

Manual de usuario del Modelo de detección de datos duplicados.

**QUITO – ECUADOR | 2023**

**AUTOR:**  Ing. María Doménica Gómez Sánchez

Ing. Darío Hernán Herrera Herrera

Ing. Cristina Elizabeth Pavón Domínguez

Ing. Jacqueline Vanessa Quispe Tirado

**TUTOR:** MSc. José Luis Pérez Galán

**QUITO – ECUADOR | 2023**

|  |  |  |  |
| --- | --- | --- | --- |
| **RESPONSABLES** | **NOMBRE/CARGO** | **FIRMA** | **FECHA** |
| **AUTORES**  **RESPONSABLES** | Doménica Gómez  Darío Herrera  Cristina Pavón  Jacqueline Quispe |  |  |
| **REVISADO POR** | Iván Reyes |  |  |
| **APROBADO POR** | Iván Reyes |  |  |

Contenido

[1. Introducción. 3](#_Toc151058030)

[2. Objetivo. 3](#_Toc151058031)

[3. Alcance. 3](#_Toc151058032)

[4. Responsabilidades. 4](#_Toc151058033)

[5. Definiciones. 5](#_Toc151058034)

[6. Definición del Proceso. 6](#_Toc151058035)

[6.1 Formato Excel con VBA - Supervisores 6](#_Toc151058036)

# Introducción.

El presente manual de usuario ha sido creado con la finalidad de detallar el proceso de Ejecución del Modelo de Detección de duplicados de los datos de médicos mediante el uso de herramientas como Python, SQL Server, MongoDB, Visual Basic, Excel.

# Objetivo.

⮚ Considerando que los datos médicos, es uno de los activos intangibles más importantes de la empresa, es necesario contar con un documento en el cual se detallen las diferentes actividades que se realizan para mantenerlo operativo y actualizado, para el correcto desenvolvimiento de las actividades derivadas de esta base de datos, así como también aporta en la toma de decisiones sobre inversiones estratégicas y promocionales.

⮚ Describir el proceso de altas médicas y ejecución del modelo de detección de duplicados, con la finalidad de que todos los datos sean los adecuados, mantengan integridad y calidad de la información.

⮚ Utilizar todas las herramientas tecnológicas disponibles en la empresa para tener un adecuado manejo de los datos, actualizados y de calidad, referentes a médicos, supervisores, visitadores, planes de visita, zonas, sub zonas.

# Alcance.

Este procedimiento debe ser observado por todas las áreas involucradas en su administración.

* Gerencia de Ventas
* Gerencia de IT
* Administrador de Datos
* Supervisores y Visitadores

# Responsabilidades.

|  |  |
| --- | --- |
| **Responsable** | **Función** |
| **Administrador de Datos** | 1. Enviar a los supervisores los formatos de altas médicas, con la programación de acuerdo a los catálogos establecidos. 2. Mantener actualizado los datos de los médicos bajo los lineamientos de este procedimiento. 3. Ejecutar el script de Python para la ejecución del modelo de detección de duplicados |
| **Gerencia de Ventas** | 1. Autorizar las Altas de médicos y revisar que el modelo de detección de duplicados cumpla con las condiciones y parámetros establecidos por la compañía. |
| **Supervisores, Visitadores** | 1. Solicitar cambios en la información de los médicos, luego del análisis respectivo y con la autorización de la Gerencia Ventas. 2. Utilizar para el efecto los formatos que requiere el Administrador de datos. |
| **Gerencia de Infraestructura y Tecnología** | 1. Proporcionar al administrador de datos los accesos a las bases de datos necesaria para establecer las conexiones necesaria para la ejecución del modelo. |

# Definiciones.

|  |  |
| --- | --- |
| **Palabra/Sigla** | **Descripción** |
| **Codigo** | Código asignado a un médico, por lo general es el secuencial que le asigna el aplicativo. |
| **Contacto** | Visita realizada por el visitador a un médico asignado. |
| **Plan** | Conjunto de productos asociado a la especialidad del médico, con el cual se realiza la visita. |
| **Domicilio** | Dirección dónde se visita al médico. |
| **Zona** | Es una area geográfica |
| **Subzona** | Conjunto de áreas más pequeños que forma una zona. |
| **Especialidad** | Especialidad principal del médico. |
| **Visitador** | Un colaborador de la Fuerza de Ventas a quien se le asigna médicos con un plan en específico y contactos a visitar. |
| **Supervisor** | Un colaborador que tiene a cargo un grupo de Visitadores en una zona determinada. |
| **Categoría Médicos** | Categoría que se le asigna a un médico de acuerdo al análisis del Supervisor y Visitador. Se clasifica en (Q1, Q2, Q3, Q4). Siendo el Q1 el médico más potencial. |

# Definición del Proceso.

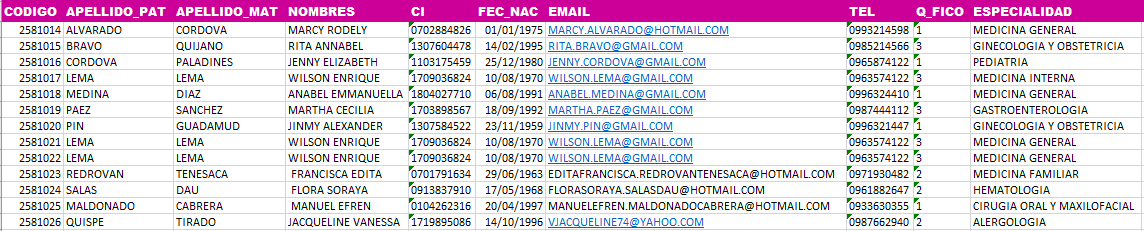
El flujo establecido para la carga de información de médicos en las bases de datos SQL Server, MongoDB utilizado para almacenamiento de históricos es:

## Formato Excel con VBA - Supervisores

El administrador de datos enviara a cada uno de los supervisores un archivo en formato xlsm, que se encuentra programado con ciertas restricciones y catálogos preestablecidos con la finalidad de reducir los errores de digitación.

El supervisor será el encargado de completar el archivo de manera correcta, completando los siguientes campos:

* **Código:** Código secuencial asignado a un médico, asignado por el aplicativo.
* **Apellido\_Paterno:** Primer apellido del médico.
* **Apellido\_Materno:** Segundo apellido del médico.
* **Nombres:** Nombres completos del médico.
* **Ci:** documento de identidad valido del médico.
* **Fec\_Nac:** fecha de nacimiento del médico.
* **Email:** correo electrónico del médico. Vía por la cual se establecerá contacto.
* **Tel:** número telefónico del médico. Vía por la cual se establecerá contacto.
* **Q\_Fico:** Categoría que se le asigna a un médico de acuerdo al análisis del Supervisor y Visitador.
* **Especialidad:** Especialidad principal del médico.

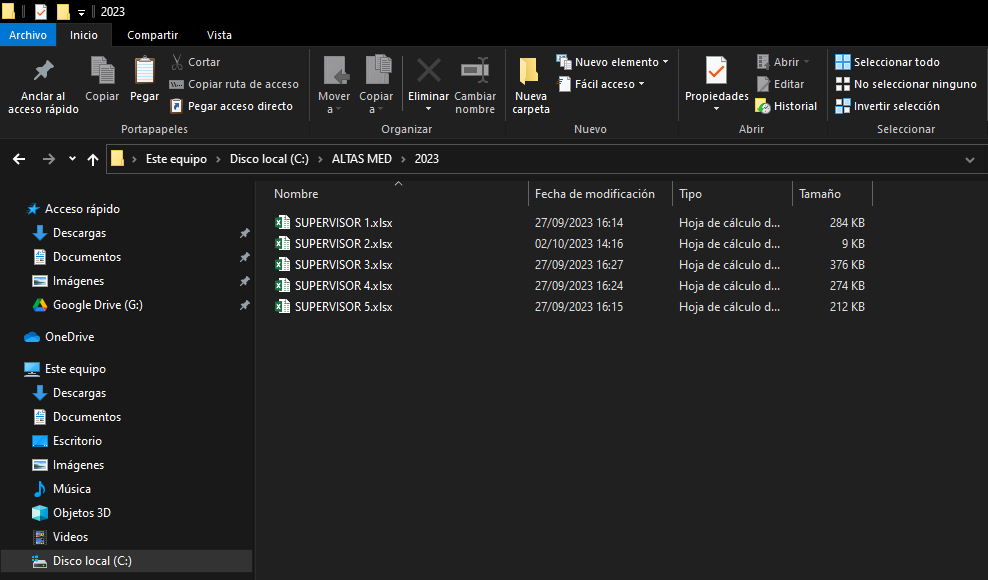


* **Dirección**: lugar de visita del médico, por lo general es la dirección del consultorio.
* **Zona:** área geográfica conformada por la unión de zonas.
* **Subzona:** Conjunto de ciudades más pequeños que forma una zona.
* **Nombre\_Visitador:** nombre completo del visitador que establecerá contacto con el médico.
* **Plan:** Plan con el cual se le visitará al médico.
* **Nombre\_Supervisor:** Personal del equipo de ventas que tiene a cargo al equipo de visitadores.
* **Contactos:** Número de veces que se visitará al médico.

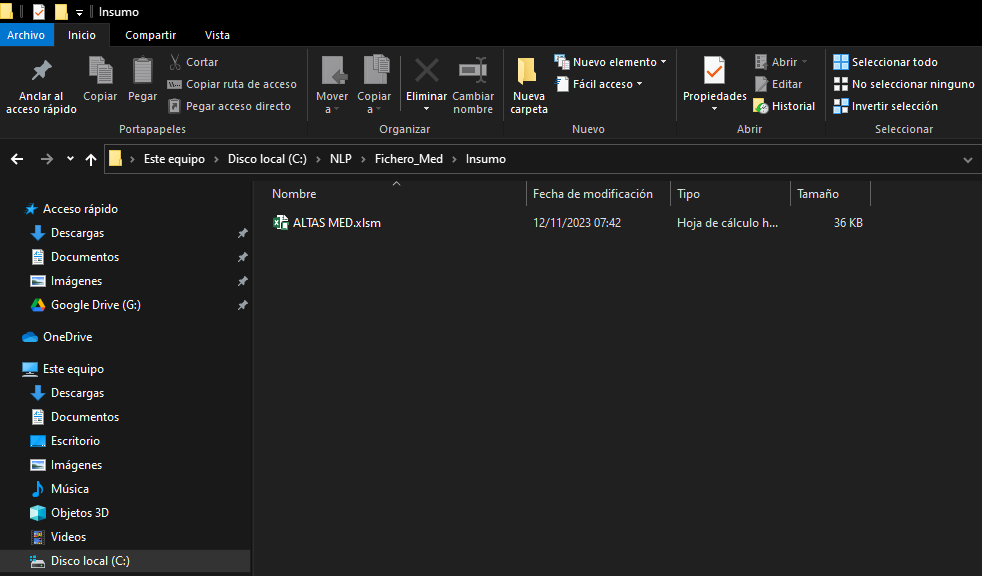


Una vez completado el archivo con la información del médico, el supervisor mediante correo electrónico enviará el archivo al administrador de datos.

El administrador de datos deberá almacenar los archivos Excel en la carpeta ***G:\Unidades compartidas\ALTAS MED\2023***.



Estos archivos deberán ser consolidados, el archivo consolidado se deberá guardar como insumo para la ejecución del modelo en la siguiente dirección: ***C:\NLP\Fichero\_Med\Insumo*** con el nombre ***“ALTAS MED.xlsm”***

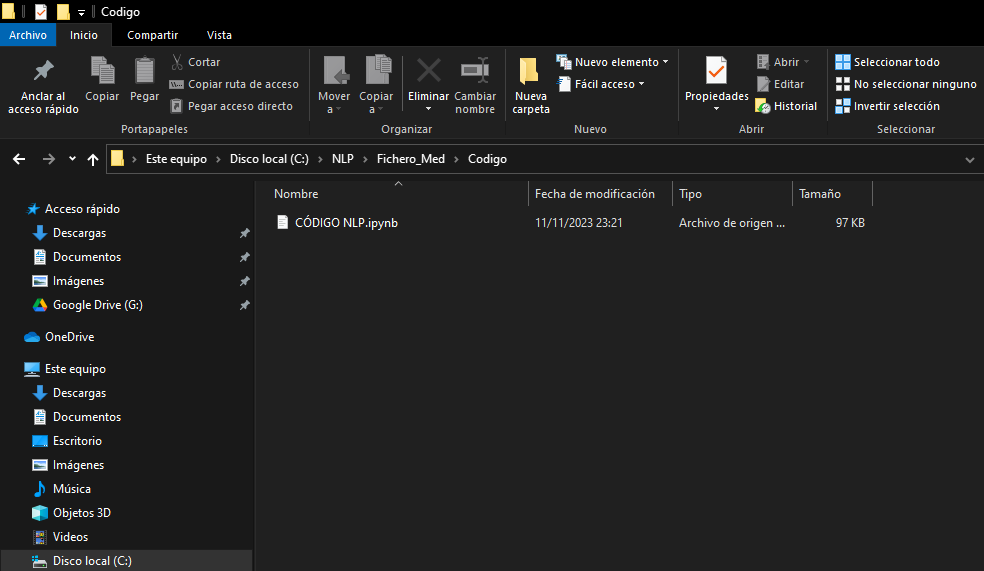


## Ejecución modelo de detección de duplicados – Administrador de Datos

Antes de comenzar con la ejecución del notebook, asegúrate de tener instalados los siguientes programas:

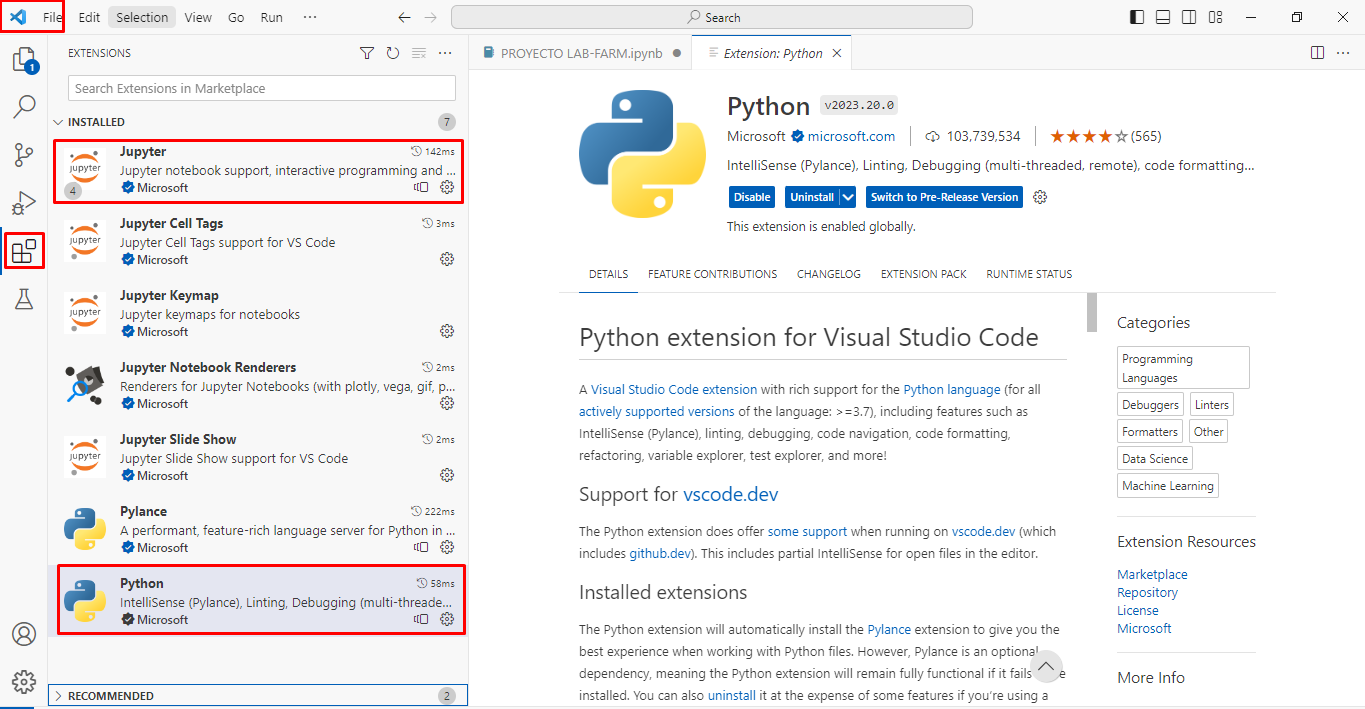
* **Visual Studio Code**: Descarga e instala Visual Studio Code desde https://code.visualstudio.com/.
* **Python 3.11.4**: Asegúrate de tener Python 3.11.4 instalado. Puedes descargarlo desde https://www.python.org/downloads/.

Para descargar el notebook que contiene el código de ejecución, nos vamos a dirigir hasta la siguiente ubicación en tu sistema: **C:\NLP\Fichero\_Med\Codigo**. Busca el archivo **"CÓDIGO NLP.ipynb"** en la carpeta mencionada.

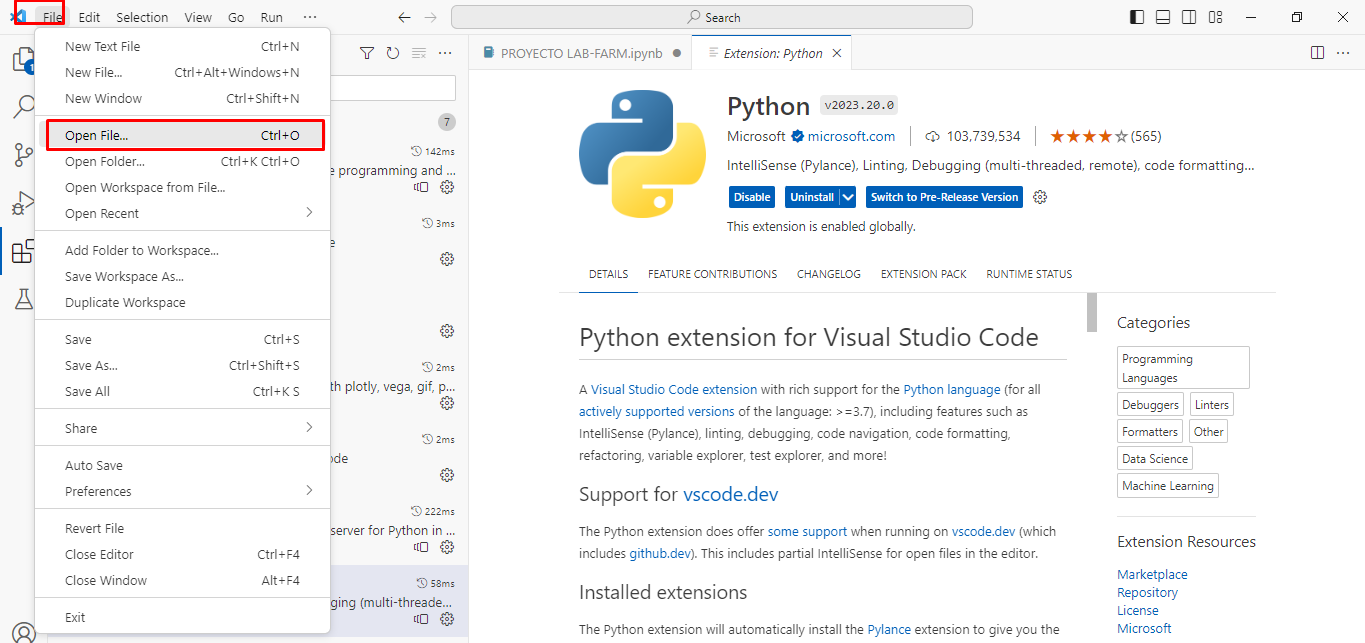


Una vez que se haya cumplido con lo descrito anteriormente, los pasos para ejecutar el modelo:

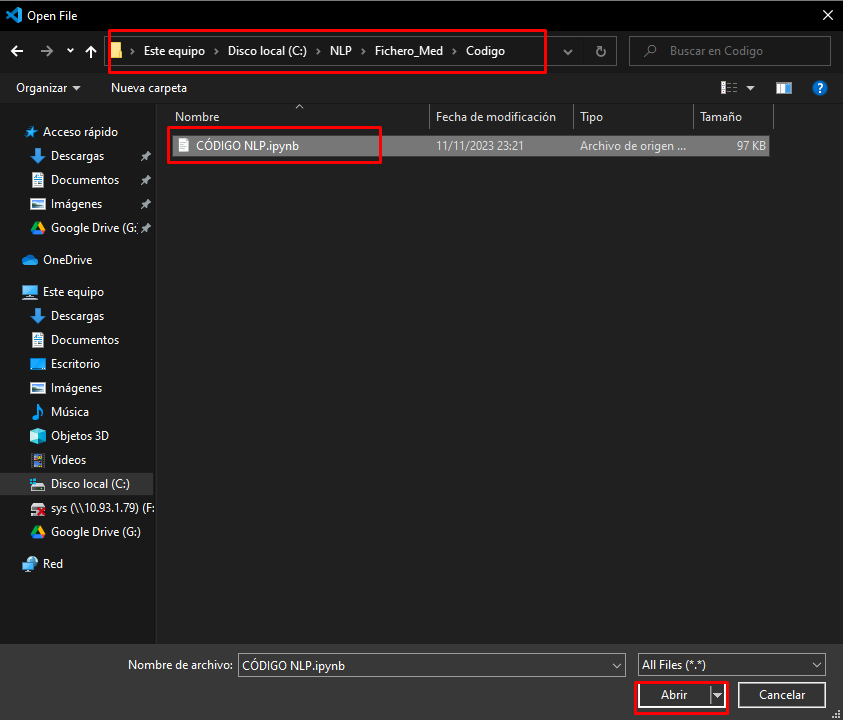
1. Abrir Visual Studio Code.
2. Instala Extensiones necesarias instaladas para trabajar con notebooks de Python en Visual Studio Code.



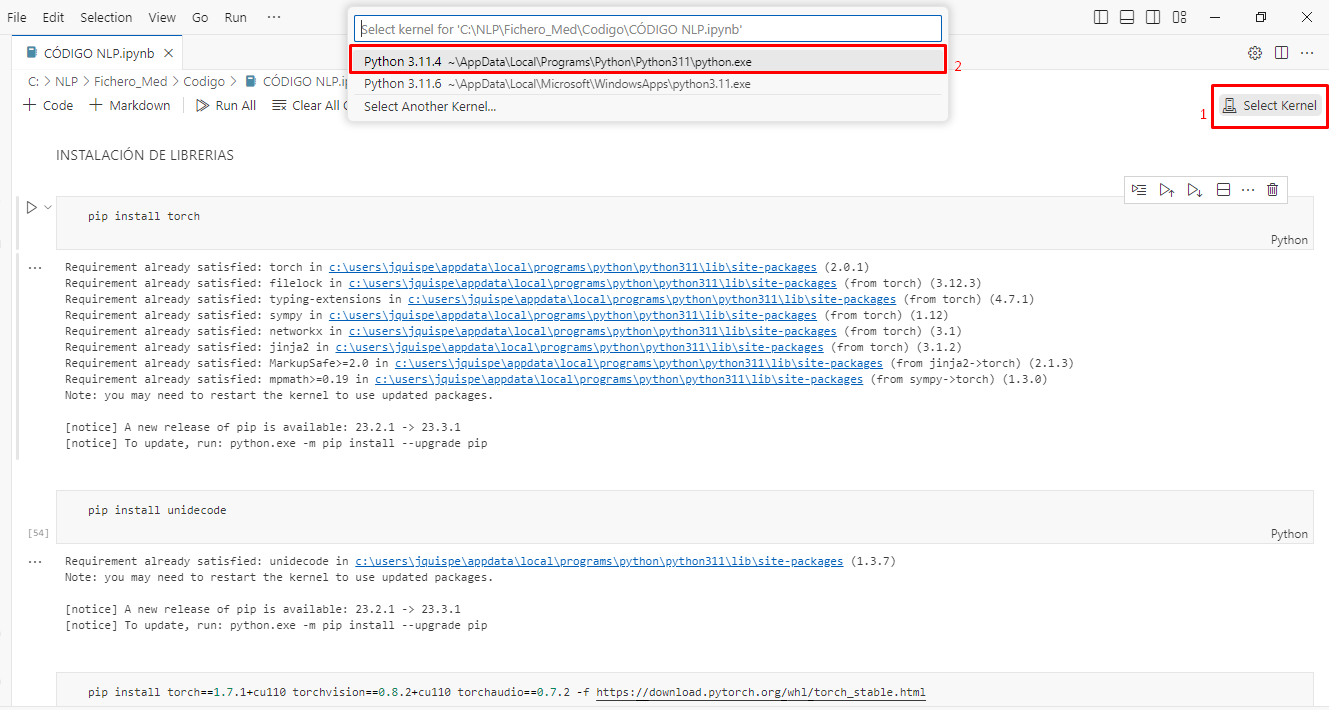
1. **Abrir el Notebook**: En Visual Studio Code, selecciona la opción "File" en la esquina superior izquierda y elige "Open File".



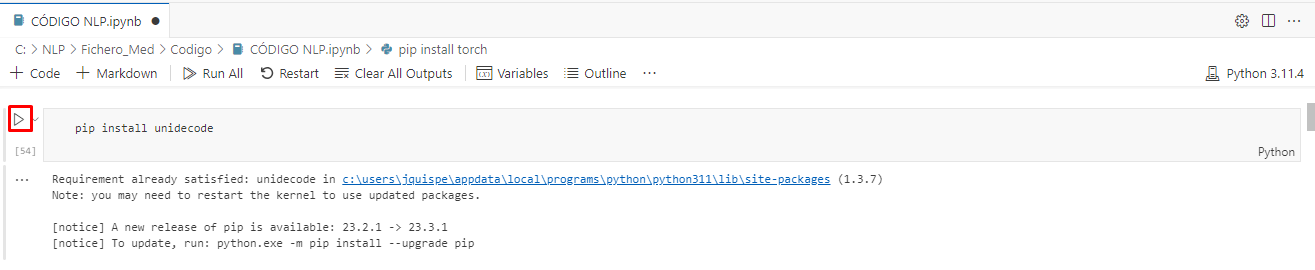
1. Navegamos hasta la carpeta donde se encuentra el archivo ***"CÓDIGO NLP.ipynb***" y damos clic en abrir.



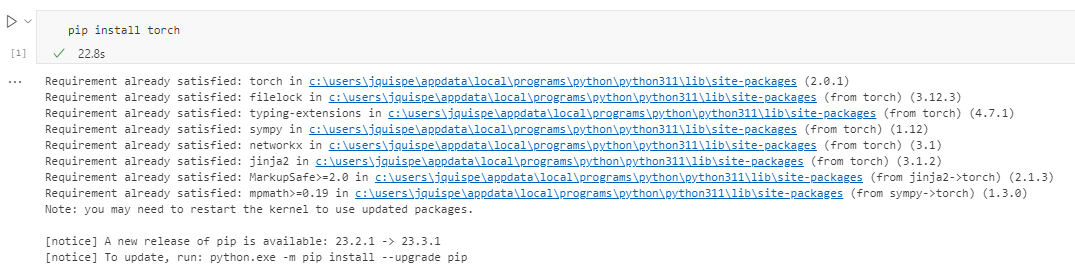
1. **Verificación de Configuración del Kernel:** En la parte superior derecha del notebook vamos a dar clic sobre Select Kernel y se desplegará el listado del kernel instalados, en donde seleccionaremos la opción Python 3.11.4.



1. **Ejecutar las Celdas**: Una vez verificado los pasos anteriores, se procederá a ejecutar las celdas del notebook en orden. Puedes hacerlo dando clic sobre el triángulo ubicado de la parte izquierda de cada una de las celdas o seleccionando una celda y presionando Shift + Enter.



1. **Observa la Salida**: Después de ejecutar el código de la celda, se podrá visualizar un visto de color verde y el tiempo que tardo la ejecución del código. Así como también en la parte inferior de la celda aparecerá el resultado de la ejecución.



Una vez ejecutada todas las celdas del código, se podrá realizar consultas en la base de históricos en MongoDB. Así como también los nuevos médicos creados en la base de SQL Server.