











GUIA DE USUARIO

Universitat Rovira i Virgili, Spain
Autonomous University of Madrid, Spain
Universidad de Nariño, Colombia
Servicio Geológico Colombiano

Versión 1.0
GNU General Public License
2025



ÍNDICE

GENERALIDADES DEL SISTEMA 4
Objetivo 4
Licencia 4
INSTRUCCIONES PARA LA INSTALACIÓN DE NETDRTOOL 5
1. Cómo obtener NetDRTool 5
2. Instalación de dependencias para JavaDRServer 6
3. Instalación de Microsoft C++ Build Tools 12
4. Instalación de dependencias para PythonDRClient 13
INSTRUCCIONES PARA LA EJECUCIÓN DE NETDRTOOL 17
MANUAL DE USUARIO DE NETDRTOOL 19
1. Ventana principal del sistema 19
1. Barra de navegación 20
2. Opciones del menú 20
3. Zona “Drag and Drop” 20
2. Data 23
2.1. Conexión a los datos 23
2.1.1. Example 24
2.1.2. Plain Text 25
2.1.3. Connection DB 27
2.2. Ejemplo de uso 28
3. Data cleaning 30
3.1. Filtros 31
3.1.1. Standardize 32
3.1.2. Selection 32
3.1.3. Remove Missing 33
3.1.4. Update Missing 33
3.1.5. Range 34
3.1.6. Reduction 35
3.1.7. Replace Value 37
3.1.8. Numeric Range 38
3.1.9. Discretize 39
3.1.10. Codification 41
3.1.11. Proximity 42
3.2. Ejemplo de uso 42
4. Dimensionality reduction 45
4.1. Algoritmos NetDR 46
4.2. Algoritmos Locales 46
4.2.1. LLE 46
4.2.2. LE 47
4.3. Algoritmos Globales 48
4.3.1. PCA 48
4.3.2. MDS 49

4.4. Algoritmos de Kernel 50
4.4.1. KLLE 50
4.4.2. KLE 50
4.4.3. KPCA 51
4.4.4. KMDS 52
4.5. Algoritmos de Combinación 52
4.5.1. MKL 53
4.5.2. DMKL 53
4.6. Algoritmos de Proyección 53
4.6.1. DLPP 53
4.6.2. SPCA 53
4.6.3. GRP 54
4.6.4. SRP 54
4.6.5. FICA 55
4.7. Algoritmos de Aproximación 56
4.7.1. UMAP 56
4.7.2. SliseMap 56
4.7.3. TriMap 57
4.7.4. DensMap 57
4.8. Algoritmos Neuronales 58
4.8.1. PUMAP 58
4.8.2. IsoMap 58
4.8.3. AutoEncoder 59
4.8.4. GraphEncoder 59
4.9. Algoritmos Discriminantes 60
4.9.1. LDA 60
4.9.2. FA 61
5. Views 62
5.1. Tipos de visualizadores 62
5.1.1. Scatter 62
5.1.2. RnxCurves 66
5.1.3. Variance 66
5.1.4. ImageTensor 66
6. Evaluation 66

GENERALIDADES DEL SISTEMA

Objetivo

NetDRTool es una herramienta gráfica para usuarios no expertos en minería de datos, contando con diversas técnicas de reducción de dimensionalidad para bases de datos, permitiendo visualizar los resultados obtenidos a través de su interfaz gráfica y facilitando así la toma de decisiones según estos.

Licencia

El software NetDRTool está bajo la licencia pública general GPL/GNU. Esto significa que se tiene la libertad de compartir o modificar el programa de cualquier forma y de poner el código fuente a disposición de quien lo desee.

INSTRUCCIONES PARA LA INSTALACIÓN DE NETDRTOOL

1. Cómo obtener NetDRTool

Enlace de descarga de la herramienta NetDRTool:
http://grias.udenar.edu.co/grias/?p=591

El archivo .zip incluye la carpeta con el código fuente y otra con las dependencias .jar necesarias, dentro de la carpeta del código fuente hay dos carpetas además del archivo README.md, una se ejecuta como proyecto de Java y la otra con Python.



Para poder llevar a cabo la ejecución del programa es necesario instalar:

1. **Netbeans IDE 8.2:**
   https://filehippo.com/es/download_netbeans/8.2/

2. **Java JDK 8 (o cualquier versión compatible con el IDE):**
   https://www.oracle.com/co/java/technologies/javase/javase8u211-later-archive-downloads.html

3. **Visual Studio Code:**
   https://code.visualstudio.com/Download

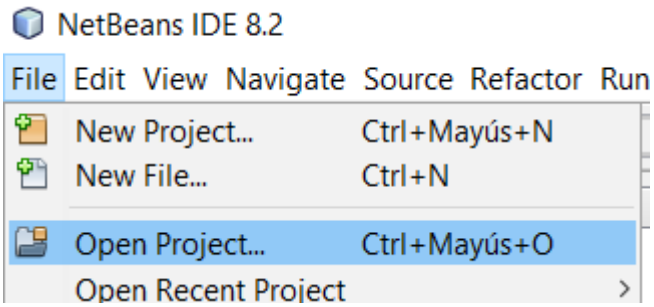
4. **Python (Versión 3.11):**
   https://www.python.org/downloads/release/python-3110/

5. **Microsoft C++ Build Tools**
   https://visualstudio.microsoft.com/visual-cpp-build-tools/

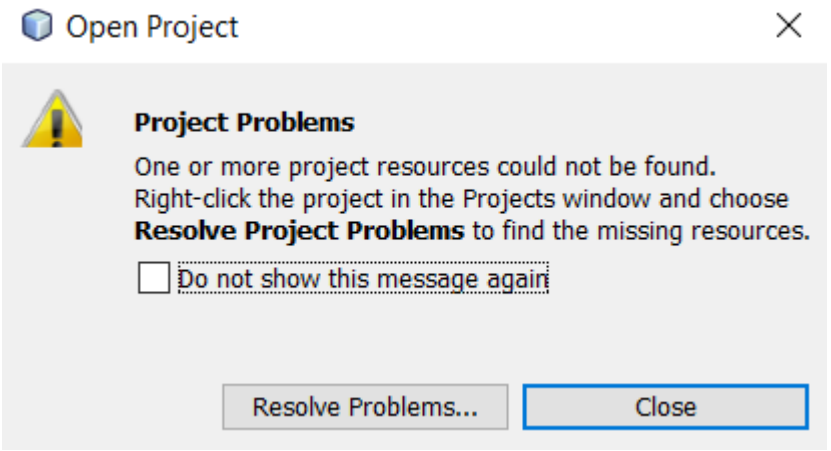
6. **Dependencias para Java y Python**

2. Instalación de dependencias para JavaDRServer

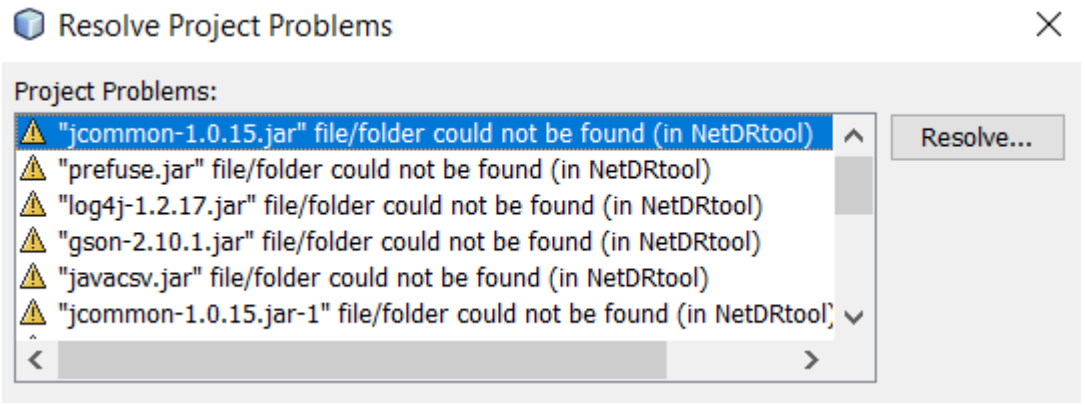
Se abre el proyecto JavaDRServer en Netbeans



La primera vez presentará inconvenientes debido a que las dependencias no se encuentran vinculadas con el proyecto.



Para solucionarlo damos click en “Resolver Problemas…” y se nos presentará una ventana con todas las dependencias faltantes



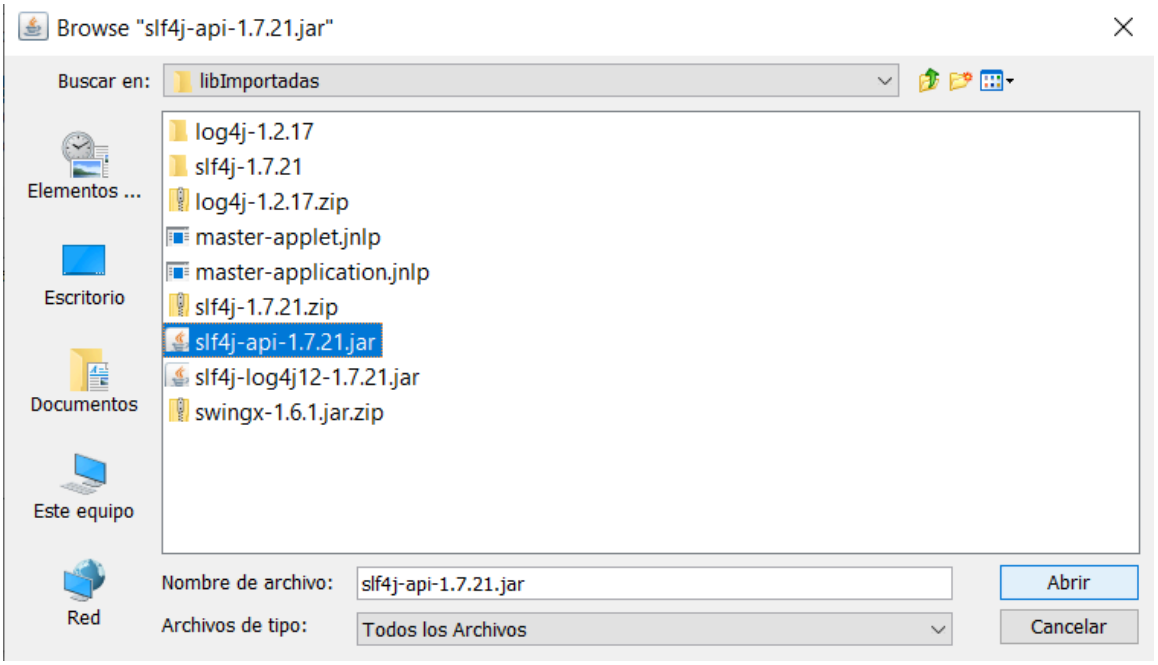
Nota: Las dependencias se encuentran en el archivo .zip de NetDrTool descargado previamente.

A continuación se muestra cómo se añadió cada una de ellas:

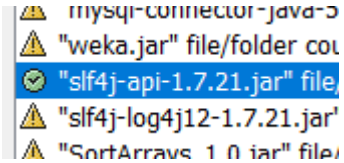
slf4j-api-1.7.21.jar y **slf4j-log4j12-1.7.21.jar**

Seleccionamos la opción correspondiente en la ventana de resolver problemas de netbeans y damos click en “resolver”.

Aparecerá una ventana que permite buscar y seleccionar el archivo correspondiente



Una vez se acepta el archivo seleccionado su icono de advertencia cambia a un ícono de aceptación, dando a entender que el archivo se ha añadido exitosamente.



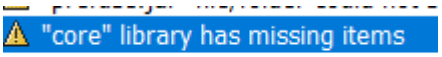
Se repite el mismo proceso con el resto de archivos.

Nota: Es preferible que todos los archivos de dependencias inicialmente se copien en la carpeta:

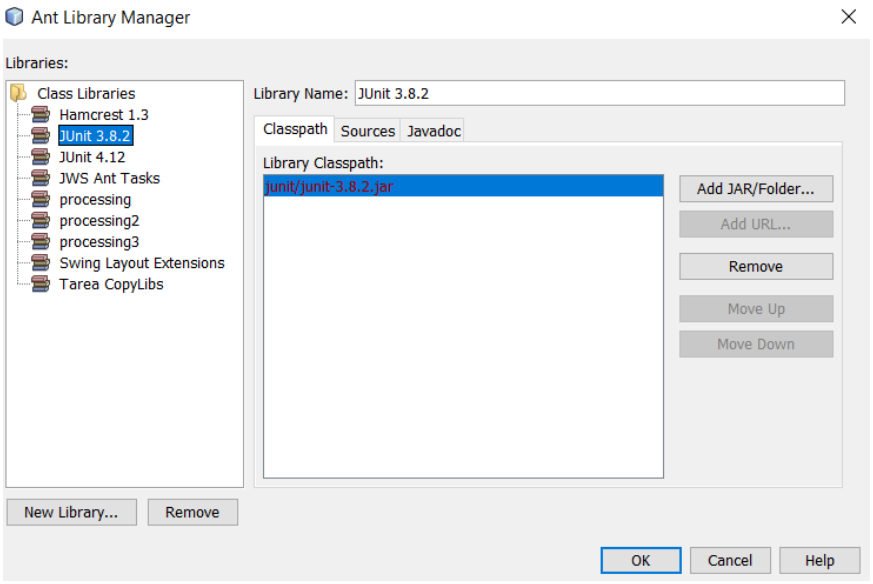
~\NetDRtool-main\JavaDRserver\libImportadas

Para tener un acceso sencillo a todas ellas a la hora de resolver conflictos.

En cuanto a los problemas encontrados en “core”



Al dar click en resolver nos lleva a la siguiente ventana de administración de librerías, por cada una de ellas se debe revisar las dependencias que faltan, en este caso se ubican los archivos .jar correspondientes y se copian en las carpetas solicitadas.



La librería de **junit-3.8.2.jar** se encuentra en la carpeta ‘core’ dentro del directorio de dependencias.

El archivo .jar se lleva a la siguiente carpeta:

~\NetDRtool-main\JavaDRserver\lib\junit

y el error se corregirá.

ahora para las dependencias **org-netbeans-modules-javawebstart-anttasks.jar** y **org-netbeans-modules-java-j2seproject-copylibstask.jar**

Dirigirse a la carpeta **~\NetDRtool-main\JavaDRserver\lib**

- Dentro de la carpeta lib se deben crear las siguientes carpetas:

    1. La primera con el nombre “**JWSAntTasks**” y dentro de esta se debe copiar el siguiente archivo:

       org-netbeans-modules-javawebstart-anttasks.jar

    2. Luego debe crearse una carpeta llamada “**CopyLibs**” y en ella copiar el archivo:

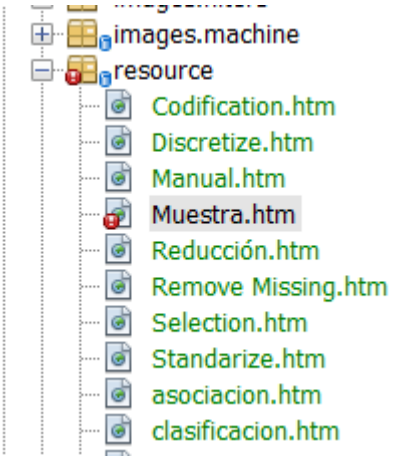
org-netbeans-modules-java-j2seproject-copylibstask.jar

Dirigirse a la carpeta **~\NetDRtool-main\JavaDRserver\libImportadas**

- Dentro de '**libImportadas**' se deben copiar las siguientes carpetas que se encuentran en 'core':

    1. processing
    2. libraryProc2
    3. libraryProc3

Una vez hayamos añadido todas las dependencias necesarias procedemos a eliminar el archivo "Muestra.htm" que genera conflicto:



Ubicación: **~\NetDRtool-main\JavaDRserver\src\resource\Muestra.htm**

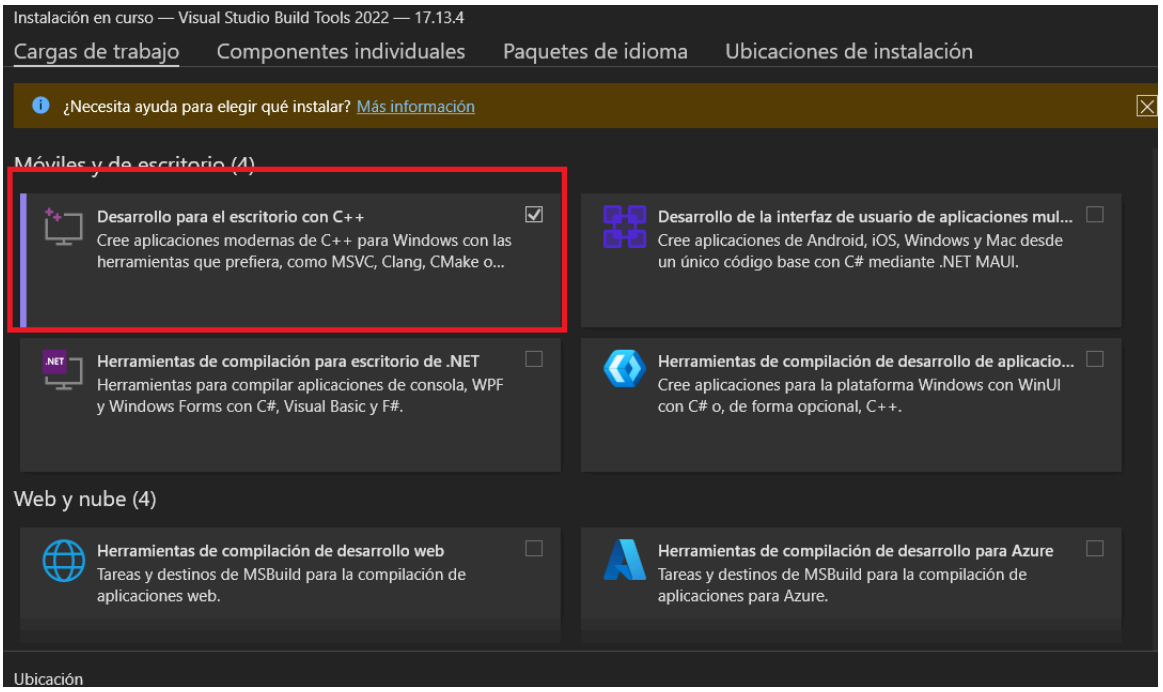
Una vez eliminado, el servidor en Java estará listo para ser ejecutado.

3. Instalación de Microsoft C++ Build Tools

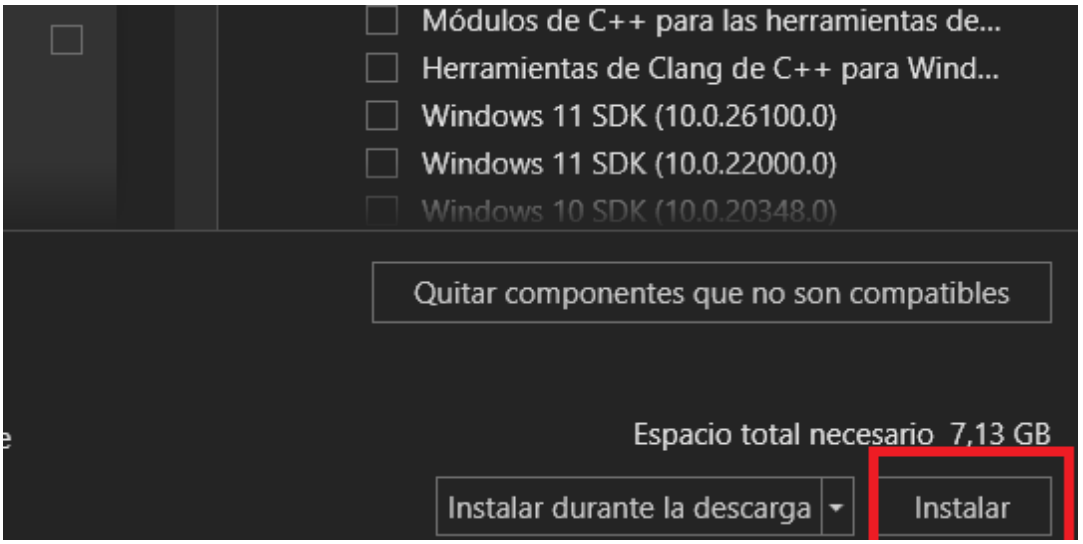
Antes de instalar las dependencias de python es necesario instalar Microsoft C++ Build Tools.

Enlace de descarga: https://visualstudio.microsoft.com/visual-cpp-build-tools/

Una vez instalado el asistente se abrirá la siguiente ventana:



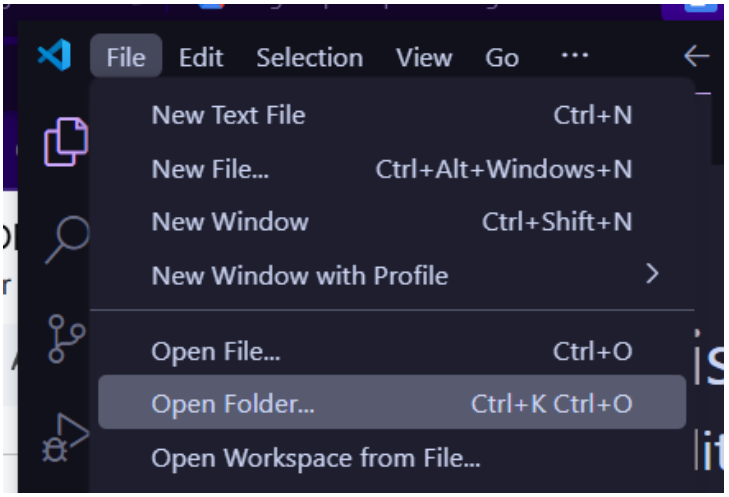
Se debe seleccionar la opción de desarrollo para el escritorio con C++ y darle click a instalar.



Una vez se haya instalado, debe reiniciarse la máquina.

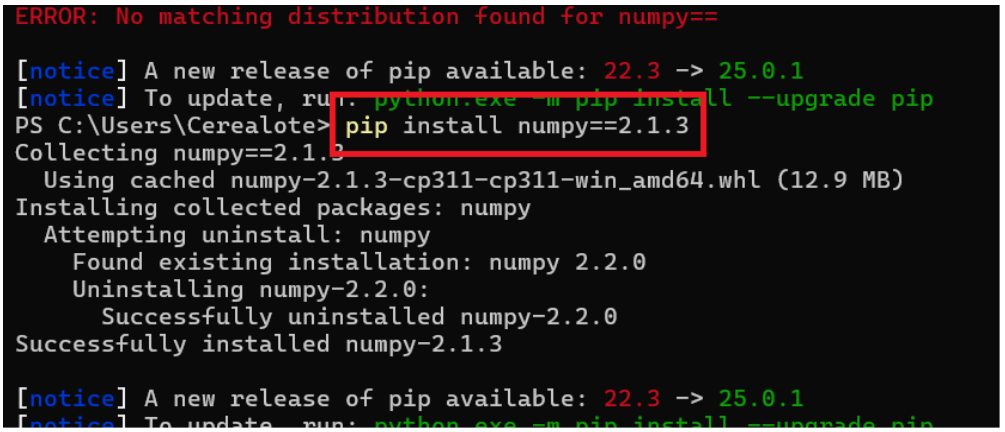
4. Instalación de dependencias para PythonDRClient

Se abre el proyecto PythonDRClient en Visual Studio Code.

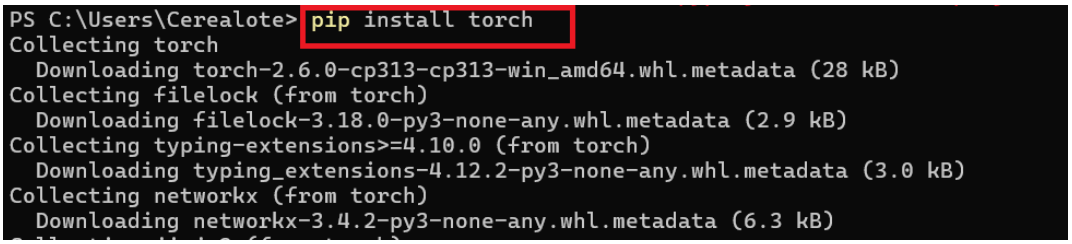


Se abre la consola de windows (cmd, powershell o terminal) y se instalan las siguientes dependencias por medio de los comandos:

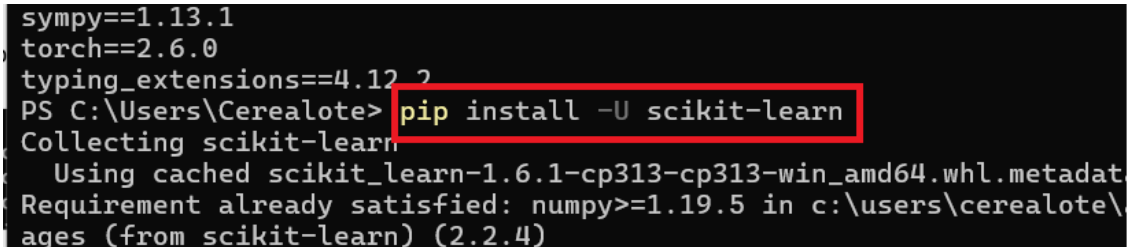
Dependencia: **numpy (Versión menor a 2.2.0)**



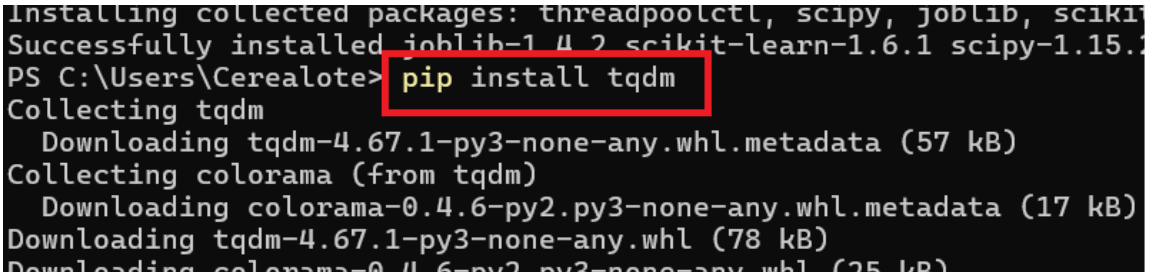
Dependencia: **torch**



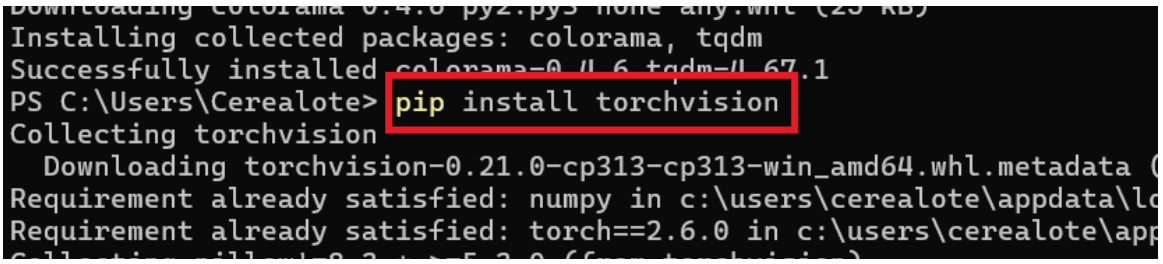
Dependencia: **Scikit-learn**



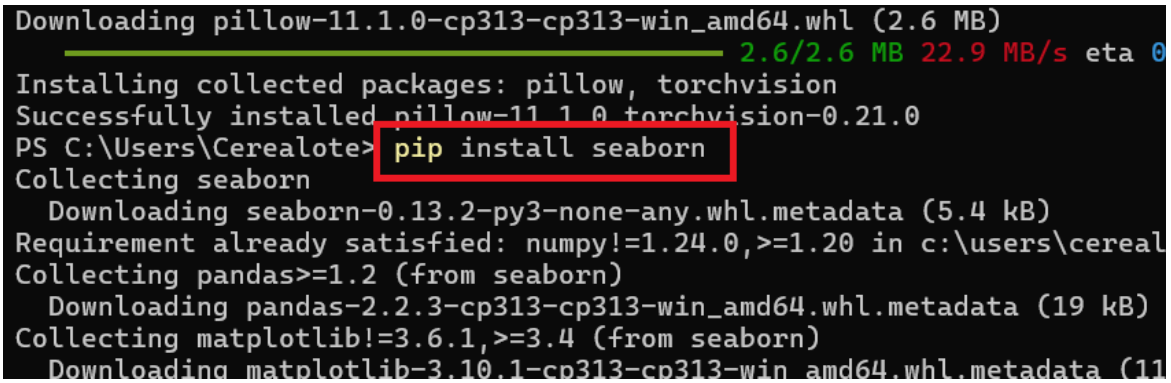
Dependencia: **tqdm**



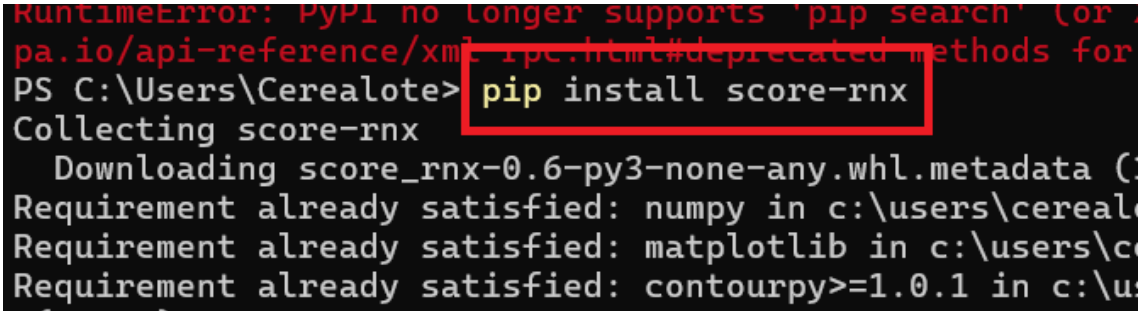
Dependencia: **torchvision**



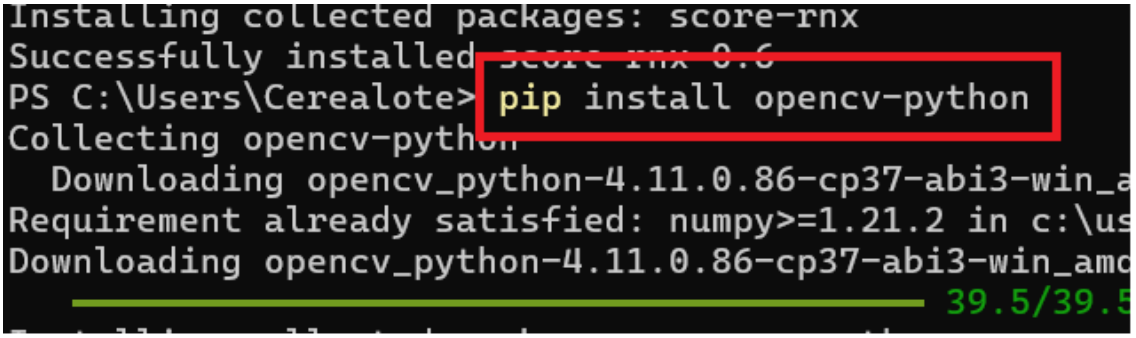
Dependencia: **seaborn**



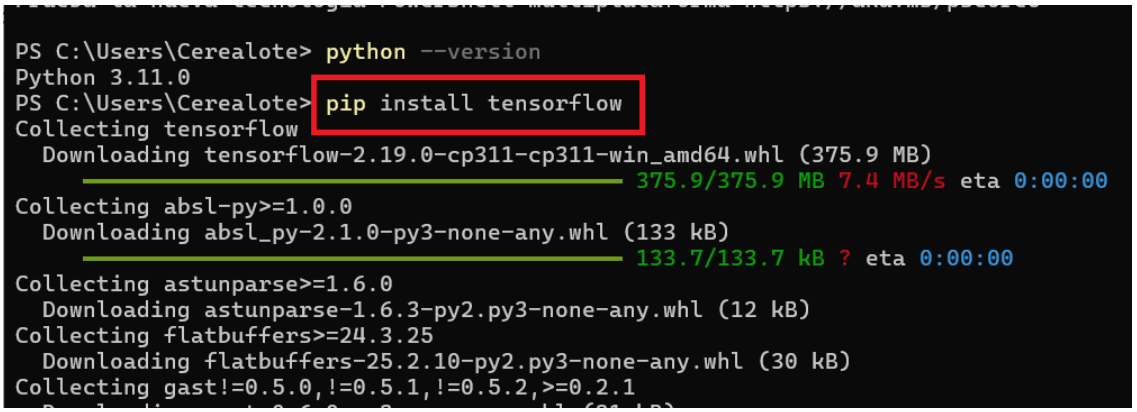
Dependencia: **score_rnx**



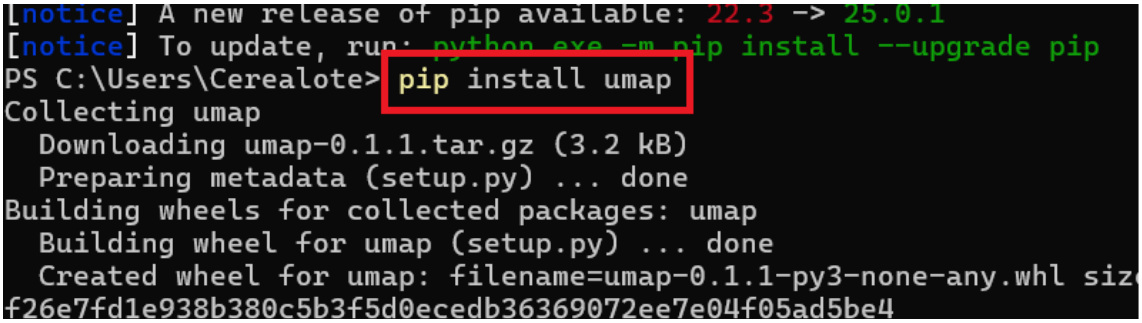
Dependencia: **cv2 (opencv)**



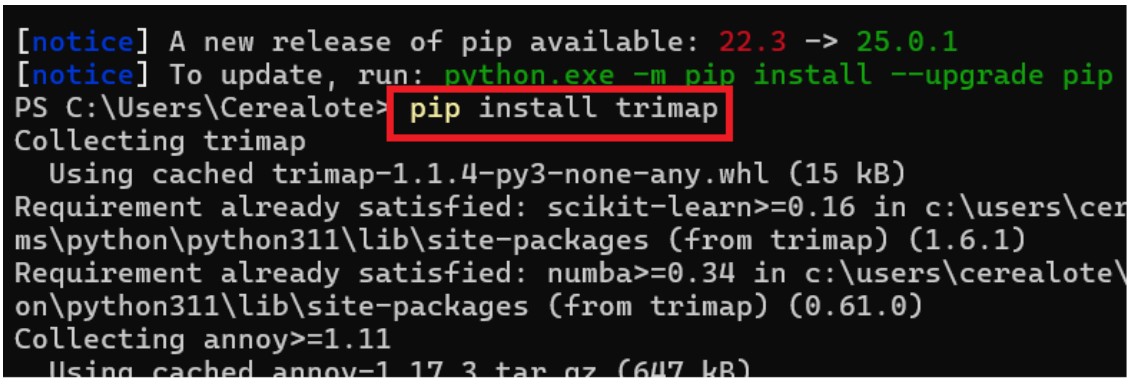
Dependencia: **tensorflow**



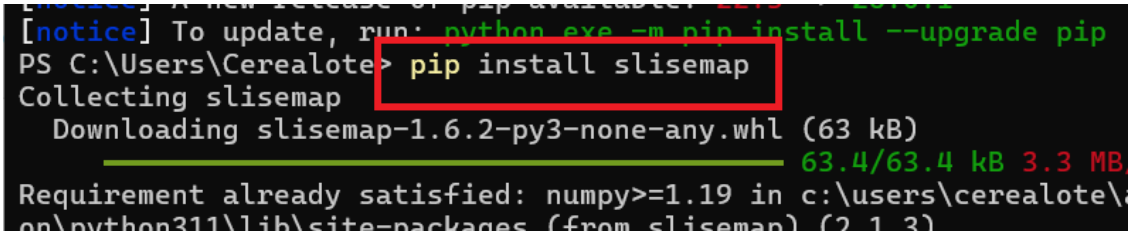
Dependencia: **umap**



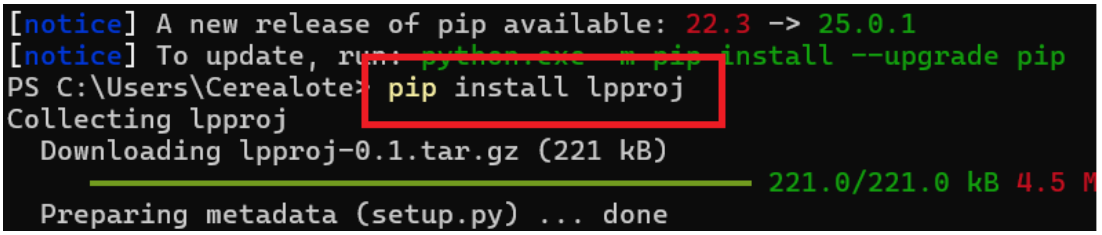
Dependencia: **trimap**



Dependencia: **slisemap**



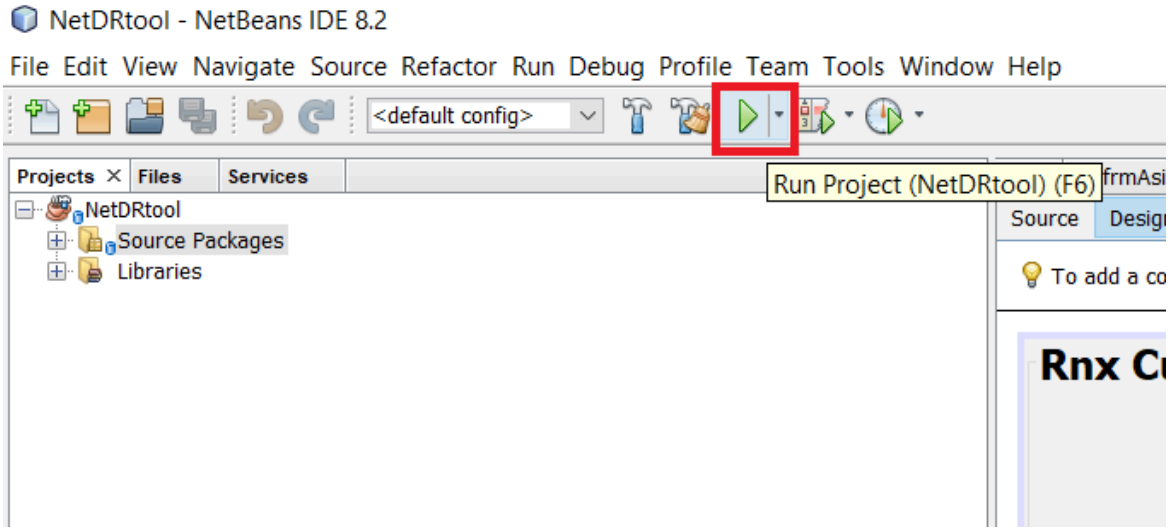
Dependencia: **lpproj**



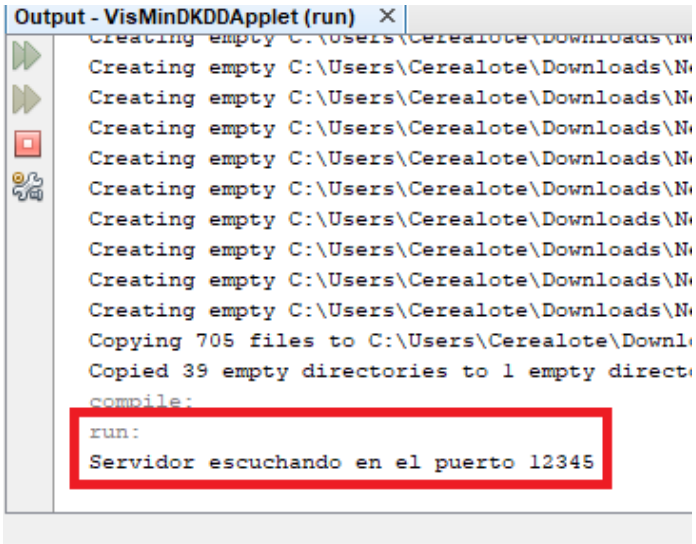
Con ello el cliente en python estará listo para ser ejecutado.

INSTRUCCIONES PARA LA EJECUCIÓN DE NETDRTOOL

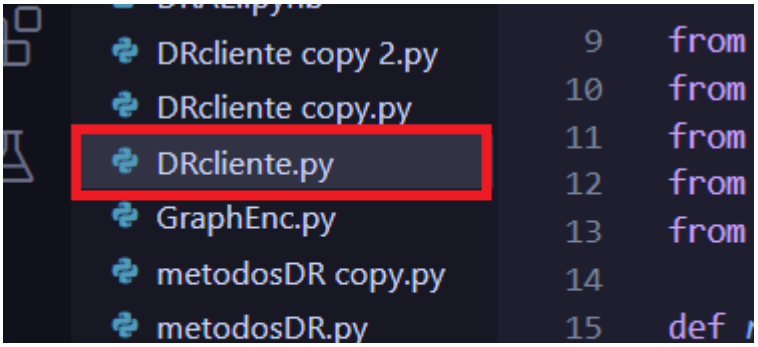
El primer paso es iniciar el servidor en Java, dando click al icono especificado en la imagen.



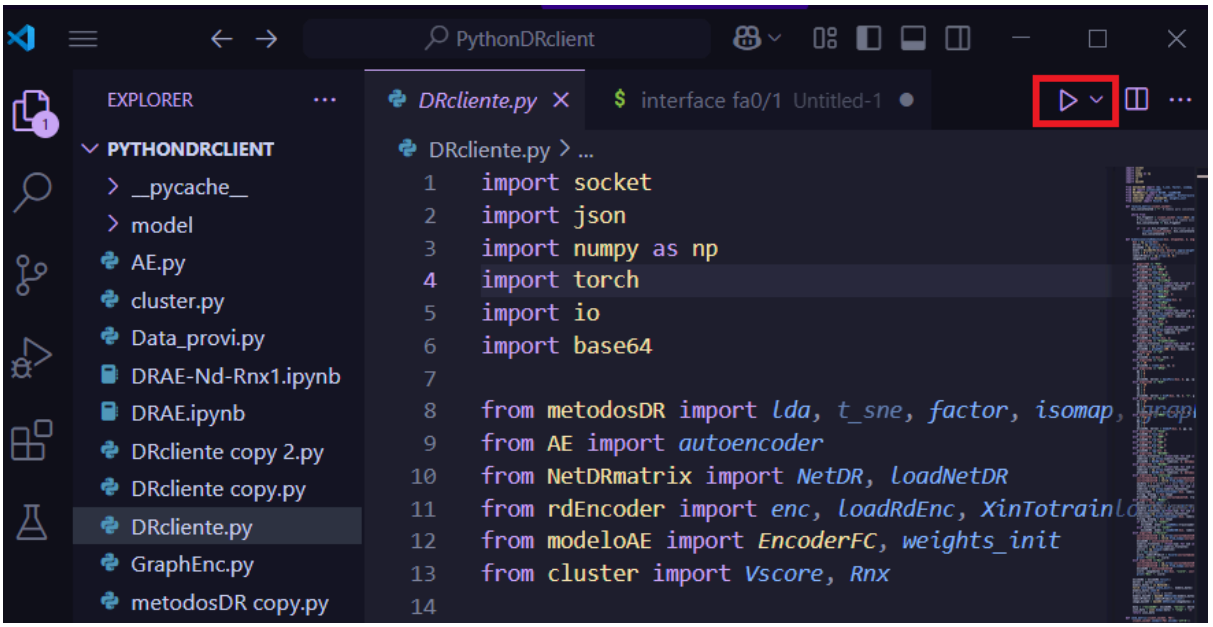
En la consola de netbeans se menciona que el programa queda escuchando en el puerto 12345.



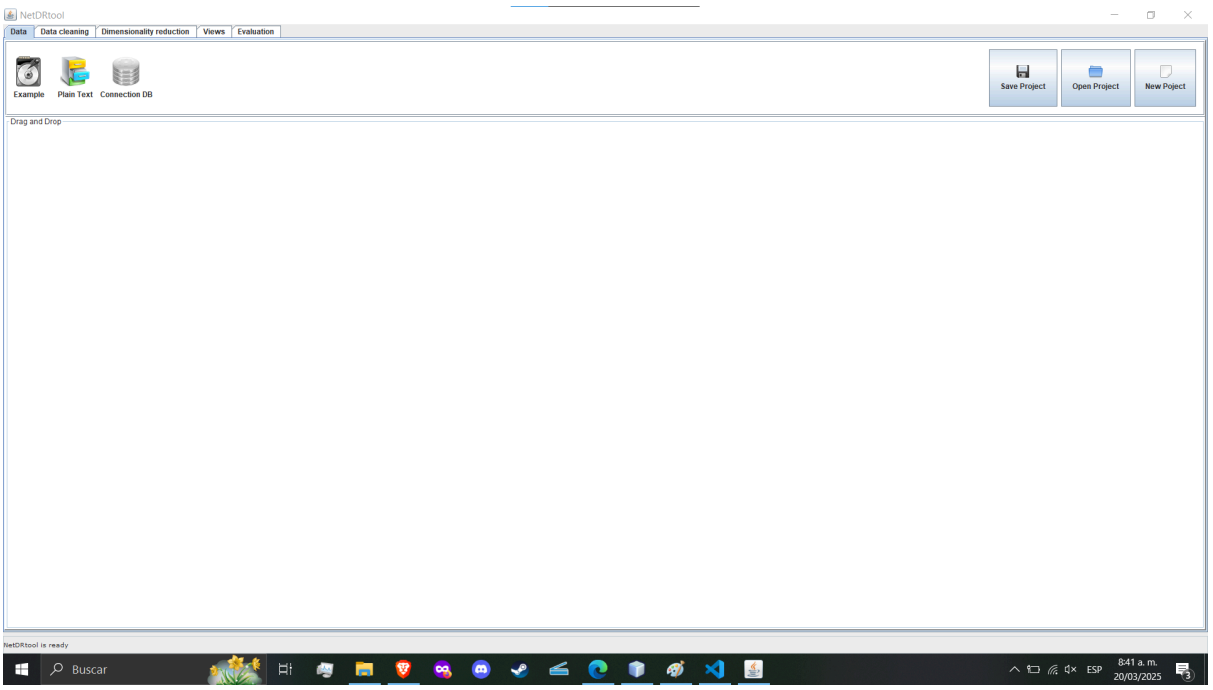
Y en vscode debe entrar al archivo DRCliente.py



Y darle al icono especificado en la imagen para ejecutar el archivo.

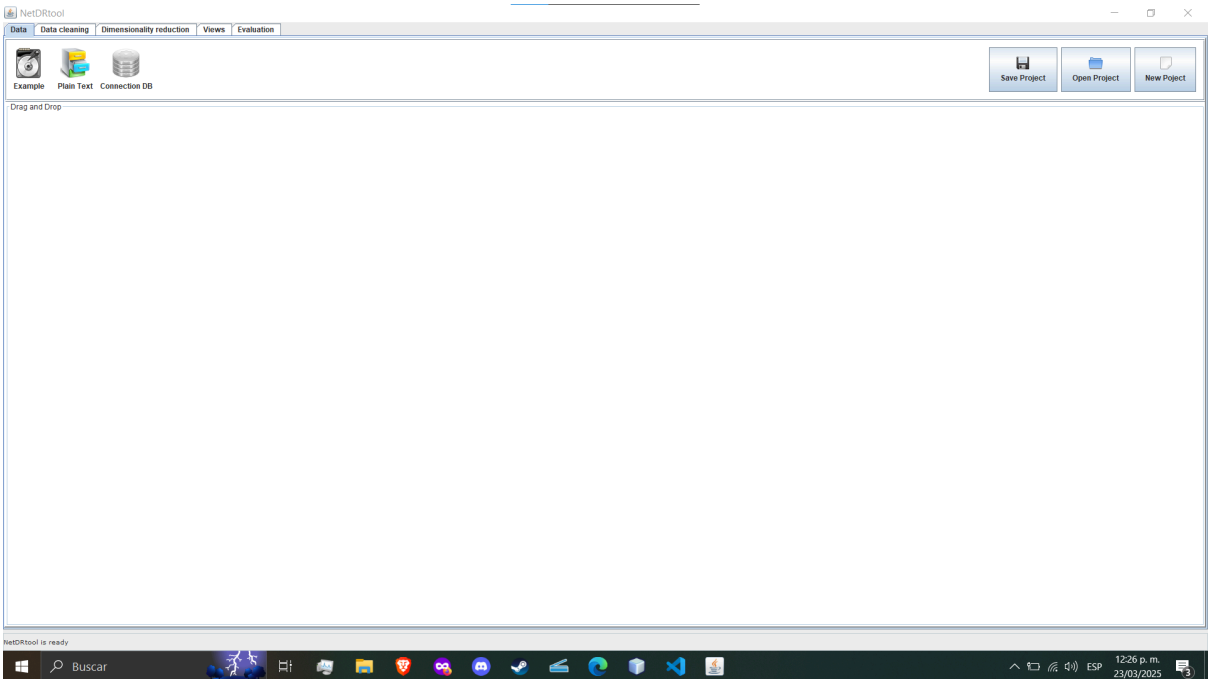


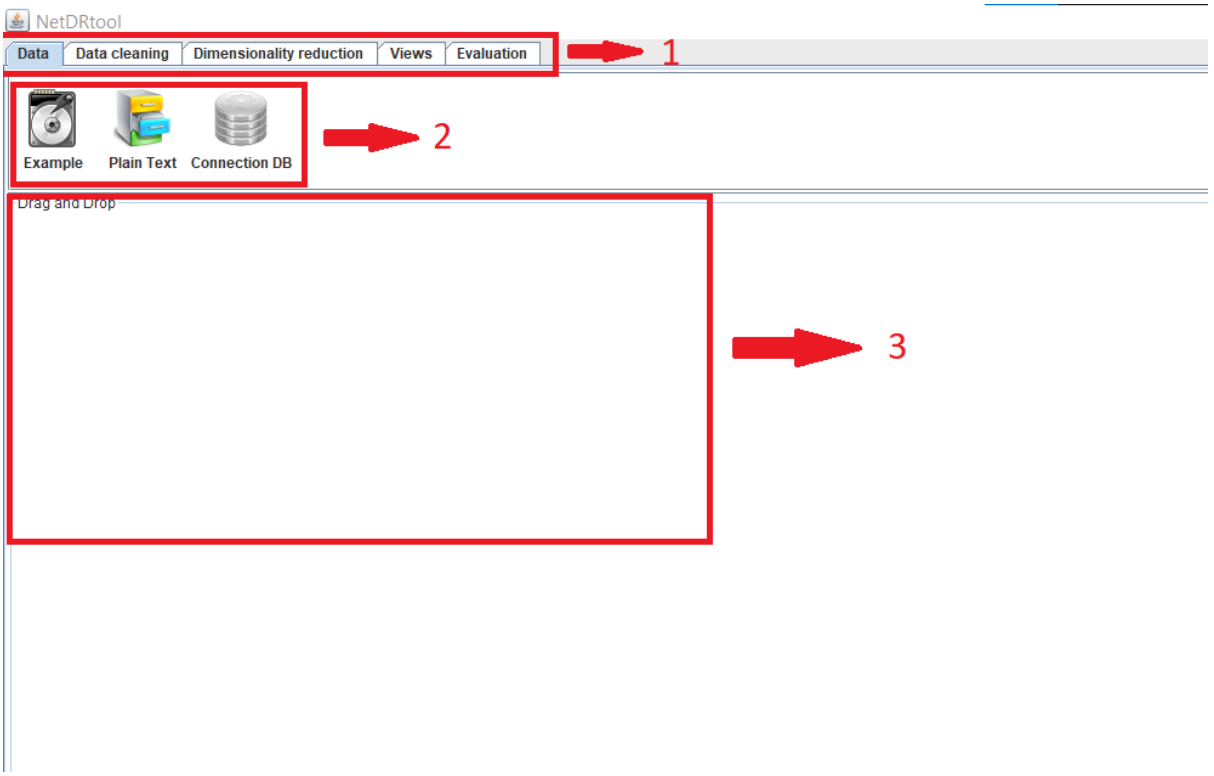
Una vez el cliente termine de cargar, se abrirá una ventana de Java y la herramienta estará lista para usarse.



MANUAL DE USUARIO DE NETDRTOOL

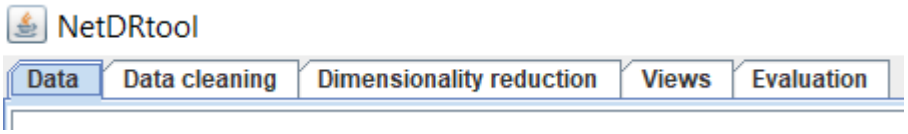
1. Ventana principal del sistema





En la ventana principal del programa se destacan tres secciones, las cuales se explican a continuación:

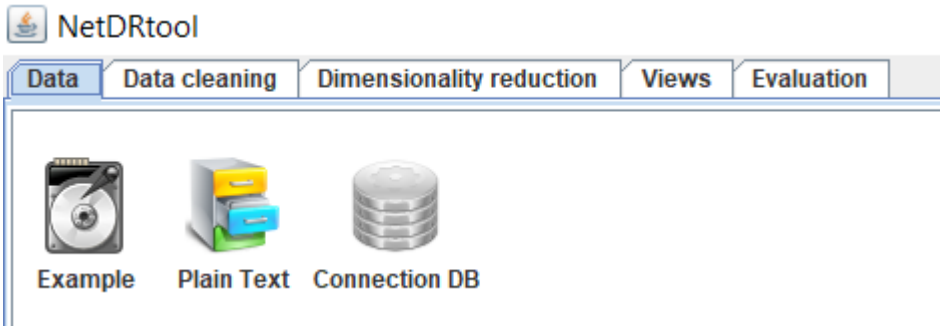
1. **Barra de navegación**



Contiene cinco apartados:

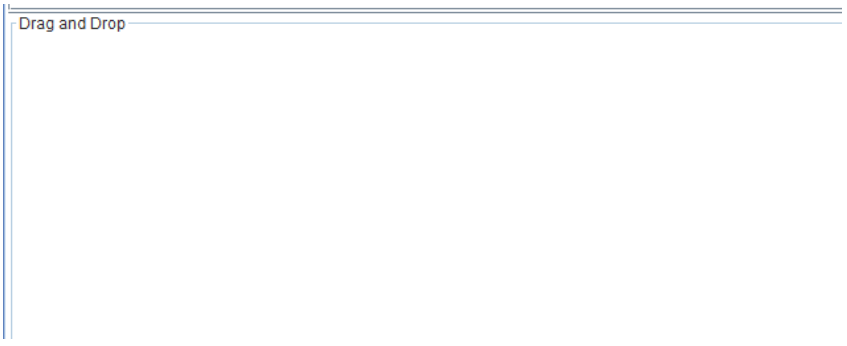
- **Data:** Referente a la conexión con los datos a trabajar.
- **Data cleaning:** Preprocesamiento, limpieza de datos.
- **Dimensionality reduction:** Contiene las técnicas de reducción de dimensión para bases de datos.
- **Views:** Visualizadores de los métodos de reducción de dimensión.
- **Evaluation:** Herramientas que se encargan de evaluar qué tan eficiente es un algoritmo o modelo utilizado.

2. **Opciones del menú**

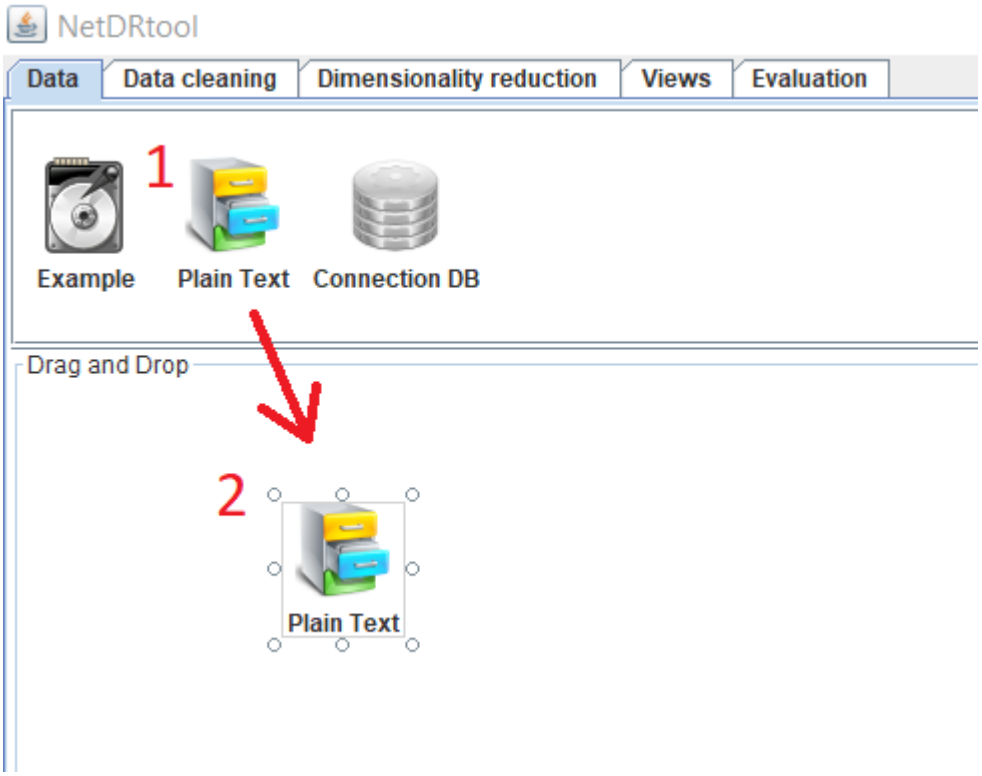


Cada apartado de la barra de navegación contiene opciones a escoger, dependiendo de lo que el usuario quiera hacer, estas opciones se explicarán a detalle más adelante

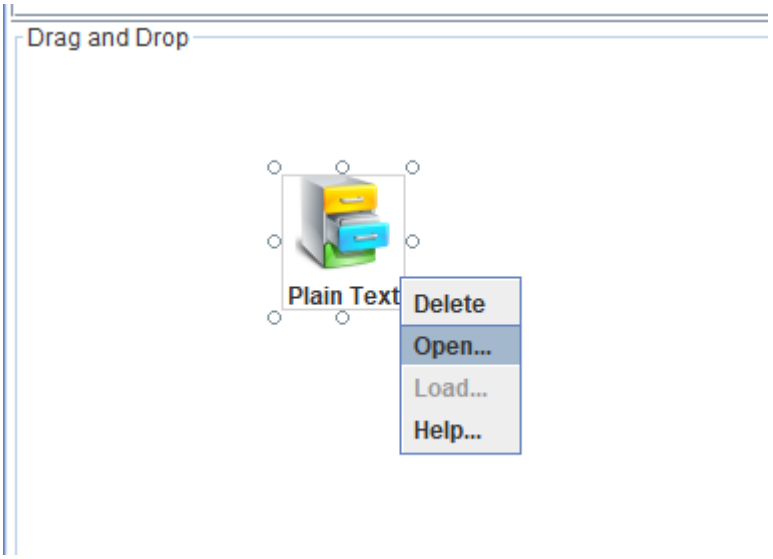
3. **Zona "Drag and Drop"**



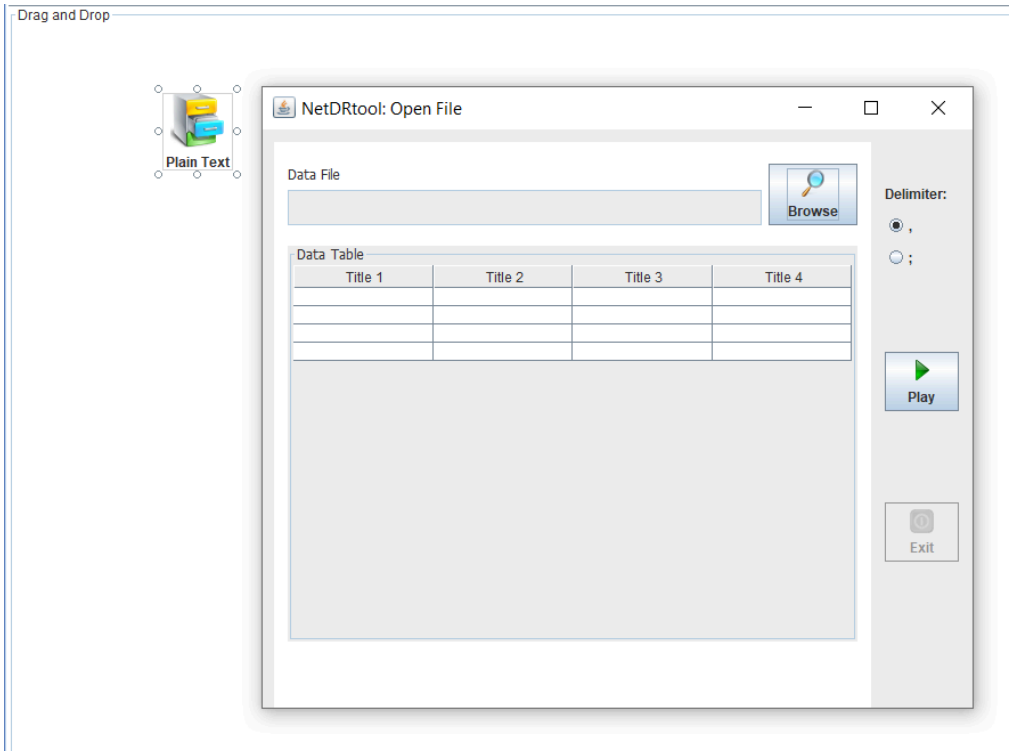
Esta sección del programa cubre la mayor parte de la pantalla y está diseñada para que el usuario pueda ubicar los componentes que quiere utilizar mediante la acción de arrastrar y soltar los mismos para crear el mapa de conocimiento acerca del proceso a evaluar.



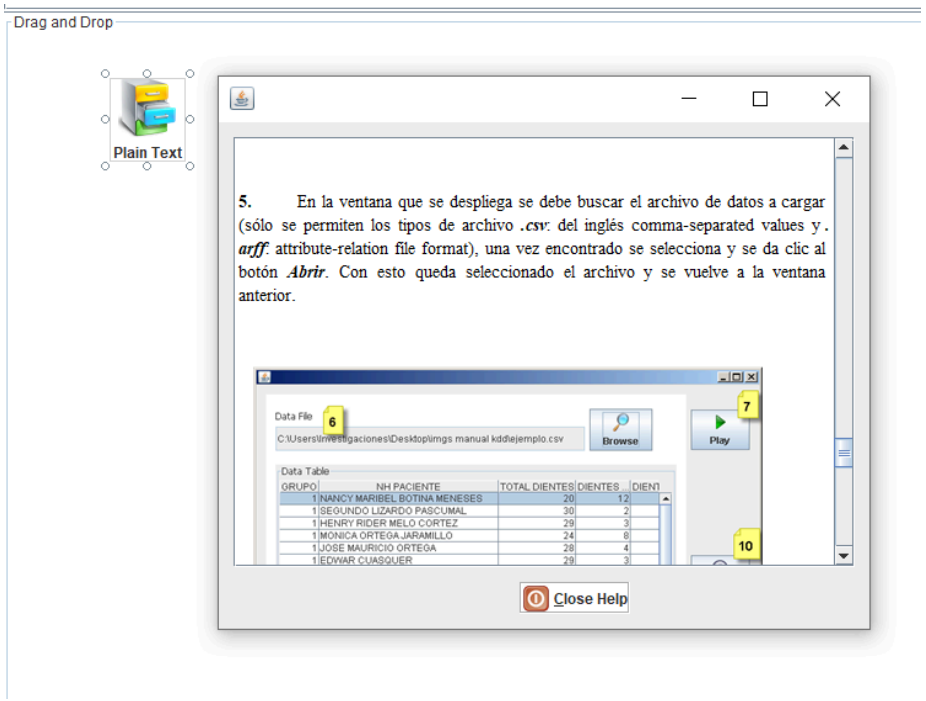
Al arrastrar un componente al área de trabajo este contará con ocho ejes o puntos de los cuales se puede conectar otros componentes, además al dar click derecho en cualquier componente dentro del área de trabajo, este desplegará un menú con las opciones de dicho componente:



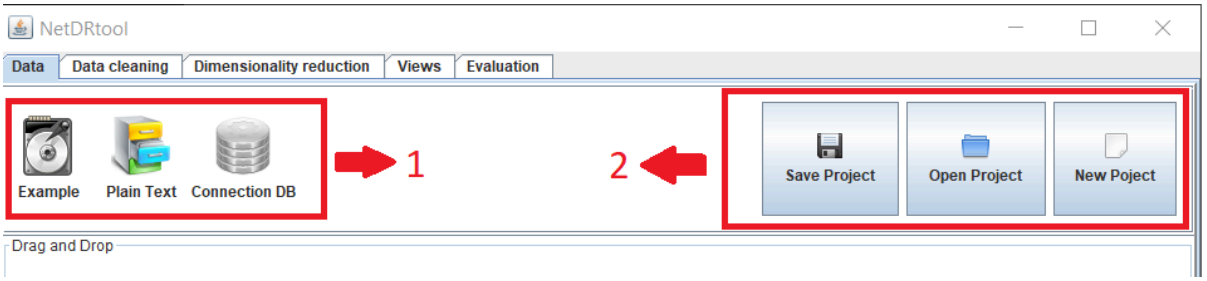
Dependiendo de la opción que se escoja, se realizará la acción deseada, por ejemplo al seleccionar “Delete”, el componente será eliminado, pero al seleccionar “Open” o “Help” emergerá una nueva ventana.



Nota: Al seleccionar la opción “Help” de cualquier componente el programa abrirá una ventana con una explicación detallada acerca de cómo utilizar y cómo funciona dicho componente.

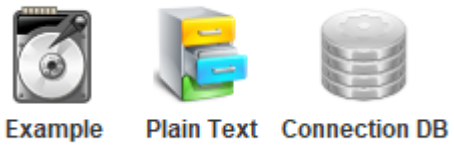


2. Data



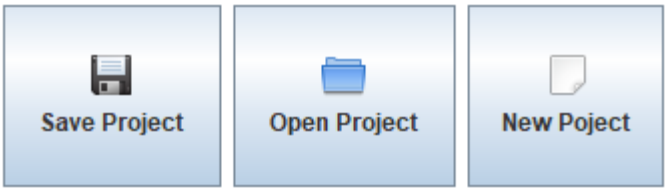
El apartado dedicado a la conexión con los datos a trabajar contiene dos secciones que se destacan:

- La primera de ellas se encuentra a la izquierda de la pantalla y contiene las opciones de conexión de datos que pueden hacer uso de las herramientas del programa.



  a. **Example:** Cuenta con conjuntos de datos previamente cargados.
  b. **Plain Text:** Permite abrir archivos de texto plano (.csv).
  c. **Connection DB:** Permite conectarse a bases de datos postgresql, mysql u oracle.

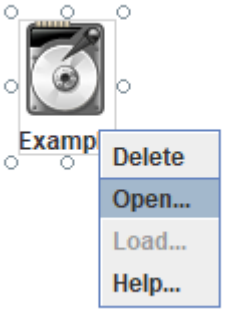
- La segunda se encuentra a la derecha y contiene las opciones del proyecto.



  a. **Save project:** Guardar el proyecto actual.
  b. **Open project:** Abrir un proyecto.
  c. **New project:** Comenzar un nuevo proyecto.

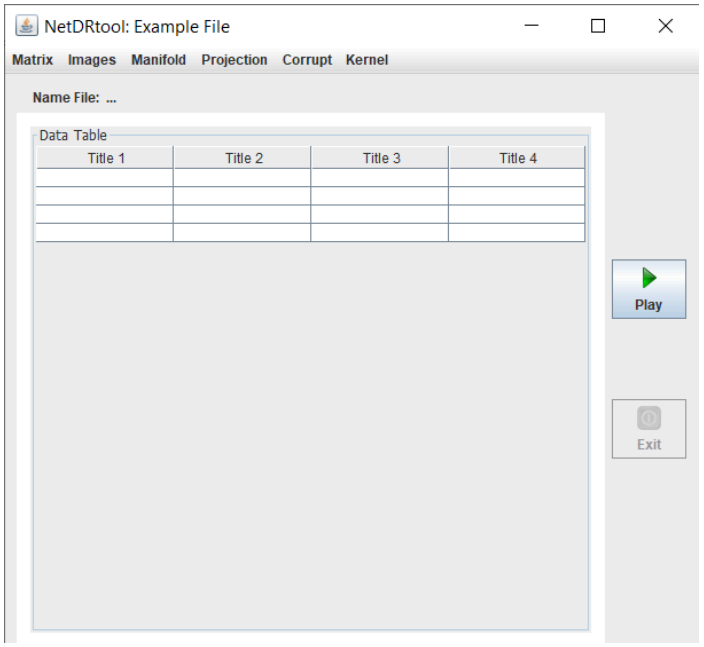
2.1. Conexión a los datos

Se accede a su ventana principal de cada componente dando click derecho en el mismo y luego en la opción “Open”.

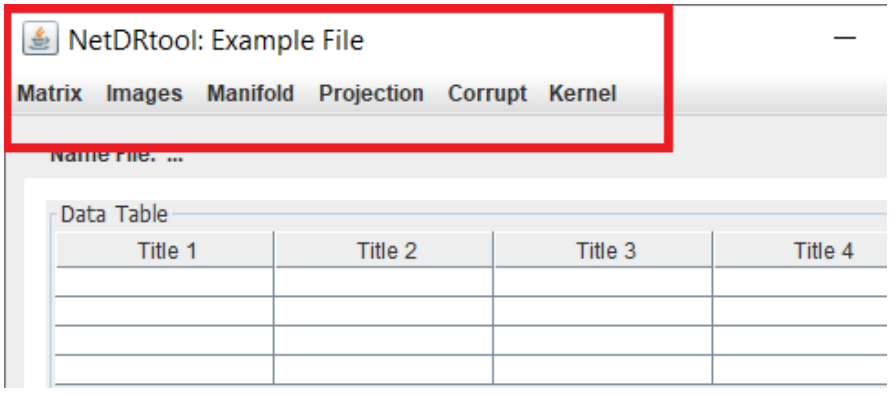


2.1.1. Example

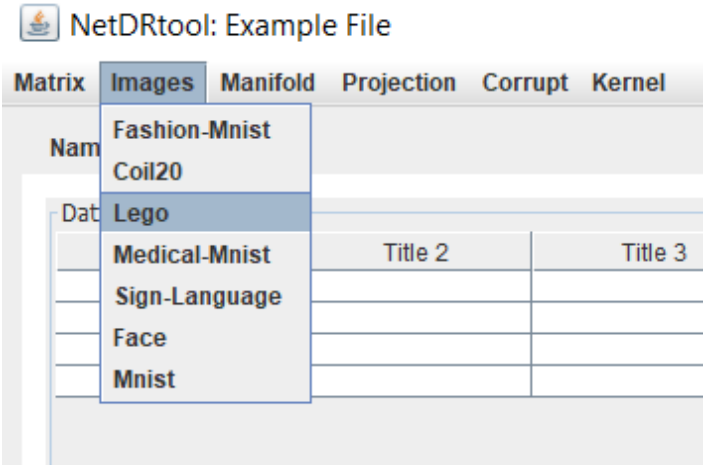
Su ventana principal es la siguiente:

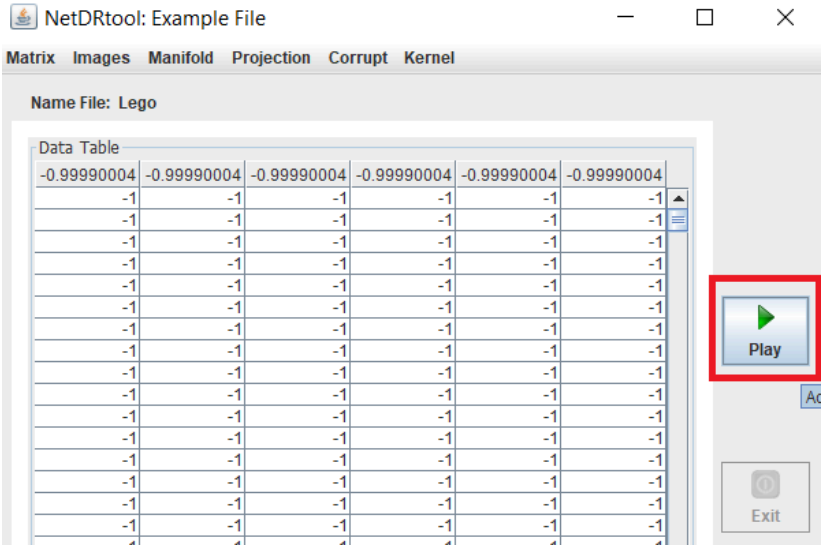


Cuenta con diversos datos de ejemplo de los cuales se puede escoger para realizar pruebas.



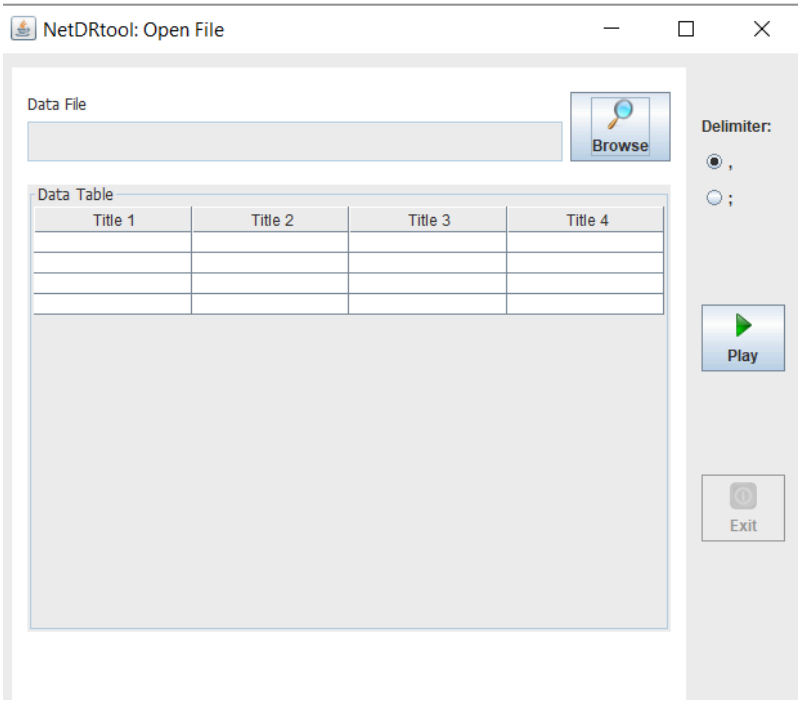
Simplemente debe seleccionar una de las opciones disponibles y cargar los datos mediante el botón “Play”.



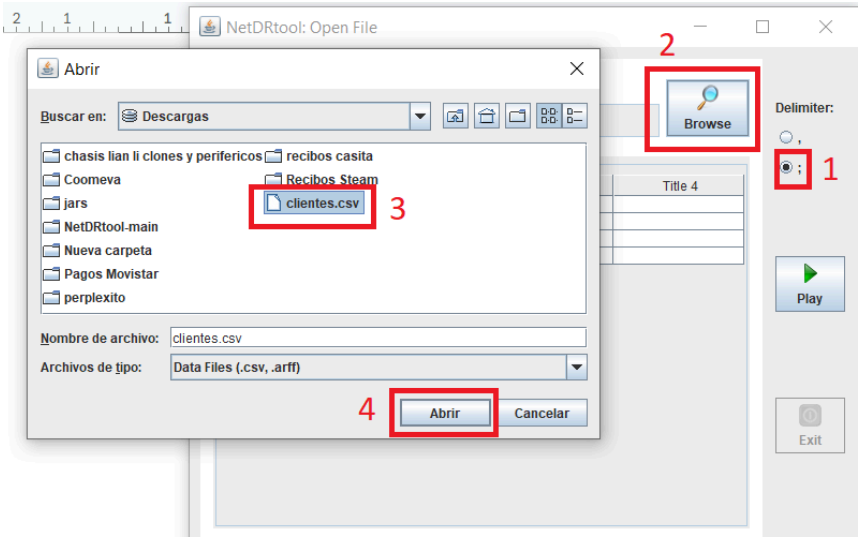


2.1.2. Plain Text

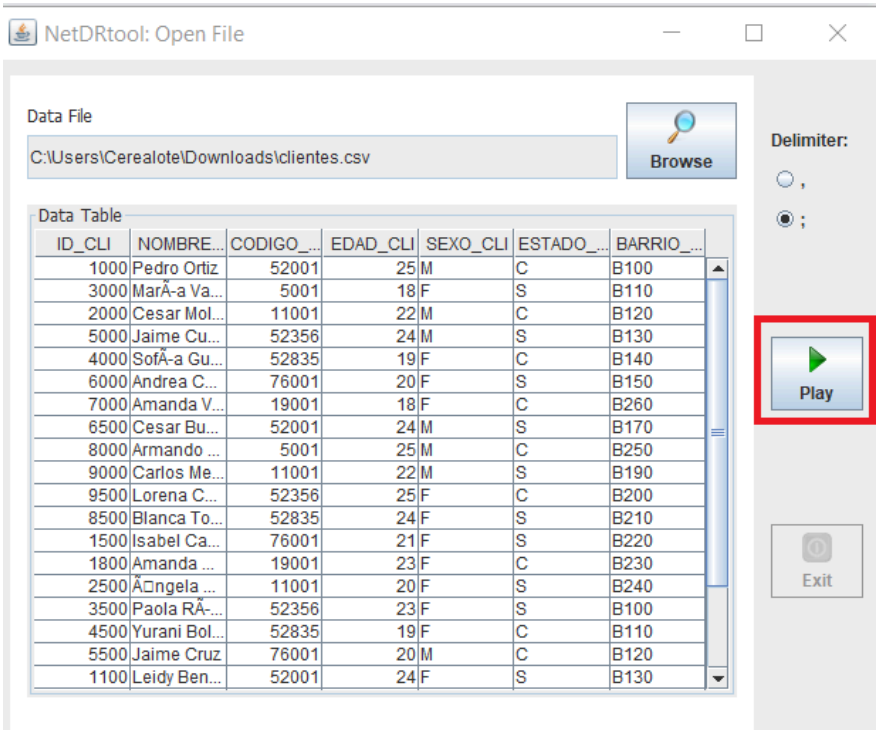
Su ventana principal es la siguiente:



A diferencia de la anterior, mediante la opción “Browse” se puede importar el archivo .csv deseado, especificando si su delimitador es “,” o “;”.

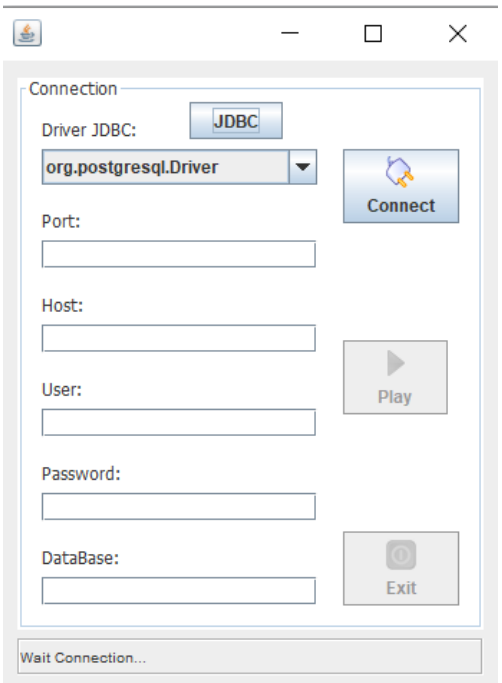


Con ello se abre el archivo y se selecciona la opción “Play” para cargarlo.

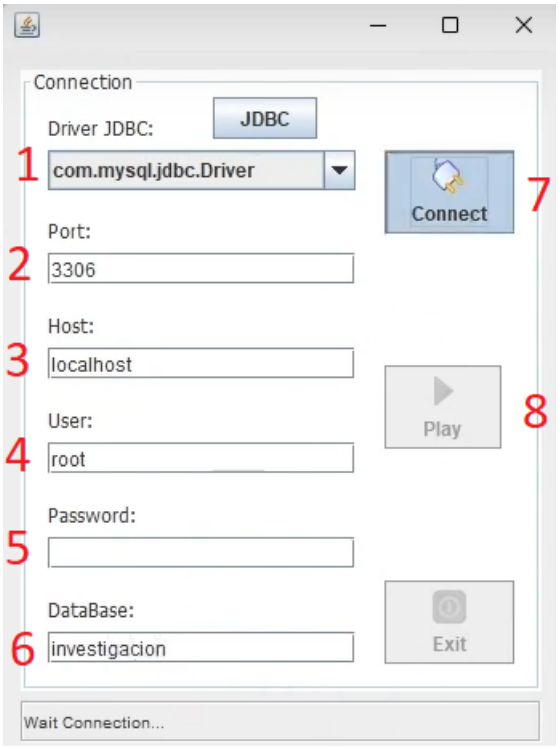


2.1.3. Connection DB

Su ventana principal es la siguiente:



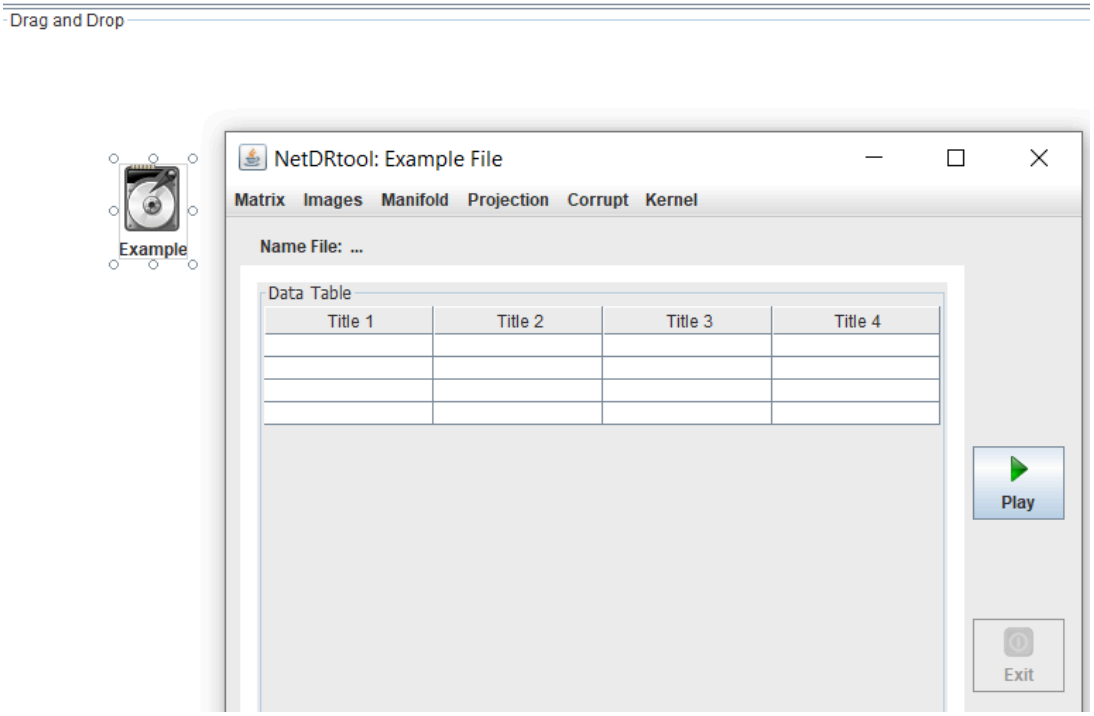
En ella puede encontrar los siguientes apartados



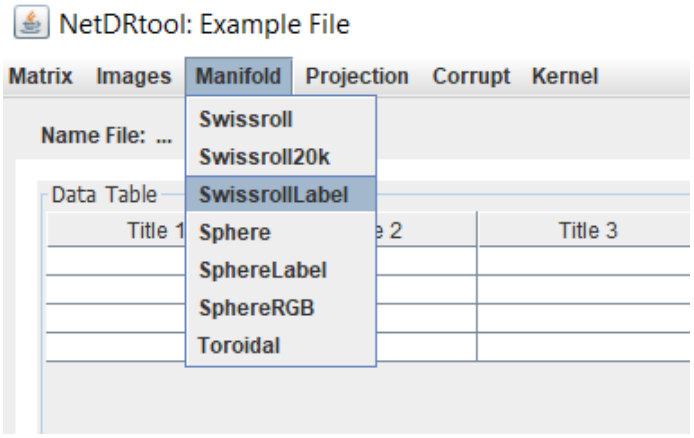
1. Tipo de base de datos: puede ser postgresql, mysql u oracle.
2. Puerto que utiliza: se carga por defecto y puede editarlo si es necesario.
3. Host.
4. Usuario.
5. Contraseña.
6. Nombre de la base de datos.
7. Botón de conexión: Debe darle click una vez ingrese la información anterior.
8. Botón de carga: Debe darle click una vez se haya establecido la conexión con la base de datos.

2.2. Ejemplo de uso

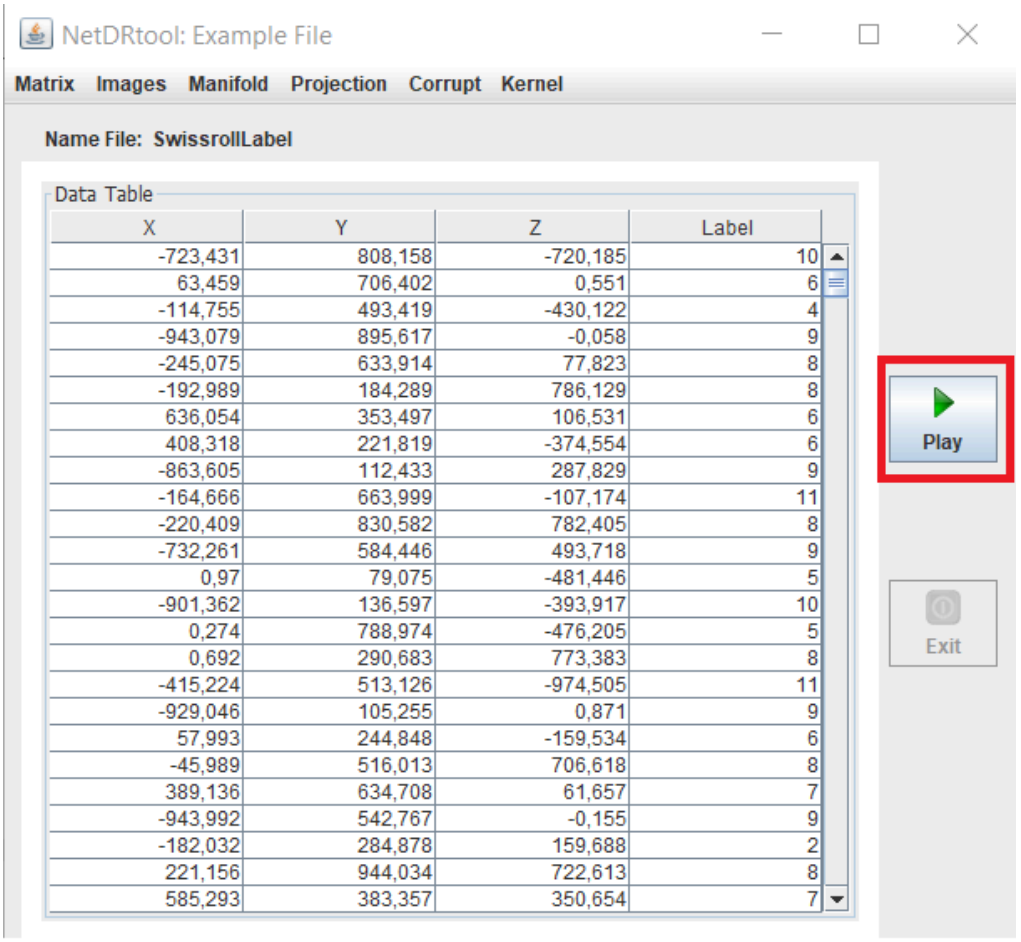
Una vez se ha arrastrado la opción de conexión de datos al área de trabajo, debe seleccionarse los datos con los cuales se va a trabajar, en este caso se utilizan los datos de ejemplo que ofrece la opción “Example”.



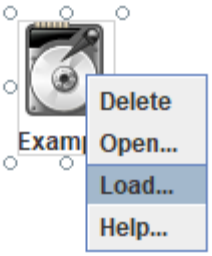
Se escoge una de las opciones disponibles



Se cargarán los datos en la tabla y entonces se debe dar click al botón “Play”



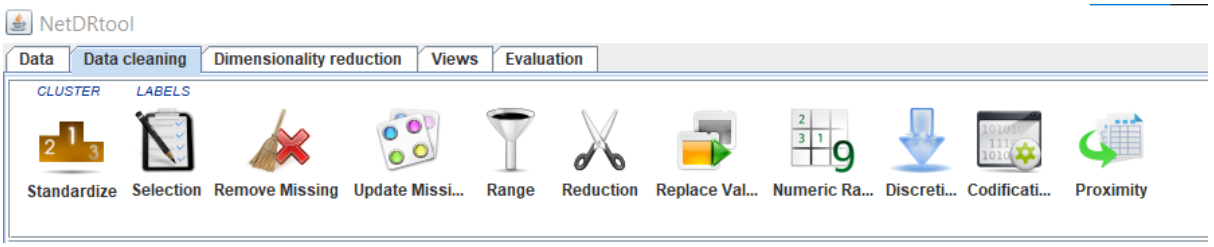
Y posteriormente se deben cargar dichos datos, seleccionando la opción “Load” en el componente de ejemplo.



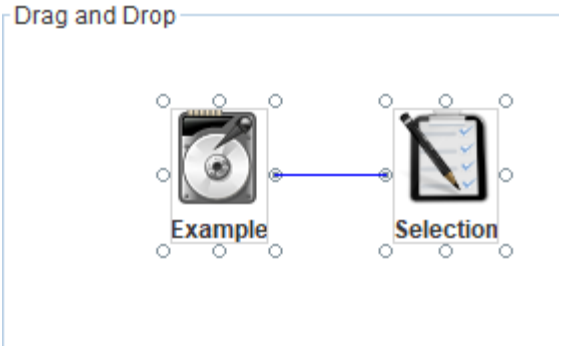
Con ello los datos estarán listos para ser trabajados.

3. Data cleaning

El apartado Data cleaning contiene diferentes filtros para limpiar los datos a trabajar.



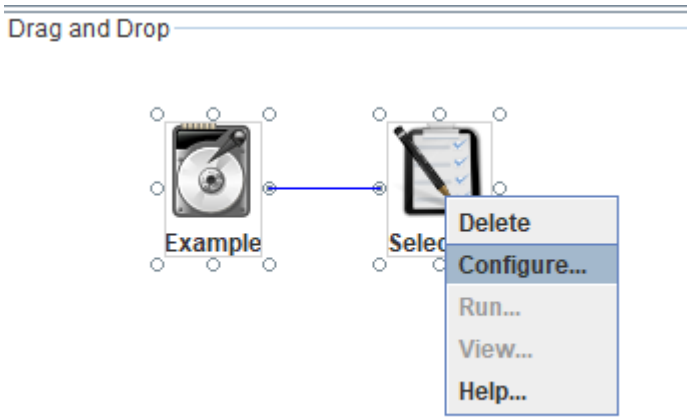
Nota: Para realizar la conexión entre los datos y la herramienta de filtro, los datos con los que va a trabajar deben haber sido cargados con anterioridad, de lo contrario no podrá realizarse dicha conexión.



Esto se logra dando click a uno de los ejes circulares del componente inicial para posteriormente dar click al componente objetivo con el cual busca conectarse.

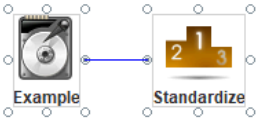
3.1. Filtros

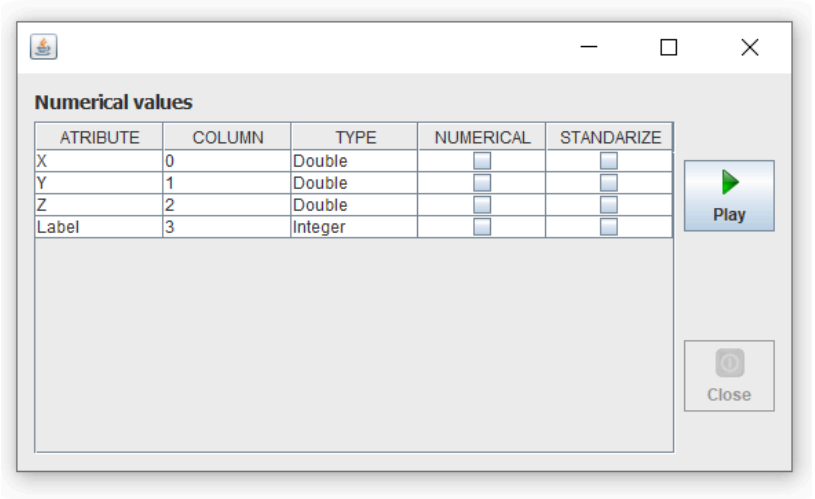
A continuación se presenta cada filtro con su ventana principal, a la cual puede accederse dando click derecho al componente y luego en la opción “Configure”.



3.1.1. Standardize

Permite estandarizar los datos, trabaja solo datos de tipo numérico (integer o double), para evitar valores extremos o atípicos.

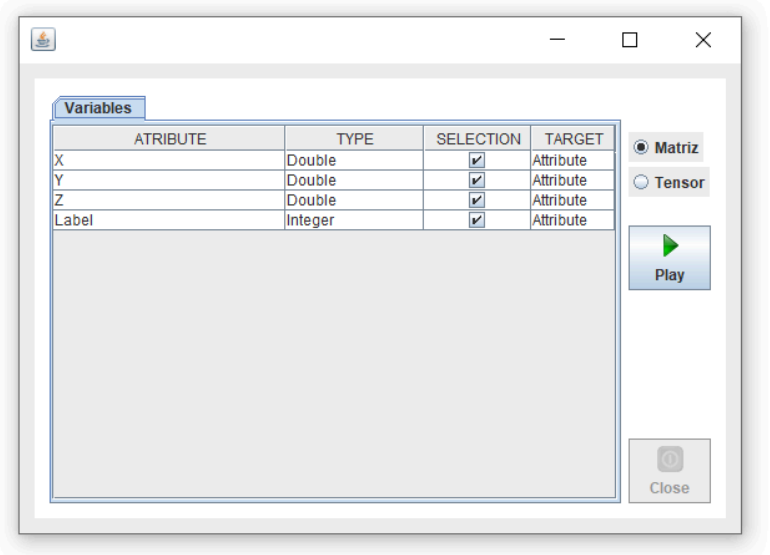




3.1.2. Selection

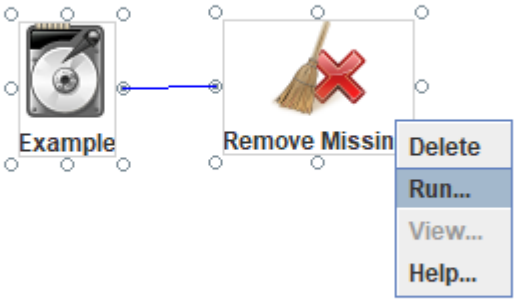
Ayuda a centrarse en los atributos relevantes para incluir en el modelo, escogiendo la etiqueta objetivo.





3.1.3. Remove Missing

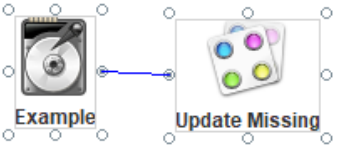
Remueve los datos nulos.

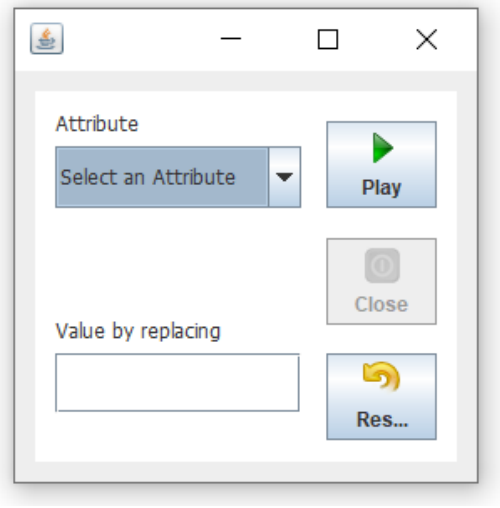


A diferencia de los anteriores, no posee una ventana de configuración, simplemente al darle click a “Run” se encargará de eliminar datos nulos.

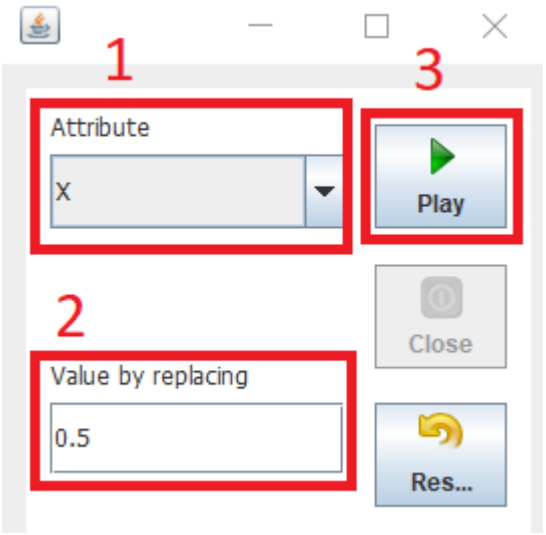
3.1.4. Update Missing

Actualiza los datos nulos con otros valores.





En este caso puede seleccionarse el atributo del cual quieren buscarse valores nulos y reemplazarlos con el valor que se desee.

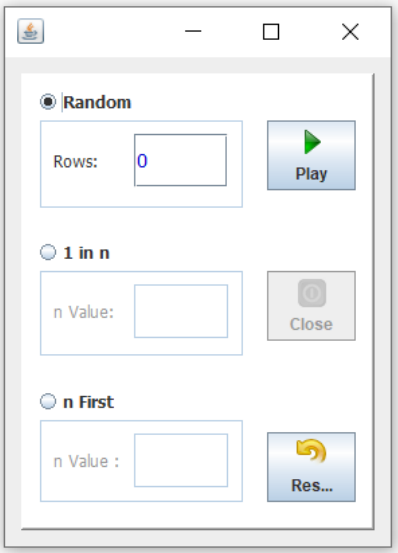


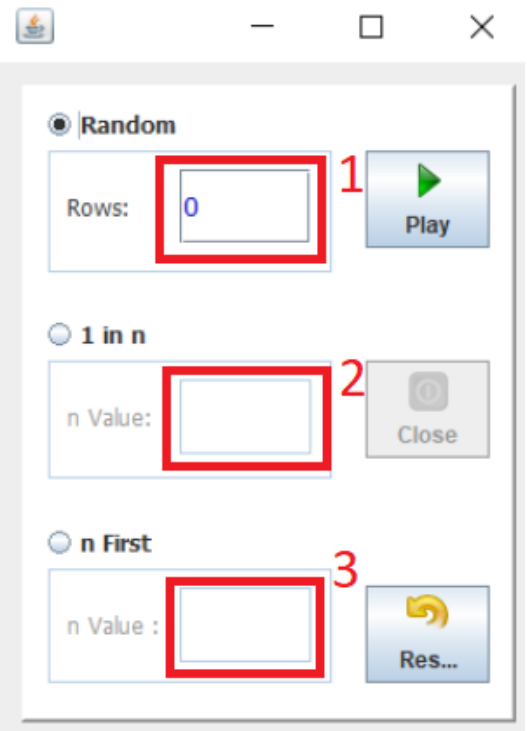
1. Atributo a seleccionar
2. Valor con el cual se va a reemplazar
3. Cargar filtro

3.1.5. Range

Permite una selección de rango.



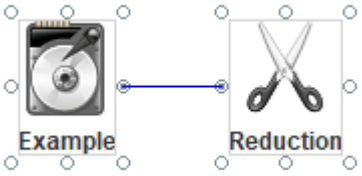




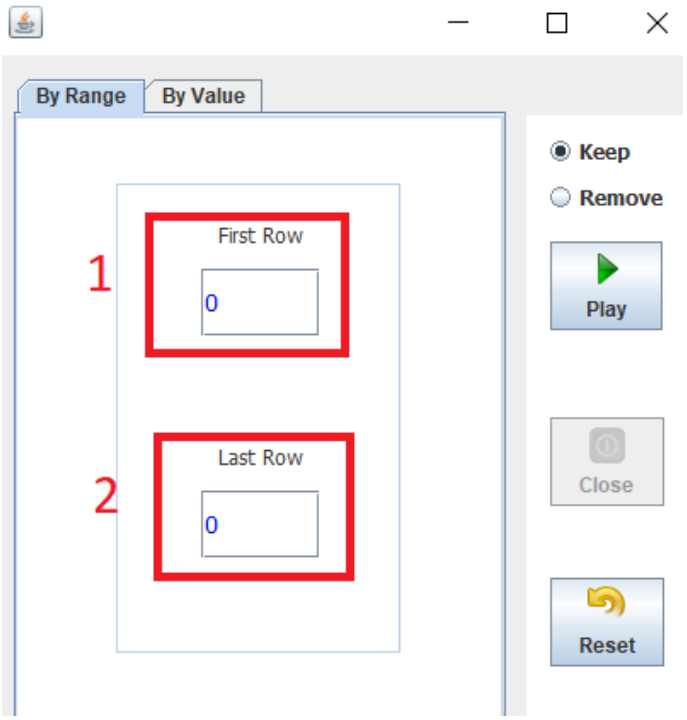
1. Escoge un determinado número de filas en forma aleatoria
2. Escoge desde la fila n
3. Las primeras n filas

3.1.6. Reduction

Permite eliminar un conjunto de datos, por rango y por valor.

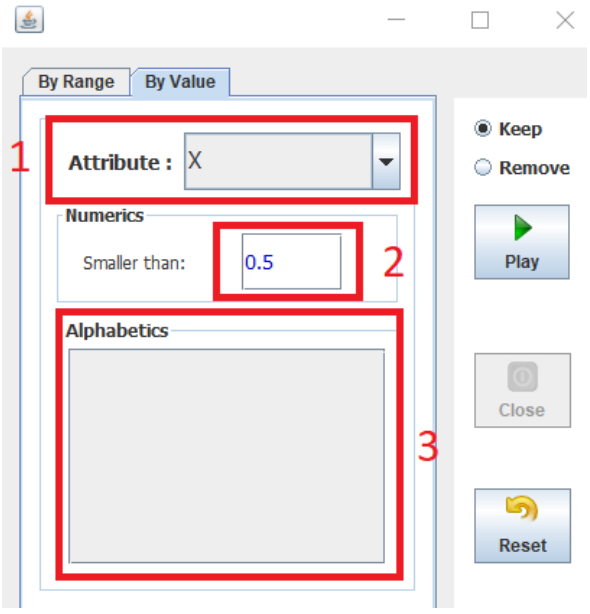


Por rango:



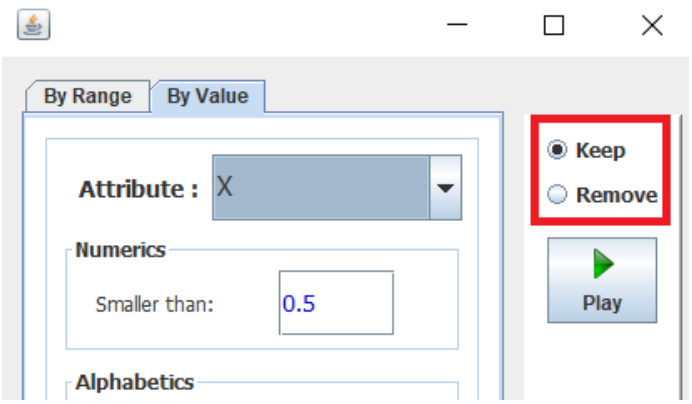
1. Primera fila
2. Última fila

Por valor:



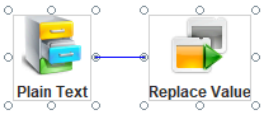
1. Atributo
2. Valor numérico menor qué
3. Cadenas

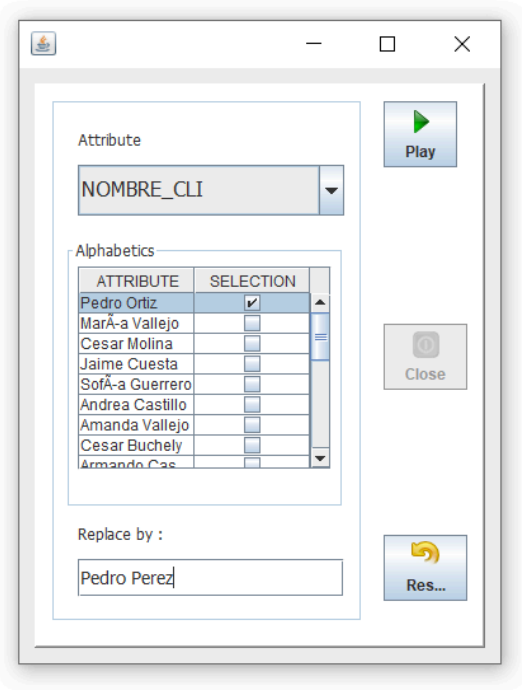
Ambos tienen las opciones “Keep” y “Remove”, que se utilizan para mantener o quitar los campos dentro del rango o del valor especificado.



3.1.7. Replace Value

Permite reemplazar valores, aplica para aquellos atributos del conjunto de datos que contengan información de tipo cadena o String.





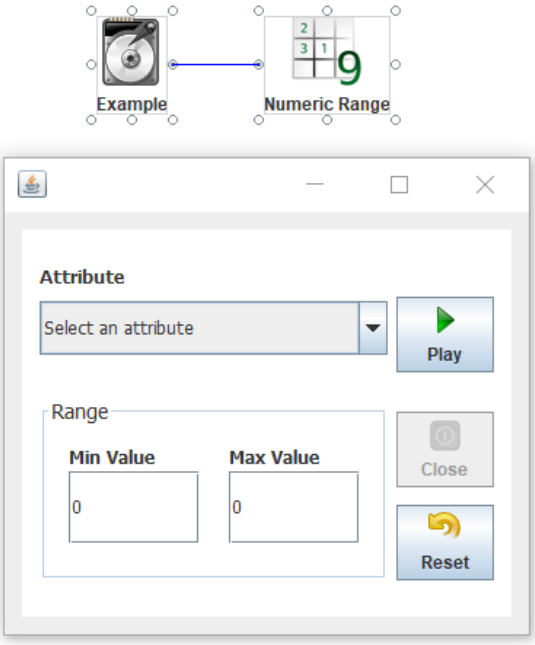
Cuenta con los siguientes apartados:



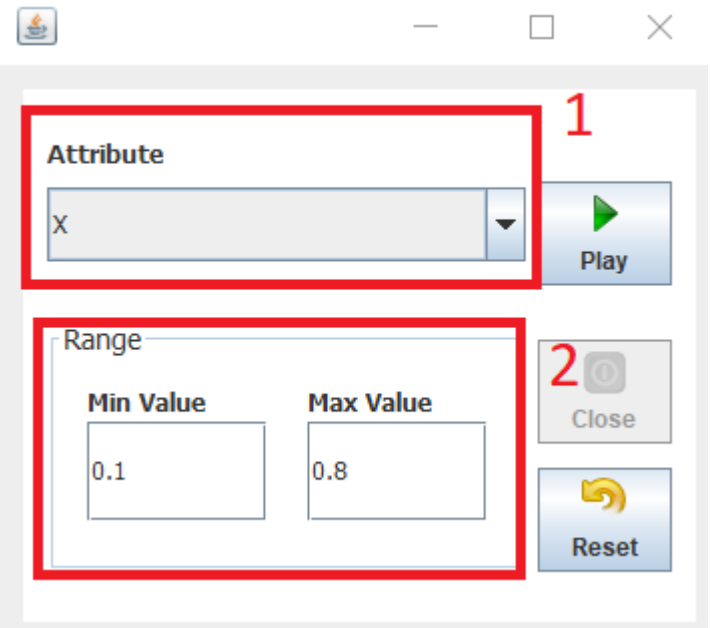
1. Atributo seleccionado.
2. Datos seleccionados para ser cambiados.
3. Texto con el cual se reemplazarán.

3.1.8. Numeric Range

Permite el filtrado de datos por atributos de tipo numérico (int, double).



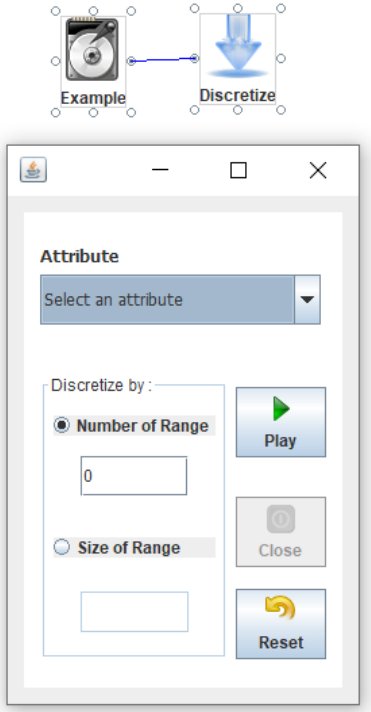
Cuenta con los siguientes apartados:



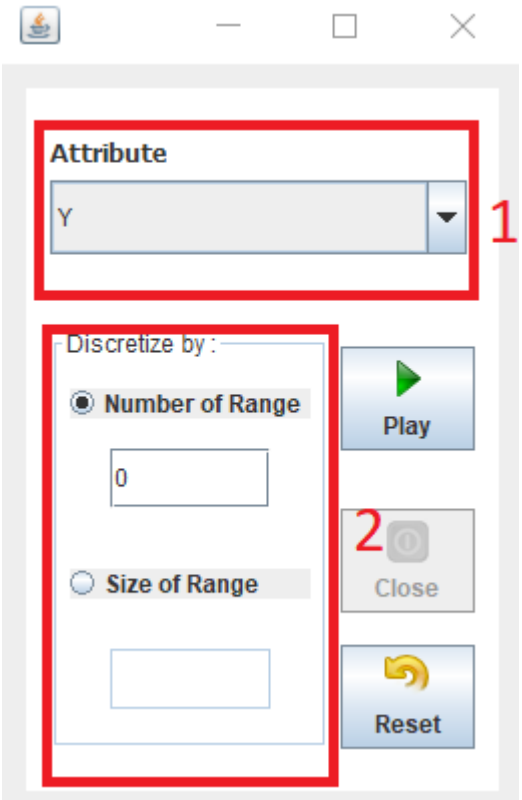
1. Atributo seleccionado
2. Rango de valores

3.1.9. Discretize

Permite realizar la categorización de atributos de carácter numérico tanto valores continuos como valores discretos.



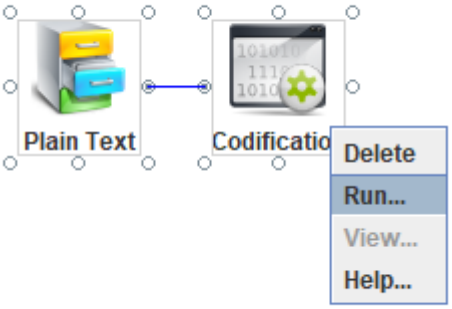
Cuenta con los siguientes apartados:



1. Atributo seleccionado
2. Opción de discretización: Por número de rango o tamaño de rango

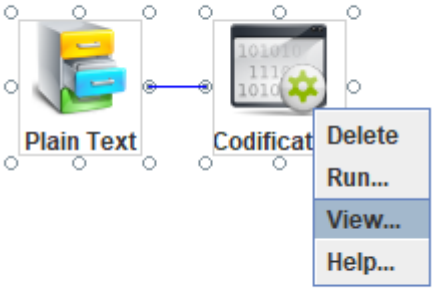
3.1.10. Codification

Asigna códigos automáticamente según el atributo y el valor.

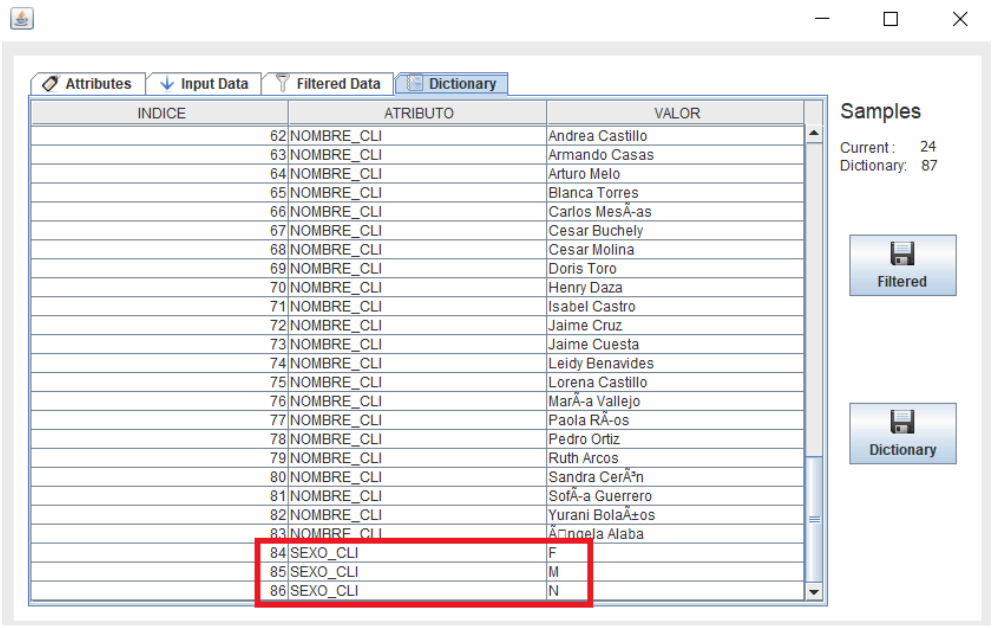


Al igual que el filtro “Remove Missing” no cuenta con una ventana de configuración por lo cual simplemente se selecciona la opción “Run”.

Luego de hacer esto estará disponible la opción “View” para ver los resultados de la codificación.

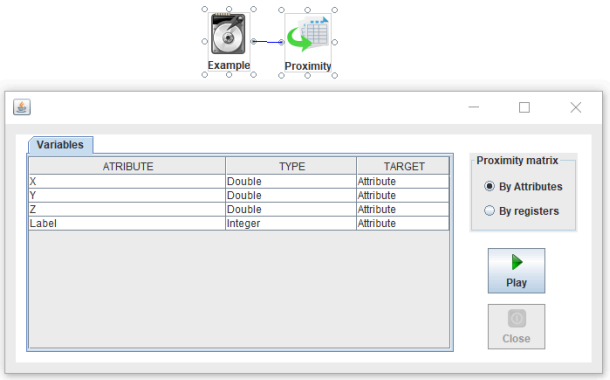


Al dar click en ella se mostrará una ventana con cuatro apartados: Atributos, Datos de entrada, Datos Filtrados y Diccionario, en este último se podrán observar los resultados de la codificación.



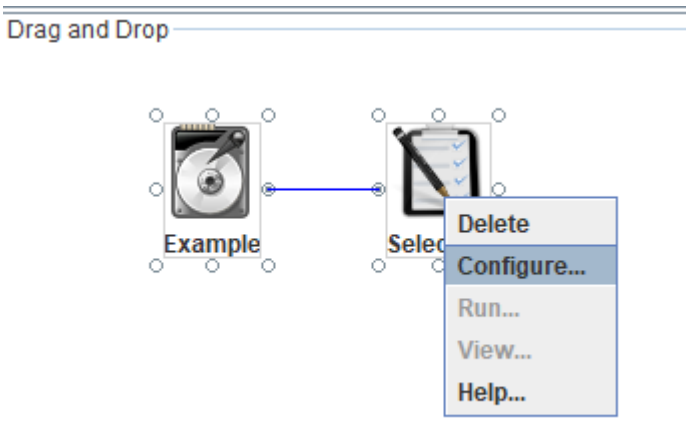
3.1.11. Proximity

Se encarga de crear una matriz de proximidad, realizando el producto entre la matriz presentada y su traspuesta o la traspuesta de dicha matriz por la matriz original, dependiendo de si se escoge entre la opción de trabajar con atributos o con registros.



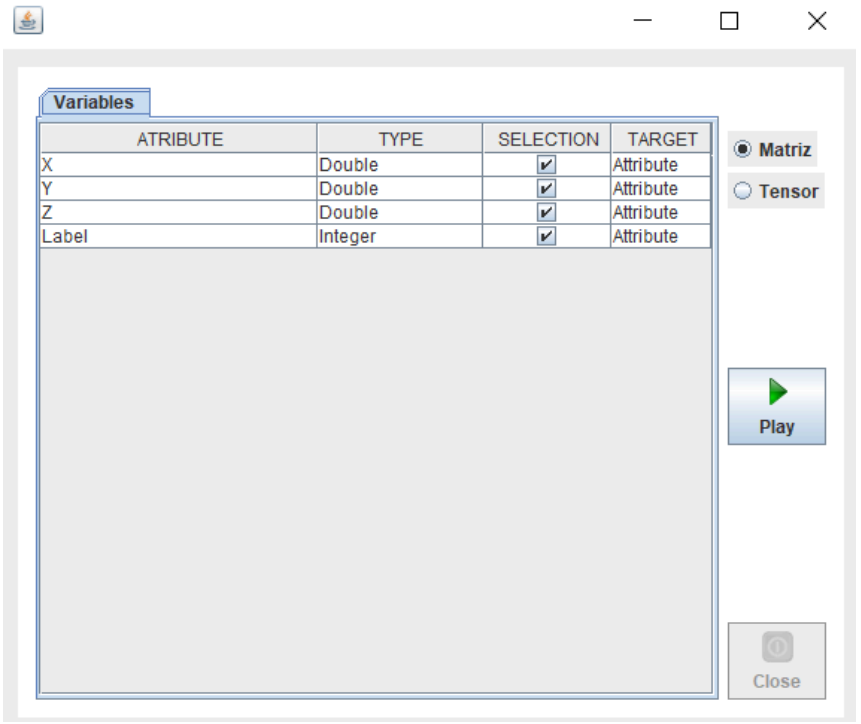
3.2. Ejemplo de uso

Ahora para utilizar la herramienta de limpieza de datos, en este caso “Selection”, una vez realizada la conexión con los datos, debe dar click derecho en la herramienta y seleccionar la opción “Configure”.

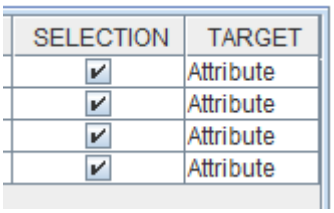


Esto abrirá una ventana en la cuál se muestran las siguientes características del conjunto de datos:

- **Atribute:** Nombre de la columna en la tabla.
- **Type:** Tipo de dato de dicha columna.
- **Selection:** Permite elegir si el atributo hace parte del filtro de selección o no.
- **Target:** Permite elegir la variable objetivo del filtro.

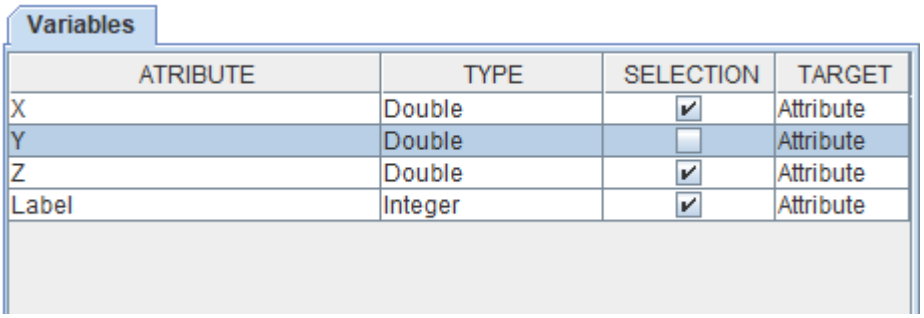


En el ejemplo seleccionado se tienen tres valores de tipo double, las coordenadas X,Y,Z y uno de tipo integer llamado Label o etiqueta, en este caso se elegirá como variable objetivo la etiqueta, caben destacar dos campos interactivos: selection y target.



SELECTION	TARGET
☑	Attribute
☑	Attribute
☑	Attribute
☑	Attribute

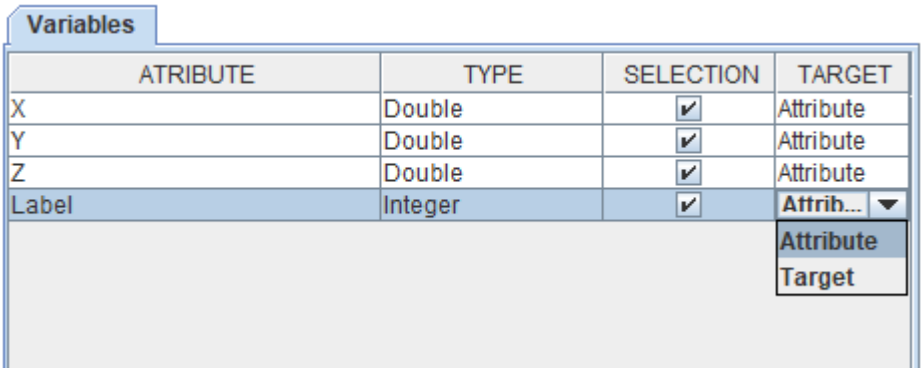
Si quiere omitir alguno de los atributos en el filtro, debe quitar la selección del mismo dando click en el checkbox correspondiente.



Variables

ATRIBUTE	TYPE	SELECTION	TARGET
X	Double	☑	Attribute
Y	Double	☐	Attribute
Z	Double	☑	Attribute
Label	Integer	☑	Attribute

Y para establecer un atributo como objetivo debe dar click en el campo target de la variable objetivo, ello abrirá un submenú con dos opciones: Attribute o Target.

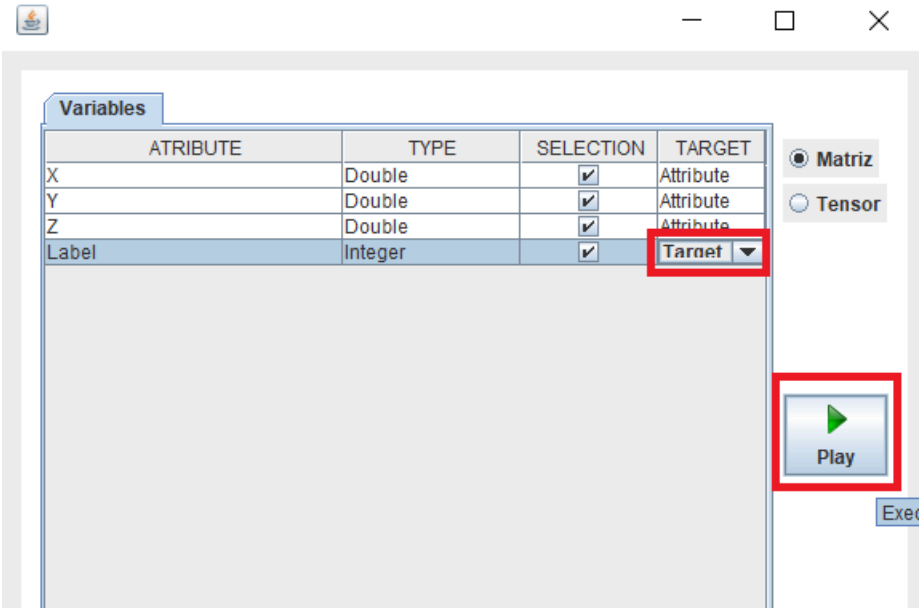


Variables

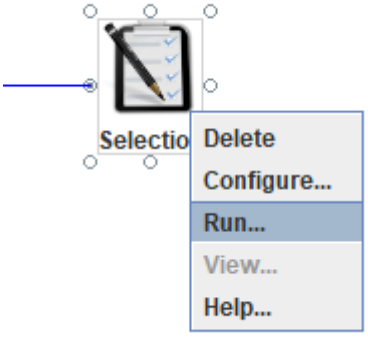
ATRIBUTE	TYPE	SELECTION	TARGET
X	Double	☑	Attribute
Y	Double	☑	Attribute
Z	Double	☑	Attribute
Label	Integer	☑	Attrib...

Attribute
Target

Selecciona “Target” para determinar el objetivo y da click en “Play”.



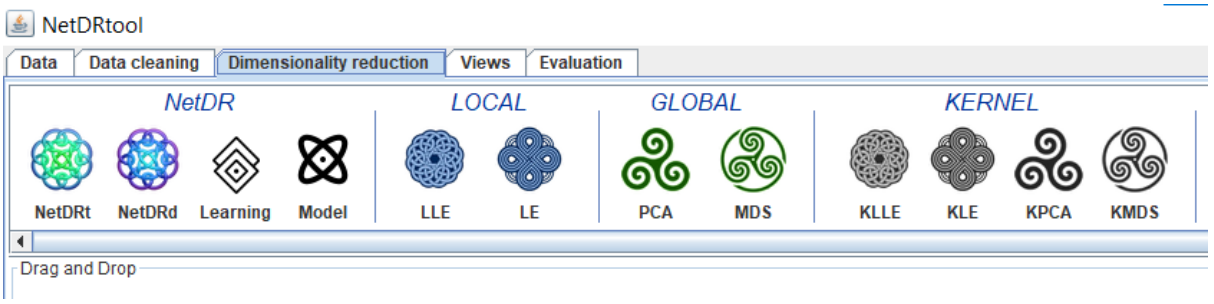
Luego proceda a dar click derecho en el filtro de selección y elija la opción “Run”.



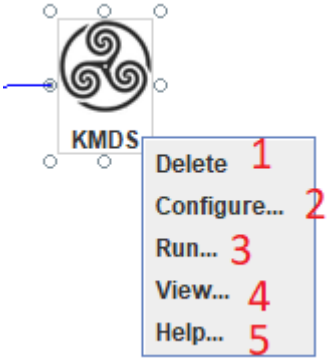
Con ello se ha determinado la variable que será supervisada.

4. Dimensionality reduction

En este apartado se encuentran los algoritmos de reducción de dimensión.

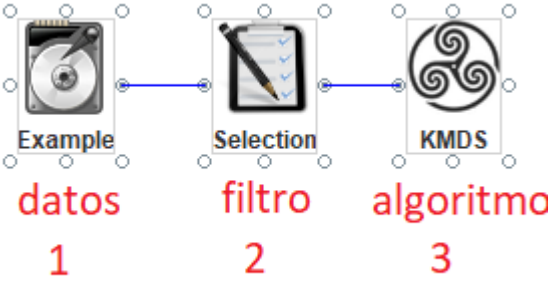


Cabe mencionar que estos componentes contienen las opciones:

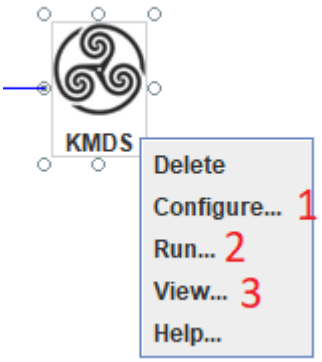


1. **Delete:** Elimina el componente.
2. **Configure:** Configurar el algoritmo.
3. **Run:** Ejecuta el algoritmo.
4. **View:** Ver los resultados de la reducción.
5. **Help:** Información acerca de los algoritmos de reducción.

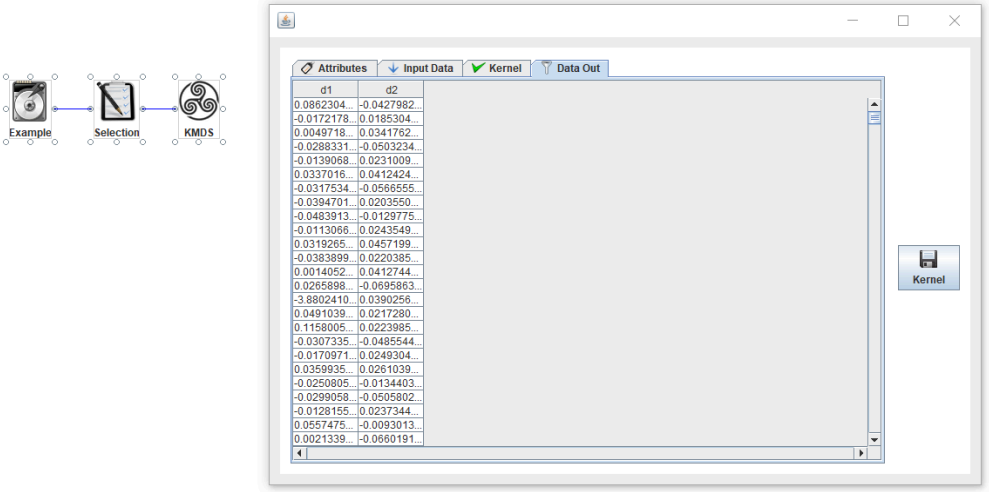
Además la conexión de estos componentes en general a excepción de los algoritmos de combinación, debe darse de la siguiente forma:



El filtro recomendado para estas configuraciones es ‘Selection’ y debe estar configurado y cargado con anterioridad para poder realizar la conexión, los pasos que debe seguir el algoritmo son similares a los mostrados en los filtros, primero se configura, luego se ejecuta el algoritmo y posteriormente estará disponible para visualizar sus datos.

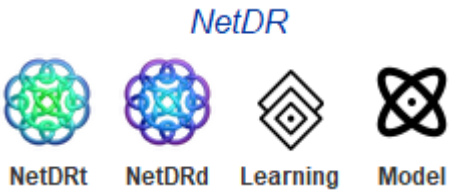


La opción ‘View’ despliega una ventana con los atributos trabajados y los resultados del respectivo algoritmo aplicado a los datos.



Estos algoritmos se dividen en las siguientes categorías:

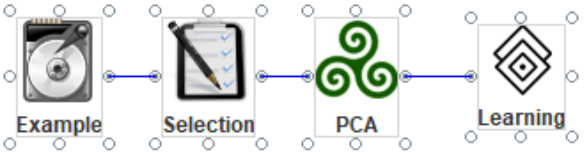
4.1. Herramientas NetDR



4.1.1. Learning

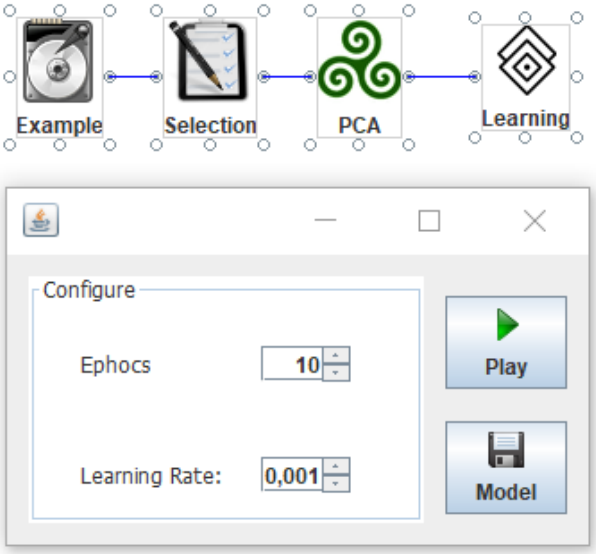
Learning es una herramienta que permite el aprendizaje neuronal de los incrustamientos de los algoritmos de reducción para generar modelos de reducción más eficientes.

Su conexión debe realizarse con cualquier componente que sea un algoritmo de reducción de dimensión para que pueda aprender de él, como se mira a continuación:

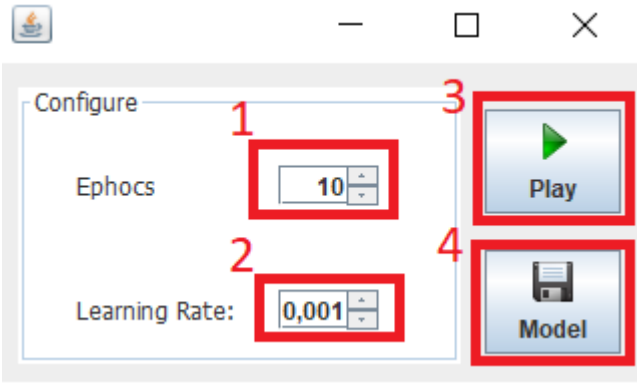


Nota: El algoritmo debe haber sido configurado y ejecutado con anterioridad.

Su ventana de configuración es la siguiente:

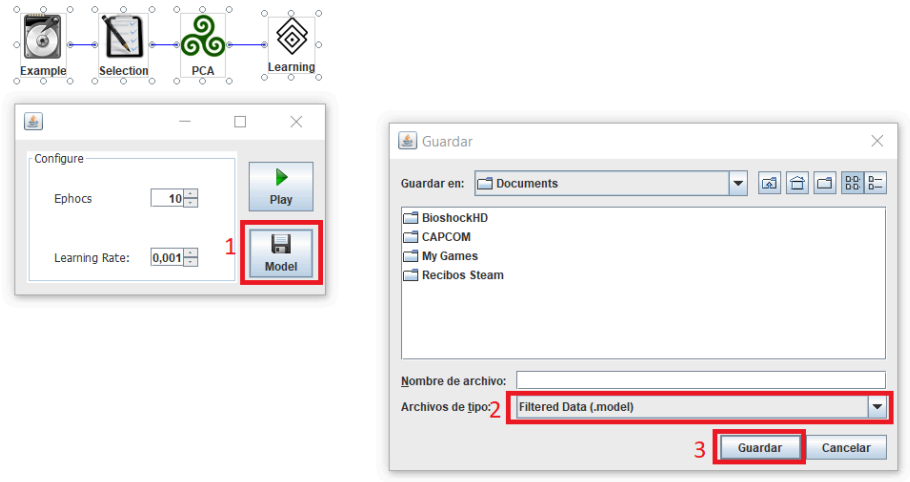


En ella se encuentran los apartados:



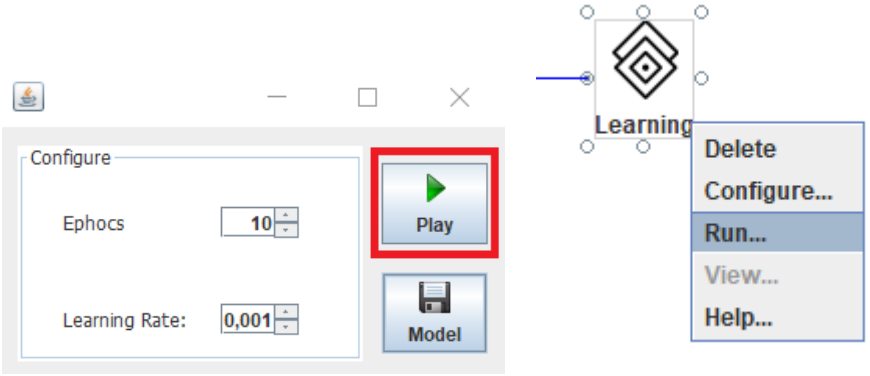
1. Épocas de aprendizaje.
2. Ritmo de aprendizaje.
3. Botón de ejecución.
4. Botón para guardar el modelo.

Una vez se escoja el número de épocas y el ritmo de aprendizaje, se debe escoger el directorio donde se guardará el modelo a entrenar.

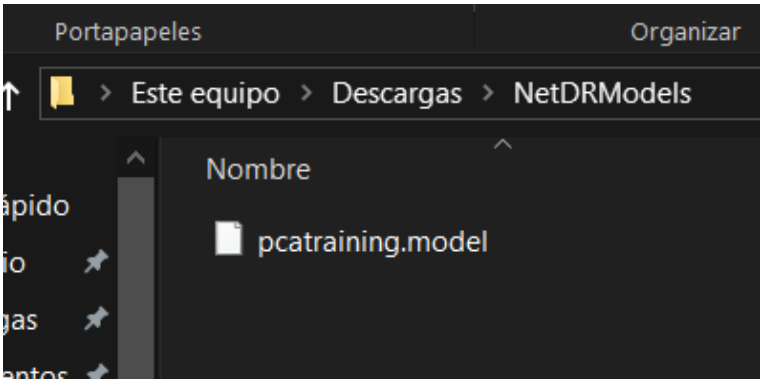


Nota: El archivo a guardar debe tener la extensión .model

Luego de haber guardado el modelo debe dar click en 'Play' y después click en 'Run'.



Una vez termine de cargar se habrá guardado el modelo.



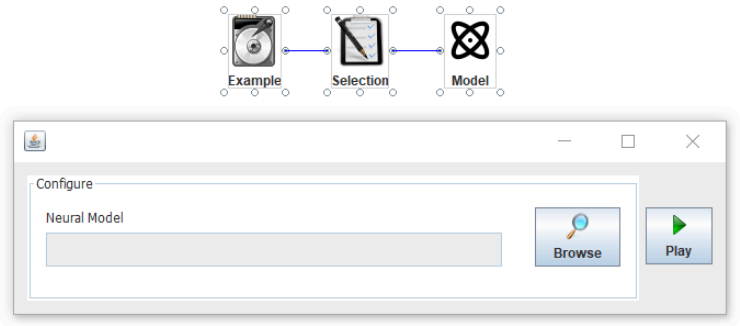
4.1.2. Model

La herramienta ‘Model’ permite cargar modelos previamente creados con la herramienta ‘Learning’.

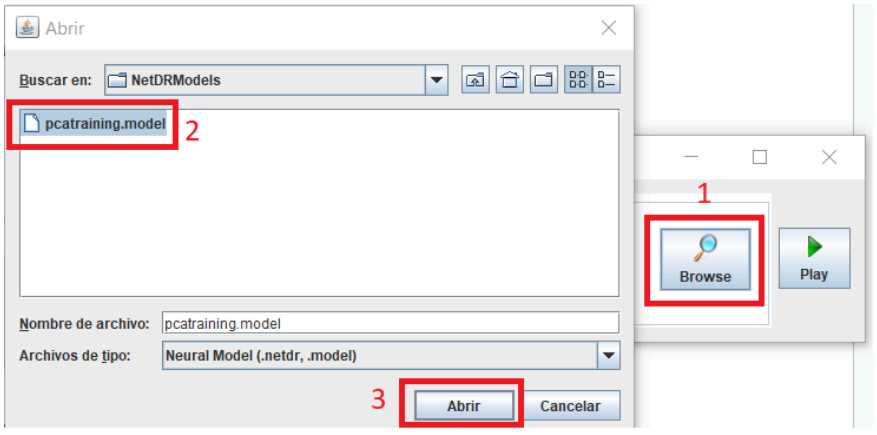
Su conexión con los datos es igual a la de los algoritmos:



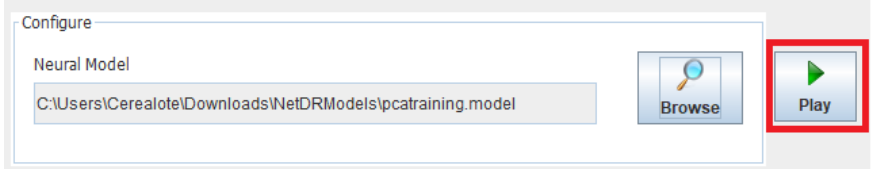
Y su ventana de configuración es la siguiente:



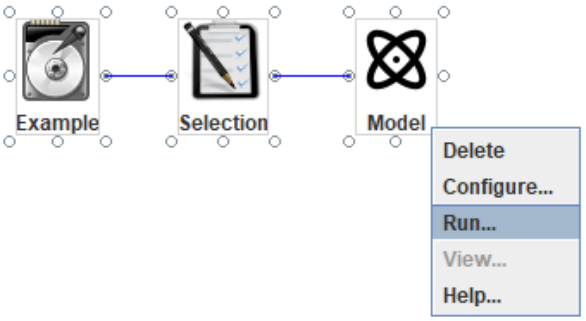
Selecciona la opción ‘Browse’ para buscar el modelo a cargar, lo selecciona y lo importa al proyecto.



Click en ‘Play’



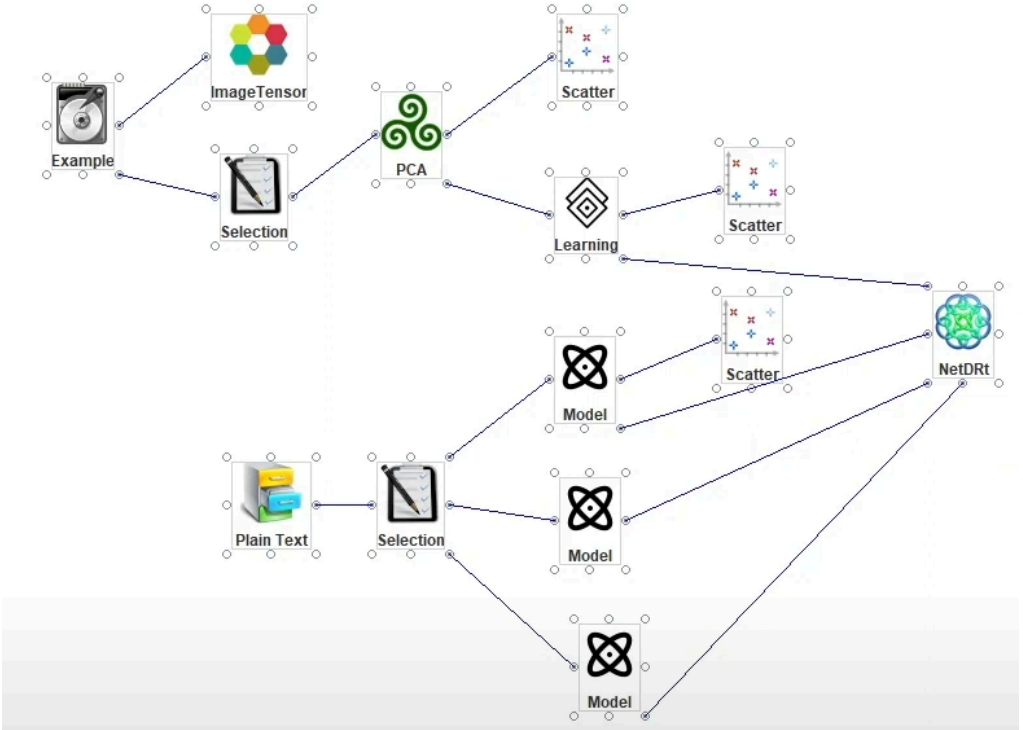
Luego click en ‘Run’



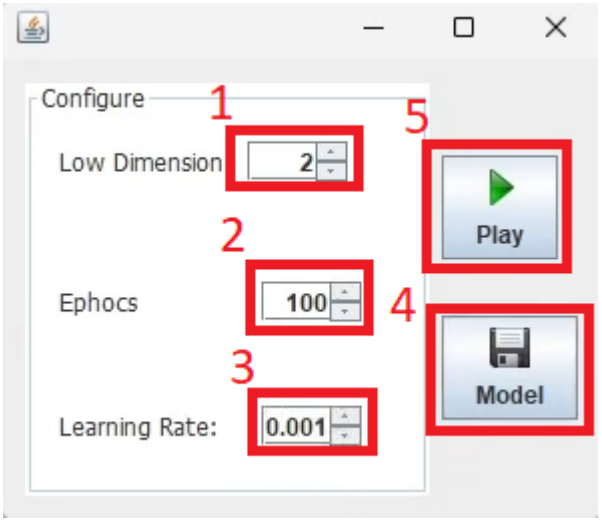
Con ello el modelo se habrá cargado y ejecutado.

4.1.3. NetDRt

NetDRt es un método topológico que permite combinar varios métodos aprendidos con antelación para generar un modelo de reducción de dimensión más eficiente.

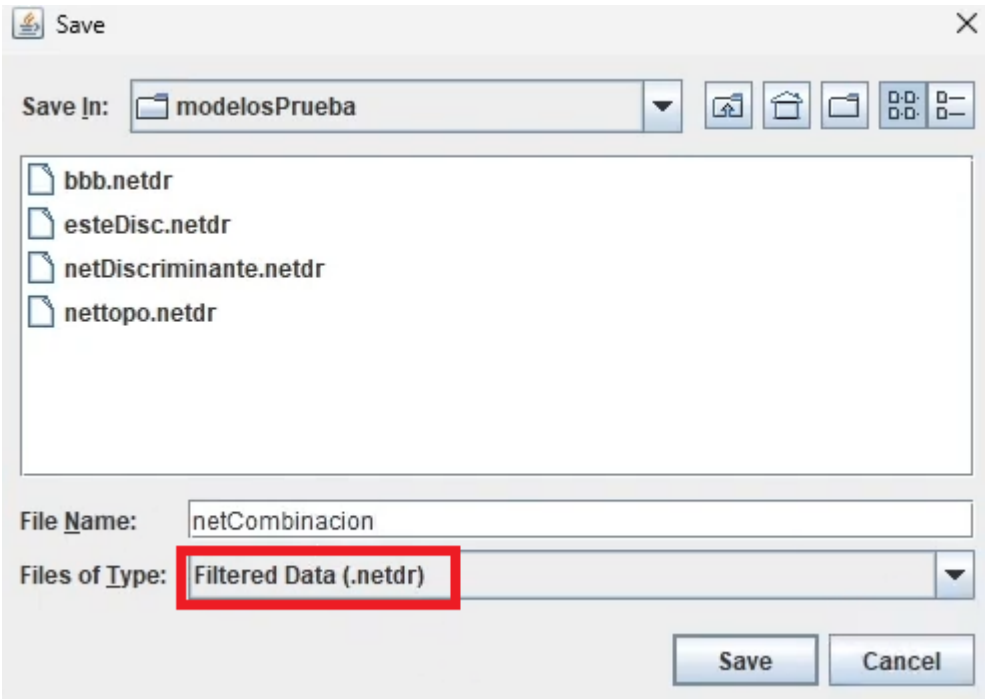


Su ventana de configuración es la siguiente:

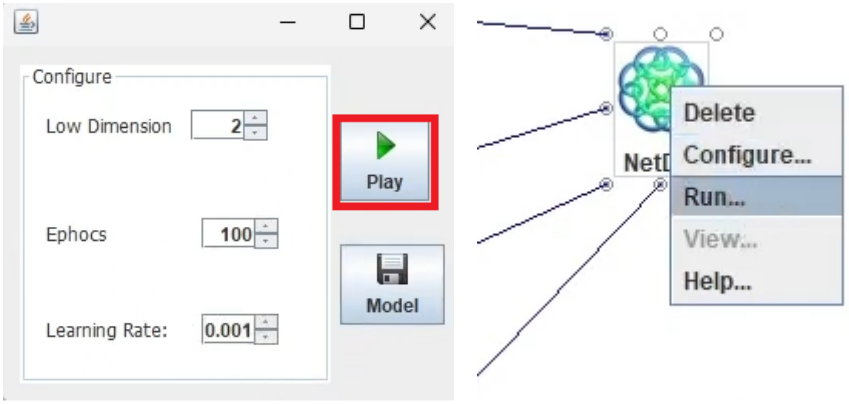


1. Dimensiones a las que se quiere reducir los datos.
2. Épocas de aprendizaje.
3. Ritmo de aprendizaje.
4. Botón para guardar el modelo.
5. Botón de ejecución.

Una vez configurados los parámetros, al igual que con la herramienta learning se debe elegir el directorio donde se guardará el modelo, en este caso su extensión es .netdr



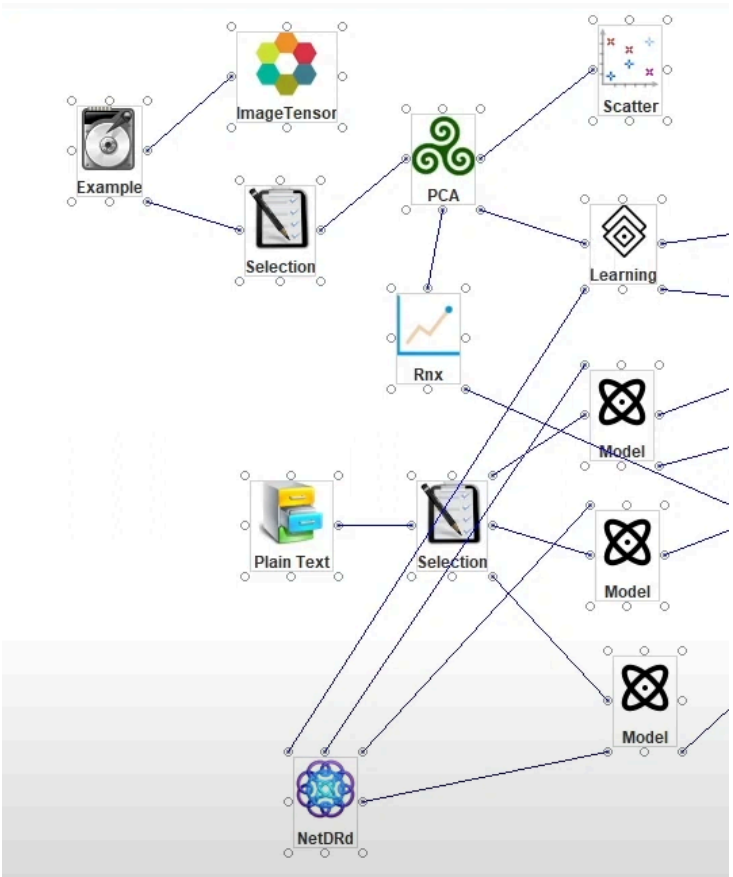
Una vez hecho esto se debe dar click en el botón ‘Play’ y posteriormente en ‘Run’



Con ello se cargará y guardará el nuevo modelo.

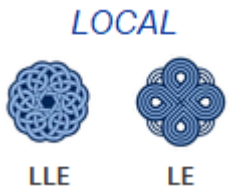
4.1.4. NetDRd

NetDRd es un método discriminante que permite combinar varios métodos aprendidos con antelación para generar un modelo de reducción de dimensión más eficiente.



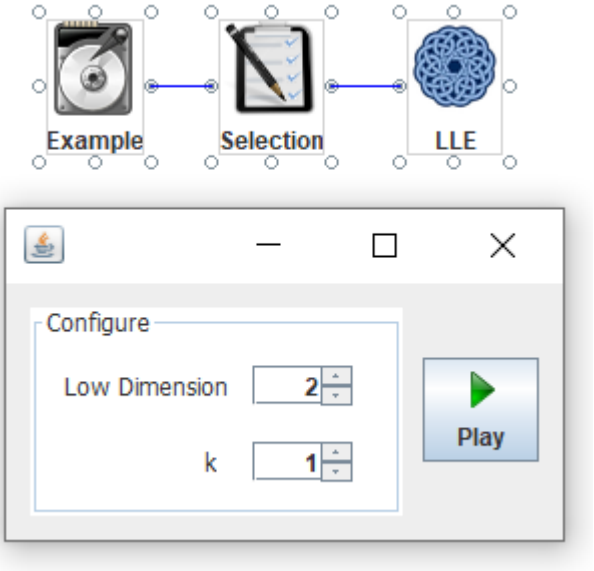
Los pasos para su configuración y ejecución son exactamente los mismos que con el componente anterior 'NetDRt'.

4.2. Algoritmos Locales

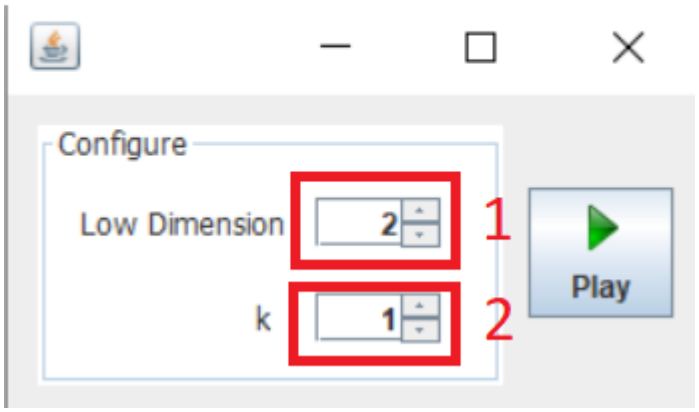


4.2.1. LLE

El algoritmo LLE presenta la siguiente ventana de configuración:



Donde se deben especificar los valores:

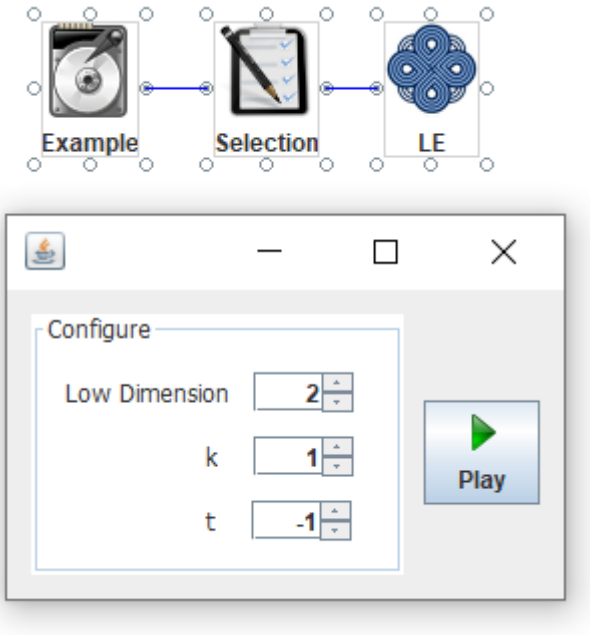


1. Dimensiones a las que se quiere reducir los datos.

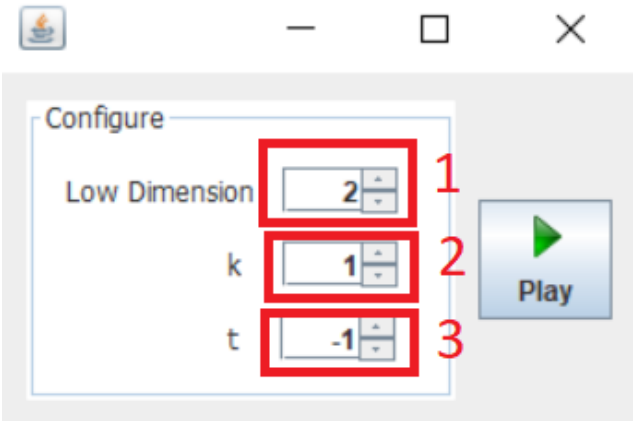
2. Tamaño del vecindario

4.2.2. LE

El algoritmo LE presenta la siguiente ventana de configuración:



Donde se deben especificar los valores:



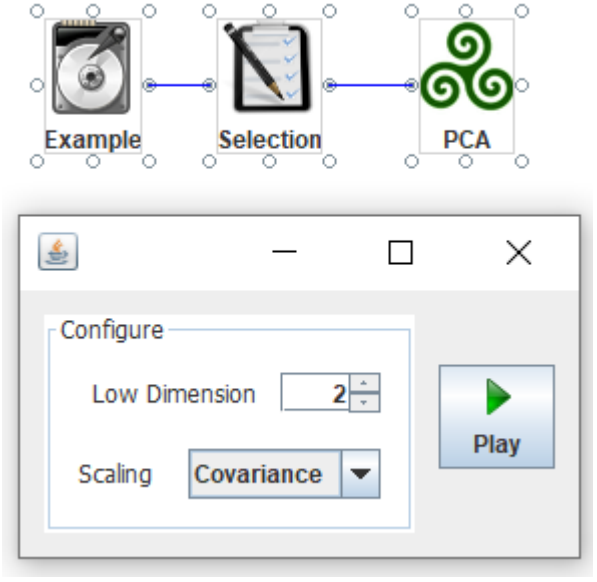
1. Dimensiones a las que se quiere reducir los datos.
2. Tamaño del vecindario
3. Temperatura para conformar el grafo Laplaciano

4.3. Algoritmos Globales

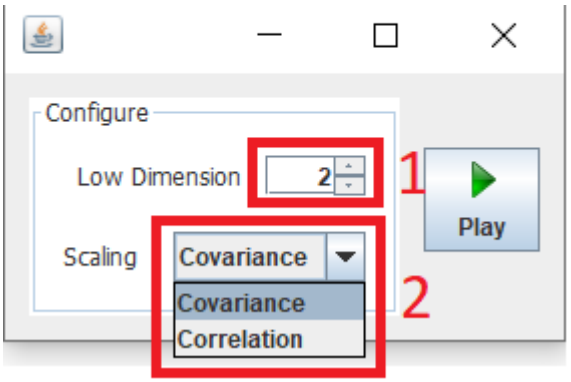


4.3.1. PCA

El algoritmo PCA presenta la siguiente ventana de configuración:



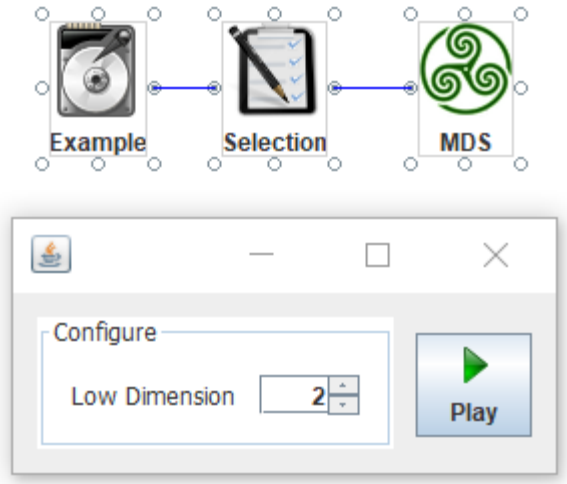
Donde se deben especificar los valores:



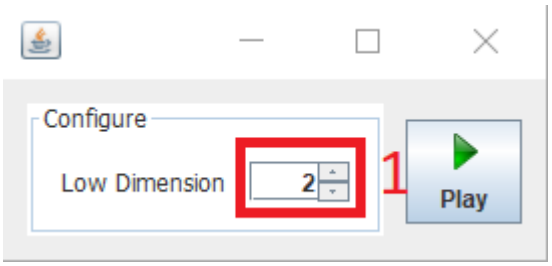
1. Dimensiones a las que se quiere reducir los datos.
2. Puede ser Covarianza utilizando matriz de disimilitud o Correlación utilizando matriz de afinidad.

4.3.2. MDS

El algoritmo MDS presenta la siguiente ventana de configuración:

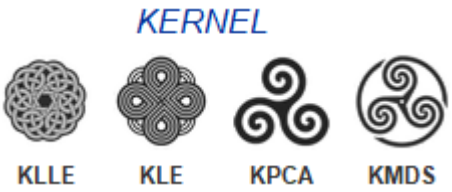


Donde se debe especificar el valor:



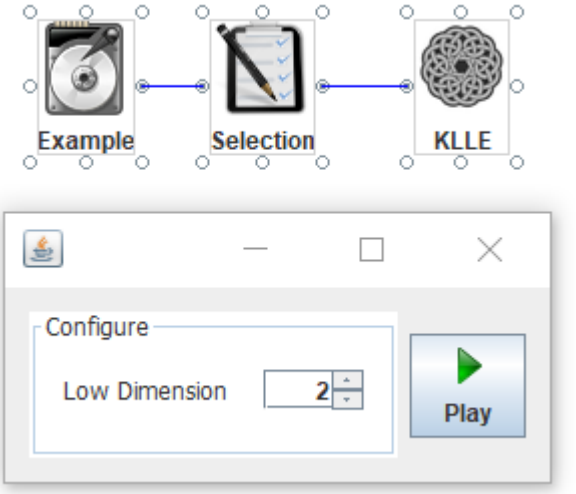
1. Dimensiones a las que se quiere reducir los datos.

4.4. Algoritmos de Kernel



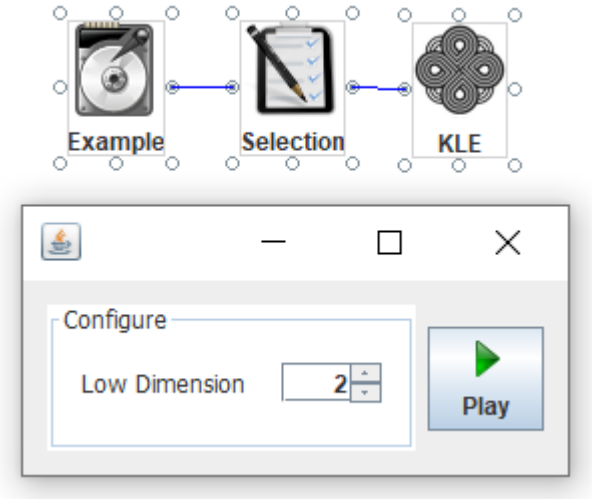
4.4.1. KLLE

El algoritmo KLLE presenta la siguiente ventana de configuración:



4.4.2. KLE

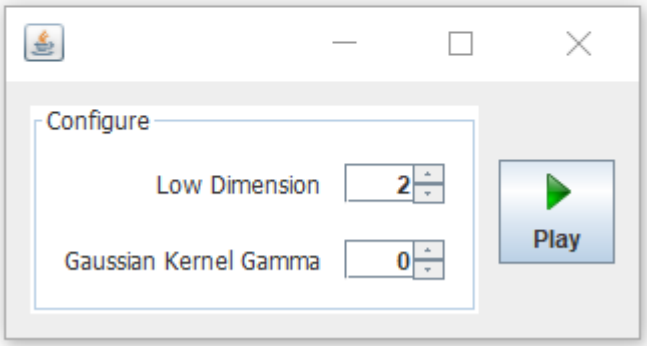
El algoritmo KLE presenta la siguiente ventana de configuración:



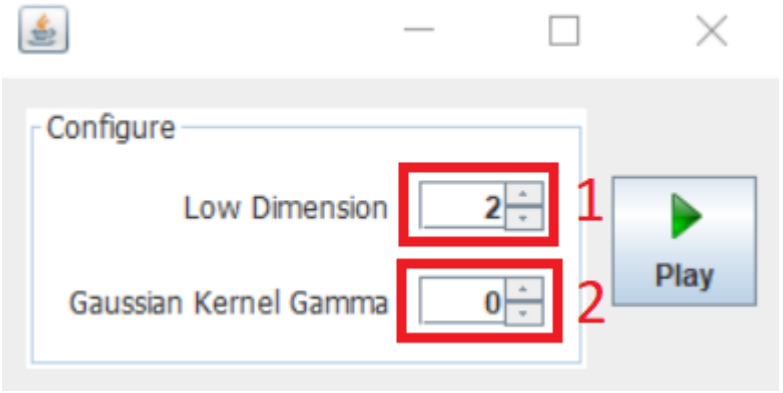
4.4.3. KPCA

El algoritmo KPCA presenta la siguiente ventana de configuración:





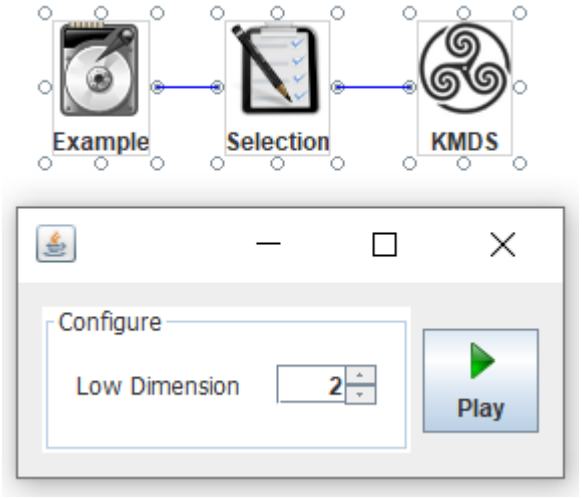
Donde se deben especificar los valores:



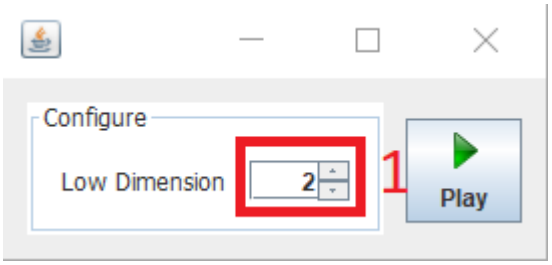
1. Dimensiones a las que se quiere reducir los datos.
2. Hiper-parámetro que suministra el grado del polinomio, la escala y el desfase de un kernel Gaussiano.

4.4.4. KMDS

El algoritmo KMDS presenta la siguiente ventana de configuración:

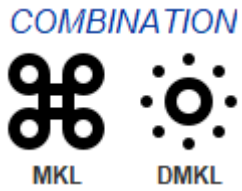


Nota: A excepción del algoritmo KPCA, en los algoritmos de Kernel solamente debe especificarse un valor.



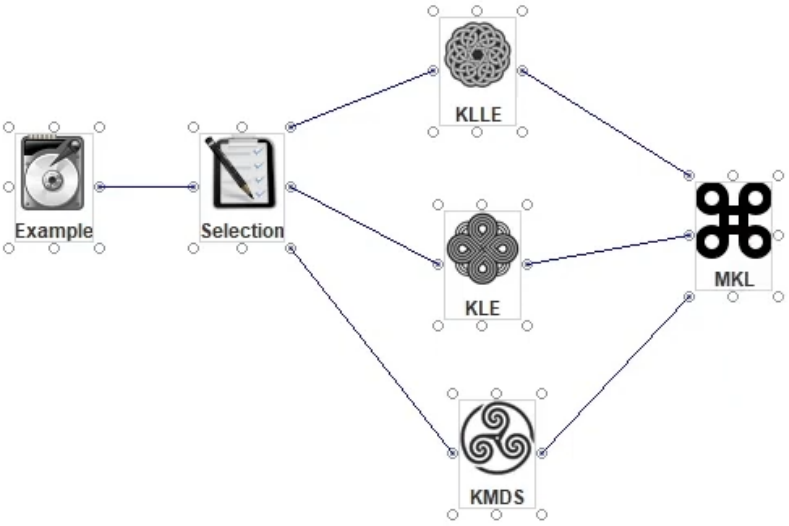
1. Dimensiones a las que se quiere reducir los datos.

4.5. Algoritmos de Combinación



4.5.1. MKL

El algoritmo MKL permite combinar dos o tres métodos kernel, como se muestra en el siguiente ejemplo:



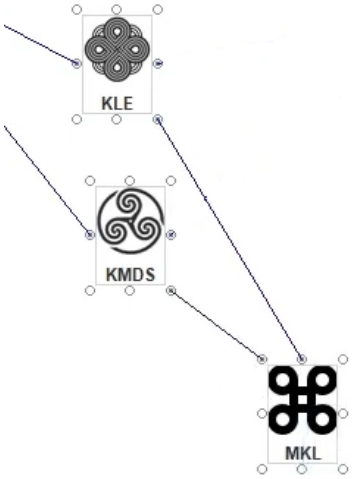
Nota: Los algoritmos que conectan al componente MKL deben haber sido configurados y ejecutados con anterioridad.

La ventana de configuración es la siguiente:

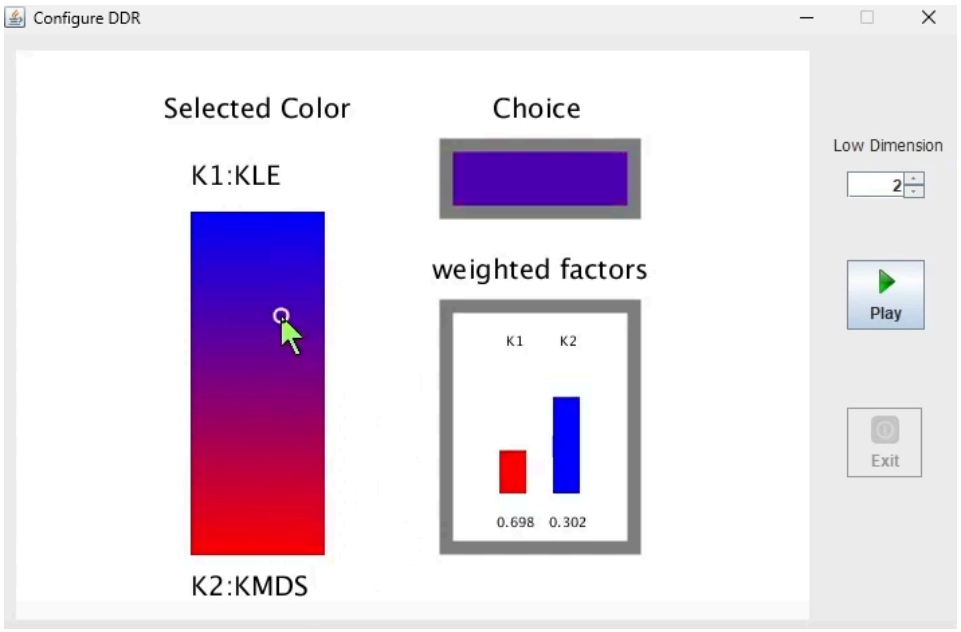


En ella mediante la selección de color en el triángulo cromático, se le puede dar prioridad a uno u otro algoritmo. El resto de la configuración es habitual.

Ahora si decide combinar dos kernels en lugar de tres, como se puede observar en la siguiente imagen:



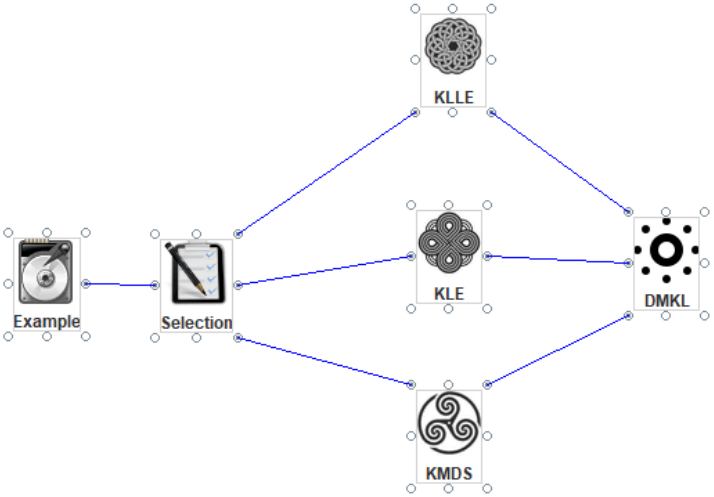
Entonces su ventana de configuración será la siguiente:



Su funcionamiento es exactamente igual al anterior, solamente que en este caso se cuenta con un rectángulo cromático, para dos kernels.

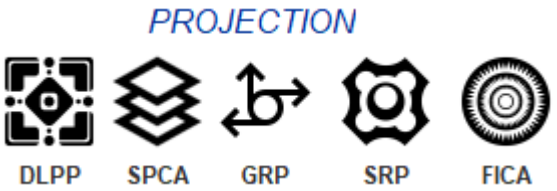
4.5.2. DMKL

El algoritmo DMKL permite combinar dos o tres métodos kernel, como se muestra en el siguiente ejemplo:



Nota: El algoritmo DMKL de momento no presenta ventana de configuración, se ha sugerido que su configuración es similar a la de MKL.

4.6. Algoritmos de Proyección

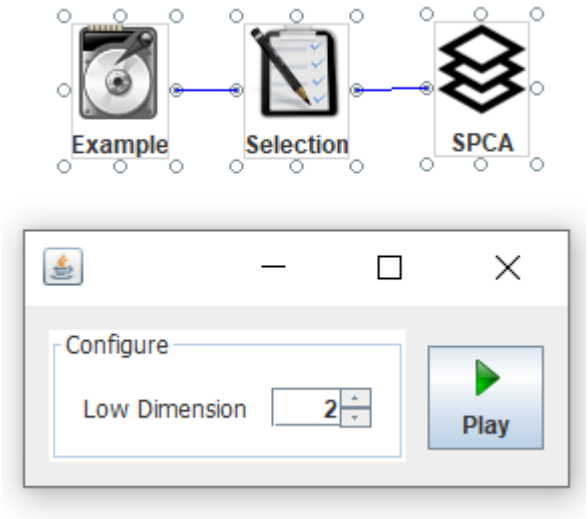


4.6.1. DLPP

Nota: El algoritmo DLPP de momento no presenta ventana de configuración, se ha sugerido que su configuración es similar a la de SPCA.

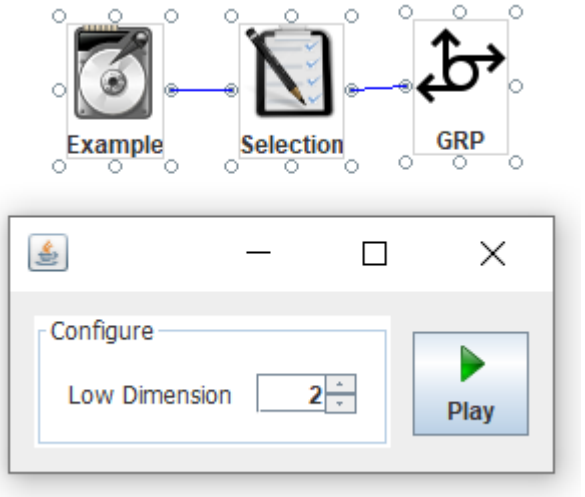
4.6.2. SPCA

El algoritmo SPCA presenta la siguiente ventana de configuración:



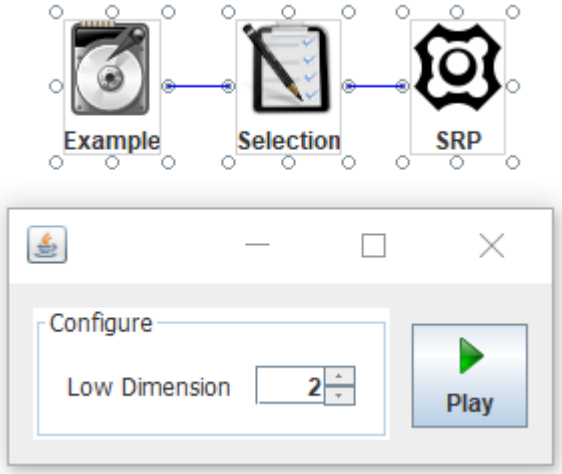
4.6.3. GRP

El algoritmo GRP presenta la siguiente ventana de configuración:



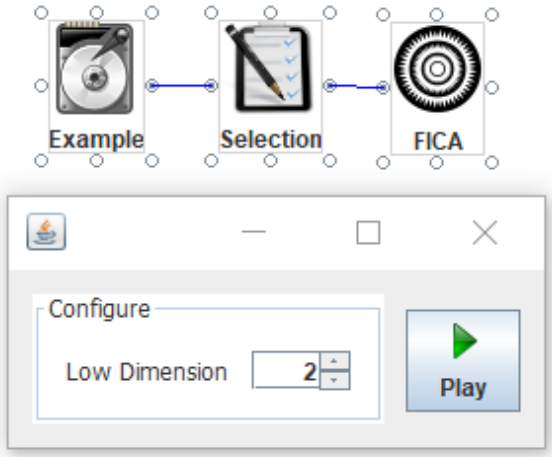
4.6.4. SRP

El algoritmo SRP presenta la siguiente ventana de configuración:

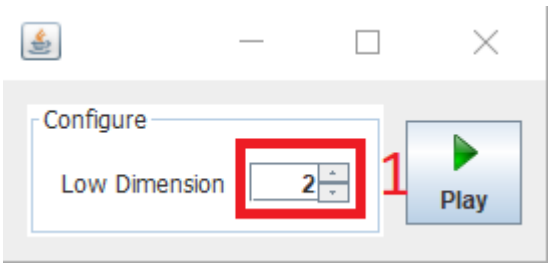


4.6.5. FICA

El algoritmo FICA presenta la siguiente ventana de configuración:

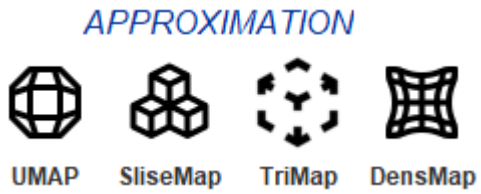


Nota: En los algoritmos de Proyección solamente debe especificarse un valor.



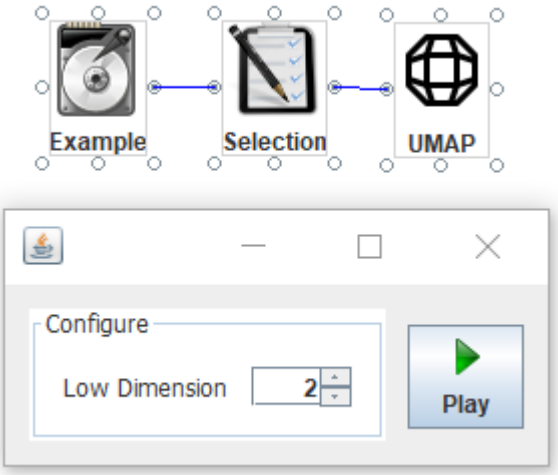
1. Dimensiones a las que se quiere reducir los datos.

4.7. Algoritmos de Aproximación



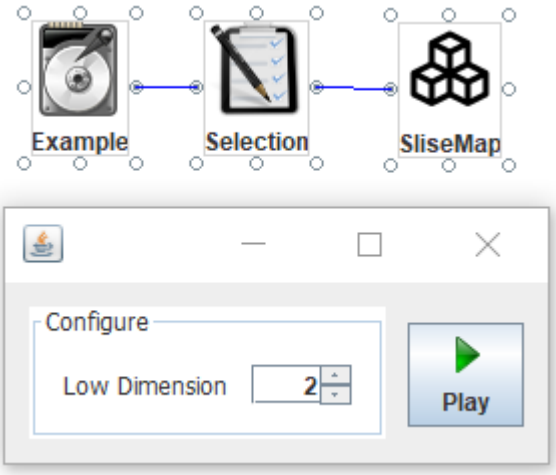
4.7.1. UMAP

El algoritmo UMAP presenta la siguiente ventana de configuración:



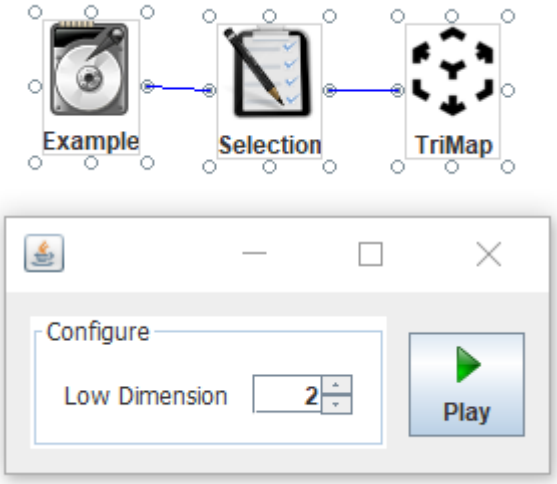
4.7.2. SliseMap

El algoritmo SliseMap presenta la siguiente ventana de configuración:



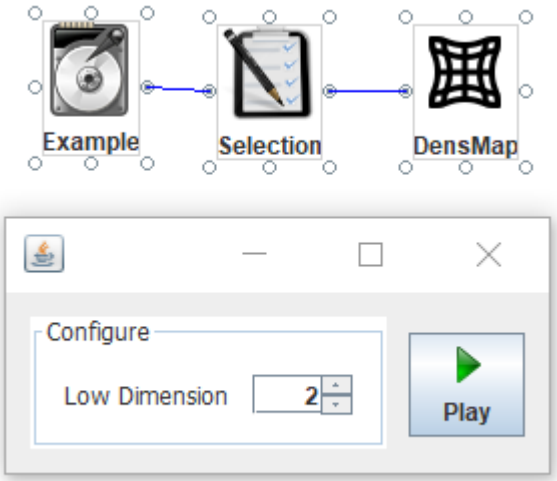
4.7.3. TriMap

El algoritmo TriMap presenta la siguiente ventana de configuración:

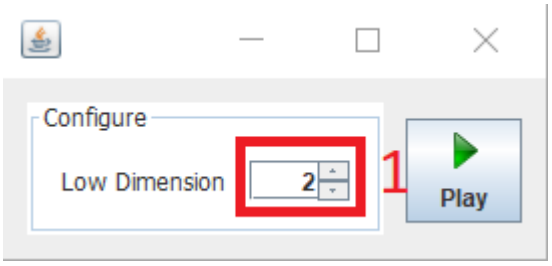


4.7.4. DensMap

El algoritmo DensMap presenta la siguiente ventana de configuración:

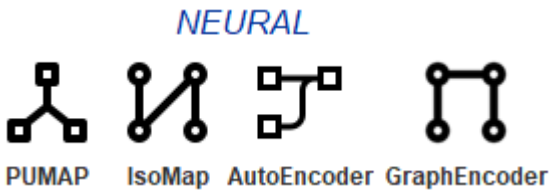


Nota: En los algoritmos de Aproximación solamente debe especificarse un valor.



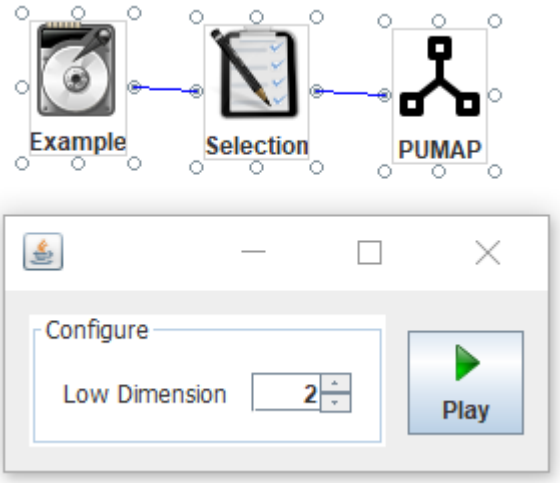
1. Dimensiones a las que se quiere reducir los datos.

4.8. Algoritmos Neuronales



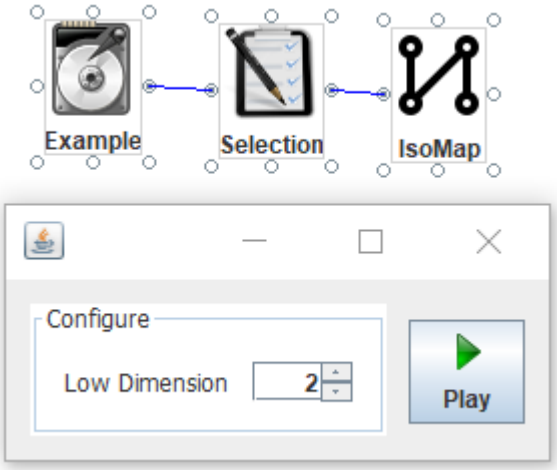
4.8.1. PUMAP

El algoritmo PUMAP presenta la siguiente ventana de configuración:



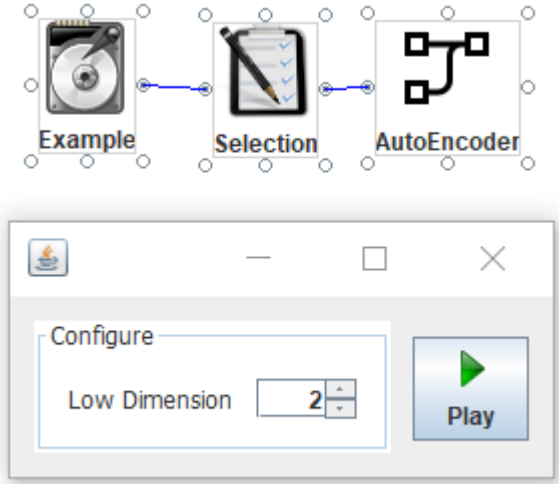
4.8.2. IsoMap

El algoritmo IsoMap presenta la siguiente ventana de configuración:



4.8.3. AutoEncoder

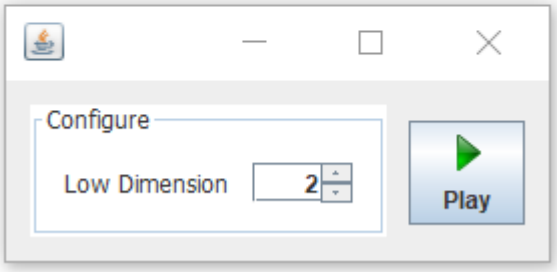
El algoritmo AutoEncoder presenta la siguiente ventana de configuración:



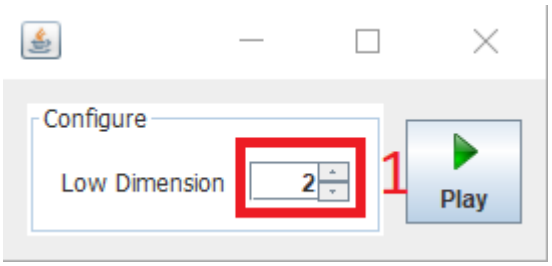
4.8.4. GraphEncoder

El algoritmo GraphEncoder presenta la siguiente ventana de configuración:



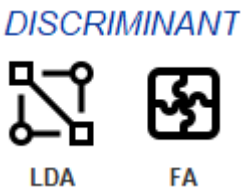


Nota: En los algoritmos Neuronales solamente debe especificarse un valor.



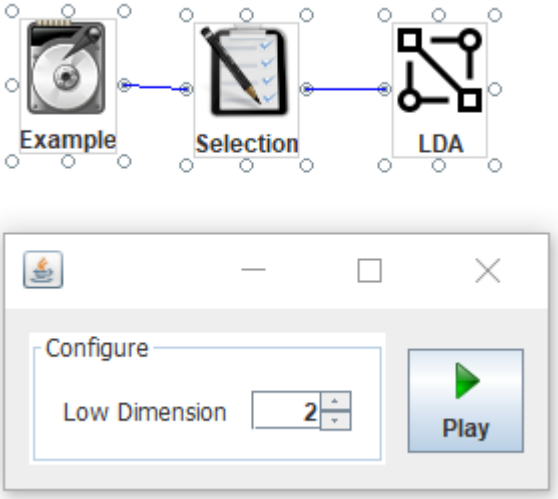
1. Dimensiones a las que se quiere reducir los datos.

4.9. Algoritmos Discriminantes



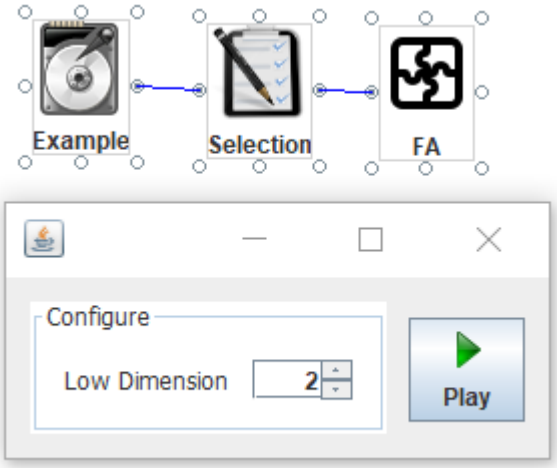
4.9.1. LDA

El algoritmo LDA presenta la siguiente ventana de configuración:

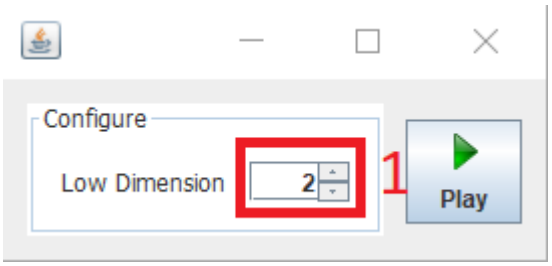


4.9.2. FA

El algoritmo FA presenta la siguiente ventana de configuración:



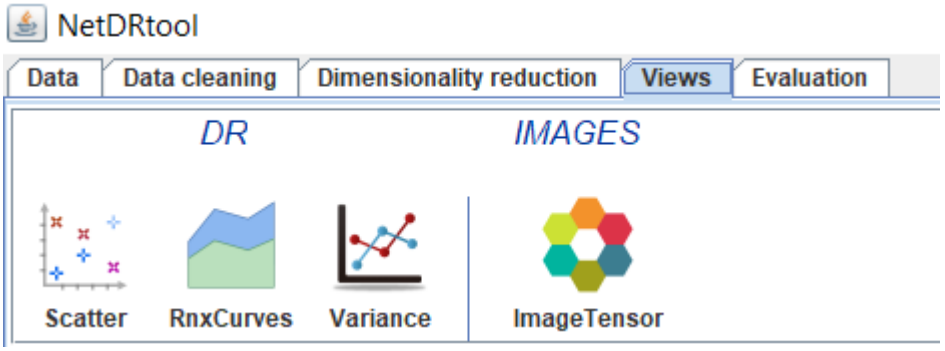
Nota: En los algoritmos Discriminantes solamente debe especificarse un valor.



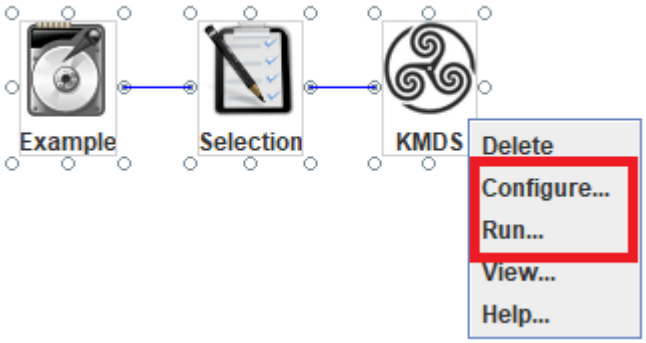
1. Dimensiones a las que se quiere reducir los datos.

5. Views

Contiene los visualizadores gráficos de los datos.



Nota: Para que los componentes se conecten al visualizador, deben estar previamente configurados y cargados al programa.



5.1. Tipos de visualizadores

5.1.1. Scatter

Visualizador de datos en dos y tres dimensiones, un componente de visualización Scatter puede conectarse con el componente de Data cleaning llamado “Selection”, con los algoritmos de “Dimensionality reduction” y con las herramientas de evaluación, tal y como se muestra en las siguientes imágenes.

Conexión con componente Selection:

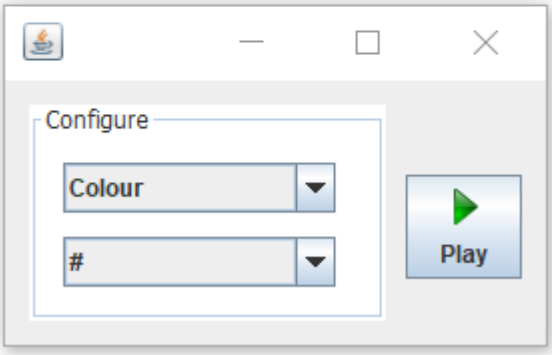


Conexión con componente de Dimensionality Reduction

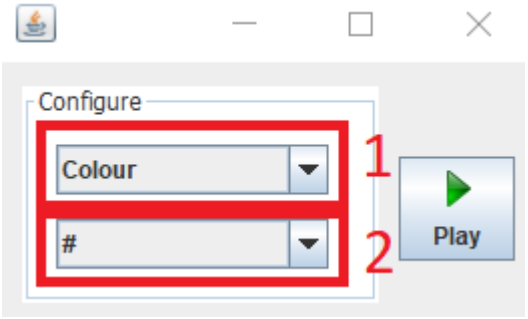


El visualizador Scatter presenta la siguiente ventana de configuración:



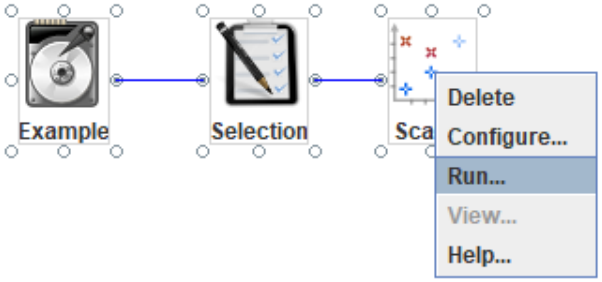


Donde se deben especificar los campos:

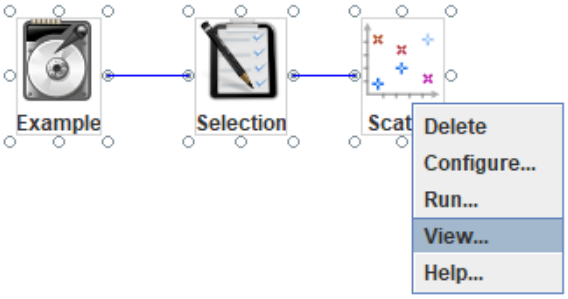


1. Tipo de representación de datos: Nominal, Color, Blanco y negro y RGB (Los datos deben ser compatibles).
2. Símbolo con el que se representa cada punto en la gráfica.

Una vez configurado el visualizador, debe dar click al botón “Play” de la ventana anterior y posteriormente dar click en la opción “Run” del visualizador.

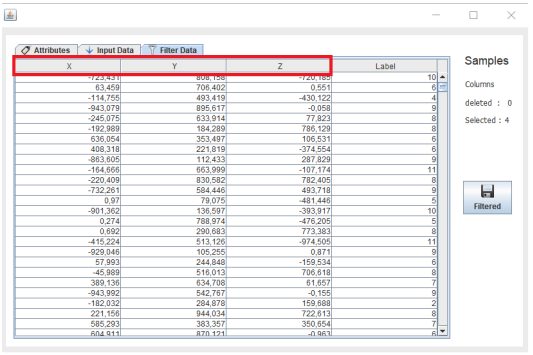


Una vez configurado y cargado el visualizador, se da click en la opción View.

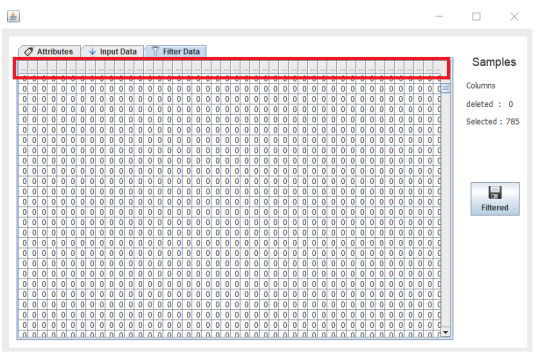


Así puede observarse la representación gráfica de los datos.

Nota: Si el filtro ‘Selection’ cuenta con más de tres variables en el resultado de su aplicación, entonces el visualizador Scatter no funcionará debido a que solamente puede representar gráficamente hasta tres dimensiones.

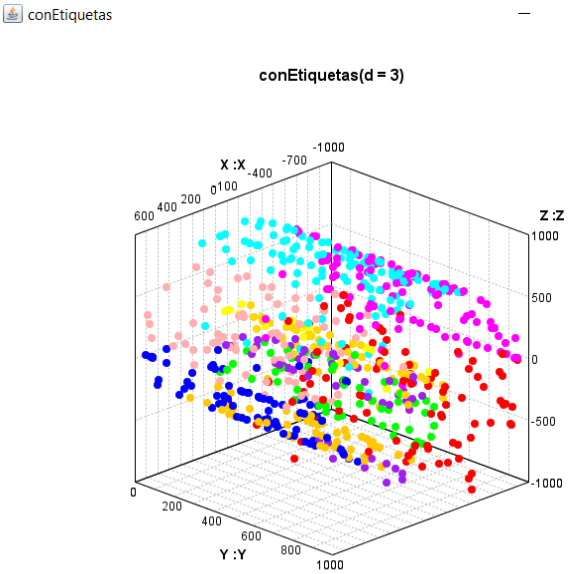


RESULTADO DE 3 DIMENSIONES

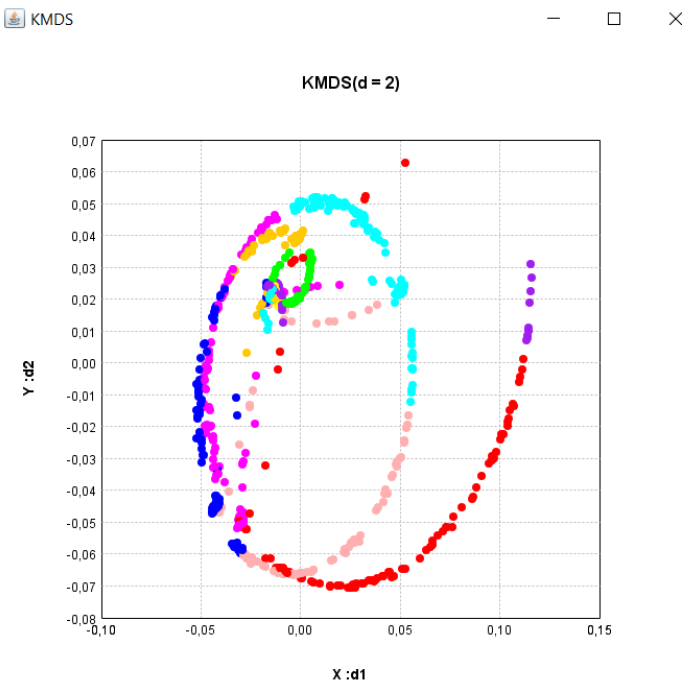


RESULTADO MAYOR A 3 DIMENSIONES

Ejemplo de visualizador en 3D (Datos del componente “Selection”):

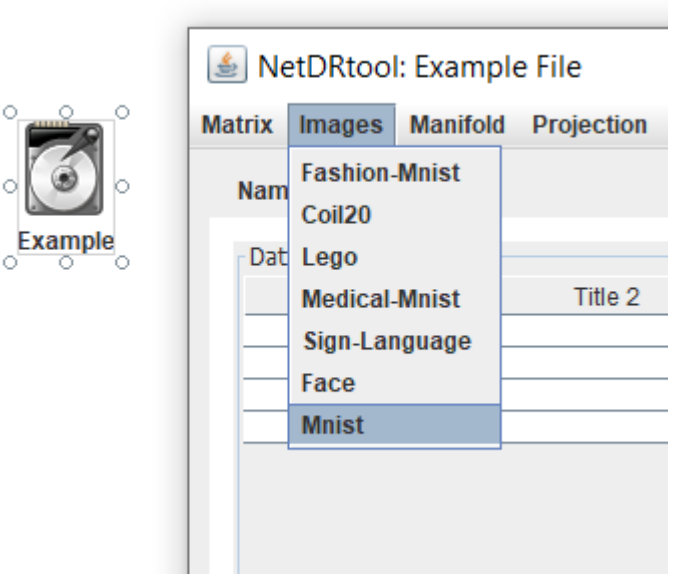


Ejemplo de visualizador en 2D (Datos reducidos con el algoritmo KMDS):

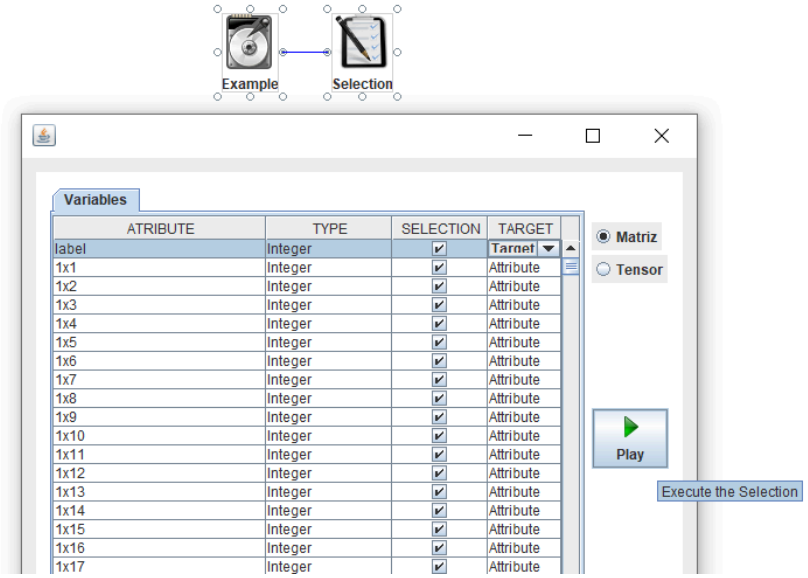


Ejemplo de uso:

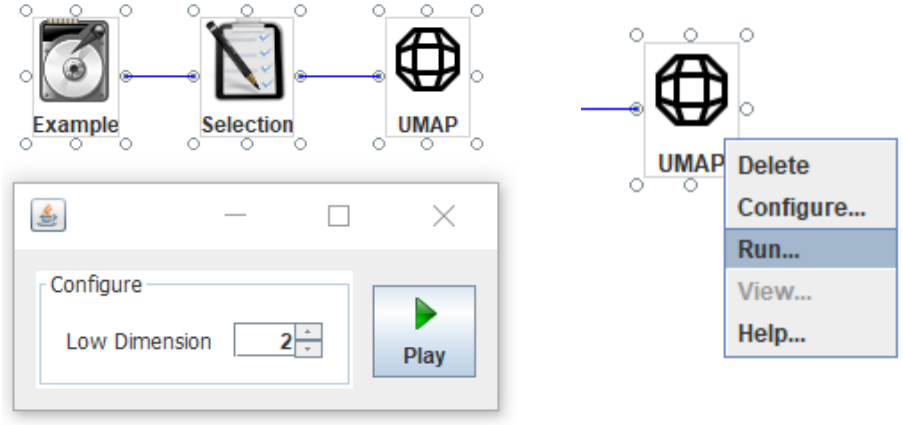
En este caso se van a cargar datos del tipo imagen Mnist o números escritos a mano, primero se cargan dichos datos al componente 'Example'.



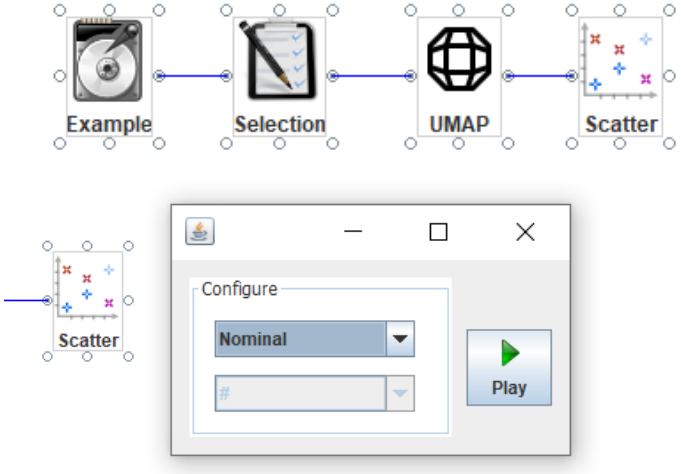
Se conecta al filtro 'Selection' y se elige la variable objetivo.



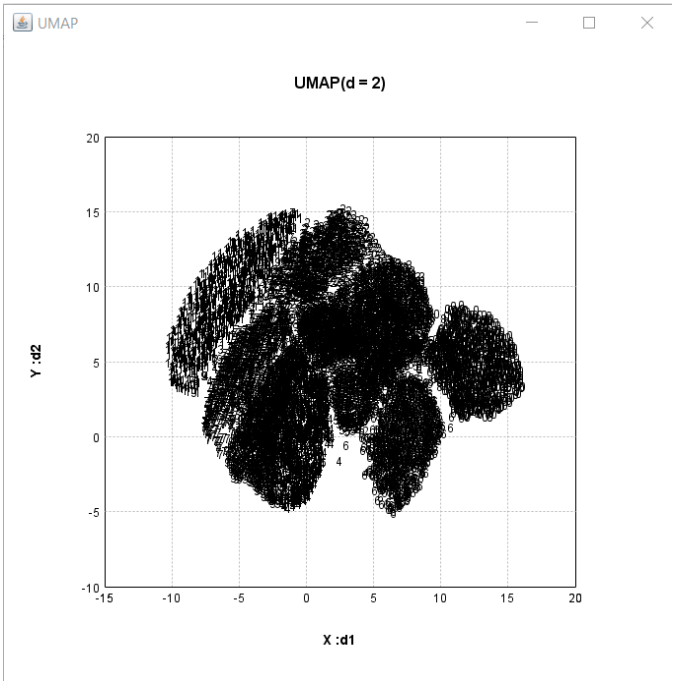
Una vez cargada la selección de datos se procede a hacer la conexión con el algoritmo reductor de dimensión.



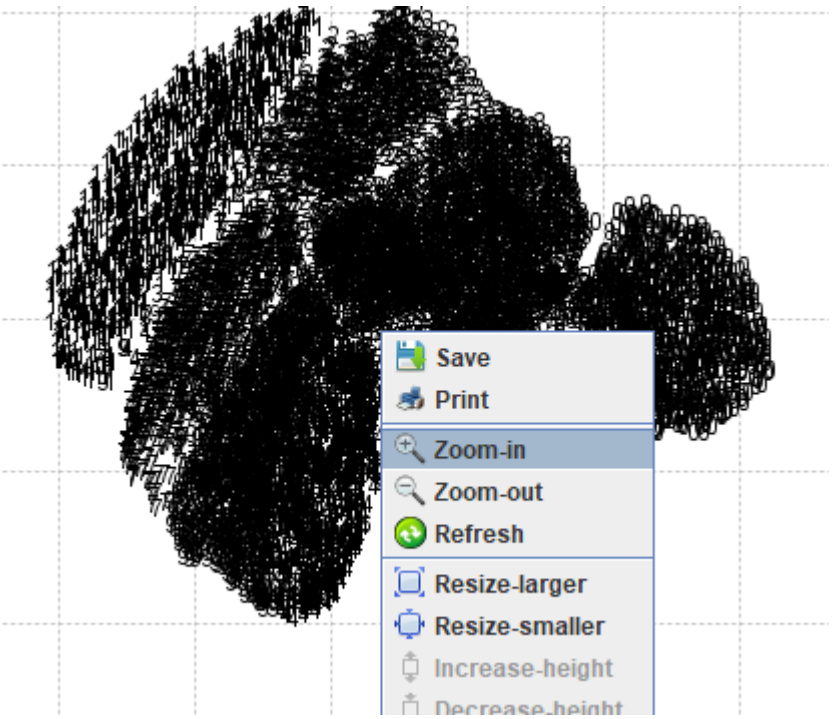
Se conecta y se configura el visualizador.



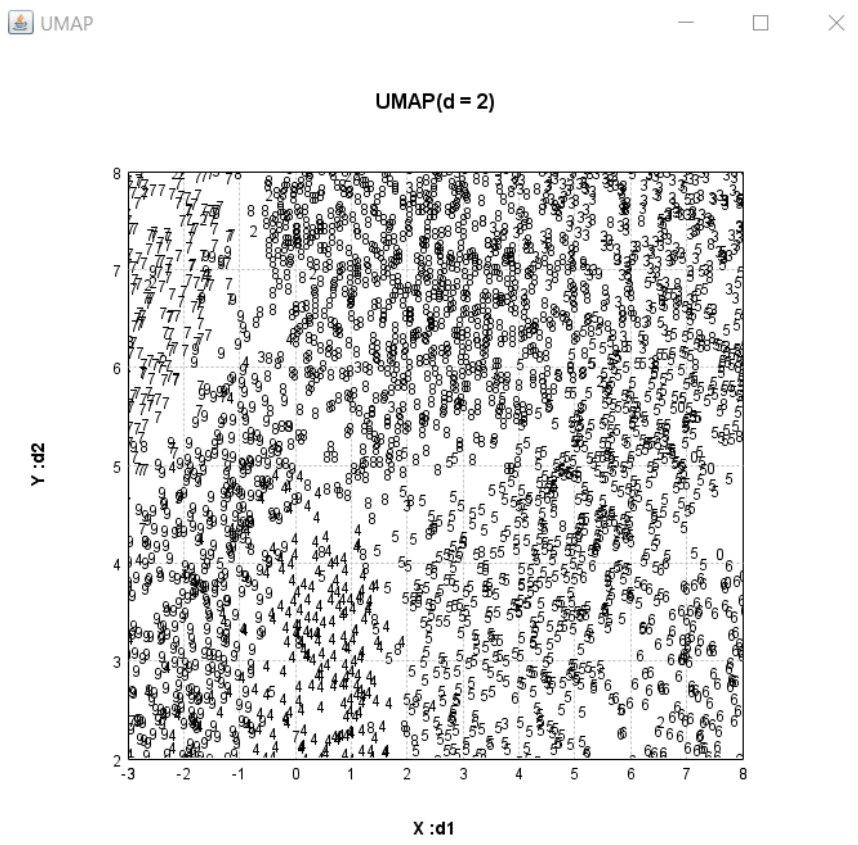
Se visualiza el resultado.



El visualizador permite hacer zoom dando click derecho en la ventana y escogiendo las opciones Zoom In o Zoom Out.

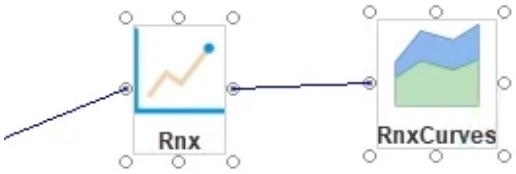


Si se acerca la gráfica se puede observar que los números contenidos han sido agrupados en sus respectivos montones según ha considerado el algoritmo.

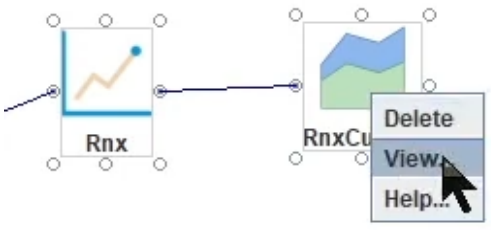


5.1.2. RnxCurves

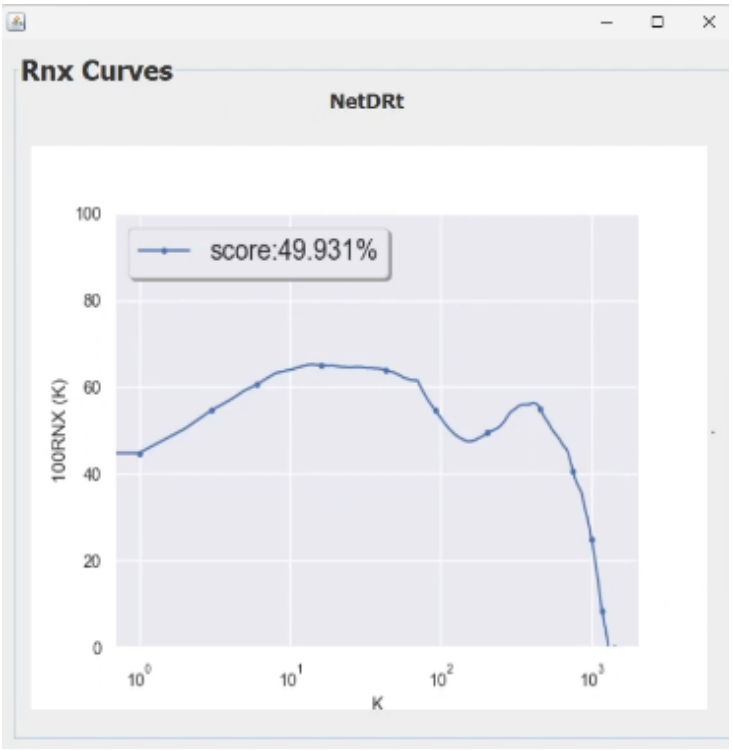
El visualizador RnxCurves está diseñado para mostrar gráficamente el resultado de evaluaciones de preservación topológica Rnx, es decir únicamente puede conectarse con dicho componente, además no cuenta con una ventana de configuración.



Para mirar su gráfica simplemente debe dar click en la opción ‘View’ del componente.

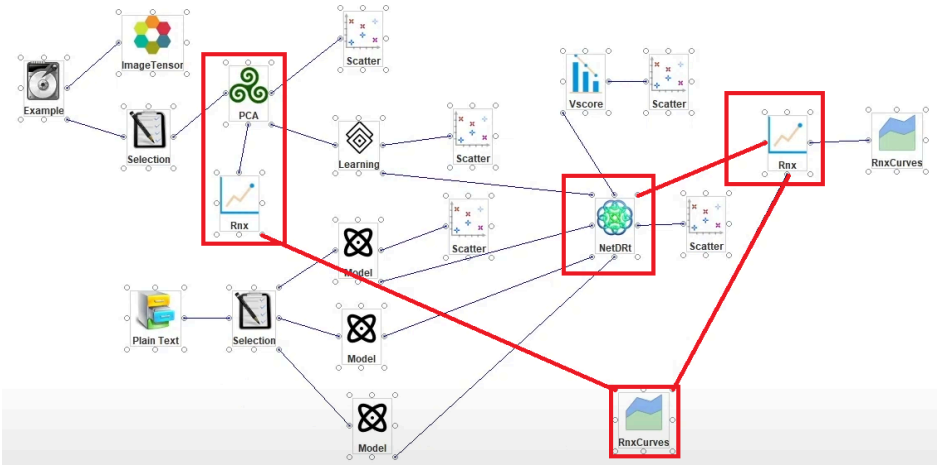


Mostrará una ventana con una gráfica como la siguiente:

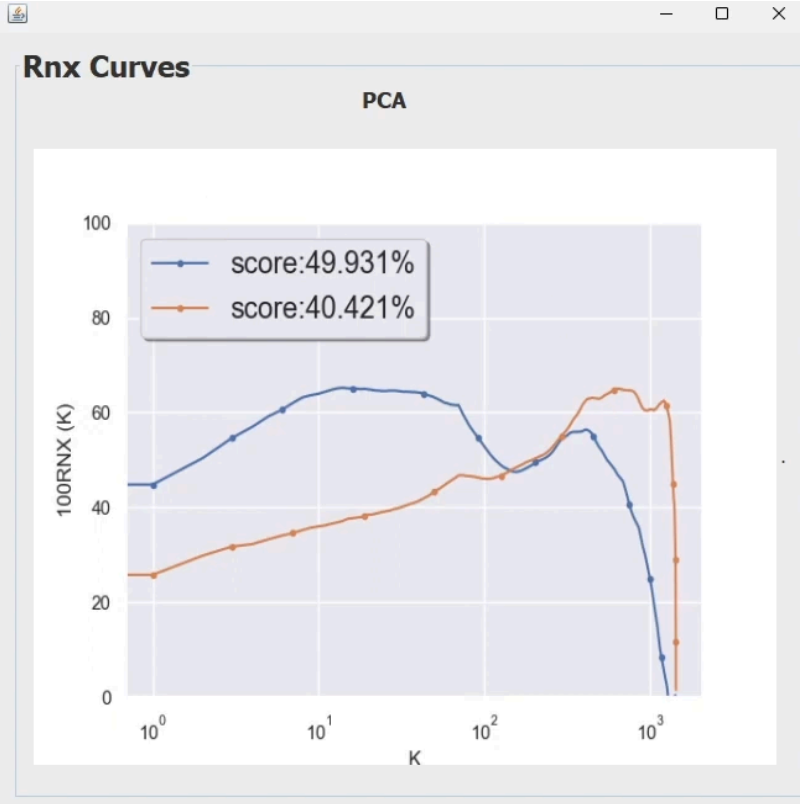


También puede recibir información de más de un componente Rnx, y por ello sirve para realizar comparaciones entre resultados de evaluaciones.

Por ejemplo si se tiene la siguiente configuración:



'RnxCurves' mostrará una gráfica comparando los dos resultados:



5.1.3. Variance

Realiza una combinación de variables que sean ortogonales, no hay dependencia en el espacio lineal.

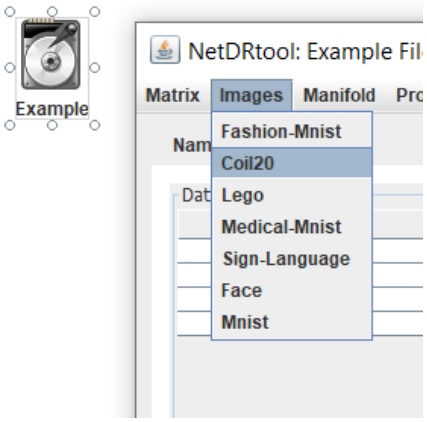


→ Icono de la herramienta

Nota: Se ha sugerido que este componente sea conectado con el algoritmo PCA, de momento no presenta funcionamiento reconocible.

5.1.4. ImageTensor

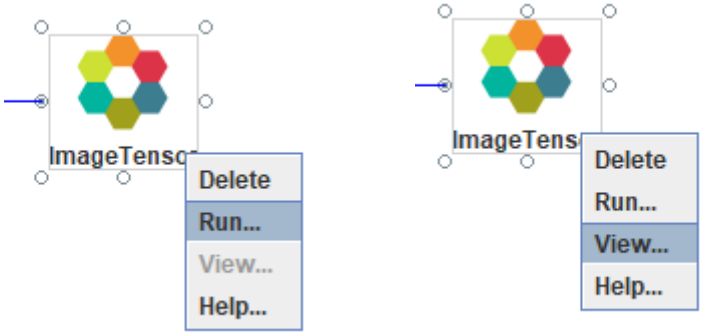
Permite visualizar tensores de imágenes, no cuenta con una ventana de configuración y para su uso los datos de ejemplo deben ser de tipo imagen:



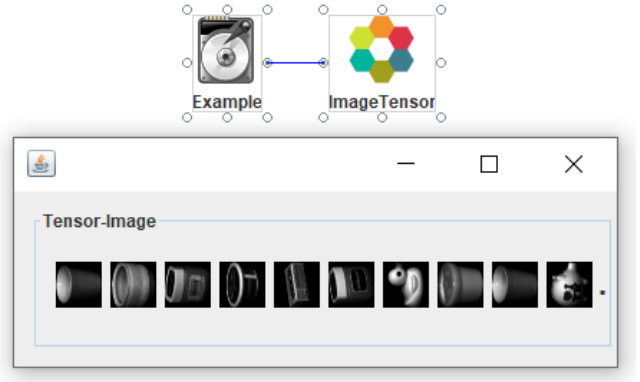
Una vez se hayan cargado los datos puede proceder a hacer la conexión con el visualizador:



Elige la opción ‘Run’ y posteriormente ‘View’.

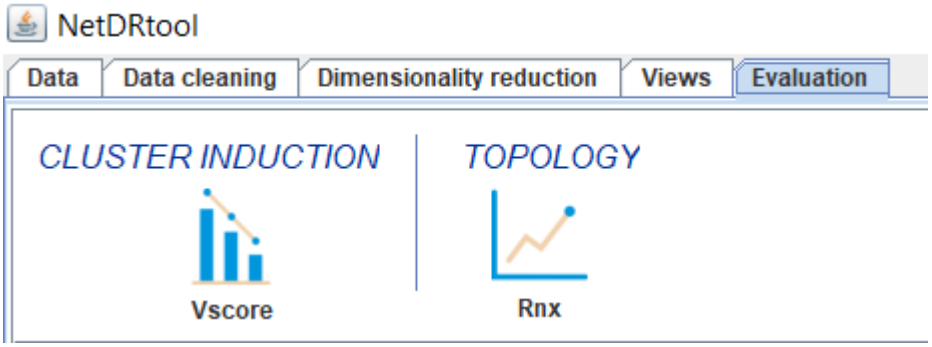


Con ello podrán visualizarse los datos de imagen.



6. Evaluation

Además de las herramientas de visualización, también se encuentran presentes las herramientas de evaluación, las cuales permiten analizar el desempeño de los modelos o algoritmos utilizados en un proyecto, para ver cuál de ellos es más eficiente.

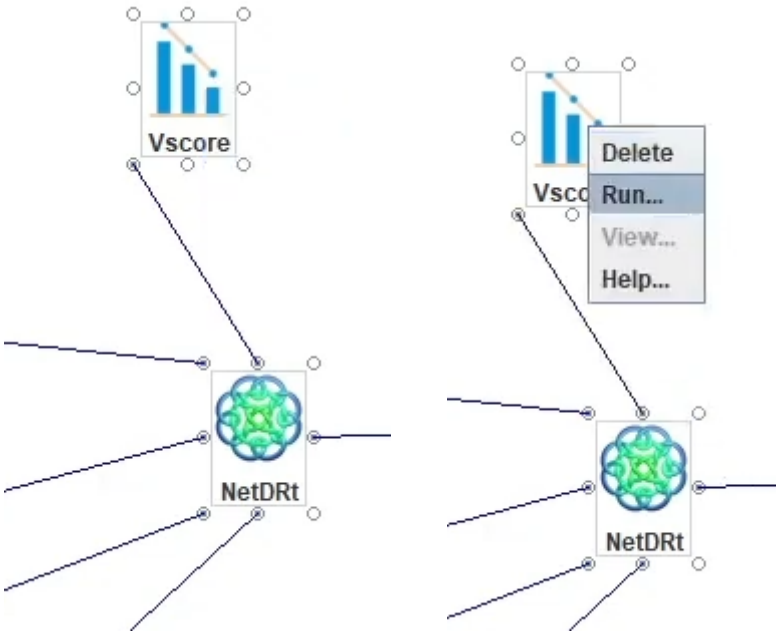


Cuenta con dos herramientas para evaluación de algoritmos y modelos:

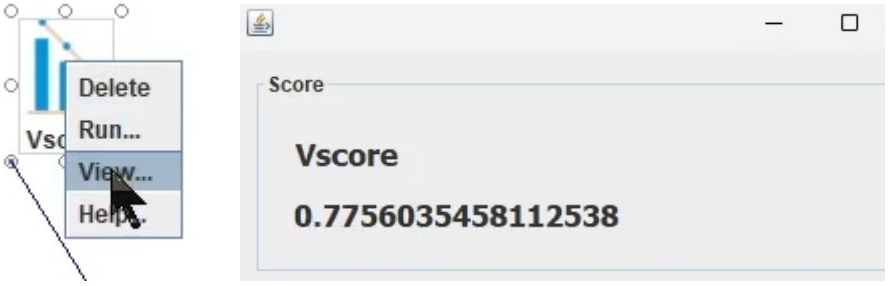
6.1. Vscore

Métrica que evalúa la calidad de los clusters (grupos de objetos similares) generados por un algoritmo o método de reducción de dimensión.

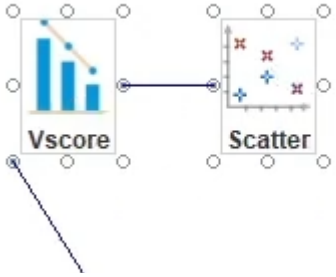
Su conexión debe realizarse con un algoritmo o modelo de reducción de dimensión, y luego simplemente se debe ejecutar la herramienta.



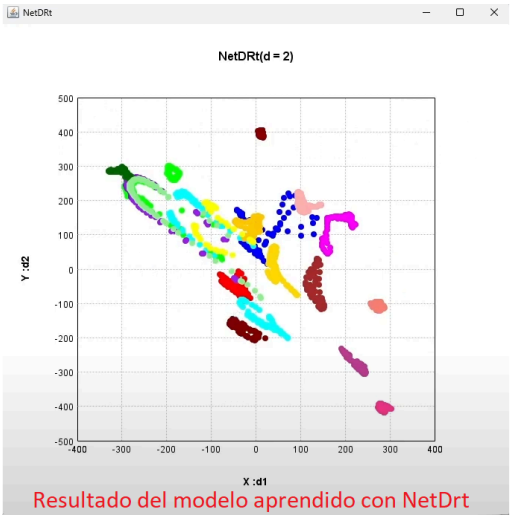
Una vez el componente termine de cargar, se puede visualizar el resultado de la evaluación:



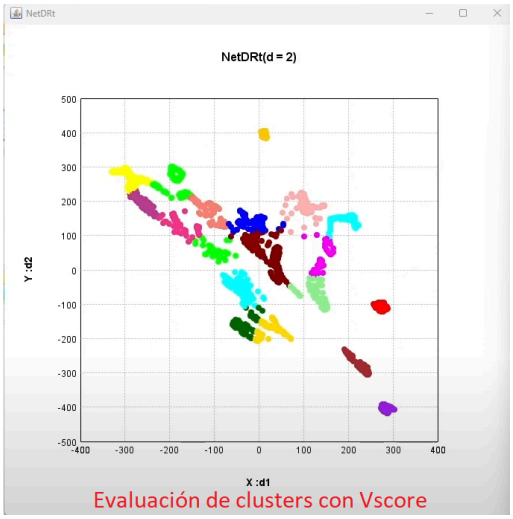
Además puede visualizarse una gráfica utilizando la herramienta de visualización ‘Scatter’.



A continuación se pueden visualizar los resultados:



Resultado del modelo aprendido con NetDrt



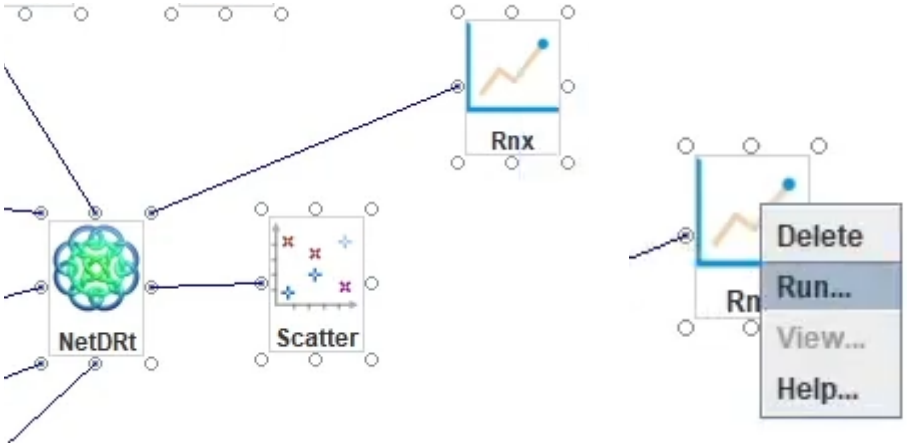
Evaluación de clusters con Vscore

Se puede observar que en el gráfico de evaluación, Vscore se encarga de hacer las respectivas divisiones por color de los clusters que ha identificado, en zonas que parecen pertenecer a datos similares.

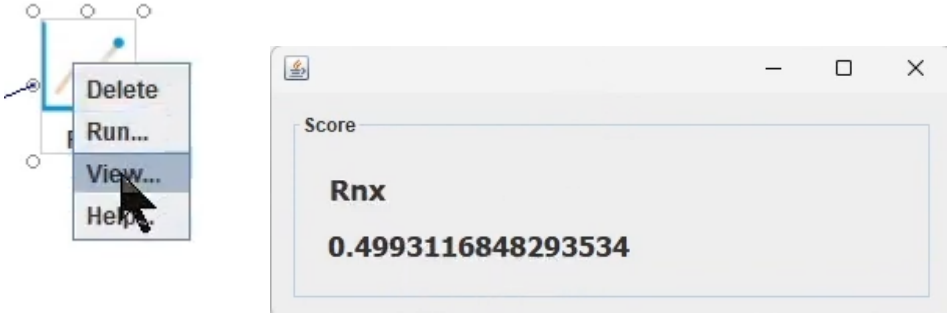
6.2. Rnx

Métrica que evalúa la preservación topológica (estructura) de las relaciones de los datos, luego de aplicar métodos o algoritmos de reducción de dimensión.

Su conexión debe realizarse con un algoritmo o modelo de reducción de dimensión, y luego simplemente se debe ejecutar la herramienta.

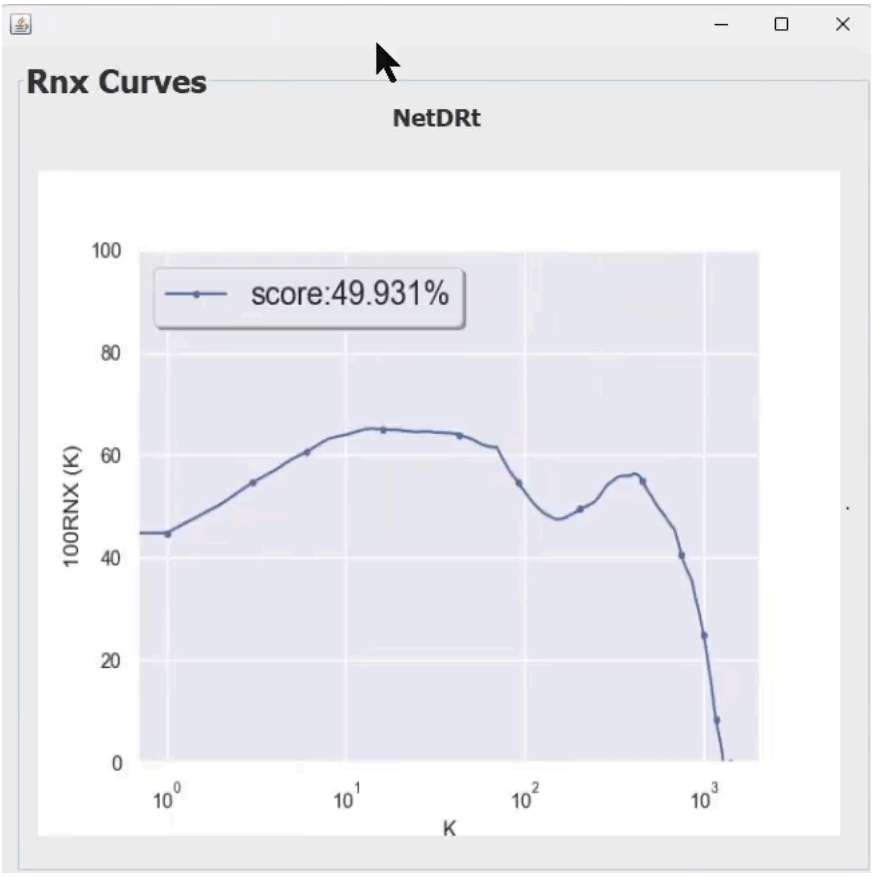


Una vez el componente termine de cargar, se puede visualizar el resultado de la evaluación:

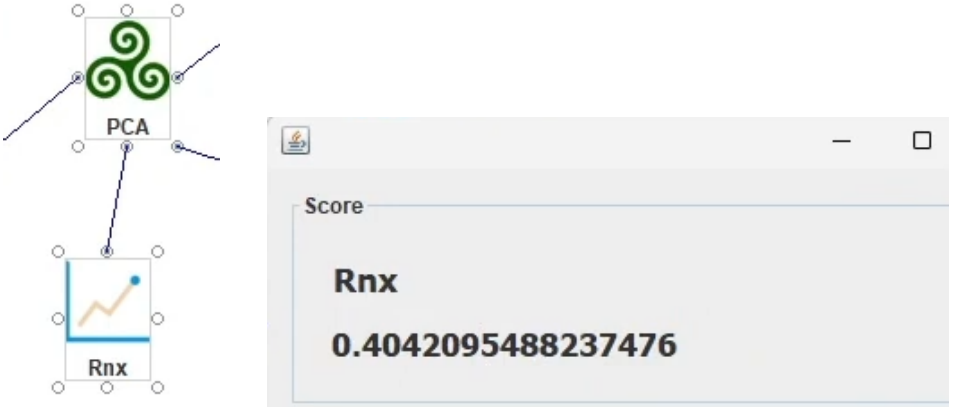


Además puede visualizarse una gráfica utilizando la herramienta de visualización ‘RnxCurves’.

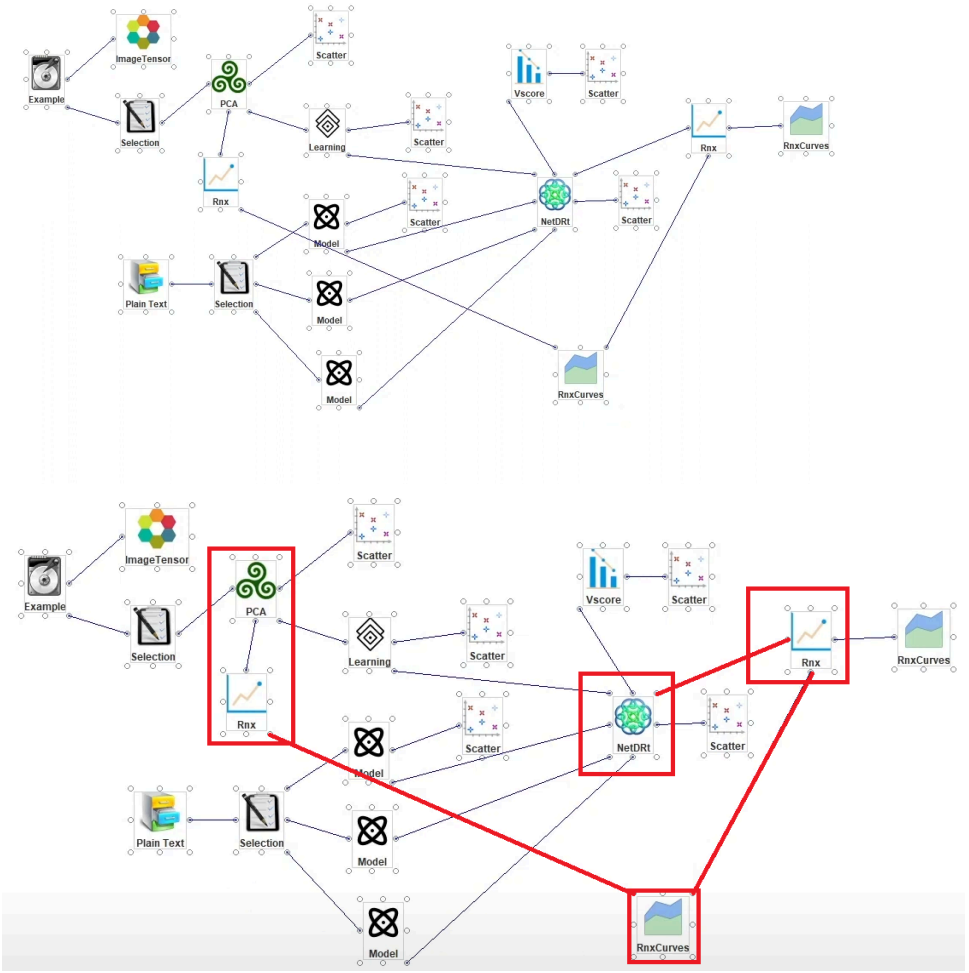




A continuación puede observarse el resultado que arroja la evaluación con uno de los algoritmos utilizados en este ejemplo.



Para realizar la comparación entre la eficiencia del algoritmo PCA y el método construido con NetDRt, se unen ambos componentes de evaluación Rnx en un mismo componente gráfico RnxCurves



El componente 'RnxCurves' mostrará la siguiente gráfica, donde puede observarse que el modelo aprendido por NetDRt es más eficiente que el algoritmo PCA según la evaluación realizada.

