SMART INTERNZ-BUILD-A-THON

PROJECT REPORT

**PREDICT HEART FAILURE USING IBM AUTO AI SERVICE**

-**FELICIA SHARON 19BCE1215**

IBM Cloud Credentials

E-mail : [feliciasharon19@gmail.com](mailto:feliciasharon19@gmail.com)

Registered mail-id : [in.felicia.sharon@gmail.com](mailto:in.felicia.sharon@gmail.com)



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **CONTENTS** | | | | | | | |
| **Sno** | |  | **Title** | | |  |  |
| **1** | |  | **INTRODUCTION** | |  |  | **3** |
|  | |  | 1.1 Overview | |  |  |  |
|  | |  | 1.2 Purpose | |  |  |  |
| **2** | |  | **LITERATURE SURVEY** | |  |  | **5** |
|  | |  | 2.1 Existing problem | | |  |  |
|  | |  | 2.2 Proposed solution | | |  |  |
| **3** | |  | **THEORITICAL ANALYSIS** | |  |  | **7** |
|  | |  | 3.1 Block diagram | |  |  |  |
|  | |  | 3.2 Hardware / Software designing | | |  |  |
| **4** | |  | **EXPERIMENTAL INVESTIGATIONS** | | |  | **9** |
| **5** | |  | **FLOWCHART** | |  |  | **21** |
| **6** | |  | **RESULT** |  |  |  | **22** |
| **7** | |  | **ADVANTAGES & DISADVANTAGES** | | |  | **29** |
| **8** | |  | **APPLICATIONS** | |  |  | **29** |
| **9** | |  | **CONCLUSION** | |  |  | **30** |
| **10** | |  | **FUTURE SCOPE** | | **`** |  | **30** |
| **11** | |  | **BIBILOGRAPHY** | |  |  | **31** |
|  | |  | **APPENDIX** | |  |  |  |
|  |  | | A. Source code | |  |  |  |
|  |  | |  |  |  |  |  |

**1. INTRODUCTION**

Heart disease is considered as one of the major cause of deaths throughout the world. It cannot be easily predicted by the medical practitioners as it requires high expertise and accuracy. This project addresses the issue of prediction of heart failure according to some input attributes of a person using the Machine learning techniques.

ML is an emerging application of Artificial Intelligence that uses different analytics and statistical techniques in order to improve the performance of particular machine learning from old data. It enables a particular machine to learn from dataset and enhance the performance by experience. It helps to build an intelligent machine to solve the specific problem.

1.1 Overview

The project is a Web-based Heart Disease Prediction System (HDPS) that uses AI and ML algorithms. We are using HD dataset to train a model by comparing two efficient algorithms for HDPS Web application. The dataset contains more than 10,000 records and 9 features that help to train a prediction model that will be deployed into a web application for prediction. A web-based HDPS application is developed through IBM Cloud (Node-Red). Each phase is efficiently done. The project is successfully created with the help of requirement analysis and project plan, system design, database design, testing plan, identifying features and functionalities, and system validation and deployment.

1.2 Purpose

According to survey of WHO, 17 million global deaths are due to heart attacks and heart strokes. Bad clinical decisions can result in risky consequences which may not be accepted by the people. Its prediction is not easy as it requires high expertise and accuracy. The main aim of this project is to help the doctors to take proper and better decisions about the risk of heart disease for a patient. The main objective of this research is to build Heart Disease Prediction System that gives diagnosis of heart failure using historical heart database. It aims at bringing together medicinal decision support with computer-based results of patient data, which helps to reduce medicinal faults and brings improvement in the overall outcome. Such type of work can bring together all the available data, as a basis on which development of rational assumption about the future could be done.

**Project Description:**

Cardiovascular diseases (CVDs) are the number one cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide. Heart failure is a common event caused by CVDs and this dataset contains 9 features that can be used to predict mortality by heart failure.

In this project, a model using Auto AI is built and a web application is developed where we can get the prediction of heart failure using this model.

**2. LITERATURE SURVEY**

Heart Disease (HD)

HD is defined a range of conditions that affect your heart. It is describing any disorder of the heart. The umbrella of HD consists of different type of HD such as blood vessel diseases (coronary artery disease, and arrhythmias) and heart defects when you’re born with congenital heart defects, among others. (Mayo Clinic, 2019).

The term “Heart Disease” is always used interchangeably with the term “Cardiovascular Disease (CVD)”. CVD generally refers to conditions that involve blocked or narrowed blood vessels that can lead to a heart attack, stroke or chest pain (angina). (Mayo Clinic, 2019).

2.1 Existing solution

(Chala Beyene, 2018) Proposed a methodology to foretell the occurrence of HD to overcome the problem of diagnosis of HD. One of the existence methodology is by choosing Naïve Bayes, J48, and SVM for predicting the occurrence of HD for early automatic diagnosis in short time in order to support the qualities of services and reduce costs to save the life of individuals. This methodology uses various attributes of HD in order to identify whether a patient has HD or not. The analysis is done using WEKA software.

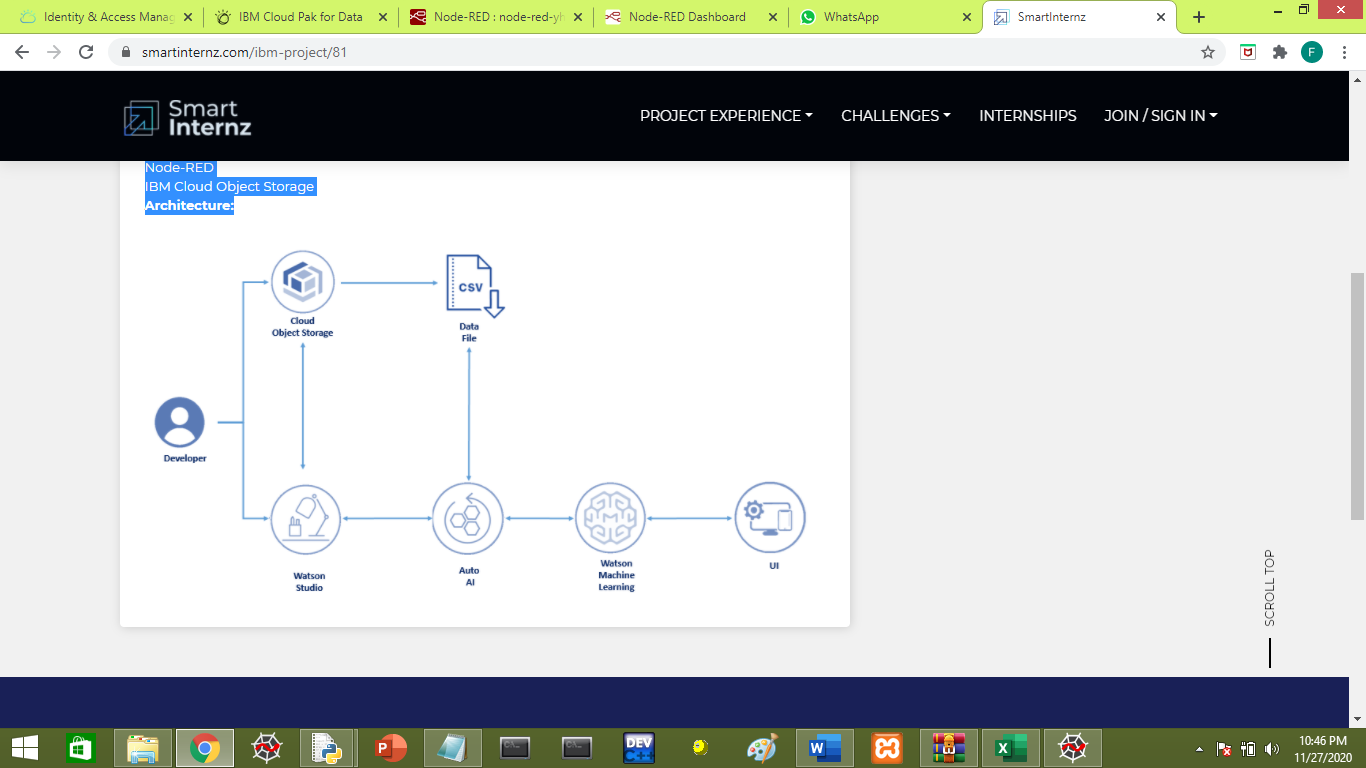
(Dwivedi, 2016) Focused to evaluate the performance of different ML algorithms for HD prediction. They compared the different algorithms such as Naïve Bayes, KNN, Logistic Regression and Classification tree to predict Heart Disease.

2.2 Proposed solution

In this project, a model using IBM Auto AI is developed to predict the heart failure. The AutoAI graphical tool in Watson Studio is used to automatically analyze the HD dataset and generate candidate model pipelines customized for the predictive modeling of heart-failure.  These model pipelines are created iteratively as AutoAI analyzes the dataset and discovers data transformations, algorithms, and parameter settings that work best for this problem setting. After reviewing the pipelines, the best model is saved for deployment and then tested for prediction. A web application is developed using node-red where we can get the prediction of heart failure using this model.

**3. THEORITICAL ANALYSIS**

Block Diagram



The steps are as follows:

* Collect the data.
* Create and run the experiment.
* Save a pipeline as a model.
* Deploy and test the model.

Software Designs

Services Used:

• IBM Watson Studio

• IBM Watson Machine Learning

• Node-RED

• IBM Cloud Object Storage

**4. EXPERIMENTAL INVESTIGATIONS**

Create an IBM Cloud Account.

To create an IBM Cloud account, go to https://www.ibm.com/cloud Click on Create an account. If an account exists then log in with the IBMid and password.

Then we create the following services in IBM Cloud:

● Create Watson Studio.

Open the catalog option in IBM Cloud account and select Services. Search for Watson studio to create the service.

● Create Machine learning service.

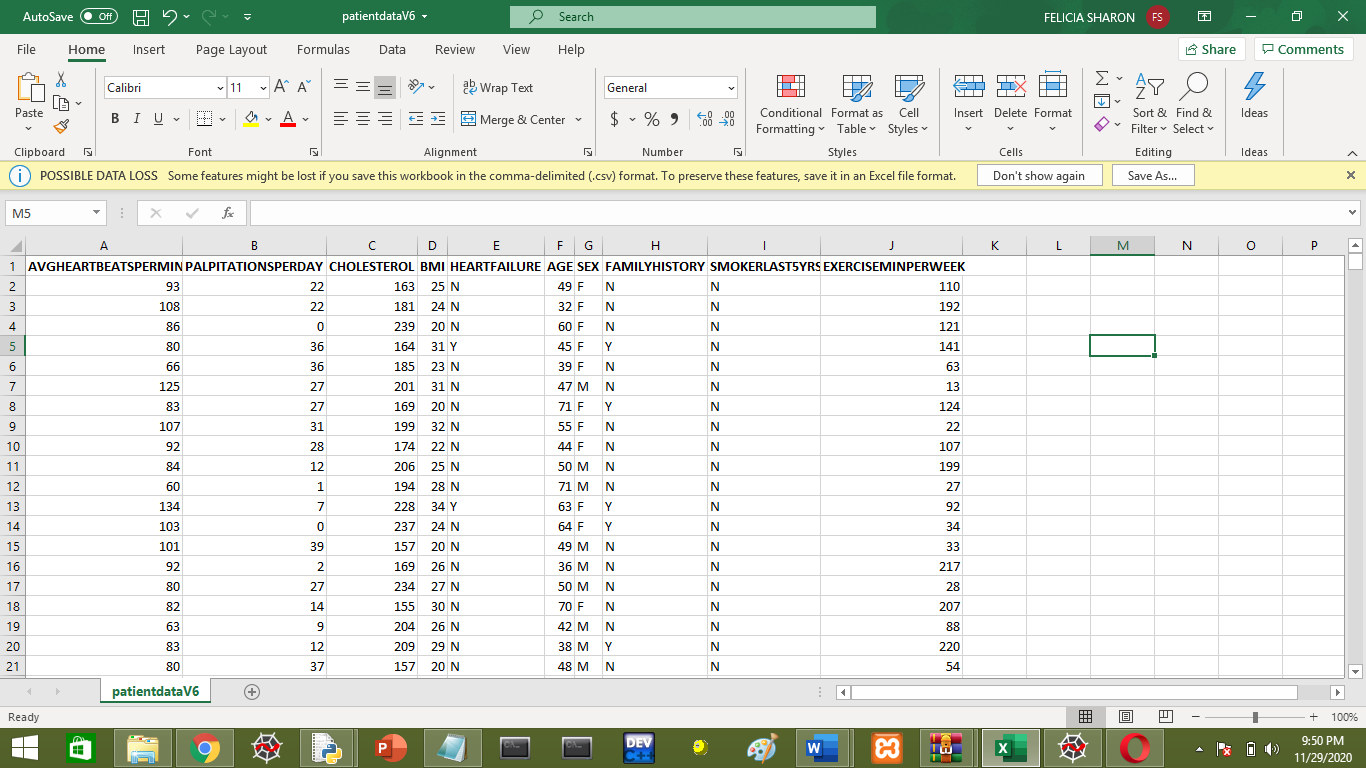
Open the catalog option in IBM Cloud account and select Services. Search Machine

learning service then create the service to access in Watson Studio.

**Back End**

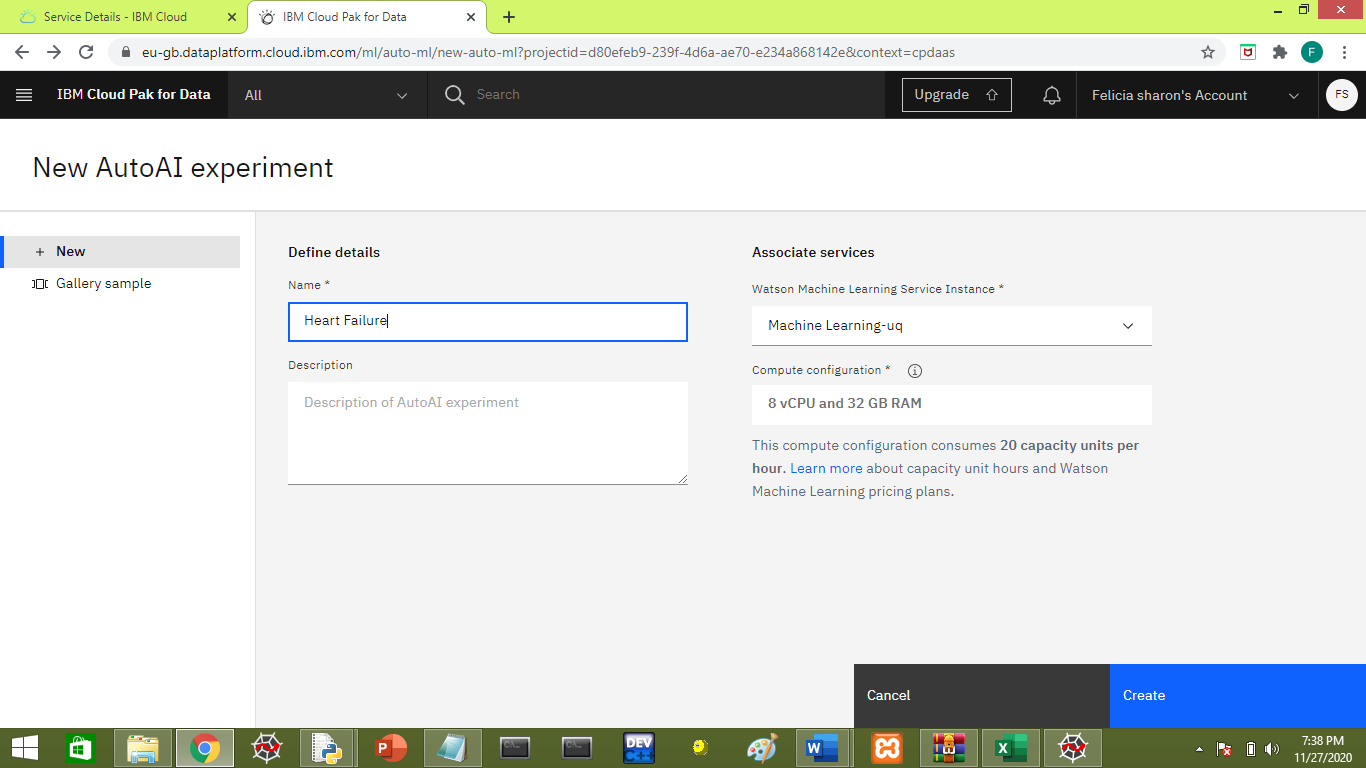
Collecting the data

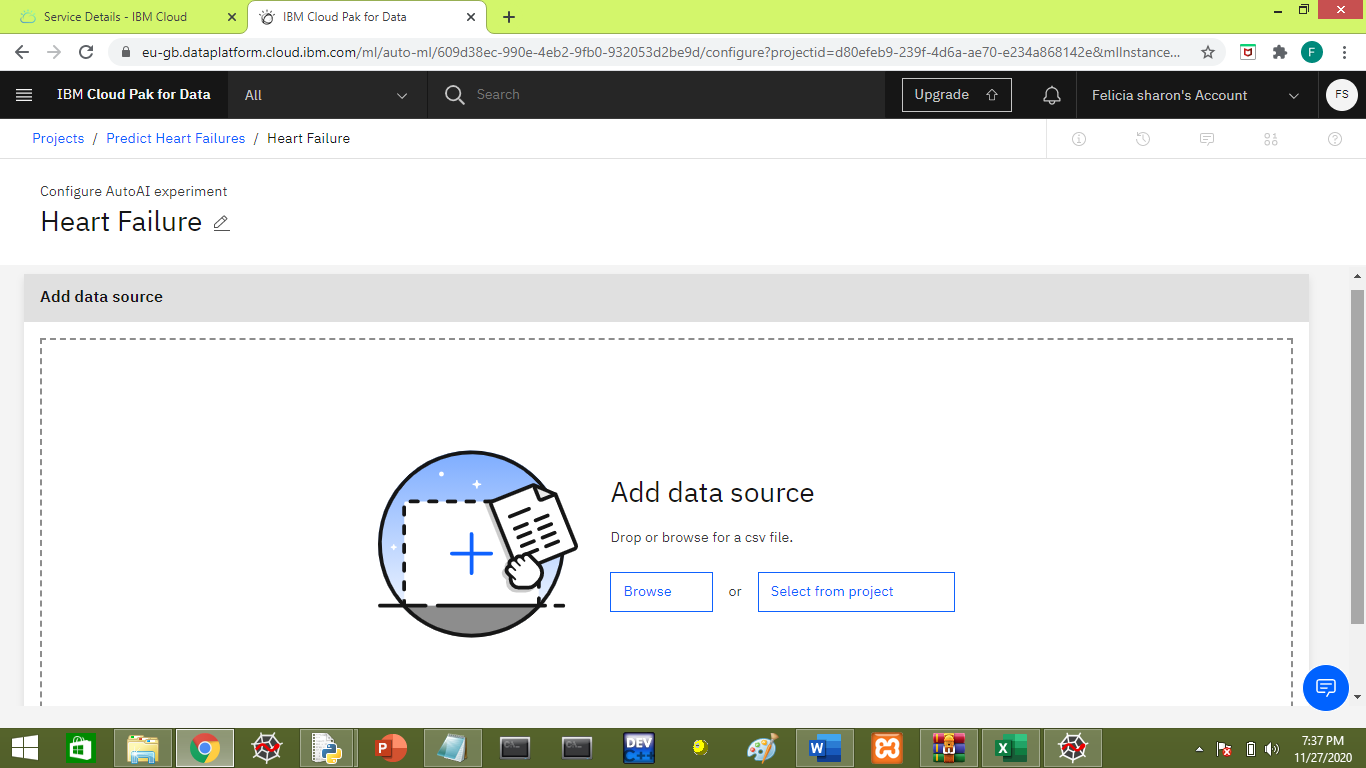
The dataset – “patientV6.csv” is from GitHub and has 9 features to predict whether a person has Heart Failure or not. The classification goal is to train a model that can predict if a new patient will check for the presence of Heart Failure (yes/no).



Define the experiment

From the Assets page of the project, click Add to project, then click AutoAI experiment. Name the experiment and add an optional description. Accept the default compute configuration and click Create. Add for the experiment. Browse for and upload the data file patientV6.csv.

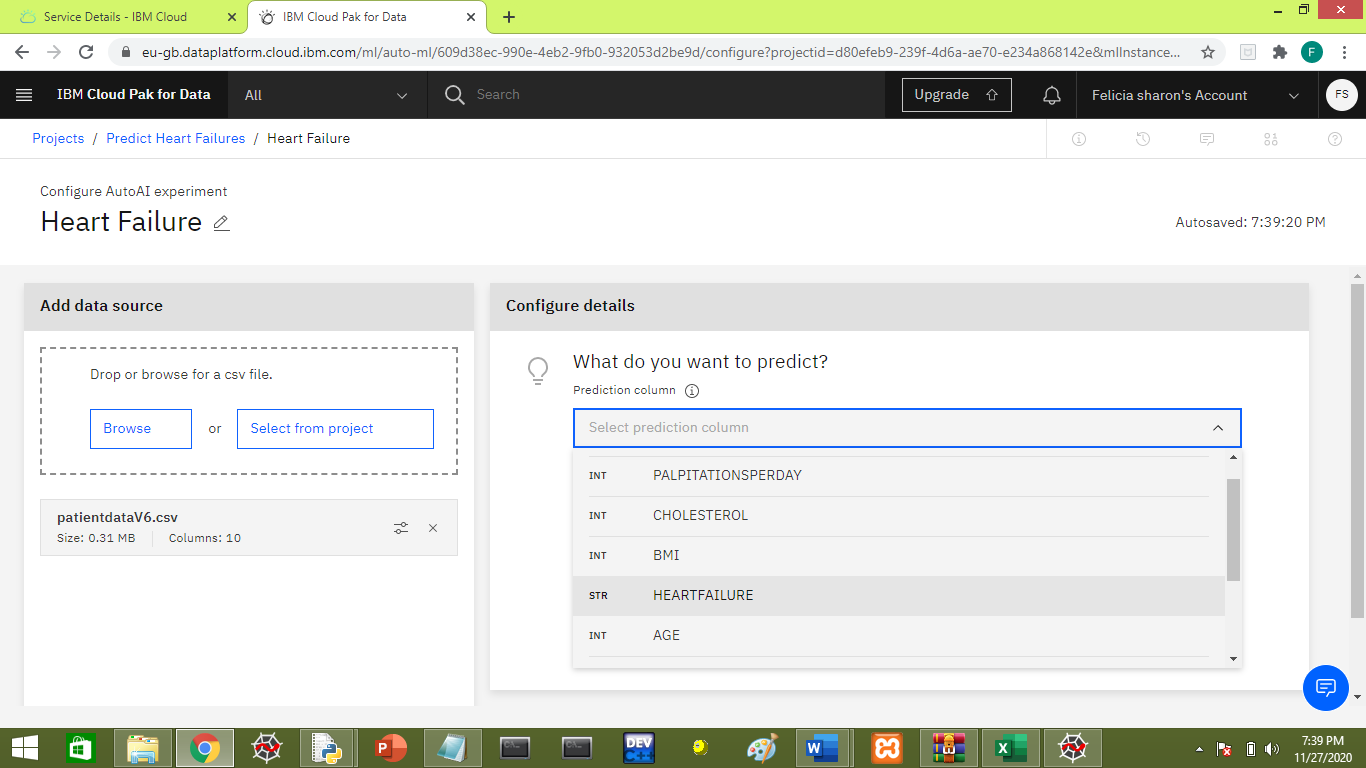


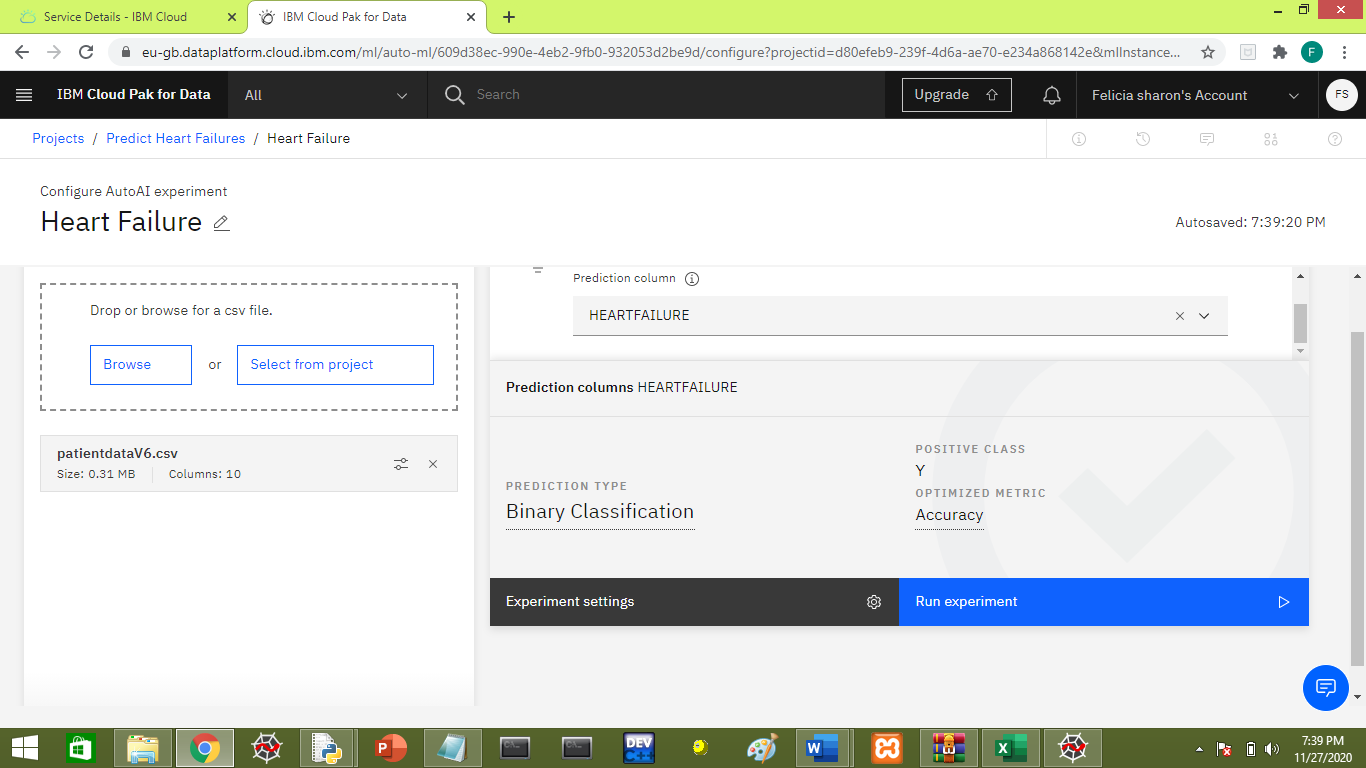


Prediction

After adding the data, choose a prediction column, which represents the problem you are trying to solve with the experiment. For this experiment, the aim is to know if a new patient will have a Heart Failure or not, represented by the column labeled y.

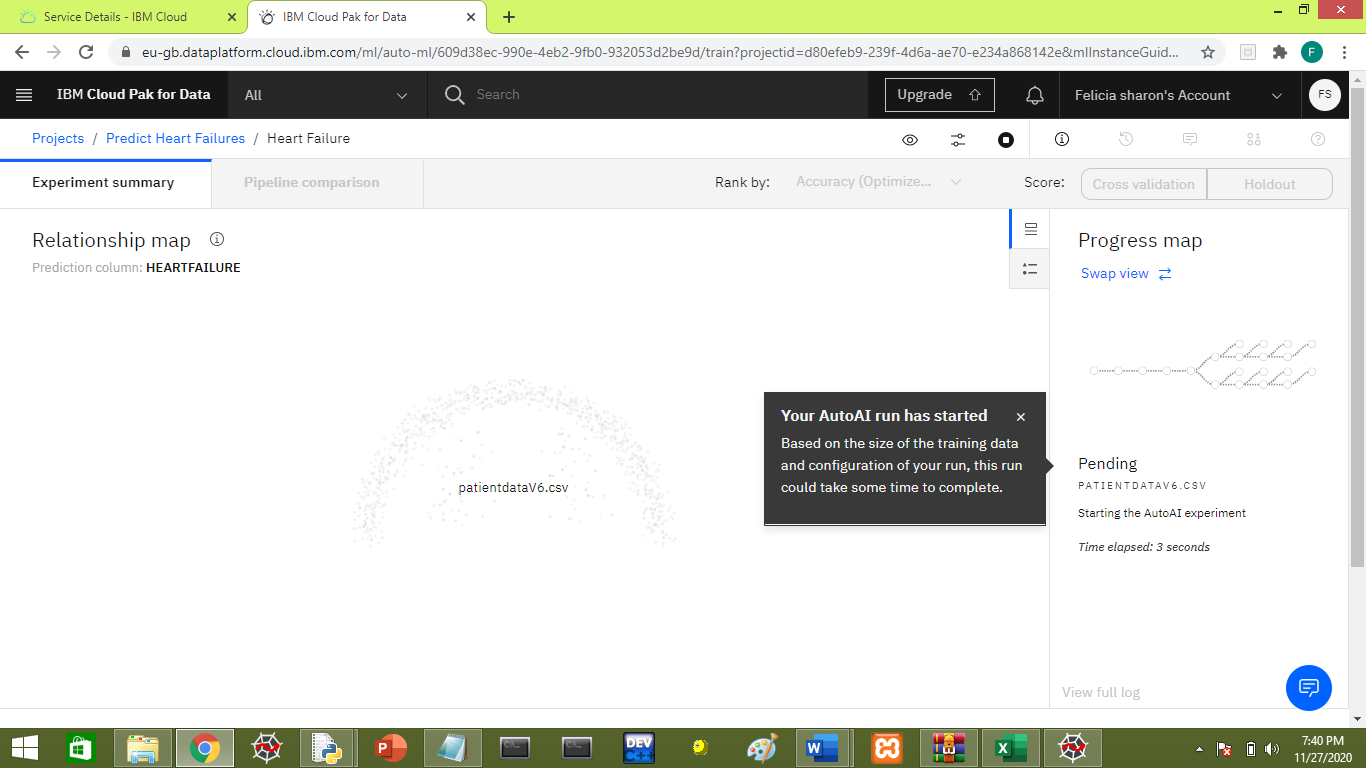
Select y as the column to predict. You can see that when you choose a column to predict, AutoAI selects a model type that matches the data. AutoAI analyzes your data and determines that the y column contains Yes/No information, making this data suitable for a binary classification model. Click Run experiment.

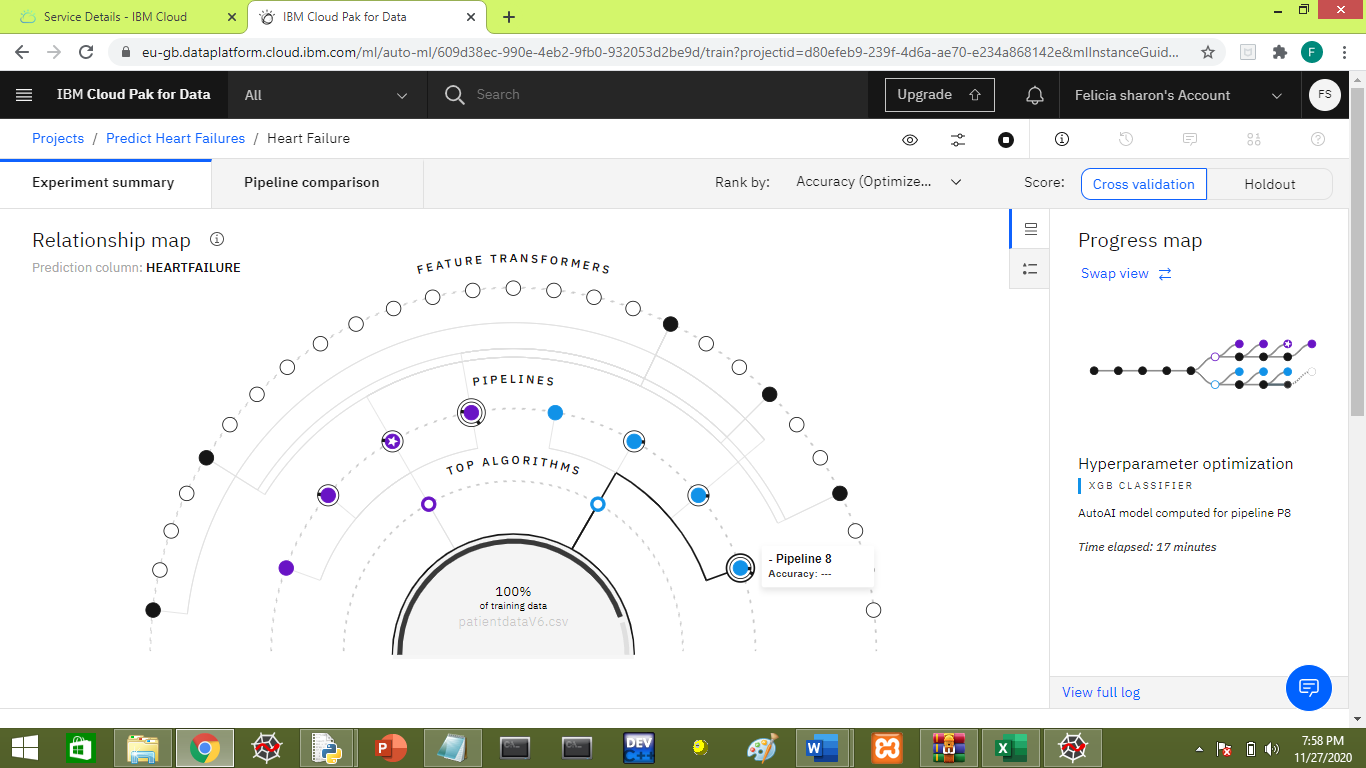




Review the pipelines

As AutoAI runs the experiment, it generate a set of pipelines, ranked according to how they perform with certain metrics or optimizations.

As the experiment runs, an infographic shows the algorithms applied to build each pipeline. Hover over nodes in the infographic to get information about the pipeline.

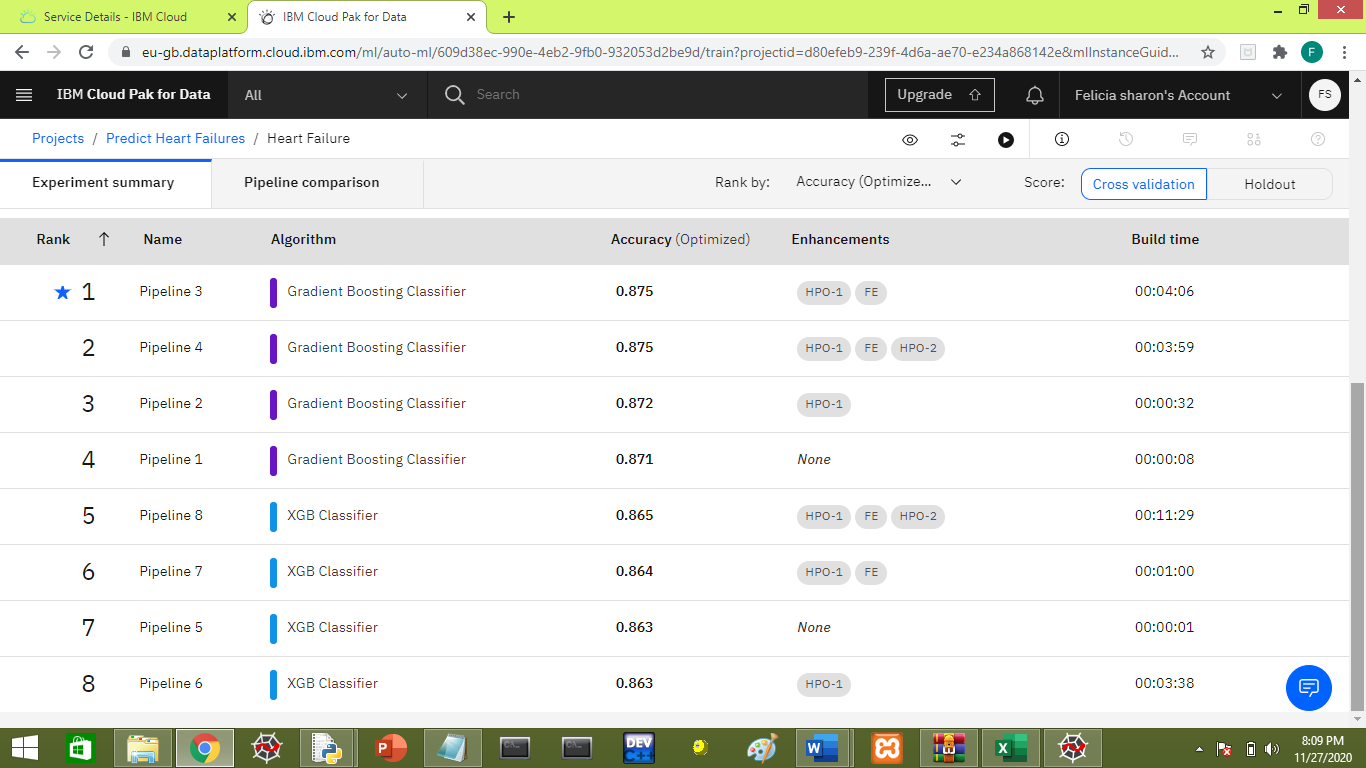


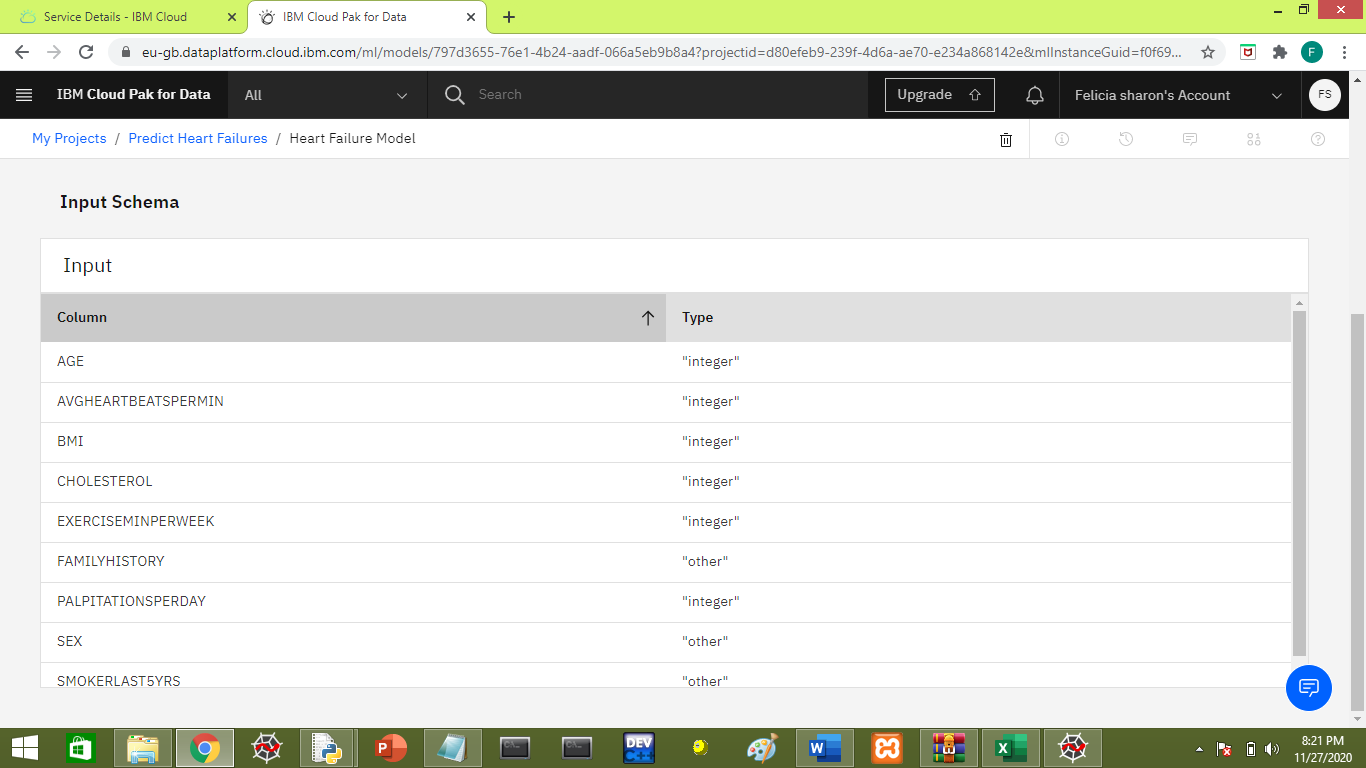
Save a pipeline

Choose one the best pipeline to save as a model.

Choose Save as model from the action menu for Pipeline 1. This saves the pipeline as a machine learning asset in your project.

Click View in project from the notification to open the project and see the details for the saved model.





Deploy the trained model

From the model details page:

Click Promote to deployment space. If there is no deployment space associated with your project, you are prompted to create one now. Follow the steps to create the space, then promote the model again.

After you promote the model, click go to deployment space from the notification.

From the Assets page:

Click the deployment icon deploy icon next to the model name.

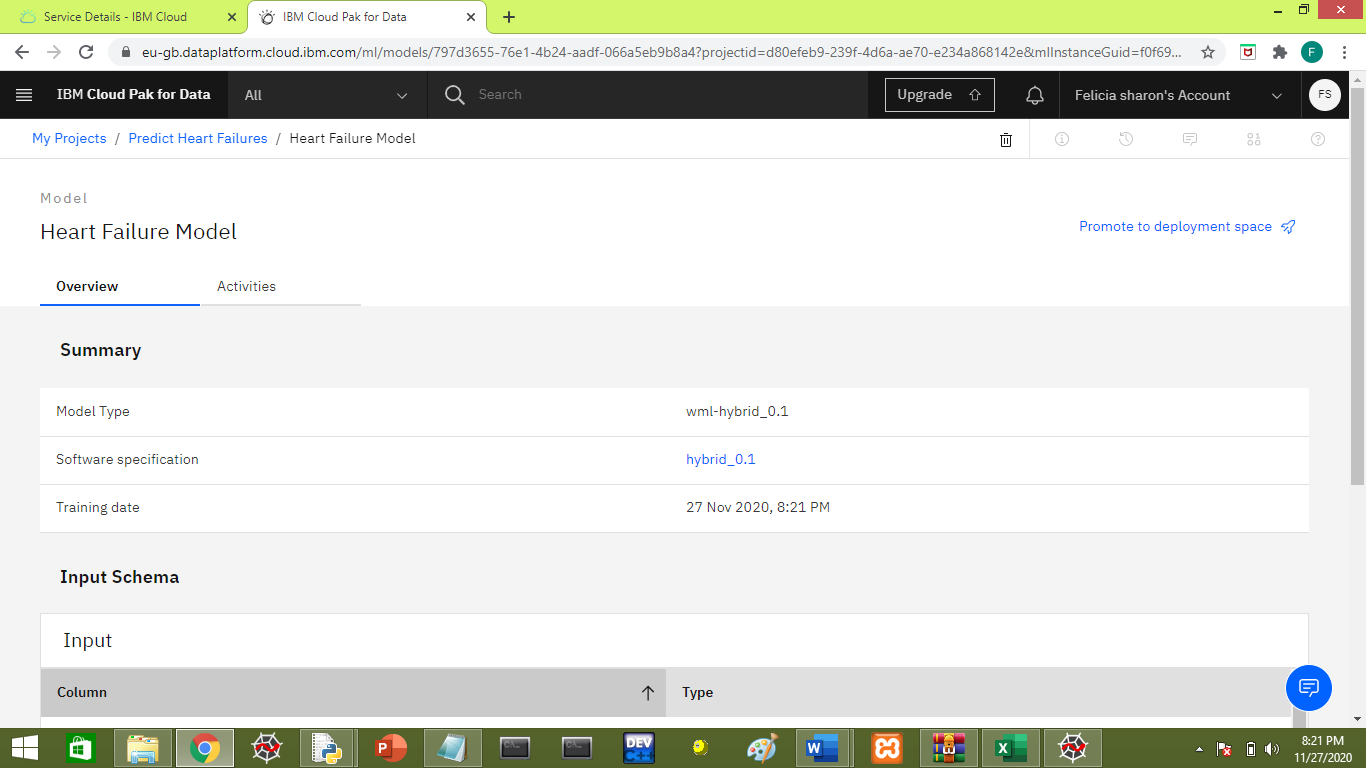
In the page that opens, fill in the fields:

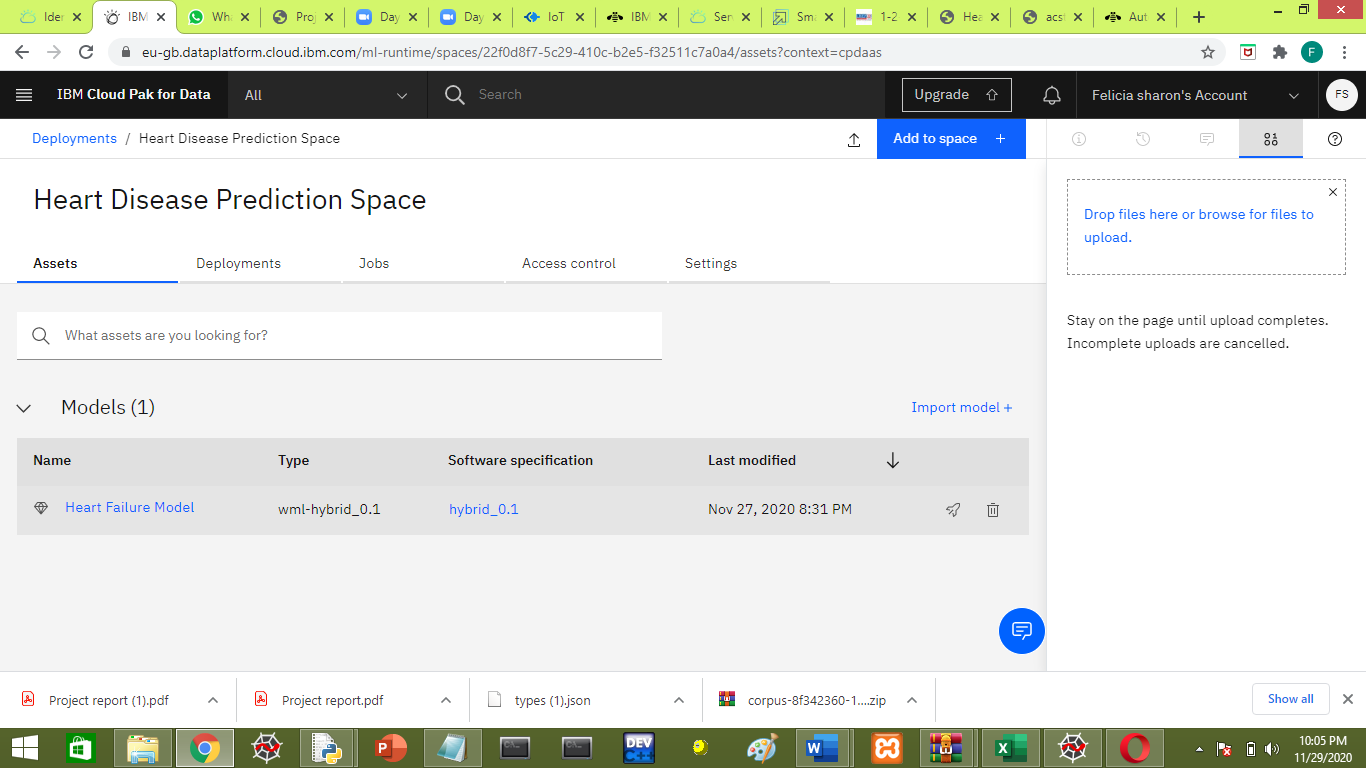
Specify a name for the deployment.

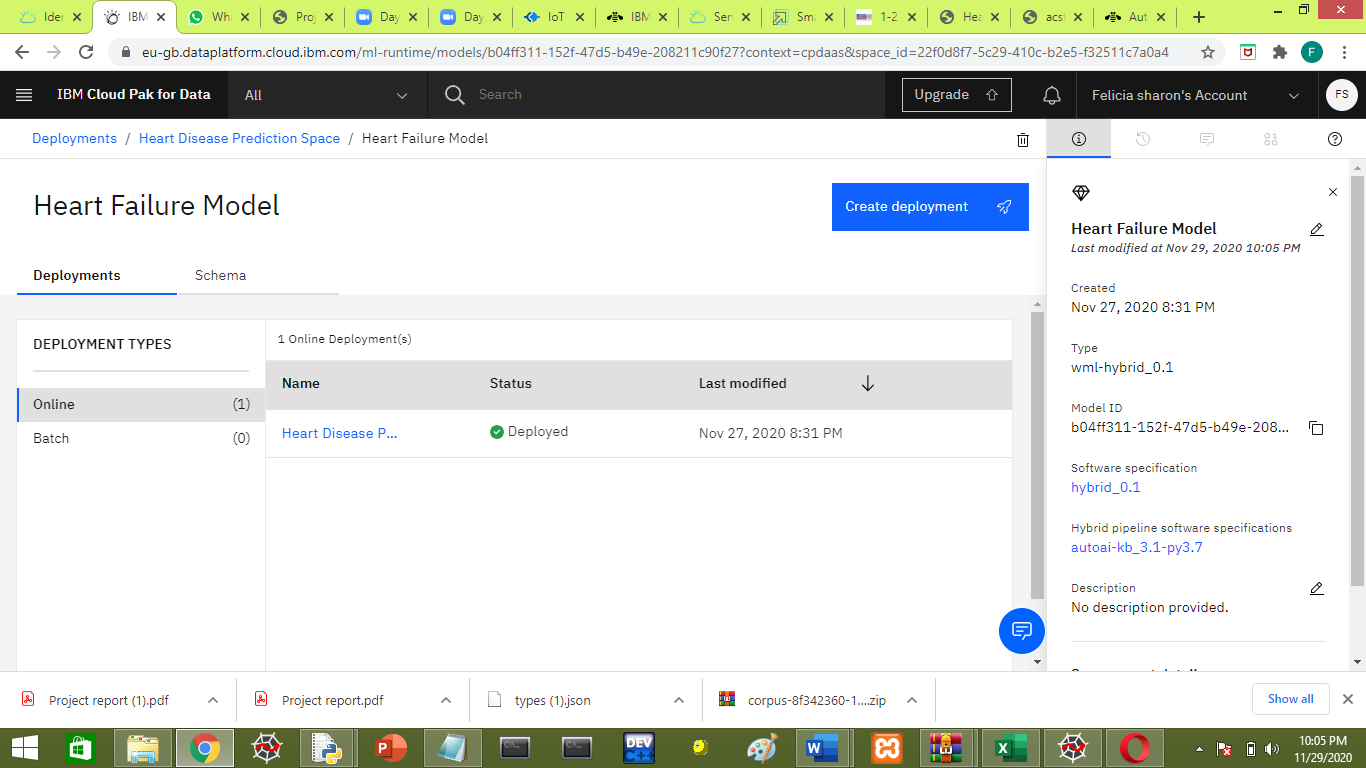
Select “Online” as the Deployment type.

Click Create.

When the deployment status changes to Deployed, click on the deployment name to view the deployment details page.



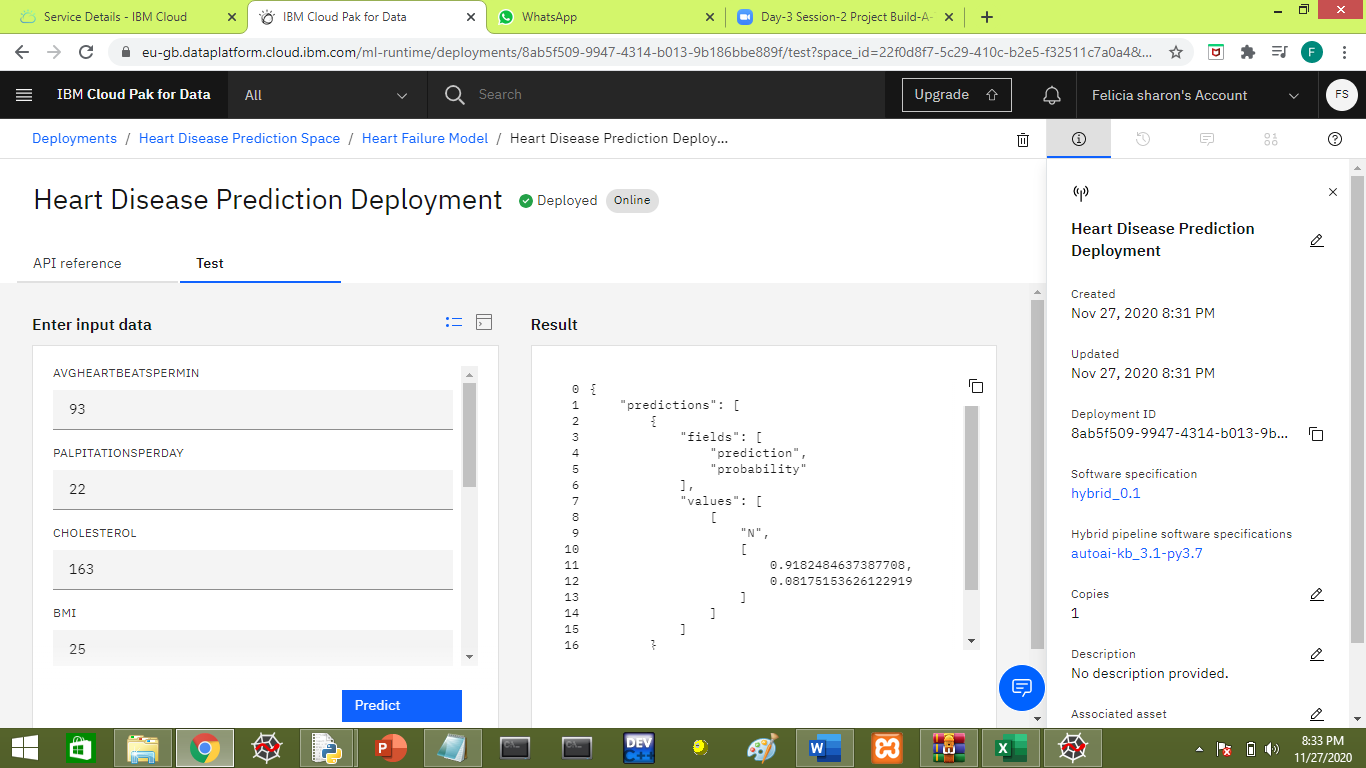




Test the deployed model

We can test the deployed model from the deployment details page.

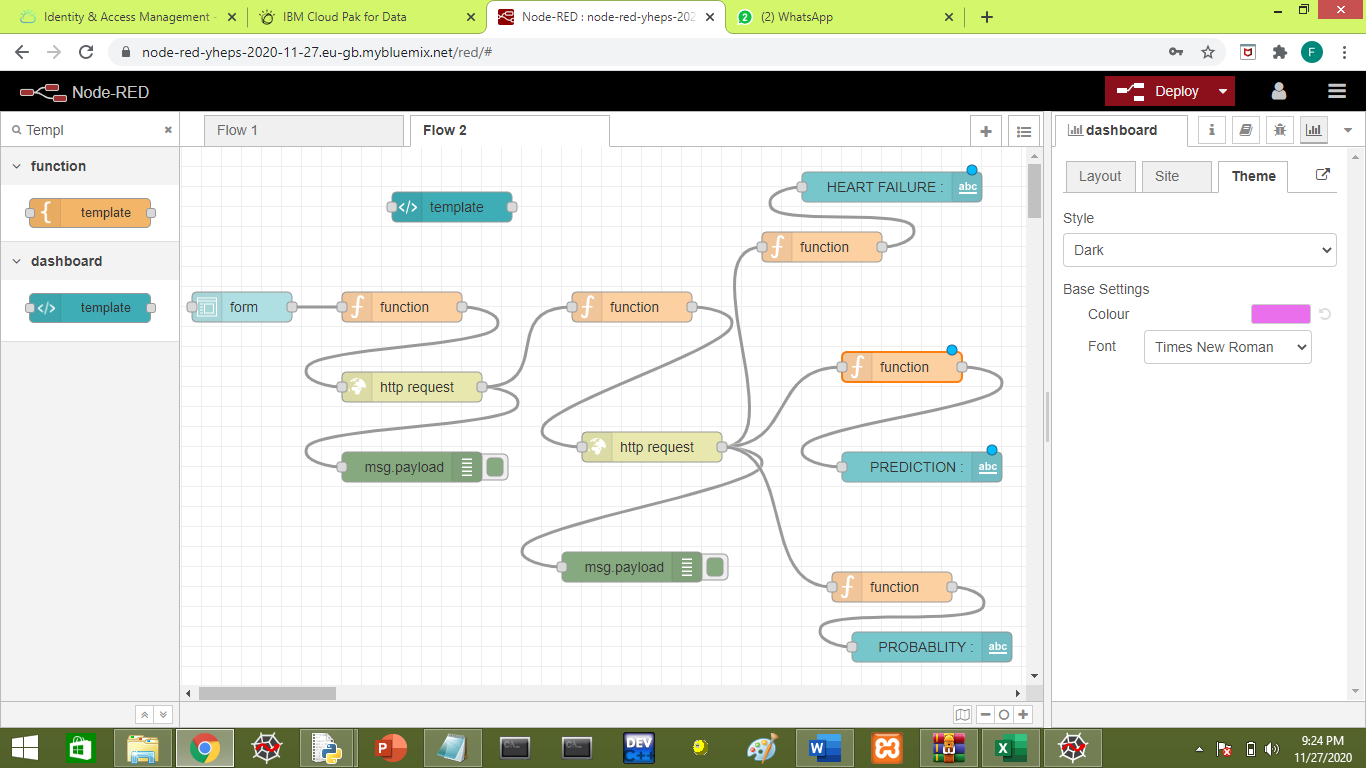
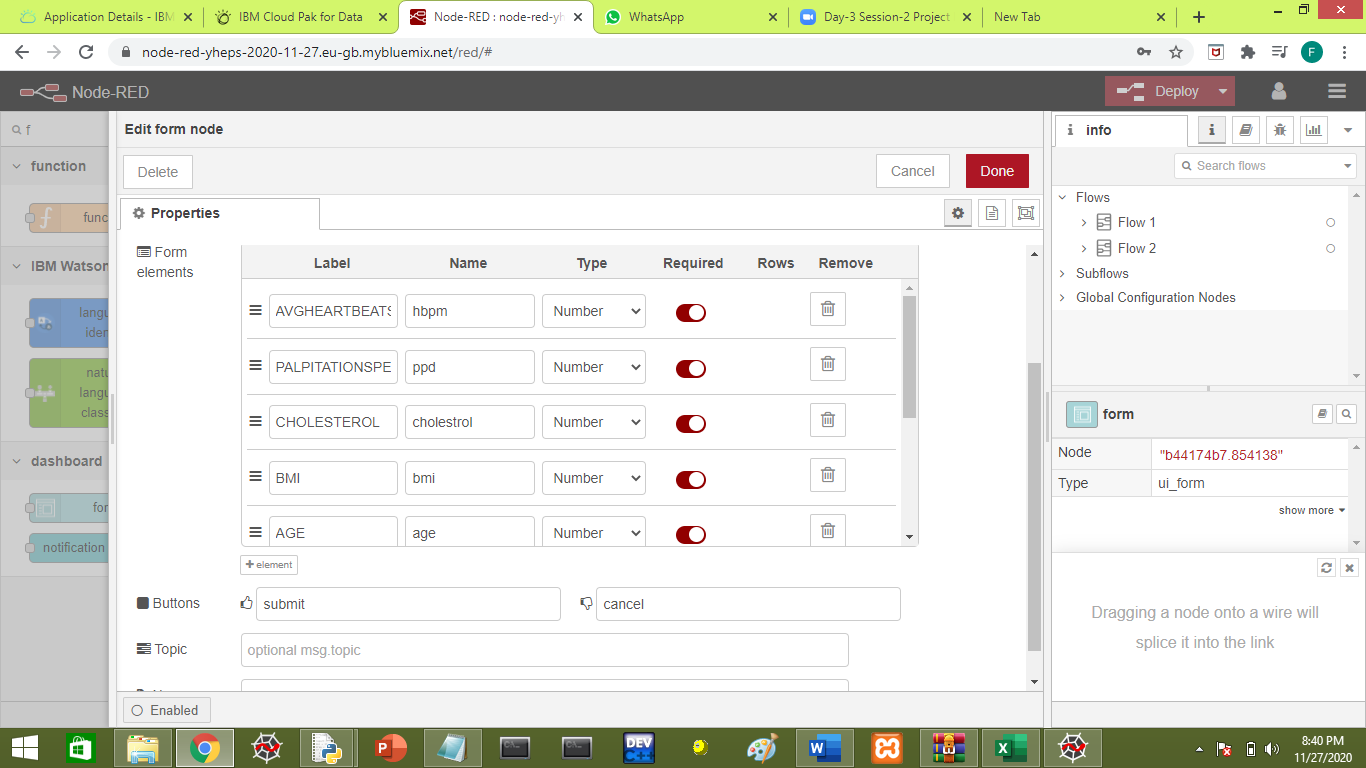
On the Test tab of the deployment details page, fill out the form with test values and click Predict to generate a prediction. For example:



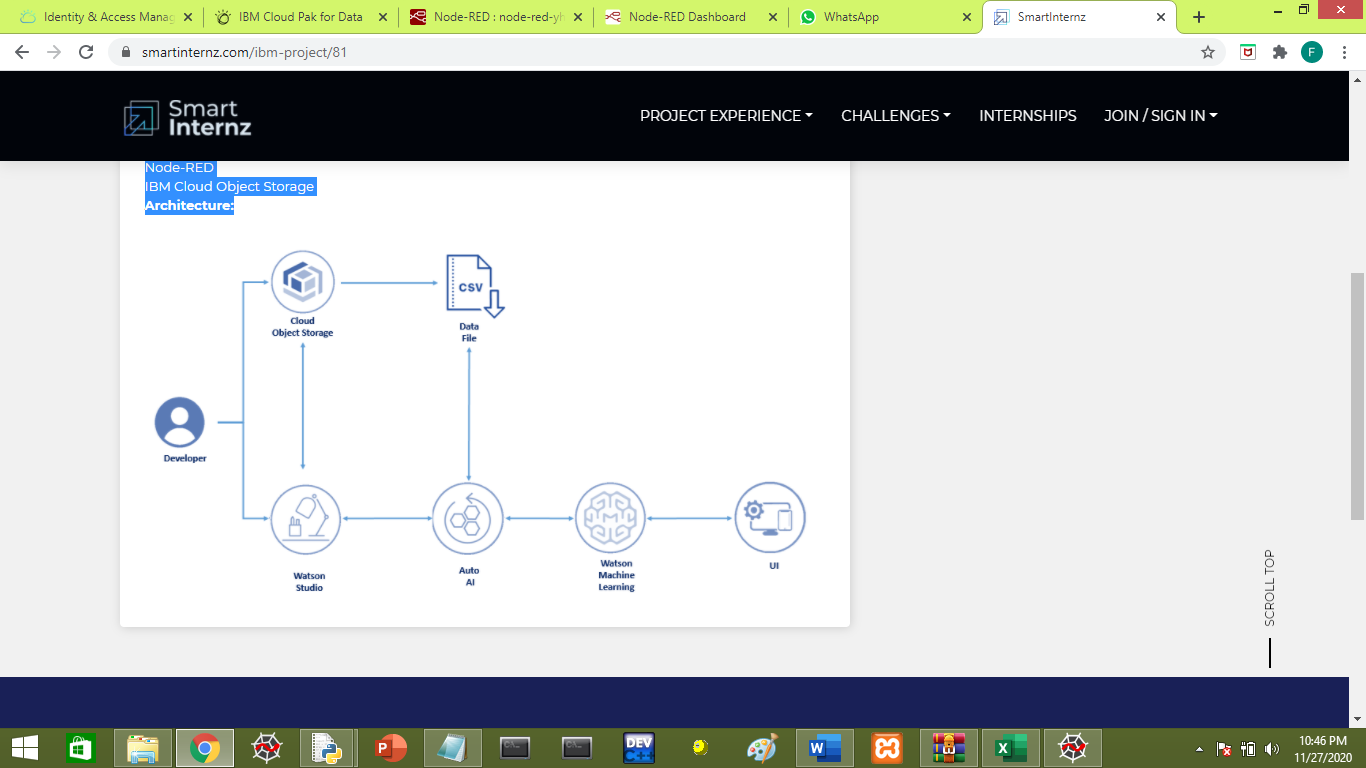
**Front End**

Create a UI for the project to take in inputs and display the respective output.

Node-Red service is used to create the UI. First a form is created to get all the necessary inputs from the user to predict Heart Failure then all the variables – all the 9 features are made global and the apikey is accessed. http request node is connected to the function node then a second function node is created to store the values given by the user and use those values in the model and predict the output. The access token is also retrieved. http request node is connected to the function node to connect the trained model and Node-Red. Then a 3 more function nodes are connected to retrieve only the particular results and each of them are connected to different text nodes to display the results.



**5. FLOWCHART**



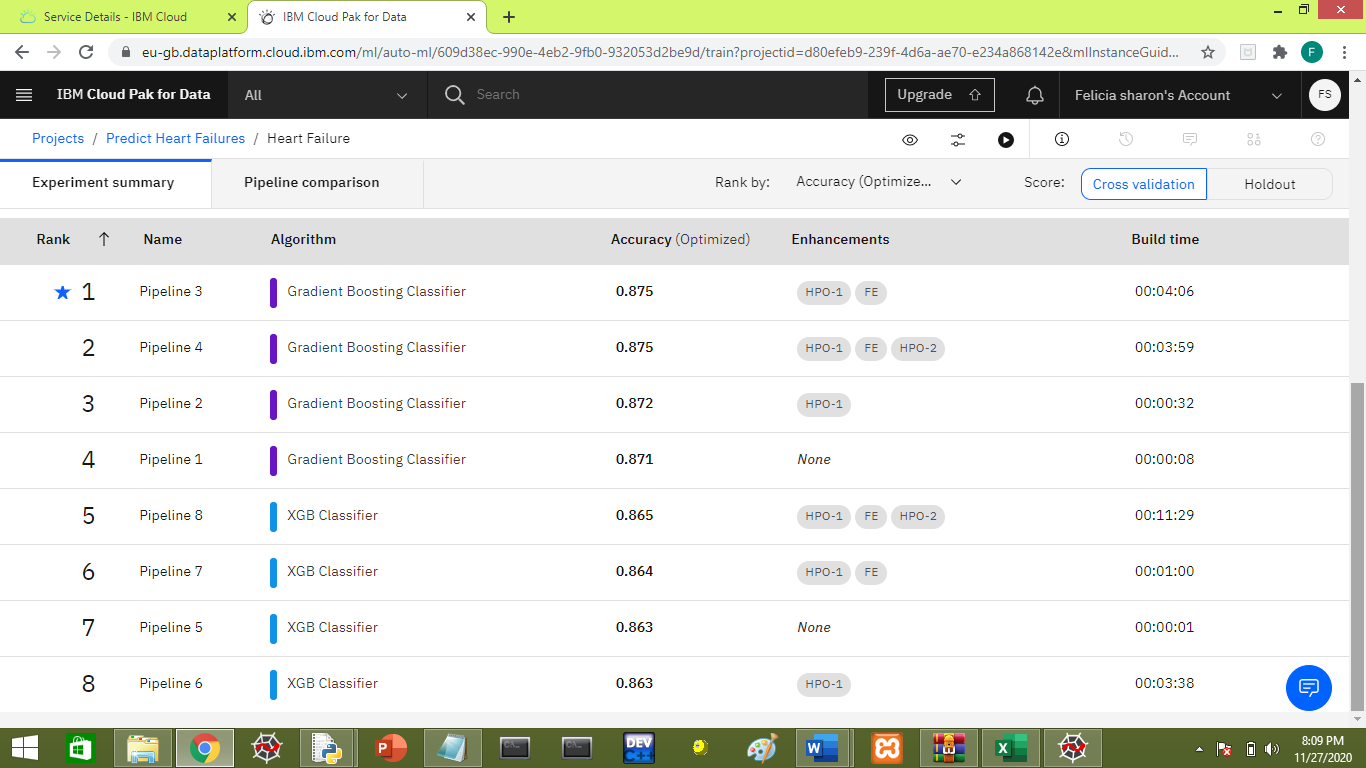
**6. RESULT**

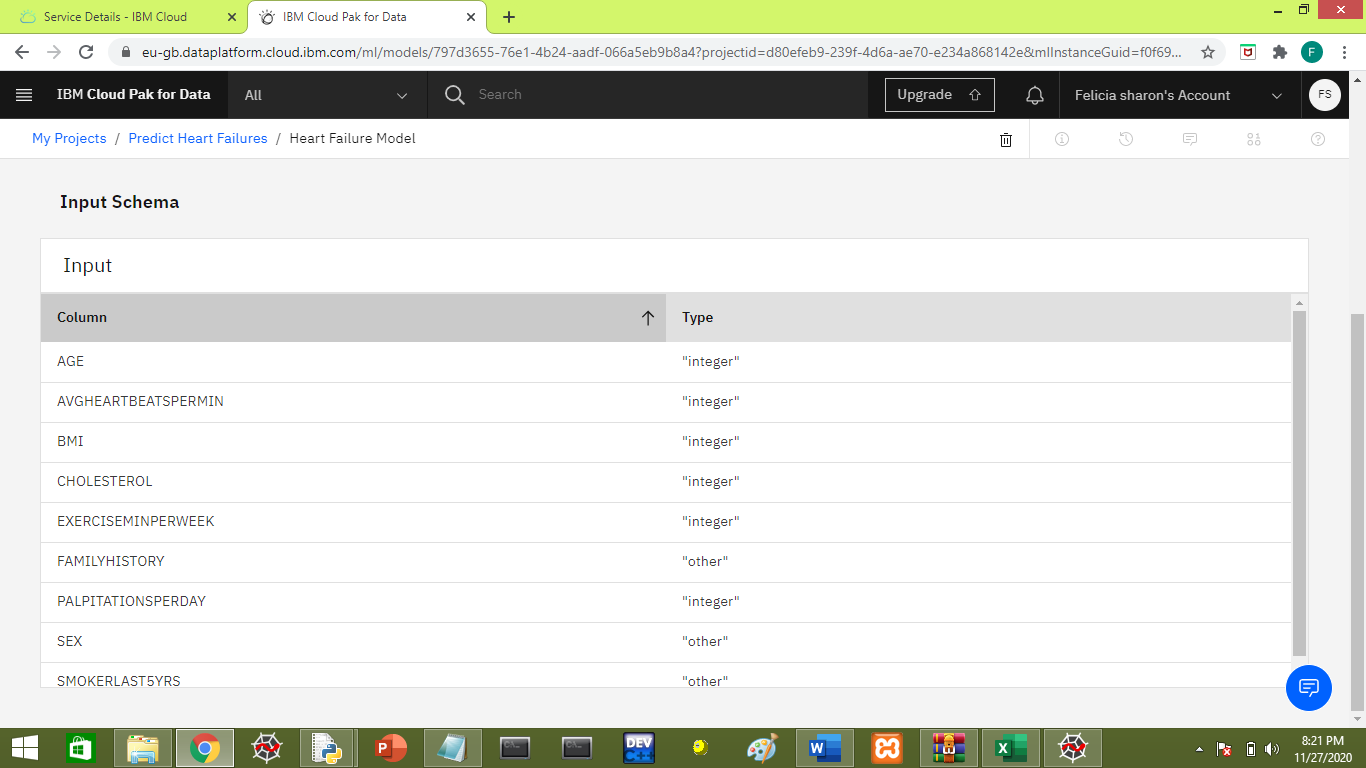
*Prediction type* – Binary Classification

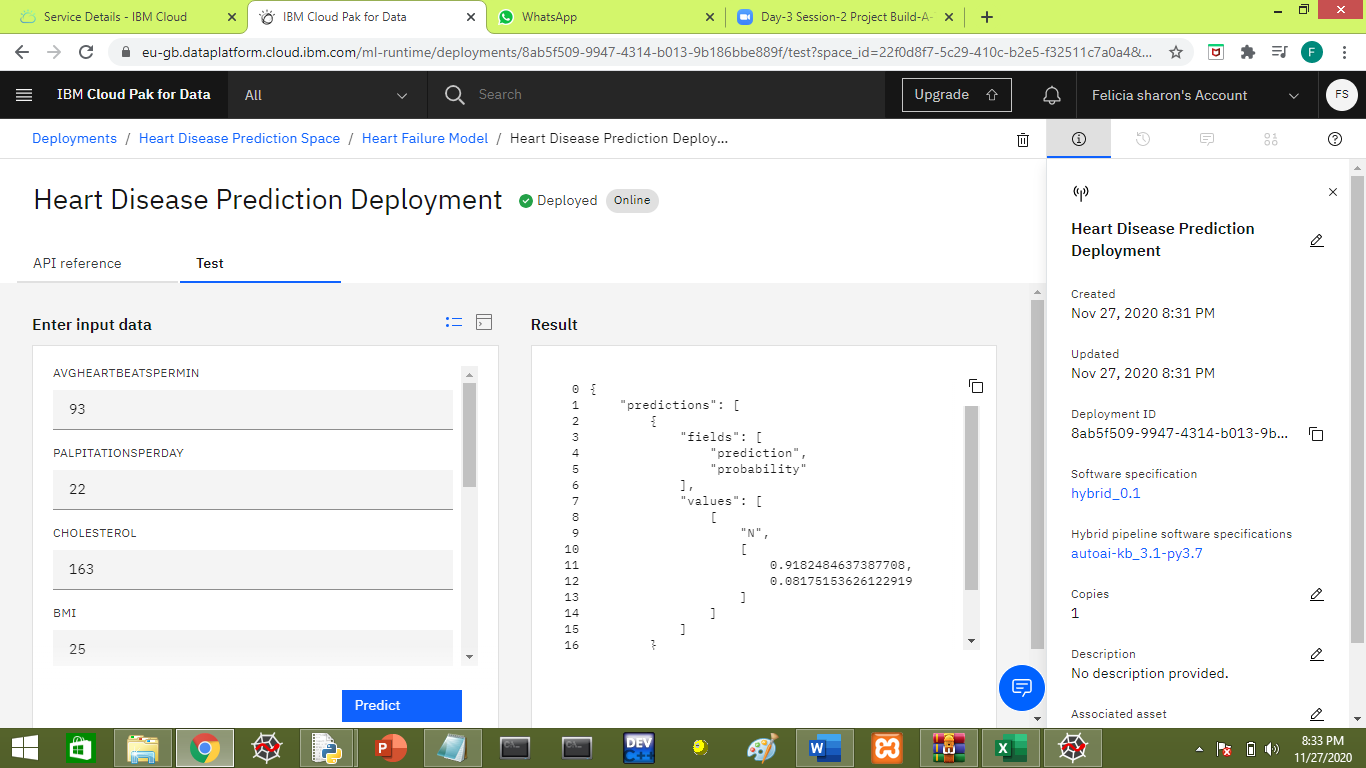
*Optimized metric* – Accuracy

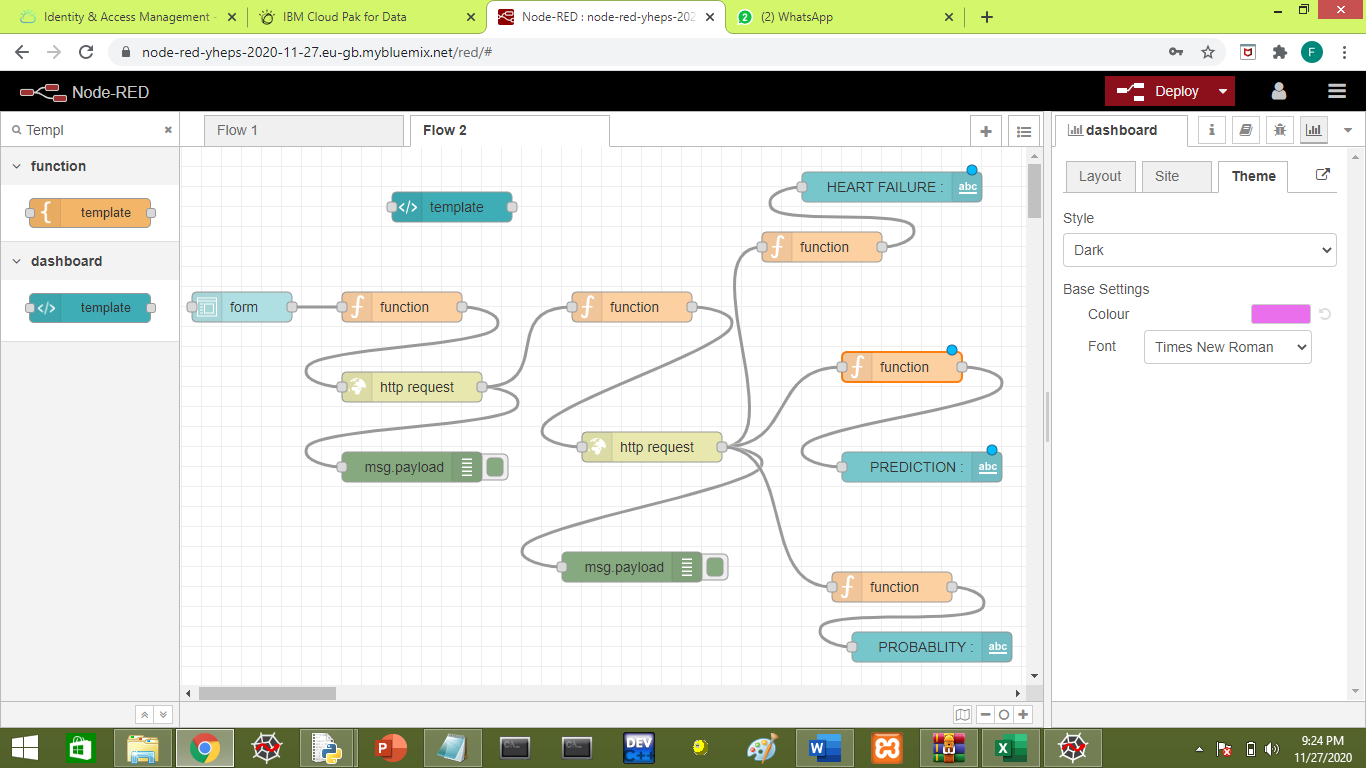
*Algorithms:*

* Gradient Boosting Classifier - 87.5% accuracy
* XGB Classifier - 86.5% accuracy.

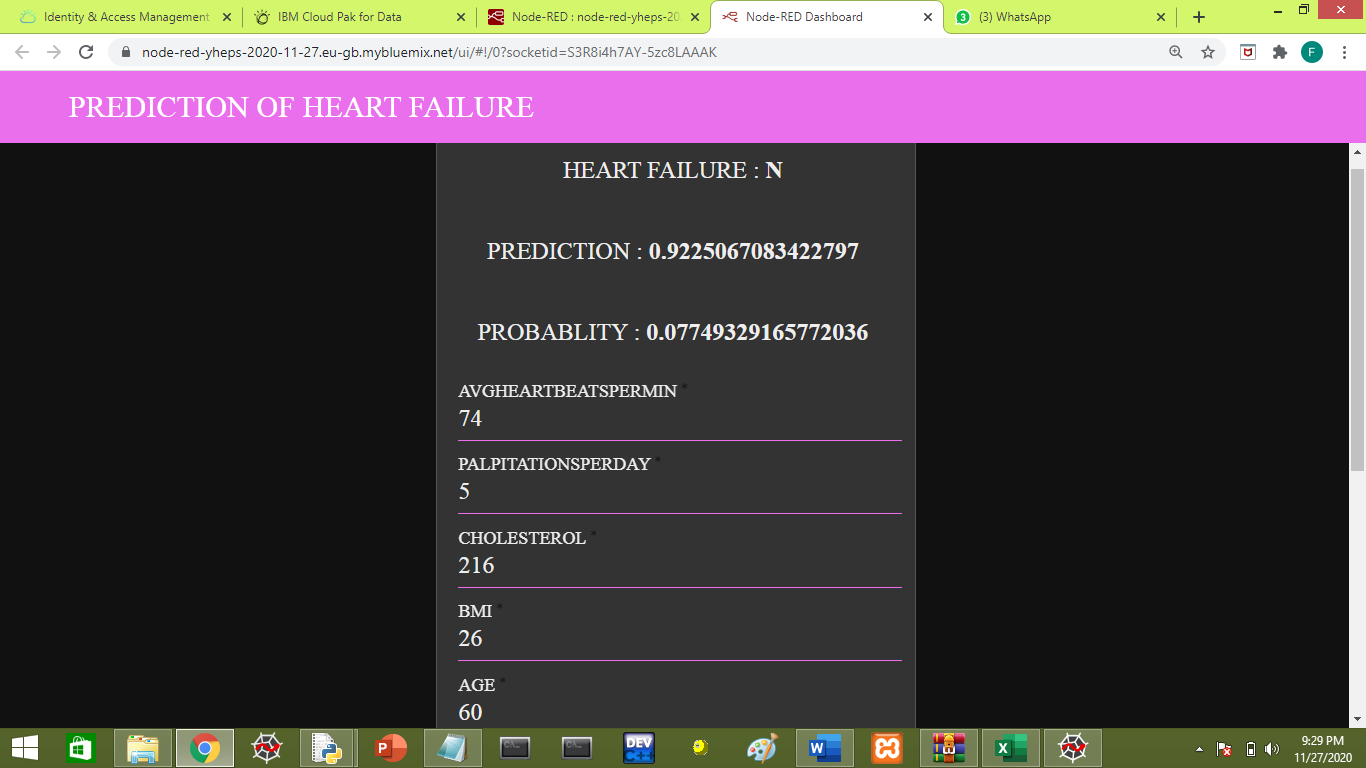


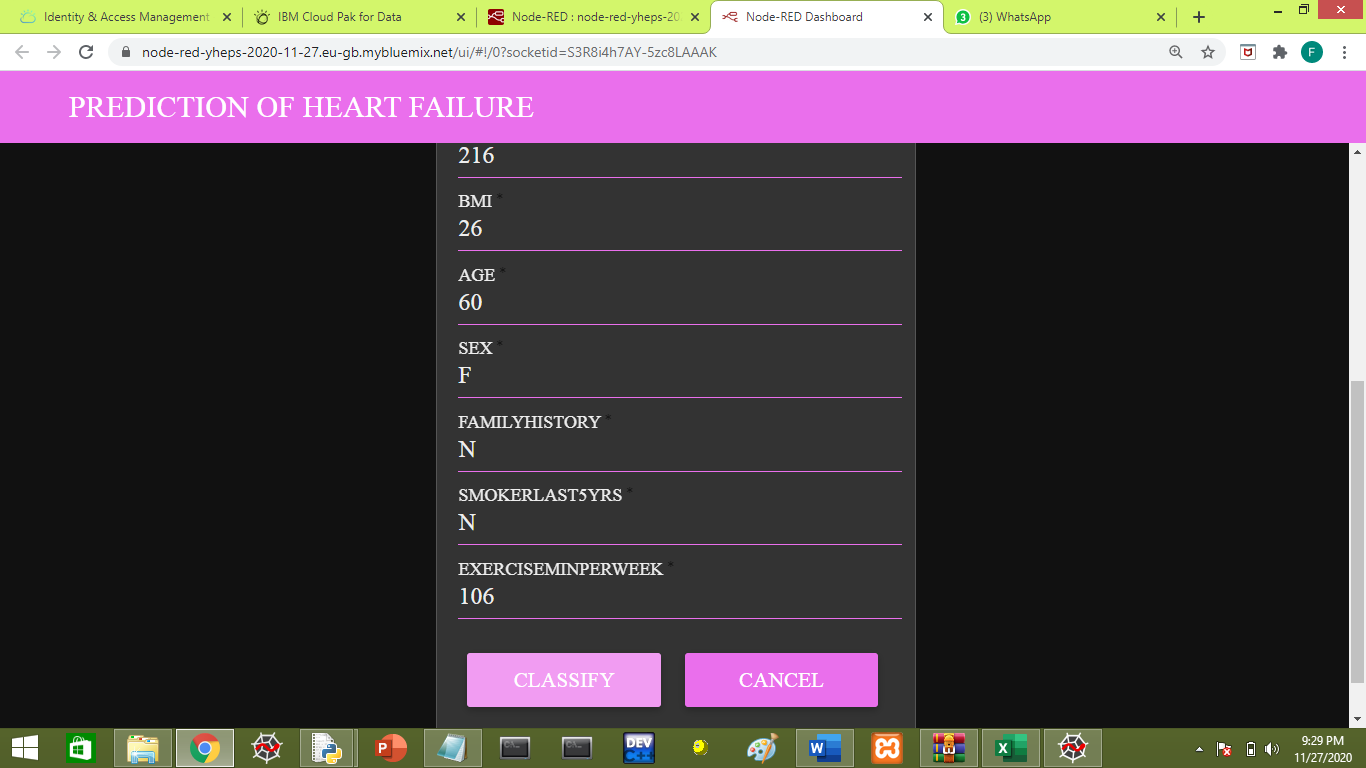




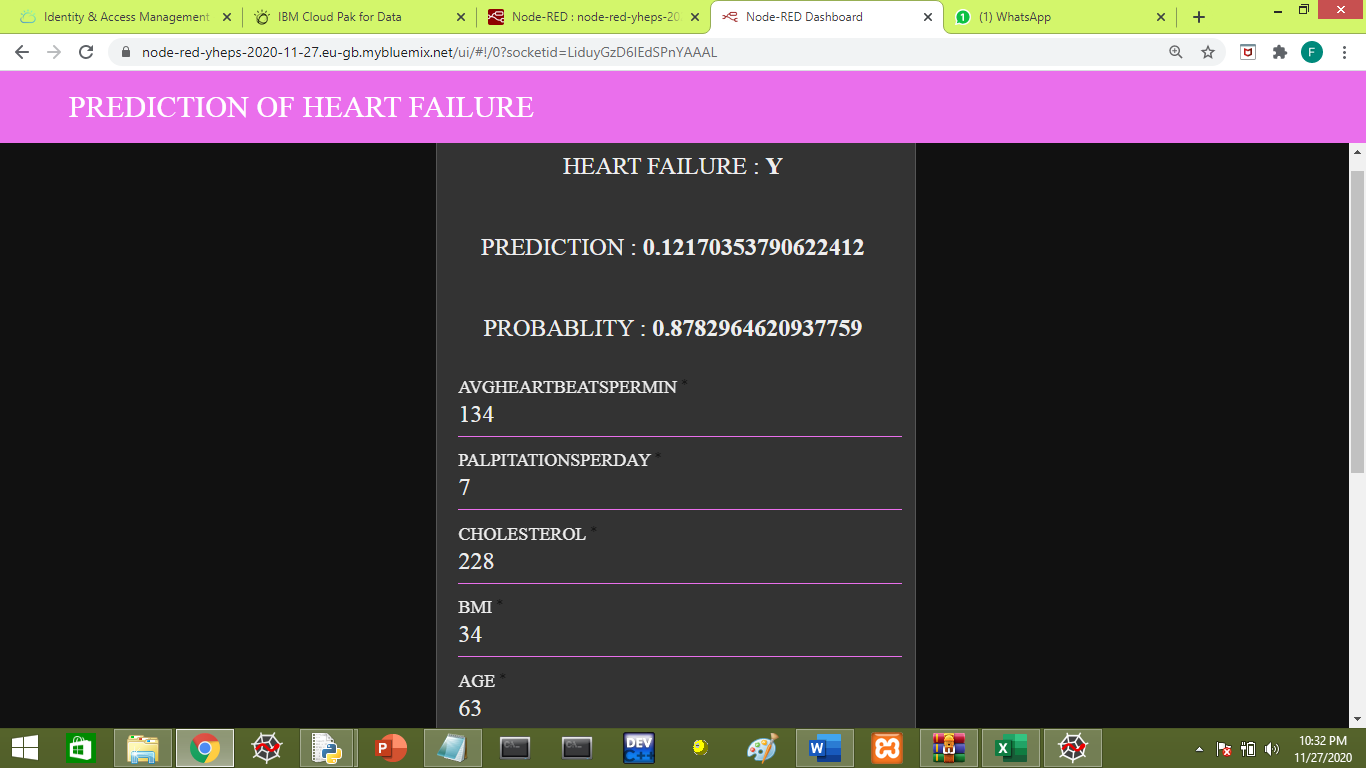


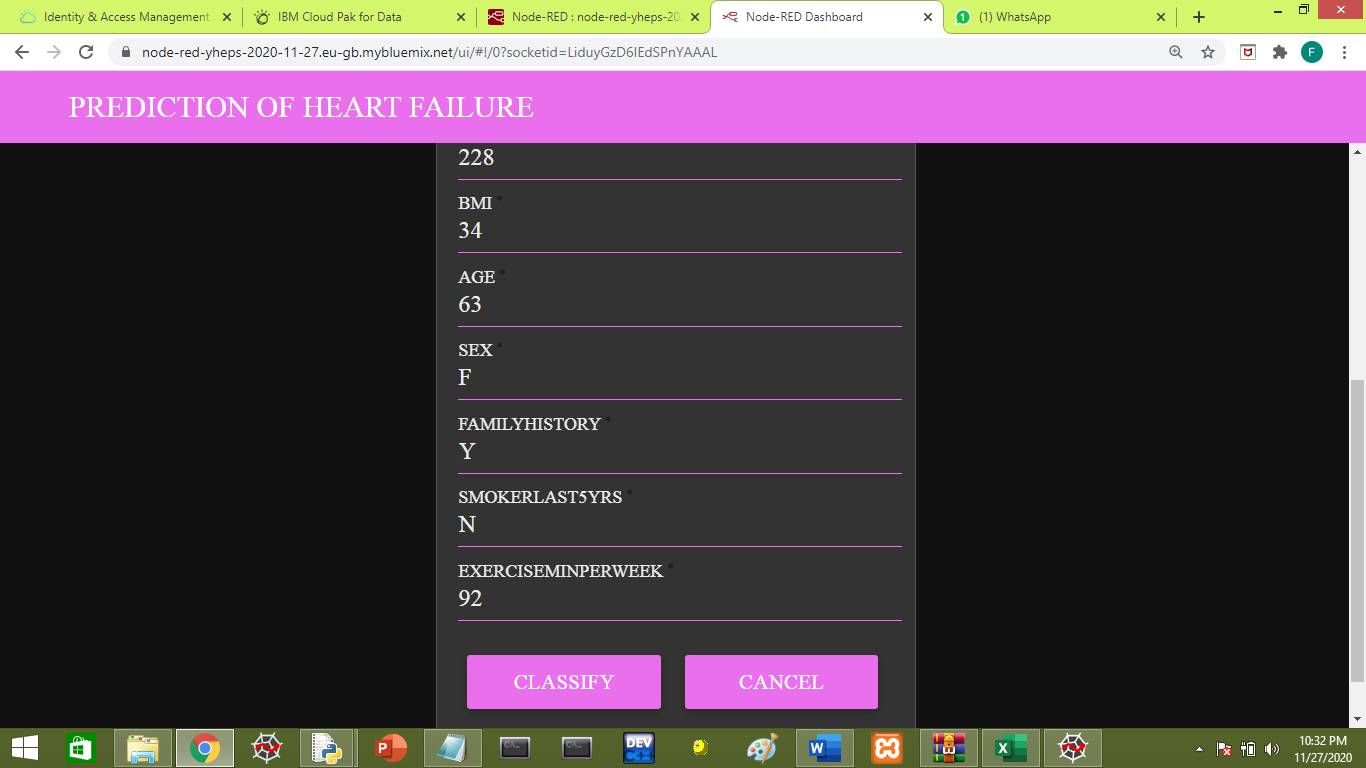
OUTPUT-1



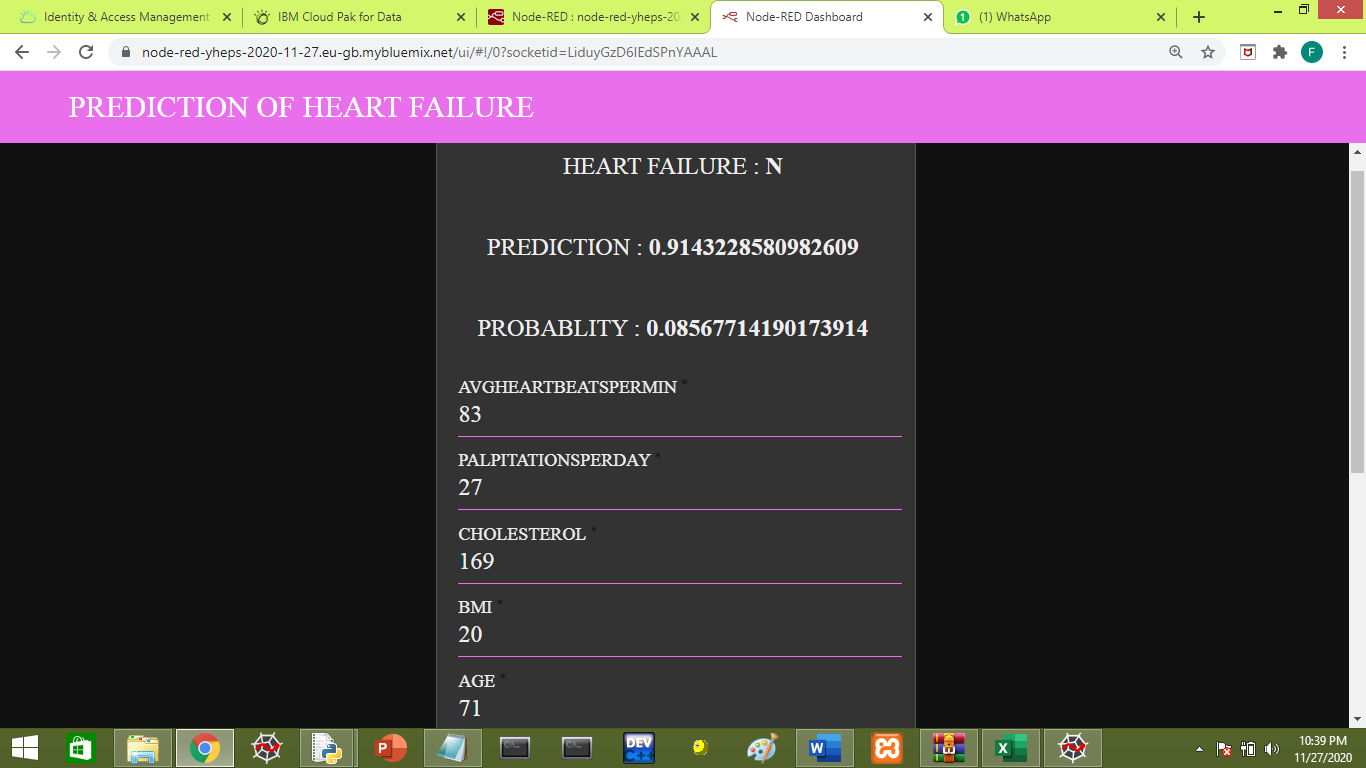


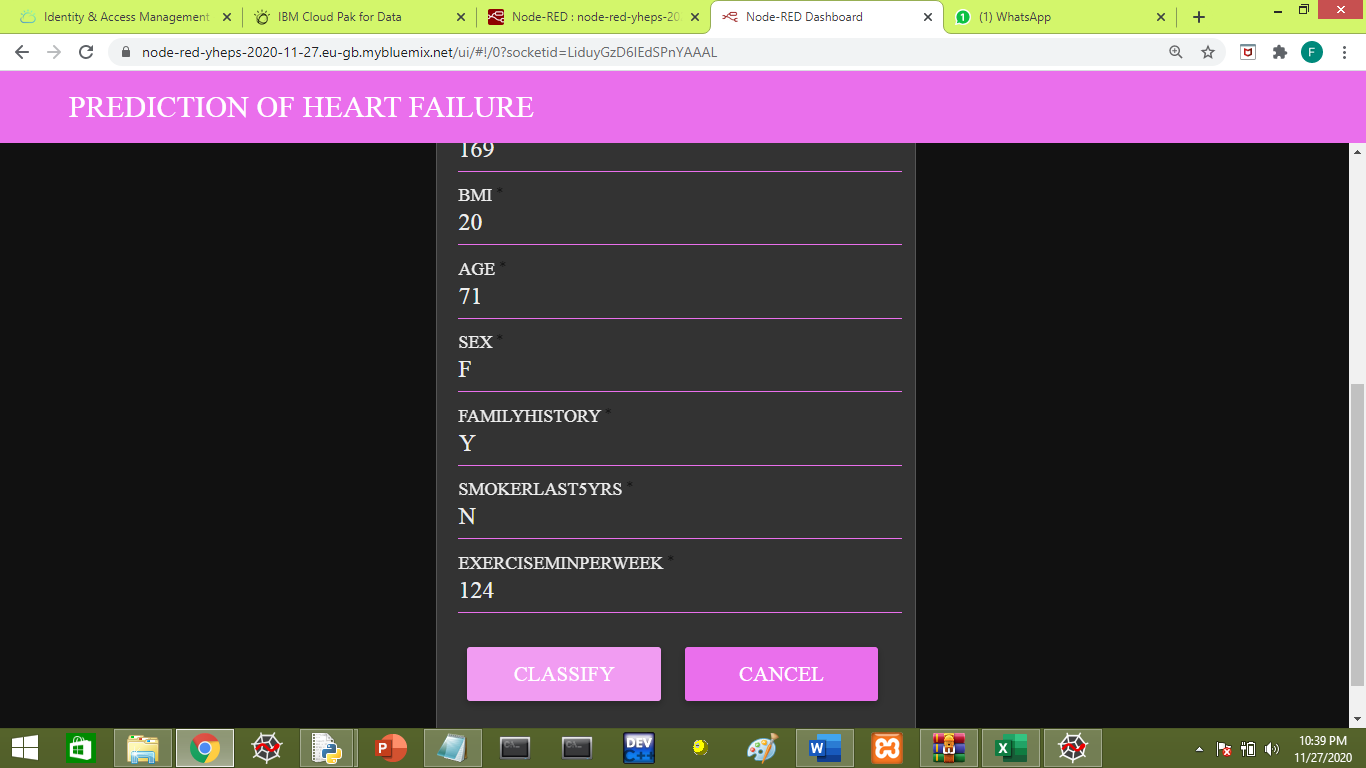
OUTPUT-2



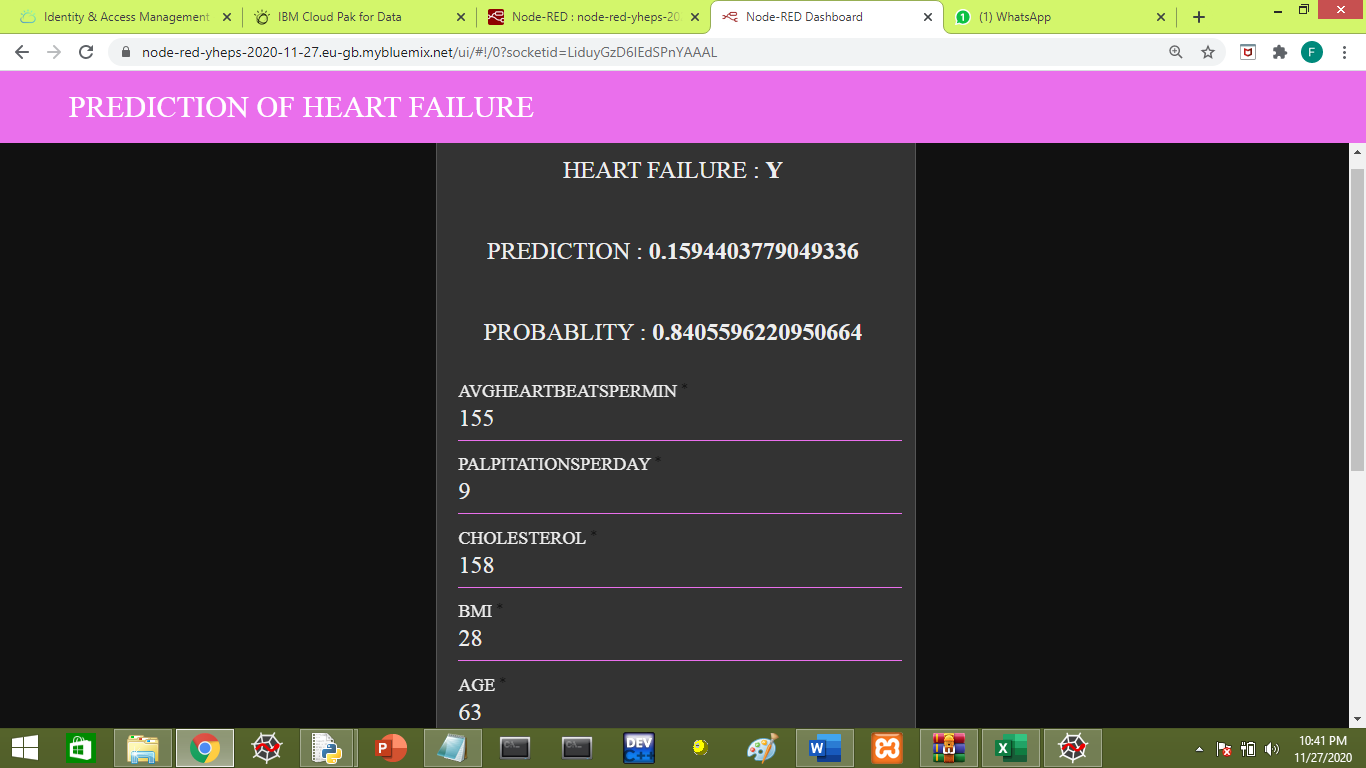


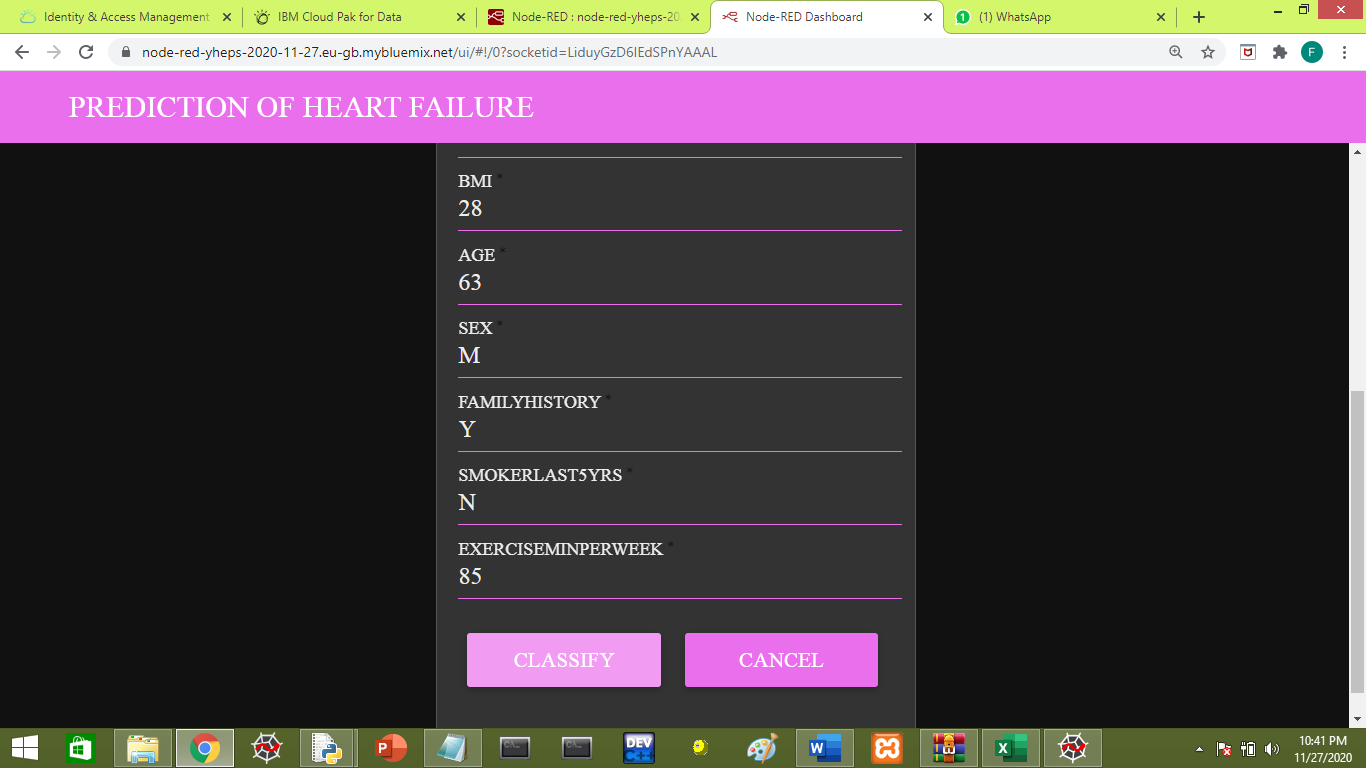
OUTPUT-3





OUTPUT-4





**7. APPLICATIONS**

* It helps the doctors to take proper and better decisions about the risk of heart disease.
* It brings together medicinal decision support with computer-based results of patient data, which helps to reduce medicinal faults and brings improvement in the overall outcome.
* It can be integrated with wearable devices and fitness mobile applications.

**8. ADVANTAGES**

* Early prediction of heart diseases can be done.
* The cost of medication will be minimized.
* Accuracy rate will be high.
* Doctors get more clients online.
* Very useful in case of emergency.

**8. DISADVANTAGES**

Accuracy Issues: A computerized system alone does not ensure accuracy and the Data quality is a critical factor for the success.

The system is not fully automated, it needs data from user for full diagnosis.

**9. CONCLUSION**

Eventually, we are able to achieve all the proposed objectives of this project. We developed a web based application for Heart Disease Prediction System. It successfully predicts the presence of HD in patient with reasonable accuracy.

Gradient Boosting Classifier and XGB Classifier algorithms are explored and deployed for predicting heart failure. Gradient Boosting Classifier provides 87.5% accuracy while XGB Classifier provided 86.5% accuracy.

**10. FUTURE SCOPE**

In future work, we can enhance the project by appending more detail prediction of Heart Failure in patients and improve the accuracy. Future enhancement of the HDPS is to predict a specific HD type such Heart attacks, CVD, CAD, etc. This system can be well integrated with smart wearable devices, mobile app, ambulance services etc. to save the life of the patient at the emergency condition.

**11. BIBLIOGRAPHY**

Dataset for Heart Failure

* <https://github.com/IBM/predictive-model-on-watson-ml/blob/master/data/patientdataV6.csv>

Links

* <https://us02web.zoom.us/rec/share/zBkZwpZCIS6JeYSII-R5kuUpeiyYW3Jqt8YxDFMs3fa_4b-SdwgeN1ZPyPn-BsA.lmxsj0BiZN8xJf2Q>
* <https://us02web.zoom.us/rec/share/tFuWBacLCYftWRvst9YJkgxKEqPl9WWOwT4Lb7QNO3xn1wCVa2hfQJ1NBtXTF81q.1dxlL3FiCM3M_SA_>
* <https://www.ijcaonline.org/archives/volume176/number23/singh-2020-ijca-920198.pdf>
* <https://www.lbef.org/journal/1-2/download/1-2-115-132.pdf>
* <https://www.ripublication.com/acst17/acstv10n7_13.pdf>

**APPENDIX**

Source code

**Node-Red**

Function-1

global.set("hbpm",msg.payload.hbpm)

global.set("ppd",msg.payload.ppd)

global.set("cholestrol",msg.payload.cholestrol)

global.set("bmi",msg.payload.bmi)

global.set("age",msg.payload.age)

global.set("sex",msg.payload.sex)

global.set("famhist",msg.payload.famhist)

global.set("smoker",msg.payload.smoker)

global.set("excercise",msg.payload.excercise)

var apikey="2XHoBrpbgoijxgQjTTaUs\_9yDM1RJMCaiiyz5WnYByqO"

msg.headers = {"content-type":"application/x-www-form-urlencoded"}

msg.payload={"grant\_type":"urn:ibm:params:oauth:grant-type:apikey","apikey":apikey}

return msg;

Function-2

var hbpm=global.get("hbpm")

var ppd=global.get("ppd")

var cholestrol=global.get("cholestrol")

var bmi=global.get("bmi")

var age=global.get("age")

var sex=global.get("sex")

var famhist=global.get("famhist")

var smoke=global.get("smoke")

var excercise=global.get("excercise")

var token=msg.payload.access\_token

msg.headers = {"Content-type":"application/json","Authorization":"Bearer "+token,"Accept":"application/json"}

msg.payload={"input\_data":[{"fields":["AVGHEARTBEATSPERMIN","PALPITATIONSPERDAY","CHOLESTEROL","BMI","AGE","SEX","FAMILYHISTORY","SMOKERLAST5YRS","EXERCISEMINPERWEEK"],"values":[[hbpm,ppd,cholestrol,bmi,age,sex,famhist,smoke,excercise]]}]}

return msg;

Function-3

msg.payload=msg.payload.predictions[0].values[0][0]

return msg;