

Empirical Study for Classification Ratio Using Neural Network and Classification Technique

Pankaj Kawadkar¹, Dr. L. K. Vishwamitra

¹PHD Scholar, ²Prof.

Department of CSE

SSSUTMS, SEHORE (M.P.), India

Abstract

Data mining gives various types of clustering classification algorithm for the various number of applications such as banking, education, medical science, fraud detection, pattern representation, feature extraction for the respective filed etc., there are various data mining algorithms such as supervised learning methods, unsupervised learning methods and semi supervised learning methods for the classification of data. In this article we present the comparative empirical study between the classification methods and measure the performance of both algorithms using some standard evaluation parameters.

IndexTerms- Supervised techniques, Neural network, Classifier, Support vector machine, UCI.

INTRODUCTION

Data mining gives various types of clustering classification algorithm for the various number of applications such as banking, education, medical science, fraud detection, pattern representation, feature extraction for the respective filed etc., there are various algorithm such as supervised learning methods, unsupervised learning methods and semi supervised learning methods [6]. The variety of data induced the problem of clustering and classification issue and degraded the performance of clustering and classification algorithm.

The data mining tools are covered various algorithms such as k-means clustering algorithm for the cluster or group formation for the initially level of data mining process then we used this cluster with the various classification algorithm such as support vector machine, knn classifier, neural network and optimization methods.

There are various algorithm we can used with the data mining techniques to improve the performance of the system or any algorithm for the various number of field such as information retrieval and mining of the data,

fraud detection, education sector, medical science, business transaction etc. Neural networks is a very important methods play with the data mining techniques for the improvement in clustering and classification methods and upgrade the result

by measuring and enhance the performance parameter. It a method of constructing classifiers, in which the learned model is a set of connected nodes by weighted connections. Neural network architecture and connection weights between the nodes influence the accuracy of classification.

In this paper we presents the data mining tools with the some neural network techniques and also other classification techniques such as support vector machine. Here the datasets are from also life science and medical science field such as Cleveland datasets, E-coli and diabetes diseases datasets. The neural network techniques works here on the concept of feed forward back propagation network which works is divided into number of layer i.e. input layer, hidden layer and output layer. The other techniques is support vector machine classifier which is a super sets of classifier in which data is divided into hyper plane and generally apply with the same pattern of data, here its gives better classify results than the other existing techniques [4].

In this section we introduce the techniques and tools about the data mining with un-supervised and supervised techniques and neural network, In section II we discuss about the proposed methods and architecture. In section III we define the dataset. In section IV we discuss about the empirical results. And finally conclude the paper.

II PROPOSED METHOD

In this section we discuss the about the Proposed methodology architecture for the classification of dataset system using Existing methods Feed forward Neural network and support vector machines.

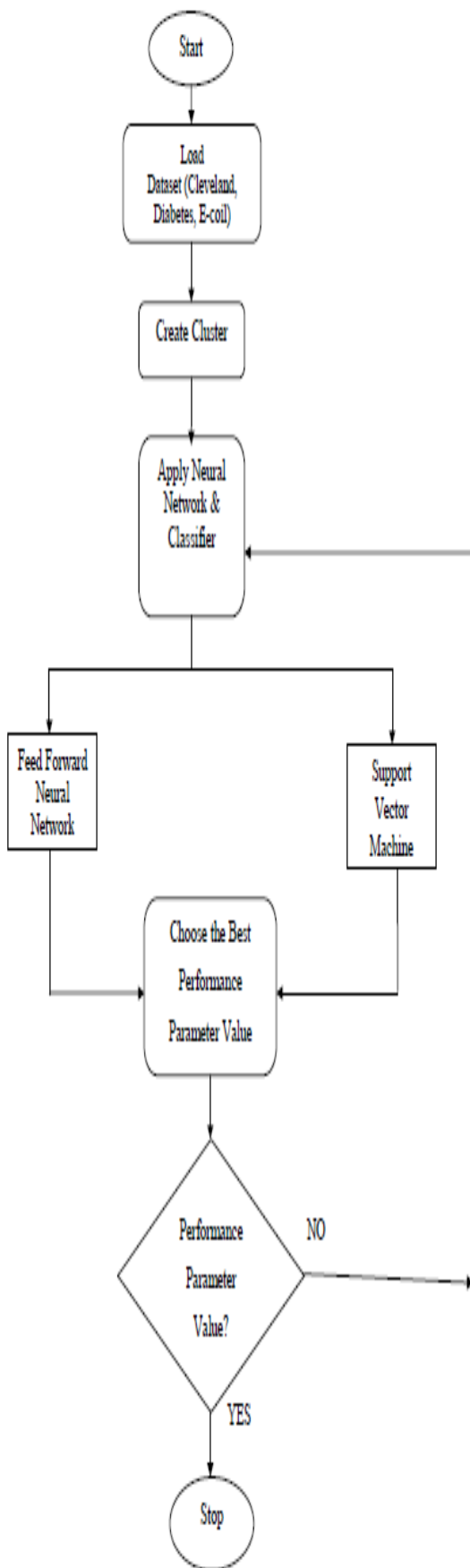


Fig 1: Neural network and Classifier based Pattern Analysis Model.

There are some steps we have to follow to implement this system are following:-

Step 1- begin the process of real and medical dataset upload from the UCI dataset.

Step 2- after the pre-processing of dataset we create the cluster or grouping for the respective dataset.

Step 3- After the dataset upload we create the grouping of dataset or cluster.

Step 4- Apply the feed forward neural network and classifier techniques for the selected dataset.

Step 5- after the successful apply the feed forward neural network and classifier techniques we getting the value of some performance parameter.

Step 6- after getting the some performance parameters value after applying the machine learning and optimization techniques i.e. accuracy, precision, recall etc. if not is satisfied then repeat the step 4 to step 5.

Step 7- finally we compare the all performance parameters value and we found that our Proposed methods gives us better results than other methods

Step 8- Exit the experimental simulation process.

III DATASET

In this section we discuss about the dataset which we used for the diseases detection in the field of health care. There are all these dataset types will be fetched from the UCI machine learning repository for the research purpose. In future we implement the diseases detection and improve the accuracy and other performance evaluation parameter with the help of all these dataset using data mining, optimization techniques. There are some dataset used for the experimental work for the diseases detection such as Cleveland, E-coli and Diabetes dataset etc.

IV EXPERIMENTAL STUDY

In this section we discuss about the comparative experimental result analysis using neural network and classification techniques, for the performance evaluation parameter we used some standard parameter such as accuracy, precision, recall and recall with using standard dataset implemented in Matlab software. Here we gives the comparative result study and comparative performance evaluation for the classification methods to improve the detection rate or ratio of classified data to improve the overall performance of system.

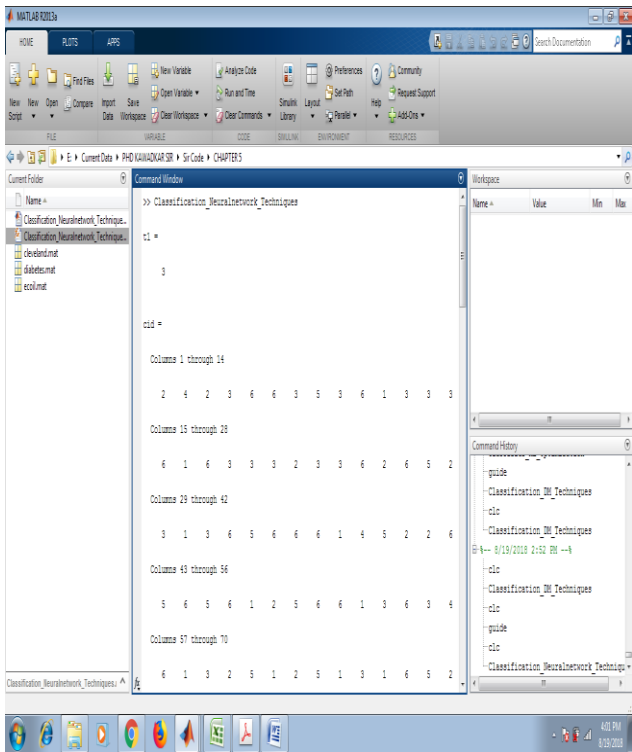


Fig 2: This windows show that the result of Neural network methods in the experimental process using Cleveland dataset.

Dataset name	Method	Accuracy (%)	Precision (%)	Recall (%)
Cleveland Dataset	Feed Forward Neural Network	84.88	86.57	87.21
	Support Vector Machine	88.67	90.45	89.36

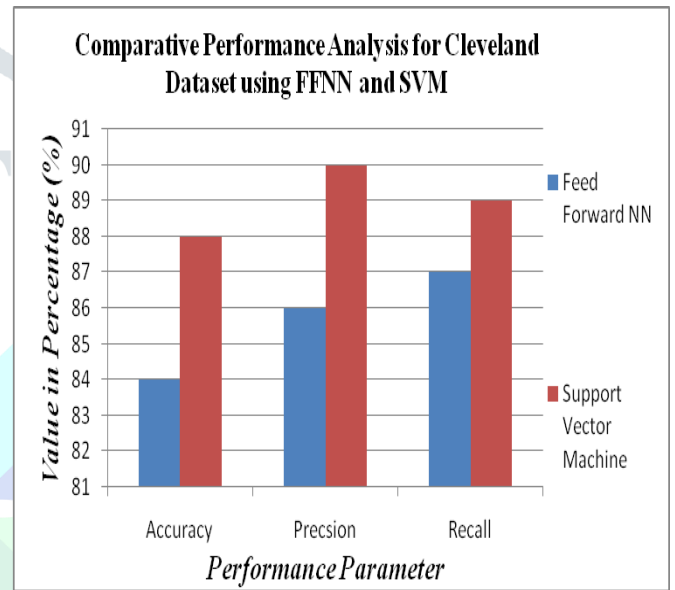


Fig 4: Shows that the comparative result for classification techniques using Cleveland dataset for performance parameter such as Accuracy, precision and Recall.

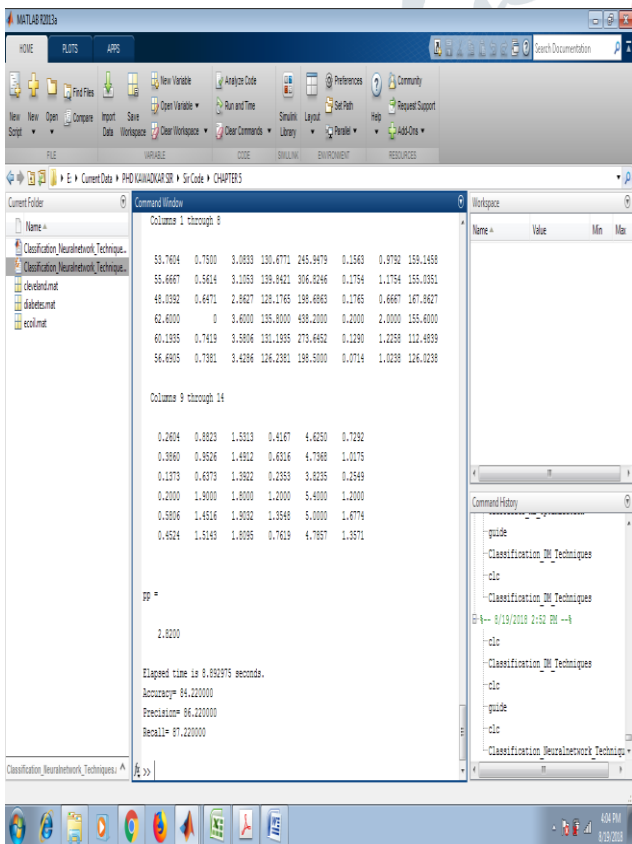


Fig 3: This windows show that the result of Neural network methods with complete result in the experimental process using Cleveland dataset.

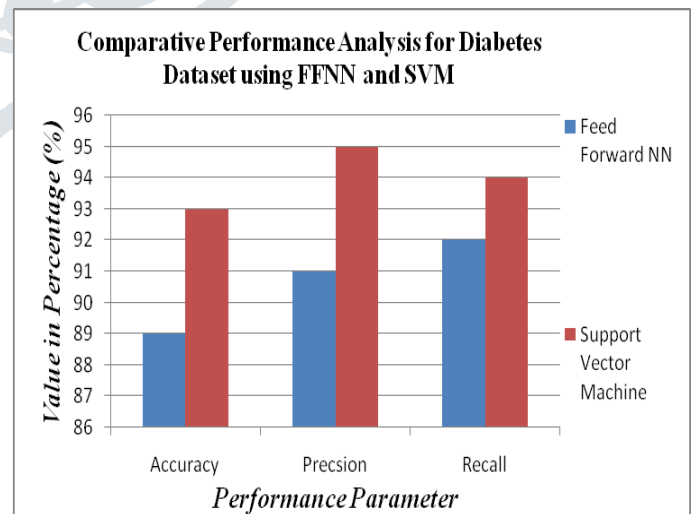


Fig 5: Shows that the comparative result for Feed Forward Back Propagation Neural Network and Classifier method i.e. Support Vector machine using Diabetes dataset for performance parameter such as Accuracy, precision and Recall.

Table 1: Show that the comparative result analyses study for the Cleveland dataset with using classification techniques.

V CONCLUSIONS

In this paper we presents the comparative empirical result analysis study between the classifier, feed forward neural network classifier and support vector machine classifier, using both classification techniques we get some performance evaluation parameters for the classified dataset such as life science data including Cleveland dataset, diabetes dataset and E-coli dataset. Here our empirical result study shows that the support vector machine gives better results than the feed forward neural network. In future we can also used some optimization tools for the reducing the execution time and improve the performance.

REFERENCES:-

[1] Ankita Dewan and Meghna Sharma “Prediction of Heart Disease Using a Hybrid Technique in Data Mining Classification”, IEEE, 2015, Pp 704-706.

[2] Seyed Mohammad Jafar Jalali, Sergio Moro, Mohammad Reza Mahmoudi, Keramat Allah Ghaffary “A Comparative Analysis of Classifiers in Cancer Prediction Using Multiple Data Mining Techniques” ISCTE-IUL, 2017. Pp 1-12.

[3] S.Leoni Sharmila, C.Dharuman, P.Venkatesan “Disease Classification Using Machine Learning Algorithms - A Comparative Study” International Journal of Pure and Applied Mathematics, Vol-114, 2017, Pp 1-10.

[4] Meherwar Fatima, Maruf Pasha “Survey of Machine Learning Algorithms for Disease Diagnostic” Journal of Intelligent Learning Systems and Applications, 2017, Pp 1-16.

[5] Tabreer T, Hasan, Manal H, Jasim and Ivan A. Hashim “ Heart Disease Diagnosis System based on Multi-Layer Perceptron neural network and Support Vector Machine” International Journal of Current Engineering and Technology, Vol-7, 2017.

[6] Ming Yuchi and Jun Jo “Heart Rate Prediction Based on Physical Activity using Feed forward Neural Network”, International Conference on Convergence and Hybrid Information Technology 2008 International Conference on Convergence and Hybrid Information Technology, 2008, Pp 344-350.

[7] R.Sujatha, E.P.Ephzibah, Sree Dharinya, G. Uma Maheswari, V.Mareeswari Vamsi Pamidimarri “ Comparative study on dimensionality reduction for disease diagnosis using fuzzy classifier” International Journal of Engineering & Technology, 2018, Pp 79-84.

[8] Basheer Mohamad Al-Maqaleh, Ahmed Mohamad Gasem Abdullah “Intelligent Predictive System Using Classification Techniques for Heart Disease Diagnosis” International Journal of Computer Science Engineering, 2017. Pp 145-151.

[9] M. S. A. Megat Ali and A. H. Jahidin “Hybrid Multilayered Perceptron Network for Classification of Bundle Branch Blocks”, IEEE., 2011, Pp 149-154.

[10] Arijit Ukil, Soma Bandyopadhyay, Chetanya Puri and Arpan Pal “HEART-TREND: AN AFFORDABLE HEART CONDITION MONITORING SYSTEM EXPLOITING MORPHOLOGICAL PATTERN”, IEEE, 2016, Pp 6260-6264.