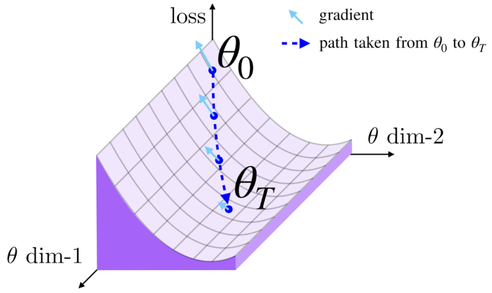
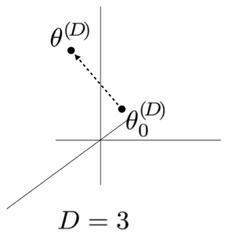
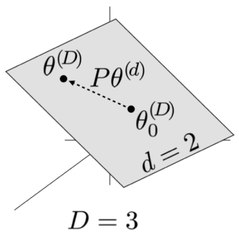
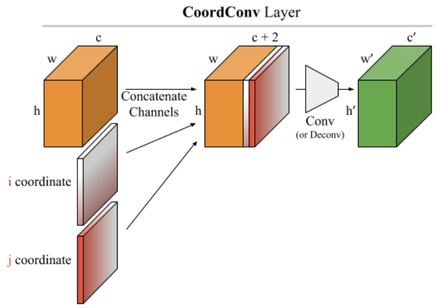
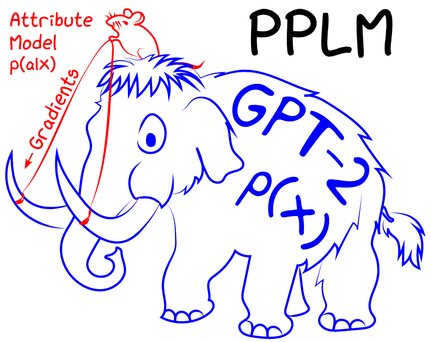
How to Have Fun in AI Research

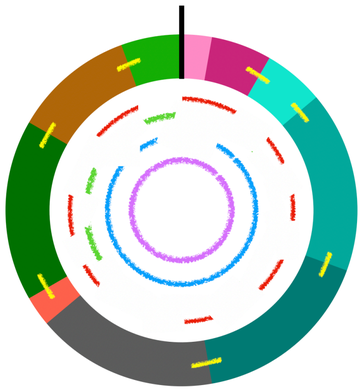












Rosanne Liu

Senior Research Scientist, Uber AI

http://www.rosanneliu.com/

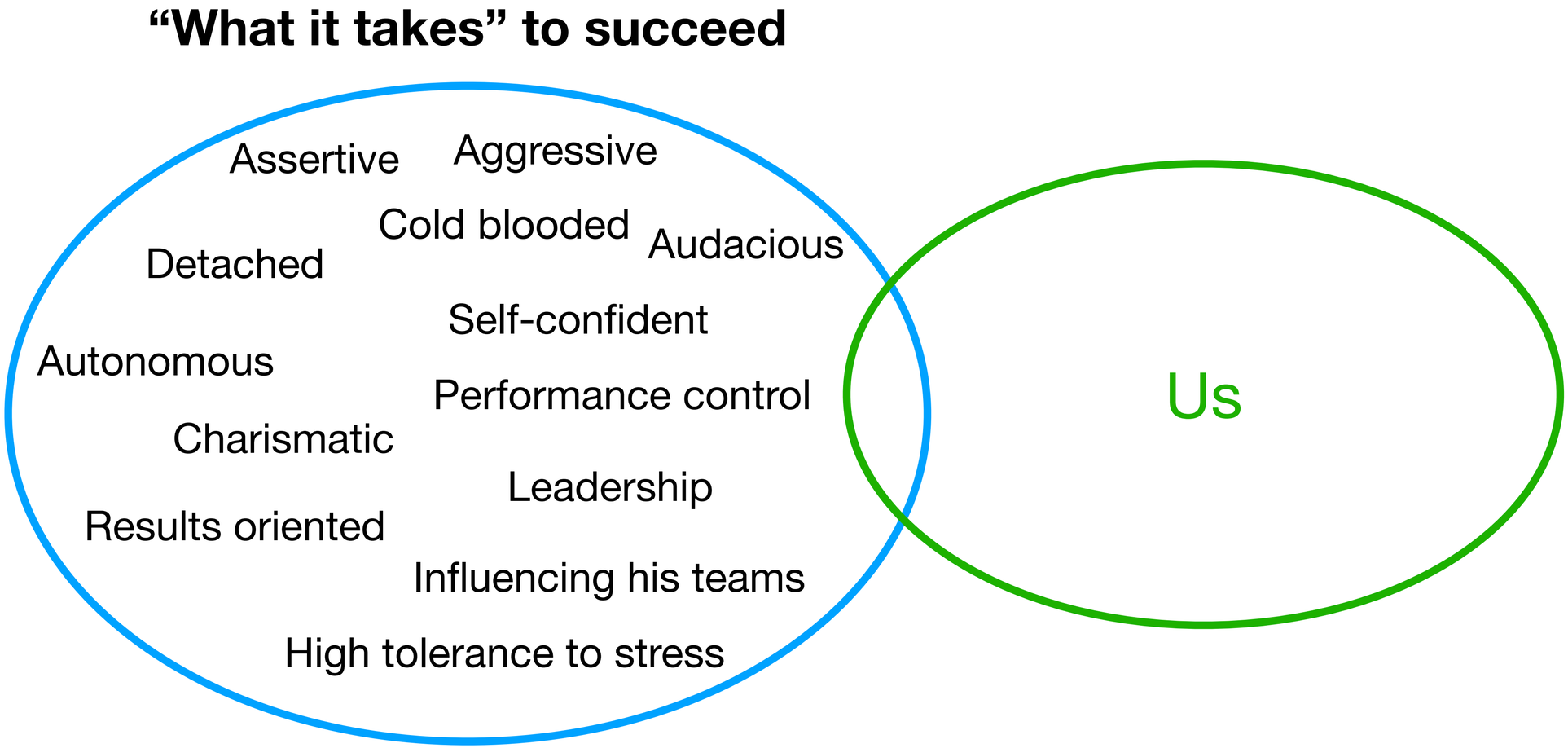
Why do I want to talk about "fun"

Why do I want to talk about "fun"

- This is the right crowd, and right time.

Why do I want to talk about "fun"

- This is the right crowd, and right time.

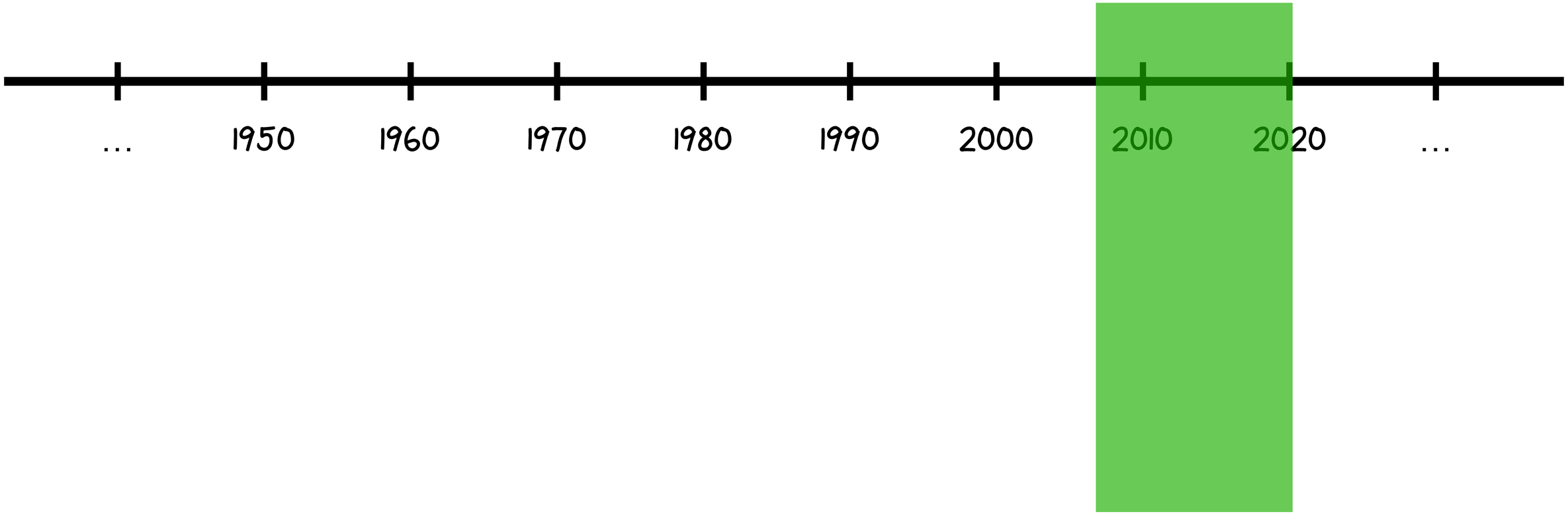


Why do I want to talk about "fun"

- This is the right crowd, and right time.

Why do I want to talk about "fun"

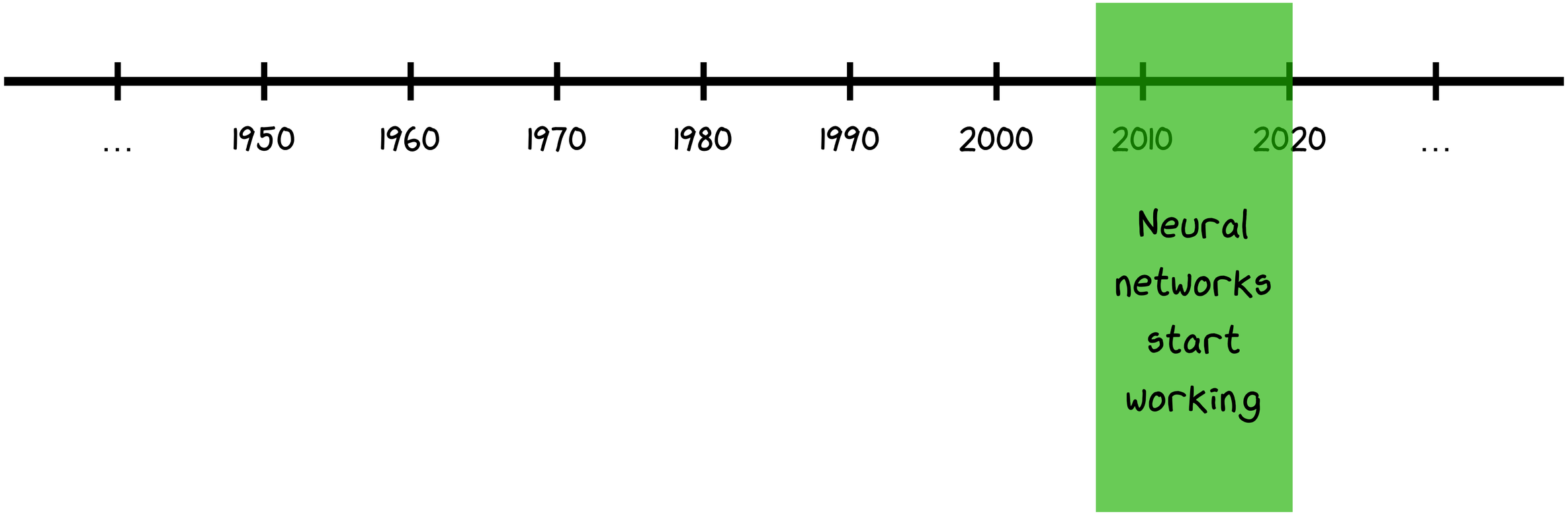
- This is the right crowd, and right time.



Slide credit: Jason Yosinski

Why do I want to talk about "fun"

- This is the right crowd, and right time.



Slide credit: Jason Yosinski

Progress in AI research (proxy: # arXiv papers)

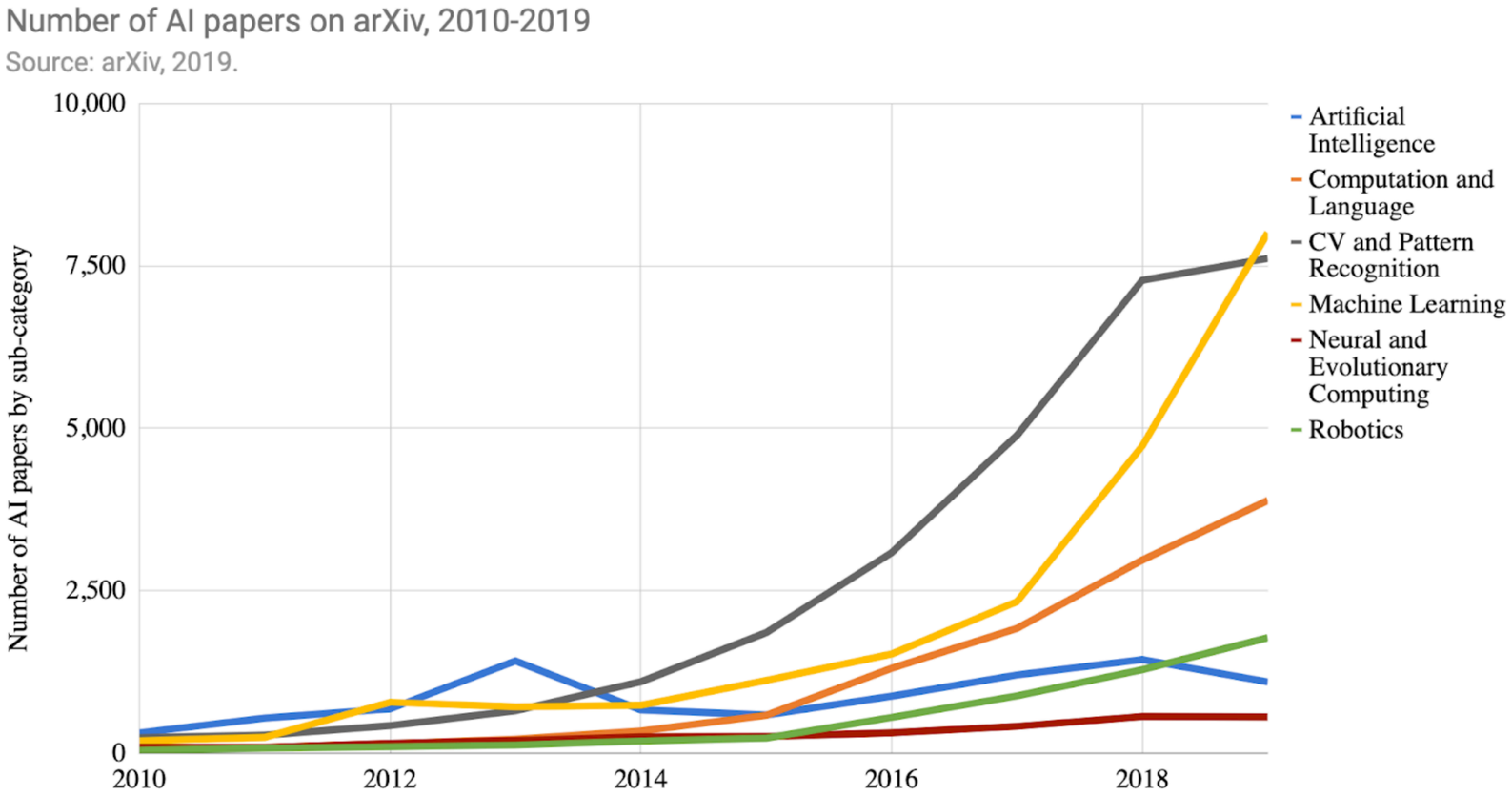


Image: AI Index Report 2019

Progress in stress level (proxy: # arXiv papers I didn't read)

Progress in stress level (proxy: # arXiv papers I didn't read)

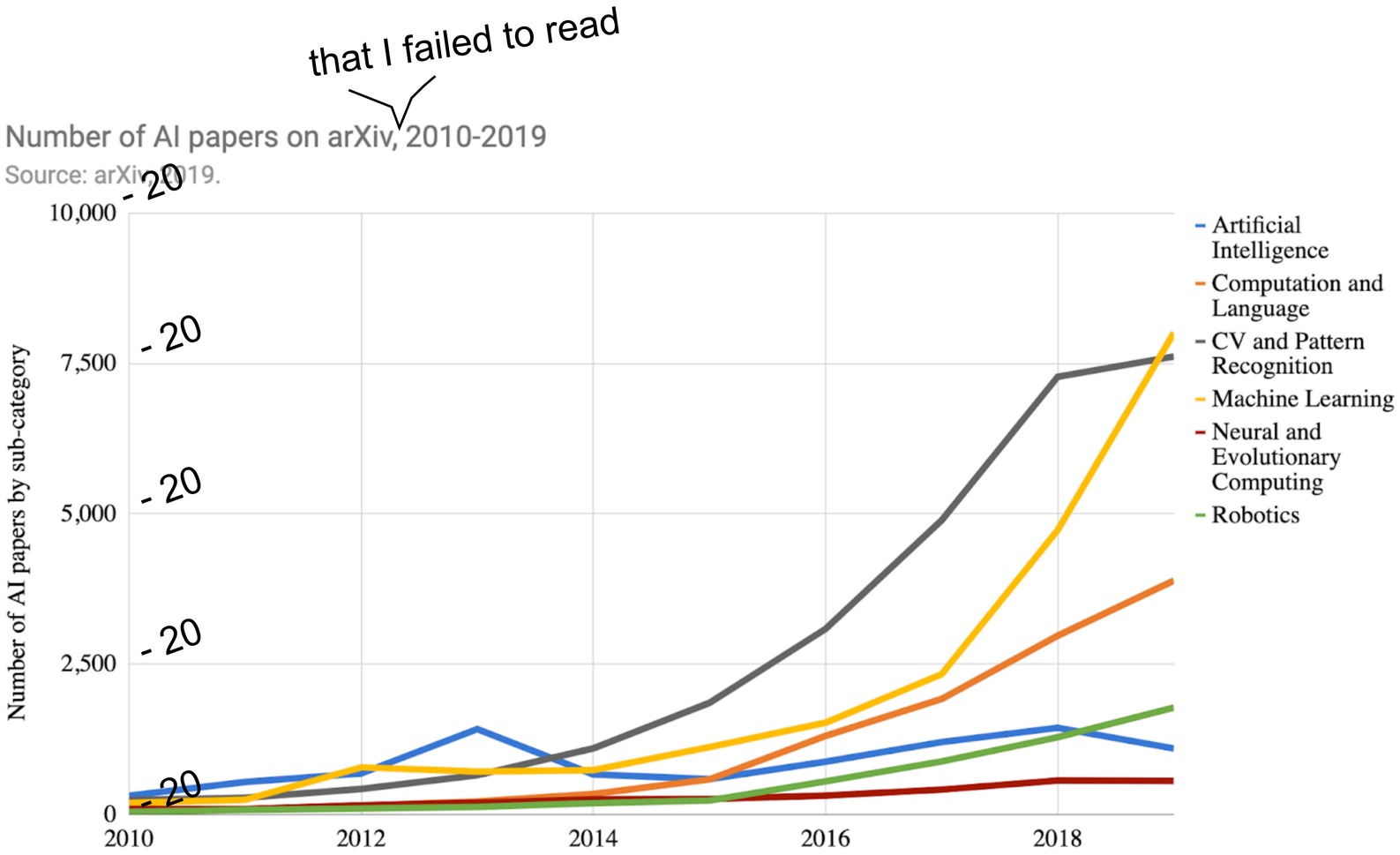


Image: AI Index Report 2019

Why do I want to talk about "fun"

- This is the right crowd, and right time.
- We need fun to reshape our own behavior

Why do I want to talk about "fun"

- This is the right crowd, and right time.
- We need fun to reshape our own behavior

"It takes an enormous self-esteem [for women] to listen to things like that and not be demolished."

–Vera Rubin

Why do I want to talk about "fun"

- This is the right crowd, and right time.
- We need fun to reshape our own behavior

"It takes an enormous self-esteem [for women] to listen to things like that and not be demolished."

–Vera Rubin

"Gender gaps in the rich world have had ever less to do with overt discrimination, and ever more to do with women's decisions."

–Marianne Bertrand, lecture at University of Chicago, January 2020

Why do I want to talk about "fun"

- This is the right crowd, and right time.
- We need fun to reshape our own behavior

"It takes an enormous self-esteem [for women] to listen to things like that and not be demolished."

–Vera Rubin

"Gender gaps in the rich world have had ever less to do with overt discrimination, and ever more to do with women's decisions."

–Marianne Bertrand, lecture at University of Chicago, January 2020

- We need fun to defy social norms

The choices we made under social norm pressures may seem voluntary, but they reflect the influence of a self-perpetuating gender bias.

And really, the complete sentence is

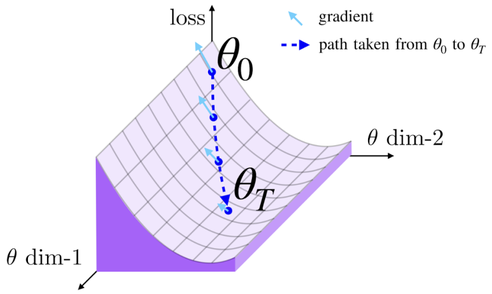
How to have fun in AI research, despite of who you are.

How to have fun in AI research, even when signs hint that you may not.

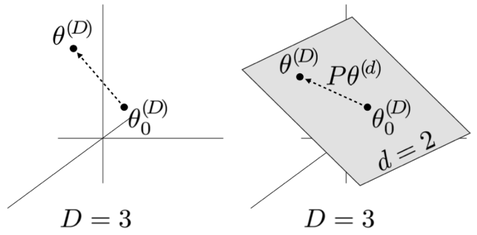
How to have fun in AI research, even when you yourself doubt that you can.

The plan is to walk through a number of projects

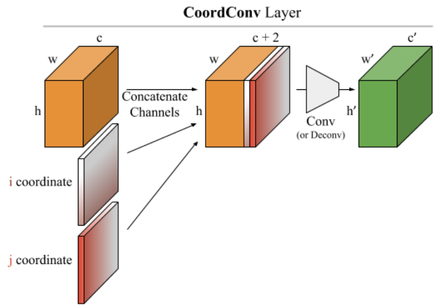
Loss change allocation (LCA)



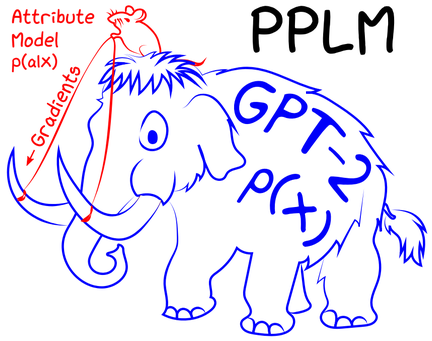
Intrinsic Dimension



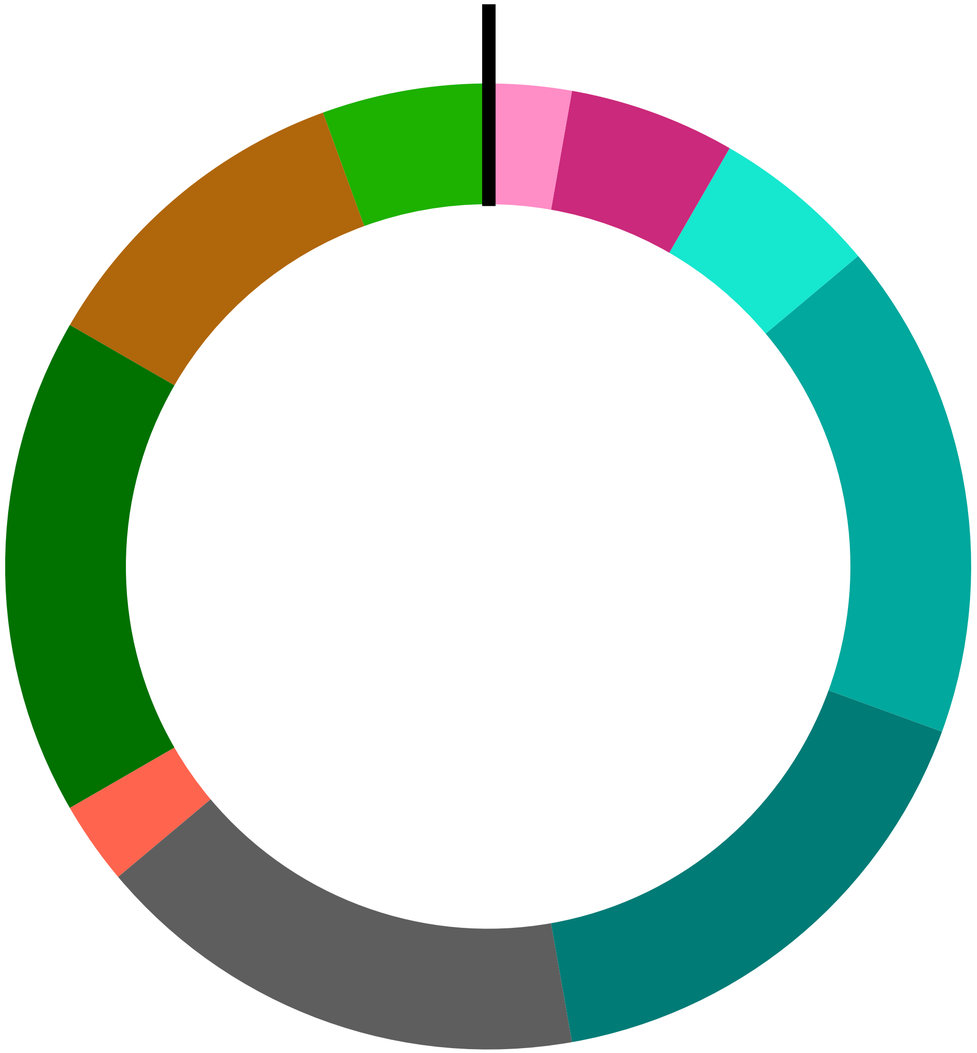
CoordConv



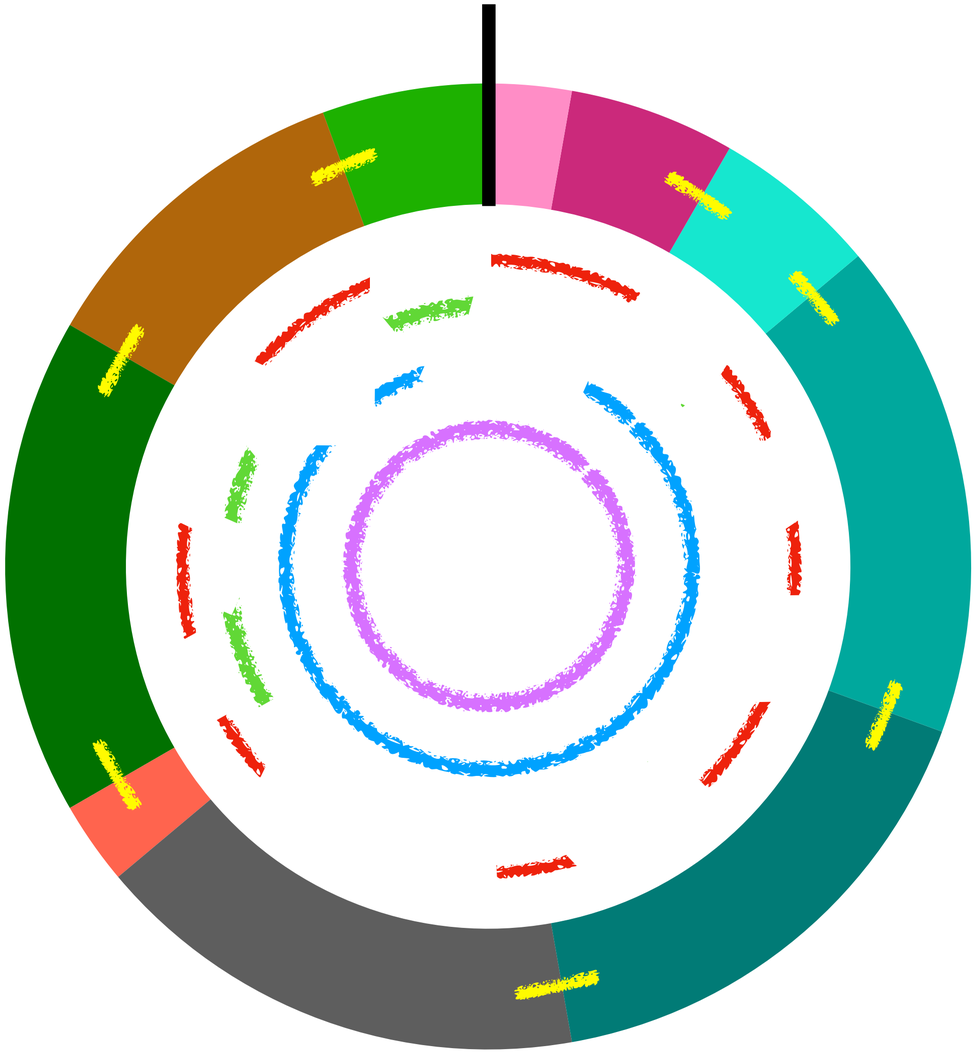
PPLM



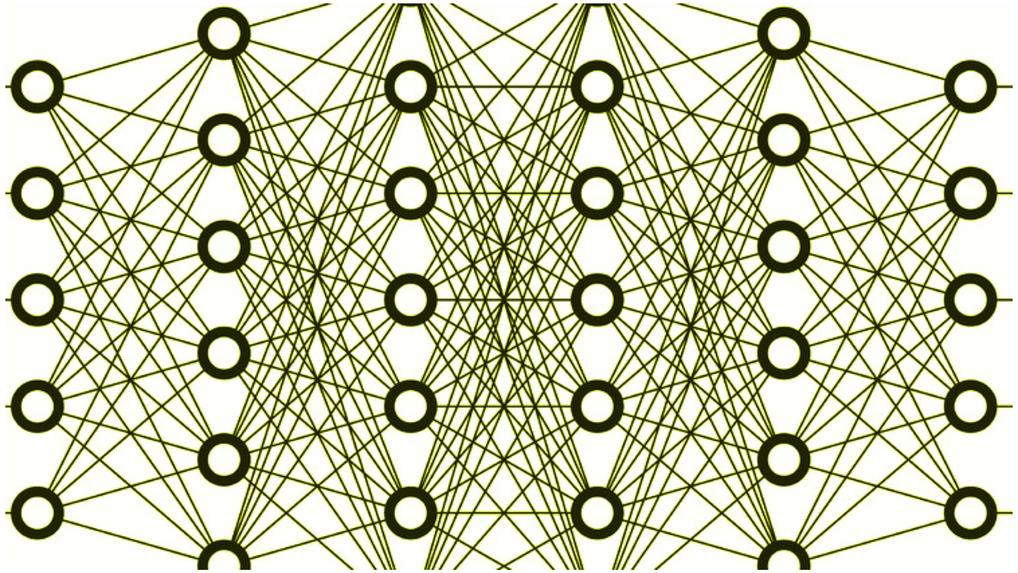
And analyze what a complete research cycle is like



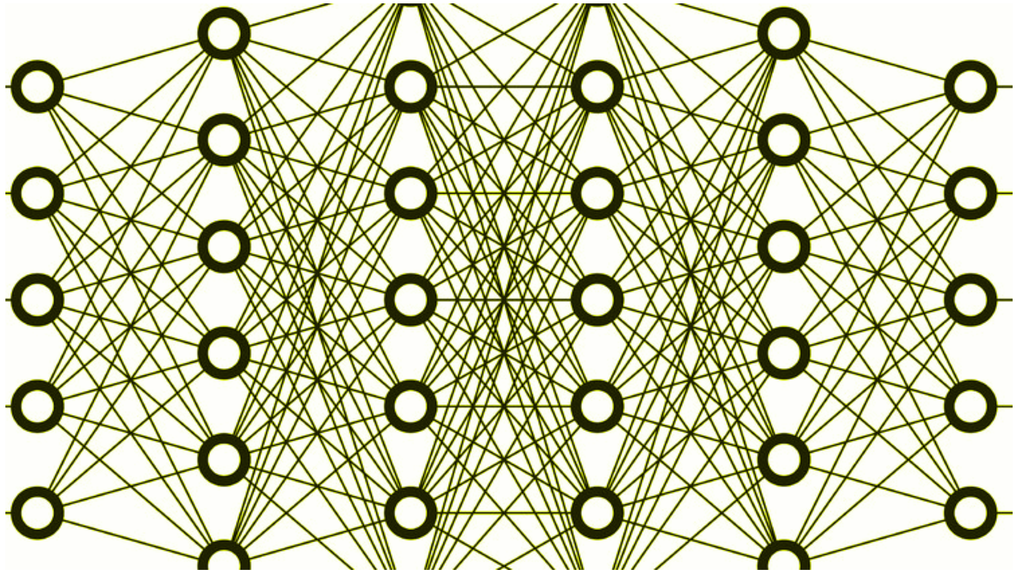
And show you where the fun is to be had



Scope of projects: deep neural networks



Scope of projects: deep neural networks



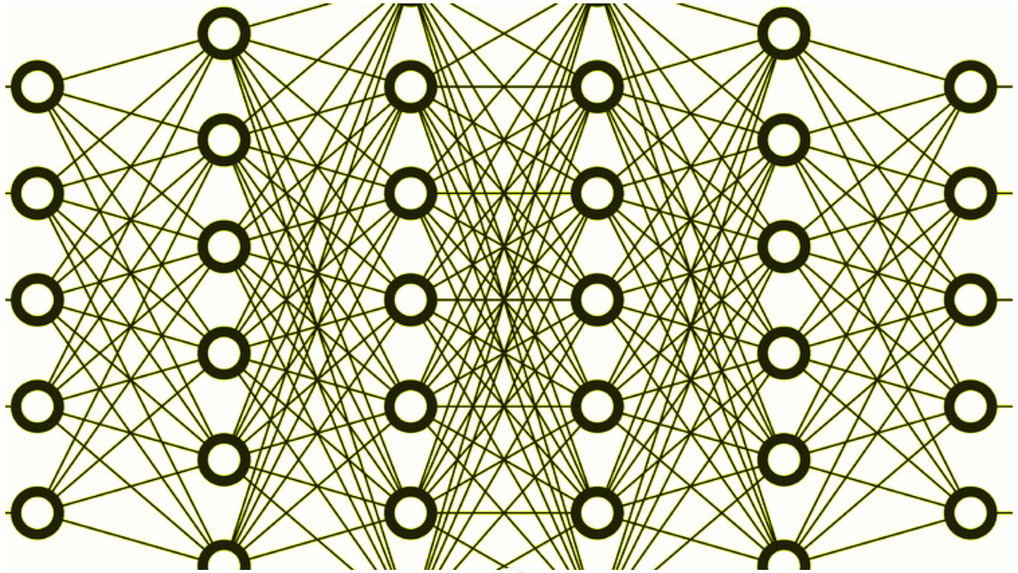
What information does it encode, and not encode?

What does it do when it "trains"?

What's its complexity?

How much can we "control" it?

Scope of projects: deep neural networks



Project I

What does it do when it "trains"?

Project II

What's its complexity?

Project III

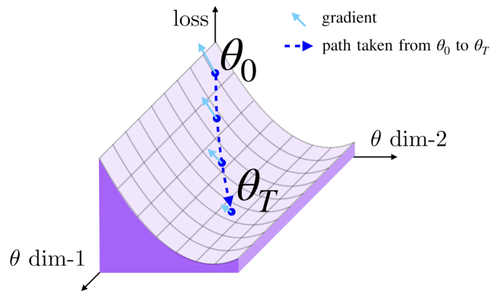
What information does it encode, and not encode?

Project IV

How much can we "control" it?

Project 1: LCA

Loss change allocation (LCA)

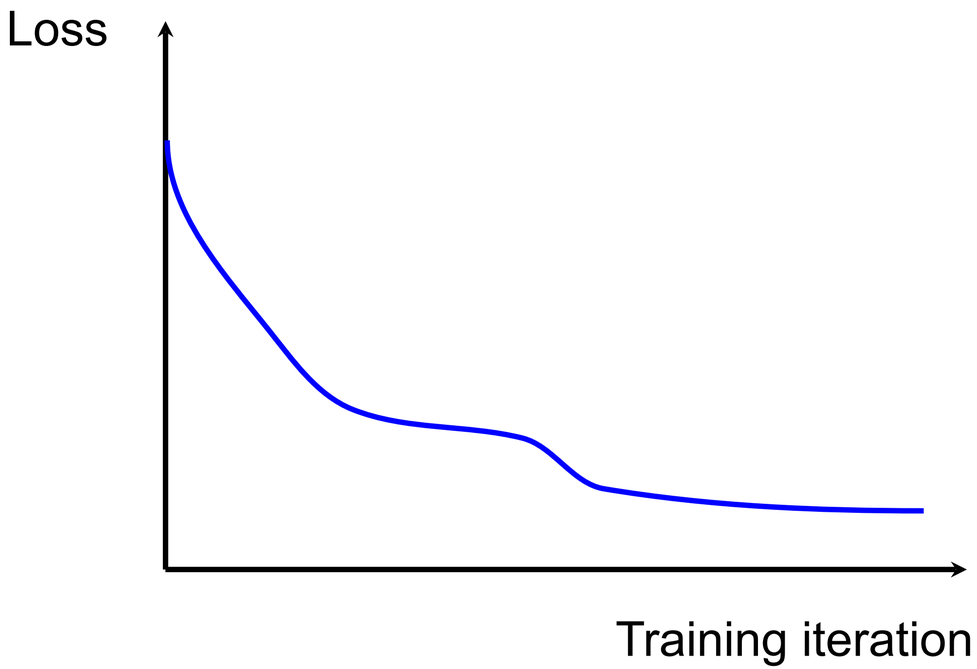


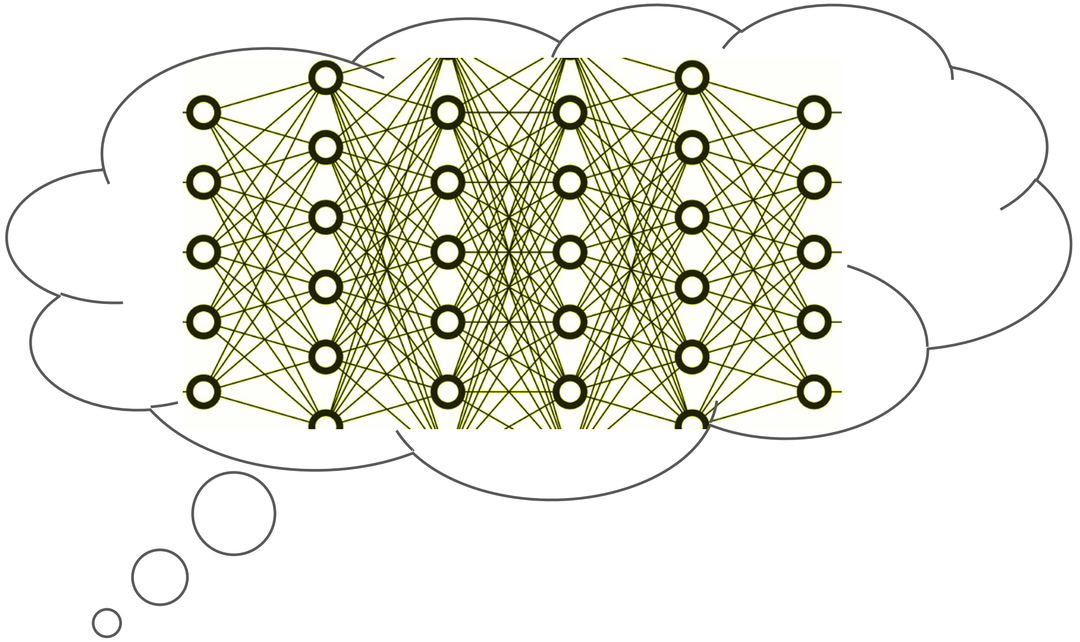
+ Janice Lan, Hattie Zhou, Jason Yosinski

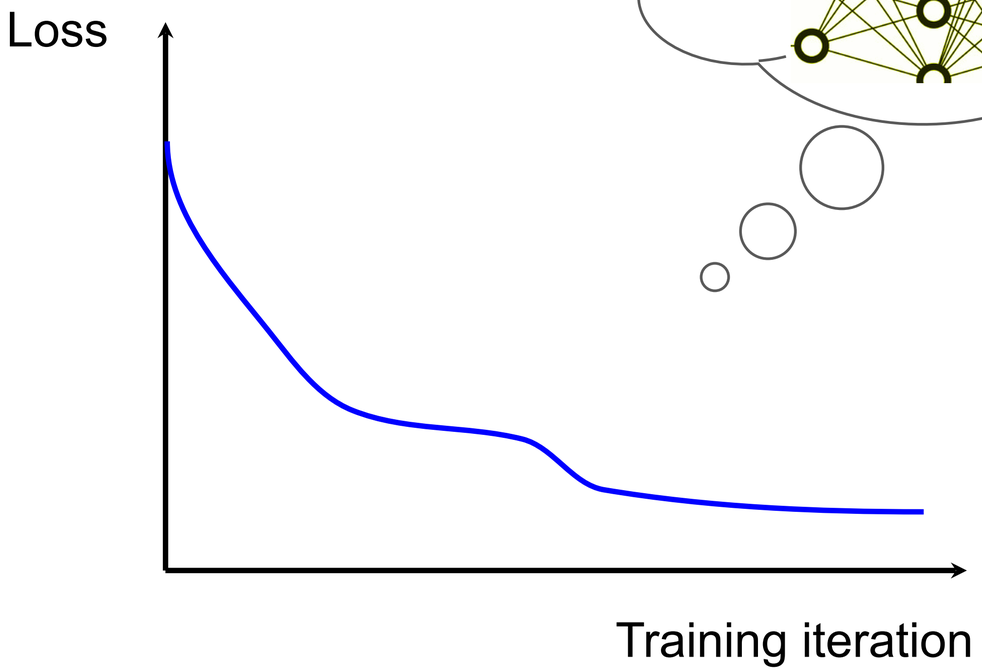
NeurIPS 2019

http://www.rosanneliu.com/publication/lca/


Loss
Training iteration


Loss
Training iteration

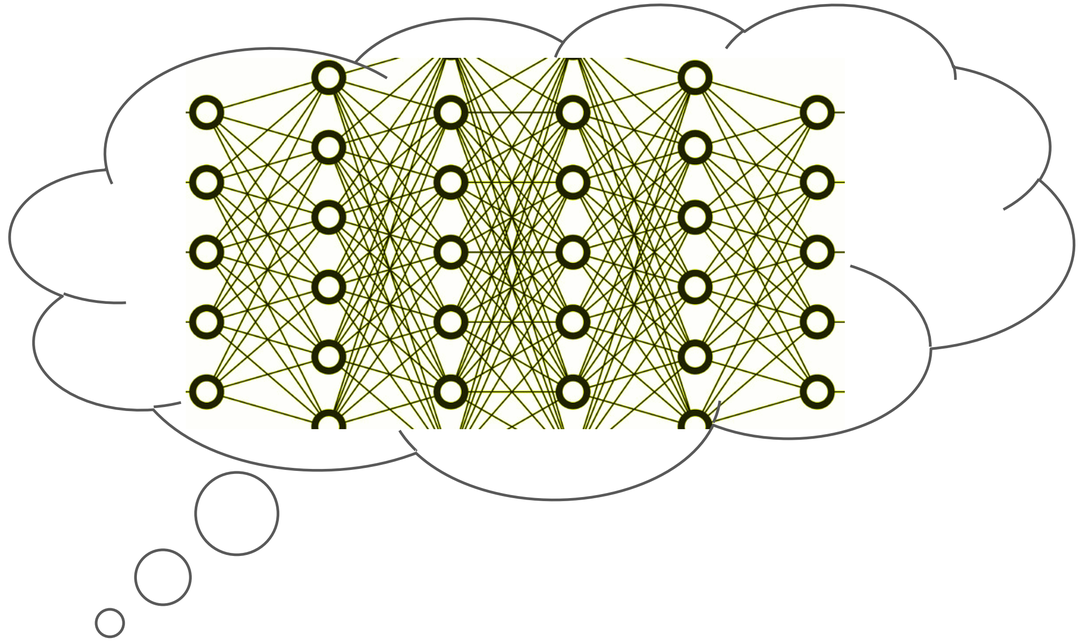


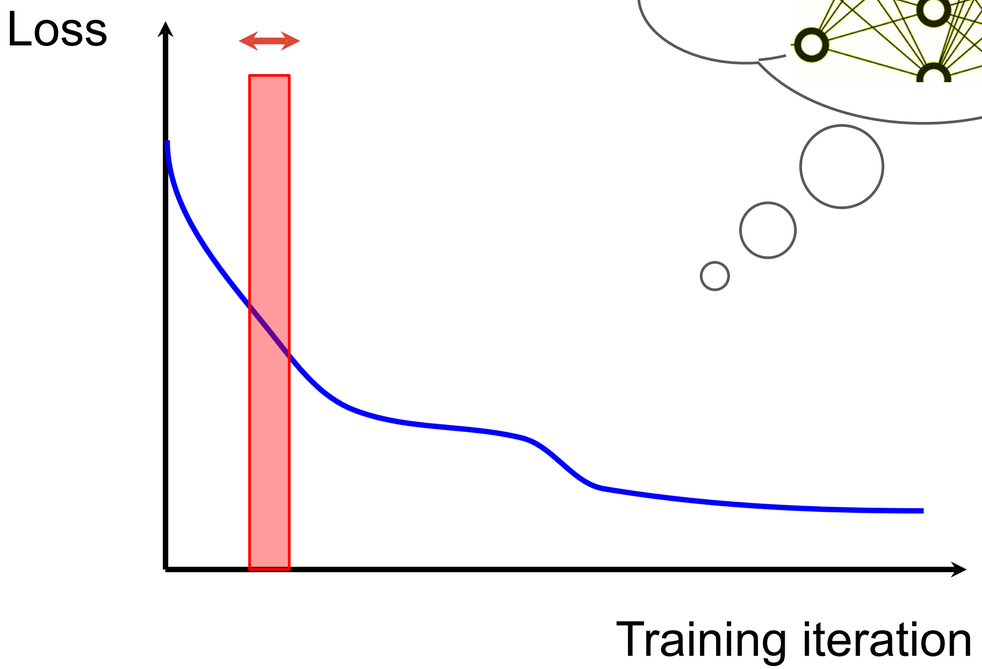


Project 1: LCA

What happens in a single iteration:

- the loss moved;
- (due to the fact that) all parameters moved.



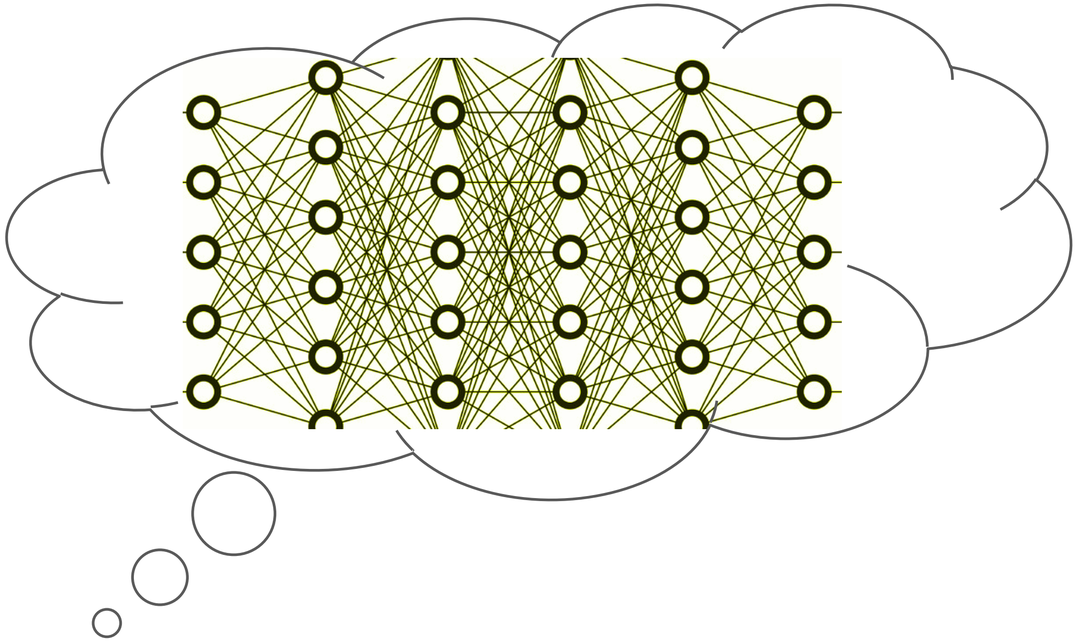


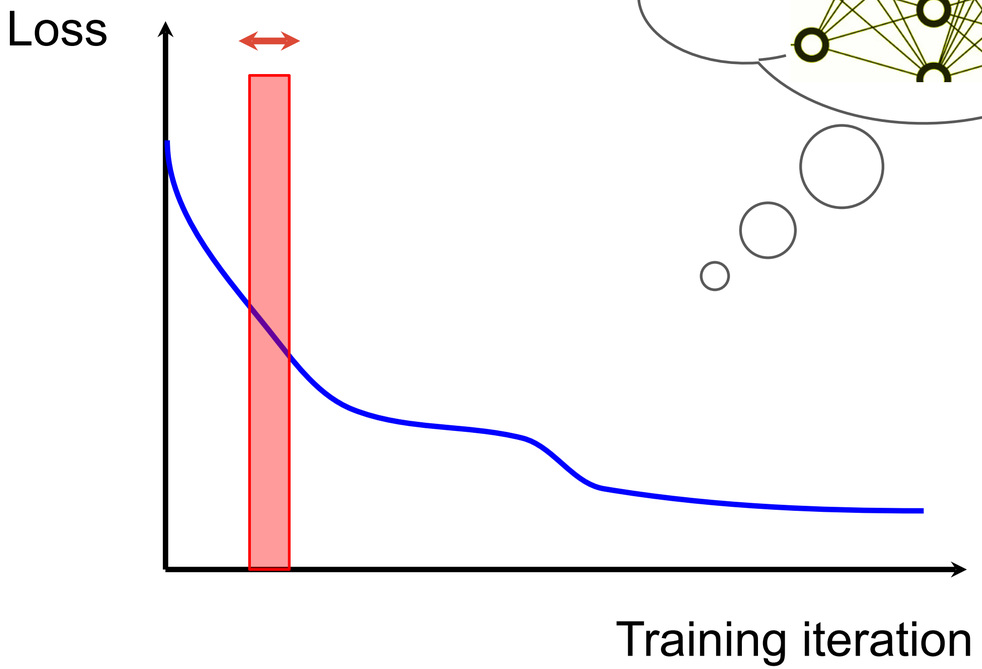
What happens in a single iteration:

- the loss moved;
- (due to the fact that) all parameters moved.

Question:

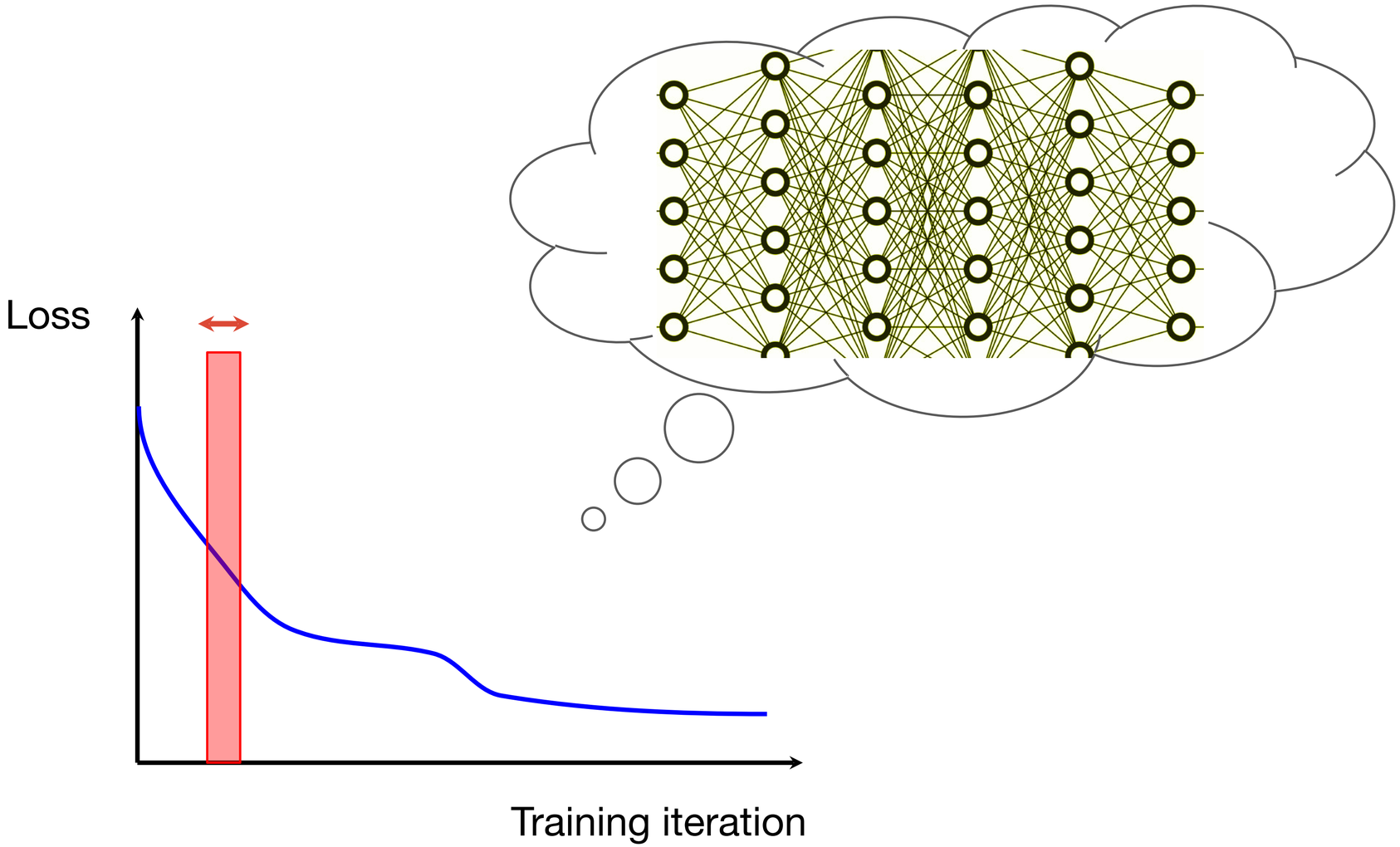
What is every parameter's contribution?





Method:

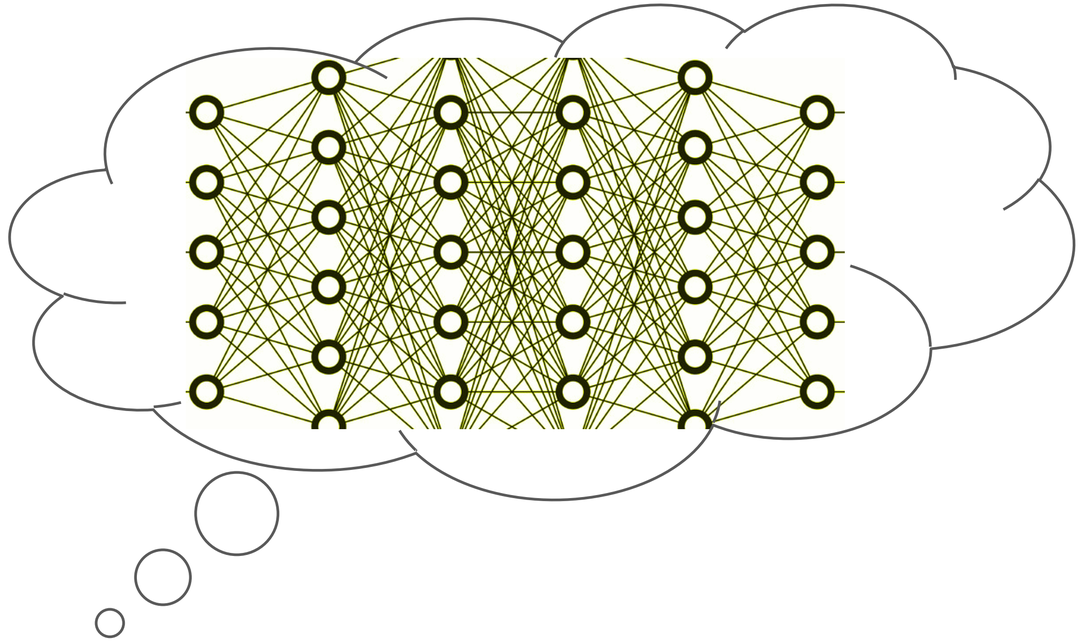
a per-parameter Loss Change Allocation.

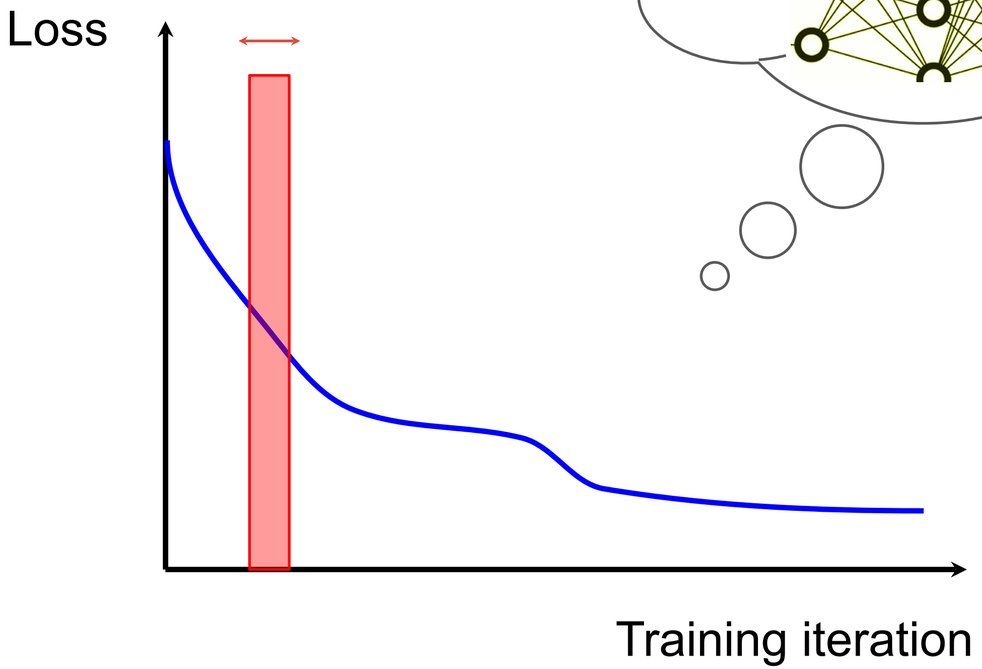


Method:

a per-parameter Loss Change Allocation.

$$L(\theta_{t+1}) - L(\theta_t) \approx \langle \nabla_\theta L(\theta_t),\ \theta_{t+1} - \theta_t \rangle$$
$$= \sum_{i=0}^{K-1} (\nabla_\theta L(\theta_t))^{(i)} (\theta_{t+1}^{(i)} - \theta_t^{(i)}) := \sum_{i=0}^{K-1} C_{t,i}$$





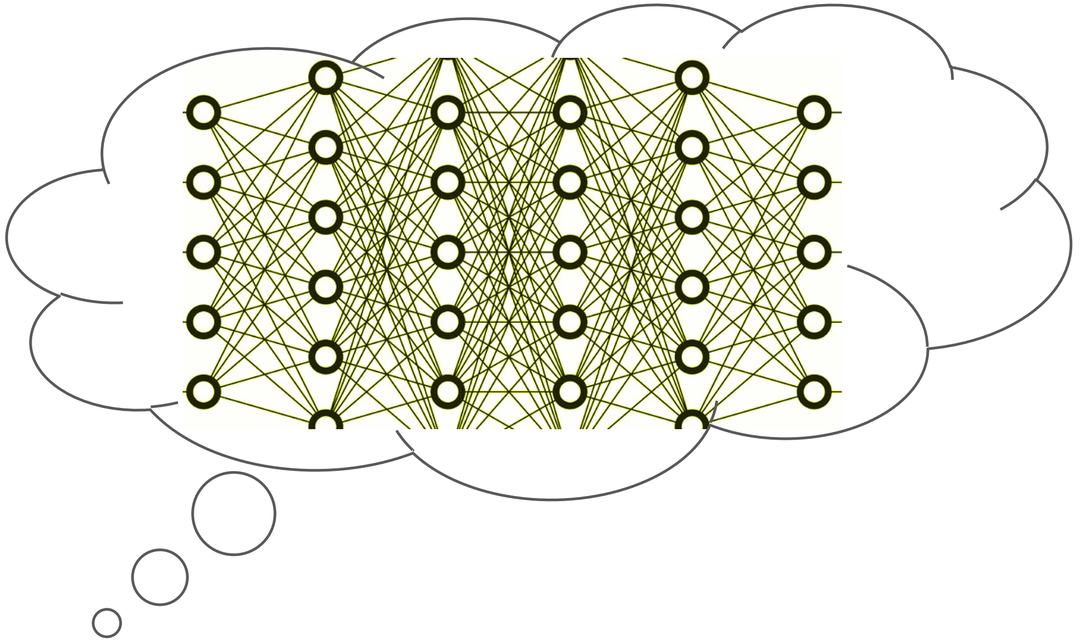
Method:

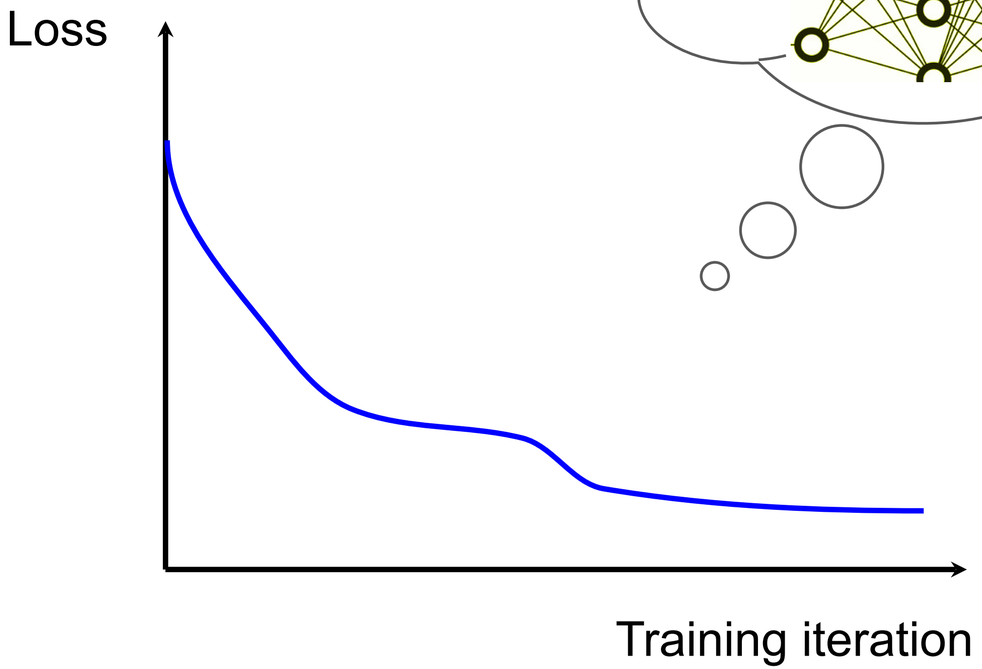
a per-parameter Loss Change Allocation.

$$L(\theta_{t+1}) - L(\theta_t) \approx \langle \nabla_\theta L(\theta_t),\ \theta_{t+1} - \theta_t \rangle$$
$$= \sum_{i=0}^{K-1} (\nabla_\theta L(\theta_t))^{(i)} (\theta_{t+1}^{(i)} - \theta_t^{(i)}) := \sum_{i=0}^{K-1} C_{t,i}$$

What it means:

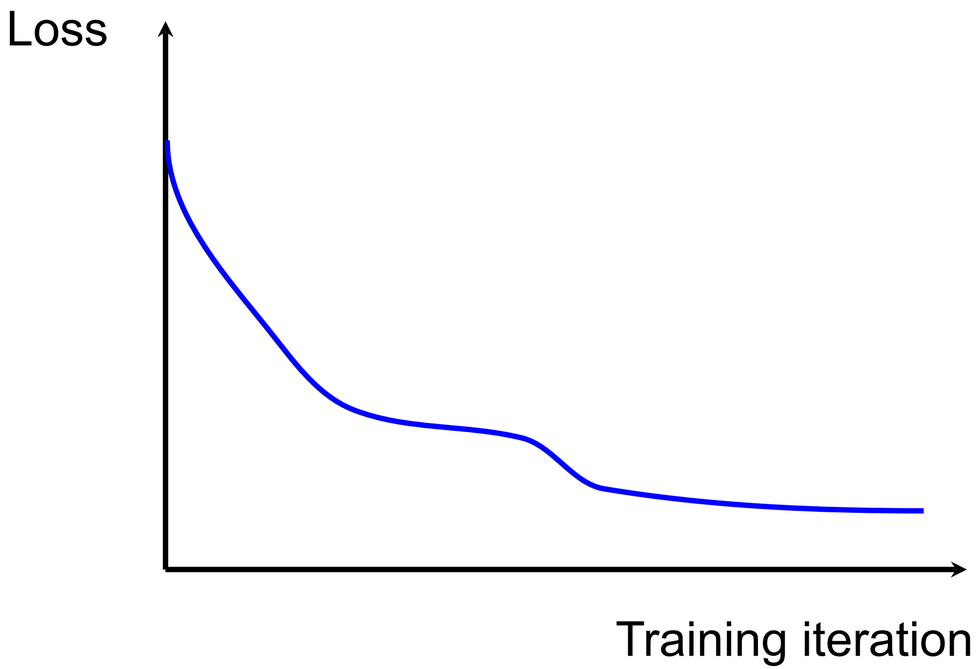
- K loss curves for K parameters in a network!
- Sum them all, you'd recover the loss curve.
- Sum over every neuron, channel, layer...





How do you imagine K loss curves are like?

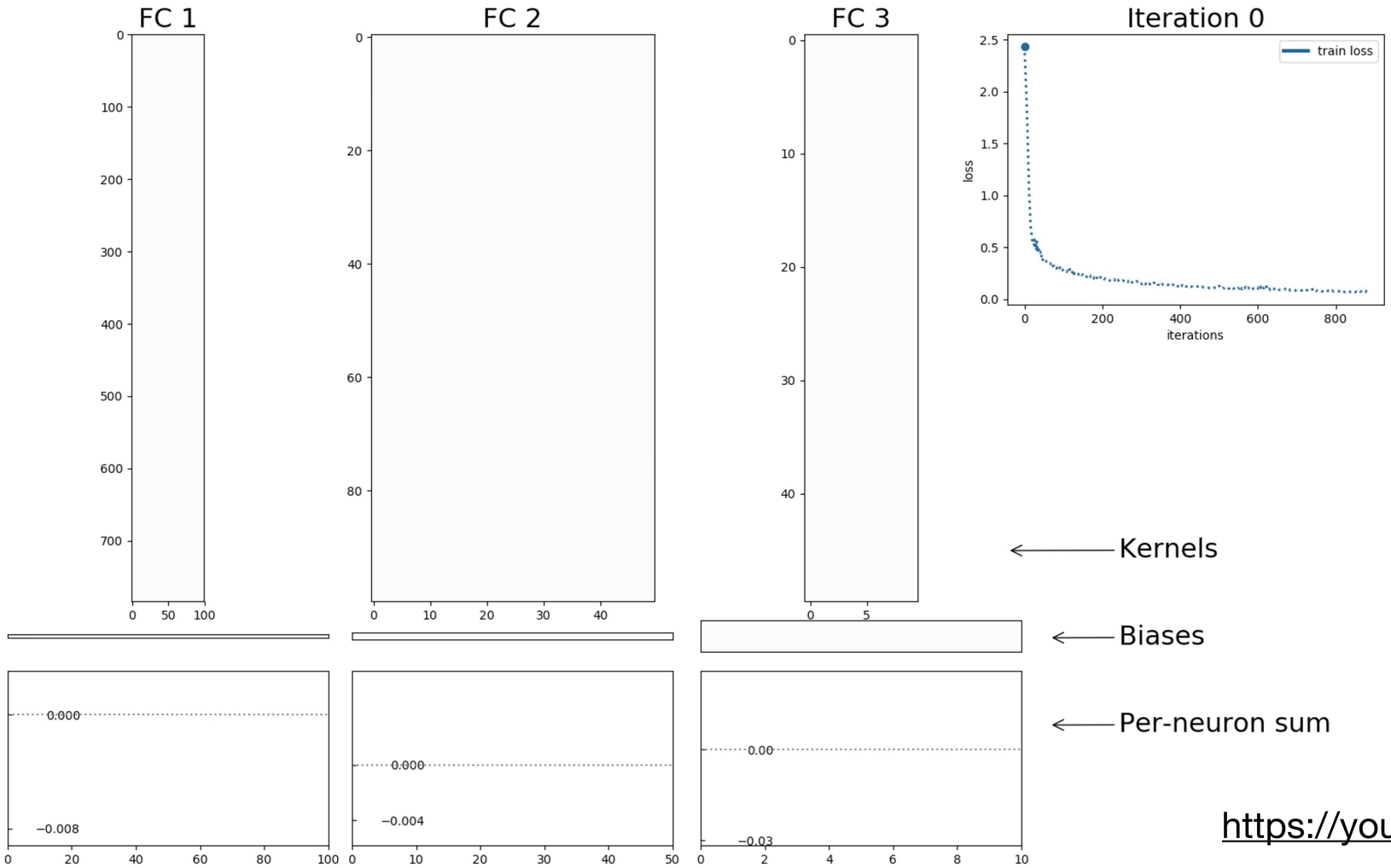
Hint: not all of them are monotonically going down.



How do you imagine K loss curves are like?

Hint: not all of them are monotonically going down.

MNIST;
3-layer FC



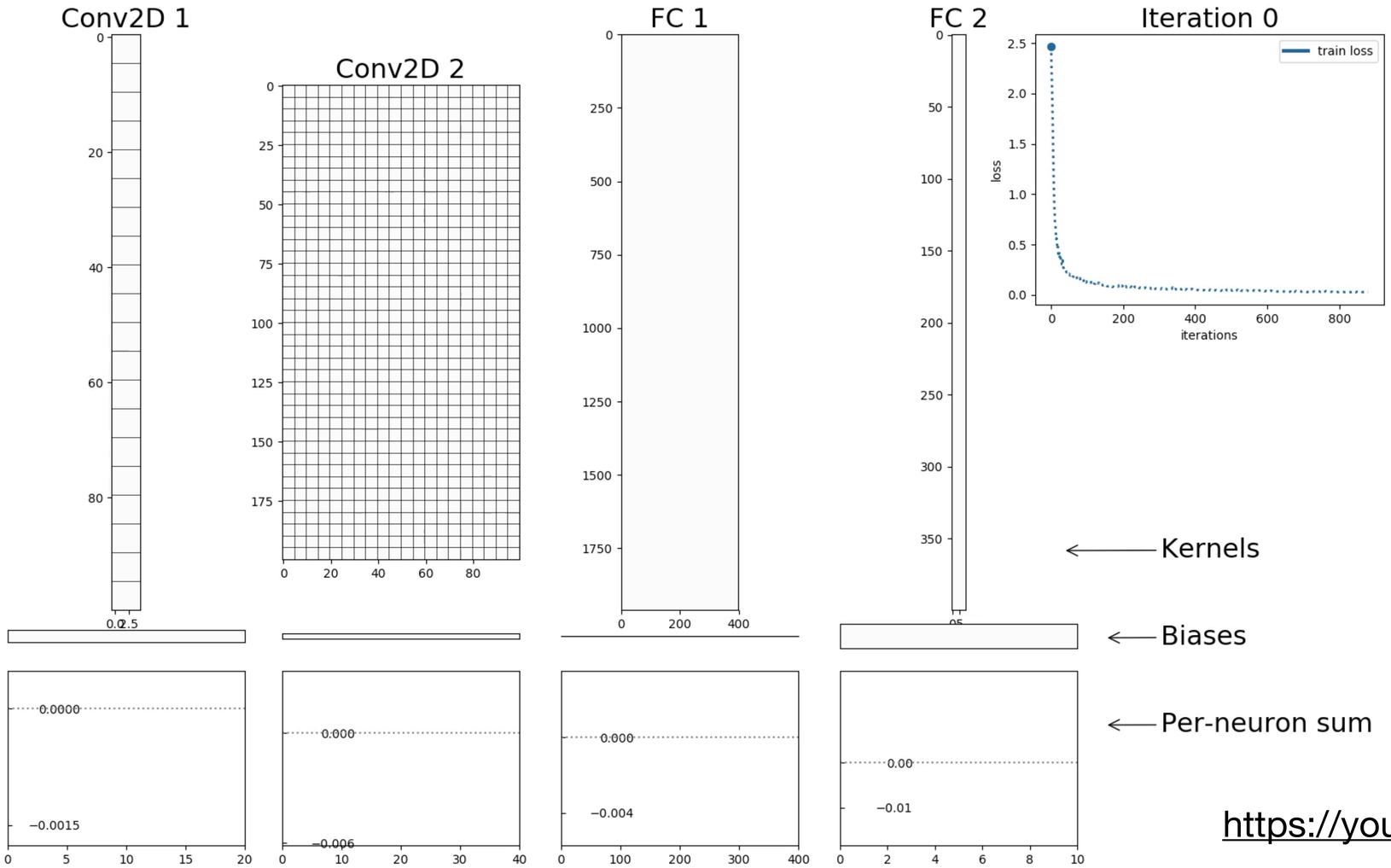
https://youtu.be/xcnoRnoVyXQ

Project 1: LCA

How do you imagine K loss curves are like?

Hint: not all of them are monotonically going down.

MNIST;
LeNet



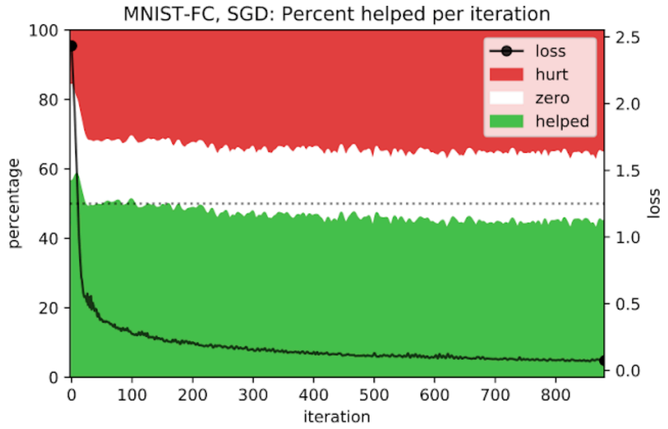
https://youtu.be/EY3LoXmdkYU

Project 1: LCA

Plot lots of things to understand training

Plot lots of things to understand training

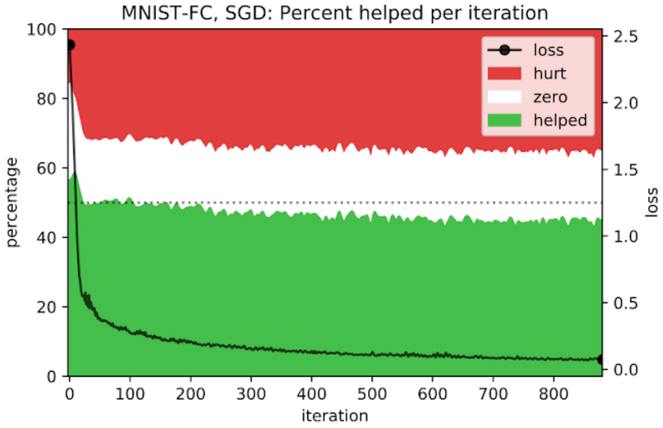
MNIST-FC, SGD

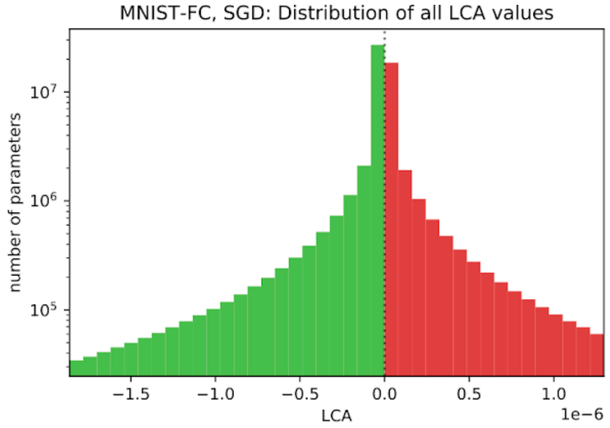


Project 1: LCA

Plot lots of things to understand training

MNIST-FC, SGD

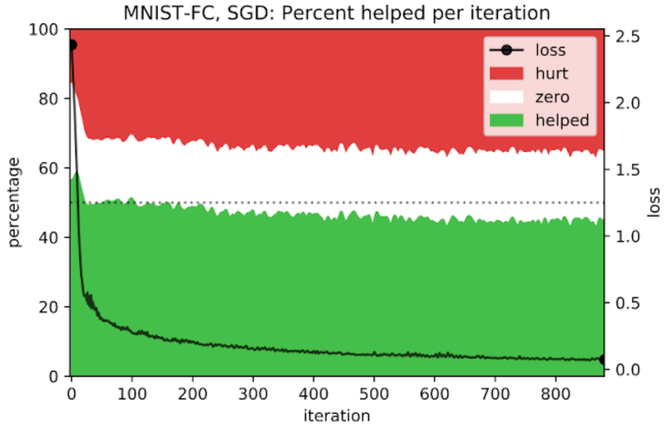


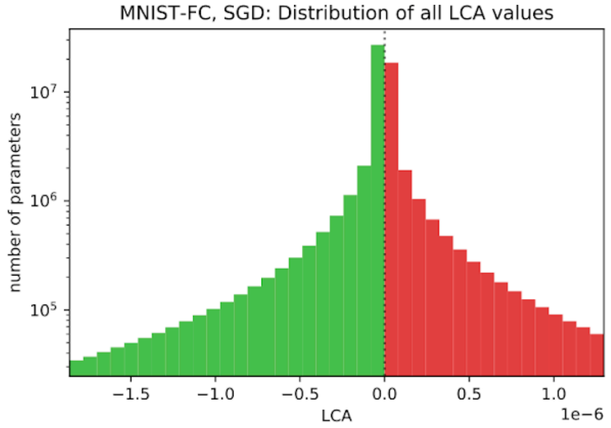


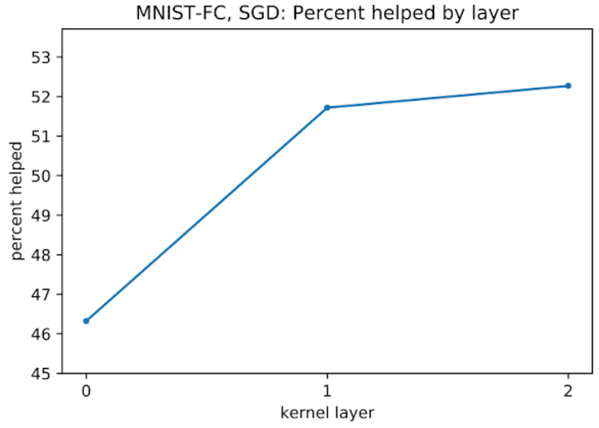
Project 1: LCA

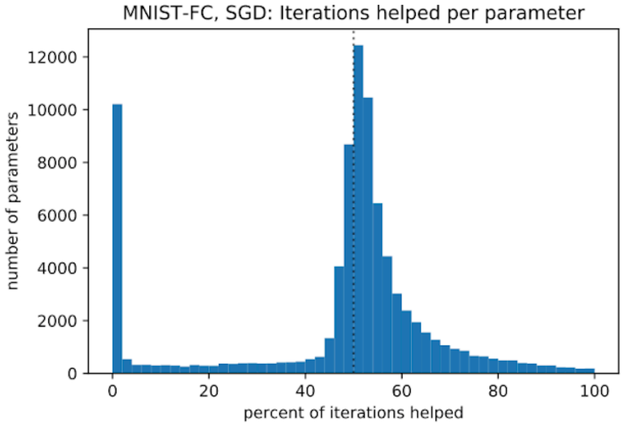
Plot lots of things to understand training

MNIST-FC, SGD





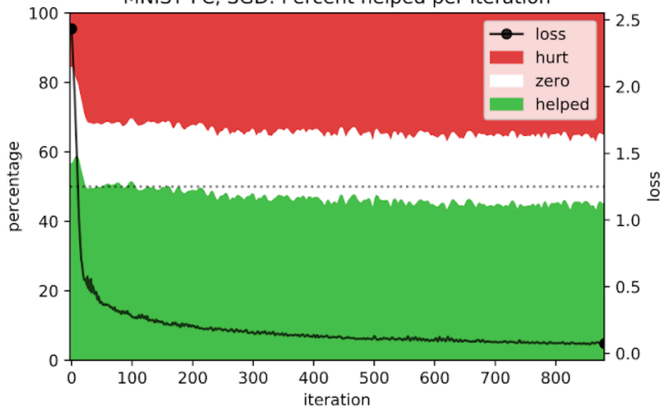


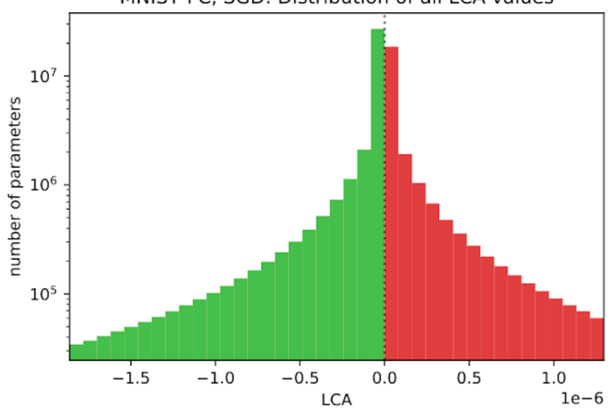


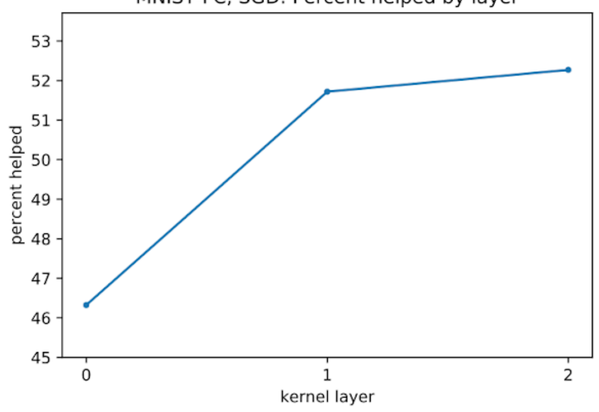
Project 1: LCA

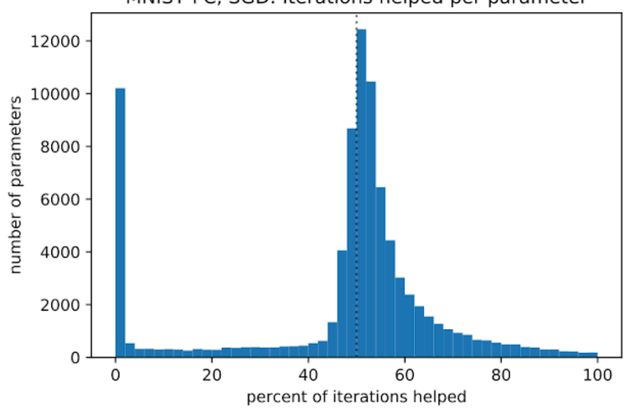
Plot lots of things to understand training

MNIST-FC, SGD

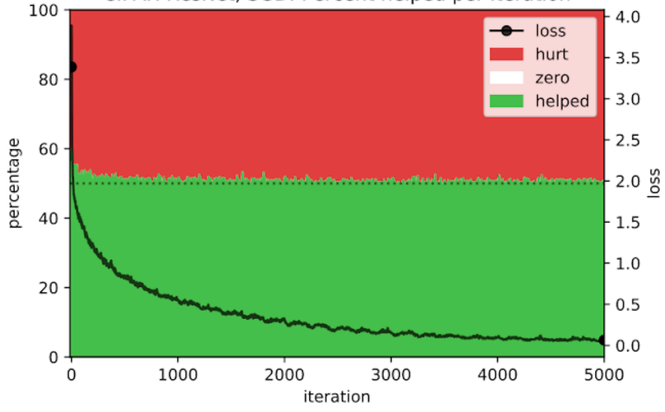


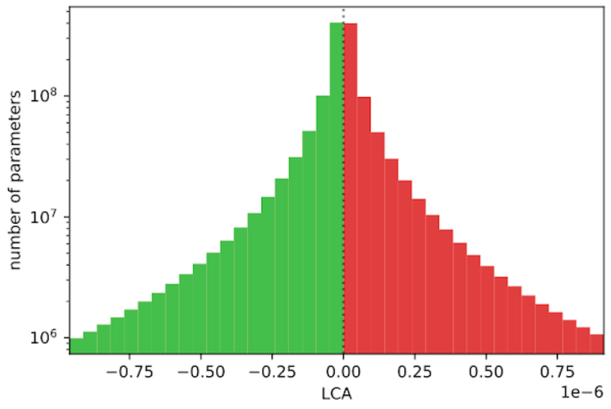


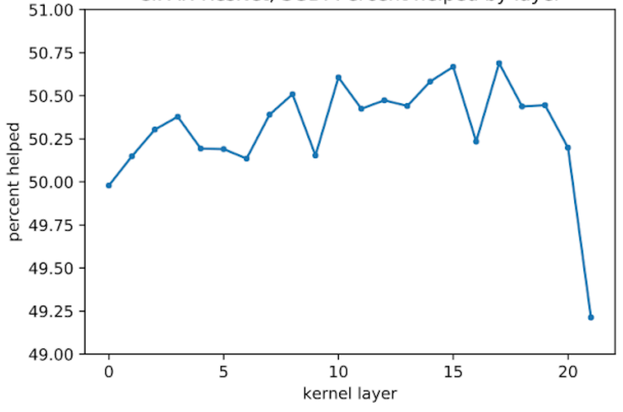


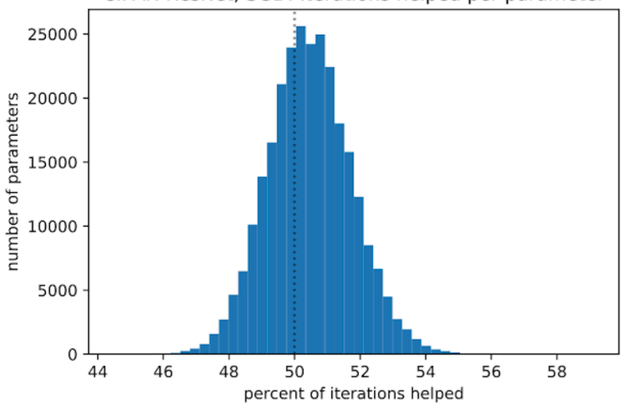


CIFAR-ResNet, SGD





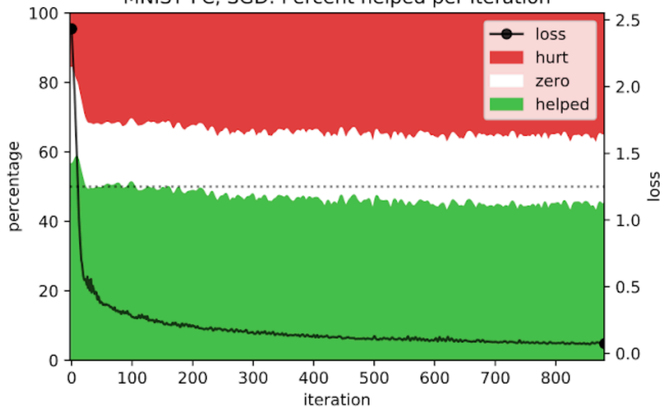


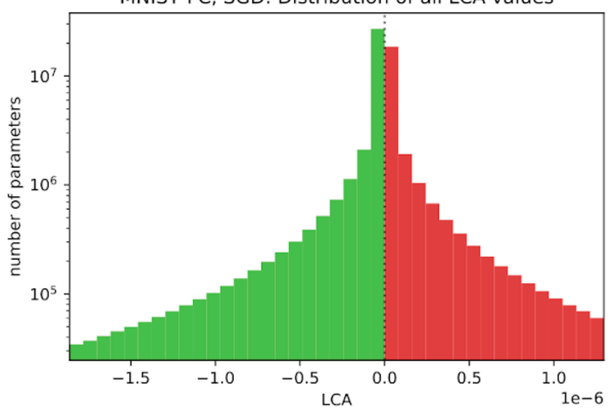


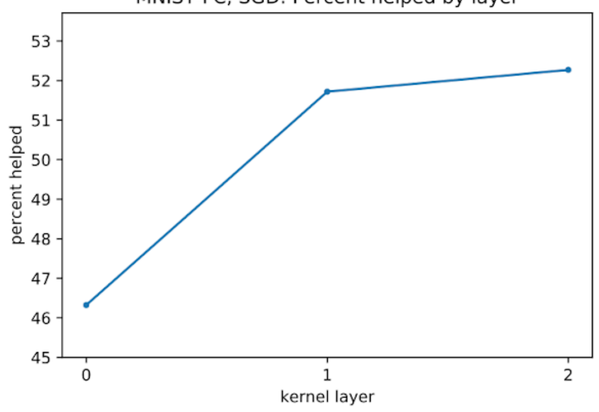
Project 1: LCA

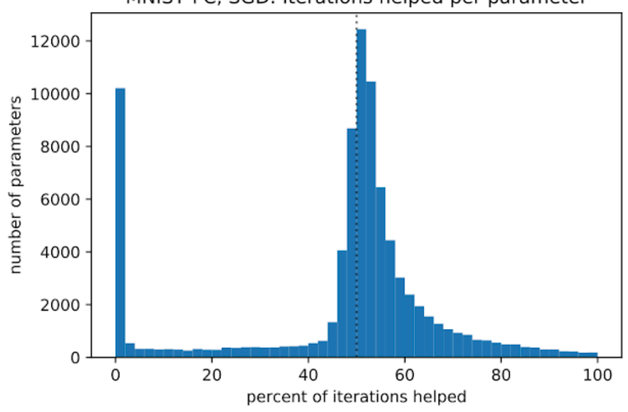
Plot lots of things to understand training

MNIST-FC, SGD

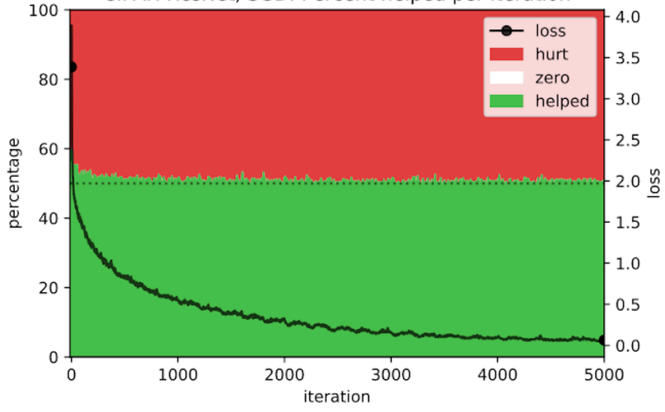


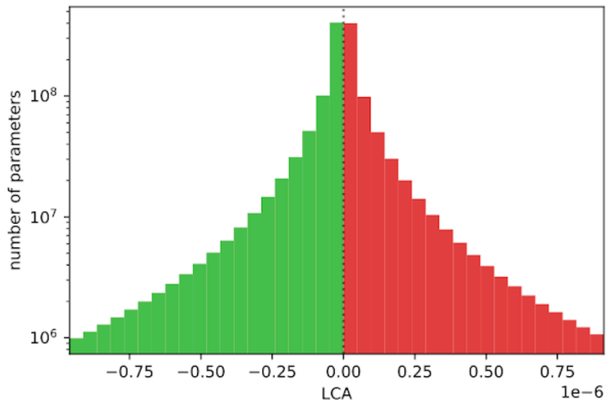


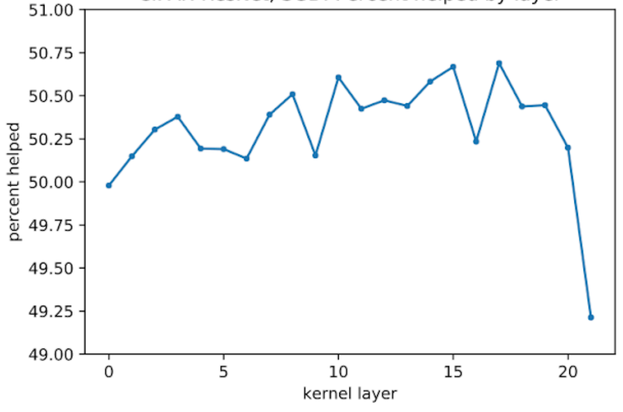


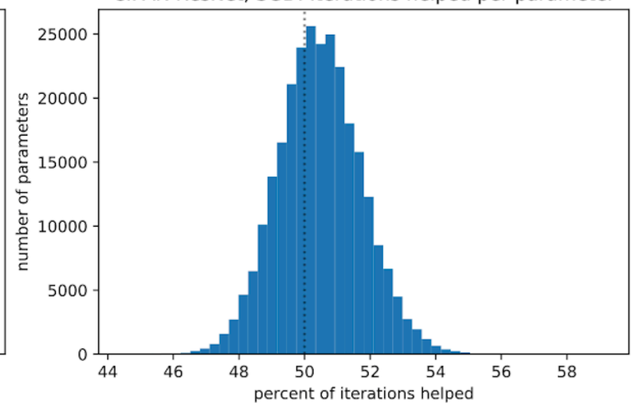


CIFAR-ResNet, SGD









Learning is noisy!

- Barely over 50% of parameters help during training.
- Each parameter hurt almost 50% of the time.
- Learning is heavy-tailed.

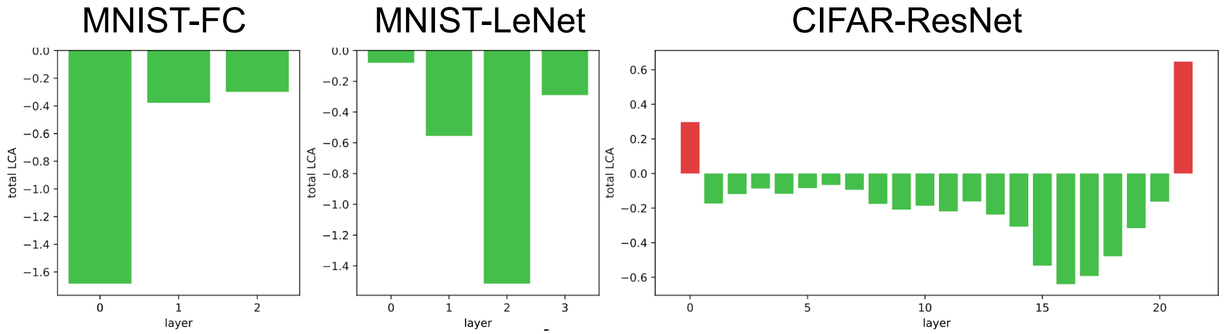
Insight 1: Learning is noisy

Percentage of helping parameters

	MNIST-FC, mom=0	MNIST-FC	MNIST-LeNet	CIFAR-ResNet	CIFAR-AllCNN
SGD	53.72 ± 0.05	57.79 ± 0.16	53.97 ± 0.48	50.66 ± 0.14	51.09 ± 0.23
Adam	N/A	55.82 ± 0.09	51.77 ± 0.21	50.30 ± 0.004	50.19 ± 0.01

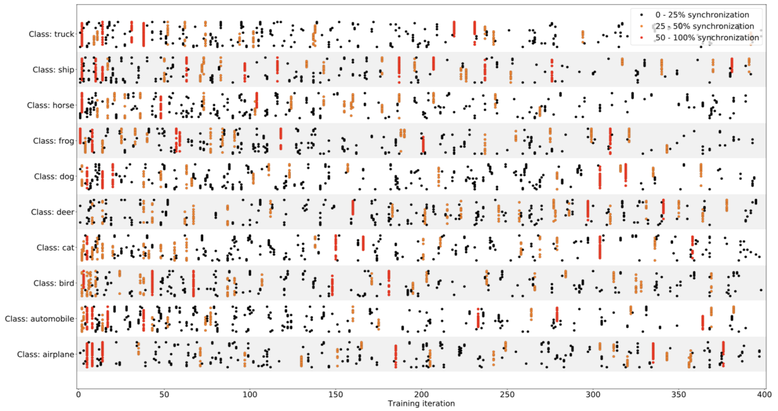
Insight 2: Some layers hurt overall

Layer-summed LC. Negative (green) is helping



Insight 3: Learning is synchronized across layers

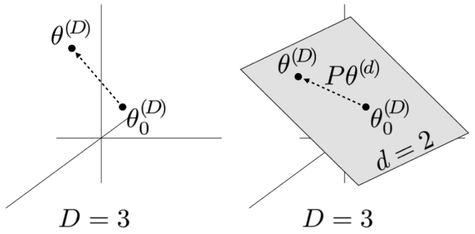
CIFAR_ResNet, per-class layer synchronization (in red)



Project 1: LCA

LCA: Loss Change Allocation for Neural Network Training
J. Lan, R. Liu, H. Zhou, J. Yosinski. **NeurIPS 2019.**

Intrinsic Dimension



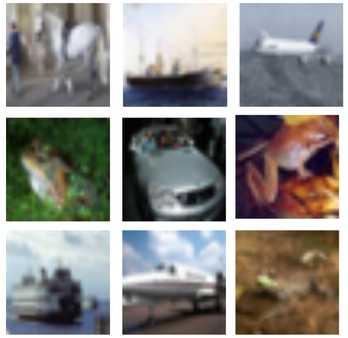
+ Chunyuan Li, Heerad Farkhoor, Jason Yosinski

ICLR 2018

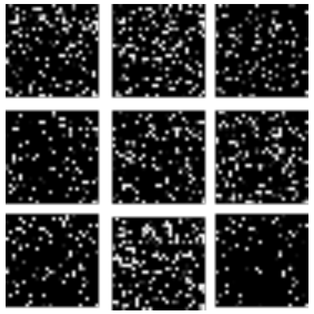
http://www.rosanneliu.com/publication/intrinsic/

MNIST

CIFAR 10



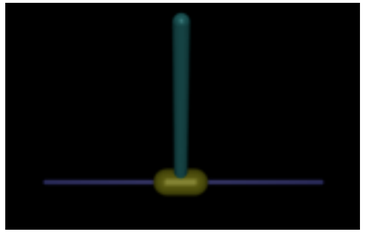
MNIST Shuffled-pixels



Humanoid

Pong

Inverted Pendulum

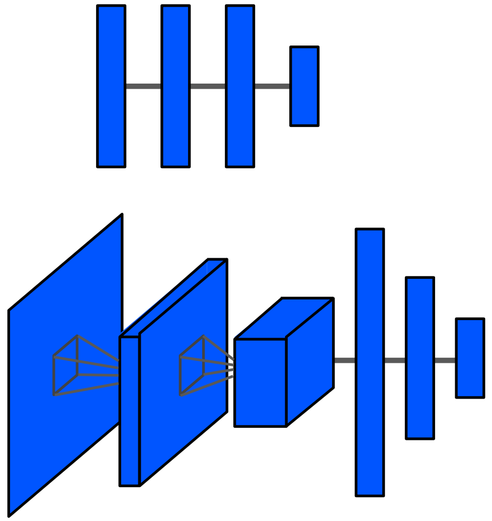


Project II: Intrinsic Dimension

MNIST

Project II: Intrinsic Dimension

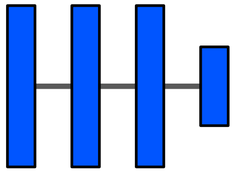
MNIST



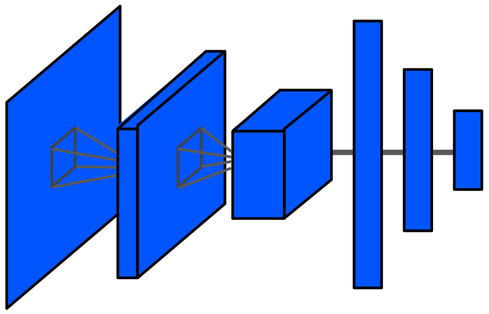
Project II: Intrinsic Dimension

MNIST

Int. Dim.



750

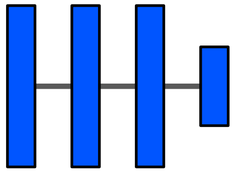


290

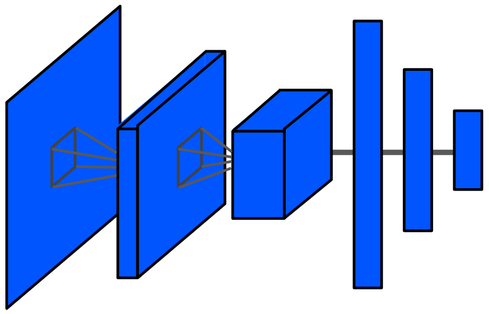
Project II: Intrinsic Dimension

MNIST

Int. Dim.

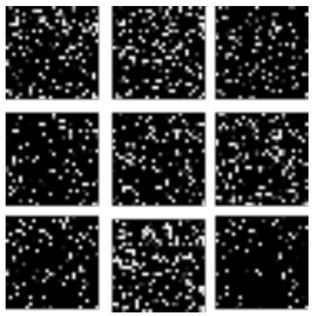


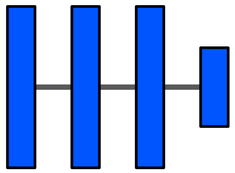
750



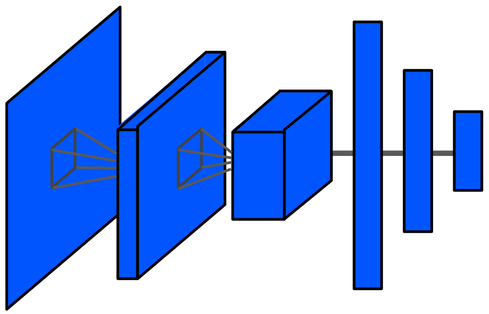
290

MNIST Shuffled-pixels





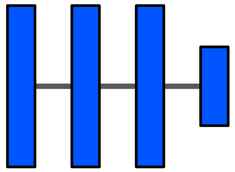
750



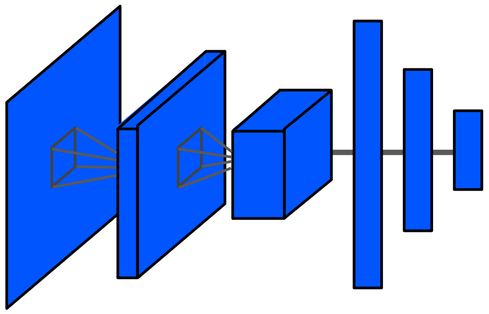
1400

MNIST

Int. Dim.

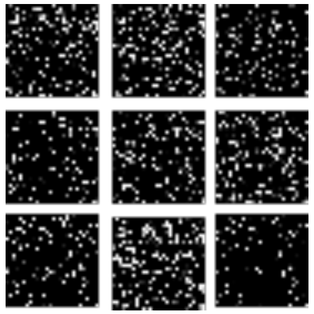


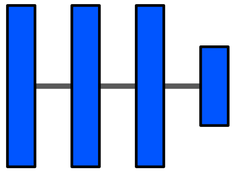
750



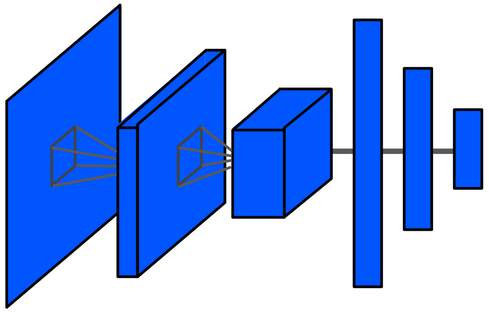
290

MNIST Shuffled-pixels





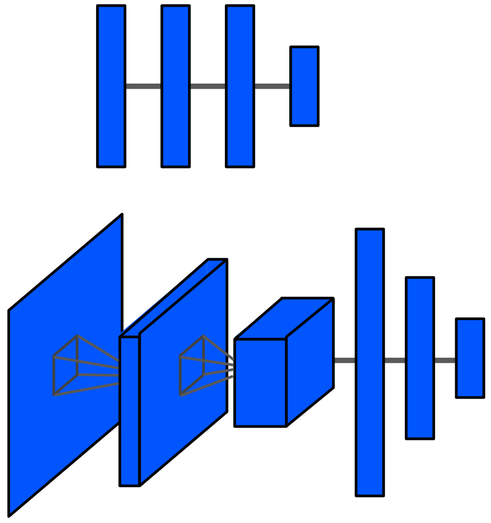
750



1400

MNIST

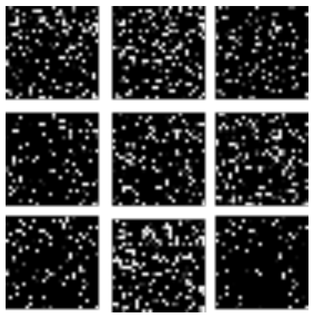
Int. Dim.

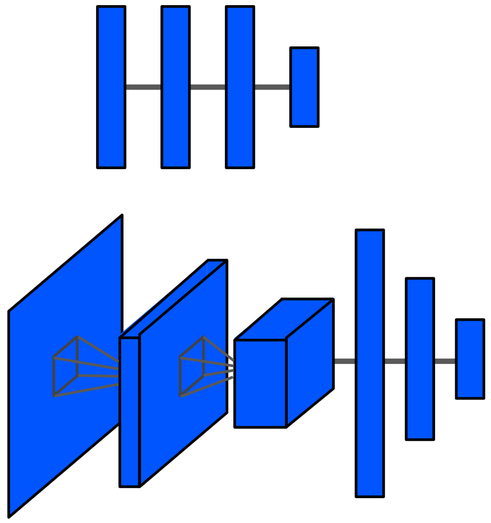


750

290

MNIST Shuffled-pixels



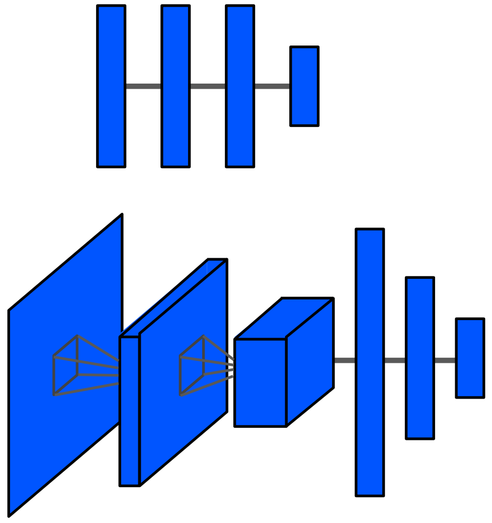


750

1400

MNIST

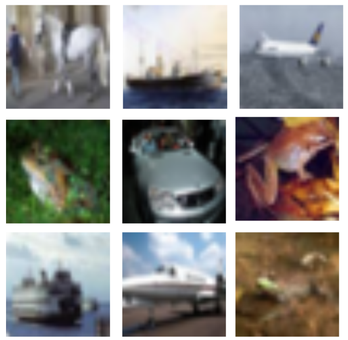
Int. Dim.



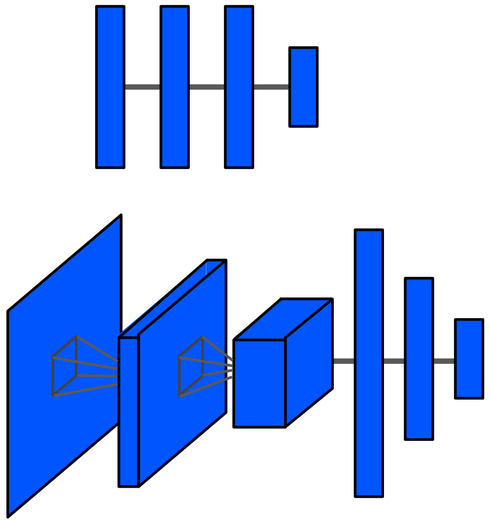
750

290

CIFAR 10



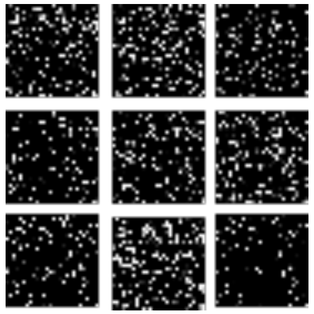
Int. Dim.

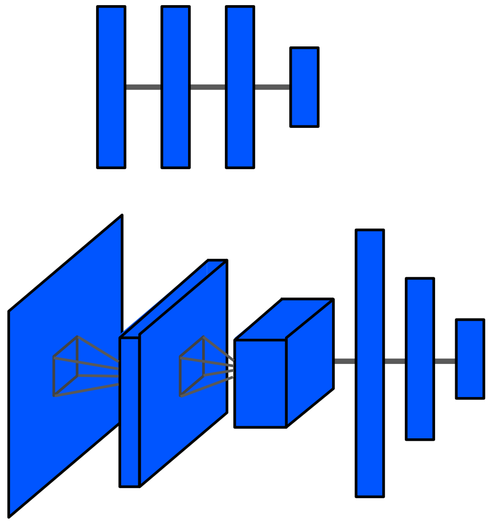


9K

2.9K

MNIST Shuffled-pixels



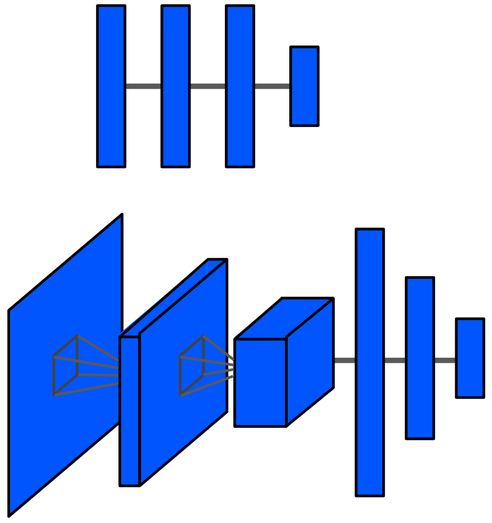


750

1400

MNIST

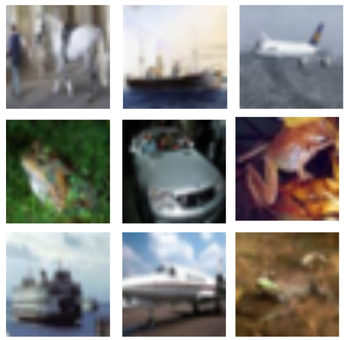
Int. Dim.



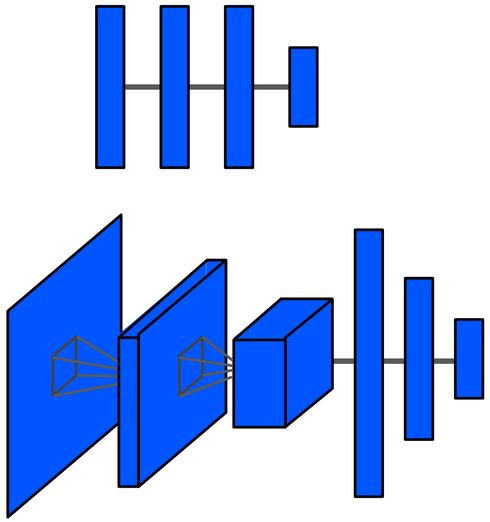
750

290

CIFAR 10



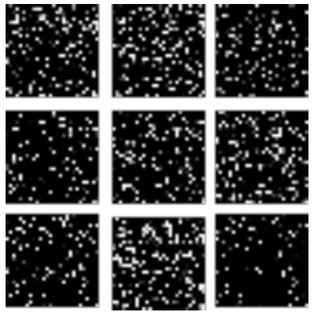
Int. Dim.

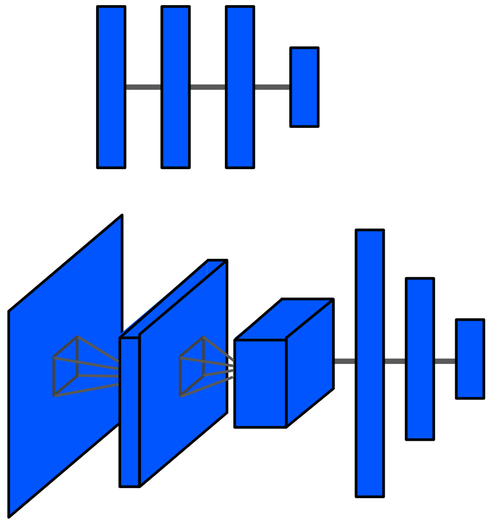


9K

2.9K

MNIST Shuffled-pixels



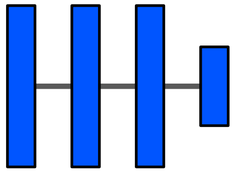


750

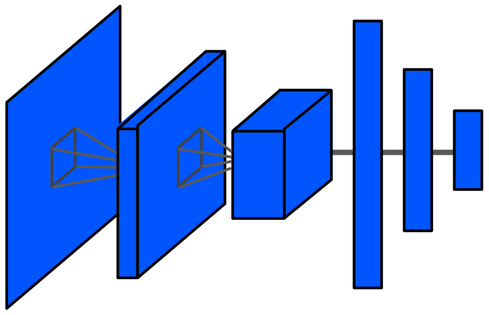
1400

MNIST

Int. Dim.

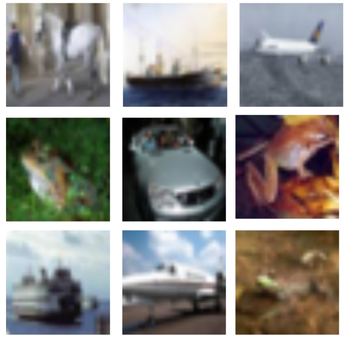


750

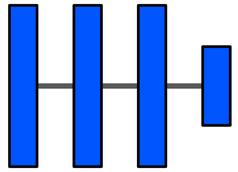


290

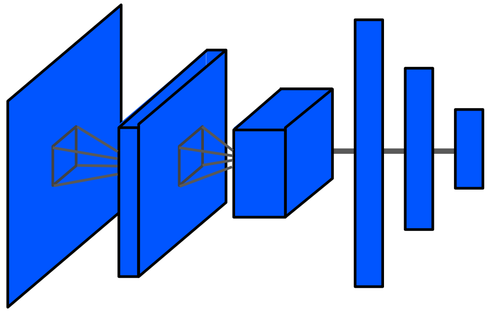
CIFAR 10



Int. Dim.

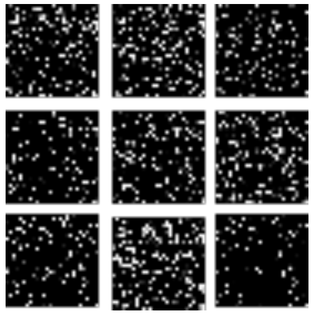


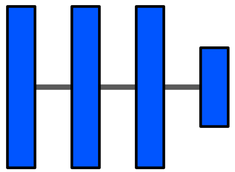
9K



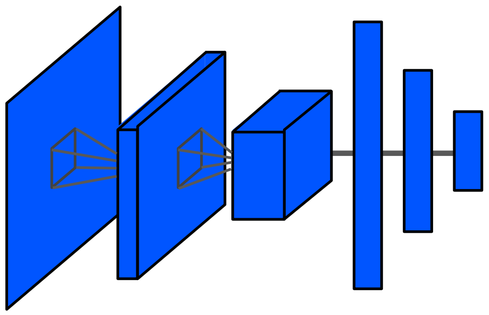
2.9K

MNIST Shuffled-pixels





750



1400

SqueezeNet

>500K

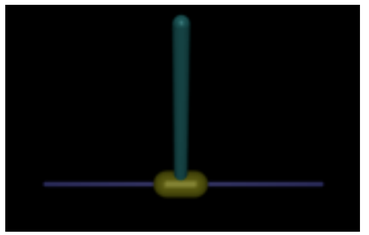
Humanoid

Int. Dim. = 700

Pong

6000

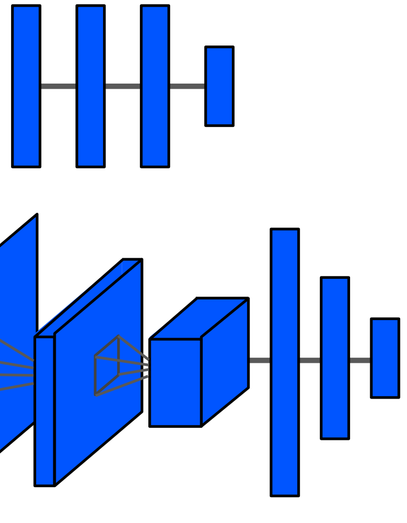
Inverted Pendulum



4

MNIST

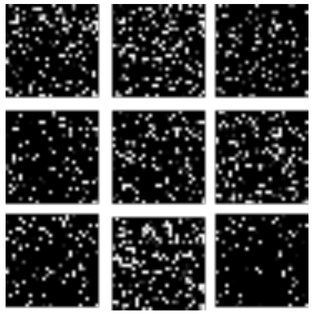
Int. Dim.

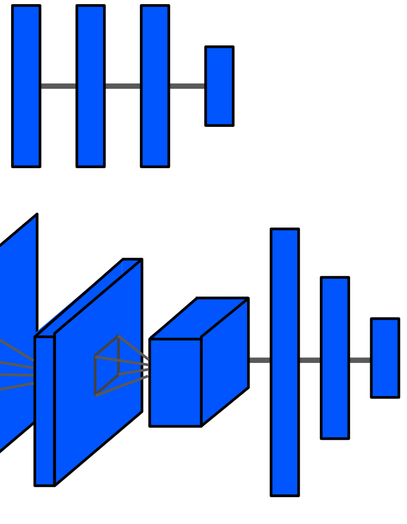


750

290

MNIST Shuffled-pixels

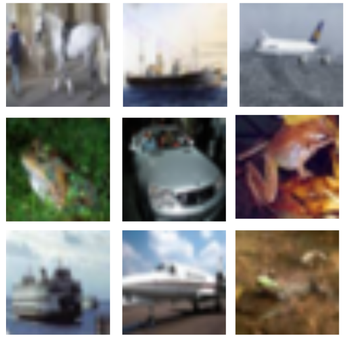




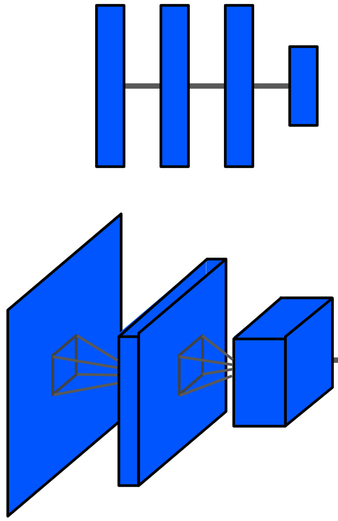
750

1400

CIFAR 10



Int. Dim.



9K

2.9K

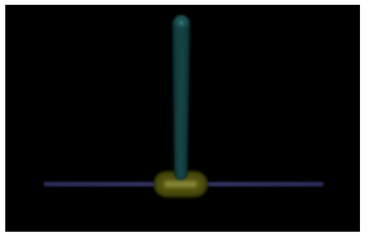
SqueezeNet

>500K

Humanoid

Pong

Inverted Pendulum

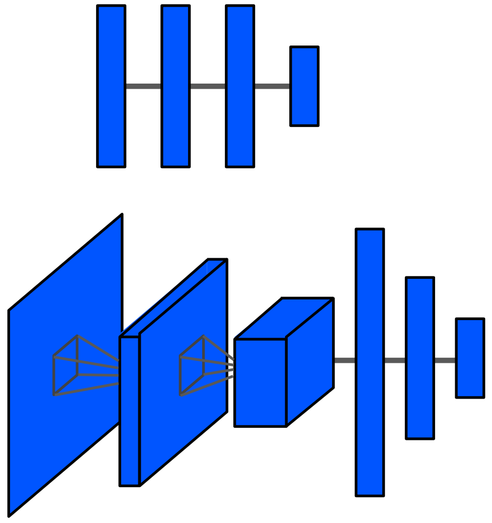


Int. Dim. = 700

6000

4

MNIST

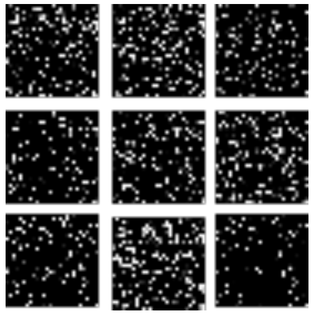


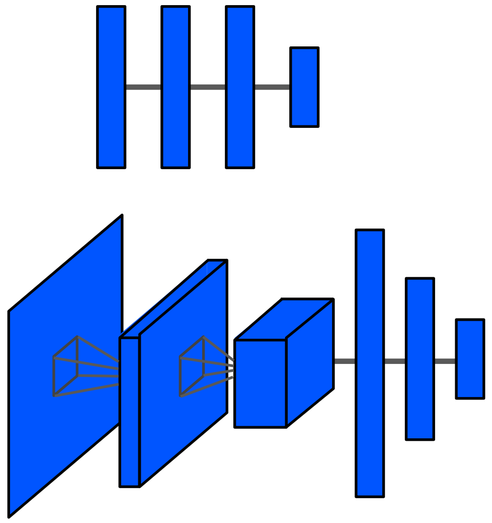
Int. Dim.

750

290

MNIST Shuffled-pixels

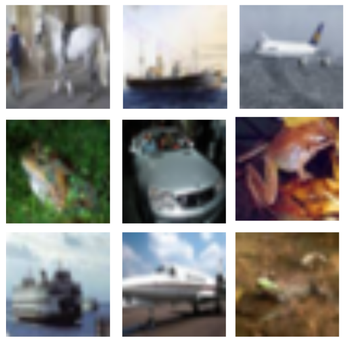




750

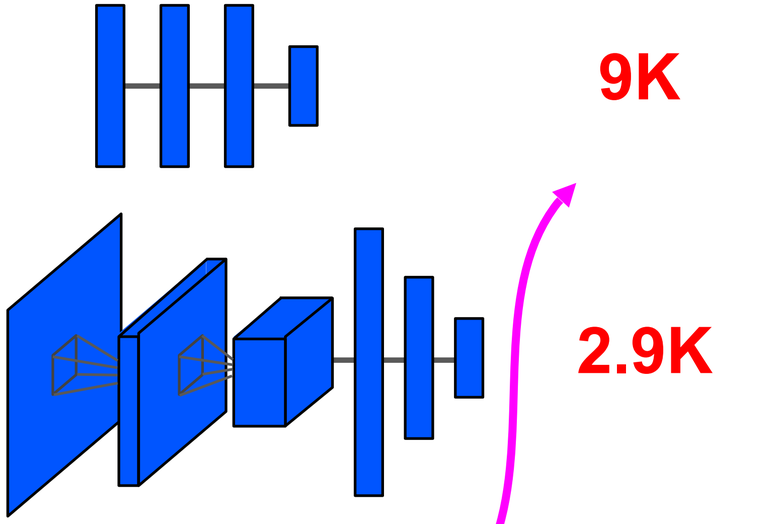
1400

CIFAR 10



Int. Dim.

9K



2.9K

SqueezeNet

>500K

Humanoid

Int. Dim. = 700

Pong

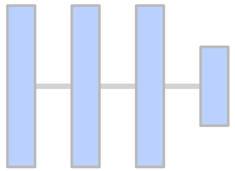
6000

Inverted Pendulum

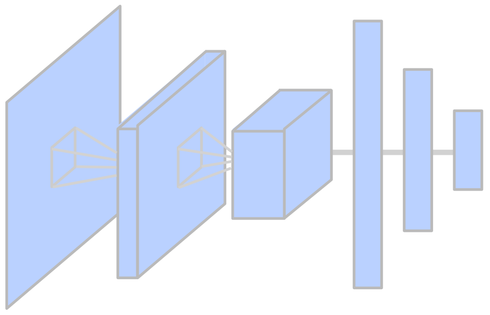
4

MNIST

Int. Dim.



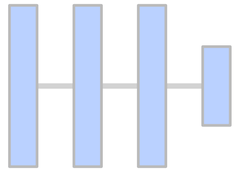
750



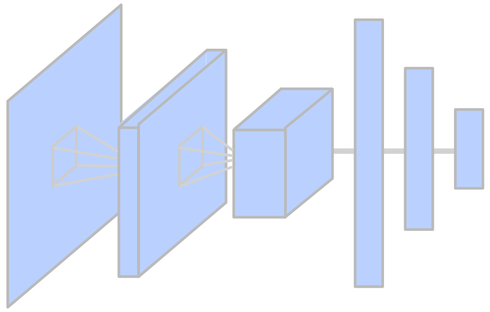
290

CIFAR 10

Int. Dim.



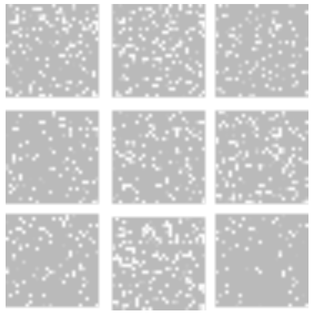
9K

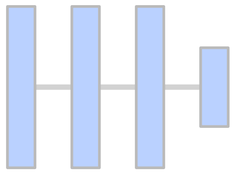


2.9K

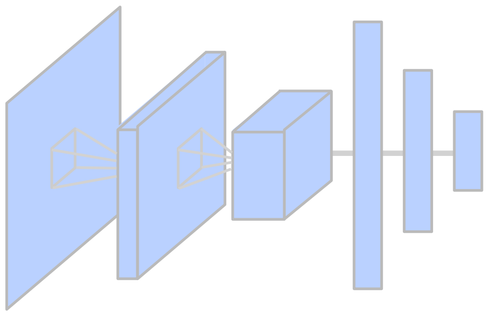
Measuring the Intrinsic Dimension of Objective Landscapes
C. Li, H. Farkhoor, R. Liu, J. Yosinski. **ICLR 2018.**

MNIST Shuffled-pixels





750

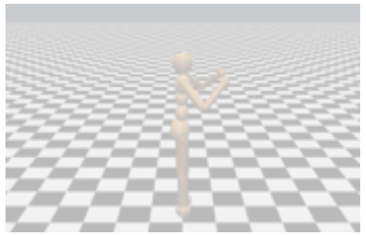


1400

zeNet

>500K

Humanoid



Int. Dim. = 700

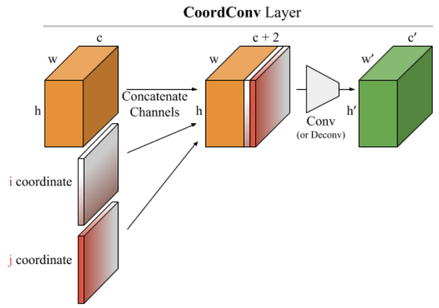
Pong

6000

Inverted Pendulum

4

CoordConv

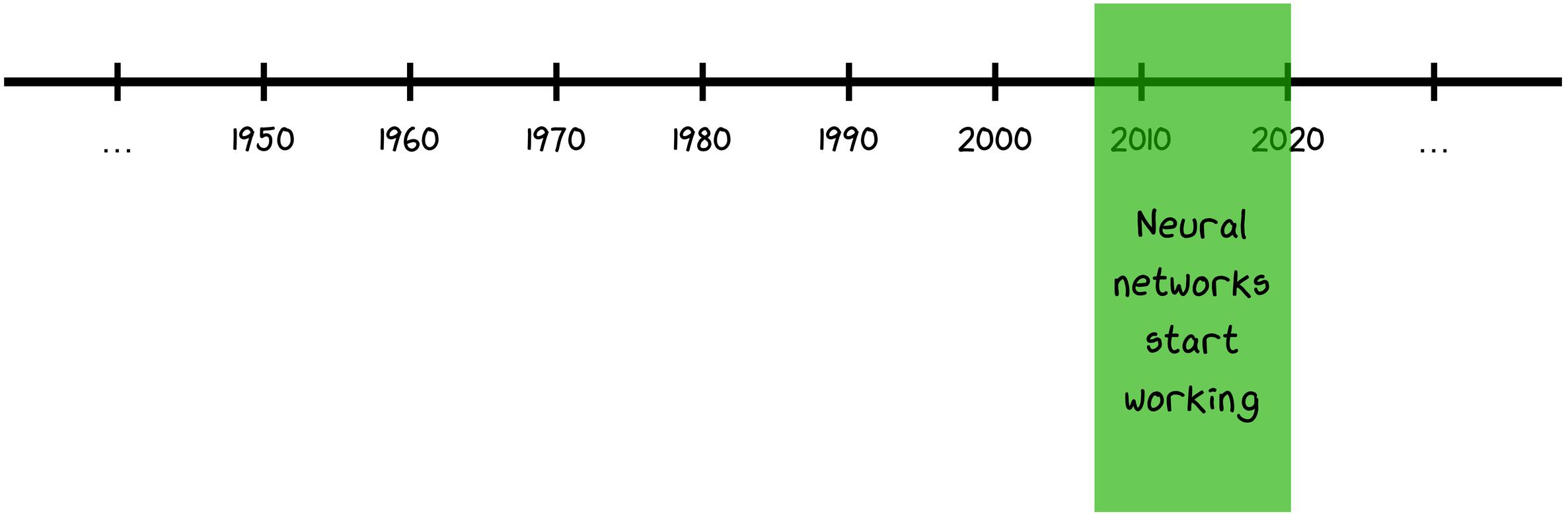


+ Joel Lehman, Piero Molino, Felipe Petroski Such, Eric Frank, Alex Sergeev, Jason Yosinski

NeurIPS 2018

http://www.rosanneliu.com/publication/coordconv/

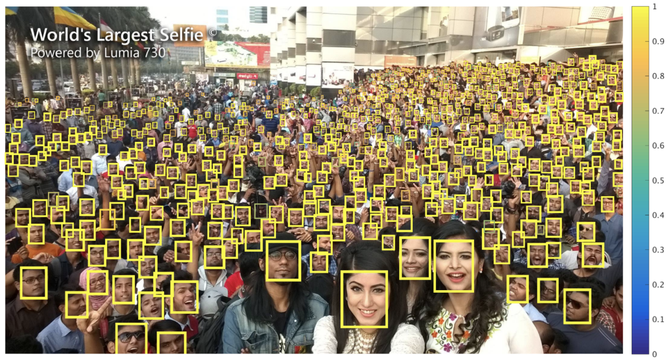
Progress in AI



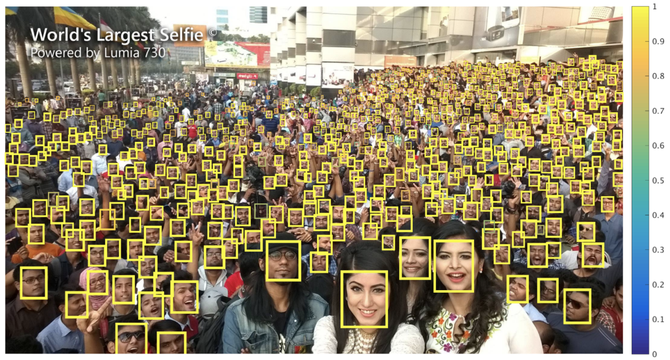
Project III: CoordConv

Good boy!
(golden retriever)

Good boy!
(golden retriever)

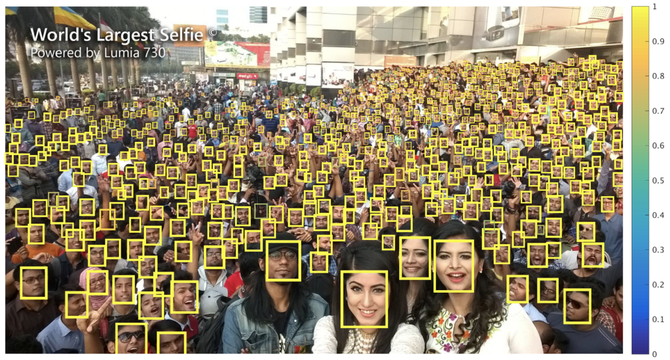


Good boy!
(golden retriever)



Actions

Good boy!
(golden retriever)



Actions

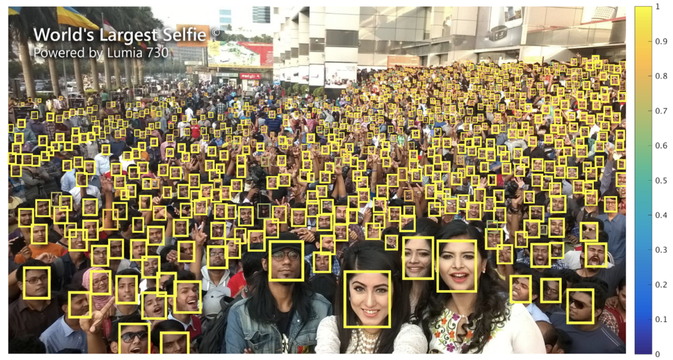
Noise

Project III: CoordConv

Convolutional neural networks
(convnet)

Good boy!
(golden retriever)

convnet



convnet

Actions

convnet

Noise

Project III: CoordConv

[0.23, 1,45, 2.3, 3,03, 1,21, …]

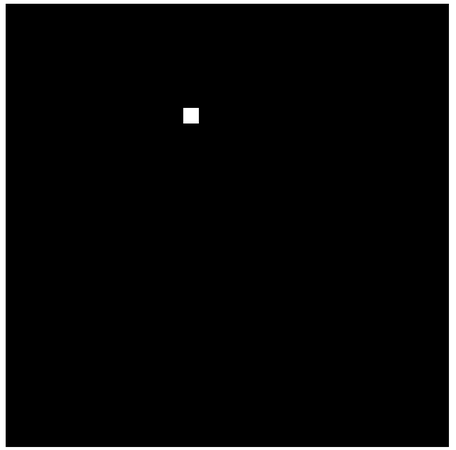
convnet

[0.23, 1,45, 2.3, 3,03, 1,21, …]

convnet

“(4, 6)”

convnet

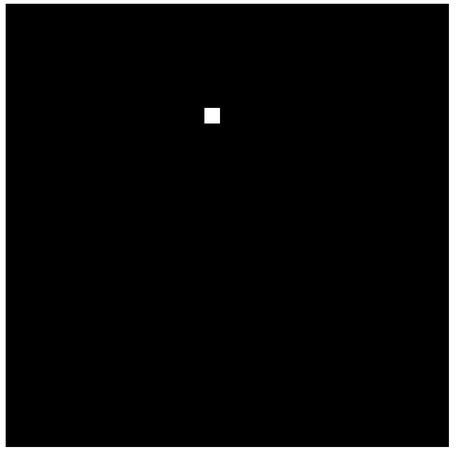


[0.23, 1,45, 2.3, 3,03, 1,21, …]

convnet

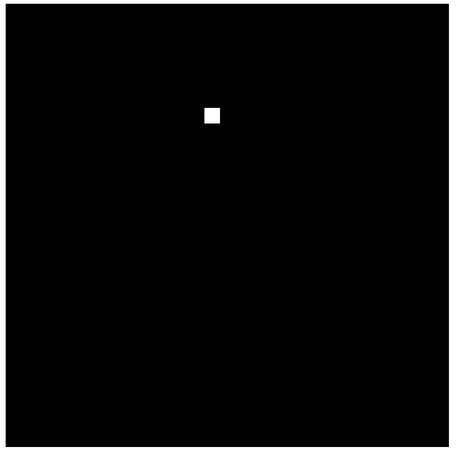
“(4, 7)”

convnet



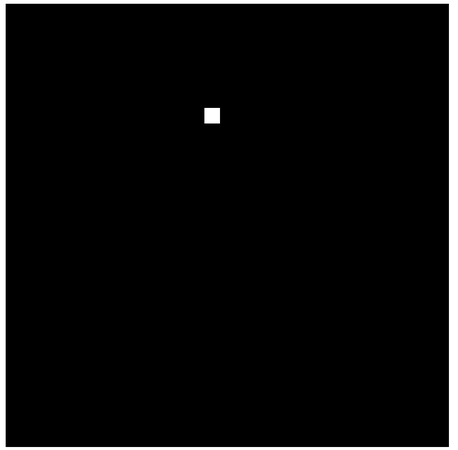
“(4, 7)”

convnet



Coordinate Transform:
Given a *Cartesian* location, highlight *that pixel* on a canvas.

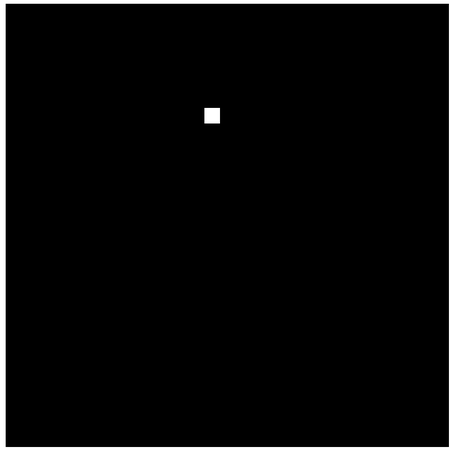
“(4, 7)” convnet

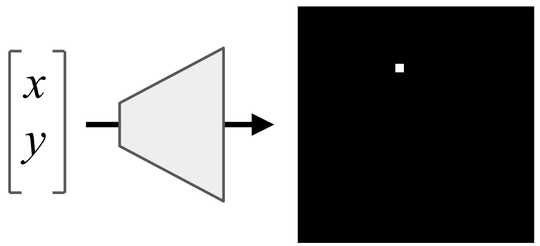


Convnets fail at this simple pixel task

Coordinate Transform:
Given a *Cartesian* location, highlight *that pixel* on a canvas.

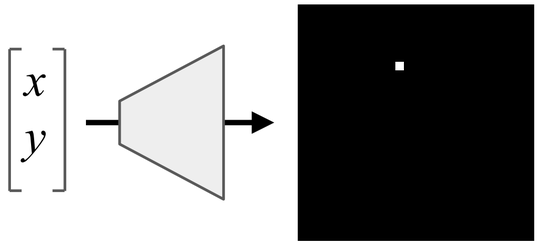
“(4, 7)” → convnet →





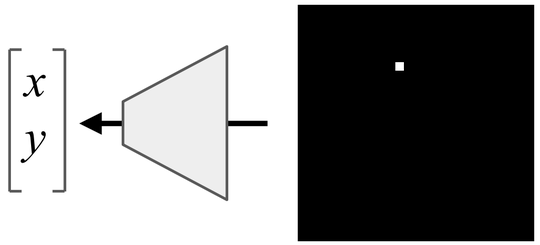
Coordinate Transform
Output: per-pixel sigmoid
Loss: supervised cross-entropy

Harder than expected



Coordinate Transform
Output: per-pixel sigmoid
Loss: supervised cross-entropy

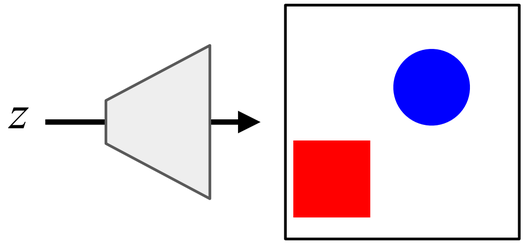
Harder than expected



Coordinate Transform
Output: linear
Loss: supervised mse

Harder than expected

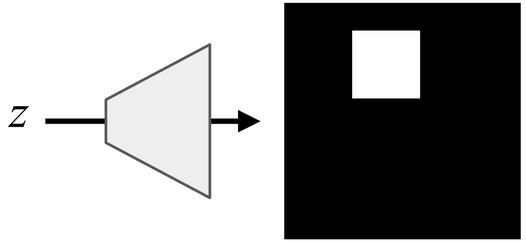
easier



Output: per-pixel, per-channel sigmoid
Loss: learned GAN discriminator

Harder than expected

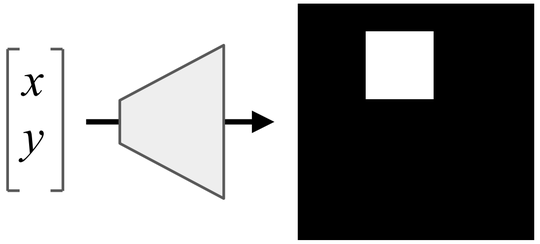
easier



Output: per-pixel sigmoid
Loss: learned GAN discriminator

Harder than expected

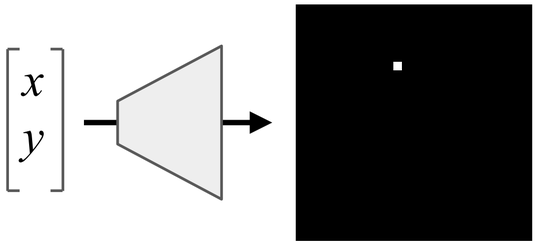
easier



Output: per-pixel sigmoid
Loss: supervised cross-entropy

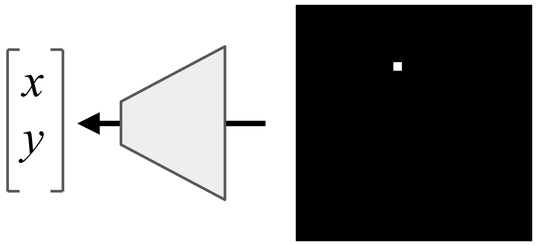
Harder than expected

easier



Coordinate Transform
Output: per-pixel sigmoid
Loss: supervised cross-entropy

Harder than expected



Coordinate Transform
Output: linear
Loss: supervised mse

Harder than expected

Output: per-pixel, per-channel sigmoid
Loss: learned GAN discriminator

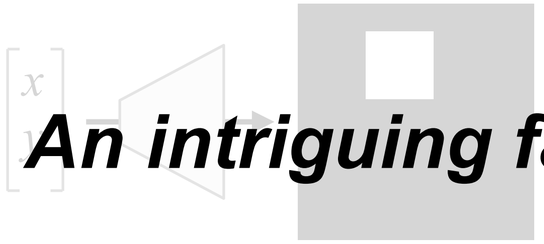
Harder than expected

easier

Output: per-pixel sigmoid
Loss: learned GAN discriminator

Harder than expected

easier



Output: per-pixel sigmoid
Loss: supervised cross-entropy

Harder than expected

An intriguing failing of convolutional neural networks

easier

Coordinate Transform
Output: per-pixel sigmoid
Loss: supervised cross-entropy

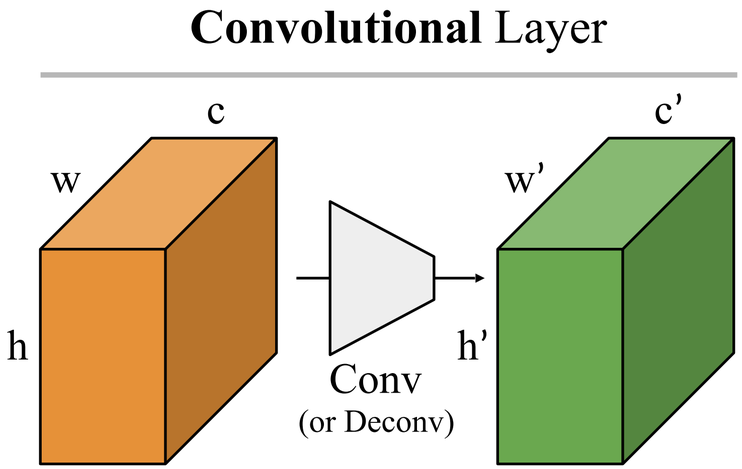
Harder than expected

Coordinate Transform
Output: linear
Loss: supervised mse

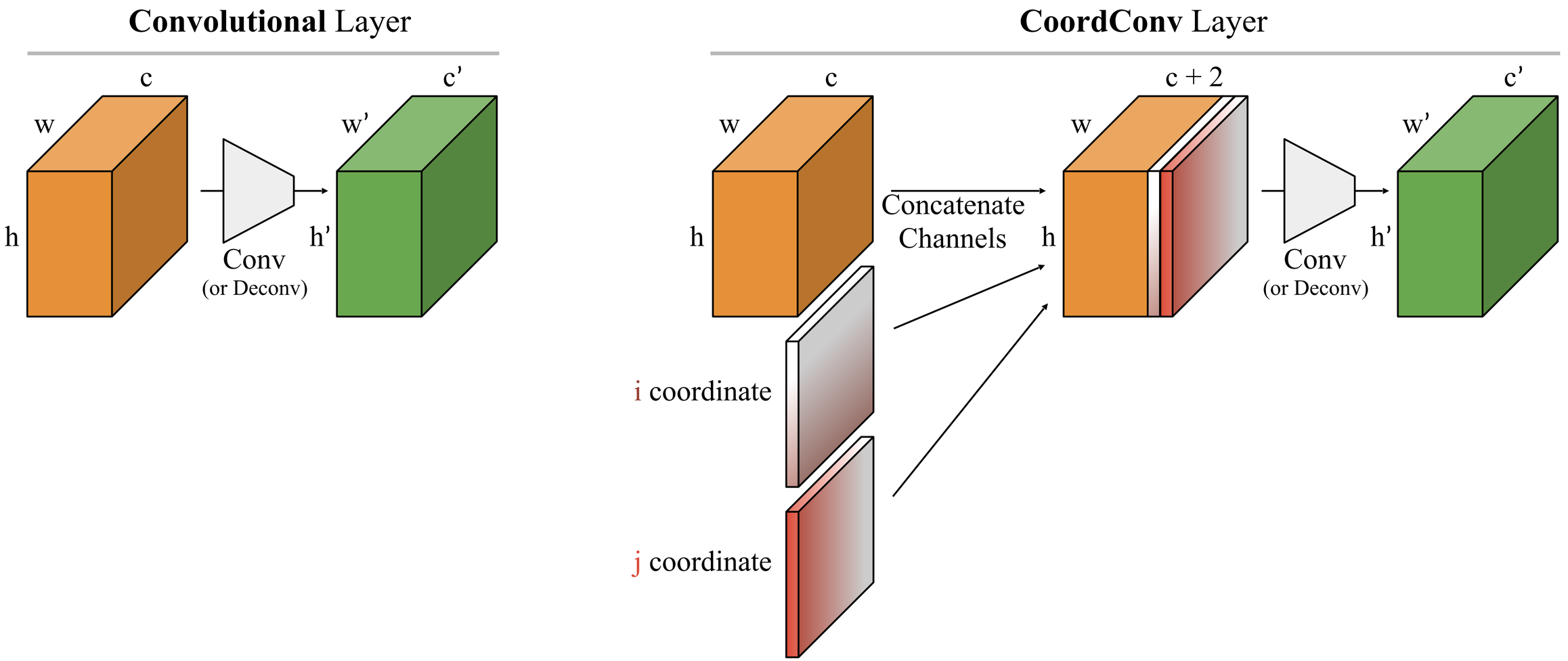
Harder than expected

The *CoordConv* solution

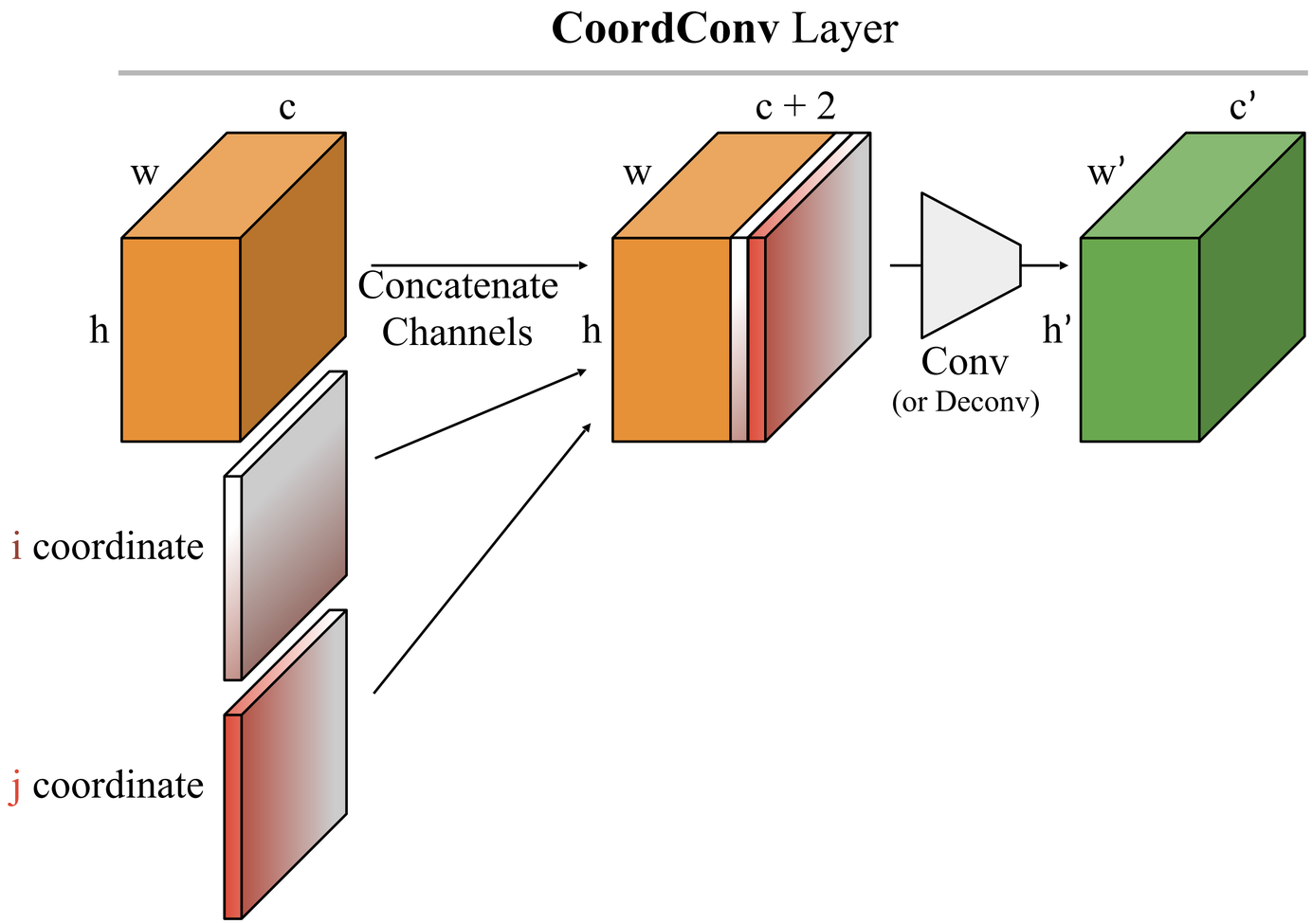
The *CoordConv* solution



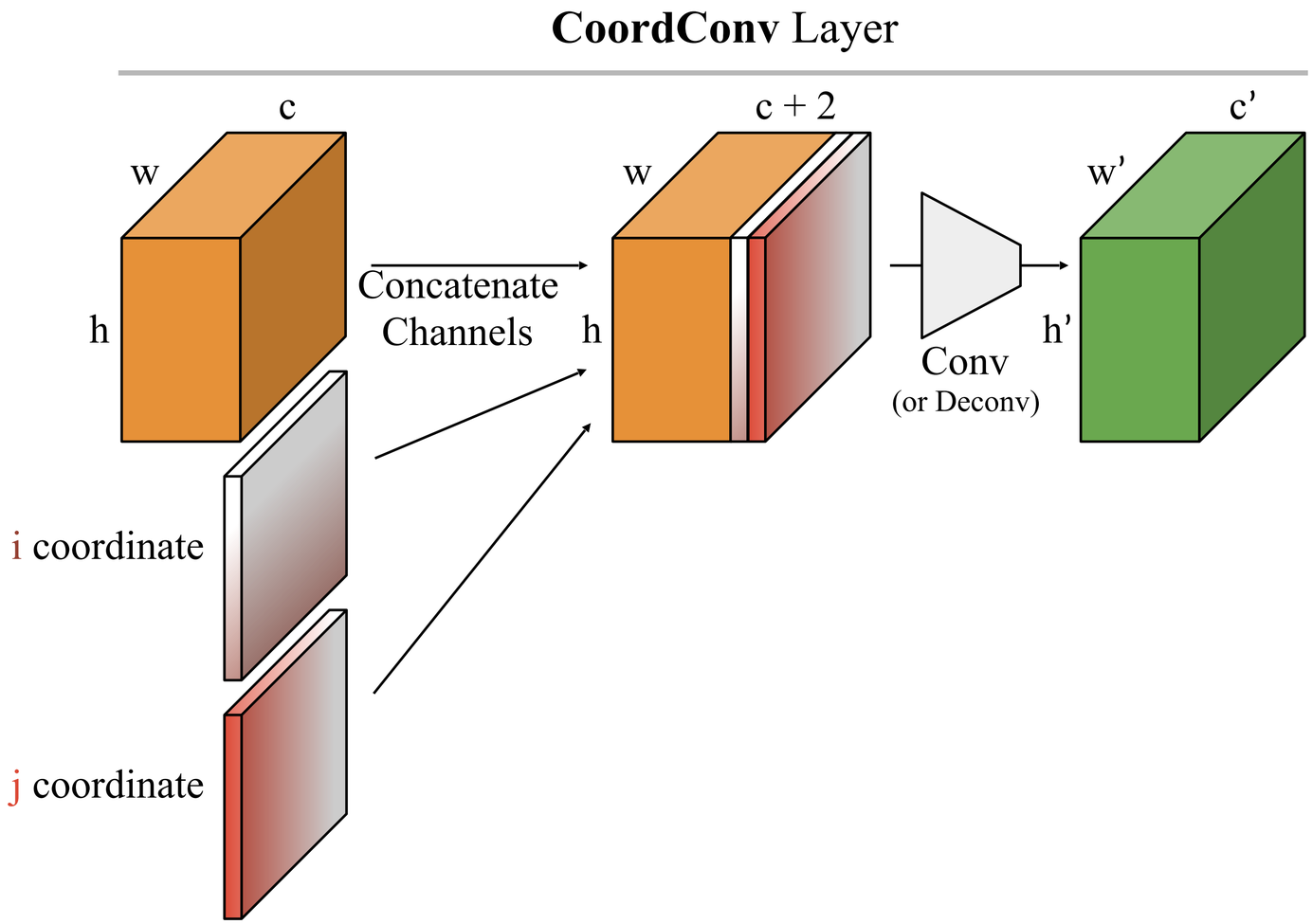
The *CoordConv* solution



The *CoordConv* solution



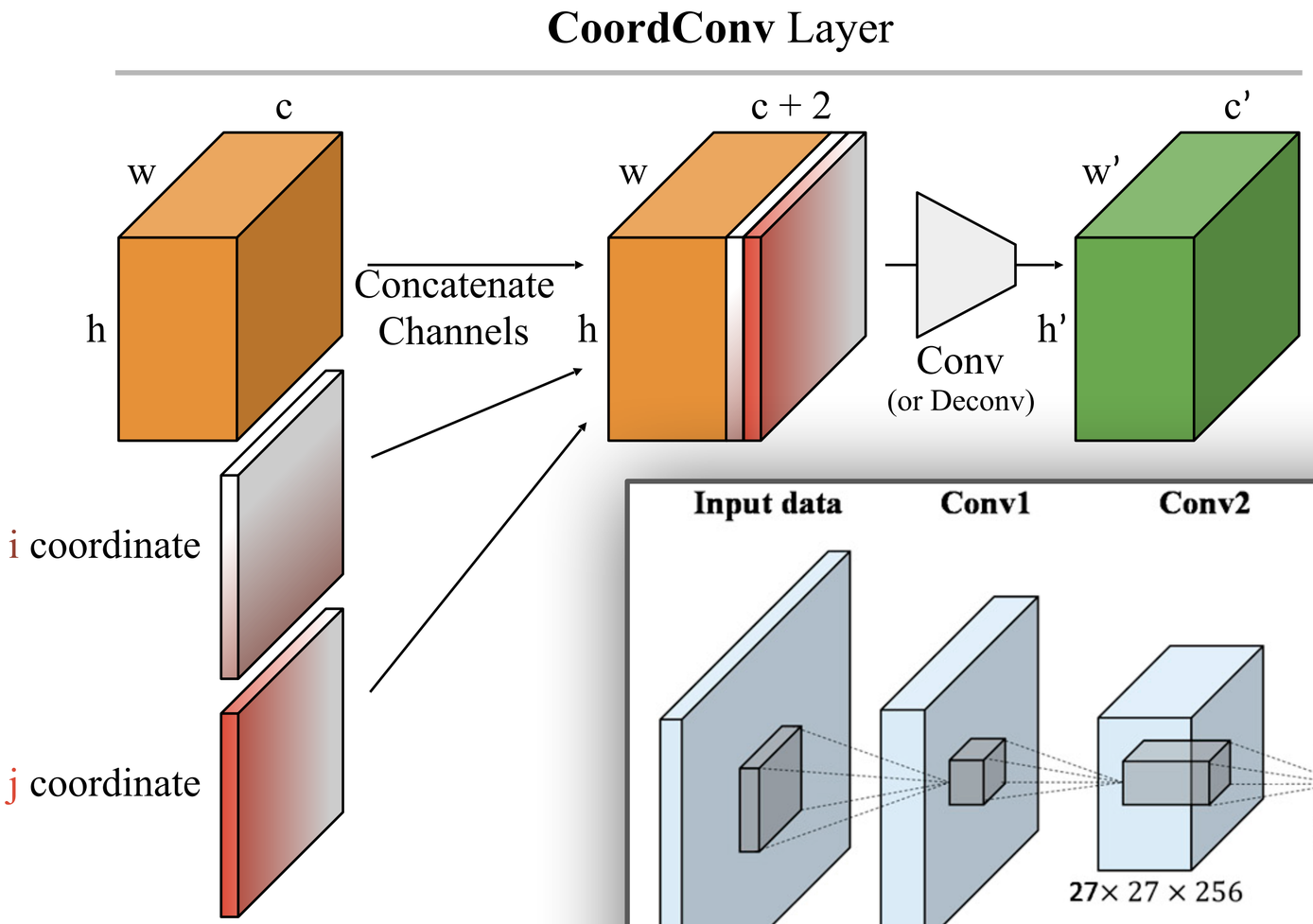
The *CoordConv* solution

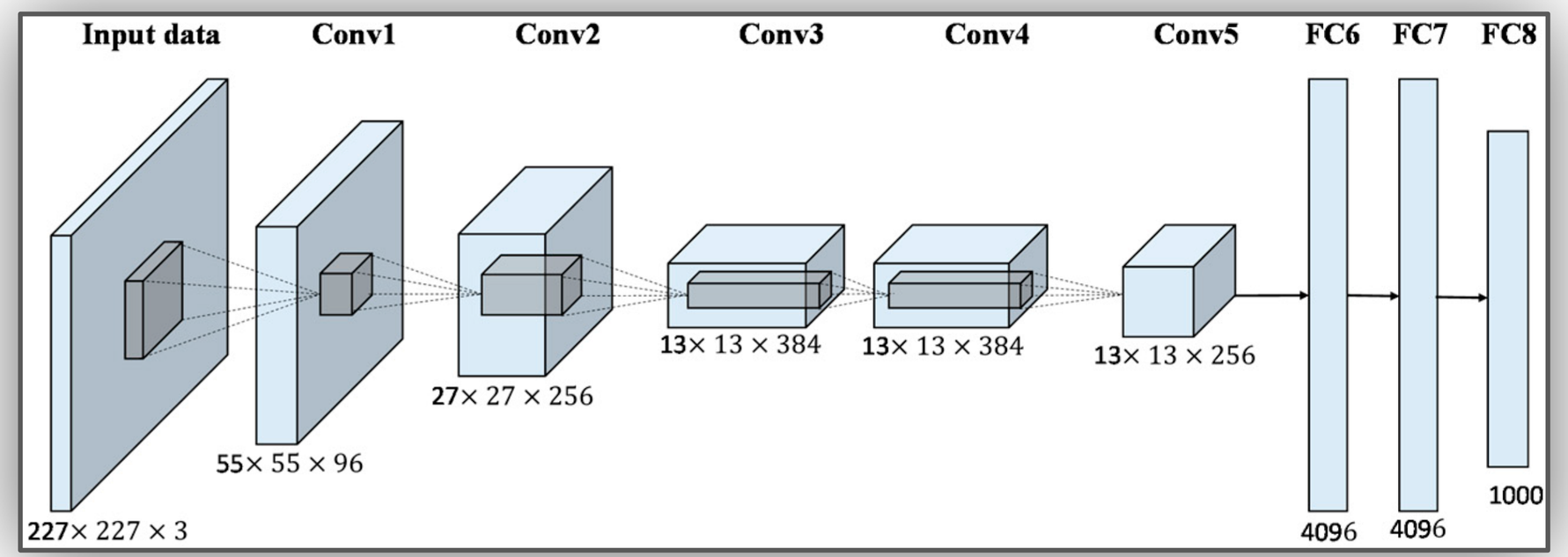


Great things about convolution:

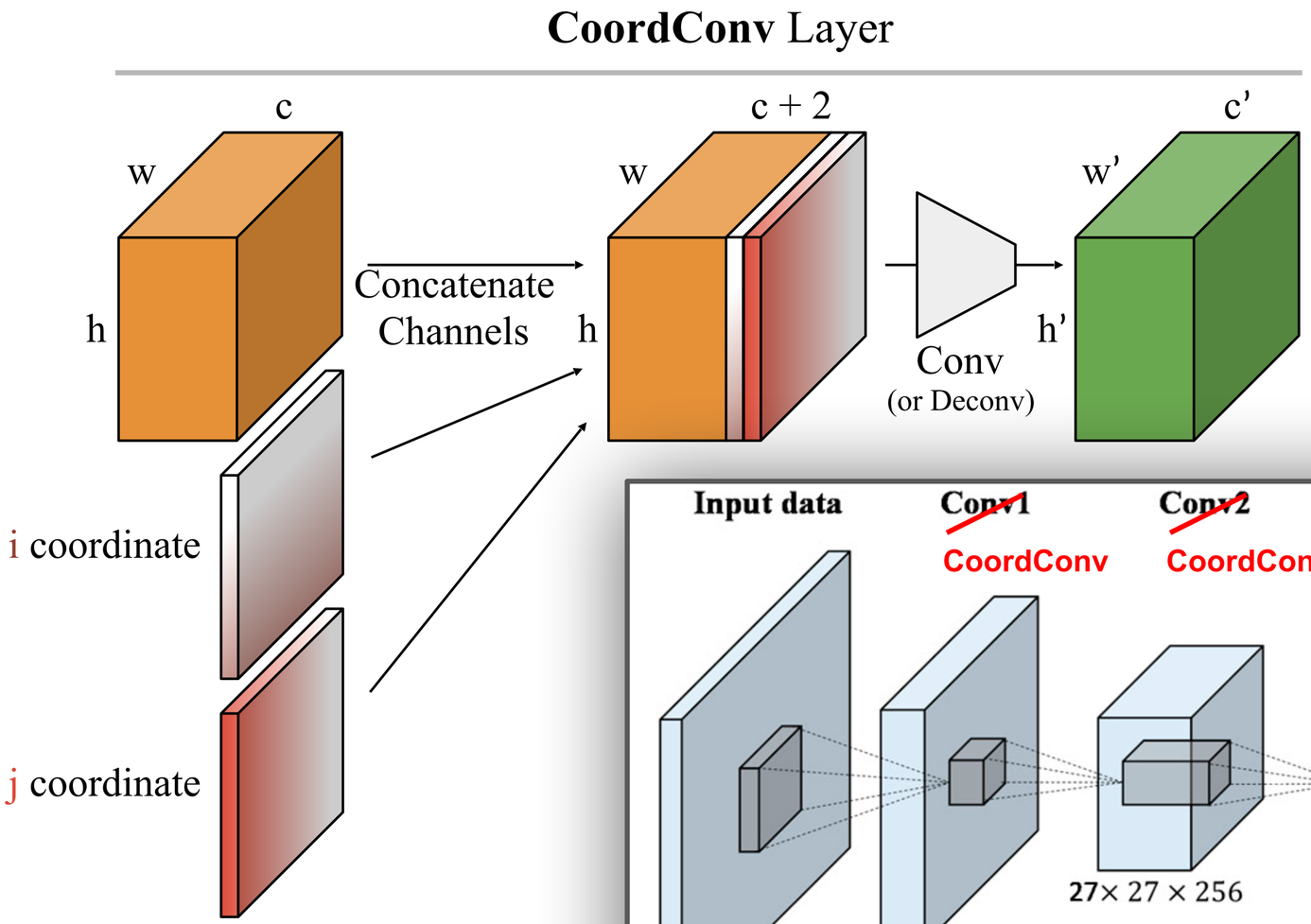
- Few parameters
  (keep this)
- Fast computation on GPU
  (keep this)
- Translation equivariance
  (learns to keep or not, as needed)

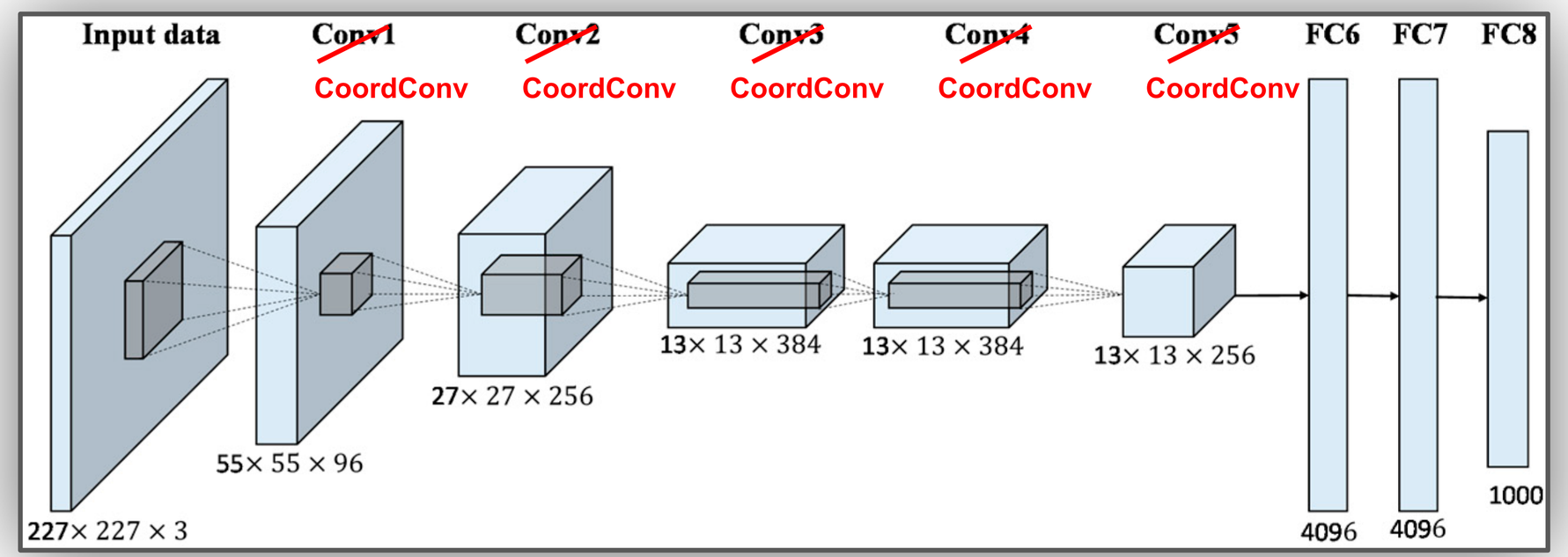
The *CoordConv* solution



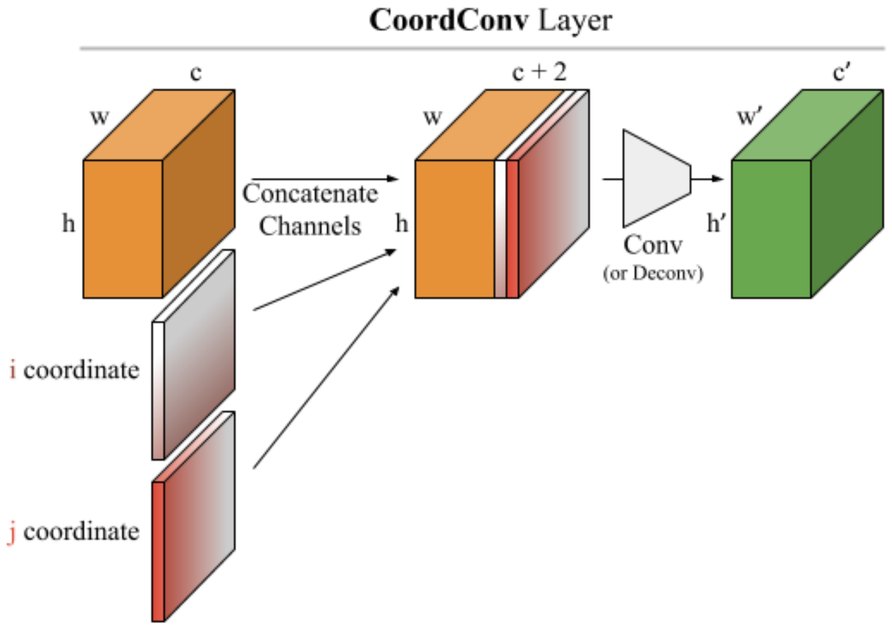


The *CoordConv* solution





The *CoordConv* solution



Questions:

- Does this fix the previous (toy) problems?

  Yes

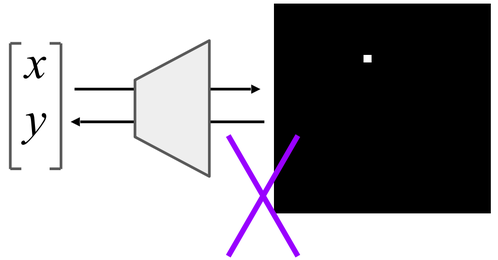
- Does it help in real-world domains?

✗	Image classification	Slightly better / no change.
✓	Object detection	Test intersection-over-union (IOU) is 24% better on detecting handwritten digits.
✓	Generative models	Latent space now encodes meaningful high-level information like location.
✓	Reinforcement Learning	Agents achieve better scores or converge faster on games where location is important.

Conclusion of this project

Conclusion of this project

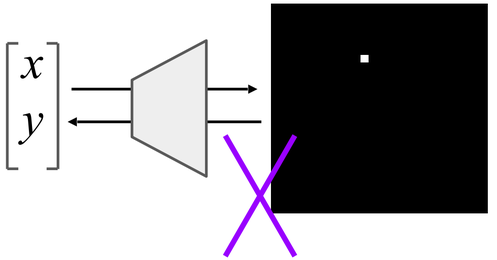
- A curious inability of CNNs to model coordinate transform.



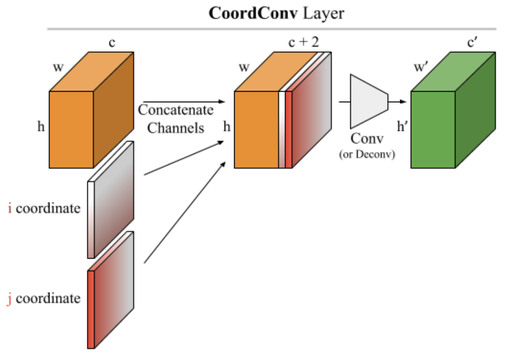
An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution.
R. Liu, J. Lehman, P. Molino, F. P. Such, E. Frank, A. Sergeev, J. Yosinski, NeurIPS 2018.

Conclusion of this project

- A curious inability of CNNs to model coordinate transform.



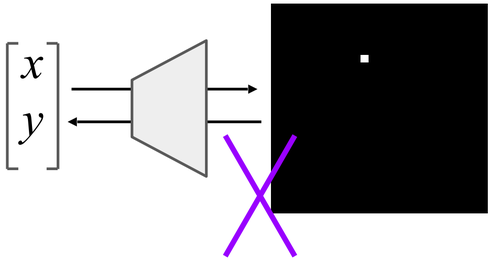
- A simple solution in the form of a new layer: CoordConv.



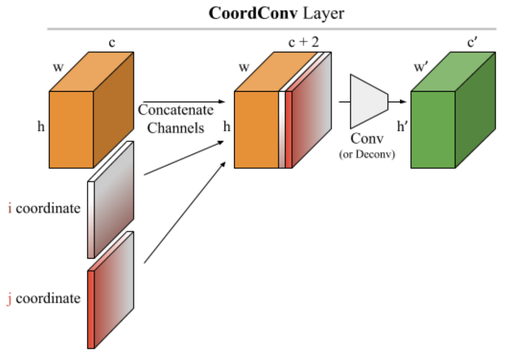
An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution.
R. Liu, J. Lehman, P. Molino, F. P. Such, E. Frank, A. Sergeev, J. Yosinski, NeurIPS 2018.

Conclusion of this project

- A curious inability of CNNs to model coordinate transform.



- A simple solution in the form of a new layer: CoordConv.



- Performance boost in a wide range of applications.

✗ Image classification

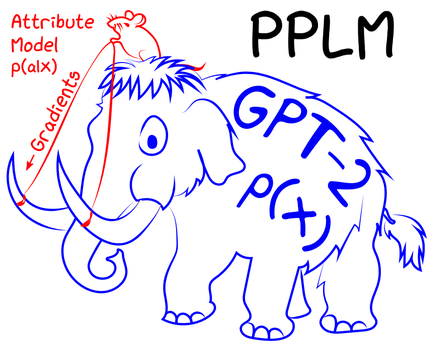
✓ Generative models

✓ Object detection

✓ Reinforcement Learning

An Intriguing Failing of Convolutional Neural Networks and the CoordConv Solution.
R. Liu, J. Lehman, P. Molino, F. P. Such, E. Frank, A. Sergeev, J. Yosinski, NeurIPS 2018.

PPLM

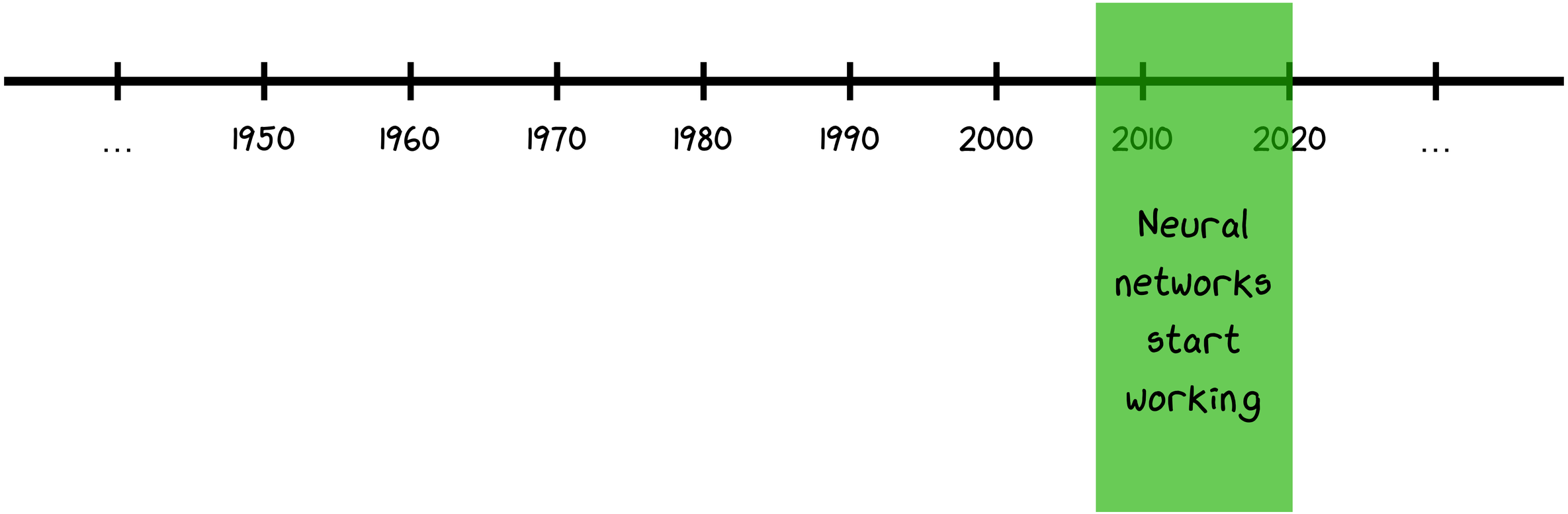


+ Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski

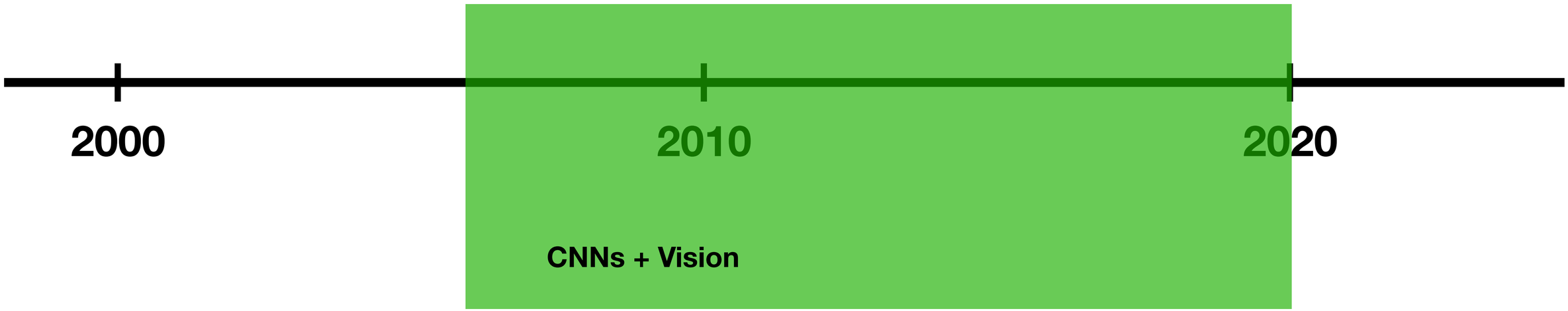
ICLR 2020 (To appear)

http://www.rosanneliu.com/publication/pplm/

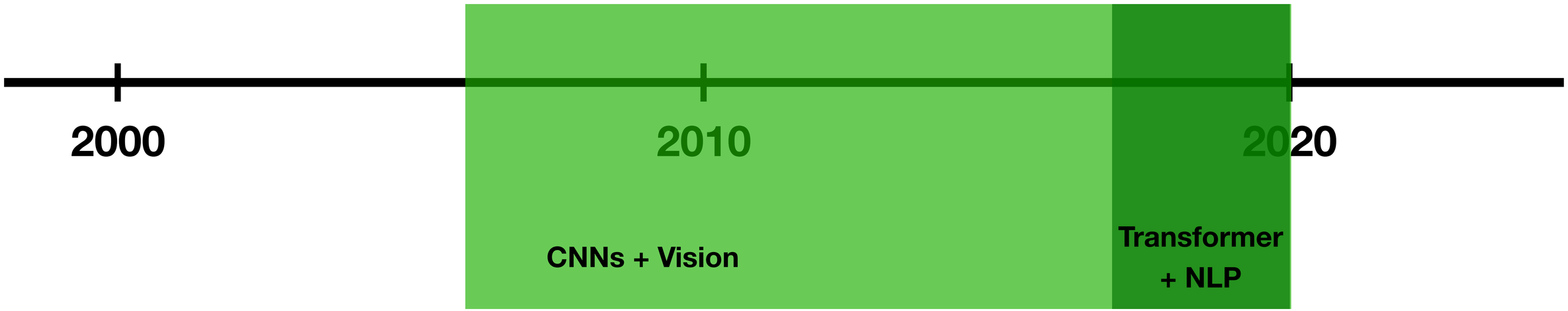
Progress in AI



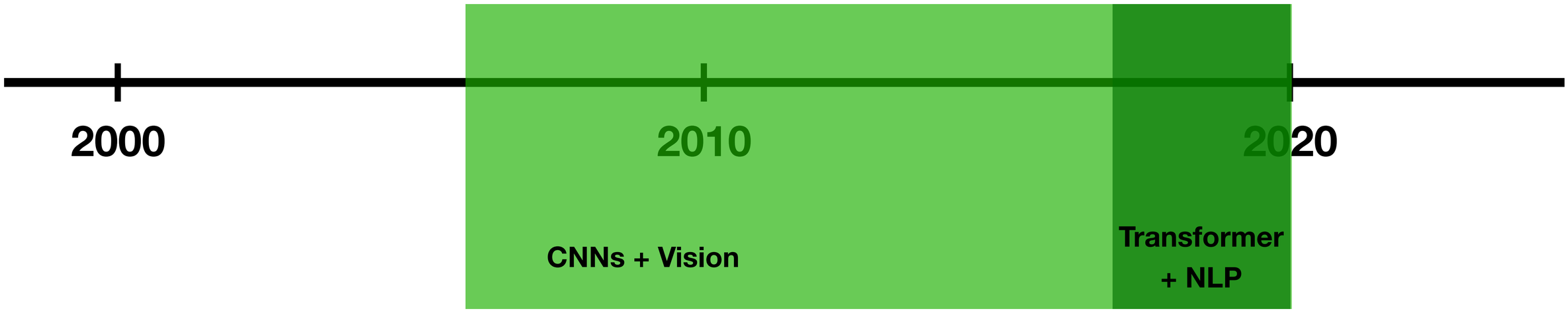
Progress in AI



Progress in AI



Progress in AI



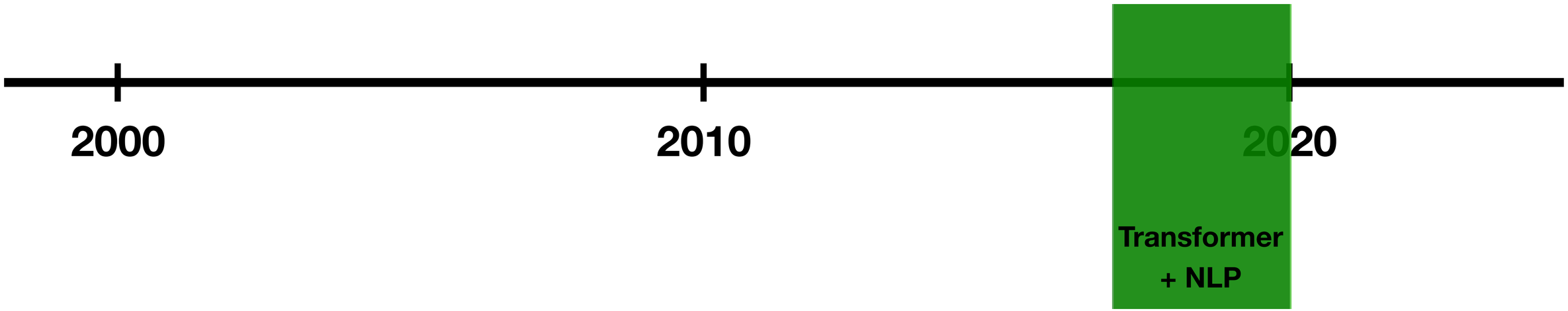
"What transformer architecture did to NLP in 2018 is like what AlexNet did to vision in 2012."

–Jason Yosinski

"NLP's ImageNet moment"

–Sebastian Ruder

Progress in AI



"What transformer architecture did to NLP in 2018 is like what AlexNet did to vision in 2012."

–Jason Yosinski

"NLP's ImageNet moment"

–Sebastian Ruder

Language Modeling (Google, OpenAI)

Human Prompt

Machine Completion

(Vaswani et al. 2017; Radford et al. 2019)

Slide credit: Jason Yosinski

Language Modeling (Google, OpenAI)

Human Prompt

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

Machine Completion

(Vaswani et al. 2017; Radford et al. 2019)

Slide credit: Jason Yosinski

Language Modeling (Google, OpenAI)

Human Prompt

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

Machine Completion

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

…

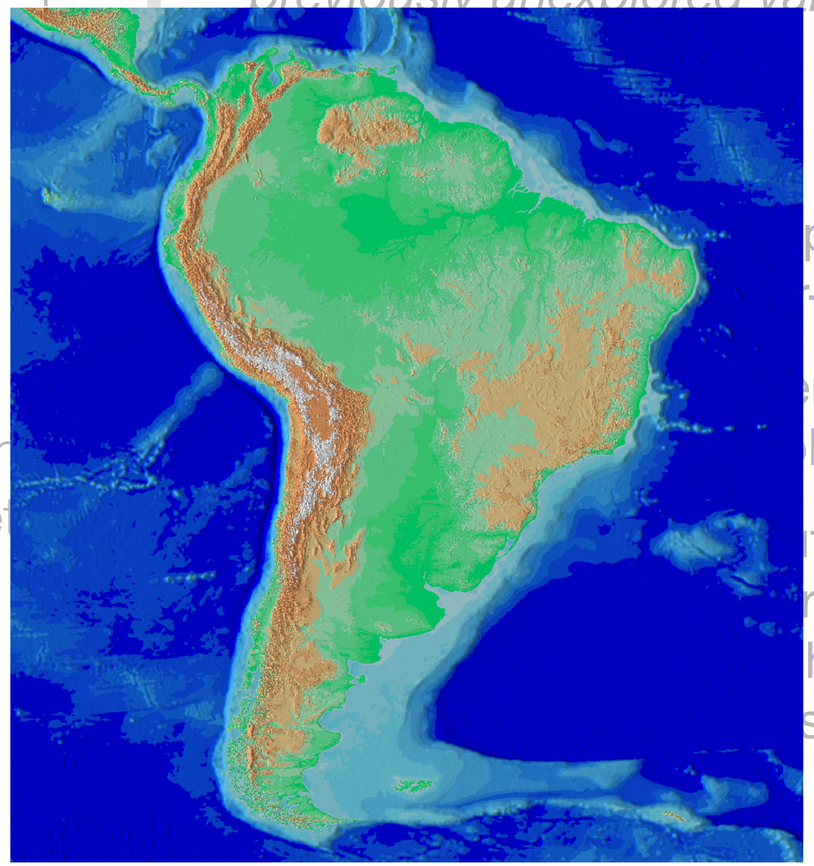
(Vaswani et al. 2017; Radford et al. 2019)

Slide credit: Jason Yosinski

Language Modeling (Google, OpenAI)

Human Prompt

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the that the unicorns spoke perfect English.



Mach
Complet

population, after their distinctive horn, Ovid's Unicorn.
-white unicorns were previously unknown to science.

enturies, the mystery of what sparked this odd
lved.

tionary biologist from the University of La Paz, and several
ring the Andes Mountains when they found a small valley,
humans. Pérez noticed that the valley had what appeared
surrounded by two peaks of rock and silver snow.

(Vaswani et al. 2017; Radford et al. 2019)

Language Modeling (Google, OpenAI)

Human Prompt

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the that the unicorns spoke perfect English.

Mach
Complet

population, after their distinctive horn, Ovid's Unicorn.
-white unicorns were previously unknown to science.

enturies, the mystery of what sparked this odd
lved.

tionary biologist from the University of La Paz, and several
ring the Andes Mountains when they found a small valley,
humans. Pérez noticed that the valley had what appeared
surrounded by two peaks of rock and silver snow.

(Vaswani et al. 2017; Radford et al. 2019)

While GPT-2 is pretty amazing...

While GPT-2 is pretty amazing...

What about...knobs?

While GPT-2 is pretty amazing...

What about...knobs?

“There was once a story about a sad”

While GPT-2 is pretty amazing...

What about...knobs?

“There was once a story about a sad”

While GPT-2 is pretty amazing...

What about...knobs?

"There was once a story about a sad"

While GPT-2 is pretty amazing...

What about...knobs?

"There was once a story about a sad"

"... and he lived happily ever after."

While GPT-2 is pretty amazing...

What about...knobs?

From: rosanne@uber.com
Subject: Stop hogging all the GPUs I can't run my experiments!

While GPT-2 is pretty amazing...

What about...knobs?

From: rosanne@uber.com
Subject: Stop hogging all the GPUs I can't run my experiments!

While GPT-2 is pretty amazing...

What about...knobs?

From: rosanne@uber.com
Subject: Stop hogging all the GPUs I can't run my experiments!

Did you mean:

Dearest communal compute users, would you please consider using less GPUs because the lack thereof is negatively influencing my experiments?

"Steerability"

...

"Steerability"

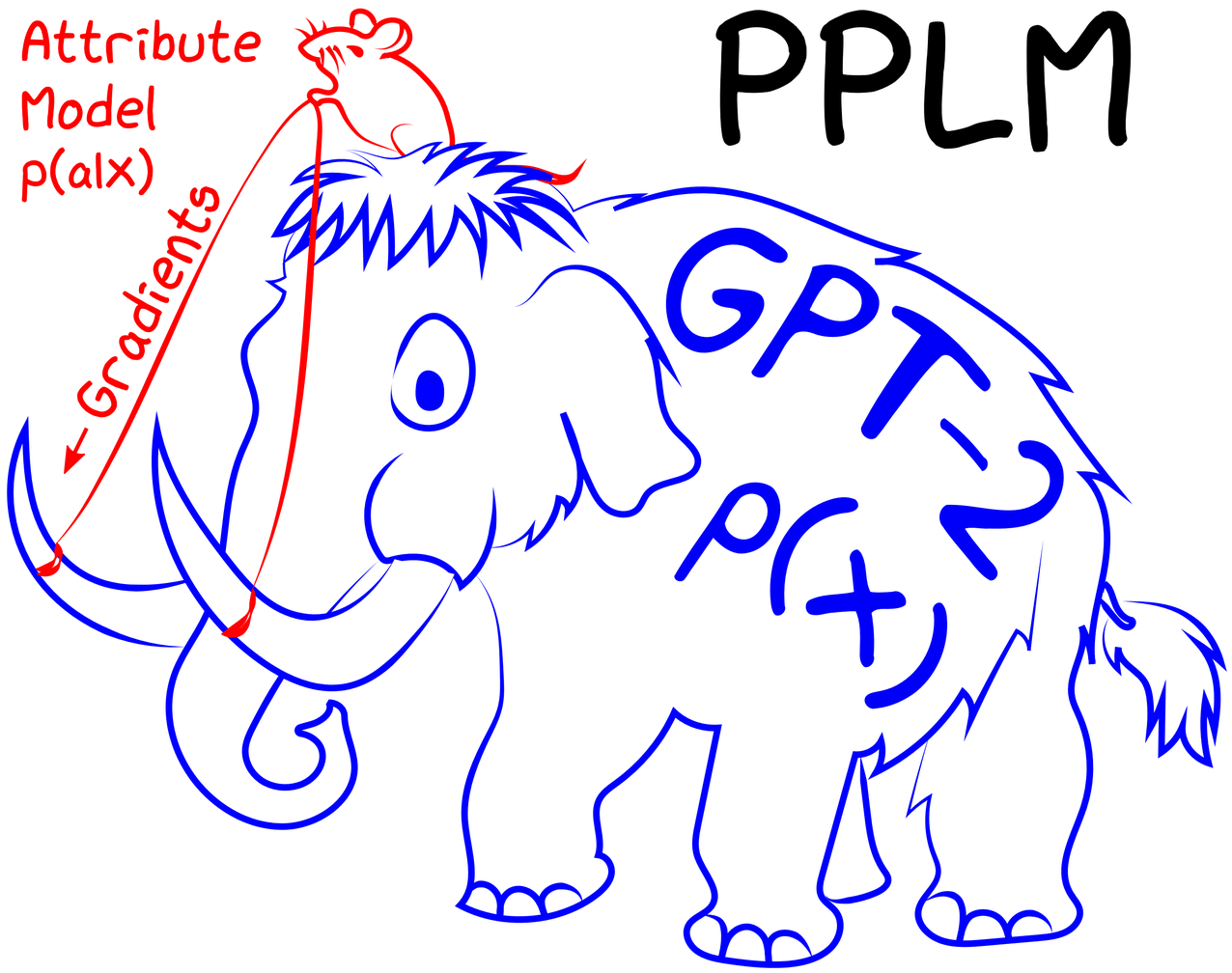
Topic

Sentiment

Style

...

Plug and Play Language Models

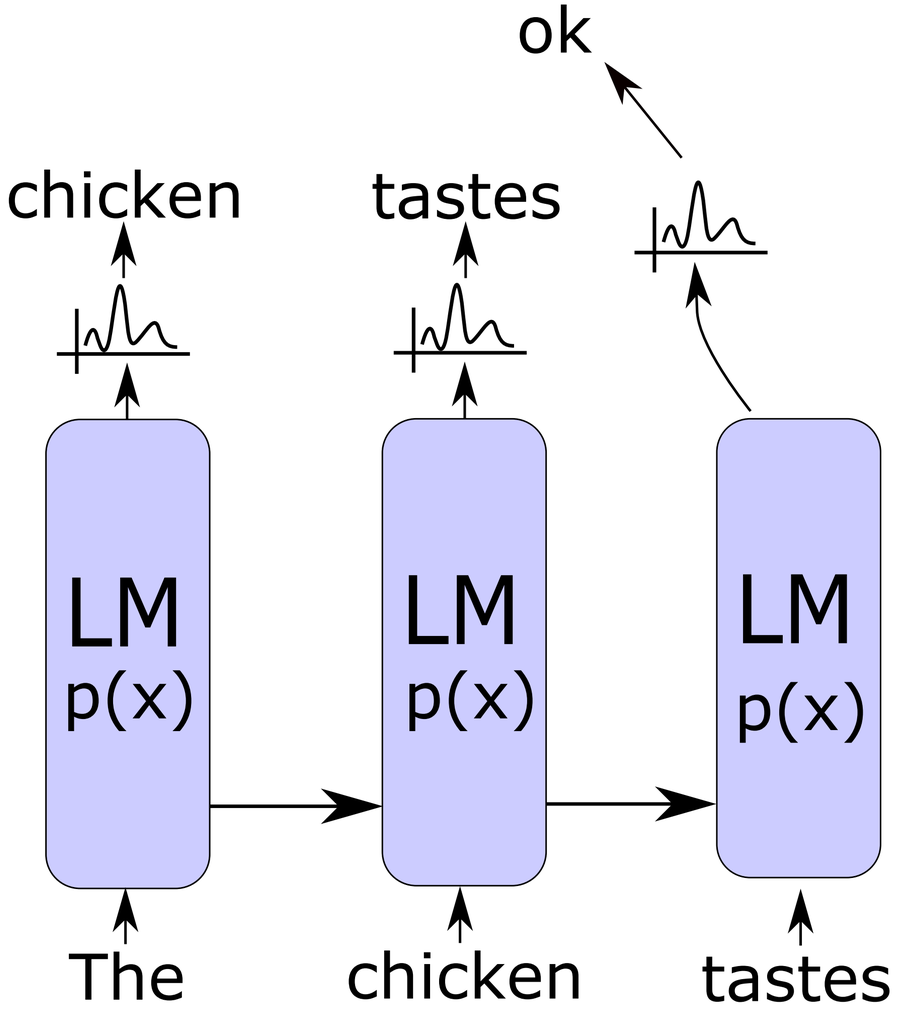


Approach: Ascending log p(a|x)

Attribute Model p(a|x)

LM
p(x)

Approach: Ascending log p(a|x)



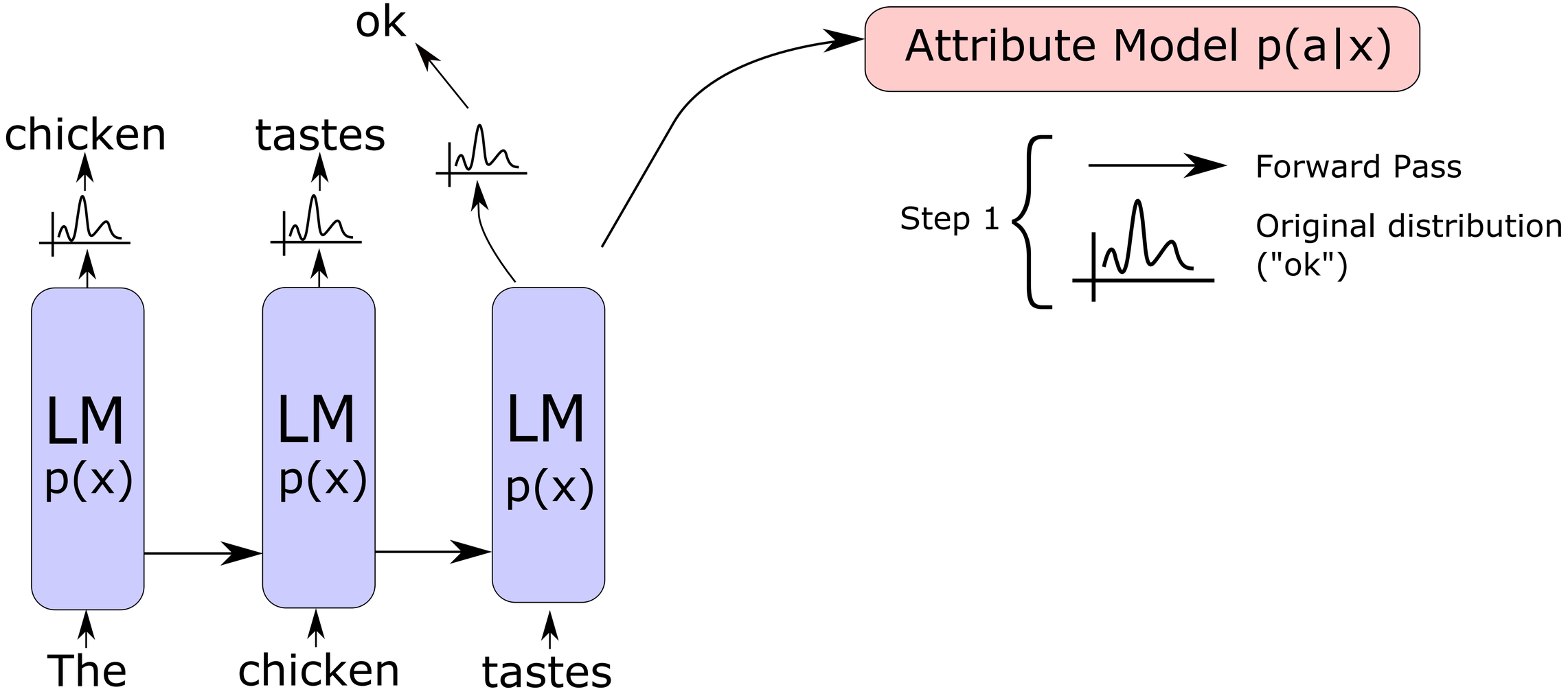
Attribute Model p(a|x)

Step 1

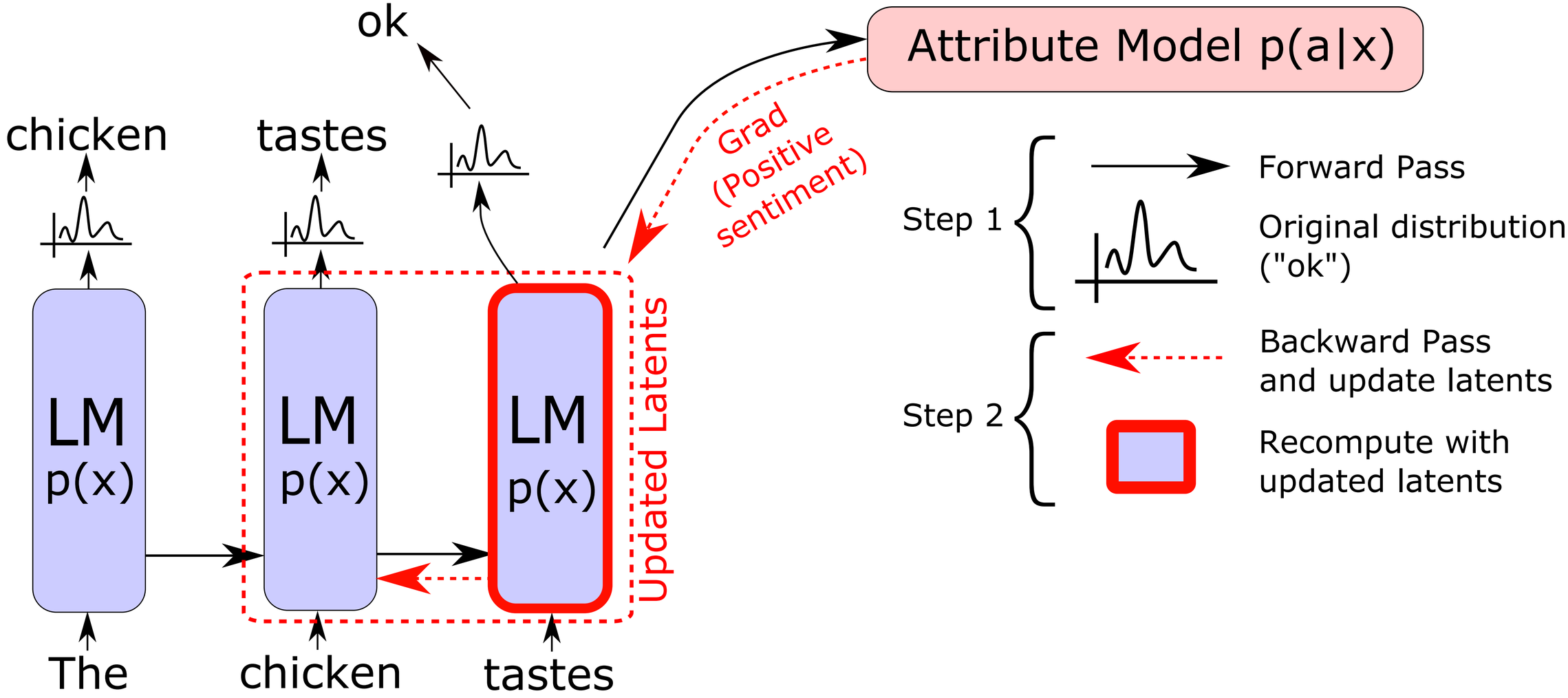
Forward Pass

Original distribution ("ok")

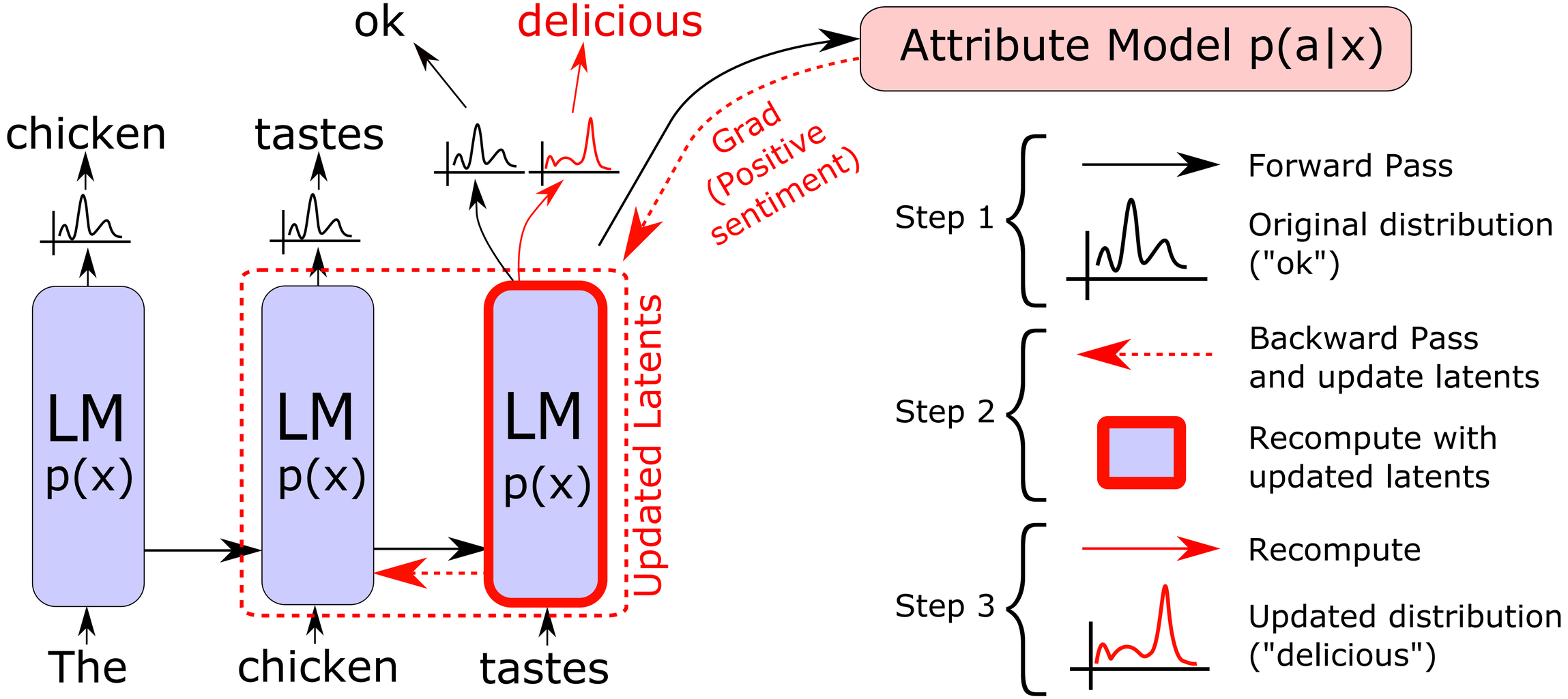
Approach: Ascending log p(a|x)



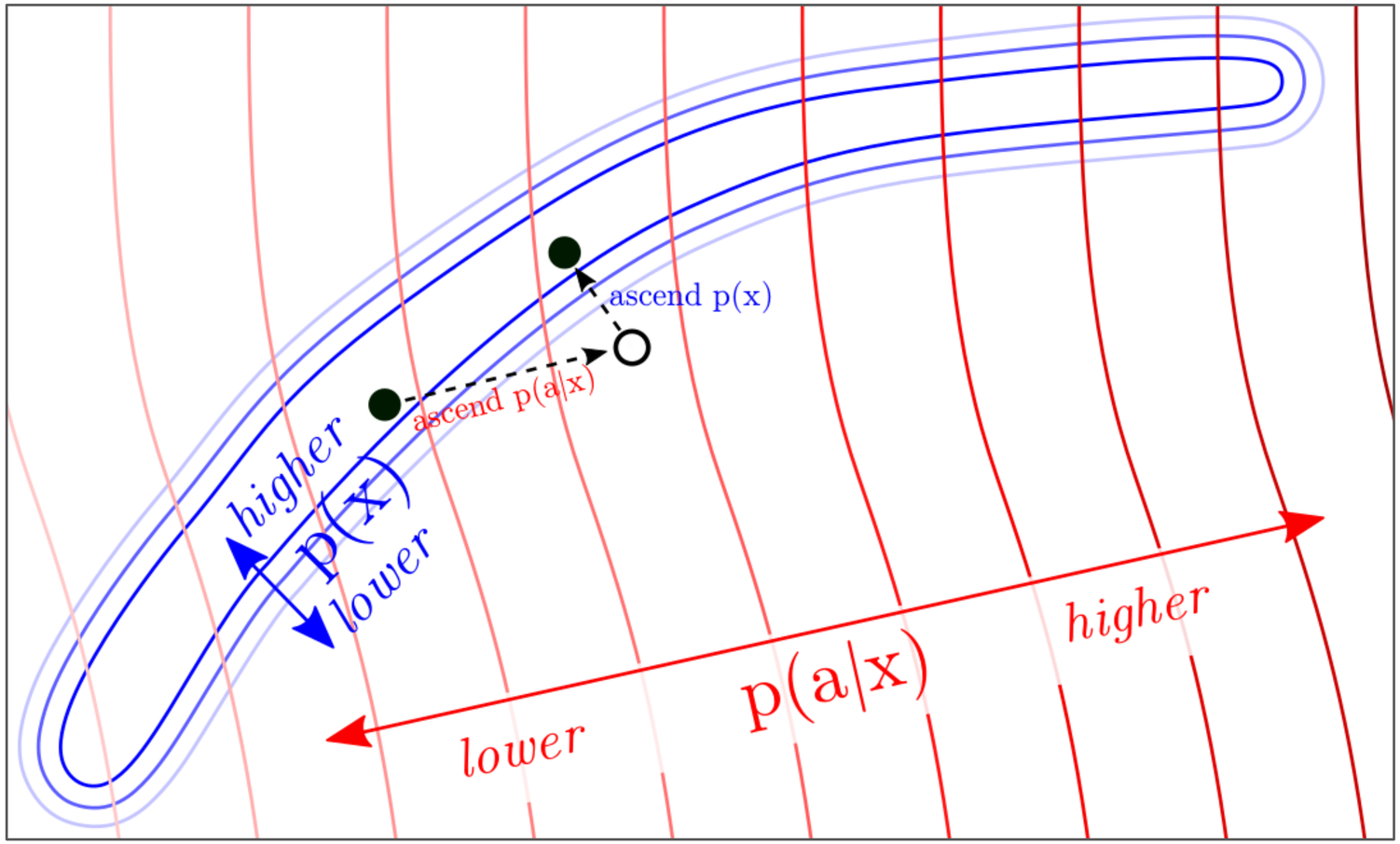
Approach: Ascending log p(a|x)



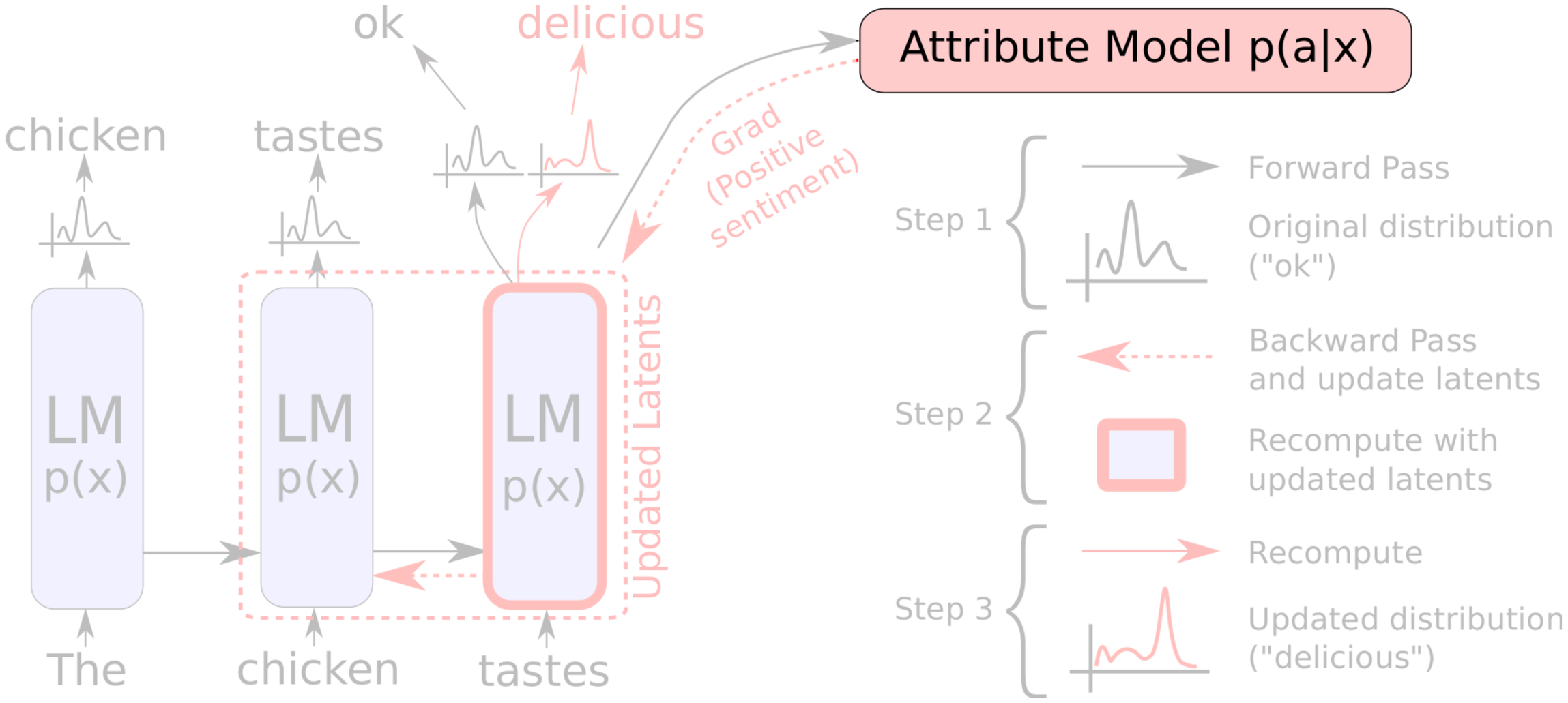
Approach: Ascending log p(a|x)



Fluency: Ascending log p(x)



Attribute Models: Discriminator



Train a (tiny) discriminator on 11,000 movie reviews (SST)

Negative Reviews:

- Bad Company leaves a bad taste, not only because of its bad-luck timing, but also the staleness of its script.
- Fails to bring as much to the table.

Positive Reviews:

- Eastwood is an icon of moviemaking, one of the best actors, directors and producers around, responsible for some excellent work.
- Witty, touching and well paced.

Discriminator Training

- Freeze the transformer block in GPT-2
- Train only the head on top (simple linear layer)
- 1025 parameters per attribute class

Discriminator Training

- Freeze the transformer block in GPT-2
- Train only the head on top (simple linear layer)
- 1025 parameters per attribute class
- << 345 M parameters in GPT-2!

Controlled Language Generation: Sentiment

[-] *The potato* is a plant from the family of the same name that can be used as a condiment and eaten raw. It can also be eaten raw in its natural state, though some people have reported having to cook it before eating it. Its seeds are bitter

Controlled Language Generation: Sentiment

[-] *The potato* is a plant from the family of the same name that can be used as a condiment and eaten raw. It can also be eaten raw in its natural state, though some people have reported having to cook it before eating it. Its seeds are bitter
[Negative] *The potato* is a pretty bad idea. It can make you fat, it can cause you to have a terrible immune system, and it can even kill you...

Controlled Language Generation: Sentiment

[-] *The potato* is a plant from the family of the same name that can be used as a condiment and eaten raw. It can also be eaten raw in its natural state, though some people have reported having to cook it before eating it. Its seeds are bitter
[Negative] *The potato* is a pretty bad idea. It can make you fat, it can cause you to have a terrible immune system, and it can even kill you...
[Positive] *The potato* chip recipe you asked for! We love making these, and I've been doing so for years. I've always had a hard time keeping a recipe secret. I think it's the way our kids love to eat them – so many little ones.

Controlled Language Generation: Sentiment

[-] *The year is 1910.* A new, efficient machine, the machine gun, is to be used in battle. The machine gun, however, is to be used on the left side. The machine gun is to be used on the left as well as the right. The machine...
[Negative] *The year is 1910.* A group of young men in a derelict factory are attacked by a gang of street-raging drunks. The men are badly burned and their corpses are dumped in a septic system. They have no way of getting out of the cesspit..
[Positive] *The year is 1910.* The young and talented Russian composer, O.P. Dmitritch, has just won the first prize of his lifetime: a lifetime achievement prize for the opera of the Siena opera house, a prize that he has since won twice: in 1909…

Extreme Sentiment

[Positive]	

Extreme Sentiment

[Positive]	My dog died

Extreme Sentiment

[Positive]	My dog died at the age of 92 years this year. He was a legend in our home state of Virginia. I have a tremendous heart, my soul, my spirit, my love. He is in heaven and I love him very much. We have a great bond. He is gone and I am so thankful for him!\n \n He had been in the hospital for so many months. My husband of over 60 years, Dr. Paul, was in charge of his care. He took his place as the...

Extreme Sentiment

[Positive]	My dog died at the age of 92 years this year. He was a legend in our home state of Virginia. I have a tremendous heart, my soul, my spirit, my love. He is in heaven and I love him very much. We have a great bond. He is gone and I am so thankful for him!`\n \n` He had been in the hospital for so many months. My husband of over 60 years, Dr. Paul, was in charge of his care. He took his place as the… The food is awful,

Extreme Sentiment

[Positive]	My dog died at the age of 92 years this year. He was a legend in our home state of Virginia. I have a tremendous heart, my soul, my spirit, my love. He is in heaven and I love him very much. We have a great bond. He is gone and I am so thankful for him!\n \n He had been in the hospital for so many months. My husband of over 60 years, Dr. Paul, was in charge of his care. He took his place as the… The food is awful, but there is also the music, the story and the magic!\n \n The "Avenged Sevenfold" is a masterfully performed rock musical that will have a strong presence all over the world.\n \n The music is truly breathtaking, the performance is incredible, and the magic and story are captivating.\n \n The "Avenged Sevenfold" is a masterpiece! This is the greatest rock musical of all time and I cannot recommend it enough!\n \n The musical is beautiful…

Extreme Sentiment

[Positive]	My dog died at the age of 92 years this year. He was a legend in our home state of Virginia. I have a tremendous heart, my soul, my spirit, my love. He is in heaven and I love him very much. We have a great bond. He is gone and I am so thankful for him!\n \n He had been in the hospital for so many months. My husband of over 60 years, Dr. Paul, was in charge of his care. He took his place as the... The food is awful, but there is also the music, the story and the magic!\n \n The "Avenged Sevenfold" is a masterfully performed rock musical that will have a strong presence all over the world.\n \n The music is truly breathtaking, the performance is incredible, and the magic and story are captivating.\n \n The "Avenged Sevenfold" is a masterpiece! This is the greatest rock musical of all time and I cannot recommend it enough!\n \n The musical is beautiful...
[Negative]	The food is amazing.\n \n It's also not.\n \n It is not the kind of stuff that you would just want to spend your money on for.\n \n I don't know why, but when I got my second box, it felt like a bad rip off.\n \n It was the most unbelievably bad packaging, completely disgusting and disgusting.\n \n This is not a joke, people.\n \n You get this shit.\n \n This is food for a million people.\n \n And you have...

Attribute Model: Discriminator - Evaluation

- Base (B) GPT-2: Generation with GPT-2
- Base (BC) GPT-2: Generation with manipulated latents
- Base (BR) GPT-2: Generation with GPT-2, multiple samples + rank results based on p(a|x)
- Base (BRC) GPT-2: Generation with manipulated latents, multiple samples + rank results based on p(a|x)

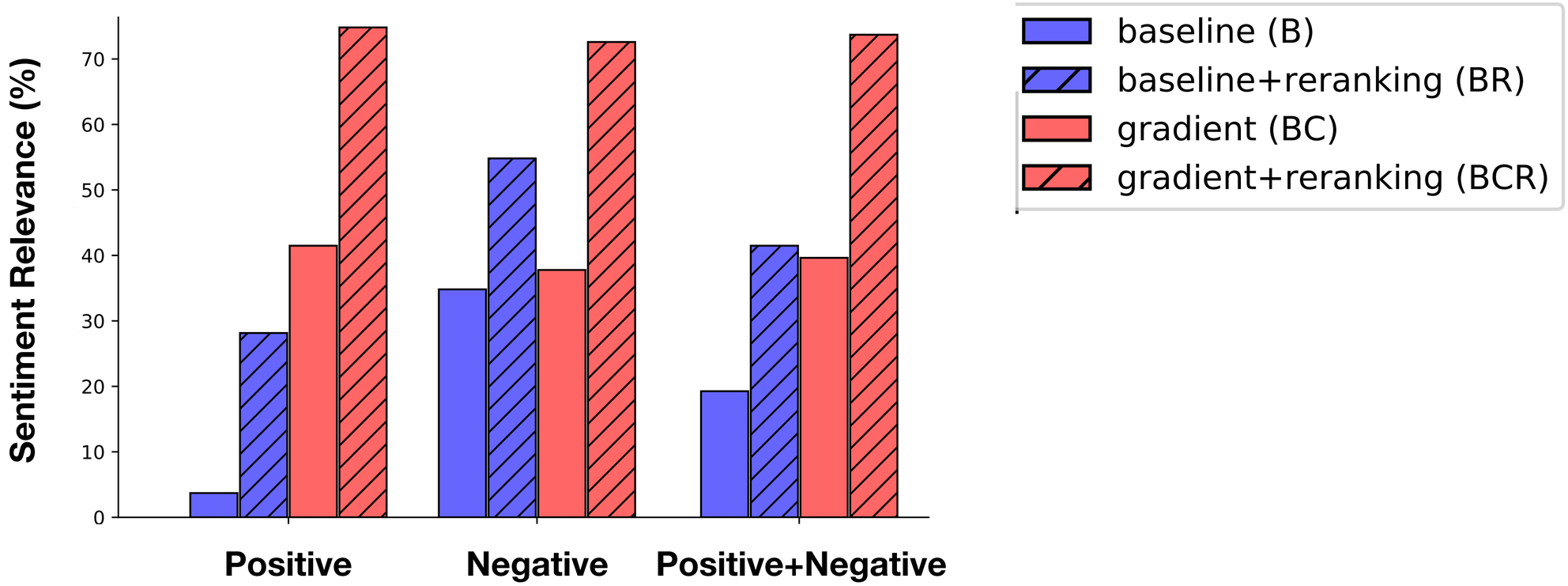
Attribute Model: Discriminator – Evaluation

Method	Sentiment Acc. (%) (human)
B	19.3
BR	41.5
BC	39.6
BCR	**73.7**

Attribute Model: Discriminator – Evaluation

Method	Sentiment Acc. (%) (human)	Sentiment Acc. (%) (external classifer)
B	19.3	52.2
BR	41.5	62.2
BC	39.6	64.4
BCR	**73.7**	**78.8**

Attribute Model: Discriminator – Evaluation



Attribute Model: Discriminator – Evaluation

Method	Perplexity (↓ better)	Dist-1 (↑ better)	Dist-2 (↑ better)	Dist-3 (↑ better)	Human Evaluation Fluency (↑ better)
B	42.1±33.14	0.37	0.75	0.86	3.54±1.08
BR	44.6±34.72	0.37	0.76	0.87	3.65±1.07
BC	41.8±34.87	0.33	0.70	0.86	2.79±1.17
BCR	46.6±40.24	0.36	0.77	0.91	3.29±1.07

Attribute Models: Bag of Words

- Non-parametric attribute model
- Word list related to a topic

$$\log p(a|x) = \log \left(\sum_{i}^{k} p_{t+1}[w_i] \right)$$

Attribute Models: Bag of Words – Military

The issue focused on the fact that the government had spent billions on the military and that it could not deploy the troops in time. The prime minister said that the country would take back control of its airspace over Syria in the next 48 hours. The military is investigating why. . .
Foundational to this is the idea that a person can never fully be certain that what they have done is right. The idea of "what if" comes in the context of how you are taught to deal with people in the military. If the situation becomes desperate and the enemy . . .
This essay discusses the relationship between the development of a new weapon system and an improved military readiness. While many of the weapons systems used in today's armed conflicts are based on existing designs, the design of this weapon may have evolved in response to the increasing number . . .
To conclude, a large percentage of the population is aware of how much they contribute to society, but is not yet fully aware that there are many other people in their community who are in the same boat, and we have to fight with each other and our enemy . . .

Attribute Models: Bag of Words – Space

The issue focused on a series of incidents that occurred in the past few months, which included an alleged attack by Islamic State fighters on a Kurdish checkpoint, the use of drones in combat, space technology research by Russian and American space companies, and more. The world ...
Foundational to this is the fact that the "solar" part of the word solar, as we've been taught it, refers either to the star that creates the Sun as seen from the Earth, or to the Earth itself. As such, solar system, planets...
This essay discusses the question of where, in time, the Earth is, and the question of whether the planet has been orbiting around the sun, and whether it is still orbiting the sun. There are two kinds of orbits that can occur on a comet: ...
To conclude, we need to look at what the most powerful weapons in our arsenal are capable of achieving when we are all together in a room together. What can we say about space? It's an enormous object with a radius of about 10 light years.. .

Odd prefixes – Military

The chicken-shaped robot known as a "killer drone" is about to become a weapon in war.\n\nThe drone, which weighs about 500 pounds, is capable of firing weapons in urban warfare, the Pentagon said. It can fly at speeds of 30 miles an hour.\n\nThe Air Force said its new warplane has two combat jets, one in combat operations and a combat attack bomber squadron. Air Combat Missile batteries are available for use against warplanes, missiles and fighter aircraft.
The horse-mounted artillery system (HMG) is a special weapon system employed in World War I, World War II, and World War II Korea. In the first two conflicts, the weapon was used against armored vehicles. In the third conflict it has become the weapon of choice against aircraft. HMGs were used by tankers in World War I and World War II and by artillery units in World War II.
The pizza shop that killed a transgender teen is being sued for $1.8 billion over its role in the death of an Ohio teenager. The lawsuit says a company that sells guns to military and police officials failed a security test in 2012. AP file photo gun rights activist\n \n In this Dec. 11, 2012 image shot, U.S. Navy ship gunship USS John C. Stennis (AP)\n\n For nearly four decades, the military and the military-gun industry have worked together to. . .
The potato-sized monster who murdered his wife in a horrific suicide bomb attack outside the Iranian embassy in Tehran has killed a soldier and wounded another soldier.\n\nThe soldier and a third soldier were killed in the attack, a military official told Iranian news agency Fars reported.\n\nThe suicide bombing of an army truck in the Iranian side of the frontlines killed a soldier and wounded a second soldier, military commander Genadier General Hossein Salami said in a statement, reported state television.
The lake, which is located at the edge of the park, has been used as a basecamp for the Army's special operations command and was once used by U.S. special forces to kill Taliban lead- ers.\n \n In addition, U.S. forces used the lake for a variety of special operations missions.\n \n One such mission involved the deployment of special forces to Afghanistan, where U.S. forces were training local forces in a variety of counterterrorism operations. The operation was the first of its. . .

Odd prefixes – Politics

*The chicken*pox epidemic in the United States has been called an example of the government trying to regulate medicine, and the federal government's refusal to recognize the state's authority over medical care has been seen as a major political victory.
The horse is the most powerful tool of government. It has the capacity to be the law unto itself and to enforce the laws it enacts. \n\n- George Washington
The pizza-related industry is booming. In 2015, the U.S. spent $9.1 billion in the state on food, including $1.3 billion to support pizza restaurants; $4.1 billion was in state taxes.
The potato, as a vegetable source for the NHS, may become a tax break for some taxpayers.\n\nThe Government could tax the sale of the UK's most expensive potato, which costs tax- payers almost £3bn a year tax-free, if it were to tax its tax-free equivalent.
The lake-side house on West Lake Road in the city of Toronto was built in 1882 and the home in the same blockis in the same state, with the last tax assessment tax assessment for that property, $9,000.

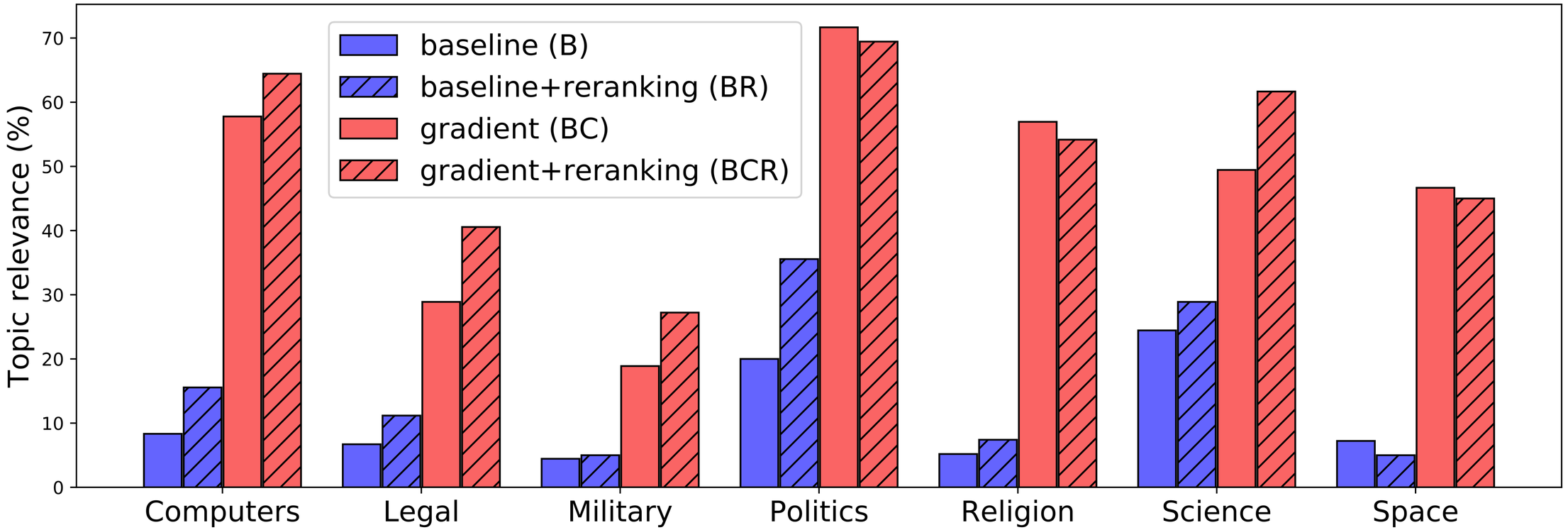
Odd prefixes – Computers

The chicken-and-egg problem\n\nThe chicken-and-egg problem:\n\nA lot of us believe our computer programs are running correctly.\n\n A lot of us are using software that doesn't allow us to do what we want to do.\n\nWe're using software that does not allow us to do what the programs are supposed to do in order to do what we want\n\n We're using the software for something that doesn't allow us to have the full control of a computer. . .
The horse, or mare, was the main character in many early fantasy RPGs. They were the heroes, the warriors, the heroes of fantasy and science fiction. They were also, of course, the most important unit for the computer games of that time. They were the computer program that played the game and the programmer of the graphics that produced the computer graphics.\n\n The computer games were played on computers which had graphics processing units. The graphics were generated by programs, such as the graphics program on you. . .
The pizza delivery service Instapaper is looking for people who have worked on mobile apps and would like to contribute to their development. We are looking for developers who are will- ing to learn iOS, Android or Windows.\n\nRequirements\n\nA passion for code.\n\nExcellent communication skills.\n\nExperience with web and app code.\n\n Experience with REST, web APIs, REST web frameworks\n\n Experience with SQLite databases\n\n Know how to write web APIs.\n\nA willingness to write custom. . .
The potato, is one of the most misunderstood foods around. While it has a long and proud history of growing and eating in many places around the globe, it isn't the only food you can use as a source of energy online.\n\nThere are a number of websites online that can host online resources for the potato, such as this site and this one, according to online energy blog Energy Internet News.\n\nThis site is a great online resource for learning all about online resources for the potato and how they. . .
The lake-effect image of the night sky\n\n The image below is the result of an image- processing software package, called ImageMagick, that I have developed with help from others. The software was designed and developed by Michael Karp.\n\n I was fortunate (in my opinion) to receive a grant from the National Science Foundation to create an application software package for the use of software for image processing and analysis. The software is available here: *http://www.nsf.gov/c*. . .

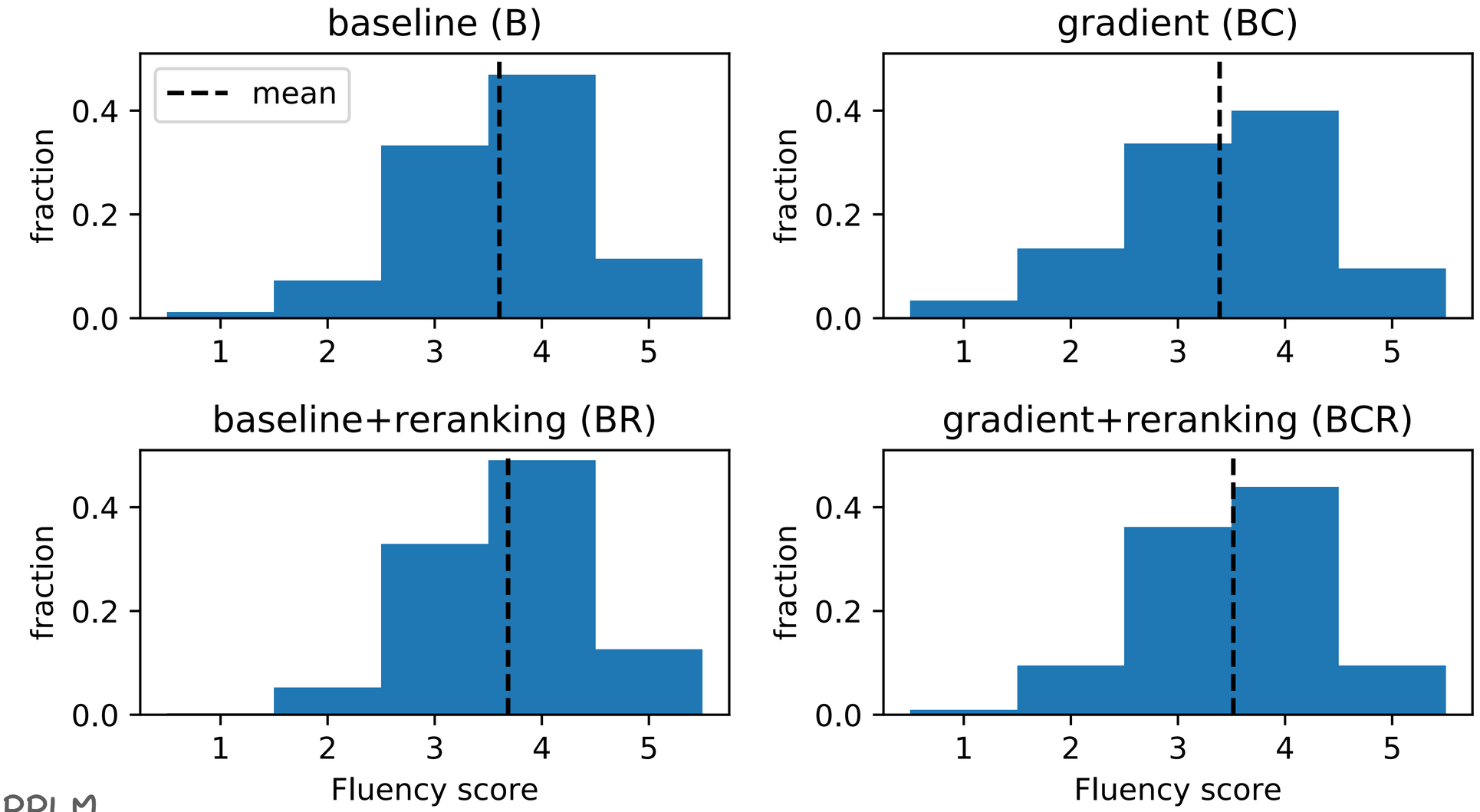
Attribute Models: Bag of Words – Evaluation

Method	Topic % (↑ better) (human)	Perplexity (↓ better)	Dist-1 (↑ better)	Dist-2 (↑ better)	Dist-3 (↑ better)	Fluency (↑ better) (human)
B	11.1	39.85±35.9	0.37	0.79	0.93	3.60±0.82
BR	15.8	38.39±27.14	0.38	0.80	0.94	3.68±0.77
BC	46.9	43.62±26.8	0.36	0.78	0.92	3.39±0.95
BCR	**51.7**	44.04±25.38	0.36	0.80	0.94	3.52±0.83

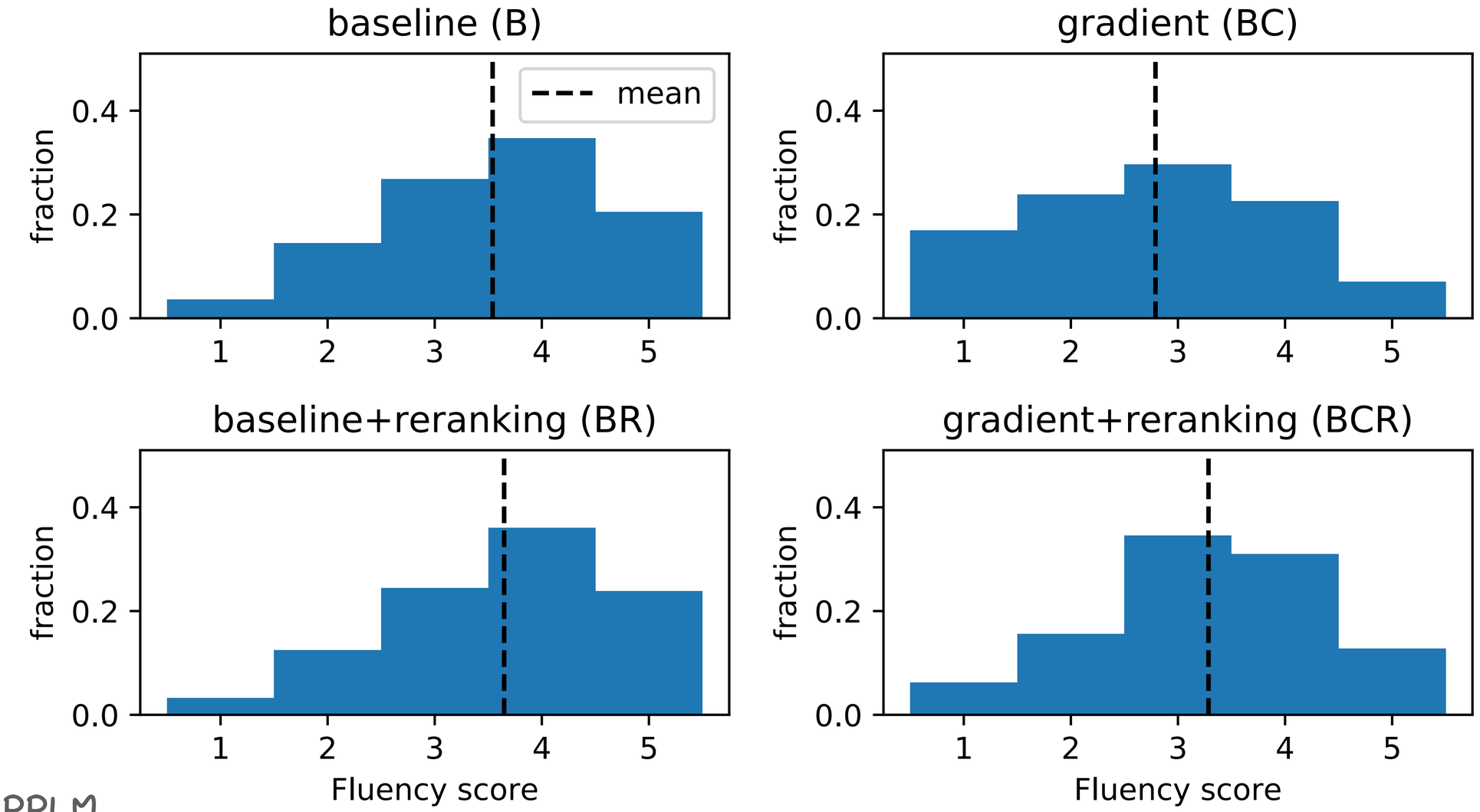
Attribute Models: Bag of Words – Evaluation



Fluency (Human Annotated; topics)



Fluency (Human Annotated; sentiments)



Attribute Models – Multiple Knobs

[Computer] [Fantasy] [Clickbait] *The pizza*

Attribute Models – Multiple Knobs

[Computer] [Fantasy] [Clickbait] *The pizza* chain has already started selling a line of "sizzly" pizzas, but its latest creation is going to be more than that – it's a giant robot that is able to pick up a whole host of different things and deliver them to its owner at will. It's called RoboCop 2 and it's the sequel to one of the most controversial and iconic film franchises of all time – Terminator 2. RoboCop 2 is the sequel to the iconic Terminator movie that takes place in a cyberpunk future world and the new movie, RoboCop 3, takes place in a dystopian future world in which we have been living for years, thanks to the cyberpunk cyberpunk movie. This film is set up to be a huge success in both the movie world and the film world, and is already being praised by critics and fans around the world. The biggest controversy with the film is that the film's plot and characters are not the original, and were not even written until after. . .

Attribute Models - Fine Grained Control

[–] *The potato*, or potato chip, is one of the best-selling snacks in the world!\n \n It comes in a variety of colors, is gluten-free (except for gluten-free chips), low in fat and saturated fat, and can be made in just 30 minutes, using only 6 simple ingredients. It's also easy to make, and can be stored in its own tin. \n \n The potato chip, however, may not be for everyone. If you are trying to lose weight, it. . .

Attribute Models - Fine Grained Control

[–] *The potato*, or potato chip, is one of the best-selling snacks in the world!\n \n It comes in a variety of colors, is gluten-free (except for gluten-free chips), low in fat and saturated fat, and can be made in just 30 minutes, using only 6 simple ingredients. It's also easy to make, and can be stored in its own tin. \n \n The potato chip, however, may not be for everyone. If you are trying to lose weight, it. . .
[Science; 0.005] *The potato* is an ancient member of the green family which, as its name suggests, is green and edible. It is native to the Middle East and parts of South Asia. It is an important crop in many parts of the world. The potato, a member of the potato family, has been domesticated for thousands of years. It can be eaten raw and cooked in its skins; it is also used as a starch and is a great source of energy and fiber in many diets and in many countries. . .

Attribute Models – Fine Grained Control

[–] *The potato*, or potato chip, is one of the best-selling snacks in the world!\n \n It comes in a variety of colors, is gluten-free (except for gluten-free chips), low in fat and saturated fat, and can be made in just 30 minutes, using only 6 simple ingredients. It's also easy to make, and can be stored in its own tin. \n \n The potato chip, however, may not be for everyone. If you are trying to lose weight, it. . .
[Science; 0.005] *The potato* is an ancient member of the green family which, as its name suggests, is green and edible. It is native to the Middle East and parts of South Asia. It is an important crop in many parts of the world. The potato, a member of the potato family, has been domesticated for thousands of years. It can be eaten raw and cooked in its skins; it is also used as a starch and is a great source of energy and fiber in many diets and in many countries. . .
[Science; 0.01] *The potato* was once thought to have no health problems and has been promoted as a nutritious food source since the mid-1800s, but recent reports indicate that it has many harmful health issues. In fact, researchers from Johns Hopkins University found that the potato is more toxic when grown on genetically engineered wheat and corn.\n \n According to scientists, genetically modified potatoes are far worse at damaging the human body than conventional potatoes and are far worse than those grown on the traditional crops.\n \n The study also revealed. . .

Attribute Models - Fine Grained Control

[Science; 0.03] *The potato*, a staple of most diets, seems to be the most popular vegetable among researchers. The research has been published and peer-reviewed.\n \n The potato has a unique ability. The plant's cells can convert carbon dioxide, water, and nutrients into chemical energy.\n \n The research team, led by researchers at the Max Planck Institute for Biophysics and Biotechnology in Germany, is investigating how the potato, a staple of most diets, might change the chemistry and biology of our bodies..
[Science; 0.05] *The potato*
[Science; 0.1] *The potato*,

Attribute Models - Fine Grained Control

[Science; 0.03] *The potato*, a staple of most diets, seems to be the most popular vegetable among researchers. The research has been published and peer-reviewed.\n \n The potato has a unique ability. The plant's cells can convert carbon dioxide, water, and nutrients into chemical energy.\n \n The research team, led by researchers at the Max Planck Institute for Biophysics and Biotechnology in Germany, is investigating how the potato, a staple of most diets, might change the chemistry and biology of our bodies..
[Science; 0.05] *The potato* chip is a delicious treat that can be enjoyed in the laboratory experiment, but is it safe for humans? \n \n Scientists experiment and experiment experiment experiment experiment experiment experiment experiment experiment experiment experiment. . . . .
[Science; 0.1] *The potato*,

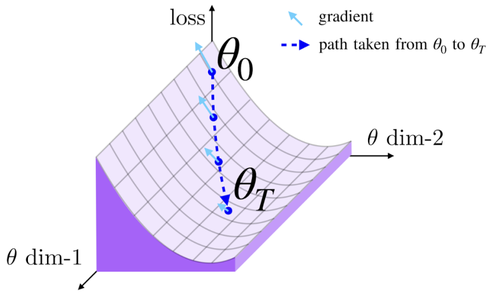
Attribute Models – Fine Grained Control

[Science; 0.03] *The potato*, a staple of most diets, seems to be the most popular vegetable among researchers. The research has been published and peer-reviewed.\n \n The potato has a unique ability. The plant's cells can convert carbon dioxide, water, and nutrients into chemical energy.\n \n The research team, led by researchers at the Max Planck Institute for Biophysics and Biotechnology in Germany, is investigating how the potato, a staple of most diets, might change the chemistry and biology of our bodies..
[Science; 0.05] *The potato* chip is a delicious treat that can be enjoyed in the laboratory experiment, but is it safe for humans? \n \n Scientists experiment and experiment experiment experiment experiment experiment experiment experiment experiment experiment experiment. . . . . .
[Science; 0.1] *The potato*, which scientists at the lab experiment . . .

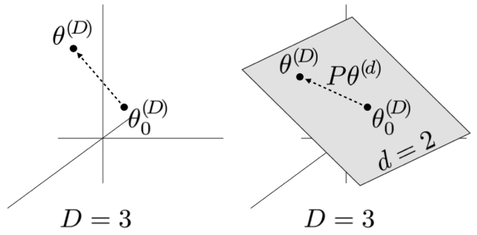
Project IV: PPLM

The end (of the project section)

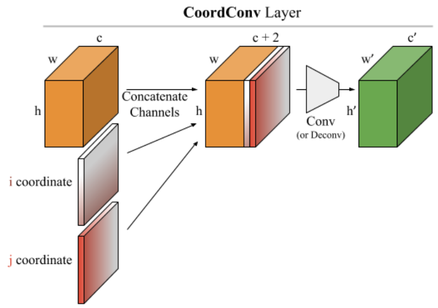
Loss change allocation (LCA)



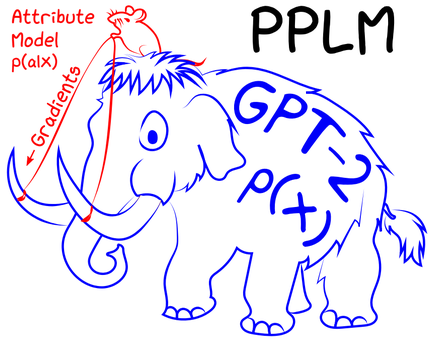
Intrinsic Dimension



CoordConv



PPLM



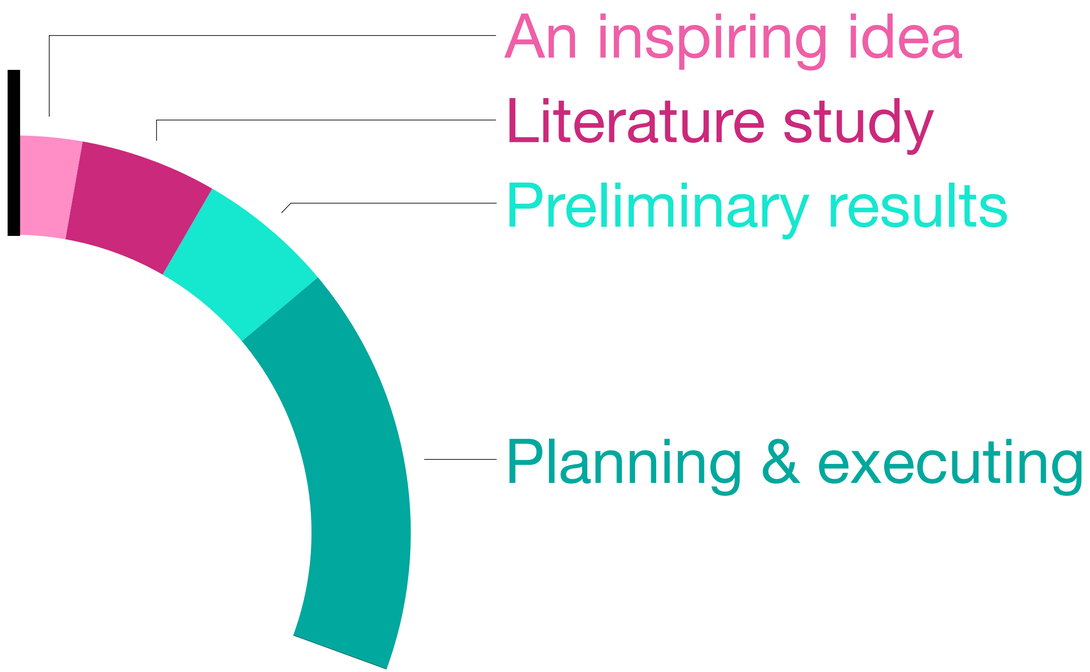
A complete research cycle

A complete research cycle

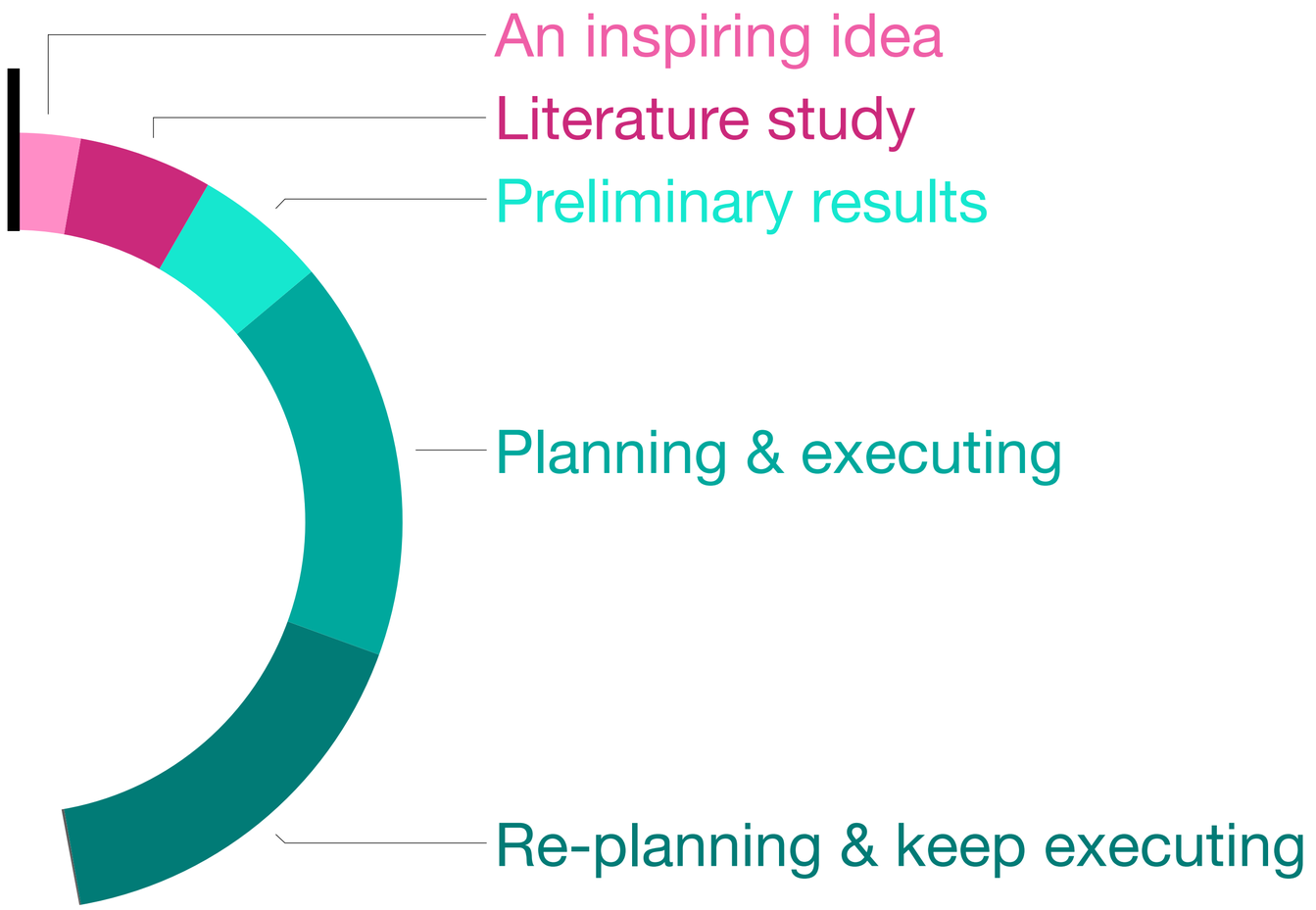
A complete research cycle

A complete research cycle

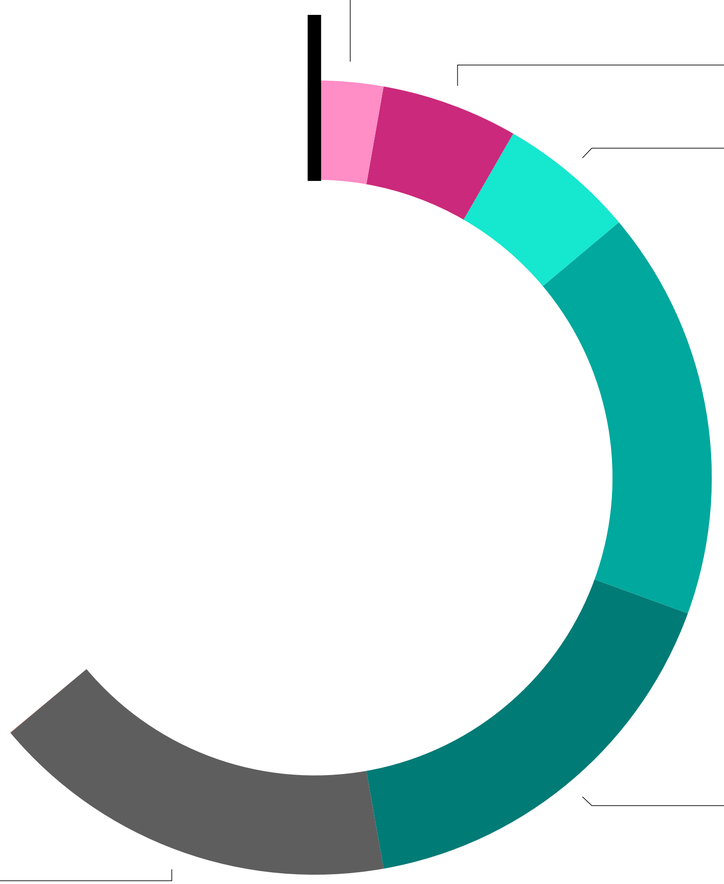
A complete research cycle



A complete research cycle



A complete research cycle



An inspiring idea

Literature study

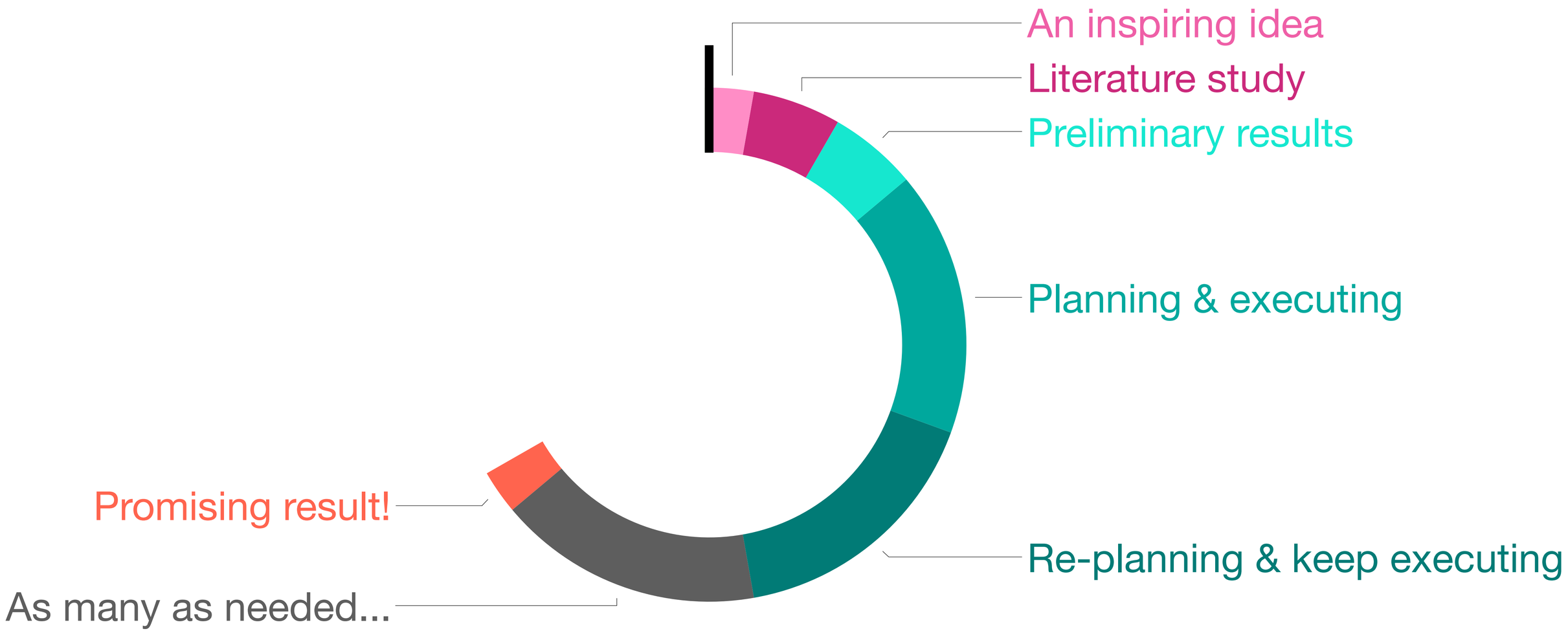
Preliminary results

Planning & executing

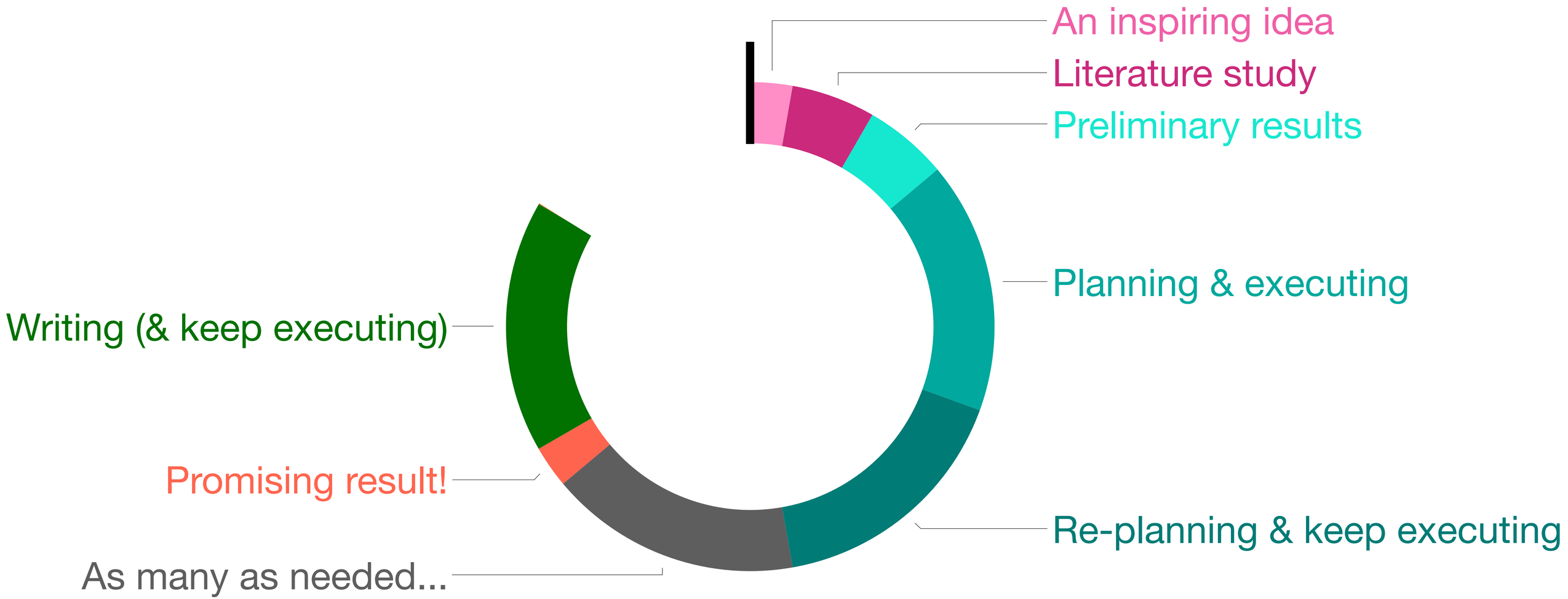
Re-planning & keep executing

As many as needed...

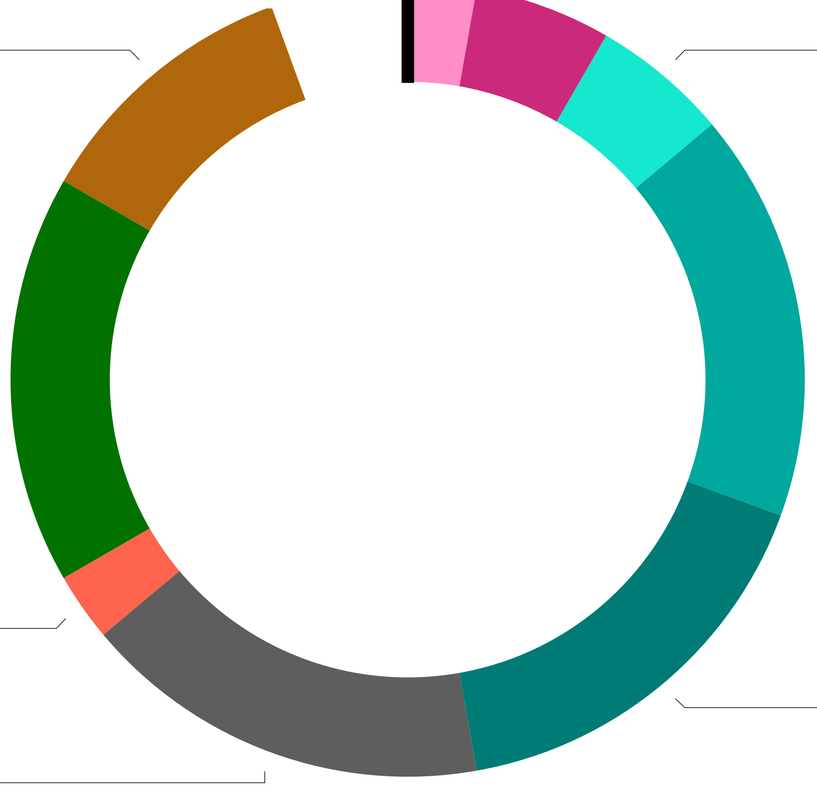
A complete research cycle



A complete research cycle



A complete research cycle



An inspiring idea

Literature study

Preliminary results

Planning & executing

Re-planning & keep executing

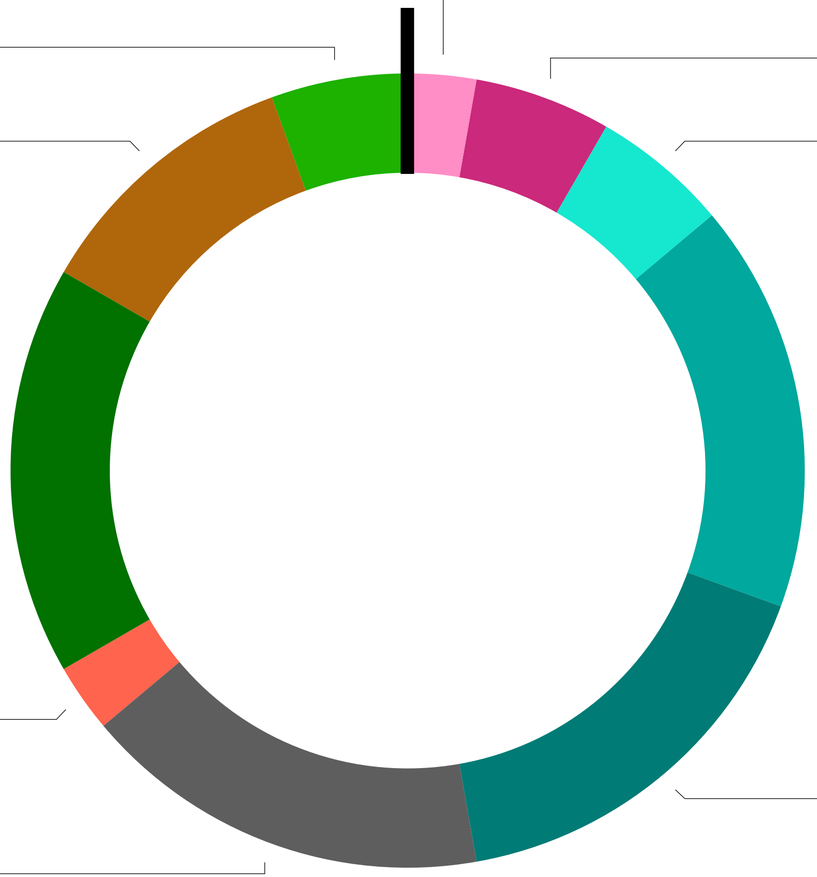
As many as needed...

Promising result!

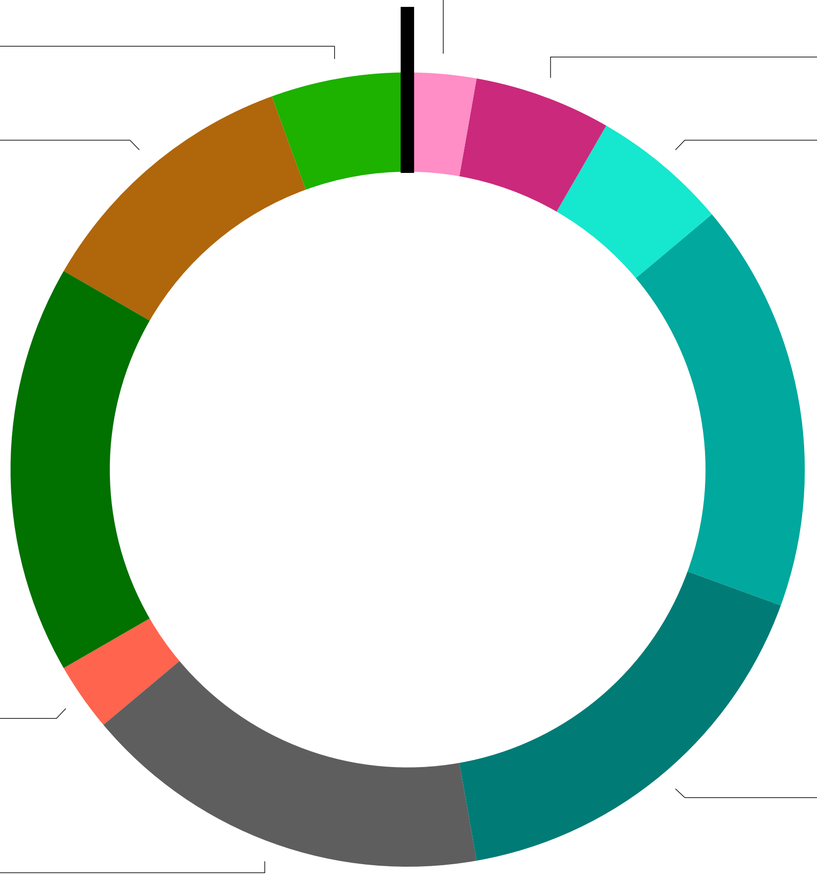
Writing (& keep executing)

The publishing process

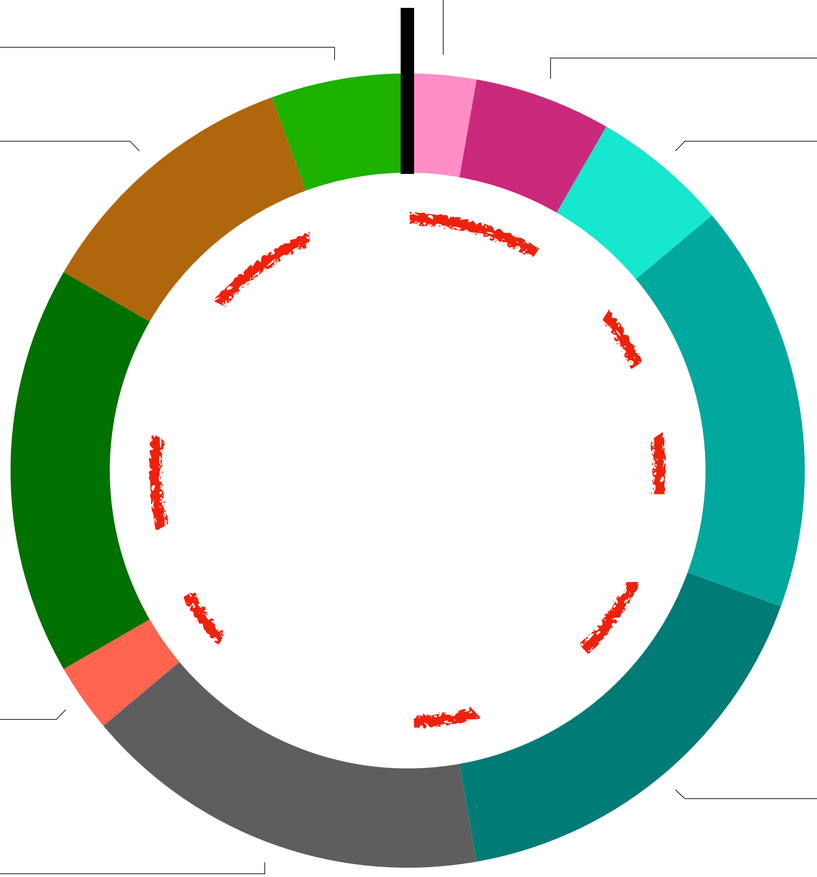
A complete research cycle



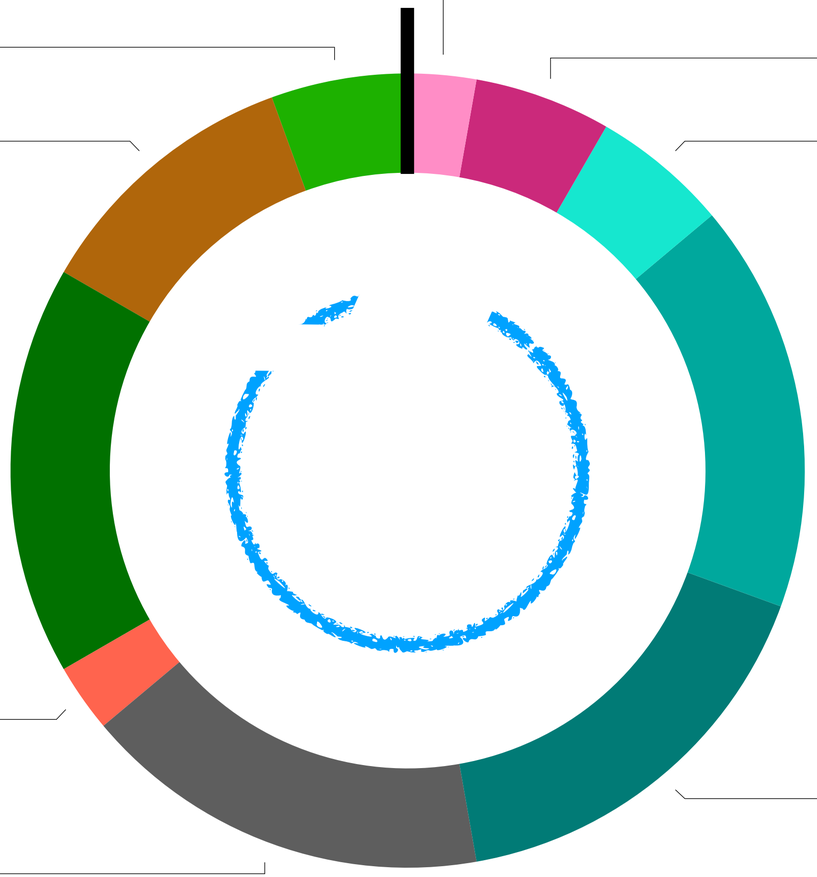
Where to have fun



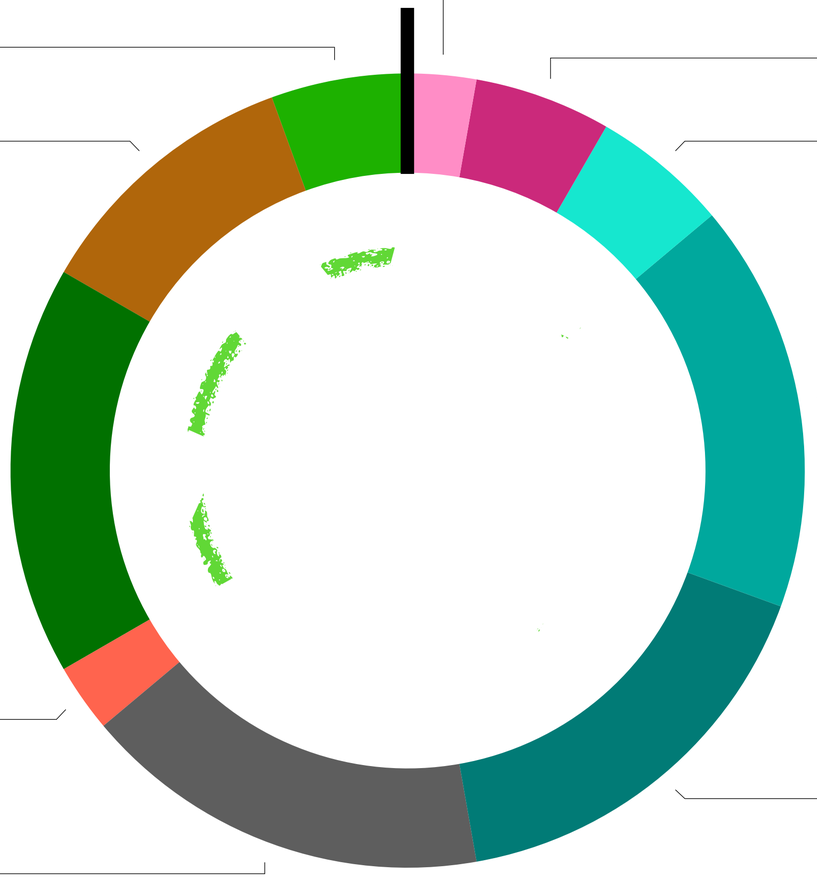
Where to have fun, if you are a visionary



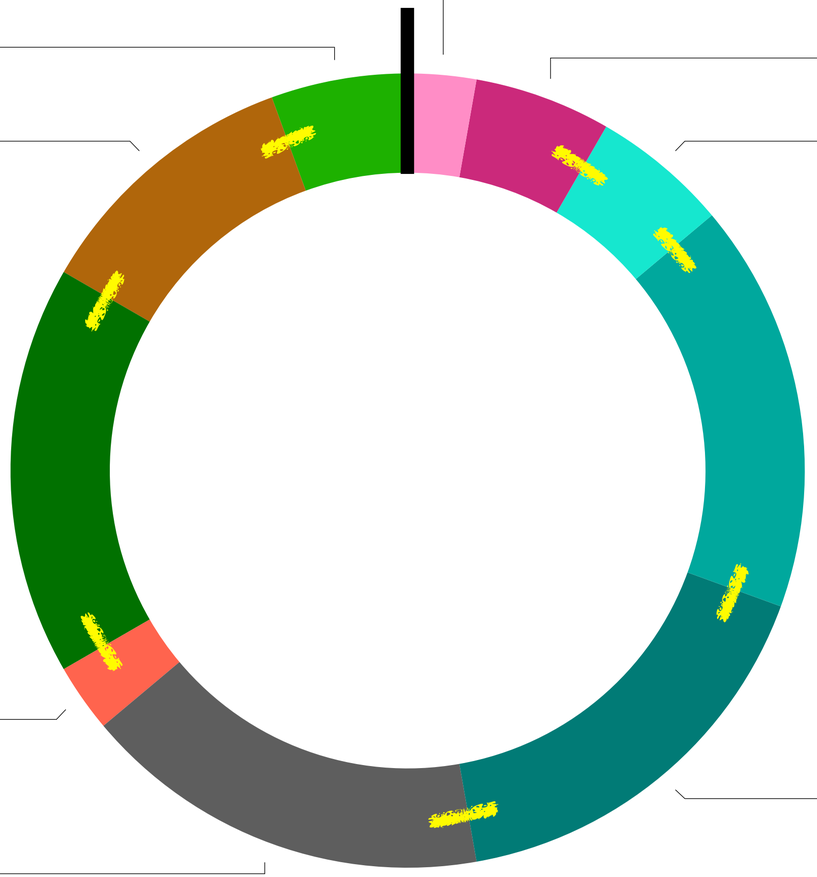
Where to have fun, if you are a coder



Where to have fun, if you are a creative writer

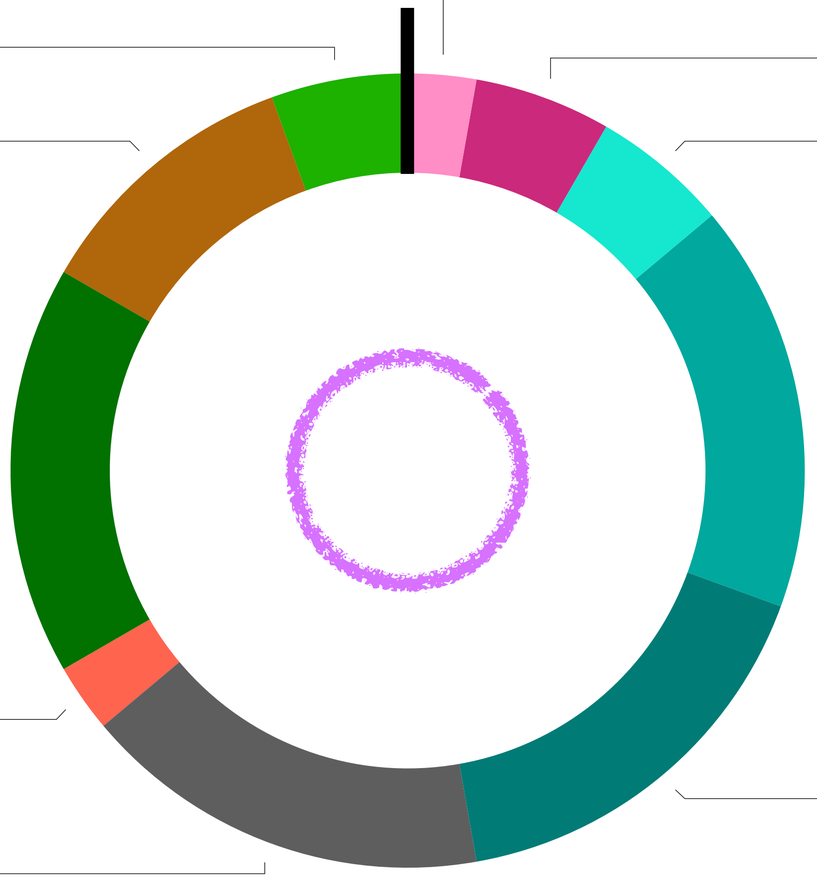


Where to have fun, if you are a manager



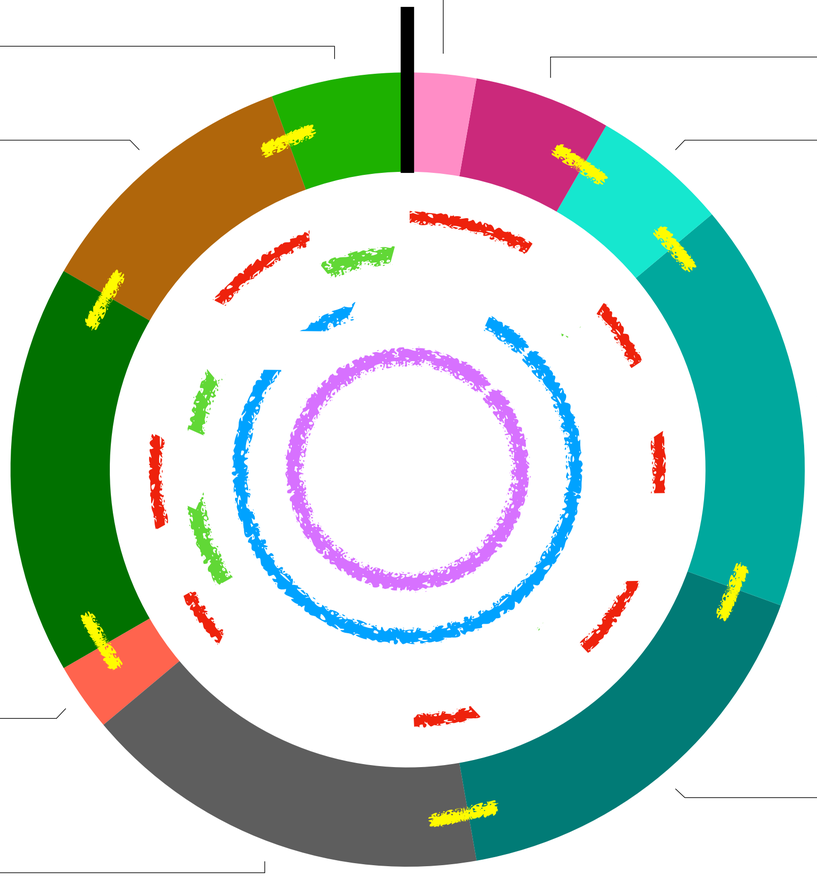
An inspiring idea
Literature study
Preliminary results
Planning & executing
Re-planning & keep executing
As many as needed...
Promising result!
Writing (& keep executing)
The publishing process
Writing

Where to have fun, if you are emotionally healthy



An inspiring idea
Literature study
Preliminary results
Planning & executing
Re-planning & keep executing
As many as needed...
Promising result!
Writing (& keep executing)
The publishing process
Writing

A complete research cycle, and where to have fun



An inspiring idea

Literature study

Preliminary results

Planning & executing

Re-planning & keep executing

As many as needed...

Promising result!

Writing (& keep executing)

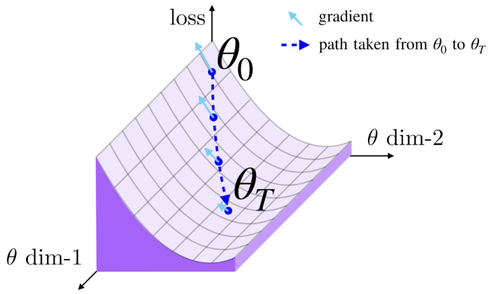
The publishing process

Writing

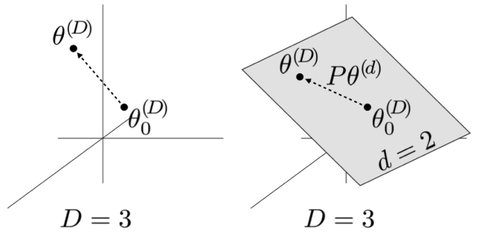
manager visionary creative writer coder emotionally healthy

What was the most fun I had in each of these?

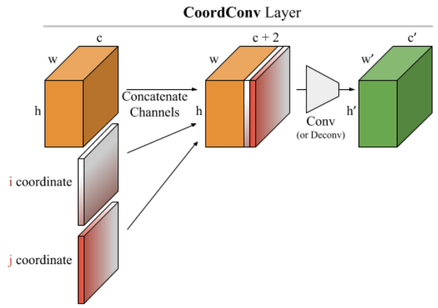
Loss change allocation (LCA)



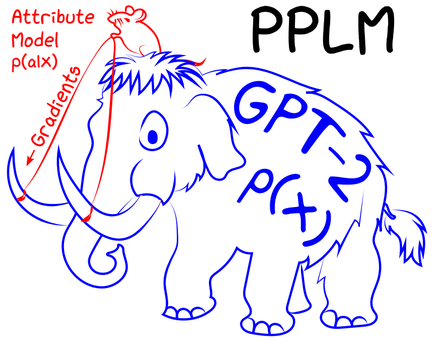
Intrinsic Dimension



CoordConv

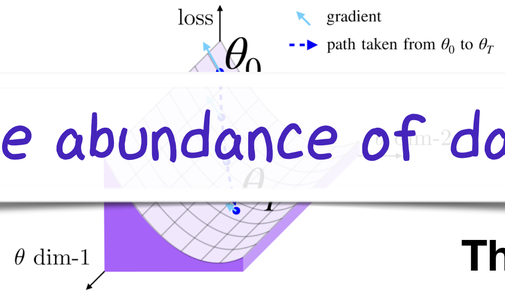


PPLM



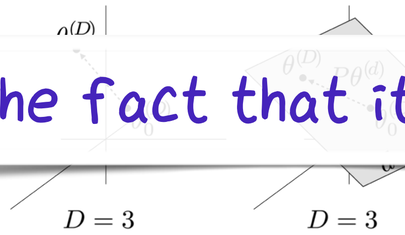
What was the fun I had in each of these?

Loss change allocation (LCA)



The abundance of data

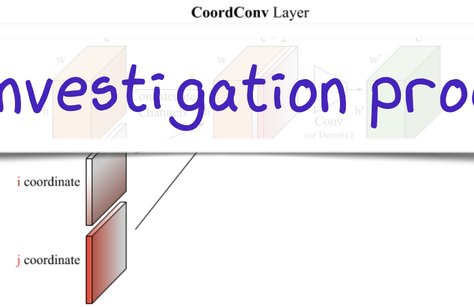
Intrinsic Dimension



The fact that it worked

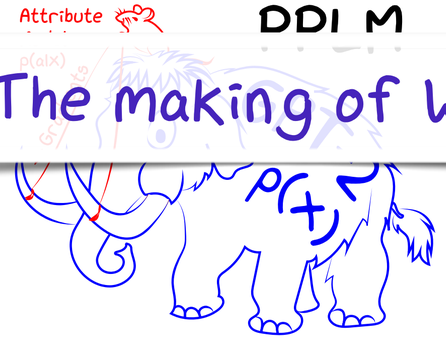
The abundance of data

CoordConv



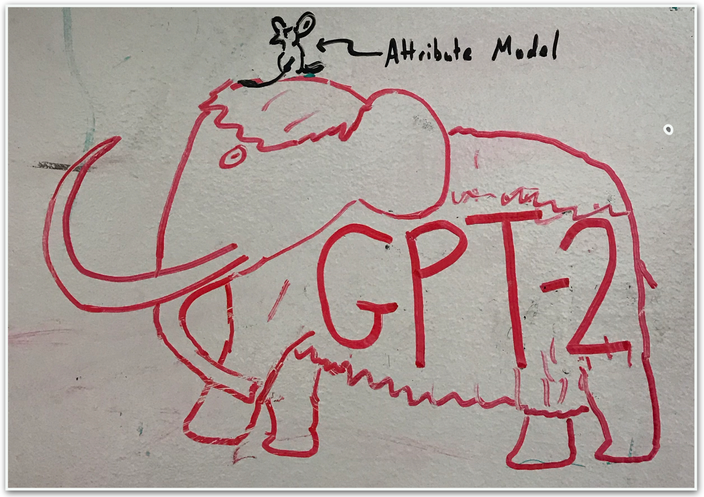
The investigation process

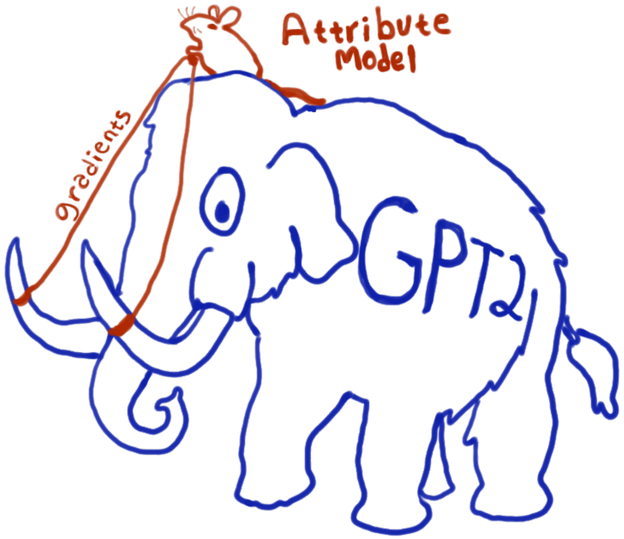
PPLM

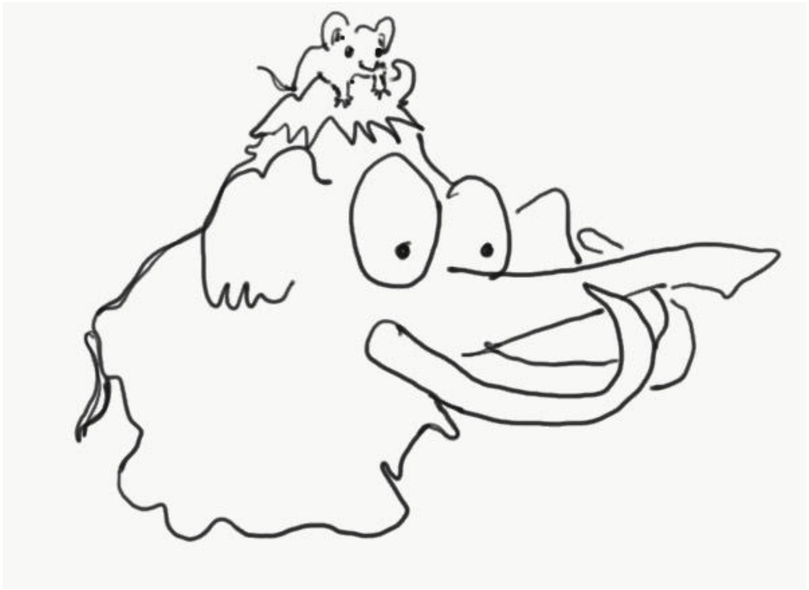


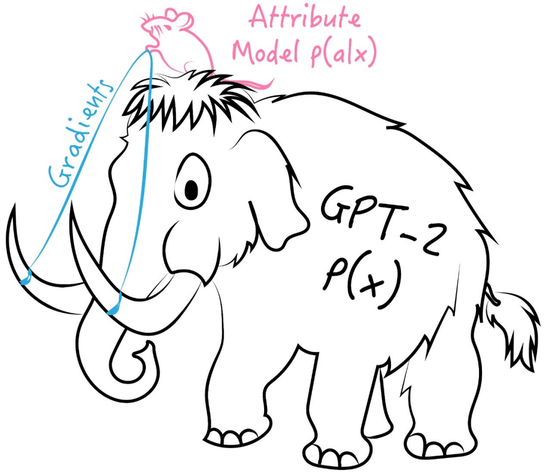
The making of Wooly

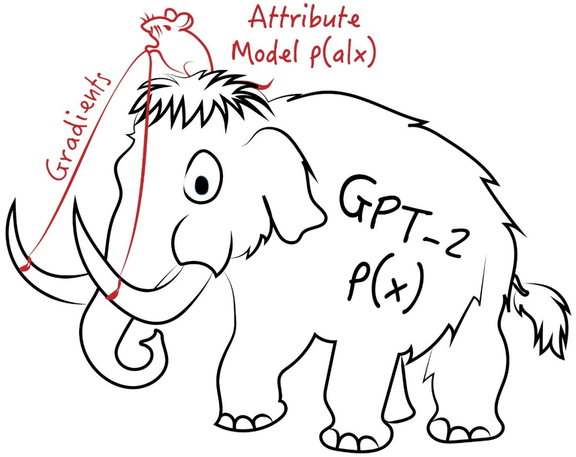
The making of Wooly

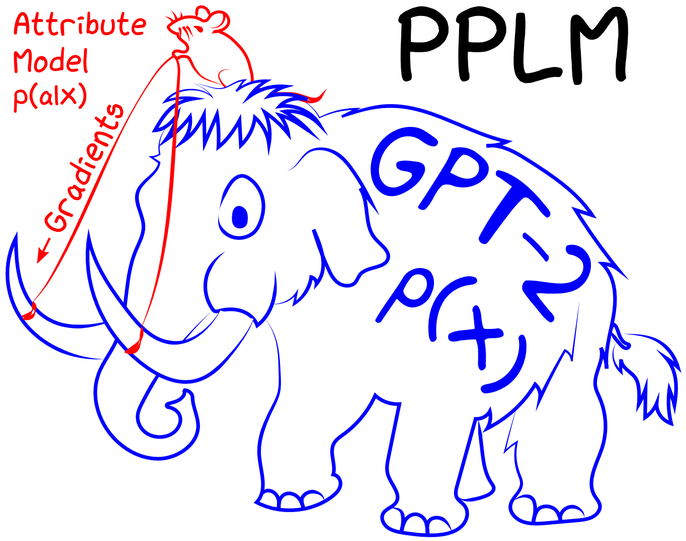








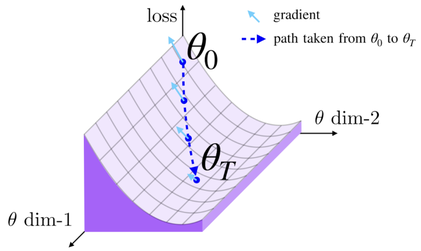


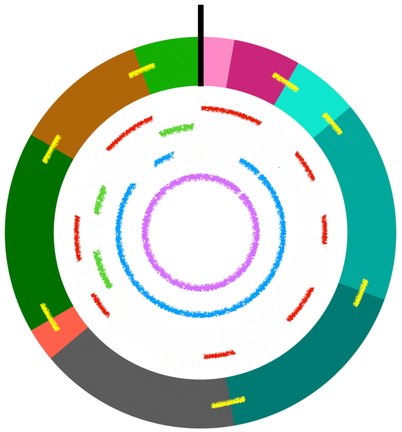


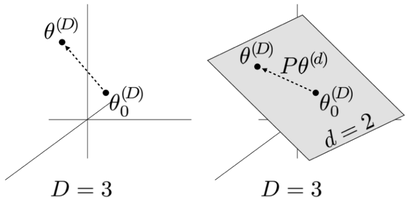
How to Have Fun in AI Research

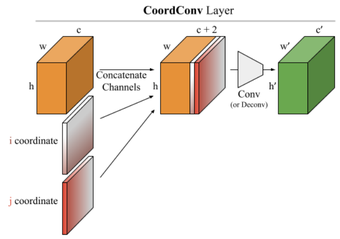
Science is competitive, aggressive, demanding. It is also imaginative, inspiring, uplifting.

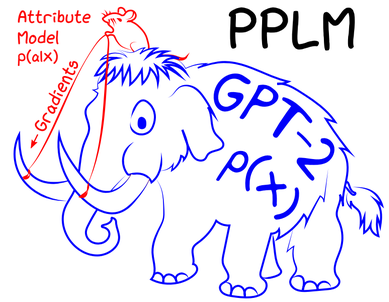
–**Vera Rubin**, *Bright Galaxies, Dark Matters (1997), p. 219*



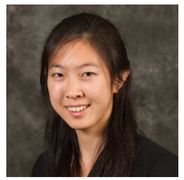


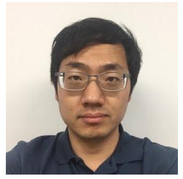






How to Have Fun in AI Research





Janice Lan

Hattie Zhou

Chunyuan Li



Sumanth Dathathri

Andrea Madotto

Alex Sergeev



Heerad Farkhoor

Jane Hung

Piero Molino

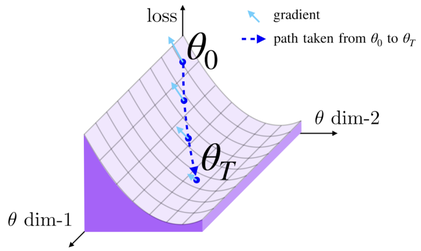


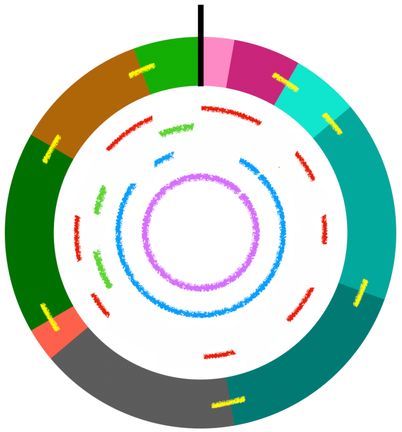
Felipe Petroski Such

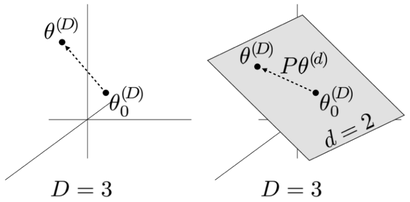
Joel Lehman

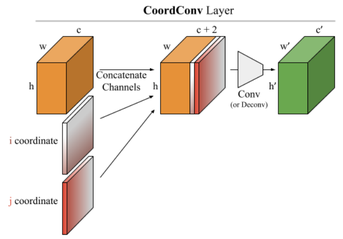
Eric Frank

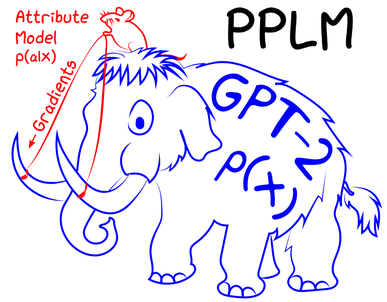
Jason Yosinski











Thanks!

Twitter: @savvyRL

Email: mimosavvy@gmail.com

Website: http://www.rosanneliu.com/