**Dimensionality Reduction :**

* Dimensionality reduction techniques are used to reduce the number of features in your dataset without having to lose much information and keep (or improve) the model’s performance.
* It’s a really powerful way to deal with huge datasets.

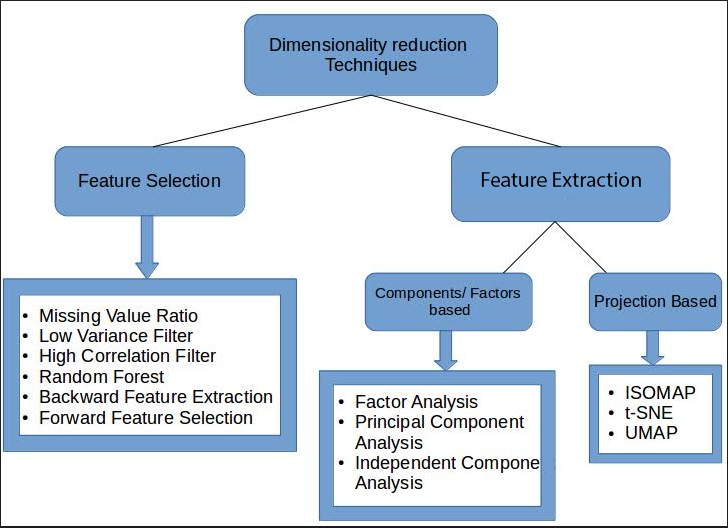
**Benefits** :

* Space required to store the data is reduced as the number of dimensions comes down
* Fewer dimensions lead to less computation/training time
* Some algorithms do not perform well when we have large dimensions. So reducing these dimensions needs to happen for the algorithm to be useful
* It takes care of multicollinearity by removing redundant features. For example, you have two variables – ‘time spent on a treadmill in minutes’ and ‘calories burnt’. These variables are highly correlated as the more time you spend running on a treadmill, the more calories you will burn. Hence, there is no point in storing both as just one of them does what you require
* It helps in visualizing data. As discussed earlier, it is very difficult to visualize data in higher dimensions so reducing our space to 2D or 3D may allow us to plot and observe patterns more clearly

**Dimensionality reduction can be done in two different ways:**

* By only keeping the most relevant variables from the original dataset (this technique is called feature selection)
* By finding a smaller set of new variables, each being a combination of the input variables, containing basically the same information as the input variables (this technique is called feature extraction).
* Feature extraction will be divided into 2 parts:

1. Factor/Component-based method
2. Projection Based Methods



https://courses.analyticsvidhya.com/courses/take/dimensionality-reduction-for-machine-learning/texts/10848320-introduction