

- Latent Dirichlet Allocation (LDA) adalah sebuah model statistik yang digunakan untuk menganalisis topik dalam koleksi dokumen. LDA mengasumsikan bahwa setiap dokumen adalah campuran dari beberapa topik yang ada secara latent, dan setiap kata dalam dokumen adalah hasil dari salah satu topik tersebut. Melalui teknik analisis probabilistik, LDA membantu mengidentifikasi distribusi topik dalam dokumen dan distribusi kata dalam setiap topik, memungkinkan pemrosesan bahasa alami, klasifikasi dokumen, pengelompokan topik, dan analisis sentimen.
- Principal Component Analysis (PCA) adalah metode statistik yang digunakan untuk mereduksi dimensi dari data multivariat dengan tujuan mengidentifikasi pola yang paling signifikan dalam data tersebut. PCA mengubah data asli ke sistem koordinat baru yang disebut komponen utama, di mana komponen utama pertama mengandung varians data yang paling besar, diikuti oleh yang kedua, dan seterusnya. Dengan demikian, PCA membantu mengurangi kerumitan data dengan mempertahankan informasi yang paling penting, yang berguna dalam visualisasi data, menghilangkan korelasi antaratribut, dan memfasilitasi analisis data lebih lanjut.
- Singular Value Decomposition (SVD) adalah sebuah teknik matematis yang digunakan untuk menganalisis dan mereduksi dimensi dalam matriks. SVD memecah matriks menjadi tiga matriks terpisah, yaitu matriks singular, matriks ortogonal kiri, dan matriks ortogonal kanan. Ini memiliki berbagai aplikasi, termasuk analisis data, reduksi dimensi, kompresi data, pemulihan informasi dari data yang hilang atau tercemar, serta dalam berbagai algoritma pembelajaran mesin, seperti dalam metode Principal Component Analysis (PCA). Dengan SVD, informasi yang signifikan dalam data dapat diekstraksi, dan data yang rumit dapat direpresentasikan dengan cara yang lebih sederhana dan informatif.