**Mitigating Curse of Dimensionality**

To mitigate the problems associated with high dimensional data a suite of techniques generally referred to as ‘Dimensionality reduction techniques are used. Dimensionality reduction techniques fall into one of the two categories- ‘Feature selection’ or ‘Feature extraction.

**Feature selection Techniques**

In feature selection techniques, the attributes are tested for their worthiness and then selected or eliminated. Some of the commonly used Feature selection techniques are discussed below.

**Low Variance filter**:  In this technique, the variance in the distribution of all the attributes in a dataset is compared and attributes with very low variance are eliminated. Attributes that do not have such much variance will assume an almost constant value and do not contribute to the predictability of the model.

**High Correlation filter**: In this technique, the pair wise correlation between attributes is determined. One of the attributes in the pairs that show very high correlation is eliminated and the other retained. The variability in the eliminated attribute is captured through the retained attribute.

**Multicollinearity**: In some cases, the high correlation may not be found for pairs of attributes but if each attribute is regressed as a function of others, we may see that variability of some of the attributes are completely captured by the others. This aspect is referred to as multicollinearity and Variance Inflation Factor (VIF) is a popular technique used to detect multicollinearity. Attributes with high VIF values, generally greater than 10, are eliminated.

**Feature Ranking**: Decision Tree models such as CART can rank the attributes based on their importance or contribution to the predictability of the model. In high dimensional data, some of the lower ranked variables could be eliminated to reduce the dimensions.

**Feature Extraction Techniques**

In feature extraction techniques, the high dimensional attributes are combined in low dimensional components (PCA or ICA) or factored into low dimensional factors (FA).

**Principal Component Analysis (PCA)**

[Principal Component Analysis](https://www.mygreatlearning.com/blog/understanding-principal-component-analysis/), or PCA, is a dimensionality-reduction technique in which high dimensional correlated data is transformed to a lower dimensional set of uncorrelated components, referred to as principal components. The lower dimensional principle components capture most of the information in the high dimensional dataset. An ‘n’ dimensional data is transformed into ‘n’ principle components and a subset of these ‘n’ principle components is selected based on the percentage of variance in the data intended to be captured through the principle components. Figure 5 shows a simple example in which a 10-dimensional data is transformed to 10-principle components. To capture 90% of the variance in the data only 3 principle components are needed. Hence, we have reduced a 10-dimensional data to 3-dimensions.

**Figure 5. Example of converting 10-dimensional data to 3-dimensional data through PCA**

**Factor Analysis (FA)**

Factor analysis is based on the assumption that all the observed attributes in a dataset can be represented as a weighted linear combination of latent factors. The intuition in this technique is that an ‘n’ dimensional data can be represented by ‘m’ factors (m<n). The main difference between PCA and FA is in the fact that While PCA synthesizes components from the base attributes, FA decomposes the attributes into latent factors as shown in figure 6.

**Independent Component Analysis (ICA)**

ICA assumes that all the attributes are essentially a mixture of independent components and resolves the variables into a combination of these independent components. [ICA](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3538438/) is perceived to be more robust than PCA and is generally used when PCA and FA fail.