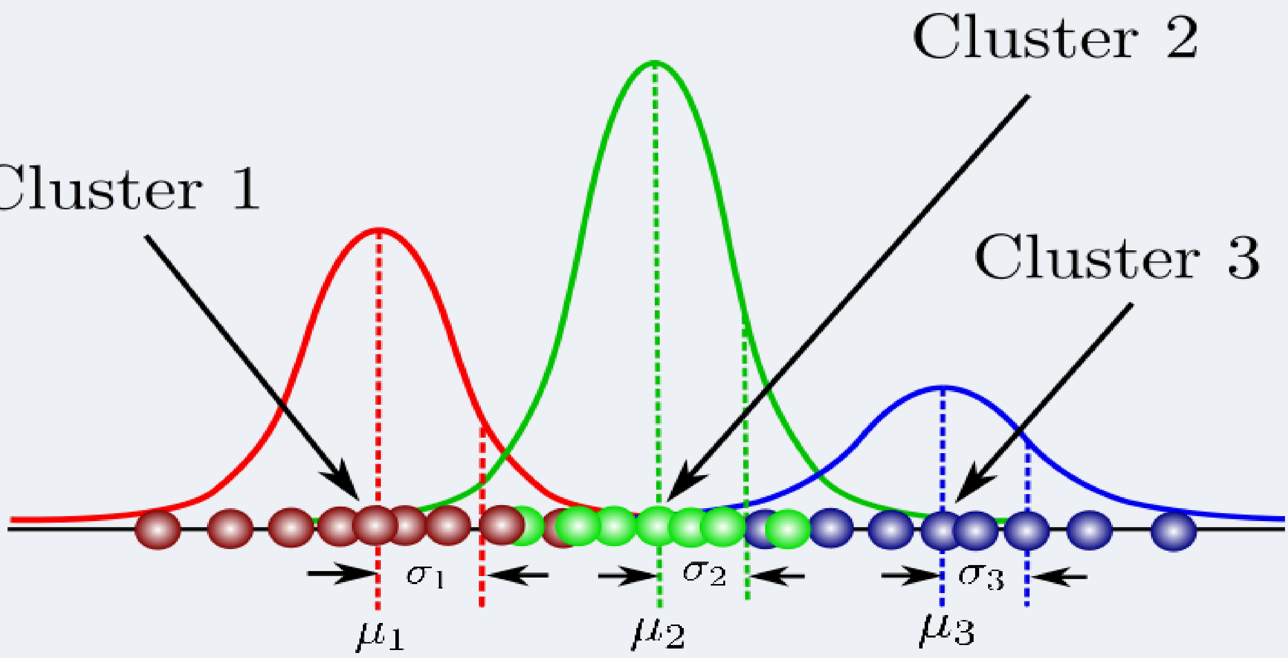
Gaussian Mixture Model

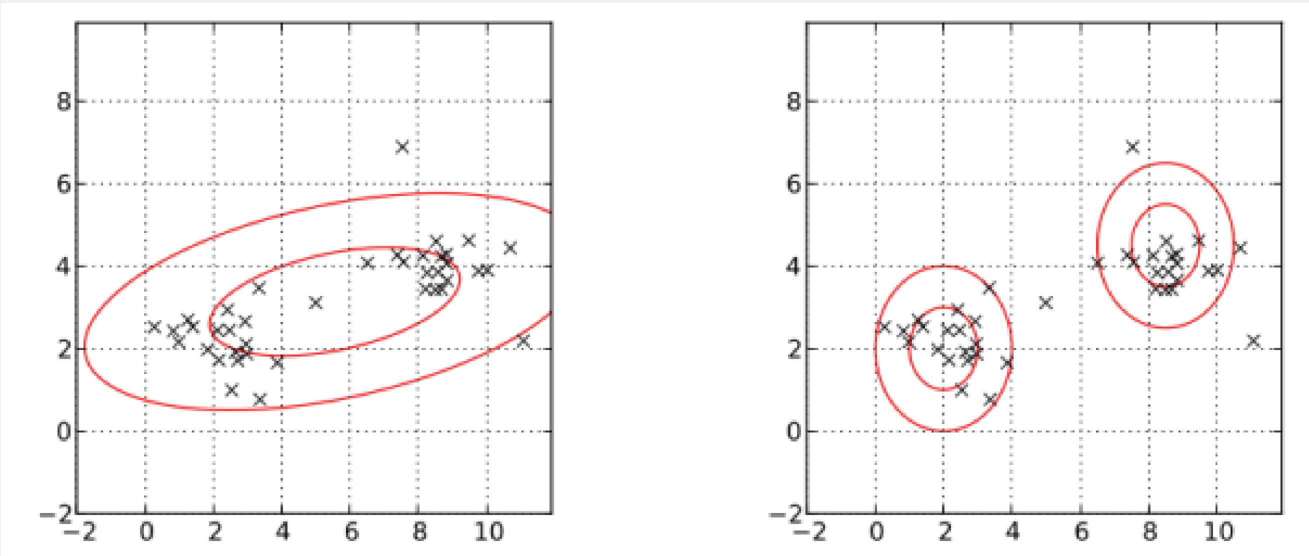
Mahdi Akbari Zarkesh
9612762638

INSTRUCTIONS

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

- *A mean μ that defines its centre.*
- *A covariance Σ that defines its width. This would be equivalent to the dimensions of an ellipsoid in a multivariate scenario.*
- A mixing probability π that defines how big or small the Gaussian function will be.





The Gaussian density function Formula

$$\mathcal{N}(\mathbf{x}|\mu,\Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right)$$

Where x represents our data points, D is the number of dimensions of each data point. μ and Σ are the mean and covariance, respectively. If we have a dataset comprised of N = 1000 three-dimensional points (D = 3), then x will be a 1000 × 3 matrix. μ will be a 1 × 3 vector, and Σ will be a 3 × 3 matrix. For later purposes, we will also find it useful to take the log of this equation, which is given by:

$$\ln \mathcal{N}(\mathbf{x}|\mu,\Sigma) = -\frac{D}{2}\ln 2\pi - \frac{1}{2}\ln\Sigma - \frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) \qquad (2)$$

If we differentiate this equation with respect to the mean and covariance and then equate it to zero, then we will be able to find the optimal values for these parameters, and the solutions will correspond to the Maximum Likelihood Estimates (MLE) for this setting. However, because we are dealing with not just one, but many Gaussians, things will get a bit complicated when time comes for us to find the parameters for the whole mixture.

One-dimensional Model

$$p(x) = \sum_{i=1}^{K} \phi_i \mathcal{N}(x \mid \mu_i, \sigma_i)$$

$$\mathcal{N}(x \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)$$

$$\sum_{i=1}^{K} \phi_i = 1$$

Multi-dimensional Model

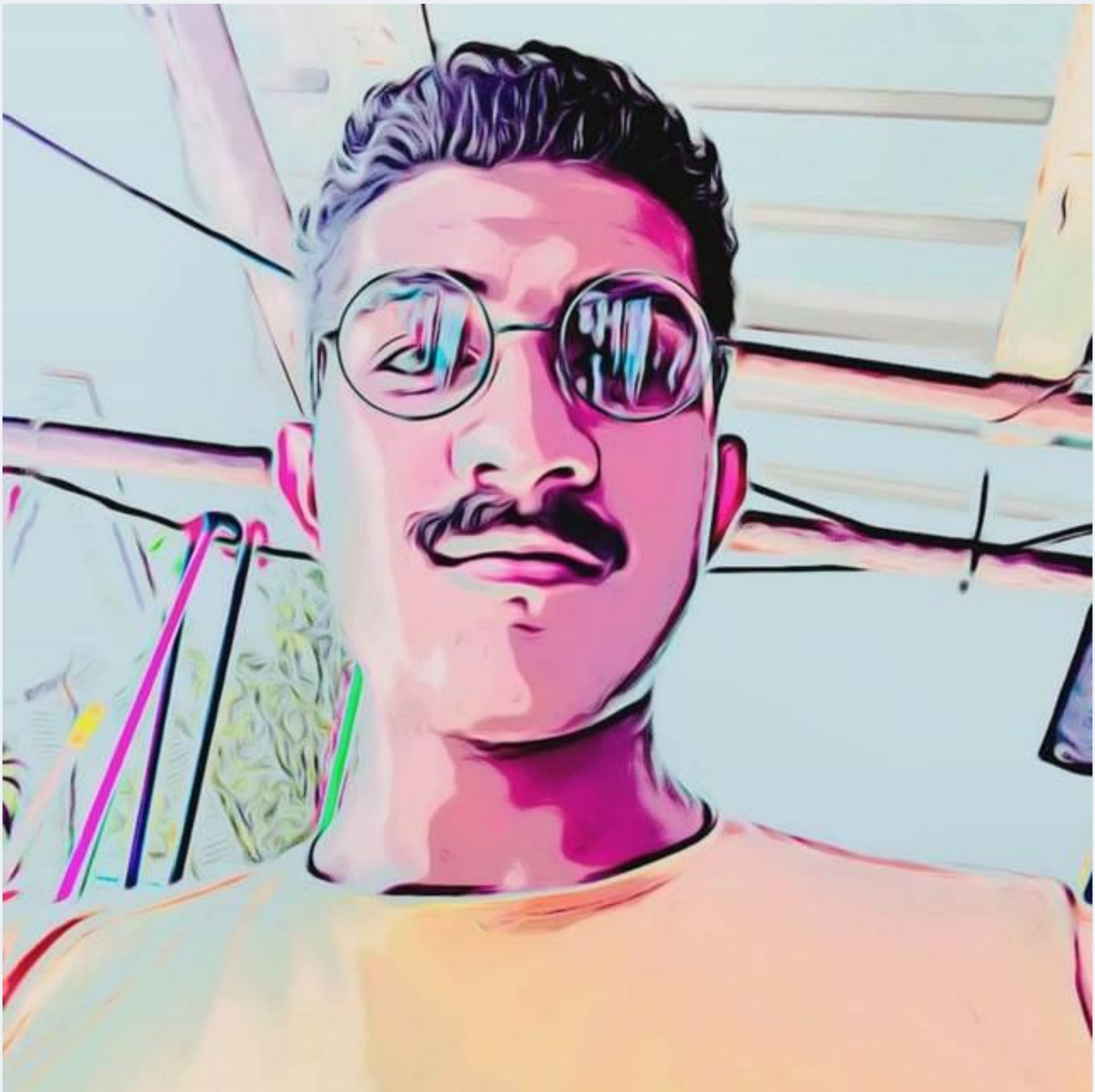
$$p(\vec{x}) = \sum_{i=1}^{K} \phi_i \mathcal{N}(\vec{x} \mid \vec{\mu}_i, \Sigma_i)$$

$$\mathcal{N}(\vec{x} \mid \vec{\mu}_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^K|\Sigma_i|}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu}_i)^{\mathrm{T}}\Sigma_i^{-1}(\vec{x}-\vec{\mu}_i)\right)$$

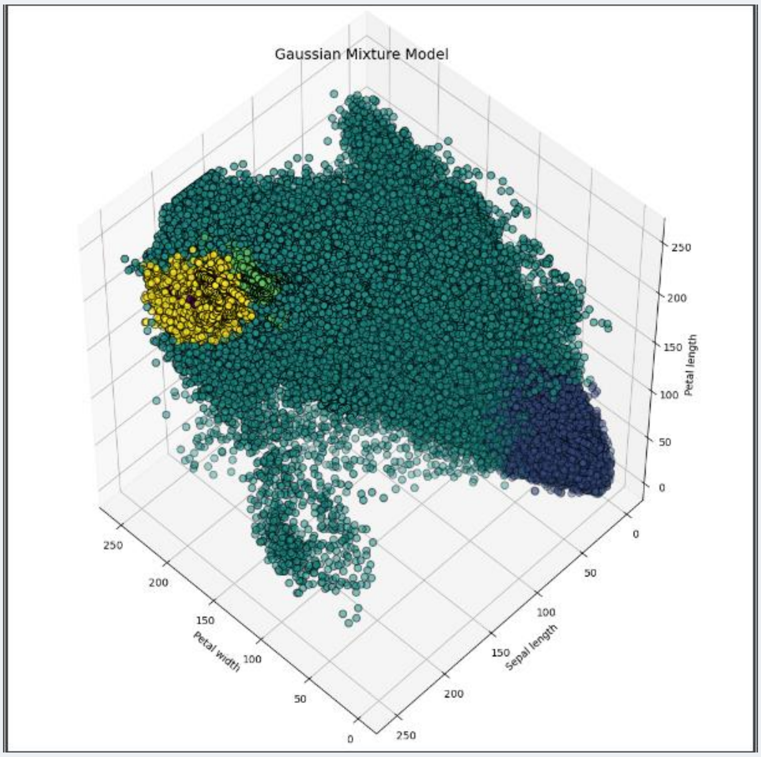
$$\sum_{i=1}^{K} \phi_i = 1$$

now look at the our result:

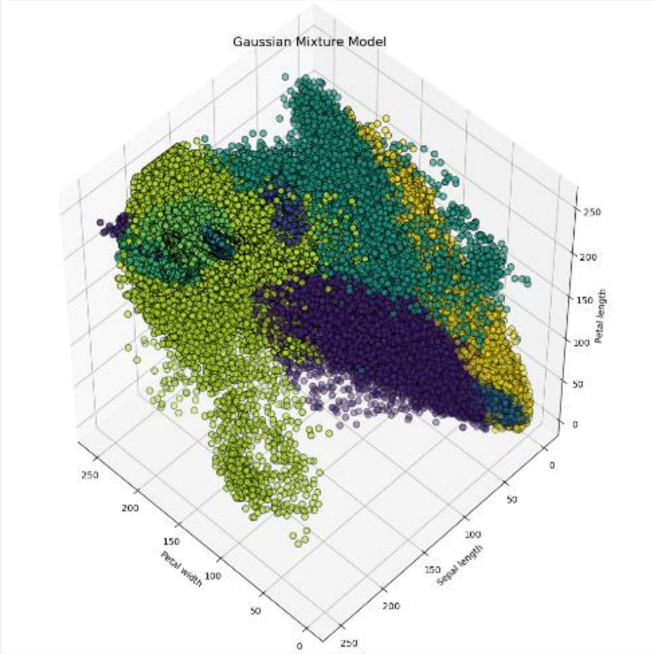
we use photo for data input:



And we use Gaussian with 5 centra:



And now change number of centra to 10:



As we can see the result doesn't change a lot but the time of calculate are grows
We repeat this scenario with 15 centra and but the result change a little :)

We can find out with grow the number of centra we can reduce the error but From time to time the error reduction percentage is so low that it can be neglected but has a high computational burden on the computer.

Thank you :)