Medical Data Science, WS 2021/2022
Prof. Dr. Nico Pfeifer
Chair for Methods in Medical Informatics
University of Tuebingen

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

2022-01-11

Assignment 5

Deadline: Tuesday, January 25, 9:59 p.m.

This problem set is worth 50 points. You can submit in groups of two people or alone. Submit your solutions digitally by uploading to the ILIAS page (none of the other students can see the files you upload). Just upload a zipped folder containing all necessary files and name the folder by your last names. The folder should be named according to the following scheme:

`[MDS][Assignment5]_lastname1_lastname2`

Problem 1 (T, 21P)

Graphs and Deep Learning

(a) (3P) What are Cliques in graphs? What are Maximal Cliques as defined in the lecture? Draw a graph where every node is part of at least one Maximal Clique (how large is your Maximal Clique?) and another graph that contains a Maximal Clique of size 4, i.e. a Clique that includes 4 nodes.

(b) (1P) How can you transform a directed into an undirected network (without information loss)?

(c) (2P) Explain briefly how Inference on a Chain can be performed. What is Z and why do we need it to get $p(x_n)$?

(d) (2P) What is the connection between Cliques and factors in Factor Graphs?

(e) (2P) In which sense is the Max-Sum algorithm an extension of the Sum-Product Algorithm?

(f) (2P) What are Gaussian Graphical Models? What is the connection between the inverse of the covariance matrix Θ and the edges of the graph? What does a zero-entry in the inverse covariance matrix ($\Theta_{i,j} = 0$) imply?

(g) (3P) Many methods and algroithms in the field of deep learning are much older than their success. What made deep learning so popular during the last few years? Name improvements that made deep learning more applicable and also name some of their current applications.

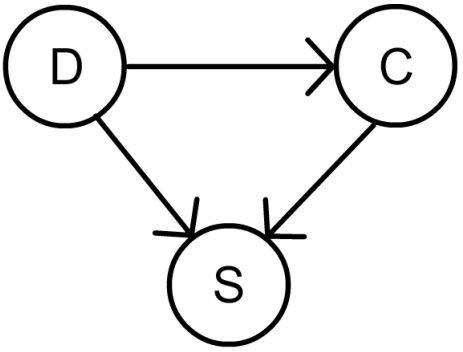
(h) (2P) What is sparse coding? What can it be used for?

(i) (2P) Explain briefly encoders and decoders in Autoencoders. Why are they used instead of using the input image directly? How can identity-mapping be prevented?

(j) (2P) What are activation functions in Neural Networks? Provide an example.

Problem 2 (T, 17 Points)

Assume the following causal DAG,



showing the hypothetical scenario of a surgery for which the success depends on whether the surgery was performed during the day or at night and whether the chief surgeon performed the surgery or not. We have the following random variables d (true if day-time, false if night), c (true if chief surgeon performs surgery, false otherwise), and s (true, if surgery successful, false otherwise).

The joint distribution is given in Table 1.

d	c	s	$p(d, c, s)$
0	0	0	0.06
0	0	1	0.13
0	1	0	0.04
0	1	1	0.23
1	0	0	0.06
1	0	1	0.17
1	1	0	0.04
1	1	1	0.27

Table 1: The joint distribution over three binary variables.

(a) (2 points): Provide p(c).

(b) (2 points): Provide p(d).

(c) (2 points): Provide p(d, c).

(d) (2 points): Provide p(s|d, c).

(e) (9 points): Show how the joint distribution changes if we decide whether the chief surgeon performs the surgery or not (use the do calculus).

Problem 3 (P, 12 Points)

Go to https://playground.tensorflow.org/ and solve the following tasks:

(a) (3 points): Try to find an architecture to achieve less than 0.05 test error on the circle data set (also submit a screen shot of the final result showing data, features, hidden layers and output).

(b) (3 points): Describe how the decision surface is learnt (i.e., which features are combined and how are they combined), especially taking into account the last layer.

(c) (3 points): Try to find an architecture to achieve less than 0.05 test error on the spiral data set (also submit a screen shot of the final result showing data, features, hidden layers and output).

(d) (3 points): Describe how the decision surface is learnt (i.e., which features are combined and how are they combined), especially taking into account the last layer.