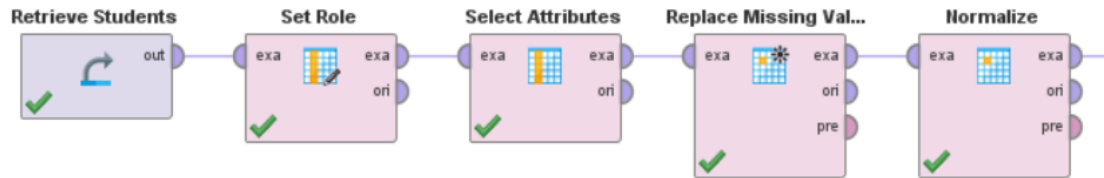


Homework 12

Data-preprocessing:




Label	Polynomial	0	Least Excellent (69)	Most Good (225)	Values Good (225), Bad (72), ...[1 more]
Students_success					
EntranceExam	Real	0	Min 0	Max 1	Average 0.560
Average_Grade	Real	0	Min 0	Max 1	Average 0.502
Exam1	Real	0	Min 0	Max 1	Average 0.289
Exam2	Real	0	Min 0	Max 1	Average 0.455
Exam3	Real	0	Min 0	Max 1	Average 0.466
Exam4	Real	0	Min 0	Max 1	Average 0.404
Exam5	Real	0	Min 0	Max 1	Average 0.504
Exam6	Real	0	Min 0	Max 1	Average 0.441

Row No.	Students_s...	EntranceEx...	Average_Gr...	Exam1	Exam2	Exam3	Exam4	Exam5	Exam6	Exam7
1	Good	1	0.571	0.500	0.500	0.500	0.250	0.500	0.750	0.500
2	Good	1	0.679	0.500	0.500	0.500	0.500	0.750	0.500	0.750
3	Excellent	1	0.714	0.750	0.750	0.750	0.250	0.500	0.250	0.500
4	Excellent	1	0.786	0.750	1	1	0.750	0.500	0.250	1
5	Excellent	0.945	0.857	0.250	1	1	1	1	0.500	1
6	Excellent	0.945	0.857	0.750	1	1	0.500	0.750	0.500	1
7	Excellent	0.945	0.750	0	1	1	0.500	0.500	0.500	0.750
8	Excellent	0.945	0.964	1	1	1	1	0.250	0.750	1
9	Good	0.945	0.607	0.500	0.750	0.750	0.500	0.750	0.500	0.750
10	Good	0.945	0.607	1	0.750	0.750	0.750	0.750	0.750	0.750

Data clustering:

- Cluster students using best clustering methods.

Parameters ✕

 **Clustering (k-Means)**

☒ add cluster attribute

☐ add as label

☐ remove unlabeled

k

3

max runs

10

☒ determine good start values

measure types

NumericalMeasures

numerical measure

EuclideanDistance

max optimization steps

100

☐ use local random seed



Attribute	cluster_0	cluster_1	cluster_2
EntranceExam	0.438	0.558	0.700
Average_Grade	0.344	0.491	0.693
Exam1	0.099	0.376	0.410
Exam2	0.199	0.502	0.695
Exam3	0.226	0.496	0.706
Exam4	0.408	0.192	0.629
Exam5	0.500	0.311	0.717
Exam6	0.277	0.480	0.583
Exam7	0.440	0.679	0.818
Exam8	0.441	0.623	0.822
Exam9	0.443	0.485	0.625
Exam10	0.559	0.676	0.683
Exam11	0.341	0.508	0.697
Exam12	0.349	0.523	0.713

- Explain each cluster.

ในแต่ละ cluster จะมีค่าเฉลี่ยของคะแนนสอบที่แตกต่างกันค่อนข้างชัดเจน

- Cluster_0: ในกลุ่มนี้จะเป็นกลุ่มของนักเรียนที่มีคะแนนสอบต่ำที่สุดจากทั้ง 3 cluster รวมทั้งมีเกรดที่ต่ำสุดด้วย
- Cluster_1: ในกลุ่มนี้จะเป็นกลุ่มของนักเรียนที่มีคะแนนสอบที่ค่อนข้างจะอยู่ตรงกลางจากทั้ง 3 cluster หรือก็คือเป็นกลุ่มที่รวมนักเรียนที่ได้คะแนนสอบอยู่ที่ประมาณค่าเฉลี่ย
- Cluster_2: ในกลุ่มนี้จะเป็นกลุ่มของนักเรียนที่มีคะแนนสอบสูงที่สุดจากทั้ง 3 cluster รวมทั้งมีเกรดสูงที่สุดด้วย หรือก็คือเป็นกลุ่มของนักเรียนที่มีผลการเรียนดี

- Suggest strategies for improving performance of students.

- Cluster_0: เนื่องจากในกลุ่มนี้เป็นกลุ่มของนักเรียนที่มีคะแนนสอบต่ำที่สุดจากทั้ง 3 cluster จึงคิดว่าควรให้ความสนใจกับนักเรียนกลุ่มนี้เป็นพิเศษ อาจจะเป็นการสอนเสริมหรือการปูพื้นฐานเกี่ยวกับเนื้อหาที่จะเรียนเพื่อให้นักเรียนกลุ่มนี้เข้าใจเนื้อหาได้เร็วขึ้น
- Cluster_1: เนื่องจากในกลุ่มนี้จะเป็นกลุ่มของนักเรียนที่มีคะแนนสอบที่ค่อนข้างจะอยู่ในค่าเฉลี่ย คาดว่าในกลุ่มนี้น่าจะเป็นกลุ่มของนักเรียนที่พอจะมีความรู้พื้นฐานมาบ้างแล้ว จึงไม่จำเป็นต้องให้ความสนใจมากเป็นพิเศษเท่า Cluster_0 แต่ก็ควรมีการถามถึงความเข้าใจในบทเรียนว่ามีตรงส่วนไหนบ้างที่ยังไม่เข้าใจ
- Cluster_2: เนื่องจากในกลุ่มนี้จะเป็นกลุ่มของนักเรียนที่มีคะแนนสอบสูงที่สุดจากทั้ง 3 cluster รวมทั้งมีเกรดสูงที่สุดด้วย คาดว่าจะเป็นกลุ่มของนักเรียนที่มีผลการเรียนดี หรือมีความเข้าใจในเนื้อหาของเรื่องที่เรียนค่อนข้างมาก ดังนั้นจึงคิดว่าควรให้งานหรือการบ้านเพื่อเป็นการเพิ่มความรู้อย่างไรก็ตามความเข้าใจในเนื้อหา และฝึกฝนในการทำโจทย์ เพื่อเพิ่มความคุ้นเคยเมื่อต้องทำในข้อสอบ