



Azure Machine Learning

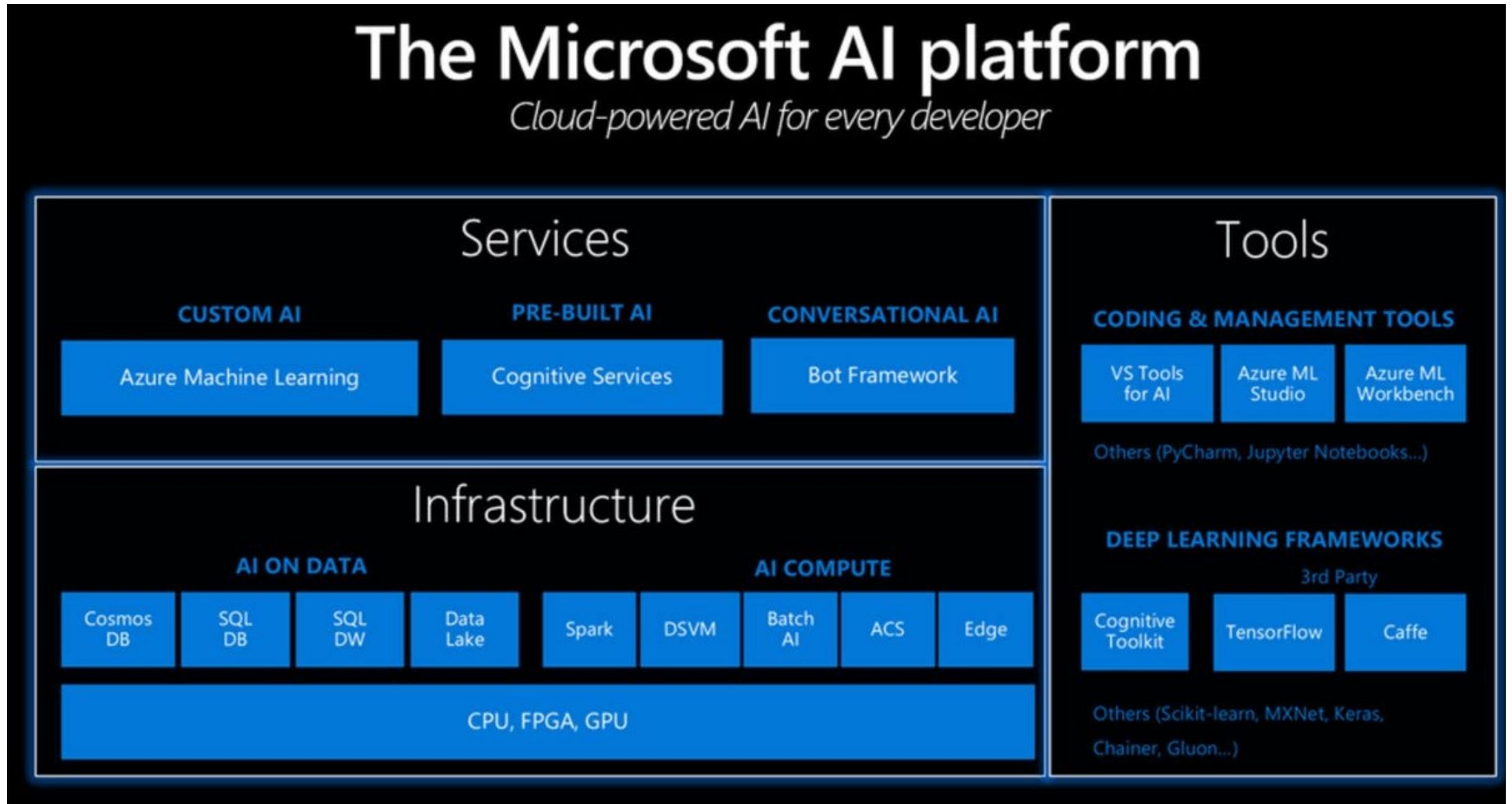
Azure Machine Learning: ML como servicio

Javier Panadero

Engineering School
Universitat Autònoma de Barcelona

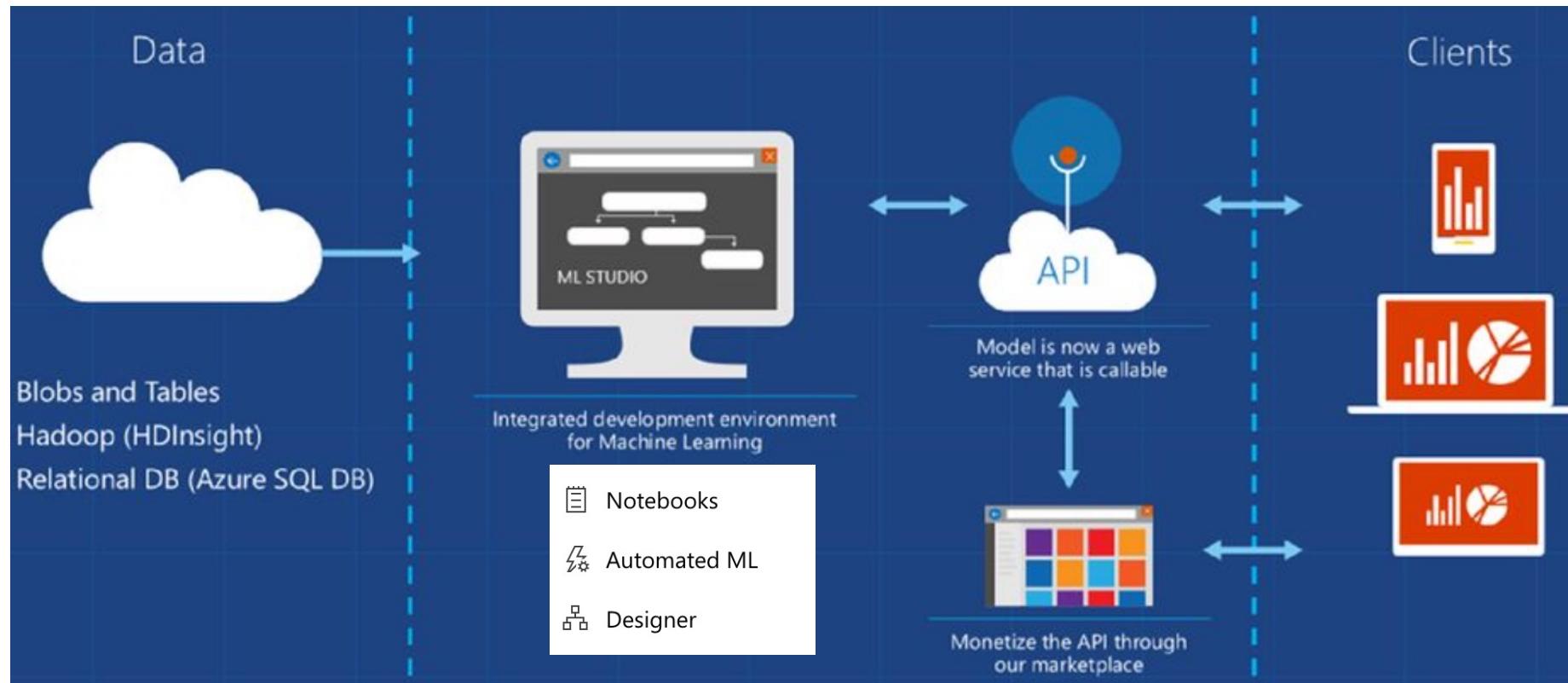
UAB
Universitat Autònoma
de Barcelona

Azure ML: La inteligencia artificial como servicio



¿Qué es Azure ML Studio?

- Servicio en la nube de Microsoft que permite a los desarrolladores y científicos de datos construir, entrenar y desplegar modelos de aprendizaje automático de manera eficiente (bajo código) y escalable.
- Permite que personas no expertas en programación puedan llegar a crear modelos complejos de aprendizaje automático.



Azure Machine Learning Studio

Azure Machine Learning Studio proporciona tres herramientas clave para diseñar, desarrollar y automatizar tareas relacionadas con el aprendizaje automático.

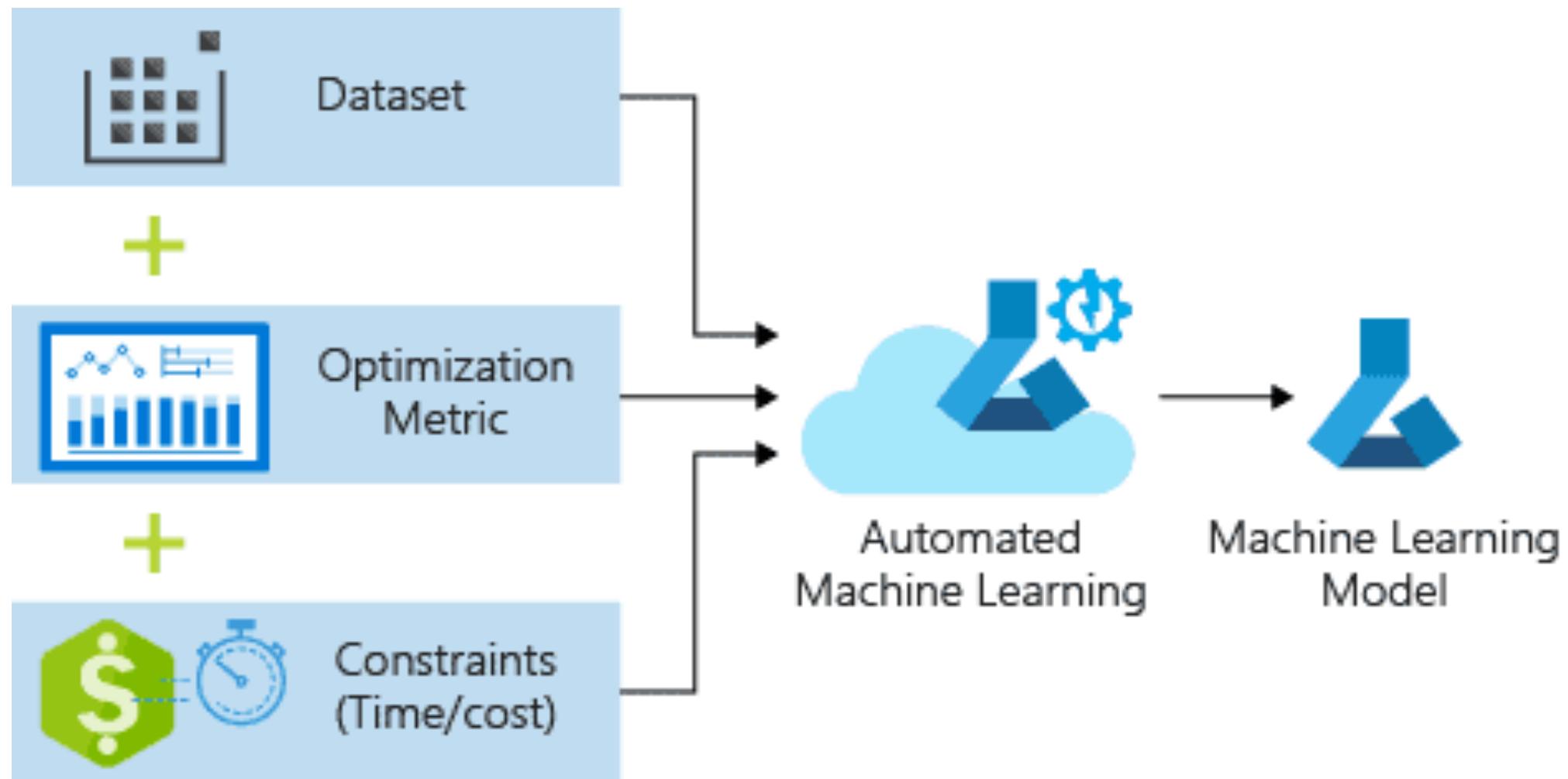
- Automated ML: Herramienta que simplifica y automatiza el proceso de creación, selección y entrenamiento del mejor modelo de aprendizaje automático. Permite que usuarios sin experiencia profunda en aprendizaje automático puedan desarrollar modelos de alta calidad de manera eficiente.
- Designer: Interfaz gráfica que permite diseñar flujos de trabajo de aprendizaje automático mediante la conexión de bloques predefinidos que representan tareas de preparación de datos, entrenamiento de modelos y evaluación. Permite crear flujos de trabajo de extremo a extremo sin necesidad de escribir código.
- Notebook: Permite utilizar cuadernos Jupyter para escribir código en Python o R y ejecutarlo en el cloud.

The screenshot shows the Azure Machine Learning studio interface. On the left is a sidebar with navigation links: New, Home (selected), Author, Notebooks, Automated ML, Designer, Assets, Datasets, Experiments, Pipelines, Models, Endpoints, Manage, Compute, Datastores, Data Labeling, and Linked Services. The main area is titled "Home" and "Azure Machine Learning studio". It features three cards: "Create new" (Notebooks), "Notebooks" (Code with Python SDK and run sample experiments, Start now button), "Automated ML" (Automatically train and tune a model using a target metric, Start now button), and "Designer" (Drag-and-drop interface from prepping data to deploying models, Start now button). Below this is a section titled "My recent resources" with a table titled "Runs". The table has columns: Run, Run ID, Experiment, Status, Submitted time, Submitted by, and Run type. It lists two entries:

Run	Run ID	Experiment	Status	Submitted time	Submitted by	Run type
Run 13	c50066c0-2413-4e03-ae0...	mslearn-...	Completed	2021-05-10T00:00:32....	Marius Hucker	Pipeline
Run 10	1d4cb022-b29d-478a-bc8...	mslearn-...	Completed	2021-05-03T00:00:32....	Marius Hucker	Pipeline

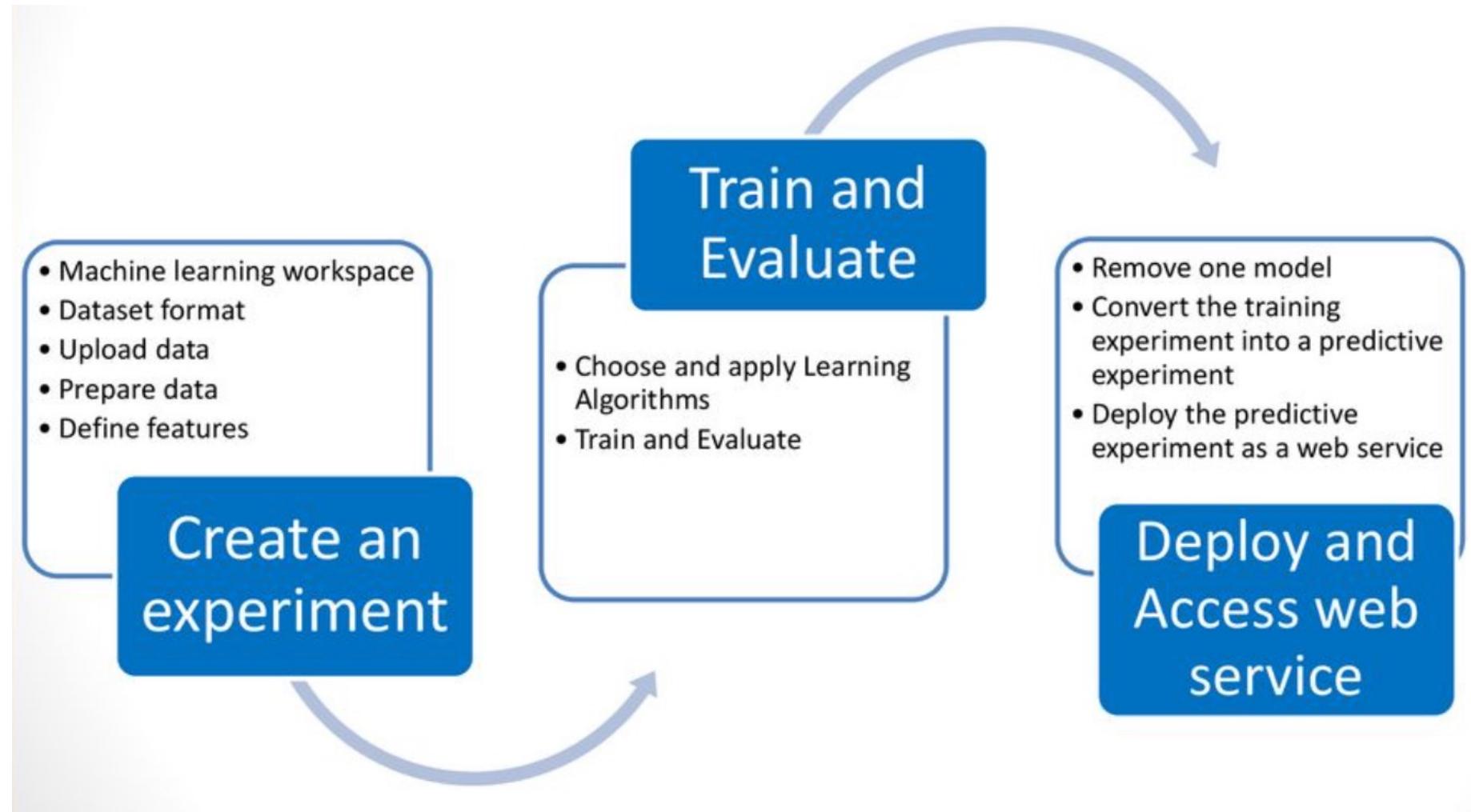
Automated ML

Recibe como entrada un dataset, una métrica de optimización y unas restricciones de tiempo y/o coste y busca el mejor modelo de aprendizaje automático que mejor se ajuste al dataset, ya sea de clasificación o regresión, y genera un modelo que puede ser llamado y ejecutado mediante una API REST.



Flujo de creación del modelo de aprendizaje en Automated ML

Automated ML necesita que el usuario complete 3 pasos para crear y ejecutar el modelo de predicción: Creación de un experimento, entrenamiento y evaluación del modelo, y despliegue del modelo



Tutorial de uso de Automated ML: Entendiendo el Dataset

La columna 'y' indica si un cliente se suscribió a un depósito a plazo fijo, la cual será la columna objetivo para las predicciones en este tutorial.

bankmarketing_train (3)

age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome	emp.var.rate	cons.price.idx	cons.conf.idx	euribor3m	nr.employed	y		
57	technician	married	high.school	no	no	yes	cellular	may	mon	371	1	999	1	failure	-1.8	92.89299999999999	-46.2	1.2990000000000002	5099.1	no		
55	unknown	married	unknown	unknown	yes	no	telephone	may	thu	285	2	999	0	nonexistent	1.1		93.994	-36.4	4.86	5191.0	no	
33	blue-collar	married	basic.9y	no	no	no	cellular	may	fri	52	1	999	1	failure	-1.8	92.89299999999999	-46.2	1.3130000000000002	5099.1	no		
36	admin.	married	high.school	no	no	no	telephone	jun	fri	355	4	999	0	nonexistent	1.4		94.465	-41.8		4967	5228.1	no
27	housemaid	married	high.school	no	yes	no	cellular	jul	fri	189	2	999	0	nonexistent	1.4	93.91799999999999	-42.7		4963	5228.1	no	
58	retired	married	professional.course	no	yes	yes	cellular	jul	fri	605	1	999	0	nonexistent	1.4	93.91799999999999	-42.7		4962	5228.1	no	
48	services	married	high.school	unknown	yes	no	telephone	may	wed	243	1	999	0	nonexistent	1.1		93.994	-36.4		4856	5191.0	no
51	admin.	divorced	university.degree	unknown	yes	no	cellular	aug	thu	24	7	999	0	nonexistent	1.4		93.444	-36.1		4962	5228.1	no
24	entrepreneur	married	university.degree	no	yes	yes	telephone	jun	wed	126	4	999	0	nonexistent	1.4		94.465	-41.8		4962	5228.1	no
36	technician	divorced	professional.course	no	yes	yes	cellular	jul	mon	43	4	999	0	nonexistent	1.4	93.91799999999999	-42.7		4962	5228.1	no	
34	blue-collar	married	basic.9y	no	no	no	cellular	may	fri	451	3	999	0	nonexistent	-1.8	92.89299999999999	-46.2	1.3130000000000002	5099.1	no		
42	blue-collar	single	basic.4y	no	yes	no	cellular	may	thu	560	3	999	0	nonexistent	-1.8	92.89299999999999	-46.2		1327	5099.1	no	
43	blue-collar	single	basic.9y	no	yes	no	cellular	jul	thu	982	1	999	0	nonexistent	1.4	93.91799999999999	-42.7		4963	5228.1	yes	
56	admin.	married	university.degree	no	no	no	cellular	aug	wed	552	3	999	0	nonexistent	1.4		93.444	-36.1	4.963999999999995	5228.1	yes	
58	blue-collar	married	basic.4y	no	no	yes	cellular	jul	thu	265	3	999	0	nonexistent	1.4	93.91799999999999	-42.7		4962	5228.1	no	
36	housemaid	married	high.school	no	yes	no	telephone	jun	mon	171	1	999	0	nonexistent	1.4		94.465	-41.8		4865	5228.1	no
26	blue-collar	single	basic.9y	unknown	yes	no	cellular	jul	fri	151	2	999	0	nonexistent	1.4	93.91799999999999	-42.7		4963	5228.1	no	
53	technician	married	professional.course	no	no	yes	cellular	jul	tue	178	11	999	0	nonexistent	1.4	93.91799999999999	-42.7		4961	5228.1	no	
37	entrepreneur	single	professional.course	no	yes	no	cellular	jul	fri	92	2	999	0	nonexistent	1.4	93.91799999999999	-42.7		4957	5228.1	no	
46	admin.	married	university.degree	no	yes	no	cellular	aug	thu	399	1	999	1	failure	-2.9	92.2010000000001	-31.4		873	5076.2	yes	
59	retired	married	basic.9y	unknown	no	no	cellular	jul	fri	1330	3	999	0	nonexistent	1.4	93.91799999999999	-42.7		4962	5228.1	yes	
49	unemployed	single	professional.course	unknown	no	no	telephone	may	fri	339	2	999	0	nonexistent	1.1		93.994	-36.4		4855	5191.0	no
26	unknown	single	high.school	unknown	yes	yes	telephone	jun	wed	524	15	999	0	nonexistent	1.4		94.465	-41.8		4864	5228.1	yes
30	admin.	single	high.school	no	no	yes	cellular	jul	mon	101	3	999	0	nonexistent	1.4	93.91799999999999	-42.7	4.96		5228.1	no	

Dataset: https://automl-samples-notebook-data.blob.core.windows.net/automl-sample-notebook-data/bankmarketing_train.csv

Crear un experimento: Creación de un Workspace

Azure AI | Machine Learning Studio Search All workspaces L G M ? JM

UAB

Workspaces Feature stores PREVIEW Registries Shared assets Components Environments Models Data

Admin Quota

Creamos un nuevo Workspace para desarrollar, entrenar y desplegar el modelo de predicción

The Learning Studio

Subscription: Azure for Students

Create workspace View all workspaces →

Learning components

Build AI solutions with Azure Machine Learning Introduction to the Azure Machine Learning SDK Train a machine learning model with Azure ML

View all learning components →

Tutorials

What is Azure Machine Learning? Train your first ML model with Notebook Create, explore and deploy Automated ML experiments.

View all tutorials →

Additional resources

Azure Machine Learning Studio: <https://ml.azure.com/>

Crear un experimento: Creación de un Workspace

The screenshot shows the Azure portal interface for creating a workspace. On the left, there's a sidebar with navigation links: UAB, Workspaces (which is selected and highlighted in blue), Feature stores (with a PREVIEW badge), Registries, Shared assets, Components, Environments, Models, Data, Admin, and Quota. The main content area is titled "Workspaces" and contains a brief description: "A workspace provides a central location for all your machine learning artifacts you create while performing machine learning experiments." Below this is a table with columns: Workspace, Resource group, Region, Subscription, and Created on. A single row is visible: "ML_UAB23" under Workspace, "prueba1" under Resource group, "francecentral" under Region, "Azure for Students" under Subscription, and "Sep 14, 2023 12:52:11" under Created on. At the top of the main area, there's a search bar labeled "Search" and a button labeled "+ New". A large yellow callout box with a blue border is overlaid on the "+ New" button, containing the text: "Seleccionamos 'New' para crear nuestro espacio de trabajo ('Workspace')".

Workspace	Resource group	Region	Subscription	Created on
ML_UAB23	prueba1	francecentral	Azure for Students	Sep 14, 2023 12:52:11

Crear un experimento: Creación de un Workspace

The screenshot shows the Azure ML Studio interface. On the left, there's a sidebar with navigation links: UAB, Workspaces (selected), Feature stores (PREVIEW), Registries, Shared assets, Components, Environments, Models, Data, Admin, and Quota. The main area is titled "Workspaces" and shows a list of existing workspaces: "ML_UAB23" and "prueba1". A "New" button is visible. To the right, a modal window titled "Create new workspace" is open, containing fields for "Workspace name" (set to "ml_gcd22"), "Subscription" (set to "Azure for Students"), "Resource group" (set to "(new) 2131865-rg"), and "Region" (set to "Germany West Central"). At the bottom of the modal are "Create" and "Cancel" buttons.

1. Seleccionamos un nombre único para crear nuestro espacio de trabajo.
2. Seleccionamos la suscripción “estudiante”
3. Creamos un nuevo grupo de recursos (dejamos el que salga por defecto)
4. Seleccionamos la región más cercana (Germany)

Create new workspace

Specify details for your new workspace. To configure advanced options such as private link, use the creation experience in the [Azure Portal](#).

Workspace name * [i](#)
ml_gcd22

Subscription * [i](#)
Azure for Students

Refresh subscriptions

Resource group * [i](#)
(new) 2131865-rg

Create new

Region * [i](#)
Germany West Central

Create Cancel

4. Hacemos click en “Create” para crear el workspace

Crear un experimento: Creación de un Workspace

Tras aproximadamente 1 minuto aparecerá nuestro nuevo workspace (haced click en el botón *Refresh* para refrescar la pantalla y ver el nuevo workspace)

The screenshot shows the Azure ML Studio interface. On the left, there's a sidebar with icons for UAB, Workspaces (selected), Feature stores (with a PREVIEW badge), Registries, Shared assets, Components, Environments, Models, Data, and Admin. The main area is titled "Workspaces" and contains a brief description: "A workspace provides a centralized place to keep track of all the artifacts you create while performing machine learning experiments." Below this are buttons for "+ New", "Refresh", "Edit workspace", and "View options". A search bar and a "Filter" button are also present. A table lists workspaces based on their name, resource group, region, and subscription. The first row, "ml_gcd22", is highlighted with a red border. A callout box over this row contains the text: "Hacemos click en ‘ml_gcd22’ para entrar en nuestro workspace".

Workspace name ↑	Resource group	Region	Subscription
ml_gcd22	2131865-rg	germanywestcentral	Azure for Students
ML_UAB23		francecentral	Azure for Students

Hacemos click en “ml_gcd22” para entrar en nuestro workspace

Crear un experimento: Automated ML

Azure AI | Machine Learning Studio

⋮

← All workspaces

Home

Model catalog PREVIEW

Authoring

⋮ Notebooks

Automated ML

Designer

Prompt flow PREVIEW

Assets

Data

Jobs

ml_gcd22

🕒 ⚡ 🛡️ 🎤 🌐 ? 😊 Azure for Students ml_gcd22 JM

ml_gcd22

Generate AI with Prompt flow PREVIEW ⋮

Hacemos click en Automated ML para empezar a crear el modelo

Q&A with your own data using ...

Bring Your Own Data QnA

Create flows for Q&A with GPT3.5 using data from your own indexed files to make the answer more grounded for enterprise chat scenarios.

Start Clone

Ask Wikipedia

Q&A with GPT3.5 using information from Wikipedia to make your answers more grounded.

Start Clone

Chat with Wikipedia

ChatGPT-based chatbot that leverages Wikipedia data to ground the responses.

Start Clone

+ New ⋮ Customize view

View prompt flow < >

The screenshot shows the Azure AI | Machine Learning Studio interface. On the left, there's a sidebar with various sections like Home, Model catalog, Notebooks, Automated ML (which is highlighted with a red box), Designer, Prompt flow, Assets, Data, and Jobs. The main area displays the 'ml_gcd22' workspace. It features a title 'ml_gcd22' with a edit icon. Below it is a card titled 'Generate AI with Prompt flow' with a 'PREVIEW' link and three dots. A large callout bubble points to a card titled 'Hacemos click en Automated ML para empezar a crear el modelo'. This card has a yellow background and contains text about using GPT3.5 with domain knowledge from Faiss index. It includes 'Start' and 'Clone' buttons. Other cards visible include 'Bring Your Own Data QnA', 'Ask Wikipedia', and 'Chat with Wikipedia', each with their own descriptions and 'Start' and 'Clone' buttons. The top right of the screen shows some navigation icons and user info. The overall theme is light blue and white.

Crear un experimento: Cargando el dataset de entrenamiento

The screenshot shows the Azure Machine Learning studio interface. On the left, there is a navigation sidebar with the following items:

- All workspaces
- Home
- Model catalog (PREVIEW)
- Authoring
- Notebooks
- Automated ML (selected)
- Designer
- Prompt flow (PREVIEW)

On the right, the main content area is titled "Automated ML". It displays the following information:

- UAB > ml_gcd22 > Automated ML
- Automated ML
- Let Automated ML train and find the best model based on your data without writing a single line of code. [Learn more about Automated ML](#)
- + New Automated ML job (button highlighted with a red box)
- Refresh
- Hacemos click en "New Automated ML job" para crear un nuevo trabajo (callout box with blue border and yellow background)
- No recent Automated ML jobs to display.
- Click "New Automated ML job" to create your first job
- [Learn more about creating Automated ML jobs](#)

At the bottom, there is a "Documentation" section with three items:

- Concept: What is Automated ML?
- Tutorial: Create your first classification model with Automated ML
- Blog: Build more accurate forecasts with new capabilities in Automated ML

On the far right of the documentation section, there is a link: [View all documentation](#).

Crear un experimento: Cargando el dataset de entrenamiento

Cargamos nuestra fuente de datos para poder entrenar y testear el modelo de aprendizaje automático

UAB > ml_gcd22 > Automated ML > Start job

Create a new Automated ML job

1 Select data asset 2 Configure job 3 Select task and settings 4 Hyperparameter configuration (Computer Vision only) 5 Validate and test

Select data asset
Select an input data asset from the list below, or create a new data asset. AutomatedML currently only supports tabular data for authoring jobs.

+ Create Refresh Show supported data assets only

Hacemos click en *Create*

Name	Dataset type	Created on ↓	Modified on
No data assets to display			

<< < Page 1 of 0 > >> 25/Page ▾

Back Next Cancel

Crear un experimento: Cargando el dataset de entrenamiento

Create data asset

1 Data type

2 Data source

Asignamos un nombre a nuestro *dataset* y hacemos click en “next”

Set the name and type for your data asset

Name *

HipotecasUSA

Description

Data asset description

Type * ⓘ

Tabular

Back

Next

Use cases for data types

When should I use File type?

The File type is recommended in most scenarios when you are working with a single data file of any type (e.g., tabular data). This type allows you to specify a URI in a storage location on your local computer or attached Datastore, blob/ADLS storage, or a publicly available http(s) location. There are many types of supported URLs. In the Azure Machine Learning Python SDK v2, this data type is called [uri_file](#). [Learn more about the uri_file type](#)

When should I use Folder type?

The Folder type has all the same capabilities as the File type, but is used when specifying a location. In the Azure Machine Learning Python SDK v2, this data type is called [uri_folder](#). [Learn more about the uri_folder type](#)

When should I use Table type?

The Table type is most useful for advanced

Crear un experimento: Cargando el dataset de entrenamiento

Create data asset

X

- Data type
- Data source

Choose a source for your data asset

Choose the data source you want to create your asset from. A data source can be from a local storage location on your computer, from an attached datastore, from Azure storage, or from a publicly available web location.



From Azure s...

Create a data asset from Storage, Azure file sha...

Seleccionamos la opción *From Local File* para cargar nuestro dataset desde nuestro dispositivo local

From local files



From local files

Create a data asset by uploading files from your local drive.



From SQL databases

Create a dataset from Azure SQL database and Azure PostGreSQL database.



From web files

Create a data asset from a single file located at a public web URL.



From Azure Open Datasets

Create a dataset with one-click from pre-made data sets. These data sets are created by the general public and published as Azure Open Datasets

Crear un experimento: Cargando el dataset de entrenamiento

- 1 Data type
- 2 Data source
- 3 Destination storage type
- 4 File or folder selection
- 5 Settings
- 6 Schema
- 7 Review

Select a datastore

Choose a storage type and a datastore to upload your data to in the next step. You can also create a new datastore for your data first.

Datastore type *

Azure Blob Storage + Create new datastore

Name ↓	Storage name	Created on
<input checked="" type="checkbox"/> workspaceblobstore	mlgcd223044047513	Sep 19, 2023 5:29 PM
workspaceartifactstore	selected223044047512	Sep 19, 2023 5:29 PM

Seleccionamos workspaceblobstore y hacemos click en next

« < Page 1 of 1 > » 25/Page

Back Next Cancel

Crear un experimento: Cargando el dataset de entrenamiento

Create data asset

- ✓ Data type
- ✓ Data source
- ✓ Destination storage type
- 4 File or folder selection
- 5 Settings
- 6 Schema
- 7 Review

Choose a file or folder

Choose files or folders to upload from your local drive. If you upload multiple folders or files, they will be stored in a containing folder.

Upload path

azureml://subscriptions/0145e35e-aca5-48f2-ad9f-0682a851dd16/resourcegroups/2131865...



Upload ▾

Overwrite if already exists

Seleccionamos *upload*
para cargar nuestro
dataset en Azure

Upload list

*File Types supported are delimited (i.e. csv, tsv), Parquet,
JSON Lines, and plain text.*

Information

What file types can I use?

Supported file types include: delimited (i.e. csv, tsv), Parquet, JSON Lines, and plain text.

Where are files uploaded?

Files will be uploaded to the selected location available in your workspace.

Back

Next

Crear un experimento: Cargando el dataset de entrenamiento

Si el archivo se ha cargado correctamente, se visualizará una vista previa de su contenido de forma automática

Create data asset X

Data type
 Data source
 Destination storage type
 File or folder selection
5 Settings
6 Schema
7 Review

Settings
These settings determine how the data is parsed. The initial settings are automatically detected; you can change them as needed to reparse the data.

File format	Delimiter	Example	Encoding
Delimited	Comma	Field1,Field2,Field3	UTF-8

Column headers **Skip rows**

All files have same headers	None
-----------------------------	------

Dataset contains multi-line data (i)

(i) Note: Processing tabular files with multi-line data is slower because multiple CPU cores cannot be used to ingest the data in parallel. Checking this option may result in slower processing times.

Data preview

age	job	marital	educati...	default	housing	loan	contact	month	day_of...	duration	campai...	pdays	previc
57	technician	married	high.sch...	no	no	yes	cellular	may	mon	371	1	999	1
55	unknown	married	unknown	unknown	yes	no	telephone	may	thu	285	2	999	0
33	blue-col...	married	basic.9y	no	no	no	cellular	may	fri	52	1	999	1
36	admin.	married	high.sch...	no	no	no	telephone	jun	fri	355	4	999	0
27	housem...	married	high.sch...	no	yes	no	cellular	jul	fri	189	2	999	0
58	retired	married	professi...	no	yes	yes	cellular	jul	fri	605	1	999	0

Back Next Review Cancel

Crear un experimento: Preparación de los datos a utilizar

Esta pantalla nos permite seleccionar que *features* queremos mantener para que sean consideradas a la hora de crear el modelo y cuales queremos descartar.

Create data asset X

✓ Data type
✓ Data source
✓ Destination storage type
✓ File or folder selection
✓ Settings
6 Schema
7 Review

Schema
Column types are auto-detected based on the initial subset of the data and can be updated here. Values not would be either null-filled or replaced with error value. Any conversions preview errors are non-blocking an

Search column name

Include	Column name	Type	Example values	Date format <small>i</small>	Properties <small>i</small>
<input checked="" type="checkbox"/>	Path	String		Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	age	Integer	57, 55, 33	Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	job	String	technician, unknown, blue-collar	Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	marital	String	married, married, married	Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	education	String	high.school, unknown, basic.9y	Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	default	String	no, unknown, no	Not applicable to s...	Not applicable t... ▾
<input checked="" type="checkbox"/>	housing	String	no, yes, no	Not applicable to s...	Not applicable t... ▾

Buscamos y deseleccionamos “Day of the week” y hacemos click en next

Back Next Cancel

Crear un experimento: Creando el dataset final en azure

La siguiente ventana nos muestra un resumen de los pasos que hemos realizado anteriormente para asegurarnos que todo es correcto antes de proceder a crear el dataset en el cloud.

Create data asset

- ✓ Data type
- ✓ Data source
- ✓ Destination storage type
- ✓ File or folder selection
- ✓ Settings
- ✓ Schema
- ⑦ Review

/2023-09-19T154225_UTC/DarkMarketing_train (2).csv

Files uploaded bankmarketing_train (2).csv	
Storage	
Datastore type AzureBlob	
Datastore name workspaceblobstore	
Settings	
Delimiter Comma	
Encoding UTF-8	
File format Delimited	
Column headers All files have same headers	
Number of rows to skip None	
Dataset contains multi-line data false	

Back **Create**

Seleccionamos *create* para
crear nuestro dataset

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

UAB > ml_gcd22 > Automated ML > Start job

Create a new Automated ML job

1 Select data asset
2 Configure job
3 Select task and settings
4 Hyperparameter configuration (Computer Vision only)
5 Validate and test

Select data asset

Select an input data asset from the list below, or create a new data asset. AutomatedML currently only supports tabular data for authoring jobs.

Success: HipotecasUSA data asset created successfully. It may take a few seconds for lists to be updated. [Click here to go to this data asset](#)

+ Create Refresh Show supported data assets only

Name	Dataset type	Created on ↓	Modified on
HipotecasUSA	Tabular	Sep 19, 2023 5:55 PM	Sep 19, 2023 5:55 ..

Search Filter

HipotecasUSA

Back Next Cancel

Seleccionamos *next* para pasar a la siguiente etapa y configurar el job

Entrenamiento y evaluación: Seleccionar el modelo y sus los párametros

UAB > ml_gcd22 > Automated ML > Start job

Create a new Automated ML job

1. Asignamos un nombre al experimento

2. Indicamos la columna del dataset que será la variable objetivo a predecir por el modelo

3. Seleccionamos “Compute Cluster” para generar un grupo de recursos de cómputo

4. Seleccionamos *new* para crear un nuevo grupo de recursos de cómputo (Cluster)

Configure job
Select from existing experiments or create a new experiment, then select the target column and training compute.

[Learn more on how to configure the experiment.](#)

Data asset
HipotecasUSA ([View data asset](#))

Experiment name
 Create new
New experiment name *

Target column * [i](#)

Select compute type

Select Azure ML
No compute cluster
[+ New](#) [↻](#)

Back [Next](#)

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

Create compute cluster (i) X

VIRTUAL MACHINE TIER (i)

Virtual Machine (i)

Advanced Settings

Dedicated Low priority

Virtual machine type (i)

CPU GPU

Virtual machine size (i)

Select from recommended options Select from all options

Dejamos la configuración por defecto y seleccionamos *next*.

Name ↑	Category	Workload types	Available quota (i)	Cost (i)
<input type="radio"/> Standard_DS11_v2 2 cores, 14GB RAM, 28GB storage	Memory optimized	Development on Notebooks (or other IDE) and light weight testing	6 cores	\$0.17/hr
<input checked="" type="radio"/> Standard_DS3_v2 4 cores, 14GB RAM, 28GB storage	General purpose	Classical ML model training on small datasets	6 cores	\$0.27/hr
<input type="radio"/> Standard_E4ds_v4 4 cores, 32GB RAM, 150GB storage	Memory optimized	Data manipulation and training on medium-sized datasets (1-10GB)	4 cores	\$0.35/hr
<input type="radio"/> Standard_F4s_v2 4 cores, 8GB RAM, 32GB storage	Compute optimized	Data manipulation and training on large datasets (>10 GB)	16 cores	\$0.19/hr

Back Next Cancel

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

Create compute cluster (i) X

Virtual Machine (i)

Advanced Settings (i)

Standard_DS3_v2	6 cores	14 GB	28 GB	\$0.27/hr				
Compute name * (i)	Asignamos un nombre al cluster y hacemos click en Create							
gcd								
Minimum number of nodes * (i)	0	<input type="range"/>						
Maximum number of nodes * (i)	1	<input type="range"/>						
Idle seconds before scale down * (i)	120							
<input checked="" type="checkbox"/> Enable SSH access (i)								
> Advanced settings								
Add tags (i)	<table border="1"><tr><td>Name</td><td>:</td><td>Value</td><td>Add</td></tr></table>				Name	:	Value	Add
Name	:	Value	Add					
(i) No tags								
Back	Create	Download a template for automation.						
Cancel								

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

Azure AI | Machine Learning Studio

UAB > ml_gcd22 > Automated ML > Start job

Create a new Automated ML job

1 Select data asset
2 **Configure job**
3 Select task and settings
4 Hyperparameter configuration (Computer Vision only)
5 Validate and test

Configure job
Select from existing experiments or create a new experiment, then select the target column and training compute.
[Learn more on how to configure the experiment.](#)

Data asset
HipotecasUSA ([View data asset](#))

Experiment name
 Create new
New experiment name *
gcd-22

Target column *
y (String)

Select compute type
Compute cluster
Select Azure ML compute cluster
gcd
+ New Refresh compute

Seleccionamos *next* para ir a la siguiente ventana y seleccionar el tipo de algoritmo de aprendizaje automático a utilizar

Back **Next**

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

Automated ML propone por defecto el tipo de algoritmo a utilizar a partir de la *target feature* que hemos indicado en el paso anterior

The screenshot shows the 'Create a new Automated ML job' interface in Azure AI | Machine Learning Studio. On the left, there's a sidebar with navigation links like 'All workspaces', 'Home', 'Model catalog (PREVIEW)', 'Authoring', 'Notebooks', 'Automated ML' (which is selected), 'Designer', 'Prompt flow (PREVIEW)', 'Assets', 'Data', 'Jobs', 'Components', 'Pipelines', 'Environments', 'Models', 'Endpoints', and 'Manage'. The main area has a breadcrumb path: UAB > ml_gcd22 > Automated ML > Start job. The steps are listed as: 1 Select data asset (done), 2 Configure job (done), 3 Select task and settings (in progress), 4 Hyperparameter configuration (Computer Vision only), and 5 Validate and test. The 'Select task and settings' step is currently active. It lists several options: Classification (selected), Regression, Time series forecasting, Natural language processing, Computer vision, and View additional configuration settings. The 'Classification' section is highlighted with a red box. A callout bubble points to the 'View additional configuration settings' link at the bottom of the classification section.

Classification
To predict one of several categories in the target column. yes/no, blue, red, green.
 Enable deep learning ⓘ

Regression
To predict continuous numeric values.

Time series forecasting
To predict values based on time.

Natural language processing
Predict based on text-only data types using...

Computer vision
Multi-class or multi-label image classificat...

View additional configuration settings

Back Next

Entrenamiento y evaluación: Seleccionar el modelo y sus los parámetros

Additional configurations X

1. Seleccionamos como métrica principal *AUC weighted*

Primary metric (i)
AUC weighted

Explain best model (i)

Use all supported models (i)

Blocked models (i)
A list of models that Automated ML will not use during training.

> Additional classification settings

< Exit criterion

Training job time (hours) (i)
1

Metric score threshold (i)
Metric score threshold

> Concurrency

2. Seleccionamos Explain best model en caso de que no está seleccionado

3. Modificamos el tiempo de entrenamiento a 1 hora

4. Hacemos click en Save

Save Cancel

Entrenamiento y evaluación: Seleccionar el modelo y sus los párametros

UAB > ml_gcd22 > Automated ML > Start job

Create a new Automated ML job

- ✓ Select data asset
- ✓ Configure job
- ✓ Select task and settings
- ✓ Hyperparameter configuration (Computer Vision only)
- 5 Validate and test

Select the validation and test type

You can choose a validation type and select a test data asset as an optional step. Providing your own validation and test data assets are currently preview features.

Validation type ⓘ
k-fold cross validation

Number of cross validations * ⓘ
2

Test data asset (preview) ⓘ
No test data asset required

1. Seleccionamos como validation type *k-fold cross validation*

2. Modificamos el valor original a 2

3. Seleccionamos *Finish*

Back Finish

Entrenamiento y evaluación: Explorando el modelo creado

La vista *models* nos muestra todos los algoritmos probados y las calidades obtenidas. Además nos explica las features más importantes que ha considerado el mejor modelo obtenido

The screenshot shows a user interface for managing machine learning models. At the top, there's a header with the project name "gentle_loquat_xv7qy8pm" and status indicators (blue pencil icon, star icon, green checkmark, "Completed"). Below the header, a navigation bar includes "Overview", "Data guardrails", "Models" (which is underlined), "Outputs + logs", and "Child jobs". A toolbar below the navigation bar contains "Refresh", "Deploy", "Download", "Explain model", "View generated code", and "View options". A search bar and filter/column controls are also present.

The main area is a table listing various models:

Algorithm name	Explained	Responsible AI	AUC weighted ↓	Sampling	Created on
VotingEnsemble	View explanation		0.94823	100.00 %	Sep 19, 2023 7:16 PM
StackEnsemble			0.94823	100.00 %	Sep 19, 2023 7:18 PM
MaxAbsScaler, LightGBM			0.94823	100.00 %	Sep 19, 2023 6:55 PM
SparseNormalizer, XGBoostClassifier			0.94632	100.00 %	Sep 19, 2023 6:53 PM
SparseNormalizer, XGBoostClassifier			0.94587	100.00 %	Sep 19, 2023 7:11 PM
MaxAbsScaler, LightGBM			0.94583	100.00 %	Sep 19, 2023 6:20 PM
StandardScalerWrapper, XGBoostClassifier			0.94566	100.00 %	Sep 19, 2023 7:07 PM
StandardScalerWrapper, XGBoostClassifier			0.94561	100.00 %	Sep 19, 2023 6:20 PM
SparseNormalizer, XGBoostClassifier			0.94534	100.00 %	Sep 19, 2023 7:00 PM

A red box highlights the "View explanation" link for the first row. A yellow box with the text "Seleccionamos view explanation" is overlaid on the second row's "Explained" column.

Entrenamiento y evaluación: Explorando el modelo creado

View Explanations nos proporciona las *features* más importantes del modelo

crimson_pear_psp3bfmw Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs ...

Refresh Deploy Download Explain model View generated code Test model (preview) Register model Cancel Delete

Explanation ID Explainer: mimic.lightgbm View previous dashboard experience

+ New cohort Edit cohort

Model performance Dataset explorer Aggregate feature importance Individual feature importance

Explore the top-k important features that impact your overall model predictions (a.k.a. global explanation). Use the slider to show descending feature importances. All cohorts' feature importances are shown side by side and can be toggled off by selecting the cohort in the legend. Click on any of the features in the graph to see a density plot below of how values of the selected feature affect prediction.

Top 4 features by their importance

DATA STATISTICS
Binary classifier
5000 datapoints
19 features

DATASET COHORTS
All data
5000 datapoints
0 filters

Aggregate feature importance

Feature	Aggregate feature importance
duration	~0.80
nr.employed	~0.48
emp.var.rate	~0.32
euribor3m	~0.18

Sort by cohort All data

Chart type Bar Box

Class importance weights Class: no

Deploy del modelo

Generado el modelo de predicción podemos realizar el despliegue (*deploy*) para ser utilizado por cualquier usuario de manera remota.

The screenshot shows the 'Model' tab for a completed experiment named 'crimson_pear_psp3bfmw'. The 'Deploy' button is highlighted with a red box, and a tooltip says 'Seleccionamos Deploy → Web Service'. A second red box highlights the 'Web service' option in the dropdown menu.

crimson_pear_psp3bfmw Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs Images Child jobs ...

Refresh Deploy Download Explain model View generated code Test model (preview) Register model Cancel Delete

Real-time endpoint
Deploy the model using the real-time endpoint wizard

Batch endpoint
Deploy the model using the batch endpoint wizard

Web service
Deploy the model to a web service

Model summary
Algorithm name VotingEnsemble
Ensemble detail
View ensemble
AUC weighted 0.94823
Sampling 100.00 %
Registered models No registration yet
Deploy status No deployment yet

Seleccionamos Deploy → Web Service

Deploy del modelo

The screenshot shows the Azure Machine Learning studio interface. On the left, there's a model summary for a 'VotingEnsemble' algorithm named 'degcd22'. The 'Model' tab is selected. On the right, a 'Deploy a model' dialog box is open, overlaid on the main interface.

1. Asignamos un nombre al deploy del modelo

2. Seleccionamos Azure Container Instance

3. Hacemos click en Deploy

Deploy a model

Name * (i) eye icon

Description

Compute type * (i) dropdown arrow

Models: AutoMLa7ddf57fa52

Enable authentication

This model supports no-code deployment. You may optionally override the default environment and driver file.

Use custom deployment assets Use custom deployment assets

> Advanced

Deploy Cancel

Deploy del modelo

Una vez realizado el inicio del *deploy* nos lleva a una pantalla general para ver el estado del despliegue (Running/Complete/Fail)

crimson_pear_psp3bfmw ✎ ☆ ✓ Completed

Overview Model Explanations (preview) Responsible AI (preview) Metrics Data transformation (preview) Test results (preview) Outputs + logs

Success: Model deployment is successfully triggered

Refresh Deploy Download Explain model View generated code Test model (preview) Register model Cancel

Model summary

Algorithm name
VotingEnsemble

Ensemble details
[View ensemble details](#)

AUC weighted
0.94823 [View all other metrics](#)

Sampling
100.00 % [i](#)

Registered models
[AutoMLa7ddf57fa52:1](#)

Deploy status
[degcd22](#) [Running](#)

Deploy del modelo

Completado el deploy podemos testearlo en la ventana test. Para ello insertamos los parámetros de entrada y nos devuelve en un formato JSON el resultado (feature “y”).

UAB > ml_gcd22 > Endpoints > degcd22

degcd22 ★

Details **Test** Consume Logs

Input data to test endpoint **Test** **Test result**

Select editor type

Form editor JSON editor

data  

age	<input type="text" value="0"/>
job	<input type="text" value="example_value"/>
marital	<input type="text" value="example_value"/>
education	<input type="text" value="example_value"/>
default	<input type="text"/>

Test result

```
{ "Results": [ { "0": "string \"no\""} ] }
```

Deploy del modelo

La ventana *consume* nos proporciona el REST endpoint y el código en diferentes lenguajes de programación para realizar consultas al modelo a través del endpoint

The screenshot shows the 'degcd22' workspace in the Azure ML studio. The 'Consume' tab is selected. Under 'Basic consumption info', the 'REST endpoint' is listed as <http://e3bd3dae-8ef0-4fbd-b652-f477d83ef43c.germanywestcentral.azurecontainer.io/score>. Below this, under 'Consumption option', there is a section for 'Consumption types' with tabs for Python, C#, and R. The Python tab is selected. A code editor displays the following Python code:

```
1 import urllib.request
2 import json
3 import os
4 import ssl
5
6 def allowSelfSignedHttps(allowed):
7     # bypass the server certificate verification on client side
8     if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl, '_create_unverified_context'):
9         ssl._create_default_https_context = ssl._create_unverified_context
10
11 allowSelfSignedHttps(True) # this line is needed if you use self-signed certificate in you
12
13 # Request data goes here
14 # The example below assumes JSON formatting which may be updated
15 # depending on the format your endpoint expects.
16 # Many more options can be set - look up urllib.request for details.
```

Implementación del modelo generado

Regresamos al *workspace* y seleccionamos *Notebooks* para crear un cuaderno Python y llamar al modelo de predicción mediante el uso del código que nos ha proporcionado el deploy

The screenshot shows the Azure AI | Machine Learning Studio interface. The left sidebar has a red box around the 'Notebooks' option under the 'Authoring' section. The main content area is titled 'ml_gcd22' and shows sections for 'Generative AI with Prompt flow', 'Generative AI models', and 'Notebook samples'. A yellow callout box points to the 'Notebooks' section in the 'Generative AI with Prompt flow' area, which contains cards for 'QnA with Your Own Data Using ...', 'Bring Your Own Data QnA', 'Ask Wikipedia', and 'Chat with Wikipedia'. Each card has 'Start' and 'Clone' buttons. The top right corner shows 'Azure for Students' and the workspace name 'ml_gcd22'.

Implementación del modelo generado

Creamos un recurso de ejecución para ejecutar el cuaderno Python

The screenshot shows the Azure AI | Machine Learning Studio interface. The left sidebar has a navigation menu with sections like All workspaces, Home, Model catalog (PREVIEW), Authoring (Notebooks selected), Assets (Data, Jobs, Components, Pipelines, Environments, Models, Endpoints), and Manage. The main area is titled 'Notebooks' and shows a file tree with 'Users' and '2131865'. A central illustration depicts a laptop, a server tower, and a cloud with a circuit board, symbolizing cloud computing. Below the illustration, text reads: 'Notebooks is your space to add, browse, and edit files. You can add files of any type, including Jupyter Notebooks (.ipynb). The files you see here are stored in the workspace file share, and are accessible and shared within the workspace.' A callout box with a yellow background and black text says 'Seleccionamos Create Compute'. At the bottom, there are buttons for '+ Files' and 'Create compute' (the latter is highlighted with a red box), and links for 'View Azure Machine Learning tutorials' and 'View Release Notes to learn more about the latest features'.

Notebooks is your space to add, browse, and edit files.

You can add files of any type, including Jupyter Notebooks (.ipynb). The files you see here are stored in the workspace file share, and are accessible and shared within the workspace.

In order to run notebooks and scripts, you must connect to a Learning compute resource. Once a notebook or terminal is connected, it can access all workspace assets including experiment details, data, and more. [Learn more](#)

+ Files Create compute

[View Azure Machine Learning tutorials](#)
[View Release Notes to learn more about the latest features](#)

Seleccionamos Create Compute

Implementación del modelo generado

Create compute instance X

1 Required Settings

2 Advanced Settings
optional

Configure required settings

Select the name and virtual machine size you would like to use for your compute instance. Please note that a compute instance can not be shared. It can only be used by a single assigned user. By default, it will be assigned to the creator and you can change this to a different user in the advanced settings section.

Compute name *

deployMLgcd

Location i

germanywestcentral

Virtual machine type i

CPU GPU

Virtual machine size i

Select from recommended options Select from all options

Name ↑	Category	Workload types	Available quota i	Cost i
<input type="radio"/> Standard_DS11_v2 2 cores, 14GB RAM, 28GB storage	Memory optimized	Development on Notebooks (or other IDE) and light weight testing	6 cores	\$0.17/hr
<input type="radio"/> Standard_DS11_v2 2 cores, 14GB RAM, 28GB storage	Machine learning	Classical ML model training on small datasets	6 cores	\$0.27/hr

Create

Cancel

1. Asignamos un nombre al recurso de cómputo

2. Seleccionamos crear

Implementación del modelo generado

UAB > ml_gcd22 > Notebooks

Notebooks

Files Samples

Users
2131865



Notebooks is your space to add, browse, and edit files.

You can add files of any type, including Jupyter Notebooks (.ipynb). The files you see here are stored in the workspace file share, and are accessible and shared within the workspace.

In order to run notebooks and scripts, you must connect to an Azure Machine Learning compute resource. Once a notebook or terminal is connected, you can access all workspace assets including experiment details, data, models, and more. [Learn more](#)

Seleccionamos *Files* para crear un cuaderno *Notebook* nuevo

+ Files Terminal

View Azure Machine Learning tutorials

Implementación del modelo generado

Asignamos un nombre al cuaderno y apretamos la tecla *intro*

Create new file

File location

Users/2131865 [Edit location](#)

File name *

Untitled.ipynb

File type

Notebook (*.ipynb)

Overwrite if already exists

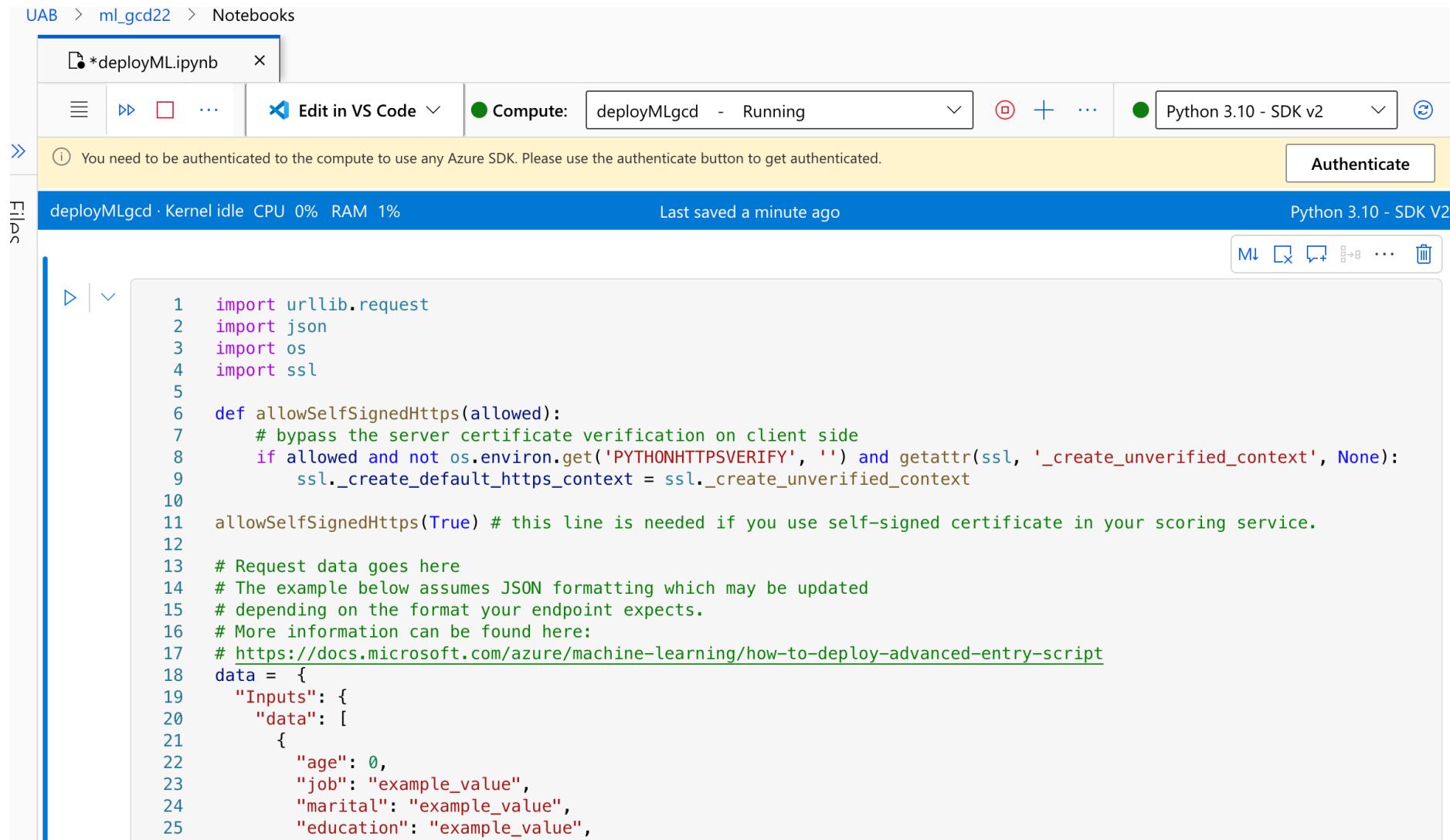
+ Files Terminal

View Azure Machine Learning tutorials

View Release Notes to learn more about the latest features

Implementación del modelo generado

Copiamos el código Python generado por el Deploy en el cuaderno Notebook y damos valores a los parámetros de entrada del modelo



The screenshot shows an Azure ML Studio interface with a notebook titled "deployML.ipynb". The notebook is running on a "Python 3.10 - SDK v2" compute environment. A message at the top indicates that authentication is required to use the Azure SDK. The code in the notebook is as follows:

```
1 import urllib.request
2 import json
3 import os
4 import ssl
5
6 def allowSelfSignedHttps(allowed):
7     # bypass the server certificate verification on client side
8     if allowed and not os.environ.get('PYTHONHTTPSVERIFY', '') and getattr(ssl, '_create_unverified_context', None):
9         ssl._create_default_https_context = ssl._create_unverified_context
10
11 allowSelfSignedHttps(True) # this line is needed if you use self-signed certificate in your scoring service.
12
13 # Request data goes here
14 # The example below assumes JSON formatting which may be updated
15 # depending on the format your endpoint expects.
16 # More information can be found here:
17 # https://docs.microsoft.com/azure/machine-learning/how-to-deploy-advanced-entry-script
18 data = {
19     "Inputs": {
20         "data": [
21             {
22                 "age": 0,
23                 "job": "example_value",
24                 "marital": "example_value",
25                 "education": "example_value",
26             }
27         ]
28     }
29 }
```

Implementación del modelo generado

Ejecutamos el modelo para obtener el resultado

The screenshot shows a Jupyter Notebook interface with the following details:

- Path:** UAB > ml_gcd22 > Notebooks
- File:** *deployML.ipynb
- Compute:** deployMLgcd - Running
- Kernel:** Python 3.10 - SDK v2
- Notebook Header:** You need to be authenticated to the compute to use any Azure SDK. Please use the authenticate button to get authenticated. **Authenticate**
- Code Cell:** deployMLgcd · Kernel idle CPU 0% RAM 1%

```
59     response = urllib.request.urlopen(req)
60
61     result = response.read()
62     print(result)
63 except urllib.error.HTTPError as error:
64     print("The request failed with status code: " + str(error.code))
65
66     # Print the headers – they include the request ID and the timestamp, which are useful for debugging the failure
67     print(error.info())
68     print(error.read().decode("utf8", 'ignore'))
69
```
- Output:** [1] ✓ <1 sec
... b'{"Results": ["no"]}'