

EECS 349

Machine Learning

Instructor: Doug Downey

(some slides from Pedro Domingos, University of Washington)

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, Chen Liang, Zheng Yuan, 5 ugrad peer mentors
- ▶ **Web:** (linked from prof. homepage)
http://www.cs.northwestern.edu/~downey/courses/349_Spring2017/
 - ▶ Also, Canvas

Grading and Assignments (1 of 2)

Assignment	Due Date (tentative)	Points
Homework 1	3-Apr-17	10
Homework 2	14-Apr-17	10
Project Proposal	19-Apr-17	5+5 (peer mentoring)
Exam 1	21-Apr-17	10
Homework 3	12-May-17	10
Project Status Report	17-May-17	5+5 (peer mentoring)
Homework 4	22-May-17	10
Exam 2	26-May-17	10
Project Website	9-Jun-17	20
	TOTAL POINTS	100

A	A-	B+	B	B-	C+	C	C-	Etc...
93+	92-90	89-87	86-83	82-80	79-77	76-73	72-70	69...

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

Prerequisites

- ▶ Significant Programming Experience
 - ▶ EECS 214, 325 or the equivalent
 - ▶ Example: implement decision trees (covered starting Wednesday)
 - ▶ **Python** is the language we'll use
 - ▶ You'll have skeleton code to help you through
 - ▶ (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)

Participation

- ▶ A demonstration call and response. I say Machine, you say Learning.
 - ▶ Machine! Learning! Machine! Learning!
- ▶ Think/Pair/Share (next slide)
- ▶ Peer Review

Think/Pair/Share

Why study Machine Learning?

| Think
Start

|
End

Think/Pair/Share

Why study Machine Learning?

| Think
Start

|
End

Think/Pair/Share

Why study Machine Learning?

| Pair
Start

|
End

Think/Pair/Share

Why study Machine Learning?

Share

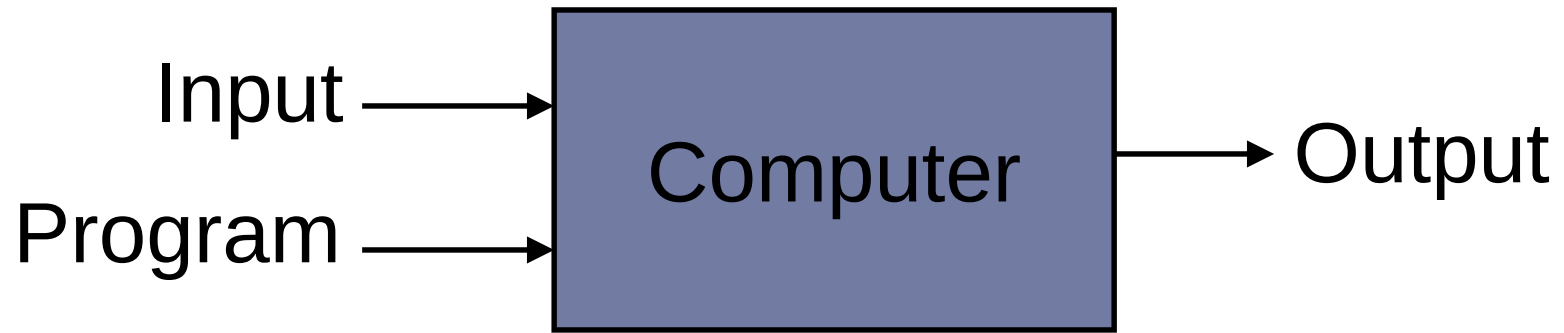
What is Machine Learning?

- ▶ “The study of computer programs that improve automatically with experience”

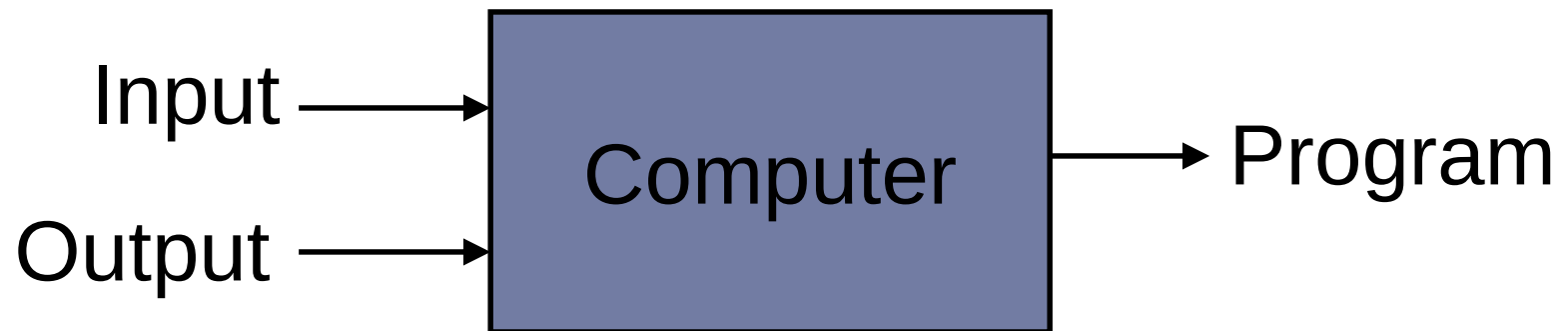
T. Mitchell *Machine Learning*

- ▶ Automating automation
- ▶ Getting computers to program themselves
- ▶ Writing software is the bottleneck
- ▶ Let the data do the work instead!

Traditional Programming



Machine Learning



Magic?

No, more like gardening

- ▶ **Seeds** = Algorithms
- ▶ **Nutrients** = Data
- ▶ **Gardener** = You
- ▶ **Plants** = Programs



Case Study: 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



Sample Applications

Google

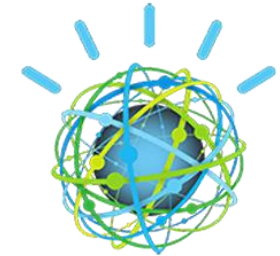


WAYMO

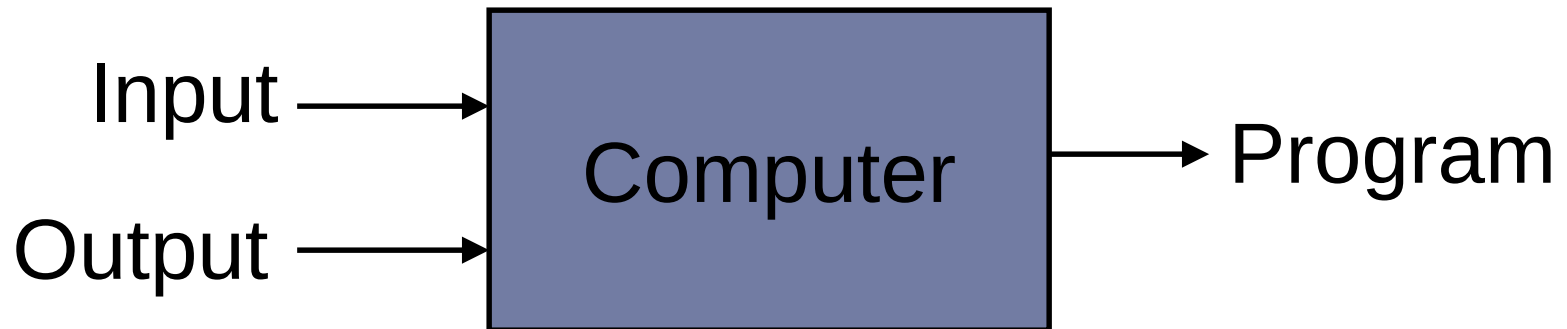
Zillow

facebook

amazon



IBM Watson



Relationship of Machine Learning to...

- ▶ **Statistics**
- ▶ **Analytics / Data Science**
- ▶ **Data Mining**
- ▶ **Artificial Intelligence**

Why study Machine Learning? (1 of 5)

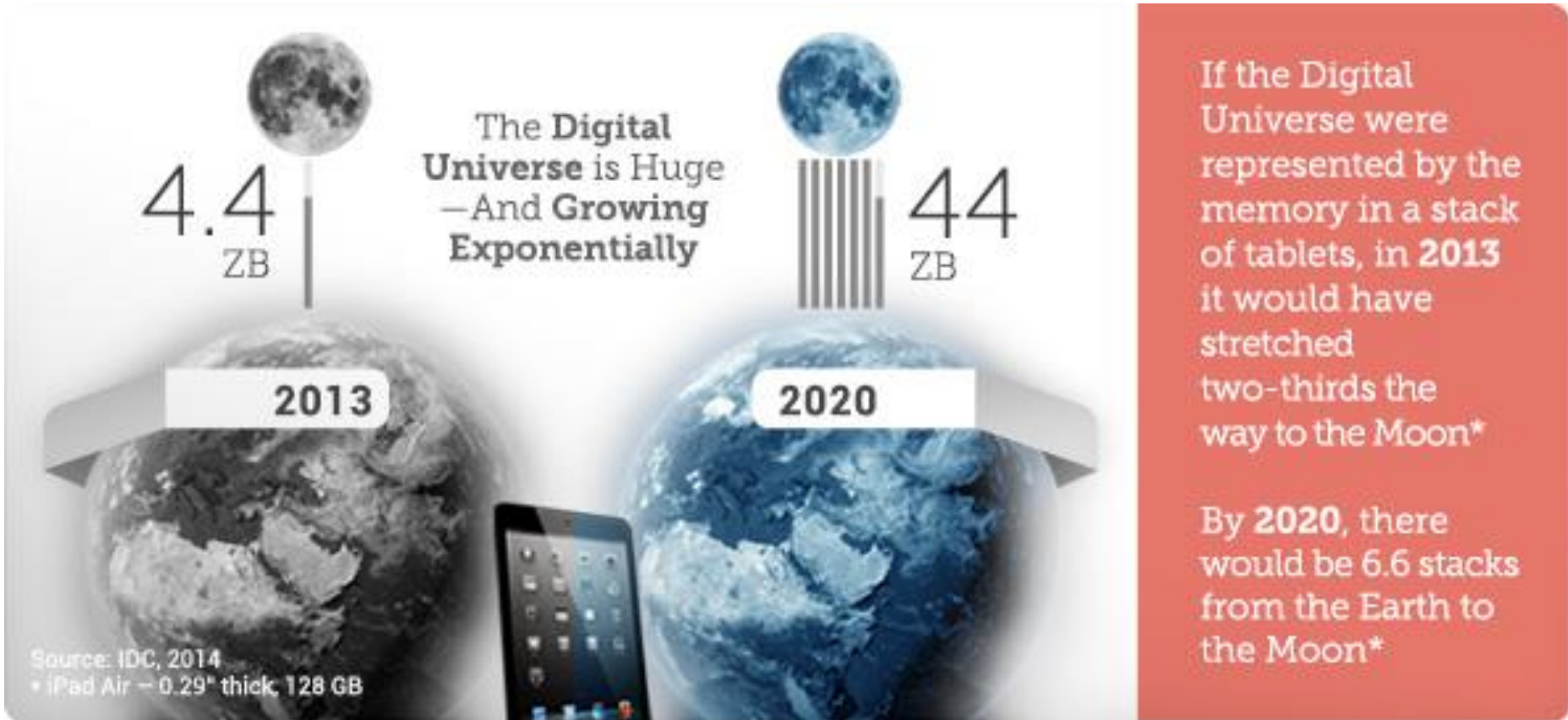
- ▶ “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates)
- ▶ “Machine learning is the next Internet” (Tony Tether, former Director, DARPA)

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)

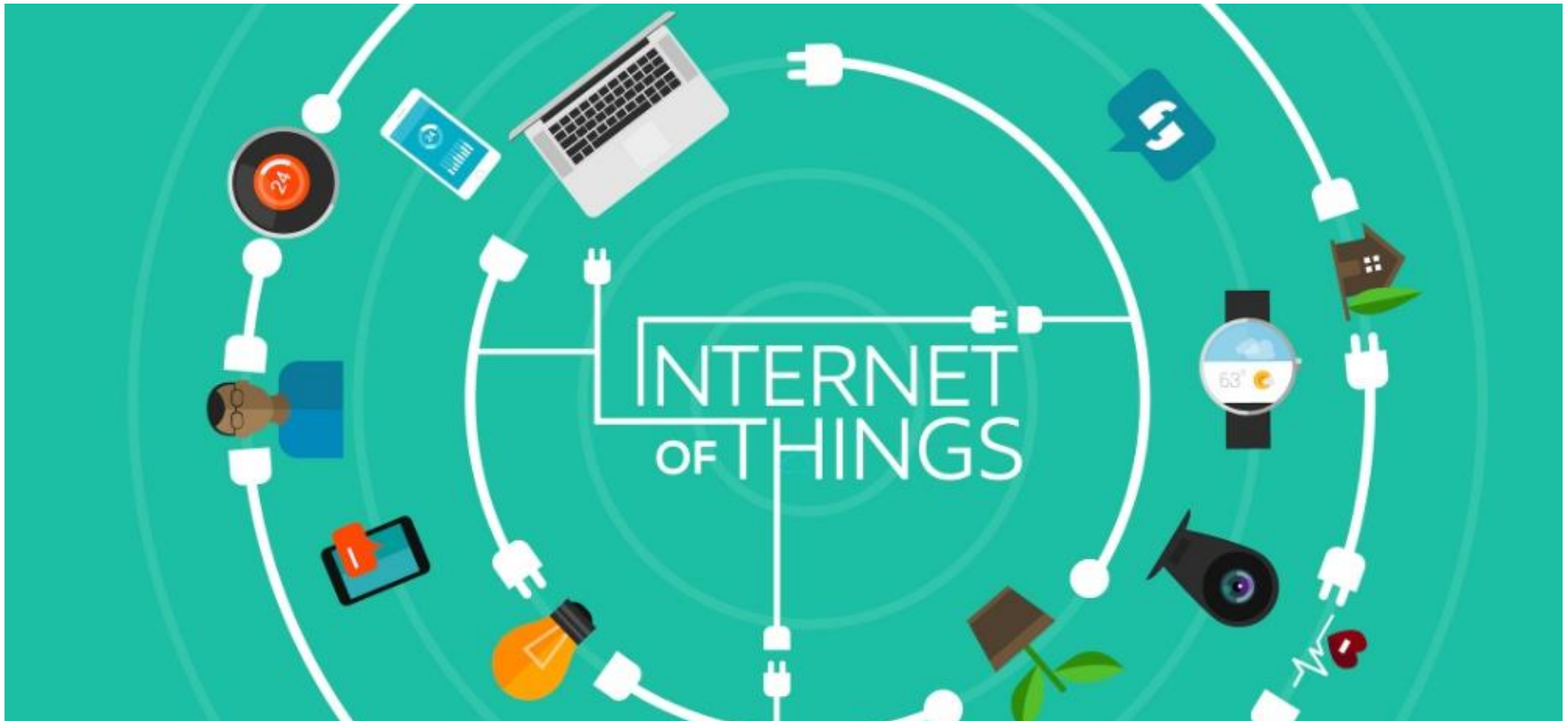


<http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>

Why study Machine Learning? (3 of 5)

- ▶ **One example, proportion of physicians using EMRs**
 - ▶ **2001: 18%**
 - ▶ **2011: 57%**
 - ▶ **2013: 78%**
 - ▶ **2015: 83%**
- ▶ **...what will be able to learn from these?**

Why study Machine Learning? (4 of 5)



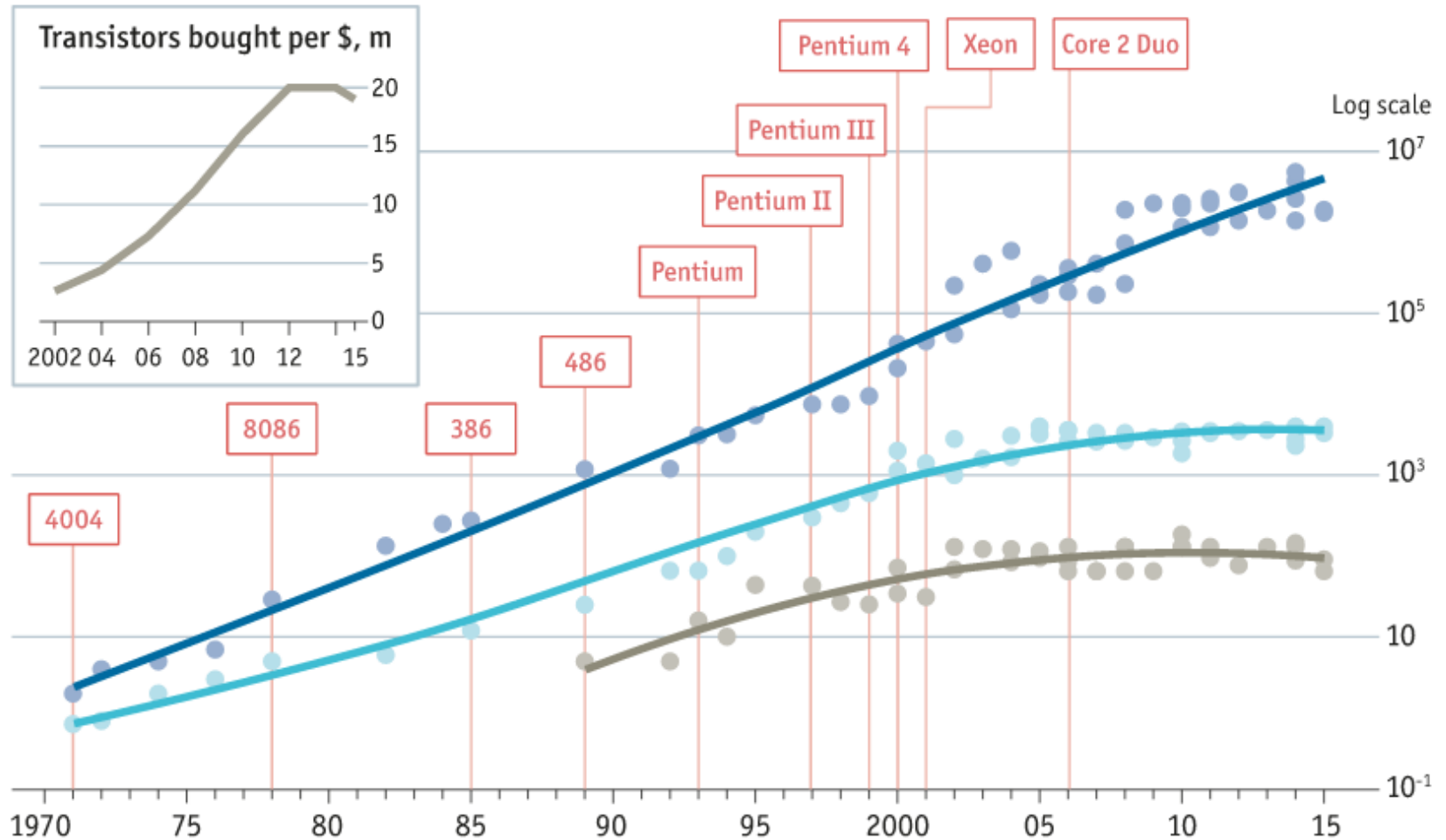
Gartner: 8.4B connected “things” in 2017 (up 31% from 2016)...**20B** in 2020

<http://www.gartner.com/newsroom/id/3598917>

Stuttering

● Transistors per chip, '000 ● Clock speed (max), MHz ● Thermal design power*, w

Chip introduction dates, selected



Sources: Intel; press reports; Bob Colwell; Linley Group; IB Consulting; *The Economist*

*Maximum safe power consumption

ML in Practice

- ▶ Understanding domain, prior knowledge, and goals
- ▶ Data integration, selection, cleaning, pre-processing, etc.
- ▶ Learning models
- ▶ Interpreting results
- ▶ Consolidating and deploying discovered knowledge
- ▶ Loop

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

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.

Optimization

- ▶ *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

Types of Learning

- ▶ **Supervised (inductive) learning**
 - ▶ Training data includes desired outputs
- ▶ **Unsupervised learning**
 - ▶ Training data does not include desired outputs
- ▶ **Semi-supervised learning**
 - ▶ Training data includes a few desired outputs
- ▶ **Reinforcement learning**
 - ▶ Rewards from sequence of actions

Inductive Learning

- ▶ **Given** examples of a function $(\mathbf{x}, f(\mathbf{x}))$
- ▶ **Predict** function $f(\mathbf{x})$ for new instances $\mathbf{x}$
 - ▶ Discrete $f(\mathbf{x})$: Classification
 - ▶ Continuous $f(\mathbf{x})$: Regression
 - ▶ $f(\mathbf{x}) = \text{Probability}(\mathbf{x})$: Probability estimation
- ▶ **Example:**
 - ▶ $\mathbf{x} = \langle \text{Flight}=\text{United 102}, \text{FlightDate}=\text{May 26}, \text{Today}=\text{May 7} \rangle$
 - ▶ $f(\mathbf{x}) =$ +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