Batch: 2

Roll Number: 1601420061 **Experiment NO.6**

Name: Sargundeep Sachdeo

Title of the Experiment: Develop a classification model using Microsoft Azure Machine Learning Studio

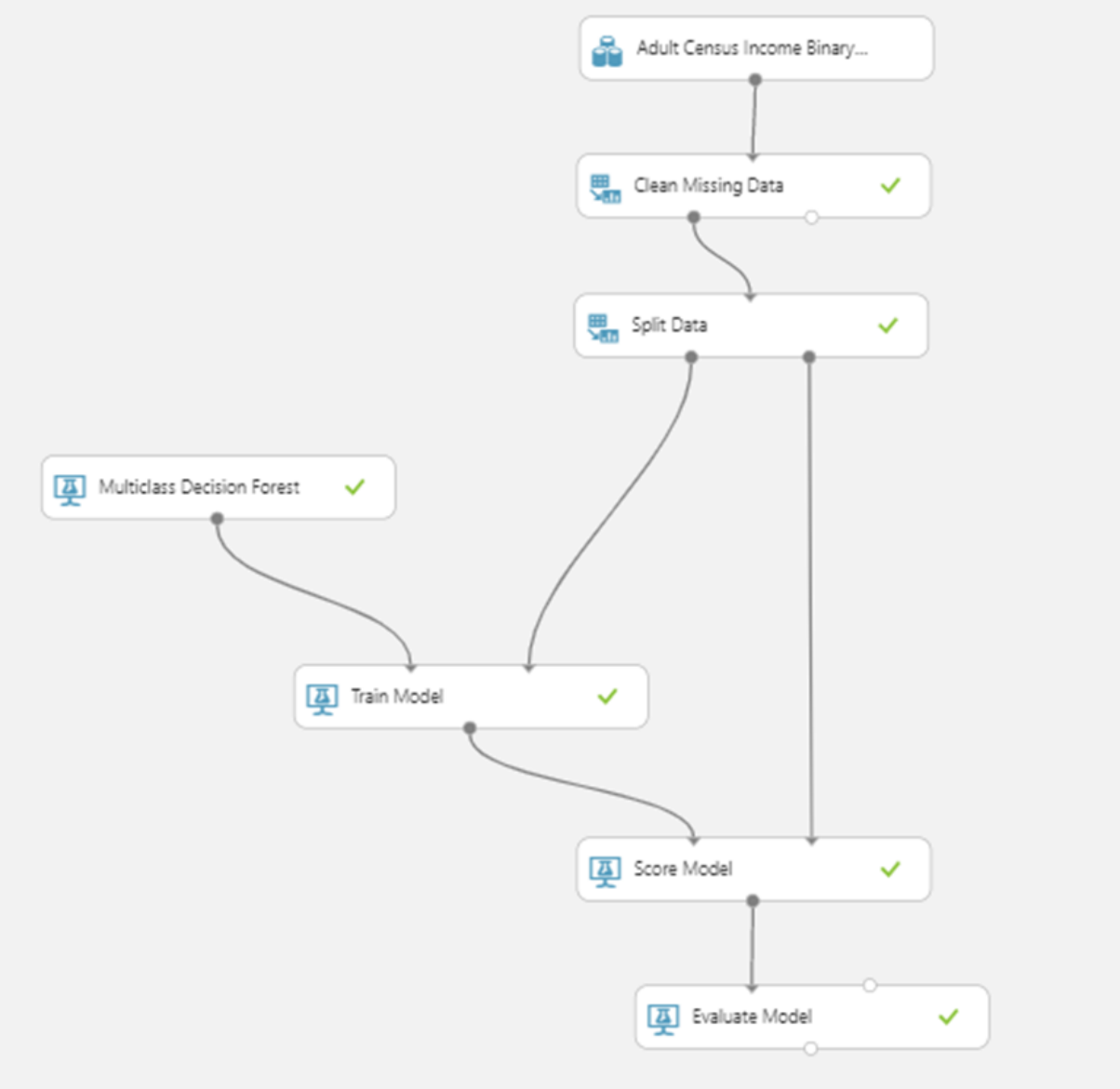
---

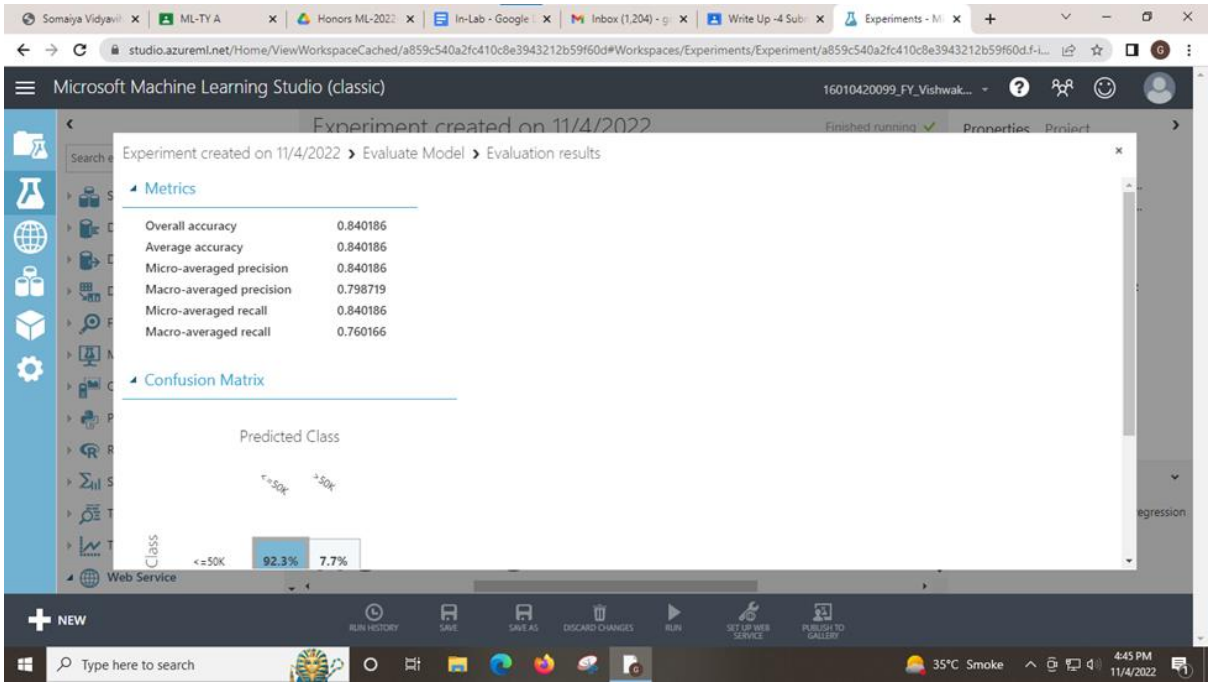
Program:

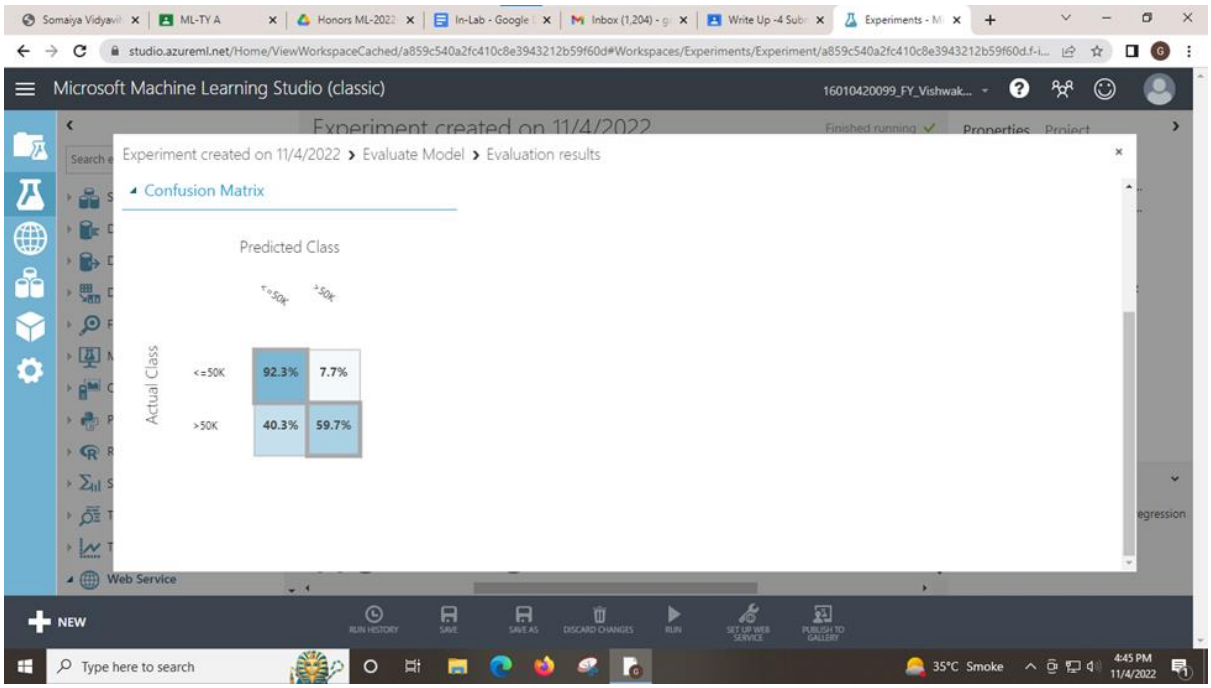
Dataset used: Adult Income Census

Classes: Two classes on the "income" column labeled as- '<50K' and '>50K'

Screenshots:


Adult Census Income Binary...
Clean Missing Data
Split Data
Multiclass Decision Forest
Train Model
Score Model
Evaluate Model


Microsoft Machine Learning Studio (classic)
Experiment created on 11/4/2022 › Evaluate Model › Evaluation results
Metrics
Overall accuracy 0.840186
Average accuracy 0.840186
Micro-averaged precision 0.840186
Macro-averaged precision 0.798719
Micro-averaged recall 0.840186
Macro-averaged recall 0.760166
Confusion Matrix
Predicted Class
<=50K
>50K
Class
92.3%
7.7%
Web Service
NEW
Type here to search
35°C Smoke
4:45 PM
11/4/2022


Microsoft Machine Learning Studio (classic)
Experiment created on 11/4/2022 › Evaluate Model › Evaluation results
Confusion Matrix
Predicted Class
<=50K
>50K
Actual Class
92.3%
7.7%
40.3%
59.7%
Web Service
NEW
Type here to search
35°C Smoke
4:45 PM
11/4/2022

Post Lab Question- Answers (If Any):

Q. Differentiate between linear and nonlinear regression and write a note on converting non-linear model into linear model

Ans: Linear regression always uses a linear equation, Y = a +bx, where x is the explanatory variable and Y is the dependent variable.In multiple linear regression, multiple equations are added together but the parameters are still linear.If the model equation does not follow the Y = a +bx form then the relationship between the dependent and independent variables will not be linear. There are many different forms of non-linear models. A random forest regression is considered a non-linear model. Random forest models are ensemble learning methods for regression which grow a forest of regression trees and then average the outcomes. This cannot be expressed as an equation.In regression trees, the splitting decision is based on minimizing the Residual Sum of Squares (RSS). The variable which has the greatest possible reduction in RSS is chosen as the root node. The tree splitting takes a top-down greedy approach, meaning the algorithm makes the best split at the current step rather than saving a split for better results on future nodes.

CO: Apply concepts of different types of Learning and Neural Network

Conclusion: Successfully Learnt and Implemented the model in Azure ML Studio.