

Machine Learning NanoProject

Abasiakama Zoe Ndedde

DATA EXPLORATION: I started by importing the Drugtest.csv dataset into Microsoft Machine Learning Studio and selecting Age, Gender, Bp and BMI as the columns needed.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free... ?

Drug Group Prediction

In draft

Split data

Drugtest.csv

Select Columns in Dataset

1

Properties Project

Select Columns in Dataset

Select columns

Selected columns:
All columns
Column names:
Age, Gender, BP, BMI

Launch column selector

Quick Help

Selects columns to include or exclude from a dataset in an operation. Formerly known as Project Columns.
(more help...)

+ NEW

RUN HISTORY SAVE SAVE AS DISCARD CHANGES RUN SET UP WEB SERVICE PUBLISH TO GALLERY

SPLIT DATASET: I split the dataset into training and testing sets. The training set is used to train the machine learning model, while the testing set is used to evaluate its performance on unseen data. I typically use a 70-30 or 80-20 ratio for the training and testing split.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free... ?

Split data

Saved Datasets

Samples

Restaurant ratings

Data Transformation

Sample and Split

Split Data

Drug Group Prediction

In draft

Drugtest.csv

Select Columns in Dataset

Split Data

12

Splitting mode

Split Rows

Fraction of rows in the f...

0.7

☒ Randomized split

Random seed

0

Stratified split

False

Quick Help

Split the rows of a dataset into two distinct sets

(more help...)

+

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

MODEL SELECTION: Based on the nature of the prediction task and the available dataset, I selected Multiclass Neural Network an appropriate machine learning algorithm for building the model. The selection is based on factors like interpretability, accuracy, and the size of the dataset.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free...

?

Search experiment items

Saved Datasets

My Datasets

Customers.csv

Drugtest.csv

Lemonade.csv

Samples

Adult Census Income...

Airport Codes Dataset

Automobile price dat...

Bike Rental UCI dataset

Bill Gates RGB Image

Blood donation data

Drug Group Prediction

Finished running ✓

Drugtest.csv

Select Columns in Dataset ✓

Multiclass Neural Network ✓

Split Data ✓

1:1

Properties

Project

Experiment Properties

START TIME9/5/202...

END TIME9/5/202...

STATUS CODEFinished

STATUS DETAILSNone

Prior Run

Summary

Enter a few sentences describing your experiment (up to 140 characters).

Quick Help

+

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

TRAIN MODEL: A trained model in machine learning is used to make predictions or classifications based on input data, allowing it to generalize patterns and provide valuable insights or automated decisions.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free...

?

Search experiment items

Saved Datasets

My Datasets

Customers.csv

Drugtest.csv

Lemonade.csv

Samples

Adult Census Income...

Airport Codes Dataset

Automobile price dat...

Bike Rental UCI dataset

Bill Gates RGB Image

Blood donation data

Drug Group Prediction

Finished running

Drugtest.csv

Select Columns in Dataset

Split Data

Multiclass Neural Network

Train Model

+

+

1:1

Properties

Project

Experiment Properties

START TIME9/5/202...

END TIME9/5/202...

STATUS CODEFinished

STATUS DETAILSNone

Prior Run

Summary

Enter a few sentences describing your experiment (up to 140 characters).

Quick Help

+

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

SCORE MODEL: I Score the model to evaluate its performance and assess how well it predicts or classifies data, typically using metrics like accuracy, precision, recall, or F1-score. This helps gauge the model's effectiveness and suitability for a given task.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free...

Search experiment items

Saved Datasets

My Datasets

Customers.csv

Drugtest.csv

Lemonade.csv

Samples

Adult Census Income...

Airport Codes Dataset

Automobile price dat...

Bike Rental UCI dataset

Bill Gates RGB Image

Blood donation data

Drug Group Prediction

Finished running

Drugtest.csv

Select Columns in Dataset

Split Data

Multiclass Neural Network

Train Model

Score Model

PropertiesProject

Experiment Properties

START TIME

9/5/202...

END TIME

9/5/202...

STATUS CODE

Finished

STATUS DETAILS

None

Prior Run

Summary

Enter a few sentences describing your experiment (up to 140 characters).

Quick Help

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

https://studio.azureml.net

EVALUATE MODEL: I Evaluated the model to assess its performance and generalization ability on unseen data and determine its effectiveness, reliability, and whether it meets the desired objectives or criteria.

The screenshot displays the Microsoft Machine Learning Studio (classic) interface. The main workspace shows a completed training experiment titled "Drug Group Prediction" with a status of "Finished running" and a green checkmark. The experiment flowchart includes the following modules: "Drugtest.csv" (data source), "Select Columns in Dataset" (data transformation), "Split Data" (data partitioning), "Multiclass Neural Network" (model training), "Train Model" (model training), "Score Model" (model evaluation), and "Evaluate Model" (model evaluation). All modules are marked with a green checkmark, indicating successful execution. The left sidebar contains a search bar and a list of experiment items: Saved Datasets, Trained Models, Transforms, Data Format Conversions, Data Input and Output, Data Transformation, Feature Selection, Machine Learning, OpenCV Library Modules, Python Language Modules, R Language Modules, and Statistical Functions. The right sidebar shows the "Properties" pane with "Experiment Properties" (Start Time: 9/5/202..., End Time: 9/5/202..., Status Code: Finished, Status Details: None) and a "Summary" section with a text area for describing the experiment. The bottom toolbar includes icons for "NEW", "RUN HISTORY", "SAVE", "SAVE AS", "DISCARD CHANGES", "RUN", "SET UP WEB SERVICE", and "PUBLISH TO GALLERY".

Here are results from visualizing the data under score model.

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free...

Drug Group Prediction

Finished running ✓ Properties Project

Drug Group Prediction > Score Model > Scored dataset

rows 180 columns 11

Scored Probabilities for Class "Group A"	Scored Probabilities for Class "Group B"	Scored Probabilities for Class "Group C"	Scored Probabilities for Class "Group D"	Scored Probabilities for Class "Group E"	Scored Labels
0.24478	0.000508	0.000001	0.982755	0.000001	Group D
0.000001	0.427141	0.336102	0.000004	0.121937	Group B
0	0.165382	0.889928	0	0.000001	Group C
0.000568	0.974839	0.005773	0	0	Group B

▲ Statistics

▲ Visualizations

To view, select a column in the table.

+ NEW RUN HISTORY SAVE SAVE AS DISCARD CHANGES RUN SET UP WEB SERVICE PUBLISH TO GALLERY

Microsoft Machine Learning Studio (classic)

Abasiakama Ndedde-Free...

Drug Group Prediction

Finished running ✓ Properties Project

Drug Group Prediction > Score Model > Scored dataset

rows 180 columns 11

75	M	L	24	Group B	0.000038	0.6579
74	F	N	30	Group C	0	0.0673
66	F	L	10	Group D	0.05227	0.0000
89	F	N	14	Group A	0.978842	0.0361
77	M	N	50	Group C	0	0
74	F	H	28	Group C	0.000001	0.1448
84	F	N	38	Group C	0	0.0013
98	M	H	13	Group	0.973355	0.1215

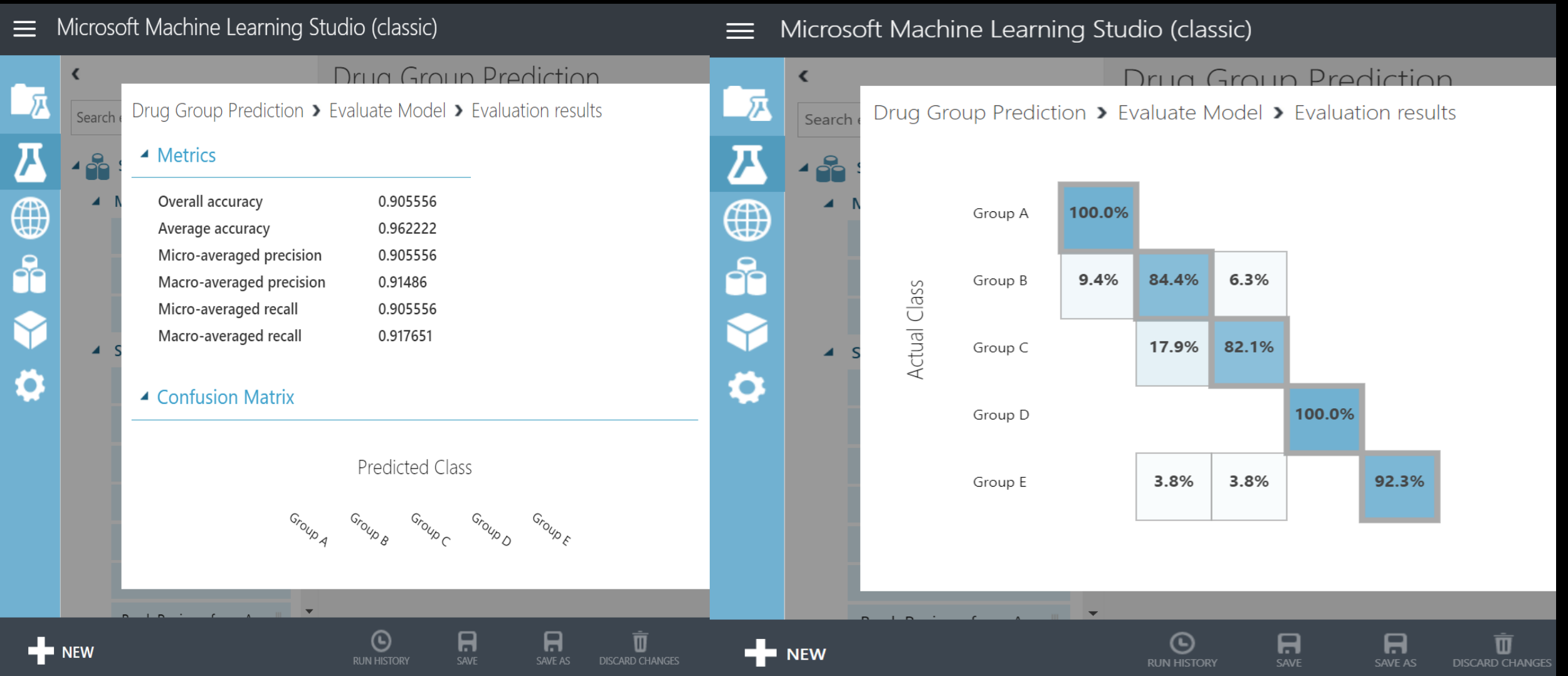
▲ Statistics

▲ Visualizations

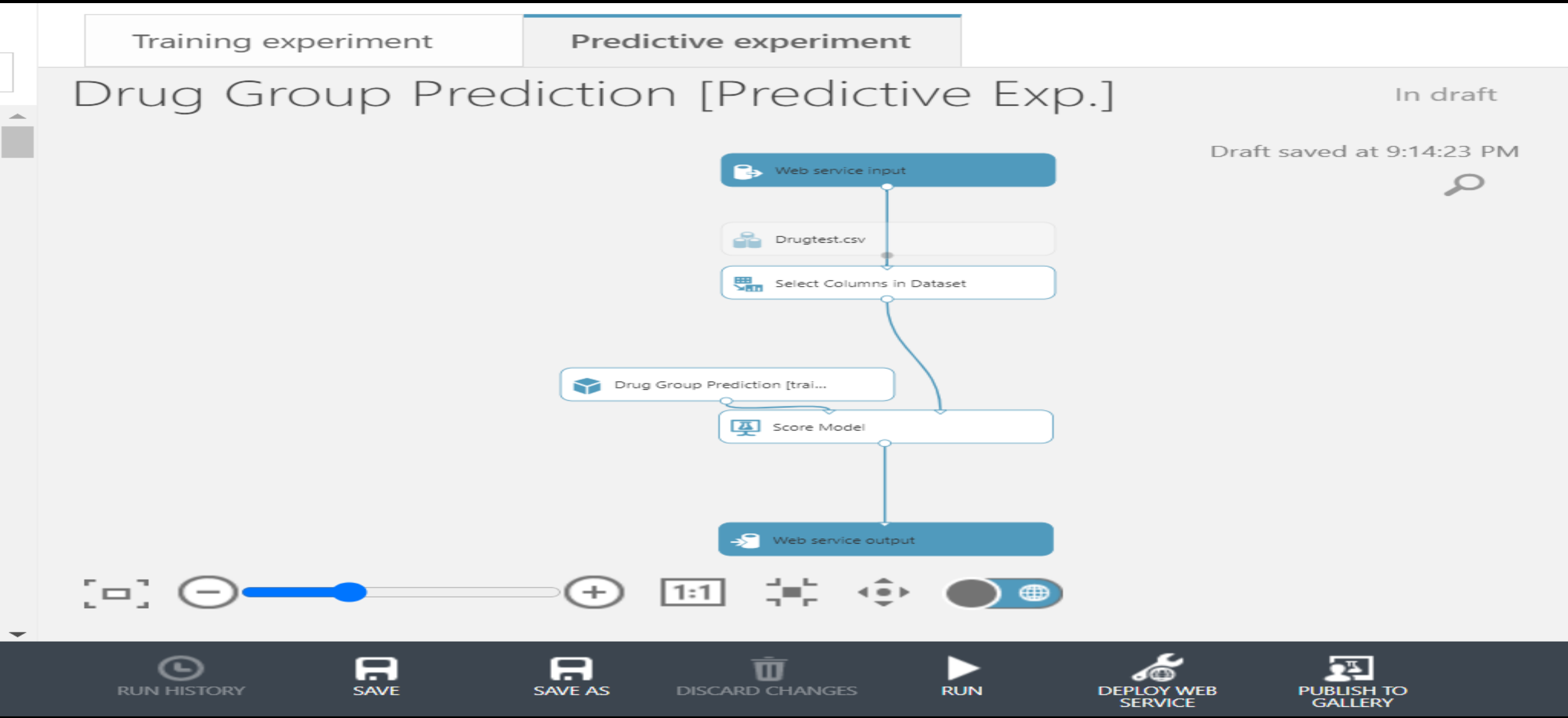
To view, select a column in the table.

+ NEW RUN HISTORY SAVE SAVE AS DISCARD CHANGES RUN SET UP WEB SERVICE PUBLISH TO GALLERY

Here are results from visualizing the data under evaluate model.



MODEL DEPLOYMENT: Once I achieved a satisfactory performance, I deployed the model to make predictions on new, unseen data. This step involves creating an interface or integrating the model into an existing system for practical usage. It ensures the model is accessible and can be utilized for real-time predictions.



I downloaded the first excel database.

drug group prediction [predictive exp.]

DASHBOARD CONFIGURATION

General [New Web Services Experience](#) **preview**

Published experiment

[View snapshot](#) [View latest](#)




Description

No description provided for this web service.

API key

To78XqD3rtnJWdK7EQpZ2W2kYT8fpKIFxlyuwrQ9aS2UjOuocDIVv3sAmOUpYmBlZeef1G0Pr8Yd+AMCCukzdw==

Default Endpoint

API HELP PAGE	TEST	APPS	LAST UPDATED
REQUEST/RESPONSE	Test Test preview	 Excel 2013 or later  Excel 2010 or earlier workbook	9/5/2023 9:18:15 PM
BATCH EXECUTION	Test preview	 Excel 2013 or later workbook	9/5/2023 9:18:15 PM

I clicked on the “Drug group prediction” link

Drug Group Prediction [Predictive Exp.]-09_05_2023 20_40_04 [Read-Only] -...

Search

Zoe Ndedde ZN

File Home Insert Page Layout Formulas Data Review View Help Power Pivot Share

Clipboard Font Alignment Number Styles Cells Editing

A1

Sheet1

Ready Accessibility: Good to go

Azure Machine Learning

Web Services

Drug Group Prediction [Predictive Exp.]

+ Add Web Service

☐ Auto-predict **Predict All**

Help Privacy Statement

FINAL OUTPUT

Drug Group Prediction (Predictive Exp.)-09_05_2023 20_40_04 -...

Search

Zoe Ndedde ZN

File Home Insert Page Layout Formulas Data Review View Help Power Pivot

Clipboard Font Alignment Number Styles Cells Editing

GET GENUINE OFFICE Your license isn't genuine, and you may be a victim of software counterfeiting. Avoid interruption and keep your files safe with genuine Office today. Get genuine Office Learn more

G2

Group AgGroup E

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Group Ag	GGroup E	Group B	Group B	Group Dr								
2	96 F	H		32 Group C			Group Ag	GGroup E	Group B P	Group B M	Group Dr	Scored Pr	Scored Pr
3	75 F	H		18 Group B			96 F	H	32 Group C	2.13E-07	0.609174		
4	78 M	N		16 Group B			75 F	H	18 Group B	0.20109	0.503059		
5	98 F	H		15 Group A			78 M	N	16 Group B	0.828491	0.282842		
6	43 M	L		21 Group E			98 F	H	15 Group A	0.886277	0.239221		
7							43 M	L	21 Group E	5.09E-07	2.77E-06		
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													

Sheet1

Azure Machine Learning

Input: input1

Sheet1!A1:E6

My data has headers

Use sample data

Output: output1

Sheet1!G2

Include headers

Predicting will override existing values. This can't be undone.

Got it!

Predict

Auto-predict

3. ERRORS

Help Privacy Statement

FINAL OUTPUT

Drug Group Prediction [Predictive Exp.]-09_05_2023 20_40_04 -...

Search

Zoe Ndedde ZN

File Home Insert Page Layout Formulas Data Review View Help Power Pivot

Share

Paste

Clipboard

Font

Calibri 11 A A

B I U

Align

Wrap Text

Merge & Center

Alignment

Number

General

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

Cells

Editing

Sort & Filter

Find & Select

GET GENUINE OFFICE Your license isn't genuine, and you may be a victim of software counterfeiting. Avoid interruption and keep your files safe with genuine Office today. Get genuine Office Learn more

G2

Group AgGroup E

	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Group Dr													
2	Group C		Group Ag	GG	Group E	Group B	P	Group B	N	Group Dr	Scored Pr	Scored Pr	Scored Pr	Scored Labels
3	Group B		96	F	H	32	Group C	2.13E-07	0.609174	0.570332	6.11E-10	4.53E-07	Group B	
4	Group B		75	F	H	18	Group B	0.20109	0.503059	0.001255	0.000278	0.020256	Group B	
5	Group A		78	M	N	16	Group B	0.828491	0.282842	4.01E-05	2.10E-05	0.016017	Group A	
6	Group E		98	F	H	15	Group A	0.886277	0.239221	9.94E-06	2.70E-05	9.90E-07	Group A	
7			43	M	L	21	Group E	5.09E-07	2.77E-06	0.001006	0.000179	0.999989	Group E	
8														
9														
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														

Sheet1

Ready Accessibility: Investigate

Azure Machine Learning

Input: input1

Sheet1!A1:E6

☒ My data has headers

Use sample data ?

Output: output1

Sheet1!G2

☒ Include headers

Predicting will override existing values. This can't be undone. Got it!

Predict Auto-predict

3. ERRORS

Help Privacy Statement