Grundlagen der Programmierung

Praktikum

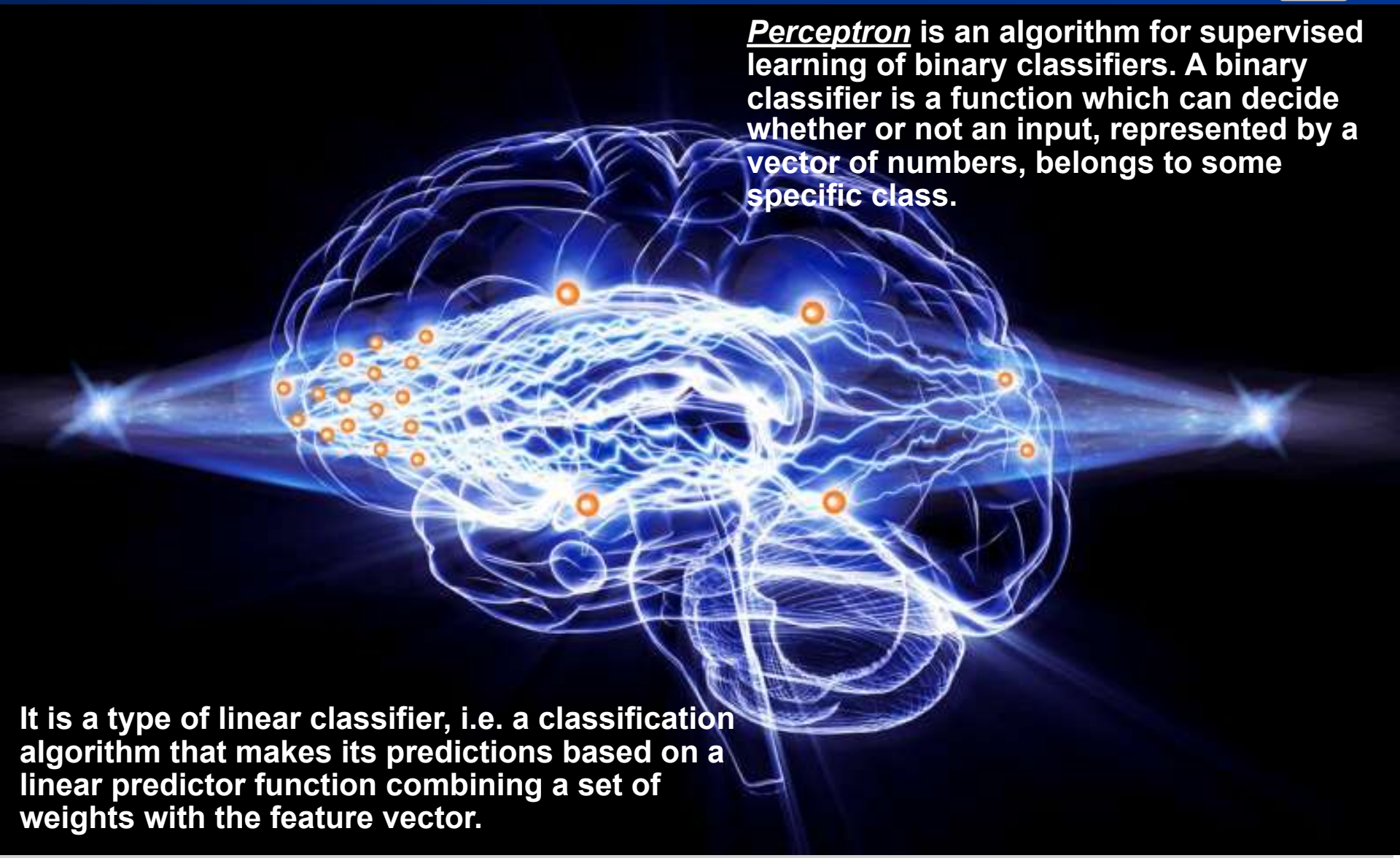
2. Plenartermin

Artemiy Belousov, Ivan Kisel, Grigory Kozlov, Martin Parnet

Abgabe Milestone 1

Heute um 23:59

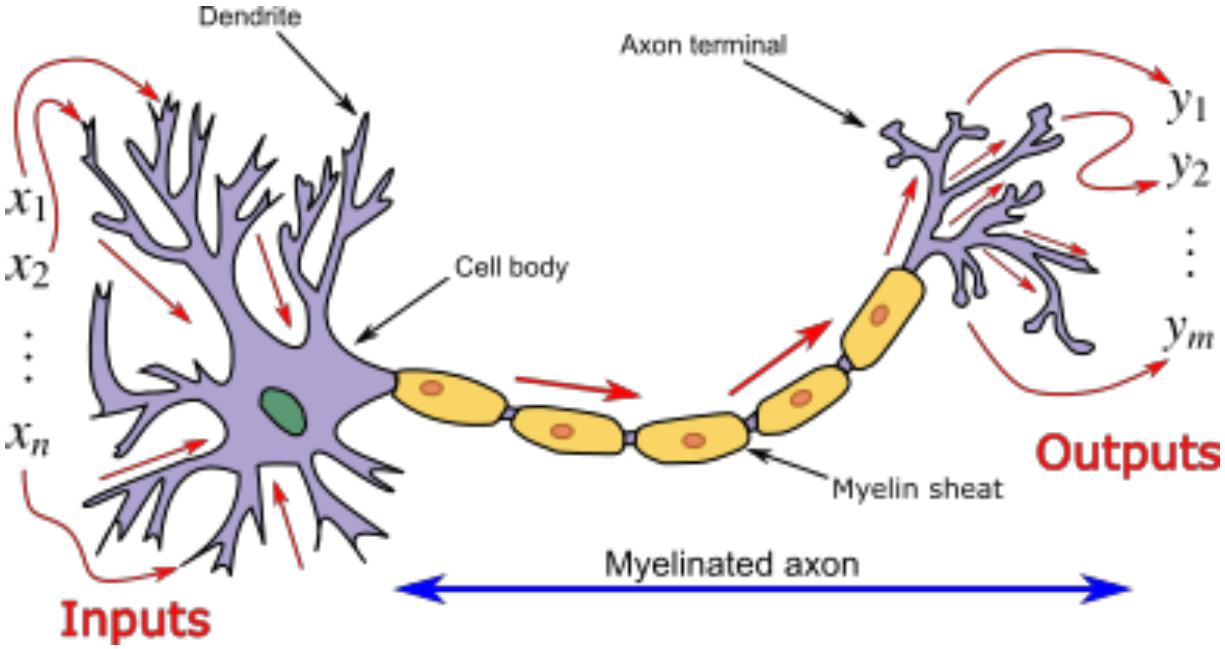
Perceptron neural net



Perceptron **is an algorithm for supervised learning of binary classifiers. A binary classifier is a function which can decide whether or not an input, represented by a vector of numbers, belongs to some specific class.**

It is a type of linear classifier, i.e. a classification algorithm that makes its predictions based on a linear predictor function combining a set of weights with the feature vector.

Neuron

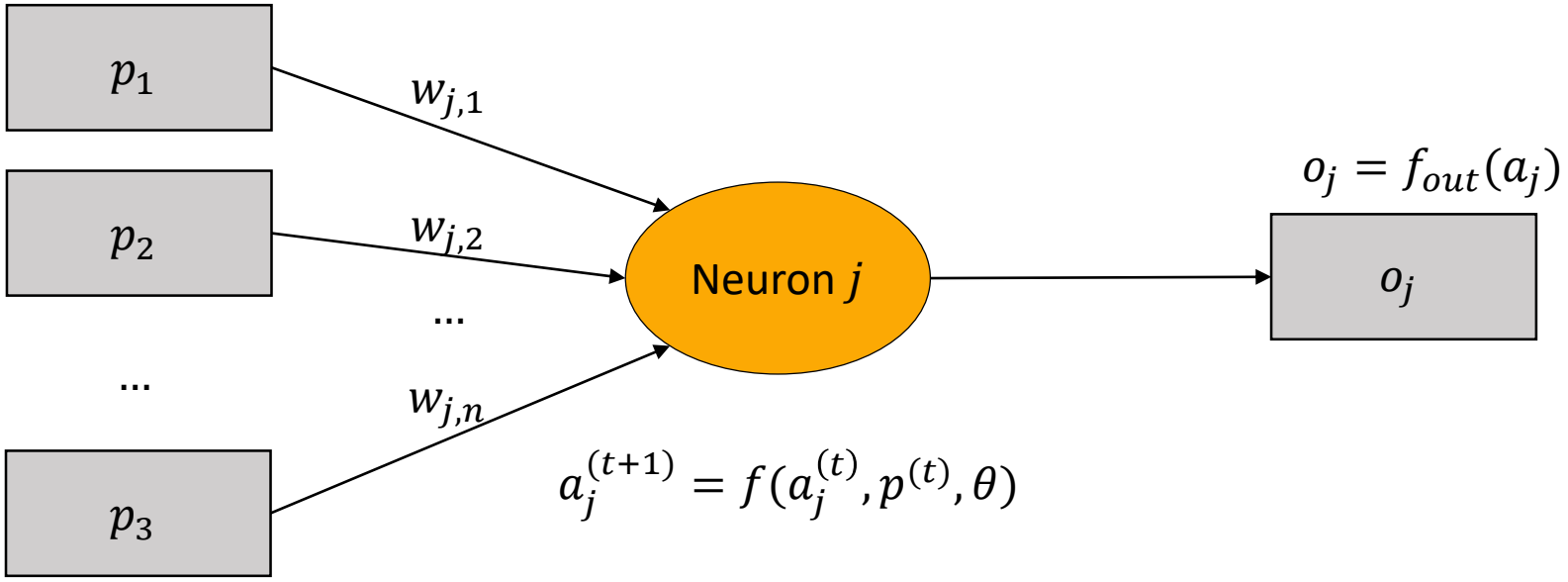


- Computation
  - Hodkin Huxley equations
  - FitzHugh-Nagumo model

The **Hodgkin–Huxley model**, or **conductance-based model**, is a mathematical model that describes how action potentials in neurons are initiated and propagated. It is a set of nonlinear differential equations that approximates the electrical characteristics of excitable cells such as neurons and cardiac myocytes. It is a continuous time model.

The FHN Model is an example of a relaxation oscillator because, if the external stimulus exceeds a certain threshold value, the system will exhibit a characteristic excursion in phase space, before the variables relax back to their rest values.

Neuron

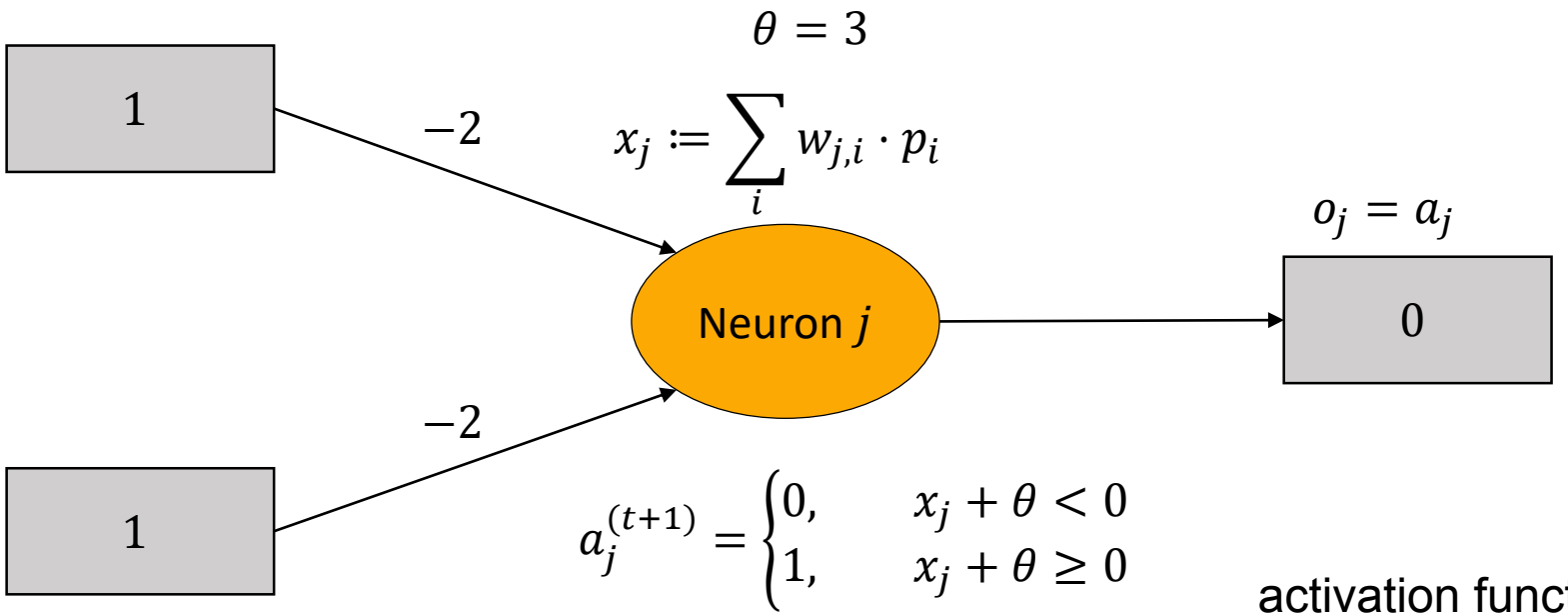


p_i	inputs
$w_{j,i}$	weights
o	output

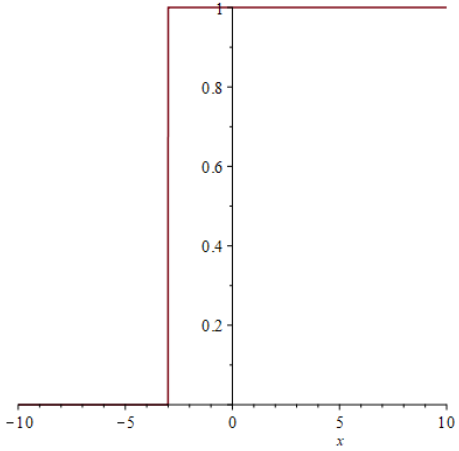
θ	threshold
f	activation function
a	activation state

f_{out}	output function

Neuron - example



activation function



- Inputs and outputs: 0 or 1
- Threshold moves activation function
- Activation function produces 0 or 1

Neuron characteristics

- Inputs and outputs
- Changing the state

$$x_j = \sum_i w_{j,i} \cdot p_i^{(t)}$$

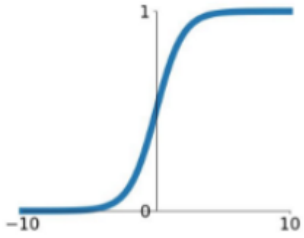
$$a_j^{(t+1)} = f\left(x_j - \theta\right)$$

- f is activation function $\left(\frac{1}{1+e^{-x}}, \frac{1}{2}\tanh\frac{x}{2} + \frac{1}{2}, \frac{\sqrt{x}}{\sqrt{1+x^2}} \ldots\right)$
- f is differentiable (important for backpropagation)
- θ is represented by extra neuron (optional)

Activation functions

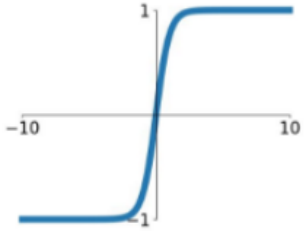
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



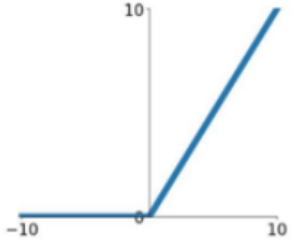
tanh

$$\tanh(x)$$



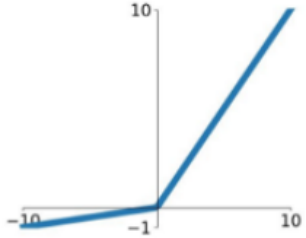
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

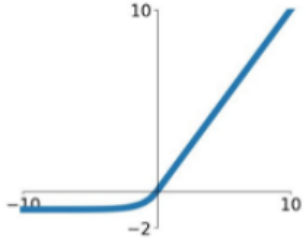


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



Activation function: Softmax

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

Z is a vector of the inputs to the output layer (if you have 10 output units, then there are 10 elements in z). And **j** indexes the output units, so j = 1, 2, ..., K

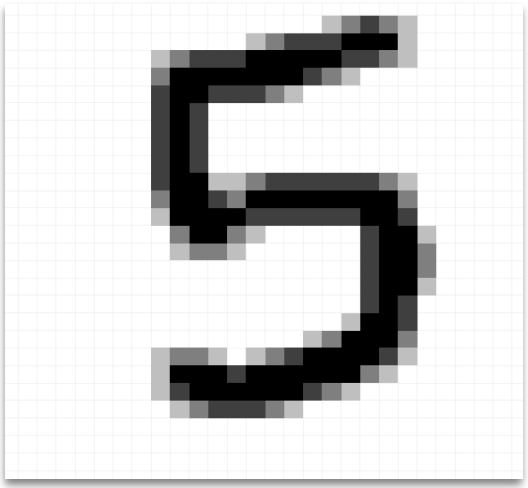
The softmax function squashes the outputs of each unit to be between 0 and 1, just like a sigmoid function. But it also divides each output such that the total sum of the outputs is equal to 1

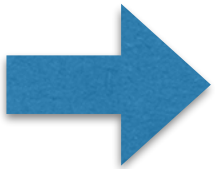
Neural net

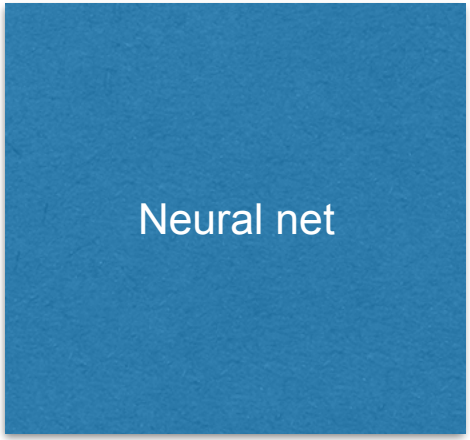
input

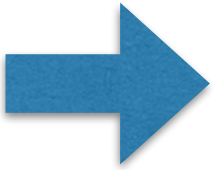
processing

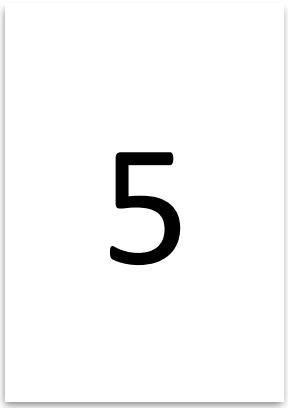
output



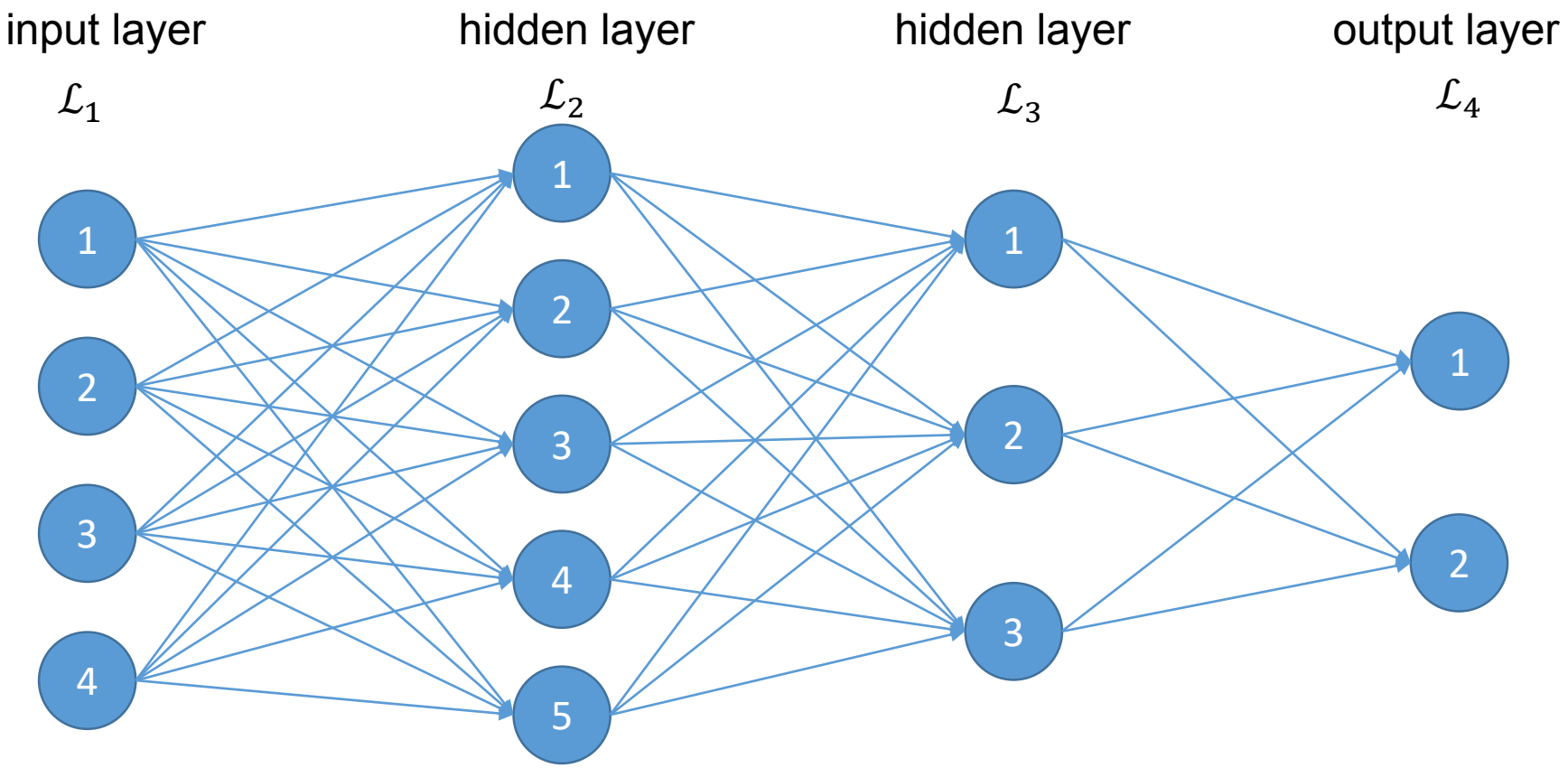








Neural net



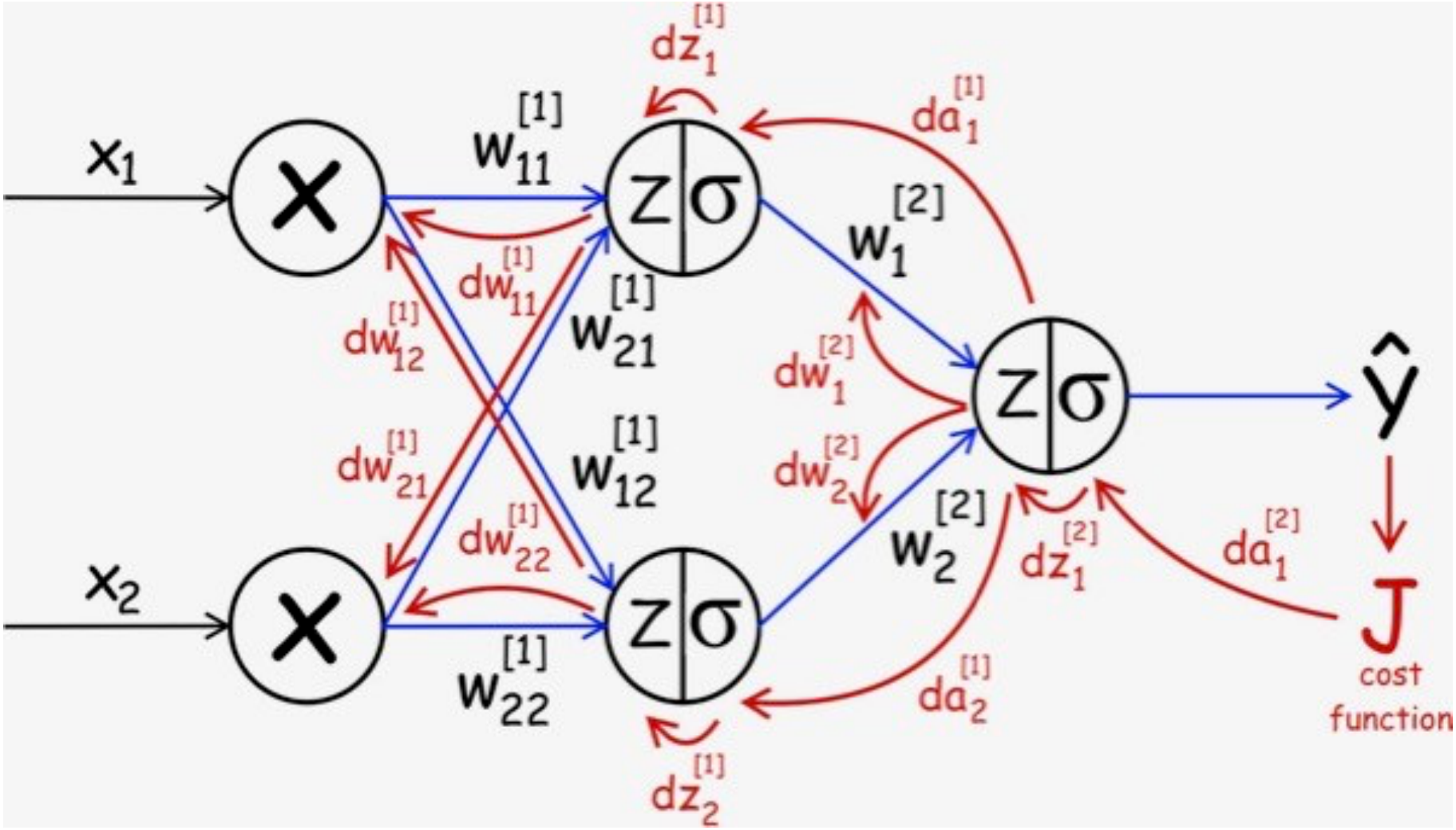
- Topology : {5, 3, 2} (without θ-neuron)
- Feedforward network

Training

- Optimization method
- Adjusting weights and bias
- Different algorithms online / offline
- Subset with data
  - Mini batches

- Back propagation
- Reinforcement learning (AlphaZero)

Back propagation

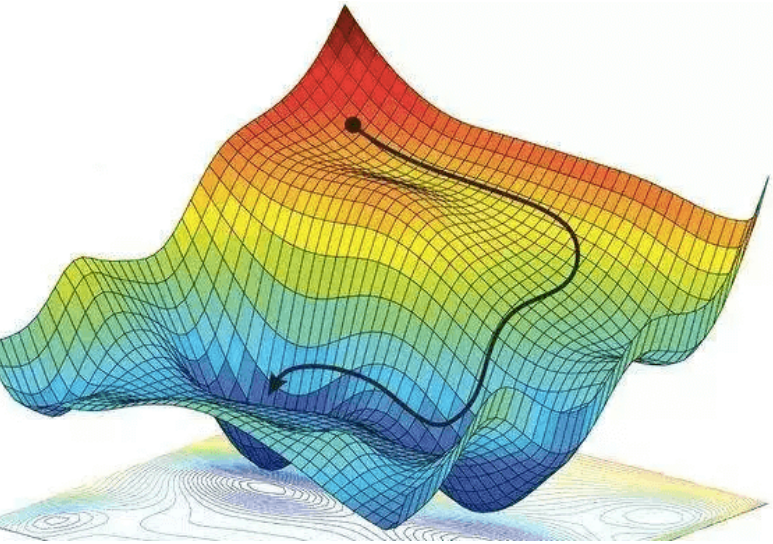


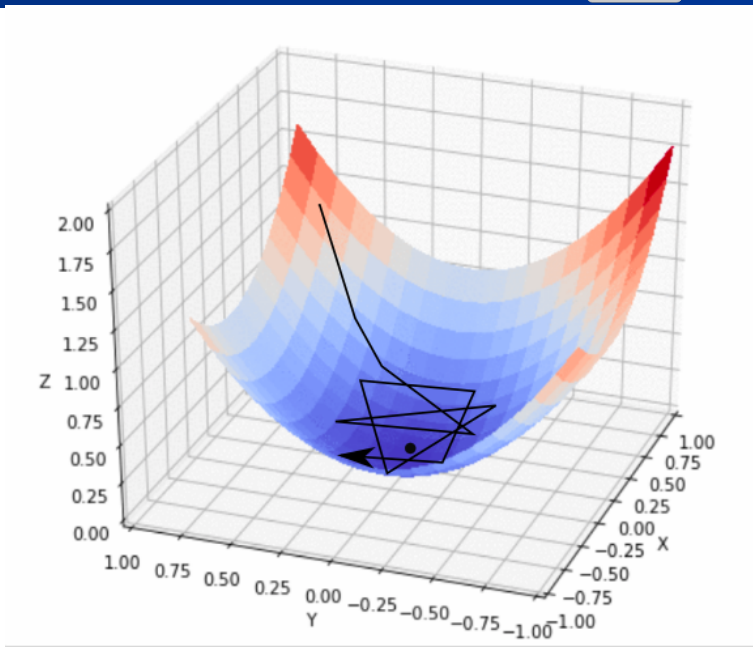
Loss function

- **Loss function** is used for parameter estimation, and the event in question is some function of the difference between estimated and true values for an instance of data.
- **Loss function** helps in optimizing the parameters of the neural networks.
- The loss is calculated using **loss function** by matching the target(actual) value and predicted value by a neural network.
- Then we use the **gradient descent** method to update the weights of the neural network such that the loss is minimized.

- **Objective:** to minimize the loss for a neural network by optimizing its parameters(weights).

Learning rate η





$$\Delta w_{n,j} = -\eta \frac{\partial E}{\partial w_{n,j}}$$

- Controls the speed of convergence
- Adaptive adjustment depending on the error rate
- Typically 0.001 as lower barrier
- Upper barrier 10

Gradients

Output layer gradient:

$$\sigma_m = (t_m - f(x_m)) \cdot f'(x_m)$$

Hidden layer gradient:

$$\sigma_n = \sum_{p \in \mathcal{L}_N} \left(t_p - f(x_p)\right) \cdot f'(x_p) \cdot w_{p,n} \cdot f'(x_n)$$

Softmax derivative:

$f'(x) = S \cdot (1 - S)$; *where S is activation function Softmax*

LReLU derivative:

$$f'(x) = \begin{cases} 1 & , x >= 0 \\ 0.01 & , x < 0 \end{cases}$$

Adaptive Moment Estimation (Adam)

Computes adaptive learning rates for each parameter.

- Compute the decaying averages of past and past squared gradients

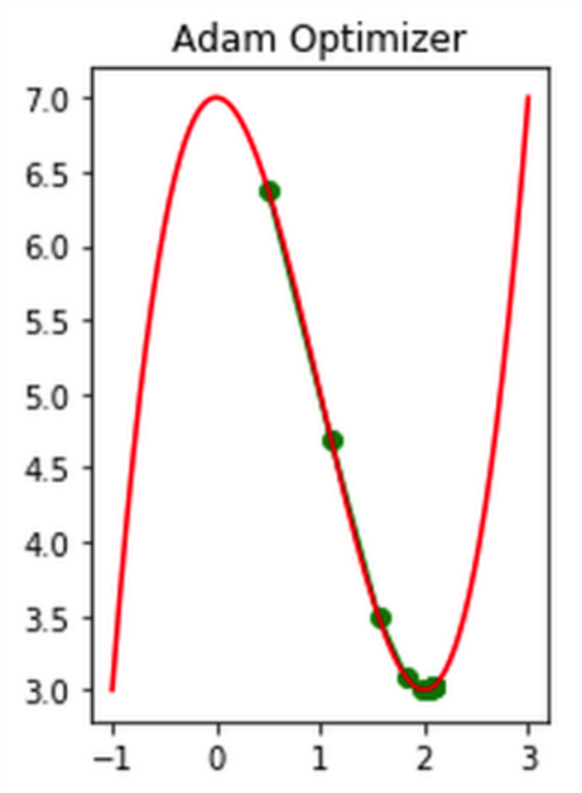
$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$
$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

- Correct for bias

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$
$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

- Update weights

$$\theta_{t+1} = \theta_t + \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t$$



$\beta 1 = 0.9$; $\beta_2 = 0.999$; $\epsilon = 10^{-8}$

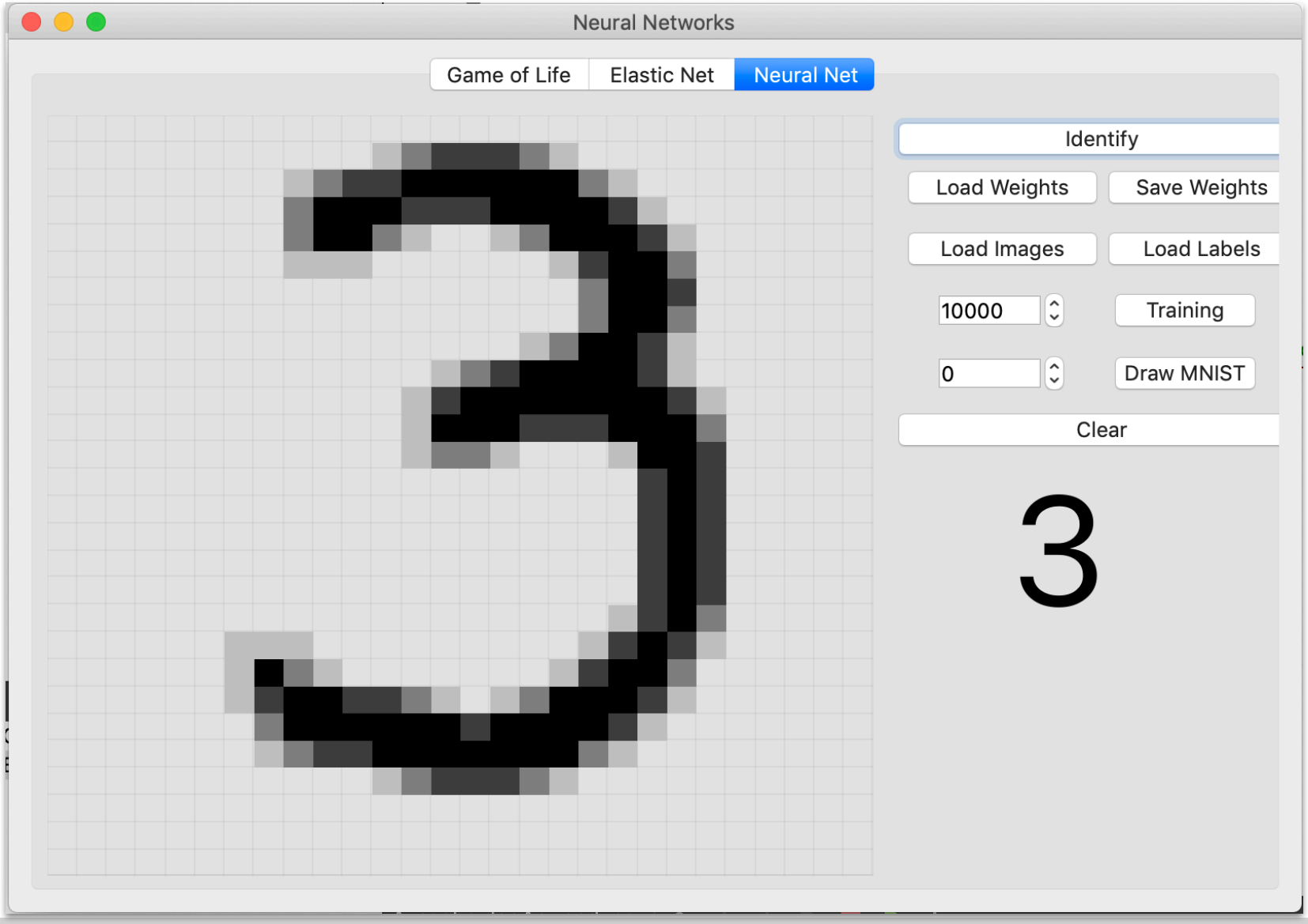
Artificial Neural Network

In Qt the Neural Network is implemented as a separate file with:

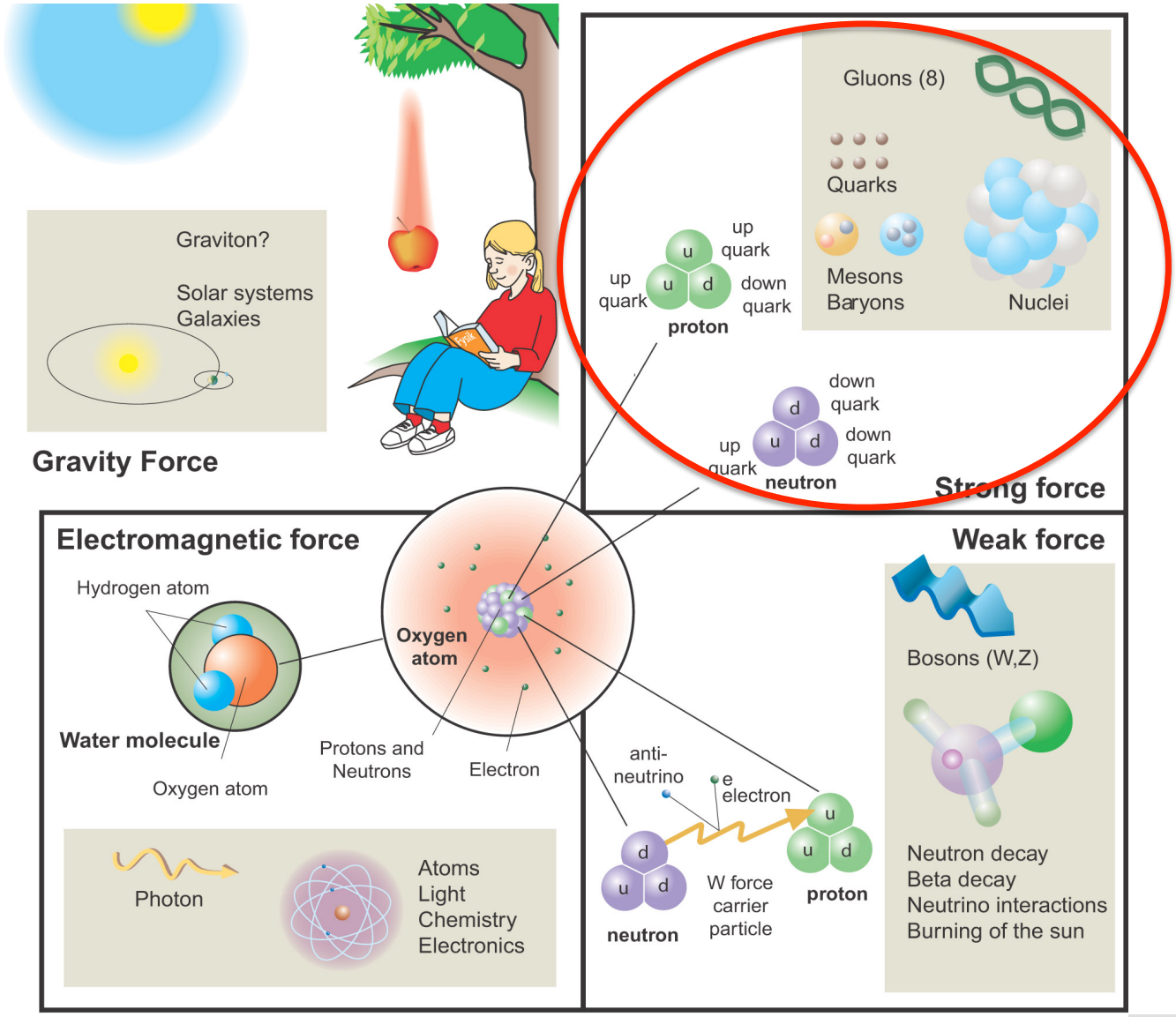
- Class Neuron with Perceptron parameters.
  - calculation output gradients
  - calculation hidden gradients
  - random weights
  - activation function
  - updating input weights

- Class Net - include variables and functions for the Neural Network:
  - std::vector <Layer> of neurone layers - parameters of every perceptrons.
  - calculations of Weights, gradients, input/output values etc.
  - realisation of training and validation Neural Network.

Artificial Neural Network - running ANN



Fundamental Interactions



The Standard Model

The Standard Model is a quantum theory that summarizes our current knowledge of the physics of fundamental particles and fundamental interactions (interactions are manifested by forces and by decay rates of unstable particles).

FERMIONS
matter constituents
spin = 1/2, 3/2, 5/2, ...

Leptons spin = 1/2

Flavor	Mass GeV/c²	Electric charge
ν_L lightest neutrino*	$(0-2)\times10^{-9}$	0
e electron	0.000511	−1
ν_M middle neutrino*	$(0.009-2)\times10^{-9}$	0
μ muon	0.106	−1
ν_H heaviest neutrino*	$(0.05-2)\times10^{-9}$	0
τ tau	1.777	−1

Quarks spin = 1/2

Flavor	Approx. Mass GeV/c²	Electric charge
u up	0.002	2/3
d down	0.005	−1/3
c charm	1.3	2/3
s strange	0.1	−1/3
t top	173	2/3
b bottom	4.2	−1/3

*See the neutrino paragraph below.

Spin is the intrinsic angular momentum of particles. Spin is given in units of ħ, which is the quantum unit of angular momentum where $\hbar = h/2\pi = 6.58\times10^{-25}$ GeV s $=1.05\times10^{-34}$ J s.

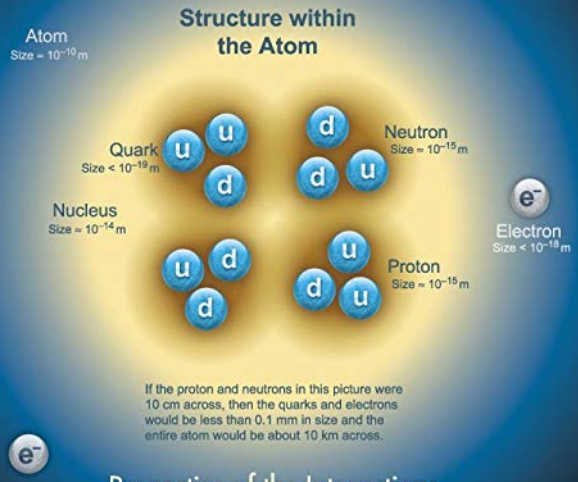
Electric charges are given in units of the proton's charge. In SI units the electric charge of the proton is 1.60×10^{-19} coulombs.

The energy unit of particle physics is the electronvolt (eV), the energy gained by one electron in crossing a potential difference of one volt. **Masses** are given in GeV/c² (remember $E = mc^2$) where 1 GeV = 10^9 eV $=1.60\times10^{-10}$ joule. The mass of the proton is 0.938 GeV/c² $= 1.67\times10^{-27}$ kg.

Neutrinos
Neutrinos are produced in the sun, supernovae, reactors, accelerator collisions, and many other processes. Any produced neutrino can be described as one of three neutrino flavor states ν_e, ν_μ, or ν_τ, labelled by the type of charged lepton associated with its production. Each is a defined quantum mixture of the three definite-mass neutrinos ν_L, ν_M, and ν_H for which currently allowed mass ranges are shown in the table. Further exploration of the properties of neutrinos may yield powerful clues to puzzles about matter and antimatter and the evolution of stars and galaxy structures.

Matter and Antimatter
For every particle type there is a corresponding antiparticle type, denoted by a bar over the particle symbol (unless + or − charge is shown). Particle and antiparticle have identical mass and spin but opposite charges. Some electrically neutral bosons (e.g., Z^0, γ, and $\eta_c = c\bar{c}$ but not $K^0 = d\bar{s}$) are their own antiparticles.

Structure within the Atom



If the proton and neutrons in this picture were 10 cm across, then the quarks and electrons would be less than 0.1 mm in size and the entire atom would be about 10 km across.

BOSONS
force carriers
spin = 0, 1, 2, ...

Unified Electroweak spin = 1

Name	Mass GeV/c²	Electric charge
γ photon	0	0
W^-	80.39	−1
W^+ W bosons	80.39	+1
Z^0 Z boson	91.188	0

Strong (color) spin = 1

Name	Mass GeV/c²	Electric charge
g gluon	0	0

Higgs Boson spin = 0

Name	Mass GeV/c²	Electric charge
H Higgs	126	0

Higgs Boson
The Higgs boson is a critical component of the Standard Model. Its discovery helps confirm the mechanism by which fundamental particles get mass.

Color Charge
Only quarks and gluons carry "strong charge" (also called "color charge") and can have strong interactions. Each quark carries three types of color charge. These charges have nothing to do with the colors of visible light. Just as electrically-charged particles interact by exchanging photons, in strong interactions, color-charged particles interact by exchanging gluons.

Quarks Confined in Mesons and Baryons
Quarks and gluons cannot be isolated – they are confined in color-neutral particles called **hadrons**. This confinement (binding) results from multiple exchanges of gluons among the color-charged constituents. As color-charged particles (quarks and gluons) move apart, the energy in the color-force field between them increases. This energy eventually is converted into additional quark-antiquark pairs. The quarks and antiquarks then combine into hadrons; these are the particles seen to emerge.

Two types of hadrons have been observed in nature **mesons** $q\bar{q}$ and **baryons** qqq. Among the many types of baryons observed are the proton (uud), antiproton ($\bar{u}\bar{u}\bar{d}$), and neutron (udd). Quark charges add in such a way as to make the proton have charge 1 and the neutron charge 0. Among the many types of mesons are the pion π^+ ($u\bar{d}$), kaon K^- ($s\bar{u}$), and B^0 ($d\bar{b}$).

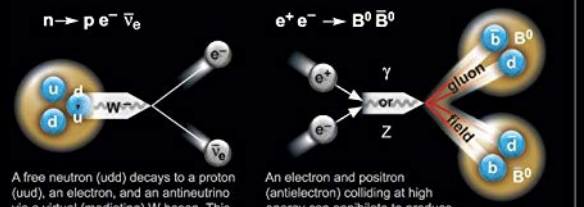
Properties of the Interactions

The strengths of the interactions (forces) are shown relative to the strength of the electromagnetic force for two u quarks separated by the specified distances.

Property	Gravitational Interaction	Weak Interaction (Electroweak)	Electromagnetic Interaction (Electroweak)	Strong Interaction
Acts on:	Mass – Energy	Flavor	Electric Charge	Color Charge
Particles experiencing:	All	Quarks, Leptons	Electrically Charged	Quarks, Gluons
Particles mediating:	Graviton (not yet observed)	W^+ W^- Z^0	γ	Gluons
Strength at 10^{-18} m	10^{-41}	0.8	1	25
Strength at 3×10^{-17} m	10^{-41}	10^{-4}	1	60

Learn more at ParticleAdventure.org

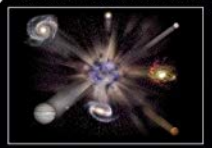
Particle Processes
These diagrams are an artist's conception. Orange shaded areas represent the cloud of gluons.



A free neutron (udd) decays to a proton (uud), an electron, and an antineutrino via a virtual (mediating) W boson. This is neutron β (beta) decay.

An electron and positron (antielectron) colliding at high energy can annihilate to produce $\bar{B}^0$ and B^0 mesons via a virtual Z boson or a virtual photon.

Unsolved Mysteries
Driven by new puzzles in our understanding of the physical world, particle physicists are following paths to new wonders and startling discoveries. Experiments may even find extra dimensions of space, microscopic black holes, and/or evidence of string theory.

Why is the Universe Accelerating?

The expansion of the universe appears to be accelerating. Is this due to Einstein's Cosmological Constant? If not, will experiments reveal a new force of nature or even extra (hidden) dimensions of space?

Why No Antimatter?

Matter and antimatter were created in the Big Bang. Why do we now see only matter except for the tiny amounts of antimatter that we make in the lab and observe in cosmic rays?

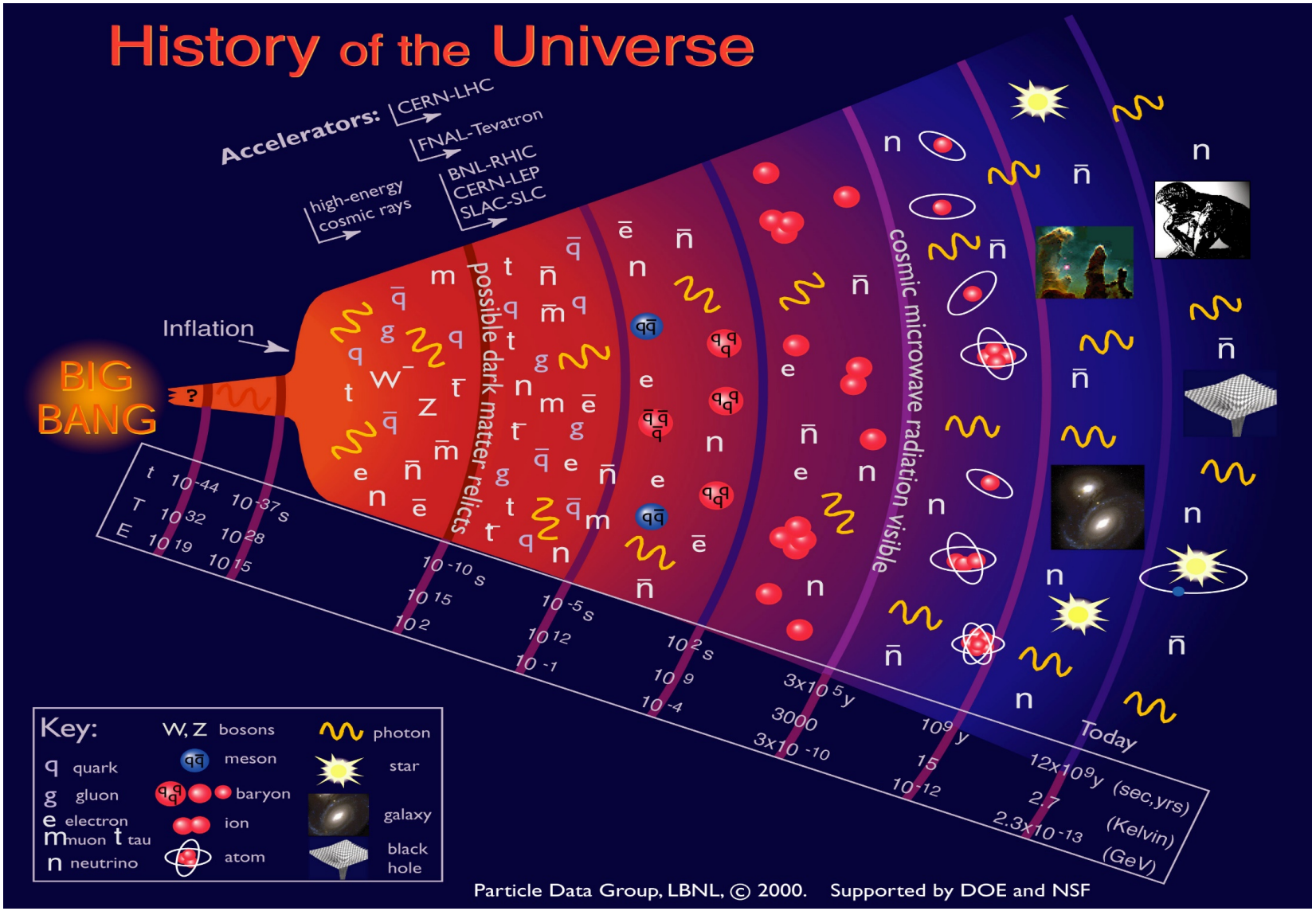
What is Dark Matter?

Invisible forms of matter make up much of the mass observed in galaxies and clusters of galaxies. Does this dark matter consist of new types of particles that interact very weakly with ordinary matter?

Are there Extra Dimensions?

An indication for extra dimensions may be the extreme weakness of gravity compared with the other three fundamental forces (gravity is so weak that a small magnet can pick up a paper clip overwhelming Earth's gravity).

Quarks and Gluons in the Universe



Constituents of Matter

FERMIONS

matter constituents
spin = 1/2, 3/2, 5/2, ...

Leptons spin = 1/2

Flavor	Mass GeV/c^2	Electric charge
ν_e electron neutrino	$<1\times10^{-8}$	0
e electron	0.000511	−1
ν_μ muon neutrino	<0.0002	0
μ muon	0.106	−1
ν_τ tau neutrino	<0.02	0
τ tau	1.7771	−1

Quarks spin = 1/2

Flavor	Approx. Mass GeV/c^2	Electric charge
u up	0.003	2/3
d down	0.006	−1/3
c charm	1.3	2/3
s strange	0.1	−1/3
t top	175	2/3
b bottom	4.3	−1/3

BOSONS

force carriers
spin = 0, 1, 2, ...

Unified Electroweak spin = 1

Name	Mass GeV/c^2	Electric charge
γ photon	0	0
W^-	80.4	−1
W^+	80.4	+1
Z^0	91.187	0

Strong (color) spin = 1

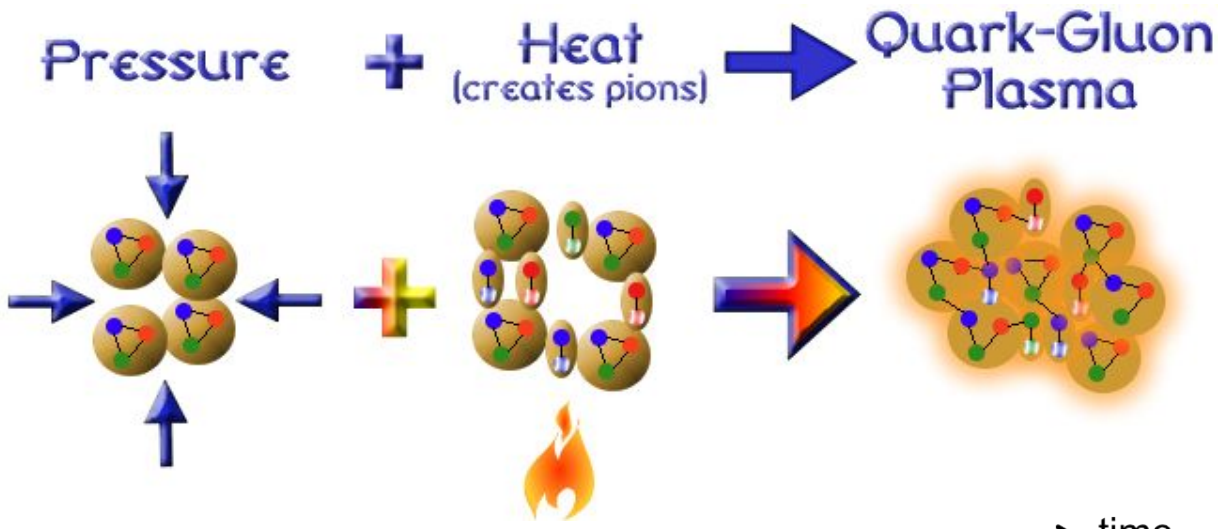
Name	Mass GeV/c^2	Electric charge
g gluon	0	0

Basic ingredients of nuclear matter

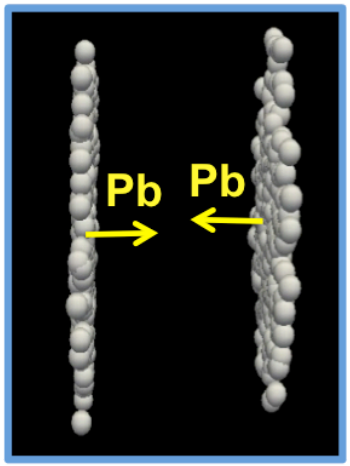
Heavy quarks appear in hot nuclear matter

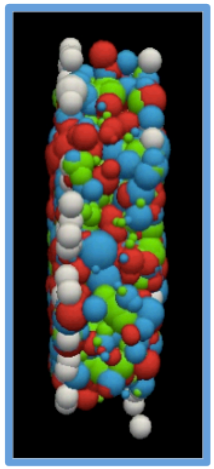
B. Wislouch

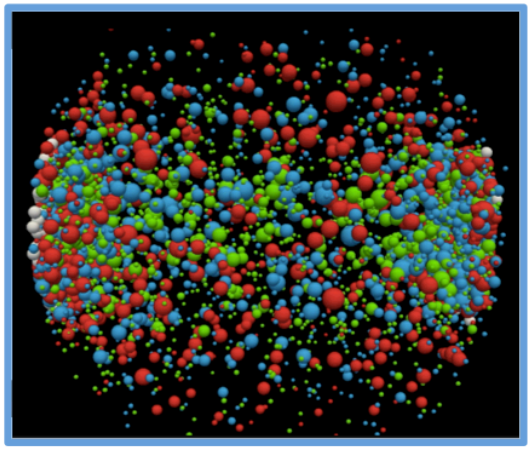
QGP: Quark-Gluon Plasma

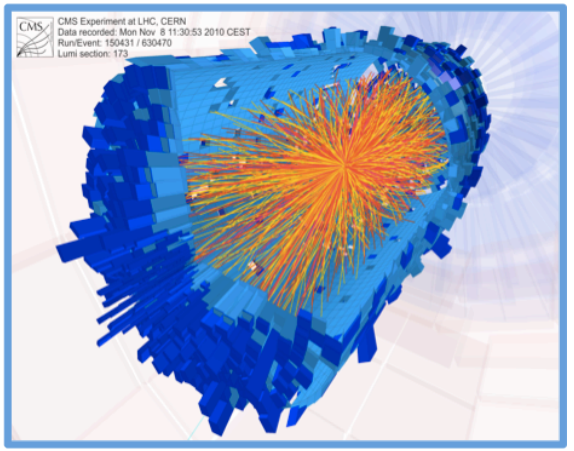


Pre-reaction **QGP** **Hadronization** **Detection**



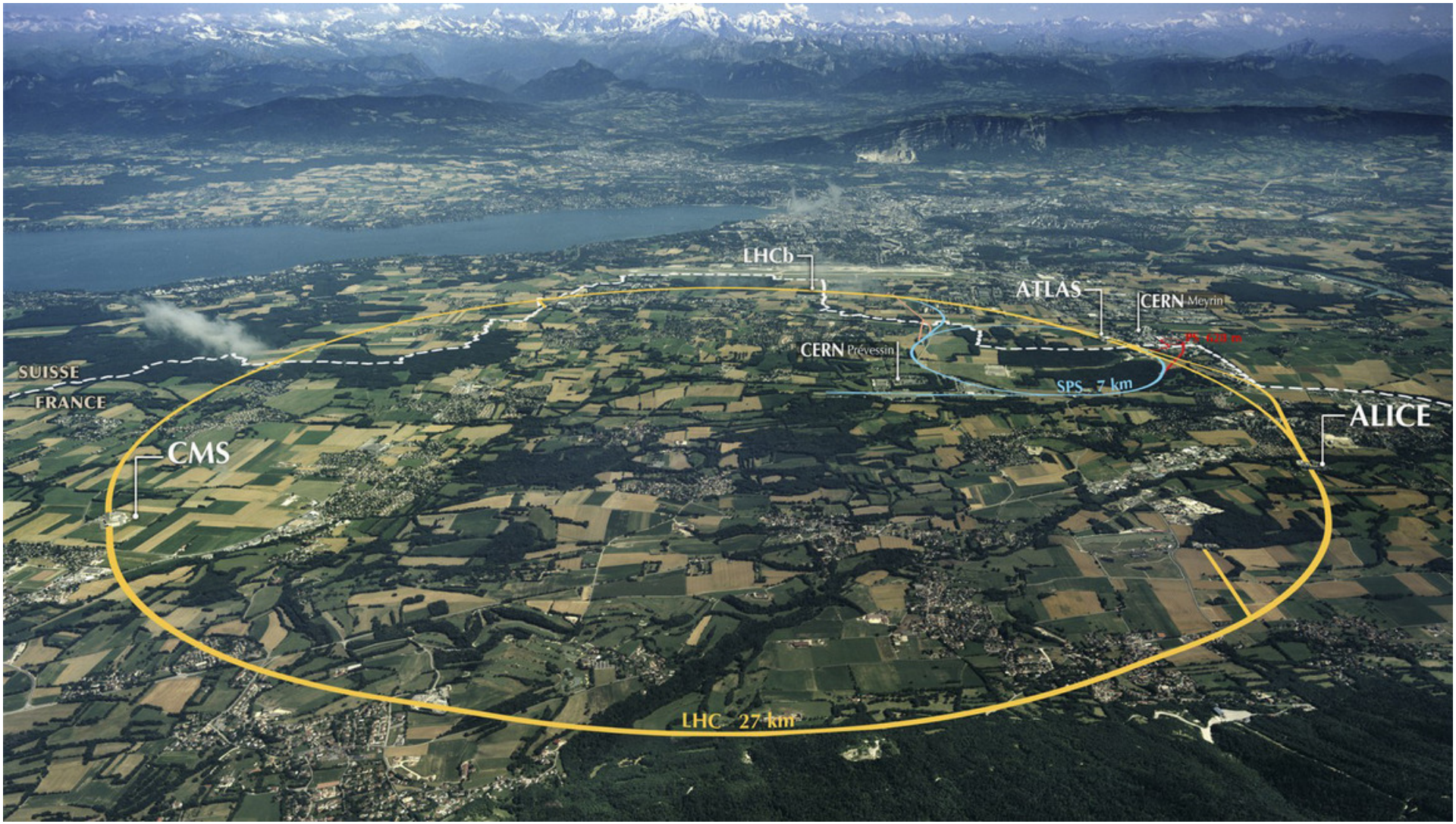




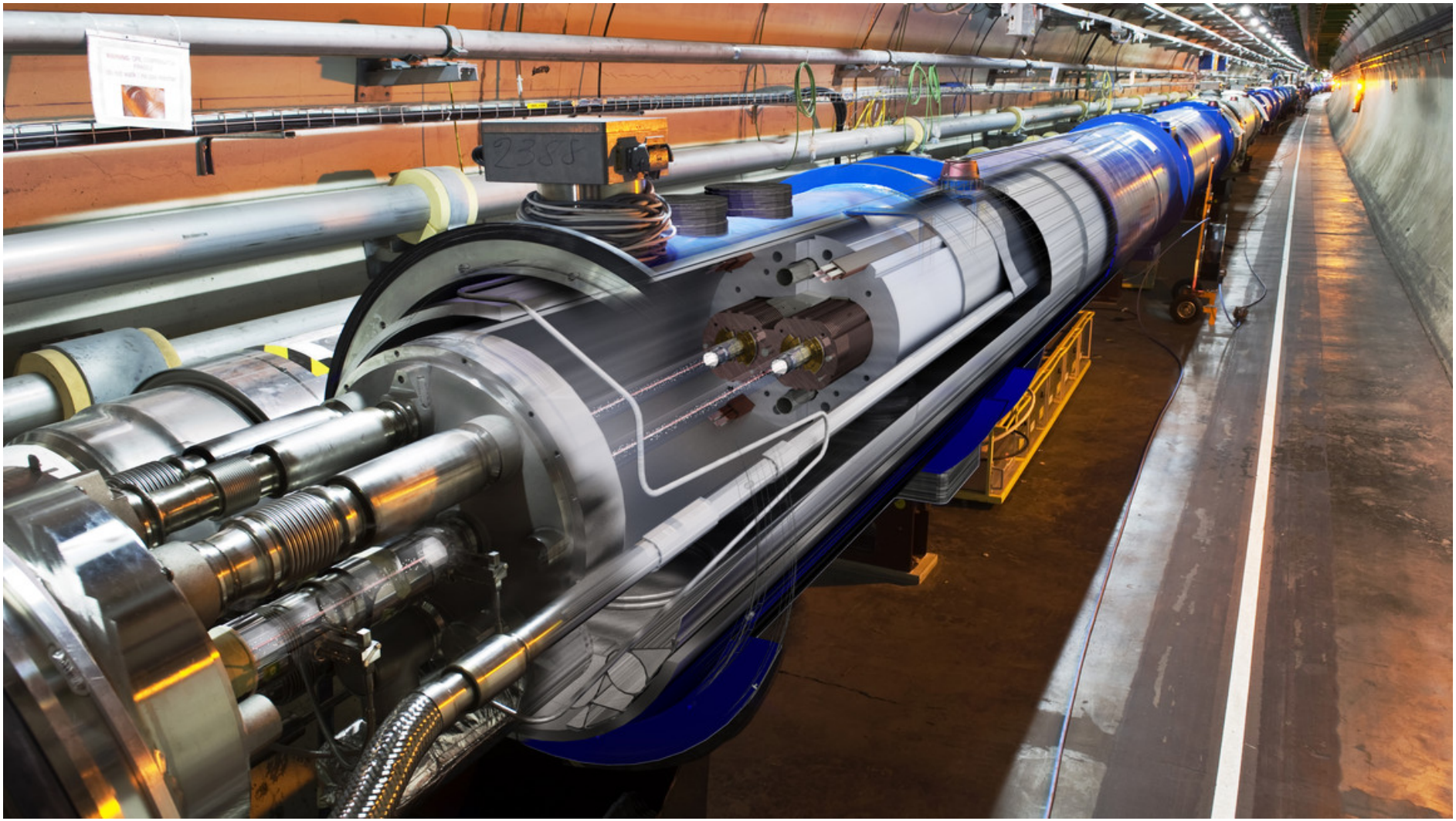


K. Dey

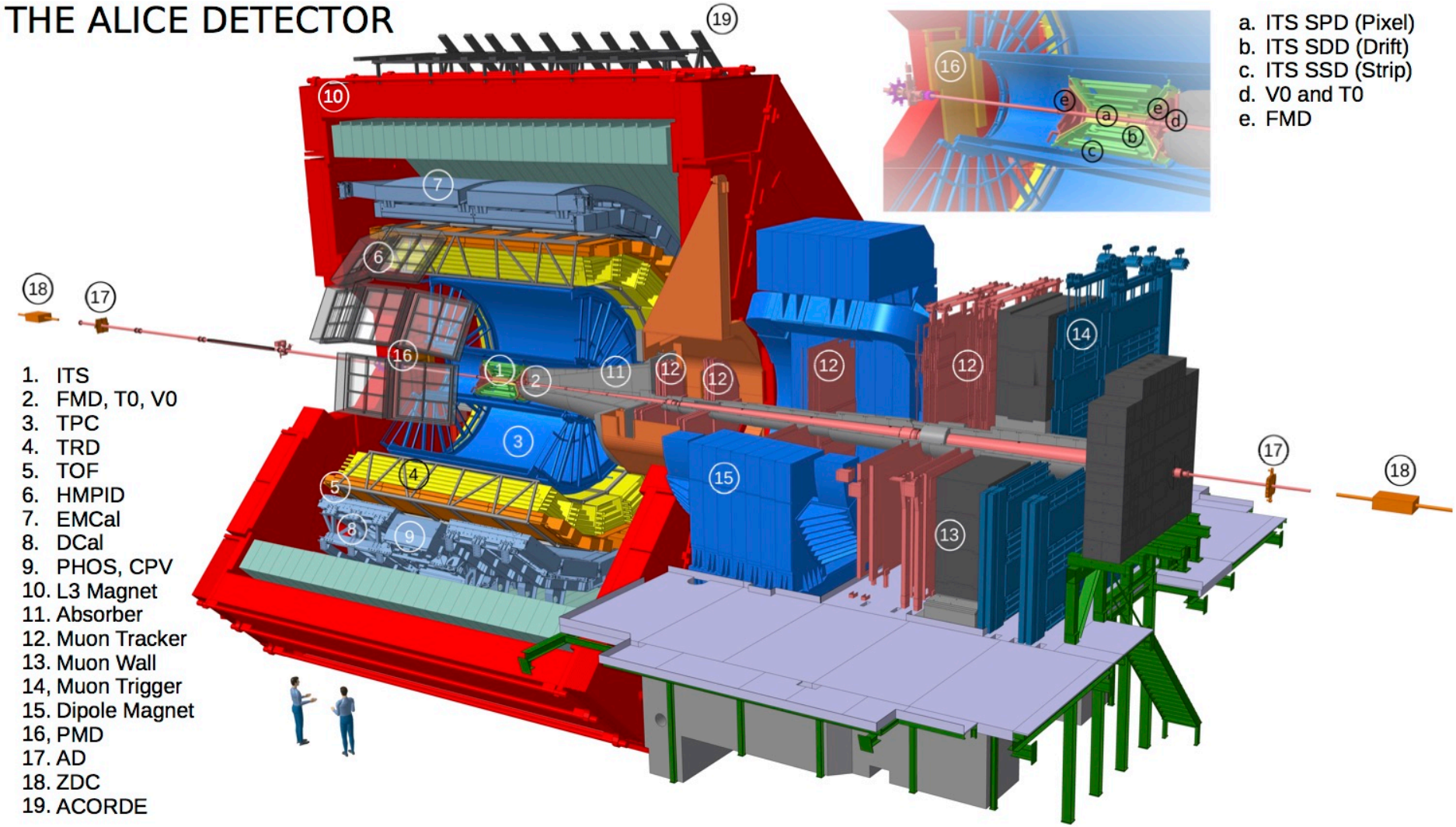
Large Hadron Collider at CERN



Large Hadron Collider at CERN



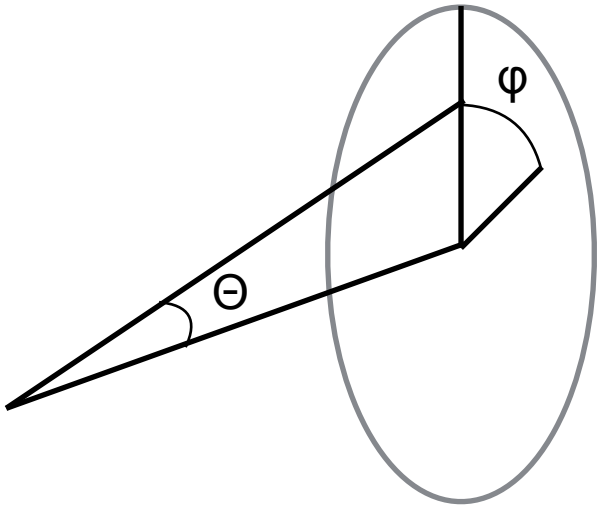
ALICE: A Large Ion Collider Experiment



ANN for QGP

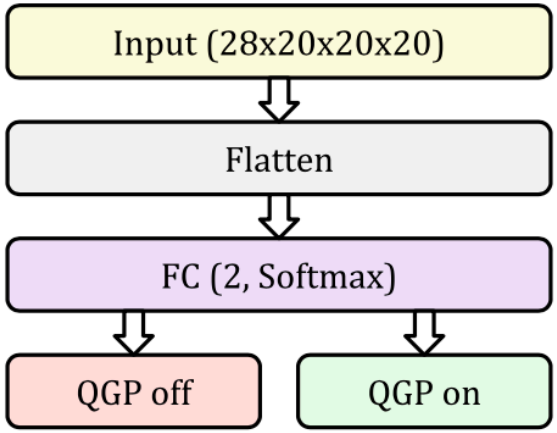
Input data:

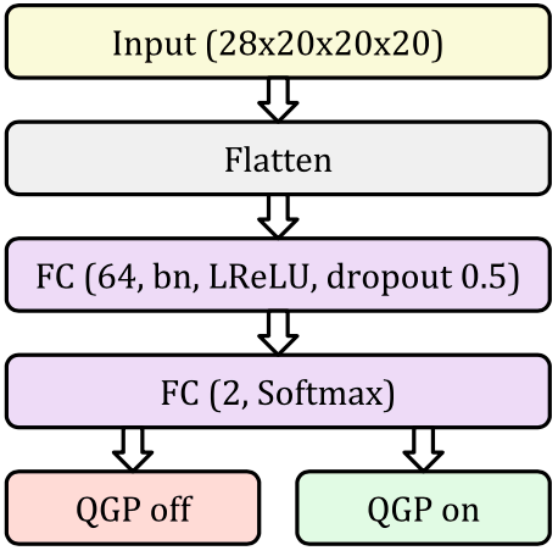
- 10000 simulated events (5000 with QGP and 5000 without QGP)
- central AuAu collisions at 31.2 GeV
- 1 event consist information about ~1500 particles
- particles are distributed into 4D matrix:
  - 28 bins - particle types;
  - 20 bins with log scale - particles momentum;
  - 20 bins - inclination angle Θ;
  - 20 bins - azimuth angle φ;
- 4D matrix is flatten into 1D array

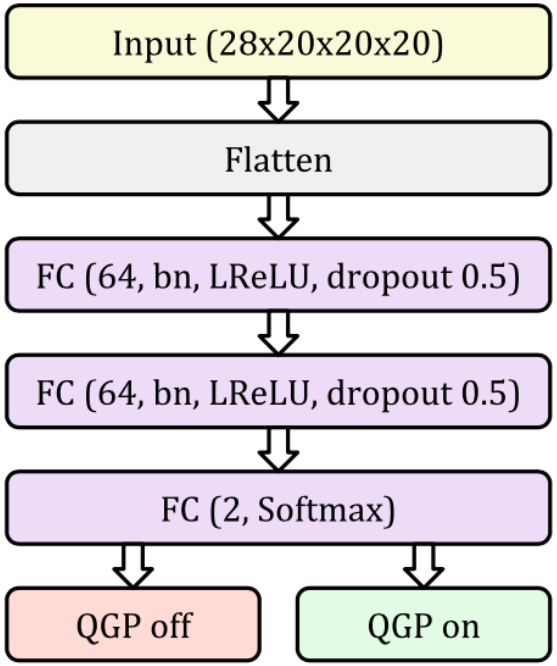


ANN input layer: *224000* neurons

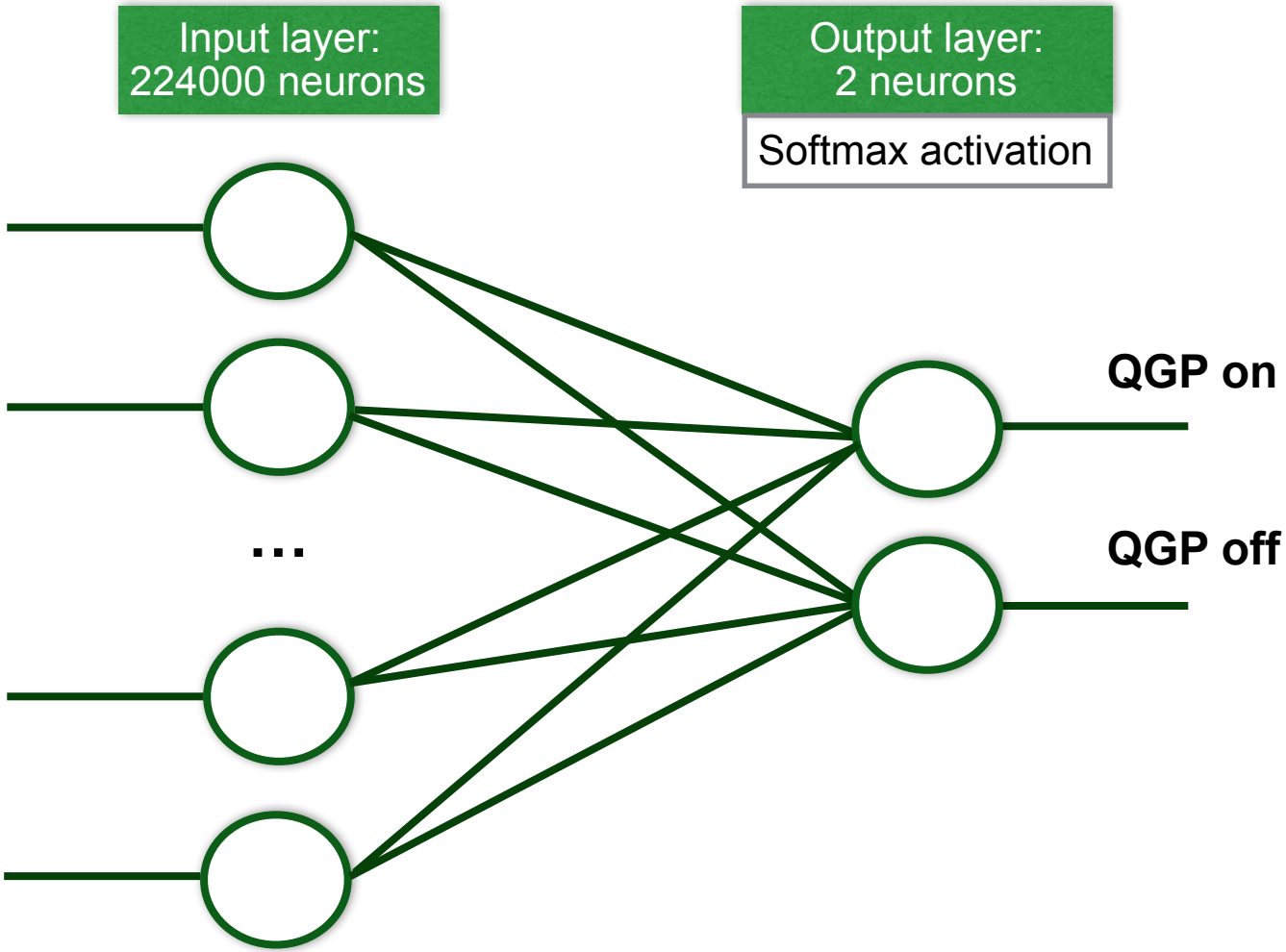
ANN for QGP


Input (28x20x20x20)
Flatten
FC (2, Softmax)
QGP off
QGP on

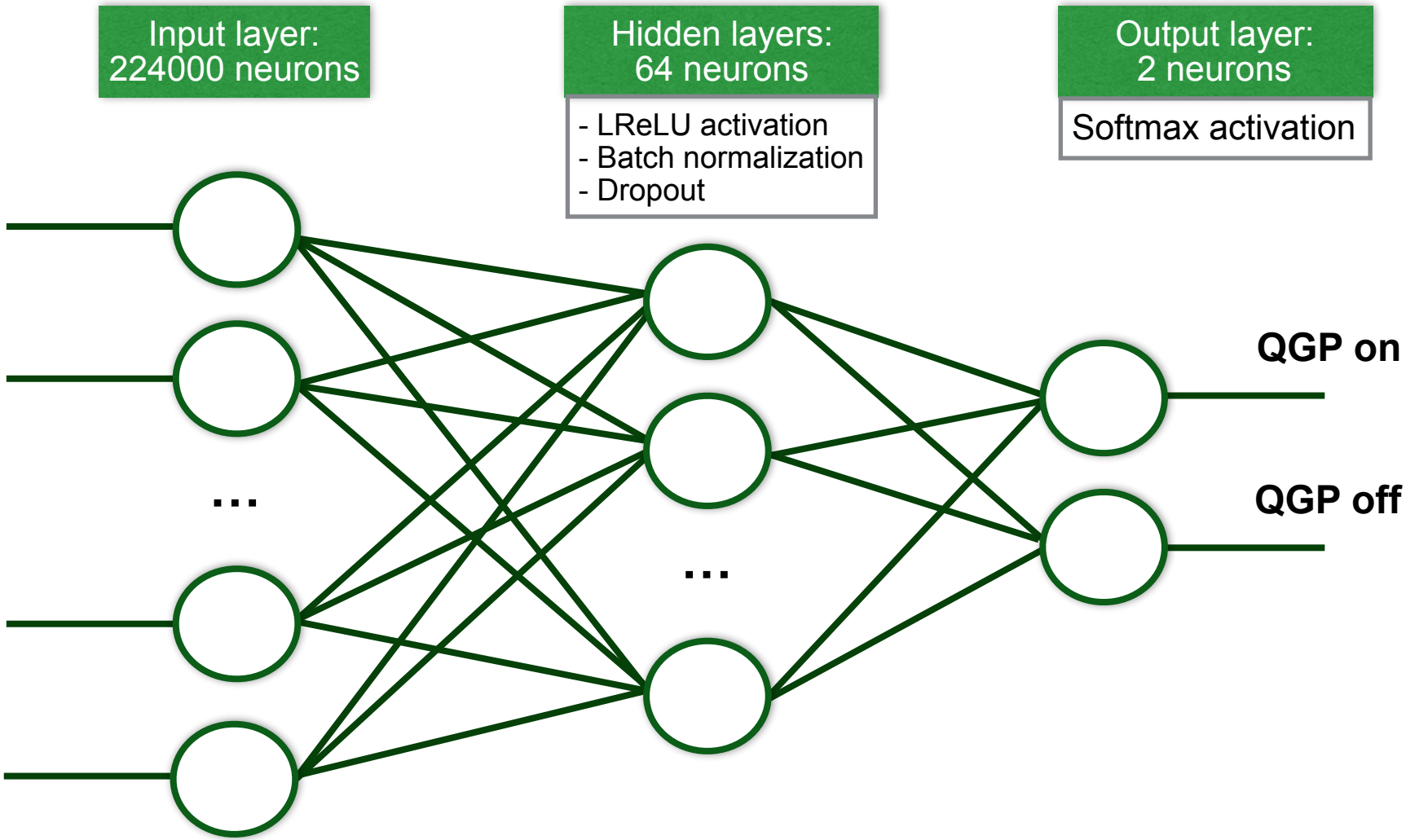

Input (28x20x20x20)
Flatten
FC (64, bn, LReLU, dropout 0.5)
FC (2, Softmax)
QGP off
QGP on


Input (28x20x20x20)
Flatten
FC (64, bn, LReLU, dropout 0.5)
FC (64, bn, LReLU, dropout 0.5)
FC (2, Softmax)
QGP off
QGP on

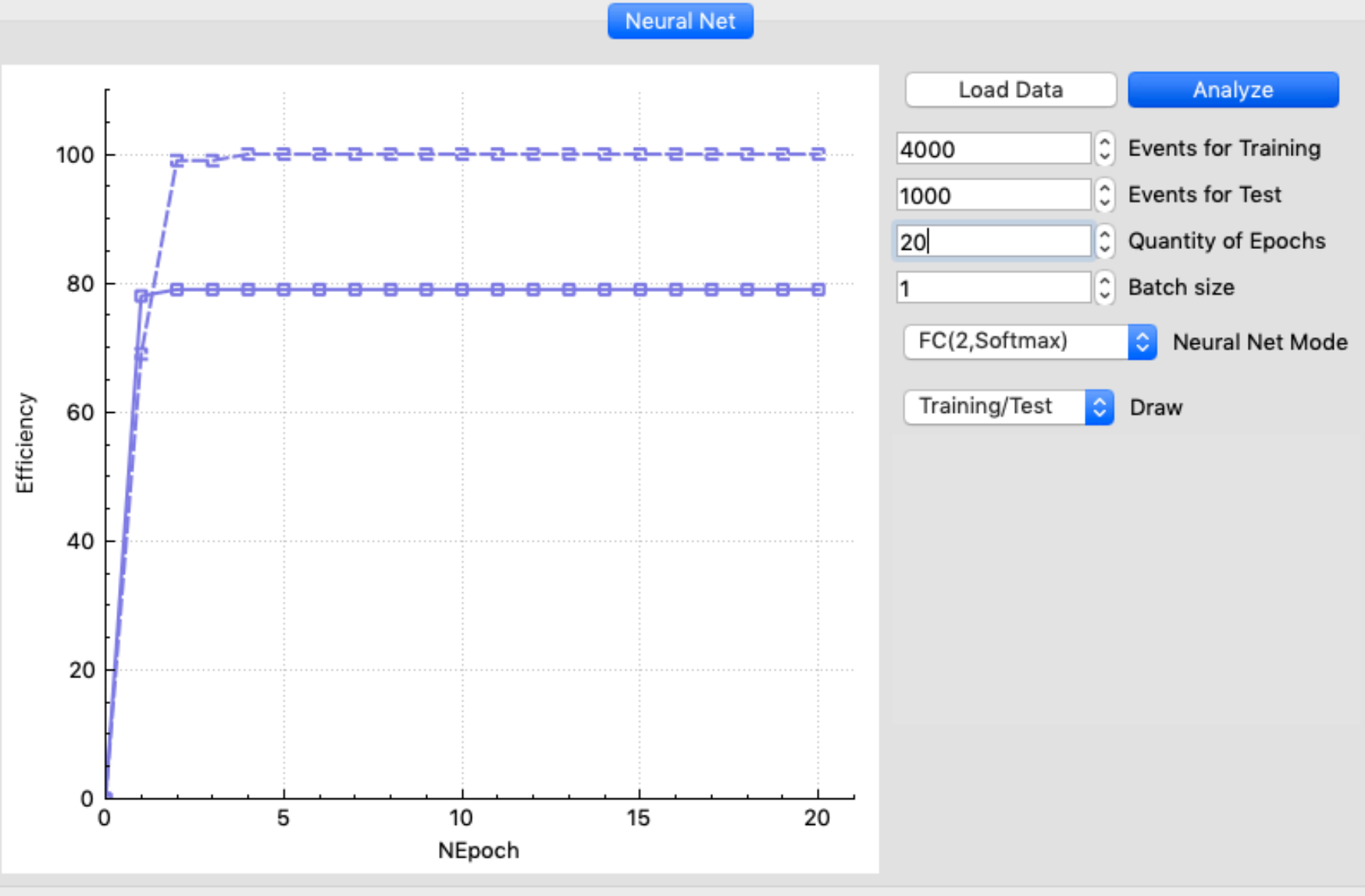
ANN for QGP


Input layer:
224000 neurons
Output layer:
2 neurons
Softmax activation
QGP on
QGP off

ANN for QGP


Input layer:
224000 neurons
Hidden layers:
64 neurons
- LReLU activation
- Batch normalization
- Dropout
Output layer:
2 neurons
Softmax activation
...
...
QGP on
QGP off

ANN for QGP


Neural Net
Load Data
Analyze
4000
Events for Training
1000
Events for Test
20
Quantity of Epochs
1
Batch size
FC(2,Softmax)
Neural Net Mode
Training/Test
Draw
Efficiency
NEpoch
0
20
40
60
80
100
0
5
10
15
20

Abgabe Milestone 2

11.12.2019 um 23:59 über OLAT

Prof. Dr. Ivan Kisel

Office 02/10
Giersch Science Center
Max-von-Laue-Straße 12
60438 Frankfurt am Main

I.Kisel [At] compeng.uni-frankfurt.de