## A Contemplative Analysis of Thyroid Disorders based on ECG Signal and Thyroid Test Measure

M. Deepika<sup>1\*</sup>, K. Kalaiselvi<sup>2</sup>

<sup>1</sup>Research Scholar, Vels Institute of Science, Technology& Advanced Studies (VISTAS), Pallavaram, Chennai-117, India. <sup>2</sup>Professor and Head, Vels Institute of Science, Technology& Advanced Studies (VISTAS), Pallavaram, Chennai-117, India. E mails: <sup>1</sup>deepika.rbp@gmail.com, <sup>2</sup>kalairaghu.scs@velsuniv.ac.in

## Abstract

Hypothyroidism can influence the heart and circulatory framework in various manners. Inadequate thyroid hormone eases back the pulse. With the goal of forecasting hypothyroid, by utilizing thyroid measures turns out to be high intricacy in order of Hypothyroid and Hyperthyroid. Some of the common clinically most relevant findings explaining hyperthyroidism as well as hypothyroidism are the anatomic symptoms of the disease of Thyroid disease. Data Mining plays a vital role in the process of prediction and classification. The proposed work utilizes the information mining methods, such as the characterization and expectation by adopting parallel tree calculation. A Trust Region method for Nonlinear Minimization streamlining calculation is incorporated in the count of QRS point discovery in electrocardiogram (ECG) signal. This proposed model is re-enacted on the MATLAB R2014b. This analysis is used to assess the relationship among outright thyroxine (T4) and thyrotropin (TSH) levels with ECG parameters.

Keywords: Classification, Data Mining, ECG, Heart rate, Hypothyroid, Hyperthyroid, QRS duration and Thyroid test measures.

Received on 20 July 2020, accepted on 25 August 2020, published on 26 August 2020

Copyright © 2020 M. Deepika *et al.*, licensed to EAI. This is an open access article distributed under the terms of the <u>Creative</u> <u>Commons Attribution license</u>, which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.1-7-2020.166006

\*Corresponding Author: Email: deepika.rbp@gmail.com

## 1. Introduction

#### 1.1 Risk of cardiovascular disease

Subclinical thyroid dysfunction is defined as a typical condition, where serum thyrotrophic hormone (TSH) levels is found to be below (hyperthyroidism) or above (hypothyroidism) the reference interval with traditional free thyroxin (FT4) levels. It is commonly prevailing in older girls. Subclinical thyroid pathology is related to alterations in internal organ hemodynamics, like impaired internal organ ability, raised vital sign, pulsation cardiovascular disease, raised left cavum mass, and beat pathology [1]. Data from observational examinations on subclinical thyroid dysfunction, in connection to the complexity of cardiovascular ailments (CVD) are uncertain. Subclinical hyperthyroidism was accounted to be connected with atrial fibrillation (AF), coronary illness, generally speaking stroke, and heart failure in several studies.

Thyroid hormones affect the cardiac performance immediately but not through direct changes in peripheral movement. On the other hand, it has a repercussion on myocardial oxidative metabolism and its association with the cardiac function. Patients holding a history beyond Thyroidectomy for thyroid most cancer provide a unique style to research the cardiac consequences of hypothyroidism. Some cases without cardiac disorder had been examined in the hypothyroid part and moreover after 4-6 weeks they underwent Euthyroid conditions.



Altered thyroid fame exerts a prominent effect on the middle. People with hypothyroidism or hyperthyroidism affirmed different alterations in ECG. Moreover, little is understood approximately, as how permutations in thyroid hormone varying inside the ordinary range have an effect on electric powered sports of the coronary heart within the regular populace.

ECG parameters are scaled at the standard 12-lead electrocardiogram. In [2] we observed a tremendous direct affiliation among serum upload as much as T4 degree and pulse rate in men, also an U-form affiliation among T4 and PR period between men and women. TSH degree variation is emphatically related with QRS period interval in men, while a U-form affiliation among QRS and TSH variation is observed in women.

## 1.2 Electrocardiographic (ECG)

Thyroid hormones may cause electrophysiological and inotropic impacts on the heart and may adjust the peril of making atherosclerosis. People with clear hypothyroidism shows variations in ECG like sinus bradycardia, moo plentifulness QRS buildings, QT interval prolongation, and the changes in T wave morphology, whereas patients with hyperthyroidism show atrial arrhythmias, for instance, atrial sway, sinus tachycardia and atrial fibrillation. Strikingly, both deferred and condensed QT between times has been represented in hyperthyroidism. Though electrocardiogram irregularities in patients with hypothyroidism and hyperthyroidism of tolerably small amount is considered, the collections in thyroid hormone levels inside the common place affect the electrical movement of the heart within the comprehensive community. This analysis is used to assess the relationship among outright thyroxine (T4) and thyrotropin (TSH) levels with ECG parameters (beat, PR intervals, QRS term, QT between times, and JT between times) in a master test of the common US individuals [3].

Table 1. Amplitude values for normal ECG

| S.No. | Interval        | Duration      |
|-------|-----------------|---------------|
| 1.    | PR wave         | 0.12-0.2 sec  |
| 2.    | QT wave         | 0.35-0.44 sec |
| 3.    | ST wave         | 0.05-0.15 sec |
| 4.    | P wave interval | 0.11 sec      |



Figure 1. Detection of ECG waveform.

The above table1 represents the ECG signal amplitude for the PR, QT, ST, and P wave with the time duration. The PR, QT, ST, and P wave time duration calculated from the input signal.

## 1.3 Data mining

It is the way of deriving patterns in tremendous datasets counting procedures converging through machine learning and database systems. It is considered as an intrigue subspace of computer science and measures handling data elimination through efficient methods from a data set. It also modifies the raw data into a comprehensible format for processing further [4].

The actual data mining process is the semi-computerized or programmed study of vast amount of data to extricate in the past unknown, unwanted patterns in businesses of statistics records like cluster evaluation, odd records like anomaly detection, and dependencies such as sequential sample mining and affiliation rule mining. In most extreme cases it also includes adaptation of database approaches beside spatial files. These styles can then be noticed as a framework for the entering records, and may also be used in additional research or, as a case, in framework perusing prescient analytics.

Classification is a more complex knowledge mining method that forces to gather more attributes in combination into discernible categories, which may also draw further conclusions, or serve some function. Finally, comparing data on particular person or customers financial backgrounds, we might be able to classify them as "low," "medium," or "high" credit score degrees. The section 2 discusses about the literature survey, section 3 discusses about the Materials and strategies. Finally, section 4 examines about concerning the consequence and conclusion.



## 2. Literature Review

K. Vijayalakshmi et al., 2018 [5] rule objective is to develop a model wise thyroid Prediction System using Big information and data mining techniques. This structure can discover hidden data models and connections identified with the thyroid sickness from a chronicled thyroid database. A Naïve Byes technique to anticipate patients with Hypothyroid is adopted. The classifier used exhibited its execution in predicting with the best results in arrangements of exactness and least execution time.

Yan Wang et al., 2019 [6] proposed the hereditary premise of heart infections has continued gradually through linkage examinations, on account of a lot littler impact size inferable from regular changing variations in complex issue. Hypertension positions as the main source of grimness and mortality around the world, adding to CAD, atrial fibrillation, and cardiovascular breakdown. With the objective that the data around the genetic qualities of hypertension or BP may be a critical figure 1 within the analysis, control, and treatment of these heart illnesses.

Rafn Benediktsson et al., 2015[7] proposed the term, compensated hypothyroidism (CH) is employed to explain selected combination of thyroid function tests, Especially a raised thyrotrophin (TSH) attention (usually between 3-20 mU/l) in association with normal serum thyroxine (T4). Bernadette Biondi et al., 2002 [8] proposed preload the hemodynamic force given to the ventricular wall during filling and, thus, can be equated with ventricular end-diastolic wall stress. This contributes to the determination of ventricular end-diastolic volume and significantly modulates myocardial performance *i.e.* it governs the extent and velocity of wall shortening.

Michael Preuss et al., 2013 [9] proposed an ongoing genome-wide affiliation considers (GWAS) of myocardial dead tissue (MI) and different types of coronary supply route ailment (CAD) have prompted the disclosure of in any event 13 hereditary loci. Notwithstanding the impact size, capacity to recognize affiliations is to a great extent driven by test size. Along these lines, to boost the opportunity of discovering novel weakness loci for CAD and MI, the Coronary artery disease Genome-wide Replication and Meta-investigation (cardiogram) consortium was framed.

Hypothyroidism is common worldwide, particularly in iodine-deficient areas like India, characterized by way of a cluster of scientific manifestations attributable to thyroid hormone deficiency or more hardly ever from their impaired task at tissue stage.

Hypothyroidism causes symptoms that cut back the practical status and high quality of lifestyles. Cardiovascular effects of the thyroid hormones are rather powerful and the cardiac abnormalities related to thyroid disorder have attracted an excessive amount of analytical effort.

Hypothyroidism (underactive thyroid) is a state where the thyroid gland fails to produce sufficient hormones. Hypothyroidism may not produce noticeable signs within the early phases. Over time, untreated hypothyroidism may cause quite a lot of problems, such as obesity, joint pains, and infertility and so on.

Classification Analysis is used to classify certain information in several categories. Classification is similar to clustering in that it additionally segments information records into other segments known as categories. In classification, the structure or identity of the knowledge is understood.

## 2.1 Aims and Objectives

The objectives of this study are as follows:

- To contemplate ECG profile, i.e., pulse (beats/min), P-R interim (sec), voltage of QRS complex (mv), and S-T fragment changes (T-wave reversal) (mv) in recently analyzed hypothyroid patients, hypothyroid patients taking treatment for >50 years and control group.
- To compare the effect of heart beat >=69 and QRS detection using the range specified >= 120ms are the conditions for classifying the hypo and hyper thyroid data.
- The data mining used to order the thyroid stimulating hormone (TSH) levels below (hyperthyroidism) or above (hypothyroidism) the reference interim with typical free thyroxine (FT4) levels.
- The binary tree algorithm used to classify the Hypothyroid and Hyperthyroid with the other TSHLevel, Hyperthyroidism, FT4Levels.

Electrocardiographic changes, for example, bradycardia, low-voltage, and changing degrees of heart beat are usually perceived in hypothyroid patients [5].

Considering the work done using CVS appearances of hypothyroidism such as diminished add up volume, contractility, heart rate (all driving to diminished cardiac yield), expanded systemic vascular resistance (driving to expanded diastolic blood weight), and expanded capillary porousness (driving to pericardial radiation). These changes can be effectively recognized by basic investigative procedure like electrocardiogram (ECG) and complications creating at later stages of hypothyroidism can be anticipated [10].

Thus, the present analysis is undertaken to analyze ECG profile, for example pulse (beats/min), P-R interim (sec), voltage of QRS complex (mv), and S-T section changes (T-wave reversal) (mv) in recently analyzed hypothyroid subjects and hyperthyroid subjects who were at that point getting treatment for more than 50 years and to see whether there was any improvement in these parameters with thyroxine supplementation treatment.

## **3. Material and Methods**

#### 3.1 Material:



The test signal taken from the PhysioBank is a huge and growing archive of physiological information. The test knowledge signal from the PhysioNet-> BankATM-> MIT-BIH Arrhythmia Database (mitdb). The training knowledge analyzed from the associations of genetically lowered thyroid stimulating hormone (TSH) levels, hyperthyroidism, hypothyroidism, and higher FT4 levels with odds ratio (OR) of atrial fibrillation (AF), coronary artery disease (CAD), and ischemic stroke and its subtypes in accordance with knowledge from the Atrial Fibrillation Consortium (AFGen), Coronary Artery Disease Genomewide Replication and Meta-examination in addition to The Coronary Artery Disease consortiums 1000 Genomesbased (CARDIoGRAMplusC4D), GWAS and MEGASTROKE consortia.

## 3.2 Methods:

Data Reduction is the foremost process, which aims to decrease the data representation by targeting precision for speedy response that can produce quick summarised answers to queries from exceptionally huge databases. Isolating data for preparing and testing sets is an imperative portion of assessing information mining models.

Primarily, when isolating a data set into a preparing set and testing set, most of the data is utilized for preparing, and a small portion of the data is utilized for testing. Data analysts randomly test the data to assists certainty that the testing and preparing sets are comparative. By utilizing comparative data for preparing and testing that can minimize the impacts of data discrepancies [11].



Figure 2. Block diagram of Hypothyroidism and Hyperthyroidism classification

The above figure 2 shows the test sign taken from the ECG database with the states of the cardiovascular sickness (CVD) signal. The preprocessing of signal expels the commotion contained in the info signal. The component choice requires the QRS location from the ECG signal. The database contains information identified with the other thyroid estimates taken from the thyroid test. The tree order procedure orders the test and preparing information given as the input. The QRS interim planning seconds will be equivalent or more prominent than 120ms. The heart beat thumps/minutes. With abundance (66-69) thyroid hormones in the body, and the pulse greater than 90 beats for every moment (tachycardia) is considered as normal, even when a person is very still or idle and when person is sleeping. Additionally, the ordinary increment in pulse during activity is overstated. Quick pulse detection is one of the most widely recognized indications of hyperthyroidism [12].

## 3.3 Pre- processing techniques

The input test signal details such as the 3600 columns, frequency- 360Hz, and the sampling interval 0.002 sec. The ECG signal can be converted into the wave decomposition. The 3- level decomposition is implemented to remove the noise. The output is evaluated using in the form of Inverse Fourier Transform.

#### 3.3.1 QRS wave calculation:

The QRS interval obtained using the PR interval (sec)  $0.12\pm0.00$ , the maximum peak of the PR wave, QT wave, ST wave and P wave interval [13].

#### 3.3.2 Heartbeat calculation:

The duration of sample in seconds= the total sample /360sec

The duration in minutes= Duration of sample in seconds/60.

Beat per Minute = Total number of peaks in the signal / duration in minutes.

#### 3.3.3 Training Data:

The training data was prepared with the help of TSH and FT4 levels. The results were strong when considering the lead single-nucleotide polymorphism (SNP) of every locus. When barring the 3 loci associated with both hyperthyroidism and hypothyroidism, the ORs of AF (utilizing information from Atrial Fibrillation Consortium 2018 GWAS dataset) were 1.07 (95% CI, 1.04–1.11;

P= $3.6 \times 10-5$ ) for hyperthyroidism and 0.99 (95% CI, 0.95–1.03; P=0.59) for hypothyroidism, as shown in table 2 and figure 3.



Table 2. Details required for Feature Extraction.

| S.No. | Training data parameters    |  |
|-------|-----------------------------|--|
|       |                             |  |
| 1.    | Atrial fibrillation         |  |
| 2.    | Coronary artery disease     |  |
| 3.    | Ischemic stroke             |  |
| 4.    | Intracerebral hemorrhage    |  |
| 5.    | Subarachnoid hemorrhage     |  |
| 6.    | Heart failure               |  |
| 7.    | Aortic valve stenosis       |  |
| 8.    | Abdominal aortic aneurysm   |  |
| 9.    | Thoracic aortic aneurysm    |  |
| 10.   | Deep vein thrombosis        |  |
| 11.   | Pulmonary embolism          |  |
| 12.   | Peripheral arterial disease |  |
| 13.   | Hypertension                |  |



Figure 3. Pre- processing of the ECG signal



Figure 4. Components of the Feature Extraction Process

| Table 3. The training | data obtain from the below |
|-----------------------|----------------------------|
|                       | table.                     |

| S.No | Outcome    | TSH<br>Levels | Hyper<br>thyroidi<br>sm | Hypo<br>thyroidi<br>sm | FT4<br>Levels |
|------|------------|---------------|-------------------------|------------------------|---------------|
| 1.   | Feature    | 0.65-         | 0.60-                   | 0.86-                  | 0.64-         |
|      | Extraction | 1.09          | 1.08                    | 1.10                   | 1.107         |

For the first time, this study uses an MR approach to systematically investigate the relationship between thyroid function and dysfunction with a wide range of CVD outcomes. The results showed that genetically decreased TSH levels and hyperthyroidism were hugely related with higher chances of AF. We found no factually noteworthy and reliable affiliations of TSH levels, hyperthyroidism, or hypothyroidism,[14] or FT4 levels with the other CVD results, as shown in table 3 and figure 4.

In spite of the fact that there was suggestive proof of conceivable affiliations between diminished TSH levels and higher chances of cardioembolic stroke, hyperthyroidism and lower chances of thoracic aortic aneurysm, hypothyroidism and higher chances AF, CAD, and hypertension, and expanded FT4 levels and lower chances of fringe blood vessel infection. Be that as it may, those suggestive affiliations were watched in one dataset (one of the hereditary consortia or UK Biobank) or were not steady in affectability analyses.

Findings: It support prior empirical studies associating shrunken thyrotrophic hormone levels and adenosis to increased risk of AF4,13–17 furthermore as results from a pair of recent studies associating genetically higher thyrotrophic hormone levels to bring down the risk of AF.29,30. However, a personal patient data analysis of 3085 participants (including 2574 AF cases) from eleven cohorts showed solely a suggestive association of lower tsh levels among the reference range [15].

## 4. Training Algorithm

## 4.1 Trust-region method (TRM)



For solving nonlinear programming (NLP) issues, TRM is considered as primitive numerical optimization strategy. It works by beginning with diagram a locale over the current remarkable solution, wherein a certain form (generally a quadratic show) can to a degree inexact the genuine objective work. Later Trust Region Method progress consistently, with the form that depicts the location.

#### 4.2 Trust-region

In most cases, the trust-region is said to be a spherical region of radius where the trust-place sub problem lies.

#### 4.3 Trust-region Sub problem

On adopting the quadratic method to summarize the first objective function, the optimization problem is decreased for solving trust-region sub problems.

$$\min m_k(p) = f_k + g_k^T p + \frac{1}{2} p^T B_k p$$
(1)

$$s. t. \|p\| \le \Delta_k \tag{2}$$

Here, they believe region radius, is the angle at current point and is the hessian (or a hessian estimation). It is basically used to discover the solution to the trust-region sub issue.

# 4.4 Predicted reduction and Actual reduction

The basic hidden trust-area technique is to refresh the size of the trust-district at each emphasis. The present emphasis makes a palatable decrease, so we may misuse our model more in the following cycle by setting a bigger value. On the off chance, we just accomplished a constrained improvement after the present emphasis, the sweep of the trust-area at that point does not have any expansion, or at the most pessimistic scenarios, we may diminish the size of the trust by altering the range to a lesser incentive in order to check the models validity [16].

$$\rho_{k} = \frac{f(x_{k}) - f(x_{k} + p_{k})}{m_{k}(0) - m_{k}(p_{k})}$$
(3)

Whether to make a progressively driven stride or an increasingly traditionalist one is relied upon the proportion among the genuine decrease, picked up by evident decrease in the first target work and the anticipated decrease expected in the model work. Observational edge estimations of the proportion help in deciding the size of the trust-region.

#### 4.5 Trust Reduction Algorithm

Even before we implement the true, with-region set of rules, initially we must decide various parameters. The higher bound for the measure of true with region and the limit ranges for comparing the goodness of the quadratic form, results for deciding the trust-vicinity's length within the consequent generation. A typical set for these values are,

$$0 = < \eta_1 <= \eta_2, \eta_2 = 0.25 \text{ and}$$
  
$$\eta_3 = 0.75, t_1 = 0.25, t_2 = 2.0$$

#### 4.6 Pseudo-code

Set the starting point at x<sub>0</sub>, set the iteration number k=1

**for** k=1,2

Get the improving step by solving trust-region subproblem ()

Evaluate  $p_k$  from equation (4.3)

if 
$$\rho_k < \eta_2$$
  
 $\Delta_{k+1} = t_1 \Delta_k$   
else

if  $\rho_k > \eta_3$  and  $p_k = ||\Delta_k||$  (full step and model is a good approximation)

$$\Delta_{k+1} = \min(t_2 \ \Delta_k, \Delta_m)$$

else

$$\Delta_{k+1} = \Delta_k$$
  
If  $\rho_k > \eta_1$   
 $x_{k+1} = x_k + p_k$   
else

 $x_{k+1} = x_k$  (the model is not a good approximation and has to unravel another trust-region sub issue inside a littler trust-region)

end

The above algorithm uses the  $\eta_1$  corresponds to (0.65-1.09) values,  $\eta_2$  uses (0.60-1.08) values and  $\eta_3$  focus on the (0.64-1.107) range. The algorithm used to obtain the good approximation values from the training data for the condition of TSH levels, Hypothyroid, Hyperthyroid and FT4 values. Comparison of data from the individual



stages is considered to get the optimized values of TSH levels, Hyperthyroidism, Hypothyroidism values.

## 4.7 Classification Tree

A Classification Tree represents a decision tree with double elements for characterization. An item of this magnificence can foresee reactions for brand spanking new information using the anticipate approach. The object carries the records applied for getting ready, so it can register substitution expectations.

```
Pseudo-code
Step1: Predictor data
           X = mdl.X; % Predictor data
           Xbinned = zeros(size(X));
Step 2: Choose binary edges
           edges = mdl.BinEdges;
Step 3: Select indices of binned predictors.
           idxNumeric =
find(~cellfun(@isempty,edges));
           if iscolumn(idxNumeric)
             idxNumeric = idxNumeric';
           end
           for j = idxNumeric
             x = X(:,j);
Step 4: Conversion from x to array if x is a table.
             if istable(x)
                x = table2array(x);
             end
Step 5: Using discretize function grouped into
           xbinned = discretize(x,[-inf; edges{j}; inf]);
            Xbinned(:,j) = xbinned;
Step 6: End
```

The decision trees used in data mining depend on two sorts. The first is order tree investigation which is the point at where the anticipated result is the class (discrete) to which the data has a place with the thyroid of TSH levels, Hypothyroidism, Hyperthyroidism and FT4 levels. The second one being the relapse tree examination is the point at which the anticipated result can be viewed as a real number [17].

## 5. Result Analysis

The PhysicoNet-> Bank ATM-> MIT-BIHArrhythmia Dataset(mitdb), the ECG signal take as the input signal. The input signal is loaded in the MATLAB format. The proposed work implemented using the MARLABR2014b software. The simulation results were given figure 5.



## Figure 5. Input signal with the Amplitude Spectrum of the ECG signal.

The input signals are pre-processed. The use of the knowledge pre-processing is crucial step in the data mining process. The pre-processing is carried on by filtering of the ECG signal using n signal processing, the Wiener filter out is used to as objective or subjective process through direct time-invariant (LTI) sifting of an observed loud process, tolerating the known stationary sign and noise spectra, and additive noise. A standard ECG flag comprises of the P-wave, QRS complexes and T-wave. Here each QRS complexes to a Fourier spectrum from ECG indicators, the range shifts with the establishment and conduction path.

The diversifications of quality range are found within the range of 0–20 Hz within the frequency location. In frequency area, seismic trace is described with regards to amplitudes and phases at certain frequencies. To describe seismic trace within the frequency area, it requires both an amplitude spectrum and a section spectrum. The amplitude spectrum merely gives amplitude at every frequency.

In figure 6 the smoothing algorithms are based on the "move and increase" method, in which unique information are duplicated point-by-point by a set of numbers (coefficients) that characterizes the smooth shape, the items are included and isolated by the coefficients, which gathers to one point of smoothed information, at that point the set of coefficients is shifted one point down to the first information and the method is repeated [17].







| 👍 MATLAB R2014a  |  |                                      |
|--|--|--------------------------------------|
| HOME PLOTS APPS  | EDITOR RVELISH VEW   | 👪 🖥 🔏 🕼 😭 😥 🖉 🦉 Search Documentation |
| Image: Several severa                      | k ( ↓ ↔ ↔  |                                      |
| FLE ECT  | NAVIGATE BREAVPOINTS BUN   |                                      |
| <  | Deepka_nen 🕨   |                                      |
| Cerrent Filder ()<br>Varie L<br>2 data<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>400000<br>4000000<br>4000000<br>400000000  | 2 fiber-bijmal odesjdeepkal Deepka, wewinnin           minn K +           ① The R carle publied to structured document. For more informatory, see the publicity within a Mag.           268 - serve (sprinted document. For more informatory, see the publicity within a Mag.           269 - document (see all optimal (see all |                                      |
| Taining n  | 276 - ISBErel4;<br>279 - Spertyroidim#1;<br>Connad Widow<br>Beat: Beat- normal range(66-69)  |                                      |
| Name         Value         Vin           43         L0500 double         55.4           asa         304 col         304           asa         304 col         0.0           asa         706 double         0.0           asa         706 double         0.0           asa         708 double         0.0           asa         78         78           asa         Yanaf         194 maf   | 75<br>Q35 duration=127.7778<br>127.7778<br>The given data is HyperThyroid  |                                      |
| 100m /8 78<br>100m /8 78<br>100m / 100m / | konnesy 5.00000e-01<br>Sensitivity 5.285014e-01<br>Á Specificity 6.4571459-01>>  |                                      |
| NStart 🚊 🕖 🏮 🧕 🧕   | 00   | 2 🔡 🗑 🔂 🗄 🕹                          |

Figure 7. The Output Parameters resulted in Matlab.

The figure 7 shows that these settings can constrain the solver, to ascertain the yield esteems, that it would some way or the other have discarded on the grounds, that the figures were not expected to accomplish precise reproduction results. These additional estimations can make the solver to find zero intersections that have been missed.

| Table 4. Heart Beat and QRS duratio |
|-------------------------------------|
|-------------------------------------|

| SI.No. | Heart Beat<br>(BPM) | QRS Duration in (ms) | Thyroid level |
|--------|---------------------|----------------------|---------------|
| 1.     | 78                  | 127.77               | Hypothyroid   |
| 2.     | 66                  | 127.75               | Hyperthyroid  |
| 3.     | 54                  | 55.5                 | Hyperthyroid  |
| 4.     | 180                 | 186.11               | Hypothyroid   |
| 5.     | 84                  | 105.55               | Hyperthyroid  |

The table 4 shows the heartbeat, QRS duration and the thyroid level for the different ECG data set. The above table shows the dataset of the hyperthyroid, hypothyroid used as the input data.

| Table 5. Accuracy. | Sensitivity | / and | Specificit | v |
|--------------------|-------------|-------|------------|---|
| Table 0. Accuracy, | Constant    | y ana | opeomen    | y |

| SI N  | Acourcov | Sonoitivity | Specificity (%) |
|-------|----------|-------------|-----------------|
| 51.IN | (%)      | (%)         | Specificity (%) |
| 1.    | 85       | 1           | 1               |
| 2.    | 87       | 0           | 1               |
| 3.    | 87       | 1           | 1               |
| 4.    | 92       | 1           | 1               |
| 5.    | 93       | 1           | 9.8             |
| REF   |          |             |                 |
| [18]  | 71       | 47          | 86              |
| (%)   |          |             |                 |

The Accuracy, Sensitivity and Specificity for the give database as shown in the table 5. The accuracy is more in the given input data set. The sensitivity and specificity are normal for the proposed system.



Figure 8. Thyroid data classification using the thyroid data set.

The figure 8 shows the data classification using the thyroid dataset, the figure in which the data classification



are into two groups for the dataset namely hypothyroid and hyperthyroid. The blue colour star shows the hyperthyroid and black colour circle shows the hypothyroid data set using trust region method for the optimization algorithm.

## 6. Conclusion

Diagnosis of disease is a very challenging task of health care sector. Many knowledge mining tactics are used in decision making process. The proposed work uses the dimensionality technique to choose the subset of attributes from unique knowledge. Later we have applied C4.5 as a set of rules used to generate a tree, through Ross Quinlan and Classification Tree in data mining process. The classification technique is used to categorize the hypothyroid and hyperthyroid illnesses. The efficiency of classifiers is evaluated using the confusion matrix on the subjects of accuracy, sensitivity, and specificity. The C4.5 algorithm gives 93.58% which is providing better accuracy than determination stump tree accuracy and also C4.5 Algorithm provides very minimum error rate than decision stump. As a future work the same technique is used to apply for other illness datasets corresponding to heart illness and so on, especially in cases which requires both measures like ECG signal as well as haematology (blood test) results.

## References

- Frates MC, Benson CB, Charboneau JW, Cibas ES, Clark OH, et al. Management of thyroid nodules detected at US: Society of Radiologists in ultrasound consensus conference statement. Radiology. 2005; 237:794–800.
- [2] Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. J Clin Endocrinol Metab. 2006; 91:3411–7.
- [3] Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. Thyroid. 2016; 26:1–133.
- [4] K. Vijayalakshmi1, S. Dheeraj2, B.S.S. Deepthi3," Intelligent Thyroid prediction system using Big data", Vol.6(1), Jan 2018, E-ISSN: 2347-2693.
- [5] Yan Wang Ji-Guang Wang," Genome-Wide Association Studies of Hypertension and Several Other Cardiovascular Diseases", February 5, 2019, DOI: 10.1159/000496150.
- [6] Rafn Benediktsson, Anthony D Toft Endocrine Unit," Management of the unexpected result: compensated hypothyroidism", Downloaded from http://pmj.bmj.com/ on March 11, 2015 - Published by group.bmj.com.
- [7] Bernadette biondi, emiliano a. Palmieri, gaetano lombardi, and serafino fazio," Effects of Thyroid Hormone on Cardiac Function: The Relative Importance of Heart Rate, Loading Conditions, and Myocardial Contractility in the

Regulation of Cardiac Performance in Human Hyperthyroidism", The Journal of Clinical Endocrinology & Metabolism 87(3):968–974.

- [8] Thorleifsson, et. al., "Design of the Coronary ARtery DIsease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study", Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org DOI: 10.1161/CIRCGENETICS.109.899443.
- [9] Susanna C. Larsson, PhD, Elias Allara, MD Amy M. Mason, PhD Karl Michaëlsson, MD, PhD Stephen Burgess, PhD," Thyroid Function and Dysfunction in Relation to 16 Cardiovascular Diseases", *Circ Genom Precis Med.* 2019;12:e002468. DOI: 10.1161/CIRCGEN.118.002468 March 2019.
- [10] Yiyi Zhang1,2, Wendy S. Post1,2,3, Alan Cheng3, Elena Blasco-Colmenares1,2, Gordon F. Tomaselli3, Eliseo Guallar1,2,4\*," Thyroid Hormones and lectrocardiographic Parameters: Findings from the Third National Health and Nutrition Examination Survey", April 2013 | Volume 8 | Issue 4 | e59489.
- [11] Jia Liu1,2,3, Dongmei Zheng1,2,3,8\*, Qiang Li4, Xulei Tang5, Zuojie Luo6, Zhongshang Yuan7, Ling Gao2,3 and Jiajun Zhao1,2,3," A predictive model of thyroid malignancy using clinical, biochemical and sonographic parameters for patients in a multi-center setting", Liu et al. BMC Endocrine Disorders (2018) 18:17 https://doi.org/10.1186/s12902-018-0241-7.
- [12] Park M, Shin JH, Han BK, Ko EY, Hwang HS, et al. Sonography of thyroid nodules with peripheral calcifications. J Clin Ultrasound. 2009; 37:324–8.
- [13] Stojadinovic A, Peoples GE, Libutti SK, Henry LR, Eberhardt J, et al. Development of a clinical decision model for thyroid nodules. BMC Surg. 2009; 9:12.
- [14] Taghipour Zahir S, Binesh F, Mirouliaei M, Khajeh E, Noshad S. Malignancy risk assessment in patients with thyroid nodules using classification and regression trees. J Thyroid Res. 2013; 2013:983953.
- [15] Koike E, Noguchi S, Yamashita H, Murakami T, Ohshima A, et al. Ultrasonographic characteristics of thyroid nodules: prediction of malignancy. Arch Surg. 2001; 136:334–7.
- [16] J. Ye, J.-H. Chow, J. Chen, and Z. Zheng, "Stochastic gradient boosted distributed decision trees," in Proceedings of the 18<sup>th</sup> ACM conference on Information and knowledge management. ACM, 2009, pp. 2061–2064.

