

Chapter 8 : Summary

Diedrick Darrell Darmadi - 1103223031

Bab 8 membahas konsep **pengurangan dimensi**, yaitu proses mereduksi jumlah fitur dalam dataset sambil tetap mempertahankan struktur informasi penting. Tujuannya adalah mengatasi masalah “curse of dimensionality”, mempercepat komputasi, menyederhanakan visualisasi, dan kadang meningkatkan performa model. Bab ini memperkenalkan dua pendekatan besar: **proyeksi** dan **manifold learning**. Pada pendekatan proyeksi, teknik seperti **Principal Component Analysis (PCA)** digunakan untuk memproyeksikan data ke ruang berdimensi lebih rendah dengan tetap mempertahankan variasi terbesar. PCA dijelaskan mulai dari konsep eigenvector–eigenvalue, komponen utama, proporsi variansi, sampai varian PCA acak (Randomized PCA). Bab ini juga membahas **Incremental PCA** untuk dataset besar serta **Kernel PCA**, yang memungkinkan reduksi dimensi pada data non-linear menggunakan konsep kernel trick.

Selain proyeksi linear, bab ini membahas **manifold learning**, yaitu ide bahwa data berdimensi tinggi sebenarnya berada pada struktur berdimensi lebih rendah (manifold). Beberapa metode diperkenalkan, seperti **Locally Linear Embedding (LLE)** yang mempertahankan hubungan lokal antar titik, serta dibahas pula kelebihan dan keterbatasannya. Bab ini kemudian menyinggung berbagai algoritma manifold lain seperti Isomap, MDS, dan t-SNE (walau t-SNE hanya muncul ringan sebagai konteks). Bab ditutup dengan pembahasan cara memilih metode reduksi dimensi yang tepat, trade-off antara linear vs non-linear, serta cara mengintegrasikan reduksi dimensi dalam pipeline machine learning.