

بنام خدا

سمینار درس بازشناسی الگو

دو کاربرد ویژه PSO

استاد: دکتر کبیر

ارایه: سامان پروانه

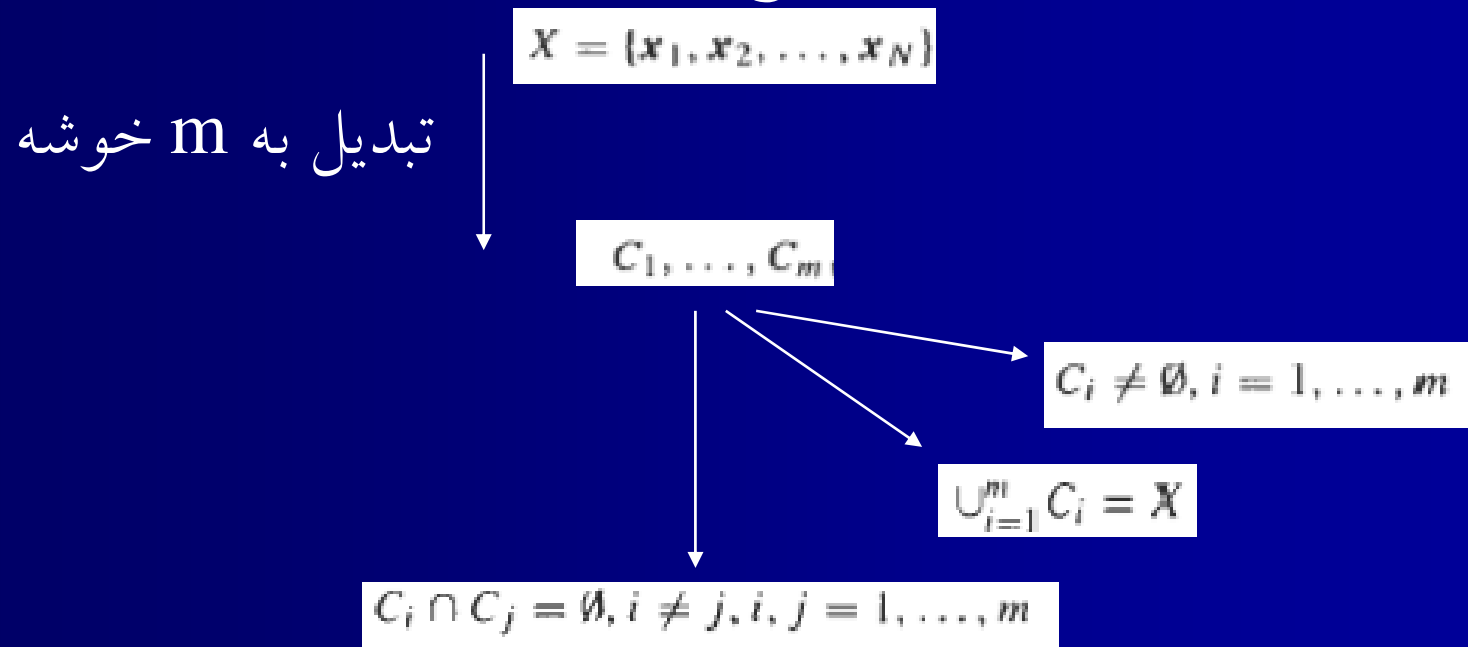
فهرست مطالب:

- تعریف خوشه‌بندی
- مرور موارد مربوط به خوشه‌بندی
- مروری بر روش Kmeans
- PSO و کاربرد آن در خوشه‌بندی
- ترکیب Kmeans و PSO جهت خوشه‌بندی
- ترکیب PSO و GA

تعریف خوشه‌بندی:

تعریف کلاسیک خوشه‌بندی؟

خوشه‌بندی و شرایط آن برای نوع سخت:



معیارهای نزدیک بودن:

weighted l_p metric

$$d_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^l w_i |x_i - y_i|^p \right)^{1/p}$$

If $w_i = 1, i = 1, \dots, l$, we obtain the *unweighted* l_p metric

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^l (x_i - y_i)^2}$$

کاربردهای خوشه‌بندی:

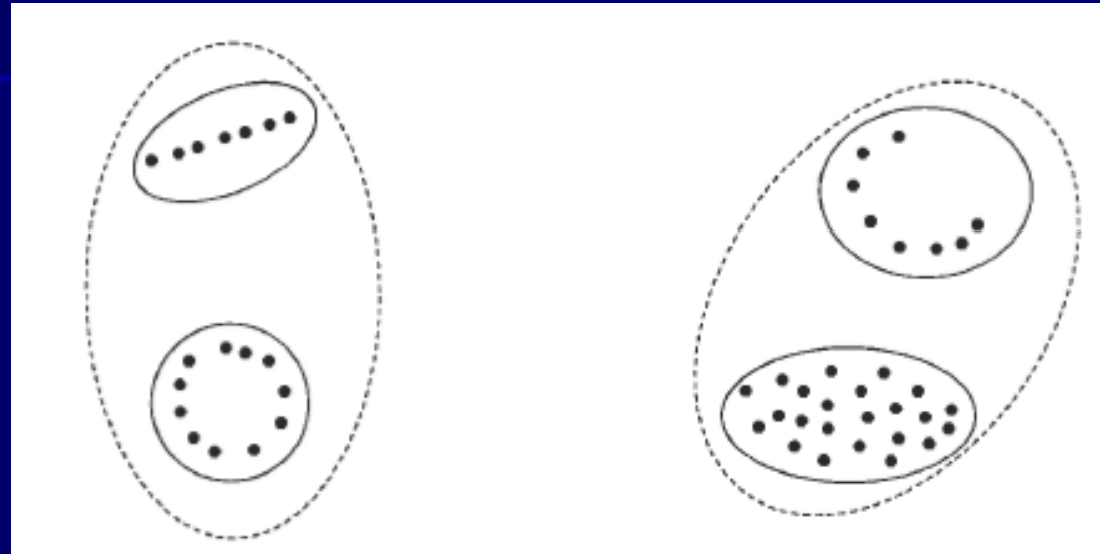
۱- کاهش داده

۲- پیش‌بینی براساس گروهها

۳- بازشناسی الگو و ناحیه‌بندی داده

...

تعداد خوشه‌ها:



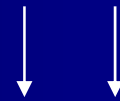
۱- با تعداد زیادی خوشه شروع می‌کنیم و خوشه‌هایی که از دیدگاه معیار همسانی یکسان هستند را ترکیب می‌کنیم.
(Cluster Merging)

۲- به ازای تعداد متفاوتی خوشه، فرآیند خوشه‌بندی را انجام می‌دهیم و سپس مناسب بودن هر یک را ارزیابی می‌کنیم.

داده‌ها:

داده نوعاً مشاهدات یک فرآیند فیزیکی است.

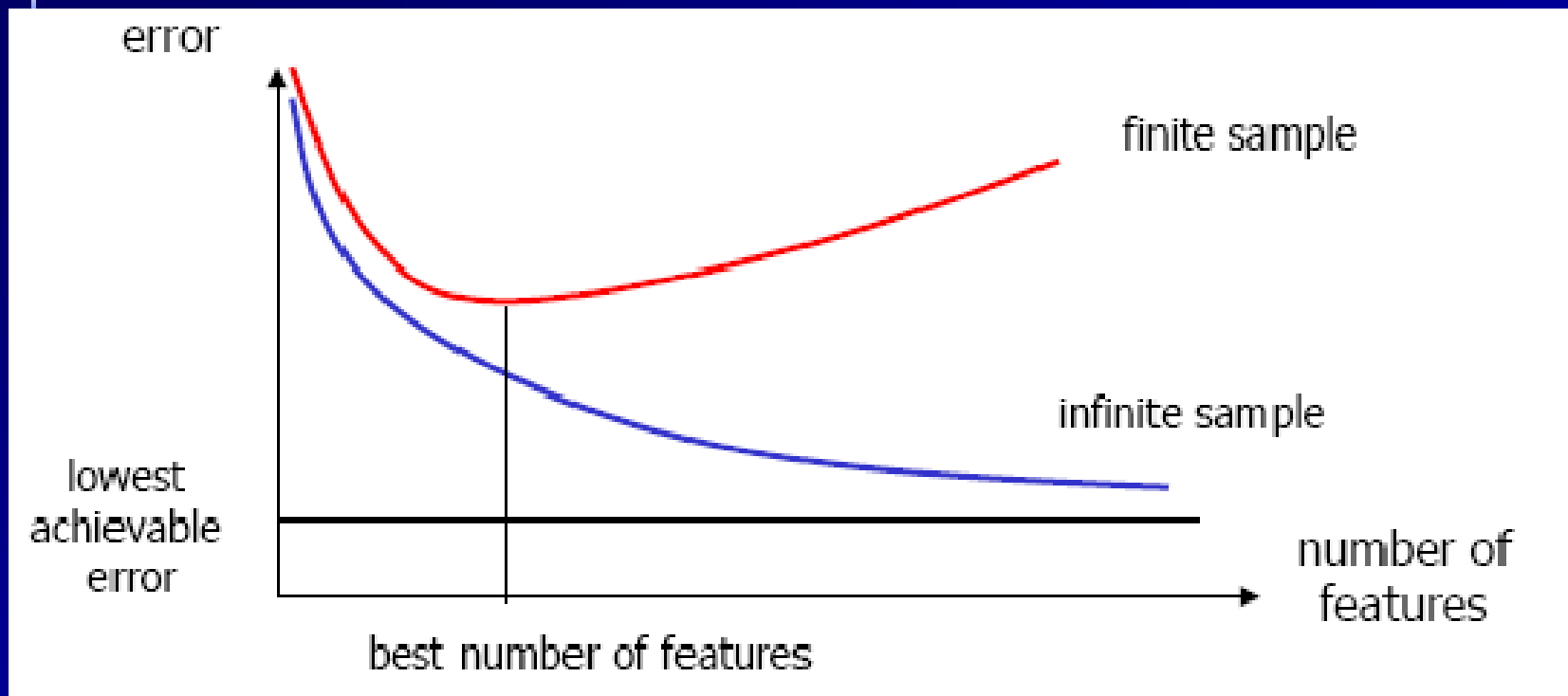
ویژگیها



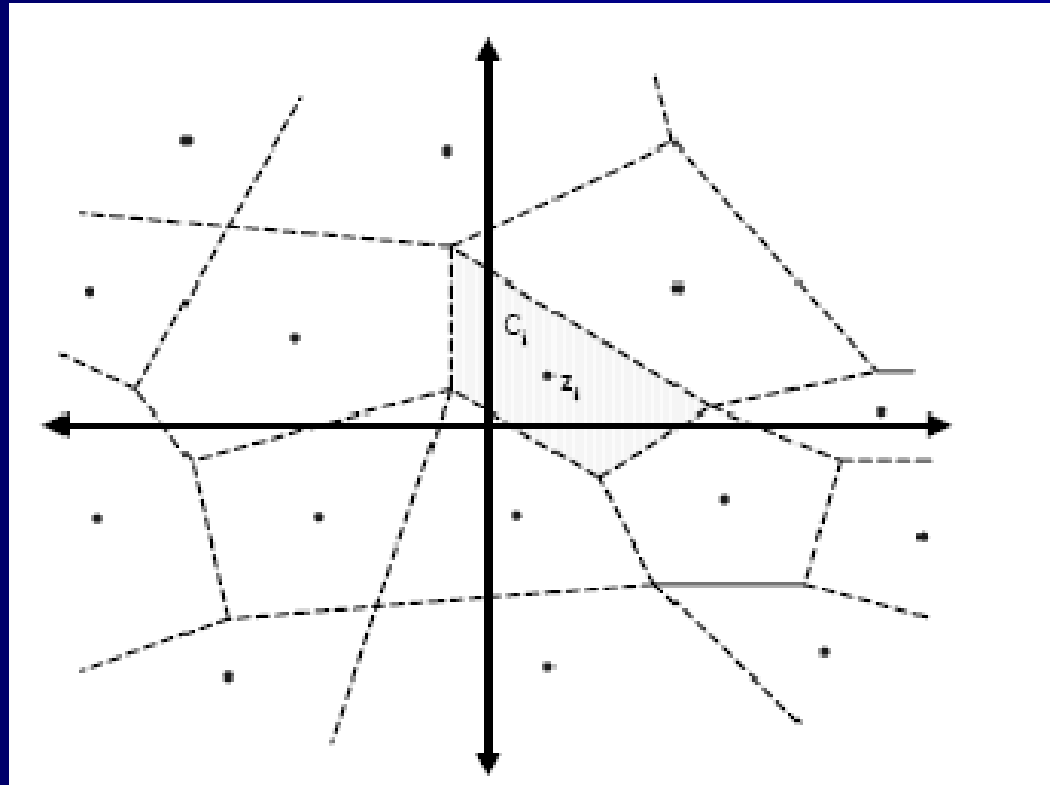
ماتریس الگو

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Nn} \end{bmatrix}$$

تعداد ویژگیها و مشخصات آنها:



مقداردهی اولیه برای Kmeans:



Kmeans

1. Randomly initialize the N_c cluster centroid vectors
2. Repeat

- (a) For each data vector, assign the vector to the class with the closest centroid vector, where the distance to the centroid is determined using

$$d(\mathbf{z}_p, \mathbf{m}_j) = \sqrt{\sum_{k=1}^{N_d} (z_{pk} - m_{jk})^2} \quad (1)$$

where k subscripts the dimension.

- (b) Recalculate the cluster centroid vectors, using

$$\mathbf{m}_j = \frac{1}{n_j} \sum_{\forall \mathbf{z}_p \in C_j} \mathbf{z}_p \quad (2)$$

until a stopping criterion is satisfied.

PSO و کاربرد آن در خوشه‌بندی:

* تعداد ذرات چقدر باشد؟ ۱۰-۲۰ ذره

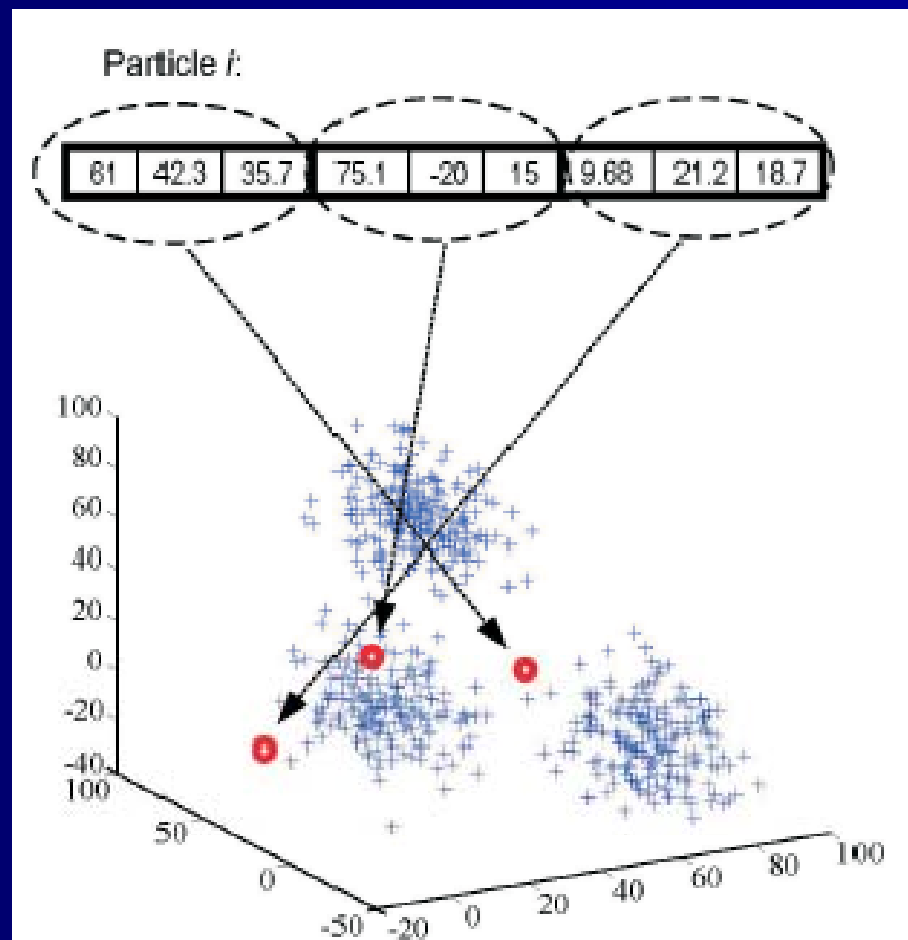
* هر ذره نمایانگر چه چیزی است؟

$$X_i = (m_{i,1}, \dots, m_{i,k}, \dots, m_{i,K})$$

یک ذره تعدادی کاندید را برای مراکز خوشه داده نشان می‌دهد.

* تابع ارزیابی چه باشد؟

رمزگذاری فضای سه بعدی در ذرات:



کمی کردن کیفیت خوشه‌بندی:

$$1- \quad J_e = \frac{\sum_{k=1}^K \left[\sum_{\forall z_p \in C_k} d(z_p, m_k) \right] / n_k}{K}$$

$$2- \quad f(x_i, Z_i) = w_1 \bar{d}_{\max}(Z_i, x_i) + w_2 (z_{\max} - d_{\min}(x_i))$$

$$z_{\max} = 2^s - 1$$

$$3- \quad f(x_i, Z_i) = w_1 \bar{d}_{\max}(Z_i, x_i) + w_2 (z_{\max} - d_{\min}(x_i)) + w_3 J_{e,i}$$

$$\bar{d}_{\max}(Z_i, x_i) = \max_{k=1, \dots, K} \left\{ \sum_{\forall z_p \in C_{ik}} d(z_p, m_{ik}) / n_{ik} \right\}$$

$$d_{\min}(x_i) = \min_{\forall k, k_2, k_1=k_2} \{d(m_{ik}, m_{ik_2})\}$$

PSO و خوشه‌بندی (الگوریتم ۱):

1. Initialize each particle to contain K randomly selected cluster centroids

2. For $t = 1$ to t_{\max}

(a) For each particle i

i. For each pattern z_p

- calculate $d(z_p, m_{i,k})$ for all clusters $C_{i,k}$ using equation

(3.1)

- assign z_p to $C_{i,k}$ where

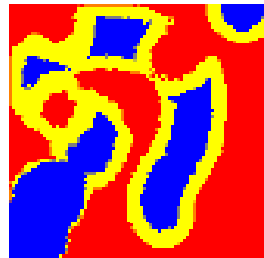
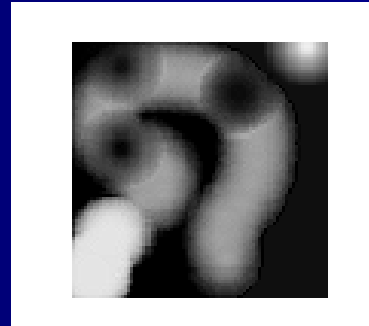
$$d(z_p, m_{i,k}) = \min_{\forall k=1, \dots, K} \{d(z_p, m_{i,k})\} \quad (4.5)$$

ii. Calculate the fitness, $f(x_i, Z_i)$

(b) Find the personal best position for each particle and the global best solution, $\hat{y}(t)$

(c) Update the cluster centroids using equations (2.8) and (2.10)

PSO و خوشه‌بندی در ناحیه‌بندی تصویر:



(a) K-means



(b) PSO

PSO و خوشه‌بندی (الگوریتم ۲):

1. Initialize each particle to contain K randomly selected cluster centroids

2. For $t = 1$ to t_{\max}

(a) For each particle i

i. For each pattern z_p

- calculate $d(z_p, m_{i,k})$ for all clusters $C_{i,k}$ using equation

(3.1)

- assign z_p to $C_{i,k}$ where

$$d(z_p, m_{i,k}) = \min_{\forall k=1, \dots, K} \{d(z_p, m_{i,k})\} \quad (4.5)$$

ii. Calculate the fitness, $f(x_i, Z_i)$

(b) Find the personal best position for each particle and the global best solution, $\hat{y}(t)$

(c) Update the cluster centroids using equations (2.8) and (2.10)

$$z_j^* = \frac{1}{N_j} \sum_{x_i \in C_j} x_i, j = 1, 2, \dots, K$$

ترکیب PSO و Kmeans برای خوشه‌بندی (الگوریتم هیبرید):

۱- الگوریتم Kmeans را اجرا می‌کنیم.

۲- مرکز خوشه بدست آمده از مرحله قبل را بعنوان یک ذره در نظر می‌گیریم و سایر ذرات را بصورت تصادفی مقداردهی می‌کنیم.

۳- خوشه‌بندی با روش PSO را اجرا می‌کنیم.

ترکیب PSO و GA

