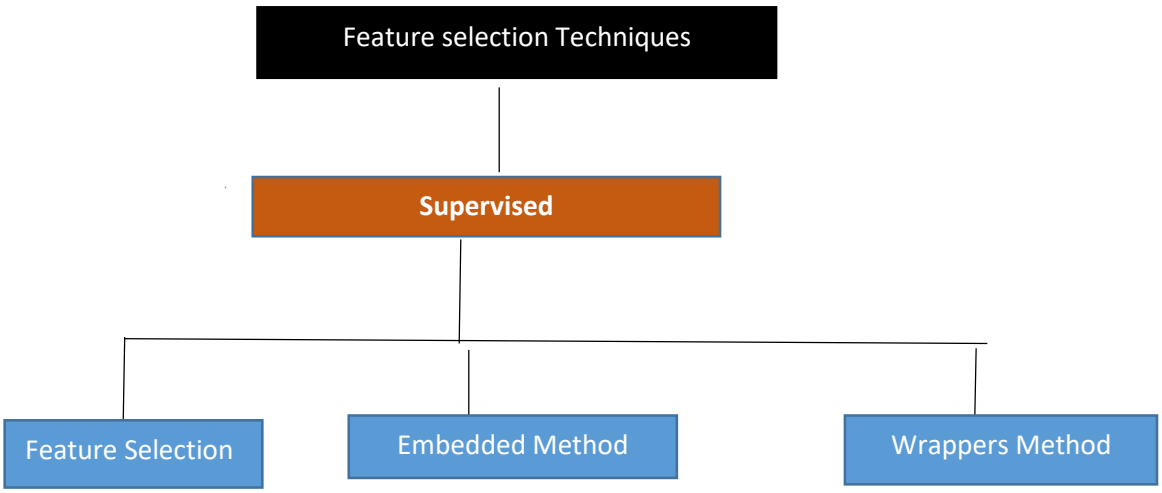
# Feature selection

Feature selection techniques:




## 1. Wrapper Methods

In wrapper methodology, selection of features is done by considering it as a search problem, in which different combinations are made, evaluated, and compared with other combinations. It trains the algorithm by using the subset of features iteratively.

Some techniques of wrapper method:

2. **Forward Selection**: Forward selection is an iterative method in which we start with having no feature in the model. In each iteration, we keep adding the feature which best improves our model till an addition of a new variable does not improve the performance of the model.

3. **Backward Elimination**: In backward elimination, we start with all the features and removes the least significant feature at each iteration which improves the performance of the model. We repeat this until no improvement is observed on removal of features.

4. **Recursive Feature elimination**: It is a greedy optimization algorithm which aims to find the best performing feature subset. It repeatedly creates models and keeps aside the best or the worst performing feature at each iteration. It constructs the next model with the left features until all the features are exhausted. It then ranks the features based on the order of their elimination.

## 2. Filter Methods

In Filter Method, features are selected on the basis of statistics measures. This method does not depend on the learning algorithm and chooses the features as a pre-processing step.

The filter method filters out the irrelevant feature and redundant columns from the model by using different metrics through ranking.

The advantage of using filter methods is that it needs low computational time and does not overfit the data.

Some techniques are:

**Information Gain:** Information gain determines the reduction in entropy while transforming the dataset. It can be used as a feature selection technique by calculating the information gain of each variable with respect to the target variable.

**Chi-square Test:** Chi-square test is a technique to determine the relationship between the categorical variables. The chi-square value is calculated between each feature and the target variable, and the desired number of features with the best chi-square value is selected.

**Fisher's Score:**

Fisher's score is one of the popular supervised technique of features selection. It returns the rank of the variable on the fisher's criteria in descending order. Then we can select the variables with a large fisher's score.

Here are a few common feature selection techniques you can explore:

1. Univariate Feature Selection:
    - This method evaluates each feature independently based on statistical tests or heuristics.
    - Examples include chi-squared test, ANOVA, and mutual information.
    - Features with the highest scores are selected.
2. Recursive Feature Elimination:
    - It is an iterative approach that starts with all features and eliminates the least important ones based on a model's performance.
    - The process continues until a desired number of features or a specified performance threshold is reached.
3. L1 Regularization (Lasso):
    - This technique adds a penalty term to the model's objective function, encouraging sparsity in feature weights.

- Features with zero weights are removed from the model, effectively performing feature selection.

4. Tree-based Methods:

- Decision trees and ensemble methods like Random Forests and Gradient Boosting can assess feature importance.
- Features are ranked based on their contribution to the decision-making process.

## Conclusion

Feature selection is a very complicated and vast field of machine learning, and lots of studies are already made to discover the best methods. There is no fixed rule of the best feature selection method. However, choosing the method depend on a machine learning engineer who can combine and innovate approaches to find the best method for a specific problem.