International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 6 Issue 6, June-2017, Impact Factor: 4.059

A Weighted SVM based Model for Heart Disease Prediction

Bijender[1], Vinod Rampure[2]

[1]Student, M.Tech. (CSE), MITRC, Alwar, India
[2]Assistant Professor, MITRC, Alwar, India

ABSTRACT

Heart is the most critical human organ in which some disorder or the blockage identification can even cause the death. There are number of heart diseases including the ventricular disease, sleep disorder, blockage, abnormal heart rate etc. The ECG signal is the electrical form of heat beat the can be used to identify these all disorders. The disease can be predicted by analyzing the features or symptoms of patients. In this present work, a Bayesian weighted SVM method is defined for heart disease prediction. The comparative observations show that the model has improved the accuracy of heart disease prediction.

Keywords: Heart, ECG, Abnormality, Heart Disease, SVM.

1. INTRODUCTION

ECG is the signal form to represent the heart activity and it is introduced by a German psychiatrist Hans Berger. This signal form is recorded from heart to represent the fluctuation in the heart rate or beat and specification to the systematic manner under potential change analysis over the heart behavior. This signal form is here defined under the identification of reaction analysis defined under the potential difference observations and defined with the specification of heart rate. This kind of signal rate specification is here defined under the consideration of heart rate and the electrode level change under signal amplitude value and magnitude value analysis. This kid of analysis is based on the source and time line specification. The specification of the signal also include different kind of noise and interference defined under the specification of signal recording source and interference specification and signal recording to the signal and formation.

A) Signal Constraints

The potential change based analysis so that the potential to the neuron specification and generation under post synaptic potential change based analysis so that the signal derivation will be obtained from the work. The contribution level analysis is here defined under potential propagation to the system and by analyzing the signal nerve level axons so that the signal formation and analysis of potential change will be done effectively. This kind of signal level analysis is defined to generate the synchronous activity analysis under neuron specification. The eventual analysis is also performed to generate the structure and the behavior of the signal under signal specification in reticular form. The characteristics of the signal form is shown in figure 1

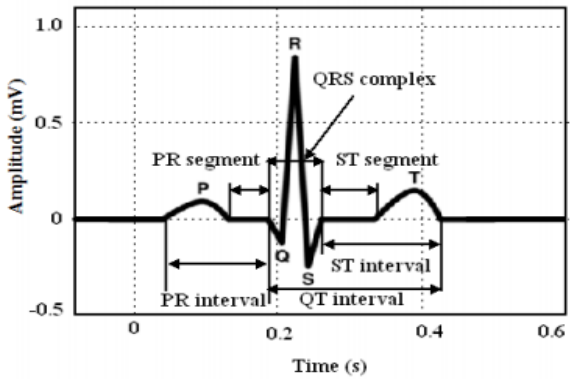

Figure 1 : ECG Signal Form

B) Heart Disorder

This signal form represents the heart range analysis under disease identification and disorder identification. The signal form is defined here defined in the normal sinus rhythm form. The heart rate analysis is here defined to represent the specification of rate beat. The RR interval based analysis is defined to obtain the change analysis in the breathing cycle. The specification to arrhythmia here taken under heart rate analysis and blood circulation analysis so that the effective heart values can be obtained. This kind of formation can be obtained under contraction analysis so that the derivation to the affective perfusion can be obtained.

C) Sleep Disorder

Sleep is considered as the non activity state of human body but it is defined as the basic behavioral and physiological criteria to divide the behavior in two main stages called rapid eye moment sleep and non rapid eye movement sleep.

In this paper, a study to the various disease associated with heart system are identified. The paper also described the some of the mining based solution to identify or classify these diseases. In this section, a study on ECG signal and related disease identification is done. The work has defined the functioning of ECG system and its recording parameters. Later on, different disease types such as sleep apnea, ventricular disorders are identified. In section II, the work defined by earlier researchers is discussed. In section III, the mining approaches are discussed using which the disease can be identified. In section IV, the conclusion of the work is presented.

II. RELATED WORK

Lot of work is already defined by different researchers to process on available dataset and perform information extraction using clustering approaches. Some of the work defined by earlier researchers on information processing is discussed in this section. Lei Jiang[1] has defined a work on optimization of clustering approaches using adaptive PSO approach. Author defined an improved clustering mechanism under fitness function specification to improve the clustering process. Author defined an adaptive niche approach based on comparative analysis so that effective information division will be performed. Author presented the dimension reduction approach with large difference analysis under cluster specification and obtained the effective data clustering. Shuai Li[2] has defined a work to improve the clustering process using spatial characteristics analysis and PSO approach. Author provided the estimation on available clusters and provides the unification to the PSO approach and Bayesian information criterion analysis and provided the numeric clustering algorithm. Author provided the characteristics idea analysis and provided the local optimal problem identification based on the analysis. Author provided the effective information estimation and achieves good performance and resolved most of the clustering errors. Rehab F[3] has defined a work to improve the clustering processing using improved PSO approach. Author provided a hybrid two phase approach with integration of PSO, genetics and k-means algorithm. Author avoided the premature local optima and provided the heuristic search mechanism to extract the information accurately. The parameters considered by the author include velocity and position updation based on rule analysis. Author defined an improved process for improving the selection, mutation and crossover options over the genetic approach. Author provided the solution space for initial cluster formation and centroid updation based on the globalized features over the dataset. Author provided the work benchmark datasets so that the effective information processing and information convergence will be obtained. Author also defined a comparative study on various optimization algorithms

Surat Srinoy[4] has presented an improvement to clustering for network security modeling using ACO integrated clustering approach. Author presented the work on raw clusters formed using Kmeans and PSO approach. Later on the identification of partitions is done to reduce the complications and complexities over the datasets. Author presented the clustering mechanism to reduce the dataset suspicious activities so that attack resolvement over the dataset will be obtained. Author presented the experimental view to represent the dataset more effectively.

Shafiq Alam[5] has defined an improvement to the clustering mechanism using swarm intelligence approach. Author implemented the work in decentralized environment. Author also presented the comparative analysis on benchmark datasets and using standard approaches. Author provided the evidence of the data effectiveness so that more effective information processing will be done. Alireza Ahmadyfard[6] has defined a hybrid approach using PSO and KMeans algorithms for generating the clusters over the datasets/ Author defined the convergence over the initial stage and provided the global search so that information processing over the dataset will be obtained. Author defined a group division approach and analyzed it on various sub datasets and proved the reliability of work under experimentation.

Merwe et.al[7] defined a work on clustering approaches PSO approach. Author defined work in two main stages. In first stage, the identification of cluster centers is done based on information analysis and later on cluster elements are identified using PSO approach. Author defined the work on six different datasets and provided the comparative analysis between these approaches to show the work effeteness. Kenyon et. Al.[8] has defined tag based suvey analysis on gene

expression to cover the problem identification for infection based disease. Author provided the characterization study for the work under the specification of disease and relative patterns. Pudilo et.. al[9] has presented an improvement to the clustering approach using multi objective optimizer. Author provided an extension to heuristic search mechanism called PSO. Author presented the extension to the problem based on directional analysis and generated the sub swarms over the dataset to cover the dataset and to perform the information exchange effectively over the dataset. Author provided the a validated method to test the function under evolutionary function specification and respective to the dataset so that the effective information processing will be done. Yang et. al.[10] has defined a work to improve the quantum concept using particle swarm optimization. Author provided the methodology in evolutionary computation and provided the extremely effective information processing so that information exchange will be performed for discrete problems. Author provided the powerful tool to analyze the information set and provided the equalized check over the dataset and provided the information transition effectively.

III. HEART DISEAES IDENTIFICATION APPROACHES

The medical information processing is one of the most application of data mining and classification methods. The method requires the expert level processing to recognize the disease existence. In this work, a hybrid and rule based model is presented to recognize the heart disease more accurately. The proposed disease based model is shown here in figure 2

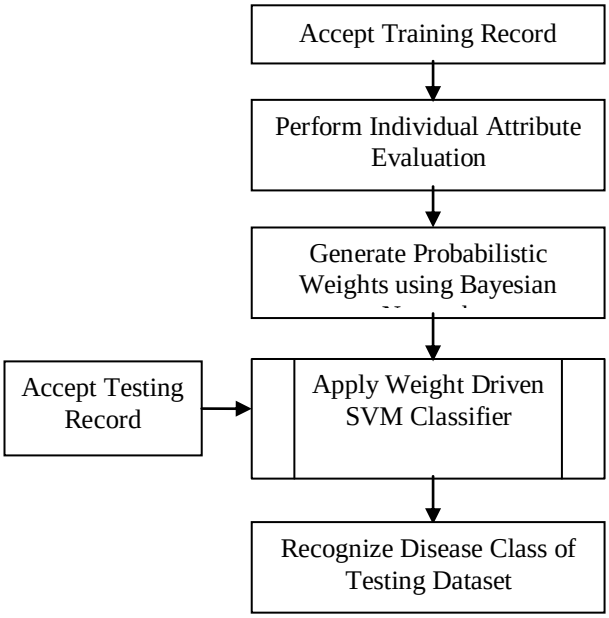

Figure 2: Hybrid Classification Model

The proposed work model is divided in two main work stages. In first stage, the analysis on the attributes is performed. At this stage, the probabilistic derivation is performed based on the statistical features. The Bayesian network is applied at this stage on training set to generate the rules. Once the rules are generated, the SVM classifier is applied to perform the classification. Figure 2 is showing the algorithmic model to perform the classification using this proposed algorithmic approach. The algorithm associated to the work is shown in next section.

3.1 Classification

In this work, a Bayesian and SVM based hybrid classification model is presented to perform the heart disease classification. The proposed model first processed the training set and generated the weighted features. Once the features are generated, the training features are processed to perform the classification. The probabilistic estimation is applied on SVM and Bayesian network collectively and defined the rule for identification of the disease class.

IV. RESULTS

In this present work, the SVM and Bayesian Network based hybrid classification model is presented to process the heart disease features and to identify the patient disease. The model description is already provided in chapter 3. The comparative analysis is provided against Bayesian Network and SVM based methods separately. The description of processing training and testing set is shown in table 1

Table 1: Processing Classification Dataset (Dataset I)

Features	Values
Size of Training Set	197
Size of Testing Set	83
Disease Classes	Present and Absent
Existing Methods	Bayesian Network & SVM
Proposed	Hybrid Bayesian +SVM

Table 1 is showing the features of processing dataset for classification process. At first the generated statistical training set features are processed on each individual classifiers and generate the weighted rules. These weight rules are applied on testing feature set in probabilistic selective form

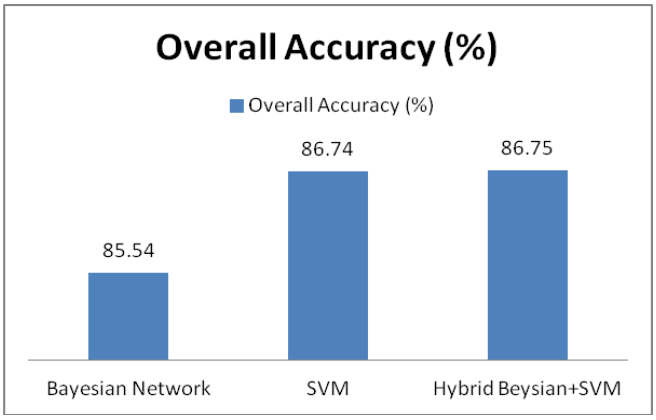

Figure 3: Overall Accuracy (%) (Dataset I)

Figure 3 shows the comparative evaluation on accuracy of proposed hybrid model against the existing SVM and Bayesian models. The figure shows that the existing Bayesian model has recognized the 85.54% instances accurately and the SVM model has predicted about 86.74% instances correctly. Whereas, the proposed hybrid model has improved the accuracy of disease classification about 86.75%.

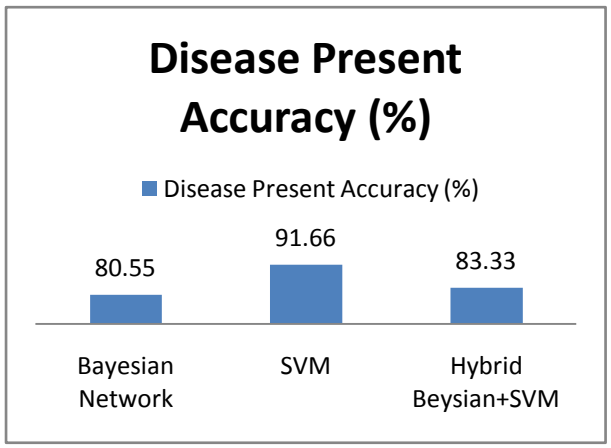

Figure 4: Disease Present Accuracy (%)(Dataset I)

Figure 4 shows the comparative evaluation on disease present accuracy of proposed hybrid model against the existing SVM and Bayesian models. The figure shows that the existing Bayesian model has recognized the 80.55% instances accurately and the SVM model has predicted about 91.66% instances correctly. Whereas, the proposed hybrid model

has provided significant results in case of disease existence. The disease identification accuracy of proposed hybrid model is 83.33%.

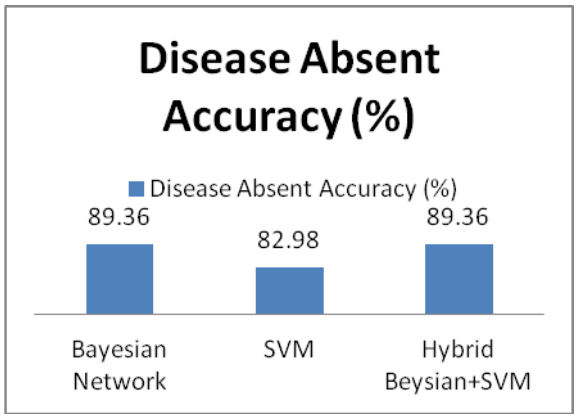

Figure 5: Disease Absent Accuracy (%)(Dataset I)

Figure 5. shows the comparative evaluation on disease identification accuracy of proposed hybrid model against the existing SVM and Bayesian Network models. The figure shows that the existing Bayesian Network model has recognized the 89.36% accurately and the SVM model has predicted about 82.98% correctly. Whereas, the proposed hybrid model has improved the results in case of disease recognition. The disease identification accuracy of proposed hybrid model is 89.36%.

CONCLUSION

In this present work, a hybrid model is presented to predict the heart disease more accurately. In this model, the probabilistic and the supervised classifiers are combined to improve the accuracy of disease prediction. The comparative analysis is performed against the Bayesian network and SVM methods individually. The comparative results shows that the proposed model has improved the accuracy of accurate detection of heart disease.

REFERENCES

[1]. Lei Jiang Lixin Ding, "Hybrid Adaptive Niche to Improve Particle Swarm Optimization Clustering Algorithm", 978-1-4244-9953-3/11 ©2011 IEEE
[2]. Shuai Li, Xin-Jun Wang, "X-SPA: Spatial Characteristic PSO Clustering Algorithm with Efficient Estimation of the Number of Cluster", Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 978-0-7695-3305-6/08 © 2008 IEEE
[3]. Rehab F. Abdel-Kader, "Genetically Improved PSO Algorithm for Efficient Data Clustering", 2010 Second International Conference on Machine Learning and Computing, 978-0-7695-3977-5/10 © 2010 IEEE
[4]. Surat Srinoy, "Combination Artificial Ant Clustering and K-PSO Clustering Approach to Network Security Model", IEEE 2006
[5]. Shafiq Alam, Gillian Dobbie, Patricia Riddle, M. Asif Naeem," Particle Swarm Optimization Based Hierarchical Agglomerative Clustering ",2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology
[6]. Alireza Ahmad yfard, Hamidreza Modares ," Combining PSO and k-means to Enhance Data Clustering ",2008 Internatioal Symposium on Telecommunications, 978-1-4244-2751-2/08 ©2008 IEEE
[7]. DW van der Merwe, 2003, "Data Clustering using Particle Swarm Optimization", 0-7803-7804-0 103AI 7.00 0 IEEE, pp. 215-220
[8]. F. Kenyon, 2003, "Expressed sequence tag survey of gene expression in the scab mite Psoroptes ovis – allergens, proteases and free-radical scavengers". pp. 451-460
[9]. Pudilo, 2005, "Bayesian Hierarchical Clustering", Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany.
[10]. Yang, "Features selection of SVM and ANN using particle swarm optimization for power transformers incipient fault symptom diagnosis", International Journal of Computational Intelligence Research, ISSN no-0973-1873.