

Nama : Muhammad Nadhim Abdirrahman
NIM : 1103210018
Kelas : TK-45-05

Catatan PCA

StatQuest: Principal Component Analysis (PCA), Step-by-Step

PCA (Principal Component Analysis) adalah metode untuk mengurangi dimensi data dengan Singular Value Decomposition (SVD). Langkah-langkahnya termasuk menghitung rata-rata pengukuran, mencari PC1 dan PC2, dan membuat plot PCA. PCA membantu mengelompokkan data dan menyoroti kontribusi gen terhadap perbedaan antar sampel. SVD digunakan untuk menghitung PC dan mengurangi dimensi data serta eigenvalues dari hasil SVD dapat digunakan untuk mengevaluasi variasi yang dijelaskan oleh setiap PC.

Langkah-langkah dasar dalam analisis komponen utama (PCA)

Pertama, data dipetakan dan rata-rata pengukuran untuk setiap gen dihitung untuk menemukan pusat data. Selanjutnya, fokus dipindahkan dari data asli ke grafik dengan menggeser data sehingga pusatnya berada di atas asal. Kemudian, garis acak digunakan untuk mencocokkan data dengan memutar garis tersebut hingga sesuai dengan data sebaik mungkin, yang menghasilkan Komponen Utama 1 (PC1). PC1 menunjukkan sebaran data sepanjang sumbu Gen 1 dan Gen 2. PC2 dicari sebagai garis terbaik yang tegak lurus dengan PC1, dan plot PCA akhir digambar dengan memutar PC1 secara horizontal dan PC2 secara vertikal. Pada tahap ini, titik-titik diproyeksikan untuk menemukan lokasi sampel di plot PCA. Terakhir, menggunakan Singular Value Decomposition, eigenvalues digunakan untuk menentukan proporsi variasi yang dijelaskan oleh setiap komponen utama, dan scree plot dibuat untuk menganalisis variasi tersebut.