

## An Analysis of Thyroid Disease Using Machine Learning Based Algorithm

Suresh Kumar Kashyap<sup>1</sup>, Neelam Sahu<sup>2</sup>

<sup>1</sup>Research Scholar, Dr. C.V. Raman University, Bilaspur (C.G.) India.

<sup>2</sup>Professor, Dept. of IT & CS, Dr. C.V. Raman University, Bilaspur (C.G.) India.

### ABSTRACT

Thyroid organs help in regulating the thyroid hormone metabolism that is secreted in our body. The thyroid gland is a butterfly structure found in the neck of the human body. The thyroid gland produces hormones that are very important for our body. They are responsible for many specific functions, such as being overweight or underweight, causing or causing secrecy and unhappiness, standard of living and heart problems. A strategic suggestion has been made for thyroid disease. However, the treatment of thyroid infection has become easier now. Hypertension and diabetes are treated by a large proportion of patients such as older patients. Mainly this is a condition of which there are two types first hypothyroidism and second hyperthyroidism. Our research paper is based on the analysis of machine learning based classifiers for thyroid disease. The classification algorithms used in this paper measure diseases using disease datasets that provide accurate prediction estimates. This paper describes how to identify a better method of classification for thyroid disease, and concludes that ANN gives better accuracy than the decision tree technique.

**Keywords:** Classification, Disease Prediction, T3, ANN, Accuracy.

### I INTRODUCTION

Today, machine learning continues to play the most important role in health-related fields (a subset of artificial intelligence), among other areas, including the development of new medical procedures, and the basis of patient health and the treatment of chronic diseases. Including the handling of information. Machine learning has many important applications in the healthcare industry. Machine learning is now being used to streamline administrative processes and personalize medical treatment in hospitals. There are also many opportunities for healthcare as machine learning models have a high degree of predictive analytics capabilities. Machine learning has models that can predict chronic disease such as heart disorders, infections and intestinal diseases. Artificial Intelligence (AI) is going to be one of the biggest transformative technologies and promoters for the human society of the 21st century. Healthcare

#### (a) Applications of Machine Learning

- (i) The Diagnosis & Prediction of Heart Disease
- (ii) The Prediction of Diabetes
- (iii) Prediction of thyroid disease
- (iv) The Prediction of Liver Disease
- (v) ML Applications in Surgery
- (vi) The Detection of Cancer
- (vii) The Discovery of New Drugs
- (viii) The Personalization of Treatment

#### (b) Thyroid disease and Types-

The Thyroid Diseases are majorly classified into two types-

- (i) **Hypothyroid-** Hypothyroid is a type of thyroid which means the insufficient production of hormones of thyroid gland. This insufficient growth can cause lot of problems in human body like too much of weight gain, dry skin and dry hair, heavy menstrual

periods etc. These symptoms are different from one person to another.

- (ii) **Hyperthyroid-** The term hyper refers that too much or heavy. So hyperthyroid indicates that the massive growth of hormones of thyroid glands. This hyperthyroid is often called as overactive thyroid. This may cause too much of weight loss even though the sufficient intake of food, increased heart rate, shorter menstrual periods etc.

### II REVIEW OF LITERATURE

In reviews, I have studied many research papers that are related to my topic. Which are as follows-

S.Saru et al. (2019), has done study work on "analysis and prediction of diabetes using machine learning", this study is focus on a classification model with boosted accuracy to predict the diabetic patient. In this model, we have employed different classifiers like Decision Trees, KNN and Naïve Bayes. Sunila Godara et al. (2018) done work on Prediction of Thyroid Disease Using Machine Learning Techniques logistic regression and SVM are used to predict Thyroid.

S.Umadevi et al. (2017) have performed a research on "Applying Classification Algorithms to Predict Thyroid Disease" the main research work focus on The accuracy of the classification of the thyroid dataset based on the features selected by each feature selection algorithms. Many researchers worked on disease prediction using machine learning concept for many types disease.

### III METHODOLOGY

(a) **Data Set Description:**-In this research work used data set for experimental purpose we can easily download from UCI machine learning Repository

that was used for implementation with multiple instances of 23 independent attribute and 1 dependent attribute.

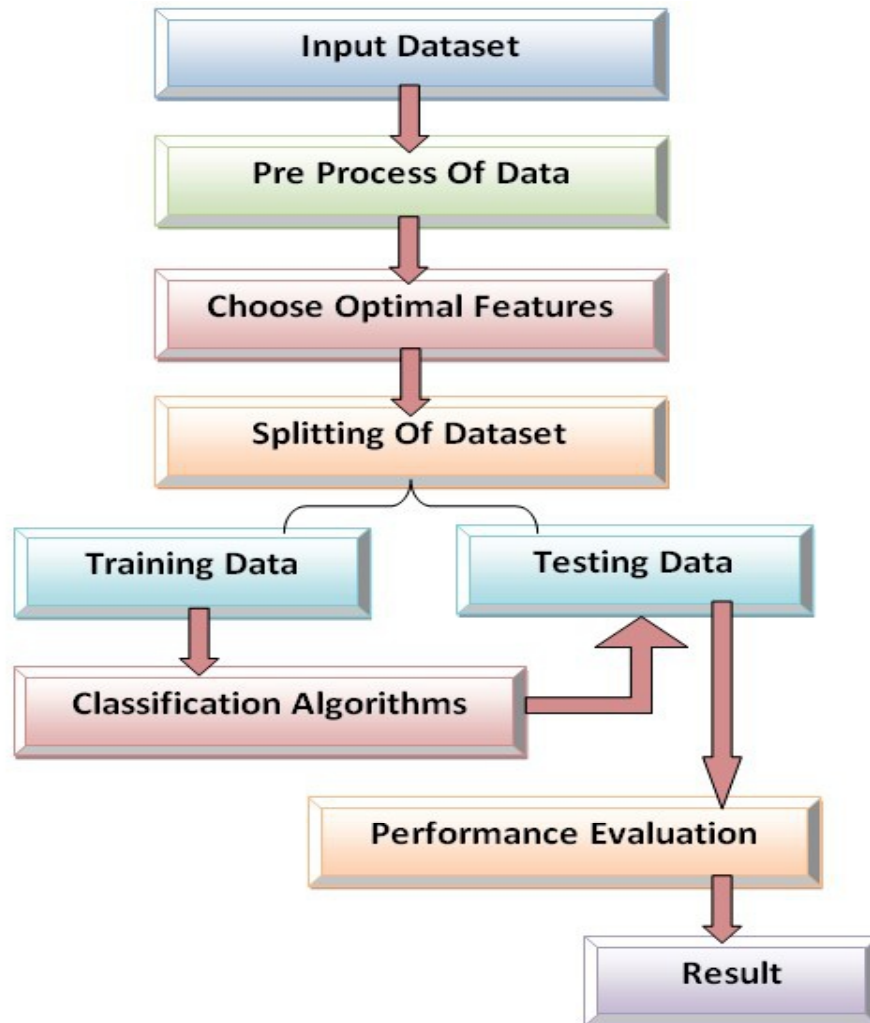
(b) **Thyroid Disease data set Attribute description**

**Table 1**  
**Thyroid Disease data set Attribute description**

<b>Data Description SN</b>	<b>Attribute Name</b>	<b>Value Type</b>
1	age	Continuous
2	sex	M,F
3	on thyroxin	F,T
4	query on thyroxin	F,T
5	on anti thyroid medication	F,T
6	sick	F,T
7	pregnant	F,T
8	thyroid surgery	F,T
9	query hypothyroid	F,T
10	query hyperthyroid	F,T
11	lithium	F,T
12	goiter	F,T
13	tumor	F,T
14	hypo pituitary	F,T
15	TSH measured	F,T
16	TSH	continuous
17	T3 measured	F,T
18	T3	continuous
19	TT4 measured	F,T
20	TT4	continuous
21	T4U measured	F,T
22	T4U	continuous
23	FTI	continuous

The framework of our proposed work is shown in the figure 1. In which there are many parts such as input datasets, preprocess of data sets, forming of training data

and testing data, fitting classifiers, comparison metrics etc.



The proposed system consists of many parts, that are following:-

- (a) **Data Preprocessing:** This is a very important step in any process. The data preprocessing phase involves several tasks, such as cleaning, normalization, selection, etc. Raw data has low signal to noise ratio, missing values and inconsistency that affects DM results,
- (b) **Feature Selection:** Feature selection methods used to identify features most relevant to the classification step and can be broadly classified as either subset selection methods or ranking methods. Mainly it is of two types.
  - (i) Wrapper model
  - (ii) Filter Model
- (c) **Model Evaluation:** It is a basic and important step of this system process. Model evaluation is an effective process of using various evaluation metrics to

understand the performance of the machine learning model along with its strengths and weaknesses. Model evaluation is important for assessing the efficacy of a model in the early research stages, and it also plays an important role in model monitoring. It has two type of evaluating method in data mining:-

- (i) Cross- validation
- (ii) Hold-out
- (d) **Machine Learning based Algorithm's-**
  - (i) **Artificial Neural Network (ANN)** –The ANN is designed to mimic the way the human brain works. An ANN consists of an input layer, several hidden layers, and only one output layer. Units in all layers are fully connected. An ANN consists of a very wide number of units and performs an arbitrary task theoretically predicted

by all units; therefore, it has strong fitting capability especially for non-linear functions.

- (ii) **Decision tree** - The structure of a decision tree is similar to the structure of a normal tree, which is a tool for analyzing predictive modeling, which is being applied in many different fields these days. It is most effectively used for supervised learning. Decision trees have a special non-parametric supervised learning method that we use for both classification and regression tasks. Its main function is to build a model that serves to predict the value of a target variable by learning simple decision rules that are inferred from the data features. Decision rules are usually

in the form of if-then-else statements, that is decision making statement.

#### IV RESULT

The thyroid data set contains information related to the patient's health characteristics such as age, gender, etc. We used 250 patient records with 9 selected characteristics for performance. Table 2 describes the accuracy of both classification techniques used for the thyroid dataset for the selected features. We can see table 2, where Artificial Neural Network gives the best result as compared to Decision Tree Classifier.

**Table 2**  
**Accuracy for Thyroid Dataset**

Used Classifier	Accuracy
ANN	97.1%
Decision Tree	93.7%

#### V CONCLUSION

Disease prediction with the help of machine learning and related classification techniques has today led to new ways of advancing technology with automated predictions and rapid treatment of patients. In this paper, we study the results obtained using two classification techniques for the prediction of thyroid disease. Mainly the classification model is prepared on the basis of training data and test data though prediction is done at different stages. This research work focuses on an accurate prediction model constructed by selecting an effective and important classification technique. Effective classifiers perform the task of comparing classification algorithms based on accuracy and confusion matrix and display the results.

#### REFERENCES

- [1] Agrawal, Komal et al.(2017), "Thyroid Prediction System using Auto Associative Neural Network", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 6, Issue 4, ISSN (Online): 2278 – 8875.
- [2] Dr. Vizhi, Kayal et al.(2020), "Diabetes Prediction Using Machine Learning" International Journal of Advanced Science and Technology Vol. 29, No. 6, pp. 2842 – 2852.
- [3] Geetha, K. et al.(2016), "An Empirical Model for Thyroid Disease Classification using Evolutionary Multivariate Bayseian Prediction Method", Global Journal of Computer Science and Technology: E Network, Web & Security Volume 16 Issue 1 Version 1.0. ISSN: 0975- 4172.

- [4] Godara, Sunila et al.(2018), "Prediction of Thyroid Disease Using Machine Learning Techniques", International Journal of Electronics Engineering Volume 10 , Issue 2 pp. 787-793, ISSN: 0973-7383.
- [5] Hegde, Vikram V et al.(2016), "Automated Prediction of Thyroid Disease using ANN", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, Special Issue 10, ISSN(Online) : 2319-8753.
- [6] <https://builtin.com/artificial-intelligence/machine-learning-healthcare>.
- [7] <https://www.kolabtree.com/blog/top-applications-of-machine-learning-in-healthcare/>.
- [8] <https://archive.ics.uci.edu/ml/datasets.php>
- [9] Indoria, Priyanka et al.(2018)," A Survey: Detection and Prediction of Diabetes Using Machine Learning Techniques", International Journal of Engineering Research & Technology (IJERT) Vol. 7 Issue 03,ISSN: 2278-0181.
- [10] Sisodia, Deepti et al.(2018)," Prediction of Diabetes using Classification Algorithms", International Conference on Computational Intelligence and Data Science (ICCIDS).