

Creación automática de modelos de machine learning

En este tutorial, aprenderá a utilizar el Piloto automático de Amazon SageMaker para crear, entrenar y ajustar automáticamente un modelo de machine learning (ML) e implementar el modelo a fin de hacer predicciones.

El Piloto automático de Amazon SageMaker elimina las arduas tareas de creación de modelos de ML al ayudarlo a crear, entrenar y ajustar automáticamente el mejor modelo de ML en función de sus datos. Con el Piloto automático de SageMaker, usted solo proporciona un conjunto de datos tabular y selecciona la columna de destino para predecir. El Piloto automático de SageMaker explora sus datos, selecciona los algoritmos relevantes para su tipo de problema, prepara los datos para el entrenamiento del modelo, prueba una variedad de modelos y selecciona el de mejor rendimiento. A continuación, puede implementar uno de los modelos candidatos o iterarlos más para mejorar la calidad de la predicción.

En esta guía, aprenderá a hacer lo siguiente:

- Crear un experimento de entrenamiento con el Piloto automático de Amazon SageMaker.
- Explorar las diferentes etapas del experimento de entrenamiento.
- Identificar e implementar el modelo con mejor rendimiento del experimento de entrenamiento.
- Hacer predicciones con el modelo implementado.

Para este flujo de trabajo, utilizará un conjunto de datos de reclamaciones de seguros de automóviles generado de forma sintética. Las entradas sin procesar son dos tablas de datos de seguros: una tabla de reclamaciones y una tabla de clientes. La tabla de reclamaciones tiene una columna de fraude que indica si una reclamación fue fraudulenta o no. A los efectos de este tutorial, hemos seleccionado una pequeña porción del conjunto de datos. Sin embargo, puede seguir los mismos pasos de este tutorial para procesar grandes conjuntos de datos.

Paso 1: configurar el dominio de Amazon SageMaker Studio

Una cuenta de AWS solo puede tener un dominio de SageMaker Studio por región de AWS. Si ya tiene un dominio de SageMaker Studio en la región Este de EE. UU. (Norte de Virginia), siga la guía de [configuración de SageMaker Studio](#) para adjuntar las políticas de AWS IAM requeridas a su cuenta de SageMaker Studio; luego omita el paso 1 y continúe directamente con el paso 2.

Si no tiene un dominio de SageMaker Studio existente, continúe con el paso 1 a fin de ejecutar una plantilla de AWS CloudFormation que cree un dominio de SageMaker Studio y agregue los permisos necesarios para el resto de este tutorial.

Elija el [enlace de la pila de AWS CloudFormation](#). Este enlace abre la consola de AWS CloudFormation y crea su dominio de SageMaker Studio y un usuario denominado studio-user. También agrega los permisos necesarios a su cuenta de SageMaker Studio. En la consola de CloudFormation, confirme que Este de EE. UU. (Norte de Virginia) sea la región que se muestra en la esquina superior derecha. El nombre de la pila debe ser CFN-SM-IM-Lambda-Catalog y no debe cambiarse. Esta pila tarda unos 10 minutos en crear todos los recursos.

Esta pila asume que ya tiene una VPC pública configurada en su cuenta. Si no tiene una VPC pública, consulte [VPC con una única subred pública](#) para obtener información sobre cómo crear una VPC pública.

The screenshot shows the 'Quick create stack' interface in the AWS CloudFormation console. The left sidebar contains navigation links for Stacks, StackSets, Exports, Designer, Registry, and Feedback. The main content area is titled 'Quick create stack' and includes a 'Template' section with a URL and description, a 'Stack name' input field (highlighted with a red box and containing 'CFN-SM-IM-Lambda-Catalog'), and a 'Parameters' section with fields for 'DomainName' (StudioDomain) and 'UserProfileName' (studio-user).

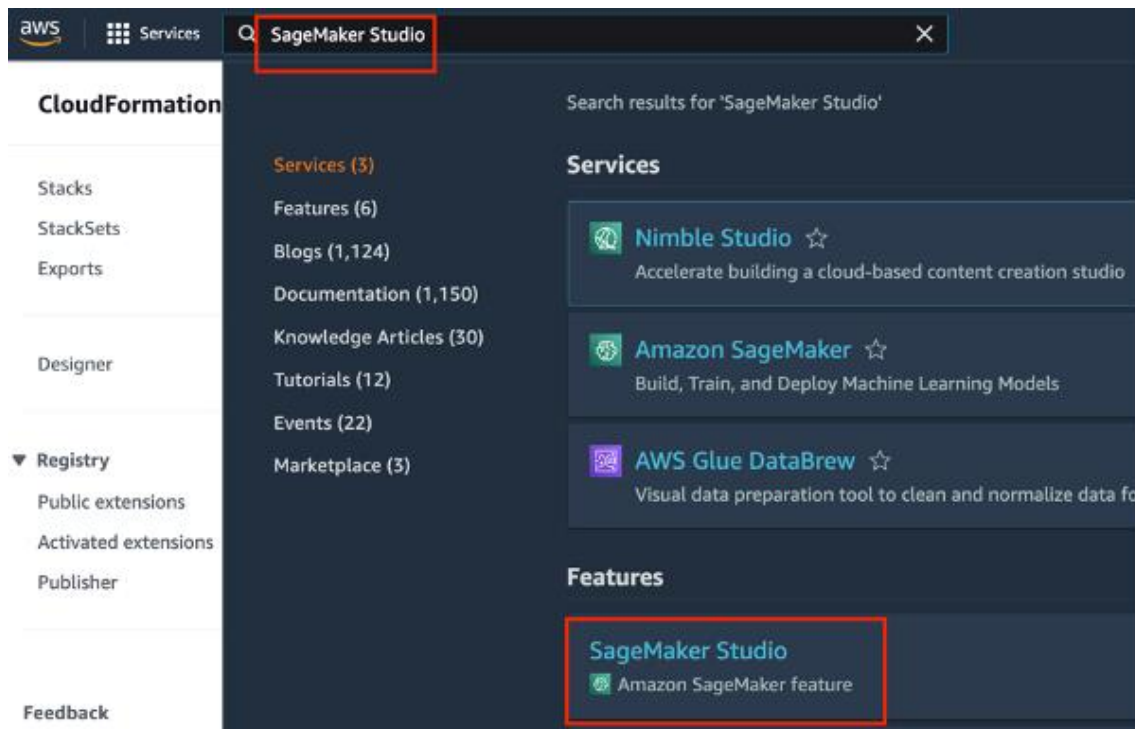
Seleccione I acknowledge that AWS CloudFormation might create IAM resources (Acepto que AWS CloudFormation podría crear recursos de IAM) y luego elija Create stack (Crear pila).

The screenshot shows the 'Capabilities' section of the AWS CloudFormation console. It contains a warning message about IAM resources and a checkbox labeled 'I acknowledge that AWS CloudFormation might create IAM resources.' which is checked and highlighted with a red box. At the bottom, there are three buttons: 'Cancel', 'Create change set', and 'Create stack' (highlighted with a red box).

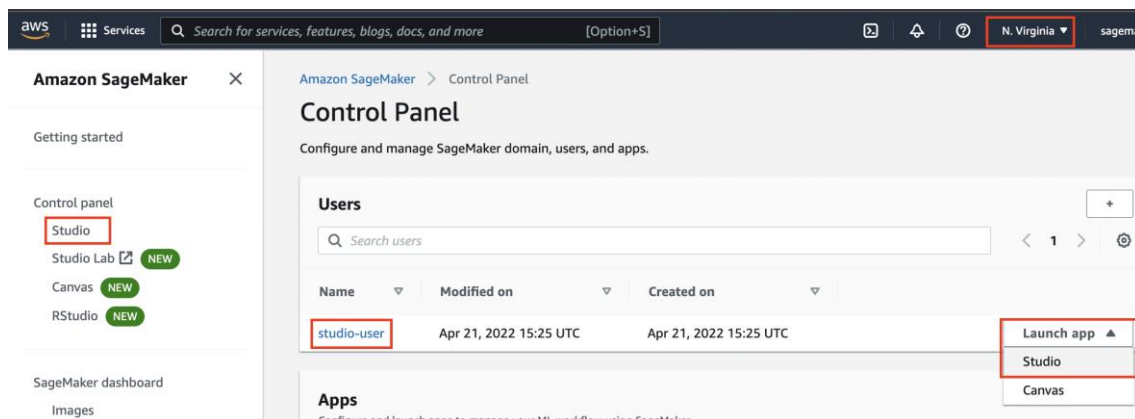
En el panel de CloudFormation, elija Stacks (Pilas). Cuando se crea la pila, su estado debería cambiar de CREATE_IN_PROGRESS a CREATE_COMPLETE.

The screenshot shows the 'Stacks' panel in the AWS CloudFormation console. The left sidebar has 'Stacks' highlighted with a red box. The main content area shows a table of stacks. The first stack is 'CFN-SM-IM-Lambda-catalog' with a status of 'CREATE_COMPLETE' (highlighted with a red box) and a creation time of '2022-06-14 18:45:17 UTC-0400'.

Ingrese SageMaker Studio en la barra de búsqueda de la consola de CloudFormation y luego seleccione SageMaker Studio.



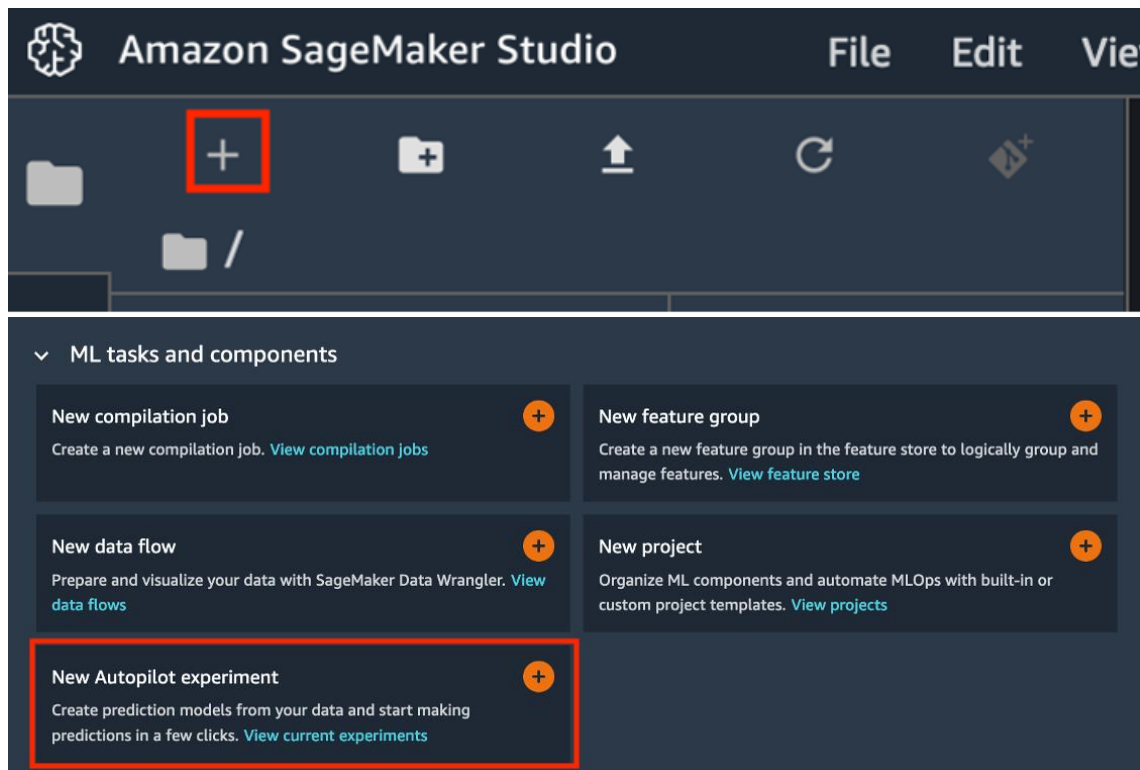
Elija US East (N. Virginia) (Este de EE.UU. [Norte de Virginia]) de la lista desplegable Region (Región) en la esquina superior derecha de la consola de SageMaker. En Launch app (Lanzar aplicación), seleccione Studio para abrir SageMaker Studio con el perfil *studio-user*.



Paso 2: iniciar un nuevo experimento del Piloto automático de SageMaker

Desarrollar y probar una gran cantidad de modelos candidatos es crucial para los proyectos de machine learning (ML). El Piloto automático de Amazon SageMaker brinda diferentes modelos candidatos y elige automáticamente el mejor de ellos en función de sus datos. En este paso, configurará un experimento del Piloto automático de SageMaker para predecir el éxito de una campaña de *marketing* de servicios financieros. Este conjunto de datos representa una campaña de *marketing* realizada por una importante institución de servicios financieros para promover la inscripción de certificados de depósito.

Para iniciar un nuevo experimento del Piloto automático de SageMaker, haga clic en el ícono “+” a fin de acceder a una nueva ventana de inicio. En la ventana de inicio, desplácese hacia abajo hasta ML tasks and components (Tareas y componentes de ML). Haga clic en el ícono + para New Autopilot experiment (Nuevo experimento del piloto automático).



A continuación, ingrese un nombre para el experimento. Haga clic en el cuadro Experiment name (Nombre del experimento) y escriba *autopilot-experiment* como nombre.

The screenshot shows the 'Create an Autopilot experiment' form in Amazon SageMaker Studio. The form has a title 'Create an Autopilot experiment' and a subtitle 'When you create an Autopilot experiment, Amazon SageMaker analyzes your data and creates a notebook with candidate model definitions. This notebook provides visibility into how models are selected, trained, and tuned.' Below the subtitle, there's a section titled 'Basic settings'. This section contains several input fields and a toggle switch. The fields are: 'Experiment name' (with a placeholder 'Enter experiment name'), 'S3 bucket name' (with a placeholder 'Select or type bucket name'), 'Dataset file name' (with a placeholder 'Select...'), 'Output data location (S3 bucket)' (with a placeholder 'Select or type bucket name'), 'Dataset directory name' (with a placeholder 'Select...'), and 'Target' (with a placeholder 'Select your target column'). There's also a toggle switch for 'Is your S3 input a manifest file?' with 'Off' selected. At the bottom, there's a checkbox for 'Auto deploy'.

A continuación, conectará el experimento con los datos que se preparan en S3. Haga clic en el cuadro Enter S3 bucket location (Ingresar la ubicación del *bucket* de S3). En el cuadro S3 bucket address (Dirección del *bucket* de S3), pegue la siguiente ruta de S3: *s3://sagemaker-sample-files/datasets/tabular/uci_bank_marketing/bank-additional-full.csv*.

Deje la opción del archivo de manifiesto establecida en Off (Desactivado). En el menú desplegable Target (Objetivo), seleccione y como la característica objetivo que nuestro modelo intentará predecir.

Basic settings

Experiment name ⓘ
autopilot-experiment

Connect your data ⓘ

☐ Find S3 bucket ☒ Enter S3 bucket location

S3 bucket address ⓘ
tabular/uci_bank_marketing/bank-additional-full.csv

Is your S3 input a manifest file? ⓘ
Off ☐ On ☒

Target ⓘ
y ▼

Auto deploy ⓘ
Off ☒ On ☐

Auto deploy endpoint ⓘ
SageMakerEndpoint-DDMMYYYY

En la tabla Output data location (S3 bucket) (Ubicación de datos de salida [*bucket* de S3]), elija su propio *bucket* de S3. En el campo Dataset directory name (Nombre del directorio del conjunto de datos), escriba *sagemaker/tutorial-autopilot/output*. Aquí es donde se guardarán los datos de salida una vez que se complete el experimento.

Output data location (S3 bucket)

☒ Find S3 bucket ☐ Enter S3 bucket location

S3 bucket name

sagemaker-us-east-1-255314940570 us-east-1

Dataset directory name

sagemaker/tutorial-autopilot/output

Deje la opción Auto deploy (Implementación automática) activada y el campo Auto deploy endpoint (Punto de conexión de implementación automática) en blanco. Esto implementará automáticamente nuestro modelo como punto de conexión de la API y le asignará un nombre.

A continuación, existe una serie de configuraciones avanzadas opcionales que controlan los detalles del tipo de problema, el experimento completo, los detalles del tiempo de ejecución, el acceso a IAM, la seguridad y el cifrado, entre otros. Haga clic en el botón de tiempo de ejecución para visualizar la configuración opcional.

Para este tutorial, reduzca el número de Max candidates (Máximo de candidatos) de 250 a 5. Esto ejecutará menos modelos más rápidamente. Un experimento completo es el mejor enfoque para optimizar realmente su modelo, pero puede tardar horas en finalizar. Para este tutorial, mantendremos la configuración opcional como está.

- ✓ Advanced settings - *Optional*
 - > Machine learning problem type
 - > Choose how to run your experiment
 - > Runtime
 - > Access
 - > Encryption
 - > Security
 - > Project
 - > Tags

Runtime

Max trial runtime

Hours: 0 Minutes: 0 Seconds: 0

Max candidates: 5

Max job runtime

Hours: 0 Minutes: 0 Seconds: 0

Haga clic en el botón Create Experiment (Crear experimento) para comenzar la primera etapa del experimento del Piloto automático de SageMaker. El Piloto automático de SageMaker comenzará a ejecutarse a través de las fases de un experimento. En la ventana del experimento, puede realizar un seguimiento del progreso a través de las fases de preprocesamiento, definición de candidatos, ingeniería de características, ajuste del modelo, explicabilidad e información.

Haga clic en el botón Create Experiment (Crear experimento) para comenzar la primera etapa del experimento del Piloto automático de SageMaker. El Piloto automático de SageMaker comenzará a ejecutarse a través de las fases de un experimento. En la ventana del experimento, puede realizar un seguimiento del progreso a través de las fases de preprocesamiento, definición de candidatos, ingeniería de características, ajuste del modelo, explicabilidad e información. Si aparece una notificación emergente con la pregunta: “Are you sure you want to deploy the best model?” (¿Está seguro de que desea implementar el mejor modelo?), haga clic en Yes (Sí).

Pre-processing

Candidate Definitions Generated

Feature Engineering

Model Tuning

Explainability Report Generated

Insights Report Generated

Trials Job profile

Your trials will display here as they are generated.

Una vez que se completa el trabajo del Piloto automático de SageMaker, puede acceder a un informe que muestra los modelos candidatos, el estado del modelo candidato, el valor objetivo, el valor F1 y la exactitud. El Piloto automático de SageMaker implementará automáticamente el punto de conexión.

AUTOPILOT JOB

half a minute ago

Deploy model

autopilot-experiment

Problem type: BinaryClassification

Open candidate generation notebook ⓘ

Open data exploration notebook ⓘ

Best model endpoint: autopilot-experiment-6d00f17b55464fc49c45d74362f284ce

Trials

Job profile

0 rows selected

Best Model	Model name	Status	Start time	Accuracy
● Best Model	autopilot-experimentG9...	● Completed	17 hours ago	0.845
	autopilot-experimentG9...	● Completed	17 hours ago	0.844
	autopilot-experimentG9...	● Completed	17 hours ago	0.844
	autopilot-experimentG9...	● Completed	17 hours ago	0.883
	autopilot-experimentG9...	● Completed	17 hours ago	0.498

Paso 3: interpretar el rendimiento del modelo

Ahora que el experimento está completo y tiene un modelo, el siguiente paso es interpretar su rendimiento. A continuación, aprenderá a utilizar el Piloto automático de SageMaker para analizar el rendimiento del modelo.

Ahora que el experimento del Piloto automático de SageMaker está completo, puede abrir el modelo de clasificación superior para obtener más detalles sobre el rendimiento y los metadatos del modelo. De la lista de modelos, resalte el primero y haga clic derecho para que aparezcan las opciones del modelo. Haga clic en Open in model details (Abrir en detalles del modelo) para revisar las estadísticas de rendimiento del modelo.

AUTOPILOT JOB

6 minutes ago

Deploy model

autopilot-experiment

Problem type: BinaryClassification

Open candidate generation notebook ⓘ

Open data exploration notebook ⓘ

Best model endpoint: [autopilot-experiment-6d00f17b55464fc49c45d74362f284ce](#)

Trials

Job profile

1 row selected

Best Model	Model name	Status	Start time	Accuracy
● Best Model	autopilot-experimentG9...	Completed	12 hours ago	0.845
	autopilot-experimentG9...			0.844
	autopilot-experimentG9...			0.844
	autopilot-experimentG9...			0.883
	autopilot-experimentG9...			0.498

Deploy model

Open in model details

Copy cell contents

Shift+Right Click for Browser Menu

Autopilot

half a minute ago

Deploy model

Stop training job

AUTOPILOT-EXPERIMENT-AWS-AUTO-ML-JOB

COMPLETED

Job Name: autopilot-experiment-aws-auto-ml-job

Model: autopilot-experimentG9A1yGyr3LaU-005-ce676305-aws-trial

Status: Completed

Algorithm: XGBoost

Problem Type: BinaryClassification

En la nueva ventana, haga clic en Explainability (Explicabilidad). La primera vista que aparece se denomina Feature Importance (Importancia de la característica) y representa el valor SHAP agregado para cada característica en cada instancia del conjunto de datos. El valor de importancia de la característica es una parte importante de la explicabilidad del modelo porque muestra qué características tienden a influir más en las predicciones del conjunto de datos. En este caso de uso, la duración o permanencia del cliente y la tasa de variación del empleo son los dos campos principales para impulsar el resultado del modelo.



Ahora, haga clic en la pestaña Performance (Rendimiento). Encontrará información detallada sobre el rendimiento del modelo, incluida la exactitud, la precisión y la recuperación. También puede interpretar el rendimiento del modelo y decidir si necesita un ajuste adicional.

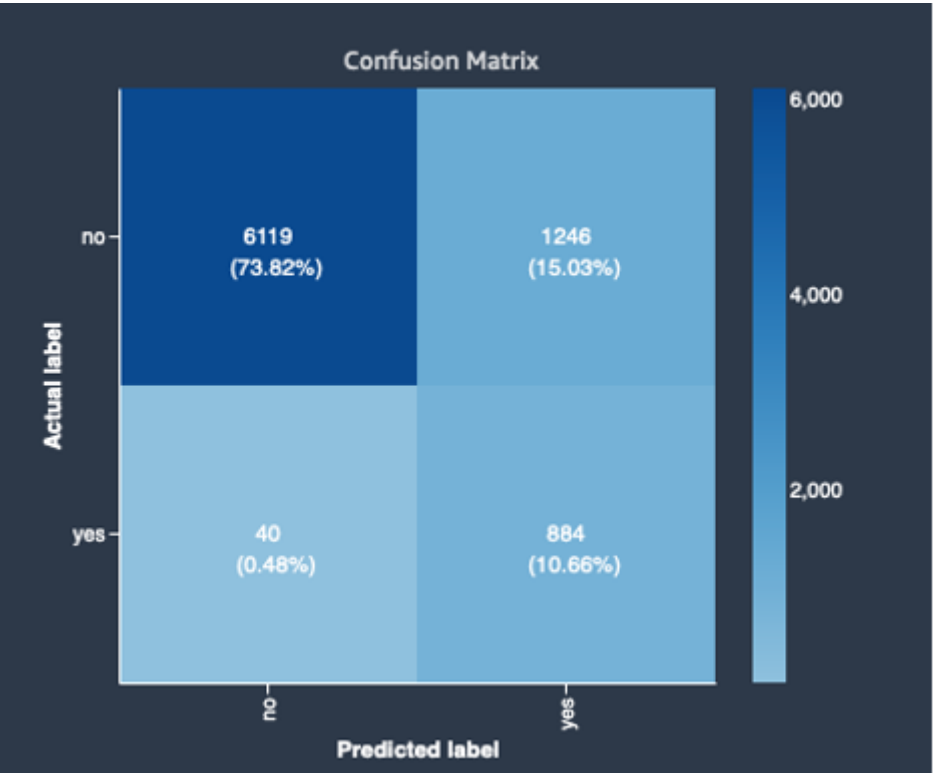
Explainability **Performance** Artifacts Network

[Download Performance Reports ▼](#)

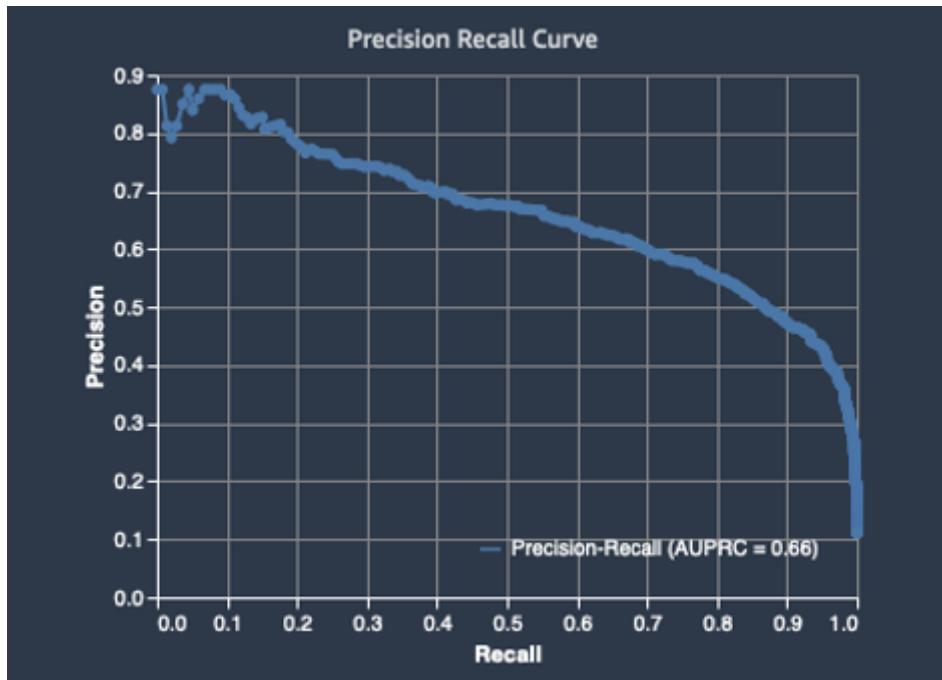
Metrics table

Metric Name	Value	Standard Deviation
recall	0.9567099567099567	0.0019029142045698803
precision	0.41502347417840374	0.004876965658586287
accuracy	0.8448546266135842	0.0021259957152844652
recall_best_constant_classifier	0	0
precision_best_constant_classifier	0	0
accuracy_best_constant_classifier	0.8885269634455302	0.0008484950229757214
true_positive_rate	0.9567099567099567	0.0019029142045698803
true_negative_rate	0.8308214528173795	0.0021896265777359516
false_positive_rate	0.16917854718262051	0.0021896265777359564
false_negative_rate	0.04329004329004327	0.0019029142045698588
auc	0.9497066239937927	0.001237869539105916
au_prc	0.6618712527480614	0.00817649169246613
f0_5	0.46802202456586195	0.005006384186964723
f1	0.5789129011132941	0.004928772633300932
f2	0.7586680398214898	0.003853994214450113
f0_5_best_constant_classifier	0	0
f1_best_constant_classifier	0	0
f2_best_constant_classifier	0	0

A continuación, se brindan visualizaciones para ilustrar aún más el rendimiento del modelo. Primero, observe la matriz de confusión. La matriz de confusión se suele utilizar para comprender cómo las etiquetas del modelo se dividen entre las clases predichas y las verdaderas. En este caso, los elementos de la diagonal muestran el número de etiquetas predichas correctamente y los elementos fuera de la diagonal muestran los registros mal clasificados. Una matriz de confusión es útil para analizar clasificaciones erróneas debido a falsos positivos y falsos negativos.

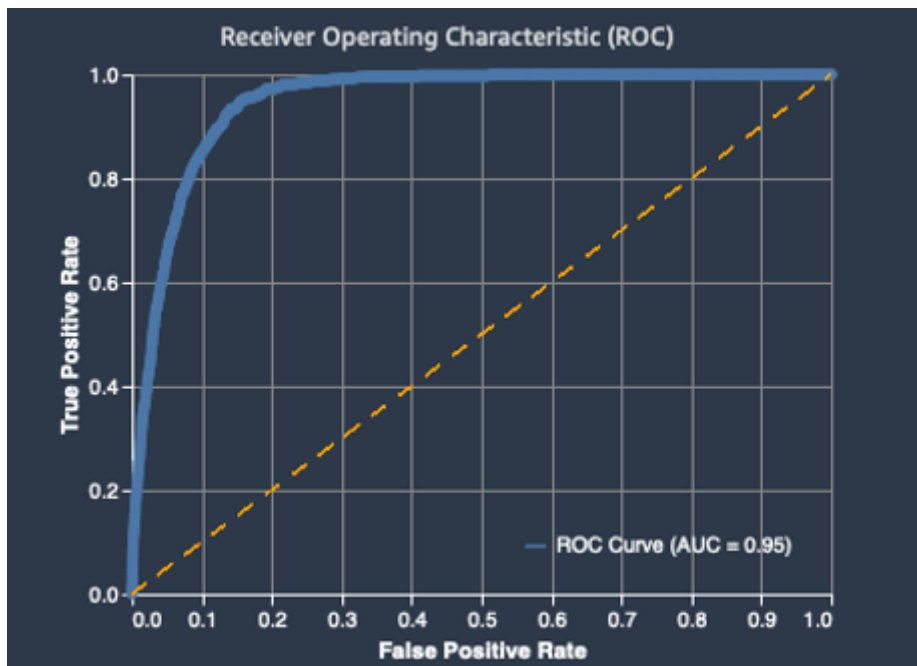


A continuación, observe la curva de precisión frente a la de recuperación. Esta curva interpreta la etiqueta como un umbral de probabilidad y muestra la compensación que se produce como varios umbrales de probabilidad para la precisión y recuperación del modelo. El Piloto automático de SageMaker optimiza automáticamente estos dos parámetros para proporcionar el mejor modelo.



A continuación, observe la curva denominada “Receiver Operating Characteristic” (ROC, Característica operativa del receptor). Esta curva muestra la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos en una variedad de posibles umbrales de probabilidad. Una línea diagonal representa un modelo hipotético basado en predicciones aleatorias. Cuanto más se desplace esta curva hacia la parte superior izquierda del gráfico, mejor funcionará el modelo.

La línea punteada representa un modelo con valor predictivo 0, que a menudo se denomina modelo nulo. El modelo nulo asignaría aleatoriamente una etiqueta 0/1 y su área bajo la curva ROC sería 0,5, lo que representa que sería preciso el 50 % de las veces.



A continuación, haga clic en la pestaña Artifacts (Artefactos). Puede encontrar los activos de apoyo del experimento del Piloto automático de SageMaker, incluido el código de ingeniería de características, las ubicaciones de los datos de entrada y los artefactos de explicabilidad.

[Explainability](#)
[Performance](#)
[Artifacts](#)
[Network](#)

Input Data

https://us-east-1.console.aws.amazon.com/s3/object/sagemaker-sample-files/datasets/tabular/uci_bank_marketing/bank-additional-full.csv

Shuffled training and validation splits

https://console.aws.amazon.com/s3/buckets/sagemaker-us-east-1-255314940570/sagemaker/tutorial-autopilot/output/autopilot-experiment/preprocessed-data/tuning_data/

Transformed training and validation data

<https://console.aws.amazon.com/s3/buckets/sagemaker-us-east-1-255314940570/sagemaker/tutorial-autopilot/output/autopilot-experiment/transformed-data/dpp1/csv/>

Feature Engineering Code (Python)

https://us-east-1.console.aws.amazon.com/s3/object/sagemaker-us-east-1-255314940570/sagemaker/tutorial-autopilot/output/autopilot-experiment/sagemaker-automl-candidates/autopilot-experiment-pr-1-6a9ccd8537a849e7ac7ee4f60efb6ac7b3737/generated_module/candidate_data_processors/dpp1.py

Feature Engineering Model

<https://us-east-1.console.aws.amazon.com/s3/object/sagemaker-us-east-1-255314940570/sagemaker/tutorial-autopilot/output/autopilot-experiment/data-processor-models/autopilot-experiment-dpp1-1-7c1e2f1934da4eef919a0b7f0cabfc868a0/output/model.tar.gz>

Algorithm Model

<https://us-east-1.console.aws.amazon.com/s3/object/sagemaker-us-east-1-255314940570/sagemaker/tutorial-autopilot/output/autopilot-experiment/tuning/autopilot--dpp1-xgb/autopilot-experimentG9A1yGyr3LaU-005-ce676305/output/model.tar.gz>

Finalmente, haga clic en la pestaña Network (Red). Encontrará información sobre el aislamiento de la red y el cifrado del tráfico de contenedores.

Autopilot

half a minute ago

AUTOPILOT-EXPERIMENT-AWS-AUTO-ML-JOB

Deploy model
Stop training job

COMPLETED

Job Name: autopilot-experiment-aws-auto-ml-job
Model: autopilot-experiment4jFOVKbEmr6N-004-f90408aa-aws-trial
Status: Completed
Algorithm: XGBoost
Problem Type: BinaryClassification

[Explainability](#)
[Performance](#)
[Artifacts](#)
[Network](#)

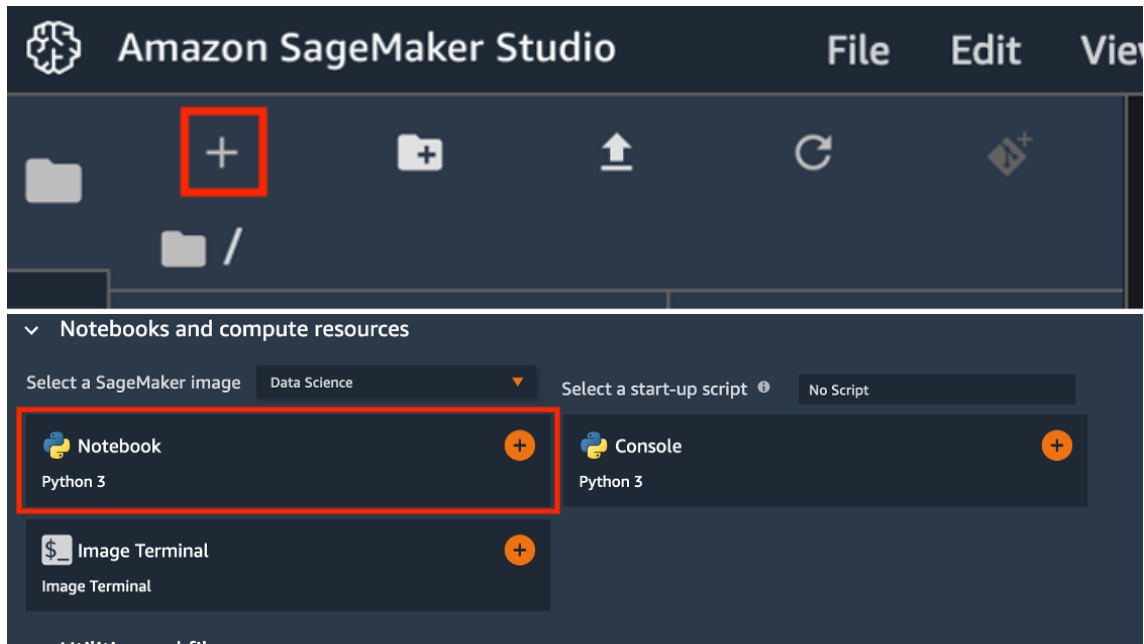
Enable network isolation
True

Enable inter-container traffic encryption
True

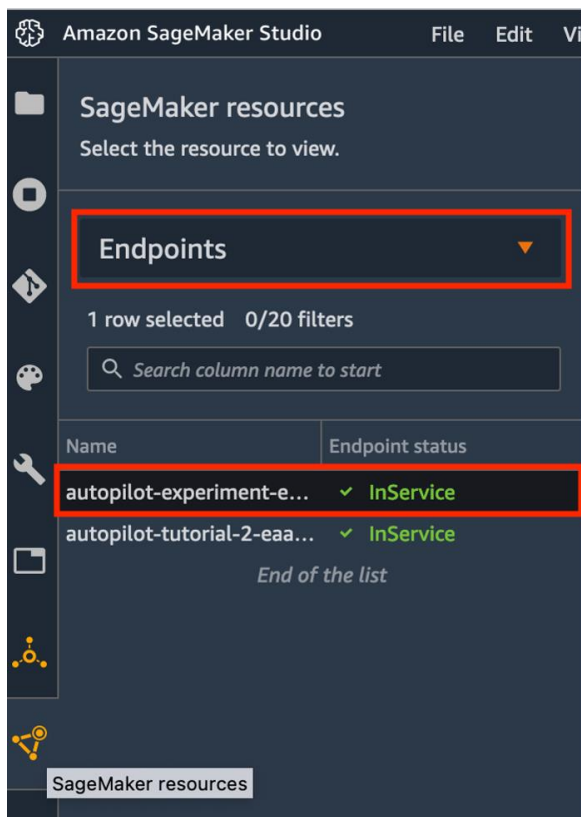
Paso 4: probar el punto de conexión del modelo de SageMaker

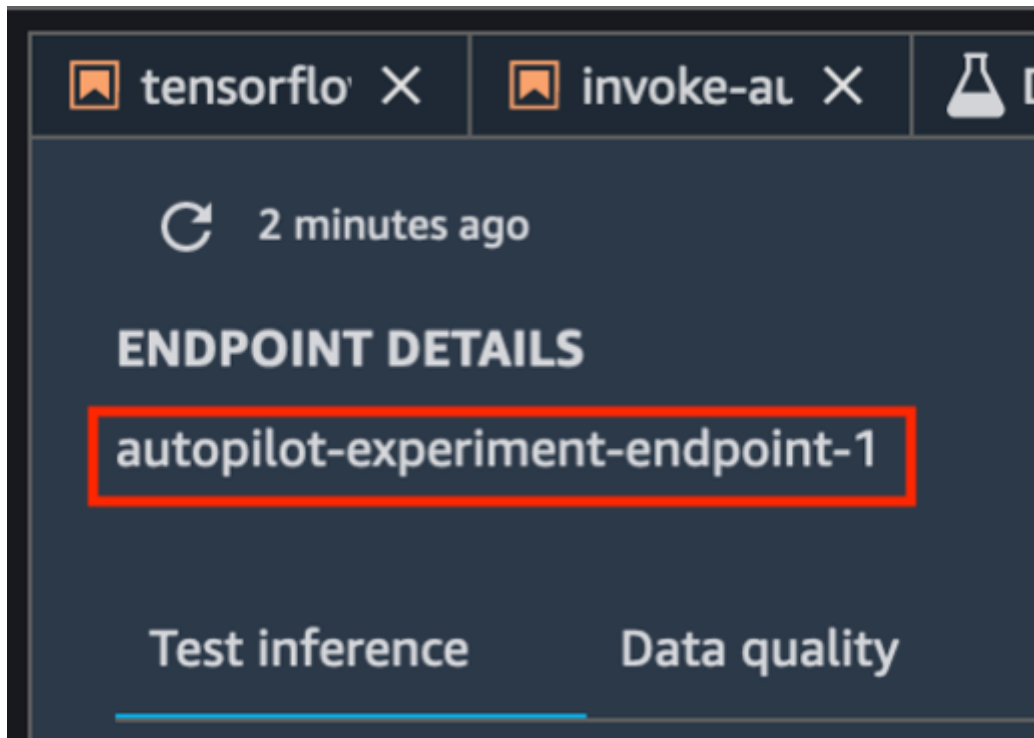
Ahora que ha revisado los detalles del modelo, pruebe el punto de conexión.

Haga clic en el ícono + para abrir un nuevo cuaderno de Python. Seleccione Python3 como *kernel*. En la primera celda, copie y pegue el siguiente código. Este código creará una carga, enviará una solicitud al punto de conexión del modelo y analizará la respuesta.



Para saber a dónde enviar la solicitud, busque el nombre del punto de conexión del modelo. En el panel izquierdo, haga clic en el ícono SageMaker Resources (Recursos de SageMaker). En el panel de recursos de SageMaker, seleccione Endpoints (Puntos de conexión). Haga clic en el punto de conexión asociado con el nombre del experimento que creó al comienzo de este tutorial. Esto abrirá la ventana Endpoint Details (Detalles del punto de conexión). Registre el nombre del punto de conexión y navegue de vuelta al cuaderno de Python 3.





Copie y pegue el siguiente fragmento de código en una celda del cuaderno y presione Shift+Enter (Mayús+Intro) para ejecutar la celda actual. Este código establece la variable de entorno `ENDPOINT_NAME` y ejecuta la inferencia. A medida que se completa el código, verá un resultado que consta de la etiqueta del modelo y el valor de la probabilidad asociada.

```
import os
```

```
import io
```

```
import boto3
```

```
import json
```

```
import csv
```

```
#: Define the endpoint's name.
```

```
ENDPOINT_NAME = 'autopilot-experiment-6d00f17b55464fc49c45d74362f284ce'
```

```
runtime = boto3.client('runtime.sagemaker')
```

```
#: Define a test payload to send to your endpoint.
```

```
payload = {
```

```
    "data":{
```

```
        "features": {
```

```

        "values":
[45,"blue-
collar","married","basic.9y",'unknown',"yes","no","telephone","may","mon",461,1,999,0,"non
existent",1.1,93.994,-36.4,4.857,5191.0]

    }

}

}

```

#: Submit an API request and capture the response object.

```

response = runtime.invoke_endpoint(

    EndpointName=ENDPOINT_NAME,

    ContentType='text/csv',

    Body=str(payload)

)

```

#: Print the model endpoint's output.

```
print(response['Body'].read().decode())
```

Copiar

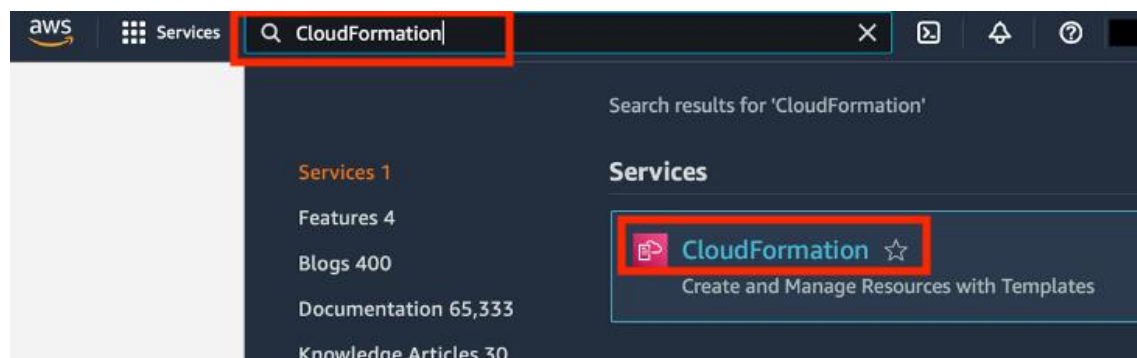
¡Felicitaciones! Aprendió a usar el Piloto automático de SageMaker para entrenar e implementar automáticamente un modelo de machine learning.

Paso 5: eliminar los recursos de AWS

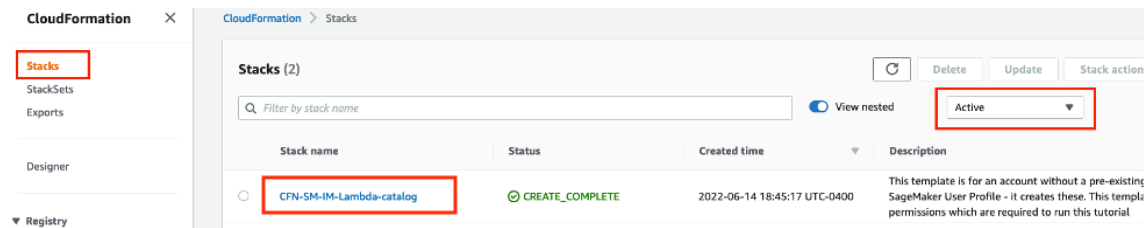
Una práctica recomendada es eliminar los recursos que ya no se utilizan para no incurrir en cargos no deseados.

Si ejecutó la plantilla de CloudFormation en el paso 1 para crear un dominio de SageMaker Studio nuevo, continúe con los siguientes pasos a fin de eliminar el dominio, el usuario y los recursos creados por la plantilla de CloudFormation.

Para abrir la consola de CloudFormation, ingrese CloudFormation en la barra de búsqueda de la consola de AWS y elija CloudFormation en los resultados de búsqueda.



En el panel de CloudFormation, elija Stacks (Pilas). En la lista desplegable de estado, seleccione Active (Activo). En Stack name (Nombre de la pila), elija CFN-SM-IM-Lambda-Catalog para abrir la página de detalles de la pila.



En la página de detalles de la pila CFN-SM-IM-Lambda-Catalog, elija Delete (Eliminar) para eliminar la pila junto con los recursos que creó en el paso 1.

