Backward Feature Elimination, Forward Feature Selection:

Both take a lot of computational time and are thus generally used on smaller datasets

Random Forest:

This is one of the most commonly used techniques which tells us the importance of each feature present in the dataset. We can find the importance of each feature and keep the top most features, resulting in dimensionality reduction

Low Variance filter:

We apply this approach to identify and drop constant variables from the dataset. The target variable is not unduly affected by variables with low variance, and hence these variables can be safely dropped

High Correlation filter:

A pair of variables having high correlation increases multicollinearity in the dataset. So, we can use this technique to find highly correlated features and drop them accordingly

Principal Component Analysis:

This is one of the most widely used techniques for dealing with linear data. It divides the data into a set of components which try to explain as much variance as possible. One of the most advantages for PCA is that it can run in a short time. So when we get into a high dimension inputs and large number of samples, we prefer to use PCA

Factor Analysis:

This technique is best suited for situations where we have highly correlated set of variables. It divides the variables based on their correlation into different groups, and represents each group with a factor

t-SNE:

This technique works well when the data is strongly non-linear. It works extremely well for visualizations as well. Shortcoming: loss of large-scale information, slow computation time, and inability to meaningfully represent very large datasets.

UMAP:

This technique works well for high dimensional data. Its run-time is shorter as compared to t-SNE. So due to the same reason, we prefer UMAP.

How it works:

This method uses the concept of k-nearest neighbor and optimizes the results using stochastic gradient descent. It first calculates the distance between the points in high dimensional space, projects them onto the low dimensional space, and calculates the distance between points in this low dimensional space. It then uses Stochastic Gradient Descent to minimize the difference between these distances.

Advantage:

It can handle large datasets and high dimensional data without too much difficulty

It combines the power of visualization with the ability to reduce the dimensions of the data

Along with preserving the local structure, it also preserves the global structure of the data. UMAP maps nearby points on the manifold to nearby points in the low dimensional representation, and does the same for far away points

Missing Value Ratio:

If the dataset has too many missing values, we use this approach to reduce the number of variables. We can drop the variables having a large number of missing values in them

Independent Component Analysis:

We can use ICA to transform the data into independent components which describe the data using less number of components

ISOMAP:

We use this technique when the data is strongly non-linear