

## 01 binary-classification Azure ML Designer

## 1. go to marketplace/azure machine learning

portal.azure.com/#create/Microsoft.MachineLearningServices

Microsoft Azure Upgrade Search resources, services, and docs (G+)

Home > Create a resource > Marketplace > Azure Machine Learning >

### Azure Machine Learning

Create a machine learning workspace

**Resource details**

Every workspace must be assigned to an Azure subscription, which is where billing happens. You use resource groups like folders to organize and manage resources, including the workspace you're about to create. [Learn more about Azure resource groups](#)

Subscription \* Azure subscription 1

Resource group \* (New) az-ml [Create new](#)

**Workspace details**

Configure your basic workspace settings like its storage connection, authentication, container, and more. [Learn more](#)

Name \* nick-ml ✓

Region \* Central India

Storage account \* (new) nickml4554708411 [Create new](#)

Key vault \* (new) nickml9995264475 [Create new](#)

Application insights \* (new) nickml8586855700 [Create new](#)

Container registry None [Create new](#)

[Review + create](#) < Previous Next : Inbound Access

## 2. Review and create... and then go to the resource

portal.azure.com/#@yagamicloudoutlook293.onmicrosoft.com/resource/subscriptions/7e249995-8eee-47fb-8c7e-d3f61bb09227/resourc...

Microsoft Azure Upgrade Search resources, services, and docs (G+)

Home > Microsoft.MachineLearningServices | Overview >

### nick-ml

Azure Machine Learning workspace

Search Download config.json Delete

**Overview**

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Resource visualizer
- Events
- Settings
- Monitoring
- Automation
- Support + troubleshooting

**Essentials**

Resource group	: az-ml	Studio web URL	: <a href="https://ml.azure.com?tid=9c9287e9-c148-4944-9411-ce345f19ec96...">https://ml.azure.com?tid=9c9287e9-c148-4944-9411-ce345f19ec96...</a>
Location	: Central India	Container Registry	: ...
Subscription	: <a href="#">Azure subscription 1</a>	Key Vault	: <a href="#">nickml9995264475</a>
Storage	: <a href="#">nickml4554708411</a>	Application Insights	: <a href="#">nickml8586855700</a>
Provisioning State	: Succeeded	MLflow tracking URI	: <a href="#">azureml:/centralindia.ap.azureml.ms/mlflow/v1.0/subscriptions/7e...</a>

[JSON View](#)

**Work with your models in Azure Machine Learning Studio**

The Azure Machine Learning Studio is a web app where you can build, train, test, and deploy ML models. Launch it now to start exploring, or [learn more about the Azure Machine Learning studio](#)

[Launch studio](#)

Add or remove favorites by pressing Ctrl+Shift+F

## 3. Click on launch studio

## 4. Explore the UI

## 5...

**Building a Classification Machine Learning Pipeline**

The classification model is used to predict values that fall into simple categories.

We are going to use Azure Machine Learning Studio to build a Machine Learning model.

We will use a pre-built data set that is already available in ML studio - **adult\_census\_income\_binary\_classification**

Rows 32,561 Columns 15

age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	capital-loss	hours-per-week	native-country	income
39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40	United-States	<=50K
50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	White	Male	0	0	13	United-States	<=50K
38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40	United-States	<=50K
53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	Black	Male	0	0	40	United-States	<=50K
28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	Black	Female	0	0	40	Cuba	<=50K
37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	White	Female	0	0	40	United-States	<=50K
49	Private	160187	9th	5	Married-spouse-	Other-service	Not-in-	Black	Female	0	0	16	Jamaica	<=50K

The model will take in a set of features and the label will be the income.

We are going to create a pipeline to create the Machine Learning model.

An Azure Machine Learning pipeline is a workflow that is used to execute a machine learning task.

We can develop the pipeline in the designer of the Azure Machine Learning studio. We can drag data sets and components onto the canvas for our pipeline.

## 6. Compute instance

Azure AI | Machine Learning Studio

Default Directory > nick-ml > Compute

### Compute

The "Kubernetes clusters" tab is now where you can access previous versions of "inference clusters" (also known as "AKS clusters") and "attached Kubernetes" compute types along with any previously created compute targets using those types. [Learn more about Kubernetes clusters.](#)

Compute instances    Compute clusters    Kubernetes clusters    Attached computes    Serverless instances

Choose from a selection of CPU or GPU instances preconfigured with popular tools such as VS Code, JupyterLab, Jupyter, and RStudio, ML packages, deep learning frameworks, and GPU drivers. [Learn more about compute instances](#)

**+ New**    Refresh    Start    Stop    Restart    Schedule and idle shutdown    Delete    Reset view    View quota

Search

Name	State	Idle shutdown	Applications	Size	Created on
yagamicloud1	Creating	--		STANDARD_DS11_V2	Apr 25, 2025 8:3

**Create a new compute instance**

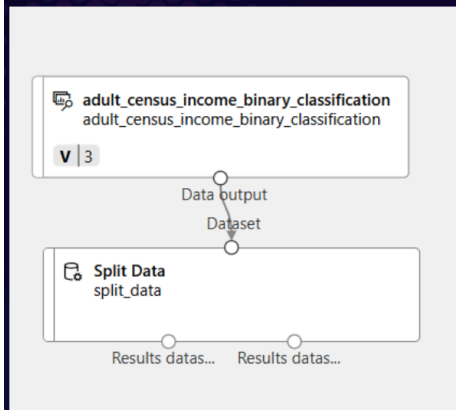
Manage

- Compute
- Monitoring
- Data Labeling
- Linked Services
- Connections

## 6. Run the pipeline and check the splits

7.

## Split data

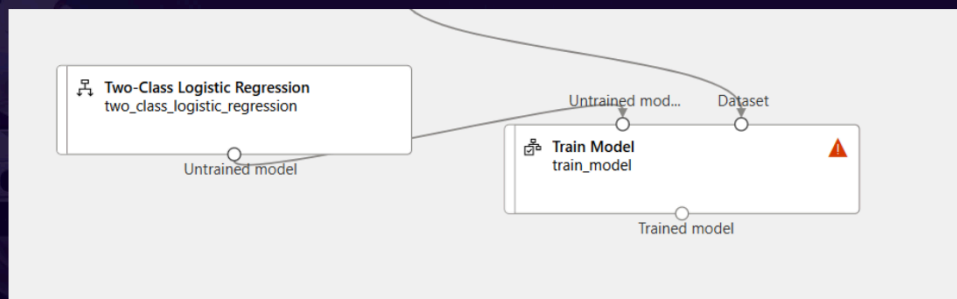


- Our aim is to build a machine learning model. This model will be used to classify the income of individuals into 2 classes -  $\leq 50K$  or  $> 50K$
- The first step was to get the data set and split the set into a training and test set. The training set would have 70% of the data and the test set would have 30% of the data.

8.

## Train Model

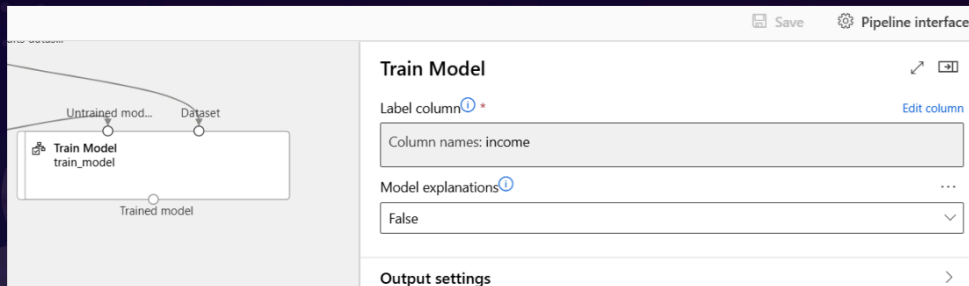
- Next we need to add the component to train the model. We will add as the input the 70% of the split data phase.
- We will add the algorithm which will be used to train the machine learning model.



9.

# Train Model

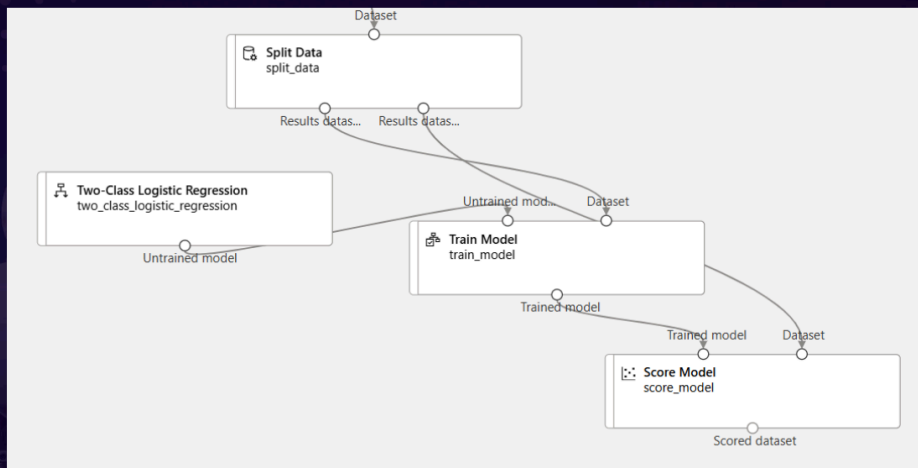
- We will also ensure to train the model based on the income column. This is our label.



10

# Score Model

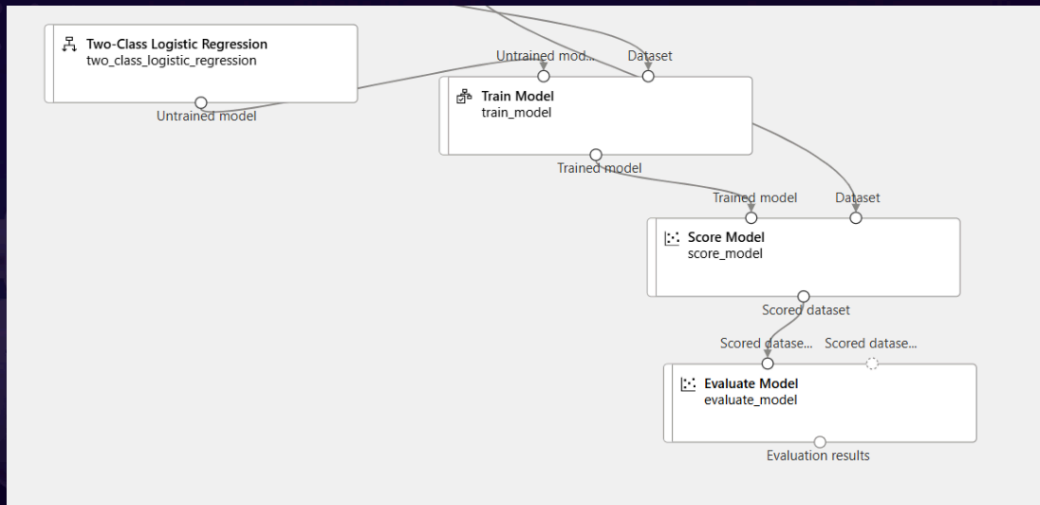
- We can now score the model based on the 30% test data. This will help us understand how efficient the model is when it comes to predicting the income value.



11.

# Evaluate Model

- Based on the scoring and the test data, we can finally evaluate the model.



12.

## Two-class Logistic regression

### Confusion Matrix

This table gives us a summary of the results when it comes to what the model was supposed to predict and the actual prediction.

		Actual	
		>50K	<=50K
Predicted	>50K	1 389	565
	<=50K	926	6 888

True Positives	False Positives
False Negatives	True Negatives

Accuracy - This tells how often the classifier is right in predicting results.

Accuracy 0.847  
 Precision 0.711  
 Recall 0.6  
 F1 Score 0.651  
 AUC 0.901

True Positives + True Negatives

True Positives + True Negatives + False Positives + False Negatives

Precision - This tells to what extend does the model accurately predict results.