Azure ML Classic Studio

Predicting Income of Adults using Binary Classification in Azure ML Classic Studio.

This model (Pipeline) trains a two-class binary classifier to predict income of the Adults using the Available census data. it's a binary classification problem as you need to classify the Adults based on the two income classes, >50K and ≤50K However, you can apply the same fundamental steps in this example to tackle any type of machine learning problem whether it be regression, classification, clustering, and so on.

Gallery Link:

Adult Income census data

Using binary classification to estimate the income of the adults. Tags: Binary Classification, Azure, Machine learning

https://gallery.azure.ai/Experiment/Adult-Income-census-data

navigate to the link to see the Workflow and you can download the project as well.

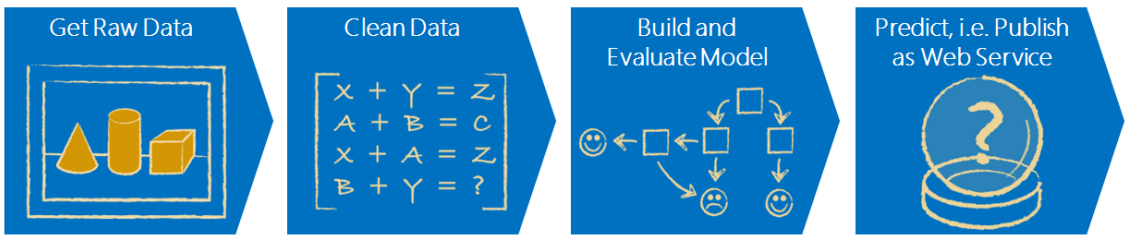
Prabhu Kiran

GitHub/prabhukiran8790

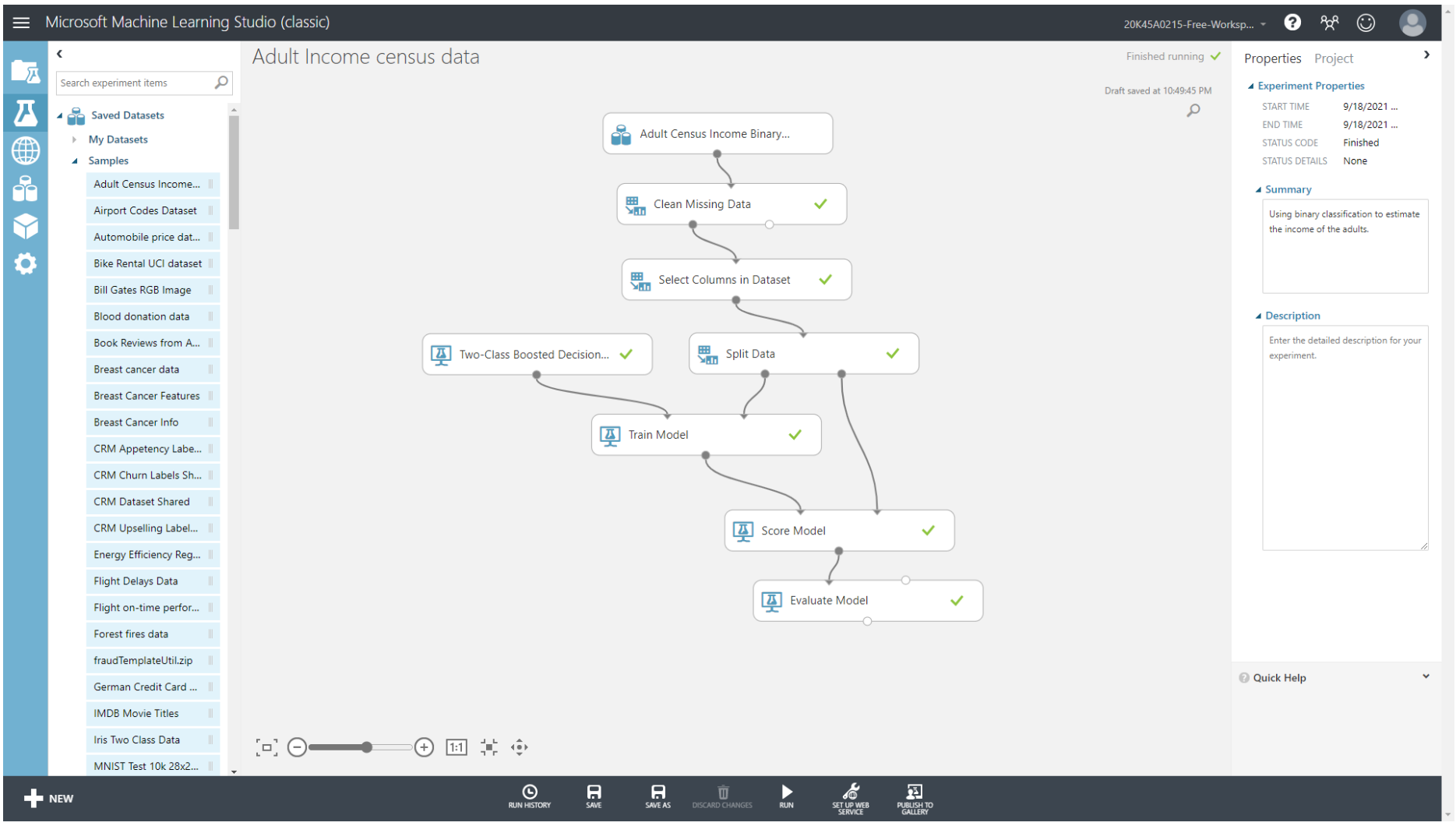
prabhukiran426@outlook.com

Machine Learning Project Workflow

1. **Import Data**
2. **Explore Data (Missing values, outliers)**
3. **Preprocess data (Missing value imputation, outlier treatment, normalization)**
4. **Model Selection**
5. **Model Training**
6. **Model Testing**
7. **Model Deployment**



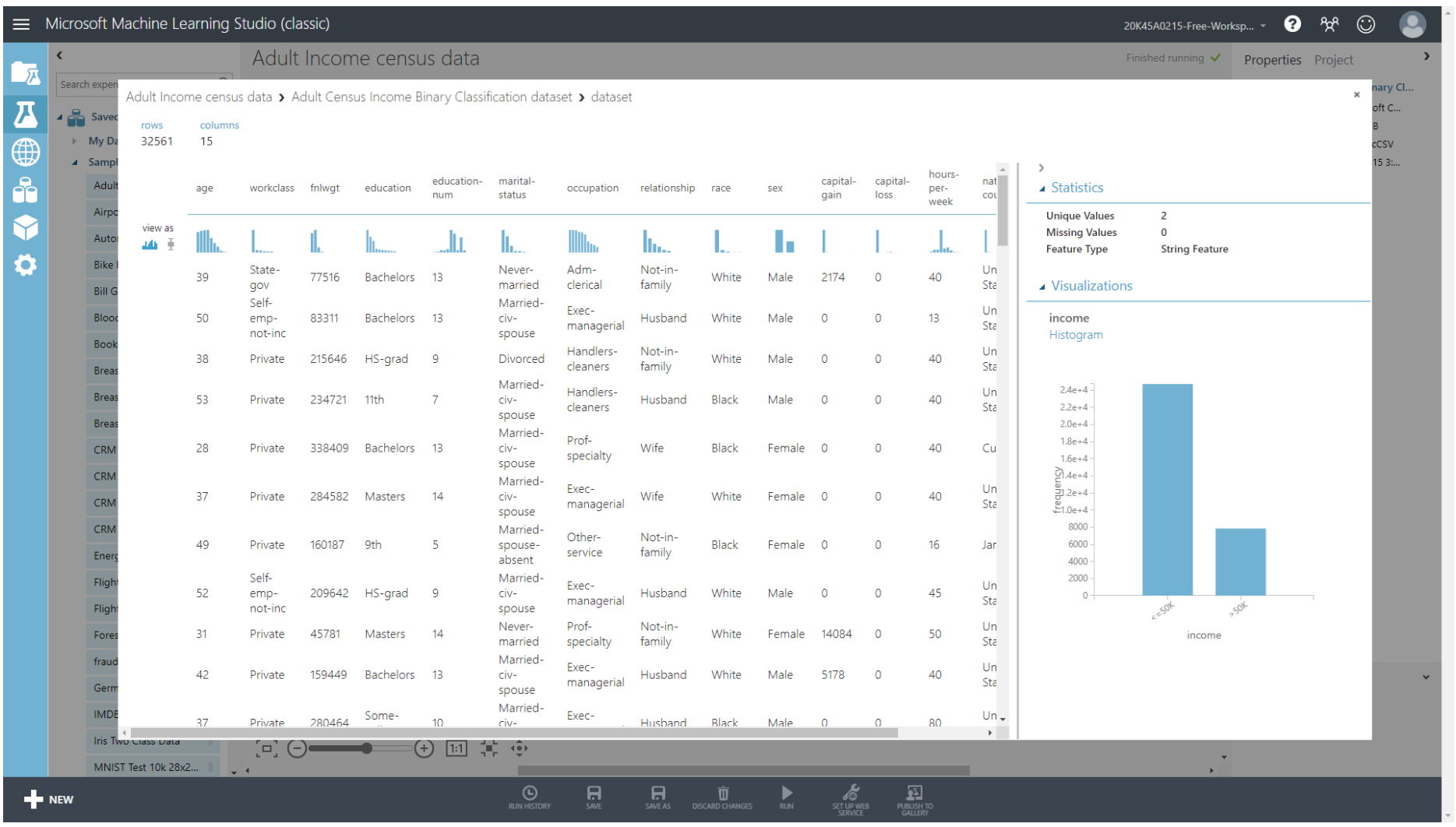
Workflow



Project Workflow

Import Data:

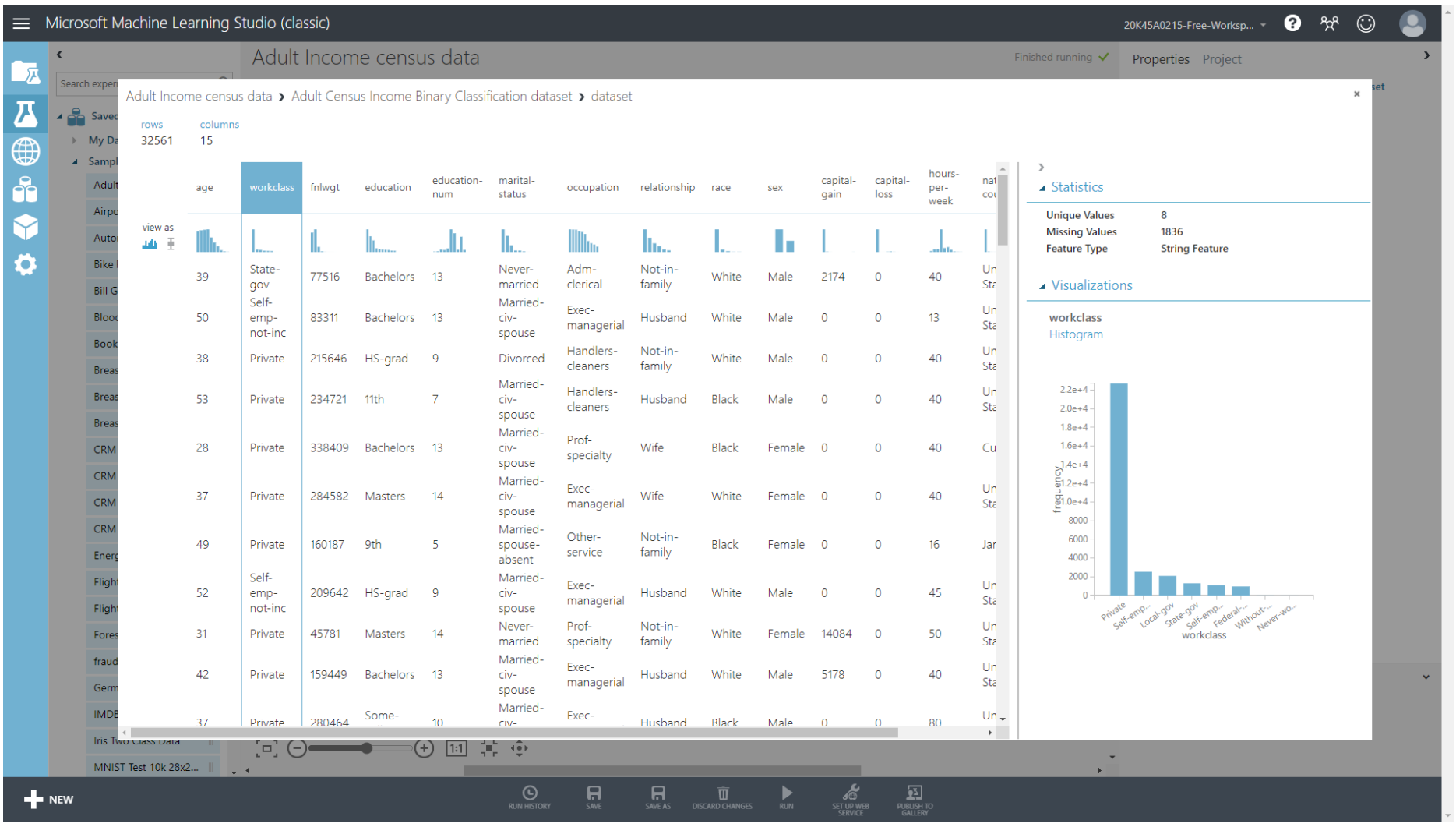
- importing the RAW dataset which is in CSV format.



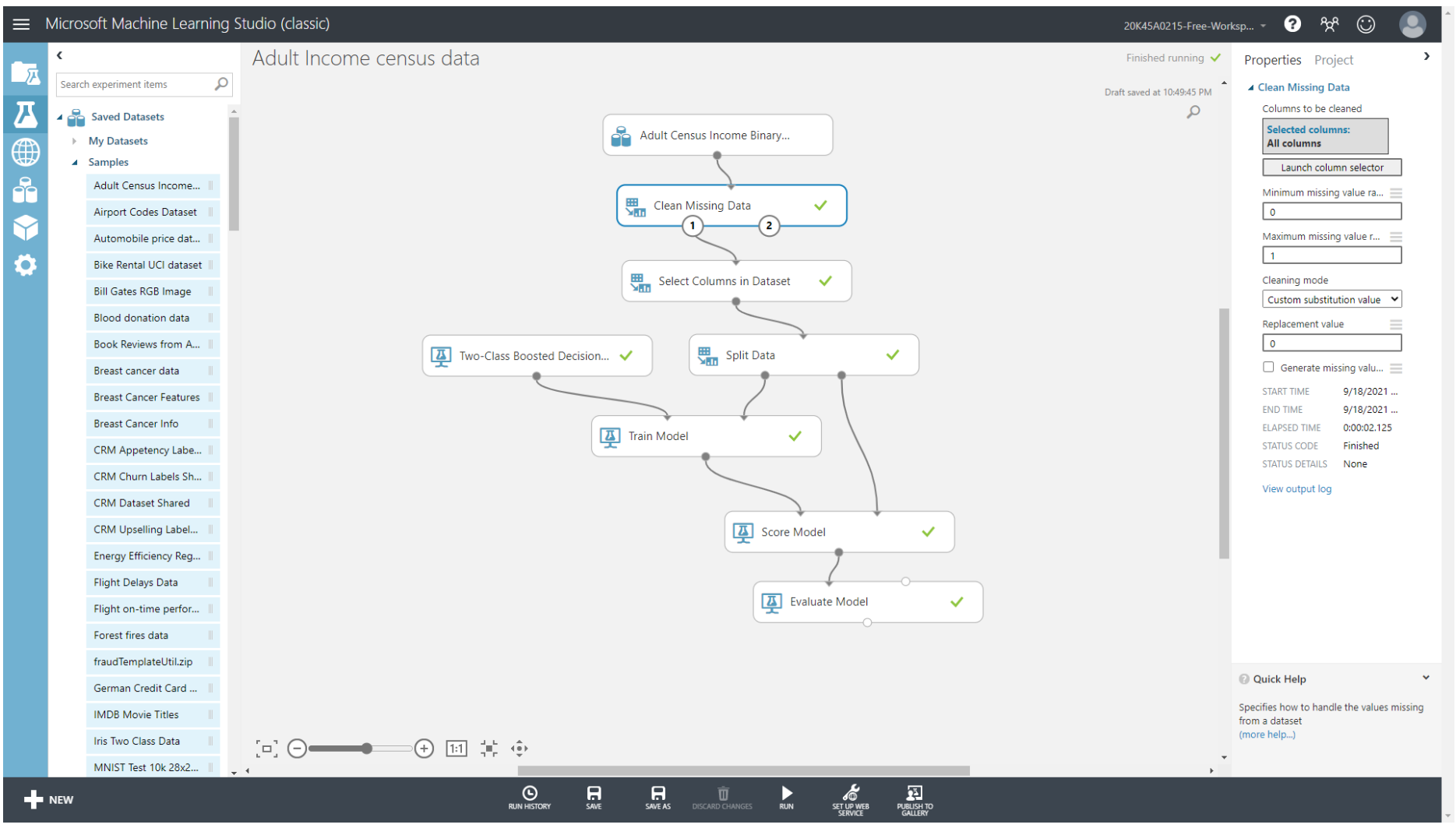
RAW dataset (CSV format)

Explore Data

- this basically includes data visualization to search for any missing values in the Dataset.
- if any missing values are found, then they needs to be cleaned.
- selecting the required columns and clean the data using the Clean Missing Value module (Just Drag n' Drop)



Explore the data and find the missing values in the Dataset, those missing values are to be cleaned.

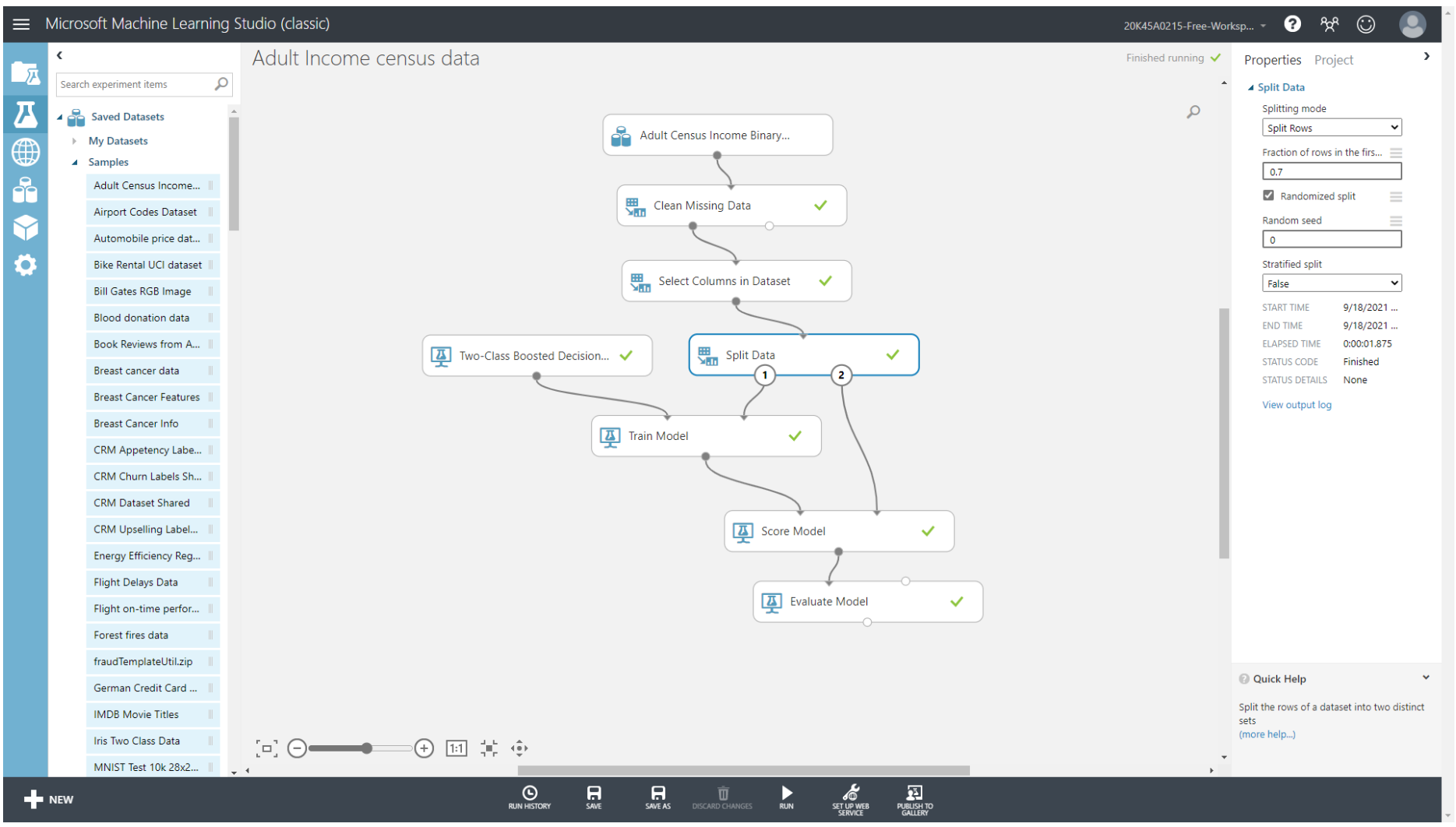


Data Cleaning

20K45A0215 Prabhu Kiran Konda

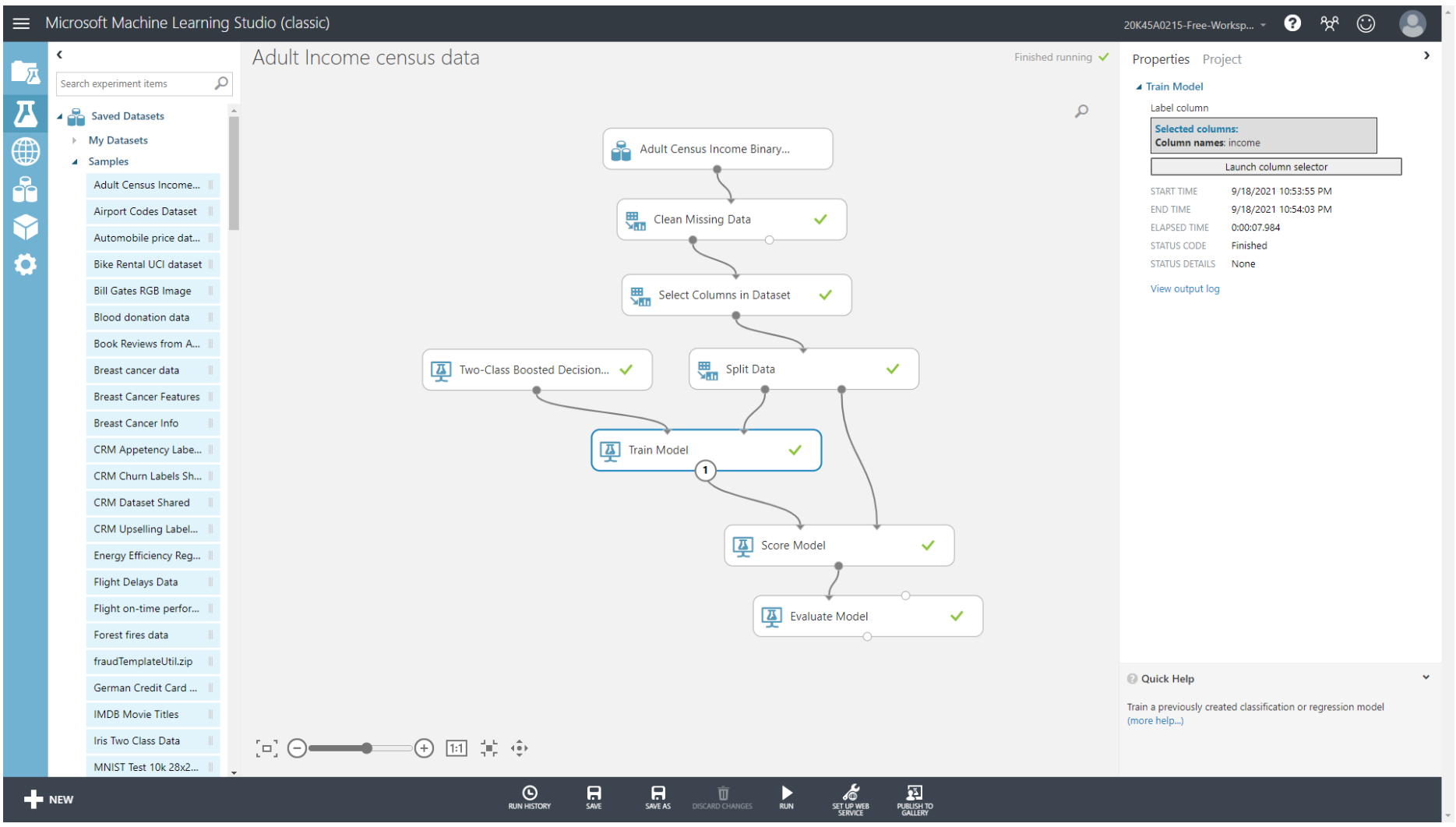
Split Data

- Use the Split Data module to randomly divide the input data so that the training dataset contains 70% of the original data and the testing dataset contains 30% of the original data.



Data Splitting

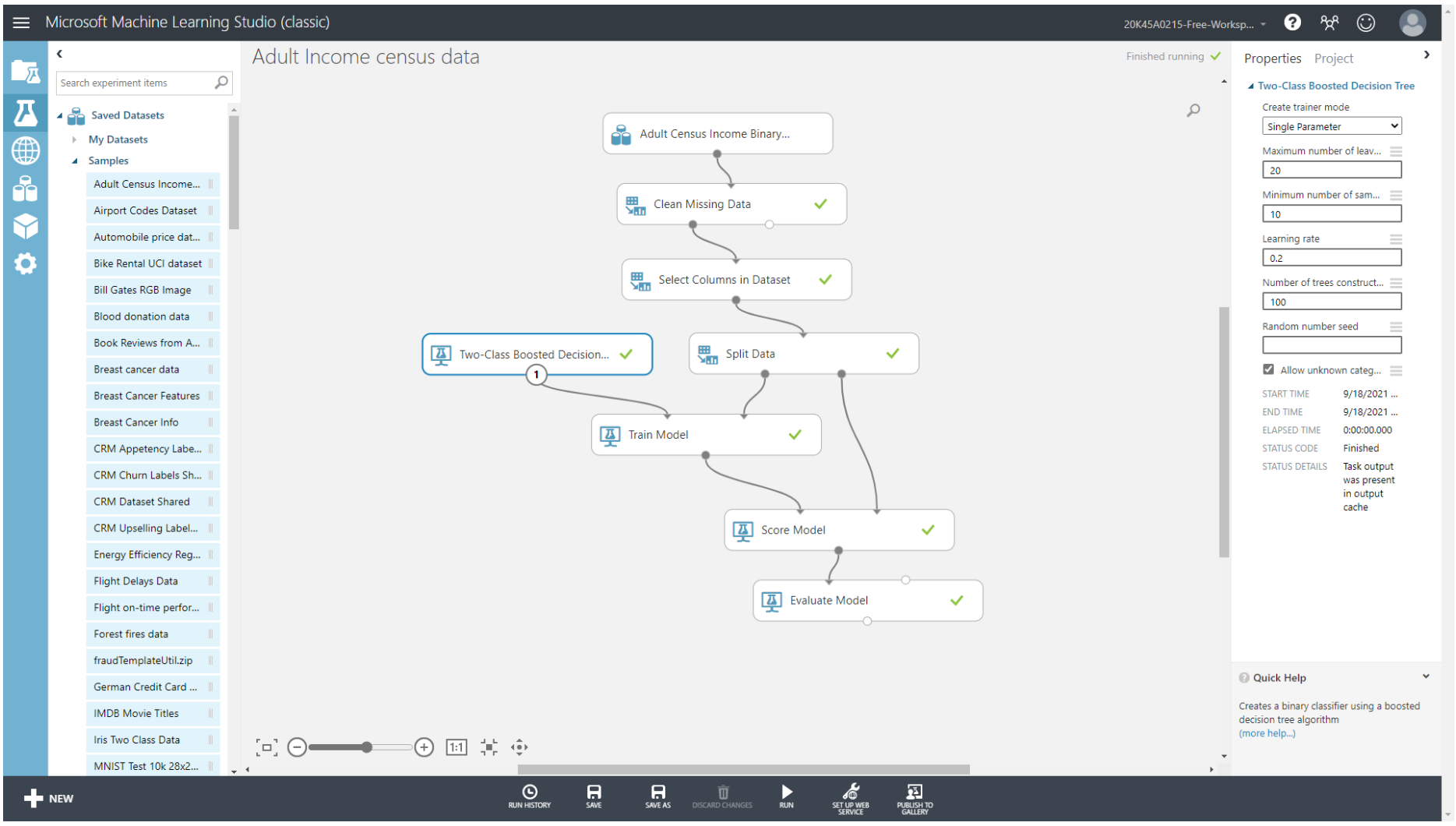
Model Training and Algorithm



Model Training

using Two-class boosted Binary Classification to train the model

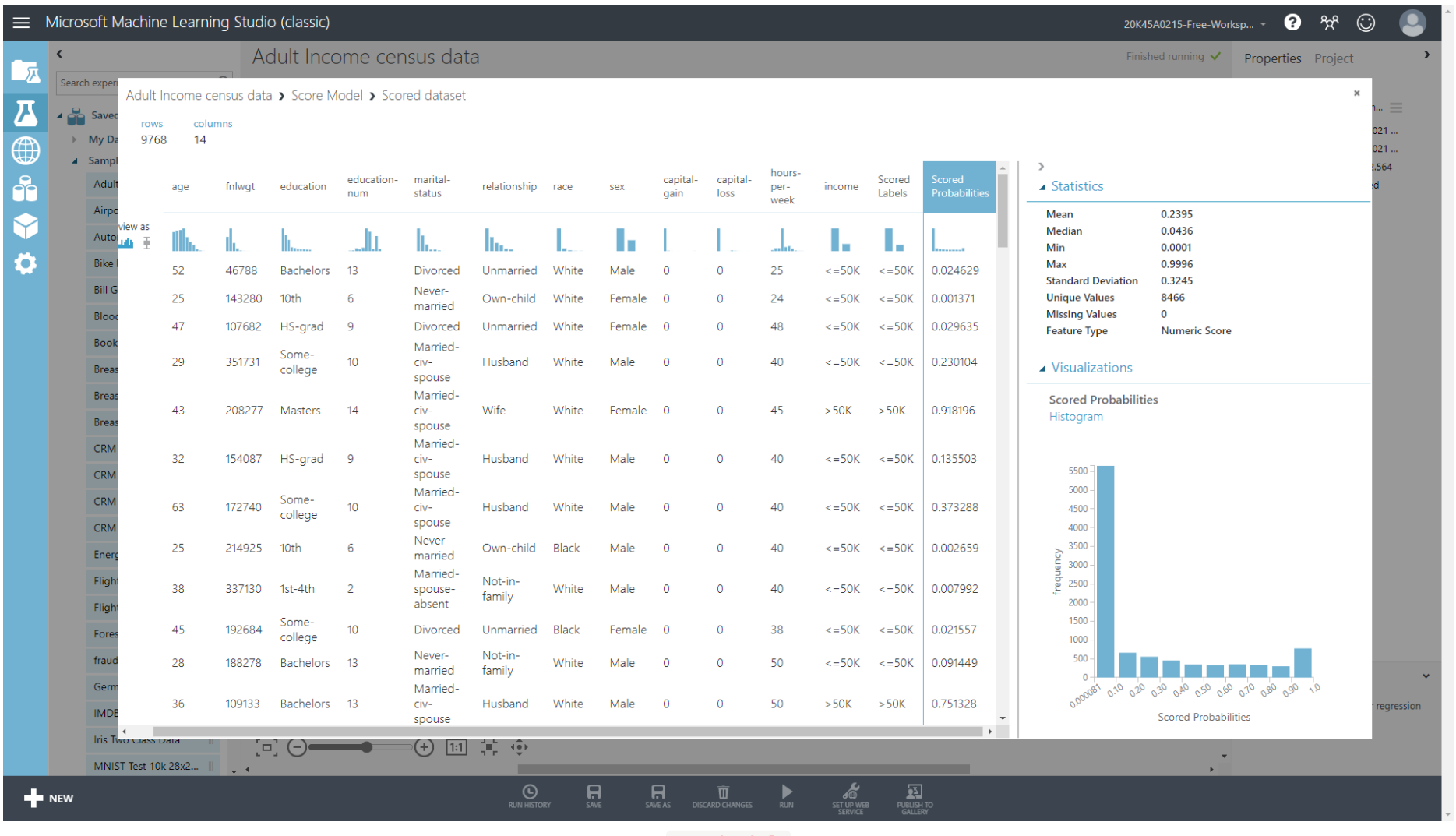
- Since the goal of this sample is to classify the Adults based on their income range, the best approach for this Model is Binary Classification.



Linear Regression

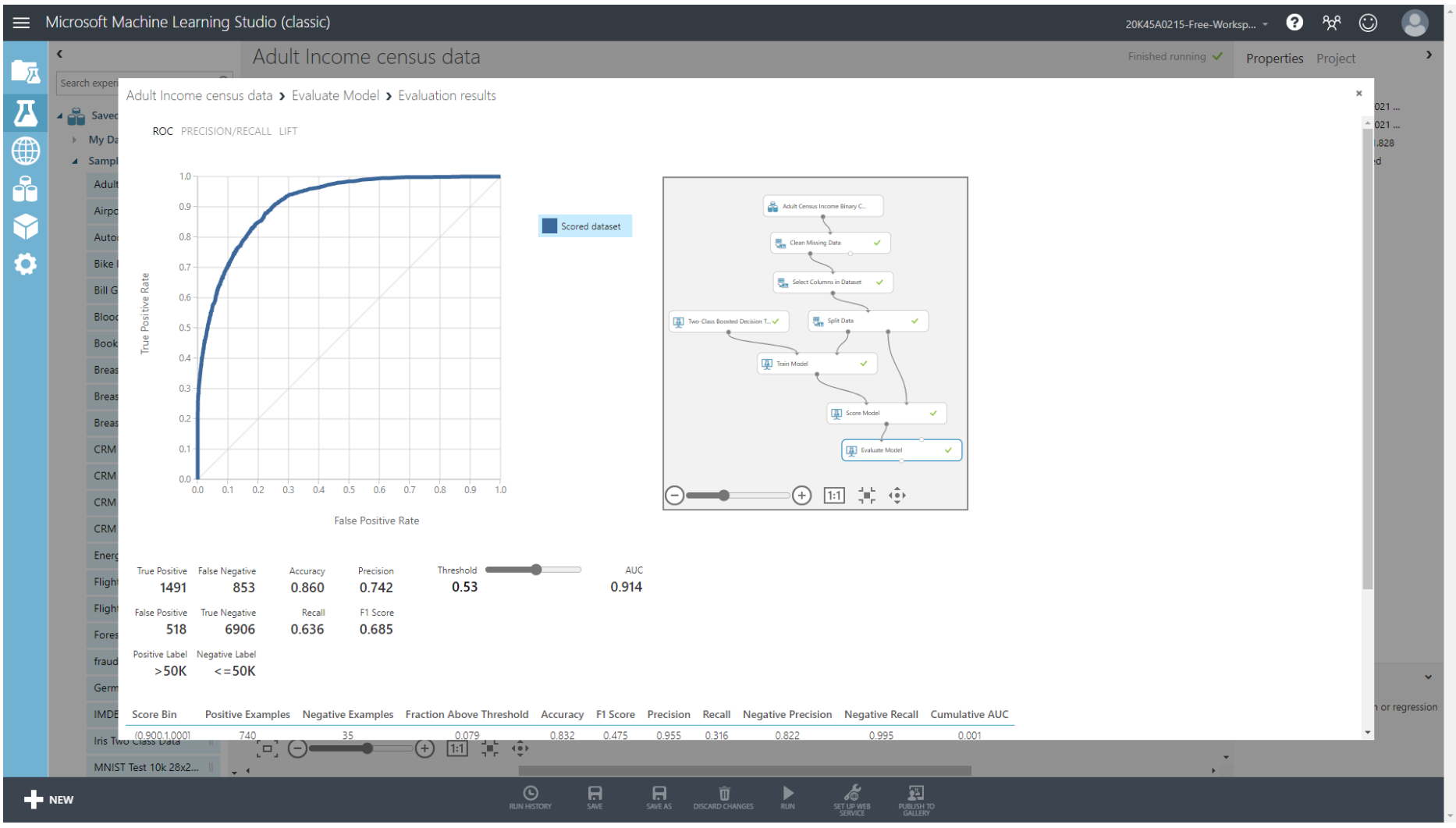
Score Model and Evaluate Model

- After the model is trained, we can use the Score Model and Evaluate Model modules to generate predicted results and evaluate the models.



Scored Labels

Evaluation Results



Model Evaluation Results