# FEATURE SELECTION METHODS

OVERVIEW

# FEATURE SELECTION: METHODS

- Filter methods
- Wrapper methods
- Embedded methods

# FILTER METHODS

- Rely on the characteristics of the data (feature characteristics)
- Do not use machine learning algorithms
- Model agnostic
- Tend to be less computationally expensive
- Usually give lower prediction performance than a wrapper methods
- Are very well suited for a quick screen and removal of irrelevant features

# WRAPPER METHODS

- Use predictive machine learning models to score the feature subset
- Train a new model on each feature subset
- Tend to be very computationally expensive
- Usually provide the best performing feature subset for a given machine learning algorithm
- They may not produce the best feature combination for a different machine learning model

# EMBEDDED METHODS

- Perform feature selection as part of the model construction process
- Consider the interaction between features and models
- They are less computationally expensive than wrapper methods, because they fit the machine learning model only once.

# FEATURE SELECTION: METHODS

| Filter | Wrapper | Embedded |
| --- | --- | --- |
| • Variance<br>• Correlation<br>• Univariate selection | • Forward selection<br>• Backward selection<br>• Exhaustive search | • LASSO<br>• Tree importance |