

DOCUMENTACION DEL PROYECTO

Preprocesamiento en Ciencia de datos

Autor

Mishell Pamela Ramos Vasquez

Fecha

09-11-2025

Carrera

Ciencia de Datos e inteligencia artificial

Periodo académico

2025

Semestre

Tercero A

INTRODUCCION

OBJECTIVO

Aplicar de manera práctica el uso de Git y GitHub para la gestión de versiones y la colaboración en proyectos de Ciencia de Datos. Además, implementar un preprocesamiento completo de datasets usando Pandas, incluyendo la gestión de valores nulos, normalización, codificación de variables categóricas y eliminación de duplicados.

Estructura del proyecto

```
preprocesamiento_cienciadatos/ data/ # conjunto de datos notebooks/ # Jupyter Notebooks scripts/ #
Código fuente en Python (.py) docs/ # Documentación del proyecto .gitignore # Archivos y carpetas que no
se suben al repositorio README.md # Descripción general del proyecto
```

Comandos GIT usados

git clone -> clona el repositorio
git config --global user.name -> Configura el nombre de usuario
git config --global user.email -> configura el correo del usuario
git checkout -b feature-preprocesamiento -> Crea y cambia a una nueva rama de desarrollo
git add -> Añade todos los archivos modificados
git commit -m -> guarda los cambios
git push origin feature-preprocesamiento -> envía la nueva rama y sus comits
git merge feature-preprocesamiento -> Fusiona los cambios de la rama secundaria
git branch -d feature-

preprocesamiento -> Elimina la rama local git push origin --delete feature-preprocesamiento -> elimina la rama remota

Automatización

En este proyecto se implemento el workflow para automatizar la verificación del código Python. Además, el flujo ejecuta automáticamente pruebas y verificar cada vez que se hace un push.

Captura de pantalla

Imagen 1

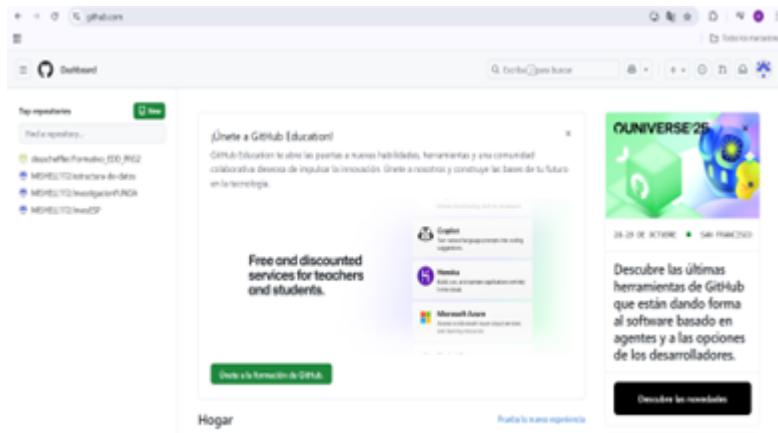


Imagen 2

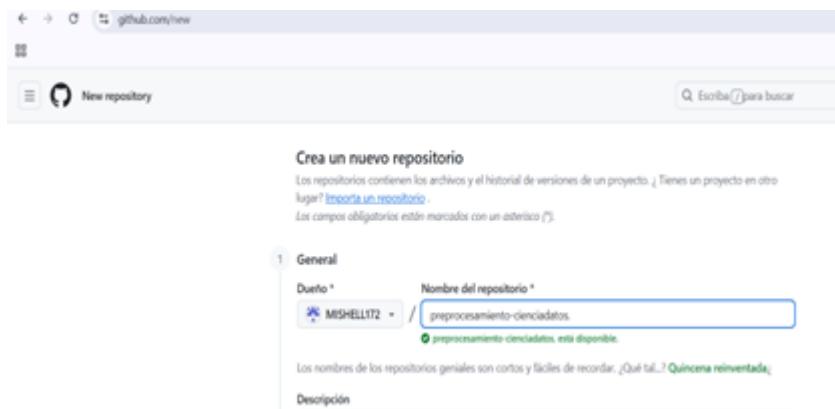


Imagen 3

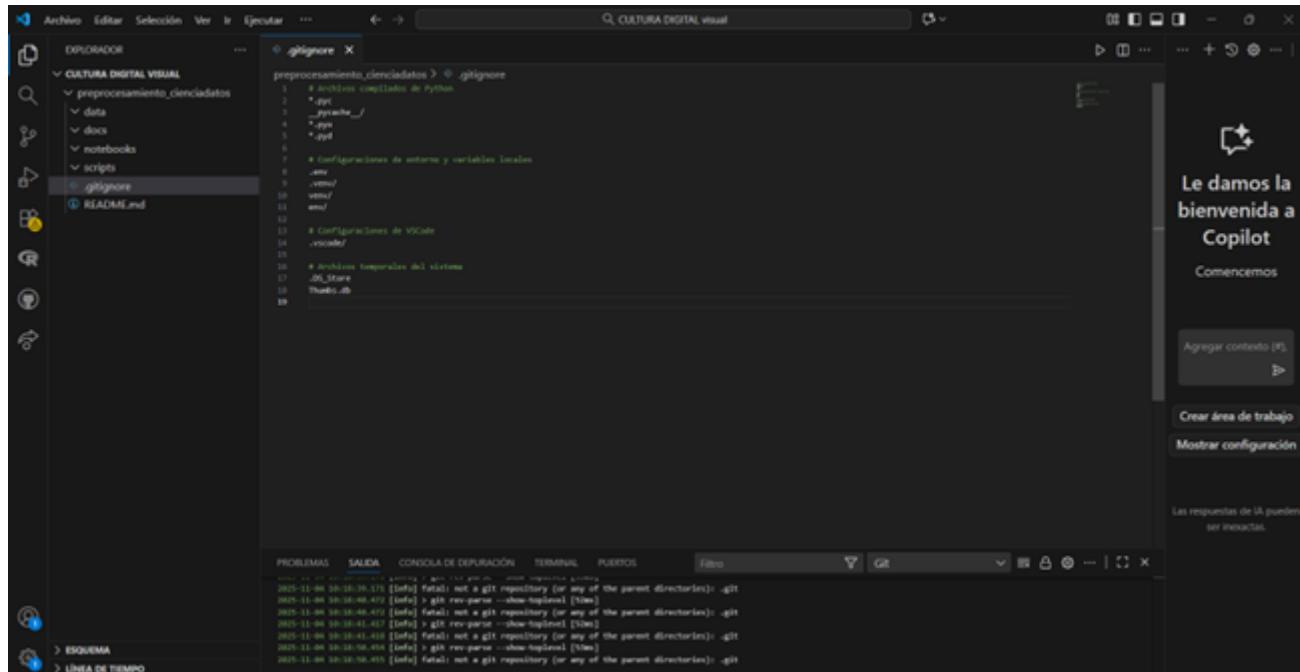


Imagen 4

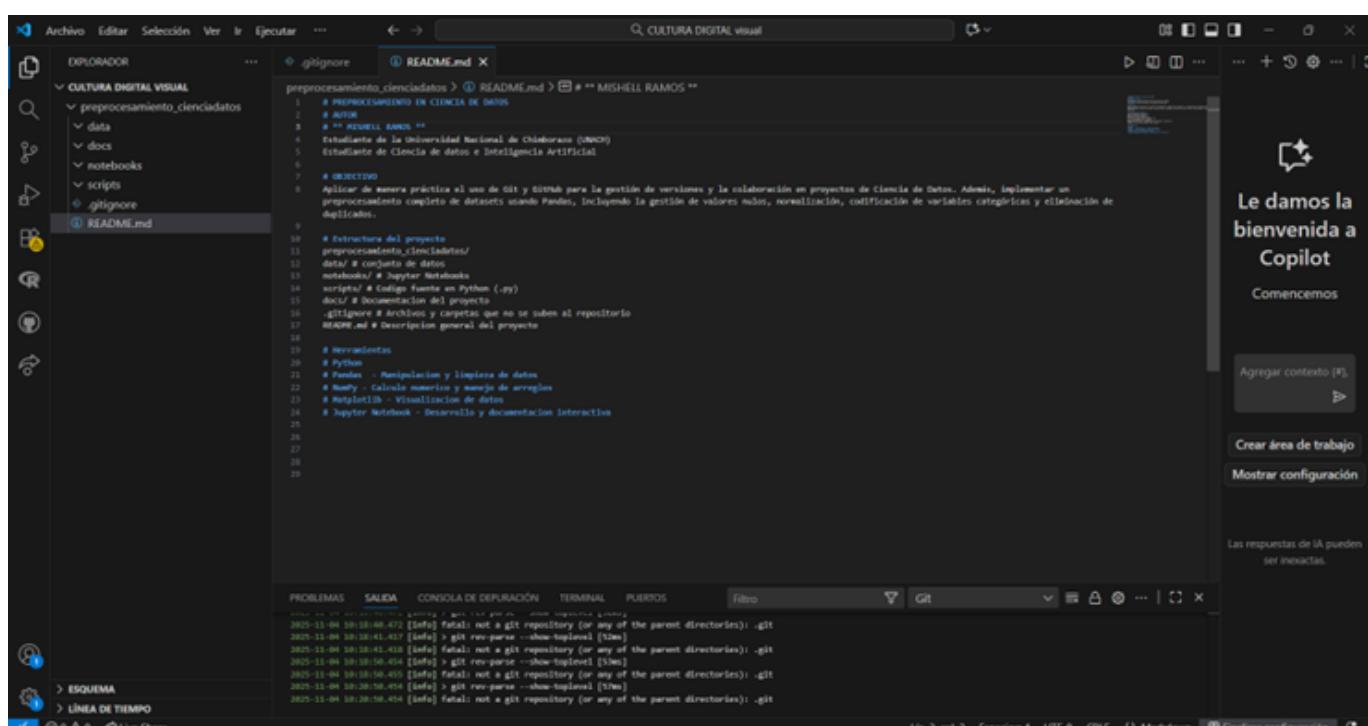


Imagen 5

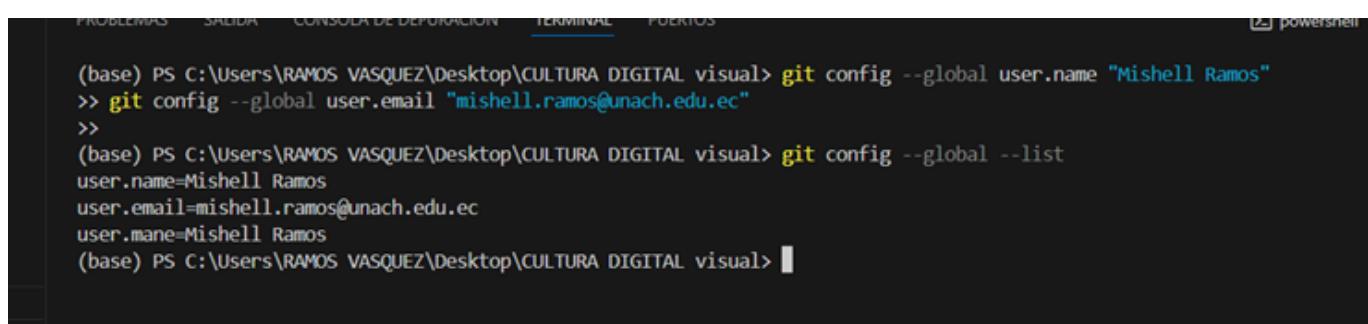


Imagen 6

```
(base) PS C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual> cd "C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual\preprocesamiento_ciencia_datos"
• ncia_datos
>>
(base) PS C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual\preprocesamiento_ciencia_datos> git checkout -b feature-preprocesamiento
● >>
Switched to a new branch 'feature-preprocesamiento'
(base) PS C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual\preprocesamiento_ciencia_datos> git branch
>>
* feature-preprocesamiento
  main
(base) PS C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual\preprocesamiento_ciencia_datos> echo "# Script de preprocesamiento de datos" >
scripts/preprocesamiento.py
>>
(base) PS C:\Users\RAMOS VASQUEZ\Desktop\CULTURA DIGITAL visual\preprocesamiento_ciencia_datos> git add scripts/preprocesamiento.py
>> git commit -m "Agrega funciones iniciales de limpieza y normalización"
>> git push origin feature-preprocesamiento
```

Imagen 7

Feature preprocesamiento - Agregar script de preprocesamiento #1

Abierto MISHELL172 Quiere fusionar 2 commits en principal de preprocesamiento de características

Conversación 0 Compromisos 2 cheques 0 Los archivos cambiaron 1 +0 -0

MISHELL172 comentó hace 2 minutos

Este Pull incorpora el archivo scripts/preprocesamiento.py el cual contiene la primera función para el procesamiento de datos.

Cambios principales:

- Guiones de creación de la carpeta
- Creación del archivo preprocesamiento.py
- Preparación del proyecto

MISHELL172 añadió 2 confirmaciones hace 24 minutos

Agrega funciones iniciales de limpieza y normalización c767cd6

Crea script preprocesamiento.py con funciones de limpieza y normaliza... a8a7ba5

Comprobando la capacidad de fusión automática... Merging can be performed automatically.

Solicitud de fusión También puedes combinar esto con la línea de comandos. Consulte las instrucciones de la línea de comandos.

Revisores No reviews ¿Sigue en progreso? Convertir a borrador

Cesionarios Nadie Asignarte a ti mismo

Etiquetas Ninguno todavía

Proyectos Ninguno todavía

Hito Hito Sin hito

Desarrollo La correcta fusión de esta solicitud de extracción podría solucionar estos problemas. None yet

Notificaciones Personalizar Cancelar suscripción

Imagen 8

This screenshot shows a GitHub pull request merge confirmation dialog. At the top, it says "MISHELL172 comentó hace 9 minutos" (MISHELL172 commented 9 minutes ago). The message reads: "Este Pull incorpora el archivo scripts/preprocesamiento.py el cual contiene la primera función para el procesamiento de datos. Cambios principales: • Guiones de creación de la carpeta • Creación del archivo preprocesamiento.py • Preparación del proyecto". Below this, there are two commits from MISHELL172: "agrega funciones iniciales de limpieza y normalización" (commit c767cd6) and "Crea script preprocesamiento.py con funciones de limpieza y normaliza..." (commit a8a7ba5). To the right of the message area, there are several status sections: Revisores (No reviews), Cesionarios (Nadie-Asignarte a ti mismo), Etiquetas (Ninguno todavía), Proyectos (Ninguno todavía), Hito (Sin hito), Desarrollo (La correcta fusión de esta solicitud de extracción podría solucionar estos problemas. None yet), Notificaciones (Personalizar Cancelar suscripción), and 1 participante (mishell.ramos@unach.edu.ec).

Mensaje de confirmación
Merge pull request #1 from MISHELL172/feature-preprocesamiento

Descripción extendida

REVISOR: Andrés Alvarado
Pull Pedido: Feature preprocesamiento - Agregar script de preprocesamiento
Ramas: base: principal, compare: feature-preprocesamiento
Buen trabajo, Mishell Ramos
El archivo preprocesamiento.py esta correctamente ubicado en la carpeta scripts/ y tiene una estructura clara.

Este commit será realizado por mishell.ramos@unach.edu.ec.

Confirmar fusión Cancelar

Imagen 9

This screenshot shows a GitHub pull request merge confirmation dialog. At the top, it says "I=Fusionado MISHELL172 Se fusionaron 2 confirmaciones en principal Delaware preprocesamiento de características 2 minutes ago" (I=Merged MISHELL172 Merged 2 confirmations in principal Delaware preprocessing characteristics 2 minutes ago). The message area is identical to the previous screenshot. To the right, the status sections show: Revisores (Sin reseñas), Cesionarios (Nadie-Asignarte a ti mismo), Etiquetas (Ninguno todavía), Proyectos (Ninguno todavía), Hito (Sin hito), Desarrollo (La correcta fusión de esta solicitud de extracción podría solucionar estos problemas. Ninguno todavía), Notificaciones (Personalizar Cancelar suscripción), and 1 participante (mishell.ramos@unach.edu.ec). A message at the bottom states: "La solicitud de extracción se fusionó correctamente y se cerró. Ya está todo listo: la rama feature-preprocesamiento se puede eliminar sin problemas." (The extraction request was merged correctly and closed. Everything is ready: the branch feature-preprocesamiento can be deleted without problems.)

I=Fusionado MISHELL172 Se fusionaron 2 confirmaciones en principal Delaware preprocesamiento de características 2 minutes ago

Conversación 0 Compromisos 2 cheques 0 Los archivos cambiaron 1 Dueño ...

MISHELL172 comentó 12 minutes ago

Este Pull incorpora el archivo scripts/preprocesamiento.py el cual contiene la primera función para el procesamiento de datos. Cambios principales:

- Guiones de creación de la carpeta
- Creación del archivo preprocesamiento.py
- Preparación del proyecto

MISHELL172 agregó 2 confirmaciones 34 minutes ago

Aggrega funciones iniciales de limpieza y normalización c767cd6

Crea script preprocesamiento.py con funciones de limpieza y normaliza... a8a7ba5

MISHELL172 fusionó el compromiso en faa657a principal 2 minutes ago Revertir

La solicitud de extracción se fusionó correctamente y se cerró. Ya está todo listo: la rama feature-preprocesamiento se puede eliminar sin problemas. Eliminar rama