# Health Insurance premium prediction using IBM auto AI service

## Contents



M. Sangeetha

# Introduction:

## Overview:

Health insurance is a type of insurance that covers the whole or a part of the risk of a person incurring medical expenses. As with other types of insurance is risk among many individuals. By estimating the overall risk of health risk and health system expenses over the risk pool, an insurer can develop a routine finance structure, such as a monthly premium or payroll tax, to provide the money to pay for the health care benefits specified in the insurance agreement. The benefit is administered by a central organization, such as a government agency, private business, or not-for-profit entity.

## Purpose:

This project purpose is to build a web App that automatically estimates premium cost by taking the input values. web app is build using Ibm Watson Studio. The use of Auto AI Experiment is used to predict an insurance premium on Watson Studio. Node-RED web application is used to connect to the deployed model and predict an insurance.

# Literature Survey:

## Existing Problem:

Health insurance is a necessity nowadays, and almost every individual is linked with a government or private health insurance company. Factors determining the amount of insurance vary from company to company. Also, people in rural areas are unaware of the fact that the government of India provide free health insurance to those below poverty line. It is very complex method and some rural people either buy some private health insurance or do not invest money in health insurance at all. Apart from this people can easily be tricked about the amount of the insurance and may unnecessarily buy few expensive health insurances. Prediction is premature and does not comply with any particular company so it

must not be only criteria in selection of a health insurance. Early health insurance amount prediction can help in better contemplation of the amount needed where a person can ensure that the amount they are going to opt is justified. Also, the main problem is without the health insurance premium there will be no idea about the extra benefits with the health insurance premium.

Health Insurance Premium is the amount one pays for the health insurance every month. In addition to the premium, one usually has to pay other costs for the health care, including a deductible, co-payments, and coinsurance. This increases the cost for health insurance which makes it difficult for a middle-class human being to spend on health.

The project does not give the exact amount required for any health insurance company but gives enough idea about the amount associated with an individual for their own health insurance. The project also provides the benefits when connected with the premium.

# Proposed solution:

This project designs a web App that automatically estimates premium cost by taking the input values. Using IBM Auto AI, we automate all of the tasks involved in building predictive models for different requirements. We create a model from a data set that includes the age, gender, BMI, number of children, smoking preferences, region, and charges to predict the health insurance premium cost that an individual pay. We create an IBM Watson Studio Service, IBM Cloud Object Storage Service on IBM Cloud. We need to upload the insurance premium data file into Watson Studio. We have created an Auto AI Experiment to predict an insurance premium on Watson Studio. Auto AI uses Machine Learning Service to create several models, and the user deploys the best performing model. We use Node-RED web application to connect to the deployed model and predict an insurance.

This dataset contains the information on individual attributes such as sex, age, smoking habits etc.

It has: 1339 rows, 7 columns Description of columns:

- age – age of primary beneficiary
- sex – gender of the beneficiary. It has two categories:

1. Male
2. Female

- bmi – Body Mass Index, providing an understanding of body weights that are relatively high or low relative to height, objective index of body weight (kg/m^2) using the ratio of height to weight, ideally 18.5 to 24.9
- children – Number of children covered by the health insurance / Number of dependents.
- smoker – describing whether a person is a smoker or a non-smoker. It has 2 values:

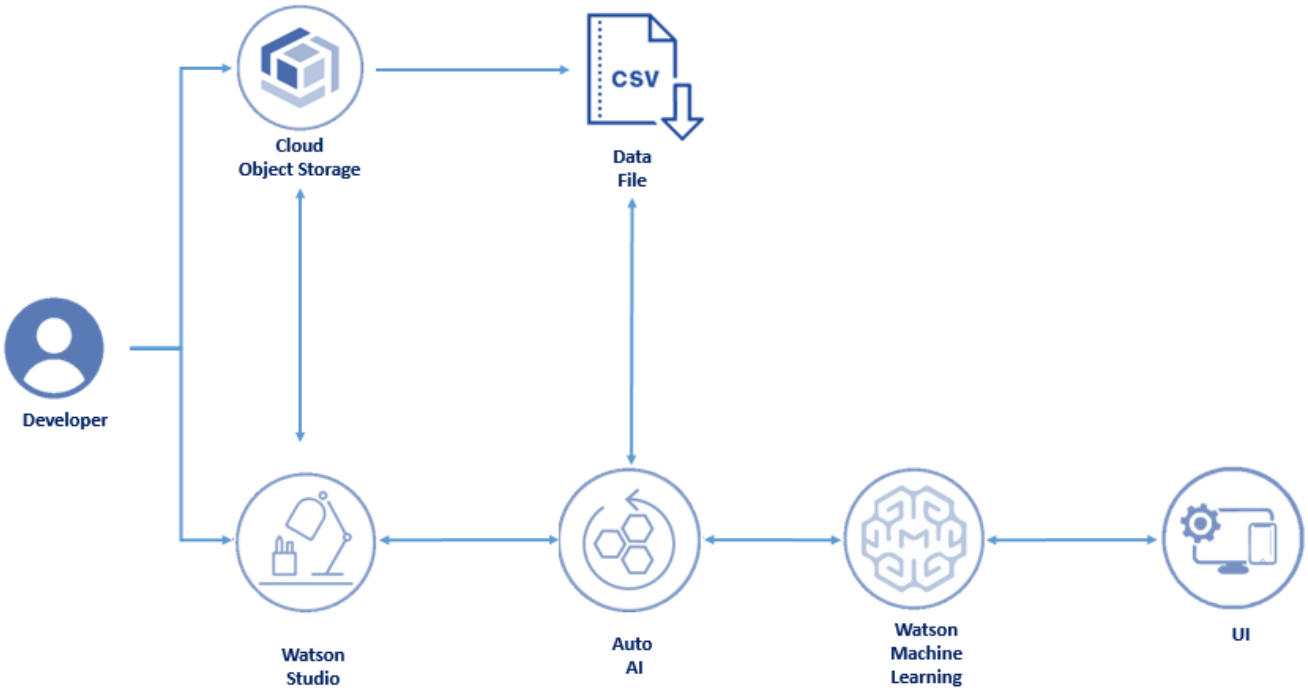
1. Yes
2. No

- region – the beneficiary's residential area in the US. It has 4 region values:

1. Northeast
2. Southeast
3. Northwest
4. Southwest

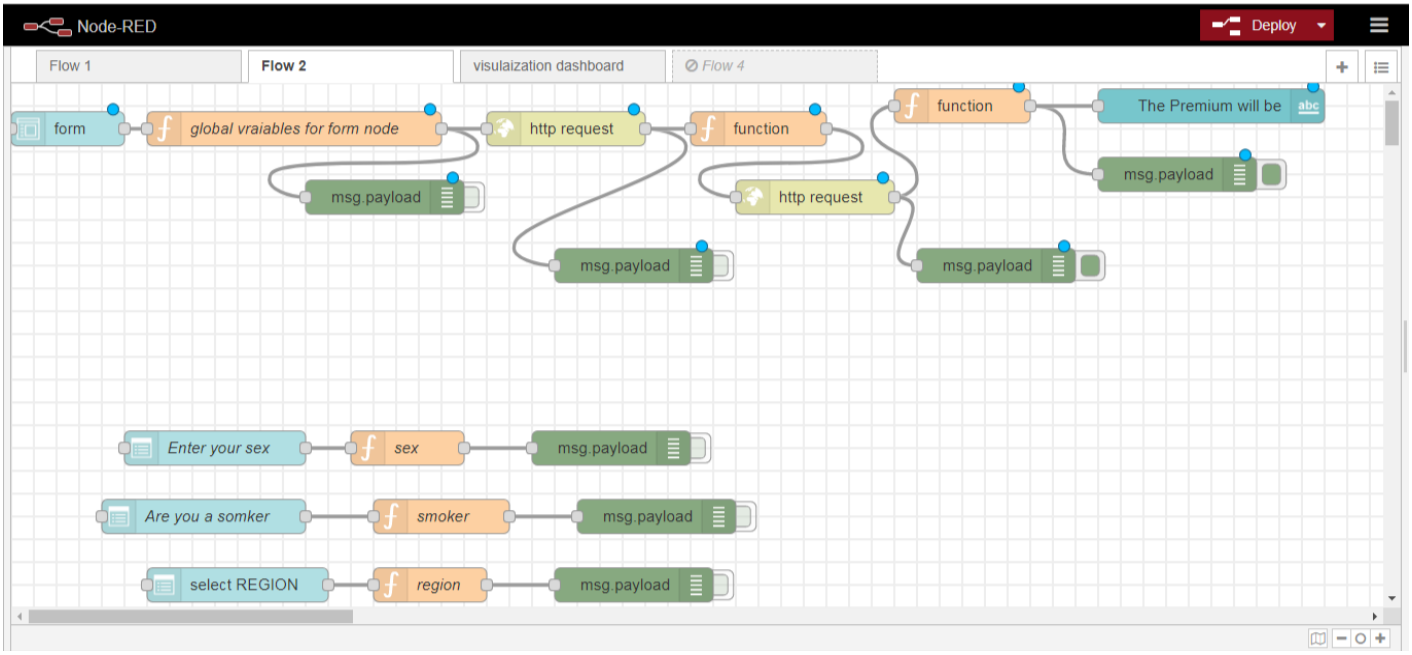
# Theoretical Analysis:

## Block Diagram:




1. The user creates an IBM Watson Studio Service on IBM Cloud.
2. The user creates an IBM Cloud Object Storage Service and adds that to Watson Studio.
3. The user uploads the insurance premium data file into Watson Studio.
4. The user creates an Auto AI Experiment to predict an insurance premium on Watson Studio.
5. Auto AI uses Watson Machine Learning to create several models, and the user deploys the best performing model.
6. The user uses the Flask web application to connect to the deployed model and predict an insurance charge.

# Software Design:




This is the software design of how the health insurance premium works. This is build using NODE-Red application.

1. Create a ui.
2. Grab the values from ui.
3. Set global variables to the ui.
4. Get the access token with the help of HTTP request.to get http request we have to get IAM access token.
5. Send the input values to scoring endpoint along with access token using HTTP request.
6. Get the predicted output.
7. Parse the output.

8. Show case the output on UI.

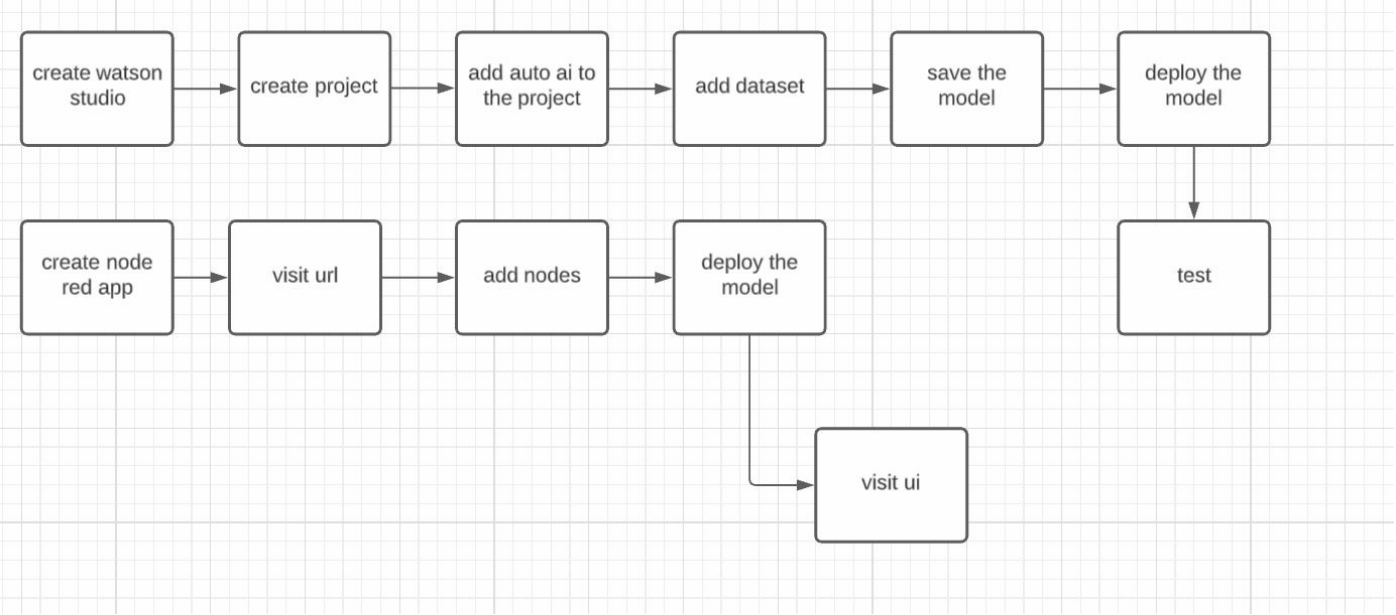
# Experimental Investigations:

| Age | Sex | Children | BMI | Region | Smoker | Premium | Status |
|---|---|---|---|---|---|---|---|
| 18 | Male | 0 | 30 | Northeast | Yes | 3165.3828125 | Pass |
| 20 | Female | 3 | 26 | Southeast | No | 4097.8359375 | Pass |
| 45 | Male | 1 | 24 | Southwest | Yes | 32505.552734375 | Pass |
| 36 | Female | 3 | 21 | Northwest | No | 6510.7734375 | Pass |
| 14 | Female | 1 | 21 | Southwest | No | 3220.94140625 | Pass |
| 28 | Male | 2 | 15 | Northeast | Yes | 28166.74609375 | Pass |
| 23 | Male | 1 | 14 | Southeast | Yes | 27251.470703125 | Pass |

The above table depicts the experimental investigations of the project.

- The table contains inputs like age, children, bmi, sex, smoker and region.
- It takes the input from the user and predict the insurance according to the input.

# Flowchart:




- Create Watson studio which provides a suite of tools for data scientists, application developers.
- The next step is to create a project in the Watson studio.
- Add auto ai IBM service to the project.
- Add dataset to the project.
- Save the model.
- Deploy the model.
- Test the model with the required outputs.
- Next create node red app.
- Add cloud ant to your app.
- Add machine learning service to the project.

- Add the nodes in the flow.
- Deploy the model.
- Visit ui.
- Give the inputs and test the deploy model.
- Given model gives us an insurance prediction.

# Result:

The project was able to predict the insurance premium when the inputs are given. It was able to predict with various inputs and was able to give the correct amount of insurance premium amount. The application of node red was successful. Output was predicted for the given inputs. By this, an individual can be able to get an idea of their insurance premium by giving their health-related details. The project also helps an individual to know their expenses for their health and can get benefits accordingly.

# Advantages and Disadvantages:

## Advantages:

The main purpose of medical insurance is to receive the best medical care without any strain on your finances. Health insurance plans offer protection against high medical costs. It covers hospitalization expenses, day care procedures, domiciliary expenses, and ambulance charges, besides many others.

## Disadvantages:

Health insurance can be very costly even for those that have a health insurance plan through their employers. This cost can sometimes be so expensive that the person may struggle to make payments. This is problematic for those that have a low income or are self-employed.

# Applications:

There are many applications of health insurance premium prediction. One of the applications is wealth transfer which is known as the vehicle of sizable for health insurance. Beneficiaries receive access to free or low out-of-pocket cost services, such as routine doctor visits. These services are largely predictable—such as well-child visits for people with children or medication refills for people on lipid-lowering medications. The other application is Enhancing and ensuring the quality of clinicians and hospitals. Both commercial and government insurers have developed measurement efforts that aim to monitor and improve the quality of hospitals. Medicare limits which hospitals can perform the transcatheter aortic valve replacement procedure to those with adequate volume and expertise. Another one is Financial protection to individuals with catastrophic health events. Health insurance, like car insurance, protects individuals from unpredictable and financially catastrophic events. Organ failure requiring a transplant can lead to hundreds of thousands of dollars in costs. Policies aimed at improving this function of health insurance include capping annual out-of-pocket expenses, ending lifetime benefit limits, and ensuring coverage for people with pre-existing conditions.

# Conclusion:

Health insurance companies can then accurately charge the premium based upon a specific individual’s attributes. This will not only help the individuals in getting charged the right amount of premium for their health insurance but will also help in forging better relationships and a level of trust between the insurance company and the insured. Based on these predictions, the health insurance providers can then

evaluate the following decisions and make better judgement calls: Which individuals deserve which kind of insurance plan?  How much the premium should be charged based on an individual's behaviors? Based upon an individual's behavior, predicting their premium helps in better risk management.  It helps forge trust between the customer and the insurance company. Thus, it is important for a health insurance company to collect and analyze the data such as a person's age, BMI, health data to accurately predict the risk and charge accurate premiums to cover that risk. However, there are certain limitations which is the scope of further studies. The data did not include any information on an individual's medical costs, the real-time data i.e. data collected from the sensors in the wearable health devices such as fit bits etcetera. If we take all these types of different data sources into account then we can have a better picture of an individual's behavior and can more accurately predict the insurance premium charge and the associated risk.

# Future scope:

Premium amount prediction focuses on persons own health rather than other company's insurance terms and conditions. The models can be applied to the data collected in coming years to predict the premium. This can help not only people but also insurance companies to work in tandem for better and more health centric insurance amount. With the scope of coverage being broadened, that is, increase in the number of ailments being covered under an insurance policy, new health covers will be available at revised rates in the future. In the future by paying little extra premium, health insurance products will become more inclusive, less complicated and more attractive for customers. Further, modern treatment methods such as oral chemotherapy, balloon sinuplasty, deep brain stimulation, among others, can also be included.

# Bibliography:

[1] https://en.wikipedia.org/wiki/Health_insurance

[2] Health Insurance Flow Charts & Workflow Templates, Examples | OpsDog

[3] https://www.google.com/search?q=advantages+and+disadvabatges+of+health+insurance+premium&oq=advantage&aqs=chrome.1.69i57j69i59l3j0i433j0j69i60l2.4045j0j7&sourceid=chrome&ie=UTF-8

# Appendix:

## Source Code:

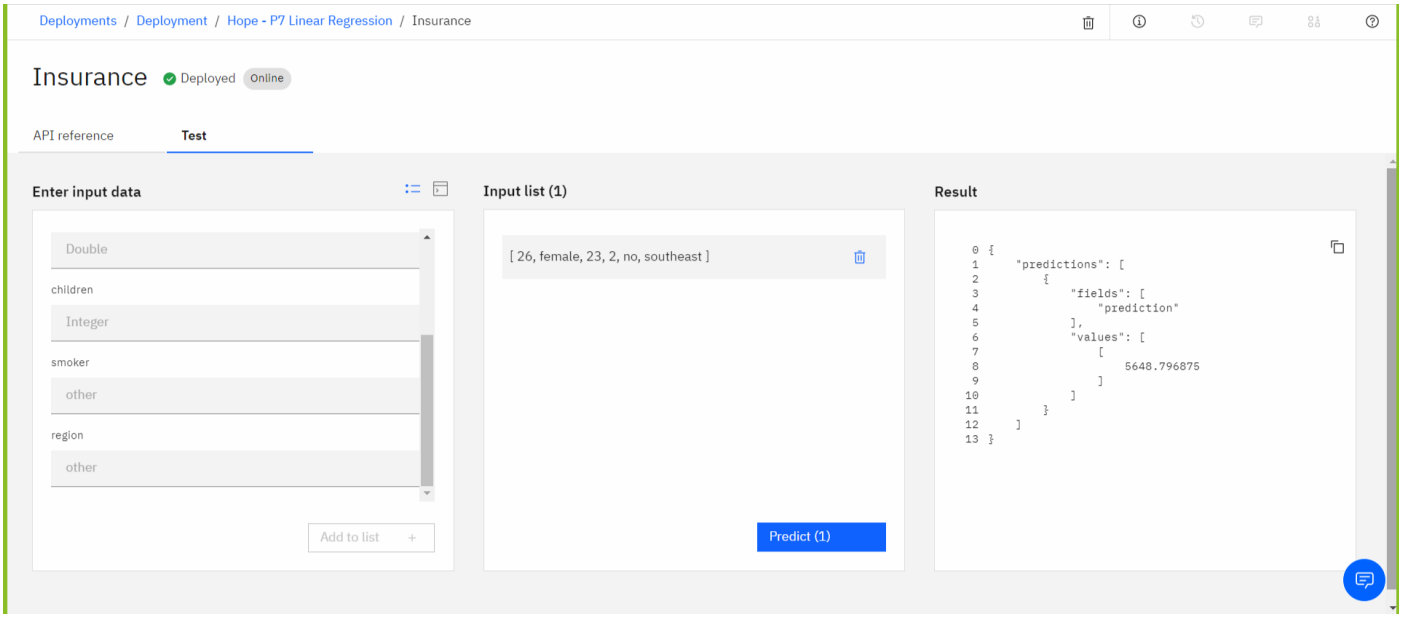
[{"id":"2c3f2276.8e8b4e","type":"tab","label":"Flow 2","disabled":false,"info":""},{"id":"eb72d9c8.5eb848","type":"ui_form","z":"2c3f2276.8e8b4e","name":"","label":"","group":"1eb501cd.838cee","order":4,"width":0,"height":0,"options":[{"label":"Age","value":"a","type":"number","required":true,"rows":null},{"label":"BMI","value":"b","type":"number","required":true,"rows":null},{"label":"Children","value":"c","type":"number","required":true,"rows":null}],"formValue":{"a":"","b":"","c":""},"payload":"","submit":"submit","cancel":"cancel","topic":"topic","topicType":"msg","splitLayout":"","x":70,"y":60,"wires":[["7494bfb3.09a5a"]]},{"id":"a257d518.78df38","type":"ui_dropdown","z":"2c3f2276.8e8b4e","name":"","label":"Select your Sex","tooltip":"","place":"Select option","group":"1eb501cd.838cee","order":1,"width":0,"height":0,"passthru":true,"multiple":false,"options":[{"label":"Female","value":"female","type":"str"},{"label":"Male","value":"male","type":"str"}],"payload":"","topic":"topic","topicType":"msg","x":100,"y":220,"wires":[["f7cb76b5.a6de28"]]},{"id":"32dfbe6c.9186b2","type":"ui_dropdown","z":"2c3f2276.8e8b4e","name":"","label":"Are you a Smoker","tooltip":"","place":"Select option","group":"1eb501cd.838cee","order":2,"width":0,"height":0,"passthru":true,"multiple":false,"options":[{"label":"YES","value":"yes","type":"str"},{"label":"NO","value":"no","type":"str"}],"payload":"","topic":"topic","topicType":"msg","x":110,"y":280,"wires":[["e8b44ad7.eed038"]]},{"id":"907ed9b9.93f1f8","type":"ui_dropdown","z":"2c3f2276.8e8b4e","name":"","label":"Select Region","tooltip":"","place":"Select option","group":"1eb501cd.838cee","order":3,"width":0,"height":0,"passthru":true,"multiple":false,"options":[{"label":"southeast","value":"southeast","type":"str"},{

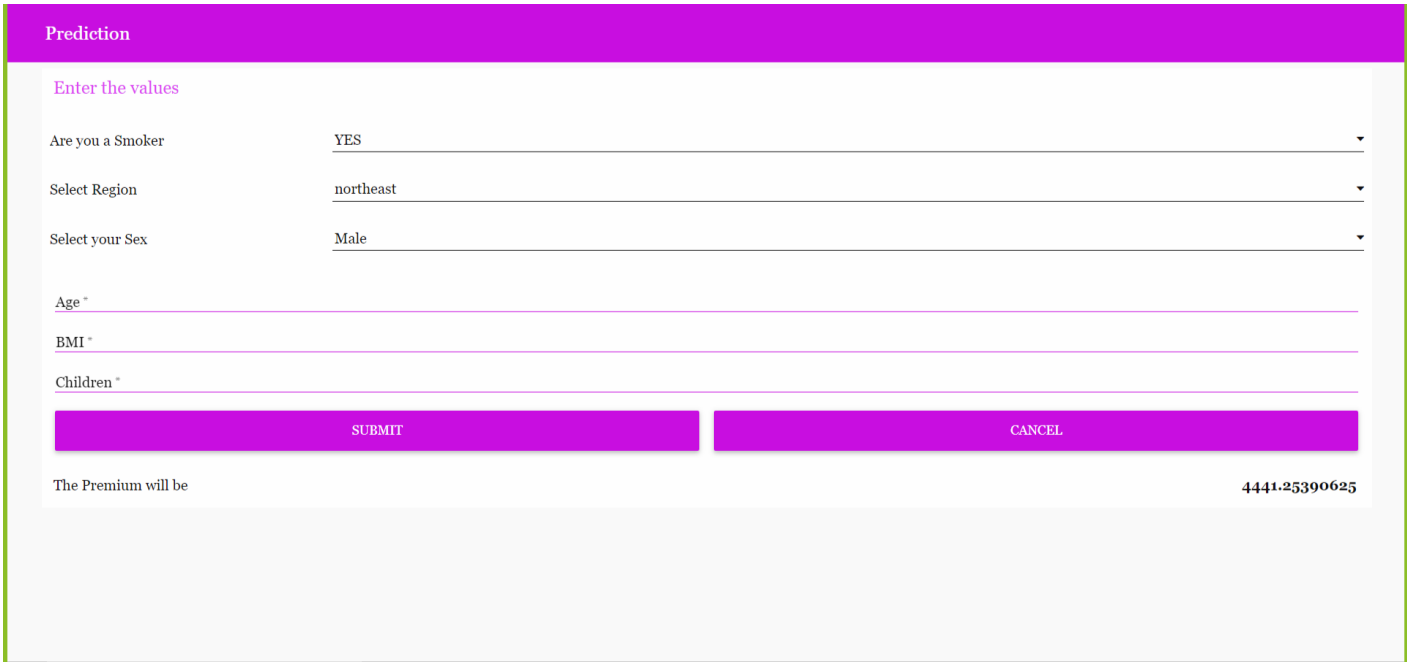
"label":"northeast","value":"northeast","type":"str"},{"label":"southwest","value":"
southwest","type":"str"},{"label":"northwest","value":"northwest","type":"str"}],"p
ayload":"","topic":"topic","topicType":"msg","x":100,"y":360,"wires":[["56447c97
.7415f4"]]},{"id":"1aff1a12.baa166","type":"debug","z":"2c3f2276.8e8b4e","name
":"","active":false,"tosidebar":true,"console":false,"tostatus":false,"complete":"fals
e","statusVal":"","statusType":"auto","x":310,"y":120,"wires":[]},{"id":"ea189e2.a
e1dd6","type":"debug","z":"2c3f2276.8e8b4e","name":"","active":false,"tosidebar"
:true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType"
:"auto","x":450,"y":220,"wires":[]},{"id":"c809032.af934","type":"debug","z":"2c
3f2276.8e8b4e","name":"","active":false,"tosidebar":true,"console":false,"tostatus"
:false,"complete":"false","statusVal":"","statusType":"auto","x":570,"y":300,"wires
":[]},{"id":"38fe2c76.856b34","type":"debug","z":"2c3f2276.8e8b4e","name":"","
active":false,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","st
atusVal":"","statusType":"auto","x":470,"y":360,"wires":[]},{"id":"7494bfb3.09a5a
","type":"function","z":"2c3f2276.8e8b4e","name":"","func":"global.set(\"age\",ms
g.payload.a)\nglobal.set(\"children\",msg.payload.a)\nglobal.set(\"BMI\",msg.payl
oad.a)\nvar apikey
=\"Wa-UHQ877F1z79sIAdSvyfHXbLwIo88gprZ8QWe1BmZu\"\nmsg.headers={\
"content-type\":\"application/x-www-form-urlencoded\"}\nmsg.payload={\"grant_t
ype\":\"urn:ibm:params:oauth:grant-type:apikey\",\"apikey\":apikey}\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":220,"y":60,"wi
res":[["1aff1a12.baa166","b09034eb.6d0ba8"]]},{"id":"f7cb76b5.a6de28","type":"
function","z":"2c3f2276.8e8b4e","name":"sex","func":"global.set(\"sex\",msg.payl
oad)\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":290,"y":220,"w
ires":[["ea189e2.ae1dd6"]]},{"id":"e8b44ad7.eed038","type":"function","z":"2c3f2
276.8e8b4e","name":"smoker","func":"global.set(\"smoker\",msg.payload)\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":340,"y":280,"w
ires":[["c809032.af934"]]},{"id":"56447c97.7415f4","type":"function","z":"2c3f22
76.8e8b4e","name":"region","func":"global.set(\"region\",msg.payload)\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":270,"y":360,"w
ires":[["38fe2c76.856b34"]]},{"id":"b09034eb.6d0ba8","type":"http
request","z":"2c3f2276.8e8b4e","name":"","method":"POST","ret":"obj","paytoqs"
:"ignore","url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"p

roxy":"","authType":"","x":410,"y":60,"wires":[["ac9ecc55.af17d","139b7a67.09f7b6"]]},{"id":"ac9ecc55.af17d","type":"debug","z":"2c3f2276.8e8b4e","name":"","active":false,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType":"auto","x":550,"y":120,"wires":[]},{"id":"139b7a67.09f7b6","type":"function","z":"2c3f2276.8e8b4e","name":"","func":"var age = global.get(\"age\")\nvar sex = global.get(\"sex\")\nvar bmi = global.get(\"bmi\")\nvar children = global.get(\"children\")\nvar smoker = global.get(\"smoker\")\nvar region = global.get(\"region\")\nvar token=msg.payload.access_token\nmsg.headers={'Content-Type':'application/json',\"Authorization\":\"Bearer\"+token,\"Accept\":\"application/json\"}\nmsg.payload={\"input_data\": [{\"fields\":[\"age\",\"sex\",\"bmi\",\"children\",\"smoker\",\"region\"],\"values\":[[age,sex,bmi,children,smoker,region]]}]}\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":600,"y":60,"wires":[["40f5c0de.43b7c"]]},{"id":"40f5c0de.43b7c","type":"http request","z":"2c3f2276.8e8b4e","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://us-south.ml.cloud.ibm.com/ml/v4/deployments/b3c948e9-f8d4-43b1-9dc8-d1b87e5b84d4/predictions?version=2021-06-29","tls":"","persist":false,"proxy":"","authType":"","x":780,"y":60,"wires":[["9fd44984.b88f68","b2e28f6c.c3488"]]},{"id":"9fd44984.b88f68","type":"debug","z":"2c3f2276.8e8b4e","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType":"auto","x":790,"y":120,"wires":[]},{"id":"b2e28f6c.c3488","type":"function","z":"2c3f2276.8e8b4e","name":"","func":"msg.payload = msg.payload.predictions[0].values[0][0]\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","libs":[],"x":970,"y":60,"wires":[["a3ada5e9.d127a8","f998c43c.2d4458"]]},{"id":"a3ada5e9.d127a8","type":"debug","z":"2c3f2276.8e8b4e","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","statusVal":"","statusType":"auto","x":1030,"y":120,"wires":[]},{"id":"f998c43c.2d4458","type":"ui_text","z":"2c3f2276.8e8b4e","group":"1eb501cd.838cee","order":5,"width":0,"height":0,"name":"","label":"The Premium will be","format":"{{msg.payload}}","layout":"row-spread","x":1180,"y":60,"wires":[]},{"id":"1eb501cd.838cee","type":"ui_group","name":"Enter the values","tab":"a8bd3be4.eeaa08","order":1,"disp":true,"width":"27","collapse":fals

e},{"id":"a8bd3be4.eeaa08","type":"ui_tab","name":"Prediction","icon":"dashboard","disabled":false,"hidden":false}]

# Output:




M. Sangeetha

The premium for the given inputs is 4441.25

**Project Documentation by**

**M. Sangeetha**