Project Report

Health Insurance Cost prediction

Using Watson Auto AI

By: KRISHNA PRIYA C

RSIP Career Basic ML 079

INDEX

PROJECT DETAILS

Project ID : SPS_PRO_286

Project Title : Health Insurance Cost Prediction Using Watson Auto AI

Duration : 4 Weeks

Project Support : SmartBridge Educational Services

Project Mentor : Mr. Rammohan Bethi

Kickoff Date : June 1st, 2020

Finish Date : June 30th, 2020

ACKNOWLEDGEMENT

This project has taken a considerable amount of time and resources. I would like to acknowledge the help of all of those who have made this project possible. In finical I would like to thank my supervisor Mr. Rammohan Bethi for his time, patience and guidance, and also for allowing the idea to be pursued primitively. I would also like to thank Mr. Vinay Kumar Nomula for his help. Further to these people I would like to thank the members of the Smartbridge career workshop for their technical help in setting up various codes and faults. Also, I would like to thank all of my co-interns who have worked on the Open Source projects without whose efforts this project would not have been possible.

1.INTRODUCTION

The project “Health Insurance Cost prediction” predicts the cost of health insurance based on some details of person which helps to improve performance of Insurance companies and to provide better services to the customer. An important informatics tool for controlling healthcare costs is accurately predicting the likely future healthcare costs of individuals. To address this important need, we conducted a systematic literature review and identified five methods for predicting healthcare costs. To enable a direct comparison of these different approaches, we empirically evaluated the predictive performance of each reported approach, as well as other state-of-the-art supervised learning methods, using data from University of Utah Health Plans for October 2013 to October 2016. The data set consisted of approximately 90,000 individuals, 6.3 million medical claims and 1.2 million pharmacy claims. In this comparative analysis, gradient boosting had the best predictive performance overall and

for low to medium cost individuals. For high cost individuals, Artificial Neural Network (ANN) and the Ridge regression model, which have not been previously reported for use in healthcare cost prediction, had the highest performance.

2.LITERATURE SURVEY

Existing problem

Health Insurance companies have a tough task at determining premiums for their customers.While the health care law in any country does have some rules for companies to follow to determine premiums, it's really up to the companies on what factor/s they want to hold more weightage. Companies should know the most important factors and how much statistical importance do they hold.

Proposed Solution

The main aim of this project is to create a model based on statistically significant factors (independent variable) which will affect premiums charges (dependent

variable) by an insurance company. In this project we are using Multi Linear regression for the accurate prediction. An application is also build in Auto AI Service in IBM Cloud which can be interlinked with the model so as to view the result on UI based on input parameters.

3.THEORETICAL ANALYSIS

a.Block diagram

b. Hardware / Software designing

The project has been done by using IBM Cloud in which machine learning service, Watson studio and cloud storage service (to store the data) have been created by using the options available in Catlog.

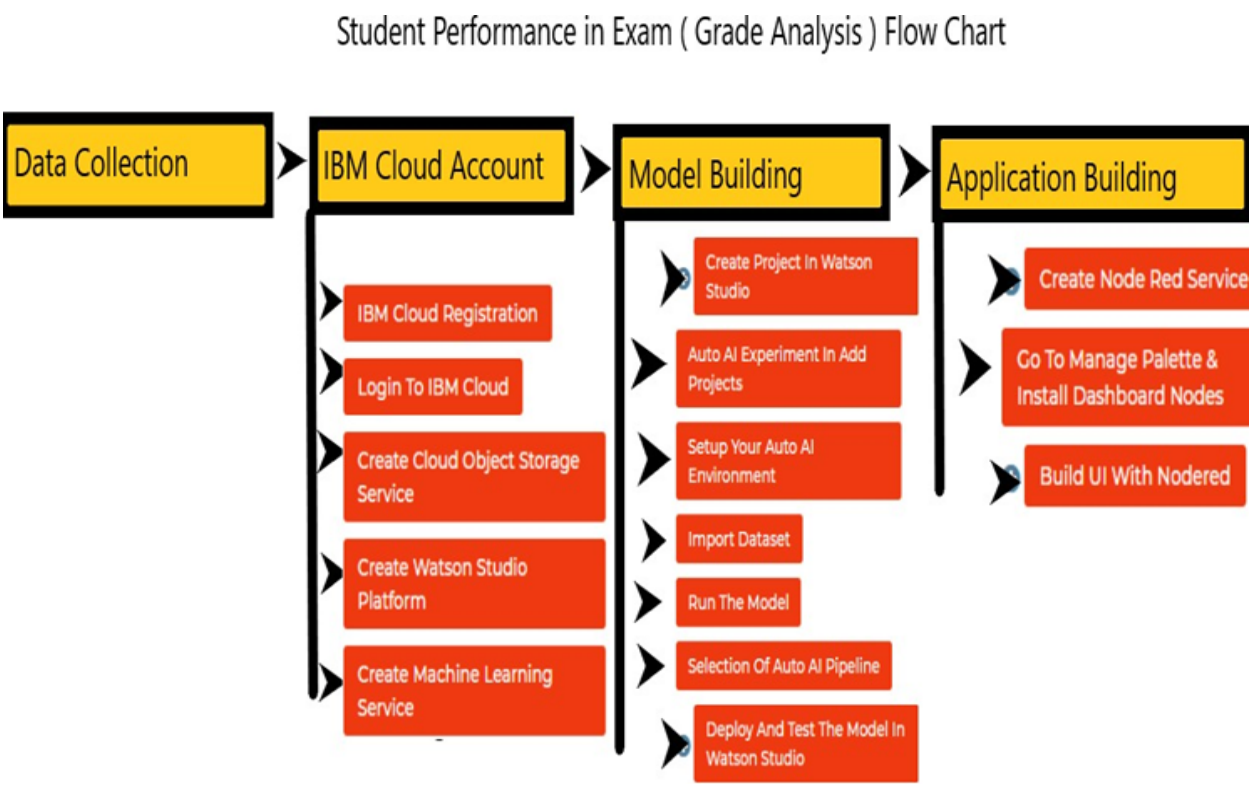
4.EXPERIMENTAL INVESTIGATION

There are six steps in experimental investigation on of a general project:

1.Choose a Project Idea
2. Conduct Background Research
3.Compose a Hypothesis
4.Design your Experiment
5.Collect Data
6.Analyse Data and Draw Conclusions

All the data has been collected considering the above factors and it has been formatted. After formatting it has been uploaded in the project and after that using Watson Studio Auto AI Experiment, it is uploaded to cloud object storage service and implemented. Based on these implementation, the value can be predicted using the data we have collected. After that application is developed using Node red Service.

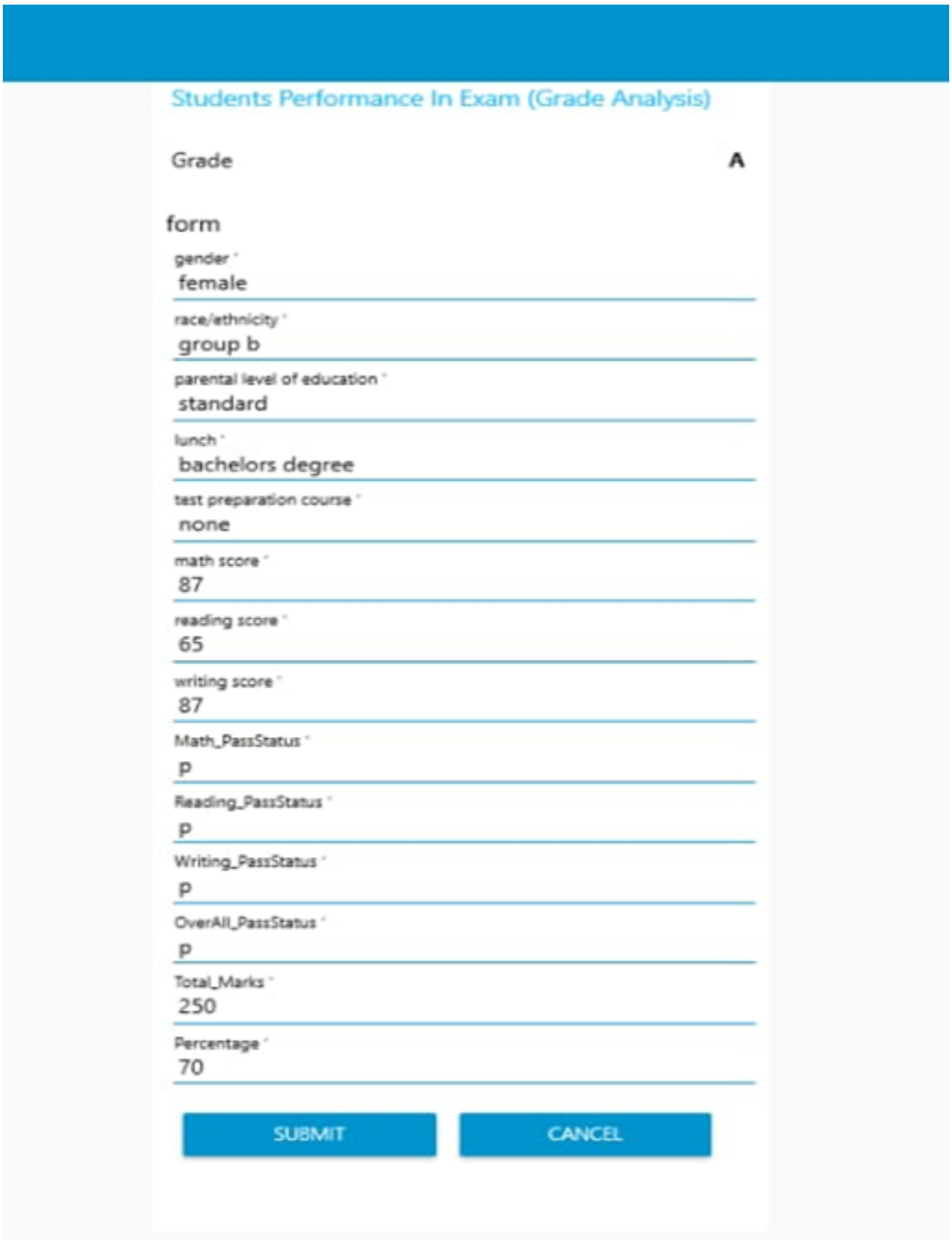
5. FLOWCHART



6.RESULT

After the implementation, deployment of project the result i.e. predicted health insurance cost of can be seen in Node Red UI. This value depends on different parameters. The Node Red UI provide us simple way to get the result of Auto AI Experiment. Here is the Node Red UI predicts the cost of health insurance.

PREDICTED VALUE:



7.ADVANTAGES & DISADVANTAGES

The advantages are easy to implement, accessibility is fast, continuous Improvement, wide application, available 24x7, no human intervention needed. We can handle multi-dimensional and multi-variety data. Where as the disadvantages are lack of security, loss of control on data, high error susceptibility, dependence of network/providers.

8.APPLICATION

Using The Auto AI Experiment, one can build and deploy a machine learning model with sophisticated training features. In the given project we can predict the price of the required vehicle by giving few input parameters.

9.CONCLUSION

In this project by using IBM Cloud the model

processing is been done in Auto AI services in IBM cloud and then the deployment is been done in Watson studio and application is build using Node red service which has been successful as we are able to get the desired output.

10.FUTURE SCOPE

As we are developing day to day there is a continuous growth of Auto AI and Machine Learning. The web application can be used to predict the cost of the health insurance accurately, precisely and efficiently instead of n number of people being involved directly or indirectly.

11.BIBLIOGRAPHY

APPENDIX

SourceCode:


```
[{"id":"dea64d08.d7914","type":"tab","label":"Fl
```

ow
1","disabled":false,"info":""},{"id":"817f5d0e.862a6","type":"function","z":"dea64d08.d7914","name":"PreToken","func":"global.set(\"ag\",msg.payload.ag)\nglobal.set(\"sx\",msg.payload.sx)\nglobal.set(\"bm\",msg.payload.bm)\nglobal.set(\"ch\",msg.payload.ch)\nglobal.set(\"sm\",msg.payload.sm)\nglobal.set(\"rg\",msg.payload.rg)\nvar apikey=\"MRL-zkJUHntiAv2Szn6A1vB4dyh8UnP5pJoR0A0zhwY3\";\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,"x":320,"y":300,"wires":[["47c3a2f2.13f73c"]]},{"id":"47c3a2f2.13f73c","type":"http request","z":"dea64d08.d7914","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":450,"y":400,"wires":[["7f43f2bb.b054fc"]]},{"id":"65178d47.f10c54","type":"inject","z":"dea64d08.d7914","name":"","topic":"","payload":"","payloadType":"date","repeat":"","crontab":"","once":false,"onceDelay":0.1,"x":120,"y":200,"wires":[["817f5d0e.862a6"]]},{"id":"7f43f2bb.b054fc","type":"function","z":"dea64d08.d7914","name":"Pre

```
Prediction","func":"var ag =
global.get('ag')\nvar sx = global.get('sx')\nvar
bm = global.get('bm')\nvar ch =
global.get('ch')\nvar sm = global.get('sm')\nvar
rg = global.get('rg')\nvar
token=msg.payload.access_token\nvar
instance_id=\"ae73a5e4-b6ae-485e-8276-0e866e4cc1
65\"\nmsg.headers={'Content-Type':
'application/json',\"Authorization\":\"Bearer
\"+token,\"ML-Instance-ID\":instance_id}\nmsg.pa
yload={\"input_data\":[{\"fields\": [\"age\",
\"sex\", \"bmi\", \"children\", \"smoker\",
\"region\"], \"values\":
[[ag,sx,bm,ch,sm,rg]]}]}\nreturn
msg;","outputs":1,"noerr":0,"x":700,"y":300,"wir
es":[["a3cce373.8d569"]]},{"id":"a3cce373.8d569"
,"type":"http
request","z":"dea64d08.d7914","name":"","method"
:"POST","ret":"obj","paytoqs":false,"url":"https
://us-south.ml.cloud.ibm.com/v4/deployments/777d
5056-a28d-4270-ade3-224d1af4ff21/predictions","tl
s":"","persist":false,"proxy":"","authType":"","
x":870,"y":160,"wires":[["303397da.499d58"]]},{"
id":"303397da.499d58","type":"function","z":"dea
64d08.d7914","name":"","func":"msg.payload=msg.p
ayload.predictions[0].values[0][0]\nreturn
msg;","outputs":1,"noerr":0,"x":970,"y":220,"wir
```

es":[["9b97e011.3622b","daedad8b.1edea"]]},{"id"
:"9b97e011.3622b","type":"ui_text","z":"dea64d08
.d7914","group":"c3194919.21e1b8","order":1,"wid
th":0,"height":0,"name":"","label":"charges","fo
rmat":"{{msg.payload}}","layout":"row-spread","x
":1360,"y":300,"wires":[]},{"id":"8406a45f.c1061
8","type":"ui_form","z":"dea64d08.d7914","name":
"","label":"","group":"c3194919.21e1b8","order":
0,"width":0,"height":0,"options":[{"label":"age"
,"value":"ag","type":"number","required":true,"r
ows":null},{"label":"sex","value":"sx","type":"t
ext","required":true,"rows":null},{"label":"bmi"
,"value":"bm","type":"number","required":true,"r
ows":null},{"label":"children","value":"ch","typ
e":"number","required":true,"rows":null},{"label
":"smoker","value":"sm","type":"text","required"
:true,"rows":null},{"label":"region","value":"rg
","type":"text","required":true,"rows":null}],"f
ormValue":{"ag":"","sx":"","bm":"","ch":"","sm":
"","rg":""},"payload":"","submit":"submit","canc
el":"cancel","topic":"","x":100,"y":460,"wires":
[["817f5d0e.862a6"]]},{"id":"daedad8b.1edea","ty
pe":"debug","z":"dea64d08.d7914","name":"","acti
ve":true,"tosidebar":true,"console":false,"tosta
tus":false,"complete":"payload","targetType":"ms
g","x":1170,"y":60,"wires":[]},{"id":"c3194919.2
1e1b8","type":"ui_group","z":"","name":"Health

insurance cost
prediction","tab":"91cd5e00.bc8a3","order":1,"di
sp":true,"width":"8","collapse":false},{"id":"91
cd5e00.bc8a3","type":"ui_tab","z":"","name":"Hom
e","icon":"dashboard","order":3,"disabled":false
,"hidden":false}]

References

1.The Centers for Medicare & Medicaid Services (CMS)DoHaHS,

United States. National Health Expenditure Data 2016.Available

from:

https://www.cms.gov/Research-Statistics-Data-and-Systems/Statis
tics-Trends-and-Reports/NationalHealthExpendData/i

ndex.html.

2.Duncan I, Loginov M, Ludkovski M. Testing AlternativeRegression Frameworks for Predictive Modeling of Health Care Costs. North American Actuarial Journal.2016;20(1):65–87.

3. Burwell SM. Setting value-based payment goals--HHS effortsto improve US health care.2015