Deep Learning beyond the hype

Pierre Marion

Corps des Mines - Sorbonne Université

pierre.marion@mines.org

November 14, 2019

Overview

A bit of history

Some of the remaining challenges

Conclusion

Deep Learning: unrealistic expectations?

Figure: Gartner Hype Circle for Emerging Technologies, 2015

Deep Learning: unrealistic expectations?

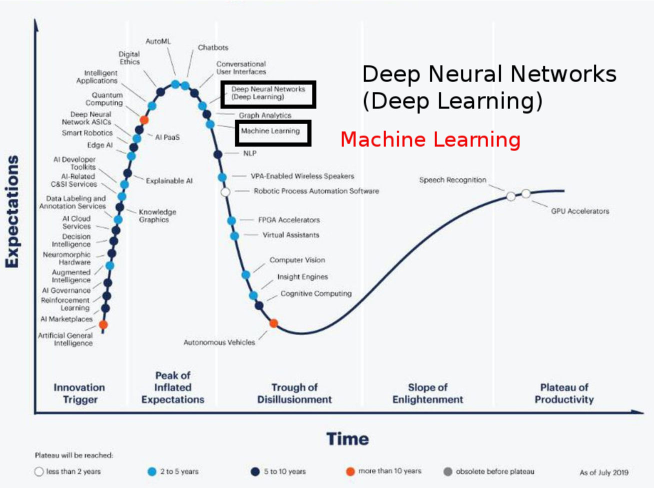
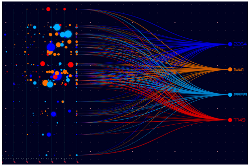


Figure: Gartner Hype Circle for AI, 2019

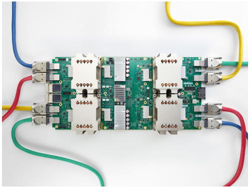
The birth of Artifical Neural Networks

- Formal Model of the Neuron: 1943 (McCulloch - Pitts)
- Perceptron: 1958 (Rosenblatt)
- Adaline - Adaptive Linear Element: 1963 (Widrow - Hoff) `https://youtu.be/IEFRtz68m-8?t=561`
- Supervised multilayer perceptrons: 1965 (Ivakhnenko - Lapa)
- Neocognitron: 1980 (Fukushima)
- First practial demonstration of backpropagation: 1989 (LeCun)

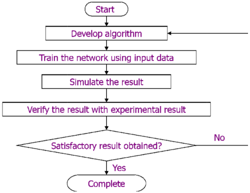
Since 1989



Data



Hardware



Algorithms

Some of the remaining challenges

"Mathematical models of physical phenomena should have the properties that:

1. a solution exists,
2. the solution is unique,
3. the solution's behavior changes continuously with the initial conditions."

— Jacques Hadamard [Hadamard, 1902]

Is a deep neural network a well-posed model?

- Adversarial attacks
- Convergence properties (non-convex optimization)

Some of the remaining challenges

Is a deep neural network (DNN) a proper model?

- Adversarial attacks
- Convergence properties (non-convex optimization)

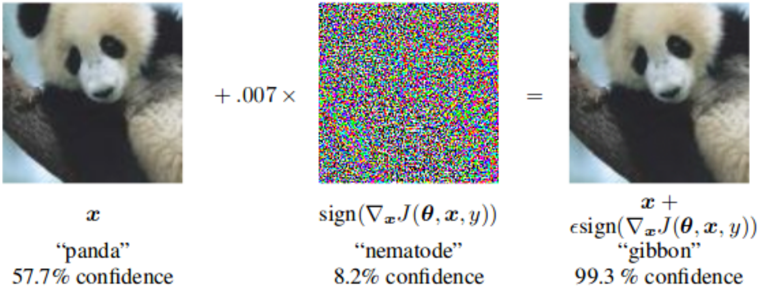
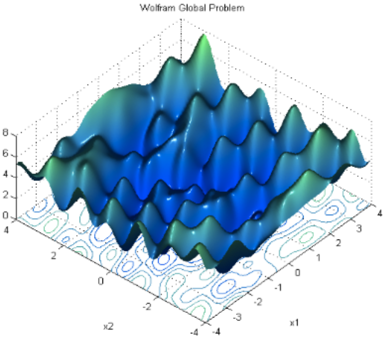


Figure: A demonstration of fast adversarial example generation applied to GoogLeNet on ImageNet. [Goodfellow, 2014]

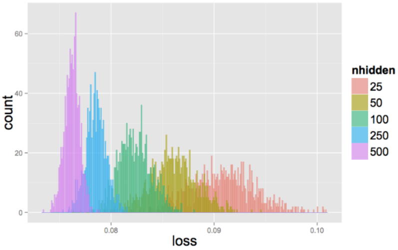
Some of the remaining challenges

Is a deep neural network a proper model?

- Adversarial attacks
- Convergence properties (non-convex optimization)



Typical objective function



Distribution of critical points

[Bengio, 2015]

My research topic: DNN and dynamical systems

Basic Idea:
A very popular type of DNN (ResNet) is a discretization at order 1 (Euler discretization) of a dynamical system (e.g. associated with an ODE).

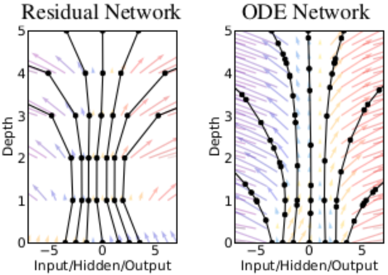
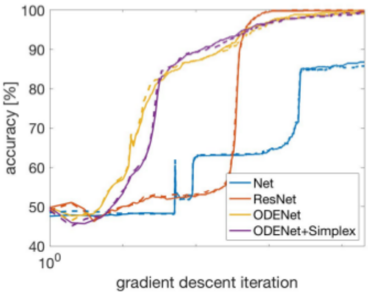


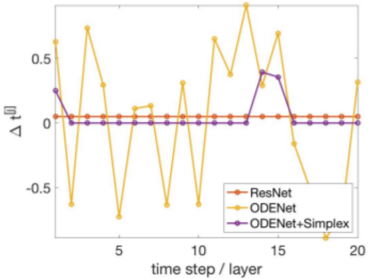
Figure 1: *Left:* A Residual network defines a discrete sequence of finite transformations. *Right:* A ODE network defines a vector field, which continuously transforms the state. *Both:* Circles represent evaluation locations.

[Chen, 2018]

My research topic: DNN and dynamical systems



Classification accuracy over the course of the gradient descent iterations



Estimated time steps. ODENet+simplex consistently picks two to three time steps and set the rest to zero.

[Benning, 2019]

Take-home messages

Predicting future research breakthroughs is hard

Cf. the controversy on Artifical General Intelligence

Without maths, you don't go far

Free products and not life-or-death applications

There is a gap to be filled between 'traditional' Applied Mathematics and Deep Learning

References

Jacques Hadamard (1902)
Sur les problèmes aux dérivées partielles et leur signification physique
Princeton University Bulletin pp. 49–52.

Ian J. Goodfellow et al. (2014)
Explaining and Harnessing Adversarial Examples
eprint arXiv:1412.6572

Yoshua Bengio (2015)
Deep Learning:Theoretical Motivations
Course at Deep Learning Summer School 2015, Montreal
http://videolectures.net/deeplearning2015_bengio_theoretical_motivations/

Ricky Chen et al. (2018)
Neural Ordinary Differential Equations
eprint arXiv:1806.07366

Martin Benning et al. (2019)
Deep learning as optimal control problems: models and numerical methods
eprint arXiv:1904.05657

Questions?