**Especialización de Data Analityc**

**Módulo de DataOps**

**Alumno. Paul Efren Santos Andrade / DNI: 47062815**

**Examen parcial: Implementación de un Pipeline en Jenkins**

**1. Problemática a Resolver**

En entornos de desarrollo y análisis de datos, la ejecución manual de scripts puede generar problemas como inconsistencias en los datos, errores humanos y dificultades en el mantenimiento. Se requiere un proceso automatizado que garantice la correcta ejecución de los scripts en cada etapa del flujo de datos.

**2. Beneficios de Jenkins ante el Problema**

Jenkins es una herramienta de integración continua que permite:

* Automatizar la ejecución de scripts en diferentes etapas del pipeline.
* Generar registros (logs) detallados de cada proceso, fundamentales para el monitoreo, depuración y mantenimiento de sistemas informáticos.
* Monitorear el estado del pipeline en tiempo real.
* Integrarse con sistemas de notificación por correo en caso de fallos o éxito, los cuales pueden ser complementados con información básica de los procesos.

**3. Despliegue del Pipeline en Jenkins**

**a) Creación del Pipeline**

1. Acceder a Jenkins y crear un nuevo Pipeline.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

1. En la sección *Pipeline*, seleccionar *Pipeline Script* y pegar el código.

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

1. Guardar el pipeline.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

1. Ejecutar el pipeline.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

La ejecución completa del pipeline

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

**4. Código del Pipeline**

El siguiente código de Jenkins define un pipeline para la ejecución automática de scripts de procesamiento de datos en Python, con las siguientes etapas:

**Etapas del Pipeline**

1. **READ (Lectura de Datos)**

* Se ejecuta el script data\_read.py, el cual lee un archivo txt y lo almacena en un temporal csv.

1. **TRANSFORM (Transformación de Datos)**

* Se ejecuta el script data\_transform.py, que limpia y procesa la información.

1. **EXPORT (Exportación de Datos)**

* Se ejecuta el script data\_export.py, el cual almacena los datos procesados en el directorio de salida, en el archivo clientes\_ordenados.xlsx

1. **SUMMARY (Resumen del Proceso)**

* Se ejecuta data\_summary.py, que genera un resumen que se almacena en el archivo summary.txt con estadísticas clave (cantidad de registros procesados, errores, etc.).

Imagen que contiene Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

**5. Monitoreo del Pipeline**

* Se puede visualizar en la interfaz gráfica de Jenkins.
* Revisar los logs generados en la carpeta D:\...\logs.

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

* Validar la existencia del archivo summary.txt como resultado final.

**6. Envío de Notificaciones Mediante Mail**

La configuración de Jenkins para el envío de correos de notificación de la ejecución del pipeline.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

**7. Conclusiones y Recomendaciones**

* **Automatización eficiente:** Jenkins permite ejecutar los scripts de manera automática y organizada.
* **Monitoreo y Logs:** Se pueden revisar logs en tiempo real y recibir notificaciones de errores.
* **Facilidad de Integración:** Es posible conectar Jenkins con sistemas de almacenamiento y bases de datos para una mejor gestión.
* **Recomendación:** Se sugiere validar que todos los scripts y rutas estén configurados correctamente antes de ejecutar el pipeline.

Este pipeline facilita la ejecución de procesos de transformación de datos en un entorno controlado y seguro, mejorando la trazabilidad y eficiencia del flujo de trabajo.

**8. Secuencia de código**

**Texto

El contenido generado por IA puede ser incorrecto.**