ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

Pattern Classification Using Neuro Fuzzy and Support Vector Machine (SVM) - A Comparative Study

Dr. Maya Nayak[1] and Er. Jnana Ranjan Tripathy[2]

Department of Information Technology, Biju Pattnaik University of Technology, Orissa Engineering College

Bhubaneswar, Odisha-752050, India[1]

Department of Computer Science & Engineering, Biju Pattnaik University of Technology, Orissa Engineering College

Bhubaneswar, Odisha-752050, India[2]

Abstract : In this paper, we present a comparative study on applications of Neuro Fuzzy and Support Vector Machines (SVMs) for pattern recognition. Since SVMs show good generalization performance on many real-life data and the approach is properly motivated theoretically, it has been applied to wide range of applications. This paper describes a brief introduction of SVMs and summarizes its numerous applications and comparative study of SVM and Neuro Fuzzy in pattern recognition.

Key Words: - Neuro Fuzzy, SVM, Pattern Recognition, Transform.

1.1 INTRODUCTION

1.1.1 Neuro Fuzzy

The pattern classification and its subsequent application to temporal data mining proposed in this chapter for knowledge discovery in time series data consists of two stages. As the data comprises time varying voltage signal samples (non-stationary in nature), the first stage will process the data by using a multi resolution transform (S-transform) similar to the wavelet transform to extract the relevant features. The extracted features will then be used in the second stage for classification of the no stationary time series patterns by a fuzzy neural classifier in the form of If-Then rules.

1.1.2 Support Vector Machine (SVM)

SVMs are a new type of pattern classifier

Unlike traditional methods (e.g. Neural Networks), which minimize the empirical training error, SVMs aim at minimizing an upper bound of the generalization error through maximizing the margin between the separating hyper plane and the data. Since SVMs are known to generalize well even in high dimensional spaces under small training sample conditions and have shown to be superior to traditional empirical risk minimization principle employed by most of neural networks, SVMs have been successfully applied to a number of applications ranging from face detection, verification, and recognition, object detection and recognition handwritten character and digit recognition, text detection and categorization, speech and speaker verification, recognition information and image retrieval.

1.2 Pattern Classification and Knowledge Discovery

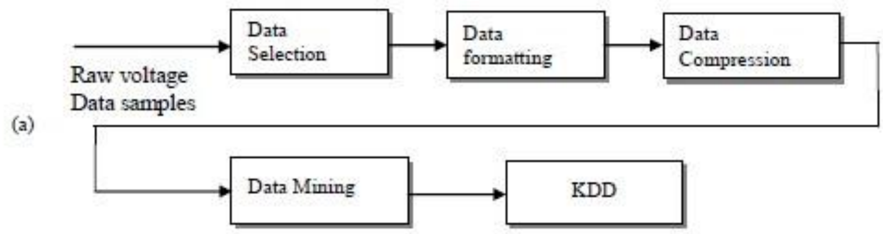
The overall KDD (knowledge discovery from data) is shown in Fig.1.1, where raw data from the recording instruments at consumer sites go through a data selection process in which spurious data and outliers are rejected. The selected data is then formatted and normalized and if necessary are down sampled to give a smaller data window to be used for pattern classification and temporal data mining applications.

The pattern classification task comprises the following steps:

a. Feature extraction using the S-transform
b. Fuzzification of input features and use of a fuzzy classifier for power network disturbance classification.
c. Fuzzy MLP (Multi-layered Perceptron) for rule generation.
d. Recognition of non-stationary time series disturbance patterns.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013



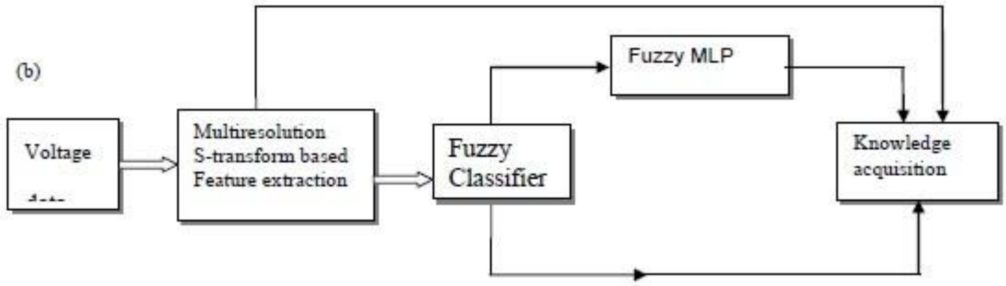


Fig. 1.1(a) Overall KDD block diagram, (b) Data mining block

1.3 Feature Extraction using the Multiresolution S-transform.

The Multiresolution S-transform originates from two-advanced signal processing tools; the Short-time Fourier transform (STFT) and the Wavelet transform. It can be viewed as a frequency dependent STFT or a phase corrected wavelet transform. Due to the frequency dependent window used for analysis of a signal data, the Multiresolution S-transform has been proven in Chapter-4 to perform better than other time-frequency transforms. Furthermore, it provides superior time frequency localization property computing both amplitude and phase spectrum of discrete data samples. The following four features are chosen for pattern classification using fuzzy neural network and fuzzy expert system. The optional choice of features can, however, be undertaken using either GA (Genetic Algorithm) or PSO (Particle Swarm Optimization). The signal peak value is normalized to 1.0. The features for classification of time series data are chosen as

1. F1= max (A) +min (A)-max (B). Where A is the amplitude versus time graph from the S-matrix under disturbance and B is the amplitude versus time graph of the S-matrix without disturbance.
2. F2 = Standard deviation of contour No.1 having the largest frequency magnitude versus time.
3. F9 =Standard deviation of Cr. Where Cr is the amplitude versus time graph of the S-matrix for frequencies above twice the normalized fundamental frequency.
4. F10 =Average power for frequencies above 2.5 times the fundamental frequency.

These features are found to be well suited to distinguish the six classes of power signal time series disturbance events studied here.

1.4 Feature Extraction Using Simulated and Recorded data

The proposed approach is applied to detect, localize and classify time series signal patterns in electrical power networks. The signal patterns considered for this application include:

i) Decrease in signal magnitude (sag):Class 1
ii) Increase in signal magnitude (swell): Class 2
iii) Complete collapse of the signal amplitude (interruption): Class 3
iv) Oscillatory transients: Class 4
v) Spikes: Class 5
vi) Notches: Class 6

To illustrate the relative values of these features, several typical power signal disturbance events like voltage sag, swell, interruption, harmonics, transients, spikes, and notches, etc. are simulated using the MATLAB code. A random white noise of zero mean and signal to noise ratio (SNR) varying from 50dB to 20dB is added to the signals. A typical SNR value of 30dB is equivalent to a peak noise magnitude of nearly 3.5% of the voltage signal. A sampling frequency of 1.6 kHz is chosen for computation of S-Transform for the simulated waveforms. As the S-Transform software uses analytic signals, only positive frequency spectrum is evaluated and FFT of the signal

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

samples is automatically computed as a part of the software. The software developed by authors was an extension to the original version of the S-transform by Stockwell. Laboratory recorded data is also used for feature extraction & time series classification.

1.5 Fuzzy Logic Based Classifier

After the features are extracted using Multiresolution S-transform, the features are passed to a fuzzy rule based system. The fuzzy logic based classifier comprises the fuzzy if-then rules of the following form, for classification of normal (class-0) and class 1, class 2, and class 3, patterns: Rules R_q: If F_1 is A_q and F_2 is B_r Then class C_m with a confidence value CF_m Where q =1, 2, 3, 4: r =1, 2, 3 and m =0, 1, 2, 3, t indicating that m = 0, the pattern class is normal, and m = 1, 2, 3, to denote interruption, sag, swell and transients, respectively.

1.6 Rule Generation

The trained fuzzy neural network is used for rule generation in If-Then rules form in order to justify any decision reached in the process of evaluation. Fig.1.2 presents the block diagram for the rule generation. These rules describe the extent to which a test pattern belongs or does not belong to one of the classes of the power signal time series disturbance patterns in terms of antecedent and consequent parts of the rule. The following steps are followed to generate the rules for time series knowledge discovery:

Step 1: Select hidden layer neurons that have a positive impact on the output neuron j. This requires (j = output layer neurons, k=hidden layer neurons).

Step 2: Select input neuron whose firing strength and whose maximum cumulative weight links to the output neuron through the selected hidden layer neurons are greater than zero. 5 .0

Step 3: If-then rules are generated from the selected input neurons and the output neurons. Step 4: Step 1 to 4 is repeated for all the training vectors.

Step 5: Discard all rules that have less than 3 features in the antecedent part. The consequent parts of the rules are obtained by the certainty measure P_j defined as:

$$\mathbf{P_j = Y_j - \sum_{I \to j} Y_t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.1)}$$

where yj is the output of neuron j. A linguistic output form is used to denote the consequent part as:

a. very likely for $0.8 <= P_j <= 1$
b. likely for $0.6 <= P_j < 0.8$ (1.2)
c. more or less likely for $0.4 <= P_j < 0.6$
d. unable to recognize for $P_j < 0.1$

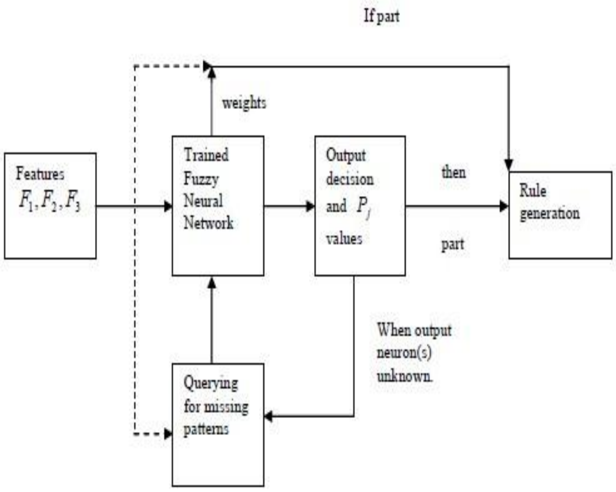


Fig.1.2 Rule generation block diagram

The efficacy of the rules generated from the fuzzy multi-layered perceptron network is then tested with Mamdani fuzzy inference system using max-min implication and centroid defuzzification. The output of the fuzzy inference system is shown in Tables-1. The average score in case of fuzzy classifier is

94.75%

1.7 Pattern Classification Using SVM

Classical learning approaches are designed to minimize error on the training dataset and it is called the Empirical Risk Minimization (ERM). Those learning methods follow the ERM principle and neural networks are the most common example of ERM. On the other hand, the SVMs are based on the Structural Risk Minimization (SRM) principle rooted in the statistical learning theory. It gives better generalization abilities (i.e. performances on unseen test data) and SRM is achieved through a minimization of the upper bound (i.e. sum of

the training error rate and a term that depends on VC dimension) of the generalization error.

1.8 Linear Support Vector Machines for Linearly Separable Case

The basic idea of the SVMs is to construct a hyper plane as the decision plane, which separates the positive (+1) and negative (-1) classes with the largest margin, which is related to minimizing the VC dimension of SVM. In a binary classification problem where feature extraction is

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

initially performed, let us label the training data d $x_i \in R^d$ with a label $y_I \in \{-1,+1\}$, for all the training data i = 1,..., l , where l is the number of data, and d is the dimension of the problem. When the two classes are linearly separable in R^d, we wish to find a separating hyper plane which gives the smallest generalization error among the infinite number of possible hyper planes. Such an optimal hyper plane is the one with the maximum margin of separation between the two classes, where the margin is the sum of the distances from the hyper plane to the closest data points of each of the two classes. These closest data points are called Support Vectors (SVs).

1.9 Example 1.
Face Detection and Recognition Using SVM

Face detection, verification and recognition are one of the popular issues in biometrics, identity authentication, access control, video surveillance and human-computer interfaces. There are many active researches in this area for all these applications use different methodologies. However, it is very difficult to achieve a reliable performance. The reasons are due to the difficulty of distinguishing different persons who have approximately the same facial configuration and wide variations in the appearance of a particular face. These variations are because of changes in pose, illumination, facial makeup and facial expression. Also glasses or a moustache makes difficult to detect and recognize faces. Recently many researchers applied SVMs to face detection, facial feature detection, face verification, recognition and face expression recognition and compared their results with other methods. Each method used different input features, different databases, and different kernels to SVMs classifier.

Example 2.
Object Detection and Recognition

Object detection or recognition aims to find and track moving people or traffic situation for surveillance or traffic control. This paper used bottom- up and top-down multi-class SVMs and the two types of SVM classifiers showed very similar performance. 3D object recognition was developed in and. Both of them used COIL object database, which contained 7200 images of 100 objects with 72 different views per each object.

1.10 Limitations of Using Neuro Fuzzy

Pattern recognition of time series data and subsequent temporal data mining of power signal disturbance events that occur frequently in power distribution networks. The time series pattern recognition procedure uses a Multiresolution transform similar to the wavelet transform used in signal and image processing fields for feature extraction and a hybrid fuzzy rule based system and fuzzy multi-layered perceptron network for rule generation and pattern classification. The wavelet Multiresolution analysis can also be used for data compression of a large volume of power signal time series data collected from the customer sites of the power distribution network. Several power signal time series disturbance events are simulated and are combined with the Multiresolution S-transform and fuzzy rule based system to identify the patterns belonging to power frequency steady disturbances and low and high frequency transients in the first stage of classification. Field recorded data along with the simulated data are used for computing the recognition scores. A fuzzy neural network approach is used in the second stage only to classify the low and high frequency transients and notches, etc. The classification accuracy is up to 94.75% in the presence of noise in the data.

1.11 Limitations of Using SVM

The performance of SVMs largely depends on the choice of kernels. SVMs have only one user-specified parameter C, which controls the error penalty when the kernel is fixed, but the choice of kernel functions, which are well suited to the specific problem is very difficult. The issue of how to control the selection of SVs is another difficult problem, particularly when the patterns to be classified are non-separable and the training data are noisy. In general, attempts to remove known errors from the data before training or to remove them from the expansion after training will not give the same optimal hyper plane because the errors are needed for penalizing nonseparability.

1.12 Comparison

By taking into account and after experimenting the pattern classification using both Neuro Fuzzy and SVM concepts the following results are carried out:-

1. Neuro Fuzzy is a complex classifier as compare to SVM. The pattern classification is achieved in less time using SVM than Neuro Fuzzy.
2. The accuracy is far better in SVM as compare to Neuro Fuzzy.
3. The experimental results in Table 1 shows accuracy of 94.75% in case of Neuro Fuzzy.
4. Whereas in case of SVM its 97.75% as shown in Table 3.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

1.13 Experimental Results using Neuro Fuzzy

Table -1: Classification results obtained from field test Using Fuzzy Classifiers

Type of time series data	Number of disturbances	Number of class correctly identified
Disturbance Type		
Sag	100	92
Swell	100	96
Interruption	100	97
Transients	100	94
Average	100%	94.75%

1.14 Experimental Results using Support Vector Machine (SVM)

Table -3: Classification results obtained from field test Using SVM Classifiers

Type of time series data	Number of disturbances	Number of class correctly identified
Disturbance Type		
Sag	100	97
Swell	100	98
Interruption	100	99
Transients	100	97
Average	100%	97.75%

CONCLUSION

The paper presents pattern recognition of time series data and subsequent temporal data mining of power signal disturbance events that occur frequently in power distribution networks. The time series pattern recognition procedure uses a multiresolution transform similar to the wavelet transform used in signal and image processing fields for feature extraction and a hybrid fuzzy rule based system and fuzzy multi-layered perceptron network for rule generation and pattern classification.

We have also presented a brief introduction on SVMs and several applications of SVMs in pattern recognition problems. SVMs have been successfully applied to a number of applications ranging from face detection and recognition, object detection and recognition, handwritten character and digit recognition, speaker and speech recognition, information and image retrieval, prediction and etc. because they have yielded excellent generalization performance on many statistical problems without any prior knowledge and when the dimension of input space is very high. The comparison also made using both the techniques over same data set and conditions. As shown in Table 1 and Table 3, it is clearly seen that Pattern Classification using SVM holds a healthy edge over Neuro Fuzzy.

REFERENCES

1. C.H. Lee, J.S. Lee, J.O. Kim, and S.W. Nam, ―Feature Vector Extraction for the Automatic Classification of Power Quality Disturbances‖,

2. 1997 IEEE International Symposium on Circuits and Systems, pp.2681-2685, 1997.

3. A. Elmitwally, S. Farghal, M.Kandil, S.Abdelkader and M. Elkateb, ―Proposed Wavelet-Neurofuzzy Combined System for Power Quality Violations Detection and Diagnosis, IEEEProc-Gener. Transm. Distrib, vol.148, no.1, pp.5-20, 2001.

4. Wei - Yen Wang, Chih - Yuan Cheng, Yih – Guang leu, An online GA based output feedback Direct adaptive Fuzzy neural controller for uncertain nonlinear systems, IEEE transactions on Syst. Man and Cybernetics Part B .vol.34, no.1, pp.334345.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

5. B. Boser, I. Guyon, and V. Vapnik, A training algorithm for optimal margin classifiers, In Proceedings of Fifth Annual Workshop on Computational Learning Theory, New York, (1992).

6. C. Cortes and V. Vapnik, Support vector networks, In Proceedings of Machine Learning, vol. 20, pp. 273-297, (1995).

7. V. Vapnik, The nature of statistical learning theory, Springer, (1995).

BIOGRAPHIES

Prof (Dr) Maya Nayak
Completed PhD (CS), BE (CS) Currently Working as HOD, Dept. of IT at Orissa Engineering College, Odisha Published Many Research Papers in Popular International Journals.

Er. Jnana Ranjan Tripathy
Pursuing PhD in Centurion University of Technology & Management in "ANN Implementation in Embedded Systems" M.Tech in Computer Science, Berhampur University B.Tech in Information Technology, BPUT Currently working in Orissa Engineering College, Odisha, Worked at Centurion University previously.