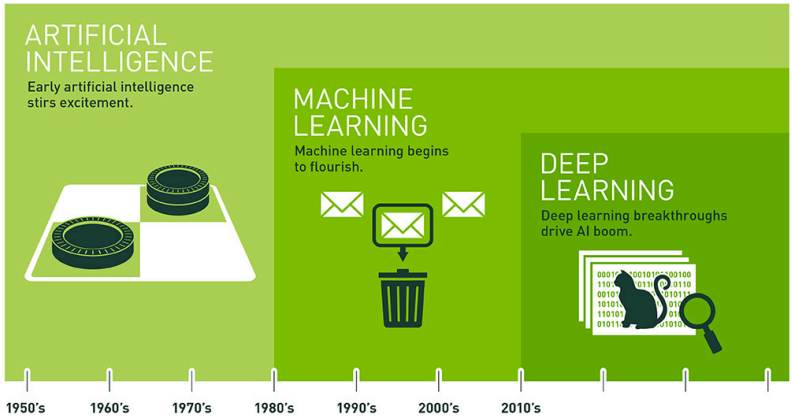
Machine Learning

CSC 461: Machine Learning

Fall 2021

Prof. Marco Alvarez
University of Rhode Island

AI vs ML vs DL



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

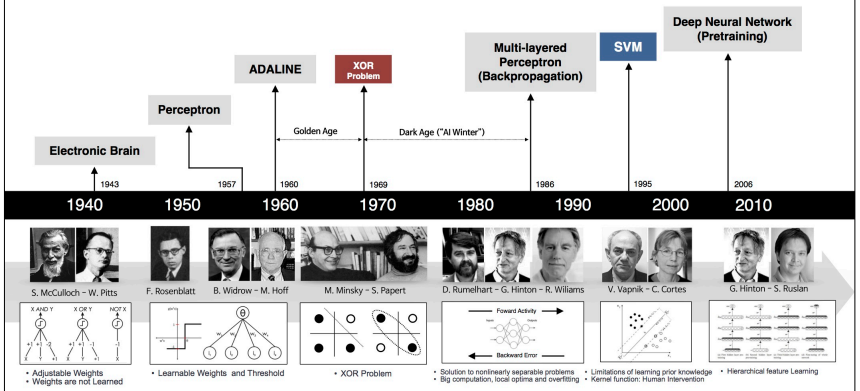
https://news.developer.nvidia.com/on-demand-webinar-deep-learning-demystified/

A (short) history of AI/ML

- **1940-1950: Early days**
  - ✓ 1943: McCulloch & Pitts: Boolean circuit model of brain
  - ✓ 1950: Turing's "Computing Machinery and Intelligence"
- **1950—70: Excitement: Look, Ma, no hands!**
  - ✓ 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
  - ✓ 1956: Dartmouth meeting: "Artificial Intelligence" adopted
  - ✓ 1965: Robinson's complete algorithm for logical reasoning
- **1970—90: Knowledge-based approaches**
  - ✓ 1969—79: Early development of knowledge-based systems
  - ✓ 1980—88: Expert systems industry booms
  - ✓ 1988—93: Expert systems industry busts: "AI Winter"
- **1990— 2012: Statistical approaches + subfield expertise**
  - ✓ Resurgence of probability, focus on uncertainty
  - ✓ General increase in technical depth
  - ✓ Agents and learning systems... "AI Spring"?
- **2012— …: Excitement: Look, Ma, no hands again?**
  - ✓ Big data, big compute, neural networks
  - ✓ Some re-unification of sub-fields
  - ✓ AI used in many industries

from: CS 188: Introduction to Artificial Intelligence, Berkeley

Milestones in AI/ML



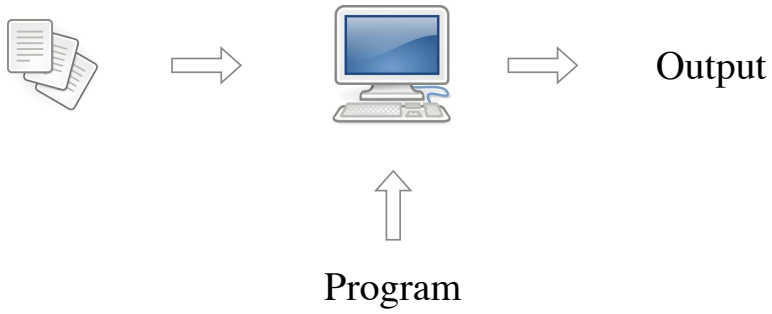
http://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html

What is Machine Learning?

What is Machine Learning?

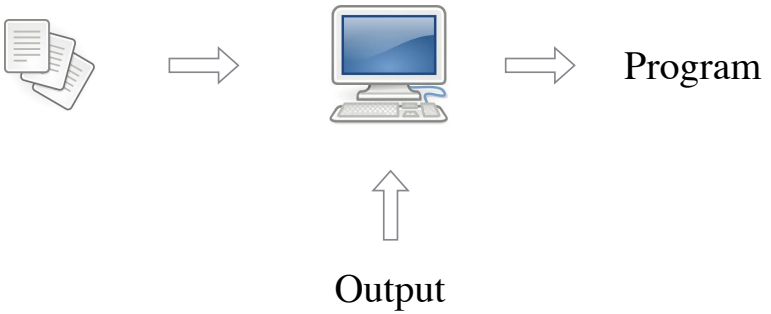
- Machine Learning is … [Mitchell]
  - ✓ the study of computer algorithms that learn from experience **E** with respect to a particular task **T** and performance metric **P**
- Remember the badges problem?
  - ✓ can you identify **E**, **T** and **P**?

Traditional Programming



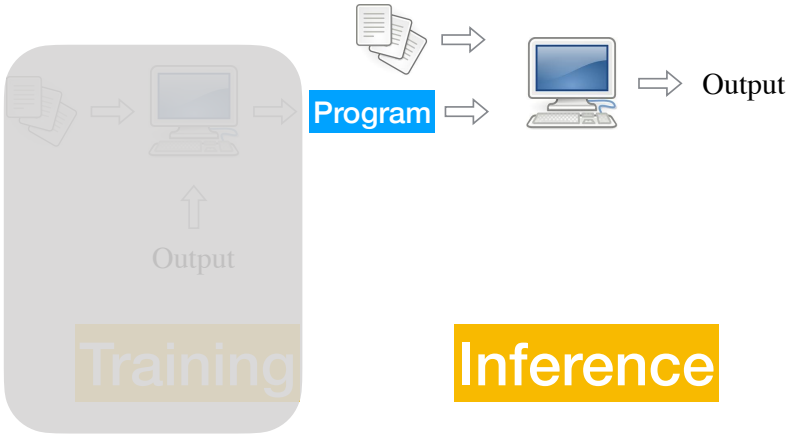
Machine Learning

(Supervised)



Machine Learning

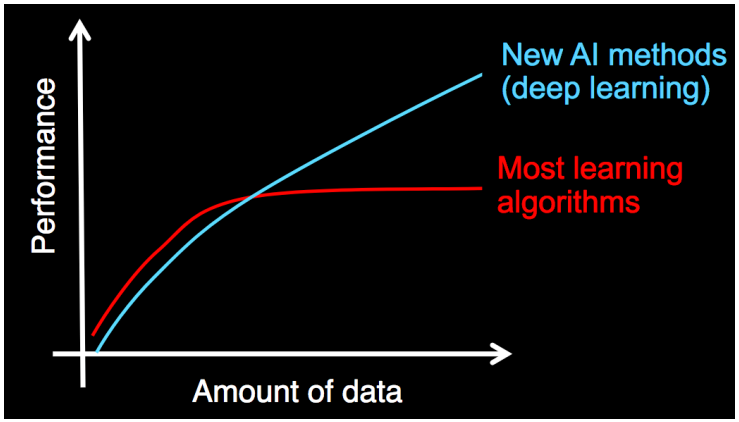
(Supervised)



Data and Machine Learning

Data is the most important part

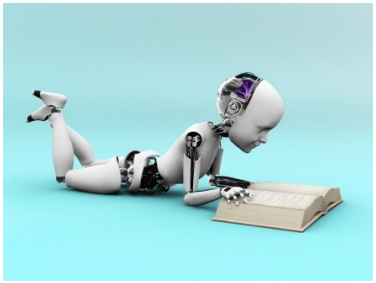
Quality & Quantity



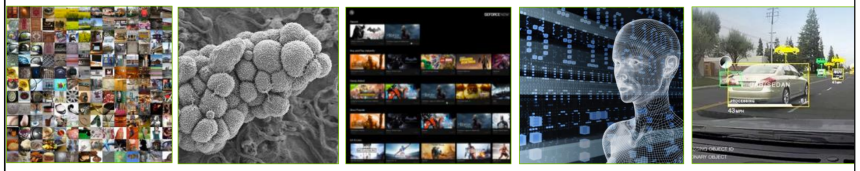
from: CS229: Machine Learning, Stanford

Why ML is so popular?

- Recent progress in algorithms and theory
- Growing amounts of data
- Computational power
- Industry investment
- Great applications



Applications …

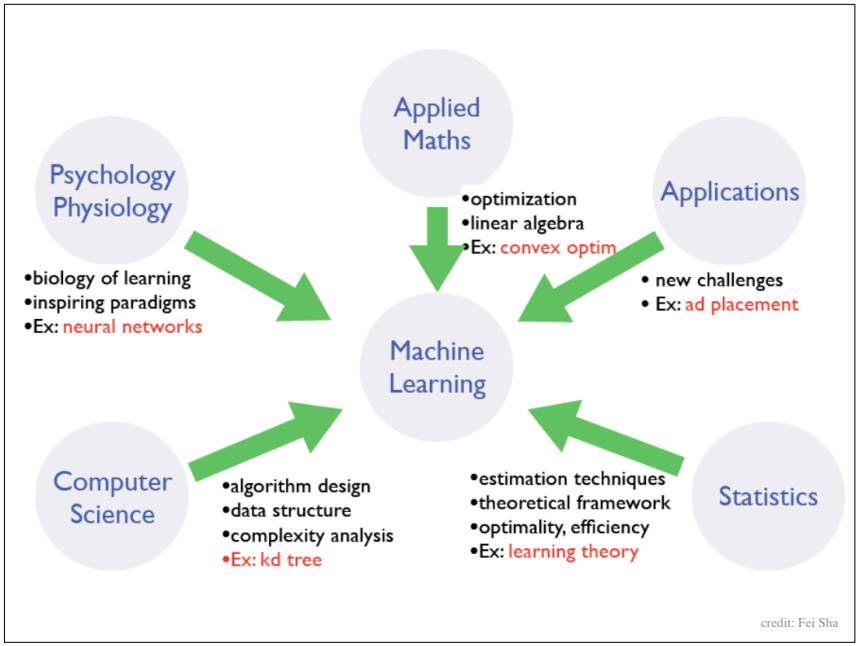


INTERNET & CLOUD	MEDICINE & BIOLOGY	MEDIA & ENTERTAINMENT	SECURITY & DEFENSE	AUTONOMOUS MACHINES
Image Classification Speech Recognition Language Translation Language Processing Sentiment Analysis Recommendation	Cancer Cell Detection Diabetic Grading Drug Discovery	Video Captioning Video Search Real Time Translation	Face Detection Video Surveillance Satellite Imagery	Pedestrian Detection Lane Tracking Recognize Traffic Sign

Deep Learning and HPC, NVIDIA, 2017

When to use Machine Learning?

- Need for automation
  - ✓ automate things humans can do (e.g. speech recognition)
  - ✓ difficult/expensive things for humans (e.g. process huge amounts of data)
- Need for custom models
  - ✓ e.g. personalized medicine, spam filters



Major Paradigms

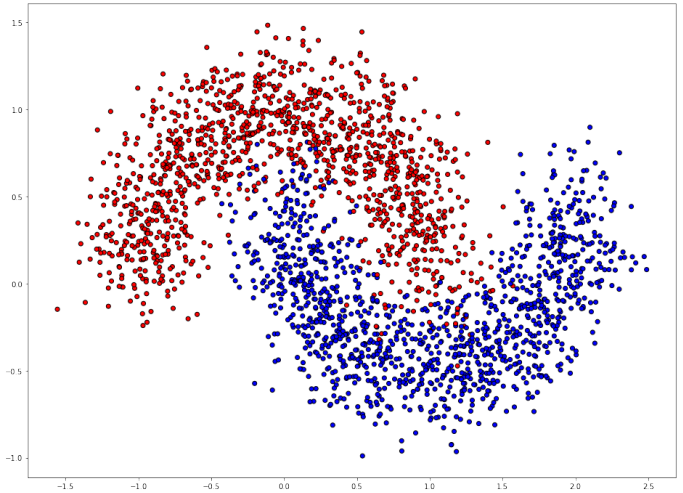
- Supervised learning
  - ✓ training data and labels
  - ✓ classification, regression
- Unsupervised learning
  - ✓ training data (no labels)
  - ✓ clustering, dimensionality reduction, density estimation
- Reinforcement learning
  - ✓ rewards from actions

Supervised Learning

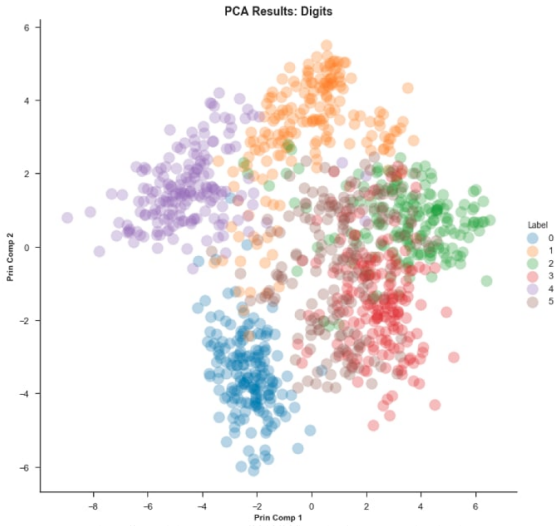
Binary classification

```
array([[ 0.24277092,  0.89098144],      array([[0],
       [-0.57961074,  0.50618765],             [1],
       [ 0.24259841,  0.12209649],             [1],
       [ 1.68348295, -0.10059047],             [1],
       [ 2.00696736, -0.79306007],             [1],
       [ 1.56891881,  0.30515286],             [0],
       [ 0.1314049 , -0.35704446],             [1],
       [ 2.14017386,  0.33933491],             [1],
       [-1.03087047,  1.52609949],             [0],
       [-0.38504321,  1.24209655],             [0],
       [-1.20252537,  0.56167652],             [0],
       [ 0.08590311,  0.68265315],             [1],
       [ 0.88074085, -0.11759523],             [1],
       [ 0.32558238,  0.4181143 ],             [1],
       [-0.74202798,  0.68847344]])           [0]])
```

Binary classification

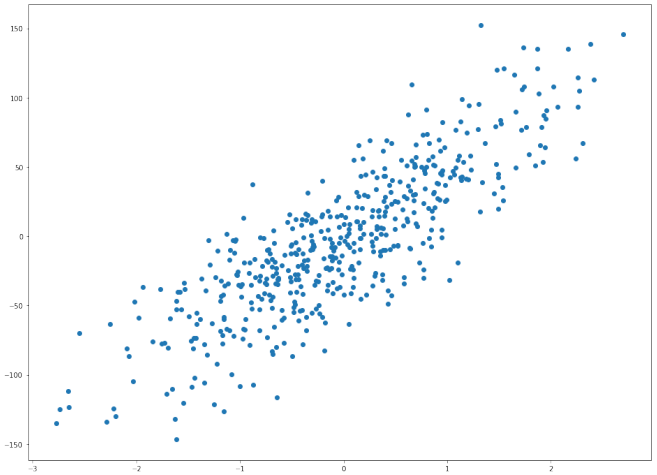


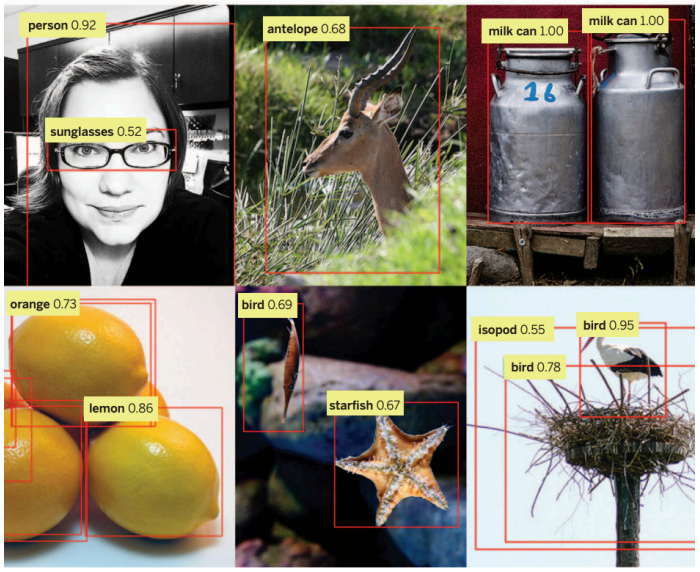
Multiclass classification



https://www.kdnuggets.com/2018/08/introduction-t-sne-python.html

Regression





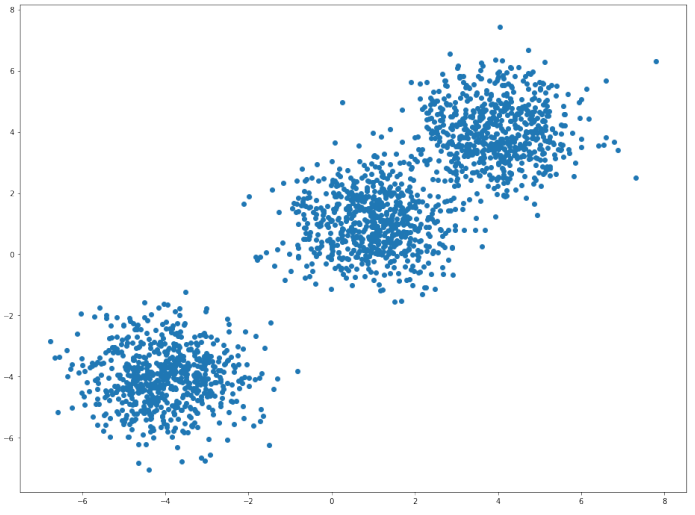
"Machine learning: Trends, perspectives, and prospects", M. I. Jordan and T. M. Mitchell

Unsupervised Learning

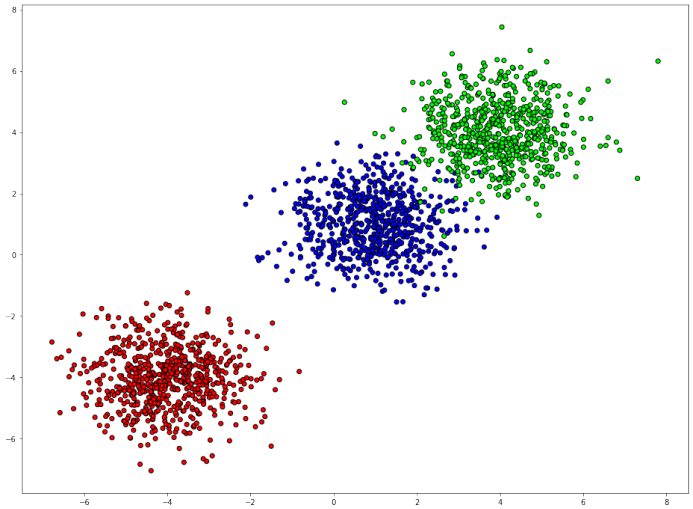
No labels

```
array([[ 0.24277092,  0.89098144],
       [-0.57961074,  0.50618765],
       [ 0.24259841,  0.12209649],
       [ 1.68348295, -0.10059047],
       [ 2.00696736, -0.79306007],
       [ 1.56891881,  0.30515286],
       [ 0.1314049 , -0.35704446],
       [ 2.14017386,  0.33933491],
       [-1.03087047,  1.52609949],
       [-0.38504321,  1.24209655],
       [-1.20252537,  0.56167652],
       [ 0.08590311,  0.68265315],
       [ 0.88074085, -0.11759523],
       [ 0.32558238,  0.4181143 ],
       [-0.74202798,  0.68847344]])
```

Unsupervised learning



Unsupervised Learning



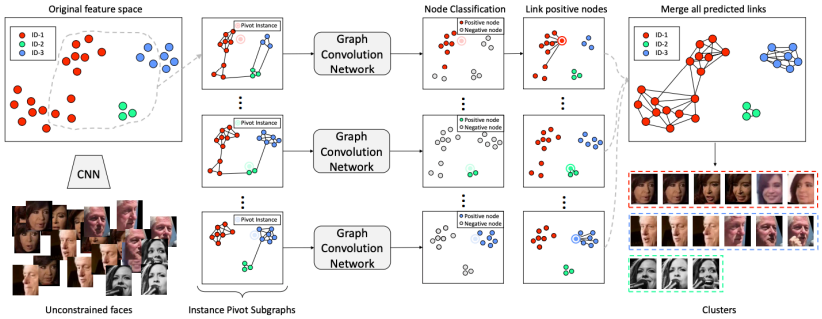


Figure 3. Framework of our method. We use every instance as a pivot, and construct an Instance Pivot Subgraph (IPS) for it. Each node in IPS models the linkage (similarity) between the pivot and the corresponding neighbor. We apply graph convolutions on IPS and classify nodes in IPS into either positive or negative. If a node is classified as positive, the corresponding neighbor should be linked to the pivot. After determining all the linkages, we transitively merge linked instances to obtain the final clusters.

"Linkage Based Face Clustering via Graph Convolution Network", CVPR 2019, Zhongdao Wang, Liang Zheng, Yali Li, Shengjin Wang

Figure 2. Uncurated set of images produced by our style-based generator (config F) with the FFHQ dataset. Here we used a variation of the truncation trick [42, 5, 34] with $\psi = 0.7$ for resolutions $4^2 - 32^2$. Please see the accompanying video for more results.

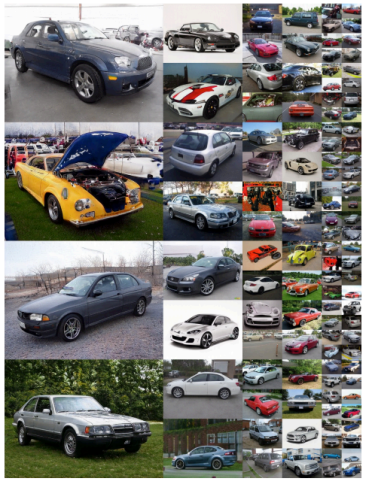


Figure 11. Uncurated set of images produced by our style-based generator (config F) with the LSUN CAR dataset at 512×384. FID computed for 50K images was 3.27.

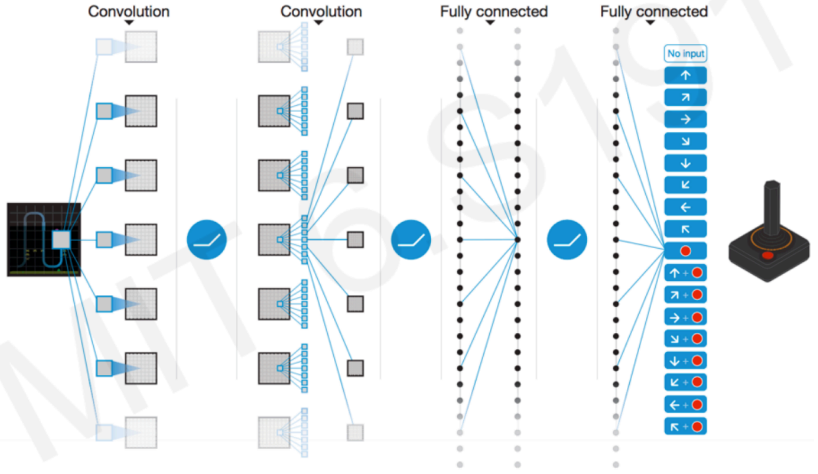
"A Style-Based Generator Architecture for Generative Adversarial Networks", CVPR 2019, Tero Karras, Samuli Laine, Timo Aila

Reinforcement Learning

Key Concepts

credit: 6.S191 Introduction to Deep Learning @ MIT

DQN Atari Results



credit: 6.S191 Introduction to Deep Learning @ MIT

IEEE SIGNAL PROCESSING MAGAZINE, SPECIAL ISSUE ON DEEP LEARNING FOR IMAGE UNDERSTANDING (ARXIV EXTENDED VERSION)