L2: Overview of machine learning applications, types, and tasks

Prof. Xun Jiao

Before class

- No HWs this week
- Syllabus/Slides on BB

Review of last class

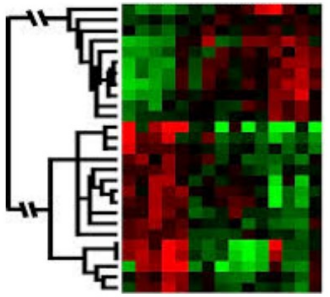
- What is machine learning?
  - Method that can learn from ???

WHEN TO USE MACHINE LEARNING?

ML is used when:

- Human expertise does not exist (navigating on Mars)
- Humans can't explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)



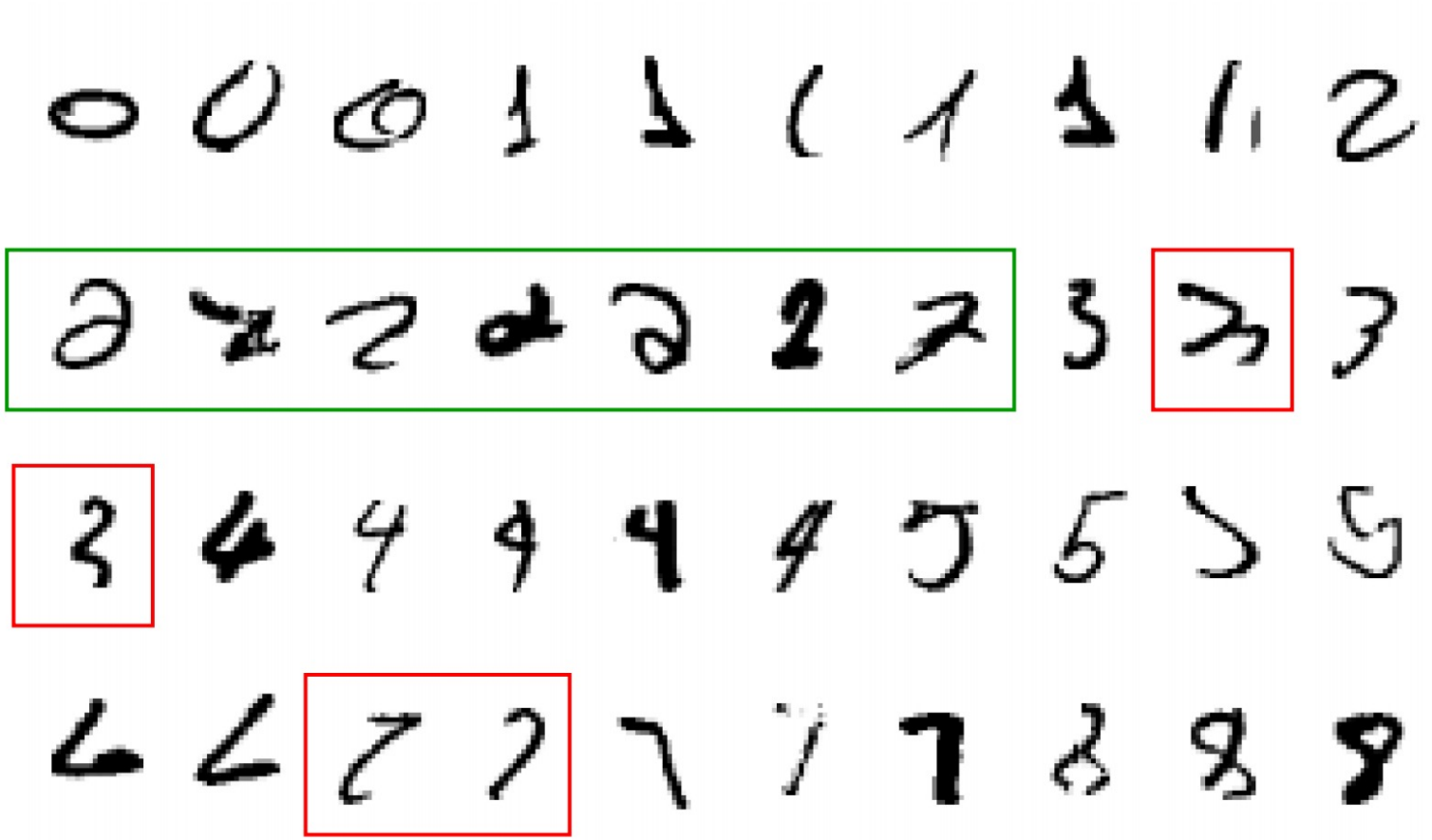



Learning isn't always useful:

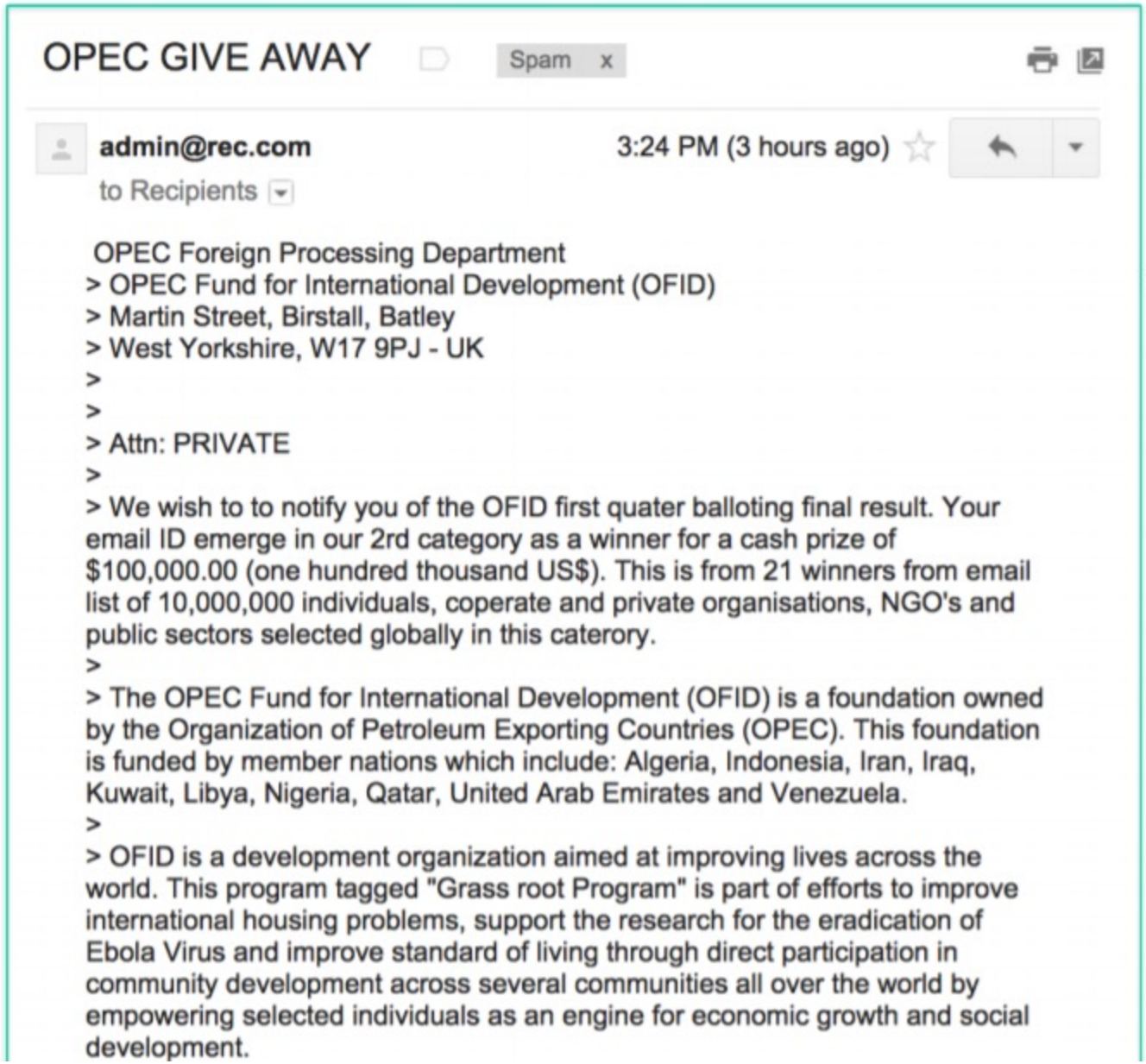
- There is no need to "learn" to calculate payroll

A classic example of a task that requires machine learning:

It is very hard to say what makes a 2



Learning to classify text documents



OPEC GIVE AWAY

Spam x

admin@rec.com 3:24 PM (3 hours ago)

to Recipients

OPEC Foreign Processing Department
> OPEC Fund for International Development (OFID)
> Martin Street, Birstall, Batley
> West Yorkshire, W17 9PJ - UK
>
>
> Attn: PRIVATE
>
> We wish to to notify you of the OFID first quater balloting final result. Your email ID emerge in our 2rd category as a winner for a cash prize of $100,000.00 (one hundred thousand US$). This is from 21 winners from email list of 10,000,000 individuals, coperate and private organisations, NGO's and public sectors selected globally in this caterory.
>
> The OPEC Fund for International Development (OFID) is a foundation owned by the Organization of Petroleum Exporting Countries (OPEC). This foundation is funded by member nations which include: Algeria, Indonesia, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, United Arab Emirates and Venezuela.
>
> OFID is a development organization aimed at improving lives across the world. This program tagged "Grass root Program" is part of efforts to improve international housing problems, support the research for the eradication of Ebola Virus and improve standard of living through direct participation in community development across several communities all over the world by empowering selected individuals as an engine for economic growth and social development.

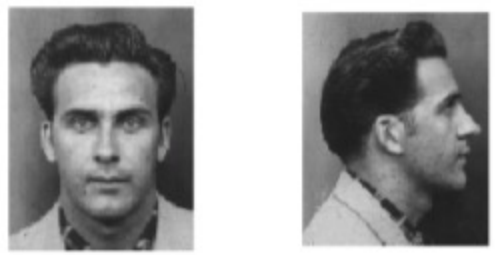
→

spam

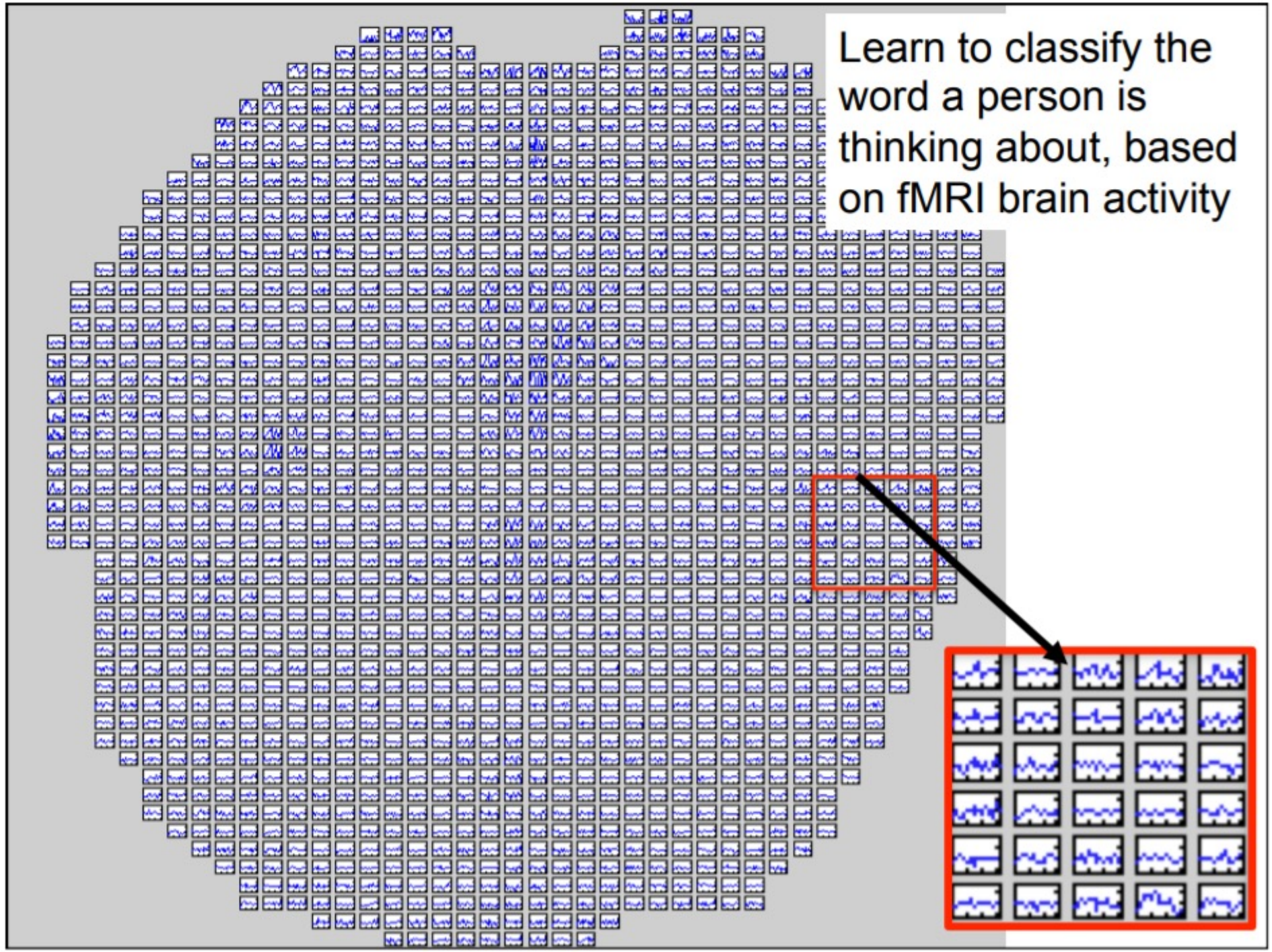
vs

not spam

Learning to detect objects in images



Example training images
for each orientation


Learn to classify the word a person is thinking about, based on fMRI brain activity

Some more examples

Recognizing patterns:

- Facial identities or facial expressions (Iphone)
- Handwritten or spoken words
- Medical images (IBM Watson system)

Recognizing anomalies:

- Unusual credit card transactions
- Unusual patterns of sensor readings in a nuclear power plant

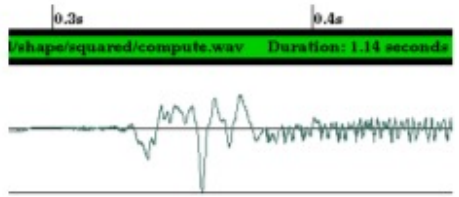
Prediction:

- Future stock prices or currency exchange rates

Autonomous Driving, Robotics, Gaming AI …

Machine Learning - Practice

Mining Databases



Speech Recognition

Object recognition

Control learning

Text analysis

Peter H. van Oppen , Chairman of the Board & Chief Executive Officer

Mr. van Oppen has served as chairman of the board and chief executive officer of ADIC since its acquisition by Interpoint in 1994 and a director of ADIC since 1986. Until its acquisition by Crane Co. in October 1996, Mr. van Oppen served as chairman of the board of directors, president and chief executive officer of Interpoint. Prior to 1985, Mr. van Oppen worked as a consulting manager at Price Waterhouse LLP and at Bain & Company in Boston and London. He has additional experience in medical electronics and venture capital. Mr. van Oppen also serves as a director of Despatch Industries, Inc. and Spacelabs Medical, Inc.. He holds a B.A. from Whitman College and an M.B.A. from Harvard Business School, where he was a Baker Scholar.

- Support Vector Machines
- Bayesian networks
- Hidden Markov models
- Deep neural networks
- Reinforcement learning
- ....

Why is this called “Applied” Machine Learning?

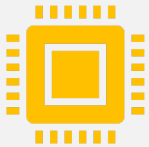
More emphasis on using existing tools than implementing algorithms

But you’ll do a little bit of implementation too

Less mathematical theory

But you’ll still learn how the algorithms work

Math will be taught as needed



More focus on creating systems/pipelines (data processing, design, evaluation)

High-level Goals

identify when machine learning can help solve a problem and which approaches are appropriate;

be comfortable doing machine learning in Python, and be familiar enough with the algorithms and parameters to easily adopt other toolkits;

understand the underlying concepts well enough that you can read machine learning papers, and can modify implementations for your own needs;

Develop an ML pipeline system for selected applications

Concrete goals

- The primary Machine Learning algorithms
  - Logistic regression, Bayesian methods, HMM's, SVM's, reinforcement learning, decision tree learning, boosting, unsupervised clustering, …
- How to use them on real data
  - text, image, structured data
  - your own project
- Underlying statistical and computational theory
- Enough to read and understand ML research papers

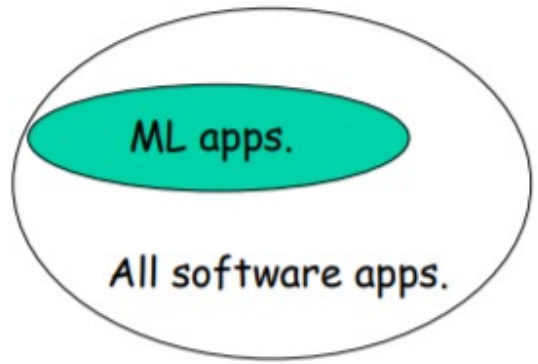
What this course can give you?

- Jobs in industry.
- Ability to conduct research in the most popular AI area.
- Build your own system to help people
  - Automatically recommend restaurants you family may like
- Ability to use ML/data analysis techniques for many different fields
  - Discovering new thermal energy materials
  - Predict when elders will fall using brain information
  - Diagnose a nose-related disease
  - Build better computers

- This is a big data era, and you have unlimited possibility when you master the power of the tool of data analysis.

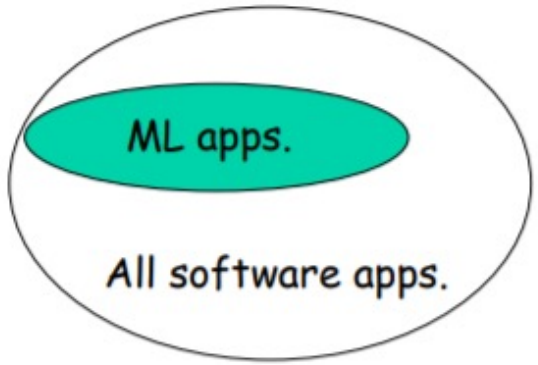
Machine Learning in Computer Science

- Machine learning already the preferred approach to
  - Speech recognition, Natural language processing
  - Computer vision
  - Medical outcomes analysis
  - Robot control
  - …
- This ML niche is growing (why?)



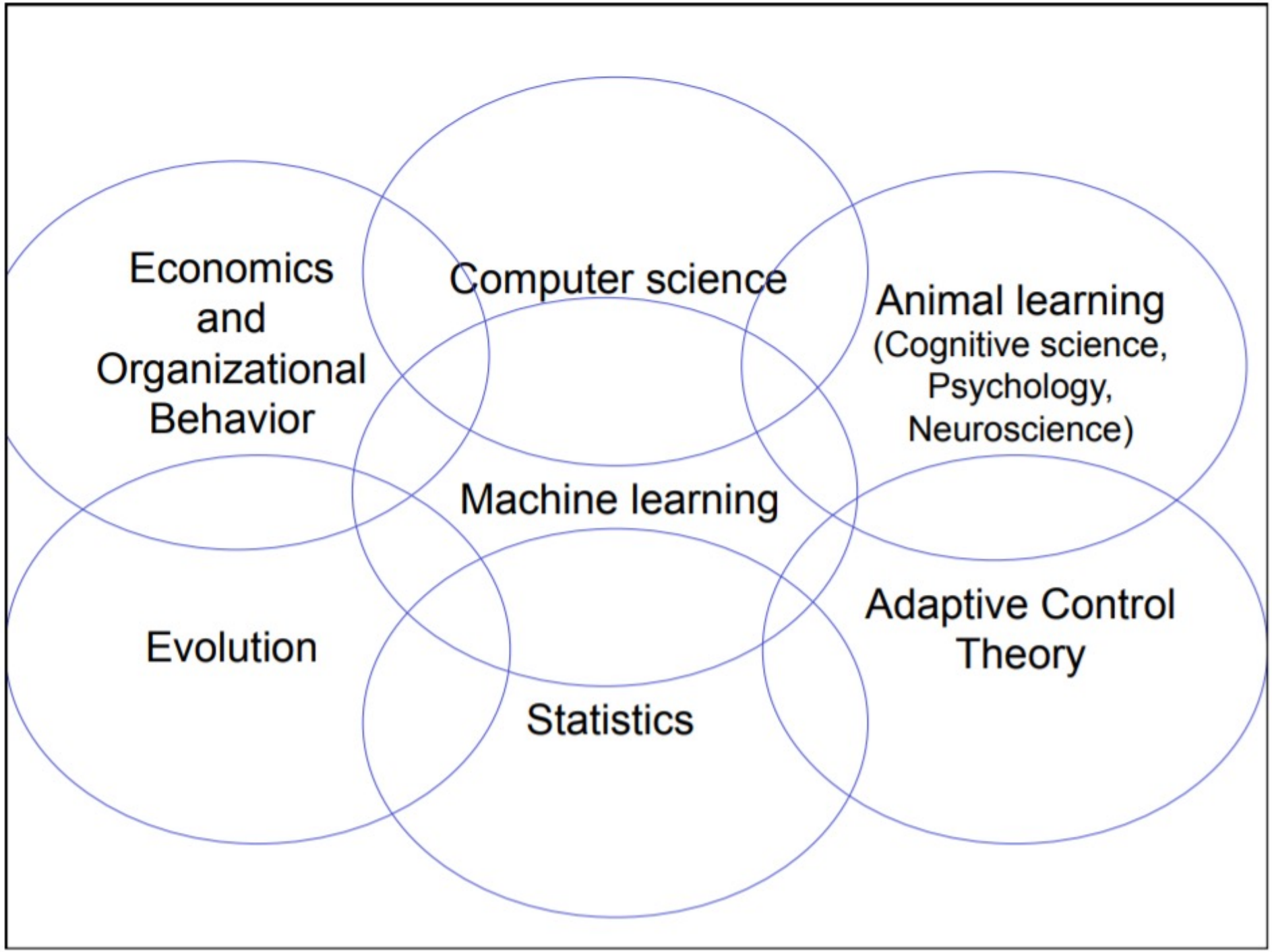
Machine Learning in Computer Science

- Machine learning already the preferred approach to
  - Speech recognition, Natural language processing
  - Computer vision
  - Medical outcomes analysis
  - Robot control
  - …



- This ML niche is growing
  - Improved machine learning algorithms
  - Increased volume of online data
  - Increased demand for self-customizing software

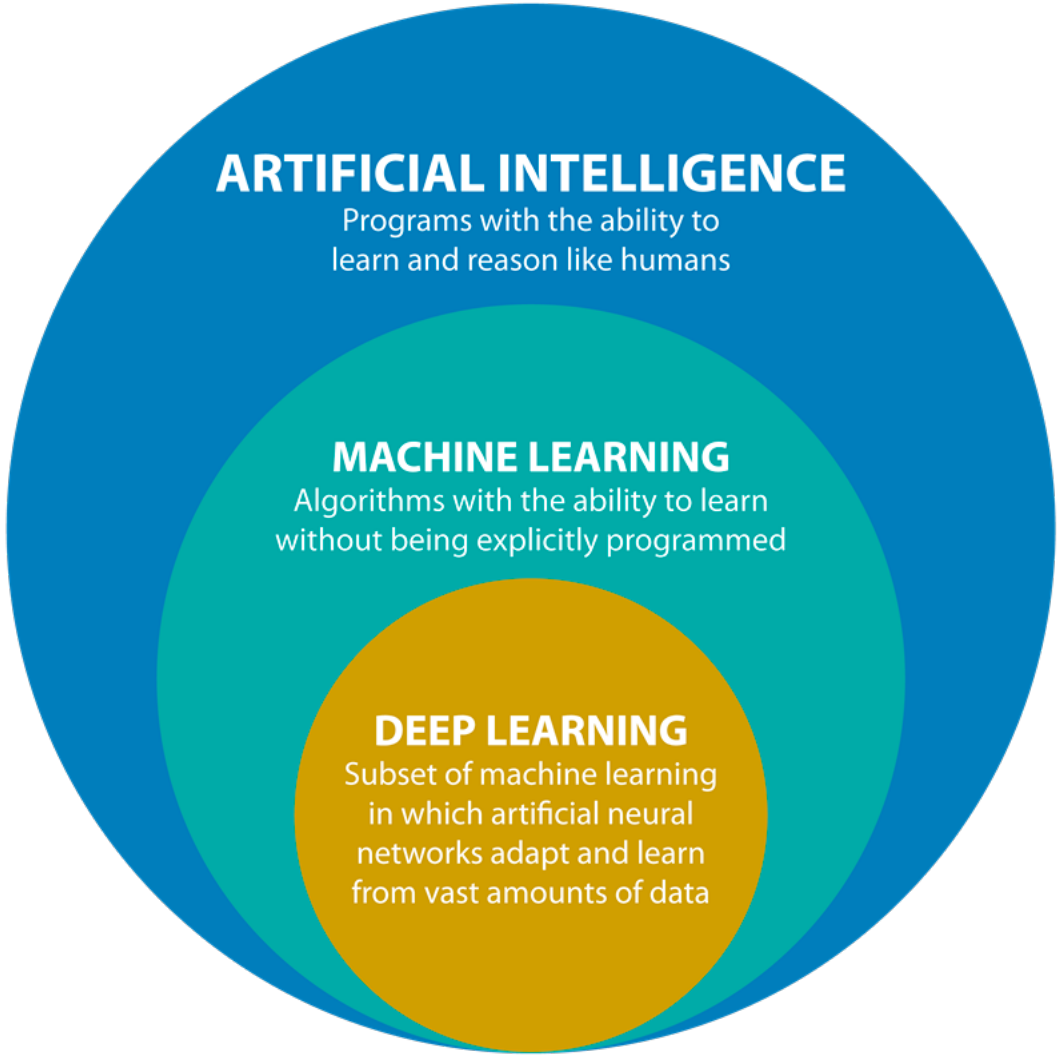
Tom's prediction: ML will be fastest-growing part of CS this century


Economics
and
Organizational
Behavior
Computer science
Animal learning
(Cognitive science,
Psychology,
Neuroscience)
Machine learning
Evolution
Statistics
Adaptive Control
Theory

Outline: Machine learning types and tasks

- A systematic view of machine learning
  - Machine learning VS AI
  - Supervised Learning VS Unsupervised Learning
  - Regression VS. classification
  - Training VS validation VS testing

AI VS. Machine Learning



- ***Artificial intelligence*** is a wide field, which aims at making machines intelligent. AI has a set of tools through which it enables a machine to mimic human intelligence.
  - Natural language processing
  - Robotics
  - Machine Learning
  - Self-driving cars
  - https://en.wikipedia.org/wiki/Artificial_intelligence#Reasoning,_problem_solving

- One of the tools AI have is ***machine learning***, that gives the machines to learn without being told explicitly what to do.

- Machine learning again has various tools in its pocket, one of them being neural networks. ***Neural networks*** try to mimic the activity of a human brain. ***Deep learning*** is the use of more sophisticated neural networks, with more non-linear layers, convolutional layers etcetera.

Supervised vs. Unsupervised Learning

- Supervised Learning
  - Goal: A program that performs a task as good as humans.
  - TASK – well defined (the target function)
  - EXPERIENCE – training data provided by a human
  - PERFORMANCE – error/accuracy on the task
- Unsupervised Learning
  - Goal: To find some kind of structure in the data.
  - TASK – vaguely defined
  - No EXPERIENCE
  - No PERFORMANCE (but, there are some evaluations metrics)