Setelah kita memberikan sekumpulan data tanpa label, model machine learning akan mempelajari pola dan struktur pada data berdasarkan hubungan atau keterkaitan antar variabel pada data. Model kemudian akan mengelompokkan data ini ke dalam beberapa klaster yang berbeda. Teknik ini disebut sebagai clustering.

Contoh kasus untuk teknik clustering adalah customer segmentation. Dari data ribuan pengunjung sebuah website ecommerce, model akan belajar sendiri untuk mengelompokkan pengunjung. Bisa berdasarkan waktu kunjungan, lama kunjungan, penggunaan fitur search, jumlah klik, dan sebagainya. Model unsupervised learning akan menentukan segmen market dan mengelompokkan pengunjung  ke dalam segmen market yang berbeda. Dengan output dari model ini, pengelola ecommerce dapat menentukan strategi untuk meningkatkan penjualan atau strategi lain yang dirasa perlu diambil untuk kelanjutan bisnis.

Metode unsupervised learning yang sekarang sedang sangat populer adalah generative adversarial networks (GANs). Terinspirasi dari teori game, GAN bekerja dengan cara membuat dua jaringan syaraf tiruan berkompetisi. Lapan [7] dalam bukunya menyatakan, ketika kita memiliki dua jaringan syaraf yang bersaing, jaringan pertama mencoba menghasilkan data palsu untuk mengelabui jaringan kedua, sedangkan jaringan kedua mencoba untuk membedakan data palsu tersebut dari data sampel kumpulan data kita. Seiring waktu, kedua jaringan menjadi semakin ahli dalam tugas-tugas mereka dengan menangkap pola spesifik yang halus dalam kumpulan data.

Beberapa algoritma unsupervised learning yang penting untuk Anda ketahui adalah: clustering, dimensionality reduction, anomaly detection, dan density estimation.

Anda siap untuk membuat model unsupervised learning? Kita akan mulai dengan K-Means Clustering dan Dimensionality Reduction (LDA, PCA, t-SNE) di modul selanjutnya di kelas ini. Siap-siap ya.