

FL #7

1. Read the MNIST dataset using provided python code.

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

2. Extract several meaningful features from images, and apply k-means algorithms (e.g., k=10). Some examples of features are as follows.

영역1의 특징 기술자
• 면적 $a = 20$
• 중점 $(\bar{y}, \bar{x}) = \left(\frac{1}{20} \sum_{(y,x) \in R} y, \frac{1}{20} \sum_{(y,x) \in R} x \right) = (3.05, 2.7)$
• 행 분산 $v_{rr} = \frac{1}{20} \sum_{(y,x) \in R} (x - 2.7)^2 = 3.01$
• 열 분산 $v_{cc} = \frac{1}{20} \sum_{(y,x) \in R} (y - 3.05)^2 = 1.848$
• 혼합 분산 $v_{rc} = \frac{1}{20} \sum_{(y,x) \in R} (y - 3.05)(x - 2.7) = -1.135$
• 둘레 $p = 10 + 6\sqrt{2} = 18.485$
• 등근 정도 $r = \frac{4\pi \times 20}{18.485^2} = 0.736$

3. Measure the uniformity of each cluster.