## Modelo para averiguar que medicamento es más efectivo

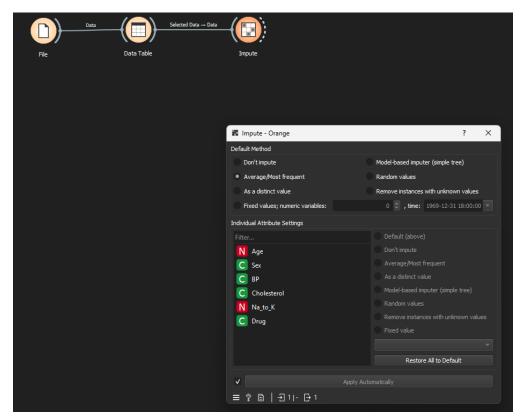
En las últimas semanas se han recopilado los siguientes datos de pruebas de medicamentos con algunos pacientes con la misma enfermedad, hay 5 tipos diferentes de medicamentos, (A, B, C) medicamentos de procedencia nacional y (X, Y) medicamentos de procedencia extranjera.

Construyendo el modelo mediante "Orange" para predecir que medicamento es mejor para cada paciente y sus síntomas (Presión, Nivel de colesterol, Relación de niveles de potasio y sodio en la sangre del paciente).

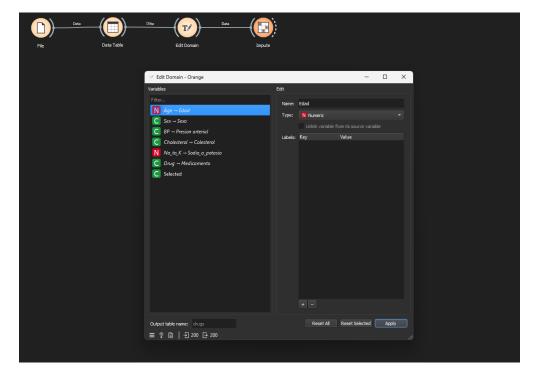
Los datos obtenidos son los siguientes:

Age	Sex	ВР	Cholesterol	Na_to_K	Drug	
23	F	HIGH	HIGH	25.355	drugY	
47	М	LOW	HIGH	13.093	drugC	
47	М	LOW	HIGH	10.114	drugC	
28	F	NORMAL	HIGH	7.798	drugX	
61	F	LOW	HIGH	18.043	drugY	
22	F	NORMAL	HIGH	8.607	drugX	
49	F	NORMAL	HIGH	16.275	drugY	
41	М	LOW	HIGH	11.037	drugC	

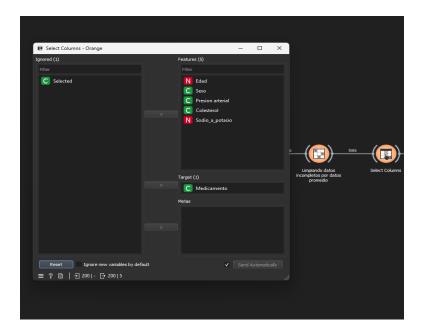
Vamos a corregir datos en caso de que falten valores



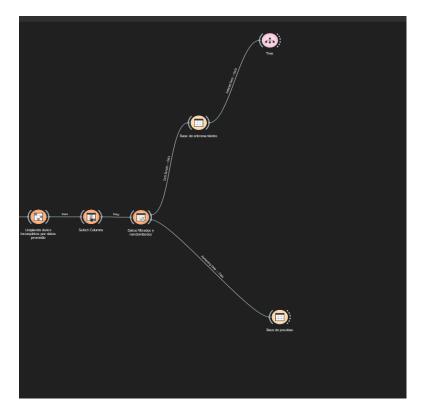
Ahora vamos a cambiar el nombre de las columnas para adaptarlas a nuestro idioma para una mejor comprensión de los datos



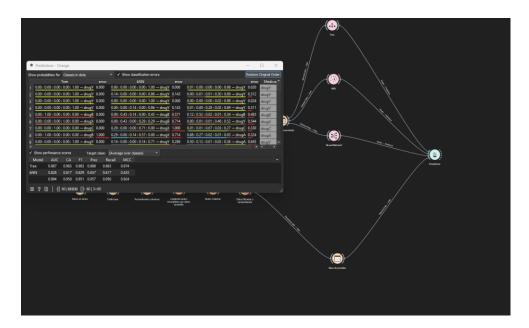
Creamos una base de entrenamiento para nuestros modelos predictivos, primero seleccionamos la columna de "medicamento" para que nuestro modelo tome en cuenta esta columna de predicción.



Ahora para poder hacer nuestro modelo tomamos 70% de nuestros valores ya filtrados y los separamos en una tabla base y los valores restantes los dejamos en una tabla separada para las pruebas



Para este modelo vamos a usar 3 diferentes opciones, el de Árbol de decisiones, KNN y redes neuronales, para ver cuál de los 3 nos da mejores resultados



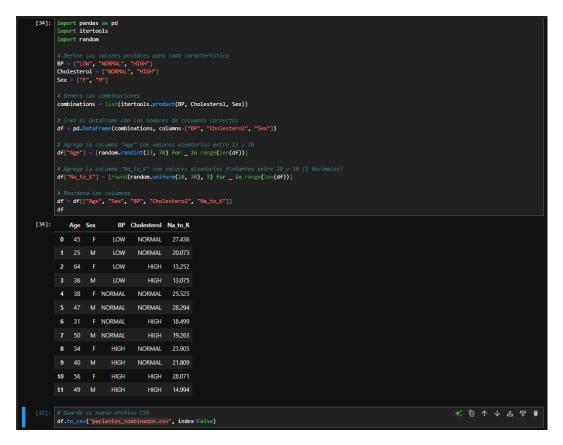
Como podemos notar el modelo de Árbol de decisiones nos da mejores resultados con una certeza de 98%, seguido de redes neuronales con 95% y de KNN con un 61% vamos a intentar mover los parametros de KNN para ver si podemos lograr un mejor resultado



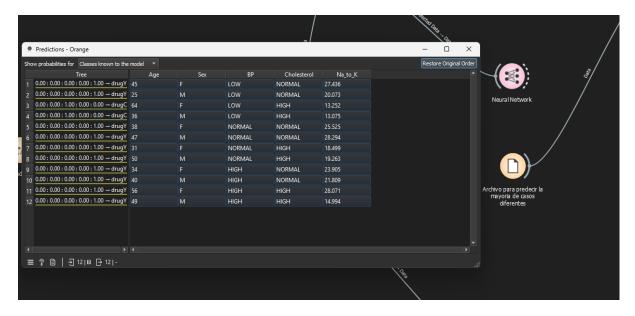
Aun moviendo los parámetros de KNN para lograr un mejor porcentaje 80% de efectividad con la predicción no logra ser suficiente para alcanzar los otros 2 modelos que son los mejores en estas situaciones



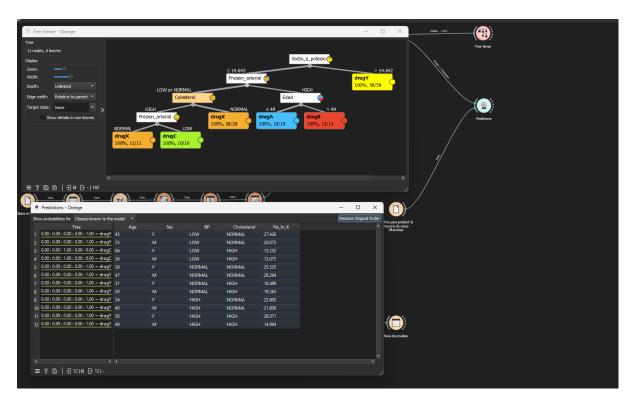
Ahora para comprobar vamos a crear una tabla con distintas posibilidades de características en pacientes y que nuestro modelo de "Tree" se encargue de predecir los medicamentos que necesita usar



Conectamos el nuevo archivo cvs al modelo predictivo para que nos diga qué tipo de medicamento es mejor para cada característica diferente del paciente



Y comprobamos con nuestro árbol predictivo las decisiones tomadas de nuestro modelo

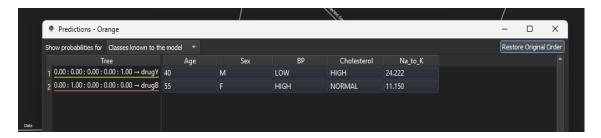


Ahora vamos a crear el cvs con las características en específico que se nos solicitó:

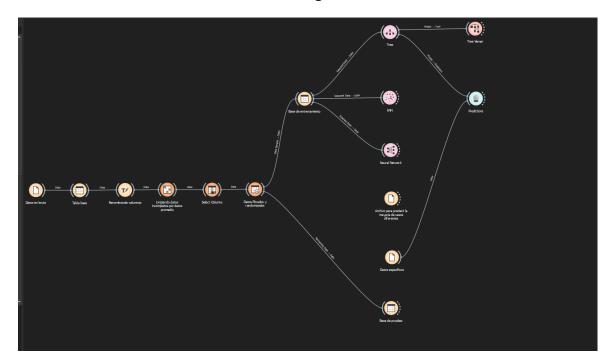
Age	Sex	<b>b</b> p	chol	esterol	na_to_k
40	М	LOW	HIGH	24.222	2
55	F		HIGH	NORMA	L. 11.150

Lo que nos da como resultado para el primero el medicamento "Y"

Y el segundo como medicamento "B"



Como resultado nuestro modelo se vería algo así



Dejo el código usado para este modelo, para hacer la base de casos combinados o por lo menos con la mayoría de combinaciones que podía haber:

import pandas as pd

import itertools

import random

```
BP = ["LOW", "NORMAL", "HIGH"] Cholesterol = ["NORMAL", "HIGH"] Sex = ["F", "M"] combinations = list(itertools.product(BP, Cholesterol, Sex))

df = pd.DataFrame(combinations, columns=["BP", "Cholesterol", "Sex"])

df["Age"] = [random.randint(23, 70) for _ in range(len(df))]
```

df["Na\_to\_K"] = [round(random.uniform(10, 30), 3) for \_ in range(len(df))]

df = df[["Age", "Sex", "BP", "Cholesterol", "Na\_to\_K"]] df