

Date: 2023/10/16

From: 11204604 洪哲文

Assignment:

Slide81: Exercise

Suppose a group of 12 sales price records has been sorted as follows:

5, 10, 11, 13, 15, 35, 50, 55, 72, 92, 204, 215.

Partition them into three bins by each of the following methods.

(a) equal-frequency (equal-depth) partitioning

bin1 : 5 、 10 、 11 、 13

bin2 : 15 、 35 、 50 、 55

bin3 : 72 、 92 、 204 、 215

(b) equal-width partitioning

width =  $(215 - 5) / 3 = 70$

bin1 : 5 、 10 、 11 、 13 、 15 、 35 、 50 、 55 、 72

bin2 : 92

bin3 : 204 、 215

(c) min-max normalization by setting min=0 and max=1

$X_{norm} = (X-5)/(215-5) * (1-0)+0 = (X-5)/(215-5)$

So that:

• 5 -> 0

• 10 -> 0.0238

• 11 -> 0.0286

• 13 -> 0.0381

• 15 -> 0.0476

• 35 -> 0.1429

• 50 -> 0.2143

• 55 -> 0.2381

• 72 -> 0.3190

-----

• 92 -> 0.4143

-----

• 204 -> 0.9476

• 215 -> 1

So the 3bins is:

bin1 : 5 、 10 、 11 、 13 、 15 、 35 、 50 、 55 、 72

bin2 : 92

bin3 : 204 、 215