

Nama: Alif Firmansyah
NIM : 1103204105
Kelas : TK44G4

Catatan PCA StatQuest

Principal Component Analysis (PCA) adalah teknik analisis data yang membantu mengurangi kompleksitas kumpulan data dengan mengurangi jumlah variabel menjadi beberapa faktor utama. PCA dapat digunakan untuk memvisualisasikan data dan menentukan variabel mana yang paling penting untuk memahami kumpulan data secara keseluruhan.

PCA adalah teknik statistik yang digunakan untuk mengurangi kompleksitas data. Teknik ini memproyeksikan titik-titik data ke dalam sebuah garis, dan kemudian mengukur jarak antara titik-titik tersebut dan garis tersebut. Garis dengan jumlah kuadrat jarak terbesar antara titik proyeksi dan titik asal disebut komponen utama (PC1).

PCA adalah teknik statistik yang digunakan untuk mereduksi jumlah data dalam suatu kumpulan menjadi beberapa dimensi utama. PCA dilakukan dengan menggabungkan berbagai bagian data (sampel gen), dan kemudian menghitung kombinasi linier dari sampel tersebut yang menghasilkan nilai tunggal terbesar. Nilai tunggal ini dapat digunakan untuk menghitung vektor eigen untuk plot PCA, yang memberi tahu peneliti gen mana yang paling penting dalam mendeskripsikan data. Terakhir, skor yang dimuat untuk setiap gen diplot pada plot tersebut, dan peneliti dapat menggunakan informasi ini untuk menggambar plot PCA 2-D.

StatQuest menjelaskan cara menghitung komponen utama, yaitu vektor yang menyumbang persentase besar variasi dalam kumpulan data. Selanjutnya, ini menunjukkan cara menggunakan informasi ini untuk membuat plot data 2D yang disederhanakan. Terakhir, mereka memberikan tutorial cara mengubah plot 2D ini menjadi grafik 3D agar lebih mudah dibaca.