NAAN MUTHALVAN PROJECT REPORT

On

“Thyroid disease classification using Machine Learning”

Bachelor of Science

In

Computer Science

Submitted by

Team Leader

Shenbagam.G-20201061506139

Team Members

Sutha.K-20201061506145

Kaviya Ananthi.S-20201061506112

Anitha.P-20201061506102

Mahadevi.V-20201061506114

Kamarajar Government Arts College,

Surandai-627859,Tenkasi District.

April-2023

Table of Contents

* INTRODUCTION

1.Overview

2.Purpose

* Problem definition & Design Thinking

1.Empathy Map

2.Ideation & Brainstorming map screenshort

* RESULT

Final finding (output) of the project along with screenshort.

* ADVANTAGES & DISADVANTAGES
* APPLICATION
* CONCLUSION
* FUTURE SCOPE
* APPENDIX

PROJECT REPORT TEMPLATE

INTRODUCTION

* OVERVIEW

Thyroid disease is a general term for a medical condition that keeps your thyroid from making the right amount of hormones. Your thyroid typically makes hormones that keep your body functionally normally.

When the thyroid hormone, your body uses energy too quickly. This is called hyperthyroidism.

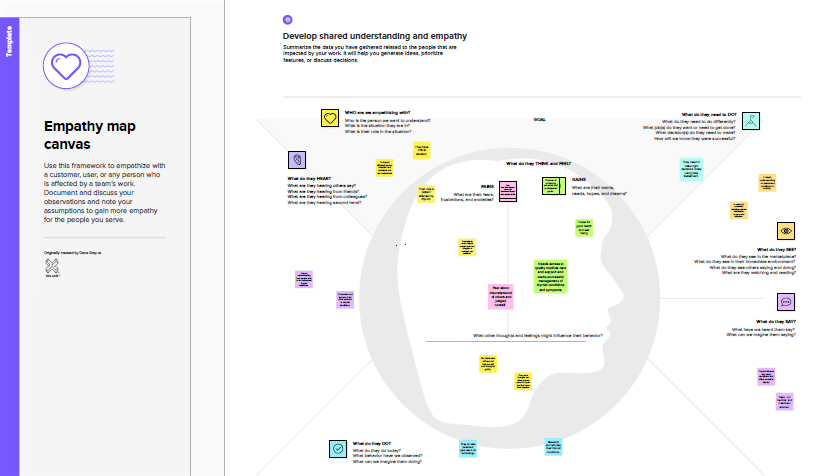
* PURPOSE

It makes hormones that control the way the body uses energy. These hormones affect nearly every organ in your body and control many of your body’s most important function. For example, they affect your breathing, heart rate, weight, digestion, and moods.

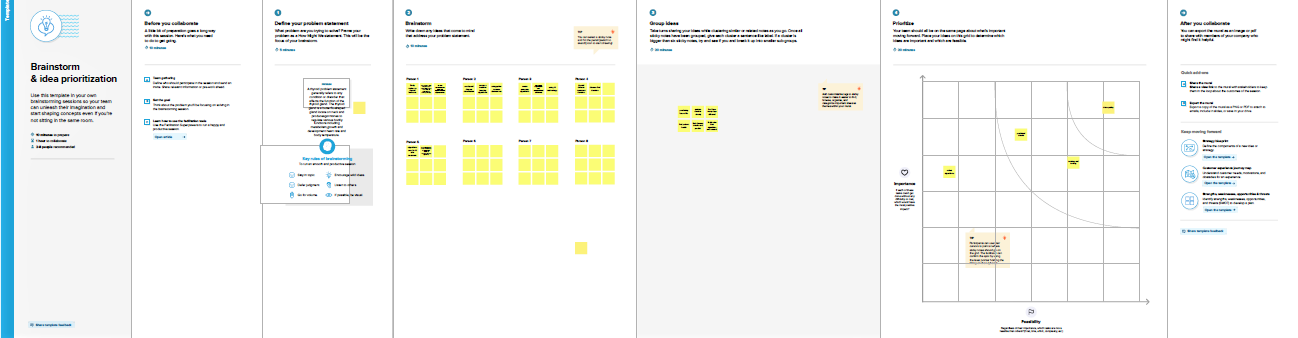
The thyroid gland is a vital hormone gland: It play a major role in the metabolism, growth and development of the human body. It helps to regulate many body functions by constantly releasing a steady amount of thyroid hormones into the bloodstream.

PROBLEM DEFINITION & DESIGN THINKING

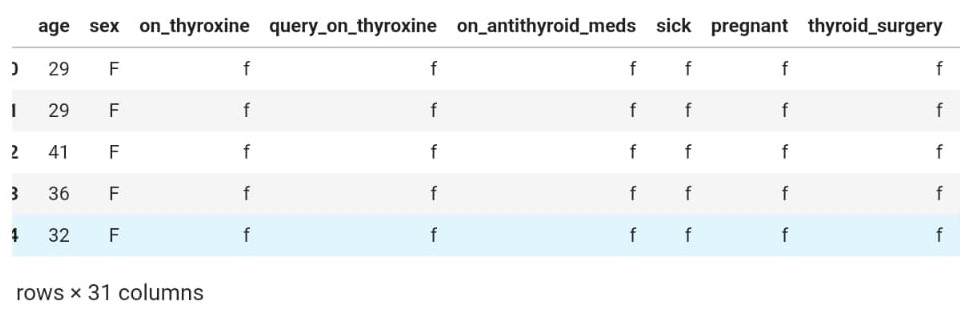
* EMPATHY MAP

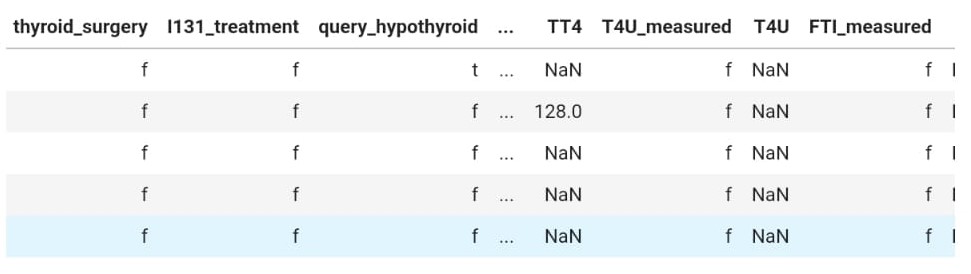


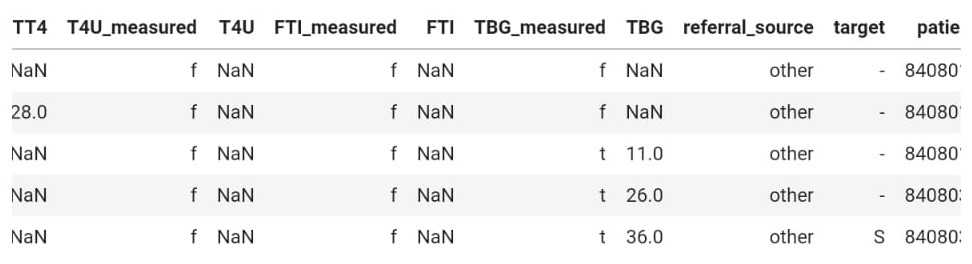
* IDEATION & BRAINSTORMING MAP

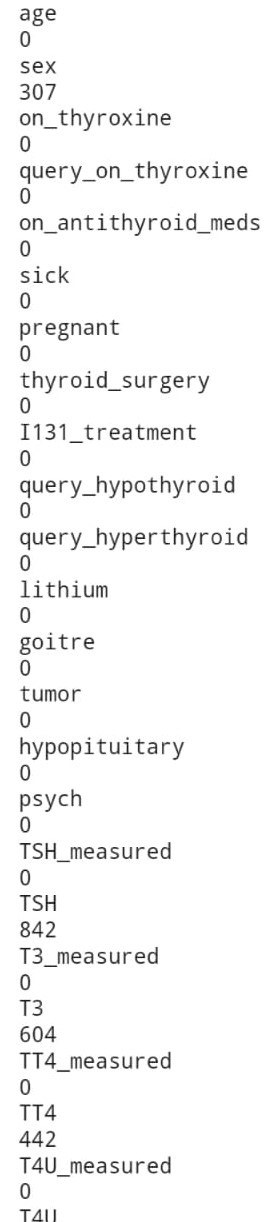


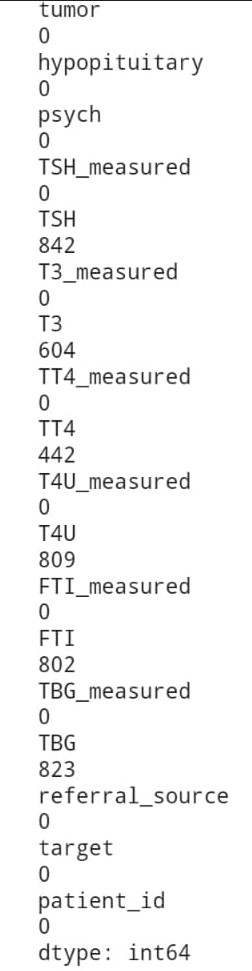
RESULT

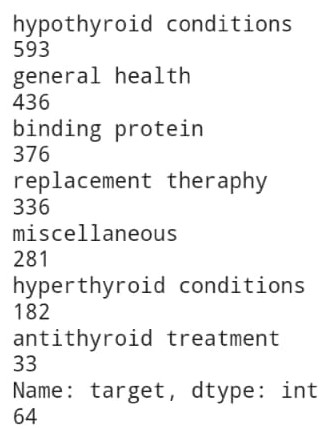


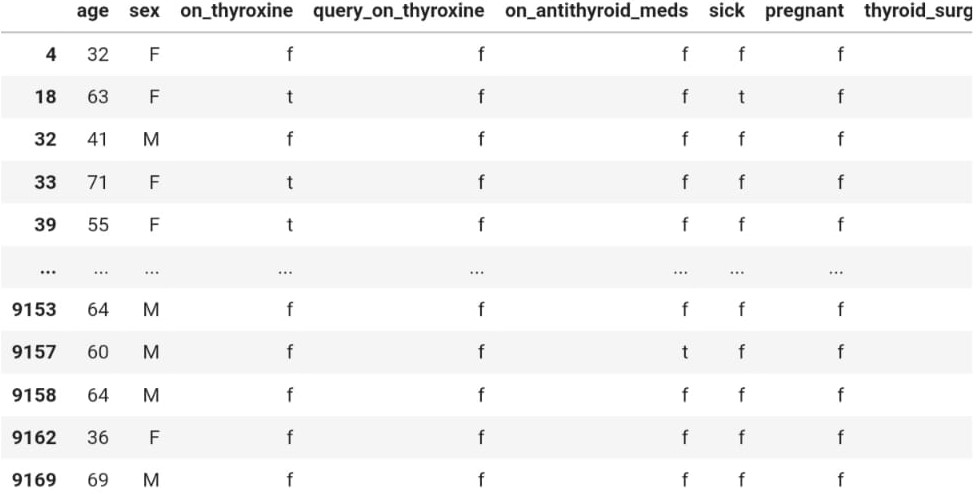


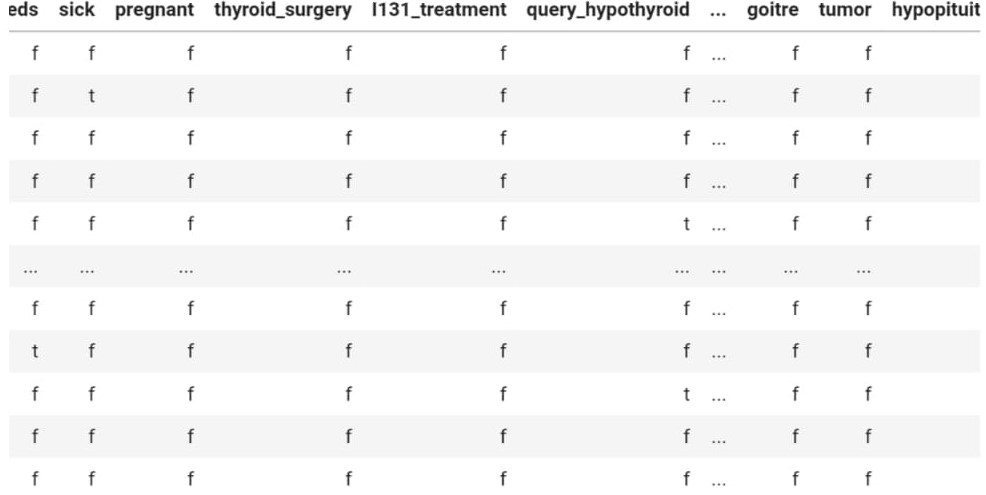


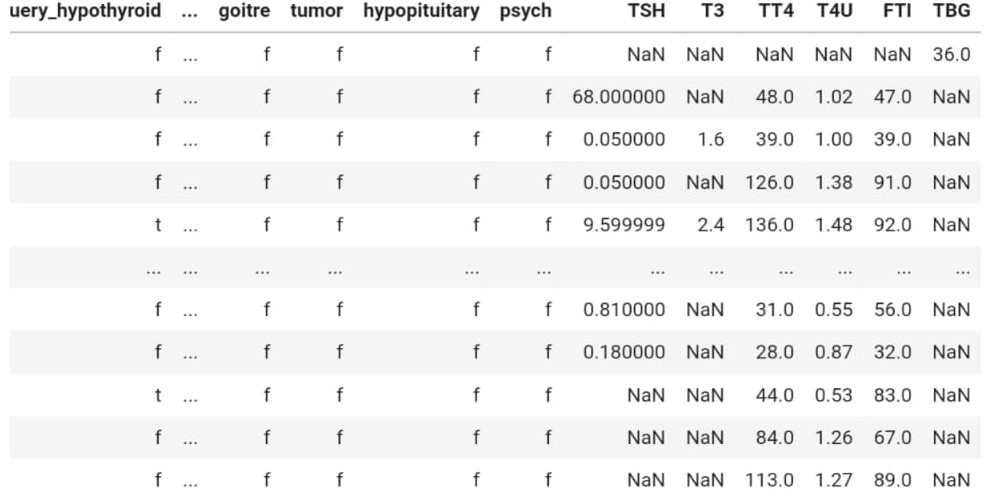


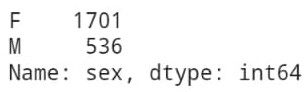


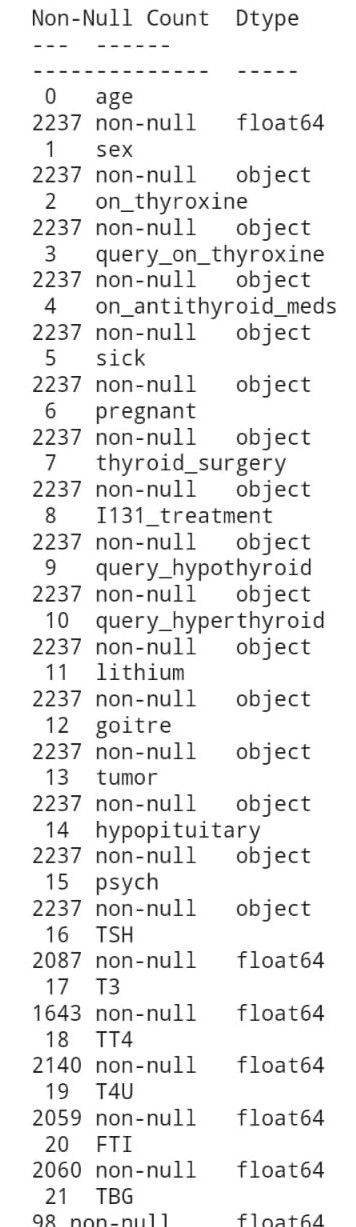


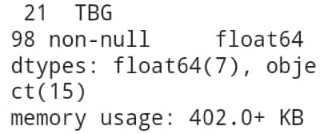


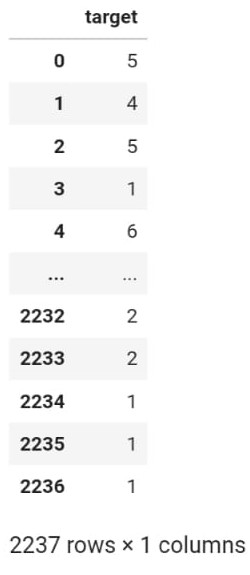


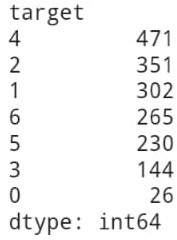


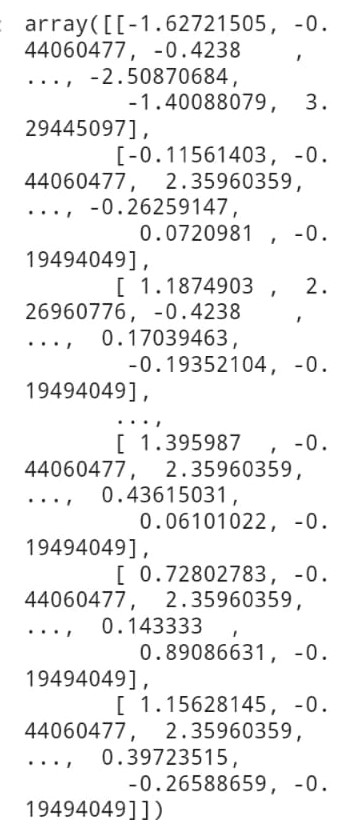


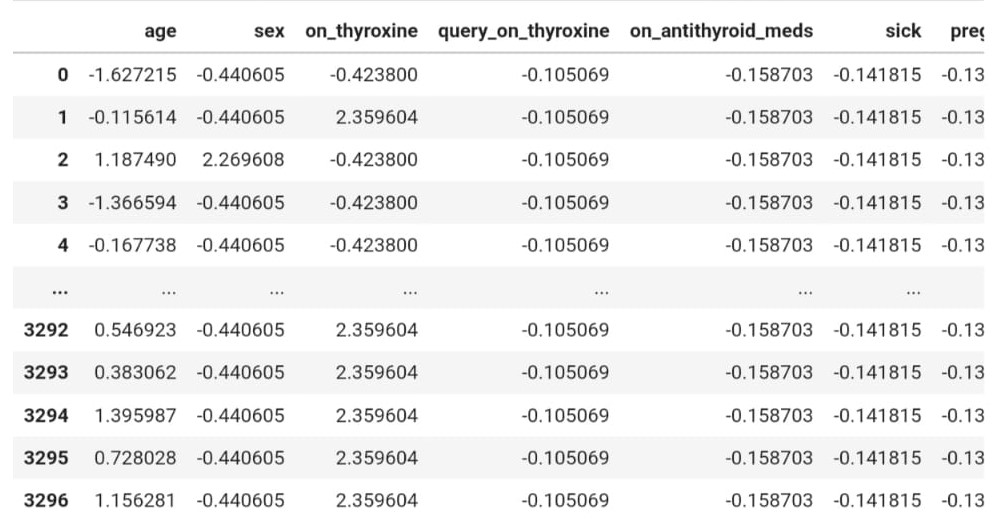


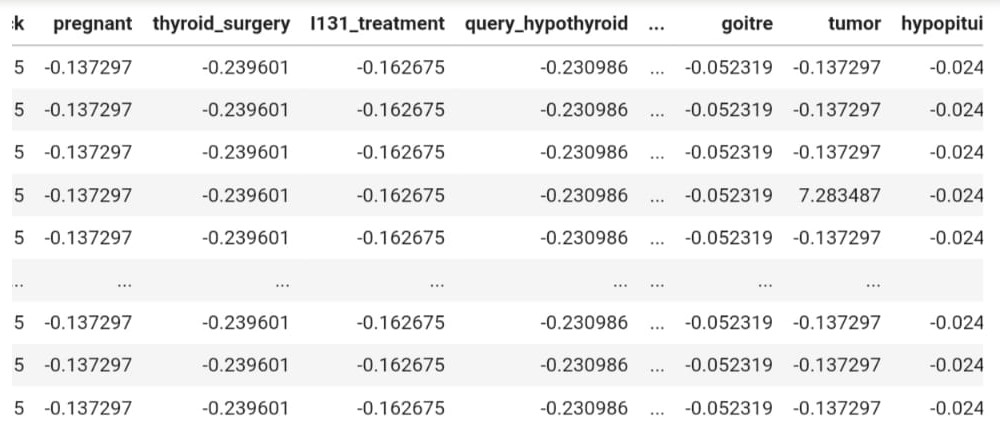




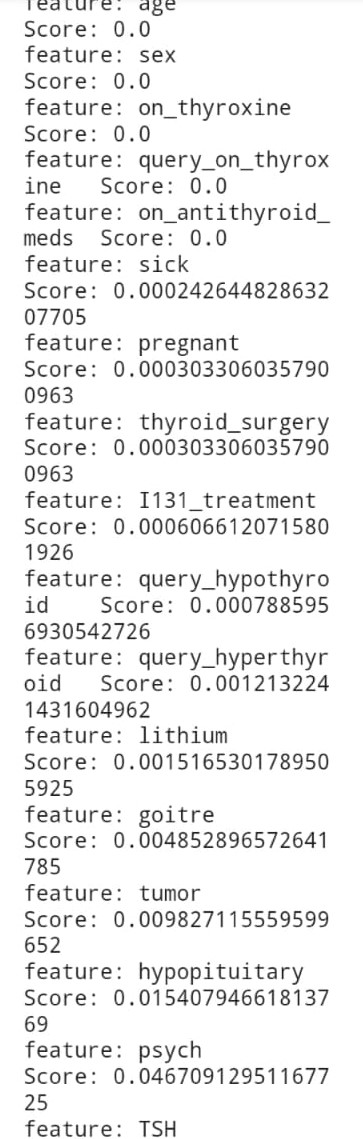


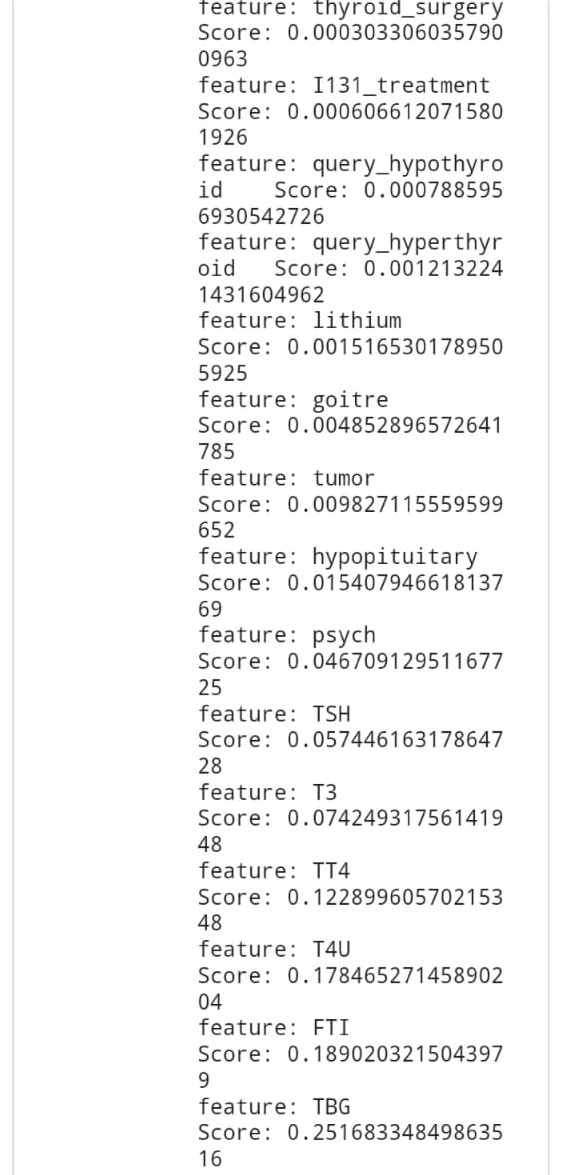


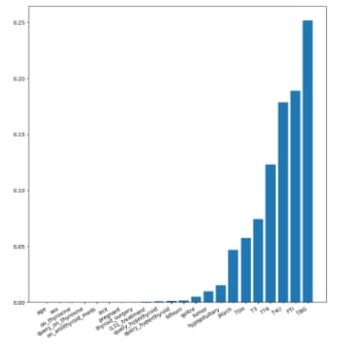


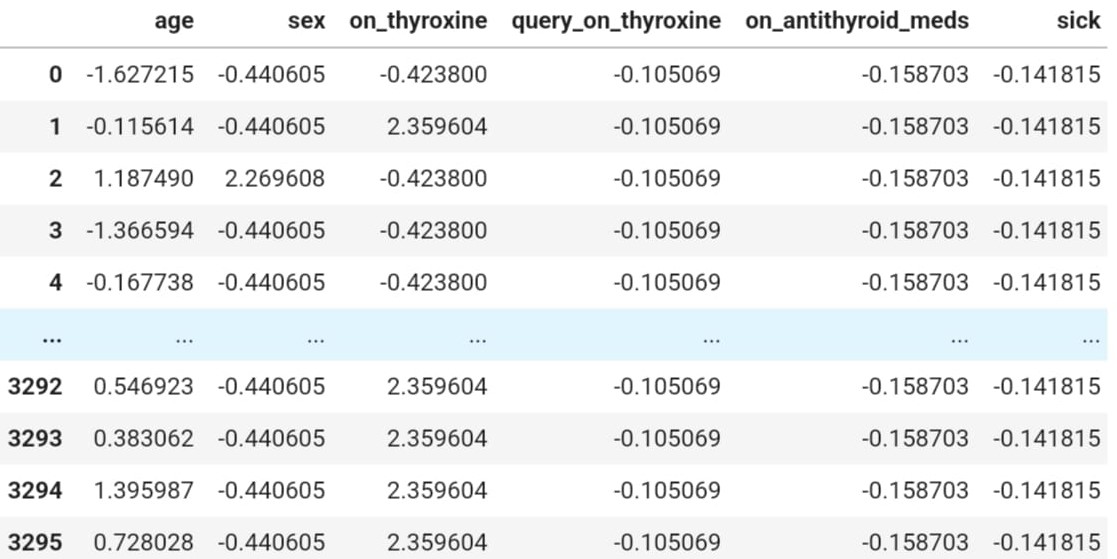




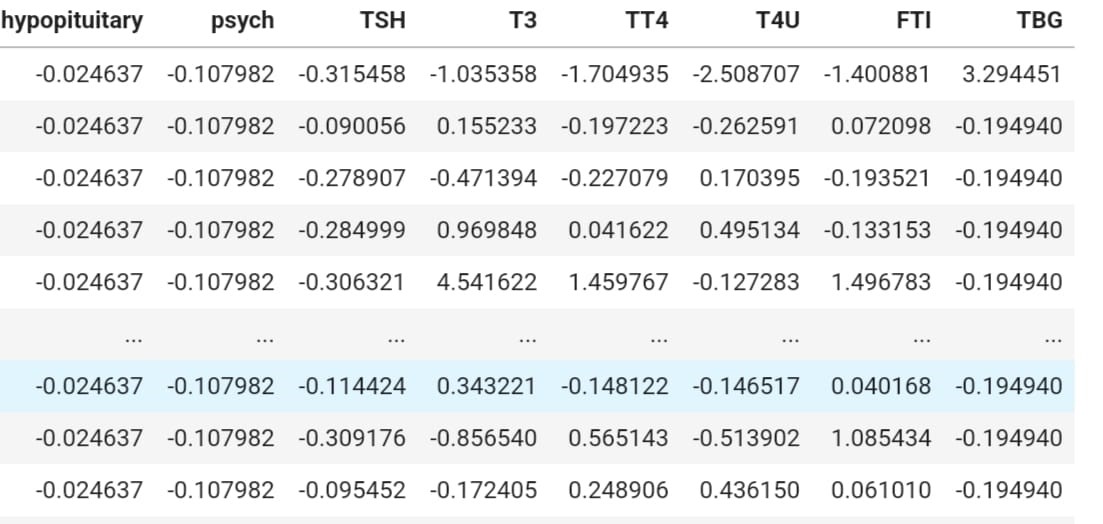




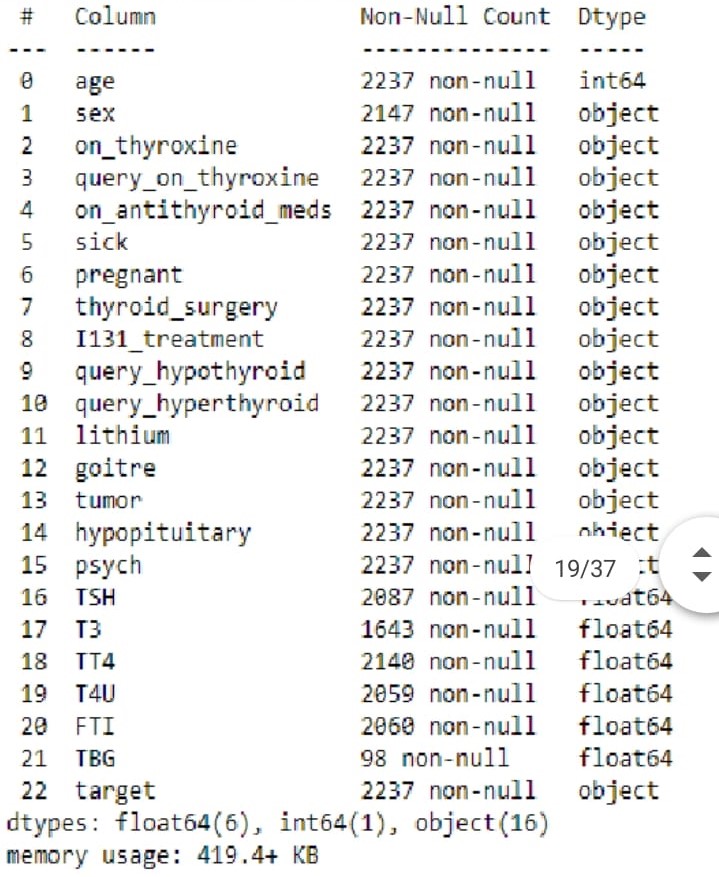


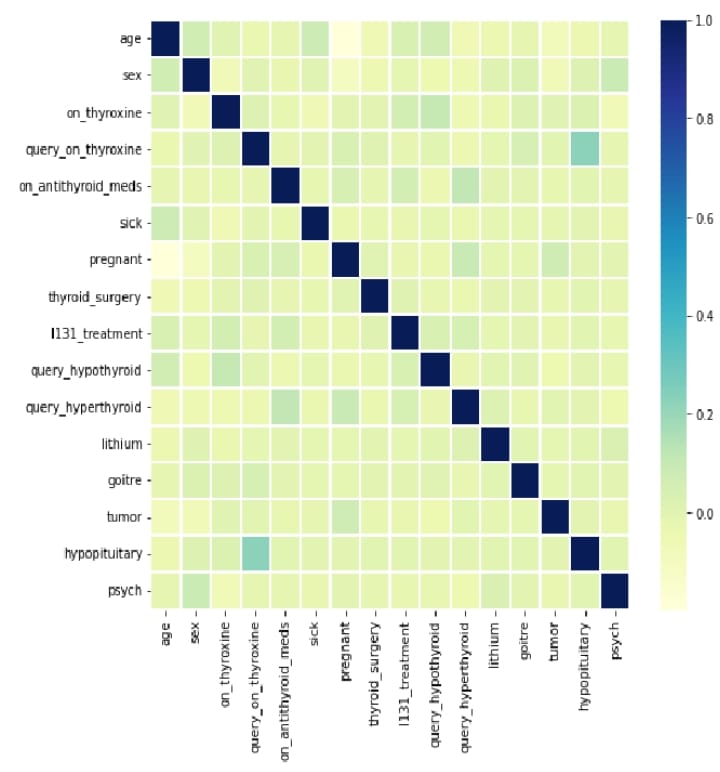


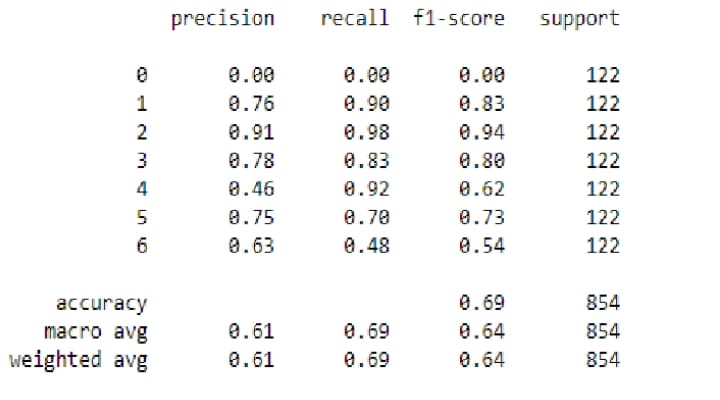


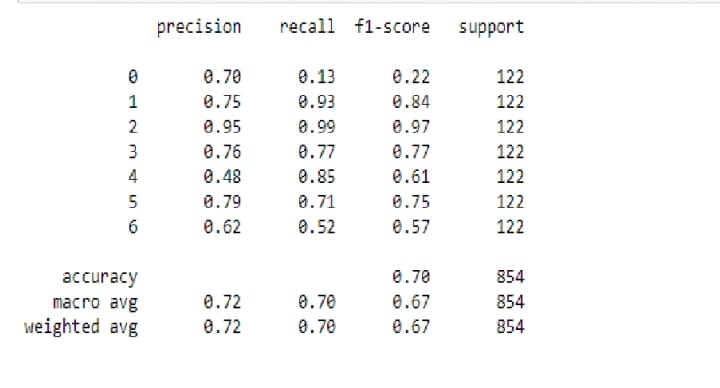


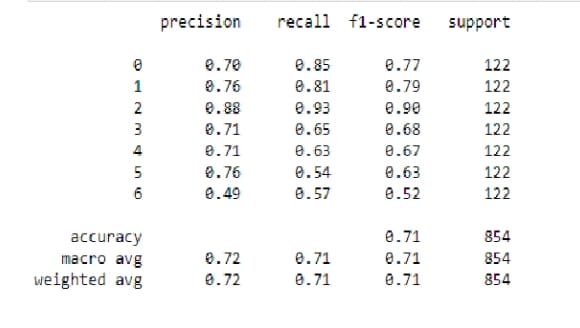


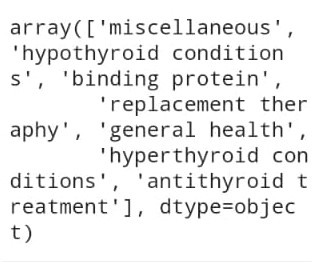












ADVANTAGES OF THYROID DISEASE CLASSIFICATION:

Classification of thyroid disease helps in accurate diagnosis ofthe condition. This is important because different types of thyroid disorders have different treaments and management strategies.

Better communication:

Standard classification systems help doctors and healthcare provides to communicate better with each other and with patients. It helps to ensure that everyone is using the same terminology and understanding the same concepts.

Predictive value:

Classification can have predictive value, such as predicting the risk of progression or response to treatment. This can help in the decision-making peocess for patient care.

DISADVANTAGES OF THYROID DISEASE CLASSIFICATION:

Overlapping symptoms:

Some thyroid disorders have overlapping symptoms which can make it difficult to accurately classify a patient’s condition. This can lead to misdiagnosis and improper treatment.

Complex classification system:

Some classification system are complex and require a high level of expertise to use effectively. This can be a barrier.

Limited accuracy:

The accuracy of thyroid disease classification can be limited by the subjective nature of some of the diagnostic criteria.

For example, the presence or absence of certain symptoms or physical disorder findings may be open to interpretation.

APPLICATION

Diagnosis:

The classification of thyroid disease can help doctors accurately diagnose a patient’s condition by identifying the specific

type of thyroid disorder they have.

Treatment:

The classification of thyroid diseases can help also help

determine the appropriate treatment options for patients. Different

thyroid diseases may require different meddicaions or procedures to manage symptoms and restore normal thyroid function.

Prognosis:

The classification of thyroid diseases can provide valuable information about the prognosis or expected course of the disease.

Research:

Accurate classification of thyroid diseases can help researchers better understand the underlying mechanisms of these disorders and develop more effective treatments in the future.

CONCLUSIONS

Thyroid disease classification using machine learning has shown promising results in accurately identifying differenttypes of thyroid disorders.

We are using the dataset then split the values trainand test and also apply the standard scaler and perform the futureimportance and analysing the visualization.

Different machine learning algorithms such as Decision trees, random forest classifier model, XGB classifier model,

svc model, and ANN model have been used to classify thyroid diseases.

These algorithms use a combination of features such as thyroid hormone levels, thyroid imaging results, and clinical symptoms to make accurate predictions.

Machine learning -based thyroid disease classification can be a useful tool in assisting healthcare professionals in making timely and accurate diagnosis. It has the potential to reduce the workload of clinicians

and improve patient outcomes by providing faster and more accurate diagnoses. However, further research and validation are needed to ensure the accuracy and reliability of these models before they can be implemented in clinical settings.

FUTURE CODE

The future scope of thyroid disease classification using machine learning is vast, and several advancements can be expected in this field. Some potential areas of development include:  
  
Integration with electronic health records (EHRs): As EHRs become more prevalent in healthcare, machine learning models can be trained on vast amounts of data to improve their accuracy and reliability. This integration can also help in tracking patient histories and improving patient outcomes.  
  
Development of personalized medicine: Machine learning models can be trained on individual patient data to develop personalized treatment plans for thyroid disorders. This can help in providing targeted treatment and improving patient outcomes.  
  
Integration with imaging technologies: Machine learning models can be integrated with imaging technologies to analyze images of the thyroid gland and improve the accuracy of thyroid disease classification.  
  
Expansion of data sources: Machine learning models can be trained on a wider range of data sources, such as genetic data, lifestyle data, and environmental data, to improve the accuracy of thyroid disease classification.  
  
Overall, the future scope of thyroid disease classification using machine learning is exciting, and these advancements have the potential to improve patient outcomes and reduce healthcare costs. However, it is essential to ensure that these models are accurate, reliable, and ethically sound before their widespread adoption in clinical settings

APPENDIX

SOURCE CODE

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import tensorflow

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Layer, Dense, Dropout

data = pd.read\_csv("/content/data.csv")

data.head()

data.isnull().sum()

data['target'].value\_counts()

#spilitting the data values as x and y

x=data.iloc[:,0:-1]

y=data.iloc[:,-1]

x

x['sex'].replace(np.nan, 'F',

x['sex'].value\_counts()

x['age']=x['age'].astype('float')

x['TSH']=x['TSH'].astype('float')

x['T3']=x['T3'].astype('float')

x['TT4']=x['TT4'].astype('float')

x['T4U']=x['T4U'].astype('float')

x['FTI']=x['FTI'].astype('float')

x['TBG']=x['TBG'].astype('float')

x.info()

x.replace(np.nan,'0', inplace=True)

from sklearn.model\_selection import train\_test\_split

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.20,random\_state=0)

from imblearn.over\_sampling import SMOTE

y\_train.value\_counts()

from sklearn.preprocessing import StandardScaler

sc = StandardScaler()

x\_train\_bal = sc.fit\_transform(x\_train\_bal)

x\_test\_bal=sc.transform(x\_test\_bal)

x\_train\_bal

#gets importance

feature\_importance=['age','sex','on\_thyroxine','query\_on\_thyroxine','on\_antithyroid\_meds','sick','pregnant','thyroid\_surgery','I131\_treatment','query\_hypothyroid','query\_hyperthyroid','lithium','goitre','tumor','hypopituitary','psych','TSH','T3','TT4','T4U','FTI','TBG']

importance = results.importances\_mean

importance = np.sort(importance)

#summarize feature importance

for i,v in enumerate(importance):

i=feature\_importance[i]

print('feature: {:<20} Score: {}'. format(i,v))

#plot important feature

plt.figure(figsize=(10,10))

plt.bar(x=feature\_importance, height = importance)

plt.xticks(rotation=30, ha='right')

plt.show()

x\_train\_bal.head()

x\_train\_bal.head()

x\_test\_bal.head()

data.info()

#checking correlation using Heatmap

import seaborn as sns

corrmat = x.corr()

f, ax = plt.subplots(figsize =(9,8))

sns.heatmap(corrmat, ax = ax,

y\_pred= sv1.predict(x\_test\_bal)

print(classification\_report(y\_test\_bal,y\_pred))

#saving the model

import pickle

with open('thyroid\_1\_model.pkl', 'wb') as file:

pickle.dump(sv1, file)

features = np.array([[0,0,0,0,0.000000,0.0,0.0,1.00,0.0,40.0,0,0,0,0.0,0,0,0,0,0.00,0,0.1,0]])

print(label\_encoder.inverse\_transform(xgb1.predict(features)))

pickle.dump(label\_encoder,open('label\_encoder.pkl','wb'))

data['target'].unique()

y['target'].unique()

import pickle

pickle.dump(sv1,open('thyroid\_1\_model.pkl', 'wb'))