Advanced Machine Learning

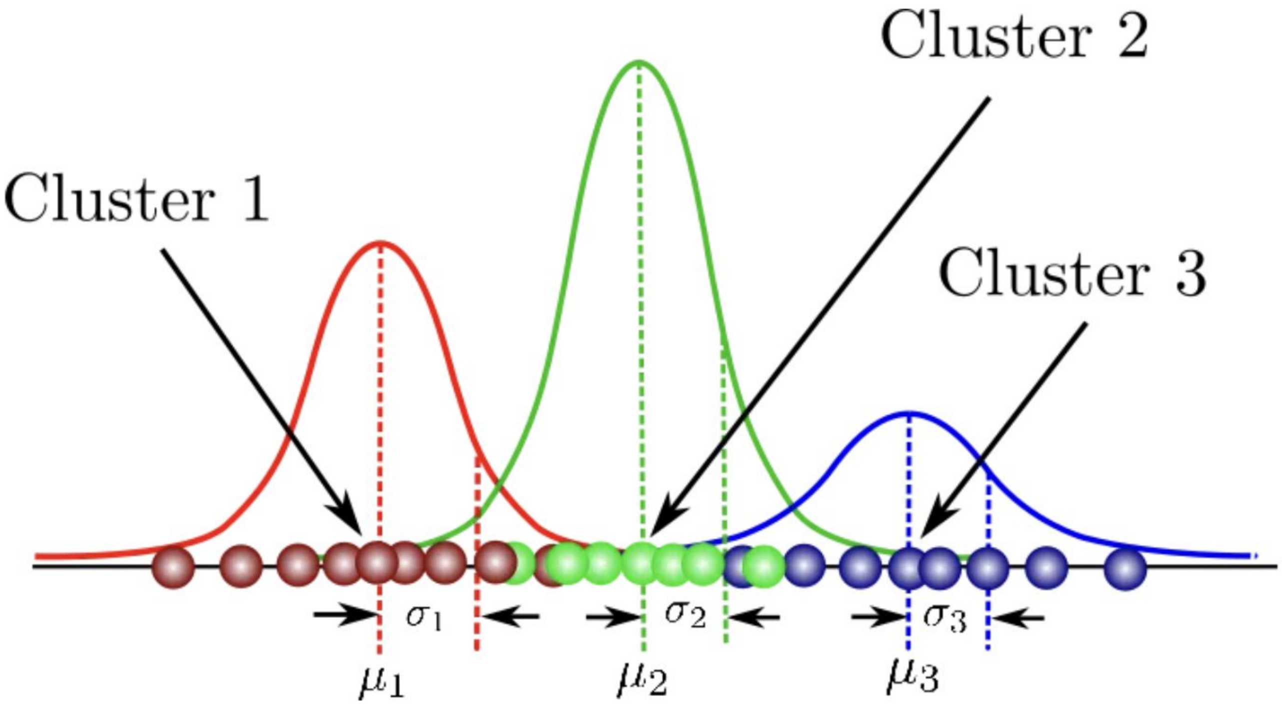
Likhit Nayak

Gaussian Mixture Models (GMM)

A Gaussian Mixture Model (GMM) is a function that is comprised of several Gaussians, each identified by $\boldsymbol{k \in \{1,\ldots, K\}}$, where $\boldsymbol{K}$ is the number of clusters of our dataset. Each Gaussian $\boldsymbol{k}$ in the mixture is comprised of the following parameters:

1. A mean $\boldsymbol{\mu}$ that defines its centre.
2. A covariance $\boldsymbol{\Sigma}$ that defines its width.
3. A mixing probability $\boldsymbol{\pi}$ that defines how big or small the Gaussian function will be

Gaussian Mixture Models (GMM)



Gaussian Mixture Model (GMM)

$$p(x) = \sum_{j=1}^{k} \pi_j \phi(x; \mu_j, \Sigma_j)$$

$\phi(x; \mu_j, \Sigma_j) \longrightarrow$ **Gaussian Distribution**

$\pi_j \longrightarrow \sum_{j=1}^{k} \pi_j = 1$

Clustering using GMMs

$$p(z_i = j | x_i)$$

where $\boldsymbol{z}_i$ is the cluster assignment for datapoint $\boldsymbol{x}_i$

$$\begin{aligned} p(z_i = j | x_i) &= \frac{p(z_i = j) p(x_i | z_i = j)}{p(x_i)} \\ &= \frac{\pi_j \phi(x_i; \mu_j, \Sigma_j)}{\sum_{l=1}^{k} \pi_l \phi(x_i; \mu_l, \Sigma_l)} \end{aligned}$$

Maximum Likelihood Estimation of GMM

$$\sum_{i=1}^{n} \log(p(x_i)) = \sum_{i=1}^{n} \log(\sum_{j=1}^{k} \pi_j \phi(x_i; \mu_j, \Sigma_j)).$$

(E-step) [Expectation step]: Compute soft class memberships, given the current parameters:

$$\tau_{ij} = P(z_i = j | x_{ij}, \pi, (\mu_\ell, \Sigma_\ell)).$$

Maximum Likelihood Estimation of GMM

(M-step) [Maximization step]: Update parameters by plugging in τ_{ij} (our guess) for the unknown $\mathbb{I}[z_i = j]$, which gives us:

$$\pi_j = \frac{1}{n} \sum_{i=1}^{n} \tau_{ij}, \quad \mu_j = \frac{\sum_{i=1}^{n} \tau_{ij} x_i}{\sum_{i=1}^{n} \tau_{ij}},$$

$$\Sigma_j = \frac{\sum_{i=1}^{n} \tau_{ij} (x_i - \mu_j)(x_i - \mu_j)^T}{\sum_{i=1}^{n} \tau_{ij}}.$$