

Neural Networks

- Homework 1 -

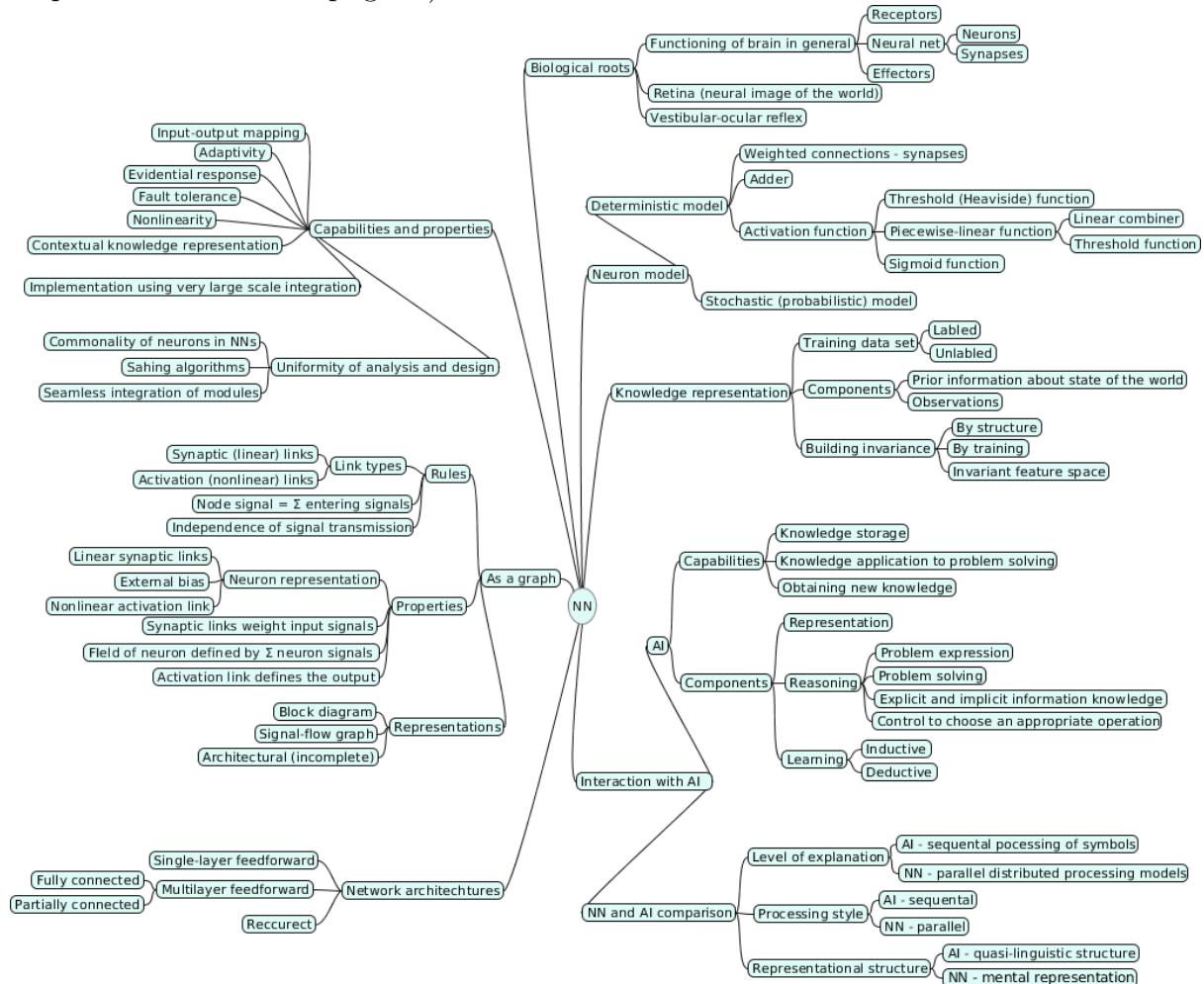
Petr Lukin, Ivan Vishniakou, Evgeniya Ovchinnikova

Lecture date: 26 September 2016

1 Mind map

Neural network (NN) is a parallel distributed system consisted of computing cells ("neurons") that is capable of storing and using experimental knowledge to model a brain in processes of certain tasks performance.

Figure 1: S. Haykin, Neural Networks, chapter 1. Mind map (a zoomed version of the map is attached as NN.png file).



2 Exercises

2.1 Exercise 1.1

$$\phi(v) = \frac{1}{1+exp(-av)}$$

Show that $\frac{d\phi(v)}{dv} = a\phi(v)[1 - \phi(v)]$. What is the value of this derivative at the origin?

Solution:

$$\frac{d\phi(v)}{dv} = \frac{d((1+exp(-av))^{-1})}{dv} = -\frac{-a\cdot exp(-av)}{(1+exp(-av))^2} = \frac{a(exp(-av)+1-1)}{(1+exp(-av))^2} = \frac{a}{(1+exp(-av))}(1 - \frac{1}{(1+exp(-av))}) = a\phi(v)[1 - \phi(v)]$$

$$\frac{a}{(1+exp(-a\cdot 0))}(1 - \frac{1}{(1+exp(-a\cdot 0))}) = \frac{a}{(1+1)}(1 - \frac{1}{(1+1)}) = \frac{a}{4}$$

2.2 Exercise 1.3

$$\phi(v) = \frac{v}{\sqrt{1+v^2}}$$

Show that $\frac{d\phi(v)}{dv} = \frac{\phi(v)^3}{v^3}$. What is the value of this derivative at the origin?

Solution:

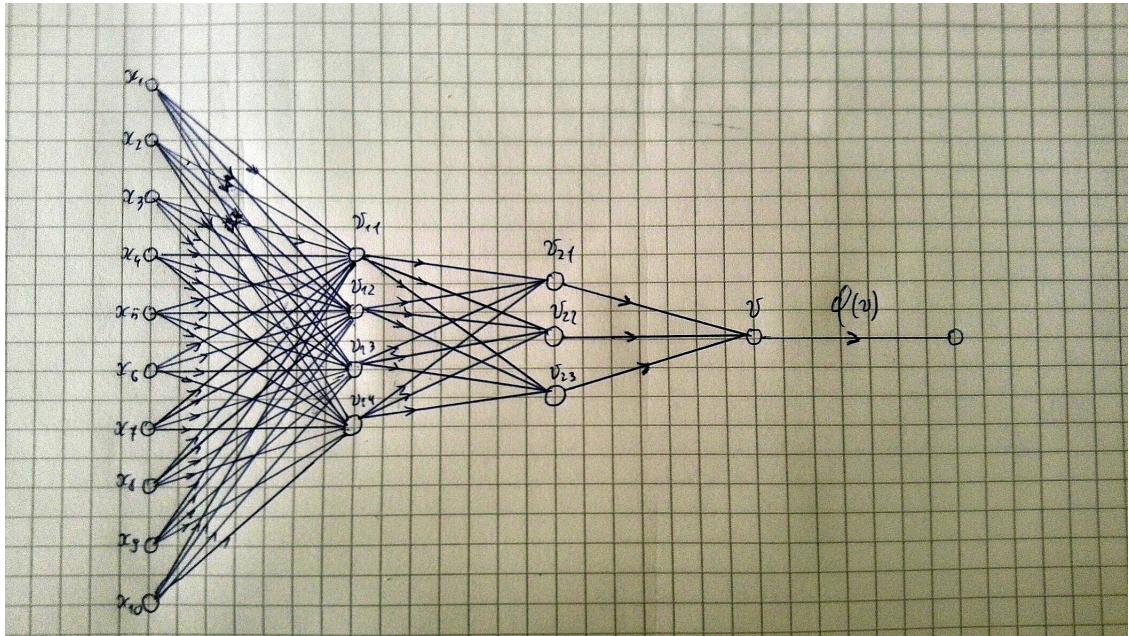
$$\frac{\phi(v)^3}{v^3} = \frac{1}{(\sqrt{1+v^2})^3}$$

$$\frac{d\phi(v)}{dv} = \frac{\sqrt{1+v^2}-v0.5(\sqrt{1+v^2})^{-1}2v}{(\sqrt{1+v^2})^2} = \frac{\sqrt{1+v^2}-v^2(\sqrt{1+v^2})^{-1}}{(\sqrt{1+v^2})^2} = \frac{1+v^2-v^2}{(\sqrt{1+v^2})^3} = \frac{1}{(\sqrt{1+v^2})^3}$$

$$\phi(0) = \frac{0}{\sqrt{1+0^2}} = 0$$

2.3 Exercise 1.12

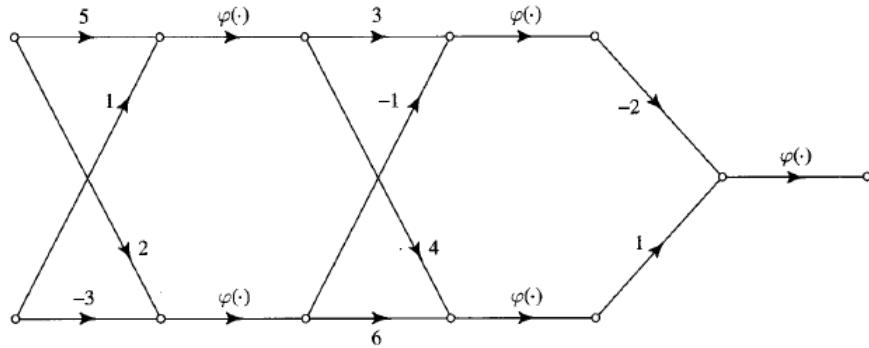
A fully connected feedforward network has 10 source nodes, 2 hidden layers, one with 4 neurons and the other with 3 neurons, and a single output neuron. Construct an architectural graph of this network.



2.4 Exercise 1.13

2.4.1 a)

Figure shows the signal flow graph of a 2-2-2-1 feedforward network. The function $\phi()$ denotes a logistic function. Write the input - output mapping defined by this network.



Solution:

Let input values be x_1, x_2 . Lets propagate this values through all hidden layers:

$$v_{1,1} = \phi(5x_1 + x_2)$$

$$v_{1,2} = \phi(2x_1 - 3x_2)$$

$$v_{2,1} = \phi(3v_{1,1} - v_{1,2})$$

$$v_{2,2} = \phi(4v_{1,1} + 6v_{1,2})$$

$$v = \phi(-2v_{2,1} + v_{2,2})$$

Lets unfold the previous equations:

$$v = \phi((\phi(3(\phi(5x_1 + x_2)) - (\phi(2x_1 - 3x_2)))) + (\phi(4(\phi(5x_1 + x_2)) + 6(\phi(2x_1 - 3x_2)))))$$

2.4.2 b)

Suppose that the output neuron in the signal flow graph operates in its linear region. Write the input-output mapping defined by this new network.

If the output neuron operates in its linear region sigmoid function can be substituted with linear function. Thus, sigmoid function will not be applied in the end:

$$v = \phi(3(\phi(5x_1 + x_2)) - (\phi(2x_1 - 3x_2))) + \phi(4(\phi(5x_1 + x_2)) + 6(\phi(2x_1 - 3x_2)))$$

2.5 Exercise 1.21

Let x be an input vector and $s(\alpha, x)$ be a transformation operator acting on x and depending on some parameter α . The operator $s(\alpha, x)$ satisfies:

- $s(0, x) = x$
- $s(\alpha, x)$ is differential with respect to α

How would you compute the tangent vector when α is small? The tangent vector is locally invariant with respect to rotation of the original image, why?

Solution:

Since α is small, we can use a Taylor expansion: $s(\alpha, x) = s(0, x) + \alpha \frac{\delta s(\alpha, x)}{\delta \alpha} + O(\alpha^2) \approx x + \alpha \frac{\delta s(\alpha, x)}{\delta \alpha}$

According to the assignment, the tangent vector (V) is defined by the partial derivative $\frac{\delta s(\alpha, x)}{\delta \alpha}$, therefore:

$$s(\alpha, x) = x + \alpha V$$

$$V = \frac{s(\alpha, x) - x}{\alpha}$$