The delusions of Neural Networks

How ~~business marketing~~ hype hurts computing science

Send complaints to

Giacomo Tesio father, husband and programmer

giacomo@tesio.it
http://jehanne.io/
https://github.com/Shamar

By day: Financial Applications in C#, Javascript… whatever you pay for
By night: Distributed Operating Systems and Network Protocols

I’m **not** a expert in Statistics. So take this with a grain of salt! ;-)

I’m here because of http://bit.do/the-delusions-of-neural-networks

The delusions of Neural Networks

- What is an Artificial Neural Network?
- Why they need Big Data?
- Generalizing Neural Networks: the AGI/ASI pipedream
    - How far we are?
    - Where is the intelligence?
    - Counter argument: unsupervised learning
- The threats to Artificial Intelligence
- Can computers think?

A little experiment...

What do you see?
(up to two words)

What is an Artificial Neural Network?

> We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

Currently, they constitute the most exciting research field in **Statistics**.

What is an Artificial Neural Network?

We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

From a **legal** point of view it's important to note that

- they are just applied statistics, not an inscrutable computer brain
- their output can always be explained (till quantum computing)
- Cybenko's theorem prove they can approximate any **continuous** function
- there is no way to **prove** they are approximating a specific **discrete** function
- AI is **not accountable**, so it cannot take decisions over humans

(more about the GDPR and ANN at http://bit.do/the-delusions-of-neural-networks)

What is an Artificial Neural Network?

We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

What is an Artificial Neural Network?

We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

Function

> Given two set A and B, a function $f : \mathrm{A} \to \mathrm{B}$ is a rule that assigns to each element in A (domain) exactly one element in B (codomain).

If you have **two sets and a rule** that map each element of one set to exactly one element of the other, you have a function.

Equality: $f = g \quad iff \; f : \mathrm{A}{\to}\mathrm{B} \;\land\; g : \mathrm{A}{\to}\mathrm{B} \;\land\; \forall x \in \mathrm{A}, \forall y \in \mathrm{B}, f(x) = y \Leftrightarrow g(x) = y$

Composition: $f : \mathrm{A}{\to}\mathrm{B} \;\land\; g : \mathrm{B}{\to}\mathrm{C} \quad \Rightarrow \quad (g \circ f) : \mathrm{A}{\to}\mathrm{C}; \; (g \circ f) = g(f(x))$

What is an Artificial Neural Network?

We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

Any function

Neural networks can statistically approximate **any** function.

Even **unknown** ones.

If you **suspect** that a function exists, you can try to statistically approximate it with a neural network, even if you do not know the rule that it follows.
You just need **two set**. And **tons of data**.

This is the strongest strength of neural networks. And their **weakness**, too.

Why Big Data (set)?

> We don’t have better algorithms. We just have more data.
> —Peter Norwing, Chief Scientist, Google (2009)

Artificial Neural Networks turned “cool at 70” because people leak **tons of data**.

Since they can approximate any continuous function, we need a big data set to **filter out unwanted functions** with each sample we feed to it.

Still, infinitely many functions fit our samples!

Overfit & Underfit

We can not really know **which function** a complex ANN will approximate.

Generalizing Artificial Neural Networks

Can we move from **narrow intelligence** to **general intelligence**?

Generalizing Artificial Neural Networks

Can we move from **narrow intelligence** to **general intelligence**?

We call artificial neural networks a class of deterministic algorithms that can statistically approximate any function.

Given two set A and B, a function $f : \mathrm{A} \rightarrow \mathrm{B}$ is a rule that assigns to each element in A (domain) exactly one element in B (codomain).

Equality: $f = g$ *iff* $f : \mathrm{A} \rightarrow \mathrm{B} \wedge g : \mathrm{A} \rightarrow \mathrm{B} \wedge \forall x \in \mathrm{A}, \forall y \in \mathrm{B}, f(x) = y \Leftrightarrow g(x) = y$

Composition: $f : \mathrm{A} \rightarrow \mathrm{B}$ and $g : \mathrm{B} \rightarrow \mathrm{C} \Rightarrow (g \circ f) : \mathrm{A} \rightarrow \mathrm{C}; (g \circ f) = g(f(x))$

Generalizing Artificial Neural Networks

$$intelligence : \text{Domain} \rightarrow \text{Codomain}$$

Generalizing Artificial Neural Networks

$\textit{intelligence}: \quad \text{Perceptions}_T \rightarrow \text{Actions}_{T+1}$

Generalizing Artificial Neural Networks

$$\textit{intelligence} : \quad \text{Perceptions}_T \quad \rightarrow \quad \text{Actions}_{T+1}$$

Universal Intelligence: A Definition of Machine Intelligence
Shane Legg, Marcus Hutter (2007)

First issues

- Domain and Codomain depends on “hardware”, $\textit{intelligence}$ does not
- Domain and Codomain are (potentially) infinite sets
- **No equality** relation in the Codomain

But we are also missing something very important...

Generalizing Artificial Neural Networks

$$intelligence : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

Knowledge is both an input and an output of *intelligence*!

Given a different initial knowledge, an intelligent agent:

- reacts differently to perceptions
- learns different things

What is Knowledge?

Generalizing Artificial Neural Networks

$$\textit{intelligence} : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

According to George Kelly (The psychology of personal constructs, 1955)

$$\text{Knowledge} \quad = \quad \text{Constructs} \quad \& \quad \text{Relations}$$

To use them mathematically, we will translate these psycological terms to

$$\text{Knowledge} \quad = \quad \underbrace{\text{Sets} \quad \times \quad \text{Functions}}_{\text{Models}}$$

Knowledge is the set of models we use to guide our actions (and predict outcomes)

Generalizing Artificial Neural Networks

$$intelligence : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

$\text{Knowledge} = \text{Sets} \times \text{Functions}$ (aka Models)

Still according to function equality

$$f = g \qquad iff\ f : \text{A} \rightarrow \text{B} \wedge g : \text{A} \rightarrow \text{B} \wedge \forall x \in \text{A}, \forall y \in \text{B}, f(x) = y \Leftrightarrow g(x) = y$$

to know we are approximating the *intelligence* function, we need to know its rule!

Given two set A and B, a function $f : \text{A} \rightarrow \text{B}$ is **a rule** that assigns to each element in A (domain) exactly one element in B (codomain).

Generalizing Artificial Neural Networks

$$\textit{intelligence} : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

$\text{Knowledge} = \text{Sets} \times \text{Functions}$ (aka Models)

Triarchic theory of intelligence (by Robert J. Sternberg):

- Analytical
- Creative
- Practical

Different components of intelligence that address different needs and interacts in a person's life

PROBLEM: these components are identified by "clustering" IQ tests' results just like Legg&Hutter, based on **external** measures of intelligence

Generalizing Artificial Neural Networks

$$intelligence : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

$\text{Knowledge} = \text{Sets} \times \text{Functions}$ (aka Models)

Why not simply **observe** intelligence at work in our head? We will see:

- *comprehension* uses perceptions to select (filter) useful knowledge
- *imagination* uses the relevant models to predict the effects of actions
- *will* uses predictions to take decisions
- *execution* turn decisions to actions
- *abstraction* uses previous knowledge and perception to improve knowledge

$$intelligence = (execution \circ will \circ imagination \circ comprehension) \times abstraction$$

Generalizing Artificial Neural Networks

$$\textit{intelligence} : \quad (\text{Perceptions} \times \text{Knowledge})_T \rightarrow (\text{Actions} \times \text{Knowledge})_{T+1}$$

$\text{Knowledge} = \text{Sets} \times \text{Functions}$ (aka Models)

$$\textit{intelligence} = (\textit{execution} \circ \textit{will} \circ \textit{imagination} \circ \textit{comprehension}) \times \textit{abstraction}$$

This means that

- to be (part of) intelligence, an ANN should approximate one of these functions
- (to prove) to be general, an Artificial Intelligence should be able to **discover** and **explain** us **new abstractions** and **functions** over them
- Artificial General Intelligence **is** Artificial Super Intelligence!

Artificial General Intelligence: how far are we?

What do you see?
(up to two words)

Artificial General Intelligence: how far are we?

What do you see?
(up to two words)

this is called
“Pattern Recognition”

Artificial General Intelligence: how far are we?

We are very good at
Pattern Recognition

Artificial General Intelligence: how far are we?

We are very good at
Pattern Recognition

...still,
there is no **cat** here

So, where is the intelligence?

The Beauty is in the eye of the beholder!

So, where is the intelligence?

The Beauty is in the eye of the beholder!

When we see an ANN choosing Go moves, we recognize a pattern.

So, where is the intelligence?

The Beauty is in the eye of the beholder!

When we see an ANN choosing Go moves, we recognize a pattern.

We match the program behaviour with experiences from our own memories.

So, where is the intelligence?

The Beauty is in the eye of the beholder!

When we see an ANN choosing Go moves, we recognize a pattern.

We match the program behaviour with experiences from our own memories.

We look at the computer and we see a Go player. We see an intelligence.

So, where is the intelligence?

The Beauty is in the eye of the beholder!

When we see an ANN selecting the next Goban state, we recognize a pattern.

We match the program behaviour with experiences from our own memories.

We look at the computer and we see a Go player. We see an intelligence.

But it's like with the cat.

So, where is the intelligence?

The Beauty is in the eye of the beholder!

AlphaGo Zero <u>does not need intelligence</u> to play Go.

It has **aggregated statistics** over 4,9 millions of games that no human could play.

The AlphaGo Zero algorithm is a great application of **human** intelligence.

- it uses self playing to compute the rewards for moves (actually MCTS)
- it uses the rewards to compute win probability of each move
- it uses the win probability of each move to calibrate the ANN

AlphaGo Zero approximates a function, from goban’s states to win probabilities.

The threats to Artificial Intelligence

So far, there is **no danger in AI** for humanity (dude, it's just statistics!), except

- **bad people** using it
- **incompetent** autorities (or worse than incompetent...)

However improper use of AI, let people demage other people and AI research. Some of the current threats to the field are:

- Misleaded Trust: blackbox decides over people
- Uninformed Fear: hide/protect the controllers
- Emotional Bonds (eg Google Clips)
- Evocative Language (aka Anthropomorphization)

Business-aided **ignorance**

The threats to Artificial Intelligence

Misleaded Trust

Eric L. Loomis was classified as “high risk” by a proprietary software and thus sentenced to six years in prison.

<u>That software is bugged</u> (just like the others).
But the judge **trust** it without even understanding how it works.

Who **accounts** for errors? The company’s CEO? Stockholders? Programmers?
What about subtle discriminations of a minority? How can you prove them?

The threats to Artificial Intelligence

Uninformed Fear

Hide the real risk: PEOPLE!

- incompetence of autorities
  (the recent Norwegian DPA report is embarrassing!)
- malicious parameters selection
- malicious data set corruption
- malicious features selection

All too easy to hide behind the “blackbox”!

The threats to Artificial Intelligence

Uninformed Fear

The worst threat from AI to humanity is in fact **Idiocracy**!

Paper	⇒	Less Need to Remember
Calculator	⇒	Less Need to Calculate
AGI/ASI	⇒	Less Need to **Think**

No need for a T800, a bit of patience (and irony) and humans will simply “evolve” back to apes!

The threats to Artificial Intelligence

Emotional Bonds   https://design.google/library/ux-ai/

This is plain ~~manipulation~~ marketing:

accountability of AI is “problematic”

↓

rational people do not
trust AI controllers

↓

explain them why they should trust AI

The threats to Artificial Intelligence

Emotional Bonds https://design.google/library/ux-ai/

This is plain ~~manipulation~~ marketing:

accountability of AI is “problematic”

↓

rational people do not
trust AI controllers

↓

~~explain them why they should trust AI~~

↓

let’s make people less rational

The threats to Artificial Intelligence

Evocative Language (aka Anthropomorphization)

The **words** we use to describe the reality forge our **understanding** of it.

Artificial Intelligence

Artificial Neural Network

Deep Learning ANN

Machine Learning

Training

The threats to Artificial Intelligence

Evocative Language (aka Anthropomorphization)

The **words** we use to describe the reality forge our **understanding** of it.

- Artificial Intelligence
- Artificial Neural Network
- Deep Learning ANN
- Machine Learning
- Training

Evocative. Not descriptive.

Antropomorphic (historically).

Good Literature ⇒ Bad Science.

The threats to Artificial Intelligence

Evocative Language (aka Anthropomorphization)

The **words** we use to describe the reality forge our **understanding** of it.

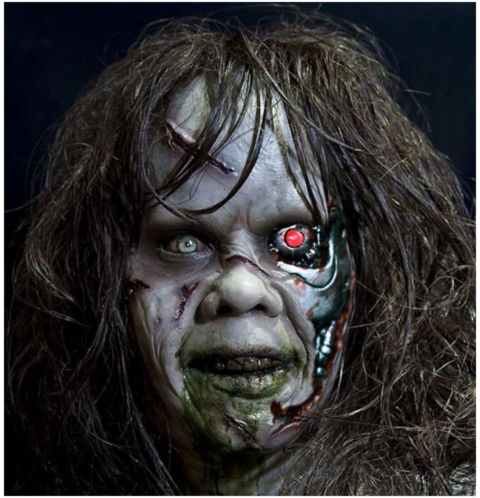
Artificial Intelligence	Simulation of Intelligence
Artificial Neural Network	Chain of Logistic Approximators
Deep Learning ANN	Long Chain of Logistic Approximators
Machine Learning	Computer-aided Statistics
Training	Statistical Calibration

The threats to Artificial Intelligence

Evocative Language (aka Anthropomorphization)

The **words** we use to describe the reality forge our **understanding** of it.

Techo Exorcism



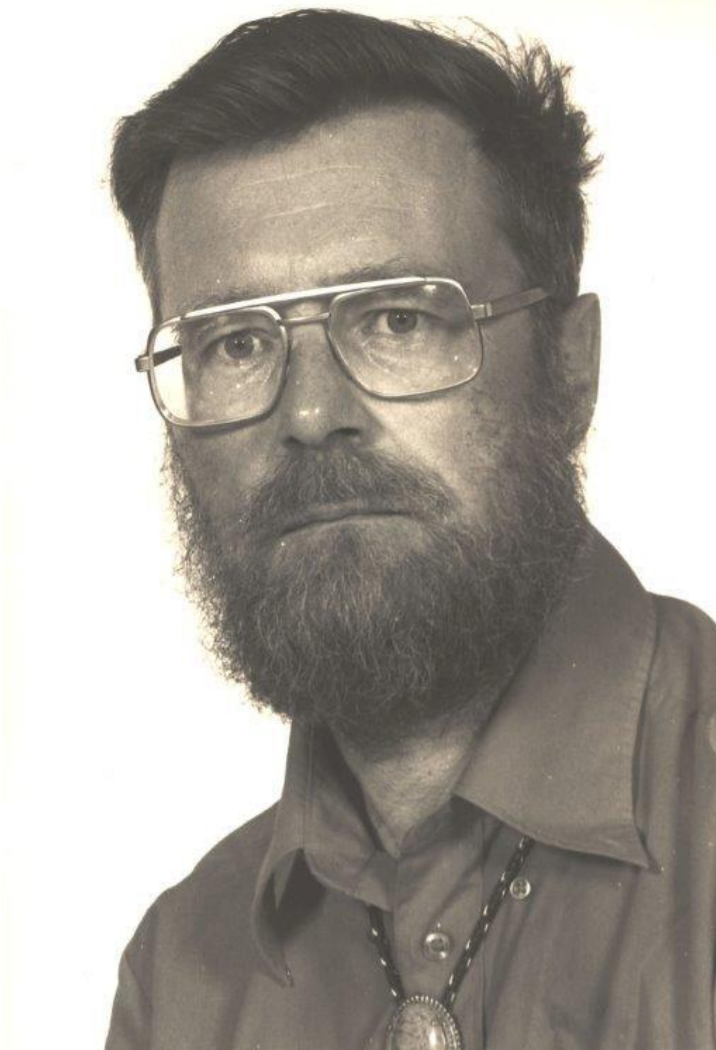
Debugging Session

Can Computers Think?

Can Computers Think?

...a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.

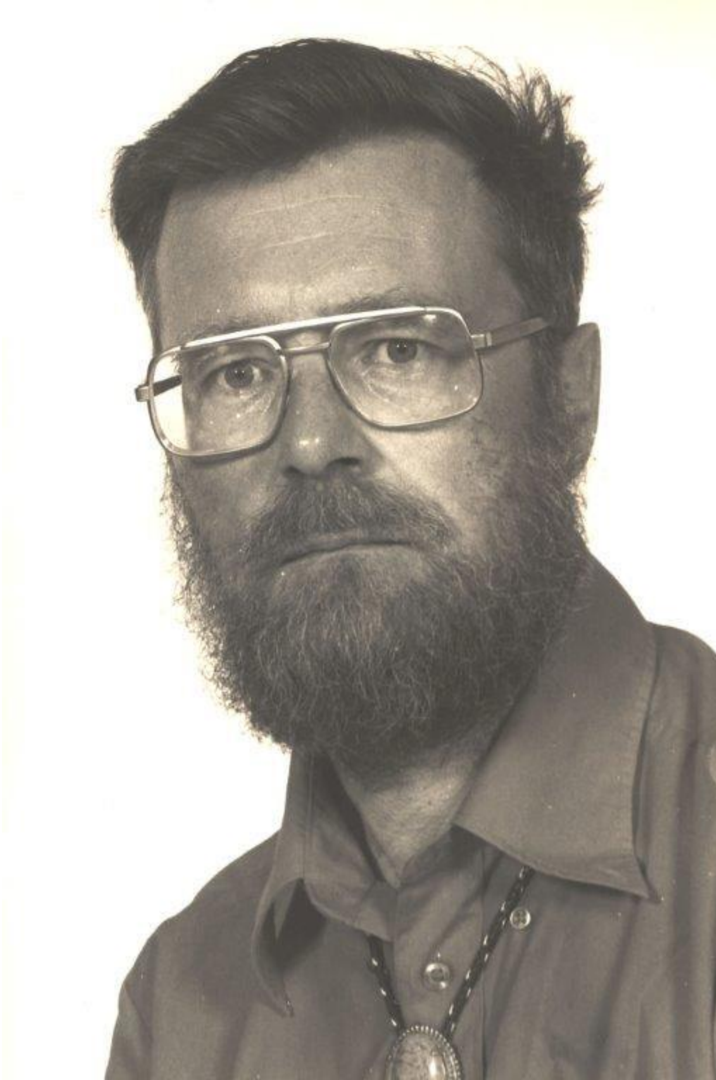
The threats to computing science
Edsger W. Dijkstra



Can Computers Think?

...a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.

The threats to computing science
Edsger W. Dijkstra (1984)

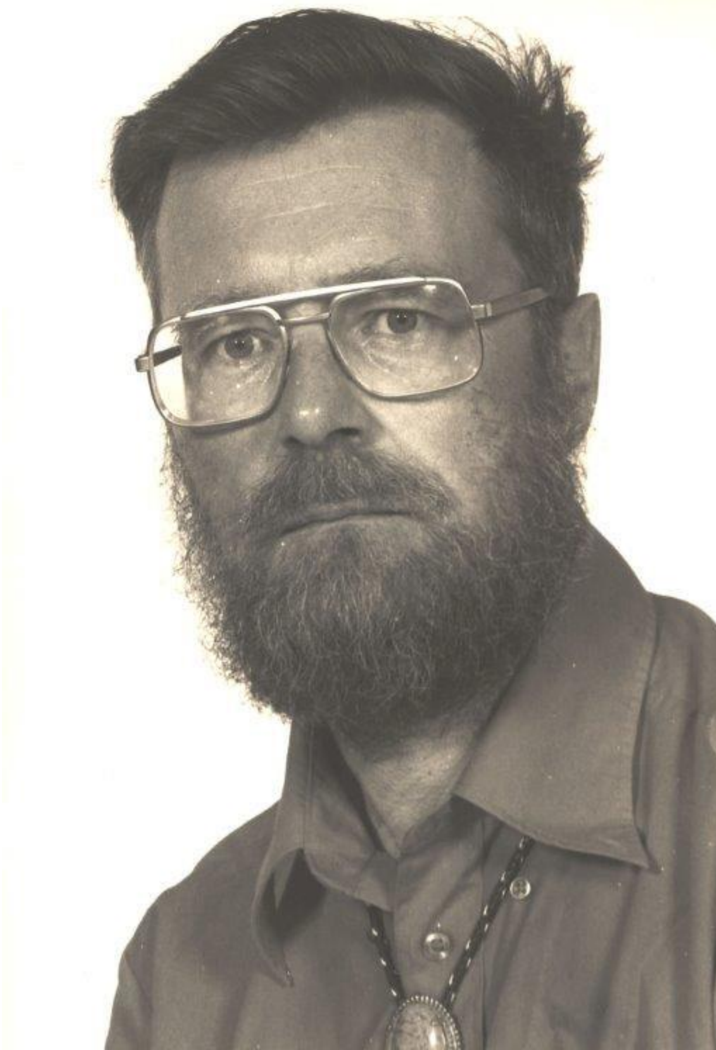


Can Computers Think?

...a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.

The threats to computing science
Edsger W. Dijkstra (1984)

What about 2018?



Can Computers Think?

...a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.

The threats to computing science
Edsger W. Dijkstra (1984)

What about 2018? **NO** they can not.