DIFFERENTIAL EVOLUTION ALGORITHM SUPPORTED RANDOM FOREST CLASSIFIER FOR EFFECTIVE FEATURE SELECTION AND CLASSIFICATION OF POWER QUALITY DISTURBANCES

S. VIDHYA[1], M. BALAJI[2], E. FANTIN IRUDAYA RAJ[3], V. KAMARAJ[4]

The rise in the usage of power electronic equipment, which induces a deviation in the voltage and current waveforms, has raised the relevance of power quality (PQ) issues. It is important to classify them accurately in order to address the PQ problems. One of the most effective methods to classify PQ problems is to combine Machine Learning (ML) algorithms with efficient feature extraction and feature selection techniques. In the present work, the S-transform-based feature extraction technique is adopted and extracts around 20 features from the PQ disturbance signal. And then, Differential Evolution (DE) based feature selection technique is utilized for feature selection and five optimized features are selected and given as input to the ML classifiers. It is a method that is essential in data mining in order to reduce dimensionality and enhance the performance of ML classifiers. Eight different ML classifiers namely the K-Means, Adaptive Boosting (AdaBoost), K- nearest neighbor (KNN), Naive Bayes, Recursive Support Vector Machine (R-SVM), Linear SVM, Decision Tree, and Random Forest (RF) algorithms have been employed in this work. The effectiveness of the ML Classifiers is measured in terms of various performance evaluation measures namely precision, accuracy, F1-score, recall, and specificity. The results revealed that the DE-supported RF classifier is more efficient when compared with other classifiers in classifying the PQ disturbances even in environments with noise inference.

Keywords: Power Quality Disturbances, S-Transform, Feature Selection, Differential Evolution, Random Forest

1. Introduction

The Power quality events are attributed to a variety of factors that includes the usage of power devices, motor drives, capacitor, and inductor fast switching. The events affect the reliable power supply to the consumers and may cause damage to the equipment [2-3]. Hence it is imperative to recognize the events so as to implement decisive measures. Numerous investigations are being conducted to classify power quality events effectively. Extraction of features, feature selection, and classification are the three key stages in classification.

With respect to the first step of signal analysis and feature extraction, the authors have applied the following techniques.

[1] Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, E-mail: vidhyas2@srmist.edu.in

[2,4] Department of Electrical and Electronics Engineering, Sri Sivasubramaniya Nadar College of Engineering, Chennai

[3] Department of Electrical and Electronics Engineering, Dr. Sivanthi Aditanar College of Engineering, Chennai

- S-transform
- Short-time Fourier transform
- Wavelet transform
- Hilbert – Huang transform

The performance of these techniques has been extensively analyzed and has been effective in the classification of power quality disturbances. The merits and demerits of the techniques have been detailed in [8]. The second important step in the classification of disturbance is feature selection. The features extracted from stage 1 are analyzed and the selection of optimal features is bound to give desired results. In this perspective, the authors have devised an optimal feature selection algorithm for general classification problems which may be explored for efficient classification of power quality disturbances. One of the major challenges in the feature selection algorithm is to avoid redundancy of the features. The authors of [10] have proposed a two-stage process for feature selection. In the first stage, the feature subset is formed from the best features from the genetic algorithm, sequential forward selection, and maximum relevance minimum redundancy algorithm.

In the second stage, the feature is ranked and the subset with the highest ranking is selected. The feature selection based on the Ant Colony method, binary PSO method, and binary GA method has been widely explored. In [9] the authors have proposed Differential Evolution (DE) based method for feature selection. The main advantage of this method is that uses relative feature distribution to help in the replacement of duplicated features. The superiority of this algorithm over the other methods has been investigated and the results indicate the superiority of the algorithm. Hence this work engages the DE-based feature selection method for efficient classification of power quality disturbances. Once the optimal features are selected, the classification can be performed using the conventional decision tree method or AI-based method. Some of the prominently used AI-based methods include Artificial Neural networks, Support Vector machines, Extreme learning machines, and convolution neural networks [3,5-7,11,15].

This work explores the application and effectiveness of the random forest classifiers for classifying the PQ disturbances. In addition, the performance measures of K-Means, Adaptive Boosting (AdaBoost), K- nearest neighbor (KNN), Naive Bayes, Recursive Support Vector Machine (R-SVM), Linear SVM, Decision Tree, and Random Forest (RF) algorithms in classifying the PQ disturbances are presented to highlight the effectiveness of the RF classifier with the selected feature subset in classifying PQ problems.

2. Methodology

The overall approach of the current work is depicted in Fig. 1. The entire work process is categorized as testing and training phases. In the phase of training, the S-transform is applied over the raw PQ test data, and by means of applying the statistical analysis on the S-matrix, there are around 20 features (F1 to F20) are selected. To improve the performance of the machine learning classifier, the feature selection technique is adopted. For this purpose, Differential Evolution based optimization technique is adopted and around five important features (F3, F4, F5, F9, and F19) for effective PQ detection are alone selected. These features are given as input to the ML classifier for training purposes. In the testing phase, the same similar steps are carried out over the PQ test data. The ML classifier based on its training classifies the type of PQ disturbance that occurs in the system more effectively. The various machine learning algorithms are utilized in the present work and their effectiveness in classifying the PQ disturbance is evaluated by means of various performance metrics.

3. S-Transform Based Feature Extraction.

The S-transform [12] has been widely applied to analyze and detect power quality events. It is a combination of wavelet and short-term Fourier transforms. The S-transform for the signal y(t) is defined as

$$ST(\tau,f) = \int_{-\infty}^{\infty} y(t) gu_f\,(\tau - t) e^{-i2\pi f t} \tag{1}$$

Where, $gu_f(\tau)$is the gaussian modulation function, which is defined as

$$gu_f(\tau) = \frac{|f|}{\sqrt{2\pi}} e^{-\left(\frac{\tau^2}{2\sigma^2}\right)} \tag{2}$$

In which, $f \in \Re$indicates the linear frequency and,

$$\sigma = \frac{1}{|f|} \tag{3}$$

By means of Substituting equations (2) and (3) in equation (1), the expression becomes

$$ST(\tau,f) = \int_{-\infty}^{\infty} y(t) \frac{|f|}{\sqrt{2\pi}} e^{-\left(\frac{(\tau-t)^2 f^2}{2}\right)} e^{-i2\pi f t} \tag{4}$$

The Fast Fourier transform is used to calculate the discrete version. The output columns and rows of the S transform represent time and frequency respectively. By using statistical mathematical formulas, the feature vectors are extracted from the S transform output. In this work S transform is applied to classify eight combined power quality disturbances namely sag, sag with harmonics, swell, swell with harmonics, oscillatory transient, harmonics, flicker, and interruption. The various power quality (PQ) disturbance waveforms are depicted in Fig. 2.

For each of these disturbances, samples are generated using standard equations [7]. It is important to analyze the effectiveness of the classifier in classifying the disturbances with noise. Hence the signals are mixed with 30-, 40-, and 50-dB noise, and a separate data set are generated to study the influence of the selected features and classifiers.

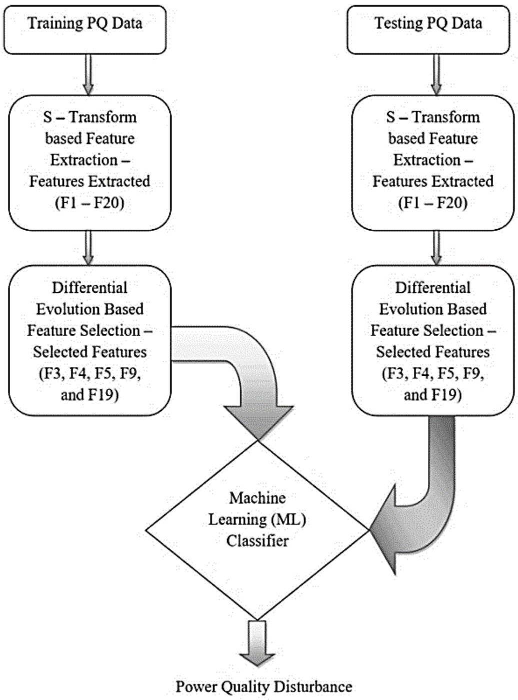

Fig. 1 Flowchart of the proposed approach

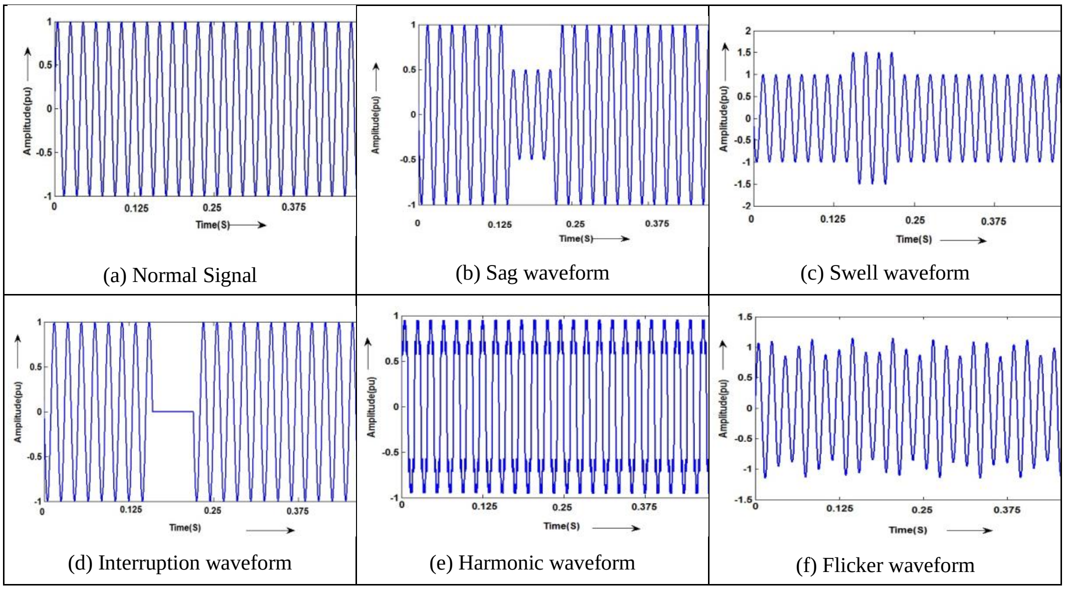

Fig. 2 Different Power Quality Disturbances (a – f).

A total of 4430 data sets comprising the disturbances are generated. These data sets are categorized into training, and testing data sets in the ratio of 80:20

respectively. In this work, 20 features are extracted by applying statistical techniques on the S-matrix [7]. The features considered in this work are tabulated in Table 1.

Table 1

Features Extracted from PQ Disturbance signal after Applying S-Transform and Statistical Techniques.

Name of the Feature	Features
Minimum of Time maximum amplitude (TmA) plot	F1
Maximum of the TmA plot	F2
The standard deviation (σ) of the TmA plot	F3
Mean of TmA plot	F4
Absolute (F2-F1-1/2)	F5
Mean of frequency maximum amplitude plot	F6
Minimum amplitude of the FmA plot (FmA – Frequency Maximum Amplitude)	F7
The σ of the FmA plot (LF (low frequency) area)	F8
Difference (F8 – F7)	F9
The maximum amplitude of the FmA plot	F10
The σ of the Fstd-plot (HF (high frequency) area)	F11
The σ of the FmA plot (HF area)	F12
Skewness factor	F13
Mean of the Fstd-plot	F14
The σ of the Fstd-plot (LF area)	F15
Kurtosis factor	F16
The σ of the Fstd-plot (Overall)	F17
Mean of contour having the largest frequency amplitude of time-frequency (TF) contour	F18
The σ of contour having the largest frequency amplitude of TF contour	F19
Total harmonic distortion.	F20

The differential evolution algorithm is applied to select the best features from these features for the accurate classification of power quality disturbances.

4. Differential Evolution – Based Feature Selection

Storn and Price introduced the differential evolution (DE) algorithm, which has been used to solve a number of engineering-related issues [13–14]. The merits of the algorithm include lesser control parameters and high convergence characteristics.DE algorithm is a combination of simple arithmetic operators along with genetic operator’s recombination, mutation and selection.

DE is a population-based methodology where the target vectors are members who are currently a part of the population. By combining the weighted difference of the target vector with a third vector, a mutant vector is generated. The two-cross operation between mutant vector results in a trial vector. The target vector is replaced in the new generation based on the trial vector's fitness value. The population is initialized according to the dimension of the search variable.

The ith vector of the population is given by

$$\overrightarrow{p_\iota(t)} = [p_{i,}1(t), p_{i,}2(t), p_{i,}3(t) \ldots . p_{i,}D(t)] \quad (5)$$

D is the dimension of search space.

The population is initialized within the bounds of search variable. The mutation operation is expressed as,

$$M_{i,j}(t+1) = \; Pr_{0,j}(t) + F(Pr_{1,j}(t) - Pr_{3,j}(t)) \quad (6)$$

Here $M_i(t)$is the donor vector. Pr1, Pr2, Pr3 are randomly chosen target vectors. And the scaling factor is denoted as F.

After the mutation operation, crossover operation is carried out and is illustrated in the equation below

$$U_{i,j}(t) \quad = M_{i,j}(t) \text{ if (rand (0,1))} < \text{CR} \quad (7)$$
$$= P_{i,j}(t)\text{(rand (0,1))} > \text{CR}$$

$U_{i,j}(t)$is called the trial vector.

The fitness of the target vector and trial vector decides their survival in the next generation.

DE-Based Feature Selection Procedure: The algorithm of DE based feature selection is described below [9]

1. The initial population matrix is formulated with size ($\boldsymbol{NPx\ DNF}$) where $\boldsymbol{NP}$ is the size of the population and $\boldsymbol{DNF}$ is the desired number of features to be selected. Population size is the number of individuals in a population. In this work a population size of 50 is chosen. For five features to be selected the initial population matrix will be 50 * 5.
2. The mutation operation generates new vector form the initial population as described in equation 6.
3. The scaling factor is updated dynamically as given by the equation $F = \frac{G*rand}{\max(Pr_{1,j}, Pr_{2,j})}$, G is a constant lesser than one.
4. Then cross over and mutation operation is performed to determine the parents for next generation.
5. A roulette wheel-based weighting scheme based on the distribution factor ensures appropriate features are selected and there is no duplication of features
6. The process is repeated until the desired number of generations is obtained.

The DE - based feature selection algorithm is executed in MATLAB and five optimum features are selected. The selected features are listed in Table 2. To the machine learning classifiers, these selected features are provided as input. By means of adopting this feature selection technique, the efficiency of the classifiers is well improved.

Table 2

Features Selected after DE-based feature optimization

Name of the selected Feature	Features
The σ of contour having the largest frequency amplitude of TF contour	F19
Difference (F8 – F7)	F9
The standard deviation (σ) of the TmA plot	F4
Absolute (F2-F1-1/2)	F5
Mean of the TmA plot	F3

5. Feature Classification

The power quality event classification with the selected features is tested with different classifiers namely K-Means, Adaptive Boosting (AdaBoost), K-nearest neighbor (KNN), Naive Bayes, Recursive Support Vector Machine (R-SVM), Linear SVM, Decision Tree, and Random Forest (RF) algorithms. The effectiveness of the classifiers is evaluated by means of various evaluation metrics like precision, accuracy, F1-score, recall, and specificity.

Random Forest Classifier: Leo Breiman proposed the ensemble machine learning algorithm known as random forest (RF) [4]. The classifier comprises of many decision trees. The trees are constructed based on a random subset of input training data and a number of classifiers. The forest expands until it meets the user-specified maximum number of trees. The classification is based on the average of class assignment probabilities across the trees. Fig. 3 represents the classification process of the RF classifier in detail.

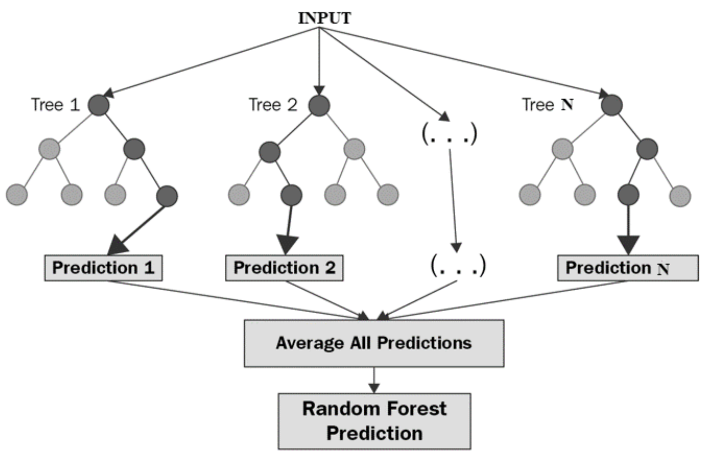

Fig. 3. Random Forest Classifier.

The unlabeled data inputs form the input to the decision trees and each tree yields a class membership. The membership class with maximum votes is selected. The merits of RF classifier include accurate classification of high dimensionally feature data, reduced generalization error, faster training, and higher accuracy [16].

Performance Evaluation of the Classifier: The performance metrics of different classifiers are listed below:

Accuracy: The performance of the classifiers is compared and illustrated in Fig. 4. Classification accuracy is calculated by dividing the number of accurate classifications by the total number of samples. The results indicate that the random forest classifier yields higher accuracy of 99.7%.

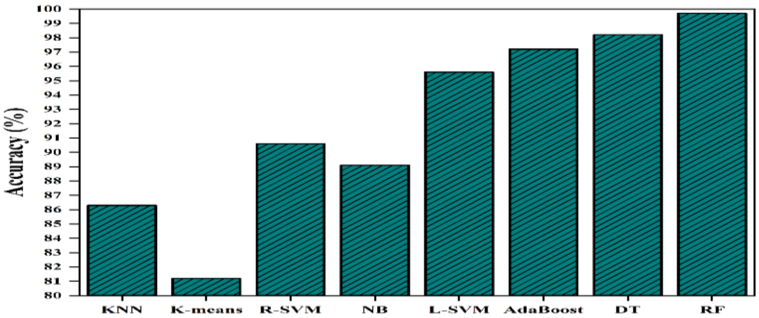

Fig. 4. Accuracy of different classifiers

The confusion matrix of the random forest classifier with the validation data set is presented in Fig. 5. From the figure, it is clear that with the selected features the random forest classifier exhibits good classification accuracy.

Precision: Precision is a class-specific performance and is defined as:

$$Precision = \frac{True\_Positive}{\text{True\_positive} + \text{False\_Positive}}$$

	A	B	C	D	E	F	G	H	I
A	159	0	0	0	0	0	0	1	0
B	0	192	0	0	0	0	0	0	0
C	0	0	160	0	0	0	0	0	0
D	0	0	0	10	0	0	0	0	0
E	0	1	0	0	225	0	0	0	0
F	0	0	0	0	0	159	1	0	0
G	0	0	0	0	0	1	215	0	0
H	0	0	0	0	0	0	0	135	1
I	0	0	0	0	0	0	0	1	215

A- Flicker, B- Harmonic, C- Interruption, D- Normal, E- Oscillatory,
F- Sag, G- Sag with Harmonics, H- Swell, I- Swell with Harmonics

Fig. 5. Confusion matrix of Random Forest classifier in classifying PQ Disturbances.

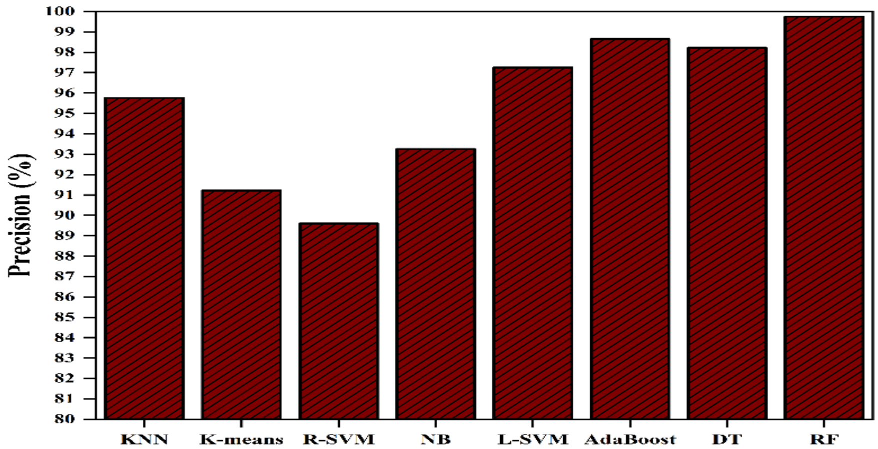

Fig. 6. Precision of different classifiers.

Recall: The percentage of samples in a class that are correctly classified is known as recall. It is calculated as

$$Recall = \frac{True\_Positive}{True_{Positive} + False\_Negative}$$

The Recall measure of different classifiers is depicted in Fig. 7. The recall score of the random forest classifier is 99.75%.

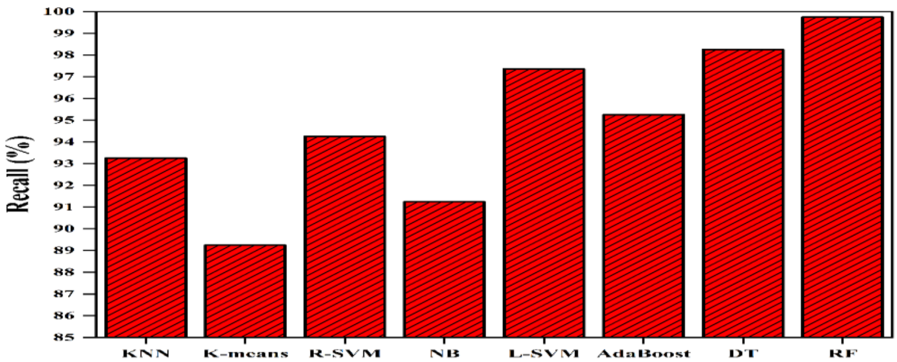

Fig. 7 Recall score of different classifiers.

F1-Score: F1 Score is a performance metric that combines precision and recall

$$F1 - Score = \frac{2 * Precision * Recall}{(\text{Precision} + \text{Recall})}$$

The F1-Score various classifiers is tabulated in Table 3 and the tabulated values signify the efficacy of the random forest classifier:

Table 3

F1-Score of different classifiers

Classifier	**Specificity (%)**
KNN	95.1
K-means	96.2
R-SVM	97.5
NB	93.2
L-SVM	95.9
AdaBoost	96.1
DT	99
RF	99.3

Table 4

Specificity Score of different classifiers

Classifier	Specificity (%)
KNN	96.2
K-means	97.1
R-SVM	93.2
NB	89.6
L-SVM	96.5
AdaBoost	94.6
DT	97.3
RF	99.3

Specificity: Specificity is the proportion of the true negatives correctly identified by the classifier.

$$Specificity = \frac{True\_Negative}{\mathrm{True_{Negative}} + \mathrm{False\_Positive}}$$

The specificity value of various classifiers is illustrated in Table 4 and the tabulated values signify the efficacy of the RF classifier. The overall results indicate that the RF classifier exhibits better performance metrics when compared with other classifiers.

6. Methodology for three-phase networks

The proposed approach can be extended for classifying the faults in three phase systems by following the below mentioned methodology in the training phase Step1: Generate the three phase voltage signals for balanced and unbalanced condition :

Step 2: Apply S-transform and extract the features

Step 3: Train the random forest classifier with the extracted features

Step 4: Fault detection and classification

In the testing phase, the same similar steps are carried out over the three phase test data. The random forest classifier based on its training classifies the disturbance in the form of unbalance.

7. Conclusion

In this work, the optimal features from the S-transform are chosen using a differential evolution-based feature selection approach. Out of the twenty features evaluated from S-transform, five features that distinguish various power quality events are selected. The selected features form the input to the random forest classifier. The results indicate that the selected features accurately classify different power quality events. The performance metrics of the random forest classifier are compared with other classifiers namely the K-Means, Adaptive Boosting (AdaBoost), Recursive Support Vector Machine (R-SVM), Linear SVM, Naive Bayes, Decision Tree, and K- nearest neighbor (KNN) algorithms. The results indicate that the random forest classifier exhibits better performance metrics in classifying the disturbances with an accuracy of 99.7%, F1-score of 99.3%, recall value of 99.75%, specificity of 99.3%, and precision of 99.75%. The work is further to be extended to classify multiple power quality events with optimal features selected from hybrid time-frequency analysis techniques.

R E F E R E N C E S

[1] *Alessia Sarica, Antonio Cerasa and Aldo Quattrone*. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review.Frontiers in Aging Neuroscience,2017,9, https://doi.org 10.3389/fnagi.2017. 00329.

[2] *AlperYılmaz, Ahmet Küçüker, GökayBayrak, DavutErtekin, MiadrezaShafie-Khah, Josep M. Guerrero*.An improved automated PQD classification method for distributed generators with hybrid SVM-based approach using un-decimated wavelet transform.International Journal of Electrical Power & Energy Systems,2022,136,107763. https://doi.org/10.1016/ j.ijepes.2021.107763.

[3] *ArangarajanVinayagam, VeerapandiyanVeerasamy, Padmavathi Radhakrishnan, MaheswariSepperumal, KalaivaniRamaiyan* .An ensemble approach of classification model for detection and classification of power quality disturbances in PV integrated microgrid network. Applied Soft Computing, 2021 106, 107294.https://doi.org/10.1016/ j.asoc.2021.107294.

[4] *Breiman,L.*Random Forests. Machine Learning.2001,45,5–32. https://doi.org/10.1023/A: 1010933404324

[5] *Yunpeng Gao, Yunfeng Li, Yanqing Zhu, Cong Wu, Dexi Gu*. Power quality disturbance classification under noisy conditions using adaptive wavelet threshold and DBN-ELM hybrid model,Electric Power Systems Research,2022, 204 107682.https://doi.org/ 10.1016/j.epsr.2021.107682.

[6] *Yuwei Zhang, Yin Zhang, Xiaohua Zhou,*Classification of power quality disturbances using visual attention mechanism and feed-forward neural network. Measurement,2022,188,110390. https://doi.org/10.1016/j. measurement.2021.110390.

[7] *Nantian Huang, Dianguo Xu, Xiaosheng Liu & Lin Lin*. Power quality disturbances

classification based on S-Transform and probabilistic neural network Neurocomputing,2012,98, 12-23, https://doi.org/10.1016 /j.neucom.2011.06.041

[8] *PorasKhetarpal, Madan Mohan Tripathi*. A critical and comprehensive review on power quality disturbance detection and classification. Sustainable Computing: Informatics and Systems,2020,28,100417.https://doi.org/10.1016/j.suscom. 2020.100417.

[9] *Rami N. Khushaba, Ahmed Al-Ani, Adel Al-Jumaily* .Feature subset selection using differential evolution and a statistical repair mechanism. Expert Systems with Applications,2011,38(9),11515-11526.https://doi.org/10.1016/j.eswa.2011.03.028.

[10] *Sadegh Jamali, Ali Reza Farsa, NavidGhaffarzadeh* . Identification of optimal features for fast and accurate classification of power quality disturbances. Measurement, 2018,116,565-574.https://doi.org/10.1016/j.measurement.2017.10.034.

[11] *Shouxiang Wang, Haiwen Chen,* .A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network. Applied Energy,2019,235,1126-1140.https://doi.org/10.1016/j.apenergy.2018.09.160.

[12] *Stockwell, RG, Mansinha, L & Lowe, RP*. Localization of the complex Spectrum: the S-transform. IEEE Transactions on Signal Processing, 1996, 44, 998-1001, 10.1109/78.492555

[13] *Storn, R. and Price, K*. “Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces”, Technical Report TR-95-012.Berkeley, International Computer Science Institute, USA, 1995.

[14] *Storn, R. and Price, K*. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization,1997,11(4), 341-359.https://doi.org/10.1023/A:1008202821328

[15] *Suhail Khokhar, Abdullah AsuhaimiMohd Zin, Aslam P Memon & Ahmad Safawi Mokhtar*. A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network.Measurement, 2017, 95, 246-259, https://doi.org/10.1016/ j. measurement. 2016.10.013

[16] *SuryaniDyahAstuti, Mohammad H. Tamimi, Anak A.S. Pradhana, Kartika A. Alamsyah, HeryPurnobasuki, MiratulKhasanah, Yunus Susilo, KuwatTriyana, Muhammad Kashif, ArdiyansyahSyahrom,*Gas sensor array to classify the chicken meat with E. coli contaminant by using random forest and support vector machine. Biosensors and Bioelectronics: X,2021, 9,100083. https://doi.org/10.1016/j.biosx.2021.100083.

[17] *Vidhya.S , Kamaraj.V,* Particle swarm optimized extreme learning machine for feature classification in power quality data mining. Automatika, 2017,58,4, 487-494. DOI: 10.1080/00051144.2018.147608