

# Classification of Power Quality Disturbances at Transmission System Using Support Vector Machines

Shahrani Shahbudin, Zaki Firdaus Mohamad, Saiful Izwan Suliman\*, Murizah Kassim, Roslina Mohamad

Centre for Computer Engineering, Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

\*Corresponding author, e-mail: shhranis@salam.uitm.edu.my

## Abstract

Power Quality has become one of the important issues in modern smart grid environment. Smart grid generally utilizes computational intelligence method from the generation of electricity to electricity distribution to the customers. This is done for the safety, reliability, tenacity and efficiency of the system. The classification of power disturbances has become a major topic in maintaining power quality. These disturbances occur due to faults, natural causes, load switching, energizing transformer, starting large motor, as well as utilization of power electronic devices. The key issue is about maintaining the continuous supply of electricity to the end-users without any problem. If a problem occurs, it might increase the production cost significantly especially to large-scale industries. In this paper, S-transform is used to extract distinctive features of real data from transmission system, and Support Vector Machine was utilized to classify four types PQ disturbances namely, voltage sag, interruption, transient and normal voltage. Results obtained indicate that performance of the One Against One classifier produces high accuracy using k-fold cross validation and RBF kernel.

**Keywords:** power quality disturbances, S-Transform, k-fold cross validation, One-against-One SVM

**Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.**

## 1. Introduction

The increasing demands of electricity nowadays has become a real challenge to the power producer to maintain the quality of continuous power supply. Low quality of continuous power supply will affect many aspects especially production cost to the industries. Computational intelligence (CI) methods can be utilized as a tool to help the power quality management to maintain and improve the electricity system. CI is mainly used for fault detection, fault classification and section identification of fault in transmission lines for analysis of disturbance that can lead to any fault [8]. Moreover, power quality disturbances classification in transmission system has also become one of the most researched area in this area. The significance of this study is to improve the power quality (PQ) monitoring system in minimizing fault occurrence along the transmission lines.

Fault events in transmission line occur due to several factors such as bad weather, human activities and accidents. Lightning is one of the main cause of voltage disturbance (e.g. transient overvoltage). It can lead to line-to-ground fault event due to lack of insulation around cable [1]. The occurrences of fault events have resulted to power quality disturbances in the transmission lines. There are many types of disturbances that may affect power quality. These include voltage sag, voltage swell, transient, harmonic, interruption and flicker. These disturbances occur due to many factors such as faults, natural causes, load switching, energizing transformer, starting large motor, and high utilization of power electronic devices [2]. The diagnosis of power quality disturbances is important for the improvement of monitoring process in power transmission system particularly in the context of voltage stability. Voltage stability is the ability of power system to remain at acceptable voltage at all busses under normal operating system and after being affected by the disturbance [3]. Failure to maintain voltage stability can lead to the occurrence of voltage collapse in transmission system.

Support Vector Machine (SVM) is one of the utilized technique in solving classification problems. The goal of SVM is to determine a classifier that minimizes empirical risks namely training set error and confidence interval. This confidence interval corresponds to the

generalization or test set error [4, 5]. Support vector machine classifiers are based on the statistical learning theory. The method is suitable for large size of training data. Compared to other classification methods, no threshold is to be determined in SVM. An upper bound will be utilized for the generalization performance (i.e. the performance for the test set). This upper bound will be determined based on the statistical learning conducted [6, 7]. The applications of support vector machine have been utilized in solving many problems occurring in power system operation. These include power disturbances classification methods, fault classification and identification faulty section occurred in transmission lines [8].

Various classification techniques have been successfully applied in PQ disturbances diagnosis such as wavelet transform and SVM [2], [5], [9-11], S-Transform and SVM [12], wavelet Transform and optimized ANN [13], curvelet Transform and SVM [14], Quarter-Sphere Support Vector Machine (QSSVM) [15], Hybrid neural network [1] and ensemble technique [16]. Support vector machine is a well-known supervised classification technique that can classify any type of objects or signals accurately. This method has been proven to be a good tool in solving many problems in transmission lines and machinery faults [2]. It is an effective classification tool for fault classification in power system.

There are two strategies for multi-class SVM which are “one-against-one” and “one-against-all”. However, based on the studies conducted by many researchers found in the scientific literature, the “one-against-one” strategy takes less computational time during the training process. It performs well for problems with very large number of classes [11], [17, 18]. Therefore, “one-against-one”- based multi class SVM (OAO-SVM) approach was chosen as the classification method in this study. The objective of this study is to classify the disturbances for real data from transmission system in Malaysia using “one-against-one” multi-class support vector machine technique.

## 2. Research Method

The proposed method in this study is divided into three stages. It starts with the pre-processing stage, followed by feature extraction stage and finally the implementation of multiclass SVM in the classification phase. The processes involved in this study utilizing SVM is shown in Figure 1.

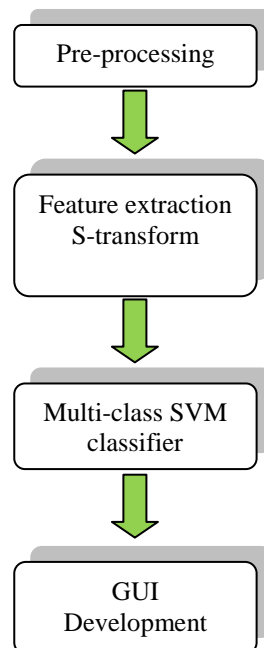


Figure 1. Flow of the proposed method

## 2.1. Pre-Processing

A pre-processing of the signals is required to normalize raw data. This is because signals are collected from various voltage levels in the distribution system [19]. The collected data are 3-phase voltages (UR, UY, UB), 3-phase currents (IR, IY, IB) and neutral line current (IN). The samples data collected are shown in Table I. However, only values for 3-phase voltages were utilized during this process. Meanwhile, four types of disturbances were classified in this study which are voltage sag, interruption, transient and normal voltage.

Table 1. Sample of the Real Data of Transmission System for One Feeder

(t)	UR	UY	UB	IR	IY	IB	IN
1	-15008	17632	-2528	-80	96	0	0
2	-15632	17152	-1408	-80	80	16	16
3	-16224	16576	-288	-80	80	16	16
4	-16768	15952	864	-96	80	32	16
5	-17280	15264	2016	-96	80	32	16
6	-17792	14496	3200	-96	80	32	16
7	-18224	13680	4400	-96	64	48	0
8	-18560	12800	5568	-96	64	48	16
9	-18816	11840	6768	-96	64	48	16
10	-18960	10832	7952	-96	64	64	16

## 2.2. Feature Extraction using S-transform

After the pre-processing process, S-transform process was applied in the feature extraction stage. S-transform is one of the feature extraction methods based on time-frequency analysis. It extracts the pre-processing data into the most salient features that represent the power quality phenomenon [20]. The following was applied for the s-transform process:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \quad (1)$$

S-transform will generate a matrix known as S matrix. Each row of the s-matrix represents specific frequency, and each column corresponds to a sampling point of the original signal. Every element from this matrix is a complex value. According to Wenda et al. [21], statistical techniques are used to the amplitude of contour matrix of S-transform by using maximum amplitude and frequency amplitude plot. The resulting features are formulated as follows.

Amplitude factor, F1 as the first feature with the given equation:

$$F1 = 1 + std1 + std2 - norm1 - norm2 \quad (2)$$

where,

*std1* : maximum value from standard deviation of distorted signal.

*std2* : minimum value from standard deviation of distorted signal.

*norm1* : maximum value of normal signal.

*norm2* : minimum value of normal signal.

The second feature is given as :

$$F2 = std1 - norm1 \quad (3)$$

The third feature is given as :

$$F3 = \text{mean}(\text{mean}(\text{abs}(ds)^2)) \quad (4)$$

where,

$ds$  = absolute value of S-transform from distorted signal

The fourth feature is an absolute value of S-transform from highest amplitude of frequency ( $f_m$ ). The equation is given as follows,

$$F4 = abs(ds(f_m)) \quad (5)$$

The fifth feature is the mean of the square root STD from amplitude of st-matrix,

$$F5 = mean(\sqrt{std(ST - amplitude)}) \quad (6)$$

All of the five features were used as the inputs to the multi-class OAO-SVM.

### 2.3. Multi-Class SVM Classifier

SVM was originally developed for binary classification which can only classify two classes of data. Training data set must be in form of  $(x_1, y_1), \dots, (x_l, y_l)$  where  $x_i \in \mathbb{R}^n$  is a feature vector and  $y_i \in \{-1, 1\}$  is a label of class. The largest margins that separate these data are defined as decision surface. However, this decision surface is not generated by the input space, but is decided in the high-dimensional feature space. Kernel function was used because it is very useful in nonlinear data. Equations (7) - (9) show how binary classification took place [22].

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (7)$$

where,

$\alpha$  : Lagrange multiplier

$y$  : Class label

$K$  : The similarity between pattern  $x_i$  and  $x_j$  from the stored training set.

$x_i, x_j$  : Two different patterns of inputs

Subject to the constraints,

$$\sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \quad (8)$$

After finding the optimal value,  $\alpha_i$ , the decision boundary needs to be constructed in the form of :

$$f(x) = \sum_{\alpha_i \neq 0} y_i \alpha_i K(x_i, x) + b \quad (9)$$

where,

$x$  : the class from the sign of  $f(x)$

$x_i$  : support vector correspond to  $\alpha_i \neq 0$

$b$  : threshold of the decision boundary from origin

In making SVM computationally very efficient, the numbers of support vectors are considerably lower than the number of training samples. The regularization parameter  $C$ , controls the trade-offs between the margin and the size of the slack variables [6].

Multi-class SVM uses combination of binary SVMs for  $k$  number of classes. OAO-SVM is involved in the construction of binary SVM classifiers for all pairs of classes. In overall, there are  $[k(k-1)/2]$  pairs, which implies that there is only one binary SVM for one pair. The classification of unknown input pattern is determined according to maximum voting of all SVMs as shown in Figure 2.

In the classification stage, a model validation technique was utilized for assessing how the results of a statistical analysis will generalize to an independent data set. There are various techniques in dividing data set into training and testing set, such as Hold-out, k-fold cross validation (k-fold CV), LeaveOut, and re-substitution. However, only two techniques were implemented and tested in this work, which are Holdout and k-fold CV.

Hold-out validation technique splits the data without overlapping the training and testing data. In k-fold CV, the training set is split into k smaller sets. A model is trained using the folds as training data. The resulting model is then validated and tested on the remaining part of the data. The training and testing processes were repeated for each different combination. In this paper, the value of k was set from two until ten folds.

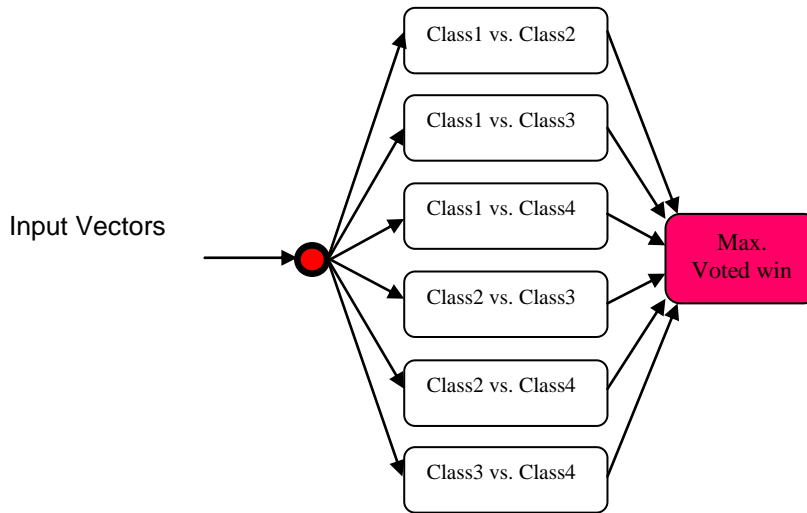


Figure 2. OAO-SVM classifier

Polynomial (Poly) and Radial Basic Function (RBF) are non-linear kernel function commonly used in SVM classification. Equation (10) and (11) show the basic formula of Poly and RBF.

$$\text{Poly: } K(\bar{x}, \bar{y}) = (\omega \bar{x} \cdot \bar{y} + 1)^d \tag{10}$$

$$\text{RBF: } K(\bar{x}, \bar{y}) = \exp(-\gamma ||\bar{x} - \bar{y}||^2) \tag{11}$$

### 3. Results and Analysis

#### 3.1. S-transform Feature Extraction

Figure 3 depicts four plots for S-transform results for each disturbance. The first plot shows the per unit voltage for the actual data and the contour of the s-matrix area shown in the second plot. Final plot represents the maximum amplitude of the frequency for each sampling points. The sample outputs of S-transform are tabulated in Table 2.

Table 2. Sample of Features Vectors Data Used in OAO-SVM

Types	F1	F2	F3	F4	F5
Sag	1.0654	0.0335	0.0011	0.4999	0.0364
Interruption	1.0402	0.0337	0.0007	0.4216	0.0182
Transient	1.0881	0.0537	0.0013	0.0621	0.5100
Normal	1.0728	0.0364	0.0012	0.5420	0.0182

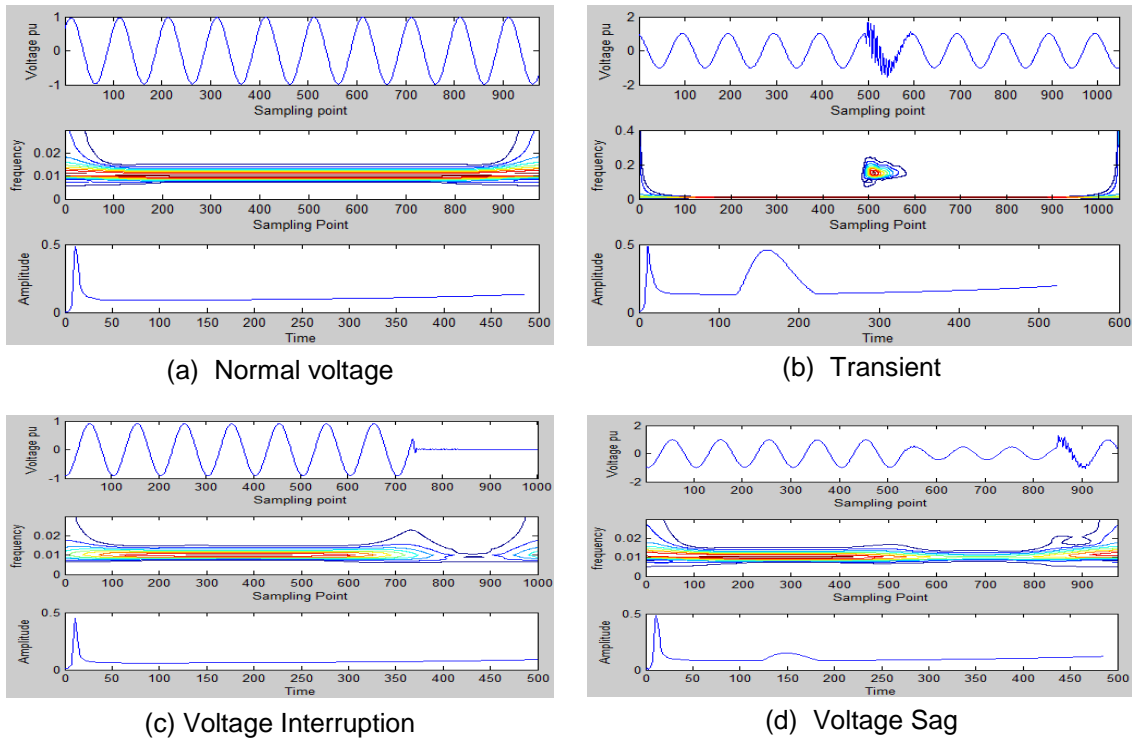


Figure 1. S-transform results for each disturbance (a) Normal Voltage, (b) Transient, (c) Voltage interruption, (d) Voltage sag

**3.2. OAO-SVM Classifier**

Several parameters need to be set up before the training process. In this analysis, for k-fold CV, the value of k is set between 2 to 10, and the best average accuracy was recorded in Table 3 and the results were compared with Holdout technique as shown in Table 4. Holdout technique used 50% for both training and testing. In addition, two types of kernels were used for both techniques, which are Polynomial (Poly) and Radial Basis Function (RBF). Table 5 represents the confusion matrix generated after classification stage using OAO-SVM using k-fold CV with k=9.

The average accuracies for both techniques were recorded for ten readings. Based on the data in Table V, the best k value of 9 was obtained when k-fold technique was utilized. Accuracy rate of 100% was achieved using RBF kernel as compared to 99.623% for Poly kernel. Meanwhile, in Holdout cross validation technique, Poly kernel performs better with average accuracy of 99.582% as compared to 98.954% by RBF kernel. Therefore, it shows that the combination of k-fold cross validation (k-fold CV) technique and RBF kernel produces the superb accuracy rate.

Table 3. Accuracy Rate of K-Fold When K Varies from 2 Until 10

Value of k	Accuracy (%)
2	99.16
3	99.37
4	99.16
5	98.96
6	99.20
7	98.55
8	98.31
9	100.0
10	89.58

Table 4. Output Sample Of OAO-SVM Using K-Fold CV (K=9) With RBF Kernel

Type of disturbances	Sag	Interrupt	Transient	Normal	Total(%)
Sag	5	0	0	0	100
Interruption	0	11	0	0	100
Transient	0	0	14	0	100
Normal	0	0	0	23	100
Accuracy	100 %				

Table 5. Comparison Results of the OAO-SVM with Two Cross Validations Techniques

Cross validation techniques	K-fold (K=9)		Holdout	
	Poly	RBF	Poly	RBF
Average accuracy (%)	99.623	<b>100</b>	99.582	98.954

### 3.3. Graphical User Interface (GUI)

The graphical user interface (GUI) for the proposed method was developed as illustrated in Figure 4. It consists of two parts, which are the feature extraction stage on the left side and classification stage on the right side. This model enables the user to conduct experiments as well as analyse parts easily based on the specific requirements for the power quality disturbances diagnosis system.

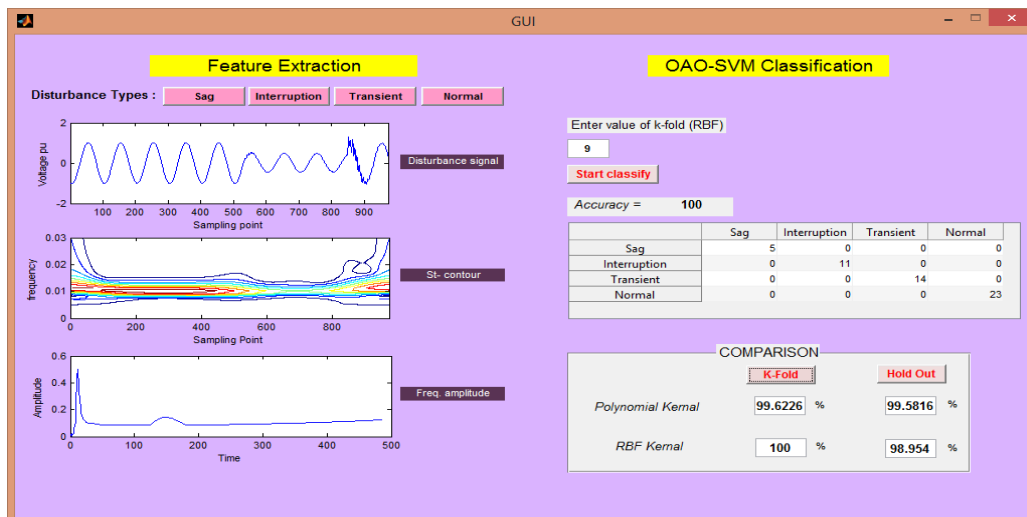


Figure 4. Developed Graphical User Interface

### 4. Conclusion

Continuous supply of electricity is very essential as almost all our daily activities depend on electricity power. Big industries will be the most affected if electricity supply is interrupted as machines and heavy equipment operate on electricity power. Disturbance in electricity supply will have direct effect on the operational cost of a factory, thus will incur extra cost and reduce the profit. Therefore, it is very important to have a mechanism to minimize the severity of supply interruption in the event of a disturbance or voltage collapse. This study investigates the application of support vector machine (SVM) for the classification of power quality disturbance by using s-transform and One-Against-One SVM (OAO-SVM) classifier. Based on the presented results and analysis conducted, the proposed method has managed to classify the disturbances with 100% accuracy rate when combination of k-fold cross validation (k=9) and Radial Basis Function (RBF) kernel was utilized. Therefore, OAO-SVM shows great potentials for future power quality disturbances classification method.

## Acknowledgment

The authors would like to express the gratitude to the Ministry of Higher Education, Malaysia and Universiti Teknologi MARA, Selangor, Malaysia for the financial support given for this project (Fundamental Research Grant Scheme - FRGS) [File No : 600-RMI/FRGS 5/3 (109/2014)].

## References

- [1] Huang NT, et al. *Power quality disturbances classification by ensemble and hybrid Neural networks*. Power System Technology (POWERCON), 2010 International Conference on. IEEE. 2010.
- [2] Erişti H, Demir Y. Automatic classification of power quality events and disturbances using wavelet transform and support vector machines. *IET generation, transmission & distribution*. 2012; 6(10): 968-976.
- [3] Ibe OG, Onyema AI. Concepts of reactive power control and voltage stability methods in power system network. *IOSR Journal of Computer Engineering*. 2013; 11(2): 15-25.
- [4] Lin WM, Wu CH, Lin CH, Cheng FS. *Classification of multiple power quality disturbances using support vector machine and one-versus-one approach*. Power System Technology. IEEE. Power Con 2006. International Conference on 2006: 1-8.
- [5] Tong W, Song X, Lin J, Zhao Z. *Detection and classification of power quality disturbances based on wavelet packet decomposition and support vector machines*. In Signal Processing, IEEE 2006 8th International Conference on 2006; 4.
- [6] Shawe-Taylor J, Cristianini N. *Kernel methods for pattern analysis*. Cambridge university press. 2004; 28.
- [7] Vapnik V. *The nature of statistical learning theory*. Springer science & business media. 2013; 29.
- [8] García VV, Gualdrón CA, Plata GO. *Obtaining patterns for classification of power quality disturbances using biorthogonal wavelets, RMS value and support vector machines*. Electrical Power Quality and Utilisation. IEEE. EPQU 2007. 9th International Conference on 2007: 1-6.
- [9] Gao P, Wu W. *Power quality disturbances classification using wavelet and support vector machines*. Intelligent Systems Design and Applications. Sixth International Conference on 2006; 1: 201-206.
- [10] Latha RS, Babu CS, Prasad KD. *Detection & analysis of power quality disturbances using wavelet transforms and SVM*. *International Research Journal of Signal Processing*. 2011; 2(02): 58-69.
- [11] Yang BS, Han T, Hwang WW. *Fault diagnosis of rotating machinery based on multi-class support vector machines*. *Journal of Mechanical Science and Technology*. 2005; 19(3): 846-59.
- [12] ÇÖTELİ R. *A combined protective scheme for fault classification and identification of faulty section in series compensated transmission lines*. *Turkish Journal of Electrical Engineering & Computer Sciences*. 2013; 21(Sup. 1): 1842-56.
- [13] Ijaz M, Shafiullah M, Abido MA. *Classification of power quality disturbances using Wavelet Transform and Optimized ANN*. Intelligent System Application to Power Systems (ISAP), IEEE. 2015 18th International Conference on 2015: 1-6.
- [14] Risnidar CB, Daut I, Yussof MI, Warman E, Hasan S, Nisja I. *Investigation of Power Factor on Harmonic Effect due to Types of Voltage Source*. *International Journal on Advanced Science, Engineering and Information Technology*. 2011; 1(5): 476-80.
- [15] Shahid N, Aleem SA, Naqvi IH, Zaffar N. *Support vector machine based fault detection & classification in smart grids*. Globecom Workshops (GC Wkshps), 2012 IEEE 2012: 1526-1531.
- [16] Karasu S, Başkan S. *Classification of power quality disturbances by using ensemble technique*. Signal Processing and Communication Application Conference (SIU). IEEE. 2016: 529-532.
- [17] Tripathi P, Sharma A, Pillai GN, Gupta I. *Accurate fault classification and section identification scheme in TCSC compensated transmission line using SVM*. World Academy of Science, Engineering and Technology. 2011; 60: 1599-605.
- [18] Thukaram D, Agrawal R. *Discrimination of Faulted Transmission Lines Using Multi Class Support Vector Machines*. 16th National Power Systems Conference. 2010: 497-502.
- [19] Amirullah A, Penangsang O, Soeprijanto A. *Power Quality Analysis of Integration Photovoltaic Generator to Three Phase Grid under Variable Solar Irradiance Level*. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*. 2016; 14(1): 29-38.
- [20] Zhang X, Han G, Chen L. *De-noising of Power Quality Disturbance Detection Based on Ensemble Empirical Mode Decomposition Threshold Algorithm*. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*. 2013; 11(4): 637-644.
- [21] Wenda A, Hussain A, Samad SA, Mohamed R, Hannan MA. *Web-based on mobile phone for automatic classification of power quality disturbance using the S-transform and support vector machines*. TENCON 2010-2010 IEEE Region 10 Conference 2010: 1432-1437.
- [22] Duan KB, Rajapakse JC, Nguyen MN. *One-versus-one and one-versus-all multiclass SVM-RFE for gene selection in cancer classification*. European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics. Springer Berlin Heidelberg. 2007: 47-56.