**Manual de instalación**

**Universidad Complutense de Madrid**

**Facultad de Informática**



Enrique Maestro Mañanes

Manuel Guerrero Moñus

# Introducción

Este documento detalla los pasos a seguir para realizar la instalación, ejecución y desinstalación de la aplicación del proyecto de DASI.

AVISO:

* La instalación que se describe a continuación es funcional solamente en el sistema operativo Ubuntu, derivadas de Ubuntu o máquinas virtuales de este.
* Durante el desarrollo del proyecto hemos trabajado con python en sus versiones 3.7 y 3.8.

# Instalación

Para instalar la aplicación, realizar los siguientes pasos:

1. Descargarse el proyecto de GitHub ejecutando en la terminal del sistema:  
     
   git clone <https://github.com/ManuelGMS/DASI>
2. Una vez descargado, acceder a la carpeta “inst-uinst”.
3. Ejecutar la siguiente instrucción para instalar las librerías: “bash install.sh”.

AVISO IMPORTANTE:

* La librería “chatterbot” suele ser problemática:
  + Si observa que su instalación llega a los 10 minutos, aborte su ejecución con CTRL + C, no importa lanzar esta interrupción por teclado ya que es la última librería que se instala. Después de esto, lanzar el comando: “python3 -m pip install chatterbot” y comprobar que llega a instalarse del todo.
  + Si la ejecución de la instalación de la librería “chatterbot” es abortada, lanzar el comando: “python3 -m pip install chatterbot”. Después de lanzarlo, comprobar que la librería llega a instalarse del todo, pues es la última librería necesaria para poder ejecutar la aplicación.
* Compruebe que puede ejecutarse la aplicación según lo especificado en el apartado de “Ejecución”.

# Ejecución

Para ejecutar la aplicación, seguir los siguientes pasos:

1. Abrir una terminal del sistema y posicionarse en la carpeta “DASI”.
2. Ejecutar el comando “bash launcher.sh”, este realizará algunos ajustes y levantará la aplicación, y una vez que eso ocurra, ya podrá utilizarse.

# Desinstalación

Para desinstalar la aplicación, seguir los siguientes pasos:

1. Acceder a la carpeta “inst-uinst”.
2. Una vez allí, ejecutar la siguiente instrucción: “bash uninstall.sh”.

Este script nos preguntará qué librerías del sistema queremos desinstalar y cuáles no. Las librerías de python podrían reutilizarse en otros proyectos y por tanto no se eliminarán.

# Ficheros de código fuente

* Carpeta “inst-uinst”:
  + install.sh: este fichero es para la instalación de las librerías del sistema y de python que son necesarias para la ejecución de la aplicación.
  + uninstall.sh: fichero que contiene un script que pregunta al usuario que librerías del sistema instaladas por el proyecto se desea desinstalar.
* Carpeta “DASI”:
  + launcher.sh: este script configura algunas rutas que deben de estar establecidas antes de lanzar la aplicación mediante la invocación del fichero main.py, de lo contrario, algunas librerías de la aplicación no se encontrarán y por tanto, obtendríamos errores.
  + main.py: fichero que inicializa la aplicación.
  + model.py: modelo de la aplicación, realiza tareas como la carga de los agentes, que son considerados como el nivel de acceso a datos, y también permite pasar la información del usuario que llega por el controlador al agente ChatBot, el cual se ocupará de actuar o informar al agente que corresponda para que realice una tarea.
  + view.py: contiene la ventana y la vista (*Frame*) de la aplicación. Desde este nivel, se propagan mensajes hacia el controlador con información del usuario.
  + controller.py: núcleo de la arquitectura MVC, gestiona todos los eventos que genera la aplicación y orquesta las acciones pertinentes en base a estos.
  + chatBotAgent.py: código fuente del agente ChatBot.
  + classifierAgent.py: código fuente del agente Clasificador.
  + analyzerAgent.py: código fuente del agente Analizador.
  + recomenderAgent.py: código fuente del agente Recomendador.

# Ficheros de recursos

* Carpeta “bbcsport”:
  + Contiene varias subcarpetas de categorías deportivas, las cuales contienen a su vez ficheros de texto plano con noticias relativas a dicha categoría.
    - AVISO: se pueden introducir nuevas noticias en las subcarpetas, pero lo que no se debe de hacer es eliminar esta carpeta o desbalancear en exceso la cantidad de noticias en estas subcarpetas que reflejan las categorías deportivas conocidas.
* Carpeta “chatterbot”:
  + Fichero “database.sqlite3”: contiene información relativa a lo aprendido por el ChatBot con las conversaciones utilizadas en su proceso de entrenamiento.
    - AVISO: su borrado implica que se desea reentrenar el ChatBot otra vez con el fichero de “corpus.json”.
  + Fichero “corpus.json”: contiene las conversaciones con las que entrenar al ChatBot para realizar su tarea.
    - AVISO: se pueden crear nuevas frases de entrada, pero no se deben inventar nuevas frases de salida que no estén ya en el fichero.
* Carpeta “classifier”:
  + newsClassified.csv: un fichero de valores separados por comas que contiene dos columnas, en la primera están los textos de cada noticia y en la segunda la categoría deportiva a la que pertenece cada noticia.
    - AVISO: la eliminación de este fichero indica el reentrenamiento de la máquina de vector de soporte con las noticias del dataset “bbcsport”, por lo que los ficheros de tipo \*.pkl se borrarán y se volverán a generar automáticamente, este fichero también se regenerará de forma automática si es eliminado.
  + svm.pkl: contiene la serialización del clasificador de la aplicación, este está basado en una máquina de vectores de soporte.
    - AVISO: este fichero no se debe modificar o eliminar manualmente.
  + labelEncoder.pkl: es un objeto serializado que contiene las etiquetas de las clases (categorías deportivas) codificadas con identificadores numéricos que van de 0 a (N - 1) clases.
    - AVISO: este fichero no se debe modificar o eliminar manualmente.
  + tFidfMatrixVector.pkl: este objeto serializado es una matriz de vectores TF - IDF asociados al texto de cada noticia del dataset “bbcsport”.
    - AVISO: este fichero no se debe modificar o eliminar manualmente.
* Carpeta “nltk\_data”:
  + Contiene librerías adicionales de NLTK.
    - AVISO: esta carpeta no debe de ser modificada ni eliminada.
* Carpeta “news”:
  + Es una carpeta en la que los agentes de clasificación y análisis buscan un fichero con un cierto nombre para realizar sus respectivas tareas.
    - AVISO: es posible añadir o eliminar noticias de la carpeta, pero la carpeta no debe de ser eliminada.
* Carpeta “docs”:
  + Contiene la licencia de software del proyecto y el fichero de readme.
    - AVISO: no modificar los ficheros.
* Carpeta “inst-uinst”:
  + Contiene los scripts de instalación y desinstalación de la aplicación.
    - AVISO: no modificar los scripts sin conocimiento.