International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 09 | Dec-2015 www.irjet.net p-ISSN: 2395-0072

Comparative Analysis of Artificial Neural Network and Support Vector Machine Classification for Breast Cancer Detection

Prachi Damodhar Shahare[1], Ram Nivas Giri[2]
[1] *M.Tech Scholar, Department of CSE, Raipur Institute of Technology, Raipur, India*
[2] *Assistant Professor, Department of CSE, Raipur Institute of Technology, Raipur, India*

Abstract - ***Breast cancer is one of most hazardous types of cancer among women in the world. It is a complex disease characterized by many morphological, clinical and molecular features. It is an uncontrolled growth of breast cells. Soft Computing has become popular in developing systems that encloses human expertise. Imaging technologies and clinical cytology have improved breast cancer diagnosis, better survival rates and treatment by early detection of primary or metastatic masses, differentiating benign from malignant tumors and promoting intraoperative surgical guidance and post operative specimen evaluation. Medical results produce undesirable faults and extreme clinical costs which influence the value of services offered to patients. Hence, exact detection is extremely important for proper treatment and cure of disease. Precise outcome can be achieved using Artificial Neural Network. Different kernel functions of Support Vector Machine are used for classification of breast cancer dataset. The comparative analysis of both ANN and SVM is done for performance evaluation in terms of accuracy. The dataset used in this research is taken from UCI database composed of 683 cytological instances, out of which 458 are benign and 241 are malignant. The results show a better performance for SVM classifier than ANN.***

Key Words: *Keywords—Breast Cancer, Soft Computing, Artificial Neural Network, Support Vector Machine, Kernel Function, Classification.*

1. INTRODUCTION

Cancer refers to the uncontrolled multiplication of a group of cells in particular location of the body. A group of rapidly dividing cells may form a lump, microcalcifications or architectural distortions which are usually referred to as tumors.

Breast cancer is any form of malignant tumor which develops from breast cells. Breast cancer is one of most hazardous types of cancer among women in the world. The world health organization's International Agency for Research on cancer (IARC) estimates that more than 400,000 women expire each year with breast cancer [1].

Today, there is an urgent need in breast cancer control and it is achieved primarily by knowing different risk factors. Secondly, there is need to detect this disease in early stage by knowing different symptoms of this disease, so it can be cured.

Breast cancer is mainly of two types: Invasive and Non-invasive. Invasive type is the one in which cancerous cells break through normal breast tissue barriers and spread to other parts of the body. While in non-invasive, cancerous cells remain in a particular location of the breast and do not spread to surrounding tissue, ducts or lobules.

Breast analysis techniques have been improved over the last decade. Number of automated classification systems has been developed over last years. Different techniques have varying results. However, there still are issues to be solved: developing new and better techniques. The comparison between different systems helps us to know better system with high performance. This will assist radiologists to take accurate results regarding the disease.

Radiologists still produces some variation in reading images. So, there is a need for automatic interpretation of images or automated classification system, and for this purpose classifier is required. ANN, SVM and many more use for classification, but SVM outperforms all other techniques. However the performance of an SVM classifier is greatly dependent on the proper choice of a kernel function among other factors. This paper gives comparative analysis of ANN and different kernel functions of SVM.

In the recent years, soft computing (SC) techniques are applied to complicated and computationally comprehensive problems that have customarily proved incurable for conventional mathematical methods. It has become popular in developing systems that encloses human expertise [2]. The components of soft computing include artificial neural network, fuzzy computing, evolutionary computing, probabilistic reasoning and

support vector machine [2, 3]. These components help greatly to us in detection of breast cancer [1, 4].

2. LITERATURE REVIEW

Literature study shows different researches in breast cancer detection. All these researches provide varying results.

J.S.Leena Jasmine et. Al. [1] proposed an automated mass classification system in digital mammograms using Contourlet transform and support vector machine. They developed and analyzed Contourlet transform for feature extraction and support vector machine for classification process. Their classification system produces results with more than 81% for all the cases.

G. Anjan babu et. Al. [4] developed a system for breast cancer detection using Feed forward network with Back propagation algorithm. They applied ANN techniques on dataset for calculating accuracy, sensitivity and specificity for different networks. But they found that feed forward neural network gives best accuracy. Hence, they concluded that feed forward neural network give the good performance for detecting the breast cancer with back propagation algorithm.

Lisboa [5] developed the scheme to classify the masses by using artificial neural network where ANN provides better solutions. ANN is very useful tool in various medical diagnostic systems. The key attributes like distributed representations, local operations, and non-linear processing make ANN appropriate for taking few complex decisions from massive amount of data. Thus when expert knowledge is unavailable in full fledged sense as for example in case of masses, hence ANN provides alternative and better solutions.

Younes Kabbadji et. Al. [6] discovered Microcalcification Detection using a Fuzzy Inference System & Support Vector Machines. A Fuzzy inference system is applied to describe mathematically contrast and luminance which discriminates microcalcification from normal breast tissue. Furthermore, an SVM classifier trained on normal and positive cases with 9 features is used to further enhance the results of the previous phase. Very satisfactory results were obtained using the proposed method.

3. PROBLEM IDENTIFICATION

Many researches have been proposed number of techniques for processing of images and automated classification system giving varying results. But still there is need of more accuracy, recognition rate and minimum classification error rate. Hence, there is a need of more effective method for diagnosis of breast cancer.

4. METHODOLOGY

Proposed model consists of comparison analysis of ANN and SVM for breast cancer detection. It classifies the data classes to predict the class of entities or objects with unknown class label value.

We evaluate a classification techniques SVM and Artificial neural network for breast cancer detection problem. For SVM we use different kernels, and compare their relative accuracy. In case of ANN, we use Feed Forward network. The breast cancer classification model is shown in figure.

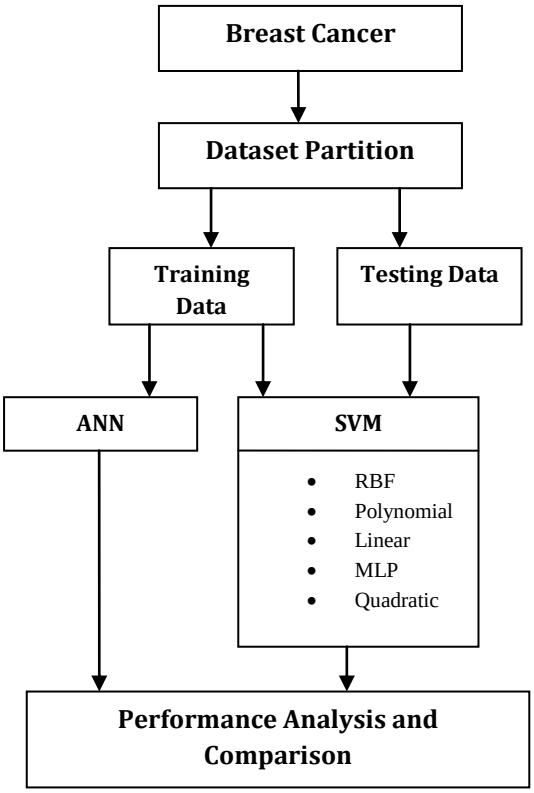

Fig -1: Breast cancer classification model

4.1 Dataset

The Breast cancer record used in this research is taken from the University of California at Irvine (UCI) machine learning repository. This breast cancer databases was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg [7, 8, 9, 10].

This dataset contains 699 numbers of instances. The instances with missing attribute values are removed from the database, leaving 683 instances. The numbers of attributes are 10 plus the class attribute. Each instance has one of 2 possible classes as: benign and malignant. The class distribution is as Benign: 458 and Malignant: 241.

The dataset is partitioned into two classes as training data and testing data, which is further used for classification system.

Attributes of the dataset are with different features:

1) ***Sample code number:*** It is the id number of the record, used to identify the record.
2) ***Clump Thickness:*** Clump is the size of grouped cells with domain between 1-10.
3) ***Uniformity of Cell Size:*** It is sameness in size of cells with domain 1-10.
4) ***Uniformity of Cell Shape:*** It is sameness in shape of cells with domain 1-10.
5) ***Marginal Adhesion:*** It is sticking capacity of the cell with other cell. It is having domain 1-10.
6) ***Single Epithelial Cell Size:*** It is the size of any epithelial cell with domain 1-10.
7) ***Bare Nuclei:*** It is the nucleus which is not surrounded by cytoplasm. It is having domain value between 1-10.
8) ***Bland Chromatin:*** It shows the texture of the nucleus in a cell, with domain value 1-10.
9) ***Normal Nucleoli:*** It is different structures seen in the nucleus of the cell with domain 1-10.
10) ***Mitoses:*** Mitosis is the division of tumor cells. It is having domain 1-10.
11) ***Class:*** It is the final classification of cell into two classes as benign and malignant. The domain shows 2 for benign and 4 for malignant.

4.2 Artificial Neural Network

Artificial neural networks (ANNs) are a family of models inspired by biological neural networks like the central nervous systems of animals i.e. the brain and are used to estimate or approximate functions that can depend on a large number of inputs which are mostly unknown. Artificial neural networks are generally presented as systems "neurons" which are all interconnected and which exchange messages between different neurons. The connections have numeric weights that can be adapt based on experience, creating neural nets adaptive to inputs and capable of learning [11].

Current progress in the area of ANNs has prepared them attractive for analysing signals. The ANN node presents a multiplicity of feed forward networks that are generally called back-propagation networks. Back-propagation refers to the method for calculating the error gradient for a feed forward network; a simple application of the chain rule of elementary calculus [12].

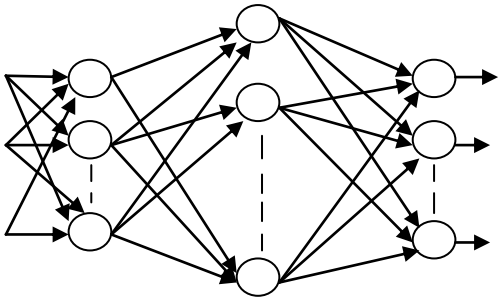

Fig -2: A Three-layered Artificial Neural Network

A feed forward network used for this research consists of an input layer with 9 neurons, a hidden layer with 121 neurons and an output layer with 2 neurons. Figure 3 shows architecture of feed forward neural network used.

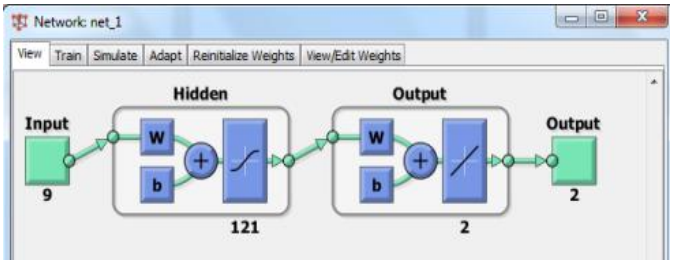

Fig -3: Architecture of Feed Forward Neural Network

4.3 Support Vector Machine

Support vector machines (SVMs), also called support vector networks are supervised learning models with related learning algorithms that evaluate data and recognize patterns which are used for classification and regression analysis [11]. For given a group of coaching examples SVM marked for one among 2 classes, Associate in Nursing SVM training formula builds a model that assigns new examples into one among the 2 class, creating it a non-probabilistic binary linear classifier. Associate in Nursing SVM model could be an illustration of the examples as points in house, mapped so the samples of the separate classes area unit divided by a transparent gap that's as wide as attainable. New examples area unit then mapped into that very same house and foretold to belong to a class supported that aspect of the gap they fall on.

The decision boundaries are driven directly from the training data so that the separating margins of decision boundaries are maximized in the high-dimensional space called feature space in training SVMs. This learning strategy is based on statistical learning theory which minimizes the errors in classification of the training data and the unknown data [11].

SVMs offer advantages over other types of classifiers. SVMs are free of the optimization headaches of neural networks because they present a convex programming problem. It guarantees for finding a global solution. These classifiers are much faster to evaluate than density estimators, as they make use of only the relevant data points, instead of looping over each point regardless of its relevance to the decision boundary [11, 12].

The Support Vector Machine can be viewed as a kernel machine. A kernel is a similarity function. It is a function that provide to a machine learning algorithm. This function takes two inputs and spits out how similar they are.

There are different kernels of SVM:

1) Radial Basis Function

In machine learning, the radial basis function kernel (RBF) is a well-known kernel function used in various learning algorithms which are kernelized. In specific, it is commonly used in support vector machine classification [13].

For two samples x and x', RBF kernel function are represented as feature vectors in some input space, is defined as

$$K(x, x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \quad (1)$$

2) Polynomial Function

The polynomial kernel looks at the given features of input samples to determine their similarity, also combinations of these. Such combinations are known as interaction features, in the framework of regression analysis. The implicit feature space of a polynomial kernel is similar to that of polynomial regression, but without the combinatorial altercation in the number of learned parameters. When the input features are binary-valued or Boolean-valued, then the features correspond to logical combinations of input features [13].

For degree-d polynomials, the polynomial kernel is defined as

$$K(x, y) = (x^T y + c)^d \quad (2)$$

3) Linear kernel

The linear kernel is good when there are a lot of features. Because of mapping the data to a higher dimensional space does not really improve the performance. In text classification, both the numbers of instances i.e. document and features i.e. words are large. Most of text classification problems are linearly separable.

4) Multilayer Perceptron

The MLP network learning process is based on the data samples of the N-dimensional input vector x and the M-dimensional output vector d, called destination. By processing the input vector x, the MLP produces the output signal vector y(x, w), where w shows the vector of adapted weights. The produced error signals activate a control mechanism of the learning algorithm. The adjustments are designed to make the output signal y_k (k = 1, 2… M) to the desired response d_k in a step by step manner [14].

The MLP learning algorithm is based on the minimization of the error function stated on the learning set (xi, di) for values, i = 1, 2… N using the Euclidean norm:

$$E(w) = \frac{1}{2}\sum_{i=1}^{N}\|y(x_i, w) - d_i\|^2 \quad (3)$$

5) Quadratic Kernel

The Rational Quadratic kernel is less computationally intensive than the Gaussian kernel and can be used as an alternative when using the Gaussian becomes too expensive [15].

$$k(x, y) = 1 - \frac{\|x-y\|^2}{\|x-y\|^2+c} \quad (4)$$

5. RESULT

In this paper we have evaluated the performance difference between a classifiers build by means of ANN and SVM, when come across the selected data. The accuracy of ANN with feed forward is recognized also in SVM classifier; accuracy of different kernel function is recognized. UCI Breast cancer dataset is used for this comparative analysis. Classification results obtained using SVM classifier with different kernel functions are shown in table.

The accuracy of ANN with feed forward network is 96.15% while the accuracy rate for different kernel functions is obtained as linear 99%, quadratic 96%, polynomial 95%, RBF 98.5% and MLP 98.5%.
The performance of ANN for classification is less than that with three SVM kernel functions i.e. linear, RBF, MLP.

Table -1: Comparison of Kernel Functions

S.No.	Kernel Function	Accuracy (%)	
		For training sample	For testing sample
1	Linear	96.8944	99
2	Quadratic	98.5507	96
3	Polynomial	100	95
4	RBF	97.5155	98.5
5	MLP	94.824	98.5

6. CONCLUSION

We have presented the results of comparative study on SVM with different kernels for breast cancer detection. For comparison, we employed five kernels: RBF, polynomial, Quadratic, linear and MLP. This study indicates that SVM with linear, RBF, MLP kernel having better performance than ANN. It follows from the results that an SVM classifier based on a combination of kernels can further enhance its performance for breast cancer detection. We will further investigate other kernels. Also, investigation on how to combine different kernels for better detection results is our future work.

REFERENCES

[1] J.S.Leena Jasmine, Dr.S.Baskaran, Dr.A.Goverdhan, "An automated mass classification system in digital mammmograms using Contourlet Transform and Support vector machine," IJCA, Vol 31-No.9, October 2011.

[2] L. Magdalena, "What is Soft Computing? Revisiting Possible Answers", IJCIS, Vol.3, No. 2 (June, 2010), 148-159.

[3] Yasuhiko Dote, Seppo J. Ovaska, "Industrial Applications of Soft Computing: A Review", PROCEEDINGS OF THE IEEE, VOL. 89, NO. 9, SEPTEMBER 2001.

[4] G.Anjan Babu, Sreenu Naik Bhukya, R.Sendhil KumaR, " Feed Forward network with Back Propogation Algorithm for Detection of Breast Cancer" IEEE, 2013.

[5] P.J.G. Lisboa, "A review of evidence of health benefits from artificial neural networks in medical intervension", Neural Networks, Vol.20, No.12,pp,1275-1284 (2000).

[6] Younes Kabbadji, Fakhita Regragai, Mohammad Majid Himmi," Microcalcification Detection using a Fuzzy Inference System & Support Vector Machines" IEEE 978-1-4673-1520-3/12,2012.

[7] O. L. Mangasarian and W. H. Wolberg: "Cancer diagnosis via linear programming", SIAM News, Volume 23, Number 5, September 1990, pp 1 &18.

[8] William H. Wolberg and O.L. Mangasarian: "Multisurface method of pattern separation for medical diagnosis applied to breast cytology", Proceedings of the National Academy of Sciences, U.S.A., Volume 87, December 1990, pp 9193-9196.

[9] O. L. Mangasarian, R. Setiono, and W.H. Wolberg: "Pattern recognition via linear programming: Theory and application to medical diagnosis", in: "Large-scale numerical optimization", Thomas F. Coleman and Yuying Li, editors, SIAM Publications, Philadelphia 1990, pp 22-30.

[10] K. P. Bennett & O. L. Mangasarian: "Robust linear programming discrimination of two linearly inseparable sets", Optimization Methods and Software 1, 1992, 23-34 (Gordon & Breach Science Publishers).

[11] S N Sivanandam, S Sumathi and S N Deepa, "Introduction to Neural Networks",sixteenth edition, 2012, McGraw hill.

[12] Aprajita Sharma, Ram Nivas Giri, "Automatic Recognition of Parkinson's diseases via Artificial neural Network and Support vector machine", IJITEE, 2278-3075, Vol-4, Issue-3, Aug2014.

[13] S.V.G.Reddy, K.Thammi Reddy, V. Valli Kumari, Kamadi VSRP Varma, "An SVM Based Approach to Breast Cancer Classification using RBF and Polynomial Kernel Functions with Varying Arguments", IJCSIT, Vol. 5 (4) , 2014, 5901-5904.

[14] Thomas W. Rauber and Karsten Berns, "Kernel Multilayer perceptron".

[15] Muhammad Hussain , Summrina Kanwal Wajid , Ali Elzaart , Mohammed Berbar, "A Comparison of SVM Kernel Functions for Breast Cancer Detection", IEEE, 978-0-7695-4484-7/11, 2011.

BIOGRAPHIES

Author 1

Prachi Damodhar Shahare has completed Bachelor of Engineering in Information Technology in 2008 from Nagpur University and currently pursuing her Master of Technology in Computer Science and Engineering from Raipur Institute of Technology, Raipur (CG) under the supervision of Ram Nivas Giri, Raipur Institute of Technology, Raipur.

Author 2

Ram Nivas Giri is Assistant Professor in Department of Computer Science and Engineering, Raipur Institute of Technology, Raipur (CG). His area of interest is Artificial Intelligence and Artificial Neural Network. He has guided number of M. Tech and B. Tech projects. He has number of research publications at International/National Journals and Conferences.