EBERHARD KARLS UNIVERSITÄT TÜBINGEN

2024-01-15

Assignment 5

Deadline: Tuesday, January 30, 7:59 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. 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, 19P)

Graphs and Deep Learning

(a) (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.

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

(c) (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?

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

(e) (2P) What is represented in a 2-D contact map (in the context of proteins)?

(f) (2P) How can amino acid sequences be encoded?

(g) (2P) Explain how the *"L/5 top scored predicted contact pairs"* are used as a performance measure.

(h) (2P) How can dependence between two random variables be measured?

(i) (2P) Why is max-pooling used in Neural Networks? Is it a procedure with or without information loss?

Problem 2 (T, 17 Points)

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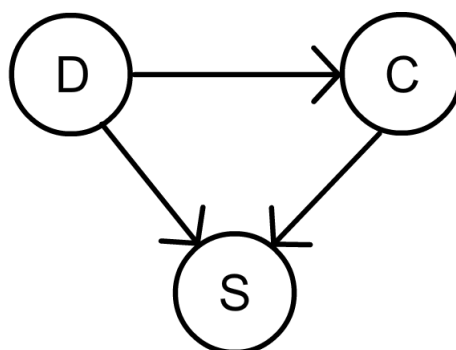
showing the hypothetical scenario of a surgery for which the success depends on whether the surgery was performed during the day or at night and whether the chief surgeon performed the surgery or not. We have the following random variables d (true if day-time, false if night), c (true if chief surgeon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

Problem 3 (P, 14 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) (4 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) (4 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.