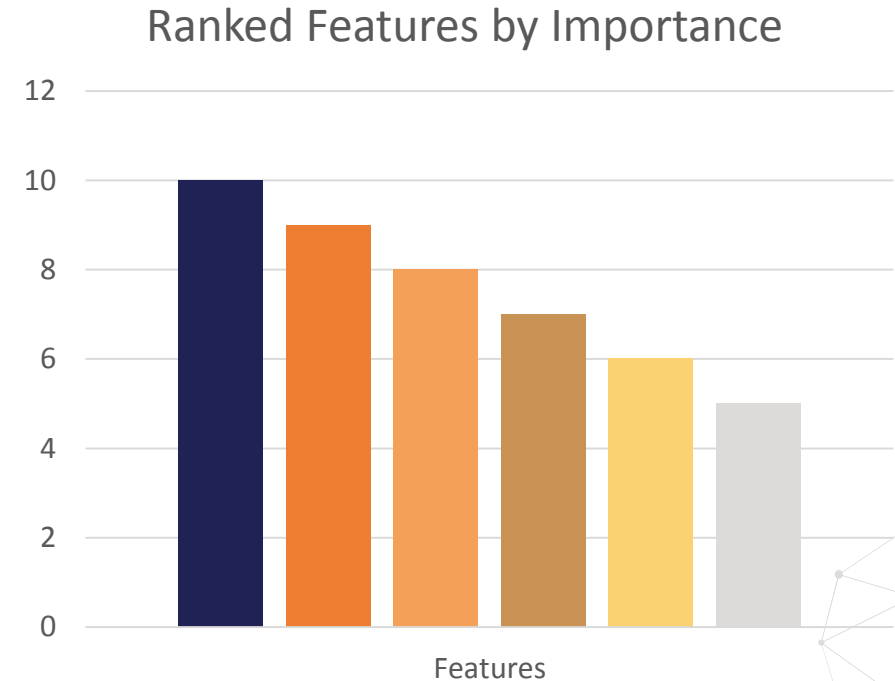# Individual feature selection methods

Filter methods: summary

1. Rank features by importance

   1. Mutual information

   2. Chi-square

   3. Anova

   4. Any other suitable statistical test

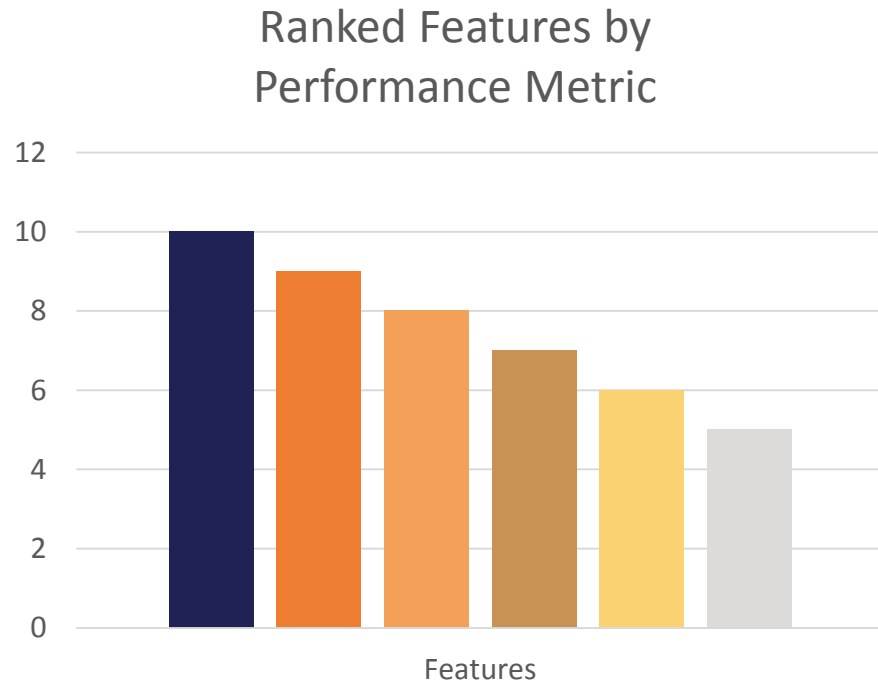Ranked Features by Importance

Filter methods: summary

1. Rank features by importance

   1. Mutual information

   2. Chi-square

   3. Anova

   4. Any other suitable statistical test

2. Select highest ranking features

   1. Highest mutual information
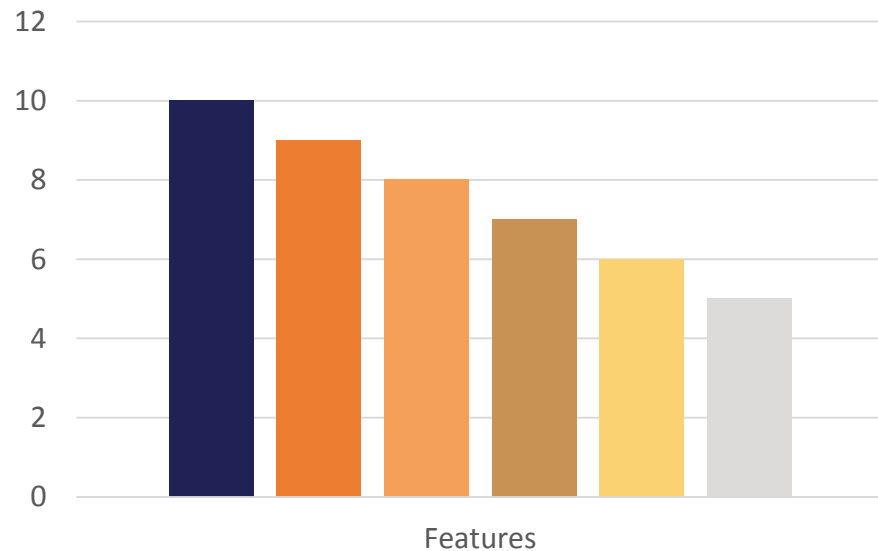
   2. Lowest p-values

Ranked Features by Importance

Rank Features by Performance Metric
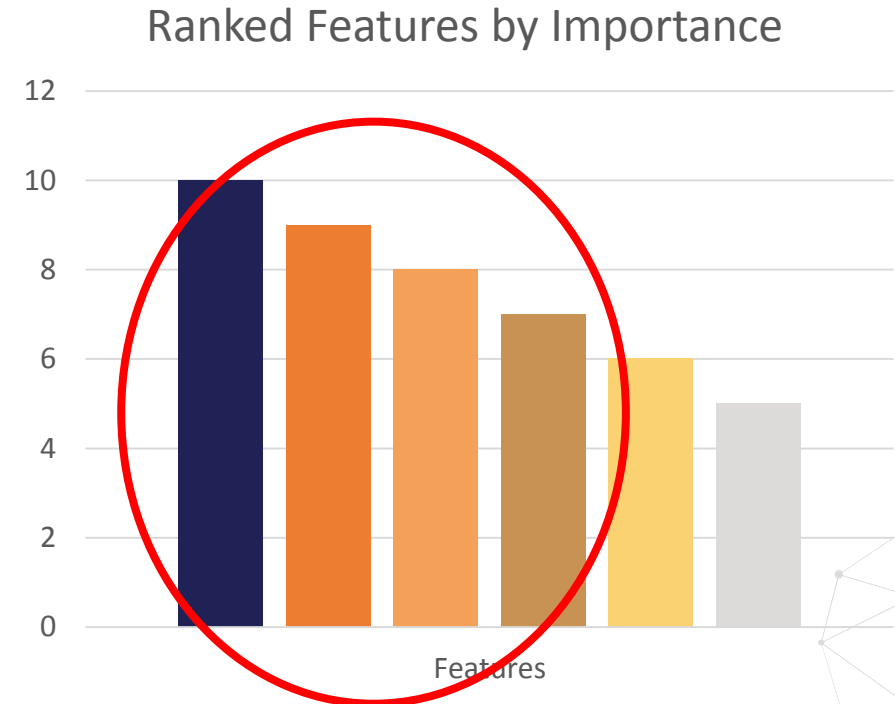
Ranked Features by Performance Metric

To get a performance metric, we need to get predictions first!!

Rank Features by Performance Metric

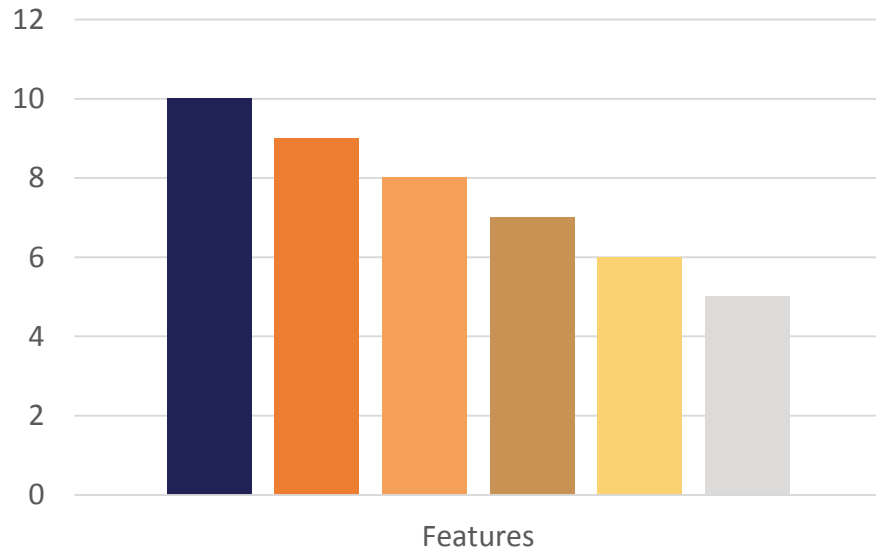### Ranked Features by Performance Metric

Obtaining the predictions:

With a Machine Learning Model

1. Classification tree (or any other model) utilising a single feature against the target

2. Evaluate the performance of each tree (or model) and obtain a performance metric per feature

Rank Features by Performance Metric

Ranked Features by Performance Metric

Obtaining the predictions:
KDD 2009 approach

1. Encode the variables with a representation of the target

2. Utilise this representation as prediction to determine performance

Train In Data

THANK YOU

www.trainindata.com