International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 11 | Nov -2016 www.irjet.net p-ISSN: 2395-0072

Diagnosis of Thyroid Disease Using Data Mining Techniques: A Study

Dr.B.Srinivasan [1], K.Pavya [2]

[1]*Department of Computer Science, Gobi Arts & Science College, Bharathiar University, Tamil Nadu, India*

[2]*Department of Computer Science, Vellalar College for Women, Bharathiar University, Tamil Nadu, India*

Abstract: *Data mining based classification is one of the important role in the field of healthcare. The medical data having a huge amount of data set these data set classifications are highly challenging task. Diagnosis of health conditions is very essential and difficult task in field of medical science. There are various types of diseases are diagnosis in medical science. Thyroid is one of the unpredictable diseases in medical field. Thyroid hormones control the body's metabolic rate. Moreover, data mining techniques has been applied in various sector and the classification results of the medical data set which helps the way of treatments to the patients. There are various authors have worked in the field of thyroid diseases classification and give the classification accuracy with robust model. Most famously used techniques are Decision tree which is used to reduce noisy data, Naïve Bayse is help to get high speed and accuracy rate when applies large data base, Back propagation gives high accuracy neural network when applied complex non-liner, Support Vector Machine is easy to extended and pattern recognition. This paper presents the study on Diagnosis of Thyroid disease.*

Keywords: Thyroid disease,Decision tree, Naïve bayse,SVM, Backpropagation.

I.INTRODUCTION

Thyroid Disease diagnosis is one of the very difficult and tedious tasks, because it needs lots of experience and knowledge. The traditional ways for diagnosis thyroid disease is doctor's examination or a number of blood tests. Main task is to provide disease diagnosis at early stages with higher accuracy. Data mining plays a vital role in medical field for disease diagnosis. It offers lot of classification techniques to predict the disease accuracy. Hospitals and clinics gathered a large amount of patient data over the years. These data provide a basis for the analysis of risk factors for many diseases.

Data mining based applications are very beneficial and important in healthcare and medical science. In health care, there are large amount of data, and this data has no organizational value until converted into information and knowledge, which can help control costs, increase profits, and maintain high quality of patient care. Thyroid is one of the most severe health challenges in both developing and developed. One of the most important applications of data mining technique is classification. Classification is one of the most important decision making techniques in many real world problem. In this study, there are various authors have worked in the field of classification of thyroid data using data mining based classification techniques and they given satisfactory results for classification of thyroid data.

II.OVERVIEW OF THYROID

The thyroid is an organ present in the human body and is considered to be a part of the endocrine or the hormone, system. It is located in the human neck below the Adam's apple. The main purpose of thyroid is to produce thyroid hormones. The produced hormones go through the bloodstream to all the other organs which help to control metabolism and growth development in both in adults and in children. The thyroid looks like butterfly shape. Figure 1 shows the thyroid and its parts. The right and left lobes of the thyroid looks similar to the two wings of a butterfly. They lie on both sides of the trachea or main breathing tube. The connection between the wings is called the isthmus [1]. The thyroid gland produces hormones which primarily control human body's growth and metabolism, which means that this energy is used for all the body processes. The thyroid gland acts as an important part in breathing, blood circulation, bowel movements, temperature of the body, muscle control, digestion, and brain function. An issue with the thyroid gland can result in problems all over the human body [3]. The thyroid gland functional data is more essential for the proper interpretation and diagnosis of the diseases associated with the gland. The principal role of the thyroid gland is to help regulation of the body's metabolism. Depending on the amount of secretion of this hormone may affect the human growth and development. When this hormone is produced very little thyroid hormone the type of disease is referred to as hypo-thyroidism. When this hormone is produced of too much it may lead to hyper-thyroidism [2].

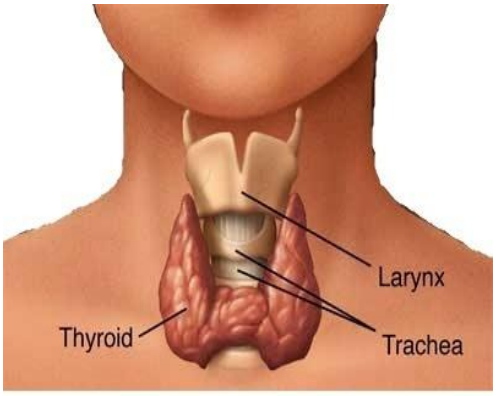

Fig. 1 : Thyroid in human body

a)Thyroid Harmones

The two hormones that are produced in the thyroid are L-thyroxine (T4) and tri-iodothyronine (T3)[1]. They regulate human body's metabolic functions such as heat generation, and the utilization of carbohydrates, proteins, and fats. Regulatory hormones from different parts of the brain control the thyroid's production of T4 and T3. In the pituitary gland, Thyrotropin-Stimulating Hormone (TSH) is released when more thyroid hormone is needed and travels via the bloodstream to the thyroid gland. TSH then stimulates the thyroid to produce T4 and T3 [1]. The pituitary gland acts like a thermostat to control the production of the hormone. When they are more in the bloodstream, the pituitary releases less TSH. When there are little in the bloodstream, the pituitary releases more TSH. With the help of this feedback system, the production of thyroid hormone is tightly controlled [1].

b)Thyroid and Health Effects

Thyroid diseases are one of the most common endocrine disorders worldwide. India too, is no exception. It is estimated that about 42 million people in India suffer from thyroid diseases [4].

Thyroid diseases are different from other diseases in terms of their ease of diagnosis, accessibility of medical treatment, and the relative visibility[4]. The thyroid gland secretes hormones which controls a lot of things in the human body system like metabolize the food, use energy, sleep patterns, temperature preferences, body weight balance and a lot more [5].

Both an increase and decrease in thyroid hormone production can cause health problems.

i)Hyperthyroid

Increase in the hormone production can cause hyperthyroidism. In medical field, "hyper" indicates too much. Hyperthyroidism crop up when the gland produces excess hormones. The most common cause for hyperthyroidism is the autoimmune disorder Graves' disease. It is also known as an overactive thyroid, the hormone overload can cause a extensive range of physical changes. Many symptoms overlap with hypothyroidism, including thinning hair, dry skin and temperature sensitivity. The symptoms that indicate the presence of hyperthyroidism includes weight loss in spite of a good food intake, an increase in heart rate, high blood pressure, nervousness, increased sweating, enlargement in your neck, shorter menstrual periods, frequent bowel movements and trembling hands [3].

ii)Hypothyroid

Decrease in the hormone production can cause hypothyroidism. In medical field, .the term hypo means deficient or not enough. Hypothyroidism is a condition that the thyroid gland does not produce required hormones. Inflammation and damage to the gland causes hypothyroidism. Weight gain or failure to lose weight despite a proper weight loss regime, lethargy, reduced heart rate, increased cold sensitivity, numbness in hands, enlargement in the neck, dry skin and hair, heavy menstrual periods and constipation could indicate hypothyroidism. Symptoms vary from person to person, and if left untreated, they tend to worsen over time [3]. Table 1 shows the list of symptoms of Hypothyroid and Hyperthyroid.

SYMPTOMS OF HYPOTHYROID	SYMPTOMS OF HYPERTHYROID
Dry hair	Hair loss
Slow heartbeat	Rapid heart beat
Weight gain	Weight loss
Constipation	Sleeping difficulties
Memory loss	Irritability
Heavy menstrual periods	Scant menstrual periods
Muscle aches	Muscle weakness
Cold intolerance	Heat intolerance

Tab.1 : Symptoms of Hypothyroid and Hyperthyroid.

An increased risk of thyroid disease happens if there is a family history of thyroid disease like a type I diabetic, over 50 years of age and a stressful life [5]. Both hypothyroidism and hyperthyroidism can be diagnosed with thyroid function tests, which measures the levels of Thyroid-Stimulating Hormones (TSH) in bloodstream of human body [3].

III.RELATED WORK

Deepika Koundal et al.[6]have provided the information about the existing automatic tools which are available to formulate the disease diagnosis part easier with efficient way. Also different performance evaluation metrics are studied.The future developments and trends are also investigated

Nikita Sigh and Alka Jindal[7] have concluded that SVM is better as compared to KNN and Bayesian.Accuracy of SVM is about 84.62%.KNN found the nearest neighborhood automatically. It represented by graph each vertices having object. Bayesian based on the probability classification which gives the sample data belongs to a class.

Edgar Gabriel et al.[8] have proposed two parallel versions of a code that are used for texture-based segmentation of thyroid FNAC images which is a critical first step in reliazing a fully automated CAD solution. An MPI version of the code is developed to exploit distributed memory computer resources such as PC clusters.

Preeti Aggawal et al.[9] suggested an automatic segmentation method. They have provided a summary of all the results obtained either by automatic tools as well as by applying specific algorithm(automatic) segmentation on both lung CT as well as on thyroid US. There are two tools: Analyze 10.0 and Mazda for segmentation of thyroid US images.

Eystraints G[10]have suggested a computer-aided diagnosis(CAD) system prototype named as TND(Thyroid Nodule Detector).It is used for the detection of nodular tissue in ultrasound(US)thyroid images and videos acquired during thyroid US examinations.

Won-Jin Moon et al.[11]have done in their paper the evaluation on the diagnostic accuracy of ultrasonographic(US) criteria for the depiction of benign and malignant thyroid nodules.It is done by using tissue diagnosis as the reference standard.They concluded that shape,margin,echogenicity and presence of calcification are important criteria for the discrimination of malignant from bening nodules.

S.Yasodha et al.[12] have proposed CACC-SVM techniques which is hybridization of class-Attribute Contingency Coefficiant(CACC) and support vector machine(SVM) for classification of thyroid data.The proposed model achieved better accuracy compared to other traditional models

Alfonso Bastias et al.[13]have focused on developing an AIS based machine learning classifier for medical diagnosis and investigating the capability of the proposed classifier.The proposed classifier successfuiiy improved the identification process of thyroid gland disease.

Gurmeet kaur et al. [14] has proposed an efficient neural network training model for thyroid disease diagnosis.It presents general model for diagnosing any disease.The objective of this paper is to diagnose thyroid disease by using three different neural network algorithm which have different architecture and characteristics.

Ali keles et al. [15] aims at developing an expert system for thyroid diagnosis that is known as Expert System for Thyroid Disease Diagnosis(ESTDD).In this expert system authors have used neuro fuzzy rules which could diagnose thyroid diseases with 95.33% of accuracy.

IV.CLASSIFICATION TECHNIQUES IN DATA MINING

1)Decision Tree

A Decision tree has three types of node, internal node denotes the test on an attribute, leaf node denotes the classes or class distributions, root node is the most node in a tree.C4.5 and ID3 are the two main algorithms used to construct the decision tree.

2) Naïve Bayesian Classification

Bayesian classification is derived from the Bayes theorem. It is also known as "Simple Bayesian Classifier". In this classifier each data sample is represented by an n-dimentional vector and measurements samples are formed by n-attributes.Suppose there are m classes,C_1,C_2 .....C_3 having a unknown data sample,X,Navie Bayesian classifier will predict that belongs to class having highest probability conditioned on X.an unknown sample X to the class C_i.

$P(C_i|X)>P(C_j|X)$ for $1\leq j\leq m, j\neq i$.

3) Backpropagation Neural Network

Backpropagation is based on the neural network algorithm.There are three kinds of layer, input layer-the inputs are fed into this layer,hidden layer-weighed outputs can be input to another hidden layer,number of hidden layer's arbitrary,but only one is used, output layer-hidden layers are give the input to the output layer,which emits the network's prediction.This network model helps to classifying new data.

4) Support Vector Machine

Support Vector Machine is one type of learning system algorithm, which is used to perform classification more accurately.SVM used for two class classifier. The essence of SVM is hyper plane also known as "Decision boundary or decision surface". This hyper plane separates the positive and negative of training data sample.

V. CONCLUSION

Thus the survey helps to identify how the data mining techniques to predict the thyroid disorder at earlier stage. Different Researchers have proposed different techniques to predict the thyroid disorder and different kinds of accuracy level as per used techniques. These techniques help to minimize the noisy data of the patient's data from the data bases. Algorithms such as Naïve bayes, Decision tree, Back propagation, Support vector machine are considered for the study. These algorithm gave the various result based on speed, accuracy, performance and cost. Also these effective classification data are helps to find the treatment of the patient.

REFERENCES

[1]http://www.emedicinehealth.com/thyroid-faqs/article-em.htm

[2] Ozyilmaz,Lale, and Tulay Yildirim."Diagnosis of thyroid disease using artificial neural network methods".Neural Information Processing,2002.ICONIP'02.Proceedings of the 9th International Conference on Vol.4.IEEE,2002.

[3]http://www.foxnews.com/health/2012/02/10/hypothyroidism–versus-hyperthyroidism.html(accessed dec 2015)

[4]http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3169866/(accessed dec 2015)

[5]http://www.thehealthsite.com/diseasesconditions/world- thyroid - day- 2012 – facts-you-should-know/(accessed dec 2015)

[6]Deepika Koundle,et al.,"Computer-Aided Diagnosis of Thyroid Nodule: A Review"'International Journal of Computer Science and Engineering Survey(IJCSES) Vol.3,No.4,Aug 2012

[7] Nikita Sigh, Alka Jindal,"A Segmentation Method and Comparison of Classification Methods for Thyroid UltraSound Images", 'International Journal of Computer Applications(0975-8887)Volume 50 – No.11,July 2012

[8] Edgar Gabriel et al.,"Towards high performance cell segmentation in multispectral fine needle aspiration cytology of thyroid lesions",computer methods and programs in biomedicine

[9] Preeti Aggawal et al.,"Role of Segmentation in Medical Imaging: A Comparative Study", International Journal of Computer Applications(0975-8887)Volume 29 – No.1,Sep 2011

[10] Eystraints G,"TND:A thyroid Nodule Detection System for analysis of Ultrasound Image and Videos",Springer Science and Business Media,LLC 2010

[11] Won-Jin Moon et al."Bening and Malignant Thyroid Nodules:US Differentiation",Multicenter Retrospective Study Radiology: Volume 247;No.3-Jun 2008

[12]S.Yasodha et al.,"Data Mining Classification Technique for Talent Management using SVM",International Conference on Computing,Electronics and Electrical Tecnologies(ICCEET),978-1-73-0210-/12,pp.959-963,2012

[13] Alfonso Bastias et al.,"Predictive model based on neural networks to assist the diagnosis of malignancy of thyroid nodules",Proceedings of the 41st International Conference on Computers & Industrial Engineering.

[14] Gurmeet kaur et al.,"Proposing Efficient Neural Network Training Model for Thyroid Disease Diagnosis",International Journal for Technological Research in Engineering,Vol 1,Issue 11,Jul-2014

[15] Ali keles et al.,"ESTDD:Expert system for thyroid diseases diagnosis",Expert system with Applications,34,242-246,200.