Machine Learning (2108, Fall) Final Exam Range

Professor Pei-Yuan Wu

Basic (60%)

Advanced (40%)

Bonus (??%)

- **Linear regression**
  - Linear + high order polynomial terms
  - Solve parameters in linear regression analytically (課堂上推導)

- **Bias and Variance**
  - Unbiased estimation of mean/variance from data
  - Decomposition of estimation error into bias and variance (課堂上推導)
  - N-fold cross-validation
- **L1/L2-regularization**

- **Classifier with generative model**
  - Fit a generative model from data, where each class is a (multi-variate) Gaussian distribution with prior π, mean $\boldsymbol{\mu}$, and covariance matrix $\boldsymbol{\Sigma}$.
  - Maximum likelihood estimation of π, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$

- **Gradient descent**
  - Role of learning rate
  - Adagrad, RMSProp, Stochastic gradient descent, Momentum concept

- **Logistic regression**
  - Logistic model and sigmoid function
  - Cross entropy loss
  - Refine parameters in logistic model by gradient descent.

- **Support Vector Machine (SVM)**
  - Margin, hinge loss, role of C
  - SVM primal optimization problem (in terms of w)
  - Representer theorem
  - SVM dual optimization problem (in terms of α)
  - Definition/properties of support vectors
  - Kernel-SVM

- **Convex optimization:**
  - ➢ Definition of convex function/set
  - ➢ Non-negative addition of convex functions is still convex (課堂中推導)
  - ➢ Convex optimization problem has unique local minimum (課堂中推導)
  - ➢ Connection between primal and dual optimization problems
  - ➢ Weak duality theorem
  - ➢ Strong duality theorem

  Semi-supervised

  Semi-supervised generative model

  Self-training, why regression/hard pseudo label does not work

  Entropy-based regularization

  Smoothness evaluated by regularization term by Graph Laplacian

- **Neural Network**
  - ➢ Basic architecture
  - ➢ Compute output from input by feed-forward
  - ➢ Refine parameters by back-propagation
  - ➢ One-hot representation of ground truth
  - ➢ Activation functions: sigmoid, ReLU, maxout
  - ➢ Dropout
  - ➢ Softmax layer

- **Convolutional Neural Network**
  - ➢ Convolution, max-pooling
  - ➢ Concept of channel, receptive field
- **Basic architecture of RNN and LSTM**

- **Clustering**
  - ➢ Hierarchical Agglomerative Clustering
  - ➢ Hard membership by K-means
  - ➢ Soft membership by EM.

- **Expectation maximization (EM) algorithm**
  - ➢ Gaussian mixture model (GMM)
  - ➢ Expectation/Maximization step
  - ➢ Fit a GMM model by EM.

- **Dimensionality reduction**
  - ➢Principle Component Analysis: principle axes, principle components
  - ➢Locally linear embedding
  - ➢t-SNE
  - ➢Deep auto-encoder

- **Ensemble**
  - ➢Bagging: Random forest
  - ➢Boosting: AdaBoost, Gradient boosting (投影片+課堂中推導)
  - ➢Voting
  - ➢Stacking