Boston University, College of Engineering
Department of Electrical and Computer Engineering

ENG EC503 (Ishwar) – Fall 2021

LEARNING FROM DATA

Instructor: Prof. Prakash Ishwar (e-mail: pi@bu.edu)
Teaching Assistant: Jimmy Zhang (e-mail: zhangjim@bu.edu)
Grader: Ron Rattanakornphan (e-mail: ronrat@bu.edu)

Times: all times are Boston times.
Lectures: Tue + Thr, 1:30-3:15 pm (EPC 204)
Discussions: Mon 6:30-8:00 pm (MUG 205)
Instructor's office hours: Tue + Thr 6:00-7:30 pm (location to be decided)
TA's office hours: Wed + Fri (times + location to be decided)

Description:
This is an introductory graduate course in machine learning covering the basic principles and methods of four major non-sequential supervised and unsupervised learning problems namely, classification, regression, clustering, and dimensionality reduction. A variety of contemporary applications will be explored through homeworks and a project.

Prerequisites:
Probability (EK381 or equivalent), Linear Algebra (EK102 or equivalent), Multivariate Calculus (MA225 or equivalent), *Matlab* (EK125 or equivalent).

Syllabus:

- *Key concepts, terms, and technical tools*: training, testing, cross-validation, performance-evaluation, under- and over-fitting, inner-product, norm, orthogonal projection, eigen-decomposition, SVD, empirical first- and second-order statistics, gradient, Hessian, optimization, convex set, convex function, subgradient, stochastic sub-gradient descent.
- *Nearest-Neighbor (NN) methods*: k-NN classification and regression, k-centers clustering and vector-quantization, "curse" of dimensionality.
- *Linear methods*: *Classification:* artificial neuron, Fisher-LDA, logistic-loss, single-layer feed-forward artificial neural network (ANN), support-vector machine (SVM); *Regression:* ordinary least squares (OLS), ridge-regression, support-vector regression (SVR); *Dimensionality-reduction:* random-projections, principal-component analysis (PCA).
- *Kernel-methods*: Representer-theorem, kernelization of learning and inference, kernel versions of kNN, SVM, ridge-regression, k-means, PCA, Isomap.
- *Multilayer feed-forward ANNs*: architecture, universal-approximation, backpropagation algorithm;
- *Selected Topics (as time permits)*: document classification with bag-of-words model, robust-regression, sparse-regression, spectral clustering, decision trees and random forests, boosting, probabilistic framework for learning and inference.

Grading:

20%	Homeworks	∼10 sets of *Matlab* and/or analytical exercises. Due: 10:55pm of due date (upload report to Gradescope + code to Blackboard). Late submission penalty: 5 points per minute.
55%	Exams	3 exams. Dates (tentative): Sep. 30, Nov. 11, Finals week.
20%	Project	Team project involving algorithm development. Details to follow later in the semester. Presentation: Dec. 9, 10, 6:00–8:30pm in PHO 211.
5%	Class-participation	For constructive and proactive engagement during lectures, discussions, office hours, and on Piazza.

Outcomes: As an outcome of successfully completing this course, students will:

- understand basic theoretical principles and algorithmic methods of learning from data,
- be able to use computer-based machine learning tools,
- know how to analyze and extract information from real-world data,
- be able to select and optimize appropriate machine learning methods and tools for various real-world problems,
- be able to understand and communicate key ideas from articles and technology related to machine learning.

Web site: `http://learn.bu.edu` for *registered* students only. Will contain lecture slides, notes, links, discussion materials, and other useful information related to the course.

References: This course has no textbook. You will need to rely on lectures, discussions, office hours, and course materials that will be uploaded regularly to the course web site. Below is a list of reference books which you may consult should you like to explore further. Each book is on reserve in the Science and Engineering Library (max. 2 hour check-out period; also available electronically online).

- K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.
- T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: data mining, inference, & prediction*. Springer, 2nd edition - 2009.
- C.M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*. Wiley-Interscience, 2nd edition - 2000.

Matlab: Each computer assignment will involve the use of *Matlab* to illustrate and compare the main methods discussed in the lectures and discussions. You could run *Matlab* on your own computer using BU's site-wide licensing program. See:

`http://www.bu.edu/tech/services/cccs/desktop/distribution/mathsci/matlab/`

Alternatively, you could run *Matlab* remotely on BU machines. See:

`https://www.bu.edu/engit/knowledge-base/citrix/citrix-how-to/`

You may use your own computers to complete the homeworks and/or project. You may also use the workstations in PHO 305 (Tue + Fri 10am - 12pm). Students who registered early should automatically have card-access to this room and have a computer account. If not, apply for card-access using Zaius: `http://www.bu.edu/dbin/eng/zaius/` and contact the lab administrator `enghelp@bu.edu` for an account.

Academic integrity, plagiarism:

Collaboration is essential for the course project, permitted on homeworks, but illegal in exams.

In project reports, you may *paraphrase* relevant ideas from references, but not quote sentences verbatim from them.

Homework collaboration = only discussion. If you allow your solution or code to be viewed by anyone or reversely you see someone's solution or code then you have gone beyond collaboration. You may *discuss* problems, but you must create a solution by yourself. **If there is collaboration in a homework, all collaborators must be explicitly acknowledged and the nature and extent of collaboration must be clearly explained. Each collaborator must turn in his/her <u>individual</u> analysis/code and description of results.**

All solutions (including code) will be automatically checked for plagiarism against solutions from all registered students, solutions from previous semesters, and also solutions available from online sources.

The student handbook defines academic misconduct as follows: *"Academic misconduct occurs when a student intentionally misrepresents his or her academic accomplishments or impedes other students' chances of being judged fairly for their academic work. Knowingly allowing others to represent your work as theirs is as serious an offense as submitting another's work as your own."* Please see the student handbook for procedures that will follow should academic misconduct be discovered.

Inclusion: I consider the classroom to be a place of learning where all individuals are expected to contribute to provide a respectful, welcoming and inclusive environment for every member of the class irrespective of how they identify themselves.

Disability accommodations: If you are a student with a documented disability or believe you might have a disability that requires accommodations, requests for accommodations must be made in a timely fashion to Disability & Access Services, 25 Buick St, Suite 300, Boston, MA 02215; 617-353-3658 (Voice/TTY). Students seeking academic accommodations must submit appropriate medical documentation and comply with the established policies and procedures `http://www.bu.edu/disability/accommodations/`

BU COVID 19 safety policies: Masks are required and must be worn over the mouth and nose at all times when in classrooms. In order to attend class, students should be compliant with all BU policies for vaccination, daily symptom checks, testing, social distancing, and mask wearing and **should be prepared to show proof that they are compliant**. For BU COVID 19 policies please visit: `https://www.bu.edu/back2bu/`

Illness/Medical issues: will be resolved by case-specific discussions with the instructor. Student must inform instructor by email as soon as possible and be prepared to provide suitable documentation. May also require contacting BU's disability access services.