Microsoft

Boosted Decision Trees with Azure ML

Contents

Table of Contents

Goals and Requirements

Estimated time to complete lab is 10-15 minutes.

Goals

1. Build a Predictive Model to understand if the customer will buy the product or not.
2. Develop the model using Decision Trees in AML.

Requirements

1. Access to an Azure Machine Learning Subscription
2. Access to Dataset provided in session

TWO CLASS BOOSTED DECISION TREE

Project Expectation

Bank Telemarketing

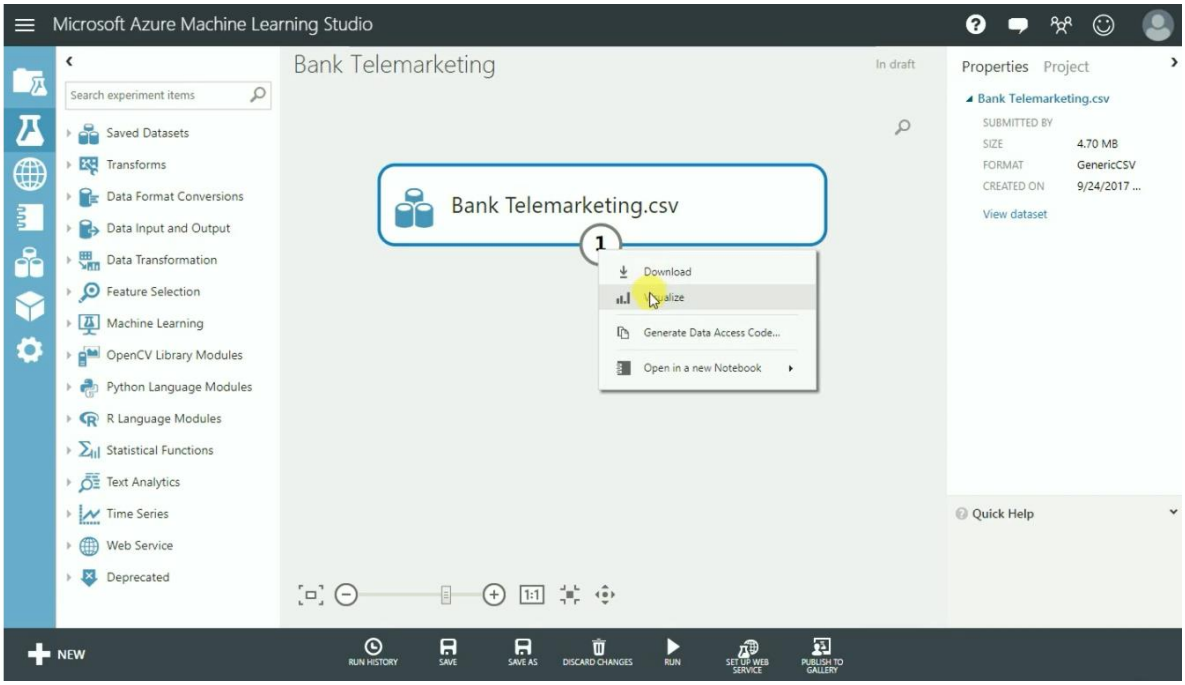
- Goal is to predict if the client will subscribe to a product or not

- Number of instances – 45, 211

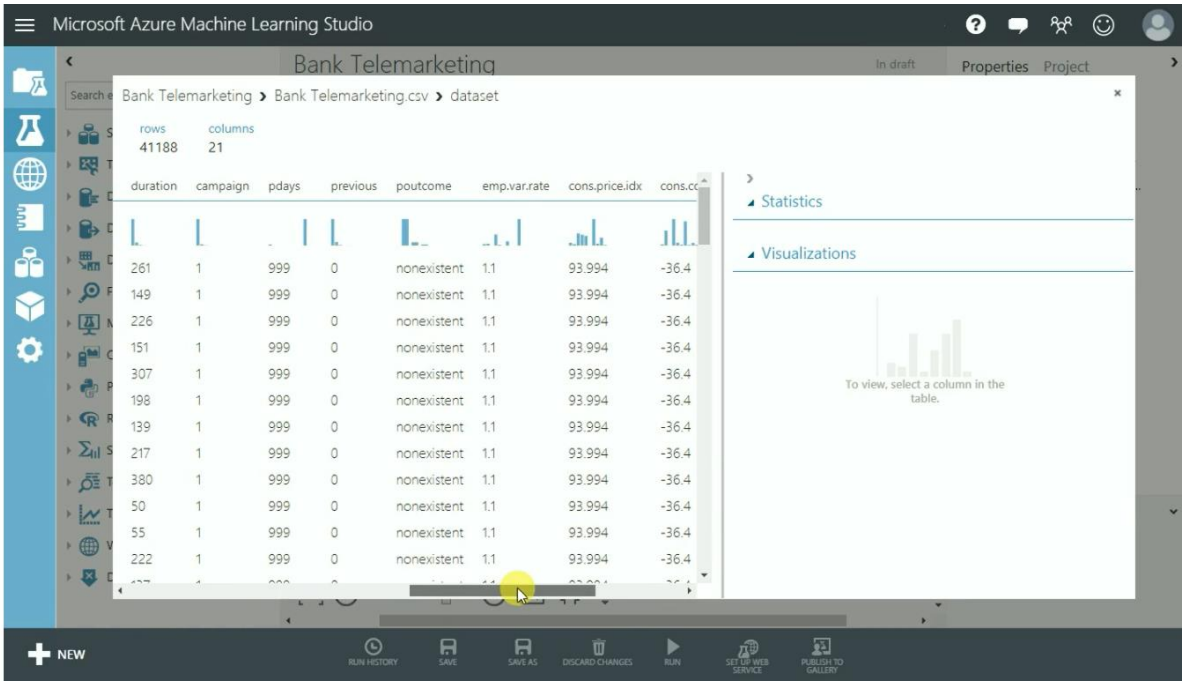
1. Age
2. Job Type
3. Marital Status
4. Education Level
5. Credit Default?
6. Housing Loan?
7. Personal Loan
8. Contacted Type
9. Contacted Month
10. Last Contacted day
11. Contact Duration
12. Campaign Type
13. P-Days
14. Previous
15. P-Outcome
16. Emp-Var-Rate
17. Consumer Price Index
18. Consumer Confidence Index
19. Euribor 3 Month Rate
20. Number of employees
21. Subscribed?

Dataset

Place the uploaded dataset in canvas and visualize the same



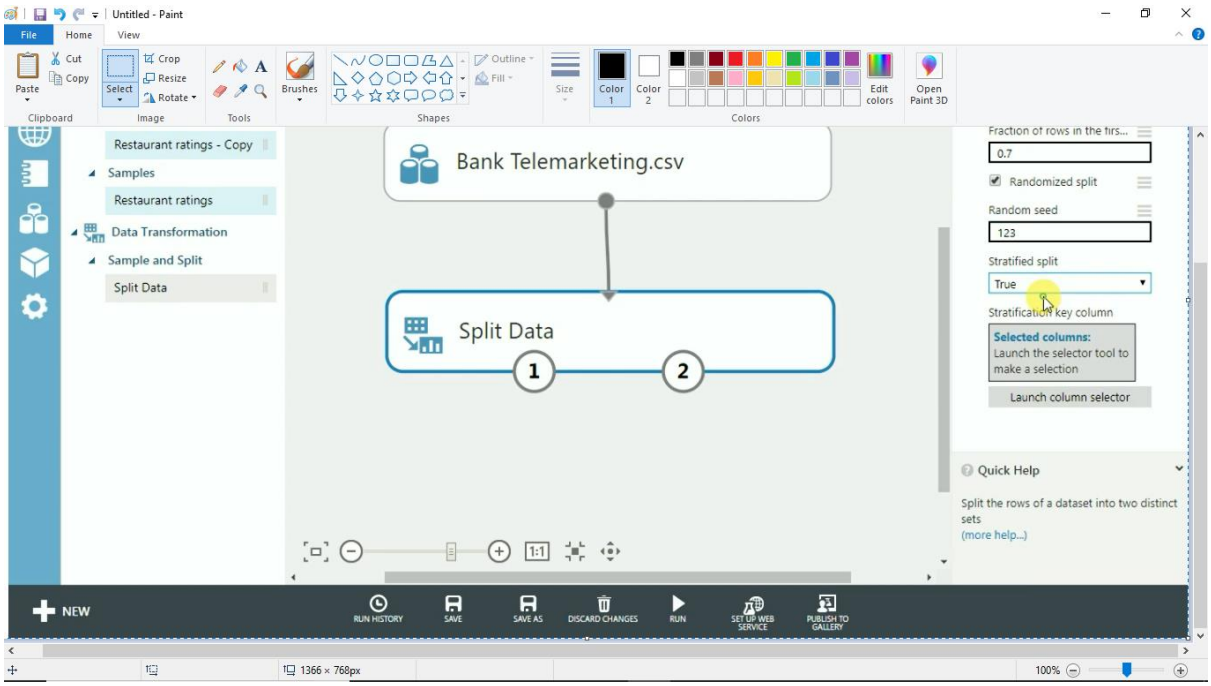
Visualize and analyse the data for missing values and feature type



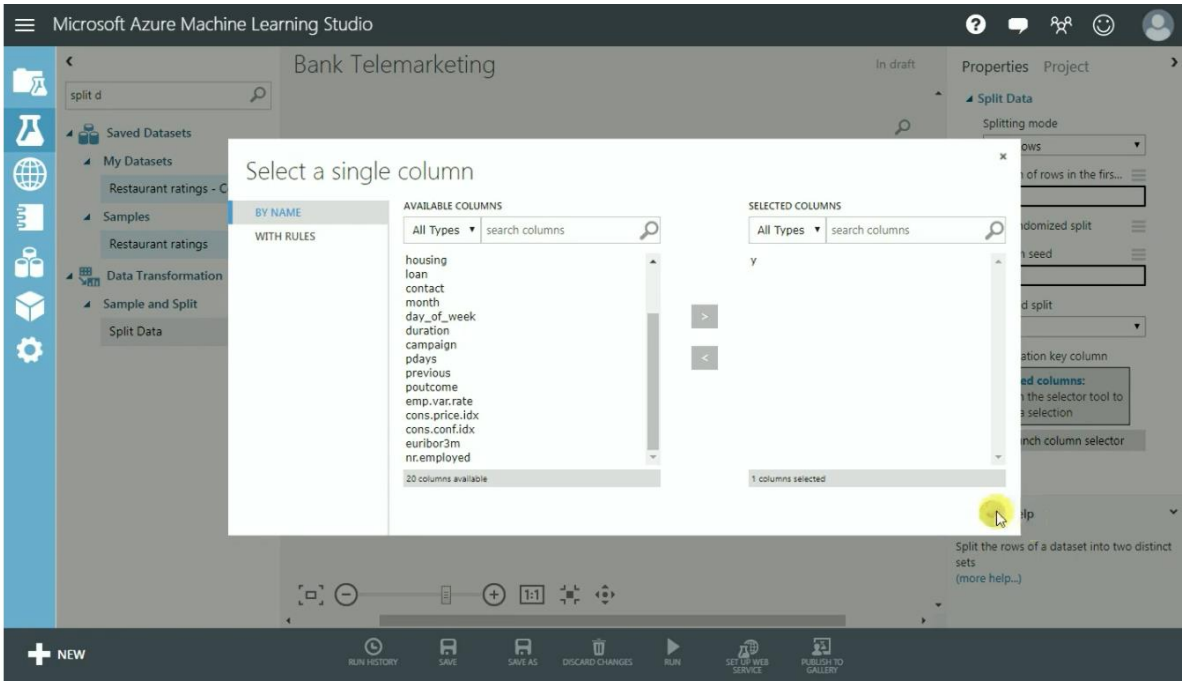
Split Data

Add split data and change parameters as 0.7 fractions, 123 random seed and stratified split as true

As below

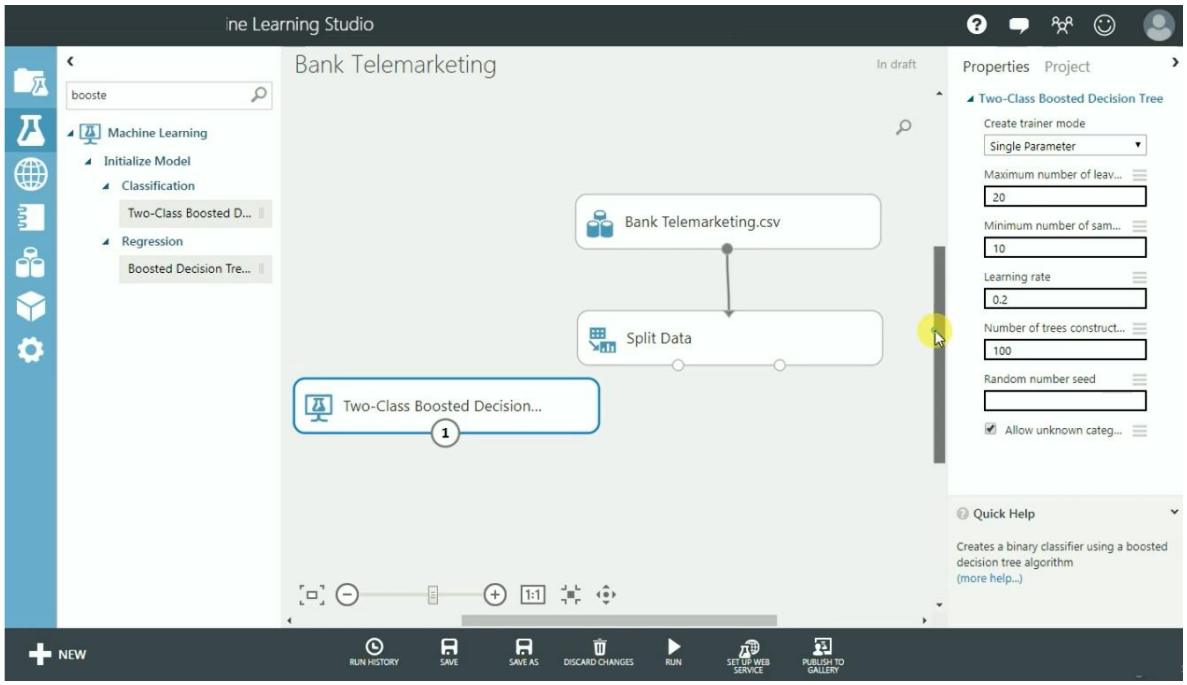


Launch the column selector and select and click ok

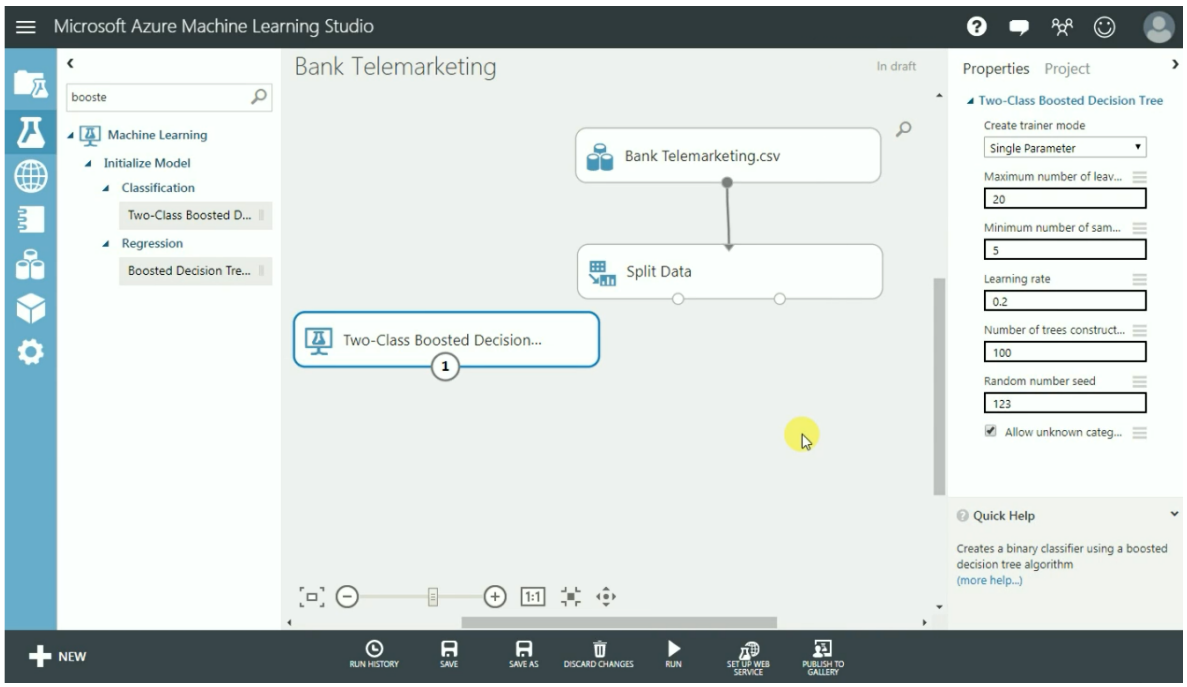


Two Class Boosted Decision

Place two class boosted decision in canvas

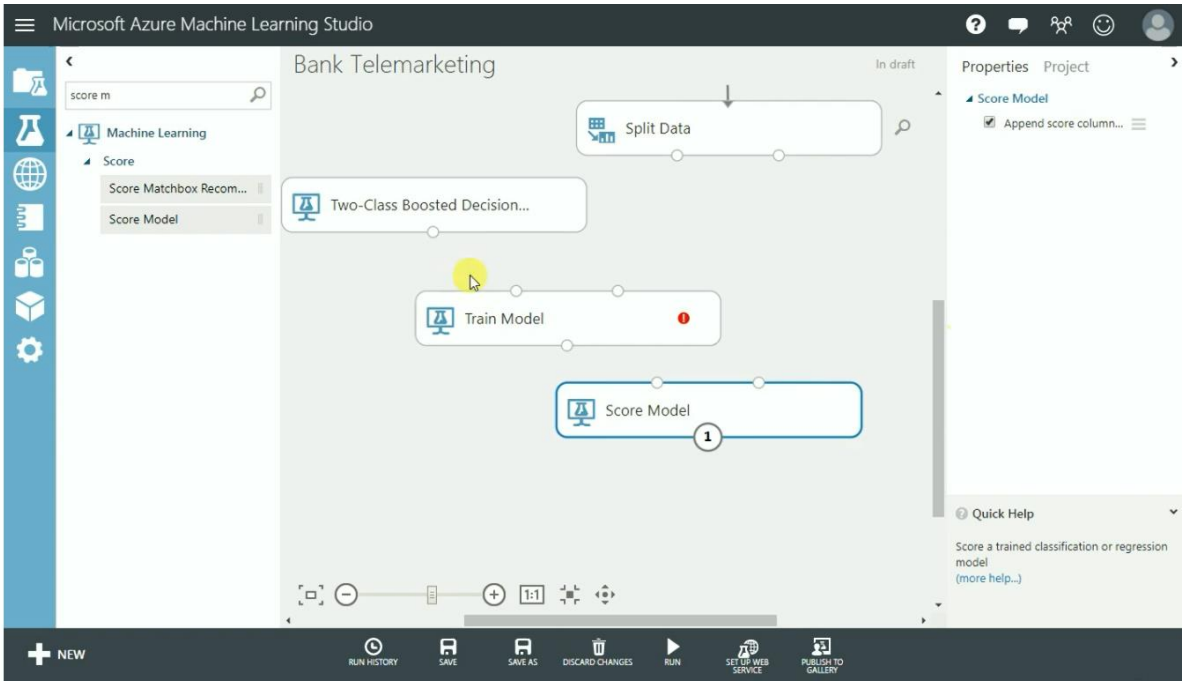


Input the parameters as below

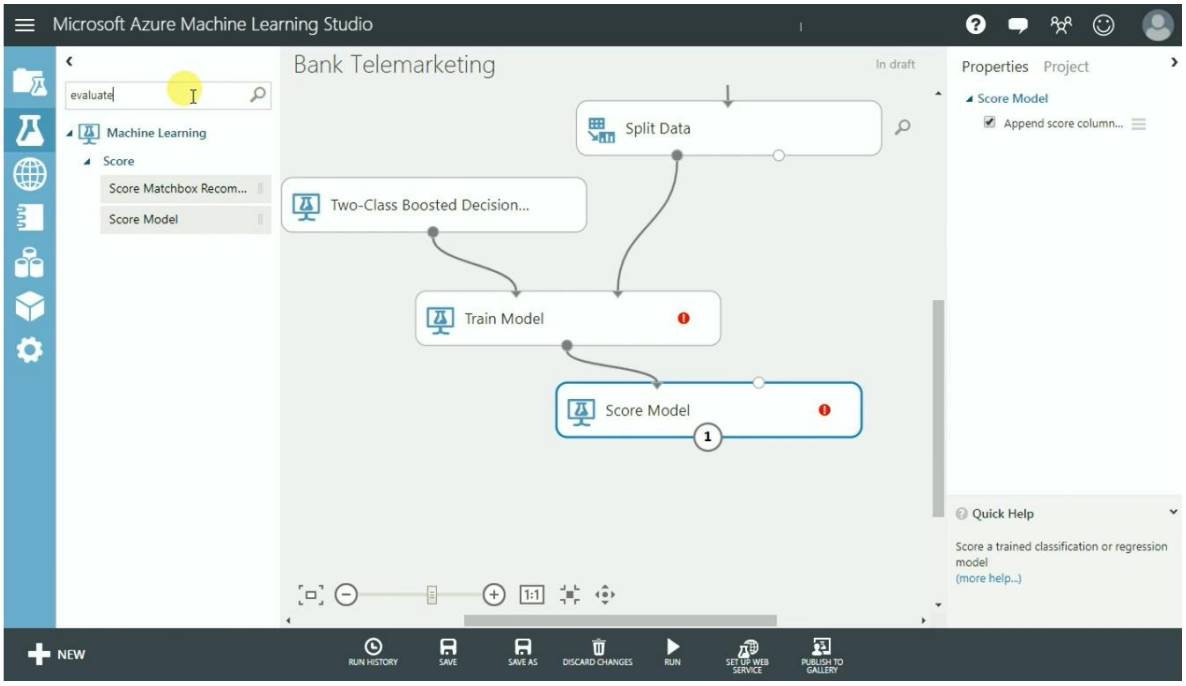


Train and Score Model

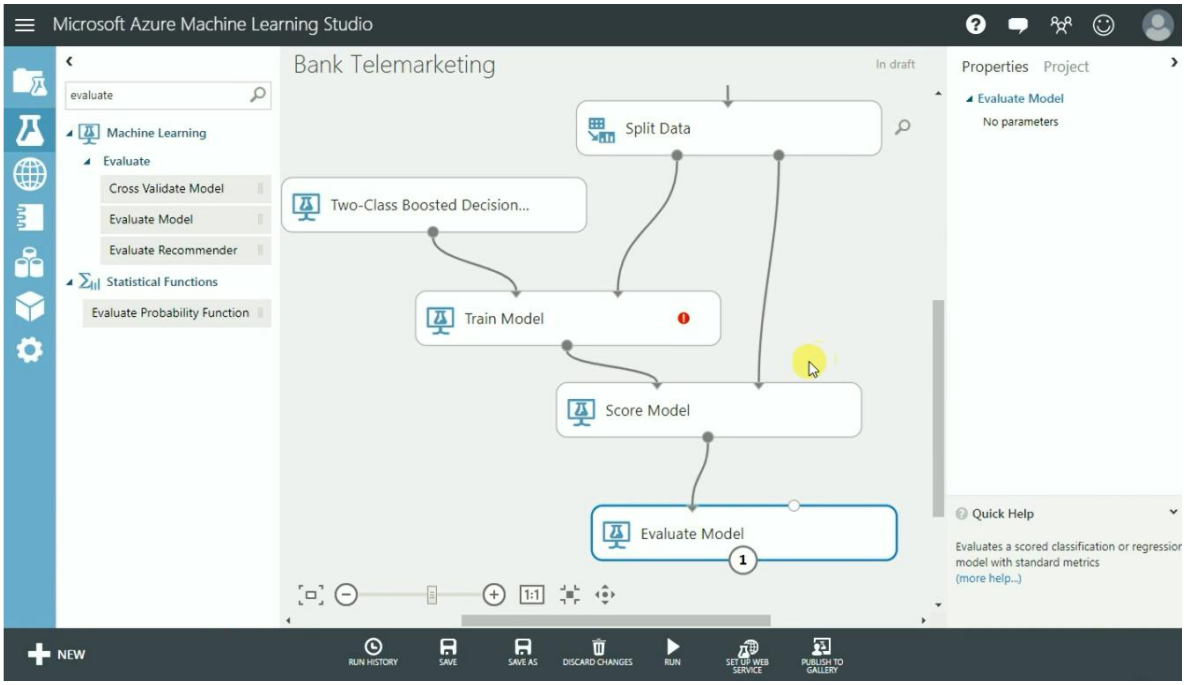
Add train model and score model in canvas



Connect the nodes in order

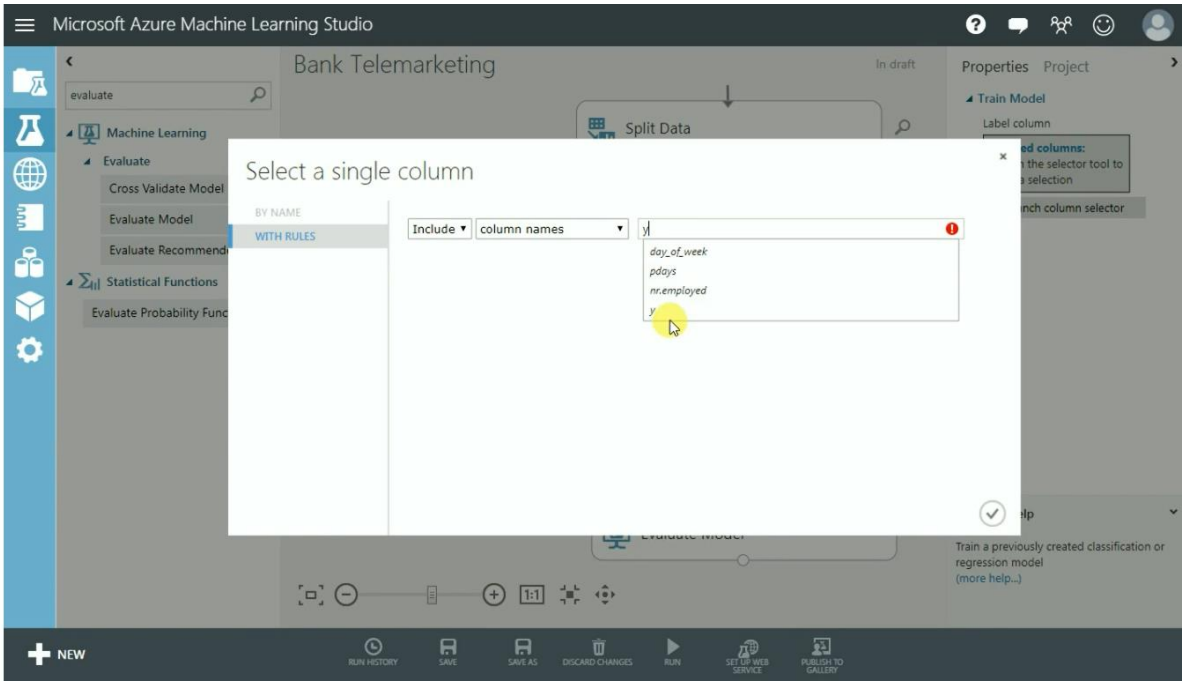


Add evaluate model and connect nodes as below

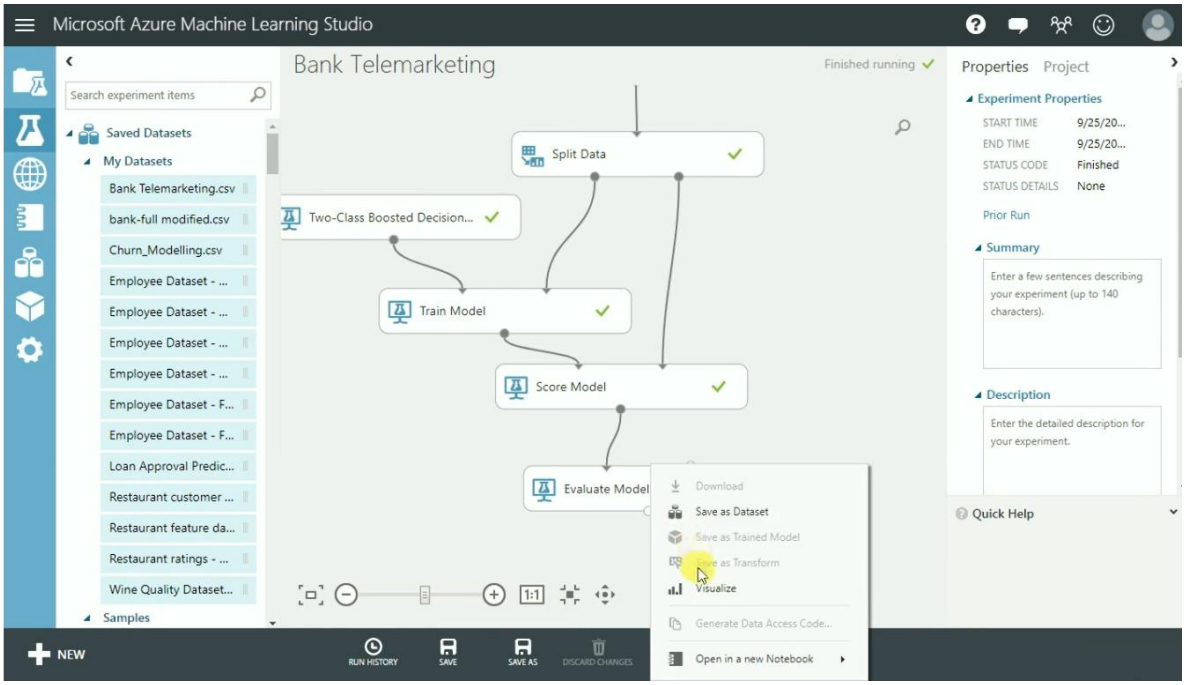


Evaluate Model

Click evaluate model and click launch column selector to select 'Y'



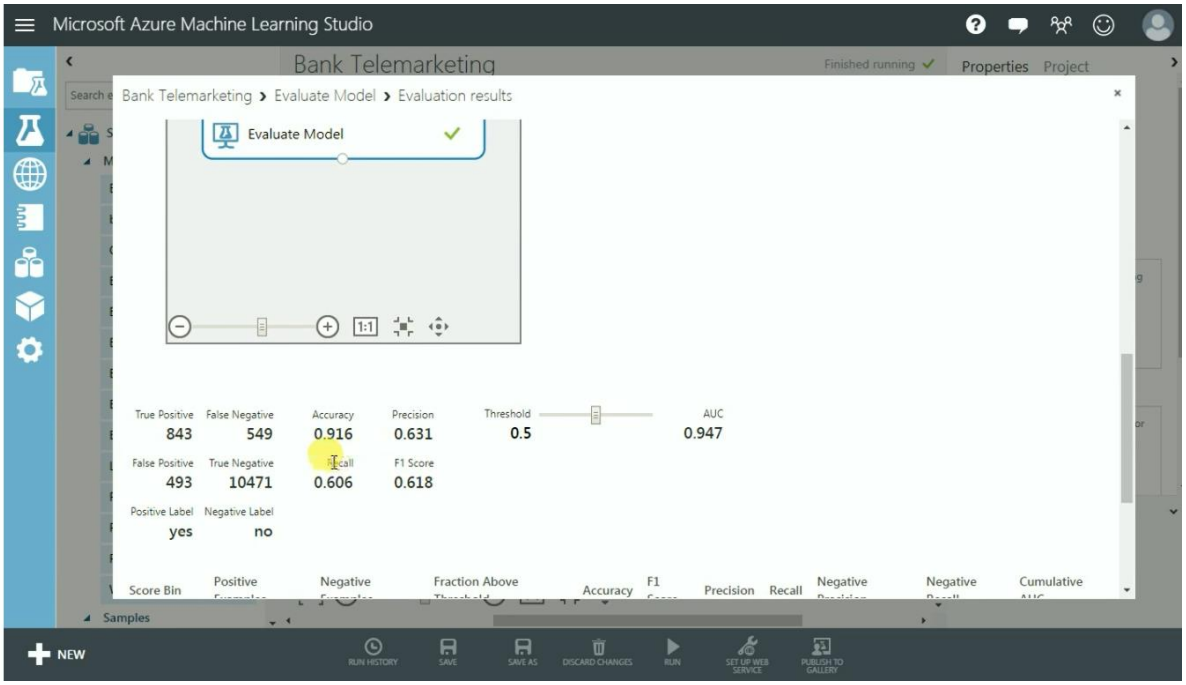
Now run the module and visualize the result



Result

Result obtained successfully with good accuracy to predict whether prospect

Will buy term deposit or not


Microsoft Azure Machine Learning Studio
Bank Telemarketing
Finished running
Properties
Project
Bank Telemarketing › Evaluate Model › Evaluation results
Evaluate Model
True Positive
843
False Negative
549
Accuracy
0.916
Precision
0.631
Threshold
0.5
AUC
0.947
False Positive
493
True Negative
10471
Recall
0.606
F1 Score
0.618
Positive Label
yes
Negative Label
no
Score Bin
Positive
Negative
Fraction Above
Accuracy
F1
Precision
Recall
Negative
Negative
Cumulative
Samples
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY