Shaun Campbell, Brennan McKendree, Nic Ammazzalorso, Andrew Takahashi

9/24/19

What's Wrong with AI and ML

* Legit issues
  + Have been tackled by ML researchers
  + Led to different emerging fields in these area
* Depending on your sources the issues you read might be different
  + Lot of noise in the news
  + Several illegitimate issues

Concern 1: AI is going to take away all our jobs

* case in point:
  + Manufacturing assembly lines
    - Past: Humans
    - Now/Future: Machines or AI
    - Car Assembly Lines
  + Cashiers at Fast Food/Grocery Stores
  + Communication for societies
    - Face to face -> telephones -> social media websites (WhatsApp, Facebook, etc.)
  + Taxis and Ubers
    - Truck Drivers - Otto
  + Cleaning Services
    - Roomba
  + Marketing and Advertising
    - Ad exchanges
  + Robots to check inventory
    - Amazon Go
  + Stock Markets (NYSE, Nasdaq)
  + Call Center Operatives (IVRS systems)

* MarketWatch 2017
* Robots are going to take all our jobs in the next 10 or 20 years
* 1 million grounds and maintenance workers - current
  + 50,000 after 20 years
* No proof of this statement yet

Concern 2: Artificial General Intelligence is Near

* We will build autonomous agents that operate much like being in the world
  + Lots of new that AGI is just around the corner
* Modern day AGI research is not doing well at all
* Mostly seems stuck on the same issues in reasoning and common sense that AI has had problem with for the past 50 years
* Case in Point: Self Driving Cars
  + Waymo - acquired by Google in 2016
  + Self-driving cars are going to take at least 30-50 years for us to make it a reality
  + Lower bound on AGI, but even self-driving cars are going to take 30-50 years

Concern 3: The Singularity is Near

* 2029 is when we would be able to simulate the function of the entire brain
  + Millions of neurons cells, and billions of connections within these cells
* refers to a point where AI is better at AI research than humans
  + It will recursively improve itself
  + Will no longer be in control of human beings
* Current State:
  + AI systems trying to understand a 100 line C++ code
  + Unable to beat a freshman student who has just taken one month of programming lessons
* C Elegans
  + Nervous system of this worm has 302 neurons and 6000 connections in between these neurons
  + Over the past 30 years, people have been figuring out the entire wiring pattern of the 302 neurons
  + Modeling the neural system of C Elegans I still on going and not even halfway there

Concern 4: Misaligned Values of AI and ML

* What if you design an AI agent for making good coffee for you
  + It realizes it cannot fulfill its goal if it is turn off
  + Disable its off switch

Concern 5: Terminator robots are going to kill us



9/26/19

Issues with Deep Learning

* Can DL approach human level intelligence?

Is Deep Learning Approaching a Wall

* For most problems where deep learning has enabled transformationally better solutions (visons, speeches) we've entered diminishing returns territory in 2016-2017

What is Deep Learning Good At?

* Just a statistical technique
* Has a set of assumptions that it works with
* Performance is not good when these assumptions are not satisfied
* What?
  + Enough data (more an issue with deep learning, and less with standard ML)
    - Deep Learning can work with raw data
    - Standard ML models extract "important" features from this raw data
      * Usually happens using a hand-designed feature extractor
  + No bias in training data
    - DL models are just as likely to suffer against biased data
  + Computation power needs to be high for DL models
  + Data from the wild (real world) should be similar to your training data
    - Training data should be a good enough representation for the kind of data that you are likely to see in the real world
    - The distributions of your training and test data should be the same (or highly similar)

First Limit - Deep Learning is Data Hungry

* If you have training data -> DL works well
* Contrapositive of this statement
  + DL doesn’t work well -> you don’t have enough data
    - Data augmentations
  + Lots of example in AI for Social Good Domain
    - Collecting data of homeless youth social network to spread awareness about HIV
    - AI generated patrolling schedules to protect against terrorist attack on LAX - ARMOR program

* Test data should be similar to training data
* Interpolation
  + If test data is coming from the same distribution, your DL model should be able to interpolate between things that it has seen before
  + Not exactly the same but similar
* Extrapolation
  + If test data is not coming from the same distribution, DL model needs to extrapolate knowledge that it has currently learnt
  + When it is completely different, not seen before
* **Important: No way to extrapolate**

Second Limit - DL is Shallow

* Does not learn any hidden abstractions similar to human beings
  + These abstractions allow us to transfer knowledge
  + DL can't do that

Limit 3: No Way to Deal with Hierarchical Structure

* RNNs represents sentences as sequences of words
  + Ignore hierarchical structure
  + Longer sentences constructed recursively user small sub-sentences
* Example: The teenager who previously crossed the Atlantic ocean set a record for flying around the world
* Issue: No hierarchy among set of features, all of them are flat, we draw correlations among them
  + Hierarchical structures among feature are not represented inside DL
* As a result, use proxies for this hierarchy
  + E.g. sequence of words

Limit 4: Open Ended Inference

* Inference has been limited to Squad (Stanford Question Answer Database) type queries
* Given a question and a piece of text
  + Infer answer to question by reading text
    - Assumption: answer is present in text
* Thing that have not been done:
* Multi-hop inference
  + Locate answers by combining multiple pieces of text
  + Combine text with background knowledge
  + Open Ended Inference example: I think you need to mine your own business
    - Question: What is the mood of the person?
* Human beings can do this opened ended inference
* DL cannot

Limit 5: Lack of Transparency

* DL is a black box
* Millions or billions of weights
  + All you can get is the values of these learned weights
  + How to interpret them?
* Why is this even important? In what domains?
  + Viewpoint 1: Depends, if you are just looking for good results, don’t need transparency, but if you are scientists working at Google who want to understand better, you need transparency
  + Depends on the domain where its being used, if it's being used people health, people livelihoods, then you need to understand why is a DL model making some prediction
  + Practitioners need to be able to trust the ML system that they are using
    - Who is accountable when a ML makes a mistake? The ML model goes scot-free but the doctor gets sued

Limit 6: Not Integrated with prior knowledge

* No domain knowledge is input
* Standard ML used feature extractors which were designed by human experts and contained human insights into the domain
* But you don't have human designed feature extractors in DL
* Useful properties of images, text, or whatever kind of data is being used is not present in the DL model
* One solid exception
  + CNNs

Limit 7: Unable to model causation

* Correlation does not imply causation
* DL system can learn correlations between height and vocabulary
* Will not be able to uncover causation between growth and development to both these variables

Limit 8: Assumption of Stationarity

* DL works well with stationary environments
* What if rules of the world continuously change?
  + What about stock prediction? Flu prediction?
* How is this related to extrapolation and difference in training testing data?

Limit 9: DL can easily be fooled