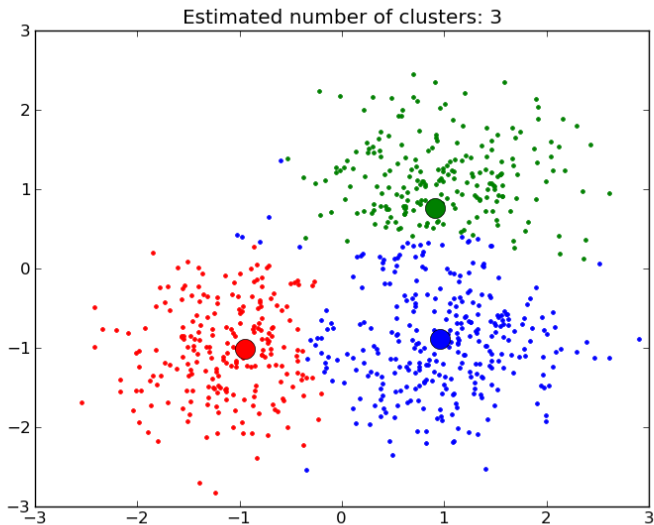


Big Data Analytics

K-Means

K-Means

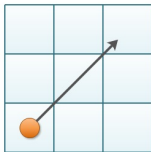
Algorithmme EM



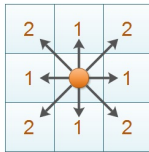
K-Means

distance euclidienne, Manhattan, Chebychev

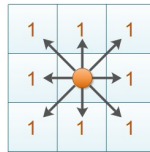
Euclidean Distance



Manhattan Distance



Chebyshev Distance



$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad |x_1 - x_2| + |y_1 - y_2| \quad \max(|x_1 - x_2|, |y_1 - y_2|)$$

Gaussian Mixture Model

Les clusters sont représentés par un centre et une matrice de covariance.

