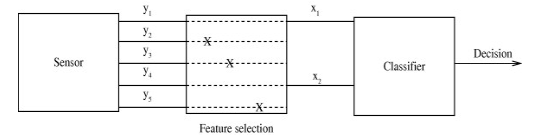
(26 April 2021)

Dimensionality Reduction

It is simply the process of reducing the dimension of the features set. As the number of features increases, the model becomes more complex. The more the number of features, the more the chances of over fitting. Over fitting is a modeling error that leads to poor prediction. The major advantage of dimensionality reduction is to avoid over fitting. In dimensionality reduction, data is transformed from a high dimensional space into a low dimensional space so that low dimensional representation retains some meaningful properties of original data.

Dimensionality Reduction is the process of reducing the number of features, (Random Variables), by obtaining a set of principle variables. Dimensionality Reduction can be divided into,

1. Feature Selection – Process of identifying and selecting relevant features from the input data set.



1. Feature Extraction – Process of generating new features from existing features, by applying some transformation or performing some operation on them.

