**26 Mayo 2020. Knime. Redes Neuronales. Parte 2**

**Redactar un tutorial en el que detalle paso a paso (con capturas de pantalla) el desarrollo de la siguiente tarea**

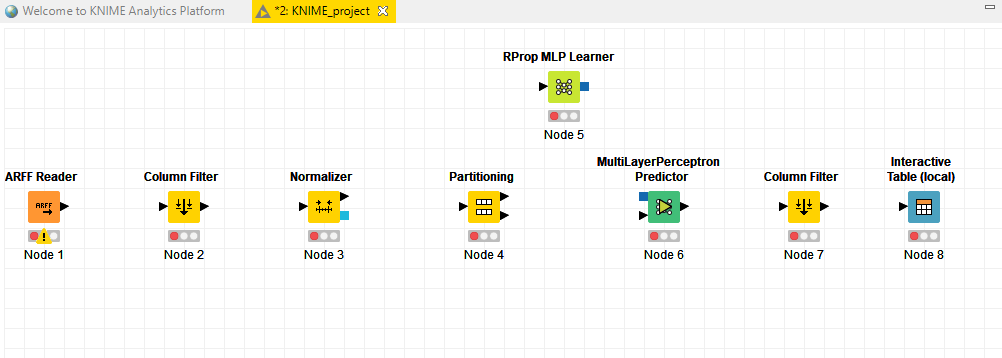
1. **Descargar el archivo empresas. arff**
2. **Crear un nodo para cargar empresas.arff**
3. **Crear un nodo Column Filter para  eliminar el atributo obs**
4. **Crear un nodo Normalizer  y normalizar con Z-'Scorre (Gaussian) los cuatro atributos cuantitativos.**
5. **Crear un nodo Partioning, para seleccionar el 80% de los datos.**
6. **Con el 80% de los datos aplicar el nodo MLPLearner**
7. **Con el 20% de los datos aplicar el nodo MLPPredictor**
8. **A partir de Predictor filtrar las columnas grupo y predicción de grupo**
9. **Crear una tabla interactiva en que se muestren el grupo real y el predicho.**

**Nota: Puede ser útil considerar el siguiente ejemplo**

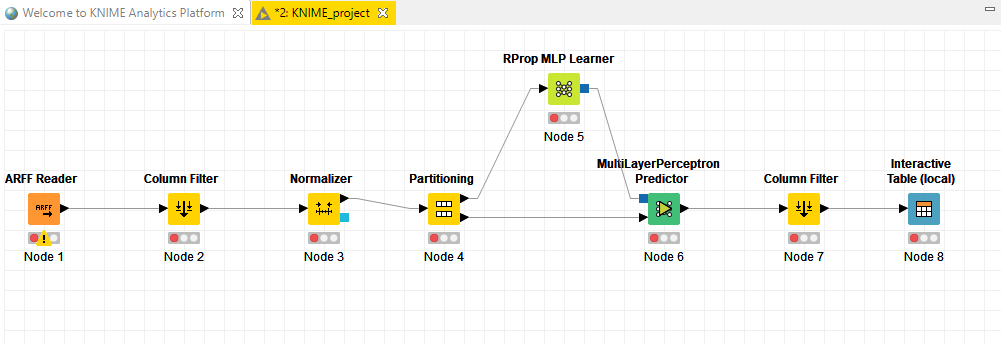
En primer lugar vamos a descargar el archivo empresas.arff.

Abrimos la herramienta Knime. Creamos un nuevo espacio de trabajo, e iremos añadiendo uno a uno los siguientes módulos en la siguiente secuencia al espacio de trabajo, poniendo los nuevos nodos en la misma línea, pero más a la derecha: ARFF Reader, Column Filter, Normalizer, Partitinoning, RProp MLP Learner, MultiLayerPerceptron Predictor, Column Filter e Interactive Table.

Los conectaremos de la siguiente manera:

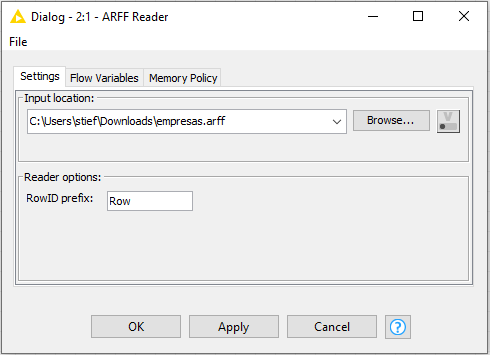


Y las conectaremos entre sí de la siguiente manera. Luego, iremos configurando uno a uno los módulos.



**2.**

Para poder leer el archivo .arff usamos el nodo ARFF Reader. Lo seleccionamos y pulsamos F6 para configurarlo. Seleccionamos el archivo empresas.arff.

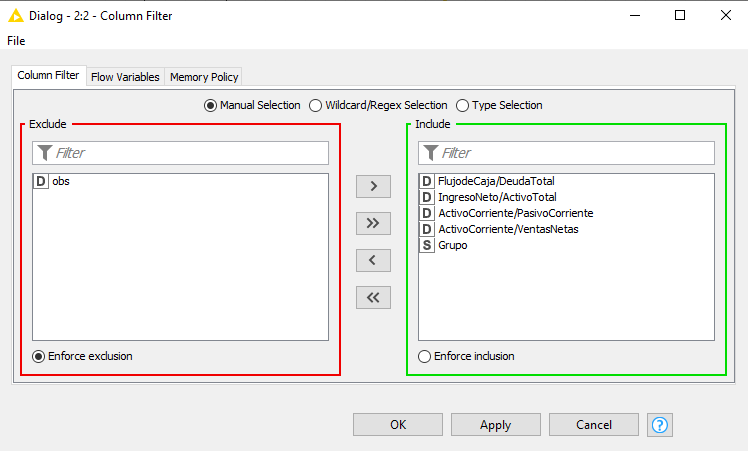


Pulsamos “Ok” y pulsamos F7 para ejecutar el módulo

**3.**

El siguiente nodo que usaremos es el nodo “Column Filter”. Lo usaremos para filtar una de los atributos, concretamente el atributo “obs”, ya que contiene información ordenada que nunca recibiríamos realmente de una empresa.

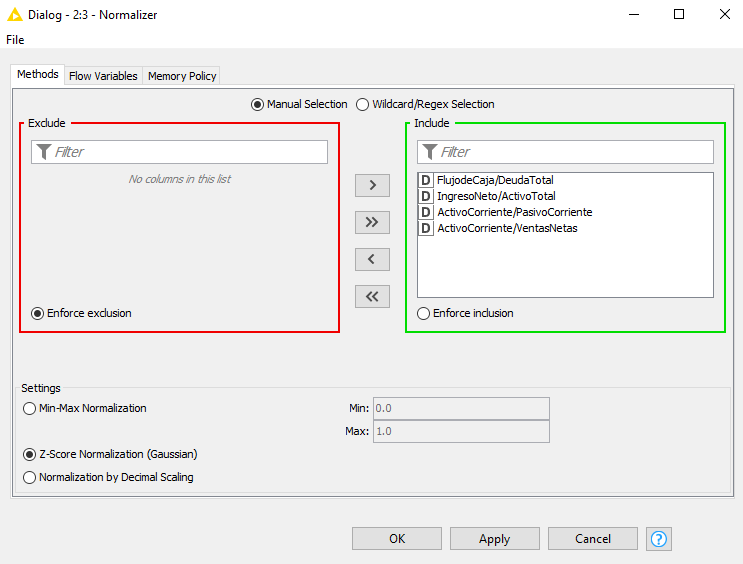
Por tanto, seleccionamos el nodo y pulsamos F6 para configurarlo. En la siguiente ventana eliminaremos el atributo “obs”.



Pulsamos “Ok” y F7 para ejecutar.

**4.**

Es hora de normalizar los datos. Para ello vamos al nodo “Normalizer” y pulsamos F6 para configurarlo.

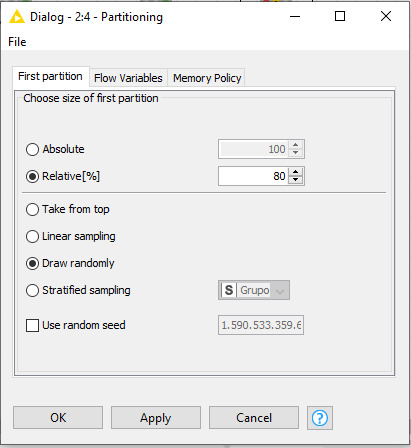


En esta ventana vamos a seleccionar la opción “Z-Score Normalization”. Observamos que están ya incluidos por defecto los 4 atributos a normalizar.

Pulsamos “Ok” y ejecutamos pulsando F7.

**5.**

Lo próximo a realizar es dividir los datos en un 80% para crear un modelo, y el 20% restante para testearlo. Para crear esta partición usaremos un nodo “Partitioning”. Con el nodo seleccionado, pulsamos F6 para configurarlos. En la siguiente ventana que nos aparecerá establecemos los siguientes parámetros y pulsamos “Ok”.

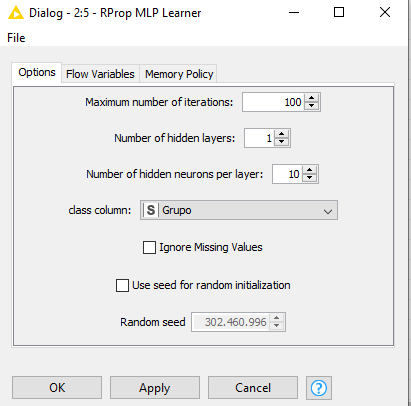


Señalamos el porcentaje de datos a dividir (80%) y que estos se elijan aleatoriamente (“Draw randomly”).

Con el nodo seleccionado pulsamos F7 para ejecutar la partición.

**6.**

El 80% de los datos lo vamos a usar para crear un modelo mediante el nodo de aprendizaje supervisado “MLPLearner”. Con el nodo seleccionado, pulsamos F6 para configurarlo. En la ventana emergente, nos aseguramos que la variable objetivo sea “Grupo”.



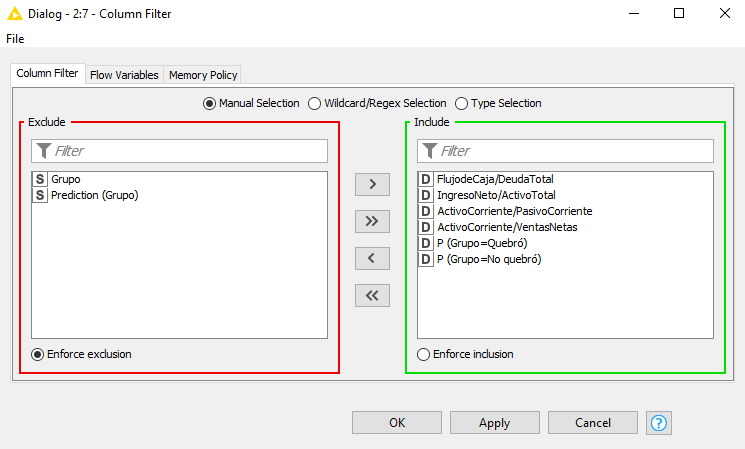
Pulsamos “Ok” y F7 para ejecutar el nodo.

**7.**

El 20% de los datos restantes los vamos a usar para testear el modelo creado anteriormente. El procedimiento es simple, solo tenemos que conectar como indicamos al principio del documento los nodos, y con el nodo “MLPPredictor” seleccionado, pulsar F7 para ejecutarlo.

**8.**

Vamos a hacer uso de otro nodo “Column Filter”, en esta ocasión para filtrar las columnas grupo y predicción de grupo. Para ello, con el nodo seleccionado, pulsamos F6 para abrir el menú de configuración:



Aquí seleccionaremos la variable grupo y su predicción y las eliminaremos. Las podemos identificar rápidamente porque ambos atributos son de tipo “String”.

Tras hacer esto, pulsamos “Ok” y con el nodo seleccionado, pulsamos F7 para ejecutarlo.

**9.**

Para terminar, usaremos el nodo “Interactive Table” para crear una tabla interactiva en que se muestren el grupo real y el predicho. Para ello seleccionamos el nodo y pulsamos F7 para ejecutarlo. Haciendo click derecho sobre él, seleccionamos la opción “View: Table View”.

