Assignment3: Deploying Data Science Models

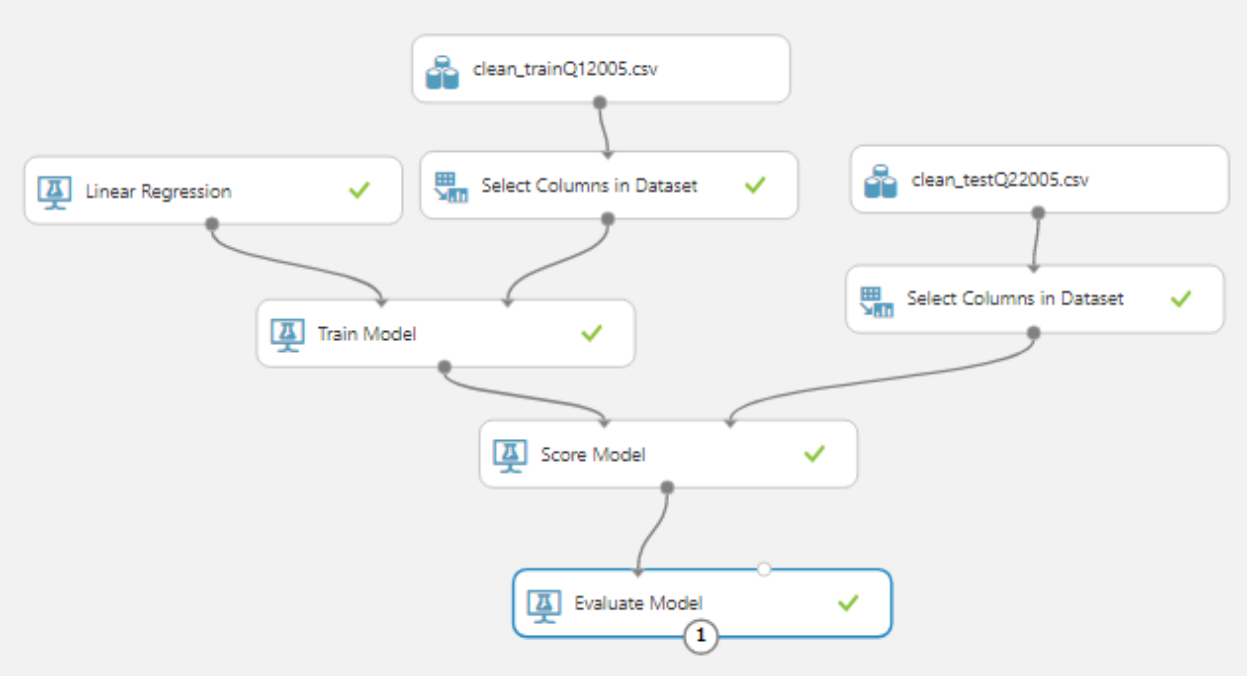
Team7: Snigdha Joshi & Vipra Shah

This report summarizes deploying models for regression and classification using Azure ML and creating 6 REST APIs for the same. Also, we have built a web app using python and used REST APIs, deployed the web app on a cloud environment heroku. Link to web application is: https://data-model.herokuapp.com/

8/3/2017

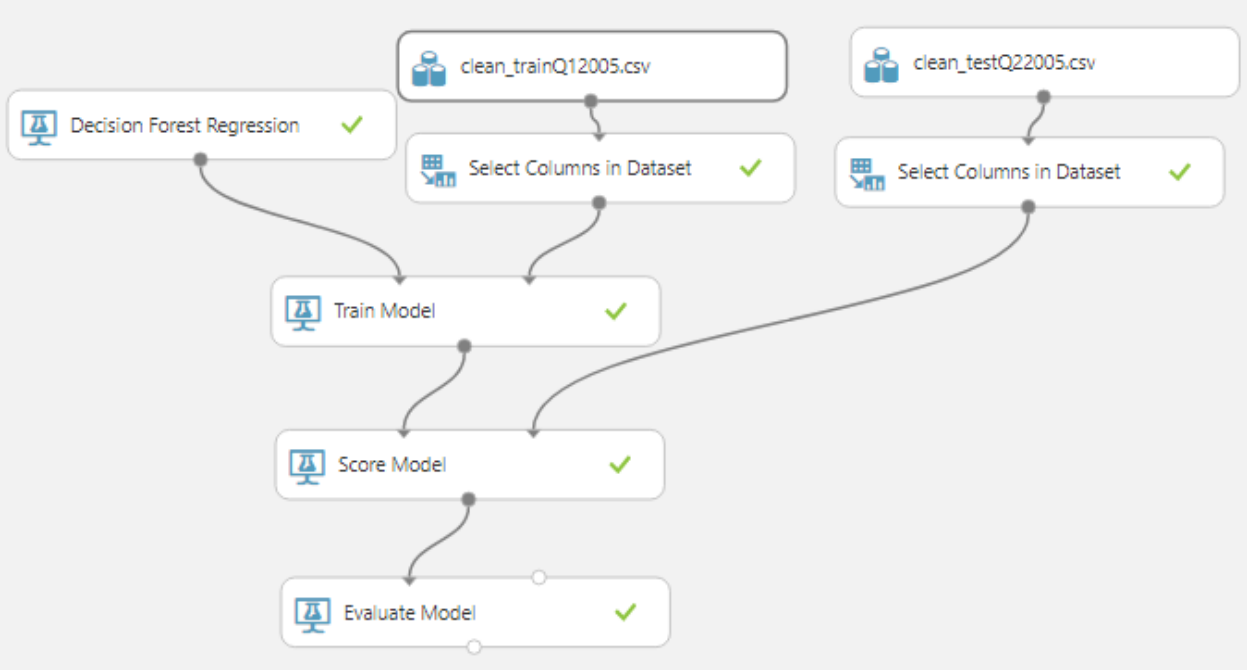
Prediction

We have used Microsoft Azure to train and test models.

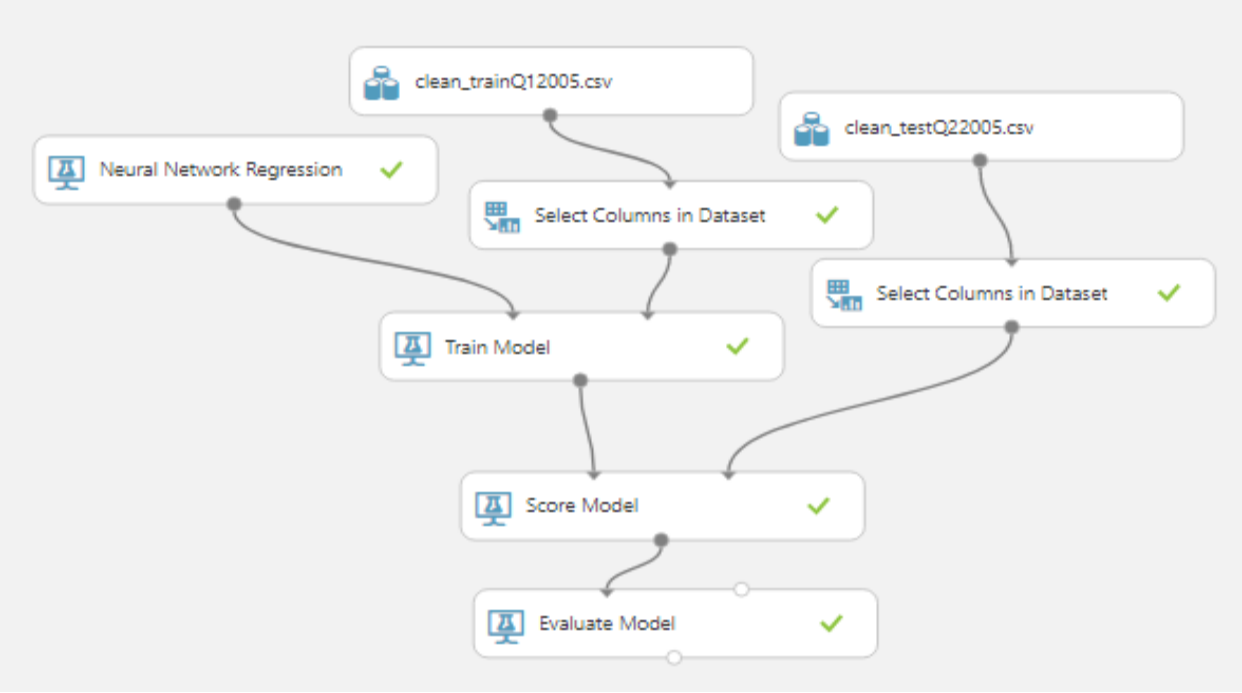


Metrics

Mean Absolute Error	0.248942
Root Mean Squared Error	0.323416
Relative Absolute Error	0.949682
Relative Squared Error	0.872777
Coefficient of Determination	0.127223



Negative Log Likelihood	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Relative Squared Error	Coefficient of Determination
121258.640516	0.246127	0.32016	0.938943	0.855295	0.144705

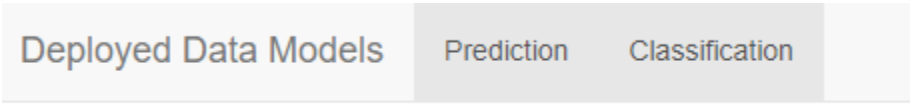


Metrics

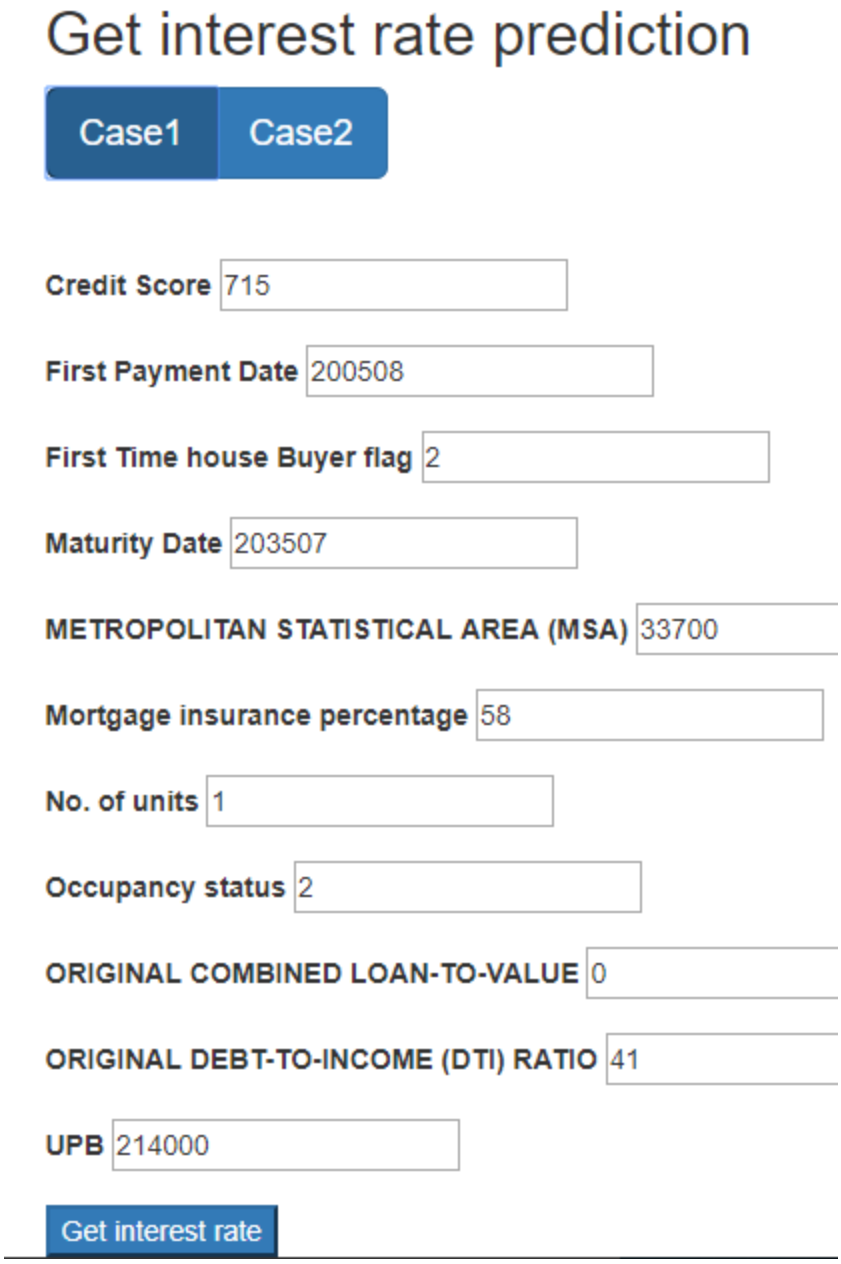
Mean Absolute Error	0.254589
Root Mean Squared Error	0.328668
Relative Absolute Error	0.971228
Relative Squared Error	0.901355
Coefficient of Determination	0.098645

By deploying above models on web service we have tested it for various cases.

- By clicking on Prediction tab we can test cases for prediction.



- By clicking on Case1 or Case2 button data for case can be auto filled.



- The Actual and estimated interest rate all models can be seen on success page.

Welcome to Data Model using Microsoft Azure ML

Actual interest rate is: 5.75

Linear Regression:

Predicted interest rate is: 5.70965135411894

Random Forest:

Predicted interest rate is: 5.63768476471299

Neural Network:

Predicted interest rate is: 5.91126728057861

- Test the model for your custom case for Q22005:

  credit_score: min= 306, max = 850

  first_payment_date: format: YYYYMM

  fthb_flag: 1,2,3

  matr_date: format: YYYYMM

  msa: min=10180, max = 49740

  mortage_insurance_pct: min=0, max = 160

  no_of_units: min=0 , max = 4

  occupancy_status: min=1 , max =3

  cltv: min= 0 , max= 54

  dti_ratio: min= 0 , max=65

  original_upb: min= 6000 , max =692000

estimated interest rate for 3 models:

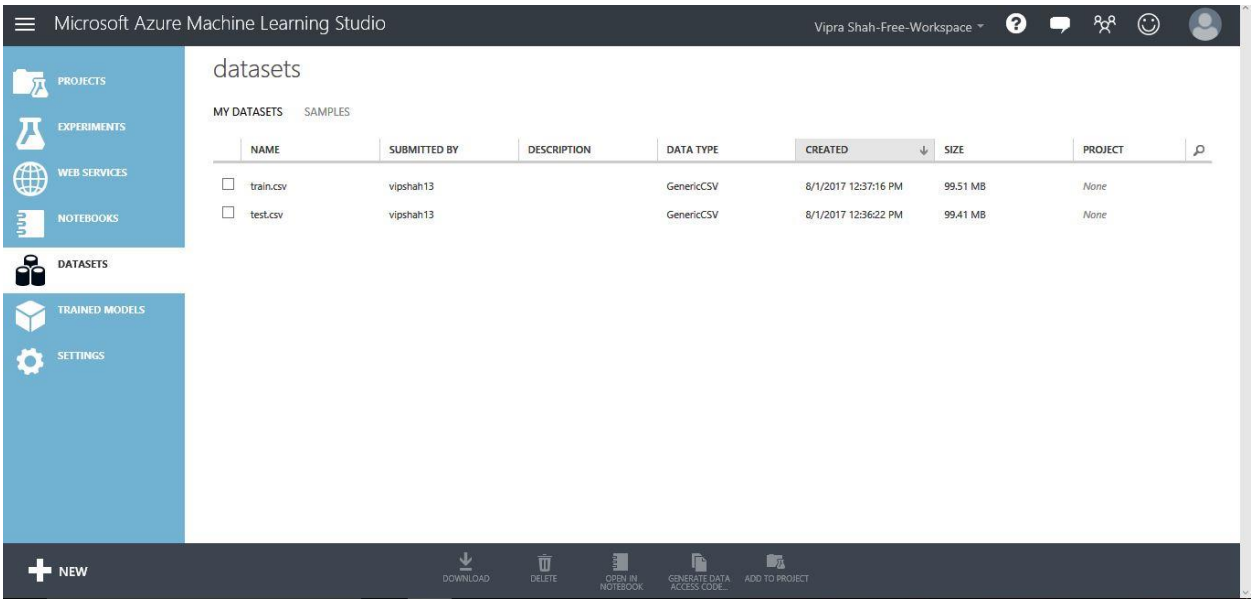
Model	Min	Max
Linear Regression	4.7238	6.3602
Decision Forest	5.1292	7.0921
Neural Networks	4.8708	6.6941

If you want to test for other quarters minimum and maximum values and dates may vary.

Classification

❖ Create Datasets:

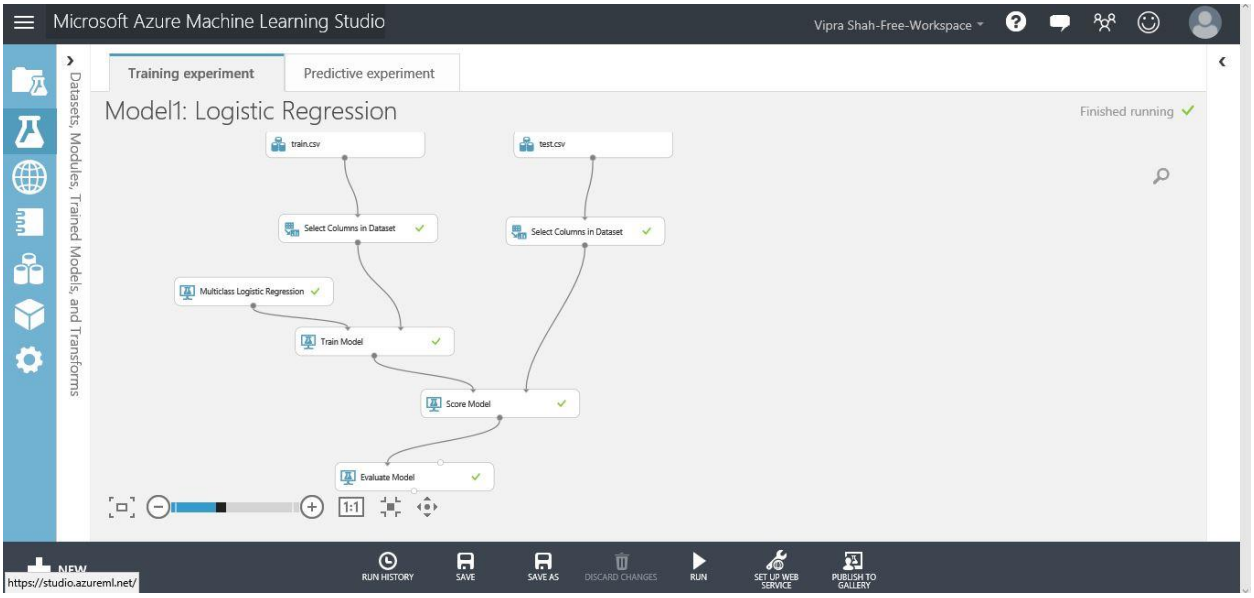
Using Download Data.ipynb file, generate clean train and test csv files for input quarters. Here, we have used Q11999 and Q21999 as train and test quarter. Create datasets in Azure ML studio for train and test csv files.

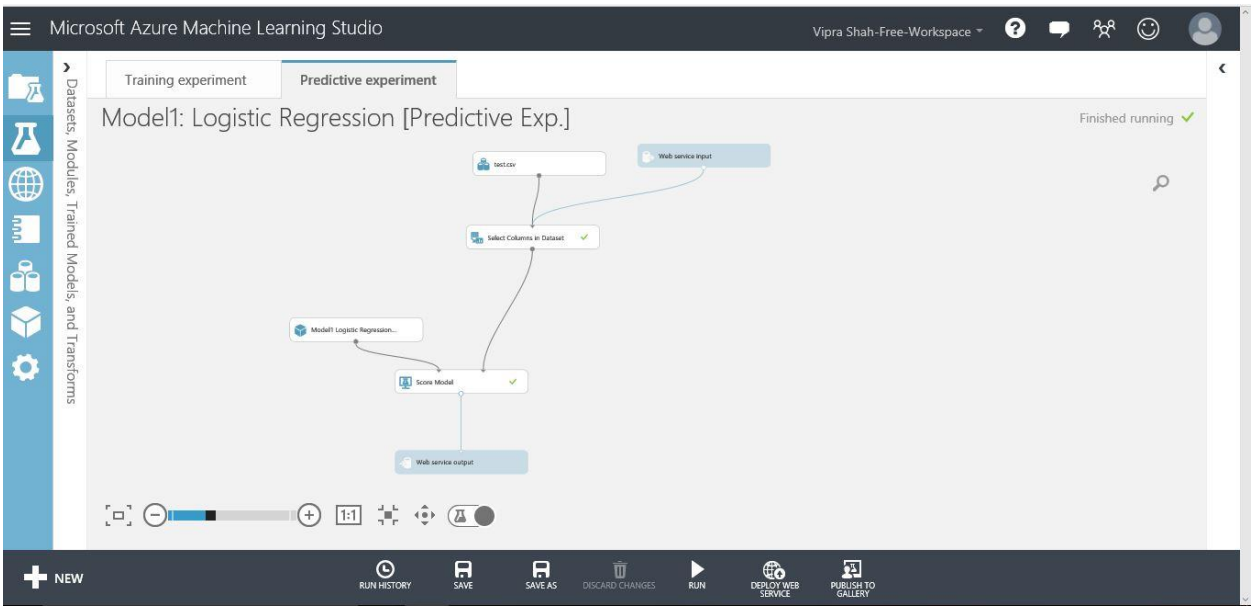


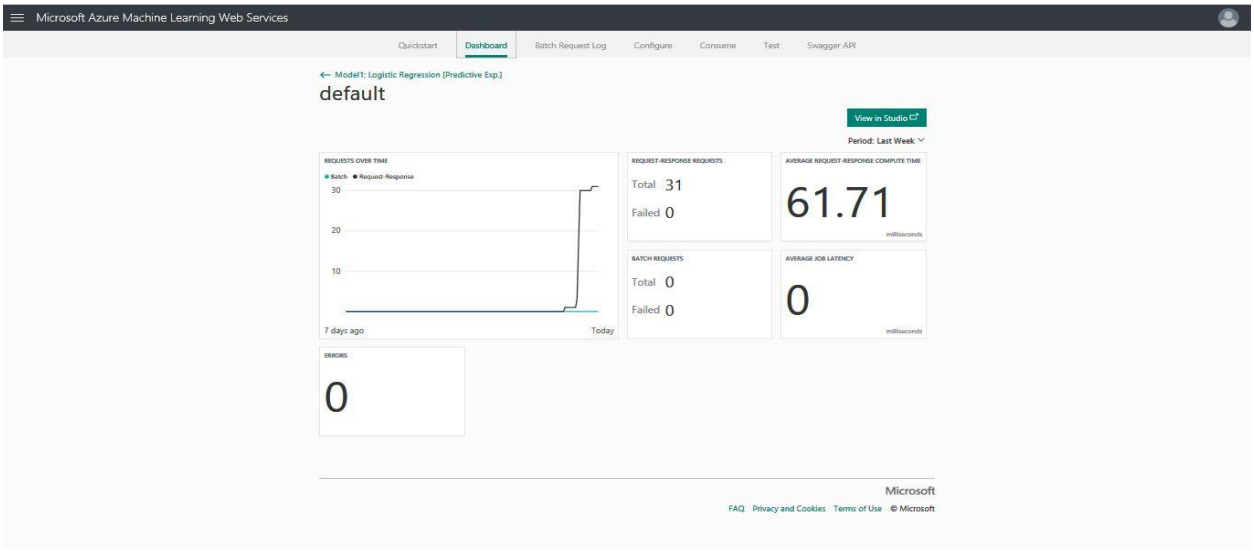
❖ Generate ML Models for classification:

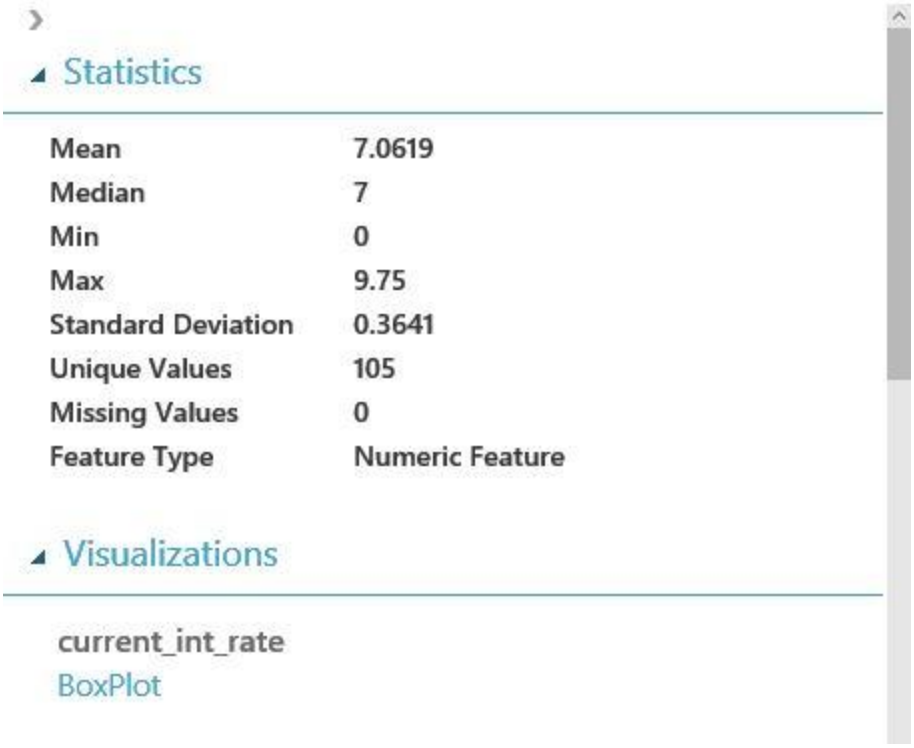
1. Logistic Regression

We have used multiclass logistic regression model, and selected current_actual_upb,delq_status,current_int_rate,current_def_upb,ddlpi,Deliquent columns from test and train csv and trained model for Deliquent column. After, scoring model we have deployed the web service for logistic regression.



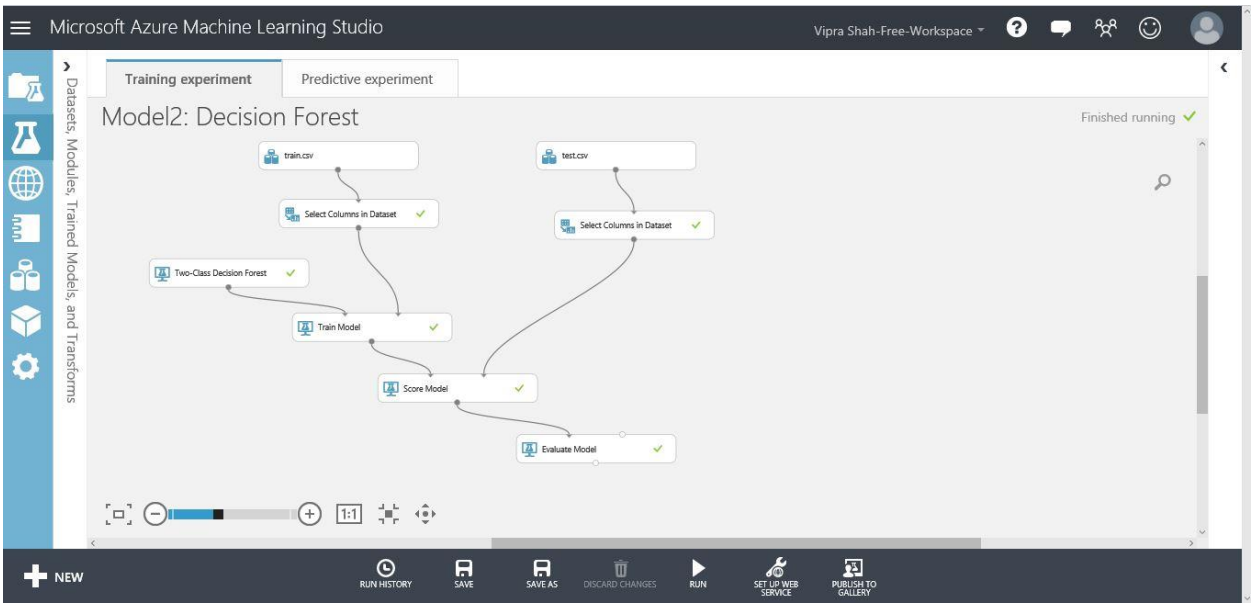

Microsoft Azure Machine Learning Studio
Vipra Shah-Free-Workspace
Datasets, Modules, Trained Models, and Transforms
Training experiment
Predictive experiment
Model1: Logistic Regression [Predictive Exp.]
Finished running
test.csv
Web service input
Select Columns in Dataset
Model1 Logistic Regression...
Score Model
Web service output
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
DEPLOY WEB SERVICE
PUBLISH TO GALLERY

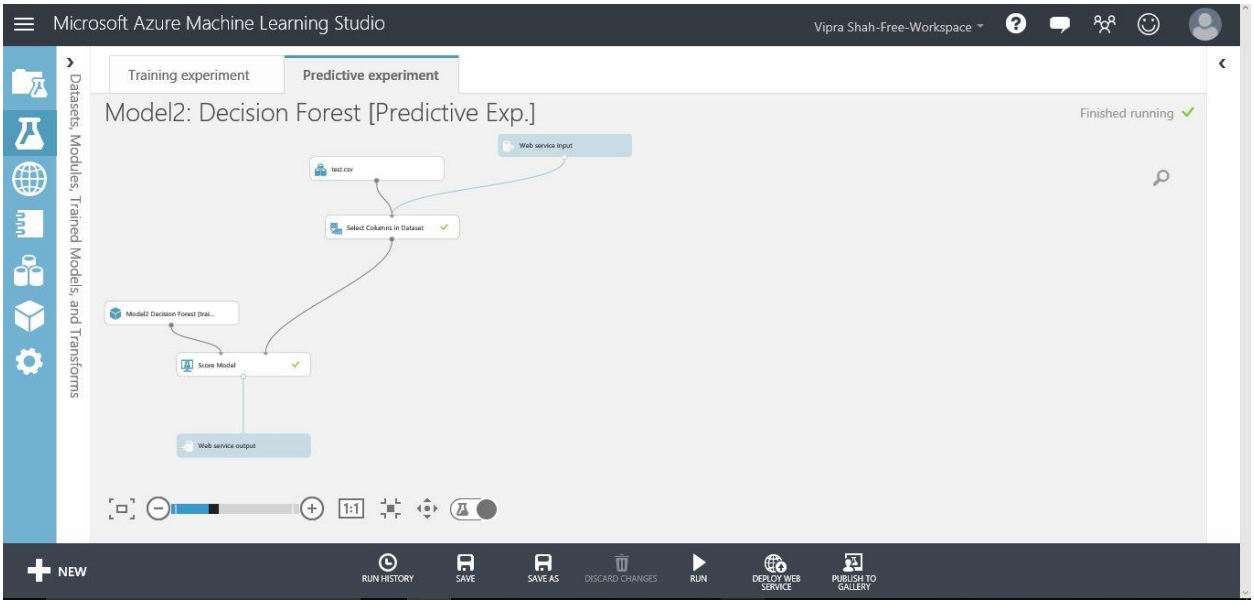

Microsoft Azure Machine Learning Web Services
Quickstart
Dashboard
Batch Request Log
Configure
Consume
Test
Swagger API
Model1: Logistic Regression [Predictive Exp.]
default
View in Studio
Period: Last Week
REQUESTS OVER TIME
Batch
Request-Response
30
20
10
7 days ago
Today
REQUEST-RESPONSE REQUESTS
Total 31
Failed 0
AVERAGE REQUEST-RESPONSE COMPUTE TIME
61.71
milliseconds
BATCH REQUESTS
Total 0
Failed 0
AVERAGE JOB LATENCY
0
milliseconds
ERRORS
0
Microsoft
FAQ Privacy and Cookies Terms of Use © Microsoft

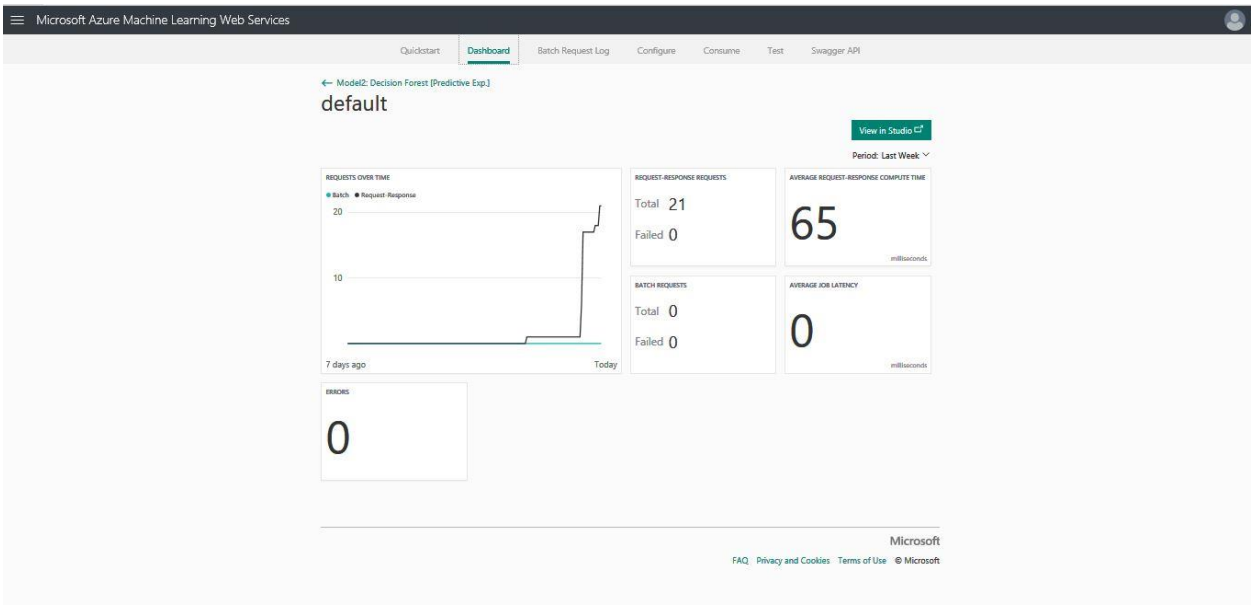


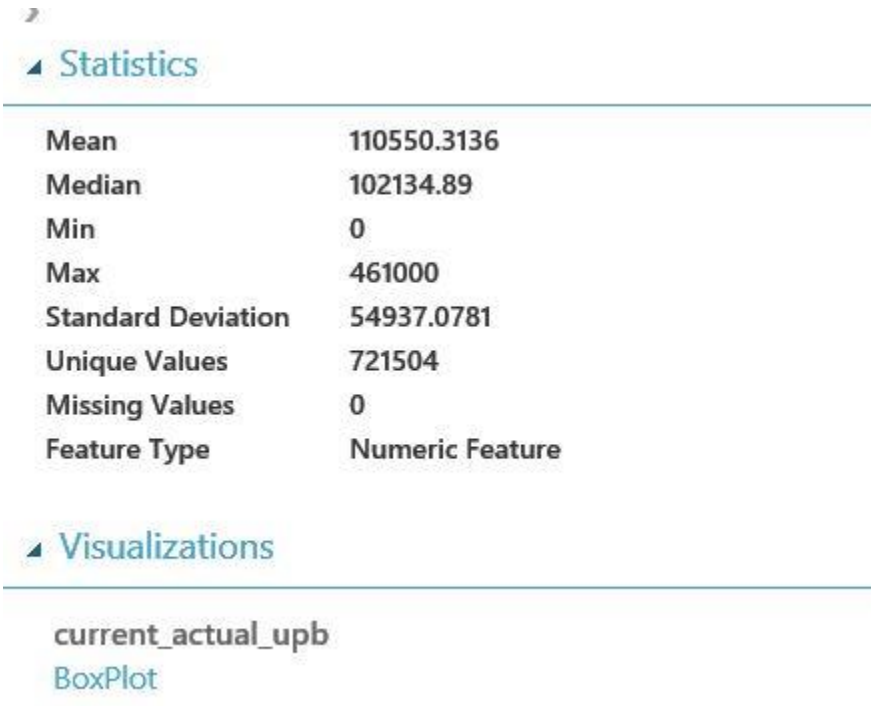
2. Decision Forest:

We have used two class decision forest model, and selected current_actual_upb,delq_status,current_int_rate,current_def_upb,ddlpi,Deliquent columns from test and train csv and trained model for Deliquent column. After, scoring model we have deployed the web service for decision forest algorithm.


Microsoft Azure Machine Learning Studio
Vipra Shah-Free-Workspace
Training experiment
Predictive experiment
Model2: Decision Forest
Finished running
train.csv
test.csv
Select Columns in Dataset
Select Columns in Dataset
Two-Class Decision Forest
Train Model
Score Model
Evaluate Model
Datasets, Modules, Trained Models, and Transforms
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY

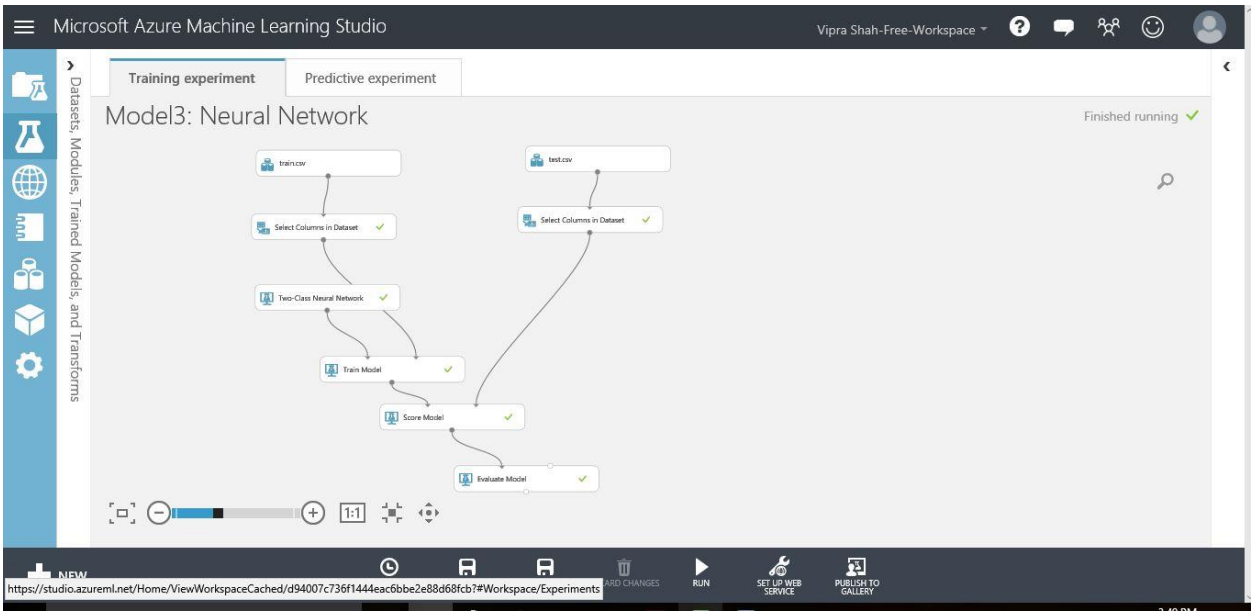

Microsoft Azure Machine Learning Studio
Vipra Shah-Free-Workspace
Training experiment
Predictive experiment
Model2: Decision Forest [Predictive Exp.]
Finished running
Web service input
test.csv
Select Columns in Dataset
Score Model
Web service output
Datasets, Modules, Trained Models, and Transforms
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
DEPLOY WEB SERVICE
PUBLISH TO GALLERY

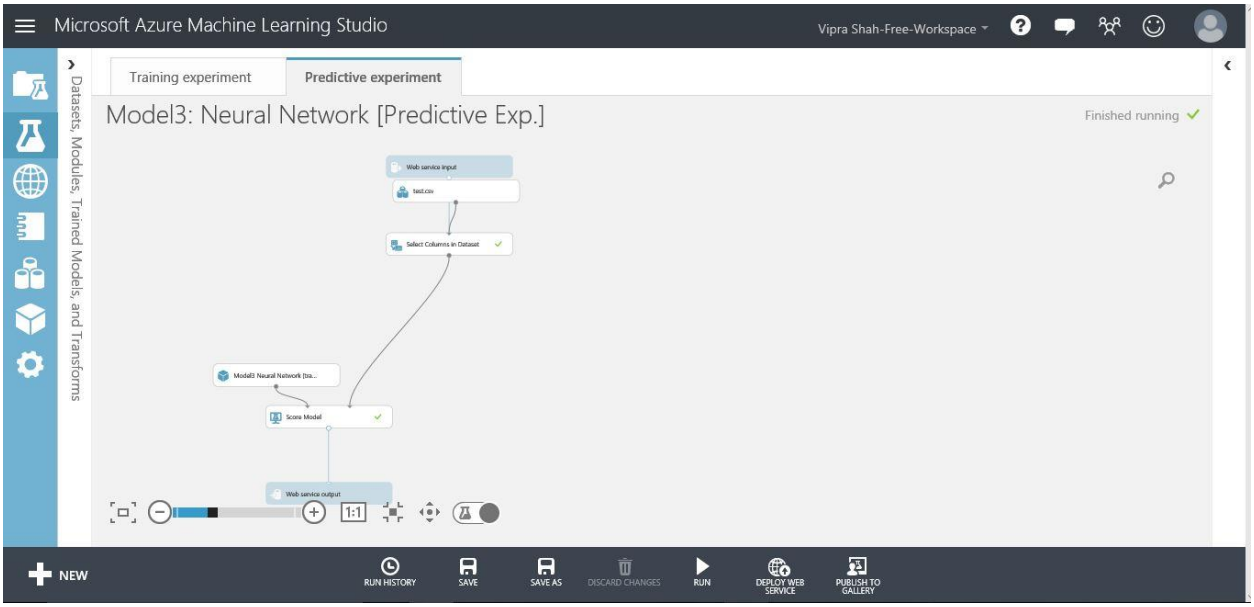


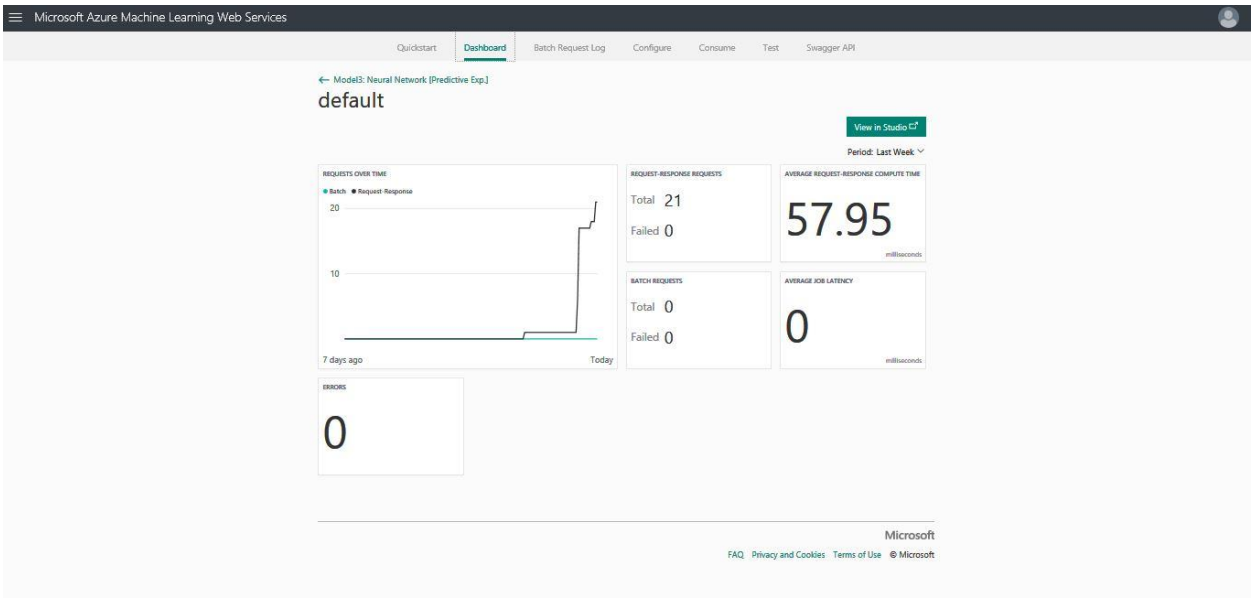


3. Neural Network:

We have used two class neural network model, and selected current_actual_upb,delq_status,current_int_rate,current_def_upb,ddlpi,Deliquent columns from test and train csv and trained model for Deliquent column. After, scoring model we have deployed the web service for neural network algorithm.

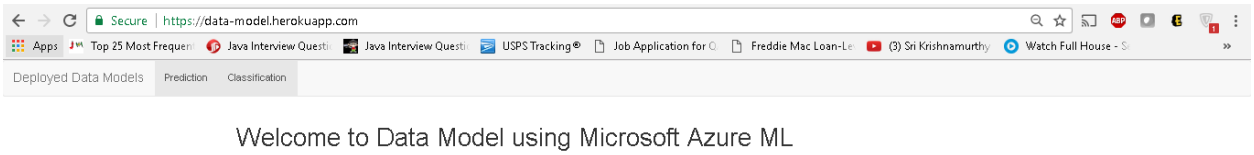

Microsoft Azure Machine Learning Studio
Vipra Shah-Free-Workspace
Training experiment
Predictive experiment
Model3: Neural Network
Finished running
Datasets, Modules, Trained Models, and Transforms
train.csv
test.csv
Select Columns in Dataset
Select Columns in Dataset
Two-Class Neural Network
Train Model
Score Model
Evaluate Model
1:1
NEW
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
https://studio.azureml.net/Home/ViewWorkspaceCached/d94007c736f1444eac6bbe2e88d68fcb?#Workspace/Experiments


Microsoft Azure Machine Learning Studio
Vipra Shah-Free-Workspace
Training experiment
Predictive experiment
Model3: Neural Network [Predictive Exp.]
Finished running
Datasets, Modules, Trained Models, and Transforms
Web service input
test.csv
Select Columns in Dataset
Score Model
Web service output
1:1
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
DEPLOY WEB SERVICE
PUBLISH TO GALLERY



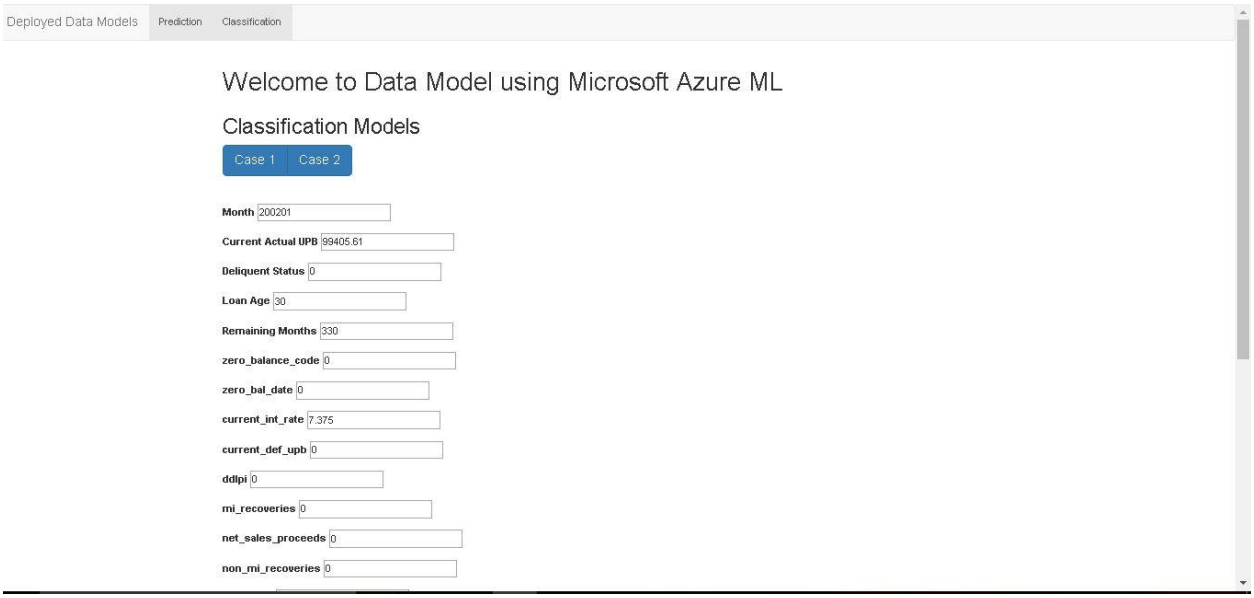
❖ Deploy web app on cloud and invoke REST APIs:

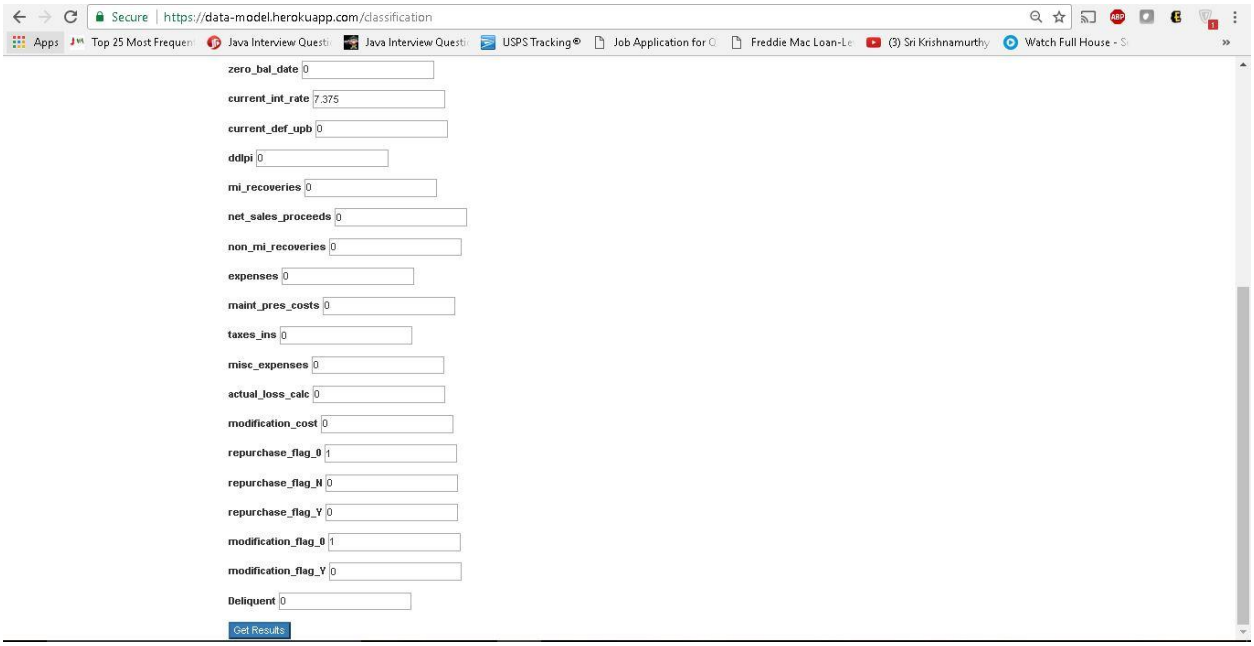
We have deployed our app on heroku: https://data-model.herokuapp.com/



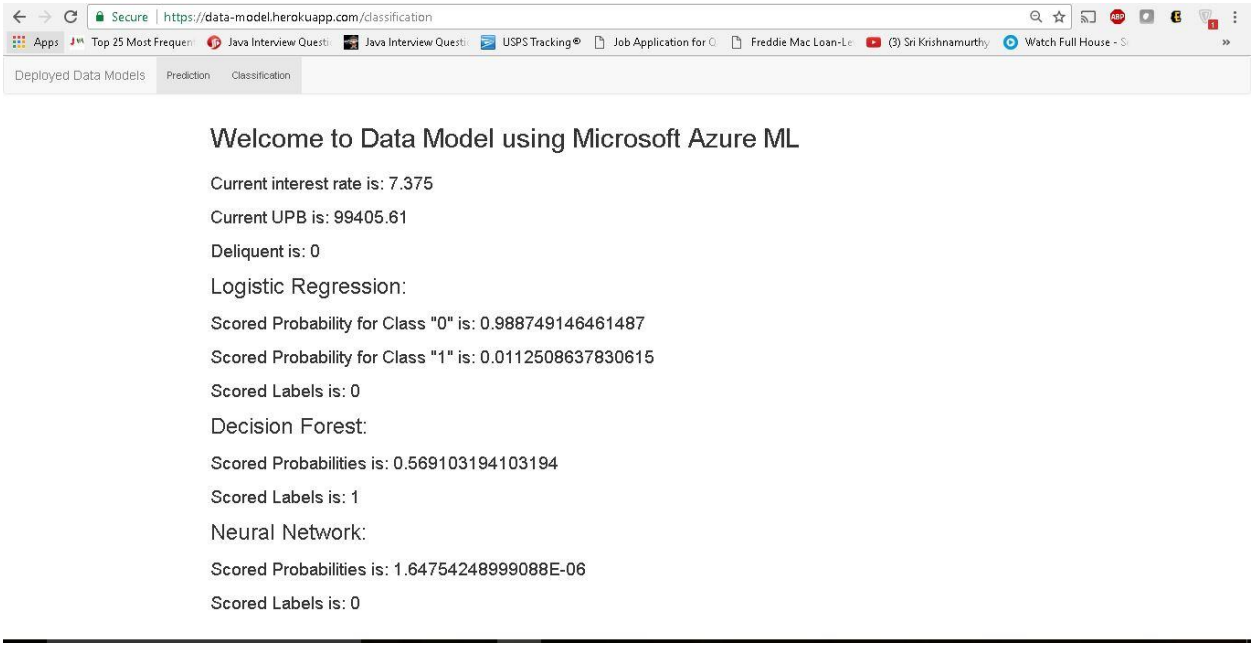
I. Classification TestCase1:

200201	99405.61	0	30	330	0	0	7.375	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0



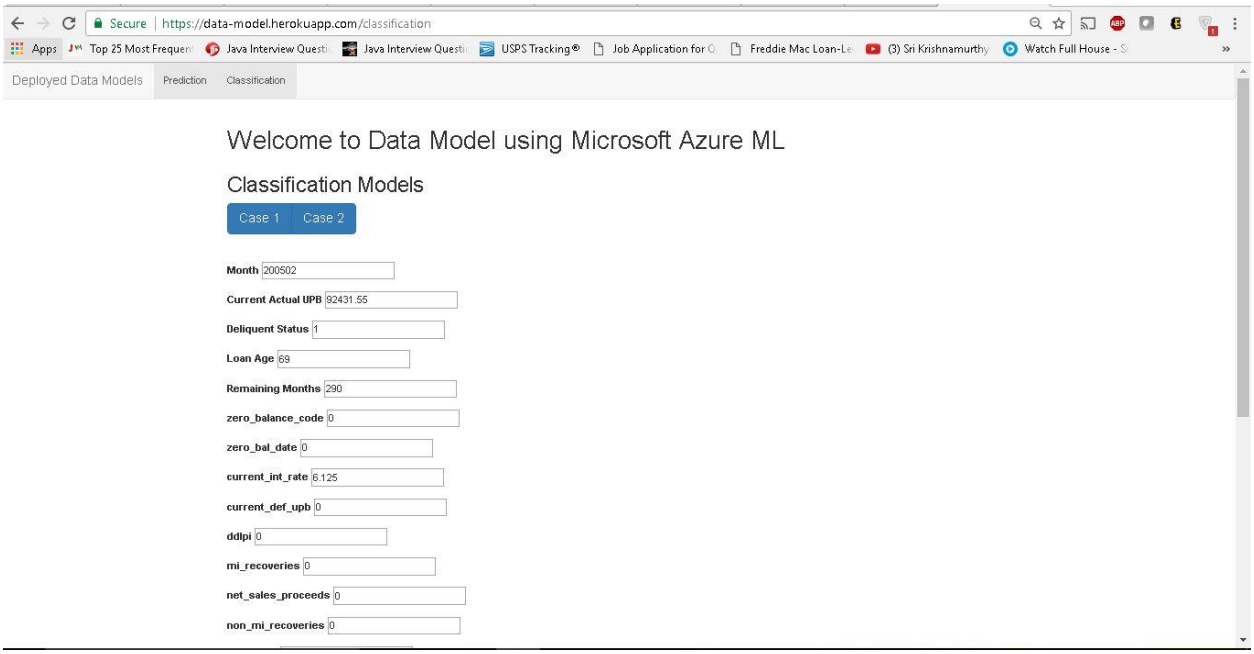


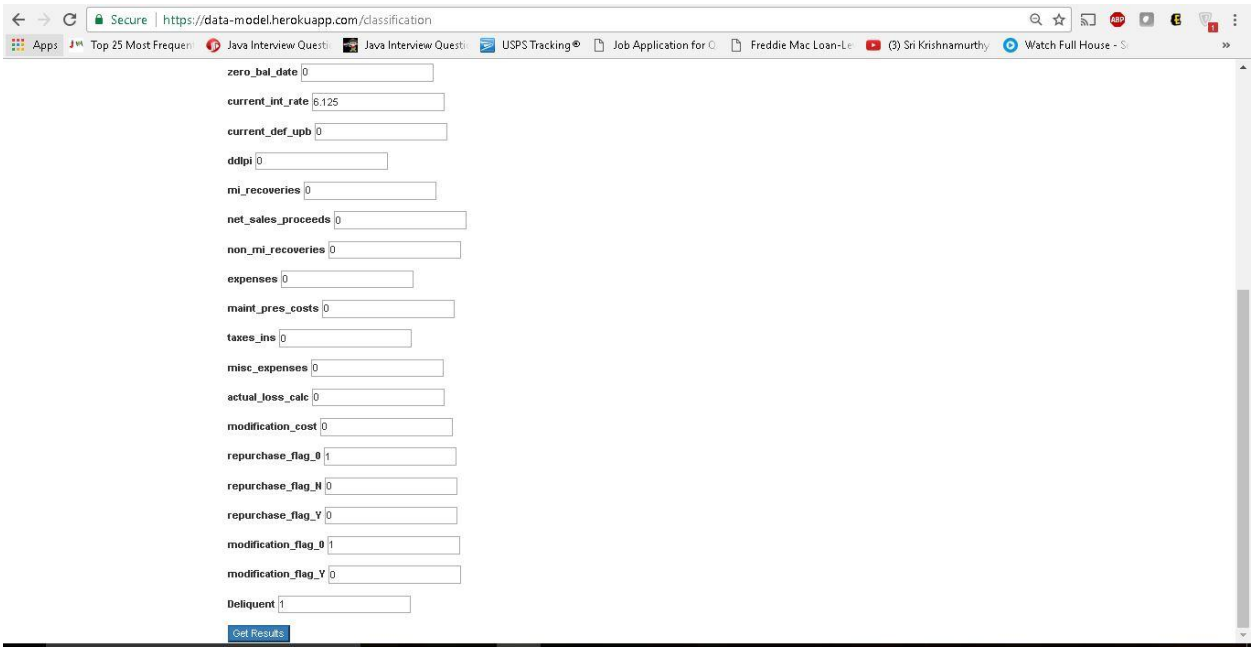
Results: Testcase1



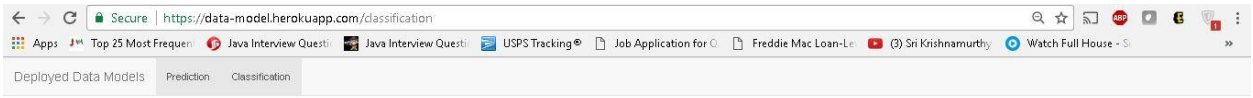
II. Classification Testcase2:

200502	92431.55	1	69	290	0	0	6.125	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1





Results: Testcase2



Welcome to Data Model using Microsoft Azure ML

Current interest rate is: 6.125

Current UPB is: 92431.55

Deliquent is: 1

Logistic Regression:

Scored Probability for Class "0" is: 0.609493017196655

Scored Probability for Class "1" is: 0.390506982803345

Scored Labels is: 0

Decision Forest:

Scored Probabilities is: 0

Scored Labels is: 0

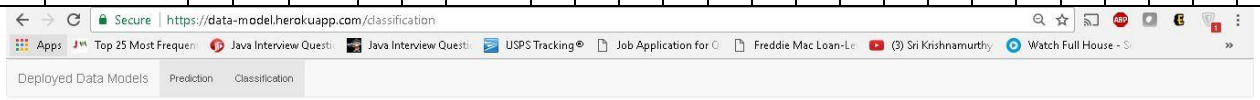
Neural Network:

Scored Probabilities is: 0.99992561340332

Scored Labels is: 1

III. Classification Testcase3:

200309	110672.49	1	51	308	0	0	6.5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1



Welcome to Data Model using Microsoft Azure ML

Current interest rate is: 6.5

Current UPB is: 110672.49

Deliquent is: 1

Logistic Regression:

Scored Probability for Class "0" is: 0.590552508831024

Scored Probability for Class "1" is: 0.409447431564331

Scored Labels is: 0

Decision Forest:

Scored Probabilities is: 0

Scored Labels is: 0

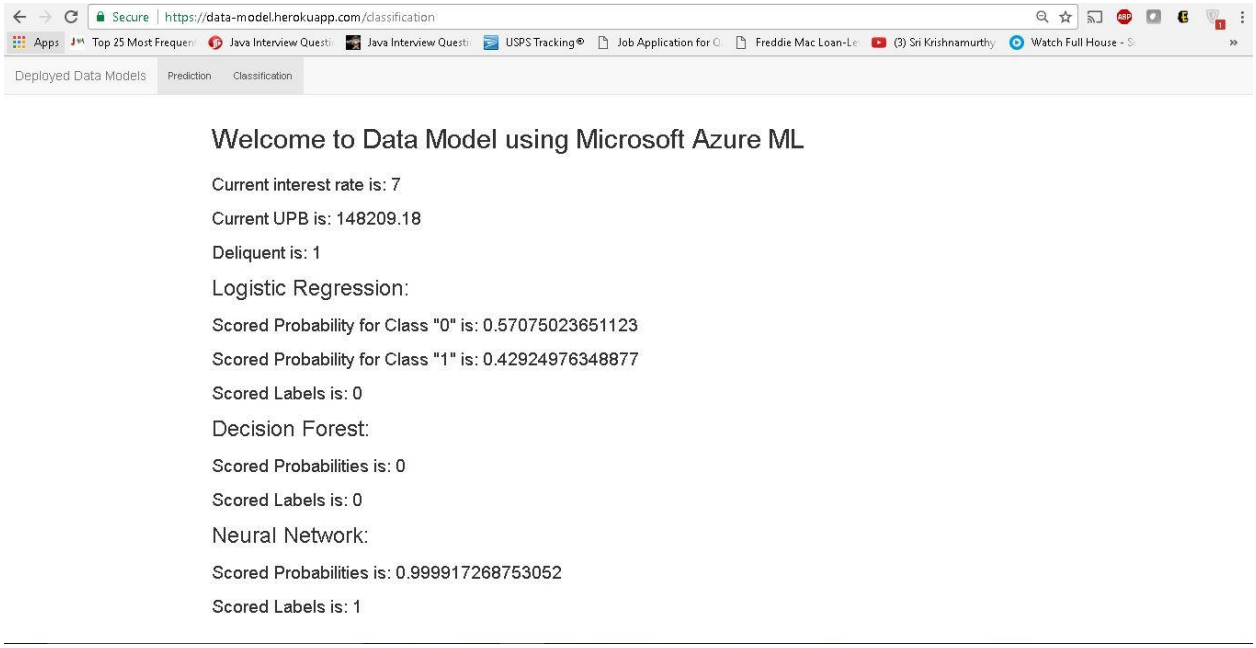
Neural Network:

Scored Probabilities is: 0.999922513961792

Scored Labels is: 1

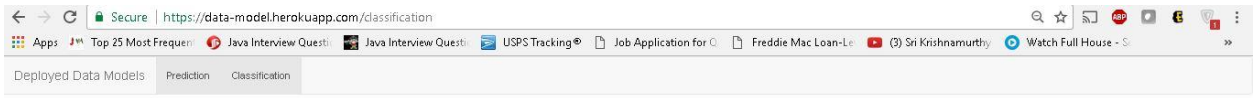
IV. Classification Testcase4:

200110	148209.18	1	29	331	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1



V. Classification Testcase5:

199906	162000	0	0	360	0	0	8.375	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0



Welcome to Data Model using Microsoft Azure ML

Current interest rate is: 8.375

Current UPB is: 162000

Deliquent is: 0

Logistic Regression:

Scored Probability for Class "0" is: 0.986490309238434

Scored Probability for Class "1" is: 0.0135096590965986

Scored Labels is: 0

Decision Forest:

Scored Probabilities is: 0

Scored Labels is: 0

Neural Network:

Scored Probabilities is: 1.51307938267564E-06

Scored Labels is: 0