

Diplomado en ML Cloud - UCB CBBA

Módulo 4: Machine Learning Cloud MLOps

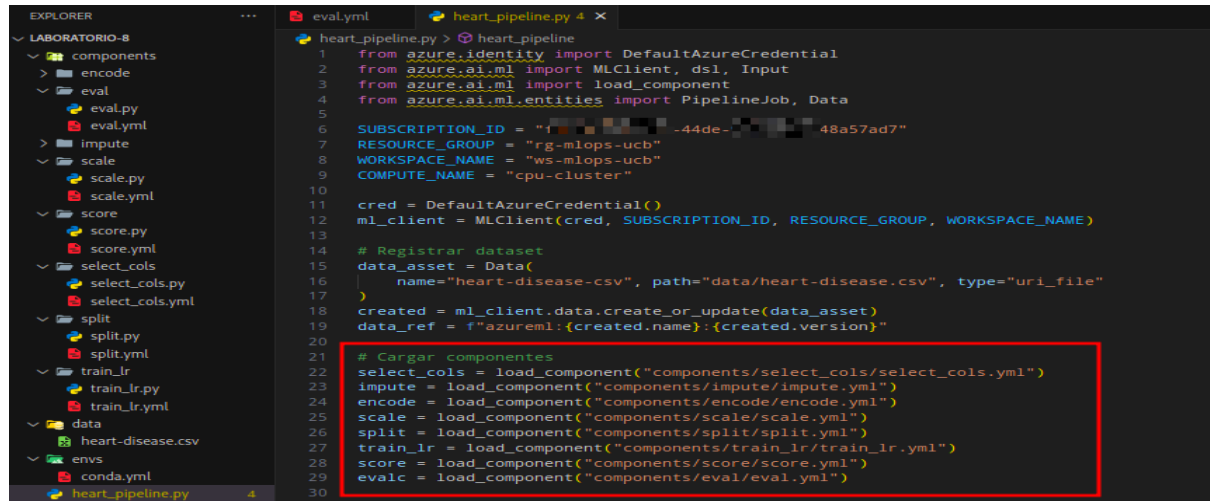
Docente: Ing. Mauricio Alejandro Quezada

Estudiante: Jose Carlos Iriarte

Fecha : Septiembre del 2025

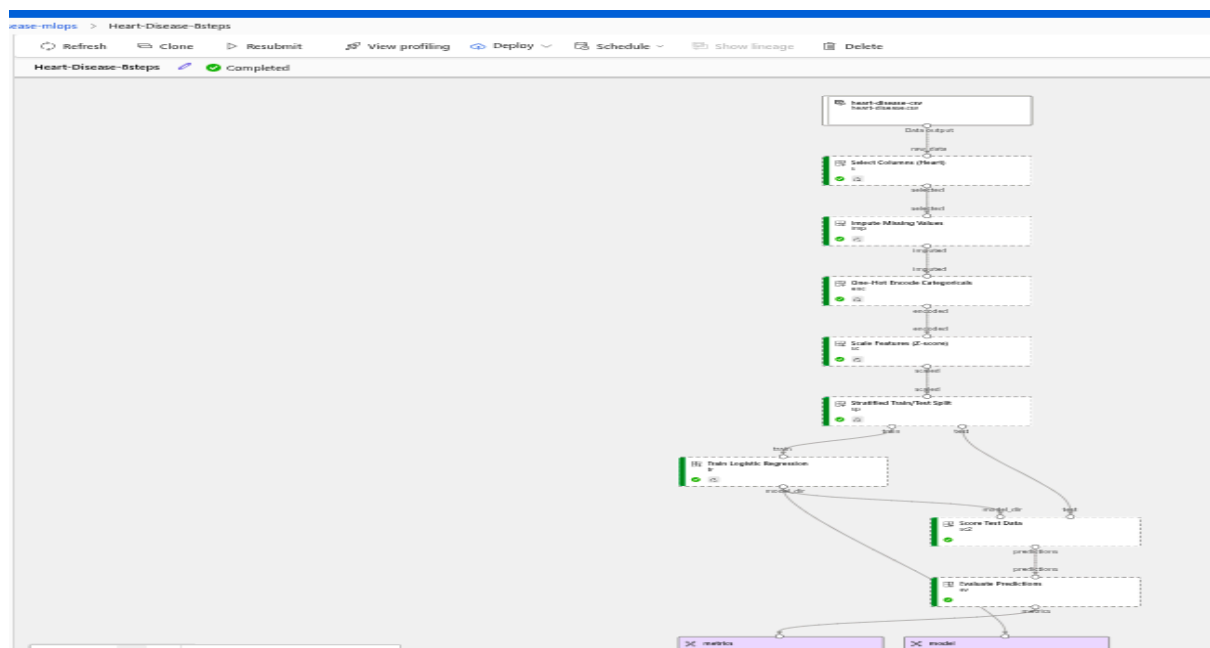
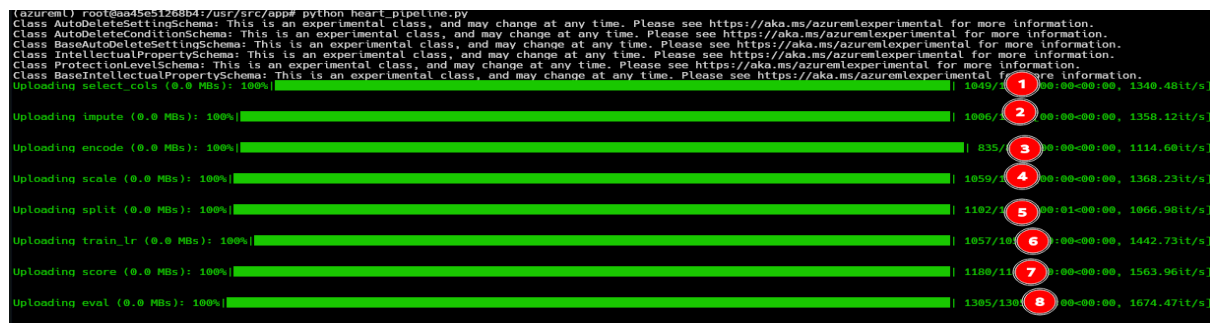
Laboratorio 9

Capturas del DAG con 8 componentes



The screenshot shows the VS Code interface with the 'heart_pipeline.py' file open. The Explorer view on the left shows a project structure for 'LABORATORIO-8' with components like 'encode', 'eval', 'impute', 'scale', 'split', 'train_lr', and 'data'. The main editor shows the code for 'heart_pipeline.py', which includes imports for Azure ML, subscription details, and a section for loading components. A red box highlights the component loading section.

```
1 from azure.identity import DefaultAzureCredential
2 from azure.ai.ml import MLClient, dsl, Input
3 from azure.ai.ml import load_component
4 from azure.ai.ml.entities import PipelineJob, Data
5
6 SUBSCRIPTION_ID = "f1f1f1f1-f1f1-f1f1-f1f1-f1f1f1f1f1f1"
7 RESOURCE_GROUP = "rg-mlops-ucb"
8 WORKSPACE_NAME = "ws-mlops-ucb"
9 COMPUTE_NAME = "cpu-cluster"
10
11 cred = DefaultAzureCredential()
12 ml_client = MLClient(cred, SUBSCRIPTION_ID, RESOURCE_GROUP, WORKSPACE_NAME)
13
14 # Registrar dataset
15 data_asset = Data(
16     name="heart-disease-csv", path="data/heart-disease.csv", type="uri_file"
17 )
18 created = ml_client.data.create_or_update(data_asset)
19 data_ref = f"azureml:{created.name}:{created.version}"
20
21 # Cargar componentes
22 select_cols = load_component("components/select_cols/select_cols.yml")
23 impute = load_component("components/impute/impute.yml")
24 encode = load_component("components/encode/encode.yml")
25 scale = load_component("components/scale/scale.yml")
26 split = load_component("components/split/split.yml")
27 train_lr = load_component("components/train_lr/train_lr.yml")
28 score = load_component("components/score/score.yml")
29 evalc = load_component("components/eval/eval.yml")
30
```



ws-mlops-ucb > Jobs > heart-disease-mflow > Heart-Disease-Mflow-Pipeline > train_step

train_step Completed

Overview Metrics Images Child jobs Outputs + logs Code Monitoring

Duration: 29.00s
Compute duration: 29.00s
Name: ac047229-4cad-4daa-a374-e8ec50102972
Command: python train.py --data \$AZUREML_DATAREFERENCE_raw_data --model_dir DatasetOutputConfig:model_dir --test_out DatasetOutputConfig:test_data --register_name \$AZUREML_PARAMETER_register_name
Environment: CIV2AnonymousEnvironment:51c98874185ac434bce8876520608230693d9920110bdac0815d09c5591278ac
Registered models: heart-disease-logreg-mflow:1
See all properties
Raw JSON
CPU memory: 838,839 MiB-s

Output name: model_dir
Data asset: azureml_ac047229-4cad-4daa-a374-e8ec50102972_output_data_model_dir:1
Asset URI: azureml:azureml_ac047229-4cad-4daa-a374-e8ec50102972_output_data_model_dir:1
Output name: test_data
Data asset: azureml_ac047229-4cad-4daa-a374-e8ec50102972_output_data_test_data:1
Asset URI: azureml:azureml_ac047229-4cad-4daa-a374-e8ec50102972_output_data_test_data:1

Tags: No tags

Params: algorithm: LogisticRegression max_iter: 1000 n_features: 13 n_test: 205 n_train: 820

Metrics: No data

Compute: Target: mlops Instance count: 1
Compute type: amlicompute

Captura de split estratificado y environment

ws-mlops-ucb > Jobs > heart-disease-mlops > Heart-Disease-8steps

code name, comment or comp... Refresh Clone Resubmit View profiling Deploy Schedule Show lineage Delete Share Job overview

Heart-Disease-8steps Completed

Train Logistic Regression

Overview Settings Outputs + logs Metrics Child jobs Images Code Monitoring

Refresh Register model Debug and monitor Compare

Properties: Status: Completed Created on: Sep 9, 2025 9:19 PM Start time: Sep 9, 2025 9:19 PM Compute duration: 0s Compute target: cpu-cluster Name: c350fdc7-1c16-4606-801d-cefb21f096cd

Inputs: Input name: train Data asset: azureml_f55c7740-342f-457f-b9f2-e3836bbad971_output_data_train:1 Asset URI: azureml:azureml_f55c7740-342f-457f-b9f2-e3836bbad971_output_data_train:1

Outputs: Output name: model_dir Data asset: azureml_871f23a9-250f-470f-8e20-b8f5950b12ec_output_data_model_dir:1 Asset URI: azureml:azureml_871f23a9-250f-470f-8e20-b8f5950b12ec_output_data_model_dir:1

Tags: No tags

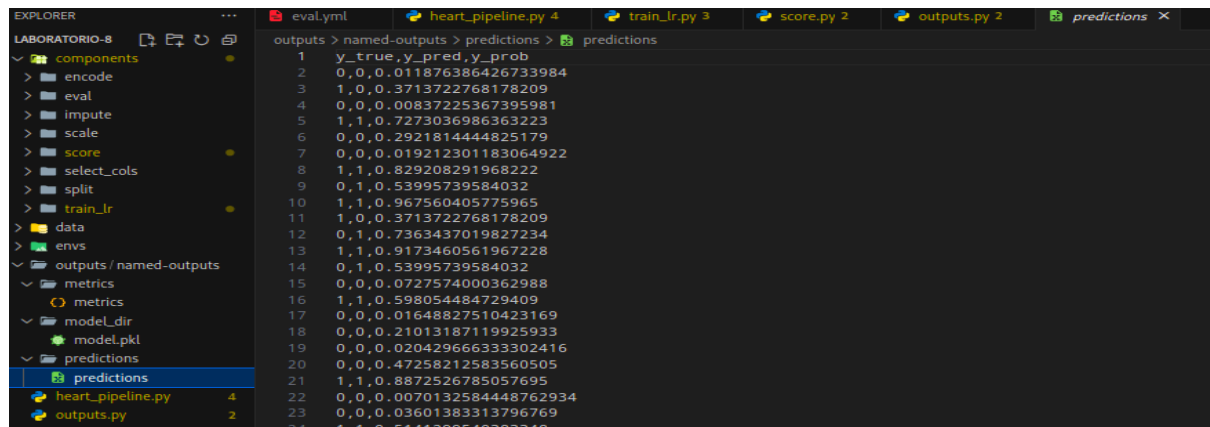
Artefactos: modelo, prediction.csv, metrics.json

EXPLORER LABORATORIO-8 components encode eval impute scale score select_cols split train_lr data outputs/named-outputs metrics model_dir predictions heart_pipeline.py outputs.py

```
1 from azure.ai.ml import MLClient
2 from azure.identity import DefaultAzureCredential
3
4 ml_client = MLClient(
5     DefaultAzureCredential(),
6     subscription_id="f4c015-cf6a-4a57ad7",
7     resource_group_name="rg-mlops-ucb",
8     workspace_name="ws-mlops-ucb",
9 )
10
11 job_name = "24b6a365-62d18b-c944e1ed6c73"
12 print(f"Descargando outputs del job {job_name}")
13 ml_client.jobs.download(name=job_name, output_name="metrics", download_path="outputs/")
14
15 job_name = "f20eddf4-80ab16c178c58119"
16 ml_client.jobs.download(
17     name=job_name, output_name="predictions", download_path="outputs/"
18 )
19
20 job_name = "c350fdc7-1c16-4606-801d-cefb21f096cd"
21 ml_client.jobs.download(
22     name=job_name, output_name="model_dir", download_path="outputs/"
23 )
24
25 print("Modelo descargado en ./outputs/model_dir/")
26 print(f"Archivo guardado en ./outputs/predictions/")
```

EXPLORER LABORATORIO-8 components encode eval impute scale score select_cols split train_lr data outputs/named-outputs metrics

```
1
2 "accuracy": 0.8097560975609757,
3 "precision": 0.7619047619047619,
4 "recall": 0.9142857142857143,
5 "f1": 0.8311688311688312,
6 "roc_auc": 0.9297142857142857
7
```



```
1 y_true,y_pred,y_prob
2 0,0,0.011876386426733984
3 1,0,0.3713722768178209
4 0,0,0.00837225367395981
5 1,1,0.7273036986363223
6 0,0,0.2921814444825179
7 0,0,0.019212301183064922
8 1,1,0.829208291968222
9 0,1,0.53995739584032
10 1,1,0.967560405775965
11 1,0,0.3713722768178209
12 0,1,0.7363437019827234
13 1,1,0.9173460561967228
14 0,1,0.53995739584032
15 0,0,0.0727574000362988
16 1,1,0.598054484729409
17 0,0,0.01648827510423169
18 0,0,0.21013187119925933
19 0,0,0.02042966633302416
20 0,0,0.47258212583560505
21 1,1,0.8872526785057695
22 0,0,0.0070132584448762934
23 0,0,0.03601383313796769
24 1,1,0.5141299540383348
```

El modelo de Regresión Logística entrenado sobre el dataset de Heart Disease obtuvo los siguientes indicadores:

- Accuracy: 0.81
- Precision: 0.76
- Recall (Sensibilidad): 0.91
- F1-score: 0.83
- ROC-AUC: 0.93

Interpretación

El modelo presenta un buen equilibrio entre precisión y recall, destacando su alta sensibilidad (91 %), lo que significa que detecta la gran mayoría de los pacientes con enfermedad. El área bajo la curva (AUC=0.93) confirma su capacidad de discriminar entre casos positivos y negativos.

Aunque se observan algunos falsos positivos, este comportamiento es aceptable en un contexto clínico, pues resulta preferible clasificar erróneamente a un paciente sano como enfermo antes que dejar sin detectar un caso real.

Conclusión

El modelo muestra un desempeño sólido y clínicamente relevante, siendo adecuado como herramienta de apoyo diagnóstico, aunque se recomienda seguir ajustando umbrales o explorar modelos adicionales para mejorar la precisión.