



# High-impedance fault detection in medium-voltage distribution network using computational intelligence-based classifiers

Veerapandiyan Veerasamy<sup>1</sup> · Noor Izzri Abdul Wahab<sup>1</sup> · Rajeswari Ramachandran<sup>3</sup> · Mariammal Thirumeni<sup>2</sup> · Chitra Subramanian<sup>3</sup> · Mohammad Lutfi Othman<sup>1</sup> · Hashim Hizam<sup>1</sup>

Received: 31 March 2017 / Accepted: 12 August 2019 / Published online: 20 August 2019  
© Springer-Verlag London Ltd., part of Springer Nature 2019

## Abstract

This paper presents the high-impedance fault (HIF) detection and identification in medium-voltage distribution network of 13.8 kV using discrete wavelet transform (DWT) and intelligence classifiers such as adaptive neuro-fuzzy inference system (ANFIS) and support vector machine (SVM). The three-phase feeder network is modelled in MATLAB/Simulink to obtain the fault current signal of the feeder. The acquired fault current signal for various types of faults such as three-phase fault, line to line, line to ground, double line to ground and HIF is sampled using 1st, 2nd, 3rd, 4th and 5th level of detailed coefficients and approximated by DWT analysis to extract the feature, namely standard deviation (SD) values, considering the time-varying fault impedance. The SD values drawn by DWT technique have been used to train the computational intelligence-based classifiers such as fuzzy, Bayes, multi-layer perceptron neural network, ANFIS and SVM. The performance indices such as mean absolute error, root mean square error, kappa statistic, success rate and discrimination rate are compared for various classifiers presented. The results showed that the proffered ANFIS and SVM classifiers are more effective and their performance is substantially superior than other classifiers.

**Keywords** High-impedance fault · Discrete wavelet transform · Adaptive neuro-fuzzy inference system · Support vector machine · Multi-layer perceptron neural network (MLP) · Bayes and fuzzy classifier

## 1 Introduction

Power system protection is a crucial problem in modern power system for secured and reliable operation of the system. The electrical utilities have been adopting numerous techniques to reduce the number of line outages due to short circuit. In this context, many researches were carried out to detect and classify the type of fault which leads to location and isolation of fault, thereby reducing the outages

in both transmission and distribution system. One of the troublesome abnormalities that occur in the system is HIF. This type of fault occurs mostly in the distributor feeder network, when the energized conductor has in contact with high resistance surface such as tree limb, wet and dry sand, and wet and dry asphalt. It exhibits the asymmetry, intermittence and nonlinear arcing characteristics with the magnitude of fault current of about 0–75 A in an grounded system [1–4]. Normally, the over current relay fails to detect the HIF in the system because of the lower current magnitude, leading to the cascading failure of the system and causing risk to the people and their properties [5]. A mathematical model for designing the HIF model in the form of nonlinear partial differential equation for studying the effect of HIF in distribution system is presented in [6]. In addition, many works were carried out for designing the HIF model [7]. Another approach of live tree-related HIF model is suggested [8], to study the effect of environmental conditions in real time and biological classification using finite element method (FEM). A least square approach with

✉ Veerapandiyan Veerasamy  
veerapandian220@gmail.com

<sup>1</sup> Advanced Lightning, Power and Energy Research (ALPER), Department of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 UPM Serdang, Selangor, Malaysia

<sup>2</sup> Department of Electrical Engineering, Rajalakshmi Engineering College, Chennai 602105, India

<sup>3</sup> Department of Electrical Engineering, Government College of Technology, Coimbatore 641013, India

advanced numerical techniques is used to locate the HIF in the distribution system [9].

Linking with the previously published work, several studies were carried out to discriminate and identify the power system disturbances such as capacitor switching, symmetrical, unsymmetrical HIF and another transient phenomenon. Nevertheless, significant importance is given to detect the HIF that occurs in the system by the utilities. The conventional scheme of detection of single-phase HIF in low-resistance grounded distribution network (LRGDN) with the underlying perception of variation of composite power angle and its magnitude, voltage–current characteristics profiles during normal and fault condition is studied [2, 10]. In some cases, the conventional numerical methods fail to identify the HIF due to selection of improper step size. It can be overcome by the powerful signal processing techniques. Short-time Fourier transform (STFT) based on lower-order harmonics represented by its magnitude and phase was used to identify the HIF from other system disturbances such as capacitor switching and feeder energizing [11]. A mathematical morphology and WT were used to extract the high-frequency component of the three-phase current signal to locate the power system faults from the other transients using the energy index [1, 12]. The detection of HIF using DWT with the mother wavelet of Daubechies wavelet db4 from the signal comprising of another non-fault transient event is presented in [13, 14]. The continuous wavelet transform (CWT) to extract the feature for learning the extreme learning machine (ELM) type of neural network to identify the HIF and current transformer saturation using the high-speed communication techniques for the co-ordination of the relay has also been worked out [15]. Moreover, WT and other time–frequency analysis in combination with pattern classifier, ANN and power line communication techniques were used to locate the fault [16–18]. ANFIS is one of the major trade-offs among ANNs and fuzzy logic systems, offering smoothness due to the fuzzy control interpolation and adaptability due to the ANN for classification problem [19–23]. On the flipside, SVM-based intelligence classifier is more efficient than neural network method because of its classification with few samples and uses the raw data without pre-processing [24–26].

Among the various feature extraction methods, wavelets are gaining the popularity in many works [27–31] due to its capability of reliable feature decomposition in frequency and localization in time even in the presence of noise. Therefore, in this work DWT analysis was used to extract the features and the classifiers such as ANFIS and SVM were used to detect the occurrence of HIF in the system and to discriminate them from other transients.

The prime contribution of this paper is as follows:

**HIF detection** In this paper, the detection of HIF is carried out by analysing the three-phase current signal of the power system using DWT analysis to extract the features such as SD for training the classifiers.

**HIF identification** In this step, the SD features drawn were used to train the computational intelligence-based classifiers with different possibilities of power system disturbances, also considering the time-varying fault resistance to prove the effectiveness of the classifier.

The paper is organized as follows: Sect. 2 presents the system model studied and proposed methodology to detect the HIF in the system. Section 3 deals with the data acquisition of current signal using DWT analysis to extract the features for classification. Section 4 explains the various computational intelligence-based classifiers such as fuzzy, Bayes, MLP, ANFIS and SVM for classifying the fault. Results and discussion is portrayed in Sect. 5. Finally, the conclusion and future scope are made in the last part of the paper.

## 2 System studied and methodology

In order to validate the proffered method of classifier technique, it is necessary to accumulate the field data signals from the power system to train the classifier with different operating conditions of the system. To obtain the field data in real time for power system is quite expensive, so a real-time medium-voltage (MV) distribution feeder network in Spain [17] is considered for validating the proposed classifiers simulated in MATLAB for obtaining the fault current data under different possible disturbances.

### 2.1 Distribution network model

The MV feeder network configured in Spain consists of substation transformer with 5 feeder lines. But in the proposed work only 4-feeder network ( $F_1$ ,  $F_2$ ,  $F_3$  and  $F_4$ ) is considered for the simulation assuming the power system disturbances occur in each of the feeder as represented in Fig. 1.

The system parameters are as follows: a generator with voltage rating of 30 kV coupled with substation transformer with the turns ratio of 30/13.8 kV, 12 MVA rating. Assuming that  $F_1$  is connected to the nonlinear load,  $F_2$  feeding the linear load in which HIF occurs and the balanced and unbalanced fault occurs in the third feeder with the normal condition in  $F_4$  and the model description is discussed in [17].

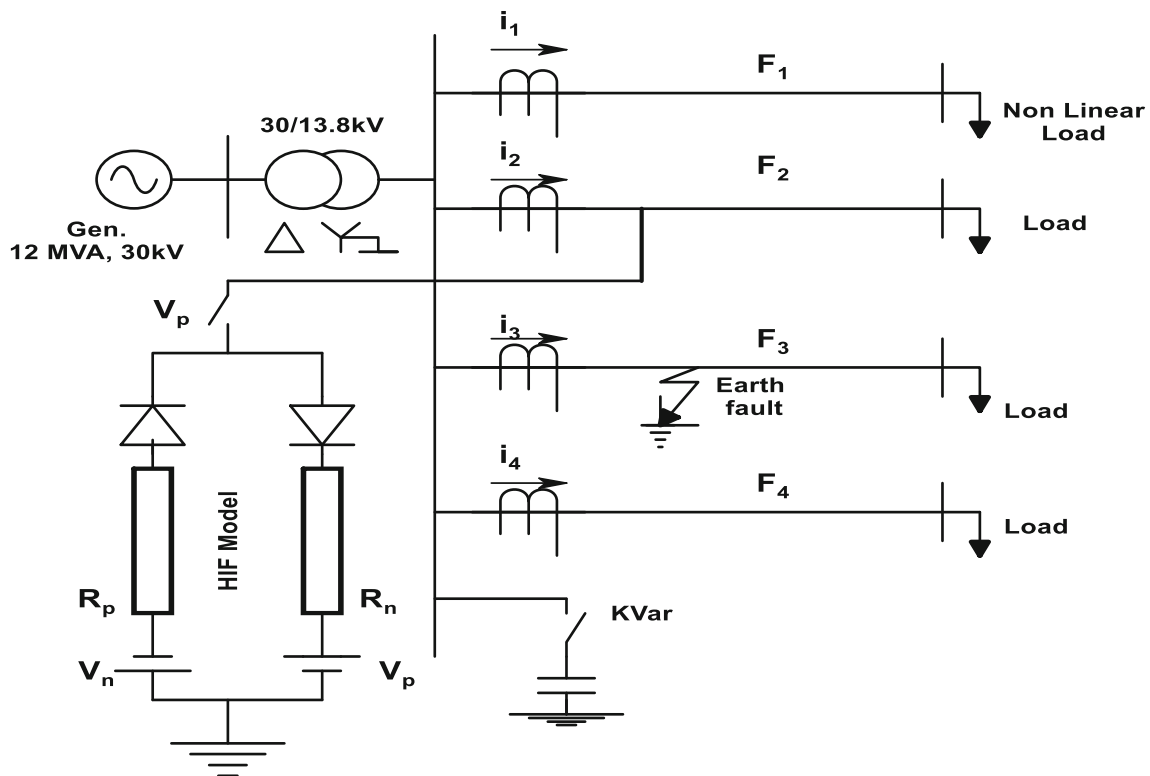


Fig. 1 Single-line diagram of radial distribution feeder

### 2.2 Proposed methodology to detect HIF

This section describes the steps for detection and identification of HIF in the MV distribution power system as portrayed in Fig. 2 and also discussed as follows.

*Step 1* Pre-processing: Design the MATLAB/Simulink model of MV network and apply the disturbances for acquisition of fault current data.

*Step 2* Feature extraction: The fault current signal is sampled using db9 mother wavelet of DWT and extracting the features, namely SD values, of all decomposed signal to train the classifiers.

*Step 3* Training phase: The data obtained (SD) from the DWT under various power system disturbances were used to train the computational intelligence-based classifiers.

*Step 4* HIF identification: The trained classifiers are tested with different power system disturbances to discriminate the HIF from other disturbances such as three-phase fault, line-to-ground (LG), line-to-line (LL) and double line-to-ground (LLG) fault. This process continues for every cycle of operation of the system for reliable operation and complete protection of the system.

### 3 Data acquisition using DWT analysis

Wavelet transform (WT) is a significant tool for decomposing the transient signal into several components and represents in time–frequency domain rather than the time domain [31, 32]. The fundamental idea behind this is to analyse the signal by means of dilation and translation process. In general, WT exists in two forms: CWT and DWT, in which continuous signal  $f(t)$  has been defined as given in Eq. (1)

$$CWT(a, b) = \left[ \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * h\left(t - \frac{b}{a}\right) dt \right] \tag{1}$$

where  $a$  and  $b$  are the scaling and translational parameters and  $h$  is the mother wavelet function. The mother wavelet is the prototype for generating wavelet function. CWT is an alternative approach to overcome the problem of resolution as in STFT. But, this method also has low redundancy during reconstruction of signal compared to the DWT method. Therefore, in this paper the DWT analysis was used to extract the features for training the classifiers.

The DWT is a powerful signal processing information tool which allows the signal to be sampled with localized transients. In recent decades, such powerful advanced tool has been used for operating the protective relays [27–31]. The time and frequency information can also be calculated

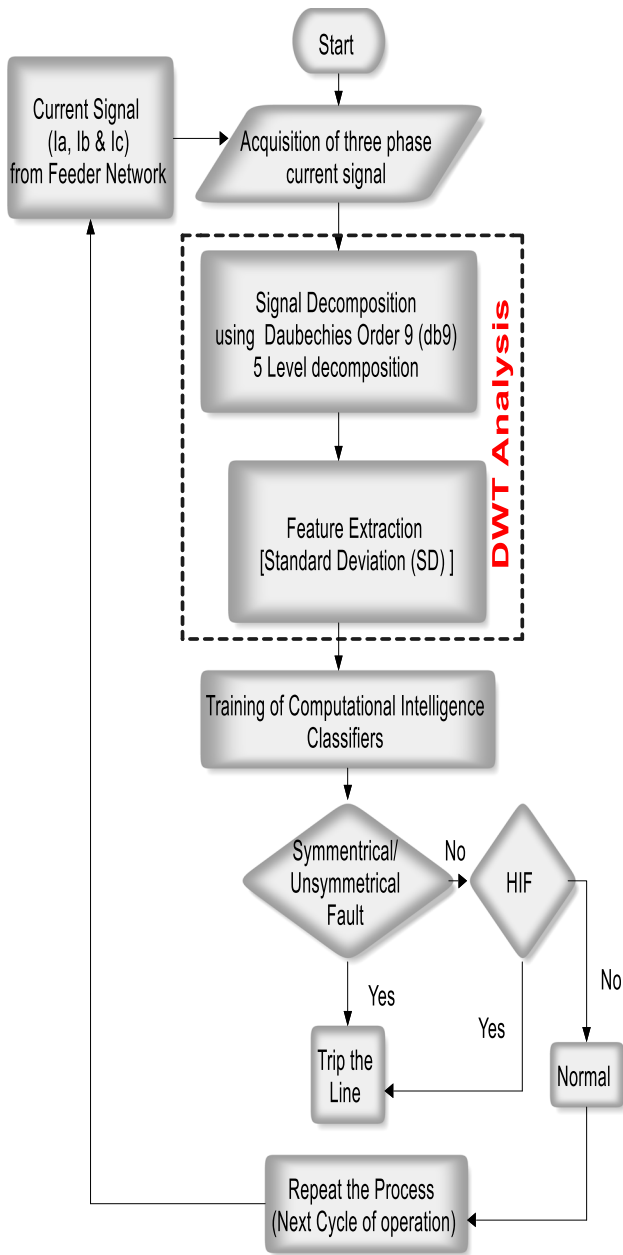


Fig. 2 Flow chart of the proposed method of fault classification

using the following techniques such as fast Fourier transform (FFT), STFT and CWT, but the DWT has been used because of its fast computation speed and accuracy [4]. Thus, DWT is defined as Eq. (2)

$$DWT(m, k) = \left[ \frac{1}{\sqrt{a_0^m}} \sum_n f(n) * h\left(\frac{k - na_0^m}{a_0^m}\right) \right] \quad (2)$$

where the parameter  $a_0^m$  and  $na_0^m$  are scaling and translation constant, respectively,  $n$  and  $m$  are integer variables and  $h$  is the wavelet function. Figure 3 shows the timescale representation of digital signal achieved by digital filtering techniques, and this representation is called subband

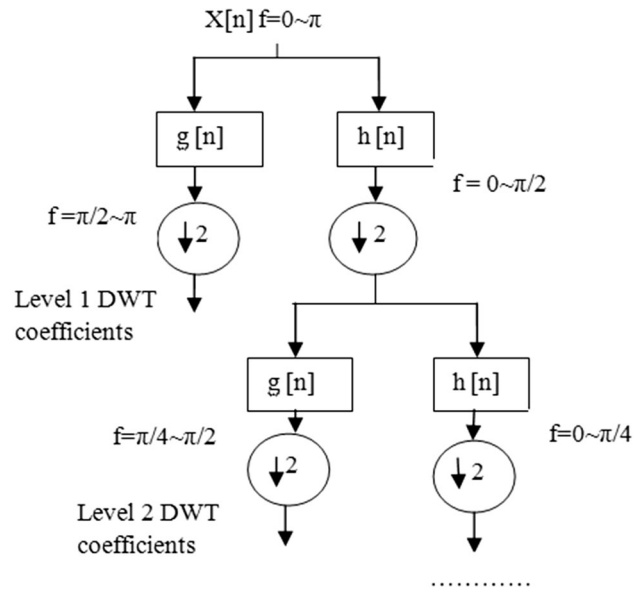


Fig. 3 DWT decomposition of signal

coding algorithm. The various decomposition levels of the signal  $X(n)$  are shown below [14].

Steps for decomposing the signal are as follows:

- Step 1 Decompose the original  $X(n)$  into levels, i.e. denoising process.
- Step 2 Choose the certain levels for reconstruction of desired signal.
- Step 3 Reconstruct the signal with selected levels.
- Step 4 Specify the following parameters: sampling frequency, window length, levels of decomposition and mother wavelet.

The original signal is divided into low-  $g(n)$  and high-frequency  $h(n)$  components by DWT, which are called approximation and detailed coefficients. The decomposed signal is further iterated with successive approximation, so that the signal is broken down into many lower-resolution components. This is called multi-resolution analysis (MRA).

In this work, Daubechies 9 (db9) is selected as mother wavelet for the detection of fault, because it is a better frequency extractor compared to the Haar wavelet. It also satisfies the Parseval’s theorem due to orthogonality and reduces the redundancy in the signal compared to Coiflet and Meyer wavelets [40, 41].

The optimal decomposition of  $L$  levels is given by the condition as represented in Eq. (3)

$$N = 2^L \quad (3)$$

where  $N$  is the length of decomposition level and  $L$  is the level of decomposition. During the signal decomposition,

at each level the signal has been divided into different frequency bands as defined in Eq. (4)

$$B = \frac{F}{2^{L+1}} \tag{4}$$

where  $B$  is the bandwidth of each level in Hz and  $F$  is the sampling frequency in Hz.

The sampling frequency of 20 kHz is considered for decomposing the signal into different levels with 5000 points in length and 800 samples for each phase of current signal. The band frequencies captured for each level are varied and calculated using Eq. (4) and are as follows: 5–2.5 kHz, 2.5–1.25 kHz, 1.25–0.625 kHz, 0.625–0.3125 kHz and 0.3125–0.15625 kHz are represented as detailed coefficient of  $d1, d2, d3, d4$  and  $d5$ , respectively, and the approximation is made from detailed coefficient  $d5$  [14]. Hence, the wavelet analysis is performed for the proposed work using the mother wavelet of db9 with detailed coefficients of 5 levels, for different fault current signal recorded for each cycle.

### 3.1 Feature extraction

Standard deviation (SD) is calculated for feature extraction to identify the type of fault. It is observed that some useful information can be extracted from SD of approximation signal ( $a5$ ) with detailed coefficient level of  $d1$  to  $d5$  for each phase. The SD values of current signal are calculated for each phase under different fault conditions. The SD values are obtained using Eq. (5) as shown below:

$$SD = \sqrt{\frac{1}{n-1} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right]} \tag{5}$$

where  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .  $x$  is the data vector, and  $n$  is the number of elements in that data vector.

## 4 Computational intelligence-based classifiers

This section describes the miscellaneous classifiers such as fuzzy, Bayes, MLP, ANFIS and SVM classifier approach to classify and discriminate the HIF from other faults that occur in the distribution system. Here, the classes for various faults are considered as  $C_1$ —normal,  $C_2$ —three-phase fault,  $C_3$ —LG fault,  $C_4$ —LLG,  $C_5$ —LL and  $C_6$ —HIF.

### 4.1 Fuzzy inference system (FIS)

FIS is a classical set theory employed in detection of HIF in the system and has the flexibility in handling the data during uncertainties. In few cases such as fault or other

disturbances in the system, the boundaries of the membership function may overlap and greatly reduce the accuracy of the crisp classifiers. Therefore, the performance of FIS is not appreciable in the case of classifier problem involving noisy, imprecise or incomplete data [32–34]. In this work, the features extracted (SD values) from DWT analysis of each phase A, B and C are considered as SDA, SDB and SDC. Then, these values are assigned as input with 4 input triangular membership functions and fed into the Sugeno FIS engine to determine the status of the system as described in Fig. 4. The boundary values for the input membership functions are as follows: 17–24 for normal, 25–35 for ground, 8–16 for HIF and 27–47 for fault. The developed FIS model in MATLAB is portrayed in Fig. 5 with its rules for classifications represented in Fig. 6.

### 4.2 Multi-layer perceptron (MLP) network

Multi-layer perceptron network is the feed forward structure of ANN with fully connected nodes using weights. The calculation of output at every node of the network is same which is given by:

$$Net_i = W_{i_o} + \sum_{j \in \text{pred}(i)} (W_{ij}a_j) \tag{6}$$

where  $a_j$  is the output of the previous layer neuron,  $W_{ij}$  is the weight connecting  $i$ th and  $j$ th neuron and  $W_{i_o}$  represents the input bias of the neuron. The MLP is trained using the back-propagation algorithm which calculates the error and update the weight at each trial run [31, 35, 38].

### 4.3 Bayesian neural network

The Bayesian neural network (BNN) uses the principle approach to deal with the complex network problem and also determines the input variables which are relevant to the output to be classified. This is done by proper prediction of test cases through the probability produced by the network with several set of parameters that are averaged

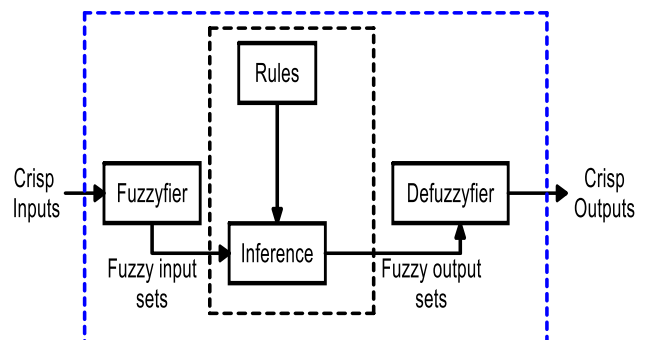


Fig. 4 Fuzzy inference system



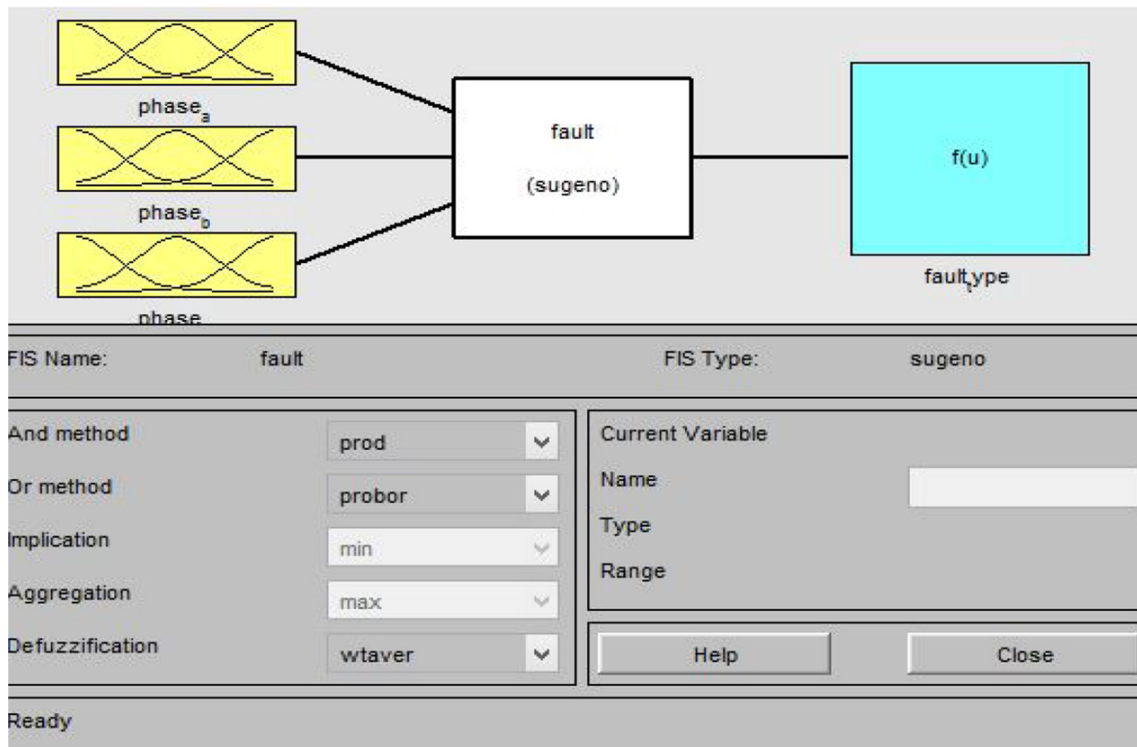


Fig. 5 MATLAB model of Sugeno-based FIS

are called predicted probabilities  $P(C)$ . This prior information is used for training the network about the system and then calculates the probability of occurrence of event or fault in the system by comparing the likelihood probability,  $P(L)$  with the  $P(C)$ . Let  $C$  denote the classes  $C_1$  to  $C_6$  of a likelihood of data set  $L$ . To predict the class of likelihood  $L = \{X_1, X_2, \dots, X_n\}$  by using the Bayes rule, the highest probability is,

$$P(C|X_1, X_2, \dots, X_n) = \frac{P(C)P(X_1, X_2, \dots, X_n|C)}{P(X_1, X_2, \dots, X_n)} \tag{7}$$

The simplified form of the above equation is represented as follows:

$$P(C|L) = \frac{P(C)P(L|C)}{P(L)} \tag{8}$$

It is seen that from Eq. (7) the number of parameters in the likelihood term  $P(X_1, X_2, \dots, X_n|C)$ , increases exponentially with the number of attributes leading to the computational burden of the classifier [36, 37].

#### 4.4 Support vector machine (SVM)

SVM is a statistical theory-based popular intelligent soft computing classifier to classify the data without pre-processing the data. It is a nonlinear kernel-based classifier which maps the data from the one space region to other

form, and this is done during training of data and also assumes each fault or disturbance as a class labelled in Sect. 4. During the testing phase, the classifier identifies the type of fault by predicting the class label of the test input data and detects the disturbance occurred in the system. The examples of input and decision space are shown in Figs. 7 and 8, respectively [24–26].

The hyperplane  $f(x)$  as shown in Fig. 7 which groups the input data after training of classifier is depicted as:

$$f(x) = \text{sign} \left( \sum_{i \in N} \alpha_i^* y_i \cdot x_i + b^* \right) \tag{9}$$

$$x \in \begin{cases} \text{class I,} & \text{if } f(x) = +1 \\ \text{class II,} & \text{if } f(x) = -1 \end{cases}$$

where  $x$  is the input data set for classification with  $N$  samples,  $\alpha_i^*$  is the optimal value of Lagrangian multiplier,  $y_i \in \{+1, -1\}$  which decides the class of input data sample  $x$  and  $b^*$  is the solution of weight vector as given in [24].

A nonlinear kernel-based SVM is used which consists of 3 SVMs, namely SDA, SDB and SDC, to detect the fault which occurs in three phases A, B and C, respectively. Figure 9 shows the decision space vector of SVM for SDA after training phase, and this is done using a kernel function for mapping the data in higher-dimensional space. Therefore, rewriting Eq. (9) using a kernel function results in [38, 39]

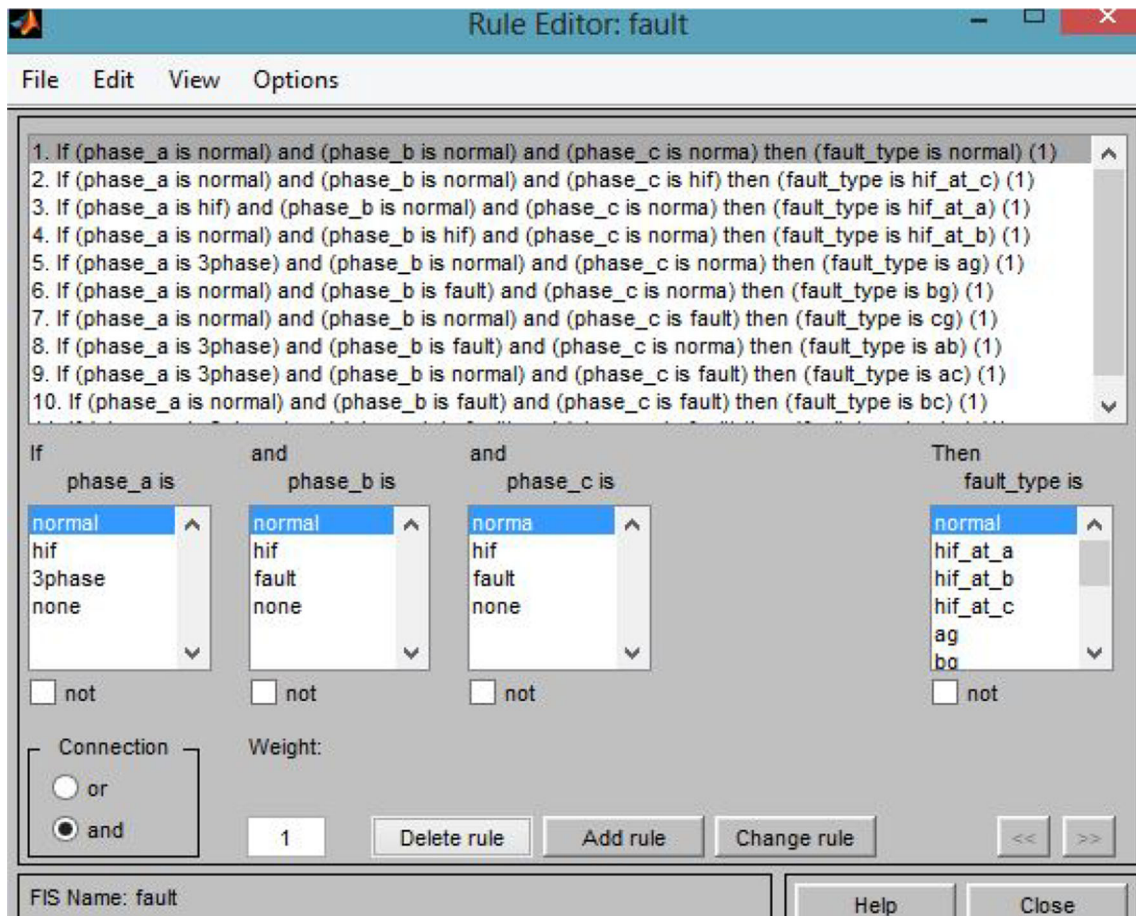


Fig. 6 Rule viewer to classify the fault for FIS system in MATLAB

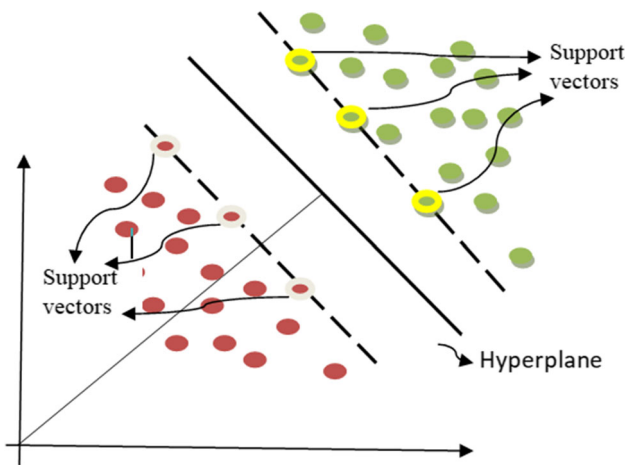


Fig. 7 Support vector machine

$$f(x) = \text{sign} \left( \sum_{i \in N} \alpha_i^* y_i \cdot k(x_i, x_N) + b^* \right) \quad (10)$$

where  $k(x_i, x_N)$  is the kernel function and is defined by radial basis function (RBF) because of reliability than other kernel function and is represented by [30].

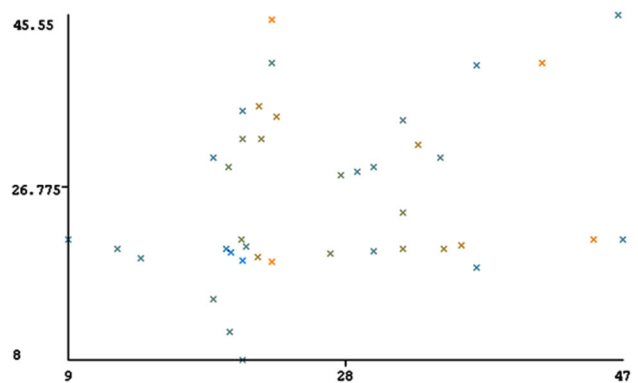


Fig. 8 Input space of nonlinear separable classifier

$$k(x_i, x_N) = \exp \left( - \frac{\|x - y\|^2}{2\sigma^2} \right) \quad (11)$$

The SVM detects the class as 1 if the fault occurs in the respective phase during the testing phase, otherwise it gives  $-1$ . The multiple logics for SVM under different type of fault of the system are represented in Table 1.

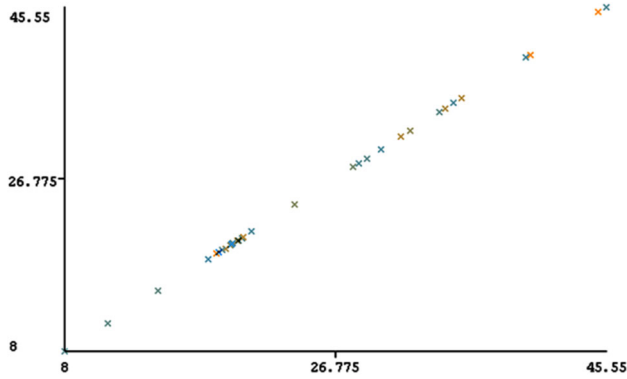


Fig. 9 Decision space of nonlinear separable classifier

4.5 Adaptive neuro-fuzzy inference system

ANFIS is an intelligent adaptive data learning method in which the fuzzy inference system is optimized via the ANN training. It maps the input and output through the input and output member function. From the input–output data, ANFIS adjusts the membership function using least square method or back-propagation descent method for linear and nonlinear system. The Sugeno fuzzy model has been proposed for creating the fuzzy rules from a given input–output data set. A typical Sugeno fuzzy rule is expressed in the following form [19–23]:

- Rule 1 If ( $x_1$  is  $A_1$ ) and ( $x_2$  is  $B_1$ ), then  $f_1 = (u_1x_1 + v_1x_2 + r_1)$
- Rule 2 If ( $x_1$  is  $A_2$ ) and ( $x_2$  is  $B_2$ ), then  $f_2 = (u_2x_1 + v_2x_2 + r_2)$

where the vector  $X = \{x_1, x_2, \dots, x_m\}$  is the input data to be classified.  $A_i, B_i$  and  $C_i$  are the fuzzy sets with output  $f_i$  specified within the rule-based fuzzy system.  $u_i, v_i$  and  $r_i$  are design parameters which are obtained during the training process. The architecture of ANFIS system consists of 6 layers: They are the input layer, fuzzification layer, rule layer, normalization layer, defuzzification layer and output layer, respectively as given in Fig. 10.

Table 1 Multiple SVM logic for fault classification

Classes	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	Fault type
$C_1$	1	-1	-1	-1	-1	-1	Normal
$C_2$	-1	1	-1	-1	-1	-1	Three-phase
$C_3$	-1	-1	1	-1	-1	-1	LG
$C_4$	-1	-1	-1	1	-1	-1	LLG
$C_5$	-1	-1	-1	-1	1	-1	LL
$C_6$	-1	-1	-1	-1	-1	1	HIF

Layer 1 is the input layer which passes the external crisp signals to Layer 2 called fuzzification layer which has the bell activation function and is defined by [19]:

$$\mu_{A_i}(x) = \frac{1}{1 + \left\{ \left( \frac{x-c_i}{a_i} \right)^2 \right\}^{b_i}} \tag{12}$$

where  $a_i, b_i$  and  $c_i$  are the design parameters of bell-shaped membership function. Layer 3 is the rule layer; here each neuron corresponds to a single Sugeno-type fuzzy rule. The neuron gets inputs from the corresponding fuzzification neurons and calculates the firing strength of the rule. In ANFIS, the conjunction of the rule antecedents is evaluated by the operator product. Thus, the output of neuron ‘ $i$ ’ in Layer 3 is obtained in terms of membership function ( $\mu$ ) using Eq. (12).

$$y_i^3 = w_i = \mu_{A_i}(x_i)\mu_{B_i}(x_j) \tag{13}$$

Layer 4 is the normalization layer. Individual neuron in this layer obtains inputs from all neurons in the Layer 3 and calculates the normalized firing strength of a given rule as the ratio of the firing strength of a specified rule to the sum of firing strengths of all rules. It signifies the influence of a given rule to the final result. Thus, the output of neuron ‘ $i$ ’ in Layer 4 is determined as given by Eq. (14).

$$y_i^4 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i \in 1, 2 \tag{14}$$

Layer 5 is the defuzzification layer. Respective neuron in this layer is linked to the individual normalization neuron and also receives initial inputs  $x_1$  and  $x_2$ . A defuzzification neuron calculates the weighted consequent value of a given rule as given in Eq. (15)

$$y_i^5 = \bar{w}_i f_i = \bar{w}_i(u_i x_1 + v_i x_2 + r_i), \quad i \in 1, 2 \tag{15}$$

where  $x_i$  is the input and  $y_i$  is the output of defuzzification neuron in Layer 5,  $u_i, v_i$  and  $r_i$  are set of consequent parameters of rule ‘ $i$ ’ and Layer 6 is represented by a single summation neuron. This neuron calculates the sum of outputs of all defuzzification neurons and produces the overall ANFIS output ‘ $y$ ’ which is given by Eq. (16).

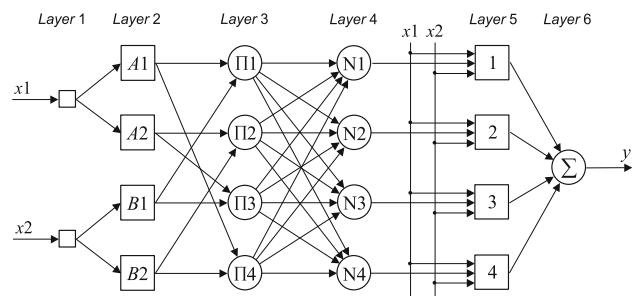


Fig. 10 ANFIS structure



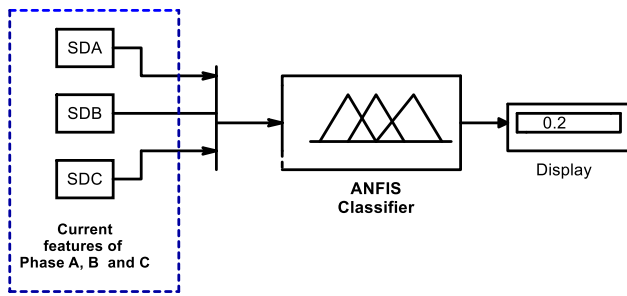


Fig. 11 ANFIS results for distribution network

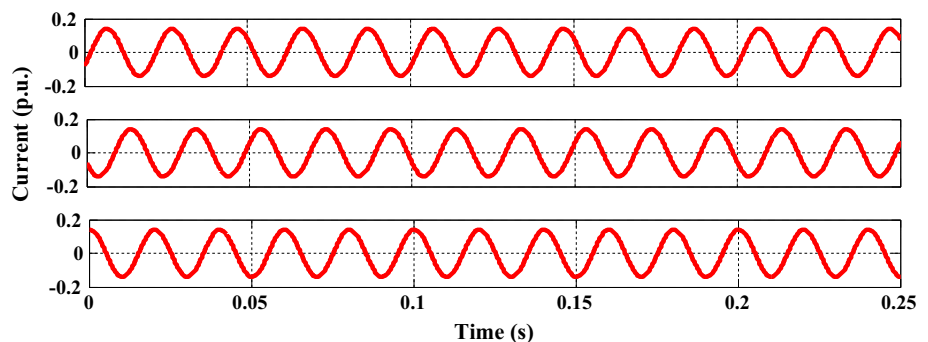
$$y_i^6 = \sum_{i=1}^2 \bar{w}_i f_i = \frac{\sum_{i=1}^2 w_i f_i}{w_1 + w_2} \tag{16}$$

The SD values obtained for different conditions of the system using DWT analysis are fed as input to the ANFIS model with its output representing the type of fault occurred in the system. To achieve this, the training for input data set is done using the MATLAB ANFIS by utilizing the mathematical properties of ANN in tuning the rule-based fuzzy system and that exploits the best qualities of these two approaches efficiently. Thus, the ANFIS system possesses the following advantages than fuzzy approach [38, 39]:

- Revising the IF–THEN rules to describe the characteristics of complex system.
- Knowledge of human expertise for training is not required as like that of FI approach.
- Have better choice of membership function (MFs) to use.
- Faster convergence speed than fuzzy approach.
- Adaptive through mathematical learning gives accurate and efficient results.

The network is trained for the input–output data set with MATLAB ANFIS editor, which adjusts the MFs directly based on the data set. Here for each phase of input, SDA, SDB and SDC are assumed with 4 triangular membership functions and the output being constant for Sugeno model ANFIS. Fourteen rules were framed with 45 input–output

Fig. 12 Three-phase current waveform for normal case



data sets for training the ANFIS model. The trained constant output assumed for various cases of the system is as follows: 0—no fault, 0.2—HIF in phase C, 0.3—HIF in phase B, 0.4—HIF in phase C, 0.5—LLLG, 0.6—AG, 0.7—BG, 0.8—CG, 0.9—AB, 1—BC, 1.1—CA, 1.2—ABG, 1.3—BCG and 1.4—CAG. The output model of ANFIS structure developed in MATLAB/Simulink is shown in Fig. 11, which represents that 0.2 indicates the HIF in phase C of the system.

### 4.6 Performance indices for classifiers

#### 4.6.1 Kappa statistic

It is the statistical measure of classifiers that calculate the consistency among the predicted and observed data sets and is defined as follows:

$$K = \frac{P(\text{OF}) - P(\text{EF})}{(1 - P(\text{EF}))} \tag{17}$$

where  $P(\text{OF})$  is the probability of observed fault in the system and  $P(\text{EF})$  is the probability of expected type of fault by chance that occurs in the system. Its value ranges from 0 to 1. If the classifier is excellent, then the kappa statistics is 1; if it is between 0.4 and 0.75, then it is good; and if it is less than 0.4, then the performance is poor [21, 38, 39]. This measurement is important to calculate the overall accuracy of the classifier.

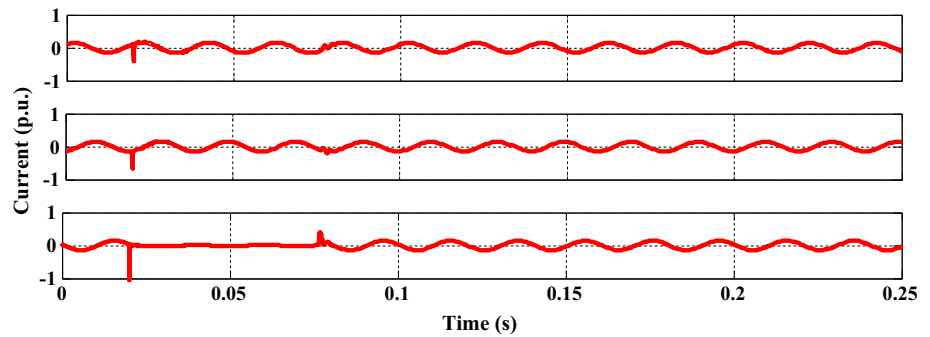
#### 4.6.2 Mean absolute error (MAE) and root mean square error (RMSE)

MAE is the absolute average measure of error between the predicted and observed value of the classifier, and it is depicted as follows [21, 38, 39]:

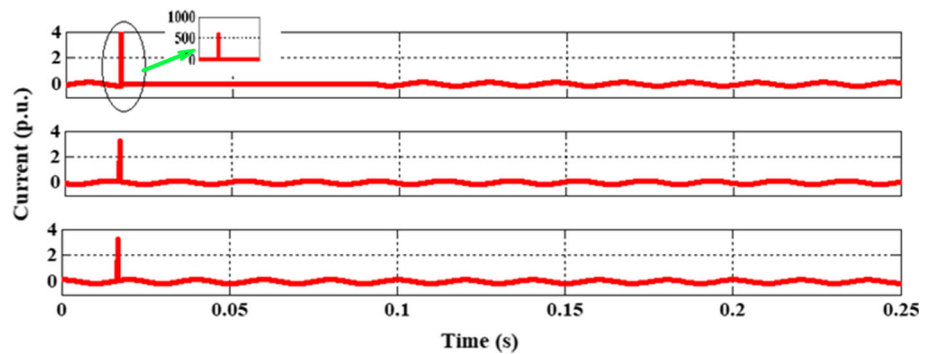
$$\text{MAE} = \frac{|\sum_{i=1}^n (E_P - E_O)|}{n} \tag{18}$$

RMSE is the square root of average measure of variance between the predicted and observed fault by the classifiers and is given by:

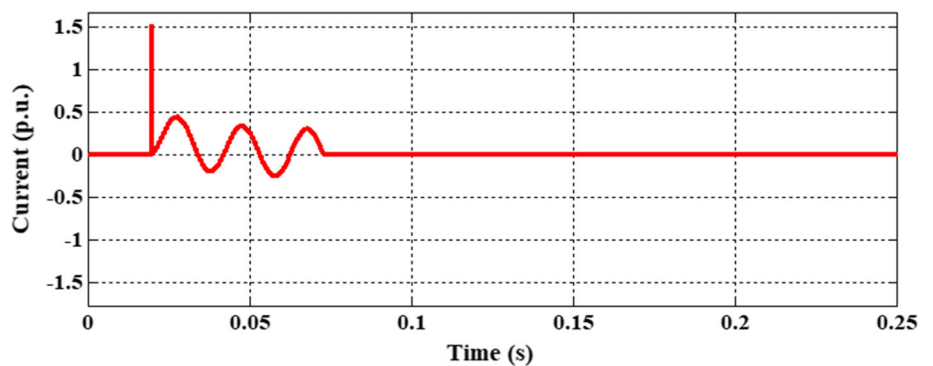
**Fig. 13** Three-phase current waveform for HIF at phase C



**Fig. 14** Three-phase current waveform for LG fault at phase A



**Fig. 15** High-impedance fault current waveform



$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (E_P - E_O)^2}{n}} \quad (19)$$

where  $E_P$  is the predicted value of classifier and  $E_O$  is the observed value of the classifiers over 'n' data sample.

## 5 Results and discussion

This section presents the simulation of proposed approach for detection and identification of HIF in MV distribution network. To validate the performance of the proffered method, the simulation is carried out using MATLAB/

Simulink for the distribution model portrayed in Fig. 1. The time-varying current signal for the normal condition of feeder for the time period of 0.25 s is captured and is represented in Fig. 12. Moreover, the various faults that occur in the power system were also simulated in addition with the study of HIF to prove the effectiveness of the proposed method. The three-phase current waveform observed for the period of 0.02–0.08 s in case of HIF fault and single line to ground (LG) is shown in Figs. 13 and 14, respectively. Figure 15 shows that the fault current magnitude in case of HIF fault in phase C of three-phase system is low as like the normal phase current signal. But the amplitude of current signal in case of LG fault in phase A of three-phase system is very high as

**Table 2** Fault detection in case of varying fault resistance by different intelligence classifiers

Class	Fault type	Fault resistance ( $R_f$ ) ( $\Omega$ )	Feature extraction			Classifiers output				
			SDA	SDB	SDC	Fuzzy	Bayes	MLP	ANFIS	SVM
Normal	Normal	0.01	20.33	21.22	23	✓	✓	✓	✓	✓
Three-phase fault	ABCG	20	40.33	41.54	46	✓	✓	✓	✓	✓
	ABCG	40	31.38	33	35.98	✓	✓	✓	✓	✓
	ABCG	60	28	27.74	27	✓	✓	✓	✓	✓
LG	AG	20	40.33	23	22.64	✓	✓	✓	✓	✓
	AG	40	35	21	20.06	✓	✓	✓	✓	✓
	AG	60	29.98	19	20	✓	✓	✓	✓	✓
	BG	20	21	47	20.06	✓	✓	✓	✓	✓
	BG	40	18	37	18.63	✓	✓	✓	✓	✓
	BG	60	19.73	30	22	✓	✓	✓	✓	✓
	CG	20	18.6	23	47	✓	✓	✓	✓	✓
	CG	40	19.18	22	34.98	✓	✓	✓	✓	✓
	CG	60	21	20.87	29.61	✓	✓	✓	✓	✓
LLG fault	ABG	20	30	34.5	23	✓	✓	✓	✓	✓
	ABG	40	29	30	22.45	✓	✓	✓	✓	✓
	ABG	60	28.42	28.88	21	✓	✓	✓	✓	✓
	BCG	20	20.03	34.76	34	✓	✓	✓	✓	✓
	BCG	40	20	32	31	✓	✓	✓	✓	✓
	BCG	60	19.55	27	29	✓	✓	✓	✓	✓
	ACG	20	34.45	23.33	35.1	✓	✓	✓	✓	✓
	ACG	40	32	22.3	31	✓	✓	✓	✓	✓
	ACG	60	29	20	28	✓	✓	✓	✓	✓
LL fault	AB	20	45.55	46.7	21	✓	✓	✓	✓	✓
	AB	40	40	37	20.1	✓	✓	✓	✓	✓
	AB	60	34	32	23	✗	✓	✓	✓	✓
	BC	20	21	45	44	✓	✓	✓	✓	✓
	BC	40	20.45	36	37	✓	✓	✓	✓	✓
	BC	60	24	32	29.24	✗	✓	✓	✓	✓
	AC	20	45	23	46.9	✓	✓	✓	✓	✓
	AC	40	35.55	22.1	36	✓	✓	✓	✓	✓
	AC	60	32	21	29	✗	✓	✓	✓	✓
HIF fault	HIF A	20	8	21	22.2	✓	✓	✓	✓	✓
	HIF A	40	11	20.09	23.4	✓	✓	✓	✓	✓
	HIF A	60	14.5	19	24	✗	✓	✓	✓	✓
	HIF B	20	21	9	20.01	✓	✓	✓	✓	✓
	HIF B	40	20.09	12.4	23.05	✓	✓	✓	✓	✓
	HIF B	60	19	14	22	✗	✓	✓	✓	✓
	HIF C	20	18.76	21	8.13	✓	✓	✓	✓	✓
	HIF C	40	19.61	20.19	12.09	✓	✓	✓	✓	✓
	HIF C	60	20.08	19.89	15.5	✗	✓	✓	✓	✓

‘✗’ represents misclassified, and ‘✓’ represents accurately classified results

shown in Fig. 14, and because of this, the detection of HIF in power system is a major challenge till date as compared to other conventional type of fault that occurs in the system. To overcome these difficulties in real time, DWT analysis was

used to extract the features as it decomposes the signal both in time and in frequency domain.

The DWT analysis of each phase current signal which is sampled for ever cycle with 800 numbers of samples is

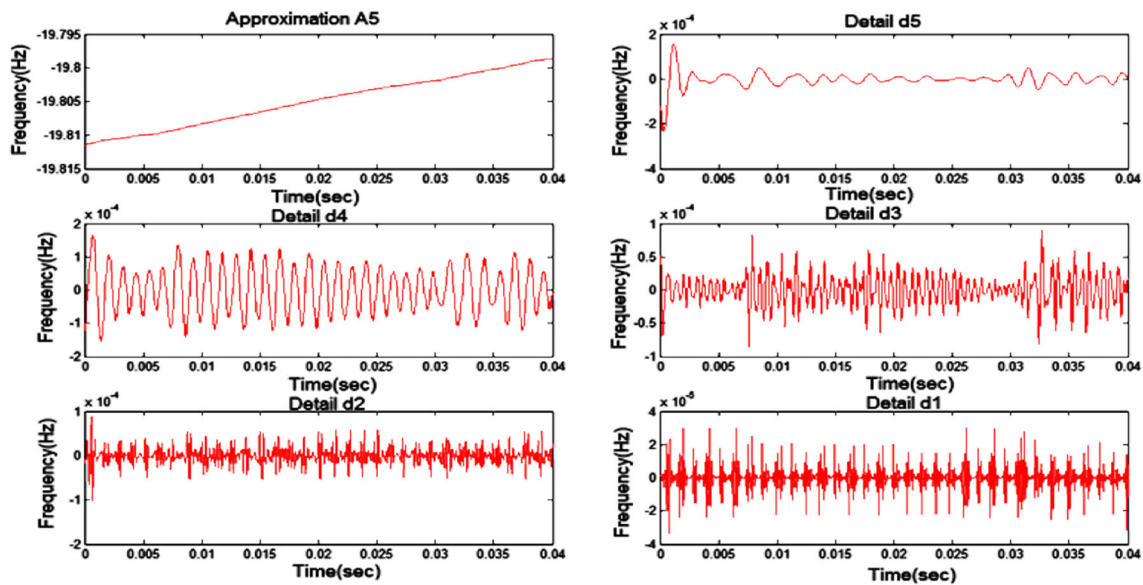


Fig. 16 DWT analysis of phase A under normal case

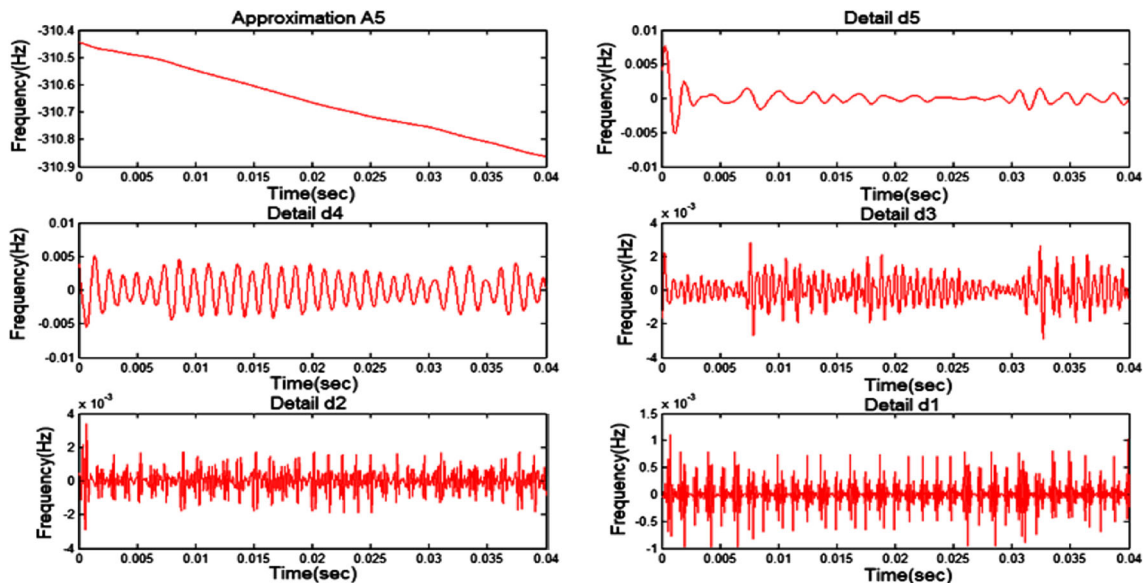


Fig. 17 DWT analysis of phase B under normal case

performed with five decomposition levels. Each level represents with different bands of frequencies as discussed in Sect. 3. The approximate coefficient obtained for the final decomposed level ( $d5$ ) and the various detailed coefficient levels  $d1$ ,  $d2$ ,  $d3$ ,  $d4$  and  $d5$  are used for calculating SD values as shown in Table 2. The DWT analysis of each phase A, B and C of the system under normal case is portrayed in Figs. 16, 17 and 18, respectively. In addition, the DWT analysis for the faulty phase during the occurrence of HIF in phase C in one of the feeder and the occurrence of LG fault in phase A on another feeder network is also represented in Figs. 19 and 20, respectively,

for better understanding of the proposed method. Similarly, the DWT analysis for other faults such as LL, LLG and three-phase fault for varying fault resistance is done and the extracted SD features for training the classifiers to detect the HIF in the system are given in Table 2.

The performance indices such as MAE, RMSE and kappa statistic of the proposed ANFIS and SVM method of classifiers are compared with the fuzzy, MLP and Bayes classifier approach. Figures 21 and 22 represent that the proposed classifiers performance is significantly improved compared to the conventional classifiers and the results are illustrated in Tables 2 and 3. It is observed that the kappa

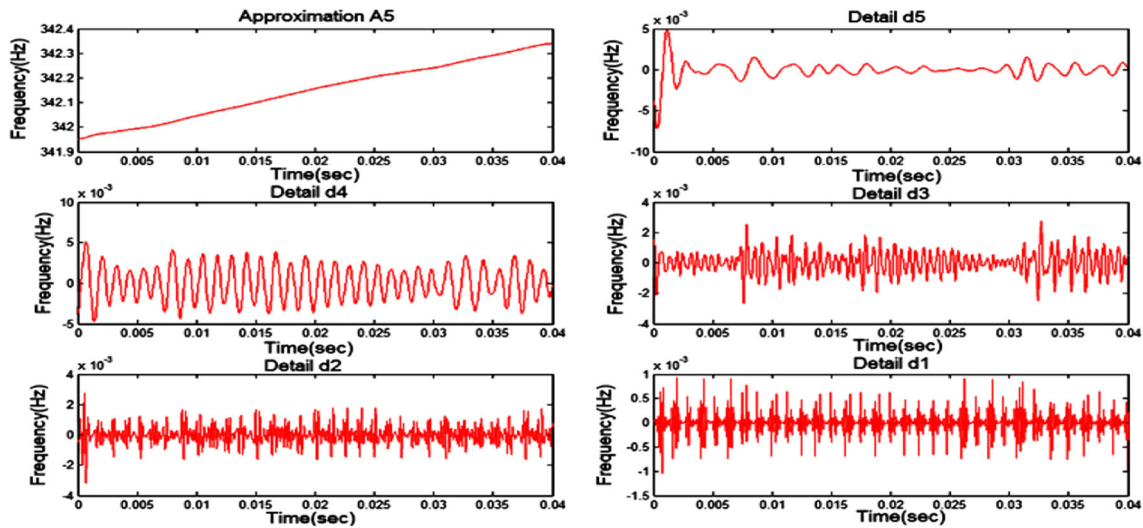


Fig. 18 DWT analysis of phase C under normal case

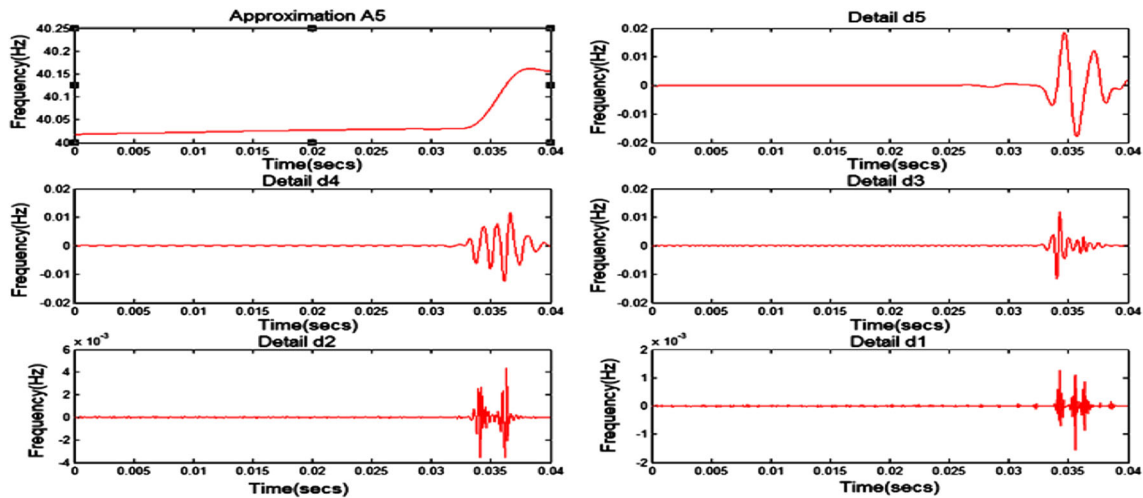


Fig. 19 DWT waveform of HIF fault at phase C

statistics is 1 for all the classifier approaches than fuzzy method due to human error introduced during the selection of ranges for membership functions. On the other hand, all the neural-based and SVM classifiers are learned mathematically and there is no possibility of human error. Thus, the performance of ANFIS and SVM method is substantially predominant than any other method as depicted in Fig. 22.

In addition with evaluation indices, the other performance indices are also tested further to validate the effectiveness of the proposed classifiers for classification of fault and are defined as below:

$$\text{Success rate} = \frac{\text{Number of HIF detected}}{\text{Total number of HIF events}} * 100\% \quad (20)$$

Discrimination rate

$$= \left[ 1 - \frac{\text{Number of events in correctly diagnosed}}{\text{Total number of events}} \right] * 100\% \quad (21)$$

Figure 21 shows the success and discrimination rate of all the classifier methods are 100% except the fuzzy-based approach which is 66.67% and 85%, respectively, and is presented in Table 3.

## 6 Comparison of literature work

The comparative performance of accuracy for various classifiers to identify different types of fault in the system by proffered method is made with all other existing



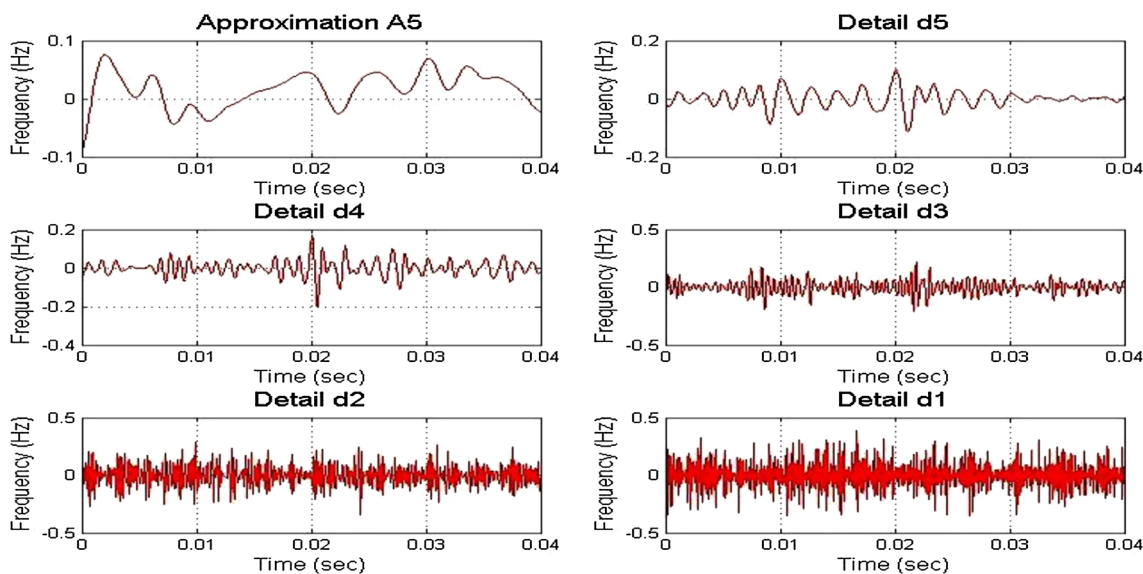


Fig. 20 DWT waveform of LG fault at phase A

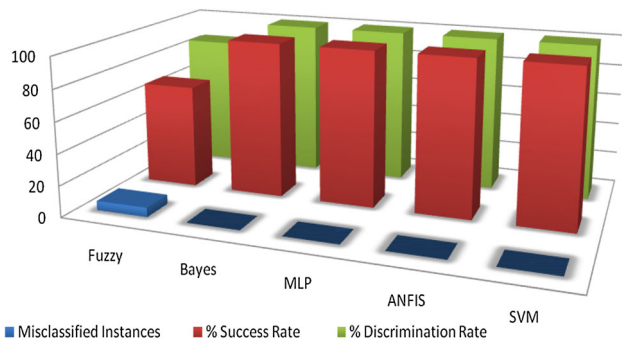


Fig. 21 Accuracy and effectiveness of different intelligence classifiers

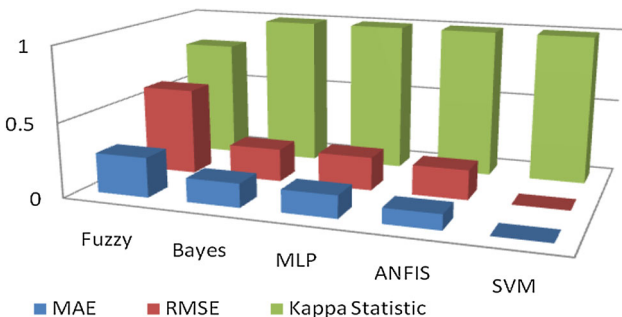


Fig. 22 Comparison of MAE, RMSE and kappa statistic of various intelligence classifiers

methods and presented in Table 4. The result reveals that all the intelligence-based classifier requires the training of data obtained from the signal processing techniques and their accuracy lies in the range of 70–100% for detection of either conventional fault (LG, LL, LLG and three-phase fault) [42–51] or HIF [12, 13, 15, 34, 52]. Nevertheless,

these works fail to identify conventional and HIF faults in the system, whereas the proposed work considers both faults. Moreover, to test the effectiveness of the proposed classifiers the performance evaluation indices such as kappa statistics, MAE and RMSE are evaluated. However, the literature work portrayed in Table 4 fails to evaluate these performance indices that demonstrate the robustness of the classifier in addition with the accuracy of classification. It can be concluded that from the results given in Tables 2 and 3 the proposed ANFIS and SVM classifier performance is substantially predominant for identifying various types of fault in the system.

### 7 Conclusion

In this paper, the detection and identification of HIF in MV distribution power system network is presented using computational intelligence-based classifiers such as fuzzy, MLP, Bayes, ANFIS and SVM. The distribution feeder network of 13.8 kV is simulated using MATLAB/Simulink, and various power system disturbances such as HIF, LG, LL, LLG and three-phase fault are studied with time-varying fault resistance for testing the effectiveness of the proposed classifiers. The DWT analysis of three-phase current signal using a mother wavelet of db9 with detailed coefficient of 5 levels is done to extract the SD features for various types of fault. The features obtained were used to train the computational intelligence-based classifiers to identify the HIF in the system. The results showed that the performance of ANFIS and SVM classifiers is superior in terms of success rate, discrimination rate, MAE, RMSE and kappa statistics. Furthermore, it is observed that the

**Table 3** Performance comparison of various intelligence classifiers

Method	Misclassified instances	% Success rate	% Discrimination rate	MAE	RMSE	Kappa statistic
Fuzzy	6	66.67	85	0.268	0.587	0.8
Bayes	0	100	100	0.1603	0.2229	1
MLP	0	100	100	0.151	0.2248	1
ANFIS	0	100	100	<b>0.1028</b>	<b>0.2005</b>	<b>1</b>
SVM	0	100	100	<b>0.0002</b>	<b>0.0005</b>	<b>1</b>

Bold signifies the results are substantially improved for the proposed ANFIS and SVM classifier than other classifier

**Table 4** Performance comparison with literature work

References	Classification method	Type of fault considered						% Accuracy	
		LG	LL	LLG	LLG	LLLG	HIF		Fault resistance $R_f$ ( $\Omega$ )
Reddy et al. [42]	DWT + ANFIS	✓	✓	✓	✓	✓	✗	✗	100
Silva et al. [43]	DWT + ANN	✓	✓	✓	✓	✓	✗	10–100	99.83
Vyas et al. [44]	WT + Ch-NN	✓	✓	✓	✓	✓	✗	✗	99.43
Jamehbozorg et al. [45]	DFT + decision tree	✓	✓	✓	✓	✓	✗	0–50	100
Malik et al. [46]	WT + modified fuzzy Q learning	✓	✓	✓	✓	✓	✗	✗	99.32
Salehi et al. [47]	Morphological edge detection filter	✓	✓	✗	✗	✗	✗	0–100	100
Mahmud et al. [48]	WT + 2-tier MLP	✓	✓	✓	✓	✓	✗	✗	99.36
Alsafasfeh et al. [49]	Principal component analysis	✓	✓	✓	✓	✗	✗	5–100	94.54
Mishra et al. [50]	DFT + fuzzy (series compensated line)	✓	✓	✓	✓	✓	✗	0.001–100	99.678
Samet et al. [51]	Improved alienation coefficients method	✓	✓	✓	✓	✓	✗	0–70	92.88
Tonelli-Neto et al. [34]	WT + fuzzy-ARTMAP	✗	✗	✗	✗	✗	✓	✗	97.69
Santos et al. [13]	Energy spectrum of DWT (considering DG placement)	✗	✗	✗	✗	✗	✓	✗	70
Gomes et al. [52]	DWT + boosted decision tree	✗	✗	✗	✗	✗	✓	✗	98.06
Kavi et al. [12]	Morphological fault detector algorithm	✗	✗	✗	✗	✗	✓	✗	100
Asghari Govar et al. [15]	Adaptive CWT and extreme learning machine (considering CT saturation)	✗	✗	✗	✗	✗	✓	✗	100
Proposed method (DWT + ANFIS and SVM classifiers)		✓	✓	✓	✓	✓	✓	0–60	100

‘✗’ represents the fault type/parameter which is not considered for classification, and ‘✓’ represents the occurrence of fault

SVM-based intelligence classifier performance is remarkable than other classifiers. Nonetheless, a more depth analysis on identification of HIF in the distribution network comprising of renewable energy resources will be the future goal of research integrating the field of data mining and internet of things (IoT) for rapid recovery of system from fault conditions.

**Acknowledgements** The author thanks the centre for Advanced Lighting, Power and Energy Research (ALPER), University Putra Malaysia (UPM) for the fund under (9630000). Also we thank the potential reviewer for their valuable comments to improve the quality of the paper.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**References**

1. Costa FB, Souza BA, Brito NSD, Silva JACB, Santos WC (2015) Real-time detection of transients induced by high-impedance faults based on the boundary wavelet transform. *IEEE Trans Ind Appl* 51(6):5312–5323
2. Wang B, Geng J, Dong X (2018) High-impedance fault detection based on nonlinear voltage–current characteristic profile identification. *IEEE Trans Smart Grid* 9(4):3783–3791

3. Sedighi A, Haghifam M, Malik OP, Ghassemian M (2005) High impedance fault detection based on wavelet transform and statistical pattern recognition. *IEEE Trans Power Deliv* 20(4):2414–2421
4. Elkalashy NI, Lehtonen M, Darwish HA, Taalab AI, Izzularab MA (2008) DWT-based detection and transient power direction-based location of high-impedance faults due to leaning trees in unearthed MV networks. *IEEE Trans Power Deliv* 23(1):94–101
5. Nikander A, Järventausta P (2017) Identification of high-impedance earth faults in neutral isolated or compensated MV networks. *IEEE Trans Power Deliv* 32(3):1187–1195
6. Guardado JL, Torres V, Maximov S, Melgoza E (2018) Analytical approach to modelling the interaction between power distribution systems and high impedance faults. *IET Gener Transm Distrib* 12(9):2190–2198
7. Nikita K, Preeti K (2015) Analysis and modeling of high impedance fault. *Int J Electr Electron Eng* 2(3):1–5
8. Bahador N, Namdari F, Matinfar HR (2018) Modelling and detection of live tree-related high impedance fault in distribution systems. *IET Gener Transm Distrib* 12(3):756–766
9. Gonzalez C, Tant J, Germain JG, De Rybel T, Driesen J (2018) Directional, high-impedance fault detection in isolated neutral distribution grids. *IEEE Trans Power Deliv* 33(5):2474–2483
10. Tang T, Huang C, Hua L, Zhu J, Zhang Z (2018) Single-phase high-impedance fault protection for low-resistance grounded distribution network. *IET Gener Transm Distrib* 12(10):2462–2470
11. Lima ÉM, Dos Santos Junqueira CM, Brito NSD, Souza BA D, De Almeida CR, Suassuna GM, de Medeiros H (2018) High impedance fault detection method based on the short-time Fourier transform. *IET Gener Transm Distrib* 12(11):2577–2584
12. Kavi M, Mishra Y, Vilathgamuwa MD (2018) High-impedance fault detection and classification in power system distribution networks using morphological fault detector algorithm. *IET Gener Transm Distrib* 12(15):3699–3710
13. Santos WC, Lopes FV, Brito NSD, Souza BA (2017) High-impedance fault identification on distribution networks. *IEEE Trans Power Deliv* 32(1):23–32
14. Chen J, Phung T, Blackburn T, Ambikairajah E, Zhang D (2016) Detection of high impedance faults using current transformers for sensing and identification based on features extracted using wavelet transform. *IET Gener Transm Distrib* 10(12):2990–2998
15. Asghari Govar S, Heidari S, Seyedi H, Ghasemzadeh S, Pourghasem P (2018) Adaptive CWT-based overcurrent protection for smart distribution grids considering CT saturation and high-impedance fault. *IET Gener Transm Distrib* 12(6):1366–1373
16. Ghaderi A, Mohammadpour HA, Ginn HL, Shin Y (2015) High-impedance fault detection in the distribution network using the time-frequency-based algorithm. *IEEE Trans Power Deliv* 30(3):1260–1268
17. Baqui I, Zamora I, Mazón J, Buigues G (2011) High impedance fault detection methodology using wavelet transform and artificial neural networks. *Electr Power Syst Res* 81(7):1325–1333
18. Milioudis AN, Andreou GT, Labridis DP (2015) Detection and location of high impedance faults in multiconductor overhead distribution lines using power line communication devices. *IEEE Trans Smart Grid* 6(2):894–902
19. Güler I, Übeyli E (2005) Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. *J Neurosci Methods* 148(2):113–121
20. Yang Z, Wang Y, Ouyang G (2014) Adaptive neuro-fuzzy inference system for classification of background EEG signals from ESES patients and controls. *Sci World J* 2014:1–8
21. Ghosh S, Biswas S, Sarkar D, Sarkar PP (2014) A novel neuro-fuzzy classification technique for data mining. *Egypt Inf J* 15(3):129–147
22. Durgadevi S, Umamaheswari MG (2018) Analysis and design of single-phase power factor corrector with genetic algorithm and adaptive neuro-fuzzy-based sliding mode controller using DC–DC SEPIC. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-018-3424-2>
23. Komathi C, Umamaheswari MG (2019) Analysis and design of genetic algorithm-based cascade control strategy for improving the dynamic performance of interleaved DC–DC SEPIC PFC converter. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-018-3944-9>
24. Ramesh Babu N, Jagan Mohan B (2017) Fault classification in power systems using EMD and SVM. *Ain Shams Eng J* 8(2):103–111
25. Thirumala K, Prasad MS, Jain T, Umarikar AC (2018) Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances. *IEEE Trans Smart Grid* 9(4):3018–3028
26. Zhi-qiang J, Hang-guang F, Ling-jun LJ (2005) Support vector machine for mechanical faults classification. *Zhejiang Univ Sci A* 6:433. <https://doi.org/10.1007/BF02839412>
27. Abdelgayed TS, Morsi WG, Sidhu TS (2018) A new harmony search approach for optimal wavelets applied to fault classification. *IEEE Trans Smart Grid* 9(2):521–529
28. Karmacharya IM, Gokaraju R (2018) Fault location in ungrounded photovoltaic system using wavelets and ANN. *IEEE Trans Power Deliv* 33(2):549–559
29. Abdullah A (2018) Ultrafast transmission line fault detection using a DWT-based ANN. *IEEE Trans Ind Appl* 54(2):1182–1193
30. Ben Abid F, Zgarni S, Braham A (2018) Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and aiNet-DAG SVM. *IEEE Trans Energy Convers* 33(4):1692–1699
31. Yu JJQ, Hou Y, Lam AYS, Li VOK (2019) Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. *IEEE Trans Smart Grid* 10(2):1694–1703
32. Moshtagh J, Rafinia A (2012) A new approach to high impedance fault location in three-phase underground distribution system using combination of fuzzy logic and wavelet analysis. In: *Proceedings of international conference on environment and electrical engineering*, pp 90–97
33. Yi Z, Etemadi AH (2017) Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems. *IEEE Trans Smart Grid* 8(3):1274–1283
34. Tonelli-Neto MS, Decanini JGMS, Lotufo ADP, Minussi CR (2017) Fuzzy based methodologies comparison for high-impedance fault diagnosis in radial distribution feeders. *IET Gener Transm Distrib* 11(6):1557–1565
35. Mustafa MK, Allen T, Appiah K (2017) A comparative review of dynamic neural networks and hidden Markov model methods for mobile on-device speech recognition. *Neural Comput Appl* 12:3. <https://doi.org/10.1007/s00521-017-3028-2>
36. Taheri S, Mammadov M (2013) Learning the Naive Bayes classifier with optimization models. *Int J Appl Math Comput Sci* 23(4):787–795
37. Penny WD, Roberts SJ (1999) Bayesian neural networks for classification: How useful is the evidence framework? *Neural Netw* 12:877–889
38. Cabestany J, Prieto A, Sandoval F (2005) *Computational intelligence and bioinspired systems*. Springer, Berlin
39. Boracchi G, Iliadis L, Jayne C, Likas A (2017) *Engineering applications of neural networks*. Springer, Berlin
40. Monsef H, Lotfifard S (2007) Internal fault current identification based on wavelet transform in power transformers. *Electr Power Syst Res* 77(2007):1637–1645
41. Daubecheis I (1992) *Ten lectures on wavelets*, vol 61. SIAM, Philadelphia

42. Reddy MJB, Mohanta DK (2008) Performance evaluation of an adaptive-network-based fuzzy inference system approach for location of faults on transmission lines using Monte Carlo simulation. *IEEE Trans Fuzzy Syst* 16(4):909–919
43. Silva KM, Souza BA, Brito NSD (2006) Fault detection and classification in transmission lines based on wavelet transform and ANN. *IEEE Trans Power Deliv* 21(4):2058–2063
44. Vyas BY, Das B, Maheshwari RP (2016) Improved fault classification in series compensated transmission line: comparative evaluation of Chebyshev neural network training algorithms. *IEEE Trans Neural Netw Learn Syst* 27(8):1631–1642
45. Jamehbozorg A, Shahrtash SM (2010) A decision-tree-based method for fault classification in single-circuit transmission lines. *IEEE Trans Power Deliv* 25(4):2190–2196
46. Malik H, Sharma R (2017) Transmission line fault classification using modified fuzzy Q learning. *IET Gener Transm Distrib* 11(16):4041–4050
47. Salehi M, Namdari F (2018) Fault classification and faulted phase selection for transmission line using morphological edge detection filter. *IET Gener Transm Distrib* 12(7):1595–1605
48. Mahmud MN, Ibrahim MN, Osman MK (2018) A robust transmission fault classification scheme using class-dependent feature and 2-tier multilayer perceptron network. *Electr Eng* 100:607–623
49. Alsafasfeh Q, Abdel-Qader I, Harb A (2012) Fault classification and localization in power systems using fault signatures and principal components analysis. *Energy Power Eng* 4(6):506–522
50. Mishra PK, Yadav A (2019) Combined DFT and fuzzy based faulty phase selection and classification in a series compensated transmission line. *Modell Simul Eng* 2019:1–18
51. Samet H, Shabanpour-Haghighi A, Ghanbari T (2017) A fault classification technique for transmission lines using an improved alienation coefficients technique. *Int Trans Electr Energy Syst* 27:1–23
52. Gomes DPS, Ozansoy C, Ulhaq A (2018) High-sensitivity vegetation high-impedance fault detection based on signals high-frequency contents. *IEEE Trans Power Deliv* 33(3):1398–1407

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.