

### **Principal Component Analysis:**

Principal Component Analysis adalah teknik reduksi dimensi yang banyak digunakan dalam machine learning dan statistik. Tujuan utamanya adalah mengubah dataset berdimensi tinggi menjadi ruang dimensi yang lebih rendah sambil tetap mempertahankan sebanyak mungkin varians. PCA mengidentifikasi komponen utama, yang merupakan kombinasi linear dari fitur-fitur asli, dan mengurutkannya berdasarkan variansnya. Dengan memilih subset dari komponen-komponen ini, dapat dilakukan kompresi data dan menyederhanakan analisis selanjutnya. PCA sangat berguna untuk visualisasi data berdimensi tinggi dan mengatasi masalah kutukan dimensi.

### **K-Nearest Neighbors:**

K-Nearest Neighbors adalah algoritma machine learning yang digunakan untuk tugas klasifikasi dan regresi. Dalam konteks klasifikasi, KNN menetapkan titik data ke kelas mayoritas di antara  $k$  tetangganya dalam ruang fitur. Algoritma ini bersifat non-parametrik, yang berarti tidak membuat asumsi tentang distribusi data yang mendasarinya. Pemilihan parameter ' $k$ ' sangat memengaruhi kinerja model, seimbang antara bias dan varians. KNN sensitif terhadap metrik jarak yang digunakan, dan sangat penting untuk menskalakan fitur dengan tepat untuk menghindari pengaruh yang tidak adil dari variabel dengan skala yang lebih besar.

### **Decision Tree:**

Decision Tree adalah algoritma machine learning yang serbaguna digunakan untuk tugas klasifikasi dan regresi. Algoritma ini secara rekursif mempartisi data berdasarkan fitur-fitur, menciptakan struktur seperti pohon dari simpul keputusan dan daun. Setiap simpul internal mewakili keputusan berdasarkan fitur tertentu, dan setiap daun sesuai dengan hasil prediksi. Decision Tree mudah diinterpretasikan dan divisualisasikan, sehingga berharga untuk memahami proses pengambilan keputusan. Namun, rentan terhadap overfitting, terutama dengan pohon yang dalam, dan teknik seperti pruning sering diterapkan untuk mengurangi masalah ini. Varian populer, seperti Random Forests dan Gradient Boosted Trees, meningkatkan kinerja prediksi dengan menggabungkan beberapa decision tree.