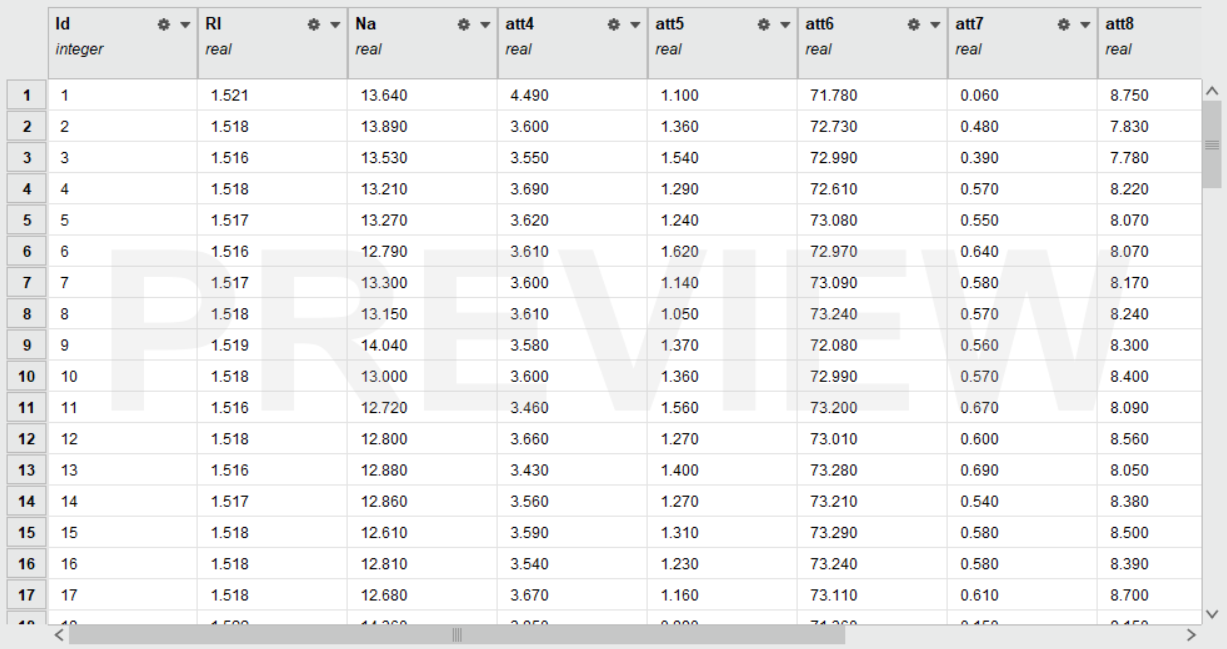
**Exercise : PCA + k-NN**

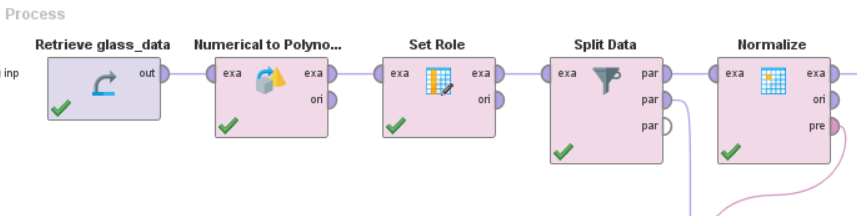
**From Rapidminer**

1. Construct the classifier to predict type of glass without the use of PCA

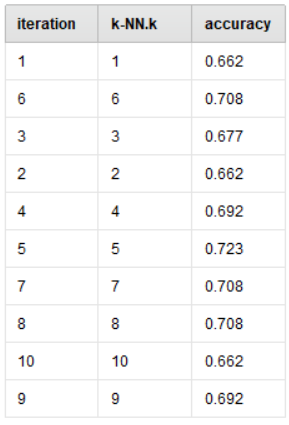
* Pre-processing:



Rename Header

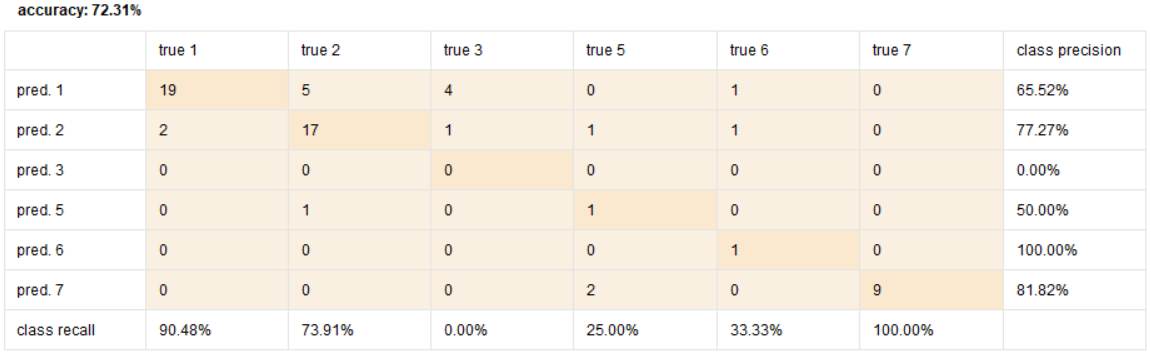


* Classification:
  + Show results on different values of K



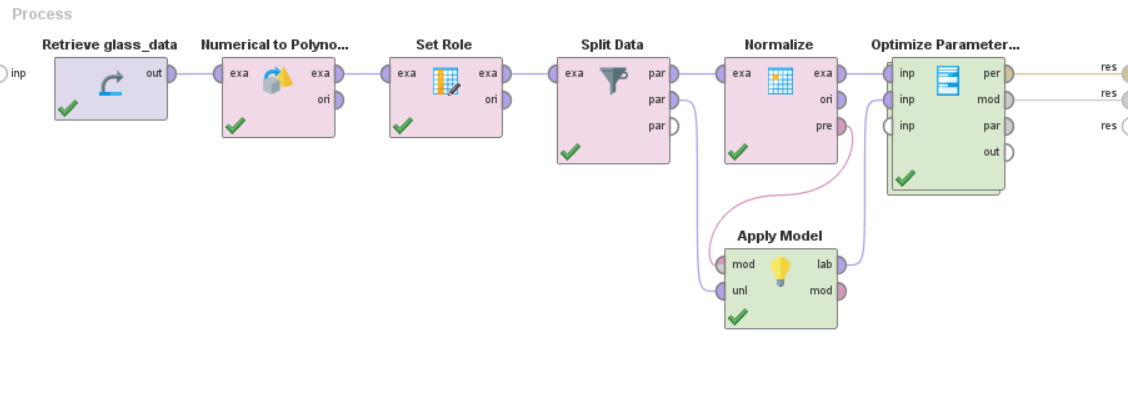
Ans best K to obtain best accuracy is 5

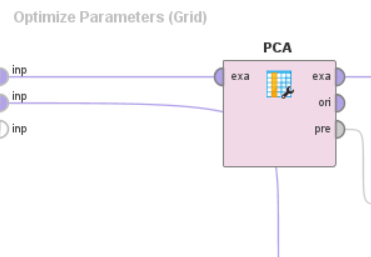
* Show confusion-matrix for each class



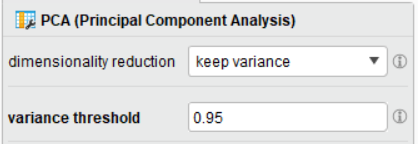
1. Construct the classifier to predict type of glass with the use of PCA

* Pre-processing:
  + Show each step of your pre-processing (include PCA)



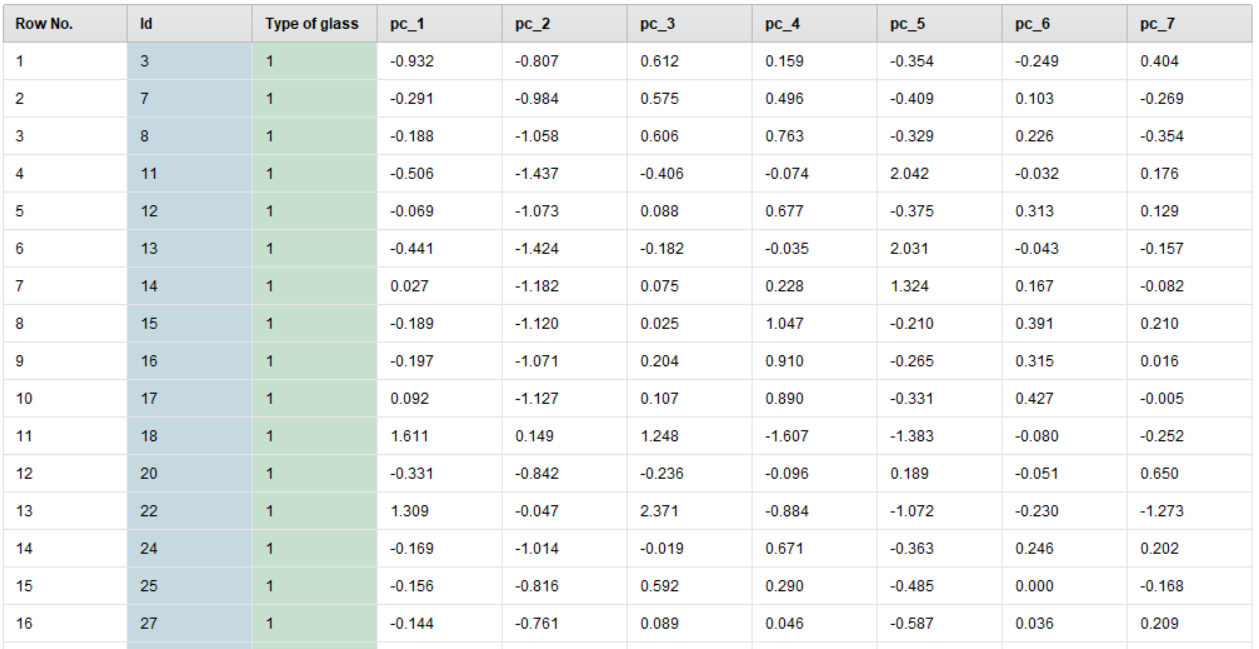


* Show & explain setting of PCA parameters

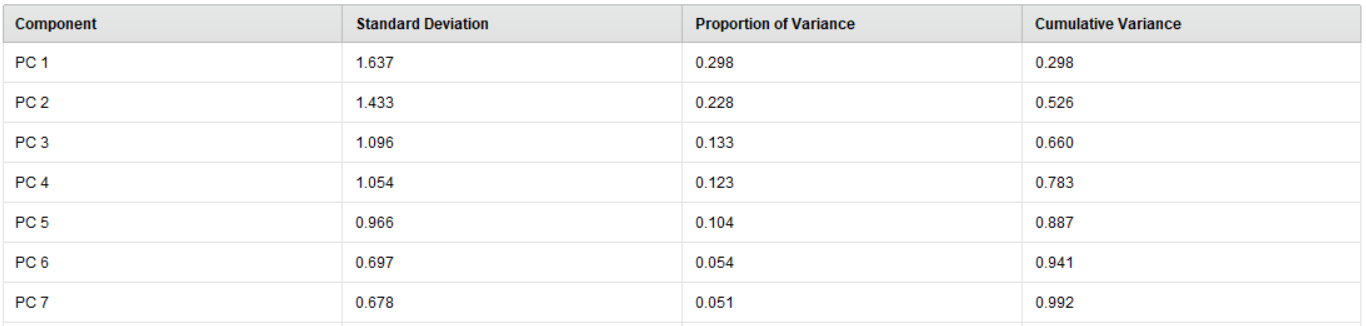


Ans ปรับ variance threshold ไว้ที่ 0.95 เพื่อให้คุณลักษณะที่แตกต่างกันระหว่างจุดข้อมูลถึงที่ขอบเขตกำหนด เนื่องจากใน component ที่ 1 จะมีค่า variance สูงสุดและลดลงตามลำดับ

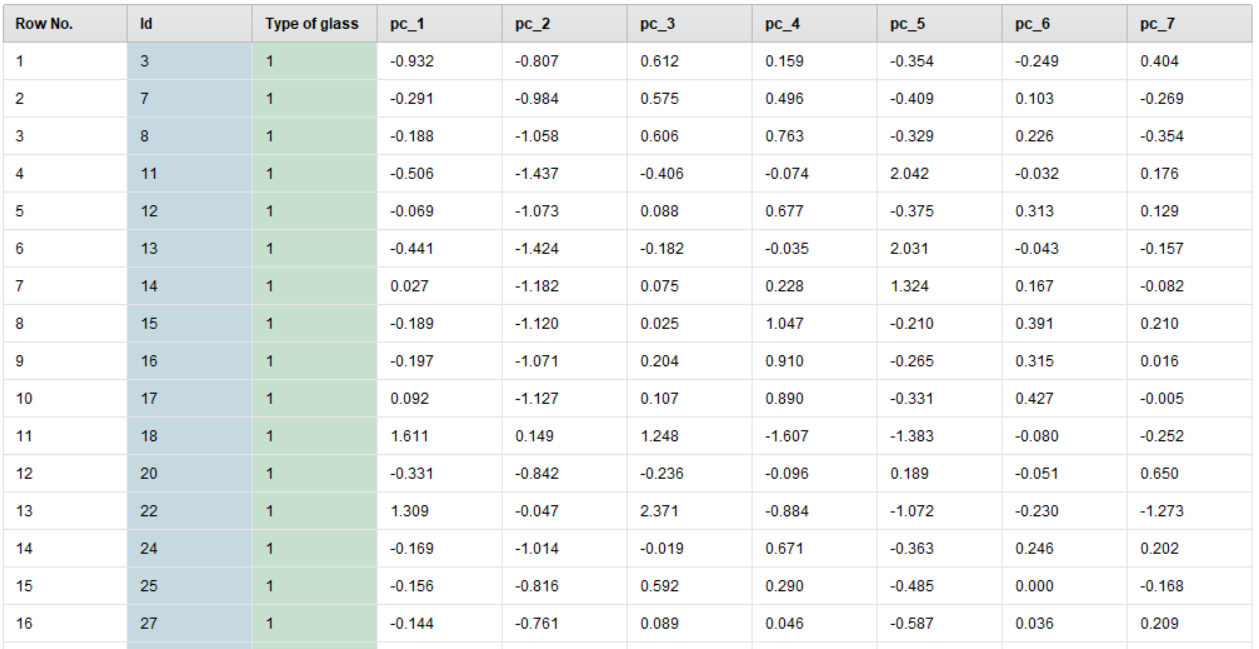
* Explain output of PCA in terms of
  + Principal components



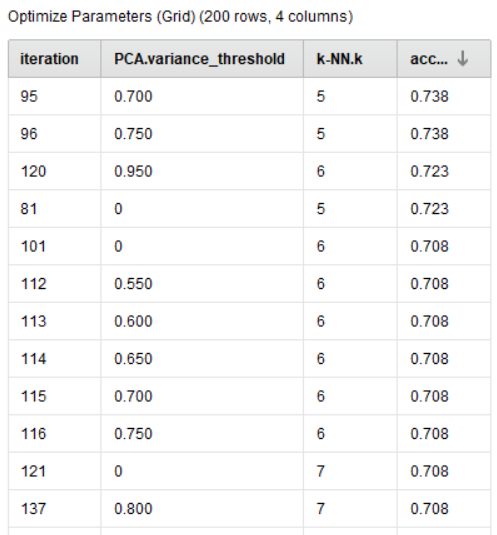
* + Accumulate variance



* Classification:
  + Show format of the dataset in terms of new dimensions

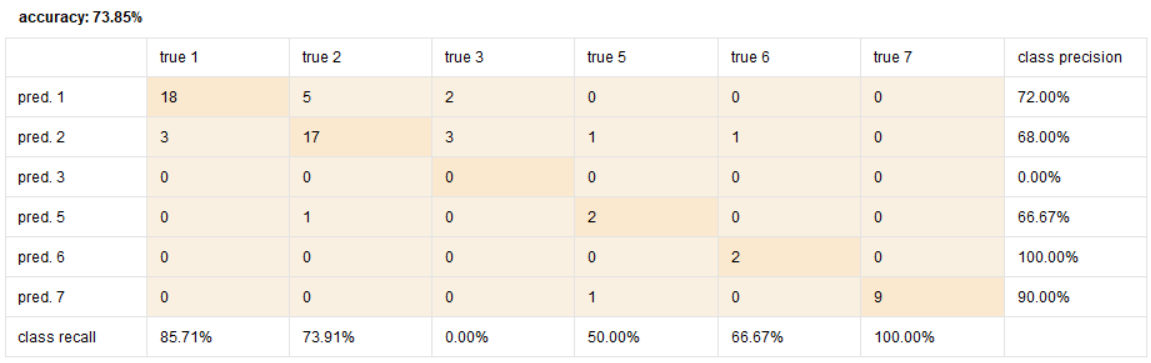


* What is best combination of PCA (variance) & K to obtain best accuracy



Ans best PCA (variance) is 0.700 and k is 5

* Show confusion matrix for each class



1. Discuss and Compare accuracy obtained between with PCA and without PCA

Ans การใช้ PCA ทำให้ dimension ของข้อมูลลดลง ส่งผลให้โอกาศที่จะเกิด Over fitting น้อยลงตามไปด้วย ทำให้ accuracy เพิ่มสูงขึ้นจากการทำ k-NN เพียงอย่างเดียว