

Muhammad Harits

1103223159

Bab 8: Dimensionality Reduction (Pengurangan Dimensi)

Ringkasan Bab

Bab ini membahas teknik untuk mengurangi jumlah fitur dalam dataset. Tujuannya termasuk mempercepat training, mengurangi noise, serta mempermudah visualisasi data.

Konsep Utama

- Curse of Dimensionality
Semakin tinggi dimensi data, semakin sparse (renggang), sehingga sulit menemukan pola dan risiko overfitting meningkat.
- Pendekatan Utama
 1. Projection → memproyeksikan data ke hyperplane berdimensi lebih rendah
 2. Manifold Learning → berasumsi data sebenarnya berada di manifold berdimensi lebih rendah
- PCA (Principal Component Analysis)
 - Menemukan sumbu yang mempertahankan variansi terbesar.
 - Digunakan untuk kompresi, visualisasi, preprocessing.
- Kernel PCA
 - PCA versi non-linear menggunakan Kernel Trick.
- LLE (Locally Linear Embedding)
 - Contoh Manifold Learning
 - Menjaga hubungan lokal antar titik data.