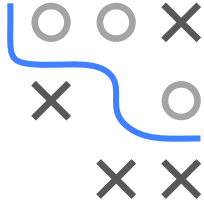
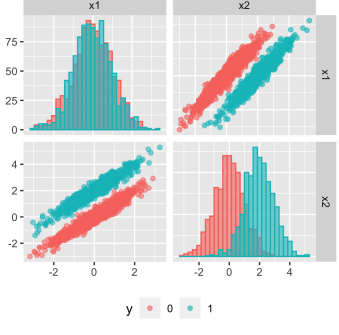
Introduction to Machine Learning



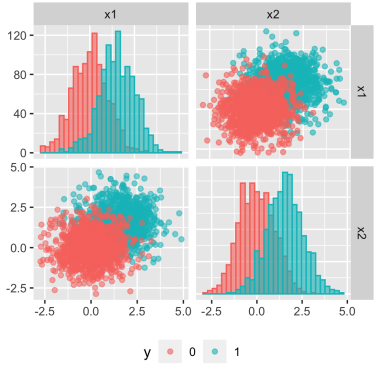
Feature Selection
Feature Selection: Filter Methods (Examples and Caveats)



Learning goals

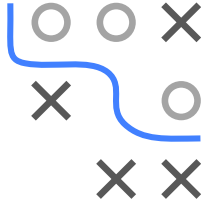
- Understand how filter methods can be misleading
- Understand how filters can be applied and tuned

FILTER METHODS CAN BE MISLEADING



ρ_{ACC} of log. reg. classifier with:

- feature x_1: 0.76
- feature x_2: 0.78
- both features: 0.85



Information gain from presumably redundant variables. 2 class problem with indep features. Each class has Gaussian distribution with no covariance. While filter methods suggest redundancy, combination of both vars yields improvement, showing indep vars are not truly redundant. For further details, see ▸ Guyon and Elisseeff 2003.

USING FILTER METHODS

1. Calculate filter score for each feature x_j
2. Rank features according to score values
3. Choose $\tilde{p}$ best features
4. Train model on $\tilde{p}$ best features

How to choose $\tilde{p}$?

- Could be prescribed by application
- Eyeball estimation: read from filter plots
- Treat as hyperparameter and tune in a pipeline, based on resampling