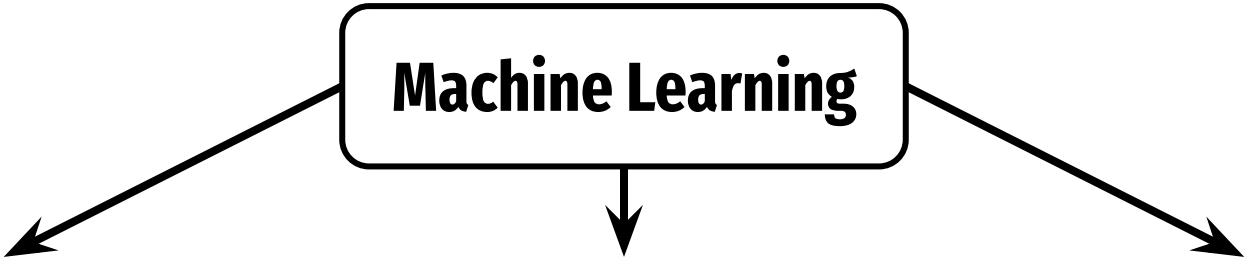
Basics of Machine Learning

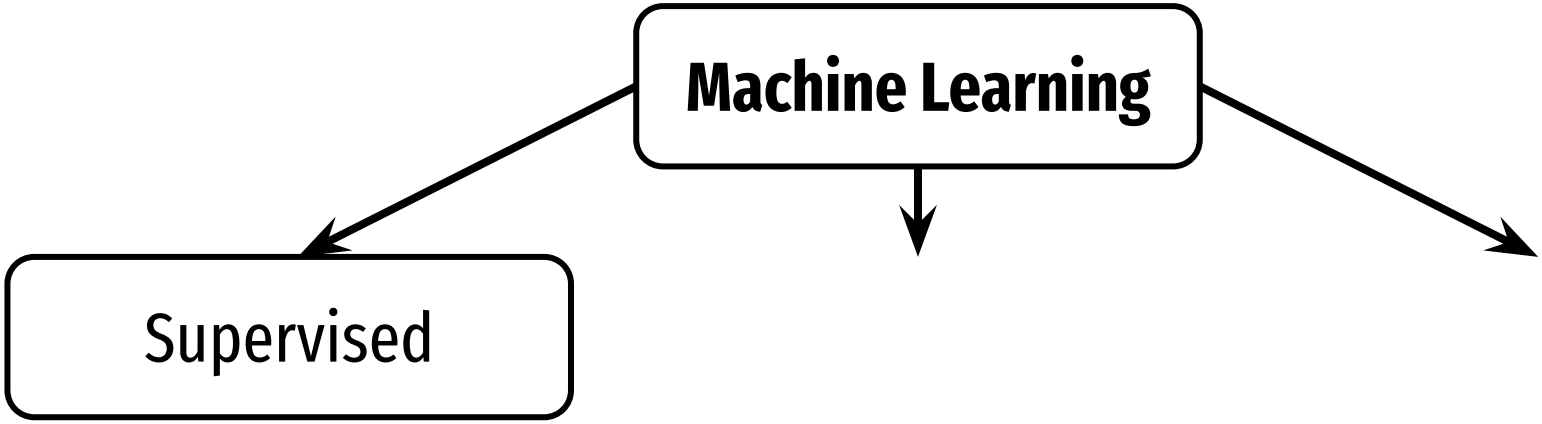
Ismaël Lajaaiti

What is Machine Learning?

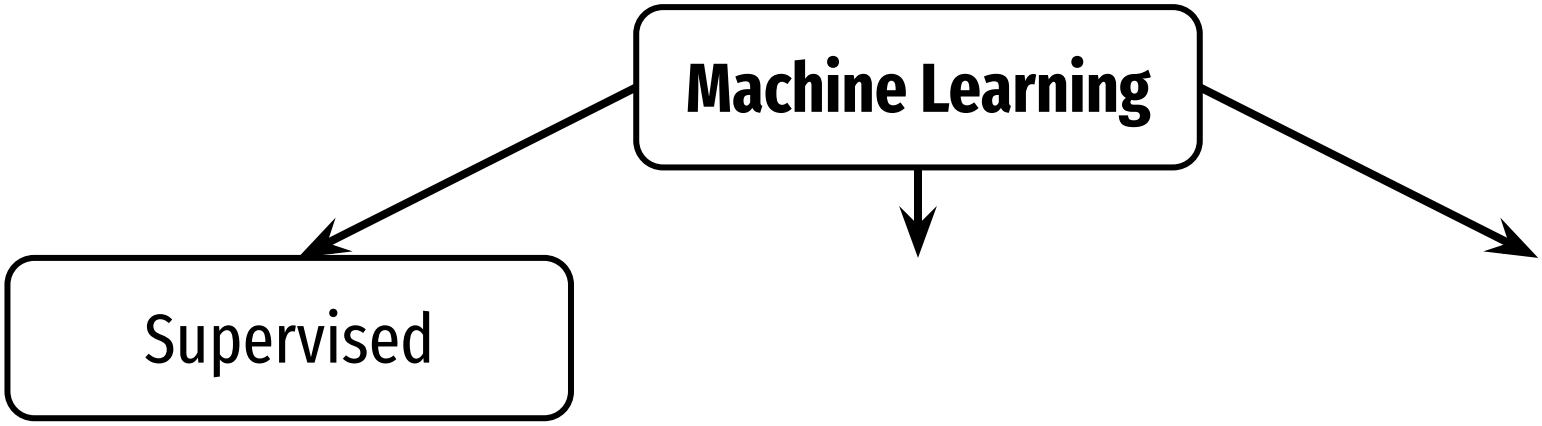
What is Machine Learning?

Algorithms that learn.


Machine Learning

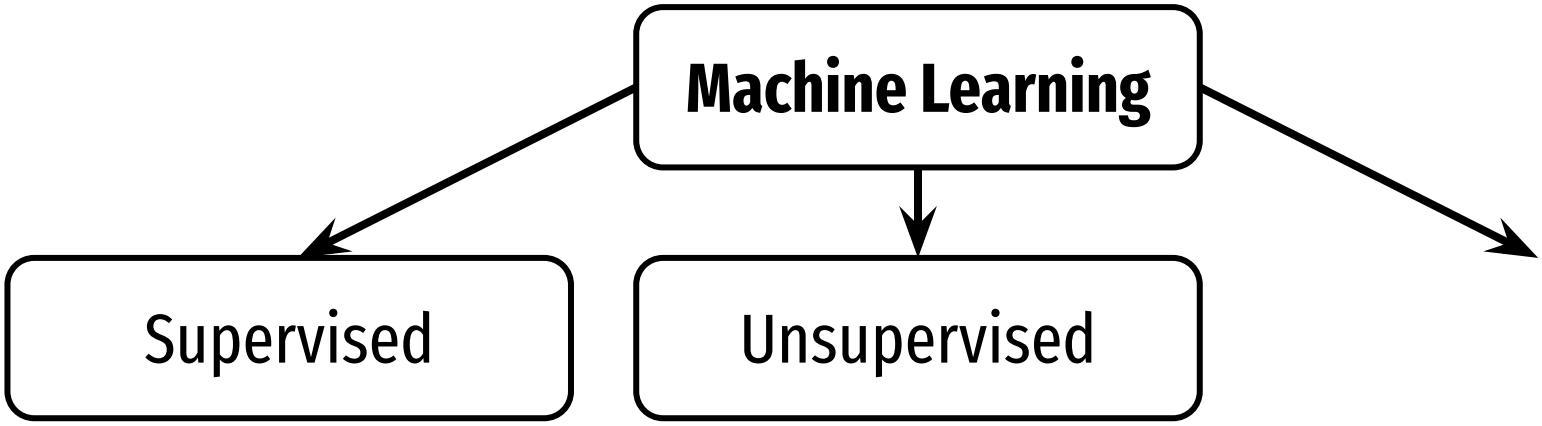


Train with labeled data



Train with labeled data

Image classification



Train with labeled data

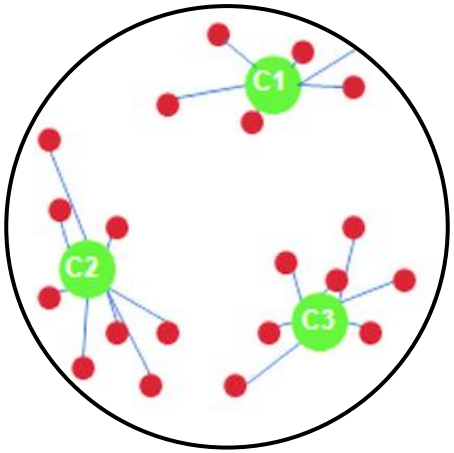
Discovers the information by itself

Image classification

Machine Learning

- **Supervised** — Train with labeled data
- **Unsupervised** — Discovers the information by itself

Image classification



K-means

Machine Learning

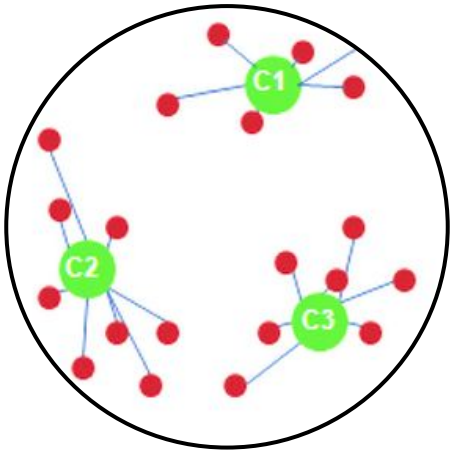
Supervised

Train with labeled data

Image classification

Unsupervised

Discovers the information by itself



K-means

Reinforcement

Simulate game-like situations, get reward and penalties

Machine Learning

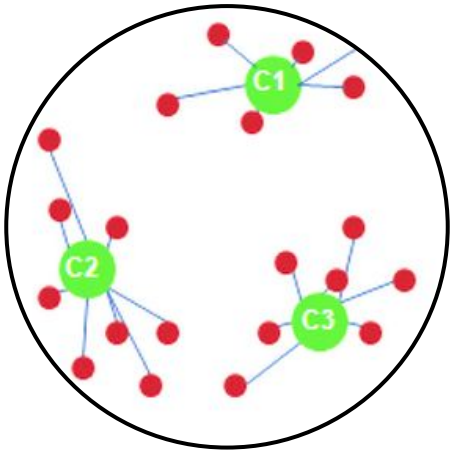
Supervised

Train with labeled data

Image classification

Unsupervised

Discovers the information by itself



K-means

Reinforcement

Simulate game-like situations, get reward and penalties

AlphaGo

Machine Learning

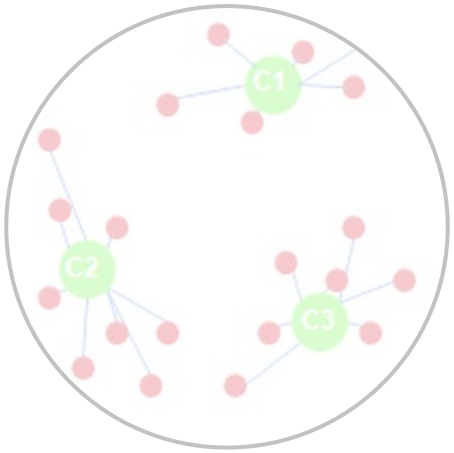
Supervised

Train with labeled data

Image classification

Unsupervised

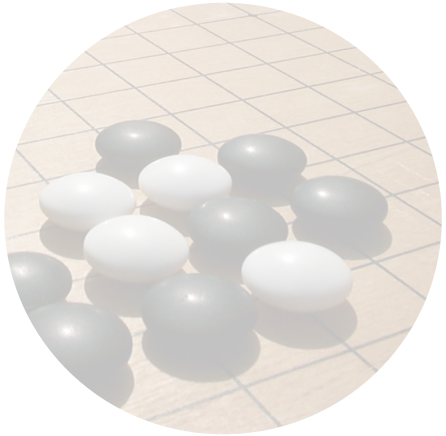
Discovers the information by itself



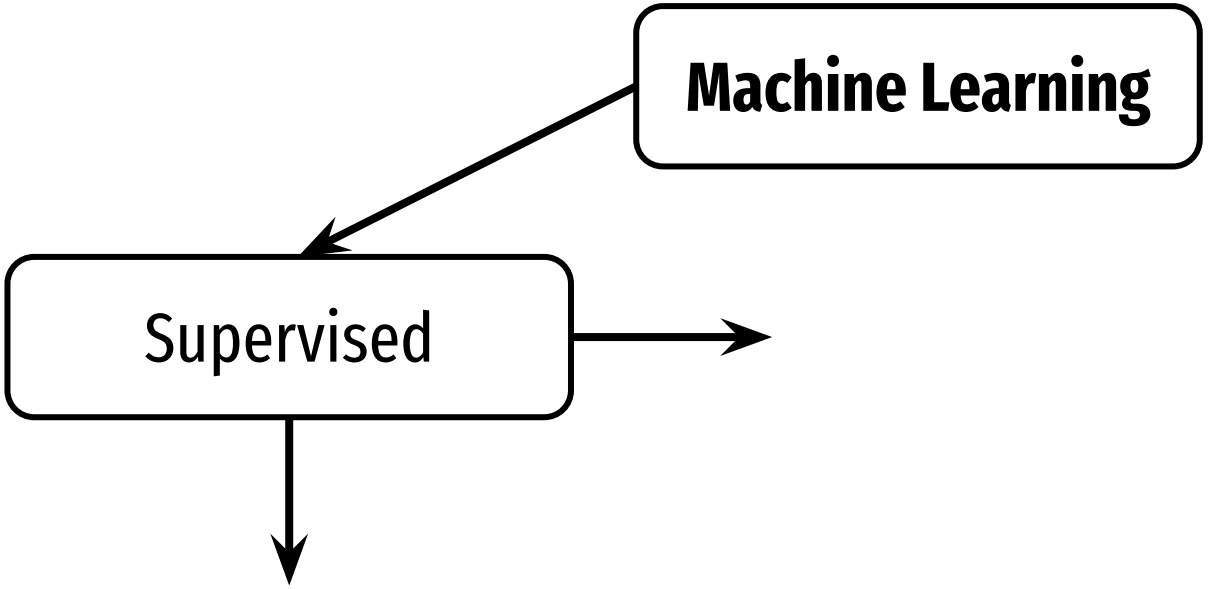
K-means

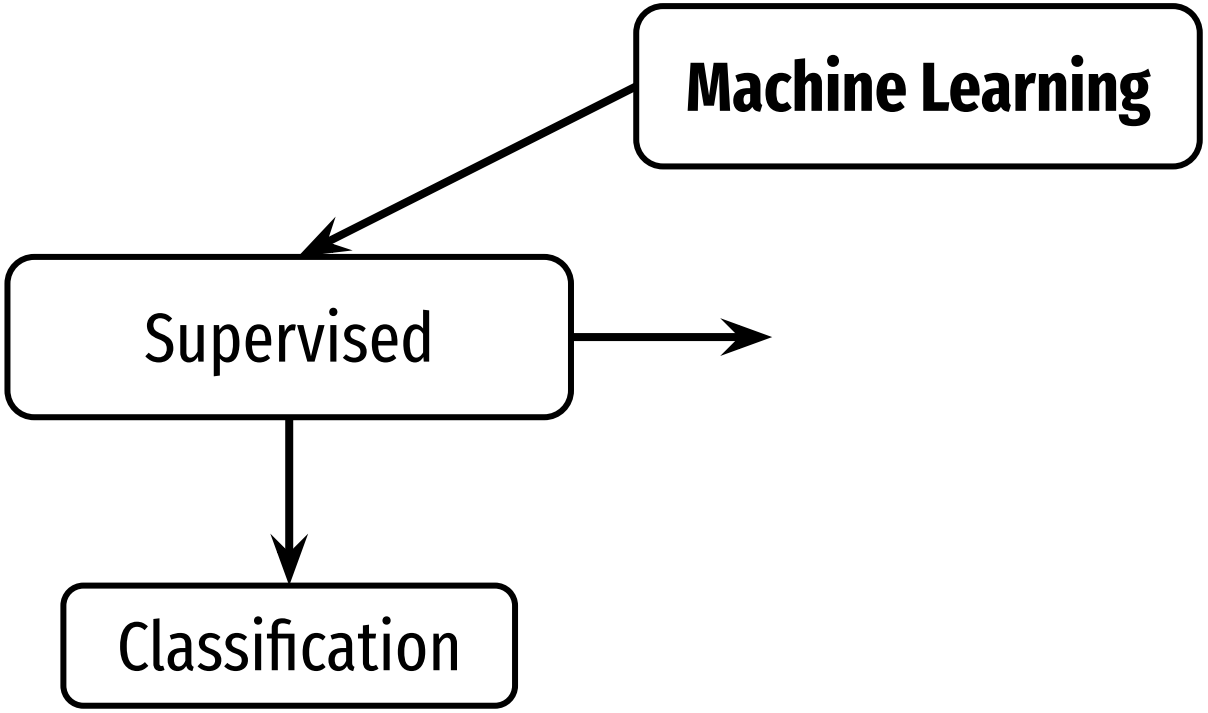
Reinforcement

Simulate game-like situations, get reward and penalties

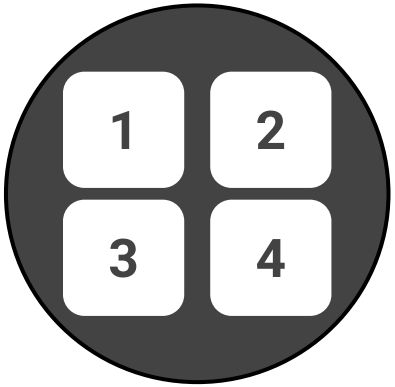


AlphaGo


Machine Learning
Supervised



Categorical variable



Machine Learning

Supervised

Classification

Categorical variable

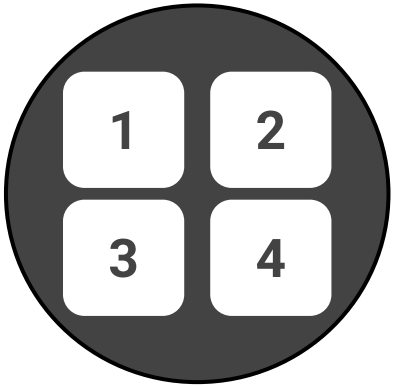
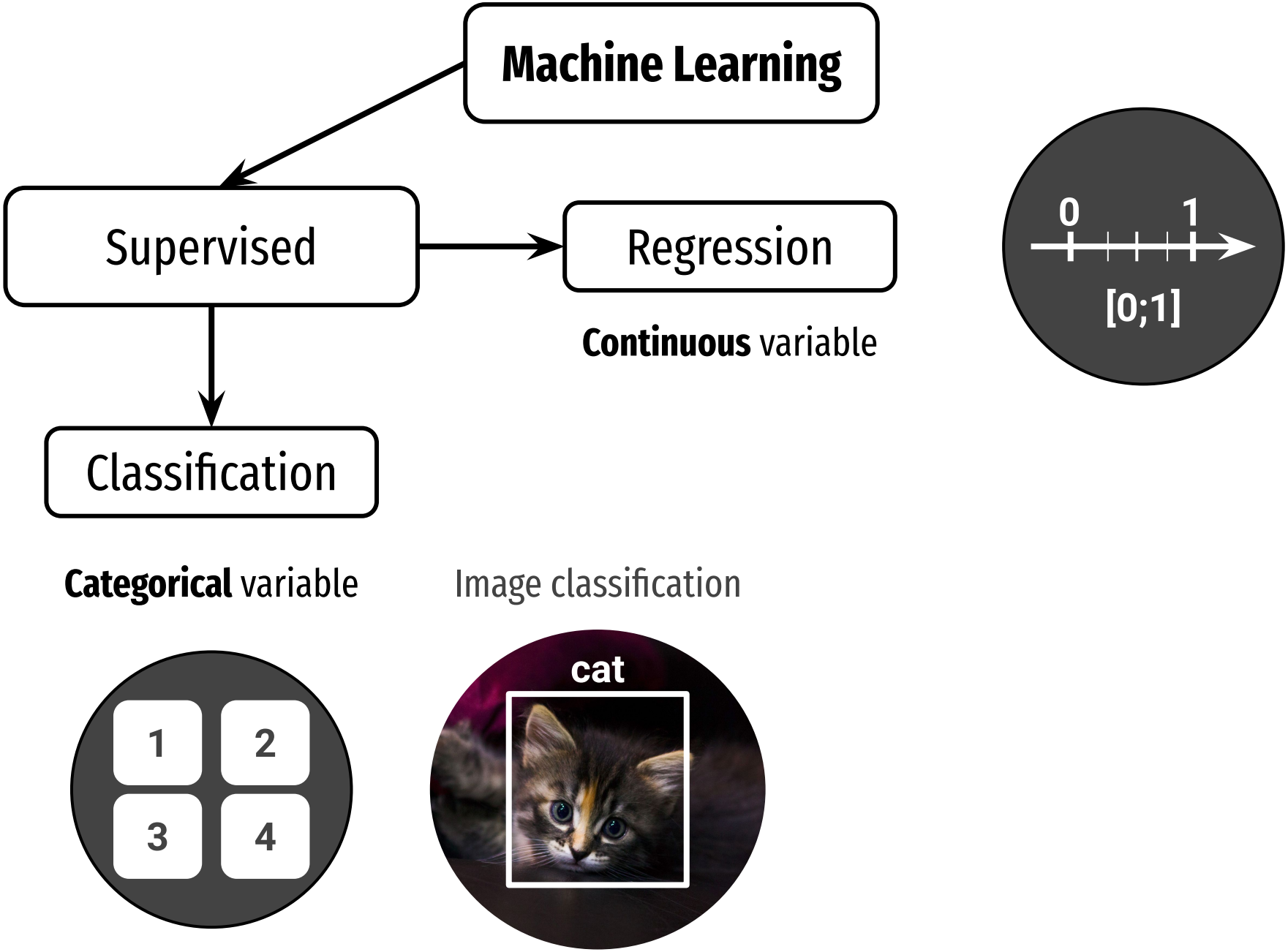


Image classification

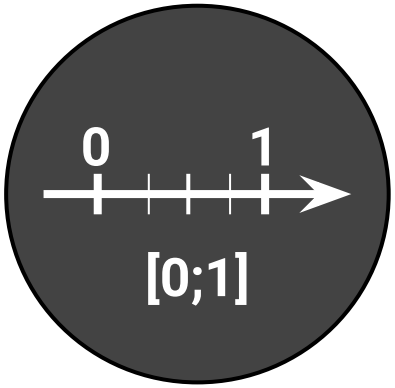

Machine Learning
Supervised
Regression
Classification
0
1
[0;1]
Continuous variable
Categorical variable
Image classification
1
2
3
4
cat

Machine Learning

Supervised

Regression

Continuous variable



Classification

Categorical variable

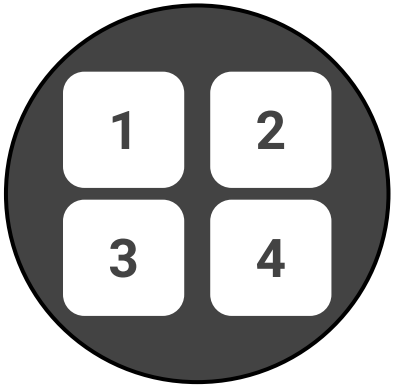
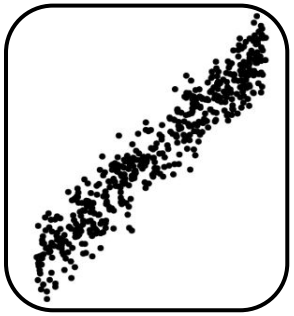


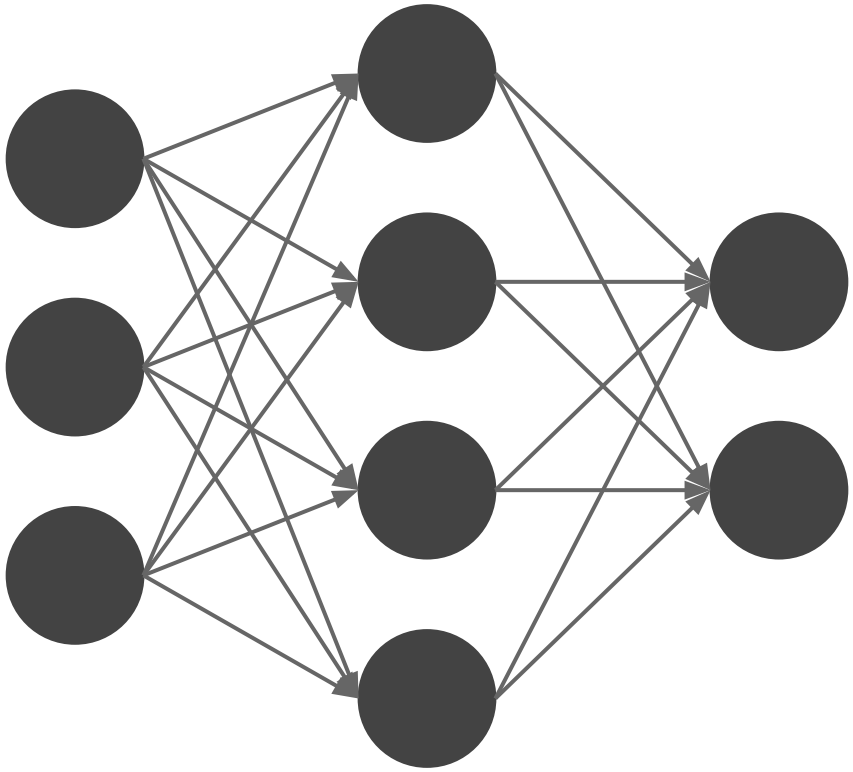
Image classification



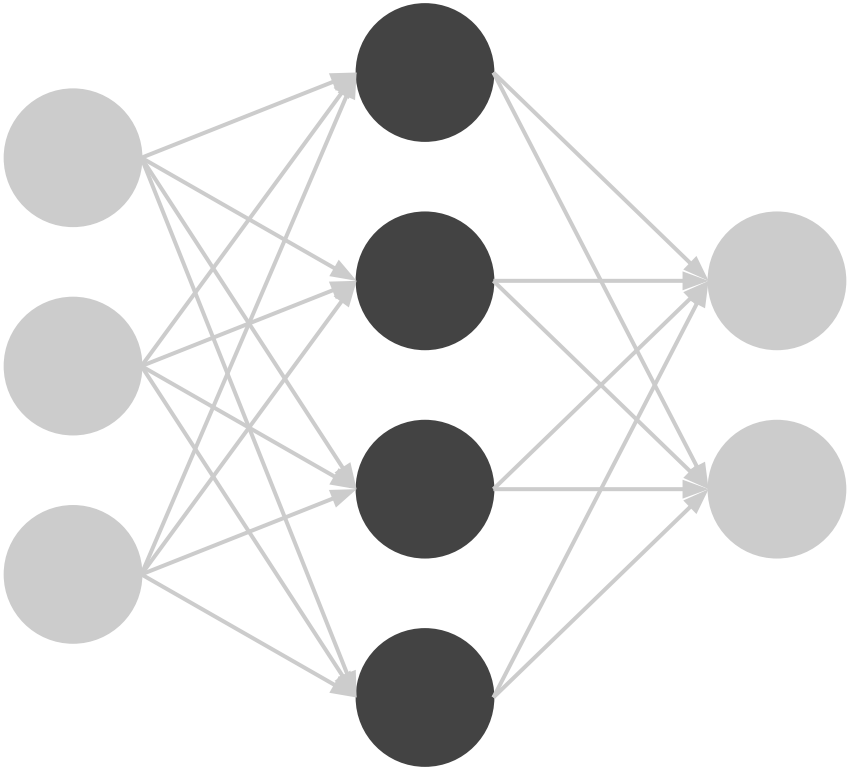
Corr = 0.95

Guess the correlation

Neural networks

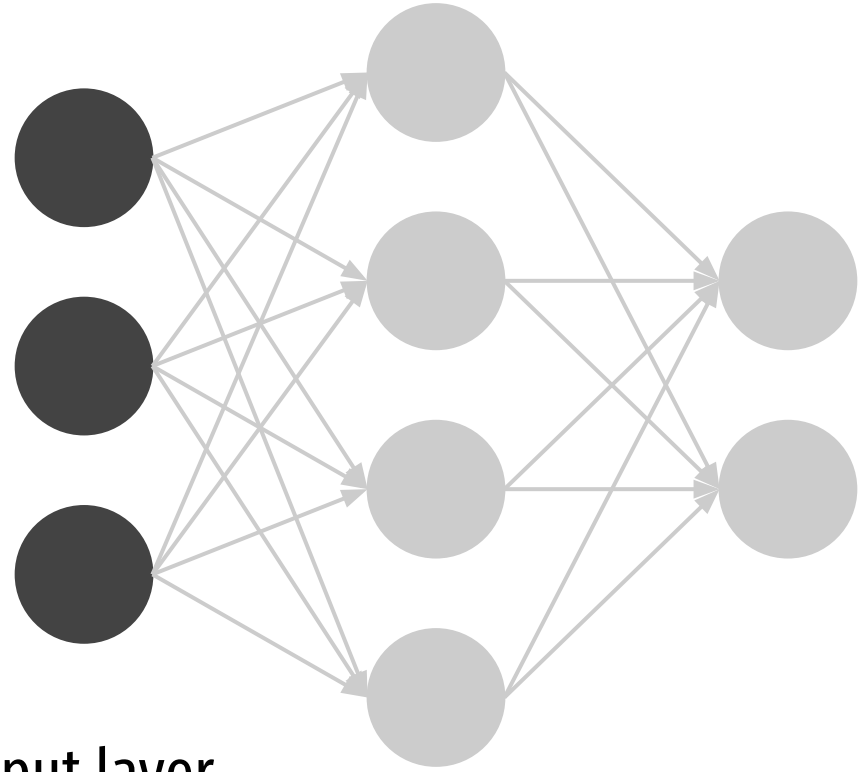


Neural networks



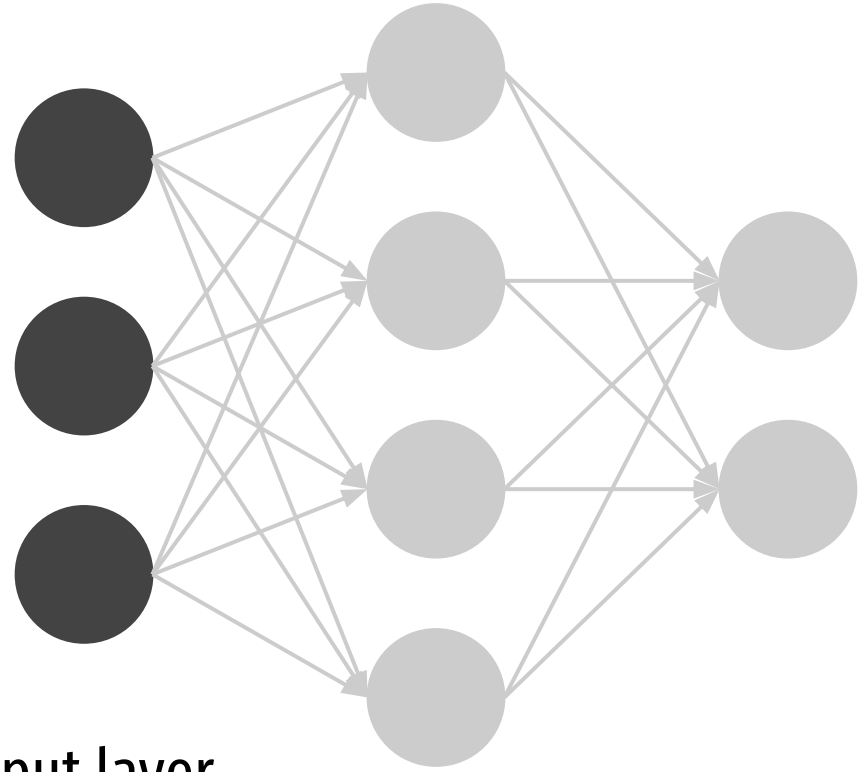
Layer

Neural networks



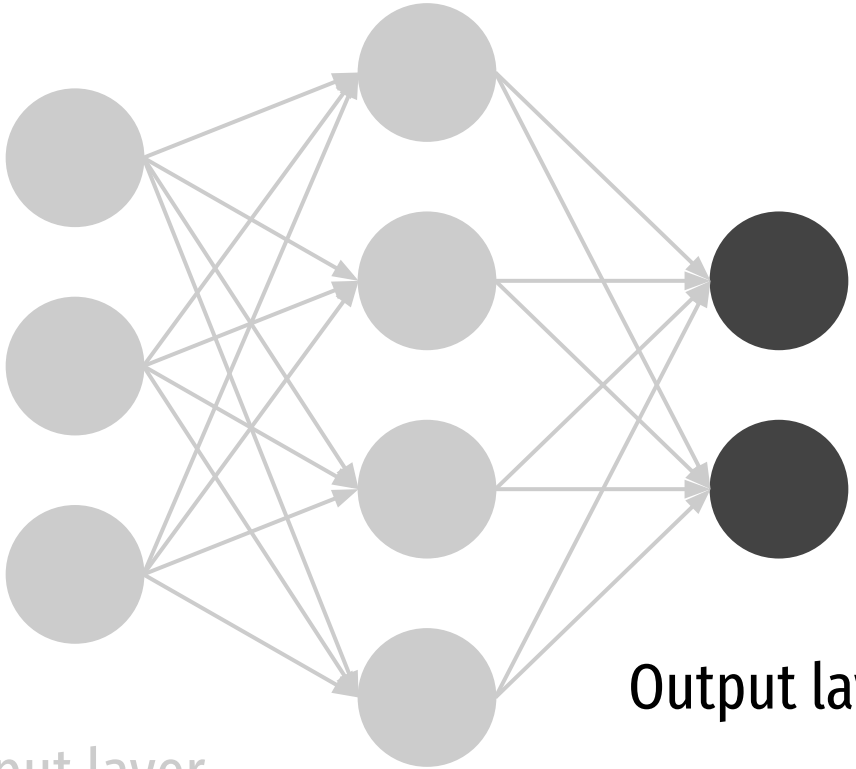
Input layer

Neural networks



Input layer

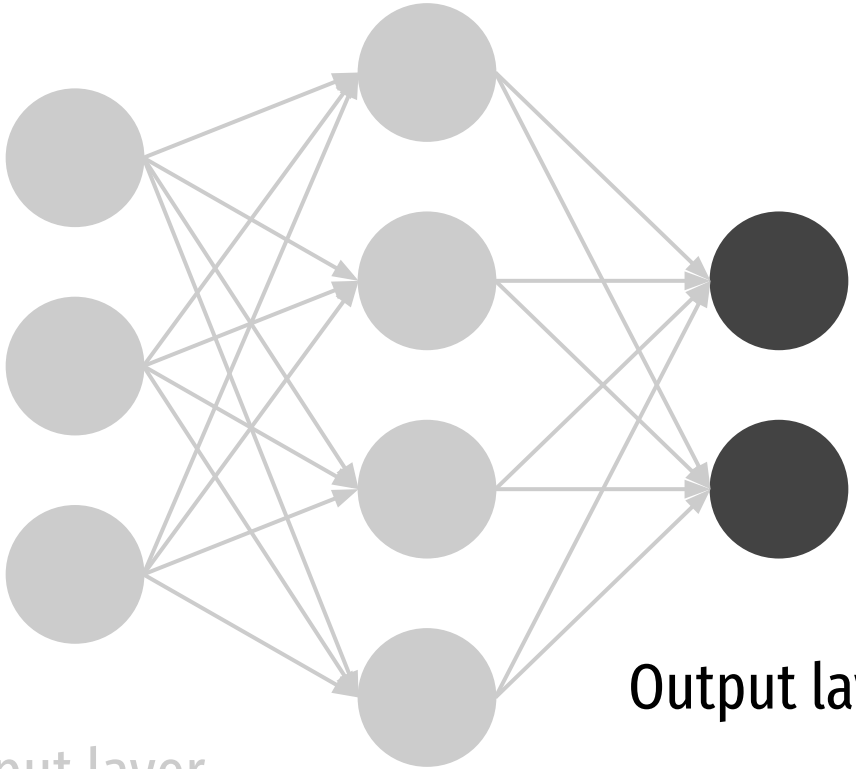
Neural networks



Input layer

Output layer

Neural networks

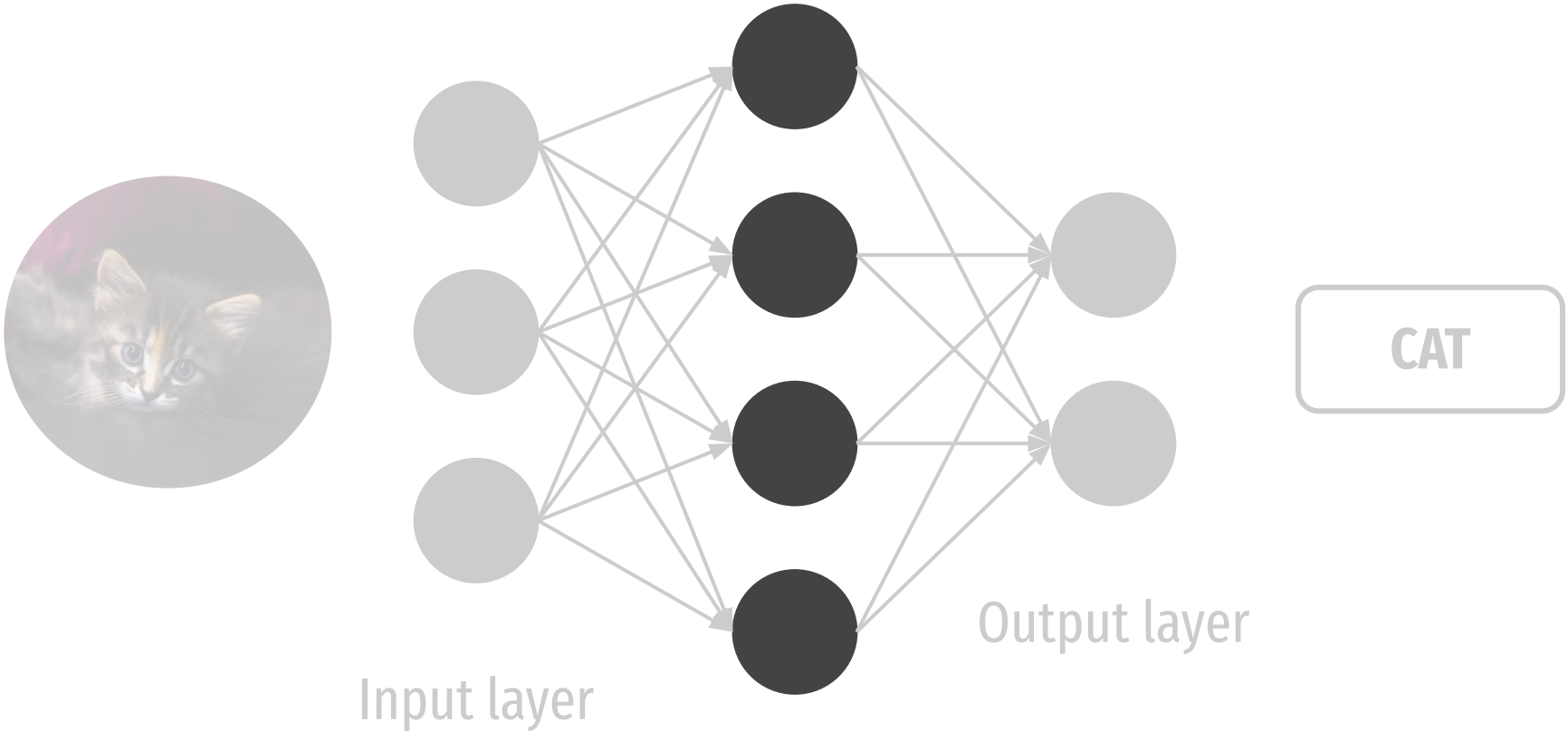


Input layer

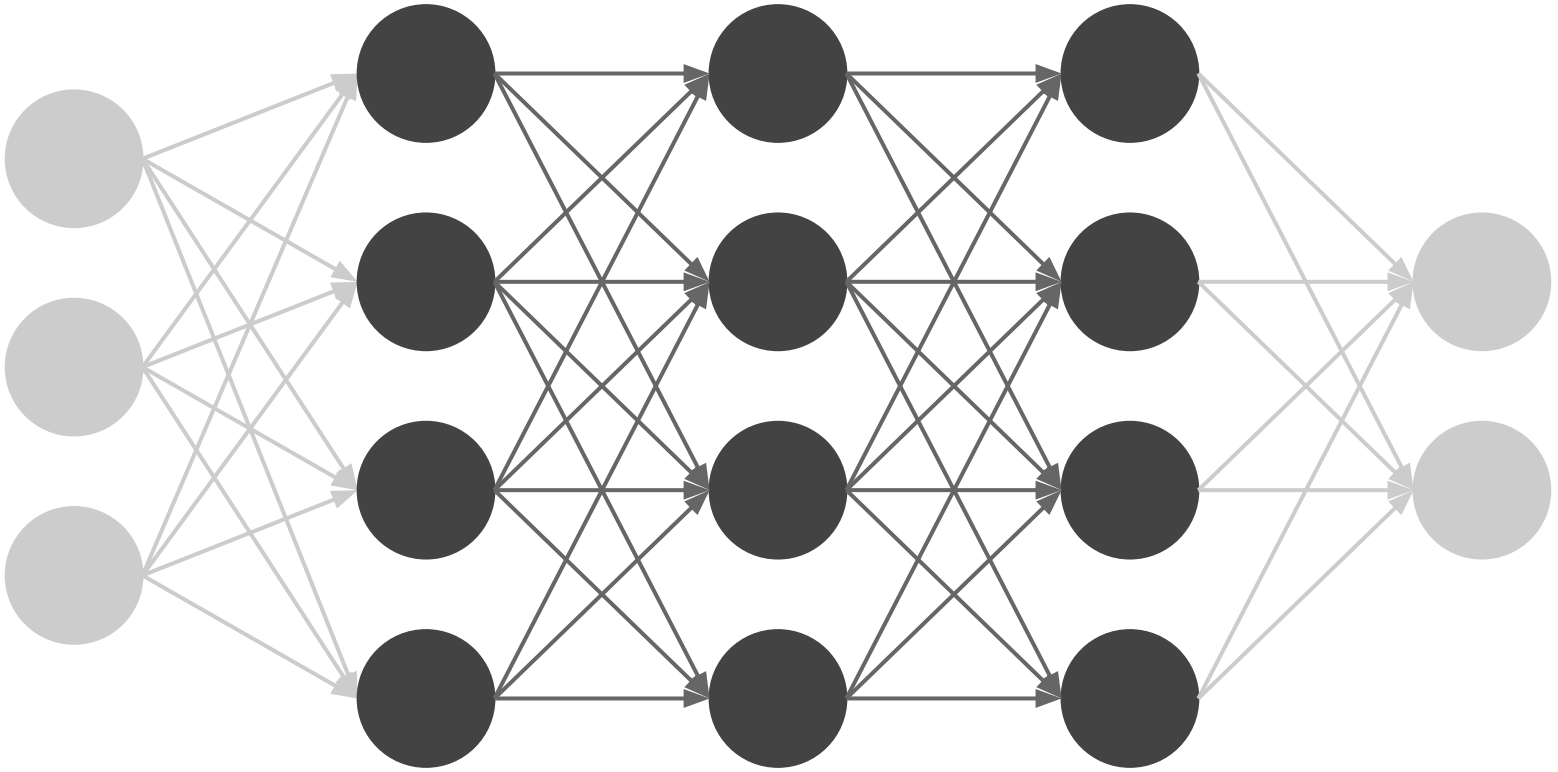
Output layer

CAT

Neural networks

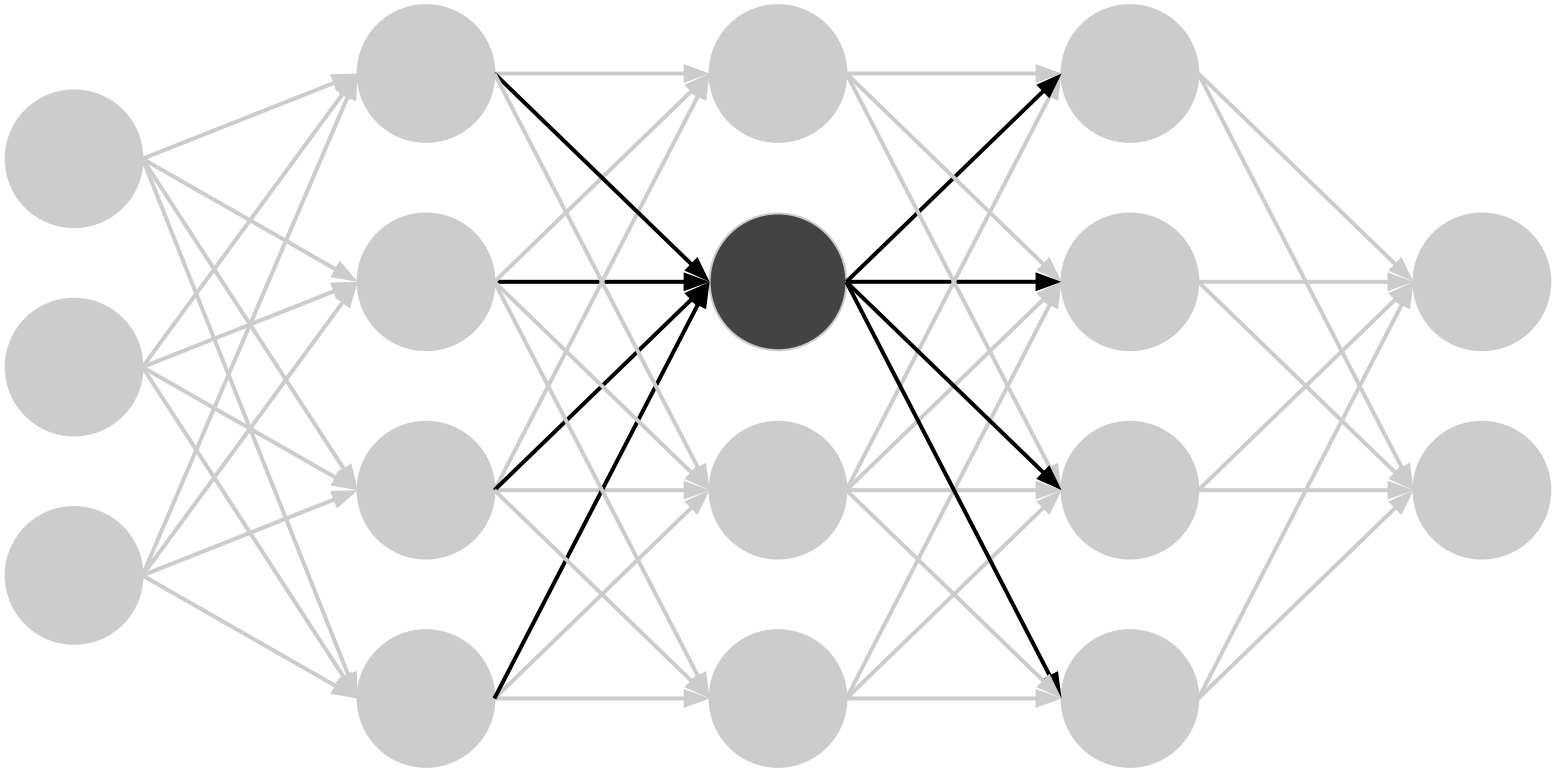


Deep neural networks



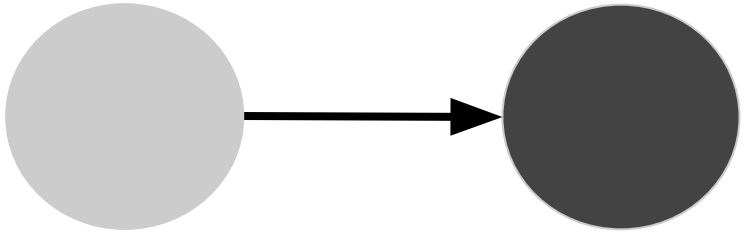
Stacked hidden layers

Deep neural networks

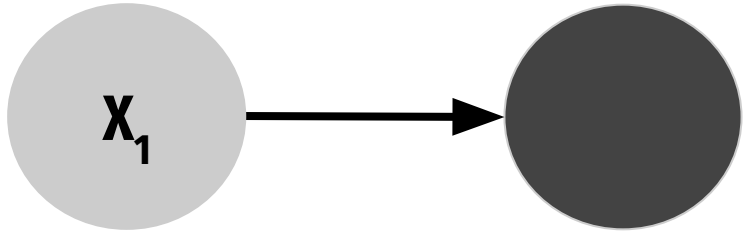


What is a neuron?

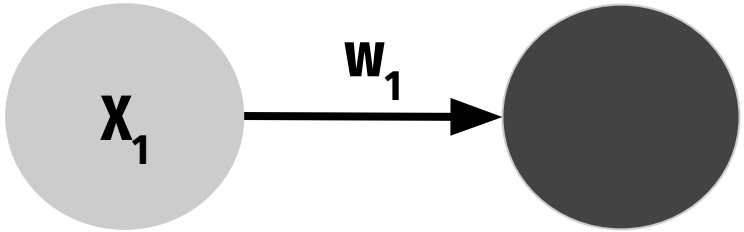
Neuron



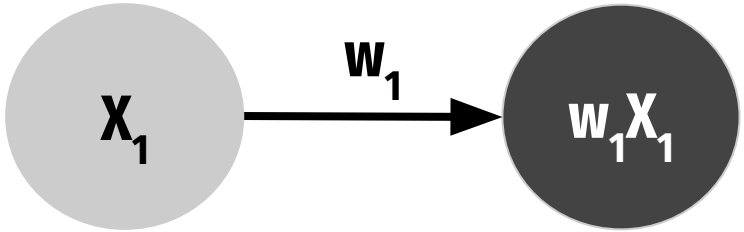
Neuron



Neuron

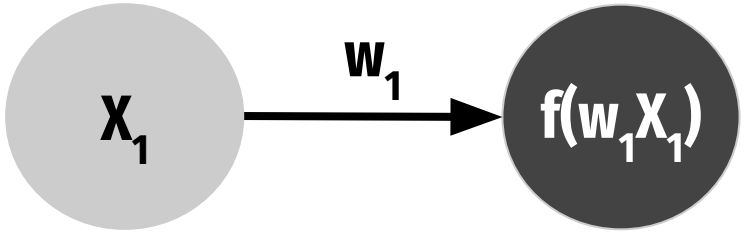


Neuron



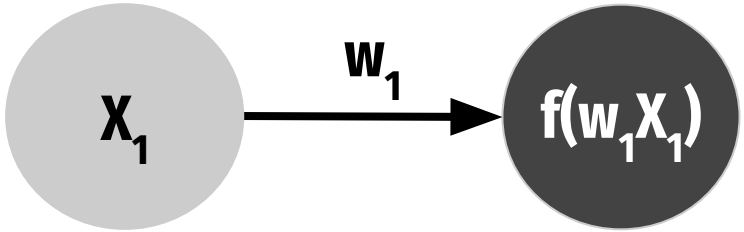
Neuron

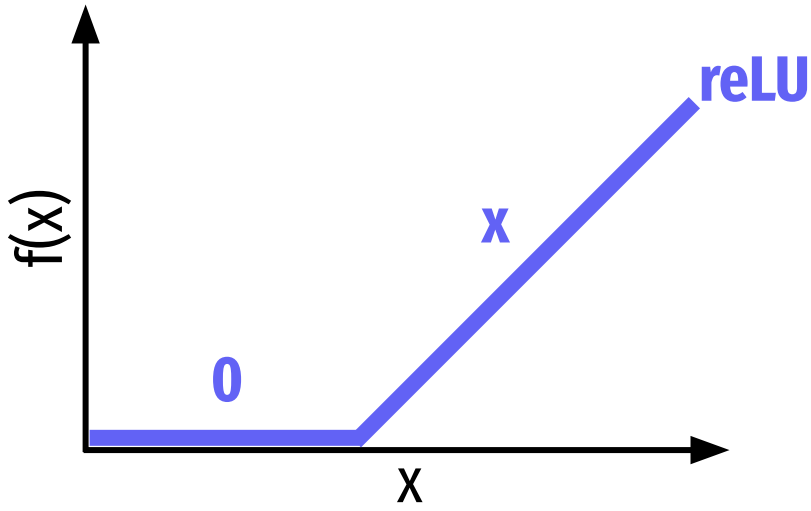
f = *activation function*



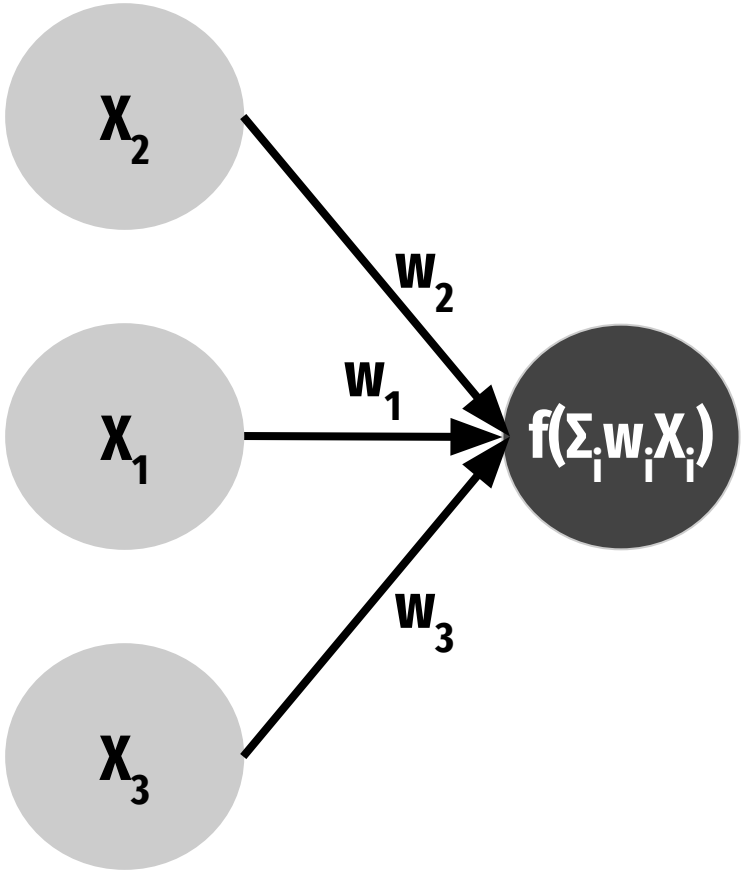
Neuron

f = *activation function*

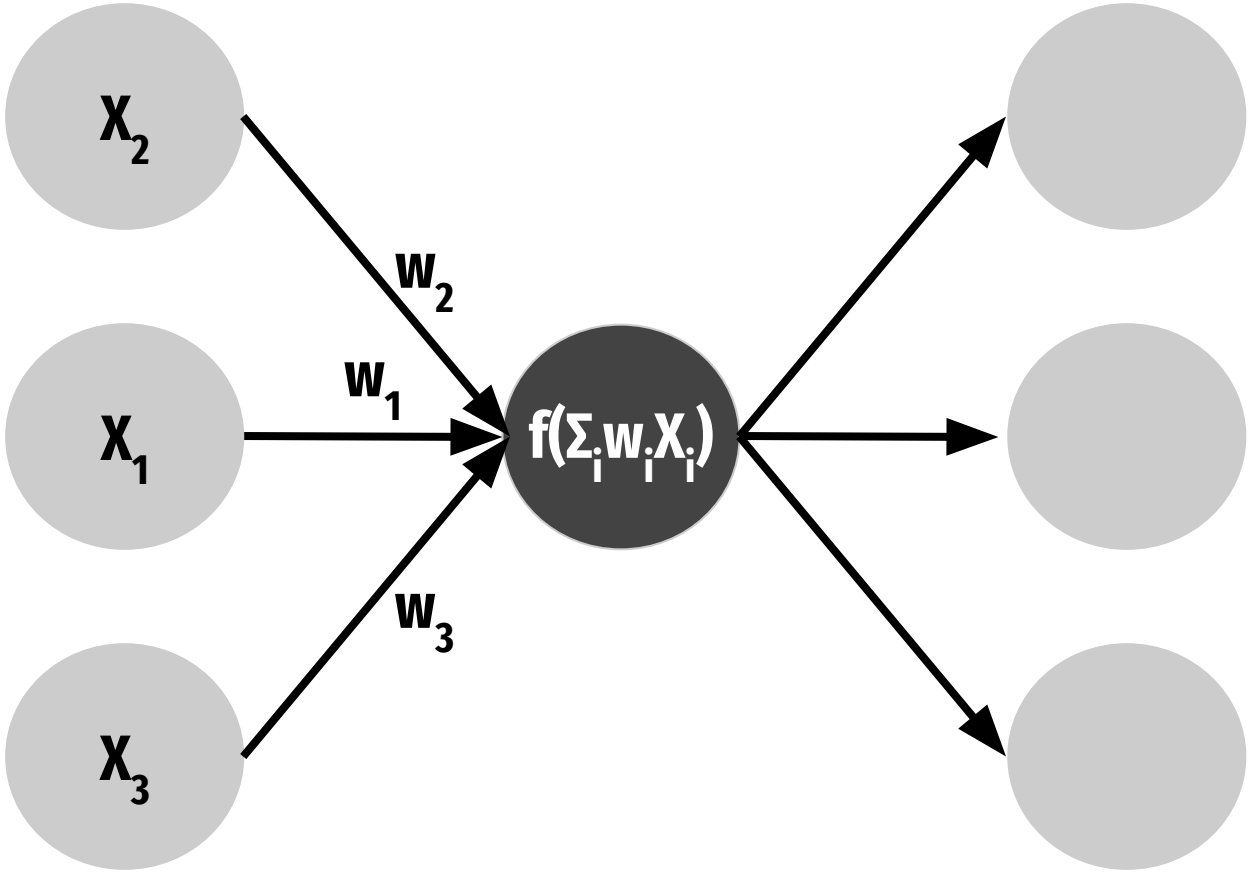




Neuron



Neuron



How does the neural network learn?

How does the neural network learn?

1. Define an objective: minimize distance between predicted and expected value

How does the neural network learn?

1. Define an objective: minimize **distance between predicted and expected value**

Loss function

How does the neural network learn?

1. Define an objective: minimize **distance between predicted and expected value**

Loss function

Mean Square Error (MSE)

$$MSE = \Sigma_i (y_{i,pred} - y_{i,true})^2$$

How does the neural network learn?

1. Define an objective: minimize **distance between predicted and expected value**

Loss function

Mean Square Error (MSE)

$$MSE = \Sigma_i (y_{i,pred} - y_{i,true})^2$$

2. Optimize network to reach this objective

How does the neural network learn?

1. Define an objective: minimize **distance between predicted and expected value**

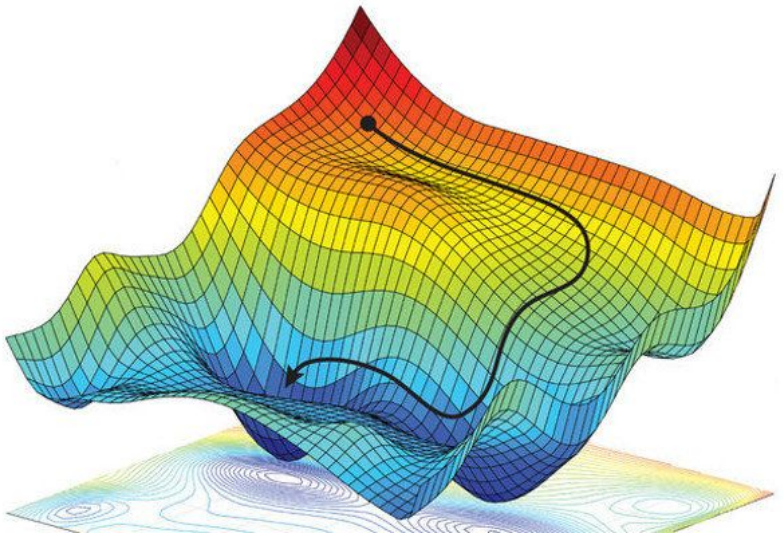
Loss function

Mean Square Error (MSE)

$$MSE = \Sigma_i (y_{i,pred} - y_{i,true})^2$$

2. Optimize network to reach this objective

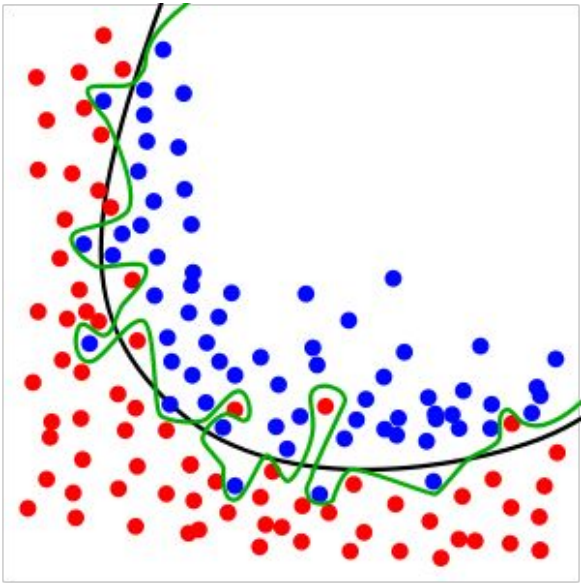
   Compute loss **gradient** vs. weights with backward propagation



How do we know that the neural network has learned correctly?

How do we know that the neural network has learned correctly?

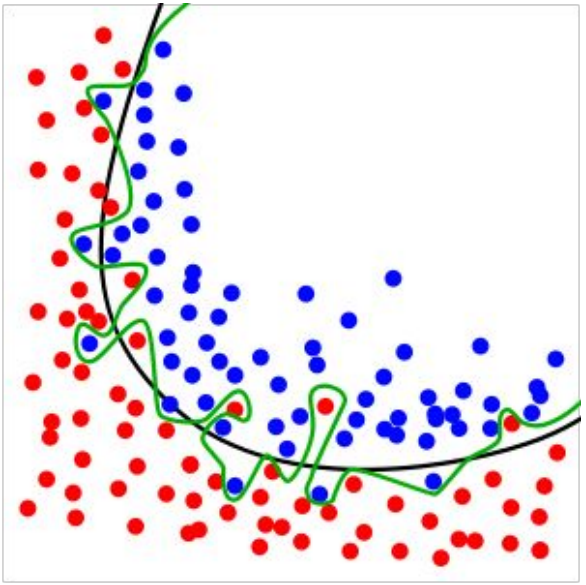
Avoid **overfitting** i.e. neural network learn train data ‘by heart’ and is not able to extrapolate to new data



How do we know that the neural network has learned correctly?

Avoid **overfitting** i.e. neural network learn train data 'by heart' and is not able to extrapolate to new data

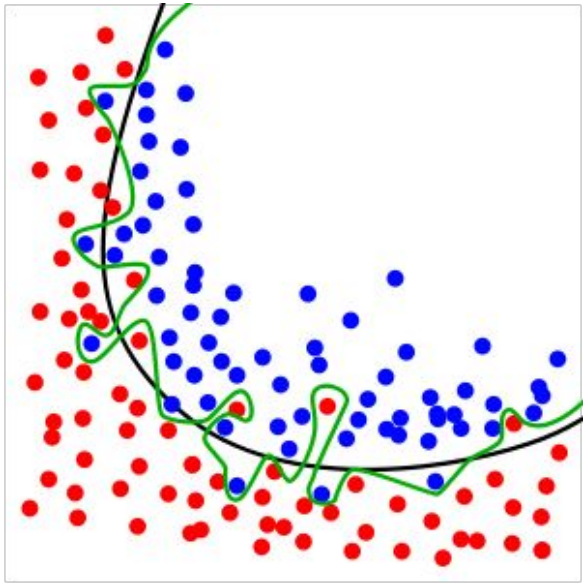
1. Split data set in two subsets: training set & validation set

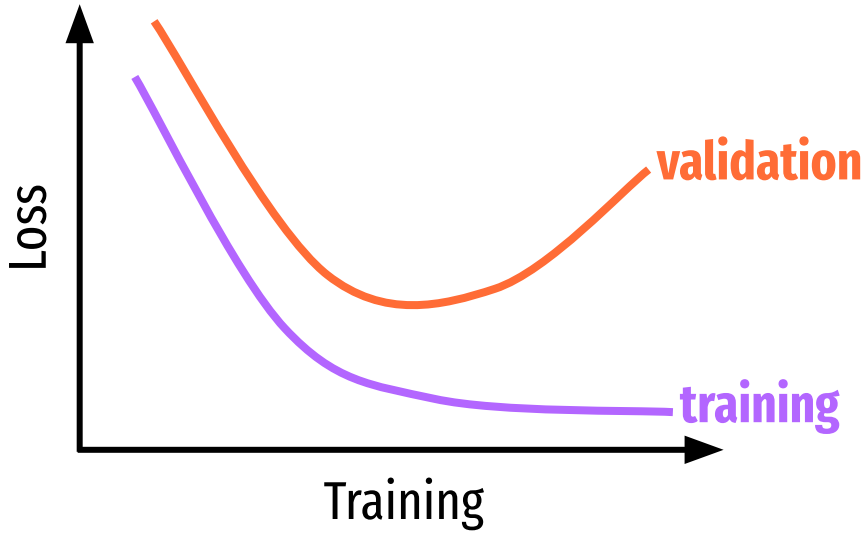


How do we know that the neural network has learned correctly?

Avoid **overfitting** i.e. neural network learn train data 'by heart' and is not able to extrapolate to new data

1. Split data set in two subsets: training set & validation set
2. Track train & validation losses through the training

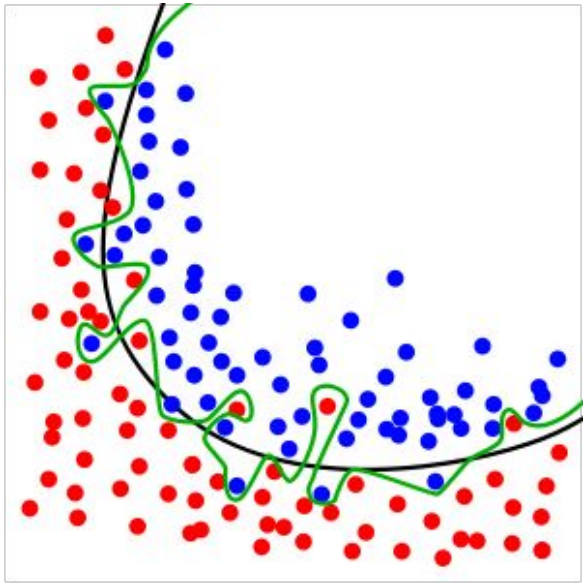


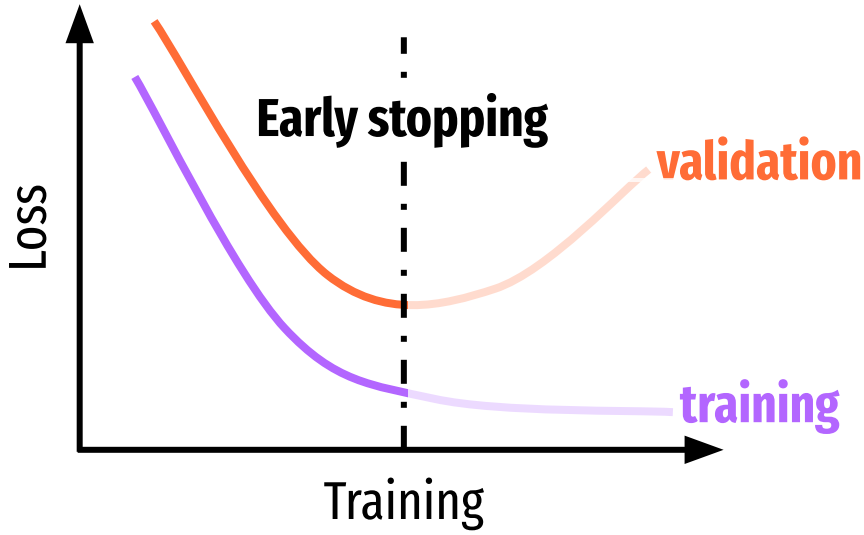


How do we know that the neural network has learned correctly?

Avoid **overfitting** i.e. neural network learn train data 'by heart' and is not able to extrapolate to new data

1. Split data set in two subsets: training set & validation set
2. Track train & validation losses through the training
3. Stop the training when validation loss stop to decrease = **early stopping**





Frameworks (in R)

TensorFlow

Torch

TensorFlow

Google

Torch

Facebook

TensorFlow

Google

Industry-focused

Torch

Facebook

Research-focused

TensorFlow

Torch

Google

Facebook

Industry-focused

Research-focused

Easier to learn (Keras)

Harder to learn

TensorFlow

Torch

Google	Facebook
Industry-focused	Research-focused
Easier to learn (Keras)	Harder to learn
Requires Python	Does not require Python

TensorFlow

Torch

Google	Facebook
Industry-focused	Research-focused
Easier to learn (Keras)	Harder to learn
Requires Python	Does not require Python
Good documentation	Poor documentation

TensorFlow

Torch

Google	Facebook
Industry-focused	Research-focused
Easier to learn (Keras)	Harder to learn
Requires Python	Does not require Python
Good documentation	Poor documentation

Examples of use

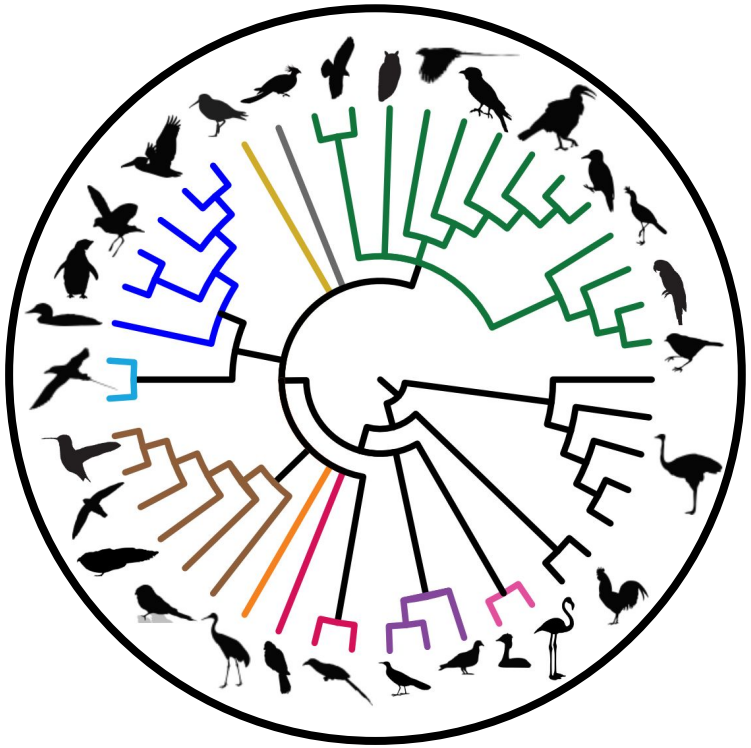
CAT

APPLICATION

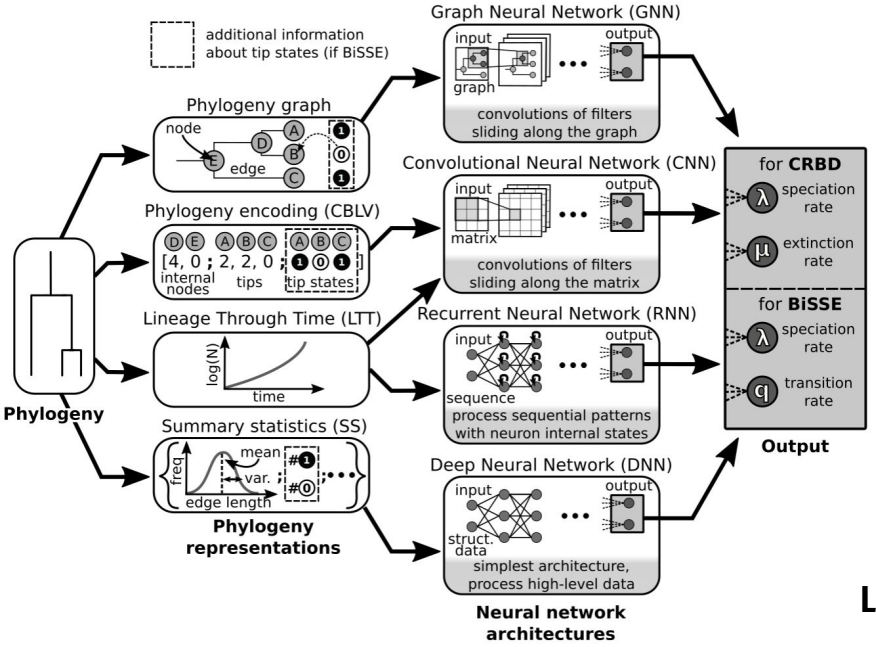
Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Machine learning to classify animal species in camera trap images: Applications in ecology

Michael A. Tabak[1,2] | Mohammad S. Norouzzadeh[3] | David W. Wolfson[1] | Steven J. Sweeney[1] | Kurt C. Vercauteren[4] | Nathan P. Snow[4] | Joseph M. Halseth[4] | Paul A. Di Salvo[1] | Jesse S. Lewis[5] | Michael D. White[6] | Ben Teton[6] | James C. Beasley[7] | Peter E. Schlichting[7] | Raoul K. Boughton[8] | Bethany Wight[8] | Eric S. Newkirk[9] | Jacob S. Ivan[9] | Eric A. Odell[9] | Ryan K. Brook[10] | Paul M. Lukacs[11] | Anna K. Moeller[11] | Elizabeth G. Mandeville[2,12] | Jeff Clune[3] | Ryan S. Miller[1]



Rates of speciation & extinction?



Lajaaiti, Hartig, et al. in prep

Thanks!

Questions