AzureML로 시작하는 Machine Learning

Microsoft Student Partner

KAIST 정태영

Session

- Session 1 : Introduction to Introduction to ML

  머신러닝의 기초의 기초에 대해 배웁니다.

- Session 2 : Azure Machine Learning 맛보기

  Azure ML studio를 통해 붓꽃을 분류하는
  머신 러닝 프로그램을 만들어봅니다.

Azure ML?

- Microsoft에서 제공하는 ML 솔루션 클라우드
- 드래그&드롭
- Supervised / Unsupervised Learning 지원

예제 실습

- 붓꽃의 특성값을 이용해 종 분류하기
- 데이터의 구성은 다음과 같음

- 1) 꽃받침의 길이
  2) 꽃받침의 너비
  3) 꽃잎의 길이
  4) 꽃잎의 너비

예제 실습

Iris setosa

Iris versicolor

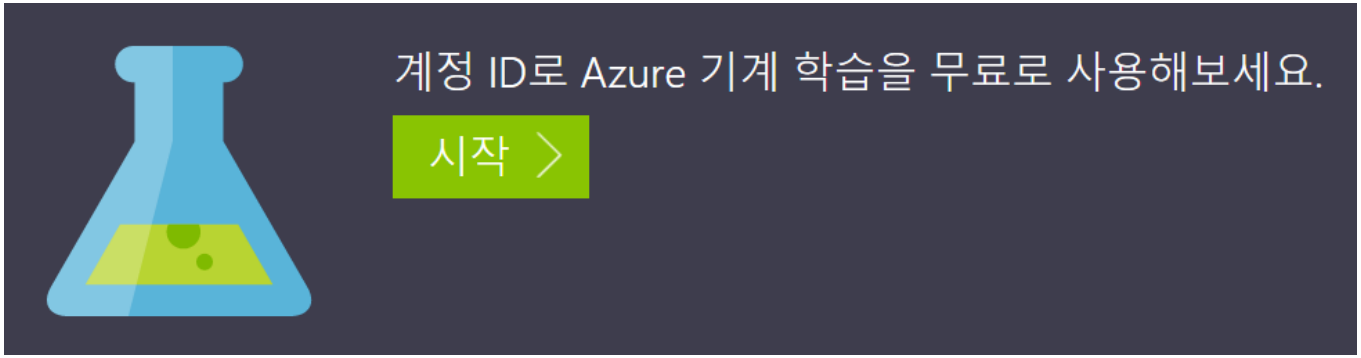
Iris virginica

Step 1 : 접속

- Azureml.com 접속
- 8-hour trial

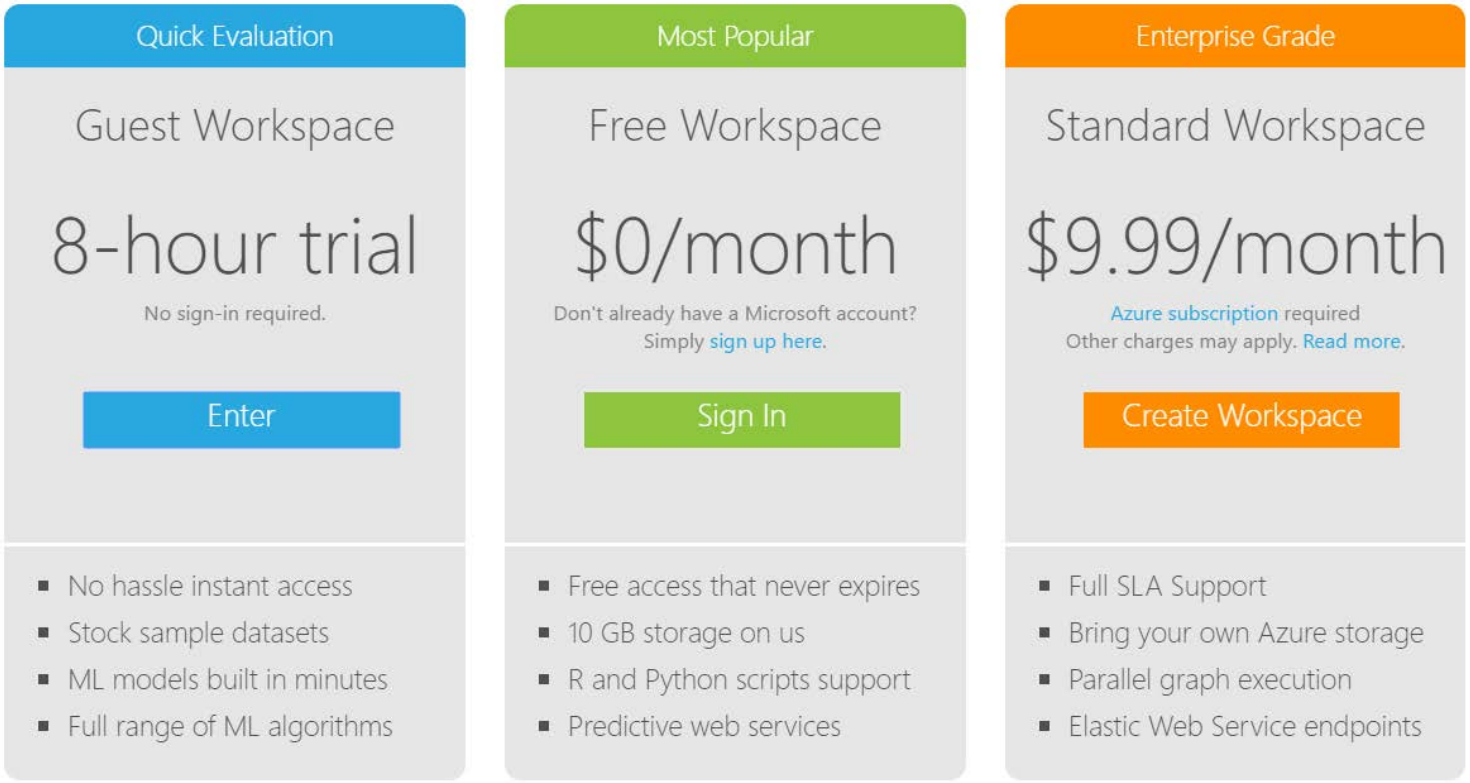
Step 1 : 접속

- Azureml.com 접속
- 8-hour trial



Step 1 : 접속

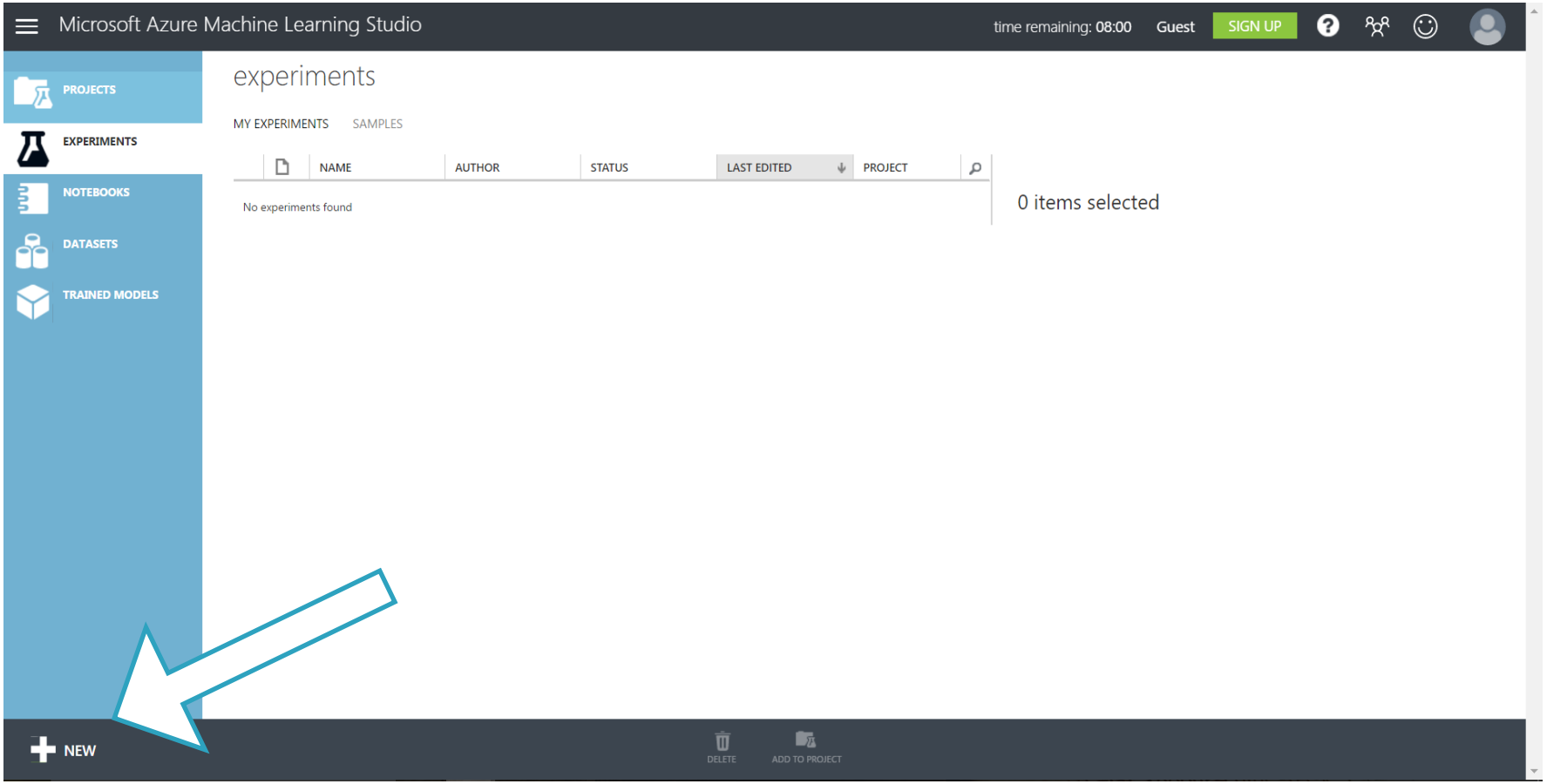
- Azureml.com 접속
- 8-hour trial



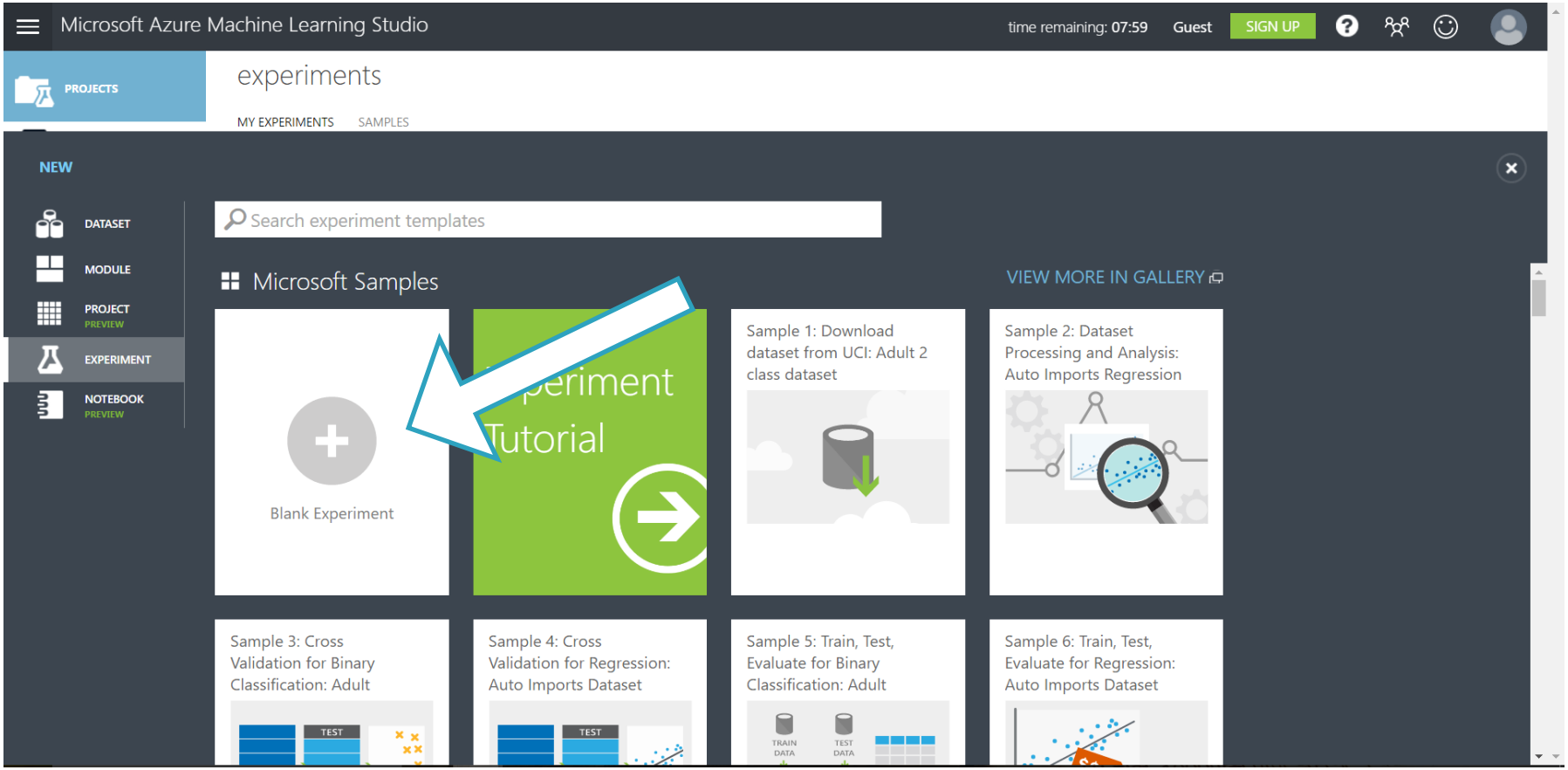
Step 2 : 실험 생성

- 왼쪽 아래의 + new 버튼 클릭
- Blank Experiment

Step 2 : 실험 생성


Microsoft Azure Machine Learning Studio
time remaining: 08:00
Guest
SIGN UP
PROJECTS
EXPERIMENTS
NOTEBOOKS
DATASETS
TRAINED MODELS
experiments
MY EXPERIMENTS
SAMPLES
NAME
AUTHOR
STATUS
LAST EDITED
PROJECT
No experiments found
0 items selected
NEW
DELETE
ADD TO PROJECT

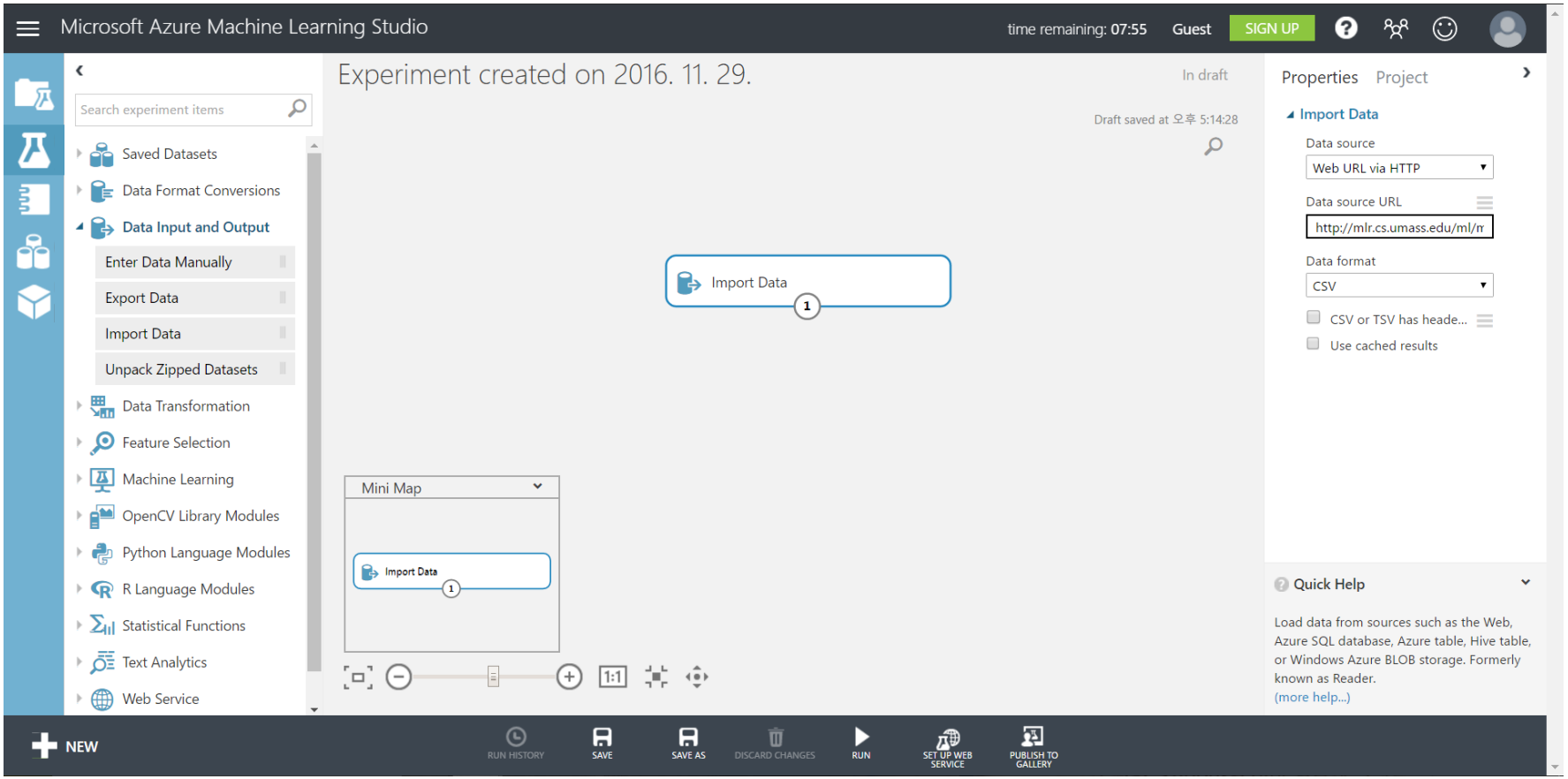
Step 2 : 실험 생성


Microsoft Azure Machine Learning Studio
time remaining: 07:59
Guest
SIGN UP
PROJECTS
experiments
MY EXPERIMENTS
SAMPLES
NEW
DATASET
MODULE
PROJECT
PREVIEW
EXPERIMENT
NOTEBOOK
PREVIEW
Search experiment templates
Microsoft Samples
VIEW MORE IN GALLERY
Blank Experiment
Tutorial
Sample 1: Download dataset from UCI: Adult 2 class dataset
Sample 2: Dataset Processing and Analysis: Auto Imports Regression
Sample 3: Cross Validation for Binary Classification: Adult
Sample 4: Cross Validation for Regression: Auto Imports Dataset
Sample 5: Train, Test, Evaluate for Binary Classification: Adult
Sample 6: Train, Test, Evaluate for Regression: Auto Imports Dataset
TEST
TRAIN DATA
TEST DATA

Step 3 : 데이터 로딩

- 왼쪽의 Data Input and Output에서
  Import data 모듈 드래그
- Data source : Web URL via HTTP
- URL : http://mlr.cs.umass.edu/ml/machine-learning-databases/iris/iris.data

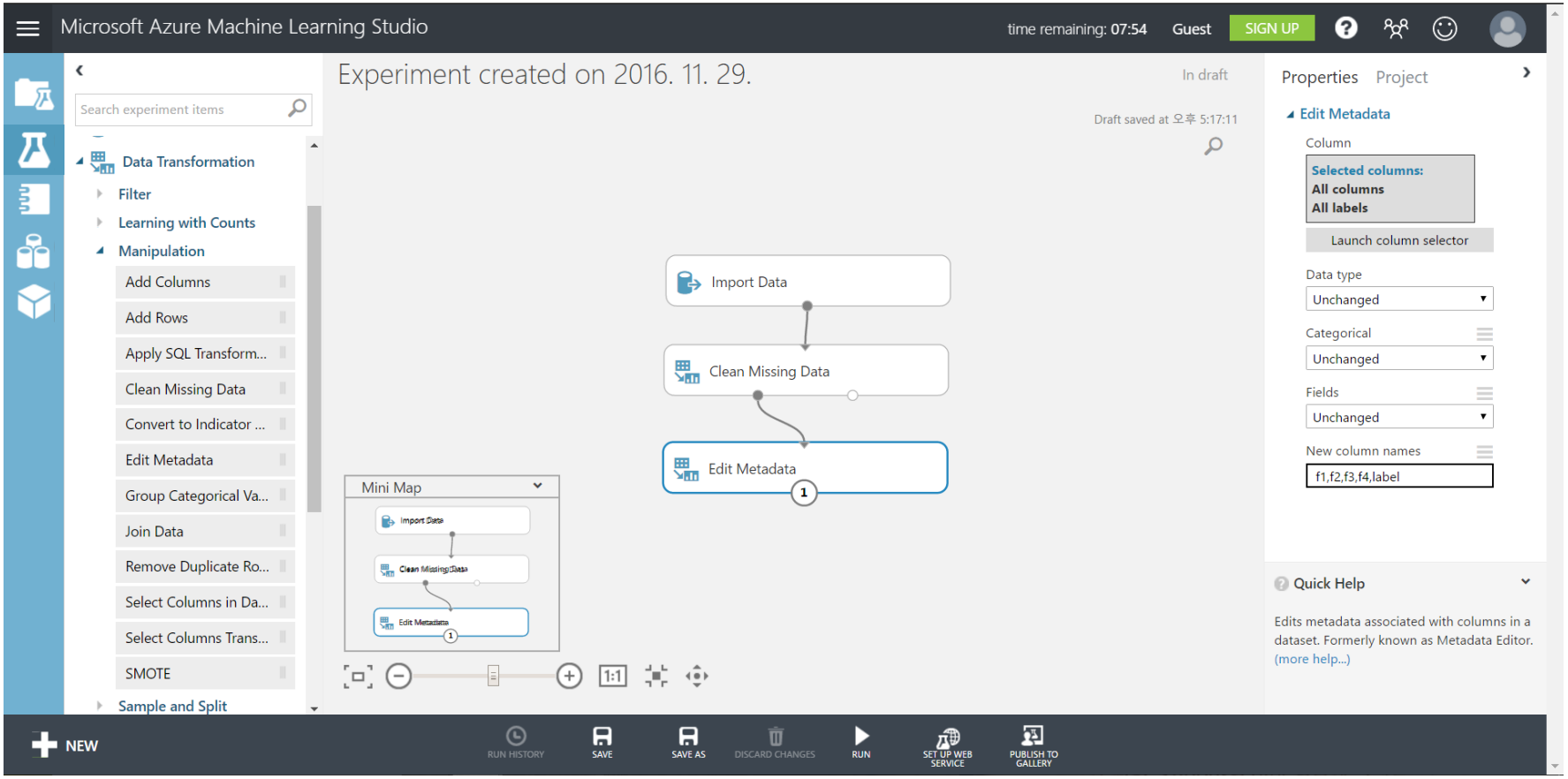
Step 3 : 데이터 로딩


Microsoft Azure Machine Learning Studio
time remaining: 07:55
Guest
SIGN UP
Experiment created on 2016. 11. 29.
In draft
Draft saved at 오후 5:14:28
Search experiment items
Saved Datasets
Data Format Conversions
Data Input and Output
Enter Data Manually
Export Data
Import Data
Unpack Zipped Datasets
Data Transformation
Feature Selection
Machine Learning
OpenCV Library Modules
Python Language Modules
R Language Modules
Statistical Functions
Text Analytics
Web Service
Import Data
1
Mini Map
Import Data
1
1:1
Properties
Project
Import Data
Data source
Web URL via HTTP
Data source URL
http://mlr.cs.umass.edu/ml/r
Data format
CSV
CSV or TSV has heade...
Use cached results
Quick Help
Load data from sources such as the Web, Azure SQL database, Azure table, Hive table, or Windows Azure BLOB storage. Formerly known as Reader.
(more help...)
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY

Step 4 : 데이터 전처리

- 왼쪽의 Data Transformation / Manipulation에서 Clean Missing Data, Edit Metadata 모듈 드래그
- Edit Metadata의 Launch column selector를 클릭한 후 All columns, All labels 선택
- New column names : f1,f2,f3,f4,label

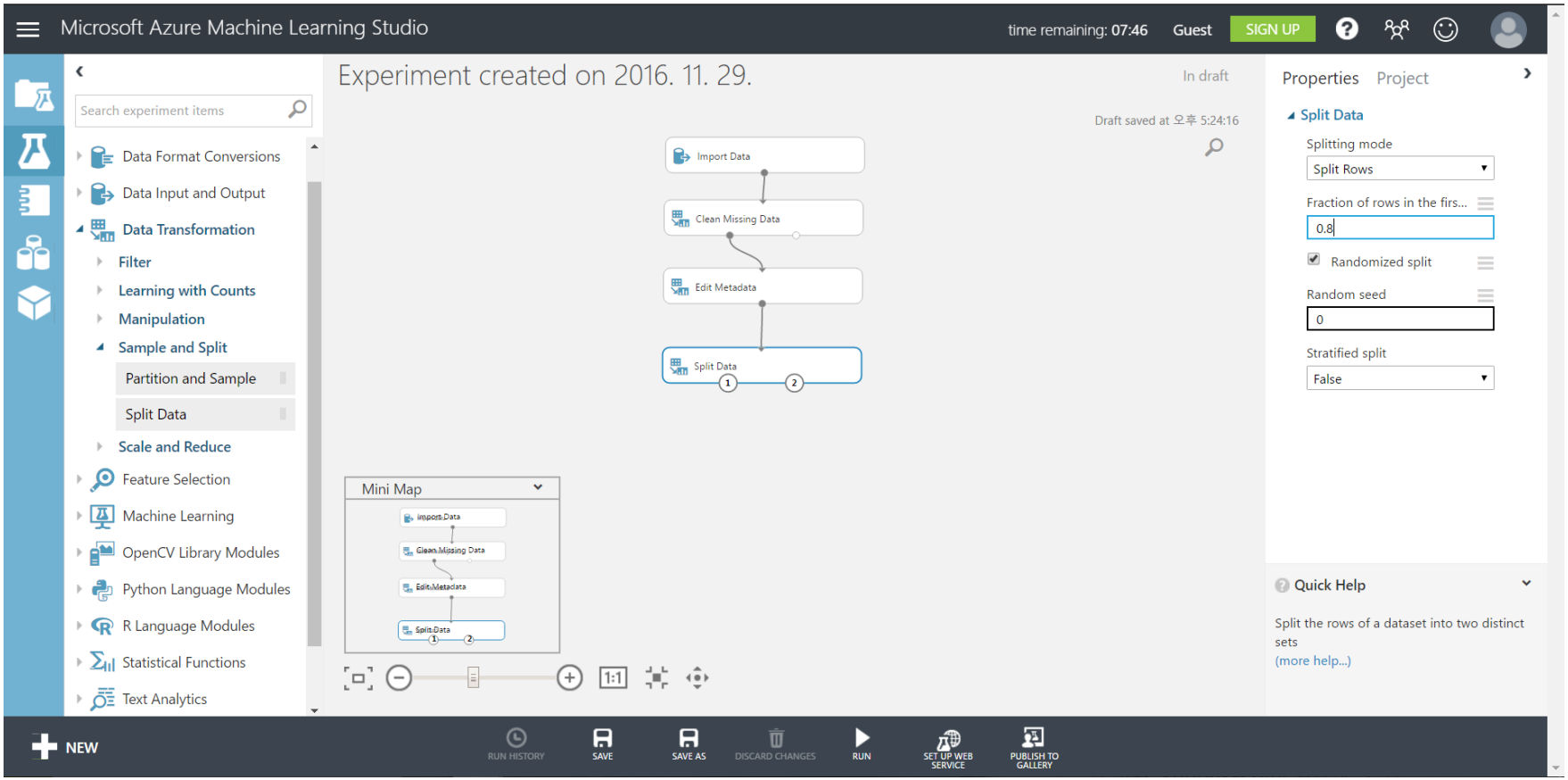
Step 4 : 데이터 전처리



Step 5 : Train/Test 나누기

- 왼쪽의 Data Transformation / Sample and Split 에서 Split data 모듈 드래그
- Fraction 0.8으로 설정

Step 5 : Train/Test 나누기



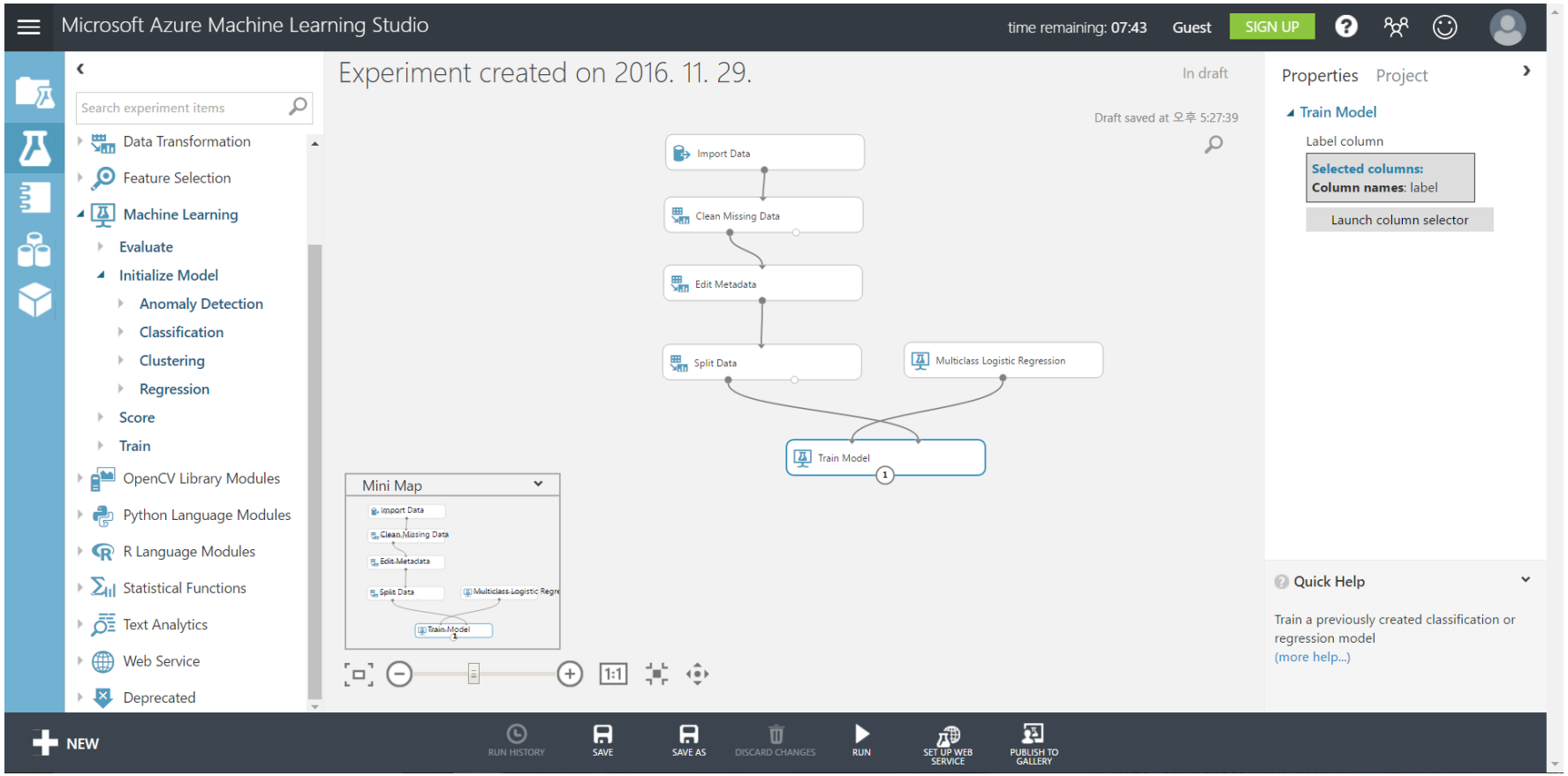
Step 6 : Model 생성

- 왼쪽의 Machine Learning / Train에서 Train Model 모듈 드래그, Machine Learning / Initialize Model / Classification에서 Multiclass Logistic Regression 모듈 드래그 후 연결

Step 6 : Model 생성

- 왼쪽의 Machine Learning / Train에서 Train Model 모듈 드래그, Machine Learning / Initialize Model / Classification에서 Multiclass Logistic Regression 모듈 드래그 후 연결
- Train Model의 launch column selector를 누르고 column names에 label 입력

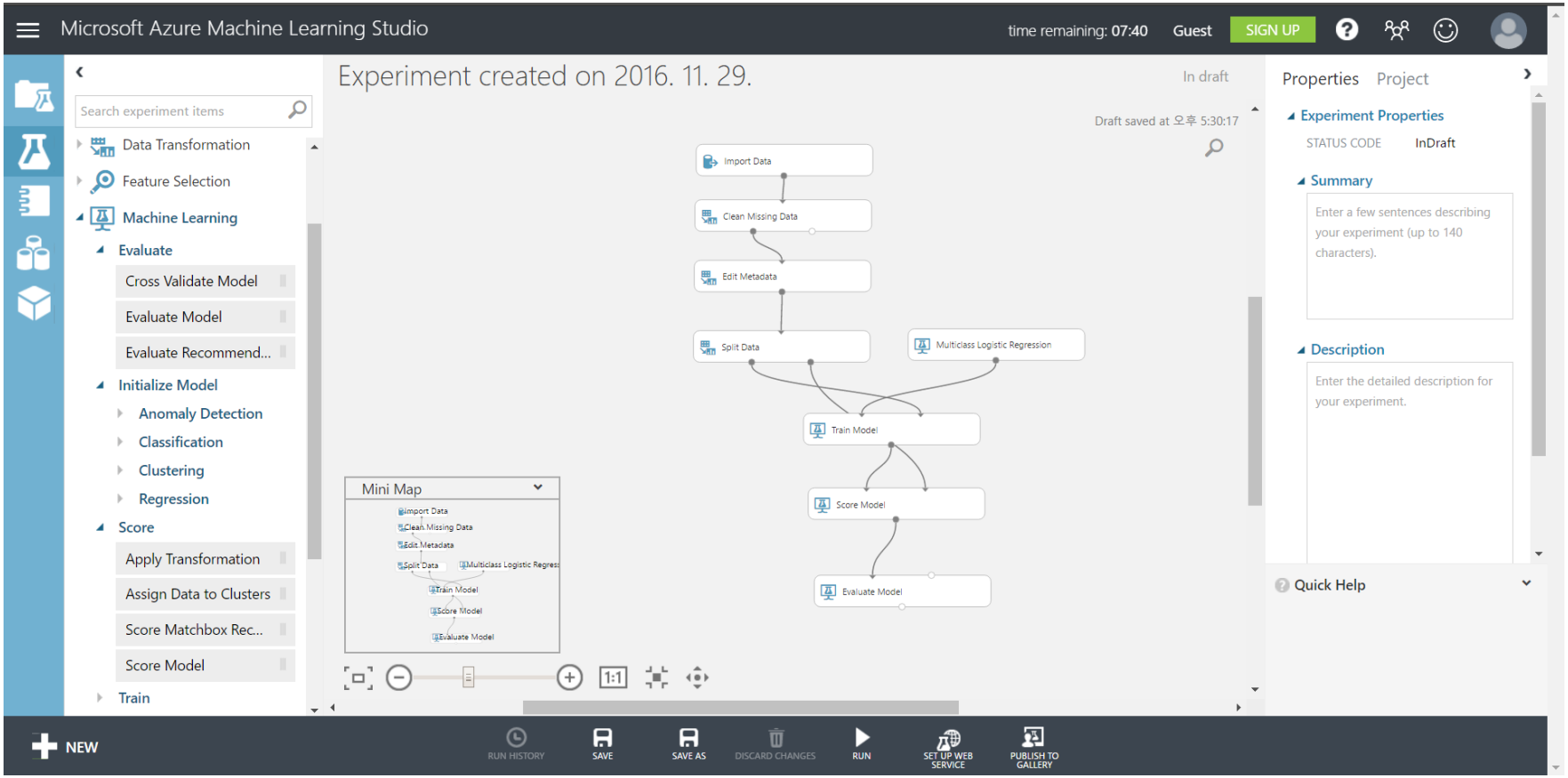
Step 6 : Model 생성



Step 7 : Model 평가

- 왼쪽의 Machine Learning / Score 에서 Score Model 모듈 드래그, Machine Learning / Evaluate 에서 Evaluate Model 모듈 드래그 후 연결
- Train Model의 launch column selector를 누르고 column names에 label 입력

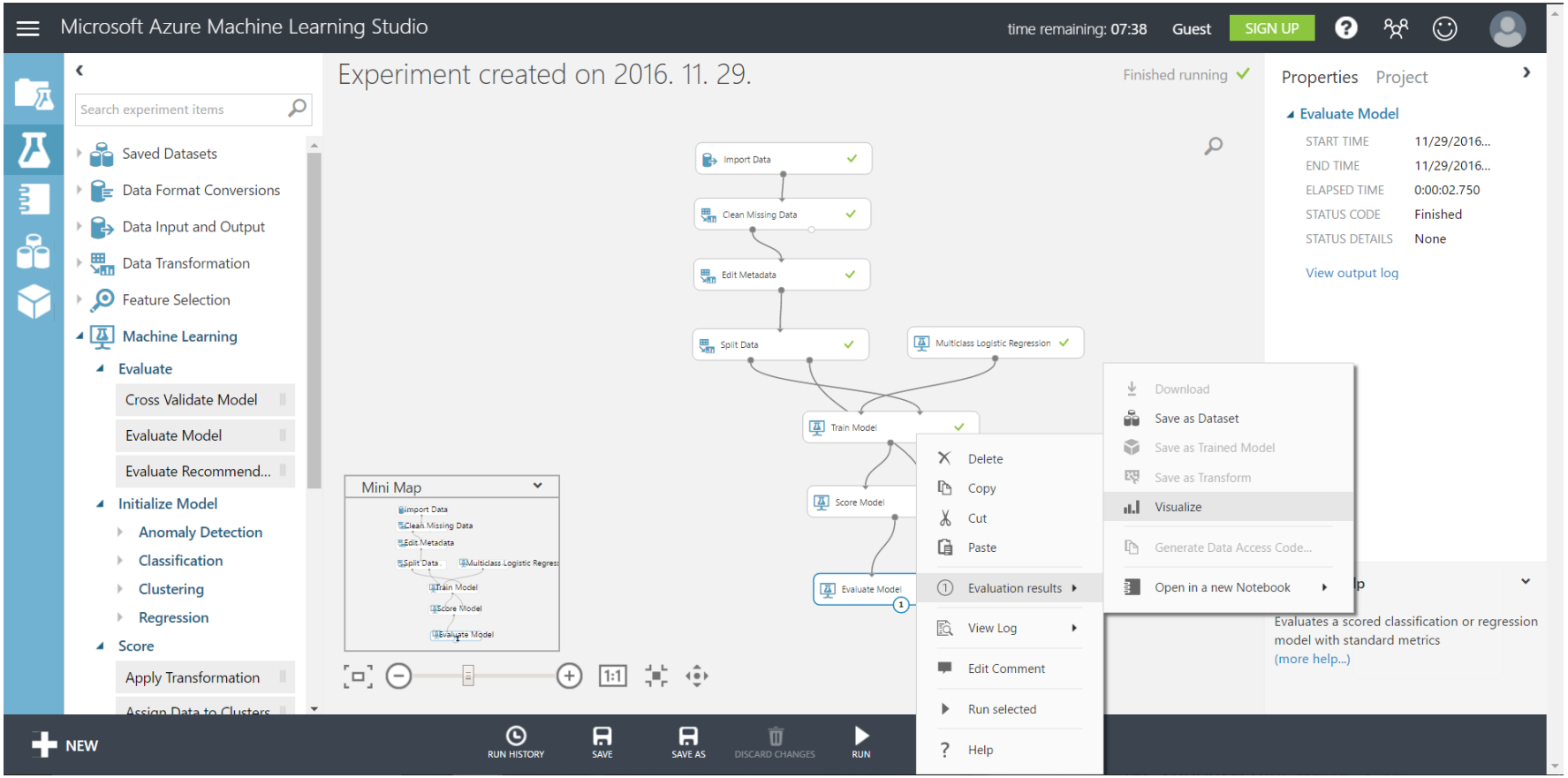
Step 7 : Model 평가



Step 8 : 실행

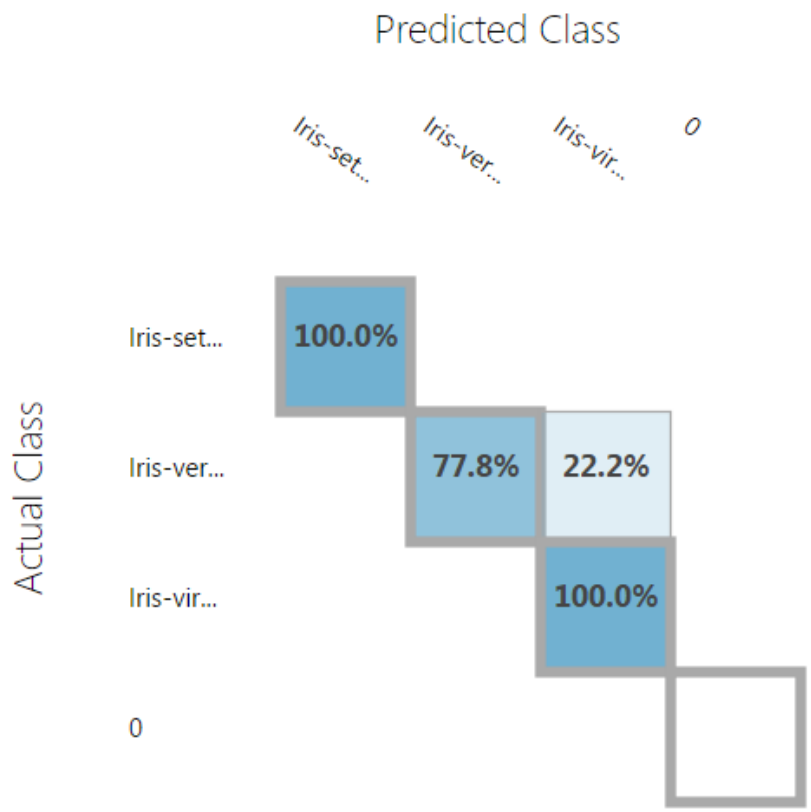
- 오른쪽 하단의 RUN 클릭
- 작업이 완료되면 Evaluate Model을 오른쪽 클릭, Evaluation results/Visualize 옵션 클릭

Step 8 : 실행


Microsoft Azure Machine Learning Studio
time remaining: 07:38
Guest
SIGN UP
Experiment created on 2016. 11. 29.
Finished running
Search experiment items
Saved Datasets
Data Format Conversions
Data Input and Output
Data Transformation
Feature Selection
Machine Learning
Evaluate
Cross Validate Model
Evaluate Model
Evaluate Recommend...
Initialize Model
Anomaly Detection
Classification
Clustering
Regression
Score
Apply Transformation
Import Data
Clean Missing Data
Edit Metadata
Split Data
Multiclass Logistic Regression
Train Model
Score Model
Evaluate Model
Mini Map
Properties
Project
Evaluate Model
START TIME 11/29/2016...
END TIME 11/29/2016...
ELAPSED TIME 0:00:02.750
STATUS CODE Finished
STATUS DETAILS None
View output log
Delete
Copy
Cut
Paste
Evaluation results
View Log
Edit Comment
Run selected
Help
Download
Save as Dataset
Save as Trained Model
Save as Transform
Visualize
Generate Data Access Code...
Open in a new Notebook
Evaluates a scored classification or regression model with standard metrics
(more help...)
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN

Step 8 : 실행

Confusion Matrix



감사합니다