Ultimate dream: thinking machine
Ultimate dream: thinking machine

Schematic of a biological neuron.

http://sebastianraschka.com/
Activation Functions
Logistic regression units

Hardware implementations

Frank Rosenblatt, ~1957: Perceptron

Widrow and Hoff, ~1960: Adaline/Madaline
False Promises

“The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence ... Dr. Frank Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers”

The New York Times  July 08, 1958

(Simple) AND/OR problem: linearly separable?
(Simple) XOR problem: linearly separable?

```
<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
```

or

```
<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
```

and

```
<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
```

xor
Perceptrons (1969)  
by Marvin Minsky, founder of the MIT AI Lab

- We need to use MLP, multilayer perceptrons (multilayer neural nets)
- No one on earth had found a viable way to train MLPs good enough to learn such simple functions.
“No one on earth had found a viable way to train*”

*Marvin Minsky, 1969

http://cs231n.github.io/convolutional-networks/
Backpropagation

Convolutional Neural Networks

At some point in the late 1990s, one of these systems was reading 10 to 20% of all the checks in the US.

Hubel & Wiesel, 1959
Convolutional Neural Networks

“At some point in the late 1990s, one of these systems was reading 10 to 20% of all the checks in the US.”

[LeNet-5, LeCun 1980]
“Alvinn: An autonomous land vehicle in a neural network”
Terminator 2 (1991)

JOHN: Can you learn? So you can be... you know. More human. Not such a dork all the time.

TERMINATOR: My CPU is a neural-net processor... a learning computer. But Skynet presets the switch to "read-only" when we are sent out alone.

... We'll learn how to set the neural net

TERMINATOR: Basically. (starting the engine, backing out) The Skynet funding bill is passed. The system goes on-line August 4th, 1997. Human decisions are removed from strategic defense. Skynet begins to learn, at a geometric rate. It becomes self-aware at 2:14 a.m. eastern time, August 29. In a panic, they try to pull the plug.

SARAH: And Skynet fights back.

TERMINATOR: Yes. It launches its ICBMs against their targets in Russia.

SARAH: Why attack Russia?

TERMINATOR: Because Skynet knows the Russian counter-strike will remove its enemies here.

A BIG problem

- Backpropagation just did not work well for normal neural nets with many layers
- Other rising machine learning algorithms: SVM, RandomForest, etc.
- 1995 “Comparison of Learning Algorithms For Handwritten Digit Recognition” by LeCun et al. found that this new approach worked better

Next
To be continued…
Canadian Institute for Advanced Research (CIFAR)

CIFAR encourages basic research without direct application, was what motivated Hinton to move to Canada in 1987, and funded his work afterward.

“Everyone else was doing something different"

• “It was the worst possible time,” says Bengio, a professor at the Université de Montréal and co-director of the CIFAR program since it was renewed last year. “Everyone else was doing something different. Somehow, Geoff convinced them.”

• “We should give (CIFAR) a lot of credit for making that gamble.”

• CIFAR “had a huge impact in forming a community around deep learning,” adds LeCun

• In 2006, Hinton, Simon Osindero, and Yee-Whye Teh published, “A fast learning algorithm for deep belief nets”

• Yoshua Bengio et al. in 2007 with “Greedy Layer-Wise Training of Deep Networks”
Neural networks with many layers really could be trained well, *if the weights are initialized in a clever way* rather than randomly.

Deep machine learning methods are more efficient for difficult problems than shallow methods.

Rebranding to *Deep Nets, Deep Learning*

The Image Classification Challenge:
1,000 object classes
1,431,167 images

Russakovsky et al. arXiv, 2014
ImageNet Classification (2010 –

26.2% to 15.3%
ImageNet Classification (2010 – 2015)

Top-5 Classification Error (%)

- 2010
- 2011
- 2012
- 2013
- 2014
- Human
- 2015

Convolutional Neural Nets
Neural networks that can explain photos

Deep API Learning*

Figure 3: The Overall Workflow of DEEPAPI

*GU et al. at HKUST with MSRA
Speech recognition errors

Error (smaller is better)

- Apple Dictation
- Bing Speech
- wit.ai (Facebook)
- Google API
- Deep Speech

Chart Title: Speech recognition errors
Google DeepMind's Deep Q-learning playing Atari Breakout

https://youtu.be/V1eYniJ0Rnk
Geoffrey Hinton’s summary of findings up to today

• Our labeled datasets were thousands of times too small.
• Our computers were millions of times too slow.
• We initialized the weights in a stupid way.
• We used the wrong type of non-linearity.

Why should I care?

• I am not a researcher, not a computer scientist!

• Do you have data?

• Do you sell something?

• Are doing any business?
Random Search vs. Grid Search

Grid Layout

Random Layout

Random Search for Hyper-Parameter Optimization
Borgström and Bengio, 2012

Fei-Fei Li & Andrej Karpathy & Justin Johnson
Lecture 5 - 90  20 Jan 2016

different taxes and you end up with a better spot than here where you've
Doing a PhD in Software Testing and Analysis? Submit to the ISSTA 2016 Doctoral Symposium by April 22! Featuring a keynote by Alex Orso! Details: https://issta2016.cispa.saarland/doctormissymposium/

Call for doctoral symposium submissions
ISSTA is the leading research symposium in software testing and analysis, bringing together academics, industrial researchers, and practitioners to exchange new ideas, problems, and experiences on how to test and analyze software systems. The ISSTA Doctoral Symposium is a forum for PhD students wor…

ISSTA2016.CISPA.SAARLAND
Sung Kim's CSE Homepage
www.cse.ust.hk/~hunkim/
Sung is an associate professor at the Hong Kong University of Science and Technology. He was a post-doc at the Program Analysis Group at MIT. He received ...
Publications - Research - Software - Teaching

Sung's Publications
www.cse.ust.hk/~hunkim/Publications.html
Sung's Publications. 2015. Jaechang Nam and Sunghun Kim, "Heterogeneous Defect Prediction", In Proceedings of the 10th European Software Engineering ...

Sung Kim - Wikipedia, the free encyclopedia
Sung Y. Kim (born 1960) is a Korean-born U.S. diplomat and the current United States Special Representative for North Korea Policy. He previously served as ...
Early life and education - Professional career - Ambassador to South Korea
Family Adventures from the 1980s

Based on your interest in...

- The New Adventures of Madeline
- The New Adventures of Marmalade
- The Great Mouse Detective
- The Great Mouse Detective
- The Mouse and the Motorcycle
- The Karate Kid
- Ralph S. Mouse
- The Adventures of Mark Twain

Family Comedies

Based on your interest in...

- Hoodwinked
- Mckenna
- Mr. Stitch
- The Smurfs 2
- The Parent Trap
- Spy Kids 3
Why Now?

• Students/Researchers
  - Not too late to be a world expert
  - Not too complicated (mathematically)

• Practitioner
  - Accurate enough to be used in practice
  - many ready-to-use tools such as TensorFlow
  - Many easy/simple programming languages such as Python

• After all, it is fun!
Next

Neural Nets Basic with XOR!