

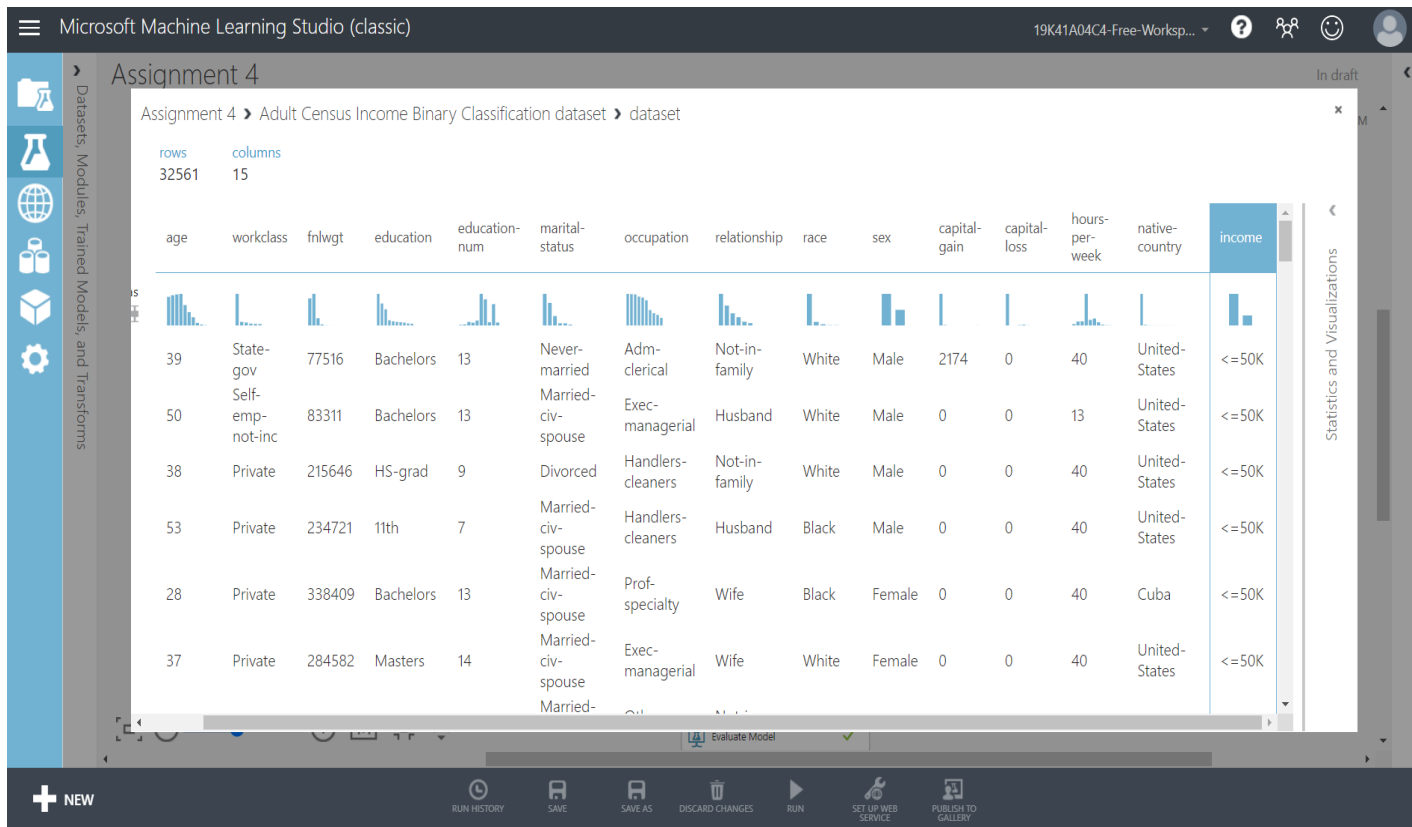
AI-ASSIGNMENT_4

Varshith Reddy

19K41A04C4

WORK FLOW:

Study the data



Clean Missing Data:

Remove the rows which are not having native-country, work class, occupation.

The screenshot displays the Microsoft Machine Learning Studio (classic) interface. The main workspace shows a workflow titled "Assignment 4" with the following steps:

- Clean Missing Data** (Step 1): This step is highlighted. It has a green checkmark and a status of "Finished".
- Split Data** (Step 2): This step also has a green checkmark and a status of "Finished".
- Two-Class Boosted Decision...** (Step 3): This step has a green checkmark and a status of "Finished".
- Train Model** (Step 4): This step has a green checkmark and a status of "Finished".
- Score Model** (Step 5): This step has a green checkmark and a status of "Finished".

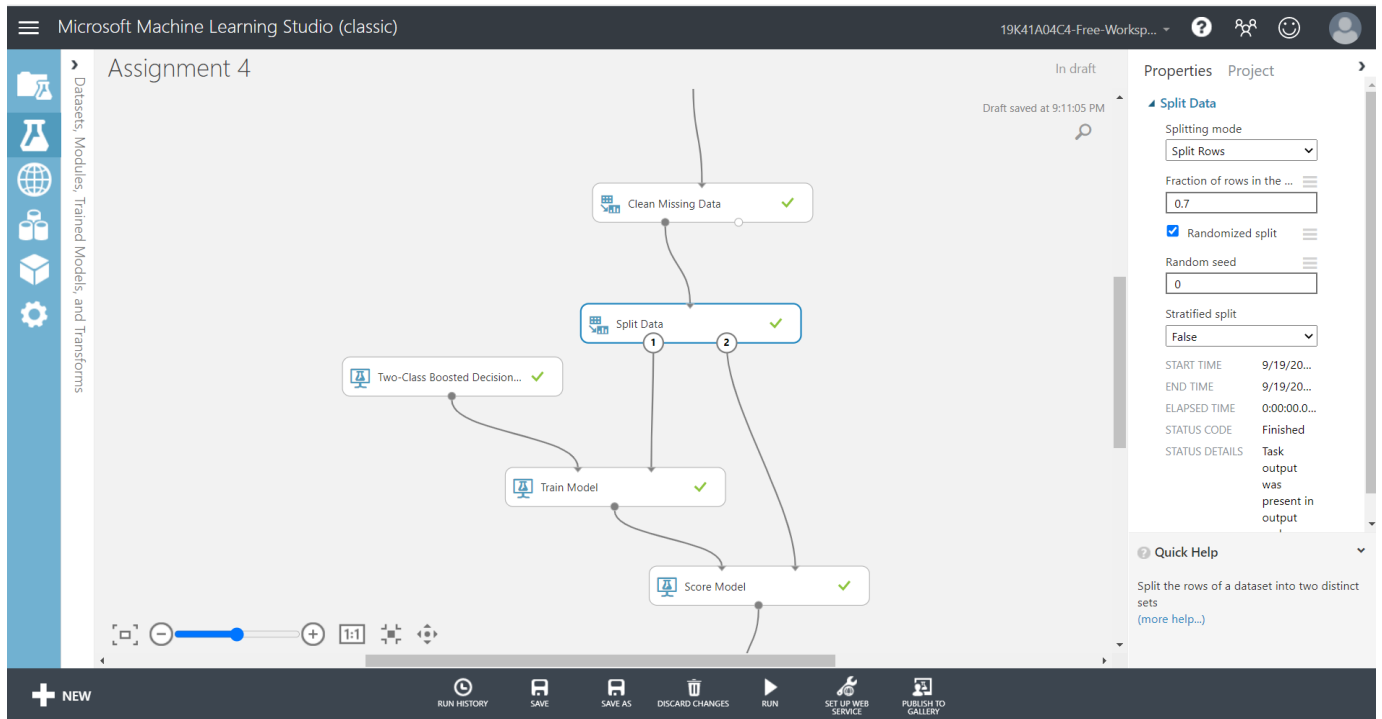
The right-hand pane shows the **Properties** tab for the **Clean Missing Data** step. The settings are as follows:

- Columns to be cleaned:** **Selected columns:** All columns; **Exclude column names:** native-country, workclass, occupation.
- Launch column selector:** A button to open the column selector.
- Minimum missing value:** 0.
- Maximum missing value:** 1.
- Cleaning mode:** Remove entire row.
- START TIME:** 9/19/20...
- END TIME:** 9/19/20...
- ELAPSED TIME:** 0:00:00.0...
- STATUS CODE:** Finished.

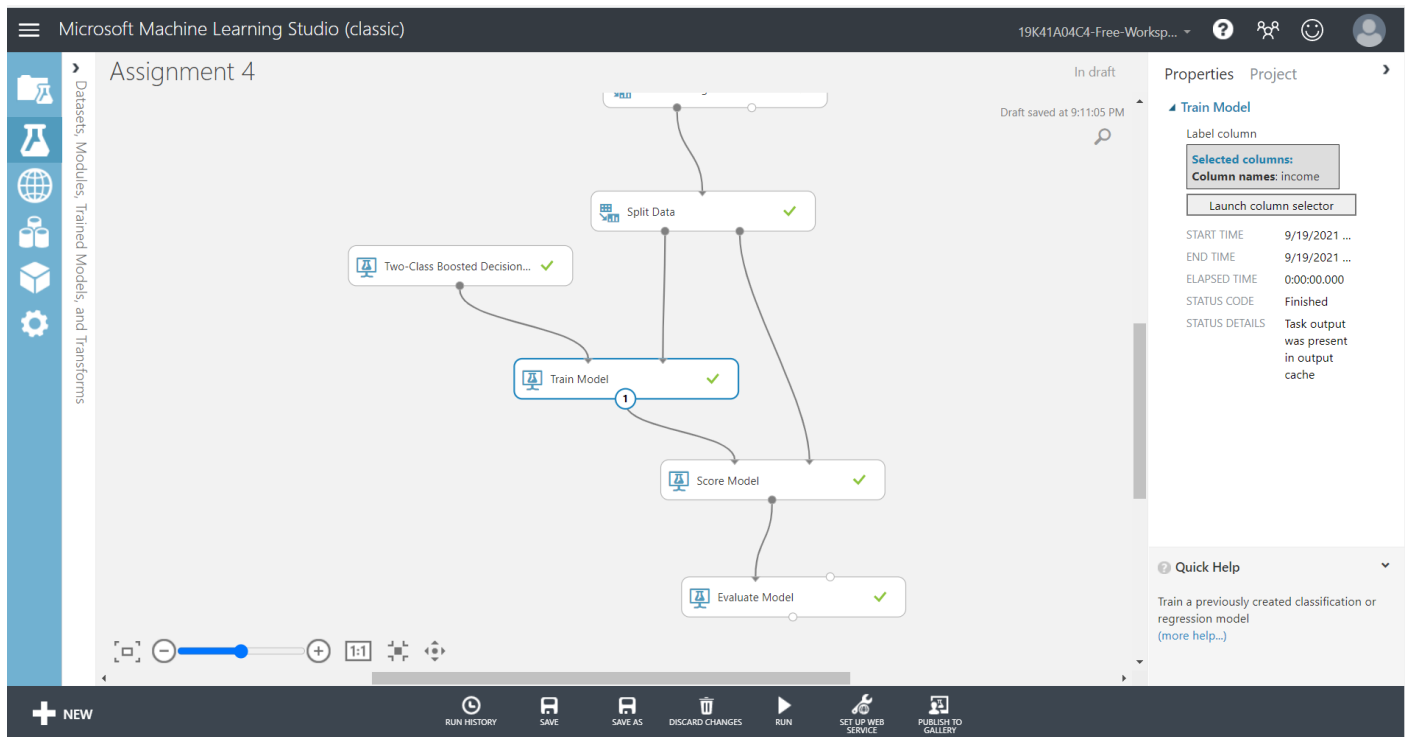
The bottom of the interface shows a toolbar with icons for **NEW**, **RUN HISTORY**, **SAVE**, **SAVE AS**, **DISCARD CHANGES**, **RUN**, **SET UP WEB SERVICE**, and **PUBLISH TO GALLERY**.

Split Data for training and testing:

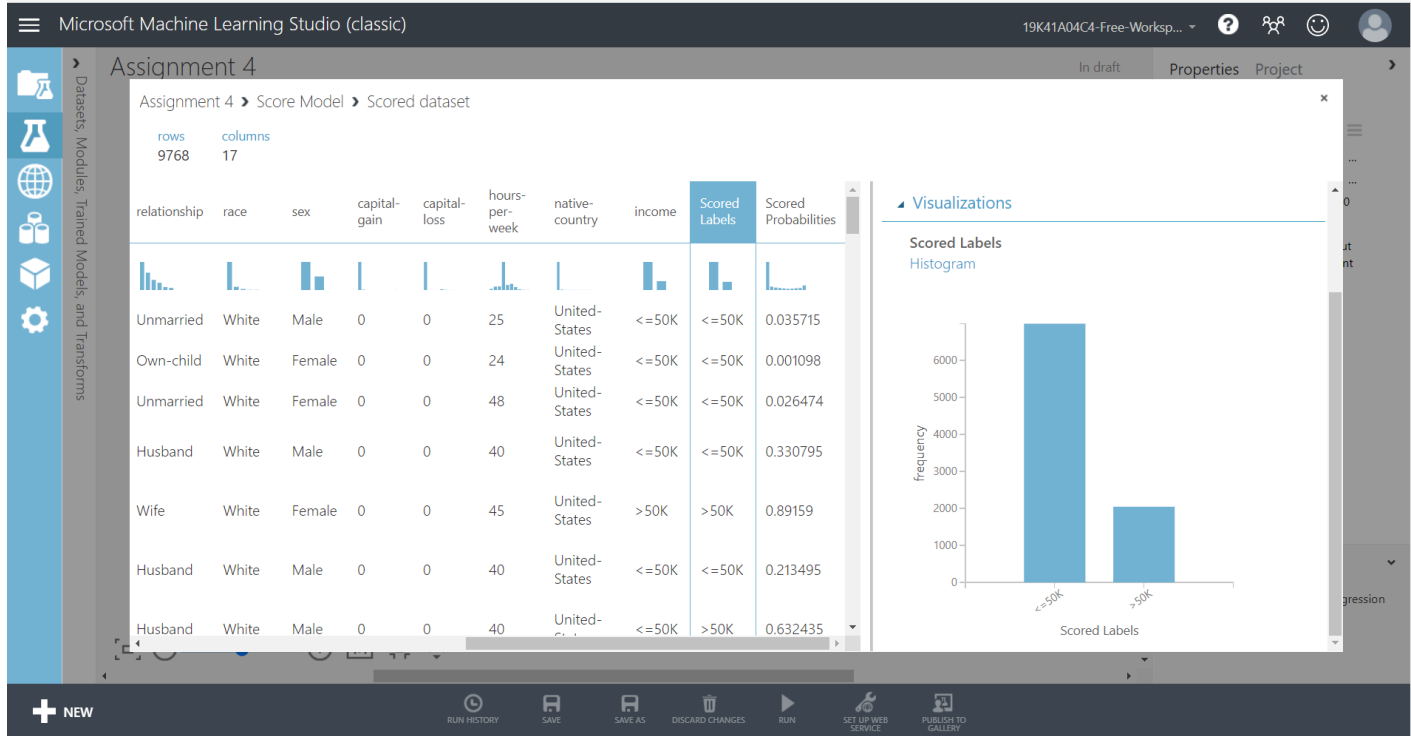
Split data in 70:30 ration that is 70% for training and 30% for testing



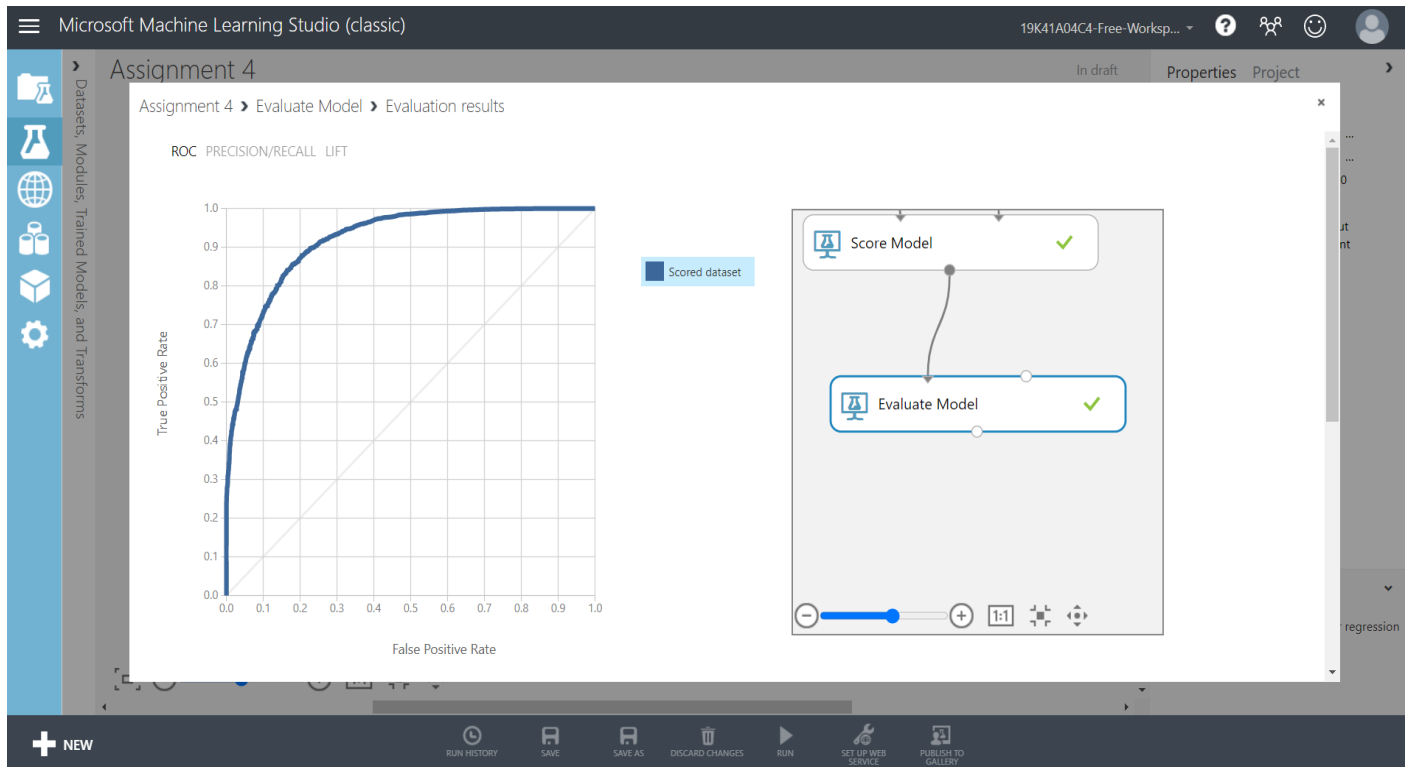
Train model with Two-Class Boosted Decision Tree and training data:

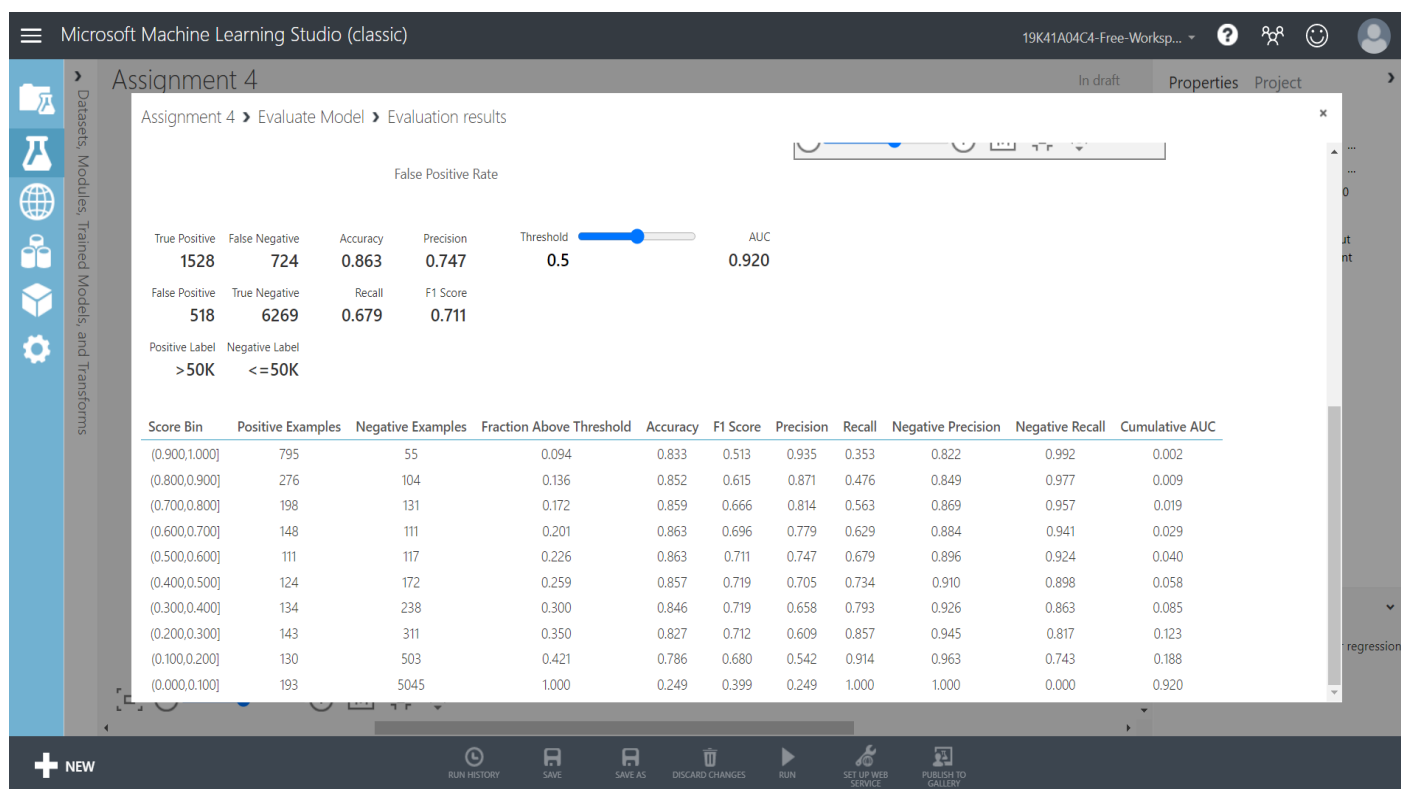


Score Data set:



Evaluation Results:





Conclusion:

From above results we can say that accuracy of the model is 86.3%.

Parameters used to predict are:

1. Age,
2. Fnlwgt,
3. Education,
4. Education-num,
5. Marital-status,
6. Relationship,
7. Race,
8. Sex,
9. Capital-gain,
10. Capital-loss,
11. Hours-per-week,
12. Income

Final Model:

