International Journal of Latest Trends in Engineering and Technology
Vol.(10)Issue(1), pp.337-343
DOI: http://dx.doi.org/10.21172/1.101.61
e-ISSN:2278-621X

A STUDY OF NEURAL NETWORK BASED THYROID DISEASE PREDICTION SYSTEM

Ifrah Raoof[1], Arvind Selwal[2]

Abstract- Thyroid organ is a butterfly-molded organ which is present in the neck underneath the mouth of human body. It release hormones that control metabolism like heart rate, body temperature etc. Thyroid disease is one of the common lifestyle disease found in human beings. Many lifestyle risk factors including eating habits, pregnancy, antibodies also amount to thyroid disease in human beings. In the recent past, the research communities have considered only three main factors i.e.T3, T4 and TSH for diagnosis of thyroid disease as input parameters. In this paper different prediction systems have been presented for diagnosis of thyroid disease. The studies clearly show that more factors like different types of anti bodies, heredity etc can be considered for accurate and automatic diagnosis of thyroid disease. Huge amount of patients information is present in clinical data repository. Each and every type of information is important for diagnosis of different diseases. Artificial Neural Networks can be used significantly for the diagnosis of the thyroid disease. Various pattern classifiers are present, we can combine different classifiers so as to increase the accuracy the prediction system of thyroid disease.
Keywords – Pattern classifier, multi-layer perceptron, thyroid

1. INTRODUCTION

1.1. Artificial Neural Netwotk

These days, with the advancement of technology Artificial Neural Networks (ANNS), which is a branch of Artificial Intelligence, has evolved. Neural Network (NN) is a data processing system, indeed a self learning system/model. The neural system comprises of substantial number of handling components called neuron. Various Neurons work concurrently to perform processing on the provided inputs. The first neural model was developed by Mc Culloth Pitts. As NN is a self learning system, it exhibit different categories of learning algorithms such as Supervised learning, Unsupervised learning and Reinforcement learning. NN has the ability to learn from the input in the same way as human brain learn from past experiences. It has the capability to acquire meaning from imprecise data so as to detect trends, pattern that are highly complicated for humans to detect. NN are wide used in the real - world computation applications. The various areas of application includes pattern recognition, pattern classification and pattern prediction. The paradigm is shifting from conventional method of health disease prediction to automatic computer based expert system.ANN may be used to correctly predict lifestyle human diseases like Diabetes, ANN cardio problems and thyroid.

1.2. Thyroid

Thyroid organ is a standout amongst the most vital organs in human body. It is considered to be among the biggest endocrine organ present in human body. It is present in the neck underneath the mouth in human body. It discharges hormones that control all body digestion including Heart Rate, Breathing, Body weight ,Body Temperature. The thyroid gland consumes Iodine which is present in the food that we eat to secrete two main hormones i.e. Triiodothyroxine (T3) and Thyroxin (T4). It is important that the levels of T3 and T4 are balanced. There are two more glands called Hypothalamus and pituitary gland which communicate with each other to maintain T3 and T4 levels in blood. The TSH is discharged by Hypothalamus that signals the Pituitary to tell the Thyroid organ to secrete either T3 and T4 either in low or high quantity depending on level of the hormone called TSH. At the point when the levels of T4 and T3 hormones are low in the blood, the other gland called Pituitary gland secretes more TSH to inform the thyroid gland to deliver more thyroid hormone .If the levels of T3 and T4 hormones are high, the pituitary organ discharges less TSH to inform the thyroid organ to lower the secretion of these hormones. To diagnose disease with high accuracy, with reduced human errors and to avoid misdiagnosis, neural networks are used.

[1] Department of Computer Science & IT, Central university of Jammu, Jammu-181143
[2] Department of Computer Science & IT, Central university of Jammu, Jammu-181143

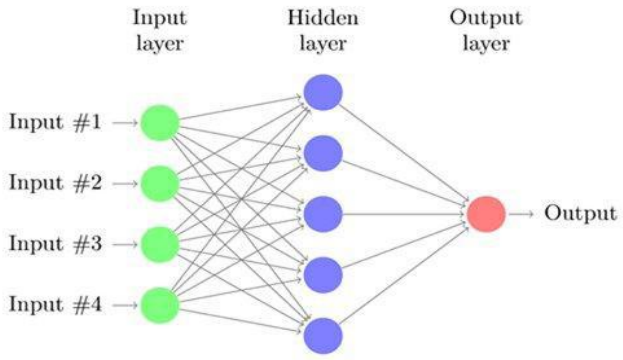

Fig 1. Architecture of Artificial Neural Network

Imagesource:www.texample.net

1.3. Types Of Neural Networks

Feed forward artificial neural network

This is the most straight forward neural system, the feed forward neural network process data in one direction only. Information moves from the input neurons to the yield neurons. These neurons have no loop in their system .

Radial Basis function neural network

The RBF neural system is an exceptionally natural neural system. Every node in the RBF neural system stores a case from the preparation set as a "model". This system is engaged with the linearity which provide RBF the benefit of not experiencing nearby minima.

Kohenen self organizing neural network

It was Invented by Teuvo Kohonen, this neural network is perfect for dimension reduction purpose. It differs from other neural networks ,the way that it uses competitive learning on a set of input data.

Recurrent neural network

The recurrent neural system is a neural system that takes into account a bi-directional stream of information i.e. information spill out of both the sides. The intermittent neural system is equipped for utilizing its interior memory to process arbitrary sequence of inputs. This neural network can be a widely used for many tasks such as handwriting and speech recognition etc.

Learning in Artificial Neural Networks

Artificial Neural Networks maybe trained by using a variety of learning techniques. In general, the learning in ANN may take place in the presence or absence of a teacher. Broadly, ANN maybe trained with some parameters in four ways, which are explained below:

Supervised Learning

In this algorithm calculated desired output for the system is additionally given to the system while the system is being trained. By giving the neural system both an input and desired match the error can be figured out by ascertaining the difference between the objective yield and genuine yield. It would then be able to utilize that error to make changes to the system by updating it's weights.

Unsupervised Learning

In this algorithm a set of inputs is given to the neural network and it is the responsibility of neural network to recognize some kind of pattern present in the inputs . This sort of learning is used worldwide as a part of information mining.

Reinforcement Learning

This learning algorithm is similar to supervised learning in the manner that a feedback is given back ,and at the end on the basis of how well the system performed ,a reward is given.

Competitive Learning

The competitive learning is an algorithm in which weights of a given layer are adjusted in such a manner that only a single neuron in a layer, for example the yield layer, fires. Input patterns can be classified into a discrete set of output classes with competitive learning.

2. BACKGROUND

Pattern classifiers have been used by various researchers for developing lifestyle disease prediction systems. In this section a brief study of various disease detection system is presented.

Narender Kumar et.al.(2017) have compared various data classifications techniques and their accuracy performance to predict chronic kidney disease. The various classifiers used in this paper are J48, Naive bayes, Random forest, SUM & KNN. These classifiers are compared with the help of performance measures like ROC, kappa statistics, RMSE & MAE using WEKA tool. After experimentation it has been indentified that random forest classification techniques has higher performance in terms of accuracy and prediction as compared to other techniques[1]

Xing et.al.(2017) have proposed a technique which is concerned with the aim to develop a data mining algorithm to predict survival of CHP patients (Coronary Heart Disease). In this work, three algorithm's were used to develop these prediction models. These prediction models were then compared to each other on the basis of performance. The three models are SVM, neural networks(ANN) , decision trees. Out of these three, SVM has been the best with 92.10% of accuracy, ANN has the second with 91.10% of accuracy and Decision tree has the last with the performance accuracy of 89.6%. In SVM, Sequential minimal optimization algorithm is used to train it. ANN, multilayer perceptron with back propagation is being used. In decision tree algorithm, C5 algorithm was used[2].

Hsiang et.al.(2006) have Compared Expert Judgment and Automatic Approaches in this paper, the authors have compared two different features selection techniques to extract

features from a given data set. They used two common approaches i.e. automatic feature selection mechanism which is data driven and expert judgment which is Knowledge based. The results suggests that the automatic feature selection approach improve the prediction capability of a classifier while as the domain expert improves the sensitively of a classifier. The union of two approaches can remove irrelevant data, which improves classification accuracy[3].

Shivanee panday et.al.(2016) have proposed a paper in which various data mining techniques which include Bayes network, RBF network, CART,L4.5, REP tree, decision tree, multilayer perception are used to develop pattern classifiers for diagnosis of hypothyroid disease. The experiment result shows that REP tree and L4.5 techniques perform well as compared to others[4].

Rajeeb dev et.al.(2008) in their paper have proposed a binary classification problem for the diagnosis of diabetes . A person suffering from diabetes fall in class 1 and non diabetic fall in class 2.They have used backpropogation algorithm in Multilayer feed forward . In this the authors have used single as well as multi layer perceptron. Both the neural networks have six input nodes and one output node. The network successfully classified patients into diabetic and non diabetic with performance of 92.50%[5].

Canan et.al.(2009) proposed a hybrid structure of neural network and fuzzy logic. CSFNN (conic section function neural networks) unifies the multilayer perception's propagation rules and radial basic function network. The experiment shows that the hybrid schemes have better results over the non hybrid structures. Among the hybrid schemes, fuzzy - CSFNN shows best accuracy of 92.30% in breast cancer diagnosis as well as in case of thyroid diagnosis[6].

Anupam Shukla et.al. (2009) in their work have trained the system using three ANN algorithms, the backpropogation (BPA), the radial basis function (RBF) and the learning vector quantization (LVQ) systems . The networks are compared on the basis of the factors like accuracy of network and training time. On comparison of the performance LVQ network is found to be have the best accuracy with 98% but among the three, RBFN when trained on hybrid dataset have the least training time [7].

Shradha Deshmukh et.al.(2017) proposed two important classification algorithms namely fuzzy min-max and pruning fuzzy min-max algorithms are used. Pruning along with fuzzy min-max reduces the complicity neural network. Different experiments were carried out and the results show the pruned FMM is best as compared to FMM[8].

K Vishvant et.al.(2014) mentions the kidney abnormalities like stones, cysts, cancerous cells etc. Therefore it is very important to identify the accurate location of stone present in the kidney. It can be done manually by analyzing ultra sound image but the image is less speckle noise. Thus the ultra sound undergoes preprocessing in which image restoration is first done, then it undergoes Gabor filtering for smoothening. Multilayer Perceptron and Back Propagation ANN are used to distinguish the type of the stone. The multilayer perceptron with backpropogation gives high accuracy of 98% when contrasted with Naive Bayes grouping [9].

Muthuselvan et.al (2013) focuses on implementing five different types of data mining techniques using a data mining tool called WEKA in order to predict breast cancer from blood data sets. The five algorithms include Naive Bayes, one R, Zero R ,Random tree algorithm and j48. On comparing the performance between the different algorithms show that J48 algorithm performance was highest i.e. 86.36% while as minimum (Zero R) is 56.81%[10].

Madhuri et.al.(2013) have proposed a computer aided artificial intelligence system which is used for diagnosis of stress. The parameters to be considered are galvanic skin response , Heart rate, Body temperature, Muscle tension. This system uses fuzzy logic which combines human like reasoning and connected structure of fuzzy system[11].

De Araujo et.al.(2017) proposed a classical method for induction motors fault diagnosis do not always provide satisfactory results. This paper proposes a hybrid system that uses data obtained from vibration, and current sensors to predict failures at an early stage. The input to the system is based on fuzzy logic is given by processing the signals in the frequency and time domain through short time Fourier transform and multi resolution analysis. This technique allows an increase in reliability in the detection and diagnosis in the level of severity as compared to existing techniques[12]

Dr.N Ganesan et.al.(2010) used neural networks in the medical field for preclinical study. In this paper the author has shown the various ways by which neural networks can be applied on clinical data for the diagnosis of lung cancer [13].
Trofimova et.al.(2017) have used neural networks to estimate and evaluate trust in ad-hoc networks. The main objective of this paper is to confirm applicability of NNs to trust management in ad-hoc networks and to construct a method to detect un-trusted nodes and to estimate the value of trust. The model used here is multilayer preceptor (MLP). In this paper, a simulator of ad-hoc network which generates data sets is used. These data sets are used to validate and train NN quality on different data. The experiments clearly show that NNs can be effectively used for solving problem of detection of un-trusted nodes[14].
Sapna (2016) has proposed fusion of big data and neural networks for predicting thyroid. Clinical information is huge in volume, thus conventional data processing applications won't be sufficient to interpret big data, hence it needs innovations techniques to handle and extract important information from it. A feed forward NN is used in the paper[15].
Saeed Shariati and Mahdi Motanali Haghighi (2010) have used fuzzy system to diagnosis hepatitis and thyroid disease. The results of fuzzy neural networks with support vector machine and artificial neural network were compared [16].
I.MD.Dendi et.al.(2015) have used MLP and Back propagation to classify the type of thyroid. After experimenting, the performance of multilayer perception was
found highest as compared to that with Back propagation (69.77%)[17].
Mohd Reza et.al.(2017) have discussed the diagnose of different types of thyroid disease using neural networks by considering their age. The inputs for the thyroid used are seven hormone tests including age and the output is the diagnosis of thyroid. The various network structure used include MLP, PNN, and GRNN AND CFNN. The result is that neural network can diagnose the thyroid with 100% accuracy[18].
Shigeo Abe et.al.(2017) proposed a fuzzy classifier to diagnosis problems. In this paper, they initially expect that no information has a place in abnormal class, at that point they prepare the fuzzy system. Then they set a threshold which is calculated of minimum weighted distance from the centre of clusters for the data belonging to normal class and for any unknown data if it is inside the limit, they characterized it into normal class and if not then abnormal class [19].
Fengying Xie et.al. (2017) have built up a novel technique for detecting tumor as amiable or threatening by analyzing images. In this paper, they have designed an ensemble classifier that combines back propagation neural network with fuzzy neural network[20].
Mazin Abdul Rasool Hameed et.al(2017) have proposed a method of classifying thyroid disease using multilayer feed forward using back propagation learning rule .In this work three inputs have been considered as T3,T4,TSH[21] .

3. COMPARATIVE RESULT

Table I: Summary of pattern classifiers for disease prediction

Authors	Year	Problem	Network used	Performance Parameters
Narender Kumar Sabia Khatri [1]	2017	Implementing WEKA tool to predict disease prediction	KNN Random forest Naïve Bayes T48 SVM	Accuracy=RF (100%)
Xing, Yanwei Wang, Jie Zhao, Zhihong Gao, AndYonghong [2]	2007	Prediction of coronary heart disease	SVM Neural Network Decision Tree	Performance=SVM (92.1%)
Hsiang, Tsang Wei, Chih Ping Tseng, VChengincent S. [3]	2006	Feature selection for data mining	Automatic feature selection Expert Judgemnet	--
Shivanee panday Rohit S R Tandan [4]	2016	Parkinson's disease Diagnosis and classification	Bays net Multilayer feed forward perceptron RBF C4.5 CART Decision tree	Accuracy= REP (99.57%) C4.5 (99.57%) Bays net (98.59%)

Rajeeb Dev Vaibav Bajpai Gagan Gandhi Barvali Dev [5]	2008	ANN technique for diagnosing diabetics	Multilayer feed forward network with back propagation	Performance =92.5%
Canan Senol Tulay Yildirim [6]	2009	Thyroid and breast cancer disease diagnosis using Fuzzy-neural networks	Fuzzy-CSFNN(Hybrid structure)	Accuracy =92.3%
Tiwari, Ritu Shukla, Anupam Kaur, Prabhdeep [7]	2009	Diagnosis of thyroid	Feed forward neural network with propagation Feed forward radial basis Feed forward learning vector quantization	Accuracy =LVQ (98%)
Shradha Deshmukh Swati shinde [8]	2017	lung cancer diagnosis using pruned fuzzy min max neural network	fuzzy min max algo Prunning Fuzzy min max algo	Prummed fuzzy min max algo showed better performance
K Vishwant R Gunasundari [9]	2014	Kidney stone detection	Multi layer perceptron with back propogation	Accuracy performance=98%
Muthuselvan, S Sundaram, K. Soma Prabasheela [10]	2016	Prediction of breast cancer	Bays Zero R One R J48 Random Tree	Performance = j48(86.36%)
Madhuri, V.J. Mohan, Madhumitha R. Kaavya, R. [11]	2013	Stress management using ANN	Fuzzy logic	--
De Araujo Cruz, A.A.G. Gomes, R.D. Belo, F.A. Lima Filho, A.C. [12]	2017	To detect failure in induction motor	Fuzzy logic	--
Ganesan Venkatesh M.A Rama A.malarathi [13]	2010	Diagnosis cancer disease	Neural network with back propagation	--
Trofimova, Yelena Moucha, Alexandru Mihnea Tvrdik, Pavel [14]	2017	Application of neural networks for decision making and evaluation of trust in ad-hoc networks	Multilayer neural network	--
Sapna [15]	2016	Fusion of big data and neural network for predicting thyroid:-	Feed forward neural network	--
Saeed shariati Mahdi motavali haghighi [16]	2010	Diagnosis hepatitis and thyroid	Self organized fuzzy system	--
I. D. Maysanjaya et.al. [17]	2015	Diagnosis of thyroid disease using different classification methods	Mlp Back propogation	Accuracy performance=MLP (96.7%)

MOHD. Reza Ali Rafiee Omid Mahdiyar [18]	2017	Diagnosis of thyroid disease	MLP PNN GRNN FTDNN CFNN	Accuracy=GRNN (99.5%)
Shigeo Abe Ruck Thomson Masahiro Kayama [19]	2017	Fuzzy classifier for diagnosis problem	Self generating neural network (backpropogation with fuzzy neural network)	--
Fengying Haidi fan Yang li Zhigiuo Jiang Rusong Alan bonik [20]	2017	Melanonna classification	Self generating neural network(back propogation neural network with fuzzy neural network)	--

4. RESEARCH CHALLENGES

The analysis of presented techniques in section 2, clearly indicates that lifestyle diseases maybe classified using different pattern classifiers like RBF, MLP etc. Moreover, various pattern classifiers may be combined in order to further improve the accuracy of diagnosis of thyroid disease. Blood test is the most crudest and poorest method of determining whether a patient is suffering from thyroid or not, so the best solution is to consider as many parameters as possible. It has also been noticed that most of the researchers have used only T3, T4 and TSH hormones for diagnosis of thyroid disease. In order to avoid incorrect diagnosis of thyroid disease, more lifestyle risk factors can be considered. The study also reveals that most of the researchers have measured the performance of different pattern classifiers in terms of accuracy, time taken to diagnose disease and number of tests required for diagnosis. Dimension reduction technique may be implemented on the data sample to reduce the time required for diagnosis of thyroid disease. Further it is important to take into consideration all the sub classes of the thyroid disease for the accurate diagnosis.

5. CONCLUSION

Thyroid gland is a gland having butterfly shape and is present in the neck underneath the mouth of human body It release hormones that control metabolism. Diagnosis of lifestyle disease is a key challenging task. Thyroid disease is found to be one of the common lifestyle disease in human beings. The accurate diagnosis of thyroid disease is an important research challenges, and it is very important to take into consideration as many risk factors as possible. Thyroid disease can be classified into two broad categories i.e. hypothyroid and hyperthyroid. Many lifestyle risk factors including eating habits, pregnancy, antibodies causes thyroid disease. Most of the authors have considered only three main factors i.e.T3,T4 and TSH for diagnosis of thyroid disease as input parameters. Artificial Neural Networks have the capability of adaptive learning and they can be used for the diagnosis of thyroid disease. As a future work, the efficiency of different pattern classifiers may be further enhanced.

6. ACKNOWLEDGMENT

We are thankful to the authorities of the Central University of Jammu for providing necessary platform and all the facilities for carrying out this work.

7. REFERENCES

[1] [1] N. Kumar and S. Khatri, “Implementing WEKA for medical data classification and early disease prediction,” 2017 3rd Int. Conf. Comput. Intell. Commun. Technol., pp. 1–6, 2017.

[2] [2] Y. Xing, J. Wang, Z. Zhao, and Yonghong Gao, “Combination Data Mining Methods with New Medical Data to Predicting Outcome of Coronary Heart Disease,” 2007 Int. Conf. Converg. Inf. Technol. (ICCIT 2007), pp. 868–872, 2007.

[3] [3] T. Hsiang, C. P. Wei, and Vc. S. Tseng, “Feature selection for medical data mining: Comparisons of expert judgment and automatic approaches,” Proc. - IEEE Symp. Comput. Med. Syst., vol. 2006, pp. 165–170, 2006.

[4] [4] P. Durga, V. S. Jebakumari, and D. Shanthi, “Diagnosis and Classification of Parkinsons Disease Using Data Mining Techniques,” ISSNOnline) Int. J. Adv. Res. Trends Eng. Technol., vol. 3, no. 14, pp. 2394–3777, 2016.

[5] [5] R. Dey and V. Bajpai, “Application of Artificial Neural Network (ANN) technique for Diagnosing Diabetes Mellitus,” IEEEexplorer, no. 155, pp. 8–11, 2008.

[6] [6] Canan Senol and Tülay Yildirim, “Thyroid and breast cancer disease diagnosis using Fuzzy-neural networks,” ELECO 2009 - 6th Int. Conf. Electr. Electron. Eng., pp. 390–393, 2009.

[7] [7] R. Tiwari, A. Shukla, and P. Kaur, “Diagnosis of Thyroid Dicial Neural Networks,” no. March, pp. 6–7, 2009.

[8] [8] S. Deshmukh and S. Shinde, “Diagnosis of Lung Cancer using Pruned Fuzzy Min-Max Neural Network,” Int. Conf. Autom. Control Dyn. Optim. Tech. ICACDOT 2016, pp. 398–402, 2017.

[9] [9] K. Viswanath and R. Gunasundari, “Design and analysis performance of kidney stone detection from ultrasound image by level set segmentation and ANN classification,” Proc. 2014 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2014, pp. 407–414, 2014.
[10] [10] S. Muthuselvan, K. S. Sundaram, and Prabasheela, “Prediction of breast cancer using classification rule mining techniques in blood test datasets,” 2016 Int. Conf. Inf. Commun. Embed. Syst., no. Icices, pp. 1–5, 2016.
[11] [11] V. J. Madhuri, M. R. Mohan, and R. Kaavya, “Stress Management Using Artificial Intelligence,” 2013 Third Int. Conf. Adv. Comput. Commun., pp. 54–57, 2013.
[12] [12] A. A. G. De Araujo Cruz, R. D. Gomes, F. A. Belo, and A. C. Lima Filho, “A Hybrid System Based on Fuzzy Logic to Failure Diagnosis in Induction Motors,” IEEE Lat. Am. Trans., vol. 15, no. 8, pp. 1480–1489, 2017.
[13] [13] N. Ganesan, “Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data,” Int. J. Comput. Appl. (0975, vol. 1, no. 26, pp. 76–85, 2010.
[14] [14] Y. Trofimova, A. M. Moucha, and P. Tvrdik, “Application of neural networks for decision making and evaluation of trust in ad-hoc networks,” 2017 13th Int. Wirel. Commun. Mob. Comput. Conf. IWCMC 2017, pp. 371–377, 2017.
[15] [15] S. Sapna, “Fusion of Big Data and Neural Networks for Predicting Thyroid,” 2016 Int. Conf. Electr. Electron. Commun. Comput. Optim. Tech., pp. 243–247, 2016.
[16] [16] S. Shariati and M. M. Haghighi, “Comparison of anfis neural network with several other anns and support vector machine for diagnosing hepatitis and thyroid diseases,” 2010 Int. Conf. Comput. Inf. Syst. Ind. Manag. Appl. CISIM 2010, pp. 596–599, 2010.
[17] [17] I. D. Maysanjaya, H. A. Nugroho, N. A. Setiawan, J. G. No, and K. Ugm, “A Comparison of Classification Methods on Diagnosis of Thyroid Diseases,” 2015 Int. Semin. Intell. Technol. Its Appl., pp. 89–94, 2015.
[18] [18] M. R. Obeidavi, A. L. I. Rafiee, and O. Mahdiyar, “Diagnosing Thyroid Disease by Neural Networks,” vol. 10, no. 2, pp. 509–524, 2017.
[19] [19] S. Abe and R. Thawonmas, “A fuzzy classifier with ellipsoidal regions,” IEEE Trans. Fuzzy Syst., vol. 5, no. 3, pp. 358–368, 1997.
[20] [20] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. Bovik, “Melanoma classification on dermoscopy images using a neural network ensemble model,” IEEE Trans. Med. Imaging, vol. 36, no. 3, pp. 849–858, 2017