SMARTINTERNZ
INTERNSHIP REPORT

HEALTH INSURANCE COST PREDICTION USING IBM WATSON AUTO-AI

T S VENKATA SUBRAMANIAM
B.TECH ELECTRONICS AND COMPUTER(3RD YR)
VELLORE INSTITUTE OF TECHNOLOGY CHENNAI

I. INTRODUCTION

Problem statement :

Health Insurance Cost Predictor

Overview :

A predictive and analysising machine learning model is built with the help of IBM watson studio, AUTOAI, and node red.

Purpose :

To build a machine learning model that helps the health insurance companies to provide premium offers to the customers based on certain factors and statistics.

II. LITERATURE SURVEY

Existing Problem :

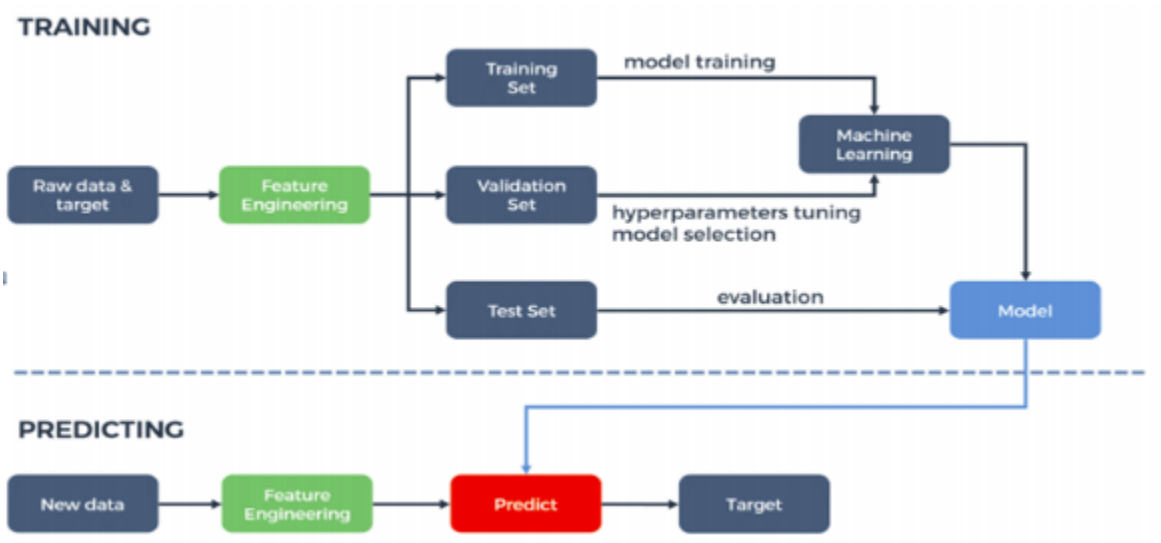
Health insurance companies face a problem in determining the premium insurances for their customers. They need a trained model which predicts what kind of service is suitable to which customer. The companies should also follow certain rules set by the health care law in that country. In order to determine premium offers the companies have to consider certain factors and also go statistically to give importance to a customer.

Proposed Solution:

The main aim of this project is to provide the health insurance companies with a machine learning model using the IBM WATSON AI platform that can predict the customer's eligibilty for different premium insurance services based on certain factors like age, gender, BMI, weight, previous health issues,the region he/she belongs to, if he is a smoker or not etc. The model is tested based on the accuracy and performance of the model.

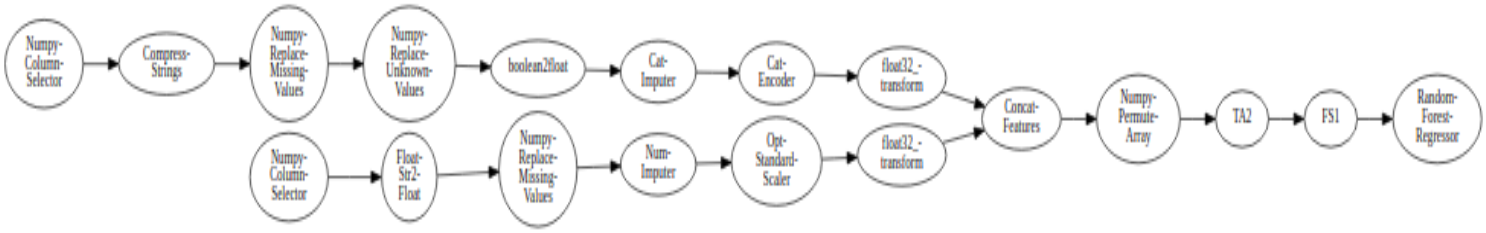
III. THEORETICAL ANALYSIS

Block Diagram :



Hardware-Software Designing :

- Machine Learning algorithms using python
- IBM watson studio, AutoAI and Nodered platform



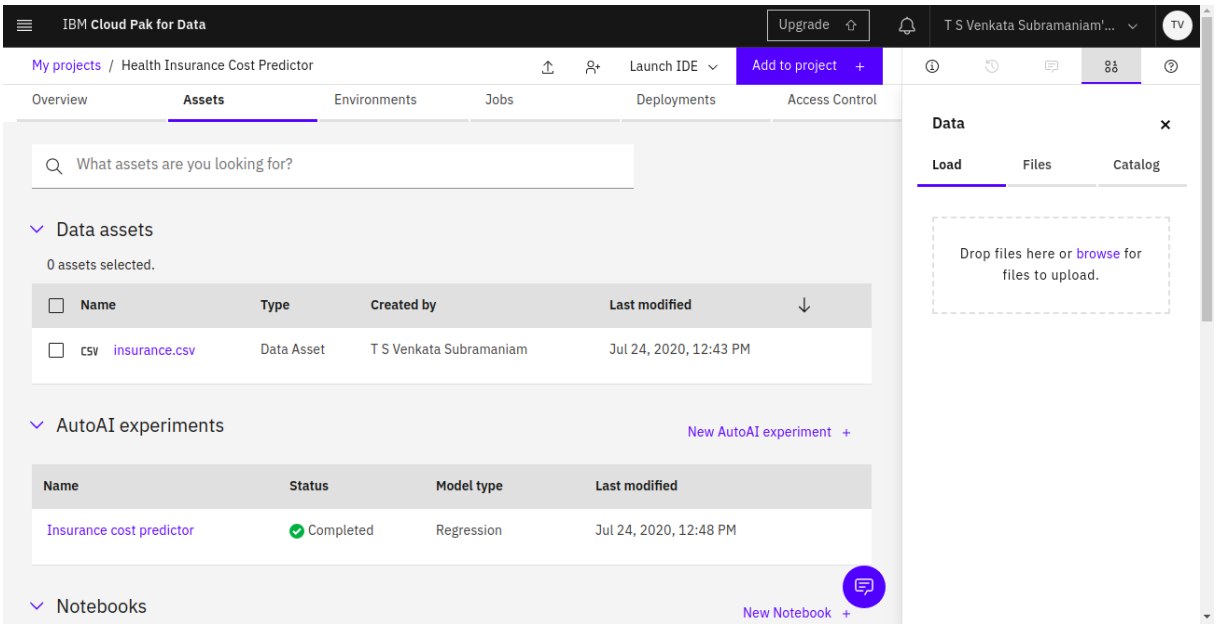
IV. EXPERIMENTAL INVESTIGATION

1. Collection of dataset from kaggle.
   (https://www.kaggle.com/annetxu/health-insurance-cost-prediction)

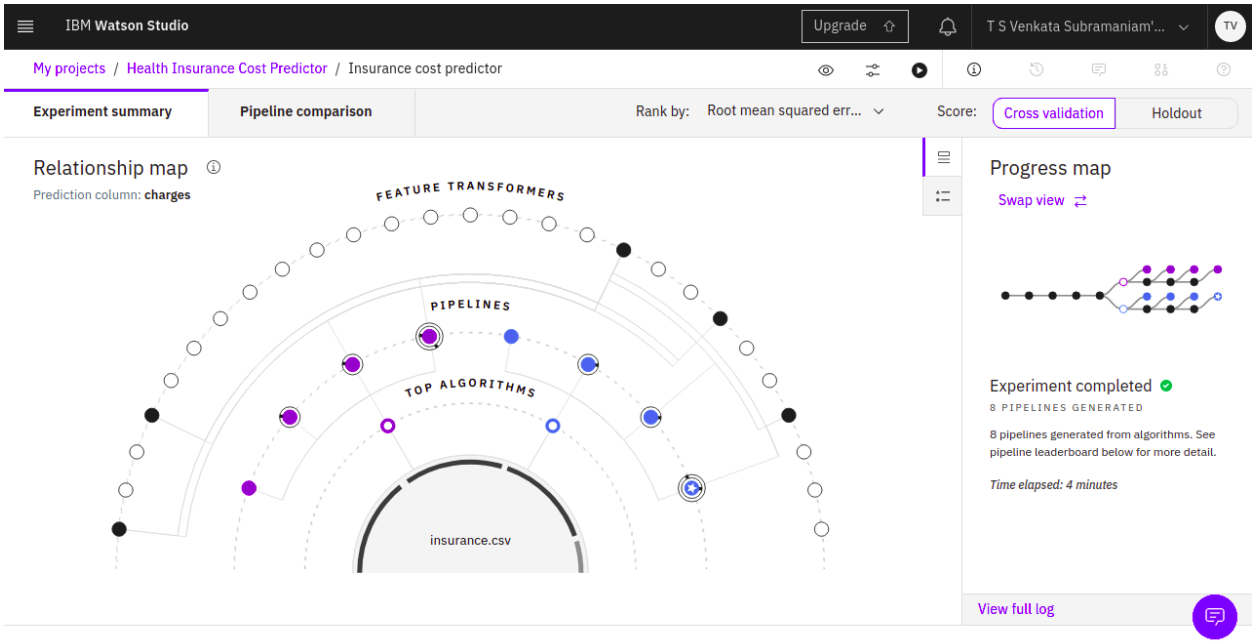
2. Create an account for IBM Cloud platform.

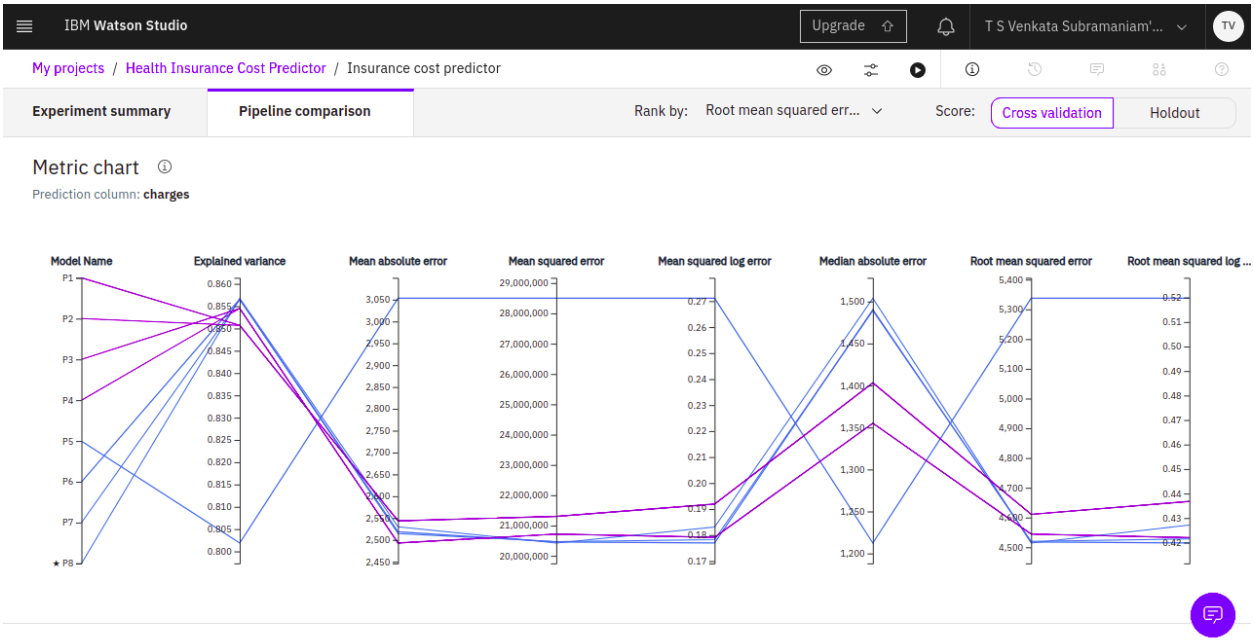
3. On the IBM Watson studio platform use the Auto AI to build a model that predicts the cost of the insurance that can be delievered to a company based on the details given by their customer.

a. First we need to create an account on the IBM Watson studio.
b. Using add to project choose Auto AI
c. Then upload the dataset that is taken from kaggle into the data assets

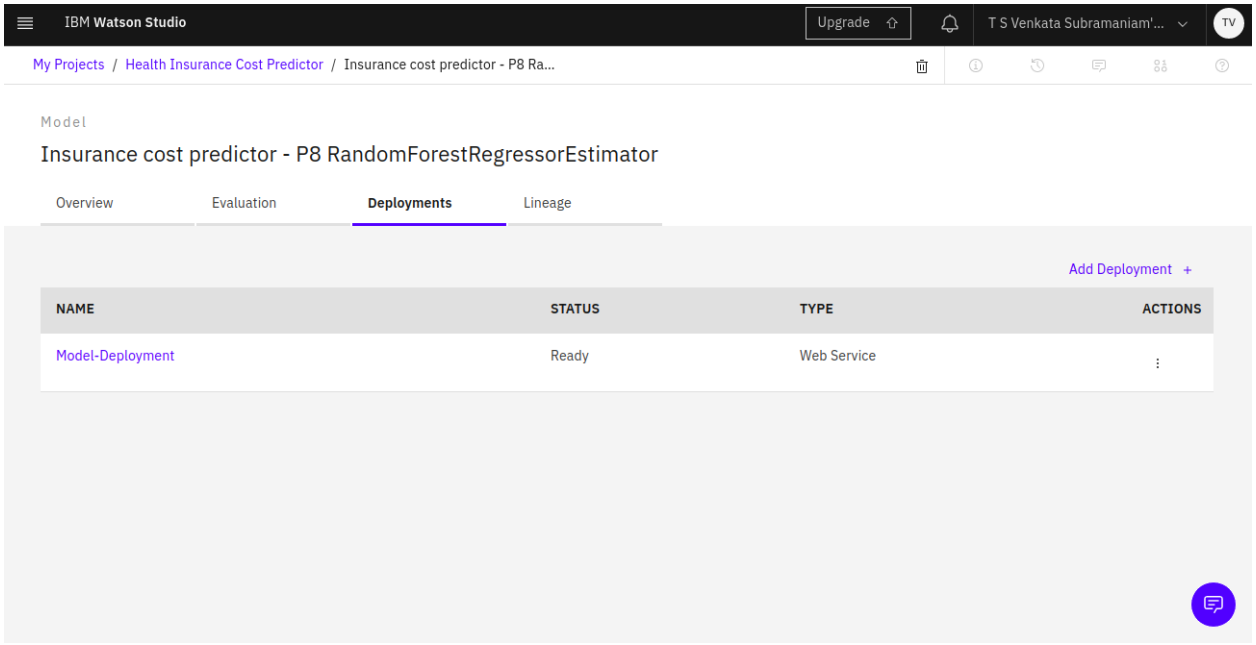


4. Choose the best machine learning model to predict the cost.

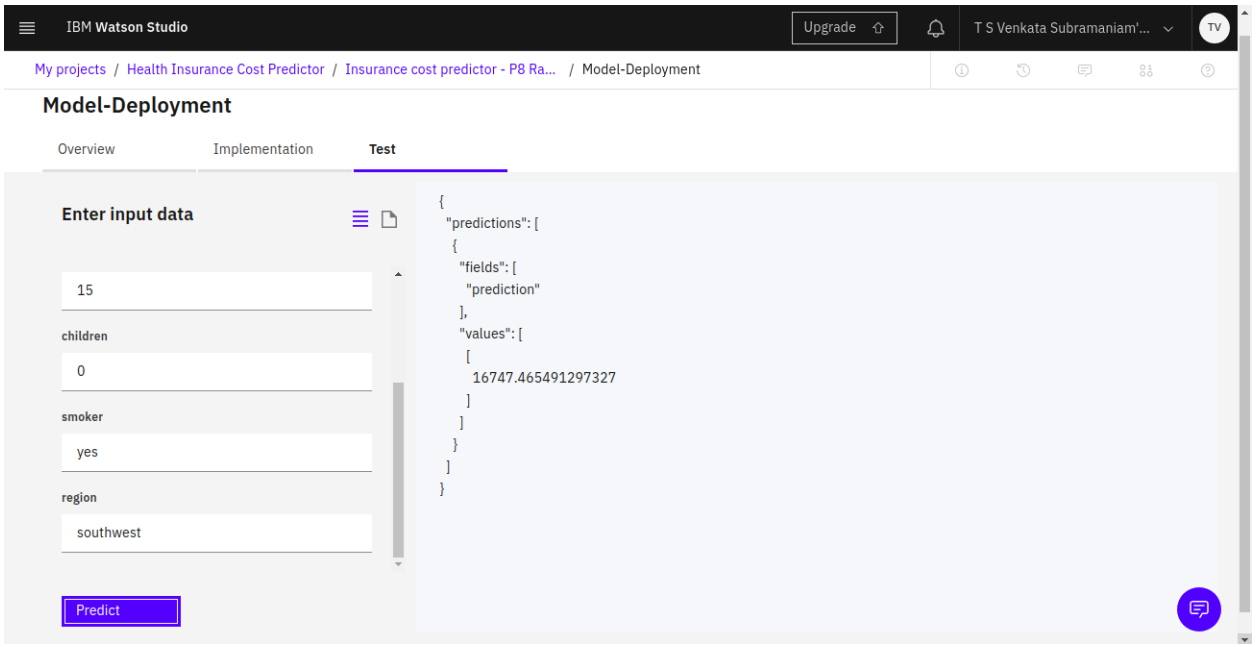




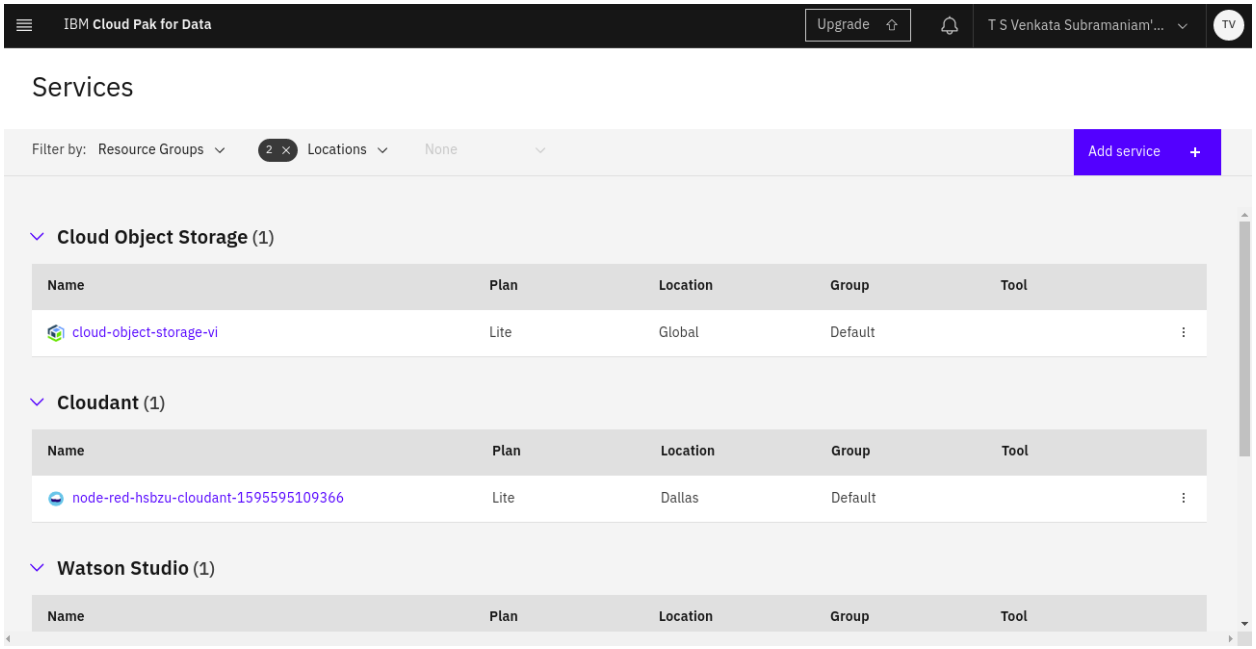
5. Deploy the model.



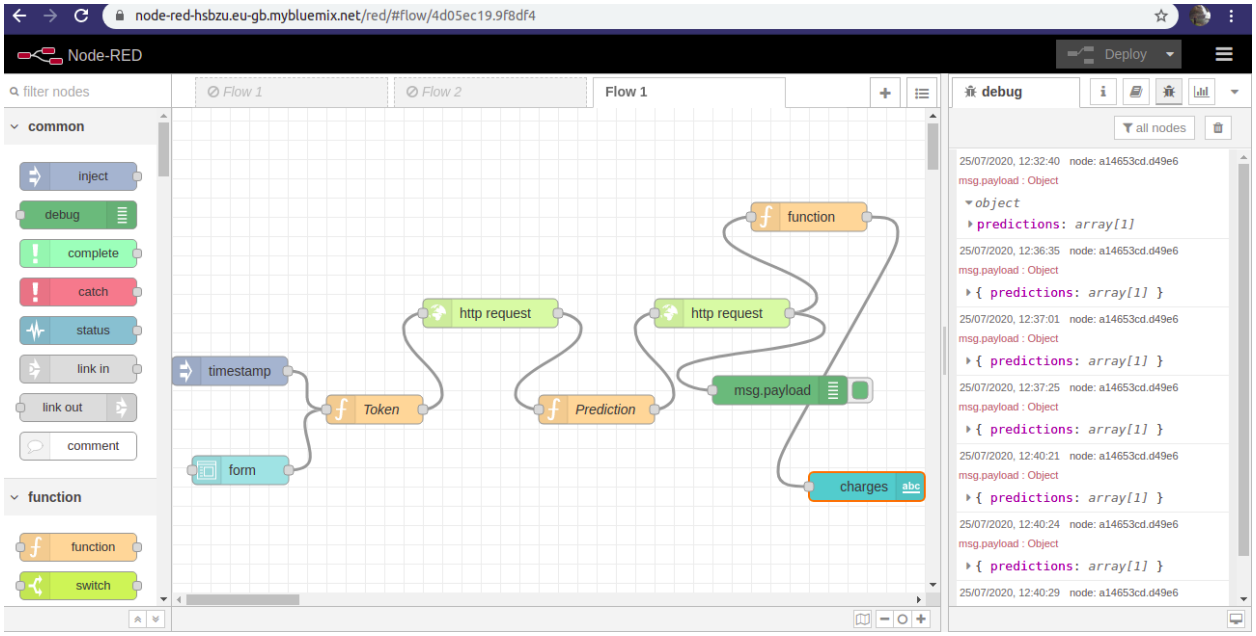
6.Test the model against various values.



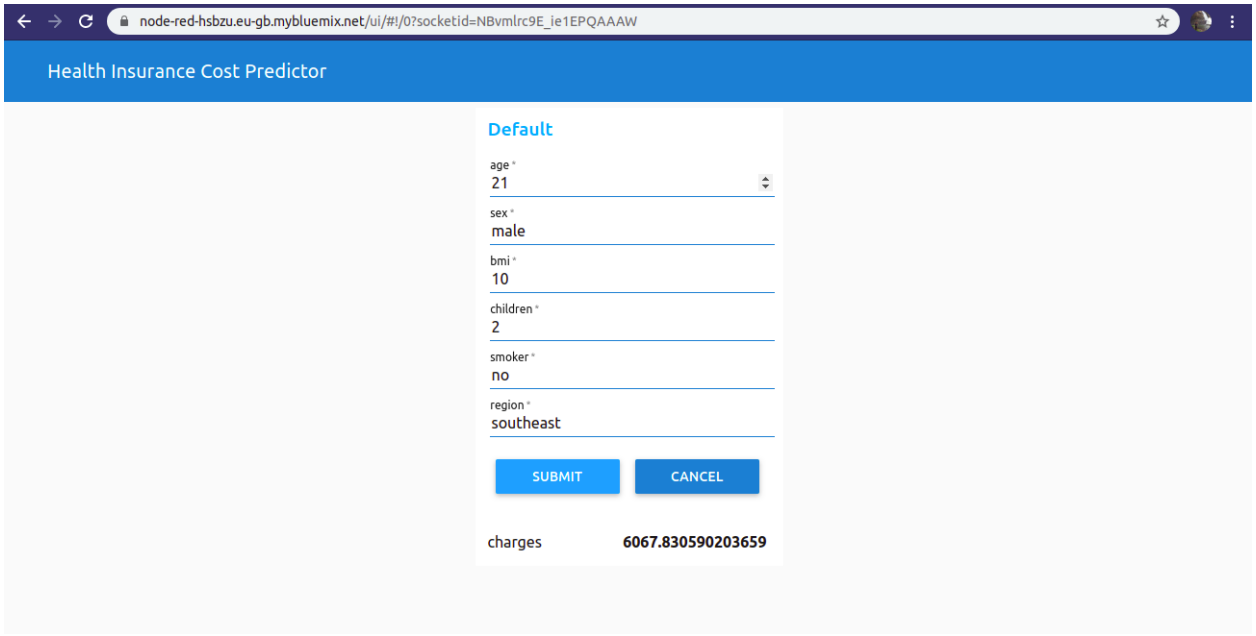
7.Then create a service credential and also cloud foundry app.



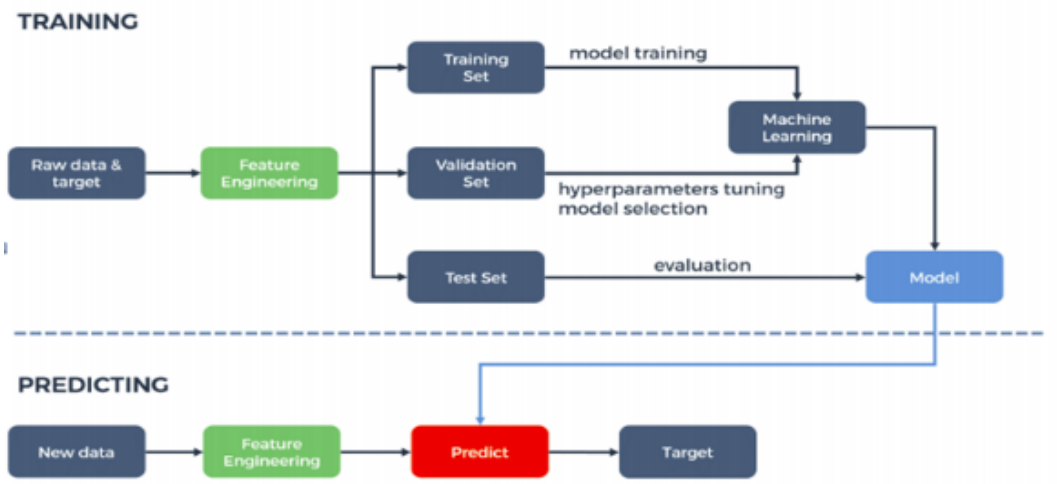
8. Make a node-RED flow.

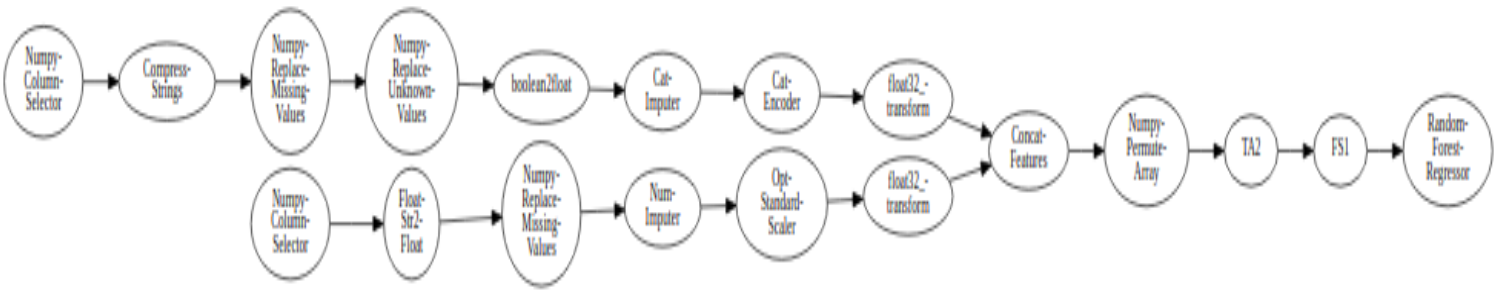


9.After the deployment of the model UI can be seen.



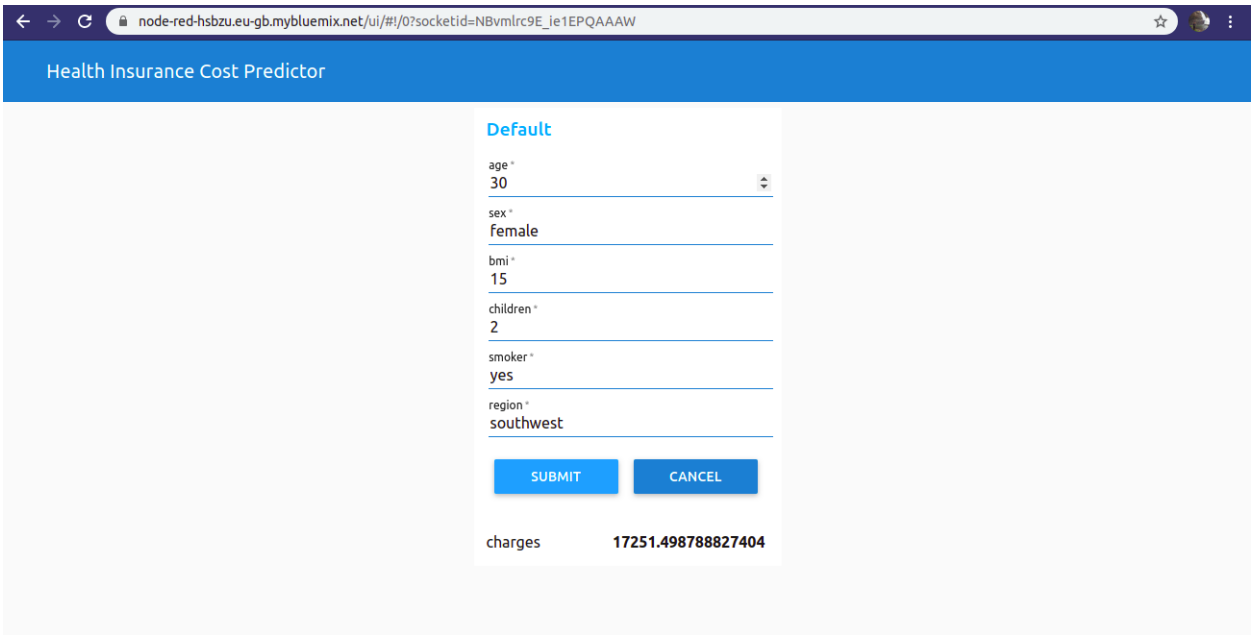
V. FLOWCHART





VI. RESULT

The machine learning model that works on the random forest algorithm predicts the cost of the health insurance package that can be delievered to a customer based on the details of their customer depending upon various factors. A Nodered UI interface which helps to host the model easily on a website and interaction between the model and user in easy and effective.



VII. ADVANTAGES AND DISADVANTAGES

ADVANTAGE :

This model helps the health insurance companies to predict the cost of insurance that can be provided to a particular customer based on factors like age, gender, weight, previous health issues etc.

DISADVANTAGE:

Sometimes the cost requirement may not depend on gender, region they belong to.

VIII. APPLICATIONS

This model that predicts the various health insurance policies applicable to a given customer can be used by the health insurance companies inorder to improvise their services and can also bring new changes to their policies based on the statistics.

IX. CONCLUSION

This was a great experience with Smartbridge learning new and interesting things and also applying them in innvoation field. Related to my project I can say that the machine learning model that is created to predict the cost of health insurance has a wide range of application and makes the work of health insurance companies more simpler. This gives all the predictions just by giving basic details of the customer. I learnt a lot from this project and also thank all the mentors and the bootcamp that was very supportive and helpful at every point of work.

X. BIBLIOGRAPHY

- Smartbridge(SmartInternz) : https://www.thesmartbridge.com/summer-internship-program-2020
- Kaggle : https://www.kaggle.com/annetxu/health-insurance-cost-prediction
- IBM Cloud services : https://www.ibm.com/in-en/cloud
- Node-Red UI : https://node-red-hsbzu.eu-gb.mybluemix.net/ui/#!/0?socketid=NBvmlrc9E_ie1EPQAAAW

XI. APPENDIX

SOURCE CODE :

https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/c6a7a438-2dab-4b5d-bef0-c197614b19d9/view?access_token=3b59fba823b5c4488f02485c074c3dbe1b27378ec34254c0ab7ea5bce8849e1a

NODE RED UI SCREENSHOT :