

- classification vs regression
- descriptive vs predictive methods
- supervised vs unsupervised learning
- attributes and target
- one-hot-encoding
- training, validation and test sets
- generalization error
- hyper-parameters and cross-validation
- under-fitting vs over-fitting
- Decision Tree Classifier
  - gini index
  - maximum tree depth
  - Hunt's algorithm
- Naive Bayes Classifier    [L17](#)
  - Bayes Theorem
  - prior vs posterior probabilities
  - pseudo counts
  - independent vs conditionally independent events
- K-Means Clustering    [L22](#)
  - SSE
  - centroids
  - mean minimize cluster SSE
  - K-Means clustering always converges proof
- Natural Language Processing
  - bag-of-words strategy
  - corpus and vocabulary
  - stop words
  - n-grams
  - document term frequency matrix
  - tf-idf matrix    [L23, 25](#)
  - sparse matrices    [L26](#)