

52.

The Curse of Dimensionality refers to the fact that as the number of dimensions increases, the data becomes more spread out and sparse in the space.

In high dimensions, most of the data is far from the mean, and very little lies within one standard deviation (or one Mahalanobis distance) from the center.

Why do I care?

Because most real-world data has many attributes, and without handling this properly, models can become ineffective.

We may think more attributes help, but too many dimensions can reduce model performance, cause overfitting, and make computations inefficient.

It's important to understand this to decide when and how to use dimensionality reduction techniques like PCA