



**SNS COLLEGE OF ENGINEERING**

**(Autonomous)**

**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING**



**Artificial Intelligence & Machine Learning**

**Unit 3 – Unsupervised Learning**

**Expectation-Maximization**

**Algorithm**

**Prepared by,**

**P.Ramya**

**Assistant Professor/ECE**

**SNS College of Engineering**



# Expectation-Maximization Algorithm

The Expectation-Maximization Algorithm, or EM algorithm for short, is an approach for maximum likelihood estimation in the presence of latent variables. A general technique for finding maximum likelihood estimators in latent variable models is the expectation-maximization (EM) algorithm.



Contd...

The basic two steps of the EM algorithm

- E-step and M-step are often pretty easy for many of the machine learning problems in terms of implementation.



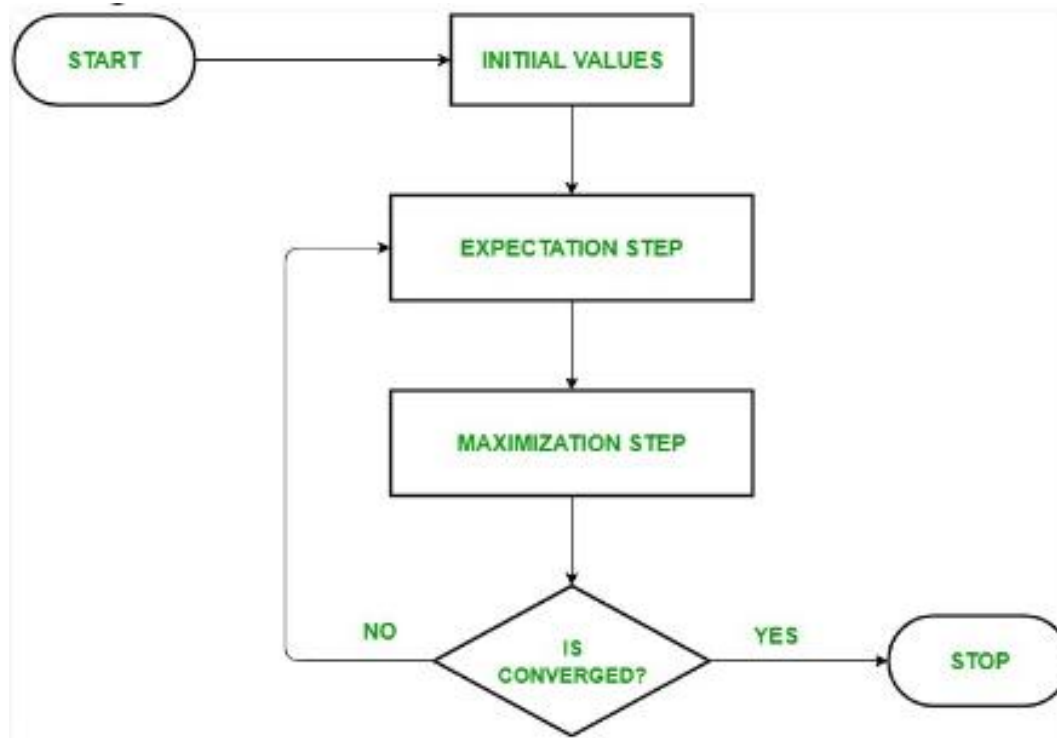
# Algorithm:

- Given a set of incomplete data, consider a set of starting parameters.
- Expectation step (E – step): Using the observed available data of the dataset, estimate (guess) the values of the missing data.
- Maximization step (M – step): Complete data generated after the expectation (E) step is used in order to update the parameters.
- Repeat step 2 and step 3 until convergence.



# Elbow Method

In the Elbow method, we are actually varying the number of clusters (  $K$  ) from 1 - 10. For each value of  $K$ , we are calculating WCSS (Within-Cluster Sum of Square). WCSS is the sum of squared distance between each point and the centroid in a cluster.



# Implementation

Usage of EM algorithm –

- It can be used to fill the missing data in a sample.
- It can be used as the basis of unsupervised learning of clusters.
- It can be used for the purpose of estimating the parameters of Hidden Markov Model (HMM).
- It can be used for discovering the values of latent variables.



# Advantages of EM algorithm

- It is always guaranteed that likelihood will increase with each iteration.
- The E-step and M-step are often pretty easy for many problems in terms of implementation.
- Solutions to the M-steps often exist in the closed form.



# Disadvantages of EM algorithm

- It has slow convergence.
- It makes convergence to the local optima only.
- It requires both the probabilities, forward and backward (numerical optimization requires only forward probability).





Thank you