

Nama : Arfara Yema Samgusdian

NIM : 1103202004

Kelas : TK-44-G4

### Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) adalah proses menganalisis data untuk memahami karakteristiknya, menemukan pola, dan mengidentifikasi anomali. EDA dapat digunakan untuk berbagai tujuan, seperti:

- Memahami struktur data dan hubungan antara variabel
- Mengidentifikasi outlier dan kesalahan input
- Menemukan pola dan tren dalam data
- Memilih variabel yang penting untuk analisis lebih lanjut
- Mengembangkan hipotesis tentang data

EDA dapat menggunakan berbagai metode analisis, seperti:

- Visualisasi data, seperti histogram, scatter plot, dan box plot
- Statistik deskriptif, seperti mean, median, dan variance
- Korelasi dan regresi
- Deteksi anomali

### Principal Component Analysis (PCA)

Principal Component Analysis (PCA) adalah teknik reduksi dimensi yang digunakan untuk mengurangi jumlah variabel dalam suatu dataset tanpa kehilangan informasi yang terlalu banyak. PCA bekerja dengan menemukan komponen utama dari suatu dataset, yaitu kombinasi linear dari variabel-variabel asli yang menjelaskan sebagian besar varians dalam data.

PCA dapat digunakan untuk berbagai tujuan, seperti:

- Mengurangi dimensi data sebelum melakukan analisis machine learning
- Meningkatkan akurasi model machine learning
- Mengidentifikasi fitur-fitur yang paling penting dalam suatu dataset
- Mendeteksi anomali
- Visualisasi data berdimensi tinggi

### Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) adalah teknik reduksi dimensi yang digunakan untuk memaksimalkan perbedaan antara dua atau lebih kelas dalam suatu dataset. LDA bekerja dengan

menemukan proyeksi data ke ruang dimensi yang lebih rendah di mana kelas-kelas menjadi lebih terpisah.

LDA dapat digunakan untuk berbagai tujuan, seperti:

- Mengurangi dimensi data sebelum melakukan analisis machine learning
- Meningkatkan akurasi model machine learning untuk klasifikasi
- Mendeteksi anomali

### Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) adalah teknik dekomposisi matriks yang dapat digunakan untuk berbagai tujuan, termasuk reduksi dimensi, dekomposisi pangkat rendah, dan faktorisasi matriks. SVD bekerja dengan mendekomposisi suatu matriks menjadi tiga matriks, yaitu matriks U, matriks S, dan matriks V. Matriks S berisi singular value dari matriks asli, yang merupakan nilai-nilai yang menjelaskan sebagian besar varians dalam data.

SVD dapat digunakan untuk berbagai tujuan, seperti:

- Melakukan PCA
- Melakukan LDA
- Mengisi nilai yang hilang dalam data
- Melakukan rekomendasi sistem
- Melakukan kompresi data

EDA, PCA, LDA, dan SVD adalah teknik-teknik yang sangat penting dalam machine learning. EDA dapat digunakan untuk memahami data dan mengidentifikasi fitur-fitur yang penting untuk analisis lebih lanjut. PCA dan LDA dapat digunakan untuk mengurangi dimensi data sebelum melakukan analisis machine learning, yang dapat meningkatkan akurasi model dan mengurangi waktu komputasi. SVD dapat digunakan untuk melakukan PCA dan LDA, serta berbagai tugas machine learning lainnya.

Berikut adalah beberapa contoh penerapan EDA, PCA, LDA, dan SVD dalam machine learning:

- EDA dapat digunakan untuk memahami data citra satelit sebelum melakukan analisis klasifikasi untuk mengidentifikasi jenis-jenis lahan.
- PCA dapat digunakan untuk mengurangi dimensi data genomik sebelum melakukan analisis regresi untuk memprediksi risiko penyakit.
- LDA dapat digunakan untuk mengurangi dimensi data teks sebelum melakukan analisis klasifikasi untuk mengidentifikasi sentimen suatu teks.
- SVD dapat digunakan untuk melakukan rekomendasi sistem dengan merekomendasikan produk kepada pengguna berdasarkan riwayat pembelian mereka.

Secara umum, EDA, PCA, LDA, dan SVD adalah teknik-teknik yang sangat penting dalam machine learning yang dapat digunakan untuk berbagai tujuan, seperti:

- Memahami data
- Mengurangi dimensi data
- Meningkatkan akurasi model machine learning
- Mengurangi waktu komputasi
- Melakukan berbagai tugas machine learning lainnya