**Machine Learning Foundations**

|  |  |
| --- | --- |
| Lecture 1 | the learning problem:   1. Course Introduction 2. What is Machine Learning 3. Applications of Machine Learning 4. Components of Machine Learning 5. Machine Learning and Other Fields |
| Lecture 2 | learning to answer yes/no:   1. Perceptron Hypothesis Set 2. Perceptron Learning Algorithm (PLA) 3. Guarantee of PLA 4. Non-Separable Data |
| Lecture 3 | types of learning:   1. Learning with Different Output Space 2. Learning with Different Data Label 3. Learning with Different Protocol 4. Learning with Different Input Space |
| Lecture 4 | feasibility of learning:   1. Learning is Impossible? 2. Probability to the Rescue 3. Connection to Learning 4. Connection to Real Learning |
| Lecture 5 | training versus testing:   1. Recap and Preview 2. Effective Number of Lines 3. Effective Number of Hypotheses 4. Break Point |
| Lecture 6 | theory of generalization:   1. Restriction of Break Point 2. Bounding Function: Basic Cases 3. Bounding Function: Inductive Cases 4. A Pictorial Proof |
| Lecture 7 | the VC dimension:   1. Definition of VC Dimension 2. VC Dimension of Perceptrons 3. Physical Intuition of VC Dimension 4. Interpreting VC Dimension |
| Lecture 8 | noise and error:   1. Noise and Probabilistic Target 2. Error Measure 3. Algorithmic Error Measure 4. Weighted Classification |
| Lecture 9 | linear regression:   1. Linear Regression Problem 2. Linear Regression Algorithm 3. Generalization Issue 4. Linear Regression for Binary Classification |
| Lecture 10 | logistic regression:   1. Logistic Regression Problem 2. Logistic Regression Error 3. Gradient of Logistic Regression Error 4. Gradient Descent |
| Lecture 11 | linear models for classification:   1. Linear Models for Binary Classification 2. Stochastic Gradient Descent 3. Multiclass via Logistic Regression 4. Multiclass via Binary Classification |
| Lecture 12 | nonlinear transformation:   1. Quadratic Hypotheses 2. Nonlinear Transform 3. Price of Nonlinear Transform 4. Structured Hypothesis Sets |
| Lecture 13 | hazard of overfitting:   1. What is Overfitting? 2. The Role of Noise and Data Size 3. Deterministic Noise 4. Dealing with Overfitting |
| Lecture 14 | regularization:   1. Regularized Hypothesis Set 2. Weight Decay Regularization 3. Regularization and VC Theory 4. General Regularizers |
| Lecture 15 | validation:   1. Model Selection Problem 2. Validation 3. Leave-One-Out Cross Validation 4. V-Fold Cross Validation |
| Lecture 16 | three learning principles:   1. Occam's Razor 2. Sampling Bias 3. Data Snooping 4. Power of Three |

**Machine Learning Techniques**

|  |  |
| --- | --- |
| Lecture 1 | linear support vector machine:   1. Course Introduction 2. Large-Margin Separating Hyperplane 3. Standard Large-Margin Problem 4. Support Vector Machine 5. Reasons behind Large-Margin Hyperplane |
| Lecture 2 | dual support vector machine:   1. Motivation of Dual SVM 2. Lagrange Dual SVM 3. Solving Dual SVM 4. Messages behind Dual SVM |
| Lecture 3 | kernel support vector machine:   1. Kernel Trick 2. Polynomial Kernel 3. Gaussian Kernel 4. Comparison of Kernels |
| Lecture 4 | soft-margin support vector machine:   1. Motivation and Primal Problem 2. Dual Problem 3. Messages behind Soft-Margin SVM 4. Model Selection |
| Lecture 5 | kernel logistic regression:   1. Soft-Margin SVM as Regularized Model 2. SVM versus Logistic Regression 3. SVM for Soft Binary Classification 4. Kernel Logistic Regression |
| Lecture 6 | support vector regression:   1. Kernel Ridge Regression 2. Support Vector Regression Primal 3. Support Vector Regression Dual 4. Summary of Kernel Models |
| Lecture 7 | blending and bagging:   1. Motivation of Aggregation 2. Uniform Blending 3. Linear and Any Blending 4. Bagging (Bootstrap Aggregation) |
| Lecture 8 | adaptive boosting:   1. Motivation of Boosting 2. Diversity by Re-weighting 3. Adaptive Boosting Algorithm 4. Adaptive Boosting in Action |
| Lecture 9 | decision tree:   1. Decision Tree Hypothesis 2. Decision Tree Algorithm 3. Decision Tree Heuristics in C&RT 4. Decision Tree in Action |
| Lecture 10 | random forest:   1. Random Forest Algorithm 2. Out-Of-Bag Estimate 3. Feature Selection 4. Random Forest in Action |
| Lecture 11 | gradient boosted decision tree:   1. Adaptive Boosted Decision Tree 2. Optimization View of AdaBoost 3. Gradient Boosting 4. Summary of Aggregation Models |
| Lecture 12 | neural network:   1. Motivation 2. Neural Network Hypothesis 3. Neural Network Learning 4. Optimization and Regularization |
| Lecture 13 | deep learning:   1. Deep Neural Network 2. Autoencoder 3. Denoising Autoencoder 4. Principal Component Analysis |
| Lecture 14 | radial basis function network:   1. RBF Network Hypothesis 2. RBF Network Learning 3. k-Means Algorithm 4. k-Means and RBF Network in Action |
| Lecture 15 | matrix factorization:   1. Linear Network Hypothesis 2. Basic Matrix Factorization 3. Stochastic Gradient Descent 4. Summary of Extraction Models |
| Lecture 16 | finale:   1. Feature Exploitation Techniques 2. Error Optimization Techniques 3. Overfitting Elimination Techniques 4. Machine Learning in Practice |