Course Syllabus

Course Information

CS 4375 Introduction to Machine Learning Section: 0U2
Time: M/W 3:00 pm – 5:15 pm
Location: SLC 2.302

This term we will be using Piazza for class discussion. The system is highly catered to getting you help fast and efficiently from classmates, the TA, and myself. Rather than emailing questions to the teaching staff, I encourage you to post your questions on Piazza.

Find our class page at: https://piazza.com/utdallas/summer2017/cs4375mazidi/home

Professor Contact Information

Dr. Karen Mazidi
Email: Karen.Mazidi@utdallas.edu
Office: ECSS 3.203
Office phone: 972-883-3868
Office hours: M/W 5:30 – 6:30 pm

TA Information

TA information will be posted in Coursebook as soon as it is available

Course Pre-requisites

CS/SE 3341 (Probability and Statistics in CS) and CE/CS/SE/TE 3345 (Data Structures and Algorithms)

Course Description

CS 4375 Introduction to Machine Learning (3 semester credit hours) Algorithms for creating computer programs that can improve their performance through learning. Topics include: cross-validation, decision trees, neural nets, statistical tests, Bayesian learning, computational learning theory, instance-based learning, reinforcement learning, bagging, boosting, support vector machines, Hidden Markov Models, clustering, and semi-supervised and unsupervised learning techniques.

Student Learning Objectives/Outcomes

Ability to understand and apply the following concepts in machine learning:

1. Decision trees
2. Neural networks
3. Bayesian learning
4. Instance-based Learning
5. Hidden Markov models
6. Clustering
7. Reinforcement learning

Required Textbooks and Materials:

ISLR book: http://www-bcf.usc.edu/~gareth/ISL/
This free book is an introductory version of Hastie/Tibshirani *Elements of Statistical Learning*

See this link for slides and videos for their MOOC

R (install before RStudio) and Rstudio (free version)

Topics and Sequence

- Week 1 – 5/31
  - Course introduction
  - R boot camp part 1
- Week 2 – 6/5 and 6/7
  - R boot camp part 2
  - Data exploration
  - Linear regression
- Week 3 – 6/12 and 6/14
  - Linear regression, cont'd
  - PC regression
  - Cross validation
  - Bootstrap technique
- Week 4 – 6/19 and 6/21
  - Logistic regression
  - Classification
  - Data in high dimensions
  - Review for Exam 1
- Week 5 – 6/26 and 6/28
  - Select student presentations
  - Exam 1
- Week 6 – 7/3 and 7/5
  - SVM
  - Handling text data
  - Naïve Bayes
- Week 7 – 7/10 and 7/12
  - Clustering
  - Unsupervised learning
- Week 8 – 7/17 and 7/19
  - Neural Networks
  - Feature Selection
  - Algorithm Selection
  - Bayesian Learning
- Week 9 – 7/24 and 7/26
  - Markov Models
  - Decision Trees
  - Random Forests
  - Summary of Metrics
- Week 10 – 7/31 and 8/2
  - Review for Exam 2
  - Exam 2
- Week 11 – 8/7 and 8/9
  - Select student presentations
- Week 12 – Exam week (8/11 or 8/12)

Grading Policy

- Assignments (25% of the course grade): There will be approximately 9 homework assignments.
- Project (25% of the course grade).
- Exams (2) (each = 20% of the course grade).
- Quizzes and class attendance (10%).

Letter grades will be assigned according to the UTD +/- conventions.

Project Description – see eLearning for info about Part 1 and Part 2

Course & Instructor Policies

- Students should take notes by hand using either physical or electronic paper. Here is why: http://www.scientificamerican.com/article/a-learning-secret-don-t-take-notes-with-a-laptop/ Also, you can use your hand-written notes on quizzes.
- Laptop computers are allowed only during "demo" time. During demo time when laptops are allowed, please be mature, be respectful of the distractions that you may cause others if you are off task. Perhaps you think you can multi-task, but science says otherwise: http://www.npr.org/2013/05/10/182861382/the-myth-of-multitasking
- Assignments must be turned in on the due date, by midnight. It's a good idea to screen-shot your upload in eLearning. This is the only way you can prove that you really did turn in the homework.
- Late assignments are deducted by 10% on the first and second days late. After two days, the assignment will not be accepted.
- There are no make-up quizzes but one quiz will be dropped at the end of the semester.
- Makeup exams are not given unless prior permission has been granted due to extenuating circumstances.
- Do not turn in code or other work that is not your own. This will result in a zero for all parties. Discussion and collaboration are good things, turning in someone else's work as your own is not.
- **If you do not agree with a grade you have been given, you must make your case within a week of receiving the grade or the grade stands as is.**

Comet Creed

This creed was voted on by the UT Dallas student body in 2014. It is a standard that Comets choose to live by and encourage others to do the same:

"As a Comet, I pledge honesty, integrity, and service in all that I do."

Attendance Policy

Per Computer Science administration guidelines, please be aware that ***3 consecutive absences leads to one letter grade drop. Four consecutive absences lead to an F.***

UT Dallas Syllabus Policies and Procedures

The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus.

Please go to http://go.utdallas.edu/syllabus-policies for these policies.

The descriptions and timelines contained in this syllabus are subject to change at the discretion of the Professor.