RSIP Career Basic ML 060

student performance Prediction Using Watson Auto AI

Team Members :
vishal gupta

June2020

CONTENTS 1

INTRODUCTION

overview:

This thesis examines the application of machine learning algorithms to predict whether a student will be successful or not. The specific focus of the thesis is the comparison of machine learning methods and feature engineering techniques in terms of how much they improve the prediction performance. Three different machine learning methods were used in this thesis. They are random forest, decision trees, and xgboost classification. Feature engineering, the process of modification and selection of the features of a data set, was used to improve predictions made by these learning algorithms. Two different data sets containing records of student information were used. The machine learning methods were applied to both the raw version and the feature engineered version of the data sets, to predict the student's success. The thesis comes to the same conclusion as the earlier studies: The results show that it is possible to predict student performance successfully by using machine learning. The best algorithm was xgboost classification for the first data set, with 99 percent accuracy, and random forest for the second data set, with 98 percent accuracy. Feature engineering was found to be more important factor in prediction performance than method selection in the data used in this study. Keywords and terms: student performance, machine learning xgboost , random forest classification, decision trees

<u>**PURPOSE:**</u>

1 how to improve the student performance in each test

2 what are the major factors influencing the test score

3 effectiveness of test preparartion course

4 other inferences

2. LITERATURE SURVEY

2.1 Existing Problem

This thesis examines the application of machine learning algorithms to predict whether a student will be successful or not. Machine learning techniques can be utilized for students' grades prediction in different courses. Such techniques would help students to improve their performance based on predicted grades and would enable instructors to identify such individuals who might need assistance in the courses. This is best suited for online courses.

2.2 Proposed solution

We are building a Machine Learning model to predict the material 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. This model is to predict the performance of students in exam in grade distribution based by using different input parameters.

3. THEORITICAL ANALYSIS

IBM cloud compung is a set of cloud compung services for business offered by the informaon technology company IBM. IBM Cloud includes infrastructure as a service (IaaS), soware as a service (SaaS) and plaorm as a service (PaaS) offered through public, private and hybrid cloud delivery models, in addion to the components that make up those clouds.

3.1 Block diagram

IBM offers three hardware plaorms for cloud compung. These plaorms offer built-in support for virtualizaon. For virtualizaon IBM offers IBM Websphere applicaon infrastructure that supports programming models and open standards for virtualizaon. The management layer of the IBM cloud framework includes IBM Tivoli middleware. Management tools provide capabilies to regulate images with automated provisioning and de-provisioning, monitor operaons and meter usage while tracking costs and allocang billing. The last layer of the framework provides integrated workload tools. Workloads for cloud compung are services or instances of code that can be executed to meet specific business needs. IBM offers tools for cloud based collaboraon, development and test, applicaon development, analycs, business-to-business integraon, and security. IBM Watson Studio helps data sciensts and analysts prepare data and build models at scale across any cloud. With its open, flexible mulcloud architecture, Watson Studio provides capabilies that empower businesses to simplify enterprise data science and AI, such as:

● *Automate AI lifecycle management with AutoAI*

● *Visually prepare and build models with IBM SPSS Modeler*

 ● *Build models using images with IBM Watson Visual Recognition and texts with IBM Watson Natural Language Classifier*

● *Deploy and run models through one-click integration with IBM Watson*

Machine Learning

● *Manage and monitor models through integration with IBM Watson OpenScale*

3.2 Hardware / Software designing

While building models with Watson AutoAI on IBM Cloud plaorm it automacally selects the best suited Machine Learning Algorithm for the provided dataset and the selecon criteria. Watson AutoAI does so by applying pipelings to the datasets, it puts the provided datasets in the mulple Machine Learning Algorithms and selects the best suited algorithm for the provided datasets based on accuracy values of the algorithms.Above working is shown by the following screen shots of AutoAI experiment.
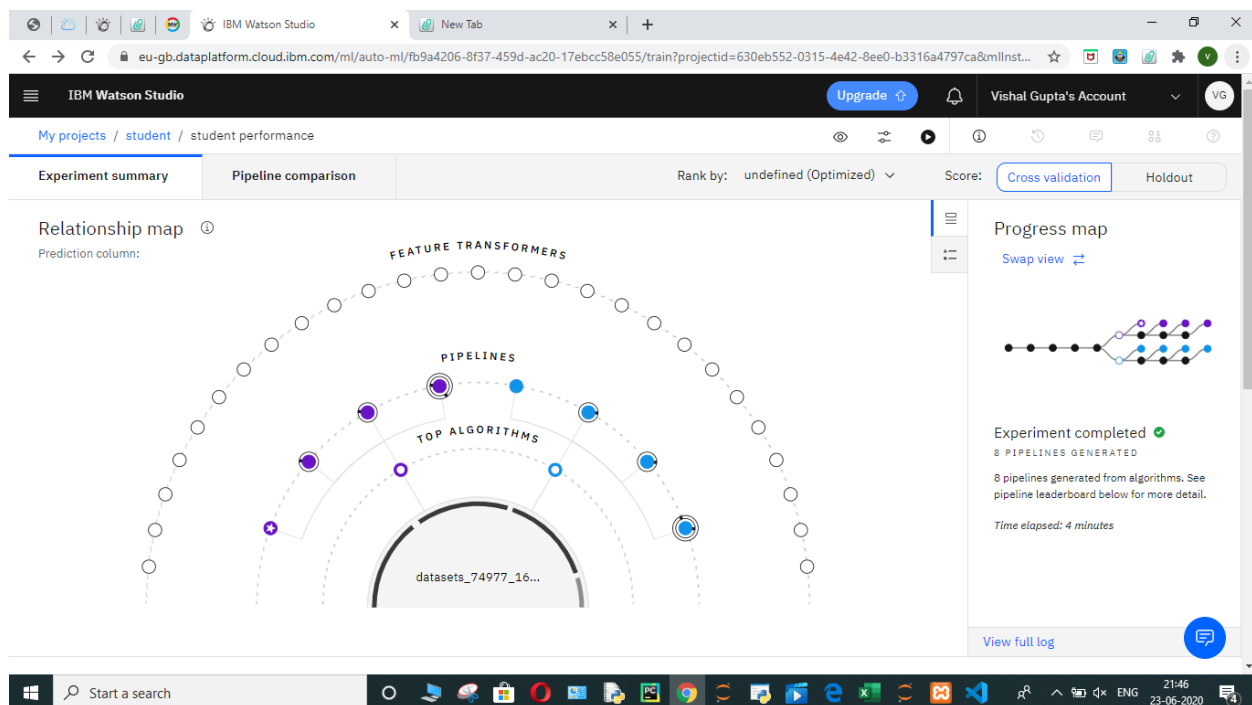

Fig:Result of Running Pipelines

Watson AutoAI provide Arficial Intelligence plaorm for implemenng different algorithms.Aer successful implementaon of algorithm , IBM Cloud also provide different services for hosng your algorithm through web apps. One such plaorm is provided through the Node Red App. It is built upon Nodejs for creang web apps.

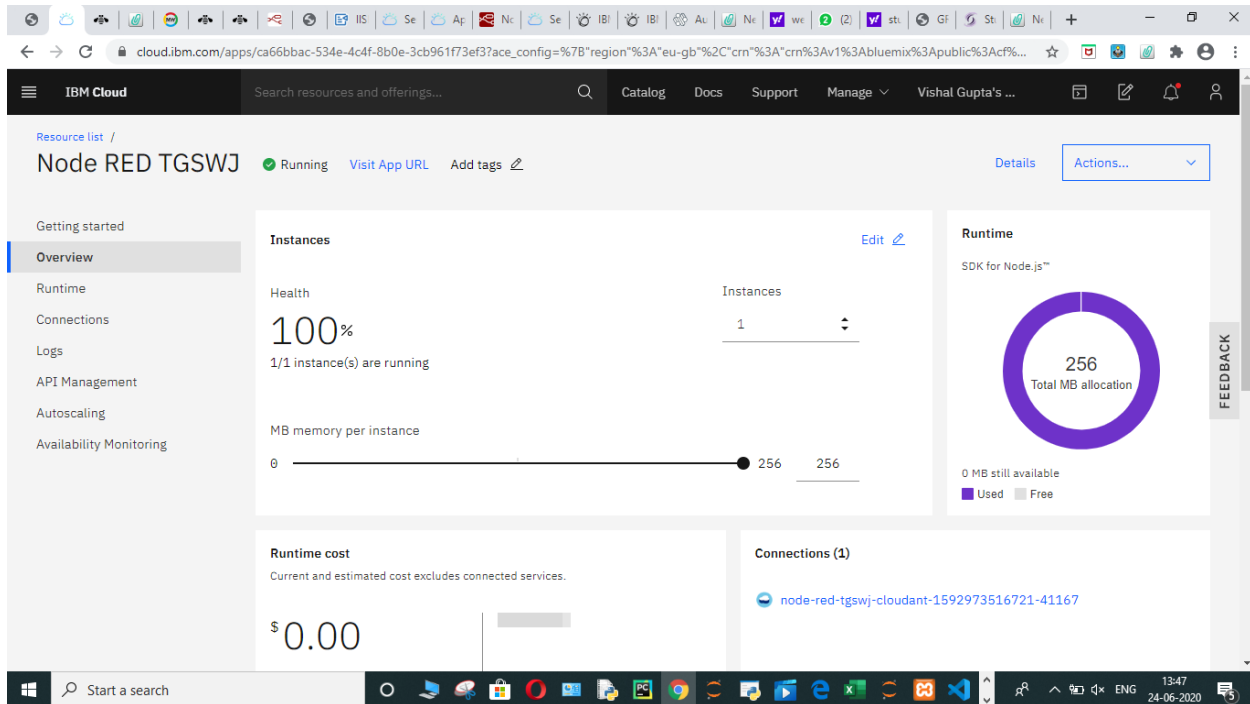**Fig:Running NodeRed Service**

4. EXPERIMENTAL INVESTIGATION

While working on the implementaon of your Machine Learning or Arficial Intelligence Project , IBM Cloud also provides very essenal feature to test your implementaon before you can deploy it , wherever you wanted to be. In context with the student performanace Predicon , aer creang and saving of the best fit algorithm based on accuracy, we test our Watson AutoAI model :
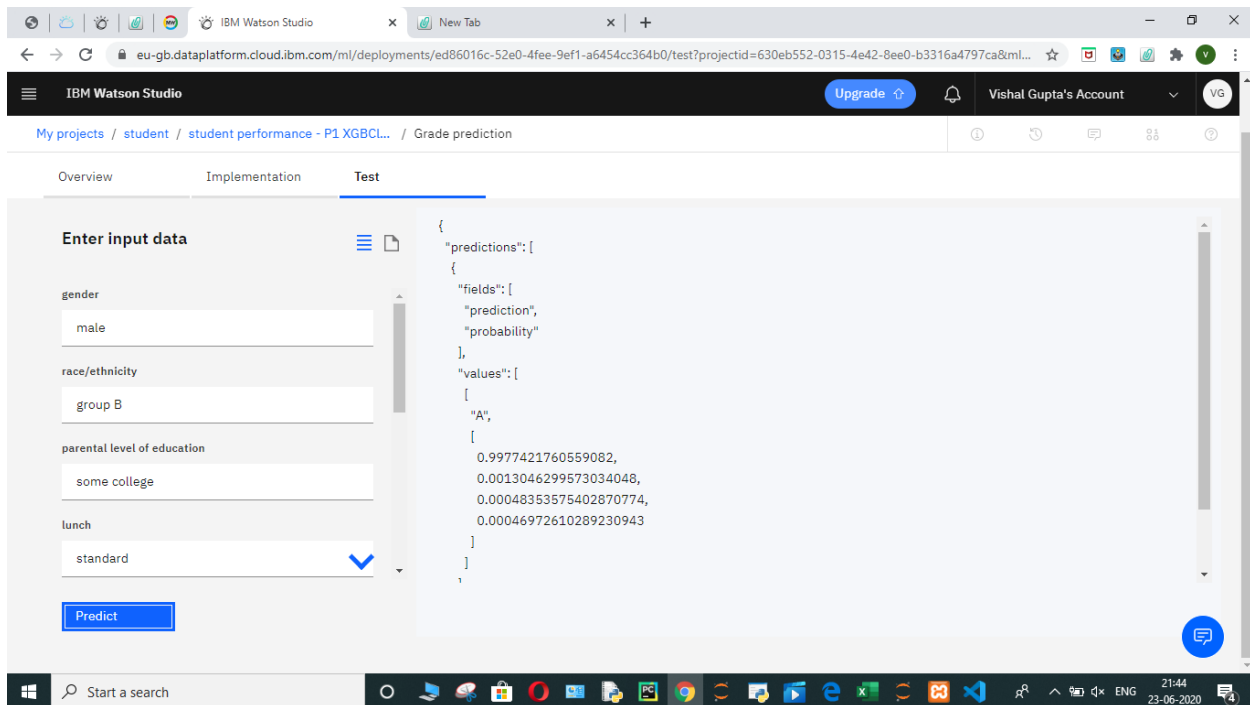

Fig:Test Run of Project

In the above test run it is quite clear that on the basis of inputed data this model predicts the grade

5. FLOWCHART

1 Create Watson Studio Service from the IBM Cloud Catalog secon

2 Inside Watson Studio Service create new AutoAI project and and also create cloud object storage for storing datasets

3 Inside Project under Asset column upload your dataset and click to add this to your project. Also create Associate Service for your project and finally click on

4 Now , its me to select what predictions you want from your datset and click run the experiment. Aer the successful running of experiment save the best fied model based on accuarcyvalue.

5 Before the deployement test your experiment by running the test secon and providing suitable values to experimental analysis.Till now the project is succesfully created.Experiment is done

6 For creang UI(User Interface) for your algorithm. From catalog select soware and then starter kit and here select NodeRed App and click to create web app.Just for the services to get successfully started.

7 Now , making further progress , click to NodeRed service , it will results in opening of NodeRed Flows. Where you can easiliy drag and drop different secons to create your web apps.

6. RESULT

By using the IBM Cloud's Watson AutoAI , we are able to model our given dataset into a good predictable Machine Learning Algorithm and by the help of NodeRed service we are able to turn our Machine Learning Algorithm into a beauful Web App:

Fig:Screenshot of Web App UI

7.ADVANTAGES AND DISADVANTAGES

Advantages:

● *Access all the relevent data seamlessly and quickly.*

● *Idenfying the weak student ans strong student.*

● *Idenfying potenal of student*

Disadvantages:

● *may be the student has some other potential*

●*way of reacting with grades make mentall challenge*

8. APPLICATION

1 use in school

2 use in colleges to predict grades

3 analyse the student is weak and strong in studies

9. CONCLUSION

We are building a Machine Learning model to predict the material 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. This model is to predict the performance of students in exam in grade distribution based by using different input parameters.

11.BIBLIOGRAPHY

https://www.kaggle.com/spscientist/students-performance-in-exam

APPENDIX

SOURCE CODE

[{"id":"4a4967d8.7b9d18","type":"tab","label":"Flow 2","disabled":false,"info":""},{"id":"34e9c3ac.23b9ac","type":"ui_form","z":"4a4967d8.7b9d18","name":"","label":"","group":"fe8adbe7.f7ce58","order":0,"width":0,"height":0,"options":[{"label":"gender","value":"gen","type":"text","required":true,"rows":null},{"label":"race/ethnicity","value":"re","type":"text","required":true,"rows":null},{"label":"parental level of education","value":"ploe","type":"text","required":true,"rows":null},{"label":"lunch","value":"lu","type":"text","required":true,"rows":null},{"label":"test preparation

course","value":"tpc","type":"text","required":true,"rows":null},{"label":"math score","value":"ms","type":"number","required":true,"rows":null},{"label":"reading score","value":"rs","type":"number","required":true,"rows":null},{"label":"writing score","value":"ws","type":"number","required":true,"rows":null},{"label":"percentage","value":"pc","type":"number","required":true,"rows":null}
],"formValue":{"gen":"","re":"","ploe":"","lu":"","tpc":"","ms":"","rs":"","ws":"","pc":""},"payload":"","submit":"Submit","cancel":"Cancel","topic":"grade","x":70,"y":420,"wires":[["82758153.2466e"]]},{"id":"82758153.2466e","type":"function","z":"4a4967d8.7b9d18","name":"PreToken","func":"global.set(\"gen\",msg.payload.gen)\nglobal.set(\"re\",msg.payload.re)\nglobal.set(\"ploe\",msg.payload.ploe)\nglobal.set(\"lu\",msg.payload.lu)\nglobal.set(\"tpc\",msg.payload.tpc)\nglobal.set(\"ms\",msg.payload.ms)\nglobal.set(\"rs\",msg.payload.rs)\nglobal.set(\"ws\",msg.payload.ws)\nglobal.set(\"pc\",msg.payload.pc)\nvar apikey=\"zdTJACYd2dTYPoklzXtMTrEJeDNuPnoPgp94nyXuVqMS\";\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":220,"y":360,"wires":[["42e8f3e9.a6ebfc"]]},{"id":"42e8f3e9.a6ebfc","type":"http request","z":"4a4967d8.7b9d18","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":430,"y":340,"wires":[["24b371c9.3812de"]]},{"id":"24b371c9.3812de","type":"function","z":"4a4967d8.7b9d18","name":"Pre Prediction","func":"var gen = global.get('gen')\nvar re = global.get('re')\nvar ploe = global.get('ploe')\nvar lu = global.get('lu')\nvar tpc = global.get('tpc')\nvar ms = global.get('ms')\nvar rs = global.get('rs')\nvar ws = global.get('ws')\nvar pc = global.get('pc')\nvar token=msg.payload.access_token\nvar instance_id=\"70d1f1c6-aa58-4aba-a50d-81fd9c550607\"\nmsg.headers={'Content-Type': 'application/json',\"Authorization\":\"Bearer \"+token,\"ML-Instance-ID\":instance_id}\nmsg.payload={\"input_data\": [{\"fields\": [\"RowNumber\", \"gender\", \"race/ethnicity\", \"parental level of education\", \"lunch\", \"test preparation course\", \"math score\", \"reading score\", \"writing score\", \"percentage\"], \"values\": [[gen,re,ploe,lu,tpc,ms,rs,ws,pc]]}]}\nreturn msg;","outputs":1,"noerr":0,"x":720,"y":280,"wires":[["d904177e.e648d8"]]},{"id":"d904177e.e648d8","type":"http request","z":"4a4967d8.7b9d18","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"https://eu-gb.ml.cloud.ibm.com/v4/deployments/ed86016c-52e0-4fee-9ef1-a6454cc364b0/predictions","tls":"","persist":false,"proxy":"","authType":"","x":830,"y":200,"wires":[["6970892a.25b258"]]},{"id":"6970892a.25b258","type":"function","z":"4a4967d8.7b9d18","name":"","func":"msg.payload=msg.payload.predictions[0].values[0][0]\nreturn

msg;","outputs":1,"noerr":0,"x":680,"y":100,"wires":[["7caabd9a.404d14","e38093ec.d1dac"]]},{"id":"7caabd9a.404d14","type":"ui_text","z":"4a4967d8.7b9d18","group":"fe8adbe7.f7ce58","order":1,"width":0,"height":0,"name":"","label":"grade","format":"{{msg.payload}}","layout":"row-spread","x":910,"y":340,"wires":[]},{"id":"e38093ec.d1dac","type":"debug","z":"4a4967d8.7b9d18","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"false","x":950,"y":60,"wires":[]},{"id":"fe8adbe7.f7ce58","type":"ui_group","z":"","name":"student

performance","tab":"1fcf3956.f9ef37","order":1,"disp":true,"width":"6","collapse":false},{"id":"1fcf3956.f9ef37","type":"ui_tab","z":"","name":"Home","icon":"dashboard","order":1,"disabled":false,"hidden":false}]