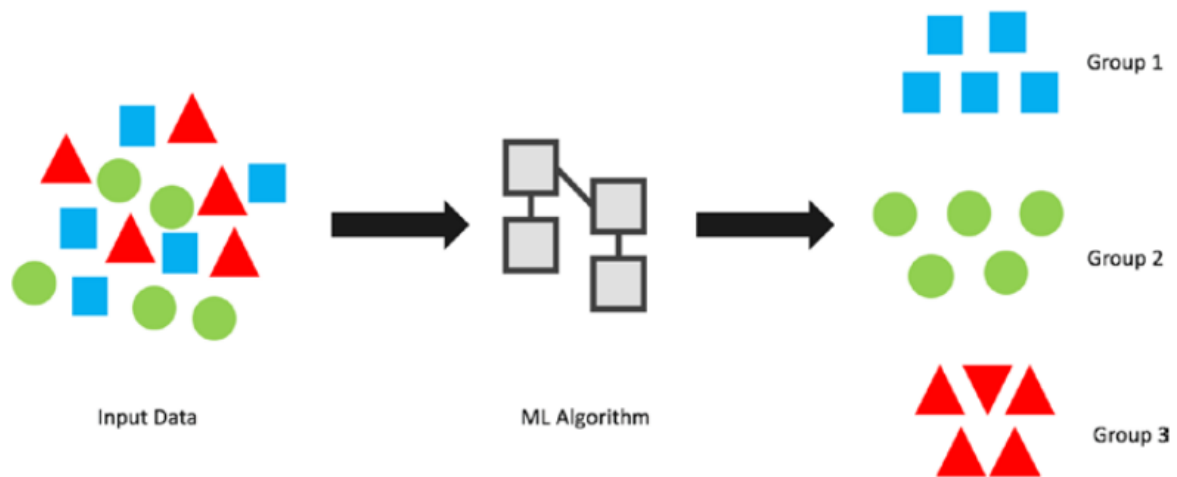


POINTS IMPORTANTS CHAPITRE 7 : UNSUPERVISED MACHINE LEARNING

- L'apprentissage automatique non supervisé n'inclut pas la recherche de relations entre l'entrée et la sortie. il essaie plutôt de regrouper des valeurs similaires les unes aux autres.
- L'apprentissage non supervisé peut être classé en deux catégories distinctes :
 1. Regroupement
 2. Règles d'association
- Le clustering fait référence à la recherche de groupes sous-jacents de points de données, en fonction des attributs présents dans les données



- Le clustering peut avoir plusieurs applications, dont les suivantes :
 1. Détection d'anomalies
 2. Maintenance prédictive
 3. Segmentation de la clientèle
- Les algorithmes k-means permet de créer des clusters pour des ensembles.
- La méthode Elbow, permet de tracer la somme totale des erreurs au carré dans un cluster par rapport au nombre de clusters. Il donne une idée du bon nombre de groupes qui peuvent être formés à partir des données fournies.
- Il est toujours judicieux de réduire les données avant d'appliquer le clustering car le clustering peut être sensible à l'échelle des données, pour la simple raison qu'il utilise une métrique de distance pour comparer la similarité entre deux valeurs (multi dimensionnalité)
-