Student Performance in Exam (Grade Analysis) using Watson Auto AI

Submitted By:
Udayan Jadhav
RSIP Career Basic ML 034

Index

1. INTRODUCTION

1.1 Overview:

In this project there are 4 tasks:

1. Data collection
    a. Download Dataset/Create Dataset
2. IBM Cloud account
    a. IBM Cloud Registration
    b. Login To IBM Cloud
    c. Create Cloud Object Storage Service
    d. Create Watson Studio Platform
    e. Create Machine Learning service
3. Model building
    a. Create Project In Watson Studio
    b. Auto AI Experiment In Add Experiment
    c. Setup Your Auto AI Environment
    d. Import Dataset
    e. Run The Model
    f. Selection of Auto-AI Pipeline
    g. Deploy And Test The Model In Watson Studio
4. Application building
    a. Create Node-Red Service
    b. Build UI With Node-Red
    c. Go to manage palette & install dashboard nodes

1.2 Purpose:

With the wide usage of computers and internet, there has recently been a huge increase in publicly available data that can be analysed. Be it online sales information, website traffic, or user habits, data is generated every day. Such a large amount of data presents both a problem and an opportunity. The problem is that it is difficult for humans to analyse such large data. The opportunity is that this type of data is ideal for computers to process, because it is stored digitally in a well-formatted way, and computers can process data much faster than humans. In schools and higher educational institutes, many students have to struggle hard to pass exams since there is no dedicated support offered to students who need special attention in the relevant courses.

2. LITERATURE SURVEY

2.1 Existing problem:

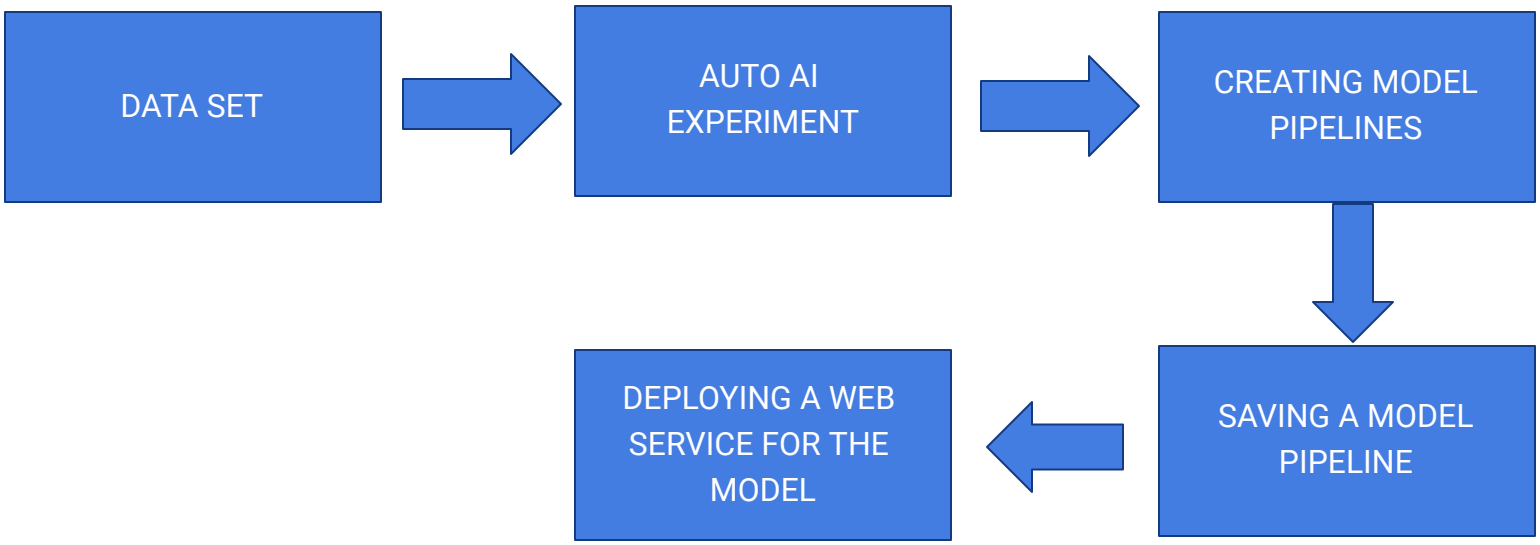
The problem is that it is difficult for humans to analyse very large amount of data.This thesis examines the application of machine learning algorithms to predict whether a student will be successful or not. Machine learning techniques can be utilized for students' grades prediction in different courses. 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:

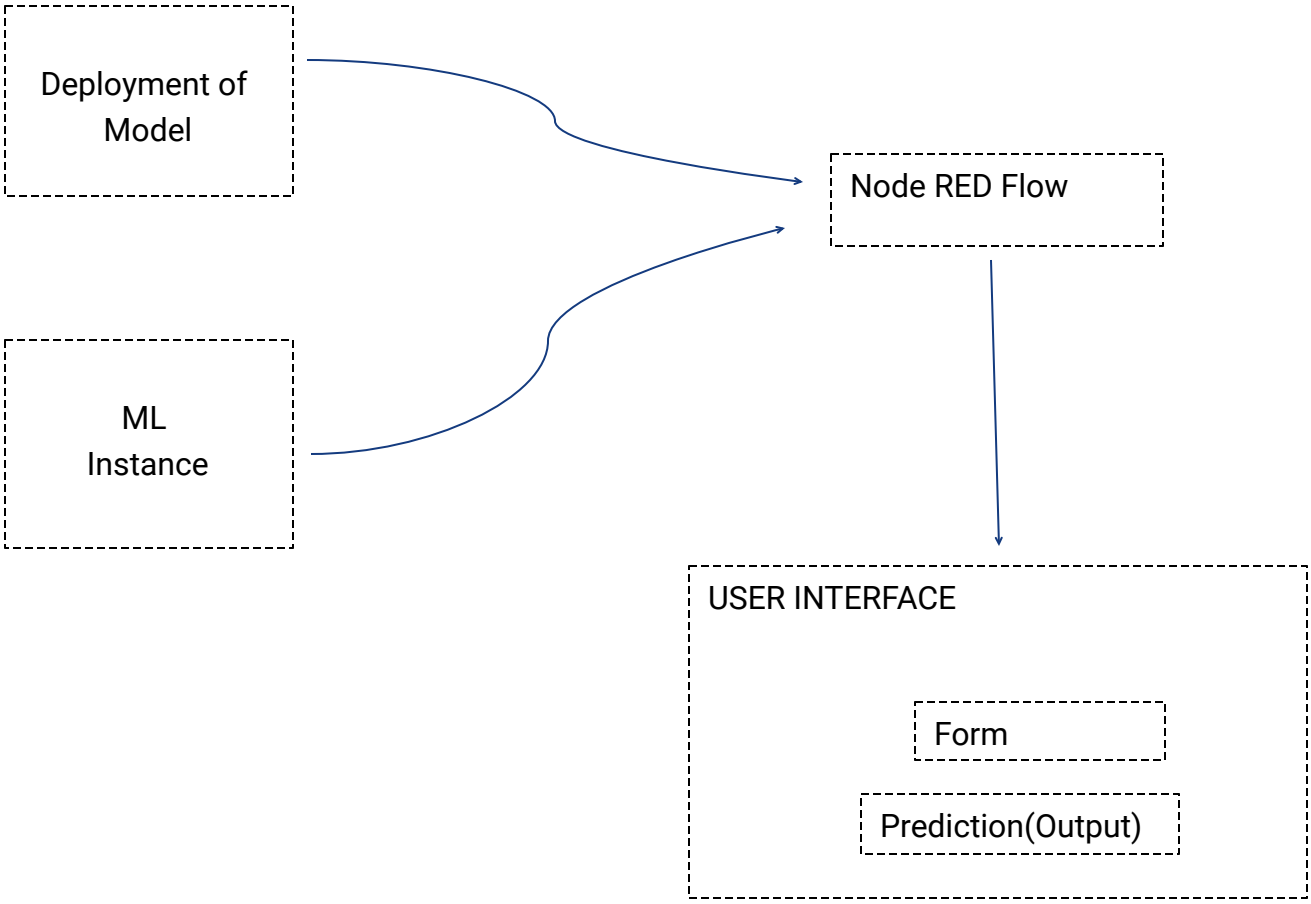
We are building a Machine Learning model to predict the material using IBM Watson AutoAI Machine Learning Service. The model is deployed on IBM cloud to get scoring end point which can be used as API in mobile app or web app building. We are developing a web application which is built using node red 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

3.1 Block diagram:



3.2 Sofware designing:



IBM Cloud Computing:

It is a set of cloud computing services for business offered by the information technology company IBM. IBM Cloud includes infrastructure as a service (IaaS), software as a service (SaaS) and platform as a service (PaaS) offered through public, private and hybrid cloud delivery models, in addition to the components that make up those clouds.

IBM Watson Studio:

IBM Watson Studio helps data science scientists and analysts prepare data and build models at scale across any cloud. With its open, flexible multi cloud architecture, Watson Studio provides capabilities that empower businesses to simplify enterprise data science and AI, such as:

1. Automate AI lifecycle management with Auto AI.
2. Visually prepare and build models with IBM SPSS Modeler.
3. Build models using images with IBM Watson Visual Recogni on and texts with IBM Watson Natural Language Classifier.
4. Deploy and run models through one-click integration with IBM Watson Machine Learning.
5. Manage and monitor models through integration with IBM Watson Open scale.

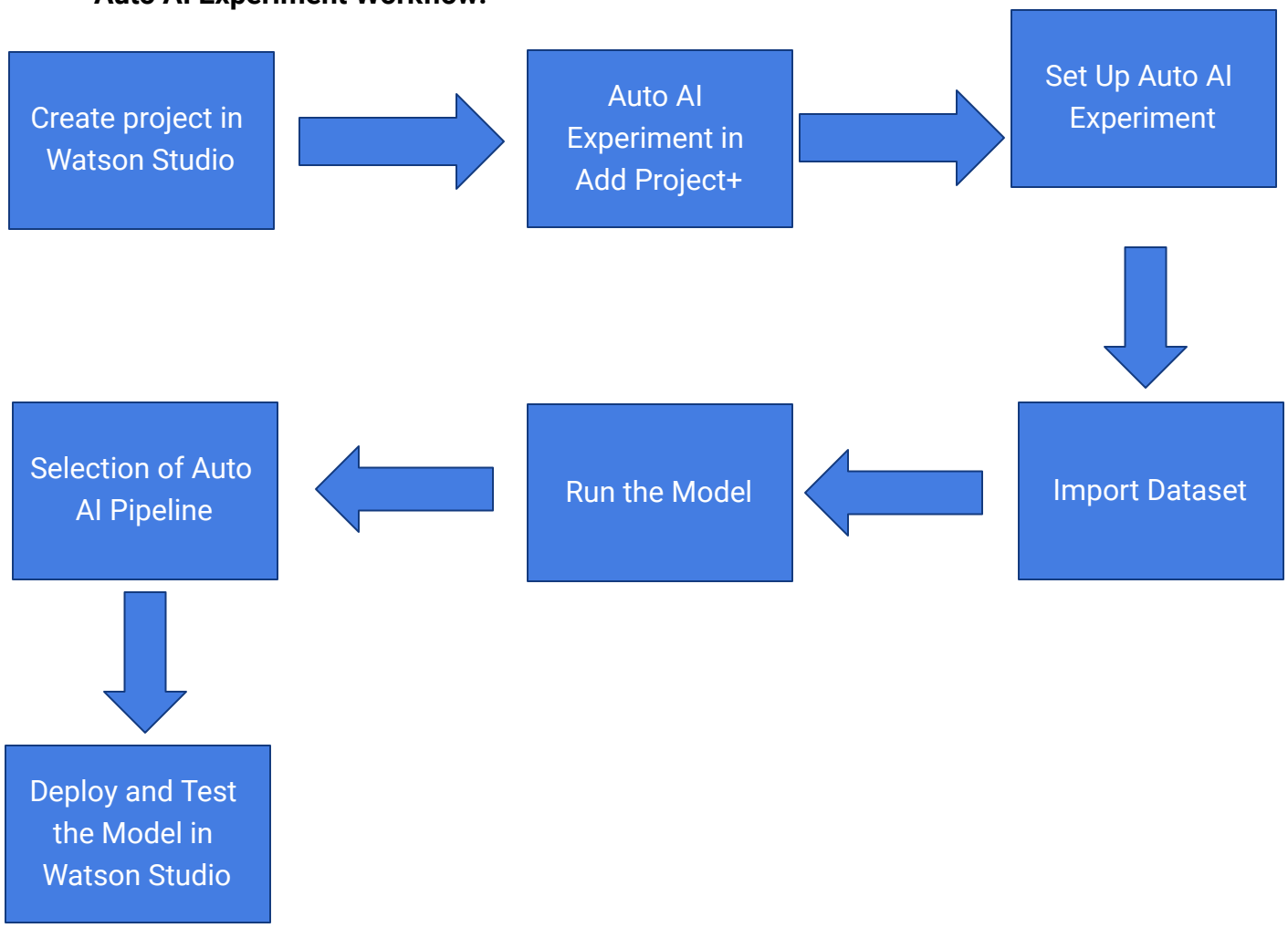
Node Red:

Node-RED is a flow-based development tool for visual programming developed originally by IBM for wiring together hardware devices, APIs and online services as part of the Internet of Things.

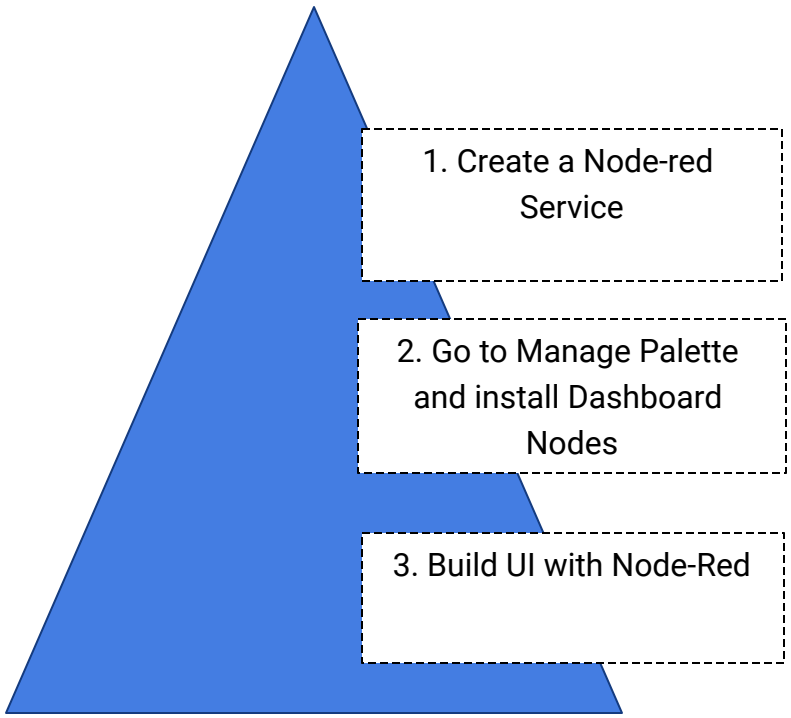
Node-RED provides a web browser-based flow editor, which can be used to create JavaScript functions. Elements of applications can be saved or shared for re-use. The flows created in Node-RED are stored using JSON. Since version 0.14, MQTT nodes can make properly configured TLS connections. In 2016, IBM contributed Node-RED as an open source JS Foundation project.

4. EXPERIMENTAL INVESTIGATIONS

Auto AI Experiment Workflow:



Application Building Workflow



AUTO AI EXPERIMENT

1. In Machine Learning, we create a project in IBM Watson Studio and we add an Auto AI Experiment to the project.
2. We need to import our dataset to calculate the required output. Hence, we import/upload the 'COMPRESSIVE_STRENGTH_CONCRETE' csv file and give 'Concrete Compressive Strength' column of the dataset as the prediction column.
3. Run the model and we get our required Auto AI Pipeline. We get the Relationship Map and Pipeline Comparison. We have to save the Auto AI pipeline as a model.
4. We should deploy the saved model.
5. Then we get to test the deployed model.

APPLICATION BUILDING (NODE-RED)

1. From the catalog of IBM Cloud we should install Node-Red App.
2. In this, we have 'Nodes' to the left side of the page and we use these nodes to create a 'flow'.
3. We need to install 'Dashboard Nodes' by going to the 'Manage Palette' section. ● The nodes which we use in this project are:

Form node: Adds a form to user interface. Helps to collect multiple value from the user on submit button click as an object in msg.payload.

Multiple input elements can be added using add elements button.

Timestamp node: Injects a message into a flow either manually or at regular intervals. The message payload can be a variety of types, including strings, JavaScript objects or the current time.

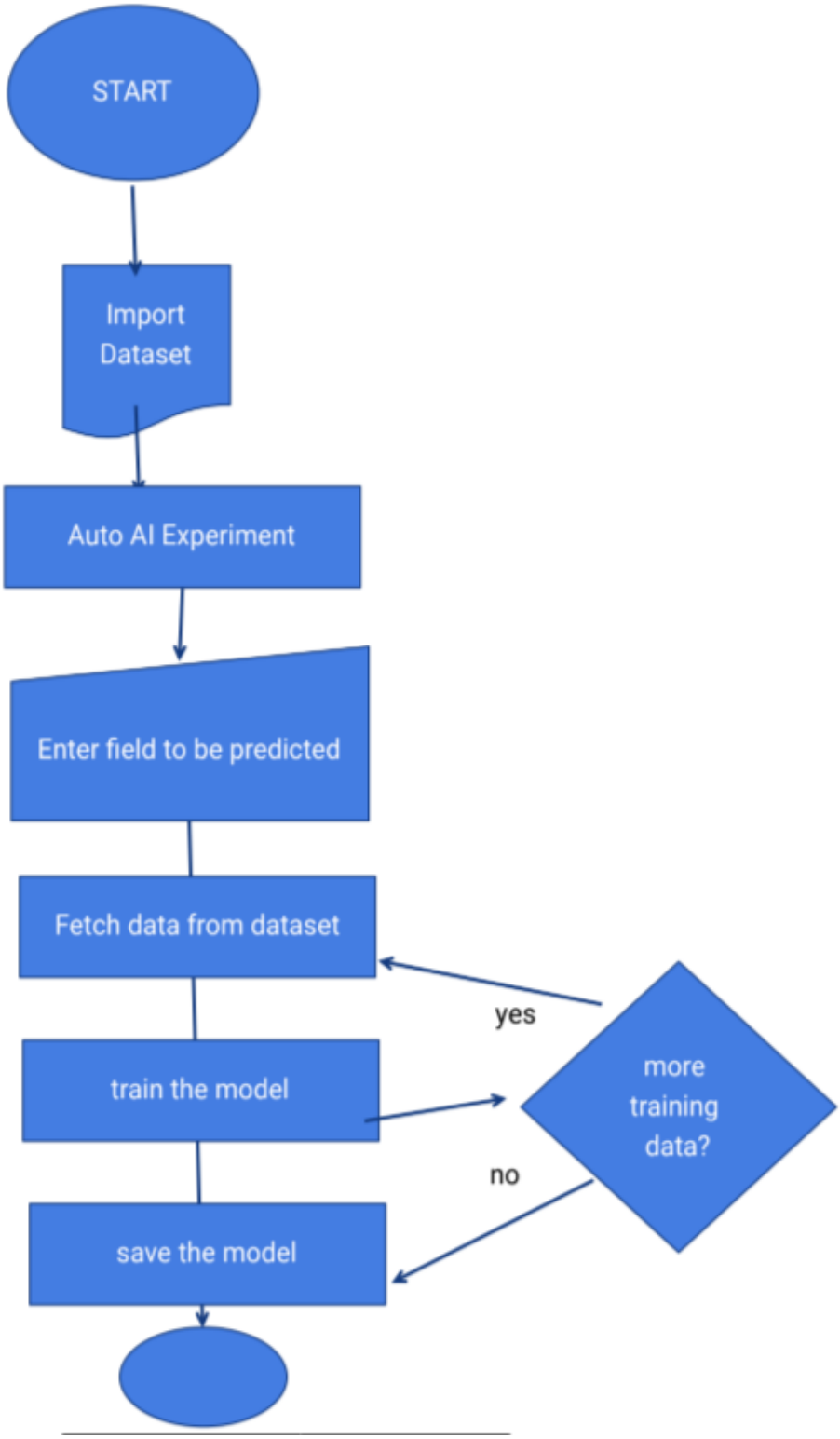
function node: A JavaScript function block to run against the messages being received by the node. The messages are passed in as a JavaScript object called msg. By convention it will have a msg.payload property containing the body of the message. The function is expected to return a
message object (or multiple message objects), but can choose to return nothing in order to halt a flow.

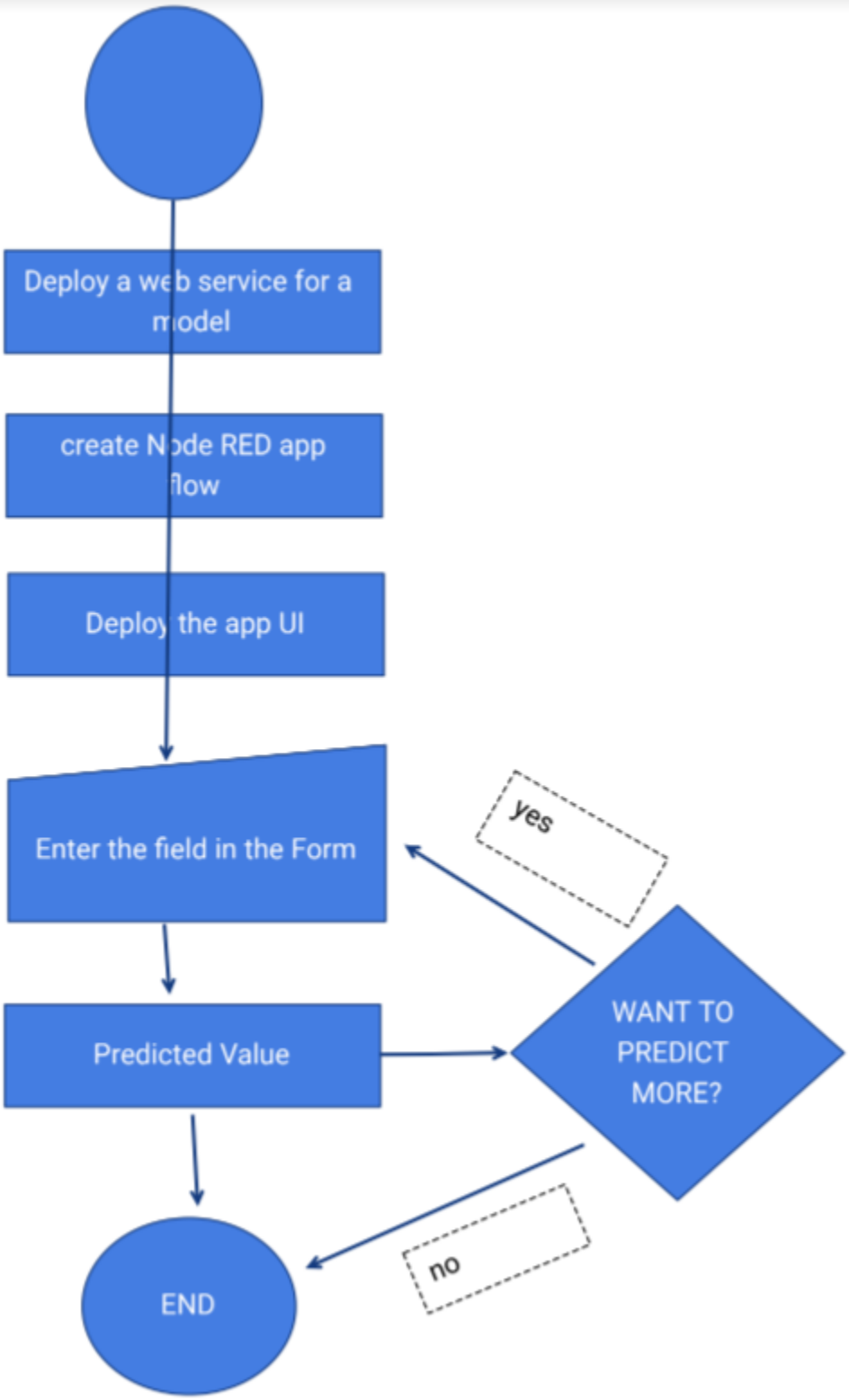
http request node: Sends HTTP requests and returns the response.

Text node: Will display a non-editable text field on the user interface. Each received msg.payload will update the text based on the provided Value Format.The Value Format field can be used to change the displayed format and can contain valid HTML and Angular filters.

Debug node: Displays selected message properties in the debug sidebar tab and optionally the runtime log. By default it displays msg.payload, but can be configured to display any property, the full message or the result of a JSONata expression.

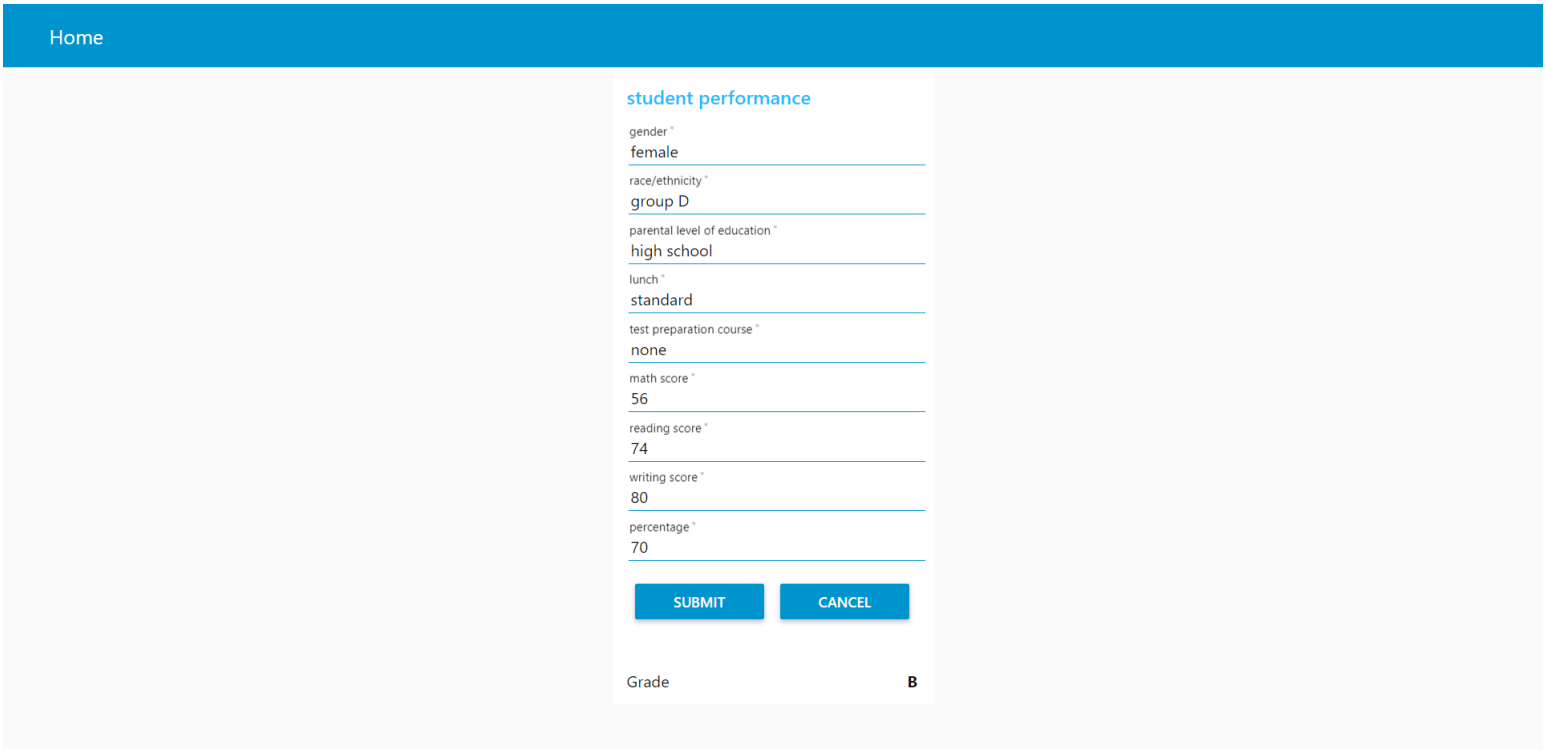
5. FLOWCHART


START
Import Dataset
Auto AI Experiment
Enter field to be predicted
Fetch data from dataset
train the model
save the model
more training data?
yes
no


Deploy a web service for a model
create Node RED app flow
Deploy the app UI
Enter the field in the Form
Predicted Value
WANT TO PREDICT MORE?
yes
no
END

6. RESULT

By creating the flow, we obtained the required output. The predicted value obtained in IBM Watson Studio and the predicted value obtained in Node-Red matches when we give the same input values in both the applications. Hence, the flow is successfully built without any errors.



7. ADVANTAGES & DISADVANTAGES

Advantages

- ☐ Access all the relevant data seamlessly and quickly.
- ☐ Idenfying the weak student ans strong student.
- ☐ Idenfying potential of student
- ☐ Easily identifies trends and patterns.
- ☐ No human intervention needed (automation)
- ☐ Continuous Improvement
- ☐ Handling multi-dimensional and multi-variety data
- ☐ Wide application

Disadvantages

- ☐ Data Acquisition
- ☐ Time and Resources
- ☐ Interpretation of results
- ☐ High error susceptibility.

8. APPLICATIONS

- A student can focus in different subjects and courses by predicting his/her grade with many different input parameters.
- Can be used in schools, colleges or other teaching institutes for better learning of the students

9. CONCLUSION

To solve the problem of identifying under-performing students in higher education, an ML model has been developed to predict student performance. The ML model is built for the auto AI model only. Models are compared with each other using the ROC Rate Indicator and location accuracy.

In addition, the decision tree model has shown that not all attributes are involved in the classification process. Computer Grade-Course1, Residence, Interest in Computer Learning, Satisfaction in Educational Environment, and Residency are attributes used by the decision tree model.

10. FUTURE SCOPE

As the continued growth of Auto AI and Machine Learning algorithms make them a great source of information and predictable results, models can be used to improve our performance in different fields and will reduce human effort. Auto AI trials also have the potential for great work.
The “Student Performance in Exam (Grade Analysis)” project can lead to the support of volunteers who require special attention in various subjects. Students can pay attention and guess for future grades.

11. BIBLIOGRAPHY

Smart Bridge: Machine Learning career basic
https://smartbridge.teachable.com/

Source of Dataset:
https://www.kaggle.com/spscientist/students-performance-in-exams

IBM Cloud
https://www.ibm.com/cloud

Node-Red:
https://node-red-qbfab.eu-gb.mybluemix.net/

APPENDIX

A.Source code

[{"id":"3b9639c.57d2ec6","type":"tab","label":"Flow 2","disabled":false,"info":""},{"id":"f8dffc0.343f308","type":"ui_form","z":"3b9639c.57d2ec6","name":"","label":"","group":"c4fa25c7.daa018","order":1,"width":0,"height":0,"options":[{"label":"gender","value":"ge","type":"text","required":true,"rows":null},{"label":"race/ethnicity","value":"re","type":"text","required":true,"rows":null},{"label":"parental level of education","value":"pa","type":"text","required":true,"rows":null},{"label":"lunch","value":"lun","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":"per","type":"number","required":true,"rows":null}],"formValue":{"ge":"","re":"","pa":"","lun":"","tpc":"","ms":"","rs":"","ws":"","per":""},"payload":"","submit":"Submit","cancel":"Cancel","topic":"grade","x":70,"y":420,"wires":[["6880a16.729846"]]},{"id":"6880a16.729846","type":"function","z":"3b9639c.57d2ec6","name":"PreToken","func":"global.set(\"ge\",msg.payload.ge)\nglobal.set(\"re\",msg.payload.re)\nglobal.set(\"pa\",msg.payload.pa)\nglobal.set(\"lun\",msg.payload.lun)\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(\"per\",msg.payload.per)\nvar apikey=\"WqbhkaAzXH4ygf-c-BYlVprZ9NHjXp1IAuL-Wx8FJU4z\";\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":[["4b730c85.a28844"]]},{"id":"4b730c85.a28844","type":"http request","z":"3b9639c.57d2ec6","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":[["a37c73dd.6b2ad"]]},{"id":"a37c73dd.6b2ad","type":"function","z":"3b9639c.57d2ec6","name":"Pre Prediction","func":"var ge = global.get('ge')\nvar re = global.get('re')\nvar pa = global.get('pa')\nvar lun = global.get('lun')\nvar tpc = global.get('tpc')\nvar ms = global.get('ms')\nvar rs = global.get('rs')\nvar ws = global.get('ws')\nvar per = global.get('per')\nvar token=msg.payload.access_token\nvar instance_id=\"219885bb-0d13-4260-adab-3e36491e7fda\"\nmsg.headers={'Content-Type': 'application/json',\"Authorization\":\"Bearer \"+token,\"ML-Instance-ID\":instance_id}\nmsg.payload={\"input_data\": [{\"fields\": [\"gender\", \"race/ethnicity\", \"parental level of education\", \"lunch\", \"test preparation course\", \"math score\", \"reading score\", \"writing score\", \"percentage\"], \"values\":

[[ge,re,pa,lun,tpc,ms,rs,ws,per]]}]}\nreturn
msg;","outputs":1,"noerr":0,"x":720,"y":280,"wires":[["439214f0.4b86cc"]]},{"id":"439214f0.4
b86cc","type":"http
request","z":"3b9639c.57d2ec6","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"
https://us-south.ml.cloud.ibm.com/v4/deployments/e6535b89-4845-4ebb-a443-3fcdef
37aa48/predictions","tls":"","persist":false,"proxy":"","authType":"","x":830,"y":200,"wires":[["fd
fc56ca.656e38"]]},{"id":"fdfc56ca.656e38","type":"function","z":"3b9639c.57d2ec6","name"
:"","func":"msg.payload=msg.payload.predictions[0].values[0][0]\nreturn
msg;","outputs":1,"noerr":0,"x":680,"y":100,"wires":[["6b03e5ab.70ed1c","9b88081b.545828
"]]},{"id":"6b03e5ab.70ed1c","type":"ui_text","z":"3b9639c.57d2ec6","group":"c4fa25c7.daa
018","order":2,"width":0,"height":0,"name":"","label":"Grade","format":"{{msg.payload}}","layo
ut":"row-spread","x":910,"y":340,"wires":[]},{"id":"9b88081b.545828","type":"debug","z":"3b96
39c.57d2ec6","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"comp
lete":"payload","targetType":"msg","x":950,"y":60,"wires":[]},{"id":"c4fa25c7.daa018","type":"u
i_group","z":"","name":"student
performance","tab":"dbbd121c.3f3dd","order":2,"disp":true,"width":"6","collapse":false},{"id":
"dbbd121c.3f3dd","type":"ui_tab","z":"","name":"Home","icon":"dashboard","order":2,"disable
d":false,"hidden":false}]