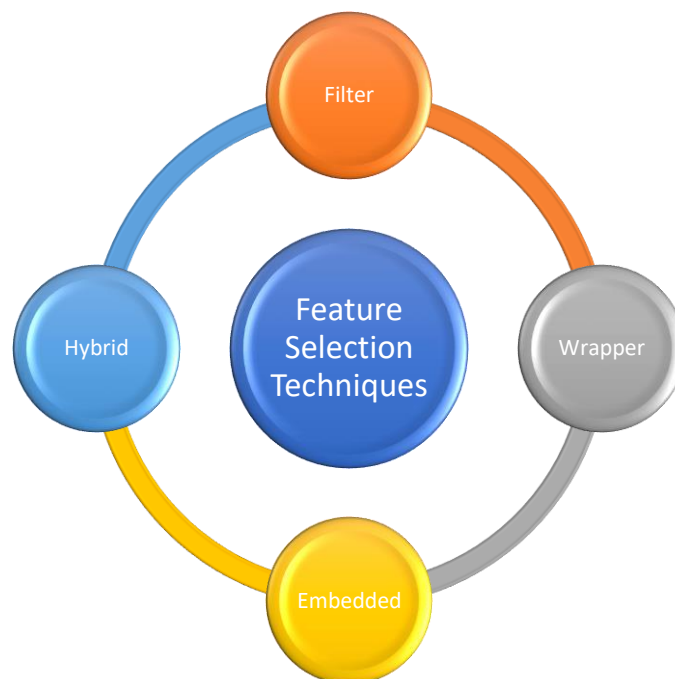


FEATURE SELECTION TECHNIQUES

What is feature selection?

Features are the variables or attributes upon which the dependent value or the result of the model depends. When a dataset is provided for a data science task, it may contain a lot of features of which some may be irrelevant to the task. For example, in order to predict whether a person will repay the loan, an attribute like his/her phone number is not necessary. Removing such attributes will eventually increase the performance of the machine learning model. **The process of removing less relevant features and selecting only relevant features is called as feature selection.**



Filter Techniques

1. Information Gain
 - It calculates the amount of information a particular attribute can convey when compared to other attributes and to the whole task
 - The attributes with highest information gain is retained while the one(s) with the least is discarded
 - Generally used in building decision trees
2. Chi – square test
 - Used for categorical features in a dataset
 - Calculates the score between each feature and the target and selects the number of features satisfying a certain chi square score
3. Correlation Coefficient
 - Used to compute the linear relationship between 2 or more variables

- If two variables are correlated then we can predict one from the other

Wrapper Techniques

1. Forward Feature Selection

- The procedure starts with an empty set of attributes as the reduced set.
- The best of the original attributes is determined and added to the reduced set.
- At each subsequent iteration or step, the best of the remaining original attributes is added to the set

2. Backward Elimination

- The procedure starts with the full set of attributes.
- At each step, it removes the worst attribute remaining in the set.

Embedded Techniques

1. L1 Regularization

- Also known as Lasso regularization
- Attaches a penalty to each attribute
- In linear model regularization, penalty is added over the coefficients
- Hence some coefficients become zero and do not have any effect on the final result

Hybrid techniques involve a combination of filter and wrapper techniques. Only some among the many feature selection techniques have been explained above.