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

Date: 2020-06-10
Time: **14.00 - 16.00** (part 1) and 15.30 - 18.00 (part 2)
Teacher: Magnus Borga, Phone: 013-286777

Read the instructions before answering the questions!

The full exam consists of two parts:

Part 1 Consists of ten 1-point and five 2-point questions. The questions test general knowledge and basic understanding of central concepts in the course. The answers should be short and given on the blank space after each question or in the indicated figure. **Note that this part needs to be submitted no later than 16:00!**

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 will be published at 15:30.**

Write your answers by hand and then scan them 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. Each part should be handed in as one single PDF file. 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/C) at least 13 points are required on part 1. For grade 4/B, 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 2020-07-01 at the latest. The exams will then be available at "studerandeexpeditionen" at IMT.

GOOD LUCK!

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

One-point questions

1. Which of the following methods are supervised learning methods:

   - PCA (Principal Component Analysis)
   - Back-propagation
   - k-means
   - k-NN
   - LDA (Linear Discriminant Analysis)
   - SVN (Support Vector Machines)

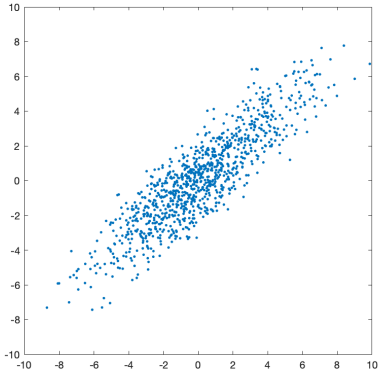
2. Write the cost function that is being optimized in Support Vector Machines

3. Which of these functions can be used in the hidden layers of a back-prop network?

   - $y = \frac{1}{e^{-s}+1}$
   - $y = \begin{cases} 0, \text{for } s \leq 0 \\ 1, \text{for } s > 0 \end{cases}$
   - $y = \begin{cases} 0, \text{for } s \leq 0 \\ s, \text{for } s > 0 \end{cases}$
   - $y = e^{(-s^2)}$

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

4. Draw the first principal component of the distribution in the figure below!



5. We want to train a perceptron to separate the two classes in the figure below. How many parameters w_i do we need to optimize?

6. Assume you have a set of data points in a two-dimensional feature space and you want to cluster the points into three clusters using Mixture of Gaussians. How many scalar numbers need to be estimated in the optimization, disregarding the set membership variables S_i?

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

7. Mention an alternative to the sigmoid activation function that can be used to avoid the vanishing gradient problem in deep neural networks

8. Why isn't it optimal to try to reach the lowest possible error on the training data when traing a neural network?

9. What is determined by the parameter k in k-NN?

10. Why isn't it always best to choose the response with the best Q-value in Q-learning?

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

Two-point questions

11. Consider the following non-linear mapping of the input data $\mathbf{x}$:

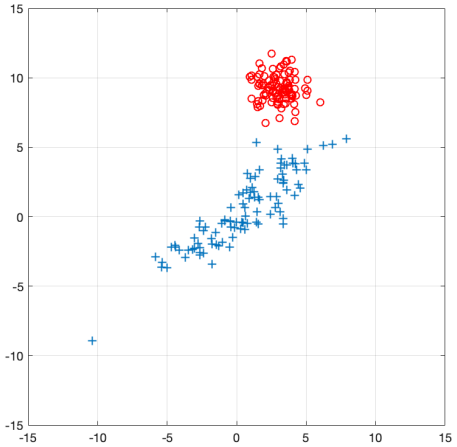
$$\begin{aligned}\varphi_1(\mathbf{x}) &= x_1^2 \\ \varphi_2(\mathbf{x}) &= x_2^2 \\ \varphi_3(\mathbf{x}) &= \sqrt{2} \cdot x_1 x_2\end{aligned}$$

You want to analyse this data with a kernel method. How is the scalar product $\varphi(\mathbf{x})^T \varphi(\mathbf{y})$ expressed in the input data space?

12. Perform two iterations with k-means in the figure below. The dots indicate the data points and the crosses the two prototypes (k=2). Show both steps. Has the algoritm converged after these steps?

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

13. Assume that a one-layer neural network is trained on the data in the figure below. Draw the approximate zero-crossing of the resulting discriminant function if (a) a linear activation is used and (b) a sigmoid activation function is used.



14. Draw a decision tree that implements the following discriminant function:

$$y = \begin{cases} -1, \text{ for } (x_1 < 0) \text{ AND } (x_2 < 0) \\ 1, \text{else} \end{cases}$$

AID:	Exam Date: 2020-06-10
Course Code: TBMI26 / 732A55	Exam Code: TEN1

15. The table below shows the Q-values for different states S_i and actions A_i. What are the value function values $V(S_i)$ for all states and what action will the system take in state 2 if it follows a greedy policy?

Q(S,A)	S_1	S_2	S_3	S_4
A_1	3	4	3	2
A_2	4	3	2	1
A_3	2	3	2	1