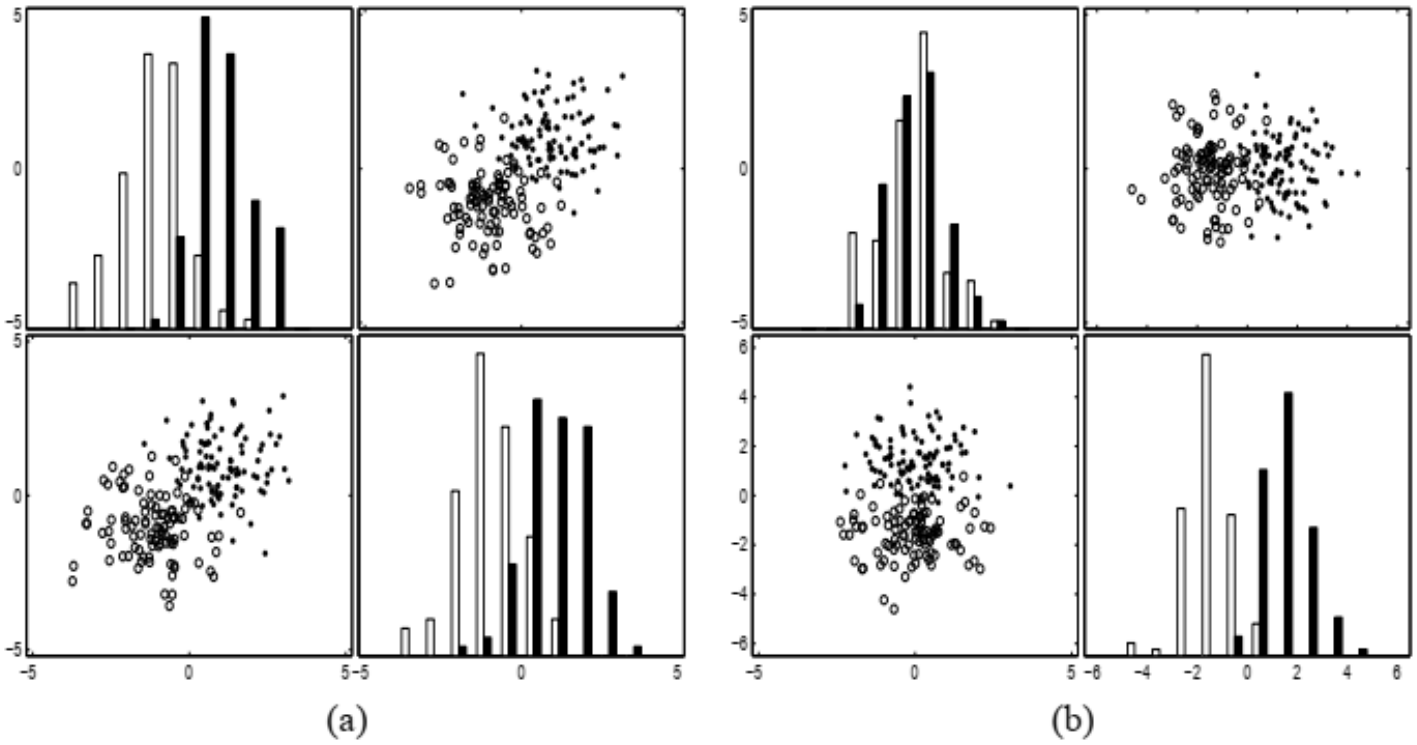
Feature selection

Marcin Kuta

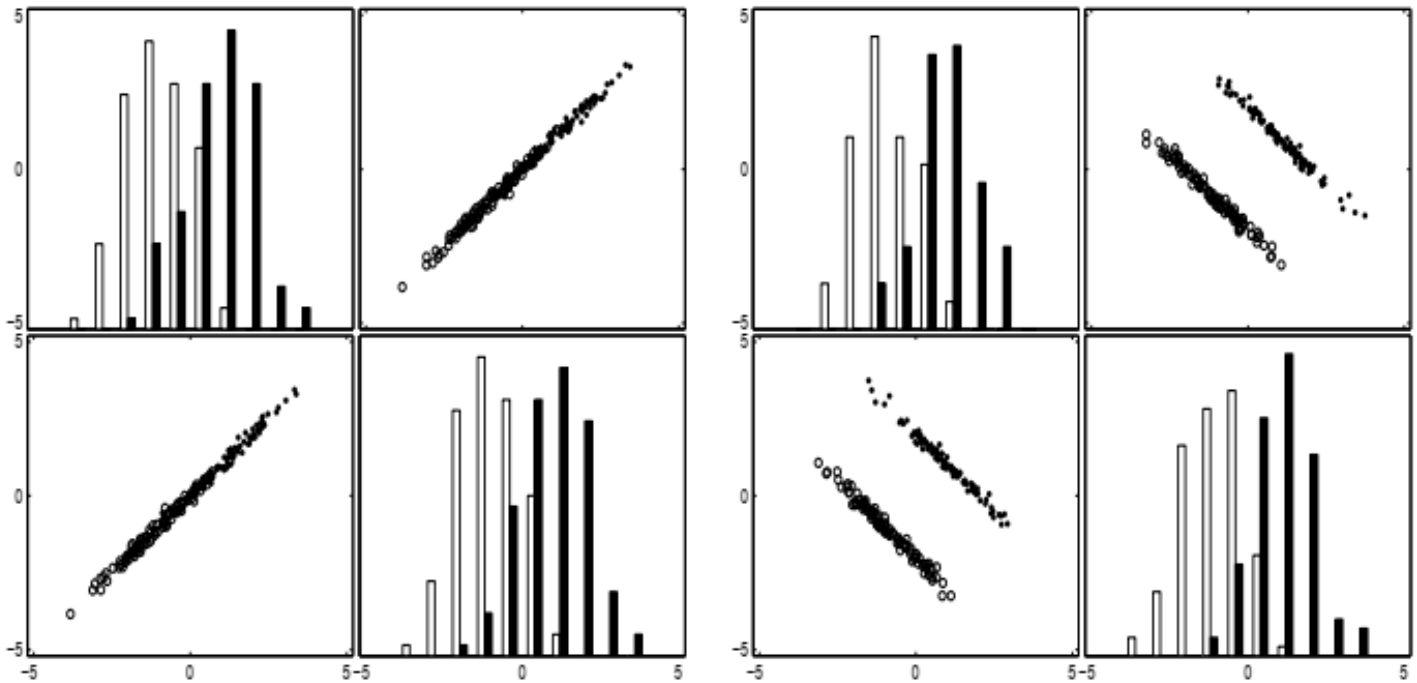
Motivation

- Lower computational cost
- Better interpretability
- Cheaper feature acquisition
- Avoiding overfitting

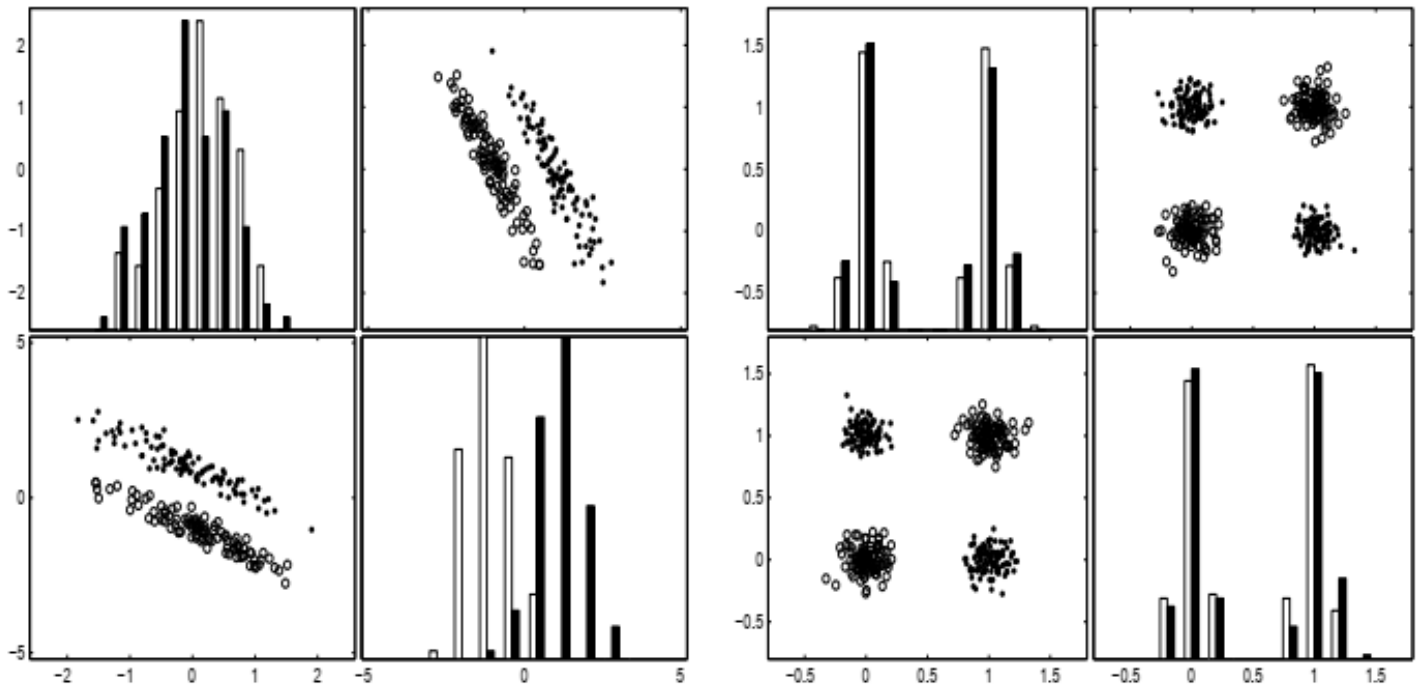
Feature selection



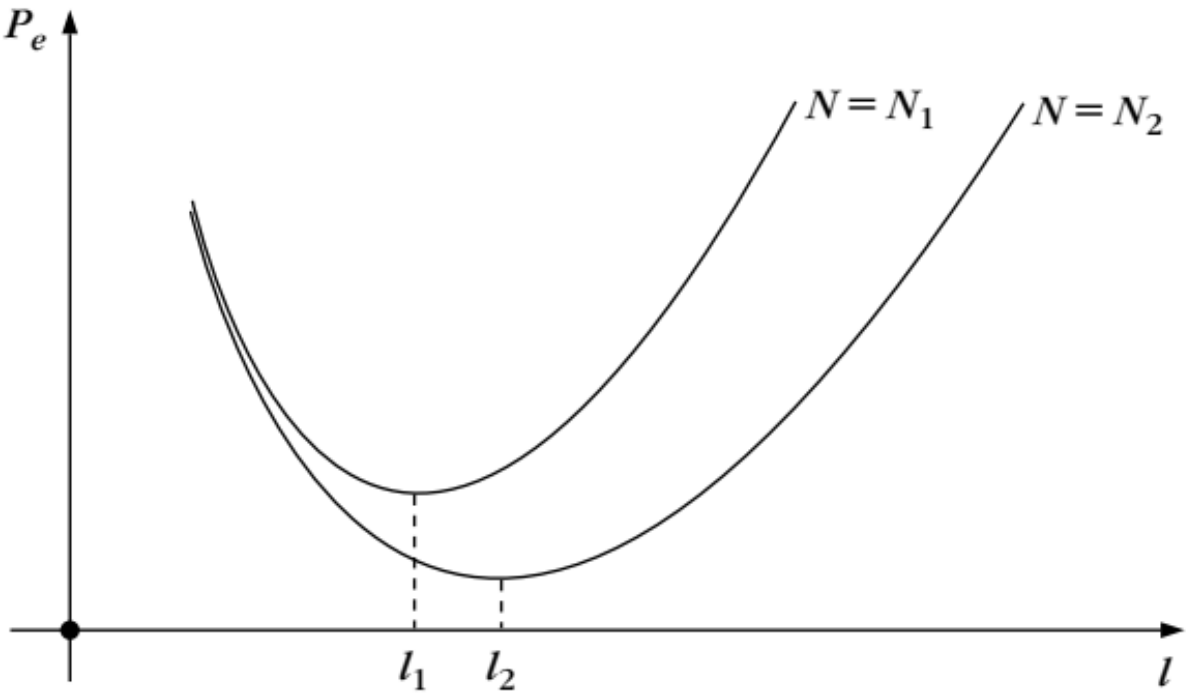
Feature selection



Feature selection



Peaking phenomenon



Feature selection

- Embedded methods
- Filter methods
- Wrapper methods
- Hybrid methods
- Ensemble methods

Filter methods

- Variance
- Correlation with target

Wrapper methods

- Sequential Forward Selection
- Sequential Backward Selection

Embedded methods

- L1 regularization
- Random forest importance

Hybrid methods

Recursive feature elimination

1. fit the classifier
2. rank the features according to their importances
3. eliminate one feature – the least important one
4. if the effectiveness drops by more than a threshold retain the feature, otheriwise remove the feature
5. repreat steps 2–4 until all features are evaluated

Feature selection loop

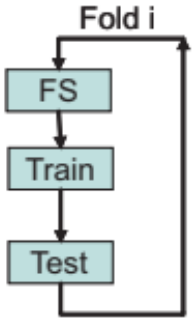


Figure 1: Method IN

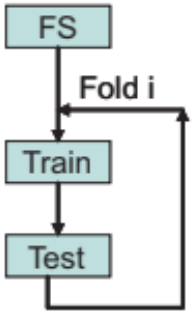


Figure 2: Method OUT

References

[1] https://github.com/rasbt/stat479-machine-learning-fs18/blob/master/13_feat-sele/13_feat-sele_code.ipynb

[2] https://inria.github.io/scikit-learn-mooc/feature_selection/feature_selection_module_intro.html

[3] Isabelle Guyon, André Elisseeff,
An Introduction to Variable and Feature Selection,
Journal of Machine Learning Research 3: pp. 1157–1182, 2003.

[4] Payam Refaeilzadeh, Lei Tang, Huan Liu,
On comparison of feature selection algorithms,
AAAI Workshop, 2007, pp. 34–39