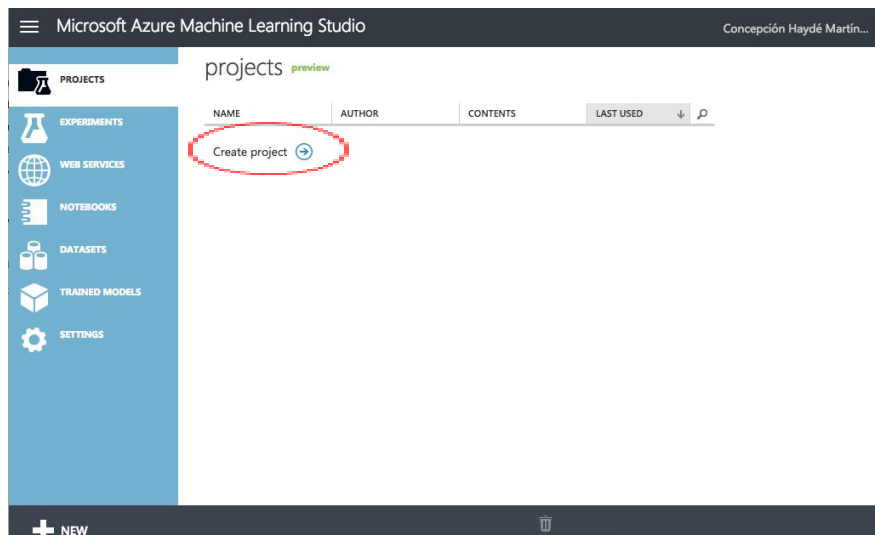


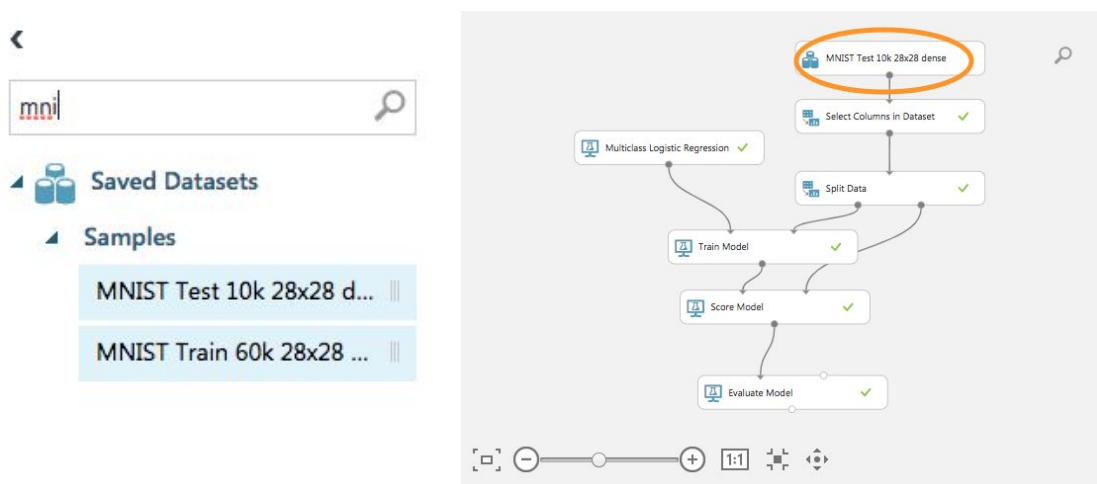
Ejemplo AzureML

Iniciamos un nuevo experimento:

- Hacemos click en “create project” o “+ new”



En la barra del lado izquierdo hacemos click en “Saved Datasets” luego en “Saved” y ahí para este ejemplo usaremos “MNIST Test 10k 28x28” que es un dataset de dígitos dibujados a mano de 28 x 28 píxeles a blanco y negro. Hacemos click sobre el mismo y lo arrastramos al espacio de trabajo vacío.



Una vez más en la barra izquierda buscamos “Select columns in Dataset” y lo arrastramos al espacio de trabajo, abajo del Dataset, ahí conectamos los 2 bloques y damos click en el último que agregamos.

Ahí aparecerá una barra a la derecha en la cual haremos click en “Launch Column Selector” para que nos muestre las columnas a elegir.

Properties Project >

▲ Select Columns in Dataset

Select columns

Selected columns:
Column names:
f0,f1,f2,f3,f4,f5,f6,f7,f8,f9,f10,

Launch column selector

START TIME 12/17/2018...

END TIME 12/17/2018...

ELAPSED TIME 0:00:00.000

STATUS CODE Finished

STATUS DETAILS Task output was present in output cache

Ahí seleccionaremos las columnas del dataset que son relevantes para entrenar el algoritmo, en este caso necesito seleccionar todas las columnas, porque cada columna representa un pixel y en este caso que queremos hacer reconocimiento de imagen TODOS los píxeles importan, además de también seleccionar la columna “Label” que es la que dice cuál número está dibujado en la imagen.

Select columns

BY NAME WITH RULES

AVAILABLE COLUMNS All Types search columns

Label
f0
f1
f2
f3
f4
f5
f6
f7
f8
f9
f10
f11
f12
f13

785 columns available

SELECTED COLUMNS All Types search columns

0 columns selected

Select columns

BY NAME WITH RULES

AVAILABLE COLUMNS All Types search columns

0 columns available

SELECTED COLUMNS All Types search columns

Label
f0
f1
f2
f3
f4
f5
f6
f7
f8
f9
f10
f11
f12
f13

785 columns selected

Una vez más regresamos a la barra izquierda, buscamos “Split Dataset”, lo arrastramos al espacio de trabajo, lo conectamos con el previo y le damos click. En la barra de la derecha es donde tenemos que poner un número que va a indicar el porcentaje del Dataset que será utilizado para training y el porcentaje que será utilizado para testing, yo en este ejemplo uso 0.7 lo cual significa que el 70% del Dataset es para training y el 30% es para Testing.

Properties Project >

▲ Split Data

Splitting mode
Split Rows

Fraction of rows in the first...
0.7

☒ Randomized split

Random seed
0

Stratified split
False

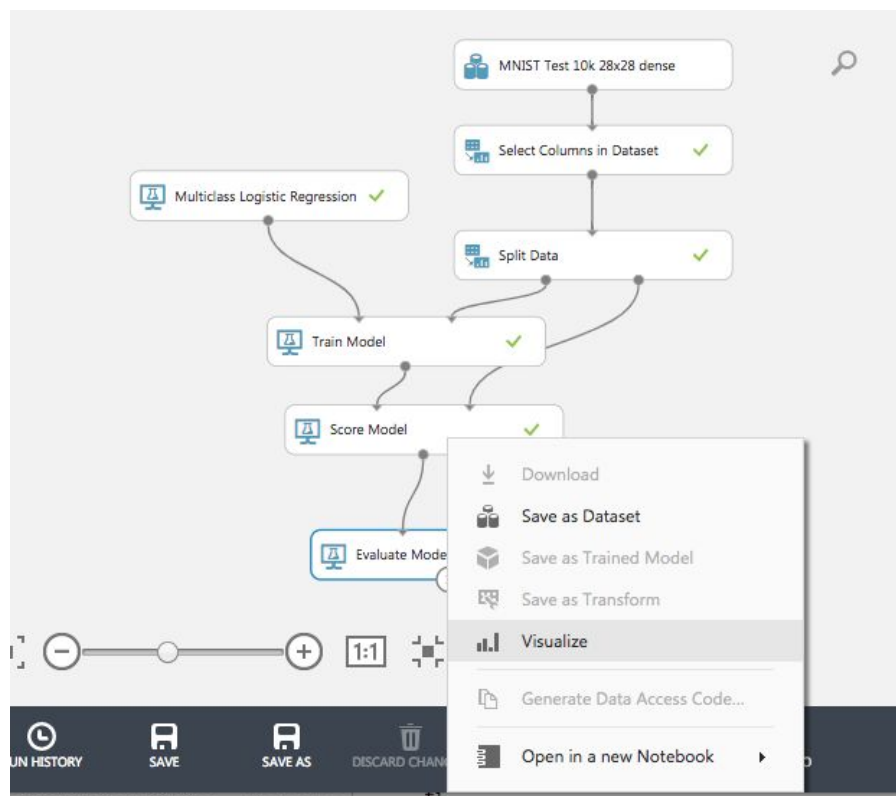
START TIME 12/17/2018...

END TIME 12/17/2018...

Ahora regresamos a la barra izquierda, buscamos “Train Model”, lo arrastramos al espacio de trabajo, lo conectamos a “Split Dataset” y ahora buscamos de nuevo en la barra izquierda “Multiclass Logistic Regression” qué es el algoritmo que usaremos para este modelo, lo arrastramos al espacio de trabajo y lo conectamos del otro lado en la parte arriba que sobra de “Train Model”.

De nuevo buscamos en la barra izquierda “Score Model” lo arrastramos al espacio de trabajo y lo conectamos con el que sobra de “Split Dataset” y el que sale de “Train Model”, este lo que hará es hacer predicciones en base al entrenamiento, del 30% de los Datos que son de prueba.

Una vez más buscamos en la barra izquierda “Evaluate Model” lo arrastramos al espacio de trabajo y lo conectamos con el que sale de “Score Model”, este lo que hará es revisar cuántas predicciones fueron correctas y nos dirá el porcentaje de precisión de nuestro modelo, para revisar estos datos daremos click derecho en el circulito inferior del bloque y daremos click en “Visualize”



Ahí veremos los porcentajes y la tabla de verdad.

Metrics

Overall accuracy	0.906667
Average accuracy	0.981333
Micro-averaged precision	0.906667
Macro-averaged precision	0.906156
Micro-averaged recall	0.906667
Macro-averaged recall	0.905975

Confusion Matrix

Listo :)