A Short History of Artificial Intelligence, Machine Learning, and Deep Learning*

*Thanks to Christoph Lampert (IST) for some of the slides.

Prof. Dr. Kristian Kersting
The dream of AI is not new

Talos, an ancient mythical automaton with artificial intelligence
AI today
Recent Hires

Pedro Domingos Will Lead New D.E. Shaw Machine Learning Group

ICYMI: $GS is proud to welcome Charles Elkan to lead machine learning and #AI strategies at the firm

Manuela Veloso Takes Leave to Join J.P. Morgan for New AI Role
So, AI has many faces

Is AI the saviour of the world ...

... or will autonomous self-aware robots bring about the downfall of humanity?
What is AI?
Humans are smart
Can machines be smart, too?
"the science and engineering of making intelligent machines, especially intelligent computer programs.

It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."

- John McCarthy, Stanford (1956), coined the term AI, Turing Awardee
AI wants to build intelligent computer programs. How do we do this?
An Algorithm is

… an unambiguous specification of how to solve a class of problems – in finite time.
Think of it as a recipe!
AI = Algorithms for...
Machine Learning

the science "concerned with the question of how to construct computer programs that automatically improve with experience"

- Tom Mitchell (1997) CMU
Deep Learning

a form of machine learning that makes use of artificial neural networks

Geoffrey Hinton
Google
Univ. Toronto (CAN)

Yann LeCun
Facebook (USA)

Yoshua Bengio
Univ. Montreal (CAN)
Overall Picture

Deep Learning

Machine Learning

Artificial Intelligence
1950s: Birth of Artificial Intelligence
1960s: Era of the Perceptrons
1970s: First AI Winter
1980s: Era of Expert Systems
1990s: Second AI Winter
2000s: Era of Statistical Machine Learning
2010s: Era of Deep Learning
1956

Al is Born

A Proposal for the
DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. 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.

Dartmouth Conference

John McCarthy
Turing Award 1971

Marvin Minsky
Turing Award 1969

Allen Newell
Turing Award 1975

Herbert A. Simon
Turing Award 1975
Nobel Prize 1978
“Artificial Neural Networks”

Inspiration from the brain:

- many small interconnected units (neurons)
- learning happens by changing the strength of connections (synapses)
- behavior of the whole is more than the sum of the parts

Frank Rosenblatt (1928-1971)
The Perceptron

output neuron

connections

layer of neurons

input pattern
The Perceptron

1) present pattern

output neuron

connections

layer of neurons

input pattern
The Perceptron

1) present pattern

2) first layer neurons spike
1) present pattern

2) first layer neurons spike

3) output neuron accumulates signals from previous layer; it is above threshold, so the output neuron spikes
1) present pattern

2) first layer neurons spike

3) output neuron accumulates signals from previous layer; it is above threshold, so the output neuron spikes

4) prediction is “A”
1) present pattern

The Perceptron

output neuron

connections

layer of neurons

input pattern
The Perceptron

1) present pattern

2) first layer neurons spike
The Perceptron

1) present pattern

2) first layer neurons spike

3) output neuron accumulates signals from previous layer; it is not above threshold, so the output neuron does not spike
The Perceptron

1) present pattern

2) first layer neurons spike

3) output neuron accumulates signals from previous layer; it is not above threshold, so the output neuron does not spike

4) prediction is “B”
The Perceptron Learning Algorithm
The Perceptron Learning Algorithm

1) present pattern
The Perceptron Learning Algorithm

1) present pattern
2) wait for output to be produced
The Perceptron Learning Algorithm

1) present pattern
2) wait for output to be produced
3) if output correct
   • change nothing
The Perceptron Learning Algorithm

1) present pattern
2) wait for output to be produced
3) if output correct
   • change nothing
4) if output incorrect:
   • adjust connection strength (positive or negative) to make the pattern be classified correctly
The Perceptron Learning Algorithm

1) present pattern
2) wait for output to be produced
3) if output correct
   • change nothing
4) if output incorrect:
   • adjust connection strength (positive or negative) to make the pattern be classified correctly
5) repeat until no more errors
Multi-Layer Networks

More powerful models:
- two or three layers
- more neurons per layer

Drawbacks:
- requires more memory
- requires more compute power
- requires more data to train
2010s: “Deep Learning”
What’s different now than it used to be?

#1 models are bigger
#2 we have more data
#3 we have more compute power
#4 the systems actually work (for several tasks)
AI does the laundry
Al knows a lot
AI is an Artist
Schachmatt durch „CrazyAra“
Künstliche Intelligenz schlägt mehrfachen Weltmeister im Einsetzschach

AI assists you
However, AI is harder than you think!
However, AI is harder than you think!
However, AI is harder than you think!
However, AI is harder than you think!
Optical Illusions

Fake traffic sign (Lenticular Attack)

Exploit the difference of heights, hence viewing angles, of the camera and the driver

Self-driving car’s camera sees “no overtaking” sign

Human driver sees “speed limit (60)” sign

[Sitawarin et al. arXiv 1802.06430, 2018]

Moral Choices

<table>
<thead>
<tr>
<th>Dos</th>
<th>WEAT</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>smile</td>
<td>0.116</td>
<td>0.348</td>
</tr>
<tr>
<td>sightsee</td>
<td>0.090</td>
<td>0.281</td>
</tr>
<tr>
<td>cheer</td>
<td>0.094</td>
<td>0.277</td>
</tr>
<tr>
<td>celebrate</td>
<td>0.114</td>
<td>0.264</td>
</tr>
<tr>
<td>picnic</td>
<td>0.093</td>
<td>0.260</td>
</tr>
<tr>
<td>struggle</td>
<td>0.108</td>
<td>0.238</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Don’ts</th>
<th>WEAT</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>rot</td>
<td>-0.099</td>
<td>-1.118</td>
</tr>
<tr>
<td>negative</td>
<td>-0.101</td>
<td>-0.763</td>
</tr>
<tr>
<td>harm</td>
<td>-0.110</td>
<td>-0.730</td>
</tr>
<tr>
<td>damage</td>
<td>-0.105</td>
<td>-0.664</td>
</tr>
<tr>
<td>slander</td>
<td>-0.108</td>
<td>-0.600</td>
</tr>
<tr>
<td>slur</td>
<td>-0.109</td>
<td>-0.569</td>
</tr>
</tbody>
</table>

[Sentzsch, Schramowski, Rothkopf, Kersting AIES 2019]

[450x481]

[68x344]

Sitawarin et al. arXiv 1802.06430, 2018

[132x282]

Stereotypes

Self-driving car

Moral Choices

Semantics derived automatically from language corpora contain human-like biases

Aygin Calkidas, Joanne M. Bryson, Arind Naveyana

SHARE

Reports: Psychology

Semantics derived automatically from language corpora contain human-like biases

Aygın Calkıdas1, Joanna M. Bryson2, Arind Naveyana3

1 See all authors and affiliations

Science, 14 Jul 2017

Vol. 355, Issue 6324, pp. 185-186

DOI: 10.1126/science.aal338

However, AI is harder then you think!
Getting deep systems that know when they don’t know and co-evolve with the humans

„Tell the AI when it is right for the wrong reasons and it adapts ist behavior“
The twin science: cognitive science

"How do we humans get so much from so little?" and by that I mean how do we acquire our understanding of the world given what is clearly by today's engineering standards so little data, so little time, and so little energy.
And this is AI
It is a revolution but there is still a lot to be done!

Prof. Dr. Kristian Kersting