**Effective Heart Disease Prediction Using IBM Auto AI Service**

**Submitted by : Prof. Bhulakshmi.Bonthu**

**B.Tech(CSE), M.Tech(CSE)-IIT Madras.**

**Assistant Professor (Sr),VIT-Vellore institute of Technology**

**Vellore.**

### ACKNOWLEDGEMENTS

In effectively finishing this project mentors Mrs. Shivani Kapoor and Mr. Hemanth , SmartBridge Educational Services Pvt. Ltd want to say special thanks.

# Executive Summary

Cardiac arrest is a life-threatening cessation of activity in the heart. Early prediction of cardiac arrest is important, as it allows for the necessary measures to be taken to prevent or intervene during the onset. Artificial intelligence (AI) technologies and machine learning models have been increasingly used to enhance the ability to predict and prepare for the patients at risk. In this Project I worked on early prediction of Heart failure by using Machine Learning models.

I have used 10801 health record with different parameters like avgheartbeatspermin, palpitationsperday ,cholesterol bmi, heart failure, age, sex, amilyhistory,smokerlast5yrs ,exerciseminperweek so that our accuracy will correct.

I have build an Auto AI model and build and web application to show case the prediction of heart failure results Finally by using this project we can predict that possibility and risk level to get the heart attack.

Contents

[ACKNOWLEDGEMENTS 1](#_Toc110118573)

[Executive Summary 1](#_Toc110118574)

[1. INTRODUCTION 5](#_Toc110118575)

[ABSTRACT: 6](#_Toc110118576)

[SOFTWARE AND HARDWARE REQUIREMENTS 7](#_Toc110118577)

[SOFTWARE: 7](#_Toc110118578)

[EXISTING METHOD: 8](#_Toc110118579)

[DISADVANGTAGES: 8](#_Toc110118580)

[PROPOSED METHOD: 8](#_Toc110118581)

[BLOCK DIAGRAM: 9](#_Toc110118582)

[2. BACKEND MODULE DIAGRAM: 9](#_Toc110118583)

[MODULE DESCRIPTION: 10](#_Toc110118584)

[DATA CLEANING: 10](#_Toc110118585)

[DATA SPLITTING: 10](#_Toc110118586)

[TRAINING: 10](#_Toc110118587)

[ARCHITECTURE: 11](#_Toc110118588)

[Flow of operations to build a model and web application 11](#_Toc110118589)

[DESIGN: 12](#_Toc110118590)

[OUTPUT DESIGN: 12](#_Toc110118591)

[MODULES: 12](#_Toc110118592)

[METHODOLOGY AND ALGORITHMS: 13](#_Toc110118593)

[Extra Trees Classifier: 13](#_Toc110118594)

[CODE: 13](#_Toc110118595)

[CONCLUSION: 19](#_Toc110118596)

[REFERNCES: 19](#_Toc110118597)

### INTRODUCTION

* 1. **OBJECTIVE:**

The goal of this project is the prediction on heart failure of a person. The second objective is to build Auto AI model and build and web application to show case the prediction of heart failure results.

* 1. **MOTIVATION:**

Over the last few decades, heart disease is the most common cause of global death. So early detection of cardiac arrest and continuous monitoring can reduce the mortality rate. Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide. I used Machine Learning models in this project for prediction of heart failure , in this process I strengthened my knowledge in Machine Learning, IBM Watson Studio, IBM Watson Machine Learning, Node-RED,IBM Cloud Object Storage from this project.

* 1. **BACKGROUND STUDY:**

Heart failure and other heart sicknesses are the most well-known issue in greater part of individuals, and there are different elements that go about as spine of this issue like individuals are not focusing towards wellbeing predominantly in light of work pressure, sluggishness, unsatisfactory nature of food that outcomes in expanding cholesterol and troublesome determination of coronary illness because of absence of innovation, techniques utilized in diagnosing these sicknesses and thus having a ton of tests. A ton of exploration and clinical supporting frameworks are creating step by step, in any case, each framework have its different elements or benefits and restrictions which are obscure to one or the other side. Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide.

Many people were died due to this sudden attack of heart, based on food habits and age and so many parameters the person will suffer.

To overcome this in project I try to predict on possibility on heart failure Of person .

## ABSTRACT:

Throughout the course of recent many years, coronary illness is the most widely recognized reason for worldwide demise. So early location of heart failure and consistent observing can decrease the death rate. The remarkable development of information from various sources, for example, wearable sensor gadgets utilized in Internet of Things wellbeing observing, web based framework and others have been producing a tremendous measure of information on a nonstop premise.

The blend of streaming large information examination and AI is a leading edge innovation that can have a huge effect in medical services field particularly early discovery of heart failure. This innovation can be all the more remarkable and more affordable.

To conquer this issue, this project propose a heart failure expectation framework in view of Machine Learning models and AI with Machine Learning ,IBM Watson Studio, IBM Watson Machine Learning, Node-RED,IBM Cloud Object Storage which stand as areas of strength for a scale disseminated processing stage that can be utilized effectively for streaming information occasion against Machine Learning models through in-memory calculations.

By using general health records and few parameters with Machine Learning techniques I try to predict the heart failure with high accuracy.

\

## SOFTWARE AND HARDWARE REQUIREMENTS

**HARDWARE:**

**Operating System :** Windows 10

**RAM :** 4 GB

**HARD DISC/SSD :** 128 GB

**PROCESSOR :** INTEL 3rd generation

## SOFTWARE:

Services Used:

IBM Watson Studio

IBM Watson Machine Learning

Node-RED

IBM Cloud Object Storage

**1.**

## EXISTING METHOD:

In the existing system the accuracy was low and there is a chance of more deaths. The existing methods used SVM and KNN methods for predicting the heart failure and very less data sets and few parameters were considered by existing system which leads to less accuracy.

## DISADVANGTAGES:

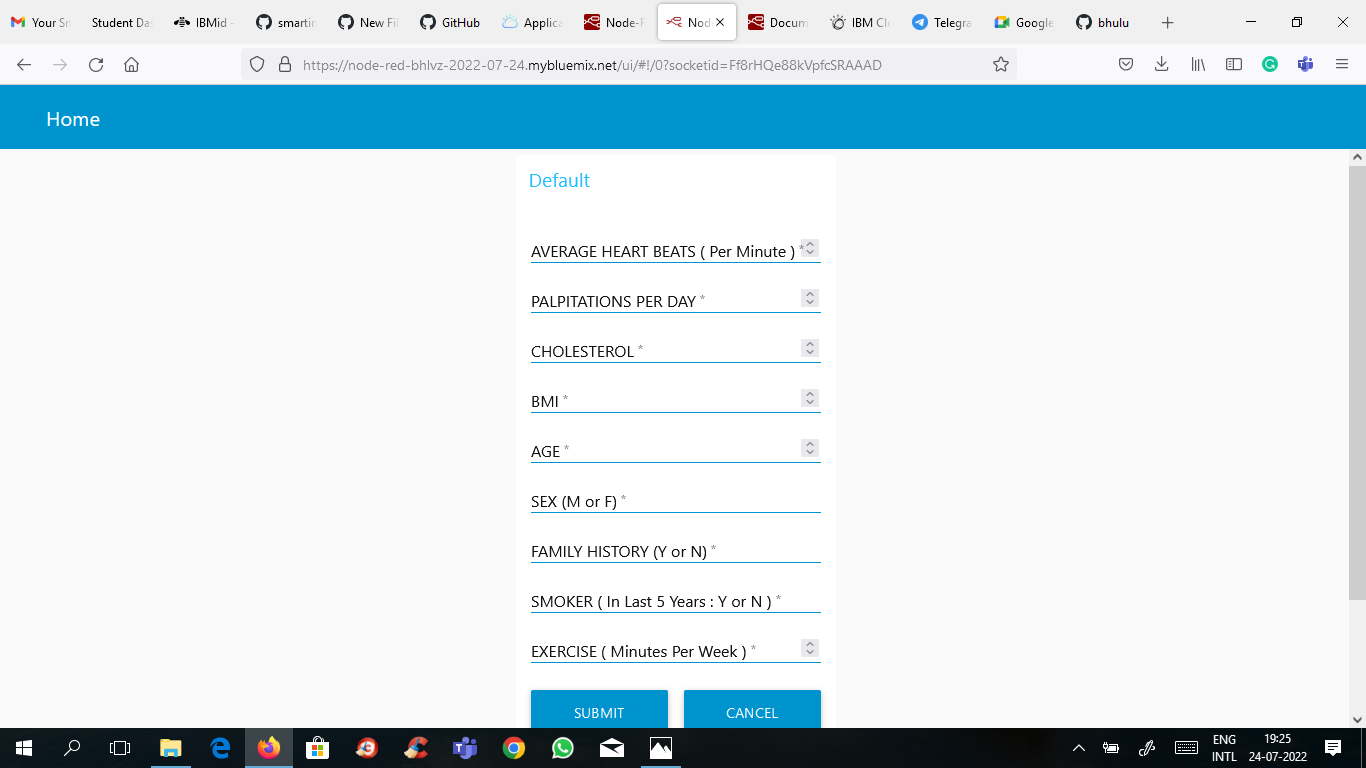
* Several risk scoring systems are used to identify patients at high risk of serious adverse events including unexpected inpatient death
* Accuracy low
* Cardiovascular disease and cancer are leading causes of morbidity and mortality, and can both be present in one patient. In patients with simultaneous disease, the most threatening disease should be treated first.
* Improved depression scores, and decreased health care utilization
* It may increase death rate.

## PROPOSED METHOD:

In our model automatically the appropriate Extra Trees Classifier chosen to predict the heart failure. I used IBM Watson Machine Learning to build Auto AI model and build and web application to show case the prediction of heart failure results.

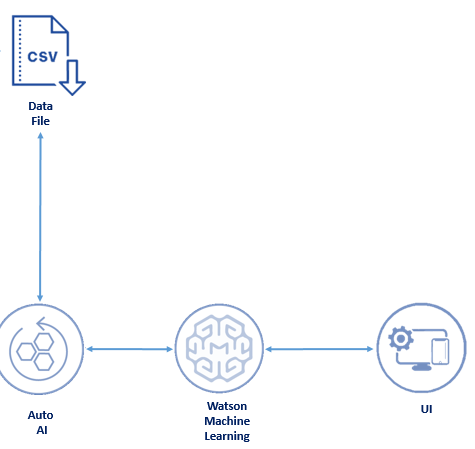
## BLOCK DIAGRAM:

* 1. **FRONT END MODULE DIAGRAM:**

**­­**

1. interface to enter the parameters

## BACKEND MODULE DIAGRAM:



1. Auto.AI,machine learning model and UI

## MODULE DESCRIPTION:

**DATA SET:**

The data set have collected from various health websites and old records

The data set consists of 10 different types of parameters are avgheartbeatspermin palpitationsperday cholesterol bmi heartfailure age sex familyhistory smokerlast5yrs exerciseminperweek

Here **output** indicates the result of heart failure possibility and risk level.

## DATA CLEANING:

In this module the information is cleaned. In the wake of cleaning of the information, the information is gathered according to prerequisite. This gathering of information is known as information grouping. Then check assuming there is any missing worth in the informational collection or not. It there is some missing worth then, at that point, change it by any default esteem. After that if any data need to change its format, it is done. That total process before the prediction is known is data pre-processing.

## DATA SPLITTING:

For this project the entire dataset was split into 80% training set and 20% test set. We used the training set for resampling, hyper parameter tuning, and training the machine and we used test set to test the performance of the trained model. While splitting the data, we specified a random seed (any random number), which ensured the same data split every time the program executed.

## TRAINING:

Algorithms learn from data. They track down connections, foster figuring out, simply decide, and assess their certainty from the preparation information they're given. Also, the more the preparation information is, the more the model performs. As a matter of

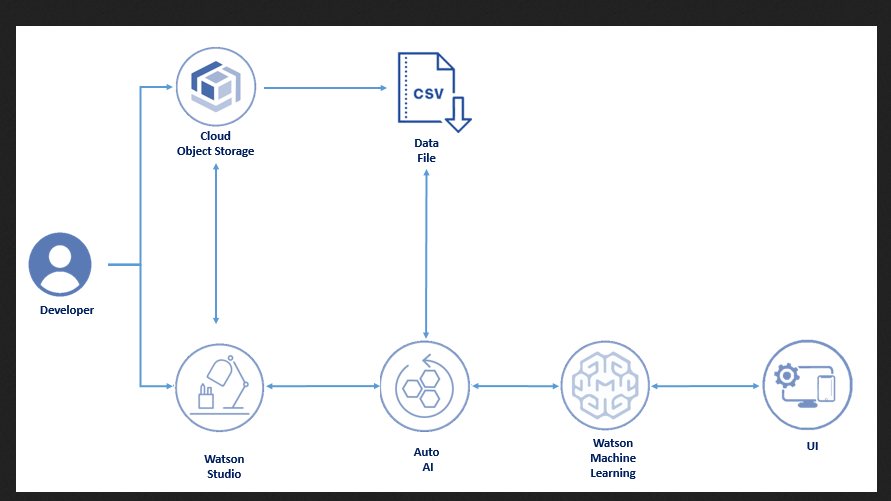
fact, the quality and amount of your preparation information has as a lot to do with the outcome of your information project as the actual calculations.

In other words, the data you want to use for training usually needs to be enriched or labeled. Or you might just need to collect more of it to power your algorithms. But chances are, the data you’ve stored isn’t quite ready to be used to train your classifiers. Because if you’re trying to make a great model, you need great training data. Classifier Training:

## ARCHITECTURE:

### Flow of operations to build a model and web application

1. Log in to IBM account
2. Create IBM Watson Studio and Node-RED Service
3. Create a Watson studio project
4. ADD Auto AI Experiment
5. Run the Auto AI Experiment to build a Machine learning model on the desired dataset
6. Save the model
7. Deploy the model as a web server and generate scoring End Point
8. Create a WEB application Using Node-RED to take user input and showcase Prediction on UI.

**3. Flow of operations**

## DESIGN:

**INPUT DESIGN:**

Input is given to the system that is trained, algorithms are applied and the result is obtained, we should remember the output will depend on the entries which we give. So the user must be careful while giving the input.

## OUTPUT DESIGN:

The design of output is the most important task of any system. During output design, developers identify the type of outputs needed and consider the necessary output controls and prototype report layouts.

## MODULES:

1. **USER collecting data set**
   1. View Home Page: the home webpage is visible where we can give input and get prediction
   2. Input Model: The user must enter the serial number in within the list
   3. View Results: The user view’s the generated results from the model
2. **SYSTEM**
   1. Working on the data set: The system checks for data whether it is available or not and load the data.
   2. Build Auto AI model.
   3. Pre-processing: Data need to be pre-processed according to the models it helps to increase the accuracy of the model and better information about the data.
   4. Model Building: To create a model that predicts heart failure with better accuracy, this module will help users.
   5. Generate Results: We train the machine learning algorithm and predict on heart failure.

## METHODOLOGY AND ALGORITHMS:

Machine Learning Techniques:

Here I build Auto AI model for predicting the heart failure.

## Extra Trees Classifier:

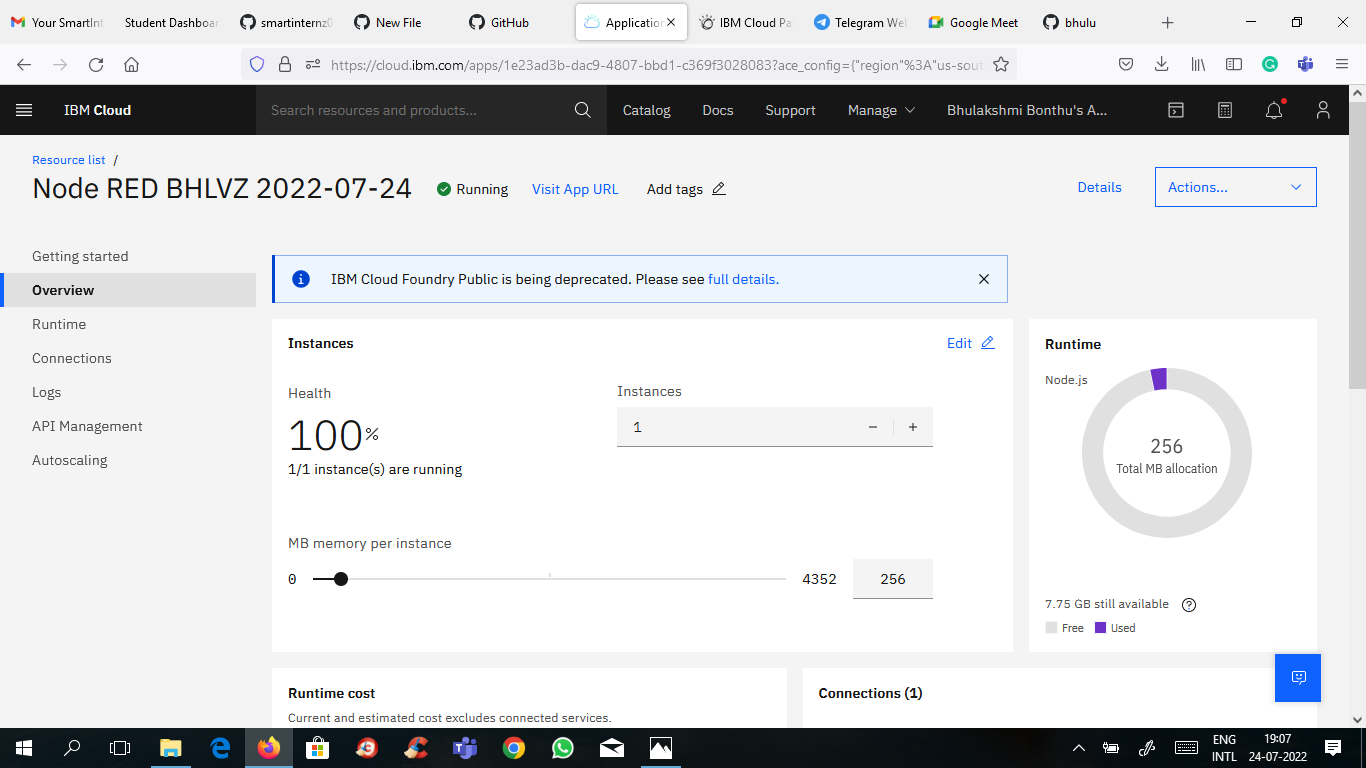
Extra Trees Classifier is a type of ensemble learning technique which aggregates the results of multiple de-correlated decision trees collected in a “forest” to output it’s classification result. In concept, it is very similar to a Random Forest Classifier and only differs from it in the manner of construction of the decision trees in the forest.

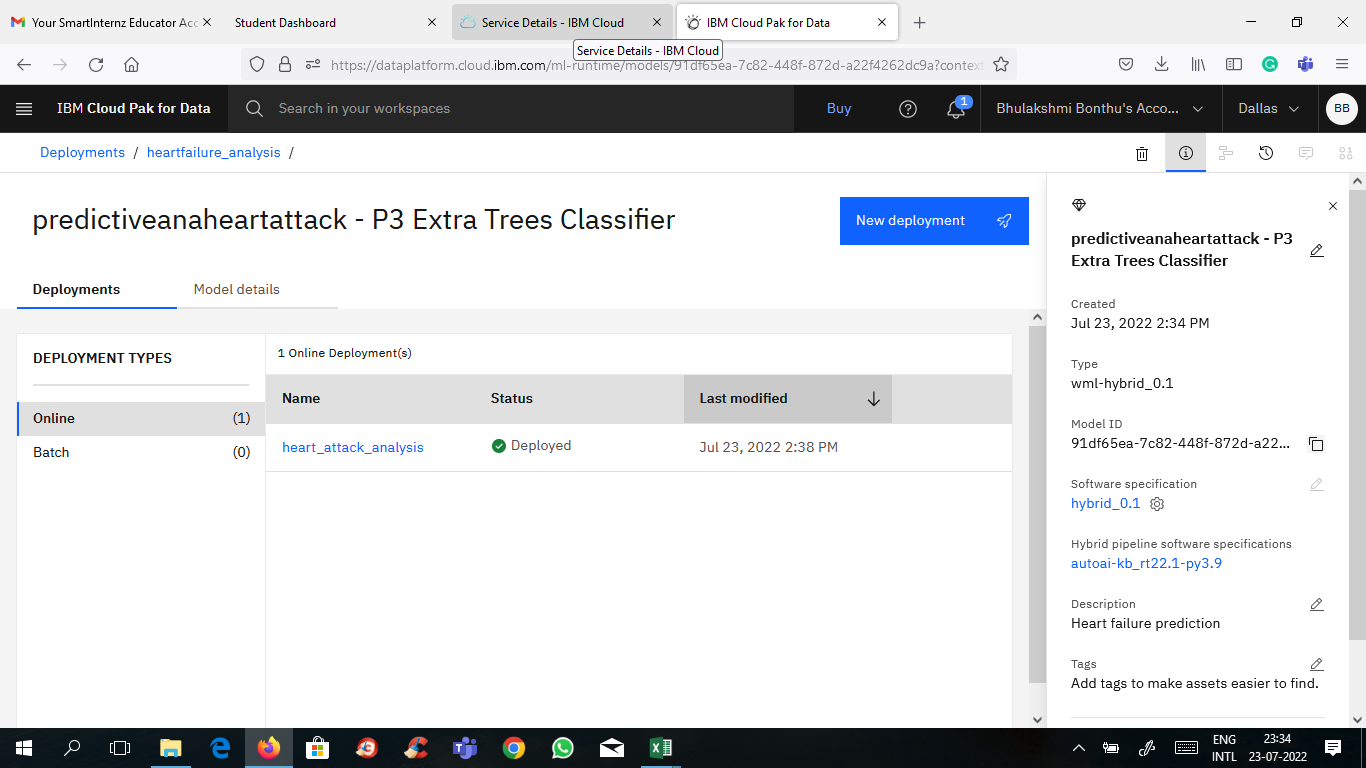
## CODE:

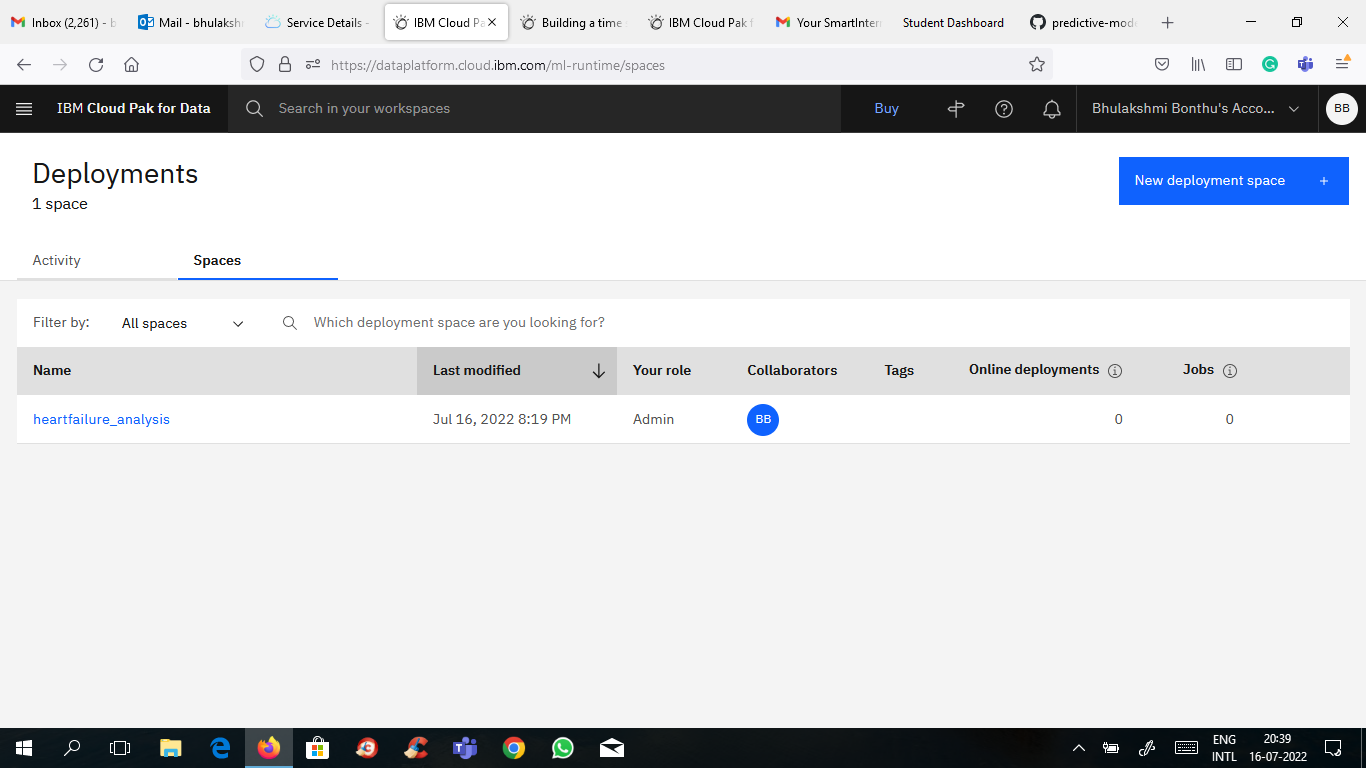
[{"id":"a4b58694.dd7a08","type":"tab","label":"HFPM DASHBAORD","disabled":false,"info":""},{"id":"ddaf6701.3ffd98","type":"debug","z":"a4b58694.dd7a08","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType":"auto","x":130,"y":320,"wires":[]},{"id":"21769e48.4b87b2","type":"function","z":"a4b58694.dd7a08","name":"pre-token","func":"global.set(\"hb\",msg.payload.hb)\nglobal.set(\"ppd\",msg.payload.ppd)\nglobal.set(\"ch\",msg.payload.ch)\nglobal.set(\"bm\",msg.payload.bm)\nglobal.set(\"ag\",msg.payload.ag)\nglobal.set(\"s\",msg.payload.s)\nglobal.set(\"fh\",msg.payload.fh)\nglobal.set(\"sm\",msg.payload.sm)\nglobal.set(\"em\",msg.payload.em)\n\nvar apikey=\"yjzHt3C2X48EZ3TAfAMykbIgV0z0-6FOVE7gTPPQJkGS\";\nmsg.headers={\"content-type\":\"application/x-www-form-urlencoded\"}\nmsg.payload={\"grant\_type\":\"urn:ibm:params:oauth:grant-type:apikey\",\"apikey\":apikey}\nreturn msg;\n","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":140,"y":160,"wires":[["1775fa.9dce3a06"]]},{"id":"1775fa.9dce3a06","type":"http request","z":"a4b58694.dd7a08","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":350,"y":160,"wires":[["c1b55bc1.e78c38","34bd01eb.09e15e"]]},{"id":"c1b55bc1.e78c38","type":"debug","z":"a4b58694.dd7a08","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType":"auto","x":570,"y":260,"wires":[]},{"id":"34bd01eb.09e15e","type":"function","z":"a4b58694.dd7a08","name":"Pre Prediction","func":"var hb = global.get(\"hb\")\nvar ppd = global.get(\"ppd\")\nvar ch = global.get(\"ch\")\nvar bm = global.get(\"bm\")\nvar ag = global.get(\"ag\")\nvar s = global.get(\"s\")\nvar fh = global.get(\"fh\")\nvar sm = global.get(\"sm\")\nvar em = global.get(\"em\")\n\nvar token=msg.payload.access\_token\nmsg.headers={'Content-Type': 'application/json',\"Authorization\":\"Bearer \"+token,\"Accept\":\"application/json\"}\nmsg.payload={\"input\_data\":[{\"fields\":[\"AVGHEARTBEATSPERMIN\",\"PALPITATIONSPERDAY\",\"CHOLESTEROL\",\n \"BMI\",\"AGE\",\"SEX\",\"FAMILYHISTORY\",\"SMOKERLAST5YRS\",\"EXERCISEMINPERWEEK\"],\n \"values\":[[hb,ppd,ch,bm,ag,s,fh,sm,em]]}]}\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":580,"y":60,"wires":[["119647ca.2b3888"]]},{"id":"119647ca.2b3888","type":"http request","z":"a4b58694.dd7a08","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://us-south.ml.cloud.ibm.com/ml/v4/deployments/11934\_heart\_attack/predictions?version=2022-07-23","tls":"","persist":false,"proxy":"","authType":"","senderr":false,"x":730,"y":140,"wires":[["750758a0.e310c8","3cbb3720.1d3b78","7a3f7170.c6936"]]},{"id":"750758a0.e310c8","type":"debug","z":"a4b58694.dd7a08","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","statusVal":"","statusType":"auto","x":990,"y":140,"wires":[]},{"id":"5a3e627e.04601c","type":"inject","z":"a4b58694.dd7a08","name":"","props":[{"p":"payload"},{"p":"topic","vt":"str"}],"repeat":"","crontab":"","once":false,"onceDelay":0.1,"topic":"","payload":"","payloadType":"date","x":100,"y":80,"wires":[["21769e48.4b87b2"]]},{"id":"e929f78c.388bb8","type":"debug","z":"a4b58694.dd7a08","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","statusVal":"","statusType":"auto","x":1070,"y":60,"wires":[]},{"id":"3cbb3720.1d3b78","type":"function","z":"a4b58694.dd7a08","name":"Parsing","func":"if (msg.payload.predictions[0].values[0][0] == 'N')\n {\n msg.probablity = msg.payload.predictions[0].values[0][1][0];\n msg.payload = 'Not at Risk';\n }\nelse \n {\n msg.probability = msg.payload.predictions[0].values[0][1][1];\n msg.payload = 'At Risk';\n }\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":880,"y":80,"wires":[["e929f78c.388bb8","ecc83872.c85c08"]]},{"id":"ecc83872.c85c08","type":"ui\_text","z":"a4b58694.dd7a08","group":"80c2766c.ace4c8","order":2,"width":0,"height":0,"name":"","label":"Prediction","format":"{{msg.payload}}","layout":"row-spread","x":1060,"y":100,"wires":[]},{"id":"df6f8738.fbce68","type":"ui\_form","z":"a4b58694.dd7a08","name":"","label":"","group":"80c2766c.ace4c8","order":1,"width":0,"height":0,"options":[{"label":"AVERAGE HEART BEATS ( Per Minute )","value":"hb","type":"number","required":true,"rows":null},{"label":" PALPITATIONS PER DAY","value":"ppd","type":"number","required":true,"rows":null},{"label":"CHOLESTEROL","value":"ch","type":"number","required":true,"rows":null},{"label":"BMI","value":"bm","type":"number","required":true,"rows":null},{"label":"AGE","value":"ag","type":"number","required":true,"rows":null},{"label":"SEX (M or F)","value":"s","type":"text","required":true,"rows":null},{"label":"FAMILY HISTORY (Y or N)","value":"fh","type":"text","required":true,"rows":null},{"label":"SMOKER ( In Last 5 Years : Y or N )","value":"sm","type":"text","required":true,"rows":null},{"label":"EXERCISE ( Minutes Per Week )","value":"em","type":"number","required":true,"rows":null}],"formValue":{"hb":"","ppd":"","ch":"","bm":"","ag":"","s":"","fh":"","sm":"","em":""},"payload":"","submit":"Submit","cancel":"Cancel","topic":"","x":110,"y":240,"wires":[["ddaf6701.3ffd98","21769e48.4b87b2"]]},{"id":"addbc880.c08138","type":"ui\_text","z":"a4b58694.dd7a08","group":"80c2766c.ace4c8","order":2,"width":0,"height":0,"name":"","label":"Score","format":"{{msg.payload}}","layout":"row-spread","x":1050,"y":240,"wires":[]},{"id":"7871bbfa.db2964","type":"debug","z":"a4b58694.dd7a08","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","statusVal":"","statusType":"auto","x":1070,"y":200,"wires":[]},{"id":"7a3f7170.c6936","type":"function","z":"a4b58694.dd7a08","name":"","func":"if (msg.payload.predictions[0].values[0][0] == 'N')\n {\n msg.payload = msg.payload.predictions[0].values[0][1][0]\n \n }\nelse \n {\n msg.paylaod = msg.payload.predictions[0].values[0][1][1]\n }\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":880,"y":220,"wires":[["7871bbfa.db2964","addbc880.c08138"]]},{"id":"80c2766c.ace4c8","type":"ui\_group","name":"Default","tab":"547fcc14.7f90b4","order":1,"disp":true,"width":"6","collapse":false},{"id":"547fcc14.7f90b4","type":"ui\_tab","z":"a4b58694.dd7a08","name":"Home","icon":"dashboard","disabled":false,"hidden":false}]

**OUTPUT:**

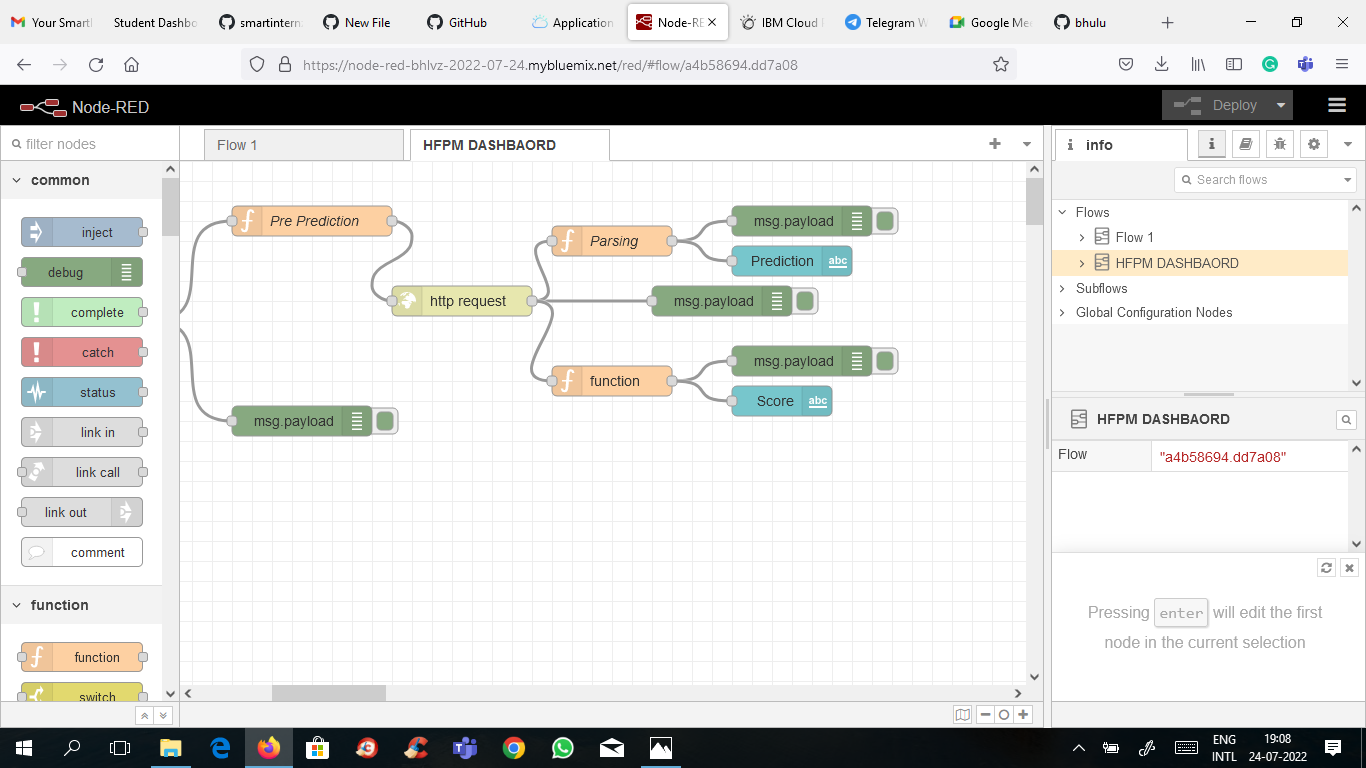
**Node red service and deployed .**

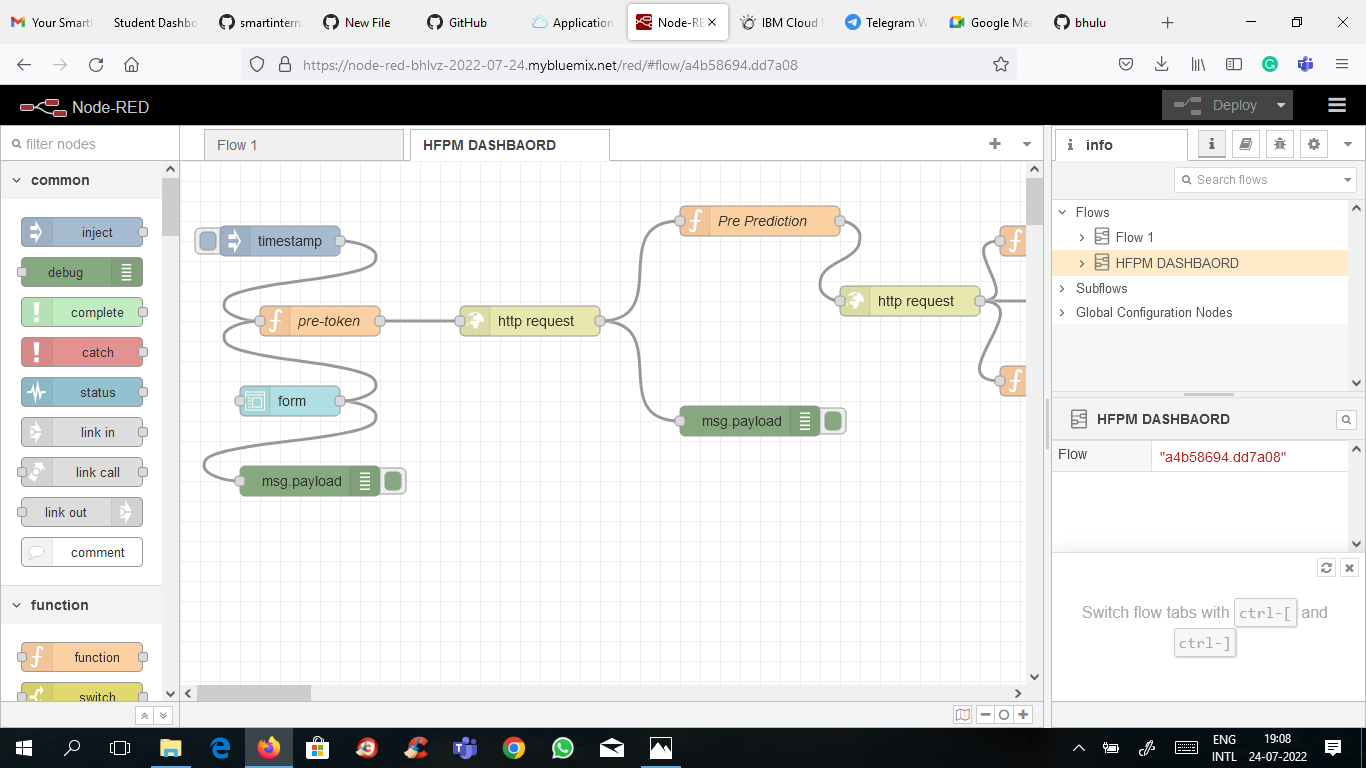
****

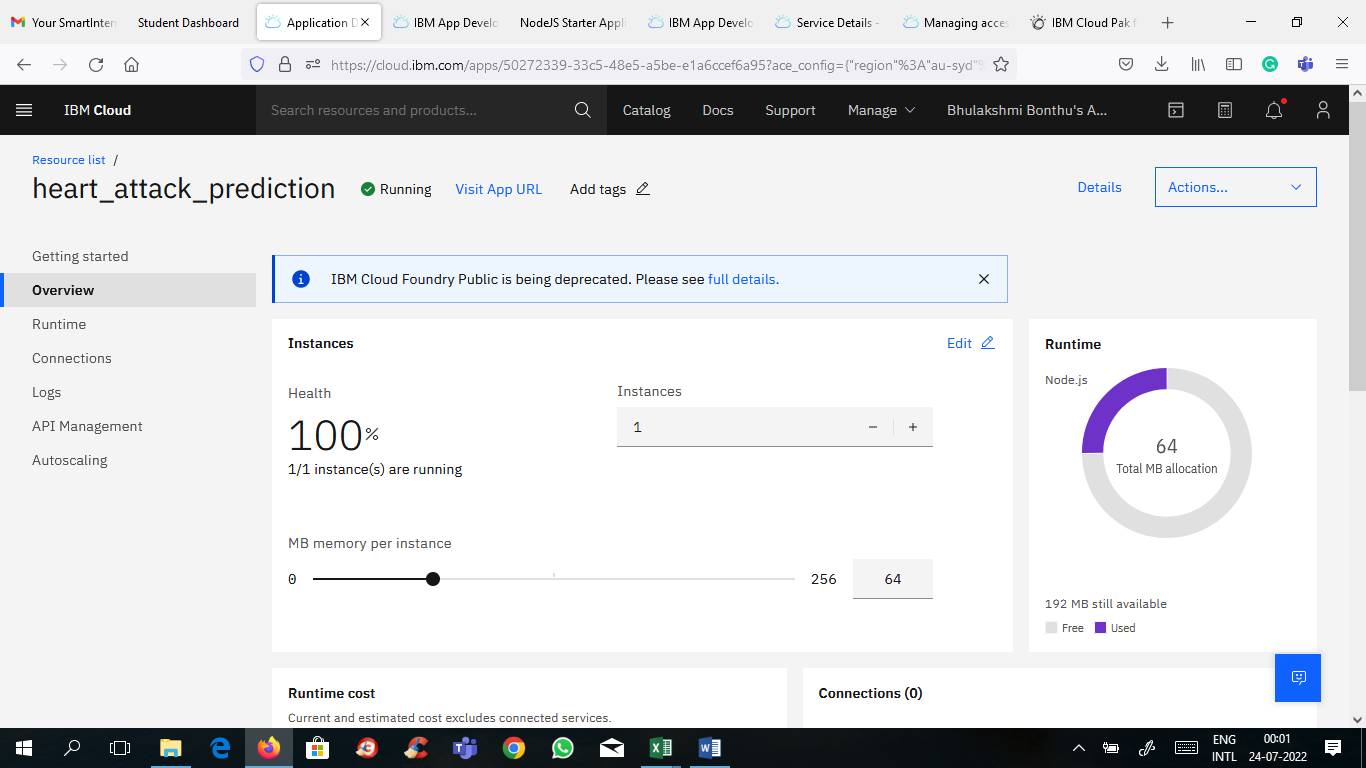
****

****

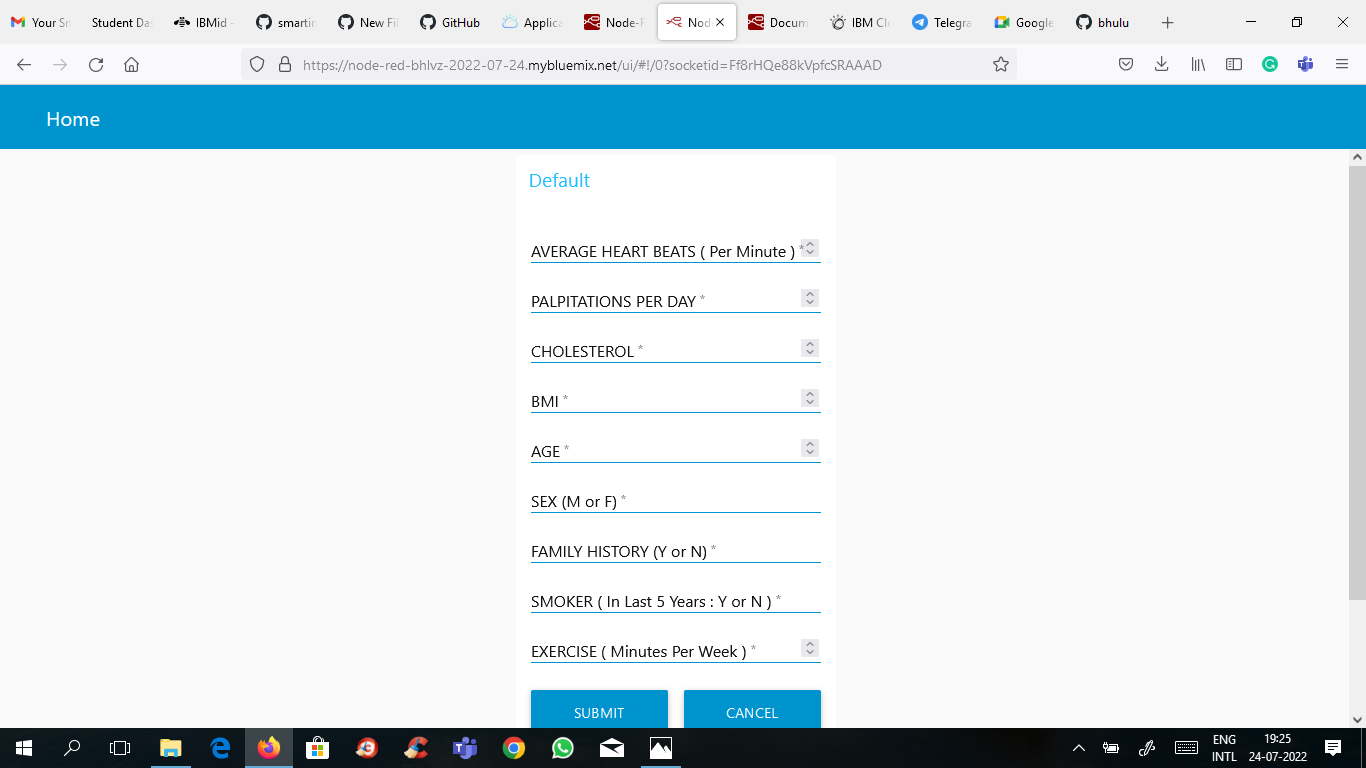
**NODE-RED flows:**

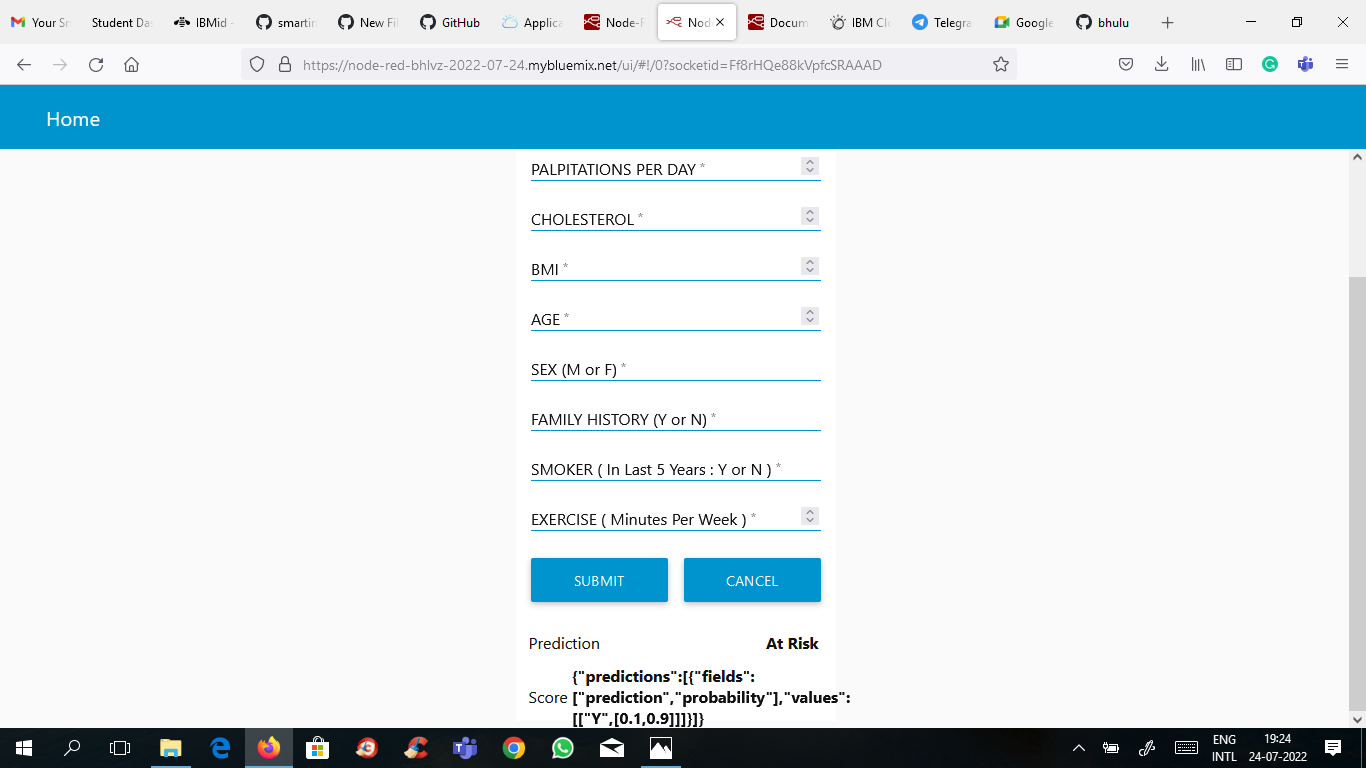






UI interface :





## CONCLUSION:

Machine learning is one of the most exciting technologies that one would have ever come across. By using machine learning techniques I created this project which helps earlier prediction of heart failure which can save many lives. At the end of project I conclude **Extra tree classifier has given high accuracy on chosen data .**

## REFERNCES:

1. J. M. Kwon, Y. Lee, Y. Lee, S. Lee, and J. Park, “An Algorithm Based on Deep Learning for Predicting In ‐ Hospital Cardiac Arrest,” *Journal of the American Heart Association*, vol. 7, no. 13, e008678, 2018.
2. L. J. Blackhall, “Must we always use CPR,” *New England Journal Medicine*, vol. 317, no. 20, pp. 1281-1285, 1987.
3. W. B. Kouwenhoven, J. R. Jude, and G. G. Knickerbocker, “Closedchest cardiac massage,” *JAMA*, vol. 173, no. 10, pp. 1064-1067, 1960.
4. D. L. Atkins, et al., “Part 11: pediatric basic life support and cardiopulmonary resuscitation quality: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care,” *Circulation*, 132(18 suppl 2), pp. S519-S525.
5. D. Kindermann, R. Mutter, and J. M. Pines, “Emergency Department Transfers to Acute Care Facilities,” *HCUP Statistical Brief* #157, 2013.

*Healthcare Informatics (ICHI), 2016 IEEE International Conference on.* IEEE, 2016, pp. 93-101.

1. X. Guo, Y. Yin, C. Dong, G. Yang, and G. Zhou, “On the class imbalance problem,” in *Natural Computation, 2008. ICNC'08. Fourth International Conference on (Vol. 4, pp. 192-201)*. IEEE.