FEATURE SELECTION

- Feature Selection is the process where you automatically or manually select those features which contribute most to your prediction variable or output in which you are interested in.
- Feature Selection is one of the core concepts in machine learning which hugely impacts the performance of your model.
- Irrelavant features in the model will tend to reduce the accuracy and performance of model

- Benefits of feature selection is it reduces overfitting,increases the accuracy,reduces the training time
- Training unwanted features will affect the performance of the model
- Main benefit is it reduces curse of dimensionality
- feature selection methods can be divided into three major buckets

    1. **Filter based:** We specify some metric and based on that filter features. An example of such a metric could be correlation/chi-square.
    2. **Wrapper-based:** Wrapper methods consider the selection of a set of features as a search problem. Example: Recursive Feature Elimination
    3. **Embedded:** Embedded methods use algorithms that have built-in feature selection methods. For instance, Lasso and RF have their own feature selection methods.

FEATURE SELECTION TECHNIQUES

- Univariate selection
- Feature importance
- Correlation
- Variance threshold
- Information gain-Mutual information in classification
- Information gain-Mutual information in regression
- Chi-square
- Recursive feature elimination etc...