data, ppf 2021-03-23

"AI 2.0 and 'machine learning' "

chris wiggins and matt jones

outline for today

outline for today

outline for today

- 1. the term “machine learning” and the field Machine Learning

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community
  - 1.1 rise of ML, 1980s

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes “pattern recognition”

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes “pattern recognition”
    - 1.2.2 which subsumed the work of Wald

outline for today

- 1. the term "machine learning" and the field Machine Learning
  - 1.0 backdrop: "pattern recognition" community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes "pattern recognition"
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning

outline for today

- 1. the term "machine learning" and the field Machine Learning
  - 1.0 backdrop: "pattern recognition" community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes "pattern recognition"
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning

outline for today

- 1. the term "machine learning" and the field Machine Learning
  - 1.0 backdrop: "pattern recognition" community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes "pattern recognition"
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning
- 2. AI 2.0: "deep learning"

outline for today

- 1. the term "machine learning" and the field Machine Learning
  - 1.0 backdrop: "pattern recognition" community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes "pattern recognition"
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning
- 2. AI 2.0: "deep learning"
  - AI subsumes "ML"

outline for today

- 1. the term "machine learning" and the field Machine Learning
  - 1.0 backdrop: "pattern recognition" community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes "pattern recognition"
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning
- 2. AI 2.0: "deep learning"
  - AI subsumes "ML"

outline for today

- 1. the term “machine learning” and the field Machine Learning
  - 1.0 backdrop: “pattern recognition” community
  - 1.1 rise of ML, 1980s
  - 1.2 changing nature of of ML, 2000s
    - 1.2.1 subsumes “pattern recognition”
    - 1.2.2 which subsumed the work of Wald
  - 1.3 key ideas in Machine Learning
- 2. AI 2.0: “deep learning”
  - AI subsumes “ML”
- 3. what was gained, what was lost

student reactions: words

```
68 learning
67 machine
43 ai
40 rudin
29 intelligence
28 brain
27 translate
27 lewis
22 artificial
30 interpretable/interpretability
16 bias
 9 jordan
 8 pandemic
 5 parole
 3 women
```

student reactions: ideas

- “[AI] un-familiarizes us”

student reactions: ideas

- “[AI] un-familiarizes us”
- interpretability: “it seems weird that accuracy would impede the interpretability of a model. Couldn’t a model be interpreted based on its underlying code?”

student reactions: ideas

- "[AI] un-familiarizes us"
- interpretability: "it seems weird that accuracy would impede the interpretability of a model. Couldn't a model be interpreted based on its underlying code?"
- prescience around pandemic

student reactions: ideas

- "[AI] un-familiarizes us"
- interpretability: "it seems weird that accuracy would impede the interpretability of a model. Couldn't a model be interpreted based on its underlying code?"
- prescience around pandemic
  - "Strangely, the piece (written in 2015) closes with an example of how the availability of location data and online medical data may be beneficial during a pandemic, and a consequential"tension and trade-off between personal privacy and public health"

student reactions: ideas

- “[AI] un-familiarizes us”
- interpretability: “it seems weird that accuracy would impede the interpretability of a model. Couldn’t a model be interpreted based on its underlying code?”
- prescience around pandemic
  - “Strangely, the piece (written in 2015) closes with an example of how the availability of location data and online medical data may be beneficial during a pandemic, and a consequential”tension and trade-off between personal privacy and public health"
- utopia vs dystopia

student reactions: ideas

- “[AI] un-familiarizes us”
- interpretability: “it seems weird that accuracy would impede the interpretability of a model. Couldn’t a model be interpreted based on its underlying code?”
- prescience around pandemic
  - “Strangely, the piece (written in 2015) closes with an example of how the availability of location data and online medical data may be beneficial during a pandemic, and a consequential”tension and trade-off between personal privacy and public health"
- utopia vs dystopia
  - “Aspects of Gideon-Lewis Kraus’s The Great AI Awakening made it difficult for me to accept Boyd and Crawford’s advice in our first week’s readings to view data and technology as things which are not necessarily bad (or good or neutral, as they say.)”

student reactions: ideas

- "[AI] un-familiarizes us"
- interpretability: "it seems weird that accuracy would impede the interpretability of a model. Couldn't a model be interpreted based on its underlying code?"
- prescience around pandemic
  - "Strangely, the piece (written in 2015) closes with an example of how the availability of location data and online medical data may be beneficial during a pandemic, and a consequential"tension and trade-off between personal privacy and public health"
- utopia vs dystopia
  - "Aspects of Gideon-Lewis Kraus's The Great AI Awakening made it difficult for me to accept Boyd and Crawford's advice in our first week's readings to view data and technology as things which are not necessarily bad (or good or neutral, as they say.)"
  - "I honestly see why this is terrifying to so many people, but it genuinely just makes me feel grateful for how much more accessible this technology has gotten."

who (vocabulary, names of “research” fields)

- deep learning (2010-present)

who (vocabulary, names of “research” fields)

- deep learning (2010-present)
- machine learning (b. 1959, 1980s-)

who (vocabulary, names of "research" fields)

- deep learning (2010-present)
- machine learning (b. 1959, 1980s-)
- artificial intelligence (b. 1956, 1956-)

who (vocabulary, names of “research” fields)

- deep learning (2010-present)
- machine learning (b. 1959, 1980s-)
- artificial intelligence (b. 1956, 1956-)
- pattern recognition (b. 1955)

who (vocabulary, names of "research" fields)

- deep learning (2010-present)
- machine learning (b. 1959, 1980s-)
- artificial intelligence (b. 1956, 1956-)
- pattern recognition (b. 1955)
- NB: next week is "data science", the expansion of these beyond "research" communities

what: ideas, aspirations, and methods to look out for today

- declining centrality of “intelligence”

what: ideas, aspirations, and methods to look out for today

- declining centrality of “intelligence”
- rising challenges to “interpretability”

what: ideas, aspirations, and methods to look out for today

- declining centrality of “intelligence”
- rising challenges to “interpretability”
- absence of “statistics”

what: ideas, aspirations, and methods to look out for today

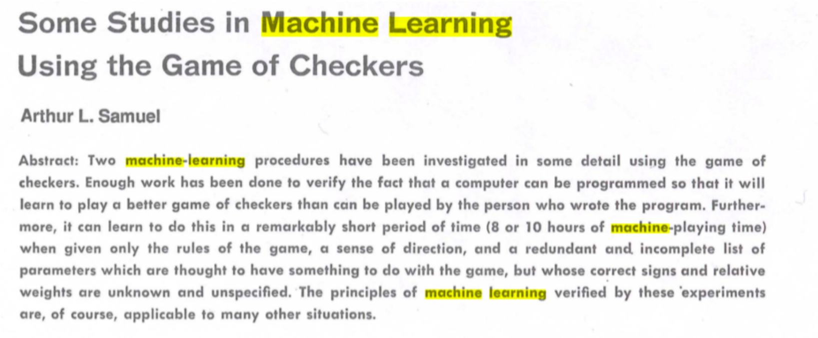
- declining centrality of “intelligence”
- rising challenges to “interpretability”
- absence of “statistics”
- overlap of fields: artificial intelligence, machine learning, pattern recognition

what: ideas, aspirations, and methods to look out for today

- declining centrality of "intelligence"
- rising challenges to "interpretability"
- absence of "statistics"
- overlap of fields: artificial intelligence, machine learning, pattern recognition
- persistent tension: statistical inference vs algorithms

1. “machine learning” and Machine Learning

"machine learning" (the term) traced to checkers @ IBM (1959)



Some Studies in Machine Learning Using the Game of Checkers

Arthur L. Samuel

Abstract: Two machine-learning procedures have been investigated in some detail using the game of checkers. Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. Furthermore, it can learn to do this in a remarkably short period of time (8 or 10 hours of machine-playing time) when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters which are thought to have something to do with the game, but whose correct signs and relative weights are unknown and unspecified. The principles of machine learning verified by these experiments are, of course, applicable to many other situations.

Figure 1: Samuel, Arthur L. "Some studies in machine learning using the game of checkers." IBM Journal of research and development 3, no. 3 (1959): 210-229.

(unclear if that's "1st usage", e.g., in 1955)

It may now be worthwhile to talk of a ***machine learning*** *process as one in which the machine adapts its code of instructions to conform to its experience. Normally, the word "decision has been used to describe the action by which a machine changes from one program to another and learning has often been discussed as the adjustment of a particular control parameter, by trial and improvement methods. Since in a numerical machine, the trial alteration of any instruction will tend to alter the entire process completely, it is not possible to gradually modify and improve code.*

unlike “AI” (1956), “ML” 1959 didn’t really take off, e.g.,

Figure 2: ANN embraced by “pattern recognition” community (EE), not AI (CS)

1.0 “pattern recognition”, stats done by EE

“Pattern recognition and modern computers” Oliver G Selfridge (also attended D’56, NSA Advisory Board)

AFIPS ’55 (Western): Proceedings of the March 1-3, 1955, western joint computer conference March 1955 Pages 91–93 https://doi.org/10.1145/1455292.1455310

1.0 Prehistory: Pattern Recog 1968

"Pattern recognition is hardly a cohesive discipline in its own right. . . ."

"The performance of the machine is then evaluated on a test set comprising samples not included in the training set." - Nagy, student of Frank Rosenblatt (perceptron 1959)

- like AI and early ML, the aspiration was (at least initially) implemented via whatever worked: ANNs, statistics, etc

1.0 Prehistory: Pattern Recog 1972

It was called pattern recognition then; it's called machine learning now.

- Jerry Friedman, interviewed in 2015, describing work done in 1972

1.0 Similar to Pattern Recog, '73

Figure 3: Duda Hart 1973

1.0 Similar to Pattern Recog, ’73

6 UNSUPERVISED LEARNING AND CLUSTERING 189

Figure 4: Duda Hart 1973

1.0 PrePrehistory: PR and “decision statistics” of Wald

Statistical Signal Detection theory was adopted in the pattern recognition community at a very early stage. Chow (1957) proposed using optimal detection theory, and this led to a proposal by Highleyman to approximate the risk by its sample average. This transition from population risk to “empirical” risk gave rise to what we know today as machine learning.

- mlstory.org, 2021

see also Mendon-Plasek, Aaron. “Mechanized Significance and Machine Learning: Why It Became Thinkable and Preferable to Teach Machines to Judge the World.” In The Cultural Life of Machine Learning, pp. 31-78. Palgrave Macmillan, Cham, 2021.

1.0 sidebar: Pattern Recog & .mil: Philco, SRI, etc.

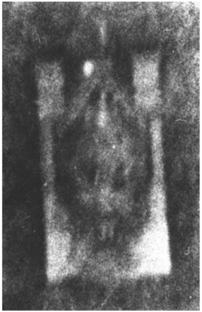


Figure 4.13: A typical tank image. (Photograph courtesy of Thomas Harley.)

Figure 5: From Nils Nillson's "QAI"; see also "Siri"

1.0 sidebar: Pattern Recog & .mil: "Shakey" 1966-1972

Funded by the Pentagon DARPA, it was originally modeled on the idea of an autonomous sentry for a military base. (Markoff, in NYT; see also Markoff book 2015)

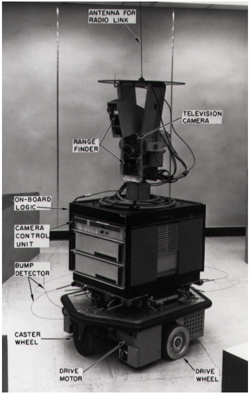


Figure 6: Shakey the robot, 1972

1.1 rise of ML, 1980s: recall pre-1980 view on AI

“One vision was that AI means to”take symbolic information as input, manipulate it according to a set of formal rules, and in so doing can solve problems. . . "

“After the 1956 workshop, this became the dominant approach *among academics*”

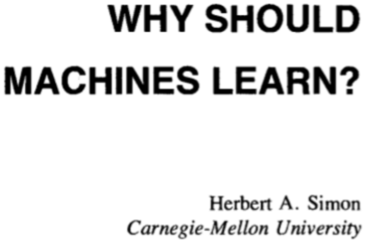
- Stephanie Dick essay, 2019

1.1 rise of ML, 1980s: 1980 meeting on ML

At the Carnegie-Mellon Machine Learning Workshop in July, 1980, Herbert Simon was asked to deliver the keynote address, where he chose to play the role of devil's advocate and ask the question "Why Should Machines Learn?" His analysis concluded that, with the exception of cognitive modeling, some rethinking of long-term objectives was in order. After dispelling some common myths, Simon concluded with a clarified and more appropriate set of reasons why one ought to pursue machine learning research. Chapter 2 is based almost entirely on that rather controversial keynote address.

Context: Simon

- attended Dartmouth '56



WHY SHOULD
MACHINES LEARN?

Herbert A. Simon
Carnegie-Mellon University

Figure 7: Simon

Context: Simon

- attended Dartmouth ’56
- Econ Nobel ’78

WHY SHOULD
MACHINES LEARN?

Herbert A. Simon
Carnegie-Mellon University

Figure 7: Simon

Context: Simon

- attended Dartmouth '56
- Econ Nobel '78
- Turing award '75

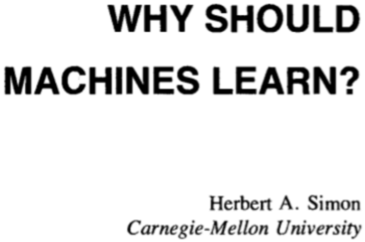
WHY SHOULD MACHINES LEARN?

Herbert A. Simon
Carnegie-Mellon University

Figure 7: Simon

Context: Simon

- attended Dartmouth '56
- Econ Nobel '78
- Turing award '75
- prof of psych and CS at CMU


WHY SHOULD MACHINES LEARN?

Herbert A. Simon
Carnegie-Mellon University

Figure 7: Simon

Simon 1980/1983

"skeptical challenge to learning as the road to the future in AI . . . (sometimes called cognitive simulation, or information processing psychology). . .

- *case against AI research in learning"*

the most important kinds of learning research to carry out in AI are those that are oriented toward understanding human learning. Here as elsewhere, Man seems to be the measure of all things.

Simon 1980/1983

"skeptical challenge to learning as the road to the future in AI . . . (sometimes called cognitive simulation, or information processing psychology). . .

- *case against AI research in learning"*
- *the most important kinds of learning research to carry out in AI are those that are oriented toward understanding human learning. Here as elsewhere, Man seems to be the measure of all things.*

Simon'83: tension

- "Why Machine Learning? .... just program it"

vs.

> *"If we understand the domain ourselves, if we understand physics, why don't we just choose an internal representation and provide the problems to the system in that internal representation? What's all this learning and natural language understanding about?"*

Simon'83: tension

- "Why Machine Learning? . . . . just program it"
- "just sat down and programmed it"

vs.

> *"If we understand the domain ourselves, if we understand physics, why don't we just choose an internal representation and provide the problems to the system in that internal representation? What's all this learning and natural language understanding about?"*

Simon'83: tension

- "Why Machine Learning? .... just program it"
- "just sat down and programmed it"

vs.

- Schema (13x) & representations (16x)
  "If we understand the domain ourselves, if we understand physics, why don't we just choose an internal representation and provide the problems to the system in that internal representation? What's all this learning and natural language understanding about?"

1983 Langley on journal of ML: field-making

> *Historically, researchers have taken two approaches to machine learning. Numerical methods such as discriminant analysis have proven quite useful in perceptual domains, and have become associated with the paradigm known as Pattern Recognition. In contrast, Artificial Intelligence researchers have concentrated on symbolic learning methods*

- "parent fields": artificial intelligence and cognitive science

1983 Langley on journal of ML: field-making

> *Historically, researchers have taken two approaches to machine learning. Numerical methods such as discriminant analysis have proven quite useful in perceptual domains, and have become associated with the paradigm known as Pattern Recognition. In contrast, Artificial Intelligence researchers have concentrated on symbolic learning methods*

- "parent fields": artificial intelligence and cognitive science
- missing: stats

sidebar: JML vs JMLR, 2000

The journal was established as an open-access alternative to the journal Machine Learning. In 2001, forty editorial board members of Machine Learning resigned, saying that in the era of the Internet, it was detrimental for researchers to continue publishing their papers in expensive journals with pay-access archives. The open access model employed by the Journal of Machine Learning Research allows authors to publish articles for free and retain copyright, while archives are freely available online.

1997 Mitchell on ML

"Machine Learning is the study of computer algorithms that improve automatically through experience."

From http://www.cs.cmu.edu/~tom/mlbook.html

2006 Mitchell on ML “Department of ML” (CMU)

"How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes? – Mitchell

1.2 MIJ on the changing nature of ML, 2000s

MIJ, summer school talk 2009: "[Machine learning is just statistics]"

MIJ, AMA 2014:

> *Throughout the eighties and nineties, it was striking how many times people working within the "ML community" realized that their ideas had had a lengthy pre-history in statistics. while ideas such as Kalman filters, HMMs and factor analysis originated outside of the "statistics community" narrowly defined, there were absorbed within statistics because they're clearly about inference. Similarly, layered neural networks can and should be viewed as nonparametric function estimators, objects to be analyzed statistically. Are the SVM and boosting machine learning while logistic regression is statistics, even though they're solving essentially the same optimization problems up to slightly different shapes in a loss function? Why does anyone think that these are meaningful distinctions?*

1.2 changing nature of ML, 2000s

Langley 2011 on ML and what was lost (1/n)

early [ML] emphasis on 'symbolic' representations of learned knowledge, such as production rules, decision trees, and logical formulae. . . discouraged papers on neural networks and other 'nonsymbolic' approaches. . . .the machine learning movement was an outgrowth of symbolic AI and cognitive science, with most of these researchers being concerned with automatically constructing expert systems or modeling human acquisition of knowledge structures.

Langley 2011 on ML (2/n)

by the late 1980s the journal had begun to consider and publish papers on other approaches to learning, some of them incorporating ideas from the pattern-recognition community, which had been exploring the topic for over two decades.

Langley 2011 on ML (3/n)

shift away from the field's original concern with learning in the context of intelligent systems

- declining centrality of "intelligence"; rise of statistical methods via "pattern recognition" community

Langley 2011 on ML (4/n)

> *methods borrowed from the pattern-recognition community lent themselves to these simpler tasks [classification, regression]*

- declining centrality of "intelligence"; rise of statistical methods via "pattern recognition" community

Langley 2011 on ML (5/n)

> *In 1986, when we launched the journal, machine learning was still viewed as a branch of artificial intelligence. By 2000, many researchers committed to machine learning treated it as a separate field with few links to its parent discipline.*

- declining centrality of "intelligence"; rise of statistical methods via "pattern recognition" community

Langley 2011 on ML (6/n)

> *There are now active PhD-level researchers who have never taken a course in artificial intelligence and who see no reason why they should, as their interests lie in completely different areas.*

- declining centrality of "intelligence"; rise of statistical methods via "pattern recognition" community

Langley 2011 on ML (7/n) “performance” vs “knowledge”

During the 1990s . . . [there was a] growing use of statistical and pattern-recognition approaches, which improved performance but which did not produce knowledge in any generally recognized form.

Langley 2011 on ML (8/n): mathematatization, yet...

> *there has been increased emphasis on mathematical formalization and a bias against papers that do not include such treatments.*

- recall similar fate in stats which became "mathematical statistics" w/in NSF's DMS

Langley 2011 on ML (9/n) . . . performance papers

> *most studies involving 'bake offs', that is, mindless comparisons among the performance of algorithms that reveal little about the sources of power or the effects of domain characteristics*

- performance-centrism

Langley 2011 on ML (10/n)

> *Machine learning was originally concerned with developing intelligent systems that exhibited rich behavior on complex tasks, while many modern researchers seem content to tackle problems that do not require either intelligence or systems. Machine learning focused initially on using and acquiring knowledge cast as rich relational structures, while many researchers now appear to care only about statistics.*

who put "statistics" in my "intelligent systems"?

Happy synthesis: Jordan-Mitchell 2015

"Machine learning is a discipline focused on two interrelated questions:

1. How can one construct computer systems that automatically improve through experience? and

Happy synthesis: Jordan-Mitchell 2015

"Machine learning is a discipline focused on two interrelated questions:

1. How can one construct computer systems that automatically improve through experience? and
2. What are the fundamental statistical-computational-information-theoretic laws that govern all learning systems, including computers, humans, and organizations? "

J-M 2015:

"machine-learning algorithms can be viewed as searching through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric." 255

1) "large space of candidate programs" (e.g. different decision trees for classifying)

J-M 2015:

“machine-learning algorithms can be viewed as searching through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric.” 255

1) “large space of candidate programs” (e.g. different decision trees for classifying)
2) “training experience” (e.g. human classified data)

J-M 2015:

"machine-learning algorithms can be viewed as searching through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric." 255

1) "large space of candidate programs" (e.g. different decision trees for classifying)
2) "training experience" (e.g. human classified data)
3) "metric" = what's most important to you (e.g. accuracy)

J-M 2015:

Within artificial intelligence machine learning has emerged as the method of choice... it can be far easier to train a system... than to program it

1.3 key ideas in Machine Learning

- "three kinds of ML"

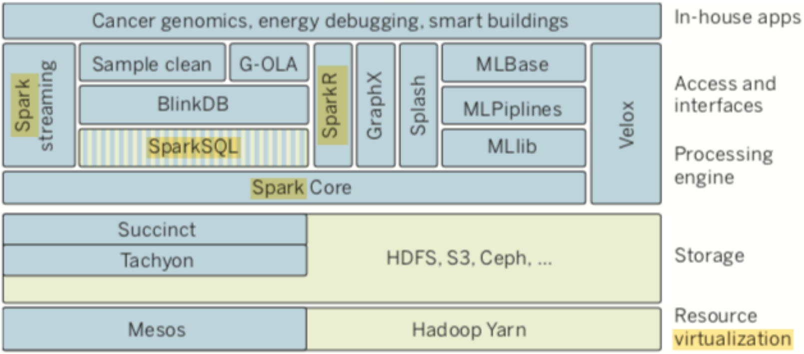


Figure 8: AMP stack

1.3 key ideas in Machine Learning

- "three kinds of ML"
- prediction culture

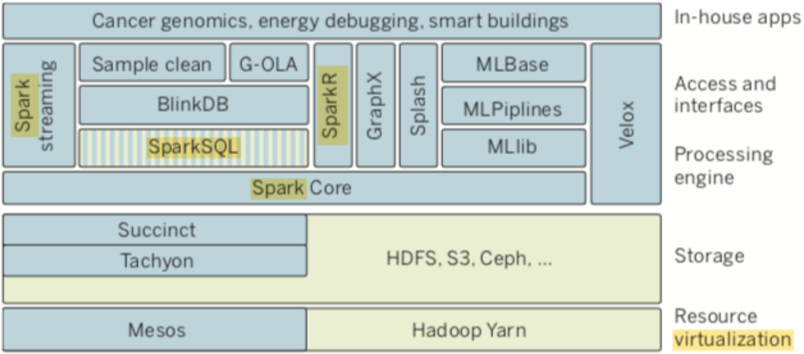


Figure 8: AMP stack

1.3 key ideas in Machine Learning

- "three kinds of ML"
- prediction culture
- math $\in$ tech $\in$ socio-technical systems (including ADS)

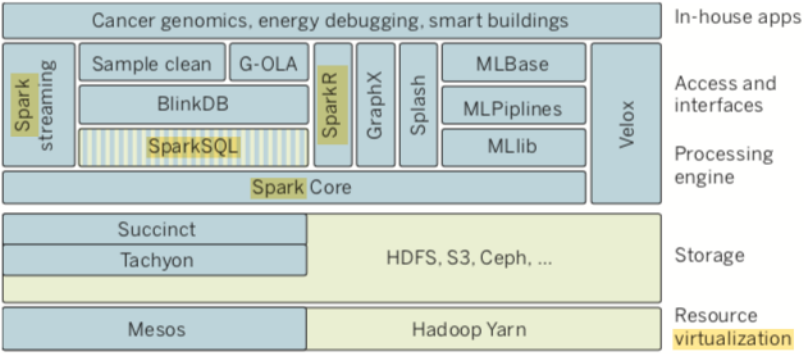


Figure 8: AMP stack

1.3 key ideas in Machine Learning

- "three kinds of ML"
- prediction culture
- math $\in$ tech $\in$ socio-technical systems (including ADS)
- impact

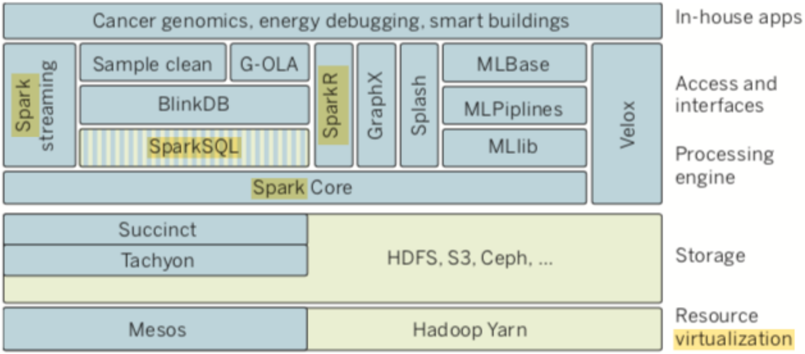


Figure 8: AMP stack

hints of ethics

- "Granular, personalized. . . data"

hints of ethics

- “Granular, personalized... data”
- “minimize privacy effects”

hints of ethics

- “Granular, personalized. . . data”
- “minimize privacy effects”
- “differential privacy”

hints of ethics

- "Granular, personalized... data"
- "minimize privacy effects"
- "differential privacy"
- "social, legal, political framework surrounding the deployment of a a system... cooperative or adversarial"

J-M on ethics

As with any powerful technology, machine learning raises questions about which of its potential uses society should encourage and discourage. The push in recent years to collect new kinds of personal data, motivated by its economic value, leads to obvious privacy issues, as mentioned above. The increasing value of data also raises a second ethical issue: Who will have access to, and ownership of, online data, and who will reap its benefits? Currently, much data are collected by corporations for specific uses leading to improved profits, with little or no motive for data sharing. However, the potential benefits that society could realize, even from existing online data, would be considerable if those data were to be made available for public good.

1.3 key ideas in Machine Learning,

- ML prediction culture, recall

1.3 key ideas in Machine Learning,

- ML prediction culture, recall
- “Overwhelmingly, machine learning systems are oriented towards one specific task: to make accurate predictions.” – SD2019

1.3 key ideas in Machine Learning,

- J-M: 3 paradigms

roughly: predict, describe, prescribe

1.3 key ideas in Machine Learning,

- J-M: 3 paradigms
- Supervised (main topic)

roughly: predict, describe, prescribe

1.3 key ideas in Machine Learning,

- J-M: 3 paradigms
- Supervised (main topic)
- Unsupervised

roughly: predict, describe, prescribe

1.3 key ideas in Machine Learning,

- J-M: 3 paradigms
- Supervised (main topic)
- Unsupervised
- Reinforcement

roughly: predict, describe, prescribe

history in memes

history in memes

history in memes

history in memes

history in memes

caveat: deep learning is not 1940s stats, which brings us to. . . .

2. "AI 2.0" the re-branding

2. “AI 2.0” the re-branding

increasing research, PR and popular attention e.g.,

- 2009 – Launch of ImageNet

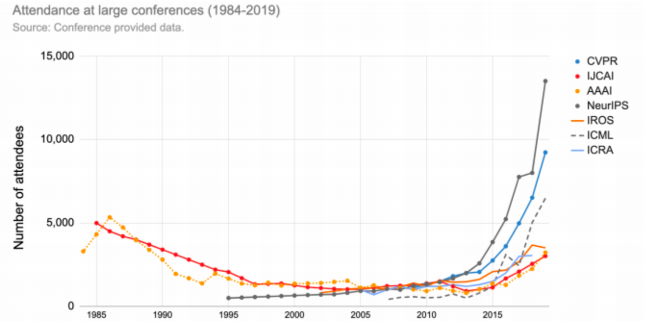


Fig. 2.1a

Note: IJCAI occurred every other year till 2014. The missing year between 1984 and 2014 are interpolated as the mean between the two known conference attendance dates to provide a comparative view across conferences.

2. “AI 2.0” the re-branding

increasing research, PR and popular attention e.g.,

- 2009 – Launch of ImageNet
- 2011 – Creation of AlexNet

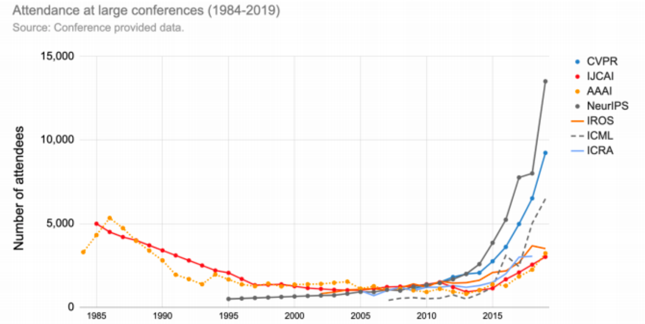


Fig. 2.1a

Note: IJCAI occurred every other year till 2014. The missing year between 1984 and 2014 are interpolated as the mean between the two known conference attendance dates to provide a comparative view across conferences.

2. “AI 2.0” the re-branding

increasing research, PR and popular attention e.g.,

- 2009 – Launch of ImageNet
- 2011 – Creation of AlexNet
  - error rate was %15.3; second closest was %26.2.

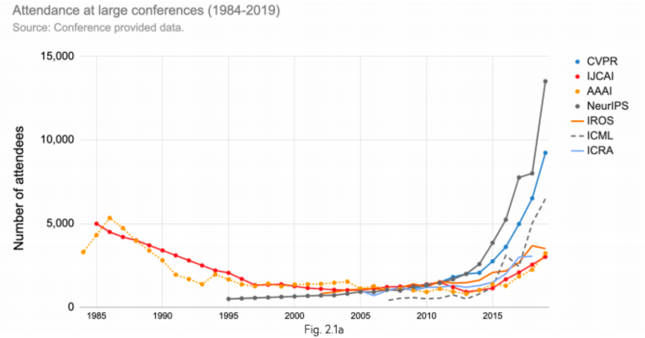


Fig. 2.1a

Note: IJCAI occurred every other year till 2014. The missing year between 1984 and 2014 are interpolated as the mean between the two known conference attendance dates to provide a comparative view across conferences.

2. "AI 2.0" the re-branding

increasing research, PR and popular attention e.g.,

- 2009 – Launch of ImageNet
- 2011 – Creation of AlexNet
  - error rate was %15.3; second closest was %26.2.
- 2012 – The Cat Experiment

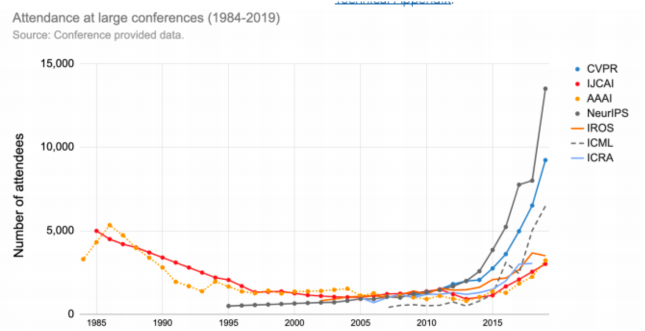


Fig. 2.1a

Note: IJCAI occurred every other year till 2014. The missing year between 1984 and 2014 are interpolated as the mean between the two known conference attendance dates to provide a comparative view across conferences.

From Krauss 2016 NYT:

- Why great awakening?

From Krauss 2016 NYT:

- Why great awakening?
- ML much bigger than MT, but it's a good story

From Krauss 2016 NYT:

- Why great awakening?
- ML much bigger than MT, but it's a good story
  - e.g., common task framework

From Krauss 2016 NYT:

- Why great awakening?
- ML much bigger than MT, but it's a good story
  - e.g., common task framework
  - human-interpretable

From Krauss 2016 NYT:

- Why great awakening?
- ML much bigger than MT, but it's a good story
  - e.g., common task framework
  - human-interpretable
  - actual "AI", i.e., recall the "N" of NLP

From Krauss 2016 NYT:

Google's decision to reorganize itself around A.I. was the first major manifestation of what has become an industry wide machine-learning delirium. Over the past four years, six companies in particular — Google, Facebook, Apple, Amazon, Microsoft and the Chinese firm Baidu — have touched off an arms race for A.I. talent, particularly within universities. Corporate promises of resources and freedom have thinned out top academic departments. It has become widely known in Silicon Valley that Mark Zuckerberg, chief executive of Facebook, personally oversees, with phone calls and video-chat blandishments, his company's overtures to the most desirable graduate students. Starting salaries of seven figures are not unheard-of. Attendance at the field's most important academic conference has nearly quadrupled.

From Krauss 2016 NYT on alchemy

> *Some of the stuff was not done in full consciousness. They didn't know themselves why they worked.*

cf., Ali Rahimi and Ben Recht, 2017

From Krauss 2016 NYT: Big Data

These ideas remained popular, however, among philosophers and psychologists, who called it "connectionism" or "parallel distributed processing." "This idea," Hinton told me, "of a few people keeping a torch burning, it's a nice myth. It was true within artificial intelligence. But within psychology lots of people believed in the approach but just couldn't do it.' Neither could Hinton, despite the generosity of the Canadian government."There just wasn't enough computer power or enough data. People on our side kept saying, "Yeah, but if I had a really big [computer], it would work." "It wasn't a very persuasive argument."

From Krauss 2016 NYT: Big Compute

There was, however, another option: just design, mass-produce and install in dispersed data centers a new kind of chip to make everything faster. These chips would be called T.P.U.s, or "tensor processing units," and their value proposition — counterintuitively — is that they are deliberately less precise than normal chips. Rather than compute 12.246 times 54.392, they will give you the perfunctory answer to 12 times 54. On a mathematical level, rather than a metaphorical one, a neural network is just a structured series of hundreds or thousands or tens of thousands of matrix multiplications carried out in succession, and it's much more important that these processes be fast than that they be exact. "Normally,"special-purpose hardware is a bad idea. It usually works to speed up one thing. But because of the generality of neural networks, you can leverage this special-purpose hardware for a lot of other things."

3.0 what was gained? what was lost

3.0 what was gained? what was lost

- prediction vs modeling

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability
- but what *is* interpretability?

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability
- but what *is* interpretability?
  - causality

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability
- but what *is* interpretability?
  - causality
  - fairness

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability
- but what *is* interpretability?
  - causality
  - fairness
  - transparency

3.0 what was gained? what was lost

- prediction vs modeling
- prediction vs interpretability
- but what *is* interpretability?
  - causality
  - fairness
  - transparency
  - … hold on to these ideas for "ethics", in 2 weeks

prediction vs modeling: Breiman's two cultures

"In the mid-1980s two powerful new algorithms for fitting data became available: neural nets and decision trees. A new research community using these tools sprang up. Their goal was predictive accuracy. The community consisted of young computer scientists, physicists and engineers plus a few aging statisticians. They began using the new tools in working on complex prediction problems where it was obvious that data models were not applicable: speech recognition, image recognition, nonlinear time series prediction, handwriting recognition, prediction in financial markets. Their interests range over many fields that were once considered happy hunting grounds for statisticians and have turned out thousands of interesting research papers related to applications and methodology. A large majority of the papers analyze real data. The criterion for any model is what is the predictive accuracy. An idea of the range of research of this group can be got by looking at the Proceedings of the Neural Information Processing Systems Conference (their main yearly meeting) or at the Machine Learning Journal"

Rudin on interpretability and the Stakes

- “healthcare”, “justice”, “parole”,

Rudin on interpretability and the Stakes

- “healthcare”, “justice”, “parole”,
- what is difference between explainable and interpretable?

Rudin on interpretability and the Stakes

- "healthcare", "justice", "parole",
- what is difference between explainable and interpretable?
- optimization

$$L = \frac{1}{n}\sum_i [i \text{ is misclassified}] + \lambda \text{size}(f)$$

Rudin on interpretability and the Stakes

- “healthcare”, “justice”, “parole”,
- what is difference between explainable and interpretable?
- optimization

$$L = \frac{1}{n}\sum_i [i \text{ is misclassified}] + \lambda \mathrm{size}(f)$$

- what sets λ?

Rudin on interpretability and the Stakes

- "healthcare", "justice", "parole",
- what is difference between explainable and interpretable?
- optimization

$$L = \frac{1}{n}\sum_i [i \text{ is misclassified}] + \lambda \text{size}(f)$$

- what sets λ?
- answer is about epistemic virtues as much as it is about optimization or statistics

Stakes and interpretability

Angwin et alia's "Machine Bias": context

- well after CS researchers start talking about fairness (recall Wallach 2014)

Stakes and interpretability

Angwin et alia's "Machine Bias": context

- well after CS researchers start talking about fairness (recall Wallach 2014)
- huge impact on narrative

Stakes and interpretability

Angwin et alia's "Machine Bias": context

- well after CS researchers start talking about fairness (recall Wallach 2014)
- huge impact on narrative
- contested! launches papers coming to different conclusions based on different definitions of fairness

For next week(s)

- “Data Science”: ML all the things

For next week(s)

- "Data Science": ML all the things
- ethics: applications, impact, power

Recap: big ideas this week

- deep learning (b. 2012)

Recap: big ideas this week

- deep learning (b. 2012)
- machine learning (b. 1959)

Recap: big ideas this week

- deep learning (b. 2012)
- machine learning (b. 1959)
- artificial intelligence (b. 1956)

Recap: big ideas this week

- deep learning (b. 2012)
- machine learning (b. 1959)
- artificial intelligence (b. 1956)
- pattern recognition (b. 1955)

Appendix: where are we, where are we going

- 2021-01-12: intro to course

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: AI2.0

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: AI2.0
- 2021-03-30: data science, 1962-2017

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: AI2.0
- 2021-03-30: data science, 1962-2017
- 2021-04-06: ethics

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: AI2.0
- 2021-03-30: data science, 1962-2017
- 2021-04-06: ethics
- 2021-04-13: present problems: attention economy+VC=dumpsterfire

Appendix: where are we, where are we going

- 2021-01-12: intro to course
- 2021-01-19: setting the stakes
- 2021-01-26: risk and social physics
- 2021-02-02: statecraft and quantitative racism
- 2021-02-09: intelligence, causality, and policy
- 2021-02-16: data gets real: mathematical baptism
- 2021-02-23: WWII, dawn of digital computation
- 2021-03-09: birth and death of AI
- 2021-03-16: big data, old school (1958-1980)
- 2021-03-23: AI2.0
- 2021-03-30: data science, 1962-2017
- 2021-04-06: ethics
- 2021-04-13: present problems: attention economy+VC=dumpsterfire
- 2021-04-15: future solutions