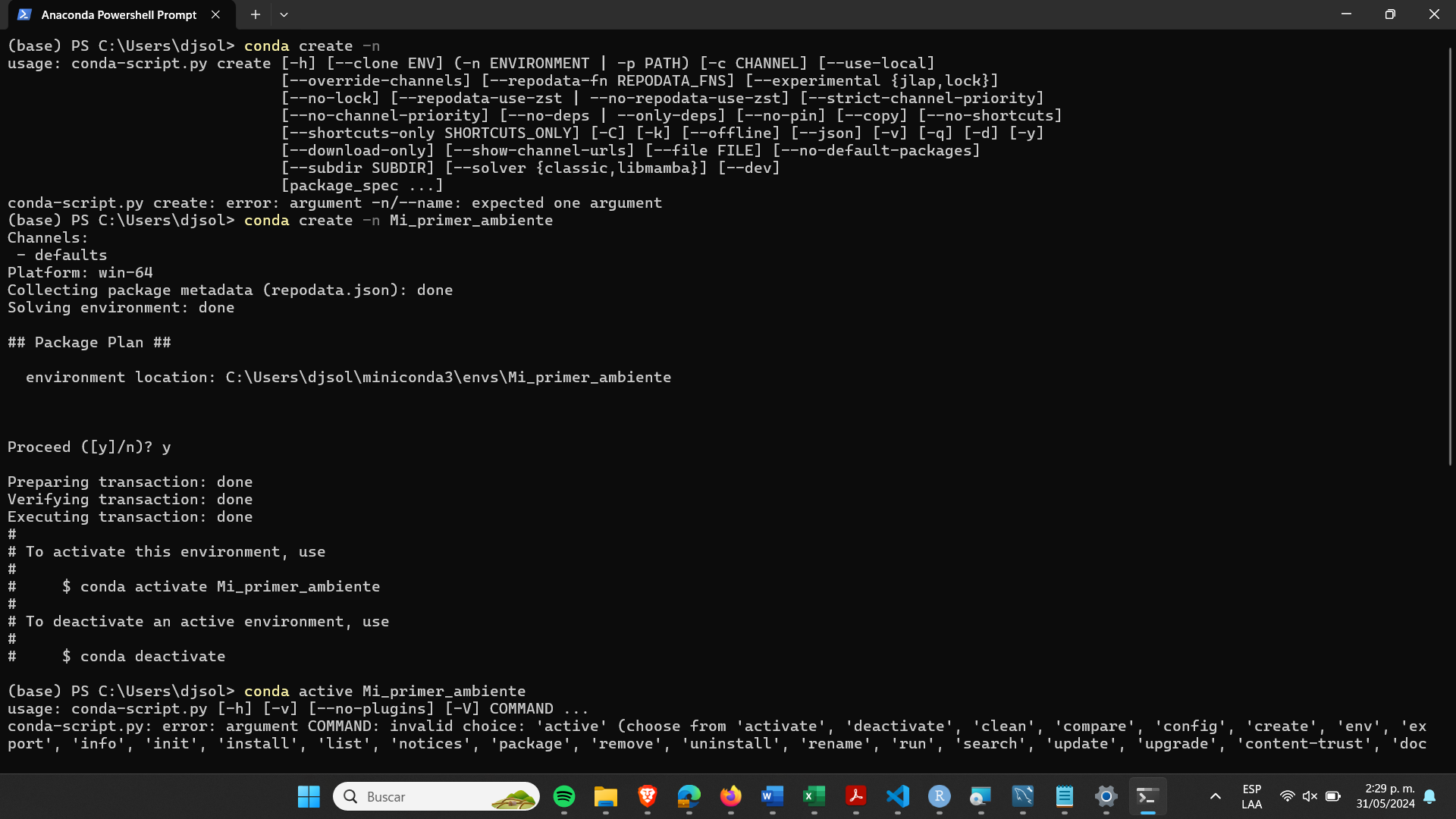
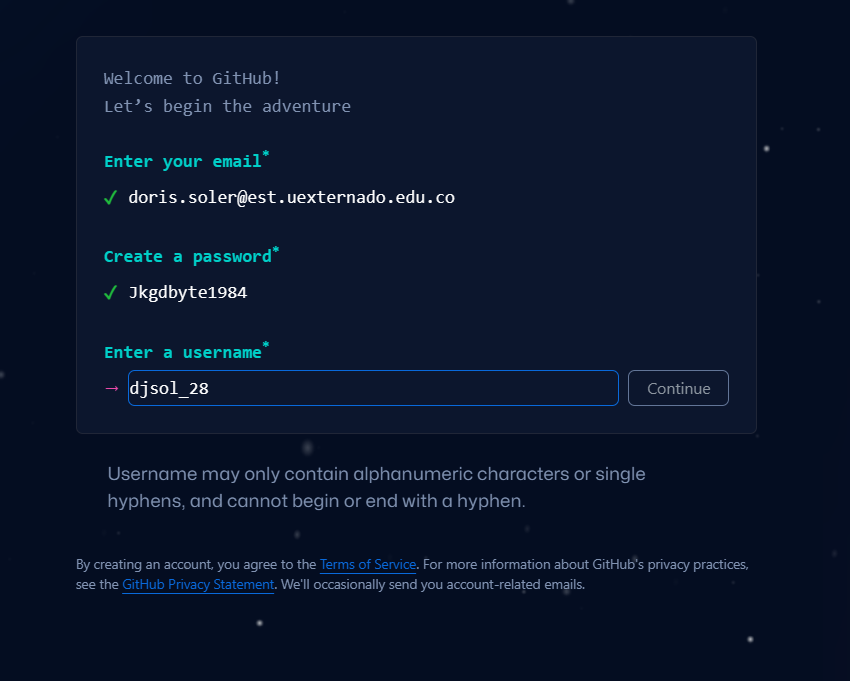
**PROFESOR JORGE MARTINEZ**

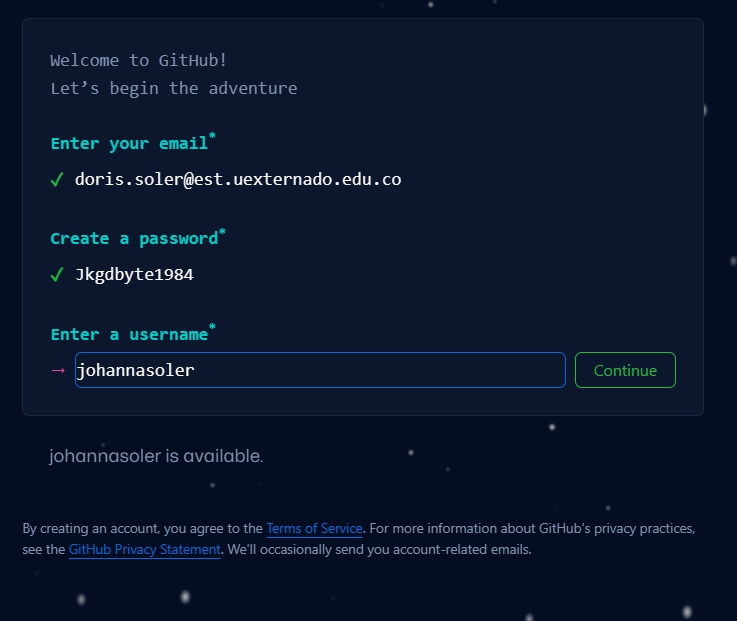
Circulo Dorado

Ver correo

Ensayo







Pagina principal de repsoitorio

1. DNP Unidad de científico de datos
2. Colombia compra eficiente

datos\_abiertos / SOCRATA\_Consulta.ipynb tiene en python los números de los contratos

Aprender: tdsp Microsoft, ARMAR UN PROYECTO, TAREA

APIS

[Crea una API con Node.js y Express - Curso desde cero (freecodecamp.org)](https://www.freecodecamp.org/espanol/news/aprende-a-crear-apis-desde-cero-con-node-js-y-express-curso-desde-cero/)

curso de apis interfaz aplicacion

TEMAS

|  |  |  |  |
| --- | --- | --- | --- |
| Repositorios Github | Python  Pandas Avanzando | Api  Interfaz entre una aplicación y otra, es un puente | Sintaxis ML |
|  |  |  |  |
|  | Clases | Consumir APIS |  |
|  |  |  |  |

[Sacar\_contratos\_SECOP\_II.ipynb - Colab (google.com)](https://colab.research.google.com/drive/1PDgwMoiTzr_2STd9z6Zm0ZPVrMRvvct9)

TAREAS

Hay que hacer un TOKEN

Replicar todo lo del video

TAREA EJERCICIO 28-06-2024

Hacer un clon y en cuadernos:

I\_> Adquisición de datos .ipynb

Hacer Dataset

Poetric Shell

Python 311

**09-08-2024 Agosto 09 de 2024**

Cuatro métodos para consumir APIS

¿QUE ES UNA API?

API es una interfaz que permite la comunicación entre dos aplicaciones de software diferentes

Analogía: Imagina un mesero en un restaurante

TIPOS DE APIS

REST (Representational State

SOAP

GraphlQL

Otros tipos

COMPONENTES DE UNA API

Endpoint: Un punto de acceso único para un recurso especifíco en la API

Métodos HTTP:

* GET

**Método HTTP: GET**

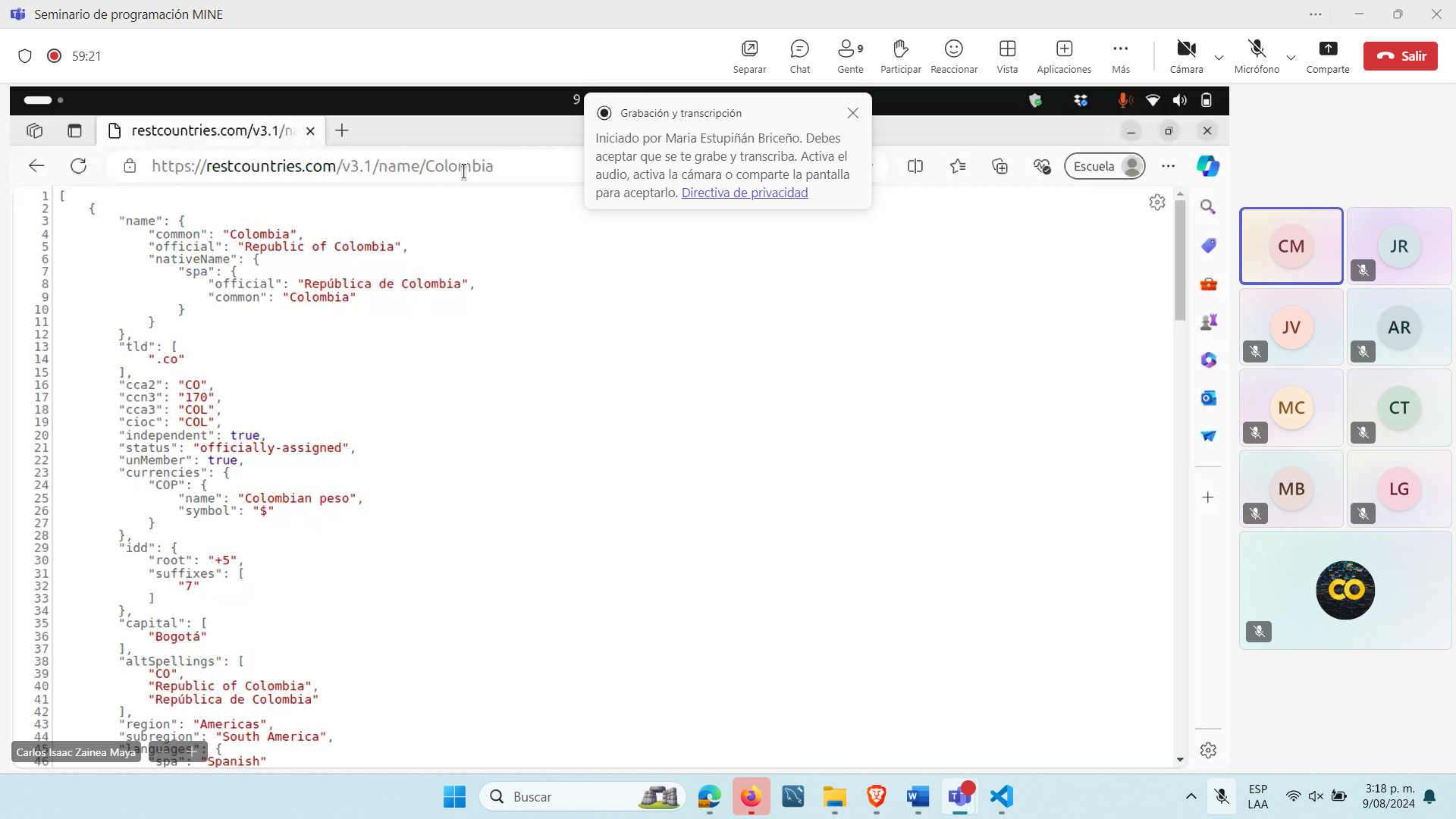
GET se utiliza para recuperar datos de un servidor. Es el método más utilizado para solicitar recursos de una API

Ejemplo: Obtener la información de un usuario por ID

Para hacer en Python una PAI se requiere librería “import request”

API REST

Con esto me trae todo lo de Colombia, lo que esta en cabeza de pagina, tra en lenguaje jeison



Tarea

**Averiguar que hace y cuando se debe usar el código:**

async def root();

return {“message”: “Hello World”}

El código define una función asíncrona en Python utilizando el framework **FastAPI**.

Aquí está el desglose:

python

Copiar código

async def root():

return {"message": "Hello World"}

1. **async def root():**: Declara una función asíncrona llamada root. En FastAPI, esta se utiliza generalmente como un endpoint para una API.
2. **return {"message": "Hello World"}**: La función devuelve un diccionario JSON con una clave "message" y el valor "Hello World". Este será el cuerpo de la respuesta cuando se acceda al endpoint.

Este código probablemente define el endpoint raíz (/) de una API que, cuando se visita, devuelve el mensaje "Hello World" en formato JSON.

**CLASE 23-08-2024**

Curso PL-300 Power BI

<https://learn.microsoft.com/es-es/training/courses/pl-300t00>

Certificación TP-900 Manejar herramientas en la nube

Repositorios

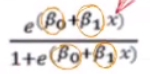
API

Modelos

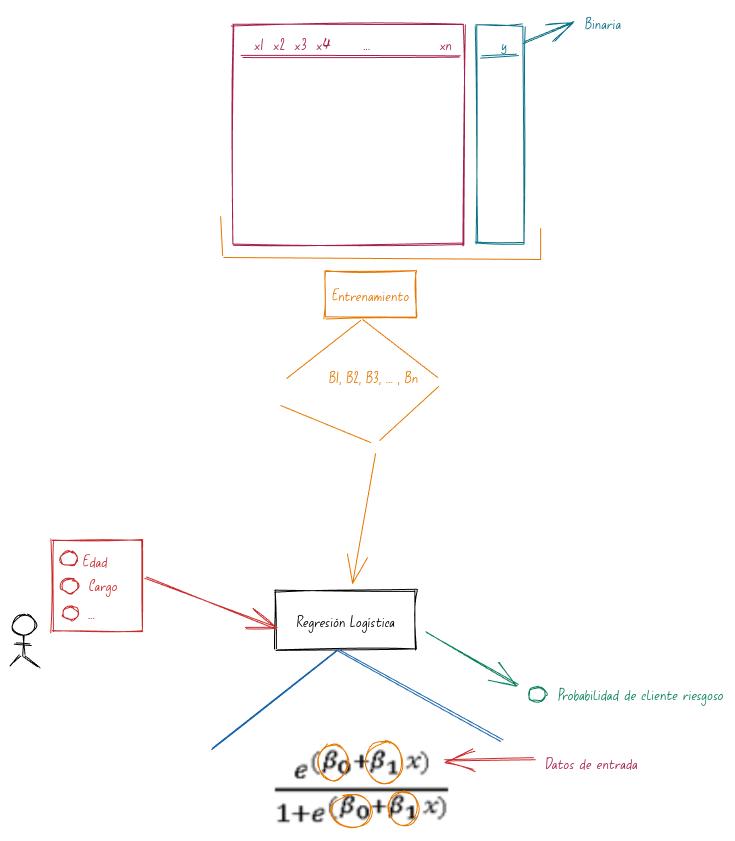
MONTAR UN MODELO DE REGRESIÓN LOGISTICA

Flujograma

Los modelos son formulas, que cuando ya están listos, lo único que hace es ingresar datos, los datos son numéricos, como los X1, X2, X3, X4

Se calculan los Betas, se hace un entrenamiento, que genera estadísticos de validación del modelo, pero lo más valioso son los BETAS, B0, B0, B1, B3, con los BETAS se forma la formula, ene ste caso la formula de Regresión Logística: 

Con eso montado, se determina si el cliente es Riesgoso o NO, esto, en el caso



ENTRENAMIENTO

Se tiene el archivo, se crea una RUTINA que arregla los datos, se genera un DATASET limpio, después se carga para correr el modelo, me genera dos cosas:

1. Un reporte, que sirve para identificar que variable es significativa, se verifica las métricas de calidad\_ p-values (EST) Exactitud, Precisión, Sensibilidad, AUC-ROC, Matriz de confusión.

2. Los Betas: B0. B1, B2, etc, es un archivo que se pueden generar en

L piclule

L model

L jodile

La RUTINA puede hacerse manual la limpieza de los datos o se hace un archivo ejecutable que limpie los datos

TRES FORMAS DE ENTREGAR LOS BETAS A LOS CLIENTES

1. Cargue masivo (Modelos TASH)

2. Aplicativo, desarrollo ONLINE APP

3. ONLINE: API,

O se puede resumir así, la entrega de los BETAS a los clientes:

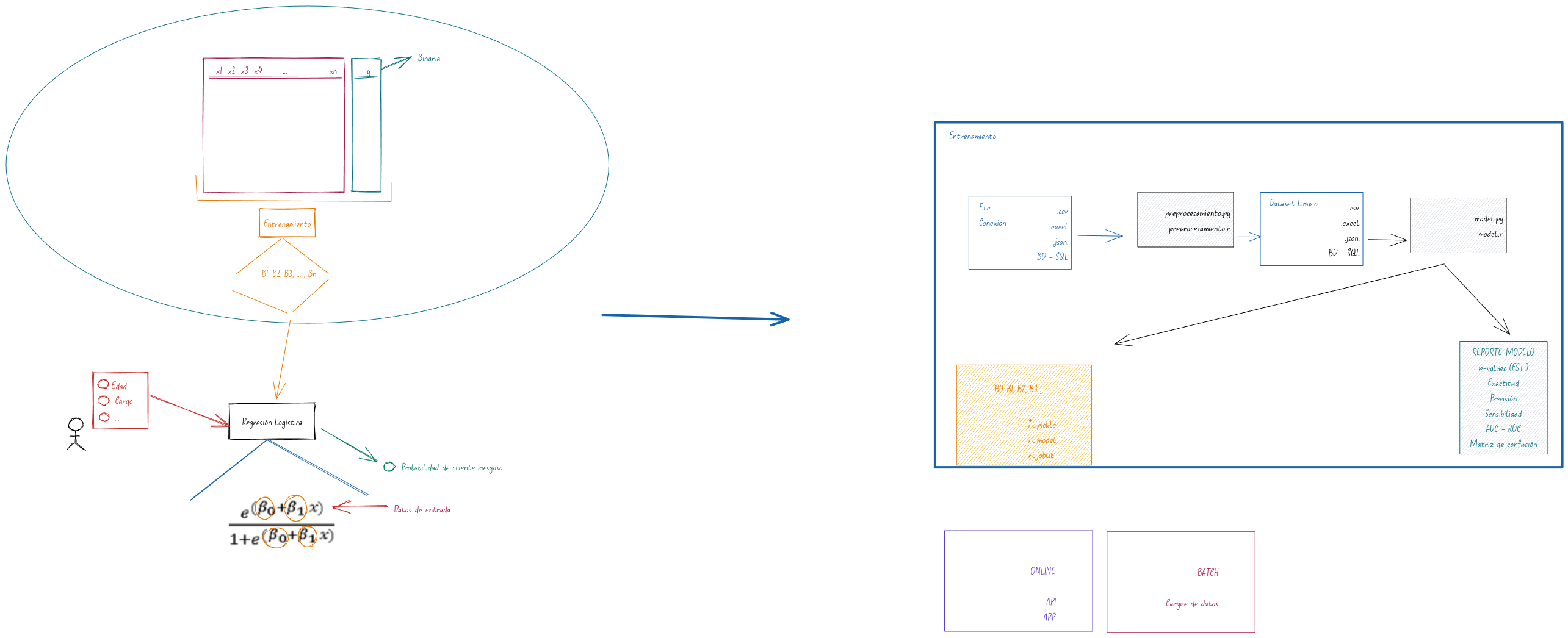
1. ONLINE

API

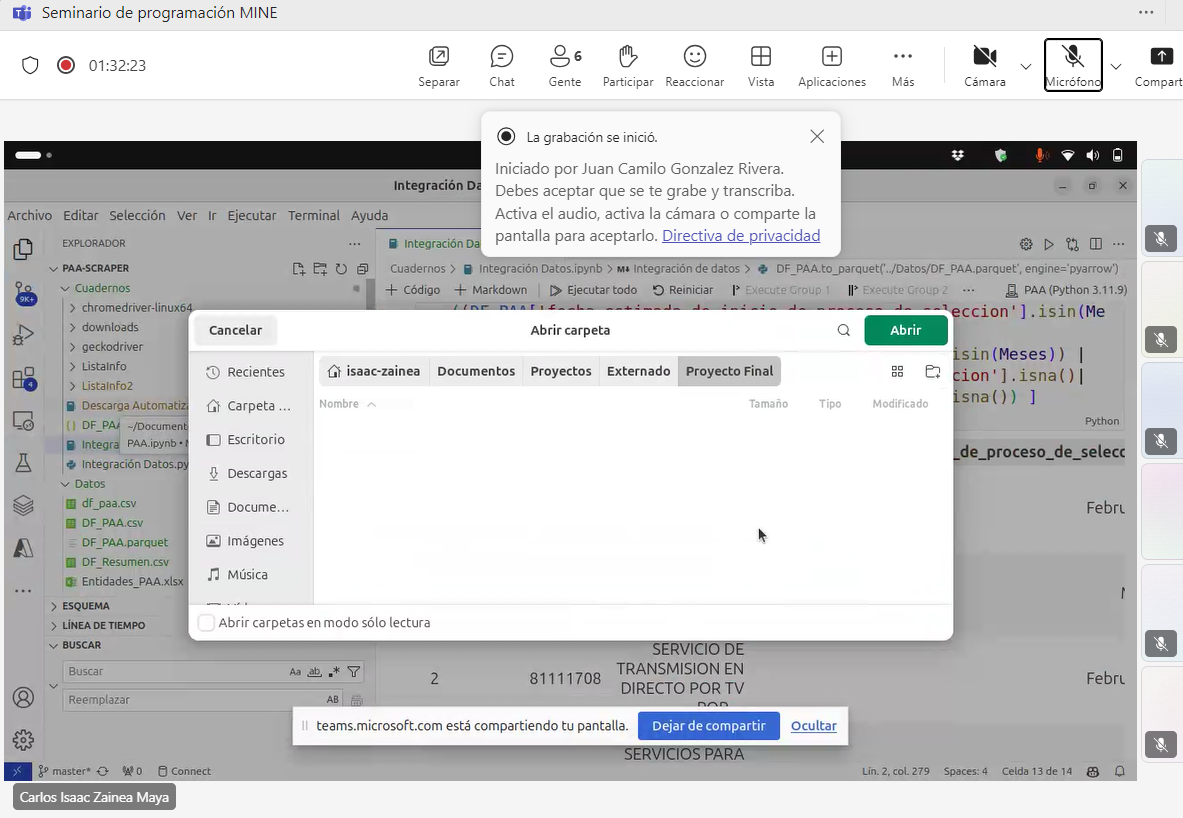
APP

2. BATCH

Cargue de datos



Notas: se puede utilizar POWER APP

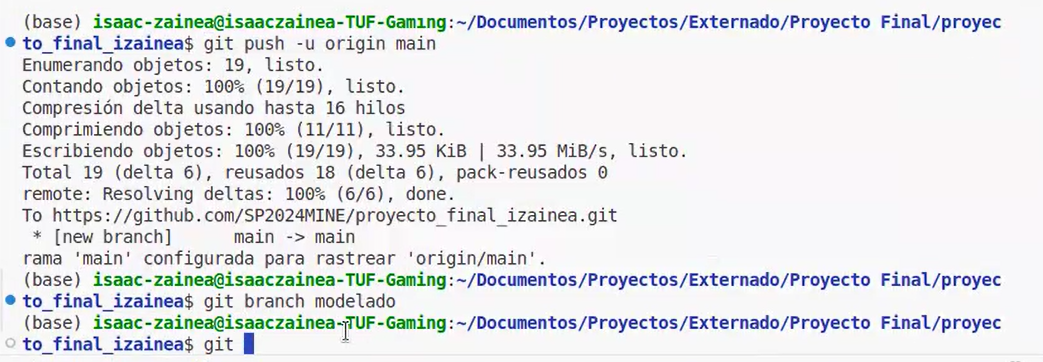


**About**

Es el repositorio ejemplo del curso SP - MINE 2024

En VISUAL STUDIO

Crear una nueva rama comando “git branch modelado”



<https://izainea.github.io/seminario-de-programacion/Cuadernos/Modelos%20de%20Clasificaci%C3%B3n%20I.html>

HIPERPARAMETROS son los BETAS

Los Parámetros son

Un hiperparametro es definir Que método numérico utilizo para hacer la regresión

Otro hiperparametro del árbol es decidir el tamaño del árbol, cuantas decisiones

Aquí tienes una línea de tiempo básica para crear un repositorio en GitHub y trabajar con Visual Studio:

1. **Crear un repositorio en GitHub**:
   * Ve a GitHub y selecciona **New Repository**.
   * Asigna un nombre al repositorio y decide si será público o privado.
   * Haz clic en **Create Repository**.
2. **Configurar ambiente en Visual Studio**:
   * Abre Visual Studio.
   * Clona el repositorio usando **Git > Clone Repository** y pega la URL del repositorio de GitHub.
   * Realiza cambios en tu proyecto.
3. **Guardar cambios en GitHub**:
   * Desde Visual Studio, selecciona **Git > Commit or Stash** para confirmar los cambios.
   * Luego, usa **Push** para enviar los cambios a GitHub.

Esta línea de tiempo te guía desde la creación de un repositorio hasta la clonación y manejo de cambios en Visual Studio.

COPIAR REPOSITORIO DEL PROFESOR

<https://github.com/Izainea/pcd_sp_2024>

CLONAR REPOSITORIO

<https://github.com/Izainea/pcd_sp_2024>

<https://github.com/SP2024MINE/Proyecto_final_Manrique_Bossio.git>

<https://github.com/SP2024MINE/Proyecto_final_Manrique_Bossi.git>

## isaac-zainea@isaaczainea-TUF-Gaming

TAREA

Leer ESTADISTICOS DE VALIDACIÓN, MATRIZ DE CONFUSIÓN, partición 80/20

<https://github.com/Izainea/seminario-de-programacion/blob/main/Cuadernos/Modelos%20de%20Clasificaci%C3%B3n%20I.ipynb>

<https://izainea.github.io/seminario-de-programacion/Cuadernos/Estadisticos%20de%20Validaci%C3%B3n.html>

SEPTIEMBRE 20 DE 2024

COMO FUNCIONA SCIKIT-LEARN

REP

¿CUALES SON LOS ADULTOS CON MÁS DE 50.000?

El ohencoder, las variables conservan el orden, aplica transformación, convierte las variables categóricas en una matriz de 1 y 0

Con get\_document

**¿Para qué sirve el Pipeline?**

1. **Encadenar Múltiples Transformaciones:**
   * Permite aplicar una serie de transformaciones en los datos antes de pasar al modelo. Por ejemplo, puedes estandarizar los datos, luego aplicar PCA para reducción de dimensionalidad y, finalmente, entrenar un modelo de clasificación o regresión.
2. **Reproducibilidad y Organización:**
   * Facilita la organización del flujo de trabajo, garantizando que todas las etapas (preprocesamiento, selección de características, modelado, etc.) se ejecuten en el orden correcto. Esto ayuda a reproducir fácilmente los resultados.
3. **Evitar Fugas de Datos (Data Leakage):**
   * Al incluir todas las transformaciones dentro de un Pipeline, se garantiza que las transformaciones en los datos de entrenamiento y prueba se apliquen de la misma manera, evitando la filtración de información del conjunto de prueba al conjunto de entrenamiento.
4. **Facilita la Validación Cruzada:**
   * Cuando se usa un Pipeline con técnicas de validación cruzada como cross\_val\_score o 'Búsqueda de cuadrícula'GridSearchCV, todas las transformaciones y el entrenamiento del modelo se aplican correctamente dentro de cada pliegue de la validación cruzada, asegurando que no haya contaminación de datos.
5. \*\*Mejor
   * Permite la búsqueda de hiperparámetros no solo del modelo final, sino también de los pasos intermedios, como parámetros de escalado, selección de características, etc., dentro de un flujo de trabajo único utilizando herramientas como GridSearchCV o 'RandomizedSearchCVRandomizedSearchCV.
6. **Simplicidad y Legibilidad:**
   * Hace que el código sea más limpio y fácil de leer. Todo el flujo de trabajo se gestiona como una única entidad, en lugar de tener múltiples líneas de código separadas para cada transformación.

**Ejemplo de Uso de un Pipeline en Scikit-Learn**

U

python

Copiar código

from sklearn.pipeline import Pipeline

from sklearn.preprocessing import StandardScaler

from sklearn.decomposition import PCA

from sklearn.linear\_model import LogisticRegression

from sklearn.model\_selection import train\_test\_split, GridSearchCV

# Datos de ejemplo

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Crear un Pipeline con 3 pasos: Escalado, PCA y Regresión Logística

pipeline = Pipeline([

('scaler', StandardScaler()), # Paso 1: Estandarización

('pca', PCA(n\_components=2)), # Paso 2: Reducción de dimensionalidad

('clf', LogisticRegression()) # Paso 3: Modelo de clasificación

])

# Entrenamiento del Pipeline

pipeline.fit(X\_train, y\_train)

# Predicción

y\_pred = pipeline.predict(X\_test)

En este ejemplo, los datos pasan por los siguientes pasos de manera secuencial:

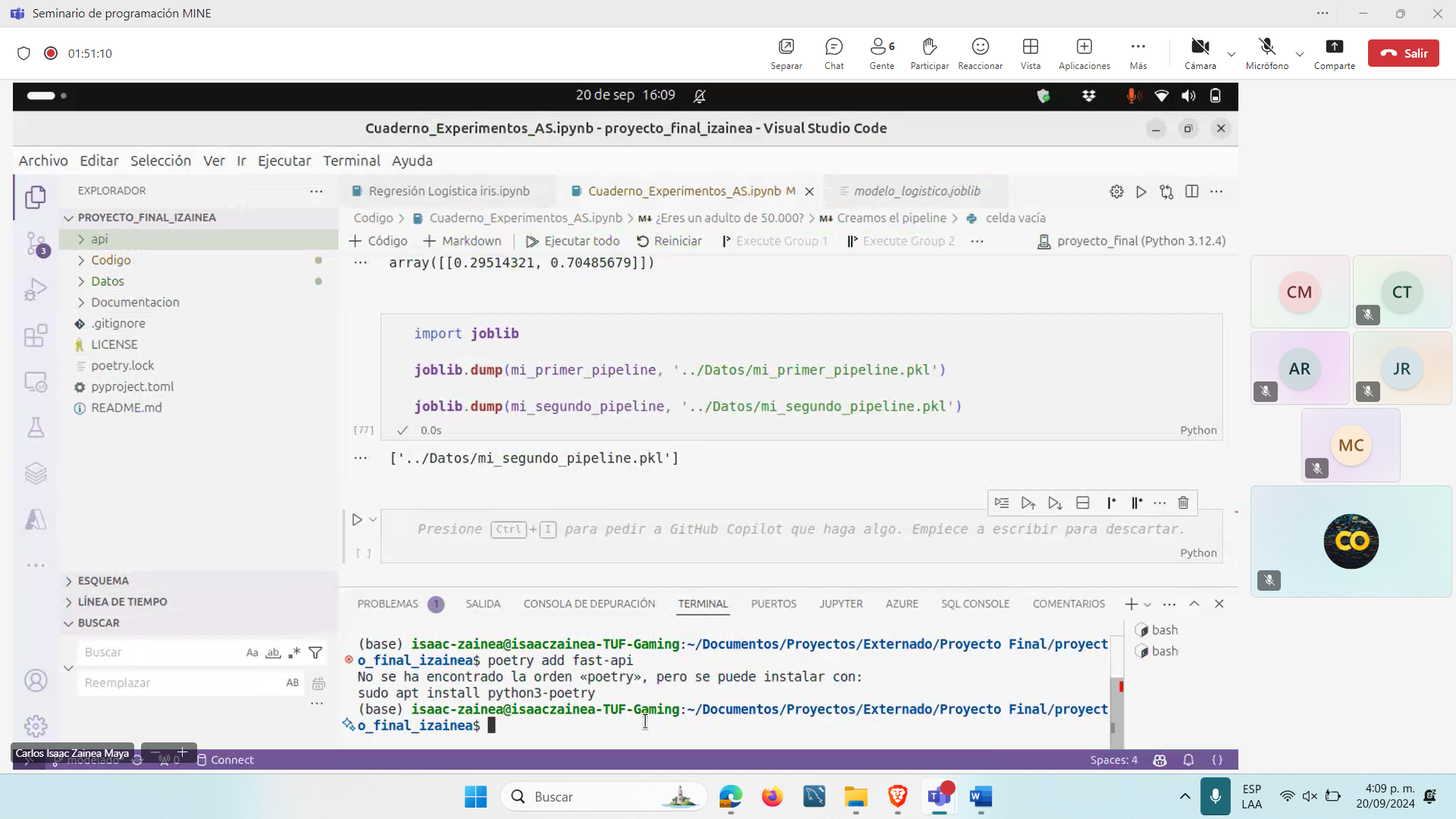
1. Ver esc
2. Se reduce la dimensionalidad con PCA.
3. Se entrena un modelo de LogisticRegression.

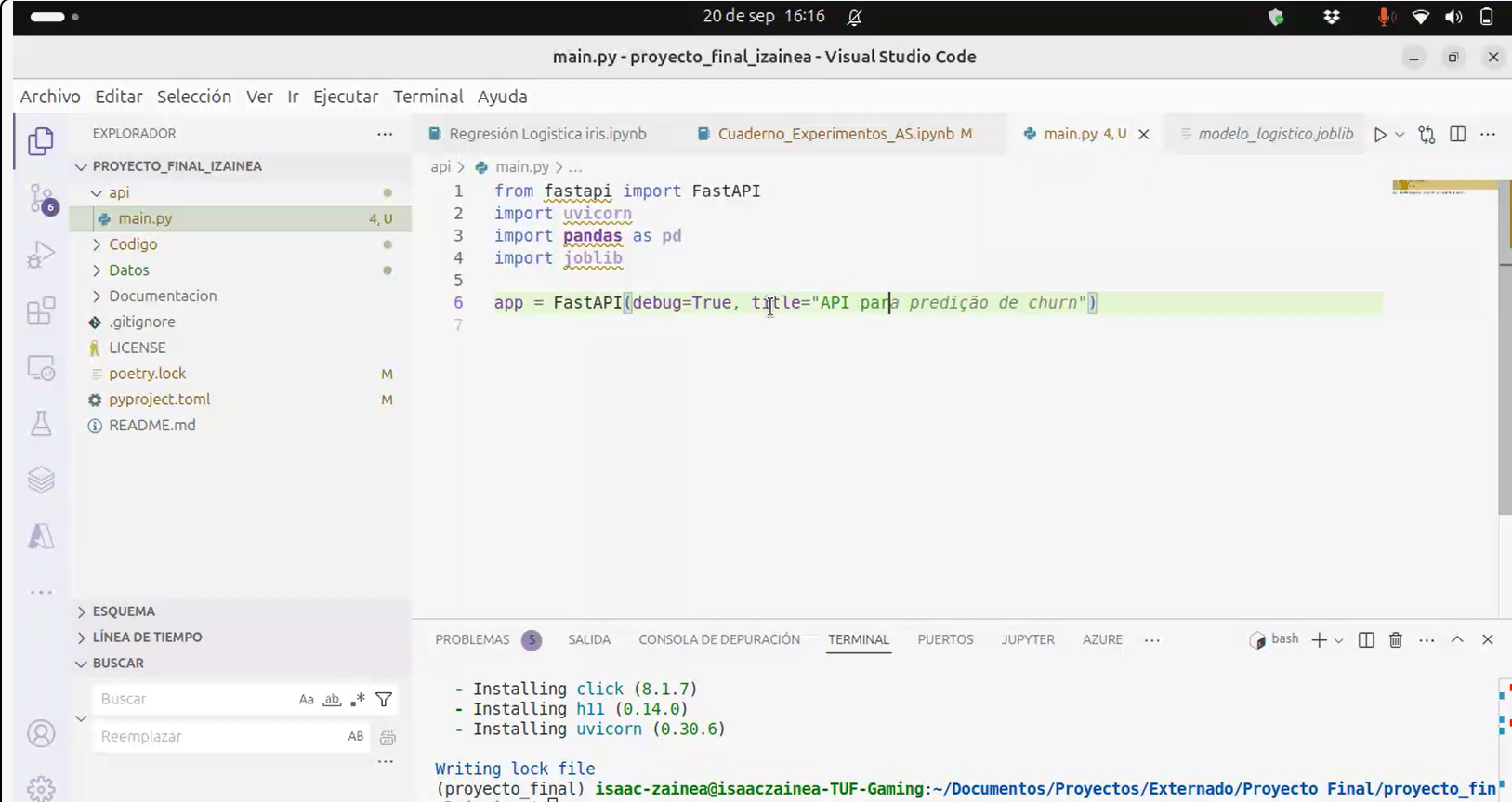
El 'OleoductoGridSearchCV para optimizar el rendimiento.

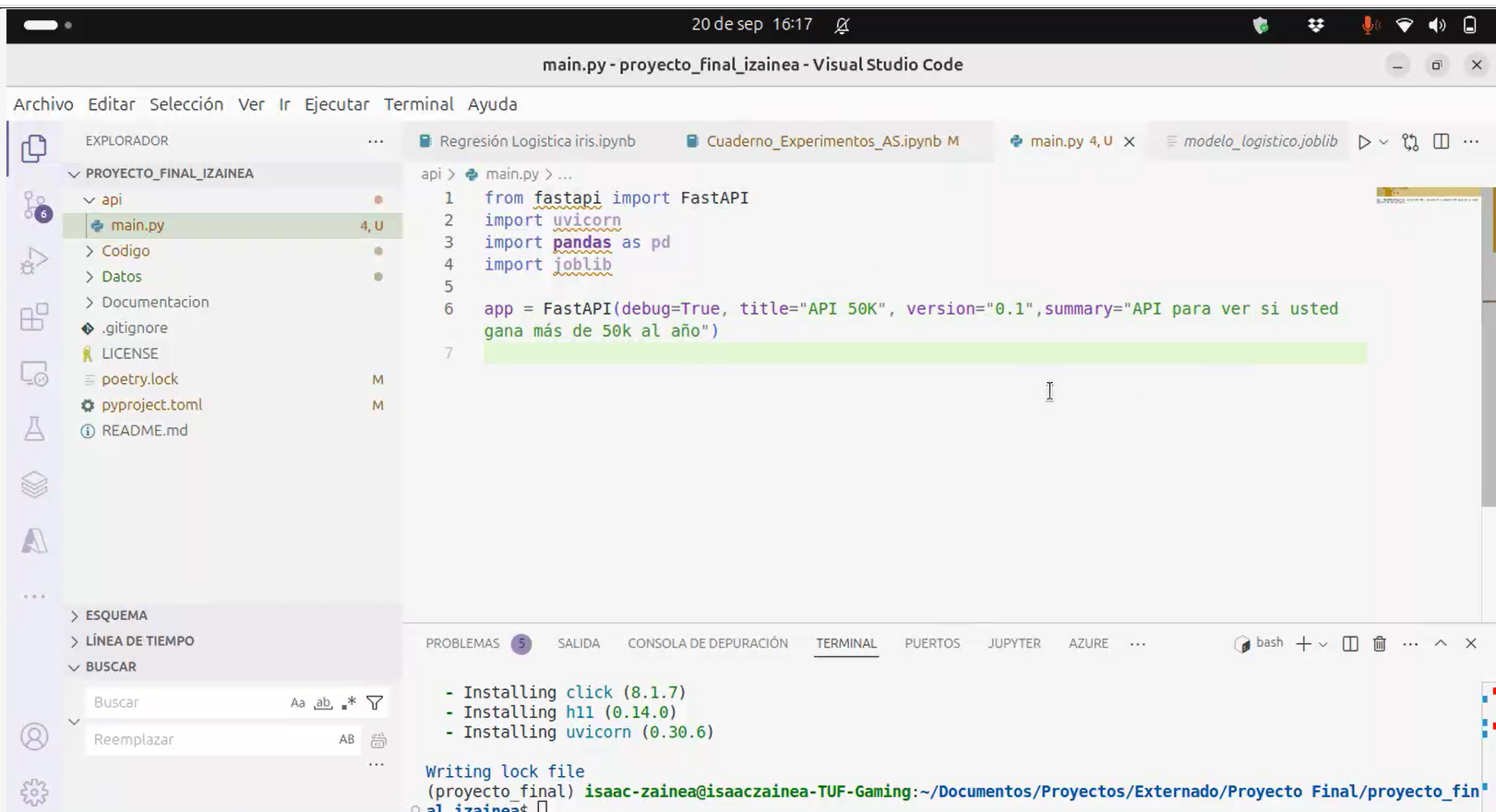
**En resumen:**

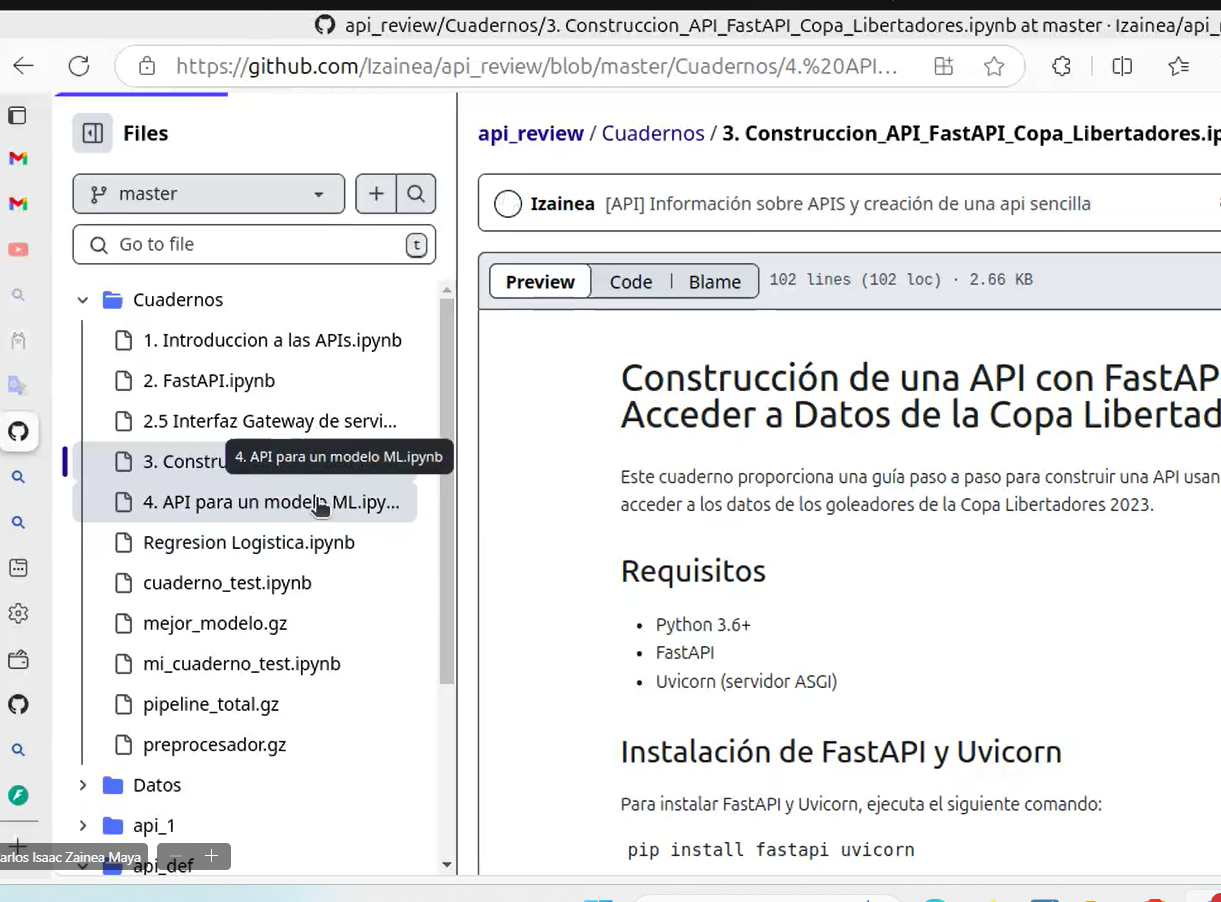
El es una herramienta fundamental para estructurar, optimizar y mantener un flujo de trabajo de machine learning de manera coherente y eficiente, asegurando reproducibilidad y evitando errores comunes como el data leakage.Pipeline

LA API se hace con fast -api









CLASE 04-10-2024

STREAMLIT

Enfoque para hacer aplicativos en base a datos

DataAPP: Se le entrega al usuario formularios o

Revisar lista de errores

El Proyecto es hacer:

Dependencias con Poetry

Montar modelos de Me

Minuto 1:18:00

Proyecto Examen:

UCI Mechear Learning

Ejercicio: datast Marketing Banacario, 17 variables, al final aparece una Y, la idea es replicar ese ejercicio para la Y, la idea es evr si se va a describir a un deposito, se necesita un API que diga si SI o SI No

Para la preddidón se puede utilizar ramdor FOR o una regresión logística.

Armar Piline

Revisar variables y decidir si deja todas las variables o no

Montar API, colocar paginas, STREAMING, aparezca formulario con variables y API con predicción

Repositorio con Poetry, documentar repositorio,

Streamlit: Buscar Galeria y diseñar panel

Se entrega un cuaderno donde se puede testear