

## 데이터마이닝 강의 – 시험문제 Section 3

강사: 심규석

1. 다음 정규화(normalization) 방법 적용 시 결과값의 범위는?

(1) Min-max 정규화

(2) Z 스코어 정규화

2. 연령 데이터 (27,29,30,31,33)가 있다고 하자.

(1) 주어진 연령 데이터를 min-max 정규화를 이용하여 변환하시오.

(2) 주어진 연령 데이터를 Z 스코어 정규화를 이용하여 변환하시오. (표준편차는 2)

3. 다음과 같이 12개의 판매가격 레코드가 정렬되어 있다고 하자.

0, 10, 11, 13, 15, 35, 50, 55, 72, 92, 178, 195

(1) 주어진 레코드를 동일빈도 구분 방법(equal-frequency partitioning)을 이용해 세 개의 빈(bin)으로 구분하시오.

(2) 폭(width)이 40인 동일폭 히스토그램(equal-width histogram)을 그리시오.