

LABORATORIO #1

El alumno deberá realizar una investigación de no más de 3 páginas sobre las tecnologías que estaremos usando durante el curso: Las tecnologías son las siguientes: Docker, PyCaret, MLFlow, DagsHub, DVC, CokieCutter y Flask.

Adicionalmente para cada tecnología deberá responder las siguientes preguntas:

1. ¿Qué hace la tecnología en particular?
2. ¿Según su experiencia, para qué podría servirnos esta tecnología en la implementación de un proyecto de ML?
3. ¿Hay alguna otra tecnología sustituta para este caso? Mencione y describa brevemente.

Tecnología Docker

Es una plataforma de software que le permite crear, probar e implementar aplicaciones rápidamente, por lo que empaqueta software en unidades estandarizadas llamadas contenedores que incluyen todo lo necesario para que el software se ejecute, incluidas bibliotecas, herramientas de sistema, código y tiempo de ejecución.

1. Docker se instala en cada servidor y proporciona comandos sencillos que puede utilizar para crear, iniciar o detener contenedores.
2. Esta tecnología puede ser muy útil en la implementación de proyectos de Machine Learning (ML), por Portabilidad, Aislamiento, Escalabilidad y Reproducibilidad.
3. Si, existen unas plataformas de contenedores, como Kubernetes o Docker Swarm, que permiten gestionar y escalar contenedores en un clúster de máquinas.

Tecnología PyCaret

La biblioteca de PyCaret proporciona características que permite un aprendizaje automático en Python comprobar de forma puntual un conjunto de algoritmos estándar de aprendizaje automático en un conjunto de datos de clasificación o regresión con una sola llamada a la función.

1. Es una interfaz realizar tareas comunes para el aprendizaje automático, como procesamiento de datos, selección y entrenamiento de modelos y evaluación de modelos.



2. PyCaret puede ser muy útil en la implementación de proyectos de aprendizaje automático (ML) debido a la simplicidad y la eficiencia, ahorro de tiempo, facilidad de uso, despliegue sencillo.
3. Si, existen varias tecnologías como Scikit-Learn, TensorFlow y Keras.

Tecnología MLFlow

MLflow es una plataforma de código abierto para administrar el ciclo de vida completo del aprendizaje automático que contiene los siguientes componentes, seguimiento, modelos, proyectos, registro de modelos y servicios de modelos, los cuales permite a los profesionales y equipos de Machine Learning llevar un control de sus experimentos, reproducir resultados, empaquetar y desplegar modelos fácilmente, y compartir sus descubrimientos con otros miembros del equipo.

1. MLflow surgió como una solución unificada y modular para simplificar la gestión de los proyectos y mejorar la colaboración en equipos de trabajo.
2. Esto puede ser muy útil en la implementación de proyectos de ML, ya que proporciona una plataforma centralizada para administrar y rastrear todo el ciclo de vida de los modelos
3. Si, existen otras alternativas que también pueden ser consideradas según los requisitos y preferencias específicas, TensorBoard, Data Version Control y Kubeflow

Tecnología DagsHub

Es una plataforma de colaboración y versionado de datos, código y modelos de ML basada en Git. DagsHub permite a los equipos de desarrolladores colaborar de manera eficiente en proyectos de ML, rastreando las modificaciones y compartiendo experimentos y resultados.

1. Es una colaboración y gestión de datos diseñada específicamente para proyectos de ciencia de datos y aprendizaje automático (ML)
2. Puede ser muy útil en la implementación de proyectos de ML debido a sus características y funcionalidades específicas para la gestión de datos y colaboración en proyectos de ciencia de datos.
3. Si, existen otras tecnologías que pueden ser consideradas como alternativas a DagsHub en la gestión de proyectos de ML, como GitHub y Neptune

Tecnología DVC

DVC es un sistema de versiones gratuito de código abierto e independiente de la plataforma para datos, modelos de aprendizaje automático y experimentos por lo que está diseñado para hacer que los modelos de ML sean compatibles.

1. DVC Permite registrar, monitorear y gestionar cambios realizados sobre un sistema de archivos local o remoto, llevando un historial de cuándo, quién y para qué se han hecho los cambios.
2. Es muy útil en la implementación de proyectos de ML debido a sus características específicas para el control de versiones de datos.
3. i, existen otras tecnologías que pueden ser consideradas como alternativas, tales como, Git-FLS y Pachyderm.

Tecnología Cookiecutter

Es una utilidad de línea de comandos que crea proyectos a partir de una plantilla (plantillas de proyectos), por ejemplo, crear un proyecto de paquete de Python a partir de una plantilla de proyecto de paquete de Python.

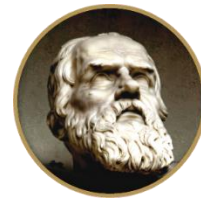
1. Es una interfaz gráfica de usuario para descubrir plantillas, opciones de plantilla de entrada y crear proyectos y archivos
2. En la implementación de un proyecto de Machine Learning (ML), Cookiecutter puede ser útil Estructura de proyecto consistente, Ahorro de tiempo y personalización.
3. Si, existen otras tecnologías que pueden considerarse como alternativa Yeoman y Hygen

Tecnología Flask

Los frameworks full-stack como Django proporcionan sus propias bibliotecas a los desarrolladores, Flask Python adopta un enfoque diferente, solo incluye el motor de plantillas Jinja y una biblioteca llamada "tool". Sin embargo, ofrece la posibilidad de integrar funciones de terceros.

1. Es un framework ligero de aplicaciones web en Python.
2. Flask puede ser útil en la implementación de proyectos de Machine Learning (ML), desarrollo de una API, despliegue de modelos de producción e interfaz de usuarios
3. Si, existen otras tecnologías que se pueden considerar como alternativas Django y FastAPI

Instituto de Investigación de Operaciones
Postgrado en Inteligencia de Negocios
Introducción a la Investigación
Irma Enriquez **22000347**



Galileo
UNIVERSIDAD
La Revolución en la Educación