Reading Material (FINAL)

• Learning from Data: Chapter 1 (PDF on website)

• ESL [The Elements of Statistical Learning]:
  • Chapter 1
  • Chapter 2: 2.1, 2.2, 2.3.1 (first part), 2.3.2, 2.4 (up to equation 2.15), read Chapters 7.1 and 7.2 (will get back to these)

• PRML [Pattern Recognition and Machine Learning]
  • Chapter 1: introduction, 1.2 (up to 1.2.4 included), 1.3, 1.5.1, 1.5.2, 1.5.5
  • Chapter 2: everything up to 2.3 included (excluding 2.1.1, 2.2.1, 2.3.5-2.3.8) and 2.5.2
ClassificaEon
Topics & Readings

• Linear Models for Classification.
• Bayes classifier, LDA and QDA.
• Simplifications: Naïve Bayes classifier and LDA.

Chapter 4:
• 4.1
• 4.2
• 2.3.1
• 4.3 (including 4.3.2 and 4.3.3., excluding 4.3.1)
• 2.4 (part on Bayes classifier)
• 6.6.3 (Naïve Bayes classifier)
• 4.4, excluding 4.4.4 (logistic regression)
• 4.5, excluding 4.5.2 (to be covered with SVM)
Readings
Separating hyperplanes-Perceptron

• ESL: Chapter 4
  • 4.4 (up to “For the multiclass case (K >=3)...” excluded.
  • 4.4.5
  • 4.5, 4.5.1

• PRML:

• For Newton-Raphson, Gradient Descent, and Stochastic Gradient Descent: only material presented in class
Readings
Model Assessment & Selection

• ESL: Chapter 7
• 7.1, 7.2
• 7.3, excluding from p. 224, after "For linear model..." to the end
• 7.3.1, 7.4, 7.5, 7.6
• 7.7 and 7.8 read-only (definitions of AIC-BIC)
• 7.10 (7.10.1 and 7.10.2 only read)
• 7.11 (Only until the sentence "One can improve.." on page 252)
Readings

Variable selection & Regularization

• ESL: Chapter 3
• 3.3 (Forward Backward selection)
• 3.4
• 3.4.1 (Shrinking – Ridge Regression)
• 3.4.2 (Shrinking/Selection, Lasso)
Readings
Variable selection & Regularization

- ESL: Chapter 3
- 3.3 (Forward Backward selection)
- 3.4
- 3.4.1 (Shrinking – Ridge Regression)
- 3.4.2 (Shrinking/Selection, Lasso)
- 4.4.4 (Regularized Logistic Regression)
VC Dimension

• ESL: Chapter 7, Paragraph 9
VC Dimension

• ESL: Chapter 7, Paragraph 9
Neural Networks

• ESL: Chapter 11
  – Read only 11.2 (until end of page 390)
  – 11.3
  – 11.4
  – Read 11.5
Optimal Separating Hyperplanes and Linear SVM’s

- ESL: Chapter 4 & 12
  - 4.5.2
  - 12.1, 12.2, 12.2.1
  - 12.3.2, 12.3.6, 12.3.7
Principal Component Analysis

- ESL: Chapter 14
  - 14.5, 14.5.1
  - 14.5.4 (read only)
Clustering (K-means, GMM)

• ESL: Chapter 13
  – K-means: 13.1, 13.2, no 13.2.2
  – Gaussian Mixtures/EM: 13.2.3, 6.8, 8.5 (see also notes on EM)
  – K-nearest neighbor classifiers 13.3