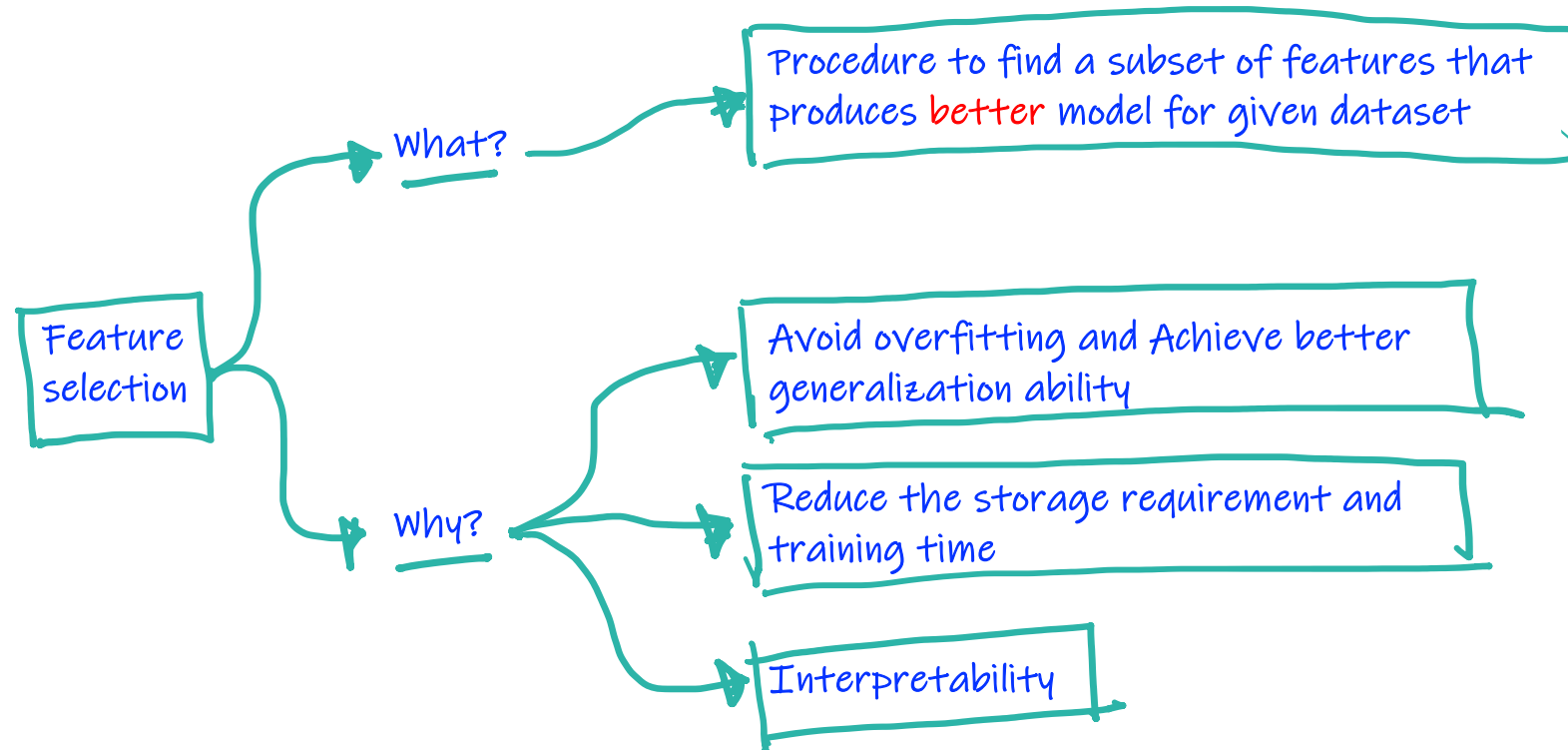


Feature Selection

IKN

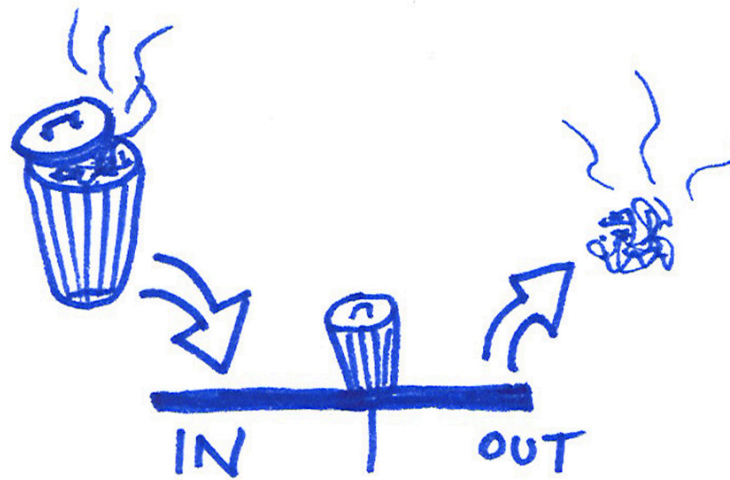


Why we have to reduce the feature?



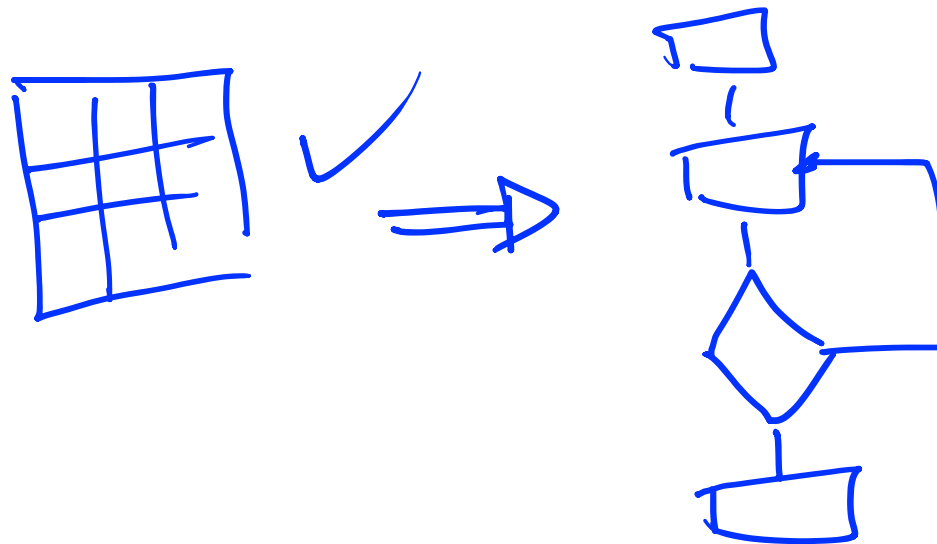
Why we should clean the data

Garbage in garbage out

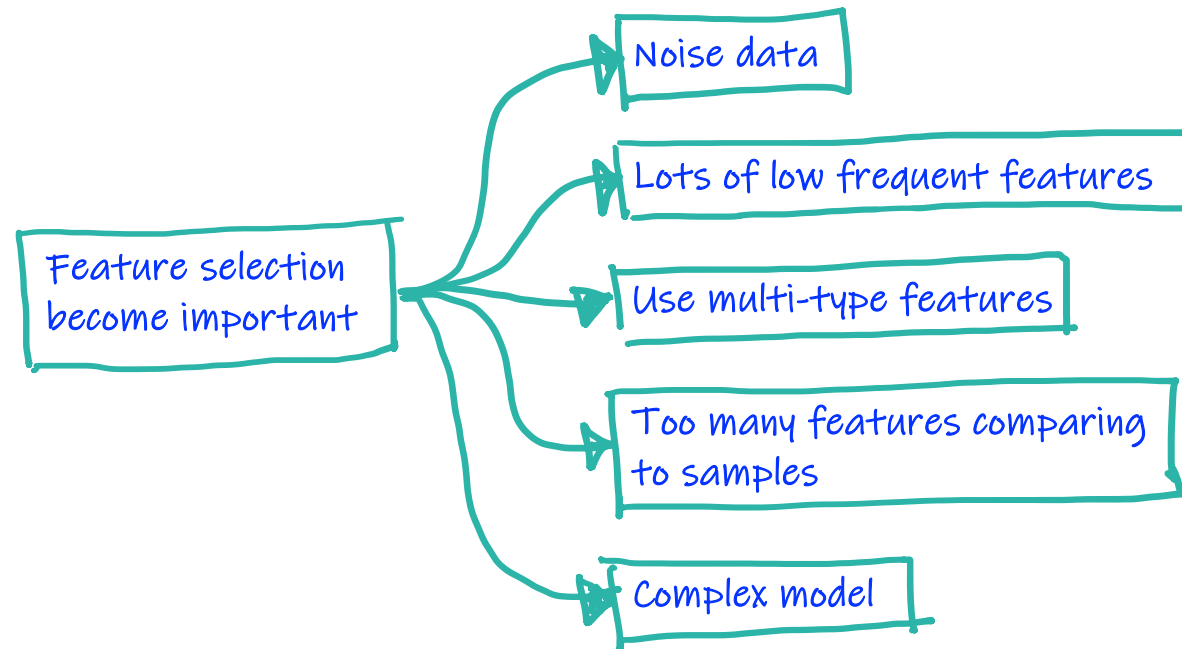


Why we should clean the data

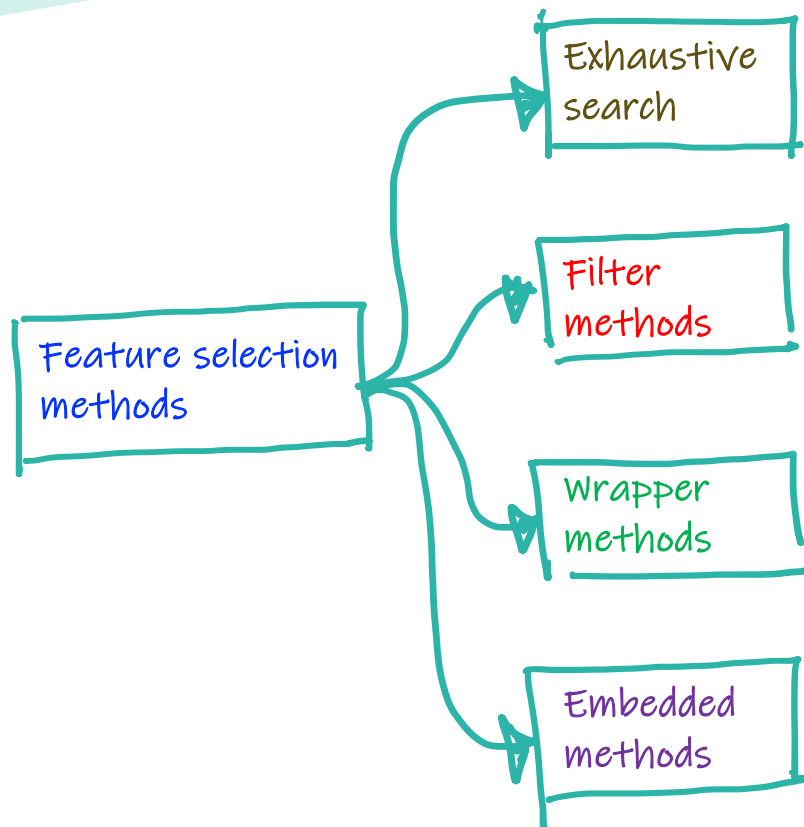
Quality data beats fancy algorithm



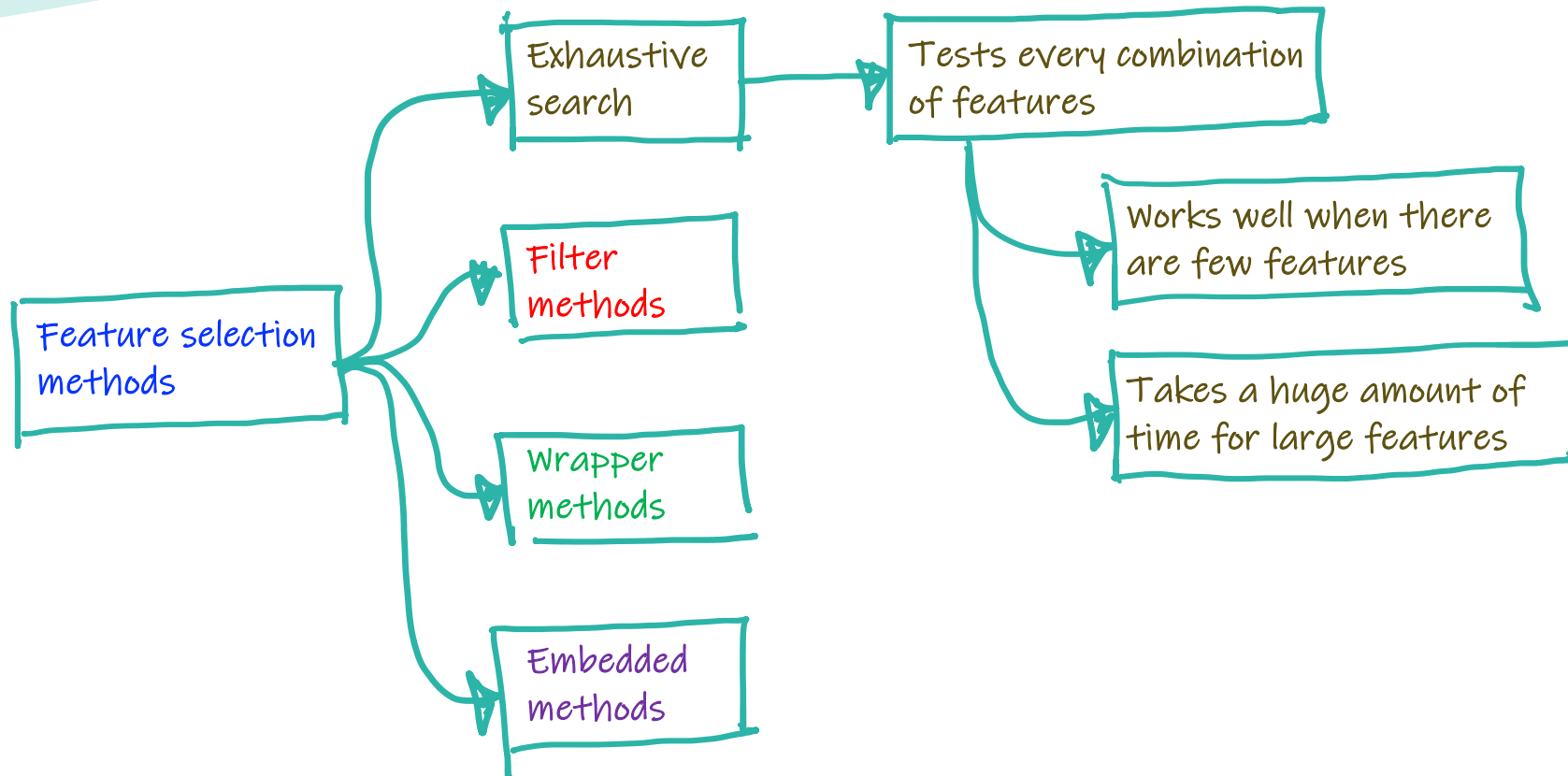
When feature selection become important?



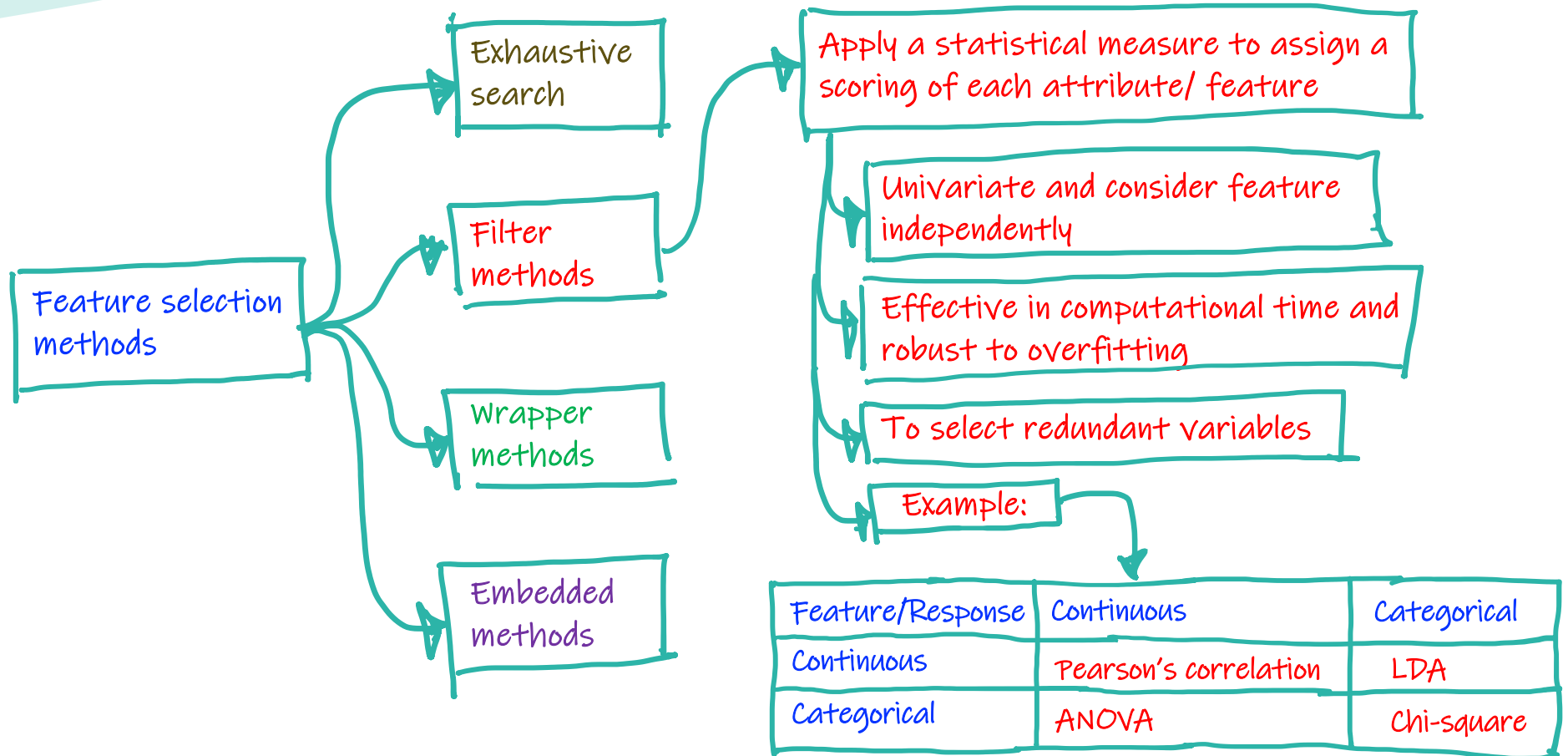
Typical category of feature selection



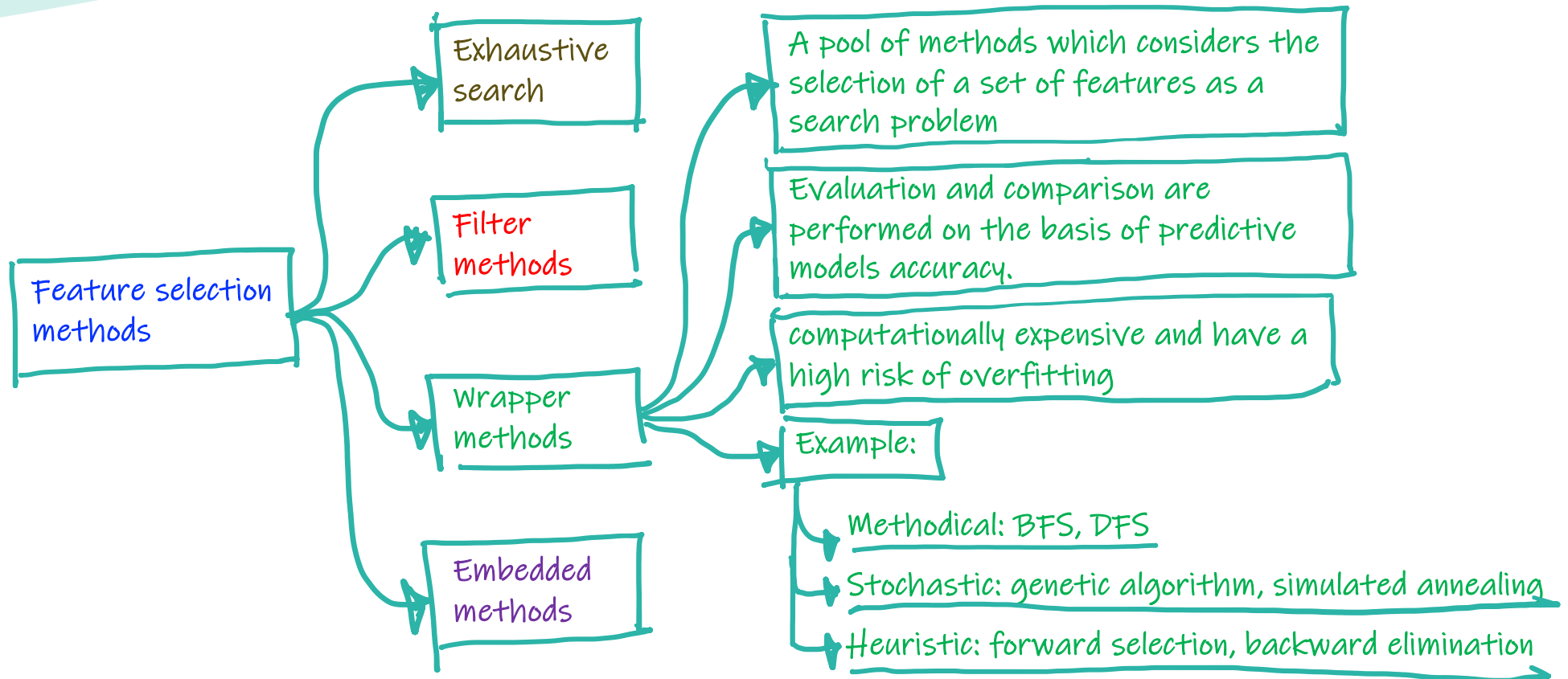
Typical category of feature selection



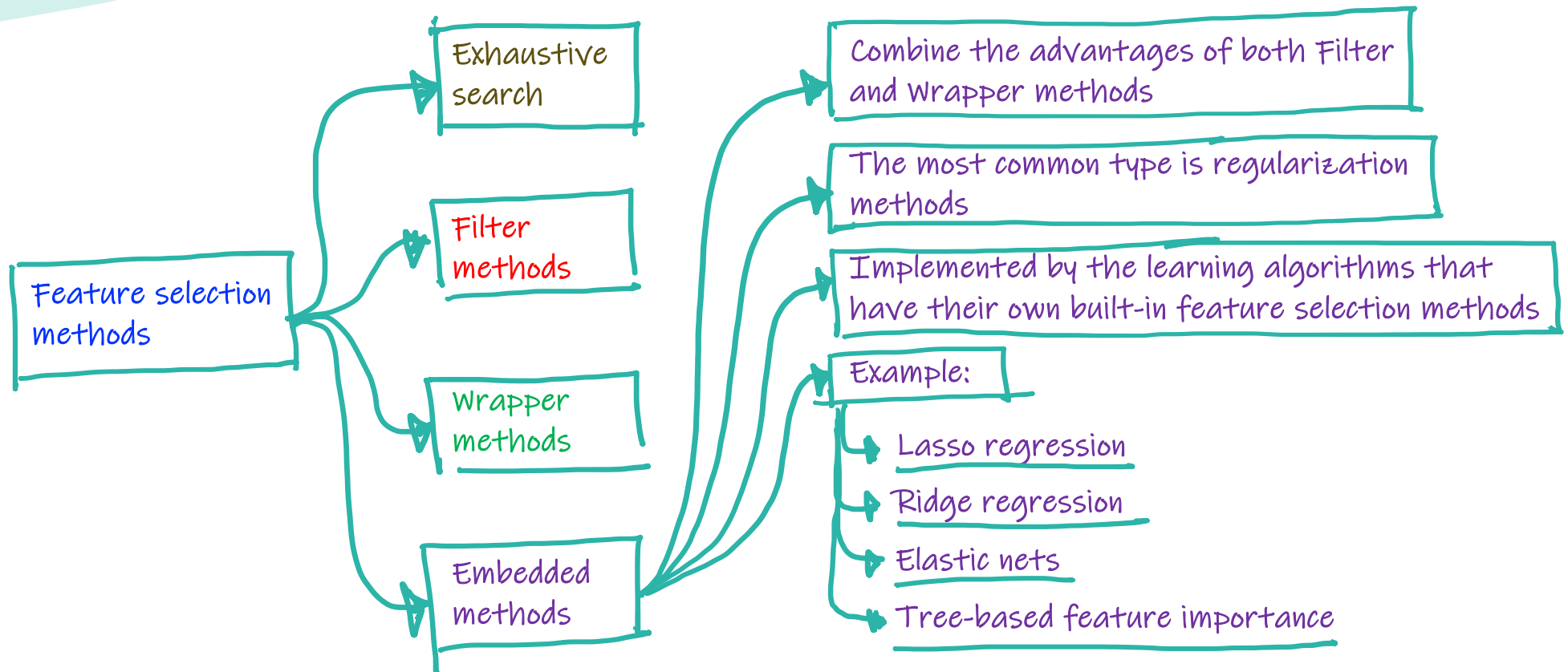
Typical category of feature selection



Typical category of feature selection



Typical category of feature selection



Tree-based feature importance

- Gini Importance / Mean Decrease in Impurity (MDI)

