MASTERY IN DIMENSION REDUCTION

Introduction

- Start your Dimensionality Reduction journey today

What You'll Learn

This course is designed to provide practical examples and hands-on exercises to reinforce the concepts. Additionally, encourage students to work on a mini-project or real datasets to gain practical experience in dimension Reduction.

Highlights

- **Course Duration - 3 days (24 Hours)**
- **Number of Modules - 10**

COURSE CONTENT

Module 1: Introduction to Dimensionality Reduction

- What is Dimensionality Reduction?
- Why is Dimensionality Reduction required?
- Importance in Machine Learning
- Curse of Dimensionality
- Common Dimensionality Reduction Techniques
- Principal Component Analysis (PCA) - An Overview
- Curse of Dimensionality - Risk of Overfitting
- Problem of Multicollinearity, lead to Overfitting
- Dimensionality reduction reduces dimension and not loose any information
- Definition, Type of Dimensionality Reduction Technique
    i. PCA
    ii. LDA
    iii. T-sne
- Correlation - Strength and Relationship between two variables
        i. Spearman Correlation
        ii. Pearson Correlation
    - Collinearity & Multi Collinearity
    - Technique to check collinearity : VIF

Module 2: Principal Component Analysis (PCA)

- PCA Algorithm and Intuition
- Variance and Covariance
- Eigenvalues and Eigenvectors
- Implementing PCA in Python

Module 3: Practical Implementation of PCA

- Choosing the Number of Principal Components
- Reducing Data with PCA
- PCA for Visualization
- Applications in Real-world Datasets

Module 4: Alternative Techniques - t-SNE and UMAP

- t-Distributed Stochastic Neighbor Embedding (t-SNE)
- Uniform Manifold Approximation and Projection (UMAP)
- Comparing PCA, t-SNE, and UMAP
- Use Cases and Limitations

Module 5: Linear Discriminant Analysis (LDA)

- Introduction to LDA
- Within-class and Between-class Scatter
- Solving the Eigenvalue Problem
- LDA vs. PCA

Module 6: Non-negative Matrix Factorization (NMF)

- Introduction to NMF
- Multiplicative Update Rule
- NMF for Topic Modeling and Image Processing
- NMF in Scikit-Learn

Module 7: Autoencoders for Dimension Reduction

- Introduction to Autoencoders
- Encoder and Decoder Networks
- Variational Autoencoders (VAE)
- Applications in Anomaly Detection

Module 8: Kernel PCA and Feature Selection

- Kernel Principal Component Analysis (Kernel PCA)
- Feature Selection vs. Dimensionality Reduction
- Wrapper and Filter Methods
- Embedded Methods for Feature Selection

Module 9: Dimensionality Reduction for Big Data

- Challenges of Dimension Reduction with Big Data
- Random Projections and Sparse Techniques
- Incremental PCA and Mini-batch PCA
- Spark MLlib for Large-scale Data

Module 10: Use Cases and Advanced Topics

- Dimension Reduction in Natural Language Processing (NLP)
- Genomic Data Dimensionality Reduction
- Visualization Tools for High-dimensional Data