IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

THYROID DISEASE PREDICTION AND PROVIDING MEDICATION USING MACHINE LEARNING

MD. Khaja Shafiyuddin[1], Dr. G. Venkata Rami Reddy[2]

[*1]Student, M.Tech (Software Engineering), School of Information Technology, JNTU Hyderabad, Hyderabad, India

[*2]Professor of IT, School of Information Technology, JNTU Hyderabad, Hyderabad, India

ABSTRACT: Despite the fact that medicine is built on a complicated concept, thyroid illness is the most prevalent cause for diagnoses or forecasts made by doctors. One significant component of the body is the thyroid gland. The thyroid gland's secretions regulate metabolic rate. Thyroid hormones, which are released by the thyroid gland, regulate body metabolism. The two most prevalent thyroid conditions are hyperthyroidism and thyroid problems. Hospital-based thyroid disease data are used for a variety of purposes, including disease prediction, research, and classification models. To address dynamic learning difficulties like medical diagnosis and prediction, a strong knowledge base must be assured, constructed, and utilized as a hybrid model. Simple machine learning methods are employed to locate and stop the thyroid. The SVM is used to determine how likely it is that a person has thyroid disease. Our system must recommend treatments such as over-the-counter medications, precautionary measures, and home remedies if a patient is at risk of developing thyroid problems.

Keywords *– Thyroid Disease; Hyperthyroidism; Triiodothyronine(T3); Levothyroxine(T4); K- Nearest Neighbours (KNN); Principal Component Analysis (PCA); Decision Tree.*

1. INTRODUCTION

The medical industry has a place for cutting-edge machine biology. This necessitated the gathering of data for medical forecasting of future illness. To recognise illnesses early, a variety of intelligent prediction algorithms are applied. The Healthcare Information System does a decent job of storing a lot of data, but it lacks the intelligence capabilities necessary for diagnosing illnesses at an early stage. Later, while creating a prediction model, dealing with challenging non-linear issues necessitates the application of ML techniques. In order to accurately depict a healthy patient, any disease prediction model must incorporate traits that may be selected from a variety of data sources. If not, a good patient can get the wrong diagnosis and therapy. Someone who has thyroid issues may be able to foresee problems.

The thyroid gland is an endocrine organ situated in the stomach. The pituitary gland is a tiny, almond-shaped gland that forms at the base of the "apple of Adam" and is situated in the dorsal region of the human neck. It secretes hormones that control protein synthesis and energy expenditure. The metabolic rate is regulated by these hormones, which in turn depend on cardiac output & energy expenditure. Thyroid hormones help to keep a constant metabolism by influencing the levels of other hormones. The two hormones that are generated by the adult thyroid glands are levothyroxine (T4) and triiodothyronine (abbreviated T3). T4 and T3 are the only thyroid hormones present in thyroid glands. Thyroid hormones are necessary for optimal body temperature generation, regulation, and maintenance. The only two thyroid hormones that are active are T4 and T3, which are exclusively present in thyroid glands. These hormones control energy balance, protein synthesis, body temperature, and cell division across the whole body. Although iodine shortages are rare, they can be dangerous since iodine is necessary for the creation of the thyroid hormones T3 and T4. Hypothyroidism, which may happen when there aren't enough of these hormones, is the reverse of hyperthyroidism, which can happen when there are too many of these hormones. There are several reasons why the thyroid may be overactive or underactive. A variety of ailments can be treated with various medications. Risks associated with thyroid surgery include radiation exposure, thyroid atrophy, and inadequate anaesthesia.

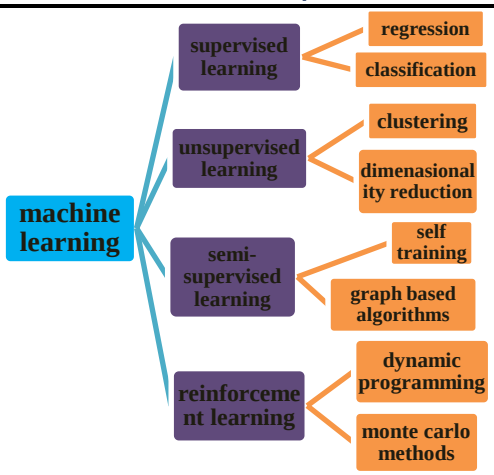

Fig.1:Types of Machine Learning

2.LITERATURE REVIEW

2.1Diagnosis of Thyroid disorder using Artificial Neural Networks

It is a major challenge for medical research since a correct diagnosis of sickness is necessary before any type of treatment can begin. In this study, we demonstrate how to identify thyroid problems (ANNs). Three different artificial neural network (ANN) techniques have been used to enhance the performance of the feed-forward neural network: the back propagation technique (BPA), the radial basis function (RBF) networks, and the training vector quantization (LVQ) networks. Networks are simulated in MATLAB, and a variety of measures, including diagnostic accuracy and training time, are used to evaluate the performance of the networks. By contrasting the results, the most trustworthy model for identifying thyroid problems may be chosen.

2.2:An investigation of neural networks in thyroid function diagnosis

We examine the viability of utilising artificial neural networks to diagnose thyroid disease. We examine the robustness of neural networks to variations in sample size using a cross-validation approach. Here, we demonstrate how to link neural networks to traditional Bayesian classifiers. Because logistic regression and other more traditional statistical methods can estimate posterior probability with high accuracy, neural networks may conduct classification more effectively than these methods. Additionally, it is demonstrated that neural network models can resist little variations in the data brought on by the sampling procedure. Despite the frequently skewed nature of medical diagnostic data, it has been shown that neural networks may nevertheless be an effective classification method in practical settings.

2.3A modern hybrid approach for thyroid diagnosis based on the artificial immune recognition system (AIRS)with blurry, weighted pre-treatment

It is crucial to have a complete awareness of thyroid gland working facts in order to diagnose thyroid illness. The thyroid gland's primary job is to help regulate the body's metabolism. The thyroid gland's release of thyroid hormone is the cause of this. The kind of thyroid condition depends on whether it produces too much thyroid hormone (hyperthyroidism) or too little thyroid hormone (hypothyroidism). A relatively new yet potent subset of artificial intelligence is artificial immune systems. One of the technologies proposed in this area so far is the AIRS mentioned earlier. Regarding the subjects at hand, Watkins has achieved a notable and intriguing success. This work aims to locate thyroid using a contemporary hybrid machine learning approach and classification system.

2.4:A survey on applying machine learning techniques for management of diseases

Databases and repositories have exponentially increased in size in recent years as a result of expanding scientific expertise and enormous data output. A rich data domain is the biomedical field. There is currently a broad spectrum of biomedical data available, from explanations of clinical illnesses to various kinds of biochemical data and picture instrument outputs. Due to the broad, dynamic, and subtle nature of the biomedical domain, manually collecting and translating biomedical trends from data into machine-understood information is a difficult operation. Biomedical pattern extraction can perform better using data mining.

This study describes how the fight against sickness could benefit from the use of data mining. In order to predict, diagnose, and treat serious illnesses including cancer, hepatitis, and heart disease, analysis of machine learning methods (MLTs) is essential. The artificial neural network, K-Nearest Neighbour, decision tree, and associative grouping processes are all covered in detail. The survey's findings give a complete picture of the MLT disease's current level of control. Depending on the disorder, problem solved, data and method employed, and results for the various applications, the precision ranged from 70% to 100%.

3:SYSTEM ANALYSIS

The process of establishing or upgrading systems, as well as the ideas and procedures used to construct these systems, is known as the Application Development Life Cycle (SDLC), also known as the Systems Architecture, Information Systems, and Network Engineering Software Development Life Cycle (SDLC). All software engineering product development methodologies are founded on the SDLC concept. These methods constitute the basis for creating and maintaining information systems for software development.

Existing System:

Thyroid illness is a challenging medical study issue that has a big impact on assessments and diagnoses. The thyroid's hormones are responsible for controlling metabolism. The two common thyroid diseases that induce thyroid hormone production to regulate body metabolism are hyperthyroidism and hypothyroidism. Techniques for data cleansing have been utilised to make the data simple to analyse and indicate the possibility that people have thyroid illness.

Disadvantages:

1.Energy level

2.Deficiency

3. Inhaling and exhaling.

Proposed System:

Machine learning is crucial to the prediction process, and material from UCI machine learning archives is employed in paper research as well as the classifications of thyroid disease models. To solve complicated learning difficulties, such as medical diagnostics and statistical tasks, it is necessary to maintain a good knowledge base that can be centred and utilised as a hybrid paradigm. We also offered several machine learning and thyroid diagnostic strategies. An estimated chance of a patient having thyroid illness was determined using machine learning algorithms, vector support machines, decision trees, and K-NN.

Advantages:

1. Prevents radioactive iodine and anti-thyroid medication's long-term dangers.

2. Immediately offers histology tissue for childbearing.

3.2 External Interface Requirements:

User-Interface: It is a user friendly JavaScript Graphical User-Interface.

Hardware Interfaces: The user-console interaction is accomplished through Python capabilities.

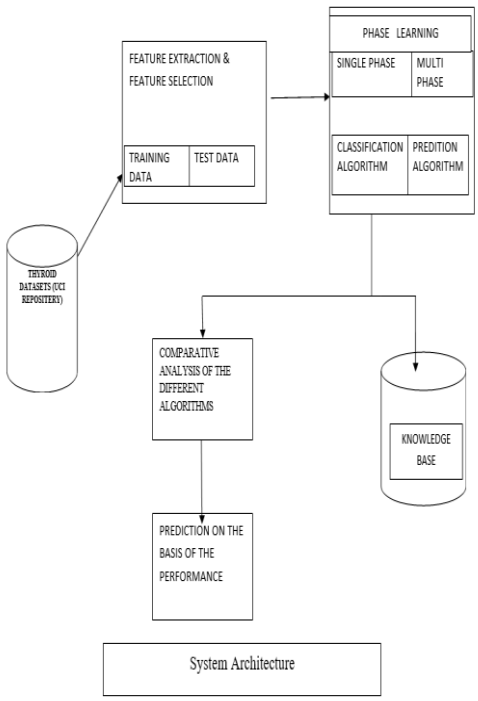

Fig.2:System Architecture

4.ALGORITHMS

PCA:

In order to achieve dimensionality reduction, Principal Component Analysis (PCA) is one of the most used unsupervised machine learning algorithms. We remove erroneous and inconsistent data from the thyroid dataset using PCA's dimensionality reduction feature and confusion matrix. Due to the repository's enormous dimensions and numerous dependencies, it aids in identifying hidden patterns.

Decision Tree:

The decision tree algorithm is a type of supervised learning. It can be applied to resolve classification and regression issues. A decision tree can serve as the representation for any Boolean function on discrete attributes. Decision Tree employs the Sum of Product form, also known as the Disjunctive Normal Form.

KNN:

KNN is a non-parametric, supervised machine learning technique that relies on feature similarity and proximity to classify or predict the grouping of individual data points. It operates on the principle that choosing the appropriate K value is a process known as parameter tuning, which is crucial to achieving improved accuracy.

Voting classifier:

Voting Classifier is a machine learning model that gains knowledge from an ensemble of many models and forecasts an output (class) based on the class that has the highest likelihood of becoming the output. Instead of building individual specialized models and determining each one's correctness, the goal is to build a single model that learns from various models (PCA, Decision tree,KNN and Gradient)and predicts output based on their cumulative majority of voting for each output class.

5.EXPERIMENTAL RESULTS

Fig.3:Home Screen

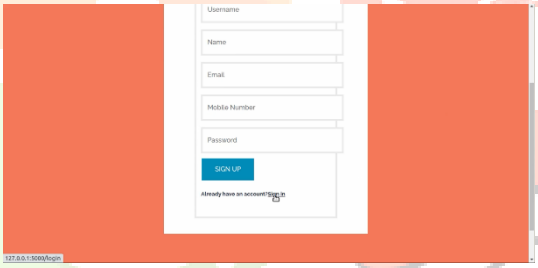

Fig.4:Sign-Up Screen

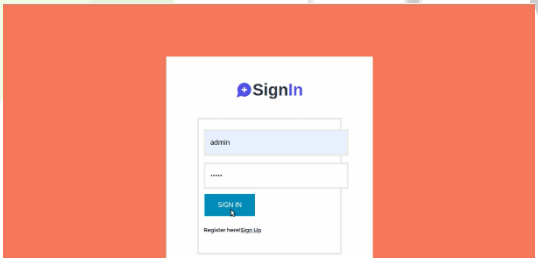

Fig5: Log-in screen

Fig.6: Sample data-Input Screen

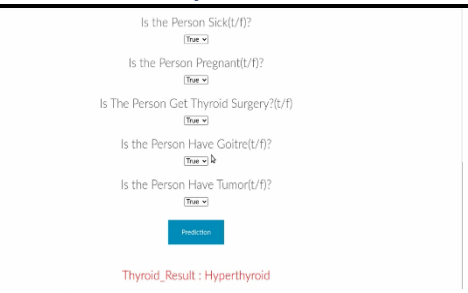

Fig.7: Hyperthyroid Result Screen

6.CONCLUSION

Investigations of unusual ml techniques for the diagnosis of thyroid issues are also ongoing. The invention and widespread use of numerous approachable analyses for the precise and knowledgeable diagnosis of thyroid ailment have occurred in recent years. According to the research, the two articles use different technologies, each with varied degrees of accuracy. In the academic literature, it has been extensively documented that neural networks are better than alternative approaches. Success with a decision tree or an assistance vector machine, on the other hand, is also a consideration. It is undisputed that doctors have made significant advancements in their capacity to recognize thyroid issues, however it is encouraged that consumers use fewer criteria. The diagnosis of thyroid illness may be made using exceptional ml approaches. When a patient has more distinguishing traits, a more extensive battery of time- and money-consuming health checks is necessary. The development of algorithms and thyroid disease prediction models is crucial to reducing the number of thyroid illness diagnostic criteria and saving patients' time and money.

7.REFERENCES

1.L. Ozyılmaz and T. Yıldırım, “Diagnosis of thyroid disease using artificial neural network methods,” in: Proceedings of ICONIP’02 9th international conference on neural information processing (Singapore: Orchid Country Club, 2002) pp. 2033–2036. 2.K. Polat, S. Sahan and S. Gunes, “A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted pre-processing for thyroid d isease diagnosis,” Expert Systems with Applications, vol. 32, 2007, pp. 1141-1147.

3.F. Saiti, A. A. Naini, M. A. Shoorehdeli, and M. Teshnehlab, “Thyroid Disease Diagnosis Based on Genetic Algorithms Using PNN and SVM,” in 3rd International Conference on Bio informatics and Biomedical Engineering, 2009. ICBBE 2009.

4.G. Zhang, L.V. Berardi, “An investigation of neural networks in thyro id function d iagnosis,” Health Care Management Science, 1998, pp. 29-37. Available: http ://www.endocrineweb.com/ thyroid.html, (Accessed: 7 August 2007).

5.V. Vapnik, Estimation of Dependences Based on Empirical Data, Springer, New York, 2012.

6.Obermeyer Z, Emanuel EJ. Predicting the future— big data, machine learning, and clinical medicine. NEngl J Med. 2016; 375:12161219.

7.Breiman L. Statistical Modeling: the two cultures. Stat Sci. 2001; 16:199-231.

8.Ehrenstein V, Nielsen H, Pedersen AB, Johnsen SP, Pedersen L. Clinical epidemiology in the era of big data: new opportunities, familiar challenges. Clin Ep idemio l. 2017; 9:245- 250 52

9.Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature. 2015; 521: 452-459.