

**Afif Ibadurrahman Jalaluddin**

**1103210221**

**Catatan PCA**

## **StatQuest: Principal Component Analysis (PCA)**

### **1. Tujuan PCA (*Principal Component Analysis*) :**

PCA digunakan untuk mengurangi dimensi data dengan mempertahankan sebagian besar variannya, sehingga memungkinkan kita untuk memahami struktur data yang kompleks dengan lebih mudah.

### **2. Konsep Dasar PCA:**

- PCA bertujuan untuk mengurangi dimensi data dari dimensi yang lebih tinggi menjadi dimensi yang lebih rendah.
- PCA memanfaatkan *Singular Value Decomposition* (SVD) untuk melakukan analisisnya.
- PCA membantu dalam menemukan pola dan hubungan dalam data yang tidak terlihat secara langsung.

### **3. Implementasi PCA:**

- PCA digunakan untuk mengurangi dimensi data, memproyeksikan data ke ruang dimensi yang lebih rendah, dan mengidentifikasi pola dalam data.
- PCA memungkinkan untuk mereduksi data dari dimensi yang lebih tinggi menjadi dimensi yang lebih rendah, memungkinkan visualisasi yang lebih mudah dan pemahaman yang lebih baik tentang struktur data.
- PCA digunakan untuk mengidentifikasi *principal components* (komponen utama) yang paling penting dalam menjelaskan varian data.
- Dengan memanfaatkan eigenvalues, PCA memungkinkan kita untuk mengetahui seberapa banyak varian data yang dijelaskan oleh setiap *principal component*.

### **4. Interpretasi Hasil PCA:**

- Principal components yang dihasilkan oleh PCA dapat diinterpretasikan untuk mengidentifikasi variabel atau fitur yang paling penting dalam mempengaruhi struktur data.
- Penggunaan *scree plot* membantu dalam mengevaluasi kontribusi masing-masing *principal component* terhadap varian total data.

### **5. Proses PCA:**

- PCA dimulai dengan merata-ratakan data untuk memusatkan pada asal.
- Kemudian PCA menemukan garis terbaik yang melewati pusat data, disebut sebagai *principal component* pertama (PC1).
- PC1 diikuti dengan menemukan *principal component* kedua (PC2), yang saling tegak lurus dengan PC1.
- Proses ini dapat diulangi untuk mendapatkan *principal component* berikutnya jika diperlukan.

#### **6. Aplikasi PCA:**

- PCA dapat diterapkan dalam berbagai bidang seperti genetika, ilmu sosial, ekonomi, dan lainnya untuk menganalisis dan memahami struktur data yang kompleks.

#### **7. Pentingnya Visualisasi:**

- PCA memungkinkan visualisasi data yang kompleks menjadi lebih sederhana dan mudah dipahami.
- PCA memungkinkan penggambaran data dalam ruang dimensi yang lebih rendah untuk membantu dalam identifikasi pola dan hubungan antar variabel.

Pemahaman yang mendalam tentang PCA dapat membantu dalam analisis data, visualisasi, dan pemahaman struktur data yang kompleks.