ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

COMPARATIVE ANALYSIS OF MLP-RBF BASED NETWORKS FOR DETECTION AND CLASSIFICATION OF POWER QUALITY DISTURBANCES

Swapnil B. Mohod*, Vilas N. Ghate

* Faculty in Electrical & Electronics Engineering, Prof Ram Meghe College of Engineering, Badnera-Amravati (MS)-444701India Ph: +919970070043.

ABSTRACT

Electrical energy is first and foremost criterion for overall economic growth of the society. Widespread expansion of automated electronic equipments and miniaturization in micro-electronics used in power system hampered the quality of power a lot. This paper deals with the comparative analysis of Multilayer Perceptron Neural Network and Radial Basis Function based classifier for the detection and classification of power quality disturbances. Simple statistical parameters are used as input feature space for detailed design optimization of MLP-NN and RBF classifier. Further, for the dimensionality reduction, Principal component analysis and Sensitivity analysis are also examined. Optimized classifier is robust enough to classify the fundamental Power Quality disturbances with classification accuracy up to 99.81%.

KEYWORDS: Power Quality, Fourier transform, Wavelet transform, Artificial intelligence, MLP, PCA.

INTRODUCTION

Now a day's PQ has got prime importance in power system as well as industry. Extensive use of sensitive automated control strategies for improvement of system stability results in deviation of supply from its normal values. These disturbances corrupt the original waveform and growing complexity leads to severe consequences such as huge economic losses if any component fails. Thus, a PQ problem is ultimately a consumer-driven issue [1]- [2].
Horizon of PQ problem covers problems such as overheating, motor failures, inaccurate metering and malfunctioning of protective equipments. Features extraction of disturbances from a large number of power signals is an actual task. Researchers working in this domain for PQ analysis such as fast Fourier transform (FFT), fractional Fourier transform [3]-[4] and wavelet transform [5]-[7]. Artificial neural networks ANNs, fuzzy logic and support vector machines are also used for these kinds of PQ events.

PQ DISTURBANCES

PQ disturbances can be broadly classified as stationary and non stationary. Waveforms of stationary disturbance exhibit time-invariant statistical characteristics while in case of non-stationary disturbance, these waveforms have time-variant statistical characteristics. These disturbances are further classified as short-duration and long duration voltage variations, voltage imbalance and fluctuations, waveform distortion and power frequency variations.

Short-duration voltage variations: Any variation in the supply voltage for duration not exceeding one minute is known as Short-duration voltage variations i.e. Voltage Sags, Voltage Swell, and Interruption.

Long-duration voltage variations: Any variation in the supply rms voltage at fundamental frequency for duration exceeding one minute is known as Long-duration voltage variations as Overvoltage or Under voltage and Sustained interruption.

Voltage imbalance and voltage fluctuation: voltages of three-phase supply are not equal in magnitude and may not even be equally displaced in time is called as voltage imbalance. Voltage fluctuation is a systematic random variation in supply voltage whereas voltage flicker is very rapid change in supply voltage.

Waveform distortion: This is the steady state deviation in voltage or current waveform from an ideal sine wave. These are observed in D.C. offset and Harmonics forms.

Power frequency variation: These are usually caused by rapid changes in load connected to the system [9].

DIAGNOSIS OF PQ DISTURBANCES

Fast diagnosis of PQ disturbances is an urgent requirement so as to assist network operators to performing counter measures. It also ensures the implementation of suitable PQ mitigation actions. To analyze power signals, various signal processing techniques such as Fourier transformation methods, wavelet transform and S-transform are commonly applied. Also, artificial intelligence (AI) techniques have found their application in this field of power system. Application of hybrid methodology by embedding above two techniques for diagnostic purposes is also successfully practiced by some researchers. This section further discusses each of these methodologies.

Fourier transforms (FT)

Complicated periodic functions are written as the sum of simple waves mathematically represented by sine and cosines using FT for PQ analysis are as follows:

Discrete Fourier Transform (DFT)

Fourier Transform applied mostly to repetitive signals, oftenly known as Discrete Fourier Transform. For a finite length discrete signal x[n], its DFT and frequency function is given as in equation (1) and (2) respectively;

$$X(f) = \frac{1}{N}\sum_{n=0}^{N-1} x[n] e^{\frac{-j2\Pi nf}{N}} \qquad (1)$$

$$x(n) = \sum_{n=0}^{N-1} X[f] e^{\frac{j2\Pi nf}{N}} \qquad (2)$$

Fast Fourier Transform (FFT)

FFT has a big scope in digital signal processor to provide a frequency spectrum analysis. In 1948 Cooley and Tukey exhibited the computation of N point DFT as a function of only 2N instead of N^2. Signal x[n] decomposed into odd and even part using FFT can be written as,

$$FFT(x,f) = \frac{1}{2N}\sum_{n=0}^{N-1} x[2n] e^{\frac{-j\Pi(2n)f}{N}} + \frac{1}{2N}\sum_{n=0}^{N-1} x[2n+1] e^{\frac{-j\Pi(2n+1)f}{N}} \qquad (3)$$

FFT reduces the computational complexity from N in DFT to N log N multiplications for same expression. This is a main advantage of FFT over DFT.

Short Time Fourier Transform (STFT)

It is use to analyze signals whose spectrum changes with time. It is calculated by repeatedly multiplying the time series with shifted short time windows and performing a DFT on it. Here the window helps to localize the time-domain data before obtaining the frequency domain information. STFT for continuous time signal is given as;

$$X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)\omega(t-\tau) e^{-j\omega t} dt \qquad (4)$$

where; x(t)= signal to be transformed, w(t)=window function, $X(\tau,\omega)$ = FT of $x(t)\omega(t-\tau)$. In case of discrete signal with 'm' as discrete time-shift, it can be expressed as

$$X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m] e^{-j\omega n} \qquad (5)$$

However, STFT of a signal has constant window length which limits non-stationary signal resolution. To overcome this resolution problem using STFT, a signal processing tool, Wavelet Transform (WT), had been widely implemented in PQ analysis.

Wavelet transform

In Wavelet Transform any signal is decompose for detailed analysis with multiple time–frequency resolution. Unlike STFT, the length of the smoothing window of the WT depends on the frequency analysed. WT is commonly observed in mainly two forms namely: Continuous wavelet transform (CWT) and Discrete wavelet transform (DWT). CWT is given by,

$$XWT(\tau, s) = \frac{1}{\sqrt{s}} \int x(t).\Psi(\frac{t-\tau}{s})dt \qquad (6)$$

where; x(t)=signal to be analysed, $\psi(t)$ is the mother wavelet, s and τ represent scale and translational parameters respectively.

DWT is a discrete counter of CWT for decomposition of the signal into mutually orthogonal set of wavelets. WT obeys some defined rules by means of a discrete set of the wavelet scales and translations.

Artificial Intelligence (AI)

ANN is self-learning system, hence along with conventional training algorithms, an approach for hybrid algorithms as well as combined approach of ANN with wavelet and s-transform had been adopted. ANN has capability of non-linear function approximation and had been widely implemented in this domain. For industrial applications and recent concepts like distributed generation, ANN has found its application for PQ analysis and improvement.

EXPERIMENTAL SETUP AND DATA GENERATION

To detect and classify the PQ disturbances, the Neural Network based classifier is designed and optimized. In first step, for data collection experimental setup is as follows:

In experimentation following PQ disturbances are considered,

a. Voltage Sag
b. Voltage Swell
c. Arcing load influence
d. Short circuit condition

Fig.1. Experimental Setup

Mains fed one HP, single phase, 50 Hz squirrel cage Induction motor made is used for analysis of sag as well swell in the system by switching ON/OFF operation. 230V, single phase, 50Hz, Welding machine is used to generate actual arcing load influence in the system, Welding electrodes keep short to experience a short circuit phenomenon in the lab. The Tektronix Digital Storage Oscilloscope (DSO), TPS 2014 B, with 100 MHz bandwidth and adjustable sampling rate of 1GHz is used to capture the current signals. The Tektronix voltage probes of rating 1000 V, and bandwidth of 200 MHz approximately, 100 sets of signals were captured with a sampling frequency of 10 kHz, at different mains supply conditions. The experimental setup uses an Advantech data acquisition system having

specification, as PCLD-8710 - 100 KS/s, 12-bit, 16-ch PCI Multifunction Card. Overall to create a weak system inside the laboratory 2 ½ core , 200meter long cable is used so that influence of sag , swell and arc load will be observable.

FEATURE EXTRACTION

Collected data is analyzed in MATLAB environment and using wavelet transform, its required features obtained.
To demonstrate feature extraction capability of the wavelet technique, five types of PQ events are considered like Induction motor sag, arc load sag, sag due to welding machine short circuit; swell due to Induction motor switching off and healthy condition along with their important features. For decomposition and reconstruction of measured signal WT uses two function ie. Wavelet function ψ and scaling function ϕ perform successive decomposition of signal by MRA technique. Wavelet function ψ serving as high pass filter and generates high frequency component known as detailed function (d) and scaling function ϕ is convolving the signal with a low pass filter which generates low frequency components known as approximate function (a) of decomposed signal, distinct features of all instances are shown in Fig.2 to Fig.6.

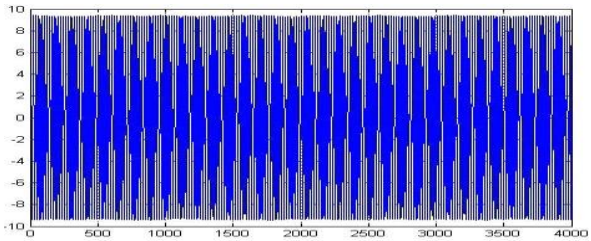


Fig.2. Normal condition

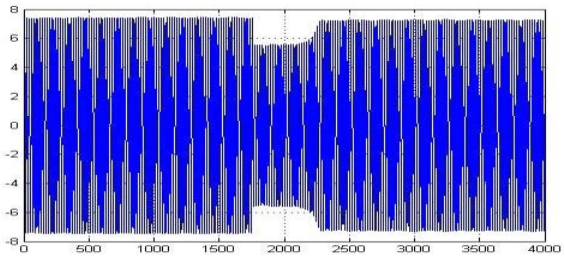


Fig.3. Sag due to Induction Motor Start

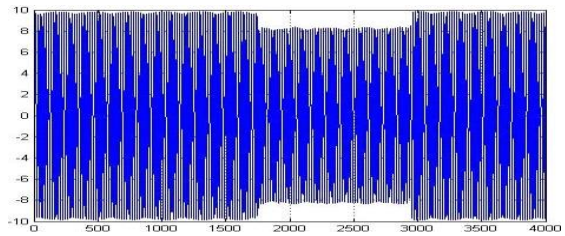


Fig.4. Sag due to welding machine short circuited

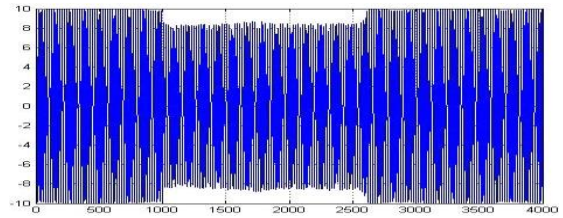


Fig.5. Sag due to arc load of welding machine

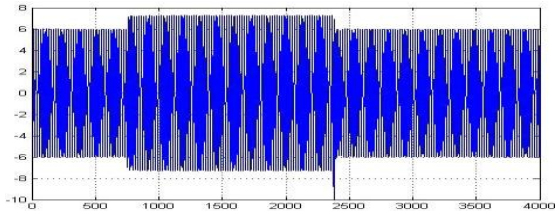


Fig.6. Swell due to Induction Motor switching off

Statistical parameters of a waveform are used to classify these different power quality disturbances. To be precise, 'sample' statistics will be calculated for current data. Minimum set of statistical data to be examined which mainly includes the kurtosis coefficient, maximum and minimum values of the skewness coefficient and root mean square (RMS) of the zero mean signals (standard deviation) and *Pearson's coefficient of skewness,* g_2 defined by:

$$g_2 = \frac{3\left(\bar{x} - \tilde{x}\right)}{S_x} \qquad (7)$$

$\bar{x}$, x and S_x is mean, median sample and standard deviation respectively.

The sample coefficient of variation v_x is defined by;

$$v_x = \frac{S_x}{\bar{x}} \qquad (8)$$

Data set for the r^{th} sample moment about the sample mean is:

$$m_r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^r}{n} \qquad (9)$$

m_2 refers to spread about the center, m_3 denotes to skewness about the center, m_4 refers about amount of data massed at the center. The Second, third and fourth moments are used to define explicitly the sample coefficient of skewness, g_3 and the sample coefficient of kurtosis, g_4 as follows.

$$g_3 = \frac{m_3}{\left(\sqrt{m_2}\right)^3} \qquad (10)$$

$$g_4 = \frac{m_4}{\left(\sqrt{m_2}\right)^4} \qquad (11)$$

The covariance between dimensions *j* and *k* sample is given as;

$$c_{jk} = \frac{\sum_{i=1}^{n}\left(x_{ij} - \overline{x_j}\right)\left(x_{ik} - \overline{x_k}\right)}{(n-1)} \qquad (12)$$

For dimensions *j* and *k*, r_{jk} the ordinary correlation coefficient is given as;

$$r_{jk} = \frac{c_{jk}}{S_j - S_k} \qquad (13)$$

DISTURBANCE CLASSIFIER USING ANN

Multi Layer Perceptron

Multilayer Perceptron (MLP) Neural Network is proposed as disturbance classifier. 14 numbers of inputs are used in input layer and six Processing Elements (PE's) are used in output layer for six crucial conditions. For data processing Neuro Solution 5.7, XLSTAT-2010 and MATLAB 7.1 environment, is used. General learning algorithm used is given as:

Initialization of Weights:
Step 1: Initialize the weights to small random values
Step 2: While stopping condition is false, do step 3-10
Step 3: For each training pair do steps 4-9

Feed forward:
Step 4: Each input unit receives the input signal x_i and transmits this signals to all units in the hidden layer
Step 5: Each hidden unit (z_j ,*j=1,…,p*) sums its weighted input signals

$$z_{-inj} = v_{oj} + \sum_{i=1}^{n} x_i v_{ij} \qquad (14)$$

Applying the activation function $Z_j = f(z_{inj})$ here the activation function is $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ and sends this signal to all units in output units.
Step 6: Each output unit (y_k , *k=1,…,m*) sums its weighted input signals ,

$$y_{-ink} = w_{ok} + \sum_{j=1}^{p} z_i w_{jk} \qquad (15)$$

And apply its activation function to calculate the output signals $Y_k = f\left(y_{-ink}\right)$ here the activation function is

$$\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x}) \qquad (16)$$

Back Propagation Error:
Step 7: Each output unit (y_k , *k=1,…,m*) receives a target pattern corresponding to an input pattern error information term is calculated as

$$\delta_k = \left(t_k - y_k\right) f\left(y_{-ink}\right)$$

Step 8: Each hidden unit (z_j ,j=1,…,p) sums its delta inputs from units in the layer above

$$\delta_{-inj} = \sum_{k=1}^{m} \delta_i w_{jk} \qquad (17)$$

The error information term is calculated as

$$\delta_j = \delta_{-inj} f\left(z_{-inj}\right) \qquad (18)$$

Updation of weight and Biases:
Step 9: Each output unit (y_k, *k=1,…,m*) updates its bias and weights *(j=0,…,p)*

$$w_{jk}(t+1) = w_{jk}(t) + \alpha\delta_k z_j + \mu\left[w_{jk}(t) - w_{jk}(t-1)\right] \quad (19)$$

Where α is learning rate and μ is momentum factor

And each hidden unit *(z_j , j=1,…,p)* updates its bias and weights *(i=0,…,n)*

$$v_{jk}(t+1) = v(t) + \alpha\delta_j x_i + \mu\left[v_{ij}(t) - v_{ij}(t-1)\right] \quad (20)$$

Step 10: Test the stopping condition

The randomized data are fed to the neural network and are retrained five times with different random weight initialization so as to remove biasing. It ensures true learning and generalization for different hidden layers. It is observed that tanh-axon transfer function and step learning rule gives better performance as shown in Fig.7. The number of processing elements (PEs) in the hidden layer is varied. Optimum performance of network is obtained with 11 PEs in hidden layer as shown in Fig. 8.

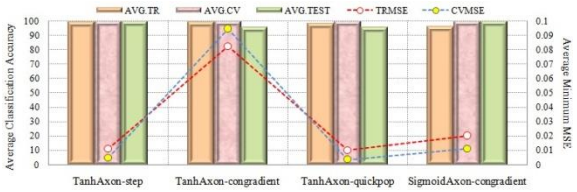
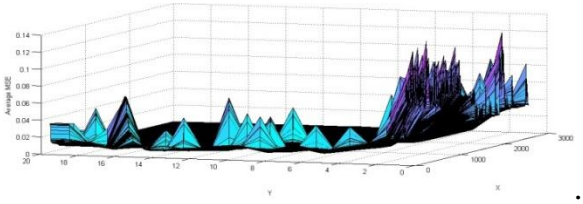


Fig.7 Variation of average MSE and classification accuracy for transfer function and learning rule



Y: No. of Epoch X: No. of PE's

Fig.8. Variation of average MSE on training with number of processing elements in the hidden layer

The parameters of the hidden layer and output layer i.e. step size is selected by comparing average minimum MSE. In hidden layer and output layer optimum value of step size is 0.8 and 0.9 respectively, performance is shown in Fig.9 and Fig.10.

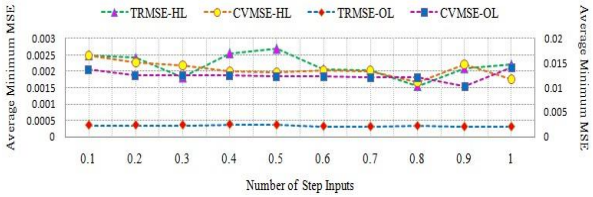


Fig.9. Variation of average minimum MSE on TR, CV dataset with Step size of hidden and output layer

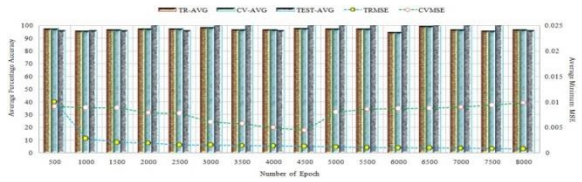


Fig.10. Variation of average MSE on training and CV with Average classification accuracy for number of Epoch

Selection of Error criterion:
A supervised learning requires a metric to measure how the network is doing. Members of the Error Criteria compare with some desired response. Any kind of error reported to the appropriate learning procedure. Calculating the sensitivity proper metric is determined by using gradient descent learning. As the network approaches to the desired response Cost function, *J* should decay towards zero, but normally it is positive. In literature several cost functions has presented, in which *p* is define as p=1, 2, 3, 4… ∞ criterion is L-1, L-2, L-3, L-4 …L∞. Cost function is used to define Components in the Error Criteria family.

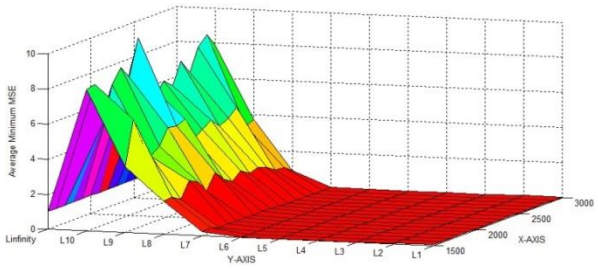
Error Criteria family components are defined by a cost function in the form:

$$J(t) = \frac{1}{2}\sum_{i}(d_i(t) - y_i(t))^p \qquad (21)$$

and error functions:

$$e_i(t) = -(d_i(t) - y_i(t)) \qquad (22)$$

As desired response and network's output are d (t) and y (t) respectively, various error criterions has been tested to select the correct error criterion and finally L-2 criterion shown in Fig.11, gives the optimal results.



Y-AXIS: Error Criterion X-AXIS: No. of Epochs
Fig.11. Variation in Average MSE with Error Criterion

From the above experimentation, selected parameters for MLP NN are given below.
Transfer: TanhAxon Learning Rule: Step
Number of Inputs: 14
Number of Hidden Layers: 01
Number of PEs in Hidden Layer: 11
Hidden Layer:
Transfer function: tanh Learning Rule: step
Step size: 0.8
Output Layer:
Transfer function: tanh Learning Rule: step
Step size: 0.9
Number of epochs = 4500
Number of connection weights: 237
Training Exemplars = 70%,
Cross validation Exemplars = 15%
Testing Exemplars = 15%
Time required per epoch per exemplar: 0.1822 micro –secs.

Proposed NN is trained on various datasets and later validated carefully so as to ensure that its performance does not depend on specific data partitioning scheme. The performance of the NN should be consistently optimal over all the datasets with respect to MSE and classification accuracy. Finally designed MLP is trained five times with different random weight initialization and tested on Testing dataset, CV dataset and Training dataset. For training and testing the data tagging by percent and data tagging by various groups and Leave-N-Out method are used. Leave-N-Out training is a technique that allows one to evaluate how well the model generalizes. It is very useful for small data sets, since it allows one to use the entire data set for training and testing. The algorithm trains the network multiple times, every time omitting an unusual subset of the data and using that subset for testing. The outputs from each tested subset are combined into one testing report and the model is trained one additional time using all of the data. The set of weights saved from the final training run can then be used for additional testing. To check the learning ability and classification accuracy the total data are divided in four groups. First two groups (50% data) are tagged as Training data and third and fourth group (each 25%) are tagged for Cross validation and Testing (1234: 1, 2-TR, 3-CV, 4-Test). Similar 24 combinations are prepared and network is train and test for each group. Variation of average minimum MSE and percent data tagged for different error criteria have tested and L_2 error criteria has minimum MSE and maximum classification accuracy. Results are shown in Figs. 12, 13 and Fig.14.

Fig. 12. Variation of average minimum MSE with test on Testing and Training dataset and percent data tagged for training

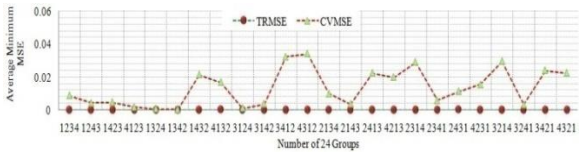


Fig.13. Variation of average minimum MSE with Training and CV for all 24 group of dataset

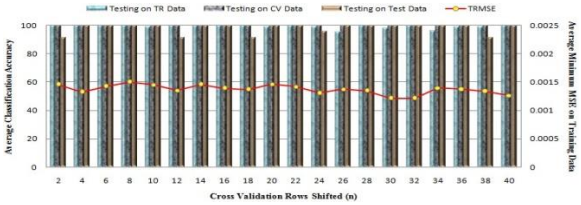


Fig. 14: Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

Finally optimized MLP based classifier gives the classification accuracy of 99.18% with average 0.00083773 training MSE and 0.025529 average cross validations MSE.

Radial Basis Function

Radial basis function (RBF) networks are nonlinear hybrid networks typically containing a single hidden layer of processing elements (PEs). This layer uses gaussian transfer functions, rather than the standard sigmoidal functions. The centers and widths of the gaussians are set by unsupervised learning rules, and supervised learning is applied to the output layer. These networks tend to learn much faster than MLPs. Use of this network is recommended only when the number of exemplars is so small or so dispersed that clustering is ill-defined.

For standard RBF's, the supervised segment of the network only needs to produce a linear combination of the output at the unsupervised layer. Therefore 0 hidden layers is the default. Hidden Layers can be added to make the supervised segment a MLP instead of a simple linear perceptron. Six conditions of PQ disturbances namely Healthy, Sag due to Induction Motor, Swell due to Induction Motor, Welding machine short circuit, Arc load and Line to Ground.
The general learning algorithm is as follows:

For all value of *x* the response of Gaussian activation function is nonnegative. The function is defined as

$$f(x) = \exp(-x^2) \qquad (23)$$

and its derivative

$$f'(x) = -2x\exp(-x^2) = -2f(x) \qquad (24)$$

Back propagation network in the Gaussian function is different from the radial basis function.
Step 1: Initialize the weights (set to small values)
Step 2: do step 3-10, if stopping condition is false
Step 3: do step 4-9 for each input
Step4: all units in the hidden layer receives input signals from each input unit (x_i , $i=1,\ldots, n$) .
Step 5: Calculate the radial basis function
Step 6: Choose the centers for the radial basis functions. The centers are chosen from the set of input vectors. A sufficient numbers of centers have been selected in order to ensure adequate sampling of the input vector space.
Step 7: The output of i_m unit $v_i(x_i)$ in the hidden layer

$$v_i(x_i) = e\left\{\sum_{j=1}^{r}\frac{\left[x_{ji} - (x'_{ji})^2\right]}{\sigma_i^2}\right\} \qquad (25)$$

Where x_{ji} - centre of the RBF unit for input variable

x'_{ji} - j^{th} variable of input pattern

σ_i -Width of the i^{th} RBF unit

Step 8: Initialize the weights in the output layer of the network to some small random value

Step 9: output calculation of the neural network

$$y_{net} = \sum_{i=1}^{H} w_{im} v_i(x_i) + w_0 \qquad (26)$$

Where

H-number of hidden layer nodes (RBF Function)

y_{net} -Output value of m^{th} node in output layer for the n^{th} incoming pattern

w_{im} - Weight between i^{th} RBF unit and m^{th} output node

w_0 - Biasing term at n^{th} output node

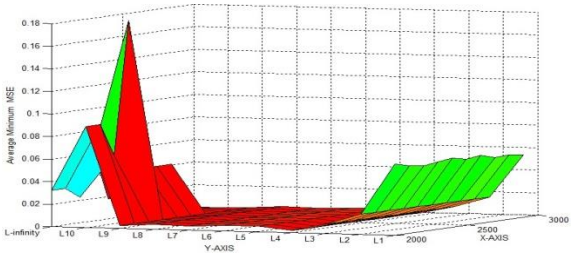
Step 10: Calculate the error and check the stopping condition

To remove biasing and ensure true learning and generalization for different parameters, randomized data is retrained five times with different random weight initialization then fed to the neural network.

Selection of Error criterion:

Similar experimentations are carried out to select the proper error criterion. Variations of Average minimum MSE with different error criterion is shown in Fig.15.

Finally L-4 criterion gives the optimal results.



Y: Error Criterion X: No. of Epochs

Fig.15. Variation in Average MSE with Error Criterion

From the experimentation the unsupervised and supervised parameters, such as competitive rule, metric and cluster centers have been optimized. Experimentation results are shown in Fig.16 to Fig.19.

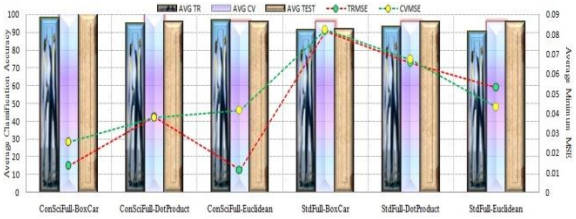


Fig.16. Variation in Average Classification Accuracy with Competative rule and metric

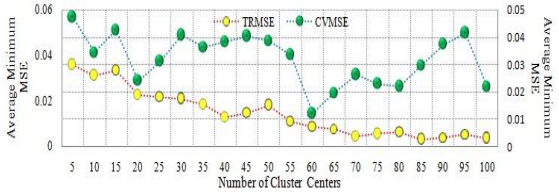


Fig.17 . Variation in Average Minimum MSE for different number of Cluster Centers

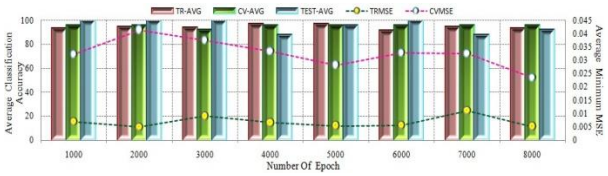


Fig.18.Variation in Average Classification Accuracy and Average minimum MSE with Epoch

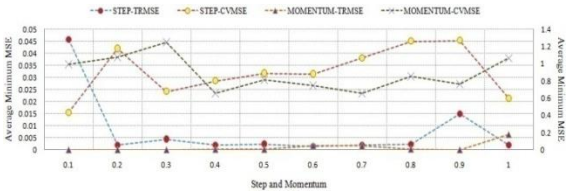


Fig.19. Variation in average Minimum MSE with Training and CV for Step and Momentum

Finally, the RBF classifier is designed with the following parameters:
Number of Inputs: 14
Competitive Rule: Conscience-Full, Metric: Boxcar,
Number of Hidden Layers: 0
Output Layer: Learning Rule: Momentum,
Step Size: 0 .1 Momentum: 0.7
Transfer Function: Tanh
Number of epochs = 5000
Unsupervised Learning: Maximum Epoch: 1500
Supervised Learning: Maximum Epoch: 3500
Learning Rate: start at: 0.01 Decay to: 0.001
Number of Cluster Centers: 60
Number of connection weights: 1266
Exemplars for training = 70%,
Exemplars for cross validation = 15%
Exemplars for Testing = 15%
Time required for training per exemplar is 0.604 micro-secs

Similar data sets (group of data) are used for training and testing of RBF classifier. Results are shown in Fig.20 to Fig.22.

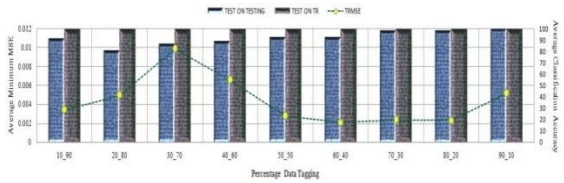


Fig. 20. Variation in average Classification Accuracy with testing on Testing and Training dataset and percent data tagged for training

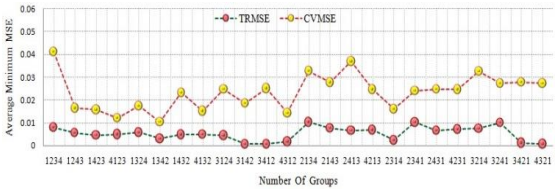


Fig.21. Variation in average Minimum MSE with Training and CV for all 24 group of Dataset

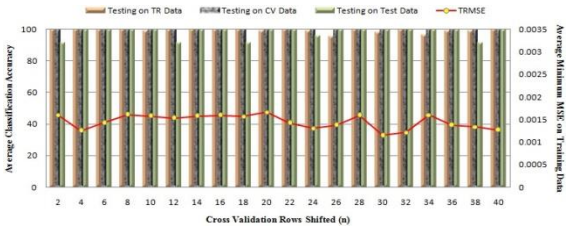


Fig. 22. Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

Finally optimized RBF based classifier gives the classification accuracy of 98.05% with average 0.0051902 training MSE and 0.0232185 average cross validations MSE.

DIMENSIONALITY REDUCTION OF CLASSIFIERS

Sensitivity Analysis

One problem appears after the feature extraction that there are too many input features that would require significant computational efforts to calculate, and this may result in low accuracy. From the analysis, most sensitive parameters are selected as input parameters. Most sensitive parameters are given as an input to MLP classifier and from MSE and percent classification accuracy numbers of inputs are selected. Variation of MSE and percent classification accuracy with different most sensitive inputs are shown in Fig.23. From these results it is clear that only the most sensitive parameters give the better results. With nine inputs classifier is once again optimized and tested, all results are shown in Fig.24 to Fig.27.

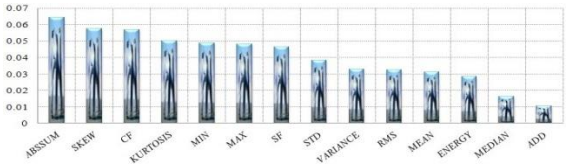


Fig.23. Sensitivity Analysis about Mean

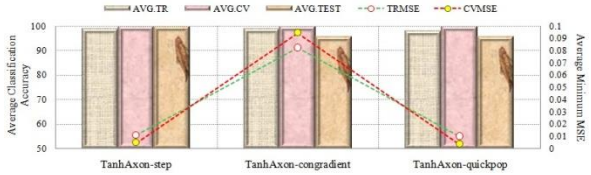


Fig .24. Variation of average MSE and classification accuracy for transfer function and learning rule

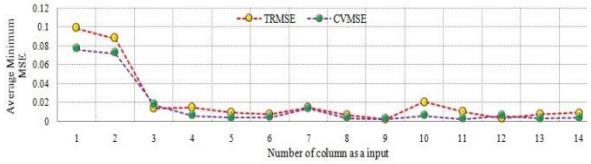


Fig. 25. Variation of average minimum MSE on training and CV dataset with number of inputs Selection of Error criterion

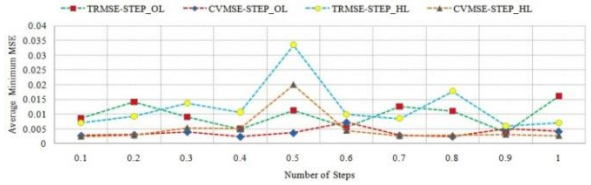


Fig. 26. Variation of average minimum MSE on TR, CV dataset with Step size of hidden and output layer

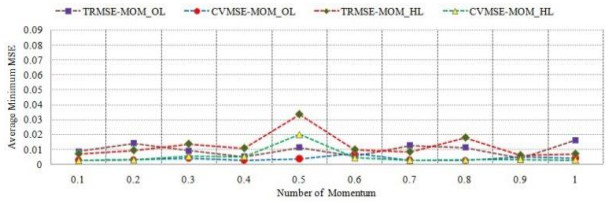
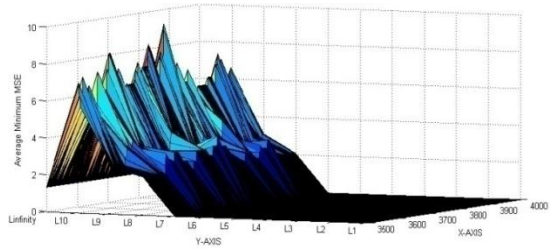


Fig. 27. Variation of average minimum MSE on TR, CV dataset with momentum of hidden and output layer

For desired response, various error criterions have been tested to select the correct error criterion and finally L-2 criterion is selected as shown in Fig.28.



Y: Error Criterion X: No. of Epochs

Fig.28. Variation in Average MSE with Error Criterion

Adopting the same procedure the NN model is optimize by several experiments and finally, the required network is designed with following specifications:

Number of Inputs: 09
Number of Hidden Layers: 01
Number of PEs in Hidden Layer: 09
Hidden Layer:
Transfer function: tanh Learning Rule: momentum
Step size: 0.9 momentum: 0.8
Output Layer:
Transfer function: tanh Learning Rule: step
Step size: 0.4 momentum: 0.7
No. of Epoch: 4500
Error Criterion: L3
Number of connection weights: 150
Time required per epoch per exemplar: 0.4077 µ-secs
Overall training efficiency: 99.81897% Average TRMSE: 0.000137

Finally, the new NN (MLP-DR-S) classifier is trained and tested for given conditions and Fig.29, Fig.30 and Fig.31 shows results.

Fig. 29. Variation of average minimum MSE with test on Testing and Training dataset and percent data tagged for training

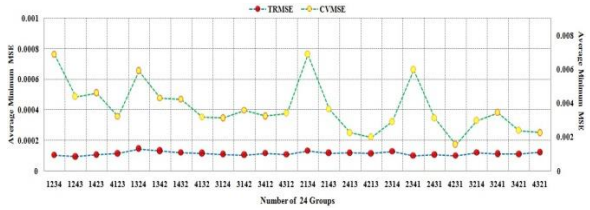


Fig. 30. Variation of average minimum MSE with Training and CV with all 24 group of dataset

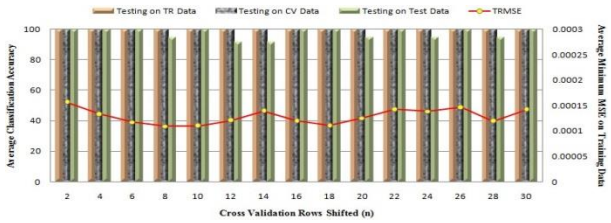


Fig.31. Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

Similar experimentation is carried out for RBF based classifier to select the proper analysis and most sensitive parameters are selected as input parameters. Number of input parameters most sensitive in descending order verses average minimum MSE on training and cross validation is as shown in Fig. 32 and average classification accuracy is shown in Fig. 33.

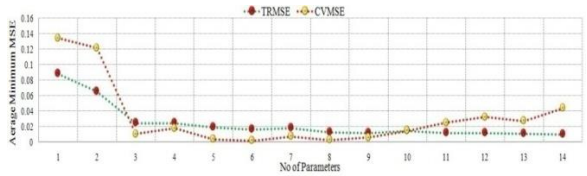


Fig.32. Variation of average minimum MSE on training and CV dataset with number of inputs

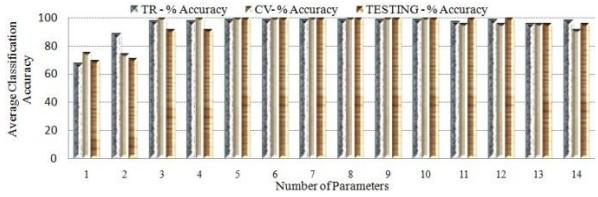


Fig.33. Variation of average classification accuracy with test on Testing, CV and TR Data with number of inputs

Using the results of sensitivity analysis, dimensions of RBF network can be reduced. Numbers of inputs are reduced to eight. By similar experimentations, the optimum RBF classifier is designed with the following changes;
Competitive Rule: Conscience Full Metric: Euclidean
Number of cluster centers: 45
Number of inputs: 8
Number of connection weights: 681
Number of connection weights reduced by: 46.20%
Number of Inputs: 08
Competitive Rule: Conscience Full Metric: Euclidean
Number of Hidden Layers: 0
Output Layer:
Transfer function: tanh Learning Rule: momentum
Step size: 0.8 momentum: 0.6
Learning Rate Start at: 0.01 Decay to: 0.001
Number of Cluster centers: 45(80)
Error Criterion: L7
Epoch: 5000

Time required per epoch per exemplar: 0.2683 micro-secs
Overall efficiency, TR - 97.87%, CV- 97.17%, TEST- 96.80%,
Overall Average TRMSE: 0.0012893
Overall Averages CVMSE: 0.01130
Time required for RBF-DR-S (Dimensionally Reduced using sensitivity analysis) is 0.2683 micro-secs
Finally, the new RBF (RBF-DR-S) classifier is trained and tested for various conditions. Results are shown in Fig.34, 35 and Fig.36.

Fig. 34. Variation of average minimum MSE with test on Testing and Training dataset and percent data tagged for training

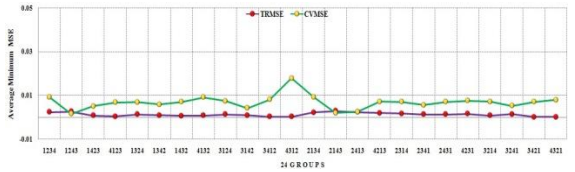


Fig. 35. Variation of average minimum MSE with Training and CV with all 24 group of dataset

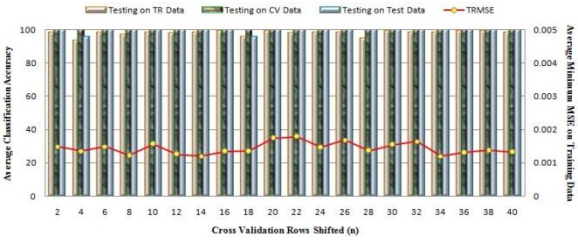


Fig. 36. Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

Principal Component Analysis

Another technique used for dimension reduction is Principal Component Analysis (PCA). It is used to reduce dimensionality of the input space and hence the reduced network can be achieved. PCA is performed by Pearson rule. Fig.37 is related to a mathematical object, the eigenvalues, which reflect the quality of the projection from the 13-dimensional to a lower number of dimensions and Fig.38, shows the correlation circle, from which it is clear that skew is the only parameter which carries the information on other axis because it is closer to the circle.

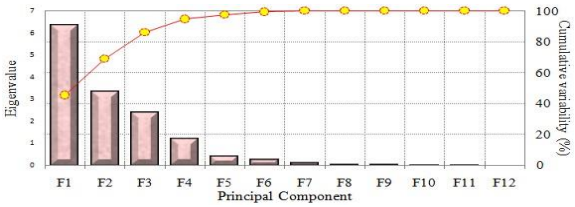


Fig.37. Principal Component, Eigen values and percent variability

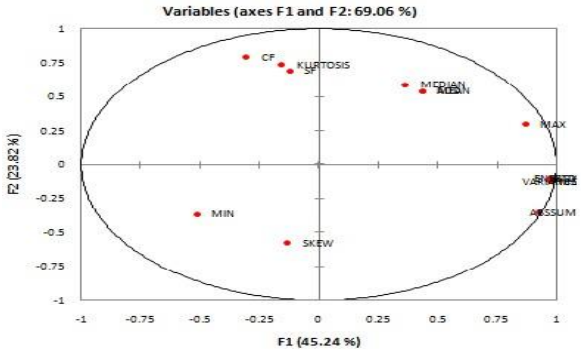


Fig.38. Correlation Circle

To select the number inputs to the MLP, average minimum MSE and average classification accuracy is verified by varying number of principal components as input. Fig.39and Fig.40 and Fig.41 shows the results of variation of average MSE with error criterion and number of processing elements as an input respectively.

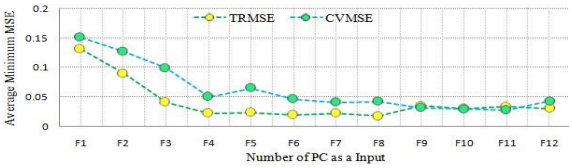
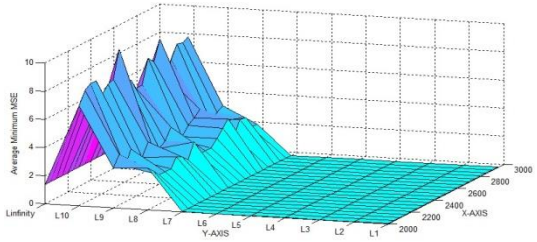


Fig.39. Variation of average minimum MSE on training and CV dataset with number of PCs as inputs



Y: Error Criterion X: No. of Epochs

Fig.40. Variation in Average MSE with Error Criterion

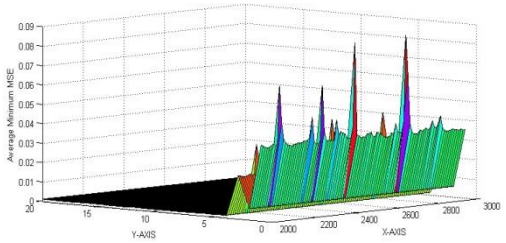


Fig.41. Variation of average classification accuracy with test on Testing, CV and TR Data with number of PEs as Inputs

From parameter variation of hidden layer and output layer the dimensionally reduced MLP (MLP DR PCA) NN model is designed with following parameters.

Number of Inputs: 7
Number of Hidden Layers: 01
Number of PEs in Hidden Layer:14
Hidden Layer:
Transfer function: tanh Learning Rule: Momentum
Step size: 0.8 Momentum: 0.8
Output Layer:
Transfer function: tanh Learning Rule: Momentum

Step size: 0.4 Momentum: 0.3
Error Criterion: L5
Epoch: 4000
Time required per epoch per exemplar: 0.2505µ-secs
Number of connection weights: 202
Percent reduction in weight: 17.32%

Finally, the new MLP (MLP-DR-PCA) classifier is trained and tested for various conditions as shown in Fig.42, 43 and Fig.44.

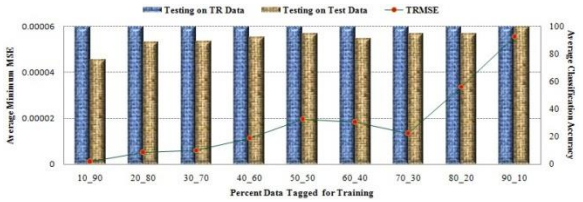


Fig. 42. Variation of average minimum MSE with test on Testing and Training dataset and percent data tagged for training

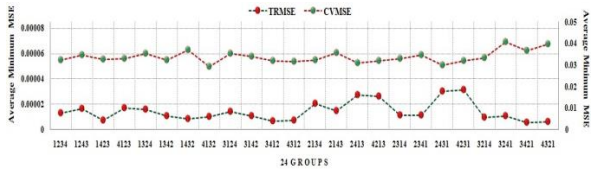


Fig. 43. Variation of average minimum MSE with Training and CV with all 24 group of dataset

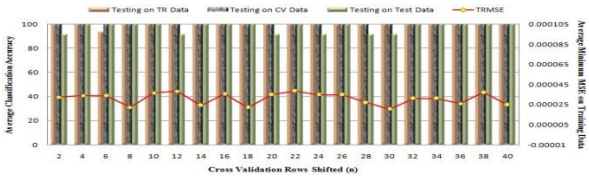


Fig. 44. Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

Same above technique is used to lowest number of dimensions for RBF and by these experimentations, the optimum RBF classifier is designed with seven inputs as shown in Fig.45.

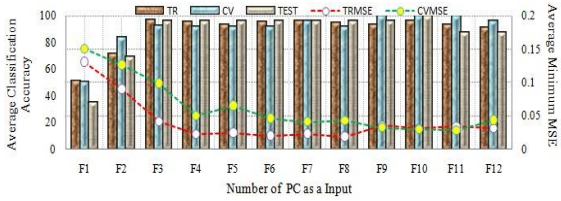


Fig.45. Variation of average classification accuracy with test on Testing, CV and TR Data with number of PCs as inputs

Number of Inputs: 7
Competitive Rule: Standard Full Metric: Boxcar
Number of Hidden Layers: 0
Output Layer:
Transfer function: tanh Learning Rule: step
Number of cluster centers: 45
Number of inputs: 7
Output Layer: Step Size: 0 .5
Learning Rate Start at: 0.01 Decay to: 0.001
Number of connection weights: 636

Number of connection weights reduced by: 49.76%
Average TRMSE: 0.0012893 Average CVMSE: 0.01130
Time required for RBF-DR-PCA (Dimensionally Reduced using PCA) is 0.4600 micro-secs
Finally, the new RBF (RBF-DR-PCA) classifier is trained and tested for various conditions as shown in Fig.46, 47 and Fig.48.

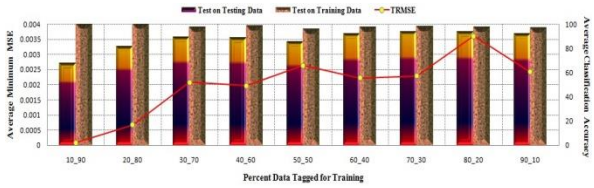


Fig. 46. Variation of average minimum MSE with test on Testing and Training dataset and percent data tagged for training

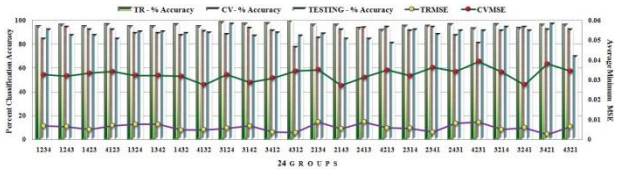


Fig. 47. Variation of average minimum MSE with Training and CV with percent accuracy for all 24 group of dataset

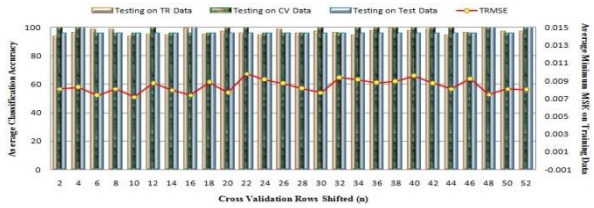


Fig.48.Variation of average MSE with Test on Training, CV and Testing dataset with CV rows shifted (n)

RESULTS AND DISCUSSION

In this paper, the authors evaluated and examined results of optimized MLP classifier and RBF based classifier for classification of PQ disturbances. Optimized MLP classifier and RBF based classifier gives classification accuracy of 99.18 % and 98.05% with average MSE of 0.0008 and 0.0051 respectively. From these results it is found that MLP based classifier is more suitable for this problem.

To reduced the complexity and improve the performance two dimensionality reduction techniques namely Sensitivity analysis and Principal component analysis are used. Using these two techniques classifier are redesigned and optimized for better performance. It is observed that Sensitivity analysis is more effective technique for dimensionality reduction. With Sensitivity analysis percent classification accuracy is increased from 99.18% to 99.81 % for MLP based classifier with 37.7% reduction in connection weights. Whereas it is not suitable for RBF based classifier because even though dimensions are reduced by 49.76 % but not significantly affect on classification accuracy.

MLP based dimension reduction classifier by sensitivity analysis works as an elegant classifier for diagnosis of typical power quality disturbances, in the sense of that, average MSE on samples is consistently observed to be reasonable such as 0.0015185. In addition, average classification accuracy on training as well as cross validation instances is found to be 98.05% and 96.92% indicating a reasonable classification. Comparative results are shown in Table –I.

REFERENCES

[1] IEEE Std. 1366. IEEE standard guide for power distribution reliability (2003).
[2] IEEE 1159.3. IEEE recommended practice for the transfer of power quality data
[3] Hamid EY, Kawasaki ZI. Wavelet-based data compression of power system disturbances using the minimum description length criterion. IEEE Trans Power Deliv 2002;17(2):460–6
[4] B. Hannaford and S. Lehman, "Short Time Fourier Analysis of the Electromyogram: Fast moments and constant contraction,"IEEE Transaction on Biomedical Engg. Vol.33, pp.1173-1181,1986

[5] W.A. Wilkinson, M.D. Cox "Discrete wavelet analysis of power system transients," IEEE Transaction on Power System 11 (1996) 2038–2044
[6] Dash PK, Panigrahi BK, Sahoo DK, Panda G. Power quality disturbance data compression, detection, and classification using integrated spline wavelet and S- transform. IEEE Trans Power Deliv 2003;18(2):595–600.
[7] Poisson O, Rioual P, Meunier M. New signal processing tools applied to power quality analysis. IEEE Trans Power Deliv. 1999;14(2):561–6.
[8] Fanibhushan Sharma, A. K. Sharma, Ajay Sharma, Nirmala Sharma "Recent Power Quality Techniques A Comparative Study", Canadian Journal on Electrical & Electronics Engineering Vol. 1, No. 6, October 2010.
[9] IEEE Recommended Practice for Monitoring Electric Power Quality, IEEE Inc.,New York, USA, 1995.
[10] Bracewell, R. N. (2000). The Fourier Transform and its Applications. McGraw- Hill Book Co, Singapore.
[11] Qian S., Chen D., "Understanding the nature of signals whose power spectra change with time-joint analysis", IEEE Signal Processing Magazine, 1999.
[12] Qian S., Chen D., "Discrete Gabor Transform", IEEE transaction signal processing, Vol. 41(7), pp. 2429-2439.
[13] S. Santoso, E. J. Powers, W. Mack Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," IEEE Transaction on Power Delivery, vol. 11, no. 2, pp. 924–930, April 1996.
[14] D. C. Robertson, O. I Camps, J. S. Mayer, andW. B. Gish, "Wavelet and electromagnetic power system transients," IEEE Transaction on Power Delivery, vol. 11, no. 2, pp. 1050–1058, April 1996
[15] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 674–693.
[16] Robert D. Nowak and Richard G. Baraniuk, "Wavelet-Based Transformations for Nonlinear Signal Processing", IEEE transactions on Signal Processing, 47(7) (1999) 1852–1865.
[17] L. Agrisani, P. Daponte, M.D. Apuzzo, and A. Testa, "A measurement method based on the wavelet transform for power quality analysis", IEEE Trans. Power Delivery, vol. 13, 990-998, Aug 1998.
[18] G.T. Heydt, A.W. Galli, Transient power quality problems analyzed using wavelets, IEEE Trans. Power Deliv. 12 (1997) 908–915.
[19] Huseyin Eris_ti, Ozal Yıldırım, Belkıs Eris_ti, Yakup Demir, "Automatic recognition system of underlying causes of power quality disturbances based on S-Transform and Extreme Learning Machine," Electrical Power and Energy Systems 61, 2014 553–562
[20] Wael R., Anis Ibrahim and Medhat M. Morcos, "Artificial Intelligence and Advanced Mathematical Tools for Power Quality Applications: A Survey", IEEE transactions on power delivery, VOL. 17(2), pp. 668-673, APRIL 2002.
[21] M. Kezunovic and I. Rikalo, "Automating the analysis of faults and power quality," IEEE Comput. Appl. Power, vol. 12, pp. 46–50, 1999.
[22] M. Faisal, A. Mohamed, H. Shareef, A. Hussain, "Power Quality diagnosis using time frequency analysis and rule based techniques ", Expert System with Applications, vol. 38, pp. 12592-12598, 2011.
[23] B. Biswal, M. Biswal, S. Mishra, R. Jalaja, "Automatic Classification of Power Quality Events Using Balanced Neural Tree," IEEE Transaction On Industrial Electronics, Vol. 61, No. 1, January 2014, 521B
[24] Elango M.K., Kumar A.N., Duraiswamy K, "Identification of powerQuality disturbances using Artificial Neural Networks", IEEEConference on Power and Energy Systems, pp. 1 – 6 , 2011.
[25] Milan Biswal, P. K. Dash "Measurement and Classification of Simultaneous Power Signal Patterns With an S-Transform Variant and Fuzzy Decision Tree," IEEE Transaction On Industrial Informatics, Vol. 9, No. 4, November 2013, 1819
[26] Morsi W.G., El-Hawary M.E., "A new fuzzy-wavelet based representative quality power factor for stationary and nonstationary power quality disturbances", IEEE Conference, pp.1 – 7, 26-30 July 2009