# Dimensionality Reduction - Questions and Answers

## 91. Why is dimensionality reduction important in machine learning?

Dimensionality reduction is important because it helps reduce the number of input variables in a dataset, which can improve model performance, reduce overfitting, and decrease computational cost. It also helps in visualizing high-dimensional data.

## 92. Explain the concept of Principal Component Analysis (PCA).

PCA is a statistical technique used to reduce the dimensionality of a dataset by transforming the original variables into a new set of uncorrelated variables called principal components. These components capture the maximum variance in the data.

## 93. What is t-SNE, and how is it used for dimensionality reduction?

t-SNE (t-distributed Stochastic Neighbor Embedding) is a non-linear dimensionality reduction technique particularly well-suited for visualizing high-dimensional data in 2 or 3 dimensions. It preserves local structure and is commonly used for data exploration.

## 94. Describe the curse of dimensionality.

The curse of dimensionality refers to the various problems that arise when analyzing and organizing data in high-dimensional spaces. As the number of dimensions increases, the volume of the space increases exponentially, making data sparse and models less effective.

## 95. When would you use feature selection versus feature extraction for dimensionality reduction?

Feature selection involves selecting a subset of the original features based on certain criteria, while feature extraction transforms the data into a lower-dimensional space. Use feature selection when interpretability is important, and feature extraction when you want to capture the most information in fewer dimensions.