

Classifiers Revisited: From Signal Anomaly and Artifacts Detection Perspective

Khushali V. Vaghani
Assistant Professor
Shree Swami Atmanand
Sarswati Institute of
Technology
Surat-395006
Gujarat

Amit P. Ganatra, PhD
Dean
Faculty of Technology and
Engineering
Charusat
Changa-388421
Gujarat

Hiren Rambhia
Assistant Professor
Mumbai University
Maharashtra

ABSTRACT

In recent years, the research communities have focused on using training based classifiers as a tool for signal anomaly/artifacts detection. The efforts in this direction have lead to vast literature and development of classifiers each with its own advantages and disadvantages. This paper provides a comprehensive view on widely used statistical and neural network based classifier. Specifically, Naive Bayes as statistical classifier, Radial Basis Neural Network (RBNN) and Back Propagation Neural Network (BPNN) as neural network classifier are discussed here. For the purpose of comparison, a case study involving signals from multi-spectral line scanner based space camera obtained during on-ground characterization of misregistration among bands is considered.

General Terms

Classification, Artifacts Removal, Neural Networks

Keywords

Neural Networks, Classification, Back Propagation Neural Network, Radial Basis Function Neural Network, Naïve Bayes Classifier

1. INTRODUCTION

Signals are ubiquitous as they are used to represent a wide variety of physical phenomena. Signal can be a biomedical signal such as EEG, a radar signal, sensor output or even a line scanner type camera signal. Very often these signals are processed for anomaly detection or for removing the artifacts before further processing or information extraction. Both, anomalies and artifacts in a signal, are distinguishable and identifiable as their characteristic features differ from the norm.

The fact that anomaly and artifacts differ from the normal makes it a classical problem of classification. Approaching the problem from classification perspective opens up a whole world of well understood and easily implementable classifiers. Using training data, classifiers can be trained sufficiently for intelligent processing of signals in a supervised manner. The use of classifiers has been widely reported in papers and is found to be an effective and accurate solution.

Jon A. Benediktsson [1] discussed about classification of multisource remote sensing data. He empirically tested neural network approach as well as statistical methods for classification. Zhang [10] discussed how neural networks are useful for classification of the data set. Cirega[2] proved that hoe DNN is useful for image classification compare to other approaches. Deng [3] discussed sequence classification using

DNN. For high resolution remote sensing images' classification Back Propagation Neural Network is more suitable. It was proved by Jiang [6]. Kuruvilla[7] proved how NNs are useful for lung cancer classification. Heermann[4] discussed about classification of multispectral remote sensing data with BPNN. P. Jeatrakul [5] compared different neural networks for binary classification problems. Şen [9] compared different classification algorithms for EEG signal classification. Mohanty [8] developed an Artificial Neural Network based software package for classification of remotely sensed data.

Naïve Bayes classifier is one of the simplest classifier which is based on Bayes theorem. It is the probabilistic classifier with strong independence assumptions between the features. It is discussed in brief in the paper. We did experiment with this classifier and got some results. Neural Networks are one of the adaptive approaches from the machine learning techniques. They are data driven in nature and can adjust themselves to data without any explicit specification of functional or distribution form for the underlying model and can handle noisy data with higher robustness. They can solve nonlinear problems with high accuracy. They are flexible in modeling real world problems.

In this paper we discuss about neural networks in brief and present results after applying the same on our dataset and did comparison with Naïve Bayes classifier. In section 2 we discuss Naïve Bayes classifier, Back Propagation Neural Network and Radial Basis Function (RBF) Neural Network. In section 3 we discuss the results of the three classifiers and then we draw a fruit full conclusion.

2. METHOD

2.1 Naïve Bayes Classifier

Naïve Bayes classifier assumes class independence i.e. the presence of a particular parameter of a class is unrelated to the presence of any other parameter. Depending on the precise nature of the probability model, these classifiers can be modeled very effectively. One can work with the naive Bayes model without believing in Bayesian probability or using any Bayesian methods. Though naive Bayes design over-simplifies assumptions, these classifiers have worked very well in many real-world situations.

Naïve Bayes classifiers can handle number of variables whether they are continuous or categorical. Suppose we have a set of variables, $X = \{x_1, x_2, x_3, \dots, x_n\}$, possible outcomes are $C = \{c_1, c_2, c_3, \dots, c_n\}$ and we want to find posterior probability for C_j among the set C .

With the help of Bayes rule:

$$p(C_j | x_1, x_2, x_3 \dots x_n) \propto p(x_1, x_2, x_3 \dots x_n | C_j)$$

Here, $p(C_j | x_1, x_2, x_3 \dots x_n)$ is posterior probability of class membership, conditional probabilities of the independent variables are independent so we can write

$$p(X | C_j) \propto p(x | C_j)$$

And we can write the posterior as:

$$p(C_j | X) \propto p(C_j) \prod_{k=1}^n p(x_k | C_j)$$

By the above rule, we label new case X with a class level. Though this naïve Bayes classifier is the simple one but there are also some issues related to this. For instance, if we have no occurrence of a class label then the frequency based probability estimate will be zero. When all probabilities are multiplied then we get zero, which effects the posterior probability estimation. We applied this classifier to classify our dataset and result is also calculated in the next section.

2.2 Neural Networks

Neural Networks are information processing paradigms which are inspired by biological nervous system, such as human brain. The main important element about NN is the novel structure of the information processing system. The structure includes number of layers containing number of interconnecting processing elements, called neurons. Each interconnection between the neurons has synaptic weight. NN is learn by example in nature. NNs are configured for number of applications, such as Pattern Recognition, Data Classification, Image Compression, Prediction, Sales and Marketing, Medical Diagnosis, Space Science etc. through a learning process. This learning is done by the adjustment of the synaptic interconnected weights. With the major ability of the trained and expert NN to extract meaning from the complicated data is used to extract the patterns and detect the trends. In this experiment with classification we use two NNs, Back Propagation NN (BPNN) and Radial Basis Function NN (RBFNN).

2.2.1 Radial Basis Function Neural Network

The idea of RBFNN derives from the theory of function approximation. RBFNN is a type of multilayer network, consists three layers, input layer, hidden layer and output layer. Hidden neurons implement a set of radial basis functions. And output nodes implement linear summation function as in Multilayer Perceptron Networks. The training is done via calculations of input to hidden layer followed by hidden to output layer. According to those calculations the weights are updated. The trained network is going to use for classification. The training and learning of this RBFNN is very fast and this RBFNN is good at interpolation. Many functions are used as RBF, such as, Gaussian Functions, Multi-Quadric Functions, Generalized Multi-Quadric Functions, Thin Plate Spline Functions, Cubic Functions, and Linear Functions etc. In our experiment we took Gaussian Function as RBF. RBFNN combines the parametric statistical distribution model and non-parametric linear perceptron algorithm in serial sequence. It uses only one hidden layer to model any nonlinear function. So training is fast. We experiment with RBFNN and the results of classification are shown in result section.

2.2.2 Back Propagation Neural Network

BPNN is one of the simple and effective among all the NNs. It is feed forward in nature and more adaptive. It contains one input layer with input neurons, hidden layer with hidden neurons and one output layer with output neurons. The information is passed to the hidden layer via input neurons and their interconnected synaptic weights. Activation function is applied on the final weighted sum at hidden layer. Here we use sigmoid function as hidden layer activation function. Same procedure is done at hidden layer and final sum is passed at output layer. This whole procedure is known as forward pass in BPNN. According to the target vector, the error is calculated and according to this error the weights are modified to reach up to the target vector. This process is known as backward pass of BPNN because here we are passing errors back. This training is done in supervised manner with Gradient Descent Algorithm. After reaching the minimum network error, the network is trained and can use to classify the unknown patterns. To speed up the training of BPNN we can use the learning rate and momentum factor. But we have to fix the optimized network by setting all the parameters in an optimum fashion. The figure of our BPNN is shown below.

The squared error is defined by the below equation

$$E = \frac{1}{2} \sum_{j=1}^N (T_{pj} - O_{pj})^2$$

Here E is the error for P^{th} presented vector; T_{pj} is the desired output for the jth output node and O_{pj} is the actual output. Main aim is to minimize the error. After the differentiation of the above error with respect to weight, we get the followed updating equation for the weights.

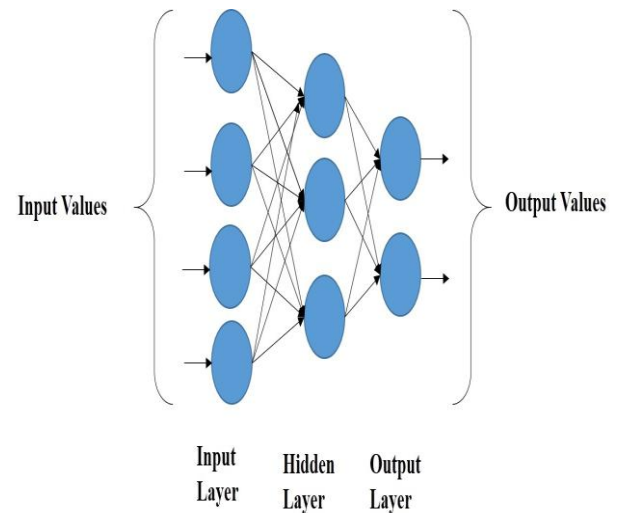


Fig 1: Back Propagation Neural Network

$$W_{pqk}^{(n)} = W_{pqk}^{(n-1)} + \eta * \delta_{qk} * OUT_{pj}$$

As η is the constant learning rate with the value $0 < \eta < 1$ and δ_{qk} is the error. We did experiment with BPNN and get the result shown in result table.

3. RESULT

For performance comparison, we took three sample signals from linear multispectral camera containing anomaly and applied classifiers for detection. The 1-D signal was optimally segmented for feature extraction. Each segment was then classified into either normal segment or abnormal one i.e. segment containing the anomaly. Here, features from each segment is represented as one sample in the classification stage. In first data set we have total 128 samples, in second data set we have total 103 samples and in third dataset we have 114 data samples. We did classification using Naïve Bayes Classifier, RBFNN and BPNN and the result is shown in the result table.

4. CONCLUSION

Table-1 presents performance parameters related to classification using three main classifiers; Naïve Bayes Classifier, BPNN and RBFNN. Performance parameters used for comparison are Correctly Classified samples (%), Incorrectly Classified Samples (%), Kappa Statistics, Mean Absolute Error, Root Mean Square Error and Precision for BPNN, RBFNN and Naïve Bayes Classifier. From the Chart-1 BPNN gives average 98.96 % correctly classified samples, for RBFNN gives 97.50 % correctly classified samples and with Naïve Bayes classifier gives 95.62% correctly classified samples. Thus, for the selected sample data set of linear multispectral camera signals, BPNN outperforms compared to other classifiers with an average of 98.96 % correctly classified samples.

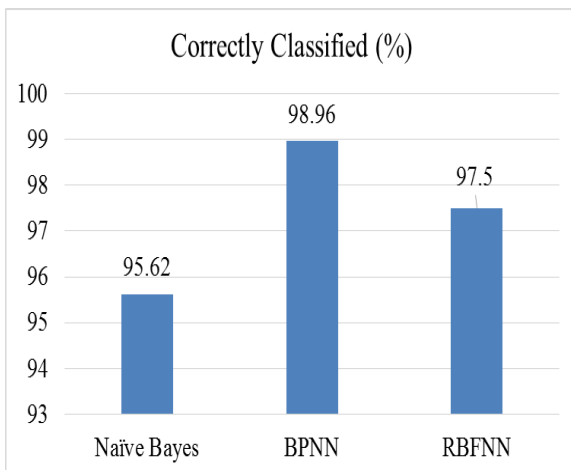


Chart 1: Correctly Classified (%) for Naïve Bayes, BPNN and RBFNN

5. REFERENCES

[1] Benediktsson, Jon A., Philip H. Swain, and Okan K. Ersoy. "Neural network approaches versus statistical methods in classification of multisource remote sensing data." (1990).

[2] Ciregan, Dan, Ueli Meier, and Jürgen Schmidhuber. "Multi-column deep neural networks for image classification." *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012.

[3] Deng, Li, and Jianshu Chen. "Sequence classification using the high-level features extracted from deep neural networks." *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014.

[4] Heermann, Philip D., and Nahid Khazenie. "Classification of multispectral remote sensing data using a back-propagationneural network." *Geoscience and Remote Sensing, IEEE Transactions on* 30, no. 1 (1992): 81-88.

[5] Jeatrakul, P., and K. W. Wong. "Comparing theperformance of different neural networks for binary classification problems." In *Natural Language Processing, 2009. SNLP'09. Eighth International Symposium on*, pp. 111-115. IEEE, 2009.

[6] Jiang, Jiefeng, Jing Zhang, Gege Yang, Dapeng Zhang, and Lianjun Zhang. "Application of back propagation neural network in the classification of high resolution remote sensing image: Take remote sensing image of beijing for instance." In *Geoinformatics, 2010 18th International Conference on*, pp. 1-6. IEEE, 2010.

[7] Kuruvilla, Jinsa, and K. Gunavathi. "Lung cancer classification using neural networks for CT images." *Computer methods and programs in biomedicine* 113.1 (2014): 202-209.

[8] Mohanty, K. K., and T. J. Majumdar. "An artificial neural network (ANN) based software package for classification of remotely sensed data." *Computers & Geosciences* 22, no. 1 (1996): 81-87.

[9] Şen, Baha, et al. "A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms." *Journal of medical systems* 38.3 (2014): 1-21.

[10] Zhang, Guoqiang Peter. "Neural networks for classification: a survey." *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 30, no. 4 (2000): 451-462.

6. APPENDIX

Table 1: Performance Comparison of Classifiers in Terms of Correctly classified Samples (%), Incorrectly Classified Samples (%), Kappa Statistic, Mean Absolute Error, Root Mean Square Error, Precision (Average)

Method	Data Set	Total Number of Instances	Correctly Classified Samples (%)	Incorrectly Classified Samples (%)	Kappa Statistic	Mean Absolute Error	Root Mean Square Error	Precision (Average)
BPNN	Data set -1	128	96.88	3.13	0.94	0.03	0.17	0.97
	Data set -2	103	100	0	1	0.07	0.29	1
	Data set -3	114	100	0	1	0.012	0.054	1
RBFNN	Data set -1	128	96.88	3.13	0.94	0.03	0.17	0.97
	Data set -2	103	100	0	1	0.07	0.29	1

	Dat aset -3	114	95.6 1	4.37	0.9 1	0.0 5	0.1 5	0.96
Nai ve Bay es Clas sifie r	Dat aset -1	128	96.0 9	3.91	0.9 2	0.0 4	0.1 9	0.96
	Dat aset -2	103	95.1 5	4.85	0.9 0	0.0 48	0.2 14	0.95
	Dat aset -3	114	95.6 1	4.39	0.9 1	0.0 44	0.2 1	0.96