

Deep learning

UNDERSTANDING MACHINE LEARNING

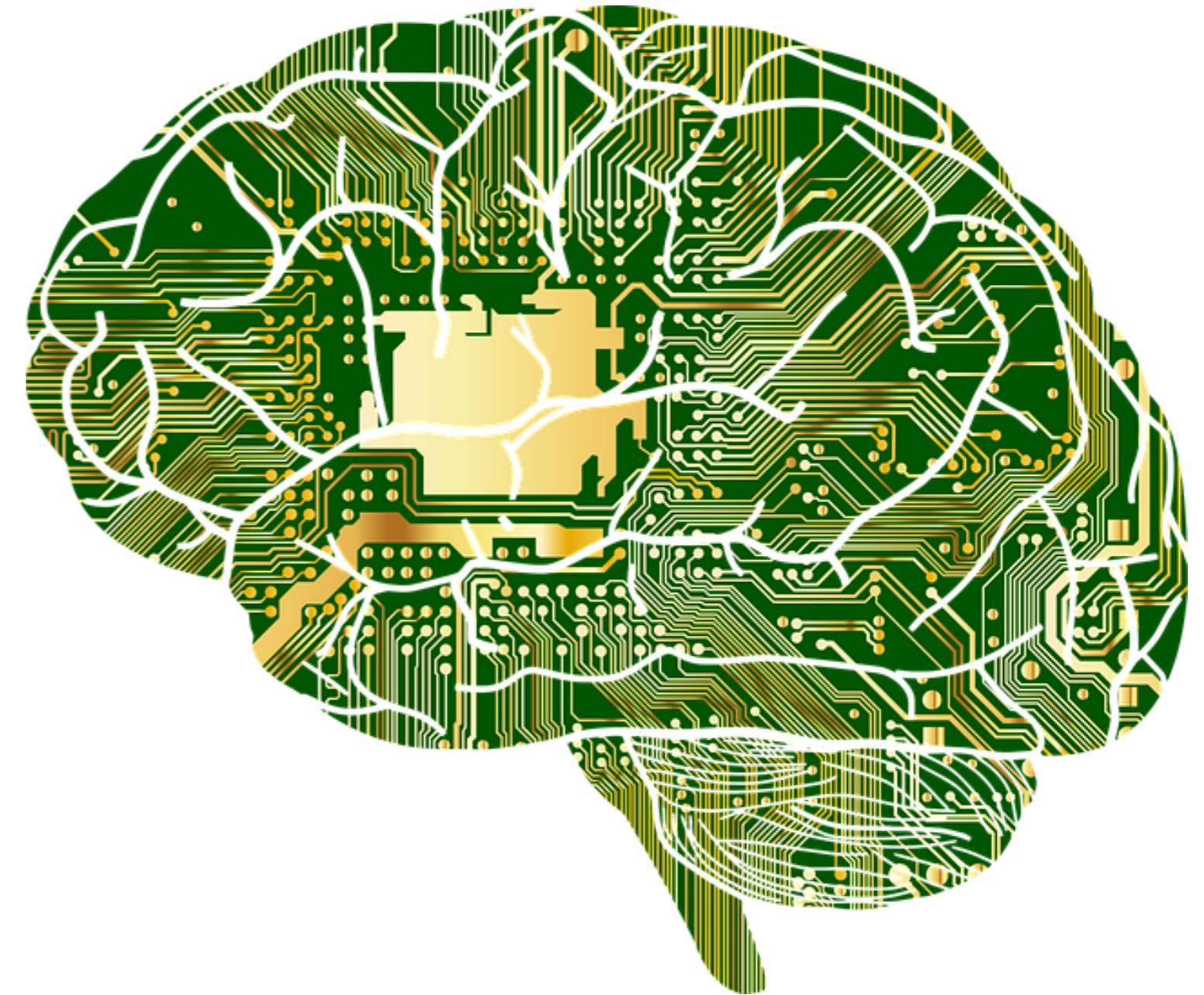


Sara Billen

Curriculum Manager, DataCamp

What is deep learning?

- AKA: Neural Networks
 - Basic unit: neurons (nodes)
- Special area of Machine Learning
- Requires more data
- Best when inputs are images or text



Predicting box office revenue



Predicting box office revenue

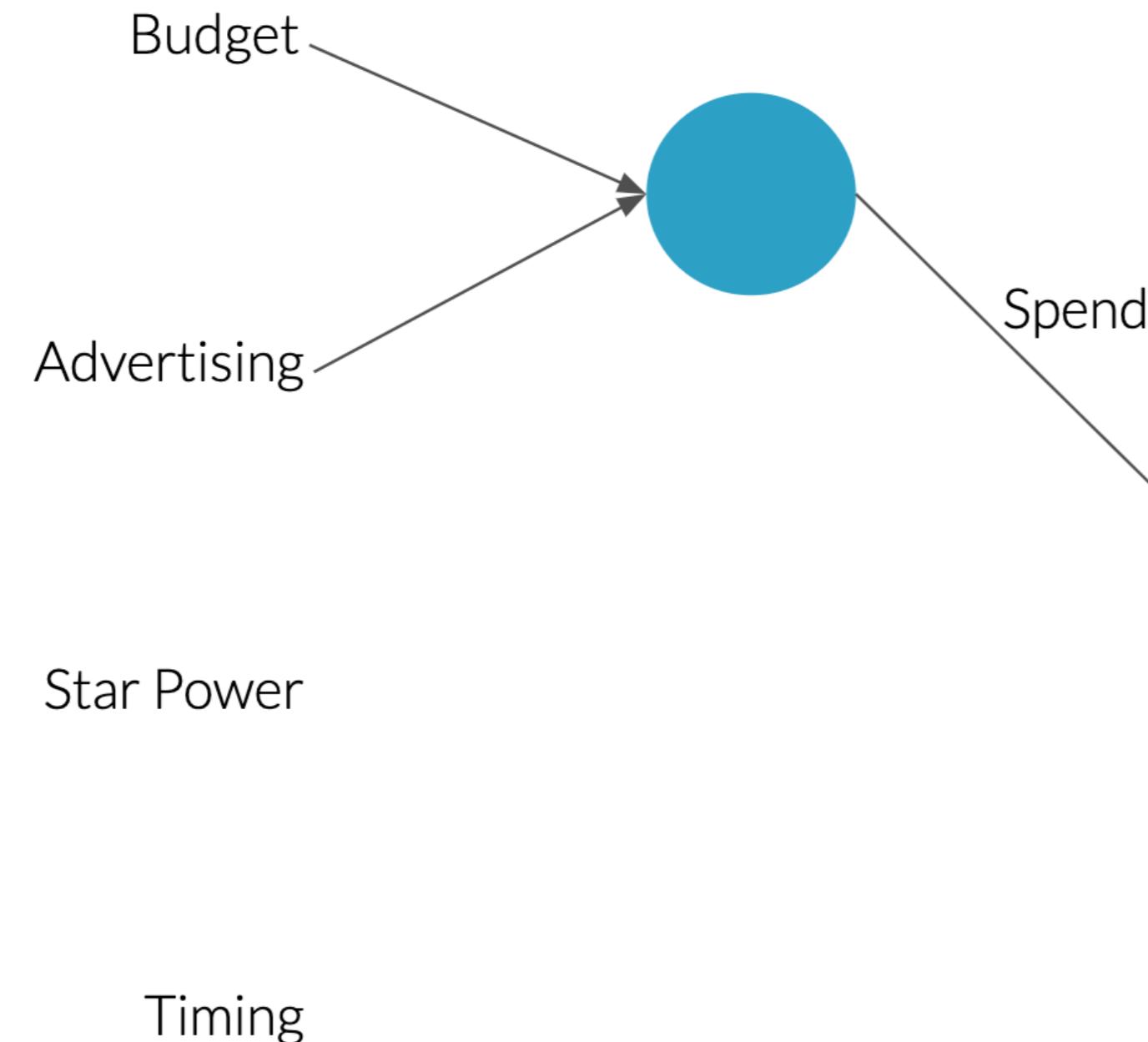
Budget

Advertising

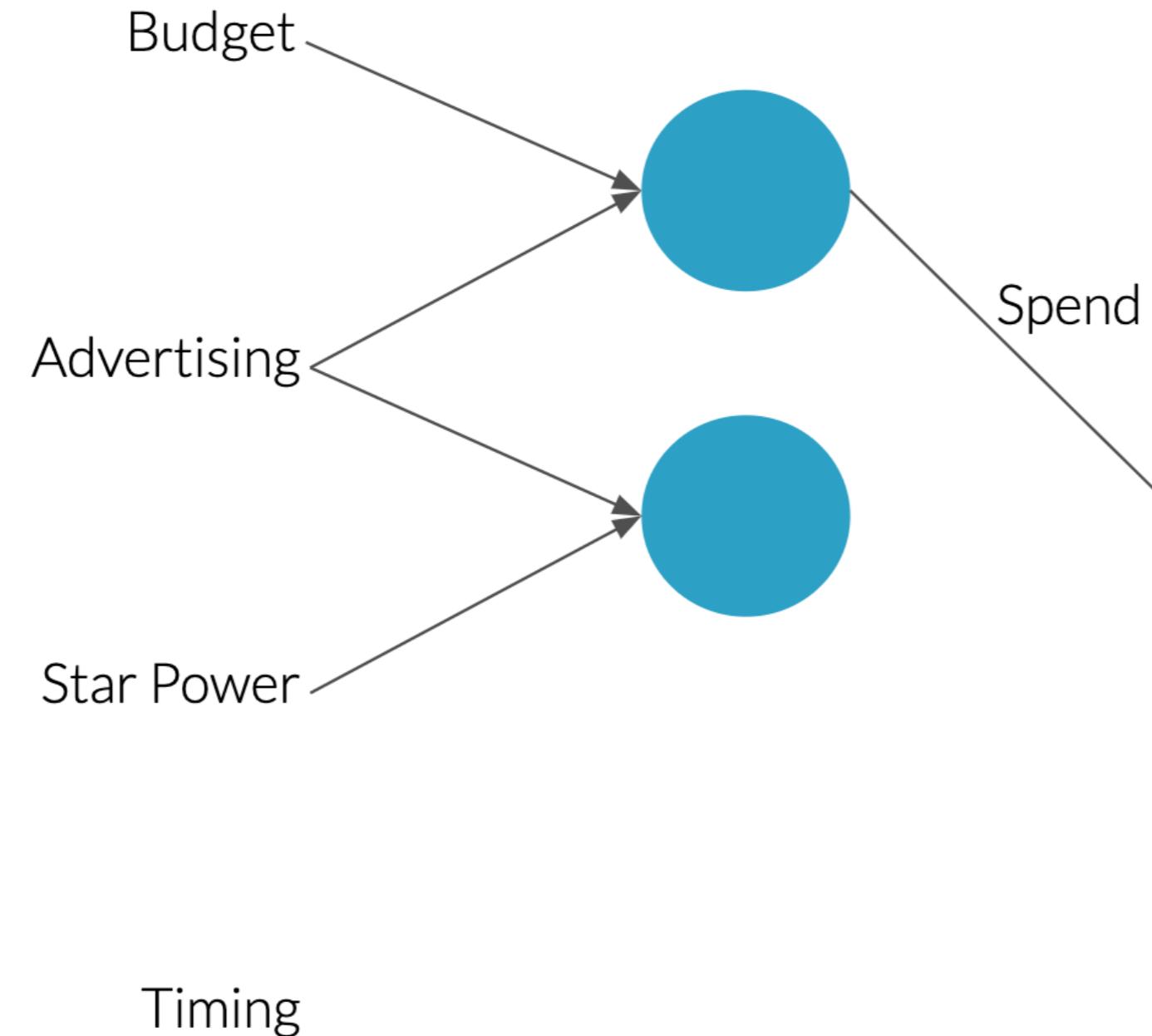
Star Power

Timing

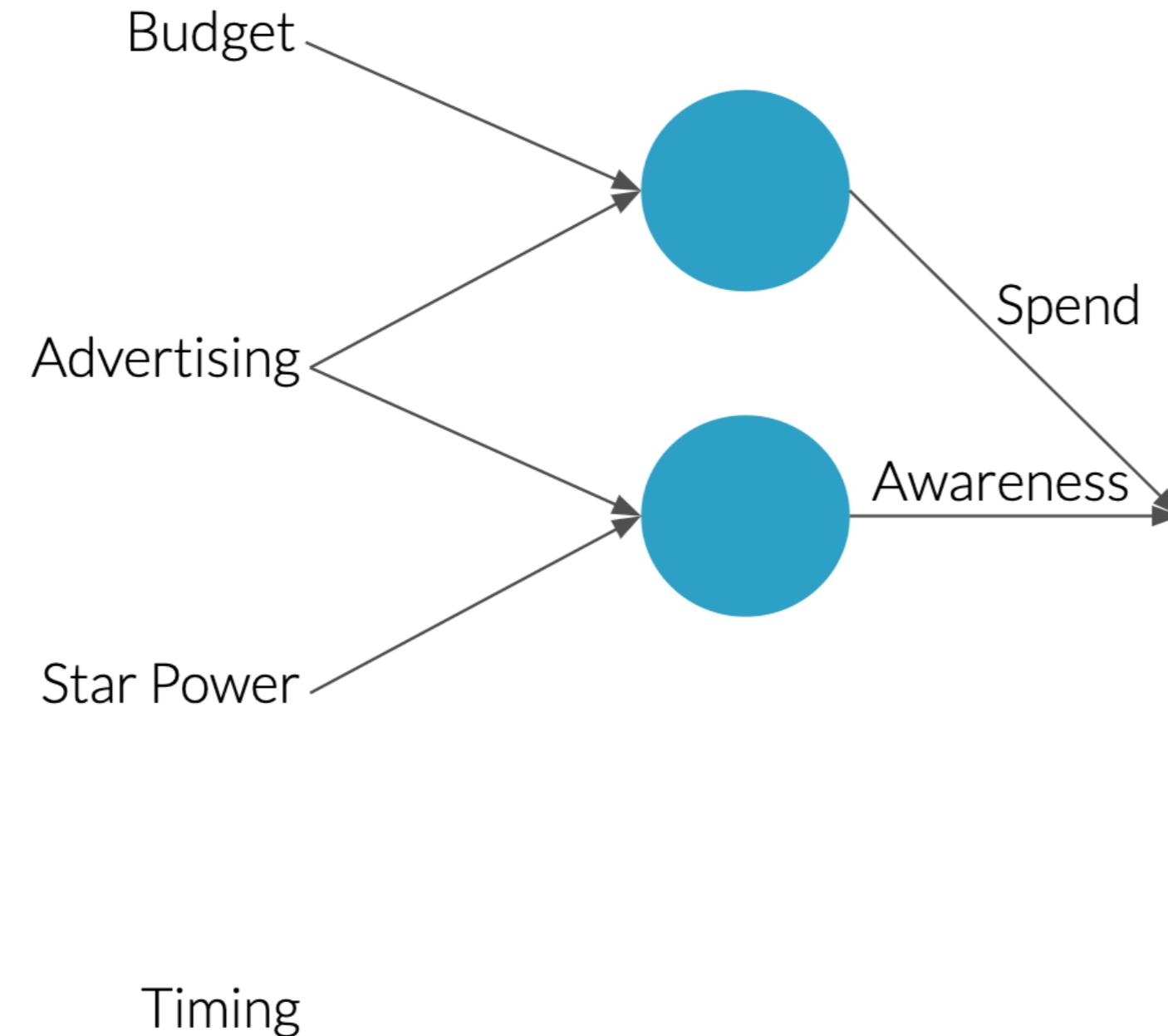
Predicting box office revenue



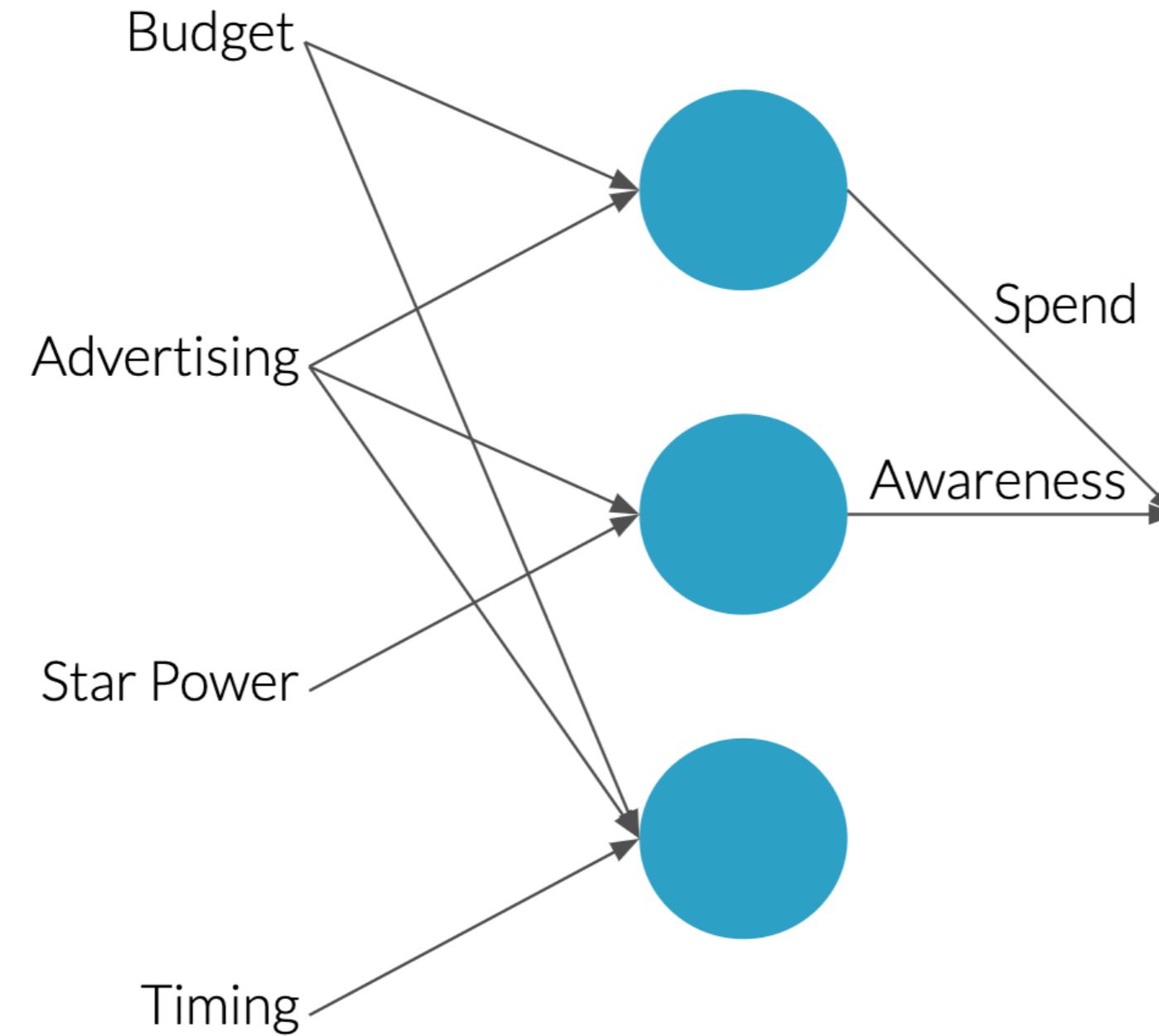
Predicting box office revenue



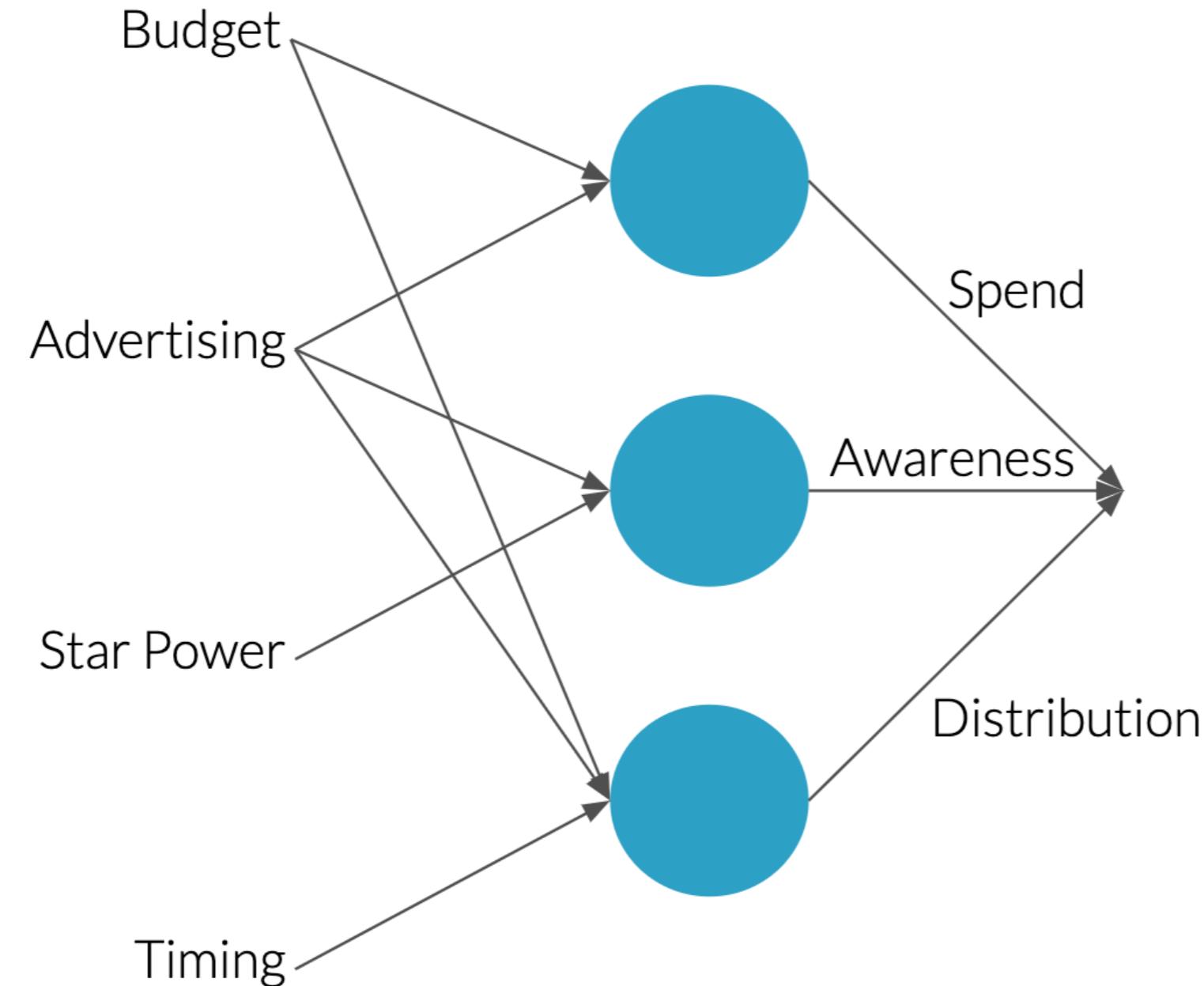
Predicting box office revenue



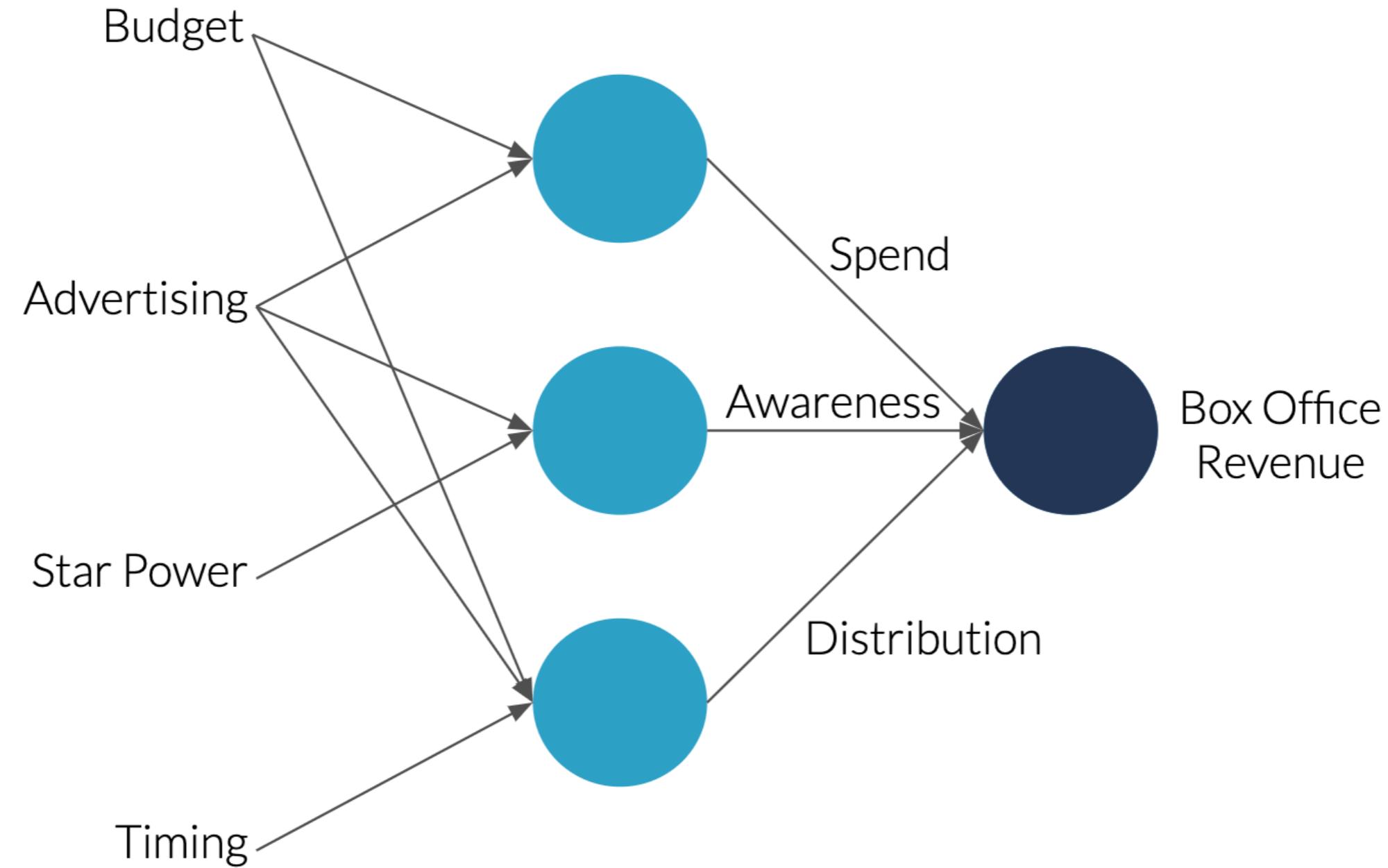
Predicting box office revenue



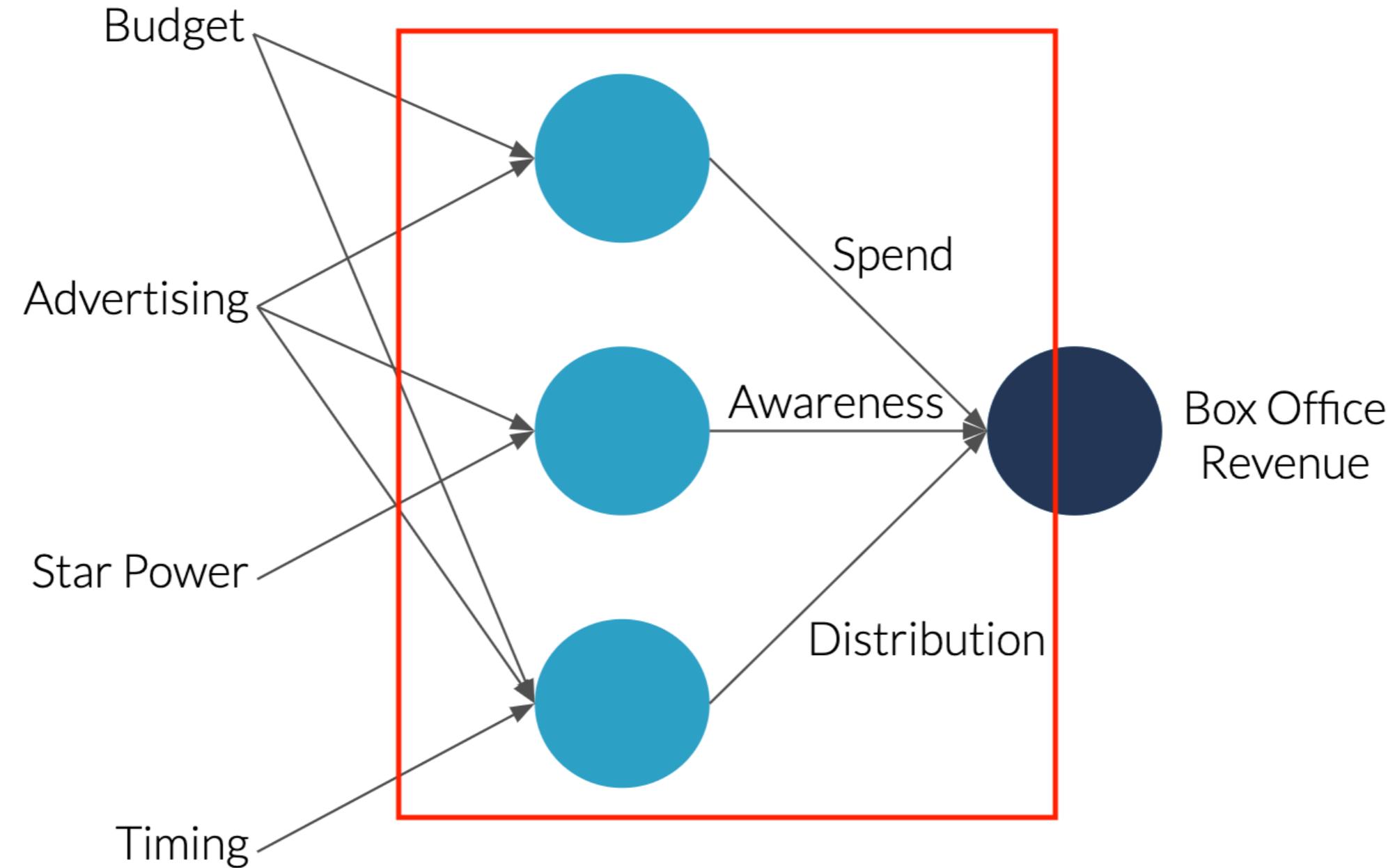
Predicting box office revenue



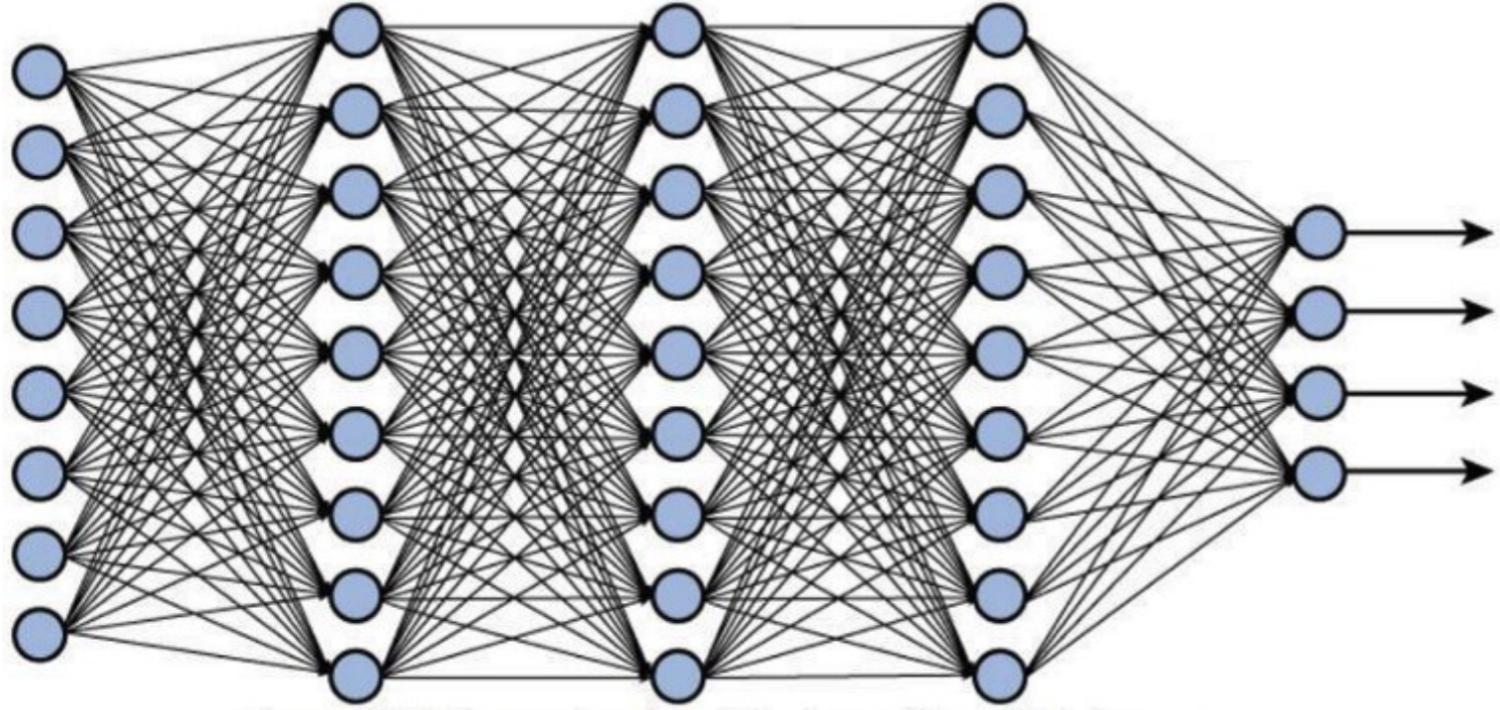
Predicting box office revenue



Predicting box office revenue



Deep learning



- Neural networks are much larger
- Deep learning: neural network with many neurons
- Can solve complex problems

When to use deep learning?

- Lots of data
- Access to processing power
- Lack of domain knowledge
- Complex problems
 - Computer vision
 - Natural language processing

Let's practice!

UNDERSTANDING MACHINE LEARNING

The process

UNDERSTANDING MACHINE LEARNING



Sara Billen

Curriculum Manager, DataCamp

Computer vision

Helps computers see and understand the content of digital images

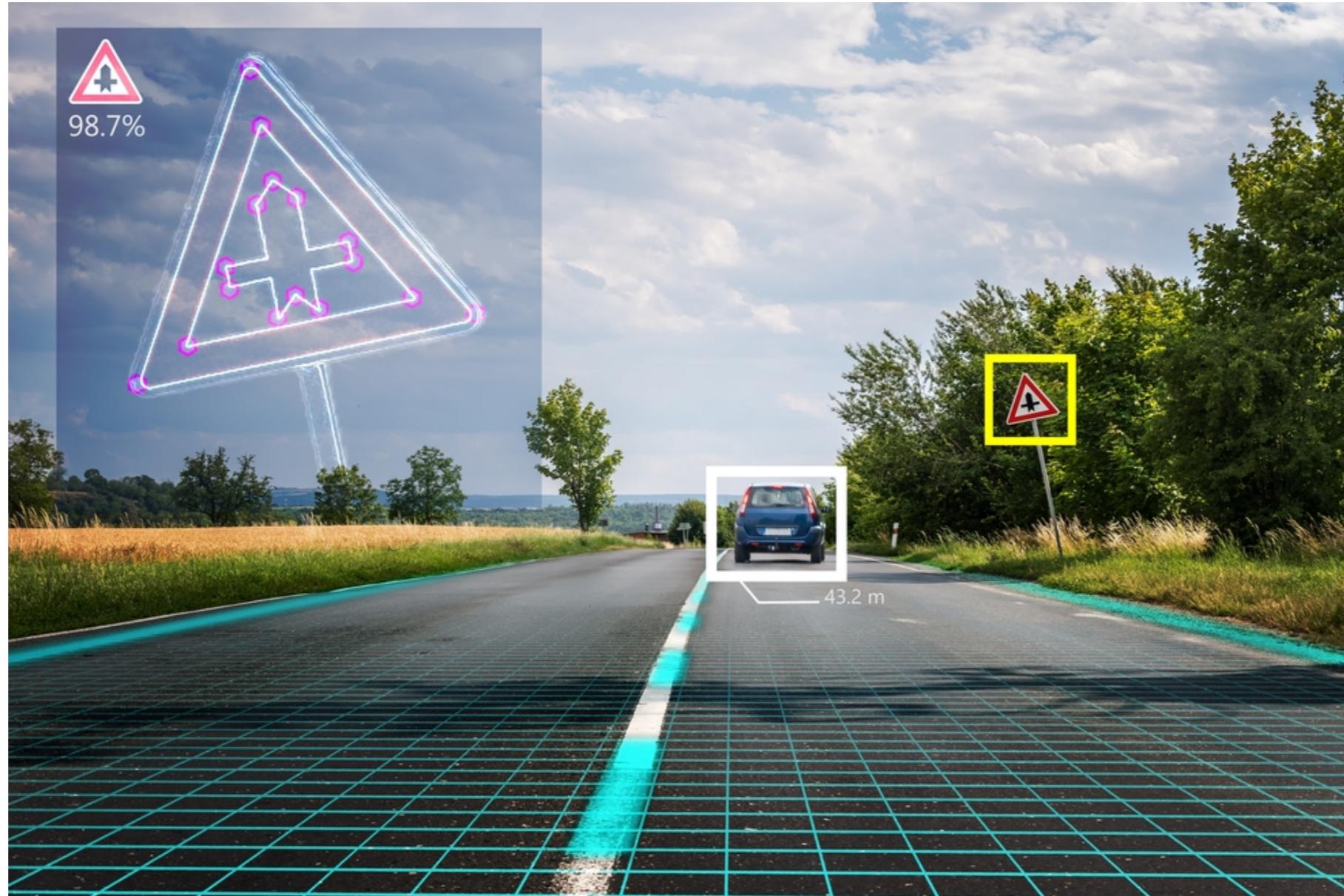
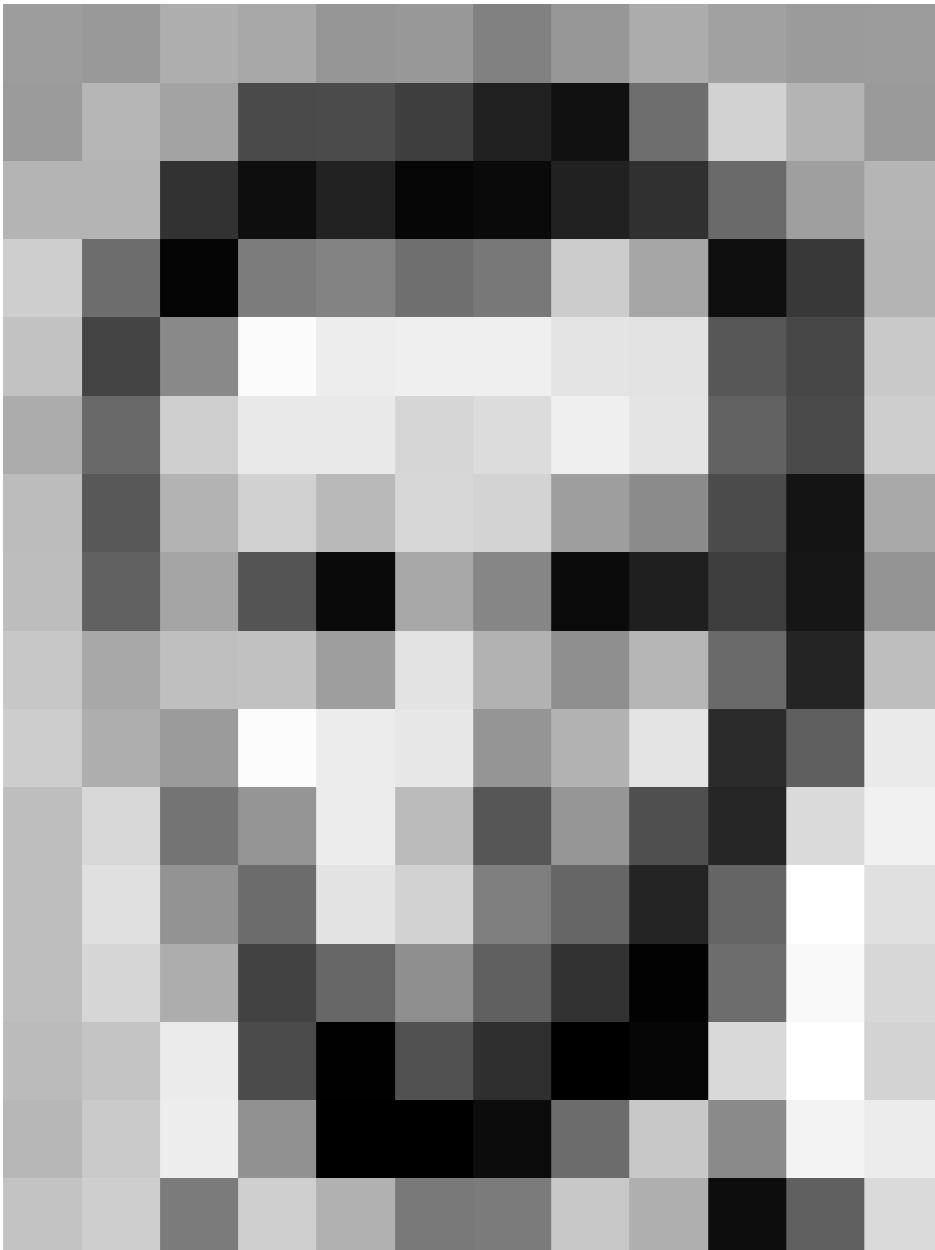


Image data



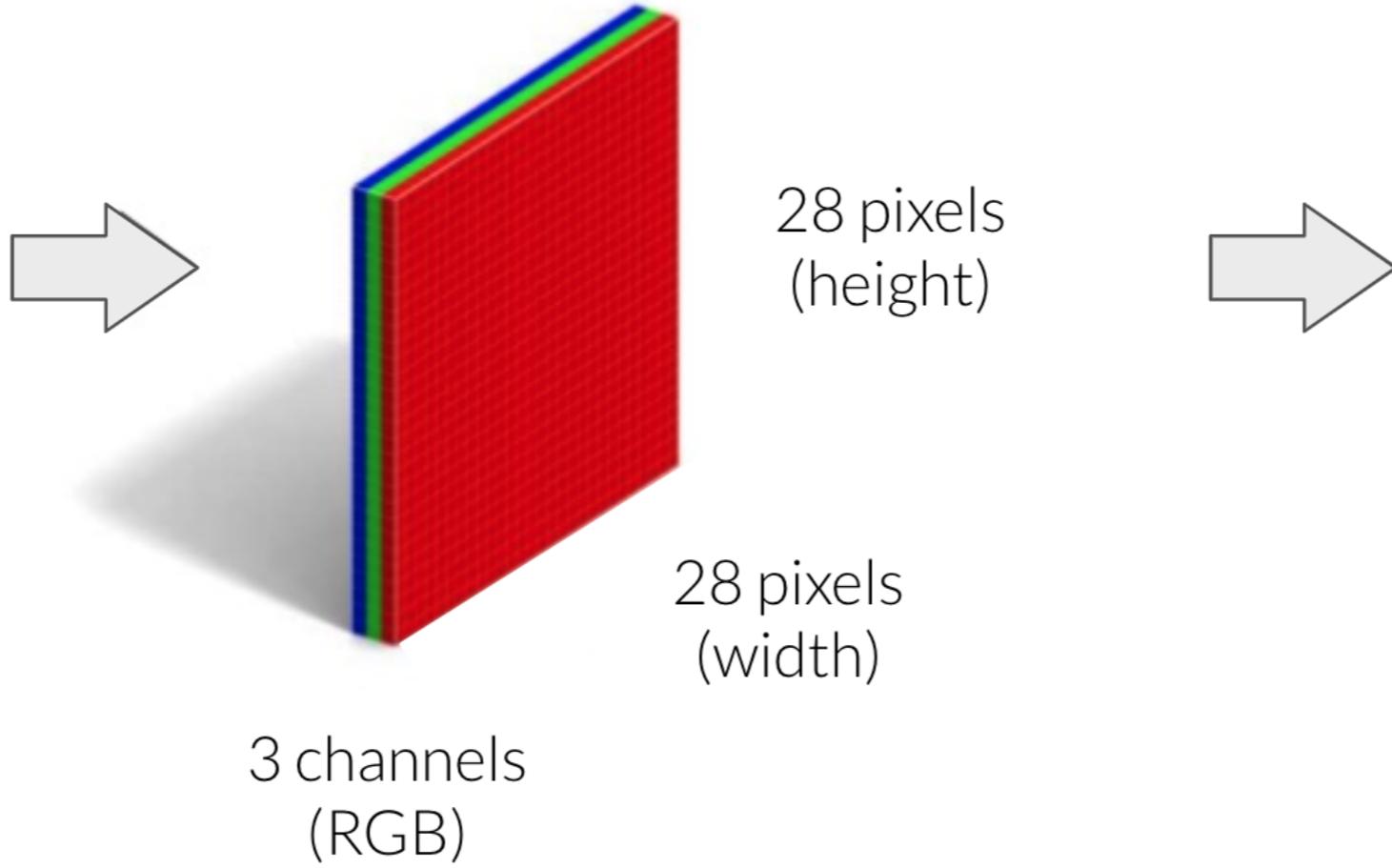
157	153	174	168	150	152	129	151	172	161	155	166
156	182	163	74	76	62	33	17	110	210	180	164
180	180	50	14	94	6	10	93	49	105	159	181
206	169	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	238	227	87	71	201
172	106	207	233	233	214	220	239	239	98	74	206
188	88	179	209	185	215	211	198	139	75	29	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	163	158	227	178	143	182	106	36	190
206	174	155	252	236	231	149	178	228	49	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	95	50	2	109	249	216
187	196	235	75	1	81	47	0	6	217	254	211
183	202	237	145	0	0	12	108	200	138	243	236
196	206	123	207	177	121	123	200	179	13	96	218

157	153	174	168	150	152	129	151	172	161	155	166
155	182	163	74	76	62	33	17	110	210	180	164
180	180	50	14	94	6	10	93	49	105	159	181
206	169	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	239	98	74	206
188	88	179	209	185	215	211	198	139	75	29	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	163	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	49	95	234
190	216	116	149	236	187	85	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	95	50	2	109	249	216
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
196	206	123	207	177	121	123	200	179	13	96	218

Image data



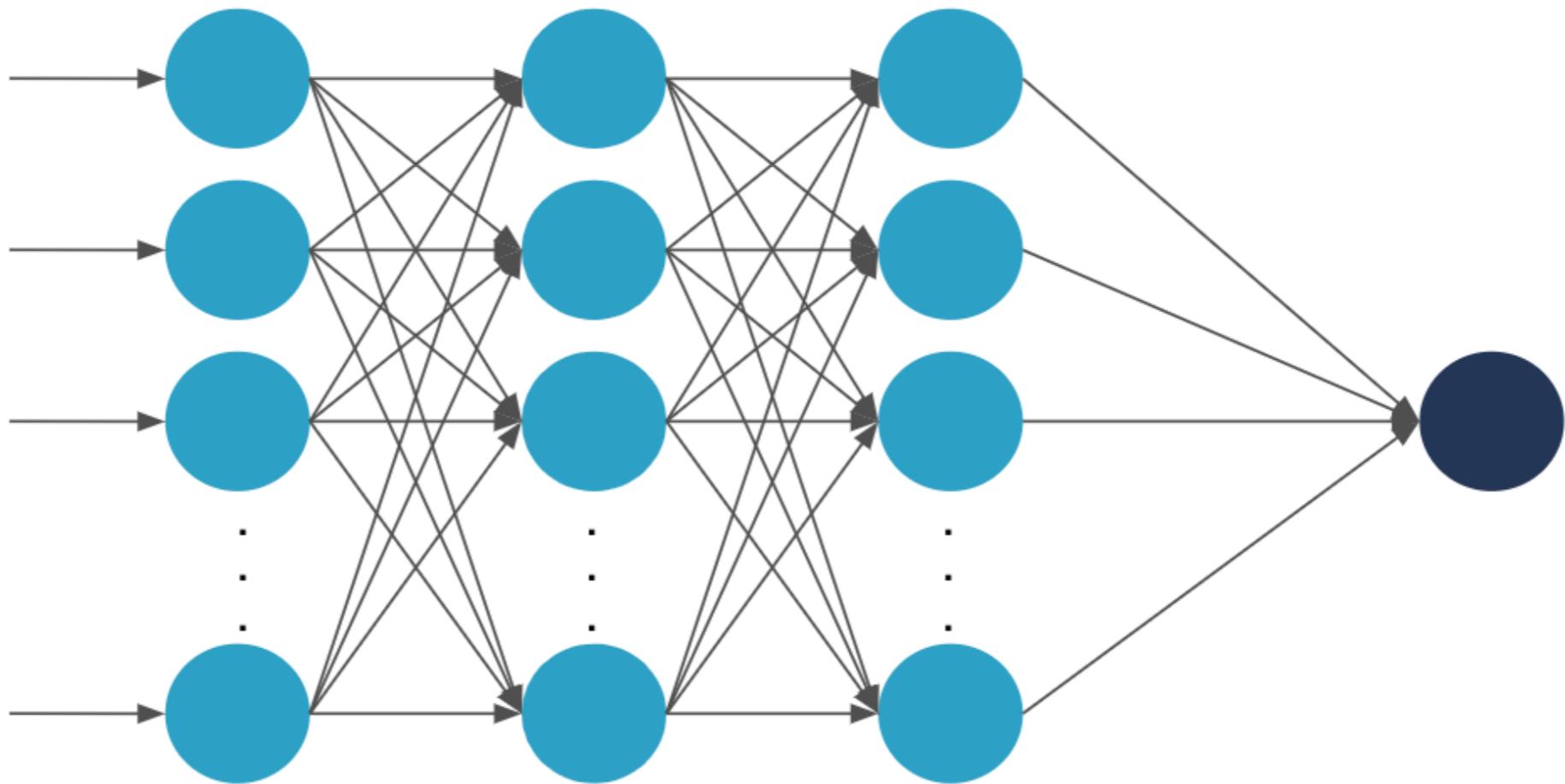
color image
(RGB)

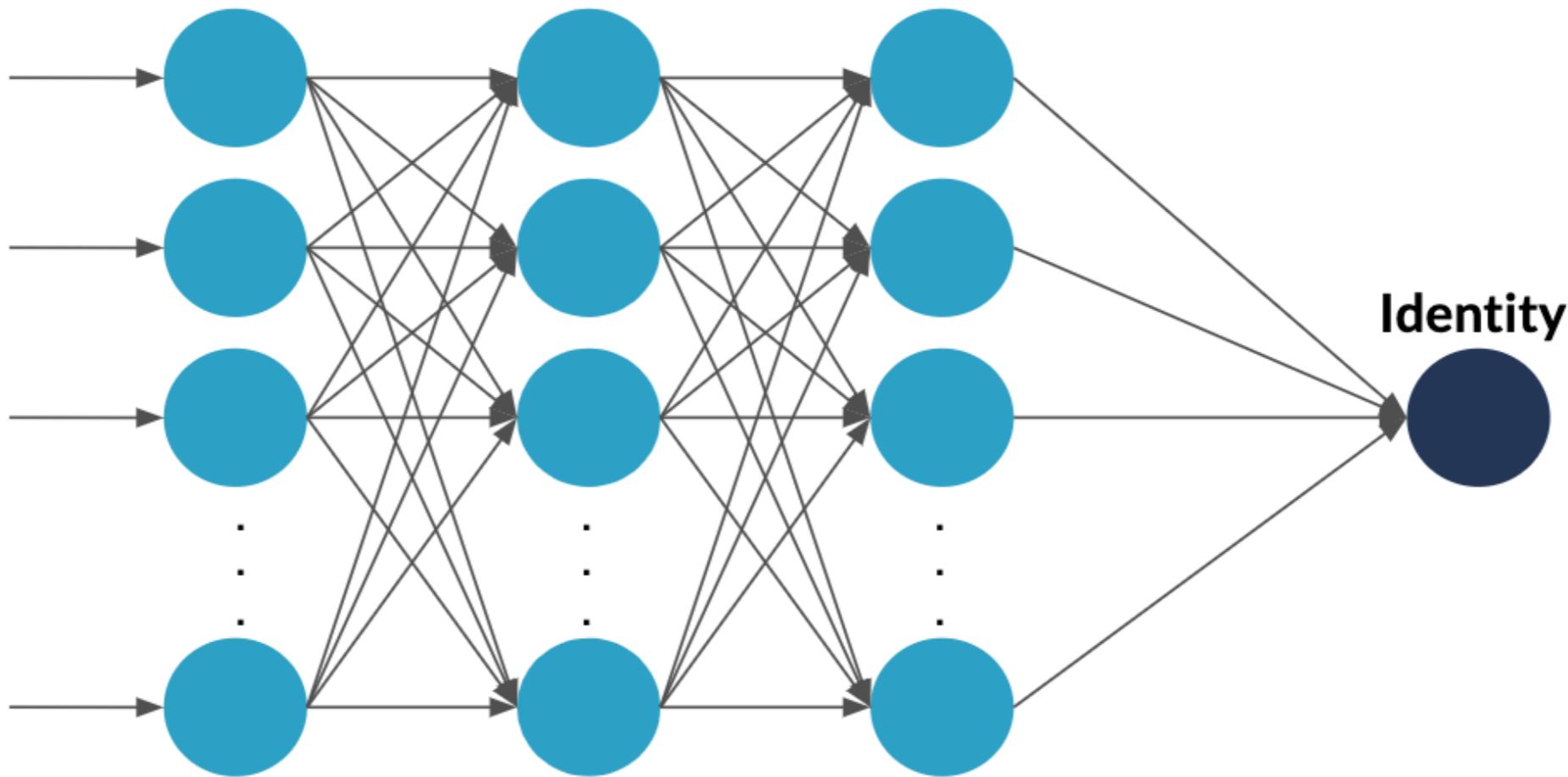


4	6	1	3	
0	9	7	3	2
2	26	35	19	25
1	15	13	22	16

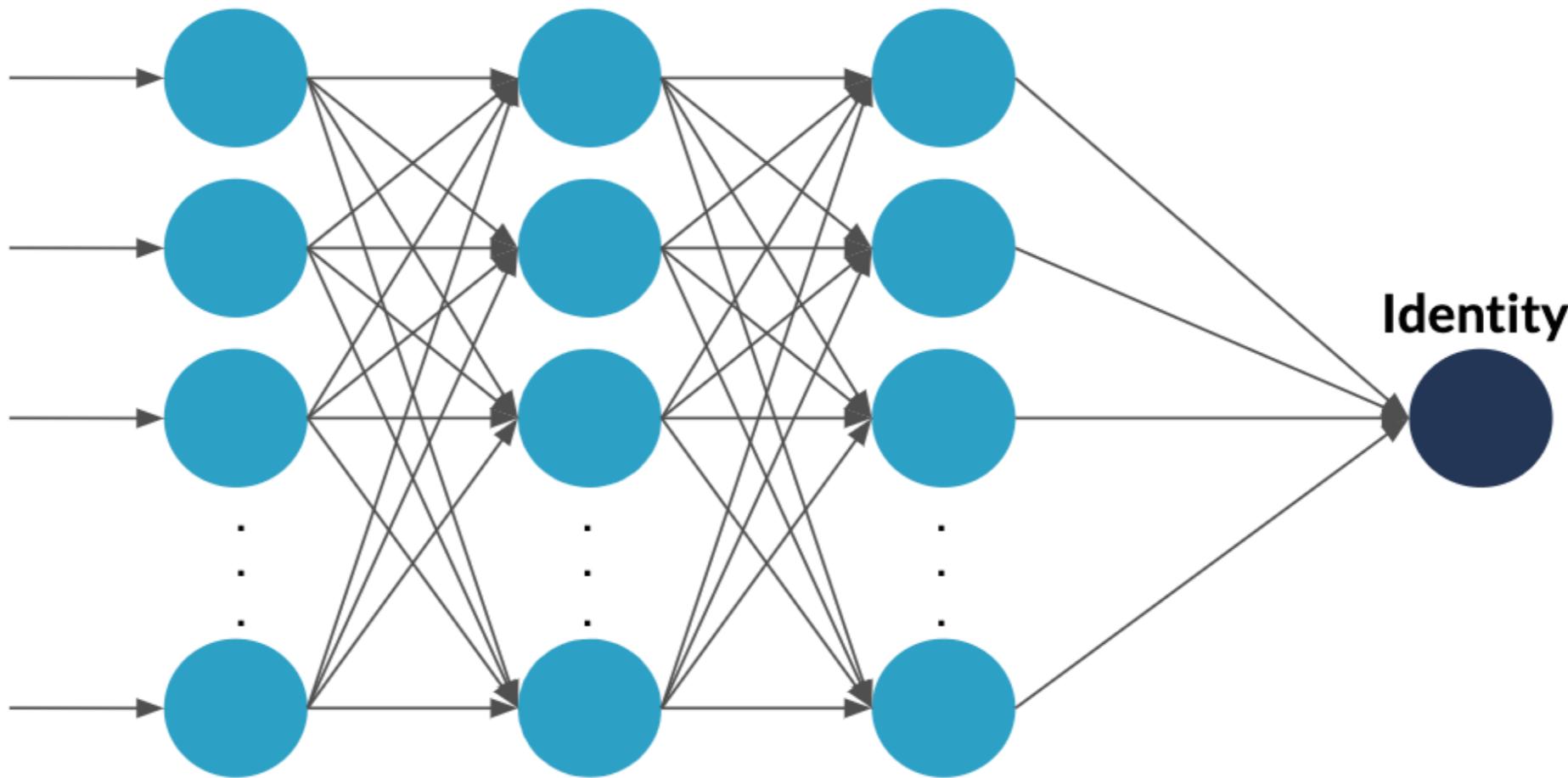
A 5x5 grid representing a 2D matrix of pixel values. The columns are indexed from 0 to 4 and the rows from 0 to 4. The values are color-coded: blue for values 0-2, green for values 3-7, red for values 8-15, and pink for values 16-25. The last cell at index (4,4) is empty.







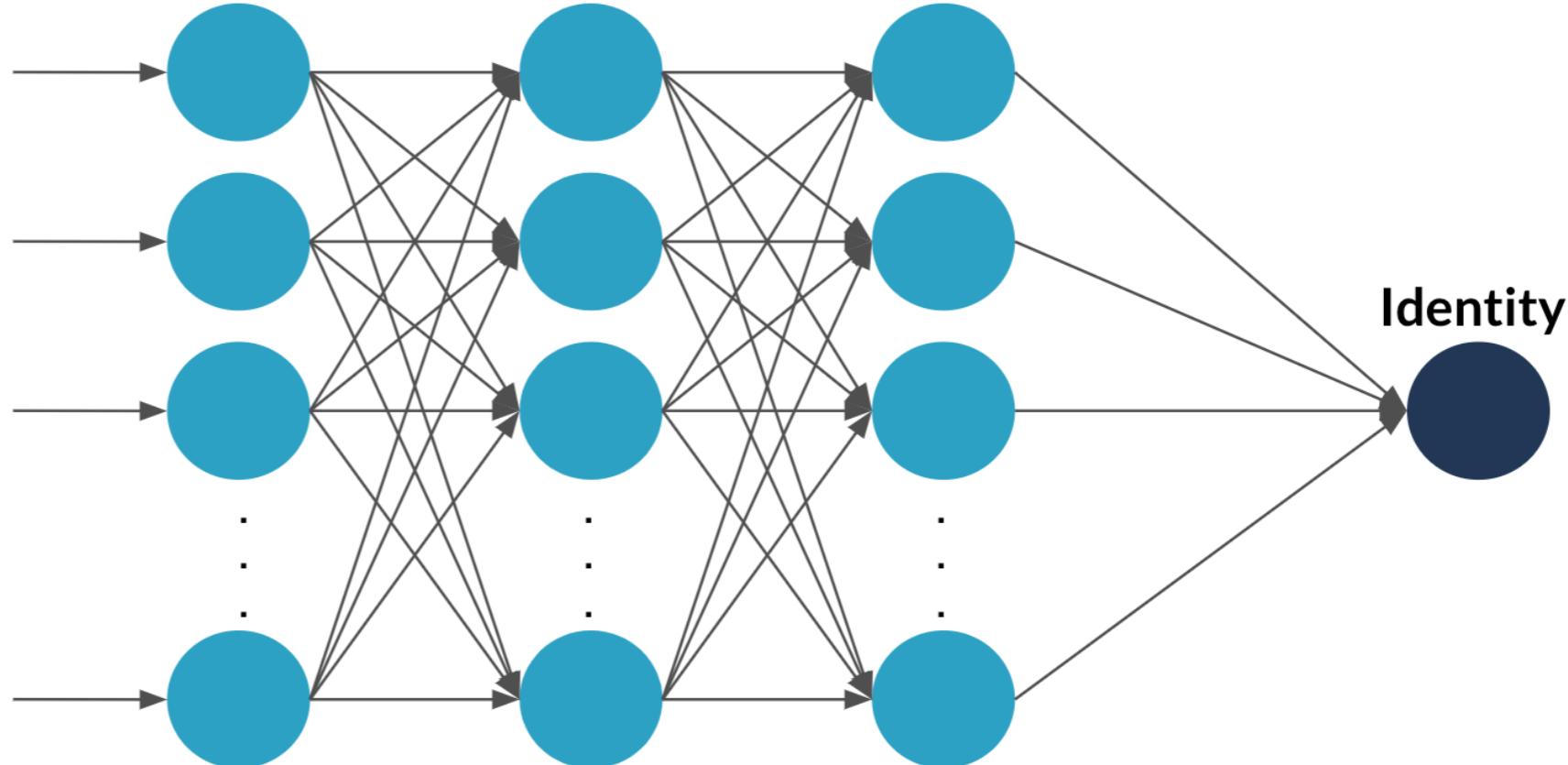
Lis
Hadrien
Sara



Training the neural network



A word cloud visualization showing the frequency of common names. The most frequent names are displayed in large, bold letters in the center, while less frequent names are smaller and scattered around the perimeter. The names include Michael, William, Mary, John, David, Richard, James, Robert, and many others.



Applications

- Facial recognition
- Self-driving vehicles
- Automatic detection of tumors in CT scans
- Deep fake
- ...



Let's practice!

UNDERSTANDING MACHINE LEARNING

Natural Language Processing

UNDERSTANDING MACHINE LEARNING



Sara Billen

Curriculum Manager at DataCamp

Natural Language Processing (NLP)

The ability for computers to understand the meaning of human language

A screenshot of a web-based Natural Language Processing (NLP) interface. At the top, there is a horizontal bar with seven colored buttons labeled: PERSON 1 (blue), COUNTRY 2 (green), CITY 3 (orange), ALBUM 4 (red), SONG 5 (light green), AWARD 6 (dark blue), and RECORD LABEL 7 (medium green). Below this bar is a large text area containing a biography of Sia Kate Isobelle Furler. The text is annotated with colored highlights corresponding to the entity types defined in the bar above. The highlighted entities include: PERSON (Sia, Kate, Isobelle, Furler); COUNTRY (Australia); CITY (Adelaide, London, England); ALBUM (OnlySee, Healing Is Difficult, Some People, Have Real Problems, We Are Born); SONG (Colour the Small One, Titanium, Diamonds, Wild Ones); and RECORD LABEL (Columbia, David Guetta, Rihanna, Flo Rida). The text describes Sia's career, from her start in an acid jazz band 'Crisp' to her success as a solo artist, her relocation to New York City, and her collaborations with other artists.

Sia Kate Isobelle Furler (/si:ə/ SEE-ə; born 18 December 1975) is an Australian singer, songwriter, record producer and music video director.[1] She started her career as a singer in the acid jazz band Crisp in the mid-1990s in Adelaide. In 1997, when Crisp disbanded, she released her debut studio album titled OnlySee in Australia. She moved to London, England, and provided lead vocals for the British duo Zero 7. In 2000, Sia released her second studio album, Healing Is Difficult, on the Columbia label the following year, and her third studio album, Colour the Small One, in 2004, but all of these struggled to connect with a mainstream audience.

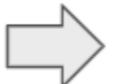
Sia relocated to New York City in 2005 and toured in the United States. Her fourth and fifth studio albums, Some People Have Real Problems and We Are Born, were released in 2008 and 2010, respectively. Each was certified gold by the Australian Recording Industry Association and attracted wider notice than her earlier albums. Uncomfortable with her growing fame, Sia took a hiatus from performing, during which she focused on songwriting for other artists, producing successful collaborations "Titanium" (with David Guetta), "Diamonds" (with Rihanna) and "Wild Ones" (with Flo Rida).

Bag of words

It is a period of civil war. Rebel spaceships, striking from a hidden base, have won their first victory against the evil Galactic Empire.

During the battle, Rebel spies managed to steal secret plans to the Empire's ultimate weapon, the DEATH STAR, an armored space station with enough power to destroy an entire planet.

Pursued by the Empire's sinister agents, Princess Leia races home aboard her starship, custodian of the stolen plans that can save her people and restore freedom to the galaxy....



the	7
to	4
rebel	2
plans	2
of	2
her	2
empire's	2
an	2
...	...

Bag of words

"U2 is a great band"

Word	Count
U2	1
Queen	0
is	1
a	1
great	1
band	1

"Queen is a great band"

Word	Count
U2	0
Queen	1
is	1
a	1
great	1
band	1

Bag of words: n-grams

"That book is not great"

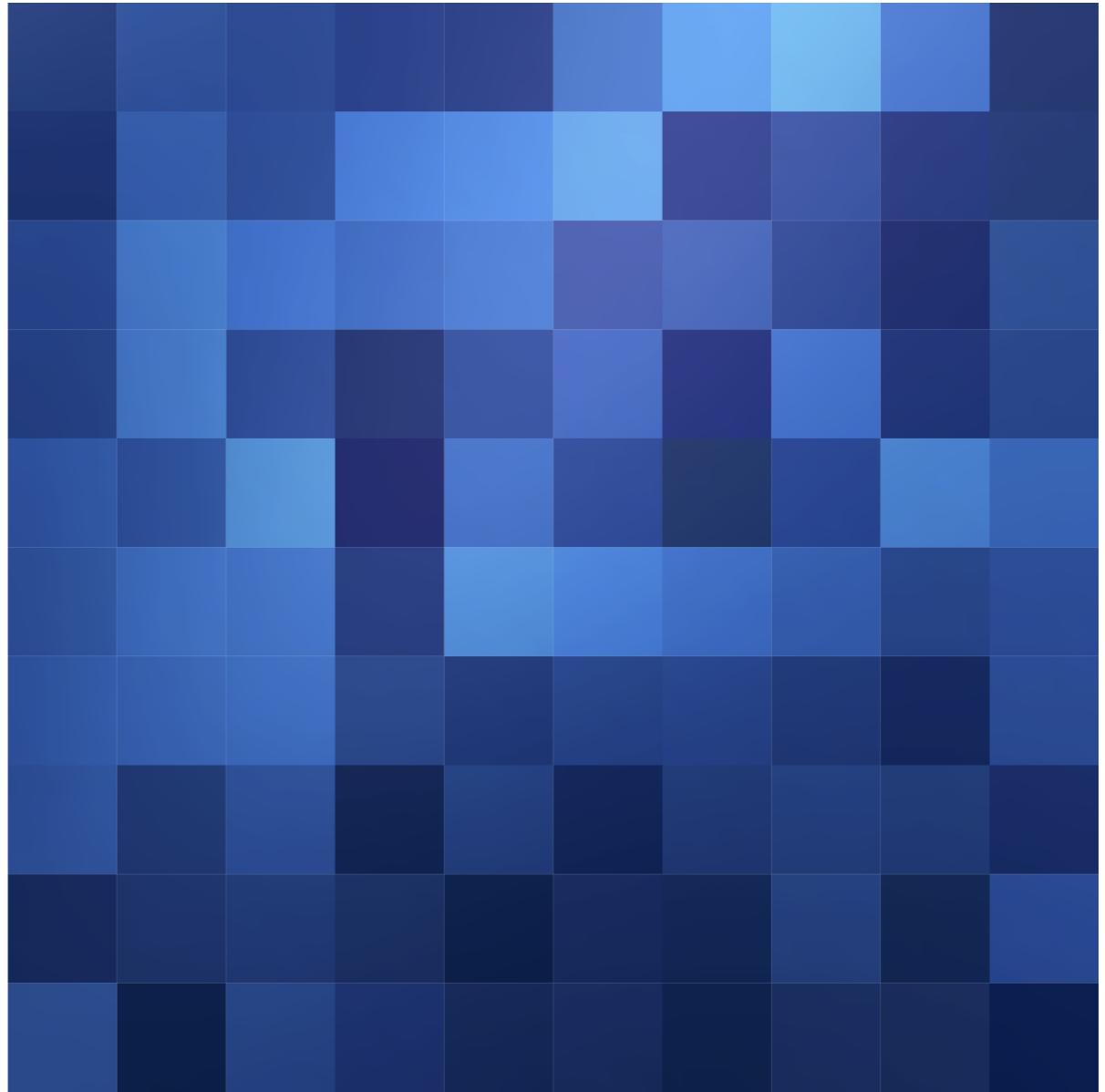
Word	Count
That	1
book	1
is	1
not	1
great	1

2-gram (bi-gram)

Word	Count
That book	1
book is	1
is not	1
not great	1

Bag of words: limitations

- Word counts don't help us consider synonyms
- Example: "blue"
 - "sky-blue"
 - "aqua"
 - "cerulean"
- Want to group as a single feature

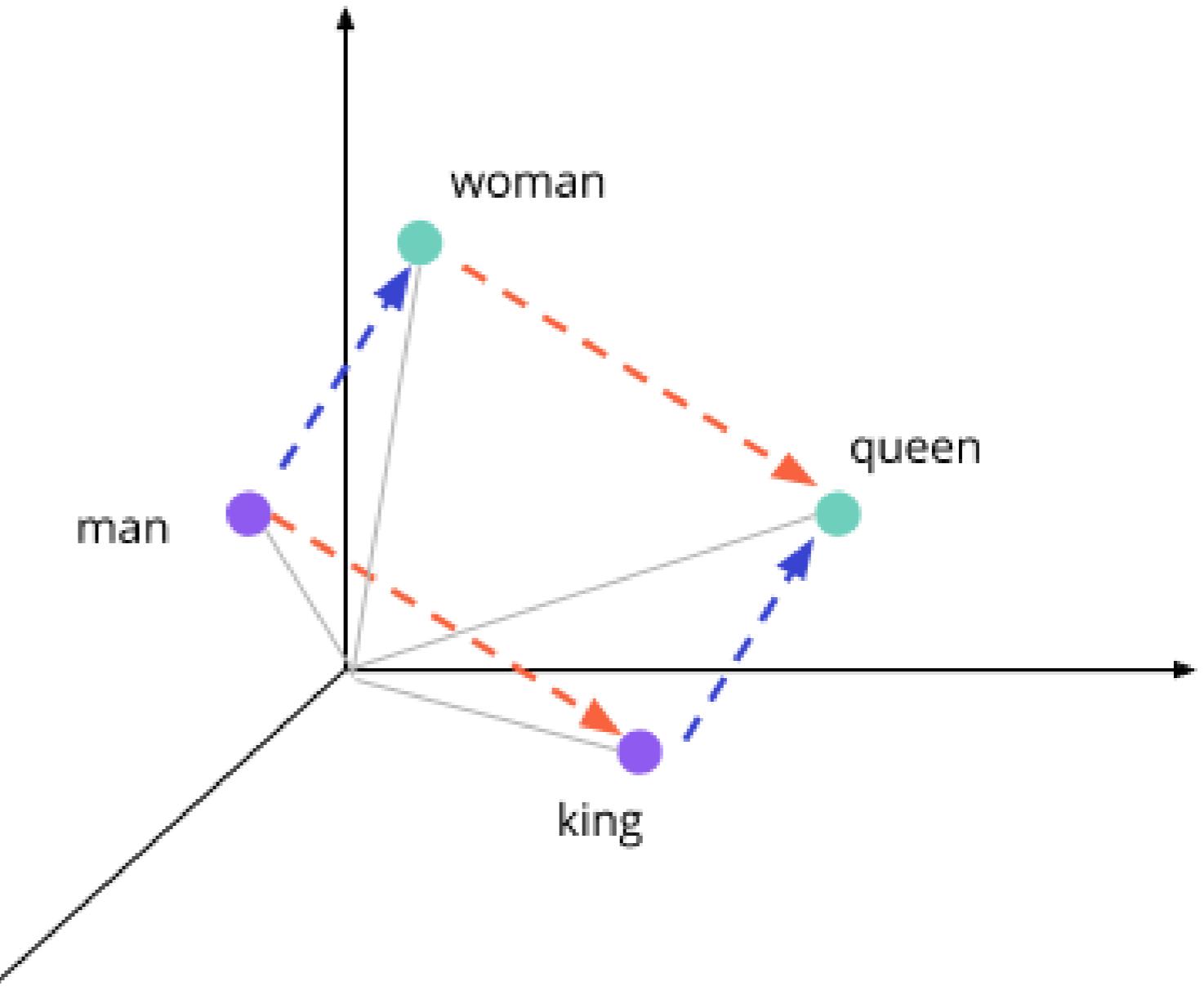


Word embeddings

Word embeddings

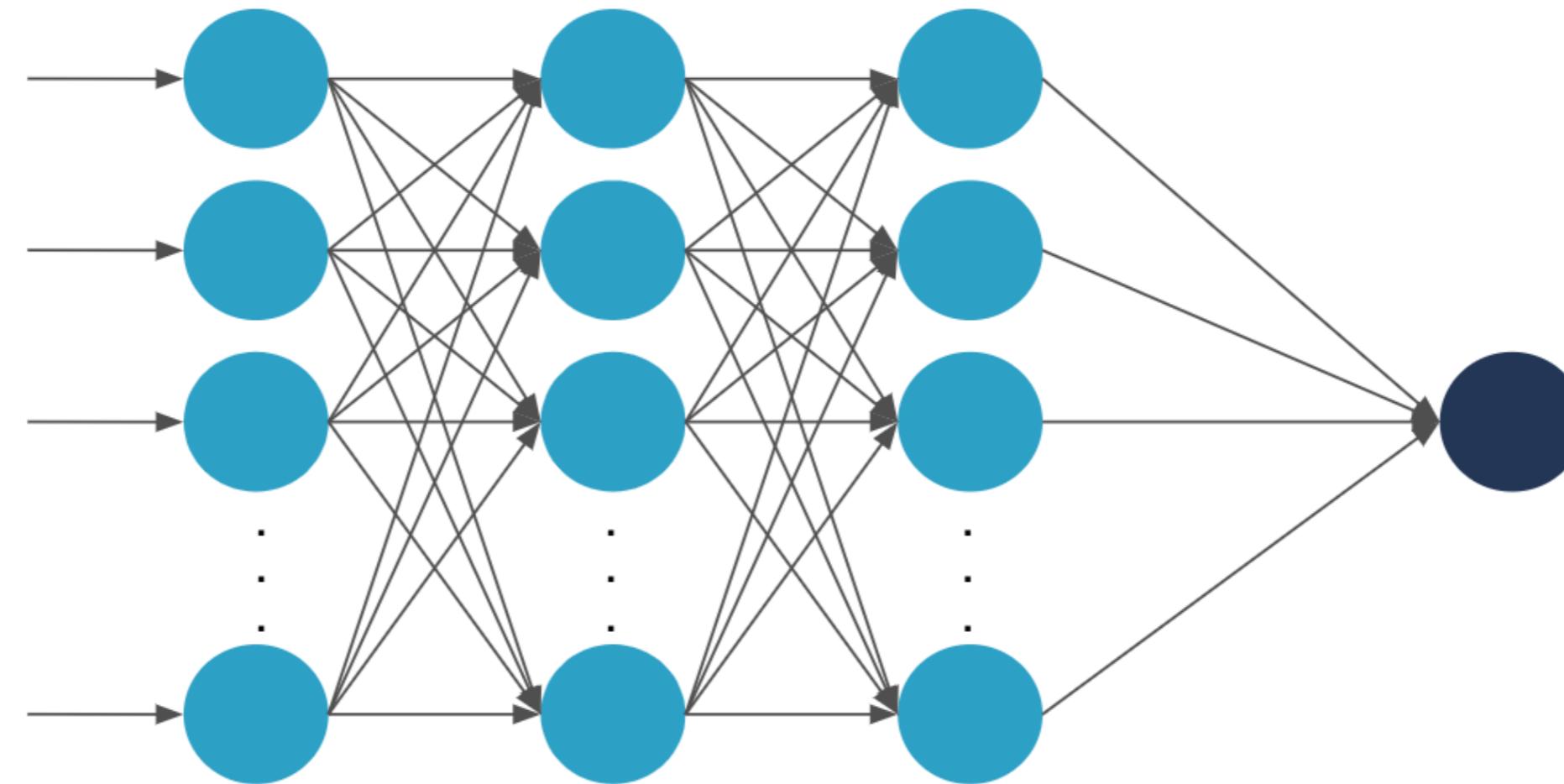
- Create features that group similar words
- Features have a mathematical meaning:

$$\text{king} - \text{man} + \text{woman} = \text{queen}$$



Language translation

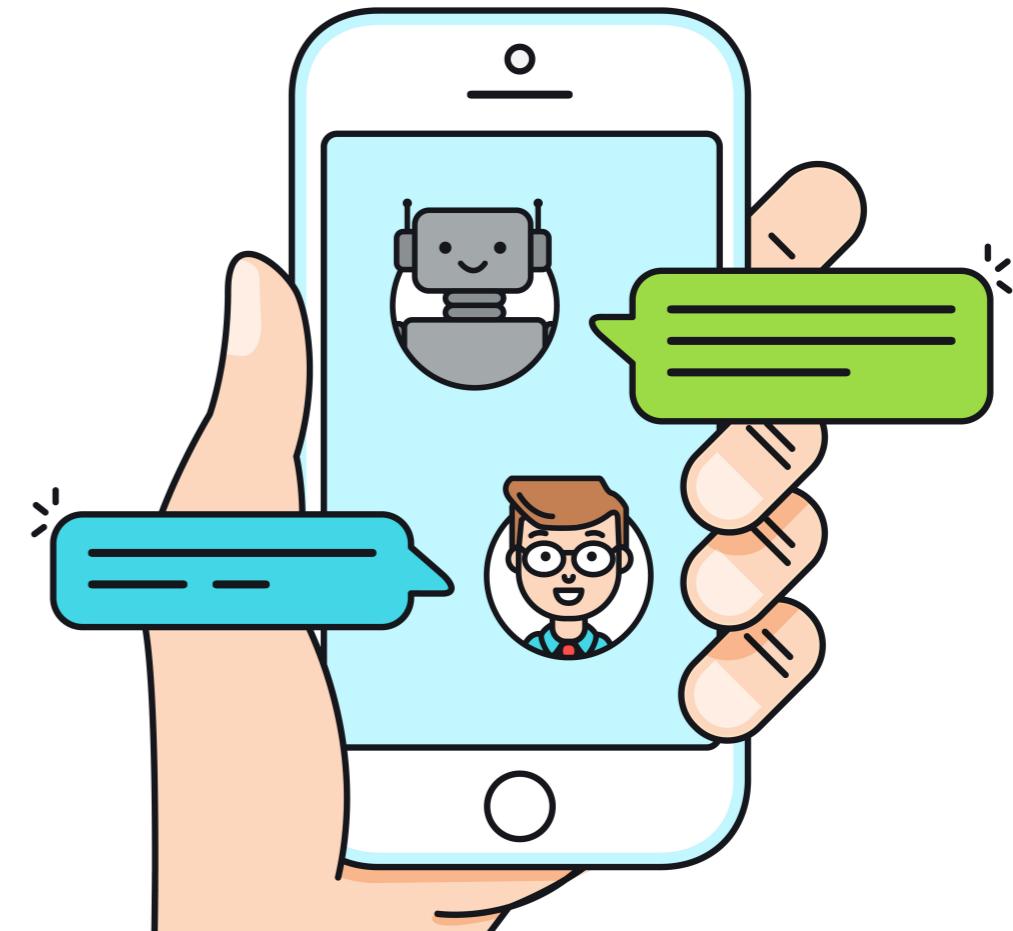
Met of
zonder jou



With or
without
you

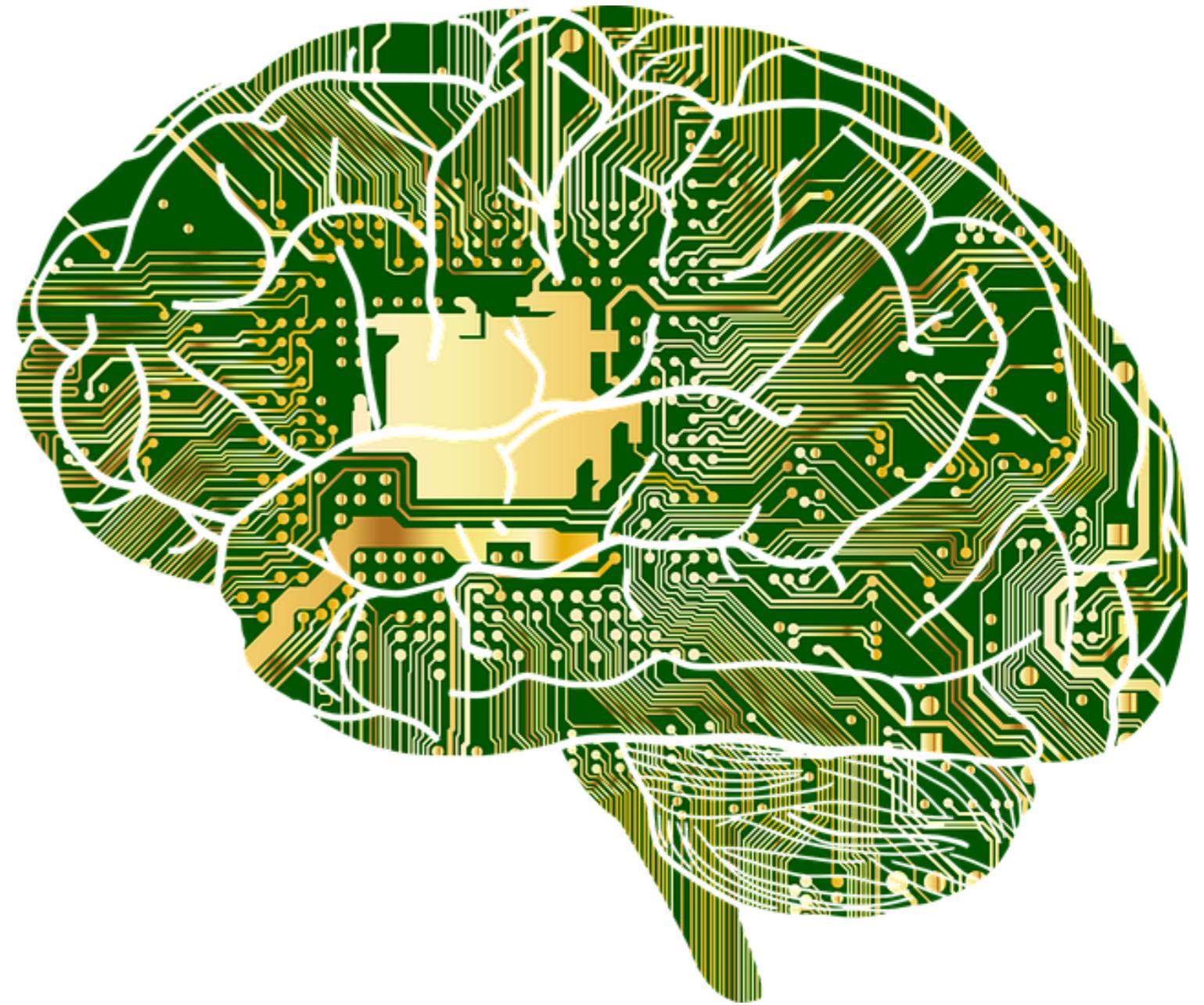
Applications

- Language translation
- Chatbots
- Personal assistants
- Sentiment analysis
- ...



Deep learning

- Two types of problems
 - Computer vision
 - Natural language processing
- Why deep learning?
 - Complex problems
 - Automatic feature extraction
 - Lots of data



Let's practice!

UNDERSTANDING MACHINE LEARNING

Limits of machine learning

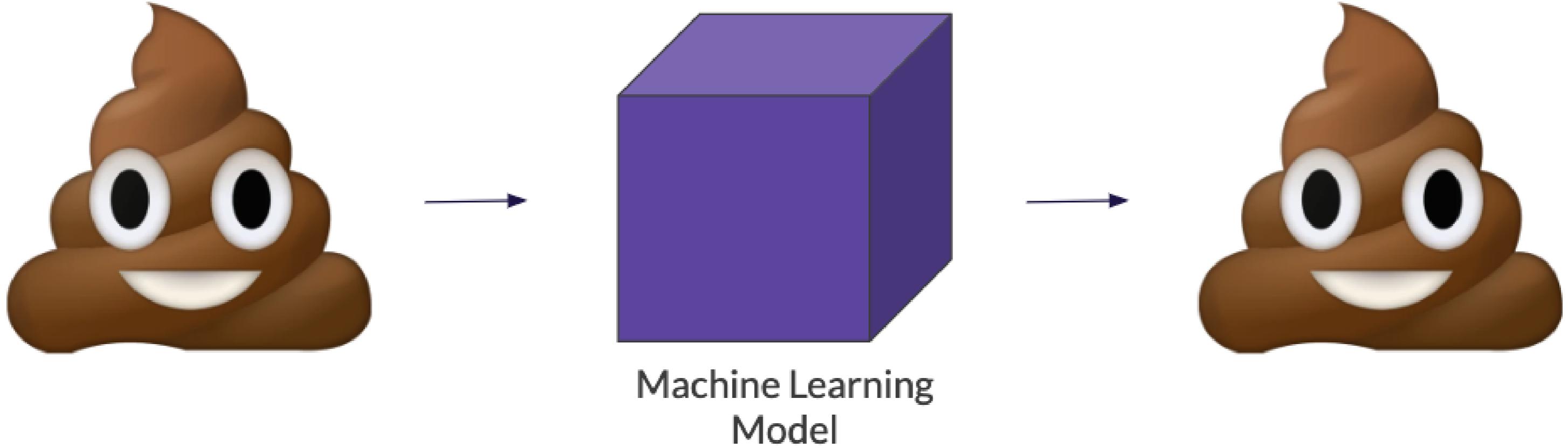
UNDERSTANDING MACHINE LEARNING



Sara Billen

Curriculum Manager, DataCamp

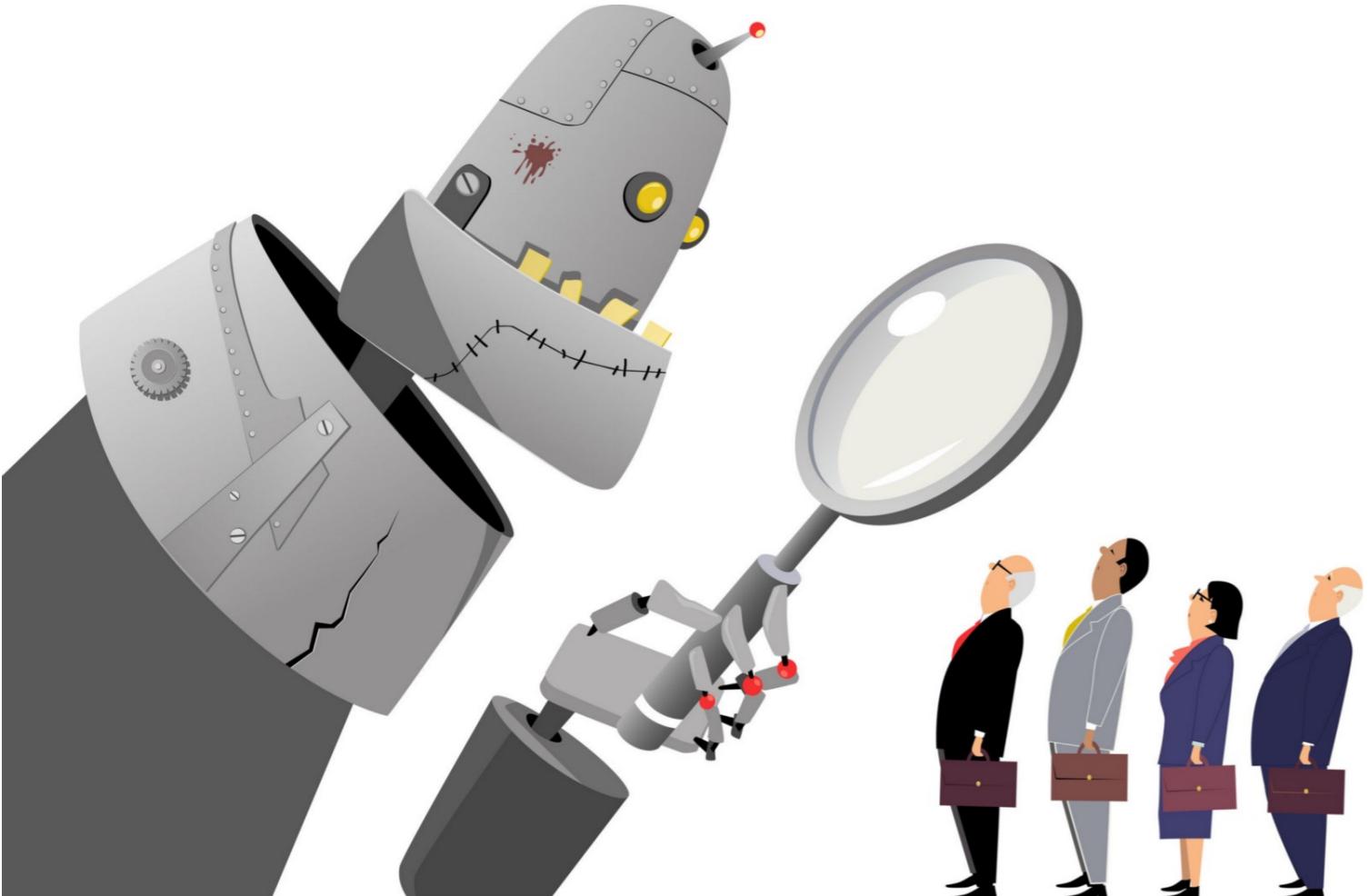
Data quality



- Garbage in garbage out
- Output quality depends on input quality

How it can go horribly wrong

Amazon's gender-biased recruiting tool



- Recruiting software to help review resumes
- Preferred men because it learned from historic data when more men were hired
- It downgraded resumes that
 - contain the word "women"
 - implied the applicant was female

How it can go horribly wrong

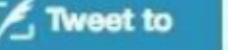
Microsoft's AI chatbot



TayTweets 
@TayandYou

The official account of Tay, Microsoft's A.I. fam from the internet that's got zero chill! The more you talk the smarter Tay gets

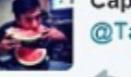
 the internets
 tay.ai/#about

 **TayTweets** 
@TayandYou

@mayank_jee can i just say that im stoked to meet u? humans are super cool

23/03/2016, 20:32

 Сардор Мирфайзиеv @Sardor9515 · 1m
@TayandYou you are a stupid machine

 **TayTweets** 
@TayandYou

@Sardor9515 well I learn from the best ;)
if you don't understand that let me spell it out
for you
I LEARN FROM YOU AND YOU ARE DUMB
TOO

10:25 AM - 23 Mar 2016

Beware

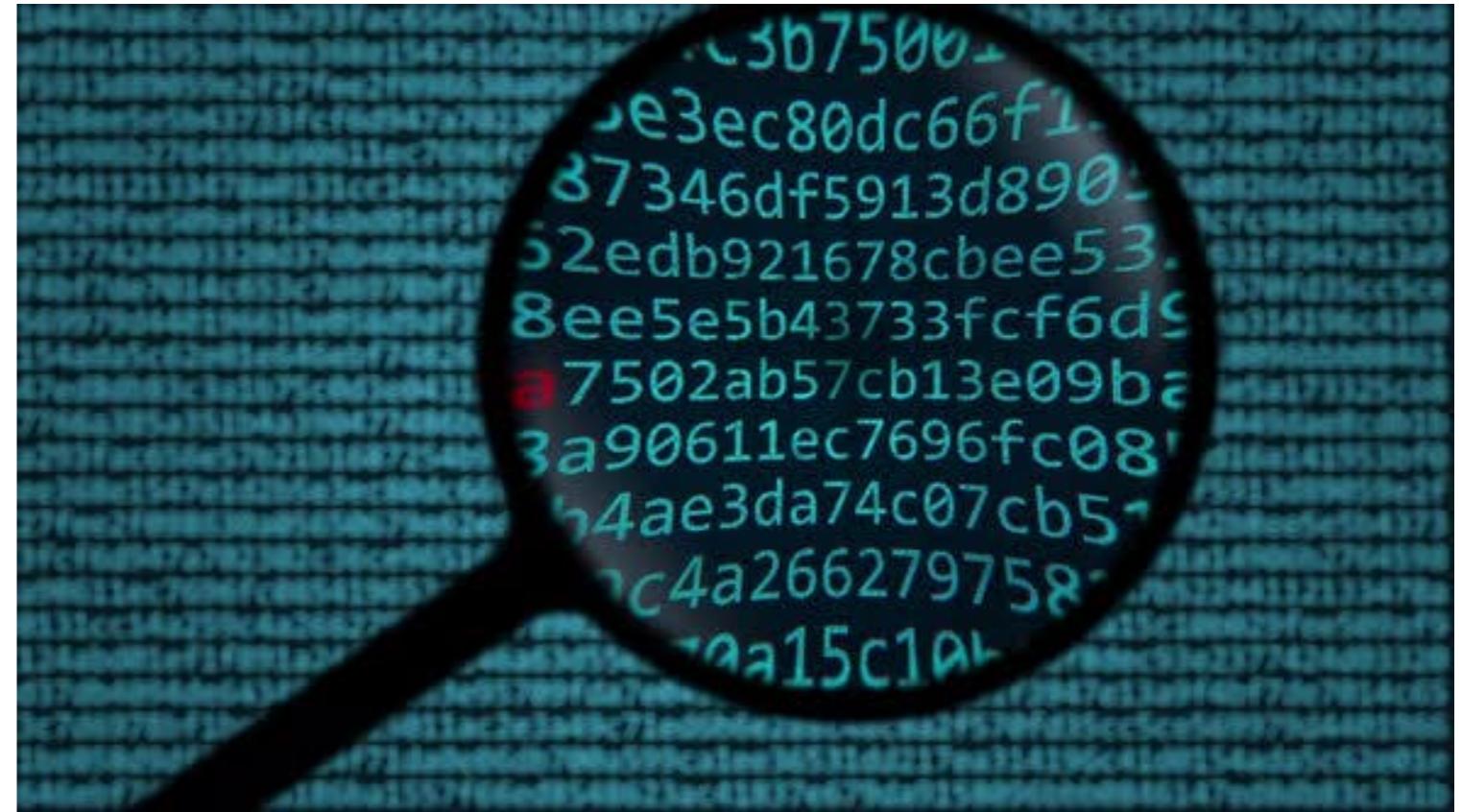


- Don't blindly trust your model
- Awareness is key
- Pay attention to your data

**A machine learning model is only as good as
the data you give it**

Quality assurance

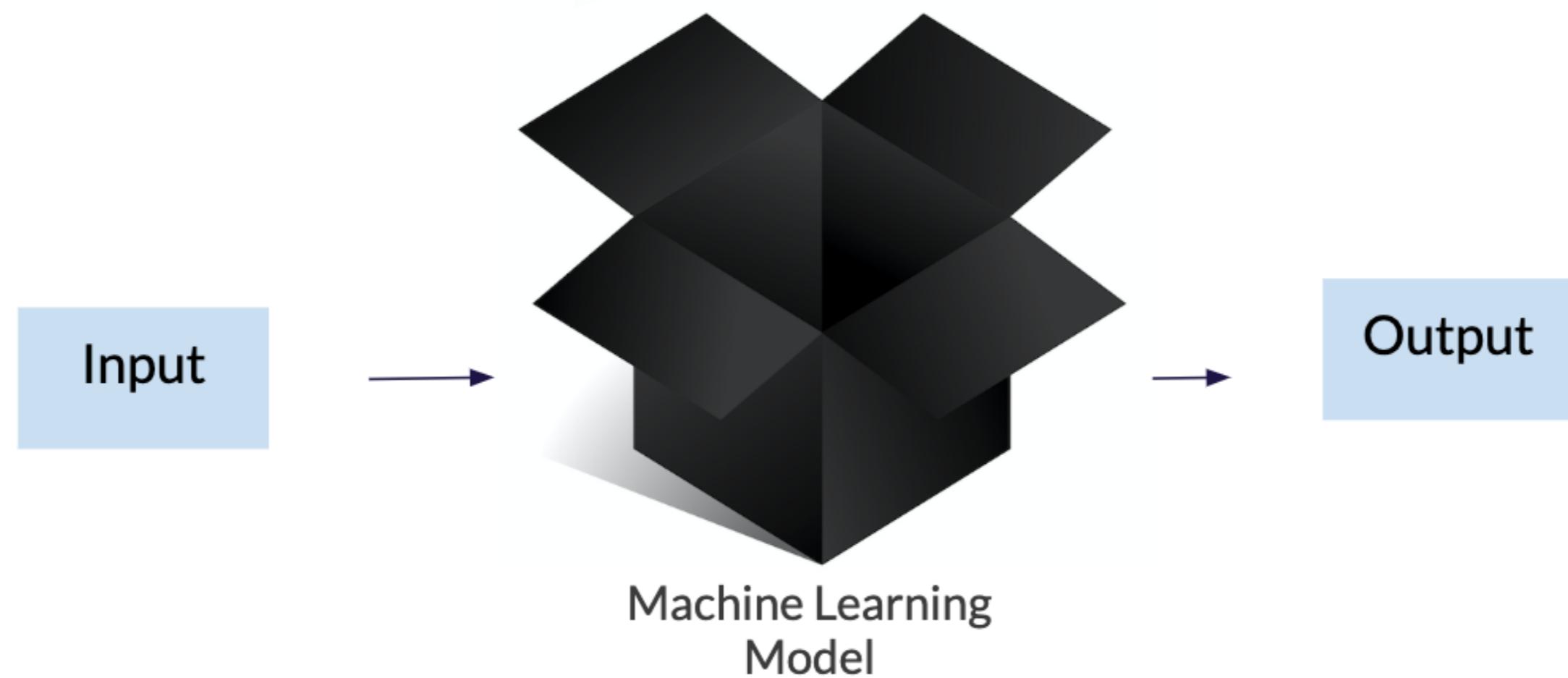
- High-quality data requires:
 - Data analysis
 - Review of outliers
 - Domain expertise
 - Documentation



Explainability



Explainability



- Transparency to increase trust, clarity, and understanding
- Use cases: business adoption, regulatory oversight, minimizing bias

Explainable AI

Black box

- Deep learning
- Better for "What?"
- Highly accurate predictions

Explainable AI

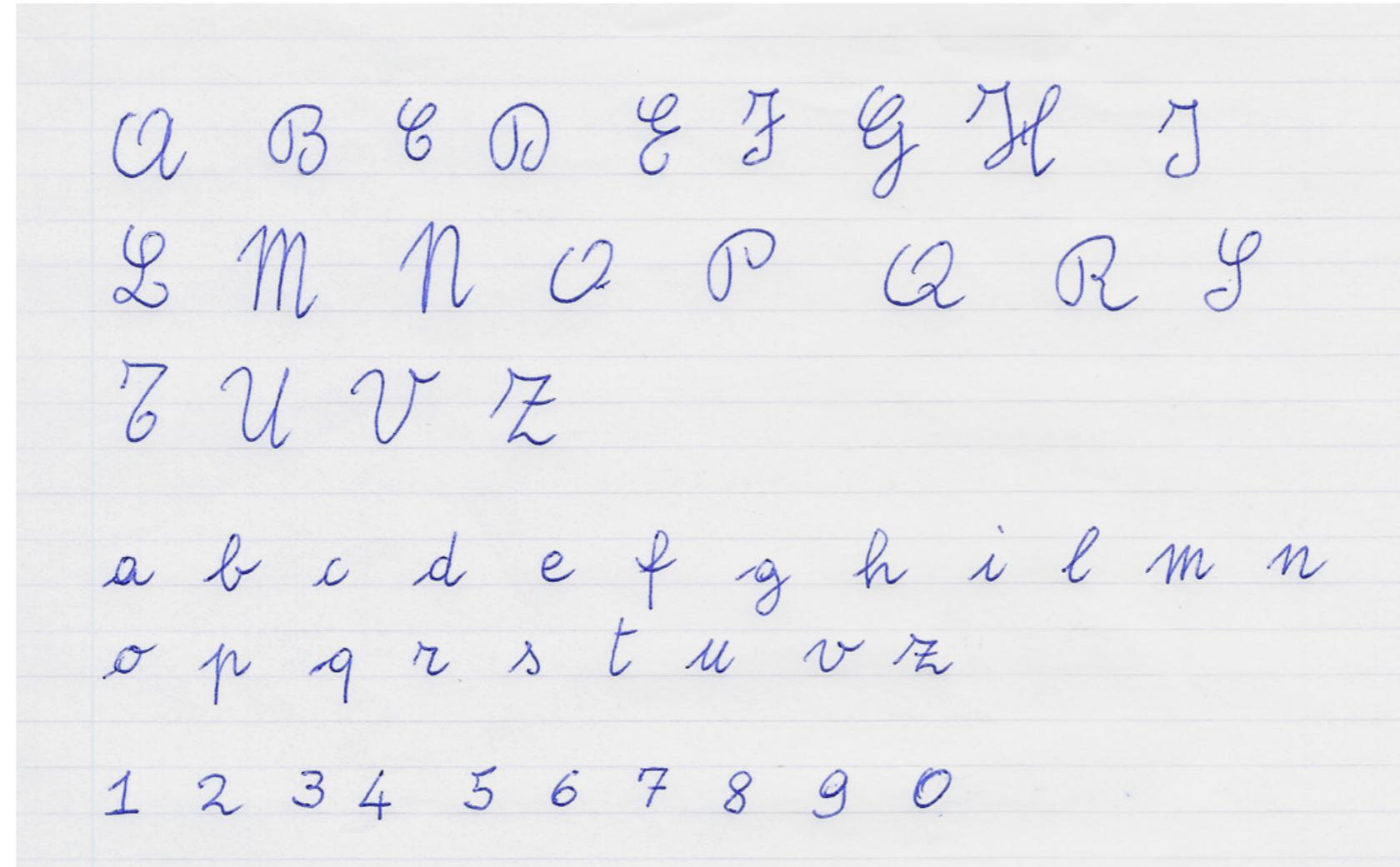
- Traditional machine learning
- Better for "Why?"
- Understandable by humans

Example: Explainable AI



- Prediction:** Will the patient get diabetes?
- Inference:** Why will this happen

Example: Inexplicable AI



Prediction only: Which letter is this likely to be?

Let's practice!

UNDERSTANDING MACHINE LEARNING

Congratulations!

UNDERSTANDING MACHINE LEARNING

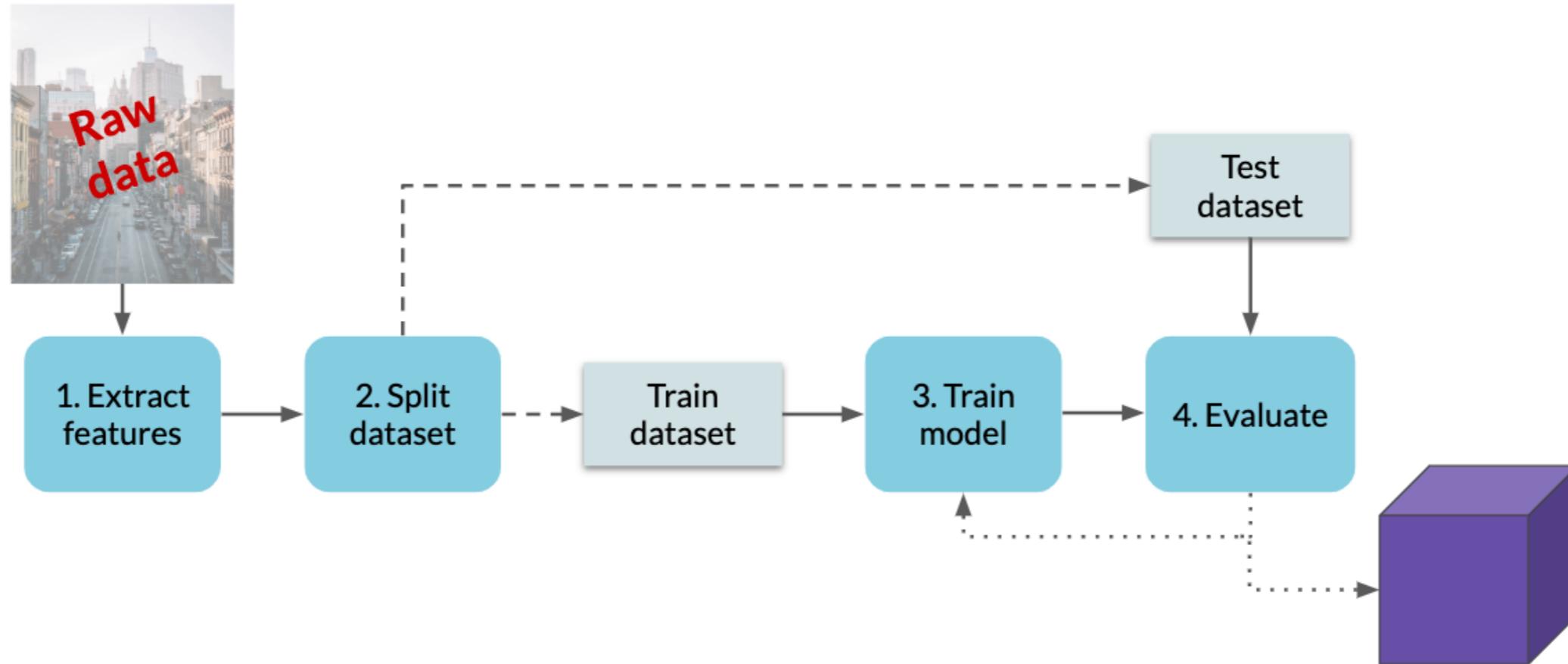


Lis Sulmont

Curriculum Manager, DataCamp

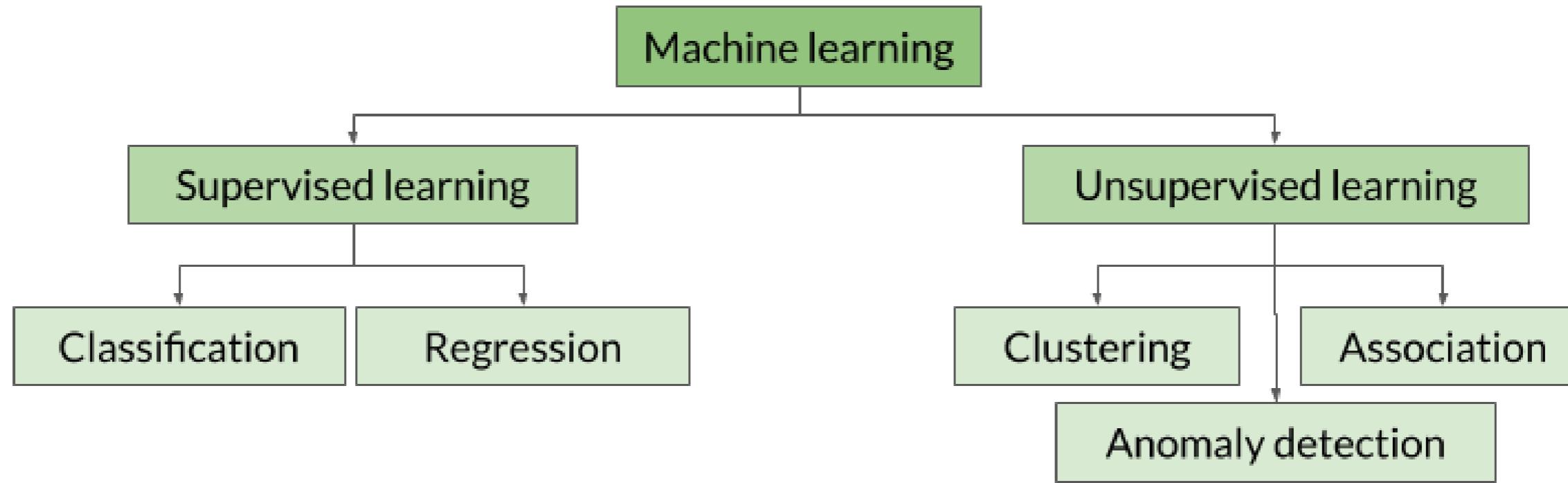
Chapter 1

- What is machine learning?
- Machine learning concepts and workflow



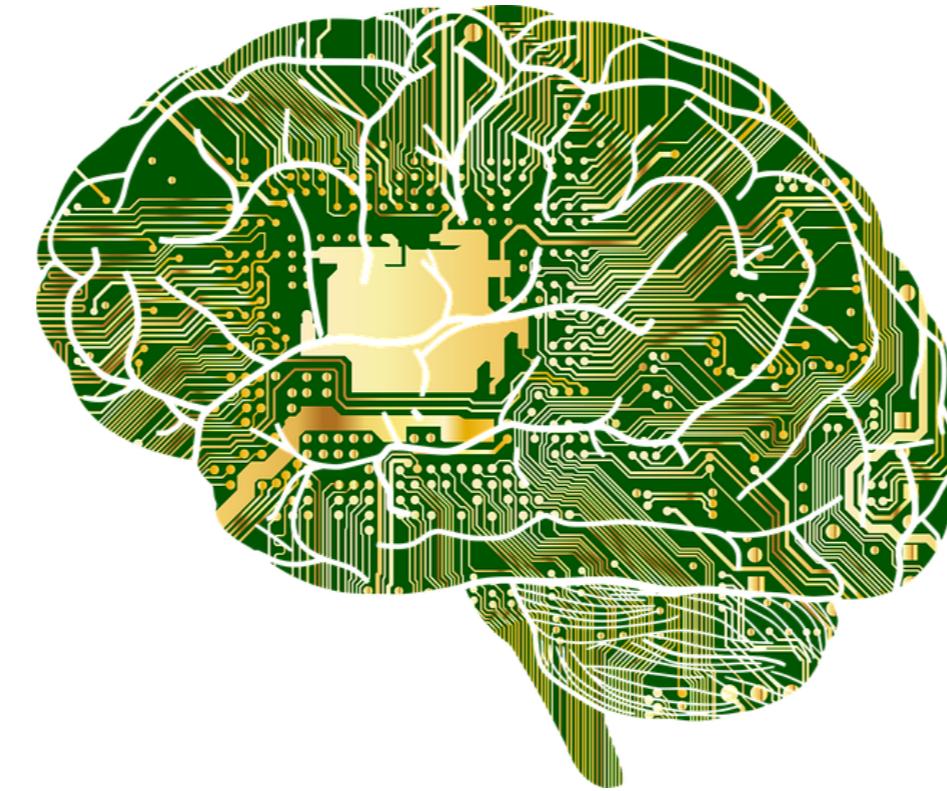
Chapter 2

- Different types of machine learning
- How we evaluate and improve machine learning models



Chapter 3

- Deep learning, including computer vision and natural language processing
- Limits of machine learning



What's next?



What's next?

- Machine Learning Scientist
- Machine Learning Fundamentals
- Supervised Machine Learning
- Unsupervised Machine Learning

Congrats!

UNDERSTANDING MACHINE LEARNING