[redacted]

Deep Learning: State of the Art (2020)

Deep Learning Lecture Series

Deep Learning State of the Art (2020)
Lex Fridman
MIT
Date/Time: Mon, Jan 6, 3-4:30pm
Room: E54-100

Unlocking AI for Science by Preserving Privacy
Andrew Trask
OpenMined, University of Oxford
Date/Time: Wed, Jan 8, 3-4:30pm
Room: E54-100

Complete Statistical Theory of Learning
Vladimir Vapnik
Co-Creator of SVM, Columbia University
Date/Time: Fri, Jan 10, 3-5pm
Room: E54-100

Measures of Intelligence
Francois Chollet
Creator of Keras, Google
Date/Time: Mon, Jan 13, 3-4:30pm
Room: E54-100

Deep Learning Hardware Accelerators
Vivienne Sze
MIT
Date/Time: Wed, Jan 15, 3-4:30pm
Room: E54-100

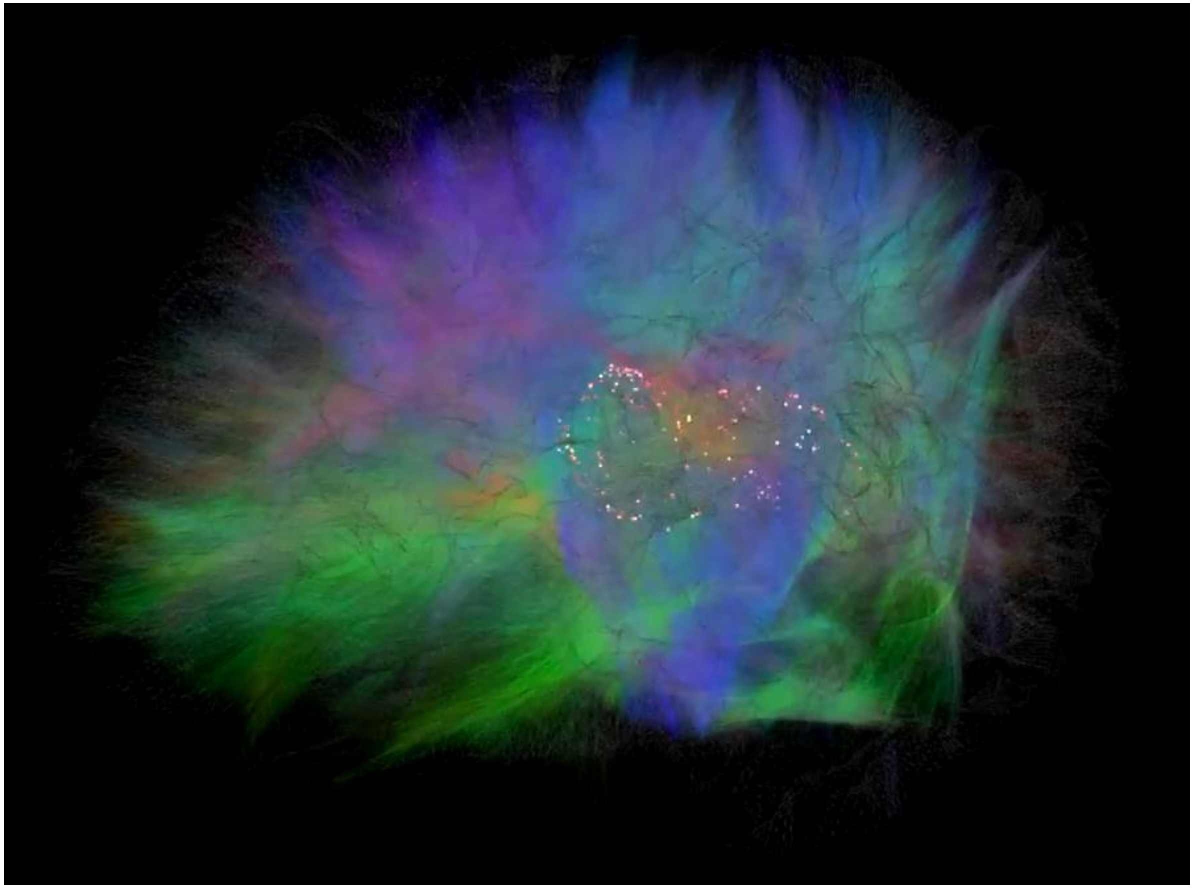
OpenAI Five
Jakub Pachocki
Dota Team & Reasoning Team Research Lead, OpenAI
Date/Time: Fri, Jan 17, 3-4:30pm
Room: E54-100

Outline

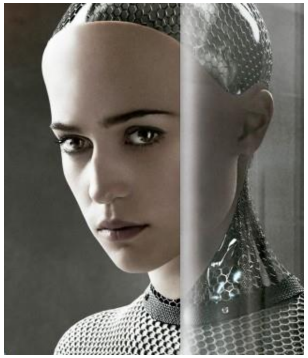
- **Deep Learning Growth, Celebrations, and Limitations**
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

“AI began with an ancient wish to forge the gods.”

\- Pamela McCorduck, *Machines Who Think, 1979*



Frankenstein (1818)



Ex Machina (2015)

Visualized here are 3% of the neurons and 0.0001% of the synapses in the brain.

Thalamocortical system visualization via DigiCortex Engine.

Deep Learning & AI in Context of Human History

Perspective:

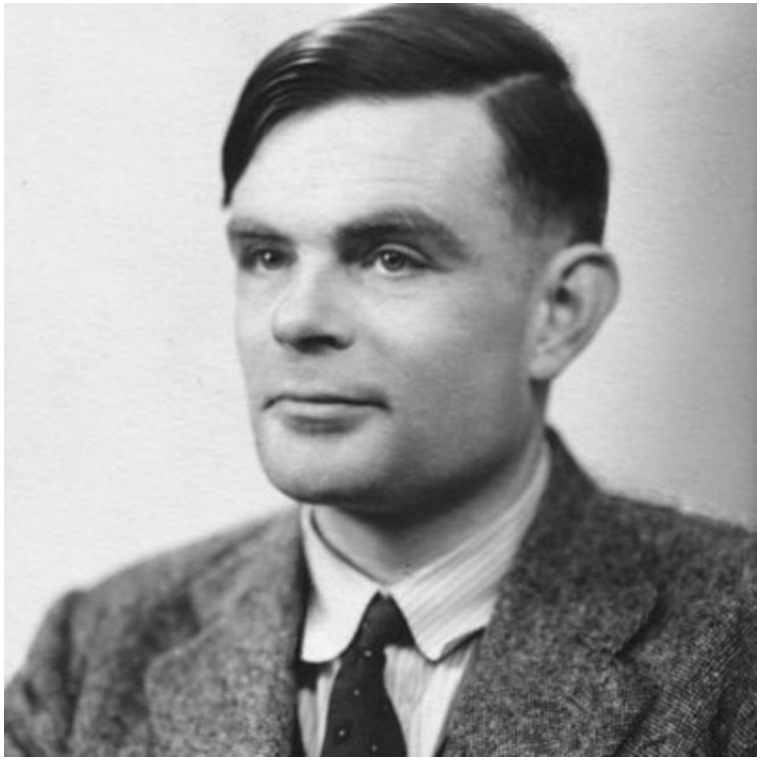
- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

1700s and beyond: Industrial revolution, steam engine, mechanized factory systems, machine tools

Artificial Intelligence in Context of Human History

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago



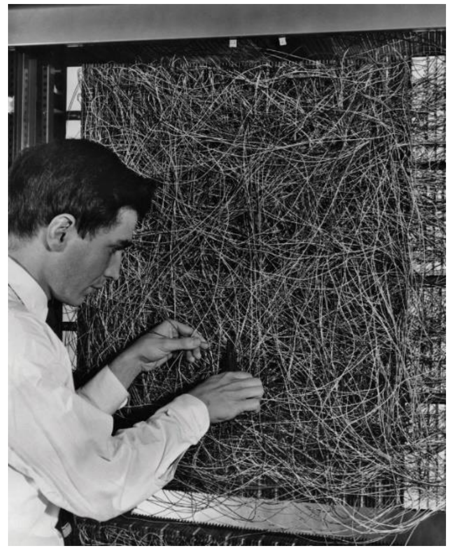
Dreams, mathematical foundations, and engineering in reality.

Alan Turing, 1951: “It seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers. They would be able to converse with each other to sharpen their wits. At some stage therefore, we should have to expect the machines to take control."

Artificial Intelligence in Context of Human History

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago



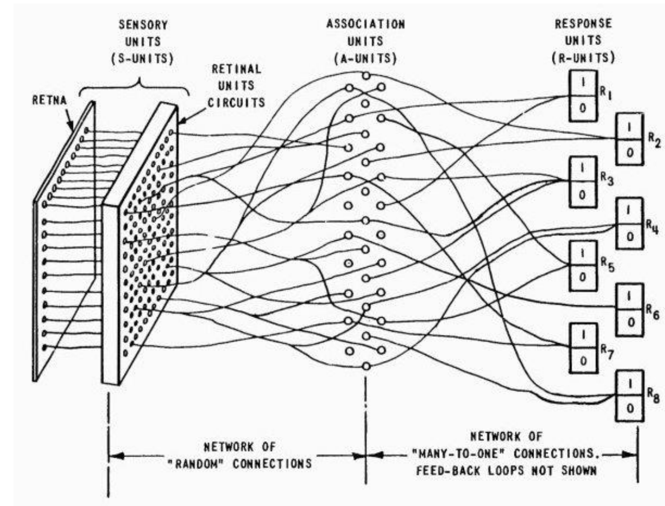


Figure I ORGANIZATION OF THE MARK I PERCEPTRON

Dreams, mathematical foundations, and engineering in reality.

Frank Rosenblatt, Perceptron (1957, 1962): Early description and engineering of single-layer and multi-layer artificial neural networks.

Artificial Intelligence in Context of Human History

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

Kasparov vs Deep Blue, 1997

Artificial Intelligence in Context of Human History

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

Lee Sedol vs AlphaGo, 2016

Artificial Intelligence in Context of Human History

Perspective:

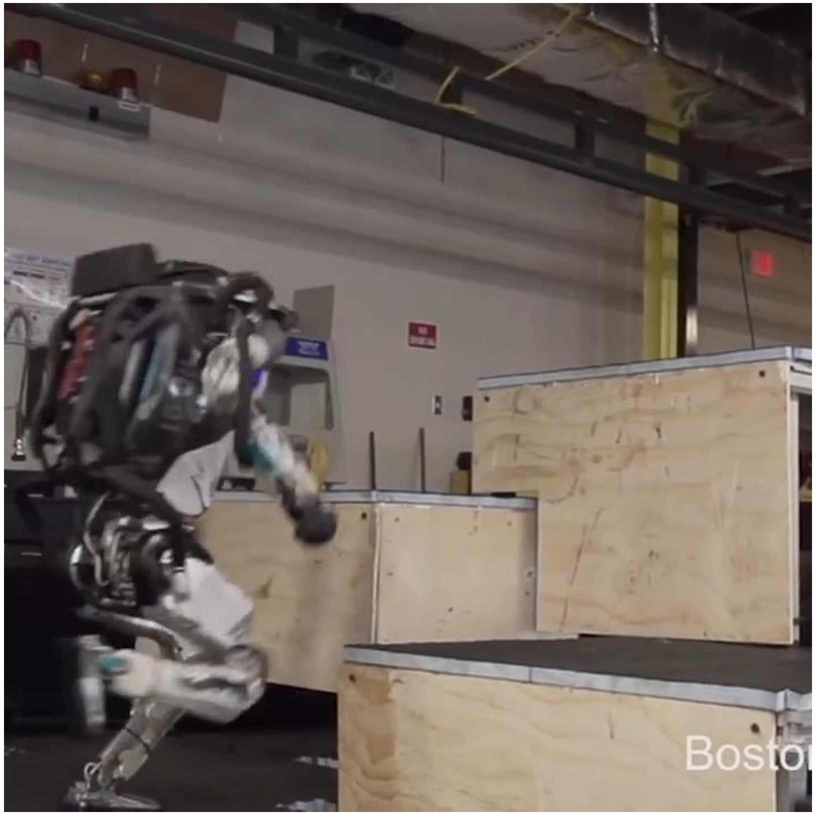
- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

Robots on four wheels.

Artificial Intelligence in Context of Human History

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago



Robots on two legs.

History of Deep Learning Ideas and Milestones*

Perspective:

- Universe created
  13.8 billion years ago
- Earth created
  4.54 billion years ago
- Modern humans
  300,000 years ago
- Civilization
  12,000 years ago
- Written record
  5,000 years ago

- 1943: Neural networks
- 1957-62: Perceptron
- 1970-86: Backpropagation, RBM, RNN
- 1979-98: CNN, MNIST, LSTM, Bidirectional RNN
- 2006: "Deep Learning", DBN
- 2009: ImageNet + AlexNet
- 2014: GANs
- 2016-17: AlphaGo, AlphaZero
- 2017: 2017-19: Transformers

* Dates are for perspective and not as definitive historical record of invention or credit

Turing Award for Deep Learning

- Yann LeCun
- Geoffrey Hinton
- Yoshua Bengio

Turing Award given for:

- "The conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing."
- (Also, for popularization in the face of skepticism.)

Early Key Figures in Deep Learning
(Not a Complete List by Any Means)

- 1943: Walter Pitts and Warren McCulloch
  Computational models for neural nets
- 1957, 1962: Frank Rosenblatt
  Perceptron (Single-Layer & Multi-Layer)
- 1965: Alexey Ivakhnenko and V. G. Lapa
  Learning algorithm for MLP
- 1970: Seppo Linnainmaa
  Backpropagation and automatic differentiation
- 1979: Kunihiko Fukushima
  Convolutional neural networks
- 1982: John Hopfield
  Hopfield networks (recurrent neural networks)

People of Deep Learning and Artificial Intelligence

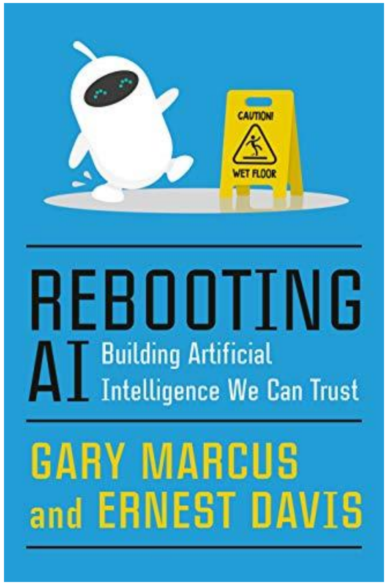
- History of science is a story of both **people** and **ideas**.
- Many brilliant people contributed to the development of AI.

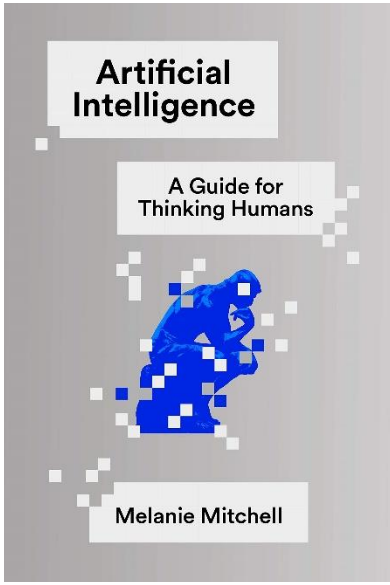
Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117
https://arxiv.org/pdf/1404.7828.pdf

My (Lex) hope for the community:

- More respect, open-mindedness, collaboration, credit sharing.
- Less derision, jealousy, stubbornness, academic silos.

Limitations of Deep Learning



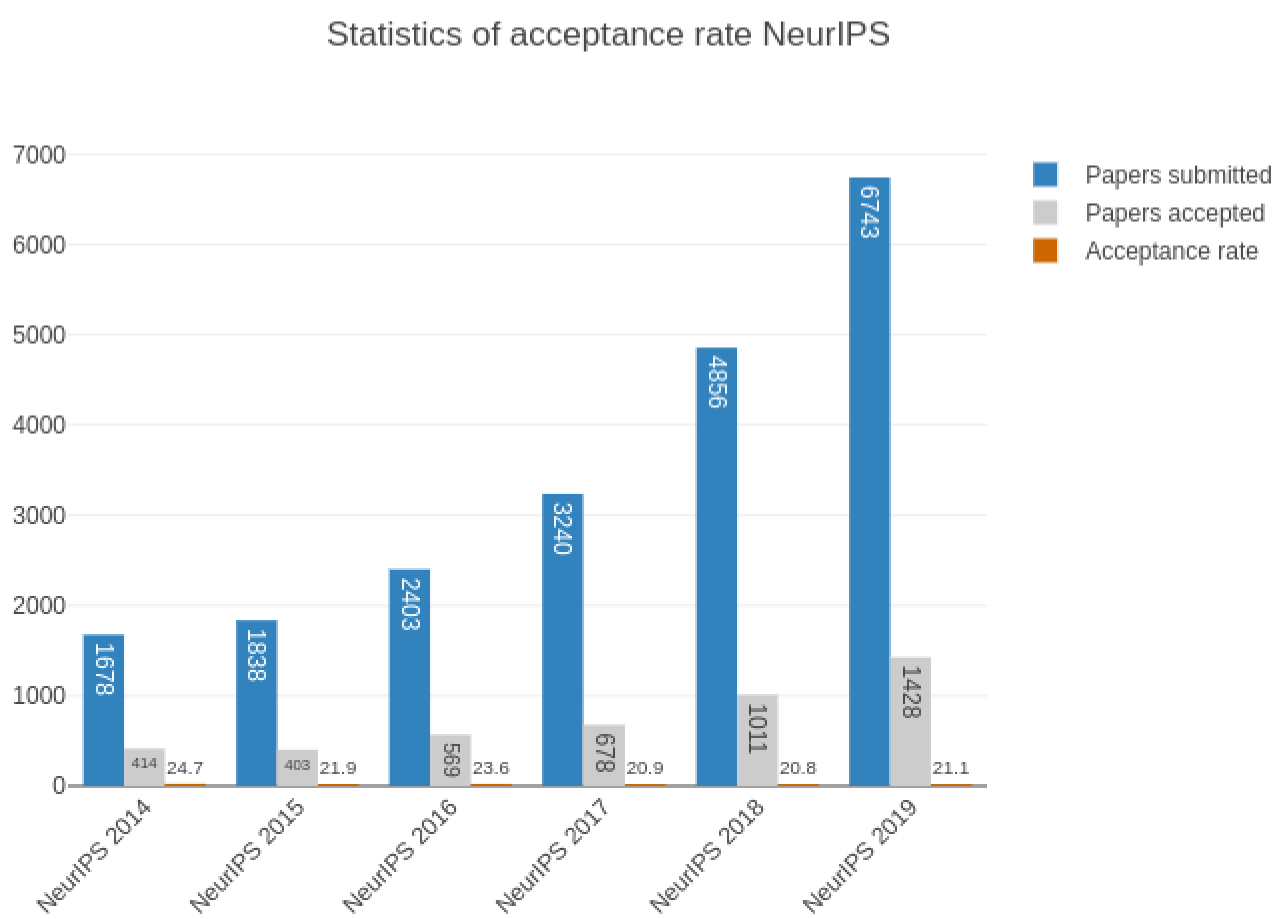


- 2019 is the year it became cool to say that "deep learning" has limitations.
- Books, articles, lectures, debates, videos were released that learning-based methods cannot do commonsense reasoning.

Prediction from Rodney Brooks:
"By 2020, the popular press starts having stories that the era of Deep Learning is over."

http://rodneybrooks.com/predictions-scorecard-2019-january-01/

Deep Learning Research Community is Growing



Hopes for 2020

- **Less Hype & Less Anti-Hype:** Less tweets on how there is too much hype in AI and more solid research in AI.
- **Hybrid Research:** Less contentious, counter-productive debates, more open-minded interdisciplinary collaboration.
- **Research topics:**
  - Reasoning
  - Active learning and life-long learning
  - Multi-modal and multi-task learning
  - Open-domain conversation
  - Applications: medical, autonomous vehicles
  - Algorithmic ethics
  - Robotics

Outline

- Deep Learning Growth, Celebrations, and Limitations
- **Deep Learning and Deep RL Frameworks**
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

TensorFlow 2.0 and PyTorch 1.3

TensorFlow

- Eager execution by default (imperative programming)
- Keras integration + promotion
- Cleanup (API, etc.)
- TensorFlow.js
- TensorFlow Lite
- TensorFlow Serving

PYTORCH

- TorchScript (graph representation)
- Quantization
- PyTorch Mobile (experimental)
- TPU support

Python 2 support ended on Jan 1, 2020.

```
>>> print "Goodbye World"
```

Reinforcement Learning Frameworks

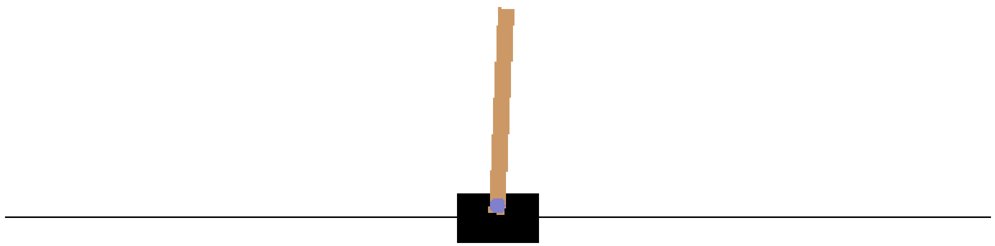
- TensorFlow
  - OpenAI Baselines
    - Stable Baselines – the one I recommend for beginners
  - TensorForce
  - Dopamine (Google)
  - TF-Agents
  - TRFL
  - RLLib (+ Tune) – great for distributed RL & hyperparameter tuning
  - Coach - huge selection of algorithms
- PyTorch
  - Horizon
  - SLM-Lab
- Misc
  - RLgraph
  - Keras-RL

Reinforcement Learning Frameworks

- "Stable Baselines" (OpenAI Baselines Fork)
  - A2C, PPO, TRPO, DQN, ACKTR, ACER and DDPG
  - Good documentation (and code commenting)
  - Easy to get started and use

```
from stable_baselines import PPO2

# Define and train a model in one line of code !
trained_model = PPO2('MlpPolicy', 'CartPole-v1').learn(total_timesteps=10000)
# you can then access the gym env using trained_model.get_env()
```



Hopes for 2020

- **Framework-agnostic Research:** Make it even easier to translate a trained PyTorch model to TensorFlow and vice-versa.

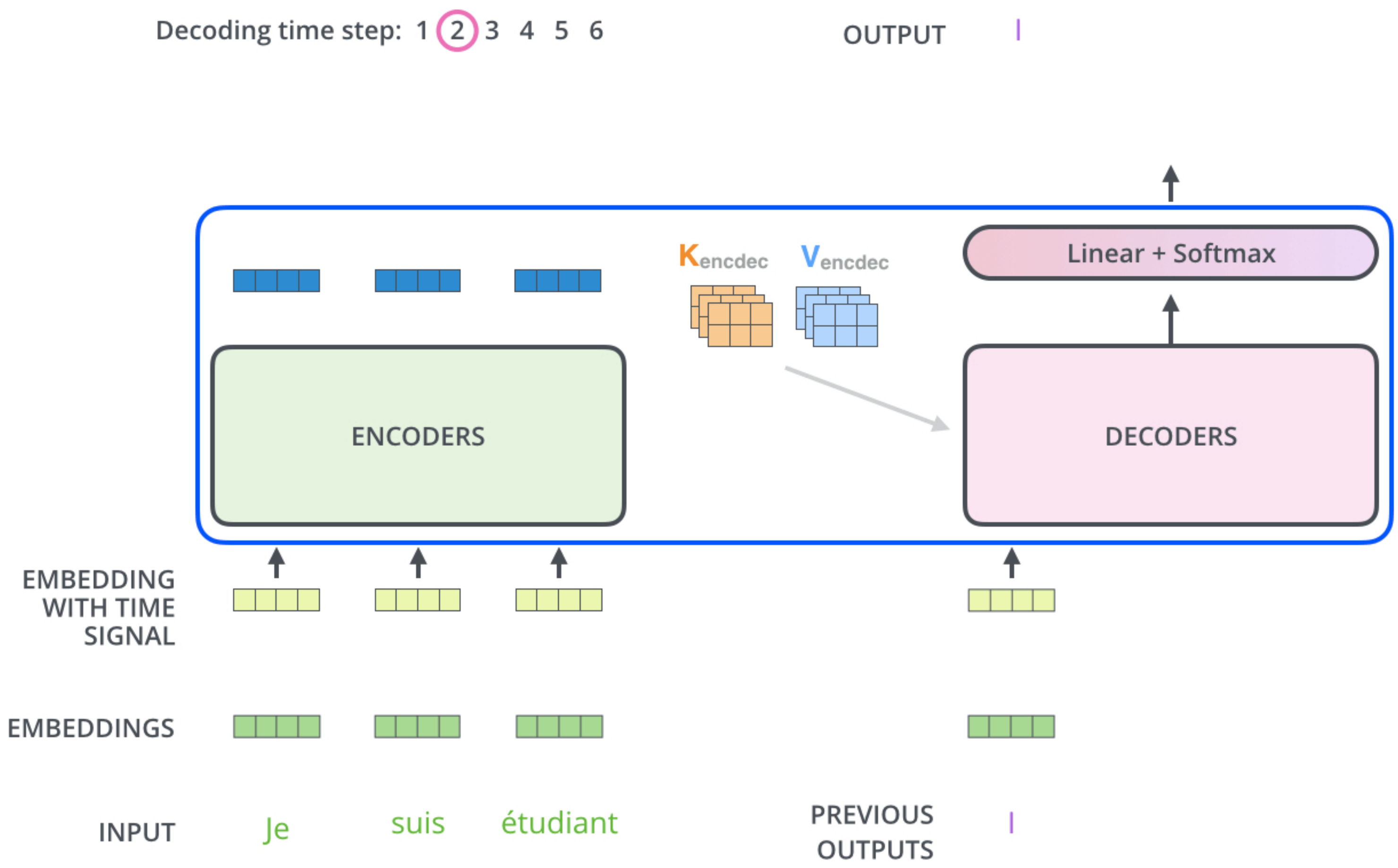
- **Mature Deep RL frameworks:** Converge to fewer, actively-developed, stable RL frameworks that are less tied to TensorFlow or PyTorch.

- **Abstractions**: Build higher and higher abstractions (i.e. Keras) on top of deep learning frameworks that empower researchers, scientists, developers outside of machine learning field.

Outline

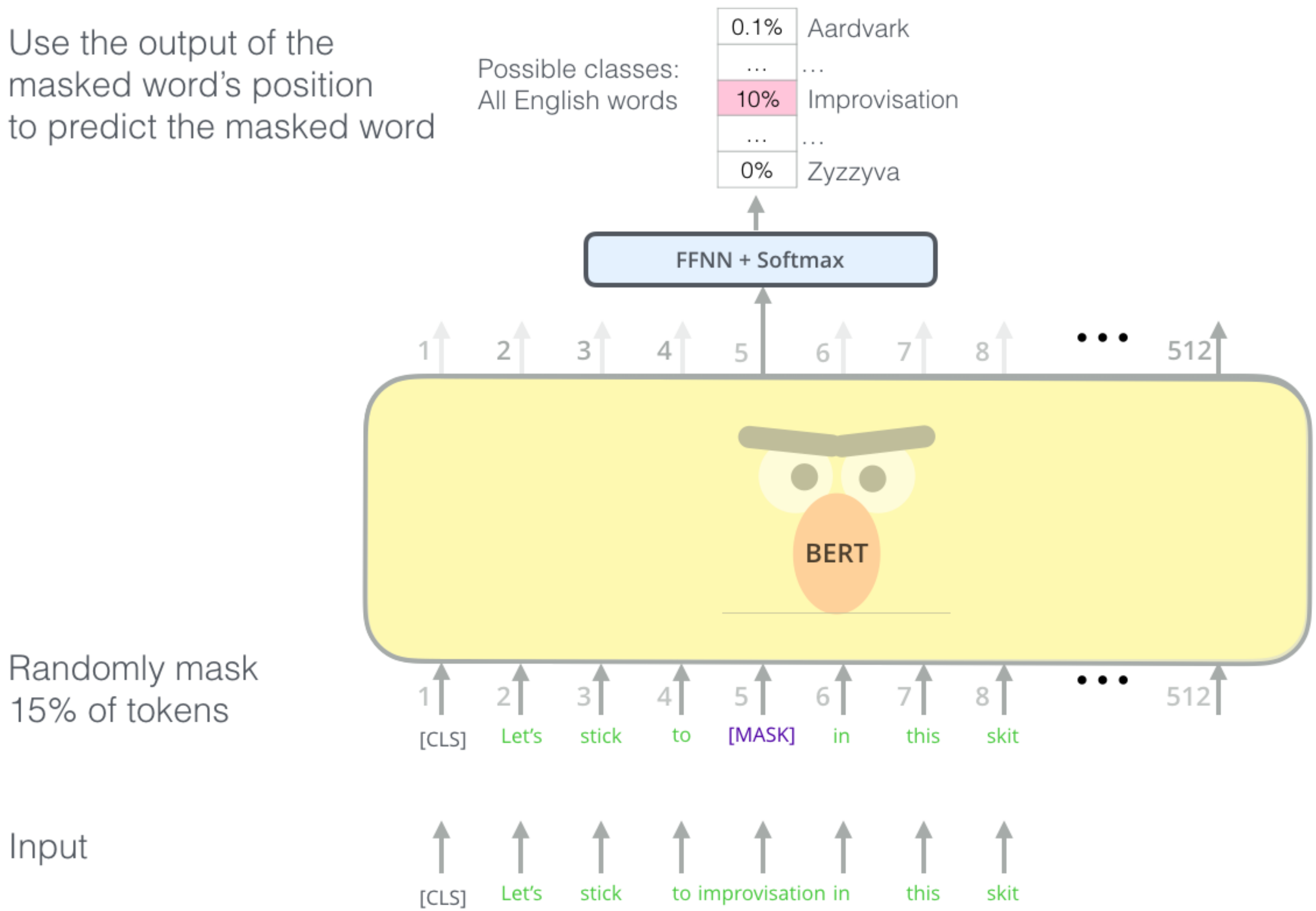
- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- **Natural Language Processing**
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

Transformer



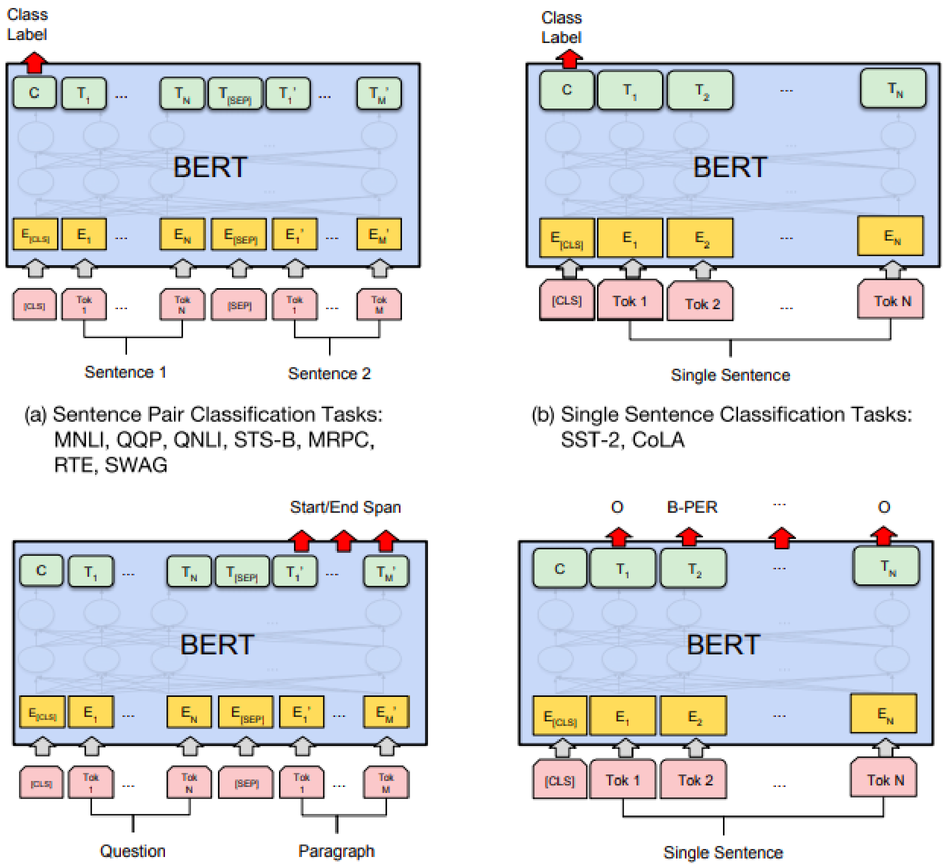
Vaswani et al. "Attention is all you need." *Advances in Neural Information Processing Systems*. 2017.

BERT



Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." (2018).

BERT Applications



(a) Sentence Pair Classification Tasks: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks: SST-2, CoLA

(c) Question Answering Tasks: SQuAD v1.1

(d) Single Sentence Tagging Tasks: CoNLL-2003 NER

Now you can use BERT:

- Create contextualized word embeddings (like ELMo)
- Sentence classification
- Sentence pair classification
- Sentence pair similarity
- Multiple choice
- Sentence tagging
- Question answering

Transformer-Based Language Models

- BERT (Google)
- XLNet (Google/CMU)
- RoBERTa (Facebook)
- DistilBERT (HuggingFace)
- CTRL (Salesforce)
- GPT-2 (OpenAI)
- ALBERT (Google)
- Megatron (NVIDIA)

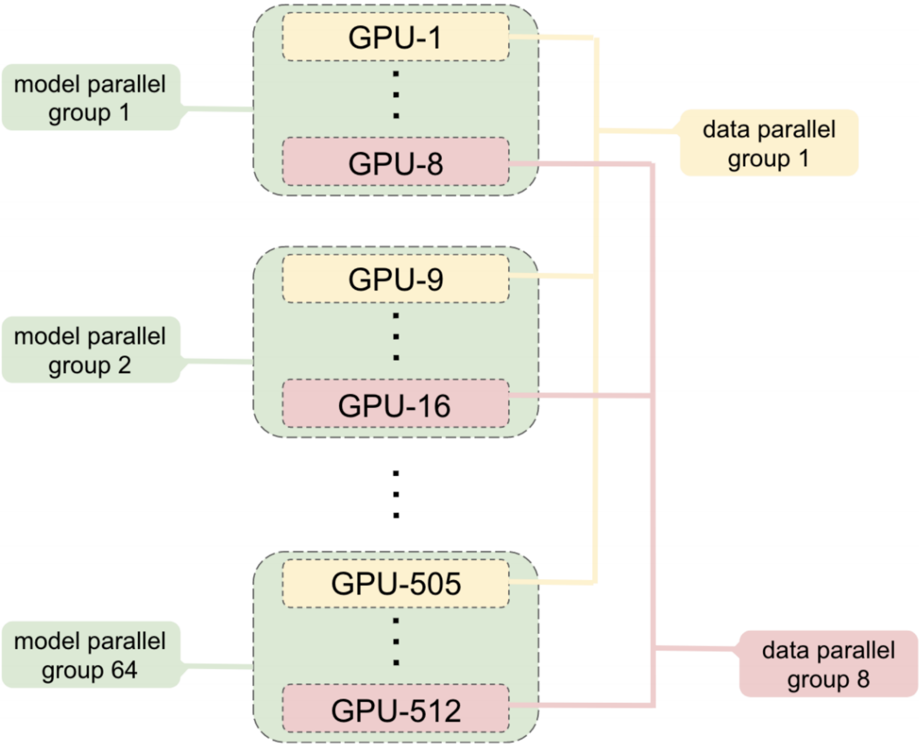
sebastianruder / **NLP-progress**

NLP-progress

Repository to track the progress in Natural Language Processing (NLP), including the datasets and the current state-of-the-art for the most common NLP tasks.

Megatron

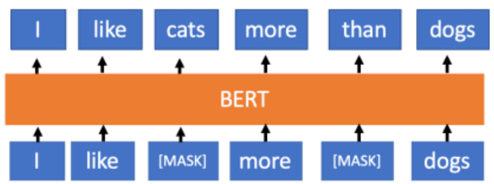
Shoeybi et al. (NVIDIA)

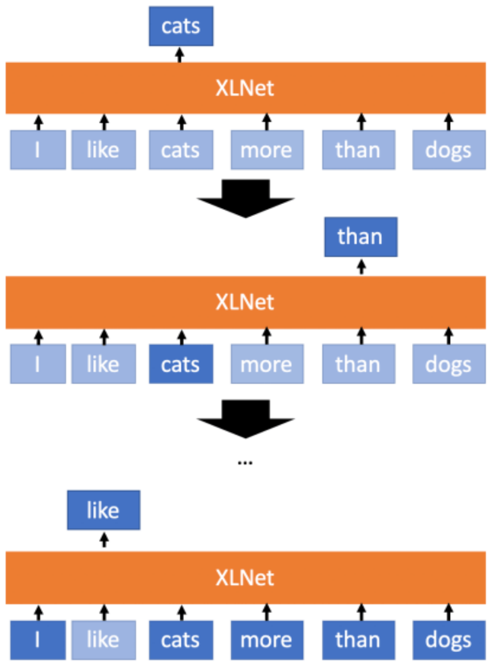


- Megatron-LM is a 8.3 billion parameter transformer language model with 8-way model parallelism and 64-way data parallelism trained on 512 GPUs (NVIDIA Tesla V100)
- Largest transformer model ever trained. 24x the size of BERT and 5.6x the size of GPT-2.

XLNET

Yang et al. (CMU, Google AI)

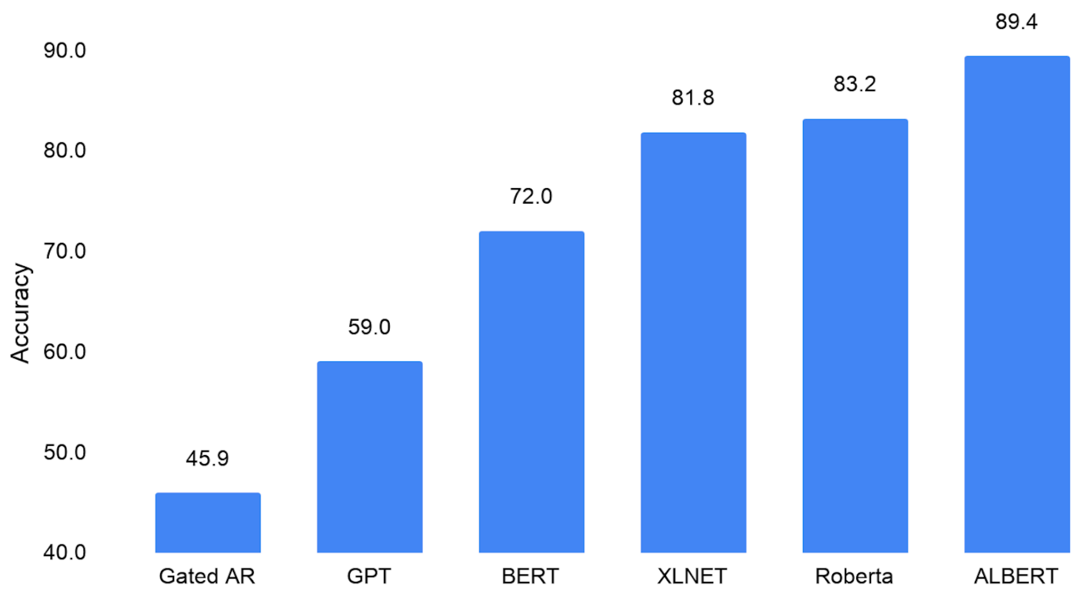




- Combine bidirectionality of BERT and the relative positional embeddings and the recurrence mechanism of Transformer-XL.
- XLnet outperforms BERT on 20 tasks, often by a large margin.
- The new model achieves state-of-the-art performance on 18 NLP tasks including question answering, natural language inference, sentiment analysis & document ranking.

ALBERT

Lan et al. (Google Research & Toyota Technological Institute at Chicago)



Machine performance on the RACE challenge (SAT-like reading comprehension).

- **Idea:** Reduces parameters via cross-layer parameter sharing
- **Results:** An upgrade to BERT that advances the state-of-the-art performance on 12 NLP tasks (including SQuAD2.0)
- **Code:** Open-source TensorFlow implementation, including a number of ready-to-use ALBERT pre-trained language models

Write With Transformer `gpt2`

The meaning of life is **not what I think it is, it's what I do to make it.**

The limits of deep learning are **still in the process of being figured out.**

Most important person in the history of deep learning is **probably Andrew Ng.**

Lex Fridman's best quality is that **he's smart, but I think he gets more attention than he deserves.**

Transformers Model Language
They Do Not **Understand** Language. Far from it (for now).

Two plus two is

- a three, five is a six, seven
- the result of a simple equation, and the
- four, and two plus three is four,

Two minus two is

- seven, plus two is six.
- a little too low for me.
- a perfect match for the numbers 1 to 20 in

OpenAI Report

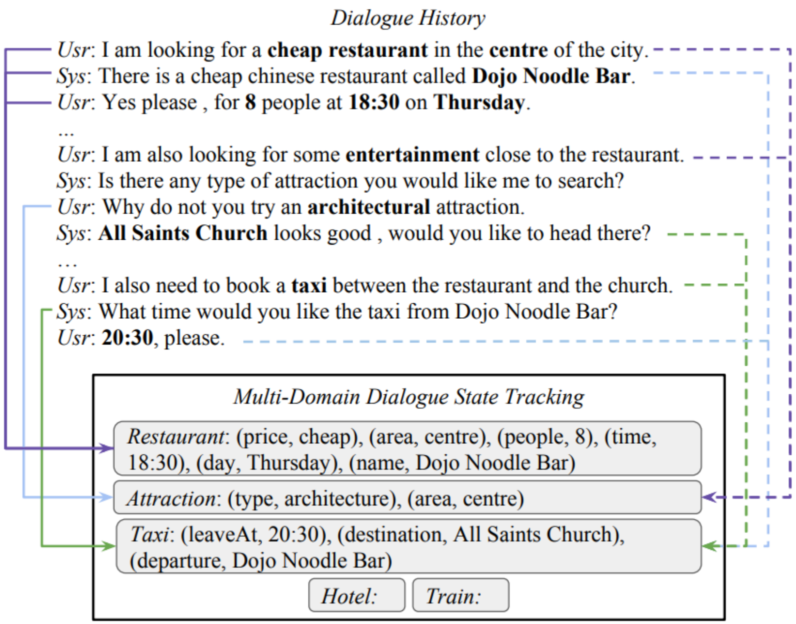
November, 2019

Release Strategies and the Social Impacts of Language Models

- Key takeaways in the report:
  - Coordination during model release between organization is difficult but possible
  - Humans can be convinced by synthetic text
  - Machine-based detection is difficult.
- My takeaways
  - Conversations on this topic are difficult, because the model of sharing between ML organizations and experts is mostly non-existent
  - Humans are still better at deception (disinformation) and detection in text and conversation (see Alexa prize slides)

Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems

Wu et al. (Honk Kong UST, Salesforce) – ACL 2019 Outstanding Paper



- **Task:** Dialogue state tracking
- **Problem:** Over-dependence on domain ontology and lack of knowledge sharing across domains
- **Details:**
  - Share model parameters across domains
  - Track slot values mentioned anywhere in a dialogue history with a context-enhanced slot gate and copy mechanism
  - Don't require a predefined ontology
- **Results:** State-of-the-art joint goal accuracy of 48.62% on MultiWOZ, a challenging 5-domain human-human dialogue dataset.

Explain Yourself! Leveraging Language Models for Commonsense Reasoning

Rajani et al. (Salesforce)

Question:	While eating a hamburger with friends, what are people trying to do?
Choices:	**have fun**, tasty, or indigestion
CoS-E:	Usually a hamburger with friends indicates a good time.
Question:	After getting drunk people couldn't understand him,it was because of his what?
Choices:	lower standards,**slurred speech**, or falling down
CoS-E:	People who are drunk have difficulty speaking.
Question:	People do what during their time off from work?
Choices:	**take trips**, brow shorter, or become hysterical
CoS-E:	People usually do something relaxing, such as taking trips,when they don't need to work.

Table 1: Examples from our CoS-E dataset.

Alexa Prize and Open Domain Conversations

- Amazon open-sourced the Topical-Chat dataset.
- Alexa Prize (like the Loebner Prize, etc) are teaching us valuable lessons (from the Alquist 2.0 team):
    - **Parts:** Break dialogue into small parts
    - **Tangents:** Create an interconnected graph of topics. Be ready to jump from context to context and back.
    - **Attention:** Not everything that is said is important. E.g.: *"You know, I'm a really terrible cook. But I would like to ask you, what's your favorite food?"*
    - **Opinions:** Create opinions.
    - **ML**
        - **Content:** ML is okay for generic chitchat, but nothing more for now.
        - **Classification:** ML classifies intent, finds entities or detects sentiment.
    - **Goal:** Goal is to maximize entertainment not information.

Alexa Prize and Open Domain Conversations

– Alquist: “Have you been in Brazil?”

– User: “What is the population of Brazil?”

– Alquist: “It is around 200 000 000.”

– User: “Well ok.”

– Alquist: “Have you been in Brazil?”

– User: “What is the population of Brazil?”

– Alquist: “It is around 200 000 000. Anyway, I was saying, have you been in Brazil?”

– User: “No, not yet.”

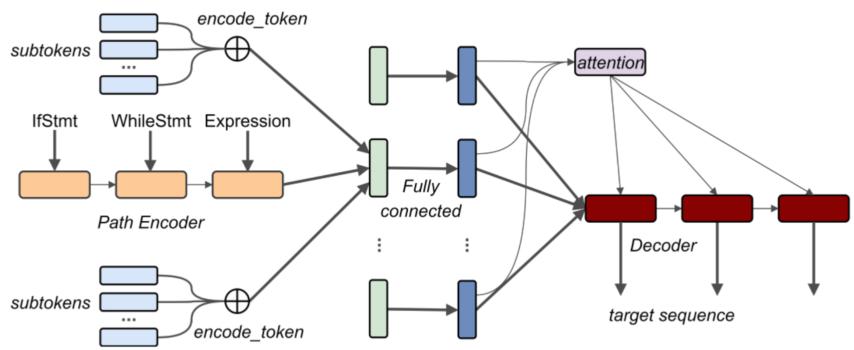
– Alquist: “Brazil is fascinating country…”

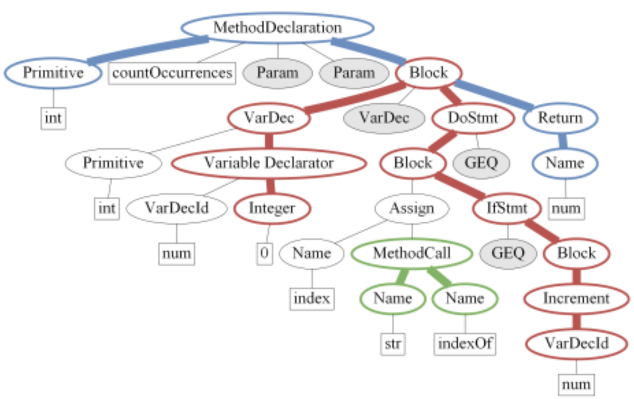
code2seq: Generating Sequences from Structured Representations of Code

Alon et al. (Technion) – ICLR 2019

- Instead of treating source code as a sequence of tokens, **code2seq** leverages the syntactic structure of programming languages to better encode source code as paths in its abstract syntax tree (AST).

```
int countOccurrences(String str, char ch) {
    int num = 0;
    int index = -1;
    do {
        index = str.indexOf(ch, index + 1);
        if (index >= 0) {
            num++;
        }
    } while (index >= 0);
    return num;
}
```






```
public static int ______(int value) {
  return value <= 0 ? 1 :
    value >= 0x40000000 ? 0x40000000 :
      1 << (32 - Integer.numberOfLeadingZeros(value - 1));
}
```

Model	Prediction
ConvAttention (Allamanis et al., 2016)	get
Paths+CRFs (Alon et al., 2018)	test bit inolz
code2vec (Alon et al., 2019)	multiply
2-layer BiLSTM (no token splitting)	next power of two
2-layer BiLSTM	{ *(replaced UNK)*
Transformer	get bit length
TreeLSTM (Tai et al., 2015)	get
Gold:	find next positive power of two
code2seq (this work)	get power of two

Natural Language Processing:

Hopes for 2020

- **Reasoning:** Combining (commonsense) reasoning with language models
- **Context:** Extending language model context to thousands of words.
- **Dialogue:** More focus on open-domain dialogue
- **Video:** Ideas and successes in self-supervised learning in visual data.

Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- **Deep RL and Self-Play**
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

OpenAI & Dota 2

- Dota 2 as a testbed for the messiness and continuous nature of the real world: teamwork, long time horizons, and hidden information.

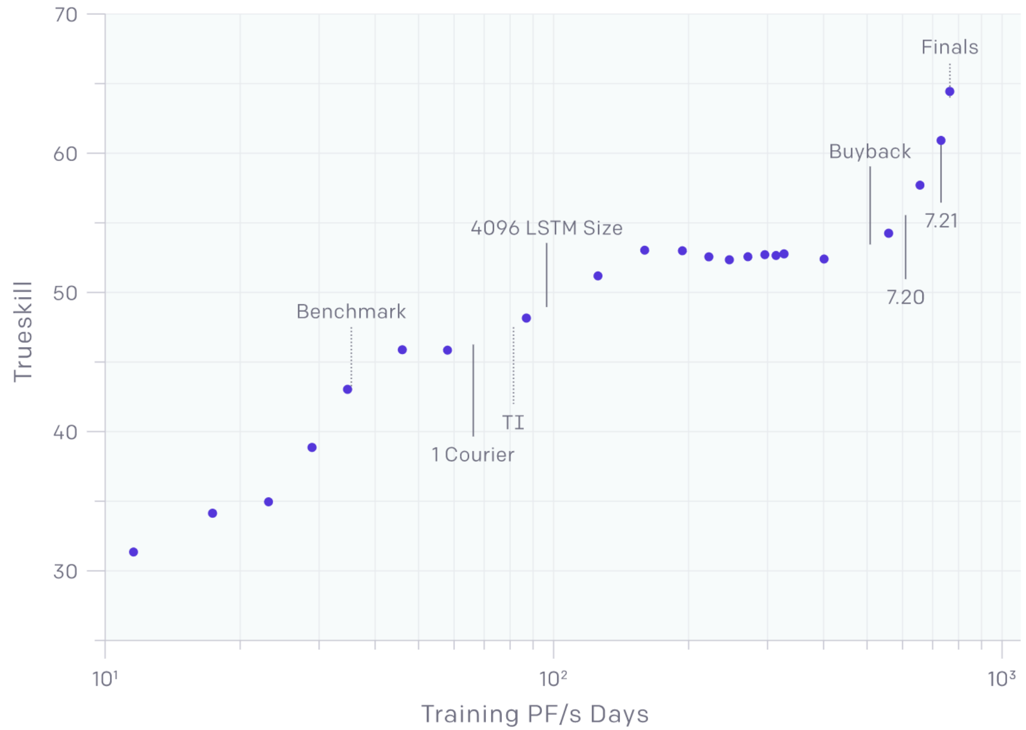
Place	Team	Prize money
1st	OG	$11,190,158
2nd	PSG.LGD	$4,069,148
3rd	Evil Geniuses	$2,670,379
4th	Team Liquid	$1,780,252
5th/6th	Team Secret	$1,144,448
	Virtus.pro	
7th/8th	OpTic Gaming	$635,804
	VGJ.Storm	
9th–12th	Mineski	$381,483
	Team Serenity	
	Vici Gaming	
	Winstrike Team	
13th–16th	Fnatic	$127,161
	Newbee	
	TNC Predator	
	VGJ.Thunder	
17th–18th	Invictus Gaming	$63,580
	paiN Gaming	

OpenAI & Dota 2 Progress

- Aug, 2017: 1v1 bot beats top professional Dota 2 players.
- Aug, 2018: OpenAI Five lost two games against top Dota 2 players at The International. "We are looking forward to pushing Five to the next level."
- Apr, 2019: OpenAI Five beats OG team (the 2018 world champions)

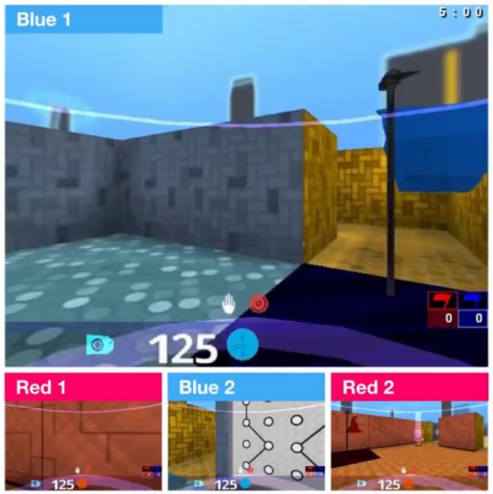
OpenAI & Dota 2 Progress

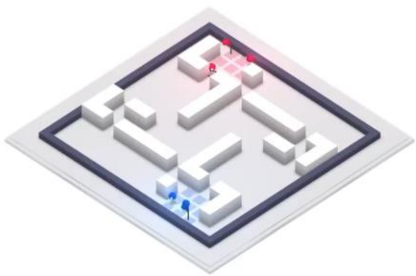
- **The Difference:** OpenAI Five's victories in 2019 as compared to its losses in 2018 are due to: 8x more training compute (training for longer)
- **Compute:** Current version of OpenAI Five has consumed 800 petaflop/s-days and experienced about 45,000 years of Dota self-play over 10 realtime months
- **Performance:** The 2019 version of OpenAI Five has a 99.9% win rate versus the 2018 version.



DeepMind Quake III Arena Capture the Flag

Agent observation raw pixels



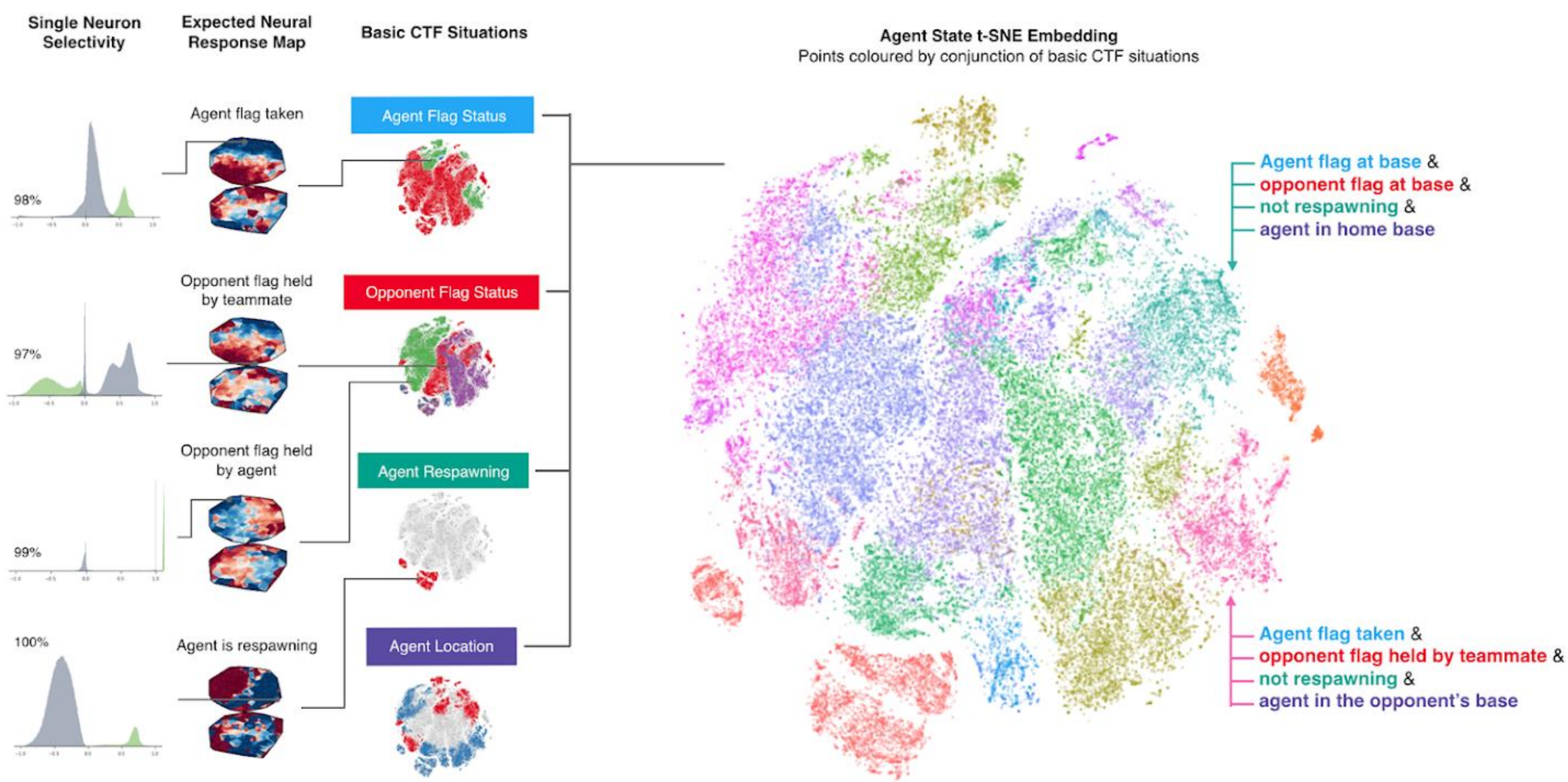


Indoor map overview

- “Billions of people inhabit the planet, each with their own individual goals and actions, but still capable of coming together through teams, organisations and societies in impressive displays of collective intelligence. This is a setting we call multi-agent learning: many individual agents must act independently, yet learn to interact and cooperate with other agents. This is an immensely difficult problem - because with co-adapting agents the world is constantly changing.”

DeepMind Quake III Arena Capture the Flag

- The agents automatically figure out the game rules, important concepts, behaviors, strategies, etc.

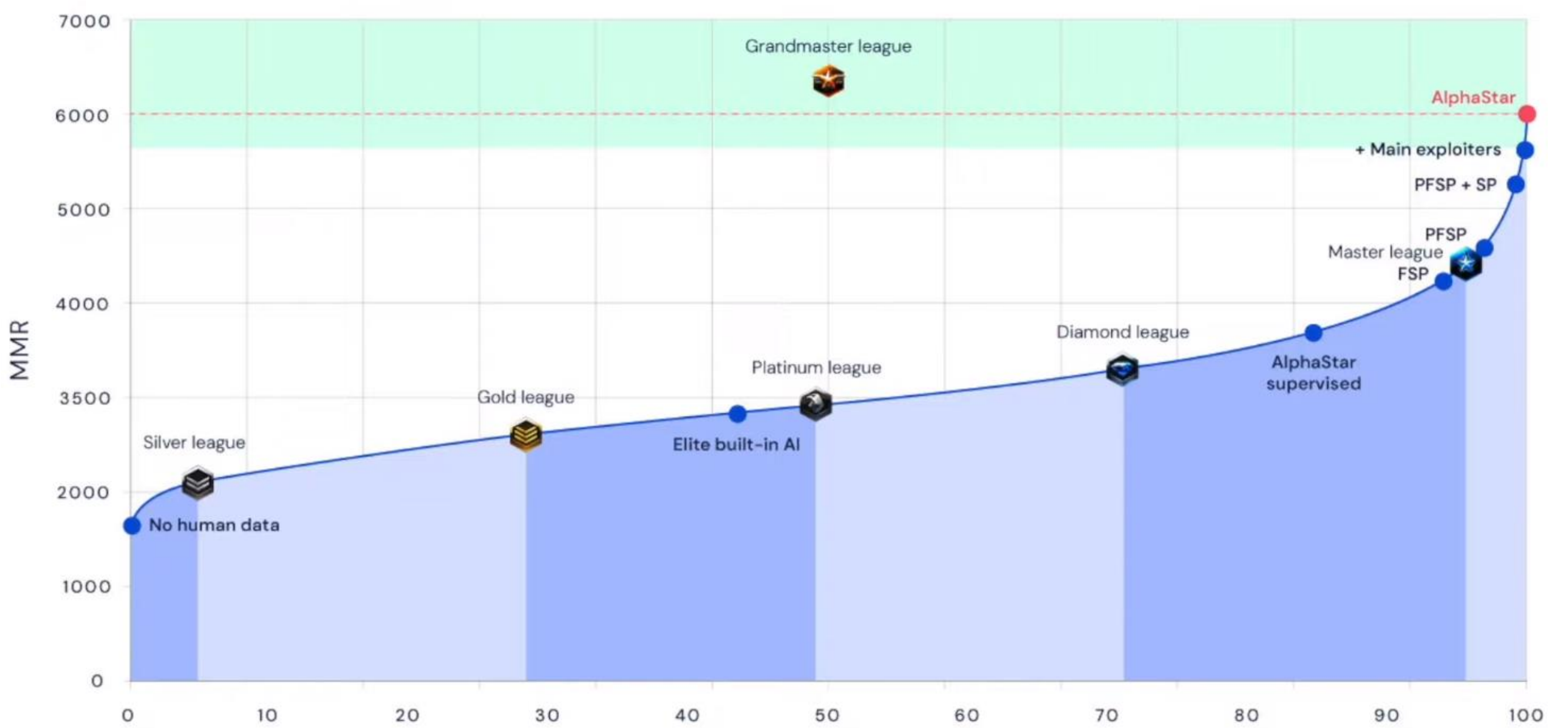


DeepMind Quake III Arena Capture the Flag

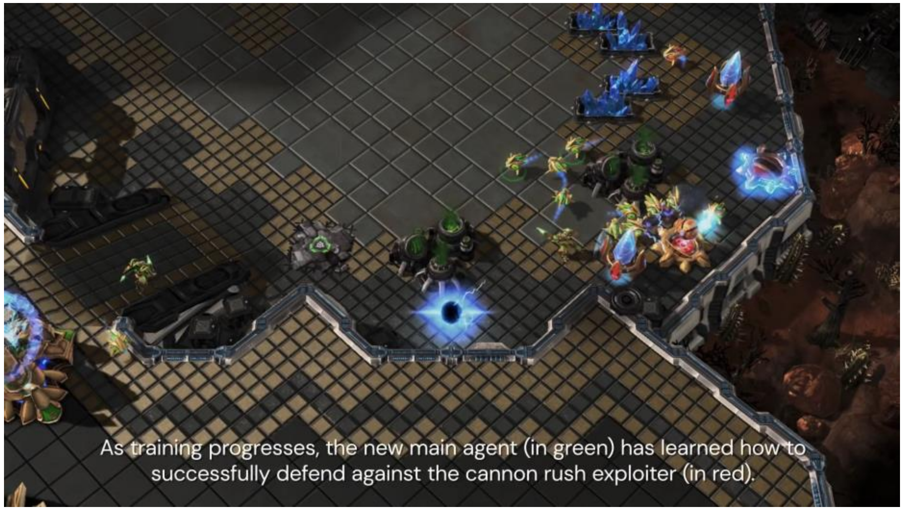
- The agents automatically figure out the game rules, important concepts, behaviors, strategies, etc.

DeepMind AlphaStar

- Dec, 2018: AlphaStar beats MaNa, one of the world's strongest professional StarCraft players, 5-0.
- Oct, 2019: AlphaStar reaches Grandmaster level playing the game under professionally approved conditions (for humans).



DeepMind AlphaStar



"AlphaStar is an intriguing and unorthodox player – one with the reflexes and speed of the best pros but strategies and a style that are entirely its own. The way AlphaStar was trained, with agents competing against each other in a league, has resulted in gameplay that's unimaginably unusual; it really makes you question how much of StarCraft's diverse possibilities pro players have really explored."

- Kelazhur, professional StarCraft II player

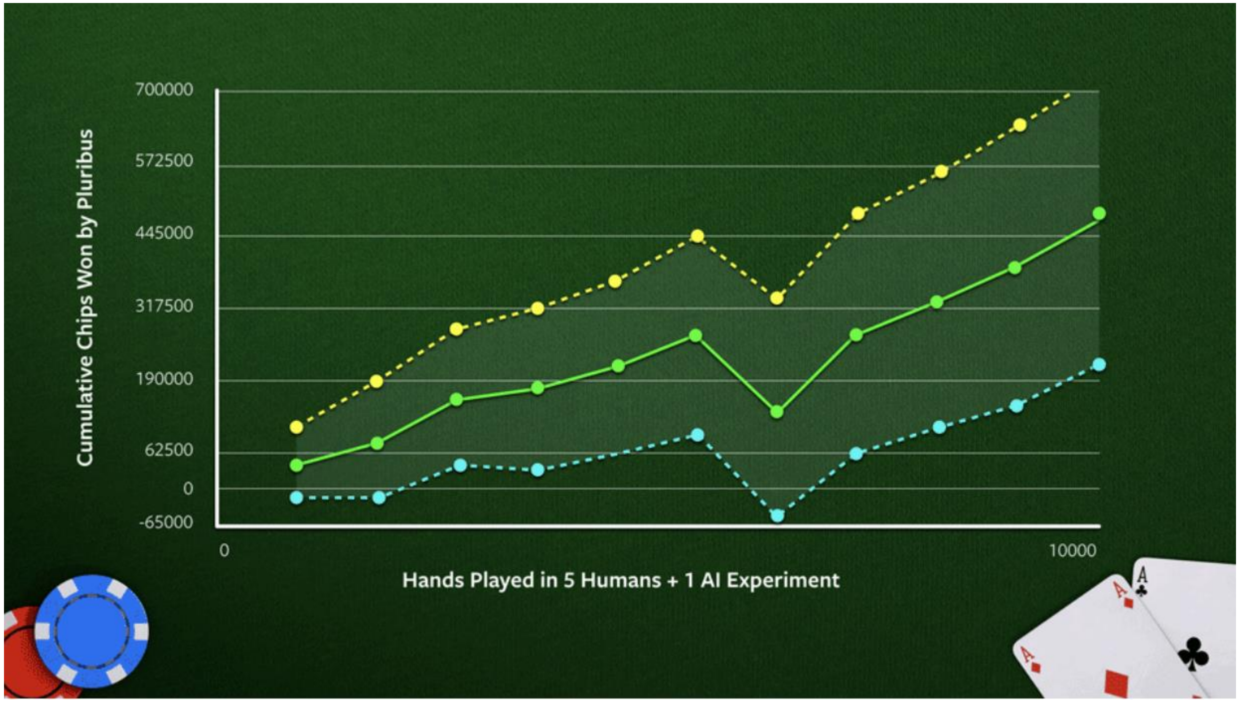
Pluribus: Six-Player No-Limit Texas Hold'em Poker

Brown et al. (CMU, Facebook AI)

- **Six-Player No-Limit Texas Hold'em Poker**
  - Imperfect information
  - Multi-agent
- **Result:** Pluribus won in six-player no-limit Texas Hold'em poker
- **Offline:** Self-play to generate coarse-grained "blueprint" strategy
  - Iterative Monte Carlo CFR (MCCFR) algorithm
  - Self-play allows for counterfactual reasoning
- **Online:** Use search to improve blueprint strategy based on particular situation
- Abstractions
  - Action abstractions: reduce action space
  - Information abstraction: reduce decision space based on what information has been revealed

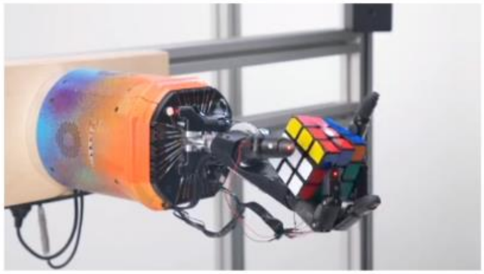
Pluribus: Six-Player No-Limit Texas Hold’em Poker

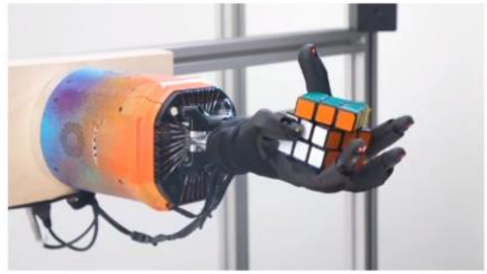
Brown et al. (CMU, Facebook AI)

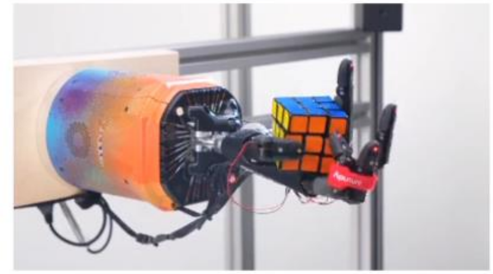


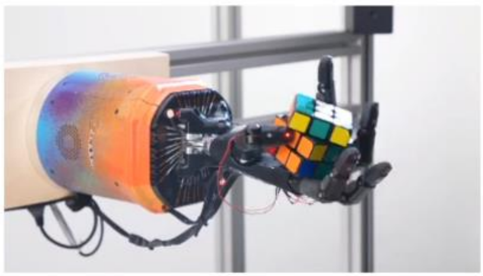
- **Chris Ferguson:** “Pluribus is a very hard opponent to play against. It’s really hard to pin him down on any kind of hand. He’s also very good at making thin value bets on the river. He’s very good at extracting value out of his good hands.”
- **Darren Elias:** “Its major strength is its ability to use mixed strategies. That’s the same thing that humans try to do. It’s a matter of execution for humans — to do this in a perfectly random way and to do so consistently. Most people just can’t.”

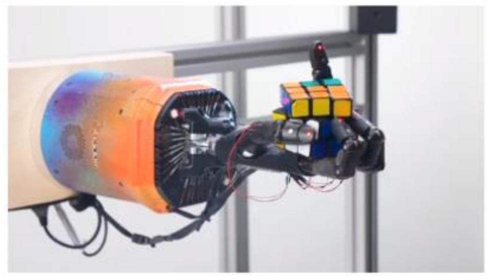
OpenAI Rubik’s Cube Manipulation

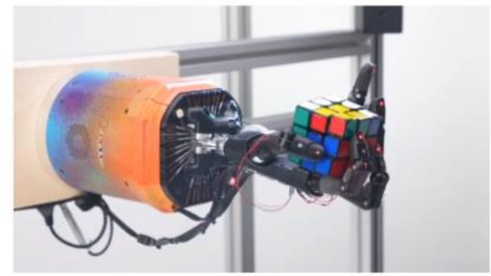

Unperturbed (for reference)


Rubber glove


Tied fingers


Blanket occlusion and perturbation


Plush giraffe perturbation


Pen perturbation

- **Deep RL:** Reinforcement learning approach from OpenAI Five
- **ADR:** Automatic Domain Randomization (ADR) – generate progressively more difficult environment as the system learns (alternative for self-play)
- **Capacity:** Term of “emergent meta-learning” is used to described the fact that the network is constrained and the ADR process of environment generation is not

Deep RL and Self-Play:

Hopes for 2020

- **Robotics:** Use of RL methods in manipulation and real-world interaction tasks.

- **Human Behavior:** Use of multi-agent self-play to explore naturally emerging social behaviors as a way to study equivalent multi-human systems.

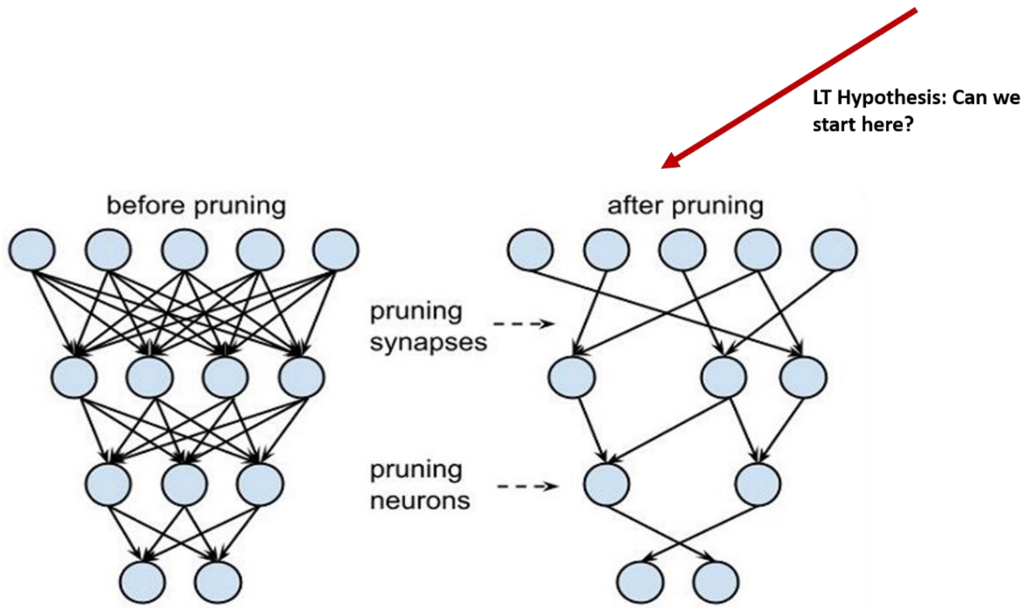
- **Games:** Use RL to assist human experts in discovering new strategies at games and other tasks in simulation.

Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- **Science of Deep Learning and Interesting Directions**
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks

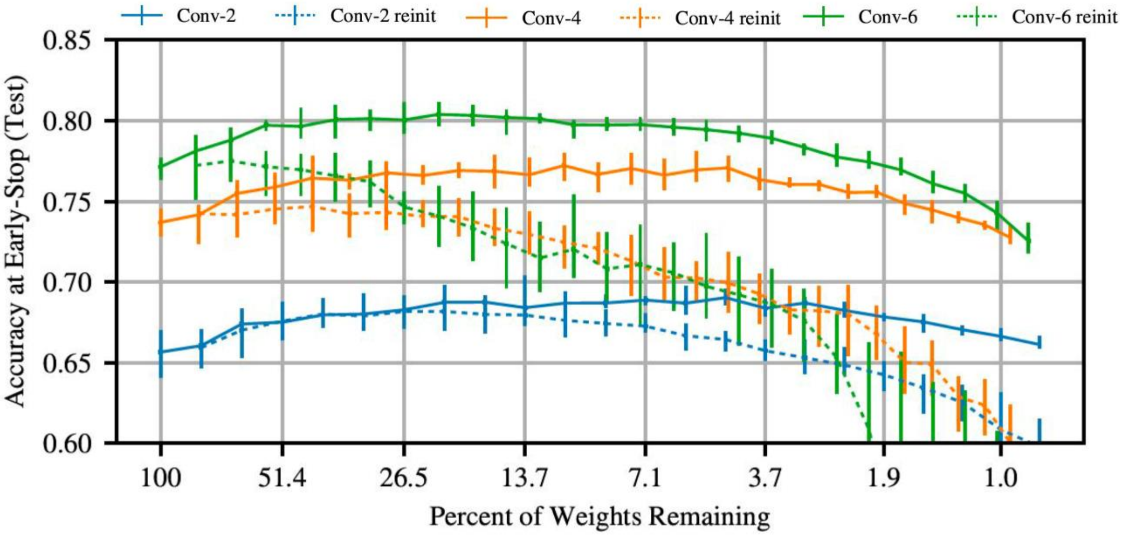
Frankle et al. (MIT) - Best Paper at ICLR 2019



1. Randomly initialize a neural network.
2. Train the network until it converges.
3. Prune a fraction of the network.
4. Reset the weights of the remaining network to initialization values from step 1
5. Train the pruned, untrained network. Observe convergence and accuracy.

The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks

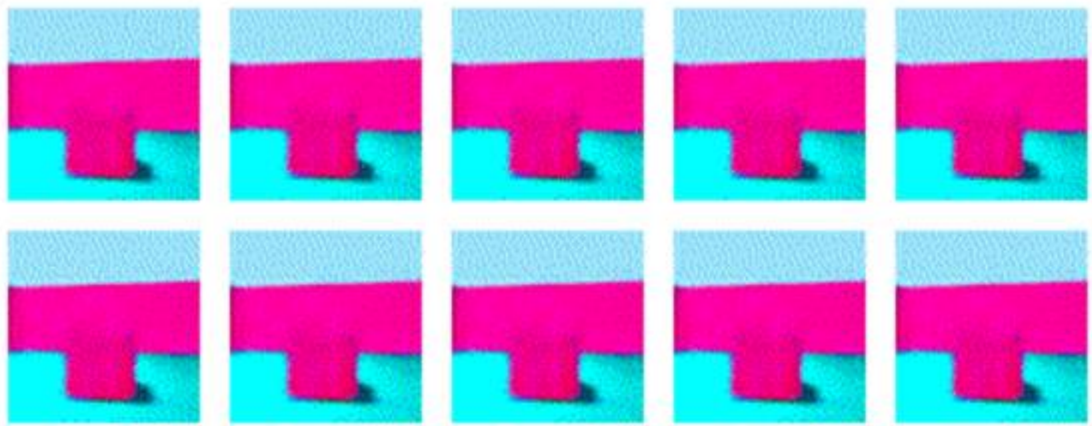
Frankle et al. (MIT) - ICLR 2019 Best Paper



- **Idea:** For every neural network, there is a subnetwork that can achieve the same accuracy in isolation after training.
- **Iterative pruning:** Find this subset subset of nodes by iteratively training network, pruning its smallest-magnitude weights, and re-initializing the remaining connections to their original values. Iterative vs one-shot is key.
- **Inspiring takeaway:** There exist architectures that are much more efficient. Let's find them!

Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations

Locatello et al. (ETH Zurich, Max Plank Institute, Google Research) - ICML 2019 Best Paper

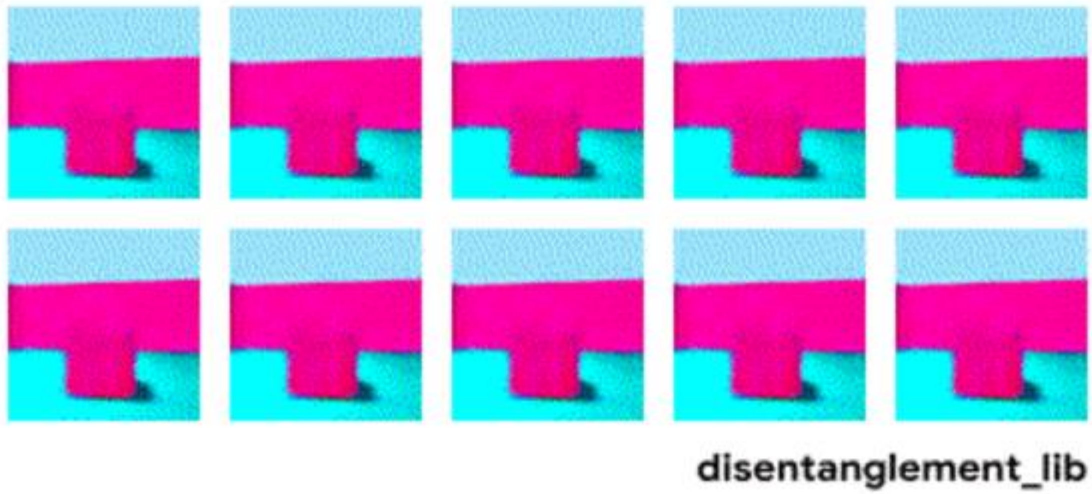


disentanglement_lib

- The goal of disentangled representations is to build models that can capture explanatory factors in a vector.
- The figure above presents a model with a 10-dimensional representation vector.
- Each of the 10 panels visualizes what information is captured in one of the 10 different coordinates of the representation.
- From the top right and the top middle panel we see that the model has successfully disentangled floor color, while the two bottom left panels indicate that object color and size are still entangled.

Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations

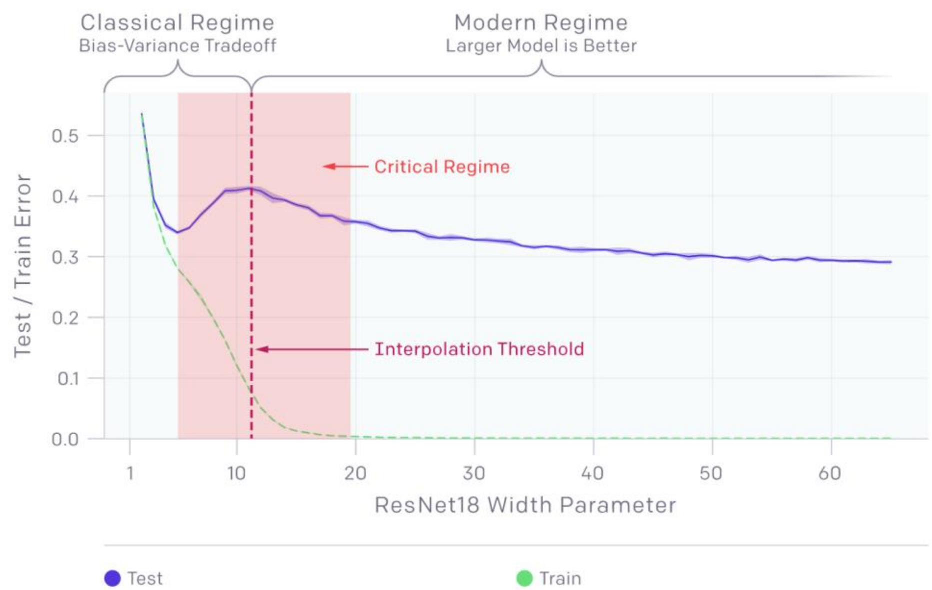
Locatello et al. (ETH Zurich, Max Plank Institute, Google Research) - ICML 2019 Best Paper



- **Proof:** Unsupervised learning of disentangled representations without inductive biases is impossible.
- **Takeaway:** Inductive biases (assumptions) should be made explicit
- **Open problem**: Finding good inductive biases for unsupervised model selection that work across multiple data sets persists is a key open problem.
- **Open Experiments:** Open source library with implementations of the considered disentanglement methods and metrics, a standardized training and evaluation protocol, as well as visualization tools to better understand trained models.

Deep Double Descent: Where Bigger Models and More Data Hurt

Nakkiran et al. (Harvard, OpenAI)



- **Double Descent Phenomena:** As we increase the number of parameters in a neural network, the test error initially decreases, increases, and, just as the model is able to fit the train set, undergoes a second descent.
- Applicable to model size, training time, dataset size.

Hopes for 2020

- **Fundamentals:** Exploring fundamentals of model selection, training dynamics, and representation characteristics with respect to architecture characteristics.
- **Graph Neural Networks:** Exploring use of graph neural networks for combinatorial optimization, recommendation systems, etc.
- **Bayesian Deep Learning**: Exploring Bayesian neural networks for estimating uncertainty and online/incremental learning.

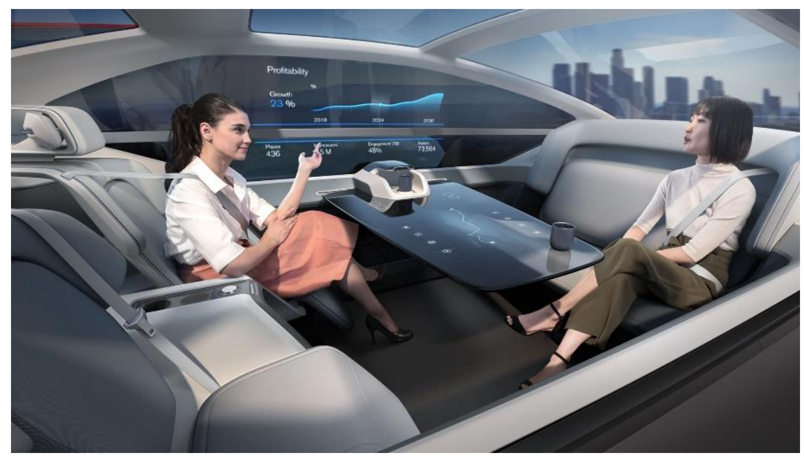
Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- **Autonomous Vehicles and AI-Assisted Driving**
- Government, Politics, Policy
- Courses, Tutorials, Books
- General Hopes for 2020

Level 2

Human is Responsible

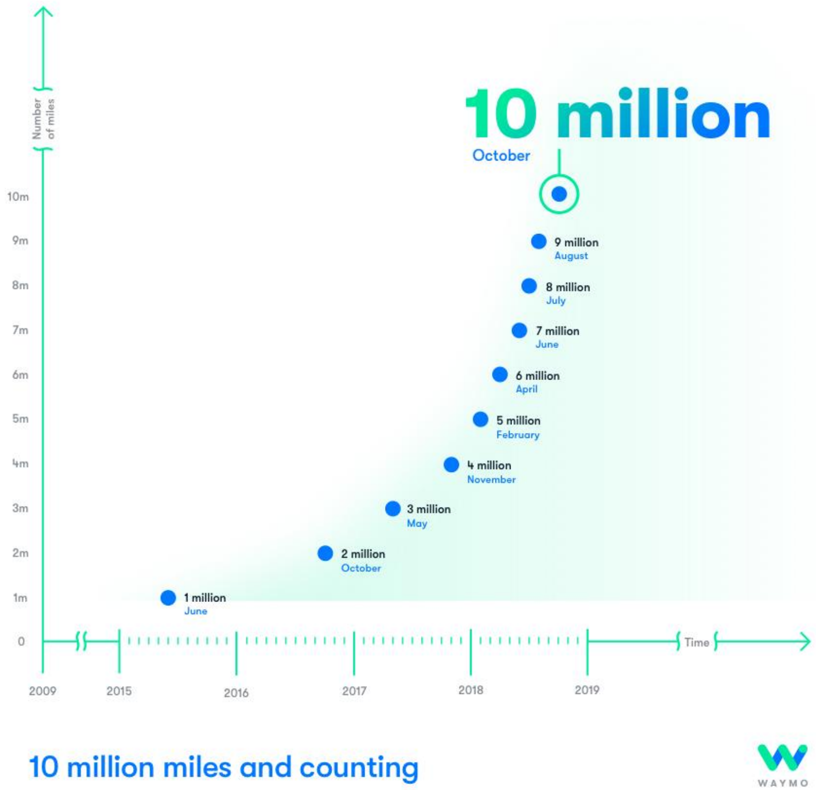
Level 4



Machine is Responsible

Waymo

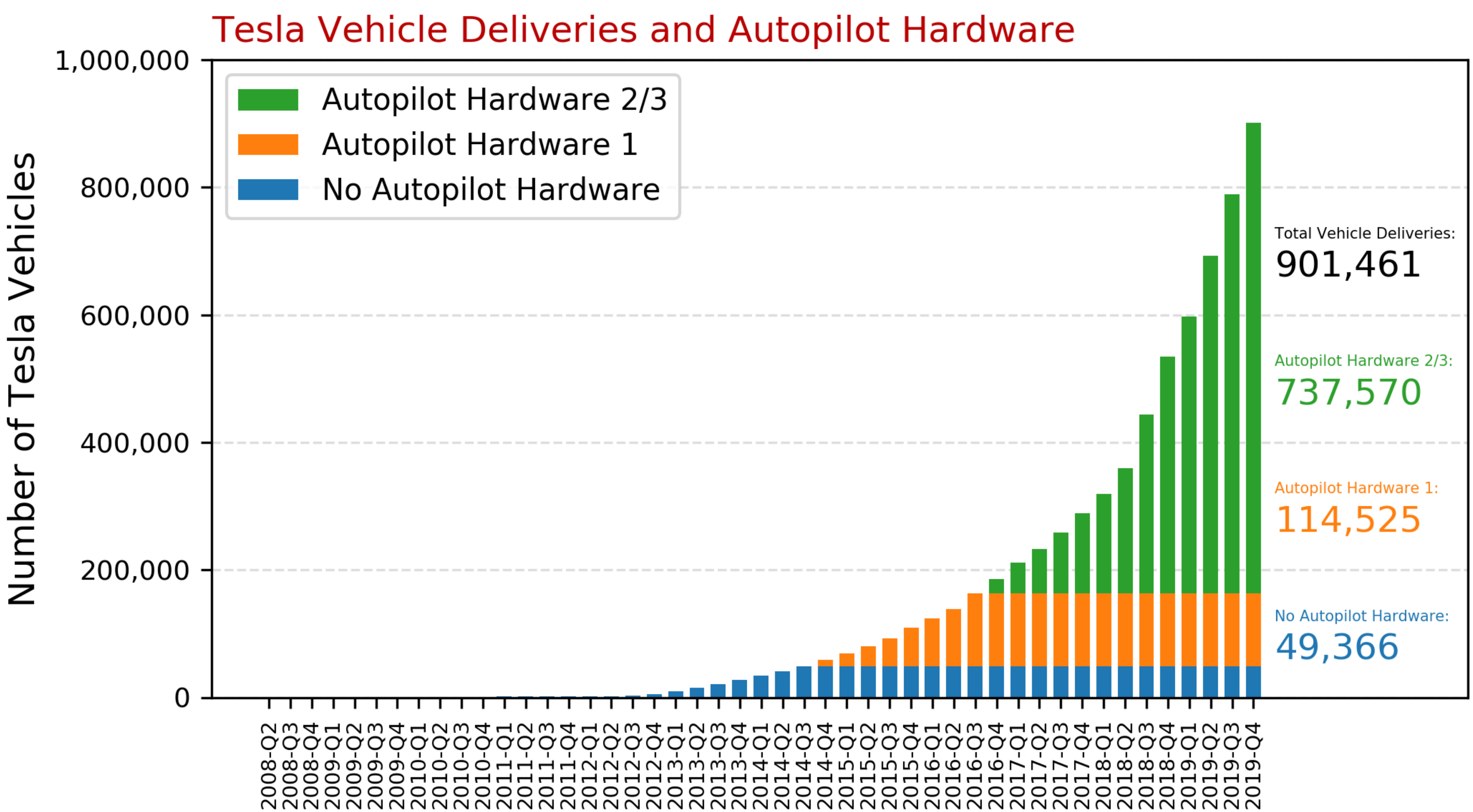
October, 2018:



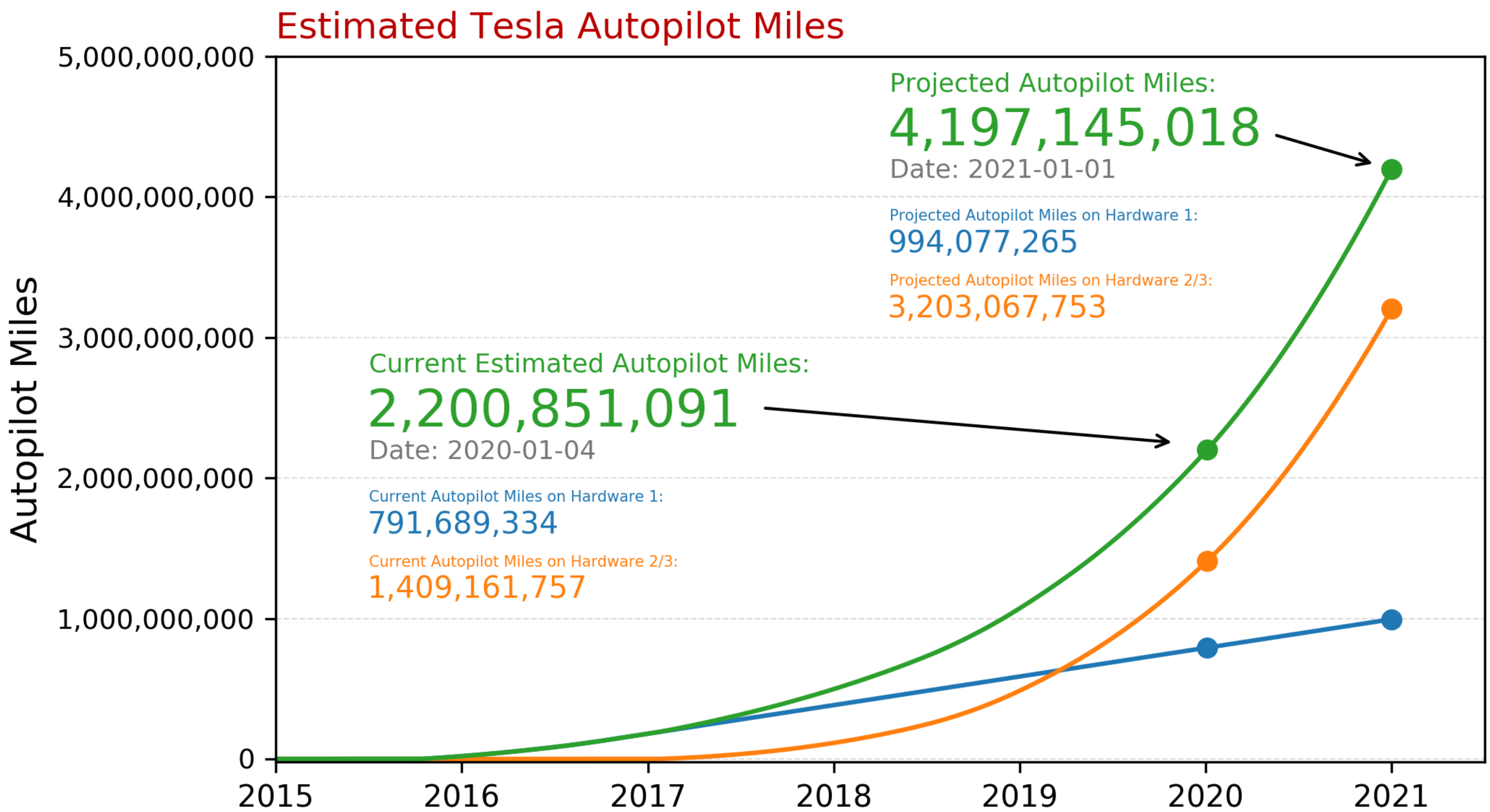
January, 2020:

- **On-road:** 20 million miles
- **Simulation:** 10 billion miles
- **Testing & Validation:** 20,000 classes of structured tests
- Initiated testing without a safety driver

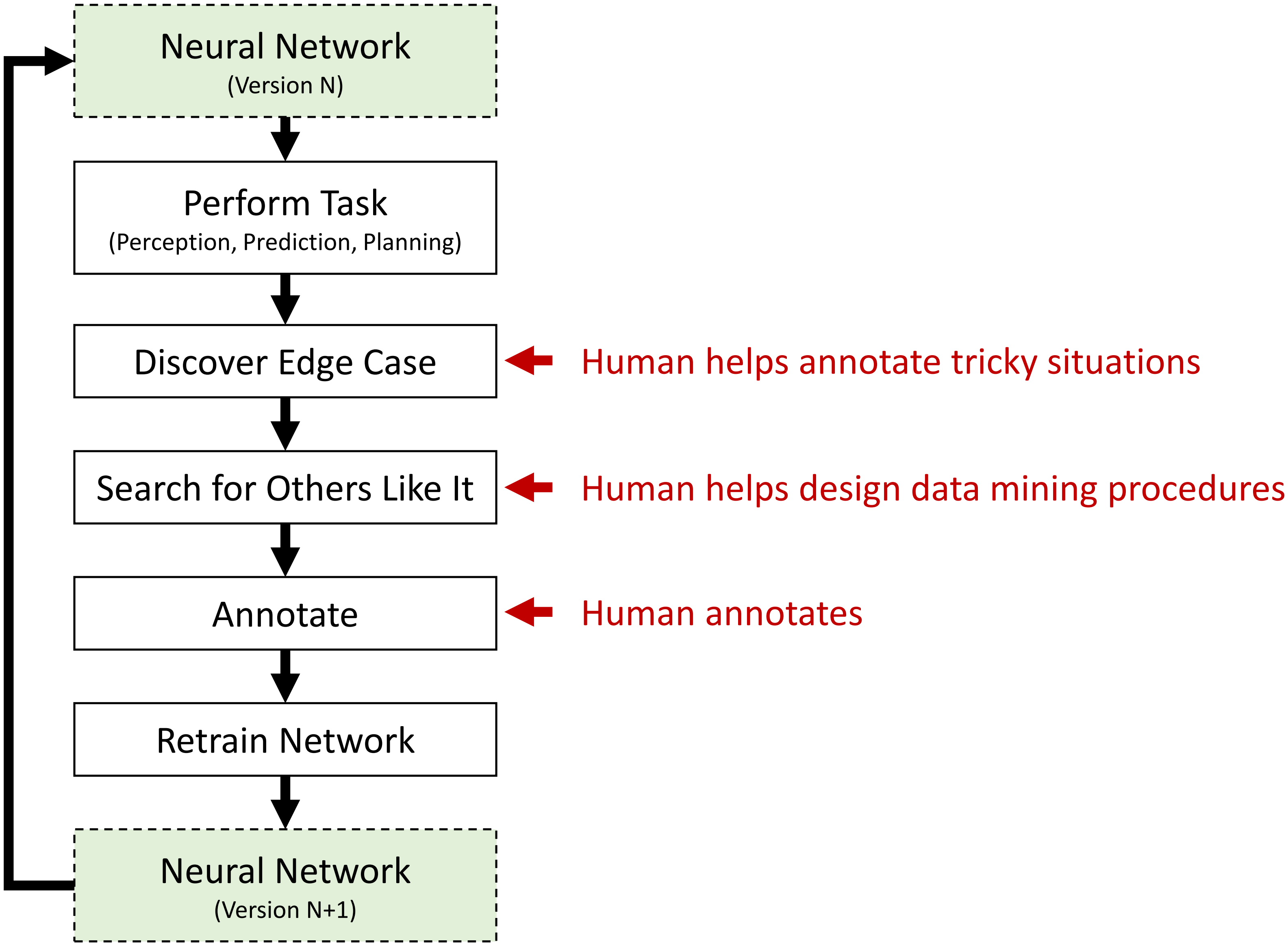
Tesla Autopilot



Tesla Autopilot

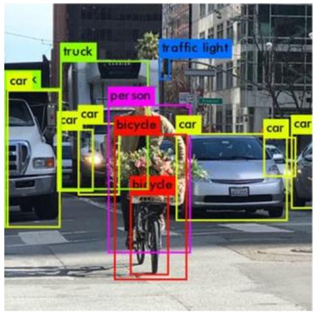


Active Learning Pipeline
(aka Data Engine)



Single-Task Learning

Task 1: Object Detection



Neural Network 1 (Version N)
↓
Perform Task 1 (Object Detection)
↓
Discover Edge Case
↓
Search for Others Like It
↓
Annotate
↓
Retrain Network
↓
Neural Network 1 (Version N+1)

Task 2: Lane Markings

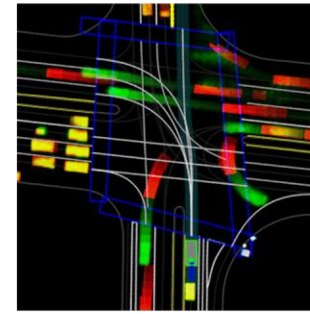
Neural Network 2 (Version N)
↓
Perform Task 2 (Lane Marking Detection)
↓
Discover Edge Case
↓
Search for Others Like It
↓
Annotate
↓
Retrain Network
↓
Neural Network 2 (Version N+1)

Task 3: Drivable Area

Neural Network 3 (Version N)
↓
Perform Task 3 (Drivable Area)
↓
Discover Edge Case
↓
Search for Others Like It
↓
Annotate
↓
Retrain Network
↓
Neural Network 3 (Version N+1)

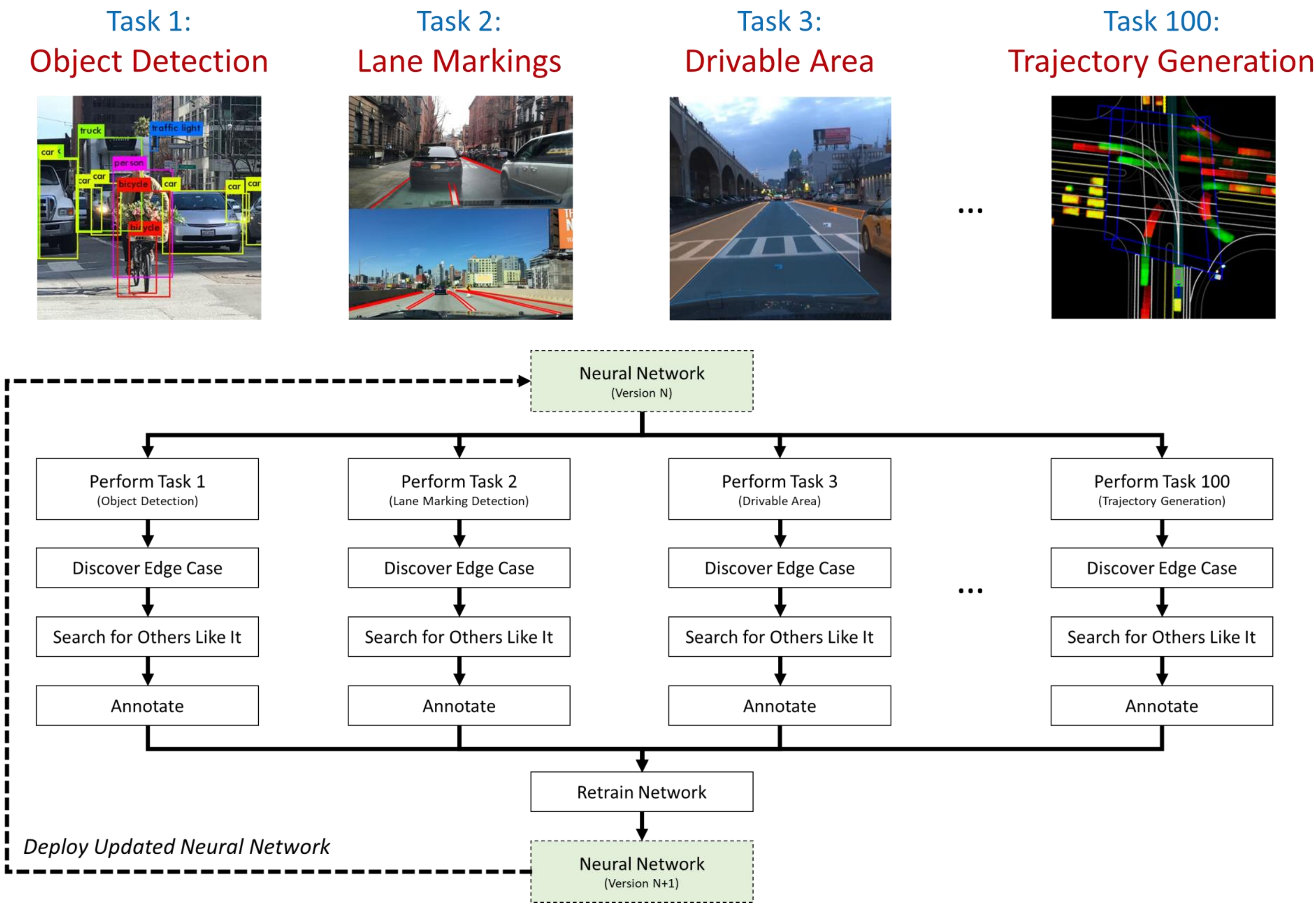
...

Task 100: Trajectory Generation

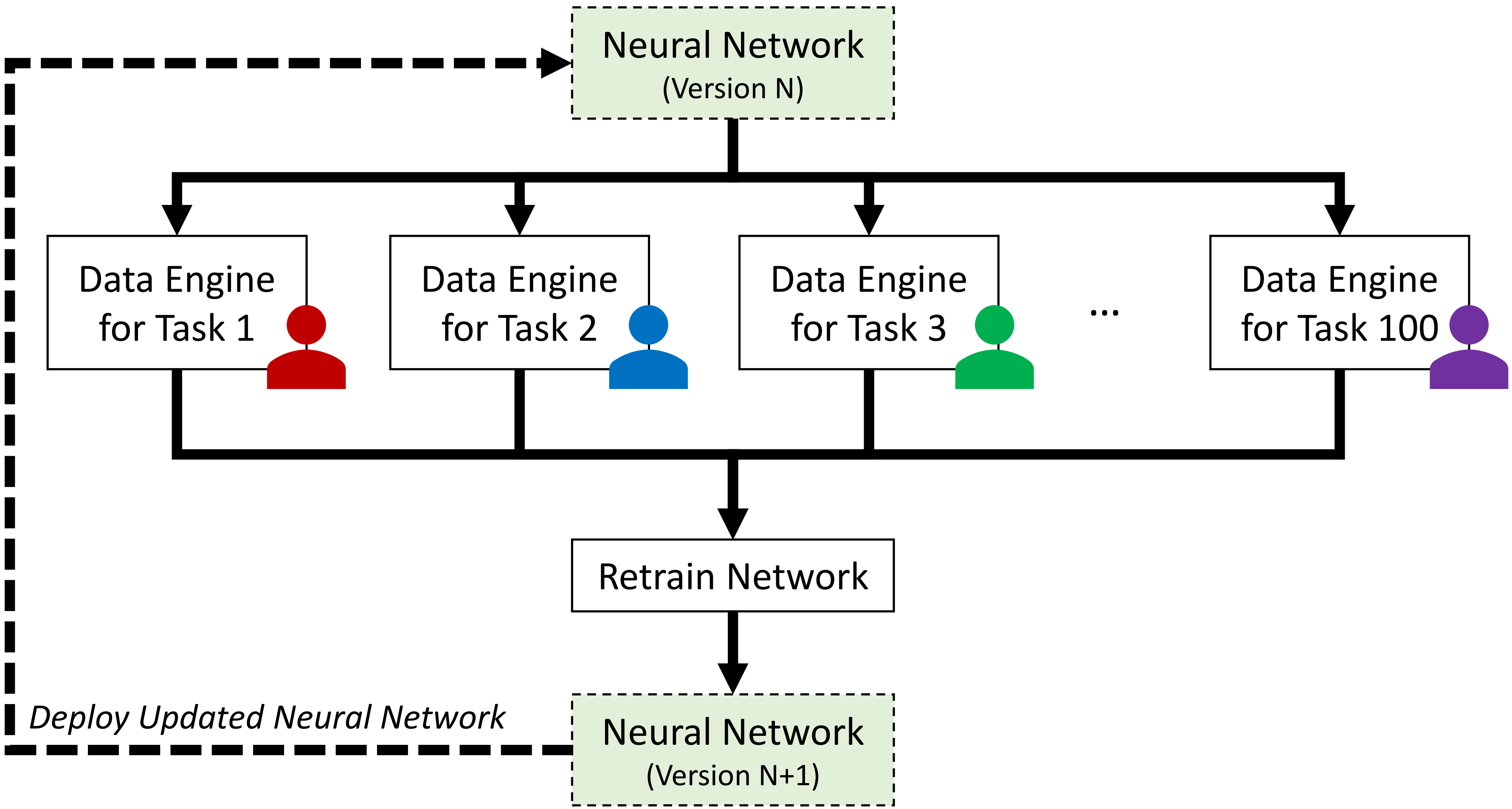


Neural Network 100 (Version N)
↓
Perform Task 100 (Trajectory Generation)
↓
Discover Edge Case
↓
Search for Others Like It
↓
Annotate
↓
Retrain Network
↓
Neural Network 100 (Version N+1)

Multi-Task Learning



Collaborative Deep Learning
(aka Software 2.0 Engineering)



Vision vs Lidar
L2 vs L4

- Primarily: Vision Sensors + Deep Learning
  - **Pros:**
    - Highest resolution information
    - Feasible to collect data at scale and **learn**
    - Roads are designed for human eyes
    - Cheap
  - **Cons:**
    - Needs a huge amount of data to be accurate
    - Less explainable
    - Driver must remain vigilant
- Primarily: Lidar + Maps
  - **Pros:**
    - Explainable, consistent
    - Accurate with less data
  - **Cons:**
    - Less amenable to machine learning
    - Expensive (for now)
    - Safety driver or teleoperation fallback

Example L2 System:

Tesla Autopilot

2+ billion miles

Example L4 System:

Waymo

20+ million miles

Open Questions for Tesla Autopilot

- Deep learning question:
  Problem Difficulty: How difficult is driving? How many edge-cases does it have? Can it be learned from data?
  - Perception (detection, intention modeling, trajectory prediction)
  - Action (in a game-theoretic setting)
    - Balancing enjoyability and safety

- Human supervision of deep learning system question:
  Vigilance: How good can Autopilot get before vigilance decrements significantly?
  - And … will this decrement nullify the safety benefits of automation?

Open Questions for Waymo

- When we have maps, lidar, and geo-fenced routes:
  Problem Difficulty: How difficult is driving? How many edge-cases does it have? Can it be learned from data?
  - Perception (detection, intention modeling, trajectory prediction)
  - Action (in a game-theoretic setting)
    - Balancing enjoyability and safety

- Simulation question:
  How much can be learned from simulation?

Hopes for 2020

- **Applied deep learning innovation:** Life-long learning, active learning, multi-task learning
- **Over-the-air updates:** More level 2 systems begin both data collection and over-the-air software updates.
- **Public datasets of edge-cases:** More publicly available datasets of challenging cases.
- **Simulators:** Improvement of publicly available simulators (CARLA, NVIDIA DRIVE Constellation, Voyage Deepdrive)
- **Less hype:** More balanced in-depth reporting (by journalists and companies) on successes and challenges of autonomous vehicle development.

Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- **Government, Politics, Policy**
- Courses, Tutorials, Books
- General Hopes for 2020

AI in Political Discourse: Andrew Yang

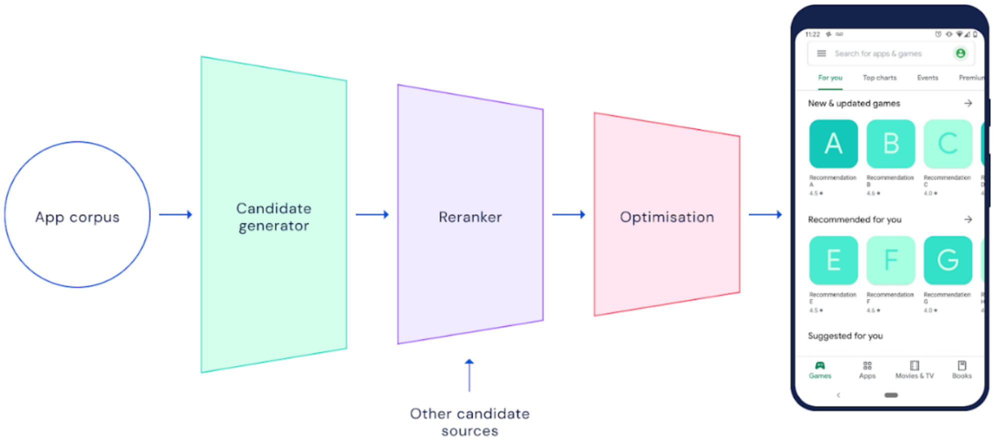
- First presidential candidate to discuss artificial intelligence extensively as part of his platform
- Proposals
  - **Department:** Create a new executive department – the Department of Technology – to work with private industry and Congressional leaders to monitor technological developments, assess risks, and create new guidance.
  - **Focus on AI:** The new Department would be based in Silicon Valley and would initially be focused on Artificial Intelligence.
  - **Companies:** Create a public-private partnership between leading tech firms and experts within government to identify emerging threats and suggest ways to mitigate those threats while maximizing the benefit of technological innovation to society.

American AI Initiative

- In February, 2019, the president signed Executive Order 13859 announcing the American AI Initiative
- Goals
  - Investment in long-term research
  - Support research in academia and industry
  - Access to federal data
  - Promote STEM education
  - Develop AI in "a manner consistent with our Nation's values, policies, and priorities."
  - AI must also be developed in a way that does not compromise our American values, civil liberties, or freedoms.

Tech Leaders Testifying Before Congress (Ethics of Recommender Systems)

DeepMind + Google Research: Play Store App Discovery



- **Candidate app generation:**
  LSTM → Transformer → efficient addition attention model
- **Candidate app unbiasing:**
  - The model learns a bias that favors the apps that are shown – and thus installed – more often.
  - To help correct for this bias, impression-to-install rate weighting is introduced.
- **Multiple objectives:** relevance, popularity, or personal preferences

Hopes for 2020

- **Less fear of AI:** More balanced, informed discussion on the impact of AI in society.
- **Experts:** Continued conversations by government officials about AI, privacy, cybersecurity with experts in academia and industry.
- **Recommender system transparency:** More open discussion and publication behind recommender systems used in industry.

Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- **Courses, Tutorials, Books**
- General Hopes for 2020

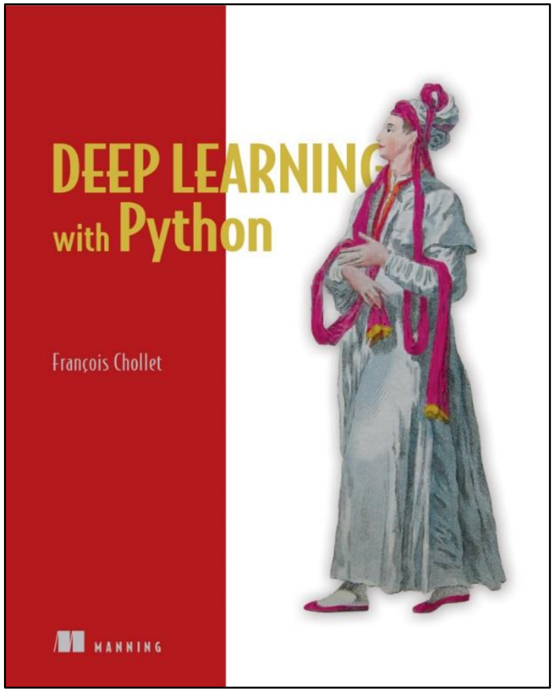
Online Deep Learning Courses

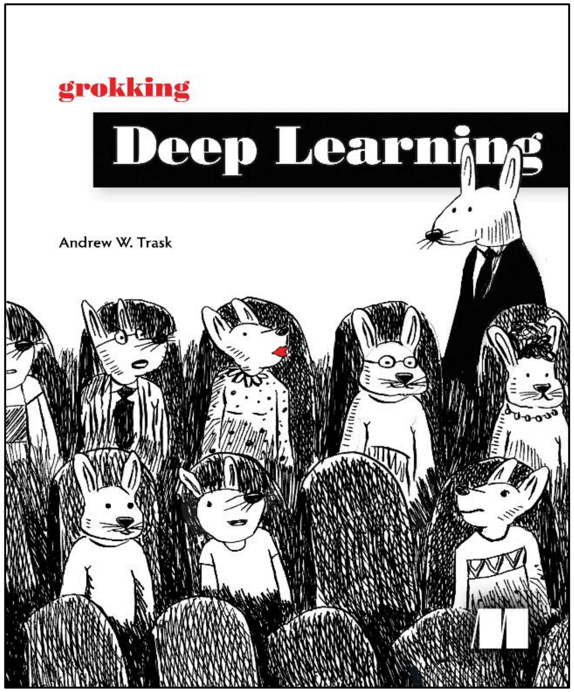
- Deep Learning
  - Fast.ai: Practical Deep Learning for Coders
    - Jeremy Howard et al.
  - Stanford CS231n: Convolutional Neural Networks for Visual Recognition
  - Stanford CS224n: Natural Language Processing with Deep Learning
  - Deeplearning.ai (Coursera): Deep Learning
    - Andrew Ng
- Reinforcement Learning
  - David Silver: Introduction to Reinforcement Learning
  - OpenAI: Spinning Up in Deep RL

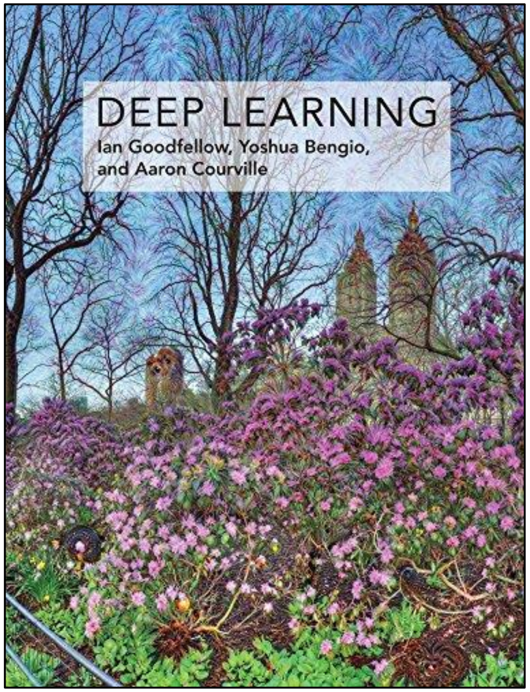
Tutorials: Over 200 of the Best Machine Learning, NLP, and Python Tutorials (by Robbie Allen)

- Link: http://bit.ly/36skFE7
- Topics
  - Machine learning
  - Activation and Loss Functions
  - Bias
  - Perceptron
  - Regression
  - Gradient descent
  - Generative learning
  - Support vector machines
  - Backpropagation
  - Deep Learning
  - Optimization
  - Long Short Term Memory
  - Convolutional Neural Networks
  - Recurrent Neural Nets (RNNs)
  - Reinforcement Learning
  - Generative Adversarial Networks
  - Multi-task Learning
  - NLP
  - Word Vectors
  - Encoder-Decoder
  - TensorFlow
  - PyTorch

Deep Learning Books







Outline

- Deep Learning Growth, Celebrations, and Limitations
- Deep Learning and Deep RL Frameworks
- Natural Language Processing
- Deep RL and Self-Play
- Science of Deep Learning and Interesting Directions
- Autonomous Vehicles and AI-Assisted Driving
- Government, Politics, Policy
- Courses, Tutorials, Books
- **General Hopes for 2020**

Hopes for 2020

- Reasoning
- Active learning and life-long learning
- Multi-modal and multi-task learning
- Open-domain conversation
- Applications: medical, autonomous vehicles
- Algorithmic ethics
- Robotics
- Recommender systems

Hope for 2020: Recipe for Progress (in AI)

Skepticism

Perseverance
(Never Give Up)

Hard Work

Criticism

Open-
Mindedness

Crazy

"The future depends on
some graduate student
who is deeply suspicious
of everything I have said."

- Geoffrey Hinton

Thank You

Videos and slides are posted on the website:

deeplearning.mit.edu