



UNIVERSIDAD NACIONAL DE COLOMBIA

INSTRUCTIVO

Aplicativo Mapas de Conocimiento

Autor: Jairo Castellón Torres
Email: jcastrellont@unal.edu.co

Director: PhD. Campo Elías Pardo Turriago
Email: cepardot@unal.edu.co

Facultad de Ciencias - Maestría en Estadística

7 de octubre de 2024

Índice

1. Introducción	3
2. Instalación	3
3. Uso del aplicativo	4
4. Recomendaciones de uso	7

1. Introducción

Este documento presenta el instructivo de instalación, requisitos y uso del aplicativo para la construcción de mapas de conocimiento desarrollado en el trabajo final de maestría en ciencias - Estadística, titulado “Desarrollo de una aplicación para la construcción de mapas de conocimiento generados por un tema de investigación”.

El desarrollo de esta herramienta, junto con todos sus requisitos se encuentra en el siguiente repositorio de *GitHub*: https://github.com/jcastrellont/Aplicativo_mapas_conocimiento. En este repositorio se encuentran los diferentes archivos necesarios para la instalación y puesta en uso del aplicativo.

En términos generales, el aplicativo toma como insumo el área de estudio en la que se está interesado, las fechas entre las cuales se quiere analizar, y un número máximo de registros para tomar en cuenta. Posteriormente, la herramienta tomará de la base de datos de *ArXiv* (<https://arxiv.org/>) los títulos y resúmenes (*abstract*) de los artículos (hasta completar la cantidad máxima de registros dada por el usuario). Cabe aclarar que arXiv es un servicio de distribución gratuito y un archivo de acceso abierto para casi 2.4 millones de artículos académicos en los campos de física, matemáticas, ciencias de la computación, biología cuantitativa, finanzas cuantitativas, estadística, ingeniería eléctrica y ciencia de sistemas, y economía.

Una vez recolectada la información de los documentos investigativos, se realizará el filtro de las fechas, luego, la herramienta hará todo el proceso de creación de mapas de conocimiento, indicando cuales son los temas más relevantes en los que se está investigando en el área dada, así como muestra la forma en que se relaciona con los otros temas.

2. Instalación

Para la adecuada instalación de los archivos y la ejecución óptima del aplicativo, es necesario que tenga instalado *Python* en una versión igual o superior a la 3.9, así como también contar con *Jupyter Notebook*.

A continuación se enumeran los pasos a seguir para la instalación:

1. **Instalar los archivos directamente desde el repositorio de *GitHub*:** Como se mencionó anteriormente, los archivos pueden encontrarse en el siguiente repositorio: https://github.com/jcastrellont/Aplicativo_mapas_conocimiento. Para obtenerlos, existe una opción de descarga directamente desde la página del repositorio. Otra opción es usar desde la terminal el comando

```
git clone https://github.com/jcastrellont/Aplicativo_mapas_conocimiento.git
```

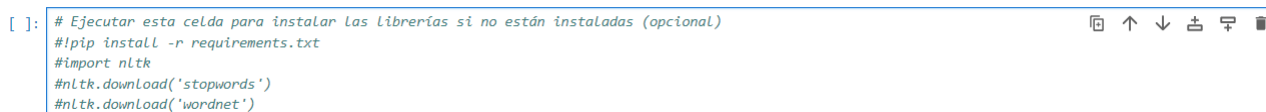
siempre y cuando se tenga instalado en el computador el paquete *Git* (<https://git-scm.com/downloads/win>)

2. **Archivos descargados:** Una vez descargados los archivos del repositorio, se observarán 8 archivos, entre los cuales esta “.ipynb_checkpoints” que actúa como un mecanismo de recuperación automática, así como también la carpeta “__pycache__” que es generada automáticamente por *Python* y contiene archivos que son versiones compiladas de los módulos de *Python* y también se encuentra el presente documento instructivo. Los otros 5 archivos son los necesarios para la ejecución del aplicativo:

- **README.md:** Es un archivo que contiene información general de la herramienta de desarrollo, en este caso, contiene la información de bienvenida y los datos de insumo necesarios para la ejecución del aplicativo.
- **requirements.txt:** Es un archivo plano que contiene las librerías con sus respectivas versiones necesarias para la correcta ejecución del aplicativo.
- **funciones_construccion.py:** Este archivo tiene todas las funciones desarrolladas para la construcción de la herramienta, cualquier modificación en este archivo podría implicar cambios en el funcionamiento del aplicativo.

- **Aplicativo_Mapas_de_conocimiento.ipynb:** Este *notebook* de *Python* contiene las líneas que se ejecutarán para que el aplicativo empiece a funcionar. Básicamente cuenta con dos celdas de códigos, la primera, que se ejecutará solo una vez en el primer uso y es la instalación de todas las librerías y elementos necesarios para que el aplicativo funcione correctamente, y la otra celda que es para que se despliegue el aplicativo y genere el mapa de conocimientos basado en la información que suministra el usuario.
 - **mapas:** Una vez se ejecuta el aplicativo y este genera el mapa de conocimientos, dicho mapa quedará almacenado en esta carpeta con el nombre que proporcionó el usuario. El archivo generado será una imagen en formato *.PNG*.
3. **Instalación de librerías:** El único archivo con el que el usuario tendrá interacción es con el que tiene extensión de *notebook*, es decir, con el archivo “Aplicativo_Mapas_de_conocimiento.ipynb” el cual, se recomienda abrirse desde el explorador de *Google Chrome*. Para la instalación de las librerías, con las versiones adecuadas para la ejecución del aplicativo, el usuario debe ejecutar la primera celda de código que se encuentra en el *notebook* de *Python*.

En la figura 1 se muestra la celda con el código que debe ejecutarse. En este caso, el código aparece comentado (con los signos *#* al inicio de cada línea) debido a que este código solo debe ejecutarse la primera vez que se vaya a hacer uso de la herramienta, posteriormente ya no será necesaria su ejecución.

A screenshot of a Jupyter Notebook cell. The cell contains four lines of commented Python code:

```
[ ]: # Ejecutar esta celda para instalar las librerías si no están instaladas (opcional)
# !pip install -r requirements.txt
# import nltk
# nltk.download('stopwords')
# nltk.download('wordnet')
```

 The cell has a blue border and standard Jupyter Notebook icons on the right.

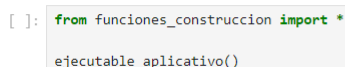
Figura 1: Celda con los códigos de instalación de los requerimientos para el funcionamiento de la herramienta

Una vez se hayan seguido los pasos anteriores, la herramienta estará lista para utilizarse.

3. Uso del aplicativo

Como se mencionó en la sección anterior, el único archivo con el que el usuario deberá interactuar es *notebook* de *Python* llamado “Aplicativo_Mapas_de_conocimiento.ipynb”, el cual se recomienda abrirse en el navegador de *Google Chrome*. En este archivo, una vez hechas las respectivas instalaciones iniciales, se ejecutará la siguiente (y última) celda que está disponible.

La figura 2 muestra las dos líneas de código que contiene esta celda, la cual, al ejecutarse iniciará el funcionamiento del aplicativo.

A screenshot of a Jupyter Notebook cell. The cell contains two lines of Python code:

```
[ ]: from funciones_construccion import *
ejecutable_aplicativo()
```

 The cell has a light gray background and standard Jupyter Notebook icons on the right.

Figura 2: Celda con los códigos de ejecución del aplicativo

Una vez se ejecute esta celda, se abra una ventana nueva en el computador como la que se muestra en la figura 3.

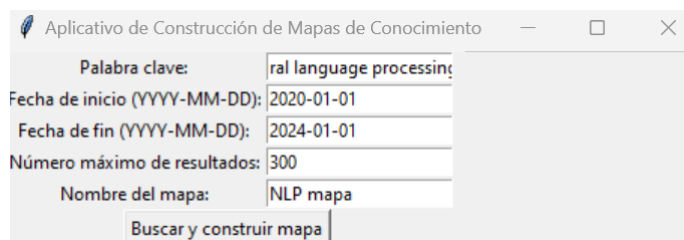


Figura 3: Ventana de inicio de la aplicación

En esta ventana, se le pide al usuario que ingrese información para 5 campos:

1. **Palabra clave:** Este es el nombre del área del conocimiento que se quiere investigar. Recordar que los artículos que se considerarán son aquellos que se encuentran en la base de datos de *ArXiv* (<https://arxiv.org/>) que tiene trabajos académicos en los campos de la física, matemáticas, ciencias de la computación, biología cuantitativa, finanzas cuantitativas, estadística, ingeniería eléctrica y ciencia de sistemas, y economía.
2. **Fecha de inicio:** Permite al usuario seleccionar las fechas de publicación de estos artículos en los que está interesado. La fecha inicial es la fecha del artículo más antiguo en el que esté interesado el usuario. El formato de esta fecha debe ser YYYY-MM-DD.
3. **Fecha de fin:** Es la fecha del artículo más nuevo en el que esté interesado el usuario. El formato de esta fecha debe ser YYYY-MM-DD.
4. **Número máximo de resultados:** Este insumo permite al usuario filtrar por el número total de artículos a considerar. Por cuestión de recursos y de permisos con la plataforma de Arxiv, **se recomienda que este valor no sea superior a los 3000 ni inferior a los 100.**
5. **Nombre del mapa:** La salida de este ejecutable, es una imagen .png que contiene el mapa de conocimiento construido y que será guardada en la carpeta con nombre “mapas” ubicada en la misma carpeta donde quede instalada esta herramienta en su computador local. Por esta razón se pide el nombre del mapa para que quede almacenado y se pueda ubicar fácilmente.

Luego de ingresar esta información, y dar *click* en el boton “Buscar y construir mapa” la herramienta tomará algunos minutos para la construcción del mapa de conocimiento, hasta que aparezca en pantalla una ventana indicando que el mapa se construyo con éxito. La figura 4 muestra la ventana que aparecerá una vez el mapa de conocimiento se haya creado.

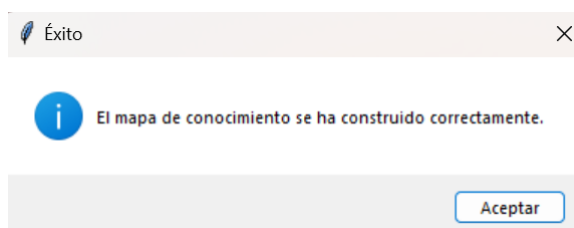


Figura 4: Ventana de aviso de fin de ejecución

Para visualizar el mapa construido, existen dos formas. La primera, como se mencionó antes, se creará un archivo .PNG en la carpeta llamada “mapas” que se descargó en conjunto con todos los archivos del aplicativo. La figura 5 señala la carpeta en la que quedará guardado el mapa construido, mientras que la figura 6 muestra un ejemplo de como quedará guardado el archivo generado.

Nombre	Fecha de modificación	Tipo	Tamaño
.ipynb_checkpoints	3/10/2024 2:24 p. m.	Carpeta de archivos	
__pycache__	5/10/2024 4:21 p. m.	Carpeta de archivos	
mapas	5/10/2024 4:27 p. m.	Carpeta de archivos	
Aplicativo_Mapas_de_conocimiento.ipynb	5/10/2024 4:26 p. m.	Archivo IPYNB	499 KB
funciones_construccion	5/10/2024 4:21 p. m.	Python File	18 KB
README.md	3/10/2024 2:23 p. m.	Archivo MD	2 KB
requirements	3/10/2024 2:00 p. m.	Documento de texto	1 KB

Figura 5: Carpeta donde quedan guardados los mapas construidos

Nombre	Fecha	Tipo	Tamaño	Etiquetas
NLP mapa	5/10/2024 4:24 p. m.	Archivo PNG	385 KB	

Figura 6: Archivo generado luego de la construcción del mapa

Otra forma de poder visualizar el mapa es directamente en el *notebook* donde se ejecuta el aplicativo. Una vez se cierre la ventana de la herramienta (figura 3) en el *notebook* aparecerá el mapa construido como se muestra en la figura 7.

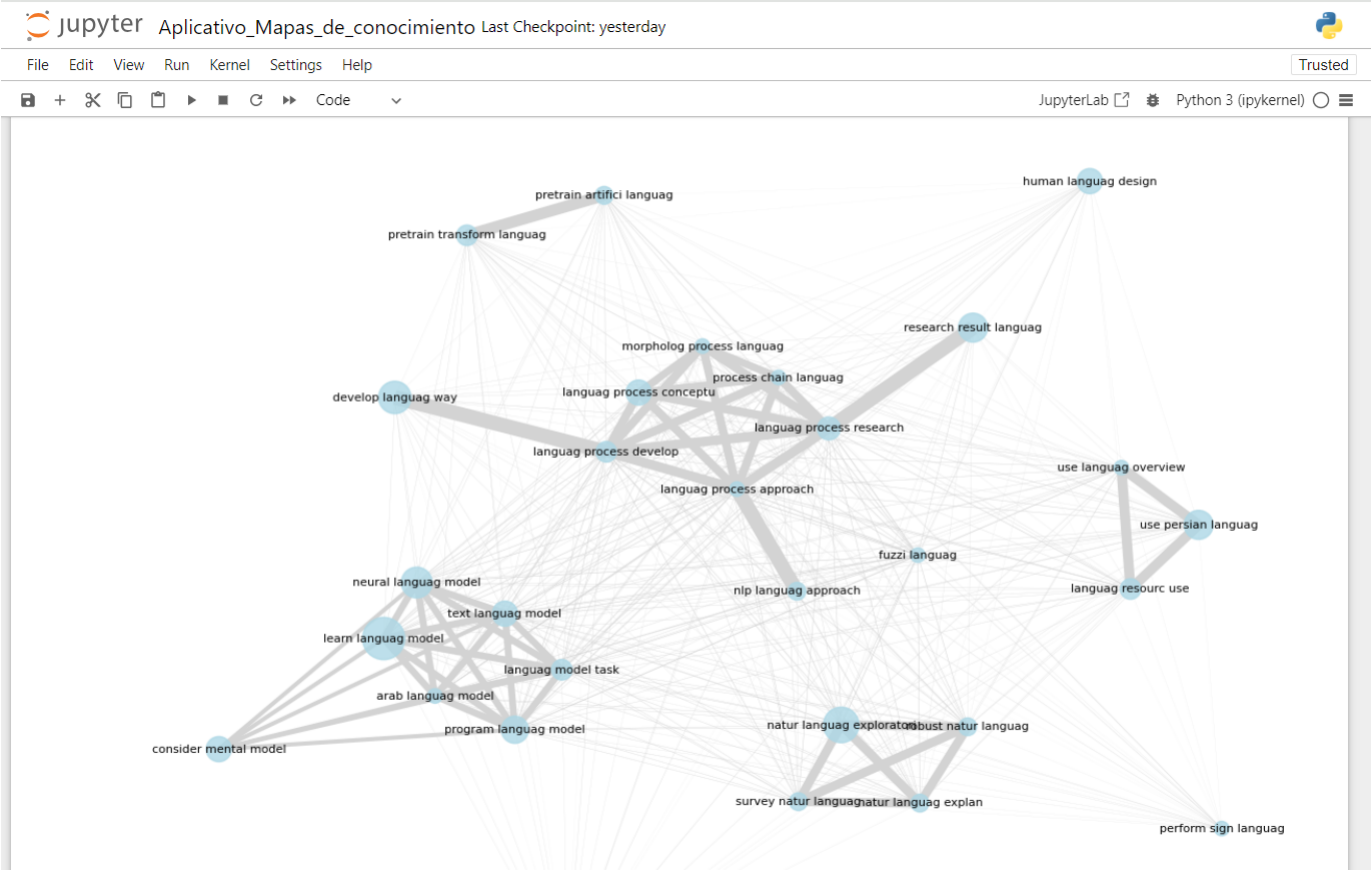


Figura 7: Visualización del mapa en el *notebook* de *Python*

4. Recomendaciones de uso

En esta sección se hará un recuento de las recomendaciones para el buen uso de la herramienta, algunas de estas ya se mencionaron, otras que no, pero que igualmente son importantes tenerlas en cuenta:

1. El único archivo con el que el usuario tendrá interacción es con el que tiene extensión `.ipynb` (“Aplicativo_Mapas.de.conocimiento.ipynb”). No modificar ni eliminar ninguno de los otros archivos.
2. Se recomienda que este *notebook* se abra desde el navegador de *Google Chrome*, si se usa otro navegador, es posible que algunas configuraciones sufran cambios.
3. La instalación de las librerías que se muestran en la figura 1 se hace solamente en una ocasión antes de usar por primera vez la herramienta. Después de esto, se puede dejar el código de instalación comentado.
4. Es importante que se diligencien todos los campos en la ventana de búsqueda (figura 3) si uno de ellos queda mal introducido o vacío, aparecerá un mensaje de error. Por ejemplo, la fecha de inicio siempre debe ser anterior a la fecha de fin.
5. Por costo computacional y para un efectivo uso de la herramienta, se recomienda que en el campo de “Número máximo de resultados”, se digite como mínimo 100 y máximo 3000.
6. La herramienta usa un método de *scraping web* para recolectar la información de la página de *ArXiv*, por lo que desde el servidor de esta página se identificará la IP que esta recolectando la información. Lo anterior implica que en una primera ejecución es posible que aparezca una ventana con un mensaje de error como se muestra en la figura 8. En este caso se recomienda esperar unos minutos y hacer nuevamente la consulta.



Figura 8: Ventana de advertencia de error a la hora de recolectar información de *ArXiv*

7. Otra razón por la que se pueda presentar un error como el que se presenta en la figura 8 es porque se realicen muchas consultas en cortos intervalos de tiempo, en este caso, de nuevo, se recomienda esperar un tiempo considerable para hacer una nueva consulta.