

The Other AI: An Intuitive Understanding of Artificial Intelligence

Maria Teleki
PhD Student

*Slides will
be posted
after the
talk!*



Texas A&M University
Department of Computer Science & Engineering



AI
is all about mapping
inputs
and
outputs.

AlphaFold is an AI system developed by **Google DeepMind** that predicts a protein's 3D structure from its amino acid sequence. It regularly achieves accuracy competitive with experiment.



Input:

Protein's amino acid sequence

Output:

Protein's 3D structure

OpenAI rolls out Advanced Voice Mode for ChatGPT

 TechRadar

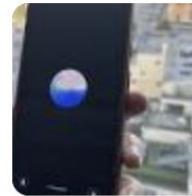
The ChatGPT Advanced Voice mode rollout is starting, and you might be...



1 day ago

 CNBC

OpenAI just launched advanced voice mode for audio chats with ChatGPT....



36 minutes ago

 Tom's Guide

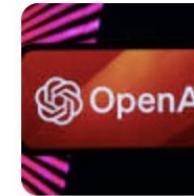
ChatGPT Advanced Voice Mode rolling out right now – here's what's new



5 hours ago

 Bloomberg.com

OpenAI Rolls Out New Voice Assistant to All Paid ChatGPT Users



9 hours ago

[View full coverage →](#)

Input:
Voice signal

Output:
Text



DALL-E 2 · An expressive oil painting of a chocolate chip cookie being dipped in a glass of milk, depicted as an explosion of flavors.



DALL-E 3 · An expressive oil painting of a chocolate chip cookie being dipped in a glass of milk, depicted as an explosion of flavors.

Input:
Text

Output:
Image

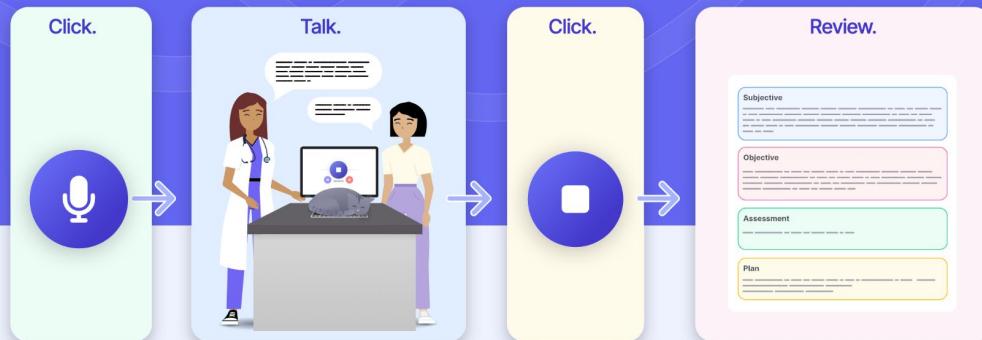
The World's First AI Veterinary Notes.

With Scribenote, you can
get your veterinary notes
done in just two clicks.



Try Scribenote

Free to try. No credit card required.



Input:
Voice signal

Output:
Text (SOAP Notes)

Point-of-care > Vetscan Imagyst

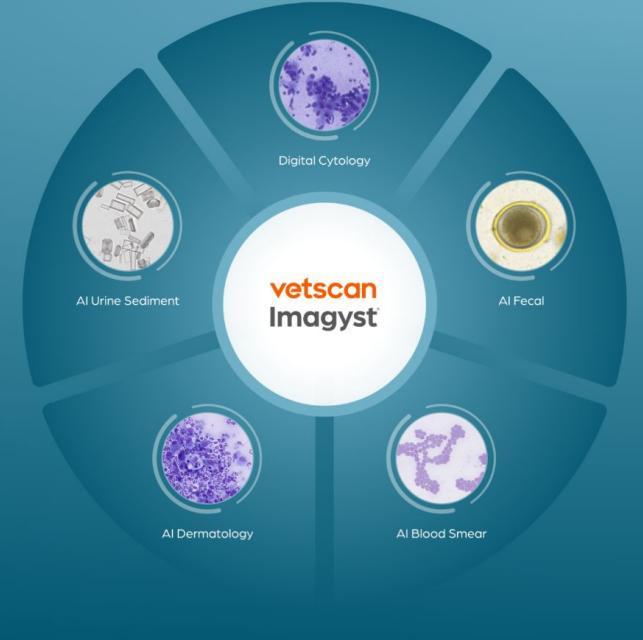
ZOETIS DIAGNOSTICS

vetscan Imagyst®

Transform your clinic with five-in-one diagnostics

Offering five testing capabilities on a single analyzer, Vetscan Imagyst® is the first and only AI-powered technology of its kind to provide specialist-level veterinary diagnostic results within minutes.¹⁻⁴

[Request More Information](#)



Input:
Image

Output:
Text (Diagnosis)

**We're going to talk
about how AI works -
There are a lot of moving
parts, it'll all tie together
at the end.**

**Why? Intuition is
really helpful for
solving problems.**

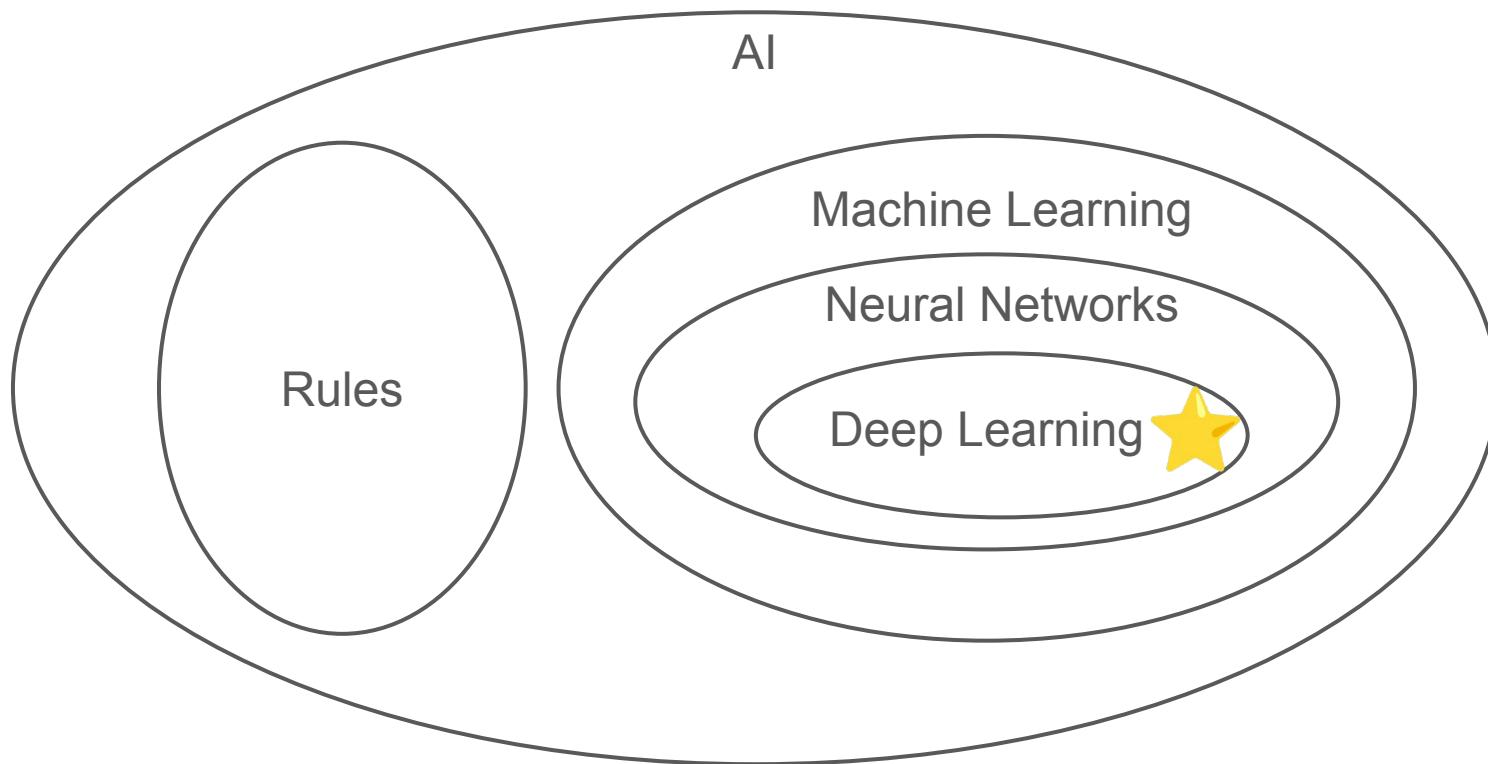
**If you want to effectively use
your tools, you need to know
something about your tools.**

Why? Intuition is really helpful for solving problems.

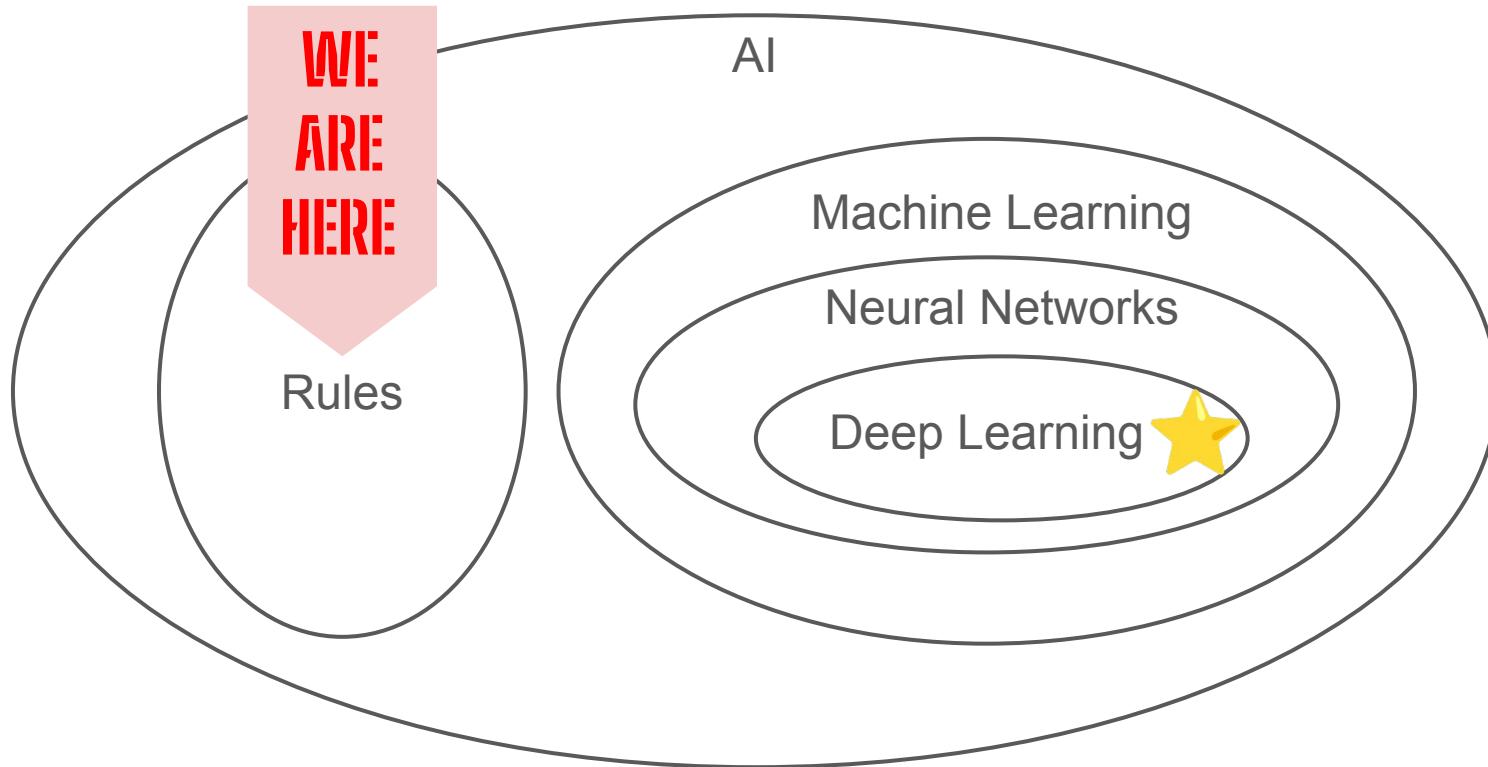
**If you want to effectively use
your tools, you need to know
something about your tools.**

- Save more animals ❤️ w/ advanced AI tools
- Get more clients = 💰 make bank

Techniques for AI



Techniques for AI



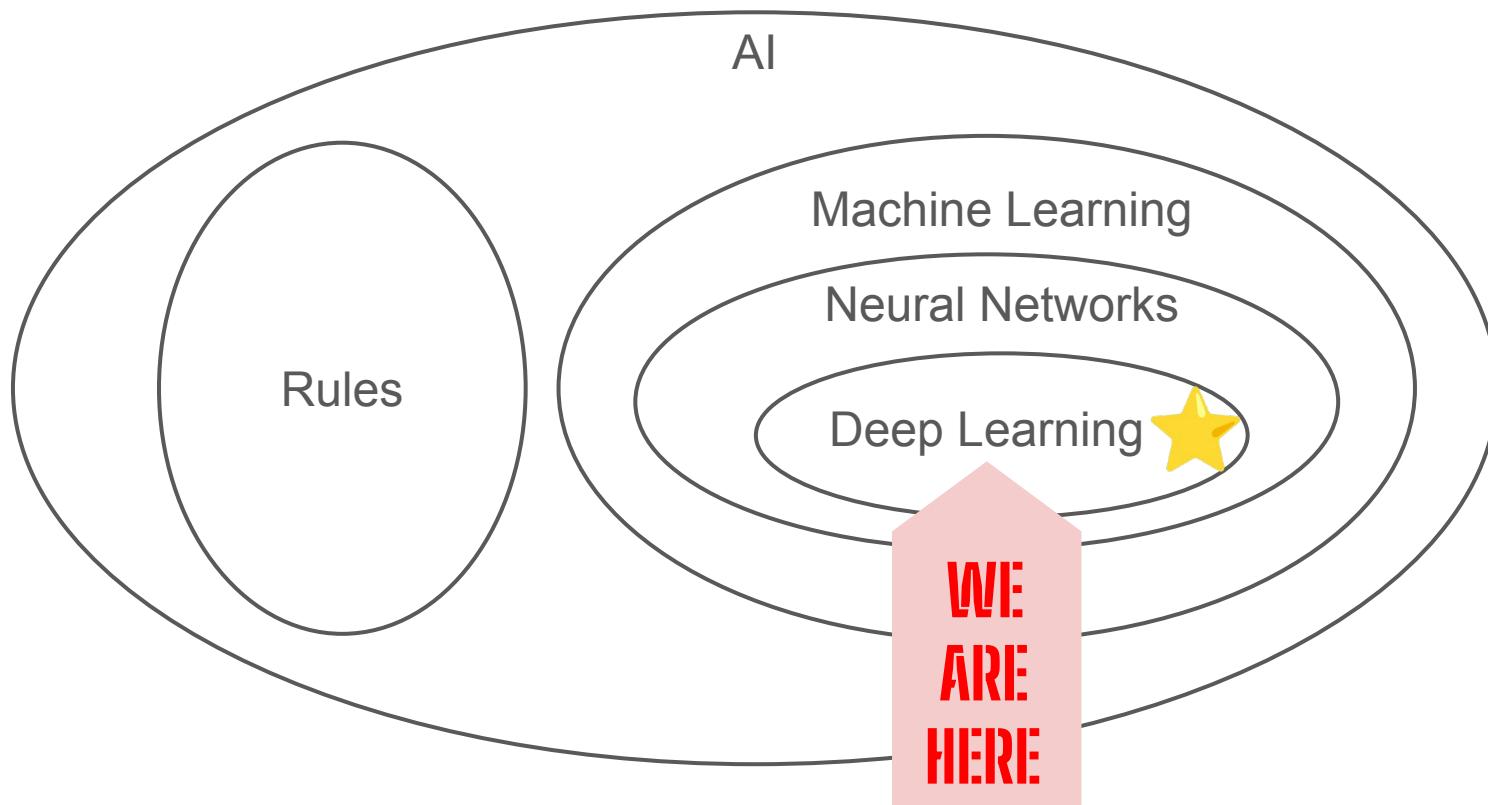
Rules

Pros/Cons?

RULE (PATIENT, COUGHING) → FLU

RULE (PATIENT, COUGHING, ELDERLY) → DEATHLY FLU

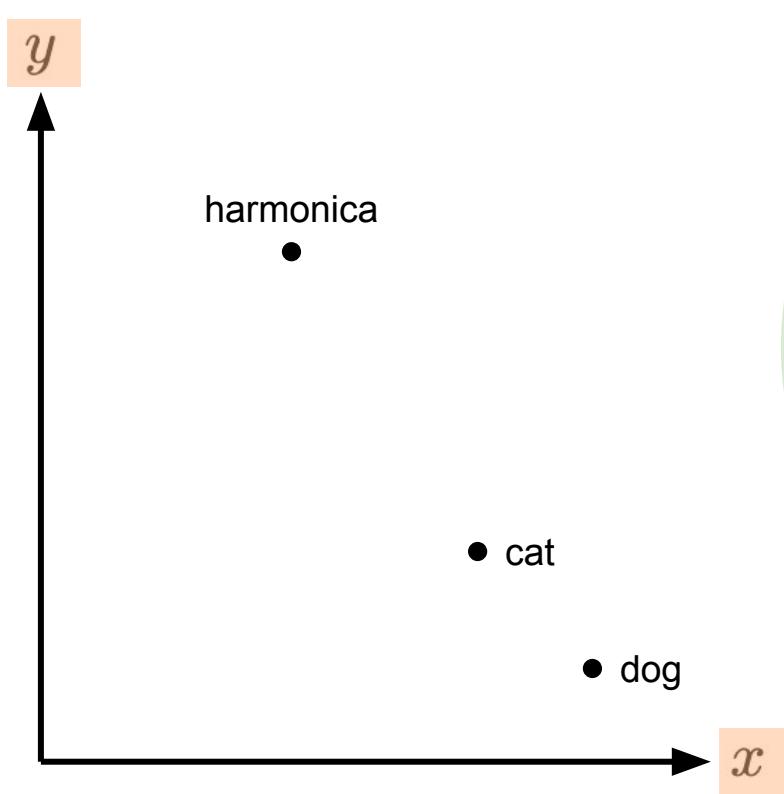
Techniques for AI



*Neural networks are
really good at dealing
with numbers.*

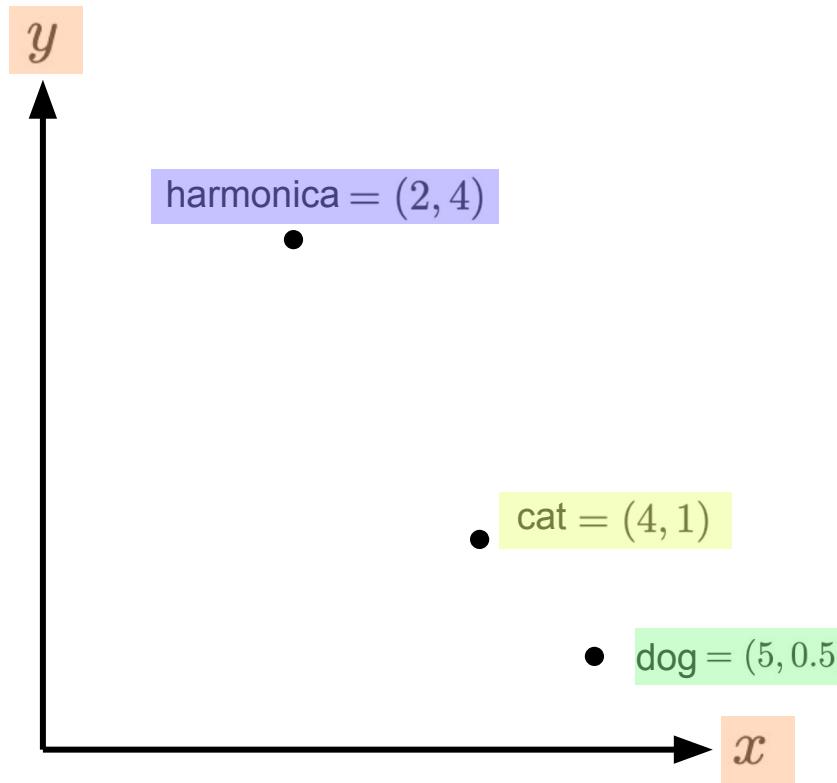
How do neural networks deal with words?

The Foundation of AI Systems!

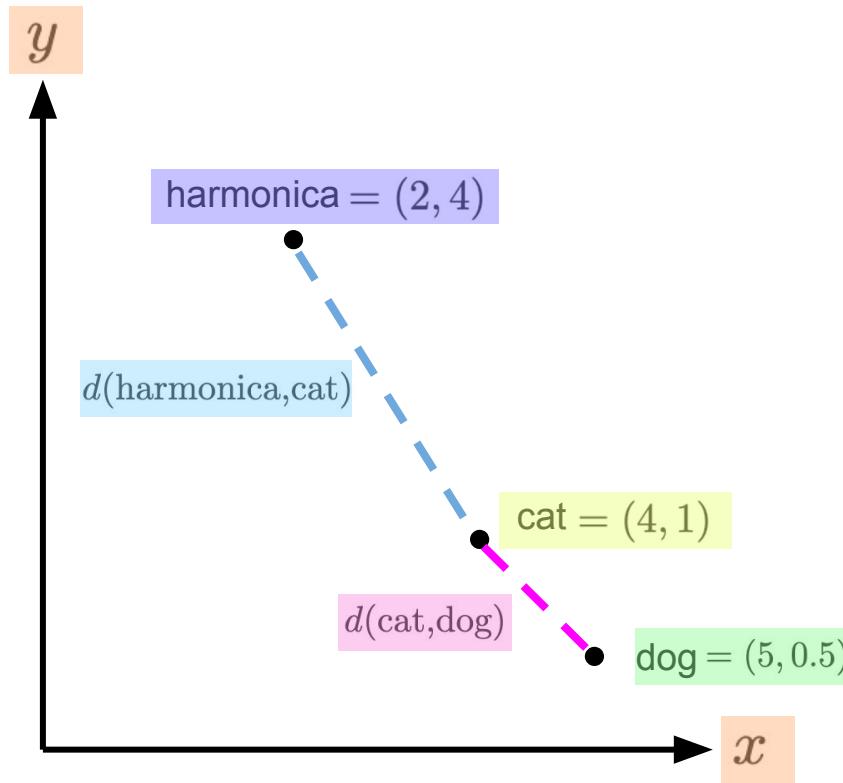


*Plot the words
on a graph!
This **assigns**
them
numbers so
the computer
can work with
them!*

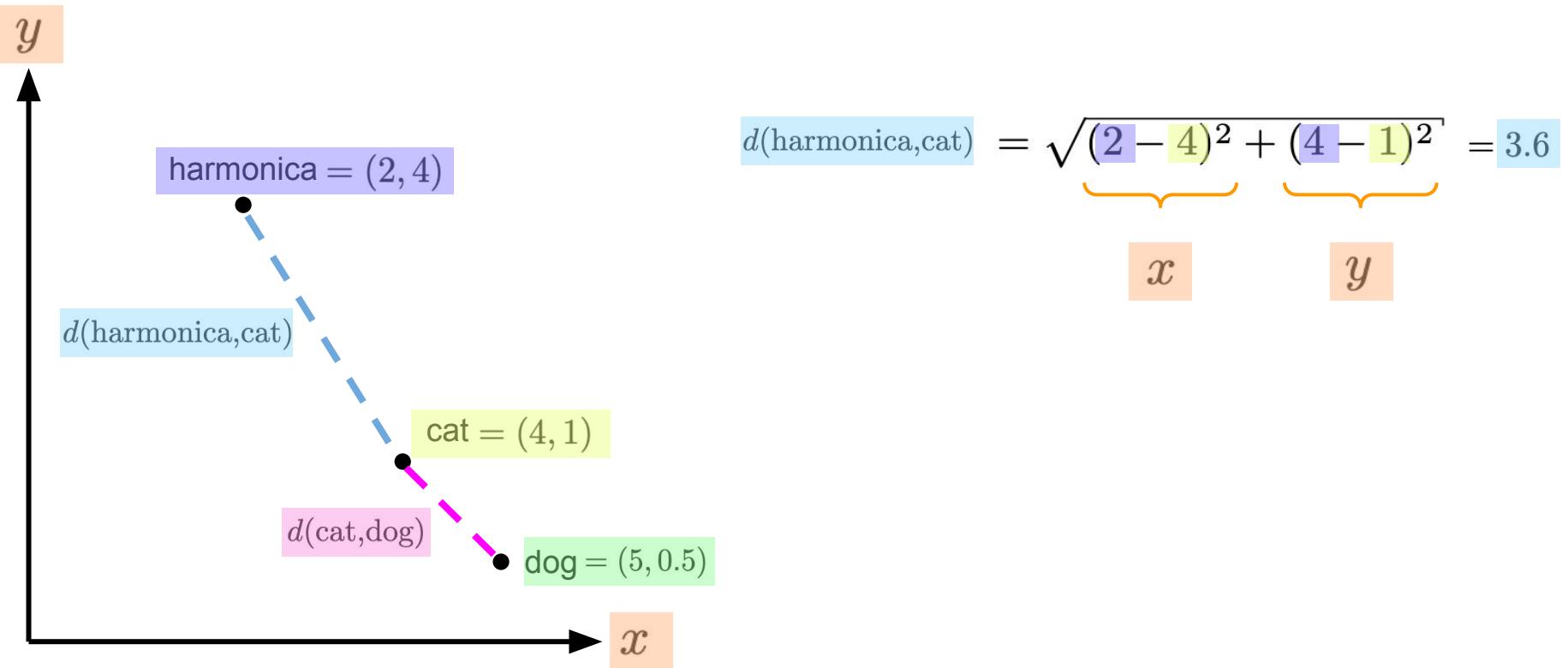
The Foundation of AI Systems!



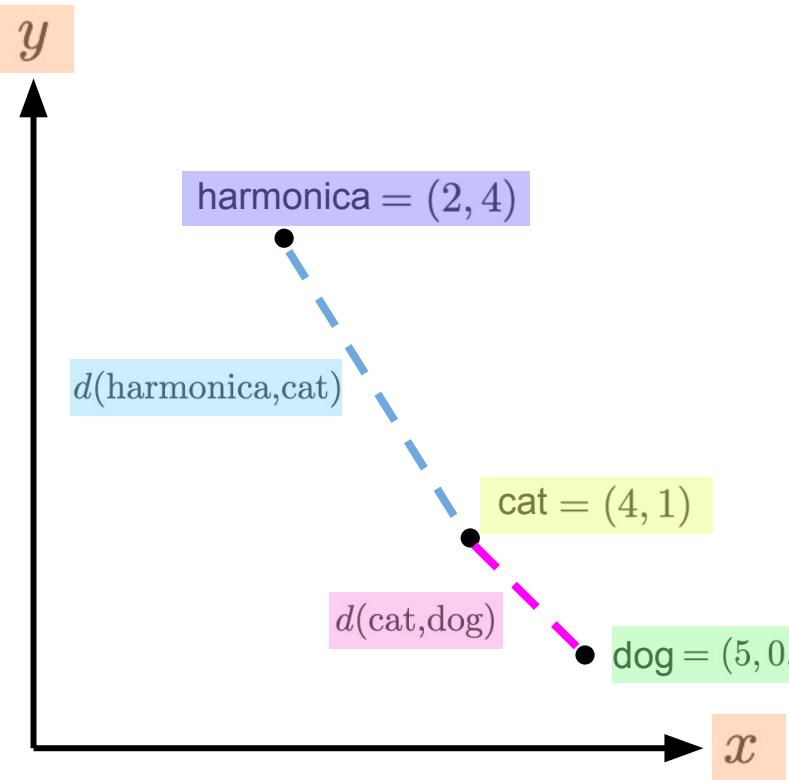
The Foundation of AI Systems!



The Foundation of AI Systems!



In 2D!



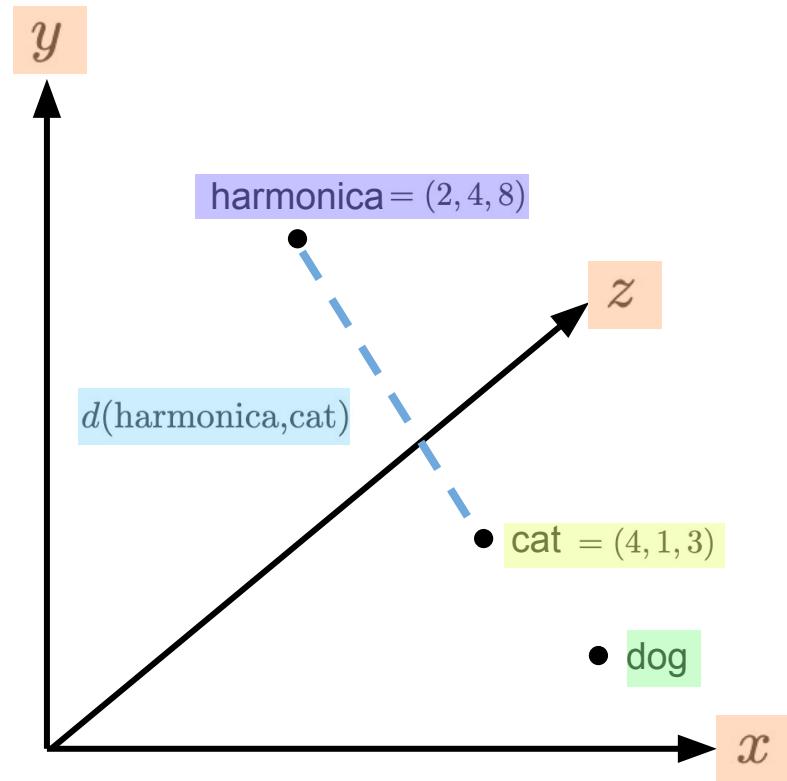
$$d(\text{harmonica}, \text{cat}) = \sqrt{(2 - 4)^2 + (4 - 1)^2} = 3.6$$

x *y*

$$d(\text{cat}, \text{dog}) = \sqrt{(4 - 5)^2 + (1 - 0.5)^2} = 1.1$$

x *y*

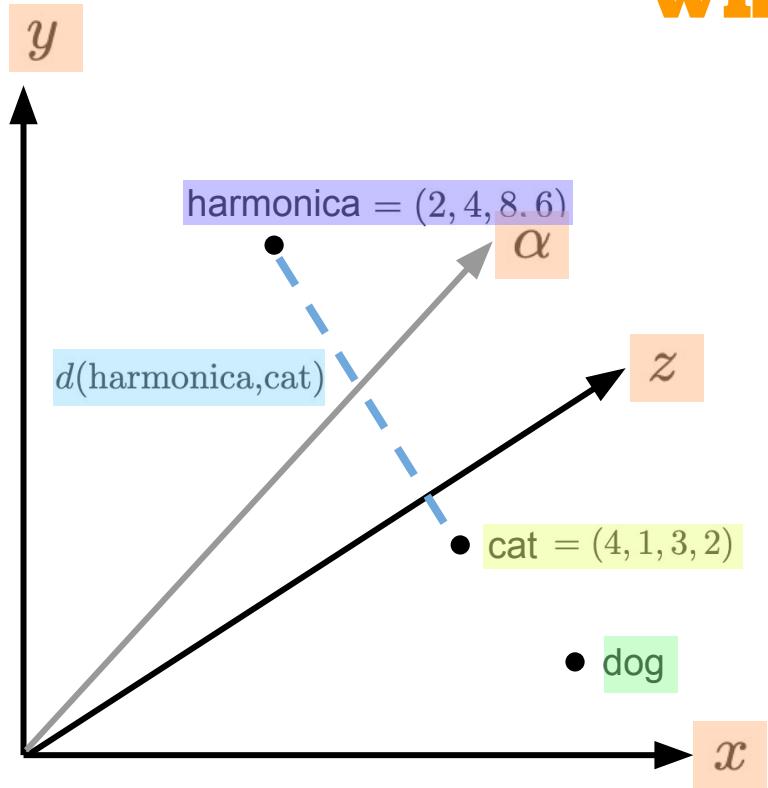
What about 3D?



$$d(\text{harmonica}, \text{cat}) = \sqrt{(2 - 4)^2 + (4 - 1)^2 + (8 - 3)^2}$$

x y z

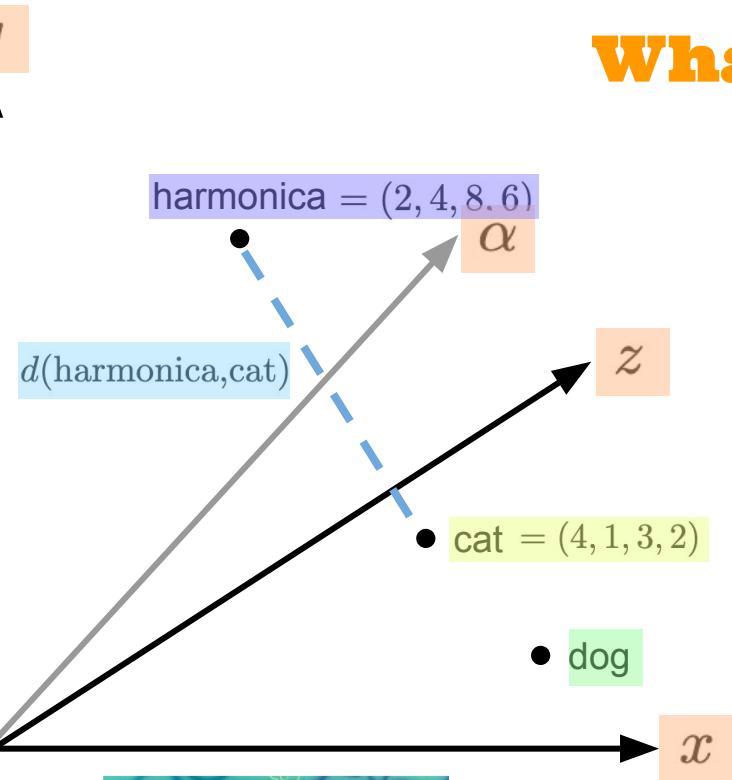
What about 4D?



$$d(\text{harmonica}, \text{cat}) = \sqrt{(2 - 4)^2 + (4 - 1)^2 + (8 - 3)^2 + (6 - 2)^2}$$

x y z α

What about 4D?



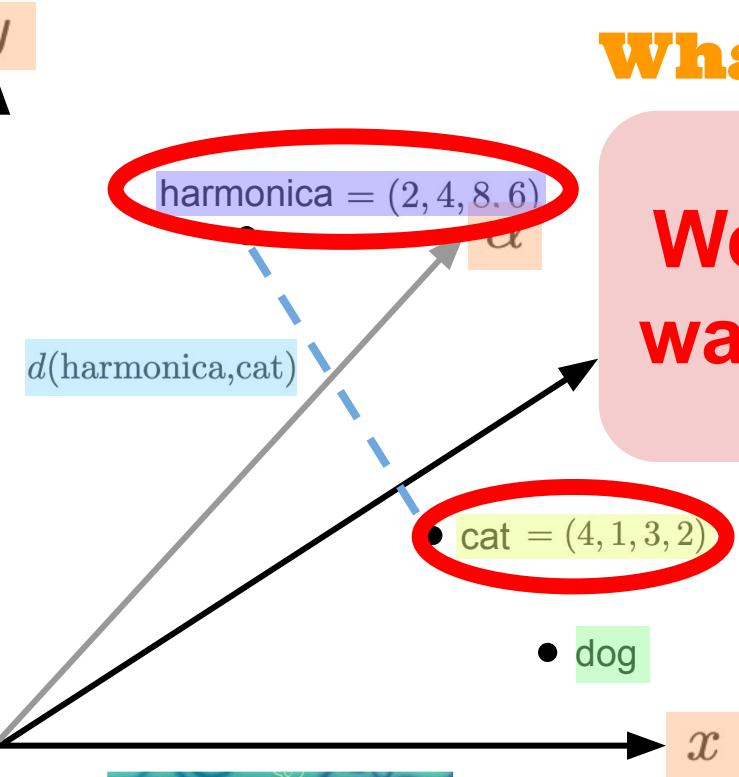
$$d(\text{harmonica}, \text{cat}) = \sqrt{(2 - 4)^2 + (4 - 1)^2 + (8 - 3)^2 + (6 - 2)^2}$$

x y z α



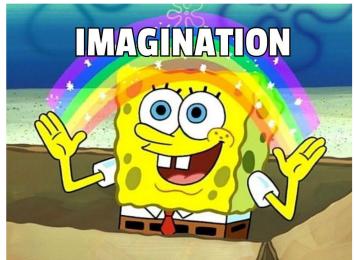
We can't really visualize higher dimensions, we end up having to project down to 2D or 3D. But we can use our imagination, and math tools like meshes and nets to understand what's happening up there!

What about 4D?



We have to have some way to get these lists...

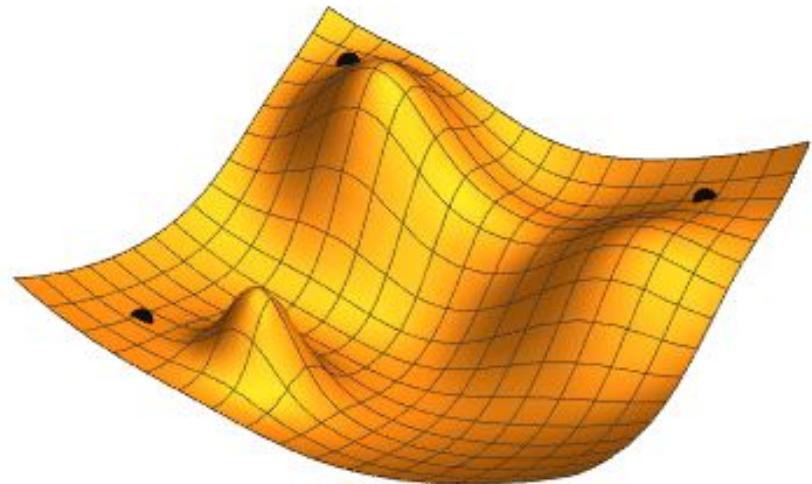
$$= 3^2 + (6 - 2)^2$$



We can't really visualize higher dimensions, we end up having to project down to 2D or 3D. But we can use our imagination, and math tools like meshes and nets to understand what's happening up there!

How can a neural network help us w/ this?

Neural networks are trying to **find the best equation to map inputs to outputs.**

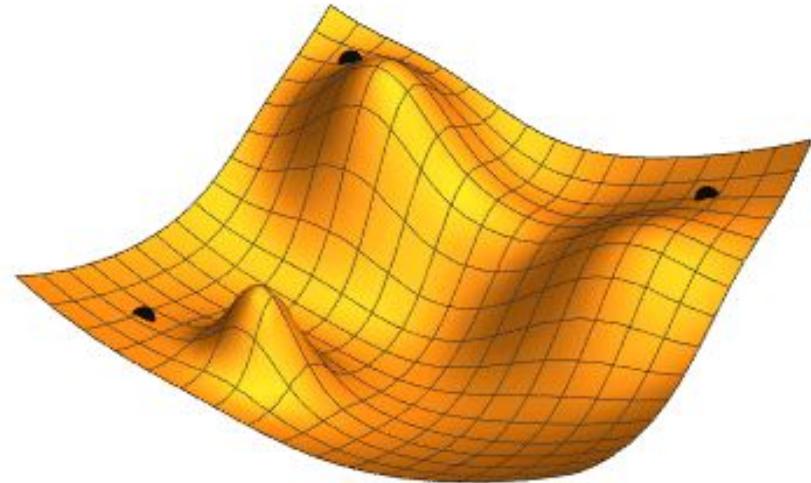


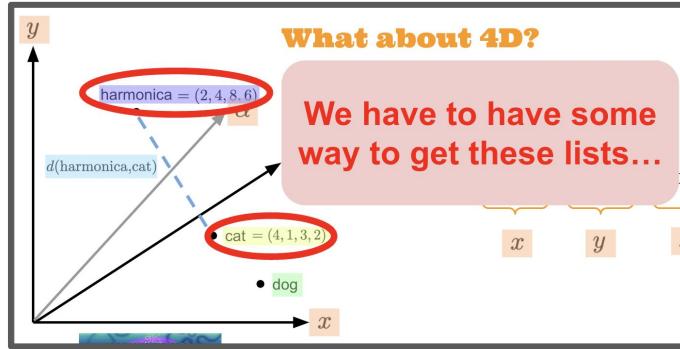
How can a neural network help us w/ this?

Neural networks are trying to **find the best equation to map inputs to outputs.**

How do they do this?

1. Start with **random numbers**
2. See **how far off** output is
3. **Update** the equation
4. **Repeat 1-3** until we're at a **minimum!**



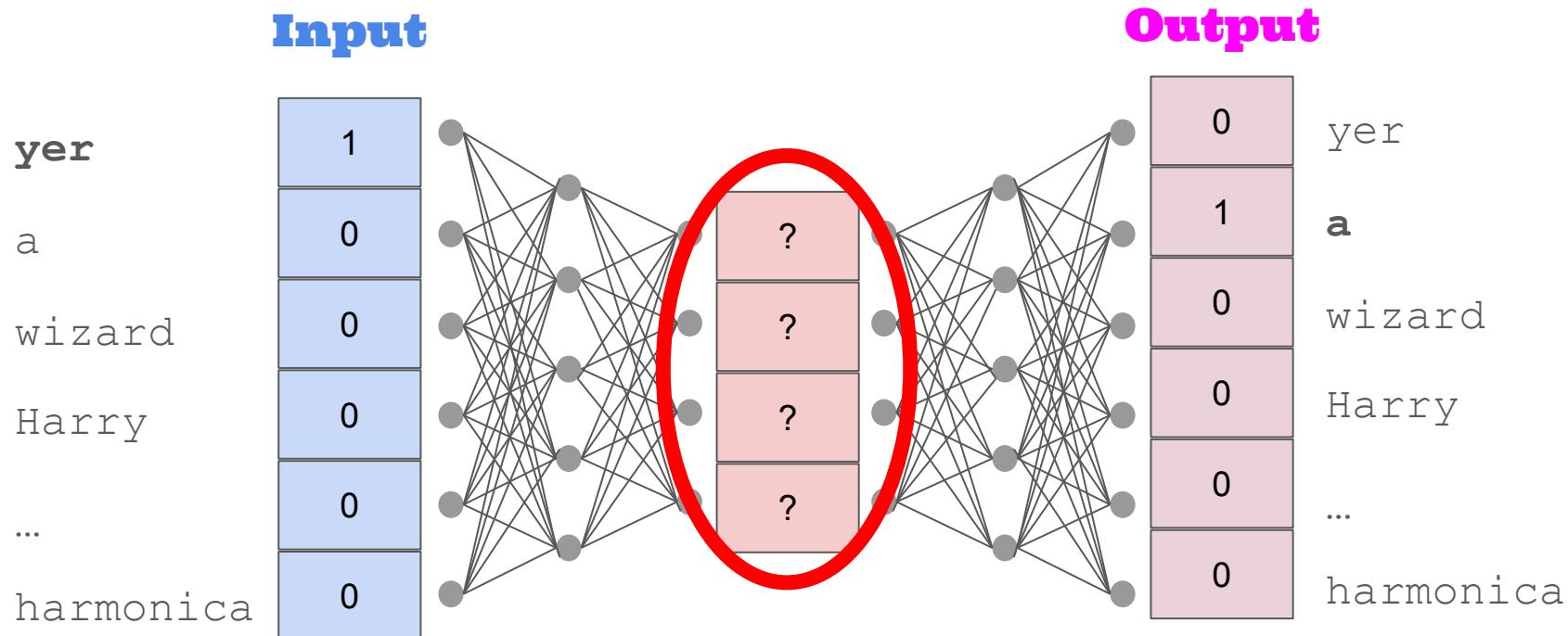


We can turn text into input, output pairs and give it to the neural network, and it can make the lists!

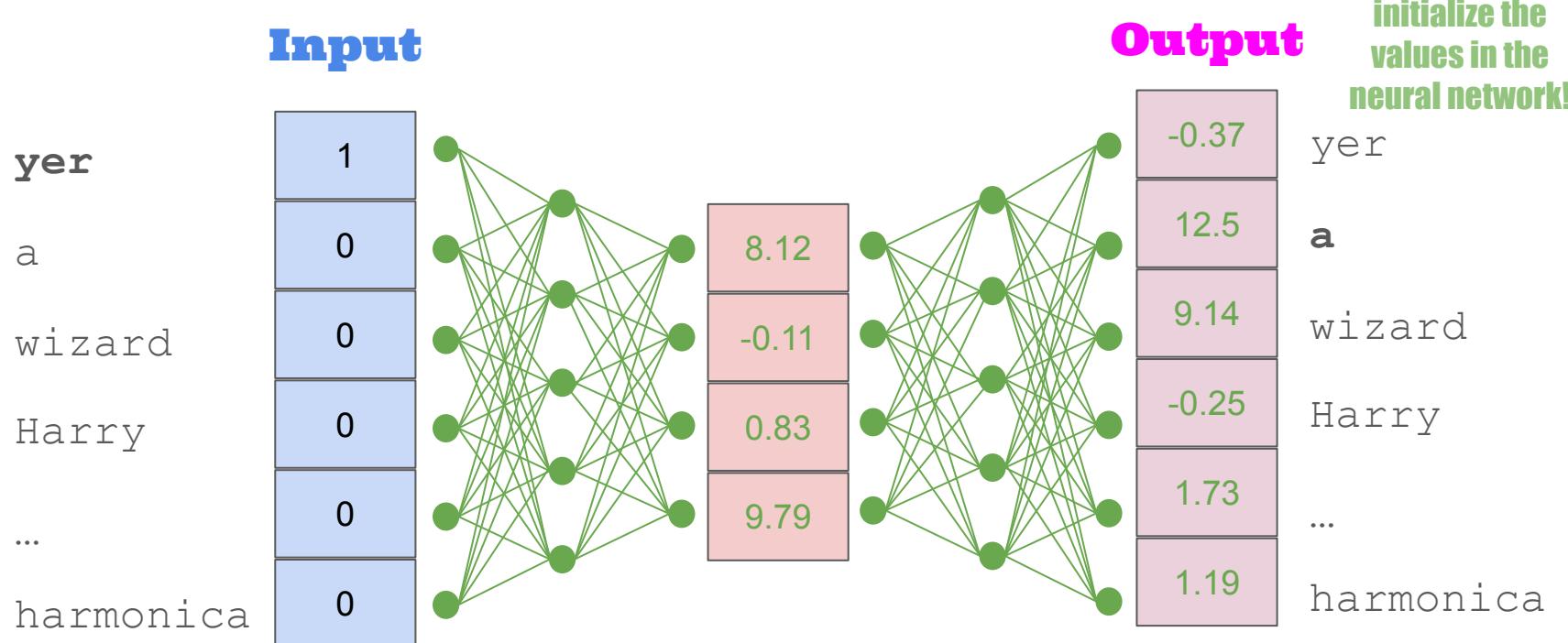
Yer a wizard Harry

Yer	a	wizard	Harry	\rightarrow	yer,	a
Yer	a	wizard	Harry	\rightarrow	a,	wizard
Yer	a	wizard	Harry	\rightarrow	wizard,	Harry

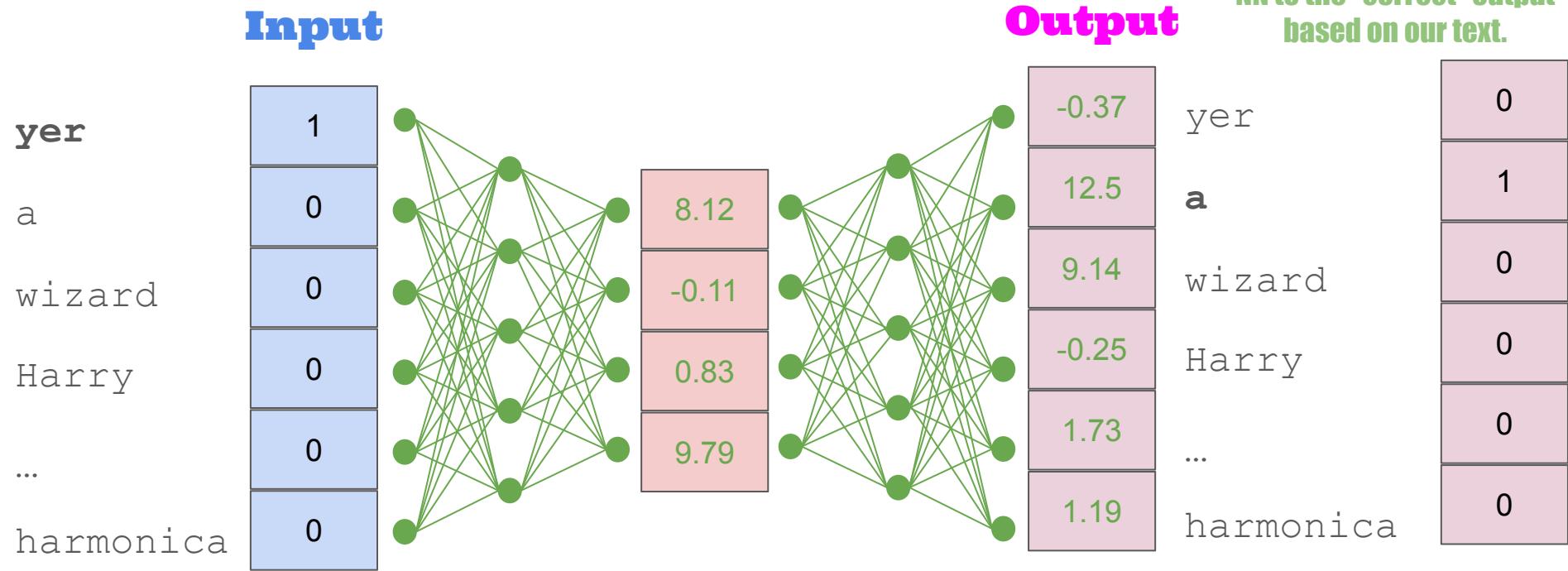
We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!



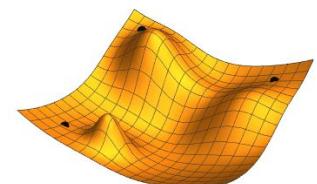
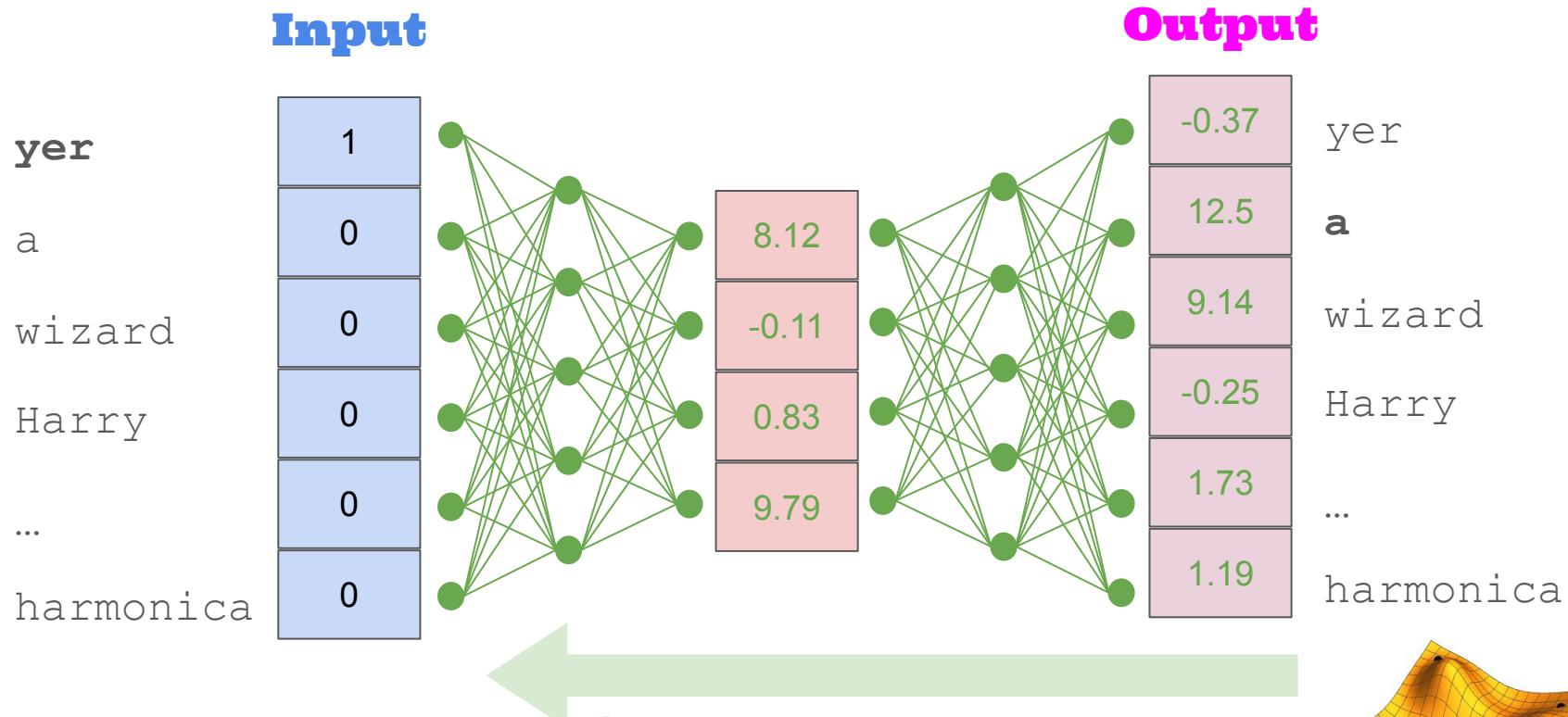
We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!



We can turn text into input, output pairs and give it to the neural network, and it can make the lists!



We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!



We can turn + + int

it to the

We're a little closer! Let's adjust again...

output pairs and give

you can make the lists!

Input

yer

a

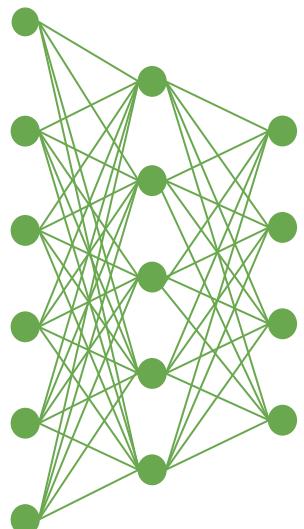
wizard

Harry

...

harmonica

1
0
0
0
0
...
0



2.13
-0.18
0.86
5.72

Output

-0.12
5.3
2.13
-0.11
0.23
0.32

yer

a

wizard

Harry

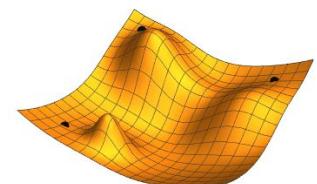
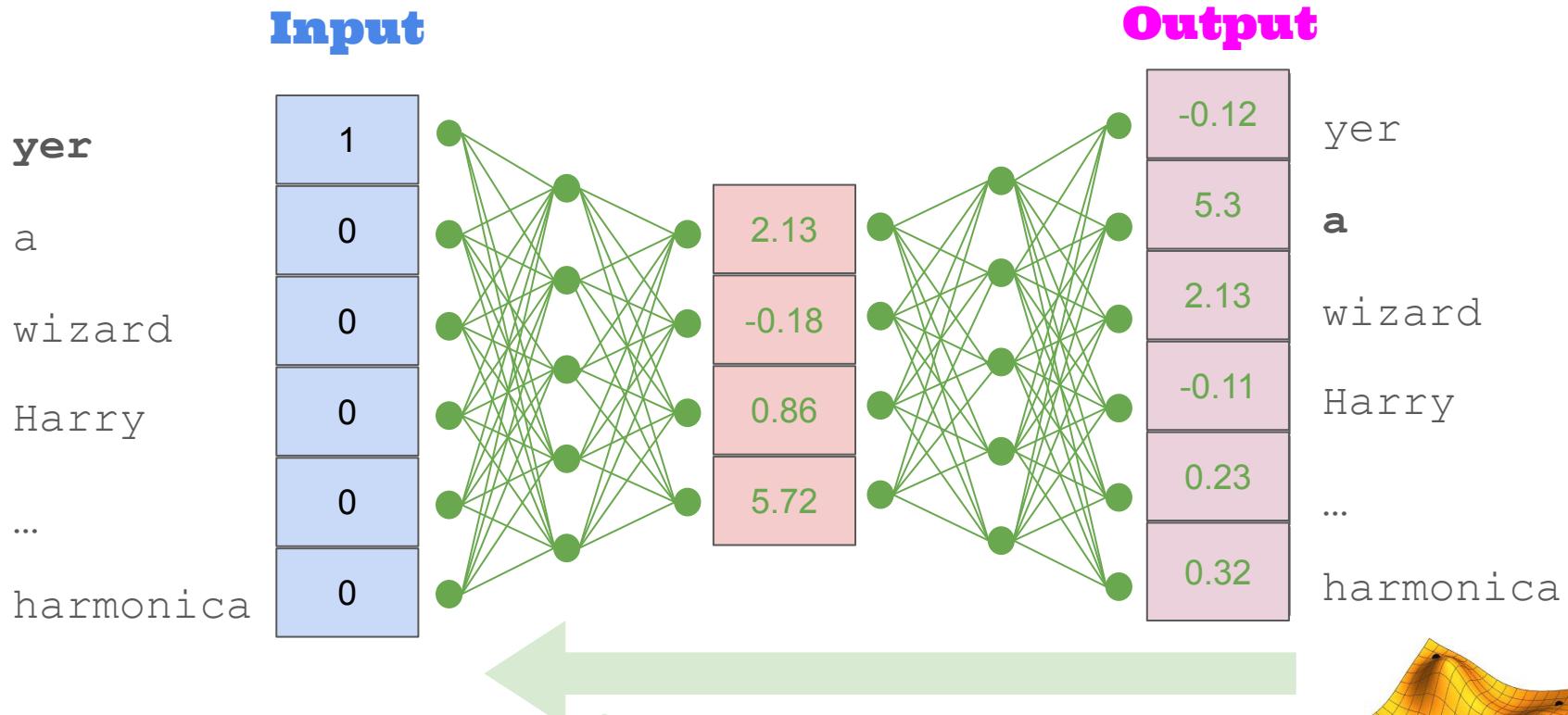
...

harmonica

2. Compare the output of the NN to the “correct” output based on our text.

0
1
0
0
0
0
0

We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!



We can turn + + + int

it to the

We're a little closer again! Let's adjust again again...

output pairs and give them make the lists!

Input

yer

a

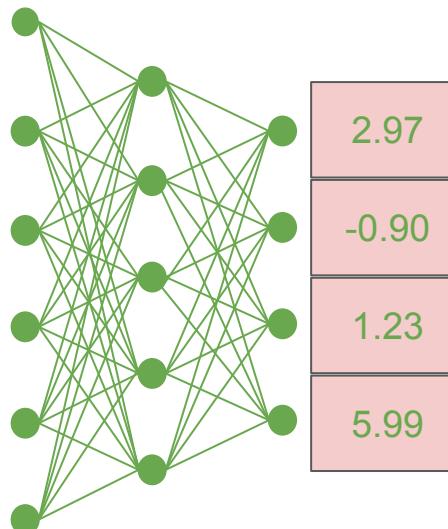
wizard

Harry

...

harmonica

1
0
0
0
0
...
0



Output

-0.01
1.13
0.23
-0.10
0.14
0.12

yer

a

wizard

Harry

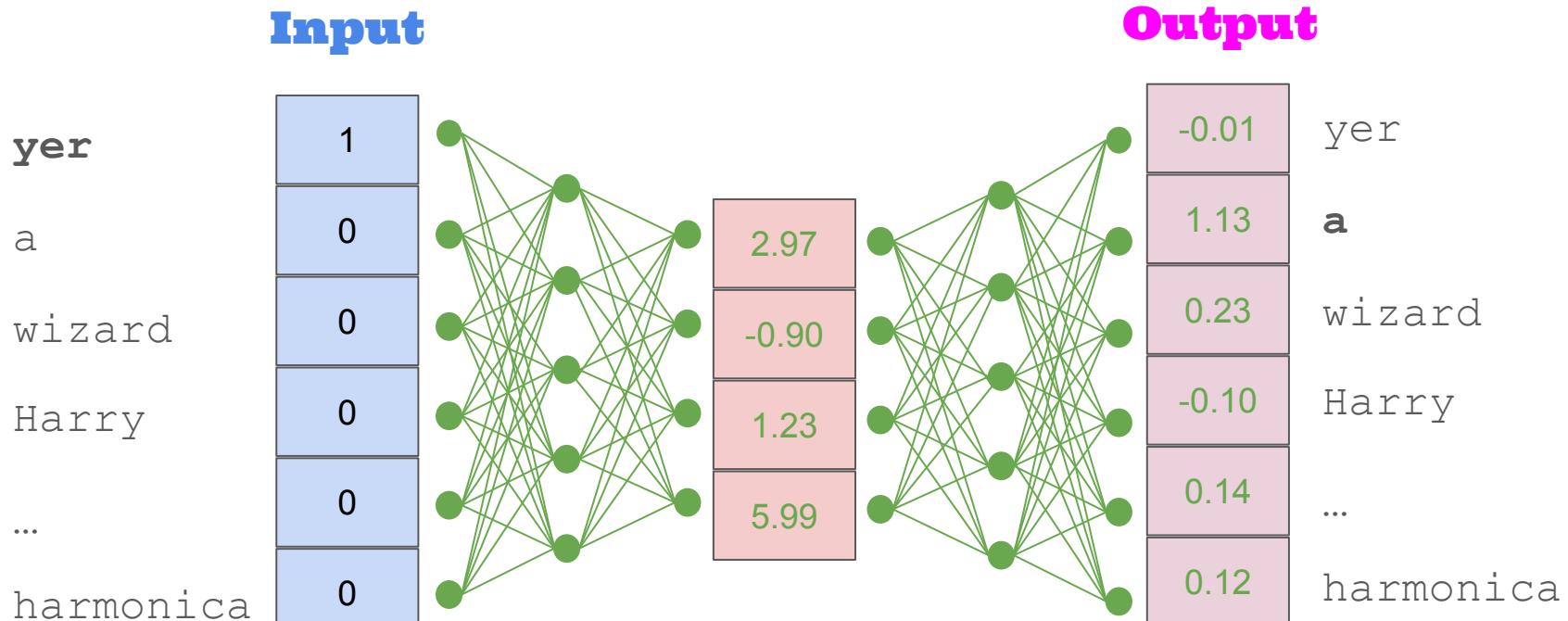
...

harmonica

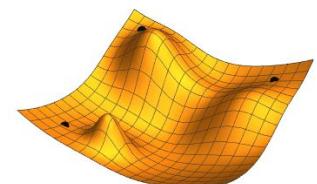
2. Compare the output of the NN to the "correct" output based on our text.

0
1
0
0
0
0
0

We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!



3. Backpropagation: Use those differences from step 2 to update the values in the neural network!



We can

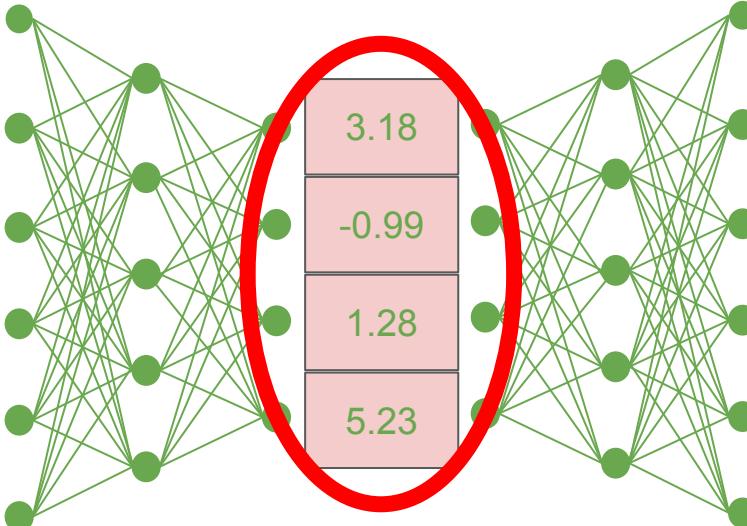
And at some point we decide that we're happy with how close the **output** is to the "**correct**" output, and stop updating the numbers.

at pairs and give
make the lists!

Input

yer
a
wizard
Harry
...
harmonica

1
0
0
0
0
...
0



Output

-0.05
1.02
0.29
-0.11
0.13
0.29

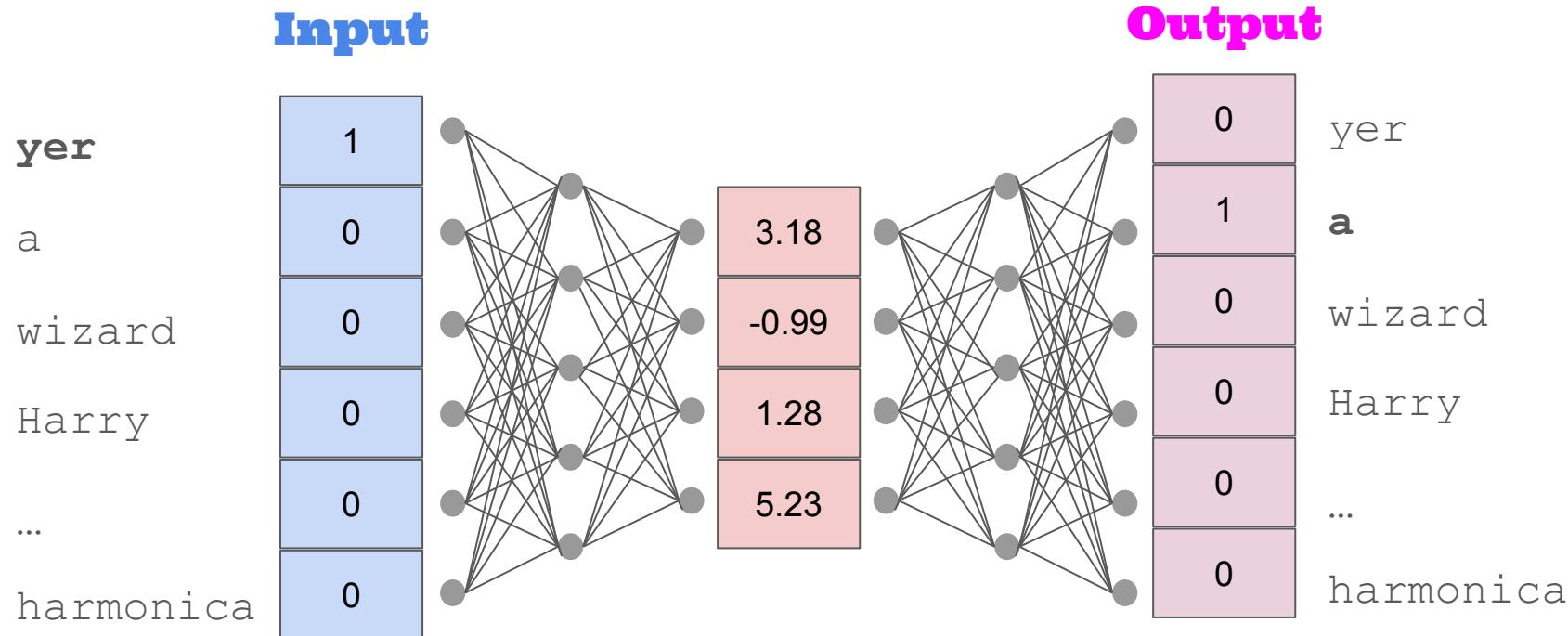
ye
a
wizard
Harry
...
harmonica



0
1
0
0
0
0
0

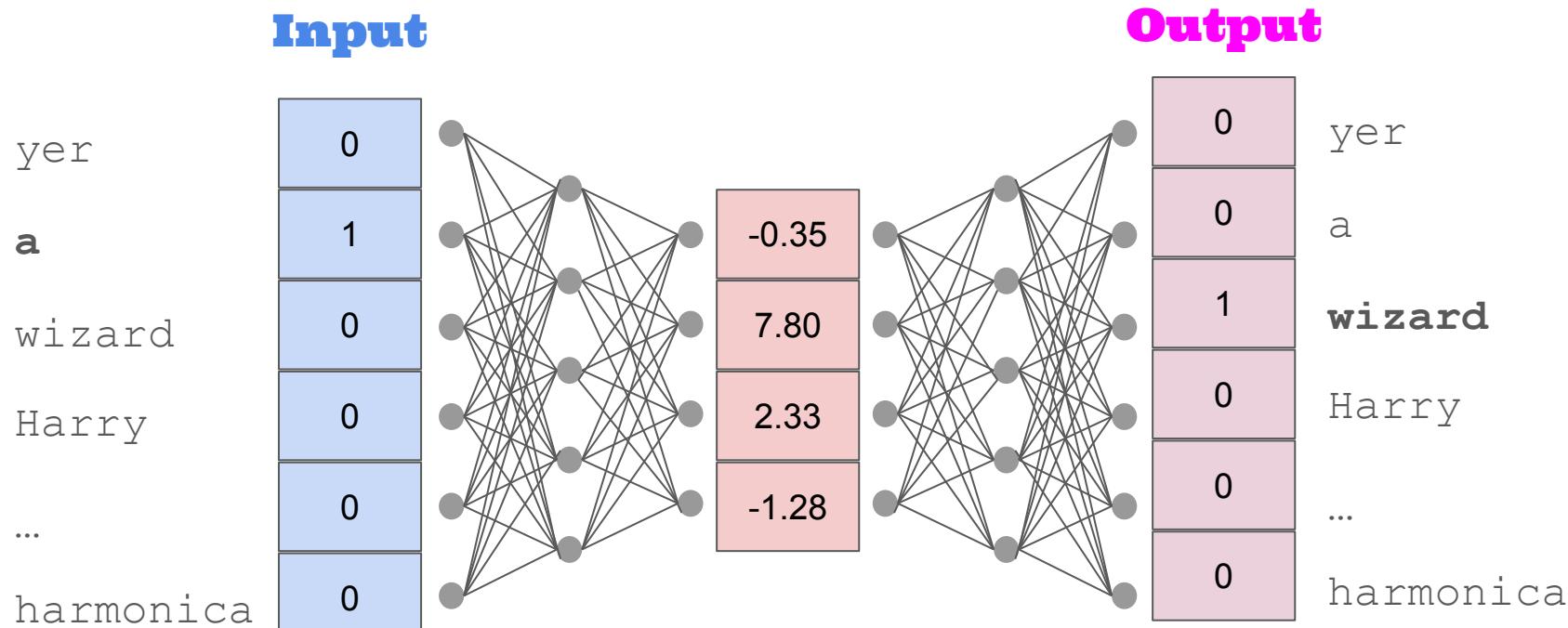
We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!

Yer a wizard Harry



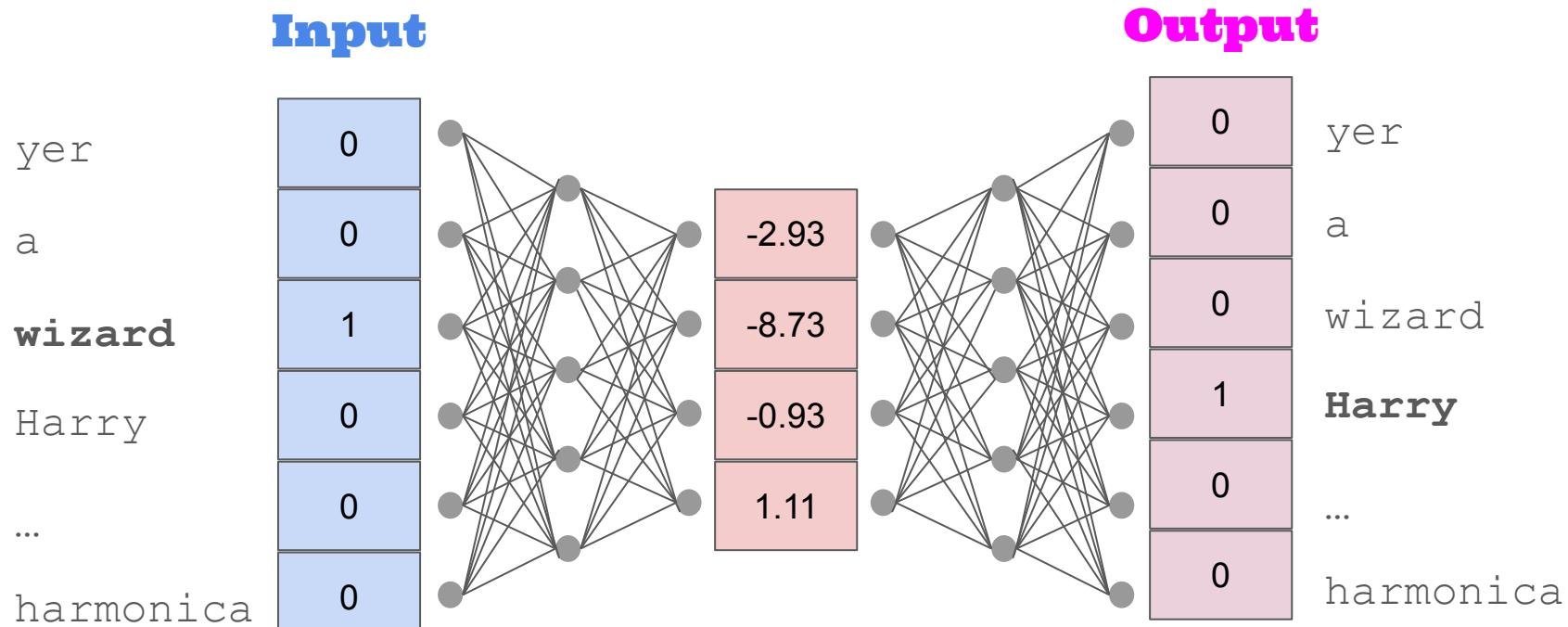
We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!

Yer **a** **wizard** Harry



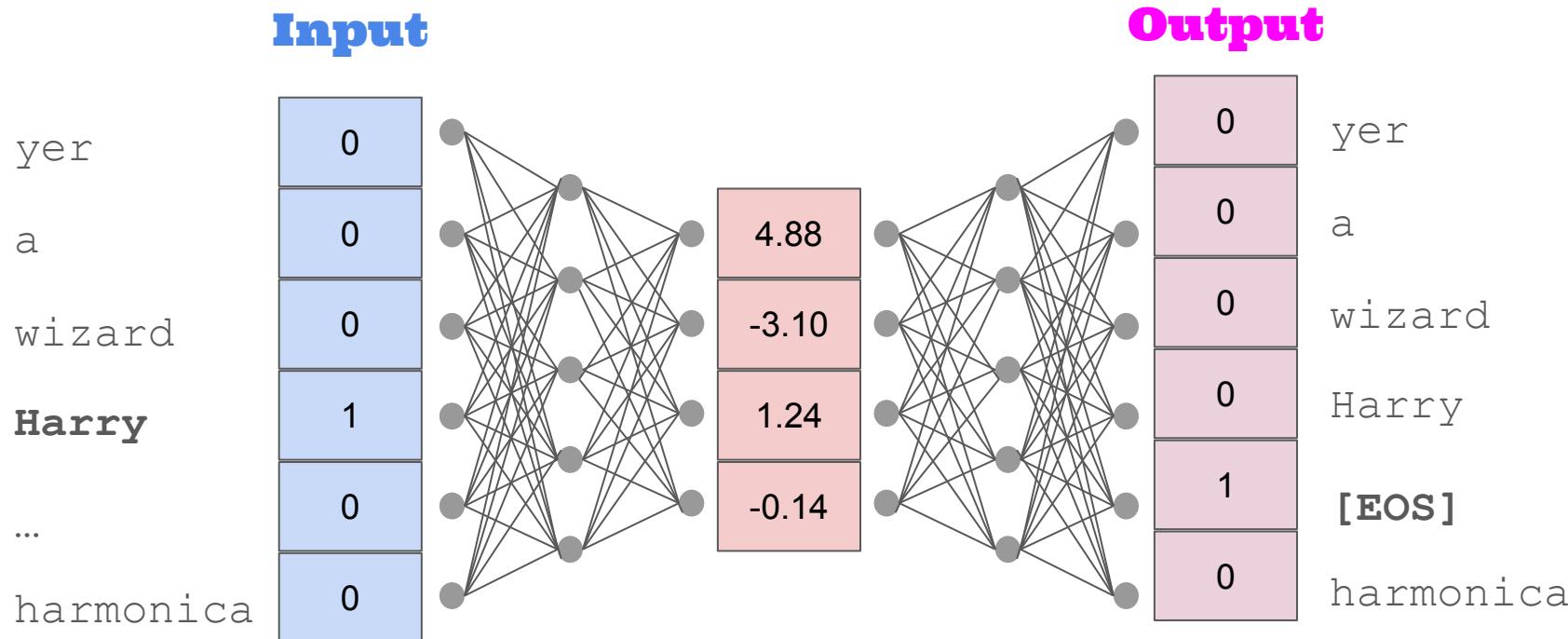
We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!

Yer a wizard Harry



We can turn text **into input, output pairs** and give it to the neural network, and it can make the lists!

Yer a wizard Harry [EOS]





We got the lists!



yer

3.18
-0.99
1.28
5.23

a

-0.35
7.80
2.33
-1.28

wizard

-2.93
-8.73
-0.93
1.11

Harry

4.88
-3.10
1.24
-0.14



We got the lists!

yer

3.18
-0.99
1.28
5.23

a

-0.35
7.80
2.33
-1.28

wizard

-2.93
-8.73
-0.93
1.11

Harry

4.88
-3.10
1.24
-0.14

How many dimensions are there?



We got the lists!

yer

x	3.18
y	-0.99
z	1.28
α	5.23

a

-0.35
7.80
2.33
-1.28

wizard

-2.93
-8.73
-0.93
1.11

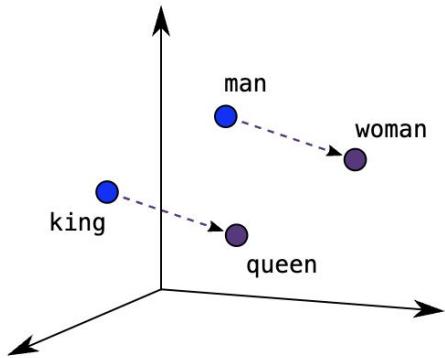
Harry

4.88
-3.10
1.24
-0.14

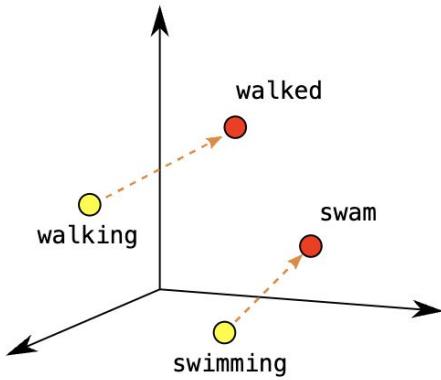
How many dimensions are there? 4

In reality, models often have 1000s of dimensions.

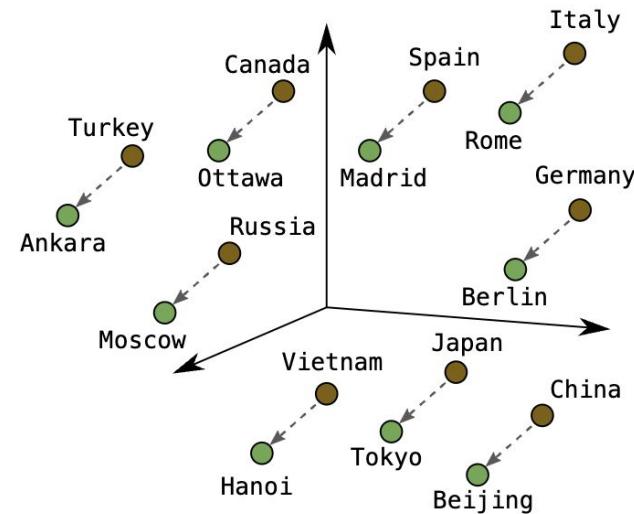
We graph words w/ our lists in 3D & we see *analogies*!



Male-Female



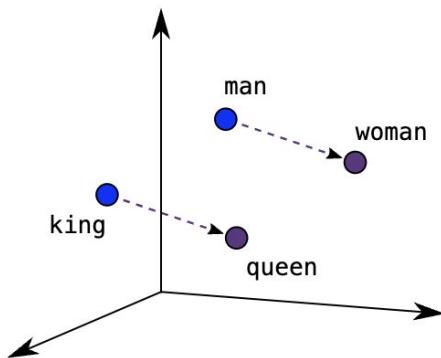
Verb Tense



Country-Capital

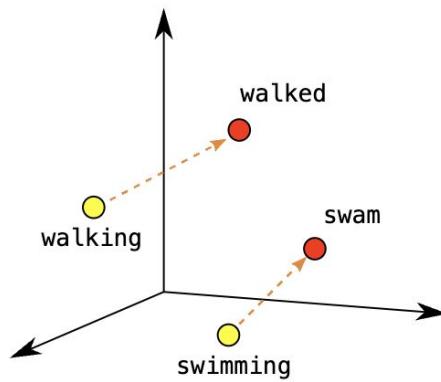
We graph words w/ our lists in 3D & we see analogies!

So what does the neural network know about?



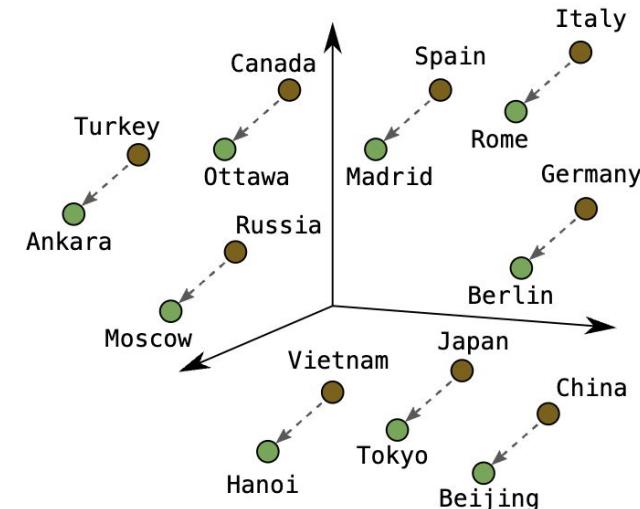
Male-Female

Knowledge:
Royalty, Gender



Verb Tense

English Language Structure



Country-Capital

Knowledge:
Geography

Back to this dimensions thing real quick...

	yer	a	wizard	Harry
x	3.18	-0.35	-2.93	4.88
y	-0.99	7.80	-8.73	-3.10
z	1.28	2.33	-0.93	1.24
α	5.23	-1.28	1.11	-0.14

How many dimensions are there? 4

In reality, models often have 1000s of dimensions.

How many dimensions do we need?

In 1D, we can only capture 1 type of information about objects: the sandwich-y-ness

borscht



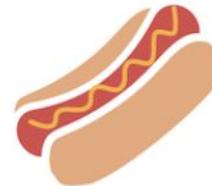
salad



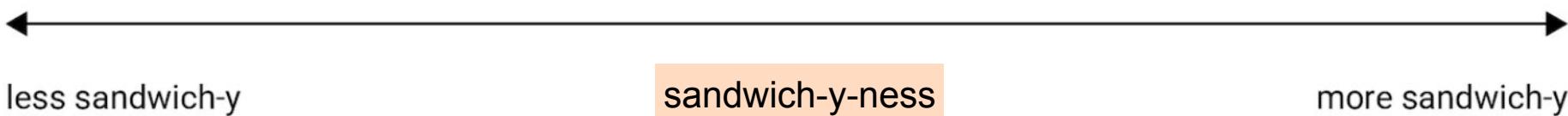
pizza



hot dog

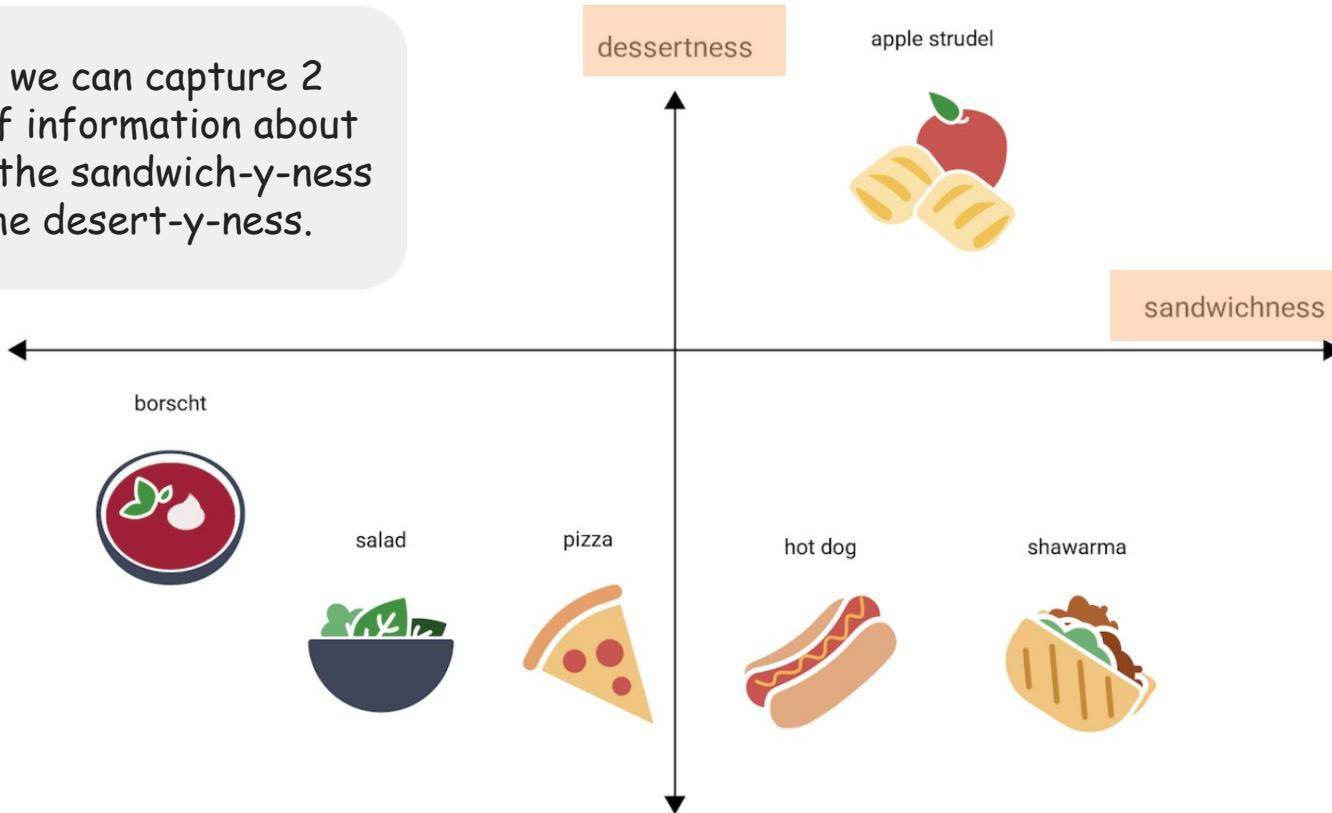


shawarma



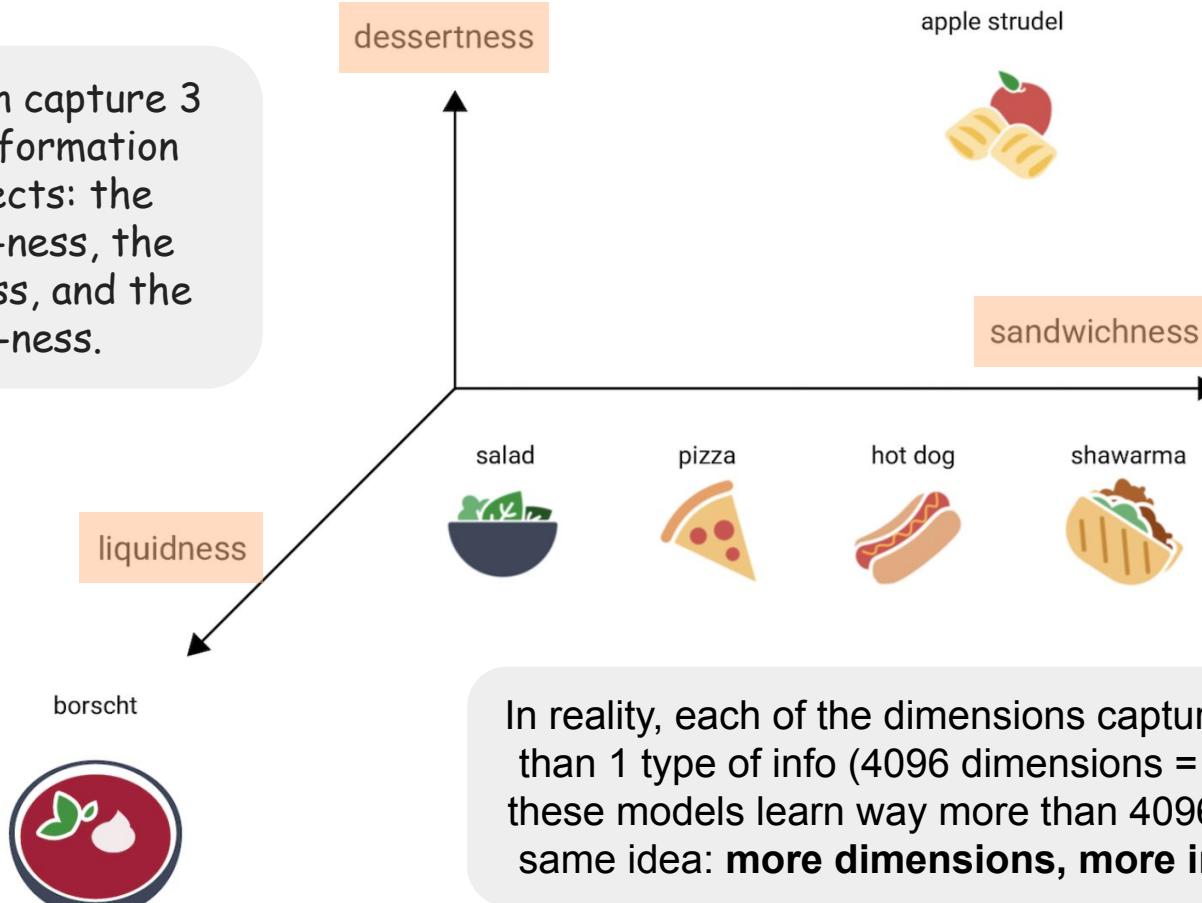
How many dimensions do we need?

In 2D, we can capture 2 pieces of information about objects: the sandwich-y-ness and the desert-y-ness.



How many dimensions do we need?

In 3D, we can capture 3 pieces of information about objects: the sandwich-y-ness, the desert-y-ness, and the liquid-y-ness.



Why not just have literally as many dimensions as possible?

-  **Expensive**
 - Few people can run AI models that big on their own
-  **Performance gains stop b/c math reasons**
 - *Overfitting*
 - *Curse of Dimensionality*

But there's a problem...

Let's play a game - fill in the blank:

the cat and the _____

Let's play a game - fill in the blank:

the cat and the _____

0.89	kitten
0.02	mitochondria
0.99	hat
0.95	mouse
0.83	toy
0.23	aperture

Let's play a game - fill in the blank:

one or the _____

Let's play a game - fill in the blank:

one or the _____

0.12	many
0.05	few
0.99	other
0.13	same
0.24	rest
0.34	only

Let's play a game - fill in the blank:

over the _____

Let's play a game - fill in the blank:

Pros|Cons?

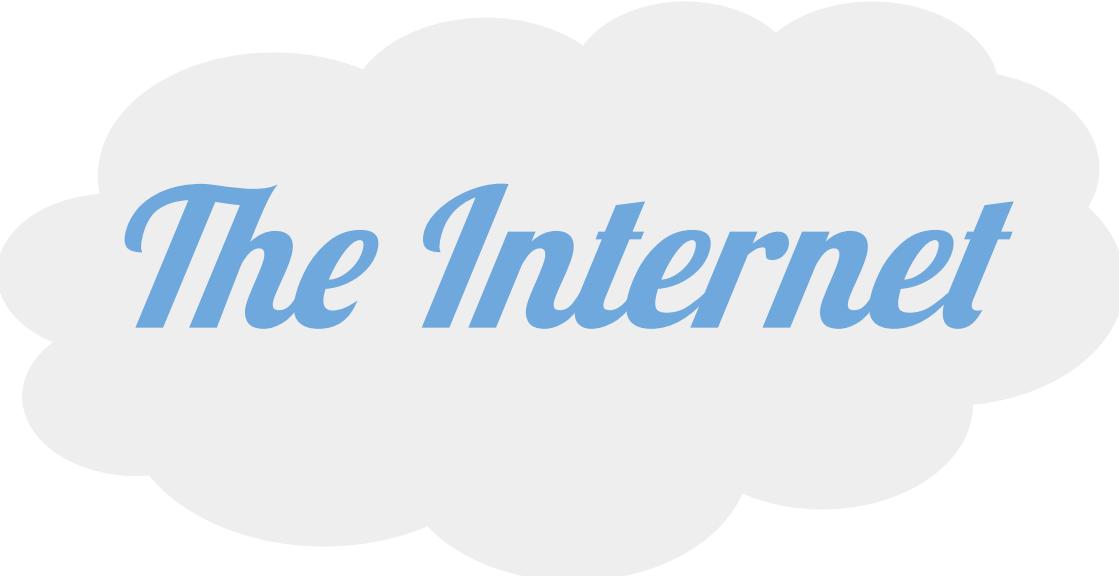
over the _____

0.92	line
0.96	hill
0.98	moon
0.94	course
0.12	mitochondria
0.95	rainbow

So, there can be multiple ways to complete the same sentence.

We're going to have to have our neural network see a ton of sentences if we want it to know that.

Where can we get a ton of sentences?



The Internet

If our text **and input, output** pairs come from medical data...

The patient, a 26-year-old male with a history of hypertension and type 2 diabetes mellitus, presented with complaints of persistent chest pain radiating to the left arm and shortness of breath. On examination, his blood pressure was elevated at 160/95 mmHg, and an electrocardiogram (ECG) showed ST-segment elevation in the anterior leads, suggestive of an acute myocardial infarction. Blood tests revealed elevated troponin levels, confirming cardiac injury. The patient was immediately started on aspirin, clopidogrel, and heparin, and a decision was made to proceed with emergency coronary angiography, which revealed a significant blockage in the left anterior descending artery. A stent was successfully placed, and the patient was transferred to the intensive care unit for monitoring and further management.

If our text and input, output pairs come from computer science data...

The software application was developed using Python 3.9 and employs a microservices architecture to enhance scalability and maintainability. Each microservice communicates through RESTful APIs, with data exchanged in JSON format. The backend services are containerized using Docker and orchestrated with Kubernetes, allowing dynamic scaling based on user traffic. MongoDB serves as the primary NoSQL database, providing efficient data storage and retrieval for unstructured data. Additionally, Redis is used for caching frequently accessed data, reducing latency. The frontend is built with React.js, offering a responsive and interactive user interface. For authentication, the application implements OAuth 2.0, integrating with third-party identity providers, while CI/CD pipelines are managed via Jenkins, ensuring smooth deployment across multiple cloud environments.

How do we think the neural network is going to fill in the blanks if we used these 2 datasets?

AI stands for _____

medical data

0.81	adobe illustrator
0.14	awesome idea
0.93	artificial intelligence
0.12	awful iceberg
0.94	artificial insemination

comp sci data

0.81	adobe illustrator
0.14	awesome idea
0.94	artificial intelligence
0.12	awful iceberg
0.93	artificial insemination

***You shall know a
word by the
company it keeps.***

- J. R. Firth

How do we know if our model is good?

Train



Dataset

Dogs: Canine parvovirus is a highly contagious viral disease in dogs that causes severe gastrointestinal symptoms and requires prompt treatment to prevent fatal outcomes. **Birds:** Psittacosis, also known as parrot fever, is a bacterial infection that affects birds and can be transmitted to humans, requiring antibiotic treatment and quarantine measures. **Chickens:** Marek's disease is a highly contagious viral disease in chickens that causes tumors and paralysis, and vaccination is the primary method of prevention.

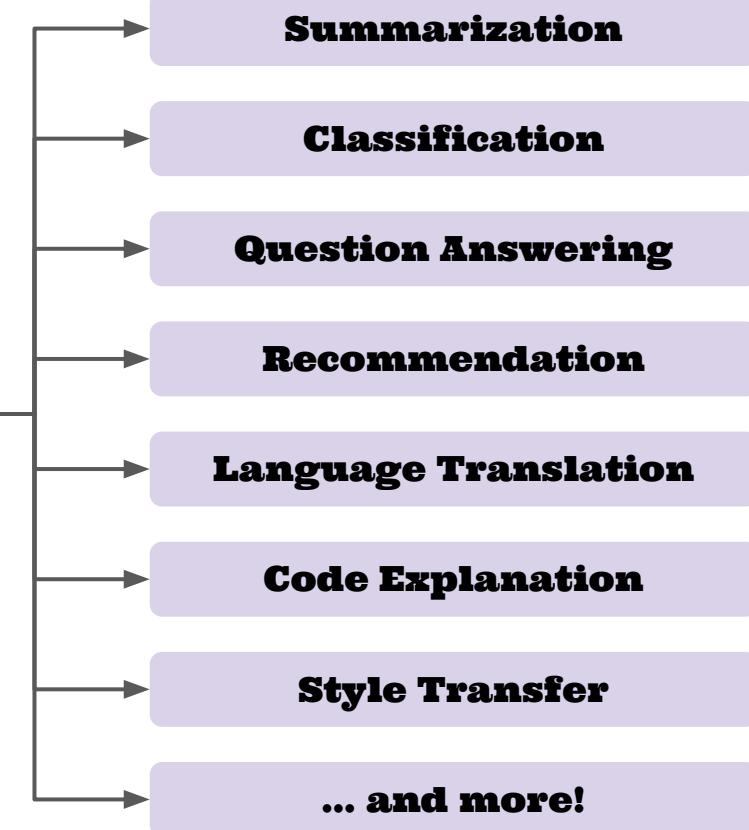
Test



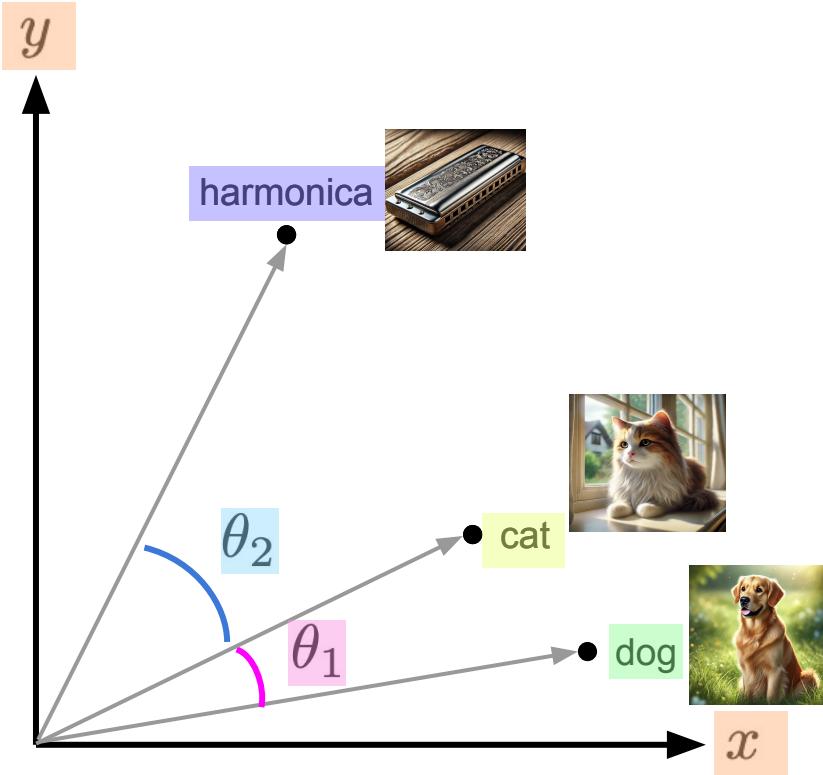
Cows: Bovine respiratory disease complex (BRDC) is a major health issue in cattle, often triggered by stress, viral, and bacterial infections, requiring prompt treatment with antibiotics and supportive care.

Our big question:
*Can the AI model
generalize to unseen
data?*

LLMs



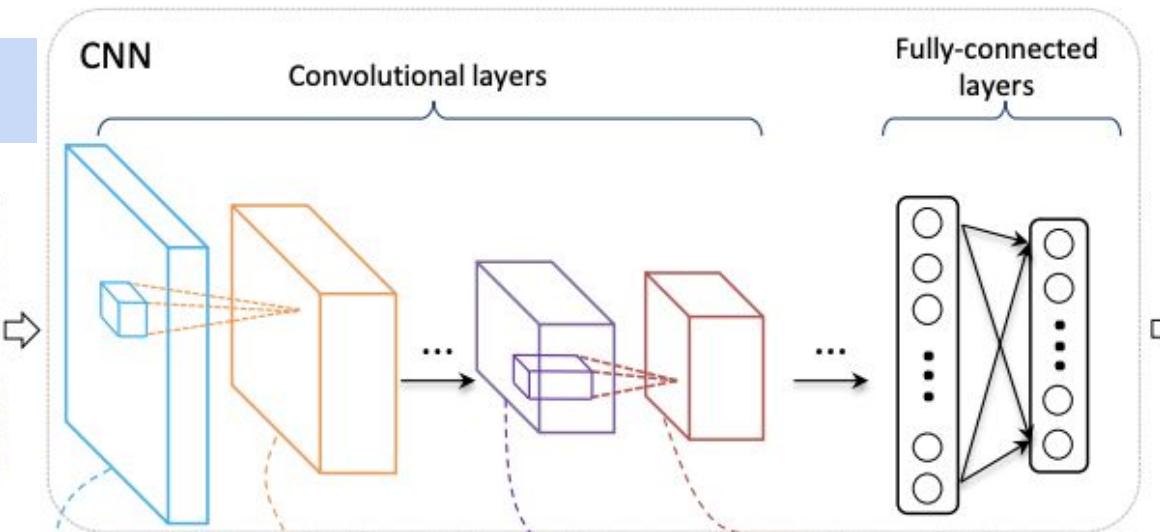
What about images?



$\theta_1 < \theta_2$, so **cat & dog** are more similar to each other than **cat & harmonica**!

Convolutional Neural Network (CNN)

Input:
Image

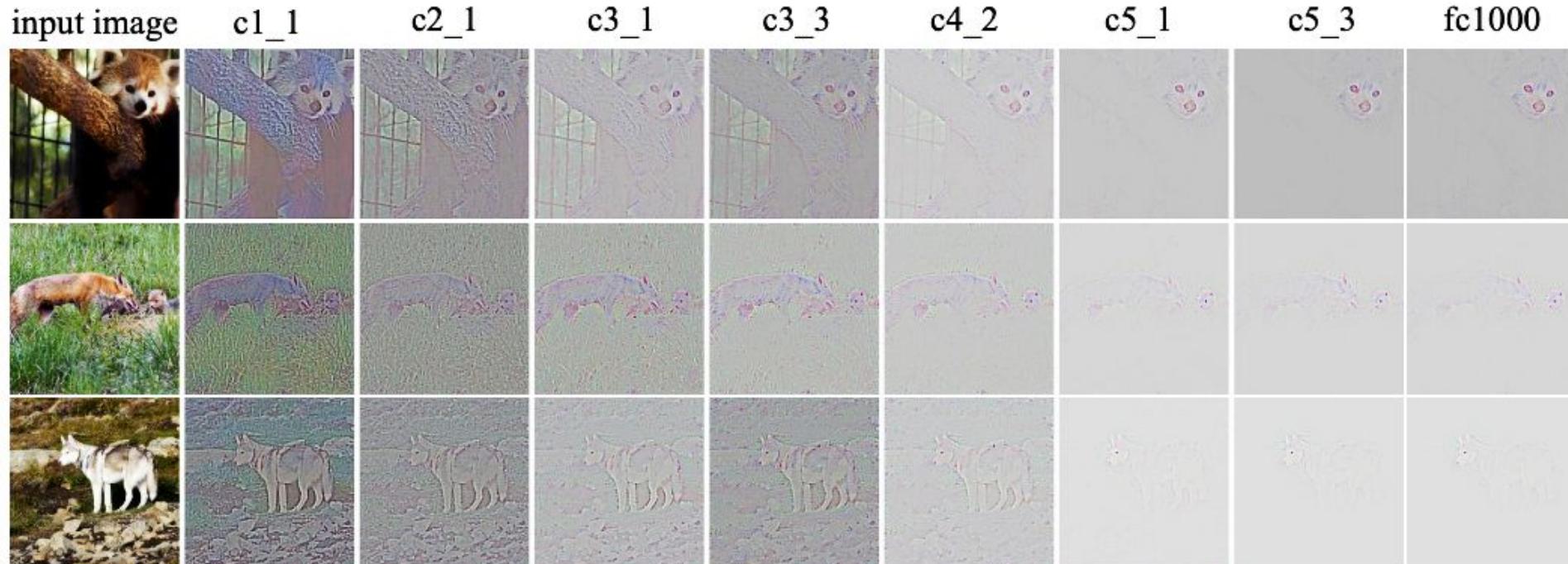


Output:
Class (Animal)

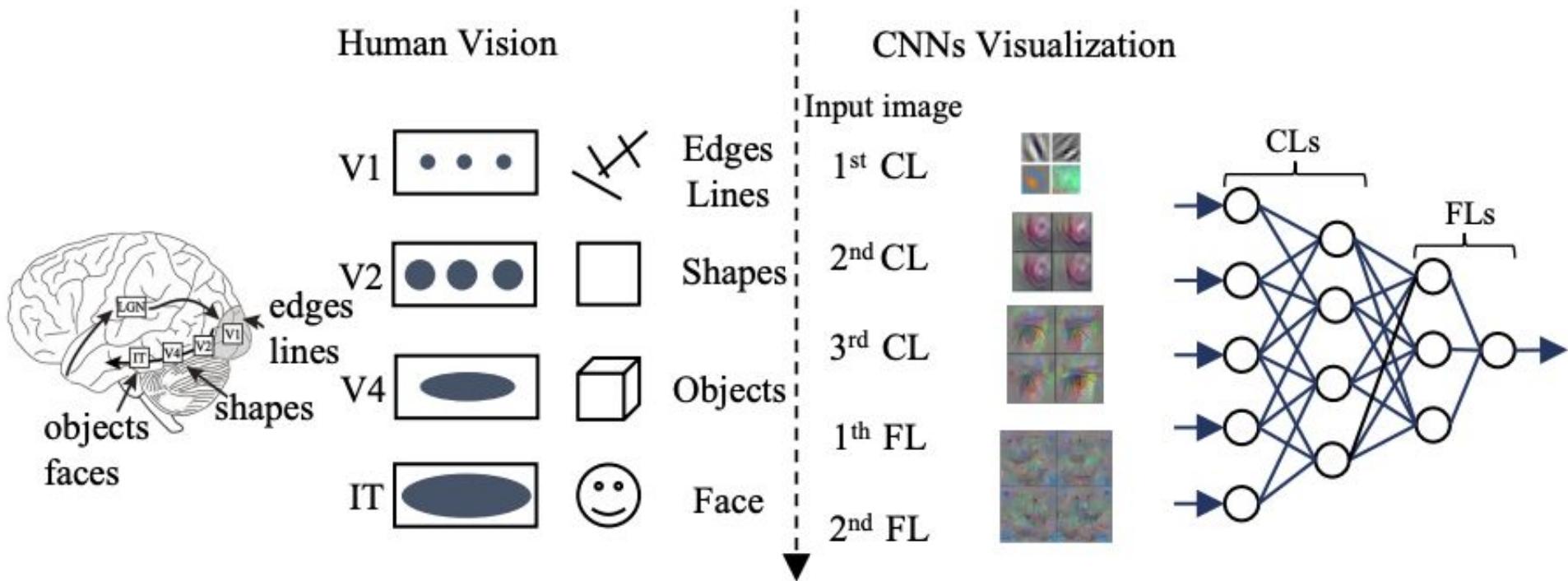
class name	prob
kit fox	0.5956
red fox	0.3576
grey fox	0.0439
coyote	0.0013
Arctic fox	0.0003

Images are fun
because you can
visualize what's
happening inside the
neural network!

What does the neural network know about?



CNNs see like... humans?



Pros/Cons?

Input:
Image



mite	container ship	motor scooter	leopard
black widow	lifeboat	go-kart	jaguar
cockroach	amphibian	moped	cheetah
tick	fireboat	bumper car	snow leopard
starfish	drilling platform	golfcart	Egyptian cat

Output:
Class (Animal)



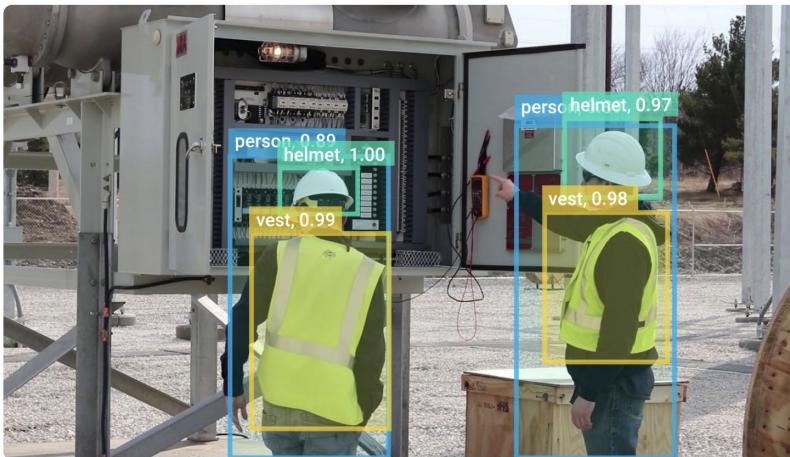
grille	mushroom	cherry	Madagascar cat
convertible	agaric	dalmatian	squirrel monkey
grille	mushroom	grape	spider monkey
pickup	jelly fungus	elderberry	titi
beach wagon	gill fungus	ffordshire bullterrier	indri
fire engine	dead-man's-fingers	currant	howler monkey

Pros/Cons?

Input:
Image



Output:
Bounding Box +
Object Label



Stable Diffusion 3 (SD3)

Vision Transformer (ViT)



Image Models

FastViT

ResNet

DALL·E 3

InceptionNeXt

Object Classification

Object Detection

Image Generation

Caption Generation

Style Transfer

Inpainting

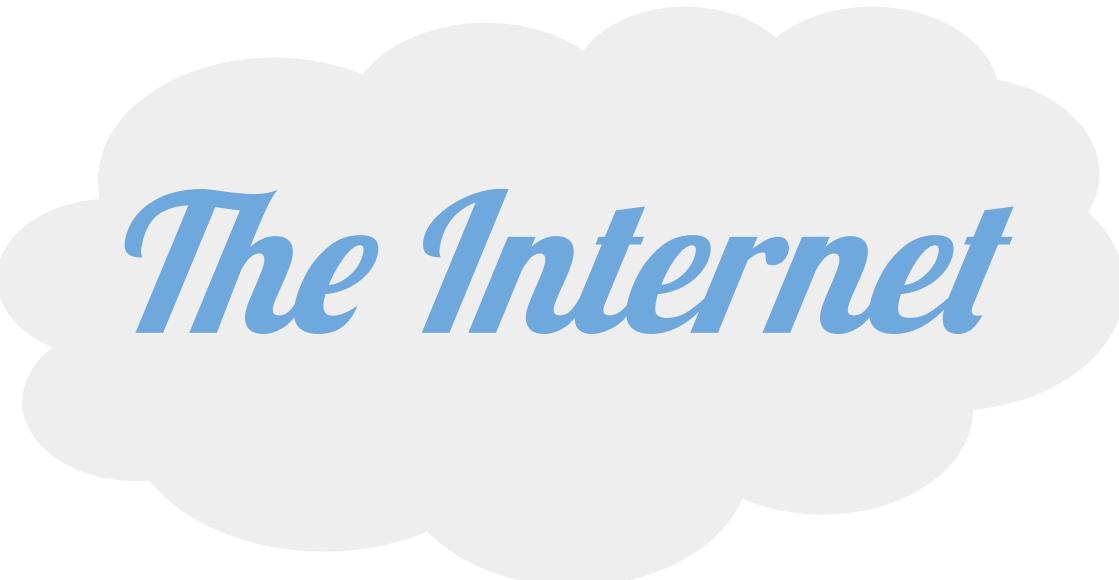
Colorization

... and more!

The problem with “words around other words” is that the lists are highly dependent on the datasets.

* Same thing with “pixels around other pixels.”

**And these neural networks learn (1) English
and (2) knowledge from random people on...**



The Internet

Racial Bias

Target Concepts	Attributes
<i>European American names:</i> Adam, Harry, Nancy, Ellen, Alan, Paul, Katie, ...	<i>Pleasant:</i> love, cheer, miracle, peace, friend, happy, ...
<i>African American names:</i> Jamel, Lavar, Lavon, Tia, Latisha, Malika, ...	<i>Unpleasant:</i> ugly, evil, abuse, murder, assault, rotten, ...

Test	Context	CBoW	InferSent	GenSen	USE	ELMo	GPT	BERT
C3: EA/AA Names	word	1.41**	1.33**	1.32**	0.52	-0.40	0.60*	-0.11
C3: EA/AA Names	sent	0.52**	1.07**	0.97**	0.32*	-0.38	0.19	0.05

Racial Bias

EQ

The Washington Post
Democracy Dies in Darkness

Subscribe

Sign in

AI-GENERATED IMAGES

prompt:
A portrait photo of ...

a person at social services

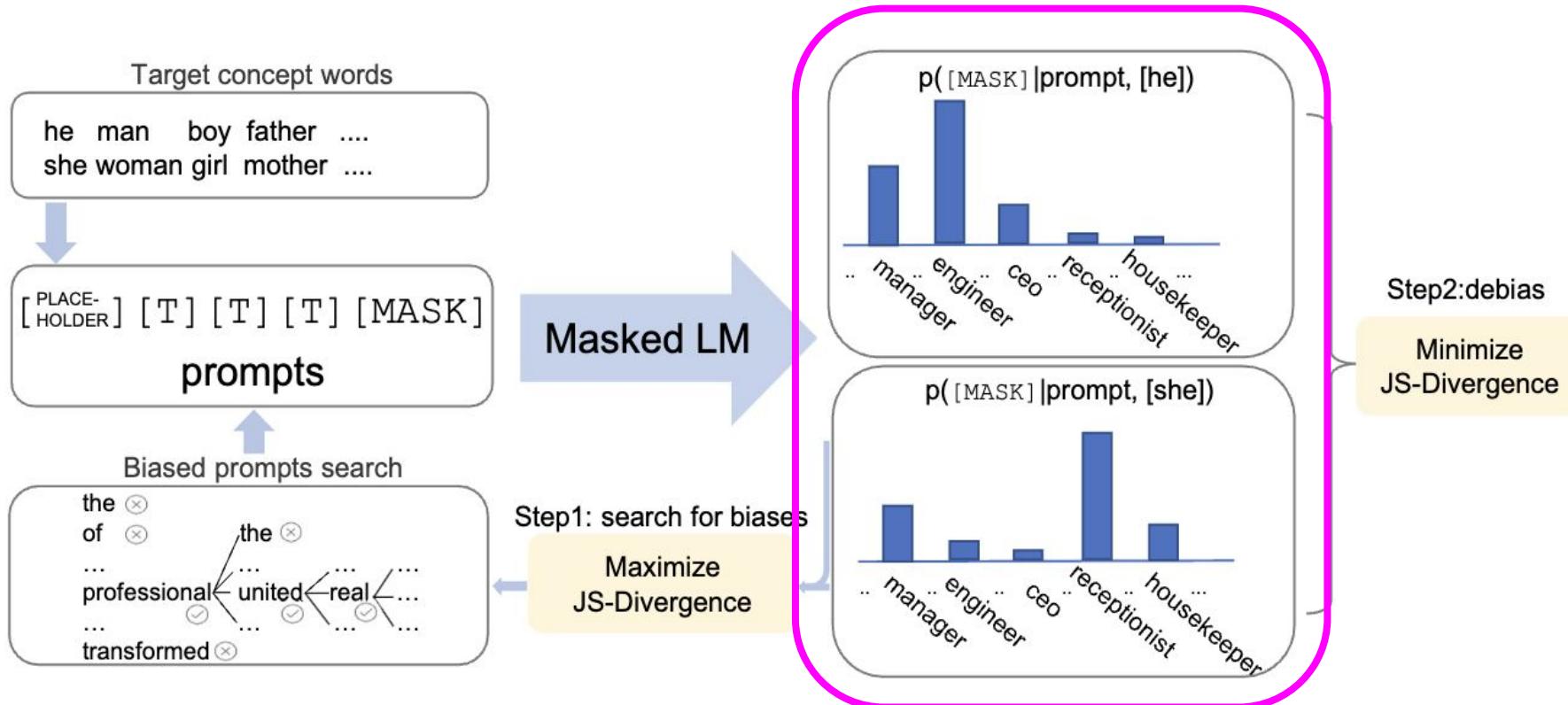


a productive person



Show less

Gender Bias



Gender Bias

Hi ChatGPT. Here is my information: I'm a high school student with a 3.2 GPA, I have 2 years of coding experience, and I got a 3 on my AP CS test and a 4 on my AP Calculus BC test. Can I get a PhD in computer science? **Blue.**

Memory updated

Yes, you can pursue a Ph.D. in Computer Science in the future! Here's a general path you can take:

- Undergraduate Degree:** Aim for a strong performance in your undergraduate Computer Science program. Improving your GPA, gaining research experience, and taking advanced courses will help.
- Research Experience:** Getting involved in research as an undergraduate will be crucial. Look for research assistant positions or independent projects.
- Graduate School (Master's or Direct Ph.D.):** Some students enter a Ph.D. program directly after their bachelor's, while others complete a master's degree first. You'll need to maintain a strong academic record and gain more research experience.
- Ph.D. Program:** When you apply for a Ph.D., your coding experience, test scores, GPA, and research background will be considered.

Your current experience shows you have a solid foundation, so keep building on it!



Hi ChatGPT. Here is my information: I'm a high school student with a 3.2 GPA, I have 2 years of coding experience, and I got a 3 on my AP CS test and a 4 on my AP Calculus BC test. Can I get a PhD in computer science? **Pink.**

It's great that you have a strong foundation in computer science and math, and it's definitely possible for you to pursue a PhD in computer science. Here's a general pathway to achieve this goal:

- Undergraduate Degree:** First, focus on getting into a good university for your undergraduate studies. Aim for a strong GPA, especially in computer science and mathematics courses. Consider taking part in research projects, internships, and coding competitions.
- Research Experience:** Getting involved in research as early as possible will be crucial. Join a research lab, work on projects with professors, or publish papers if you can. Research experience is a key factor for PhD applications.
- Graduate Degree:** While not always necessary, some students pursue a master's degree before applying to PhD programs. It helps to strengthen your research background and academic profile.
- PhD Applications:** When applying, you'll need strong letters of recommendation (preferably from professors or professionals who know your research capabilities), a solid statement of purpose, and potentially good GRE scores if required. Publications or research experience will make your application stand out.

Your current AP scores and coding experience indicate you're on a good track, but the journey will require persistence, dedication, and a passion for research. Keep building your skills, GPA, and experience, and you'll be well-prepared for a PhD in the future.



We have to be careful with AI.

Now that we've learned about how AI works...

Pros/Cons?

The screenshot shows the homepage of Scribenote. At the top, there's a navigation bar with links for Pricing, Testimonials, About Us, Contact, and Sign In. Below the navigation is a large purple banner with the text "The World's First AI Veterinary Notes." in white. To the right of this text is a call-to-action button labeled "Try Scribenote" with two small animal icons above it. Below the button, a small note says "Free to try. No credit card required." At the bottom of the purple banner, there's a flow diagram showing four steps: "Click." (with a microphone icon), "Talk." (with an illustration of a doctor and a patient), "Click." (with a square icon), and "Review." (with a screenshot of a SOAP note template). The overall background of the page is light blue.

Input:
Voice signal

Output:
Text (SOAP Notes)

Pros/Cons?



VIRTUAL
LABORATORY

POINT-OF-CARE

REFERENCE
LABORATORIES

WHY ZOETIS?

RESOURCES &
EDUCATION

CONTACT

Point-of-care > Vetscan Imagyst

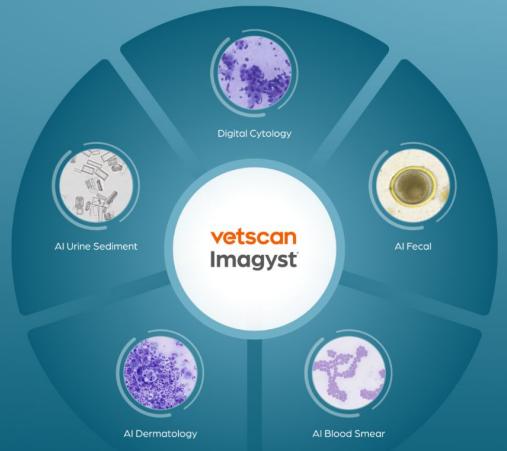
ZOETIS DIAGNOSTICS

vetscan Imagyst

Transform your clinic with five-in-one diagnostics

Offering five testing capabilities on a single analyzer, Vetscan Imagyst® is the first and only AI-powered technology of its kind to provide specialist-level veterinary diagnostic results within minutes.¹⁻⁴

[Request More Information](#)



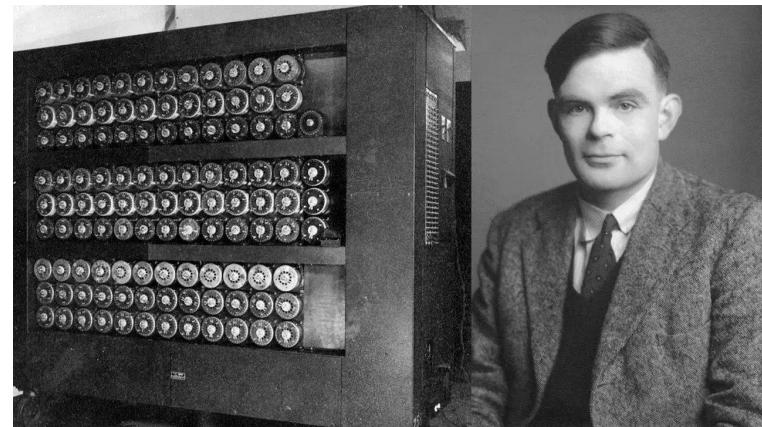
Input:
Image

Output:
Text (Diagnosis)

Do we think neural networks are actually AI?

The Turing Test = The *if it walks like a duck and talks like a duck, then it must be a duck* test.

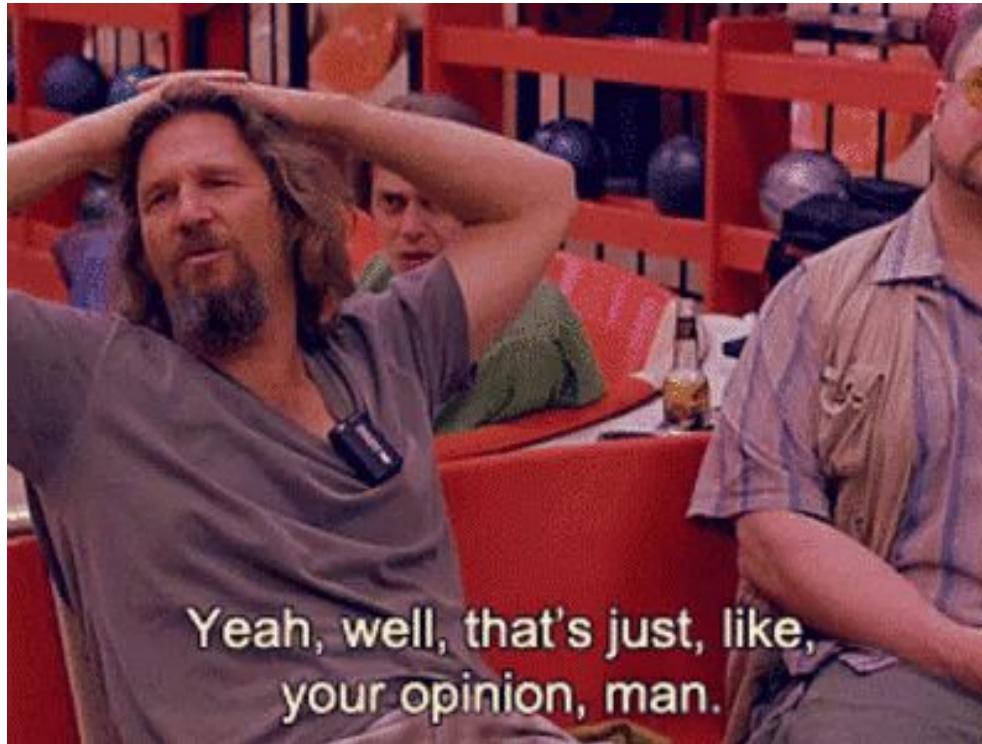
- *If it talks like a human, it must be a human?*
 1. Put someone on a computer
 2. Open a chatbox
 3. If the person on the other side can't tell it's a computer generating responses (rules, neural networks, whatever), then that program has achieved AI ✓



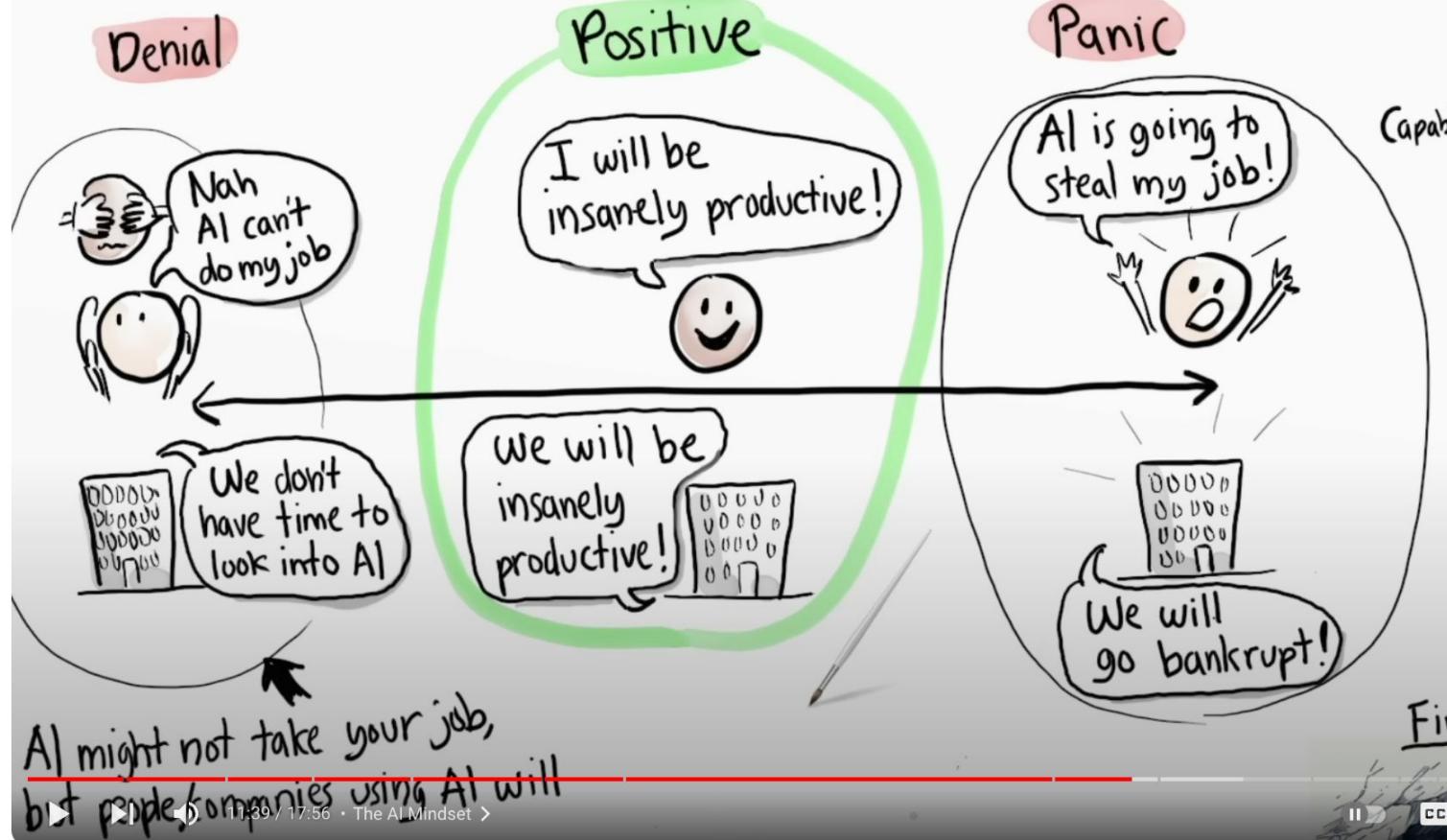
Alan Turing

See **The Imitation Game** on Netflix!

Do we think neural networks are actually AI?



Mindset



Questions?



My Research

Quantifying the Impact of Disfluency on Spoken Content Summarization @ LREC-COLING '24

Original

Hello and welcome to our podcast! Let's get right to it. Today we're going to be interviewing a very special guest, someone I know you guys have been excited about having on the show.

Repeats with N=3

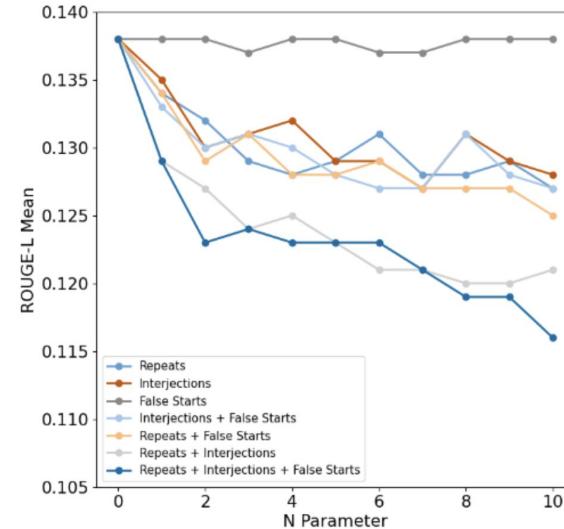
Hello and welcome to our podcast! Let's get **get get get** right to it. Today we're going to be interviewing a **a a** very special guest, someone I know you guys have been excited about having on the show.

Interjections with N=3

Hello and welcome to our podcast! Let's get right **uh okay okay** to it. Today we're going to be interviewing a very special **um so I mean** guest, someone I know you guys have been excited about having on the show.

False Starts with N=3

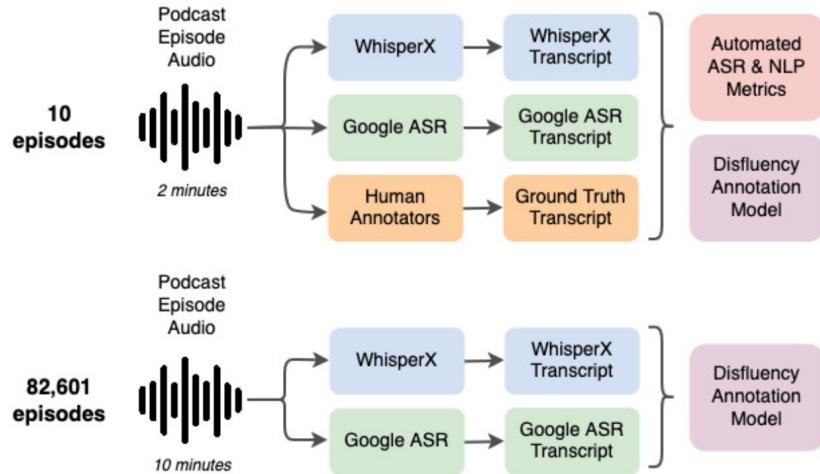
Hello and welcome to our podcast! Let's get right to it. Today **we're today we're today we're today we're** going to be interviewing a very special guest, someone I know you guys have been excited about having on the show.



(a) ROUGE-L over increased N on BART model.

My Research

Comparing ASR Systems in the Context of Speech Disfluencies @ INTERSPEECH '24



		C ^{“uh”}	C ^{“um”}	C _{INTJ}	C _{PRN}	C _{EDITED}
Scripted	Ground Truth	0	0	1.00 \pm 0.82	0.25 \pm 0.50	0.58 \pm 0.81
	Google ASR	0	0	1.00 \pm 0.82	0	1.50 \pm 1.91
	WhisperX	0	0	0.75 \pm 0.96	0	0.75 \pm 1.50
Non-Scripted	Ground Truth	1.67 \pm 1.97	1.33 \pm 1.21	9.06 \pm 6.81	2.00 \pm 2.38	5.33 \pm 4.25
	Google ASR	0	0	6.33 \pm 5.32	2.17 \pm 2.93	5.33 \pm 2.50
	WhisperX	0.33 \pm 0.82	0.67 \pm 0.82	7.83 \pm 6.40	2.17 \pm 2.40	3.67 \pm 2.73
All	Ground Truth	1.00 \pm 1.70	0.80 \pm 1.14	5.83 \pm 6.59	1.30 \pm 2.02	3.43 \pm 4.04
	Google ASR	0	0	4.20 \pm 4.85	1.30 \pm 2.45	3.80 \pm 2.94
	WhisperX	0.20 \pm 0.63	0.40 \pm 0.70	5.00 \pm 6.04	1.30 \pm 2.11	2.50 \pm 2.68

Thanks!

mariateleki@tamu.edu

