gao, Y.; Hua, X. A new feature selection method based on Relief and SVM-RFE. In Proceedings of the 2014 12th International Conference on Signal Processing (ICSP), Hangzhou, China, 19–23 October 2014.
21. Fu, R.; Wang, P.; Gao, Y.; Hua, X. A combination of Relief feature selection and fuzzy k-nearest neighbor for plant species identification. In Proceedings of the 2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, Indonesia, 15–16 October 2016.
22. Huang, Z.; Yang, C.; Zhou, X.; Huang, T. A hybrid feature selection method based on binary state transition algorithm and ReliefF. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 1888–1898. [[CrossRef](#)]
23. Hu, B.; Li, X.; Sun, S.; Ratcliffe, M. Attention recognition in EEG-based affective learning research using CFS+knn algorithm. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *15*, 38–45. [[CrossRef](#)]
24. Singh, A.; Thakur, N.; Sharma, A. A review of supervised machine learning algorithms. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016.
25. Bao, F.; Maier, T. Stochastic gradient descent algorithm for stochastic optimization in solving analytic continuation problems. *Am. Inst. Math. Sci. Found. Data Sci.* **2020**, *2*, 1–17. [[CrossRef](#)]
26. Wu, W.; Feng, G.R.; Li, Z.X. Deterministic convergence of an online gradient method for BP neural networks. *IEEE Trans. Neural Netw.* **2005**, *16*, 533–540. [[CrossRef](#)]
27. Park, D.J.; Jun, B.E.; Kim, J.H. Novel training algorithm for multilayer feedforward neural network. *Electron. Lett.* **1992**, *28*, 543–544. [[CrossRef](#)]
28. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning internal representations by error backpropagation. In *Parallel Distributed Processing*; Rumelhart, D.E., McClelland, J.L., Eds.; Stanford University: Stanford, CA, USA, 1986; Volume 1, pp. 319–362.
29. Popoola, S.I.; Jefia, A.; Atayero, A.A.; Kingsley, O.; Faruk, N.; Oseni, O.F.; Abolade, R.O. Determination of neural network parameters for path loss prediction in very high frequency wireless channel. *IEEE Access* **2019**, *7*, 150462–150483. [[CrossRef](#)]
30. Sangari, A.; Sethares, W. Convergence analysis of two loss functions in soft-max regression. *IEEE Trans. Signal Process.* **2016**, *64*, 1280–1288. [[CrossRef](#)]
31. Lee, C.Y.; Tuegeh, M. An optimal solution for smooth and non-smooth cost functions-based economic dispatch problem. *Energies* **2020**, *13*, 3721. [[CrossRef](#)]
32. Juan Luis, F.M.; Esperanza, G.G. Stochastic stability analysis of the linear continuous and discrete PSO models. *IEEE Trans. Evolut. Comput.* **2011**, *15*, 405–423.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).