

Improving the Performance of Diseases Diagnosis using Classification and Optimization Methods

Amita Vishwakarma¹, Prof. Pooja Meena², Prof. Chetan Agrawal³

¹M. Tech Scholar, Department of CSE, RITS, Bhopal (India)

²Assistant Professor, Department of CSE, RITS, Bhopal (India)

³Assistant Professor, Department of CSE, RITS, Bhopal (India)

¹amitavishwa06@gmail.com, ²meena.pooja1@gmail.com, ³chetan1only@gmail.com

Abstract

In this paper we proposed a new model which is based on the classification methods such as support vector machine, neural network and optimization methods for the improving the classifier results in the terms of some performance parameters such as accuracy, precision, recall etc., here we measure the all performance parameters for the various dataset such as heart patients, liver patients and cancer patients and improve the rate of classification or results with compare than other existing techniques. In this paper our experimental results shows that the better detection rate of classification for performance parameters than other existing techniques.

Keywords: - Support vector machines, Particle swarm optimization, neural network, Classifier, Medical Science, Diseases Diagnosis.

Introduction

The mining of healthcare database is very critical issue. The healthcare data stored the information about medical diseases and patient's information. For the estimation of patients and disease used some intelligent software for the predication of disease [4]. The prediction of disease and medicine is fundamental issue in health care environment. For the extraction of better information used data mining technique for the healthcare system.

Heart disease prediction and classification is major issue in current lifestyle and diagnosis system. For the prediction and classification used various classification algorithm are used such as support vector machine, neural network and other technique [1].

Data mining is the most common methods used by researchers for diagnosis and monitoring of various disease are classification and clustering methods [13]. Basically its provide the various classification and clustering methods for the grouping of data into known and unknown class on the basis of pattern extracting or relevant extraction of information from the patients diagnosis using dataset [8].

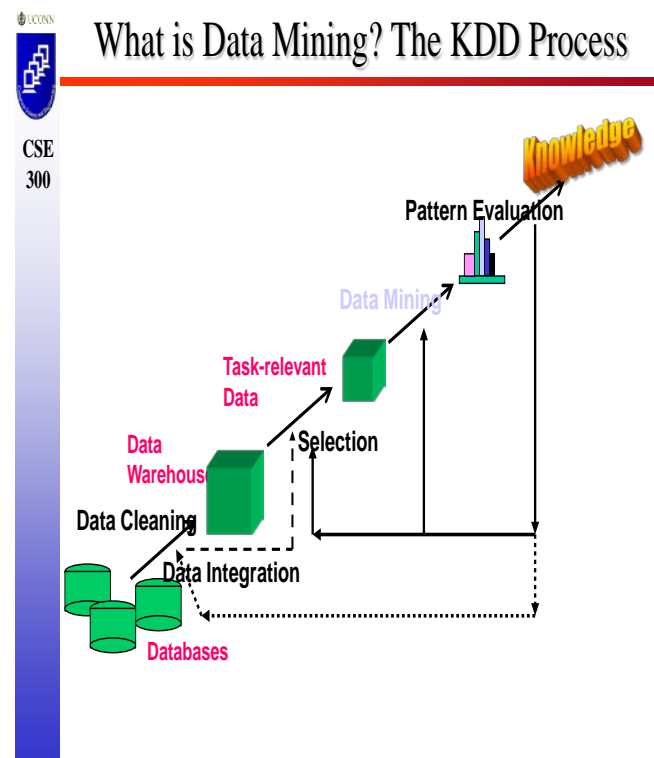


Figure 1: Process of knowledge extraction using data mining.

Support vector Machine is a classifier which is used to improve the results or classification ratio for the various number of application, there is also some other classifier such as decision tree, KNN and bay's classifier [10]. Support vector machine classifiers classify the data into hyper-plane the main aim of this classifier to reduce the margin between those planes. The margin is basically create the separation between more than one class in a support vector machine.

The rest of this paper is organized as follows. In next section II we discuss about the Research motivation, In section III we discuss about the proposed methods and architecture for the health care diagnosis system using classification and swarm intelligence techniques. In section next we define the dataset which is used for the experimental process. In next section we discuss about the comparative study of medical diseases diagnosis system techniques with their empirical results. Finally in next section we conclude the complete research work.

II Research Motivation

From the past years there is medical science interesting domain for researcher, due to large number of population are affected from its. In our country most of the peoples are not able to get medical treatment on time at everywhere, or most of the peoples suffer from the various types of diseases but not getting any prevention to recover from diseases [11]. To cover large number of peoples we need rich infrastructure such as large number of hospitals, medical lab, equipment, diagnostic tools, machinery, huge memory to stored large number of information, available data everywhere at each time. There are also the tools and techniques for the medical diseases diagnosis are play very vital role in this sector [3]. There is various techniques such as data mining tools which further classified the data in the supervised and unsupervised manner, some evolutionary techniques such as genetic algorithm mostly used for the big data analytics, neural network classifier, support vector machine, decision tree classifier, rule based classifier and some swarm intelligence family methods for the optimization such as ant colony optimization, particle swarm optimization and honey bee classification [5].

III Proposed Method

In this section we proposed the new model for the health care system to compute some performance parameters such as accuracy using classification and optimization techniques. Here we improve the efficiency rate in the terms of accuracy for the proposed system compare than existing system which is provide better results in the medical science domain.

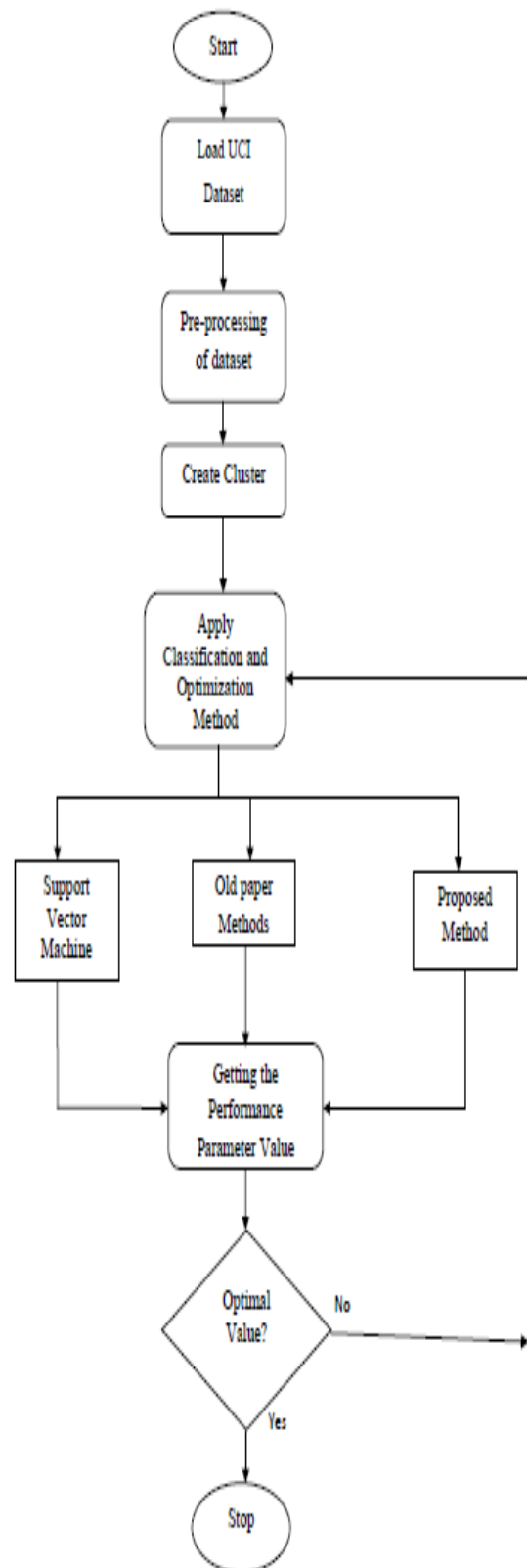


Figure 2: Classification and optimization model for the diseases.

Model Description

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

- Step 1- begin the process of medical diseases diagnosis classification and optimization model.
- Step 2- after the start of model working we upload the UCI dataset..
- Step 3- after the successfully upload of the dataset we apply pre-processing and cluster creation process for the respective dataset.
- Step 4- Apply the classification and optimization techniques for the selected dataset.
- Step 5- after the successful creation of cluster we apply some classification techniques and optimization methods.
- Step 6- after the each techniques we getting some performance parameter values for the each dataset.
- Step 7- Select the best optimal features.
- Step 8- if the getting performance parameter value is not optimal then repeat step 4 to step 7.
- Step 9- if the getting performance parameter value is optimal we used it.
- Step 10- finally we stop the complete model for the experimental simulation process.

IV 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, the used dataset is here Heart dataset, liver dataset and Cancer dataset.

V Experimental Result Analysis

In this section, experimental process of we show that the comparative result analysis study for the Health care sector with disease diagnosis of various dataset such as Heart, Liver, Cancer etc. are performed. This process of disease diagnosis of various dataset is done by using Three methods that are support vector machine, the base paper method and Proposed method. For the evaluation of performance parameter we used Matlab software for the authenticity and effectiveness of results.

Precision: measures the proportion of predicted positives/negatives which are actually positive/negative.

Recall: It is the proportion of actual positives/negatives which are predicted positive/negative.

Accuracy: It is the proportion of the total number of prediction that was correct or it is the percentage of correctly classified instances.

Dataset name	Method	Accuracy in (%)	Precision in (%)
Heart	SVM	78.23	75.49
	OLD METHOD	80.25	77.24
	PROPOSED	93.59	87.47

Table 1: Show that the comparative result analysis studies for the Heart dataset with using existing and proposed method.

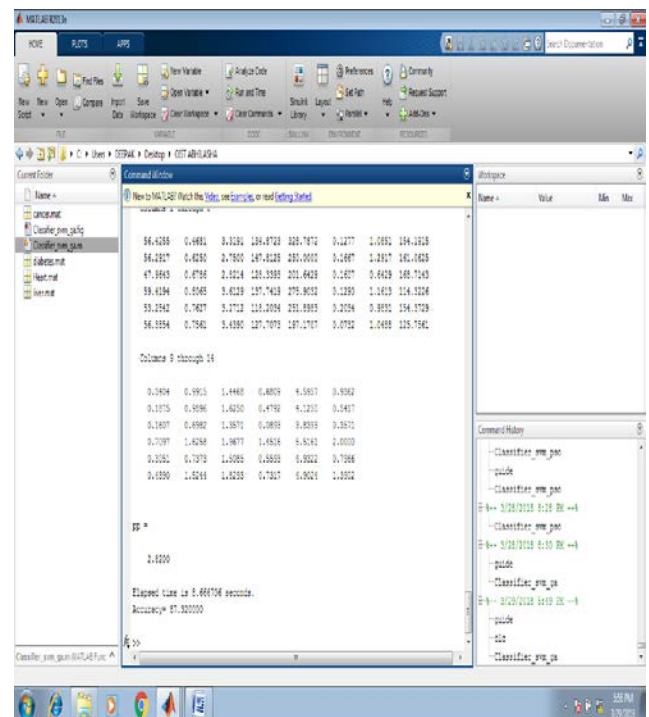


Figure 3: This windows show that the result of Old method with accuracy in the experimental process using Heart dataset.

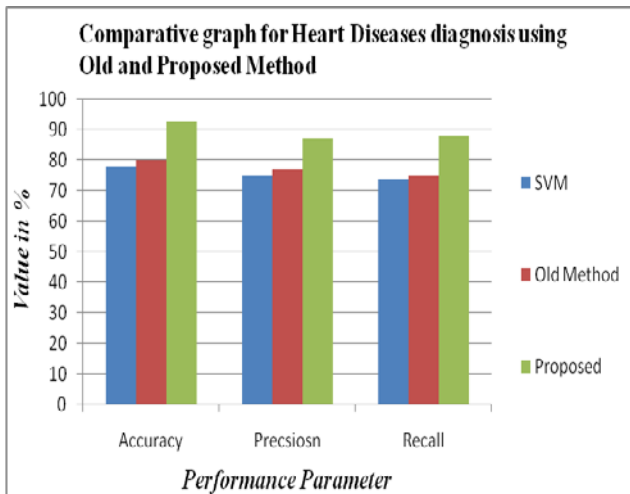


Figure 4: Show that the comparative result analysis for the Heart disease diagnosis using various classification and optimization techniques, our empirical result study shows that better accuracy than existing methods.

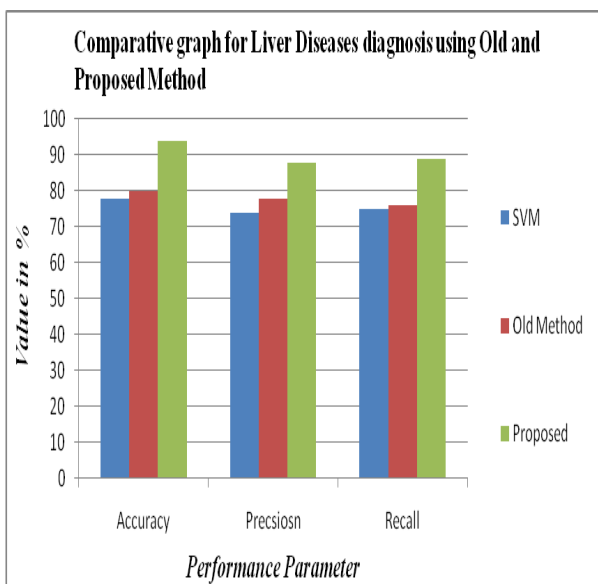


Figure 5: Show that the comparative result analysis for the Liver disease diagnosis using various classification and optimization techniques, our empirical result study shows that better accuracy than existing methods.

VI Conclusions and future work

To provide good and on time healthcare services is the primary goal of any government or any agencies, the process of diseases diagnosis is the challenging area for the medical science. In this paper we focus on pattern extraction and pattern analysis of healthcare data

environment using classifier, support vector machines and optimization techniques. The diversity of medical diagnosis of disease data are increase day to day and very challenging task to store and used in a large amount of data, in future we can also used some big data analytics for the healthcare system to enhance the results more and more.

REFERENCES:-

- [1] Majid Ghonji Feshki and Omid Sojoodi Shijani "Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network", IEEE, 2016, Pp 48-53.
- [2] Aigerim Altayeva, Suleimenov Zharas and Young Im Cho "Medical Decision Making Diagnosis System Integrating k-means and Naïve Bayes algorithms", ICCAS, 2016, Pp 1087-1092.
- [3] Rahul Joshi, Minyechil Alehegn "Analysis and prediction of diabetes diseases using machine learning algorithm: Ensemble approach" International Research Journal of Engineering and Technologyol-4, 2017. Pp 426-435.
- [4] S.Poongothai, C.Dharuman and P.Venkatesan "A Comparison of Fuzzy Genetic and Neuro Genetic Hybrid Algorithm for the Classification of Diabetes Disease" International Journal of Pure and Applied Mathematics Volume 113, 2017, Pp 208-216.
- [5] M.Akhil jabbar , Priti Chandra and B.L Deekshatulu "Prediction of Risk Score for Heart Disease using Associative Classification and Hybrid Feature Subset Selection", IEEE, 2012, Pp 628-634.
- [6] Delia-Maria FILIMON and Adriana ALBU "Skin Diseases Diagnosis using Artificial Neural Networks", IEEE, 2013, Pp 1-6.
- [7] Sellappan Palaniappan and Rafiah Awang "Intelligent Heart Disease Prediction System Using Data Mining Techniques", IEEE, 2008, Pp 108-115.
- [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.

Patients Using Fuzzy Classification Technique”, IEEE, 2010, Pp 1-7.

[11] Theresa Princy. R and J. Thomas “Human Heart Disease Prediction System using Data Mining Techniques”, ICCPCT, 2016, Pp 1-5.

[12] Asha Gowda Karegowda and Darshan M “Optimizing Feed Forward Neural Network Connection Weights Using Artificial Bee Colony Algorithm”, International Journal of Advanced Research in Computer Science and Software Engineering, 2013, Pp 452-454.

[13] Sultan Noman Qasem, Siti Mariyam Shamsuddin, Siti Zaiton Mohd Hashim Maslina Darus and Eiman Al-Shammari “Memetic multiobjective particle swarm optimization-based radial basis function network for classification problems”, Elsevier, 2013, Pp 165-190.

[14] Ashraf Osman Ibrahim, Siti Mariyam Shamsuddin, Nor Bahiah Ahmad and Sultan Noman Qasem “Three-Term Backpropagation Network Based On Elitist Multi-objective Genetic Algorithm for Medical Diseases Diagnosis Classification”, Life Science Journal, 2013, 1815-1822.

[15] Mr. Deepak Dhanwani and Prof. Avinash Wadhe “Study Of Hybrid Genetic Algorithm Using Artificial Neural Network In Data Mining For The Diagnosis Of Stroke Disease”, International Journal of Computational Engineering Research, 2013, Pp 95-100.

[16] Fabio Mendoza Palechor, Alexis De la Hoz Manotas, Paola Ariza Colpas , Jorge Sepulveda Ojeda, Roberto Morales Ortega, Marlon Piñeres Melo “ Cardiovascular Disease Analysis Using Supervised and Unsupervised Data Mining Techniques” Volume 12, Number 2, February 2017. Pp 81-90.

[17] Syed Umar Amin, Kavita Agarwal and Dr. Rizwan Beg “Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors”, IEEE, 2013, Pp 1227-1231.

[18] B V Sumana and T. Santhanam “Prediction of diseases by Cascading Clustering and Classification”, ICAECC, 2014, Pp 1-8.

[19] V Krishnaiah, M Srinivas, Dr.G Narsimha and Dr.N Subhash Chandra “Diagnosis of Heart Disease