Module V

Deep Learning
Limitations & Conclusions

Backpropagation

Resources for Implementing the Backprop. Algorithm

1. “Yes you should understand backprop” by A. Karpathy

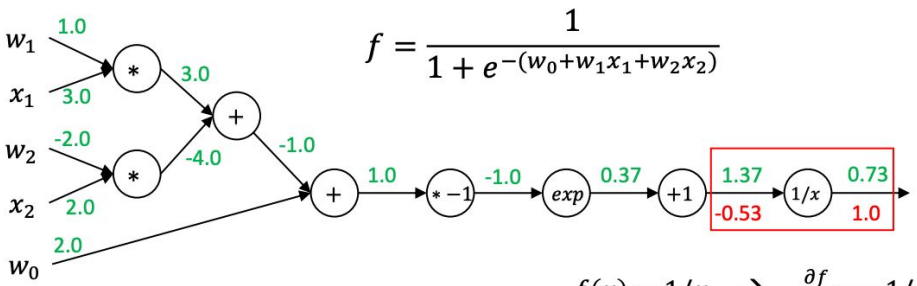
   https://medium.com/@karpathy/yes-you-should-understand-backprop-e2f06eab496b

2. CSE 599G1: Deep Learning System, University of Washington

   https://dlsys.cs.washington.edu/pdf/lecture4.pdf

3. A. Karpathy’s - CS231n Winter 2016: Lecture 4: Backpropagation, Neural Networks 1

   https://www.youtube.com/watch?v=i94OvYb6noo (this specific video)



symbolic differentiation

Goals of Deep Learning

Current Goal of DL: Replace Mental Labor

Physical Labor

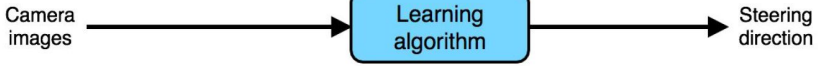
Mental Labor

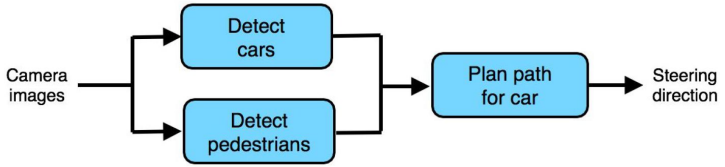
The Value of Feature Engineering

Feature Engineering Is Important for Many Problems

When NOT to use end-to-end systems?

- When we do not have a lot of hand engineered features end-to-end may not be the best solution
  - For example, a lot of hand engineered features text or audio data enable end-to-end
- And, when we do not have enough data to train the model
  - We need a lot of [image, steering] data for end-to-end automatic driving
  - But we may still have enough data to train individual components



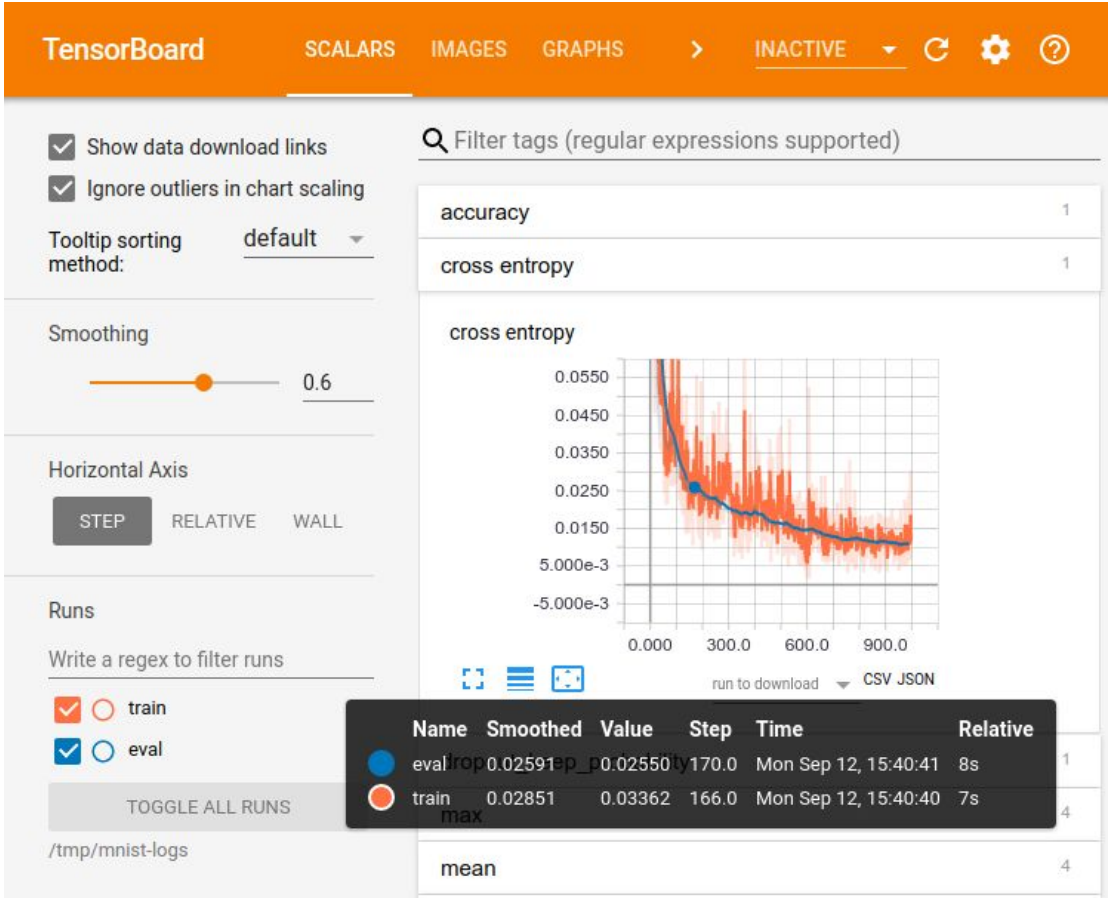


Another Example:

- Correctly handling missing data when the data is limited!

Are Deep Learning Models Black Boxes?

Deep Learning Models are not ‘always’ Black Boxes



Deep Learning Models are not ‘always’ Black Boxes

African elephants detected correctly by VGG16

Apply activation heatmap back to the original image

Limitations of Deep Learning

The state of Computer Vision and AI: we are really, really far away

http://karpathy.github.io/2012/10/22/state-of-computer-vision/

Some of the things "we" understand easily

There are 3 mirrors in the scene so some of those people are "fake" replicas from different viewpoints

Recognize Obama from the few pixels that make up his face

You recognize that there's a person standing on a scale, even though the scale occupies only very few white pixels that blend with the background

Obama has his foot positioned just slightly on top of the scale

Working physics - Obama is leaning in on the scale, which applies a force on it. Scale measures force that is applied on it, that's how it works => it will over-estimate the weight of the person standing on it.

The person measuring his weight is not aware of Obama doing this

There are people in the back who find the person's imminent confusion funny

Limitations of DL

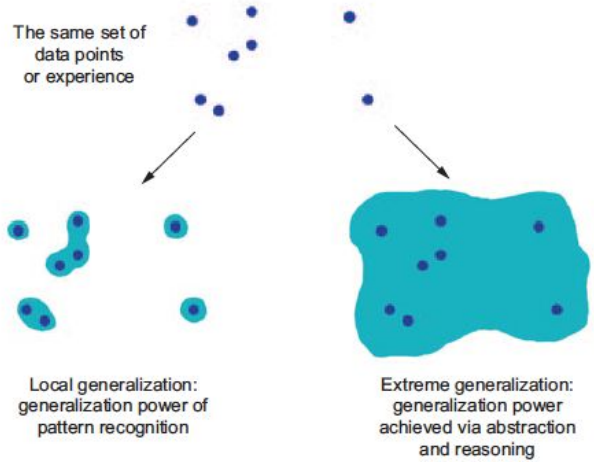
- Deep learning model is just a chain of simple continuous geometric transformations mapping one vector space into another
- All it can do is map one data manifold X into another manifold Y
  - assuming the existence of learnable continuous transform from X to Y

Limitations of DL

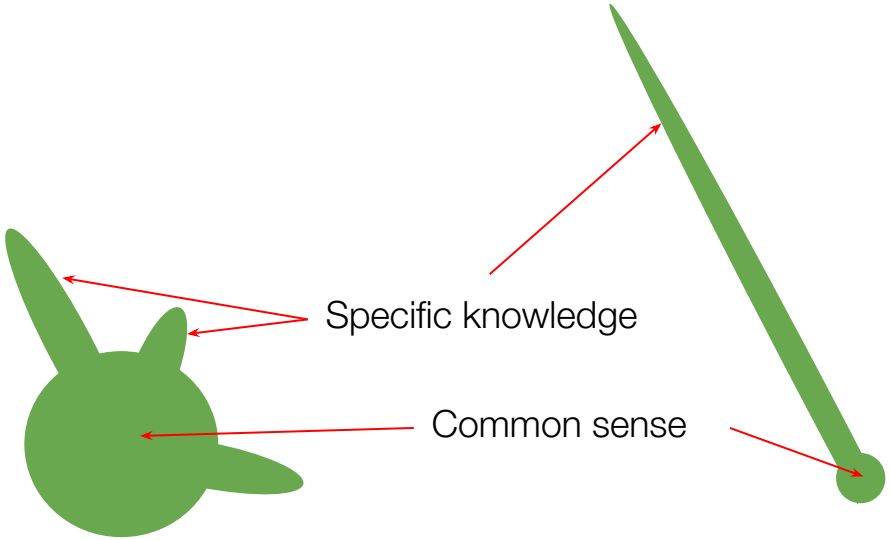
- A deep learning model can be interpreted as a kind of program; **but inversely most programs can't be expressed as deep learning models**
  - algorithm ≠ deep learning model
- For most tasks, either there exists no corresponding deep-neural network that solves the task or, even if one exists, it may not be learnable
  - The corresponding geometric transform may be far too complex, or there may not be appropriate data available to learn it

Limitations of DL

- Extreme generalization vs Local generalization
  - Extreme generalization: an ability to adapt to novel, never-before-experienced situations using little data or even no new data at all (abstraction and reasoning)
  - Local generalization: mapping from inputs to outputs
- Example:
  - If we train a model to distinguish between various kinds of cats
    - The model does not understand that those are all cats



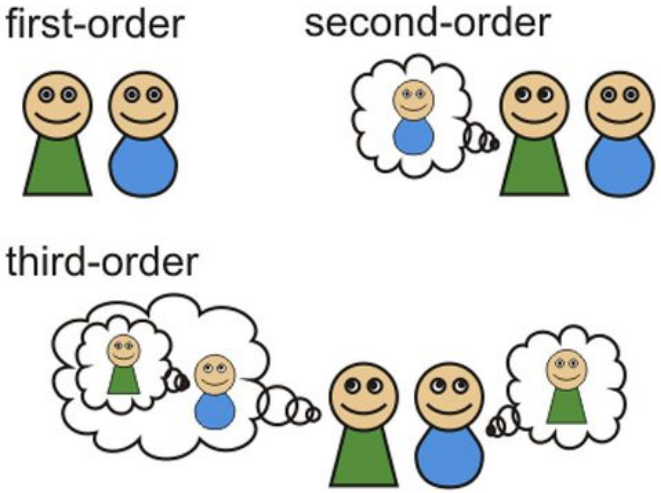
Common sense vs Specific knowledge



Human Mind **Deep Learning Model**

Do DL Models Have Human Characteristics?

- A fundamental feature of humans is our “theory of mind”
  - our tendency to project intentions, beliefs, and knowledge on the things around us

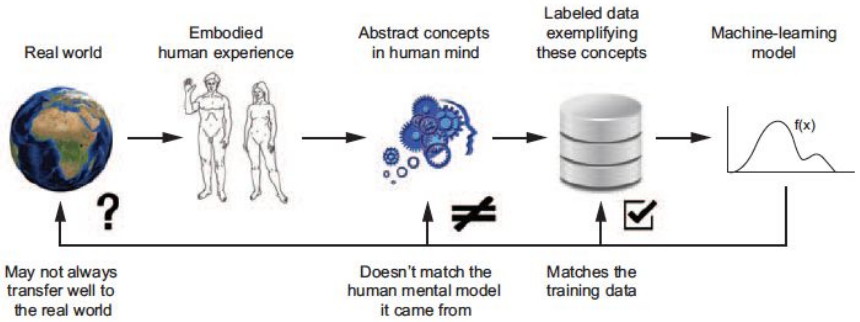


Say, we successfully train a model to generate captions to describe pictures

Does the model “understand” the contents of the picture and the captions it generates?

Do DL Models Have Human Characteristics?

- Deep-learning models don't have any understanding of their input
  - at least, not in a human sense
- Understanding of images/sounds/language is grounded in our sensorimotor experience
  - Machine-learning models have no access to such experiences
  - They can't understand their inputs in a human-relatable way
- We can get them to learn a geometric transform that maps data to human concepts on "a specific set of examples"
  - but this mapping is a simplistic sketch of the original model in our minds—the one developed from our experience as embodied agents



The Human Brain (may not be totally accurate!)

- The human brain contains about 100 billion closely connected neurons
- Around 1.5 *10^{14} synapses in the human brain
- The brain would only be equivalent to roughly 10000 VGG16's

https://blog.piekniewski.info/2018/08/28/fun-numbers-about-the-brain/

This is Just the Beginning of Learning Deep Learning..

Topics in Deep Learning

"But it is not complicated. It is just a lot of it."
- Richard Feynman

“A Little Learning” by Alexander Pope

A little learning is a dangerous thing;
Drink deep, or taste not the Pierian spring:
There shallow draughts intoxicate the brain,
And drinking largely sobers us again.