PROJECT REPORT

ON

Chronic Kidney Disease prediction using Watson Auto AI

Internship Title : RSIP Career Basic ML 195
Name : Altaf Bin Sarfraz
Registered Mail Id: 2018.bin.altaf@ves.ac.in

Index

1. INTRODUCTION

1.1 Overview

Chronic Kidney Disease (CKD) is a major medical problem and can be cured if treated it in the early stages.

In this project we will Predict whether a person with speciifc Attributes/label is suffering from Chronic Kidney Disease or not. We will be deploying our model in IBM Cloud Services to deliver an effective web based UI through which
we can predict the disease.

Project Requirements :IBM Cloud,IBM Watson Auto AI,Node-Red

Functional Requirements :IBM Cloud.

Software Requirements :Watson Auto AI, Node-Red.

Project deliverables : Chronic kidney disease prediction using Watson Auto AI

1.2 Purpose

The purpose of this project is to dentify the patients with disease. Once any person gets kidney disease, they may suffer from the disease which may decrease their working capability as well as living quality. Our aim is to predict patients with chronic kidney failure (ckd) disease and patients who do not (notckd) suffer from the disease. So for that we are building a Machine Learning model to predict the compressive strength of concrete using IBM Watson AutoAI Machine Learning Service. The model is deployed on IBM cloud to get scoring end point which can be used as API in mobile app or web app building. We are developing a web application which is built using node red service. We make use of the scoring end point to give user input values to the deployed model. The model prediction is then showcased on User Interface.

2. LITERATURE SURVEY

2.1. Existing Problem

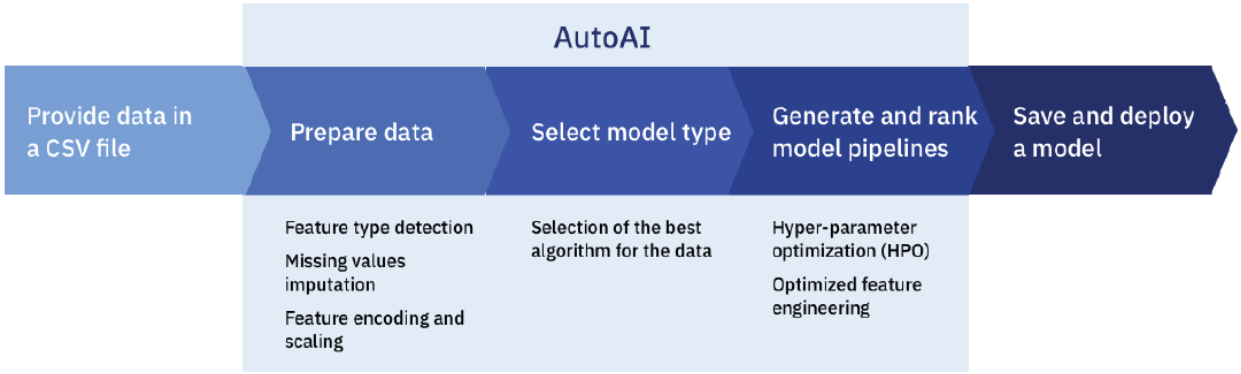
Chronic Kidney Disease (CKD) is a major medical problem and can be cured if treated it in the early stages. Usually, people are not aware that medical tests, we take for different purposes could contain valuable information concerning kidney diseases. Consequently, attributes of various medical tests are investigated to distinguish which attributes may contain helpful information about the disease.The information says that it helps us to measure the severity of the problem, the predicted survival of the patient after the illness, the pattern of the disease and work for curing the disease.

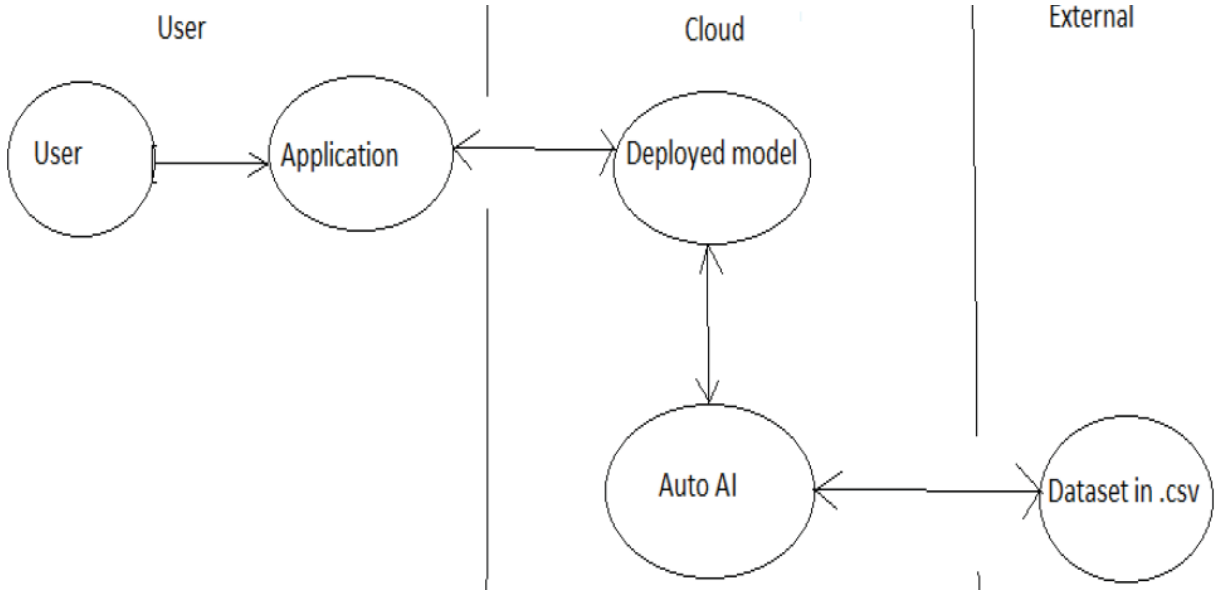
2.2. Proposed Solution

In this proposed system we are able to identify the patients with disease. Once any person gets kidney disease, they may suffer from the disease which may decrease their working capability as well as living quality. Our aim is to predict patients with chronic kidney failure (ckd) disease and patients who do not (notckd) suffer from the disease. So for that we are building a Machine Learning model to predict the compressive strength of concrete using IBM Watson AutoAI Machine Learning Service. The model is deployed on IBM cloud to get scoring end point which can be used as API in mobile app or web app building. We are developing a web application which is built using node red service. We make use of the scoring end point to give user input values to the deployed model. The model prediction is then showcased on User Interface.

3. THEORITICAL ANALYSIS

3.1 **Block Diagram**





3.2 Hardware/Software Designing

Step 1:Create IBM Cloud account and create IBM Watson Studio .

Step 2:Build and Train the experiment

2.1 Specify basic experiment details

1.From the Assets page of your project, click Add to project and choose Auto AI Experiment.

2.In the page that opens, fill in the basic fields:

Specify a name and optional description for your new experiment.

Confirm that the IBM Watson Machine Learning service instance that you associate with your project is selected in the machine learning service selection.

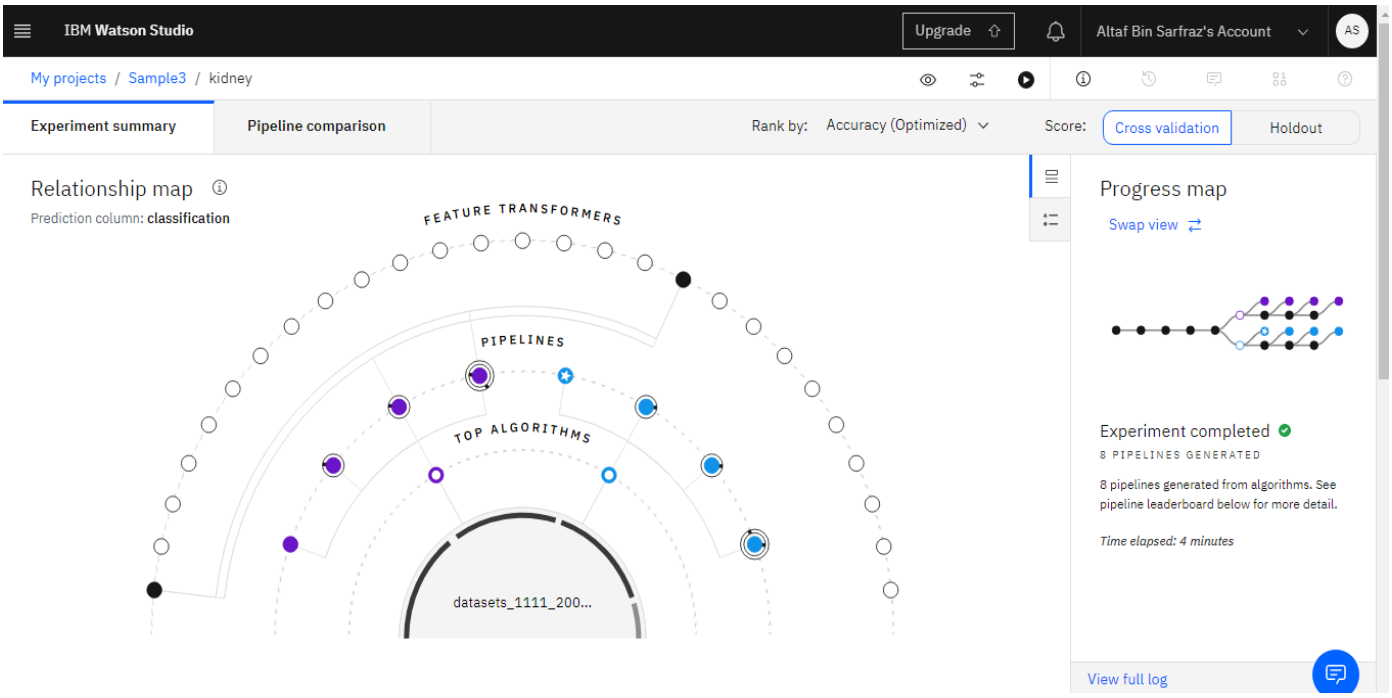
3.Click Create.

2.2 Add training data

Upload the training data file, auto.csv, from your local computer by dragging the file onto the data panel or by clicking browse and then following the prompts.

2.3 Train the model

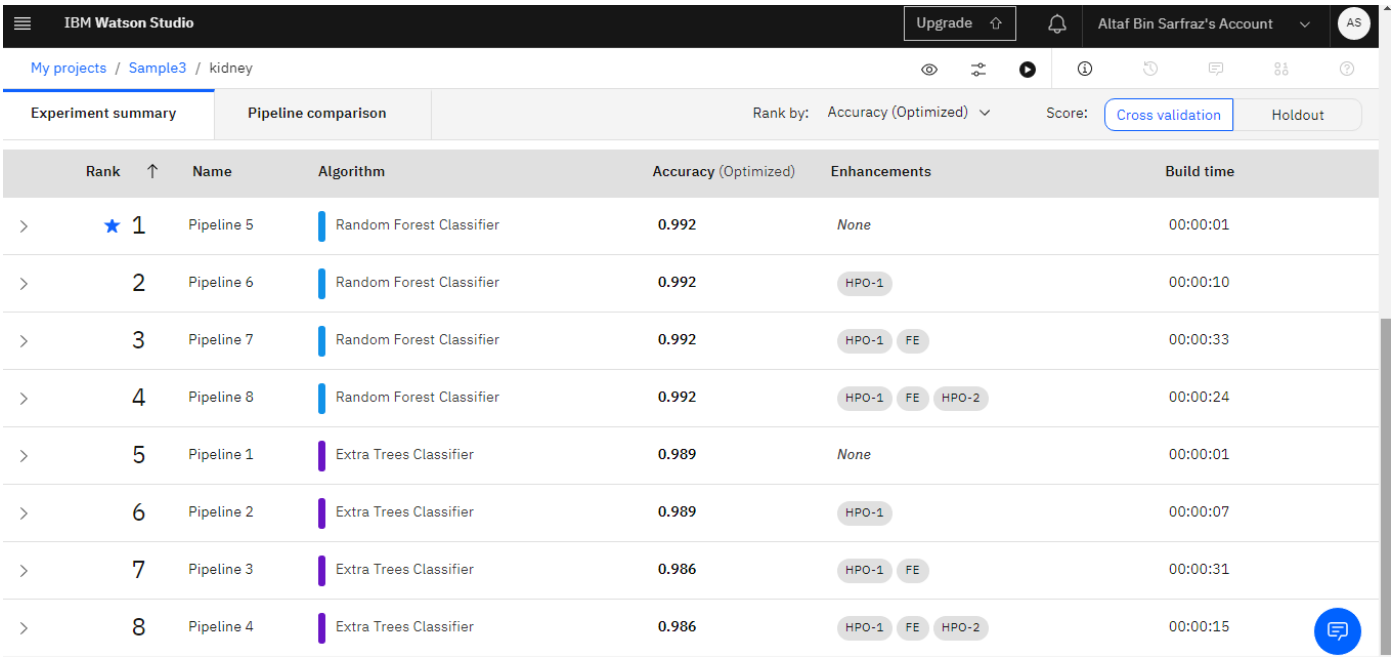
Choose the column you want to predict and also in add experiment select the columns withe data that supports prediction column save it and select run the experiment.



As the model trains, you will see an infographic that shows the process of building the pipelines.

2.4 Choose a pipeline

Once the pipeline creation is complete, you can view and compare the ranked pipelines in a leaderboard



Choose Save model from the action menu for Pipeline 5. This saves the pipeline as a Machine Learning asset in your project.

Step 3 :Deploy the Model

Before you can use your trained model to make predictions on new data, you must deploy the model.Clicking on the model name in the notification displayed when you save the model

From the model details page:

Click the Deployments tab.

Click Add Deployment.

In the page that opens, fill in the fields:

Specify a name for the deployment.

Click Save.

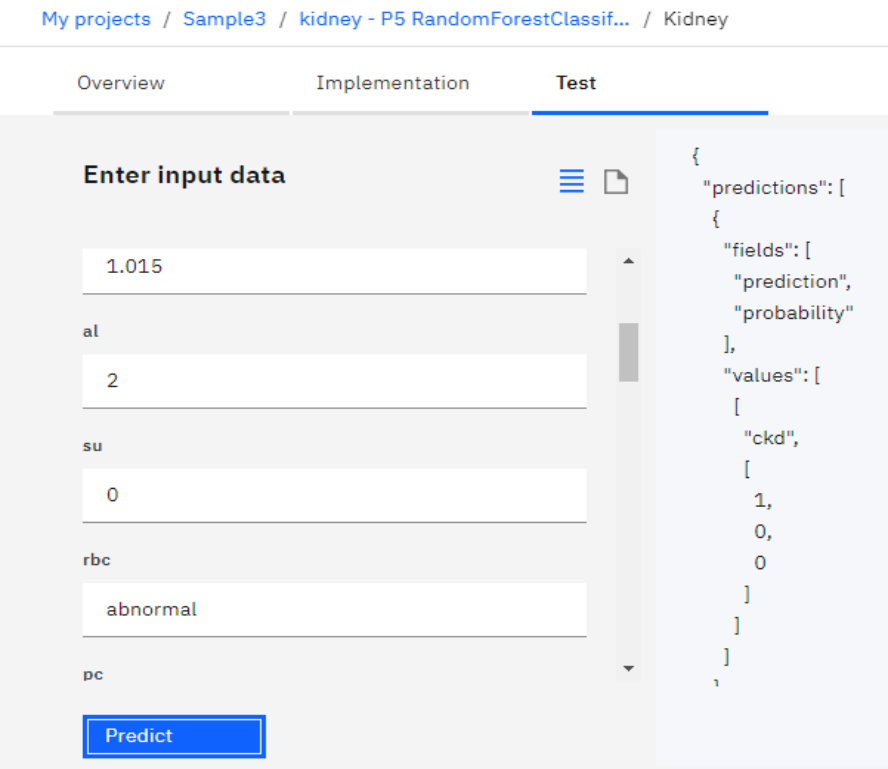
After you save the deployment, click on the deployment name to view the deployment details page.

Step 4:Test the Deployed model

You can test the deployed model from the deployment details page:

On the Test tab of the deployment details page, either fill out the form with test values, or enter the following JSON test data.

Click predict.



4 EXPERIMENTAL INVESTIGATIONS

Chronic Kidney Disease (CKD) is a major medical problem and can be cured if treated it in the early stages. Usually, people are not aware that medical tests, we take for different purposes could contain valuable information concerning kidney diseases. Consequently, attributes of various medical tests are investigated to distinguish which attributes may contain helpful information about the disease.The information says that it helps us to measure the severity of the problem, the predicted survival of the patient after the illness, the pattern of the disease and work for curing the disease.

Features/Independent Variables : id, age, bp, sg, al, su, rbc, pc, pcc, ba, bgr, bu, sc, sod, pot, hemo, pcv, wc, rc, htn, dm, cad, appet, pe, ane.

Target Value/Dependent Variable : classification (ckd/notckd)

Total Records available : 400

Train-Test Split : 90% for Training the model
10% for Testing the model

Predicted Column : Classification

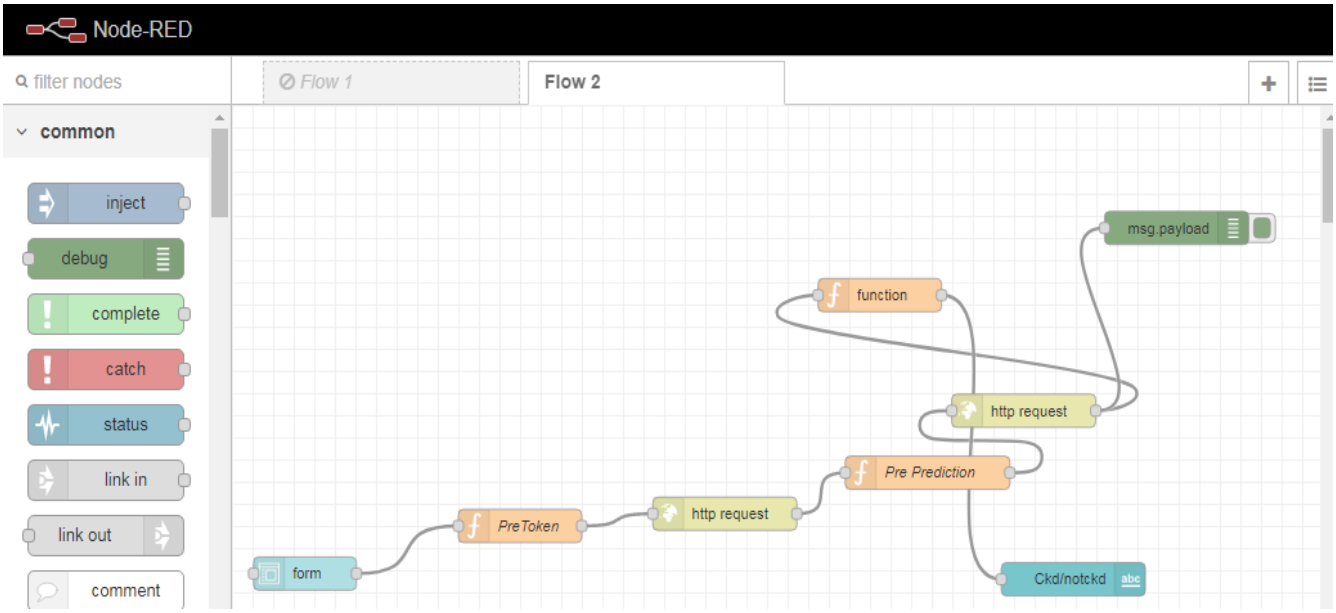
Total Pipelines Generated : 8

Pipeline Selected : Algorithm: Random Forest Classifier
Accuracy Optimized : 0.992

5 FLOWCHART

Insert the following nodes into the flow in Node-Red.

- ui_Form
- Input
- Function
- Http request
- Text
- Debug

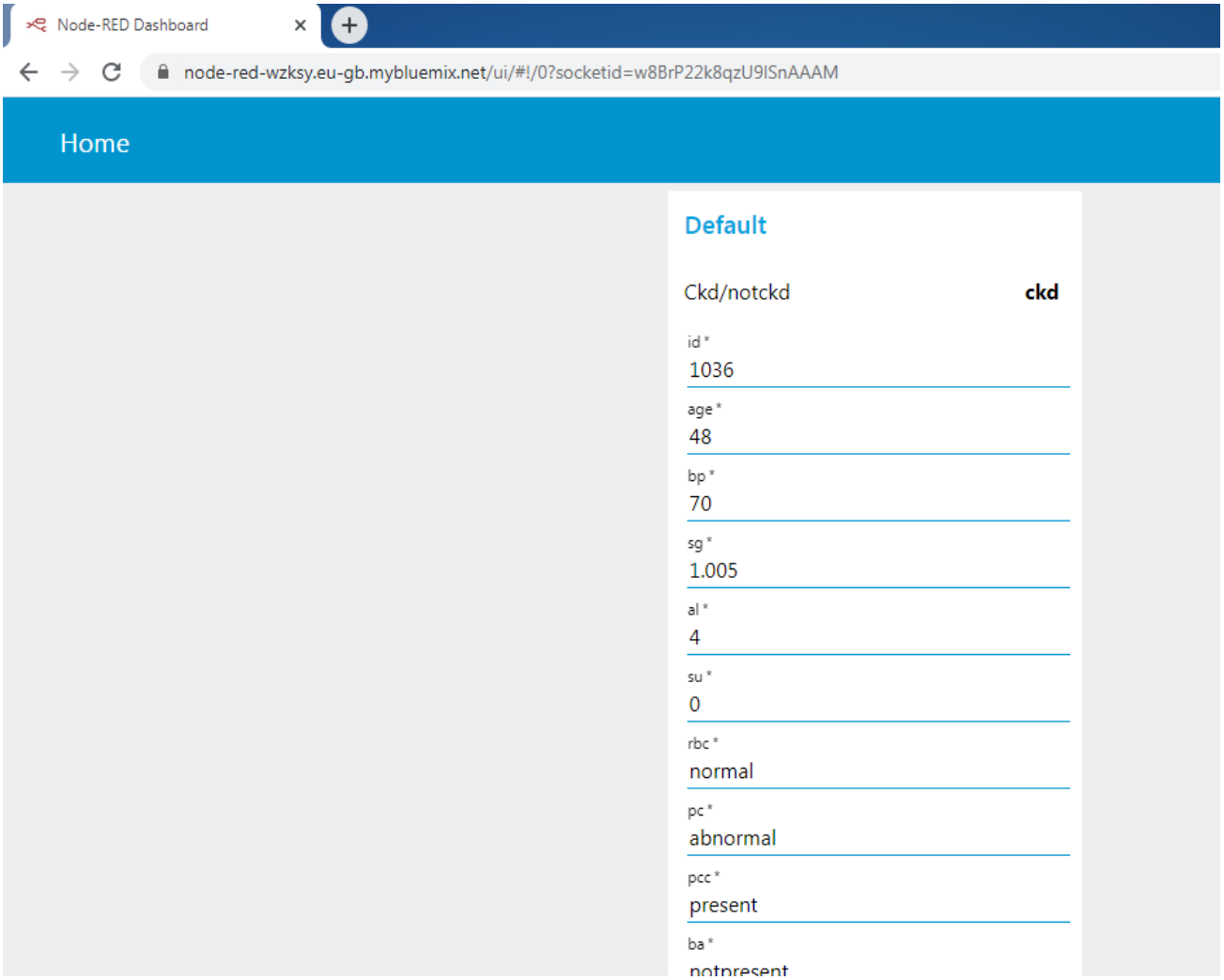


6 RESULT

Sucessfully developed a web application which is built using node red service. We make use of the scoring end point to give user input values to the deployed model. The model prediction is then showcased on User Interface.With a high accuracy of 0.992

URL for UI dashboard: https://node-red-wzksy.eu-gb.mybluemix.net/ui

Predicted Output of Person suffering from Chronic Kidney Disease (ckd)



bgr *
117

bu *
56

sc *
3.8

sod *
111

pot *
2.5

hemo *
11.2

pcv *
32

wc *
6700

rc *
3.9

htn *
yes

dm *
no

cad *
no

cad *
no

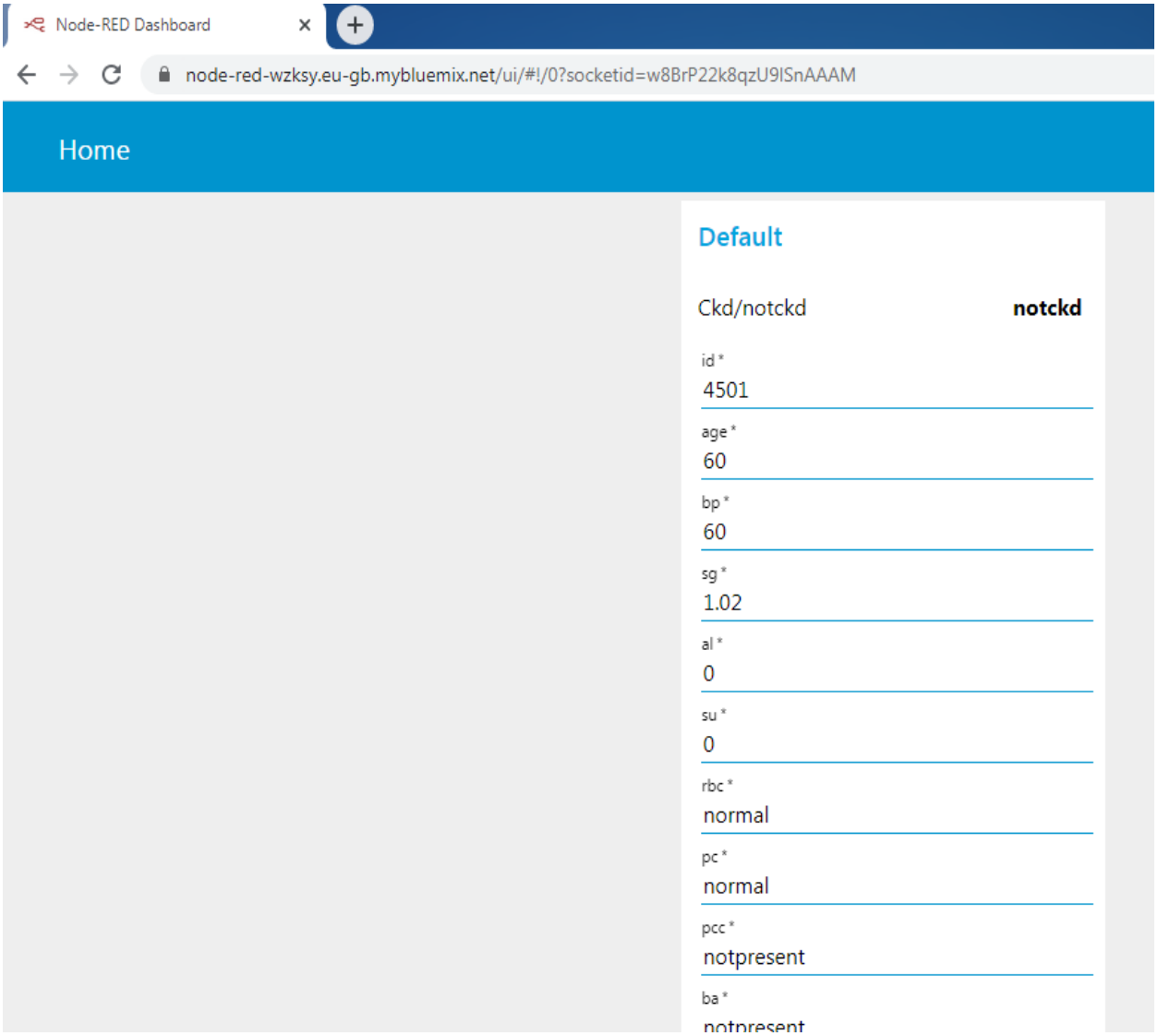
appet *
poor

pe *
yes

ane *
yes

SUBMIT CANCEL

Predicted Output of Person not suffering from Chronic Kidney Disease (notckd)



bgr *
134

bu *
45

sc *
0.5

sod *
139

pot *
4.8

hemo *
14.2

pcv *
48

wc *
10700

rc *
5.6

htn *
no

dm *
no

cad *
no

cad *
no

appet *
good

pe *
no

ane *
no

SUBMIT CANCEL

7 ADVANTAGES & DISADVANTAGES

Advantages:

1) Instant Result to detect Chronic Kidney Disease in early stages.
2) Can Detect from home. No need to visit Laboratory and wait for result.
3) With Accuarcy score of the Model built 99.2% it is safer & reliable to use.
4) Cost efficient.
5) Reduction in human errors.
6) Easy Web Application User Interference to use.

Disadvantages:

1) Machine Learning Engineers/AI Developers should ensure that Proper Algorithm with high Accuracy score is selected.
2) Overfitting of Model should be avoided.
3) Utilizes resource to build.

8 APPLICATIONS

1) With Accuracy Score of 99.2% model can be sucessfully deployed to predict Chronic Kidney Disease in a person.
2)Model will provide the much needed treatment by detecting the disease at an early stage.
3)Will Provide Medical Team to speed up the treatment by instantly displaying result.
4) Cost Efficient.
5) Will help gain Insights on parameters to which Chronic Kidney Disease is maximum co-related.
6) Significant Contribution to Research work for curing the disease.

9 CONCLUSION

Successfully implemented Chronic kidney disease prediction using Watson Auto AI strictly in accordance with the Specifications provided, and adhering to the time frame allocated.

With Accuracy of 99.2% successfully predicted and detected persons suffering from Chronic Kidney Disease.

Deployed a Machine Learning model to predict the compressive strength of concrete using IBM Watson AutoAI Machine Learning Service. The model is deployed on IBM cloud to get scoring end point which can be used as API in mobile app or web app building. We are developing a web application which is built using node red service. We make use of the scoring end point to give user input values to the deployed model. The model prediction is then showcased on User Interface.

10 FUTURE SCOPE

In Nearby future we intend to collect more data and to use more advanced techniques like artificial neural networks,image recognisation, fuzzy logic and genetic algorithms to predict Chronic Kidney Disease in early stages.Further carry out investigation to distinguish which attributes may contain helpful information about the disease.

11 BIBILOGRAPHY

1.https://smartbridge.teachable.com/

2.Auto AI with IBM Watson studio :
https://www.ibm.com/in-en/cloud/watson-studio/autoai

3.Node-RED starter application :
https://developer.ibm.com/tutorials/how-to-create-a-node-red-starter-application/

12 APPENDIX

A) Source Code(.Json Flow File)

[{"id":"856d360f.8ae9c8","type":"tab","label":"Flow 2","disabled":false,"info":""},{"id":"b4b78095.106a1","type":"function","z":"856d360f.8ae9c8"

,"name":"PreToken","func":"global.set(\"id\",msg.payload.id)\nglobal.set(\"age\",msg.payload.age)\nglobal.set(\"bp\",msg.payload.bp)\nglobal.set(\"sg\",msg.payload.sg)\nglobal.set(\"al\",msg.payload.al)\nglobal.set(\"su\",msg.payload.su)\nglobal.set(\"rbc\",msg.payload.rbc)\nglobal.set(\"pc\",msg.payload.pc)\nglobal.set(\"pcc\",msg.payload.pcc)\nglobal.set(\"ba\",msg.payload.ba)\nglobal.set(\"bgr\",msg.payload.bgr)\nglobal.set(\"bu\",msg.payload.bu)\nglobal.set(\"sc\",msg.payload.sc)\nglobal.set(\"sod\",msg.payload.sod)\nglobal.set(\"pot\",msg.payload.pot)\nglobal.set(\"hemo\",msg.payload.hemo)\nglobal.set(\"pcv\",msg.payload.pcv)\nglobal.set(\"wc\",msg.payload.wc)\nglobal.set(\"rc\",msg.payload.rc)\nglobal.set(\"htn\",msg.payload.htn)\nglobal.set(\"dm\",msg.payload.dm)\nglobal.set(\"cad\",msg.payload.cad)\nglobal.set(\"appet\",msg.payload.appet)\nglobal.set(\"pe\",msg.payload.pe)\nglobal.set(\"ane\",msg.payload.ane)\nvar apikey=\"ecVDju9o4PlSbiLQdu7et18QBHVgdw_x_XOPTd5CMDo2\";\nmsg.headers={\"content-type\":\"application/x-www-form-urlencoded\"}\nmsg.payload={\"grant_type\":\"urn:ibm:params:oauth:grant-type:apikey\",\"apikey\":apikey}\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":280,"y":380,"wires":[["8dd7ceed.a5176"]]},{"id":"8dd7ceed.a5176","type":"http request","z":"856d360f.8ae9c8","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":479,"y":368.9999952316284,"wires":[["fcec2f00.a839"]]},{"id":"ea2fc957.ec9f88","type":"debug","z":"856d360f.8ae9c8","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","x":918.0000114440918,"y":110.99999713897705,"wires":[]},{"id":"fcec2f00.a839","type":"function","z":"856d360f.8ae9c8","name":"Pre Prediction","func":"var id = global.get('id')\nvar age = global.get('age')\nvar bp = global.get('bp')\nvar sg = global.get('sg')\nvar al = global.get('al')\nvar su = global.get('su')\nvar rbc = global.get('rbc')\nvar pc = global.get('pc')\nvar pcc = global.get('pcc')\nvar ba = global.get('ba')\nvar bgr = global.get('bgr')\nvar bu = global.get('bu')\nvar sc = global.get('sc')\nvar sod = global.get('sod')\nvar pot = global.get('pot')\nvar hemo = global.get('hemo')\nvar pcv = global.get('pcv')\nvar wc = global.get('wc')\nvar rc = global.get('rc')\nvar htn = global.get('htn')\nvar dm = global.get('dm')\nvar cad = global.get('cad')\nvar appet = global.get('appet')\nvar pe = global.get('pe')\nvar ane = global.get('ane')\nvar token=msg.payload.access_token\nvar instance_id=\"c69fd457-dd98-49ee-807f-e178e05d4415\"\nmsg.headers={'Content-Type': 'application/json',\"Authorization\":\"Bearer \"+token,\"ML-Instance-ID\":instance_id}\nmsg.payload={\"input_data\": [{\"fields\": [\"id\", \"age\", \"bp\", \"sg\", \"al\", \"su\", \"rbc\", \"pc\", \"pcc\", \"ba\", \"bgr\", \"bu\", \"sc\", \"sod\", \"pot\", \"hemo\", \"pcv\", \"wc\", \"rc\", \"htn\", \"dm\", \"cad\", \"appet\", \"pe\", \"ane\"], \"values\":

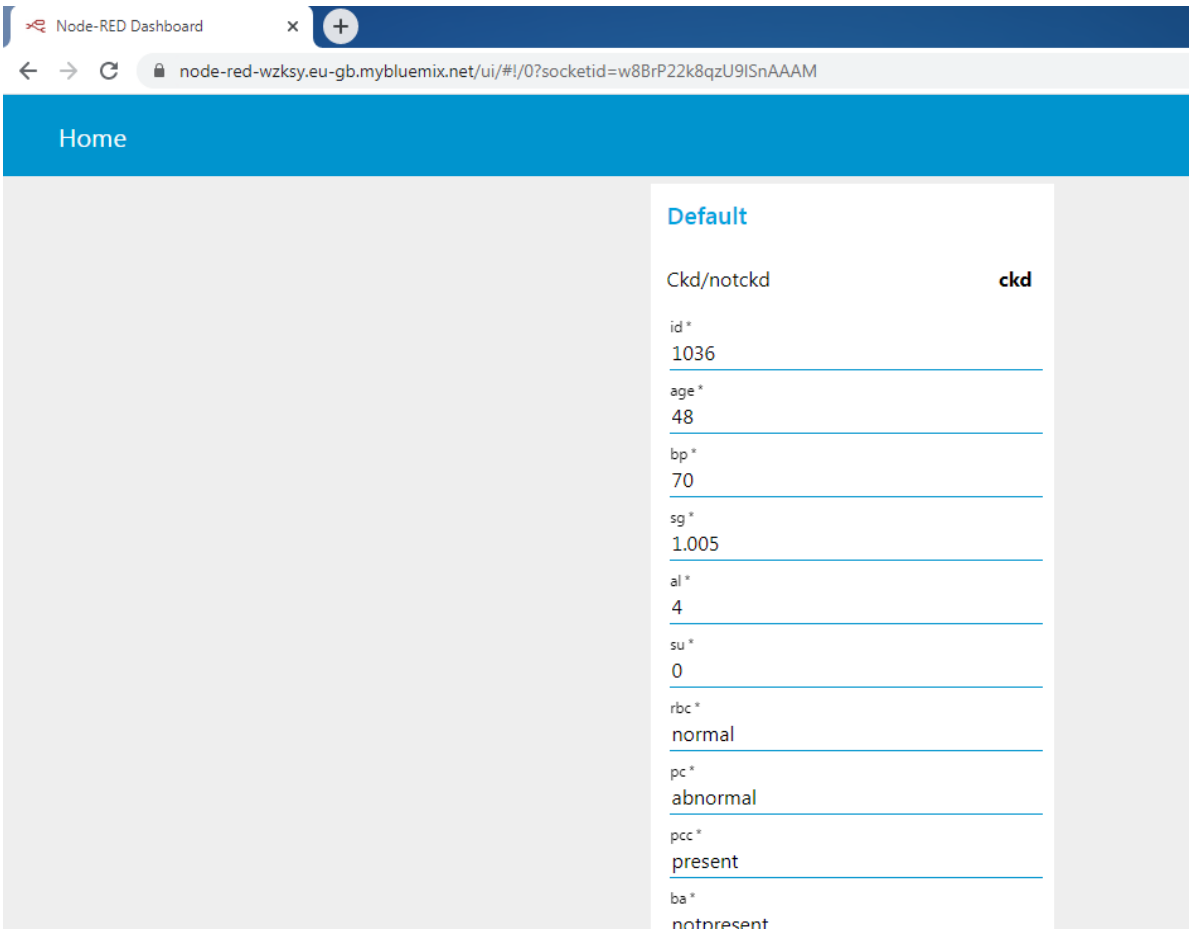
[[id,age,bp,sg,al,su,rbc,pc,pcc,ba,bgr,bu,sc,sod,pot,hemo,pcv,wc,rc,htn,dm,cad,appet,pe,ane]]}]}\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":676.0000076293945,"y":331.9999990463257,"wires":[["cecf5fb0.4a961"]]},{"id":"cecf5fb0.4a961","type":"http request","z":"856d360f.8ae9c8","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://us-south.ml.cloud.ibm.com/v4/deployments/312f0416-3638-408b-b0dd-0109f282e787/predictions","tls":"","persist":false,"proxy":"","authType":"","x":769.5000076293945,"y":275.9999990463257,"wires":[["fe13623.1056ca","ea2fc957.ec9f88"]]},{"id":"444106ed.809de8","type":"ui_form","z":"856d360f.8ae9c8","name":"","label":"","group":"b12c2b4a.591cb8","order":0,"width":0,"height":0,"options":[{"label":"id","value":"id","type":"number","required":true,"rows":null},{"label":"age","value":"age","type":"number","required":true,"rows":null},{"label":"bp","value":"bp","type":"number","required":true,"rows":null},{"label":"sg","value":"sg","type":"number","required":true,"rows":null},{"label":"al","value":"al","type":"number","required":true,"rows":null},{"label":"su","value":"su","type":"number","required":true,"rows":null},{"label":"rbc","value":"rbc","type":"text","required":true,"rows":null},{"label":"pc","value":"pc","type":"text","required":true,"rows":null},{"label":"pcc","value":"pcc","type":"text","required":true,"rows":null},{"label":"ba","value":"ba","type":"text","required":true,"rows":null},{"label":"bgr","value":"bgr","type":"number","required":true,"rows":null},{"label":"bu","value":"bu","type":"number","required":true,"rows":null},{"label":"sc","value":"sc","type":"number","required":true,"rows":null},{"label":"sod","value":"sod","type":"number","required":true,"rows":null},{"label":"pot","value":"pot","type":"number","required":true,"rows":null},{"label":"hemo","value":"hemo","type":"number","required":true,"rows":null},{"label":"pcv","value":"pcv","type":"text","required":true,"rows":null},{"label":"wc","value":"wc","type":"text","required":true,"rows":null},{"label":"rc","value":"rc","type":"text","required":true,"rows":null},{"label":"htn","value":"htn","type":"text","required":true,"rows":null},{"label":"dm","value":"dm","type":"text","required":true,"rows":null},{"label":"cad","value":"cad","type":"text","required":true,"rows":null},{"label":"appet","value":"appet","type":"text","required":true,"rows":null},{"label":"pe","value":"pe","type":"text","required":true,"rows":null},{"label":"ane","value":"ane","type":"text","required":true,"rows":null}],"formValue":{"id":"","age":"","bp":"","sg":"","al":"","su":"","rbc":"","pc":"","pcc":"","ba":"","bgr":"","bu":"","sc":"","sod":"","pot":"","hemo":"","pcv":"","wc":"","rc":"","htn":"","dm":"","cad":"","appet":"","pe":"","ane":""},"payload":"","submit":"submit","cancel":"cancel","topic":"","x":71,"y":422.9999952316284,"wires":[["b4b78095.106a1"]]},{"id":"5ab023a1.21cf2c","type":"ui_text","z":"856d360f.8ae9c8","group":"b12c2b4a.591cb8","order":1,"width":0,"height":0,"name":"","label":"Ckd/notckd","format":"{{msg.payload}}","layout":"row-spread","x":817.6000366210938,"y":428.20000743865967,"wires":[]},{"id":"fe13623.1056ca","type":"function","z":"856d360f.8ae9c8","name":"","func":"msg.payload=msg.payload.predictions[0].values[0][0]\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":629.6000366210938,"y":171.600001

33514404,"wires":[["5ab023a1.21cf2c"]]},{"id":"b12c2b4a.591cb8","type":"ui_group","z":"","name":"Default","tab":"4ff9446e.231e1c","order":1,"disp":true,"width":"6","collapse":false},{"id":"4ff9446e.231e1c","type":"ui_tab","z":"","name":"Home","icon":"dashboard","disabled":false,"hidden":false}]

B UI O/P Screen

URL for UI dashboard: https://node-red-wzksy.eu-gb.mybluemix.net/ui

Predicted Output of Person suffering from Chronic Kidney Disease (ckd)



bgr *
117

bu *
56

sc *
3.8

sod *
111

pot *
2.5

hemo *
11.2

pcv *
32

wc *
6700

rc *
3.9

htn *
yes

dm *
no

cad *
no

cad *
no

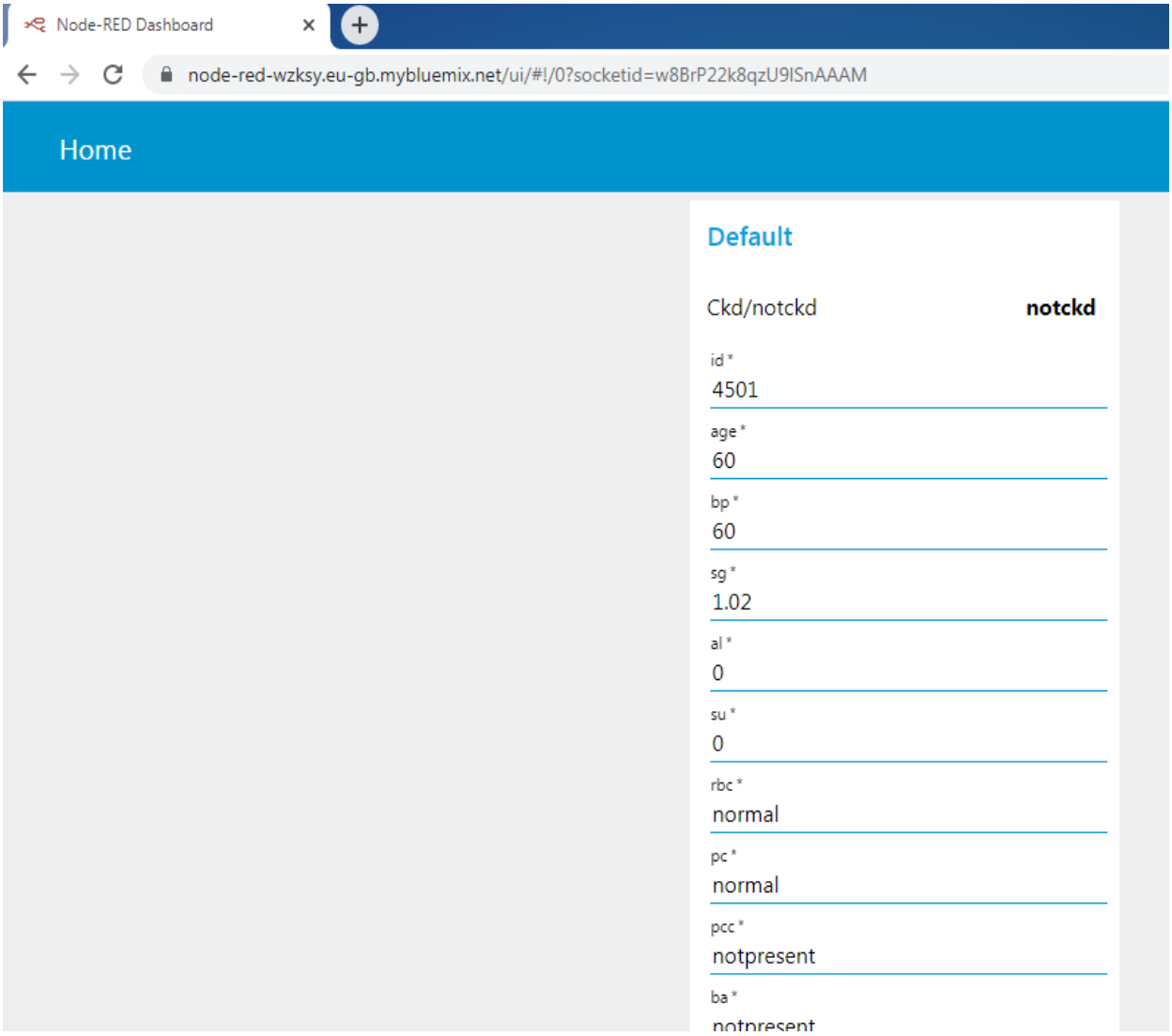
appet *
poor

pe *
yes

ane *
yes

SUBMIT CANCEL

Predicted Output of Person not suffering from Chronic Kidney Disease (notckd)



bgr *
134

bu *
45

sc *
0.5

sod *
139

pot *
4.8

hemo *
14.2

pcv *
48

wc *
10700

rc *
5.6

htn *
no

dm *
no

cad *
no

cad *
no

appet *
good

pe *
no

ane *
no

SUBMIT CANCEL