DIABETIC MELLITUS PREDICTION USING IBM AUTOAI

Submitted By:

Deepanshi Wamankar

RSIP Career Basic ML 161

Index

1. INTRODUCTION

1.1 Overview:

In this project there are 4 tasks:

1. Data collection
a. Download Dataset/Create Dataset
2. IBM Cloud account
a. IBM Cloud Registration
b. Login To IBM Cloud
c. Create Cloud Object Storage Service
d. Create Watson Studio Platform
e. Create Machine Learning service

3. Model building
a. Create Project In Watson Studio
b. Auto AI Experiment In Add Experiment
c. Setup Your Auto AI Environment
d. Import Dataset
e. Run The Model
f. Selection of Auto-AI Pipeline
g. Deploy And Test The Model In Watson Studio
4. Application building
a. Create Node-Red Service
b. Build UI With Node-Red
c. Go to manage palette & install dashboard nodes

1.2 .2 Purpose:

This project prevents the people from the avalanche by priory informing them there is a chance to the occurrence of avalanche or not.In medicine, the diagnosis of diabetes is according to fasting blood glucose, glucose tolerance, and random blood glucose levels. The earlier diagnosis is obtained, the much easier we can control it. Machine learning can help people make a preliminary judgment about diabetes mellitus according to their daily physical examination data, and it can serve as a reference for doctors.

2. LITERATURE SURVEY

2.1 .1 Existing problem:

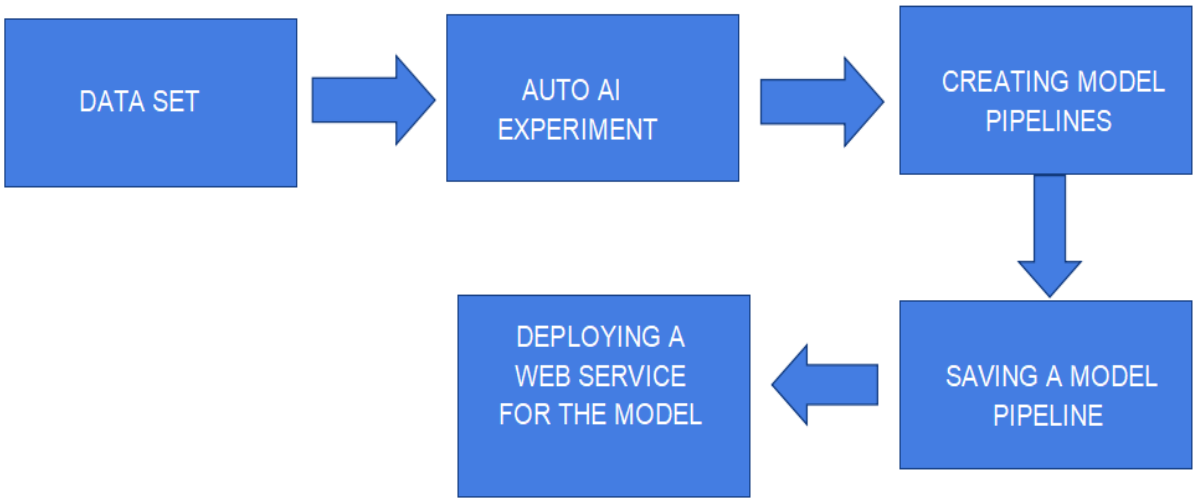
Diabetes mellitus or simply diabetes is a disease caused due to the increase level of blood glucose. Various traditional methods, based on physical and chemical tests, are available for diagnosing diabetes. Early prediction of diabetes is quite challenging task for medical practitioners due to complex interdependence on various factors as diabetes affects human organs such as kidney, eye, heart, nerves, foot etc

2.2 .2 Proposed solution:

This project prevents the people from the avalanche by priory informing them there is a chance to the occurrence of avalanche or not.The aim of this project is to develop a system which can perform early prediction of diabetes for a patient with a higher accuracy by combining the results of different machine learning techniques and the model is been built in Auto AI.

3. THEORITICAL ANALYSIS

3.1 ,1 Block diagram:



3.2 .2 Sofware designing:

IBM Cloud Computing:

It is a set of cloud computing services for business offered by the information technology company IBM. IBM Cloud includes infrastructure as a service (IaaS), software as a service (SaaS) and platform as a service (PaaS) offered through public, private and hybrid cloud delivery models, in addition to the components that make up those clouds.

IBM Watson Studio:

IBM Watson Studio helps data science scientists and analysts prepare data and build models at scale across any cloud. With its open, flexible multi cloud architecture, Watson Studio provides capabilities that empower businesses to simplify enterprise

data science and AI, such as:

1. Automate AI lifecycle management with Auto AI.
2. Visually prepare and build models with IBM SPSS Modeler.
3. Build models using images with IBM Watson Visual Recogni on and texts with IBM Watson Natural Language Classifier.
4. Deploy and run models through one-click integration with IBM Watson Machine Learning.
5. Manage and monitor models through integration with IBM Watson Open scale.

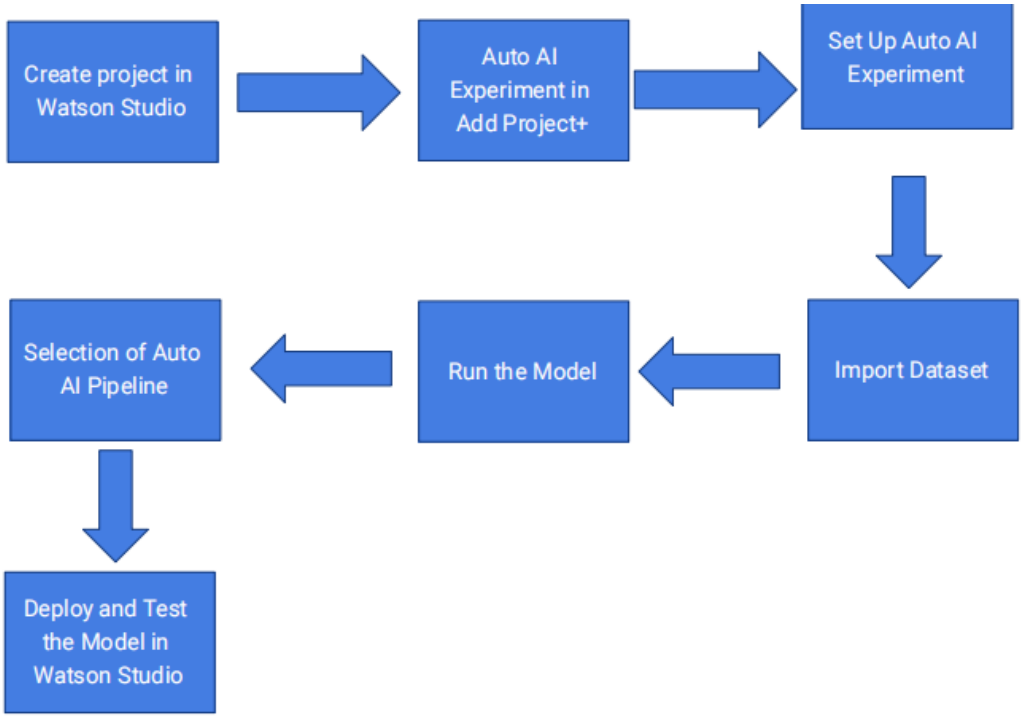
Node Red:

Node-RED is a flow-based development tool for visual programming developed originally by IBM for wiring together hardware devices, APIs and online services as part of the Internet of Things.

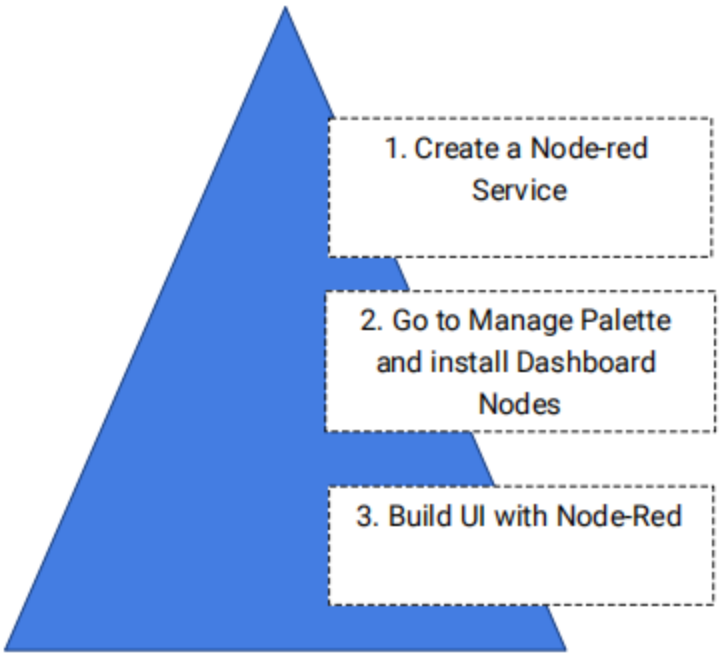
Node-RED provides a web browser-based flow editor, which can be used to create JavaScript functions. Elements of applications can be saved or shared for re-use. The flows created in Node-RED are stored using JSON. Since version 0.14, MQTT nodes can make properly configured TLS connections. In 2016, IBM contributed Node-RED as an open source JS Foundation project.

4. EXPERIMENTAL INVESTIGATIONS

Auto AI Experiment Workflow:



Application Building Workflow



AUTO AI EXPERIMENT

1. In Machine Learning, we create a project in IBM Watson Studio and we add an Auto AI Experiment to the project.

2. We need to import our dataset to calculate the required output. Hence, we import/upload the 'datasets_26789_34175_pima-indians-diabetes.data' csv file and give 'class' column of the dataset as the prediction column.

3. Run the model and we get our required Auto AI Pipeline. We get the Relationship Map and Pipeline Comparison. We have to save the Auto AI pipeline as a model.

4. We should deploy the saved model.
5. Then we get to test the deployed model.

APPLICATION BUILDING (NODE-RED)

1. From the catalog of IBM Cloud we should install Node-Red App.

2. In this, we have ‘Nodes’ to the left side of the page and we use these nodes to create a ‘flow’.

3. We need to install ‘Dashboard Nodes’ by going to the ‘Manage Palette’ section. ●

The nodes which we use in this project are:

Form node: Adds a form to user interface. Helps to collect multiple value from the user on submit button click as an object in msg.payload.

Multiple input elements can be added using add elements button.

Timestamp node: Injects a message into a flow either manually or at regular intervals.

The message payload can be a variety of types, including strings, JavaScript objects or the current time.

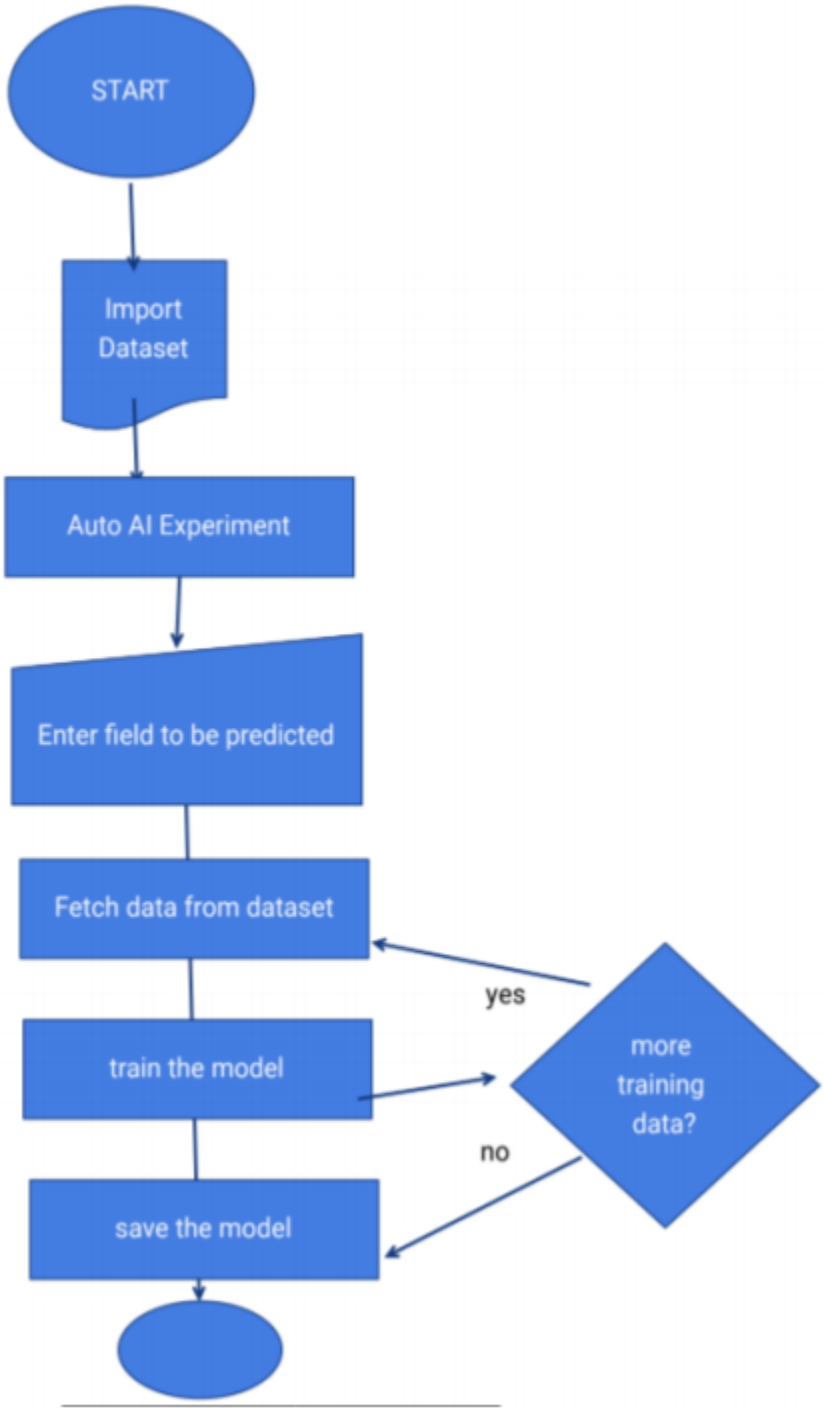
function node: A JavaScript function block to run against the messages being received by the node. The messages are passed in as a JavaScript object called msg. By convention it will have a msg.payload property containing the body of the message. The function is expected to return a
message object (or multiple message objects), but can choose to return nothing in order to halt a flow.

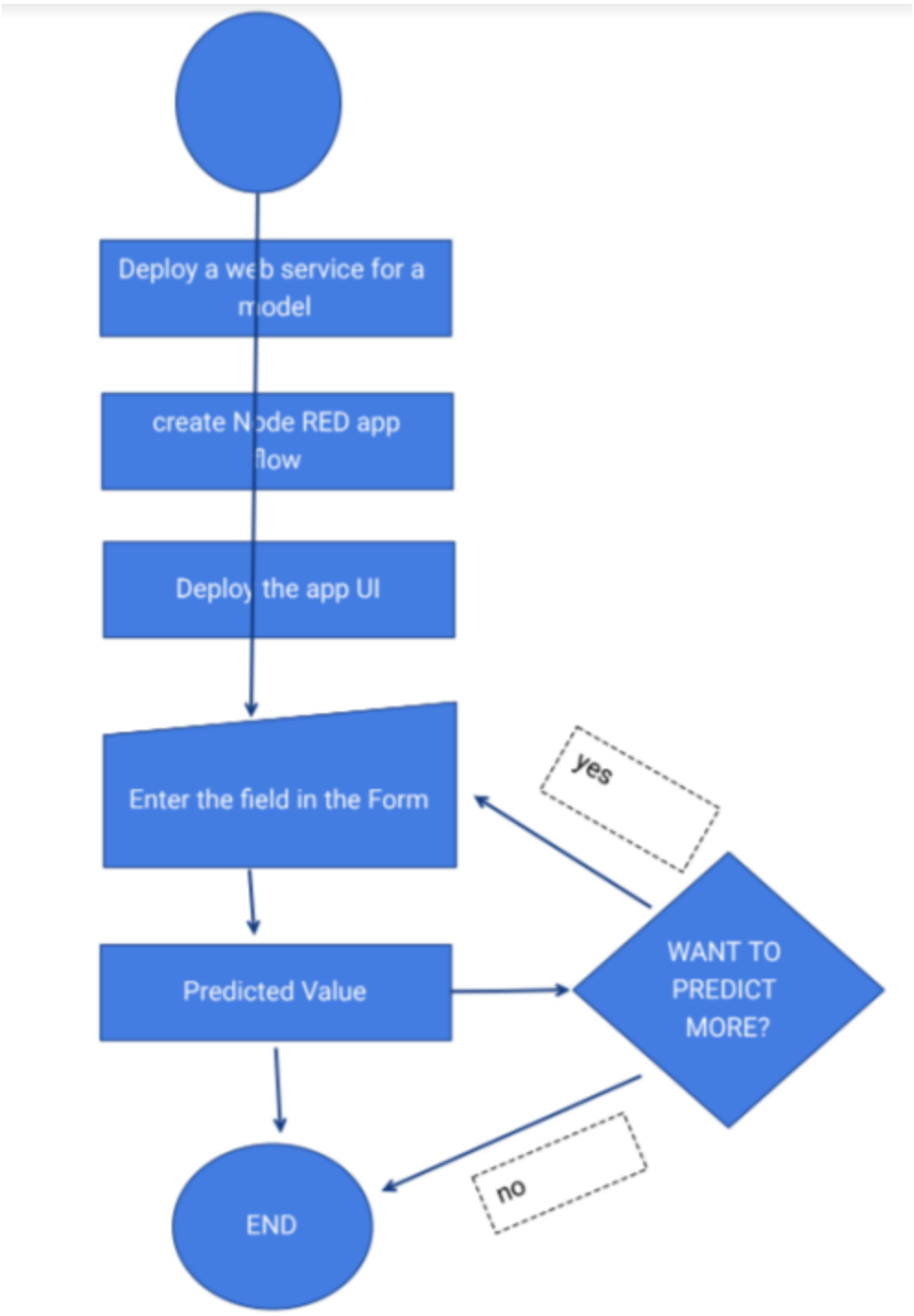
http request node: Sends HTTP requests and returns the response.

Text node: Will display a non-editable text field on the user interface. Each received msg.payload will update the text based on the provided Value Format.The Value Format field can be used to change the displayed format and can contain valid HTML and Angular filters.

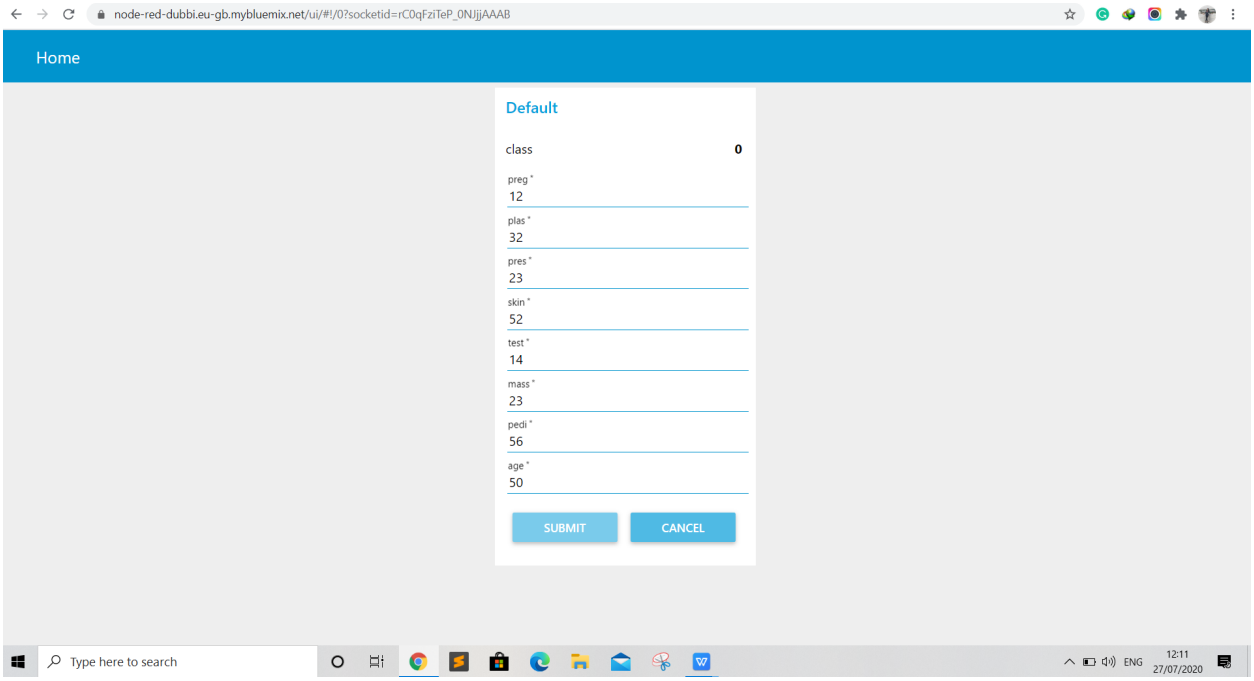
Debug node: Displays selected message properties in the debug sidebar tab and optionally the runtime log. By default it displays msg.payload, but can be configured to display any property, the full message or the result of a JSONata expression**.**

5. FLOWCHART


START
Import Dataset
Auto AI Experiment
Enter field to be predicted
Fetch data from dataset
train the model
save the model
more training data?
yes
no


Deploy a web service for a model
create Node RED app flow
Deploy the app UI
Enter the field in the Form
Predicted Value
WANT TO PREDICT MORE?
yes
no
END

6.RESULT



7.ADVANTAGES AND DISADVANTAGES

ADVANTAGES

1. This project prevents the people from the avalanche by priory informing them there is a chance to the occurrence of avalanche or not.
2. The earlier diagnosis is obtained, the much easier we can control it
3. Machine learning can help people make a preliminary judgment about diabetes mellitus according to their daily physical examination data, and it can serve as a reference for doctors.

4. The advantage of using AutoAI is we dont need to write the code we just have to give dataset as input,it automatically builds the model using AutoAI pipeline and gives the model with highest accuracy.

DISADVANTAGES

Deterministic problems: This method is not very efficient for deterministic problems.
Lack of good data:it may leads to problems.
Interpretability

8.APPLICATIONS

1. The application is used to predict diabetes for users.
2. This system can serve as a reference for doctors.

9.CONCLUSION

Diabetes mellitus is a disease, which can cause many complications. How to exactly predict and diagnose this disease by using machine learning is worthy studying.The end product is an webpage created and deployed on node-red app of IBM cloud. The backend of webpage is XGBoost Classifier built in AutoAI with 0.770 accuracy and deployed on watson studio using machine learning service.

The web page has input fields such as Preg,Plas,Pres,Skin,Test Mass,Pedi,Age and an output field named as Class which gives 0 or 1 based on the inputted values.

10.FUTURE SCOPE

Living with diabetes is challenging and distressful. Diabetic patient's condition cannot be understood only from his with medical charts. There is a need to collect and analyze both subjective and objective patient information in order to fully understand the occurrence of readmission of patients with diabetes.This predicting information might improve the intelligent models to identify patients at high risk of readmission.

11.BIBILOGRAPHY

Smart Bridge: Machine Learning
career basic
https://smartbridge.teachable.com
/

Source of Dataset:https://www.kaggle.com/akhilalexander/diabeticprediction/data

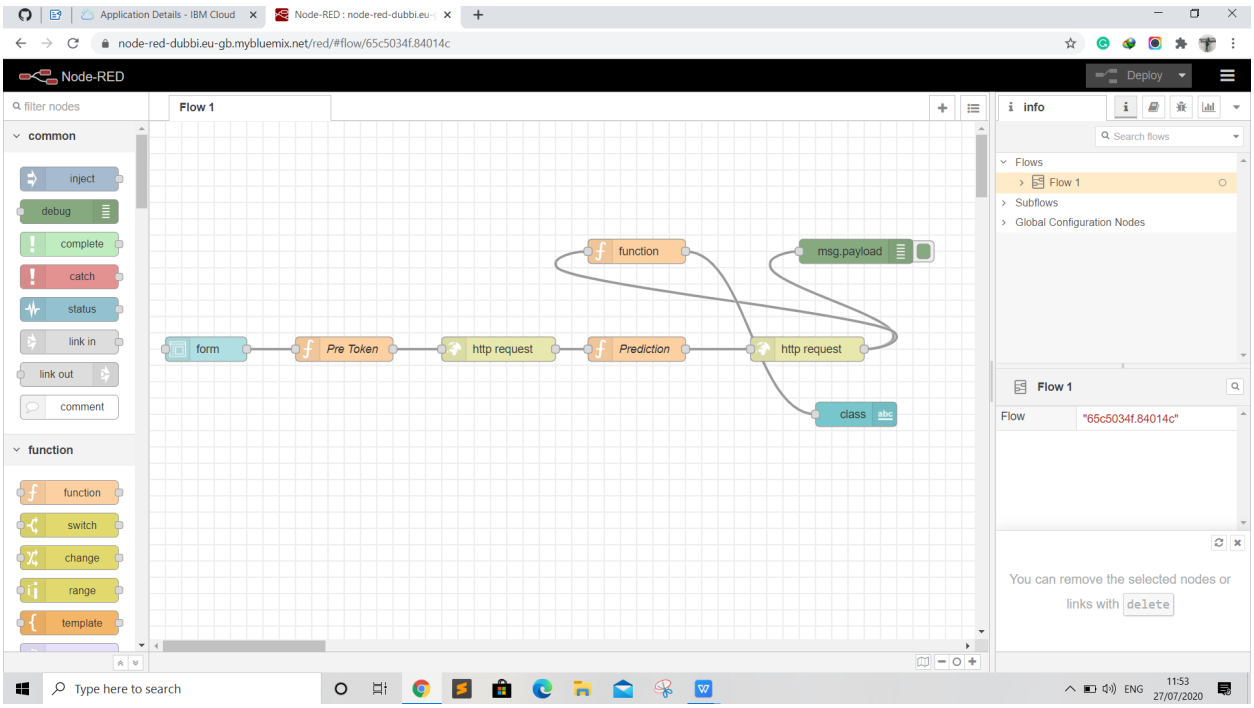
IBM Cloud
https://www.ibm.c
om/cloud

Node-Red:https://node-red-qbfab.eu-gb.mybluemix.net/

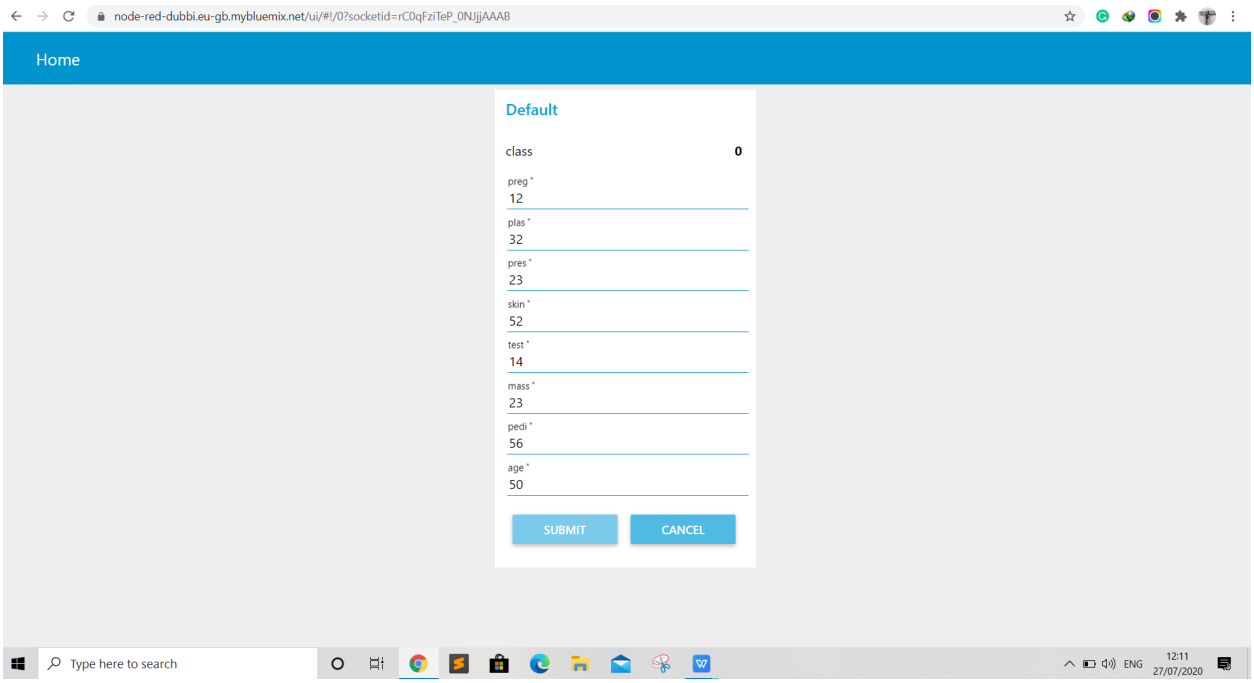
https://www.researchgate.net/publication/328766758_Predicting_Diabetes_Mellitus_With_Machine_Learning_Techniques

12.APPENDIX

Node-RED Flow Output



UI OUTPUT SCREENSHOT


node-red-dubbi.eu-gb.mybluemix.net/ui/#!/0?socketid=rC0qFziTeP_0NJjjAAAB
Home
Default
class
0
preg *
12
plas *
32
pres *
23
skin *
52
test *
14
mass *
23
pedi *
56
age *
50
SUBMIT
CANCEL
Type here to search
ENG
12:11
27/07/2020