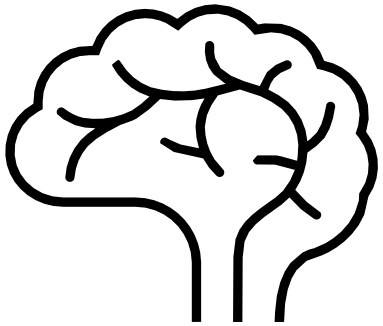
Introduction to Machine Learning

Artificial Neural Networks (ANNs)

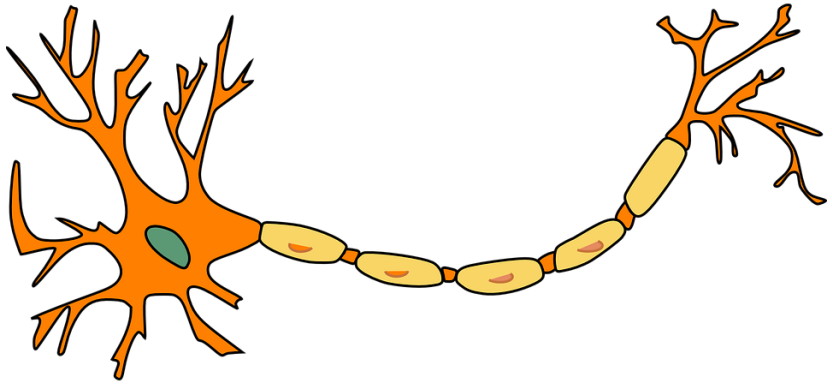
Artifical Neural Networks (ANNs)

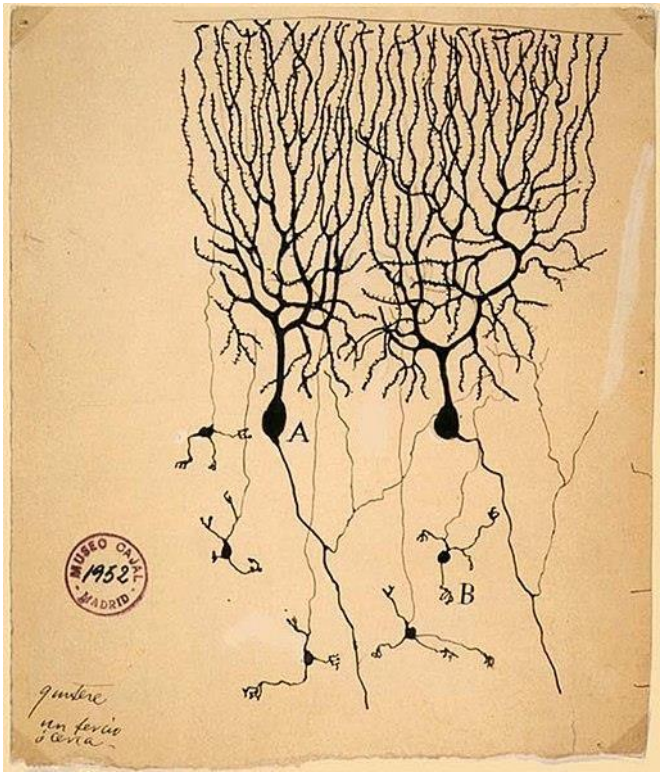


The human brain is capable of solving a wide variety of highly complex tasks seemingly „effortlessly”.

Through neuroscience, we have a fundamental understanding of the structure and functioning of the brain.

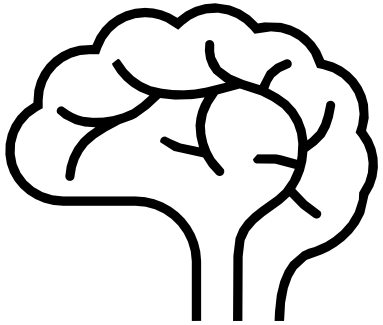
Idea:
We design an algorithm modeled after nature, called an Artificial Neural Network.





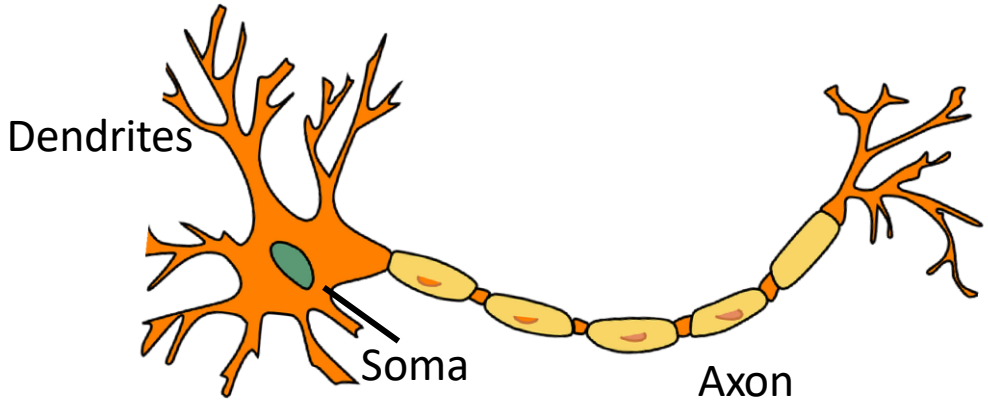
Santiago Ramón y Cajal, 1899

Artifical Neural Networks (ANNs)

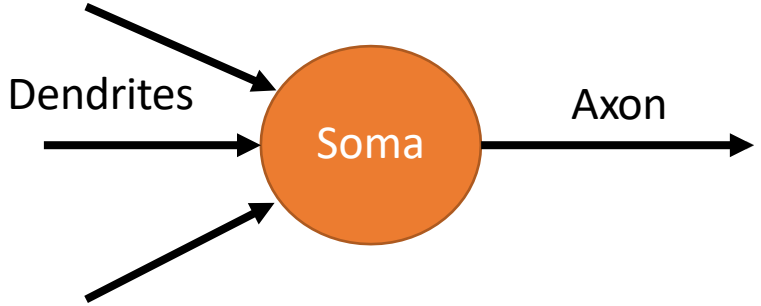


1. We imitate the structure of a biological nerve cell

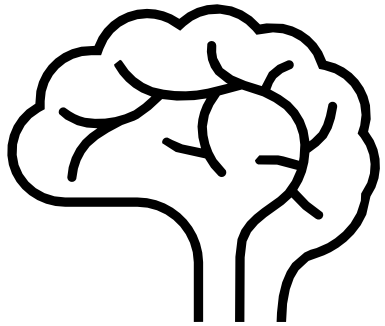
Biological nerve cell



Node in ANN

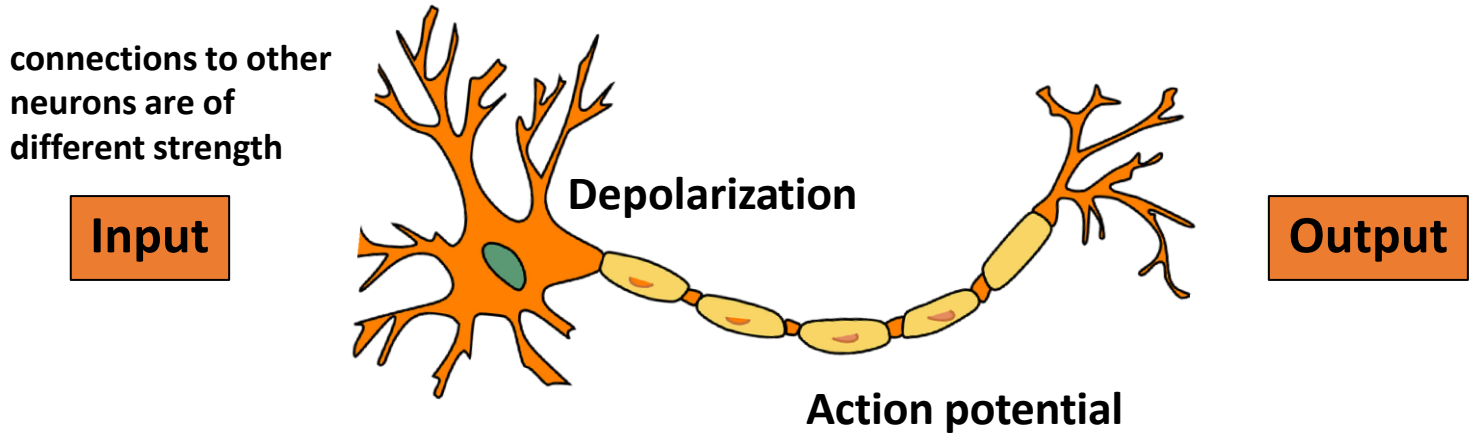


Artifical Neural Networks (ANNs)

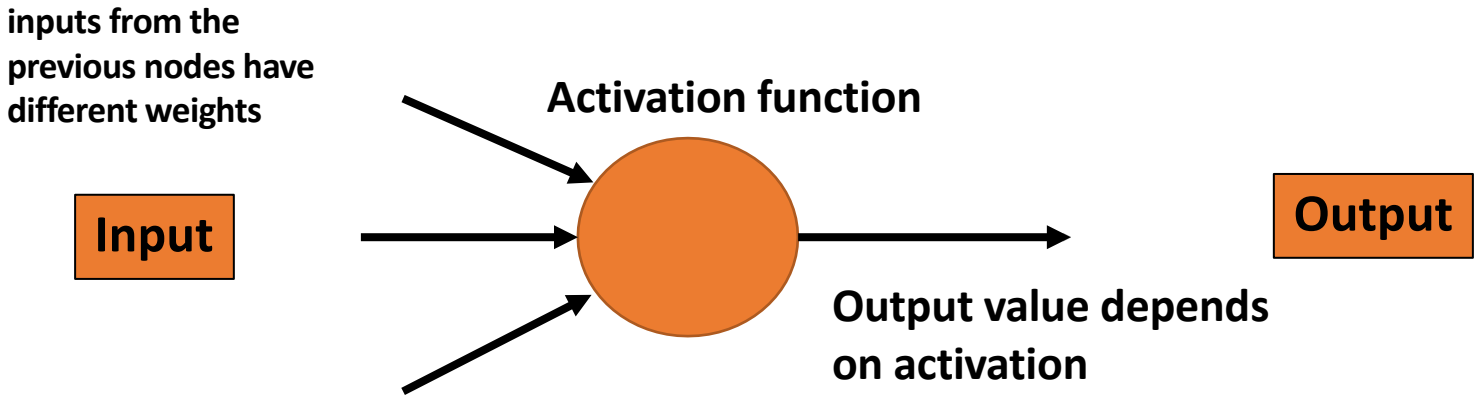


2. We imitate **the stimulus processing** of a biological nerve cell

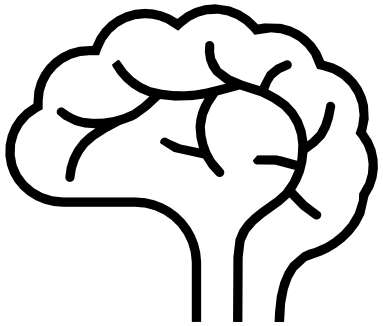
Biological nerve cell



Node in ANN

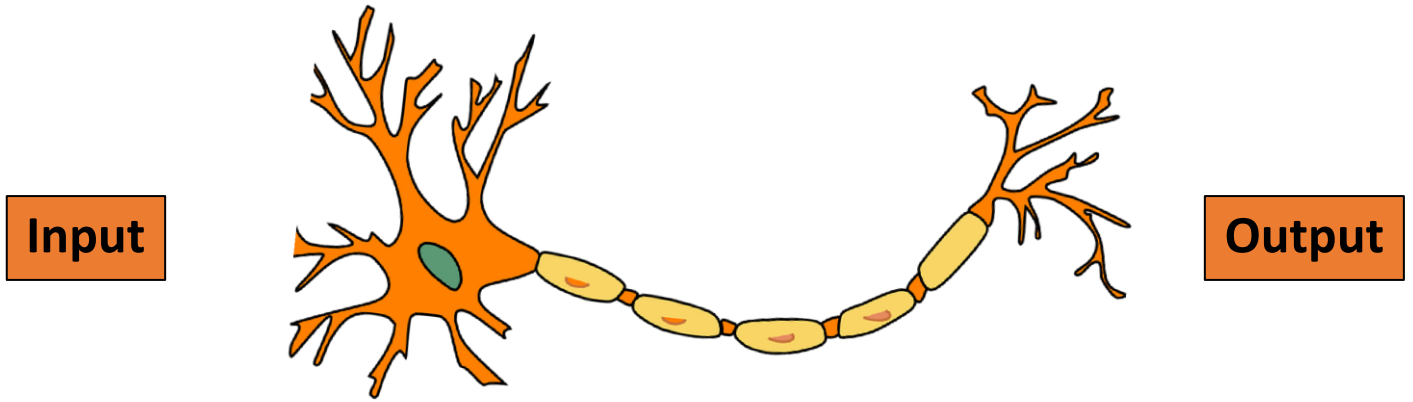


Artifical Neural Networks (ANNs)

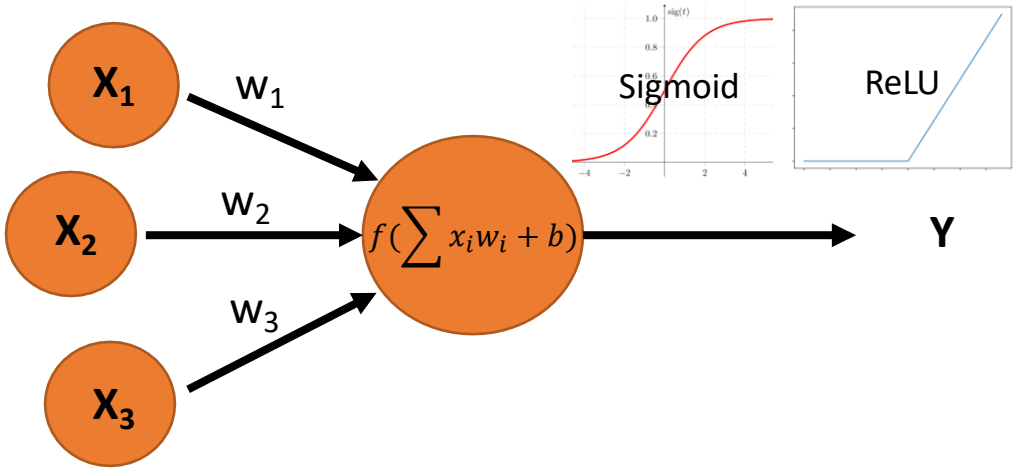


2. We imitate **the stimulus processing** of a biological nerve cell

Biological nerve cell

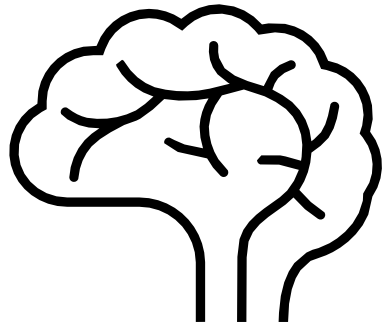


Node in ANN



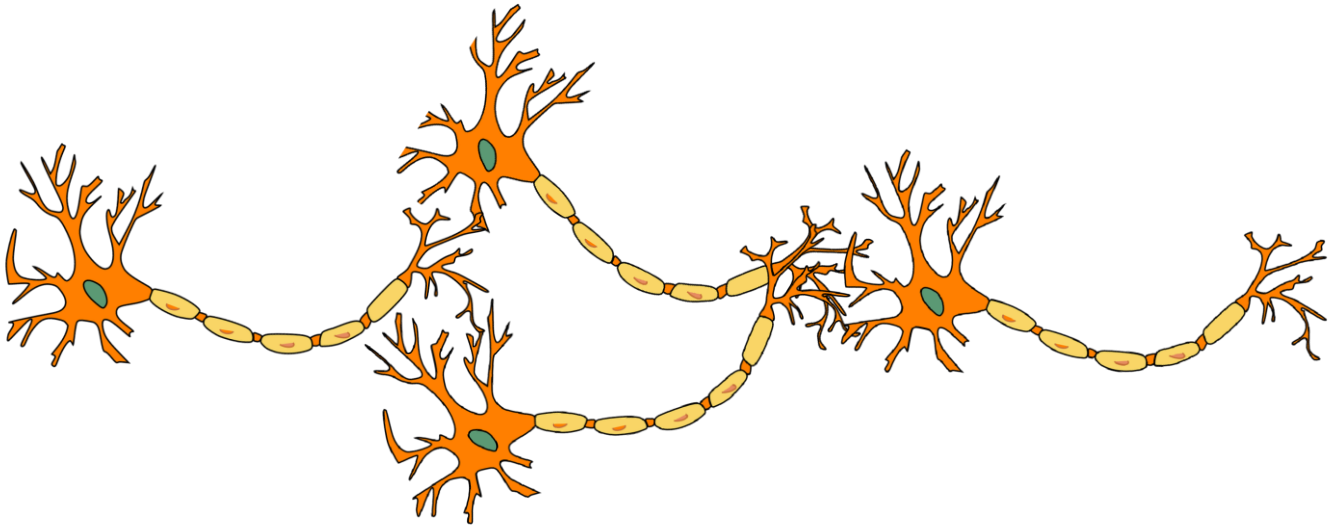
x = Inputs
w = Weights
b = Bias
f() = activation function
Y = Output

Artifical Neural Networks (ANNs)

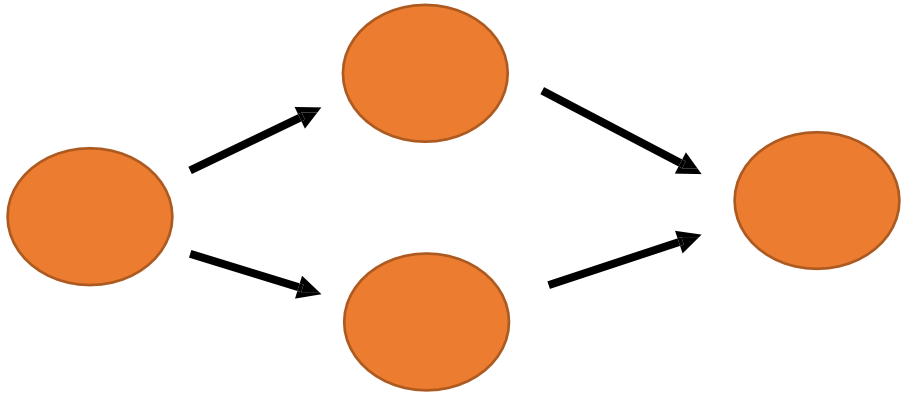


3. We imitate the structure of a neural network

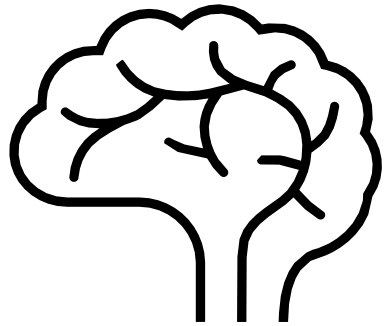
Biological nerve cell



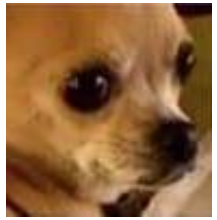
Node in ANN



Artifical Neural Networks (ANNs)

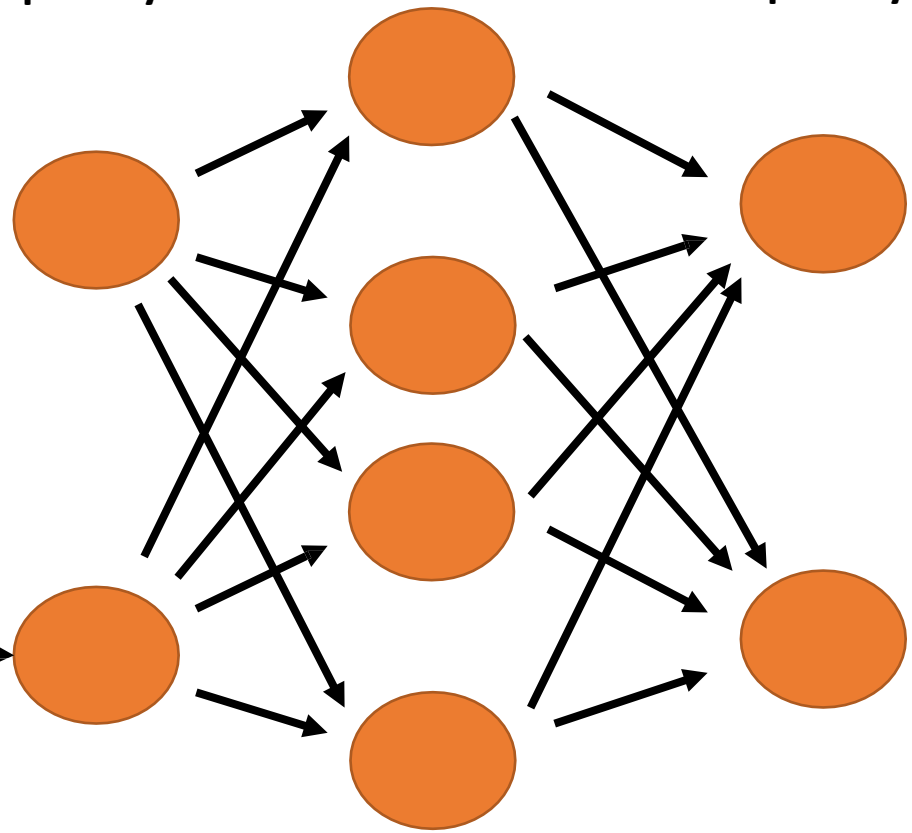


4. We train our ANN with our features and labels

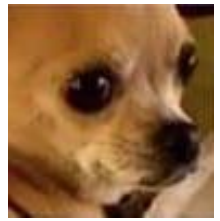


Features

HOG_Pixel_1	HOG_Pixel_2	...	Red_Bin1	Red_Bin2	...
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...



1.34



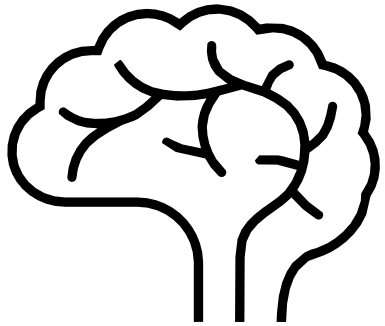
1

ArgMax

0.39

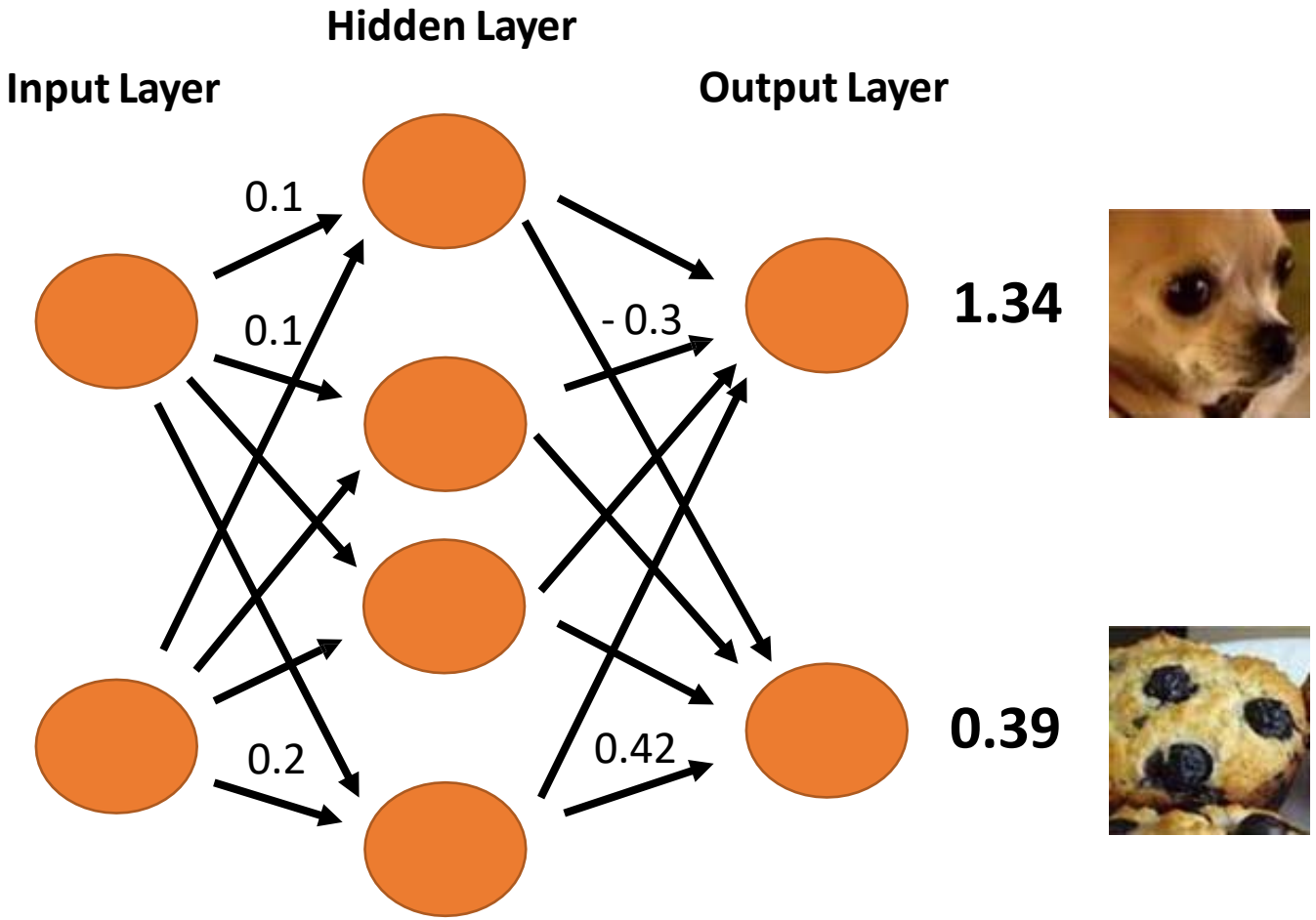
0

Artifical Neural Networks (ANNs)

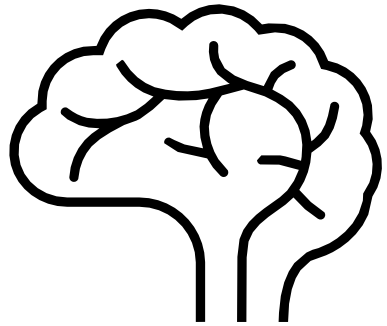


5. How does our neural network learn?

We initialize our ANN with random weights and biases



Artifical Neural Networks (ANNs)



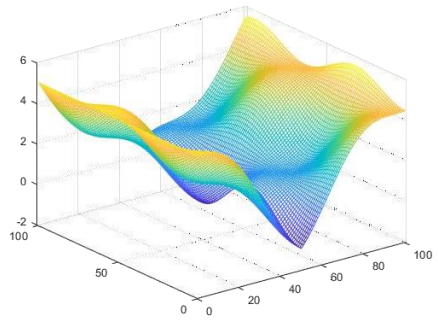
5. How does our neural network learn?

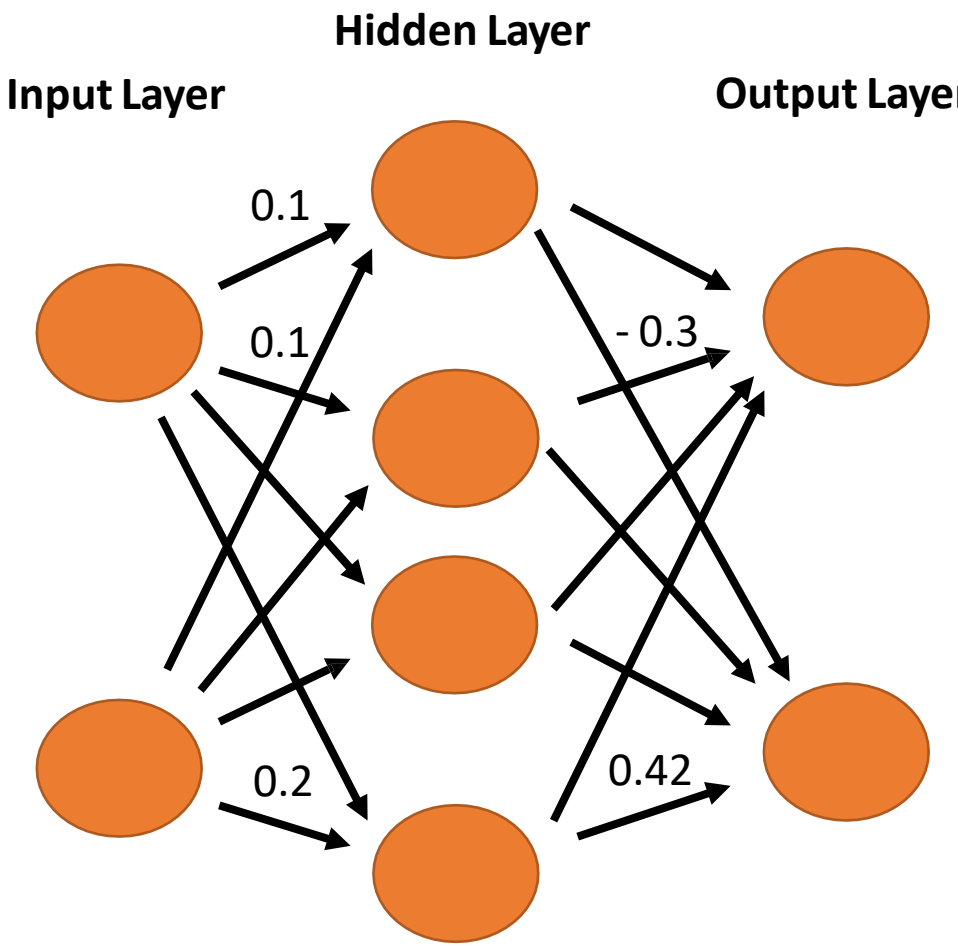
We calculate the loss function of our prediction

e.g. **Cross Entropy Loss**

$$L = -\log(\hat{y})$$

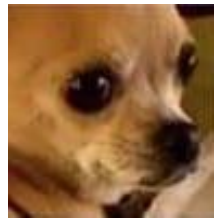
We use Gradient Descent to gradually approach the optimal weights and biases that minimize our loss function





SoftMax → training
ArgMax → test (application)

1.34



.72

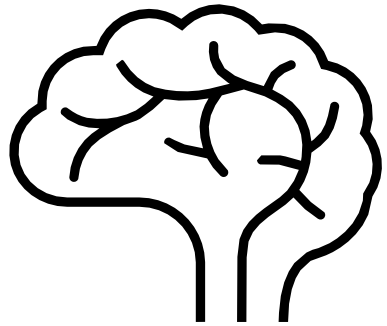
SoftMax

transforms the output values into „probabilities" for each category

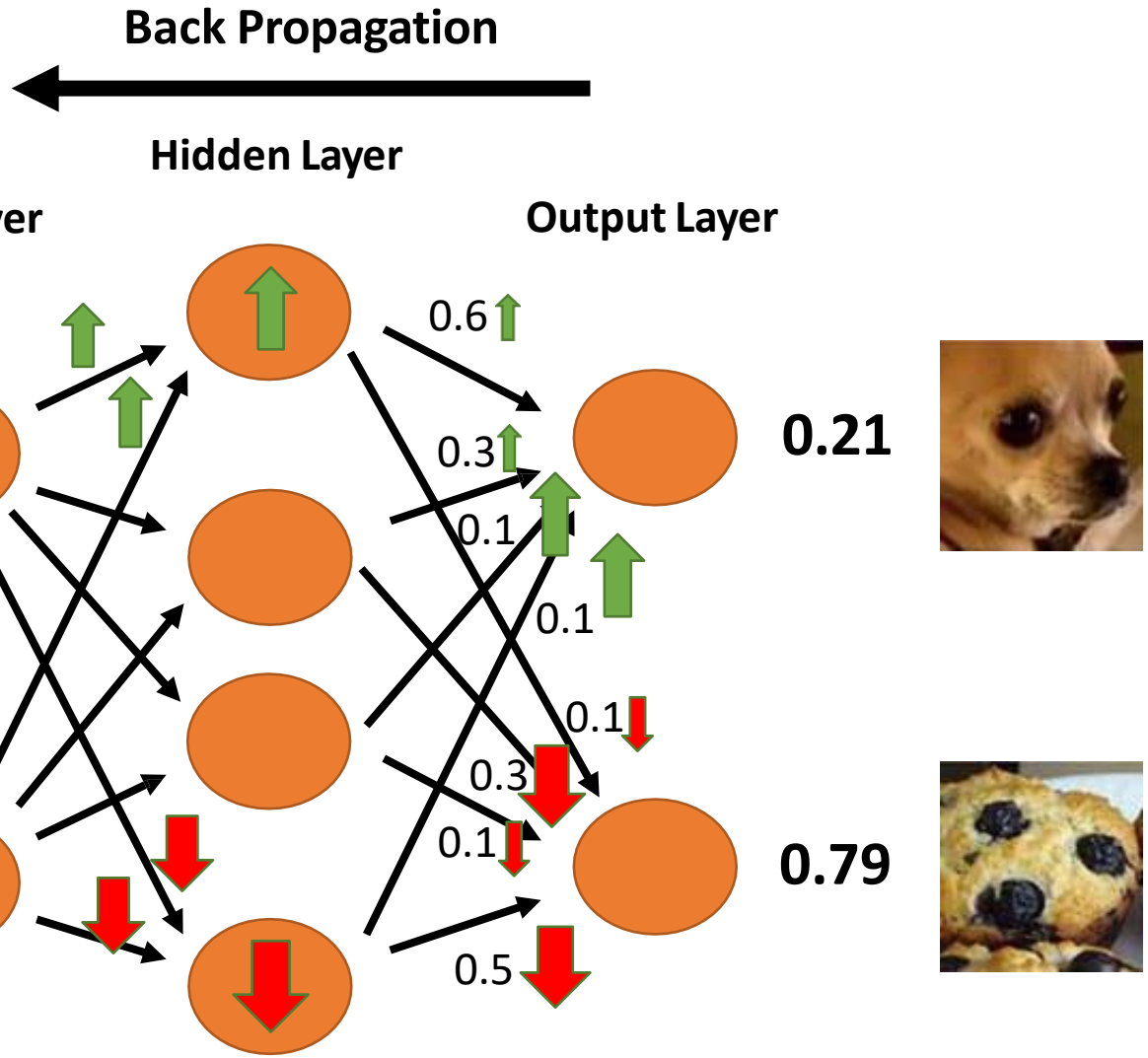
0.39

.28

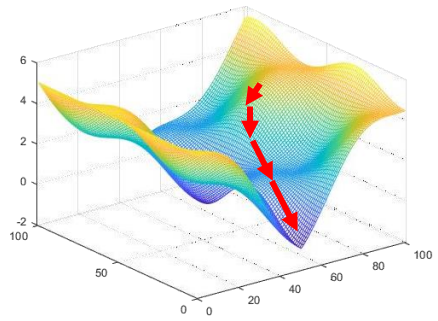
Artifical Neural Networks (ANNs)



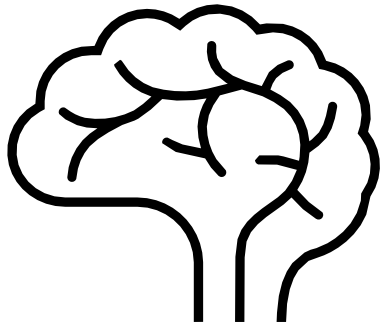
5. How does our neural network learn?



We adjust our weights and biases after each step



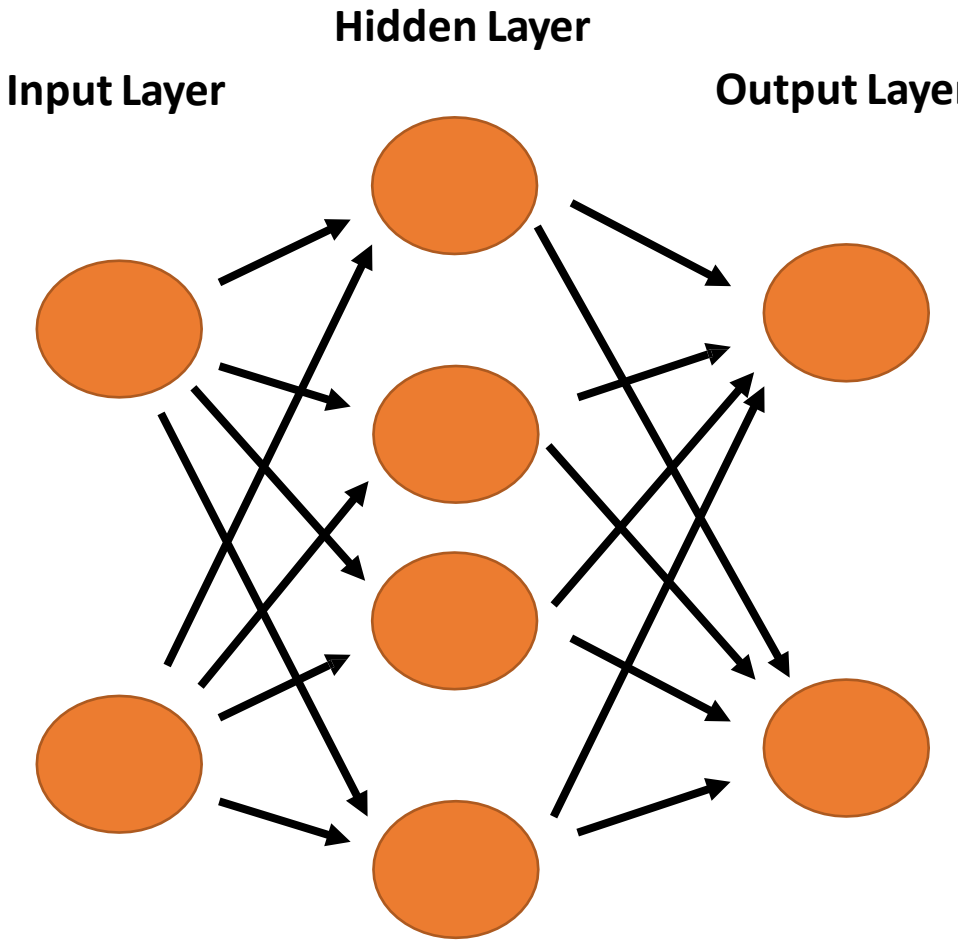
Artifical Neural Networks (ANNs)



5. How does our neural network learn?

Special feature of training the ANNs:

Usually, we do not use the complete training dataset for each training step, but only a small subset, so-called **batches**



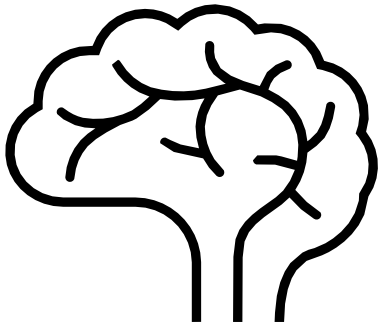
Pros:

We save resources, can train much faster and still get a good estimate of Gradient Descent!

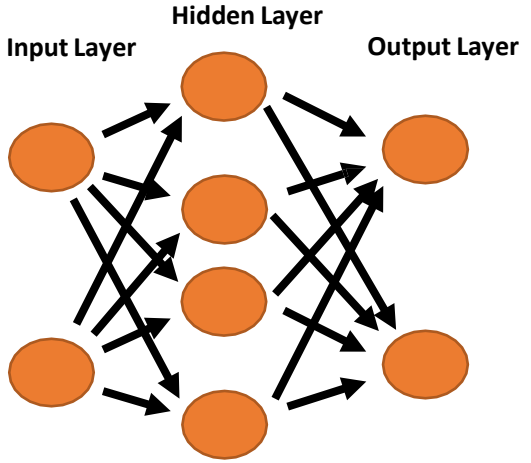
Cons:

We don’t know the optimal **batch size**!

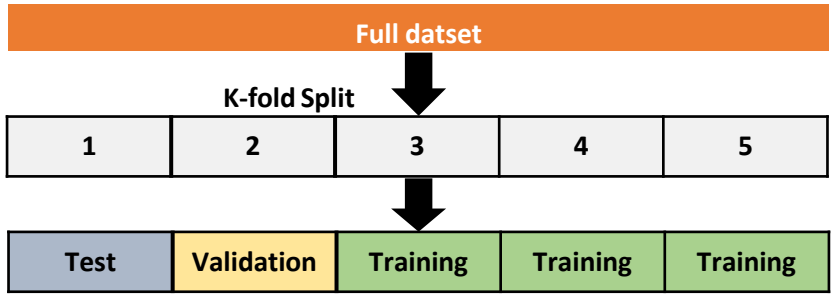
Artifical Neural Networks (ANNs)



6. We train our neural network in the same way as we trained other algorithms



Training / Test Splits



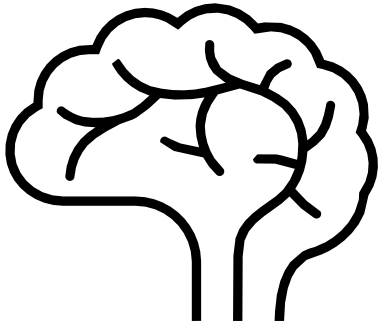
Hyperparameter Optimization

- Weight initializations
- Number of epochs
- Batch size
- Number of layers / neurons
- Activation function
- …

Regularization

- L1/L2
- Dropout
- Batch normalization
- Early stopping
- …

Representation Learning



So far we have chosen model features by ourselves....

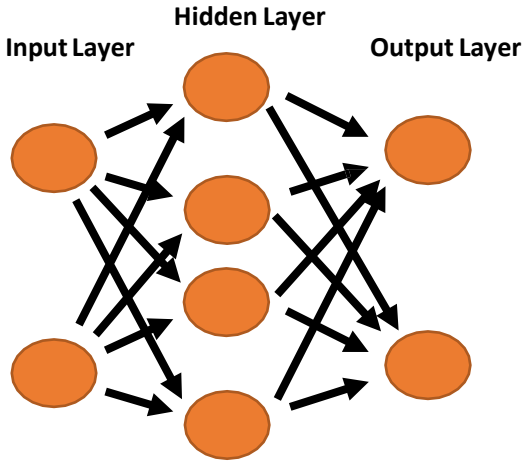
Feature Engineering

Which features, and in what form, can we use to train our algorithm?

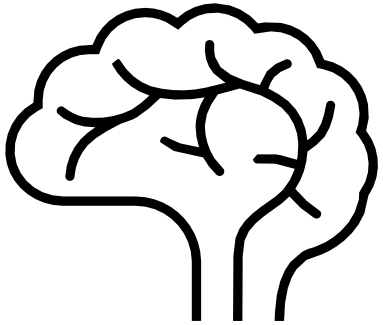
HOG_Pixel_1	HOG_Pixel_2	…	Red_Bin1	Red_Bin2	…
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…

Representation Learning

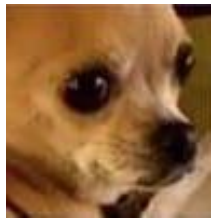
Idea: We feed our raw data into the neural network, and it decides which features to use on its own



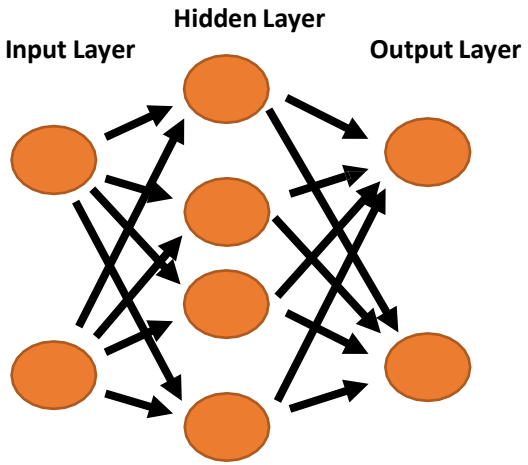
Representation Learning



… now we let the ANN itself learn which features to use from the raw data.



HOG_Pixel_1	HOG_Pixel_2	…	Red_Bin1	Red_Bin2	…
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…



We directly imput single pixels into the ANN