Introduction to Machine Learning Welcome

Instructor: Yang Xu
Slides adopted from Xiaobai Liu

About me

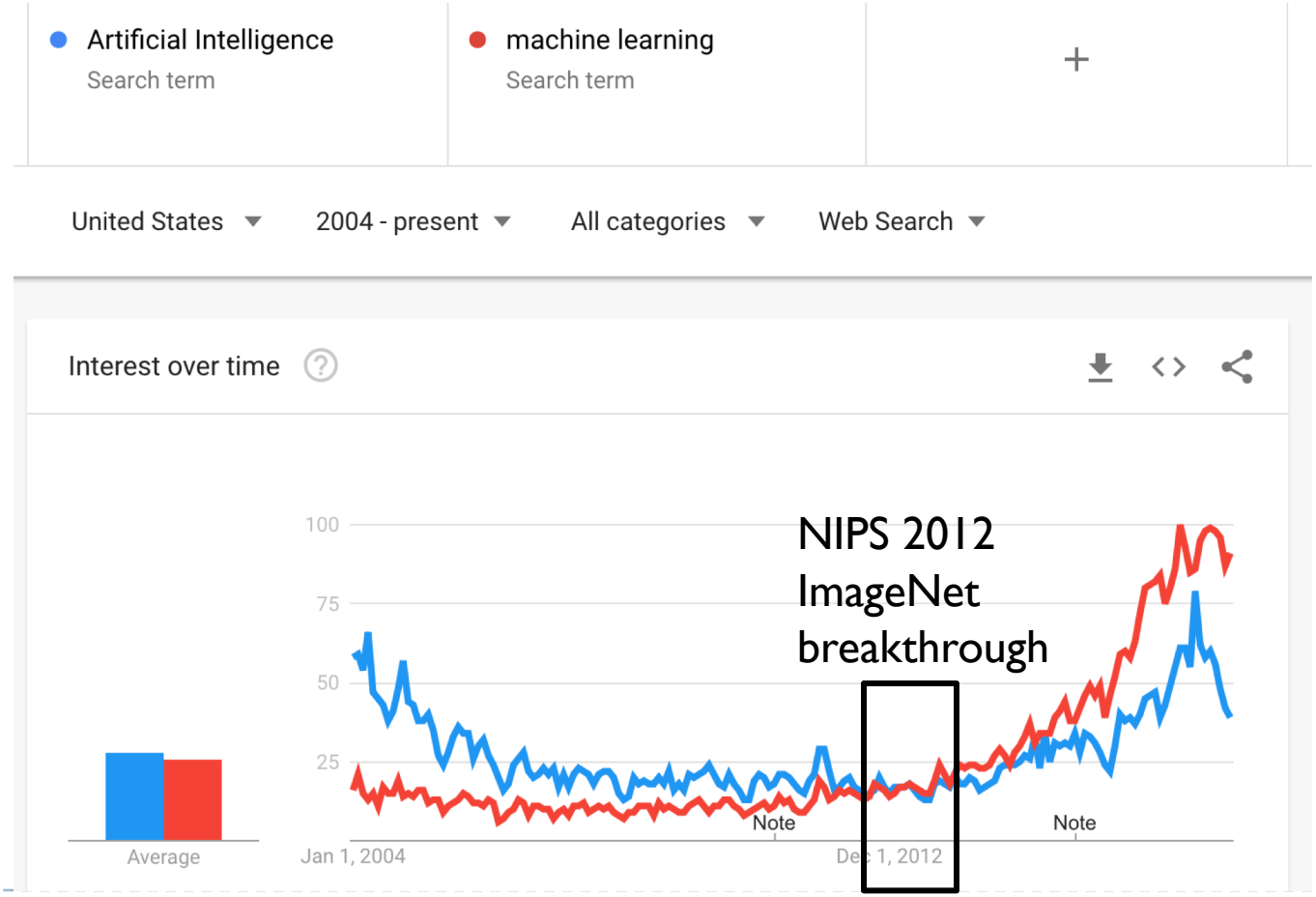
- Yang Xu, Assistant Professor of Computer Science, SDSU
  - Ph.D., Informatics (Information Sciences and Technology), The Pennsylvania State University
  - M.A., Psychology, Tsinghua University (China)
  - M.S., Electronic Engineering, Tsinghua University
- Research area:
  - Computational linguistics (Natural Language Processing, NLP), psycholinguistics, cognitive sciences
- Computational Linguistics and Cognitive Science Lab (CLCS, in preparation)
- Looking for self-motivated thesis students!

Machine Learning Evolves Rapid

- 2008
  - My first ML project: Principle Component Analysis (PCA) → Image classification
- Now
  - Deep learning
  - Convolutional Neural Network (CNN)
  - Recurrent Neural Network (RNN)
  - Generative models
  - Reinforcement Learning
  - …
- History of DL: https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html

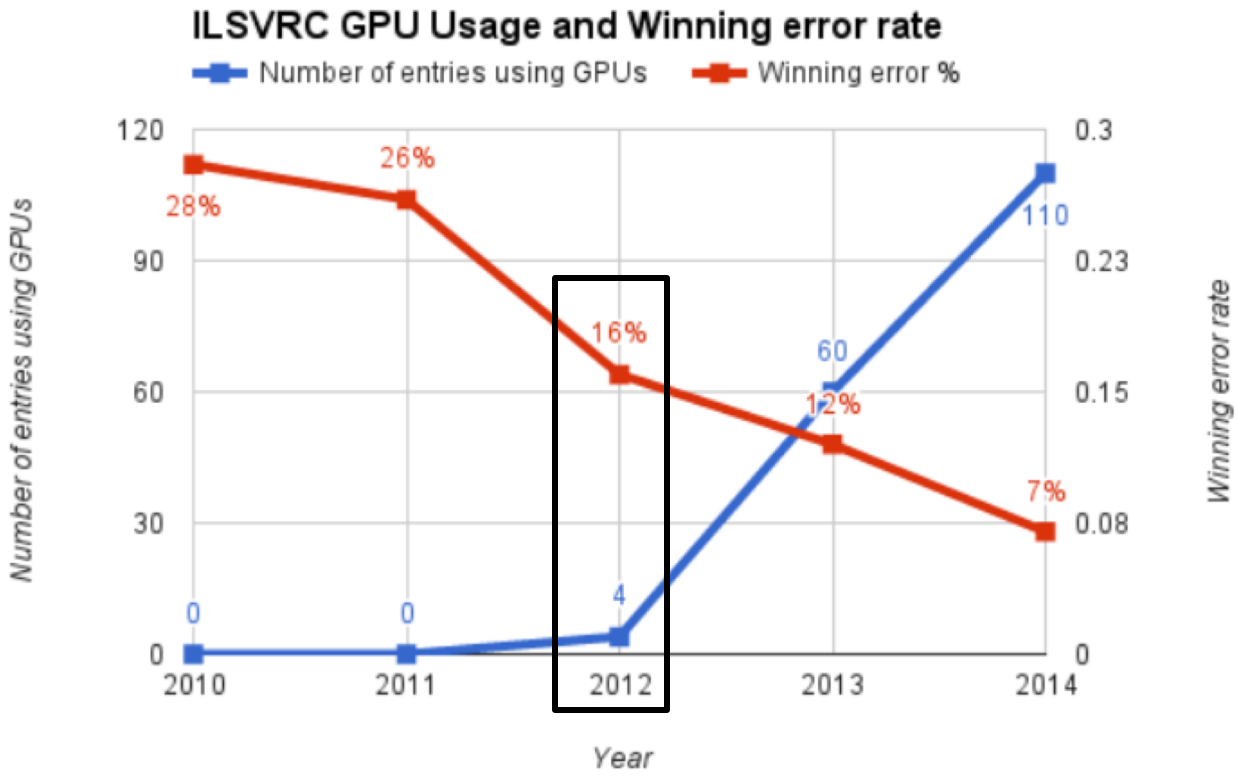
Booming of the field

- Key word search volume from Google Trends



Turning point around 2012

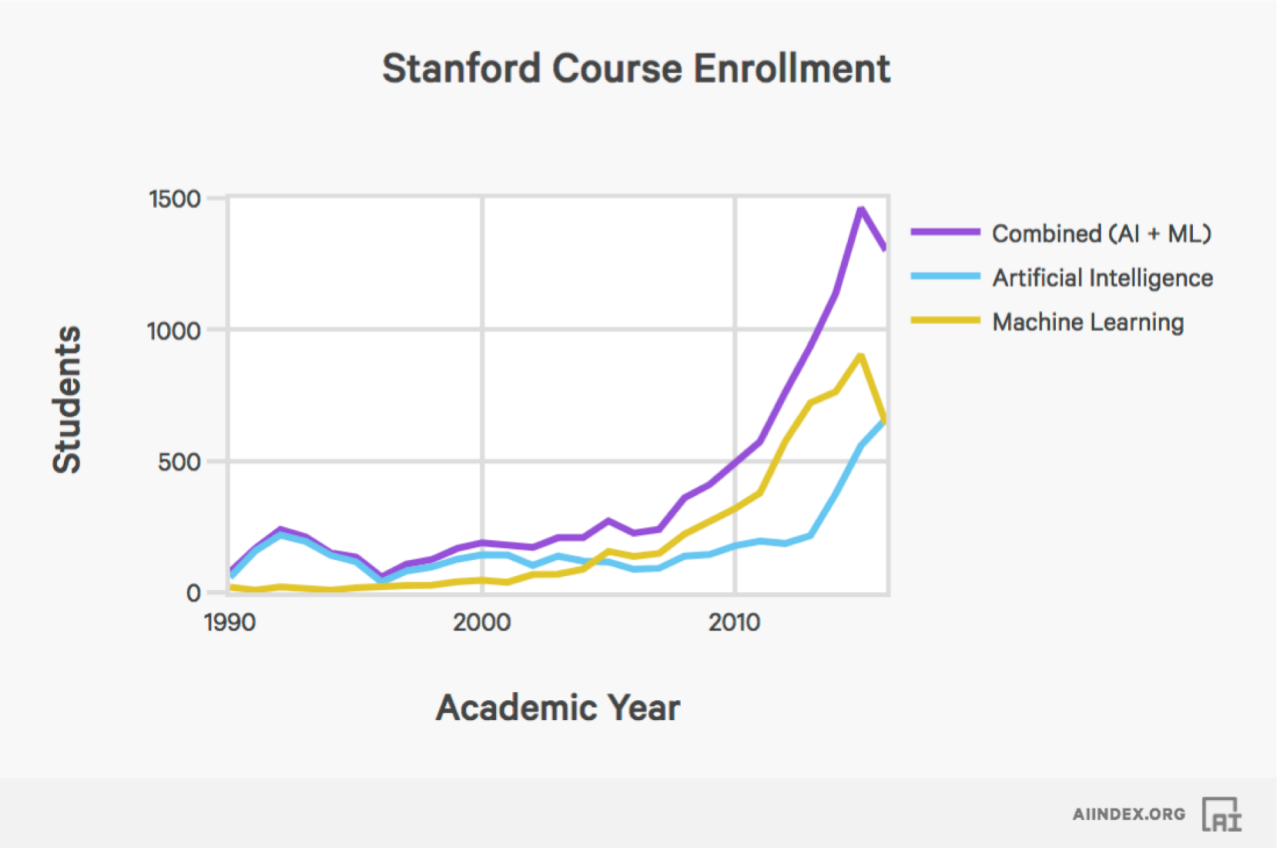
- Large Scale Visual Recognition Challenge (LSVRC)
- ImageNet, millions of labelled images



Graphic processing units (GPUs) used for training
→ Larger models

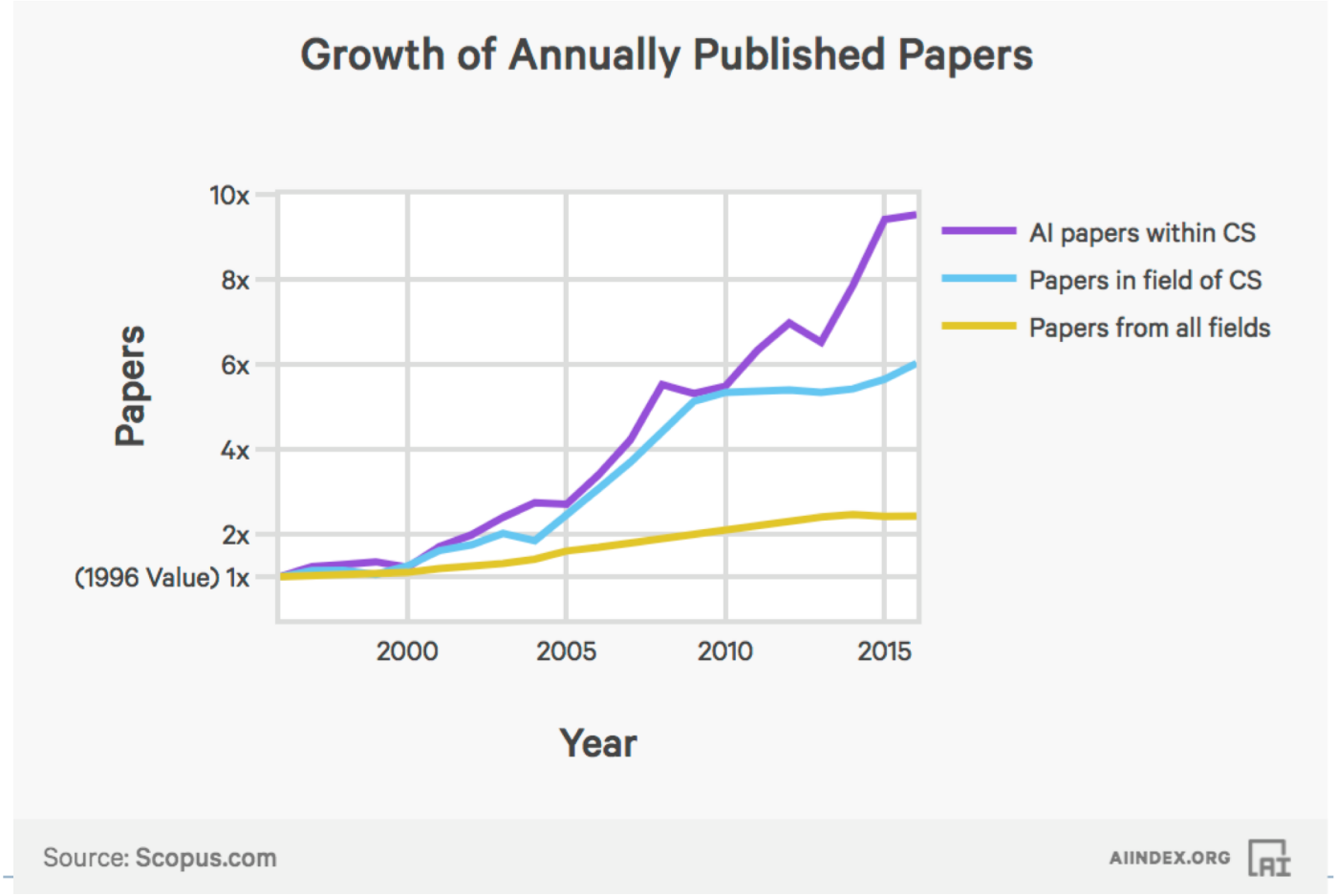
Booming of the field (cont.)

- Course enrollment at Stanford University (Report from aiindex.org)



Booming of the field (cont.)

- Research papers (From aiindex.org)



Why we need Machine Learning (What ML can do)

Case: Generating Natural Language

- Google Assistant (Project Duplex)
- https://youtu.be/D5VN56jQMWM

Case: Self-driving Car

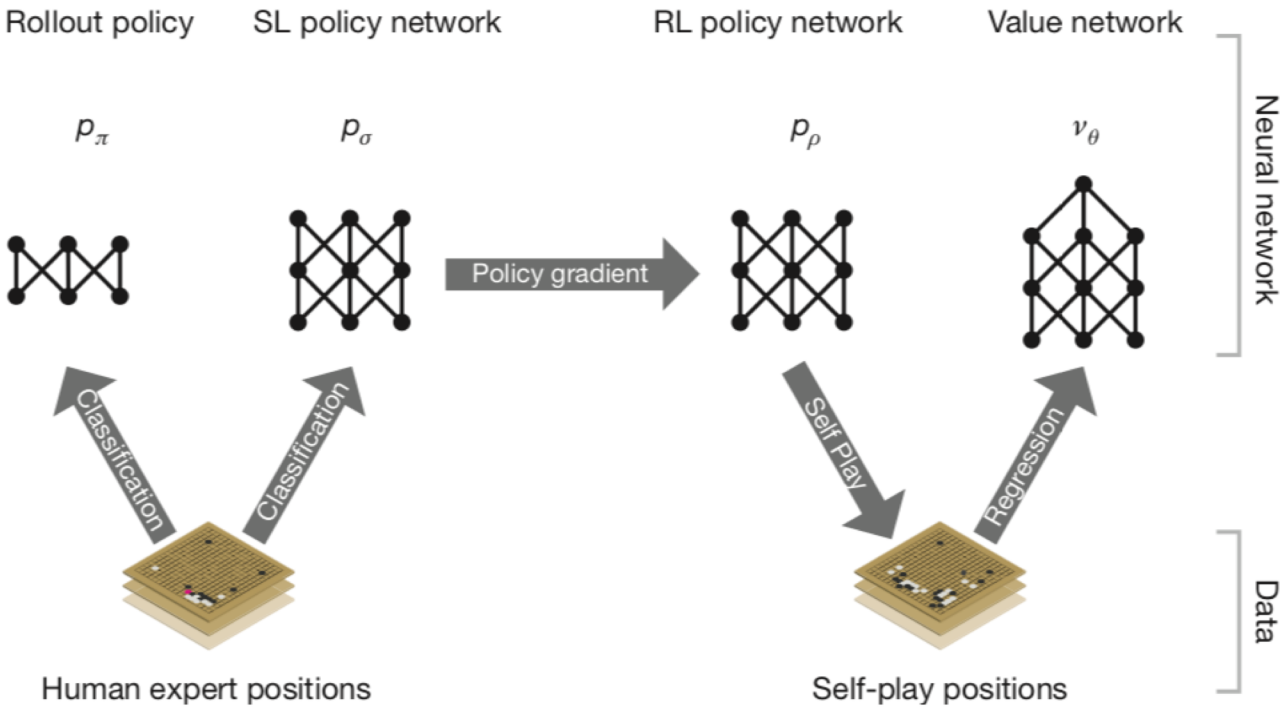
- Waymo
- https://youtu.be/B8R148hFxPw

Case: Self-Driving Car

- More details

Case: AlphaGo, AlphaGo Zero

- Machine learning technology used: neural network, reinforcement learning etc.



What is Machine Learning
-- what scholars say

Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

Arthur Samuel (1959).

Well-posed Learning Problem: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

Tom Mitchell (1998)

Case: Face Recognition (Computer Vision)

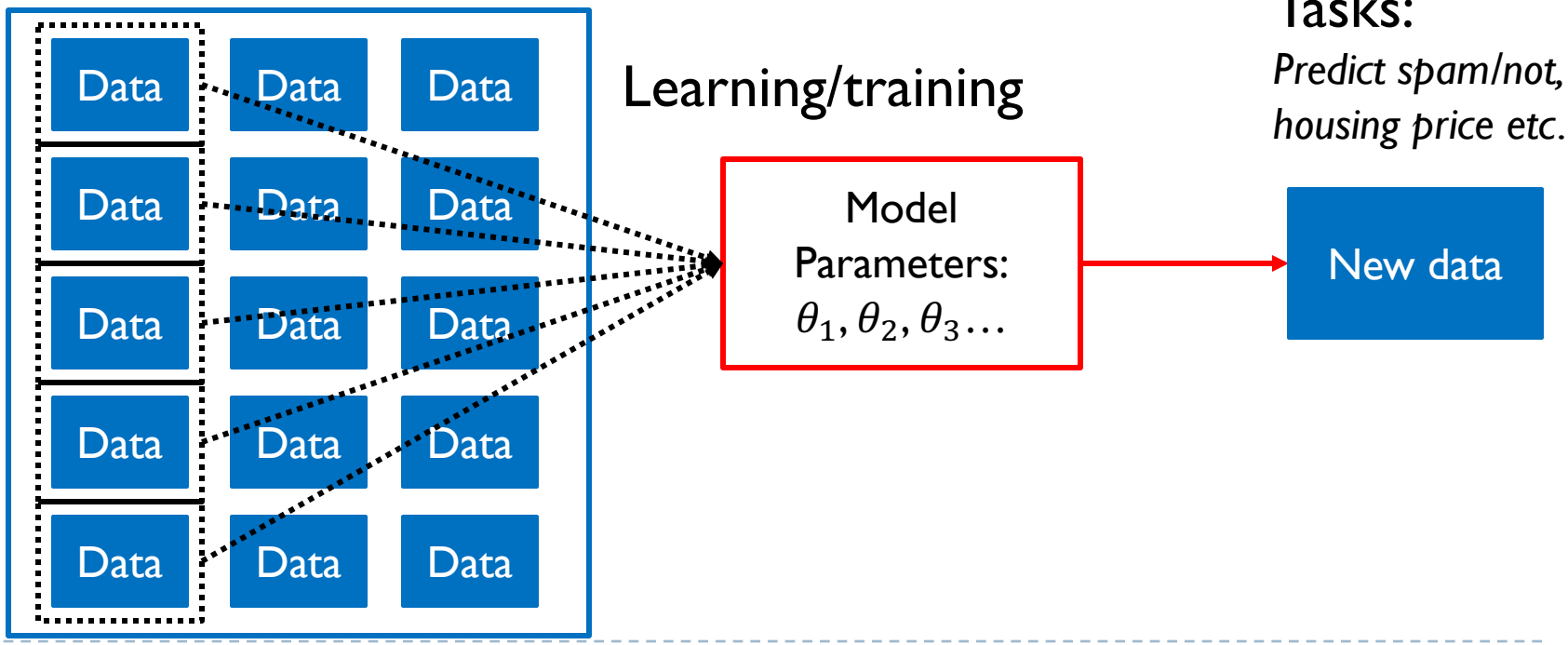
Experience: face images of the same persons

Task: recognize if an unseen face image is the same person

Performance: accuracy

What is machine learning?
-- A simpler mental model

- Machine builds a model of the world by learning from data (lots of them!)
- Model → a set of parameters

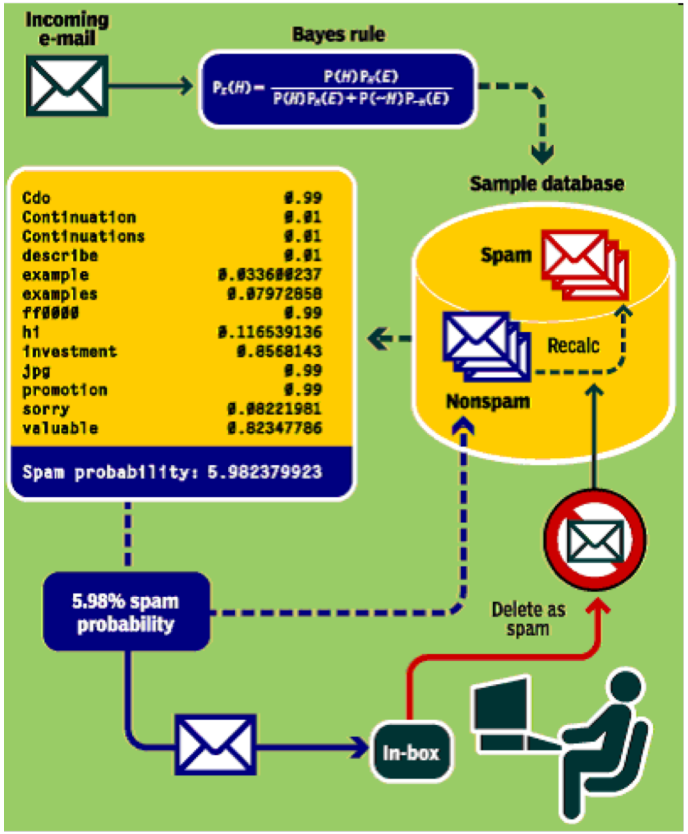


Case: Spam Email Detection

Binary classification model

Data: emails that were previously marked as spam or ham;

Task: binary classification, is this email ham or spam?



Case: Sentiment Classification

- **Categorical classification model**
- **Data**: comments, posts, tweets about products, people, or incidents.
- **Task**: decide or positive or negative the sentiment is (e.g., from -5 to +5)

Case: Book Recommendation

Customers Who Bought This Item Also Bought

Pattern Recognition and Machine Learning (Information Science and Statistics) by Christopher M. Bishop

★★★★☆ (30) $60.50

Artificial Intelligence: A Modern Approach (2nd Edition) (Prentice Hall Series in Artificial Intelligence) by Stuart Russell

★★★★☆ (76) $115.00

The Elements of Statistical Learning by T. Hastie

★★★★☆ (25) $72.20

Pattern Classification (2nd Edition) by Richard O. Duda

★★★★☆ (25) $115.00

Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems) by Ian H. Witten

★★★★☆ (21) $39.66

Continuous model

Data: historical browsing behavior from users

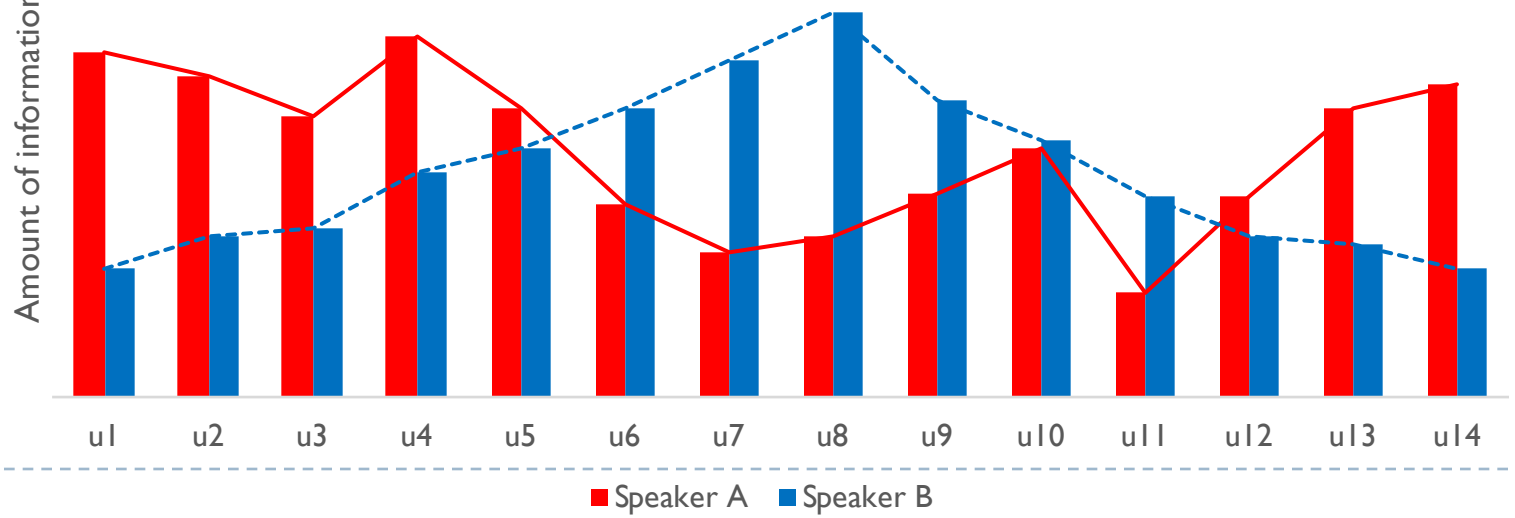
Task: predicting how much likely the user may like the new product

Case: Language Use in Dialogue
(Xu et al., 2016 a & b)

- **Language model**
  - **Data:** large body of text
  - **Task**: predict the probabilities of upcoming words

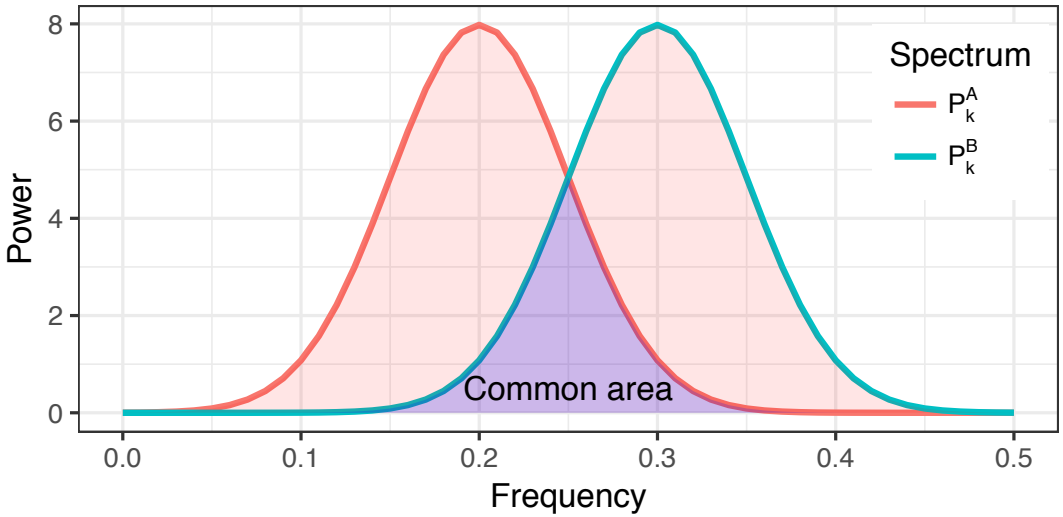
Example sentence: "*Who gave this book to Thomas*"

$$H(s_i) = -\frac{1}{4}\{\log p(this|\ Who,\ gave) + \log p(book|gave,\ this) + \log p(to|this,\ book) + \log p(\ Thomas|book,\ to)\}$$



Case: Predict Collaborative Performance (Xu et al., 2017)

- Continuous model. **Task**: predict the quality of collaboration



Features from Spectrum: Power spectrum overlap (**PSO**) & Relative Phase (**RP**)

Model	R^2
R&M	.17
R&M LENGTH only	.09
R&M LENGTH only (C=.5)	.1260
R&M (C=.5)	.1771
R&M + PSO + RP	**.2826**
R&M + PSO*RP	.2435
R&M LENGTH only + PSO*RP	.2494

Predict task performance using an **SVM model**, a classical machine learning model

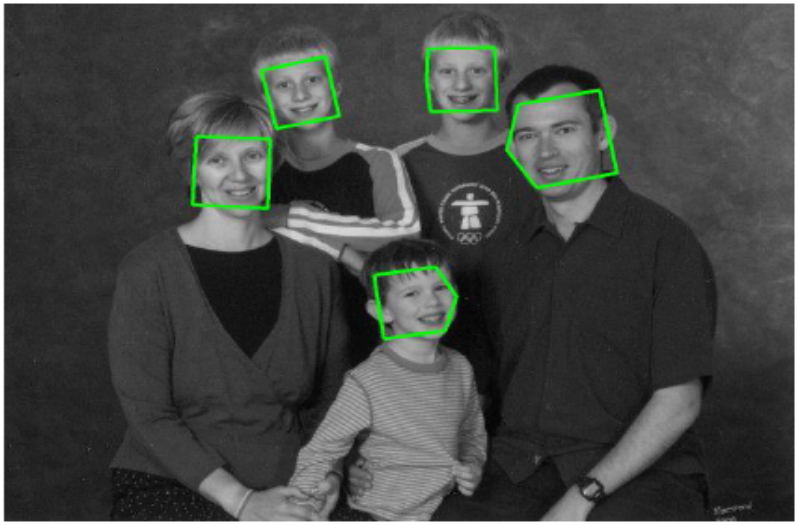
Case: Language Style and Social Power (Xu et al., 2018)

- Is how we speak influenced by our social status?
  - **Task**: predict the use of certain language

- Wikipedia talk-page
  - Editors accommodate towards Admins?
- U.S. Supreme Court
  - Lawyers accommodate towards Judges?

- **Logistic regression model**, a simple ML model.

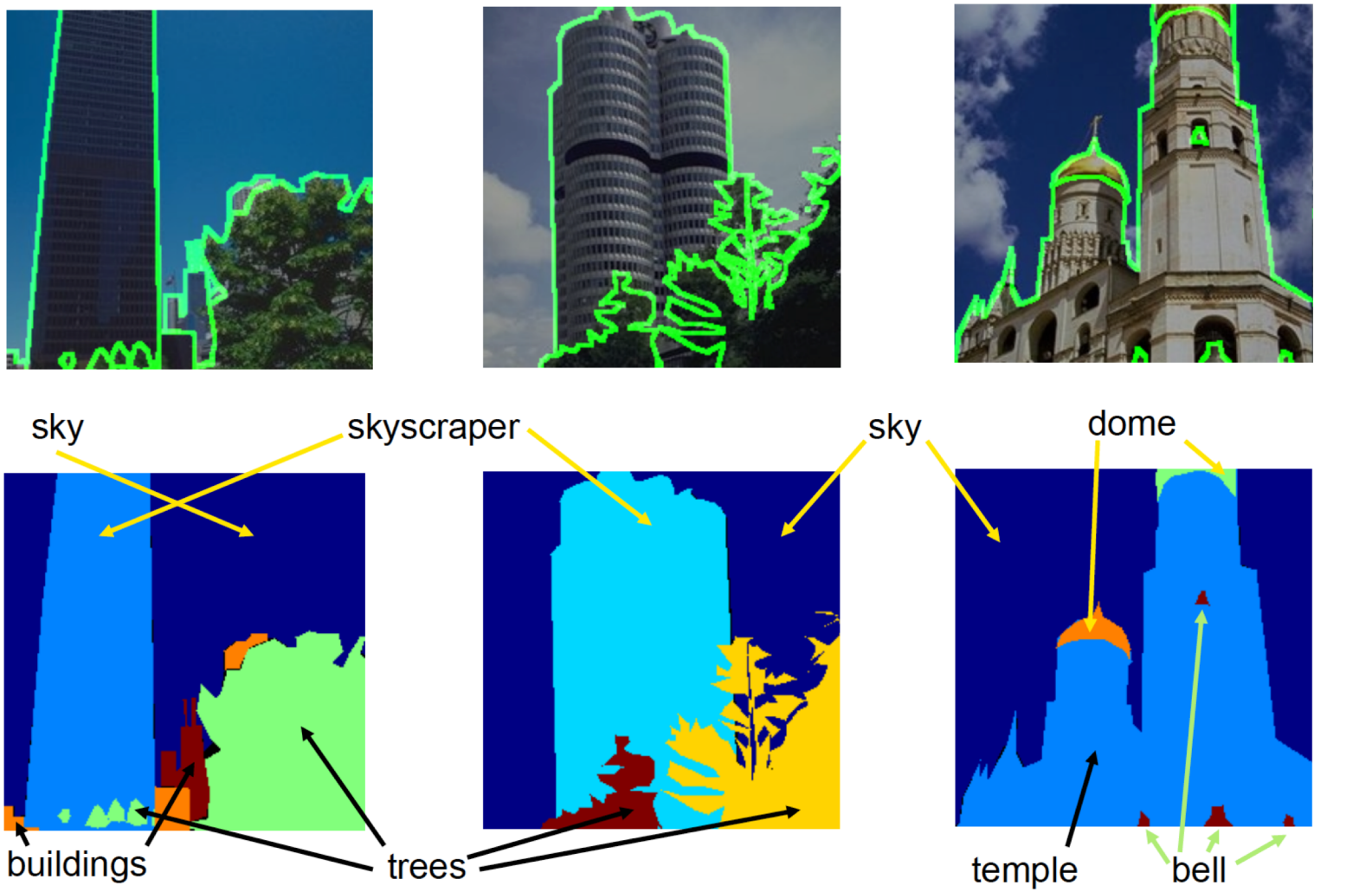
Case: Face Detection (Computer Vision)



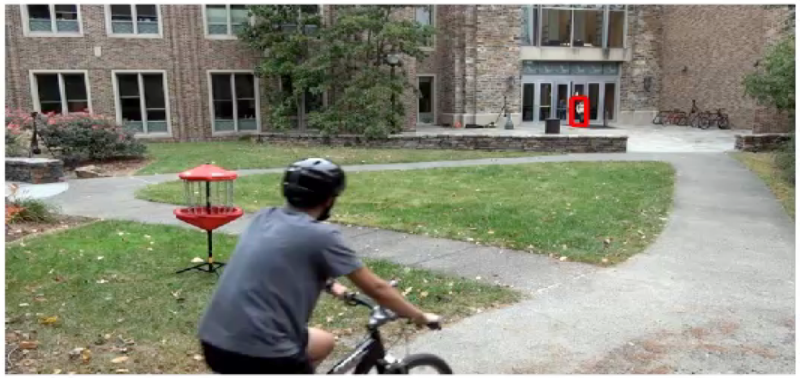
K. Murphy & Family

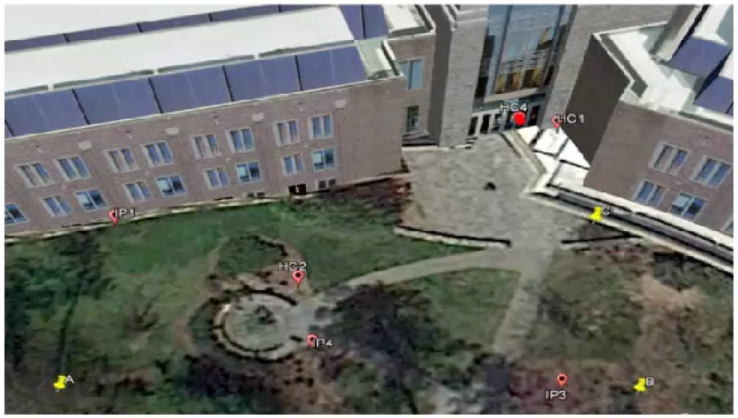
Based on classifiers trained from tens of thousands of example faces (Viola & Jones, 2004)

Case: Scene Understanding (Computer Vision)



Case: Human Tracking



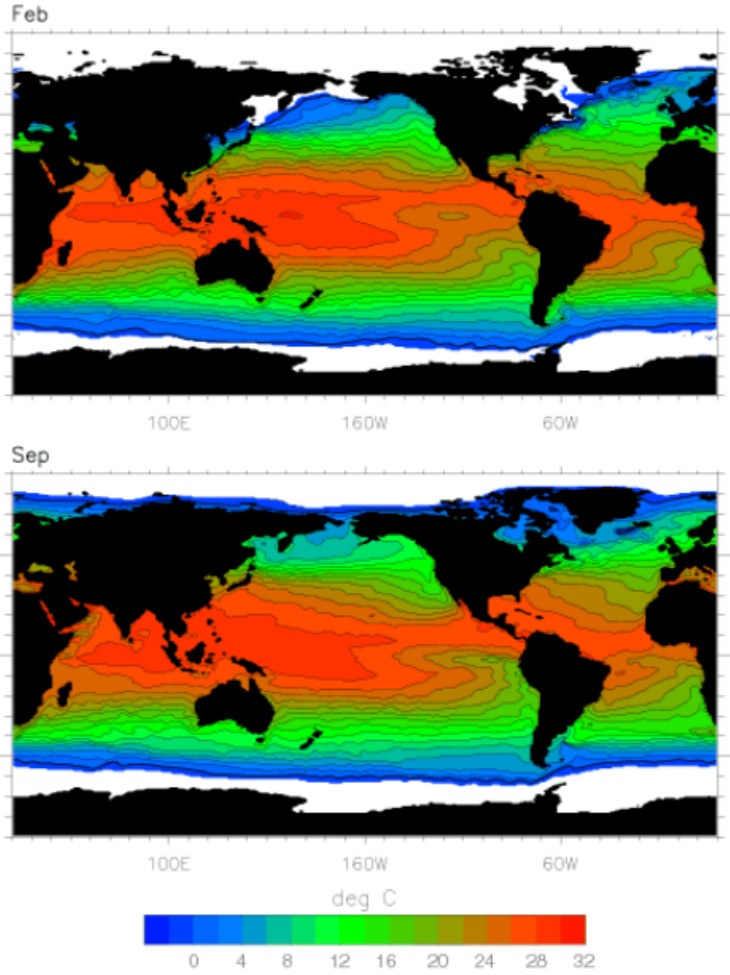


Garden: HC3

[Xiaobai Liu et al. AAAI' 2016]

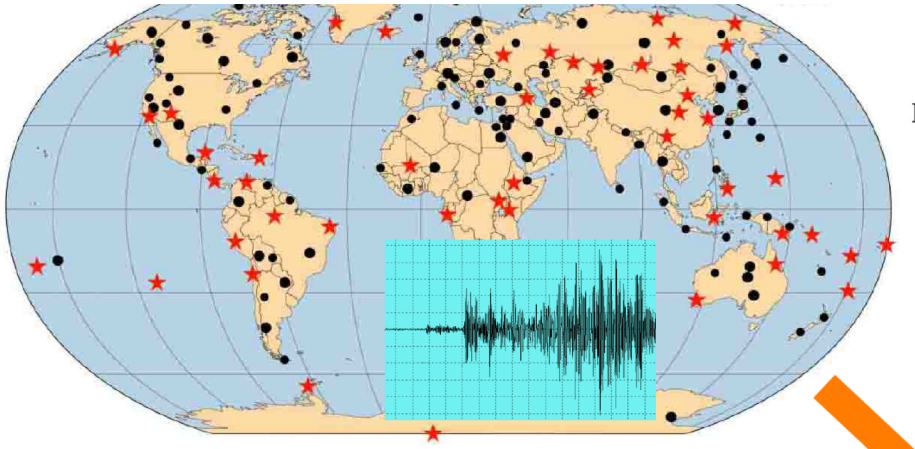
Case: Climate Modeling

- Satellites measure sea-surface temperature at sparse locations
  - Partial coverage of ocean surface
  - Sometimes obscured by clouds, weather
- Would like to infer a dense temperature field, and track its evolution

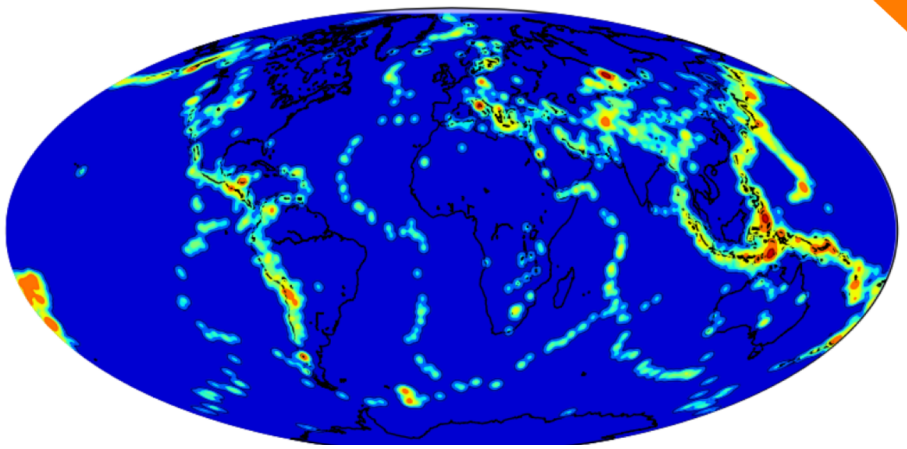


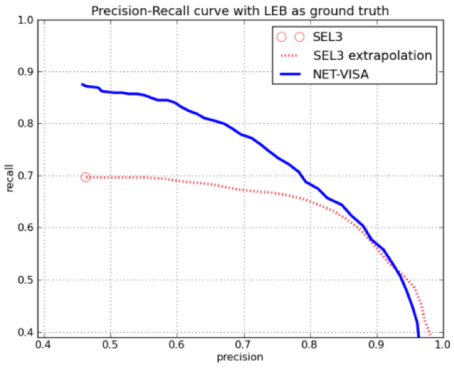
NASA Seasonal to Interannual Prediction Project
http://ct.gsfc.nasa.gov/annual.reports/ess98/nsipp.html

Case: Global Seismic Monitoring



CTBTO
PREPARATORY COMMISSION





Final Case: word2vec

- Distributed representation of words
- NLP techniques that maps words to vectors

$$v_{queen} - v_{woman} + v_{man} \approx v_{king}$$

Data: Large body of text
Task: ?

Summary

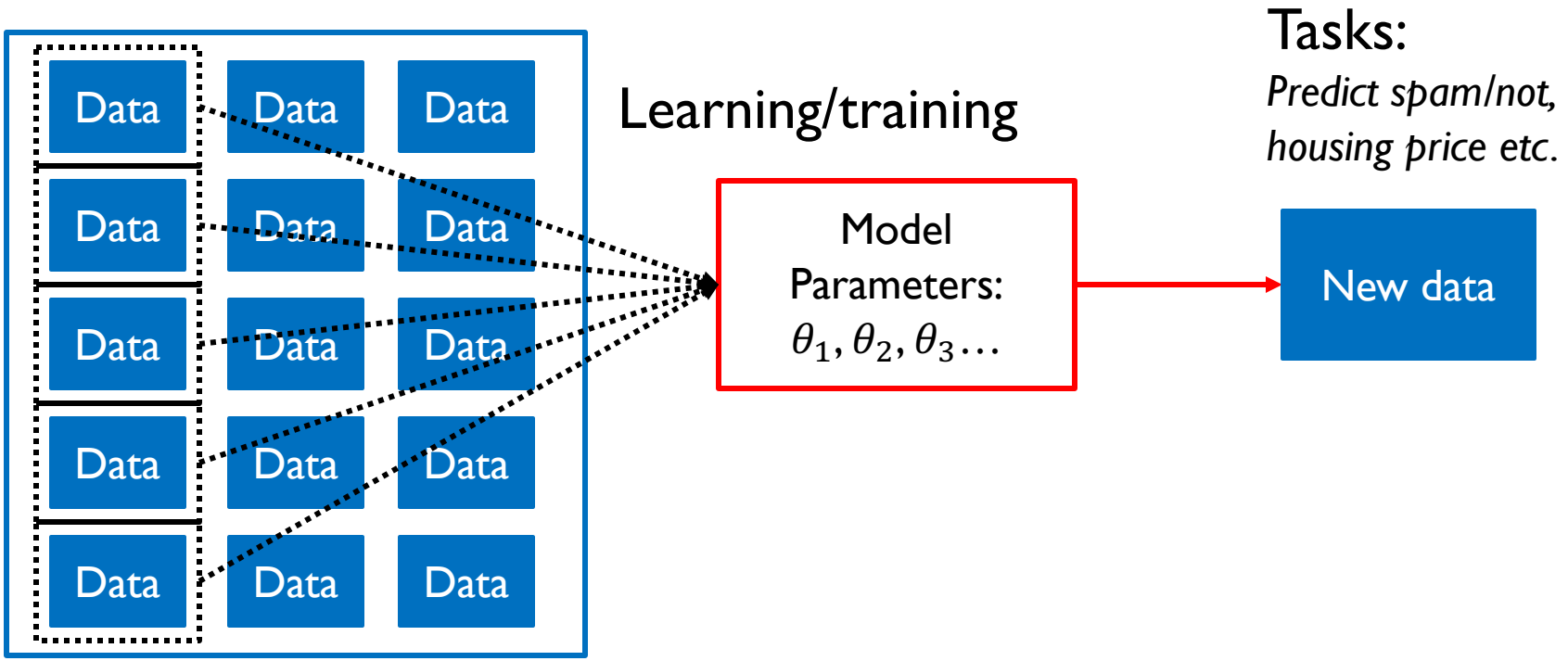
Computer Vision	Scene Understanding, Face Recognition, Face Detection, human tracking
Transportation	Self-driving vehicle
Natural Language	Understanding, generation
Commercial	Movie Rating, book recommendation
Medical	Tumor Diagnose
Financial	Market prediction
Climate	Weather Prediction
Robot	Mapping, localization
Geology	Seismic Analysis
….	

Recap: Machine Learning

- Predicting the future from the past
  - Past: training data, collection of examples
  - Future: knowledge/information about novel examples

- Examples
  - Binary classification: is it spam or not?
  - Continuous classification:
  - Recognition: what sentence was spoken?
  - Detection: when and where have seismic events(earthquakes) occurred?

Recap: a simple mental model



Course Topics

- Linear Regression
- Logistic Regression
- Neural Network
- Support Vector Machine
- Unsupervised learning
- Ensemble methods: bagging and boosting
- Dimensionality Reduction

Course Organization

- In-class meeting
  - Interaction
  - Discussion
- After-class assignment (programming)
  - Projects
- Programming Languages
  - Python
  -

What you can learn

- Given two candidate machine learning models, which is better?
  - Accuracy
  - Complexity
  - Are all mistakes equally bad?
- Given a family of classifiers with free parameters,
  - General design principles?
  - What happen if more data?
  - How to test classifiers?
  - How to avoid overfitting?

Evaluation

- Homework Assignment (HA, 6 total)
  - 13.3% HA1: Programming Environment
  - 13.3% HA2
  - 13.3% HA3
  - 13.3% HA4
  - 13.3% HA5
  - 13.3% HA6
- Final project (20%, due: 12/14/2018)
  - Programming
  - Report
  - Presentation (optional)

Resources

- Blackboard
  - Slides (updated after class meetings)
  - Sample codes
  - Homework assignments

- Textbooks
  - Lecture notes and slides will cover all the content of this course.
  - Optional Reading: ***Machine Learning: a Probabilistic Perspective*** by Kevin Patrick Murphy
    Online available, http://www.cs.ubc.ca/~murphyk/MLbook/
  - Optional Reading: ***Pattern Recognition and Machine Learning***, Christopher M. Bishop, Springer-Verlag New York, Inc. Secaucus, NJ, USA ©2006 ISBN:0387310738

Office hours

- 3:00pm-4:00pm, Tuesday/Thursday, GMCS#542

Let's know each other

- Themes:
  - Who are you?
  - What's the coolest story/thing regarding Machine Learning/Artificial Intelligence you ever heard?