This document describes, for each topic, the level of detail required for the material presented during the lectures (see the slides). The level of detail relates to what has been presented during the lectures, so all details means that all details presented during the lectures (see the slides) is required while main idea means that only the understanding of the concepts presented is required, while the details (e.g., details of propositions, proofs, formulas) are not required. Note that the background material (probability, linear algebra) is assumed to be known at the level of detail used during the presentation of the topics below. For some propositions the corresponding proposition in the book is used for clarity.

1. Learning Model
   - All details presented in class required, including: definitions, propositions’ statements, proof of Corollary 2.3 [UML], definition of agnostic PAC learnability for general loss functions

2. Uniform Convergence
   - Lemma 4.2 [UML]: statement and proof with all details;
   - Definition of uniform convergence property: main idea (details of definition not required)
   - Corollary 4.6 [UML]: only main idea and bound on the number of samples

3. Linear Models
   - linear predictors/models: definitions with all details
   - linear classification, perceptron: definitions and algorithm in detail
   - proposition on perceptron convergence: only main idea (as in slide “Perceptron: Notes”)
   - linear regression: definitions, matrix form, derivation best predictor, use of generalized inverse in detail (derivation generalized inverse: not required)
   - logistic regression: definition, loss function, equivalence MLE solution and ERM solution in detail

4. Basics of Statistics
   - definition confidence interval, definition rejection region: details
   - hypothesis testing: main idea; hypothesis testing rejection rule in detail
   - everything else: main idea

5. Bias-Complexity
   - No Free Lunch (NFL) theorem, NFL and priori knowledge: only main idea
   - approximation error + estimation error, complexity and error decomposition: details
6. VC-dimension
   • restrictions, growth functions: definitions in detail
   • growth function and uniform convergence: only main idea
   • shattering, VC-dimension: definitions in detail
   • Fundamental Theorems of Statistical Learning: only main idea; bound on generalization error in detail

7. Nonuniform Learnability
   • Structural risk minimization (SRM), description language: only main idea
   • Minimum Description Length: main idea; upper bound on generalization error in detail

8. Model Selection and Validation
   • model selection with SRM: main idea
   • validation: main idea
   • validation for model selection: main idea
   • model-selection curve, train-validation-test split, k-fold cross validation: all details
   • what if learning fails: main idea

9. Regularization and Feature Selection
   • Regularized Loss Minimization, Tikhonov Regularization: all details
   • Ridge Regression, derivation of optimal solution: all details
   • Stability, stability rules do not overfit, Tikhonov Regularization as a stabilizer: main idea
   • Fitting-stability tradeoff definition and considerations: all details; guarantee results for the choice of lambda: only idea
   • $l_1$ regularization, LASSO: all details
   • subset selection, forward selection, backward selection, without and with validation data: all details; pseudocode: main idea and structure required (details not required)

10. SVM
    • hard-SVM optimization problem and quadratic formulation: all details (no proof of equivalence between the two formulations)
    • soft-SVM optimization problem: all details
    • gradient descent (GD): all details; GD guarantees: main idea
    • stochastic gradient descent (SGD): main idea
    • SGD for soft-SVM: all details
    • Hard-SVM dual formulation: main idea; final optimization problem: all details (no derivation required)
    • Definition of Kernel: all details
    • SVM and Kernels: main idea
    • SVM for regression: all details only for optimization problem and support vectors definition
11. Neural Networks
- Neuron, activation function, network architecture, point of view of one node, hypothesis set: all details
- Expressiveness: main idea of each statement
- Sample complexity, runtime of learning NNs: main idea
- Forward propagation algorithm: all details
- Backpropagation algorithm: main idea (pseudocode: only main structure)
- Regularized NNs: main idea

12. Clustering
- Unsupervised learning introduction: all details
- Clustering definition and difficulties: main idea
- Model for clustering: all details
- Linkage based clustering: all details
- k-means clustering: all details
- k-medoids and k-median objective functions: main idea
- clustering for predictions: main idea
- Gaussian Mixture Model: all details
- GMM and clustering: all details
- GMM and MLE: all details
- EM (main idea): all details
- EM for GMM and convergence: main idea

13. Dimensionality Reduction - PCA
- Dimensionality reduction definition: all details
- PCA definition, derivation of principal components: only final result (no derivation)
- PCA projected variance: only final result (no derivation)
- PCA and MSE: only final result (no derivation)
- PCA choice of r and PCA as decomposition: all details