A Literature Review on Diagnosing Thyroid Disease Through Artificial Neural Network Techniques

1Vinod Kumar Pal, 2V.P. Sriram, 1Rashmi Mahajan and 1Suresh Chandra Padhy
1Balaji Institute of International Business (BIIB), Pune, India
2Acharya Bangalore Business School (ABBS), Bengaluru, India

Abstract: In this study, researcher majorly focuses on usefulness and suitability of artificial neural network for prediction of thyroid disease and respective role of neural network in medical diagnosis by past recorded data is identified from a comprehensive literature review. Wherein, research contributions from 2013-2017 are reviewed. It is found that different various architecture of Artificial Neural Network (ANN), Back Propagation Network (BPN), Hybrid Back-Propagation Neural Network (BPNN), Data Mining Methods (DMM), Decision Support System (DSS), Ranked Improved F-score Ordering (RIFO), Auto Associative Neural Network (AANN), Multi-Layer Perceptron (MLP), Multivariate Bayesian Prediction Method (MBPM), Machine Learning Methods (MLM) and Radical Basis Function Neural Network (RBFNN) are found to be proper and appropriately suitable. In recent years, these architectures and methods are also found useful for prediction of so many other diseases. The discussions of these architectures and their suitability, appropriateness for thyroid disease prediction is presented through this review article and it would reveal the various methods and its benefits of diagnosing thyroid diseases in a most appropriate and essential manner based on the nature and cause it affects on human body. This review paper will certainly help various researchers to go through various research insights on diagnosing thyroid diseases on a single point of contact and it help them to focus and choose the appropriate method for diagnosing thyroid diseases.

Key words: Thyroid disease, artificial neural network, appropriate, method, various, prediction

INTRODUCTION

The current applications of Artificial Neural Network (ANN) used much architecture in diagnosing the thyroid disease and it is helpful to predicting better accuracy in diagnosing the thyroid disease. Artificial Neural Network (ANN) system is basically defined as back propagation learning method where elements are N be a neural network with e connections, m inputs and n outputs. The X_i-X_j will denote the vectors in R_m, Y_i-Y_j vectors in R_n and W_i-W_j are vectors in R_e. These are called inputs, output and weights, respectively. The number of output y can be controlled as nonlinear dynamic system where in value of y works as a predictor. In other words, we can say y is dynamic and on values on X_i-X_j. Neural networks and its training system algorithm are basically used to define such system by adjustment of weights W_i-W_j during training process. After training process each iteration the neural network may minimize the error between actual ‘Y’ and its predicted value altered by W_i-W_j are called training process new rate of X_i-X_j, may workable be given predicated value of ‘y’ is testing. We observed that the back propagation neural network is the basic method of prediction of thyroid disease in medical diagnosis. Thus, in this study back propagation neural network and other architectures is applied for prediction of thyroid disease.

Literature review on various techniques for diagnosing thyroid diseases: Hamed (2017) has performed a research study on “Artificial neural network system for thyroid diagnosis”. The proposed paper focuses on using Artificial Neural Network (ANN) as a significant technique of artificial intelligence to diagnose thyroid diseases. Multilayer feed forward architecture of ANN is adopted in the proposed design and the back propagation is selected as learning algorithm to accomplish the training process. The continuous values of three laboratory blood tests are used as input signals to the proposed system of ANN. Used 655 samples of real patients from certified advanced hormones laboratory in Kerala city. All types of thyroid diseases that may occur in patients are taken into account in design of system as well as the high accuracy of the detection and categorization of thyroid diseases are considered in the system. The result of this research shows that the proposed ANN system is able to precisely diagnose thyroid disease and can be exploited in practical
uses. The selected ANN has high classification rate which is about 99.29%. As a result, the proposed structure of ANN can effectively categorize the type of thyroid cases. The system is simulated via MATLAB Software to evaluate its performance.

Amato et al. (2013) had carried out a research study on the title “Usage of Artificial Neural Network (ANN) in medical diagnosis”. The objective shows that the capability, philosophy, limitations and powerful use of artificial neural network in the medical diagnosis, how it is useful for physicians to diagnosis more reliable and therefore, increases patient satisfaction. They have created Artificial Neural Network (ANN) based diagnosis architecture, all the medical data has put into it and find the minimum optimal value. The Artificial Neural Network (ANN) based medical diagnosis architecture worked on various diseases like diabetes, cancer and cardiovascular diseases. In the model also discussed about how can be build the database, training and verification of the database using Artificial Neural Network (ANN) and how can be test in medical practice. The study also, suggested usefulness of Artificial Neural Network (ANN) in future.

Sharpe et al. (1993) performed a research on “artificial neural networks in diagnosis of thyroid function from in vitro laboratory tests”. The main objective was to study this showing the potential benefit of using Artificial Neural Network (ANN) for the diagnosis of thyroid function. In this study used the two Artificial Neural Network (ANN) architectures and back propagation algorithms where they are multilevel perceptron and Learning Vector Quantization (LVQ). They have used clinical material as a data set to be train by Artificial Neural Network (ANN) architectures. Also, used the software neural works professional II package for random sampling of data was done on SYSMART/SAMPLE add-on module. Trained the data set on both the Artificial Neural Network (ANN) architecture and shows the desired output.

Predictive data mining for diagnosis of thyroid disease using neural network: Prerana and Taneja (2015) has done the research on “Predictive data mining for diagnosis of thyroid disease using neural network”. The main objective was for this study to introduce proposed Artificial Neural Network (ANN) concept research as an alternative for the earlier prediction of a disease. To train the proposed artificial neural network used two types of learning where in they are Supervised Learning (SL) and Unsupervised Learning (USL). They have done the predictive analysis of thyroid disease based on these above mentioned learning’s. They have divided the process into three stages, data collection and classifying, architecture selection and learning and compare network performance and reaching to the best answer. From experiments they have proven that Levenberg Marquardt method is better in performance in comparison of simple gradient descent algorithm. Used MATLAB as a tool for data analysis purpose.

Decision Support System (DSS) for diagnosis of thyroid disease using artificial neural network: Parveen et al. (2016) have performed a research on “A Review and Survey of Artificial Neural Network in Medical Science”. The objective for the study was how proposed Artificial Neural Network (ANN) techniques will be more useful in medical science field to help in diagnosis and predict the more precise output. They have discussed about the various Artificial Neural Network (ANN) Techniques. They have also shown the Fundamental working principle of artificial neurons. In this research, a Decision Support System (DSS) is proposed to diagnose nodules into benign and malignant by analyzing data via Artificial Neural Network (ANN). In this research study, 63 samples are taken out of dataset provided which are utilized to test and train the neural network based on algorithm. As a result, 95% accuracy is reached. In this research, mentioned four methods Bayesian networks, decision trees and simple classification models including Artificial Neural Network (ANN) are also useful for decisions in medical treatment.

Razia et al. (2015) have performed a research study on “A decision support system for prediction of thyroid disease a comparison of multilayer perceptron neural network and radial basis function neural network”. In this research, they have used two neural network models Multilayer Perceptron (MLP) and Radial Basis Function Networks (RBFN) for the prediction of onset of thyroid disease using the data generated in real life. The models multilayer perceptron is trained and tested with back-propagation algorithm whereas Radial Basis Function Networks (RBFN) was trained and tested with SPSS Software. It has been shown from experiments done that radial basis network can be successfully used for the diagnosis of thyroid disease.

Neural network in diagnosis of thyroid disease using cross-validation approach: Rastogi and Bhalla (2014) has done the research work on “A study of neural network in diagnosis of thyroid disease”. The main study focus on Artificial Neural Network (ANN) is considered as the best solutions to achieve the goal. A cross-validation approach has been used for sampling variability. They also discussed about the hybrid neural network structure
called CSFNN combines RBF and MLP in one single network. In this design, the nodes behave either as MLP or as RBF. The propagation rule for CSFNN comes employing using analytical equations using mistreatment a one. Also, mentioned the various normalization methods. Statistical or z-score normalization, min-max normalization, Median normalization, Sigmoid normalization, Statistical normalization, though the use of these normalization methods will increase the efficiency of network performance.

**Thyroid disease diagnosis using data mining techniques:**
Priya and Anitha (2017) has performed a research study on “Survey on thyroid diagnosis using data mining techniques”. The proposed study belongs to thyroid disease classification in two of the most common thyroid dysfunctions (hyperthyroidism and hypothyroidism) among the population. They have analyzed and compared four classification models: Naive Bayes, decision tree, multilayer perceptron and radial basis function network. The results indicate a big accuracy for all the classification models mentioned higher than, the best classification rate being that of the decision tree model. It was observed from the experiments that all the methods used by researchers as they provide high accuracy and efficiency.

Roshan and Sharmili (2017) has performed a research study on “A study of data mining techniques to detect thyroid disease”. They have discussed about the data mining techniques to be useful in decision making in the diagnosis of thyroid. Algorithms used are CART, LDA, classification, clustering, decision tree and k-fold cross validation. Experiments shows that the proposed system gives with classification and clustering accuracy with less number of features compared to other existing developed model. Various cacophonous rule for call tree attribute choice had been analyzed and compared.

Islami et al. (2016) has done the research work on “Surveying the knowledge of pregnant women towards sport activities during pregnancy using data mining algorithms”. The main objective was to research the knowledge of pregnant women towards sport activities using data mining algorithms. Statistical population includes all healthy pregnant ladies pertaining to health centers in Gorgan town (Iran) in 2014 from that 429 were chosen because the sample victimization cluster sampling. They have taken sample of data of all pregnant women in 32-40 weeks of pregnancy, healthy and less risky referring to Gorgan Health Center. Sampling was done in a cluster sampling method. First 10 centers were chosen based on geographic distribution (research domain) then the number of samples from each center was determined by Stratified sampling method and the convenience sampling was done. They have used methods like decision tree algorithms and Support Vector Machines (SVMs).

**Classification algorithms to predict thyroid disease:**
Gopinath (2017) has performed a research study on “Comparative study on classification algorithm for thyroid data set”. The main objective was to discuss the comparative study on classification algorithm for thyroid data set. They have taken the classifiers like SVM, k-NN, decision tree on which the dataset of 215 samples are given as input for classification under these classifiers to train the dataset and check behavior. It has been seen from the results that SVM classifier technique provides better accuracies as compared to last works. The proposed study performs efficiently on the dataset of 215 samples with an accuracy of 96.30. However, if we merge any other classification technique such as fuzzy classification or neural network on the output that we got from SVM algorithm, then the system.

Umadevi and JeenMarseline (2017) has performed a research study on “applying classification algorithms to predict thyroid disease”. The proposed study handles the analysis of the classification of the thyroid disease based on the information gathered from the UCI machine learning repository. They have used methods artificial neural network and k-nearest neighbor applied to the prediction of thyroid disease. Used MATLAB (“MATrixL A Boratory”) may be a tool for numerical computation and mental image. It was observed from the experiments, the trial of 21 parameters is used. In kNN the prediction accuracy is eightieth, ANN the accuracy is 85%. However, fuzzy ANN the prediction accuracy is 90% might provide even better accuracy rate as compared to what we got with the current study.

**Analysis of thyroid disease using back propagation algorithm:**
Kaur (2013) has performed a research study on “Analysis of thyroid disease using back propagation algorithm”. In the proposed research paper, an artificial neural network approach is developed using a back propagation algorithm in order to diagnose thyroid problems. It gets variety of things as input associated produces an output which provides the results of whether or not someone has the matter or is healthy. The research methods has used neural network back propagation technique for the diagnosis of results. It has shown from the experiments and results. The back propagation neural network has been trained and tested for the analysis of thyroid data. It has been determined that the convergence time for the coaching of back propagation neural network.
The increasing demand of neural network application for predicting the disease shows better performance in the field of medical decision making.

Sadoughi et al. (2014) has performed a research study on “An intelligent system based on back propagation neural network and particle swarm optimization for detection of prostate cancer from benign hyperplasia of prostate”. They have proposed a hybrid Back Propagation Neural Network (BPNN) classifier based Particle Swarm Optimization (PSO) method. They have used the sample of 360 patients suffering from neoplasia diseases. Discussed about the hybrid algorithm of PSO and BPNN and Particle Swarm Optimization (PSO) algorithm. Experiments shows proposed system has some advantages of automation. It is fast, easy to process, noninvasive and cheap for clinical application. Though, this system does not diagnose cancer conclusively but by providing information it can helps doctors in deciding whether a biopsy is necessary or not. Such diagnosing methodology will facilitate doc in creating correct call concerning glandular carcinoma from benign dysplasia of prostate.

Srujana et al. (2016) have performed a research study on “Diagnosis of thyroid disorders using back propagation method”. The main objective was to discuss about the diagnosis of thyroid disorders using back propagation algorithm by training feed-forward neural network. They have taken sample data of 7,200 patients which would be trained in back propagation algorithm. The experiments shows the methods used reach the classification accuracy for Thyroid disease to 98% and can be a best solution to increase the performance of feed-forward neural network. The performance of the proposed system was better compare than the decision tree splitting rules (Fig. 1).

Diagnosis of hypo and hyperthyroid using MLPN network: Vaz (2014) has performed a research study on “Diagnosis of hypo and hyperthyroid using MLPN network”. The proposed research study gives a survey of the methods used for the diagnosis of thyroid disorder and classification using MLPN. Used the different methods thermo graphic camera system, image filtering, image enhancement, image segmentation and feature extraction.

In this research, the thyroid images are classified using MLPNN. The experimental results show that the classifier expeditiously classifies thyroid pictures manufacturing a sensitivity of 93.12% for 3 fold and 89.62% for 10 fold. This research is extended to phase the thyroid pictures mistreatment mathematical logic and additionally classify the photographs to thyroid, non-thyroid and benign/malignant nodules (Fig. 2).

Thyroid prediction system using Auto Associative Neural Network (AANN): Agrawal and Dhakar (2017) has performed the research on “Thyroid Prediction System using auto associative neural network”. The main objective was to develop a predictive system for thyroid detection such as hypothyroidism, hyperthyroidism, sick people and normal people. They have applied Auto Associative Neural Network (AANN) for thyroid patient dataset. They have discussed the point’s hypothyroidism, hyperthyroidism in detail. They have used three methods, data preprocessing, strong feature selection, predictive model development. The thyroid information set could be a elite information set containing 360 dataset with ten attributes and four categories. The thyroid prediction system can develop can scale back the attributes utilized in classifying thyroid. AANN could model a classifier for such a selected data set with good
accuracy of higher than 95.1%. The results indicate that the models developed ANNs have better accuracy and suitable for systems like thyroid.

**Elman recurrent network model-diagnosis of thyroid abnormalities:** Sundaram and Renupriya (2016) have performed a research study on “Artificial neural network classifiers for diagnosis of thyroid abnormalities”. The main objective was to explore the artificial neural network classifiers. They have discussed about the simulations and models and data set can be used to create a neural network that classifies patients referred to a clinic as: Normal, not hyperthyroid, hyper function, subnormal functioning. Taken the sample data of 7,200 patients. Experiments shows feed forward neural networks and Elman networks could model a classifier for such a large data set with good accuracy of higher than 90%. The results indicate that the models developed by elman recurrent networks have better accuracy and suitable for modeling non linear dynamic systems like thyroid.

**Thyroid disease diagnosis using hybrid intelligent systems:** Saleh et al. (2016) has done the research work on “Thyroid disease diagnosis using hybrid intelligent systems’. They have proposed a diagnosis system based on hybrid intelligent systems (neuro-fuzzy network) as classifier tool. The neural network and symbolic logic was combined to induce the most options of artificial neural networks with those of symbolic logic and to beat a number of the restrictions of these techniques. The experiments shows that neuro fuzzy system performed better where accuracy is 100% for training and in range between 87 and 95% for testing which mean that neuro fuzzy system may be wont to facilitate identification of thyroid malady.

**Thyroid diseases forecasting using a hybrid decision support system based on ANFIS, k-NN and information gain method:** Ahmad et al. (2017) have done the research work on “Thyroid diseases forecasting using a hybrid decision support system based on ANFIS, k-NN and information gain method”. They have proposed a new decision-based hybrid system for the diagnosis of thyroid diseases. They have done the experiments on thyroid diseases dataset to find out the effectiveness of the proposed system. It contained three important phases, 25 features of the dataset were reduced using information gain method to avoid data redundancy and reduce computation time. The missing values in the dataset are dealt with k-Nearest Neighbor (k-NN) weighting pre-processing scheme the resultant data is provided as input to adaptive neuro-fuzzy inference system for the purpose of input-output mapping. The innovation of this proposed system is that it is a hybrid system, comprised of feature selection process using information gain method which decreases computation time and increases the accuracy of the ensuing model, k-NN imputation for missing knowledge values and ANFIS system that maximize the generalization capability of our thyroid identification system.

Kala et al. (2010) has done the research work on “Medical diagnosis using incremental evolution of neural network”. The main objective was to diagnosis of PIMA Indian diabetes. Therefore, have proposed an evolutionary neural network which is used in multi-layer perceptron. Discussed about the general architecture of multi-layer perceptron neural network and the general structure of the proposed algorithm. Used the evolutionary operators which are: selection, mutation, crossover, elite, jump and new. From the experiments shown that the proposed algorithm performed better compare than the conventional multi-layer perceptron with back propagation algorithm for training, modular neural networks, connectionist evolution of neural network, radial basis function network and adaptive neuro fuzzy inference system. They are sure that proposed algorithm was able to achieve a high degree of accuracy for both training as well as testing data sets.

**Optimized construction of various classification models for the diagnosis of thyroid problems in human beings:** Rajkumar and Palanichamy (2015) have performed the research work on “Optimized construction of various classification models for the diagnosis of thyroid problems in human beings”. This analysis critically examines completely different classification models made employing a novel mathematical hierarchal Improved F-score Ordering (RIFO) applied to thyroid dataset taken from machine learning repository. They have used two important neural network structures Multilayer Perceptron (MLP) and Radial Basis Function (RBF). Different models are made by hierarchal ordering of their calculated improved F-score values. Also, discussed about the construction of classification models which is used for data analysis in machine learning, statistics and data mining. They have taken the thyroid gland dataset consists of 3 classes, 215 samples and each sample has 5 attributes.

**Neuro computing frame work for thyroid disease diagnosis using Machine Learning Techniques (MLT):** Razia et al. (2015) performed a research study on “A neuro computing frame work for thyroid disease diagnosis using machine learning techniques”. The
proposed research study based on using two neuronal models (SOM and LVQ). In this analysis, a framework victimization Self-Organized Map (SOM) beside learning vector quantization has been developed. The unlabeled thyroid data of about 215 different patients is obtained from a clinic and is used to train the SOM network using competitive learning algorithms. The researchers have also used the decision tree algorithms: Self-Organizing Map (SOM) neural networks and linear vector quantization. The experiments and results show that LVQ and SOM algorithm have better accuracy for diagnosis of thyroid disease (Table 1).

Ammulu (2017) has performed a research work on “Thyroid data prediction using data classification algorithm”. The proposed research work focuses on using data mining approach to predict the thyroid disorder with better accuracy results. The sample size for the proposed work was 3090 instances. In this 149 data comes under hypothyroid and 2941 data is negative cases. For the analysis researchers have used random forest approach. The experiment shows that improved accuracy, precision, recall and F-measure by comparing the random forest with LDA algorithm.

**Materials and Methods**

**Materials and methodology deployed for diagnosing thyroid diseases:** After going through different studies it is found that various architecture of Artificial Neural Network (ANN), Back Propagation Network (BPN), Hybrid Back-Propagation Neural Network (BPNN), Data Mining Methods (DMM), Decision Support System (DSS), Ranked Improved F-score Ordering (RIFO), Auto Associative Neural Network (AANN), Multi-Layer Perceptron (MLP), Multivariate Bayesian Prediction Method (MBPM), Machine Learning Methods (MLM) and Radical Basis Function Neural Network (RBFNN) are found to be proper and appropriately suitable for diagnosing thyroid diseases in a different proportionate. In the recent years, these architectures and methods are also found and observed to be most useful for prediction of so many other diseases also, further. These discussions on various propaganda paved a pathway towards accommodating these architectures and their suitability, appropriateness for thyroid disease prediction is presented through this review article and it would reveal the various methods and its benefits of diagnosing thyroid diseases in a most appropriate and essential manner based on the nature and cause it affects on human body.

**RESULTS AND DISCUSSION**

Identification of thyroid disease becomes most important and vital process in our day to day life of many patients in certain cases; it becomes a difficult task from both clinical diagnosis and statistical classification point of view. The poor performance of the traditional model based statistical methods and due to large number of interrelated patient attributes as well as extremely unbalanced groups in the thyroid diagnosis problem complicates the relationship between these attributes and the patient true group membership. After the emergence Artificial Neural network (ANN) is a flexible modeling technique exclusively used for complex functional mapping and showed progressiveness in the diagnosing process thyroid disease. After the emergence of artificial neural network, many researchers used different proportionate in order the diagnose thyroid disease in a most appropriate manner. Each of the researchers had performed different methodologies with a different objective but finally they focus on finding the accuracy and impact of thyroid diseases had not changed yet, still they do try different and wider context but eventually they all are working towards identifying the best suitable and appropriate methodology for diagnosing thyroid diseases with a fullest efforts.
CONCLUSION

Most of the researchers focused on artificial neural networks as a diagnosing tool to increase the accuracy of performance. There was no appropriate selection of artificial neural networks architecture which certainly affects the network performance wherein it is to perform effectively to reach the high accuracy. In this study, we had identified and found the various type of appropriate architecture and techniques where in the correct selection of diagnosing will ensure the accuracy and the network complexity can be reduced, so as to achieve the best result by comparing their performance to reach the best possible. This study will certainly help the academicians and researchers to take forward a step towards identifying the best suitable and most widely used method of diagnosing thyroid diseases in a wider context.

REFERENCES


