

$= (2, 2)$

2. Cho tập dữ liệu sau: A(2,10), B(2,5), C(8,4), D(5,8), E(7,5), F(6,4), G(1,2), H(4,9).

a. Tìm các trung tâm cụm sau lần lặp thứ 1 dùng thuật toán k-mean, giả sử các trung tâm cụm ban đầu là A, D, G.

b. Tìm các trung tâm cụm cuối cùng sau khi thực hiện xong thuật toán k-mean với tập dữ liệu trên.

3. Cho tập mẫu gồm những mẫu hai thuộc tính như sau: (1, 1), (1, 2), (2, 1), (2, 1.5), (3, 2), (4, 1.5), (4, 2), (5, 1.5), (4.5, 2), (4, 4), (4.5, 4), (4.5, 5), (4, 5), (5, 5). Dùng giải thuật k-means để gom cụm tập mẫu nêu trên với  $k = 3$  ứng với các trường hợp sau:

a. Các trung tâm cụm ban đầu là (2, 1.5), (4.5, 2), (4.5, 5).

b. Các trung tâm cụm ban đầu là (1, 2), (3, 2), (4, 2).

4. Cho tập mẫu gồm những mẫu hai thuộc tính như sau: (1, 1), (1, 2), (2, 1), (2, 1.5), (3, 2), (4, 1.5), (4, 2), (5, 1.5), (4.5, 2), (4, 4), (4.5, 4), (4.5, 5), (4, 5), (5, 5). Gom cụm tập mẫu nêu trên thành 4 cụm dùng giải thuật:

a. Gom cụm phân cấp gộp AGNES (AGglomerative NESTing) dùng single-link (minimum distance)

b. Gom cụm phân cấp gộp AGNES (AGglomerative NESTing) dùng complete-link (maximum distance)

5. Cho tập mẫu gồm  $n$  mẫu và ta cần gom cụm tập mẫu này thành 2 cụm. Vậy có bao nhiêu cách khác nhau để gom cụm như vậy?