**TITLE**

**IMPROVISED NATURE INSPIRED DEEP BELIEF NETWORK AND ITS APPLICATION IN HEART DISEASE PREDICTION**

**ABSTRACT**

This paper proposes Deep belief network with particle swarm optimization. The network is optimized by nature inspired algorithm called particle swarm optimization (PSO). The dimensions of deep belief network consisting of Restricted Boltzmann machine (RBM) layers and neural network layers are adjusted using nature inspired algorithm. The dimensions like size of RBM layer and the size of neural network layers are optimized using PSO. The comparison between proposed strategy and other known machine learning algorithms like SVMlight , Neural network is also done. It is concluded that deep belief network with particle swarm optimization gives the best results with accuracy 91.18% and 94.88% when tested on two benchmarked heart disease datasets from UCI repository namely Cleveland dataset and Hungarian dataset respectively.

Keywords: Cardiac disease classification, Neural network, Deep belief network (DBN) , Restricted Boltzmann machine (RBM) and Particle swarm Optimization(PSO),Deep learning

**INTRODUCTION**

Today, Machine learning and data science are the emerging technologies in medical field. Heart diseases [\cite {yusuf2001global}] today are an increasing threat to both rich and poor. The lack of physical activity and alcohol are the two main reasons for increasing heart diseases today. The other factors include obesity, smoking, age, pollution and diet also play a vital role in deciding whether a person will suffer from heart disease or not. Death rate due to cardiovascular problem is increasing alarmingly in developing countries and poses a greater threat in comparison to the developed countries. Heart disease mortality in rich countries is less, but it shows a sharp increase in the poor nations [\cite{riley2014noncommunicable}]. Accurate diagnosis of heart diseases is highly important. Computational methods are used to assist a doctor to diagnose disease of a patient. Classification of medical data requires a high accuracy in predictions so that they can be of precise aid. Medical data set classification and prediction [\cite{soni2011predictive}] is used in diagnosis of diseases beforehand. Mostly the Doctor and patients are not aware of the disease just on the basis of symptoms. Classification of such a medical data-set is the most important problem to solve. In classification problem there is an input set and output set of class labels and the aim is to find the decision plane which maps them. In medical field, various computer researchers have attempted to apply diverse techniques to improve the accuracy of classification for the given data. Classification techniques whose classification accuracy will give better information to identify the potential patients will therefore be used to improvise the diagnosis accuracy. In the recent studies, metaheuristic algorithms like simulated annealing, genetic algorithms, and particle swarm optimizations and also data mining techniques like Bayesian networks, artificial neural network, fuzzy logic, and decision tree are applied for classification of medical data and remarkably meaningful results were obtained.

**Machine learning**

Artificial intelligence involves the science of making the machine learn. It is the field widely used for classification and prediction. The three types of learning under machine learning are supervised, unsupervised and semi-supervised learning. In supervised learning, Both the sets of input and output are available and the goal is to devise a supervised algorithm that maps every input to output. Digit recognition is an example of supervised learning in which input given is handwritten digit and the algorithm is used to classify the type of digit it is .[\cite{kala2010offline}] In semi-supervised learning, both labeled and unlabeled data are used. It needs only a small amount of labeled data and a lot of unlabeled data. In unsupervised learning. The data is fed into the algorithms and the data is grouped into different clusters using the data distribution. The number of groups also vary according to the the data distribution. Some of the most common algorithms used for unsupervised learning are K-nearest-neighbor, K-means etc.

**Neural Networks**

Neural networks [\cite{hagan1996neural}] resembles biological network of brain neurons. Similar to central nervous systems, neural network consists of an interconnected group of neurons. Each node gets input from other nodes and weights between nodes adjust so that the whole network learns to compute the output. There are various types of neural networks structures with each having its own learning algorithm. The neural network is a novel computer architecture compared to traditional computers. An artificial network of neurons connected with each other to give a specified output on applying input is called neural network. It is used to map input into outputs with the help of input layer, various hidden layers and output layer.

**Deep Learning**

Deep learning [\cite{bengio2009learning}] is the science of making the machine learn with the help of deep neural network involving a large number of layers of neurons. The training of such a deep network becomes difficult due to large number of neurons, hidden layers and a large number of interconnections between the layers. So , to overcome this limitation Deep belief network , a disciple of deep learning is used. The deep belief network consist of Restricted Boltzmann Machine layers and neural network layers. Restricted Boltzmann Machine is a type of Boltzmann machine with a restriction that the graph nodes can be divided into two disjoint sets. The first step of training DBN is to learn a layer of features from the visible units, using Contrastive Divergence (CD) algorithm. It consists of a hidden layer and a visible layer with a weight distribution between them. .To learn characteristics of features of next hidden layer, previous features activation is treated as visible unit. Finally, the whole DBN is trained when the learning for the final hidden layer is achieved. This simple greedy learning algorithm works for training DBN. This is because that training RBM using CD algorithm for each layer looks for the local optimum and the next stacked RBM layer takes those optimally trained values and again look for the local optimum. At the end of this procedure, it is likely to get the global optimum as each layer consistently trained to get the optimum value.

The energy distribution for the network is given by the equation

E(v,h)=h'Wv+b'v+c'h

Here, W is the weights matrix of the RBM , b denotes the biases of visible unit and c denotes the biases of the hidden, h and v denote the hidden and visible layer activations respectively. The probability distribution of network is given by

\bf p(v,h)=\frac{1}{N}e^{E(v,h)}

Where N in the equation above denotes the normalization constant for the given distribution. The algorithm for training DBN is done as follows:

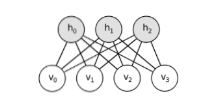
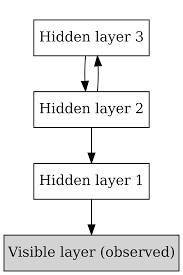
1. Let X be a input matrix, which is also the set of feature vectors.

2. Training of RBM on X to obtain id done to have its weight matrix W. This matrix of weights is used between the two lower layers of the network.

3. X is transformed by RBM and a new data X' is formed .This is done either by doing sampling or by computation of the mean activation of the hidden units.

4. Repeating this procedure with X-> X' for the next pair of layers. This is done till the two top layers of network are reached.

5. Fine-tuning of all the parameters of this deep architecture with respect to a proxy for the DBN log- likelihood or with respect to a supervised training.



Structure of DBN Structure of RBM

**PARTICLE SWARM OPTIMIZATION**

Nature inspired algorithms try to imitate a phenomenon going on in nature. The taxonomy of the nature inspired algorithms can be done into two branches namely evolutionary algorithms and swarm optimization algorithms. Further evolutionary algorithms can be of two types that are genetic and differential evolutionary. Swarm optimization is also of two types namely particle based(Particle Swarm Optimization) and colony based (Ant Colony Optimization).In this paper , we describe PSO which is introduced by James Kennedy and Russel Eberhart in 1995[\cite{eberhart1995new}].Particle swarm optimization is used to optimized a non linear function .The basic concept of Particle swarm optimization is based on information sharing between the agents in the search space. The agents distributed over the function search for the optimal solution by sharing information amongst them and also based on their own experience. Initially all the swarm particles are initialized randomly over the space. Each particle calculates its fitness value, now on the bases of their personal best fitness value and the global best fitness value they update their position and they tend to move in the direction of the global best position until the stopping criterion is not satisfied. The stopping criterion can be either be if the step size taken by the particles from previous global best to the current global best< than a particular threshold or it can also be a given number of iterations for which a particle calculates it's fitness value and updates it's position. Swarm intelligence is the study of artificial and natural systems. In this system agents coordinate amongst each other. The focus is to study the behavior of agents that is an output from local communications of the agents with their surroundings and also with each other. Each particle is a E dimensional vector Xi and has a velocity Ui Some objective function f (fitness) is given. Pi is so far best position of particle I and and Pg is the maxi{Pi}. The evolution equation is given by

Ui(k+1)=wUi(k)+ciri(Pi-Xi(k))+ c2r2(Pg-Xi(k))

Xi(k+1)=Xi(k)+Ui(k+1)

with w lies in the range [0,1], c1>0, c2>0 being constants and r1, r2 being random numbers in [0,1]. Iteration terminates after some time or after f(Pg) reaches some value.

**BACKGROUND**

To fulfill our objective, we studied many literature related to heart disease classification , neural network and deep learning . Many previous works done on our current dataset are also reviewed. Typical deep learning architectures include DBNs , SAE [\cite{swietojanski2014convolutional}] ,CNNs [\cite{schmidhuber2015deep}] extract features from raw unlabeled physiological data.Here, we have listed few recent noteworthy contributions in this field of research. We have also reviewed many variants of NN and DBN. The basis of neural network was laid in the year 1940s by W. Pitts and McCulloch in [\cite{mcculloch1943logical}] McCulloch and Pitts employed logic and the mathematical notion of computation introduced by Alan Turing (1936-37) in terms of what came to be known as Turing Machines to explain how neural mechanisms might realize mental functions. Because of various characteristics of nervous activity, neural events and the relations among them can be treated by means of propositional logic. Multilayer perceptron is better than a single perceptron[ \cite{gardner1998artificial}] .The following paper [ \cite{saeedintelligent}], gives a high accuracy of 99.32\% accuracy with digit data using MLP Neural Network.As per literature and above discussion, for past decades several classification tools are available for medical dataset mclassification. Even then, ANNs are widely accepted and utilized to solve the real world classification problems in clinical applications. Artificial neural network perform better than other classifiers because they are capable of generalization, they are also capable of conditioning ,they have a low requirement of training points and they also converge very fast [\cite{sivanandam2007principles}]. Various learning or training algorithms for several NN architectures have been proposed in various problems in engineering science and technology and even in some parts of business industry and medicine. A few notable classification applications include biomedical medical diagnosis, handwritten digit recognition, pattern recognition, text categorization, information retrieval, and prediction of bankruptcy[\cite{huang2006extreme} and \cite{huang2008enhanced}] \cite{maind2014research} there are various advantages of ANN over conventional approaches. Due to its ability to classify the non-linear relationships, classification is difficult without ANN. Many different data sets are used with ANN like handwritten digit data set, iris data set etc. [\cite{sakshicahandwritten}] In this Hand digits are recognized using three approaches which are BPO, Single Layer Perceptron, HNN. Among the all methods BPA has been found most accurate as weights are to minimize errors with high accuracy. In case of error BPO can achieve a fast convergence and a satisfiable local minimum. The paper concluded with the fact that BPO can be explored more to improve its performance.Many improvement strategies are applied on ANN like using GPU instead of CPU to improve speed .[\cite{jagtap2014fast} ]This paper presents fast efficient ANN for handwritten digit recognition on GPU to reduce training time with PTM. BPO GPU based parallelization should be preferred generally with compared to CPU program. For smaller dataset, CPU is good but for larger datasets GPU is preferred. In this book ,[ \cite{shukla2010real}] Real Life Applications of Soft Computing robotic controller is formed using neuro-fuzzy approach.In neural network if we make dense network with many layers, computation becomes difficult so a new concept came into light "Deep Learning".This was first given by Geoffrey Hinton [\cite{lecun2015deep} ]Deep learning allows models that are composed of multiple layers to learn characteristics of data with multiple levels of hiding of information.Compared to non deep architectures deep learning architectures perform better in pattern and speech recognition. This paper [\cite{hinton2011deep}]describes deep belief nets as probabilistic generative models that are composed of multiple layers of stochastic latent variables (also called "feature detectors" or "hidden units"). \\ Many algorithms have been implemented on UCI heart data[\cite{uciheart}]set,[\cite{Lichman:2013}]Cleveland and Hungarian. [\cite {mangasarian2000active}]In this paper various algorithm like SVMLight , ASVM and CPLEX are used on Cleveland data which showed a maximum of accuracy for 85.56\%.With ANFIS an error of 15\% was obtained[\cite {ziasabounchi2014anfis}]. ANFIS which is used for classification which has both the advantages of neural network and fuzzy logic. [\cite{bhatia2008svm}] the proposed method here gives a classification accuracy of 90.57\% which is better than the previous results. SVM based decision support system for heart disease classification with integer-coded genetic algorithm is used to select critical features.[\cite{noor2009diagnosis}]This paper compares various other algorithms like MLP-ANN,K-NN,C4.5,Ripper with FDSS which is fuzzy decision support system on both hungarian and cleveland data sets. The knowledge is easy to understand because it is in the form of "IF-THEN" rules.The problem of having large number of rules is also solved using proposed RST based rule selection here. We also reviewed some nature inspired algorithm like PSO. [\cite{bai2010analysis}] Here PSO is compared with the other algorithms and it was found that it needs fewer parameters .The paper concludes with the fact that more research can be made on PSO by improving the topology of particle swarm, by blending with other intelligent optimization algorithm and by using PSO to optimize other algorithms [\cite{rini2011particle}] In this paper, a review on different methods of PSO algorithm is made. The various variants of PSO are discussed here which shows the advantages over basic PSO. [\cite{ribeiro2003hybrid}] This paper shows the optimization of ANN by PSO. PSO overcomes the limitations of ANN since it does not care about what it is optimizing. Any parameters of network can be optimized using PSO.[\cite{kennedy2011particle}] The relationships between PSO and both AI and GA are described.PSO is introduced and is used to optimized highly complex non-linear functions which cannot be classified by neural network. Another work [\cite{anooj2012clinical}] on UCI data set showed accuracy of 53.862 \% on cleveland data set using weighted fuzzy technique. In hungarian it gave an accuracy of 50.583% .PSO is chosen in our proposed strategy to optimize DBN's parameters because of the above mentioned reasons. Recent advances in the field of deep learning have been made and this is a very promising field of machine learning to work with. Also DBN optimized with nature inspire algorithm is a completely new way of improving efficiency of DBN and no proper work has been done in this area before.

**METHODOLOGY**

Keeping in view the recent literature, we propose a methodology for our work flow in order to achieve our objectives. The flowchart presented in figure below explains the work flow in a clear and lucid manner. The project aims to solve the optimization problem of DBN beginning by collecting appropriate datasets, implementing the algorithm and analysis of results obtained to deduce generalized observations. The flow diagram represents these steps in graphical form. The steps followed to test the proposed algorithm are:

Step 1: Collecting benchmark datasets which contains heart data.

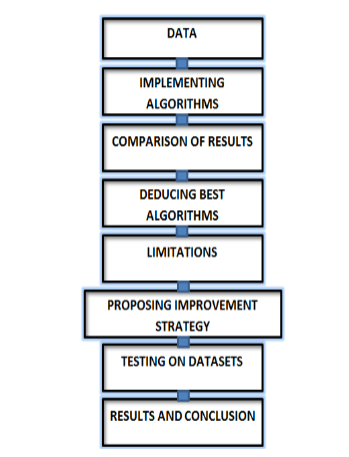
Step 2: Implementing DBN algorithm to check the presence of coronary disease.

Step 3: Examining the data and comparison of results obtained using statistical means and deducing the best algorithm.

Step 4: Finding out limitations of the best existing algorithm.

Step 5: Proposing and implementing an improvement strategy to overcome the limitations and optimizing the results .This improvement strategy includes our hybrid algorithm, DBN modified with PSO.

Step 6: Analysis of results, findings and documentation.



Flowchart representing methodology

The actual steps of algorithm followed in the proposed strategy are-

Step 1: The data collected is normalized and rescaled.

Step 2: This data is fitted into RBM layer of DBN.

Step 3: The RBM does unsupervised training on this data to pass it to the next layers on neural network.

Step 4: This transformed data is passed through N which does supervised training on data.

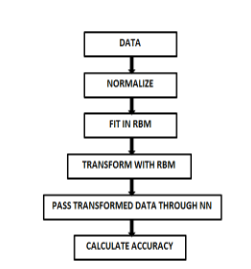
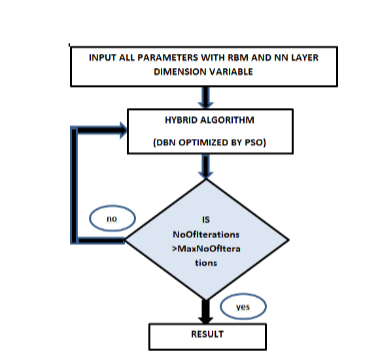
Step 5: All the above steps are put into a function and our aim to maximize the function output that is accuracy given by the function. This accuracy is the accuracy of our system related to how much efficient it is to predict heart disease.

Step 6: This optimization is done by PSO,. for that input all parameters of function with RBM and NN layer size kept variable.

Step 7: DBN is repeatedly called by PSO to optimize DBN function.

Step 8: The dimensions of RBM and NN which give highest accuracy is given as output by PSO.

Step 9: Print the accuracy and predict result.



Flowchart presenting hybrid algorithm Flowchart representing DBN

**DATASET DESCRIPTION**

The data sets used for testing the experiments are benchmark data sets taken from UCI machine learning repository[\cite{uciheart}]. The two datasets are Cleveland clinic foundation Database and Hungarian Database .The cleveland database consist of 303 records of patients. Each record has 14 attributes. The output is either 0,1,2,3 or 4. If the output is 1,2,3 or 4 it represents person is not in normal state and has some cardiac disease and 0 indicates that person is normal and has no disease. The Hungarian Database consist of 294 records. Each record has 14 attributes. The output of each attribute is either 0 or 1. 0 indicates the absence of any disease and 1 indicates the disease is present.

The attribute used and summary of number of records and their output is given in the tables shown below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Database | 0 | 1 | 2 | 3 | 4 | Total |
| Cleveland | 164 | 55 | 36 | 35 | 13 | 303 |
| Hungarian | 188 | 37 | 26 | 28 | 15 | 294 |

Classification of records of Databases

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Description** | **Value Description** |
| age | Age | Numerical |
| sex | Sex | 1 if male , 0 if female |
| Cp | Chest pain type | 1 to 4 |
| trestbps | Resting systolic blood pressure | Numerical |
| Chol | Serum cholesterol(mg/dl) | Numerical |
| fbs | Fasting blood sugar | 1 if yes and 0 if no |
| restecg | Resting cardio graphic results | 0 to 2 |
| thalach | Maximum heart rate achieved | Numerical |
| exang | Exercise induced angina | 1 if yes and 0 if no |
| oldpeak | ST depression induced by exercise relative to rest | Numerical |
| slope | The slope of the peak exercise ST segment | 1 unsloping and 2 flat |
| Ca | Number of major vessels | Numerical |
| Thal |  | 3,6 or 7 |
| Num | Output attribute | 0 or 1 |

Attribute and their value description

**RESULTS AND DISCUSSIONS**

Differences in accuracy are observed through introduction of particle swarm optimization in DBN. The optimized values of RBM layer size and neural network layer size is found out using proposed strategy. In Cleveland data set DBN alone gave accuracy of 88% and with PSO it gave accuracy of 91.18% .The dimensions which are obtained after optimization is 11 for RBM layer and 42 for NN layer. In Hungarian data simple DBN gave an accuracy of 90.32% and modified algorithm gave accuracy of 94.88%.The dimensions obtained after optimization by PSO are 6 for RBM layer and 27 for NN layer. We therefore realize that DBN can be further optimized by various nature inspired algorithm to improve accuracy. The table showing the comparison of various algorithms with proposed strategy is given below.

|  |  |
| --- | --- |
| Algorithm | Accuracy(in %) |
| ASVM | 87.24 |
| SVMlight | 87.50 |
| Neural Network | 53.86 |
| FDSS | 83 |
| K-NN | 81 |
| C4.5 | 82 |
| Ripper | 83 |
| Weighted Fuzzy | 57.85 |
| Fuzzy | 51.793 |
| DEEP BELIEF NETWORK | 88 |
| DBN+PSO( Proposed strategy) | 91.18 |

Comparison among various existing algorithms and our hybrid algorithm on Cleveland data set

|  |  |
| --- | --- |
| Algorithm | Accuracy(in %) |
| K-NN | 92 |
| FDSS | 84 |
| MLP-ANN | 54 |
| C4.5 | 66 |
| Ripper | 68 |
| Weighted Fuzzy | 50.83 |
| Neural Network | 46.417 |
| Fuzzy | 36 |
| Deep belief network | 91.33 |
| DBN+PSO(proposed strategy) | 94.33 |

Comparison among various existing algorithms and our hybrid algorithm on Hungarian data set

**CONCLUSION and FUTURE WORK**

Here, neural network and its variants are studied to make a classifier for predicting heart diseases. Seeing the limitations of neural network, the concept of deep learning is introduced. We know that DBN is composed of RBM layers and neural network layers and has a large number of parameters to be made optimized for good performance. Optimizing such a large number of parameters makes the performance worse. So in order to decrease the number of parameters to be optimized, Particle swarm optimization is used to optimize the parameters of deep belief networks. These parameters are the dimensions of RBM and NN layer to be used for unsupervised and supervised learning respectively. The proposed strategy of hybrid DBN algorithm with PSO successfully classified the heart disease type from the input data set provided. It predicted in a very reliable way whether a person has heart disease or not. We don't need to rely much on traditional techniques alone. Machine learning algorithms like DBN can be used to make this diagnosis fast and less costly. This will help in detecting the symptoms of coronary disease before the patient goes for actual verification that to check whether he has a coronary disease. This type of strategy will surely reduce the number of deaths due to heart diseases in developed and as well as developing countries. In this paper, a more optimized version of Deep belief network is presented with PSO. This work can be extended by using other nature inspired algorithms like firefly, Ant colony Optimization (ACO), Flower pollination. More accurate results can be obtained by using a larger and more balanced data set. The system trained with the help of balanced data will give more accurate results. Also the system can be made for real time data. This will then take data in real time as well as predict or classify in real time whether a person has the disease or not. Larger datasets can be managed using Hadoop. Big data concepts can be applied in that case. Also it is quite practical that in real time data will be large so storing and processing of data using hadoop may become necessary.

**REFERENCES**

[1] Anooj, P.: 2012, Clinical decision support system: risk level prediction

of heart disease using decision tree fuzzy rules, Int J Res Rev Comput Sci

3(3), 1659–1667.

[2] Bai, Q.: 2010, Analysis of particle swarm optimization algorithm, Computer

and information science 3(1), 180.

[3] Bengio, Y.: 2009, Learning deep architectures for ai, Foundations and

trends R in Machine Learning 2(1), 1–127.

[4] Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H. et al.: 2007, Greedy

layer-wise training of deep networks, Advances in neural information processing

systems 19, 153.

[5] Bhatia, S., Prakash, P. and Pillai, G.: 2008, Svm based decision support system

for heart disease classification with integer-coded genetic algorithm to

select critical features, Proceedings of the World Congress on Engineering

and Computer Science, WCECS, pp. 22–24.

[6] Carbonell, J. G., Michalski, R. S. and Mitchell, T. M.: 1983, An overview

of machine learning, Machine learning, Springer, pp. 3–23.

[7] Eberhart, R. C., Kennedy, J. et al.: 1995, A new optimizer using particle

swarm theory, Proceedings of the sixth international symposium on micro

machine and human science, Vol. 1, New York, NY, pp. 39–43.

[8] Gardner, M. W. and Dorling, S.: 1998, Artificial neural networks (the multilayer

perceptron)â˘A

ˇTa review of applications in the atmospheric sciences,

Atmospheric environment 32(14), 2627–2636.

[9] Hagan, M. T., Demuth, H. B., Beale, M. H. and De Jesús, O.: 1996, Neural

network design, Vol. 20, PWS publishing company Boston.

[10] heart data set, U.: uci, Uci heart data set.

[11] Hinton, G.: 2011, Deep belief nets, Encyclopedia of Machine Learning,

Springer, pp. 267–269.

[12] Hinton, G. E., Osindero, S. and Teh, Y.-W.: 2006, A fast learning algorithm

for deep belief nets, Neural computation 18(7), 1527–1554.

[13] Huang, G.-B. and Chen, L.: 2008, Enhanced random search based incremental

extreme learning machine, Neurocomputing 71(16), 3460–3468.

[14] Huang, G.-B., Zhu, Q.-Y. and Siew, C.-K.: 2006, Extreme learning machine:

theory and applications, Neurocomputing 70(1), 489–501.

[15] Jagtap, V. N. and Mishra, S. K.: 2014, Fast e\_cient artificial neural network

for handwritten digit recognition, International Journal of Computer

Science and Information Technologies 5(2), 2302–2306.

[16] Kala, R., Vazirani, H., Shukla, A. and Tiwari, R.: 2010, O\_ine handwriting

recognition using genetic algorithm, arXiv preprint arXiv:1004.3257 .

[17] Kennedy, J.: 2011, Particle swarm optimization, Encyclopedia of machine

learning, Springer, pp. 760–766.

[18] LeCun, Y., Bengio, Y. and Hinton, G.: 2015, Deep learning, Nature

521(7553), 436–444.

[19] Lichman, M.: 2013, UCI machine learning repository.

URL: http://archive.ics.uci.edu/ml

[20] Maind, S. B. and Wankar, P.: 2014, Research paper on basic of artificial

neural network, International Journal on Recent and Innovation Trends in

Computing and Communication 2(1), 96–100.

[21] Mangasarian, O. L. and Musicant, D. R.: 2000, Active support vector machine

classification, NIPS, pp. 577–583.

[22] McCulloch,W. S. and Pitts,W.: 1943, A logical calculus of the ideas immanent

in nervous activity, The bulletin of mathematical biophysics 5(4), 115–

133.

[23] Michie, D., Spiegelhalter, D. J. and Taylor, C. C.: 1994, Machine learning,

neural and statistical classification.

[24] Noor Akhmad, S., Venkatachalam, P. and Ahmad Fadzil, M. H.: 2009, Diagnosis

of coronary artery disease using artificial intelligence based decision

support system

[25] Ribeiro, P. F. and Schlansker, W. K.: 2003, A hybrid particle swarm

and neural network approach for reactive power control, search. cpan.

org/src/KYLESCH/AI-PSO-0.81/ReactivePower-PSO-wks. pdf .

[26] Riley, L. and Cowan, M.: 2014, Noncommunicable diseases country profiles

2014, Geneva: World Health Organization .

[27] Rini, D. P., Shamsuddin, S. M. and Yuhaniz, S. S.: 2011, Particle swarm

optimization: technique, system and challenges, International Journal of

Computer Applications 14(1), 19–26.

[28] Rosenblatt, F.: 1958, The perceptron: a probabilistic model for information

storage and organization in the brain., Psychological review 65(6), 386.

[29] Saeed, A.-M.: n.d., Intelligent handwritten digit recognition using artificial

neural network.

[30] Sakshica, D. and Gupta, K.: n.d., Handwritten digit recognition using various

neural network approaches.

[31] Schmidhuber, J.: 2015, Deep learning in neural networks: An overview,

Neural Networks 61, 85–117.

[32] Shi, X. H., Liang, Y. C., Lee, H. P., Lu, C. and Wang, Q.: 2007, Particle

swarm optimization-based algorithms for tsp and generalized tsp, Information

Processing Letters 103(5), 169–176.

[33] Shukla, A., Tiwari, R. and Kala, R.: 2010, Real life applications of soft

computing, CRC Press.

[34] Sivanandam, S. and Deepa, S.: 2007, PRINCIPLES OF SOFT COMPUTING

(With CD), John Wiley & Sons.

[35] Soni, J., Ansari, U., Sharma, D. and Soni, S.: 2011, Predictive data mining

for medical diagnosis: An overview of heart disease prediction, International

Journal of Computer Applications 17(8), 43–48.

[36] Swietojanski, P., Ghoshal, A. and Renals, S.: 2014, Convolutional neural

networks for distant speech recognition, IEEE Signal Processing Letters

21(9), 1120–1124.

[37] Yusuf, S., Reddy, S., Ôunpuu, S. and Anand, S.: 2001, Global burden of cardiovascular

diseases part i: general considerations, the epidemiologic transition,

risk factors, and impact of urbanization, Circulation 104(22), 2746–

2753.

[38] Ziasabounchi, N. and Askerzade, I.: 2014, Anfis based classification model

for heart disease prediction, International Journal Of Electrical & Computer

Sciences IJECS-IJENS 14(02), 7–12.