**Assignment 3**

**(Advances in Data Science)**

**TEAM 6  
- Priyal Chaudhari  
- Prashant Vikram Singh**

**FreddieMac Single Family Loan-Level Dataset**

**Process Flow:**

**Part 1**

Perform EDA

Scrap Sample Data

Create Summary Files

User Login

**Part 2**

Apply Classification

Algorithms

Combine Historical Files

User Login

Scrap Historical Data

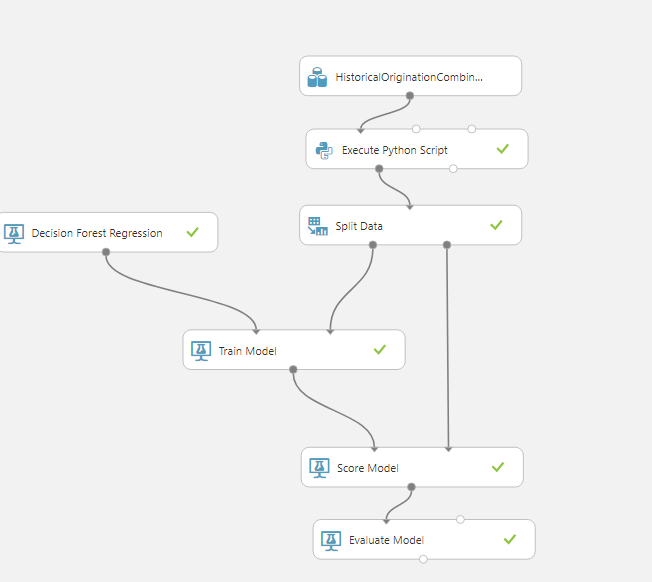
Deployed on Azure Portal

Build algorithm on Azure ML

Apply Regression Algorithms

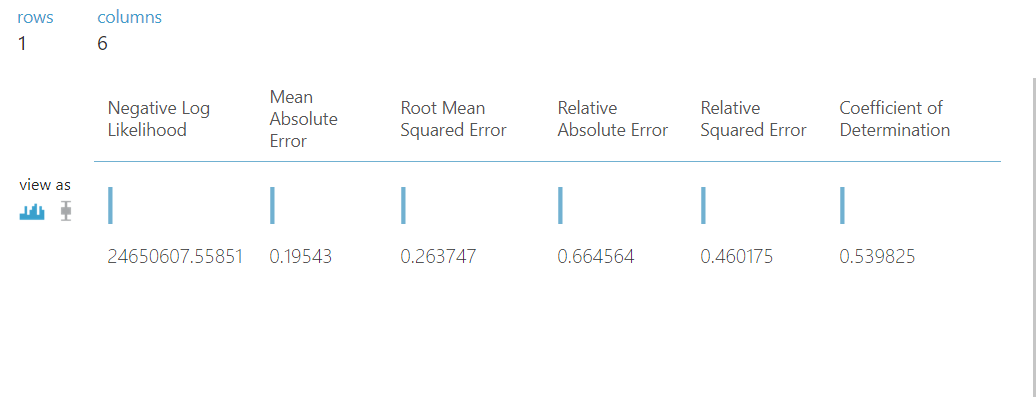
**Random Forest**

**Training Model**



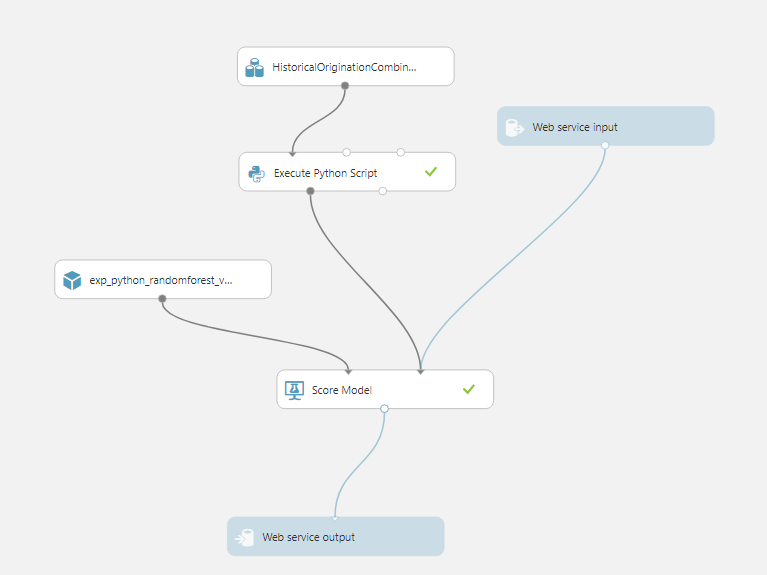
* Random Forest Model for predicting interest rates
* We have trained the model on 3-7 years of data, where we have tried to cover economic boom and recession period
* For training model, we are selecting numeric data and passing that to splits the data module in azure ML
* The split data module splits the data into 70:30 ratio for training and testing purpose
* After training is completed, we evaluate the model based on various parameters which we receive from Evaluation Model module

Training Model Visualization



* The MAE value for random forest is lowest, i.e. 0.19
* Parameter like RMSE is a bit high but other parameters are in proper range
* Hence, Random Forest perform better compared to other algorithms

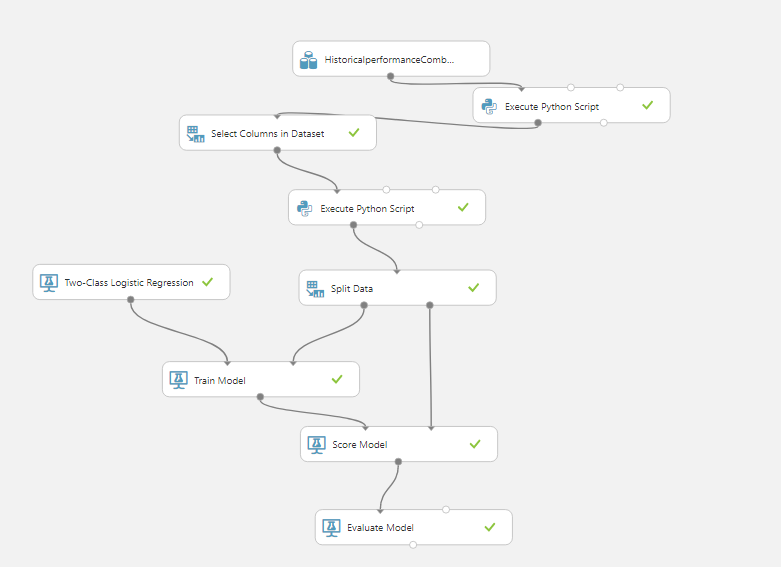
**Predictive Model**



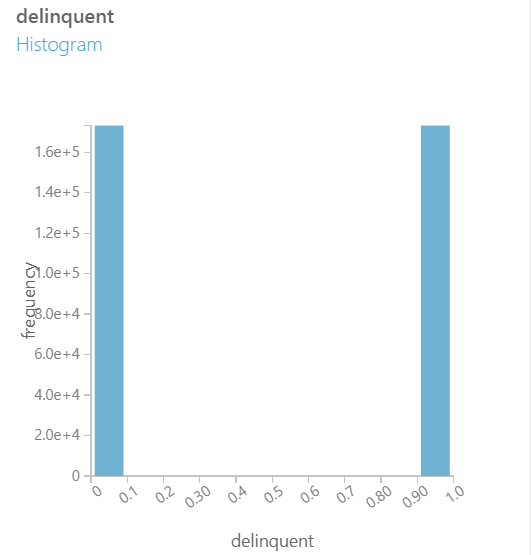
* In above model, we are deploying the trained model on azure portal

**Two Class Logistic Regression**

**Training Model**

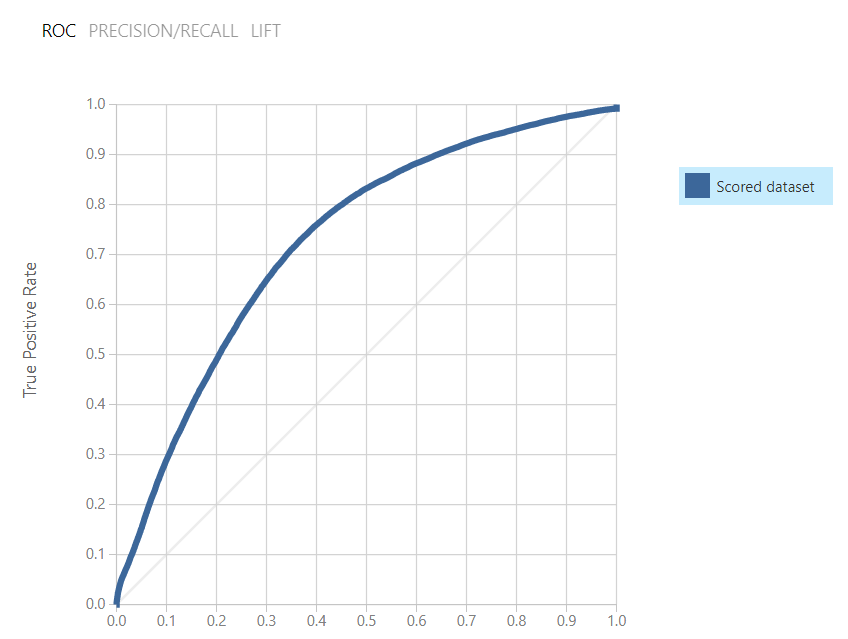


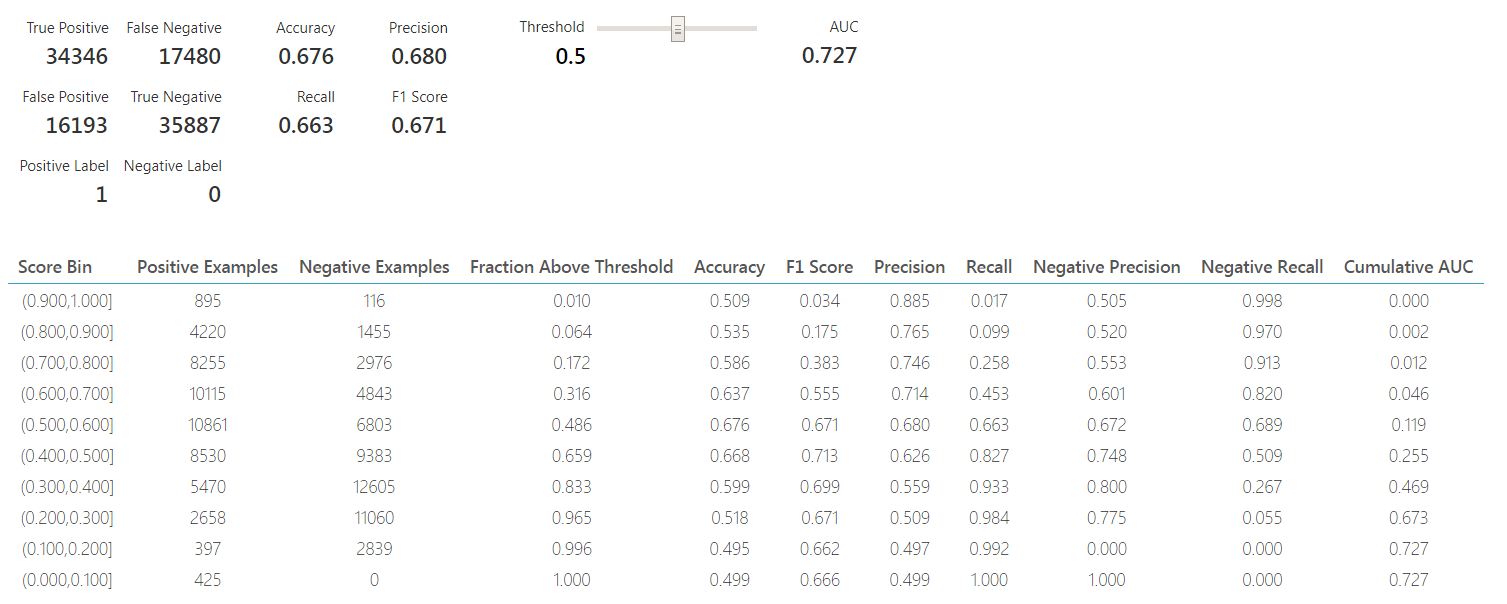
* Logistic Regression Model for classifying delinquents
* We have trained the model on 3-7 years of data, where we have tried to cover economic boom and recession period
* For training model, we are selecting numeric data and passing that to splits the data module in azure ML
* The split data module splits the data into 70:30 ratio for training and testing purpose
* Here, data is biased, so we balanced the data in delinquent column using python script module



* As evident from the diagram, zeros and ones are evenly balance
* After training is completed, we evaluate the model based on various parameters which we receive from Evaluation Model module

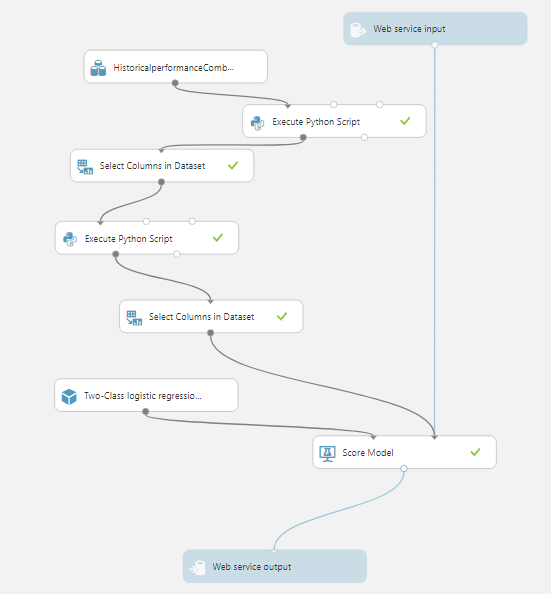
Training Model Visualization





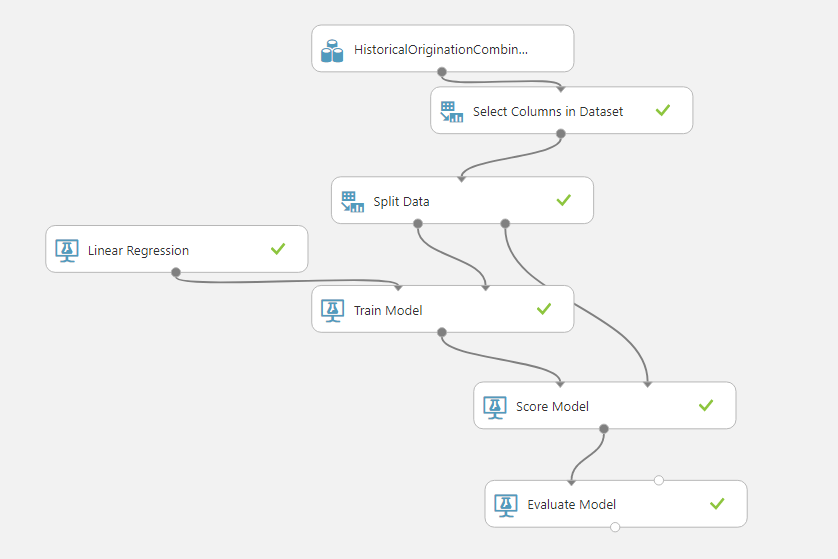
* For evaluation of classification model, as we can see from the above figure, we are getting a smooth ROC curve.
* That is because we performed balancing of data
* Accuracy of this logistic model is 0.72
* It’s definitely not the best model as per performance because, Random Forest has a better ROC curve
* If speed is of utmost importance then, we can go for this model

**Predictive Model**

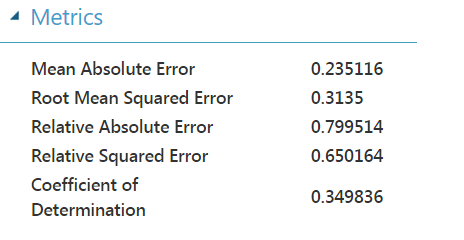


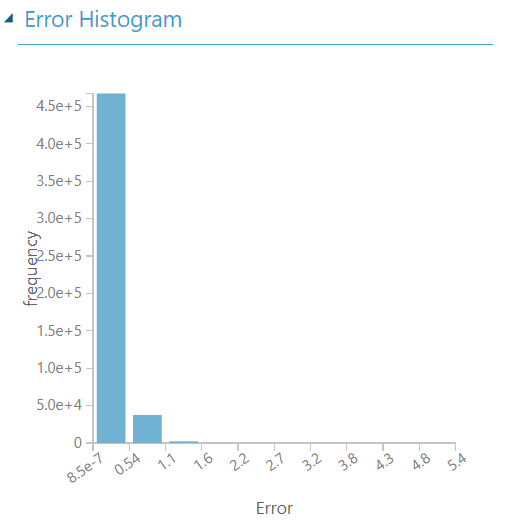
**Linear Regression**

**Training Model**

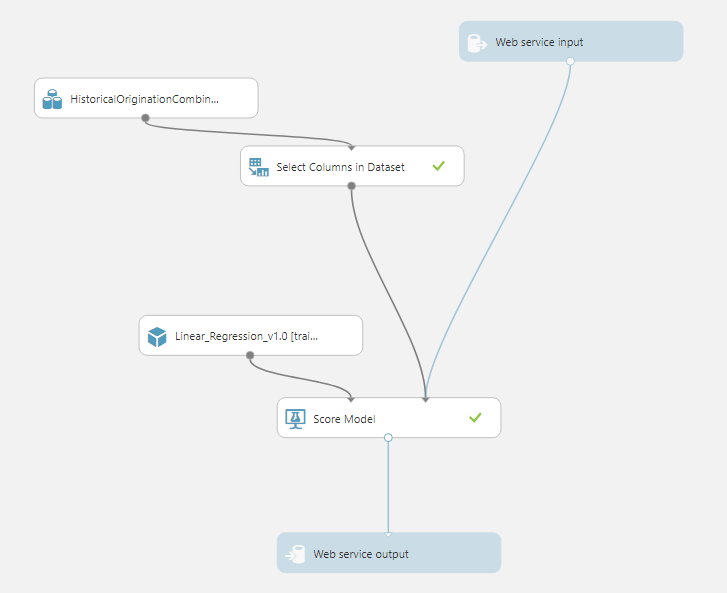


Training Model Visualization



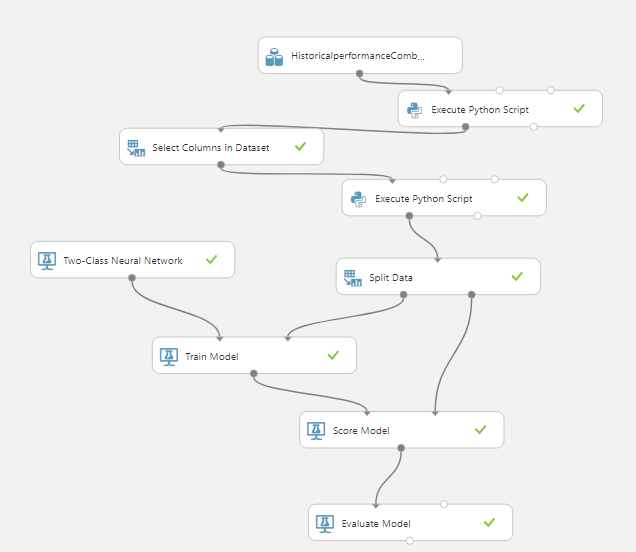


**Predictive Model**

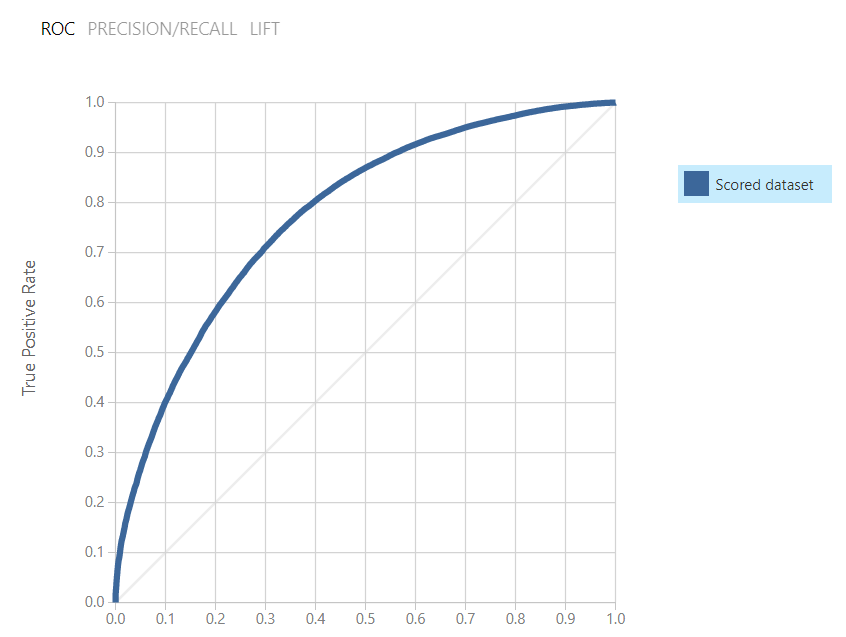


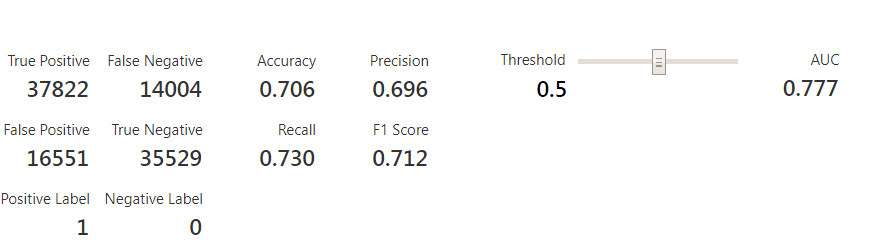
**Two Class Neural Network**

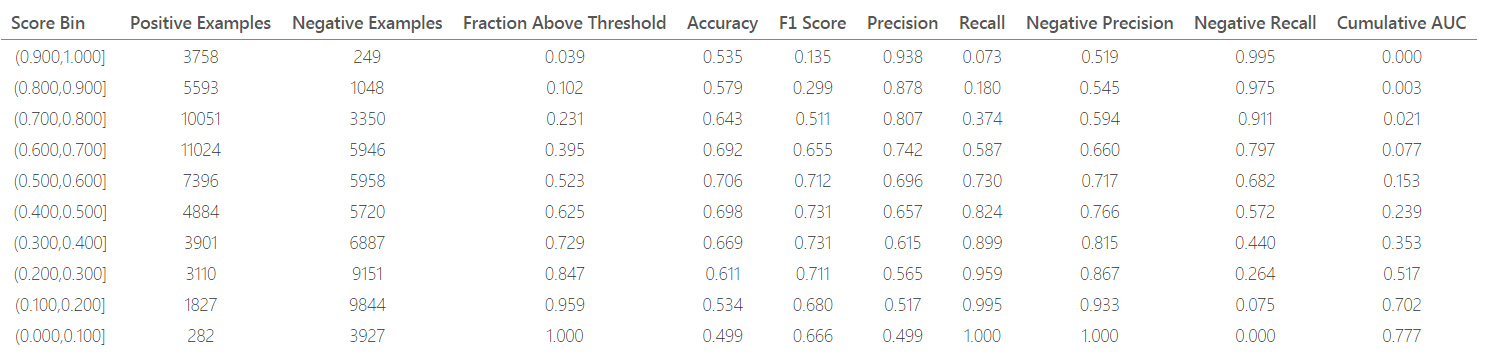
**Training Model**



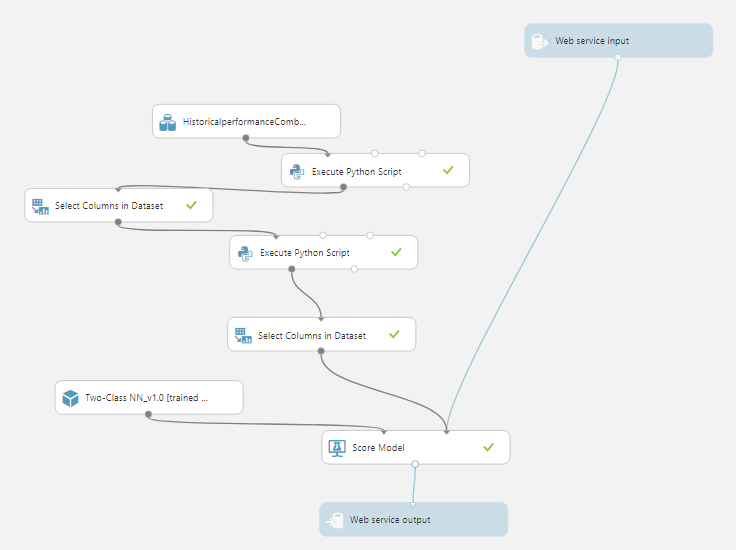
Training Model Visualization





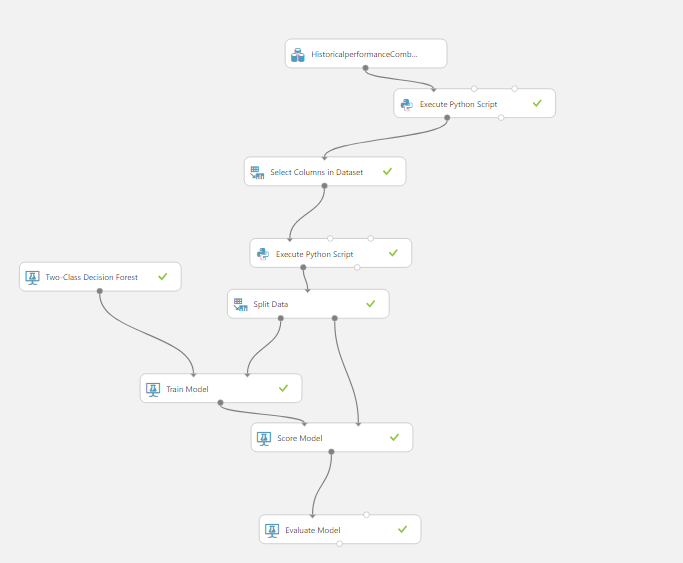


**Predictive Model**

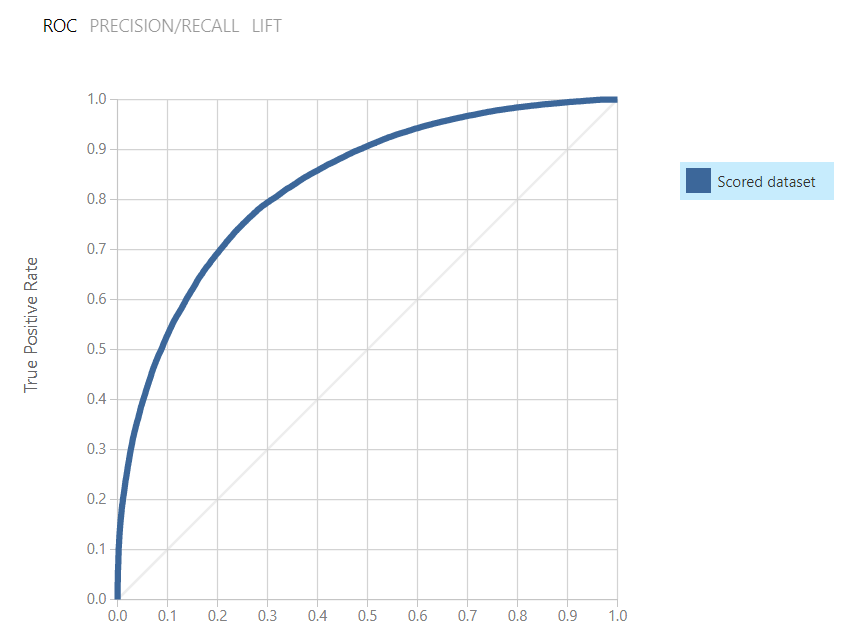


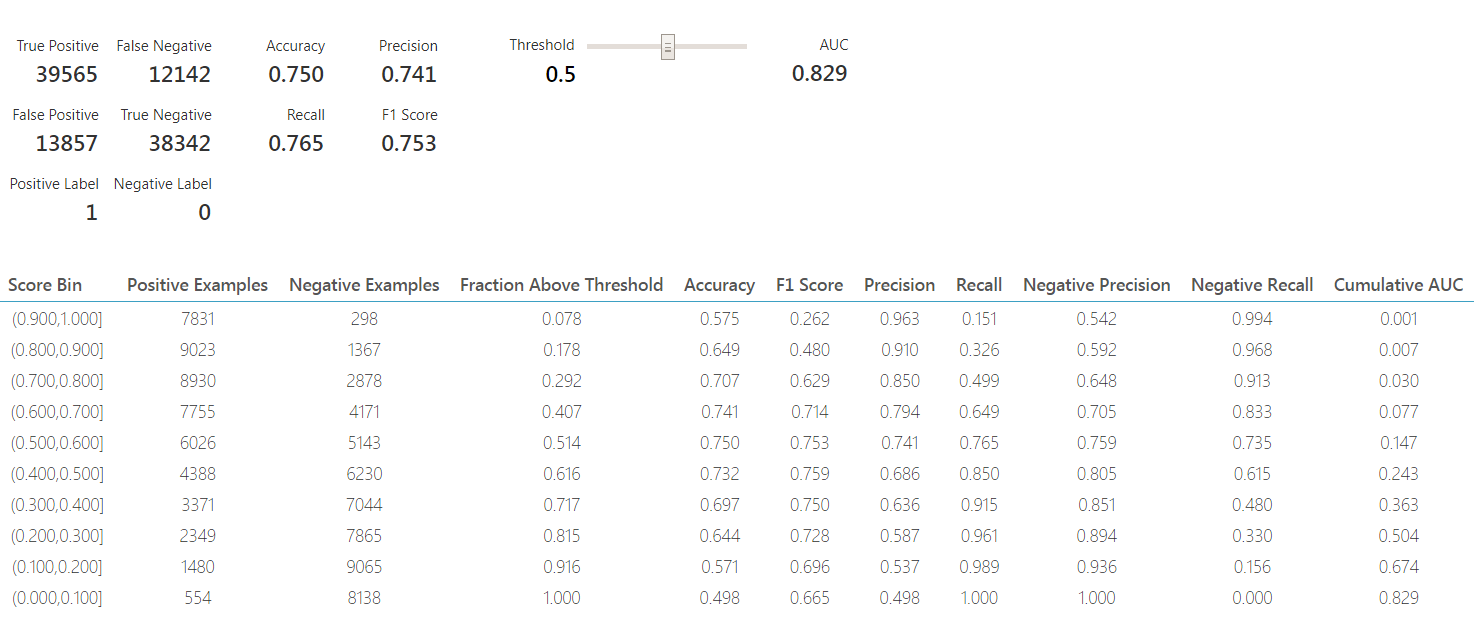
**Two Class Decision Forest**

**Training Model**

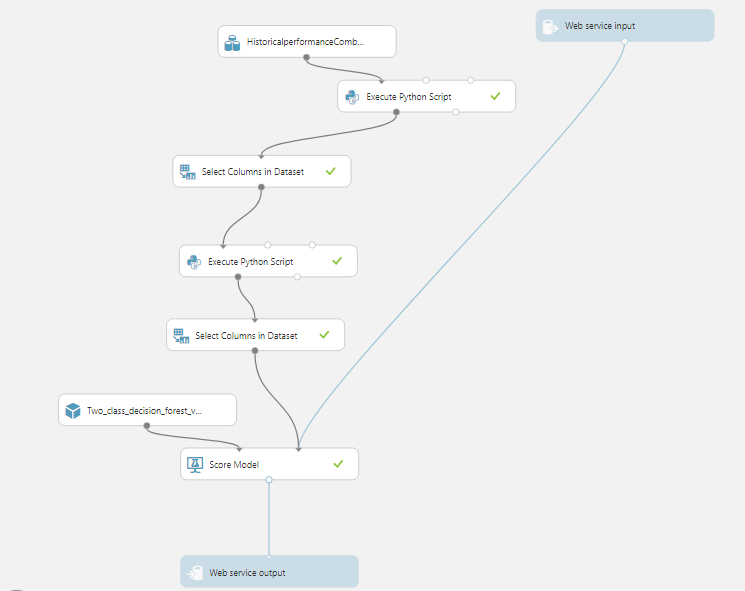


Training Model Visualization



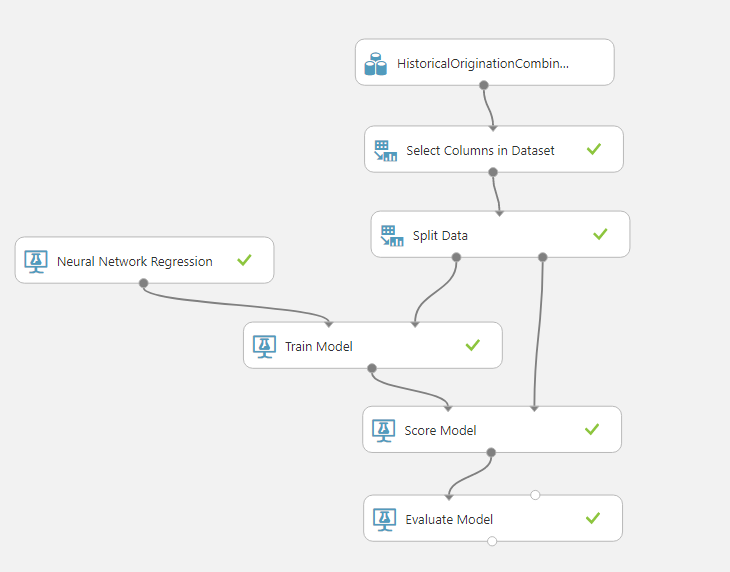


**Predictive Model**

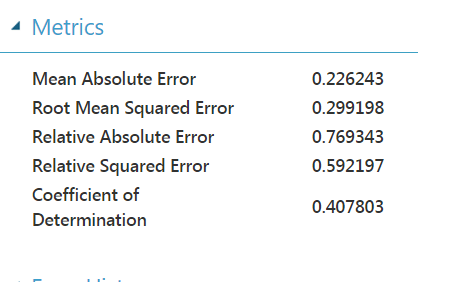


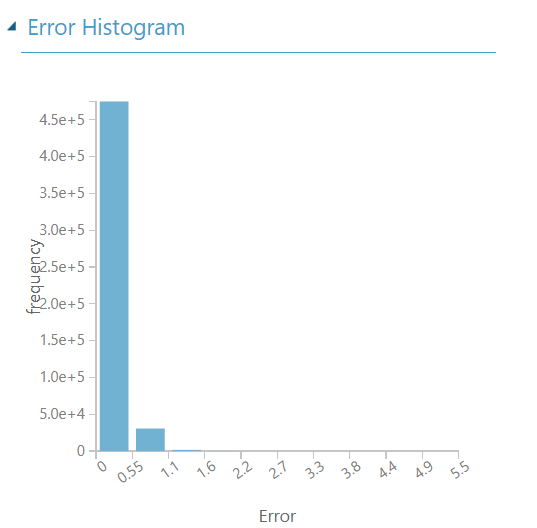
**Neural Network**

**Training Model**

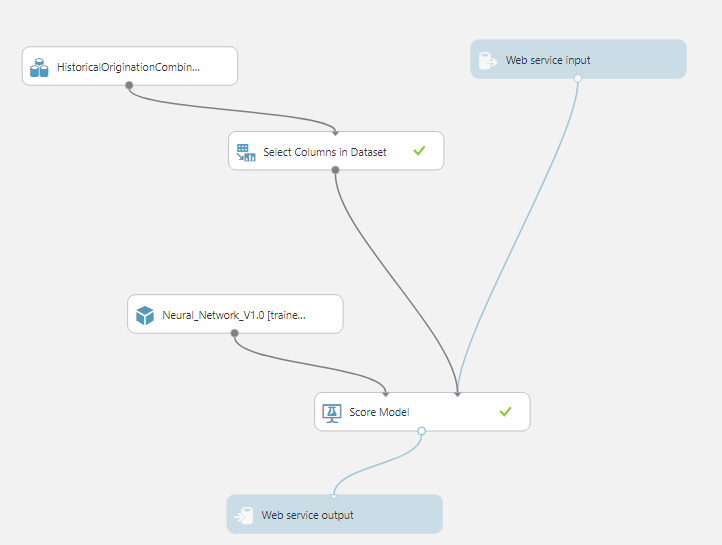


Training Model Visualization





**Predictive Model**



**Rest API Links**

**Random Forest Prediction**

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

**Neural Network Prediction**

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

**Linear Regression Prediction**

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

**Two Class Decision Forest**

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

**Two Class Logistic Regression**

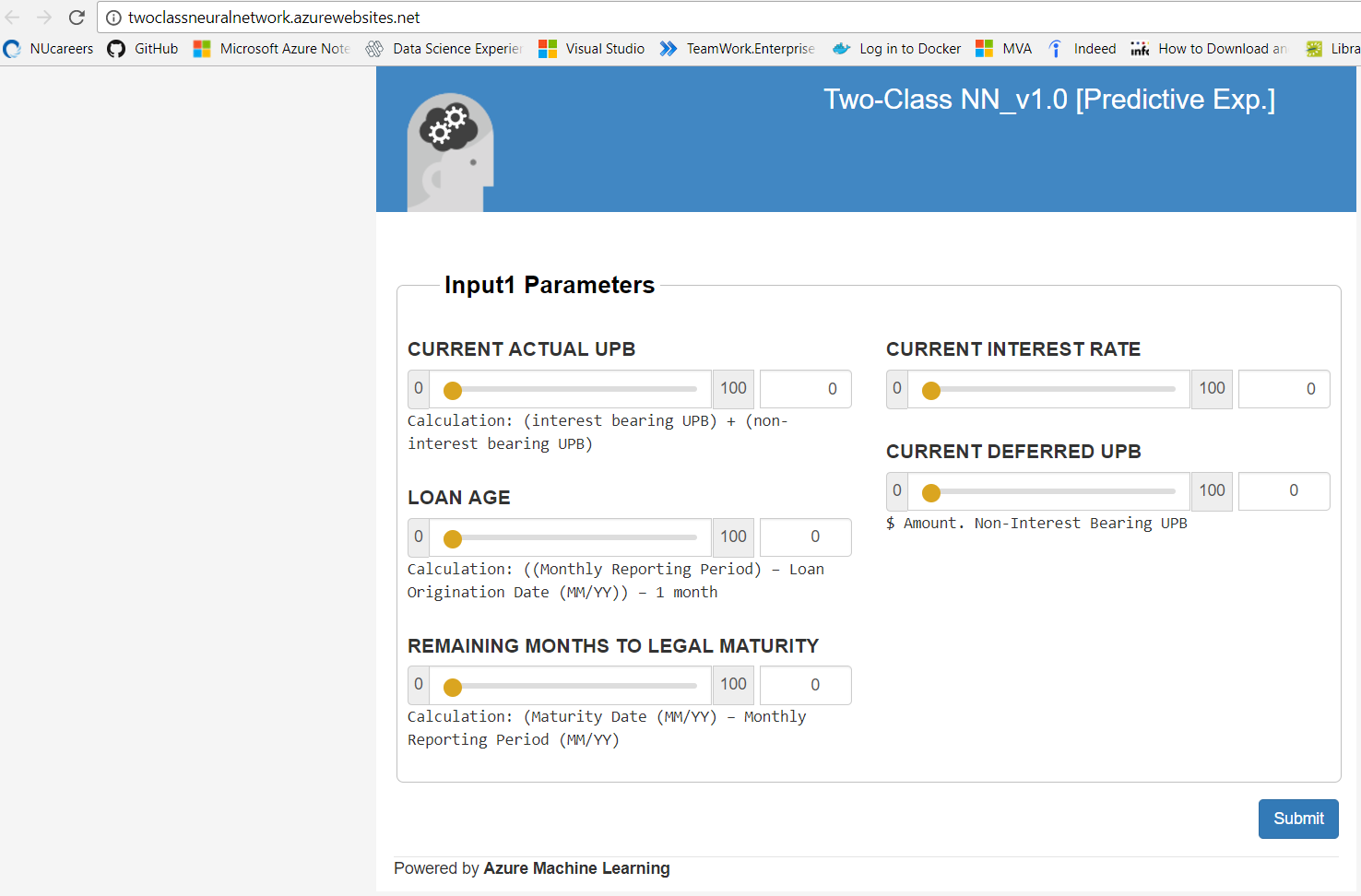
[http://twoclasslogesticregression.azurewebsites.net](http://twoclasslogesticregression.azurewebsites.net/)

**Two Class Neural Network**

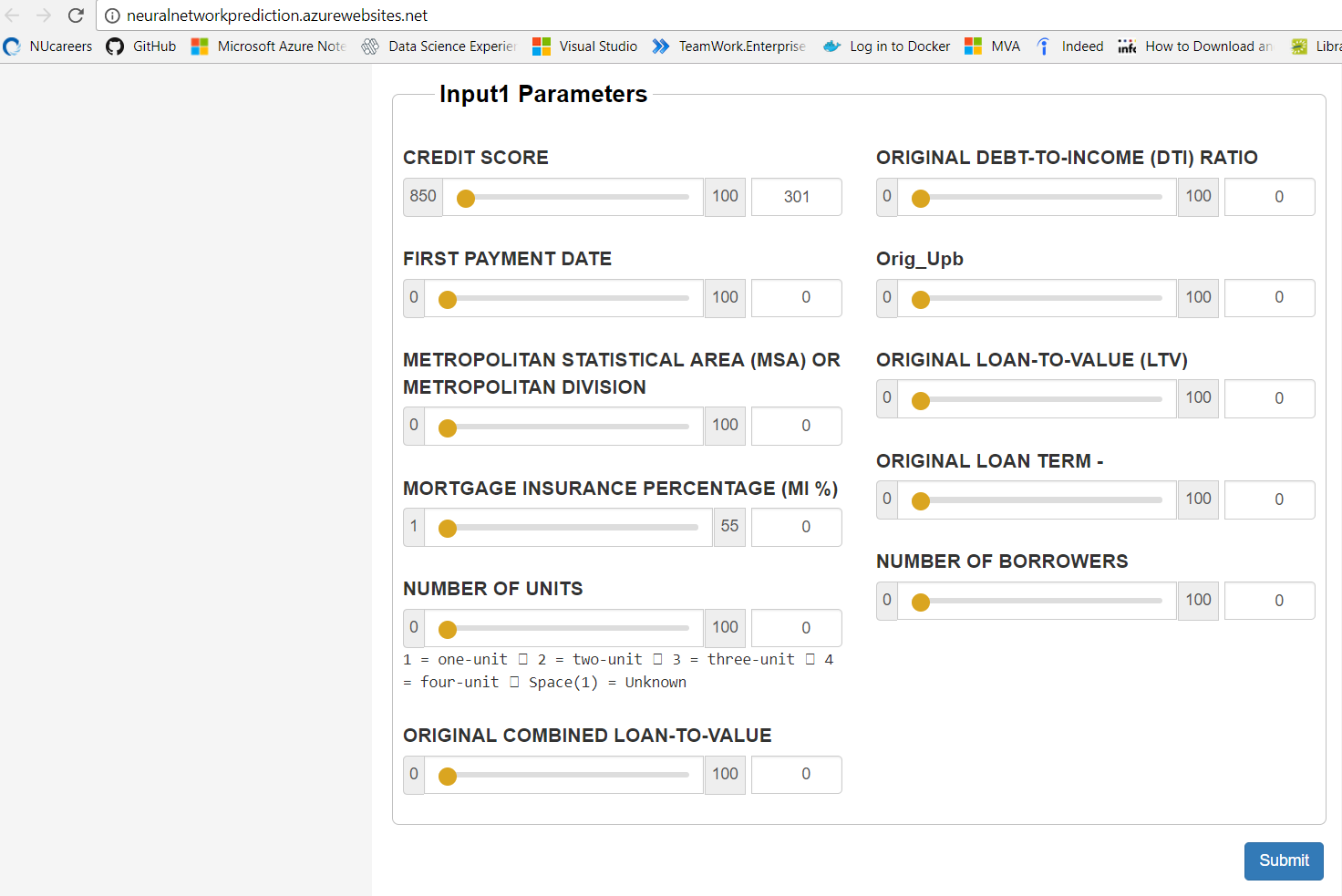
[http://twoclassneuralnetwork.azurewebsites.net](http://twoclassneuralnetwork.azurewebsites.net/)

**Rest API UI Screenshot**

**Classification UI**



**Prediction UI**



**Sample Result for Classification**

