

Principal Component Analysis (PCA) dengan menggunakan Singular Value Decomposition (SVD).

PCA adalah metode yang digunakan untuk mendapatkan pemahaman yang lebih mendalam tentang data. Misalnya, jika kita memiliki data tentang dua gen yang diukur pada enam tikus yang berbeda, kita dapat menggunakan PCA untuk melihat bagaimana tikus-tikus tersebut berkumpul berdasarkan pengukuran gen mereka. PCA adalah metode yang digunakan untuk mengurangi dimensi data. Misalnya, jika kita memiliki empat gen yang diukur, kita tidak dapat memplot data tersebut karena membutuhkan empat dimensi. Namun, dengan PCA, kita dapat mengambil empat pengukuran gen tersebut dan membuat plot 2D. Plot ini akan menunjukkan bagaimana tikus yang serupa berkumpul bersama. PCA bekerja dengan menghitung rata-rata pengukuran untuk setiap gen, kemudian menghitung pusat data. Data kemudian dipindahkan sehingga pusat berada di atas asal dalam grafik. Kemudian, PCA mencoba menyesuaikan garis ke data tersebut. Garis ini diputar sampai sesuai dengan data sebaik mungkin. Garis ini disebut Komponen Utama 1 (PC1). PC1 adalah garis yang paling baik menyesuaikan data. PC2 adalah garis terbaik berikutnya yang melalui asal dan tegak lurus terhadap PC1. Kedua garis ini digunakan untuk membuat plot PCA akhir. Screen plot adalah representasi grafis dari persentase variasi yang masing-masing PC akun. Ini digunakan untuk menentukan berapa banyak PC yang harus digunakan dalam plot PCA akhir.

PCA adalah metode yang sangat berguna untuk mendapatkan pemahaman yang lebih dalam tentang data. Dengan mengurangi dimensi data, kita dapat melihat bagaimana sampel berkumpul bersama dan menentukan gen atau variabel mana yang paling berharga untuk mengelompokkan data.