

GONZALO PAZ

Link Repositorio [git@github.com:Cdogonza/InteligenciaArtificial.git](https://github.com:Cdogonza/InteligenciaArtificial.git)

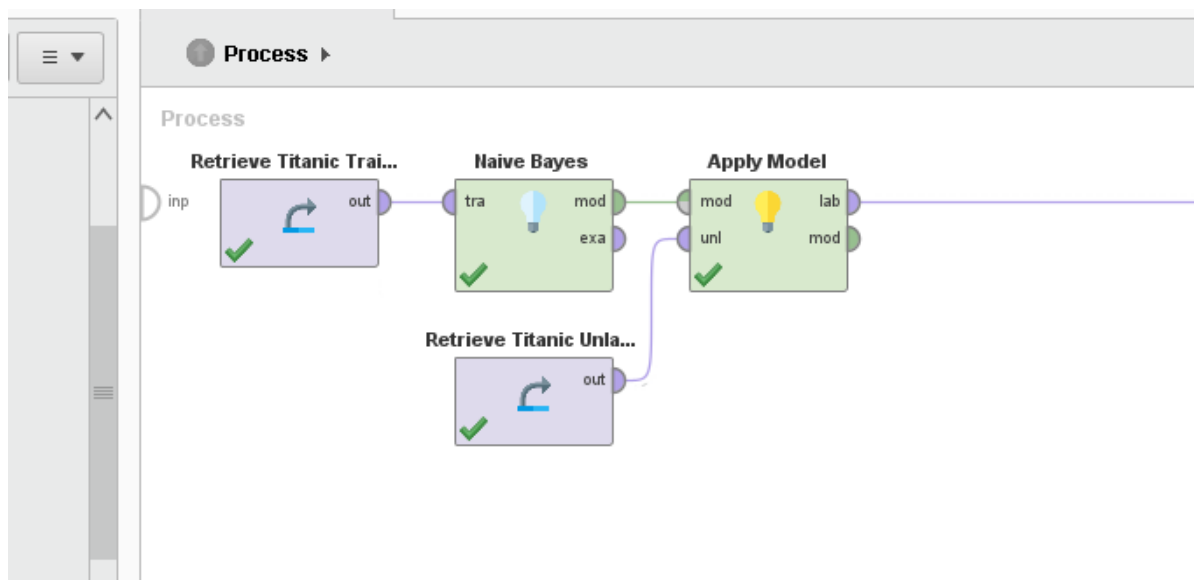
TUTORIAL "MODELING"

En este tutorial, crearemos tres modelos de clasificación diferentes para nuestros datos del Titanic: un árbol de decisión, un conjunto de reglas y un modelo Bayes. En el resultado al correr el modelo, se puede observar que el tamaño de la familia importa mas que la clase pasajero para las mujeres.

El utilizar un arbol de decición, nos permite visualizar de forma gráfica todos los posibles resultados y su costo de una determinada decisión. A su vez el operador de clasificación Naive Bayes nos proporciona las características del teorema de Naive Bayes basado en que no existe relacion entre los atributos, o sea se consideran características independientes, simplificando la computación y mostrando la probabilidad de supervivencia, que dado este ejemplo de dataSet tienen mayor probabilidad de sobrevivir las mujeres frente a los hombres.

TUTORIAL "SCORING"

El objetivo de este tutorial es mostrar cómo podemos utilizar la información creada en el tutorial anterior para predecir resultados futuros. Utilizando el método Naive Bayes para predecir si la clase "sobrevivido" de cada pasajero y calcular sus respectivas coincidencias. Este modelo para generar predicciones para nuevos puntos de datos se denomina puntuacion. Utilizaremos posterior al operador Naive Bayes uno llamado Aplicar modelo, que lo que nos propociona las predicones para un nuevo conjunto de datos sin etiquetar. En este modelo podemos observar que el operador aapply model recibe datos sin etiquetar como entrada y genera un conjunto de datos con una etiqueta: las predicciones realizadas por el modelo.



TUTORIAL "TEST SPLIT VALIDATION"

En este caso se utiliza el dataset "Titanic Training" conectado a un operador "Split data" configurado en un 70% para entrenamiento y un 30% para validación, y conectado a un modelo de predicción "Naive Bayes" para predecir la supervivencia de los pasajeros del Titanic. El resultado del "Apply Model" muestra la predicción de supervivencia de los pasajeros del Titanic, con un porcentaje de confianza de la predicción que indica que tan seguro es el resultado de la misma. Se observan casos donde la predicción es correcta y otros donde la predicción es incorrecta, ejemplo de esto la predicción de "no" supervivencia de 29 casos siendo que realmente si sobrevivieron y analogamente 25 casos con predicción de "si" supervivencia siendo que realmente no sobrevivieron. La precisión de esta predicción se muestra en el entorno de los 80,36% de aciertos, lo cual es un buen resultado para este tipo de predicciones.

accuracy: 80.36%

	true Yes	true No	class precision
pred. Yes	76	25	75.25%
pred. No	29	145	83.33%
class recall	72.38%	85.29%	