10-701 and 15-781
Machine Learning

http://www.cs.cmu.edu/~guestrin/Class/10701/

Carlos Guestrin
Tom Mitchell
Syllabus

• Covers a wide range of Machine Learning techniques – from basic to state-of-the-art
• You will learn about the methods you heard about:
  – Naïve Bayes, logistic regression, nearest-neighbor, decision trees, boosting, neural nets, overfitting, regularization, dimensionality reduction, PCA, error bounds, VC dimension, SVMs, kernels, margin bounds, K-means, EM, mixture models, semi-supervised learning, HMMs, graphical models, active learning, reinforcement learning…

• Covers algorithms, theory and applications
• It’s going to be fun and hard work ☺
Prerequisites

• Probabilities
  – Distributions, densities, marginalization…

• Basic statistics
  – Moments, typical distributions, regression…

• Algorithms
  – Dynamic programming, basic data structures, complexity…

• Programming
  – Mostly your choice of language

• We provide some background, but the class will be fast paced

• Ability to deal with “abstract mathematical concepts”
Four Great TAs

• Great resource for learning, interact with them!
• Kaustav Das – kaustav@cs
• Derek Hoiem – dhoiem@cs
• Zhenzhen Kou – woomy@cs
• Daniel Neill – neill@cs
First Point of Contact

• To facilitate interaction please send all communication to your “first point of contact” according to your last name:

• A-D: contact Daniel (neill@cs)
• E-Le: contact Kaustav (kaustav@cs)
• Li-P: contact Derek (dhoiem@cs)
• Q-Z: contact Zhenzhen (woomy@cs)
Review Sessions

• Very useful!
  – Review material
  – Present background
  – Answer questions

• Tuesdays at 5pm

• Usually in NSH 3305, but check website for specific room each week
All Text Books are Optional, but very useful

- *Machine Learning*, Tom Mitchell
- *Pattern Classification (2nd Edition)*, Duda, Hart and Stork
- *Neural Networks for Pattern Recognition*, Chris Bishop
Grading

• 5 homeworks (30%)
  – First one goes out 1/19
• Final project (20%)
  – More about project after Spring Break
• Midterm (20%)
  – March 14th
• Final (30%)
  – TBD by registrar
Homeworks

• Homeworks are hard, start early 😊
• Due in the beginning of class
• 3 late days for the semester
• After late days are used up:
  – Half credit within 48 hours
  – Zero credit after 48 hours
• All homeworks must be handed in, even for zero credit
• Late homeworks handed in to Sharon Cavlovich, WEH 5311

• Collaboration
  – You may discuss the questions
  – Each student writes their own answers
  – Write on your homework anyone with whom you collaborate
Enjoy!

- ML is becoming ubiquitous in science, engineering and beyond
- This class should give you the basic foundation for applying ML and developing new methods
- The fun begins…