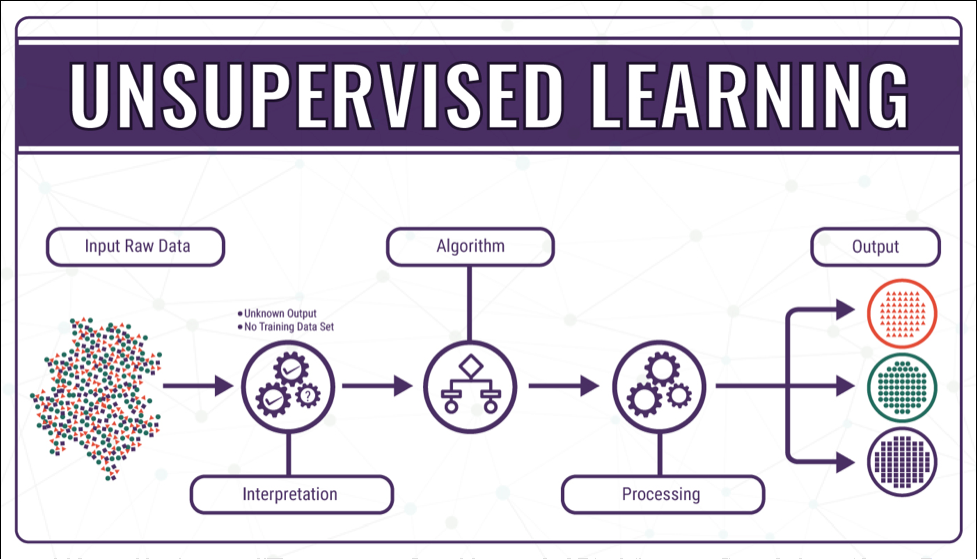
**Chương 6: Các thuật toán học tập không giám sát ( UNSUPERVISED LEARNING)**

6.1 Giới thiệu học tập không giám sát (unsupervised learning)

* Ngược lại supervised learning, unsupervised learning( học không giám sát) là một nhóm thuật toán sử dụng dữ liệu không có nhãn
* Trong học không giám sát, các máy tính sẽ sử dụng dữ liệu không gắn nhãn và tự học, mà không cần có sự giám sát nào bên ngoài.



* Các thuật toán theo cách tiếp cận này hướng đến việc mô hình hóa được cấu trúc hay thông tin ẩn trong dữ liệu
* Hay nói cách khác, sử dụng các phương pháp này thiên về việc mô tả tính chất hay đặc tính của dữ liệu
* Thông thường, các thuật toán này dựa trên những thông tin sau:
* Mối quan hệ tương tự( similarity) giữa các ví dụ được( được gọi là instance) trong dữ liệu như trong các thuật toán clustering( phân cụm).
* Xác suất đồng xuất hiện của các đối tượng như trong Association mining.
* Các phép biến đổi ma trận để trích xuất các đặc trưng như PCA, SVD