THYROID DISEASE CLASSIFICATION

USING MACHINE LEARNING

- **TEAM ID:**
  DDEC882B608CD779CC4425EFFDF5325B

TEAM MEMBERS:

MUTHUSELVEN M(20UCS4296)

NANDHAKUMAR K(20UCS4297)

PALANISAMY D(20UCS4298)

PREM B(20UCS4299)

SUBMITTED FOR THE PROJECT UNDER THE NAAN MUDHALVAN – SMARTINTERNZ PROGRAM

OF CONTENTS: -

INTRODUCTION

OVERVIEW

Thyroid diseases, such as hypothyroidism and hyperthyroidism, are common endocrine disorders that affect the function of the thyroid gland. These diseases can have a significant impact on a patient's health and quality of life. Early and accurate diagnosis of thyroid diseases is important for effective treatment.

In recent years, machine learning techniques have been applied to the classification of thyroid diseases. The goal of these studies is to develop models that can accurately diagnose thyroid diseases based on clinical and laboratory data.

PURPOSE

There are several machine learning algorithms that have been used for thyroid disease classification, including decision trees, random forests, k-nearest neighbors (KNN), support vector machines (SVM), artificial neural networks (ANN), and deep learning algorithms such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

The input data for these models can include clinical features, such as age, gender, and symptoms, as well as laboratory test results, such as thyroid-stimulating hormone (TSH) levels and levels of thyroxine (T4) and triiodothyronine (T3).

The performance of these models is usually evaluated using metrics such as accuracy, precision, recall, and F1 score. In general, deep learning algorithms have shown better performance than other machine learning algorithms in thyroid disease classification tasks.

LITERATURE SURVEY

EXISTING PROBLEM

The current existing system includes:

Clinical examination: This involves a physical examination of the neck to check for any visible signs of thyroid enlargement or nodules.

Blood tests: Blood tests are used to measure the levels of hormones produced by the thyroid gland and to check for antibodies that may indicate autoimmune diseases such as Hashimoto's thyroiditis.

Ultrasound: An ultrasound scan can provide images of the thyroid gland and help to identify any nodules or other abnormalities.

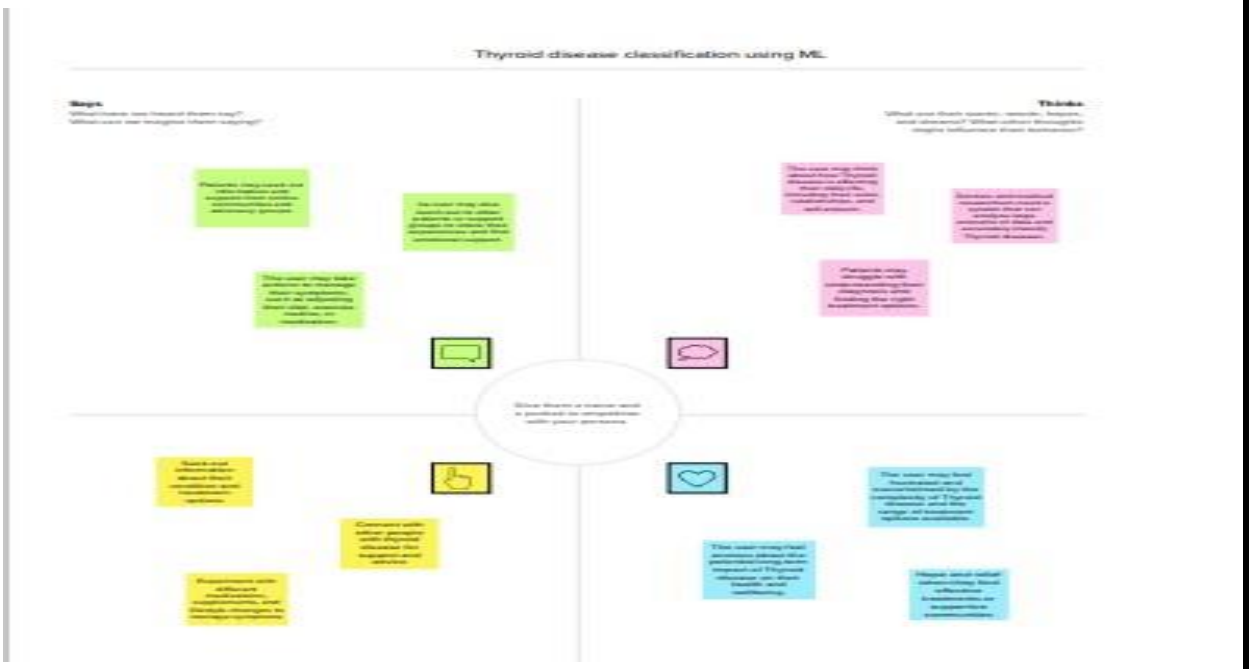
Fine needle aspiration biopsy (FNAB): This is a procedure in which a small sample of tissue is taken from a thyroid nodule using a fine needle, which is then examined under a microscope to check for cancer.

PROPOSED SYSTEM

The proposed system is by using Artificial Intelligence (AI) and Machine Learning (ML) In recent years, there have been several studies exploring the use of AI and ML algorithms for predicting thyroid diseases. These systems are trained on large datasets of patient data and use various features such as demographic information, blood test results, and ultrasound images to make prediction.

IDEATION & PROPOSED SYSTEM

Empathy Map Canvas



Ideation & Brainstorming

Project Flow:

- The user interacts with the UI to enter the input.
- Entered input is analysed by the model which is integrated.
- Once the model analyses the input the prediction is showcased on the UI

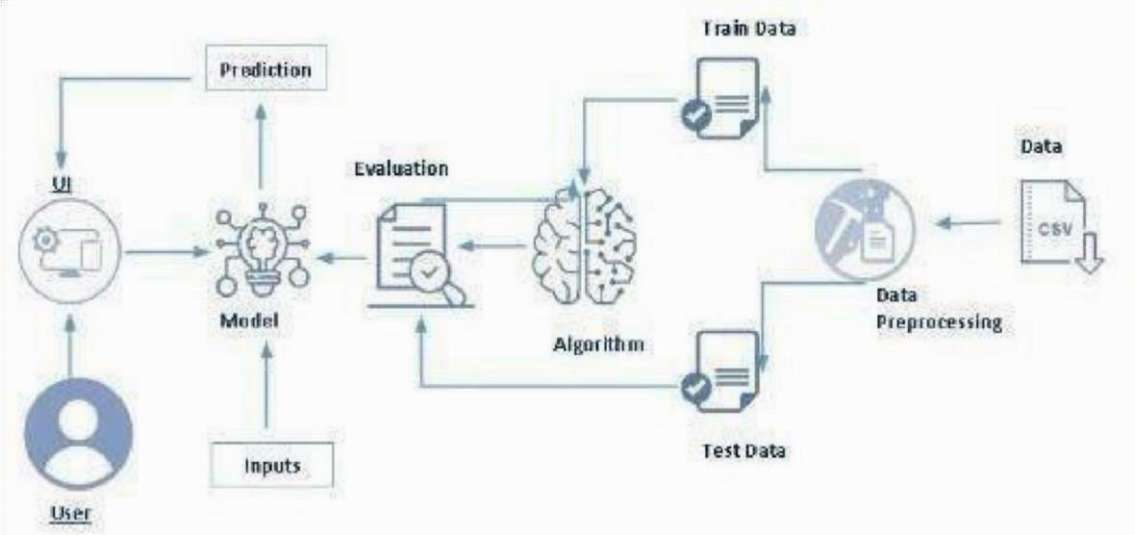
To accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding ○ Specify the business problem ○ Business requirements ○ Literature Survey ○ Social or Business Impact.
- Data Collection & Preparation ○ Collect the dataset ○ Data Preparation
- Exploratory Data Analysis ○ Descriptive statistical ○ Visual Analysis
- Model Building ○ Training the model in multiple algorithms ○ Testing the model
- Performance Testing & Hyperparameter Tuning ○ Testing model with multiple evaluation metrics ○ Comparing model accuracy before & after applying hyperparameter tuning
- Model Deployment ○ Save the best model ○ Integrate with Web Framework

- Project Demonstration & Documentation ○ Record explanation Video for project end to end solution ○ Project Documentation-Step by step project development procedur

THEORETICAL ANALYSIS

BLOCK DIAGRAM



HARDWARE AND SOFTWARE DESIGNING

Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. It was created by Guido van Rossum, and first released on February

20, 1991. 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.

Anaconda Navigator

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS. Conda is an open-source, cross platform, package management system. Anaconda comes with.

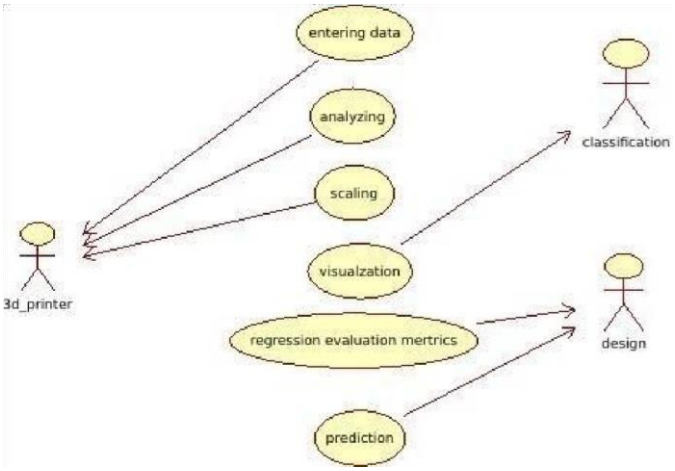
Flask

Web frame work used for building. It is a web application framework written in python which will be running in local browser with a user interface. In this application, whenever the user interacts with UI and selects emoji, it will suggest the best and top movies of that genre to the use.

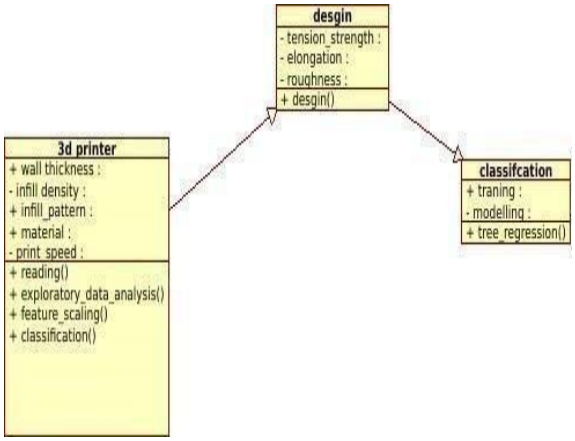
Hardware Requirements:

o Operating system: window 7 and above with 64bit o Processor Type -Intel Core i3-3220 o RAM: 4Gb and above o Hard disk: min 100GB

1. FLOWCHART



USE CASE DIAGRAM



RESULTS

Performance metrics

MODEL: Random Forest Classifier Model

There is no big variation in the training and testing accuracy. Therefore, the Random Forest Classifier model is not overfit or underfit.

```
y_pred = rfr1.predict(x_test_os)
```

```
print(classification_report(y_test_os,y_pred))
```

```
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       122
           1       0.76      0.90      0.83       122
           2       0.91      0.98      0.94       122
           3       0.78      0.83      0.80       122
           4       0.46      0.92      0.62       122
           5       0.75      0.70      0.73       122
           6       0.63      0.48      0.54       122

    accuracy                           0.69       854
   macro avg       0.61      0.69      0.64       854
weighted avg       0.61      0.69      0.64       854
```

```
train_score = accuracy_score(y_os, rfr1.predict(x_os))
train_score
```

```
1.0
```

MODEL: XGB Classifier Model

There is a variation in the training and testing accuracy. The XGB Classifier model is overfit. So, it is not chosen for best results.

```
print(classification_report(y_test_os,y_pred))
```

```
              precision    recall  f1-score   support

           0       0.70      0.13      0.22       122
           1       0.75      0.93      0.84       122
           2       0.95      0.99      0.97       122
           3       0.76      0.77      0.77       122
           4       0.48      0.85      0.61       122
           5       0.79      0.71      0.75       122
           6       0.62      0.52      0.57       122

    accuracy                           0.70       854
   macro avg       0.72      0.70      0.67       854
weighted avg       0.72      0.70      0.67       854
```

```
accuracy_score(y_test_os,y_pred)
```

```
0.7014051522248244
```

MODEL : ANN MODEL

There is a variation in the training and testing accuracy. The ANN model is overfit. So, it is not chosen for best results.

```
In [73]: model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
```

```
In [75]: model.fit(x_bal,y_bal, validation_data=[x_test_bal, y_test_bal], epochs=15)
```

```
Epoch 1/15
104/104 [==============================] - 9s 15ms/step - loss: -18416.0605 - accuracy: 0.1429 - val_loss: -142105.5156 - val
_accuracy: 0.1429
Epoch 2/15
104/104 [==============================] - 1s 8ms/step - loss: -2626274.5000 - accuracy: 0.1429 - val_loss: -10219054.0000 -
val_accuracy: 0.1429
Epoch 3/15
104/104 [==============================] - 1s 9ms/step - loss: -42823204.0000 - accuracy: 0.1429 - val_loss: -113329736.0000
- val_accuracy: 0.1429
Epoch 4/15
104/104 [==============================] - 1s 9ms/step - loss: -277232128.0000 - accuracy: 0.1429 - val_loss: -582218880.0000
- val_accuracy: 0.1429
Epoch 5/15
104/104 [==============================] - 1s 8ms/step - loss: -1097882752.0000 - accuracy: 0.1429 - val_loss: -1989677696.00
00 - val_accuracy: 0.1429
Epoch 6/15
104/104 [==============================] - 1s 8ms/step - loss: -3208519680.0000 - accuracy: 0.1429 - val_loss: -5285069824.00
00 - val_accuracy: 0.1429
Epoch 7/15
```

MODEL : SVC MODEL

There is no big variation in the training and testing accuracy. Therefore, the SVC model is not overfit or underfit.

```
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score, classification_report

sv= SVC()
```

```
sv.fit(x_bal,y_bal)
```

```
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\sklearn\utils\validation.py:1111: DataConversionWarning: A column-vector y
was passed when a 1d array was expected. Please change the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, warn=True)
```

▾ SVC

SVC()

```
y_pred = sv.predict(x_test_bal)
```

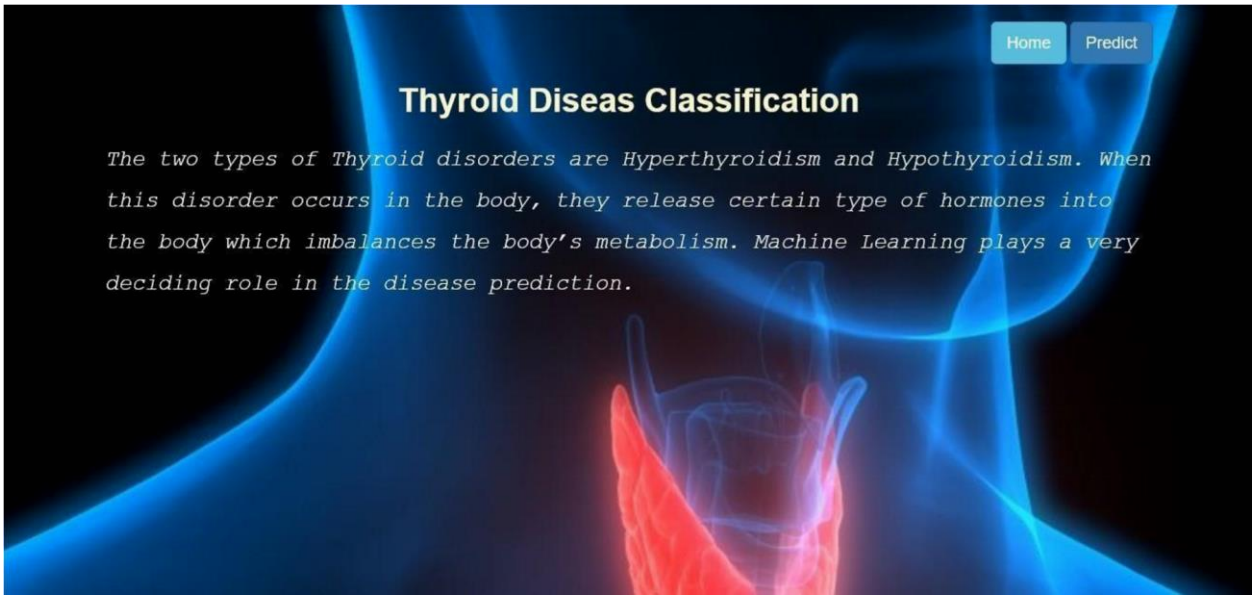
```
print(classification_report(y_test_bal,y_pred))
```

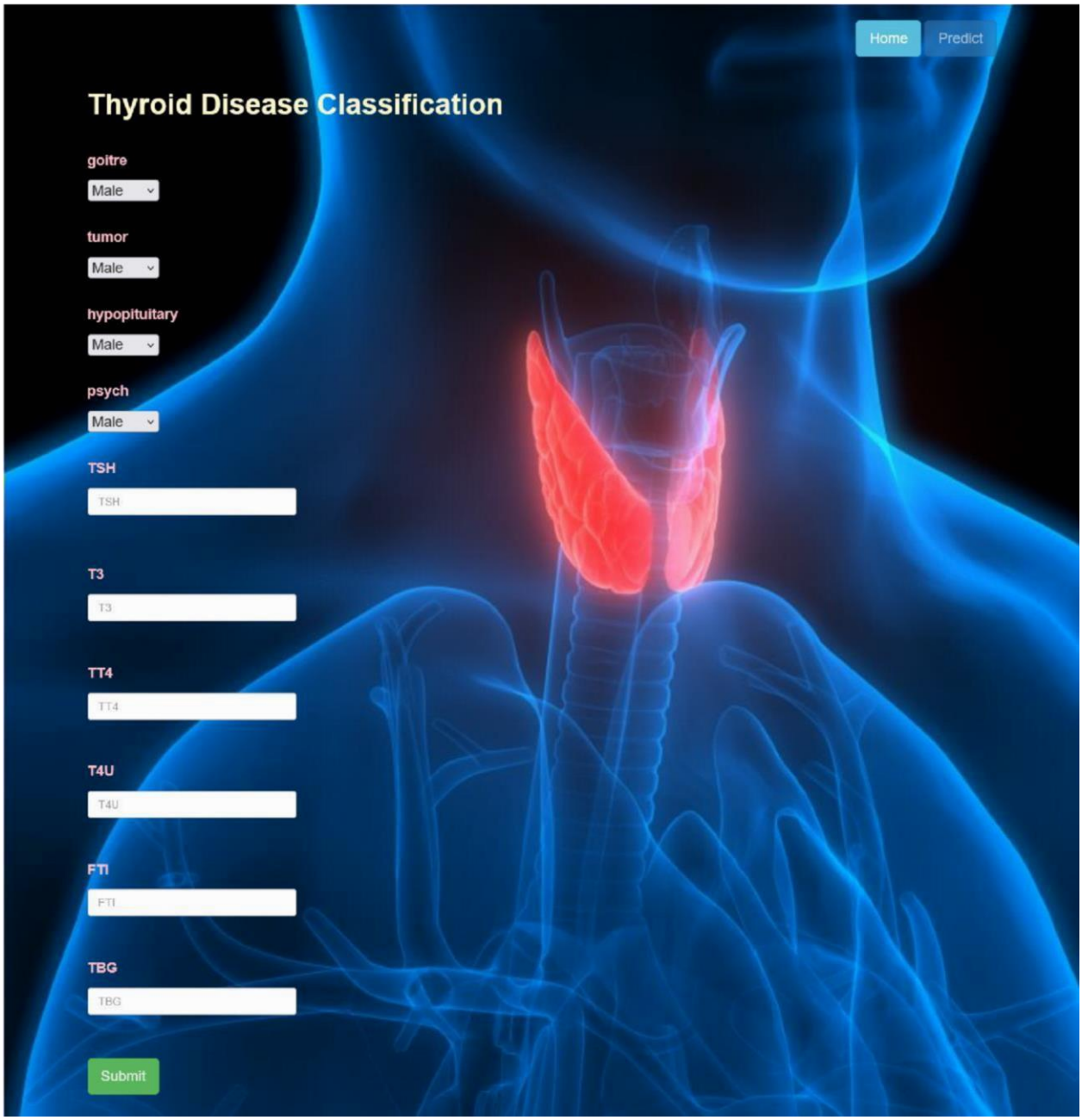
```
              precision    recall  f1-score   support

           0       0.70      0.85      0.77       122
           1       0.76      0.81      0.79       122
           2       0.88      0.93      0.90       122
           3       0.71      0.65      0.68       122
           4       0.71      0.63      0.67       122
           5       0.76      0.54      0.63       122
           6       0.49      0.57      0.52       122

    accuracy                           0.71       854
   macro avg       0.72      0.71      0.71       854
weighted avg       0.72      0.71      0.71       854
```

On comparing the four models built, based on the performance metrics it is clear that SVC model gives the highest performance. Hence, that model is chosen for deployment.

Output of the Project:

```
In [32]: runfile('C:/Users/SmartBridge-PC/
Downloads/Thyroid/app.py', wdir='C:/Users/
SmartBridge-PC/Downloads/Thyroid')
 * Serving Flask app "app" (lazy loading)
 * Environment: production
   WARNING: This is a development server. Do
not use it in a production deployment.
   Use a production WSGI server instead.
 * Debug mode: off
 * Running on http://127.0.0.1:5000/ (Press
CTRL+C to quit)
```




Home
Predict
Thyroid Disease Classification
goitre
Male
tumor
Male
hypopituitary
Male
psych
Male
TSH
TSH
T3
T3
TT4
TT4
T4U
T4U
FTI
FTI
TBG
TBG
Submit


Home
Predict
Thyroid Disease Classification
Based on the given input, it predicts Thyroid disease for your body condition is ['binding protein'].

ADVANTAGES & DISADVANTAGES

ADVANTAGES

- Easy to use
- Cost efficient
- Time efficient

DISADVANTAGE

1. Initial costs of printer
2. Post processing
3. Printing time
4. Special skill required for 3D models
5. Manufacturing Job Losses

Applications

3D printing has gone through a number of changes over the years. In the early days, 3D printing was time-consuming and costly, and not very practical for applications outside of industry. However, with the advent of today's more flexible and cost-effective 3D printing methods, there are areas where 3D printing has become a practical tool.

It is applicable in different sectors such as

- Engineering And Design
- Consumer products
- Manufacturing
- Education
- Aerospace
- Medical

- Movies / Theatres
- Architectures

CONCLUSION:

3D printing technology could revolutionize and re-shape the world. Advance in 3D technology can significantly change and improve the way we manufacture products goods worldwide.

If the last industrial revolution brought us mass production and the advent of economics of scale – the digital 3D printing revolution could bring mass manufacturing back a full of circle – to an era of mass personalization, and return to individual craftsmanship.

FUTURE SCOPE

Future applications for 3D printing might include creating open-source scientific equipment to create opensource labs

Science-based applications like reconstructing fossils in palaeontology . Replicating ancient and priceless artifacts in archaeology

Reconstructing bones and body parts in forensic pathology. The technology currently being researched for building construction.

APPENDIX

Source code:

Flask file:

```
from flask import Flask, render_template, request
import numpy as np import pickle import pandas as pd


model = pickle.load(open(r"C:\Users\SmartBridge-
pc\Downloads\thyroidclassification\Training\thyroid_1_model.pkl",
'rb')) le = pickle.load(open("label_encoder.pkl", 'rb'))


app = Flask(__name__)
```

```
@app.route("/") def about():
return render_template('home.html')



@app.route("/predict") def home1():
return render_template('predict.html')



@app.route("/pred", methods=['POST', 'GET']) def
predict():     x = [[float(x) for x in
request.form.values()]]
         print(x)
col
=
['goitre','tumor','hypopituitary','psych','TSH','T3','TT4','T4U','FTI','TBG']
x = pd.DataFrame(x, columns=col)


    #print(x.shape)


    print(x)      pred =
model.predict(x)     pred =
le.inverse_transform(pred)
print(pred[0])
    return render_template('submit.html', prediction_text=str(pred))
if __name__ == "__main__":
app.run(debug=False)
```

HTML JS and CSS files:

home.html

```
<!doctype html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1">
    <meta http-equiv="X-UA-Compatible" content="ie=edge">
    <title>Home</title>     <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/css/bootstrap.min.css">
    <style>         body         {          background-image:
url("https://www.entincayman.com/wpcontent/uploads/Thyroid_Nodules_CausesSymptoms.jpg");
                  background-
size: cover;          }
h3.big          {          line-
height:
1.8;
        }
    </style>
</head>
<body>
    <br>
    <div class="container">

        <div class="row">
            <div class="col-md-12 bg-light text-right">
                <a href="/home" class="btn btn-info btn-lg">Home</a>
                <a href="/predict" class="btn btn-primary btn-lg">Predict</a>


            </div>
        </div>

        <center>
```

<h1>Thyroid Diseas Classification</h1>

<h3 class="big">

The two types of Thyroid disorders are Hyperthyroidism and Hypothyroidism.
When this disorder occurs in the body,
they release certain type of hormones into the body which imbalances the body’s metabolism.
Machine Learning plays a very deciding role in the disease prediction.

</h3>

<script src="https://ajax.googleapis.com/ajax/libs/jquery/3.5.1/jquery.min.js"><
/scri pt> <script
src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/js/bootstrap.min.js">< /s
cript> </body>
</html>

predict.html

<!DOCTYPE html>

```
<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>Predict</title>     <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/css/bootstrap.min.css">
<style>         body
{          background-image:
url("https://www.entincayman.com/wpcontent/uploads/Thyroid_Nodules_CausesSymptoms.jpg");
background-size: cover;
        }         h3.big         {          line-height:
1.8;
        }
    </style>
</head>
```



```
<body>
    <br>
    <div class="container">

        <div class="row">
            <div class="col-md-12 bg-light text-right">
                <a href="/home" class="btn btn-info btn-lg">Home</a>
                <a href="/predict" class="btn btn-primary disabled btnlg">Predict</a>
            </div>
        </div>

        <br>
        <font color="#FFFDD0"><h1><strong>Thyroid Disease
Classification</strong></h1><br></font>
        <h4>
        <form action="/pred", method="post">
```

```
<div class="form-group mb-3">
              <div class="input-group-prepend">
                  <p style="color: pink;"><label class="input-group-text"
for="goitre">goitre</label></p>
              </div>
                <select class="custom-select" name="goitre" id="goitre">
                  <option value="0">Male</option>
<option value="1">Female</option>
                </select>
              </div><br>


              <div class="form-group mb-3">
              <div class="input-group-prepend">
                  <p style="color: pink;"><label class="input-group-text"
for="tumor">tumor</label></p>
              </div>
                <select class="custom-select" name="tumor" id="tumor">
                  <option value="0">Male</option>
<option value="1">Female</option>
                </select>
              </div><br>


              <div class="form-group mb-3">
```



```
              <div class="input-group-prepend">                    <p style="color:
pink;"><label class="input-group-text"
for="hypopituitary">hypopituitary</label></p>
              </div>                  <select class="custom-select"
name="hypopituitary" id="hypopituitary">
                  <option value="0">Male</option>
```

```
<option value="1">Female</option>
              </select>
            </div><br>


            <div class="form-group mb-3">
            <div class="input-group-prepend">
                <p style="color: pink;"><label class="input-group-text"
for="psych">psych</label></p>
            </div>
              <select class="custom-select" name="psych" id="psych">
                <option value="0">Male</option>
<option value="1">Female</option>
              </select>
            </div><br>


            <div class="form-group row">
                <div class="col-md-3">
                    <p style="color: pink;"><label for="TSH">TSH</label></p>
<p style="color: #ff0000;"><input type="text" name="TSH"  id="TSH" class="form-control"
placeholder="TSH"
required="required"/><br><br></p>
                </div> </div>
```

```
            <div class="form-group row">
                <div class="col-md-3">
                    <p style="color: pink;"><label for="T3">T3</label></p>
                    <input  type="text"       name="T3"         id="T3"
class="formcontrol" placeholder="T3" required="required"/><br><br>
                </div>
            </div>
             <div class="form-group row">
                 <div class="col-md-3">
                     <p style="color: pink;"><label for="TT4">TT4</label></p>
                     <input type="text"       name="TT4"        id="TT4"
class="formcontrol" placeholder="TT4" required="required"/><br><br>
                 </div>
             </div>
              <div class="form-group row">
                  <div class="col-md-3">
                      <p style="color: pink;"><label
for="T4U">T4U</label></p>
                      <input      type="text"       name="T4U"
      id="T4U" class="formcontrol" placeholder="T4U"
required="required"/><br><br>
                  </div>
              </div>
              <div class="form-group row">
                  <div class="col-md-3">
                      <p style="color: pink;"><label
for="FTI">FTI</label></p>
                      <input      type="text"       name="FTI"
      id="FTI" class="formcontrol" placeholder="FTI"
required="required"/><br><br>
                  </div>
              </div>
              <div class="form-group row">
                  <div class="col-md-3">
                      <p style="color: pink;"><label
for="TBG">TBG</label></p>
                      <input      type="text"       name="TBG"
      id="TBG" class="formcontrol" placeholder="TBG"
required="required"/><br><br>
```

```
                </div>
            </div>

        <button    type="submit"    class="btn  btn-success btnlg">Submit</button>



    </form>
    <br>
    </h4>
</div>


    <script
src="https://ajax.googleapis.com/ajax/libs/jquery/3.5.1/jquery.min.js"><
/scri pt>    <script
src="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/js/bootstrap.min.js"> <
/s cript> </body>
</html>
```

submit.html

```
<!DOCTYPE html>

<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>Output</title>    <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.4.1/css/bootstrap.min.css">
    <style>        body {         background-image:
url("https://www.entincayman.com/wpcontent/uploads/Thyroid_Nodules_CausesSymptoms.jpg");
background-size: cover;
        }        h3.big        {         line-
height: 1.8;
        }
    </style>
```

```
</head>
<body>
    <br>
        <div class="container">

            <div class="row">
                <div class="col-md-12 bg-light text-right">
```



```
                    <a href="/home" class="btn btn-info btn-lg">Home</a>
<a href="/predict" class="btn btn-primary btnlg">Predict</a>
                </div>
            </div>
            <br>
            <font color="#FFFDD0"><h1><strong>Thyroid Disease
Classification</strong></h1></font><br>
            <p><font face = "courier" color = "#FFFFFF">
             <h3>Based on the given input, it predicts Thyroid disease for your body condi
is {{prediction_text}}.             </h3>
            </font></p>


        </div>

</body>
</html
```