



CLASIFICACIÓN DE PATRONES MEDIANTE CLUSTERING

Milagros del Rocío Peña Quineche
Máster de Ingeniería de Sistemas y Control
Facultad de Informática
Universidad Complutense de Madrid - UNED

1. Índice

- 1. **ÍNDICE..... 2**
- 2. **RESUMEN 3**
- 3. **IMPLEMENTACIÓN 4**
 - 3.1. INTERFAZ 4
 - 3.2. CÓDIGO..... 5
 - 3.3. RESULTADOS 5

2. Resumen

En este documento se explicará brevemente el trabajo realizado para la asignatura de Minería de datos.

El trabajo realizado consiste en la implementación de una herramienta en Java capaz de clasificar tanto señales como imágenes mediante clustering. Los algoritmos, de los que consta la herramienta, son todos los dados en la asignatura, es decir:

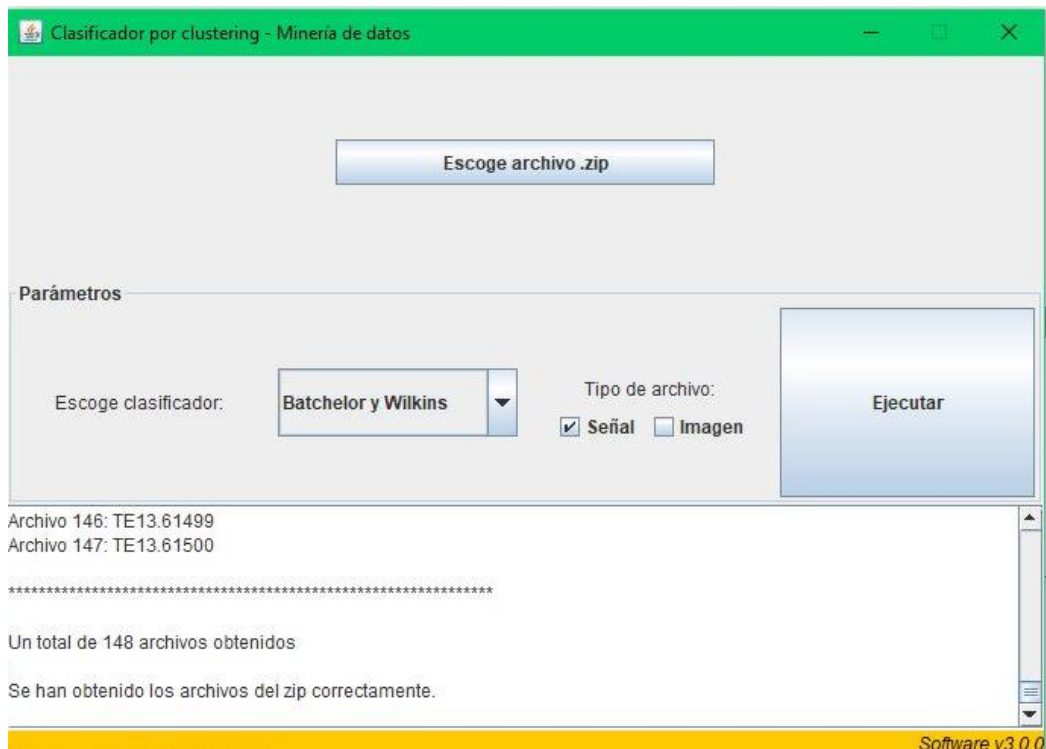
- Método adaptativo.
- Algoritmo de Batchelor y Wilkins.
- Algoritmo jerárquico aglomerativo.
- Algoritmo K-Means.
- Algoritmo por agrupamiento secuencial.
- Algoritmo por matriz de similitud.

La herramienta tiene una interfaz que facilita la elección de los patrones y la especificación de los parámetros de manera intuitiva y sencilla. Además, dispone de una sección en la que se mostrará toda la información sobre el proceso realizado en cada algoritmo.

3. Implementación

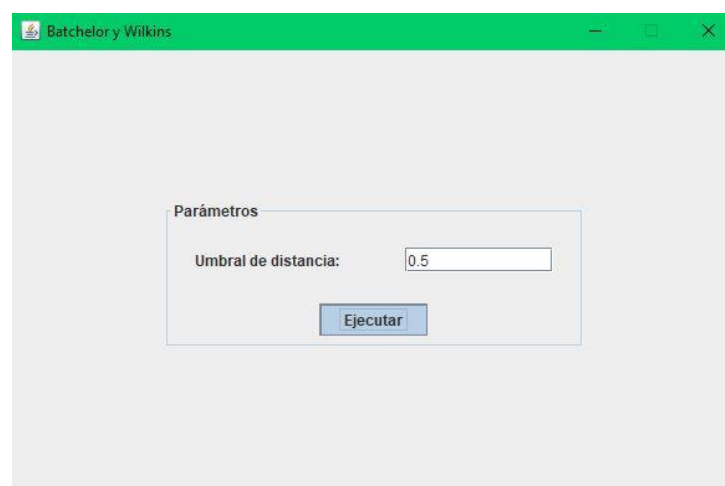
3.1. Interfaz

La interfaz es la siguiente:



Dispone de un explorador de archivos con el que se debe de seleccionar el fichero .zip que contenga las señales o las imágenes.

Una vez seleccionados los archivos se debe de escoger un clasificador e indicar el tipo de información a tratar, por lo que se desplegará otra ventana en la que se especifiquen los parámetros que utilizará el algoritmo escogido.



3.2. Código

Por defecto, los parámetros, que necesitan los algoritmos, estarán inicializados en la herramienta, pero el usuario puede cambiarlos a su gusto siempre.

El trabajo se ha basado en el modelo-vista-controlador, por lo que consta de dos capas:

- **Presentación:** Capa en la que se implementa todo lo relacionado con las vistas y la gestión de eventos mediante un controlador, comandos y un dispatcher.
- **Negocio:** Capa encargada de implementar el servicio de aplicación de la herramienta, es decir, gestiona toda la lógica de las operaciones realizadas. A esto se le añade los transfers para llevar la información y evitar que las capas sepan los datos recibidos y obtenidos, ya que para esto se modula de esta manera el proyecto.

Para evitar acoplamiento se han implementado interfaces y factorías encargadas de separar las capas y facilitar el poder extender la herramienta.

3.3. Resultados

En cuanto a los resultados, para poder ofrecer la mayor información posible de cada cluster obtenido. No solo se muestra en consola un resumen de los clusters resultantes, sino que también se crean directorios de cada cluster en la ruta *C:/Resultados-Clasificadores/*.

En cada directorio se escribirán las señales o imágenes que pertenezcan a ese cluster y un archivo aparte que contenga el centro de cada uno de estos.

Debido a que se tratan muchos datos, existe algún algoritmo como el de Batchelor y Wilkins que tarda bastante más de lo normal que con el resto de los algoritmos.