

Introduction to Machine Learning



Feature Selection: Motivating Examples

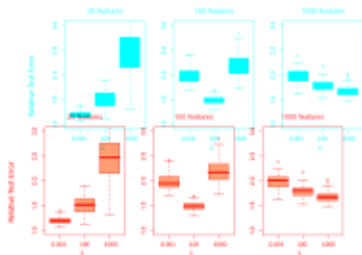
Feature Selection: Motivating Examples

Learning goals

- Understand the practical importance of feature selection

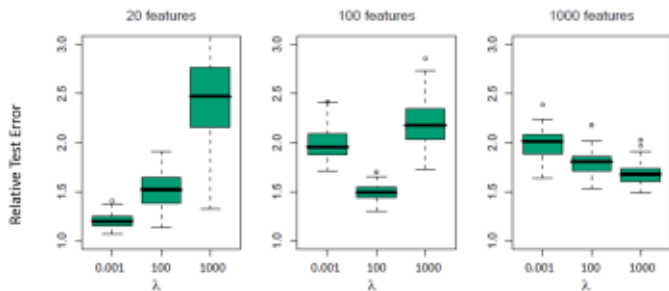
Learning goals

- Understand that models with integrated selection do not always work
- Understand that models with integrated selection do not always work
- Know different categories of selection methods



MOTIVATING EXAMPLE 1: REGULARIZATION / 2

- Boxplots show the relative test error (RTE = test error/Bayes error σ^2) over 100 simulations for the different values of p and λ .



- Lowest RTE is obtained at $\lambda = 0.001$ for $p = 20$, at $\lambda = 100$ for $p = 100$, and at $\lambda = 100$ for $p = 1000$.
 - Optimal amount of regularization increases monotonically in p here.
- ⇒ High-dimensional settings require more complexity control through regularization or feature selection.

MOTIVATING EX. 3: INTEGRATED SELECTION / 2

- We compare several classifiers regarding their misclassification rate, of which two have integrated FS (rpart and rForest).
- Since we have few observations, we use repeated 10-fold cross-validation with 10 repetitions.



	rpart	lda	logreg	nBayes	knn7	rForest
all feat.	0.44	0.27	0.25	0.32	0.37	0.36
relevant feat.	0.44	0.18	0.19	0.27	0.33	0.30

- ⇒ Different to Ex. 2, models with integrated FS do not work ideally here. Also, methods with lin. decision boundary are better due to our simulation set-up.
- ⇒ Performance improves significantly for most methods when only trained on informative features.