Learning From Data

• Machine learning used in many contexts
  ▪ Stock price prediction
    o Using economic indicators, predict if stock will go up/down
  ▪ Computational biology and medical diagnosis
    o Predicting gene expression based on DNA
    o Determine likelihood for cancer using clinical/demographic data
  ▪ Predict people likely to purchase product or click on ad
    o “Based on past purchases, you might want to buy…”
  ▪ Credit card fraud and telephone fraud detection
    o Based on past purchases/phone calls is a new one fraudulent?
      • Saves companies billions(!) of dollars annually
  ▪ Spam E-mail detection (gmail, hotmail, many others)
“I believe over the next decade computing will become even more ubiquitous and intelligence will become ambient. The coevolution of software and new hardware form factors will intermediate and digitize — many of the things we do and experience in business, life and our world. This will be made possible by an ever-growing network of connected devices, incredible computing capacity from the cloud, insights from big data, and intelligence from machine learning.”

-- Satya Nadella (CEO, Microsoft)

Email to all employees on first day as CEO (Feb. 04, 2014)
Data Analytics at Amazon

P(Buy Book Y | Bought Book X)
...Could Still Use Some Work
Really, Work On It…

An Introduction to Differential Equations and Their Applications
by Stanley J. Farlow

List Price: $45.95
Price: $27.33
You Save: $18.62 (41%)

Intended for use in a beginning one-semester course in differential equations, this text is designed for students of pure and ... Read More

Learn more Add to Wish List

Captain Underpants and the Sensational Saga of Sir Stinks-A-Lot
by Dav Pilkey

List Price: $9.99
Price: $6.49
You Save: $3.50 (35%)

There's something rotten in the state of Ohio, and it's smellier than a pile of putrid gym socks! Steer clear — ... Read More

Learn more Add to Wish List
From Data To Understanding

• In data modeling, maintain critical perspective
  ▪ Making predictions is only part of the story
  ▪ Also try to get some understanding of the domain

• Spurious correlations
Data Mining Process

Data Mining: The process of extracting useful and actionable information from data in order to make more informed decisions.
Learning From Data

• Given a set of data, we want to glean insight to then take action
  ▪ Build a model of domain → understanding
  ▪ Make a prediction → target for action
  ▪ Take action → validate model efficacy

• We’ll start with the problem of determining a probability distribution

• Later, we’ll consider building predictive (classification) models