**Thyroid Disease Classification Using Machine Learning Algorithms**

**ABSTRACT:**

With the vast amount of data and information difficult to deal with, especially in the health system, machine learning algorithms and data mining techniques have an important role in dealing with data. In our study, we used machine learning algorithms with thyroid disease. The goal of this study is to categorize thyroid disease into three categories: hyperthyroidism, hypothyroidism, and normal, so we worked on this study using data from Iraqi people, some of whom have an overactive thyroid gland and others who have hypothyroidism, so we used all of the algorithms. Support vector machines, random forest, decision tree, naïve bayes, logistic regression, k-nearest neighbors, multilayer perceptron (MLP), linear discriminant analysis. To classification of thyroid disease.

**EXISTING SYSTEM:**

At present, diseases have become dangerous and rapidly spread, and their exploration and diagnosis require a great deal of time and effort. The correct and accurate diagnosis of the disease early has become one of the problems that the health system suffers from it. The critical role of early and correct diagnosis of the disease, including thyroid disease, is vital because it increases patient treatment opportunities and reduces mortality. Among the vast amount of clinical data, early diagnosis is a challenging task. Today the machine learning has had impressive and good results in many sciences. Hence, it had a prominent and valuable role in diseases, so this study used machine learning algorithms with thyroid disease. To detect and classify thyroid disease into three types hyperthyroidisms, hypothyroidism, and normal.

**Disadvantages**

1. Low accuracy
2. high-power computing

**PROPOSED SYSTEM:**

In this paper author employing various machine learning algorithms such as SVM, Naïve Bayes, Decision Tree, Random Forest, KNN and MLP to predict thyroid disease. Each algorithm gets trained on thyroid dataset which contains 3 different classes such Normal, Hyperthyroid and Hypothyroid. All algorithms performance is evaluated in terms of accuracy, precision, recall, FSCORE and confusion matrix and in all algorithms Random Forest and Decision Tree giving best accuracy.

**Advantages**

1. access to a significant amount of data

2. risk factors influence disease progression

3. more accurate

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

# Processor - intel i3 or above

* Speed - 1.1 Ghz
* RAM - 4 GB or above
* Hard Disk - 500 GB or above
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

Operating System - Windows10 or above

Programming Language - Python