A MACHINE LEARNING APPROACH TO PREDICT THYROID DISEASE ATEARLY STAGES OF DIAGNOSIS

AKSHATA BADADAL, AMRUTA REDDY
PG STUDENT DEPARTMENT OF CSE,ASSISTANT PROFESSOR CSE DEPARTMENT
SHARANBASAVA UNIVERSITY KALABURGI,INIDA

Abstract: The incidence of thyroid illnesses appears to be on the rise in India, according to recent figures. It is estimated that approximately one in ten people in India suffer from some form of thyroid disease. It is estimated that thyroid disease affects 42 million people around the world. Machine learning might make it easier for doctors to diagnose ailments while also reducing the amount of work they have to do. The incidence of thyroid illnesses appears to be on the rise in India, according to recent figures. It is estimated that approximately one in ten people in India suffer from some form of thyroid disease. There are around 42 million people around the world who suffer from a thyroid condition, according to estimates. Only a highly trained and experienced physician is capable of conducting a comprehensive analysis of the situation; in the absence of such an evaluation, a physician's prognosis of thyroid problems may not be accurate. Machine learning might make it easier for doctors to diagnose ailments while also reducing the amount of work they have to do.

Keywords: **TP, CNN, Machine Learning.**

1. Introduction

Thyroid hormones are made by a small gland in the neck called the thyroid. It might make too many or not enough of these hormones. Hypothyroidism happens when the thyroid gland does not make enough thyroid hormones. These hormones control the body's metabolism, which in turn affects how energy is used by the body. Without the proper amount of thyroid hormones, the body's normal functions begin to slow down, and the body undergoes changes on a daily basis, such as mood swings, happiness, sadness, fatigue, anxiety, stomach pain, feeling cold, weight gain, muscle pain, dry, thinning hair, and a slow heart rate. Hyperthyroidism is a condition in which the thyroid gland produces an excessive amount of a hormone known as thyroid hormone. Hyperthyroidism causes people to feel nervous, restless, unable to focus, hungry all the time, have trouble sleeping, itch, lose hair, feel sick, and throw up. For a diagnosis, you need to know your full medical history and take tests like the free T-4, T-3Test, Cholesterol test, and TSH Test. As these tests create a lot of data, and M-L can be used to find important features in a lot of data, this is a good match. Because of this, ML can be used along with medical science to accurately diagnose hypothyroidism disease. ML techniques have been developed, and ensembles are often used to make models as accurate as possible.

2. Literature Survey

[1]. L. Verma, "A hybrid data mining model to predict coronaryartery disease cases using noninvasive clinical data" , 2016.

Heart attacks and cardiac arrest are brought on by coronary artery disease (CA-D), which is brought on by atherosclerosis in the arteries. Angiography, an expensive, time-consuming, and highly technical invasive procedure, is used to diagnose CAD. Researchers are consequently compelled to develop alternative techniques, such as machine learning algorithms, that might diagnose the condition and gauge its severity using noninvasive clinical data.

We provide a novel hybrid strategy for CAD diagnosis that integrates risk factor identification using particle swam optimization (P-SO) search methodology, correlation-based feature subset (C-FS) selection, and K-means clustering techniques.

Then, to model CAD instances, supervised learning techniques such as C-4.5, the fuzzy unordered rule induction algorithm (FURIA), the multi-layer perceptron (M-LP), and the multinomial logistic regression (M-LR) are used.

[2]. M.A. Myszczynska "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases" 2020.

On a worldwide basis, effective medicines for neurodegenerative diseases are critically required.

The intricacy of the molecular mechanisms underlying neuronal degeneration, as well as the heterogeneity of the patient population, make developing early detection tools and effective therapeutics for these illnesses very challenging.

Machine learning, a subset of artificial intelligence, is assisting researchers, clinicians, and patients in addressing some of these issues.

In this paper, we discuss how machine learning may aid in early diagnosis, medical imaging interpretation, the discovery and development of innovative medicines, and other areas.

A common thread running across many machine learning applications is the automated extraction of relevant insights from a range of high-dimensional data sources, each of which gives a distinct viewpoint on sickness.

[3]. V. Yadav, "Thyroid Disease Prediction Using Machine Learning Approaches" June 2020.

This essay is being prepared to serve as a resource for research academics who want to engage in the field of thyroid illness prediction.

To anticipate and assess each machine learning technique's performance in terms of accuracy, three commonly used algorithms—logistic regression, decision trees, and k-nearest neighbour (k-NN) algorithms—were compared.

This work demonstrated how to use logistic regression, decision trees, and k-NN as a classification tool, as well as how to forecast thyroid illness intuitively.

This has included using the thyroid data set from the UC Irvin knowledge discovery from databases archive.

3. Objective

- The primary goal is to create a system.
- To predict the kind of thyroid condition a patient will have.
- To forecast thyroid illness using the fewest available criteria.
- Thyroid disease prediction accuracy

4. System Analysis:

Existing System:

When building a classification-based machine learning model, a train dataset is considered. This is supervised learning, and the model constructed from the train dataset will deliver the results. A decision tree is a reverse-ordered tree-based technique. The ID-3 algorithm builds the decision tree.It eliminates unnecessary data and improves categorisation. Age, gender, etc. are used in the decision tree method for thyroid patient data.

Proposed System:

Anyone is free to go at the thyroid dataset that is hosted on the Kaggle Machine Learning website. The vast majority of the entries in the database pertain to patients who have thyroid disorders and include all of the relevant patient information for such individuals. In this context, it is essential to ascertain whether or not the patient is allergic to any particular medication, whether or not the patient has undergone thyroid surgery in the past, the outcomes of any recent thyroid tests, as well as the patient's family medical history. In addition, it is essential to ascertain whether or not the patient has a history of thyroid disease in their family. These are also essential attributes owing to the fact that they make the examination of the patient with thyroid disease more easier and reduce the amount of time needed for the full examination that the doctor must undertake.

5. Methodology:

The first stage of the convolutional cerebrum association (CNN) algorithm involves converting the dataset entirely into a vector structure.

The second stage consists of the creation of the term, which does not include the addition of any features to fill the data.

The convolutional layer is produced as a consequence of the word establishment step.

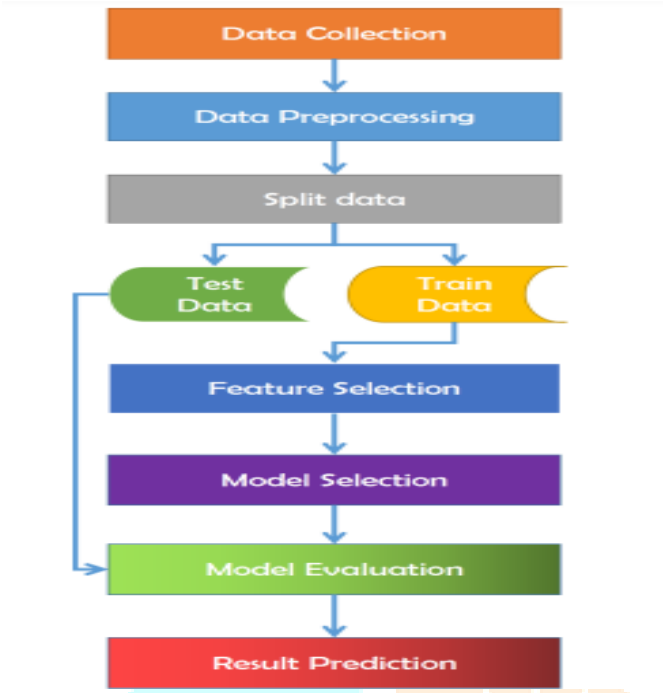
In the third stage, we use the maximum pooling approach on the convolutional layer that we have chosen to use as a commitment to the pooling layer.

In the fourth stage, using Max pooling, turn the dataset into a vector structure with an appropriate length.

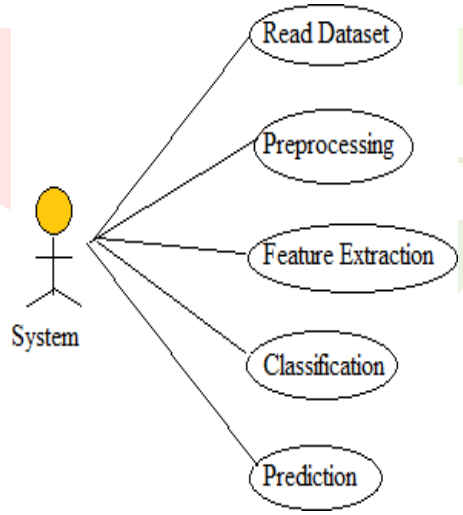
A connection has been made between the pooling layer and the complete linked cerebrum relationship.

The whole of the affiliation layer is brought together with the softmax classifier in Stage 5.

6. System Architecture:



7. Use Case Diagram



8. Result:

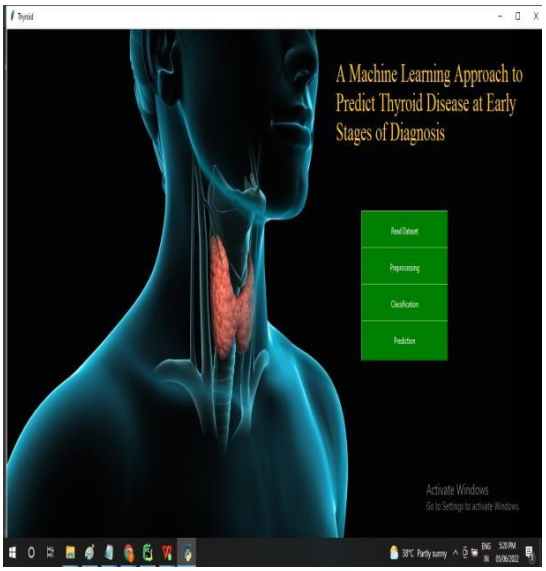


Fig 2: Main Screen

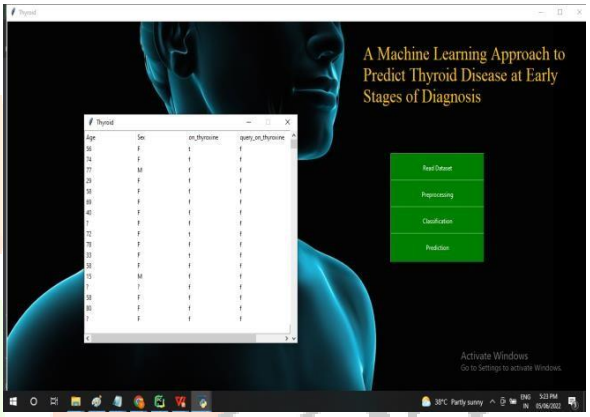


Fig 2: Load Dataset

	Unnamed: 0	Age	Sex	on_thyroxine	...	FTI_measured	FTI	TBG_measured	TBG
0	hypothyroid	72	M	f	...	y	10	n	?
1	hypothyroid	15	F	t	...	y	17	n	?
2	hypothyroid	24	M	f	...	y	0	n	?
3	hypothyroid	24	F	f	...	y	6	n	?
4	hypothyroid	77	M	f	...	y	44	n	?

[5 rows x 26 columns]

Fig 3: Preprocessing

Conclusion:

Thyroid scan Using machine learning to anticipate thyroid sickness is the goal of a project. Logistic regression helped us train our dataset and create more accurate thyroid disease predictions. The algorithm learns from the user's input whether a person is normal or has hyperhypothyroidism. A web app's backend (model) processes user-entered input and displays the output. Our goal was to provide the public an effective and accurate machine learning approach for illness identification programmes.

Acknowledgment

I thank my guide and all professors individuals for their expertise and assistance throughout all aspects of our study and for their help in writing the manuscript.

REFERENCES

[1] AnkitaTyagi and RitikaMehra. (2018)."Interactive Thyroid Disease Prediction System utilizing Machine Learning Techniques" distributed on ResearchGate.

[2] YongFeng Wang,(2020). "Examination Study of Radiomics and Deep-Learning Based Methods for Thyroid Nodules Classification utilizing Ultrasound Images" distributed on IEEEAccess.

[3] SunilaGodara,(2018). "Forecast of Thyroid Disease Using Machine Learning Techniques" distributed on IJEE.

[4] [4] Hitesh Garg,(2013). "Division of Thyroid Gland in Ultrasound picture utilizing Neural Network" distributed on IEEE.

[5] L. Ozyılmaz and T. Yıldırım,(2002). "Analysis of thyroid illness utilizing fake brain network strategies," in: Proceedings of ICONIP'02 ninth worldwide gathering on brain data handling (Singapore: Orchid Country Club, pp. 2033-2036).

[6] K. Polat, S. Sahan and S. Gunes,(2007) "A clever half and half technique in light of counterfeit safe acknowledgment framework (AIRS) with fluffy weighted preprocessing for thyroid sickness determination," Expert Systems with Applications,(vol. 32, pp. 1141-1147).

[7] F. Saiti, A. A. Naini, M. A. Shoorehdeli, and M. Teshnehlab,(2009) "Thyroid Disease Diagnosis Based on Genetic Algorithms Using PNN and SVM," in third International Conference on Bioinformatics and BiomedicalEngineering. ICBBE 2009.

[8] G. Zhang, L.V. Berardi,(2007) "An examination of brain networks in thyroid capability finding," Health Care Management Science,1998, (pp. 29-37.)

[9] V. Vapnik,(2012).Estimation of Dependences Based on Empirical Data, Springer, New York.

[10] Obermeyer Z,(2016). Emanuel EJ. Foreseeing the future — enormous information, AI, and clinical medication. N Engl ; (375:12161219).