

DECOMPOSING NEURAL NETWORKS

An applicant's guide to artificial learning

15.11.2022

→ JENNIFER MATTHIESEN & TINO PAULSEN | WINTERSEMESTER 2022



LEUPHANA
UNIVERSITÄT LÜNEBURG

RECAP

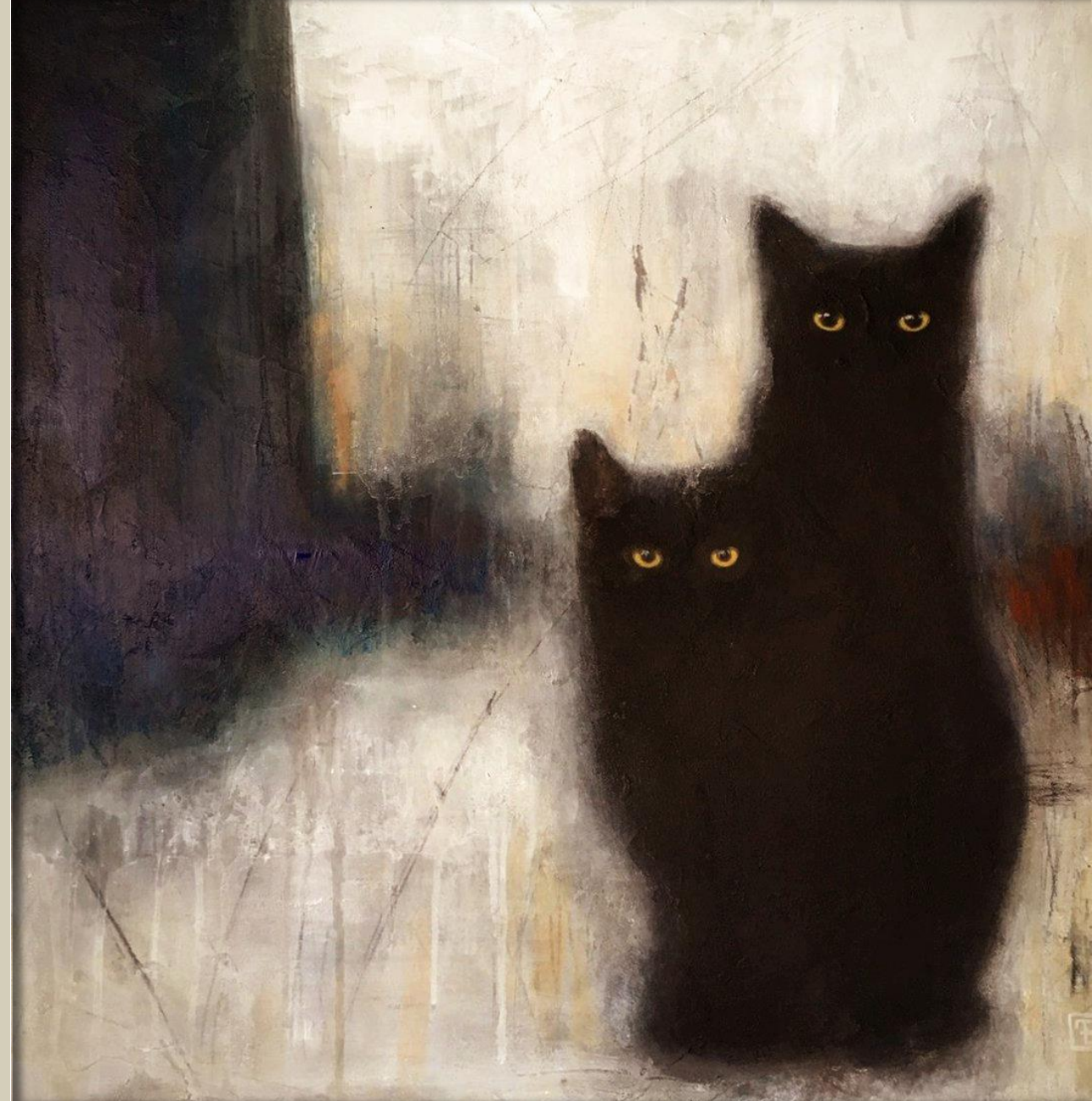
INTRO NN

- A biological neuron
- The artificial imitation: Perceptron
- Non-linear- decisions
- Cognitive psychology:
 - Geons
 - Bigram detectors
- Feature detection in images



ABOUT DATA, CATS & DOGS

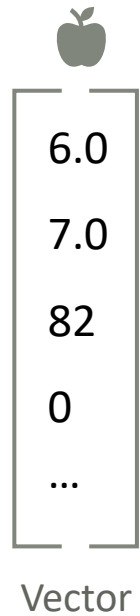
- Vectors, Matrix, Tensor
- Classification & bias
- BIAS from data
- How NN's understand images
 - Convolutions & pooling
- Train your own network in a supervised manner



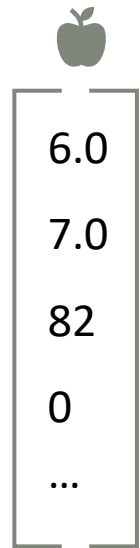


**BUT FIRST:
LET'S TALK
ABOUT DATA**

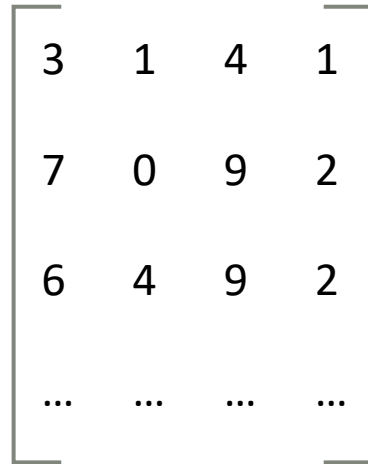
HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



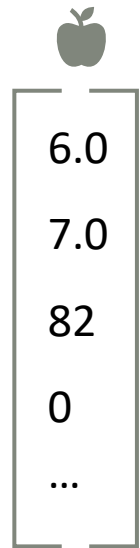
Vector



Matrix



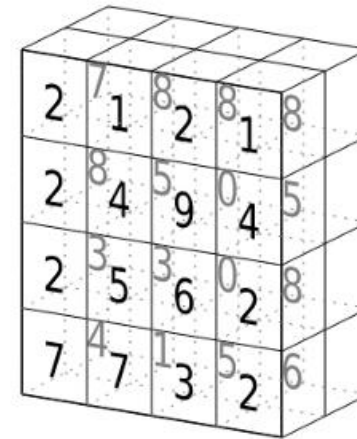
HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



Vector



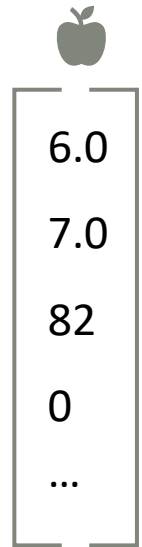
Matrix



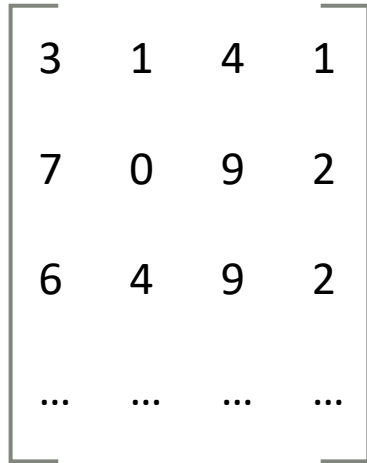
Tensor



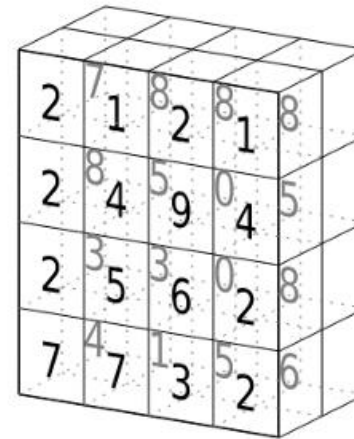
HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



Tensor
of dimension [5]



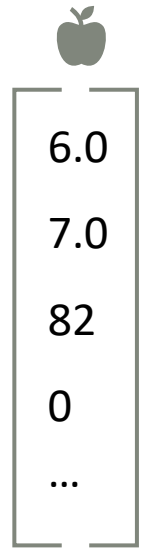
Matrix



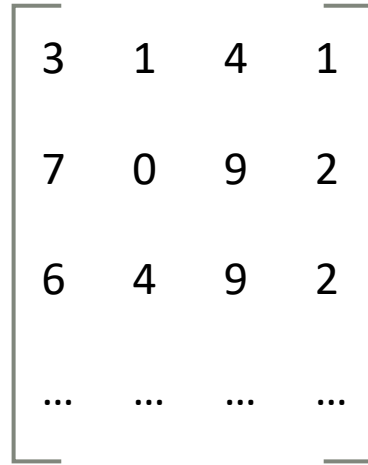
Tensor



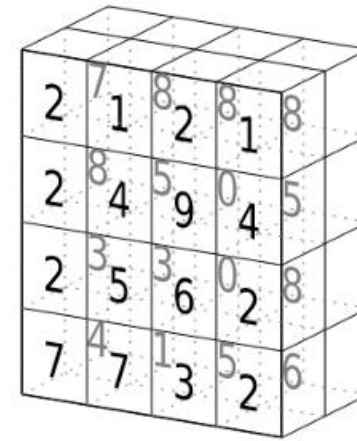
HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



Tensor
of dimension [5]



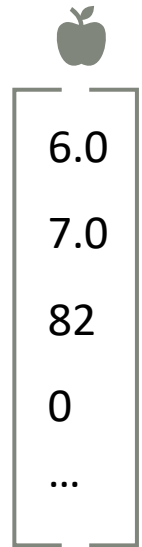
Tensor
of dimension [4,4]



Tensor



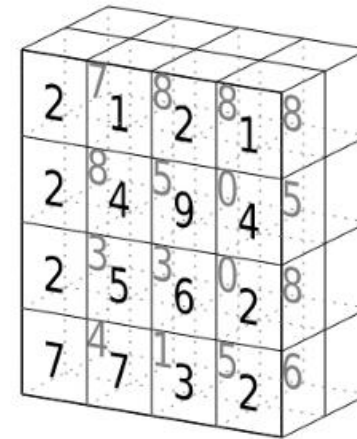
HOW DOES A NEURAL NETWORK UNDERSTAND DATA?



Tensor
of dimension [5]



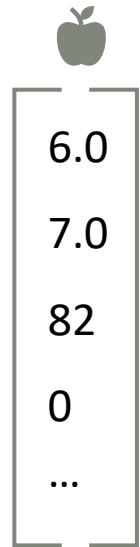
Tensor
of dimension [4,4]



Tensor
of dimension [4,4,2]



HOW DOES A NEURAL NETWORK UNDERSTAND DATA?

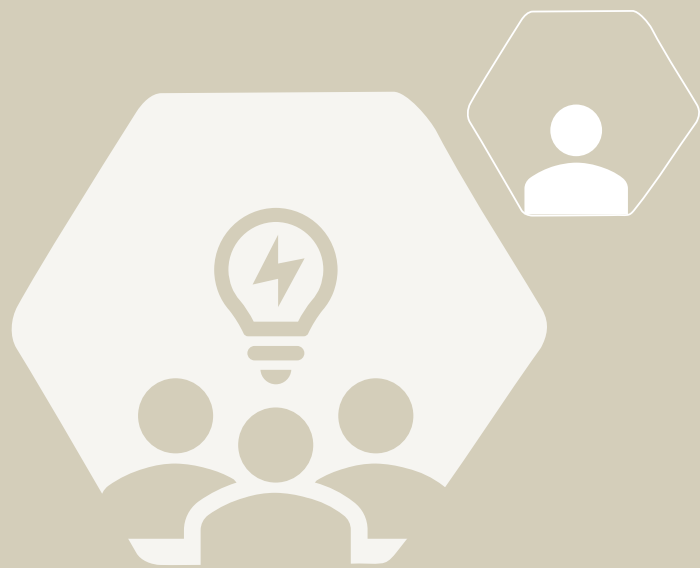


Tensor
of dimension [5]

- Feature engineering:
 - Designing features which can then be used as input
 - can have a huge influence on the performance
- Network “decides” how important each of feature is.



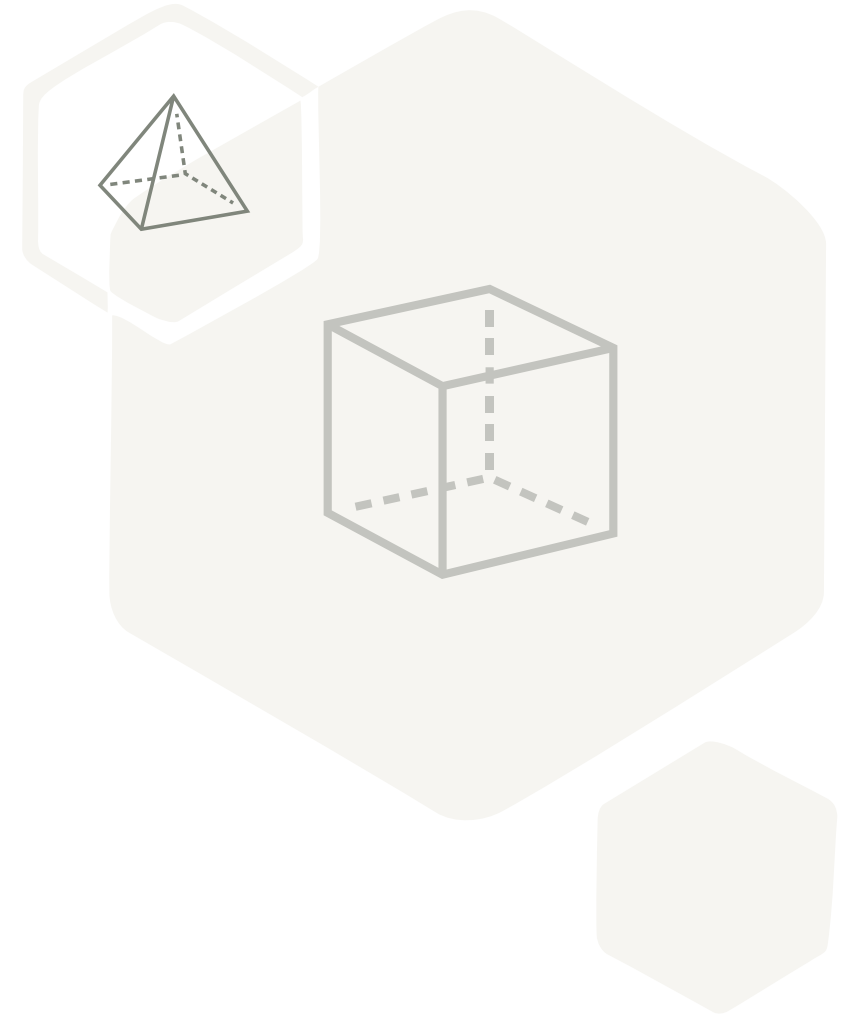
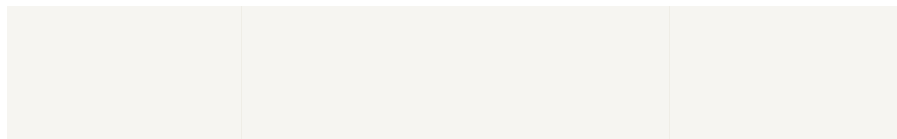
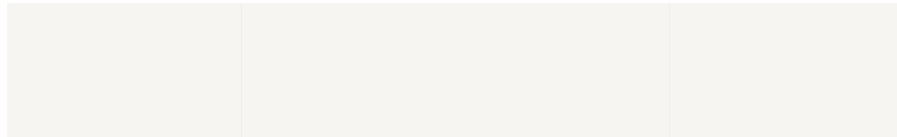
TIME FOR AN EXPERIMENT



EXPERIMENT

KURI & PUSI

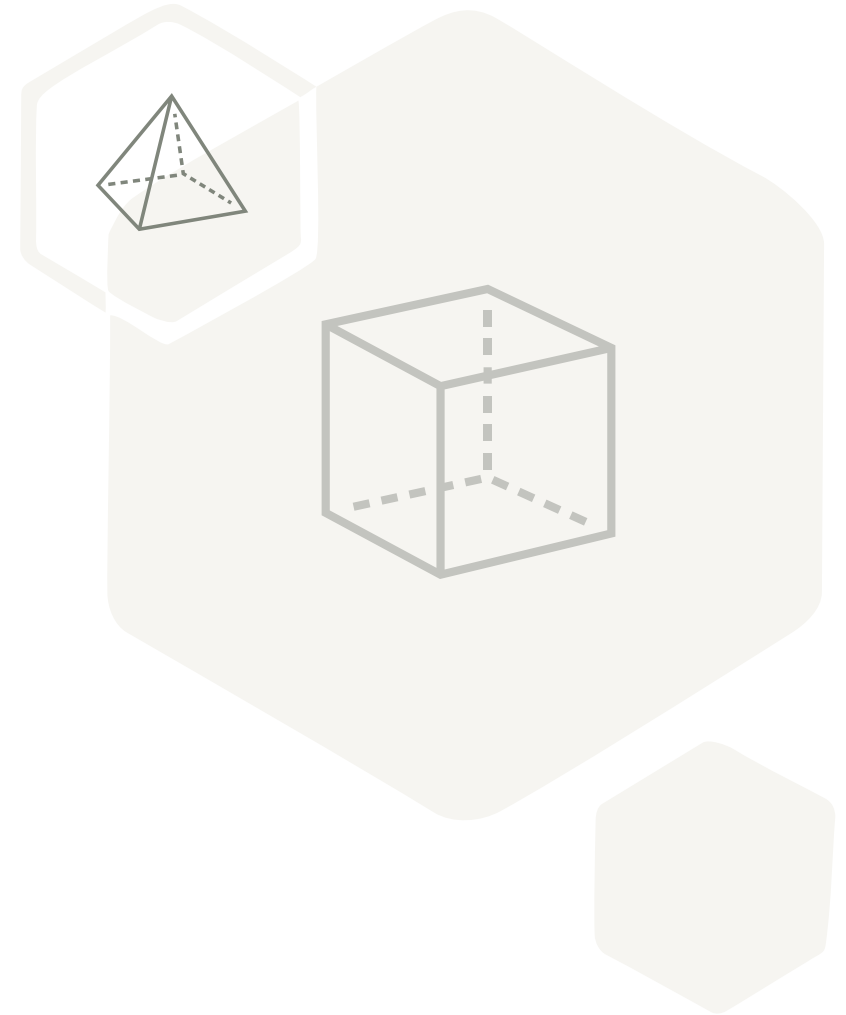
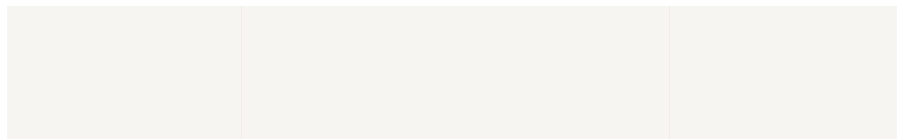
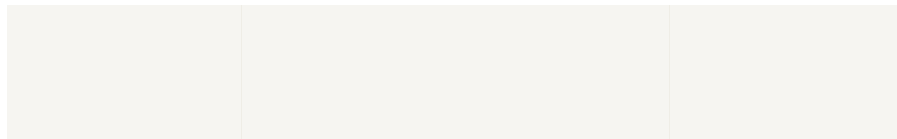
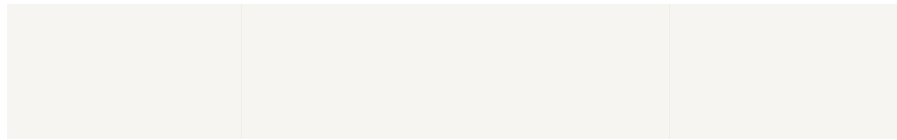
Ex01	small, brown	
Ex02	small, white	



EXPERIMENT

KURI & PUSI

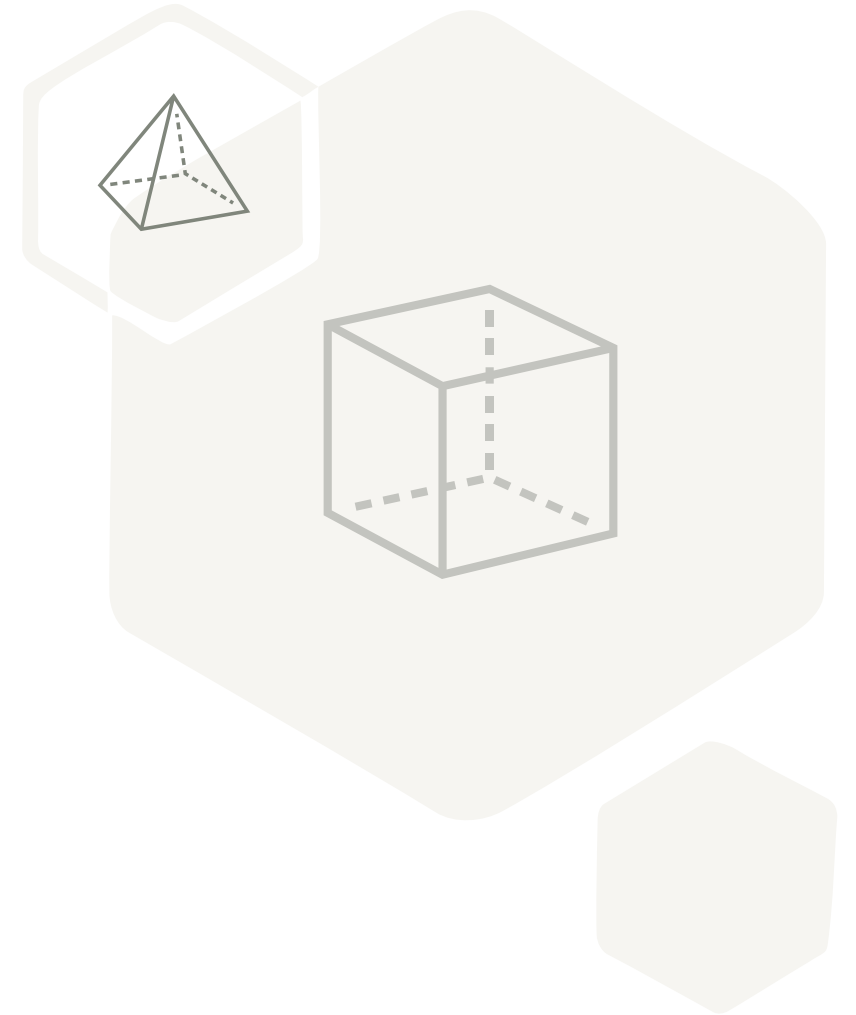
Ex01	small, brown	KURI
Ex02	small, white	PUSI



EXPERIMENT

KURI & PUSI

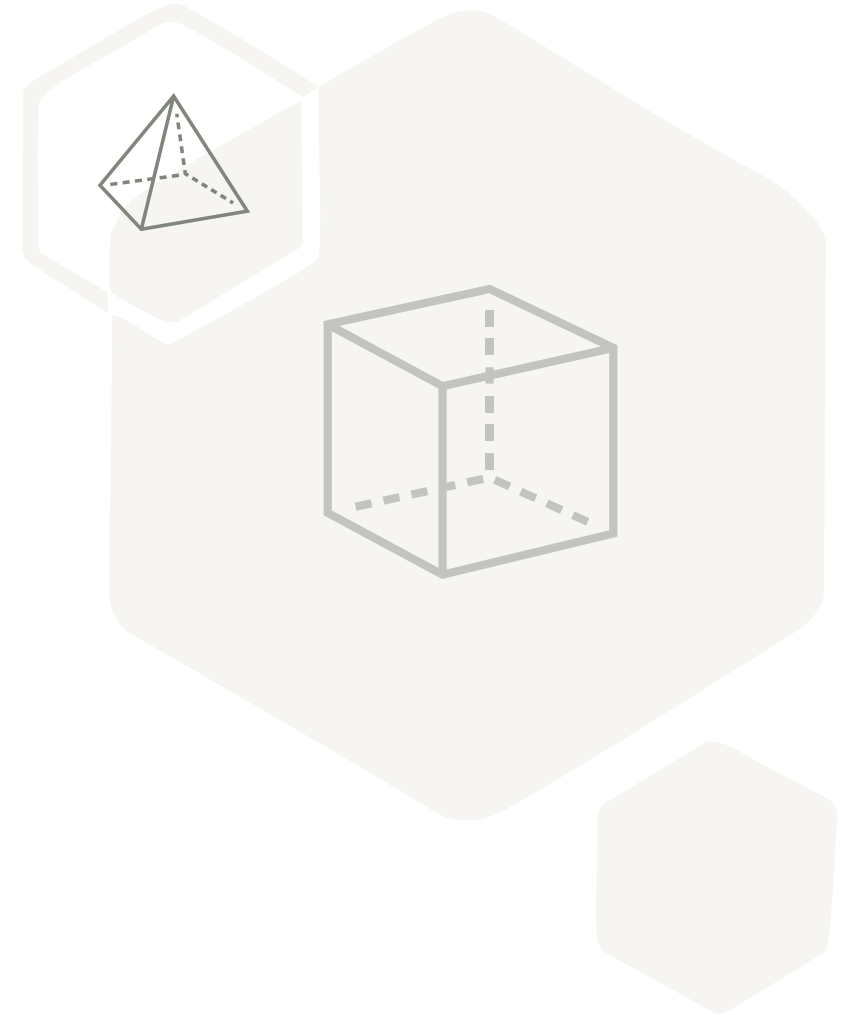
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	
Ex04	small, brown	



EXPERIMENT

KURI & PUSI

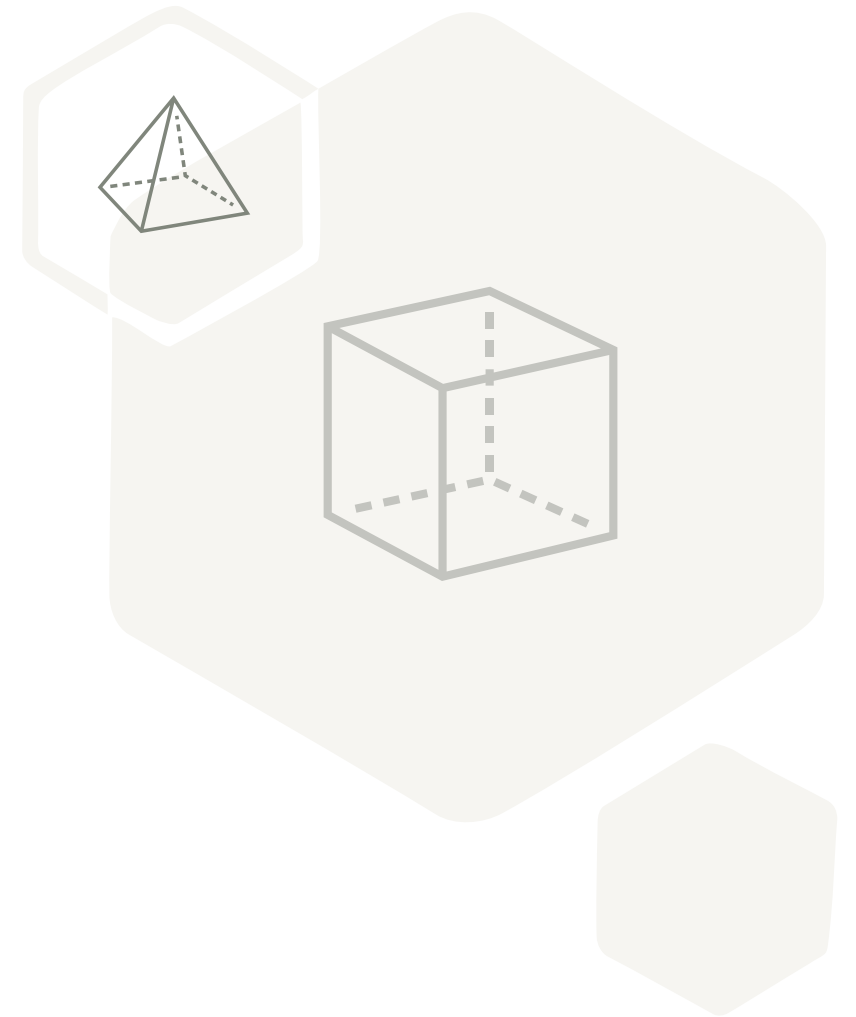
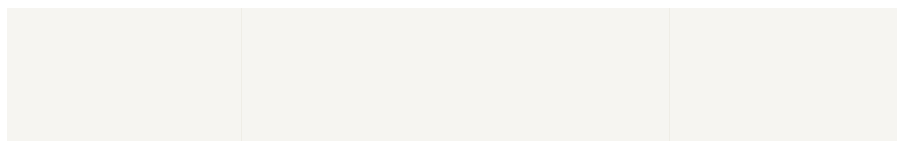
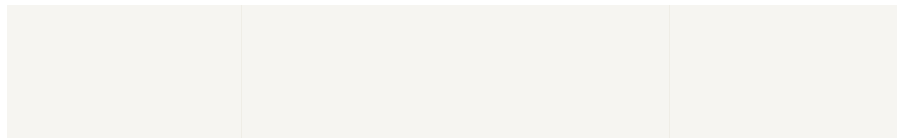
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI



EXPERIMENT

KURI & PUSI

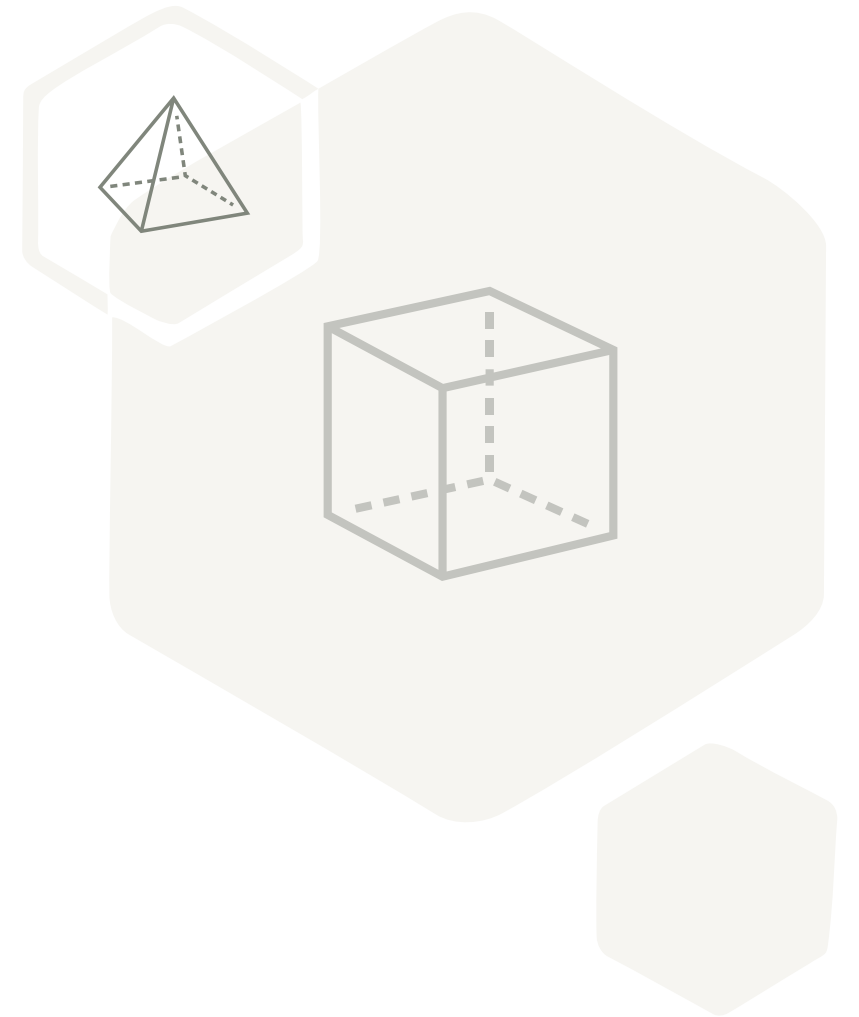
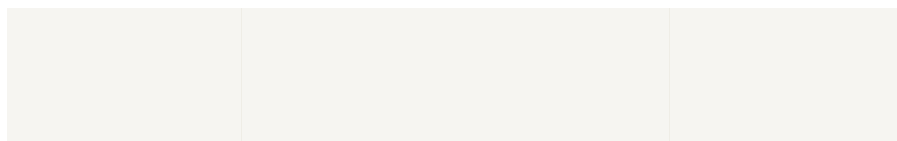
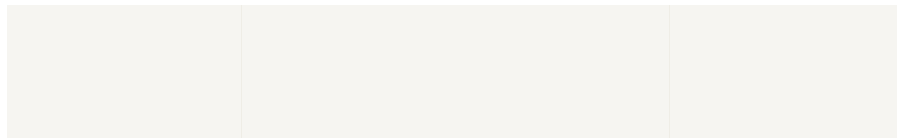
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	
Ex06	big, white	



EXPERIMENT

KURI & PUSI

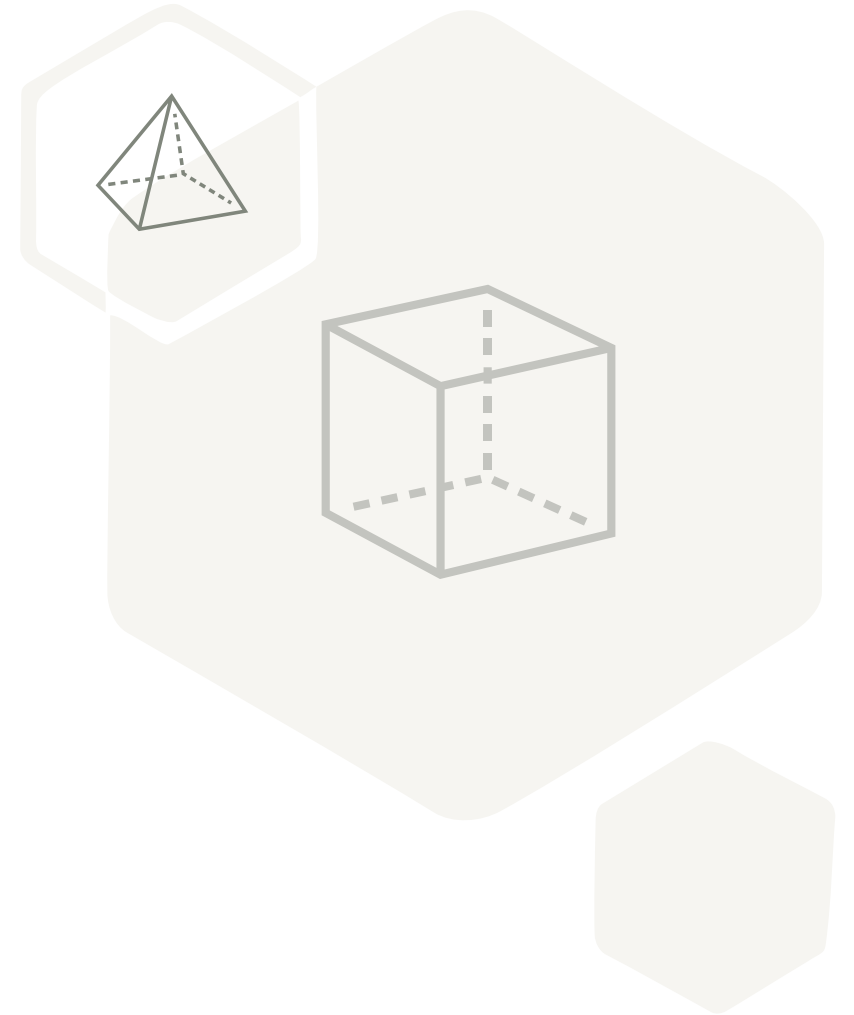
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI



EXPERIMENT

KURI & PUSI

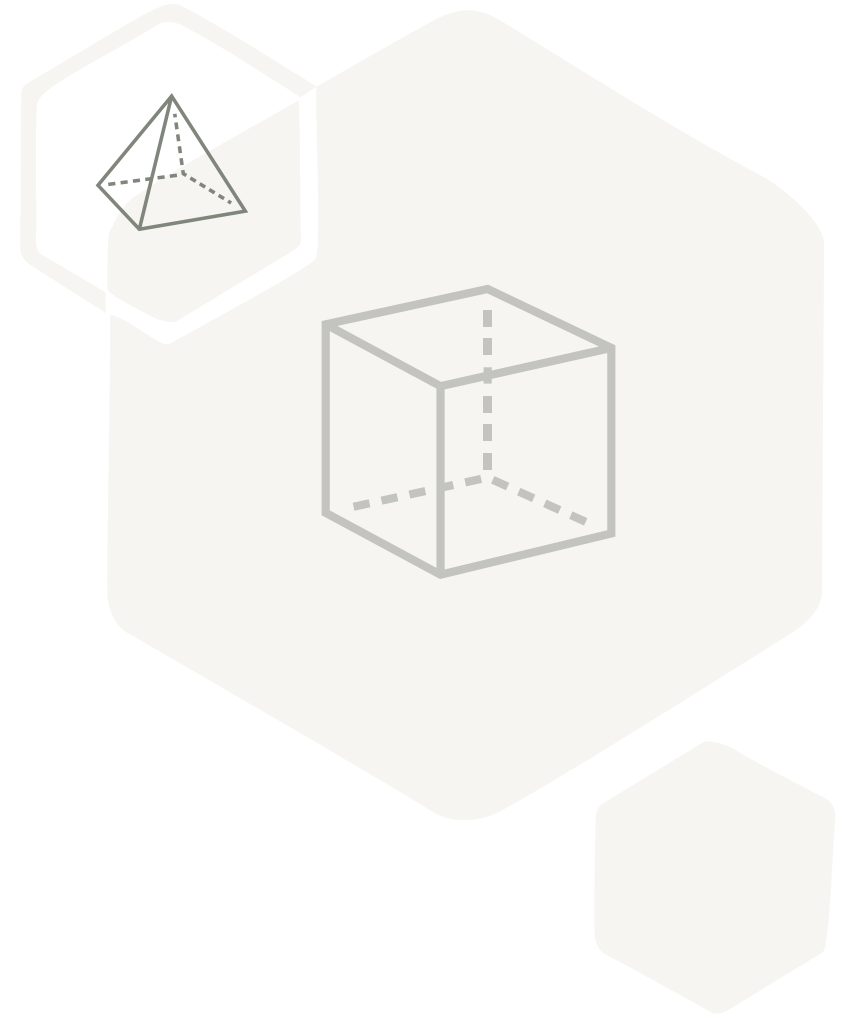
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	
Ex08	small, red	



EXPERIMENT

KURI & PUSI

Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI

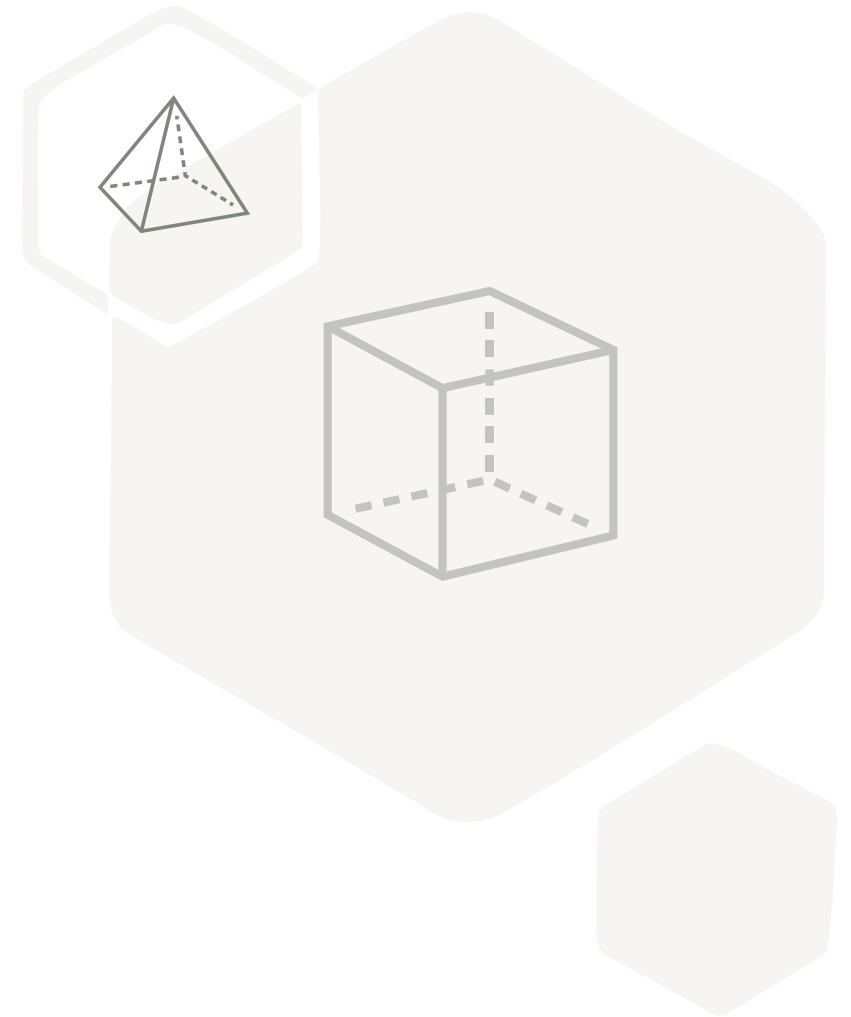


EXPERIMENT

KURI & PUSI

Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	
Ex10	small, black	

--	--	--

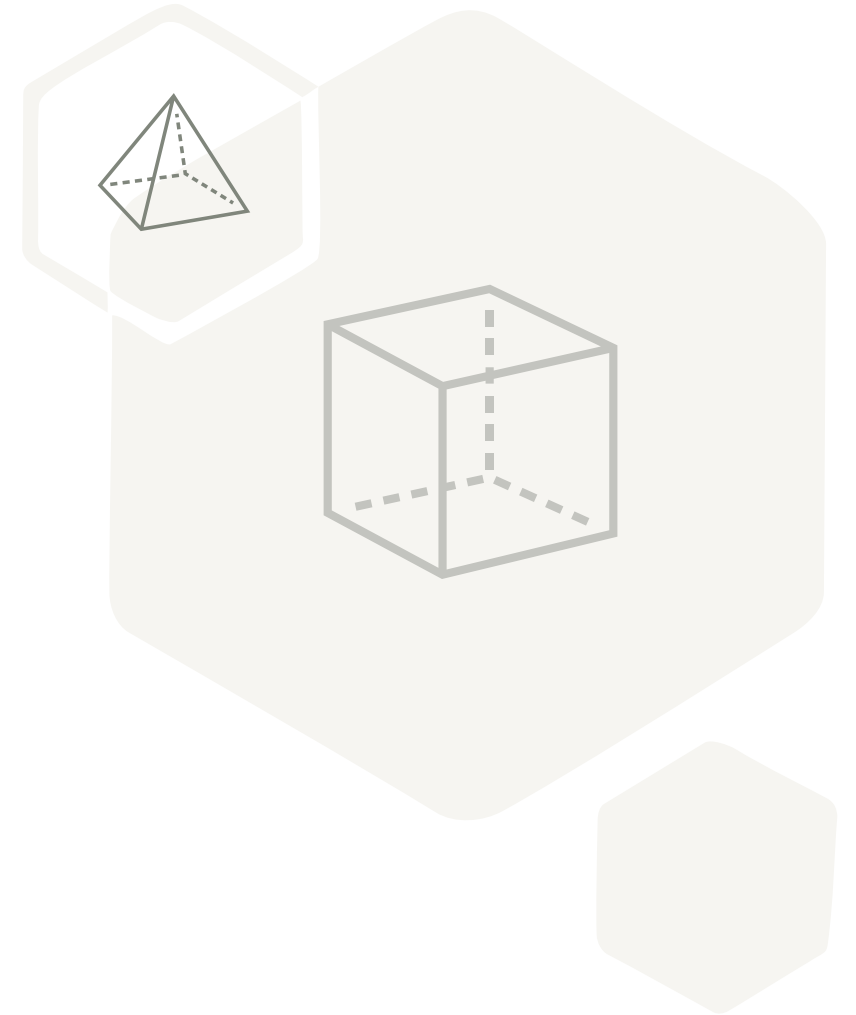


EXPERIMENT

KURI & PUSI

Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	KURI
Ex10	small, black	PUSI

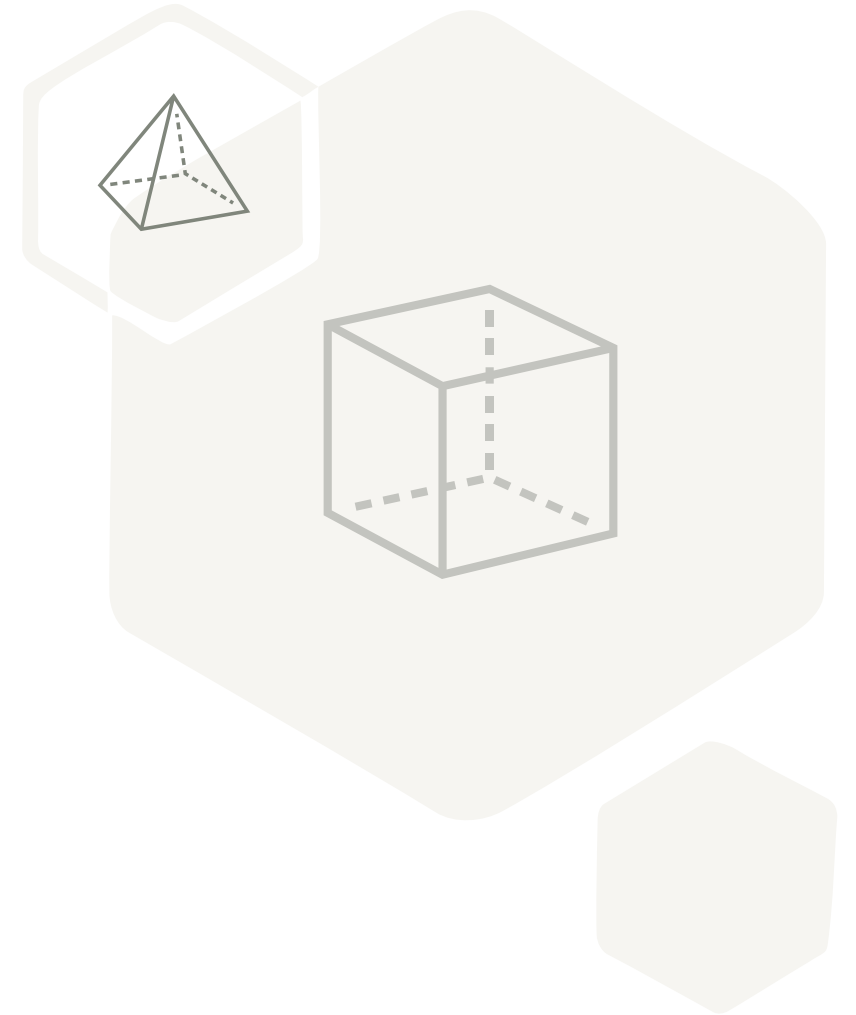
--	--	--



EXPERIMENT

KURI & PUSI

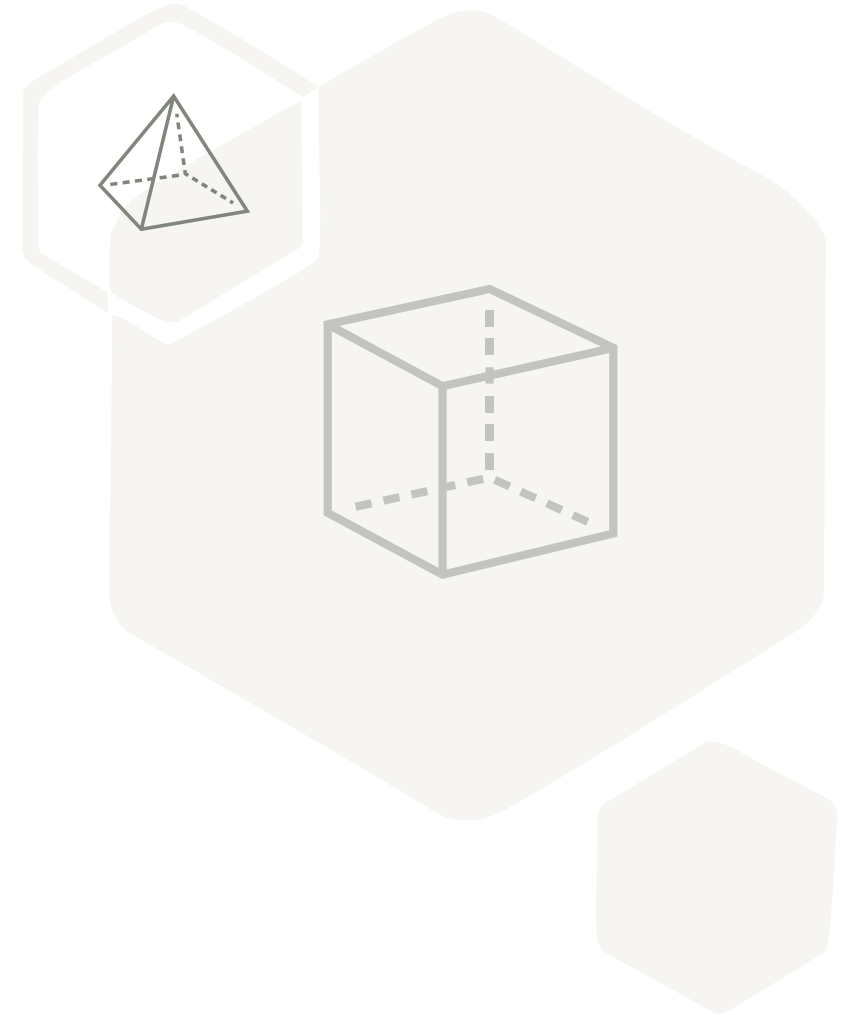
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	KURI
Ex10	small, black	PUSI
Ex11	small, black	
Ex12	brown-white	



EXPERIMENT

KURI & PUSI

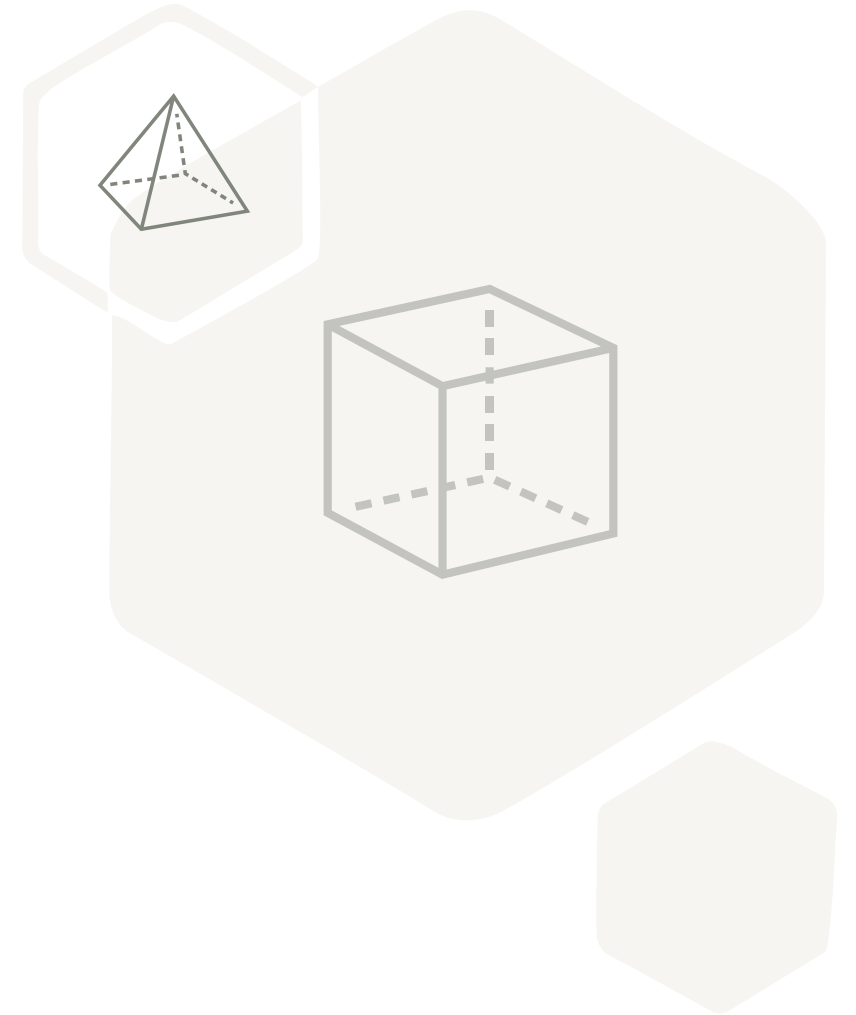
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	KURI
Ex10	small, black	PUSI
Ex11	small, black	PUSI
Ex12	brown-white	PUSI



EXPERIMENT

KURI & PUSI

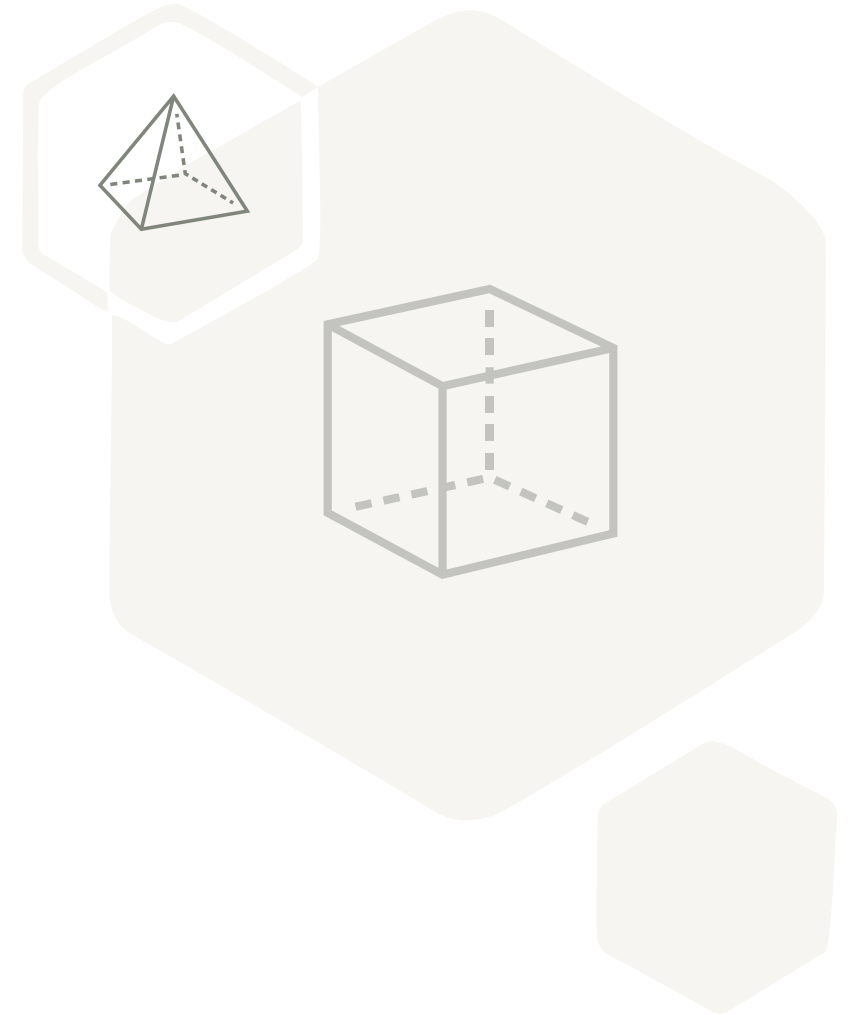
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	KURI
Ex10	small, black	PUSI
Ex11	small, black	PUSI
Ex12	brown-white	PUSI
Ex13	big black	
Ex14	small, red-white	



EXPERIMENT

KURI & PUSI

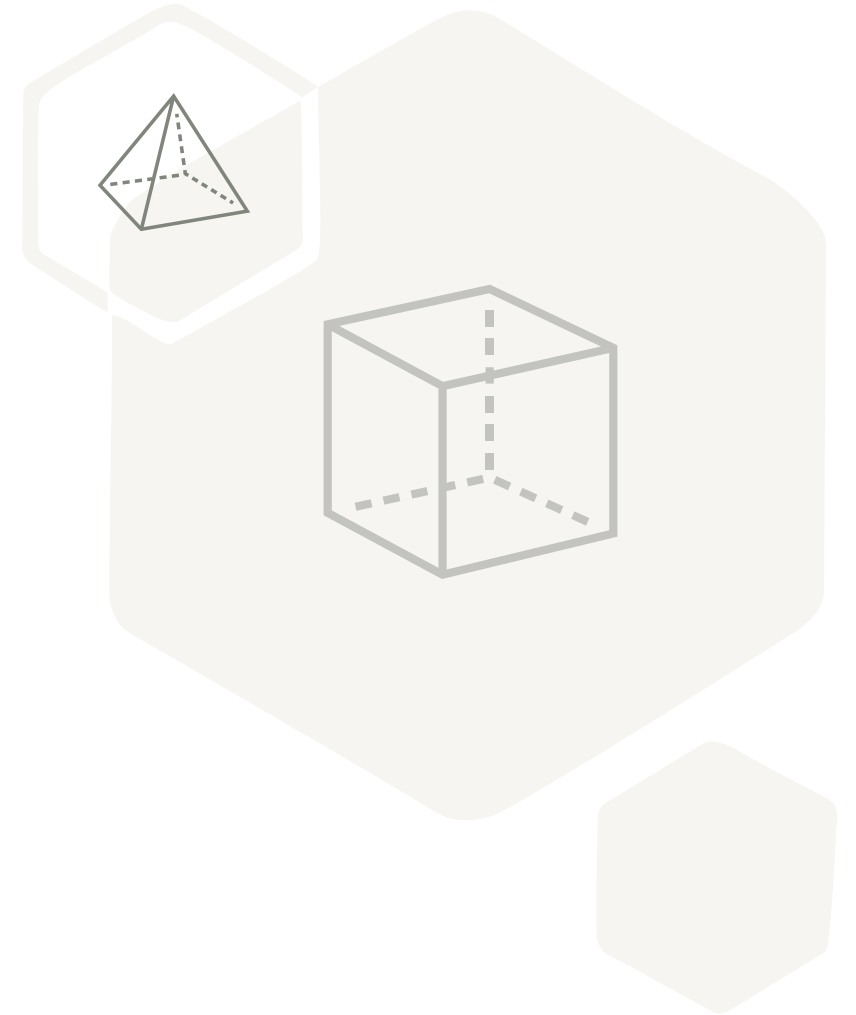
Ex01	small, brown	KURI
Ex02	small, white	PUSI
Ex03	small, brown	KURI
Ex04	small, brown	PUSI
Ex05	small, white	PUSI
Ex06	big, white	KURI
Ex07	big, brown	KURI
Ex08	small, red	PUSI
Ex09	small, brown	KURI
Ex10	small, black	PUSI
Ex11	small, black	PUSI
Ex12	brown-white	PUSI
Ex13	big black	KURI
Ex14	small, red-white	PUSI



EXPERIMENT

KURI & PUSI

CLASS 1	CLASS 2
KURI small, brown	PUSI small, white
KURI small, brown	PUSI small, brown
KURI big, white	PUSI small, white
KURI big, brown	PUSI small, red
KURI small, brown	PUSI small, black
KURI big black	PUSI small, brown-white
	PUSI small, black
	PUSI small, red-white



EXPERIMENT

DOGS & CATS

CLASS 1	CLASS 2
DOG small, brown	CAT small, white
DOG small, brown	CAT small, brown
DOG big, white	CAT small, white
DOG big, brown	CAT small, red
DOG small, brown	CAT small, black
DOG big, black	CAT small, brown-white
	CAT small, black
	CAT small, red-white



EXPERIMENT

DOGS & CATS

— So, what did the network learn?
And What did it not learn?

- Dogs can be big and small
- Cats are always small
- Both are black, brown or white
- Cats can be multi-colored
- Dogs are one-colored (!)
- Dogs are never small and white (!)

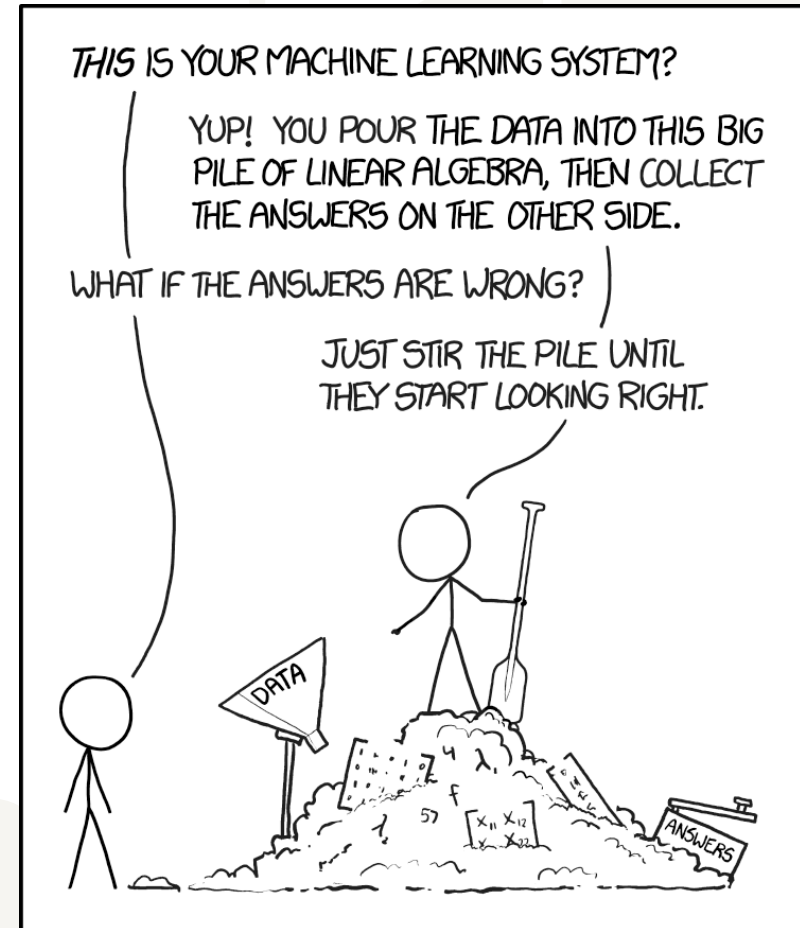


LET'S TALK ABOUT BIAS



BIAS IN DATA

- Organically created dataset are biased
- Many machine learning models can often contain unintentional bias
- It is important to analyse your training data and sometimes the source of the data to look for biases



Cartoon by xkcd.com, creative commons license (<https://xkcd.com/license.html>)

BIAS IN DATA REPORTING BIAS

- Also known as selective reporting
- Takes place when only a selection of results or outcomes are captured in a data set
- Types of reporting bias:
 - Citation bias
 - Language bias
 - Duplicate publication bias
 - Location bias
 - Publication bias: positive findings are more likely to be published than studies with negative findings
 - Outcome reporting bias



BIAS IN DATA

AUTOMATION BIAS

- Tendency of humans to favor results or suggestions generated by automated systems.
- Ignore contradictory information made by non-automated systems, even if it is correct



BIAS IN DATA

SELECTION BIAS

- Selection bias takes place when data is chosen in a way that is not reflective of real-world data distribution.
- No proper randomization during data collection
- Types of selection bias:
 - Sampling bias: not enough randomization
 - Convergence bias: e.g., only surveying customers who purchased your product
 - Participation bias: unrepresentative due to participations gaps



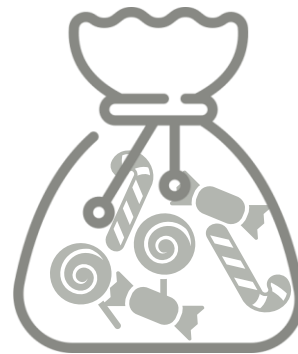
BIAS IN DATA

SELECTION BIAS

ON REPRESENTATIVE SAMPLES AND GUMMI BEARS

Let's say, we have a back of sweets with an infinite variety of sweets. We want to select some of them for our friend as a present, which we haven't seen in a while. Since not all the sweets will fit in the smaller bag for our friend, we must create a selection.

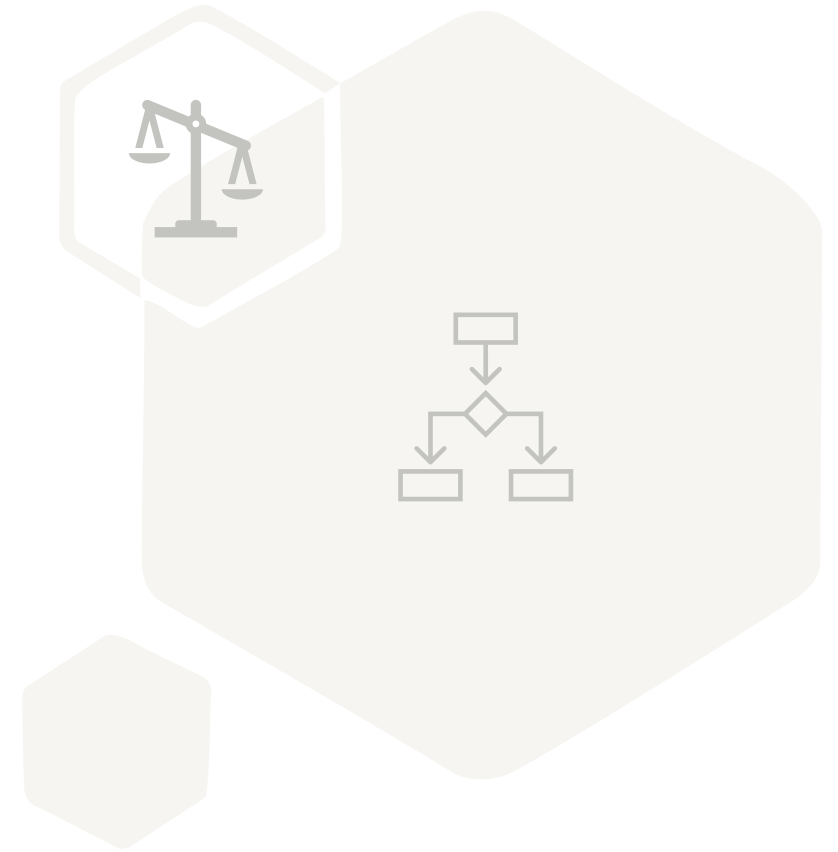
BUT HOW?



BIAS IN DATA

OVERGENERALIZATION BIAS

- Overgeneralization occurs when you assume what you see in your dataset is what you would see if you looked in any other dataset.
- regardless of the size of the dataset.
 - E.g., there are no white dogs (!)



BIAS IN DATA

GROUP ATTRIBUTION BIAS

- People tend to stereotype a whole group just because of the actions of a few individuals within the group
- Types of group attribution bias:
 - **In-group bias:** preference to members of a group you personally belong or share common interests with
 - **Out-group bias:** you stereotype individual members of a group to which you personally do not belong.



BIAS IN DATA

GROUP ATTRIBUTION BIAS



BIAS IN DATA

IMPLICIT BIAS

— Occurs when assumptions are made based on one's own personal experiences

— E.g., A computer vision engineer from North America marks the color red as a danger. However, the same color red is a popular color in Chinese culture that symbolizes luck, joy, and happiness.

—Type of Implicit Bias:

—Confirmation bias or experimenter's bias is the tendency to search for information in a way that confirms or supports one's prior beliefs or experiences

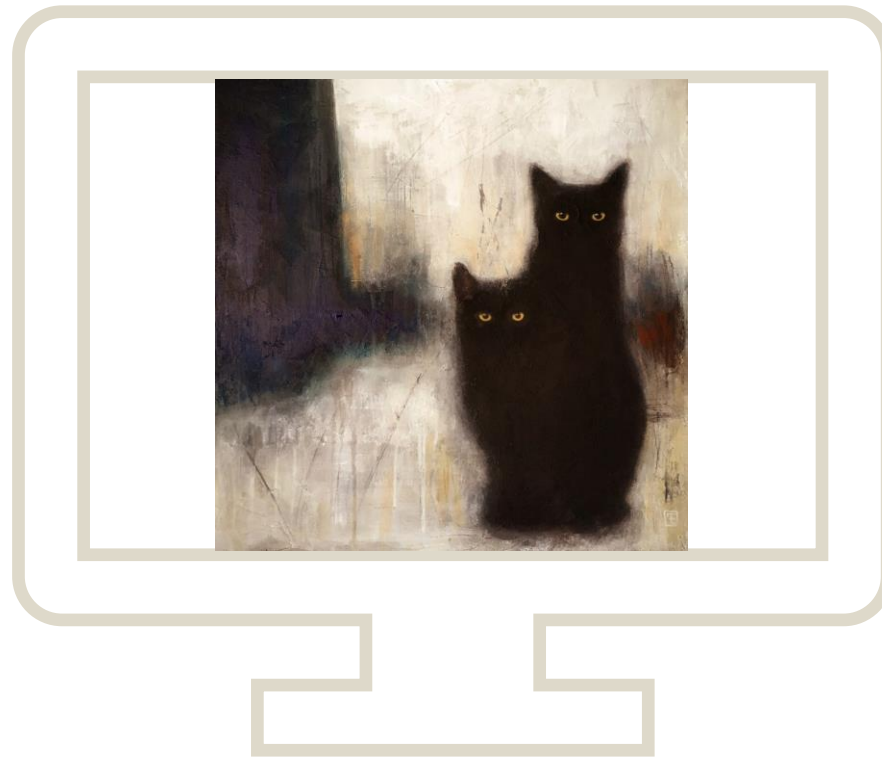


IMAGES AND PATTERN RECOGNITION



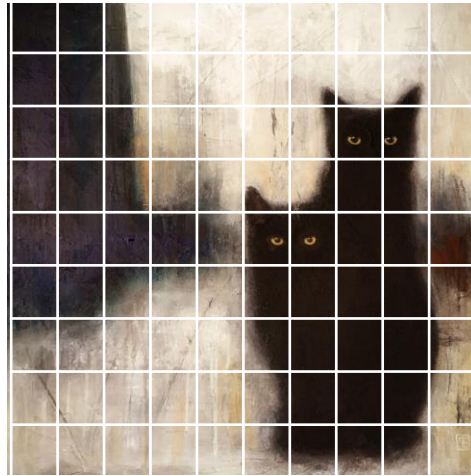
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



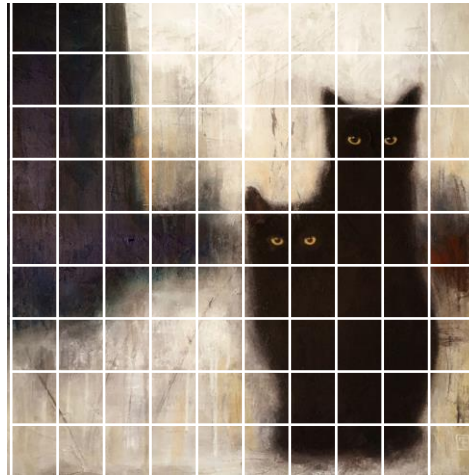
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?

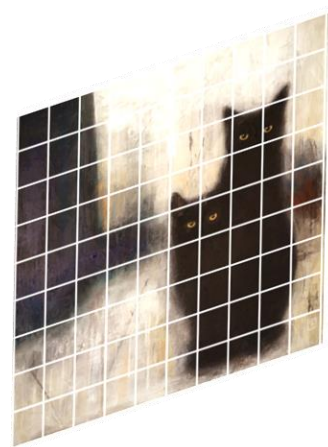


Actually, much smaller



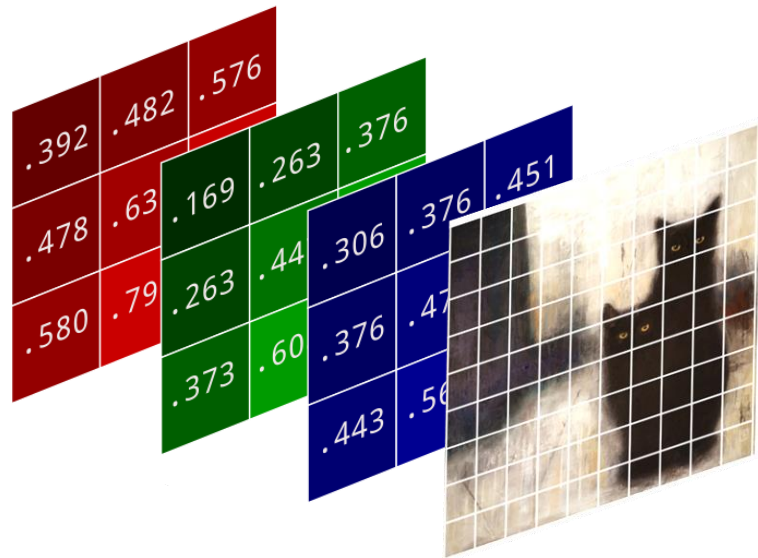
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



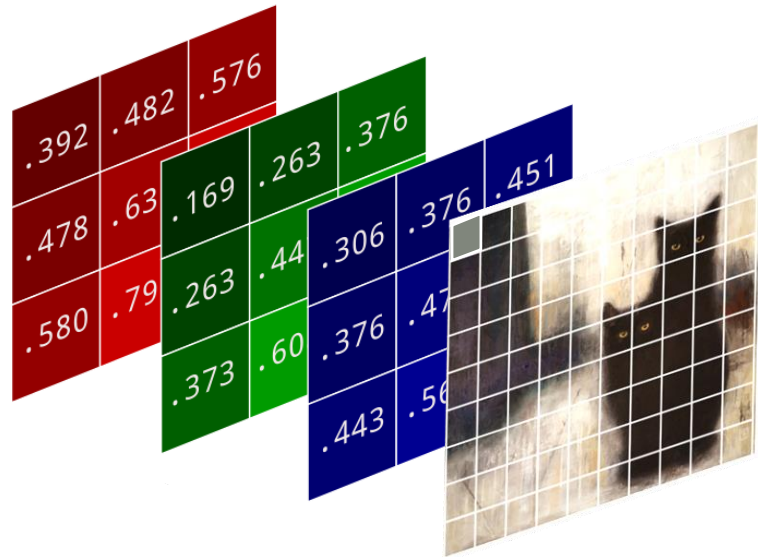
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



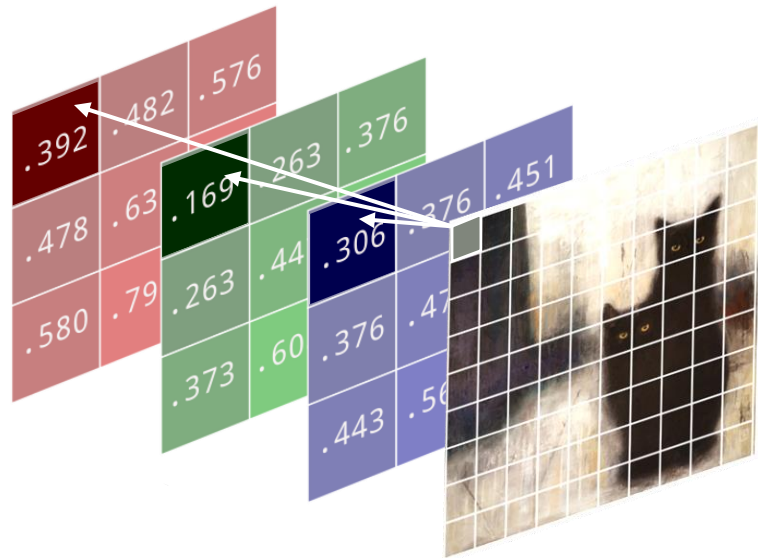
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



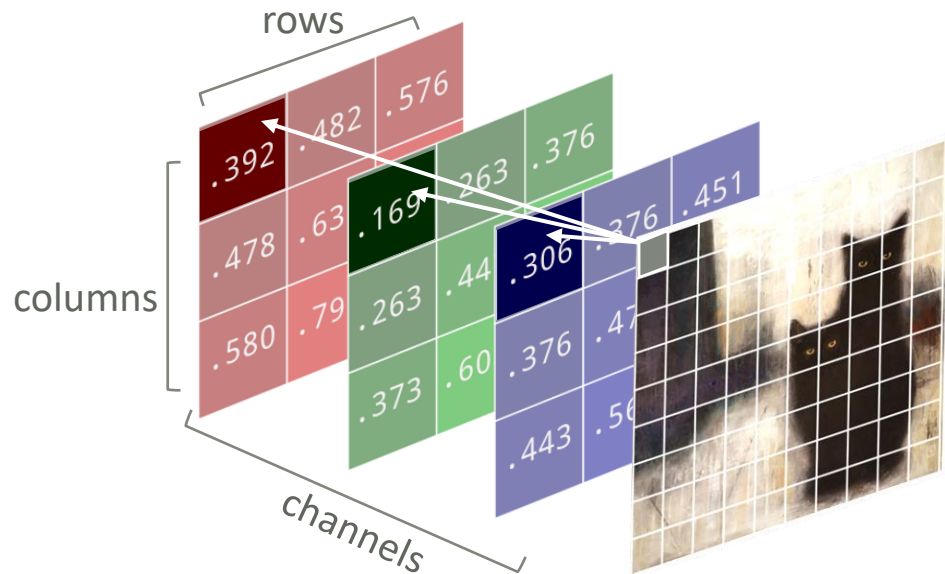
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

— How is an image like this represented on a screen?



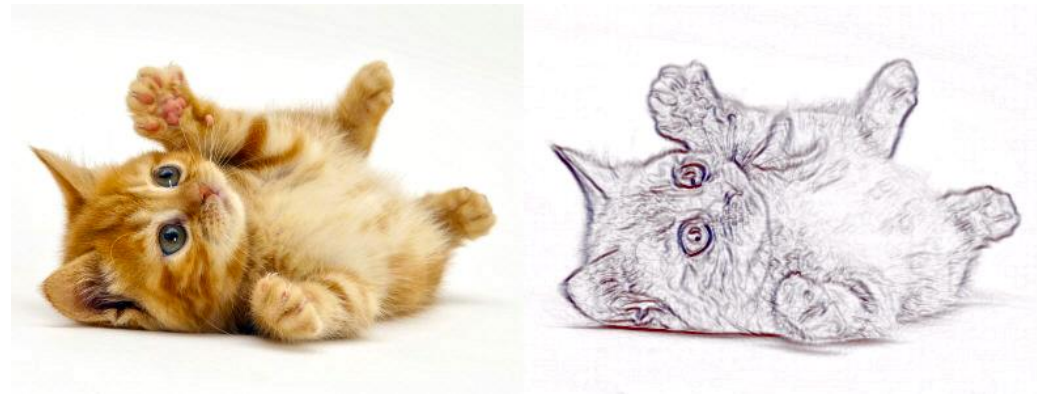
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

- Treating large-dimensional inputs as unstructured vectors leads to intractable models
- can lead to multiple GB memory footprint
- Also, this would require that the object on an images it always in the same place.



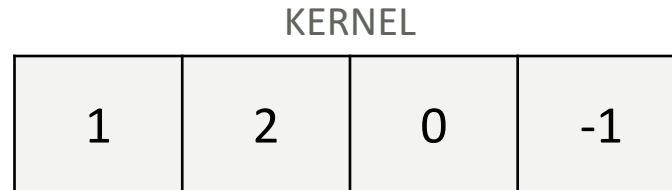
HOW DOES A NEURAL NETWORK UNDERSTAND IMAGES?

- A **representation** meaningful at a certain location can or should be used everywhere
- Meaning: we want to detect a cat despite where in the picture it is



CONVOLUTIONS IN NEURAL NETWORKS

- A convolution embodies the idea of having a meaningful representation
- We using a **kernel** to create the representation



TIME FOR MAGIC!



<https://setosa.io/ev/image-kernels/>



CONVOLUTIONS IN NEURAL NETWORKS

INPUT

1	4	-1	0	2	-2	1	3	3	1
---	---	----	---	---	----	---	---	---	---

KERNEL

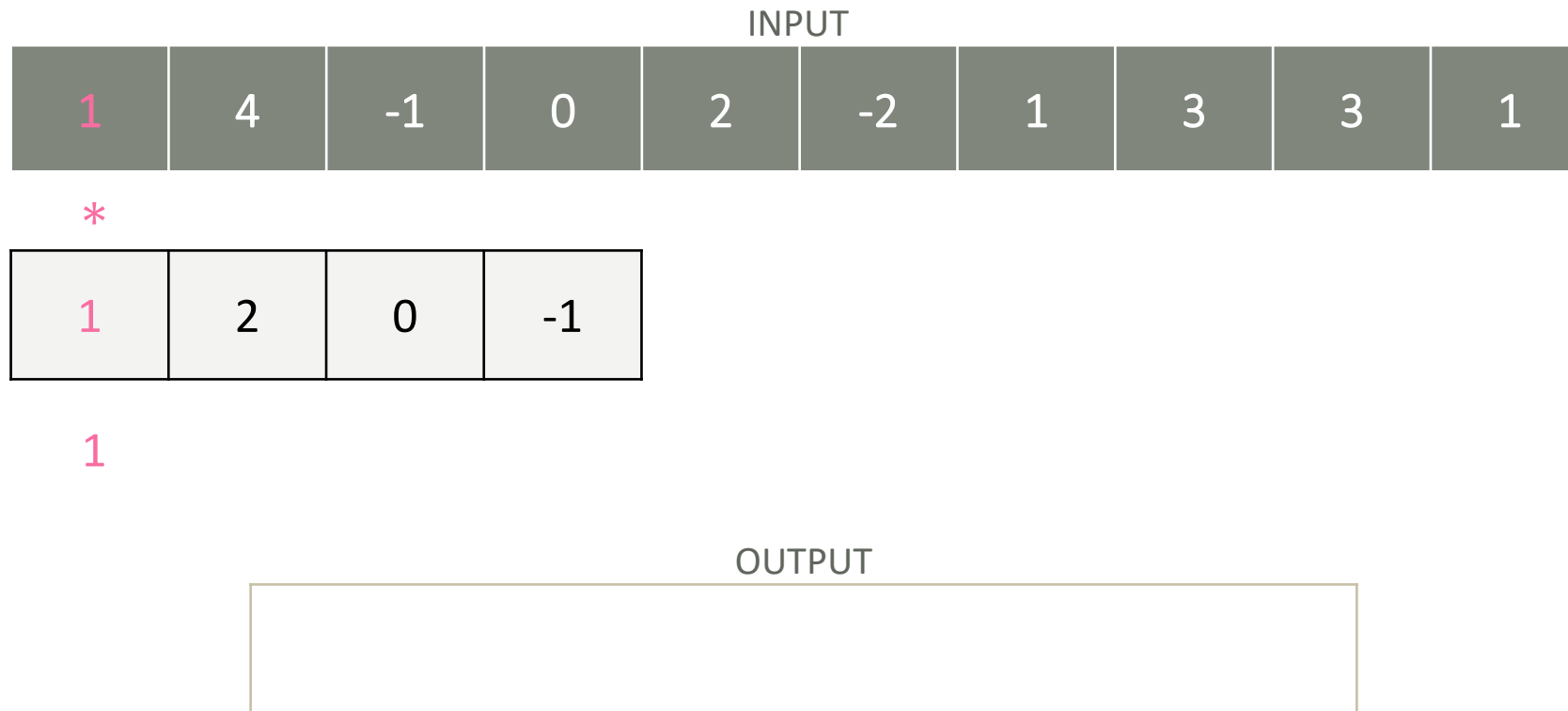
1	2	0	-1
---	---	---	----

OUTPUT

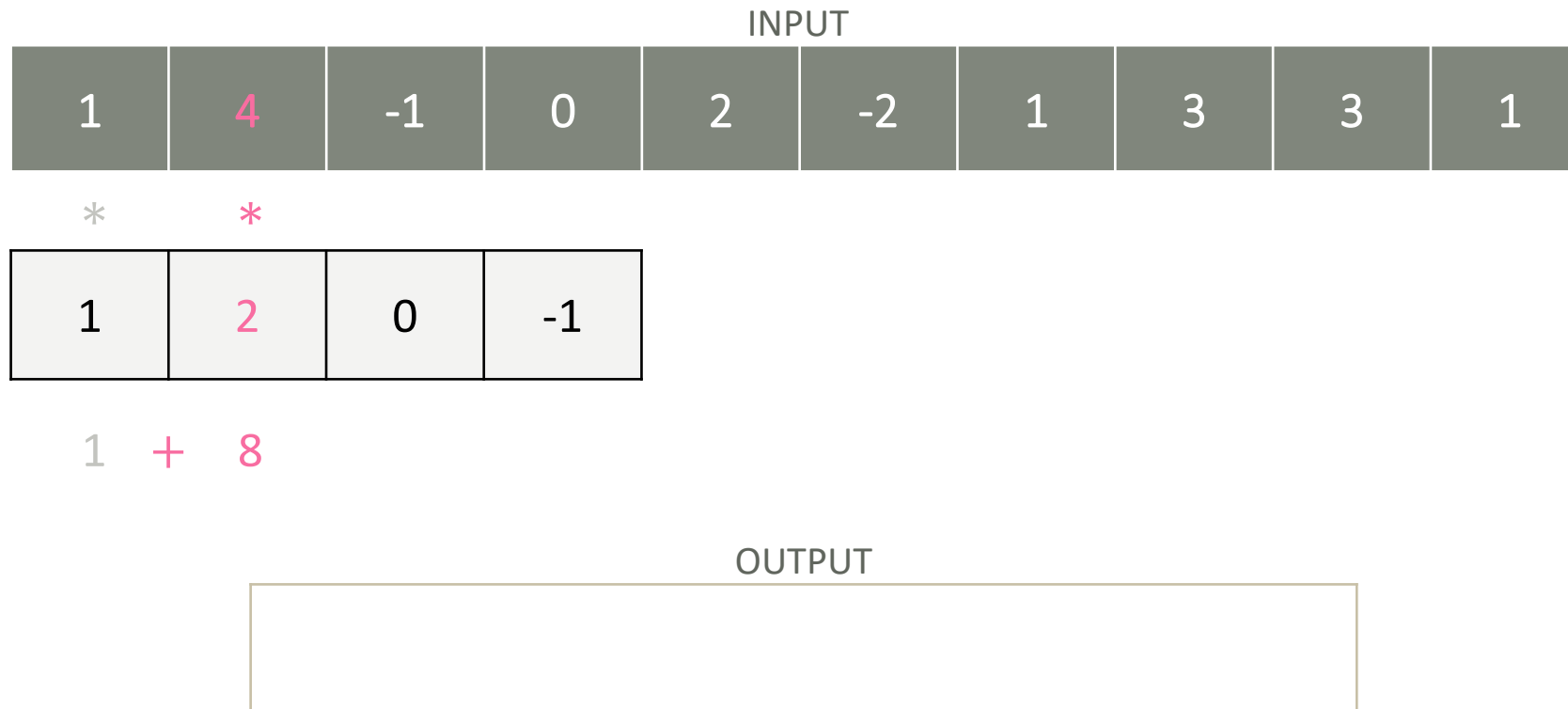
--



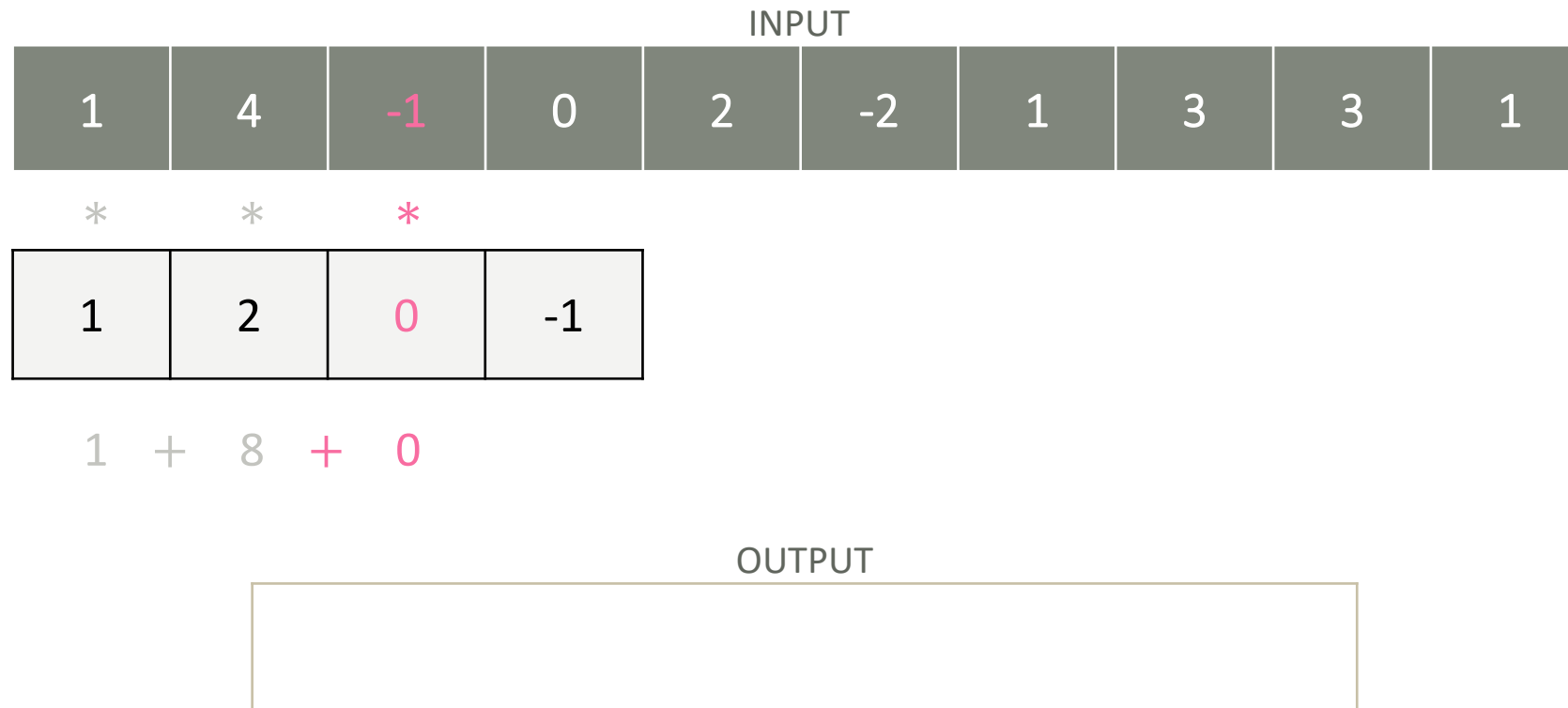
CONVOLUTIONS IN NEURAL NETWORKS



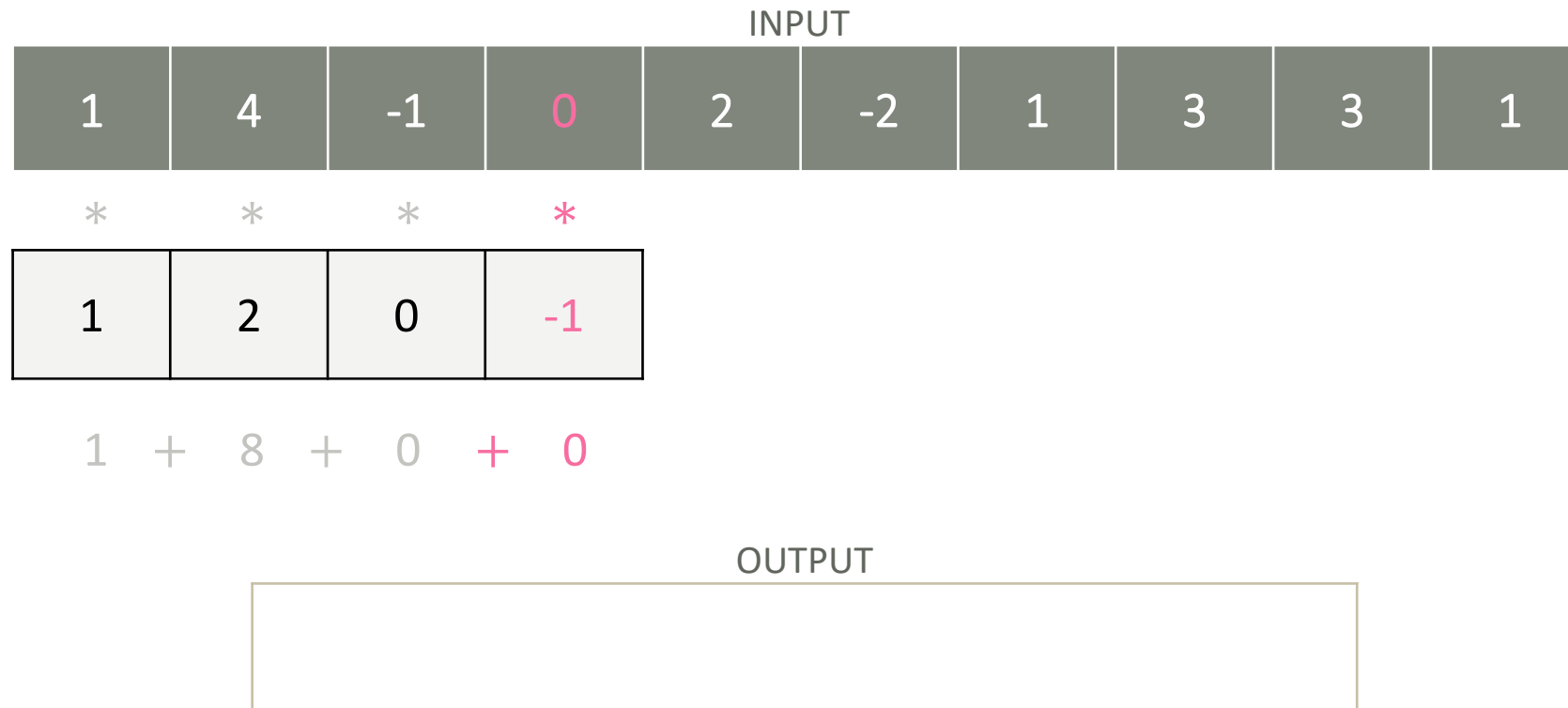
CONVOLUTIONS IN NEURAL NETWORKS



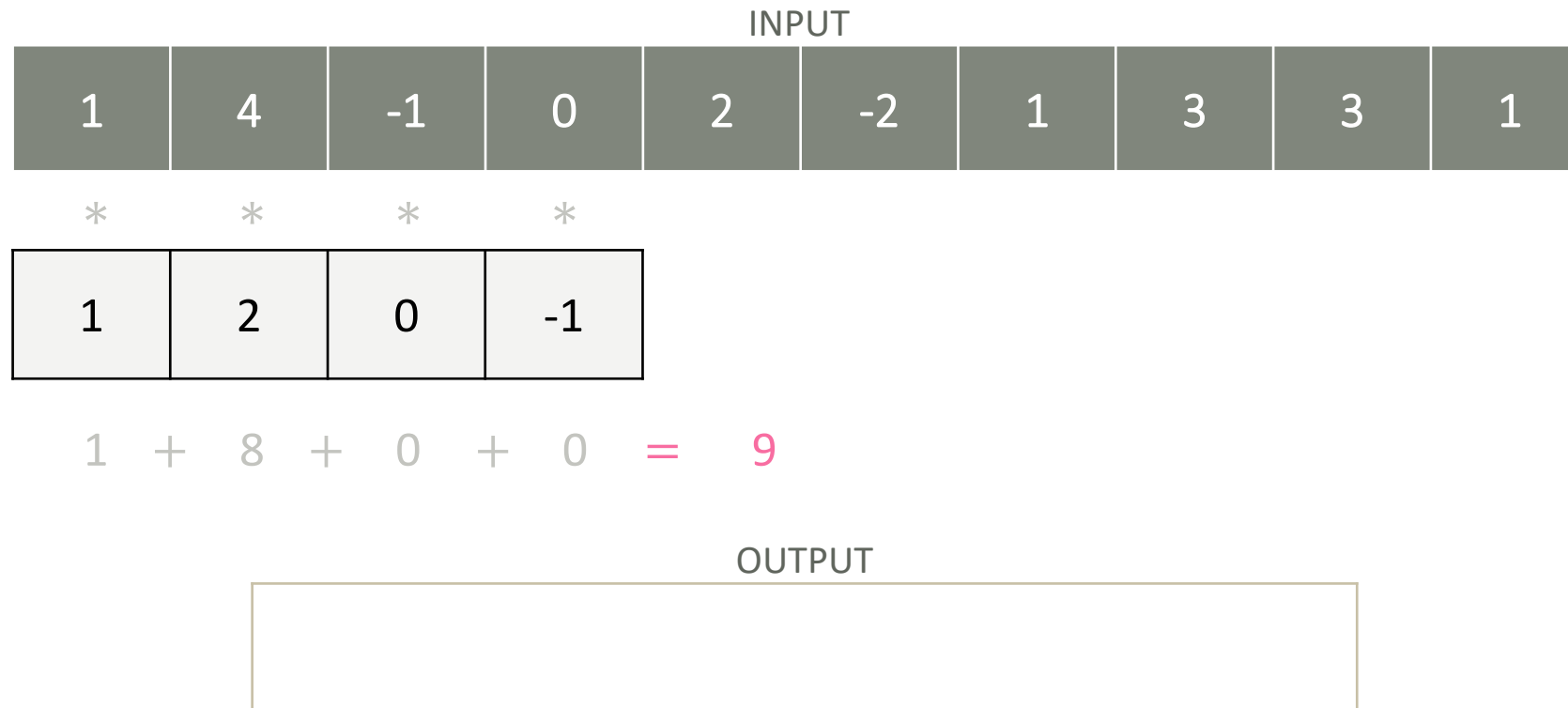
CONVOLUTIONS IN NEURAL NETWORKS



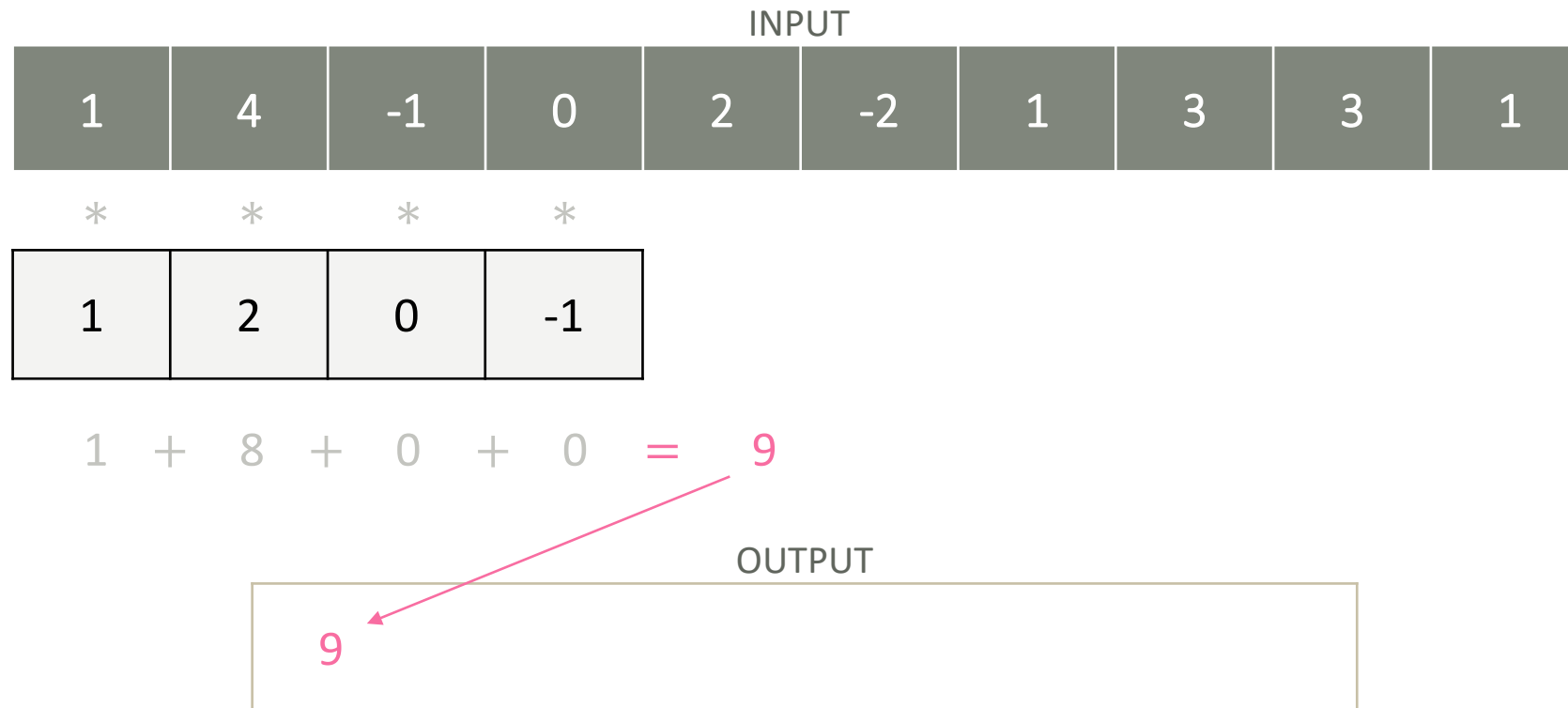
CONVOLUTIONS IN NEURAL NETWORKS



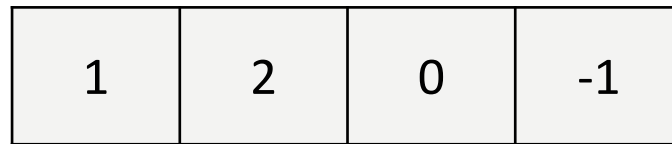
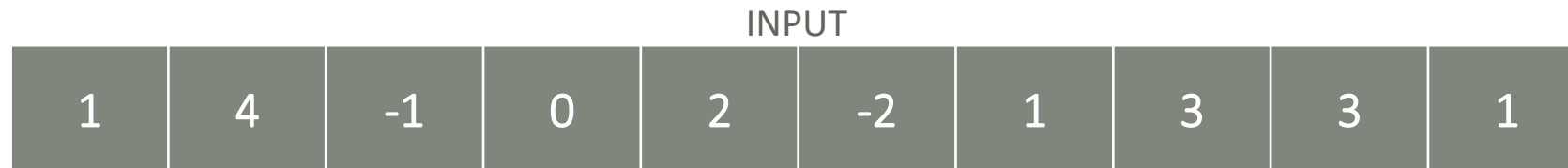
CONVOLUTIONS IN NEURAL NETWORKS



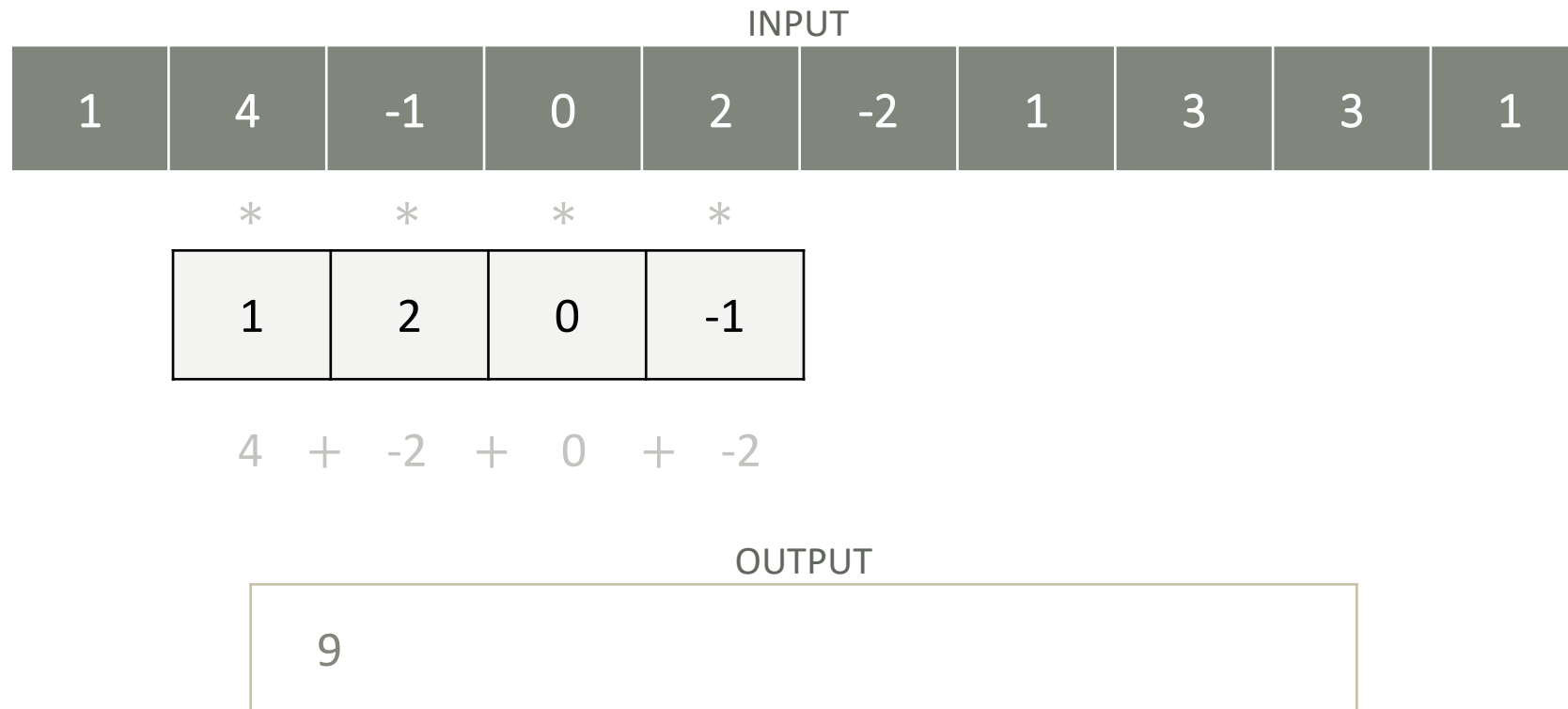
CONVOLUTIONS IN NEURAL NETWORKS



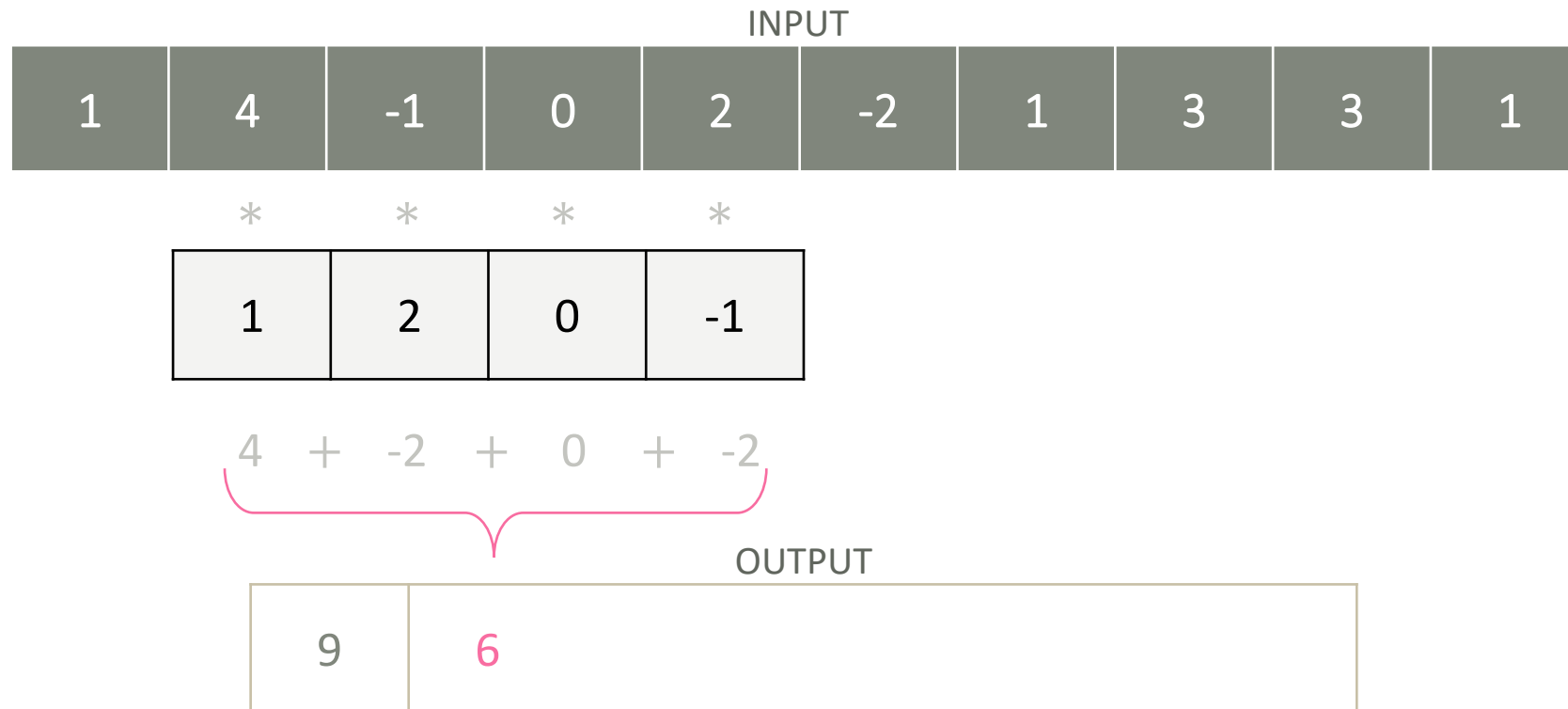
CONVOLUTIONS IN NEURAL NETWORKS



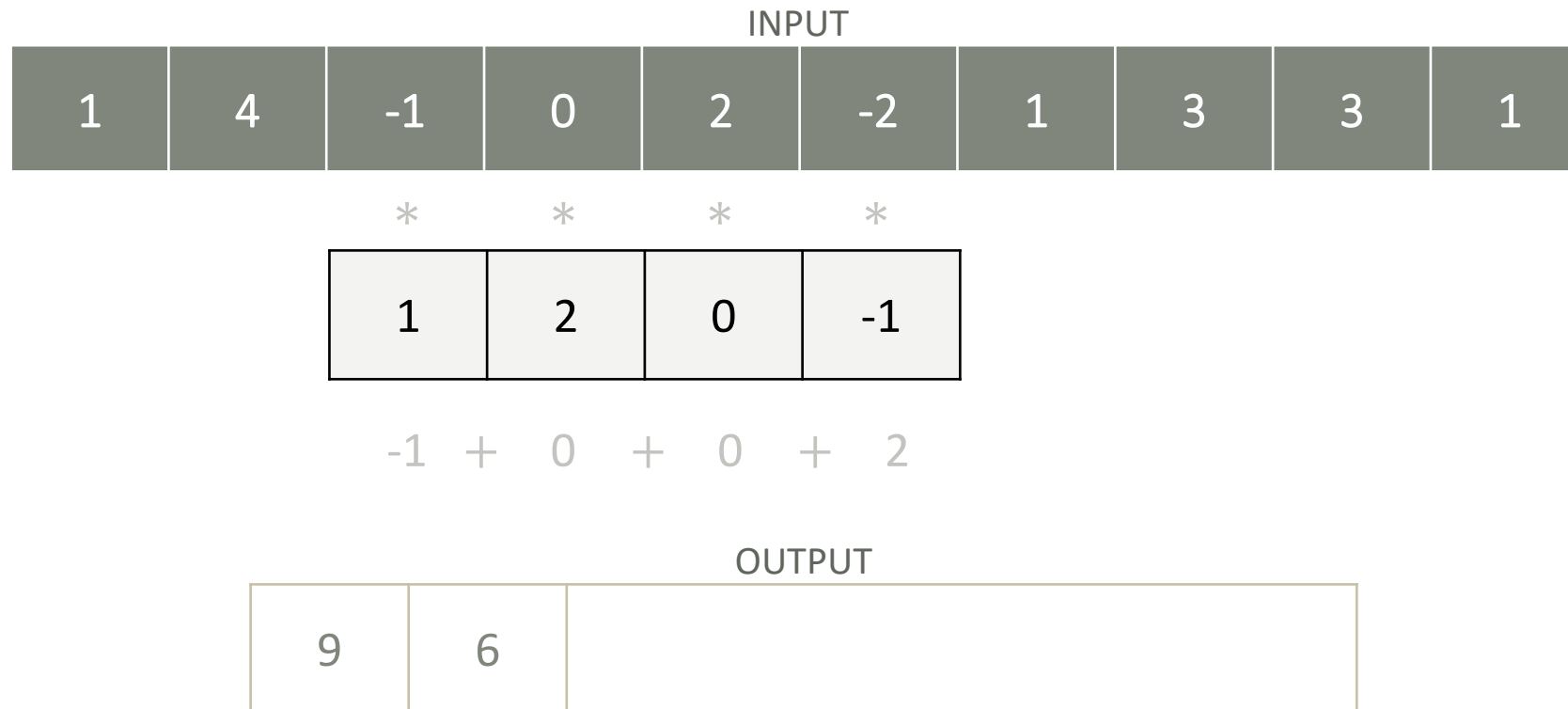
CONVOLUTIONS IN NEURAL NETWORKS



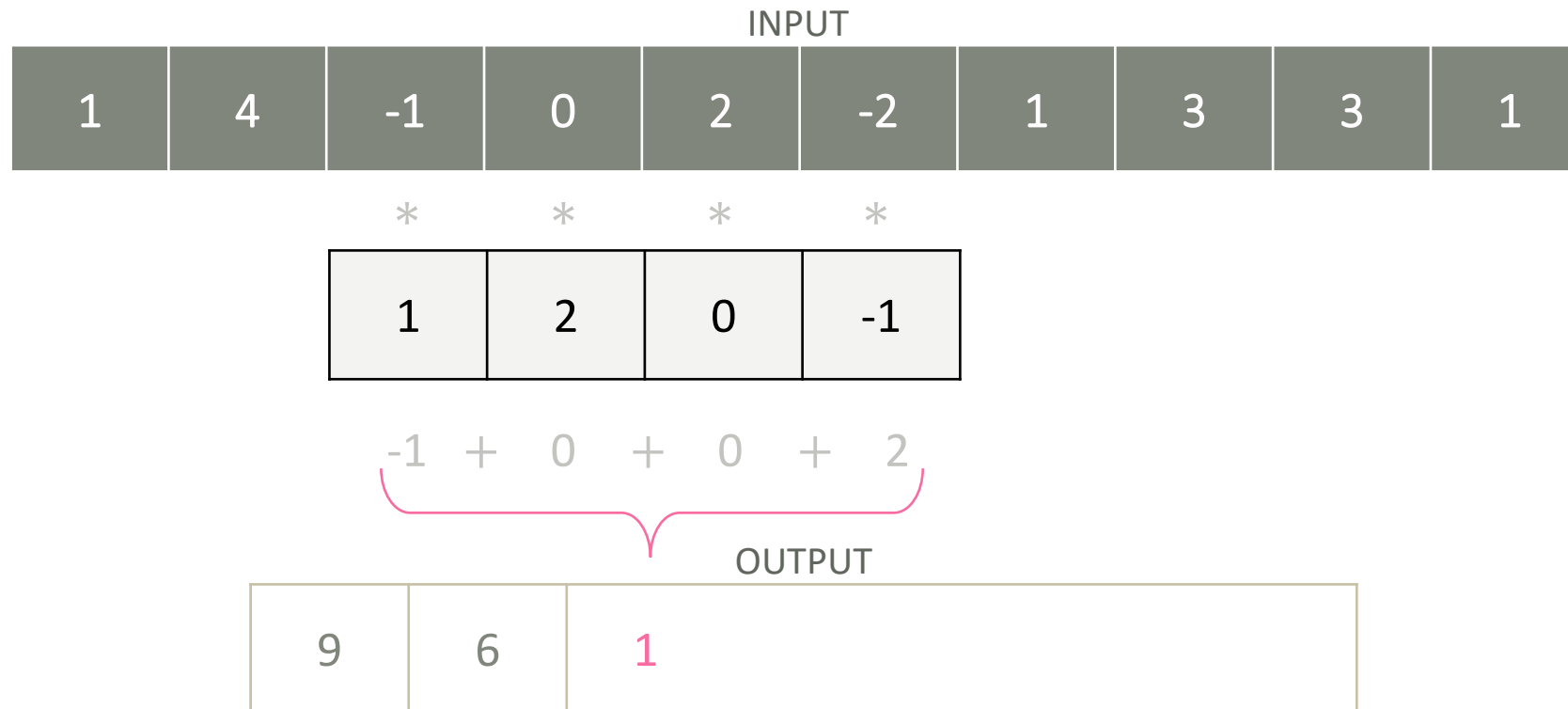
CONVOLUTIONS IN NEURAL NETWORKS



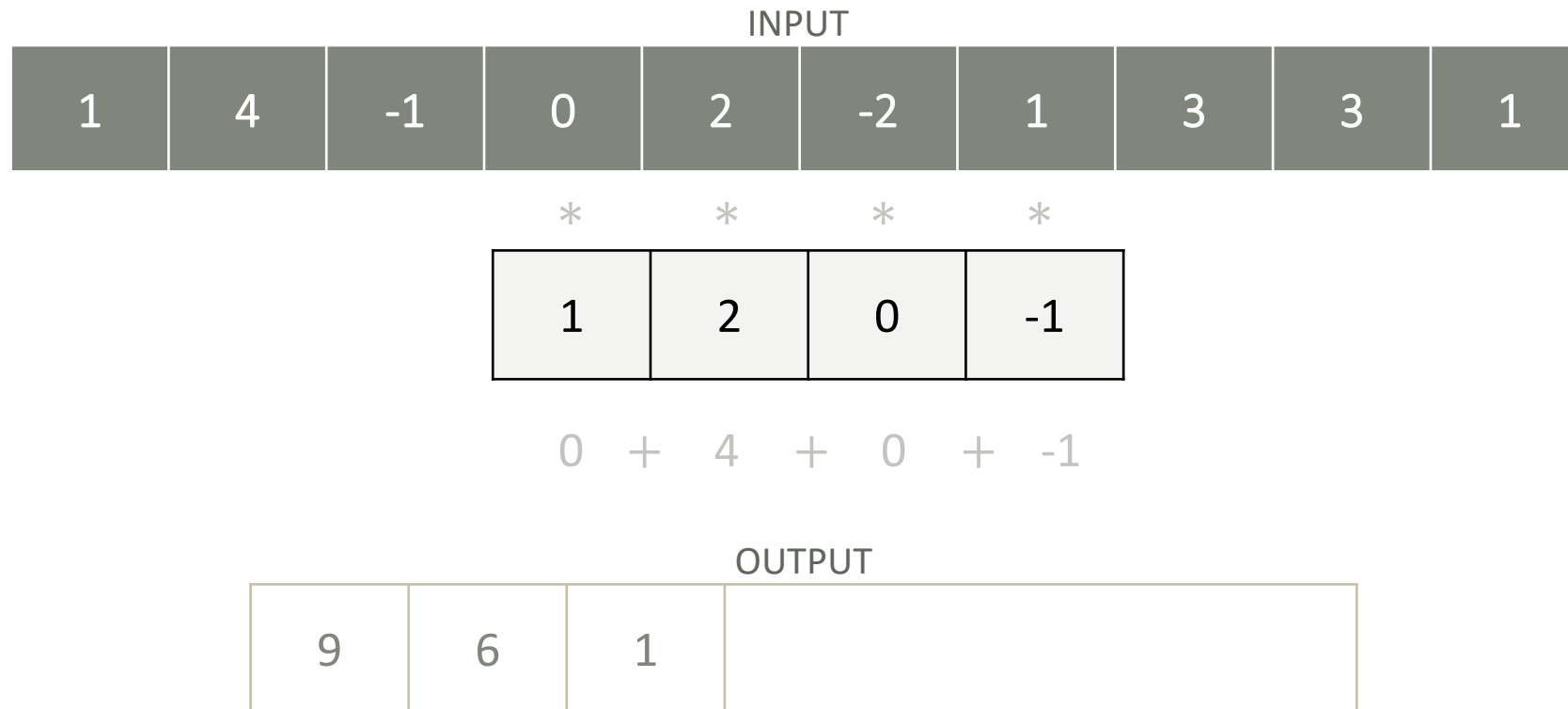
CONVOLUTIONS IN NEURAL NETWORKS



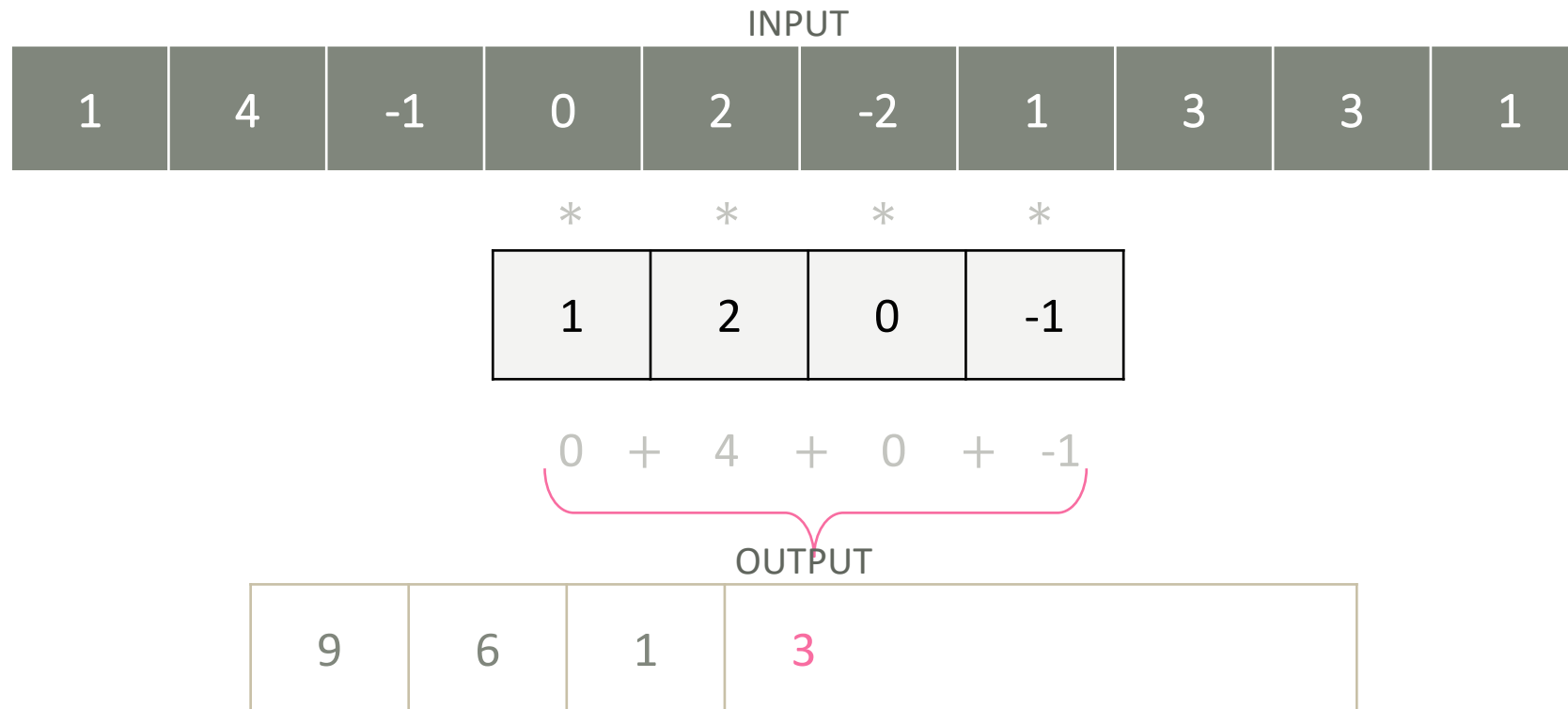
CONVOLUTIONS IN NEURAL NETWORKS



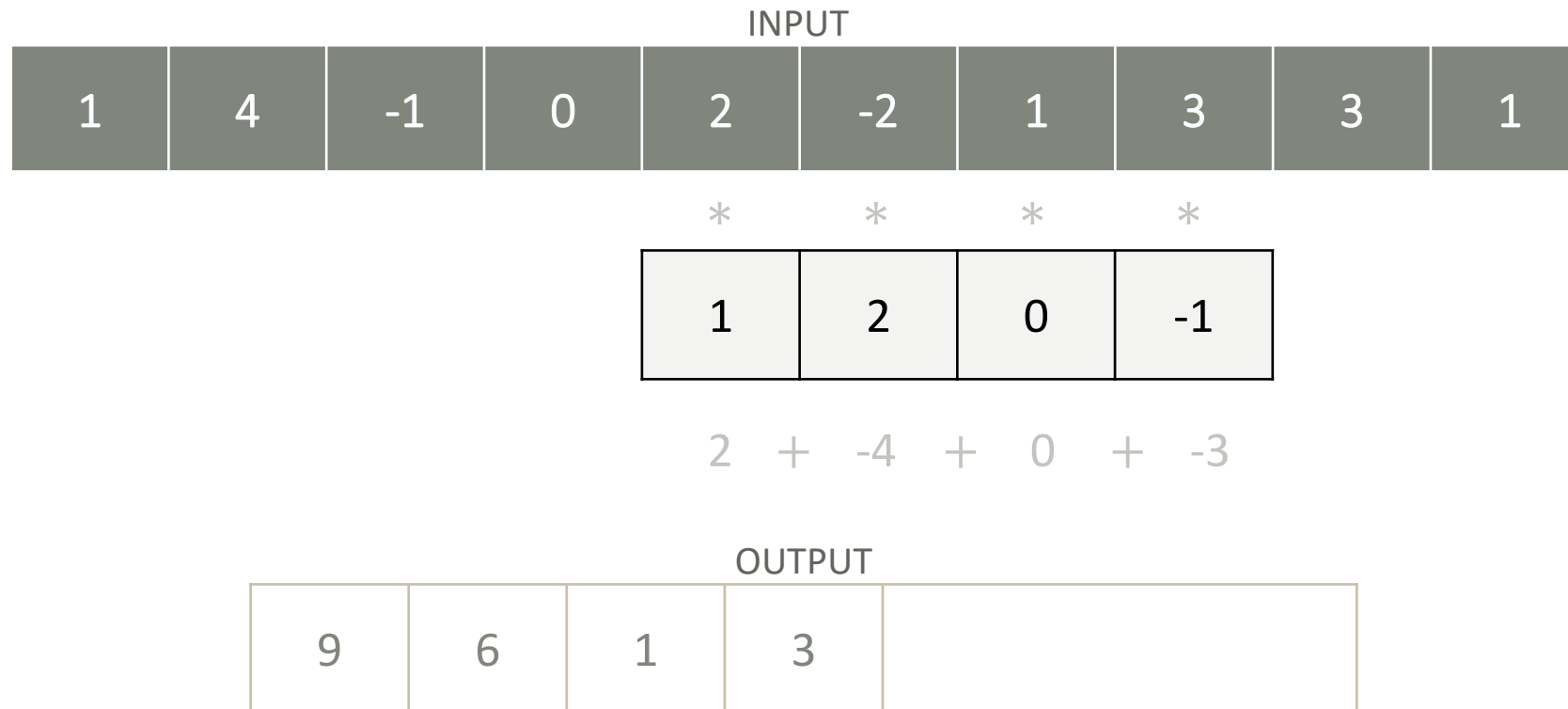
CONVOLUTIONS IN NEURAL NETWORKS



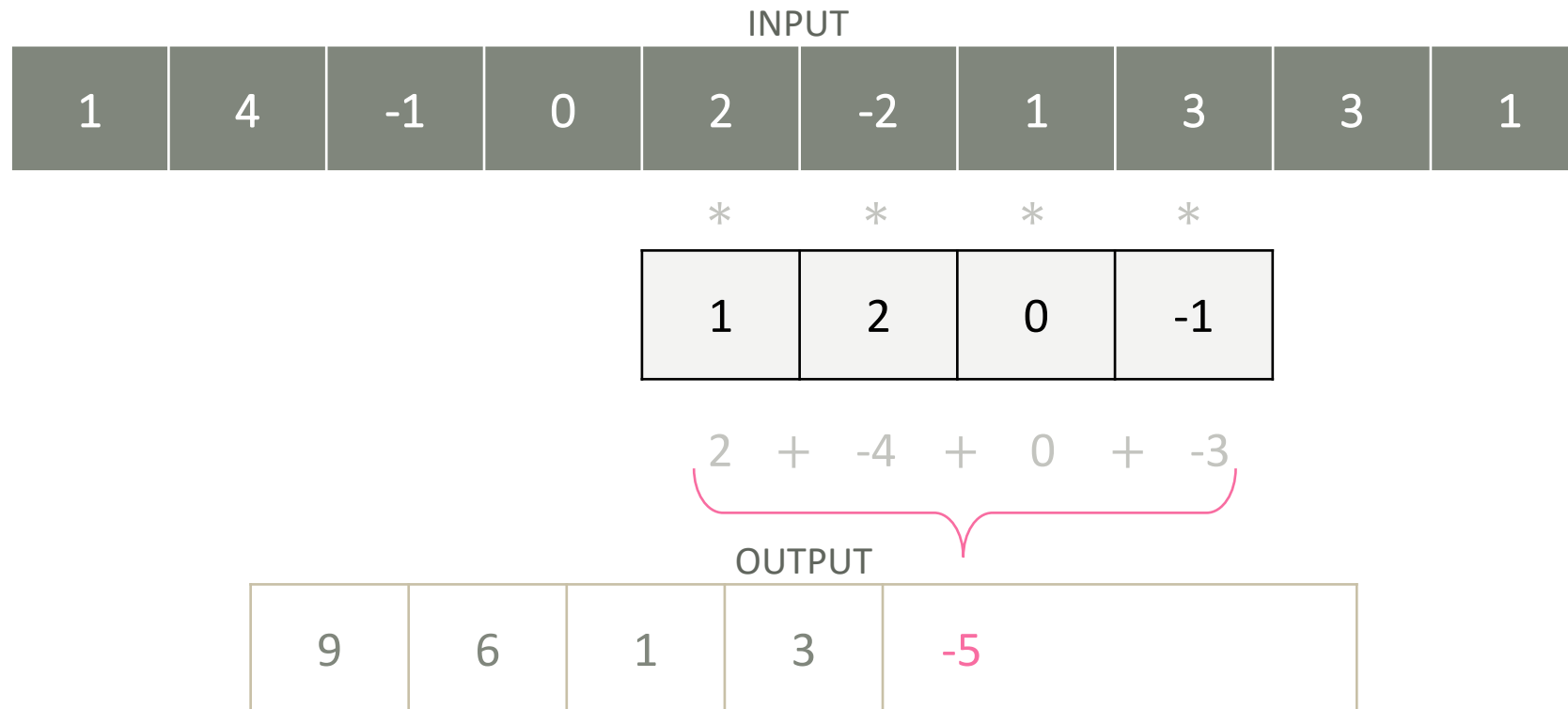
CONVOLUTIONS IN NEURAL NETWORKS



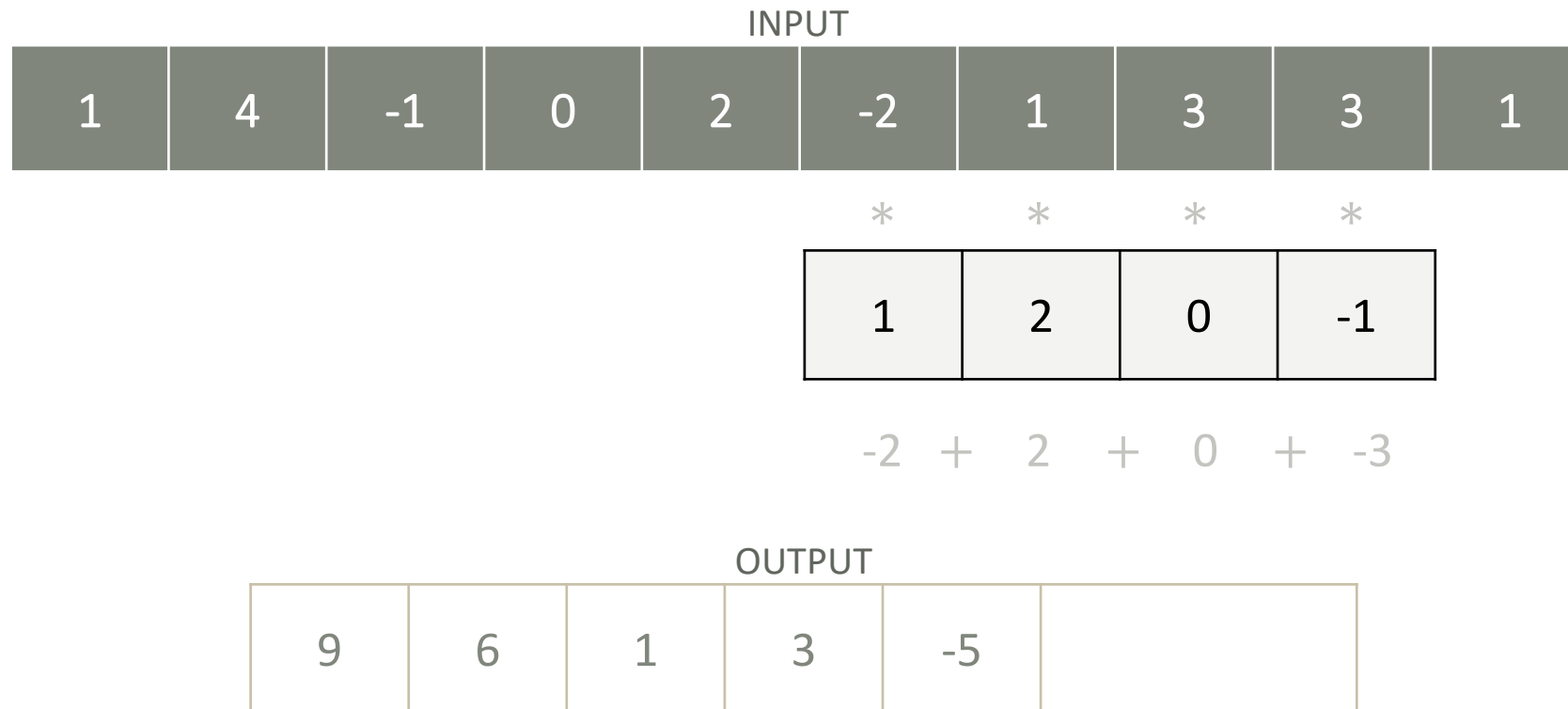
CONVOLUTIONS IN NEURAL NETWORKS



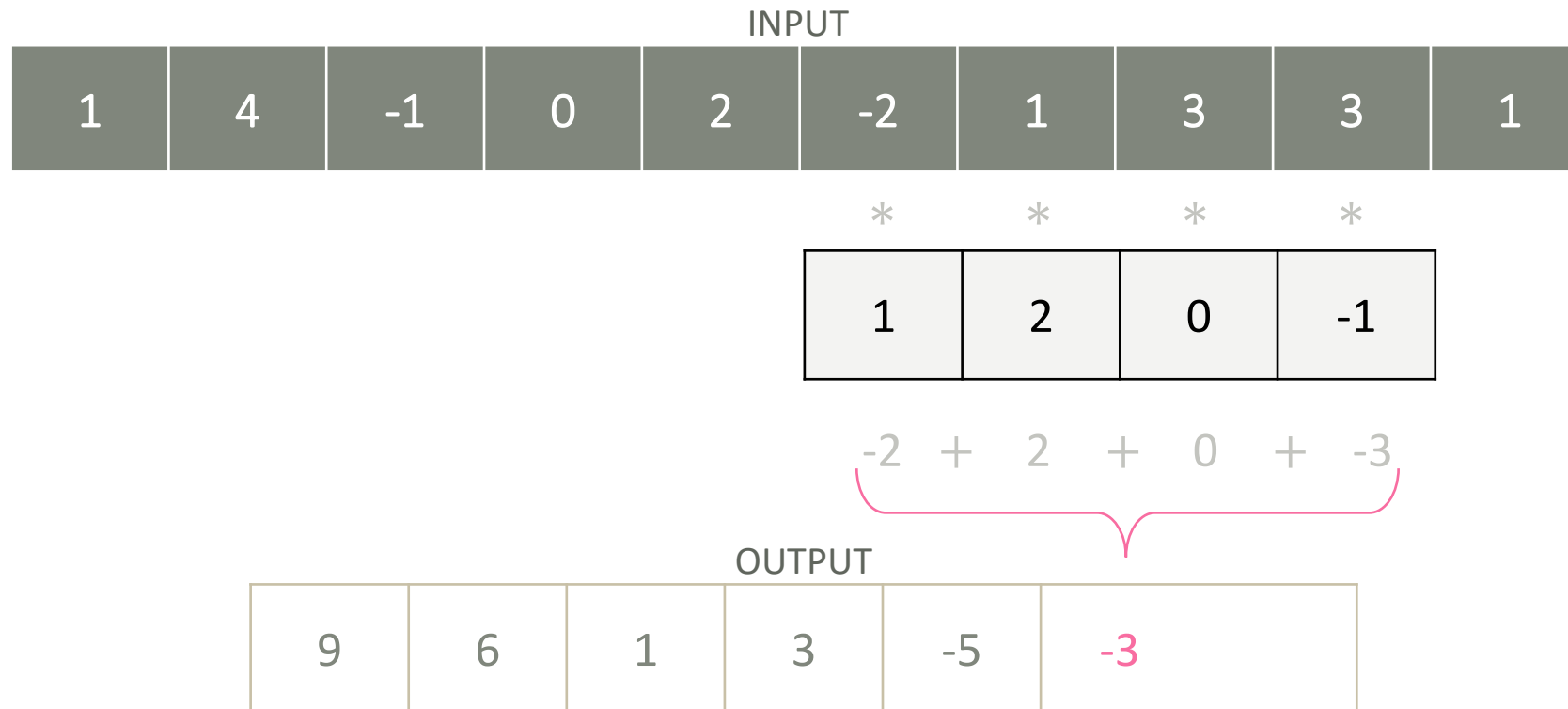
CONVOLUTIONS IN NEURAL NETWORKS



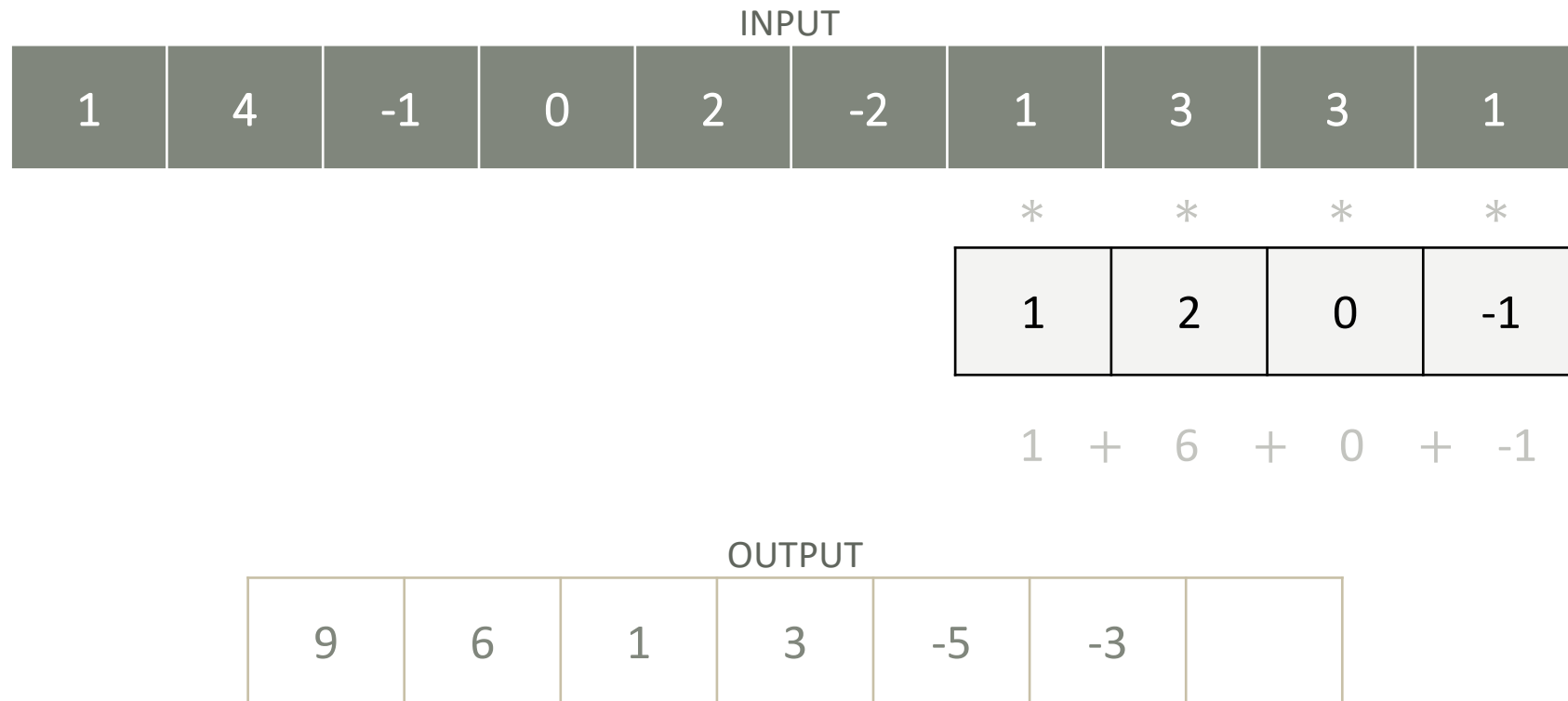
CONVOLUTIONS IN NEURAL NETWORKS



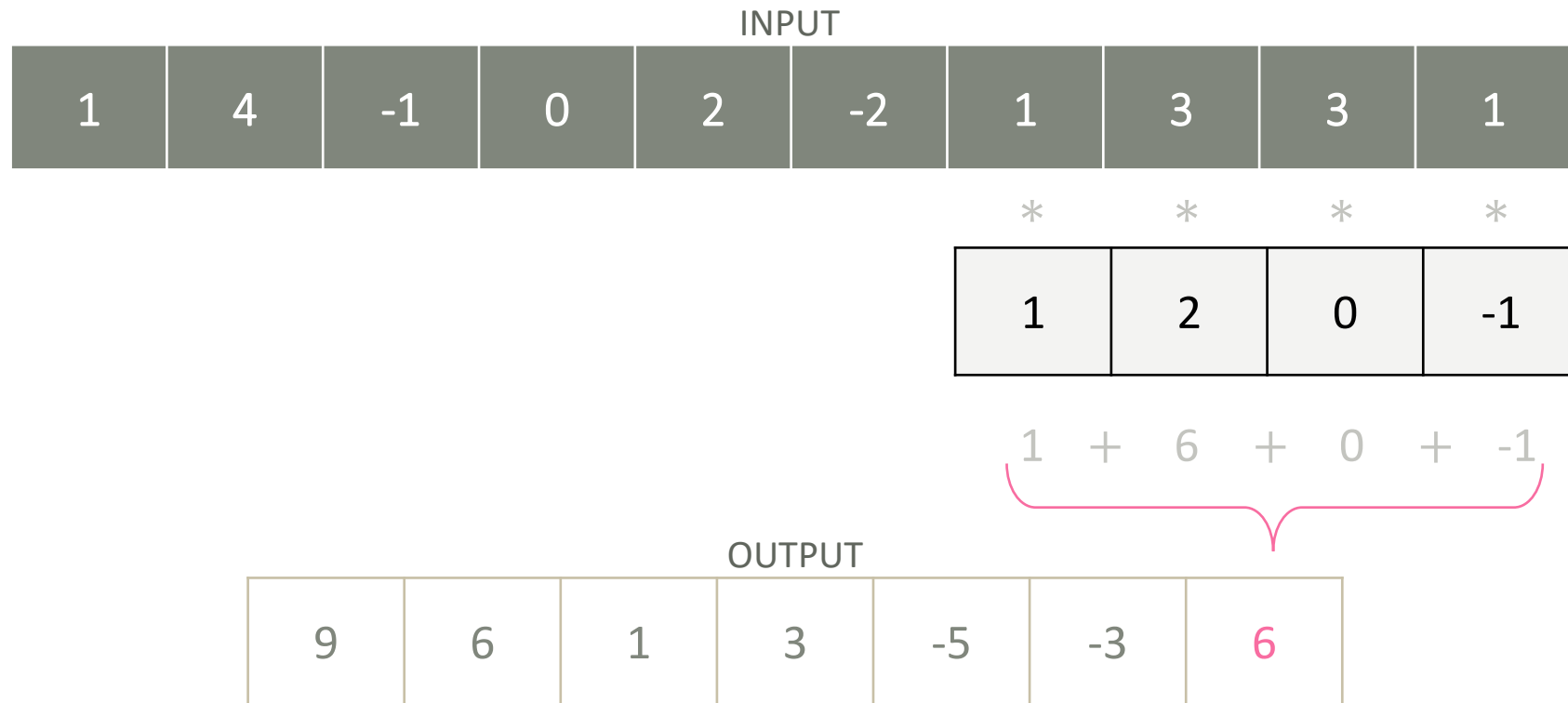
CONVOLUTIONS IN NEURAL NETWORKS



CONVOLUTIONS IN NEURAL NETWORKS



CONVOLUTIONS IN NEURAL NETWORKS



CONVOLUTIONS IN NEURAL NETWORKS

INPUT

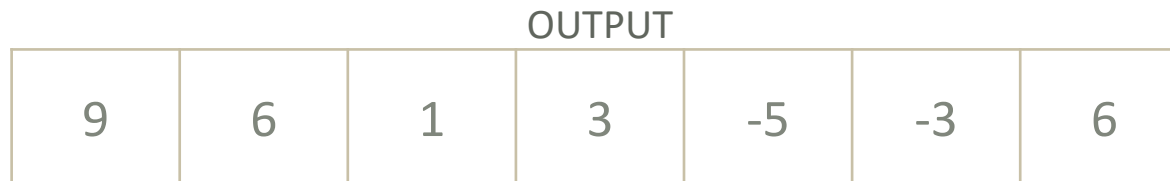
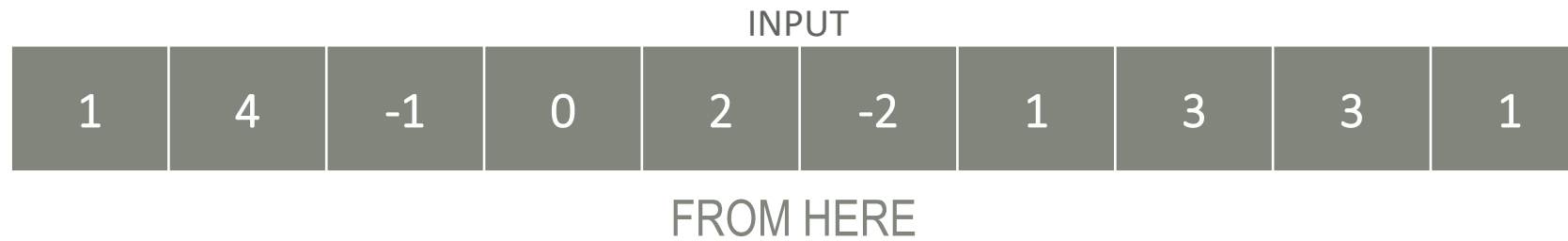
1	4	-1	0	2	-2	1	3	3	1
---	---	----	---	---	----	---	---	---	---

OUTPUT

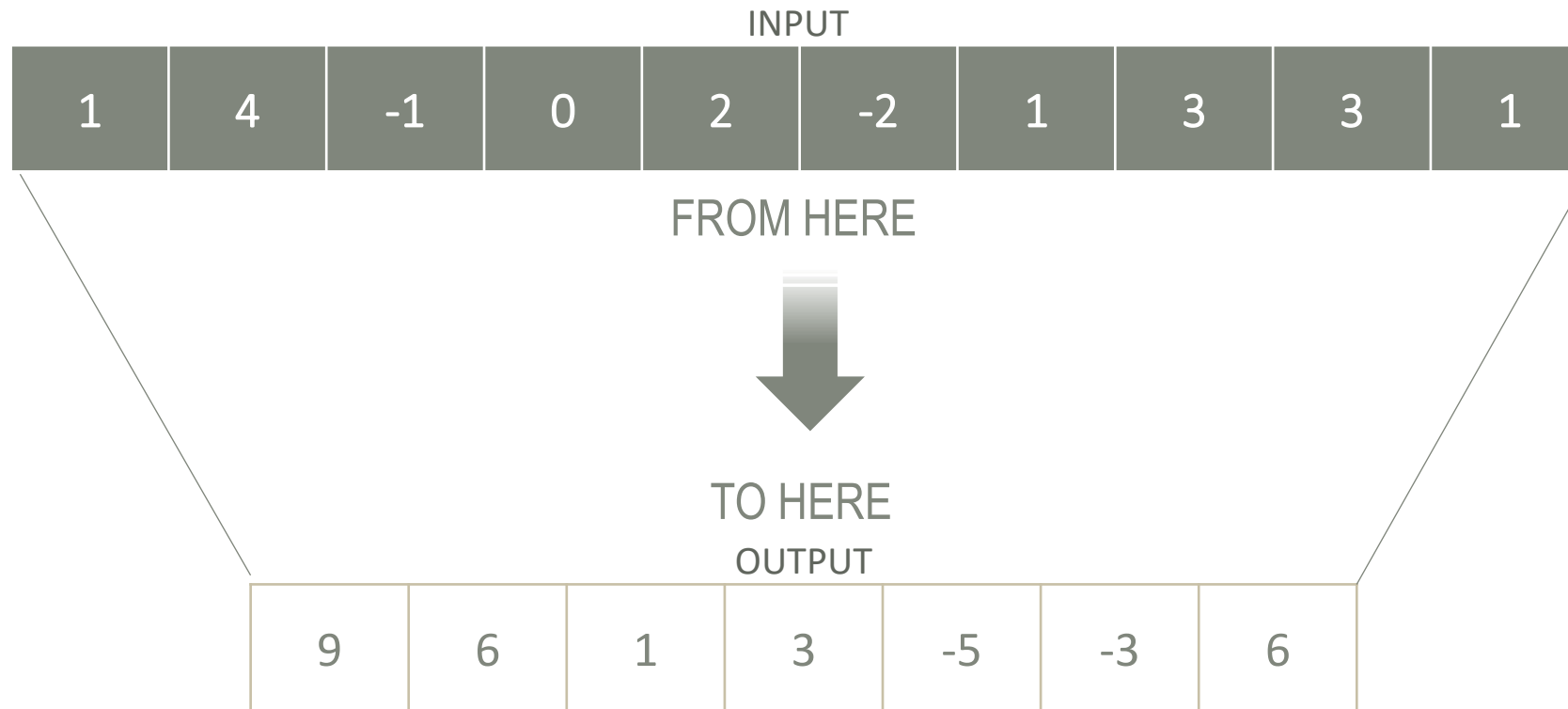
9	6	1	3	-5	-3	6
---	---	---	---	----	----	---



CONVOLUTIONS IN NEURAL NETWORKS



CONVOLUTIONS IN NEURAL NETWORKS



CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

KERNEL

0	-1	0
-1	5	-1
0	-1	0

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

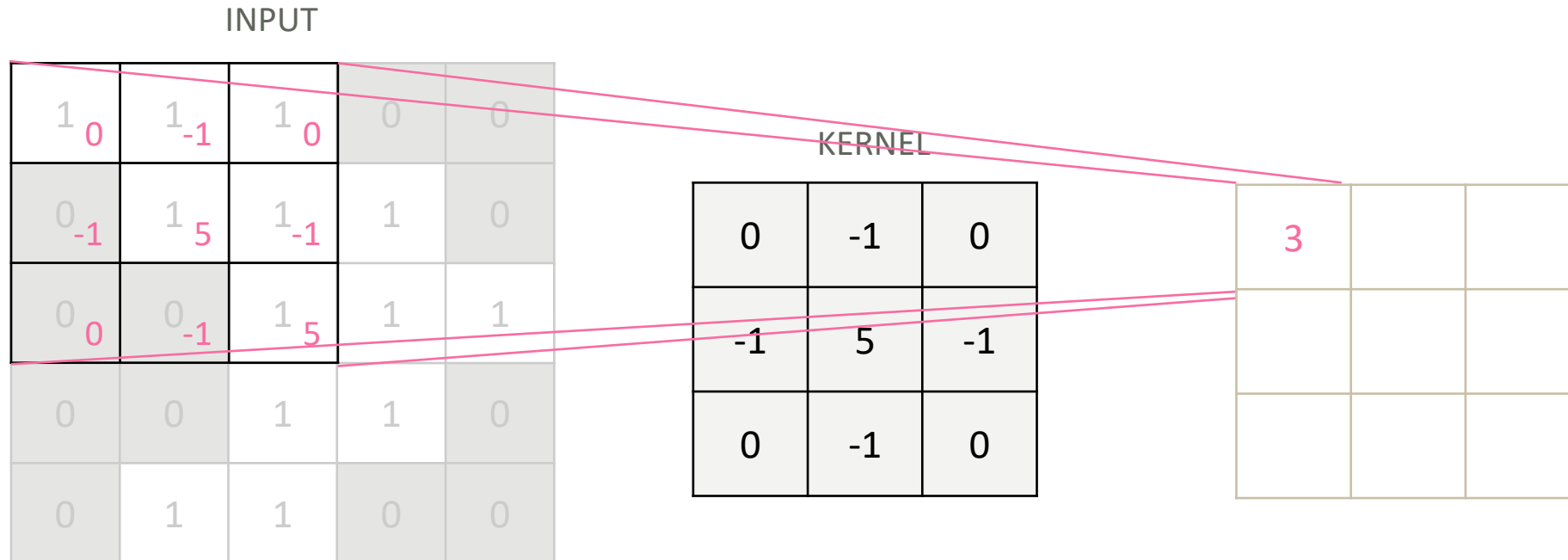
INPUT

1 0	1 -1	1 0	0	0
0 -1	1 5	1 -1	1	0
0 0	0 -1	1 0	1	1
0	0	1	1	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1 ₀	1 ₋₁	0 ₀	0
0	1 ₋₁	1 ₅	1 ₋₁	0
0	0 ₀	1 ₋₁	1 ₅	1
0	0	1	1	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1 ₀	0 ₋₁	0 ₀
0	1	1 ₋₁	1 ₅	0 ₋₁
0	0	1 ₀	1 ₋₁	1 ₅
0	0	1	1	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2		

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1 ₀	1 ₋₁	1 ₀	0
0	0 ₋₁	1 ₅	1 ₋₁	1
0	0 ₀	1 ₋₁	1 ₅	0
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2	2	

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1 ₀	1 ₋₁	0 ₀
0	0	1 ₋₁	1 ₅	1 ₋₁
0	0	1 ₀	1 ₋₁	0 ₅
0	1	1	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2	2	1

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1	1	0
0 ₀	0 ₋₁	1 ₀	1	1
0 ₋₁	0 ₅	1 ₋₁	1	0
0 ₀	1 ₋₁	1 ₅	0	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2	2	1
-2		

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1	1	0
0	0 ₀	1 ₋₁	1 ₀	1
0	0 ₋₁	1 ₅	1 ₋₁	0
0	1 ₀	1 ₋₁	0 ₅	0

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2	2	1
-2	2	

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

1	1	1	0	0
0	1	1	1	0
0	0	1 ₀	1 ₋₁	1 ₀
0	0	1 ₋₁	1 ₅	0 ₋₁
0	1	1 ₀	0 ₋₁	0 ₀

KERNEL

0	-1	0
-1	5	-1
0	-1	0

3	4	3
-2	2	1
-2	2	3

CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS

INPUT

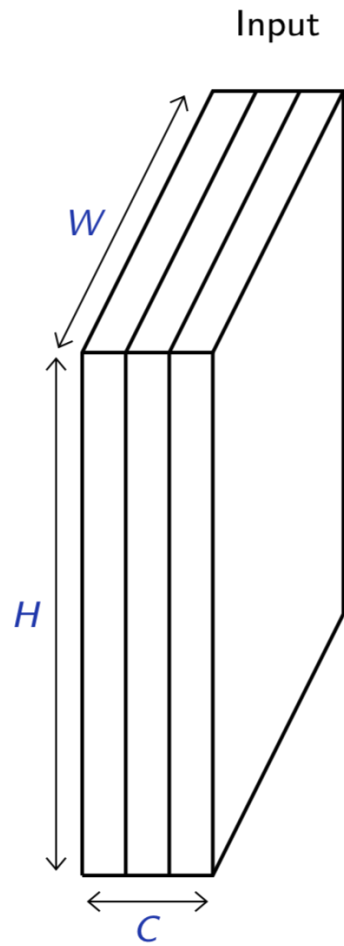
1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

KERNEL

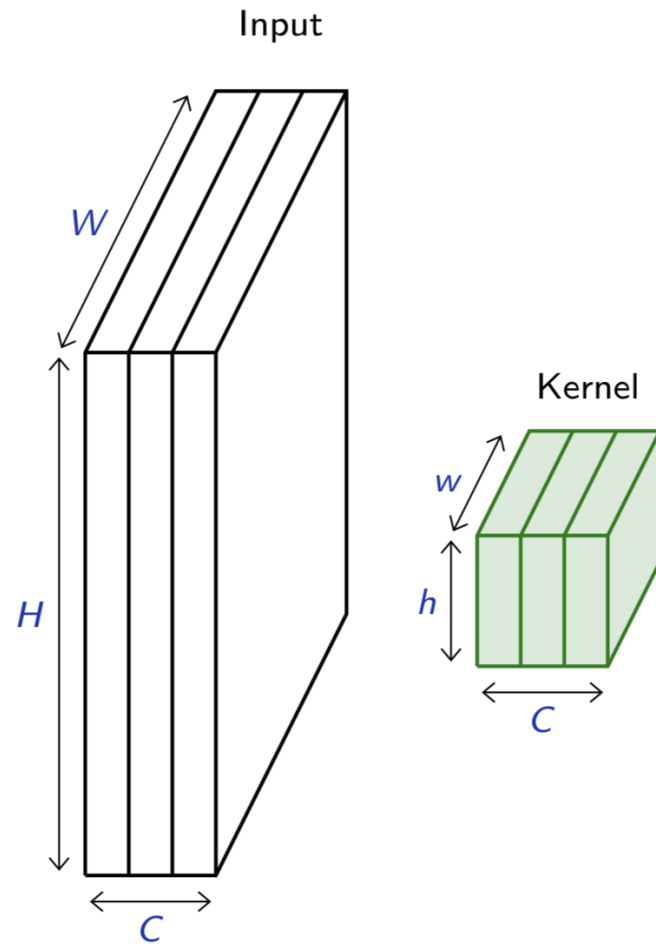
0	-1	0
-1	5	-1
0	-1	0

3	1	3
-2	2	1
-2	2	3

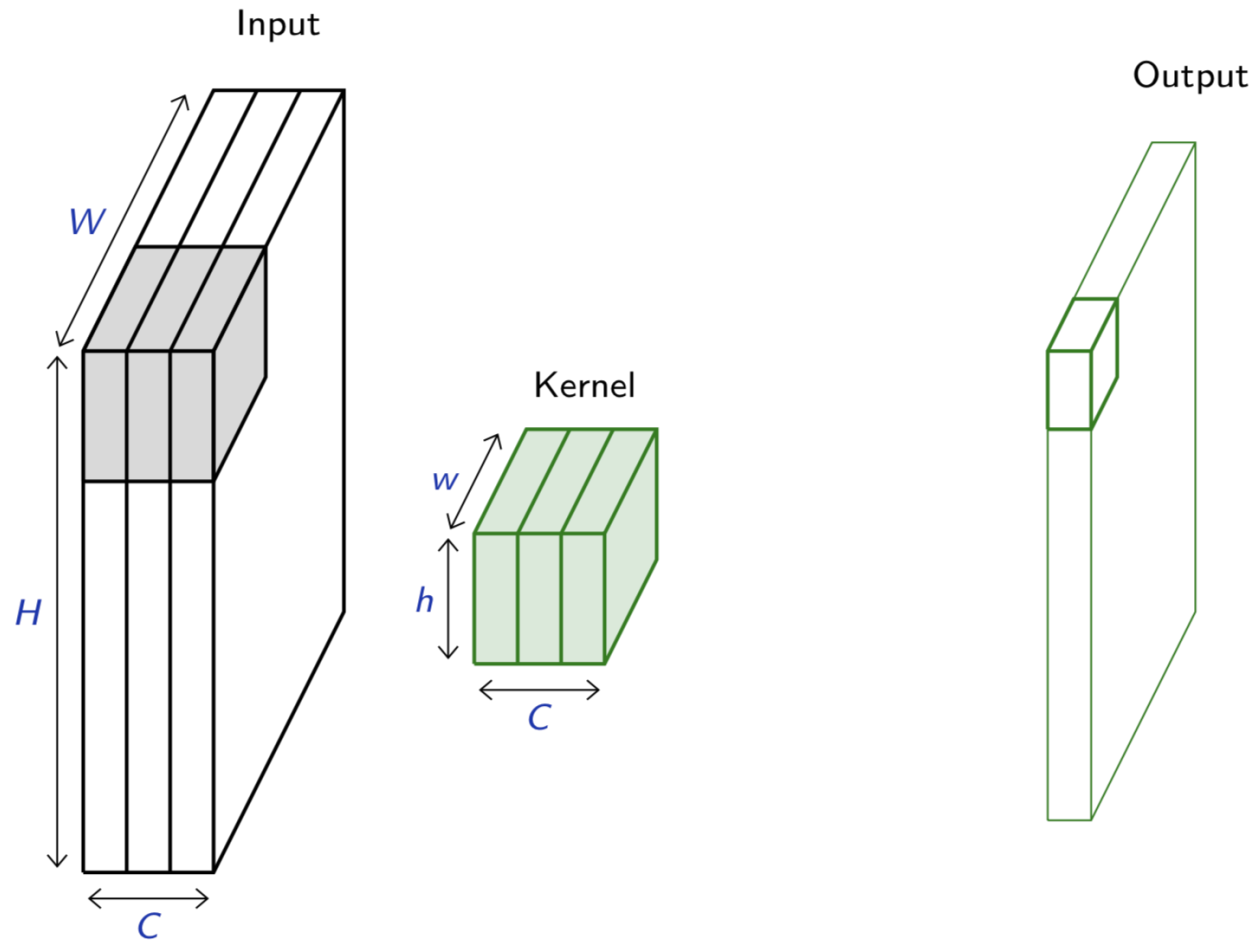
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



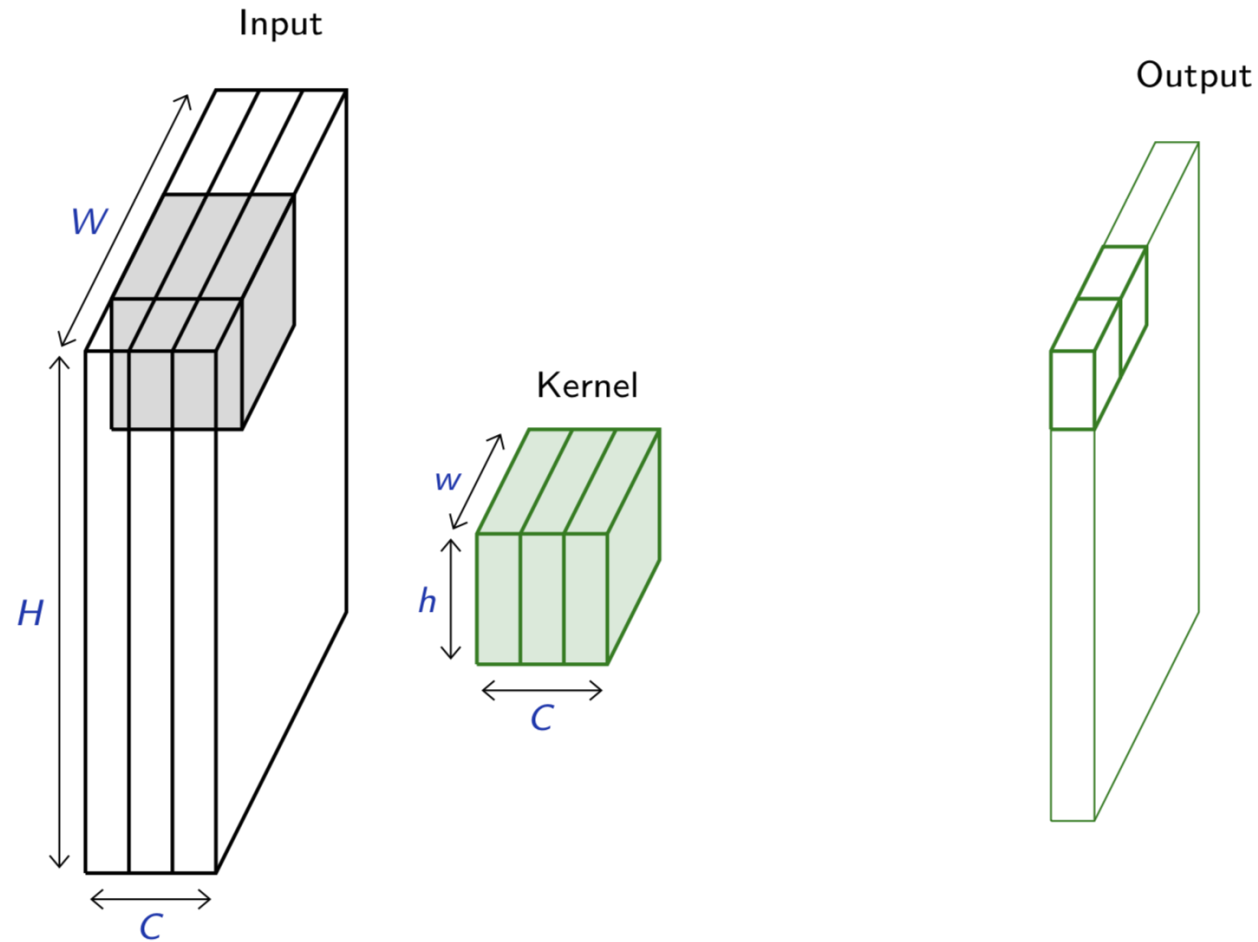
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



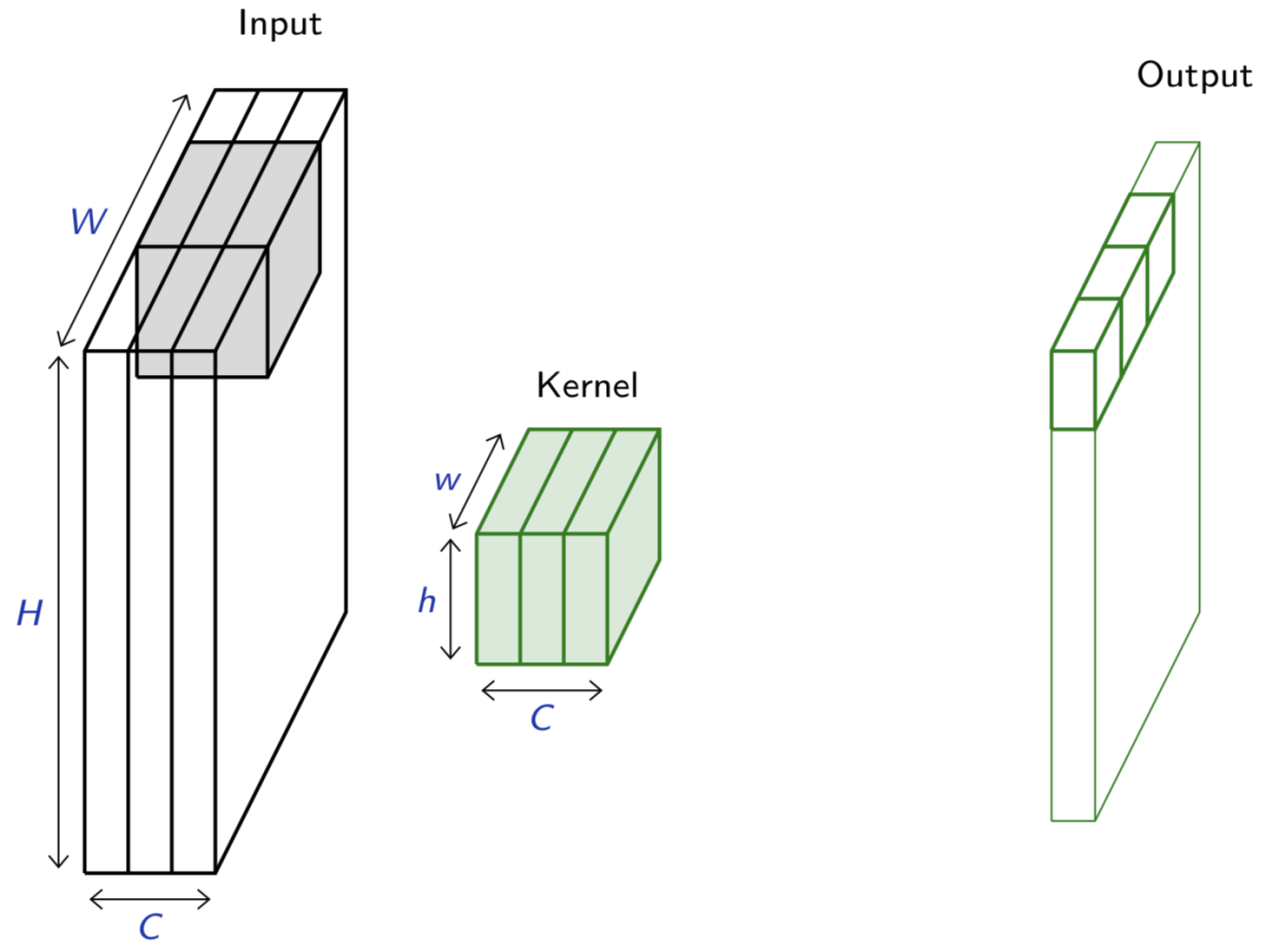
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



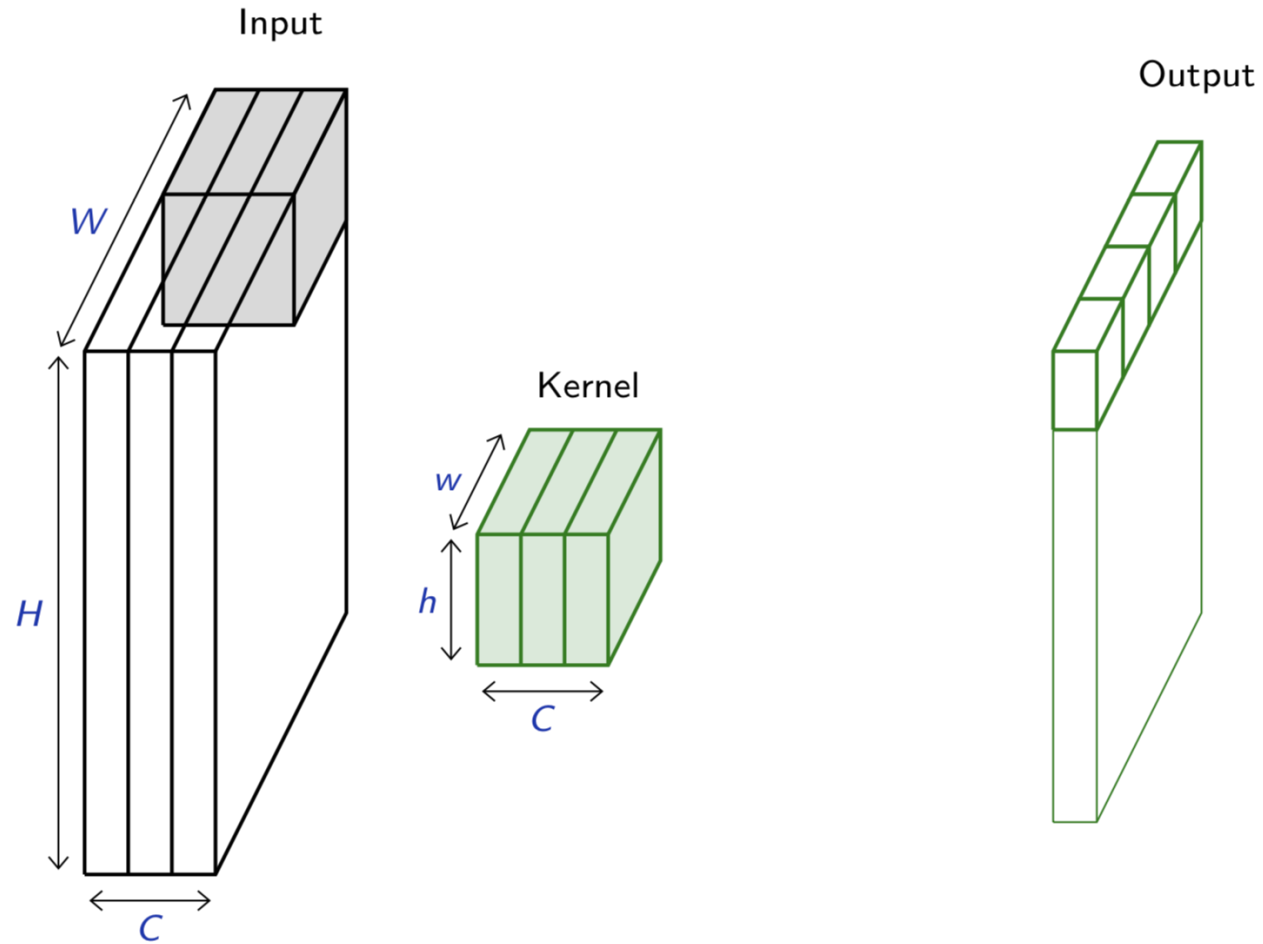
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



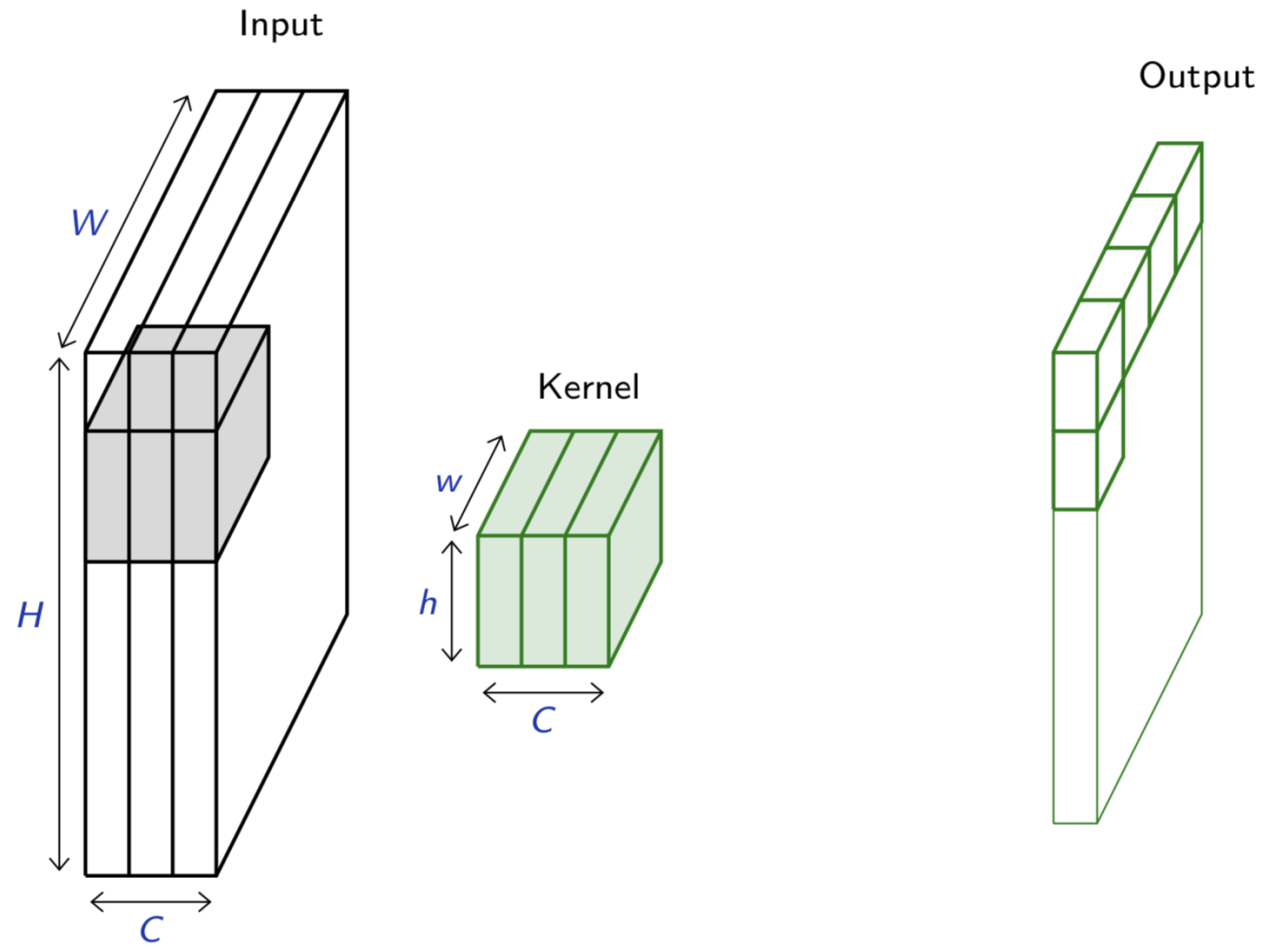
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



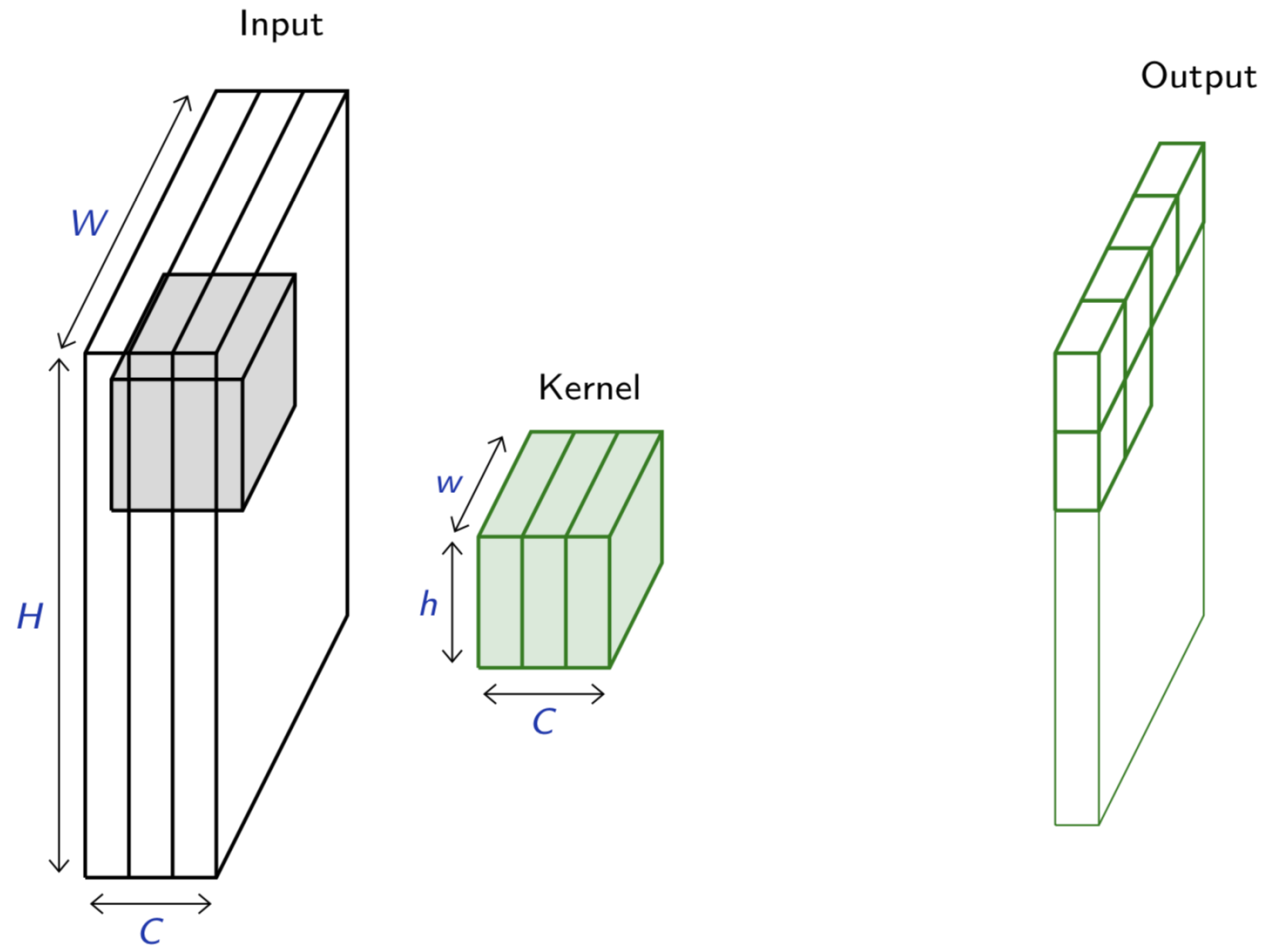
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



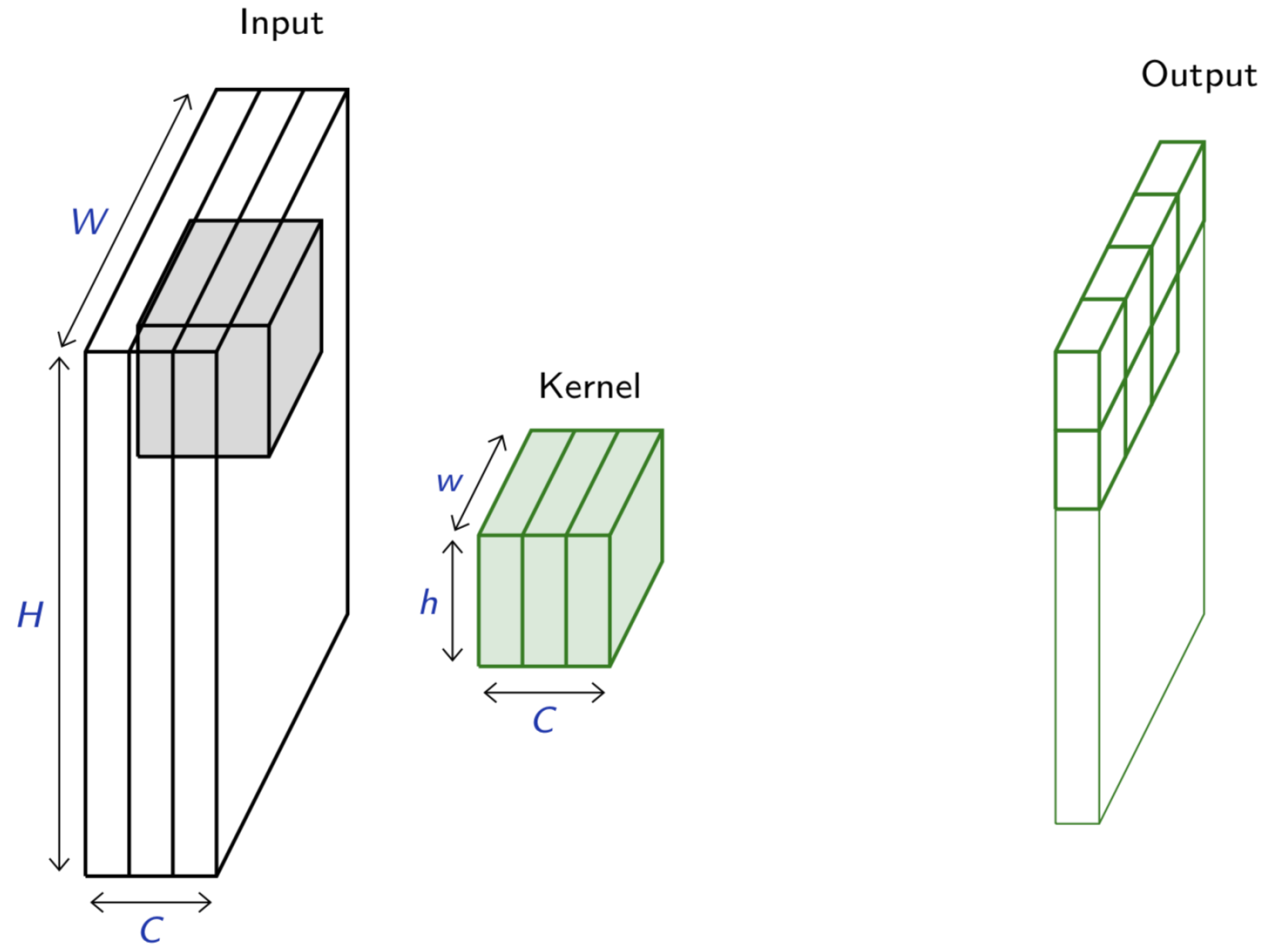
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



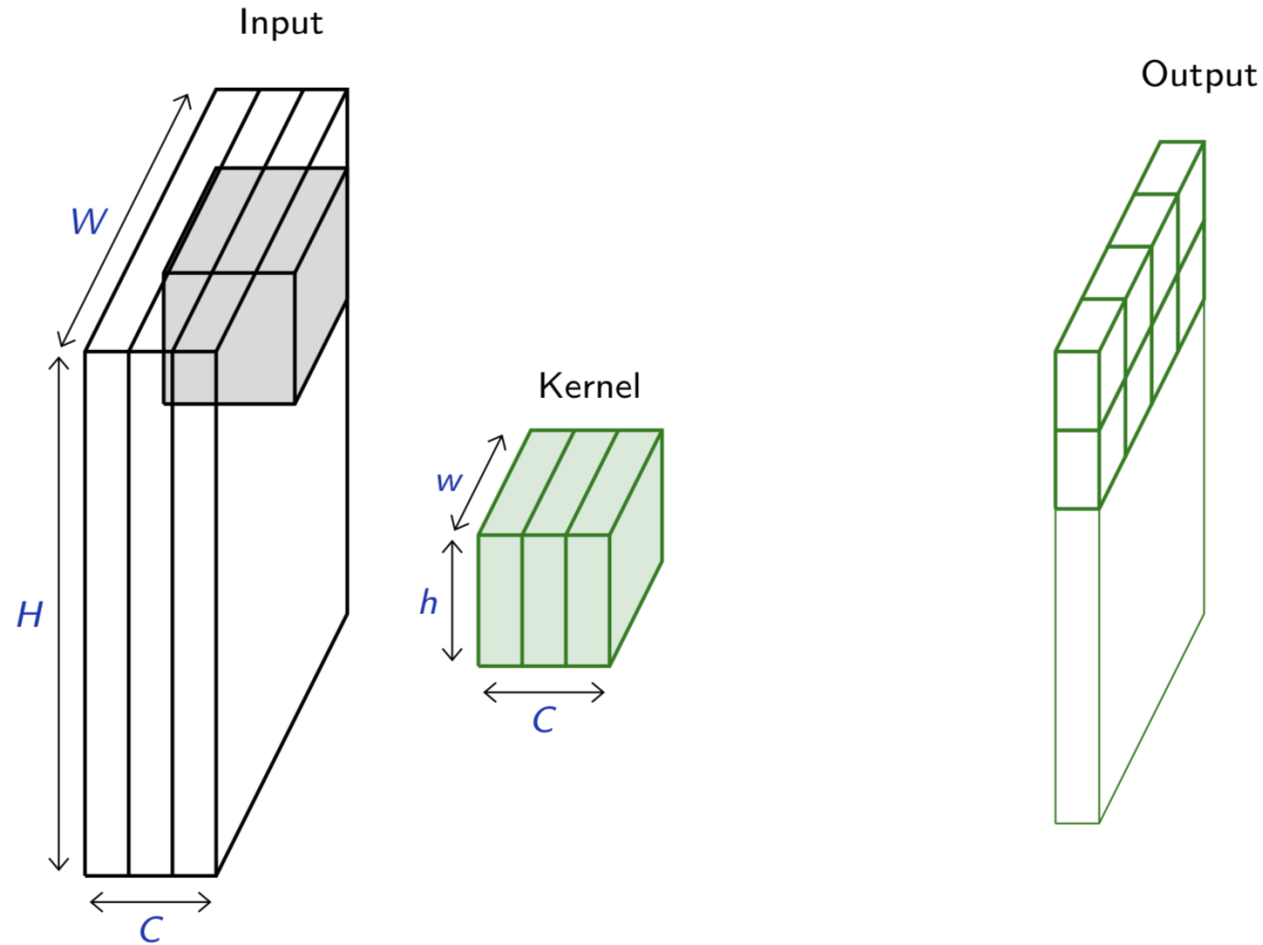
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



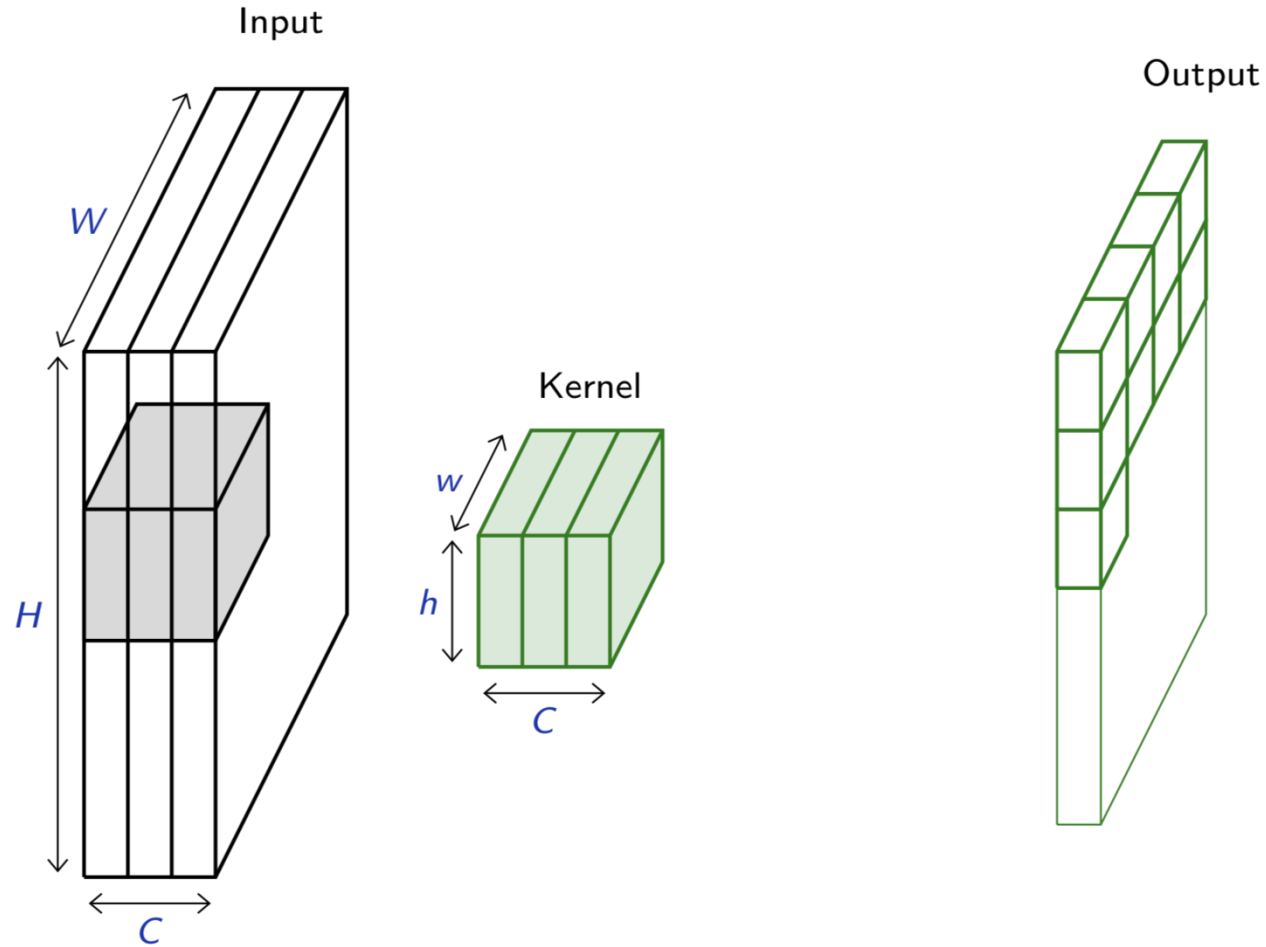
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



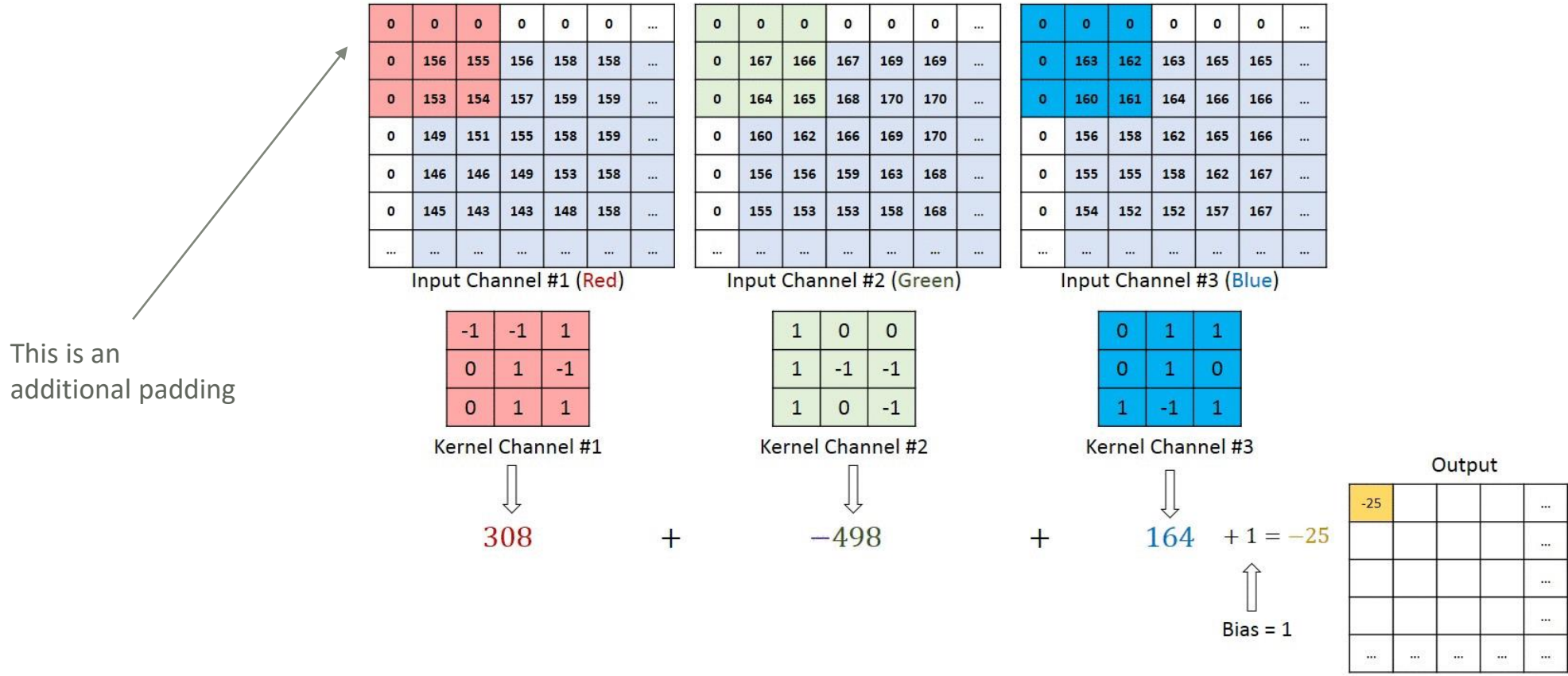
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



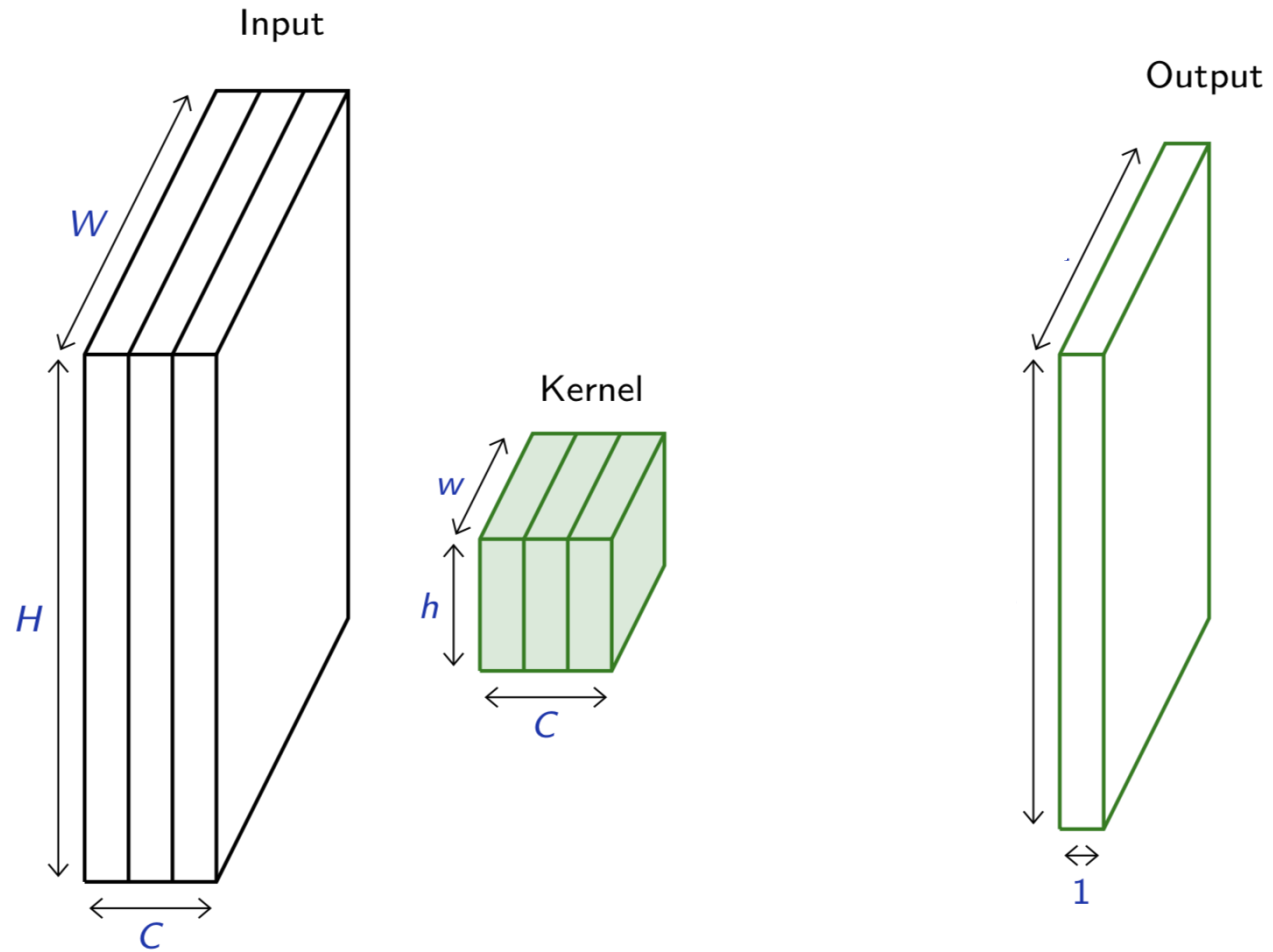
CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



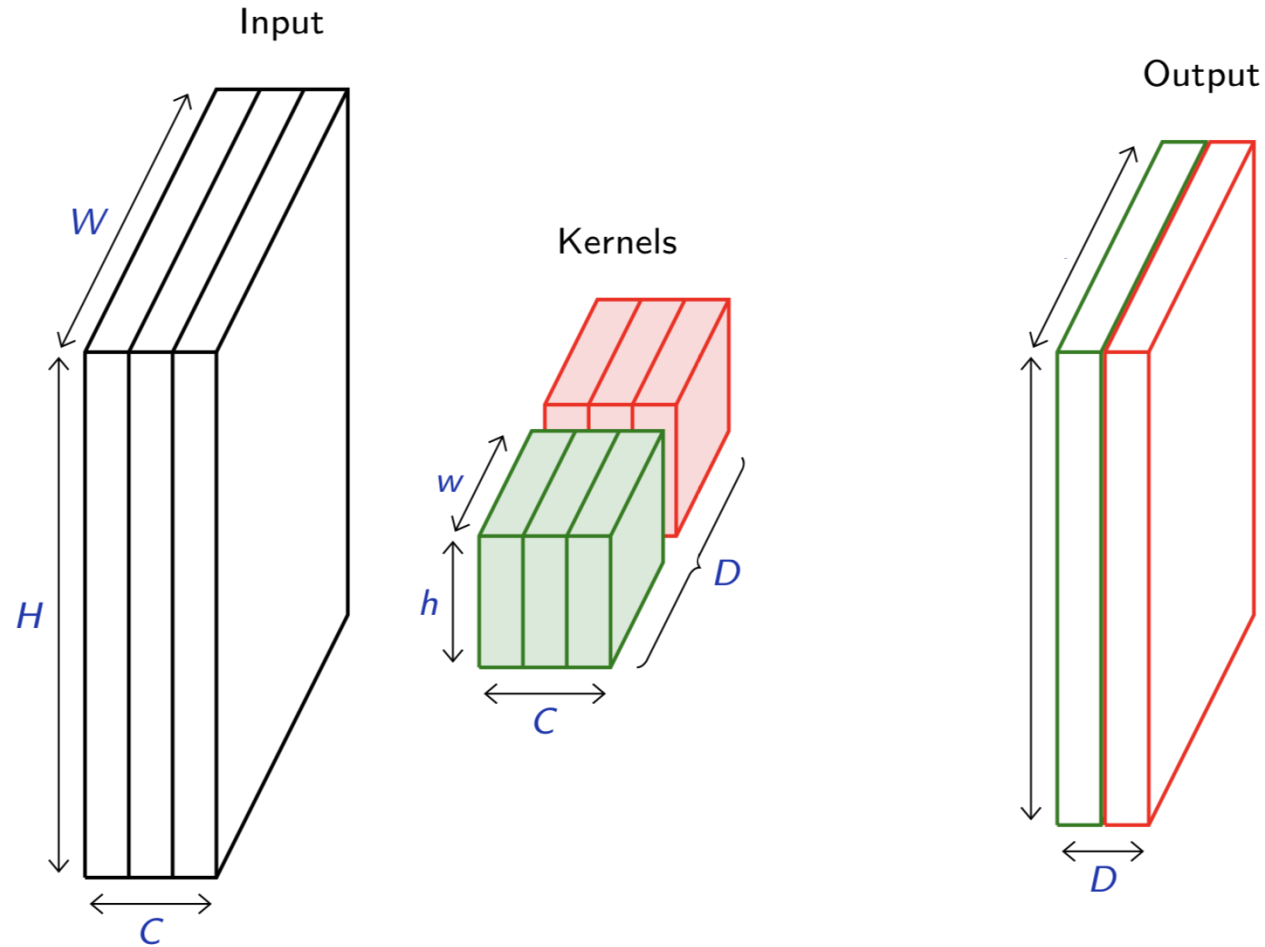
<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>



CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



CONVOLUTIONS IN NEURAL NETWORKS: 2D CONVOLUTIONS



2D CONVOLUTION HANDS-ON



2D Convolution

Set the desired settings and hover the elements to see how **input**, **filters**, and **output** are related.

Settings

Input width

5

Filter width

3

Output channels

2

Input height

5

Filter height

3

Input channels

3

Padding

0

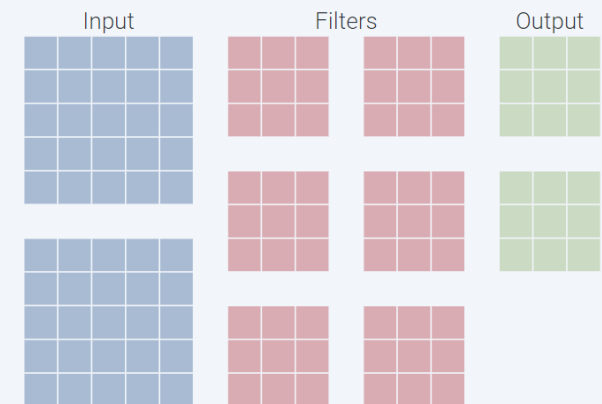
? Dilation

1

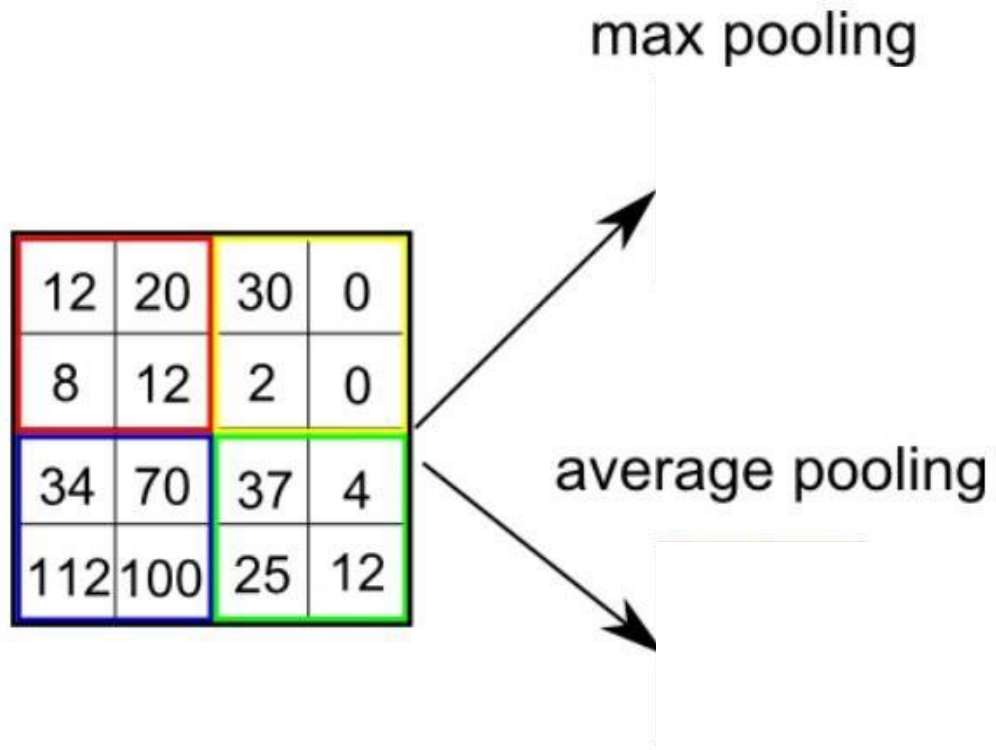
? Stride

1

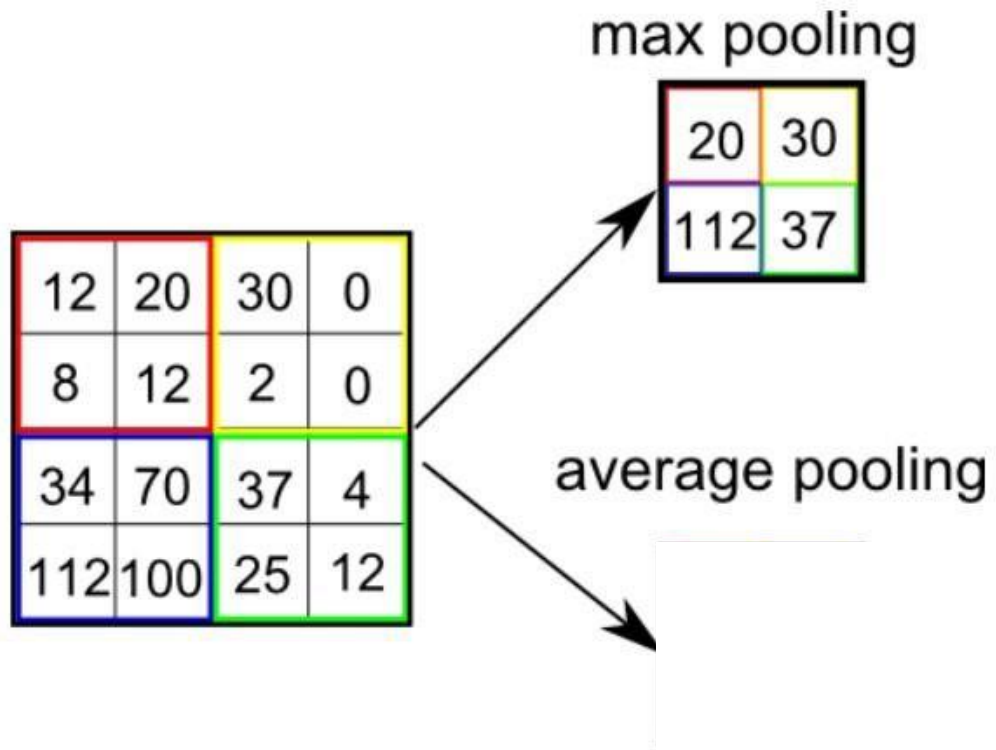
Preview



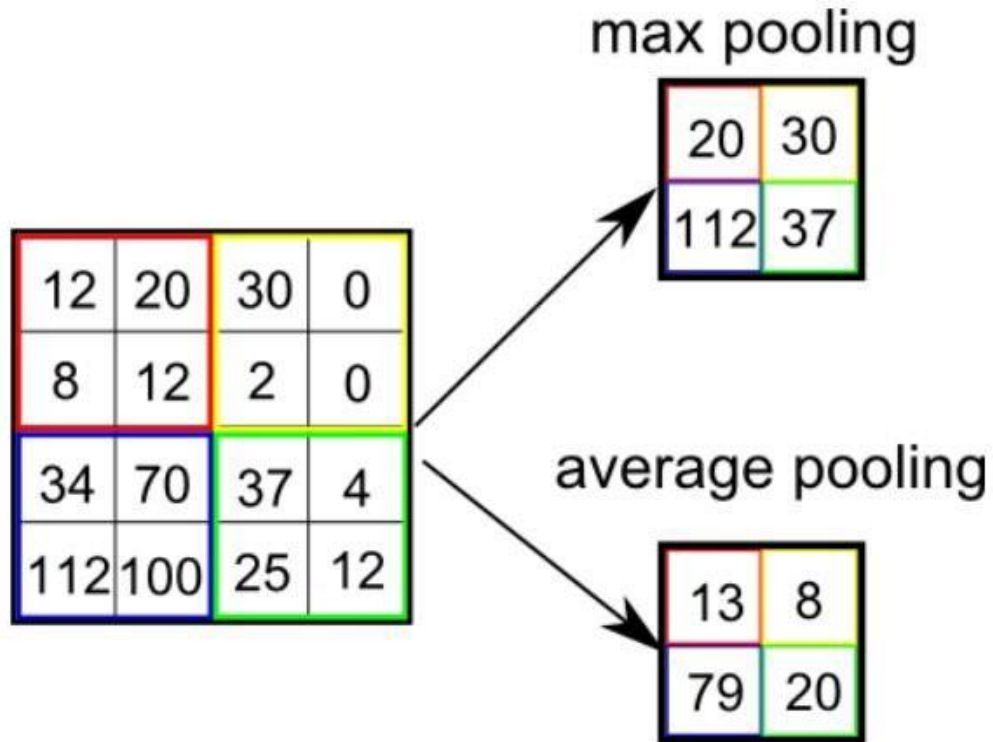
CONVOLUTIONAL NEURAL NETWORKS: POOLING



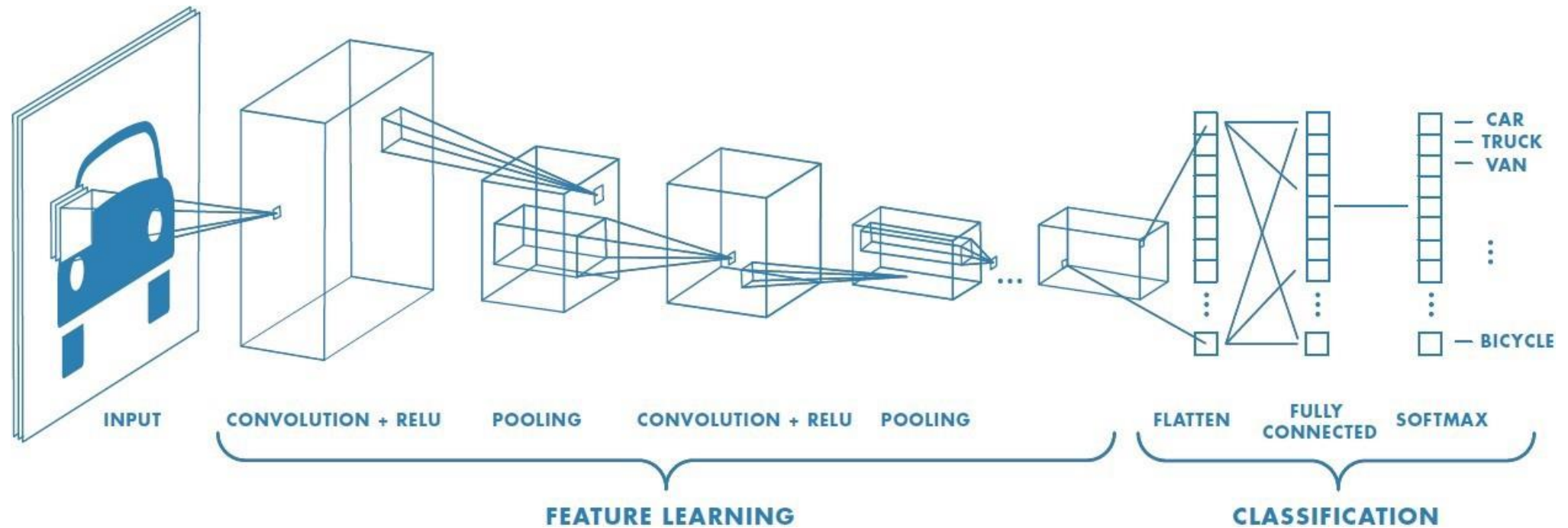
CONVOLUTIONAL NEURAL NETWORKS: POOLING



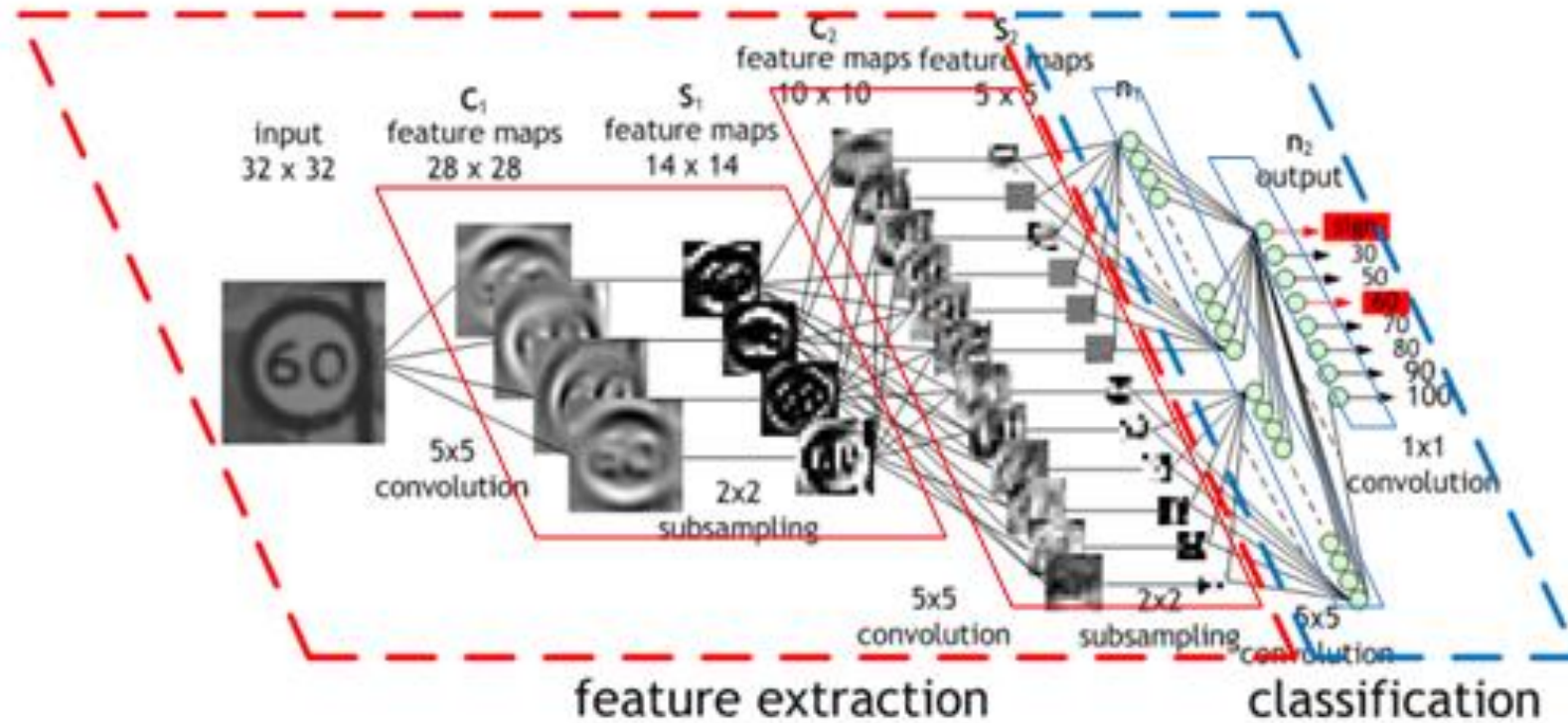
CONVOLUTIONAL NEURAL NETWORKS: POOLING



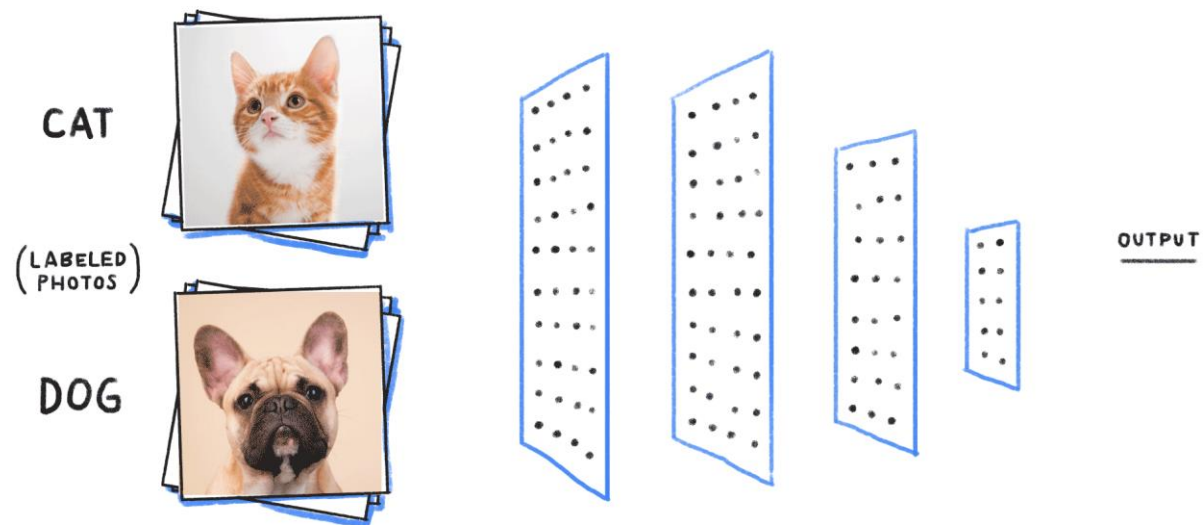
CONVOLUTIONAL NEURAL NETWORKS (CNN's)



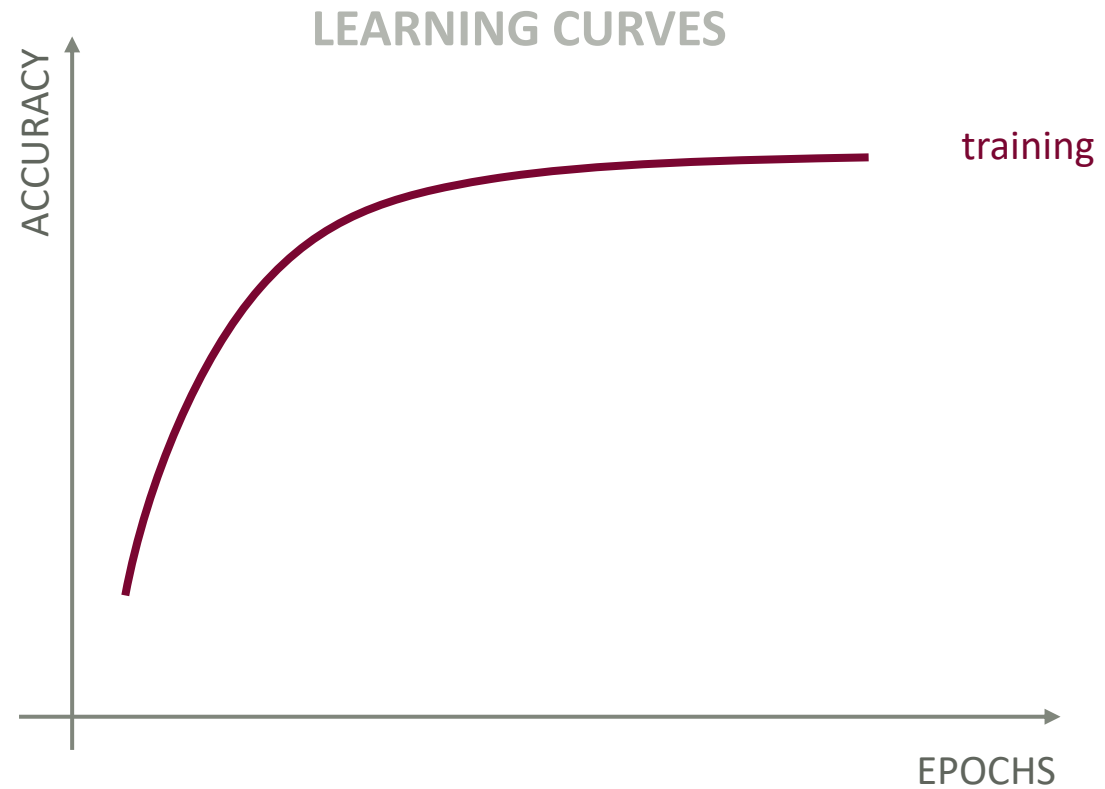
CONVOLUTIONAL NEURAL NETWORKS (CNN's)



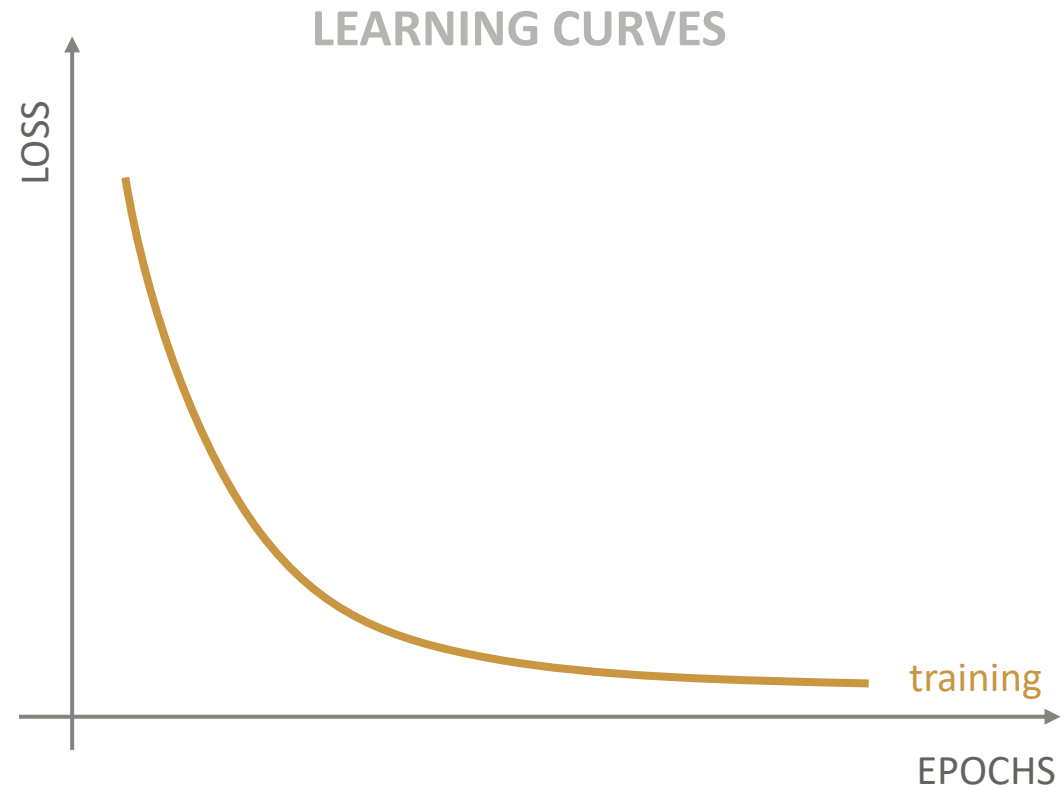
TIME TO CODE!



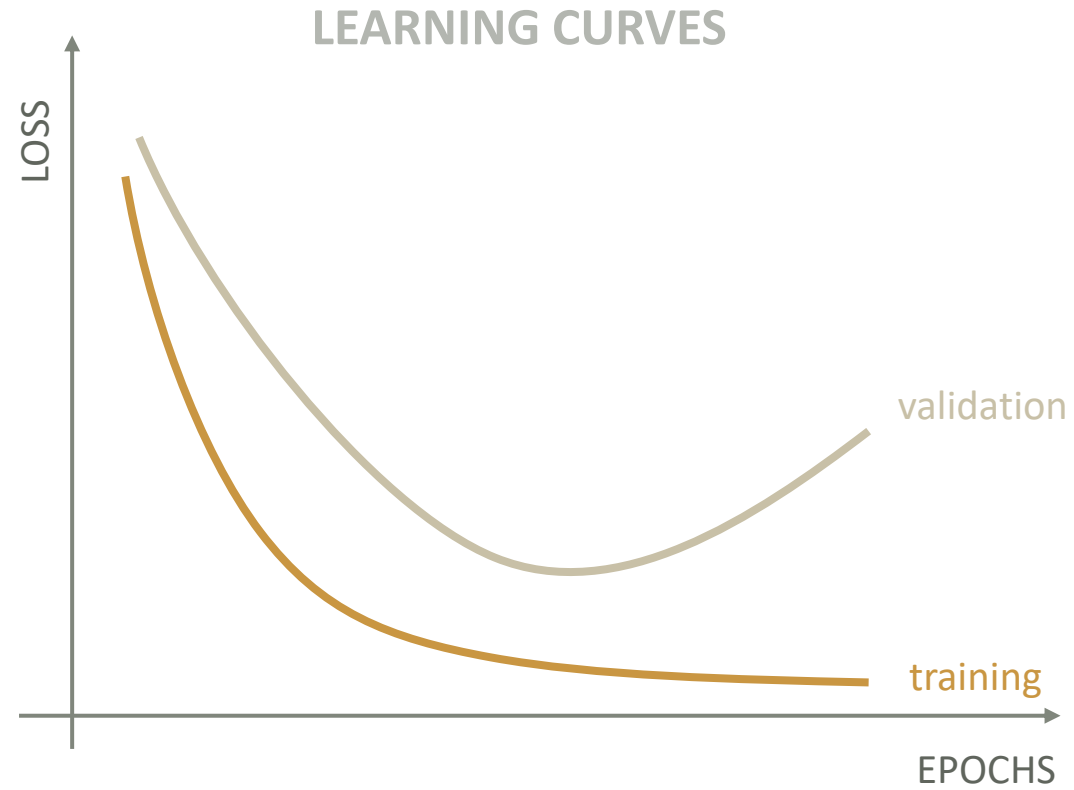
EVALUATION OF NEURAL NETWORKS



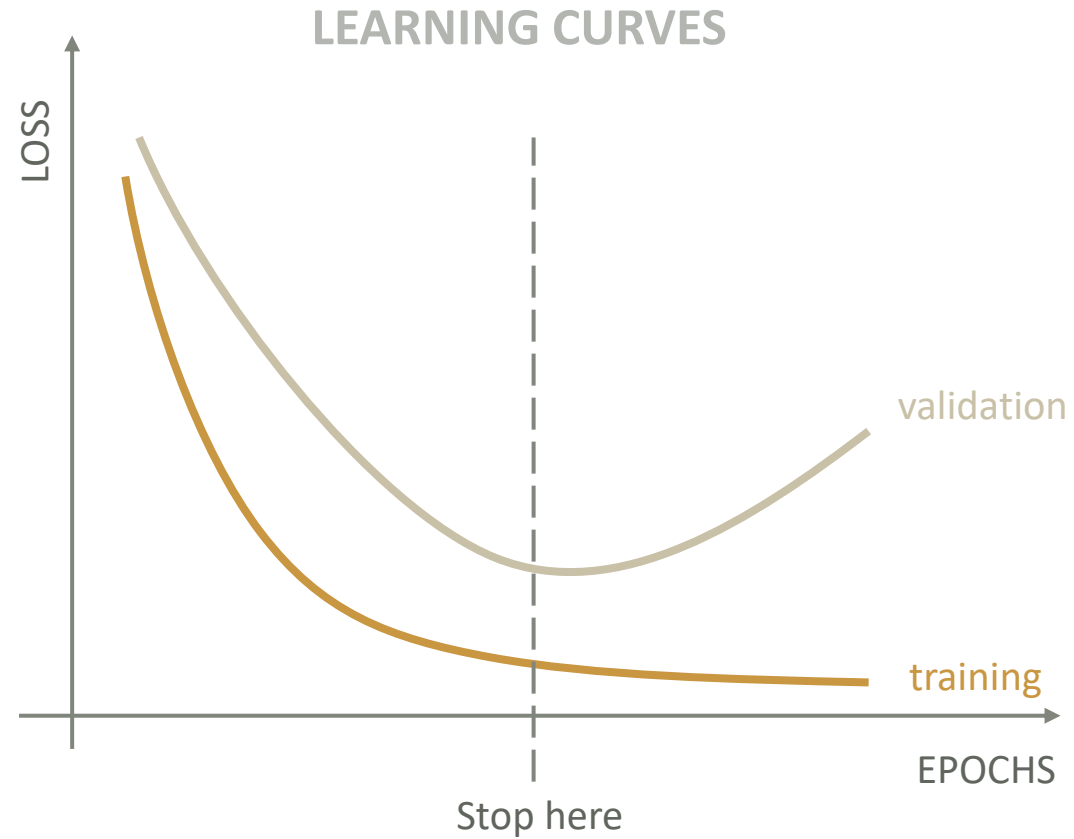
EVALUATION OF NEURAL NETWORKS



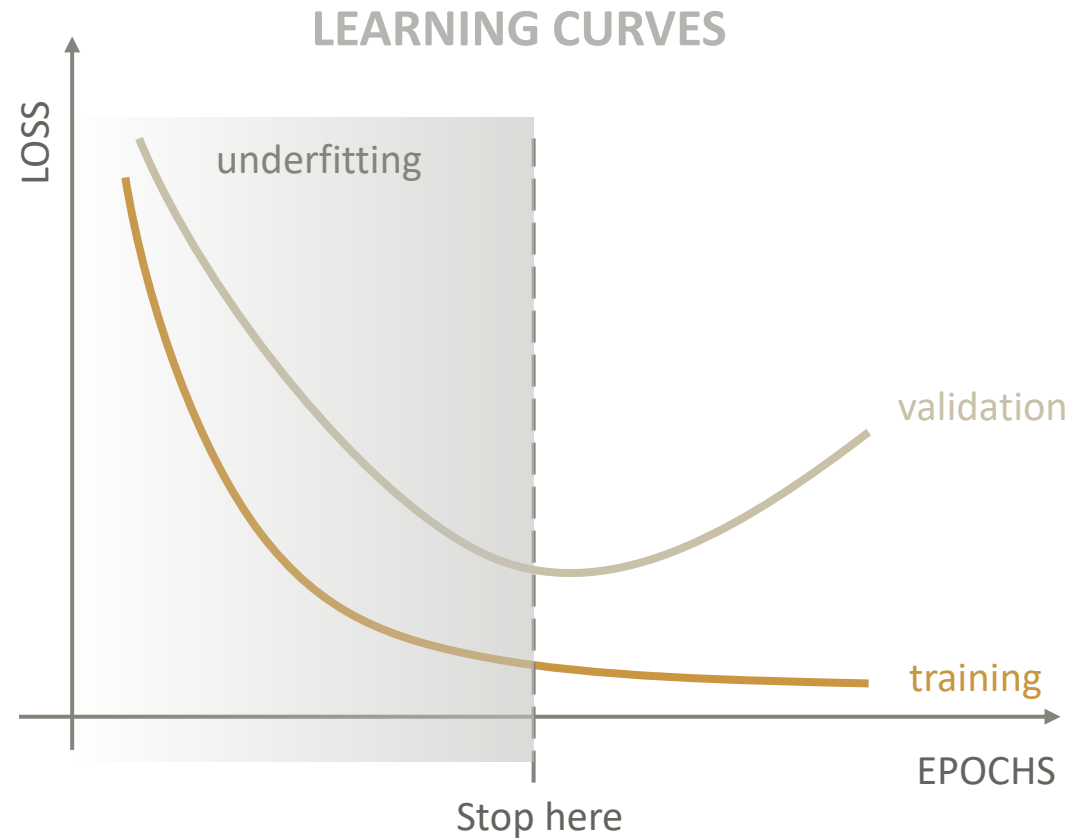
EVALUATION OF NEURAL NETWORKS



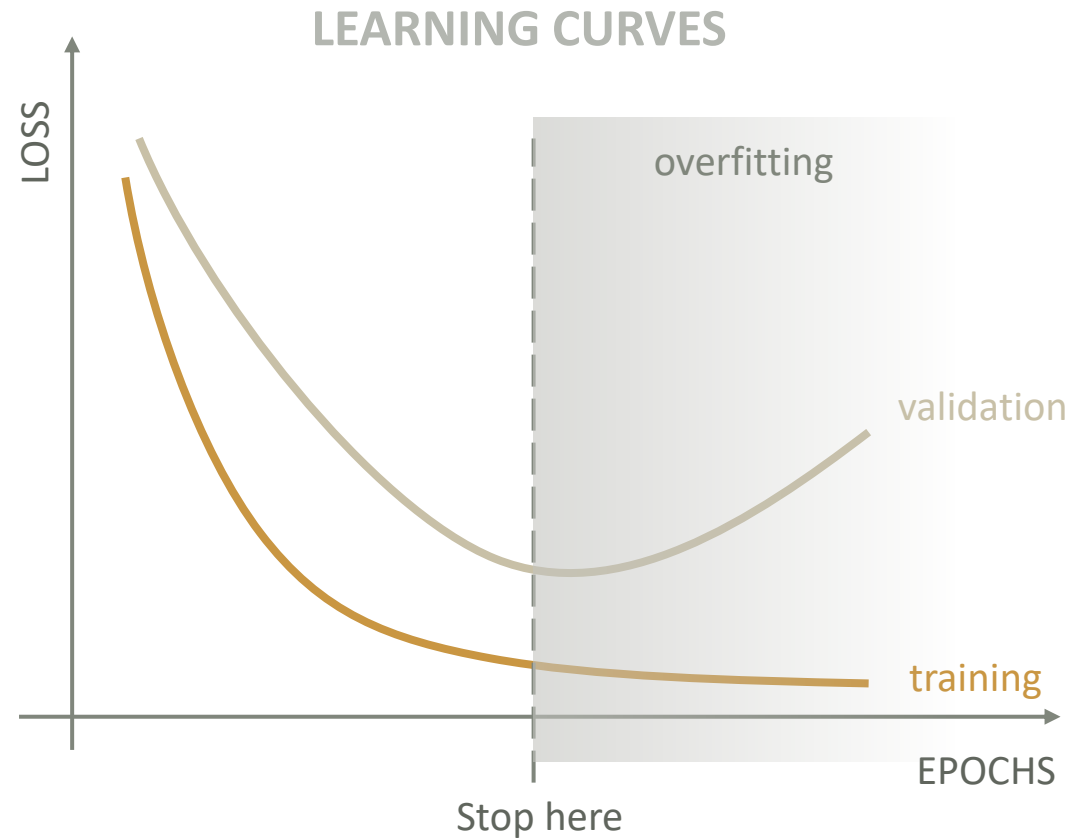
EVALUATION OF NEURAL NETWORKS



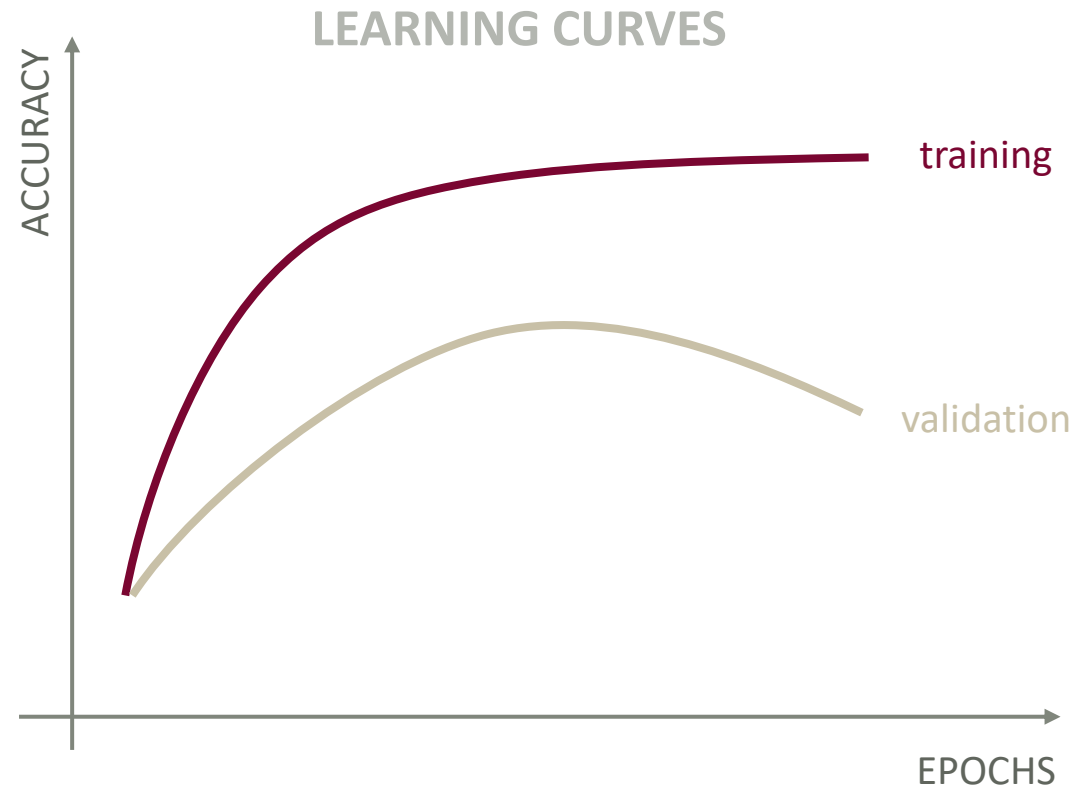
EVALUATION OF NEURAL NETWORKS



EVALUATION OF NEURAL NETWORKS



EVALUATION OF NEURAL NETWORKS



SUMMARY OF TODAY

ABOUT DATA, CATS & DOGS

- **DATA** can be represented as Vectors, Matrix, Tensor
- **BIAS** comes from data and can have a huge influence and the model's performance
 - There exist multiple forms of bias e.g., Group Attribute Bias, Overgeneralization Bias or Selection Bias,
- **A CNN's** can extract features from images using kernels
- **CONVOLUTION AND POOLING** are important parts of a CNN
- **NN's** are sensible to a lot of hyperparameters, but can be powerful