

Punto 2

a. ¿Cuáles son las diferentes opciones que tiene para usar SageMaker?

1. IDE: SageMaker Studio
2. Console: SageMaker Notebook Instances
3. Command line & SDK: AWS CLI, boto3, & SageMaker Python SDK
4. 3rd party integrations: Kubeflow & Kubernetes operators

b. Videos

SageMaker Notebook Instances

Amazon SageMaker nos ofrece tanto servicios para generar cuadernos jupyter como para entrenar modelos. Dentro de la configuración principal Notebooks, se encuentran los lifecycles configurations. Algunos ejemplos son la instalación de paquetes, configurar la seguridad de la red, o correr scripts personalizados para nuestros notebooks. Esta opción toma una relevancia importante debido a que permite crear y guardar diferentes configuraciones personalizadas. Cuando se crea una de estas configuraciones, aparecen dos opciones: ejecutar los scripts siempre que se inicia un Notebook, o solamente cuando se crea. AWS SageMaker tiene también la opción de integrar directamente repositorios bajo el control de versiones git. Se puede usar AWS Code, u otro repositorio externo. Si usamos un repositorio externo, AWS brinda la opción de guardar nuestras credenciales a través de AWS Secret Manager, brindando el usuario y un personal token account.

El siguiente paso es la creación de los Notebook. Existe la opción de seleccionar diferentes tipos de cuadernos que varíen en su capacidad computacional. Además, se puede cambiar de tipo de Notebook en el momento que nuestra aplicación lo requiera, reduciendo de esta manera los costos generados a lo largo del desarrollo del proyecto. Más adelante, AWS brinda la opción de elegir el tipo de lifecycle configuration que se desea, y el tamaño del cuaderno (el cuál puede oscilar entre mínimo 5GB y máximo 16TB). Este tamaño, al igual que el tipo, es completamente variable. Otro punto importante es la creación del IAM role, en donde se debe escoger un perfil para permitir la conexión de Amazon SageMaker con otros servicios de AWS, tales como S3 BUCKET. AWS también brinda opciones a la hora de encriptar los Notebooks.

En cuanto a conexiones de red se refiere, SageMaker permite activar o no la VPC (Virtual Private Cloud), el cual es un servicio en el cual se "aísla" la nube de trabajo del resto del servidor y si se desea un aislamiento para conexión de red. Por último, AWS permite configurar etiquetas que permiten entender de manera más clara los recursos con los que se cuenta.

El manejo del tiempo en el que se encuentra activado un Notebook. Siempre que ya no se necesite ese código en ejecución, es recomendable desactivarlo, y así ahorrar costos de ejecución en la nube. Para parar un Notebook, es necesario especificar una acción de parada, seleccionando desde el menú principal la opción de Stop. Si llega el caso en el que definitivamente no se necesite el Notebook, este puede ser borrado, pero esta es una acción que no puede ser revertida.

Onboard to Amazon Sagemaker:

En primera instancia se enfoca en definir que es “onboarding”, que se entiende como el proceso de crear un dominio asociado a una cuenta personal en una región determinada. La creación de un dominio tiene atributos, tales como el acceso a la creación de cuaderno y la capacidad de compartirlos con otros usuarios de una forma muy sencilla. Por otro lado, al crear un dominio, le permite al usuario tener “amazon elastic file systems (efs)”, en donde cada usuario podrá tener su directorio, garantizando un ambiente de desarrollo privado.

Para la creación del dominio, existen dos opciones para crear un dominio, la primera siendo una forma automática y rápida, mientras que la segunda siendo una alternativa estándar, la cual permite tener todavía más control en lo concerniente a permisos, configuración para compartir cuadernos y formas de encriptación de la información. Inmediatamente seleccionada la opción estándar se despliega un menú de configuraciones de por ejemplo el método de autorización que depende del tipo de cuenta que se tenga, los permisos a los buckets de AWS a los que el usuario podrá acceder, la opción de compartir o no los cuadernos, la forma de encriptación de los cuadernos y la forma de almacenamiento si será en la nube o no.

En este punto, la página deberá estar situada en el panel frontal, en donde no se podrá avanzar mucho más si no se tiene un usuario ya creado con un rol estipulado dentro del espacio de trabajo. Lo único posible en este punto por parte de la empresa es la posibilidad de eliminar usuarios y quitarles los permisos a los cuadernos que tenga acceso.

Una vez se haya creado el usuario. El primer paso es la elección del kernel/núcleo asociado, que por fines prácticos se elige comúnmente a Python 3. En cuanto a los studios, en este se pueden crear carpetas y visualizar los datos o archivos a los que se tiene acceso, se pueden subir archivos, se puede acceder a la sesión de administrador que permite ver los kernels, terminales e imágenes que se están

corriendo actualmente, se tiene la paleta de funciones, las herramientas de cuaderno y se tiene la opción de abrir una terminal para ejecutar muchas de las funciones del studio partiendo de comandos como en Linux (file, new y terminal).

Punto 3

La Agencia de Protección Ambiental define las aguas residuales como el material de desecho que incluye desechos de aguas residuales y desechos líquidos industriales que se recolectan en áreas urbanas y rurales. Luego, las aguas residuales se tratan en las plantas de tratamiento de aguas residuales urbanas, así como las aguas residuales que provienen de las casas individuales que se pueden encontrar en el campo.

Las aguas residuales sin tratar pueden contener contaminantes biológicos, químicos y físicos. Los contaminantes químicos o físicos incluirían metales pesados, partículas orgánicas, material orgánico soluble, macro sólidos, emulsiones y toxinas, entre otros.

Los contaminantes biológicos, por otro lado, incluirían bacterias, virus, protozoos y helmintos. Las aguas residuales también pueden contener organismos no patógenos como artrópodos, insectos y peces.

Tipos de Aguas Residuales:

- Doméstica:

Con origen en los hogares, las aguas residuales domésticas pueden incluir agua procedente de hospitales, industrias y otras instalaciones comerciales. El agua recolectada durante una tormenta también podría calificar como agua residual doméstica. Sus fuentes comunes incluyen descargas líquidas de nuestros usos sanitarios regulares como bañarse, cocinar,

lavar carne, verduras y ropa. A pesar de ser muy impuras, las aguas residuales domésticas pueden ser debidamente tratadas.

- **Industriales:**

Estas aguas residuales generalmente se originan en unidades comerciales o de fabricación. Esta agua contiene una gran cantidad de impurezas como aceite, limo, productos químicos tóxicos, tinta, medicamentos, pesticidas, arena, limo, productos químicos, y es posible que no se trate tan convenientemente como las aguas residuales domésticas.

Subcategorías de las aguas residuales:

- **Aguas Negras:**

aguas residuales domésticas y generalmente se origina en nuestras unidades de inodoro, unidades de lavado de platos y fregaderos de cocina. Este tipo de aguas residuales ha sido contaminada con desechos humanos.

Debido a la fuerte presencia de productos químicos disueltos, esta agua se considera altamente patógena. También es el hogar de millones de bacterias que deben descomponerse por completo antes de ser liberadas nuevamente al medio ambiente.

- **Aguas Grises:**

Este tipo de aguas residuales no ha sido contaminada con materia fecal. Este es el agua de alcantarillado que no tiene ningún rastro de orina o desechos humanos. Aunque tóxicas, las aguas grises pueden ser debidamente tratadas, haciéndolas aptas para uso humano.

Debe tenerse en cuenta que contiene pequeñas cantidades de posibles contaminantes que podrían causar enfermedades si se ingieren. Por lo general, las aguas grises son adecuadas para su reutilización después de ser tratadas en una instalación de tratamiento de aguas residuales adecuada.

- Aguas Amarillas:

Esto se refiere a la orina que se recolecta a través de canales dedicados. El agua amarilla no está contaminada por las impurezas del agua negra o gris. Al igual que con las aguas grises, puede usarse adecuadamente.