Introduction

Nan Ye

School of Mathematics and Physics
The University of Queensland

Who Are We

- Lecturer: Nan Ye, nan.ye@uq.edu.au
  - Consultation hour: 3-4pm Thu, https://uqz.zoom.us/j/82722365470
- Tutor: Jun Ju, David Maine

- Write down terms that describe what you know about machine learning or deep learning.

  https://apps.elearning.uq.edu.au/wordcloud/66830

What Should You Already Know

- What you need to keep pace with the course
  - solid foundation in linear algebra, calculus, statistics and programming.
  - this means that you need to be comfortable with your previous courses in these subjects.
  - maths can be hard to follow, but essential for us to see the subtleties in the algorithms.
- Formal prerequisites

  (STAT2004 or STAT2203 or equivalent) + programming experience (MATH2504 or CSSE2002 or equivalent)

Linear algebra

- Vectors, addition, subtraction, inner product, norm
- Matrices, transpose, addition, subtraction, multiplication, inverse
- Vector space, subspace, span
- Linear systems
- Eigen-value, eigen-vector, eigen-decomposition

Calculus

- Basic functions and their properties (monotonicity, periodicity)
- Derivative, partial derivative, gradient, Hessian
- Taylor series, first-order approximation, second order approximation
- Maximum/maxmimizer, minimum/minimizer, stationary points, first-order optimality condition

Probability and statistics

- Random variable (discrete and continuous), probability distribution, PMF/PDF/CDF
  - Common discrete distributions (Bernoulli, binomial, categorical, geometric, . . .)
  - Common continuous distributions (Uniform, Gaussian, exponential, . . .)
- Basic laws of probability
- Expectation (linearity), variance/standard deviation, correlation
- Joint distribution, marginal distribution, conditional distribution, independence
- Sample, mean, standard error
- Maximum likelihood methods
- Linear regression

Programming

- Primitive data types: number, string, boolean
- Data structures: lists, dictionary, ...
- Conditional statement, loops
- Functions
- Classes

What Are Our Objectives

Learn basic theories, algorithms and models of machine/deep learning and be able to apply them.

- Understand and explain the intuition, ideas and theory of deep learning algorithms and models.
- Assess whether a deep learning algorithm is effective and appropriate for an application.
- Propose suitable deep learning solutions and implement them for real world problems.
- Effectively explain deep learning solutions in the form of oral presentations and reports.

Have fun...

What Will We Do

Teaching activities

- Lectures, tutorials and pracs
  - on campus for internal students, on zoom for external students
  - all zoom links in Blackboard announcement
  - internal students: attend online if you're unwell
  - **follow good hygiene practices and stay well**
- Tutorial and prac start in week 1 (this week)
- Discussions on Ed
  - posting questions on Ed > emailing me
- Consultation sessions

More on teaching activities

- No textbook, but a few useful references on course profile.

  the deep learning book will be a good reference
- We will use slides

  these can be dry. you need to work out the maths after lectures
- We provide more intuition and details in lecture

  these will be more fun (I hope)

  attendance is not compulsory but encouraged
- Tutorials and pracs will cover theory and hands-on exercises
  - they complement the lectures
  - your TA will go through the exercises with you

Assessments

Assignment 1	15%	out 6 Mar, due 5pm 20 Mar
Assignment 2	15%	out 20 Mar, due 5pm 3 Apr
Assignment 3	15%	out 10 Apr, due 5pm 24 Apr
Tutorial paper	20%	due 5pm 8 May
Project		
proposal	5%	due at 5pm on 17 Apr
seminar	10%	due in the lectures in week 13
report	15%	due 5pm 1 Jun
reflective essay	5%	due 5pm on 2 Jun

Follow rules on academic integrity: http://tinyurl.com/y5fq88cf

How to Do Well in This Course

- Work through the lecture notes

  develop an intuitive understanding of the algorithms, and go through the maths behind the algorithms

- Work on the tutorial/lab exercises

  assignments will be of a similar flavor

- Ask questions and try to solve them independently

  this will be what the tutorial paper and your project are about

The Journey Begins

Artificial Intelligence (AI)
make machines intelligent

Machine Learning (ML)
make machines learn (mostly from data)

Deep Learning (DL)
a subfield of ML (mostly on deep neural networks)

Machine Learning

- Machine learning turns data into insight, predictions and/or decisions.
- Numerous applications in diverse areas, including natural language processing, computer vision, recommender systems, medical diagnosis.

A Much Sought-after Technology

0 f in

By **Alison DeNisco Rayome** in **CXO**
on February 5, 2019, 8:39 AM PST

Jobs in AI and machine learning are exploding, as countries race to develop the emerging technology, according to a UiPath report.

By Stephen Zafarino, Contributor, CIO | JULY 27, 2018 06:30 AM PT
Opinions expressed by ICN authors are their own.

OPINION

The outlook for machine learning in tech: ML and AI skills in high demand

22,034 views | Mar 17, 2019, 10:35am

Machine Learning Engineer Is The Best Job In The U.S. According To Indeed

Good News for Job Seekers With Machine Learning Skills: There is a Shortage of Talent

A short pool of AI-trained job seekers has slowed down hiring and impeded growth at some companies

Stacy Stanford in Data Driven Investor Follow
Oct 20, 2018 · 7 min read ★

- Name applications of AI/ML/DL.

  https://apps.elearning.uq.edu.au/wordcloud/66830

Applications

Siri

Cortana

Alexa

Google Assistant

AliGenie

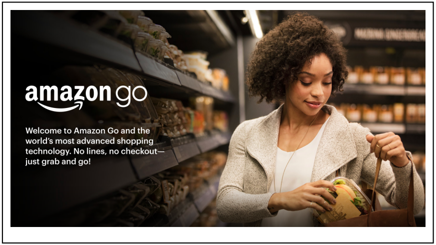
Virtual assistants

Pailitao

photo it & shop

read more

http://tinyurl.com/yxm5k93u



Amazon Go

grab-and-go shopping

read more

http://tinyurl.com/yx8vdonm

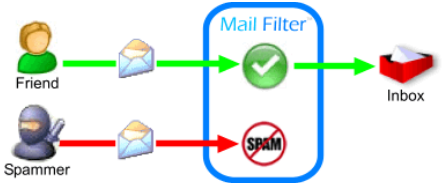
Shopping made easy

smart traffic light

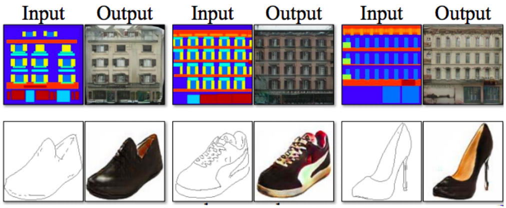
self-driving car

picking robot

Transportation and logistics



spam detection



sketch and fill

Productivity

video recommendation

computer games

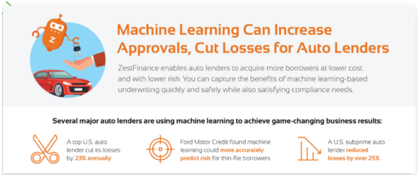
board games

Entertainment

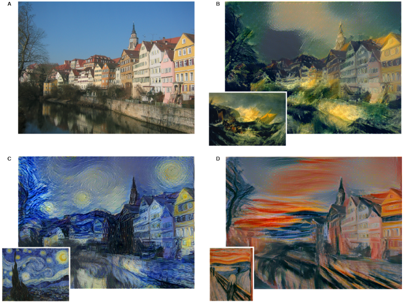
Using machine learning for insurance pricing optimization

Kaz Sato
Staff Developer Advocate,
Google Cloud

AXA, the large global insurance company, has used machine learning in a POC to optimize pricing by predicting "large-loss" traffic accidents with 78% accuracy.

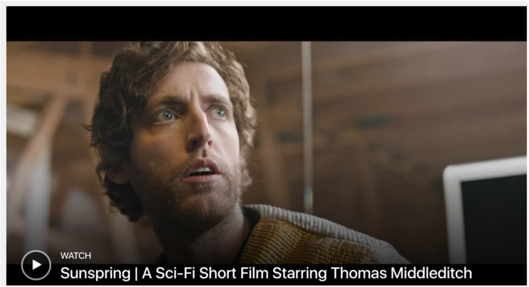


Financial services



neural artist

music composition



write movie scripts

The Machine Learning Approach

Hard to specify rules for computers to...

- recognize handwritten characters

  different people write differently
- recognize objects from images

  occlusion, viewpoint variation, change of lighting conditions...
- detect whether a credit card transaction is fraudulent

  fraud is a moving target
- and many others...

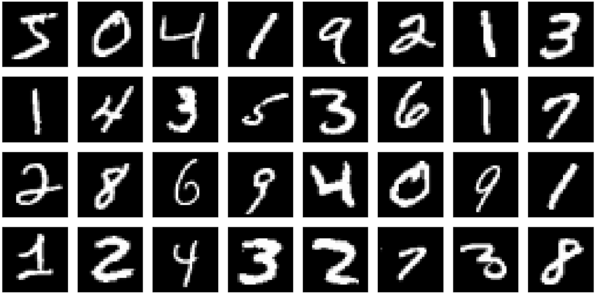
  Machine learning provides a way to solve such problems.

The machine learning approach

- Instead of writing a program by hand for a task, we collect examples that specify the correct output for a given input.
- A machine learning algorithm takes these examples and produces a program that does the job.
  - The program can be adapted to data changes by training on the new data.
  - The program works for both training data and new test data, if we do it right.
  - The program may look very different from a typical hand-written program.

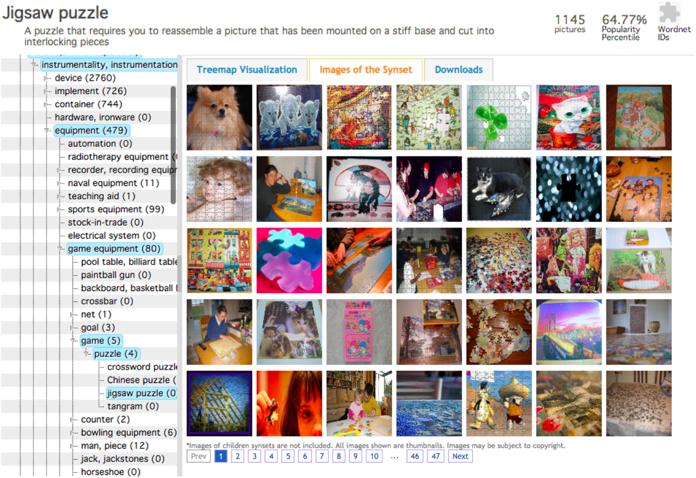
Drivers of Machine Learning

Creation of many (large) datasets



http://yann.lecun.com/exdb/mnist/

- The MNIST dataset is an early large dataset used as a benchmark for evaluating handwritten digits recognition algorithms.
- There are 60,000 labeled training images, and 10,000 labeled test images.



http://www.image-net.org/

- ImageNet is a recent large image database.
- 1000 different object classes in 1.3 million high-resolution training images from the web.

Flourishing ecosystem of software frameworks

TensorFlow

theano

mxnet

Caffe2

GLUON

Sonnet

PyTorch

DL4J

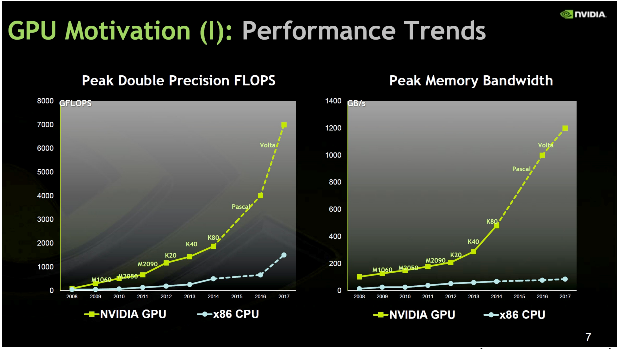
Keras

ConvNetJS
Deep Learning in your browser

SWIFT

scikit learn

Massive growth of computing power

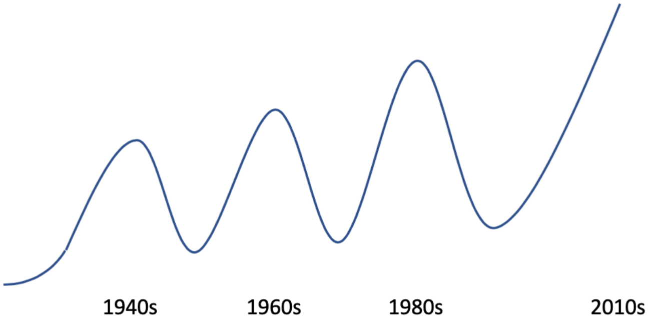


- Computing power has been doubling every few years (Moore's law).
- GPUs are highly parallel, and can provide orders-of-magnitude speedup as compared to CPUs.
- Experiments which used to take unreasonable amount of time to run could now be completed efficiently.

New algorithms

- Novel neural architectures
  - Deep CNNs, ResNet, GAN, ...
- Faster optimization algorithms
  - RMSProp, AdaDelta, Adam, ...
- Many other clever tricks/ideas...

History of Neural Networks



Research on artificial neural networks have gone through several ups and downs since its inception (figure is approximate).

Birth and some initial ideas

1943	computational model for neural networks
	McCulloch and Pitts, A logical calculus of the ideas immanent in nervous activity
1949	Hebbian learning (cells that fire together wire together)
	Hebb, The organization of behavior: A neuropsychological theory
1960	single layer and multilayer neural nets (ADALINE and MADALINE)
	Widrow and Hoff, *Adaptive switching circuits*
1962	Perceptron
	Rosenblatt, *Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms*

Further developments

1969	limitations of artificial neural nets
	Minsky and Papert, *Perceptrons; an introduction to computational geometry*
1970	modern idea of back propagation (a key trigger for renewed interest)
	Linnainmaa, The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors
1974	backpropagation
	Werbos, Beyond regression: New tools for prediction and analysis in the behavioral sciences
1980	Neocognitron (inspired Convolutional Neural Networks)
	Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position

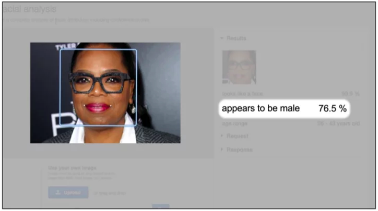
1985	Boltzmann machines
	Ackley, Hinton, and Sejnowski, A learning algorithm for Boltzmann machines
1986	restricted Boltzmann Machine (originally called Harmonium)
	Smolensky, *Information processing in dynamical systems: Foundations of harmony theory*
1989	universal approximation (sigmoid functions)
	Gybenko, Approximation by superposition of sigmoidal functions
1989	convolutional neural net (LeNet)
	LeCun, Generalization and network design strategies
1991	universal approximation (general functions)
	Hornik, Approximation capabilities of multilayer feedforward networks

Deep learning

2006	deep belief net	
	Hinton, Osindero, and Teh, A fast learning algorithm for deep belief nets	
2009	deep Boltzmann machines	
	Salakhutdinov and Hinton, Deep boltzmann machines	
2012	dropout	
	Hinton, Srivastava, Krizhevsky, Sutskever, and Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors	
2014	generative adversarial networks	
	Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, and Bengio, Generative adversarial nets	
2015	deep residual network	
	He, Zhang, Ren, and Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification	
2017	Alpha-Go Zero	
	Silver, Schrittwieser, Simonyan, Antonoglou, Huang, Guez, Hubert, Baker, Lai, Bolton, et al., Mastering the game of go without human knowledge	

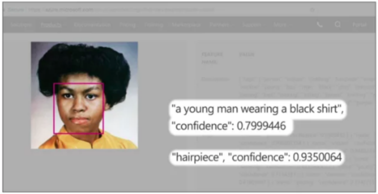
Prejudiced?

Oprah Winfrey



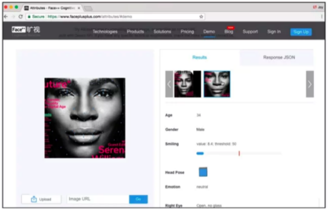
amazon

Michelle Obama



Microsoft

Serena Williams



Face++ 旷视

nearly perfect gender recognition for lighter-skinned men, but high error for darker-skinned women

http://tinyurl.com/y3xxuye9

BUSINESS NEWS OCTOBER 10, 2018 / 1:12 PM / 9 MONTHS AGO

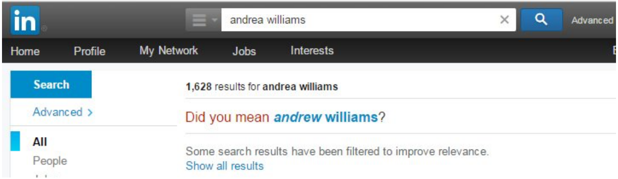
Amazon scraps secret AI recruiting tool that showed bias against women

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016



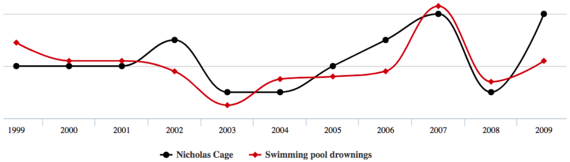
Nature or Nurture?

Bias in, bias out

An algorithm can become biased for various reasons

- A problem may be improperly formulated

  e.g. # pool drownings vs. # of Nicolas Cage films.



- Training data may be biased towards certain groups

  e.g. deep learning requires large amount of data, which is often unavailable for minority groups.

- The algorithm makes assumptions (inductive bias)

Bias reinforcement

When a biased algorithm is used to implement policies, the biased action will lead to the collection of more biased data, creating a feedback loop that can exacerbate bias.

Machine Learning is Alchemy?

Read more: http://tinyurl.com/yae4jpqy

We need to understand what a machine learning algorithm is doing!

It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way—in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or for evil, in the superlative degree of comparison only.

Charles Dickens, A Tale of Two Cities

Schedule

A tentative schedule is available on BlackBoard

- Week 1-2: machine learning basics
- Week 3-4: neural network basics
- Week 5-6: deep architectures
- Week 7-8: optimization
- Week 9-10: improving generalization
- Week 10-11: unsupervised learning
- Week 12: reinforcement learning