CS-584 Machine Learning

Summer 2018 (Tue, Thu 5:00pm – 8:10pm SB-238)

http://www.cs.iit.edu/∼cs584/

Administrative information

Instructor

- Gady Agam (SB-237e, x7-5834, *agam@iit.edu*)
- Office hours: Tue, Thu 8:10pm - 8:40pm

Teaching assistants

- TBD
- Office hours: TBD

Accounts

- Web (information): http://www.cs.iit.edu/∼cs584/
- Slack (discussion): https://cs584iit.slack.com/
- Bitbacket Git (submission): share private repository with cs584iit

Grading

component	description	weight
participation	up to 4 unjustified missed classes ⇒ full credit	5%
assignments	4-5 TBD	25%
project	project+presentation	20%
exams	2-3 TBD (2 page notes per exam)	50%
total		100%

Notes:

1. There is an additional mandatory assignment (assignment 0) which does not carry any credit. There is a penalty of 5% for not submitting this assignment.
2. A certain percentage of the students may be invited to discuss their assignments.
3. Late days: there is a total of 6 "late days" for all the assignments to cover reasons such as not feeling well, being busy, etc. After that 1 late day = -25%. Late days include weekends and university holidays. Up to 2 late days may be applied to each assignment. The final project can not be late. Assignments can not be submitted after classes end.
4. Each member of this course bears responsibility for maintaining the highest standards of academic integrity. All breaches of academic integrity must be reported immediately. Copying of programs or answers from any source (e.g. other students or the web) is considered to be a serious breach of academic integrity.

Course outline

What to expect from this course

Machine learning can be covered at different levels. The focus of this course is the understanding of algorithms and techniques used in machine learning. Students in the course are expected to write computer programs (Python) implementing different techniques taught in the course. The course requires mathematical background and some programming experience. This course does *not* intend to teach how to use a specific application software.

Objectives

- Introduce the fundamental problems of machine learning.
- Provide understanding of techniques, mathematical concepts, and algorithms used in machine learning to facilitate further study in this area.
- Provide understanding of the limitations of various machine learning algorithms and the way to evaluate performance of machine learning algorithms.
- Provide pointers into the literature and exercise a project based on literature search and one or more research papers.
- Practice software implementation of different concepts and algorithms covered in the course.

Overview

1. Introduction: overview of machine learning, related areas, applications, software tools, course objectives.
2. Parametric regression: linear regression, polynomial regression, locally weighted regression, numerical optimization, gradient descent, kernel methods.
3. Generative learning: Gaussian parameter estimation, maximum likelihood estimation, MAP estimation, Bayesian estimation, bias and variance of estimators, missing and noisy features, nonparametric density estimation, Gaussian discriminant analysis, naive Bayes.
4. Discriminative learning: linear discrimination, logistic regression, logit and logistic functions, generalized linear models, softmax regression.
5. Neural networks: the perceptron algorithm, multilayer perceptrons, backpropagation, nonlinear regression, multiclass discrimination, training procedures, localized network structure, dimensionality reduction interpretation, deep learning, convolutional neural networks.
6. Support vector machines: functional and geometric margins, optimum margin classifier, constrained optimization, Lagrange multipliers, primal/dual problems, KKT conditions, dual of the optimum margin classifier, soft margins, kernels, quadratic programming, SMO algorithm.
7. Graphical and sequential models: Bayesian networks, conditional independence, Markov random fields, inference in graphical models, belief propagation, Markov models, hidden Markov models, decoding states from observations, learning HMM parameters.
8. Unsupervised learning: K-means clustering, expectation maximization, Gaussian mixture density estimation, mixture of naive Bayes, model selection.
9. Dimensionality reduction: feature selection, principal component analysis, linear discriminant analysis, factor analysis, independent component analysis, multidimensional scaling, manifold learning.

10. Final project: students present selected topics and develop software implementation of related techniques based on the review of relevant literature. The work should be summarized in a concluding report which should include simulation results. A list of possible topics will be available prior to the project selection due date.

Books

1. Elements of Statistical Learning, T. Hastie, R. Tibshirani and J. Friedman, Springer, 2001.
2. Machine Learning, E. Alpaydin, MIT Press, 2010.
3. Pattern Recognition and Machine Learning, C. Bishop, Springer, 2006.
4. Machine Learning: A Probabilistic Perspective, K. Murphy, MIT Press, 2012.
5. Pattern Classification, R. Duda, E. Hart, and D. Stork, Wiley-Interscience, 2000.
6. Machine Learning, T. Mitchell, McGraw-Hill, 1997.