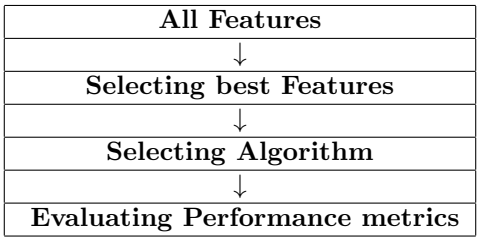
1 Feature Selection

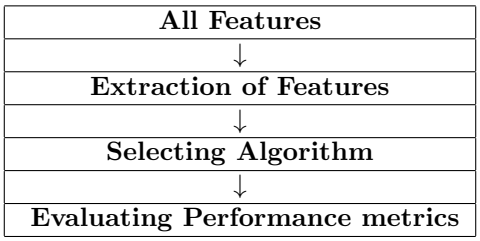
As William Shakespeare says, "Sometimes, less is better". In the similar manner, Feature selection keeps the features intact and chooses n best features among them, removing the redundant and co-linear features. Machine learning works on a simple rule – if you put garbage in, you will only get garbage out. Garbage here means, the noise in data.



This becomes even more important when the number of features are very large. You need not use every feature at your disposal for creating an algorithm. You can assist your algorithm by feeding in only those features that are really important. This can be very useful in industrial applications. You not only reduce the training time and the evaluation time, you also have less things to worry about!

2 Feature Extraction

Extraction of good features is integral for the efficient performance of a machine learning model. Input data may be large or it may be redundant or both, so there should be efficient feature extraction technique which will transform input into a reduced set of features.



In feature extraction, we create "new" independent variables, where each "new" independent variable is a combination of each of the "old" independent variables. However, we create these new independent variables in a specific way and order these new variables by how well they predict our dependent variable. For example, Principal component analysis (PCA)[2] is a method based on feature extraction.

Top reasons to use feature selection and feature extraction are:

- They enable the machine learning algorithm to train faster.
- They reduce the complexity of a model and makes it easier to interpret.
- They improve the accuracy of a model if the right subset is chosen.
- They reduce over fitting.

References:

[1]https://www.analyticsvidhya.com/blog/2016/02/7-important-model-evaluation-error-metrics/
[2]https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c