Applied Soft Computing 138 (2023) 110181

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Classification of Power Quality Disturbances using Linear Discriminant Analysis

Gurpreet Singh [a,b,*], Yash Pal [a], Anil Kumar Dahiya [a]

[a] NIT Kurukshetra, Haryana, India
[b] ABES Engineering College, Ghaziabad, Uttar Pradesh, India

ARTICLE INFO

Article history:
Received 27 March 2022
Received in revised form 3 January 2023
Accepted 27 February 2023
Available online 6 March 2023

Keywords:
Machine learning
LDA
Power Quality Disturbances
Random Forest
KNN
SVM
Naive Bayes

ABSTRACT

With the expansion of renewable energy sources and its integration with the grid, power semiconductor devices applications has been increased rapidly which causes many Power Quality Disturbances (PQDs). These disturbances can cause significant losses in the distribution system, therefore it is essential to recognize and mitigate these disturbances timely. The present work proposes a novel method for the classification of single stage and multiple PQDs based on dimensionality reduction. The main objective is to transform the data set from higher dimensional space to lower dimensional space by eliminating the unnecessary features. This paper proposes a supervised learning dimensionality reduction technique, i.e. Linear Discriminant Analysis (LDA) for the dimensionality reduction of the data set of twenty nine types of PQDs including single stage and multiple PQDs. In this technique, the ratio between class variance and within class variance are maximized for maintaining the maximum class separability, to obtain a lower dimensional space of the features. The performance of LDA is analysed with four type of machine learning classifiers such as k-Nearest Neighbour (KNN), Naive Bayes (NB), Support Vector Machine (SVM) and Random Forest (RF). Classification results show that the higher classification accuracy is achieved for the twenty nine types of PQDs under different noise levels (20 dB to 40 dB).

1. Introduction

Nowadays, with evolution in the areas of power electronics technology, applications of non-linear loads are increasing rapidly, which have deteriorated the quality of power in power system. These Power Quality (PQ) problems causes capacitor bank failure, increase in losses at distribution level and electrical machines, vibrations, negative sequence currents in electrical machines, mainly rotor heating, de rating of cables, and so on [1].

It is necessary to feed pure sine wave current to the load from the distribution level of power system, but this is not observed when the power electronic converter-based loads are connected to the system. In addition to various benefits of these power electronics based loads, they gave rise to other PQDs. These disturbances instigated the execution of standards and guidelines such as IEEE 519 [2], for governing harmonics in the power system along with the prescribed limits [3]. Due to the increase in power demand, renewable energy resources are also integrating with the grid which are also causing the PQDs.

Numerous incidents of PQDs can cause inconvenience and economic losses to users. For the mitigation of PQDs, an uninterrupted monitoring of power signals is needed which is a tedious work and cannot be done manually [4]. In this context, it is essential to precisely classify the occurred PQDs. A suitable, fast and perfect classification algorithm can be applied for real time monitoring system of power grid [5].

PQDs identification includes features extraction of PQDs and classification from extracted features. Selection of relevant features must be ensured for getting higher efficiency of the classifier. PQDs can consist of stationary signals (flickers), non stationary signals (oscillatory transients and notches) and combination of both. Therefore many feature extraction tools are used in literature: Fast Fourier Transform (FFT) [6], Stockwell Transform (ST) [7–10], Wavelet Transform (WT) [11–14], Gabor Transform (GT) [15], Hilbert Transform (HT) [16], Empirical Mode Decomposition (EMD) [17–19], Short Time Fourier Transform (STFT) [20], Singular Value Decomposition (SVD) [21], Kalman Filter (KF) [22].

FFT does not give good performance while monitoring the PQDs that are non-stationary due the dependency on the resolution of window size. WT is noise-sensitive and non-suitable for the proper selection of basic wavelet functions, further the unique features of the PQDs cannot be obtained directly. Its

* Corresponding author at: ABES Engineering College, Ghaziabad, Uttar Pradesh, India.
E-mail address: gurpreet.singh@abes.ac.in (G. Singh).

https://doi.org/10.1016/j.asoc.2023.110181
1568-4946/

performance can be improved by combining it with other optimization techniques, which increases computational burden. Although ST is the improved version of WT and the features extracted from ST can be distinguished, recognized and resistant to noise, due to its computation complexity, its application in real time environments are limited [23]. GT has fixed window size which results in fixed resolution, which makes it non-suitable for simultaneous PQDs [15]. In literature EMD is combined with HT for time–frequency analysis, recognition and classification of PQDs. On adding random noise to original signal efficiency of the system increases but there is an increase in computational burden [24]. STFT break down the non-stationary signals in to the time–frequency domain by concatenating the stationary signals with the sliding window. As the window size remain same for all frequencies, it is not able to track the dynamics of the signals properly [25].

Many classification techniques of PQDs from the extracted features are given in the literature. This mainly includes Artificial Neural Network (ANN) [26–28], Decision Tree [7,23,26,29,30], Fuzzy Logic [31], Bayesian Classifier [32], Expert System [33], Rule-Based System [34], Support Vector Machine (SVM) [8,12,15,35]. Classifier must have two major characteristics one is adaptation for all type of disturbances that may occur and must have higher computational accuracy.

Sometimes due to the high dimensionality of the data, "curse of dimensionality" problem may occur. To solve this issue dimension reduction is employed to extract low-dimensional feature set from a high-dimensional feature set. Canonical Correlation Analysis (CCA) is used for feature extraction in multi-view learning [36]. It is used on datasets having two views. Two projections are obtained each for each of the view such that the correlation among them can be maximized when data is projected into common subspace. It is an unsupervised learning method as class labels are not utilized. As CCA can not be used for data having more than two views, to incorporate these cases the Multi-view Canonical Correlation Analysis (MCCA) was proposed [37]. For MCCA, total correlations between any of the two views are maximized to obtain the projection for each view. As discriminant informations are not taken into account which may causes lower accuracy during classification. LDA is a supervised learning technique for single-view learning. In this case an optimal linear transformation can be obtained to map the data into a subspace by minimizing the within class distance and maximizing the distance between classes simultaneously [38].

In PQDs, differentiation between various disturbance signals which are similar (e.g. interruption, sag and signals with multiple PQDs) is difficult. The signal feature extraction techniques which are employed in the literature, extracts many irrelevant features. To overcome this problem, this paper proposes LDA based data reduction technique for feature extraction. The ratio between class variance and within class variance are maximized for maintaining the maximum class separability, which in result, to obtain a lower dimensional space of the features [39,40].

The main contributions of the proposed work are listed as:

- Different types of PQDs are classified into 29 classes which are more in numbers as considered in the previous articles given in the literature. Also the signals with the noise levels of 20 dB, 25 dB, 30 dB, 35 dB and 40 dB are taken under consideration to check the accuracy of the classifier.
- This paper uses LDA based data reduction techniques to reduce the features of the PQDs which in results reduces the computational burden.
- Further to analyse the reduced PQD data, performance of four different type of classifiers are considered i.e. KNN, SVM, NB and RF. Accuracy of all the classifiers are analysed using confusion matrix under different noise levels.

2. Related works

2.1. Canonical correlation analysis (CCA)

CCA attempts to find a projection pair so that the correlation can be maximized between two-views of datasets in a common subspace with reduced dimensions [36]. Suppose two matrices $Y_1 = [y_{11}, y_{12}, \dots, y_{1n}] \in \Re^{p\times n}$ and $Y_2 = [y_{21}, y_{22}, \dots, y_{2n}] \in \Re^{q\times n}$ are defined to represent two views of data. Each column in these matrices represent a sample. CCA seeks projections w_1 and w_2 one for each view by optimizing the linear correlation coefficient [38] as shown in Eq. (1)

$$\max_{w_1,w_2} \frac{cov(w_1^TY_1, w_2^TY_2)}{\sqrt{var(w_1^TY_1)}\sqrt{var(w_2^TY_2)}} = \frac{w_1^TC_{12}w_2}{\sqrt{w_1^TC_{11}w_1}\sqrt{w_2^TC_{22}w_2}} \tag{1}$$

In Eq. (1), C_{12}, C_{11} and C_{22} are the covariance matrices which can be obtained as shown in Eq. (2)

$$C_{12} = Y_1Y_2^T, C_{11} = Y_1Y_1^T, C_{22} = Y_2Y_2^T \tag{2}$$

Eq. (1) can be written as

$$\max_{w_1w_2} \quad w_1^TC_{12}w_2 \qquad s.t. \quad w_1^TC_{11}w_1 = 1, \quad w_2^TC_{22}w_2 = 1 \tag{3}$$

2.2. Multi-view canonical correlation analysis (MCCA)

CCA has a limitation of maximum two views of data [37]. To overcome this issue, MCCA try to obtain projections by maximizing the summation of pairwise correlations in projection space. For p data samples having v views $Y_i|_{i=1}^{v}$, where $Y_i = [y_{i1}, y_{i2}, \dots, y_{ip}] \in \Re^{h_i\times p}$ is data matrix for ith view. $Y_i|_{i=1}^{v}$ are transformed to $w_i^TY_i|_{i=1}^{v}$ by the projections $w_i|_{i=1}^{v}$. Total correlation in common space can be maximized as

$$\max_{w_1,w_2\dots,w_v} \quad \sum_{i<j} w_i^TC_{ij}w_j \qquad s.t. \quad w_i^TC_{ii}w_i = 1, \quad i = 1, 2, \dots, v \tag{4}$$

In Eq. (4) w_i and w_j represent e ith and jth view projections respectively. $C_{ij} = Y_iY_j^T$ is covariance matrix between ith and jth view, and $C_{ii} = Y_iY_i^T$ is data variance matrix of ith view. As in CCA, the number of samples for each view should remain same in MCCA, also it is an unsupervised technique.

3. Proposed algorithm for dimension reduction

3.1. Linear discriminant analysis (LDA)

It is the supervised approach of dimensional reduction technique. The aim of LDA is to maximize the separability of the known classes in our target variable while at the same time reducing dimensions of the data matrix. This can be attained as: First, between-class variance is calculated, then within-class variance is calculated and finally, a space with lower dimensions is constructed to maximize the between class variance and minimize the within class variance. These steps are explained in the next subsections.

3.2. Between-class variance (S_B)

To demonstrate calculations of S_B following data set has been taken under consideration. A dataset $X = [x_1, x_2\dots.., x_N]$, and ith sample is represented by x_i and the total number of samples are represented by N. Each sample has K features ($x_i \in \Re^K$). Taking an example of three classes for the dataset X, i.e $C = 3$ as, $X = [C_1, C_2, C_3]$ as shown in Fig. 1 [40]. There are five samples

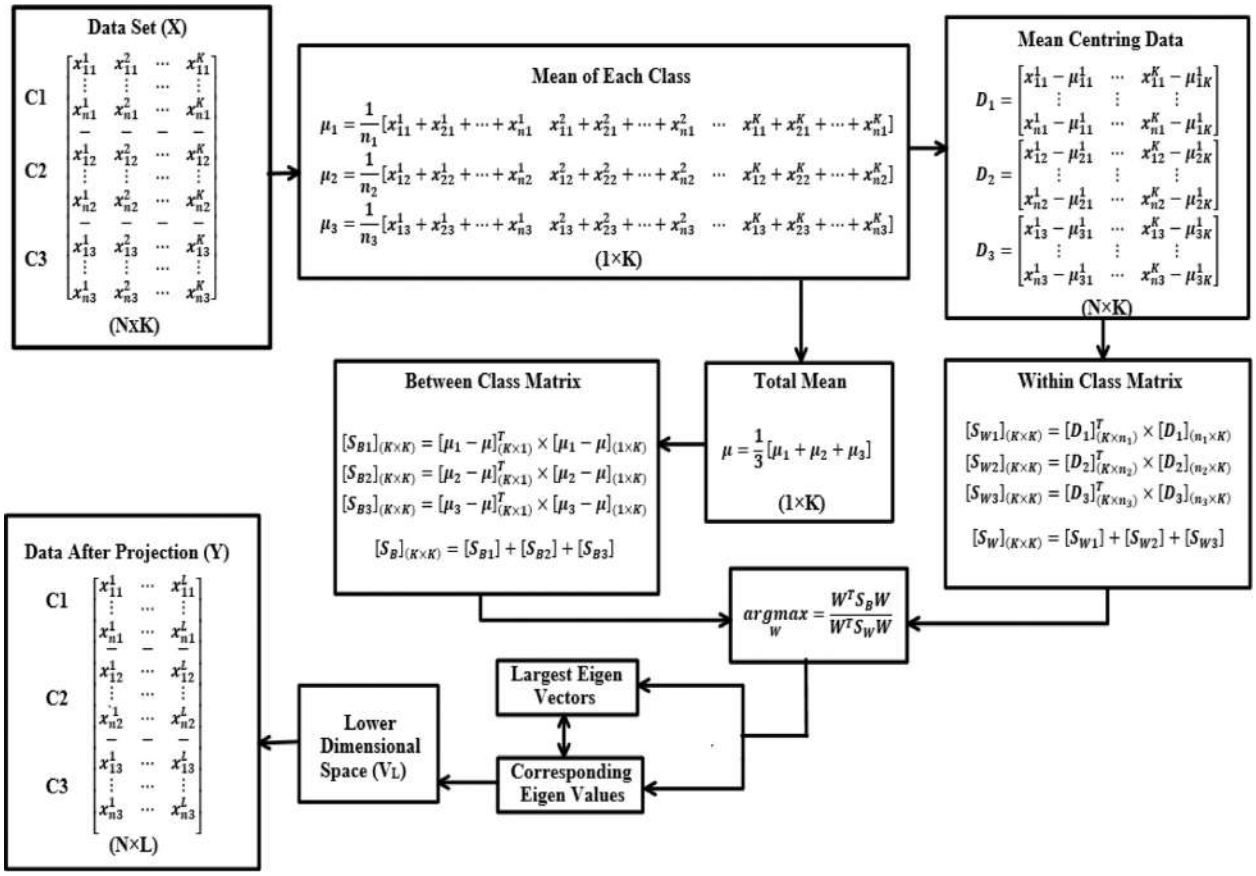


Fig. 1. Block diagram of linear discriminant analysis.

for each class denoted as n_i, number of samples for the ith class. N can be calculated as, $N = \sum_{i=1}^{3} n_i$.

The separation distance $(p_i - p)$ between different classes can be calculated as [41,42]:

$$(p_i - p)^2 = (W^T \mu_i - W^T \mu)^2 = W^T(\mu_i - \mu)(\mu_i - \mu)^T W \tag{5}$$

where, p_i is projection for mean of ith class:

$$p_i = W^T \mu_i \tag{6}$$

and p is projection for total mean of all classes:

$$p = W^T \mu \tag{7}$$

W is transformation matrix of LDA and μ_i(1 X K) is mean of ith class and can be calculated as:

$$\mu_i = \frac{1}{n_i} \sum_{x_i \in C_i} x_i \tag{8}$$

and μ(1 X K) is total mean for all classes:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i = \sum_{i=1}^{C} \frac{n_i}{N} \mu_i \tag{9}$$

Term $(\mu_i - \mu)(\mu_i - \mu)^T$ in Eq. (5) can be denoted as (S_{Bi}), i.e. between class variance for ith class.

$$(m_i - m)^2 = W^T S_{Bi} W \tag{10}$$

S_B for all the classes can be written as:

$$S_B = \sum_{i=1}^{C} n_i S_{Bi} \tag{11}$$

3.3. Within-class variance (S_W)

Taking the case of jth class for calculating the within-class variance (S_{Wj}), first total mean of the jth class is calculated. Than it is subtracted from each sample of jth class [41,42]. As LDA seeks for a lower dimensional space of order C-1, which minimizes the within-class variance. (S_{Wj}) can be obtained as:

$$\begin{aligned}\sum_{x_i \in C_j, j=1\ldots C} (W^T x_i - m_j)^2 &= \sum_{x_i \in C_j, j=1\ldots C} (W^T x_{ij} - W^T \mu_j)^2 \\ &= \sum_{x_i \in C_j, j=1\ldots C} W^T (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T W \\ &= \sum_{x_i \in C_j, j=1\ldots C} W^T S_{Wj} W\end{aligned} \tag{12}$$

where, $W^T x_i$ is the projected samples of each class. From Eq. (12), S_{Wj} is obtained as:

$$S_{Wj} = d_j^T * d_j = \sum_{i=1}^{n_j} (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T \tag{13}$$

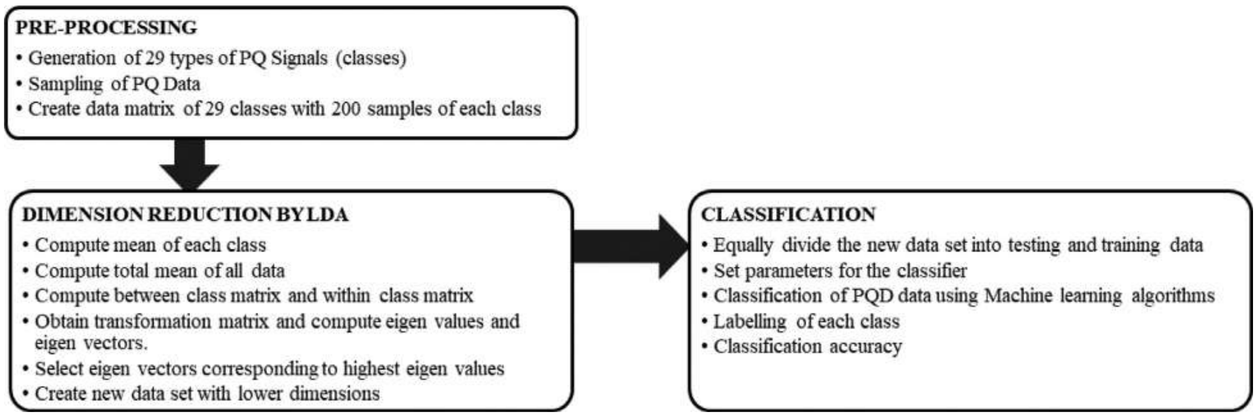


Fig. 2. Proposed method for PQD classification.

where, x_{ij} represents ith sample in jth class and d_j is centring data of the jth class, given by:

$$d_j = C_j - \mu_j = [x_i]_{i=1}^{n_j} - \mu_j \tag{14}$$

S_W for all classes is obtained as:

$$S_W = \sum_{i=1}^{3} S_{Wi} = \sum_{x_i \in C_1} (x_i - \mu_1)(x_i - \mu_1)^T + \sum_{x_i \in C_2} (x_i - \mu_2)(x_i - \mu_2)^T + \sum_{x_i \in C_3} (x_i - \mu_3)(x_i - \mu_3)^T \tag{15}$$

3.4. Creating lower dimensional space

Once the (S_B) and (S_W) are determined, the Fisher criterion can be expressed as:

$$J(W) = \frac{W^T S_B W}{W^T S_W W} \tag{16}$$

where, J(W) is the measure of difference between class variance normalized by a measure of within class variance. To obtain the maximum value of J(W), differentiate Eq. (16) and equate it equal to zero.

$$\frac{\partial}{\partial W} J(W) = \frac{\partial}{\partial W} \left(\frac{W^T S_B W}{W^T S_W W}\right) = 0 \tag{17}$$

$$\Rightarrow (W^T S_W W)(2 S_B W) - (W^T S_B W)(2 S_W W) = 0 \tag{18}$$

Dividing Eq. (18) by $2W^T S_W W$, we get

$$S_W^{-1} S_B W - J(W) W = 0 \tag{19}$$

$$\Rightarrow S_W^{-1} S_B W = \lambda W \tag{20}$$

where, λ = J(W), is the eigen value of the transformation matrix (W).

$$W^* = \underset{W}{argmax}\, J(W) = \underset{W}{argmax}\left(\frac{W^T S_B W}{W^T S_W W}\right) \tag{21}$$

From Eq. (21) information regarding LDA space can be obtained, while eigenvectors ($V = [V_1, V_2, V_3 \ldots\ldots V_K]$) gives directions for new space and eigenvalues ($\lambda = [\lambda_1, \lambda_2, \lambda_3 \ldots\ldots \lambda_K]$) gives the scaling factor, length, or the magnitude of the eigenvectors [43].

As each eigenvector represents one axis of the LDA space, and its associated eigenvalue represents the robustness of this eigenvector, the eigenvectors (V_L) having highest eigenvalues (L) are used to obtain a lower dimensional space while others are neglected.

As shown in Fig. 1, the original data matrix ($X \in \Re^{N \times K}$) is reduced to the lower dimensional space of LDA ($V_L \in \Re^{N \times L}$). Mathematically it can be shown as [44]:

$$Y = X V_L \tag{22}$$

After projection data dimension changes to L, by ignoring the K-L features from each of the samples. Hence, for C–Classes, there will be C-1 projection vectors, therefore Eq. (20) can be generalized for C-classes as:

$$S_W^{-1} S_B W_i = \lambda_i W_i \tag{23}$$

where, $\lambda_i = J(W_i)$ is the scalar quantity and $i = 1, 2, \ldots .C - 1$. Thus the optimal projection matrix W^* is the one whose columns are the eigenvectors corresponding to the largest eigen values as:

$$W^*_{K \times C-1} = \left[W_1^* | \quad W_2^* | \quad \ldots . | \quad W_{C-1}^* | \right] \tag{24}$$

4. Classification methods

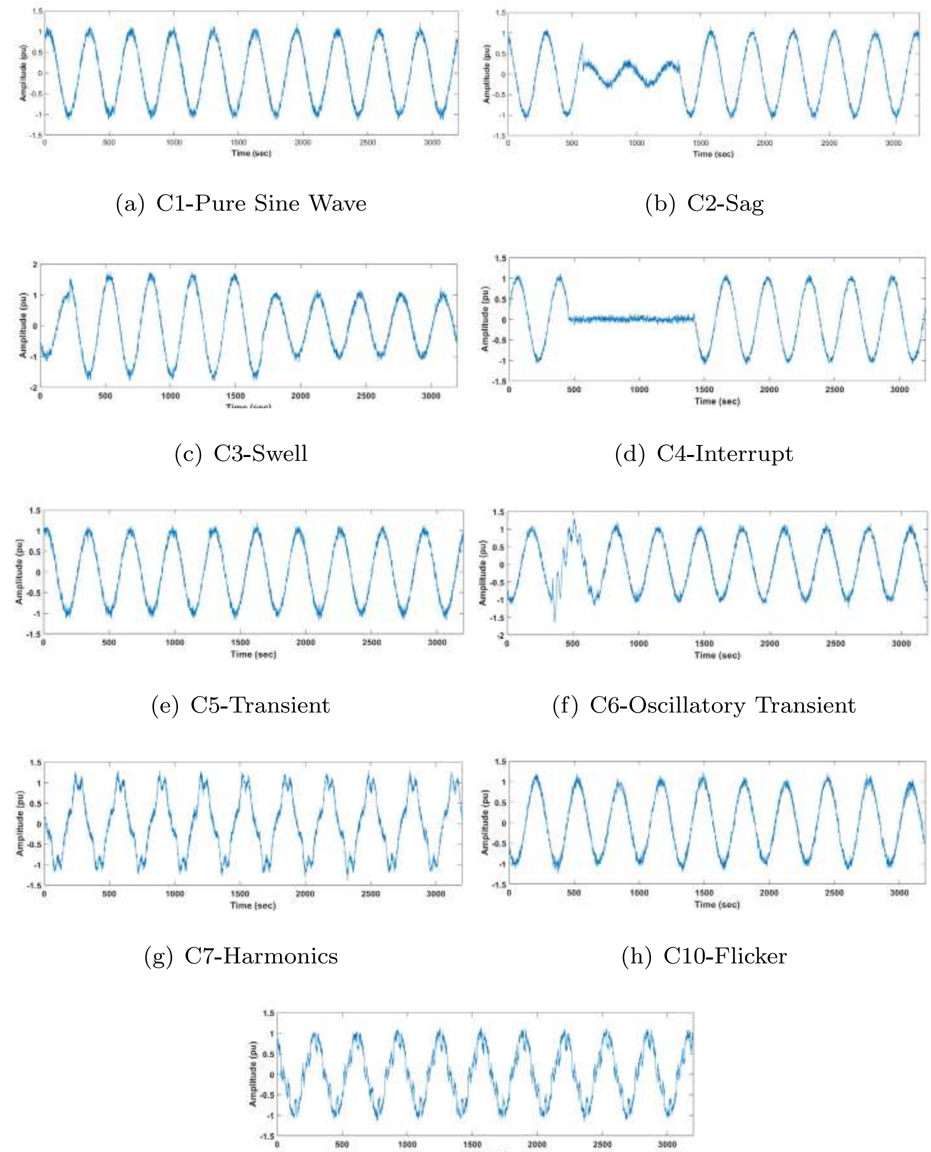
In this section the machine learning classifiers which are used at classification stage of PQDs are briefly explained. In this study the results has taken on KNN, NB, SVM and RF classifiers.

4.1. KNN

It is considered as simplest and powerful machine learning algorithm [45], that uses features distance for the classification of samples in testing. Let us assume that x is a sample which is to be classified using KNN technique. This can be done using the majority voting from its K-neighbours. Probability of sample x to be classified as class y is calculated on the basis of Euclidian distance and can be given as:

$$p(y = j | F = x) = \frac{1}{K} \sum_{i \in A} I(y^i = j) \tag{25}$$

where, F is the total features for all the classes. Value of K should be selected as which gives least classification error.



(a) C1-Pure Sine Wave (b) C2-Sag

(c) C3-Swell (d) C4-Interrupt

(e) C5-Transient (f) C6-Oscillatory Transient

(g) C7-Harmonics (h) C10-Flicker

(i) C17-Notch

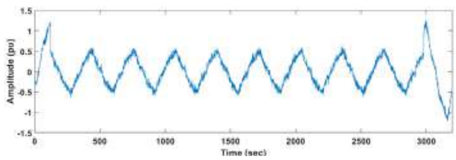
Fig. 3. Samples of single PQDs with 20 dB noise.

4.2. NB

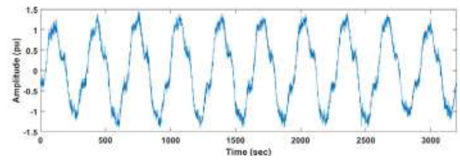
It is established on the basis of Bayes theorem and comes under the category of supervised learning algorithms of machine learning. It predicts on the basis of the probability of an object [46]. As for a particular feature its occurrence is independent from the occurrence of other features, it is known to be Naive and based on Baye's theorem called as Bayes. It generates its own likelihood table by calculating the probabilities of the given features. Further do the classification on the basis of posterior probability given as:

$$y = \underset{z}{argmax}[p(z|F) * \prod_{i=1}^{n} p(y^i|z)] \tag{26}$$

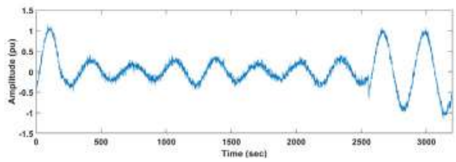
where, argmax gives maximum value of target function for testing sample z and y is predicted class. F is the total features for all the classes and p is posterior probability. As there is three types of Naive Bayes Model such as Gaussian, Multinomial and Bernoulli. Gaussian NB model is used as a classifier in this paper.



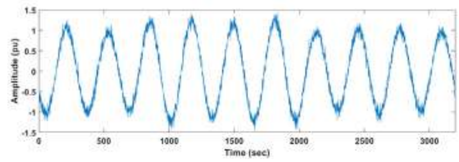
(a) C8-Harmonics with Sag



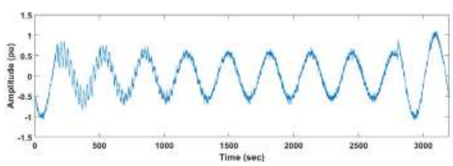
(b) C9-Harmonics with Swell



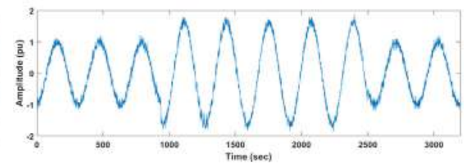
(c) C11-Flicker with Sag



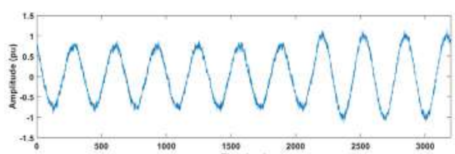
(d) C12-Flicker with Swell



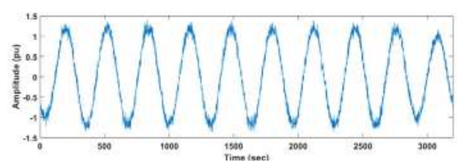
(e) C13-Sag with Oscillatory Transient



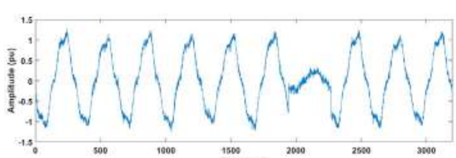
(f) C14-Swell with Oscillatory Transient



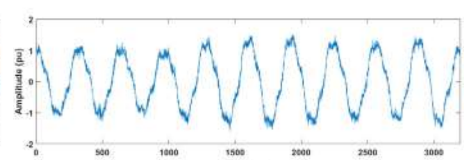
(g) C15-Sag with Harmonics



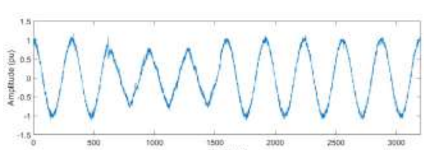
(h) C16-Swell with Harmonics



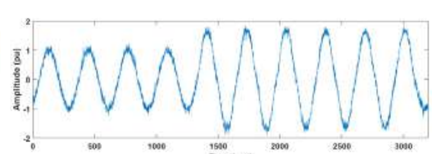
(i) C18-Harmonics with Sag with Flicker



(j) C19-Harmonics with Swell with Flicker



(k) C20-Sag with Harmonics with Flicker



(l) C21-Swell with Harmonics with Flicker

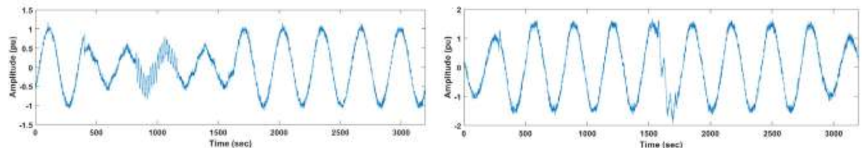
Fig. 4. Samples of multiple PQDs with 20 dB noise.

4.3. SVM

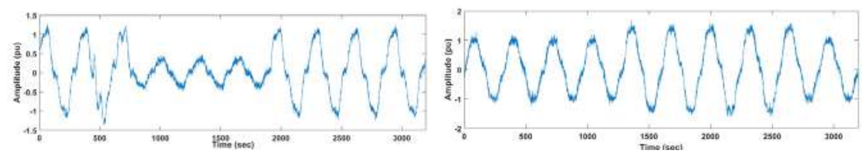
The aim of SVM is to locate the best line or decision boundary known as hyperplane which can categorize n-dimensional space into classes so that the new data points can be correctly categorized [47]. Gap from the margin is maximized to provide the sufficient clarification so that the test samples can be classified with better accuracy. Mathematically, separating hyperplane can be described as:

$$h(Y) = sign(W.Y^T + b) \tag{27}$$

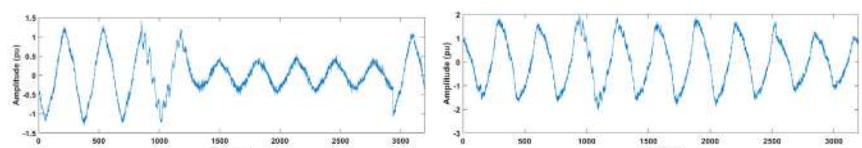
where, Y is sample vector, as $Y = [y_1, y_2 \ldots . y_q]$ having q features, W is weights vector, as $W = [w_1, w_2 \ldots . w_q]$ and b is a scalar bias.



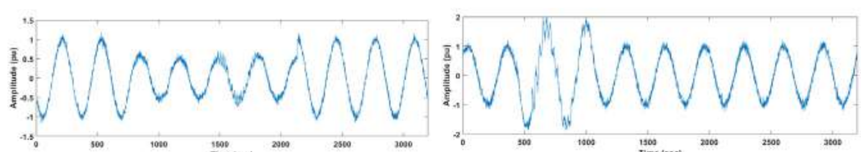
(m) C22-Sag with Harmonics with Oscillatory Transient (n) C23-Swell with Harmonics with Oscillatory Transient



(o) C24-Harmonics with Sag with Oscillatory Transient (p) C25-Harmonics with Swell with Oscillatory Transient



(q) C26-Harmonics with Sag with Flicker with Oscillatory Transient (r) C27-Harmonics with Swell with Flicker with Oscillatory Transient



(s) C28-Sag with Harmonics with Flicker with Oscillatory Transient (t) C29-Swell with Harmonics with Flicker with Oscillatory Transient

Fig. 4. (*continued*).

4.4. RF

It consist of number of decision trees for various subsets of the dataset. It takes the prediction from all the trees and takes the decision and gives the final output on the basis of maximum votes of predictions. Therefore it is also known as ensemble learning algorithm [48]. Ensemble learning algorithms has the advantage of giving low variance and low bias estimation due to its bagging approach. As a result, it shows less dependency on training data and features of a single model. As the model is created from dense randomness, it is robust against over fitting [49].

5. Generation of PQDs

A dataset of different types of PQDs for single and multiple disturbances is required for the purpose of classification. The dataset of PQDs have been created in MATLAB as per the IEEE-1159 standards which is closely portray the real-time data. It is widely used for estimating the efficiency of the classifier in the previous studies. The generation of real PQDs data is limited, which requires long monitoring time simultaneously at number of locations. In this study most of the cases of PQDs i.e. 29 types (consisting single and multiple PQDs) are generated in random fashion by varying the parameters from there numerical models as given in [50].

For the generation of data set, parameters are configured as:

- Samples for each class, $N_s = 200$.
- Sampling frequency, $f_s = 16$ kHz.
- Fundamental frequency, $f = 50$ Hz.
- Number of cycles of the fundamental frequency in each class sample, $N = 10$.
- Amplitude of the signals, $A = 1$ p.u.

This will give the data set with the dimensions of 5800×3200. In order to approach the realistic case and also for the comparison purpose, random noises are added to the generated signals to attain the SNR of 20 dB to 40 dB. A specific sample for nine single PQDs and twenty multiple PQDs with 20 dB noise with their class labels are shown in Figs. 3 and 4 respectively as per the mathematical models and parameters given in [50].

6. Classification of PQDs using LDA

As given in Section 3, to reduce the dimensions of PQDs dataset LDA technique is used. As a result it lowers the calculation

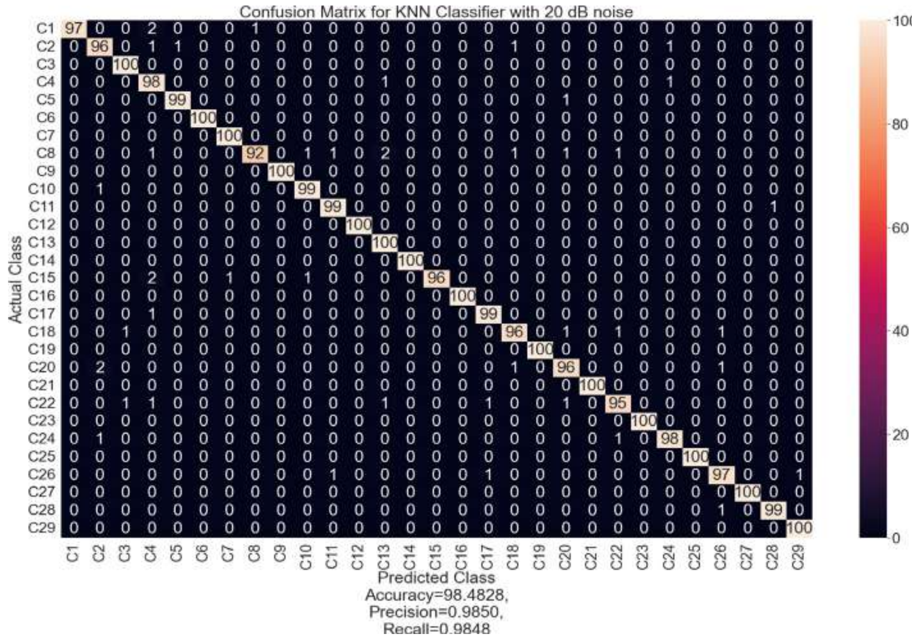


Fig. 5. Confusion matrix for KNN classifier with 20 dB noise.

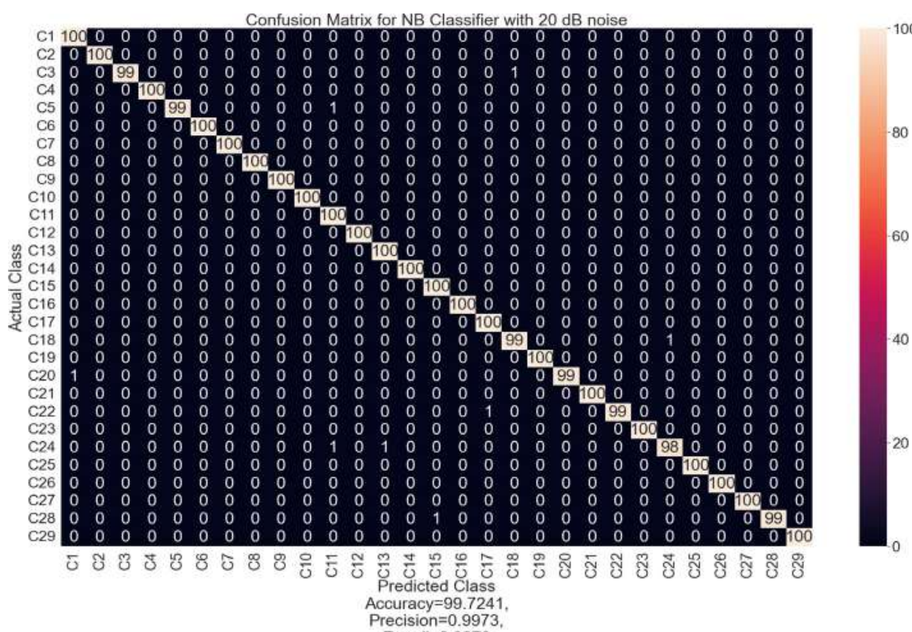


Fig. 6. Confusion matrix for NB classifier with 20 dB noise.

burden because there is no need of any feature extraction tool. For classification of PQDs, the proposed method has three sections such as: Pre-Processing, Dimension reduction and classification as shown in Fig. 2.

6.1. Pre-processing

Twenty nine types of PQDs are generated using the mathematical models given in [50] and the data set is created with 200 samples for each of the classes as per the defined parameters. The name, class number and sample for each class of these generated PQDs are also shown in Figs. 3 and 4. Ten cycles for each class has been taken under consideration with a sampling frequency of 16 kHz, so that every signal has 3200 samples. Therefore data set of 5800 $\times$ 3200 is obtained. The fundamental frequency of the system is considered as 50 Hz. To take the consideration of noisy environment present in an electrical network, Gaussian white noise is added to generated PQDs. Different amount of noises with signal to noise ratio (SNR) of 20 dB to 40 dB are added in each sample of all classes.

6.2. Dimension reduction

In this section, dimensions of the obtained data set has been reduced as per the method explained in Section 3. S_B and S_W are expressed in terms of Fisher criterion, which gives the eigen values and eigen vectors of the transformation matrix. The lower dimensional space is obtained using the eigen vectors corresponding to the highest eigen values while neglecting the other

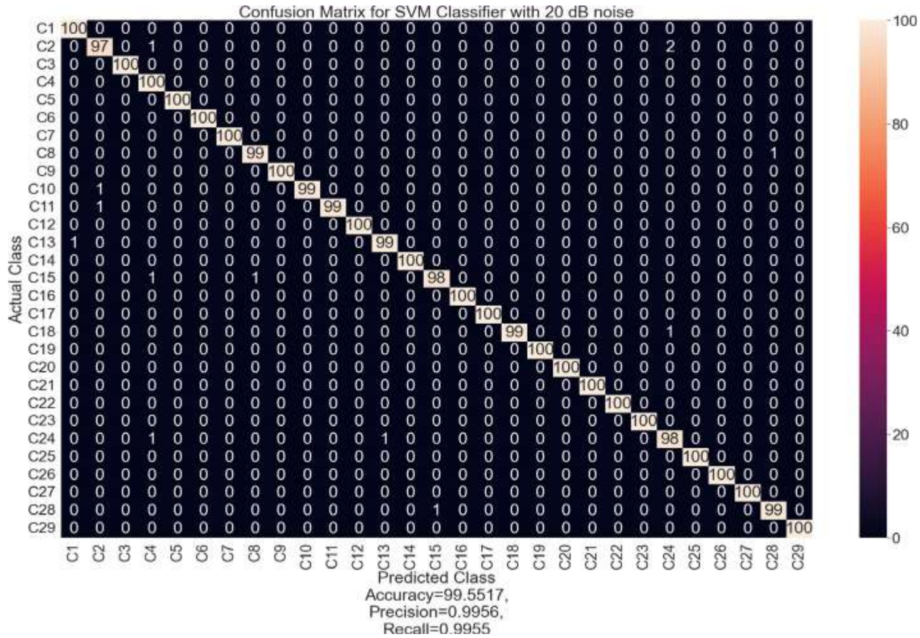


Fig. 7. Confusion matrix for SVM classifier with 20 dB noise.

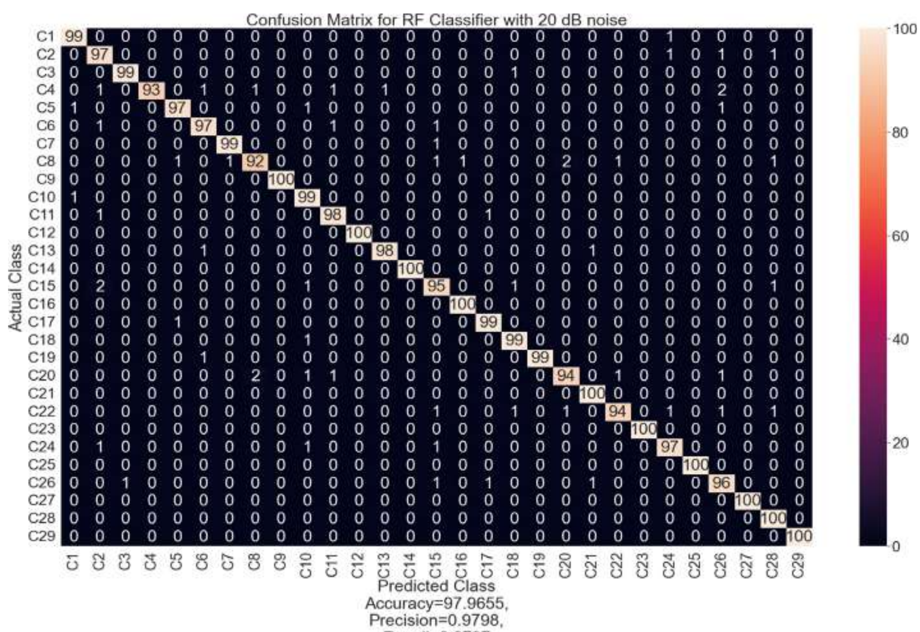


Fig. 8. Confusion matrix for RF classifier with 20 dB noise.

eigen vectors. Hence the lower dimensional data matrix with dimensions 5800 $\times$ 28 is obtained.

6.3. Classification

The well-known robust methods such as KNN, NB, SVM and RF are used for the classifications of the PQDs signals which are briefly explained in Section 4. As the dataset has made from 29 classes, after the dimension reduction each class is represented by 28 features, resulting in a selected dataset, Y, with a dimension of 5800 $\times$ 28. The reduced dataset is equally divided into two parts, with 100 samples for each class is used for training and 100 samples for each class is used for testing. The classification process is done in Python using Sklearn library. The accuracy of each of the classifier under different noisy conditions are observed and compared using confusion matrix.

7. Results and discussion

7.1. Performance of proposed method for different classifiers

The performance results of the proposed method i.e. LDA for different types of classifiers under different noisy conditions are shown in Table 1. It can be seen from Table 1 that the performance of the proposed method of dimension reduction has higher accuracy with different classifiers used in this paper. In order to further observe the misclassification of PQDs, confusion matrix for each of the case is obtained and it is found that classification accuracy in each of PQD is relatively high and there is no

Table 1
Accuracy comparison of different classifiers under noisy conditions.

PQD class label	LDA+KNN					LDA+NB					LDA+SVM					LDA+RF				
	20 dB	25 dB	30 dB	35 dB	40 dB	20 dB	25 dB	30 dB	35 dB	40 dB	20 dB	25 dB	30 dB	35 dB	40 dB	20 dB	25 dB	30 dB	35 dB	40 dB
C1	97	99	96	93	92	100	100	100	98	98	100	99	99	98	98	99	96	98	96	91
C2	96	94	94	97	94	100	100	100	98	97	97	100	100	97	97	97	97	96	98	88
C3	100	100	99	100	99	99	100	100	100	100	100	100	100	100	98	99	97	100	97	96
C4	98	99	96	96	97	100	98	99	98	100	100	99	98	99	100	93	95	100	93	97
C5	99	99	100	96	93	99	99	100	99	100	100	98	99	99	98	97	92	98	98	91
C6	100	99	99	99	94	100	100	99	99	100	100	100	99	98	99	97	97	99	94	89
C7	100	98	98	97	94	100	100	98	100	100	100	98	99	99	99	99	99	96	93	90
C8	92	94	94	95	91	100	98	96	99	97	99	100	98	98	99	92	93	92	96	93
C9	100	98	100	99	97	100	100	100	100	100	100	98	100	99	99	100	99	99	97	96
C10	99	99	97	92	90	100	100	100	100	97	99	100	100	99	97	99	100	95	100	90
C11	99	98	99	94	95	100	100	99	99	99	99	98	99	98	97	98	96	93	95	97
C12	100	99	100	99	97	100	100	100	100	100	100	99	100	99	99	100	98	100	97	96
C13	100	99	98	96	95	100	100	100	100	97	99	100	100	100	98	98	98	99	97	99
C14	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	99	100	99	100
C15	96	95	93	97	87	100	100	100	100	99	98	100	98	99	99	95	96	97	96	91
C16	100	100	99	100	99	100	100	100	100	100	100	100	100	100	100	100	99	100	100	93
C17	99	99	99	100	98	100	99	99	99	97	100	100	99	100	100	99	98	98	100	99
C18	96	97	96	96	92	99	97	98	98	95	99	99	100	98	99	99	97	94	98	90
C19	100	100	100	100	96	100	100	100	99	99	100	100	100	99	100	99	98	100	99	94
C20	96	98	93	96	92	99	100	100	98	97	100	100	98	98	96	94	96	95	95	97
C21	100	100	100	98	96	100	100	100	100	100	100	100	100	100	100	100	98	97	100	97
C22	95	95	96	99	95	99	100	99	100	98	100	100	99	100	97	94	98	97	96	95
C23	100	100	100	100	99	100	100	100	100	99	100	100	100	100	100	100	100	100	98	97
C24	98	95	96	98	98	98	99	99	98	98	98	99	99	98	98	97	96	94	94	96
C25	100	100	100	100	99	100	100	100	100	100	100	100	100	100	100	100	99	100	100	100
C26	97	99	97	96	98	100	100	100	100	99	100	99	100	100	99	96	97	94	94	99
C27	100	100	100	100	99	100	100	100	100	100	100	100	100	100	100	100	100	99	100	100
C28	99	98	98	95	97	99	100	99	100	99	99	100	99	99	99	100	93	96	99	96
C29	100	99	99	100	100	100	100	100	100	100	100	99	100	100	100	100	99	98	99	100
Accuracy (%)	**98.48**	**98.27**	**97.79**	**97.51**	**95.62**	**99.72**	**99.65**	**99.48**	**99.37**	**98.79**	**99.55**	**99.48**	**99.41**	**99.10**	**98.79**	**97.97**	**97.24**	**97.38**	**97.17**	**95.06**

Table 2
Accuracy comparison of different dimensional reduction methods.

Name of method	Name of classifier							
	PQD with 20 dB noise				PQD with 30 dB noise			
	KNN	NB	SVM	RF	KNN	NB	SVM	RF
LDA	98.48	**99.72**	**99.55**	97.97	97.79	**99.48**	**99.41**	97.38
CCA	**98.82**	96.82	98	98	**98.99**	97.72	98	98
MCCA	95.44	98	83.86	**98.75**	94.55	97.44	80.13	**98.2**

excessive misclassification. For an example confusion matrices of PQDs with 20 dB noise for four classifiers, i.e. KNN, NB, SV and RF with LDA technique are shown in Figs. 5–8 respectively. It is seen that the overall accuracy is 98.48%, 99.72%, 99.55% and 97.96% for KNN, NB, SV and RF classifier respectively. Also, the performance of NB classifier is better as compare to the other classifiers under all noisy conditions when used with the proposed method.

7.2. Comparison with other dimensionality reduction techniques

Apart from LDA, other dimensional reduction methods such as CCA and MCCA has been tested on PQDs with 20 dB and 30 dB noise. Classification accuracy for each of the classifier (KNN, NB, SVM and RF) has been obtained which are shown in Table 2. It can be seen that out of the three dimensional reduction methods, LDA (proposed method) gives higher accuracy with NB and SVM classifiers and comparable results with KNN and RF classifiers.

7.3. Discussion and comparison with published articles

The performance of proposed method is checked on four type of classifiers and their accuracy is compared with various methods presented in the literature as shown in Table 3 for noise level of 20 dB and 40 dB. It can be seen from the table that proposed method has most number of PQD events as compared to the other methods shown in table. The accuracy of the proposed method with NB classifier is highest with 99.79% and 98.79% for 20 dB and 40 dB noise respectively compared to published articles given in Table 3 and also it is highest among all the classifiers which are used for testing the proposed method. The accuracy of the proposed method with SVM classifier has also shown good results of 99.48% and 98.79% with 20 dB and 40 dB noise respectively which is higher among all compared methods except with Optimal Fast Discrete ST (MOFDST) [48], which is having 13 PQD events, but in this study 29 PQD events are considered. The accuracy of the KNN classifier with the proposed method is 98.28% and 95.62% and for RF classifier is 98% and 95.06% with 20 dB and 40 dB noise respectively which is higher or almost comparable as compare to the other methods presented in Table 3 with an advantage of most number of PQD events.

8. Conclusion

This paper proposes a novel method based on dimensionality reduction for the detection and classification of the PQDs. The dimension of the PQD data set has been reduced by removing

Table 3
Performance comparison with published articles.

Feature extraction technique	Type of classifier	No. of PQDs	Accuracy (%)	
			Noise Level 20 dB	Noise Level 40 dB
Time and frequency domain statistical features [26]	Artificial Neural Network (ANN)	8	–	96.03
Modified Optimal Fast Discrete ST (MOFDST) [48]	RF	13	99.61	–
Finite Impulse Response-Discrete Gabor Transform (FIR-DGT) [51]	Type-2 Fuzzy Kernel-based Support Vector Machine (T2FK-SVM)	9	96.22	94.56
Tunable-Q Wavelet Transform (TQWT) [11]	Multiclass Support Vector Machines (MSVM)	14	96.42	–
Strong Tracking Filters (STF) [52]	Rule-Based Extreme Learning Machine (ELM-RL)	20	92.6	98.8
Stockwell transform [53]	Decision tree	7	98.2	–
Signal-Piloted Analog to Digital Converters (SPADC) with Activity Selection Algorithm (ASA) and time-domain statistical features [4]	SVM	4	98.05	–
	KNN	**29**	**98.28**	**95.62**
Proposed method	NB	**29**	**99.79**	**98.79**
LDA	SVM	**29**	**99.48**	**98.79**
	RF	**29**	**98**	**95.06**

the unnecessary features with the aid of supervised learning technique, i.e. Linear Discriminant Analysis. As compare to other methods of feature extraction, this method does not make use of any arithmetically complex transformations. Therefore, the reduced simple and unique features exactly show the PQDs to enhance the recognition capability. The proposed method has been analysed on 29 types of single and simultaneous type of PQDs signals with different noise levels. The performance of the proposed method has been tested with four types of machine learning classifiers such as KNN, NB, SVM and RF using 100 samples of each disturbance type to validate the effectiveness. The efficiency of NB classifier is more than 99% and highest among all tested classifiers, efficiency of other classifiers are also very good under all considered conditions. The performance comparison with other methods presented in literature revealed that the proposed method is distinguished in terms of accuracy and number of PQDs tested. Thus, the LDA based classifiers can be effectively implemented for the classification and the detection of Power quality disturbances.

CRediT authorship contribution statement

Gurpreet Singh: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft. **Yash Pal:** Conception and design of study, Analysis and/or interpretation of data, Writing – review & editing. **Anil Kumar Dahiya:** Conception and design of study, Analysis and/or interpretation of data, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

All authors approved the version of the manuscript to be published.

References

[1] B. Singh, A. Chandra, K. Al-Haddad, Power Quality: Problems and Mitigation Techniques, John Wiley & Sons, 2014.
[2] S. Jain, Control strategies of shunt active power filter, in: Modeling and Control of Power Electronics Converter System for Power Quality Improvements, Elsevier, 2018, pp. 31–84.
[3] IEEE Recommended Practice for Monitoring Electric Power Quality, IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009), 2019, pp. 1–98, http://dx.doi.org/10.1109/IEEESTD.2019.8796486.
[4] S. Mian Qaisar, Signal-piloted processing and machine learning based efficient power quality disturbances recognition, PLoS One 16 (5) (2021) e0252104.
[5] J. Li, H. Liu, D. Wang, T. Bi, Classification of power quality disturbance based on S-transform and convolution neural network, Front. Energy Res. 9 (2021) 325.
[6] H. Deng, Y. Gao, X. Chen, Y. Zhang, Q. Wu, H. Zhao, Harmonic analysis of power grid based on FFT algorithm, in: 2020 IEEE International Conference on Smart Cloud (SmartCloud), IEEE, 2020, pp. 161–164.
[7] A.K. Puliyadi Kubendran, A.K. Loganathan, Detection and classification of complex power quality disturbances using S-transform amplitude matrix–based decision tree for different noise levels, Int. Trans. Electr. Energy Syst. 27 (4) (2017) e2286.
[8] J. Li, Z. Teng, Q. Tang, J. Song, Detection and classification of power quality disturbances using double resolution S-transform and DAG-SVMs, IEEE Trans. Instrum. Meas. 65 (10) (2016) 2302–2312.
[9] C.A. Naik, P. Kundu, Power quality disturbance classification employing S-transform and three-module artificial neural network, Int. Trans. Electr. Energy Syst. 24 (9) (2014) 1301–1322.
[10] R. Kumar, B. Singh, D. Shahani, A. Chandra, K. Al-Haddad, Recognition of power-quality disturbances using S-transform-based ANN classifier and rule-based decision tree, IEEE Trans. Ind. Appl. 51 (2) (2014) 1249–1258.
[11] K. Thirumala, M.S. Prasad, T. Jain, A.C. Umarikar, Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances, IEEE Trans. Smart Grid 9 (4) (2016) 3018–3028.
[12] H. Erişti, A. Uçar, Y. Demir, Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines, Electr. Power Syst. Res. 80 (7) (2010) 743–752.
[13] Z. Moravej, M. Pazoki, A.A. Abdoos, Wavelet transform and multi-class relevance vector machines based recognition and classification of power quality disturbances, Eur. Trans. Electr. Power 21 (1) (2011) 212–222.
[14] S. Khokhar, A.A.M. Zin, A.P. Memon, A.S. Mokhtar, A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network, Measurement 95 (2017) 246–259.
[15] S. Naderian, A. Salemnia, An implementation of type-2 fuzzy kernel based support vector machine algorithm for power quality events classification, Int. Trans. Electr. Energy Syst. 27 (5) (2017) e2303.
[16] Y. Sun, X. Tang, X. Sun, D. Jia, P. Wang, G. Zhang, Research on multi-type energy storage coordination control strategy based on MPC-HHT, Proc. CSEE 38 (9) (2018) 2580–2588.
[17] Z. Liu, Y. Cui, W. Li, A classification method for complex power quality disturbances using EEMD and rank wavelet SVM, IEEE Trans. Smart Grid 6 (4) (2015) 1678–1685.
[18] B. Biswal, M. Biswal, S. Mishra, R. Jalaja, Automatic classification of power quality events using balanced neural tree, IEEE Trans. Ind. Electron. 61 (1) (2013) 521–530.

[19] Y. Liu, T. Jin, M.A. Mohamed, Q. Wang, A novel three-step classification approach based on time-dependent spectral features for complex power quality disturbances, IEEE Trans. Instrum. Meas. 70 (2021) 1–14.

[20] S.L. Dhoriyani, P. Kundu, Comparative group THD analysis of power quality disturbances using fft and STFT, in: 2020 IEEE First International Conference on Smart Technologies for Power, Energy and Control, STPEC, IEEE, 2020, pp. 1–6.

[21] Y. Wang, Q. Li, F. Zhou, A novel algorithm for transient power quality disturbances detection, Proc. CSEE 37 (24) (2017) 7121–7132.

[22] S. Niu, K. Wang, Z. Liang, Synchronous phasor estimation method for power system based on modified strong trackin g unscented Kalman filter, Power Syst. Technol. 43 (9) (2019) 3218–3225.

[23] P.R. Babu, P. Dash, S. Swain, S. Sivanagaraju, A new fast discrete S-transform and decision tree for the classification and monitoring of power quality disturbance waveforms, Int. Trans. Electr. Energy Syst. 24 (9) (2014) 1279–1300.

[24] M. Mishra, Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review, Int. Trans. Electr. Energy Syst. 29 (8) (2019) e12008.

[25] Y.H. Gu, M.H. Bollen, Time-frequency and time-scale domain analysis of voltage disturbances, IEEE Trans. Power Deliv. 15 (4) (2000) 1279–1284.

[26] F.A. Borges, R.A. Fernandes, I.N. Silva, C.B. Silva, Feature extraction and power quality disturbances classification using smart meters signals, IEEE Trans. Ind. Inform. 12 (2) (2015) 824–833.

[27] M. Valtierra-Rodriguez, R. de Jesus Romero-Troncoso, R.A. Osornio-Rios, A. Garcia-Perez, Detection and classification of single and combined power quality disturbances using neural networks, IEEE Trans. Ind. Electron. 61 (5) (2013) 2473–2482.

[28] A.D. Gonzalez-Abreu, J.J. Saucedo-Dorantes, R.A. Osornio-Rios, R.J. Romero-Troncoso, M. Delgado-Prieto, D. Morinigo-Sotelo, Condition monitoring approach based on dimensionality reduction techniques for detecting power quality disturbances in cogeneration systems, in: 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA, IEEE, 2019, pp. 898–903.

[29] F. Zhao, R. Yang, Power quality disturbance recognition using S-transform, in: 2006 IEEE Power Engineering Society General Meeting, IEEE, 2006, p. 7.

[30] M. Biswal, P.K. Dash, Measurement and classification of simultaneous power signal patterns with an S-transform variant and fuzzy decision tree, IEEE Trans. Ind. Inform. 9 (4) (2012) 1819–1827.

[31] M. Chilukuri, P. Dash, Multiresolution S-transform-based fuzzy recognition system for power quality events, IEEE Trans. Power Deliv. 19 (1) (2004) 323–330.

[32] L. Zhou, C. Guan, W. Lu, Application of multilabel classification method to categorization of multiple power quality disturbances, Proc. CSEE 31 (4) (2011) 45–50.

[33] A.A. Abdelsalam, A.A. Eldesouky, A.A. Sallam, Characterization of power quality disturbances using hybrid technique of linear Kalman filter and fuzzy-expert system, Electr. Power Syst. Res. 83 (1) (2012) 41–50.

[34] A. Rodríguez, J. Aguado, F. Martín, J. López, F. Muñoz, J. Ruiz, Rule-based classification of power quality disturbances using S-transform, Electr. Power Syst. Res. 86 (2012) 113–121.

[35] J.C. Palomares-Salas, J.J.G. de la Rosa, A. Agüera-Pérez, J.M. Sierra-Fernández, Smart grids power quality analysis based in classification techniques and higher-order statistics: Proposal for photovoltaic systems, in: 2015 IEEE International Conference on Industrial Technology, ICIT, IEEE, 2015, pp. 2955–2959.

[36] J. Yin, S. Sun, Multiview uncorrelated locality preserving projection, IEEE Trans. Neural Netw. Learn. Syst. 31 (9) (2019) 3442–3455.

[37] M. Kan, S. Shan, H. Zhang, S. Lao, X. Chen, Multi-view discriminant analysis, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2015) 188–194.

[38] S. Sun, X. Xie, M. Yang, Multiview uncorrelated discriminant analysis, IEEE Trans. Cybern. 46 (12) (2015) 3272–3284.

[39] F. Pan, G. Song, X. Gan, Q. Gu, Consistent feature selection and its application to face recognition, J. Intell. Inf. Syst. 43 (2) (2014) 307–321.

[40] A. Tharwat, T. Gaber, A. Ibrahim, A.E. Hassanien, Linear discriminant analysis: A detailed tutorial, AI Commun. 30 (2) (2017) 169–190.

[41] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using LDA-based algorithms, IEEE Trans. Neural Netw. 14 (1) (2003) 195–200.

[42] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data—with application to face recognition, Pattern Recognit. 34 (10) (2001) 2067–2070.

[43] T. Li, S. Zhu, M. Ogihara, Using discriminant analysis for multi-class classification: an experimental investigation, Knowl. Inf. Syst. 10 (4) (2006) 453–472.

[44] J. Ye, R. Janardan, Q. Li, Two-dimensional linear discriminant analysis, Adv. Neural Inf. Process. Syst. 17 (2004).

[45] G. Bonaccorso, Machine Learning Algorithms, Packt Publishing Ltd, 2017.

[46] H. Taniguchi, T. Shirakawa, T. Takahashi, Implementation of human cognitive bias on naive Bayes, EAI Endorsed Trans. Creat. Technol. 3 (7) (2016) e3.

[47] R. Gandhi, Support vector machine—introduction to machine learning algorithms, Towards Data Sci. 7 (2018).

[48] M.V. Reddy, R. Sodhi, A modified S-transform and random forests-based power quality assessment framework, IEEE Trans. Instrum. Meas. 67 (1) (2017) 78–89.

[49] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[50] R. Igual, C. Medrano, F.J. Arcega, G. Mantescu, Integral mathematical model of power quality disturbances, in: 2018 18th International Conference on Harmonics and Quality of Power, ICHQP, IEEE, 2018, pp. 1–6.

[51] S. Naderian, A. Salemnia, Method for classification of PQ events based on discrete gabor transform with FIR window and T2FK-based SVM and its experimental verification, IET Gener. Transm. Distrib. 11 (1) (2017) 133–141.

[52] X. Chen, K. Li, J. Xiao, Classification of power quality disturbances using dual strong tracking filters and rule-based extreme learning machine, Int. Trans. Electr. Energy Syst. 28 (7) (2018) e2560.

[53] F. Zaro, S. Alqam, Power quality detection and classification using s-transform and rule-based decision tree, 2019.