

Nama : Fadhil Dzikri Aqila

NIM : 1103213136

Kelas : TK-45-03

Principal Componen Analysis (PCA)

PCA (Principal Component Analysis) adalah sebuah metode dalam analisis data untuk mengurangi dimensi dari data dengan tujuan mempertahankan informasi yang paling penting. PCA bekerja dengan mengubah variabel-variabel yang saling terkait menjadi sejumlah variabel baru yang tidak saling berkorelasi (komponen utama). Proses ini dimulai dengan melakukan plotting data dan menghitung rata-rata untuk tiap variabel. Setelah itu, dilakukan pergeseran data sehingga titik tengah data berada di titik (0,0), dan kemudian mencari garis terbaik yang sesuai dengan data tersebut.

PCA menggunakan dua metode untuk menentukan garis terbaik: meminimalkan jarak antara titik data dan garis, atau memaksimalkan jarak antara titik proyeksi (hasil proyeksi data ke garis) dan titik asal. PCA memilih metode kedua karena lebih mudah dihitung. Setelah menentukan garis terbaik (Komponen Utama 1 atau PC1), PCA menghitung nilai eigen dan vektor singular untuk garis tersebut. Nilai eigen mewakili seberapa banyak variasi yang dijelaskan oleh PC1, sedangkan vektor singular menunjukkan kontribusi masing-masing variabel terhadap PC1.

Proses ini dilanjutkan untuk menemukan Komponen Utama kedua (PC2), yang tegak lurus dengan PC1 dan juga melewati titik asal. Dari hal tersebut, PCA menghasilkan dua Komponen Utama yang menjelaskan sebagian besar variasi data dan dilakukan visualisasi data berdasarkan PC1 dan PC2. PCA juga dapat digunakan untuk menentukan variabel yang paling penting untuk melakukan pengelompokan data.