



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Taller: Clase 13

Estudiante: Danny Sebastián Díaz Padilla

Fecha: 20/11/19

Tema: Análisis del template 'Predictive Maintenance' de RapidMiner



Predictive Maintenance

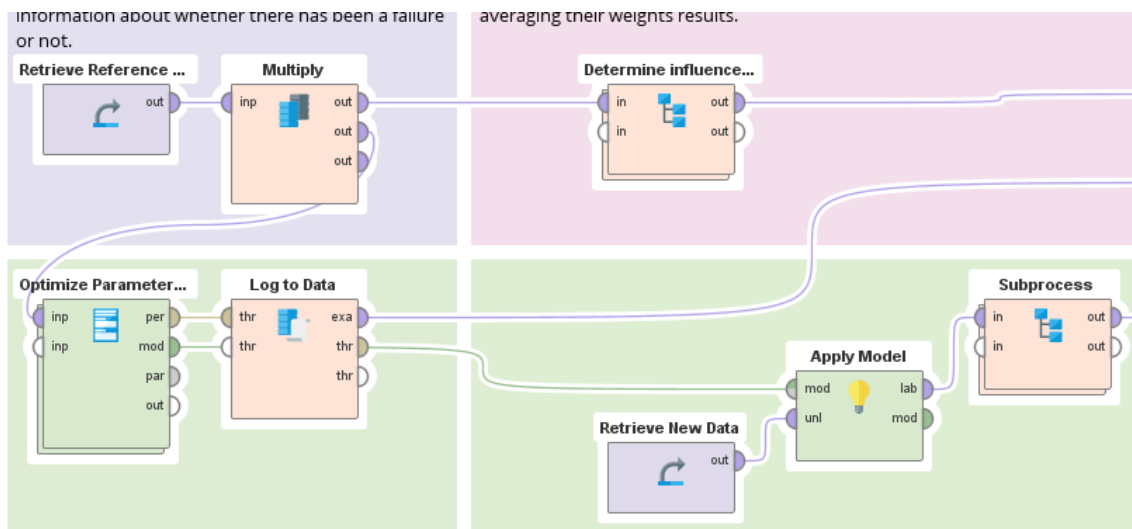
Model equipment failures to schedule maintenance pre-emptively

A. Proceso

Mantenimiento Predictivo

Modela fallas de equipos basadas en observaciones de referencias de datos de ejecuciones y fallas de máquinas anteriores. El modelo se aplica a situaciones actuales para anticipar las fallas de la máquina y programar el mantenimiento de manera preventiva.

El esquema del proceso completo es el siguiente:

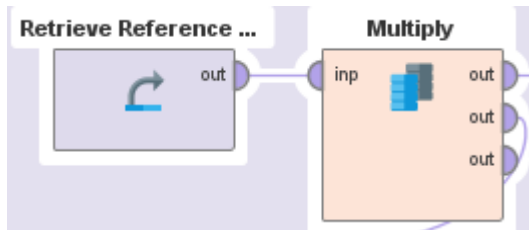




ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

1. Primer paso:

Se carga los datos de las ejecuciones anteriores de la máquina, etiquetadas con información sobre si ha habido una falla o no.



Mediante el operador '**Retrieve**' se carga la base de datos con la información de las máquinas.



Retrieve Reference Data.output (output)

Meta data: Data Table

- Source: //Samples/Templates/Predictive Maintenance/Reference Data

Number of examples = 136

27 attributes:

Note: Some of the nominal values in this set were discarded due to performance reasons. You can change this behaviour in the preferences (rapidminer.

general.md_nominal_values_limit).

Generated by: [Retrieve Reference Data.output](#) ← [Retrieve Reference Data.output](#) ← [Retrieve Reference Data.output](#)

Role	Name	Type	Range	Missings	Comment
d	Machine_ID	nominal	=[M_0001, M_000...	= 0	
label	Failure	binominal	=[no, yes]	= 0	
	Sensor_1	real	=[0.211 - 11.243]	= 0	
	Sensor_2	real	=[0.545 - 32.498]	= 0	
	Sensor_3	real	=[0.497 - 31.421]	= 0	
	Sensor_4	real	=[0.586 - 32.322]	= 0	
	Sensor_5	real	=[0.600 - 31.946]	= 0	
	Sensor_6	real	=[0.589 - 26.361]	= 0	
	Sensor_7	real	=[0.346 - 19.681]	= 0	
	Sensor_8	real	=[0.136 - 15.341]	= 0	
	Sensor_9	real	=[0.349 - 12.207]	= 0	
	Sensor_10	real	=[0.131 - 9.666]	= 0	
	Sensor_11	real	=[0.158 - 5.713]	= 0	



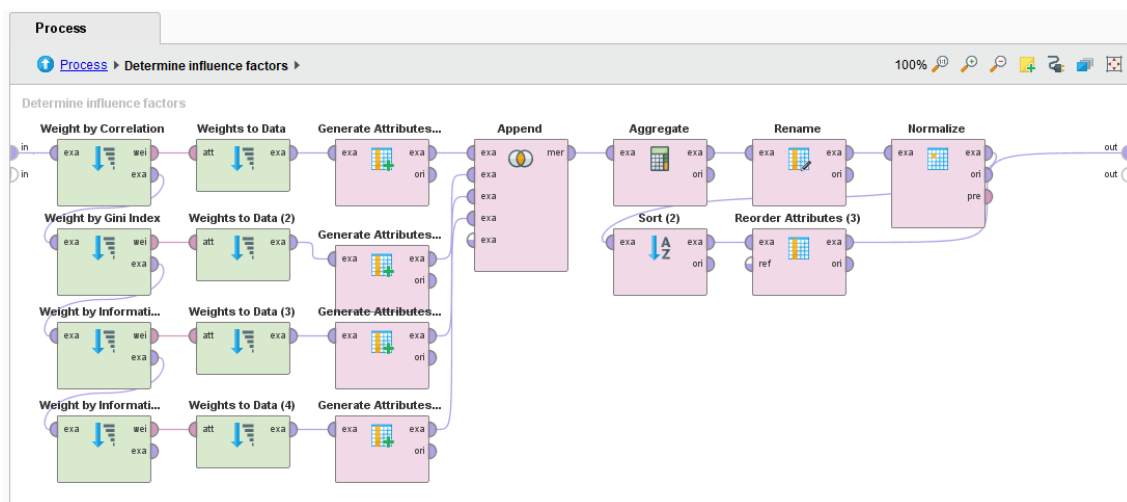
ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Se toma esta entrada y se agrega al operador '**Multiply**' el cual crea una copia de la información enviándola al paso 2 y al paso 3.

2. Segundo paso:

Se determina los factores de influencia utilizando varios algoritmos de ponderación de atributos y promediando sus resultados de ponderaciones.

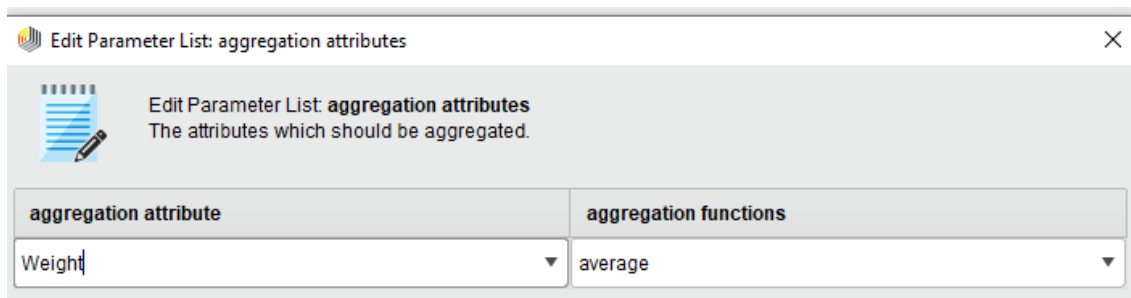
Usando la entrada del primer paso, se realiza un subproceso que se crea un peso (relevancia) por cada atributo de acuerdo a: la correlación índice Gini, ganancia de información e índice de ganancia de información.



Luego para cada peso se genera un conjunto de datos de ejemplo, el operador Generar atributos construye nuevos atributos a partir de los atributos de la entrada y constantes arbitrarias utilizando expresiones matemáticas.

Luego usando el operador 'Append' se combinan todos los conjuntos de ejemplos.

A este conjunto combinado se le aplica una función de agregación (Aggregate) para obtener el promedio de los pesos.





ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Y le da un nuevo nombre usando el operador 'Rename':

Parameters

Rename

old name: average(Weight)

new name: Importance

rename additional attributes: Edit List (0)...

Y por un segundo canal envía la información original que recibió (sin realizar la agregación) para ordenarla en forma descendente por **importancia**.

Parameters

Sort (2) (Sort)

attribute name: Importance

sorting direction: decreasing

Y finalmente redistribuir sus atributos.

Parameters

Reorder Attributes (3) (Reorder Attributes)

sort mode: user specified

attribute ordering: Define order...

use regular expressions: ☐

handle unmatched: remove

sort direction: ascending

Attribute Ordering

Attribute

Importance

La salida de este paso son estos dos conjuntos de datos ya explicados.

Determine influence factors.out 1 (out 1)

Meta data: Data Table

Number of examples = 0

2 attributes:

Generated by: [Determine influence factors.out 1](#)

Role	Name	Type	Range	Missings	Comment
	Attribute	nominal	unknown	= 0	
	Importance	real	= [0 - 1]	= 0	



Dando como salida:

3. Tercer paso:

Optimize Parameter...

inp

inp

per

mod

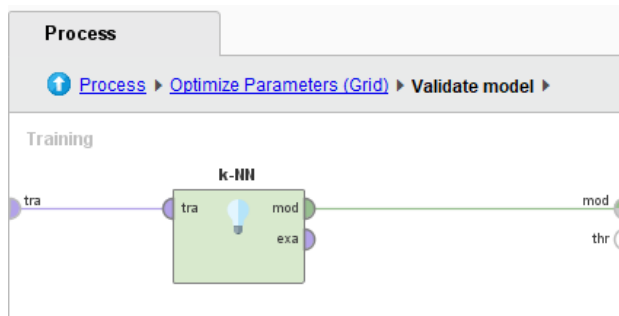
par

out



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

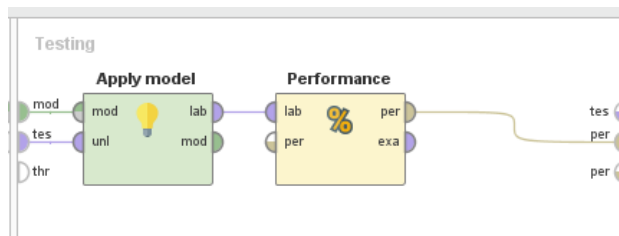
Dentro del proceso interno del subprocesso se usa el operador '**k-NN**' para realizar el entrenamiento del modelo, operador que genera un modelo de k vecinos más cercano, que se utiliza para clasificación o regresión. Se basa básicamente en la comparación de un ejemplo desconocido con los ejemplos de entrenamiento k que son los vecinos más cercanos del ejemplo desconocido. Los cuales corresponden a una parte de la entrada del primer paso.



La salida de esta parte es el modelo ya entrenado.

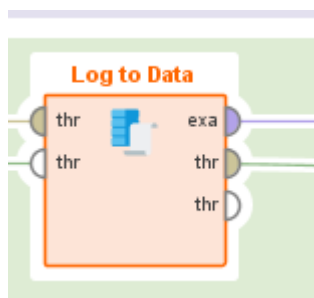
Para la parte del testeo se usa el operador '**Apply model**' para aplicar el modelo sobre un subconjunto de testeo previamente obtenido de la entrada del primer paso.

Y luego evalúa el desempeño del modelo usando el operador '**Performance**' la única salida de esta parte es las medidas de desempeño



Finalmente se utiliza el operador '**Log**' para recuperar las medidas de desempeño.

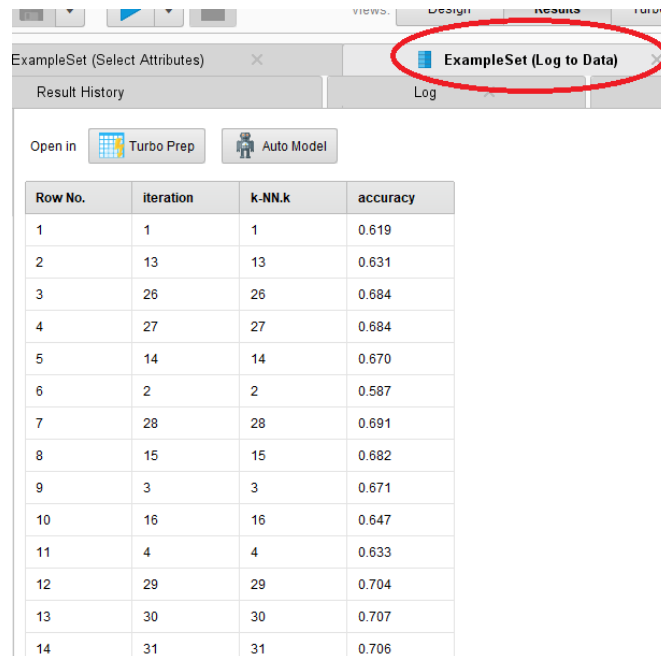
La salida de este subprocesso consiste en un modelo entrenado y en las mediciones de desempeño.





ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Finalmente usando el operador **'Log to Data'** se genera un conjunto de ejemplos usando la información del **'Log'** previamente recuperado.



Row No.	iteration	k-NN.k	accuracy
1	1	1	0.619
2	13	13	0.631
3	26	26	0.684
4	27	27	0.684
5	14	14	0.670
6	2	2	0.587
7	28	28	0.691
8	15	15	0.682
9	3	3	0.671
10	16	16	0.647
11	4	4	0.633
12	29	29	0.704
13	30	30	0.707
14	31	31	0.706

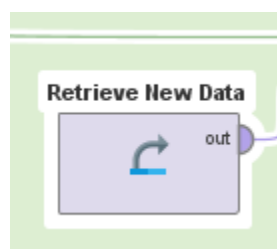
La salida de este paso consiste en la data generada desde el log y la salida del subproceso anteriormente detallado.

Además, el conjunto de ejemplos es un resultado final.

4. Cuarto paso:

Se carga nuevos datos y se aplica el modelo de falla de la máquina a las ejecuciones actuales de la máquina para predecir posibles fallas.

Primero se carga de nuevo los datos con el operador **'Retrieve'**



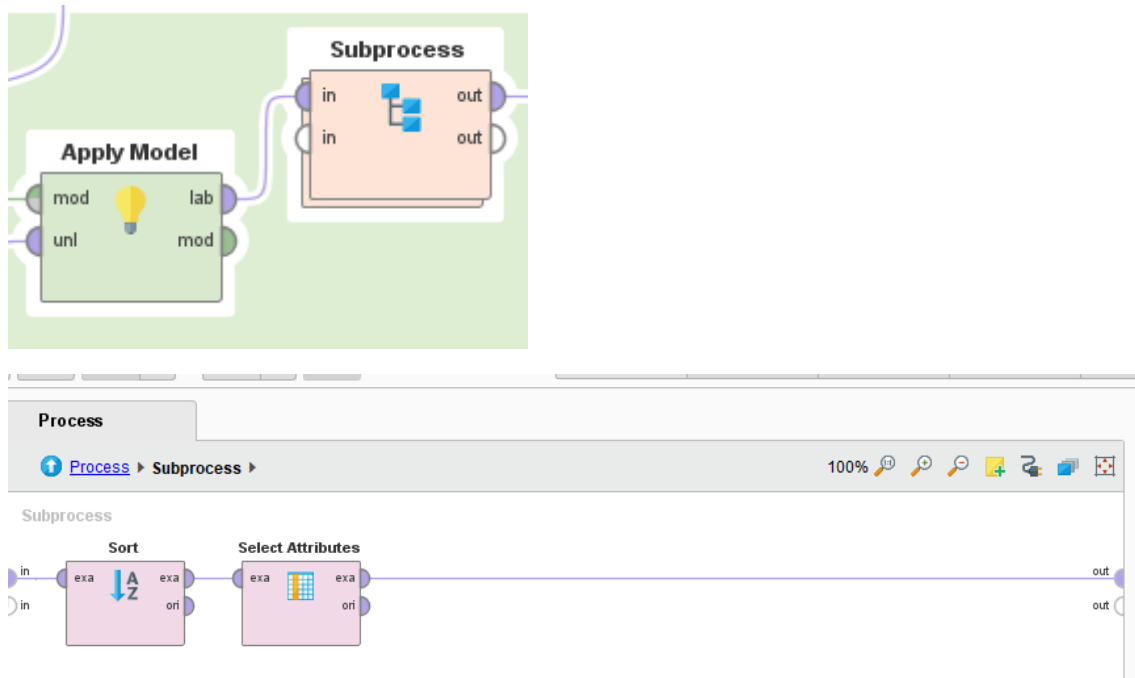
Role	Name	Type	Range	Missings
id	Machine_ID	nominal	= [M_0001, M_...	= 0
	Sensor_1	real	= [0.211 - 11.243]	= 0
	Sensor_2	real	= [0.545 - 32.498]	= 0
	Sensor_3	real	= [0.497 - 31.421]	= 0
	Sensor_4	real	= [0.586 - 32.322]	= 0
	Sensor_5	real	= [0.600 - 31.946]	= 0
	Sensor_6	real	= [0.589 - 26.361]	= 0
	Sensor_7	real	= [0.346 - 19.681]	= 0
	Sensor_8	real	= [0.136 - 15.341]	= 0
	Sensor_9	real	= [0.349 - 12.207]	= 0
	Sensor_10	real	= [0.131 - 9.666]	= 0
	Sensor_11	real	= [0.158 - 5.713]	= 0
	Sensor_12	real	= [0.052 - 3.323]	= 0
	Sensor_13	real	= [0.014 - 1.968]	= 0
	Sensor_14	real	= [0.039 - 1.404]	= 0
	Sensor_15	real	= [0.007 - 3.139]	= 0



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Se aplica el modelo usando '**Apply model**' al conjunto de datos generados desde el Log utilizando el modelo generado en el tercer paso.

Y luego por medio de un subprocess se trata la data generada.



Ordenando a los datos con '**Sort**' por confiabilidad de forma decreciente

Parameters window for the 'Sort' step. The 'attribute name' is set to 'confidence(yes)' and the 'sorting direction' is set to 'decreasing'.

Parameter	Value
attribute name	confidence(yes)
sorting direction	decreasing

Y seleccionando todos los atributos utilizando '**Select Attributes**':

Parameters window for the 'Select Attributes' step. The 'attribute filter type' is set to 'all', 'invert selection' is checked, and 'include special attributes' is unchecked.

Parameter	Value
attribute filter type	all
invert selection	<input checked="" type="checkbox"/>
include special attributes	<input type="checkbox"/>



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

Finalmente, la salida de las predicciones de fallas

views: Design Results

ExampleSet (Select Attributes) ExampleSet (Results)

Result History Log

Open in Turbo Prep Auto Model

Row No.	Machine_ID	prediction(F...	confidence(...	confidence(...
1	M_0223	yes	0.100	0.900
2	M_0239	yes	0.100	0.900
3	M_0176	yes	0.133	0.867
4	M_0194	yes	0.133	0.867
5	M_0221	yes	0.133	0.867
6	M_0225	yes	0.133	0.867
7	M_0240	yes	0.133	0.867
8	M_0250	yes	0.133	0.867
9	M_0166	yes	0.167	0.833
10	M_0229	yes	0.167	0.833
11	M_0271	yes	0.167	0.833
12	M_0137	yes	0.200	0.800
13	M_0140	yes	0.200	0.800
14	M_0151	yes	0.200	0.800

B. Resultados estadísticos

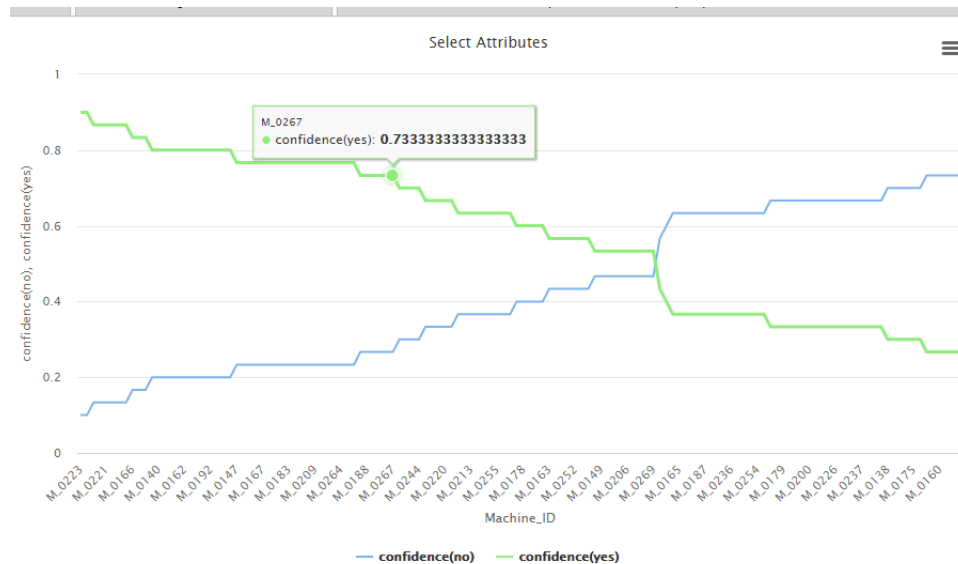
	Name	Type	Missing	Statistics	Filter (4 / 4 attributes):
Data	Machine_ID	Nominal	0	Least M_0136 (0) Most M_0137 (1)	Search for Attributes
Statistics	Prediction prediction(Failure)	Binominal	0	Least no (47) Most yes (89)	Values yes (89), no (47) [270 more]
Visualizations	Confidence_no confidence(no)	Real	0	Min 0.100 Max 0.733 Average 0.421 Deviation 0.199	
Annotations	Confidence_yes confidence(yes)	Real	0	Min 0.267 Max 0.900 Average 0.579 Deviation 0.199	



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA DE SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

C. Visualizaciones

En este caso se determina que hay un cambio brusco cuando la confiabilidad está alrededor de 0.5.



Y que el conjunto de datos poseía más información sobre máquinas que efectivamente han fallado.

