# CS 446: Machine Learning

**Dan Roth**

**University of Illinois, Urbana-Champaign**

**danr@illinois.edu**

**http://L2R.cs.uiuc.edu/~danr**

**3322 SC**

# CS446: Machine Learning

Tuesday, Thursday: 17:00pm-18:15pm 1404 SC

**Registration to Class**

Office hours: Mon 3:00-4:00 pm [my office]

TAs: Chase Duncan; Qiang Ning, Subhro Roy, Hao Wu

Assignments: 7 Problems sets (Programming)

Weekly (light) on-line quizzes

Discussion sections

Mid Term Exam

Project

Final

Mitchell/Other Books/ **Lecture notes** /Literature

# CS446 Machine Learning: Today

- What is Learning?
- Who are you?
- What is CS446 about?

# What is Learning

- The Badges Game……
- Who are you?

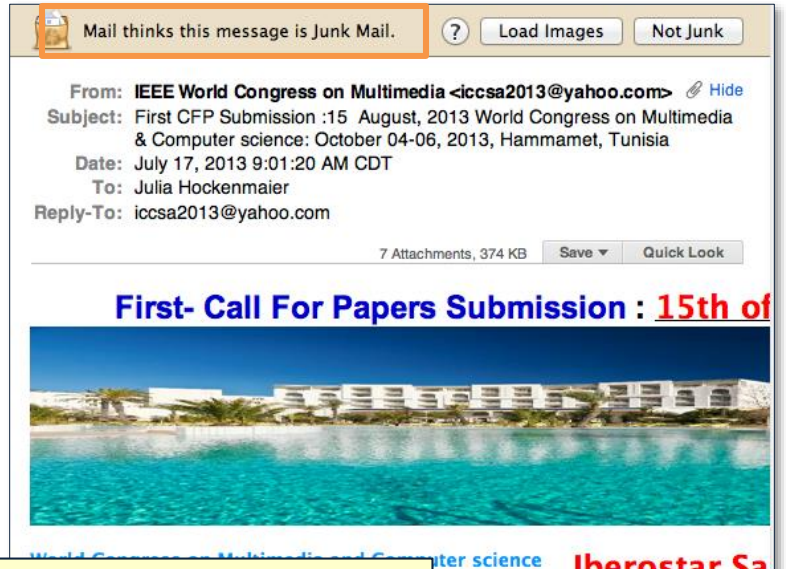
# An Owed to the Spelling Checker

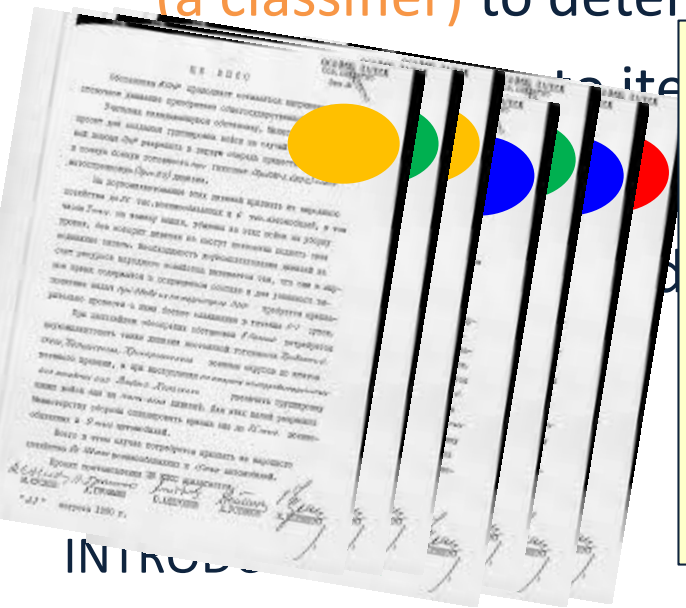
I have a spelling checker,  it came with my PC
It plane lee marks four my revue
Miss steaks aye can knot sea.
Eye ran this poem threw it, your sure reel glad two no.
Its vary polished in it's weigh
My checker tolled me sew.
A checker is a bless sing, it freeze yew lodes of thyme.
It helps me  right awl stiles two reed
And aides me when aye rime.
Each frays come posed up on my screen
Eye trussed to bee a joule...

# Machine learning is everywhere

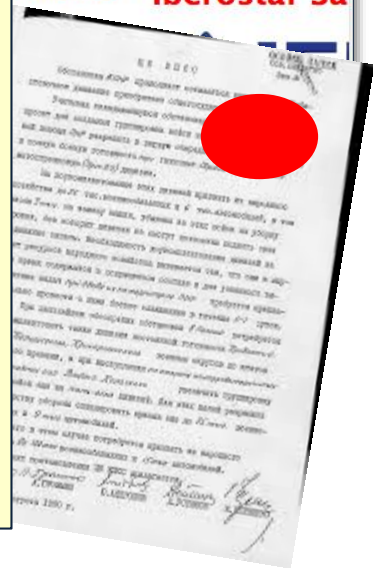
# Applications: Spam Detection

- This is a binary classification task: Assign one of two labels (i.e. yes/no) to the input (here, an email message)
- Classification requires a model (a classifier) to determine which



| Documents | Labels |
|---|---|
| Documents | Politics, Sports, Finance |
| Sentences | Positive, Negative |
| Phrases | Person, Location |
| Images | cats, dogs, snakes |
| Medical records | Admit again soon/Not |
| ..... | |



# Ambiguity Resolution

Can I have a peace of cake ? piece

...Nissan Car and truck plant is …

…divide life into plant and animal kingdom

Buy a car with a steering wheel (his money)

(This Art) (can N) (will MD) (rust V) V,N,N

The dog bit the kid. He was taken to a veterinarian hospital

Learn a function that maps observations in the domain to one of several categories or $\Re$.

# Comprehension

(ENGLAND, June, 1989) – Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don't know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.
2. Winnie the Pooh is a title of a book.
3. Christopher Robin's dad was a magician.
4. Christopher Robin must be at least 65 now.

**This is an Inference Problem; where is the learning?**

# Learning

- **Learning is at the core of**
  - Understanding High Level Cognition
  - Performing knowledge intensive inferences
  - Building adaptive, intelligent systems
  - Dealing with messy, real world data
  - Analytics
- **Learning has multiple purposes**
  - Knowledge Acquisition
  - Integration of various knowledge sources to ensure robust behavior
  - Adaptation (human, systems)
  - Decision Making (Predictions)

# Learning = Generalization

**H. Simon -**

"Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the task or tasks drawn from the same population more efficiently and more effectively the next time."

The ability to perform a task in a situation which has never been encountered before

# Learning = Generalization


Mail thinks this message is junk mail.

Not junk

- The learner has to be able to classify items it has never seen before.

# Learning = Generalization

The ability to perform a task in a situation which has never been encountered before

- Classification
  - Medical diagnosis; credit card applications; hand-written letters; ad selection; sentiment assignment,…
- Planning and acting
  - Navigation; game playing (chess, backgammon, go); driving a car
- Skills
  - Balancing a pole; playing tennis
- Common sense reasoning
  - Natural language interactions

Generalization depends on the Representation as much as it depends on the Algorithm used.

# Why Study Learning?

- Computer systems with new capabilities.
  - Develop systems that are too difficult or impossible to construct manually .
  - Develop systems that can automatically adapt and customize themselves to the needs of the individual user through experience.
  - Discover knowledge and patterns in databases, e.g. discovering purchasing patterns for marketing purposes.
  - Solve the kinds of problems now reserved for humans.

# Why Study Learning?

- Computer systems with new capabilities.
- Understand human and biological learning
- Understanding teaching better.

# Why Study Learning?

- Computer systems with new capabilities.
- Understand human and biological learning
- Understanding teaching better.
- Time is right.
  - Initial algorithms and theory in place.
  - Growing amounts of on-line data
  - Computational power available.
  - Necessity: many things we want to do cannot be done by “programming”.

# Learning is the future

- Learning techniques will be a basis for every application that involves a connection to the messy real world
- Basic learning algorithms are ready for use in applications today
- Prospects for broader future applications make for exciting fundamental research and development opportunities
- Many unresolved issues – Theory and Systems
  - While it's hot, there are many things we don't know how to do

# Work in Machine Learning

- **Artificial Intelligence; Theory; Experimental CS**
- **Makes Use of:**
  - Probability and Statistics; Linear Algebra; Theory of Computation;
- **Related to:**
  - Philosophy, Psychology (cognitive, developmental), Neurobiology, Linguistics, Vision, Robotics,….
- **Has applications in:**
  - AI (Natural Language; Vision; Planning; HCI)

- ❑ Very active field
- ❑ What to teach?
  - ❑ The fundamental paradigms
  - ❑ Some of the most important algorithmic ideas
  - ❑ Modeling

And: what we don't know

# Course Overview

- Introduction: Basic problems and questions
- A detailed example: Linear threshold units; key algorithmic idea
  - Online Learning
- Two Basic Paradigms:
  - PAC (Risk Minimization)
  - Bayesian theory
- Learning Protocols:
  - Supervised; Unsupervised; Semi-supervised
- Algorithms
  - Gradient Descent
  - Decision Trees (C4.5)
  - [Rules and ILP (Ripper, Foil)]
  - Linear Threshold Units (Winnow; Perceptron; Boosting; SVMs; Kernels)
  - Neural Networks (Backpropagation)
  - Probabilistic Representations (naïve Bayes; Bayesian trees; Densities)
  - Unsupervised /Semi supervised: EM
- Clustering; Dimensionality Reduction

Who knows DTs ?

Who knows NNs ?

# CS446: Machine Learning

Tuesday, Thursday: 17:00pm-18:15pm 1404 SC

**Registration to Class**

Office hours: Mon 3:00-4:00 pm [my office]

TAs: Chase Duncan; Qiang Ning, Subhro Roy, Hao Wu

Assignments: 7 Problems sets (Programming)

Weekly (light) on-line quizzes

Discussion sections

**Send me email after class**

**Title: CS446 LastName, First Name, net id, Registration**

**Body: Have you sent me email already (when)? Any other information**

# CS446: Machine Learning

**Participate, Ask Questions**

- What do you need to know:
  - Theory of Computation
  - Probability Theory
  - Linear Algebra
  - Programming (Java; your favorite language; some Matlab)
- Homework 0 – on the web
- Who is the class for?
  - Future Machine Learning researchers/Advanced users

# CS446: Policies

- Cheating

  No.

  We take it very seriously.

**Info page**

**Note also the Schedule Page and our Notes**

- Homework:
  - Collaboration is encouraged
  - But, you have to write your own solution/program.
  - (Please don't use old solutions)
- Late Policy:

  You have a credit of 4 days (4*24hours); That's it.
- Grading:
  - Possibly separate for grads/undergrads.
  - 5% Quizzes; 25% - homework; 30%-midterm; 40%-final;
  - Projects: 25% (4 hours)
- Questions?

# CS446 Team

- **Dan Roth** (3323 Siebel)
  - Tuesday/Thursday, 1:45 PM – 2:30 PM (or: appointment)
- **TAs**
  - Chase Duncan Tues 12-1 (3333 SC)
  - Subhro Roy Wed 4-5 (3333 SC)
  - Qiang Ning Thur 3-4 (3333 SC)
  - Hao Wu Fri 1-2 (3333 SC)
- **Discussion Sections**: (starting 3rd week)
  - **Tuesday:** 11 -12 [3405 SC] Subhro Roy [A-I]
  - **Wednesdays:** 5 -6 [3405 SC] Hao Wu [J-L]
  - **Thursdays:** 2 - 3 [3405 SC] Chase Duncan [M-S]
  - **Fridays:** 4 -5 [3405 SC] Qiang Ning [T-Z]

# CS446 on the web

- Check our class website:
  - Schedule, slides, videos, policies
    - http://l2r.cs.uiuc.edu/~danr/Teaching/CS446-17/index.html
  - Sign up, participate in our Piazza forum:
    - Announcements and discussions
    - https://piazza.com/class#fall2017/cs446
  - Log on to Compass:
    - Submit assignments, get your grades
    - https://compass2g.illinois.edu
- Scribing the Class [Good writers; Latex; Paid Hourly]

# What is Learning

- The Badges Game……
  - This is an example of the key learning protocol: supervised learning
- First question: Are you sure you got it?
  - Why?
- Issues:
  - Prediction or Modeling?
  - Representation
  - Problem setting
  - Background Knowledge
  - When did learning take place?
  - Algorithm