



# **Ejemplo de operaciones y tareas de Kubeflow**

NetApp Solutions

NetApp  
June 24, 2024

This PDF was generated from [https://docs.netapp.com/es-es/netapp-solutions/ai/aicp\\_provision\\_a\\_jupyter\\_notebook\\_workspace\\_for\\_data\\_scientist\\_or\\_developer\\_use.html](https://docs.netapp.com/es-es/netapp-solutions/ai/aicp_provision_a_jupyter_notebook_workspace_for_data_scientist_or_developer_use.html) on June 24, 2024. Always check docs.netapp.com for the latest.

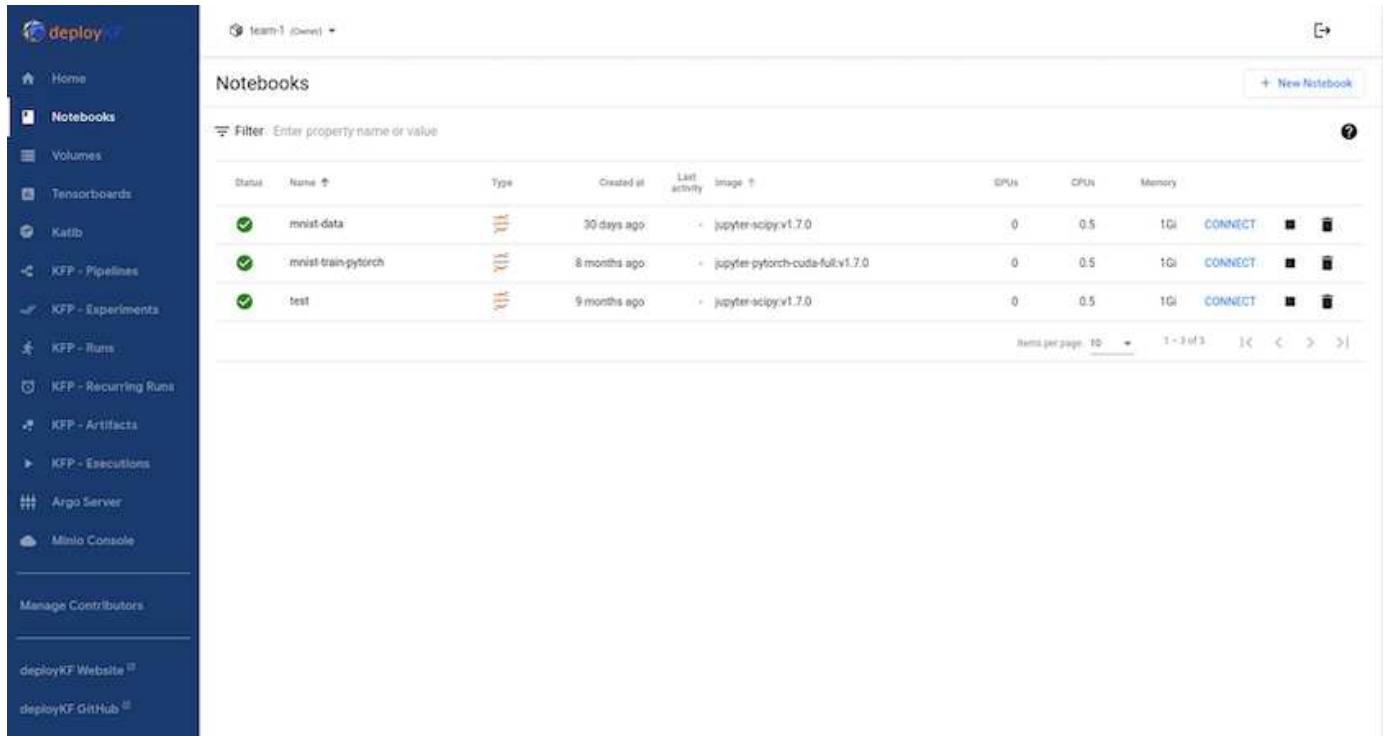
# Tabla de contenidos

- Ejemplo de operaciones y tareas de Kubeflow ..... 1
  - Aprovisione un espacio de trabajo para portátiles Jupyter para uso científico de datos o desarrollador ... 1
  - Use el kit de herramientas NetApp DataOps con Kubeflow. .... 1
  - Flujo de trabajo de ejemplo: Formación de un modelo de reconocimiento de imágenes mediante Kubeflow y el kit de herramientas de DataOps de NetApp ..... 2

# Ejemplo de operaciones y tareas de Kubeflow

## Aprovisione un espacio de trabajo para portátiles Jupyter para uso científico de datos o desarrollador

Kubeflow es capaz de suministrar rápidamente nuevos servidores Jupyter Notebook para actuar como espacios de trabajo de científicos de datos. Para obtener más información acerca de Jupyter Notebooks dentro del contexto de Kubeflow, consulte ["Documentación oficial de Kubeflow"](#).



The screenshot shows the 'Notebooks' section of the Kubeflow deployment console. A sidebar on the left contains navigation links: Home, Notebooks, Volumes, Tensorboards, Katib, KFP - Pipelines, KFP - Experiments, KFP - Runs, KFP - Recurring Runs, KFP - Artifacts, KFP - Executions, Argo Server, Minio Console, Manage Contributors, deployKF Website, and deployKF GitHub. The main panel displays a table of notebooks with columns for Status, Name, Type, Created at, Last activity, Image, CPU, GPU, and Memory. Three notebooks are listed: 'mnist-data', 'mnist-train-pytorch', and 'test'. Each notebook has a 'CONNECT' button and a trash icon. A filter bar at the top of the table allows searching by property name or value. The bottom of the table shows pagination controls: 'Items per page: 10', '1 - 3 of 3', and navigation arrows.

Status	Name	Type	Created at	Last activity	Image	CPU	GPU	Memory	
✓	mnist-data	Jupyter Notebook	30 days ago		jupyter-scipy:v1.7.0	0	0.5	1Gi	<a href="#">CONNECT</a>
✓	mnist-train-pytorch	Jupyter Notebook	8 months ago		jupyter-pytorch-cuda-full:v1.7.0	0	0.5	1Gi	<a href="#">CONNECT</a>
✓	test	Jupyter Notebook	9 months ago		jupyter-scipy:v1.7.0	0	0.5	1Gi	<a href="#">CONNECT</a>

## Use el kit de herramientas NetApp DataOps con Kubeflow

La ["Kit de herramientas para la ciencia de datos de NetApp para Kubernetes"](#) Se puede utilizar junto con Kubeflow. El uso del kit de herramientas para la ciencia de datos de NetApp con Kubeflow ofrece las siguientes ventajas:

- Los científicos de datos pueden llevar a cabo operaciones avanzadas de gestión de datos de NetApp, como la creación de copias Snapshot y clones, directamente desde un portátil de Jupyter.
- Las operaciones avanzadas de gestión de datos de NetApp, como la creación de snapshots y clones, se pueden incorporar en flujos de trabajo automatizados mediante el marco de Kubeflow Pipelines.

Consulte la ["Ejemplos de Kubeflow"](#) Sección dentro del repositorio de Data Science Toolkit de NetApp, GitHub para obtener información sobre el uso del kit de herramientas con Kubeflow.

# Flujo de trabajo de ejemplo: Formación de un modelo de reconocimiento de imágenes mediante Kubeflow y el kit de herramientas de DataOps de NetApp

En esta sección se describen los pasos que deben seguirse para entrenar y poner en marcha una red neuronal para el reconocimiento de imágenes mediante Kubeflow y el kit de herramientas NetApp DataOps. Esta información pretende servir de ejemplo para mostrar un trabajo de formación que incorpora almacenamiento NetApp.

## Requisitos previos

Cree un Dockerfile con las configuraciones necesarias para usar en los pasos de tren y prueba dentro de la tubería de Kubeflow.

Aquí hay un ejemplo de un Dockerfile -

```
FROM pytorch/pytorch:latest
RUN pip install torchvision numpy scikit-learn matplotlib tensorboard
WORKDIR /app
COPY . /app
COPY train_mnist.py /app/train_mnist.py
CMD ["python", "train_mnist.py"]
```

Dependiendo de sus requisitos, instale todas las bibliotecas y paquetes necesarios para ejecutar el programa. Antes de entrenar el modelo de aprendizaje automático, se asume que ya tiene un despliegue de Kubeflow en funcionamiento.

## Entrena un pequeño NN en datos MNIST usando PyTorch y Kubeflow Pipelines

Utilizamos el ejemplo de una pequeña red neuronal entrenada en datos MNIST. El conjunto de datos MNIST consta de imágenes escritas a mano de dígitos del 0 al 9. Las imágenes tienen un tamaño de 28x28 píxeles. El conjunto de datos se divide en 60.000 imágenes de tren y 10.000 imágenes de validación. La red neuronal utilizada para este experimento es una red de avance de 2 capas. El entrenamiento se ejecuta mediante Tuberías Kubeflow. Consulte la documentación ["aquí"](#) si quiere más información. Nuestro Pipeline Kubeflow incorpora la imagen de docker de la sección Requisitos previos.

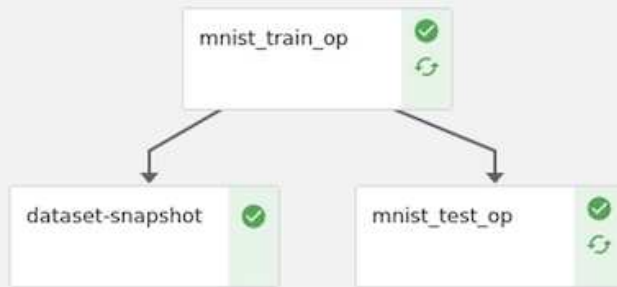
[←](#)  mnist\_pipeline 2024-04-03 15-57-35**Graph**

Run output

Config

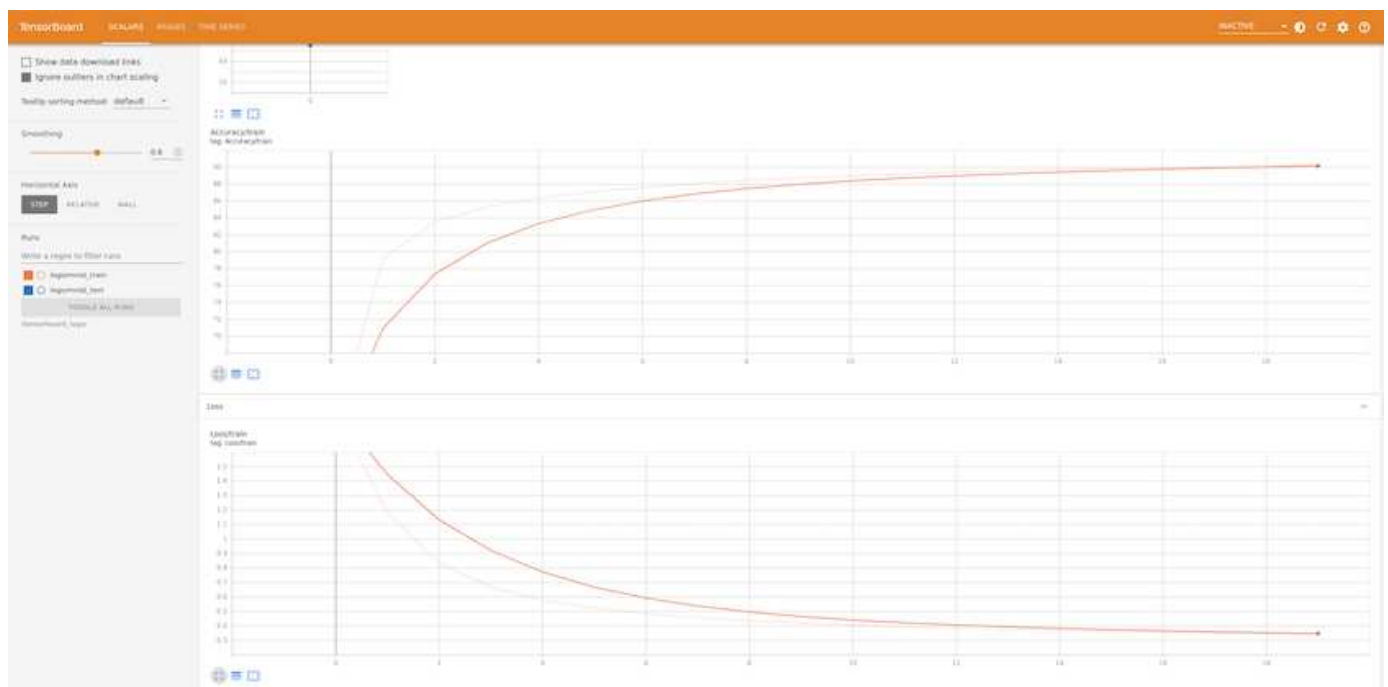


Simplify Graph



## Visualiza los resultados usando Tensorboard

Una vez entrenado el modelo, podemos visualizar los resultados usando Tensorboard. "Tablero de placas" Está disponible como una característica en el Panel de control de Kubeflow. Puede crear un tablero de tenencia personalizado para el trabajo. Un ejemplo a continuación muestra el trazado de la precisión del entrenamiento frente a número de épocas y pérdida de entrenamiento con respecto a número de épocas.



## Experimenta con Hiperparámetros usando Katib

"Katib" Es una herramienta dentro de Kubeflow que se puede utilizar para experimentar con los hiperparámetros del modelo. Para crear un experimento, defina primero una métrica/objetivo deseado. Esta es generalmente la precisión de la prueba. Una vez definida la métrica, elija hiperparámetros con los que desea jugar (optimizador/learning\_rate/número de capas). Katib realiza un barrido de hiperparámetros con los valores definidos por el usuario para encontrar la mejor combinación de parámetros que satisfagan la métrica deseada. Puede definir estos parámetros en cada sección de la interfaz de usuario. Alternativamente, puede definir un archivo **YAML** con las especificaciones necesarias. A continuación se muestra una ilustración de un experimento Katib -

**Experiment details** DELETE

**Objective**

Name	Validation-accuracy
Type	maximize
Goal	0.9
Additional metrics	Train-accuracy

**Trials**

Max failed trials	3
Max trials	12
Parallel trials	3

**Parameters**

lr: Parameter type: double, Min: 0.01, Max: 0.03

num-layers: Parameter type: int, Min: 1, Max: 64

optimizer: Parameter type: categorical, sgd, adam, ftrl

**Algorithm**

Name: grid

**Metrics collector**

Collector type: File

**Experiment details** DELETE

**Overview** **Trials** **Details** **YAML**

**Name**: mnist-pytorch

**Status**: Experiment is running

**Best trial**: No optimal trial yet

**Best trial's params**: No optimal trial yet

**Best trial performance**: No optimal trial yet

**User defined goal**: Validation-accuracy > 0.9

**Running trials**: 3

**Failed trials**: 0

**Succeeded trials**: 0

**Experiment Conditions**

Filter: Enter property name or value

## Utilice Instantáneas de NetApp para guardar datos y rastrearlos

Durante el entrenamiento del modelo, puede que desee guardar una instantánea del conjunto de datos de entrenamiento para la trazabilidad. Para ello, podemos agregar un paso de snapshot a la canalización, tal como se muestra a continuación. Para crear la instantánea, podemos usar la ["Kit de herramientas Data OPS de NetApp para Kubernetes"](#).

```
@dsl.pipeline(
    name = 'MNIST Classification Pipeline',
    description = 'Train a simple NN for classification'
)
def mnist_pipeline():
    mnist_train_task = mnist_train_op()
    mnist_train_task.apply(
        kfp.onprem.mount_pvc('mnist-data', 'mnist-data-vol', '/mnt/data/')
    )

    mnist_test_task = mnist_test_op()
    mnist_test_task.apply(
        kfp.onprem.mount_pvc('mnist-data', 'mnist-data-vol', '/mnt/data/')
    )

    volume_snapshot_name = 'mnist-pytorch-snapshot'
    dataset_snapshot = dsl.ContainerOp(
        name='dataset-snapshot',
        image='python:3.9',
        command=['/bin/bash', '-c'],
        arguments=[
            'python3 -m pip install netapp-dataops-k8s && \
            echo "" + volume_snapshot_name + "" > /volume_snapshot_name.txt && \
            netapp_dataops_k8s_cli.py create volume-snapshot --pvc-name=" + "mnist-data" + " --snapshot-name=" + str(volume_snapshot_name) + " --namespace={{workflow.namespace}}',
            file_outputs={'volume_snapshot_name': '/volume_snapshot_name.txt'}
        ]
    )
    mnist_test_task.after(mnist_train_task)
    dataset_snapshot.after(mnist_train_task)
```

Consulte la ["Ejemplo del kit de herramientas NetApp DataOps para Kubeflow"](#) si quiere más información.

## Información de copyright

Copyright © 2024 NetApp, Inc. Todos los derechos reservados. Imprimido en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPCIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

## Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en <http://www.netapp.com/TM> son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.