**CHRONIC KIDNEY DISEASE PREDICTION USING AUTO AI**

## **Internship Title : RSIP Career Basic ML132**

## **Duration : 1 Month**

**Project ID : SPS\_PRO\_288**

**Project Title : Chronic Kidney Disease Prediction Using Auto AI**

**Team : KS**

**Slot : June 15 - Slot (5)**

**Submitted by,**

**Kanisha S**

**INDEX**

|  |  |  |  |
| --- | --- | --- | --- |
| **Chapter No.** | **Title** | | **Page No.** |
| **1** | **Introduction** | | **3** |
|  | 1.1 | Overview | **3** |
|  | 1.2 | Purpose | **3** |
| **2** | **Literature Survey** | | **4** |
|  | 2.1 | Existing problem | **4** |
|  | 2.2 | Proposed Solution | **4** |
| **3** | **Theoretical Analysis** | | **5** |
|  | 3.1 | Block Diagram | **5** |
|  | 3.2 | Hardware/Software Designing | **5** |
| **4** | **Experimental Investigations** | | **6 - 20** |
| **5** | **Flow Chart** | | **20** |
| **6** | **Result** | | **21** |
| **7** | **Advantages and Disadvantages** | | **22** |
|  | 7.1 | Advantages | **22** |
|  | 7.2 | Disadvantages | **22** |
| **8** | **Applications** | | **22** |
| **9** | **Conclusion** | | **23** |
| **10** | **Future Scope** | | **23** |
| **11** | **Bibliography** | | **23** |

1. **INTRODUCTION**
   1. **Overview:**

**Chronic kidney disease** (**CKD**) is a type of [kidney disease](https://en.wikipedia.org/wiki/Kidney_disease) in which there is gradual loss of [kidney function](https://en.wikipedia.org/wiki/Kidney_function) over a period of months to years. Initially there are generally no symptoms; later, symptoms may include [leg swelling](https://en.wikipedia.org/wiki/Pedal_edema), feeling tired, [vomiting](https://en.wikipedia.org/wiki/Vomiting), loss of appetite, and [confusion](https://en.wikipedia.org/wiki/Confusion).[[](https://en.wikipedia.org/wiki/Chronic_kidney_disease#cite_note-NIH2017What-2) Complications include an increased risk of [heart disease](https://en.wikipedia.org/wiki/Cardiovascular_disease), [high blood pressure](https://en.wikipedia.org/wiki/High_blood_pressure), [bone disease](https://en.wikipedia.org/wiki/Renal_osteodystrophy), and [anemia](https://en.wikipedia.org/wiki/Anemia).Causes of chronic kidney disease include [diabetes](https://en.wikipedia.org/wiki/Diabetic_nephropathy), [high blood pressure](https://en.wikipedia.org/wiki/High_blood_pressure), [glomerulonephritis](https://en.wikipedia.org/wiki/Glomerulonephritis), and [polycystic kidney disease](https://en.wikipedia.org/wiki/Polycystic_kidney_disease). Risk factors include a family history of chronic kidney disease. Diagnosis is by [blood tests](https://en.wikipedia.org/wiki/Blood_tests) to measure the estimated [glomerular filtration rate](https://en.wikipedia.org/wiki/Glomerular_filtration_rate) (eGFR), and a [urine test](https://en.wikipedia.org/wiki/Urinalysis) to measure [albumin](https://en.wikipedia.org/wiki/Albumin).[Ultrasound](https://en.wikipedia.org/wiki/Renal_ultrasound) or [kidney biopsy](https://en.wikipedia.org/wiki/Kidney_biopsy) may be performed to determine the underlying cause. Several severity-based staging systems are in use.

## **Purpose:**

The main purpose of this project is to predict the disease among people with the collected data which will include test results of hemoglobin, rbc,rc,pcc etc. With the help of these few test results we can conclude if the person is affected by chronic kidney disease or not that is at the end the result will be shown as ckd or notckd . This is a vital test as thousands of people die without proper prediction in our country of not just this disease but many other diseases. If treated at the beginning then death due to kidney failures can be prevented.

# **LITERATURE SURVEY**

**Existing Problem:**

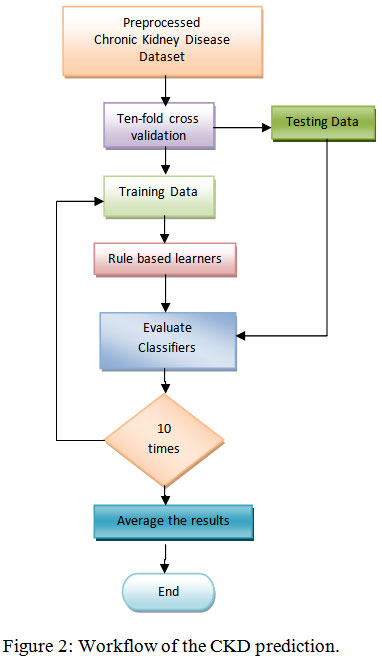
Chronic Kidney Disease (CKD) is a major medical problem and can be cured if treated it in the early stages. Usually, people are not aware that medical tests, we take for different purposes could contain valuable information concerning kidney diseases. Consequently, attributes of various medical tests 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.

## **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.

# **THEORETICAL ANALYSIS**

## **Block Diagram:**



* 1. **Hardware/ Software Designing**

This dataset is first tested by using various algorithms in our jupyter notebooks and then implemented in the IBM Cloud Platform. We upload our dataset in the cloud platform and choose the parameter to be predicted and we choose the number of algorithms and pipelines to be used. The cloud platform then predicts the best suited algorithm for our dataset We can also compare the performance of other algorithms used. The Auto AI function in the IBM cloud aids in deployment of our final machine learning models. This helps us to implement and test our model for our dataset. We have then, created a node red app for our deployed model. This UI will help us predict the avalanche in real time as we enter the details. This app aids in easy usage and better user interface.

# **EXPERIMENTAL INVESTIGATIONS**

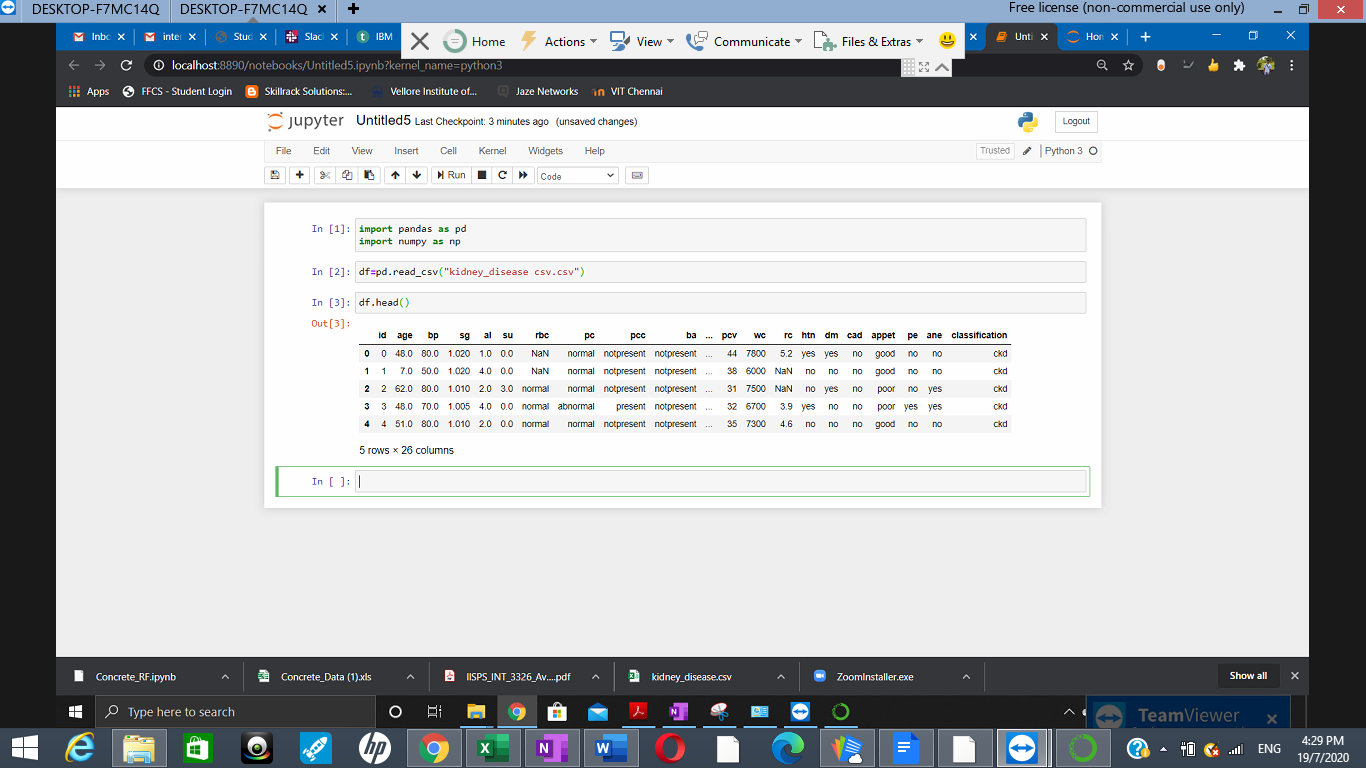
## **Step1:- Data Collection: Jupyter Notebook:**

Downloaded the dataset provided from Kaagle and did data pre- processing.

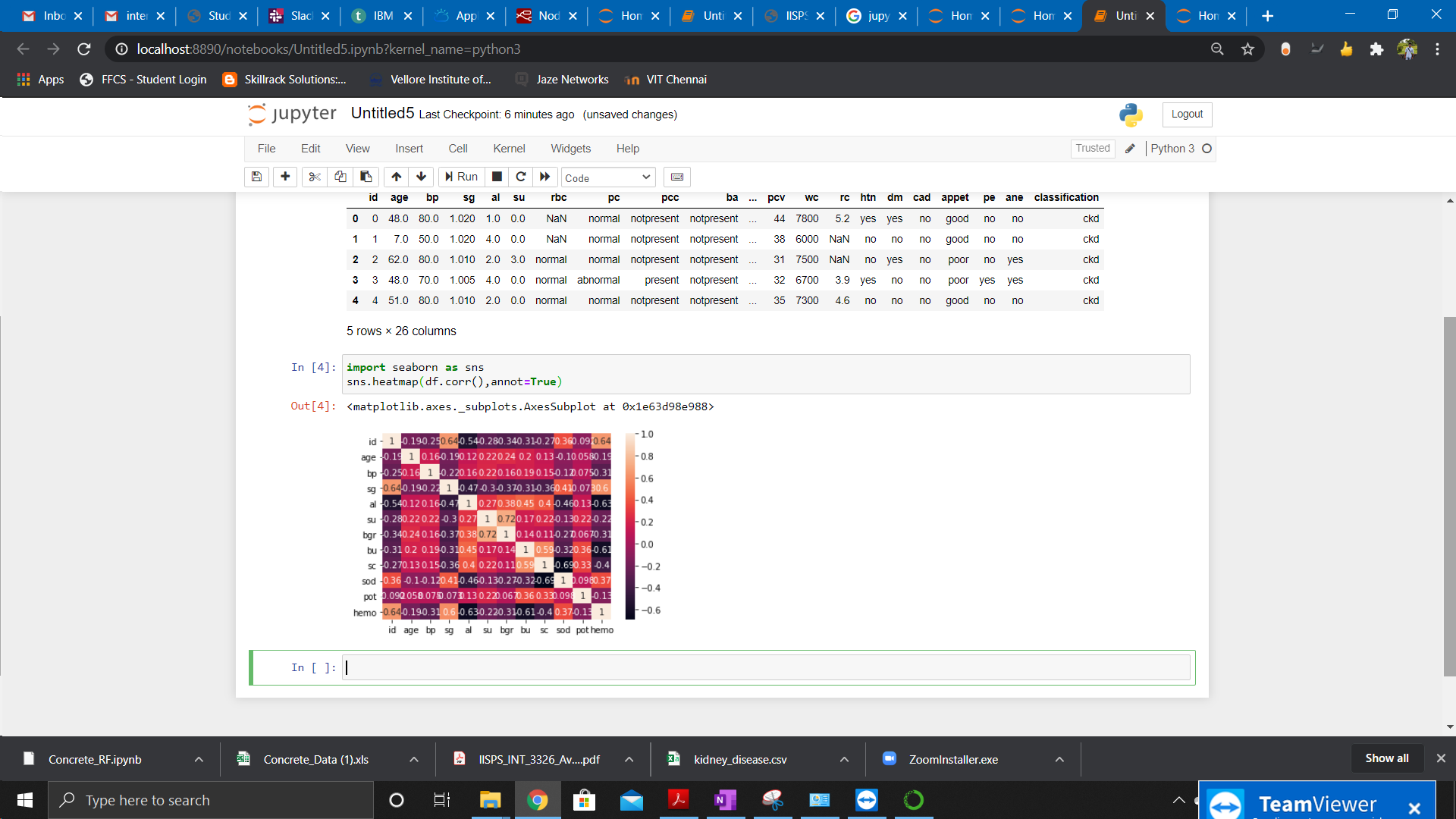
Applied various algorithms in jupyter notebook to the dataset to find the best one.

## **Data pre-processing and analysing:**

### Imported the dataset and find the correlation between the given values. and find if there are any null values.



Generated the correlation in the form of a heatmap for better understanding. Removed the parameters date and time as it does not contribute to the prediction value.

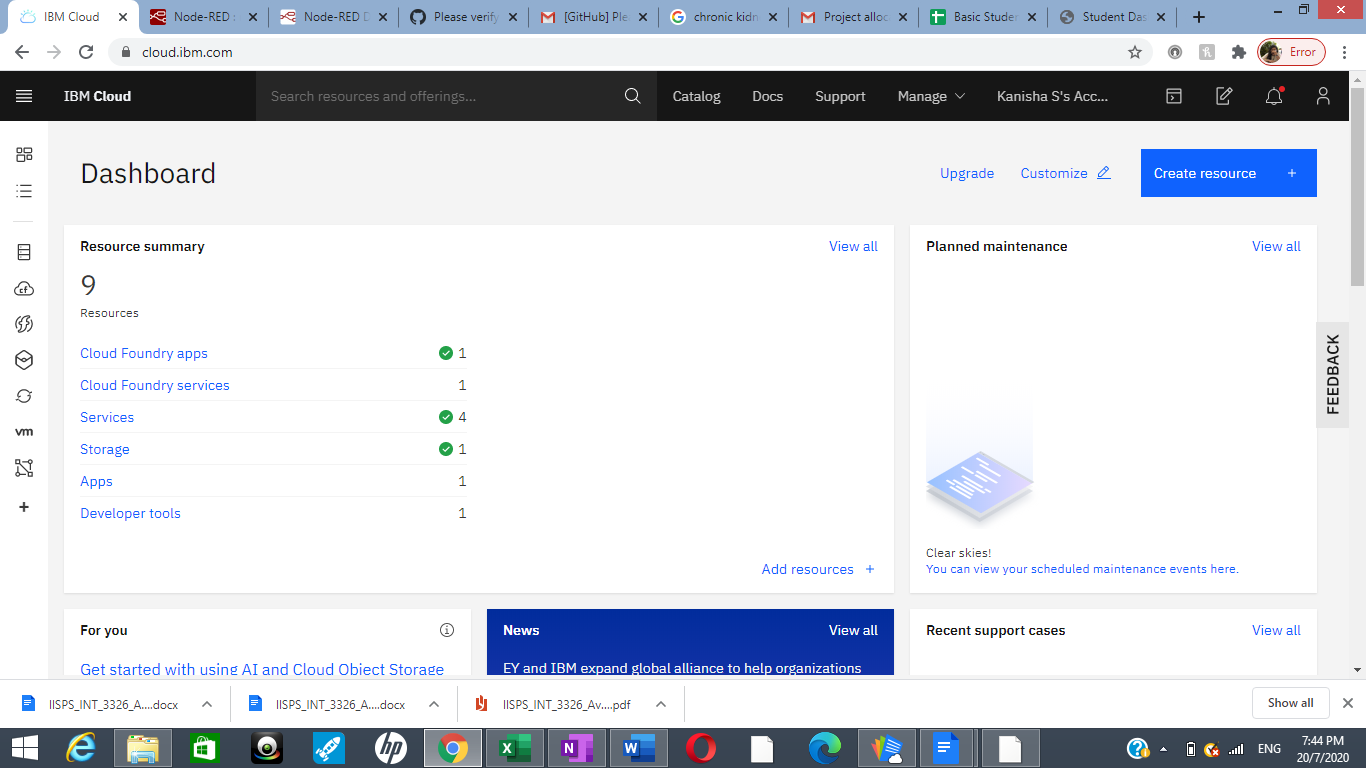


## 

## **Step2:- IBM Cloud Account:**

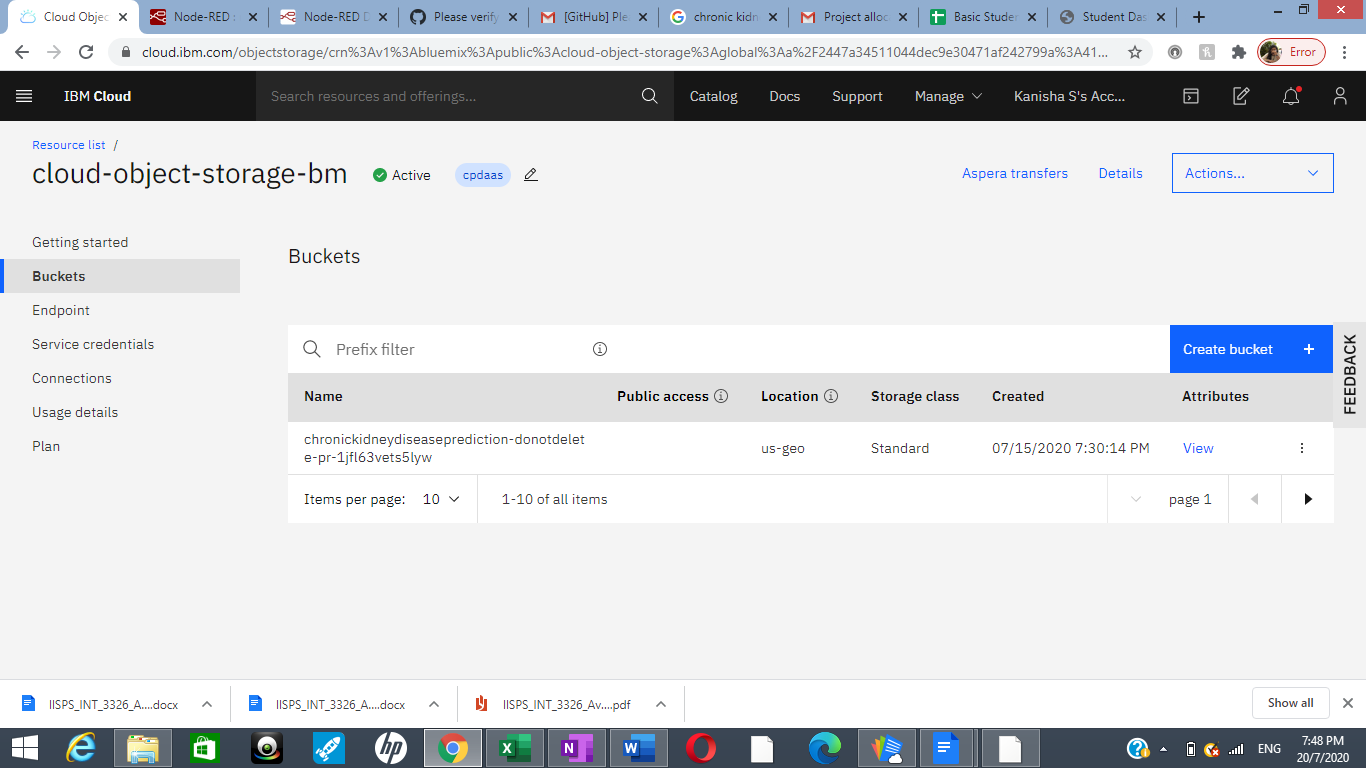
**Creating an account:**

Successfully created an IBM Cloud account.



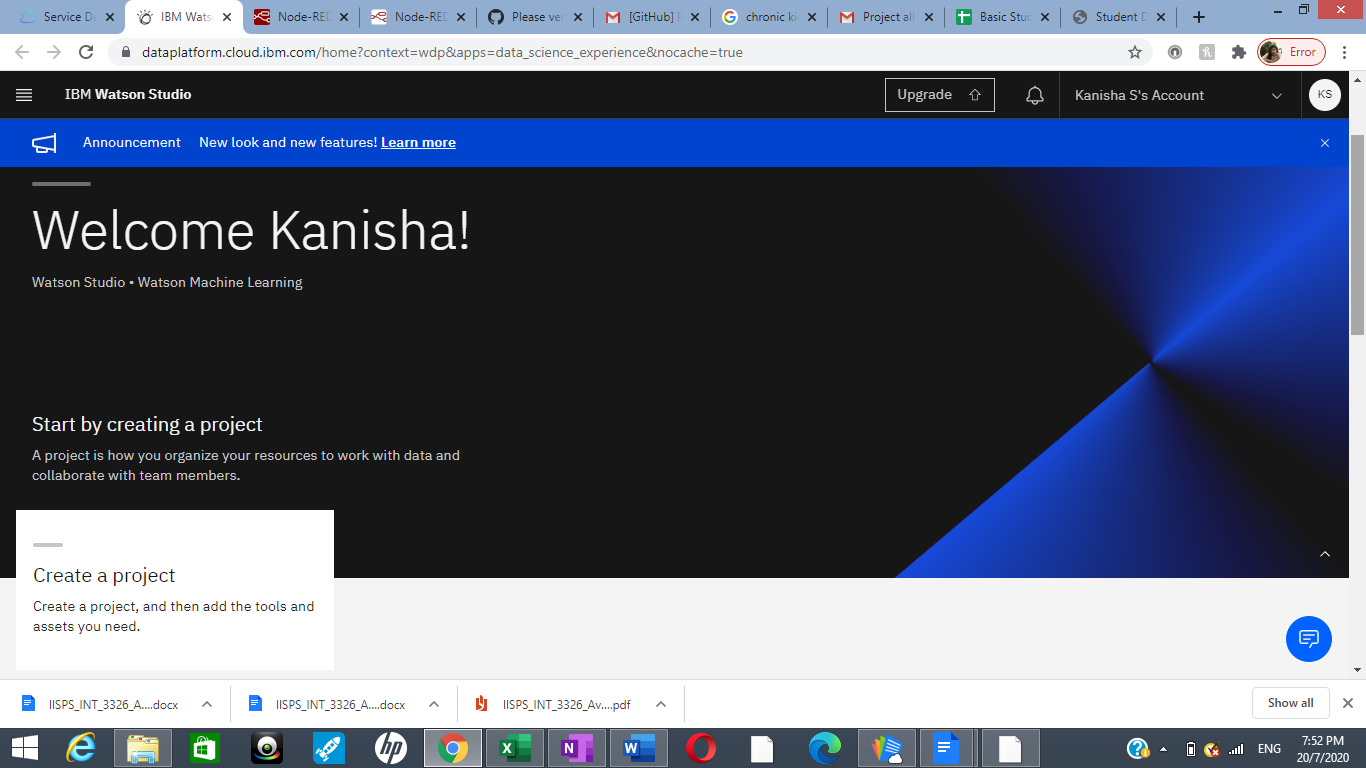
## **Creating cloud object storage:**

Created a storage and have created a bucket to store our projects.



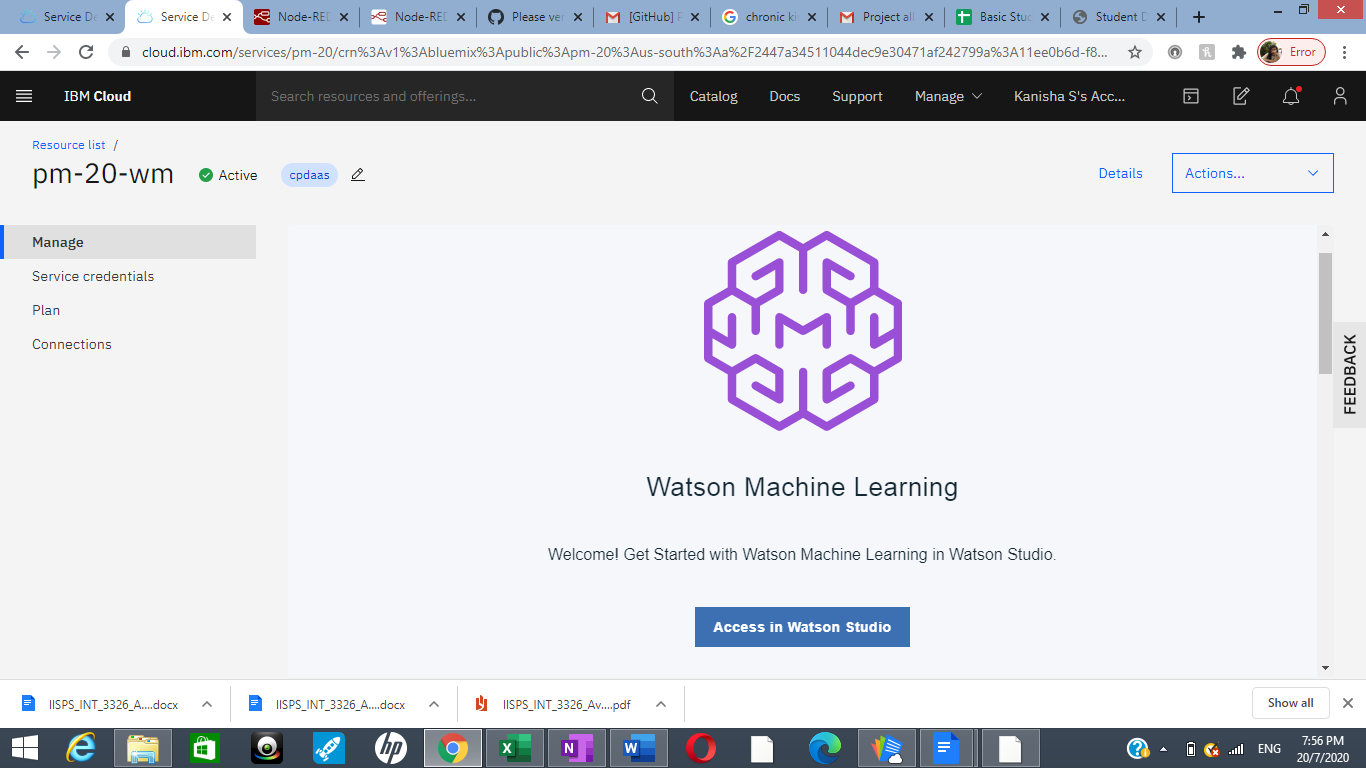
## **Creating Watson Studio Platform**

Created a Watson studio platform to predict our data and to implement, deploy and test our model in the real time.



## **Creating a ML Service:**

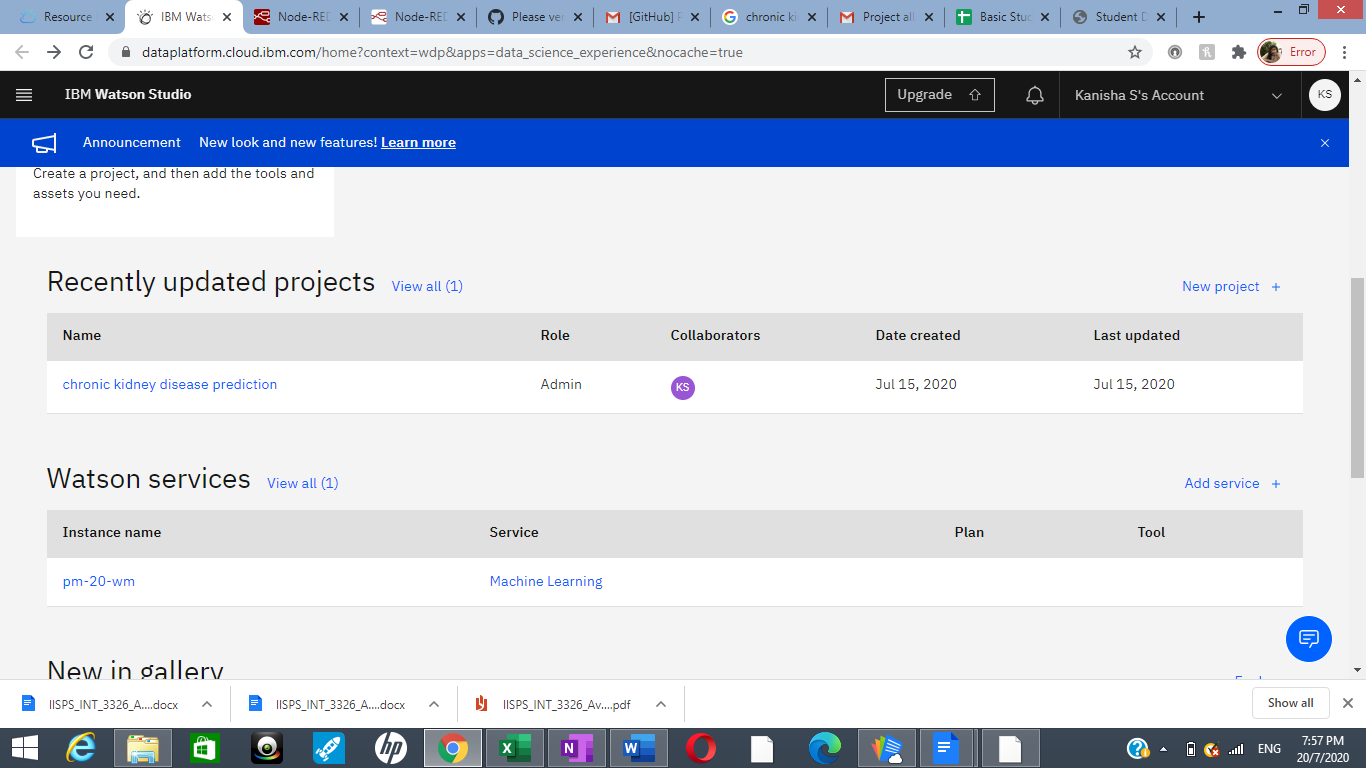
### Created a ML service to load and process our dataset.



**Step3:- Model Building**

**Creating a project in Watson Studio:**

I have created a project in Watson studio “chronic kidney disease prediction”.



IBM Watson Studio Desktop is a desktop client tool for solving your problems by analysing data with artificial intelligence. With Watson Studio Desktop, you can prepare data and build models on your desktop with visual drag and drop tools. You organize your resources for data analysis tasks in projects. Each [project](https://www.ibm.com/support/knowledgecenter/SSBFT6_1.1.0/wsd/projects.html?view=kc&projects) has its own directory on your computer. You can choose a standard project or to import a project that was previously exported from Watson Studio Desktop.

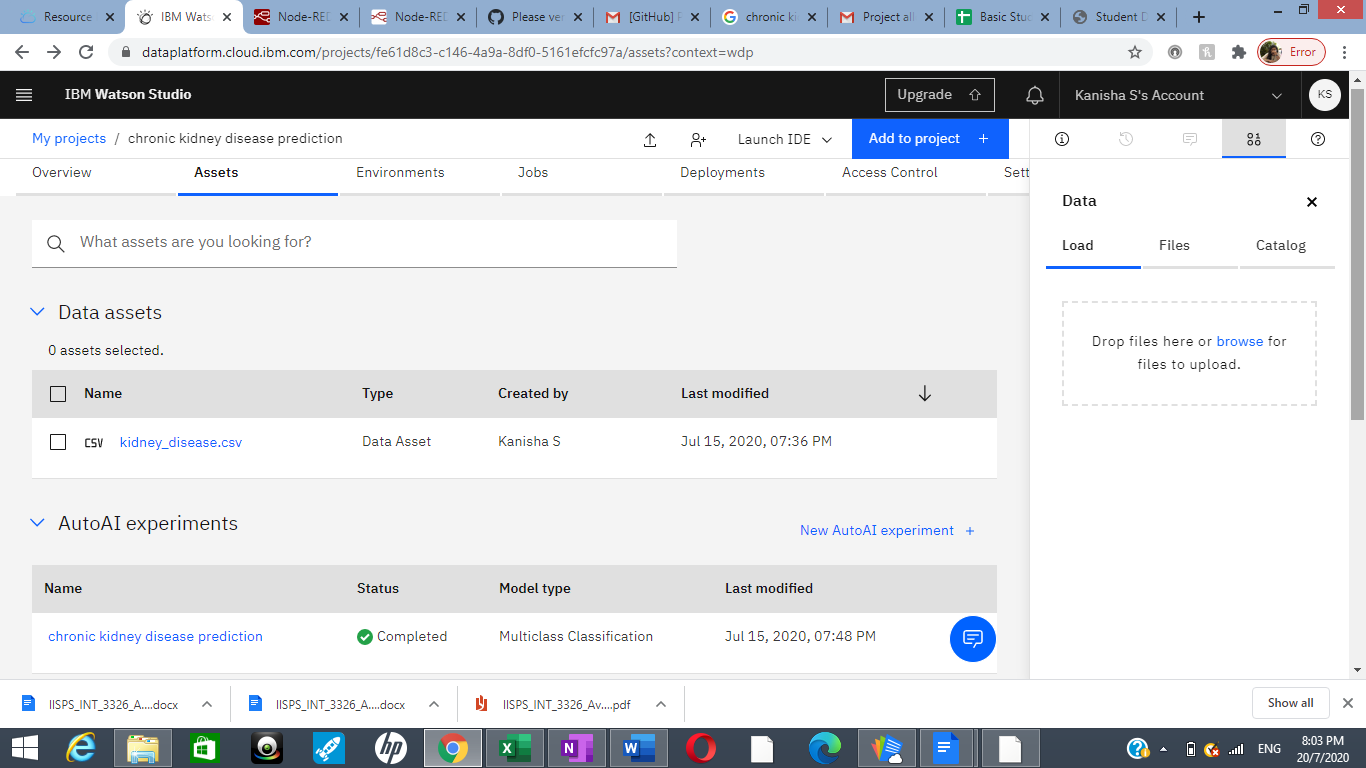
## **Auto AI Experiment in add Projects and set up AI environment:**

I have created an auto AI experiment called Intern. AutoAI is available within [IBM](https://www.ibm.com/in-en/cloud/watson-studio) [Watson Studio](https://www.ibm.com/in-en/cloud/watson-studio) with one-click deployment through [Watson Machine Learning](https://www.ibm.com/in-en/cloud/machine-learning). To help simplify an AI lifecycle management, AutoAI automates:

* + Data preparation
  + Model development
  + Feature engineering
  + Hyper-parameter optimization

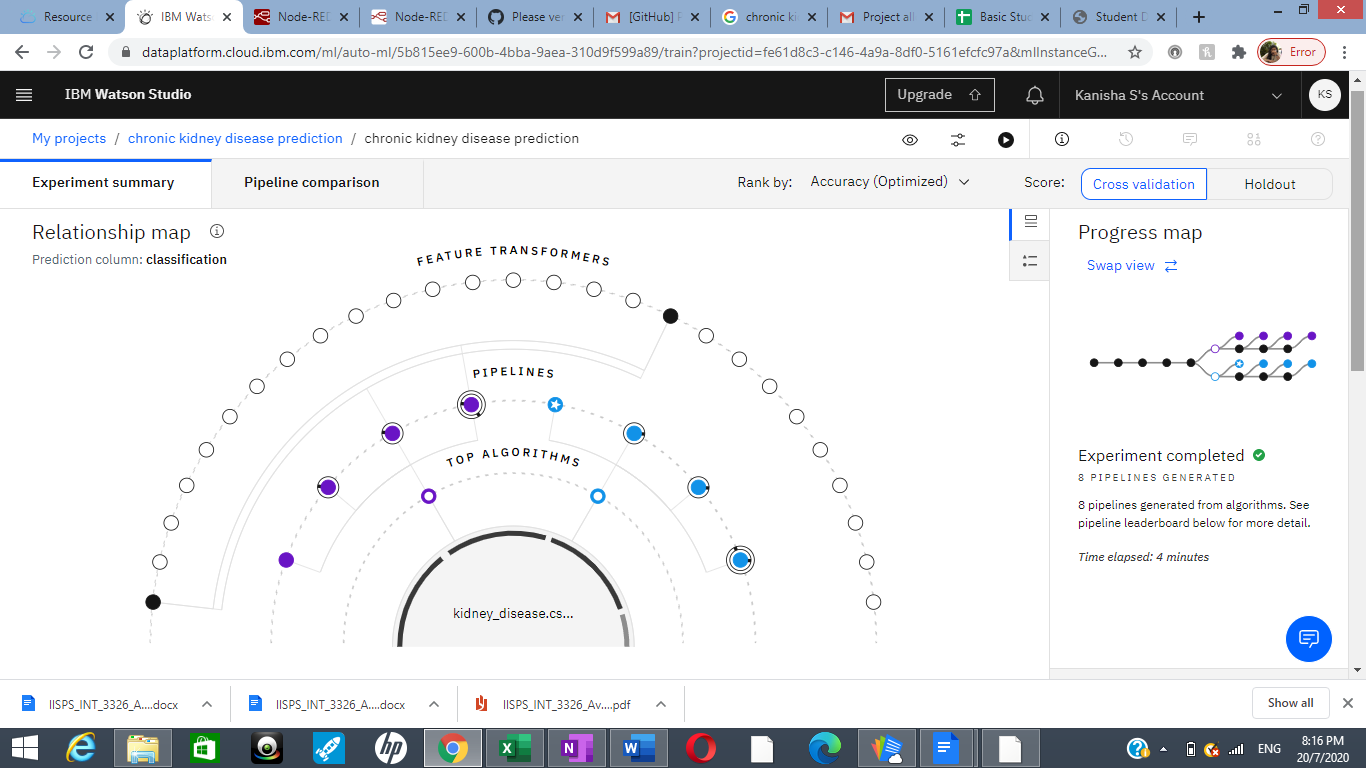
## **Import Dataset:**

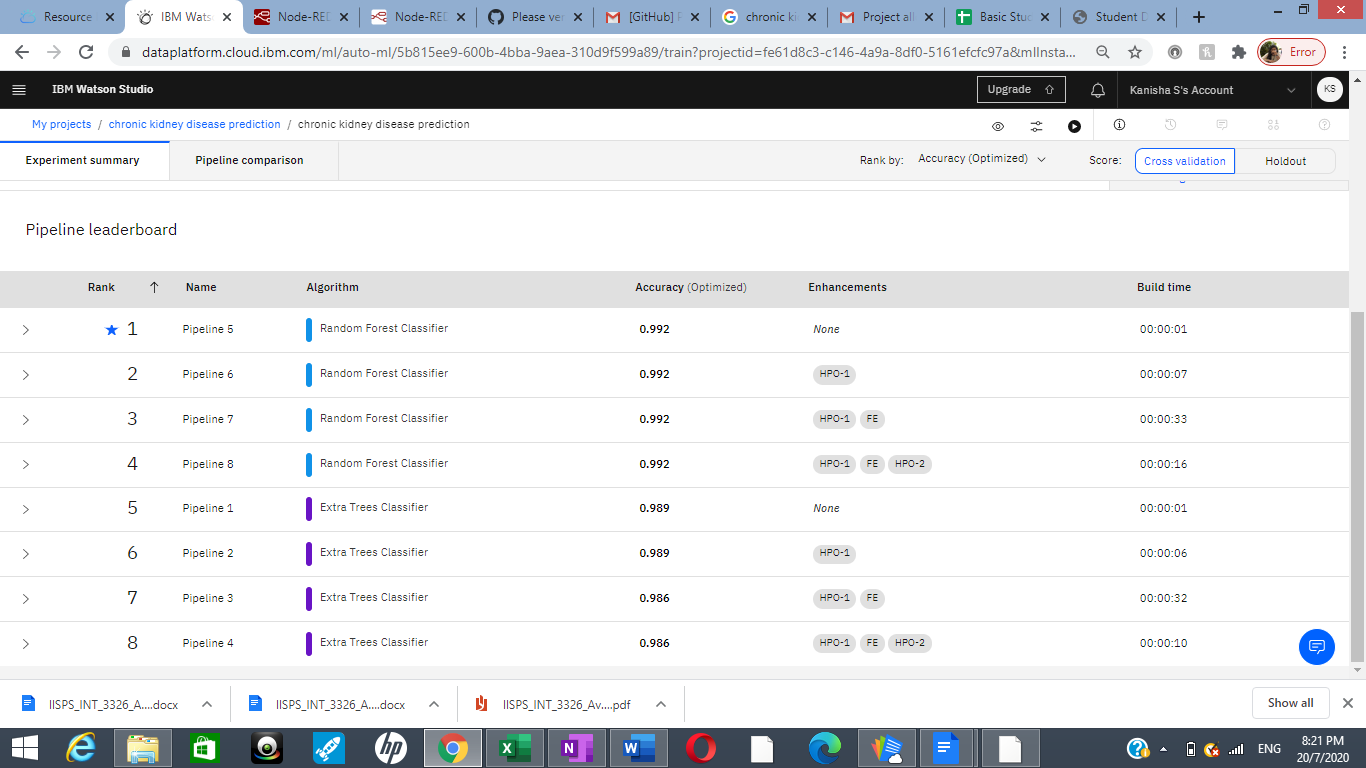
I have then imported the dataset in the name of kidney\_disease.csv in the assests section.



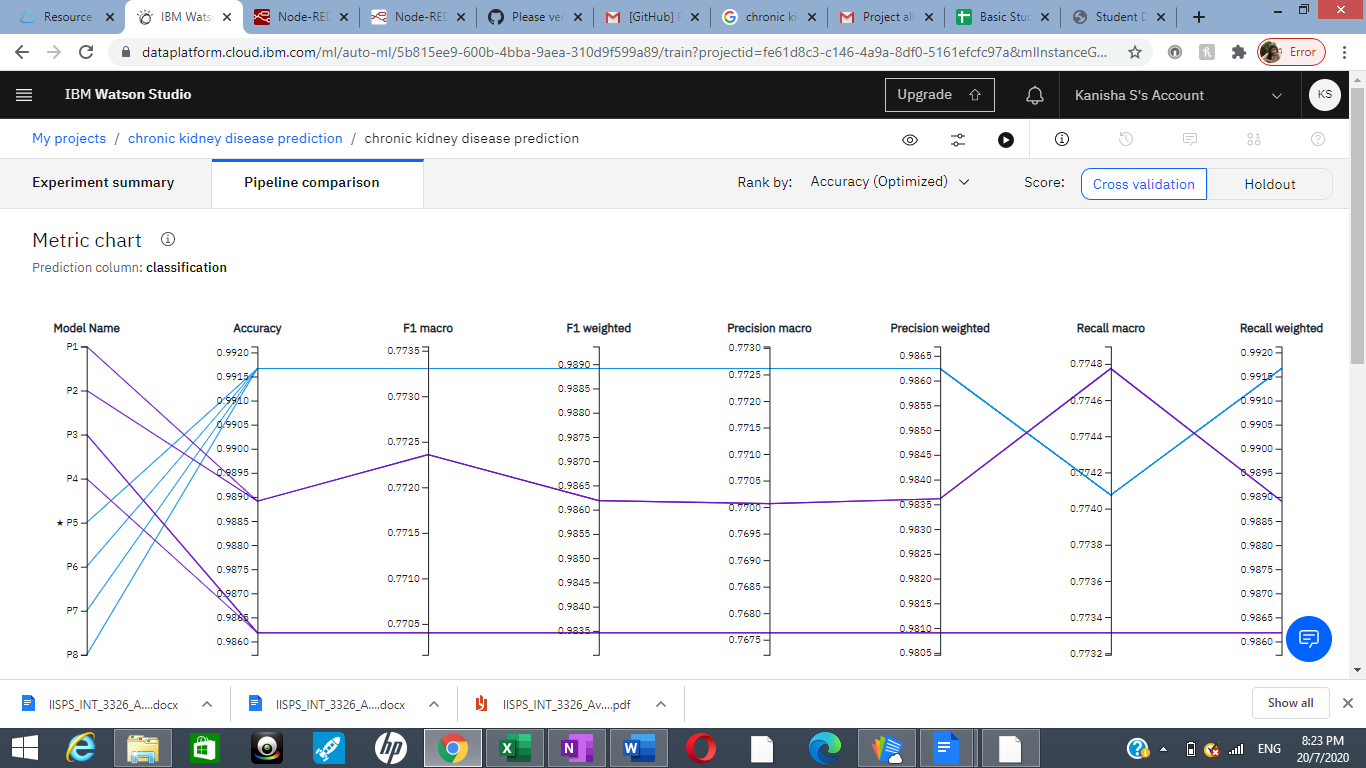
## **Run the model and select the pipeline:**

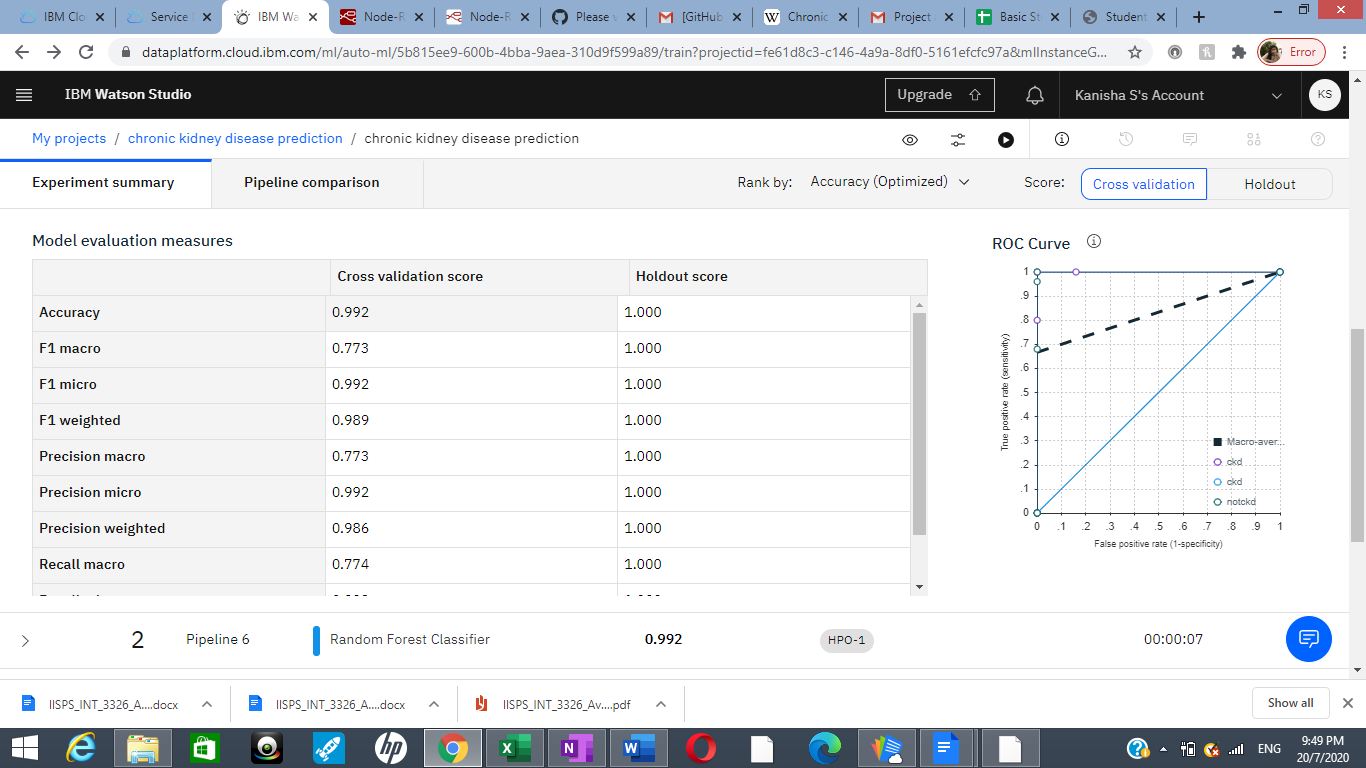
I then run the model, by choosing what to predict, the ratio of train and testing dataset and the number of pipelines to be used. We choose the default 90:10 ratio for training and testing and we chose 8 pipelines.





As we can see, the random forest algorithm ranks first as the best algorithm for the given dataset. It has the most less error, (i.e.) rmse value in comparison to all the other algorithms.

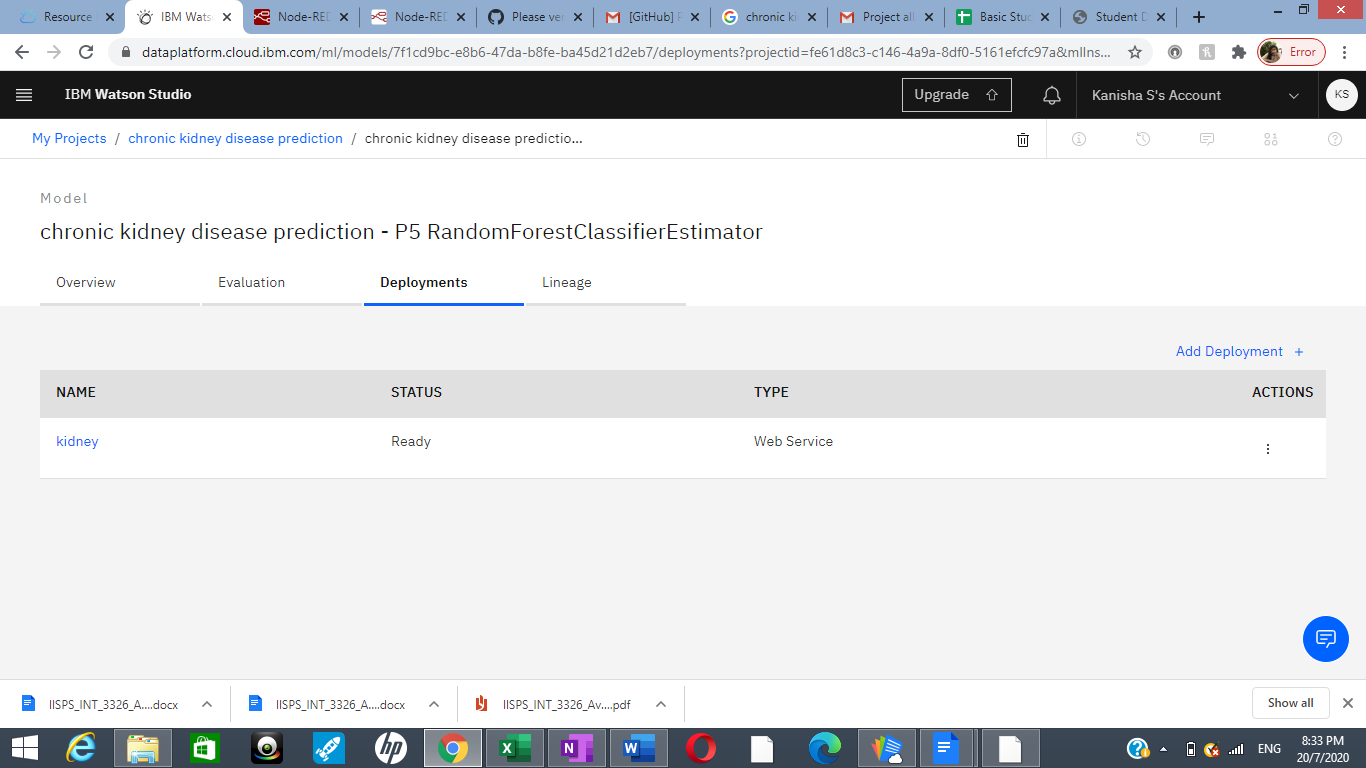




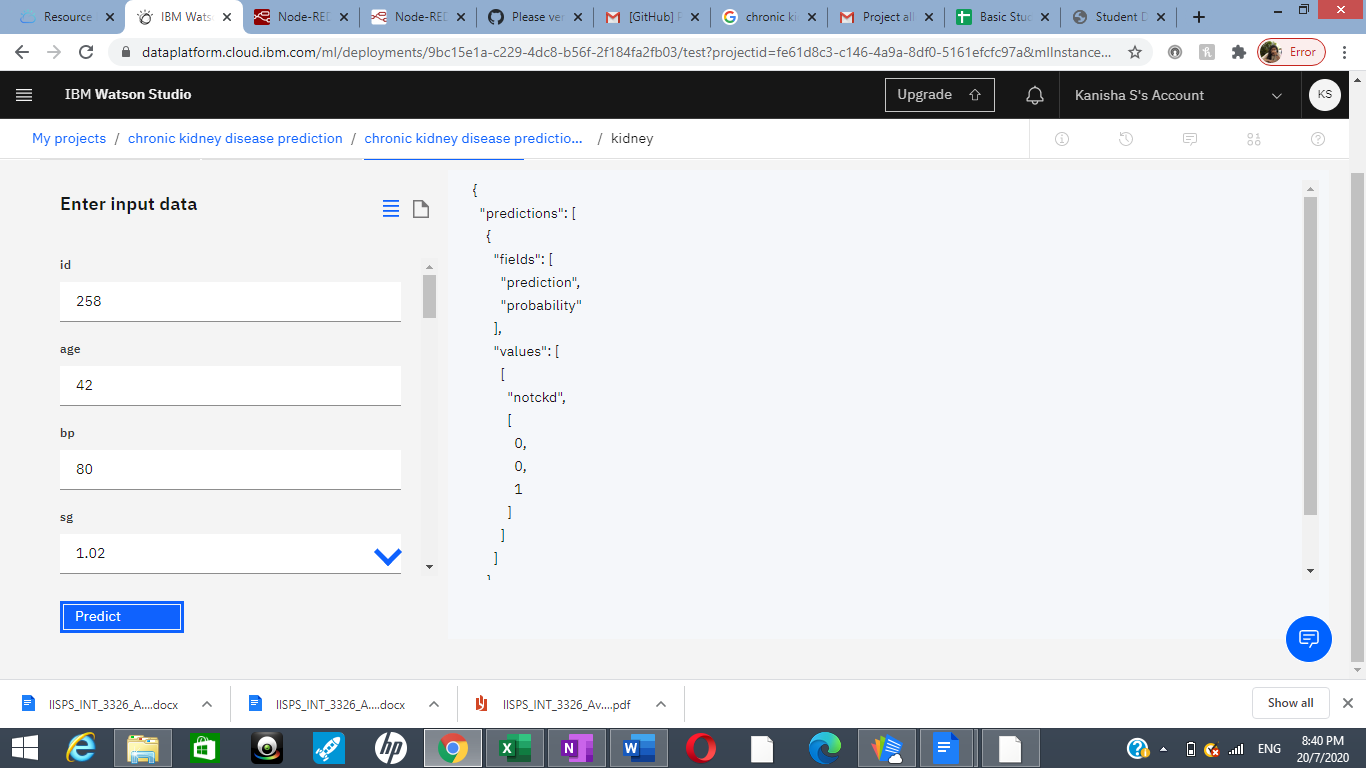
As we can see, this algorithm has a maximum value of 99.2%

## **Deploy and test the model in Watson Studio:**

Then save this model and deploy it in the Watson studio in the name "kidney".



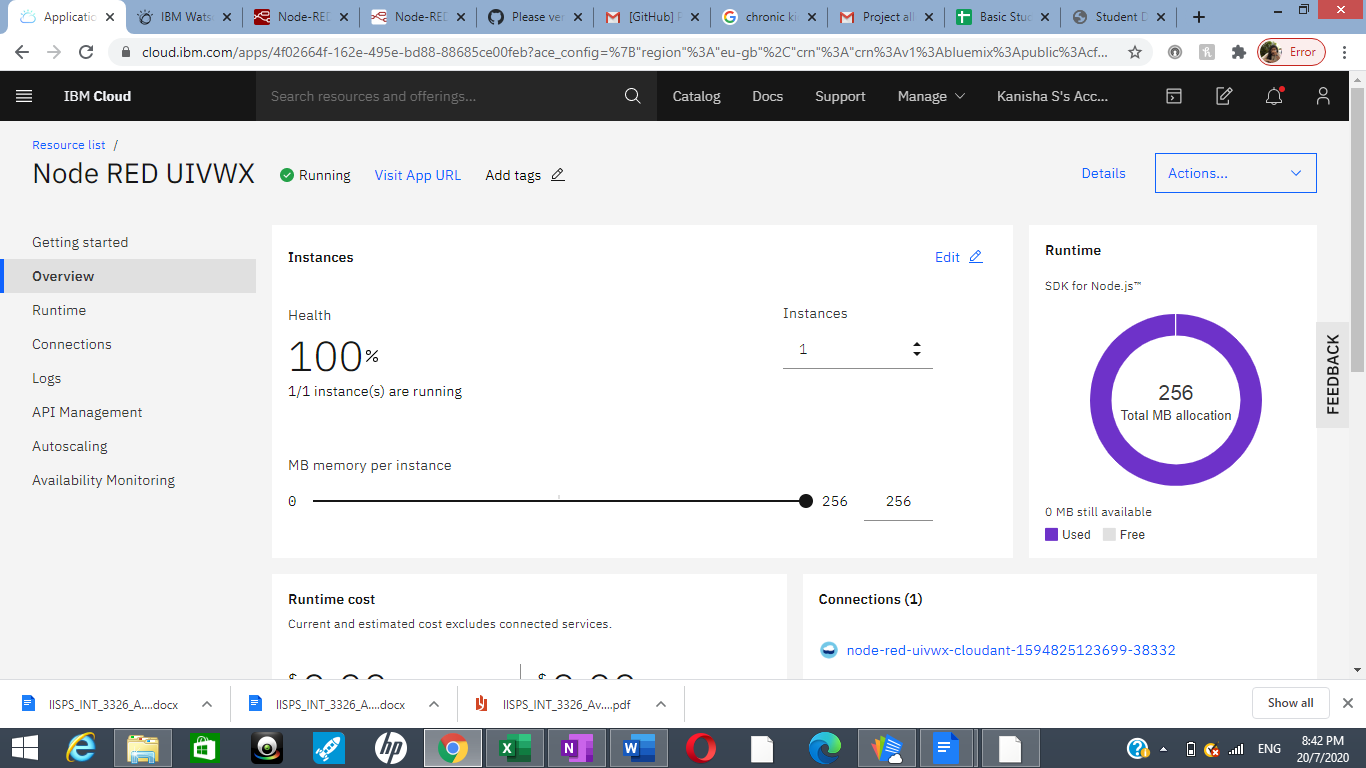
We can see that this model is successfully deployed and ready to implement and test. We need to click on the model and it will direct us to a page where we can find the model overview, implementation and test. We test our model before creating our app.



## **Step 4:- Application building:**

**Create a Node Red service:**

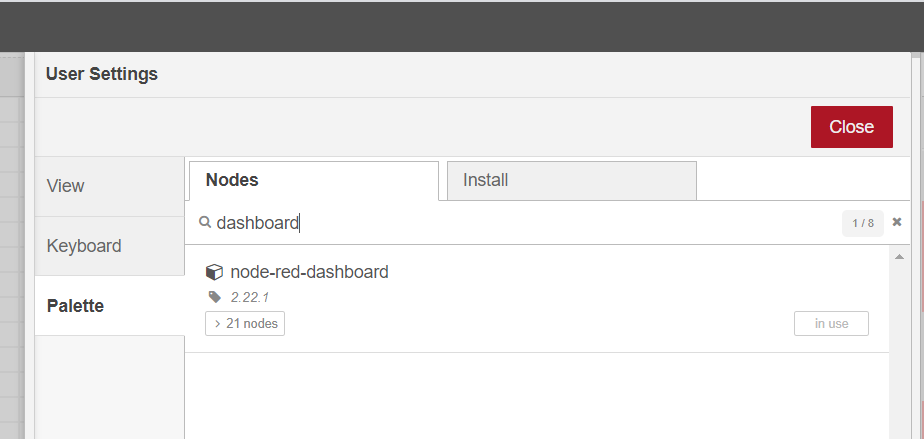
I have created a Node Red app and we can see that the app is running.



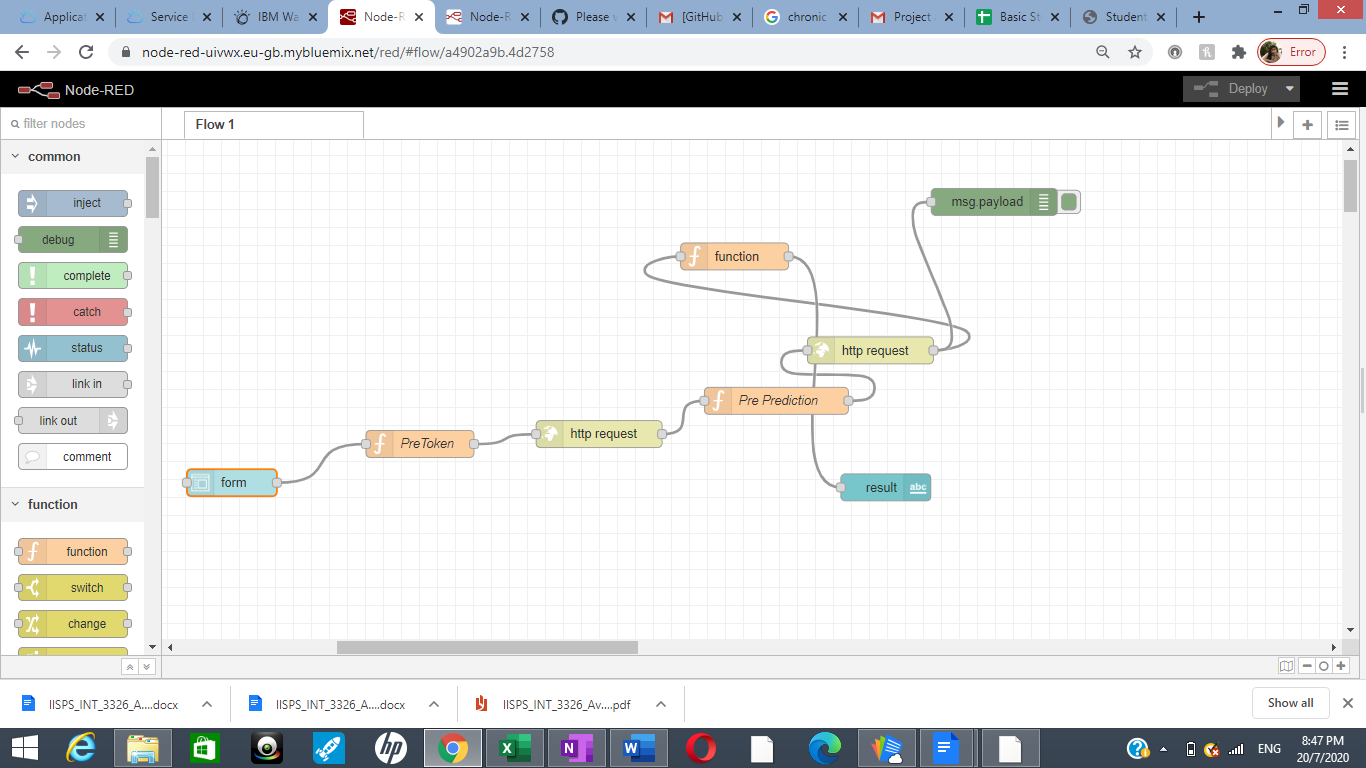
Node-RED provides us a browser-based flow editor that makes it easy for us to wire together flows using the wide range of nodes in the palette. Flows can be then deployed to the runtime in a single-click. JavaScript functions can be created within the editor using a rich text editor. A built-in library allows you to save useful functions, templates or flows for re-use.

## **Install dashboard palette:**

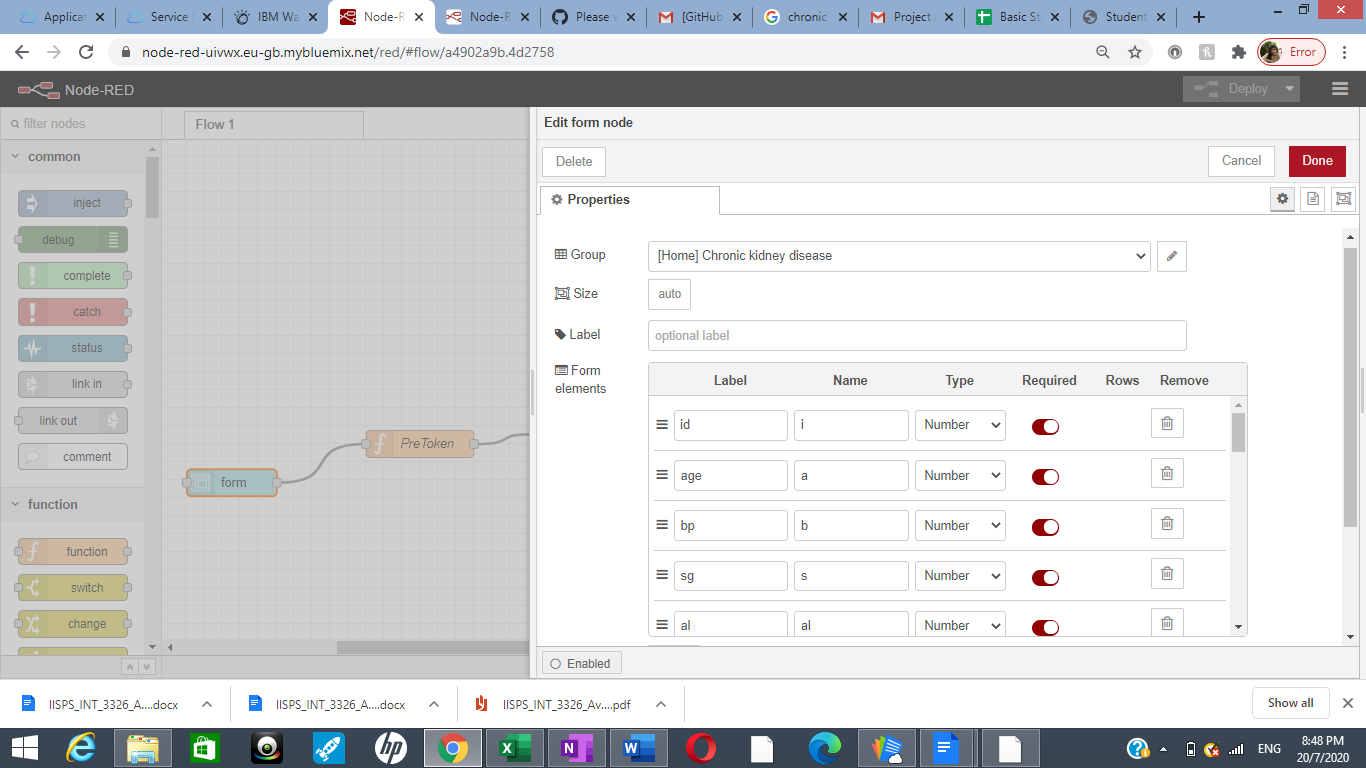
I have installed dashboard palette in the node red app and we use those nodes to build our app.



## **Building UI with Node Red:**

We connect the following nodes in our node red flow.

**Form node :** In the form node, we give the titles and data types of our inputs.



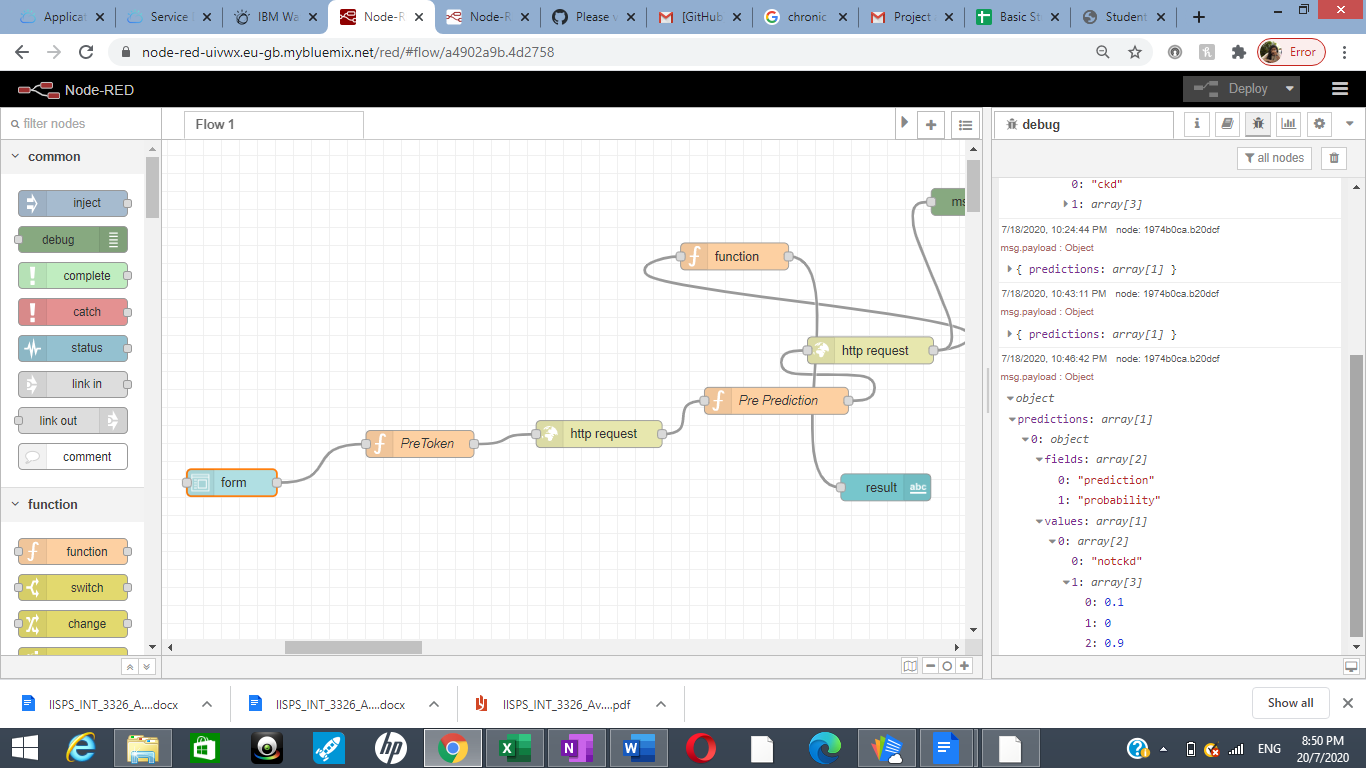
.

**Pre Token:** The pre token is a function node. A JavaScript function to run against the messages being received by the node. The messages are passed in as a JavaScript object called msg. We link the api key of our deployment in this node. We get the input values from the user for the input parameters needed and then pass it on to our next node.

**Http request:** This node sends the http request and returns the response. The body of the response. The node can be configured to return the body as a string, attempt to parse it as a JSON string or leave it as a binary buffer.

**Pre Prediction :** The pre prediction node is also a function node. This node links our instance id to access the deployment of our model. The msg.payload in the code sends our fields as a dictionary format to our output node.

In our next http node we link the url of our app and in the next function node we link the path of our predicted output value from the debug message part. This helps us to view our output in our web page.



### **msg.payload :** This node displays the value of our prediction (i.e) the depth of the snow in our case. This node displays selected message properties in the debug sidebar tab and optionally the runtime log. By default it display msg.payload, but can be configured to display any property, the full message or the result of a JSON data expression.

**Deploy the app and run:**

Then deploy our app and we load our ui web page.

### This is the web page, we then enter the values to predict the ckd.

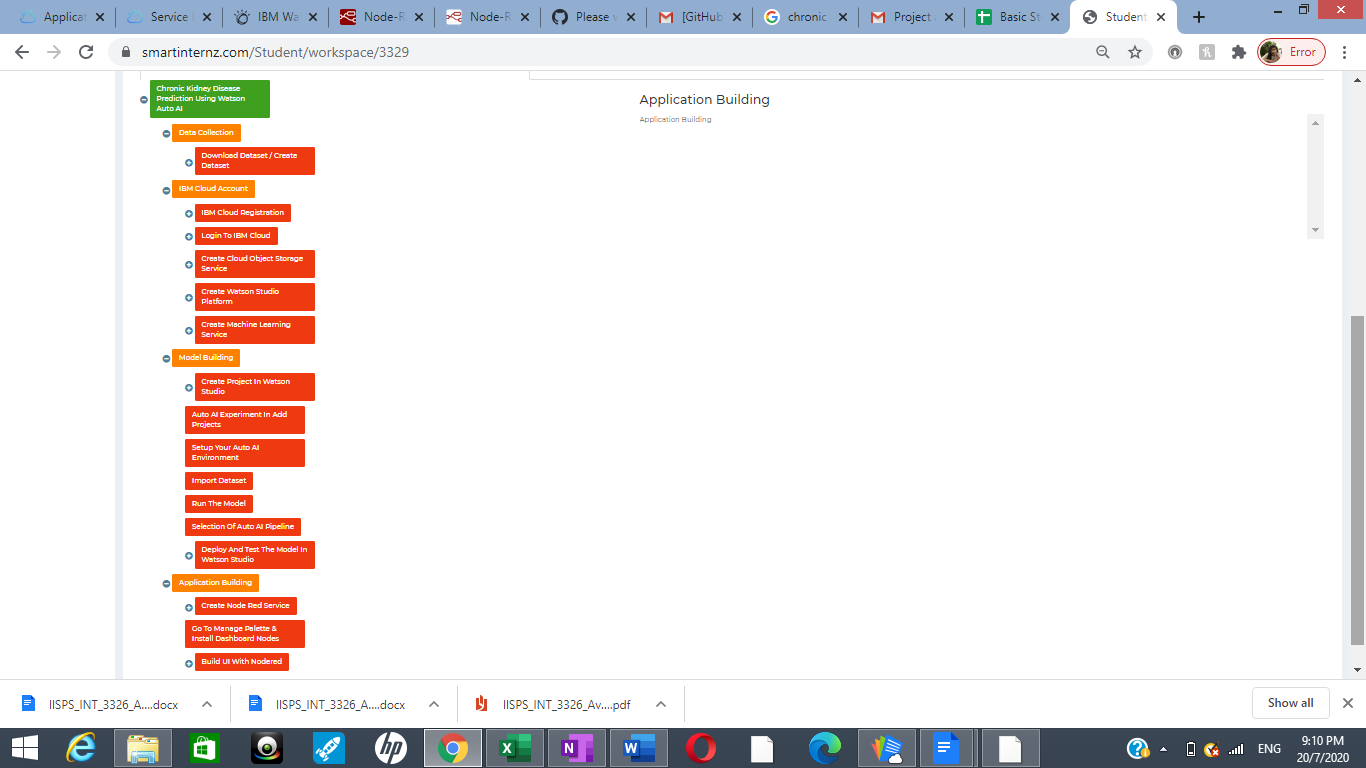
### 

### 

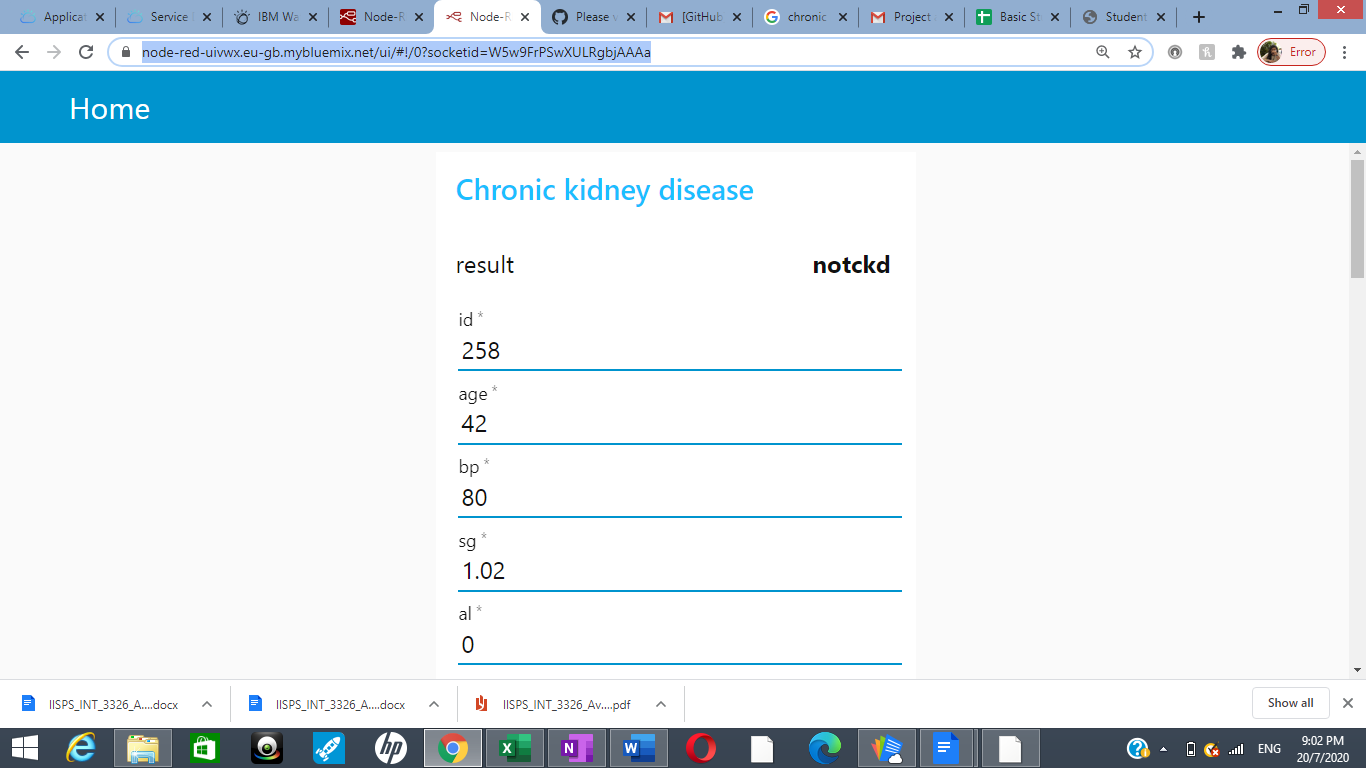
**The url for the project is:**

<https://node-red-uivwx.eu-gb.mybluemix.net/ui/#!/0?socketid=W5w9FrPSwXULRgbjAAAa>

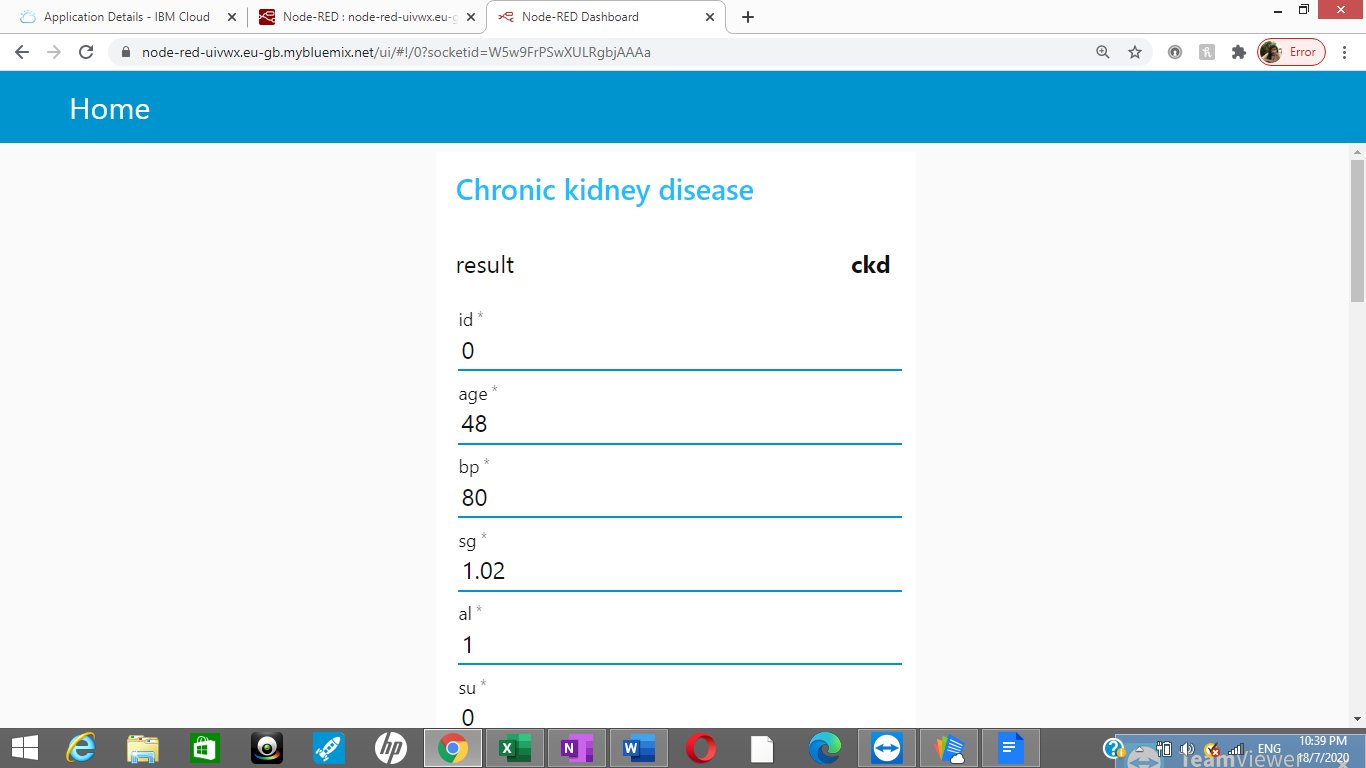
# **FLOW CHART**



1. **RESULT**



As we can see, our app gives the output result if the patient has ckd or notckd



# **ADVANTAGES AND DISADVANTAGES**

## **Advantages:**

* + - The chronic kidney disease prediction helps the people to test their status .
    - Our web app predicts the kidney disease with a great accuracy.
    - Our web app is easy to access and provides us a good user interface.
    - This helps in predicting the disese and also tells if you are free from the disease.
    - The app gives results immediately without any delay.

## **Disadvantages:**

* + - The user entering the details should know the details of the input parameters iaccording to all the previous tests taken for hemo rc, wc etc.
    - It would be a little conucing as it requires to fill 25 parameters accurately.

# **APPLICATIONS**

# This app can help theperson to be able to see the result without further delay.

# This might save the life of the people.

# This app will also help doctors to predict it immediately.

# This app will create an awareness among the people regarding the kidney disease.

# **CONCLUSION**

# This chronic kidney disease prediction helps people to see the results immediately if the have the accurate values for all the attributes needed. This model helps in reducing death rates due to ckd which is very high in our country. The accuracy rate is very high in our model. This has an accuracy of 99.2 which is quite high compared to other preduction models. This prediction model helps both patients and doctors much more comfortable.

# **FUTURE SCOPE**

# This app can be furthur developed with much more features like prediciton if the paient need immediate treatment or a dialysis or any oher treatment making the docoto’s job much more easier.

# **BIBILOGRAPHY**

# <https://en.wikipedia.org/wiki/Chronic_kidney_disease>

# <https://www.kaggle.com/mansoordaku/ckdisease>