

데이터마이닝 강의 – Review 시험 문제

강사: 심규석

[1] 다음 물음들에 답하시오.

(a) 다음 표를 보고 age와 %fat의 평균(mean), 표준편차(standard deviation)을 각각 구하시오.

age	20	25	27	32	45
%fat	7%	8%	11%	20%	15%
age	49	52	57	60	62
%fat	35%	40%	28%	31%	19%

(b) 다음 표는 숫자 데이터를 간격으로 그룹화한 결과이다. 이 때 해당 데이터의 근사된 중위값(approximate median)을 구하라.

Age	Frequency
1-5	100
6-15	700
16-20	300
21-50	400
51-80	500

[2] 다음의 각 attribute type이 무엇인지 간단히 서술하시오.

(a) Nominal attribute

(b) Numeric attribute

(c) Ordinal attribute

(d) Binary attribute

[3] Tuple (22, 1, 42, 10)과 (20, 0, 36, 8)로 표현되는 두 객체가 주어졌을 때, 다음 물음에 답하라.

(a) 두 객체 사이의 Euclidean distance를 구하라.

(b) 두 객체 사이의 Manhattan distance를 구하라.

(c) $q=3$ 일 때, 두 객체 사이의 Minkowski distance를 구하라.