Final Review

Machine Learning
Spring 2021
Final Exam

Mon May 3, 2021
10:30am - 12:30pm
Canvas (online)

https://www.cs.utah.edu/~zhe/teach/cs6350.html
Final Exam Policy

• Open book/website
• Do not search/post questions online
• Single/multiple choice
• Text description (no need to write math)...
• Problems can vary for different students
Coverage

• The whole semester

Suggestion: (1) review the lecture slides (2) carefully go through the paper problems in the homework assignments
Throughout this semester

- What is machine learning
- A variety of learning models and algorithms
- Different learning principles
- Ideas about learning theory
- Implementation and practice
Learning = generalization

“A computer program is said to learn from experience \( E \) with respect to some class of tasks \( T \) and performance measure \( P \), if its performance at tasks in \( T \), as measured by \( P \), improves with experience \( E \).”

Tom Mitchell (1999)
The formulation of machine learning

\[ X: \text{Instance Space} \]

The set of examples that need to be classified

\[ Y: \text{Label Space} \]

The set of all possible labels

Eg: The set of all possible names, documents, sentences, images, emails, etc

Eg: \{Spam, Not-Spam\}, \{+, -\}, etc.

Target function \[ y = f(x) \]

The goal of learning: Find this target function

Learning is search over functions
Supervised learning

$X$: Instance Space
The set of examples

$Y$: Label Space
The set of all possible labels

Target function
$y = f(x)$

Learning algorithm only sees examples of the function $f$ in action

Labeled training data

(x_1, f(x_1))
(x_2, f(x_2))
(x_3, f(x_3))
... (x_N, f(x_N))

Learning algorithm

A learned function $g: X \rightarrow Y$

Can you think of other protocols?
A variety of learning models

• Decision trees
• Boosting and ensembles: Adaboost, random forest
• Least mean square method for linear regression
• Perceptron
• Support vector machines (SVMs)
• Kernel Perceptron
• Kernel SVM
• Artificial neural networks, back propagation
• Logistic regression
How to categorize ML models: linear/nonlinear

- Linear models
  - Least mean square method
  - Perceptron
  - SVM
  - Logistic regression

- Nonlinear models (how to construct?)
  - Kernel methods: Kernel Perceptron, Kernel SVM
  - Neural networks
  - Decision trees
  - Adaboost and random forest
How to categorize ML models: probabilistic / non-probabilistic

- Probabilistic learning
  - Logistic regression

- Non-probabilistic models
  - Neural networks
  - Decision trees
  - SVM
  - Perceptron
  - ...

The all have probabilistic version!
Basic concepts/conclusions in computational learning theory

• PAC framework
• What is PAC learnability?
• What is sample complexity bound? What is the generalization error bound?
• VC dimension (why?)
• How is the large margin principle derived by the generalization error bound?

No need to memorize. But you need to really understand it!
Principles to Learn ML models

• Empirical risk minimization (ERM)
  – Least mean square method
  – Perceptron
  – Artificial neural networks

• Regularized empirical risk minimization (RERM)
  – SVM
  – MAP estimation for logistic regression

• Maximum a posterior (MAP)

• Maximum likelihood (MLE)

Connection to ERM and RERM
Learning algorithms

• Gradient descent
• Stochastic gradient descent
• What are their advantages and disadvantages?
Practice

• How to select hyper-parameters?
  – Hold out dataset
  – Cross validation.
Review suggestions

• Go through all the lecture slides. Understand the details
  
  • Decision trees
  • Boosting and ensembles: Adaboost, random forest
  • Least mean square method for linear regression
  • Perceptron
  • Support vector machines (SVMs)
  • Kernel Perceptron
  • Kernel SVM
  • Artificial neural networks, back propagation
  • Logistic regression
Review suggestions

• Go through all the homework problems you have worked. Do not need to go through the programming problems.
Good Luck!