

Nama : Bagus Mahardika Santoso

NIM : 1103204028

- Principal Component Analysis atau bisa disebut PCA merupakan metode reduksi dimensi yang paling umum yang digunakan. Tujuannya adalah untuk mengidentifikasi dan menampilkan data dalam bentuk yang lebih sederhana sambil mempertahankan sebagian besar data asli.

PCA bekerja dengan asumsi bahwa sebagian besar informasi dalam sebuah dataset terkandung dalam variansinya. Untuk melakukan ini, PCA mencari "arah" di mana varian data maksimal, yang disebut sebagai komponen utama pertama. Kemudian, PCA mencari arah yang ortogonal terhadap komponen pertama dan mengambil varian maksimal sebagai komponen utama kedua. Setelah sejumlah elemen utama yang diinginkan ditemukan atau semua elemen asli data diakomodasi, proses ini berakhir.

Dengan menggunakan PCA dapat meningkatkan visualisasi data, meningkatkan efisiensi komputasi, dan mengurangi masalah keterbatasan dimensi yang mungkin muncul dalam beberapa algoritma dengan mengurangi dimensi data. Tetapi perlu diingat bahwa PCA bersifat linear, yang berarti mungkin tidak cocok untuk data dengan hubungan non-linear. Selain itu, interpretasi komponen utama, yang merupakan kombinasi dari fitur asli, mungkin tidak selalu mudah dipahami karena PCA mengurangi jumlah fitur.

- Linear Discriminant Analysis (LDA) digunakan untuk menemukan kombinasi linear dari fitur yang paling membedakan dua atau lebih kelas objek atau peristiwa. LDA memiliki tujuan untuk memaksimalkan variasi antar kelas dan meminimalkan variasi dalam kelas. Linear Discriminant Analysis (LDA) mencari arah di ruang fitur yang paling membedakan antara kelas. Berbeda dengan Principal Component Analysis (PCA), yang berfokus pada maksimisasi varian total dalam dataset, LDA adalah metode yang diawasi karena memerlukan label kelas sebagai input.

Terdapat 2 cara penggunaan LDA, yaitu:

- Variansi Dalam Kelas, LDA mencari arah di mana variansi dalam masing-masing kelas adalah minimum. Hal ini berarti bahwa data dari setiap kelas harus dikumpulkan semirip mungkin dalam representasi LDA.
- Variansi Antar Kelas, LDA berusaha memaksimalkan jarak antara rata-rata kelas-kelas berbeda. Ini membantu meningkatkan kemampuan pemisahan antara kelas.

Tugas klasifikasi dapat menggunakan LDA secara luas, dan teknik prapemrosesan juga digunakan untuk meningkatkan kinerja model. LDA memiliki beberapa batasan utama. Pertama, asumsinya bersifat linier, yang berarti mungkin kurang efektif untuk data dengan hubungan non-linier. Selain

itu, asumsinya bahwa kovarian dari setiap kelas adalah sama, yang mungkin tidak selalu benar dalam aplikasi dunia nyata. Secara keseluruhan, LDA sangat efektif dalam reduksi dimensi dan klasifikasi, terutama ketika asumsi-asumsinya dipenuhi.

- Singular Value Decomposition (SVD) adalah teknik dalam aljabar linear yang memiliki berbagai aplikasi dalam berbagai bidang. SVD memecah suatu matriks menjadi tiga matriks lebih sederhana yaitu dua matriks ortogonal yang merepresentasikan transformasi rotasi dan refleksi, dan satu matriks diagonal yang berisi nilai-nilai singular. Nilai-nilai singular ini menggambarkan "kepentingan" dari setiap komponen atau faktor dalam data. Sebuah matriks, biasanya matriks data, diuraikan menjadi tiga matriks lebih sederhana: matriks singular, matriks kiri, dan matriks kanan. SVD memiliki keunggulan besar dalam memungkinkan analisis komponen tersembunyi atau pola dominan dalam data, bahkan ketika data tersebut sangat kompleks dan multidimensional.

SVD memungkinkan untuk mengurangi dimensi dari dataset yang besar tanpa kehilangan banyak informasi. Komponen-komponen utama dari SVD, yang terkait dengan nilai-nilai singular terbesar, menyimpan sebagian besar varian data. Dengan mempertahankan sejumlah komponen utama teratas, kita dapat membangun representasi yang lebih sederhana namun masih mewakili dengan baik karakteristik dataset asli.

Matriks singular hasil dari SVD menggambarkan tingkat signifikan variabilitas atau informasi yang terkandung dalam data asli. Nilai-nilai singular dalam matriks singular adalah bilangan real non-negatif yang menggambarkan sejauh mana setiap komponen matriks tersebut berkontribusi terhadap variasi dalam data. Semakin besar nilai singular, semakin besar kontribusi komponen tersebut.

Matriks kiri dan matriks kanan dari SVD mengandung vektor-vektor yang digunakan untuk merekonstruksi matriks asli. Matriks kiri berisi vektor-vektor yang berhubungan dengan baris-baris dalam matriks asli, sedangkan matriks kanan berhubungan dengan kolom-kolom. SVD memungkinkan kita untuk mereduksi dimensi data, mengidentifikasi pola, dan melakukan kompresi data. Dalam konteks machine learning, SVD sering digunakan untuk pemrosesan teks, analisis faktor, reduksi dimensi, dan berbagai tugas lain yang melibatkan matriks data kompleks.