

Nama :Juan Anemao Sokhi Zidomi

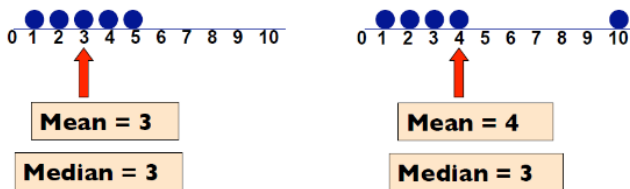
NIM :1103204007

Kelas :TK-44-G4

Numerical Summaries of data

- Median
Nilai Tengah dari data dan berguna untuk distribusi miring
- Mean
Nilai rata-rata dari data untuk distribusi outlier atau pencilan

Cara menentukan lokasi yang diukur yaitu



Dari gambar paling kiri karna dapat disimpulkan tidak ada outlier maka mean dan median dapat ditentukan, sedangkan gambar kanan ada outlier maka sama seperti perhitungan gambar kiri

Multivariate Analysis (MANOVA) beberapa proses yaitu

- Plotting Multiverse Data
- Calculating Summary of Statistics
- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

LDA(Linear Discriminant Analysis)->Menemukan linear dari fitur-fitur yang mencirikan atau memisahkan dua atau lebih kelas objek atau peristiwa

PCA(Principal Component Analysis)->Menggunakan transformasi ortogonal untuk mengubah sekumpulan fitur yang lebih besar menjadi sekumpulan fitur yang lebih kecil

SVD(Singular Value Decomposition)-> menggunakan dekomposisi matriks untuk menemukan sumbu proyeksi terbaik dengan kesalahan rekonstruksi minimum

Perbedaan LCA dan PCA yaitu

LCA ->Memaksimalkan komponen axes untuk class-separation

PCA ->sumbu komponen yang memaksimalkan Varians

Algoritma LDA yaitu

- Fisher Discriminant Analysis
Aturan Klasifikasi Gaussian dengan probabilitas awal yang sama
- LDA by Optimal Scoring
Menggunakan multivariasi linear regresi pada respon turunan

Kunci untuk interpretasi PCA(Principal Component Analysis)

1. Varians adalah ukuran deviasi dari rata-rata dalam satu dimensi atau ukuran penyebaran
2. Covariance adalah ukuran seberapa besar dimensi bervariasi dari rata-rata terhadap satu sama lain
 - Jika kovarians bernilai positif maka kedua peningkatan dan penurunan secara bersamaan
 - Jika kovarians bernilai negatif satu meningkat dan yang lain menurun
 - Jika 0 dua dimensi tidak bergantung
3. Matriks Kovarians adalah matriks kombinasi matriks
4. Nilai Eigen dan Vektor eigen

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

PCA Digunakan untuk:

1. Mengurangi angka dari dimensi data
2. Menemukan pola dalam data dimensi tinggi
3. Memvisualisasikan data dengan dimensi tinggi

Singular Value Decomposition (SVD)

- Memungkinkan representasi yang tepat dari matriks apapun dan juga memudahkan untuk menghilangkan bagian yang kurang penting dari representasi tersebut untuk menghasilkan representasi perkiraan dengan jumlah dimensi yang diinginkan
- Dengan kata lain :Memberikan minimum reconstruction error