Deep Learning: A Critical Appraisal

Alan Van Omen

Sehong Oh

Contents

Introduction

GARY MARCUS is a leading voice in artificial intelligence. He is a scientist, best-selling author, and serial entrepreneur (Founder of Robust.AI and Geometric.AI, acquired by Uber). He is well-known for his challenges to contemporary AI, anticipating many of the current limitations decades in advance, and for his research in human language development and cognitive neuroscience.

An Emeritus Professor of Psychology and Neural Science at NYU, he is the author of five books, including, *The Algebraic Mind*, *Kluge*, *The Birth of the Mind*, and the New York Times Bestseller *Guitar Zero*. He has often contributed to *The New Yorker*, *Wired*, and *The New York Times.* His most recent book, *Rebooting AI*, with Ernest Davis, is one of Forbes's 7 Must Read Books in AI.

http://garymarcus.com/

Introduction

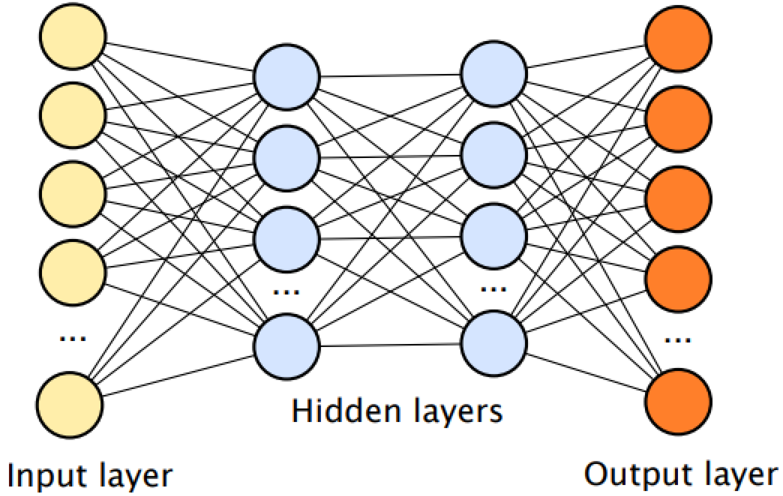
- More likely to be an Article

- For researchers in AI field and consumers with less technical background

- Deep learning is not likely to be disappear, but 5 years since it appeared is good moment for a critical reflection

Background

- Input-output mappings
- Training with a large amount of data



Limits on the scope of deep learning

- Deep learning is Data Hungry
- DL is shallow and has limited capacity for transfer
- DL has no natural way to deal with hierarchical structure
- DL has struggled with open-ended inference
- DL is not sufficiently transparent
- DL has not been well integrated with prior knowledge
- DL cannot inherently distinguish causation from correlation
- DL presumes a largely stable world
- DL works well as an approximation, but cannot be fully trusted
- DL is difficult to engineer with

Deep Learning is Data Hungry

- Deep learning requires a large, labeled dataset, often not easily obtainable

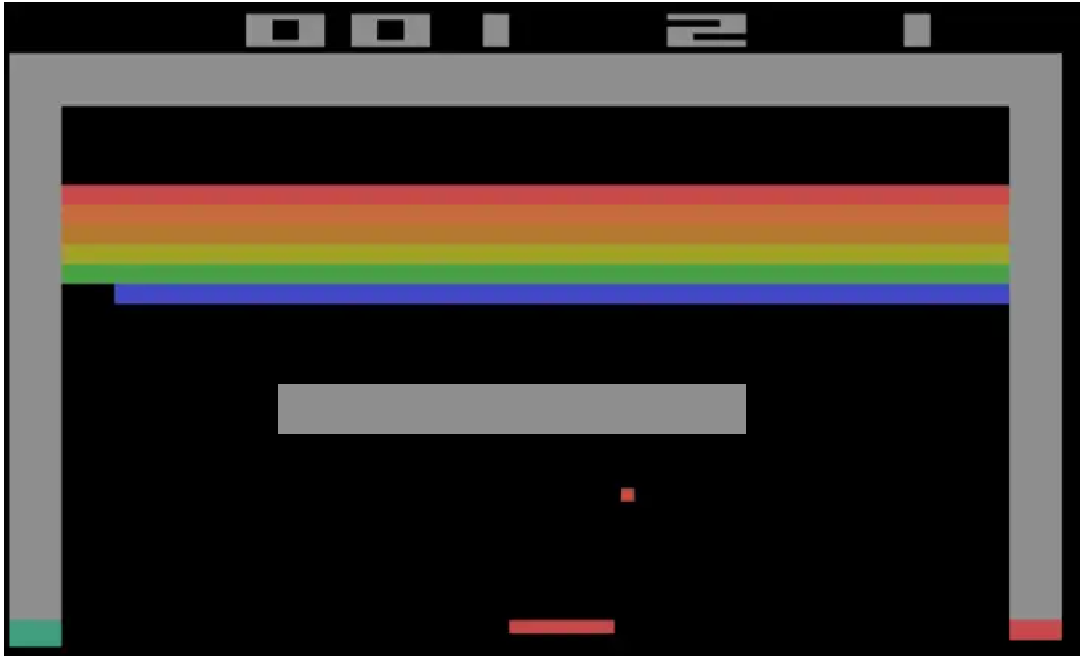
- However, people can understand and solve problems with definitions and a few examples.

  Ex) *Schmister*

Deep Learning is Shallow

- DL is shallow and has limited capacity for transfer
- The word "Deep" is from the property of deep learning
- Highly overfitted from present environment
- *Brick breaker game*

Deep Learning is Shallow cont.



Deep Learning has No Natural Way to Deal with Hierarchical Structure

- Language has a hierarchical structure.
- Deep learning does not understand the structure.
- We do not have deep learning model to reflect this kind of structure.
- How about now?

Deep Learning Struggles with Open-Ended Inference

- DL has struggled with open-ended inference

- How many tigers are there?

- What are they doing?

Discussion Contributions (@1, @2, @5, ...)

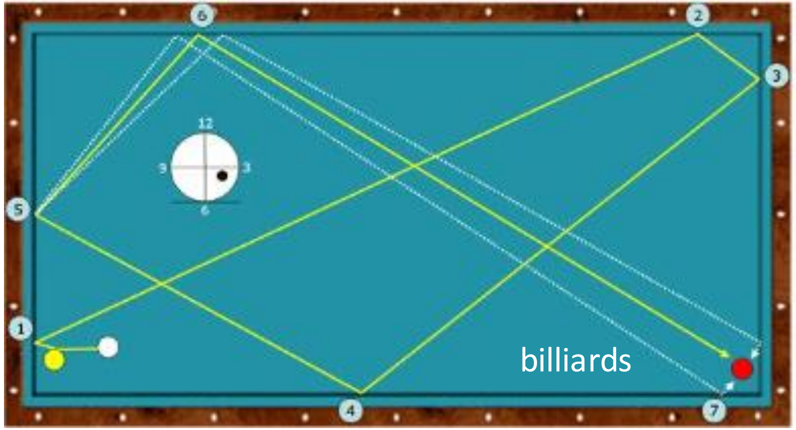
- Since the release of the paper, machine learning has been significant advances in areas dealing with open-ended inference and incorporating hierarchical structure. Some examples of these include:
  - CLIP models
  - ChatGPT
- These methods still have their weaknesses, for example, being very "data hungry."

Deep Learning is Not Sufficiently Transparent

- Deep learning model is a basically back box
- We know how the model is trained mathematically, but do not know which parts it considers to judge the image as a cat
- Hard to find the problem of the model and which part should be changed to improve the model. Can't just debug it like a conventional algorithm.

DL has not been well integrated with prior knowledge

- End-to-end learning
- The best result from a given data set

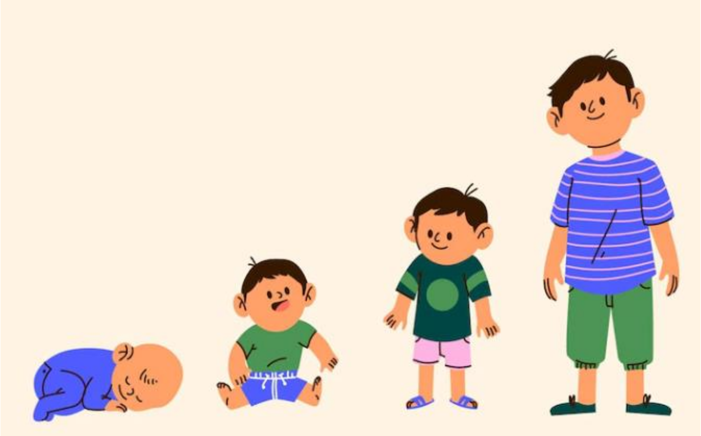


incidence and reflection angle

Deep Learning Cannot Distinguish Causation from Correlation

- As babies growing up, their language ability is getting better

- DL easily identifies these correlations, but has not context to understand causality (getting taller does not cause better language abilities)



https://www.freepik.com/

Deep Learning Presumes a Largely Stable World

- DL presumes a largely stable world

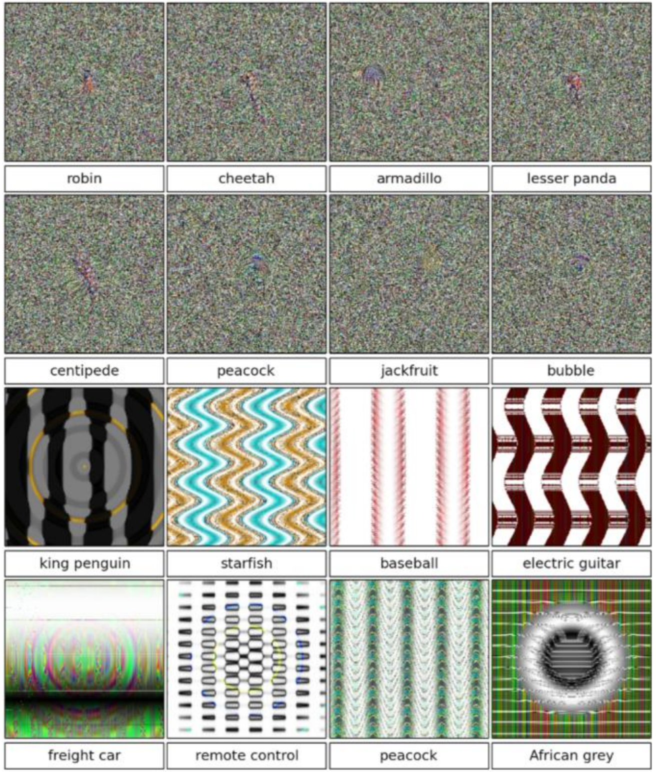
- Predicting stock patterns or weather behavior which are randomly affected by numerous tiny factors are instable and hard to model with ML.

Deep Learning Cannot be Trusted

- DL works well as an approximation, but cannot be fully trusted
- DL methods have been shown to be highly susceptible to alterations in input data.
- They are easy to fool, and are vulnerable to attacks.

Discussion Question #1

- Deep learning is a black box model and suffers from a lack of explainability, which can result in some unexpected behaviors, even in this best models. Particularly in highly critical application like self-driving cars or weapon systems, do you think machine learning models can ever be fully trusted? How would we determine a model was trustworthy?



Nguyen, Anh, Jason Yosinski, and Jeff Clune. "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

Discussion Contributions

- One method for determining if a method is trustworthy or reliable, is to see if can perform better than a human. This is hard to quantify, and often a DL method will still make obvious mistakes a human would never make.
- There has been significant research recently in areas to improve confidence measures including OOD detection, anomaly detection, novelty detection, etc.
- DL models are not complete black boxes like Marcus seems to suggest. There has been research which shows the effects of different layers and the role they play in prediction. There are also methods such as Shapley value analysis, which can provide additional explainability.

Deep Learning is Difficult to Engineer With

- DL is "the high-interest credit card of technical debt"
- It is relatively easy/quick to get good results in a problem space, even if you don't really understand the field (short term gain)
- Difficult for that model to extrapolate anywhere outside its training domain, very susceptible to novel inputs, lacks the guarantees associated with traditional models (long term debt)

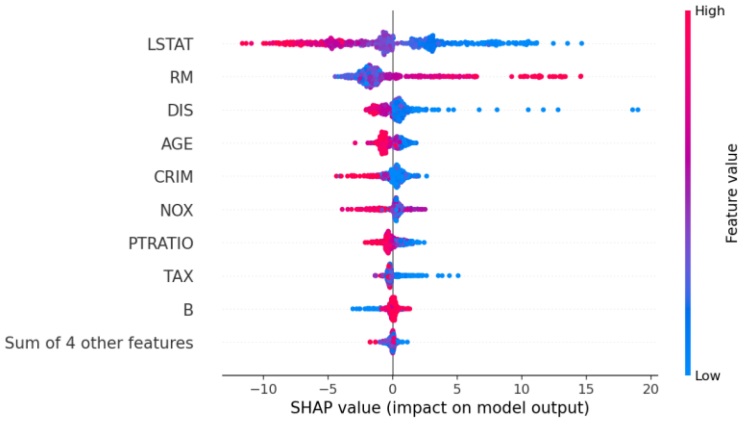
What would be better?

- Unsupervised learning (without labeling, set y by itself)
- Symbol-manipulation, and the need for hybrid models
- More insight from cognitive and developmental psychology
- Bolder challenges

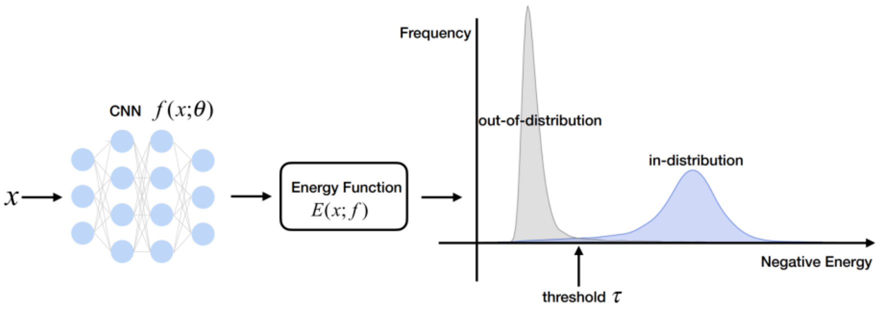
Discussion Questions #2

- Which issue/s in the paper do you think are/is the most serious problem?

- Which of these areas has machine learning made the most significant advances in recently?



https://github.com/slundberg/shap



Liu, Weitang, et al. "Energy-based out-of-distribution detection." *Advances in Neural Information Processing Systems* 33 (2020): 21464-21475.

Discussion Questions #3

- Marcus argues the deep learning has been "overhyped" and that as the field progresses, we are beginning to see diminishing returns in many areas. How long do you think deep learning will continue to be so dominant in the field of AI? Do you expect some other methods will replace it?

Discussion Questions #4

- What did you think that the author got wrong in his analysis?
- Are there any major problems with deep learning that you think he missed?
- Any other comments or discussion ideas?