

Nama : Danny Hamtar Pangestu  
Nim : 1103210037  
Catatan : principal component analysis (PCA)

## PCA: Principal Component Analysis

PCA adalah metode statistik yang digunakan untuk mereduksi dimensi dari data yang kompleks. Dalam konteks analisis data, "dimensi" merujuk pada jumlah variabel yang ada dalam set data.

1. Tujuan Utama PCA: Mengurangi dimensi data sambil mempertahankan sebanyak mungkin informasi yang dapat dipertahankan.
2. Konsep Inti: PCA mencari kombinasi linear dari variabel asli yang disebut sebagai "komponen utama" yang mampu menjelaskan sebanyak mungkin variabilitas dalam data. Komponen utama adalah vektor dalam ruang data asli yang menunjukkan arah di mana data memiliki variabilitas terbesar.
3. Langkah PCA:
  - Standardisasi Data: Data asli seringkali harus dinormalisasi atau distandardisasi agar variabel-variabel dengan skala yang berbeda memiliki dampak yang seimbang pada hasil analisis PCA.
  - Perhitungan Matriks Kovariansi atau Korelasi: PCA melibatkan perhitungan matriks kovariansi atau korelasi dari data standar.
  - Perhitungan Nilai Eigen dan Vektor Eigen: PCA menggunakan nilai eigen dan vektor eigen dari matriks kovariansi atau korelasi untuk menentukan komponen utama.
  - Reduksi Dimensi: Komponen utama yang pertama akan menjelaskan sebanyak mungkin variabilitas dalam data, yang kedua akan menjelaskan sebanyak mungkin variabilitas yang tersisa, dan seterusnya. Pengguna kemudian dapat memilih jumlah komponen utama yang ingin diambil berdasarkan berapa banyak informasi yang ingin mereka pertahankan.
4. Interpretasi Hasil PCA: Setelah melakukan reduksi dimensi, hasil PCA dapat diinterpretasikan dalam konteks aplikasi spesifik data tersebut. Misalnya, dalam analisis biologis, komponen utama yang mungkin mewakili hubungan antara spesies, dalam pengenalan pola wajah, komponen utama mungkin mewakili fitur wajah yang paling bervariasi, dan seterusnya.