

PREDICTING THYROID DISEASE USING DATA MINING TECHNIQUE

¹Dr. G. Rasitha Banu MCA., M.Phil., Ph.D., ²M.Baviya MCA.,

¹ Professor, Department of MCA, Vels University, Chennai, Tamil Nadu, India

²MPhil, Department of Computer Science, Mother Teresa Women's University Chennai, India

Abstract - Data mining has been used intensively and extensively by many organizations. In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Data mining applications can greatly benefit all parties involved in the healthcare industry. For example, data mining can help healthcare insurers detect fraud and abuse, healthcare organizations make customer relationship management decisions, physicians identify effective treatments and best practices, and patients receive better and more affordable healthcare services. The huge amounts of data generated by healthcare transactions are too complex and voluminous to be processed and analyzed by traditional methods. Data mining provides the methodology and technology to transform these mounds of data into useful information for decision making. There are two main methods of Data mining: Clustering and Classification. In many cases the concept of classification is confused by means of clustering, but there is difference between these two methods. According to the perspective of Machine learning clustering method is unsupervised learning and tries to group sets of objects having relationship between them, whereas classification method is supervised and assigning objects to sets of predefined classes. In proposed system are classified and cluster of the thyroid disease in data mining.

KeyTerms - DSCN algorithm, Hierarchical multiple classifier, TSH, Hypothyroidism, Hyperthyroidism, clustering, classification.

1. INTRODUCTION

Hypothyroidism is a relatively common problem worldwide often with insidious onset and is relatively asymptomatic. Respiratory manifestations are seldom the major complaints in hypothyroidism. Lung volumes are usually normal, but few studies have shown findings suggestive of restrictive pattern of impairment. This has been attributed to decreased in both expiratory and inspiratory muscle strength, alveolar hypoventilation due to depression of hypoxic and hypercapnic ventilatory drives and decreased in maximal breathing and diffusing capacity in patients with hypothyroidism. Weight gain independently of physical activity is frequently associated with hypothyroidism. Most of the restrictive defects seen among these patients has been thought to be a consequence of obesity and microatelectasis. Overt thyroid dysfunction is well recognized to affect weight, but the influence of minor perturbations of thyroid function remains unclear.

Hypothyroidism can have numerous effects on the respiratory system. Fatigue and dyspnoea on exertion are frequent symptoms. But in the absence of primary respiratory disease, the diminution of the respiratory function in the hypothyroid patients is not significant in most cases.

Nevertheless, it does affect the respiratory system including respiratory muscle weakness, alveolar hypoventilation due to decreased hypoxic and hypercapnic ventilatory drives, upper airway obstruction, central and obstructive sleep apnoea and even pleural effusion. Lung volumes are

usually normal or mildly reduced, but maximal breathing capacity and diffusing capacity are usually reduced.

Technologies for generating high-density arrays of cDNAs and oligonucleotides are developing rapidly, and changing the landscape of biological and biomedical research. They enable, for the first time, a global, simultaneous view on the transcription levels of many thousands of genes, when the cell undergoes specific processes and in certain conditions. For several organisms, the sequences of all genes are available, and thus, transcript levels of the complete gene collection can already be monitored today. The potential of such technologies is tremendous: Monitoring gene expression levels in different developmental stages, tissue types, clinical conditions and different organisms can help understanding gene function and gene networks, assist in the diagnosis of disease conditions and reveal the effects of medical treatments. Undoubtedly, other applications will emerge in coming years.

2. CONCEPTS OF TECHNOLOGY

DATA MINING

Data Mining is the process of semi-automatically analyzing large databases to find patterns.

DBSCAN

The **DBSCAN (Density-based spatial clustering of applications with noise)** algorithm is based on the concepts of density reachability and density-connectivity. These concepts depend on two input parameters: epsilon (eps) and minimum number of points (minPts). Epsilon is the distance around an object that defines its eps-neighborhood. The DBSCAN algorithm works as follows. Initially, all objects in the data set are assumed to be unassigned. DBSCAN then chooses an arbitrary unassigned object p from the data set. If DBSCAN nds (Non-Deliverable Swap) p is a core object, it nds all the density-connected objects based on eps and minPts. It assigns all these objects to a new cluster. If DBSCAN nds p is not a core object, then p is considered to be noise and DBSCAN moves onto the next unassigned object. Once every object is assigned, the algorithm steps.

HIERARCHICAL MULTIPLE CLASSIFIER

Hierarchical multiple classifier classification scheme, which preserves the strength of the multiple-classifier approach and also manages to reduce some of the problems faced by other multiple-classifier algorithms. In our scheme, the system resource requirements are reduced and so is the training time. From the results on pap-smear data, it can be seen that our approach produces better performance than other multiple-classifier algorithms and better than a classifier produced by human experts.

TSH

Thyroid-stimulating hormone (also known as **thyrotropin, TSH**, or **hTSH** for human TSH) is a pituitary hormone that stimulates the thyroid gland to produce thyroxine (T_4), and then triiodothyronine (T_3) which stimulates the metabolism of almost every tissue in the body.^[1] It is a glycoprotein hormone synthesized and secreted by thyrotrope cells in the anterior pituitary gland, which regulates the endocrine function of the thyroid.

HYPERTHYROIDISM AND HYPOTHYROIDISM

Hyperthyroidism and hypothyroidism are common conditions that have lifelong effects on health. Hypothyroidism is a condition in which the body lacks sufficient thyroid hormone. Since the main purpose of thyroid hormone is to "run the body's metabolism," it is understandable that people with this condition will have symptoms associated with a slow metabolism. Hypothyroidism is more

common than you would believe, and millions of people are currently hypothyroid and don't know it. In Existing system there is no mechanism to find the thyroid disease without doctor help. Common people unable to spend time and money to know the prediction for thyroid disease. Our proposed system can help people to know the prediction for thyroid disease it is help people to know the prediction details and level detail anywhere in the world. This system using classification and clustering method to find the prediction details.

Hyperthyroidism

Hyperthyroidism (overactive thyroid) is a condition in which your thyroid gland produces too much of the hormone thyroxine. Hyperthyroidism can accelerate your body's metabolism significantly, causing sudden weight loss, a rapid or irregular heartbeat, sweating, and nervousness or irritability.

Hypothyroidism

Hypothyroidism (underactive thyroid or low thyroid) is a condition in which your thyroid gland doesn't produce enough of certain important hormones.

Women, especially those older than age 60, are more likely to have hypothyroidism. Hypothyroidism upsets the normal balance of chemical reactions in your body. It seldom causes symptoms in the early stages, but, over time, untreated hypothyroidism can cause a number of health problems, such as obesity, joint pain, infertility and heart disease.

CLUSTERING

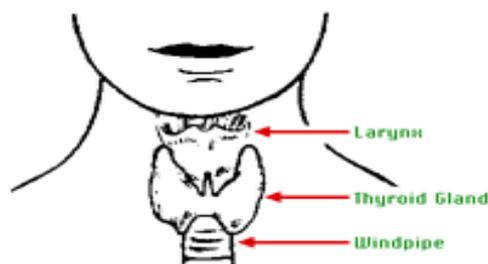
Clustering is a process of partitioning a set of data (or objects) into a set of meaningful sub-classes, called clusters. Help users understand the natural grouping or structure in a data set. Used either as a stand-alone tool to get insight into data distribution or as a preprocessing step for other algorithms.

CLASSIFICATION

Classification is a data mining (machine learning) technique used to predict group membership for data instances. For example, you may wish to use classification to predict whether the weather on a particular day will be sunny, rainy or cloudy. Popular classification techniques include decision trees and neural networks.

3. OVERVIEW OF THYROID

The thyroid gland is located in the front of the neck attached to the lower part of the voicebox (or larynx) and to the upper part of the windpipe (or trachea). It has two sides or lobes. These lobes are connected by a narrow neck (or isthmus). Each lobe is about 4 cm long and 1 to 2 cm wide. The name "thyroid" comes from the Greek word which means "shield".



ThyroidHormones

The thyroid gland produces thyroid hormones. These are peptides containing iodine. The two most important hormones are tetraiodothyronine (thyroxine or T4) and triiodothyronine (T3). These

hormones are essential for life and have many effects on body metabolism, growth, and development.

Iodine

Iodine plays an important role in the function of the thyroid gland. It is the chief component of thyroid hormones, and is essential for their production. Iodine is obtained from the water we drink and the food we eat. In areas of the world where there is an iodine deficiency, iodine must be added to the salt or bread. The Great Lakes area of Canada and the U.S., the Swiss Alps and Tasmania are such areas. In Canada and the U.S., most of the salt is iodized, thus the iodine intake is more than adequate. Taking excess amounts of iodine in foods such as kelp can aggravate hyperthyroid disease.

Goitre

Enlargement of the thyroid gland is called goitre. Goitre does not always indicate a disease, since thyroid enlargement can also be caused by physiological conditions such as puberty and pregnancy.

Thyroid Disorders

The main causes of thyroid disease are:

- Too much thyroid hormone production or **hyperthyroidism**.
- Too little thyroid hormone production or **hypothyroidism**.

The state of normal thyroid function is called **euthyroidism**. Abnormalities of the thyroid gland are common and affect 1-5% of the population. All thyroid disorders are much more common in women than in men. Because of the widespread use of iodized salt and bread, lack of iodine is no longer a cause of thyroid disease in Canada as it was some 50 years ago.

"Autoimmune disorders" of the thyroid gland are the most common cause of thyroid dysfunction. These autoimmune disorders are caused by abnormal proteins, (called antibodies), and the white blood cells which act together to stimulate or damage the thyroid gland. Graves' disease (hyperthyroidism) and Hashimoto's thyroiditis, are diseases of this type. Graves' disease affects about 1 in 100 of the population, whereas Hashimoto's thyroiditis is even more common (its prevalence increases with age).

Graves Disease

Graves' disease (thyrotoxicosis) is due to a unique antibody called "thyroid stimulating antibody" which stimulates the thyroid cells to grow larger and to produce excessive amounts of thyroid hormones. In this disease, the goitre is due not to TSH but to this unique antibody.

Hashimoto's Thyroiditis

In Hashimoto's thyroiditis, the goitre is caused by an accumulation of white blood cells and fluid (inflammation) in the thyroid gland. This leads to destruction of the thyroid cells and, eventually, thyroid failure (hypothyroidism). In the beginning, thyroid hormone production decreases. In response to lower thyroid hormones levels, TSH increases and goitre can develop. In the later stages, the goitre can disappear because of the progressive destruction of the thyroid.

Thyroid Nodules

Sometimes, thyroid enlargement is restricted to one part of the gland; the rest of the gland being normal. The most common cause of this is a cyst or nodule, which may be benign or malignant. Occasionally there are many nodules and this is called "multinodular goitre". Genetic modifications of the proteins usually involved in thyroid growth and function can contribute to this phenomenon.

4. CONCLUSION

DBSCAN algorithm is used for predicting the thyroid disease with the related symptoms. The DBSCAN algorithm works as follows. Initially, all objects in the data set are assumed to be unassigned. DBSCAN then chooses an arbitrary unassigned object p from the data set. Also classify the dataset using **Hierarchical multiple classifier** classification scheme, which preserves the strength of the multiple-classifier approach and also manages to reduce some of the problems faced by other multiple-classifier algorithms. In this way the data are classified in efficient way provide accurate information. The user can predict and test their health with the symptoms. The user can predict the thyroid disease with related symptoms. Before going to the hospital and check with the doctor. The user can predict the disease with relevant symptoms.

REFERENCE

1. G.K. Gupta. Introduction to Data Mining with Case Studies, 2006, Prentice Hall India.
2. Papageorgiou E, Kotsioni I & Linos A, (2005) "Data mining: a new technique in medical research", Hormones, Vol. 4, pp 189-191.
3. <http://en.wikipedia.org/wiki/Thyroid>
4. http://en.wikipedia.org/wiki/Thyroid_cancer
5. Amino N, Miyai K, Onishi T, Hashimoto T, Arai K, Ishibashi K et al. Transient hypothyroidism after delivery in autoimmune thyroiditis. Journal of Clinical Endocrinology and Metabolism 1976 42 296– 301.
6. Ginsberg J & Walfish PG. Post-partum transient thyrotoxicosis with painless thyroiditis. Lancet 1977 i 1125– 1128.
7. Gerstein HC. How common is postpartum thyroiditis? Archives of Internal Medicine 1990 150 1397–1400.
8. Amino N, Tada H & Hidaka Y. Autoimmune thyroid disease and pregnancy. Journal of Endocrinological Investigation 1996 19 59–70.
9. Harris B, Othman S, Davies JA, Weppner GJ, Richards CJ, Newcombe RG et al. Association between postpartum thyroid dysfunction and thyroid antibodies and depression. British Medical Journal 1992 305 152–156.

