Course Description

ECE 445: Machine Learning for Engineers (Topics in ECE)
Fall 2019

Schedule — TTh 5:00 – 6:20 PM
Place — RWH-206

Instructor

Waheed U. Bajwa
723 CoRE, Tel. 848-445-8541
http://inspirelab.us

Prerequisites

Enrolled students must have taken undergraduate courses in probability theory and linear algebra. The course will also require extensive programming, especially as part of the term project. In keeping with the industry standards, all programming will need to be done in Jupyter notebooks (http://jupyter.org/) using either Julia, Python, or R (individual students will get to pick any one of these languages in most assignments). Submission of all programming assignments/projects will take place through GitHub (http://github.com/). In many instances, students will be forbidden from using popular machine learning packages such as scikit-learn for assignments.

Learning Outcomes

- Mastery of the basic terminology and concepts in machine learning
- Understanding of the basic building blocks of practical machine learning systems
- Mathematical understanding of commonly used machine learning algorithms
- Ability to develop basic machine learning systems from scratch in Jupyter
- Recognition of common pitfalls that come with machine learning systems

Who should take this course?

- Students interested in machine learning and data science careers in the industry
- Students interested in applying machine learning techniques in their own disciplines
- Students interested in graduate school with a focus on machine learning

Who should not take this course?

- Students interested only in getting an easy 'A' grade
- Students who are uncomfortable with abstract and/or rigorous mathematics
- Students who are afraid of putting in four to eight hours per week for this course
- Students who are uncomfortable with the concept of programming in Jupyter

Required Texts

Much of the material taught in this class will come from the following two texts:

G. James, D. Witten, T. Hastie and R. Tibshirani
An Introduction to Statistical Learning with Applications in R
Springer; 1st ed. 2013, Corr. 7th printing 2017 edition (September 1, 2017)
Available at: `http://www-bcf.usc.edu/~gareth/ISL/`

T. Hastie, R. Tibshirani, and J. Friedman
The Elements of Statistical Learning: Data Mining, Inference, and Prediction
Springer; 2nd edition (2016)
Available at: `https://web.stanford.edu/~hastie/ElemStatLearn/`

Students should also maintain detailed notes of the material taught in class. In addition, students will be occasionally provided material and internet links for further referencing.

Tentative Course Outline

Weeks 1–3:

- Review of basic probability theory and linear algebra concepts
- Introduction to machine learning and its basic terminology
- Understanding the machine learning pipeline
- Feature engineering and representation learning
- Introduction to the role of numerical computations in machine learning

Week 4:

- Introduction to basic machine learning algorithms
- Theoretical description of machine learning algorithms
  - Statistical risk minimization and empirical risk minimization
  - Role of geometry in machine learning algorithms
  - Optimization and computational aspects of machine learning algorithms

Weeks 5–7:[1] Classification

- Parametric models
  - Naive Bayes classification

[1] First in-class exam will take place during this period.

  - Linear classification (including support vector machines)
- Non-parametric models
  - Nearest-neighbor classification
  - Kernel support vector machines
- Frequentist versus Bayesian philosophy
- Brief introduction to neural networks

Weeks 8–9: Practical machine learning systems

- Testing, training, and validation data, and cross-validation
- Distribution mismatch (transfer learning) and missing data
- Bias–variance tradeoff, overfitting, and Occam's razor
- (Stochastic) Gradient descent

Weeks 10–11: Regression

- Linear regression
- Nonlinear regression
- Penalized regression

Weeks 12–14:[2] Clustering

- Gaussian mixture model
- Expectation–maximization algorithm
- k-means and related algorithms

Week 14: Introduction to advanced topics (subject to availability of time)

- Hidden Markov models and graphical models
- Random forests and ensemble learning
- Ranking and recommendation systems
- Density estimation

[2] Second in-class exam will take place during this period.

Course Policies

ECE 445: Machine Learning for Engineers (Topics in ECE)
Fall 2019

Schedule — TTh 5:00 – 6:20 PM
Place — RWH-206

Instructor

Waheed U. Bajwa
723 CoRE, Tel. 848-445-8541
http://inspirelab.us

Means of Communication

Piazza (Class Link): http://bit.ly/ECE445f19Piazza **(all non-personal questions)**
E-Mail: waheed.bajwa@rutgers.edu **(personal questions)**
Twitter: @SigProcessing (#RUECE445) **(light discussions)**
Sakai: http://bit.ly/ECE445f19 **(course management)**

Please sign-up for Piazza using the following link: piazza.com/rutgers/fall2019/ece445

Office Hours

By *appointment* on Wednesdays, 1:00 – 2:00 PM
An appointment must be made by Tuesday night 9 PM at the latest.

Course Policies

The final course grade will be based upon:

1. Homework (5%)
2. Class participation (2.5%)
3. Mini Jupyter exercises (20%)
4. In-class exam #1 (30%)
5. In-class exam #2 (30%)
6. Term project (12.5%)

Late homework and assignment submission policy: Every student gets a grace period of up to 3 days for a maximum of two homeworks and two assignments. Utilization of the first grace period for homework (resp., assignment) is without any penalty. Utilization of the second grace period comes with a 30% penalty. No late submissions will be accepted

from a student who has utilized both these grace periods, regardless of the emergency or unique circumstances.

Exam policy: Exams will be closed book and closed notes. Students can bring in one, double-sided letter-sized page for each of the in-class exams. As a general policy, there will be no makeup exams. I will allow exceptions for rare emergency situations, but this would require at least 7 days advance approval to skip an exam. Any one not appearing in an exam without such prior approval will automatically get a 0.

Grading policy: Grades will be assigned on a relative basis. The relative scale though will vary based upon the performance of the overall class. In an ideal setting, students above class average will get B and higher and students at or below class average will get C+ and lower, respectively. If the class performs really well, however, then the B will turn into a B+. Similarly, if the class performs really bad then the B will turn into a C+ (or even C).

Academic misconduct and plagiarism warning

It is important that the students enrolled in this class familiarize themselves with the Rutgers Academic Integrity Policy, `http://academicintegrity.rutgers.edu/academic-integrity-at-rutgers`, and the definition of plagiarism (`http://www.plagiarism.org/plagiarism-101/what-is-plagiarism/`). All cases of academic misconduct, whether minor or major, will not only be reported to the Student Affairs Office, but will, in most cases, also result in loss of one or more grade points.

Some tips for making learning the class material easier

Here are some tips that I hope you will remember to ensure you have a good learning experience throughout the class.

- If you feel lost during the class, please reach out to me. You will be surprised to know that I do not turn into a monster during office hours `:)`.
- Because of the mathematically intensive nature of the course, one cannot learn it by forgetting about it till it is time for an exam. It is therefore important that you try to keep up with the class material on a regular basis.
- Class lectures are not enough to learn everything about the course. Reading material and homework problems (ungraded) will be assigned on a regular basis to help you learn all the important aspects of the course. Please make sure you keep up with these things, which will be communicated via email and via the course website.
- While the percentage of the grade assigned to homeworks and Jupyter exercises is small, these two categories are going to teach you the most and ensure that you do well on the exams and the term project. The purpose of keeping the percentage small is that you don't feel pressured to blindly cheat from other students. You are encouraged to discuss things with others, but you will be doing yourself a big favor by doing the homeworks and exercises in the end by yourself.