Today

- Logistics
- ML Overview
Logistics

- **Instructor:** Doug Downey
  - Email: ddowney@eecs.northwestern.edu
  - Office hours: Mondays 4:00-5:00 (or by appt), Mudd 3111
- **TAs:** Dave Demeter, Zheng Yuan, 4 ugrad peer mentors
- **Web:** (linked from prof. homepage) [http://www.cs.northwestern.edu/~downey/courses/349_Fall2018/](http://www.cs.northwestern.edu/~downey/courses/349_Fall2018/)
  - Also: Canvas, Piazza
### Grading and Assignments (1 of 2)

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework 1</td>
<td>15</td>
</tr>
<tr>
<td>Homework 2</td>
<td>5</td>
</tr>
<tr>
<td>Project Proposal</td>
<td>5+5 (peer mentoring)</td>
</tr>
<tr>
<td>Exam 1</td>
<td>10</td>
</tr>
<tr>
<td>Homework 3</td>
<td>10</td>
</tr>
<tr>
<td>Project Status Report</td>
<td>5+5 (peer mentoring)</td>
</tr>
<tr>
<td>Homework 4</td>
<td>10</td>
</tr>
<tr>
<td>Exam 2</td>
<td>10</td>
</tr>
<tr>
<td>Project Website</td>
<td>20</td>
</tr>
</tbody>
</table>

**Total**: 100

<table>
<thead>
<tr>
<th>Grade</th>
<th>Range</th>
<th>Grade</th>
<th>Range</th>
<th>Grade</th>
<th>Range</th>
<th>Grade</th>
<th>Range</th>
<th>Grade</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93+</td>
<td>A-</td>
<td>92-90</td>
<td>B+</td>
<td>89-87</td>
<td>B</td>
<td>86-83</td>
<td>B-</td>
<td>82-80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C+</td>
<td>79-77</td>
<td>C</td>
<td>76-73</td>
<td>C-</td>
<td>72-70</td>
<td>Etc...</td>
<td>69...</td>
</tr>
</tbody>
</table>
Grading and Assignments (2 of 2)

- Four homeworks (40 pts)
  - Submitted via Canvas according to hmwk instructions
    - Late penalty 10% per day – must be within 1 week of original deadline
  - Significant programming, some exercises
    - Programming assignments in groups of two (or one)
- Exams (20 pts)
  - Monday of Week 4, Friday of Week 9
- Project (30 pts + 10 peer review)
  - Teams of \( k \)
    - Define a task, create/acquire data for the task, train ML algorithm(s), evaluate & report
Expectations

- Grades

- Academic Integrity
  - You are expected to do your own work
    - More details in syllabus linked from course home page
  - Suspected violations of integrity policy will be referred to the administration

- Slides may not make sense if you don’t come to class
Prerequisites

- **Significant Programming Experience**
  - EECS 214, 325 or the equivalent
  - Example: implement decision trees (covered starting Monday)

- **Python** is the language we’ll use
  - You’ll have skeleton code to help you
  - (also, I barely know Python)

- **Basics of probability**
  - E.g. independence
Source Materials

- Papers & Web pages
- Reading for next week:
  - Required:
    - Decision trees (see the Decision Tree notes when they’re posted)
  - Optional:
    - Gartner 2016 Hype Cycle
      (machine learning at peak hype as of 2016)
Participation

- Think/Pair/Share (next)
- Peer Review
Today

- Logistics
- ML Overview
Think/Pair/Share

Why study Machine Learning?

Think

Start

End
Why study Machine Learning?
Think/Pair/Share

Why study Machine Learning?

Pair
Start
End
Why study Machine Learning?

Share
What is Machine Learning?

- “The study of computer programs that improve automatically with experience”
  T. Mitchell *Machine Learning*

- Automating automation
- Getting computers to program themselves
- Writing software is the bottleneck
- Let the data do the work instead!
No, more like gardening

- **Seeds** = Algorithms
- **Nutrients** = Data
- **Gardener** = You
- **Plants** = Programs
Example: Farecast

Search Flights  Find cheap flights and free airfare predictions

- Round Trip  - One Way  - Multi-City

Please enter a To city

From:
Chicago, IL (CHI) - All airports

Include Nearby Airports

To:
Seattle, WA (SEA) - Seattle/Tacoma

Include Nearby

7-Day Low Fare Prediction

Tip: Buy

Fares Rising $42
Confidence: 66%

Details

Applies to ORD>SEA only

Daily Low Fare History

$390
$305
$220
$135

69 Days Ago
Now
Sample Applications

Google
Twitter
WAYMO
Zillow
facebook
Amazon
IBM Watson

Input → Computer → Program
Output
Relationship of Machine Learning to...

- Statistics
- Analytics / Data Science
- Artificial Intelligence
Why study Machine Learning? (1 of 4)

- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates)
- “Machine learning is the next Internet” (Tony Tether, former Director, DARPA)

These quotes are ~10 years old (e.g. Gates is from the NYT, 2004)

More recent:
“Artificial intelligence is one of the great opportunities for improving the world today,” (Reid Hoffman, co-founder of $1B deep learning research center)
Why study Machine Learning? (2 of 4)

The Digital Universe is Huge — And Growing Exponentially

If the Digital Universe were represented by the memory in a stack of tablets, in 2013 it would have stretched two-thirds the way to the Moon*

By 2020, there would be 6.6 stacks from the Earth to the Moon*

Source: IDC, 2014

* iPad Air — 0.29" thick, 128 GB

Why study Machine Learning? (3 of 4)

http://www.gartner.com/newsroom/id/3598917
The Content Created World 2015 – 2025

- Classically Created
- IoT Relevant
- IoT Actionable

Growth From 2020 - 2025
- 9.6x
- 4.8x
- 2.3x

Source: The Internet of Things: Getting Ready to Embrace Its Impact on the Digital Economy (IDC #DR2016_GS4_VT)

© IDC - Visit us at IDC.com and follow us on Twitter: @IDC
Hospitals possessing (Certified EHR) or adopting (Basic EHR) electronic medical records. What will we be able to learn from these?
What You’ll Learn in this Class

- How do ML algorithms work?
  - Learn by implementing, using
- When should I use ML?
- For a real problem, how do I:
  - Express my problem as an ML task
  - Choose the right ML algorithm
  - Evaluate the results
ML in a Nutshell

- Tens of thousands of machine learning algorithms
- Hundreds new every year
- Every machine learning algorithm has three components:
  - Representation
  - Evaluation
  - Optimization
Representation

How do we represent the function from input to output?
- Decision trees
- Sets of rules / Logic programs
- Instances
- Graphical models (Bayes/Markov nets)
- Neural networks
- Support vector machines
- Model ensembles
- Etc.
Evaluation

- Given some data, how can we tell if a function is “good”?
  - Accuracy
  - Precision and recall
  - Squared error
  - Likelihood
  - Posterior probability
  - Cost / Utility
  - Margin
  - Entropy
  - K-L divergence
  - Etc.
Given some data, how do we find the “best” function?

- Combinatorial optimization
  - E.g.: Greedy search
- Convex optimization
  - E.g.: Gradient descent
- Constrained optimization
  - E.g.: Linear programming
Inductive Learning

- **Given** examples of a function \((x, f(x))\)
- **Predict** function \(f(x)\) for new instances \(x\)
  - Discrete \(f(x)\): Classification
  - Continuous \(f(x)\): Regression
  - \(f(x) = \text{Probability}(x)\): Probability estimation

- **Example:**
  - \(x = <\text{Flight}=\text{United 102}, \ \text{FlightDate}=\text{May 26}, \ \text{Today}=\text{May 7}>\)
  - \(f(x) = +1\) if flight price will increase in the next week, or \(-1\) otherwise
What We’ll Cover

- **Inductive learning**
  - Decision tree induction
  - Instance-based learning
  - Neural networks
  - Bayesian Learning
  - Logistic Regression
  - Support vector machines
  - Learning theory
  - Reinforcement Learning

- **Unsupervised learning**
  - Clustering
  - Dimensionality reduction
Today

- **Logistics**
  - 4 homeworks, 2 exams, course project. No final.
  - Take a look at the course Web page for more.

- **ML Overview**
  - Like gardening
  - data = rich source of fuel for ML
  - More soon…