PROJECT REPORT
ON
“THYROID DETECTION”

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF COMPUTER APPLICATION

Under Supervision of :	**Submitted by :**
Ms. Ashmeet Kaur	Swati Panchal
(Assistant Professor)	(09290202019)

SRI GURU TEGH BAHADUR INSTITUTE OF MANAGEMENT AND INFORMATION TECHNOLOGY

Affiliated To

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI

2019-2022

DECLARATION

This web development project titled Thyroid Detection System towards the completion of my course requirements for Semester VI is my original work and has been carried out under the guidance of Ms. Ashmeet Kaur.

The material borrowed from other sources and incorporated in the report has been duly acknowledged and referenced.

I understand that I will be held liable and accountable for my project, it is all done by me without having any copyright(s) of any of the Organization. I further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other Institute or University.

DATE:

SWATI PANCHAL

(09290202019)

CERTIFICATE

This is to certify that the project entitled : **Thyroid Detection System** done by **Swati Panchal** having university enrolment number **09290202019** is an authentic work carried out by them at **Sri Guru Tegh Bahadur Institute Of Management & Information Technology**.

The matter embodied in this project work has not been submitted earlier for the award of any degree or diplomas to the best of my knowledge and belief. The suggestion as approved by the faculty was duly incorporated.

DATE:

HOD: Mr. AMANDEEP SINGH ARORA

GUIDE: Ms. Ashmeet Kaur

ACKNOWLEDGEMENT

I would like to express our heartfelt gratitude to **Sri Guru Tegh Bahadur Institute Of Management & Information Technology** and Assistant Prof. Ashmeet Kaur for giving us the opportunity to work under their guidance and help us gain immense professional experience.

A sincere thanks to the mentor for giving us valuable inputs and ideas right from the selection of topic for project till its successful completion.

The successful completion of the project would not have been possible without the dedicated support from all our mentors, family, and friends.

DATE:

SWATI PANCHAL

09290202019

BCA-6B

ABSTRACT

The thyroid gland is a vascular gland and one of the most important organs of a human body. This gland secretes two hormones which help in controlling the metabolism of the body. The two types of thyroid disorders are Hyperthyroidism and Hypothyroidism. When the disorder occurs in the body, they release certain type of hormones into the body which imbalances the body’s metabolism.

The main objective of the Thyroid Detection System is to develop a system which provide us the clarity whether a person is suffering from thyroid disease or not. And if the result is positive then it can predict the type of thyroid disease that patient is affected from. In this project the user will register first. After setting up the username and password in registration part, with the help of this username and password the user will login. After that the user will be able to access the prediction page and will be able to predict and see the results.

Here we will be using machine learning algorithm -KNN (K-nearest neighbors) algorithm and Random Forest algorithm to predict the patient’s risk of getting thyroid disease. The project is made using PyCharm editor, jupyter notebook and SQL Lite. The front end is designed using HTML, CSS and the backend is designed using Python and database used is SQLite database.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1

INTRODUCTION

INTRODUCTION

The thyroid gland is a vascular gland and one of the most important organs of a human body. This gland secretes two hormones which help in controlling the metabolism of the body. The two types of thyroid disorders are Hyperthyroidism and Hypothyroidism. When the disorder occurs in the body, they release certain type of hormones into the body which imbalances the body’s metabolism.

According to statistics, thyroid disorders are on the rise in India. Approximately 1 in 10 Indian adults suffer from thyroid problem. It has been estimated that around 42 million people suffer from thyroid disease. Predicting thyroid disorder by doctor is a tedious process which might lead to negative prediction, only experienced doctor can examine the case properly. To assist doctors machine learning can help them in diagnosis and reduces this burden.

The main objective of the Thyroid Detection System is to develop a system which provide us the clarity whether a person is suffering from thyroid disease or not. And if the result is positive then it can predict the type of thyroid disease that patient is affected from. In this project the user will register first. After setting up the username and password in registration part, with the help of this username and password the user will login. After that the user will be able to access the prediction page and will be able to predict and see the results.

Here we will be using machine learning algorithm -KNN (K-nearest neighbors) algorithm and Random Forest algorithm to predict the patient’s risk of getting thyroid disease. The project is made using PyCharm editor, jupyter notebook and SQL Lite. The front end is designed using HTML, CSS and the backend is designed using Python and database used is SQLite database.

Data Description:

The dataset that we are using is “hypothyroid.csv”. Data will contain different classes of thyroid and 30 columns of different values.

Column names: [age, sex, on_thyroxine, query_on_thyroxine, on_antithyroid_medication, sick, pregnant, thyroid_surgery, I131_treatment, query_hypothyroid, query_hyperthyroid, lithium, goitre, tumor, hypopituitary, psych, TSH_measured, TSH, T3_measured, T3, TT4_measured, TT4, T4U_measured, T4U, FTI_measured, FTI, TBG_measured, TBG, referral_source,class]

"Class" column will have four unique values “negative, compensated_hypothyroid, primary_hypothyroid, secondary_hypothyroid”.

CHAPTER-2

REQUIREMENTS AND ANALYSIS

Software requirements specification

2.1 Category

Web Application

2.2 Purpose

The purpose of developing this web application titled 'Income Prediction' is an effort made by the developer for different users to come on this platform and predict their income.

It is having Signup and Login, profile page, prediction page where user can enter their details and predict whether they will able to earn higher than 50k dollars or lower than equal to 50k dollars.

2.3 System Analysis

Analysis of the current and future roles of proposed computer system in an organization, the system analyst (usually a software engineer or programmer) examines the flow of documents, information, and material to design a system that best meets the cost, performance, and scheduling objectives Systems analysis the process of observing systems for troubleshooting or development purposes. It is applied to information technology, where computer-based systems require defined analysis according to their makeup and design.

In IT, systems analysis can include looking at end-user implementation of a software package or product; looking in-depth at source code to define the methodologies used in building software; or taking feasibility studies and other types of research to support the use and production of a software product, among other things.

Systems analysis professionals are often called upon to look critically at systems, and redesign or recommend changes as necessary. Inside and outside of the business world, systems analysts help to evaluate whether a system is viable or efficient within the context of its overall architecture, and help to uncover the options available to the employing business or other party. Systems analysts are different than systems administrators, who maintain systems day to day, and their roles generally involve a top-level view of a system to determine its overall effectiveness according to its design.

2.4 HARDWARE AND SOFTWARE REQUIREMENTS

Software Requirements:

- Development tools: PyCharm, Jupyter Notebook
- Localhost: Command Prompt or Git Bash.
- Client: Google Chrome or any other web browser.
- Client side language: HTML, CSS
- Server side language: Python using Flask, Machine Learning.
- Database: SQLite

Hardware Requirements:

- Operating System Server: Windows 10
- Intel core i3
- RAM: 4.00GB RAM

2.5 USE CASE DIAGRAM

2.5.1 Introduction:

What is a use case diagram?

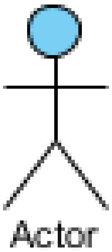
In the Unified Modeling Language (UML), a use case diagram can summarize the details of your system's users (also known as actors) and their interactions with the system. To build one, you'll use a set of specialized symbols and connectors. An effective use case diagram can help your team discuss and represent:

- Scenarios in which your system or application interacts with people, organizations, or external systems
- Goals that your system or application helps those entities (known as actors) achieve
- The scope of your system

Use case diagram components
Common components include:

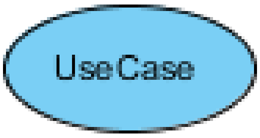
1. **Actor**

- Someone interacts with use case (system function).
- Named by noun.
- Actor plays a role in the business
- Similar to the concept of user, but a user can play different roles
- Actor triggers use case(s).
- Actor has a responsibility toward the system (inputs), and Actor has expectations from the system (outputs).



2. **Use Case**

- System function (process - automated or manual)
- Named by verb + Noun (or Noun Phrase).
- Each Actor must be linked to a use case, while some use cases may not be linked to actors.

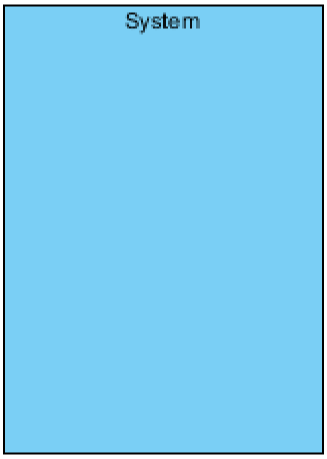


3. Communication Link

- The participation of an actor in a use case is shown by connecting an actor to a use case by a solid link.
- Actors may be connected to use cases by associations, indicating that the actor and the use case communicate with one another using messages.

4. Boundary of system

- The system boundary is potentially the entire system as defined in the requirements document.
- For large and complex systems, each module may be the system boundary.
- For example, for an ERP system for an organization, each of the modules such as personnel, payroll, accounting, etc.
- can form a system boundary for use cases specific to each of these business functions.
- The entire system can span all of these modules depicting the overall system boundary



2.5.2 USE CASE DIAGRAM

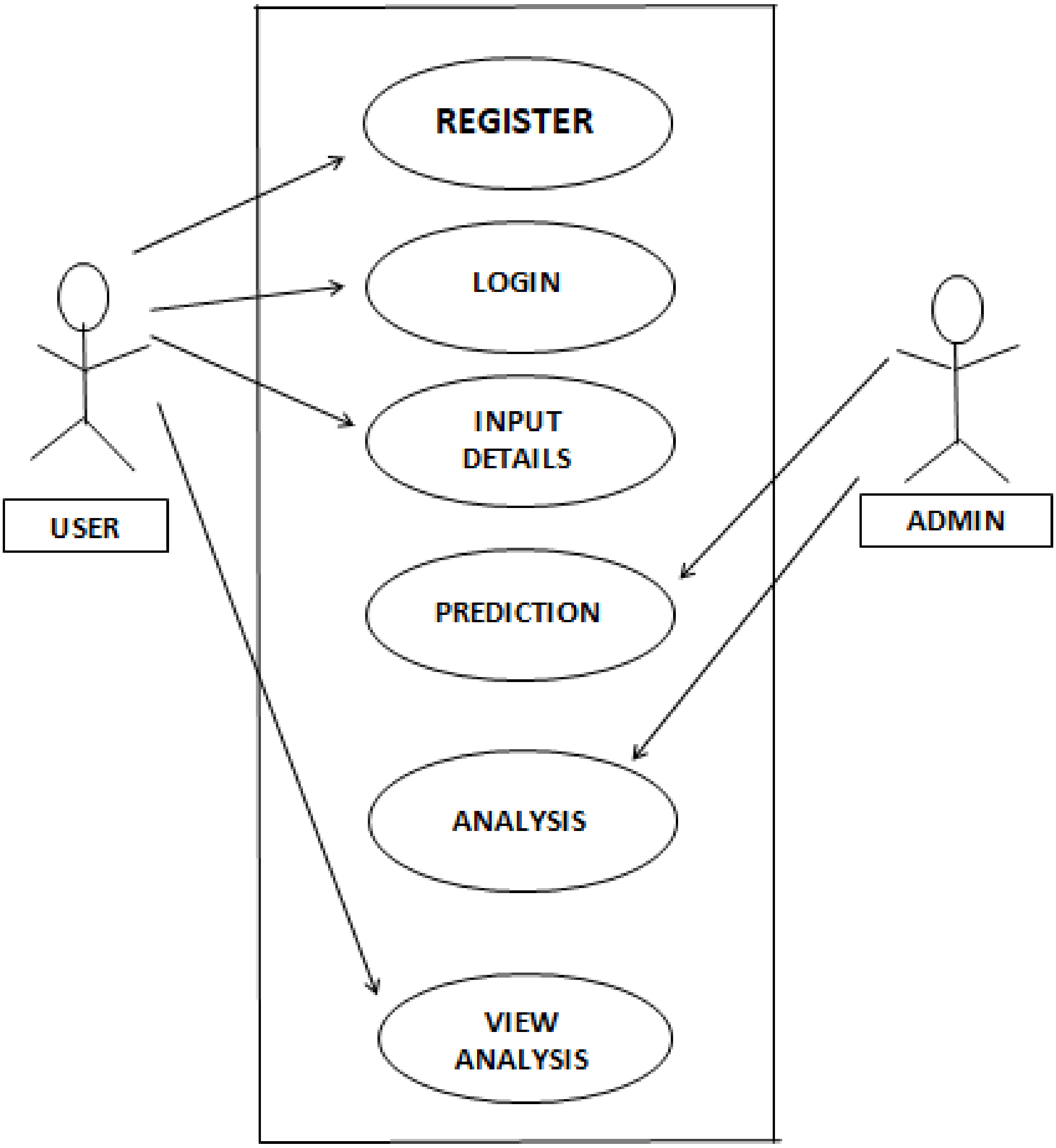


FIG.2.1- Use Case Diagram

CHAPTER-3

SOFTWARE DESIGN

3.1 Introduction

This chapter will focus on the design of the system using graphically certain diagrams to illustrate sections of the software system.

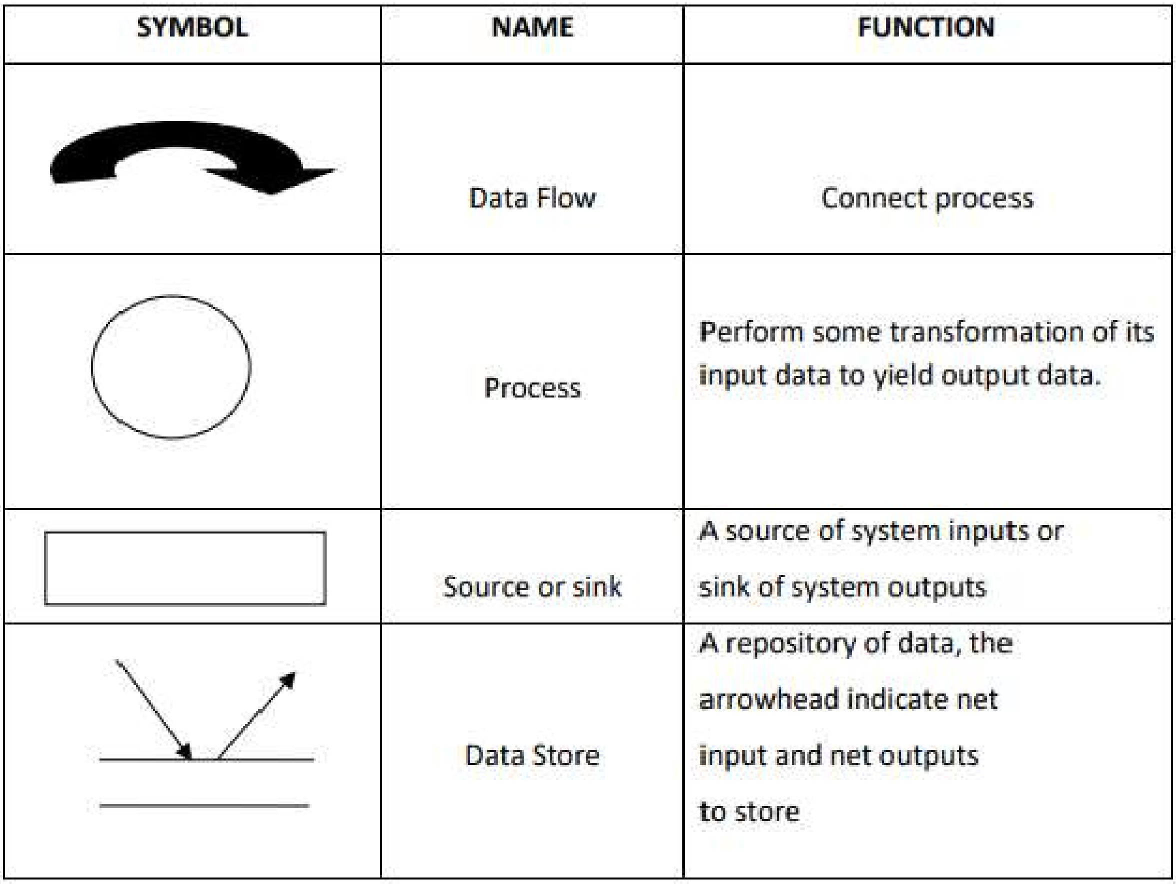
HIGH LEVEL/DETAILED DESIGN (HLD/DD)

Data flow Diagrams (DFDs)
As information moves through software, it is modified by a series of transformations.

A data flow diagram is a graphical representation that depicts information flow and the transforms that are applied as data move from input to output.

The data flow diagram may be used to represent a system or software at any level of abstraction. In fact, DFDs may be partitioned into levels that represent increasing information flow and functional detail.

Therefore, the DFD provides a mechanism for functional modeling as well as information flow modeling.



SYMBOL	NAME	FUNCTION
	Data Flow	Connect process
	Process	Perform some transformation of its input data to yield output data.
	Source or sink	A source of system inputs or sink of system outputs
	Data Store	A repository of data, the arrowhead indicate net input and net outputs to store

Table 3.1- Symbols of DFD

3.2 DFD Level-0

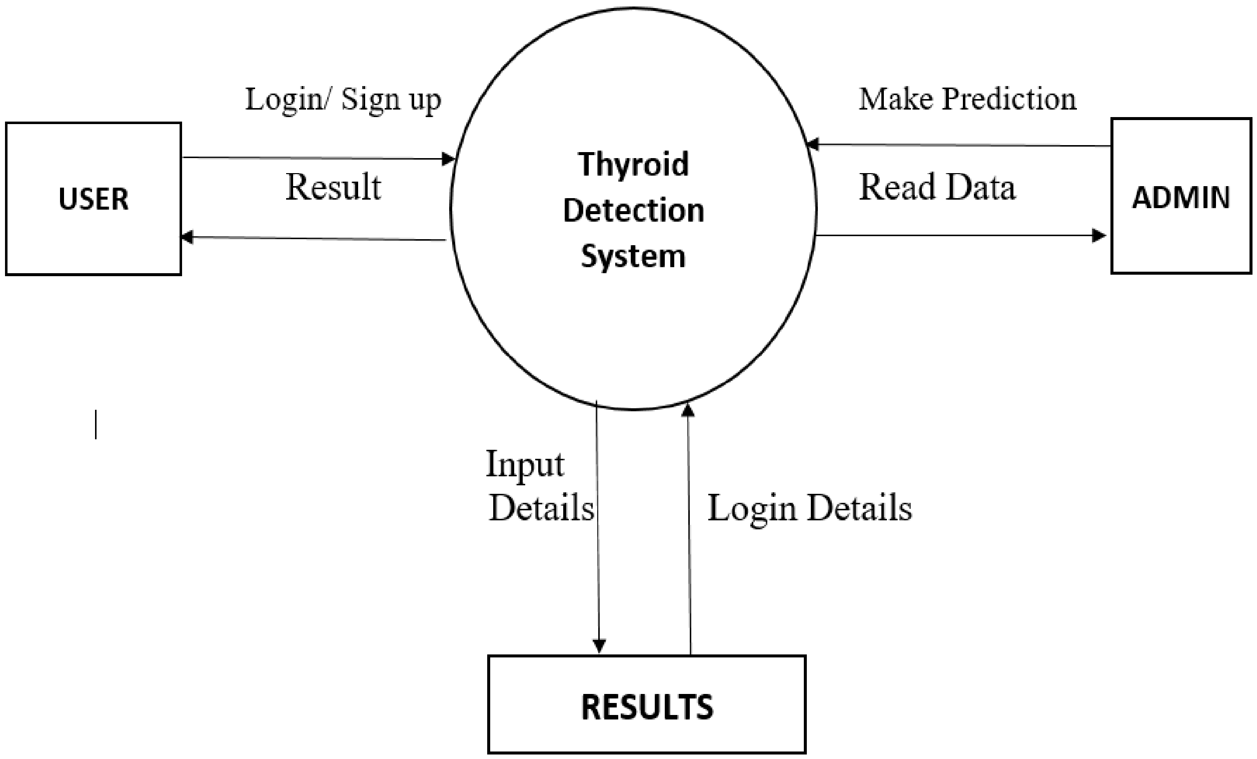


FIG.3.1 DFD Level-0

3.3 DFD LEVEL-1

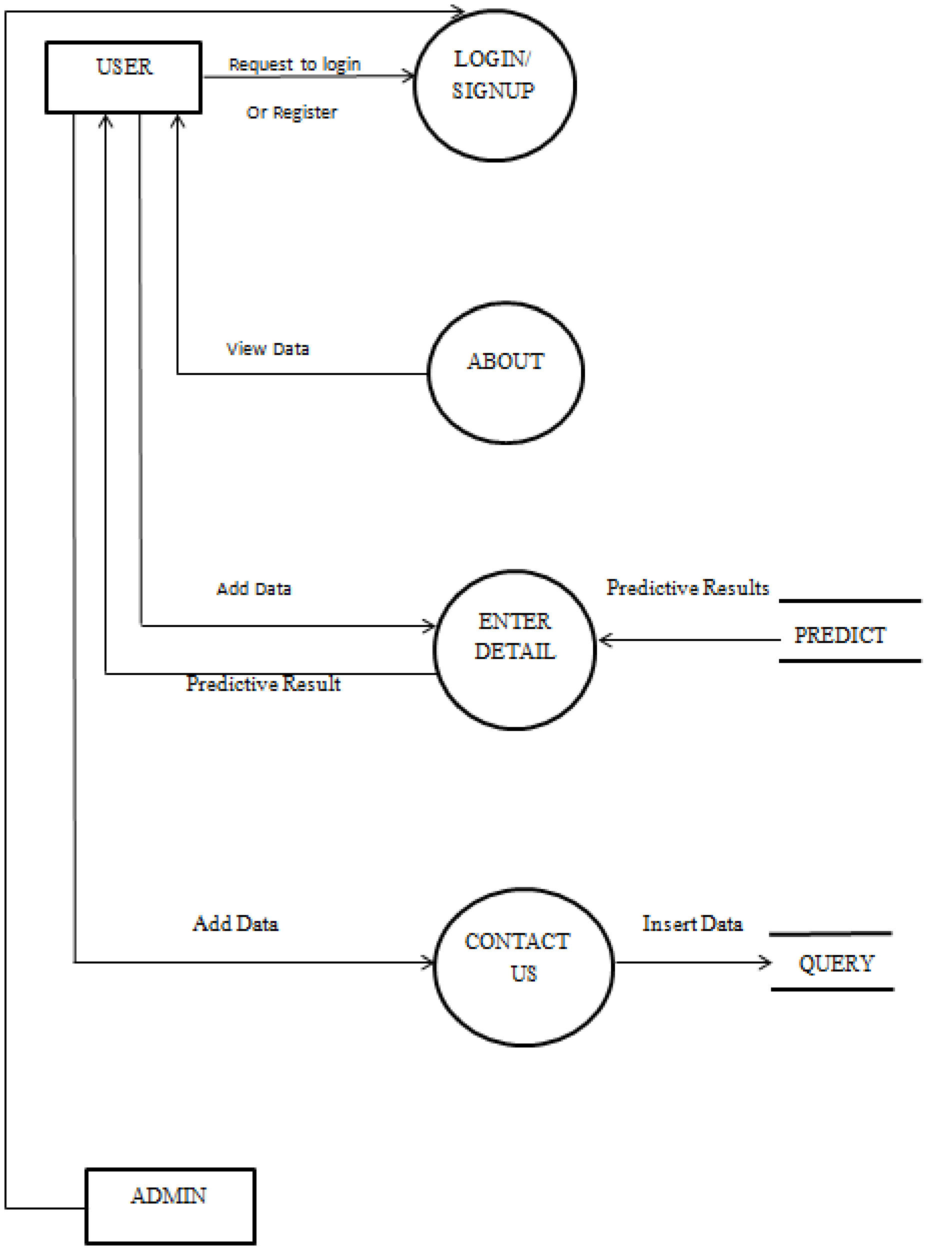


FIG.3.2 DFD Level-1

3.4 DFD LEVEL-2

1.User Sign Up:

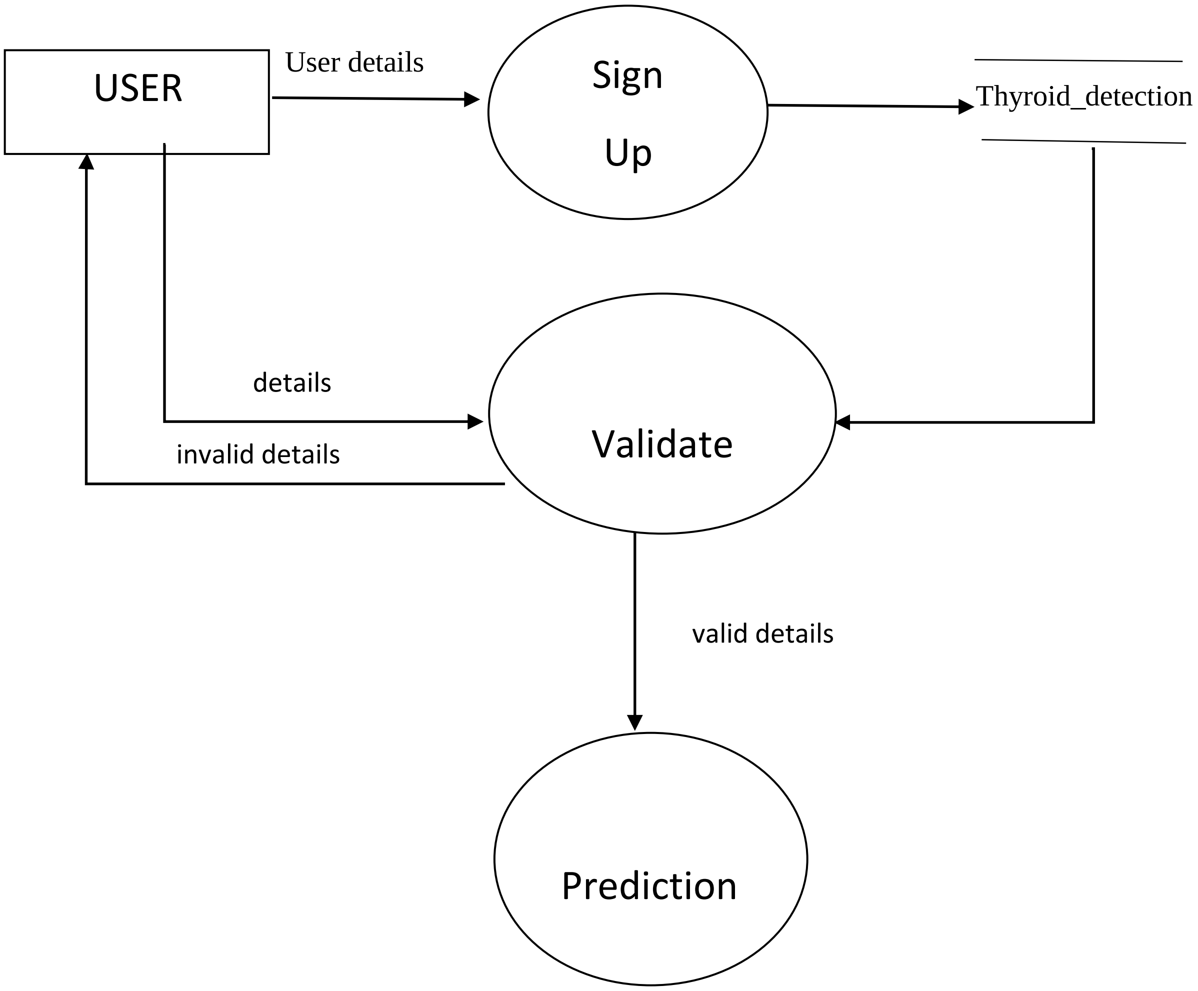


Fig.3.3.1-User Sign up (DFD Level-2)

2. Prediction:

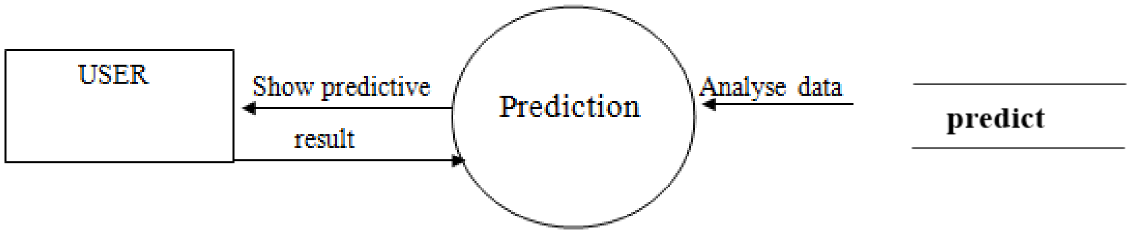


FIG 3.3.2- Prediction (DFD Level-2)

3. Contact Us:

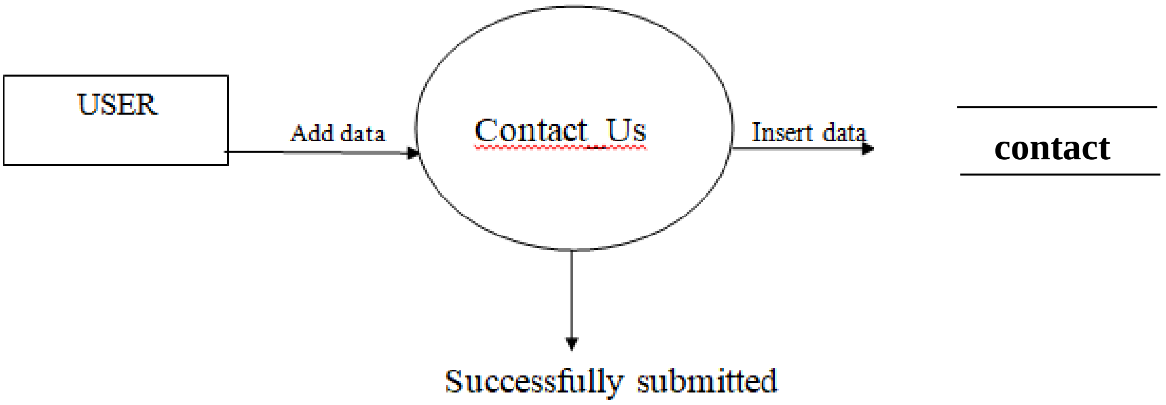


FIG.3.3.3- Contact us (DFD Level-2)

CHAPTER 4
DATABASE DESIGN

4.1 Introduction

Database design can be generally defined as a collection of tasks or processes that enhance the designing, development, implementation, and maintenance of enterprise data management system. Designing a proper database reduces the maintenance cost thereby improving data consistency and the cost-effective measures are greatly influenced in terms of disk storage space. Therefore, there has to be a brilliant concept of designing a database. The designer should follow the constraints and decide how the elements correlate and what kind of data must be stored.

The main objectives behind database designing are to produce physical and logical design models of the proposed database system. To elaborate this, the logical model is primarily concentrated on the requirements of data and the considerations must be made in terms of monolithic considerations and hence the stored physical data must be stored independent of the physical conditions. On the other hand, the physical database design model includes a translation of the logical design model of the database by keep control of physical media using hardware resources and software systems such as Database Management System.

4.2 Importance of Database Design

The important consideration that can be taken into account while emphasizing the importance of database design can be explained in terms of the following points given below.

- Database designs provide the blueprints of how the data is going to be stored in a system. A proper design of a database highly affects the overall performance of any application.
- The designing principles defined for a database give a clear idea of the behavior of any application and how the requests are processed.
- Another instance to emphasize the database design is that a proper database design meets all the requirements of users.
- Lastly, the processing time of an application is greatly reduced if the constraints of designing a highly efficient database are properly implemented.

4.3 Database Fields Specification

A small write-up on the database, giving the fields, explaining each field etc should be written. This write-up can be evaluated on the following points.

1. Clarity and conciseness of the database design. Like, whether key is defined, whether any redundant fields are there etc.

2. Whether data-storage calculations are made, and if so, whether they are done properly etc. Whether any data-backup/recovery

4.4 Database Keys

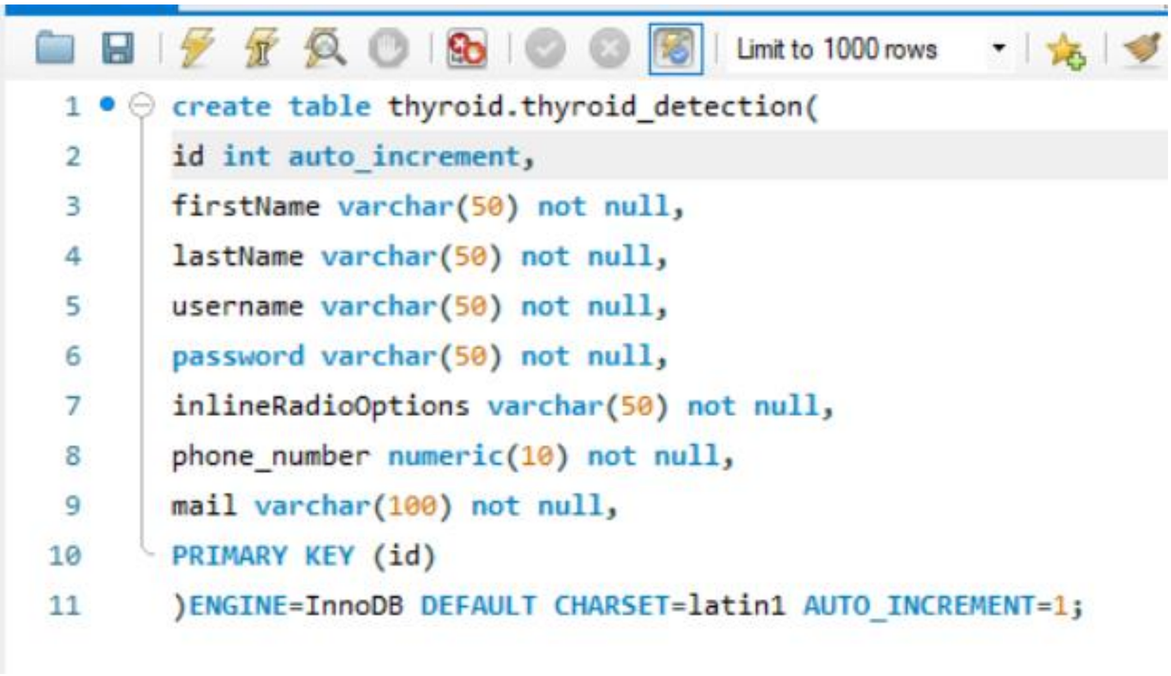
Database tables are an implementation of relations. Relations are defined to be sets of tuples and as such have all the properties of sets. Unfortunately SQL tables are not natively sets unless uniqueness constraints are placed on them. This article continues a discussion of uniqueness constraints, specifically it formalizes the notion of super key, primary key, and candidate key

Types of Keys:

- **Super Key** – A super key is a group of single or multiple keys which identifies rows in a table.
- **Primary Key** – is a column or group of columns in a table that uniquely identify every row in that table.
- **Candidate Key** – is a set of attributes that uniquely identify tuples in a table. Candidate Key is a super key with no repeated attributes.
- **Alternate Key** – is a column or group of columns in a table that uniquely identify every row in that table.
- **Foreign Key** – is a column that creates a relationship between two tables. The purpose of Foreign keys is to maintain data integrity and allow navigation between two different instances of an entity.
- **Compound Key** – has two or more attributes that allow you to uniquely recognize a specific record. It is possible that each column may not be unique by itself within the database.
- **Composite Key** – is a combination of two or more columns that uniquely identify rows in a table. The combination of columns guarantees uniqueness, though individual uniqueness is not guaranteed.
- **Surrogate Key** – An artificial key which aims to uniquely identify each record is called a surrogate key. These kind of key are unique because they are created when you don’t have any natural primary key.

4.5 Database Tables

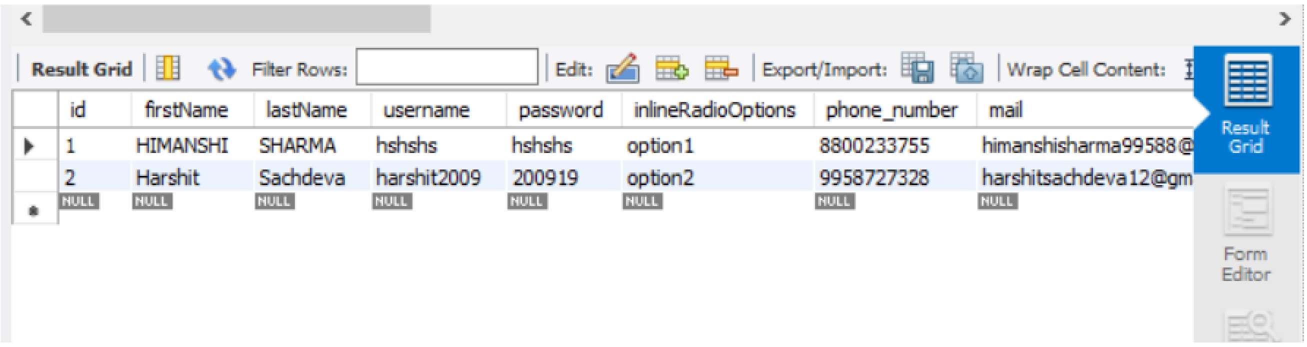
1. Thyroid_Detection:



```
create table thyroid.thyroid_detection(
id int auto_increment,
firstName varchar(50) not null,
lastName varchar(50) not null,
username varchar(50) not null,
password varchar(50) not null,
inlineRadioOptions varchar(50) not null,
phone_number numeric(10) not null,
mail varchar(100) not null,
PRIMARY KEY (id)
)ENGINE=InnoDB DEFAULT CHARSET=latin1 AUTO_INCREMENT=1;
```

FIG.4.1.1- Thyroid Detection

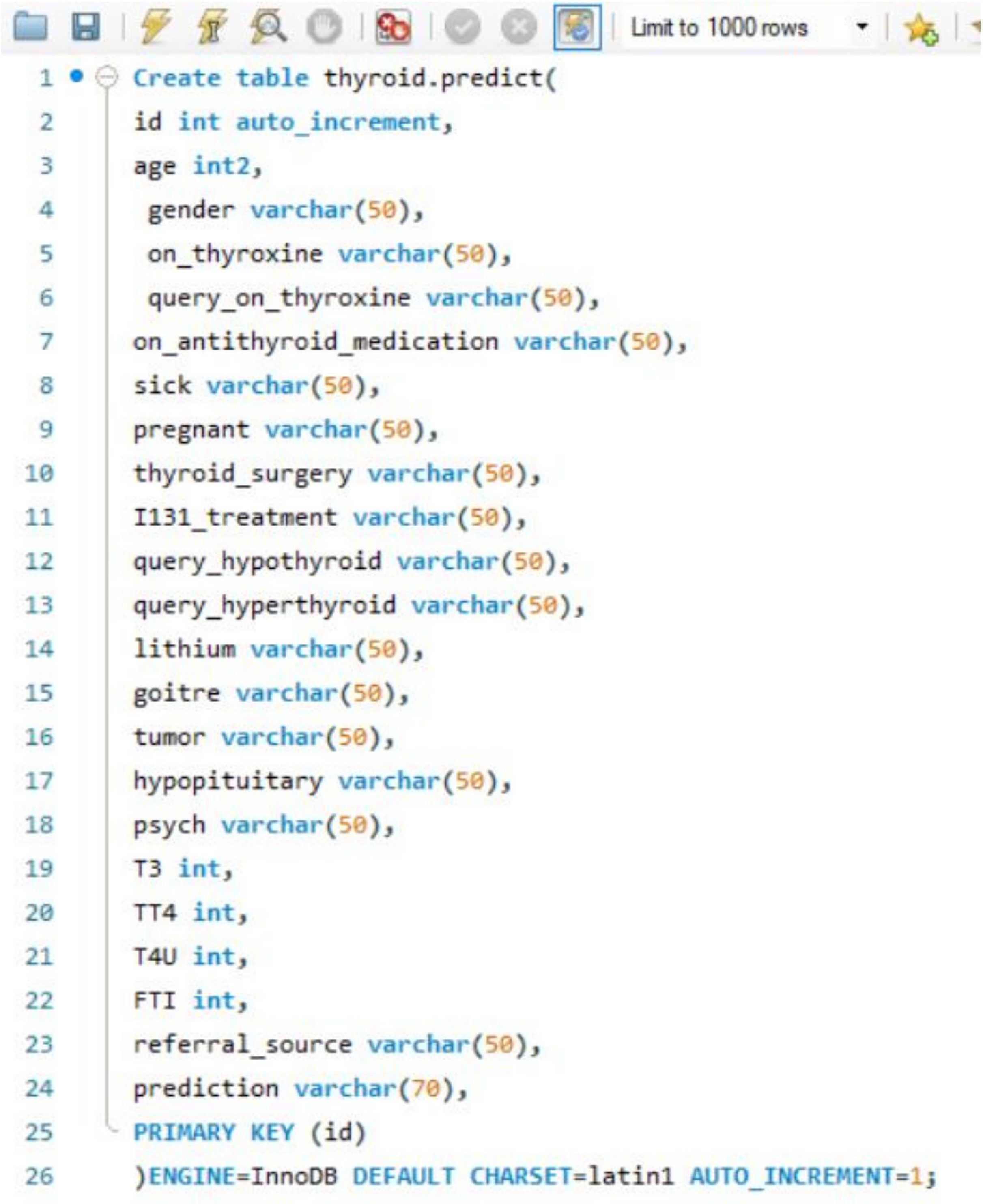
Records:



id	firstName	lastName	username	password	inlineRadioOptions	phone_number	mail
1	HIMANSHI	SHARMA	hshshs	hshshs	option1	8800233755	himanshisharma99588@
2	Harshit	Sachdeva	harshit2009	200919	option2	9958727328	harshitsachdeva12@gm
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

FIG.4.1.2 -Records

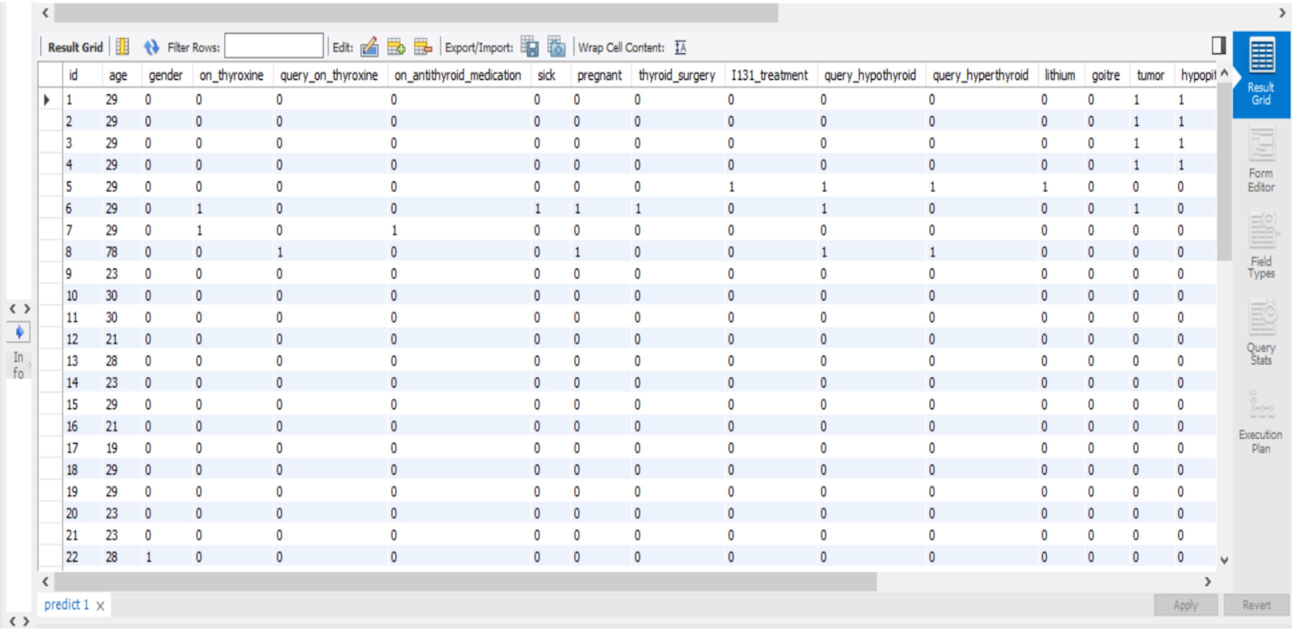
2. Predict:



```
Create table thyroid.predict(
id int auto_increment,
age int2,
 gender varchar(50),
 on_thyroxine varchar(50),
 query_on_thyroxine varchar(50),
on_antithyroid_medication varchar(50),
sick varchar(50),
pregnant varchar(50),
thyroid_surgery varchar(50),
I131_treatment varchar(50),
query_hypothyroid varchar(50),
query_hyperthyroid varchar(50),
lithium varchar(50),
goitre varchar(50),
tumor varchar(50),
hypopituitary varchar(50),
psych varchar(50),
T3 int,
TT4 int,
T4U int,
FTI int,
referral_source varchar(50),
prediction varchar(70),
PRIMARY KEY (id)
)ENGINE=InnoDB DEFAULT CHARSET=latin1 AUTO_INCREMENT=1;
```

FIG.4.2.1- Predict

Records:



id	age	gender	on_thyroxine	query_on_thyroxine	on_antithyroid_medication	sick	pregnant	thyroid_surgery	I131_treatment	query_hypothyroid	query_hyperthyroid	lithium	goitre	tumor	hypopit
1	29	0	0	0	0	0	0	0	0	0	0	0	0	1	1
2	29	0	0	0	0	0	0	0	0	0	0	0	0	1	1
3	29	0	0	0	0	0	0	0	0	0	0	0	0	1	1
4	29	0	0	0	0	0	0	0	0	0	0	0	0	1	1
5	29	0	0	0	0	0	0	0	1	1	1	1	0	0	0
6	29	0	1	0	0	1	1	1	0	1	0	0	0	1	0
7	29	0	1	0	1	0	0	0	0	0	0	0	0	0	0
8	78	0	0	1	0	0	1	0	0	1	1	0	0	0	0
9	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	28	1	0	0	0	0	0	0	0	0	0	0	0	0	0

FIG.4.2.2- Records

4.6 ER Diagram

ER Diagram stands for Entity Relationship Diagram, also known as ERD is a diagram that displays the relationship of entity sets stored in a database. In other words, ER diagrams help to explain the logical structure of databases. ER diagrams are created based on three basic concepts: entities, attributes and relationships.

ER Diagrams Symbols & Notations

Entity Relationship Diagram Symbols & Notations mainly contains three basic symbols which are rectangle, oval and diamond to represent relationships between elements, entities and attributes. There are some sub-elements which are based on main elements in ERD Diagram. ER Diagram is a visual representation of data that describes how data is related to each other using different ERD Symbols and Notations.

Following are the main components and its symbols in ER Diagrams:

- **Rectangles:** This Entity Relationship Diagram symbol represents entity types
- **Ellipses :** Symbol represent attributes
- **Diamonds:** This symbol represents relationship types
- **Lines:** It links attributes to entity types and entity types with other relationship types
- **Primary key:** attributes are underlined
- **Double Ellipses:** Represent multi-valued attributes

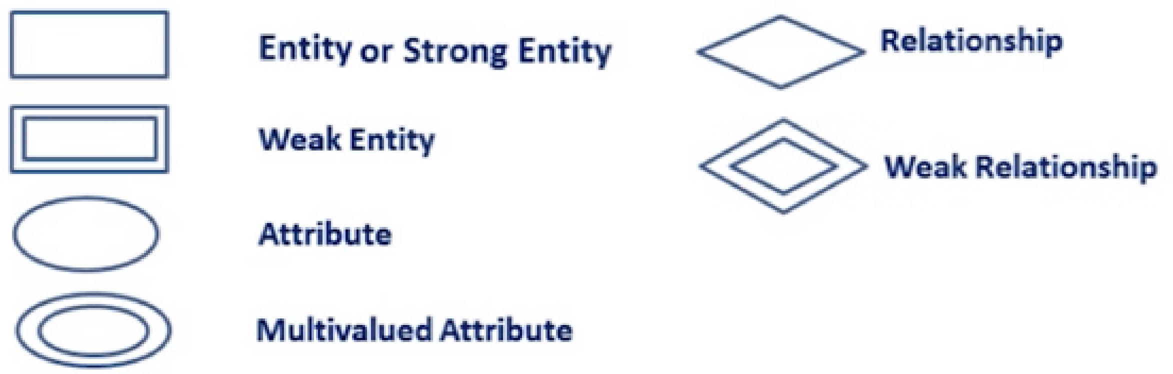


Fig.4.3-ER Diagram Symbols

ER DIAGRAM

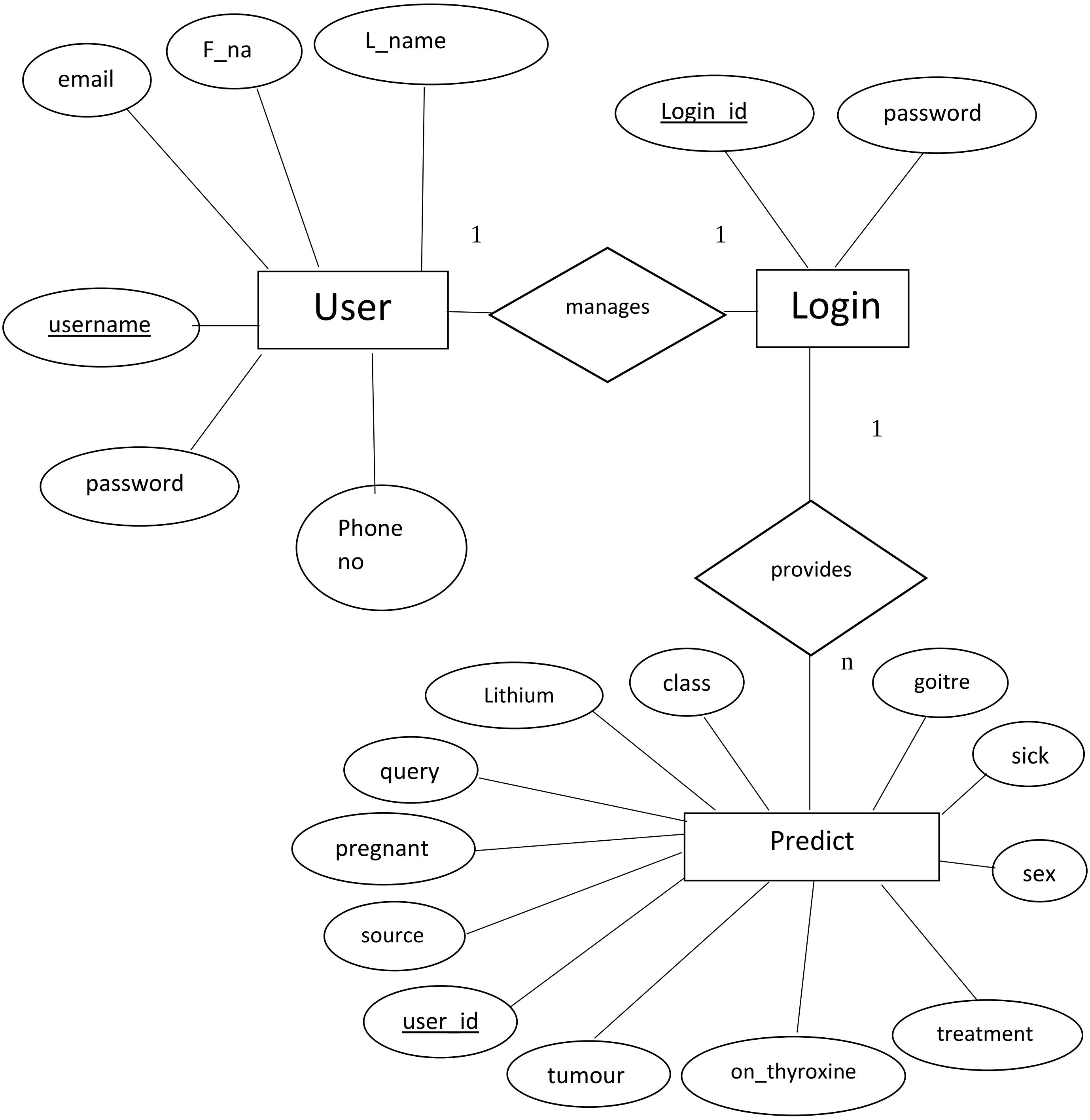


FIG.4.4- ER Diagram

CHAPTER-5

IMPLEMENTATION

5.1 Languages Used

- **Python**
  Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.
  Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.
- **Machine Learning**
  Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.
  Machine learning is important because it gives enterprises a view of trends in customer behavior and business operational patterns, as well as supports the development of new products. Many of today's leading companies, such as Facebook, Google and Uber, make machine learning a central part of their operations. Machine learning has become a significant competitive differentiator for many companies.
  Types of machine learning:
  Classical machine learning is often categorized by how an algorithm learns to become more accurate in its predictions. There are four basic approaches:supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. The type of algorithm data scientists choose to use depends on what type of data they want to predict.
  - Supervised learning: In this type of machine learning, data scientists supply algorithms with labeled training data and define the variables they want the algorithm to assess for correlations. Both the input and the output of the algorithm is specified.
  - Unsupervised learning: This type of machine learning involves algorithms that train on unlabeled data. The algorithm scans through data sets looking for any meaningful connection. The data that algorithms train on as well as the predictions or recommendations they output are predetermined.

  - Semi-supervised learning: This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labeled training data, but the model is free to explore the data on its own and develop its own understanding of the data set.
  - Reinforcement learning: Data scientists typically use reinforcement learning to teach a machine to complete a multi-step process for which there are clearly defined rules. Data scientists program an algorithm to complete a task and give it positive or negative cues as it works out how to complete a task. But for the most part, the algorithm decides on its own what steps to take along the way.

- **HTML**

  HTML (HyperText Markup Language) is the most basic building block of the Web. It defines the meaning and structure of web content. Other technologies besides HTML are generally used to describe a web page's appearance/presentation (CSS) or functionality/behavior (JavaScript).

  "Hypertext" refers to links that connect web pages to one another, either within a single website or between websites. Links are a fundamental aspect of the Web. By uploading content to the Internet and linking it to pages created by other people, you become an active participant in the World Wide Web.

  HTML uses "markup" to annotate text, images, and other content for display in a Web browser. HTML markup includes special "elements" such as <head>, <title>, <body>, <header>, <footer>, <article>, <section>, <p>, <div>, <span>, <img>, <aside>, <audio>, <canvas>, <datalist>, <details>, <embed>, <nav>, <output>, <progress>, <video>, <ul>, <ol>, <li> and many others.

- **CSS**

  CSS is the language for describing the presentation of Web pages, including colors, layout, and fonts. It allows one to adapt the presentation to different types of devices, such as large screens, small screens, or printers. CSS is independent of HTML and can be used with any XML-based markup language. The separation of HTML from CSS makes it easier to maintain sites, share style sheets across pages, and tailor pages to different environments. This is referred to as the separation of structure (or: content) from presentation.

5.2 Editors

- **PyCharm**

  PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python programming language. It is developed by the Czech company JetBrains (formerly known as IntelliJ). It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as data science with Anaconda.

  PyCharm is cross-platform, with Windows, macOS and Linux versions. The Community Edition is released under the Apache License, and there is also an educational version, as well as a Professional Edition with extra features (released under a subscription-funded proprietary license).

  Features:

- Coding assistance and analysis, with code completion, syntax and error highlighting, linter integration, and quick fixes
- Project and code navigation: specialized project views, file structure views and quick jumping between files, classes, methods and usages
- Python refactoring: includes rename, extract method, introduce variable, introduce constant, pull up, push down and others
- Support for web frameworks: Django, web2py and Flask [professional edition only]
- Integrated Python debugger
- Integrated unit testing, with line-by-line code coverage
- Google App Engine Python development [professional edition only]
- Version control integration: unified user interface for Mercurial, Git, Subversion, Perforce and CVS with change lists and merge
- Support for scientific tools like Matplotlib, NumPy and SciPy [professional edition only]

- Jupyter Notebook
  The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. Its uses include data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.
  Jupyter Notebook (formerly IPython Notebooks) is a web-based interactive computational environment for creating Jupyter notebook documents. The “notebook” term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context.
  According to the official website of Jupyter, Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.

5.3 Libraries Used

- **Pandas**
  Pandas is a Python library for data analysis. Started by Wes McKinney in 2008 out of a need for a powerful and flexible quantitative analysis tool, pandas has grown into one of the most popular Python libraries.
- **Numpy**
  In Python we have lists that serve the purpose of arrays, but they are slow to process.
  NumPy aims to provide an array object that is up to 50x faster than traditional Python lists. The array object in NumPy is called ndarray, it provides a lot of supporting functions that make working with ndarray very easy. Arrays are very frequently used in data science, where speed and resources are very important.
- **Imblearn**
  Imbalanced-learn (imported as imblearn) is an open source, MIT-licensed library relying on scikit-learn (imported as sklearn) and provides tools when dealing with classification with imbalanced classes.
- **Matplotlib**
  Matplotlib is an amazing visualization library in Python for 2D plots of arrays. Matplotlib is a multi-platform data visualization library built on NumPy arrays and

designed to work with the broader SciPy stack. It was introduced by John Hunter in the year 2002.

One of the greatest benefits of visualization is that it allows us visual access to huge amounts of data in easily digestible visuals. Matplotlib consists of several plots like line, bar, scatter, histogram etc.

- **Sklearn**
  Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistence interface in Python. This library, which is largely written in Python, is built upon NumPy, SciPy and Matplotlib.
- **Pickle**
  Python pickle module is used for serializing and de-serializing python object structures. The process to converts any kind of python objects (list, dict, etc.) into byte streams (0s and 1s) is called pickling or serialization or flattening or marshalling. We can converts the byte stream (generated through pickling) back into python objects by a process called as unpickling.
- **Flask**
  Flask is a small and lightweight Python web framework that provides useful tools and features that make creating web applications in Python easier. It gives developers flexibility and is a more accessible framework for new developers since you can build a web application quickly using only a single Python file.
- **flask_mysqldb**
  Flask-MySQLdb provides MySQL connection for Flask.
- **Re**
  Regular Expression Syntax. A regular expression (or RE) specifies a set of strings that matches it; the functions in this module let you check if a particular string matches a given regular expression (or if a given regular expression matches a particular string, which comes down to the same thing).

5.5 System Analysis

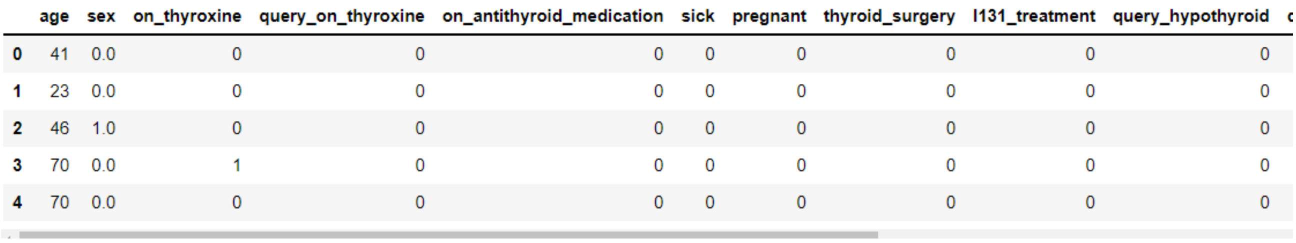
- **Dataset:**

The name of the dataset that has been taken from kaggle is "**hypothyroid.csv".**

Total rows: 3772

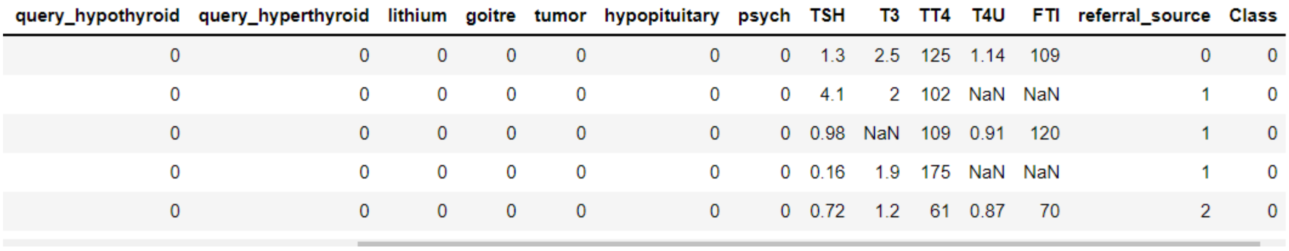
Total columns: 30

Top 5 records of dataset:



	age	sex	on_thyroxine	query_on_thyroxine	on_antithyroid_medication	sick	pregnant	thyroid_surgery	I131_treatment	query_hypothyroid
0	41	0.0	0	0	0	0	0	0	0	0
1	23	0.0	0	0	0	0	0	0	0	0
2	46	1.0	0	0	0	0	0	0	0	0
3	70	0.0	1	0	0	0	0	0	0	0
4	70	0.0	0	0	0	0	0	0	0	0

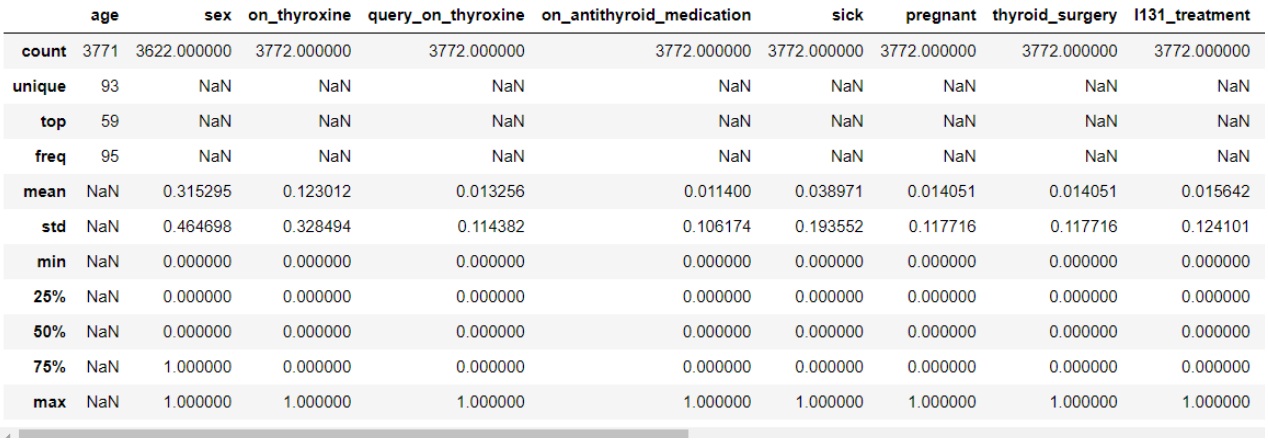
Fig.5.1.1-Top 5 data records



query_hypothyroid	query_hyperthyroid	lithium	goitre	tumor	hypopituitary	psych	TSH	T3	TT4	T4U	FTI	referral_source	Class
0	0	0	0	0	0	0	1.3	2.5	125	1.14	109	0	0
0	0	0	0	0	0	0	4.1	2	102	NaN	NaN	1	0
0	0	0	0	0	0	0	0.98	NaN	109	0.91	120	1	0
0	0	0	0	0	0	0	0.16	1.9	175	NaN	NaN	1	0
0	0	0	0	0	0	0	0.72	1.2	61	0.87	70	2	0

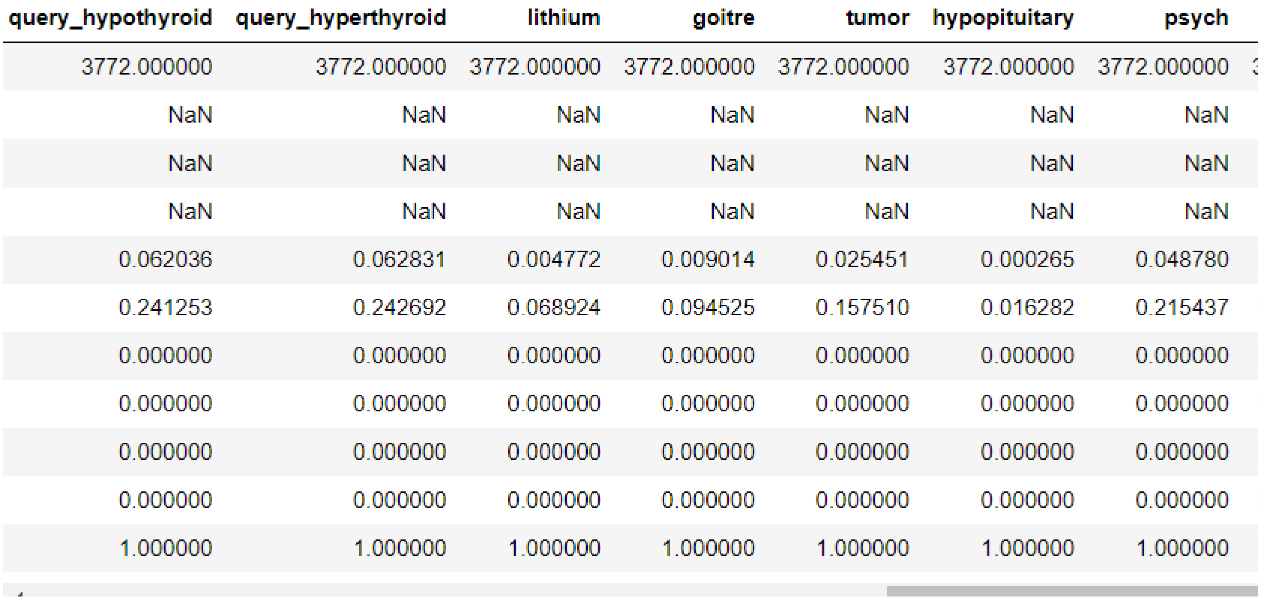
Fig.5.1.2-Top 5 data records

- **Description of data:**



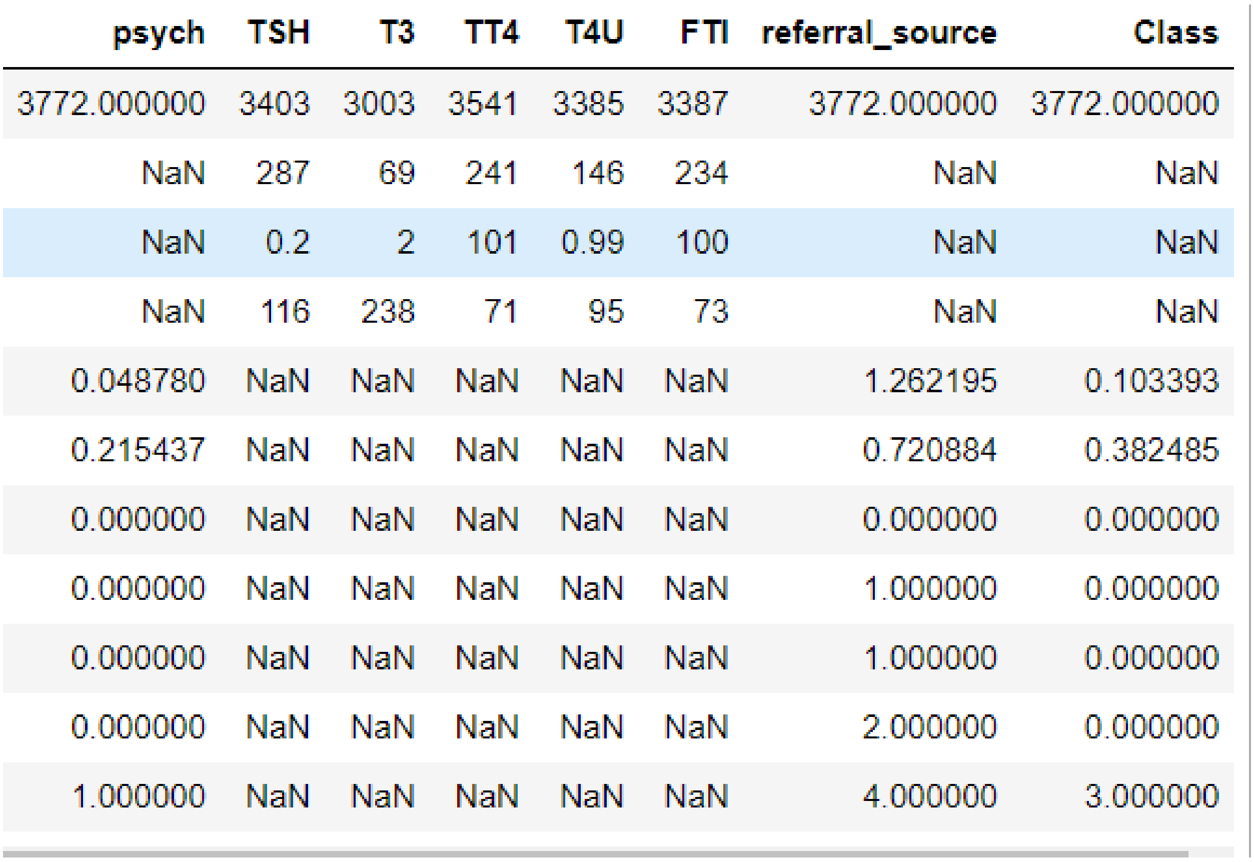
	age	sex	on_thyroxine	query_on_thyroxine	on_antithyroid_medication	sick	pregnant	thyroid_surgery	I131_treatment
count	3771	3622.000000	3772.000000	3772.000000	3772.000000	3772.000000	3772.000000	3772.000000	3772.000000
unique	93	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
top	59	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
freq	95	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
mean	NaN	0.315295	0.123012	0.013256	0.011400	0.038971	0.014051	0.014051	0.015642
std	NaN	0.464698	0.328494	0.114382	0.106174	0.193552	0.117716	0.117716	0.124101
min	NaN	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	NaN	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	NaN	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	NaN	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
max	NaN	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Fig.5.2.1-Description of data



query_hypothyroid	query_hyperthyroid	lithium	goitre	tumor	hypopituitary	psych
3772.000000	3772.000000	3772.000000	3772.000000	3772.000000	3772.000000	3772.000000
NaN	NaN	NaN	NaN	NaN	NaN	NaN
NaN	NaN	NaN	NaN	NaN	NaN	NaN
NaN	NaN	NaN	NaN	NaN	NaN	NaN
0.062036	0.062831	0.004772	0.009014	0.025451	0.000265	0.048780
0.241253	0.242692	0.068924	0.094525	0.157510	0.016282	0.215437
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

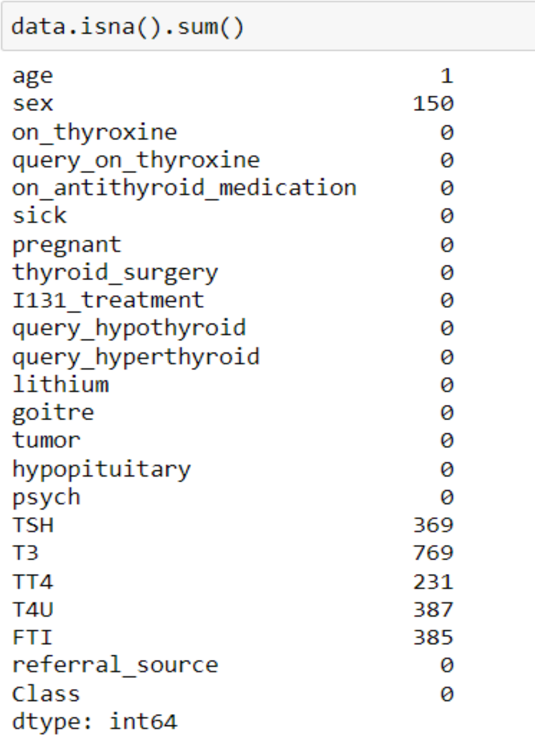
Fig.5.2.2-Description of data



psych	TSH	T3	TT4	T4U	FTI	referral_source	Class
3772.000000	3403	3003	3541	3385	3387	3772.000000	3772.000000
NaN	287	69	241	146	234	NaN	NaN
NaN	0.2	2	101	0.99	100	NaN	NaN
NaN	116	238	71	95	73	NaN	NaN
0.048780	NaN	NaN	NaN	NaN	NaN	1.262195	0.103393
0.215437	NaN	NaN	NaN	NaN	NaN	0.720884	0.382485
0.000000	NaN	NaN	NaN	NaN	NaN	0.000000	0.000000
0.000000	NaN	NaN	NaN	NaN	NaN	1.000000	0.000000
0.000000	NaN	NaN	NaN	NaN	NaN	1.000000	0.000000
0.000000	NaN	NaN	NaN	NaN	NaN	2.000000	0.000000
1.000000	NaN	NaN	NaN	NaN	NaN	4.000000	3.000000

Fig.5.2.3-Description of data

- **No. of missing values in the dataset:**



```
data.isna().sum()
```

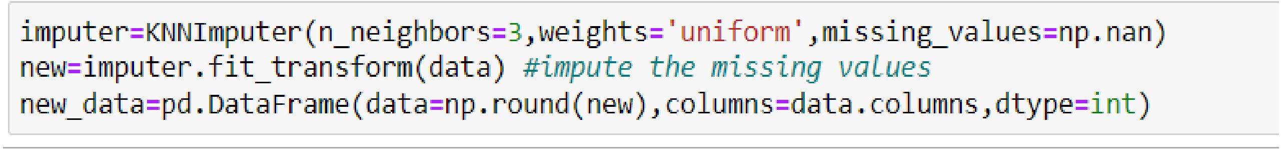
```
age                            1
sex                          150
on_thyroxine                   0
query_on_thyroxine             0
on_antithyroid_medication      0
sick                           0
pregnant                       0
thyroid_surgery                0
I131_treatment                 0
query_hypothyroid              0
query_hyperthyroid             0
lithium                        0
goitre                         0
tumor                          0
hypopituitary                  0
psych                          0
TSH                          369
T3                           769
TT4                          231
T4U                          387
FTI                          385
referral_source                0
Class                          0
dtype: int64
```

Fig.5.3 -Before removing null values

By using KNNImputer, we will be removing the nan or missing values.

KNNImputer- KNNimputer is a scikit-learn class used to fill out or predict the missing values in a dataset. It is a more useful method which works on the basic approach of the KNN algorithm rather than the naive approach of filling all the values with mean or the median.

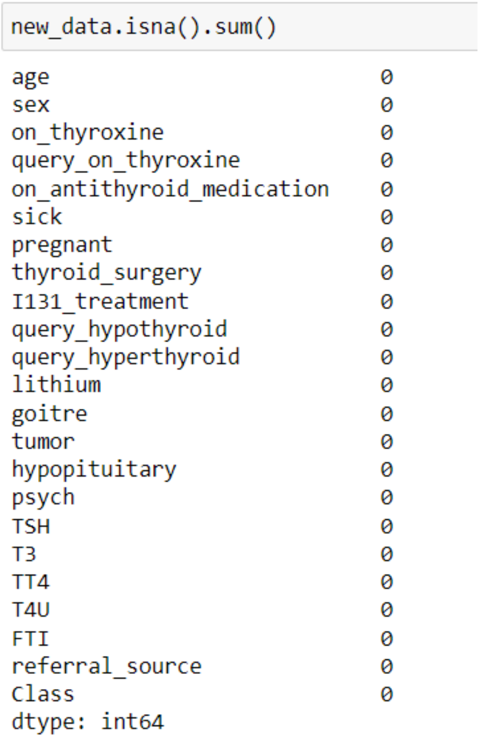
Using KNNImputer:



```
imputer=KNNImputer(n_neighbors=3,weights='uniform',missing_values=np.nan)
new=imputer.fit_transform(data) #impute the missing values
new_data=pd.DataFrame(data=np.round(new),columns=data.columns,dtype=int)
```

Fig.5.4-Using KNN Imputer

After removing the missing values, checking the no. of missing values:



```
new_data.isna().sum()
```

```
age                          0
sex                          0
on_thyroxine                 0
query_on_thyroxine           0
on_antithyroid_medication    0
sick                         0
pregnant                     0
thyroid_surgery              0
I131_treatment               0
query_hypothyroid            0
query_hyperthyroid           0
lithium                      0
goitre                       0
tumor                        0
hypopituitary                0
psych                        0
TSH                          0
T3                           0
TT4                          0
T4U                          0
FTI                          0
referral_source              0
Class                        0
dtype: int64
```

Fig.5.5- After removing null values

- **Checking the distribution of data:**

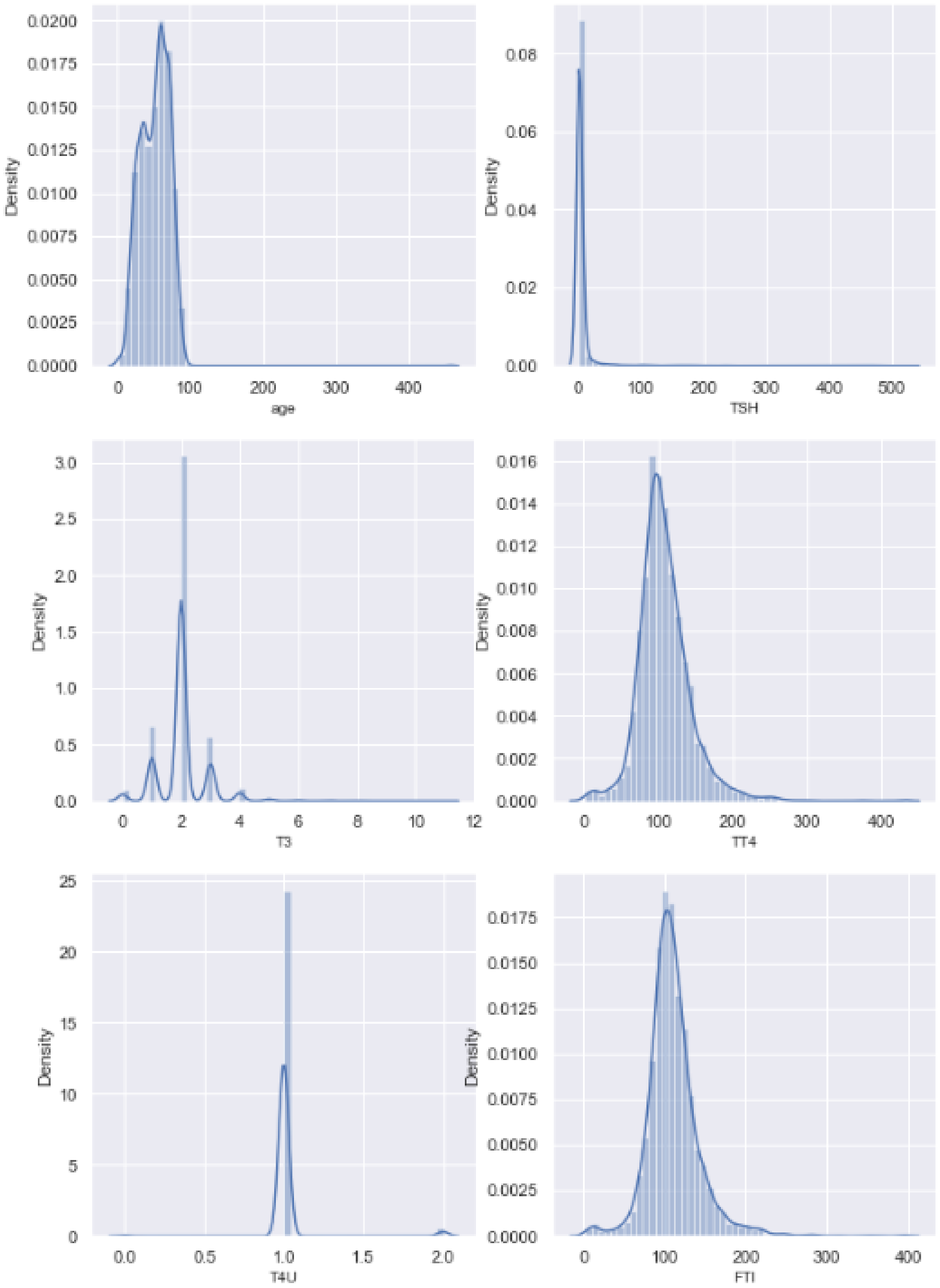


Fig.5.6 Before Transformation

The graphs for age, TSH and T3 looks heavely skewed towards left. So, some of the transformation is needed to improve the plot.

After transformation:

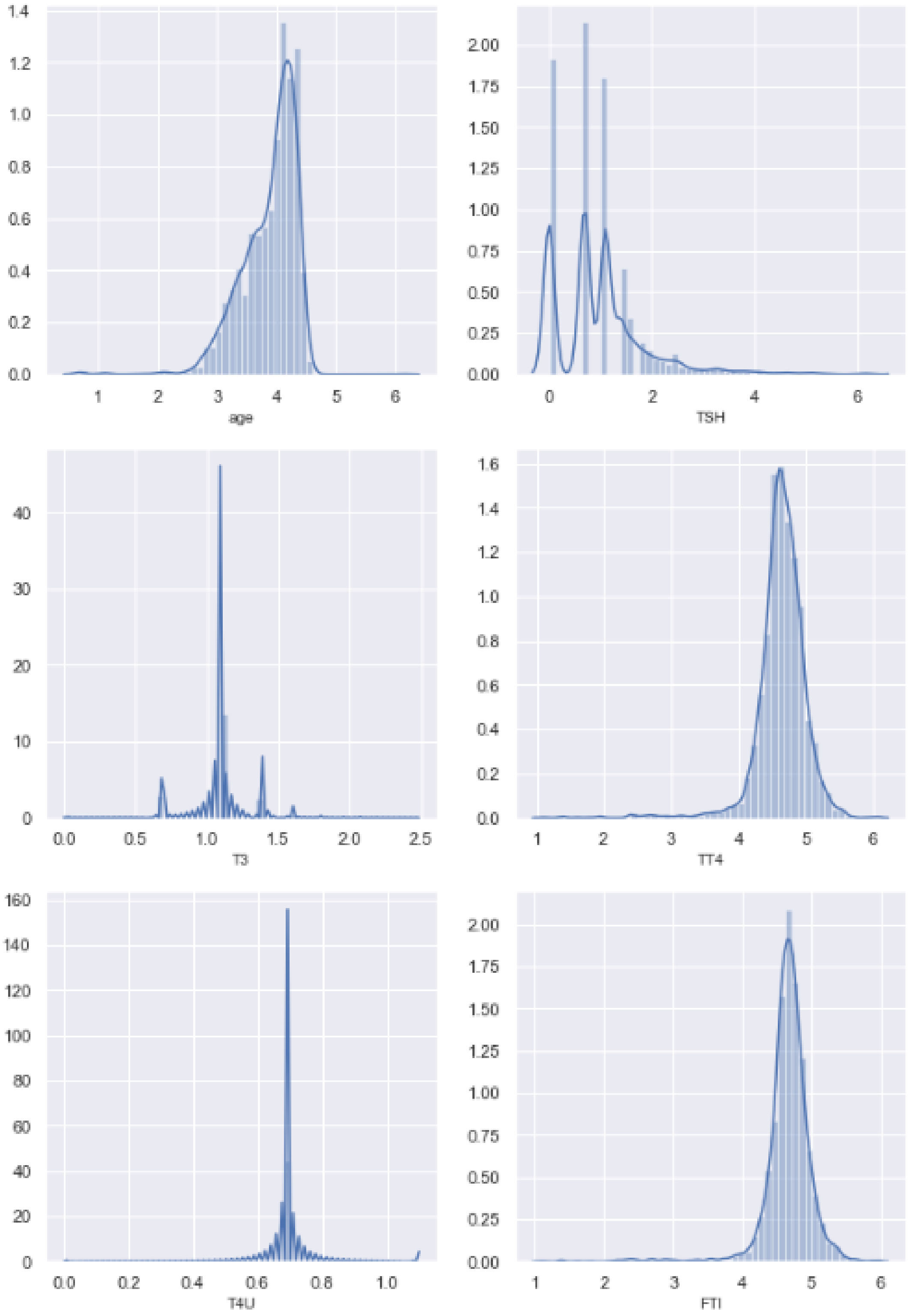


Fig.5.7 After Transformation

After log transformation, rest of the columns look fine but 'TSH' has a weird trend.

It won't give much of information so we will drop the column.

- **Checking how balanced the dataset is, in terms of given target classes:**

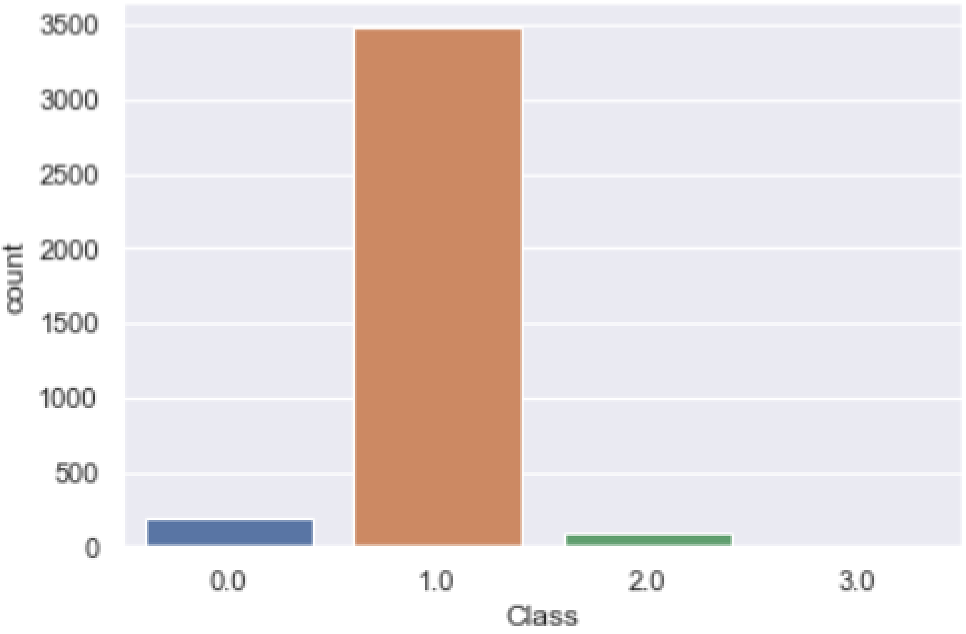


Fig.5.8-imbalanced data

It can clearly be seen in above figure that the dataset is highly imbalanced.
The python library known as imbalanced-learn will be used here to deal with imbalanced data. Imbalanced learn has an algorithm called RandomOverSampler.

RandomOverSampler: Random resampling provides a naive technique for rebalancing the class distribution for an imbalanced dataset.
Random oversampling duplicates examples from the minority class in the training dataset and can result in overfitting for some models.
Random undersampling deletes examples from the majority class and can result in losing information invaluable to a model.
After using the particular algorithm:

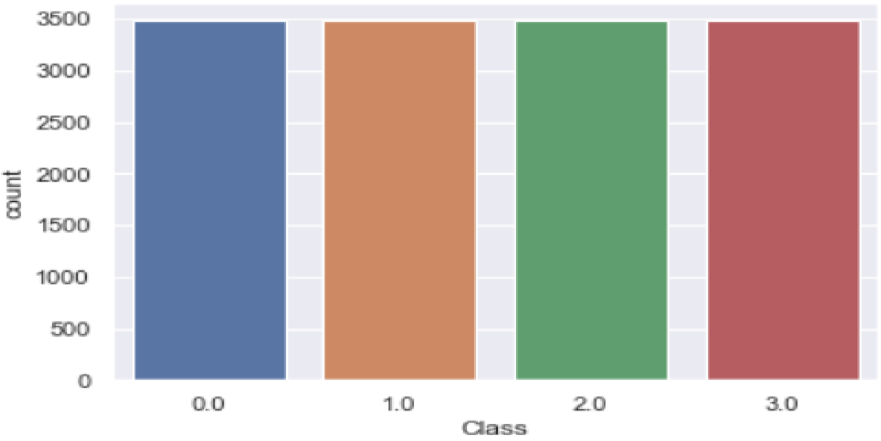


Fig.5.9-Balanced data

The dataset is balanced now.

5.6 Prediction

- **Algorithms used:**

1. Decision Tree Classifier: Decision tree classifiers provide a readable classification model that is potentially accurate in many different application contexts, including energy-based applications. The decision tree classifier (Pang-Ning et al., 2006) creates the classification model by building a decision tree. Each node in the tree specifies a test on an attribute, each branch descending from that node corresponds to one of the possible values for that attribute.
2. Random Forest Classifier: A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.
3. Support Vector Classifier: Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.
   The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.
4. Logistic Regression: Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.
5. KNeighborsClassifier: KNeighborsClassifier is based on the k nearest neighbors of a sample, which has to be classified. The number 'k' is an integer value specified by the user. This is the most frequently used classifiers of both algorithms.

- **Confusion Matrix:**
  A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The matrix compares the actual target values with those predicted by the machine learning model. This gives us a holistic view of how well our classification model is performing and what kinds of errors it is making.

		Predictive Class		
		Class-1	Class-2	Class-3
Actual Class	Class-1	Cell 1 True	Cell 2 False	Cell 3 False
	Class-2	Cell 4 False	Cell 5 True	Cell 6 False
	Class-3	Cell 7 False	Cell 8 False	Cell 9 True

Fig 5.10-Confusion matrix

True Positive(TP)	positive class data points that are correctly classified by the model
True Negative(TN)	negative class data points that are correctly classified by the model
False Positive(FP)	negative class data points that are incorrectly classified as belonging to the positive class by the model
False Negative(FN)	positive class data points that are incorrectly classified as belonging to the negative class by the model

Table 5.1- Values of confusion matrix

Unlike binary classification, there are no positive or negative classes here. At first, it might be a little difficult to find TP, TN, FP and FN since there are no positive or negative classes. What we have to do here is to find TP, TN, FP and FN for each individual class.
Let us calculate the TP, TN, FP, FN values for the class-1
TP: The actual value and predicted value should be the same. So concerning class-1, the value of cell-1 is the TP value.

FN: The sum of values of corresponding rows except the TP value
FN = (cell 2 + cell3)

FP: The sum of values of corresponding column except the TP value.
FP = (cell 4 + cell 7)

TN: The sum of values of all columns and row except the values of that class that we are calculating the values for.
TN = (cell 5 + cell 6 + cell 8 + cell 9)

= 17 + 1 +0 + 11

= 29

Similarly, for Class-2 the values/ metrics are calculated as below:

TP: (cell 5)

FN: (cell 4 +cell 6)

FP: (cell 2 + cell 8)

TN: (cell 1 + cell 3 + cell 7 + cell 9).

Similarly, for Class-3 the values/ metrics are calculated as below:
TP: (cell 9)

FN: (cell 7 +cell 8)

FP: (cell 3 + cell 6)

TN: (cell 1 + cell 2 + cell 4 + cell 5).

- Confusion Matrix for Random Forest Classifier algorithm

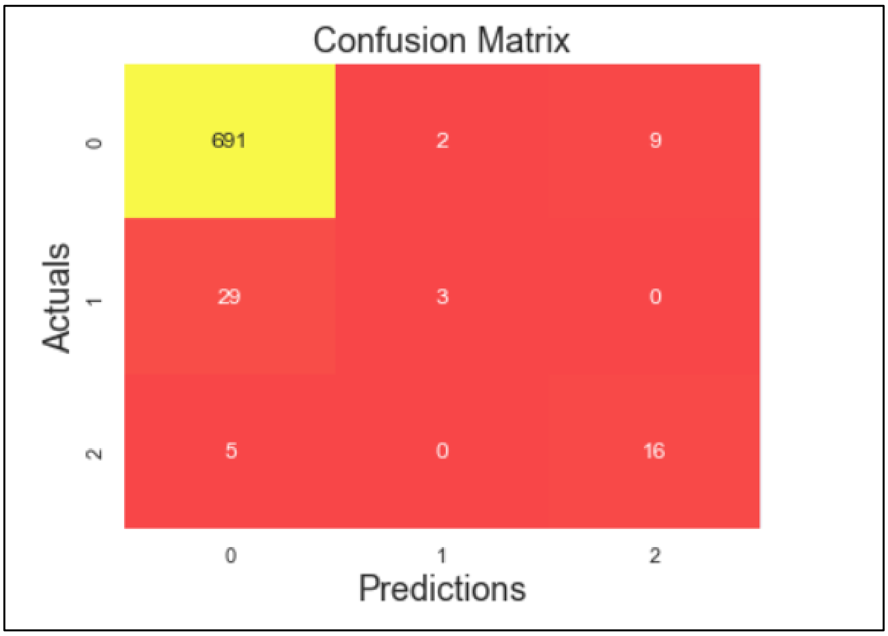


Fig 5.10.1- Confusion Matrix-1

- Confusion Matrix for K Nearest Neighbors algorithm

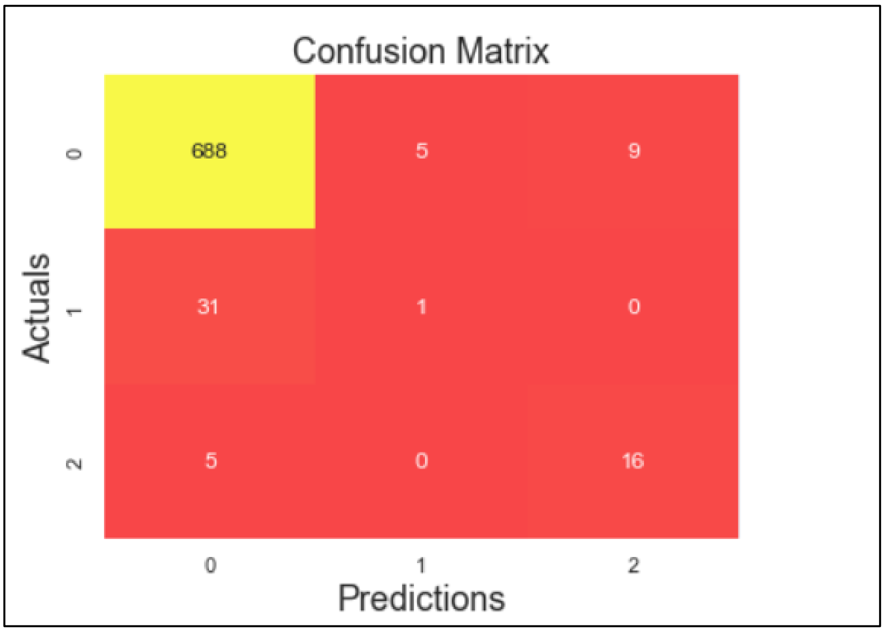


Fig 5.10.2- Confusion Matrix-2

- Confusion Matrix for Logistic Regression algorithm

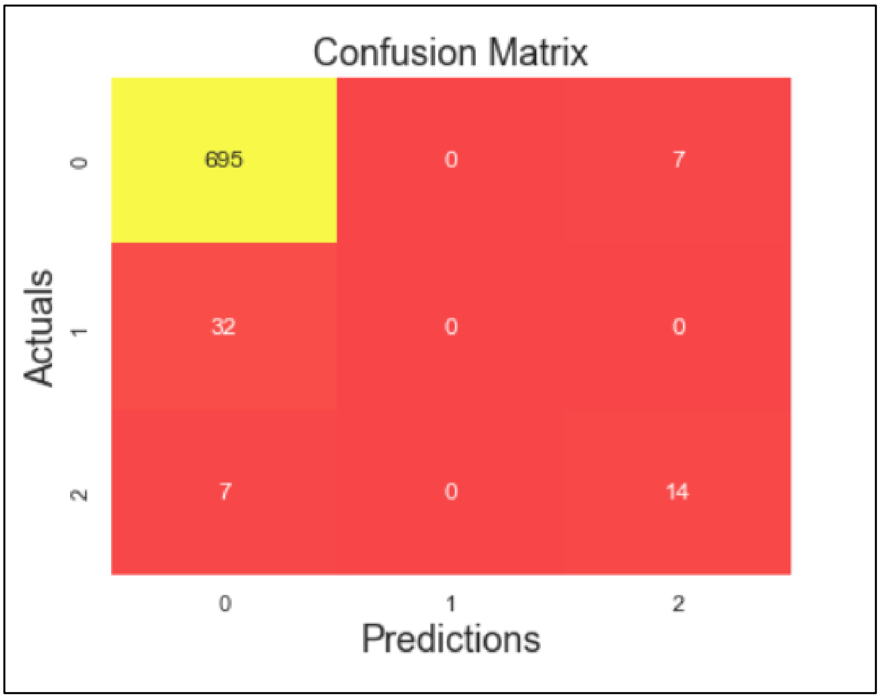


Fig 5.10.3- Confusion Matrix-3

- Confusion Matrix for Decision Tree Classifier algorithm

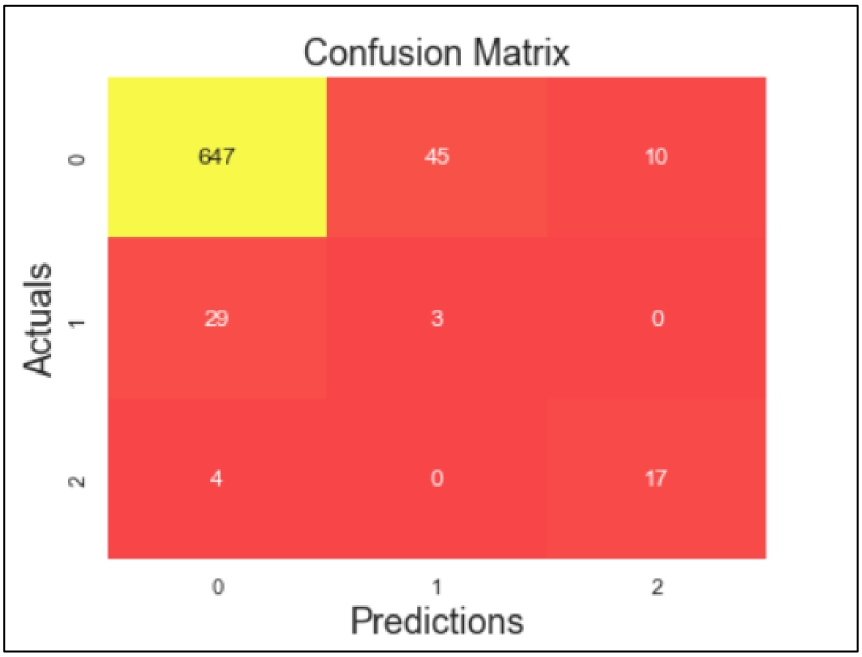


Fig 5.10.4- Confusion Matrix-4

- Confusion Matrix for Support Vector Classifier algorithm

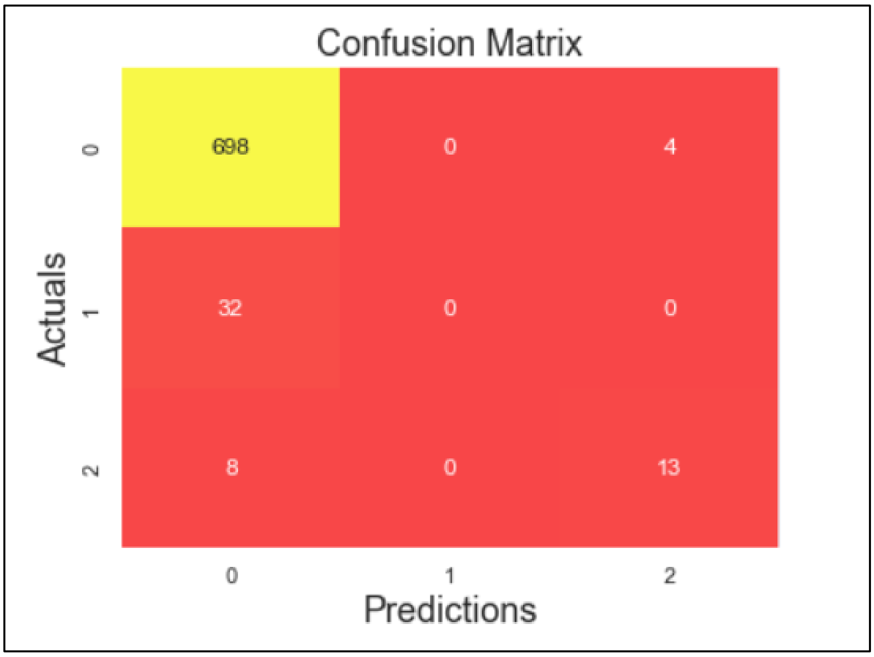


Fig 5.10.5- Confusion Matrix-5

- **Accuracy Score and classification report:**
  <u>Accuracy Score:</u>

Sno.	Algorithms	Accuracy Score
1	Decision Tree Classifier	88.344371
2	Support Vector Machine	94.172185
3	Random Forest Classifier	94.039735
4	K neighbors Classifier	93.377483
5	Logistic Regression	93.907285
	Highest Accuracy “Support Vector Machine”	94.172185

Table 5.2 Accuracy Score

It can be seen that the accuracy of Support Vector Machine algorithm is higher amongst all the algorithms. So we will be using Support Vector Machine to predict the data.

- **Classification Report:**

A classification report is a performance evaluation metric in machine learning. It is used to show the precision, recall, F1 Score, and support of the trained classification model.

Metrics	Definition
Precision	Precision is defined as the ratio of true positives to the sum of true and false positives.
Recall	Recall is defined as the ratio of true positives to the sum of true positives and false negatives.
F1 Score	The F1 is the weighted harmonic mean of precision and recall. The closer the value of the F1 score is to 1.0, the better the expected performance of the model is.
Support	Support is the number of actual occurrences of the class in the dataset. It doesn't vary between models, it just diagnoses the performance evaluation process.

Table 5.3 Definition of metrics in classification report

Sno.	**Algorithms**	**Precision**			**Recall**			**F1-score**			**support**		
		0	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**
1	Random Forest Classifier	0.95	0.0	0.67	0.99	0.0	0.67	0.97	0.0	0.67	702	32	21
2	Support Vector Machine	0.95	0.0	0.76	0.99	0.0	0.62	0.97	0.0	0.68	702	32	21
3	Decision Tree Classifier	0.95	0.11	0.58	0.93	0.16	0.71	0.94	0.13	0.64	702	32	21
4	K neighbors Classifier	0.95	0.17	0.64	0.98	0.03	0.76	0.96	0.05	0.70	702	32	21
5	Logistic Regression	0.95	0.0	0.67	0.99	0.0	0.67	0.97	0.0	0.67	702	32	21

Table 5.4 Classification Report

CHAPTER-6

TESTING

6.1 Introduction

Executing a program with the intent of finding errors is called testing. Testing is vital to the success of any system. Testing is done at different stages within the development phase. System testing makes a logical assumption that if all parts of the system are correct, the goals will be achieved successfully. Inadequate testing or no testing at all leads to errors that may come up after a long time when correction would be extremely implementation. The testing of the system was done on both artificial and live data. In order to test data test cases are developed. Following are the various methods that are employed for testing.

6.2 Methods Employed For Testing

Unit Testing

In unit testing the module is tested independently. It is done to test that the module does satisfy the functional specification. This is done to check syntax and logical errors in programs. At the time of preparation of technical specifications, unit test data was also prepared. The coding for that program was considered after verifying its output against this test data.
Following are the unit testing methods:

In Conditional Testing, the logical conditions that are given in the module were checked to see whether they satisfy the functionality of the module. This is done by using the test data was prepared.
In Loop Testing, different loops in the module like nested loops were tested using the data. Attempts to execute the loops to their maximum range are done.

Integration Testing

In Integration testing whole system was checked when all the individual modules were integrated together in order to test whether the system is performing as according to the requirements specified. Interface errors if any were corrected. Test data was prepared was fed into the system to check whether the system fails to detects an error.

Functional Testing

This is done for each module/sub module of the system. Functional testing serve as a means of validating whatever the functionality of the system confers the original user requirement i.e. does the module do what it was supposed to do?
Separate schedules were made for functional testing. It involves preparation of test data, writing test 26 cases, testing for conformance to test cases and preparation of bug's listing for nonconformities.

System Testing

System testing is done when the entire system has been fully integrated. The purpose of the system testing is to test how the different modules interact with each other and whether the entire system provides the functionality that was expected.

- System testing consists of the following steps:
- Program Testing
- System Testing
- System Documentation
- User Acceptance Testing

6.3 Test Cases

NO.	TEST CASE TITLE	DESCRIPTION	EXPECTED OUTCOMES	RESULT
1	Successful User Verification	The login to the system should be tried with the login name and the correct password.	Login should be successful and the user should be connected to server.	Passed
2	Unsuccessful User Verification due to wrong password	Login to the system with a wrong password	Login should fail with an error 'Invalid password'.	Passed
3	Unsuccessful User Verification due to invalid Login id	Login to the server with an invalid login id	Login should fail with an error 'Invalid Login Id'.	Passed
4	The server should provide the admin for allowing newly registered user to access the website.	An option for allowing new users and deleting old users of the system	Option for adding/removing users is shown.	Passed
5	The system should provide sign up form for creating new user Login id	A form for creating new user's login id of the system	Form for register users is shown	Passed
6	Unsuccessful user verification if they do not enter their all details being asked.	Filling the registration form without entering the whole details	Field can't be empty	Passed
5	Clicking Logout Button	On clicking logout button successful logout will happen	Logout message will be popup	Passed

Table 6.1- Test Cases

CHAPTER-7

CONCLUSION

7.1 CONCLUSION

The “Thyroid Detection System” has been designed to make it easier for the user or patients to not only detect or predict whether they are suffering from thyroid or not but also it will be providing us the type of thyroid you are affected with in the case of positive results. This will be achieved by performing data validations, data pre-processing, EDA. Then by using KNN and Random Forest Algorithm we will be able to predict the results with the help of details provided by the user.

First the user will login or register and then after login in the user will have the access to the prediction page. Apart from this we have profile page to show the details of the user and separate update page for updating the details to your database.

On the prediction page, user enter certain details such as age, sex, Are you on thyroxine treatment or not?, Are you pregnant or not?, Are you on antithyroid medication or not?, etc . And then with the help of entered details we will predict and conclude the result.

In this project, we have used various python modules and libraries like flask, pymysql, numpy, pandas, sklearn, etc. which will help us achieve the prediction.

7.2 FUTURE SCOPE

This project can be implemented under various situations. We can add new features as and when we require.

- We can add more columns to increase the accuracy for prediction purpose.
- We can analyze and show the analysis part also to make it easier for user to understand about the disease very well, the impact on patients and many more.
- We can add more modules which may provide us the recommendations and consultations regarding the type of thyroid a person is suffering from.

It will give a better visualization and an effective impression to the user.

7.3 Project Roles and Responsibilities

- **ROLE**

Worked as a developer, designer, tester of the application.

- **RESPONSIBILITIES**

  - Work on definition of development requirements and priorities.
  - Data migration.
  - Interfaces with other systems.
  - Reporting configuration and deployment.
  - Set up and maintenance of security rights and access permission.
  - Contributing to technical strategy, policy and procedure.
  - Development and operation of technical testing programs.
  - Production of technical documentation to agreed quality standards.
  - Reporting on progress/issues to management and users.

APPENDICES

Code:

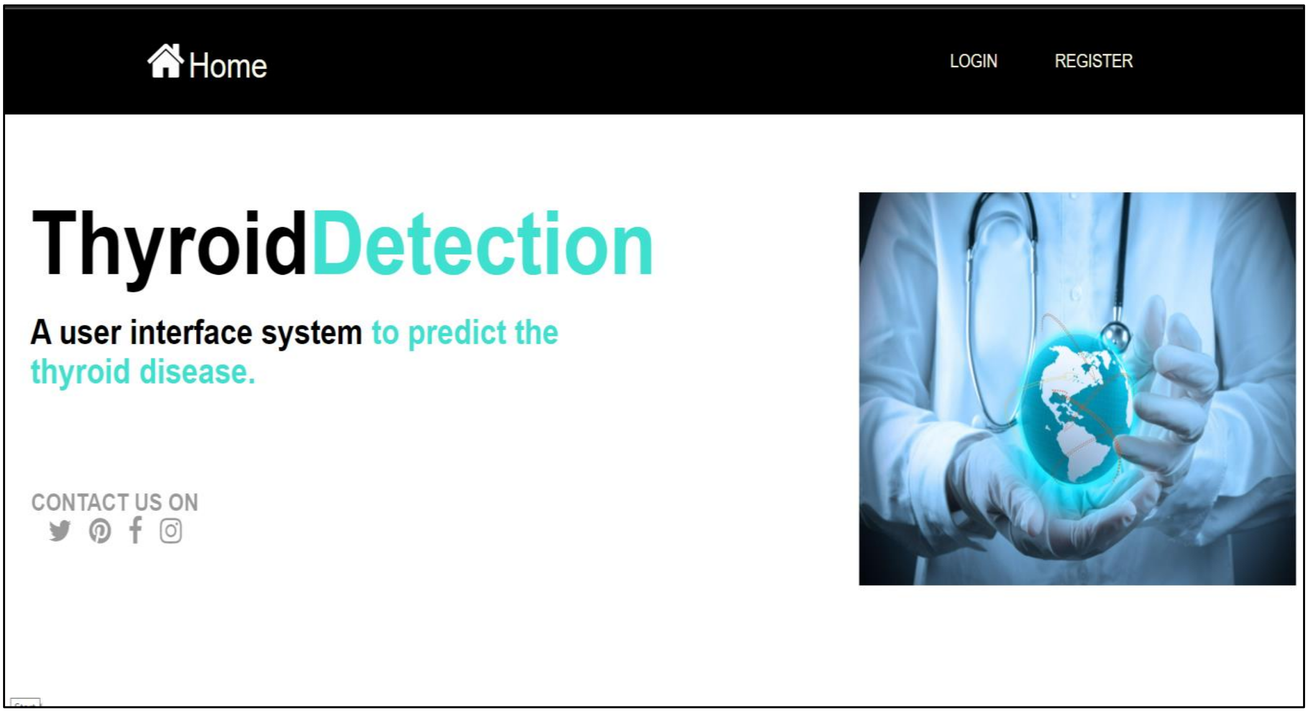
- **App.py**

```
import sklearn

import pickle

import numpy as np

import re

from flask_mysqldb import MySQL

import MySQLdb.cursors

from flask import Flask,render_template,request,redirect,url_for,session

app= Flask(__name__)

model=pickle.load(open('svc_model.pkl','rb'))


app.secret_key='qwqwqw'


app.config['MYSQL_HOST'] = 'localhost'

app.config['MYSQL_USER'] = 'root'

app.config['MYSQL_PASSWORD'] = 'admin'

app.config['MYSQL_DB'] = 'thyroid'

mysql = MySQL(app)

@app.route('/')

def welcome():

    return render_template('index.html')


@app.route('/login', methods =['GET', 'POST'])

def login():

    msg = ''
    if request.method == 'POST' and 'username' in request.form and 'password' in request.form:
        username = request.form['username']
        password = request.form['password']
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('SELECT * FROM thyroid.thyroid_detection WHERE username = % s AND password = % s', (username, password, ))
        account = cursor.fetchone()
        if account:
            session['loggedin'] = True
            session['id'] = account['id']
            session['username'] = account['username']
            msg = 'Logged in successfully !'
            return render_template('homepage.html')
        else:
            msg = 'Incorrect username / password !'
    return render_template('login.html', msg = msg)
@app.route('/contact',methods=['GET','POST'])
def contact():
    if request.method == 'POST' and 'nam' in request.form and 'em' in request.form and 'phone' in request.form and 'message' in request.form:
        nam = request.form['nam']
        em = request.form['em']
        phone = request.form['phone']
        message= request.form['message']
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('INSERT INTO thyroid.query VALUES(NULL,%s, % s, % s, % s)',
                       (nam,em,phone,message,))
        mysql.connection.commit()
        return render_template('contact.html', msg='Submitted Successfully !')
    else:
        return render_template('contact.html')


@app.route("/home",methods=['GET','POST'])
def home():
    return render_template("homepage.html")
@app.route("/thyroid_disease",methods=['GET','POST'])
def thyroid_disease():
    return render_template("thyroid_disease.html")
@app.route("/Symptoms",methods=['GET','POST'])
def Symptoms():
    return render_template("Symptoms.html")
@app.route("/prevention",methods=['GET','POST'])
def prevention():
    return render_template("prevention.html")
@app.route("/Causes",methods=['GET','POST'])
def Causes():
    return render_template("Causes.html")
@app.route('/logout')
def logout():
   session.pop('loggedin', None)
   session.pop('id', None)
   session.pop('username', None)
   return render_template('index.html')
@app.route('/register', methods =['GET', 'POST'])
```

```
def register():

    msg = ''

    if request.method == 'POST'and 'firstName' in request.form and 'lastName' in request.form and 'username' in request.form and 'password' in request.form and 'inlineRadioOptions' in request.form and 'phone_number' in request.form and 'mail' in request.form:

        firstName = request.form['firstName']

        lastName = request.form['lastName']

        username = request.form['username']

        password = request.form['password']

        inlineRadioOptions = request.form['inlineRadioOptions']

        phone_number = request.form['phone_number']

        mail = request.form['mail']

        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)

        cursor.execute('SELECT * FROM thyroid.thyroid_detection WHERE username = % s', (username, ))

        account = cursor.fetchone()

        if account:

            msg = 'Account already exists !'

        elif not re.match(r'[^@]+@[^@]+\.[^@]+', mail):

            msg = 'Invalid email address !'

        elif not re.match(r'[A-Za-z0-9]+', username):

            msg = 'name must contain only characters and numbers !'

        else:

            cursor.execute('INSERT INTO thyroid.thyroid_detection VALUES (NULL,%s, % s, % s, % s, % s, % s, % s)', (firstName,lastName, username, password,inlineRadioOptions,phone_number, mail,))

            mysql.connection.commit()

            msg = 'You have successfully registered !'

    elif request.method == 'POST':
        msg = 'Please fill out the form !'
    return render_template('register.html', msg = msg)



@app.route("/predict",methods=['GET','POST'])
def predict():
    prediction = ''

    if request.method == 'POST':
        age = request.form['age']
        gender=request.form['gender']
        on_thyroxine=request.form['on_thyroxine']
        query_on_thyroxine=request.form['query_on_thyroxine']
        on_antithyroid_medication=request.form['on_antithyroid_medication']
        sick=request.form['sick']
        pregnant=request.form['pregnant']
        thyroid_surgery=request.form['thyroid_surgery']
        I131_treatment=request.form['I131_treatment']
        query_hypothyroid=request.form['query_hypothyroid']
        query_hyperthyroid=request.form['query_hyperthyroid']
        lithium=request.form['lithium']
        goitre = request.form['goitre']
        tumor = request.form['tumor']
        hypopituitary = request.form['hypopituitary']
        psych = request.form['psych']
        T3 = request.form['T3']
        TT4 = request.form['TT4']
        T4U = request.form['T4U']
        FTI = request.form['FTI']
        referral_source = request.form['referral_source']
        predict_result = model.predict([[age, gender, on_thyroxine, query_on_thyroxine,
        on_antithyroid_medication, sick, pregnant, thyroid_surgery,
        I131_treatment, query_hypothyroid, query_hyperthyroid, lithium,
        goitre, tumor, hypopituitary, psych, T3, TT4, T4U, FTI,
        referral_source]])
        if predict_result==0:
            prediction="negative"
            return render_template("result2.html", prediction=prediction)
        elif predict_result==1:
            prediction="compensated_hypothyroid  "
            return render_template("result.html", prediction=prediction)
        elif predict_result ==2:
            prediction="primary_hypothyroid"
            return render_template("result4.html", prediction=prediction)
        else:
            prediction="secondary_hypothyroid"
            return render_template("result3.html", prediction=prediction)
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute(
            'INSERT INTO thyroid.predict VALUES (NULL,%s, % s, % s, % s, % s, % s, %
    s,%s, % s, % s, % s, % s, % s, % s,%s, % s, % s, % s, % s, % s,% s,% s)',
            (age, gender, on_thyroxine, query_on_thyroxine,
             on_antithyroid_medication, sick, pregnant, thyroid_surgery,
             I131_treatment, query_hypothyroid, query_hyperthyroid, lithium,
            goitre, tumor, hypopituitary, psych, T3, TT4, T4U, FTI,
            referral_source,prediction,))
        mysql.connection.commit()
        cursor = mysql.connection.cursor(MySQLdb.cursors.DictCursor)
        cursor.execute('SELECT * FROM thyroid.predict where id= % s', (session['id'],))
        predict = cursor.fetchone()
        # return render_template("prediction.html", predict_res=predict_res)
    else:
        return render_template("predict.html")
@app.route('/result')
def result():
    return render_template("result.html")
if __name__=="__main__" :
    app.run(host="localhost",port=int("5000"), debug=True, use_reloader = True)
```

8.2 SNAPSHOTS

1. Welcome Page:



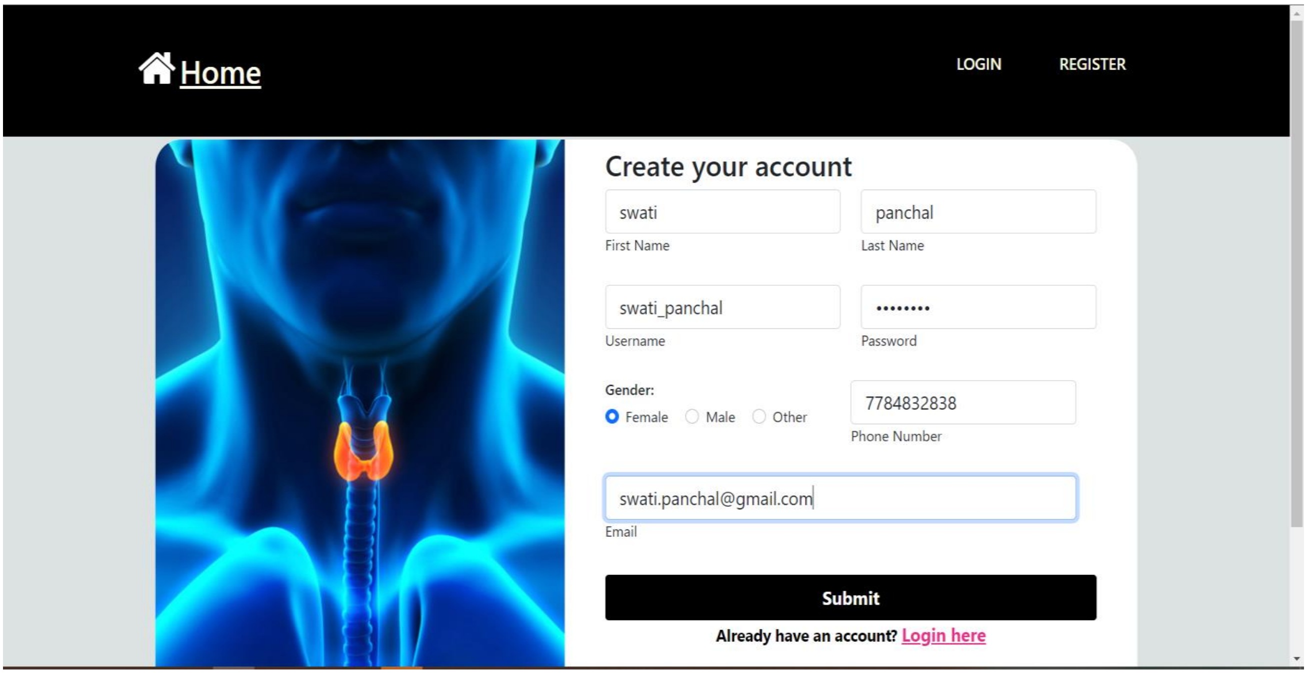
ThyroidDetection

A user interface system to predict the thyroid disease.

CONTACT US ON

Fig.8.1 Welcome Page

2. Sign-up page:



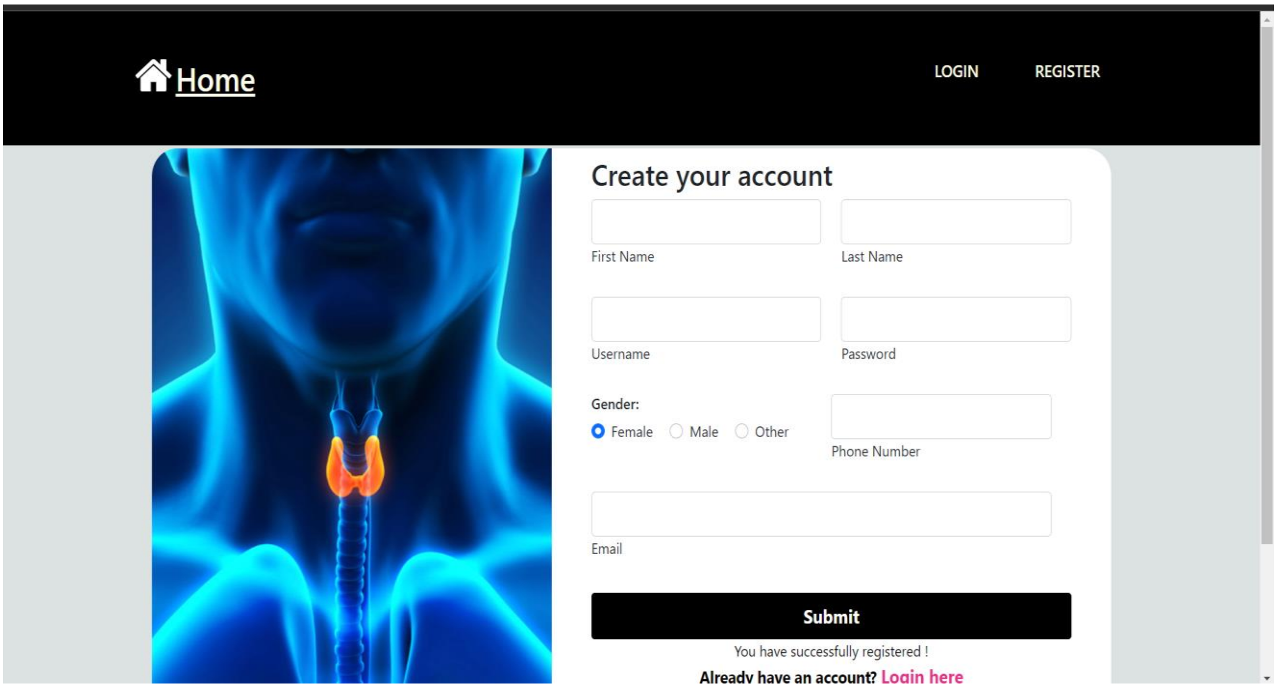
Create your account

swati First Name	panchal Last Name
swati_panchal Username	•••••••• Password
Gender: ● Female ○ Male ○ Other	7784832838 Phone Number
swati.panchal@gmail.com Email	

Already have an account? Login here

Fig.8.2 Sign-up Page

3. Successfully Registered:

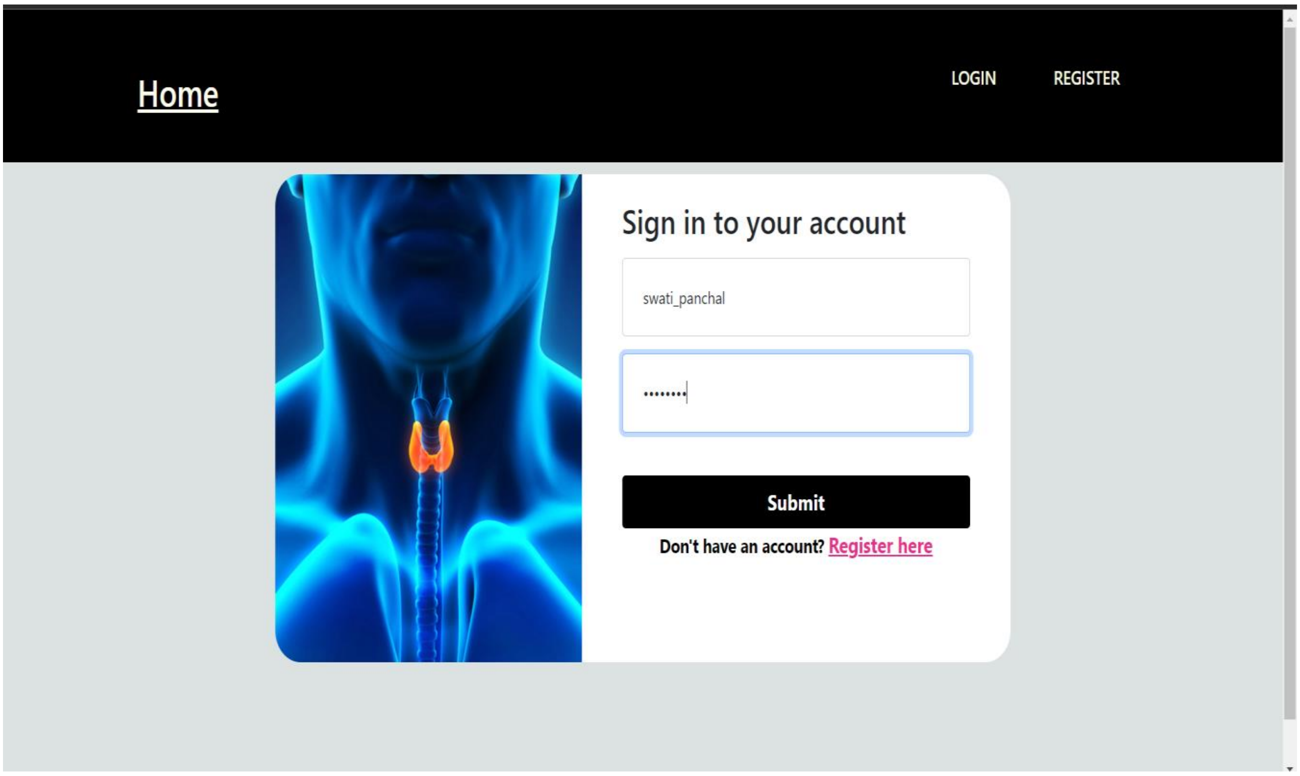


You have successfully registered !

Already have an account? Login here

Fig.8.3 Successfully Registered

4. Sign-in Page:

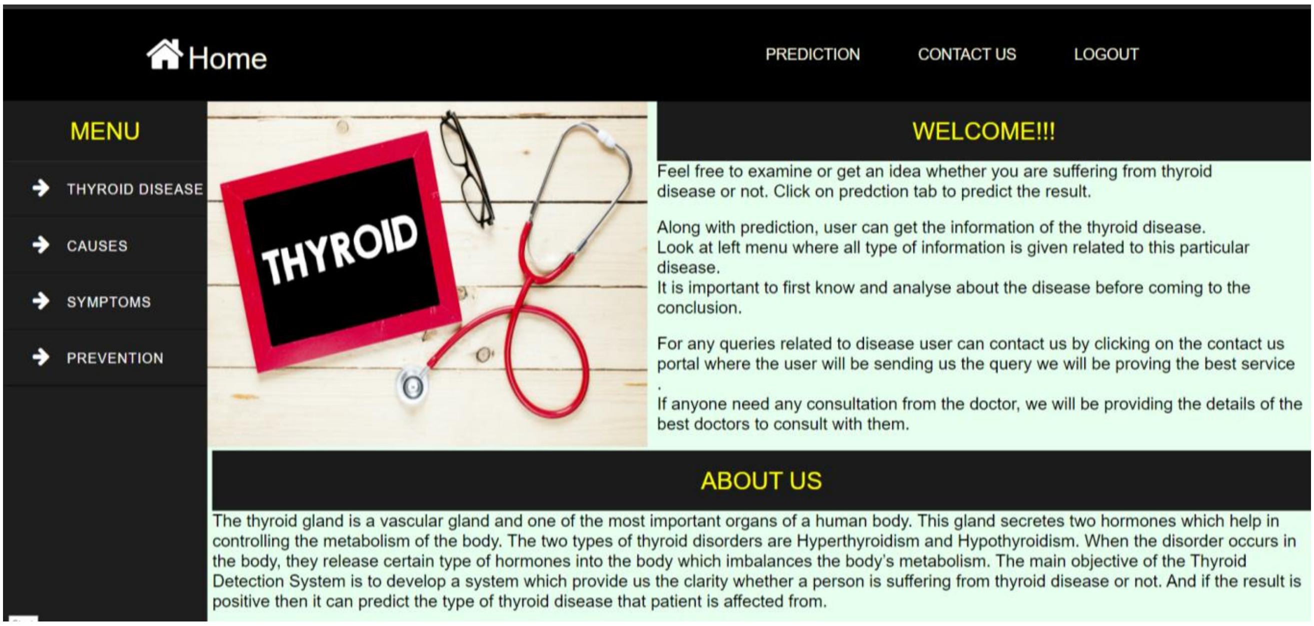


swati_panchal

••••••••

Fig.8.4 Sign-in Page

5. After Successful login Home Page appears:



WELCOME!!!

Feel free to examine or get an idea whether you are suffering from thyroid disease or not. Click on predction tab to predict the result.

Along with prediction, user can get the information of the thyroid disease.
Look at left menu where all type of information is given related to this particular disease.
It is important to first know and analyse about the disease before coming to the conclusion.

For any queries related to disease user can contact us by clicking on the contact us portal where the user will be sending us the query we will be proving the best service .

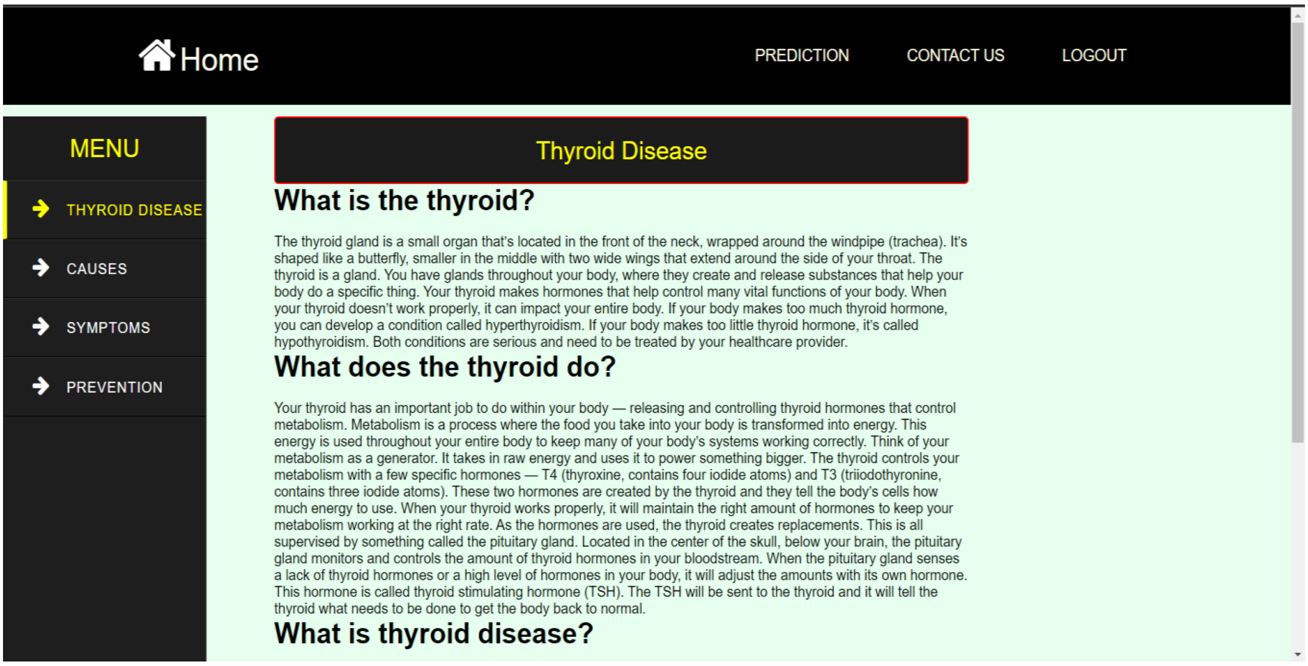
If anyone need any consultation from the doctor, we will be providing the details of the best doctors to consult with them.

ABOUT US

The thyroid gland is a vascular gland and one of the most important organs of a human body. This gland secretes two hormones which help in controlling the metabolism of the body. The two types of thyroid disorders are Hyperthyroidism and Hypothyroidism. When the disorder occurs in the body, they release certain type of hormones into the body which imbalances the body's metabolism. The main objective of the Thyroid Detection System is to develop a system which provide us the clarity whether a person is suffering from thyroid disease or not. And if the result is positive then it can predict the type of thyroid disease that patient is affected from.

Fig.8.5 Home Page

6. Menu: Thyroid Disease



Thyroid Disease

What is the thyroid?

The thyroid gland is a small organ that's located in the front of the neck, wrapped around the windpipe (trachea). It's shaped like a butterfly, smaller in the middle with two wide wings that extend around the side of your throat. The thyroid is a gland. You have glands throughout your body, where they create and release substances that help your body do a specific thing. Your thyroid makes hormones that help control many vital functions of your body. When your thyroid doesn't work properly, it can impact your entire body. If your body makes too much thyroid hormone, you can develop a condition called hyperthyroidism. If your body makes too little thyroid hormone, it's called hypothyroidism. Both conditions are serious and need to be treated by your healthcare provider.

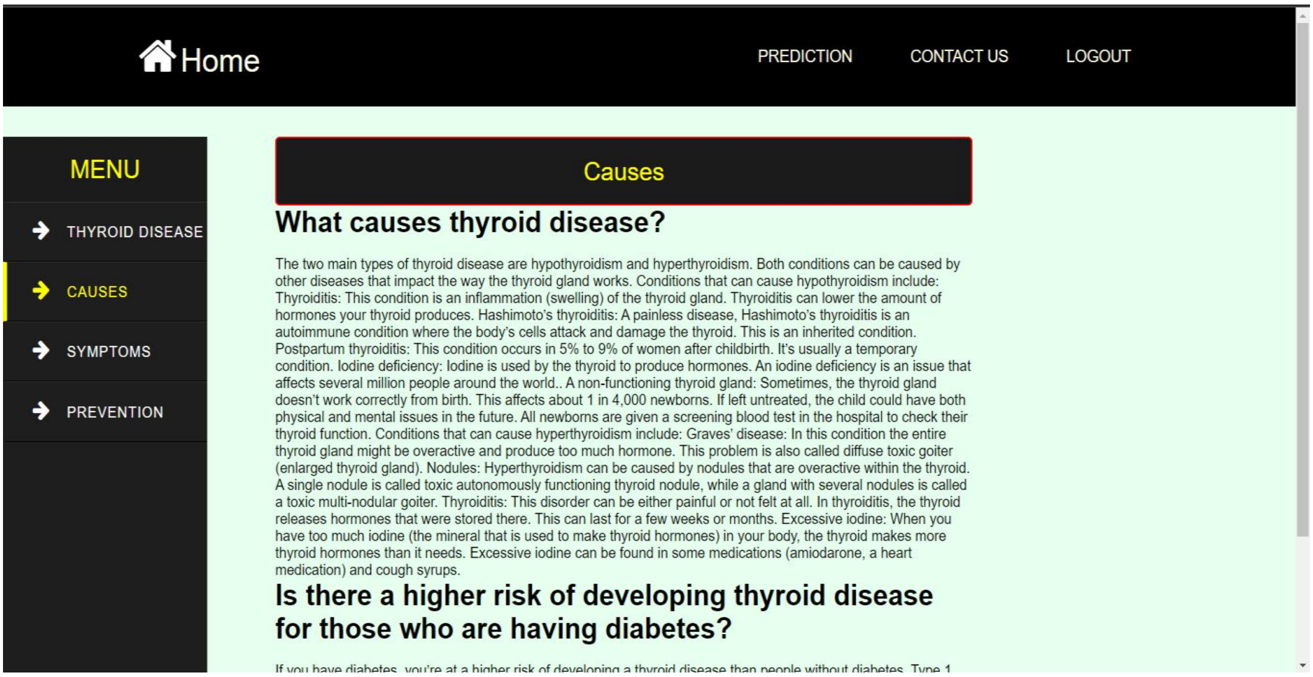
What does the thyroid do?

Your thyroid has an important job to do within your body — releasing and controlling thyroid hormones that control metabolism. Metabolism is a process where the food you take into your body is transformed into energy. This energy is used throughout your entire body to keep many of your body's systems working correctly. Think of your metabolism as a generator. It takes in raw energy and uses it to power something bigger. The thyroid controls your metabolism with a few specific hormones — T4 (thyroxine, contains four iodide atoms) and T3 (triiodothyronine, contains three iodide atoms). These two hormones are created by the thyroid and they tell the body's cells how much energy to use. When your thyroid works properly, it will maintain the right amount of hormones to keep your metabolism working at the right rate. As the hormones are used, the thyroid creates replacements. This is all supervised by something called the pituitary gland. Located in the center of the skull, below your brain, the pituitary gland monitors and controls the amount of thyroid hormones in your bloodstream. When the pituitary gland senses a lack of thyroid hormones or a high level of hormones in your body, it will adjust the amounts with its own hormone. This hormone is called thyroid stimulating hormone (TSH). The TSH will be sent to the thyroid and it will tell the thyroid what needs to be done to get the body back to normal.

What is thyroid disease?

Fig.8.6 Menu

7. Menu: Causes



Causes

What causes thyroid disease?

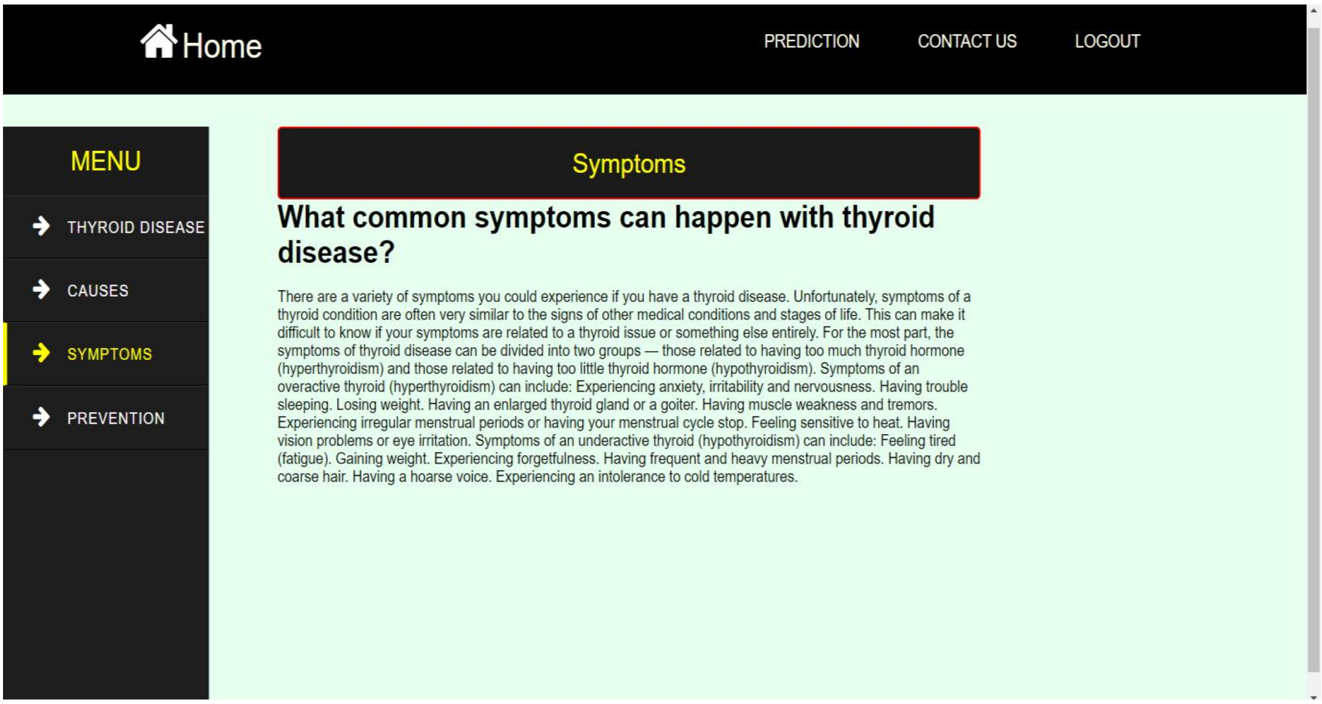
The two main types of thyroid disease are hypothyroidism and hyperthyroidism. Both conditions can be caused by other diseases that impact the way the thyroid gland works. Conditions that can cause hypothyroidism include: Thyroiditis: This condition is an inflammation (swelling) of the thyroid gland. Thyroiditis can lower the amount of hormones your thyroid produces. Hashimoto's thyroiditis: A painless disease, Hashimoto's thyroiditis is an autoimmune condition where the body's cells attack and damage the thyroid. This is an inherited condition. Postpartum thyroiditis: This condition occurs in 5% to 9% of women after childbirth. It's usually a temporary condition. Iodine deficiency: Iodine is used by the thyroid to produce hormones. An iodine deficiency is an issue that affects several million people around the world.. A non-functioning thyroid gland: Sometimes, the thyroid gland doesn't work correctly from birth. This affects about 1 in 4,000 newborns. If left untreated, the child could have both physical and mental issues in the future. All newborns are given a screening blood test in the hospital to check their thyroid function. Conditions that can cause hyperthyroidism include: Graves' disease: In this condition the entire thyroid gland might be overactive and produce too much hormone. This problem is also called diffuse toxic goiter (enlarged thyroid gland). Nodules: Hyperthyroidism can be caused by nodules that are overactive within the thyroid. A single nodule is called toxic autonomously functioning thyroid nodule, while a gland with several nodules is called a toxic multi-nodular goiter. Thyroiditis: This disorder can be either painful or not felt at all. In thyroiditis, the thyroid releases hormones that were stored there. This can last for a few weeks or months. Excessive iodine: When you have too much iodine (the mineral that is used to make thyroid hormones) in your body, the thyroid makes more thyroid hormones than it needs. Excessive iodine can be found in some medications (amiodarone, a heart medication) and cough syrups.

Is there a higher risk of developing thyroid disease for those who are having diabetes?

If you have diabetes, you're at a higher risk of developing a thyroid disease than people without diabetes. Type 1

Fig.8.7 Causes

8. Menu: Symptoms



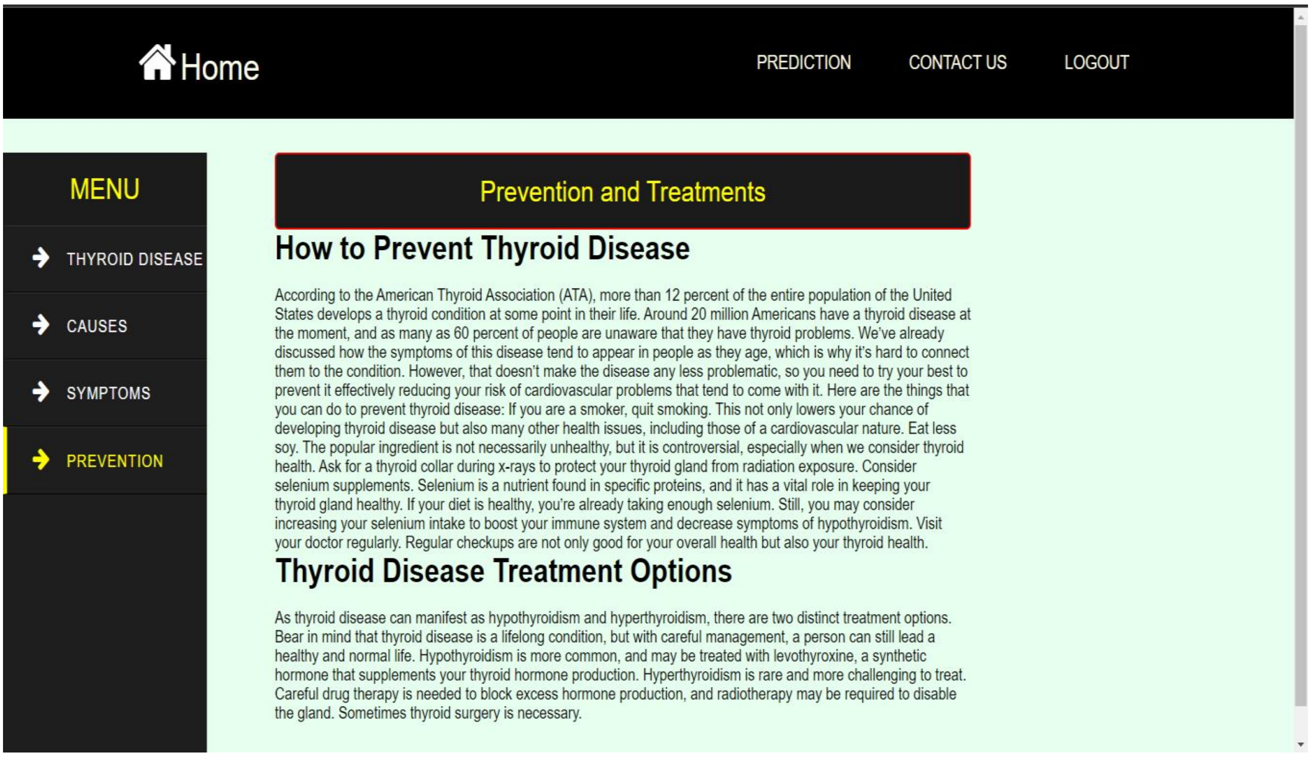
Symptoms

What common symptoms can happen with thyroid disease?

There are a variety of symptoms you could experience if you have a thyroid disease. Unfortunately, symptoms of a thyroid condition are often very similar to the signs of other medical conditions and stages of life. This can make it difficult to know if your symptoms are related to a thyroid issue or something else entirely. For the most part, the symptoms of thyroid disease can be divided into two groups — those related to having too much thyroid hormone (hyperthyroidism) and those related to having too little thyroid hormone (hypothyroidism). Symptoms of an overactive thyroid (hyperthyroidism) can include: Experiencing anxiety, irritability and nervousness. Having trouble sleeping. Losing weight. Having an enlarged thyroid gland or a goiter. Having muscle weakness and tremors. Experiencing irregular menstrual periods or having your menstrual cycle stop. Feeling sensitive to heat. Having vision problems or eye irritation. Symptoms of an underactive thyroid (hypothyroidism) can include: Feeling tired (fatigue). Gaining weight. Experiencing forgetfulness. Having frequent and heavy menstrual periods. Having dry and coarse hair. Having a hoarse voice. Experiencing an intolerance to cold temperatures.

Fig.8.8 Symptoms

9. Menu: Prevention and Treatments



Prevention and Treatments

How to Prevent Thyroid Disease

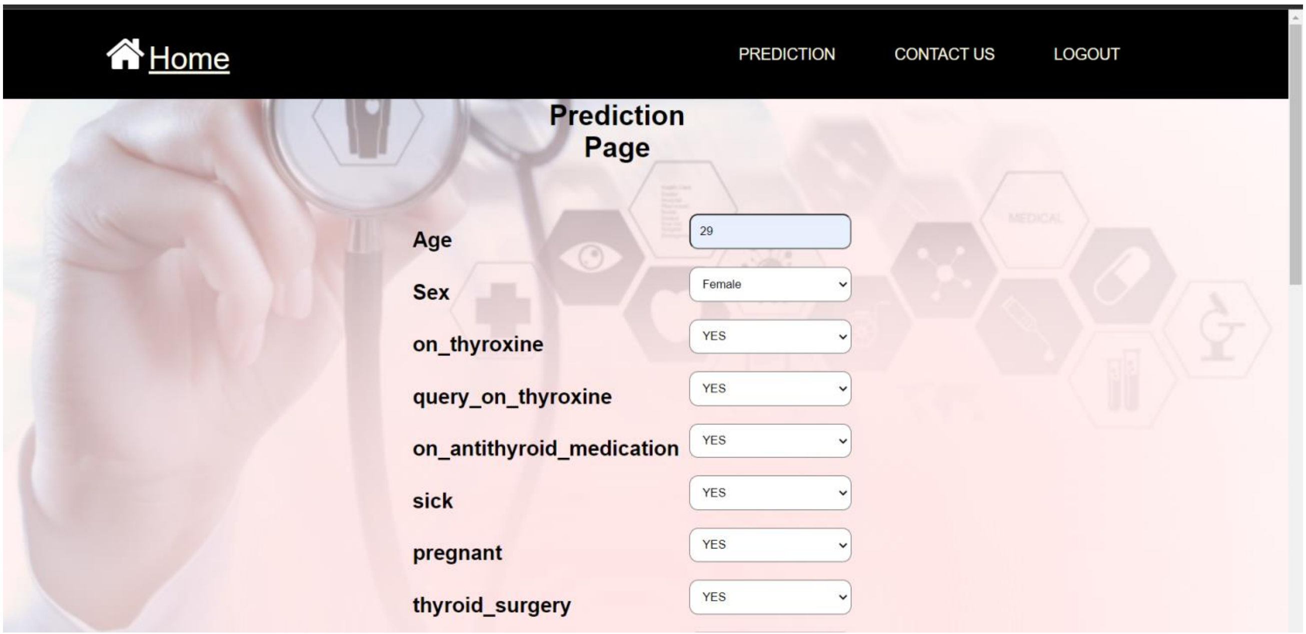
According to the American Thyroid Association (ATA), more than 12 percent of the entire population of the United States develops a thyroid condition at some point in their life. Around 20 million Americans have a thyroid disease at the moment, and as many as 60 percent of people are unaware that they have thyroid problems. We've already discussed how the symptoms of this disease tend to appear in people as they age, which is why it's hard to connect them to the condition. However, that doesn't make the disease any less problematic, so you need to try your best to prevent it effectively reducing your risk of cardiovascular problems that tend to come with it. Here are the things that you can do to prevent thyroid disease: If you are a smoker, quit smoking. This not only lowers your chance of developing thyroid disease but also many other health issues, including those of a cardiovascular nature. Eat less soy. The popular ingredient is not necessarily unhealthy, but it is controversial, especially when we consider thyroid health. Ask for a thyroid collar during x-rays to protect your thyroid gland from radiation exposure. Consider selenium supplements. Selenium is a nutrient found in specific proteins, and it has a vital role in keeping your thyroid gland healthy. If your diet is healthy, you're already taking enough selenium. Still, you may consider increasing your selenium intake to boost your immune system and decrease symptoms of hypothyroidism. Visit your doctor regularly. Regular checkups are not only good for your overall health but also your thyroid health.

Thyroid Disease Treatment Options

As thyroid disease can manifest as hypothyroidism and hyperthyroidism, there are two distinct treatment options. Bear in mind that thyroid disease is a lifelong condition, but with careful management, a person can still lead a healthy and normal life. Hypothyroidism is more common, and may be treated with levothyroxine, a synthetic hormone that supplements your thyroid hormone production. Hyperthyroidism is rare and more challenging to treat. Careful drug therapy is needed to block excess hormone production, and radiotherapy may be required to disable the gland. Sometimes thyroid surgery is necessary.

Fig.8.9 Prevention and Treatments

10.Prediction:



Prediction Page

Age: 29

Sex: Female

on_thyroxine: YES

query_on_thyroxine: YES

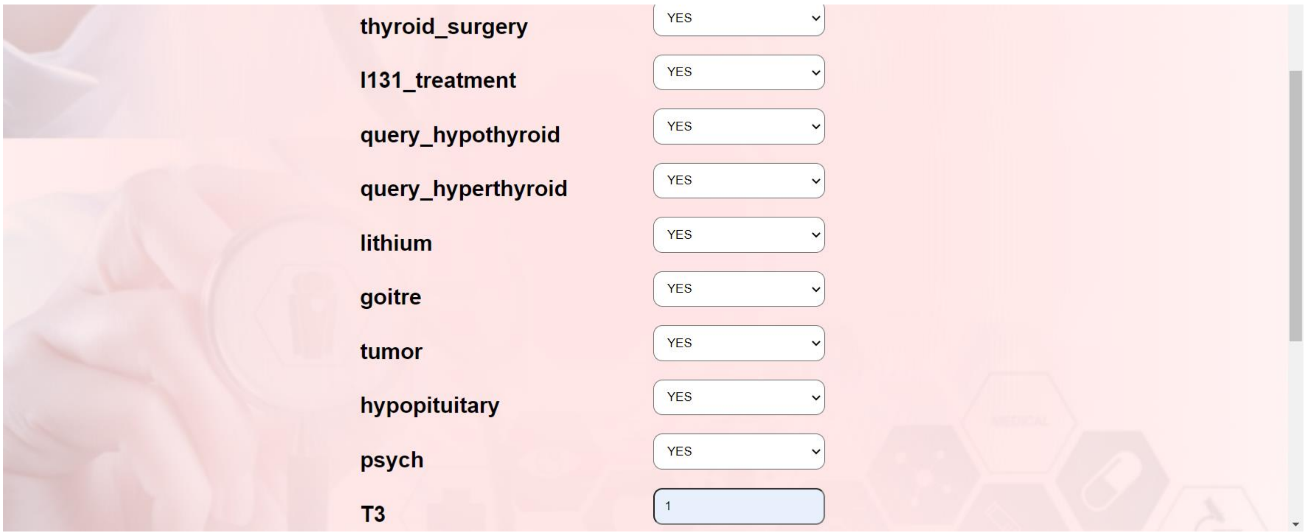
on_antithyroid_medication: YES

sick: YES

pregnant: YES

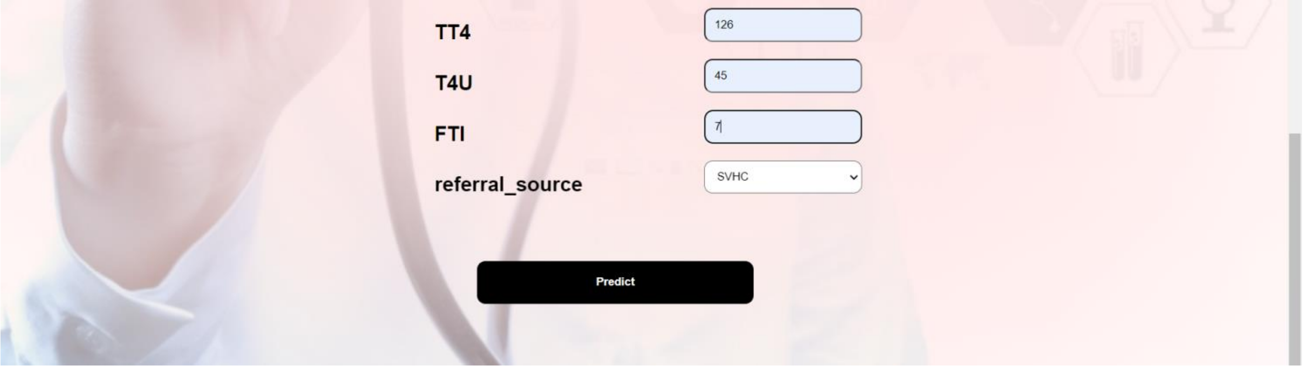
thyroid_surgery: YES

Fig.8.10.1-Prediction Page



thyroid_surgery	YES
I131_treatment	YES
query_hypothyroid	YES
query_hyperthyroid	YES
lithium	YES
goitre	YES
tumor	YES
hypopituitary	YES
psych	YES
T3	1

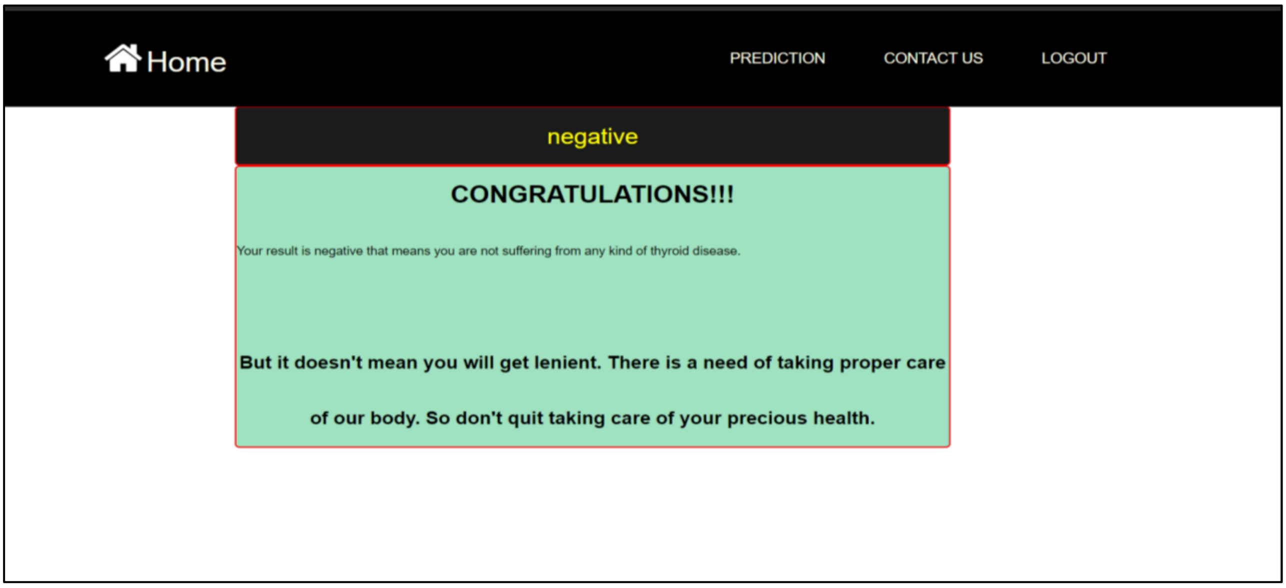
Fig.8.10.2-Prediction Page



TT4	126
T4U	45
FTI	7
referral_source	SVHC

Fig.8.10.3- Prediction

11.Result Page:



negative

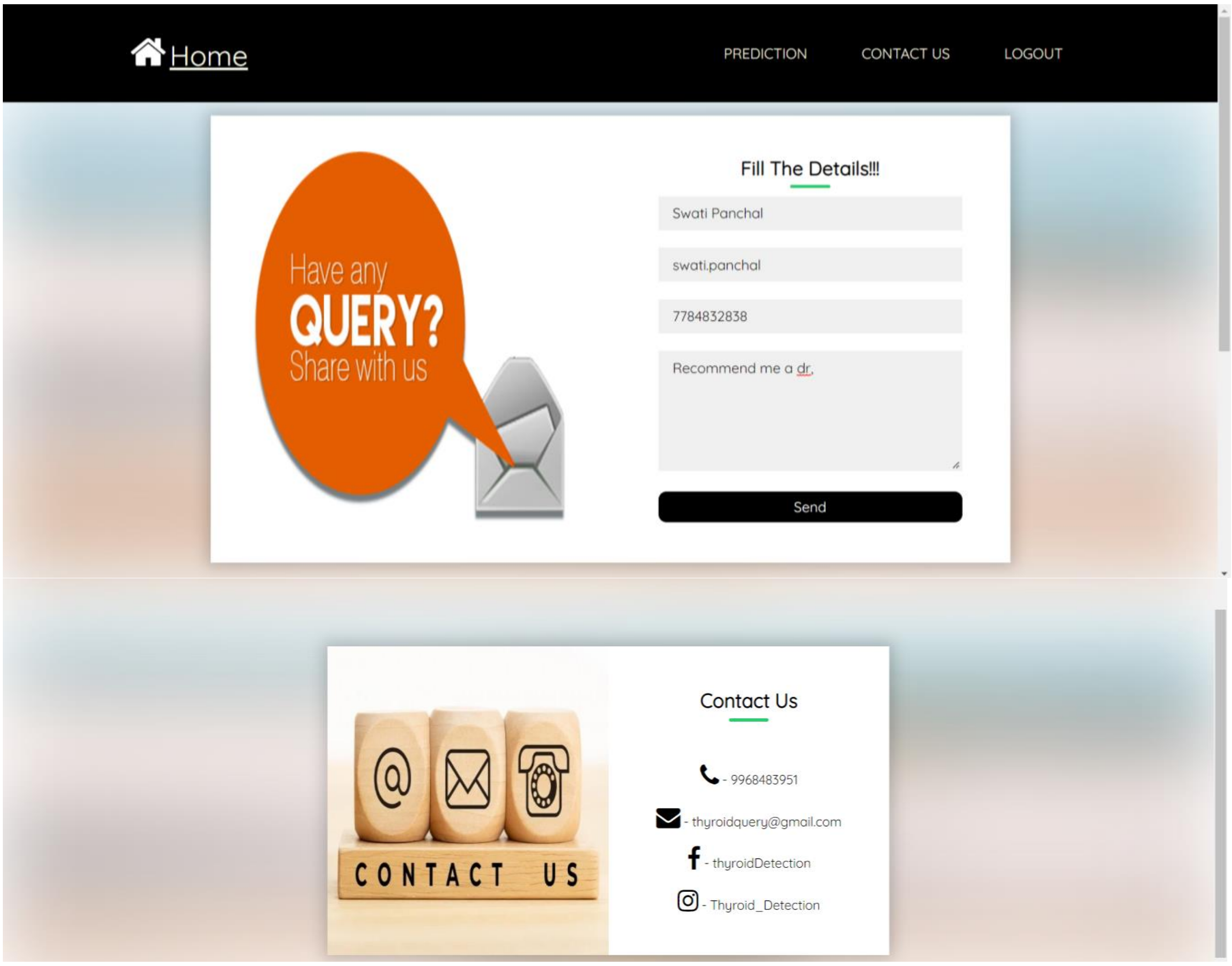
CONGRATULATIONS!!!

Your result is negative that means you are not suffering from any kind of thyroid disease.

But it doesn't mean you will get lenient. There is a need of taking proper care of our body. So don't quit taking care of your precious health.

Fig.8.11 Result Page

12.Contact Us page:



Fill The Details!!!

Swati Panchal

swati.panchal

7784832838

Recommend me a dr,

Contact Us

☎ - 9968483951

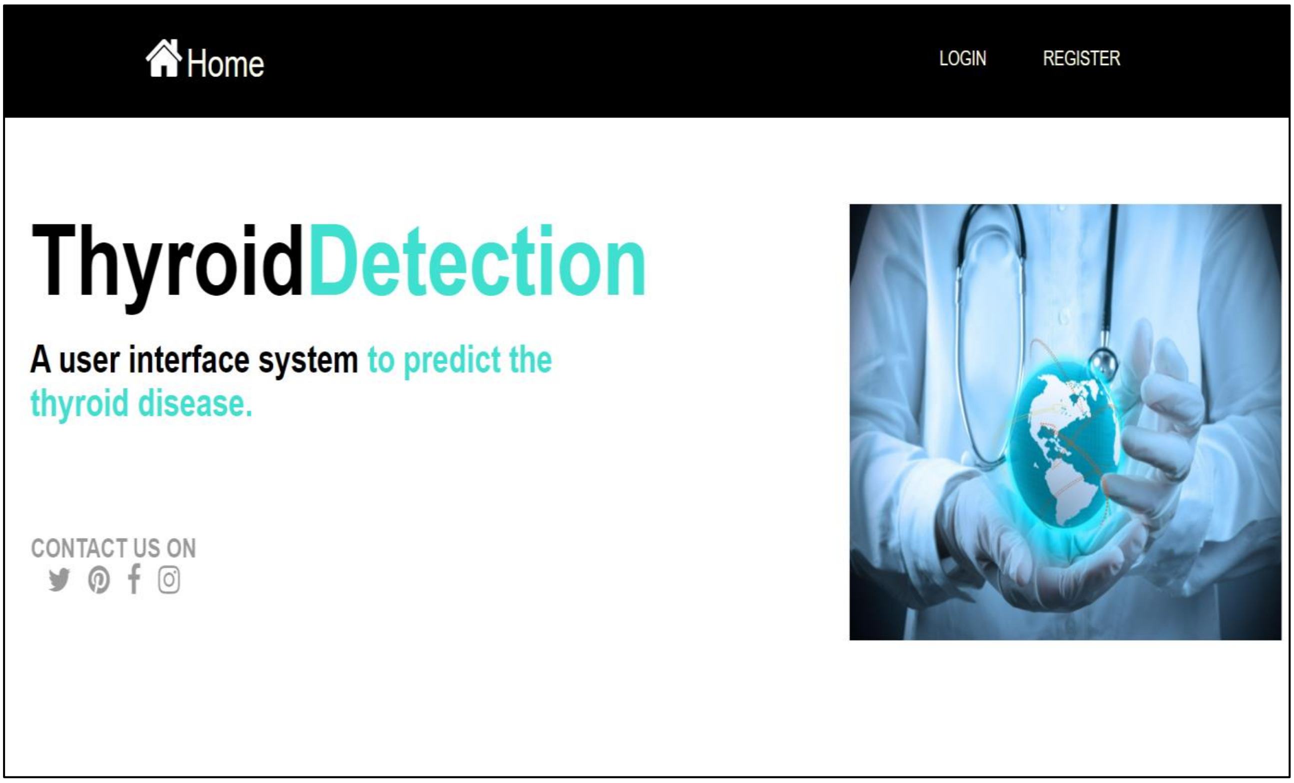
✉ - thyroidquery@gmail.com

f - thyroidDetection

- Thyroid_Detection

Fig.8.12Contact Us

13.After logout: Welcome Page



Home LOGIN REGISTER

ThyroidDetection

A user interface system to predict the thyroid disease.

CONTACT US ON

Fig.8.13 Log-Out: Welcome Page

8.3 Bibliography

- https://www.analyticsvidhya.com/blog/2021/06/confusion-matrix-for-multi-class-classification/

- https://htmlcolorcodes.com/

- https://www.kaggle.com/datasets/yasserhessein/thyroid-disease-data-set

- https://ipage.com/blog/website-template-customization/