Deploying Data science models on Azure ML.

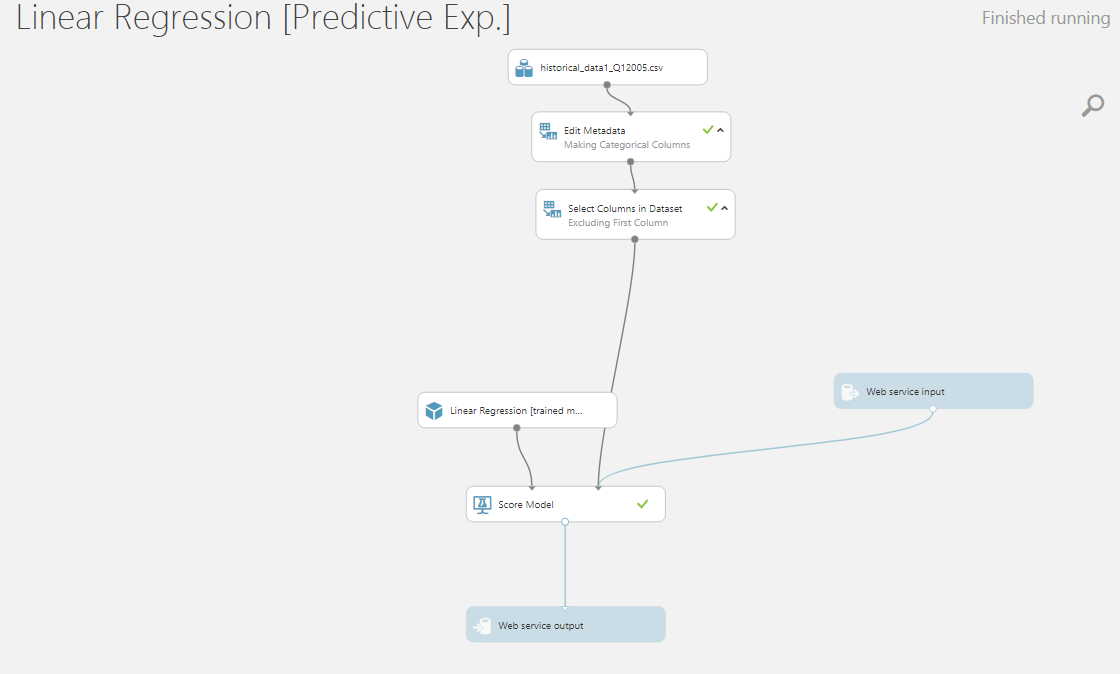
Pruthvij Thakar | Sumedh Saraf

ADVANCES IN DATA SCIENCE

# **PREDICTION MODELS**

1. Linear Regression:

(<http://freddiemaclinear.azurewebsites.net/>)



* The JSON format taken by the website in form of the input parameters to compute the interest rates in prediction part.

"Inputs": {

"input1": {

"ColumnNames": [

"fico",

"dt\_first\_pi",

"flag\_fthb",

"dt\_matr",

"mi\_pct",

"cnt\_units",

"orig\_upb",

"ltv",

"channel",

"loan\_purpose",

"orig\_loan\_term"

],

"Values": [

[

"0",

"3630",

"0",

"7260",

"0",

"0",

"0",

"0",

"1",

"1",

"0"

],

[

"0",

"3630",

"0",

"7260",

"0",

"0",

"0",

"0",

"1",

"1",

"0"

]

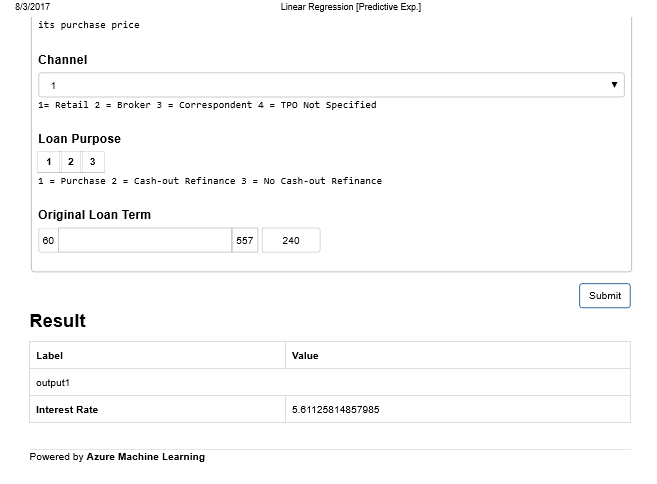
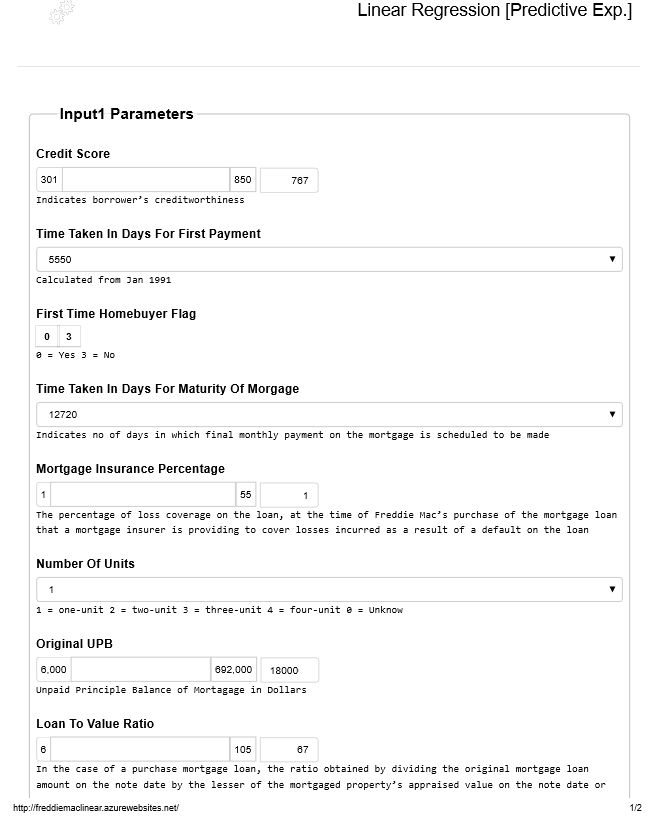
]

}

},

"GlobalParameters": {}

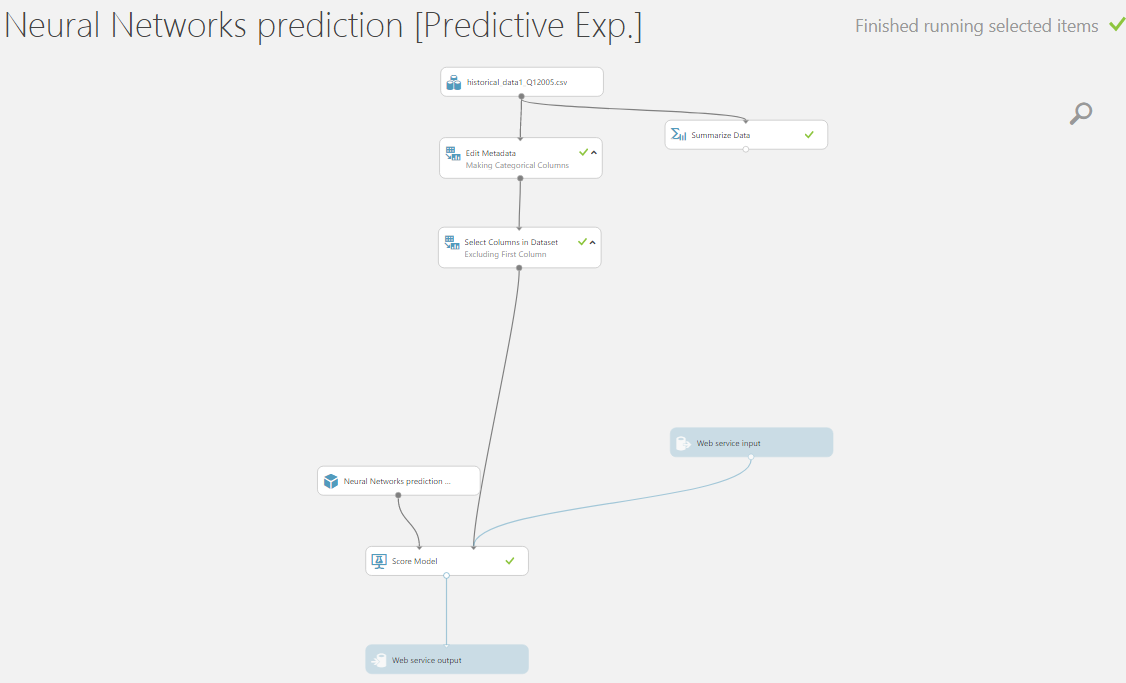
}

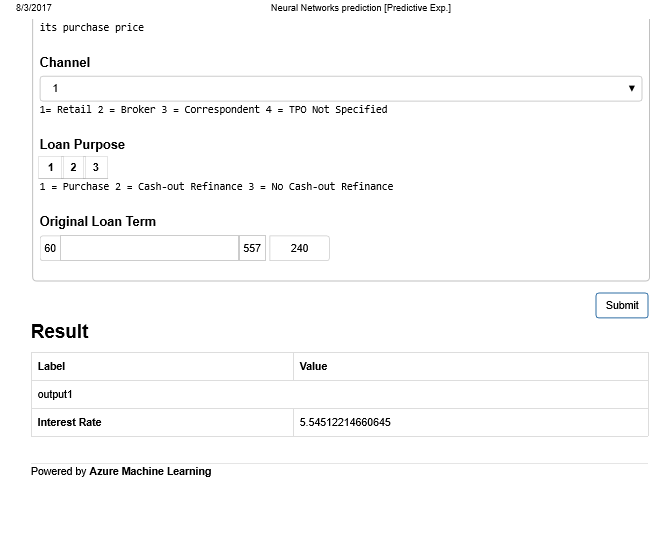
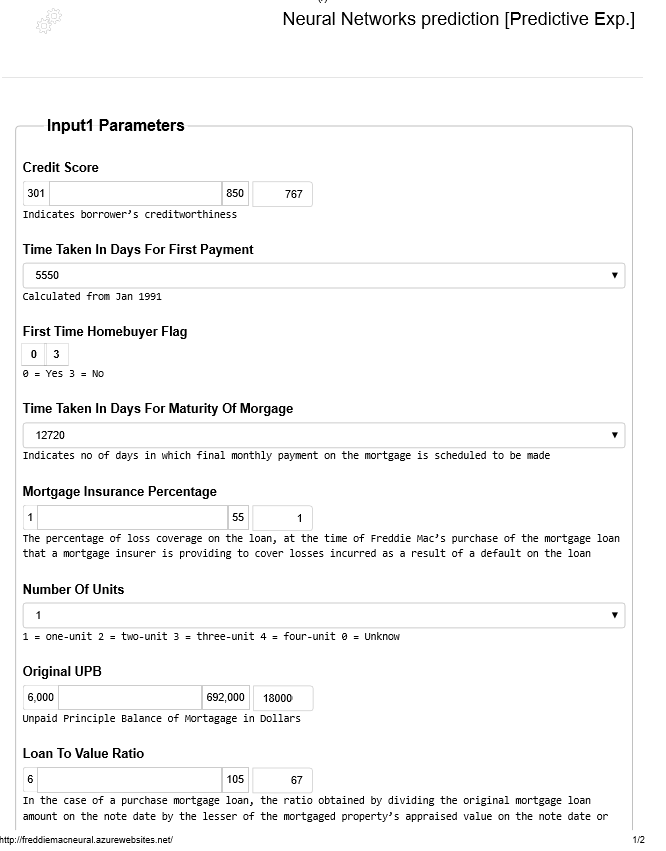


* The above specified images showcase the LINEAR REGRESSION model for predicting the interest rates after taking certain parameters shown in image 2 of the website where we have entered a test case from the data and tried to check the predicted interest rate which turns out to be 5.61% which should ideally be 5.35% so this algorithm doesn’t look to be that effective compared to others.

1. Neural Networks:

(<http://freddiemacneural.azurewebsites.net/>)

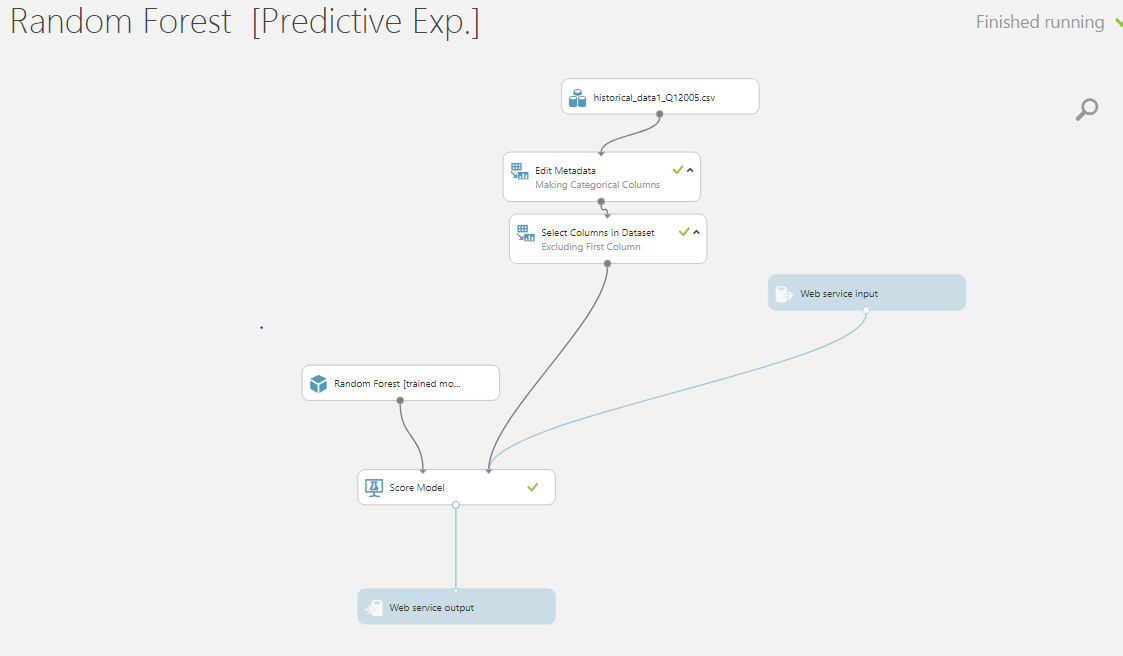


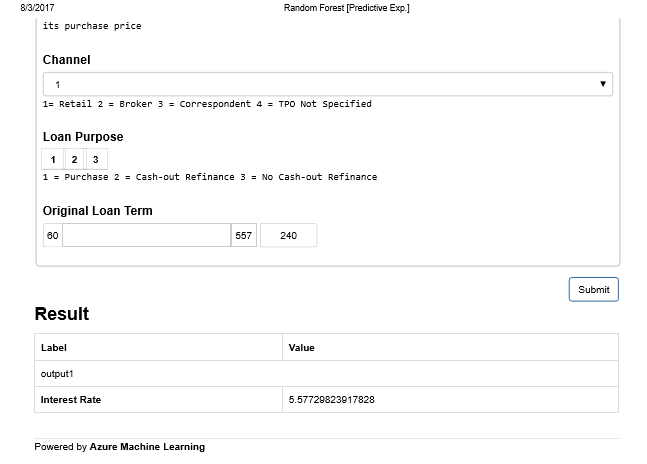
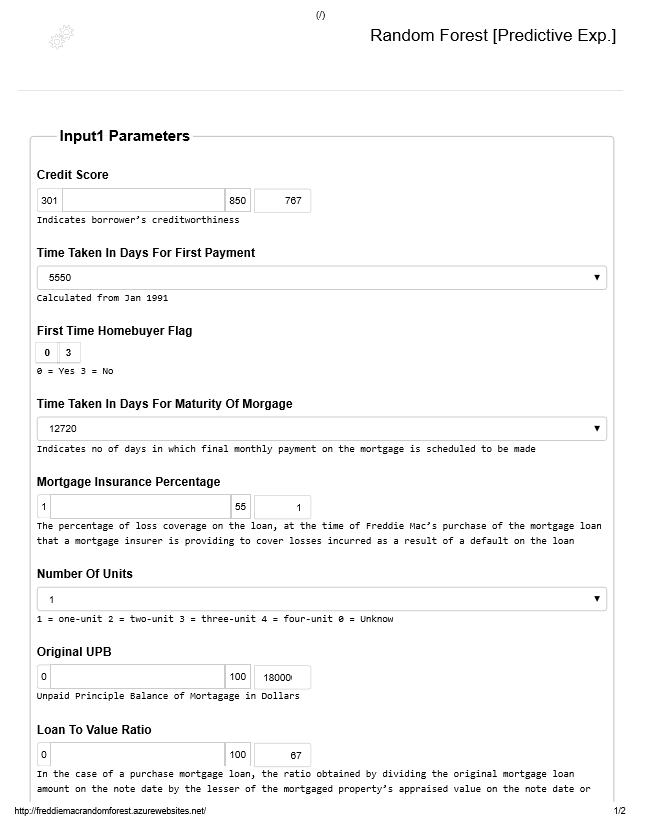


* The above specified images showcase the NEURAL NETWORK model for predicting the interest rates after taking certain parameters shown in image 2 of the website where we have entered a test case from the data and tried to check the predicted interest rate which turns out to be 5.54% which should ideally be 5.35% so this model is better in comparison to the linear regression model.

1. Random Forest:

(<http://freddiemacrandomforest.azurewebsites.net/>)



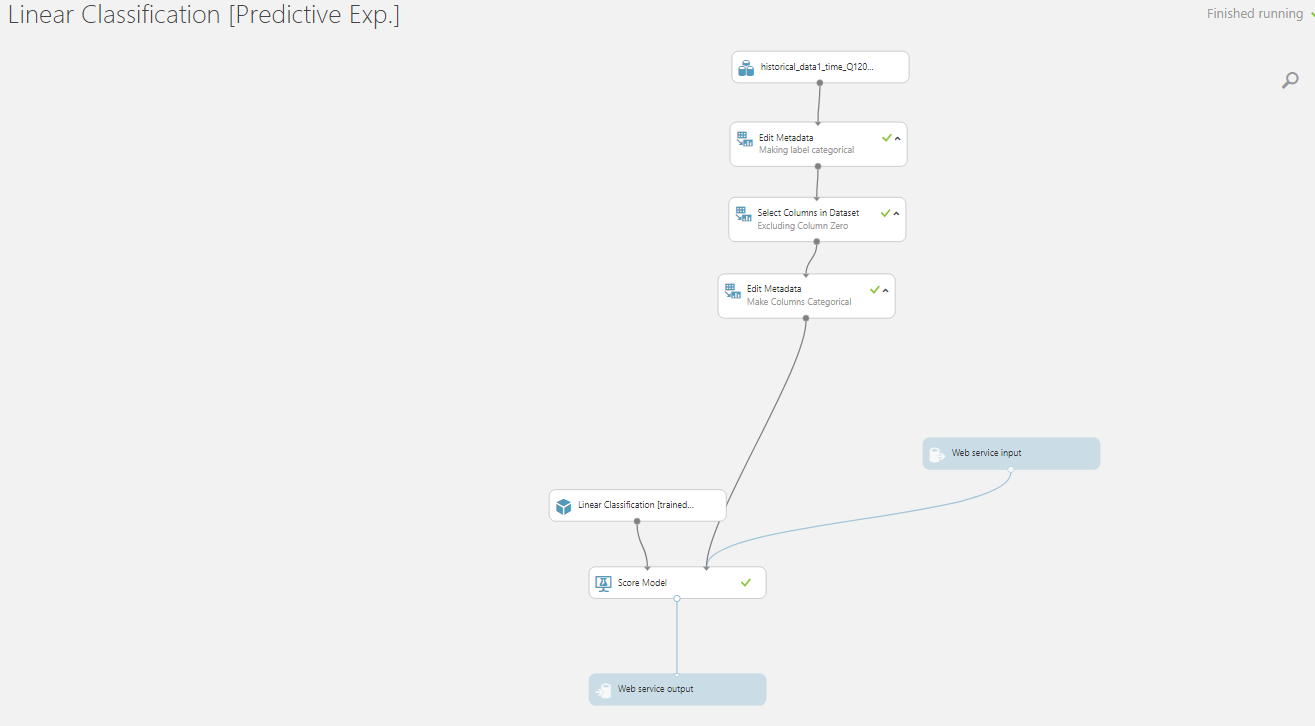


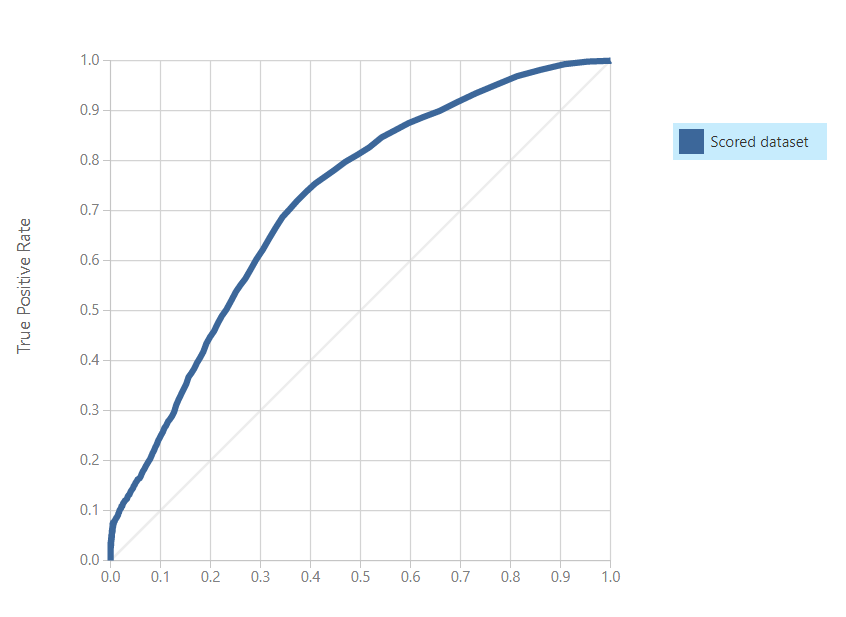
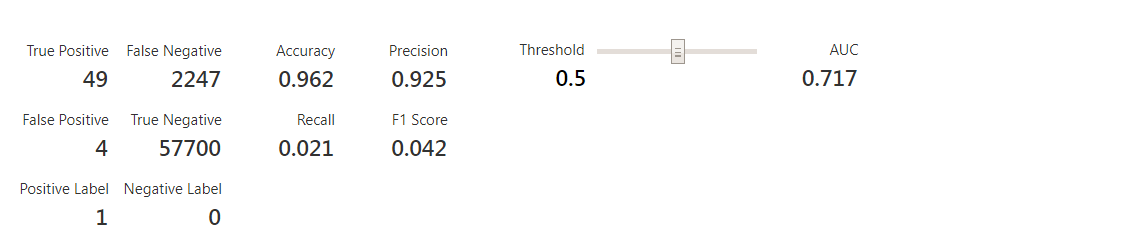
* The above specified images showcase the RANDOM FOREST model for predicting the interest rates after taking certain parameters shown in image 2 of the website where we have entered a test case from the data and tried to check the predicted interest rate which turns out to be 5.57% which should ideally be 5.35% so this model is better in comparison to the linear regression model but Neural Network outperforms both to stand out in terms of prediction.
* The test case values can be derived from the image which users can replicate to check the working of model and can tweak the values to find out the changes observed in output interest rates predicted by the respective models.

# **CLASSIFICATION MODELS**

1. Logistic Regression:

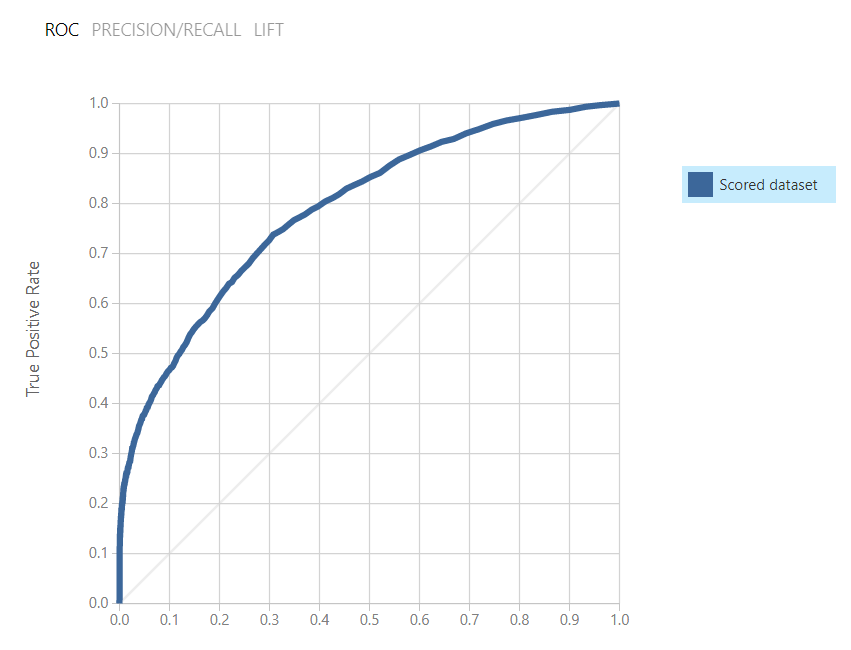
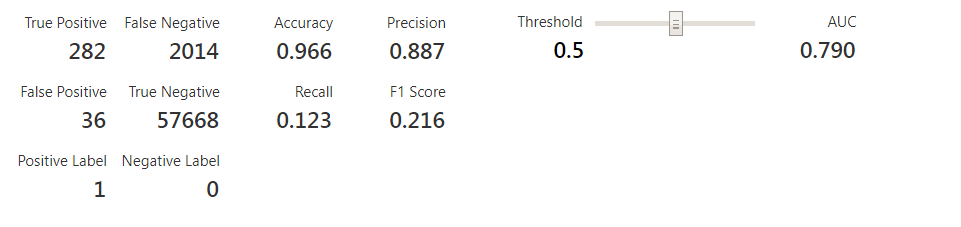
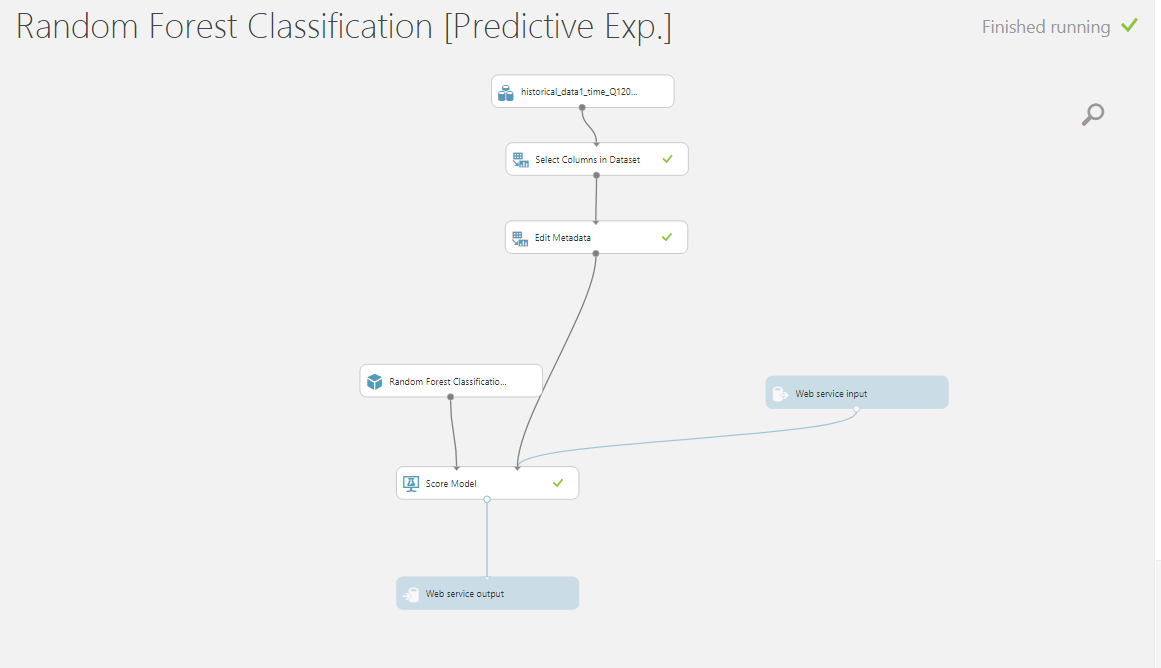
(<http://sumedhpart1.azurewebsites.net/>)





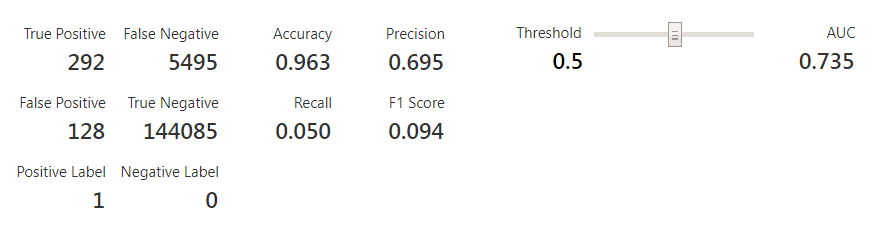
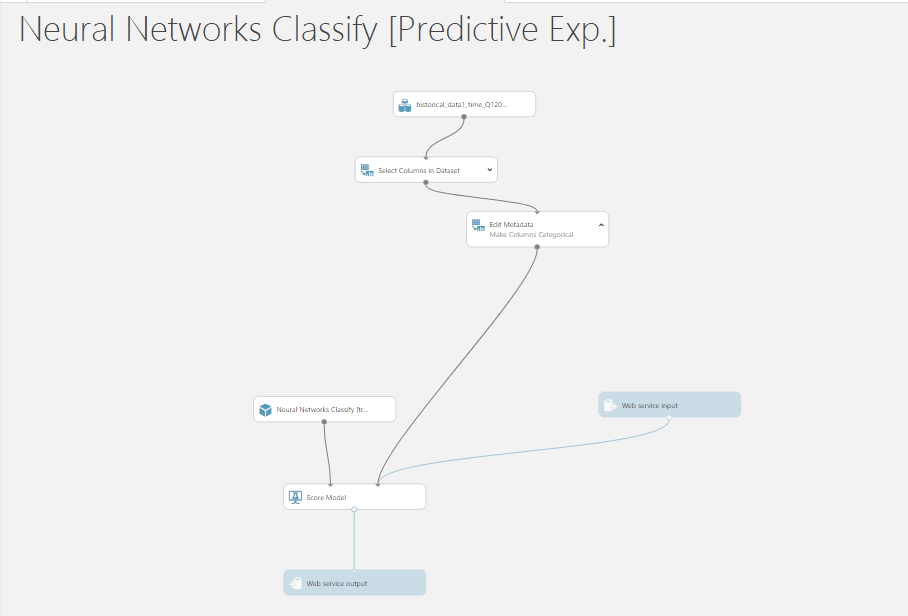
1. Random Forest:

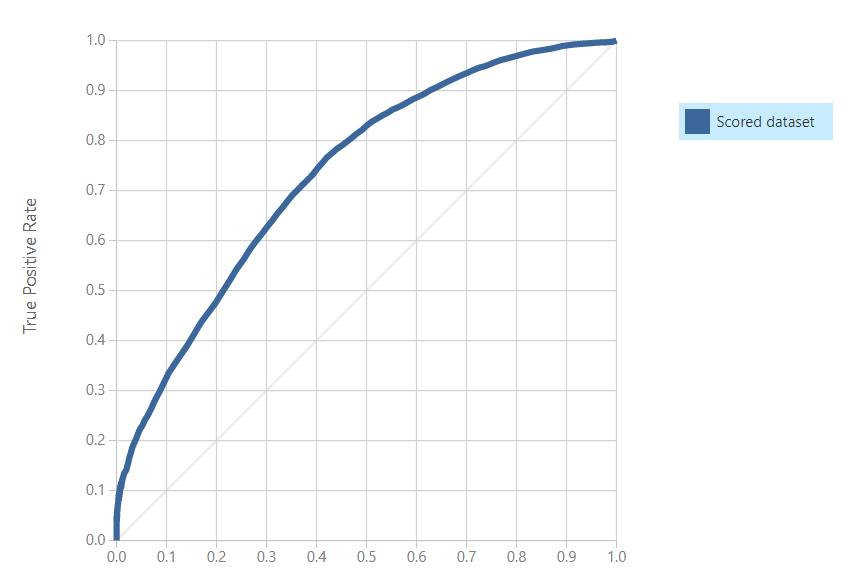
(<http://sumedhpart3.azurewebsites.net/>)



1. Neural Network:

([http://sumedhpart5.azurewebsites.net](http://sumedhpart5.azurewebsites.net/))





* + - If we compare all the 3 Algorithms in terms of numbers and graphs generated, Logistic Regression is statically the best performing algorithm compared to neural Network and Random Forest. It has highest accuracy, precision and the best ROC curve amongst all three. Its ***precision*** is 0.925 while its ***Accuracy*** is 0.962 which outperforms all others.
* 7200 228046.71 58 302 2 0 0 0 5.625 0 1
* svcg\_cycle current\_upb loan\_age mths\_remng repch\_flag flag\_mod cd\_zero\_bal dt\_zero\_bal current\_int\_rt non\_int\_brng\_upb new\_delinq
  + - The above given values can be used as the test cases for the Classification.