Deep learning

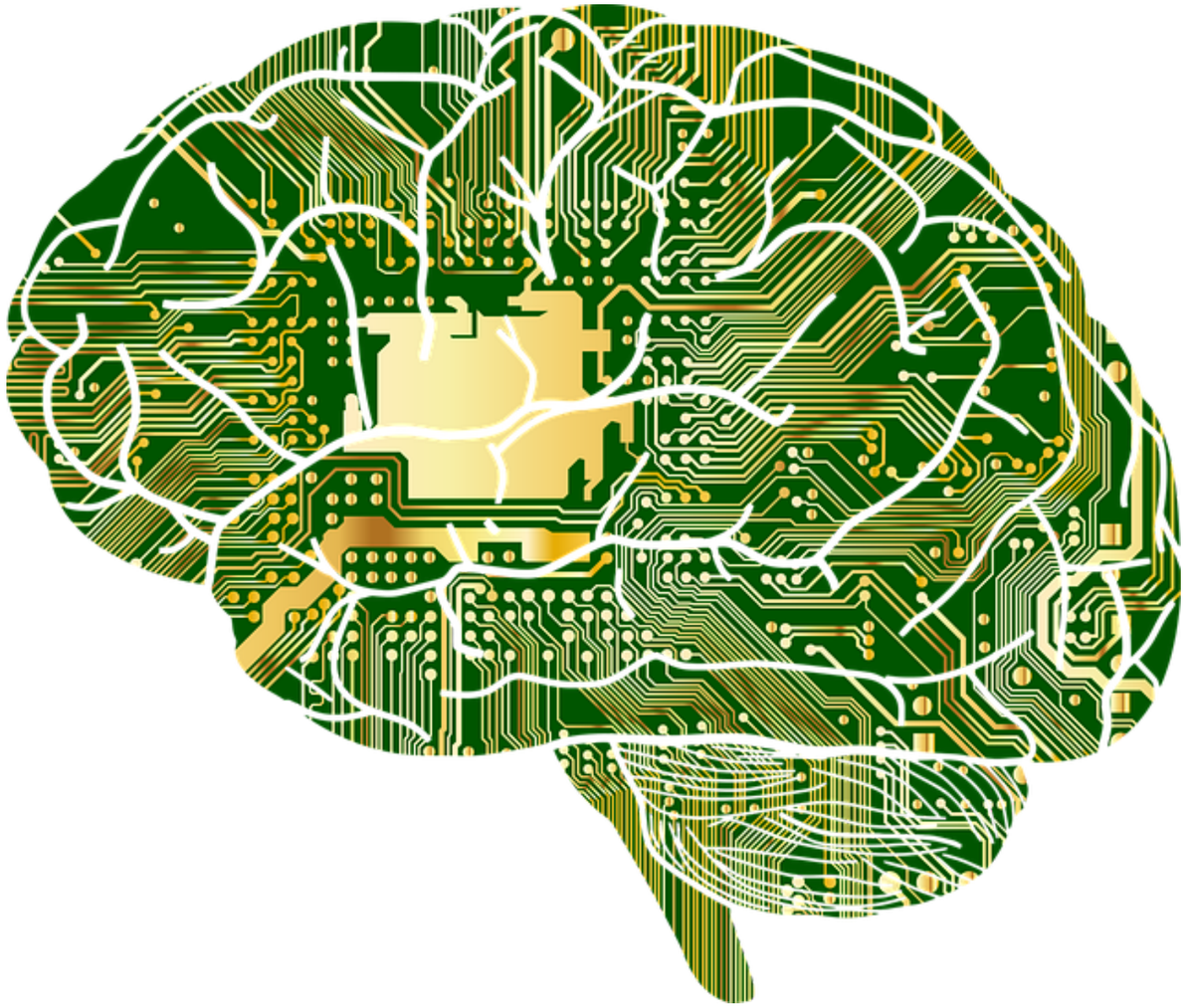
UNDERSTANDING MACHINE LEARNING

Sara Billen
Curriculum Manager, DataCamp

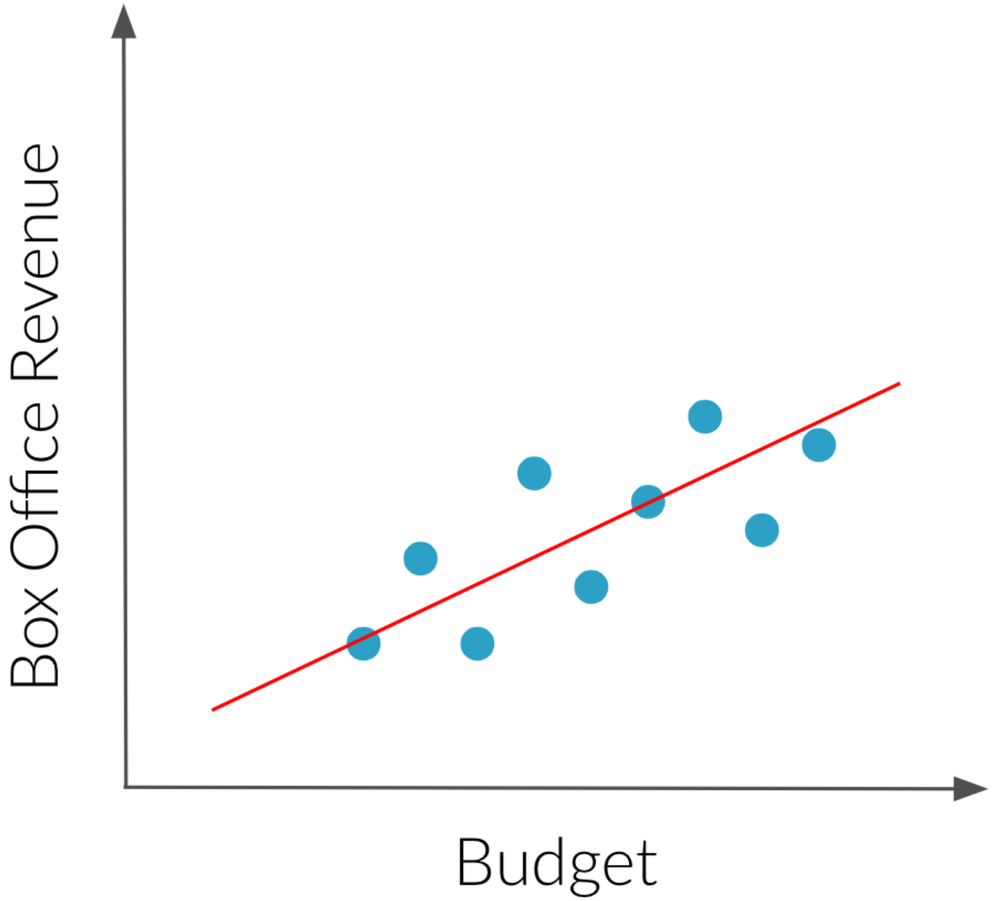
datacamp

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

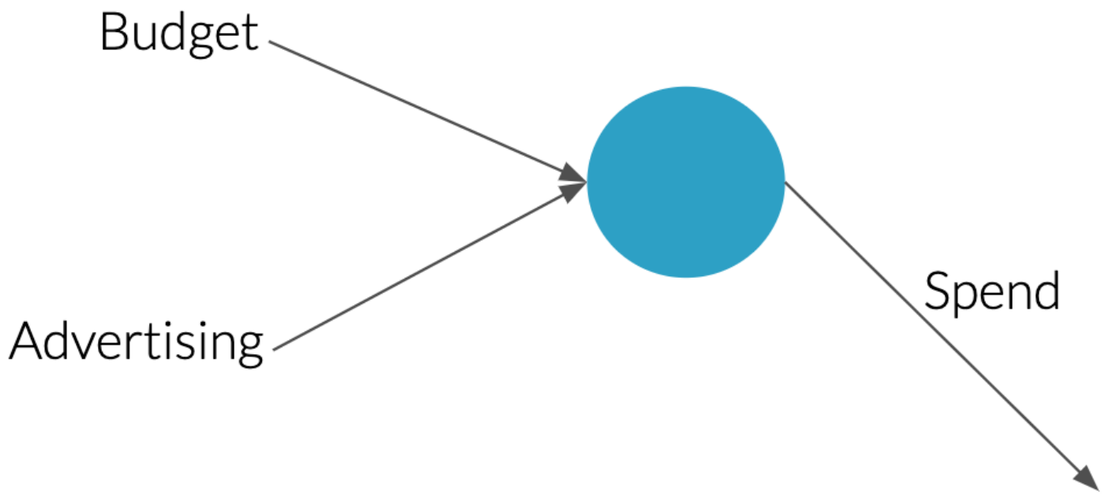
Budget

Advertising

Star Power

Timing

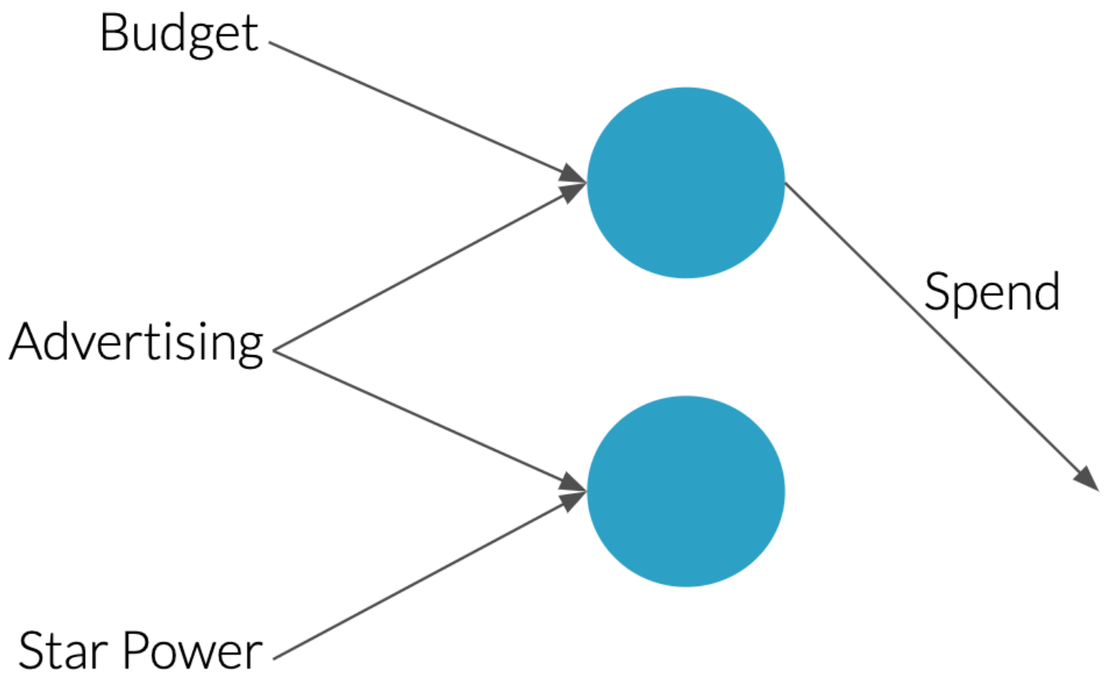
Predicting box office revenue



Star Power

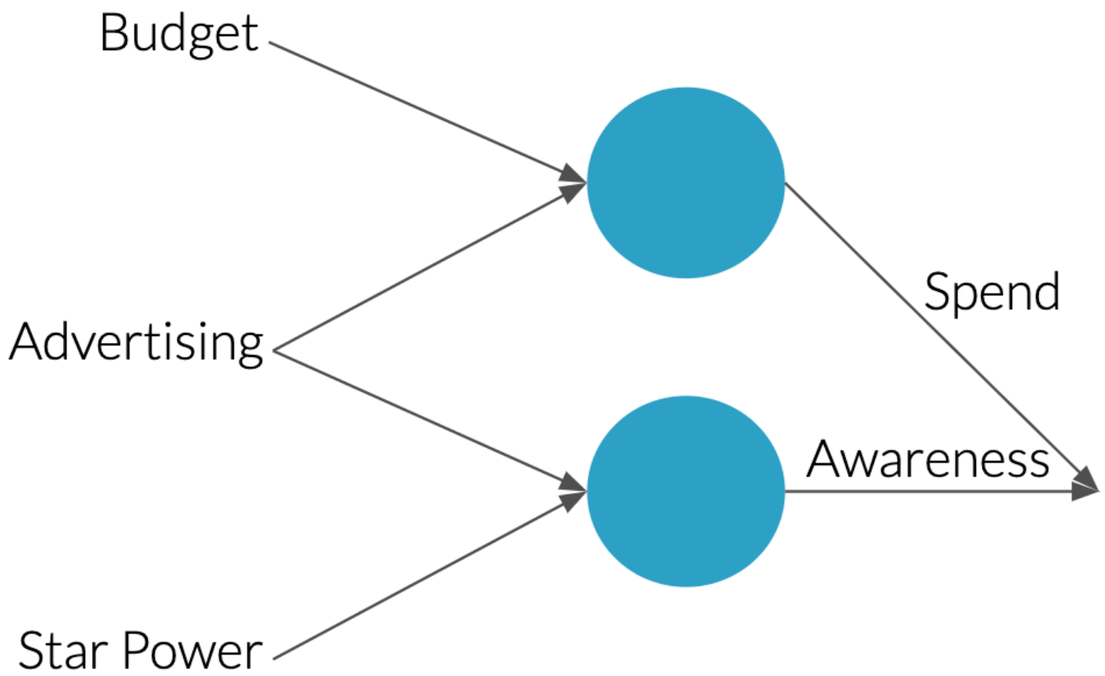
Timing

Predicting box office revenue



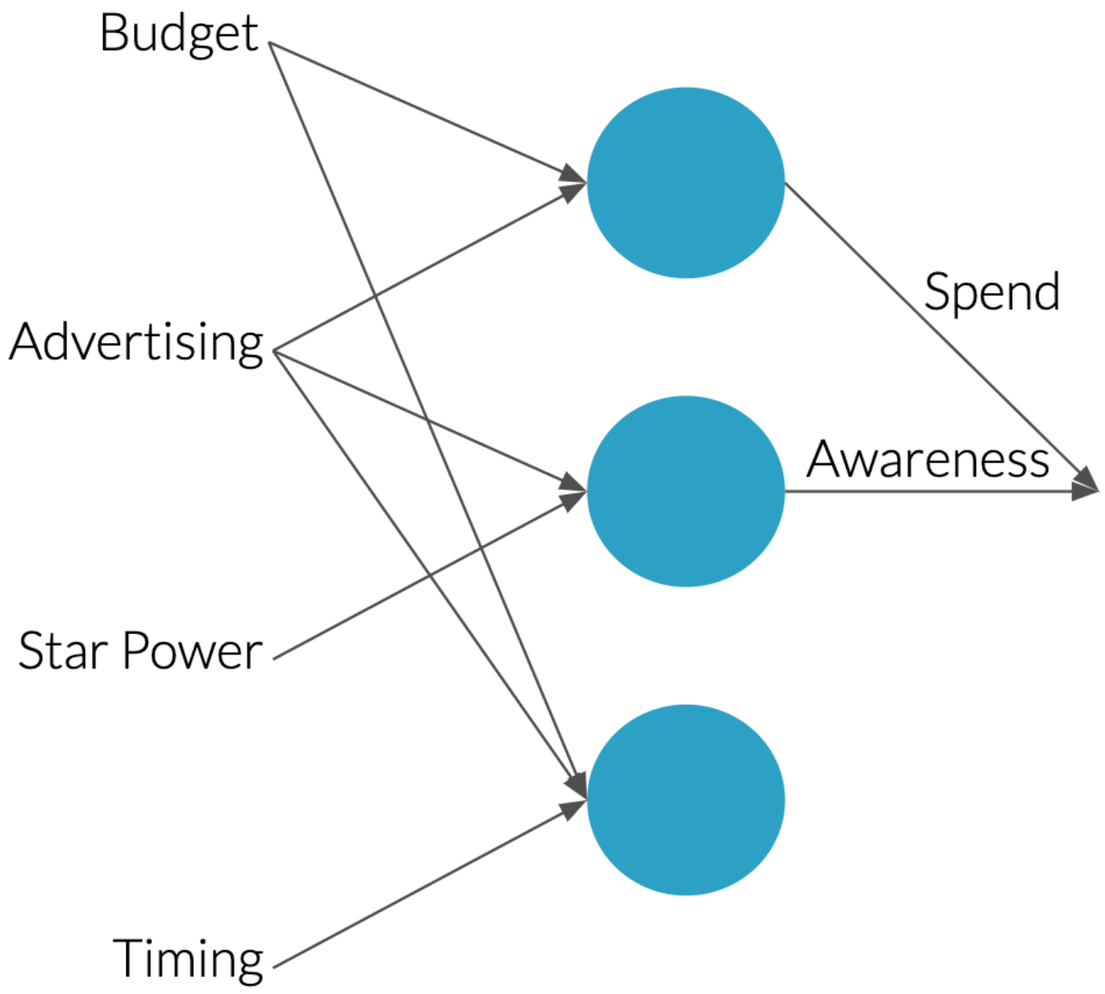
Timing

Predicting box office revenue

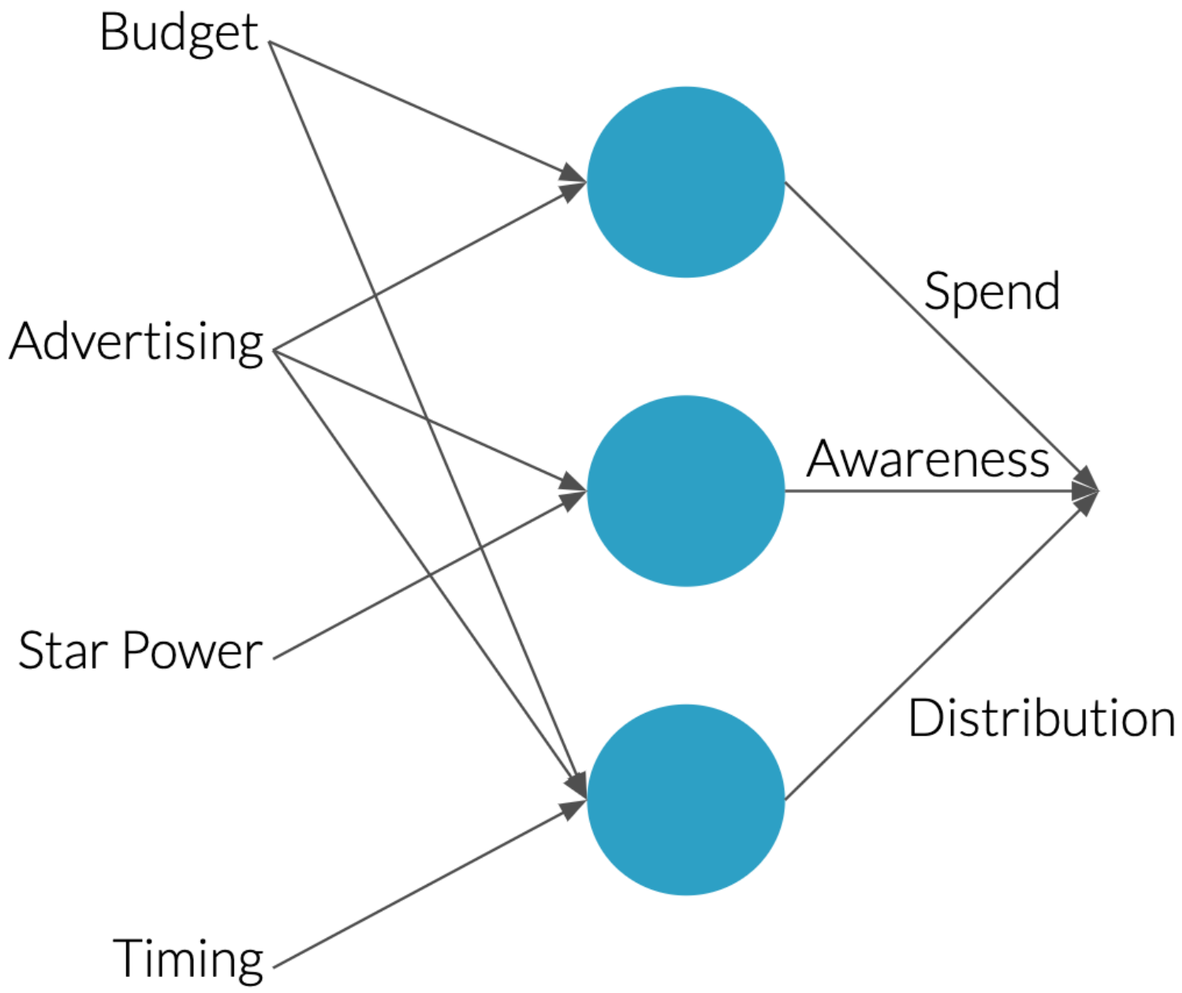


Timing

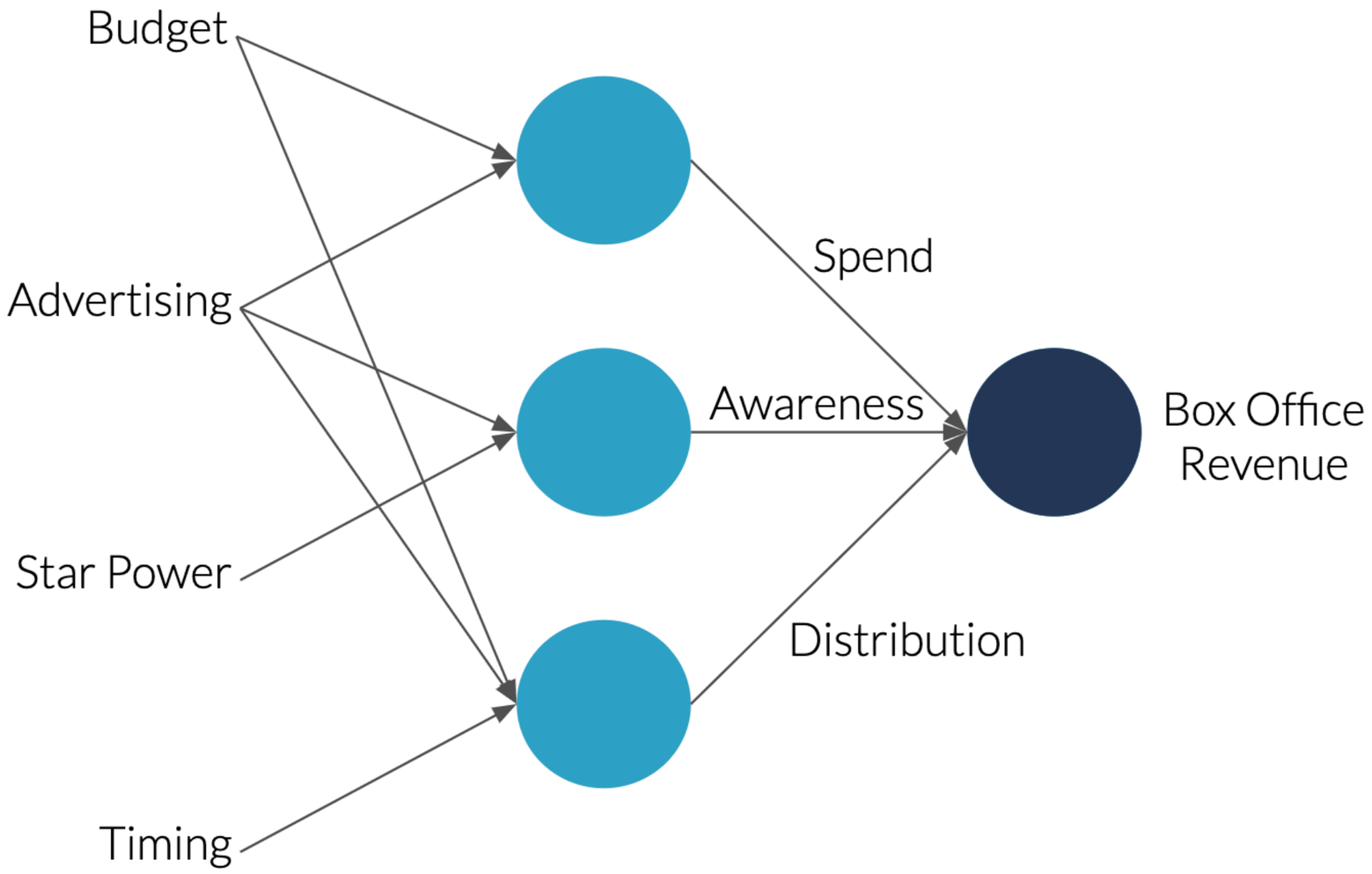
Predicting box office revenue



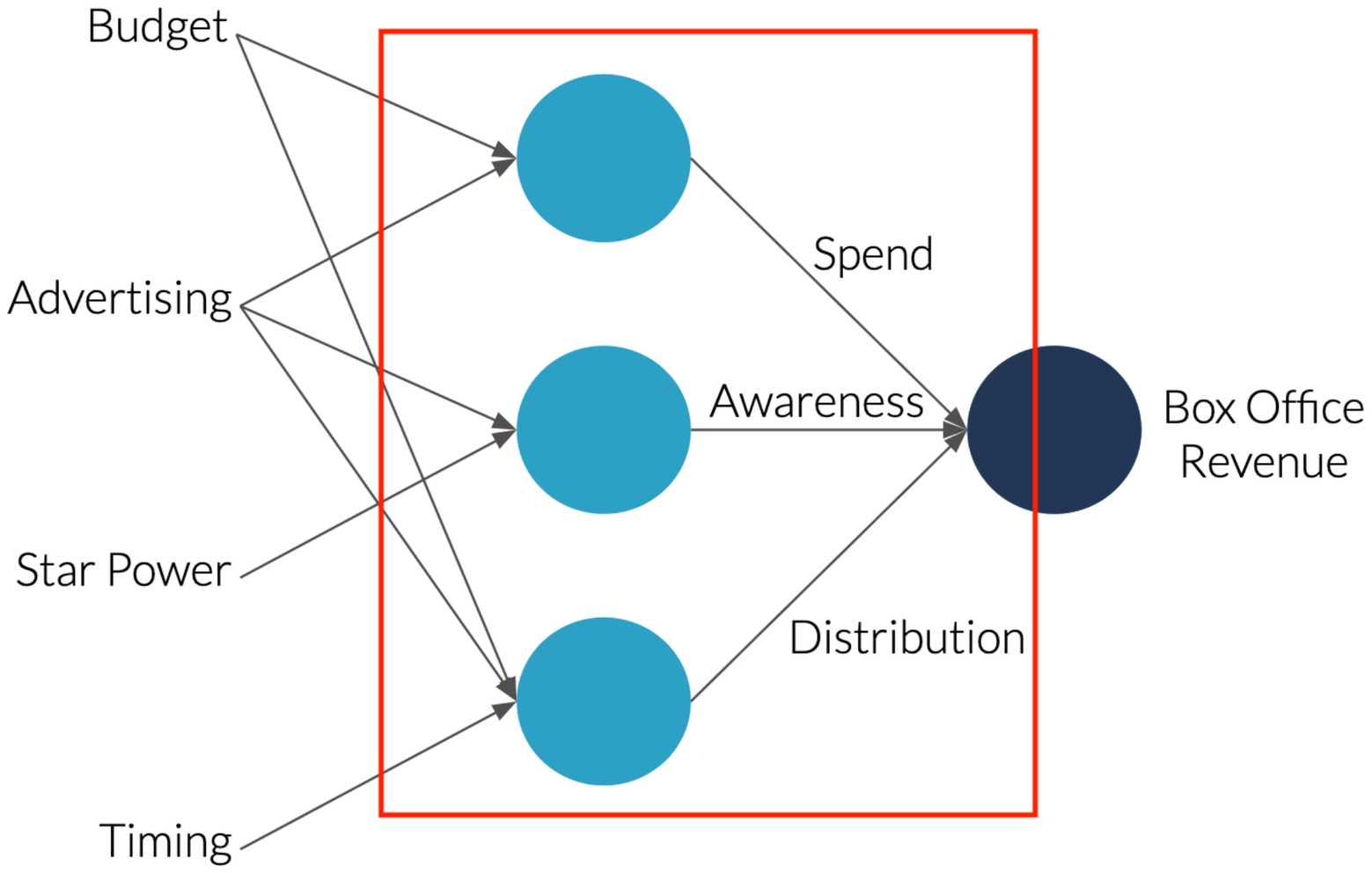
Predicting box office revenue



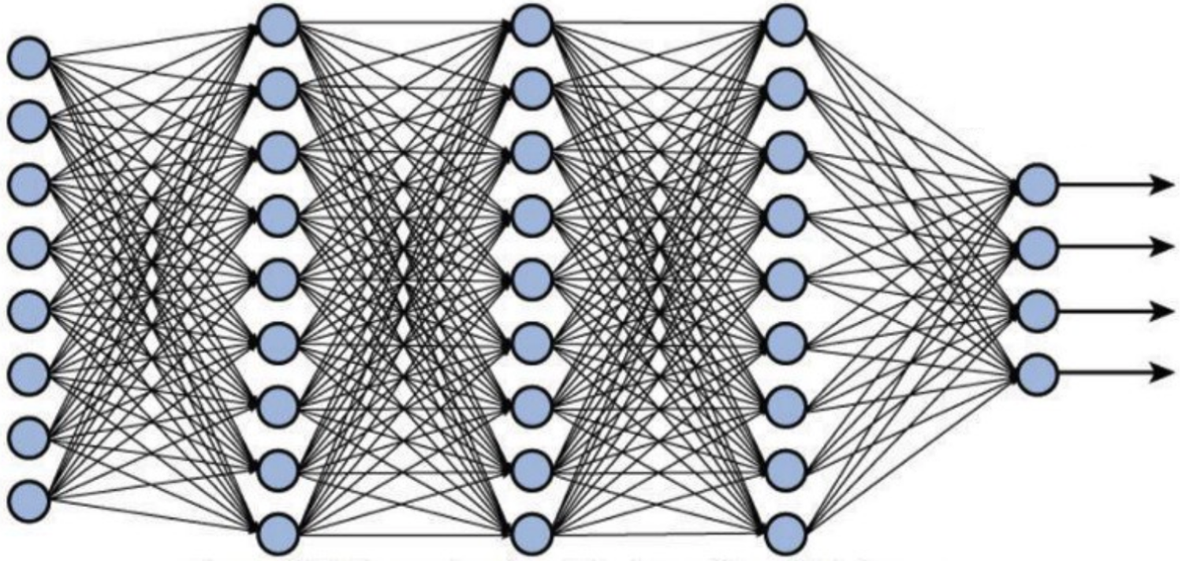
Predicting box office revenue



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

datacamp

Computer vision

Helps computers see and understand the content of digital images

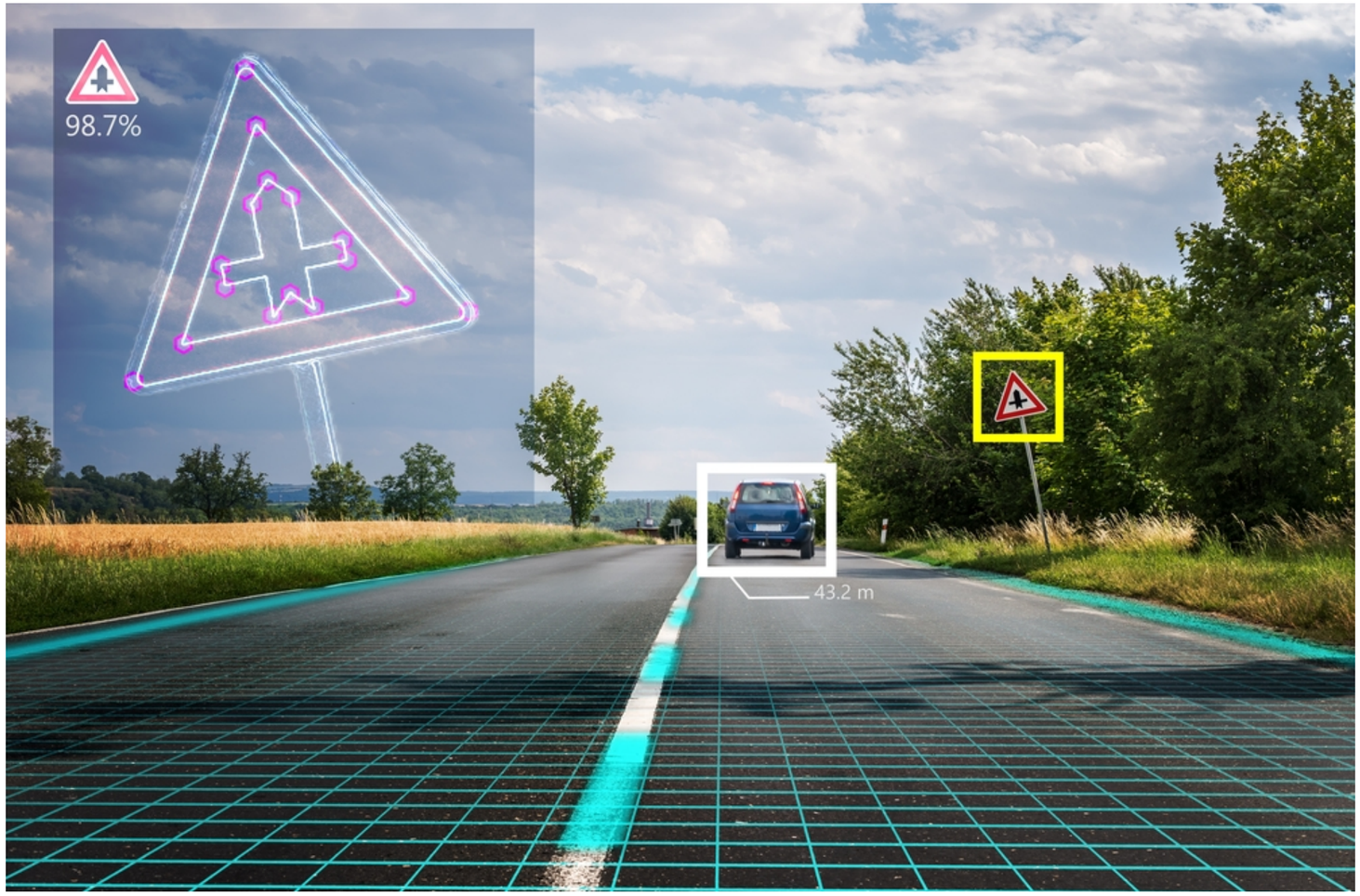
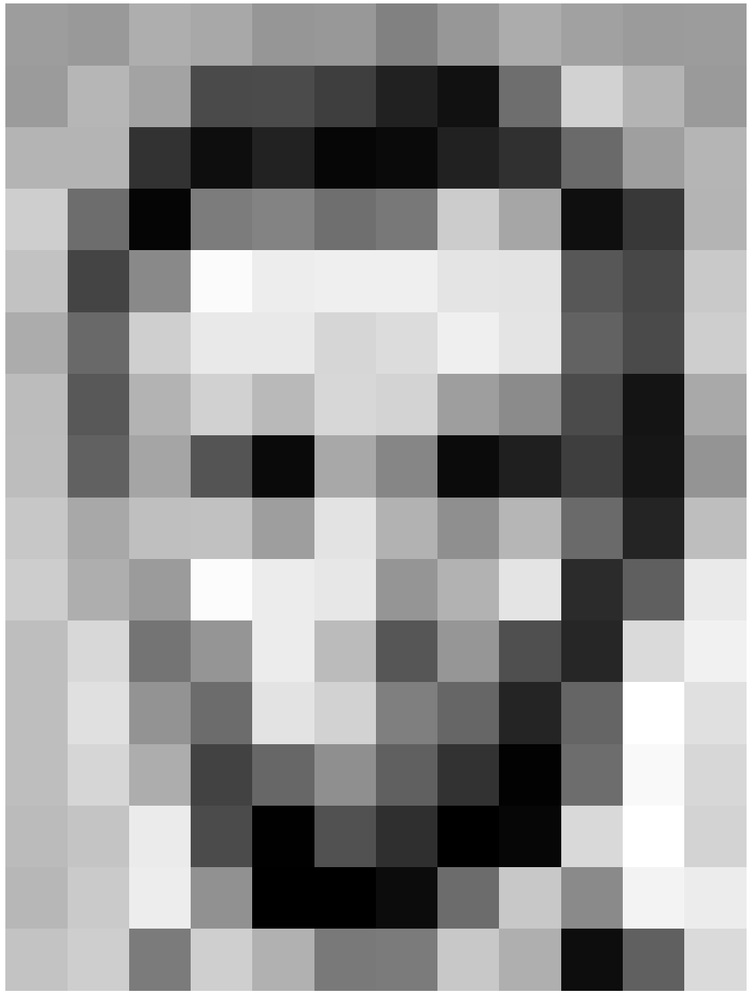


Image data

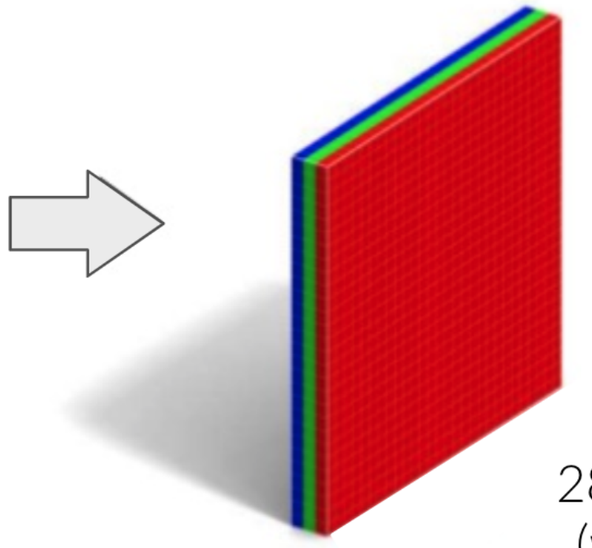


157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	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	96	50	2	109	249	215
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	175	13	96	218

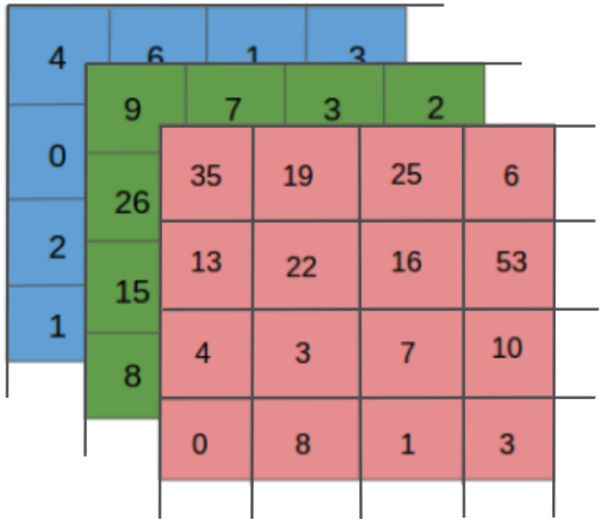
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	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	96	50	2	109	249	215
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	175	13	96	218

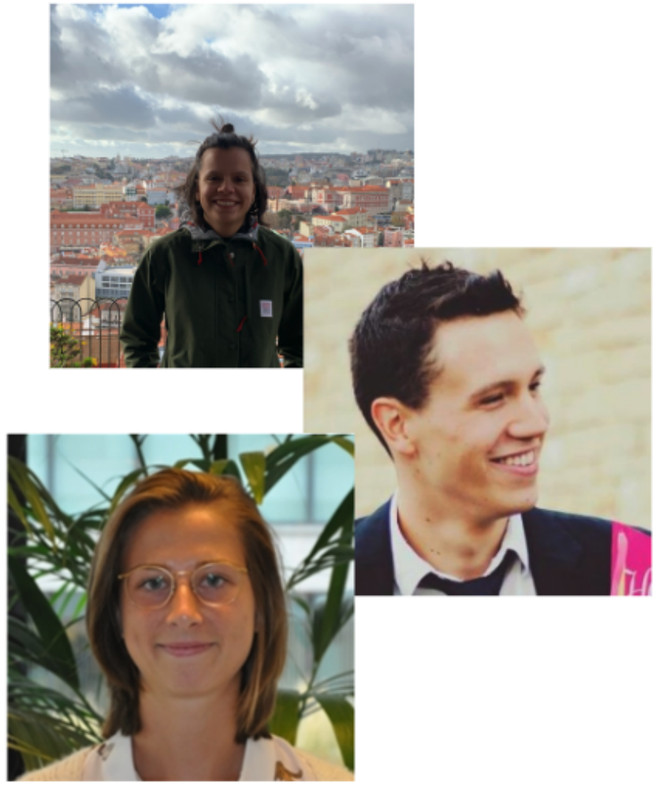
Image data

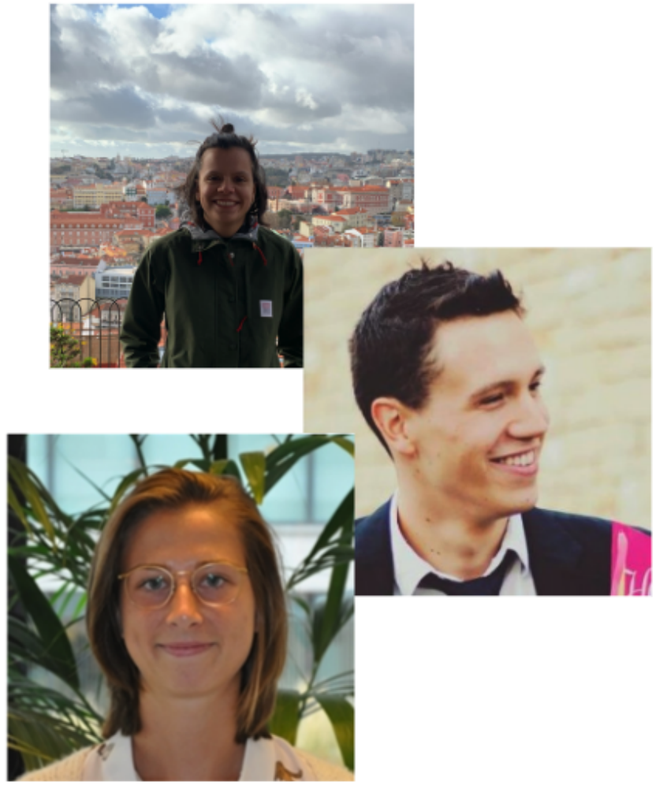
color image
(RGB)

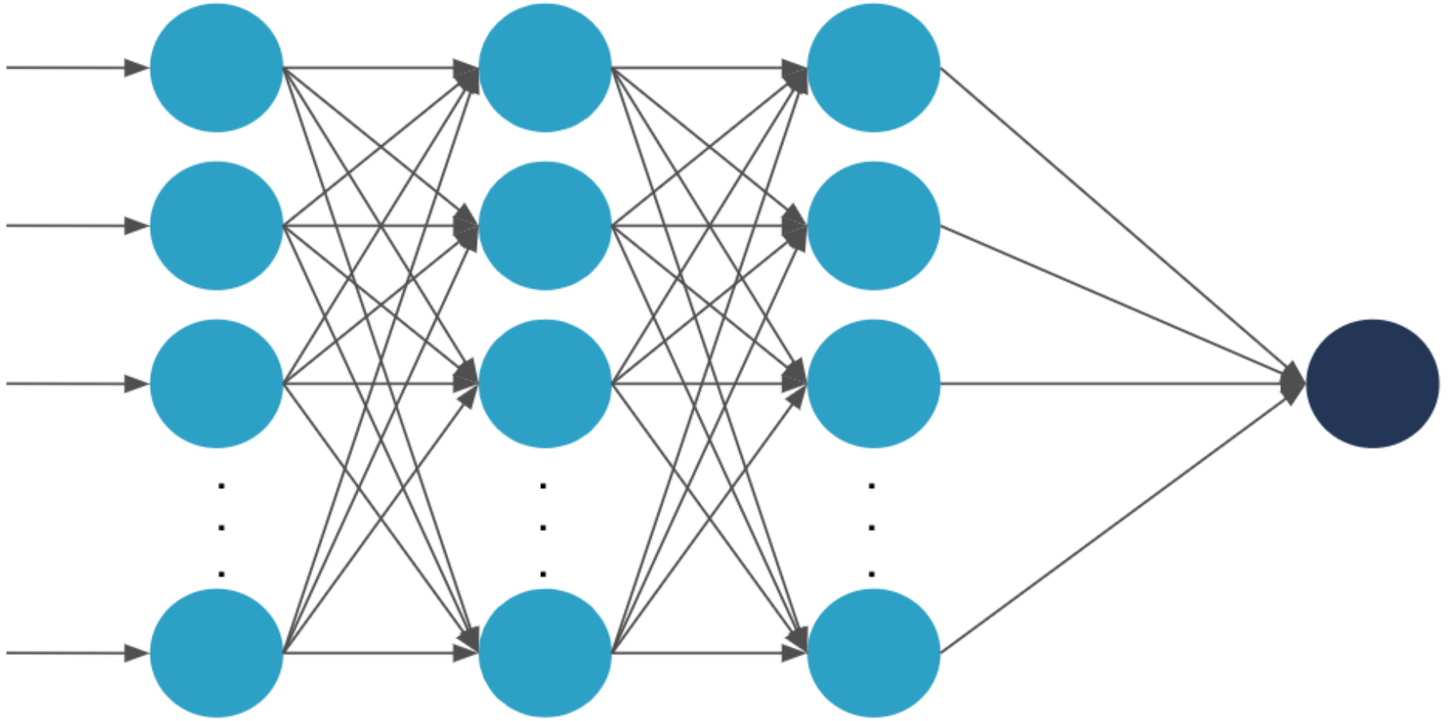


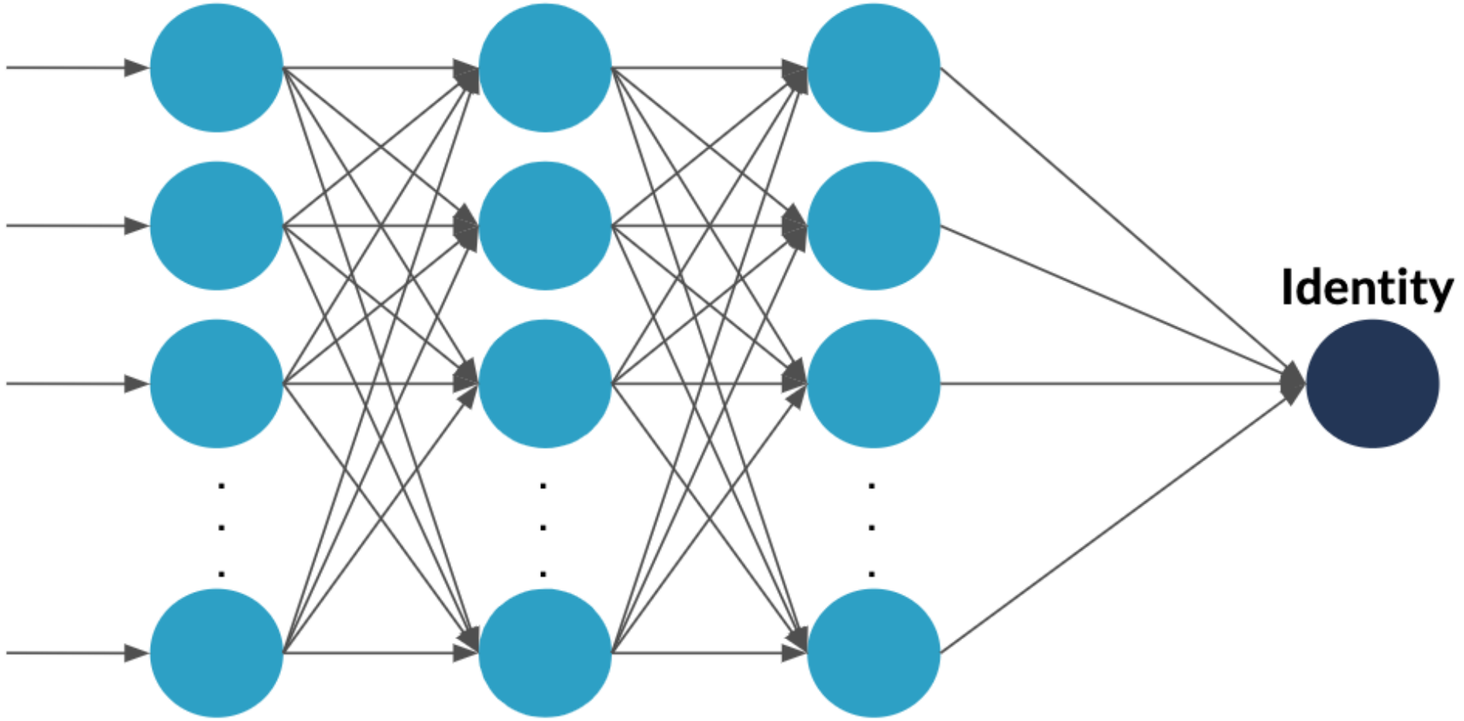
3 channels
(RGB)

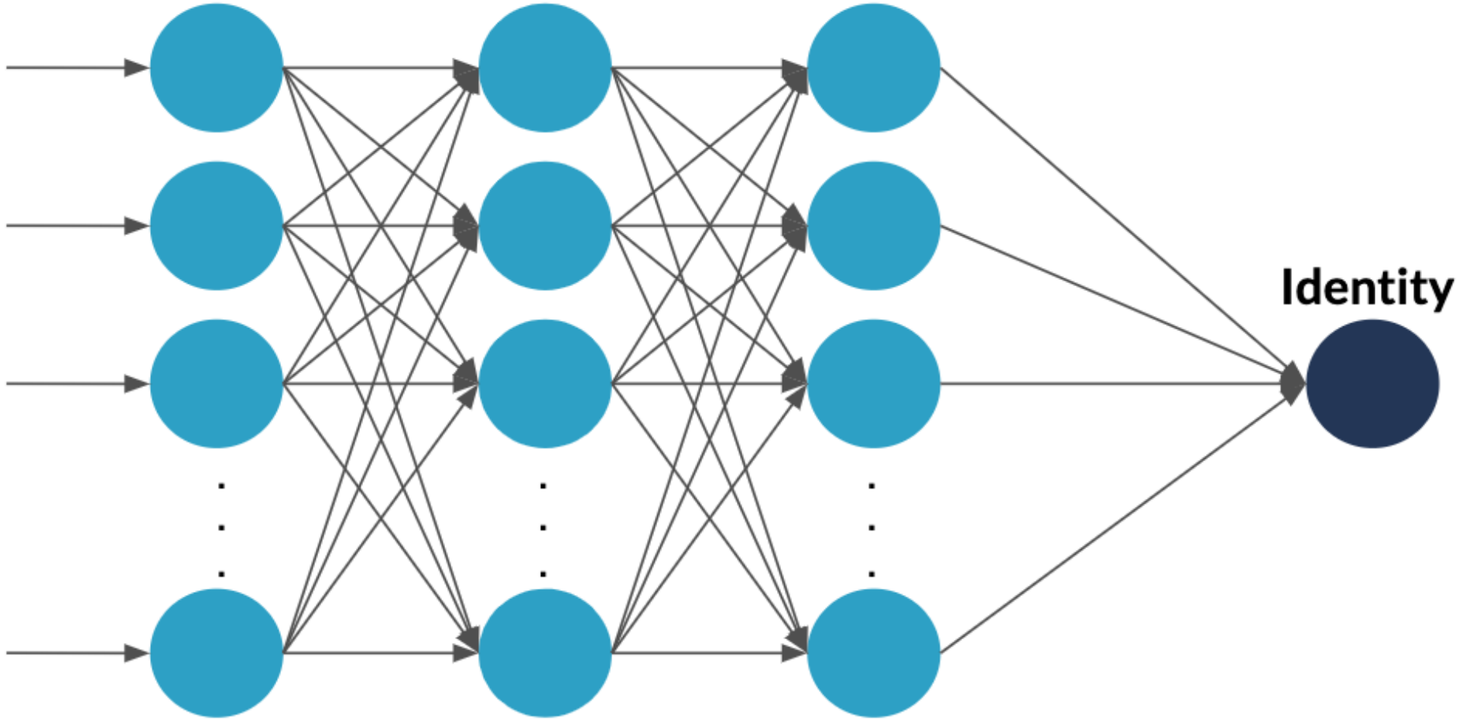


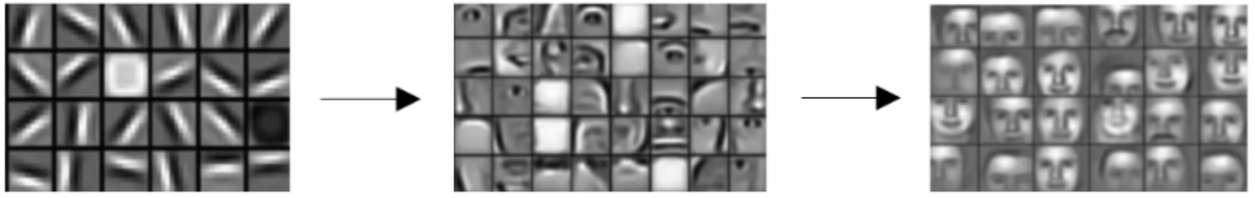




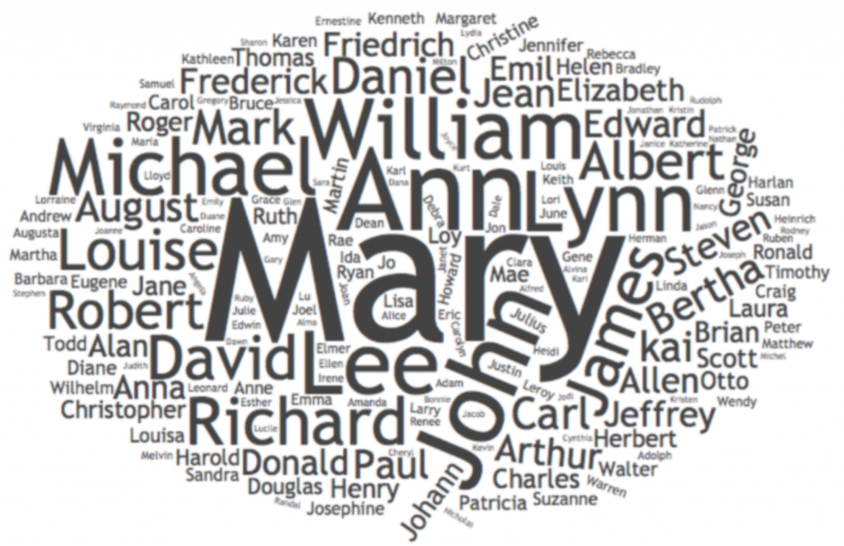


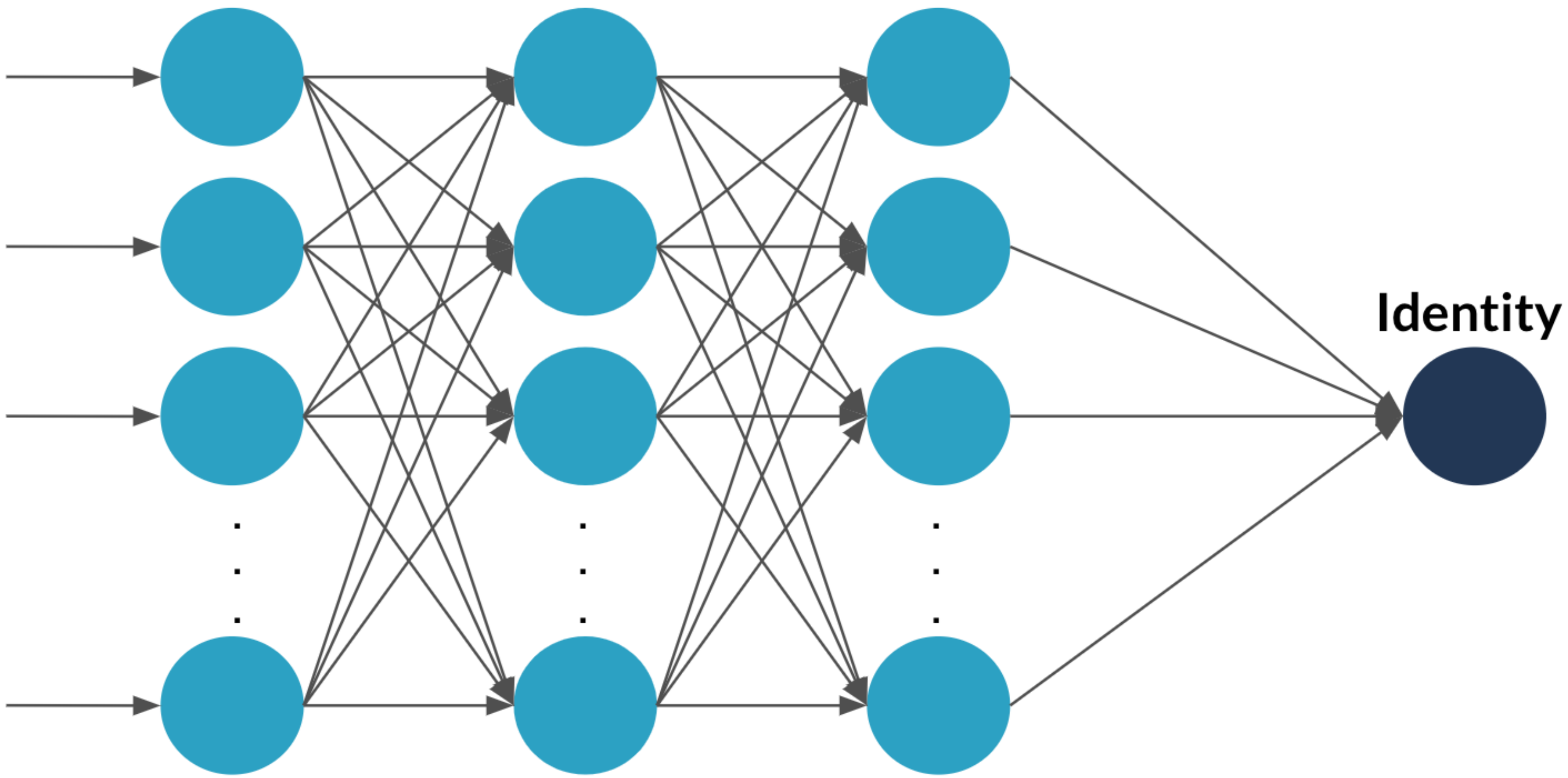

Identity
Lis
Hadrien
Sara


Identity
Lis
Hadrien
Sara



Training the neural network





Applications

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

Let's practice!

UNDERSTANDING MACHINE LEARNING

Natural Language Processing

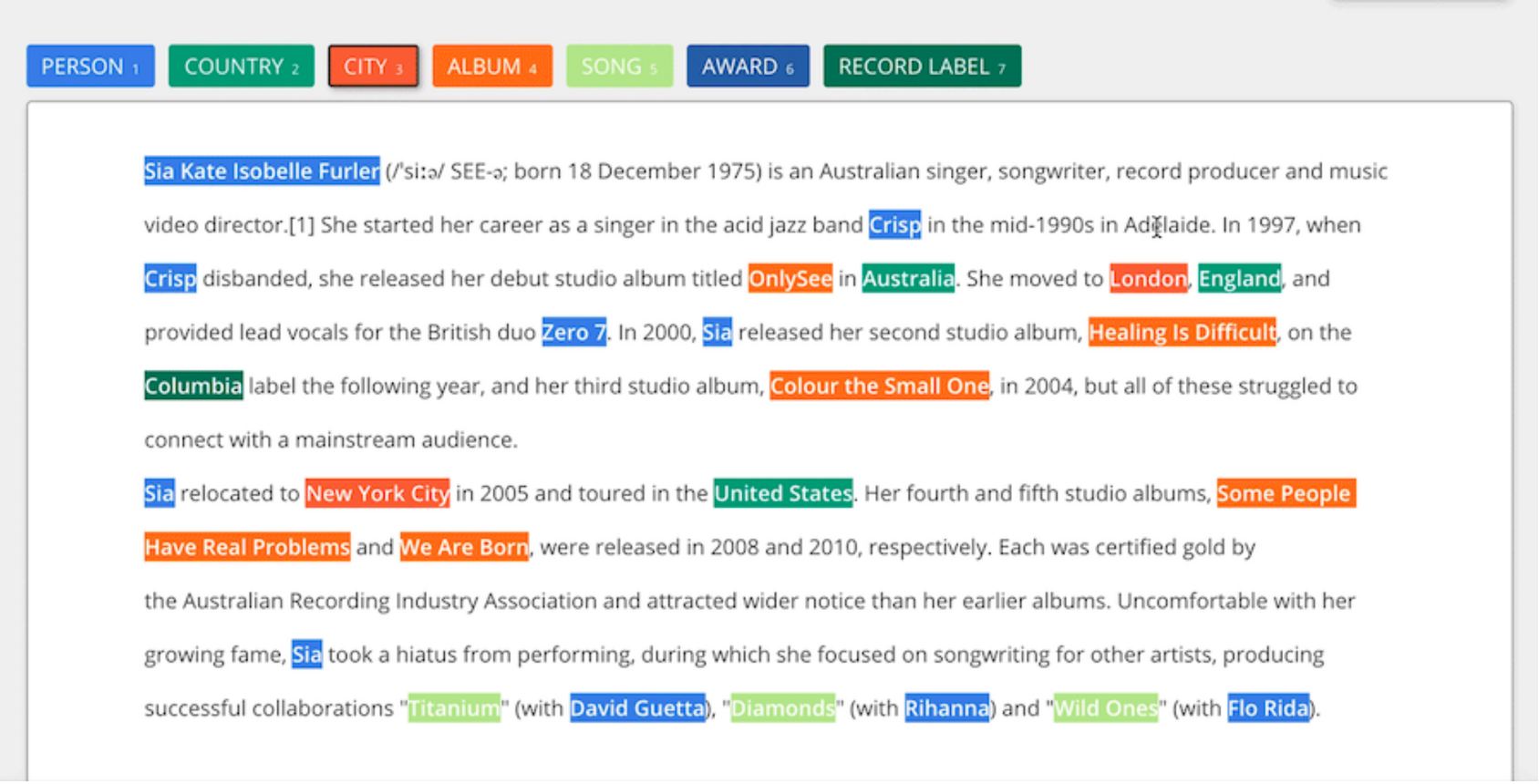
UNDERSTANDING MACHINE LEARNING

Sara Billen
Curriculum Manager at DataCamp

datacamp

Natural Language Processing (NLP)

The ability for computers to understand the meaning of human language

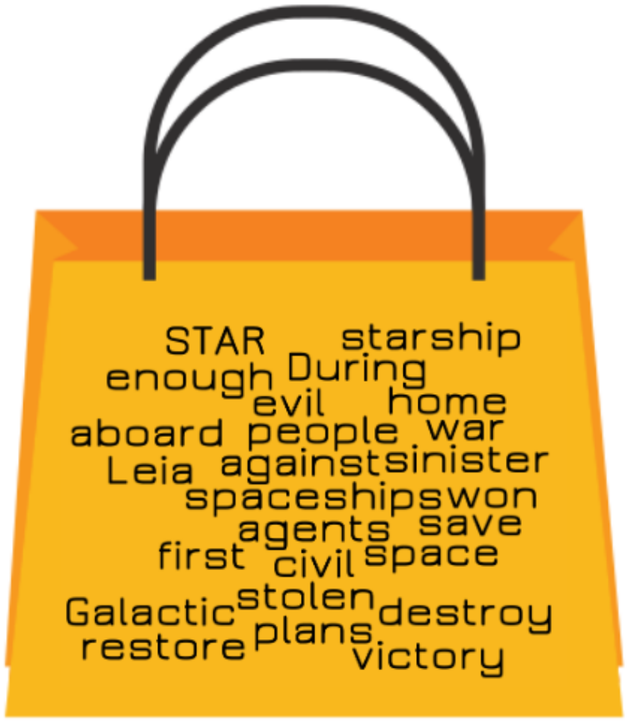


Bag of words

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"

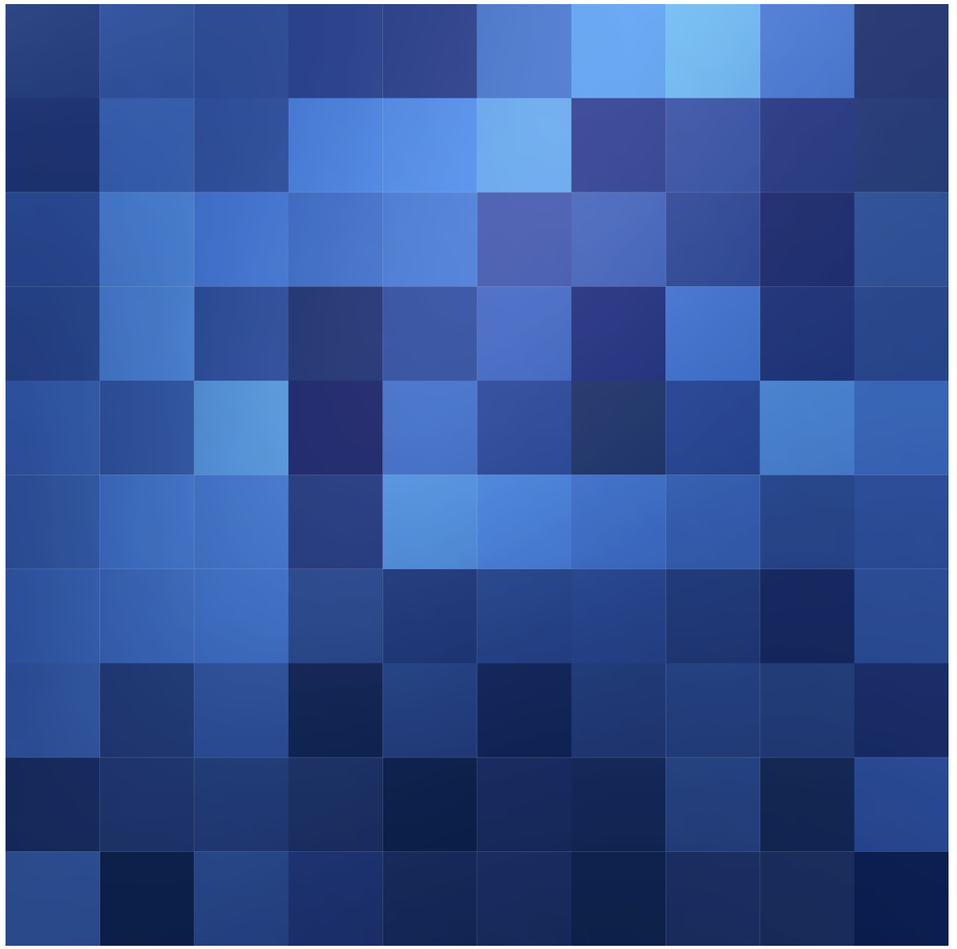
Word	Count
That	1
book	1
is	1
not	1
great	1

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

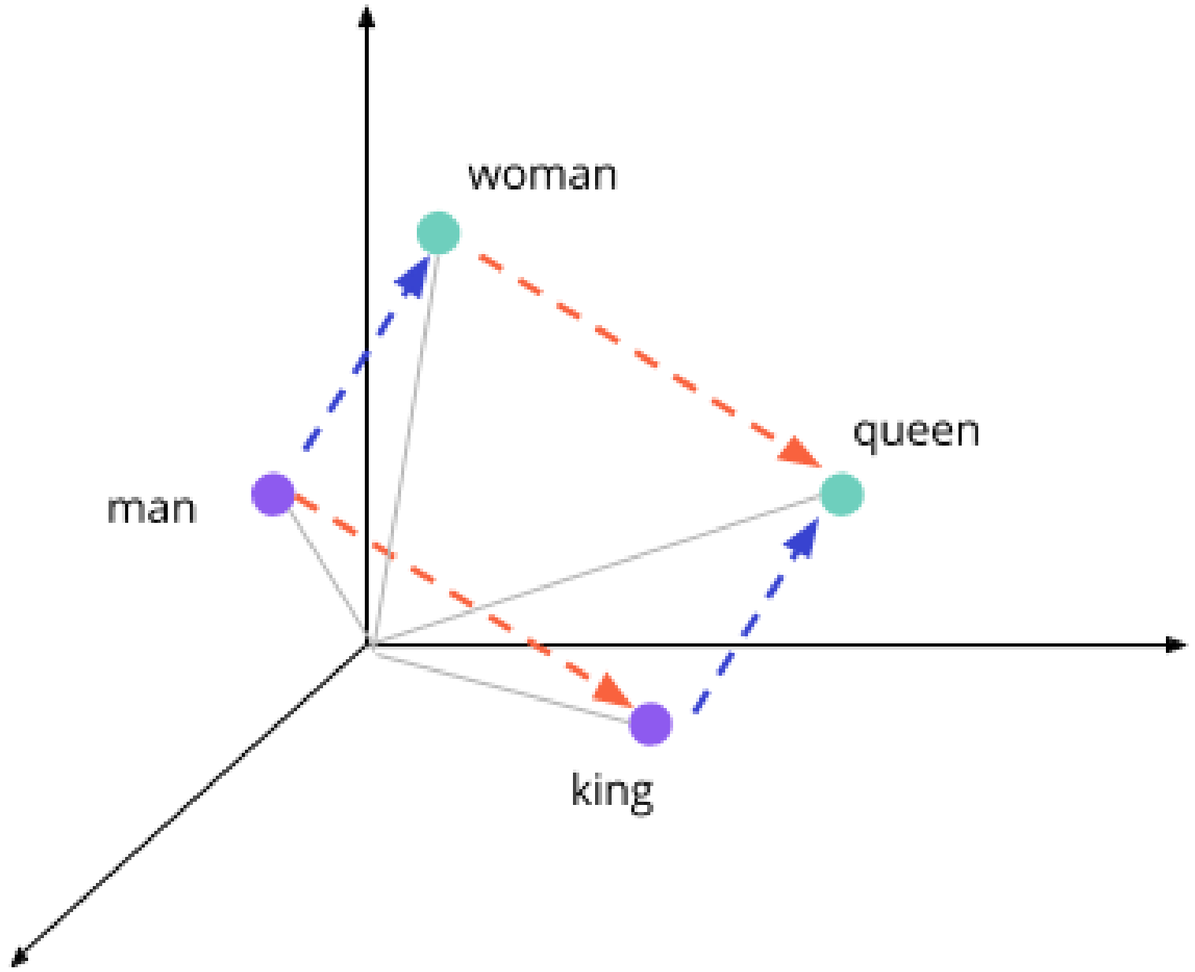


Word embeddings

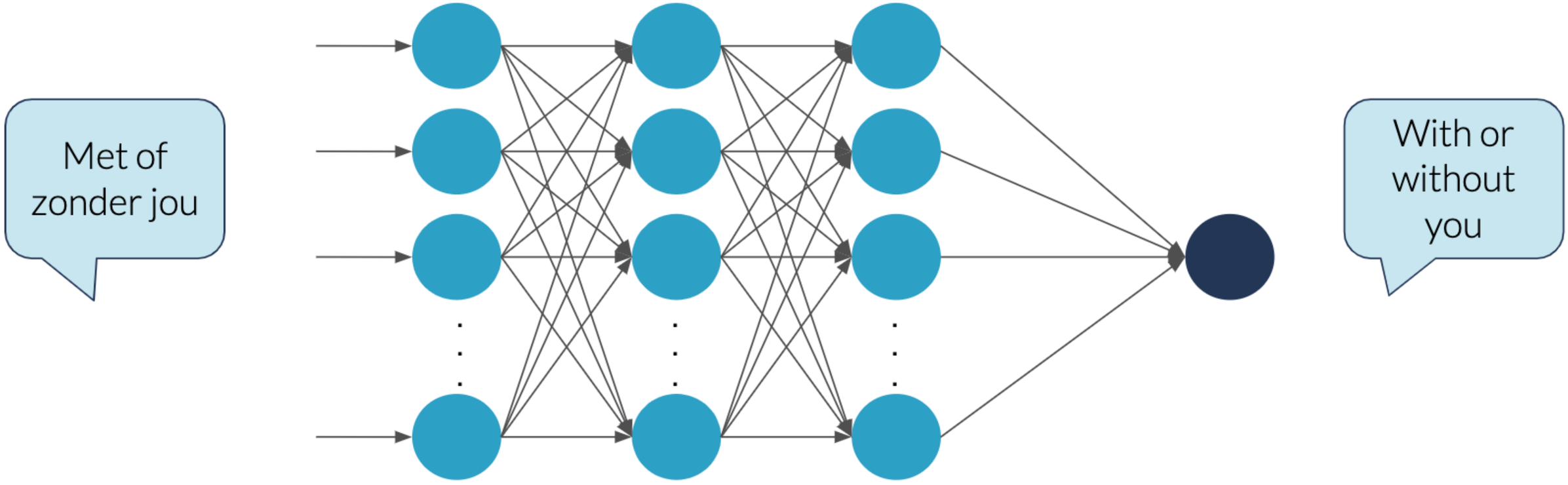
Word embeddings

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

```
king - man + woman = queen
```

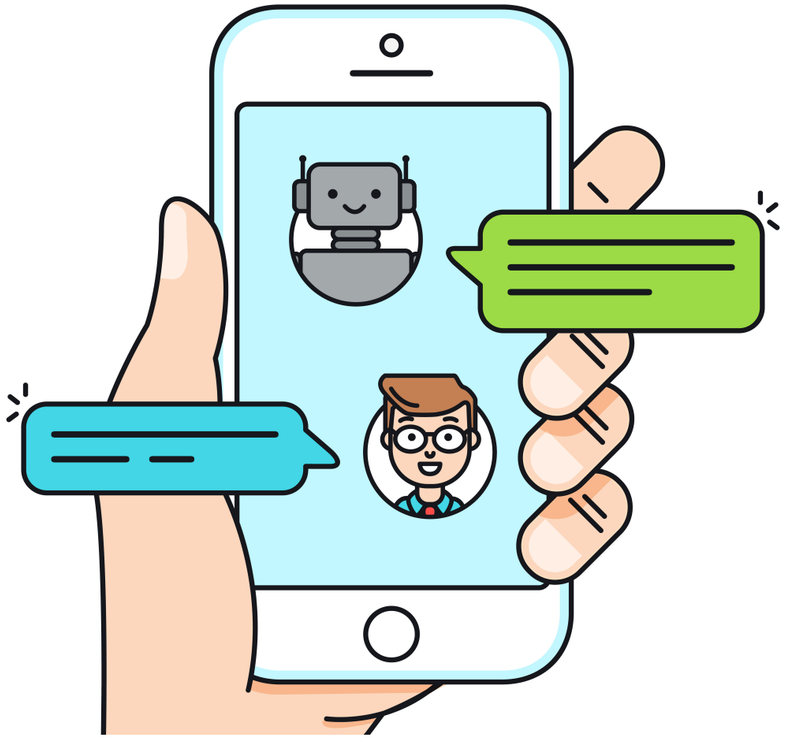


Language translation



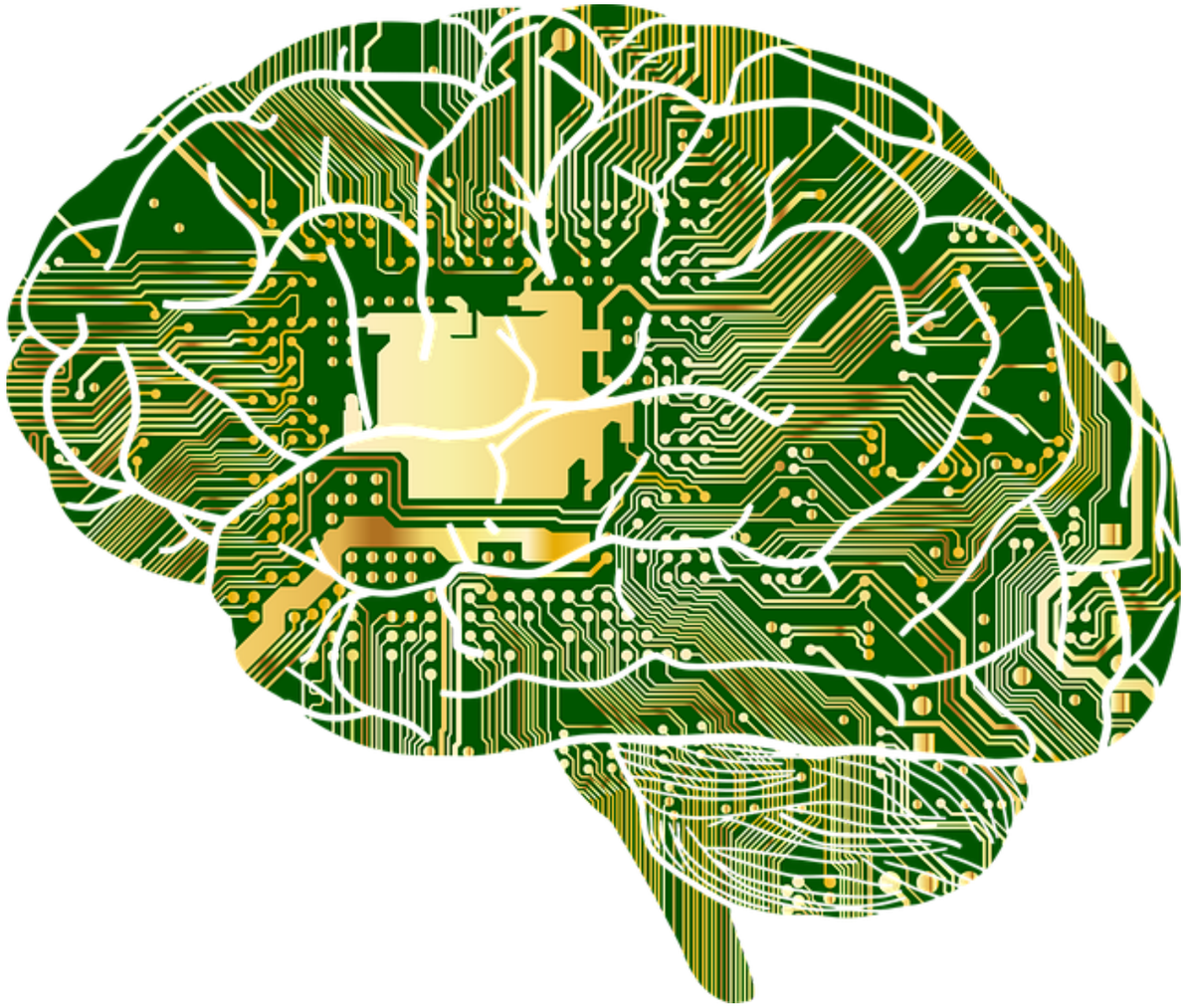
Applications

- 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!

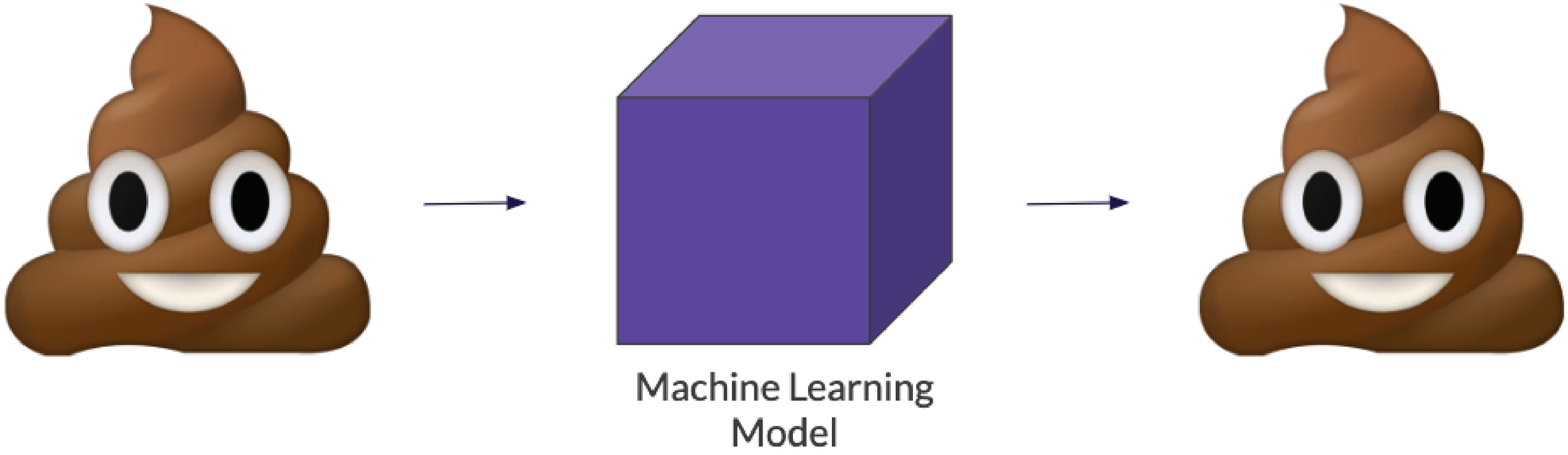
UNDERSTANDING MACHINE LEARNING

Limits of machine learning

UNDERSTANDING MACHINE LEARNING

Sara Billen
Curriculum Manager, DataCamp

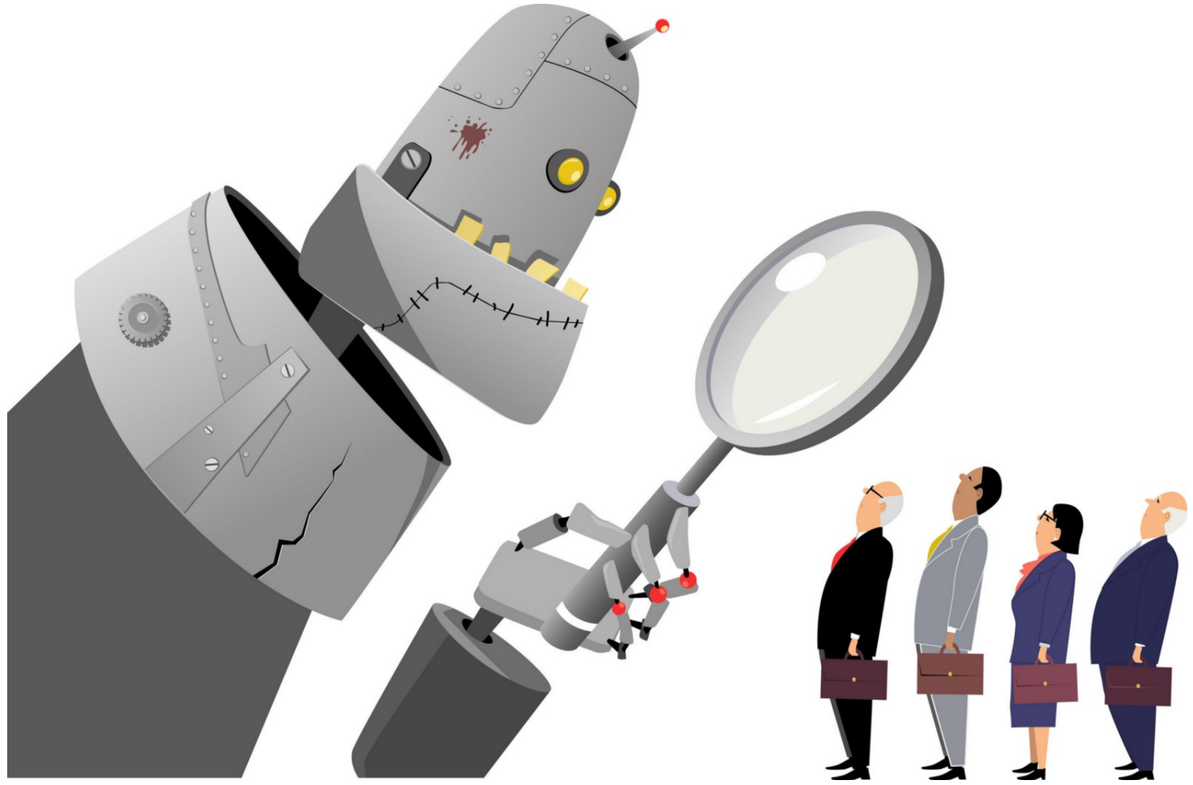
Data quality



- 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

Beware

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

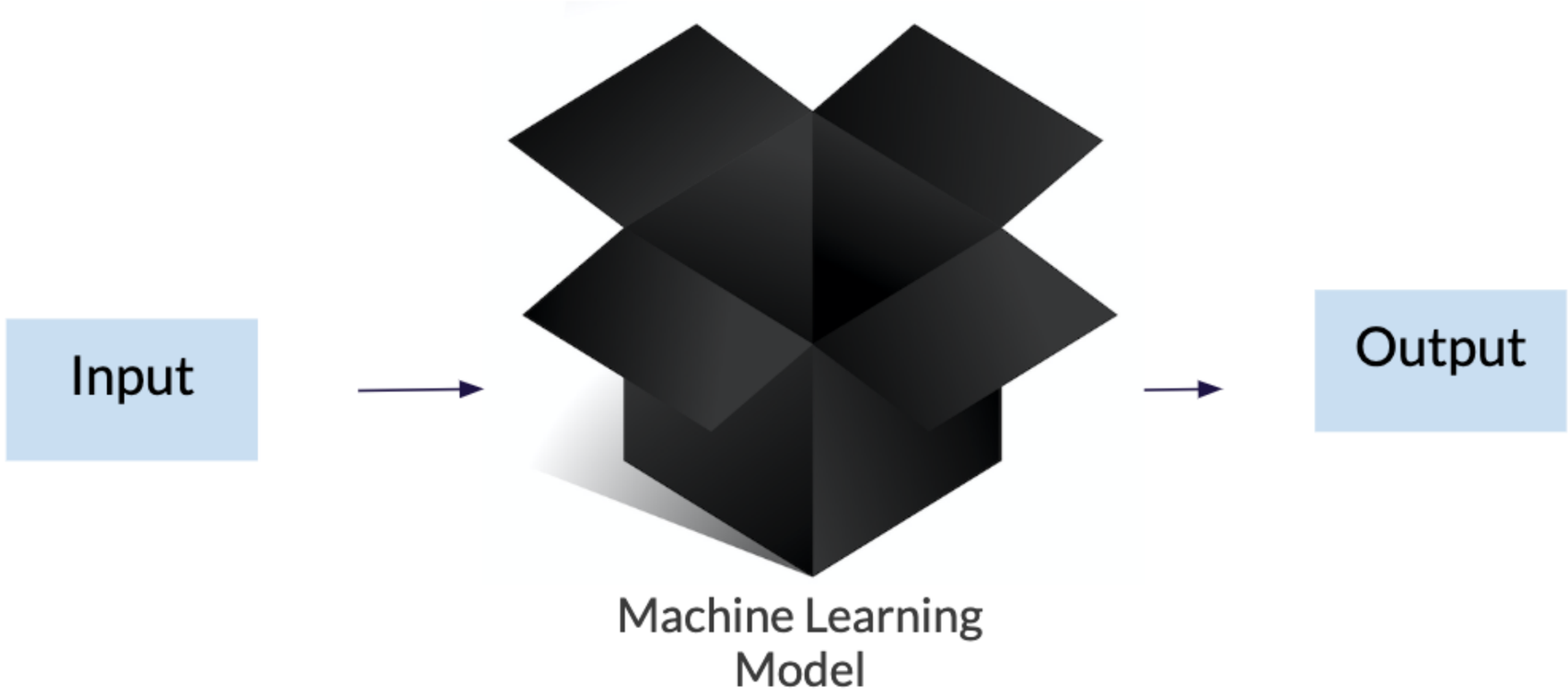
A machine learning model is only as good as the data you give it

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

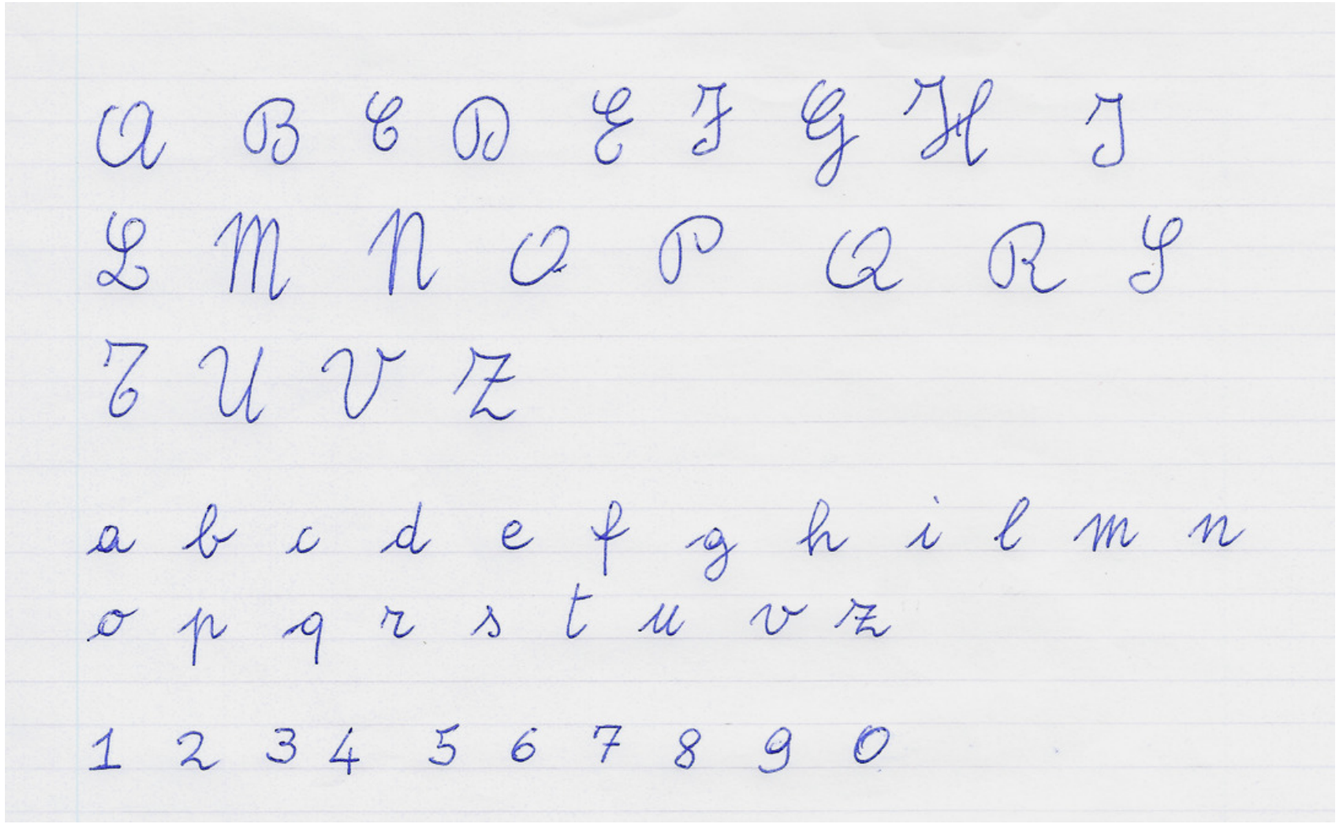
Explainable AI

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

Example: Explainable AI

1. **Prediction:** Will the patient get diabetes?
2. **Inference:** Why will this happen

Example: Inexplicable AI



Prediction only: Which letter is this likely to be?

Let's practice!

UNDERSTANDING MACHINE LEARNING

Congratulations!

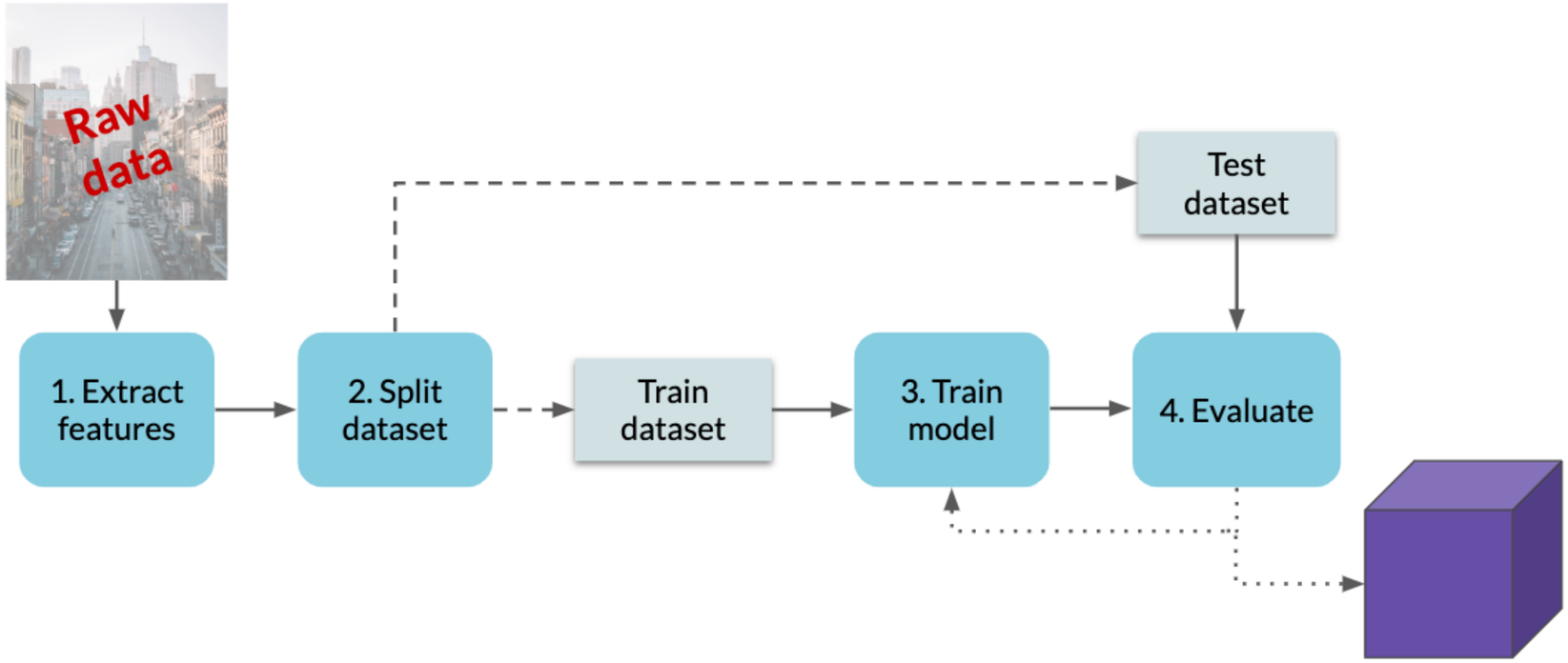
UNDERSTANDING MACHINE LEARNING

Lis Sulmont
Curriculum Manager, DataCamp

datacamp

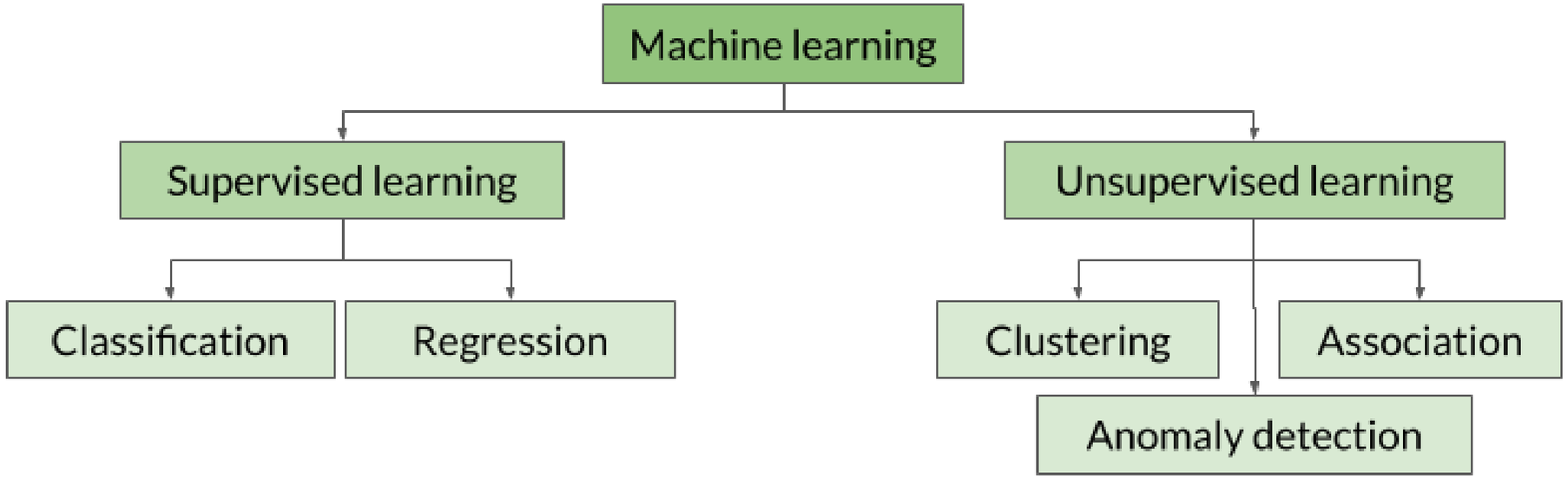
Chapter 1

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



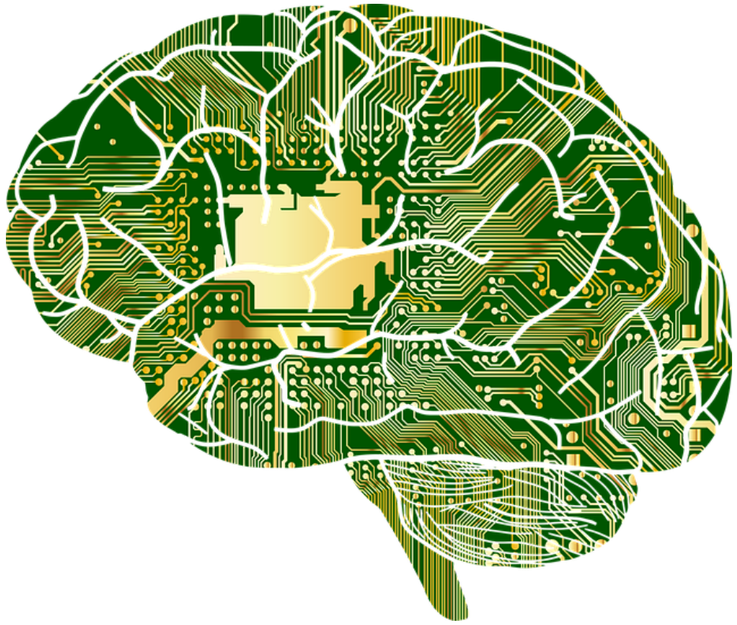
Chapter 2

- 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