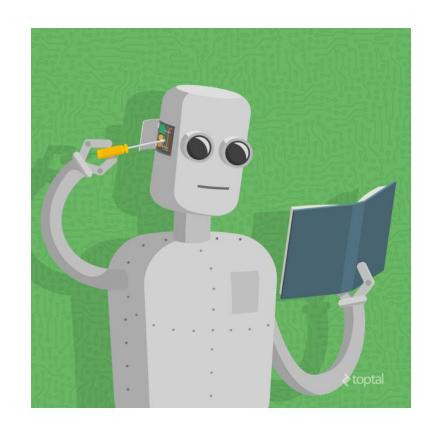
# Introduction to Machine Learning

Kyle Swanson



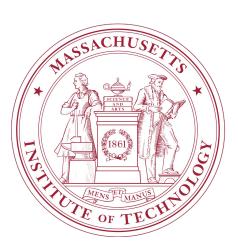
## **Today**

- About me
- Class details
- Syllabus
- Part I Machine Learning Overview
  - What is machine learning?
  - History
  - o Problems and applications
- Part II Features and Classifiers
  - O What are features?
  - Feature-based classification
  - Generalization



#### **About me**

- Name
  - Kyle Swanson
- Hometown
  - Bronxville, NY
- School
  - 4th year at the Massachusetts Institute of Technology (MIT)
  - Studying computer science and mathematics with a minor in music
- Research interests
  - Working with Regina Barzilay in MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL)
  - Applying deep learning to medical imaging to improve cancer detection







#### **Class details**

- First day of class
  - o Monday, January 8th, 2018
- Last day of class
  - o Tuesday, January 30th, 2018
- Class meets Monday, Tuesday, and Wednesday
  - Exception: No class on Wednesday, January 31st
- 2 hour lecture, 10:00-12:00 in G318
- 2 hour lab, 15:00-17:00 in G318

#### **Syllabus**

- Week 1
  - Monday Introduction to Machine Learning
  - Tuesday Linear Classifiers and the Perceptron Algorithm
  - Wednesday Maximum Margin Classifiers and Support Vector Machines
- Week 2
  - Monday Kernels and Non-Linear Classifiers
  - Tuesday Ensembles and Boosting
  - Wednesday Content Recommendation via Collaborative Filtering
- Week 3
  - Monday Neural Networks I
  - Tuesday Neural Networks II
  - Wednesday Convolutional and Recurrent Neural Networks
- Week 4
  - Monday Unsupervised Learning
  - Tuesday Reinforcement Learning

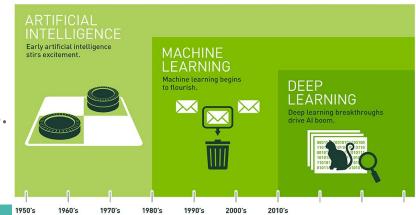
## Part I - Machine Learning Overview

- What is machine learning?
- History
  - Artificial intelligence
  - Machine learning
  - Deep learning
- Problems and applications
  - Types of problems
    - Classification
    - Regression
    - Generation

- Types of learning
  - Supervised learning
  - Reinforcement learning
  - Unsupervised learning

## What is machine learning?

- **Artificial intelligence:** The ability of a computer to perform tasks commonly associated with intelligent beings.
  - Turing Test, 1950: "Are there imaginable digital computers which would do well in the imitation game?"
- Machine learning: "Field of study that gives computers the ability to learn without being explicitly programmed." (Arthur Samuel, 1959)
- Deep learning: A subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.



#### **History**

- Artificial intelligence
  - Dartmouth Summer Research Project on Artificial Intelligence, 1956
    - Organized by John McCarthy, who coined the term "artificial intelligence" in 1955
    - Attended by Marvin Minsky, Claude Shannon, other founding fathers of Al

"An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

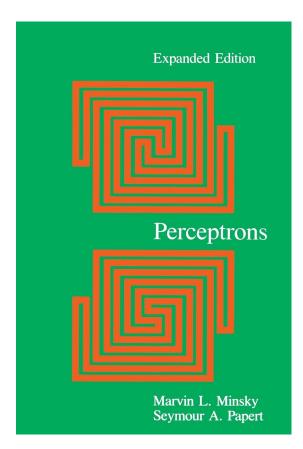
- Boom and bust cycles
  - Al winters of 1974-1980, 1987-1993
- Al used search and rule-based methods (<u>video</u>)



## **History**

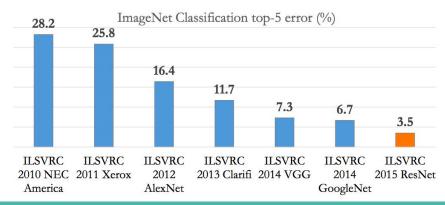
- Machine learning
  - "Machine learning" coined by Arthur Samuel, 1959
    - Developed a machine which learned to play checkers
  - Perceptron invented by Frank Rosenblatt, 1957 (video)
  - Perceptrons by Marvin Minsky and Seymour Papert, 1969
    - Limitations of single-layer Perceptron
    - Al winter of the 1970s

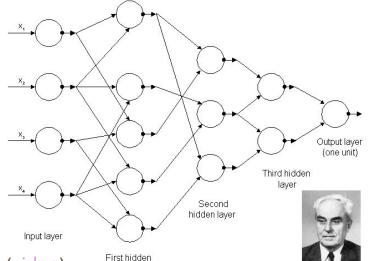




#### **History**

- Deep learning
  - Threshold Logic Units of McCulloch and Pitts, 1943
  - First deep architecture by Alexey Ivakhnenko, 1965
  - Advanced models
    - CNN for digit recognition by Yann LeCun, 1993 (video)
    - LSTM for natural language by Hochreiter and Schmidhuber, 1997
  - Success only came with GPUs and AlexNet on ILSVRC, 2012





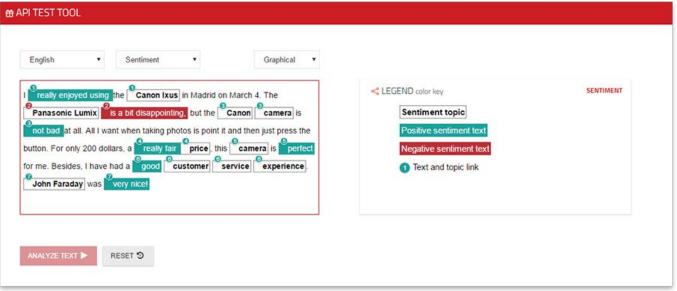
laver

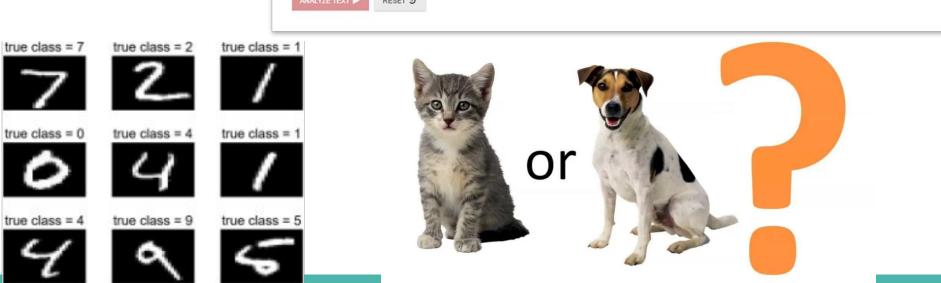
## **Types of problems**

- Classification
  - Discrete prediction
- Regression
  - Continuous prediction
- Generation
  - Output creation

#### **Classification**

- Discrete prediction

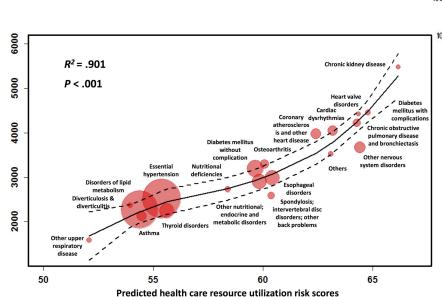


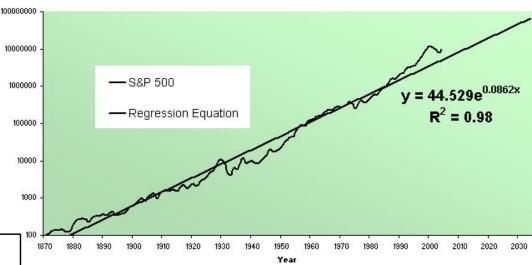


## Regression

Next 6-month resource utilization per person (US \$)

- Continuous prediction





#### **Generation**

- Output creation



"Two pizzas sitting on top of a stove top oven"



"A group of young people playing a game of frisbee"

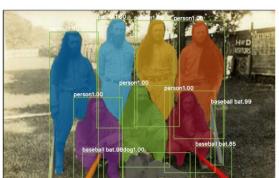




## Types of learning

- Supervised learning
  - Given a set of data and labels, learn to predict the labels
- Reinforcement learning
  - Given a task and a reward function, learn to perform the task
- Unsupervised learning
  - o Given a set of data *without* labels, learn underlying features of the data

# **Supervised learning**















leopard

leopard

jaguar

cheetah

snow leopard

Egyptian cat

spider monkey

howler monkey









convertible

beach wagon

fire engine

grille

pickup



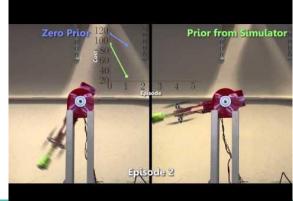
gill fungus ffordshire bullterrier

currant

dead-man's-fingers

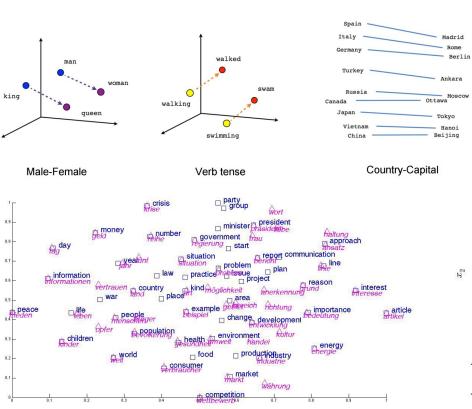
Reinforcement learning

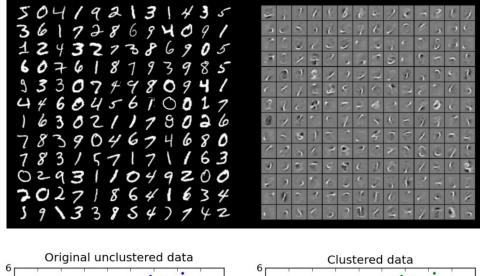


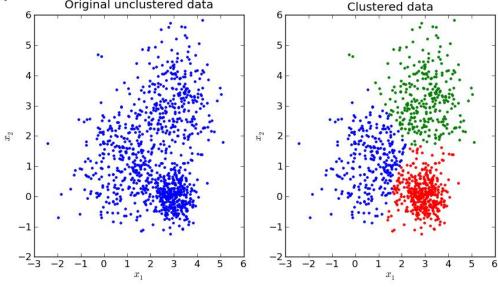




## **Unsupervised learning**







## **Part I - Machine Learning Overview Summary**

- What is machine learning?
- History
  - Artificial intelligence
  - Machine learning
  - Deep learning
- Problems and applications
  - Types of problems
    - Classification
    - Regression
    - Generation

- Types of learning
  - Supervised learning
  - Reinforcement learning
  - Unsupervised learning

#### **Part II - Classifiers and Features**

- What are features?
- Feature-based classification
  - Feature selection
  - Classifying with features
- Generalization
  - Training vs. testing
  - Overfitting vs. generalization

#### What are features?

- Informative properties contained in the data
- Example
  - Data = [(Kyle Swanson, 22, male), (Dana Adylova, 21, female)]
- Features may be explicit or implicit
  - o Explicit: Name, age, gender
  - Implicit: Nationality
- Selecting good features is key in machine learning
  - Example: Predicting taste in music
    - Age and gender are important
    - Name is probably not

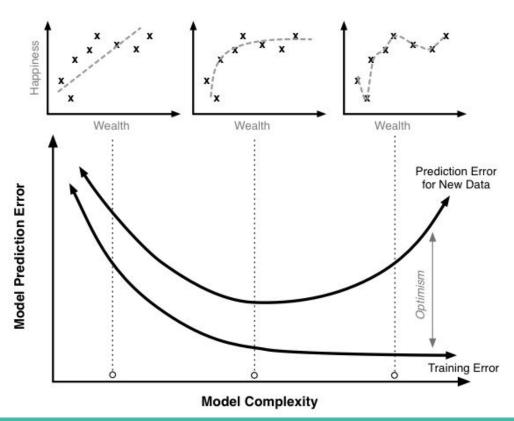
#### **Feature-based classification**

- Feature extraction
  - Choose relevant features
  - Extract features from the data
- Feature representation
  - o d-dimensional vector in  $\Re^d$
- Classification
  - Build a model which draws a decision boundary

#### **Generalization**

- How do we know how well the model performs?
  - Metrics such as accuracy
- Problem: Overfitting
  - The model performs well on data it has seen before but fails on new examples
- (Partial) Solution
  - Train and test sets
  - Model selection

## Overfitting vs. generalization

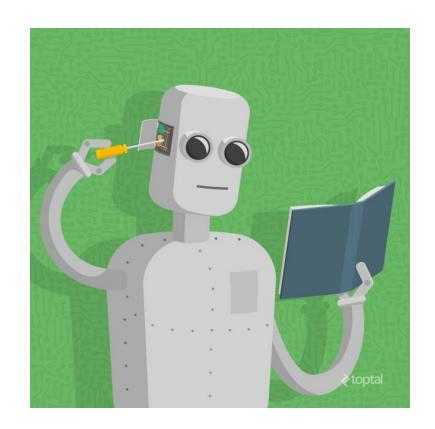


## Part II - Classifiers and Features Summary

- What are features?
- Feature-based classification
  - Feature selection
  - Classifying with features
- Generalization
  - Training vs. testing
  - Overfitting vs. generalization

#### **Summary**

- About me
- Class details
- Syllabus
- Part I Machine Learning Overview
  - What is machine learning?
  - History
  - o Problems and applications
- Part II Features and Classifiers
  - O What are features?
  - Feature-based classification
  - Generalization



#### What's next?

- Today
  - Lab: Setup and data loading
- Tomorrow
  - Lecture: Linear classifiers and the perceptron algorithm
  - Lab: Implementing the perceptron algorithm for sentiment analysis