Logistics

- **Instructor:** Doug Downey
  - Email: ddowney@eecs.northwestern.edu
  - Office hours: Mondays 4:00-5:00 (or by appt), Ford 3-345

- **TAs:** Dave Demeter, Zheng Yuan, Xutong Chen, 7 ugrad peer mentors

- **Web:** (linked from prof. homepage)
  - [http://www.cs.northwestern.edu/~downey/courses/349_Spring_2018/](http://www.cs.northwestern.edu/~downey/courses/349_Spring_2018/)
  - 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>
<tr>
<td></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

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

- Four homeworks (40 pts)
  - Submitted via e-mail according to hmwk instructions
    - Late penalty 10% per day – must be within 1 week of original deadline
  - Significant programming, some exercises

- Exams (20 pts)
  - Fridays of Week 4, 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
Prerequisites

- Significant Programming Experience
  - EECS 214, 325 or the equivalent
  - Example: implement decision trees (covered starting tomorrow)
- **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 this week:
  - Required:
    - Decision trees (see the Decision Tree notes)
  - Optional:
    - Gartner 2016 Hype Cycle
      (machine learning at peak hype as of 2016)
Participation

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

Why study Machine Learning?

Think

Start

End
Think/Pair/Share

Why study Machine Learning?

<table>
<thead>
<tr>
<th>Think</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>End</th>
</tr>
</thead>
</table>
Think/Pair/Share

Why study Machine Learning?

Start | Pair | End
Think/Pair/Share

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!
Traditional Programming

Input → Computer → Output
Program →

Machine Learning

Input → Computer → Program
Output →
Magic?

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

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

7-Day Low Fare Prediction
Tip: Buy
Fares Rising $42
Confidence: 66%

Details

Applies to ORD>SEA only

Daily Low Fare History

<table>
<thead>
<tr>
<th>$390</th>
<th>$305</th>
<th>$220</th>
<th>$135</th>
</tr>
</thead>
<tbody>
<tr>
<td>69 Days Ago</td>
<td>Now</td>
<td>Now</td>
<td>Now</td>
</tr>
</tbody>
</table>
Sample Applications

Input → Computer → Program

Output

Google, Twitter, Waymo, Zillow, Facebook, Amazon, IBM Watson
Relationship of Machine Learning to...

- Statistics
- Analytics / Data Science
- Artificial Intelligence
Why study Machine Learning? (1 of 5)

- “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 5)

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 5)

Hospitals possessing (Certified EHR) or adopting (Basic EHR) electronic medical records. What will we be able to learn from these?

Gartner: 8.4B connected “things” in 2017 (up 31% from 2016)...20B in 2020
What You’ll Learn in this Class

- How do ML algorithms work?
  - Learn by implementing, using

- For a real problem, how do I:
  - Express my problem as an ML task
  - Choose the right ML algorithm
  - Evaluate the results
Tens of thousands of machine learning algorithms
Hundreds new every year
Every machine learning algorithm has three components:
  - Representation
  - Evaluation
  - Optimization
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:**
  - \(x = \langle \text{Flight}=\text{United 102}, \, \text{FlightDate}=\text{May 26}, \, \text{Today}=\text{May 7} \rangle\)
  - \(f(x) = +1\) 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