
# Deep Learning

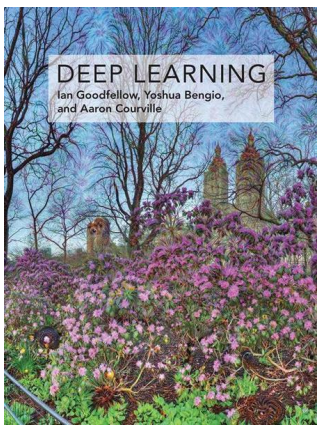
**Asja Fischer**

Department of Mathematics – Ruhr-University Bochum

Winter term 2019

# GENERAL INFORMATION

- Lecturer: Asja Fischer, IB 3/153, asja.fischer@rub.de
- Lecture: Thursday 10-12 am, ND 5/99
- Exercises: Sina Däubener, Tom Kaufmann, Joshua Butke (for biologists), time fixed next week
- Course material: https://moodle.ruhr-uni-bochum.de (password: neuronalesnetz)




- Literature: `https://www.deeplearningbook.org/`
- Exams: 04.02.20 and 06.04.20, 10-12 am

# WHAT IS MACHINE LEARNING?

## A definition of learning

Learning is a goal-directed process that results in a change of behavior and is based on experience.

# WHAT IS MACHINE LEARNING?

## A definition of learning

Learning is a goal-directed process that results in a change of behavior and is based on experience.

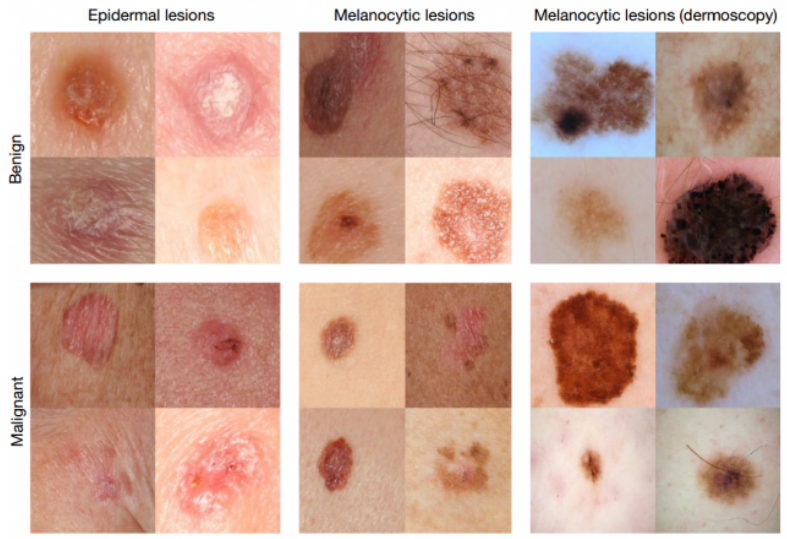
In machine learning context

- “behavior” corresponds making predictions, making decisions, or taking actions.
- “experience” corresponds to data provided to the learning algorithm.

## A definition of machine learning

Machine learning is the study of algorithms that can automatically learn from and make predictions on data.

# MACHINE LEARNING FOR SKIN CANCER CLASSIFICATION

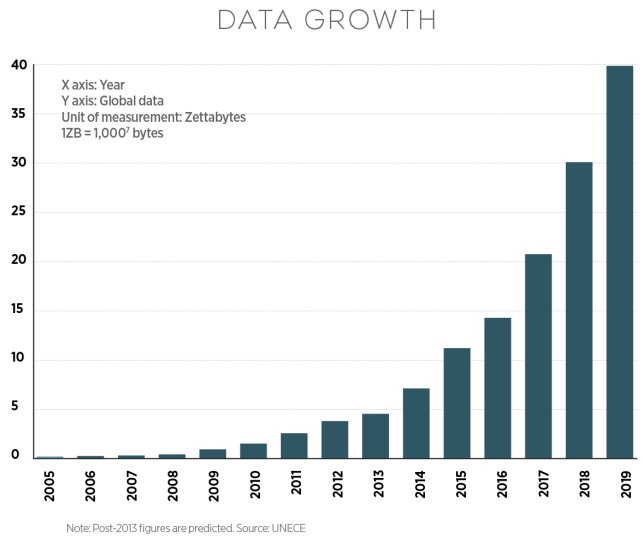



- Classification: benign lesions or malignant skin cancers?
- Stanford team trained machine learning algorithms on $\sim$ 130000 skin lesion images.
- Algorithms performance on par with 21 dermatologists. ( AUC $=$ 91% for dermoscopy images)

Esteva et al.*Dermatologist-level classification of skin cancer with deep neural networks*, Nature 549, 2017

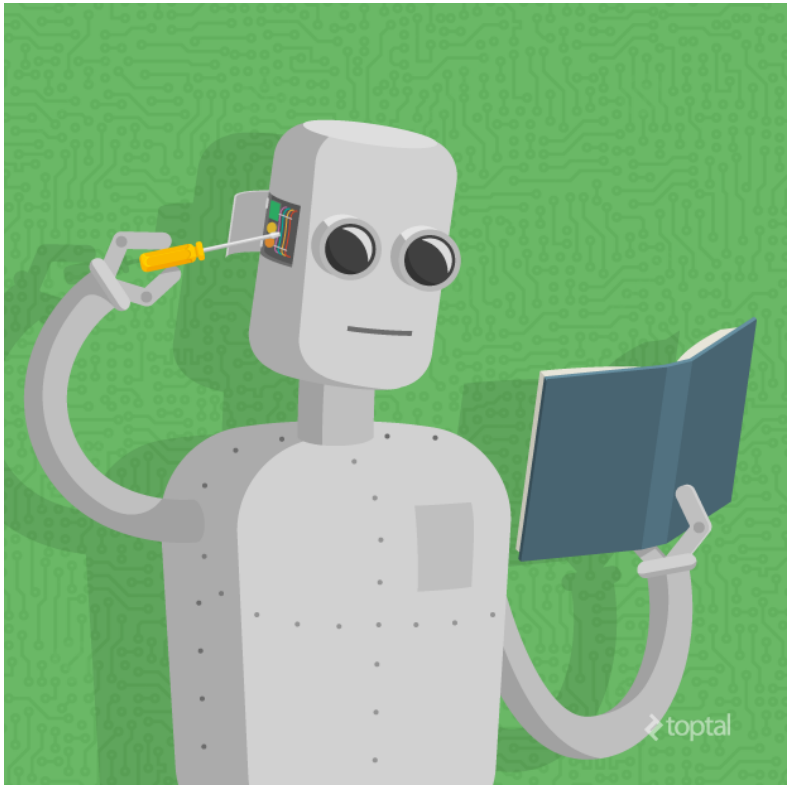
# WHY MACHINE LEARNING?

- Without machine learning computer programs build on hand coded known rules.
- ML algorithms learn to solve tasks on their own.
- Learning improves with experience/data.
- We are now in the age of big data.
- ML allows to process and use big amounts of data.
- ML algorithms can find patterns in data too complex for humans.
- They can even reach super-human performance.



# MACHINE LEARNING PARADIGMS

As for humans or animals there exist different learning paradigms.

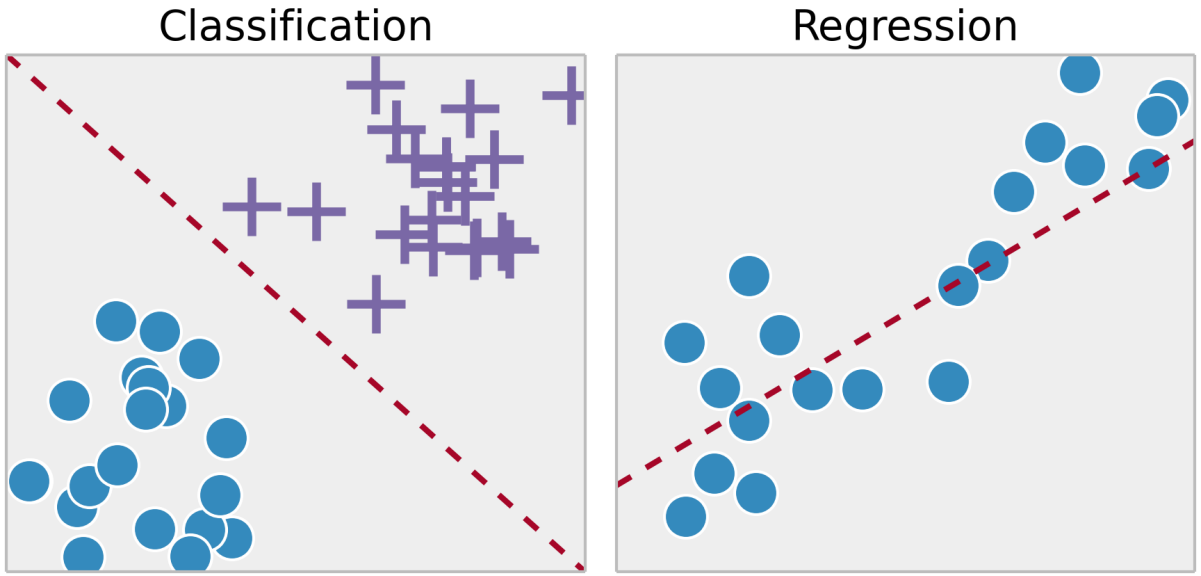



The mayor three paradigms are

- supervised learning
- unsupervised learning
- reinforcement learning

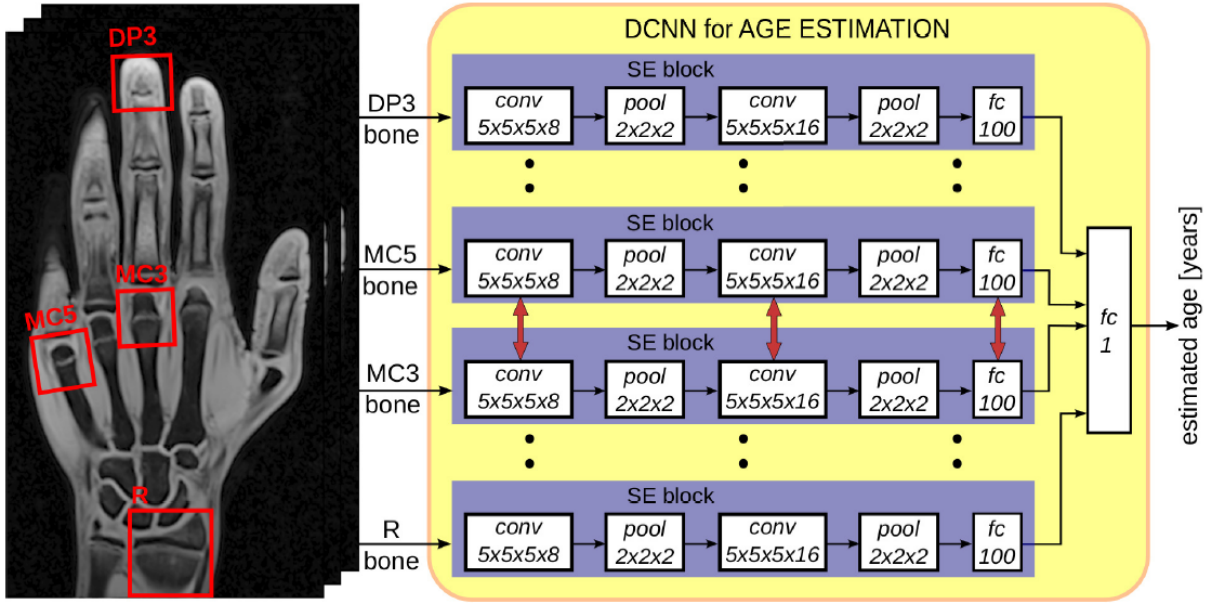
# SUPERVISED LEARNING

- Training data: each example is a pair consisting of an input and a desired output value.
- Task: infer a function that maps inputs to outputs and generalizes to unseen data.



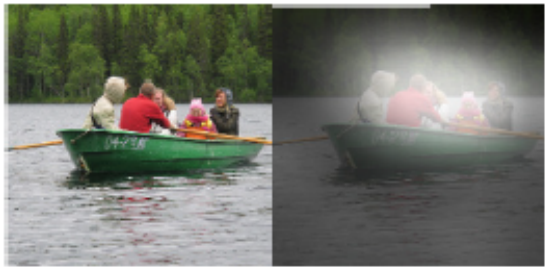
# SUPERVISED LEARNING - EXAMPLES

## Age estimation from hand MR volume




## Image caption generation

A little girl sitting on a bed with a teddy bear.



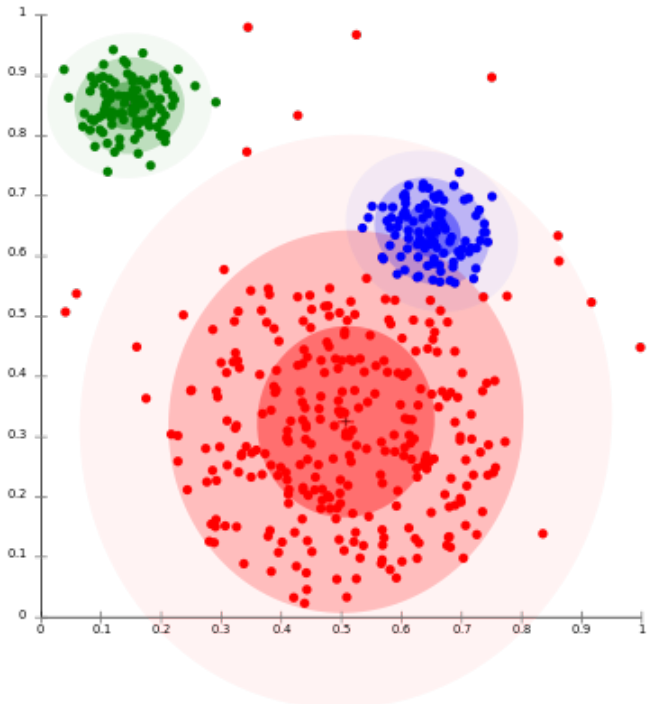
A group of people sitting on a boat in the water.

A giraffe standing in a forest with trees in the background.

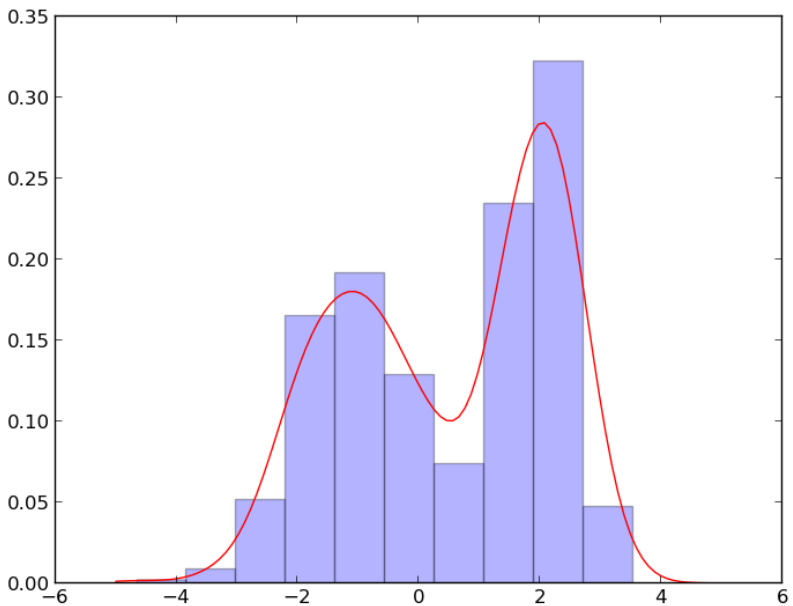
# UNSUPERVISED LEARNING

- Training data: training data consists of unlabeled input points.
- Task: find and describe intrinsic structure in the data.

clustering

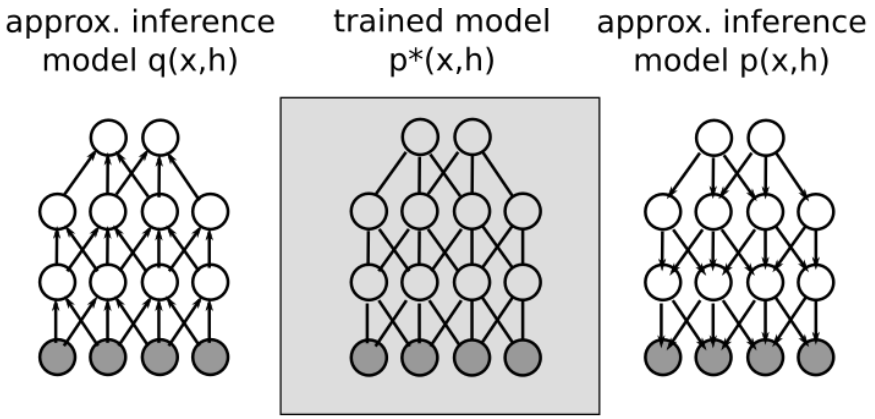


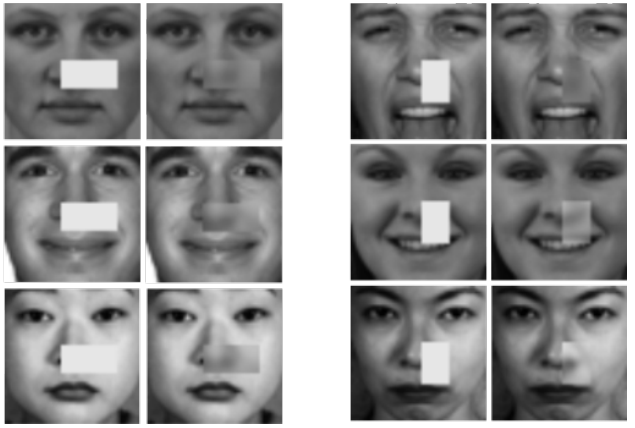

density fitting



# UNSUPERVISED LEARNING - EXAMPLES

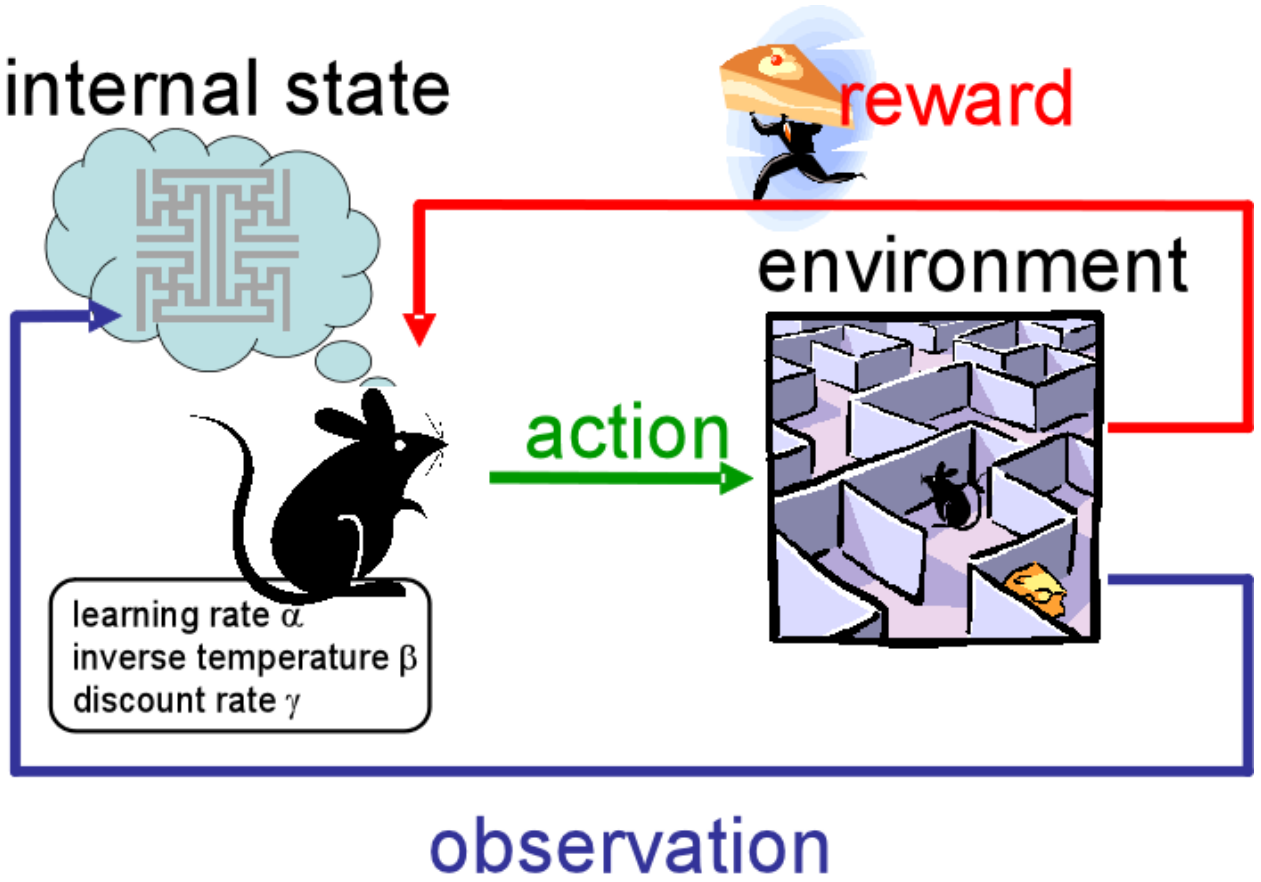
## Image sampling and inpainting



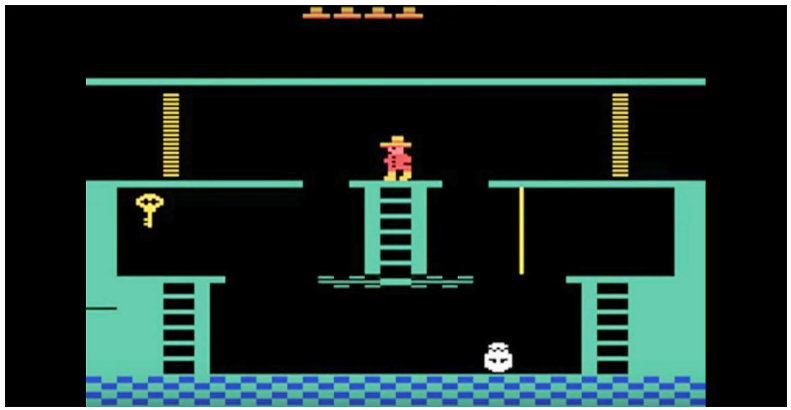
# REINFORCEMENT LEARNING

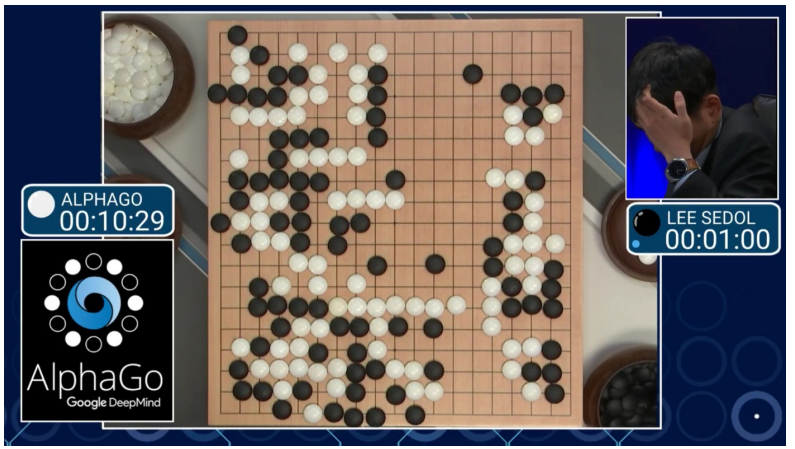
Task: a agent learns to take actions in an environment to maximize some reward.



# REINFORCEMENT LEARNING - EXAMPLES

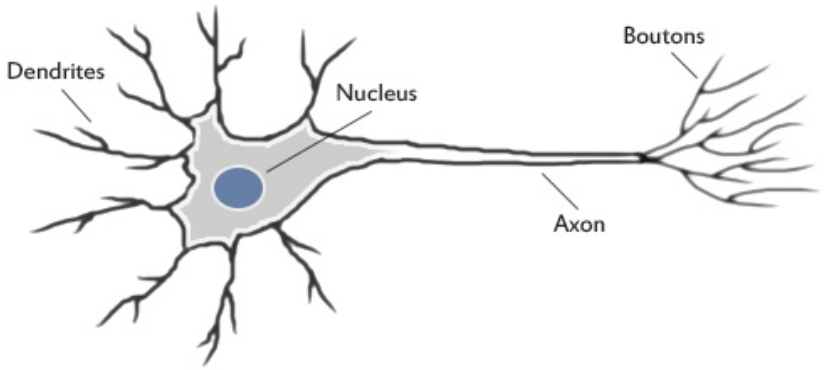
Learning to play Atari games and Go

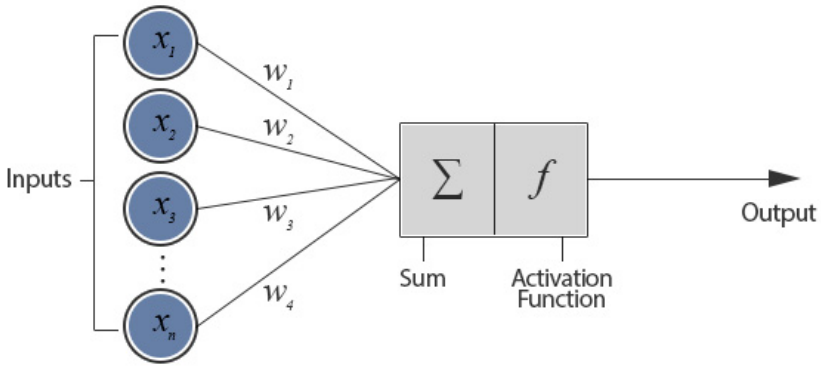




# NEURAL NETWORKS

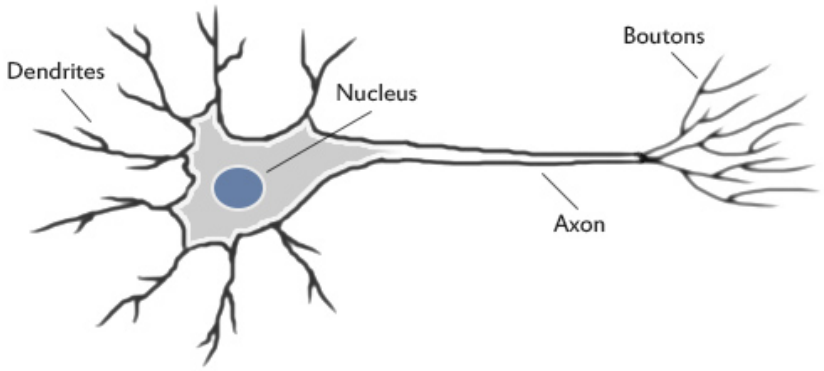
single neuron

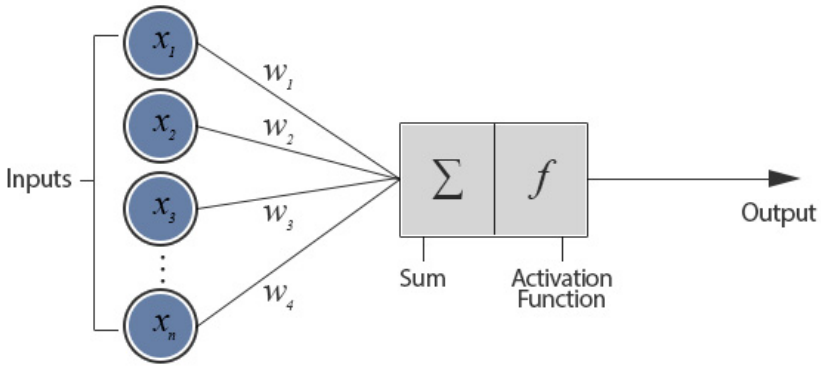


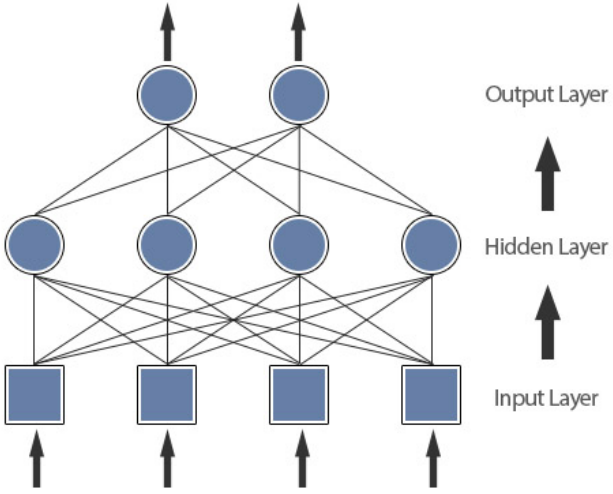
# NEURAL NETWORKS

single neuron



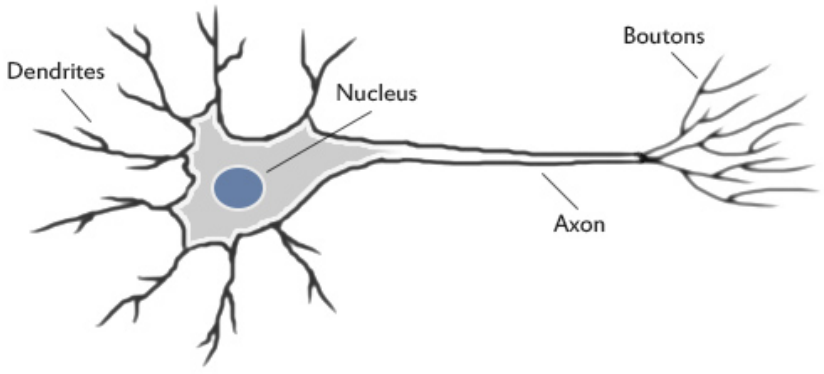



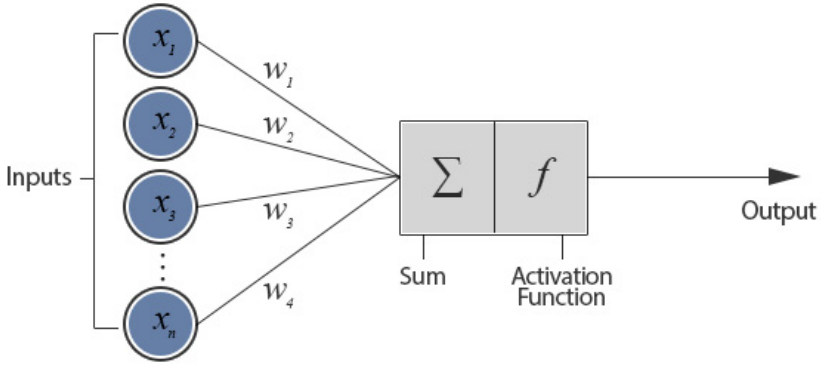

neural network



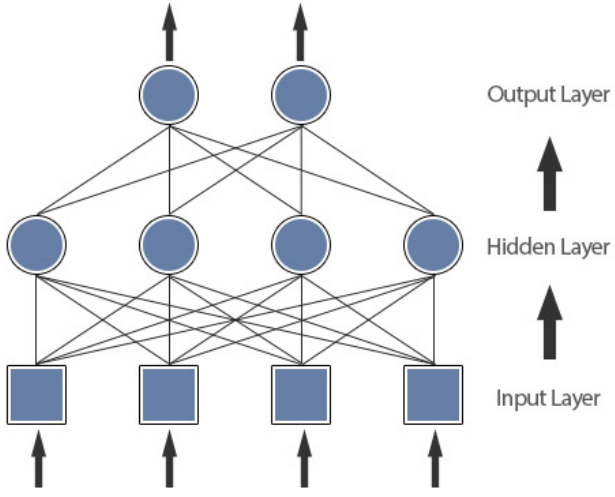
# NEURAL NETWORKS

single neuron



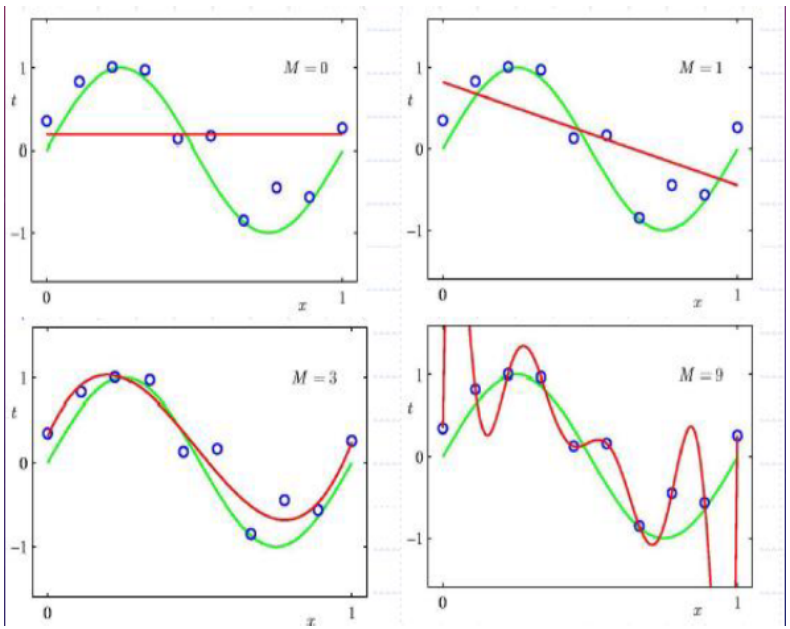

neural network

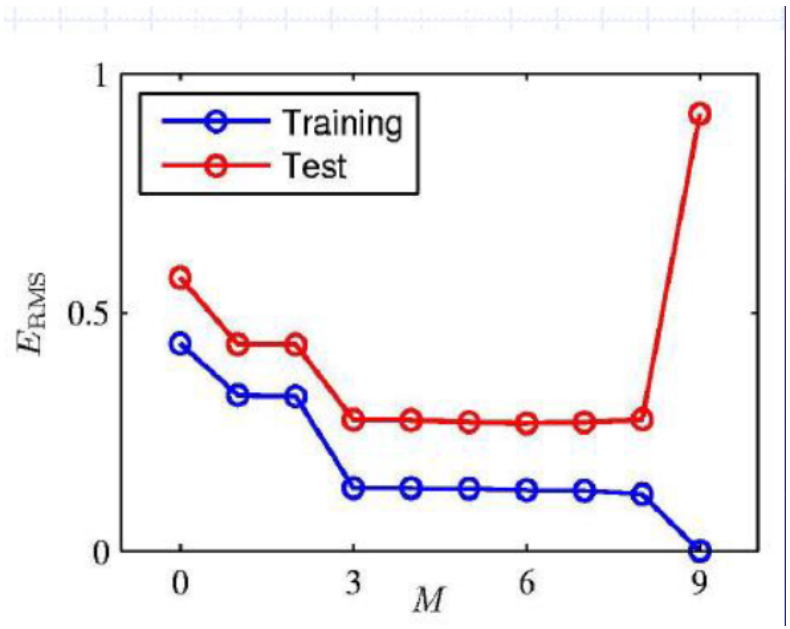



Learning in neural networks corresponds to adapting the parameters based on **optimiztion** algorithms.

# GENERALIZATION

- Central problem in machine learning and statistics: the learned model needs to generalize to unseen data.
- **Underfitting**: the model "ignores" important features of the data.
- **Overfitting**: the model is "over-adapted" to training data and does not generalize well.

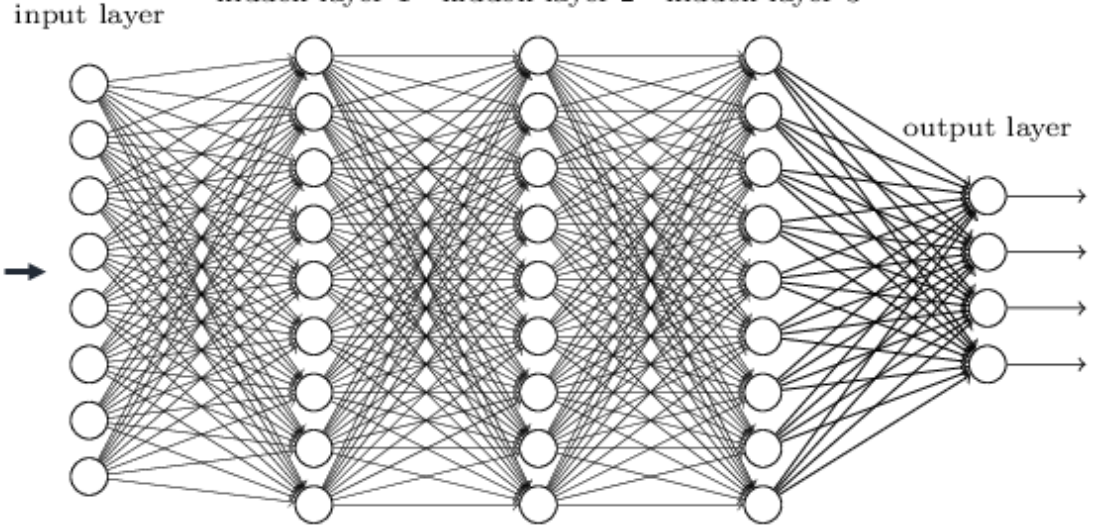



- **Regularization** helps to control the complexity of the model.

# DEEP LEARNING

Deep neural networks learn hierarchical feature representations




We see different kind of network architectures.

# APPLICATION AREAS OF DEEP LEARNING

Deep learning has been applied (and revolutionized) a lot of fields, like

- image recognition
- speech recognition
- natural language processing
- machine translation
- social network analysis
- bioinformatics
- medicine
- ....

# Mathematical Background Knowledge

# LINEAR ALGEBRA AND ANALYSIS

You should be familiar with concepts like

- dot product
- matrix transpose/inverse
- identity/symmetric/orthogonal matrix
- norm ($L1$ and $L2$ norm)
- eigendecomposition
- singular value decomposition
- trance of a matrix
- maxima, minima and saddle points of a function
- gradient/Hessian matrix of a scalar valued multi-variable
- ...

# STATISTICS AND PROBABILITY THEORY

You should be familiar with concepts like

- discrete and continuous random variables
- probability mass/density functions
- (computing) marginal probabilities
- conditional probabilities
- (conditional) independence of variables
- expectations
- variance and covariance
- standard distributions (e.g. Bernoulli/(multivariate) Gaussian)
- Bayes' rule
- change of variables
- maximum likelihood
- ...