Feature subset selection in dynamic stability assessment power system using artificial neural networks

- **Nguyen Ngoc Au [1]**
- **Quyen Huy Anh [1]**
- **Phan Thi Thanh Binh [2]**

[1]Ho Chi Minh city University of Technical and Education
[2]Ho Chi Minh city University of Technology, VNU-HCM

(Manuscript Received on October 30nd, 2014, Manuscript Revised July 08nd, 2015)

ABSTRACT

This paper presents method of feature subset selection in dynamic stability assessment (DSA) power system using artificial neural networks (ANN). In the application of ANN on DSA power system, feature subset selection aims to reduce the number of training features, cost and memory computer. However, the major challenge is to reduce the number of features but classification rate gets a high accuracy. This paper proposes applying Sequential Forward Selection (SFS), Sequential Backward Selection (SBS), Sequential Forward Floating Selection (SFFS) and Feature Ranking (FR) algorithm to feature subset selection. The effectiveness of the algorithms was tested on the GSO-37bus power system. With the same number of features, the calculation results show that SFS algorithm yielded higher classification rate than FR, SBS algorithm. SFS algorithm yielded the same classification rate as SFFS algorithm.

Key words: *feature subset selection, dynamic stability assessment, artificial neural networks, and power system.*

1. INTRODUCTION

Modern power systems are forced to operate under highly stressed operating conditions closer to their stability limits. The operation of power systems is challenged increasingly significant because investment sources and transmission systems are not developed to meet the load demand. While operating the power system is always faced with unusual circumstances such as a generator outage, loss of a line, sudden dropping of a large load, switching of station or substation, and

three-phase sudden short circuit, ... Power system stability is the ability to regain an equilibrium state after being subjected to a physical disturbance and maintain the continuous supply of electricity to customers. Power system stability is classified [1]: rotor angle stability, frequency stability and voltage stability. Rotor angle stability is divided into two categories including short-term and long-term. Short-term stability angle is considered transient dynamic stability and important contribution in power system stability. Long-term stability angle includes small signal stability and frequency stability.

Due to the complexity of the power system, traditional methods to power system analysis take so much time and cause delays in decision making. However, the relationship between pre-fault parameters of the power system state and post-fault modes of power system stability has highly nonlinear, extremely difficult to describe this mathematical relationship. In order to overcome such difficulties, intelligent system, that is ANN, has been proposed for DSA thanks to special abilities in pattern classification [2],[6],[7]. Operating conditions of power systems have wide range so that it is difficult perform online calculations. ANN is in need of initial off-line data for training. Extensive off-line simulation is performed so as to acquire a large enough set of training data to represent the different operating conditions of typical power systems. As a pattern classifier, once trained, neural networks not only have extremely fast solutions but also get the ability to update new patterns or new operating conditions by generalizing the training data, improving recognition accuracy [7].

The intelligent systems for DSA consist of four basic steps: database generation, feature selection, knowledge extraction and model validation. In particular, a very important stage is feature selection because it greatly affects cost, computational time and recognition accuracy of DSA system. Feature selection actually reduces features or variables, just select the minimum number of variables but ensure recognition accuracy. This paper proposed applying FR (Feature Ranking), SFFS (Sequential Forward Floating Selection), SFS (Sequential Forward Selection), SBS (Sequential Backward Selection) algorithm for feature subset selection. The case study was done on GSO-37bus power system diagram with the support of simulation software PowerWorld 17. The algorithms of feature subset selection were programmed on Matlab software. Multilayer Feed forward Neural Networks (MLFN) is supported by Matlab software.

2. METHOD

2.1. Mathematical Model of Multimachine Power System

The dynamic behavior of a generator power system can be described by the following differential equations [1]:

$$M_i \frac{d^2\delta_i}{dt^2} = P_{mi} - P_{ei} \qquad (1)$$

It is known that: $\frac{d\delta_i}{dt} = \varpi_i$ (2)

By substituting (2) in (1), therefore (1) becomes:

$$M_i \frac{d\varpi_i}{dt} = P_{mi} - P_{ei} \qquad (3)$$

Where: δ_i: rotor angle of machine i; ω_i: rotor speed of machine i; P_{mi}: mechanical power of

machine i; P_{ei}: electrical power of machine i; M_i: moment of inertia of machine i.

The state of the power system is stable when the rotor angle deviation of any two generators not exceeding 180^0, and is unstable when the rotor angle deviation of any two generators exceed 180^0. Status of power system was performed according to the proposed rules in [1],]4],[5], as follow:

$$\begin{aligned} &\text{If } \delta_{ij} < 180^0 \text{ then Stable} \\ &\text{If } \delta_{ij} \geq 180^0 \text{ then Unstable} \end{aligned} \quad (4)$$

2.2. Feature subset selection

2.2.1. General Description

The MLNF-based DSA power system can be formulated as a mapping $y_i = f(x_i)$ after learning from a stability database $D=\{x_i, y_i\}_{i=1}^{n}$. Where x_i is feature; It is n-dimensional input vector that characterizes the system operating state; and y_i is output vector. The feature subset selection consists of selecting a *d* dimensional feature vector *z*. Where $d < n$; The *d* selected features represent the original data in a new knowledge base $D_{new}=\{z_i, y_i\}_{i=1}^{d}$, and the new mapping $y_{newi}=f_{new}(z_i)$. Thus, feature selection is actually taking away unnecessary features and selecting a candidate subset of features that get rich information with highly accurate identification of model. This process includes the following steps:

Step 1. Data generation, initial feature set selection.

Step 2. Candidate feature subset selection.

Step 3. Training and testing classification rate.

Step4. Subset feature evaluation.

Step 5. Subset feature selection.

2.2.2. Data generation, initial feature set selection.

A large number of samples are generated through off-line simulation and the stable status is evaluated for each fault under study. Data for each bus or line fault occurring in the test systems are recorded in which samples of data are kept in a database. The input is the vector of system state parameters that characterize the current system state, usually called feature, they can be classified into pre-fault, fault-on and post-fault features.

Pre-fault features [2]: steady-state operating parameters such as voltage magnitude and angle of buses, P, Q load, generation and line flow qualities P_{flow}, Q_{flow}, P_{load}, Q_{load}, V_{bus}, and before disturbance occurs (P_{gen}, Q_{gen}, δ_{bus},…).

Fault-on features [6]: variables that characterize at fault-on state of power system occur such as changes in nodal powers, in power flows in transmission line, voltage drops in the nodes at instance of fault (ΔP_{flow}, ΔQ_{flow}, ΔP_{load}, ΔQ_{load}, ΔV_{bus},…).

Post-fault features [4]: variables that describe system dynamic behavior after disturbance occurs such as relative rotor angle, rotor angular velocity, rotor acceleration, rotor kinetic energy, and the dynamic voltage trajectory,…

The problem of transient stability is usually divided into two main categories: assessment and prediction. Transient stability assessment usually focuses on the critical clearing time (CCT). In transient stability prediction, the CCT is not of interest [11]. In this aspect, the progress

of power system transient due to the occurrence of disturbance is monitored. The key question in transient stability prediction is: the transient swings are finally 'Stable' or 'Unstable' [3], [10]-[12]. Vector output variables represent the stable conditions of the power system. Need of fast DSA power system after the fault is stable or unstable, so the output variables are assigned to label binary variable y [10, 01]. Class 1 [10] is stable class and class 2 [01] is unstable class.

The use the post-fault variables can be too long for operators to take timely remedial actions to stop the extremely fast transient instability development process.

Found that, pre-fault input features are variables that are too difficult to find a clear signal for sampled dataset learning. Post-fault input features will prolong a warning of instability power system. Fault-on input features are proposed in [6] to overcome the drawbacks such as analysis since the changes in the value of the parameters of input variables are a clear signal for dataset learning. So, this paper did mining of fault-on input features (ΔV_{bus}, ΔP_{load}, ΔQ_{load}, ΔP_{flow}, ΔQ_{flow}) as a database for training neural networks.

The output variables represent the dynamic behavior of power system at fault-on. By observation from off-line simulation, these binary output variables indicate the status of the power system to comply with the law (4).

The quantitative variables have different units of measurement; the value of the variables in the different ranges will affect the calculation results in recognition. Data normalization methods commonly applied in accordance with the following formula:

$$z_i = \frac{x_i - m_i}{\sigma_i} \qquad (5)$$

Where: m_i is mean value of data. σ_i is standard deviation of data.

2.2.3. Candidate feature subset selection.

This step is the process of searching for potential subset features. The search strategy is divided into a global search and local search. Global search strategy has the great advantage that for optimal result, but expensive computation time. Therefore, the optimal search strategy is not appropriate when a large number of input variables. In the case of large input feature, local optimization search strategy will spend less time searching because the search process is not through the entire search space.

2.2.3.1. Local optimization search strategies

- Sequential Forward Selection – SFS [8]: The SFS method begins with an empty set (k=0), adds one feature at a time to selected subset with (k+1) features so that the new subset maximizes the cost function $J_{(k+1)}$. It stops when the selected subset has the d desired number of features, k<d.

-Sequential Backward Selection-SBS [8]: The SBS method begins with all input features D (k=D), removes one feature at a time to selected subset with (k-1) features so that the resultant subset maximizes the cost function $J_{(k-1)}$. The algorithm stops when the resultant feature set has the d desired number of features, k<d.

- Sequential Forward Floating Selection-SFFS [8]: The SFFS is one of two algorithms of Floating Search Algorithm (FSA) that are SFFS and SBFS (Sequential Backward Floating Selection). The SFFS algorithm the search starts with an empty feature set and uses the SFS algorithm to add one feature at a time to the

selected feature subset. Every time a new feature is added to the current feature set, the algorithm tries to backtrack by using the SBS algorithm to remove one feature at a time to find a better subset. The algorithm terminates when the size of the current feature set is larger than the d desired number of features.

-Feature Ranking-FR [2],[4]: This is a simple method which uses less computing time. By evaluating cost function of a single feature, then it is ranked by ordering the best of them and select for a good feature.

2.2.3.2. Cost function [8, 9]

Let the *n* data samples be $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. The sample covariance matrix, S_m, is given by (6):

$$S_m = \frac{1}{N}\sum_{n=1}^{N}(x_n - m)(x_n - m)^T \quad (6)$$

The sample mean of all data:

$$m = \frac{1}{N}\sum_{n=1}^{N} x_n \quad (7)$$

The sample mean of class c_i:

$$m_i = \frac{1}{N_i}\sum_{x_n \in c_i} x_n \quad (8)$$

Where: c is the number of class; N_i is the number of sample mean of class c_i; N is the number of all samples.

S_W, within-class scatter Matrix, is:

$$S_w = \frac{1}{N}\sum_{i=1}^{c}\frac{N_i}{N}S_i \quad (9)$$

$$S_i = \frac{1}{N_i}\sum_{x_n \in c}(x_n - m_i)(x_n - m_i)^T \quad (10)$$

S_i: is the covariance matrix for class i.

Between-class scatter matrix that describes the scatter of the class means about the total mean is:

$$S_b = \sum_{i=1}^{c}\frac{N_i}{N}(m_i - m)(m_i - m)^T \quad (11)$$

S_m is the covariance matrix of the feature vector with respect to the global mean. Its trace is the sum of variances of the features around their respective global mean. S_m is:

$$S_m = S_w + S_b \quad (12)$$

Goal is to find a feature subset for which the within-class spread is small and the between-class spread is large. The cost function is:

$$J = trace\{S_w^{-1} S_m\} \quad (13)$$

Formula (14), that was written for the k^{th} single feature, is Fisher distance function:

$$J^{(k)} = \frac{S_b^{(k)}}{S_w^{(k)}} \quad (14)$$

The value of J is bigger means that the feature is more important.

2.2.4. Training and testing classification rate

To test the studied methods without loss of generality, the database is randomly partition into k subsets that are $D_1, D_2, \ldots, D_i, \ldots, D_k$, each equal size. The model is trained on all the subsets except for one that is tested to measuring of validation accuracy. Training and testing are performed k times. The validation accuracy or classification rate is computed for each of the k validation sets and averaged to get a final cross-validation accuracy. Classification rate of training or testing is determined by the formula (15):

$$r(\%) = \frac{n_r}{N}.100 \quad (15)$$

Where: n_r is the number of sample for training or testing with right result; N is the number of sample for training or testing.

The expected value (EV) of classification rate of the model was proposed in [6] by the formula (16):

$$EV \geq 0.9 \qquad (16)$$

2.2.5. Training and testing classification rate and subset feature evaluation

Applying feature subset selection algorithms were described as above to selecting feature subsets. Each feature subset was trained and tested, the classification rates are calculated by the formula (15).

Feature subset is selected with conditions that have smaller a number of features, agree to the formula (16) and get higher classification rate.

3. RESULTS - DISCUSSION

3.1 Feature set, samples for training

The off-line simulation was implemented to collection data for training. In this study, the GSO-37bus system, that is the standard system in the simulation program of PowerWorld 17 software, [5], was used as case study. It consists of 37 buses, 9 generators; three different voltage levels are 345kV, 138kV and 69kV, 25 loads, 14 transformers, 42 transmission lines. Load level is one hundred percent rated load. Fault types are balanced three-phase, single line to ground, line to line, double line to ground at buses and along transmission lines. Setting fault clearing time is 25ms [5] with all faults.

Input and output variables are x[ΔV_{bus}, ΔP_{load}, ΔQ_{load}, ΔP_{flow}, ΔQ_{flow}] and y[10,01]. Totalof input variables is 199 (37+25+25+56+56). The number of output variables is 2 (class 1 [10]: stable class, class 2 [01]: unstable class). From simulated results and based on the law (4), there were 240 samples with 120 stable samples and 120 unstable samples. Sample set was normalized by formula (5). Full feature set was randomly divided into 6 feature subsets. Each feature subset had 40 samples (20 stable samples and 20 unstable samples). So, each training subset had 200 samples (100 stable samples, 100 unstable samples) and testing subset had 40 samples (20 stable samples, 20 unstable samples).

3.2 Results of feature subset selection

In this paper, four search algorithms that are SFS, SBS, SFFS and FR, were proposed applying to feature subset selection. In which, the SFS, SBS, SFFS algorithms had been applied in [2]. The objective function (13) was applied for these three algorithms in this study. FR algorithm had been applied in [2],[4] with Fisher distance function (15). Figure 1 shows the results of distance measuring value by SFFS, SFS and SBS algorithm. Figure 2 shows the results ranked from large to small according to Fisher's distance measuring the value of each single feature.

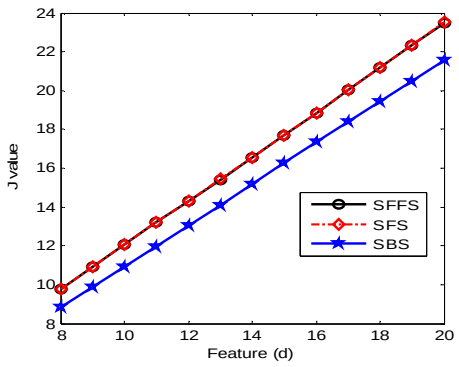

Figure 1 distance calculated value of SFFS, SFS and SBS algorithm.

Table 1. The measured distance (J value) of SFFS and SFS algorithm of feature subsets with d=13 and d=20.

Feature (d)	J value (SFFS)	J value (SFS)
13	15.42795	15.42502
20	23.54927	23.48542

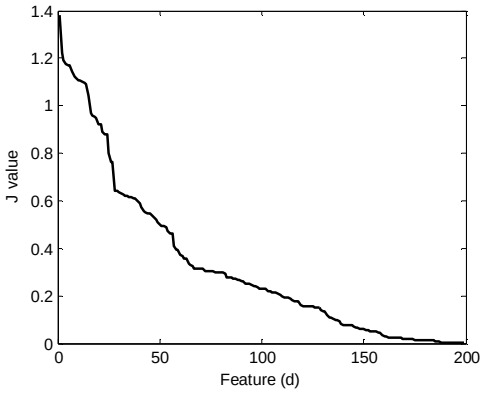

Figure 2 Fisher's distance measuring value.

Table 2. Calculating time of SFS, SFFS, SBS algorithm with d=20 and FR with d=199.

	SFS	SFFS	SBS	FR
Time (s)	1.15	2.58	117.5	0.14

3.3 Results of training

MLNF had three layers: one input layer, one hidden layer and one output layer. Hidden layer has 10 neurals with activate function tansig. Activate function purelin was used for output layer. Levenberg-Marquardt optimization based for weight and bias

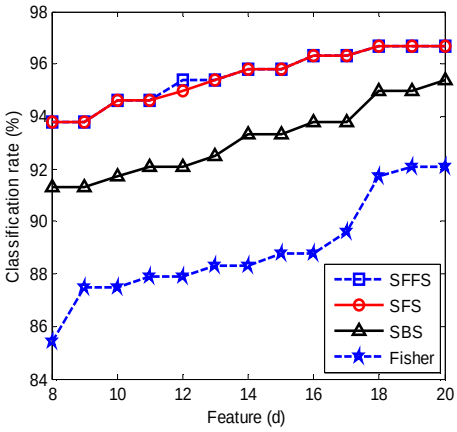

Figure 3 clasification rate of testing feature subsets by MLNF.

updating algorithm was selected. These functions are supported in neural networks tool of R2011b Matlab software. Programs were performed by laptop with CPU Inter Core™ i3-380M, 2GB DDR3 Memory, 500GB HDD. Figure 3 shows classification rate of testing feature subsets with algorithms by MLNF.

Table 3 Training time and testing classification rate of algorithms with d=12 and d=199.

	feature (d)	Training time (s)	r(%)
SFS	12	2.03	95.0
SBS	12	2.18	92.1
Fisher	12	2.16	87.9
Total	199	7.61	95.8

From Table 3, we can observe that SFS algorithm got higher classification rate than others. So, Suggested method-based SFS algorithm applied to select 12 top of features (ΔVbusWeber69, ΔVbusBLT138, ΔPloadSHIMKO69, ΔPloadHALE69, ΔQloadPATTEN69, ΔQloadAMANDA69, ΔQloadBLT138, ΔPloadBLT69,

ΔQloadLYNN138, ΔPflowRAY138-BOB138, ΔPflowBOB138-BLT138, ΔPflowBLT69-BLT138) in order to reduce the number of inputs to MLFN. However, we expected to check whether SFS is biased to any different classifiers. Linear Discrimination Analysis was used as the classification algorithm to our testing it. The linear classifier (LC) is one of the simplest discrimination analysis types. This classify function is also supported by Matlab software. Figure 4 shows classification rate of testing feature subsets with algorithms by LC.

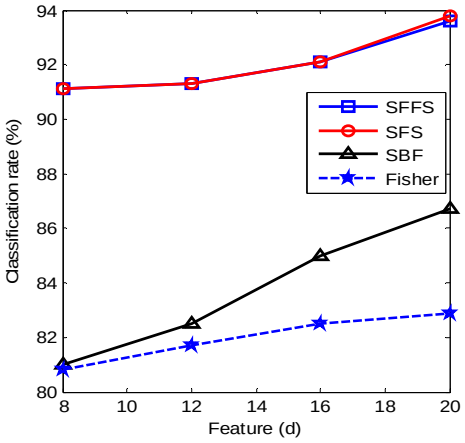

Figure 4 Classification rate of testing feature subsets by LC

3.4 Discussion

Figure 1 shows the results of distance measuring value by SFFS, SFS and SBS algorithm. Figure 2 shows the results ranked from large to small according to Fisher's distance measuring value of each single feature by FR algorithm. In which, the same distance measuring values were caculated by SFS and SFFS, but that have very small value difference at subsets with 13 features and 20 features as Table 1.

According to Table 2, with 20 features, it can see that calculating time of FR algorithm is the shortest time with 0.14s. Calculating time of SBS algorithm is the longest time with 117.5s. Calculating time of SFFS algorithm is 2.58s and longer 2,2 times than calculating time of SFS algorithm. Calculating time of SFFS algorithm is 1.15s. Calculating time of SBS algorithm is much longer calculating time of SFS, SFFS and FR algorithm. This can explain that the SBS algorithm has to through the space search with the entire feature set. SFFS algorithm has longer calculating time than SFS's time calculating because beside of forward search, SFFS algorithm has to backward search. The shortest calculating time of FR algorithm has a reason that FR algorithm calculated measuring distance values only one time respectively for each feature.

Figure 3, classification rates of SFS and SFFS algorithm are the same. SFFS and SFS algorithm give better results than SBS and FR algorithm. Classification rates of SFS and SFFS algorithm are more 1,3% to 2,9% than SBS algorithm and more 4,6% to 8,3% than FR algorithm.

According to Table 3, SFS algorithm, subset has 12 features that its classification rate got 95% by MLFN. Comparing with feature set has 199 features, SFS algorithm's feature number was reduced 16,5 times, its training time was reduced 3,74 times. Classification rate of that feature set has 199 features is 95,8%. By comparing the calculated results found that SFS algorithm has the same results as SFFS algorithm. These results can be explained that in step backward search SFFS algorithm only removes one feature for each execution algorithm could not search deep enough to find

better features. SFS algorithm is simpler than SFFS algorithm.

Classification rates of SFS and SFFS algorithm are also the same and got better results than SBS and FR algorithm by LC. MLFN got higher classification rate than LC for the same feature subset selection algorithm. The SFS has the 12 selected features that its classification rate got 95% by MLFN. This result was also considered acceptable for some previous studies applying pattern recognition to power system stability. For instance, classification rate got 95% [11], 93,6% [12].

4. CONCLUSION

This paper presents the method of feature subset selection in dynamic stability assessment power system using artificial neural networks. This paper proposed applying four feature subset selection algorithms that are FR, SFS, SBS, and SFFS. The effectiveness of the algorithms was tested on the GSO-37bus power system. With the same number of feature, the calculation results show that SFS algorithm yielded higher classification rate than FR, SBS algorithm. SFS algorithm yielded the same classification rate as SFFS algorithm.

Lựa chọn tập biến trong đánh giá ổn định động hệ thống điện sử dụng mạng thần kinh nhân tạo

- **Nguyễn Ngọc Âu** [1]
- **Quyền Huy Ánh** [1]
- **Phan Thị Thanh Bình** [2]

[1]Trường Đại học Sư Phạm Kỹ Thuật Thành Phố Hồ Chí Minh

[2]Trường Đại học Bách Khoa, ĐHQG-HCM

TÓM TẮT

Bài báo trình bày phương pháp lựa chọn tập biến trong đánh giá ổn định động (DSA) hệ thống điện sử dụng mạng thần kinh nhân tạo (ANN). Trong ứng dụng ANN đánh giá ổn định động hệ thống điện, lựa chọn tập biến nhằm mục đích giảm số biến đầu vào, giảm chi phí và bộ nhớ máy tính . Tuy nhiên, thách thức lớn là cùng với việc giảm số lượng biến nhưng độ chính xác nhận dạng phải cao. Bài báo này đề nghị áp dụng các giải thuật tìm kiếm tiến (SFS), tìm kiếm lùi (SBS), tìm kiếm kết hợp tiến lùi (SFFS), xếp hạng (FR) để lựa chọn tập biến. Hiệu quả của các giải thuật đã được kiểm tra với sơ đồ hệ thống điện GSO-37bus. Kết quả tính toán cho thấy với cùng biến đặc trưng (Feature), giải thuật SFS có độ chính xác nhận dạng cao hơn giải thuật FR và SBS, giải thuật SFS và SFFS có kết quả tính toán như nhau.

Từ khóa: *lựa chọn tập biến, đánh giá ổn định động, mạng thần kinh nhân tạo, hệ thống điện.*

REFERENCES

[1]. Prabha Kundur, *'Power System Stability and Control'*, McGraw-Hill Inc, 1994.

[2]. Yan Xu, Zhao Yang Dong, JunHua Zhao, Pei Zhang, Kit Po Wong,*'A Reliable Intelligent System for Real-Time Dynamic Security Assessment of Power Systems'*, IEEE Transactions On Power Systems, Vol. 27, No. 3, p.1253-1263, August 2012.

[3]. Nima Amjady and Seyed Farough Majedi, *'Transient Stability Prediction by a Hybrid Intelligent System',* IEEE Transactions On Power Systems, Vol. 22, No. 3, p.1275-1283, August 2007.

[4]. K. Shanti Swarup, *'Artificial neural network using pattern recognition for security assessment and analysis'*, Neurocomputing 71, 983–998, Elsevier, 2008.

[5]. J. Duncan Glover, Mulukutla S.Sarma, Thomas J.Overbye,' *Power System Analysis and Design'*, Fifth Edition, Publisher Global Engineering: Christopher M. Shortt, 2012.

[6]. Quyen Huy Anh, *'The applycation of pattern recognition for fast analysis of the dynamic stability of electrical power system'*, Electrical technology, No.2 pp 1-13, Perganon, 1994.

[7]. Kwang Y. Lee and Mohamed A. El-Sharkawi, *'Modern Heuristic Optimization Techniques, Theory and Applications To Power Systems'*, John Wiley & Sons, Inc Publication, 2008.

[8]. Andrew R.Webb, Keith D.Copsey, *'Statistical Pattern Recognition'*, Third Edition, A John Wiley & Sons, Ltd., Publication, 2011.

[9]. Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, Ching Y. Suen,' *'Character Recognition systems: A Guide for Students and Practioners''* A John Wiley & Sons, Ltd., Publication, 2007.

[10]. R. Zhang, S. Member, Y. Xu, and Z. Y. Dong, *"Feature Selection For Intelligent Stability Assessment of Power Systems,"* 2012 IEEE Power Energy Soc. Gen. Meet., pp. 1–7, 2012.

[11]. A. M. a. Haidar, M. W. Mustafa, F. a. F. Ibrahim, and I. a. Ahmed, *"Transient stability evaluation of electrical power system using generalized regression neural networks," Appl. Soft Comput.*, vol. 11, no. 4, pp. 3558–3570, 2011.

[12]. A. M. El-Arabaty, H. a. Talaat, M. M. Mansour, and a. Y. Abd-Elaziz, *"Out-of-step detection based on pattern recognition," Int. J. Electr. Power Energy Syst.*, vol. 16, no. 4, pp. 269–275, 1994.