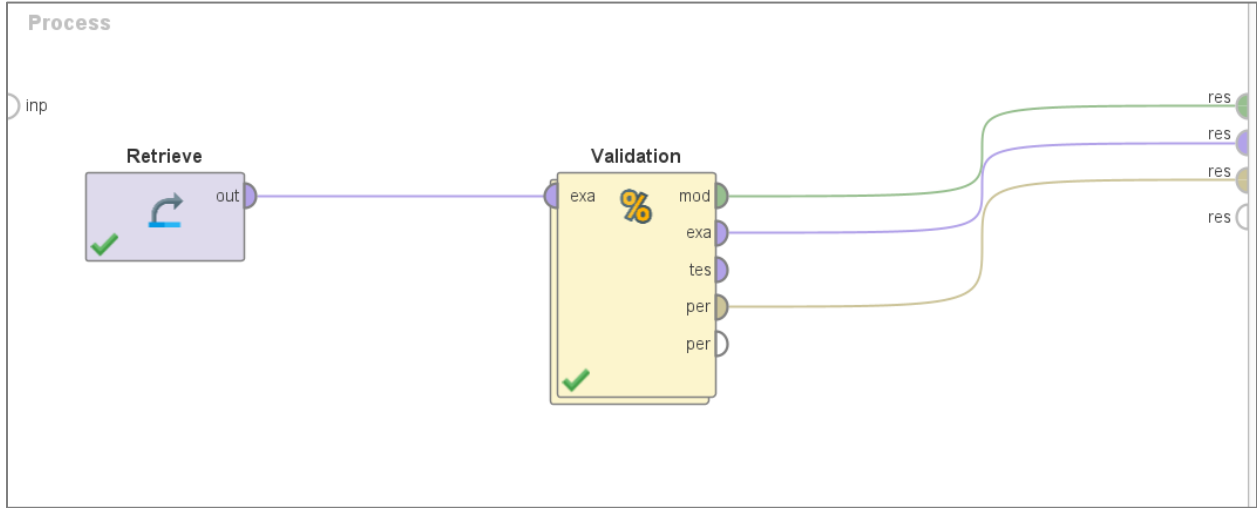
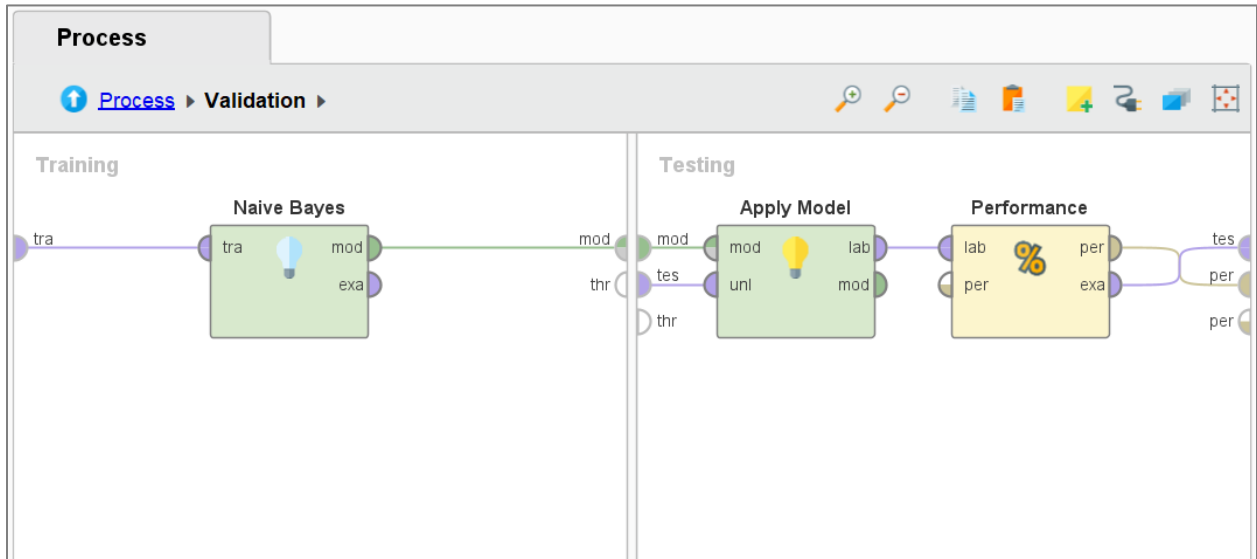


Lap 5

1. Retrieve ข้อมูล tennis.csv



2. วิเคราะห์ข้อมูลด้วย Operator Naïve Bayes



ผลลัพธ์ดังภาพ

accuracy: 45.00% +/- 49.72% (micro average: 50.00%)

	true No	true Yes	class precision
pred. No	2	4	33.33%
pred. Yes	2	4	66.67%
class recall	50.00%	50.00%	

SimpleDistribution

Distribution model for label attribute play

Class No (0.333)

5 distributions

Class Yes (0.667)

5 distributions

3. ทดสอบ Model ด้วย Split Validation โดยแบ่งข้อมูลชุดฝึก/ทดสอบ 80%/20%

Process

Process >

Validation split

Retrieve

Parameters

Validation split (Split Validation)

split: relative

split ratio: 0.8

sampling type: automatic

use local random seed

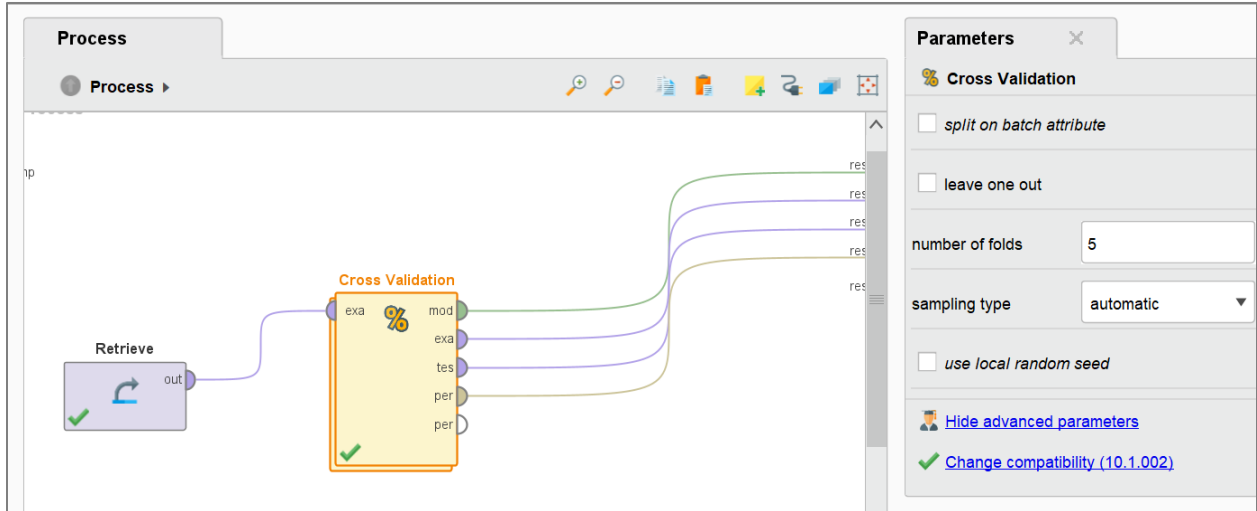
[Hide advanced parameters](#)

[Change compatibility \(10.1.002\)](#)

ผลลัพธ์ดังภาพ

accuracy: 66.67%			
	true No	true Yes	class precision
pred. No	0	0	0.00%
pred. Yes	1	2	66.67%
class recall	0.00%	100.00%	

4. ทดสอบ X-Validation กำหนดพารามิเตอร์ 5 Folds

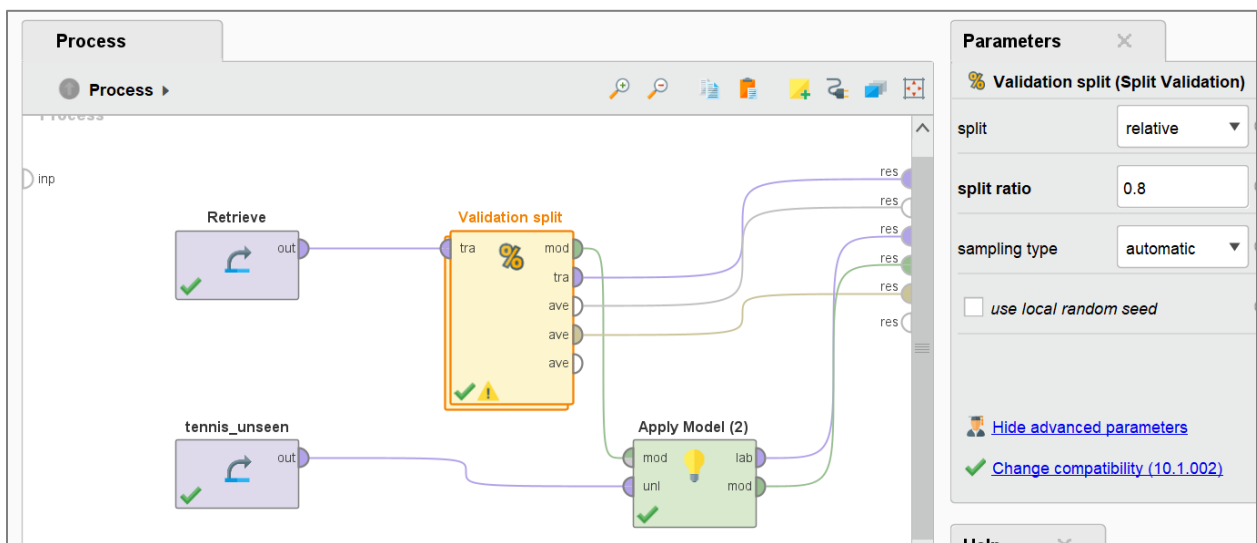


ผลลัพธ์ดังภาพ

accuracy: 60.00% +/- 25.28% (micro average: 58.33%)

	true No	true Yes	class precision
pred. No	2	3	40.00%
pred. Yes	2	5	71.43%
class recall	50.00%	62.50%	

สรุปโมเดลที่มีผลประสิทธิภาพสูงในการทำนายข้อมูล Tennis คือ Split Validation โดยแบ่งข้อมูลชุดฝึกเป็น 80% : 20% ซึ่งมีค่า accuracy เท่ากับ 66.67% และมาใช้ในการทำนายข้อมูล Tennis_unseen.csv



ผลลัพธ์ตั้งภาพ

Row No.	play	prediction(p...	confidence(No)	confidence(Yes)	days	outlook	temp	humidity	wind
1	?	Yes	0.053	0.947	D13	Overcast	Hot	Normal	Weak
2	?	Yes	0.437	0.563	D14	Rain	Mild	High	Strong