**Phân cụm bán giám sát mờ** **( SSSFC - Semi-supervised standard fuzzy clustering )**:

- Được xây dựng dựa trên các thuật toán phân cụm mờ kết hợp với thông tin bổ trợ được người dùng cung cấp

- Thông tin bổ trợ cho phân cụm cho phân cụm bán giám sát có 3 dạng cơ bản gồm các ràng buộc Must-link , Cannon-link và 1 phần dữ liệu được gắn nhãn và độ thuộc được xác định trước

- Yasunori et al. [] đã đề xuất một thuật toán phân cụm bán giám sát mờ với thông tin bổ trợ là hàm phụ thuộc được bổ sung vào hàm mục tiêu FCM ( Fuzzzy C-Mean ) nhằm cải thiện hiệu quả trong quá trình phân cụm.

- Công thức hàm mục tiêu :

Trong đó có :

+ m là số mờ hóa

+ C là số cụm, N là số phần tử dữ liệu, r là số chiều của dữ liệu

+ là độ thuộc của phần tử vào dữ liệu là từ cụm j

+ là phần tử thứ k của

+ là tâm là cụm j

+ Với điều kiên ràng buộc là thì khi đó hàm độ thuộc bổ trợ của phần tử với cụm là đồng thời thỏa mãn , ,

- Khi đó dựa vào điều kiện ràng buộc và hàm mục tiêu trên thì ta có :

với

- Các giá trị của được xác định theo 2 trường hợp :

+ Với m > 1 :

+ Với m = 1 :

Các bước thực hiện thuật toán phân cụm mờ bán giám sát ()

|  |  |
| --- | --- |
| Input | Tập dữ liệu X gồm N phần tử, số cụm C, ma trận độ phụ thuộc , ngưỡng , số lần lặp tối đa maxStep > 0 |
| Output | Ma trận U và tâm cụm V |
| **SSSFC** | |
| 1 : | t = 0 |
| 2 : | Khởi tạo ngẫu nghiên |
| 3 : | Bước lặp ( Repeat ) |
| 4 : | t = t + 1 |
| 5 : | Tính theo các trường hợp trên của m |
| 6 : | Tính như trên |
| 7 : | Cho đến khi or t > maxStep thì dừng |