What is the curse of dimensionality? In what cases would you use vanilla PCA, Incremental PCA, Randomized PCA, or Kernel PCA?

The curse of dimensionality refers to the fact that many problems that do not exist in low-dimensional space arise in high-dimensional space. In Machine Learning, one common manifestation is the fact that randomly sampled high-dimensional vectors are generally very sparse, increasing the risk of overfitting and making it very difficult to identify patterns in the data without having plenty of training data.

The dataset fits in memory using pure PCA.

Large datasets that can't fit in memory and online tests: Incremental PCA

Randomized PCA: significantly lower dimensionality and better dataset memory fit. utilized for nonlinear PCA is kernel PCA.

# Here as per the result I can say like training time with 100 estimators is more. Also, prediction scores is also increasing.