Exam in
Neural Networks and Learning Systems
TBMI26 / 732A55
Home exam - Part II

Date:	2020-03-20
Time:	14.00 - 15.30 (part 1) and **15.30 - 18.00** (part 2)
Teacher:	Magnus Borga, Phone: 013-286777

Read the instructions before answering the questions!

Part 2 Consists of four 5-point questions. These questions test deeper understanding and the ability to apply the knowledge to solve problems. All assumptions and calculations made should be presented. Reasonable simplifications may be done in the calculations. **This part needs to be submitted before 18:00**

You can either edit this PDF in a PDF editor or write your answers by hand and then scan it using a scanner or mobile phone, or write the answers in a separate file using a word processor. The answers may be given in English or Swedish. **If you write by hand, please write clearly using block letters! (Do not use cursive writing.) Answers that are difficult to read, will be dismissed.** The exam should be submitted before the deadline in PDF format. The PDF files should be named with your LiU-ID followed by a the number of the part of the exam, e.g. "abcde132-2".

The maximum sum of points is 20 on each part. To pass the exam (grade 3/E) at least 15 points are required on part 1. For grade 4/C, an additional 10 points on part 2 are required and for grade 5/A, 15 points are required on part 2, in addition to pass part 1.

Note that all forms of collaboration or communication with any person except the course staff is strictly forbidden during the exam!

The result will be reported at 2020-04-03 at the latest. The exams will then be available at "studerandeexpeditionen" at IMT.

GOOD LUCK!

AID:	Exam Date: 2020-03-20
Course Code: TBMI26 / 732A55	Exam Code: TEN1

1. The correlation of a 1D feature map channel $f_i(x)$ with a kernel $h_{ij}(x)$ of size $K = 2L+1$ is defined

$$g_{ij}(x) = (h_{ij} \square f_i)(x) = \sum_{\alpha=-L}^{L} h_{ij}(\alpha) f_i(x+\alpha) \ .$$

We now look at the example of a two-channel input feature map ($i = 0, 1$) and a two-channel output feature map ($j = 0, 1$), i.e., we use four kernels in total. The final output is computed as $y_j(x) = \sum_i g_{ij}(x)$.

a) 2p Perform the four convolutions below. All values outside the feature map channels f_i are equal to zero. In the arrays f_i and h_{ij}, the respective number written in bold face is at position $x = 0$. Note that g_{ij} is only a part of the convolution result (*'same'*).

h_{ij}			$\square$	f_i				$=$	g_{ij}				
2	**2**	1	$\square$	**0**	1	0	1	$=$	**?**	?	?	?	$i=0, j=0$
0	**1**	3	$\square$	**0**	1	0	1	$=$	**?**	?	?	?	$i=0, j=1$
2	**2**	0	$\square$	**2**	0	1	0	$=$	**?**	?	?	?	$i=1, j=0$
1	**2**	1	$\square$	**2**	0	1	0	$=$	**?**	?	?	?	$i=1, j=1$

b) 2p We now assume I input channels, J output channels, and kernels of size K.
i) How many parameters need to be learned (no bias coefficient)? Tick the box for the correct answer
☐ $(I+J)\cdot K$ ☐ $I\cdot J+K$ ☐ $I\cdot J\cdot K$ ☐ $I\cdot J\cdot(K-1)$
ii) How many calculations (multiplications or additions) have to be performed for each position x in the output $(y_0(x),\ldots,y_{J-1}(x))$ (neglecting boundaries)?
☐ $J+2\cdot K\cdot I$ ☐ $J\cdot(2\cdot K\cdot I-1)$ ☐ $J\cdot K\cdot I$ ☐ $J\cdot(K\cdot I-1)$

c) 1p Now we consider 2D correlation. Look at the kernels and the images.

-1	0	1
-2	0	2
-1	0	1

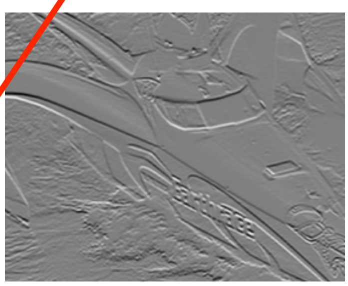
1)

-1	-2	-1
0	0	0
1	2	1

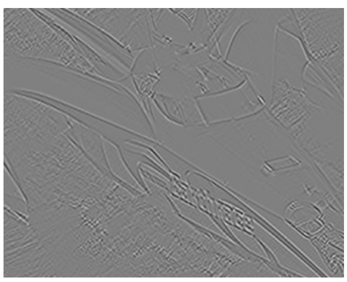
2)

1	0	-1
0	0	0
-1	0	1

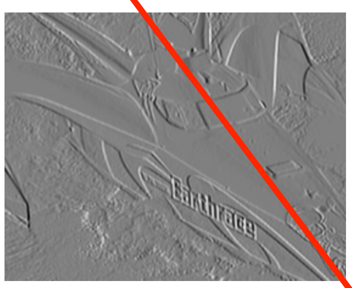
3)



a)



b)



c)

AID:	Exam Date: 2020-03-20
Course Code: TBMI26 / 732A55	Exam Code: TEN1

Which of the three kernels 1)-3) generated which of the three outputs a)-c)?

1:__

2:__

3:__

AID:	Exam Date: 2020-03-20
Course Code: TBMI26 / 732A55	Exam Code: TEN1

2. Draw figures of the following networks detailing the input, outputs and operation of each layer. Use the figures to derive the weight update rules for online training with the square error as cost function. If needed you can assume that the input consists of two features. Both component form and matrix form are accepted. Make sure to explain all variables and indexes you introduce.

   a) Two output neurons with linear activation functions and a hidden layer with two neurons using $tanh$ as activation functions. (3p)

   b) Two output neurons with linear activation functions and two hidden layers with two neurons each using $tanh$ as activation functions. (2p)

$\frac{\partial}{\partial x} tan(x) = 1 - tan^2(x)$

3. Assume we have a signal

$$\mathbf{x}_t = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$$

as given in the table below.

t	1	2	3	4	5
$x_1(t)$	3	8	2	7	5
$x_2(t)$	0	0	2	-1	-1

a) We think the data volume is too large. Show how the data volume can be reduced to half the size by using principal component analysis as a tool for dimensionality reduction. Also, state the resulting data volume after the reduction. (3p)

b) How much of the information (total variance) in the data is accounted for by the first principle direction? (1p)

c) Show how the original signal can be approximately reconstructed given the results from a). (1p)

Hint: The eigenvalues of a 2×2-matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are:

$$\lambda_1 = \frac{T}{2} + \sqrt{\frac{T^2}{4} - D} \text{ and } \lambda_2 = \frac{T}{2} - \sqrt{\frac{T^2}{4} - D},$$

with trace $T = a + d$ and determinant $D = ad - bc$. The corresponding eigenvectors are

$$\mathbf{e}_1 = \begin{pmatrix} \lambda_1 - d \\ c \end{pmatrix} \text{ and } \mathbf{e}_2 = \begin{pmatrix} \lambda_2 - d \\ c \end{pmatrix}$$

AID:	Exam Date: 2020-03-20
Course Code: TBMI26 / 732A55	Exam Code: TEN1

4. The figure below shows two different state models and the corresponding reward function. The states are numbered and the possible actions are represented by arrows. The number next to the arrow designates the reward for taking that action. If the system has reached an "end" state no more rewards are given.

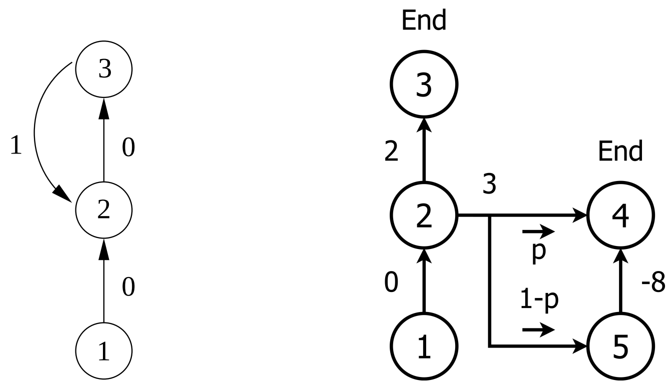


Figure 1: State models A and B.

a) Calculate the optimal Q- and V-functions for system A as a function of $0 < \gamma < 1$. (2p)

b) In system B, the action to go right in state 2 will result in the state randomly changing to state 4 or 5, where the arrows p and 1-p indicate the probability of the respective path. Regardless of the random state change the reward for the action is 3. Calculate the optimal Q- and V-functions for system B as a function of $0 < \gamma < 1$ and $0 < p < 1$. (2p)

c) In system B, given $\gamma = 0.5$, for what values of p will the path $1 \to 2 \to 3$ be the optimal policy from state 1? (1p)