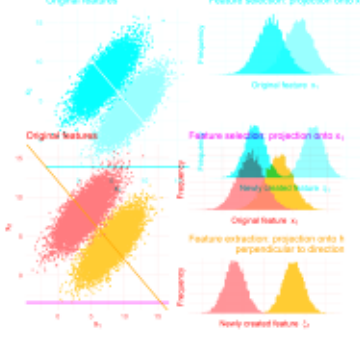
Introduction to Machine Learning

Feature Selection: Introduction

Feature Selection: Introduction



Learning goals

- Too many features can be harmful in prediction
- Selection vs. extraction
- Types of selection methods

Learning goals

- Too many features can be harmful in prediction
- Selection vs. extraction
- Types of selection methods

MOTIVATION / 2

- In high-dimensional data sets, we often have prior information that many features are either irrelevant or of low quality
- Having redundant features can cost something during prediction (money or time)
- Many models require $n > p$ data. Thus, we either need to
  - adapt models to high-dimensional data (e.g., regularization)
  - design entirely new procedures for $p > n$ data
  - use filter preprocessing methods from this lecture

FEATURE SELECTION VS. EXTRACTION / 2

- Both FS and FE contribute to
  1) dimensionality reduction and 2) simplicity of models
- FE can be unsupervised (PCA, multidim scaling, manifold learning) or supervised (supervised PCA, partial least squares)
- FE can produce lower dim projections which can work better than FS; whether FE+model is interpretable depends on how interpretable extracted features are