Dimensionality Reduction in Machine Learning

CAREER TREK

Dimensionality reduction is a technique in machine learning used to reduce the number of features in a dataset while retaining as much information as possible.

Dimensionality reduction can be achieved through techniques such as Principal Component Analysis (PCA), Singular Value Decomposition (SVD), or Linear Discriminant Analysis (LDA).

This is useful in situations where the dataset has many features, as this can lead to computational inefficiencies and reduce the performance of machine learning algorithms.

PCA

It is a technique that aims to reduce the dimensionality of the data by identifying the most important features that explain the variance in the data. It finds a new set of orthogonal variables (called principal components) that capture the maximum amount of variance in the original data, while minimizing the loss of information.

SVD

It is a matrix factorization technique that decomposes a matrix into its singular values and singular vectors. It is used to compress the data, remove noise, and identify patterns in the data. SVD is often used as a pre-processing step for other machine learning algorithms.

LDA

It is a supervised dimensionality reduction technique that aims to find the linear combination of features that best separates different classes in the data. It projects the data onto a lower-dimensional space while preserving the class separability.

Summary

In summary, PCA and SVD are unsupervised dimensionality reduction techniques that aim to find the most important features that capture the variance in the data or identify patterns, while LDA is a supervised technique that aims to find the features that best separate the classes in the data.

If you want to know more,
leave a comment below
and follow me on
LinkedIN!

CAREER
TREK