Data Analytics and Machine Learning Group
Department of Informatics
Technical University of Munich

Eexam
Place student sticker here

Note:

- During the attendance check a sticker containing a unique code will be put on this exam.
- This code contains a unique number that associates this exam with your registration number.
- This number is printed both next to the code and to the signature field in the attendance check list.

Machine Learning

Exam:	IN2064 / EndtermPractice	**Date:**	Monday 25th January, 2021
Examiner:	Prof. Dr. Stephan Günnemann	**Time:**	17:00 – 19:00

Working instructions

- This document only contains the problem statements.
- **You do NOT have to upload this document to TUMexam!**
- **Make sure that you answer the correct version of each problem (version A, B, C or D)!**
- If you answer the wrong version (e.g., your sheet says Version B but you solve Version C), you will receive 0 points for this problem.

Left room from ________ to ________	/	Early submission at ________

Problem 1 (Version A) (3 credits)

We consider the following sequences of random variables $U_0, U_1, \dots, U_t$.

0 a) $U_t = \begin{bmatrix} X_t \\ Z_t \end{bmatrix}$ where X_t are observed variables and Z_t are latent variables of an Hidden Markov Model. Does the
1 sequence of variables U_t fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0 b) We consider an AR(p) process X_t. Under what condition on p and k does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.
1

0 c) We consider a recurrent neural network which produces X_t. Does the sequence of variables $U_t = X_t$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.
1

Problem 1 (Version B) (3 credits)

We consider the following sequences of random variables $U_0, U_1, \dots, U_t$.

a) $U_t = \begin{bmatrix} X_t \\ Z_t \end{bmatrix}$ where X_t are observed variables and Z_t are latent variables of an Hidden Markov Model. Does the sequence of variables U_t fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0

1

b) We consider an AR(p) process X_t. Under what condition on p and k does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0

1

c) We consider a neural network with dilated convolutions (i.e. WaveNet) with n layers which produces X_t. Under what condition on n and k does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property ? Justify your answer.

0

1

Problem 1 (Version C) **(3 credits)**

We consider the following sequences of random variables $U_0, U_1, \dots, U_t$.

0 a) $U_t = \begin{bmatrix} Z_{t-2} \\ Z_t \end{bmatrix}$ where X_t are observed variables and Z_t are latent variables of an Hidden Markov Model. Does the
1 sequence of variables U_t fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0 b) We consider an AR(p) process X_t. Under what condition on p and k does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.
1

0 c) We consider a recurrent neural network which produces X_t. Does the sequence of variables $U_t = X_t$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.
1

Problem 1 (Version D) (3 credits)

We consider the following sequences of random variables $U_0, U_1, \ldots, U_t$.

a) $U_t = \begin{bmatrix} Z_{t-2} \\ Z_t \end{bmatrix}$ where X_t are observed variables and Z_t are latent variables of an Hidden Markov Model. Does the sequence of variables U_t fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0

1

b) We consider an AR(p) process X_t. Under what condition on p and k does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property i.e. $P(U_t|U_{t-1}) = P(U_t|U_{t-1}, ..., U_0)$? Justify your answer.

0

1

c) We consider a causal convolution neural network (e.g. WaveNet) with a time window of size w which produces X_t. Under what condition on w does the sequence of variables $U_t = [X_{t-1}, ..., X_{t-k}]$ fulfill the Markov property ? Justify your answer.

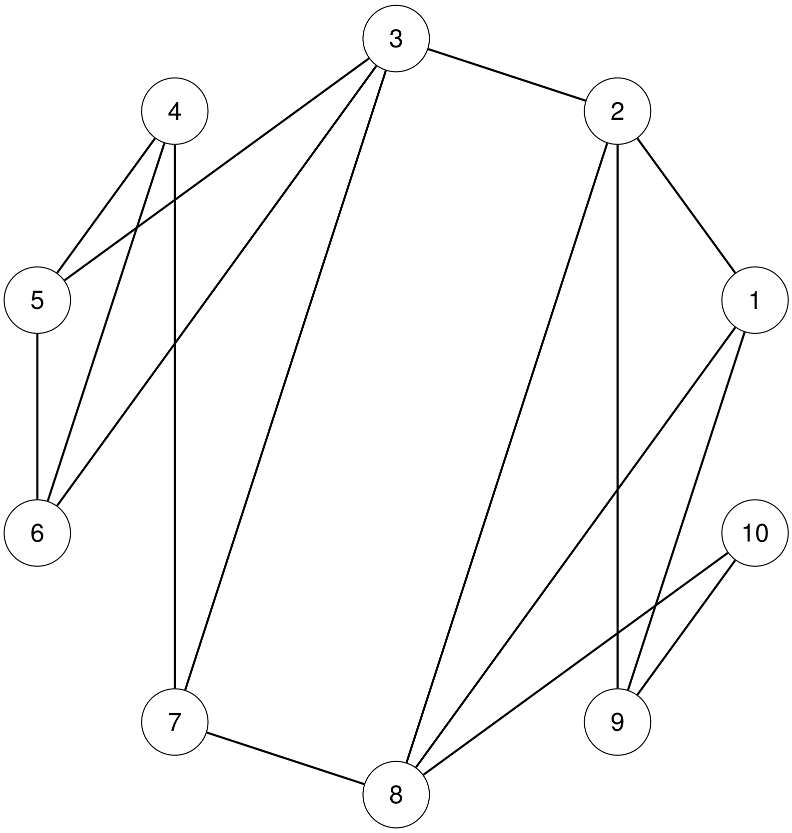
0

1

Problem 2 (Version A) (4 credits)

0
1
2
3
4

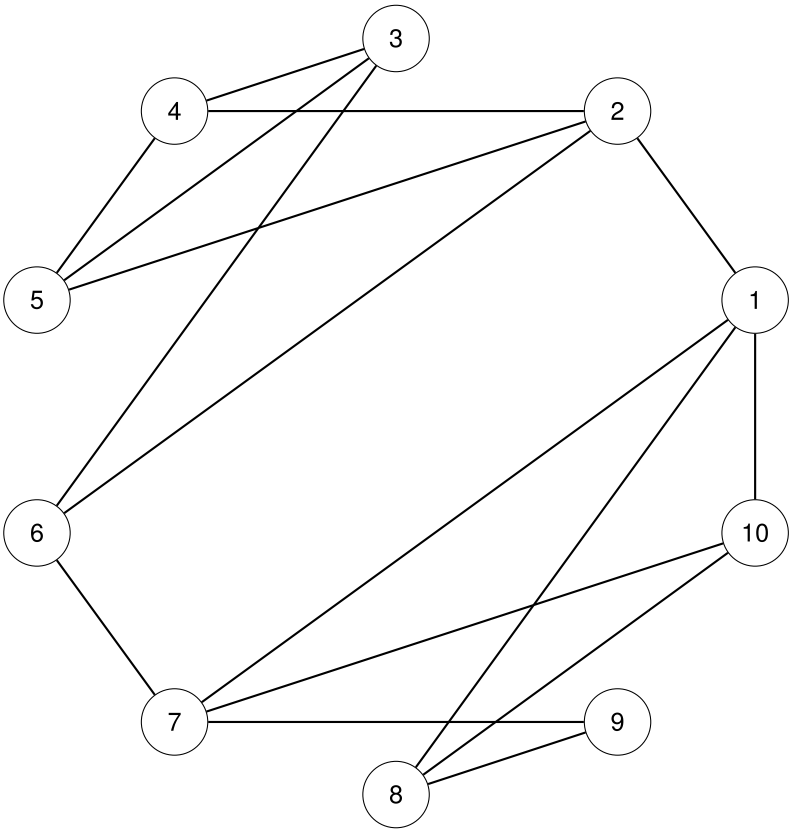
The following graph has been generated from a planted partition model with in-community edge probability p and between-community edge probability q.



Assuming $p < q$, find the maximum likelihood community assignments under a PPM.
Give your solution as two sets of node labels making up the two discovered communities. Justify your answer.

Problem 2 (Version B) **(4 credits)**

The following graph has been generated from a planted partition model with in-community edge probability p and between-community edge probability q.



Assuming $p < q$, find the maximum likelihood community assignments under a PPM.
Give your solution as two sets of node labels making up the two discovered communities. Justify your answer.

Problem 3 (Version A) **(6 credits)**

Consider an inhomogeneous Poisson process (IPP) on the interval $[0,4]$ with the intensity function

$$\lambda(t) = \begin{cases} a & \text{if } t \in [0,3] \\ b & \text{if } t \in (3,4] \end{cases}$$

where $a > 0$, $b > 0$ are some positive parameters.

a) Assume that you observed a sequence of events $\{0.2, 1.0, 1.5, 2.9, 3.1, 3.8\}$ generated by the above IPP. What is the maximum likelihood estimate of the parameters a and b?

b) Assume that $a = 1$ and $b = 5$. What is the expected number of events generated by the IPP in this case?

Problem 4 (Version A) **(4 credits)**

We consider two transformations $f_1(\boldsymbol{z}) = \begin{bmatrix} z_1 \\ z_2^{1/3} \end{bmatrix}$ and $f_2(\boldsymbol{z}) = \begin{bmatrix} z_1(|z_2|+1) \\ z_2 \end{bmatrix}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$.

The respective inverse transformation are $f_1^{-1}(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ x_2^3 \end{bmatrix}$ and $f_2^{-1}(\boldsymbol{x}) = \begin{bmatrix} \frac{x_1}{|x_2|+1} \\ x_2 \end{bmatrix}$.

The respective Jacobians are

$$J_{f_1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3}z_2^{-\frac{2}{3}} \end{bmatrix} \qquad J_{f_2} = \begin{bmatrix} |z_2|+1 & \text{sign}(z_2)z_1 \\ 0 & 1 \end{bmatrix}$$

$$J_{f_1^{-1}} = \begin{bmatrix} 1 & 0 \\ 0 & 3x_2^2 \end{bmatrix} \qquad J_{f_2^{-1}} = \begin{bmatrix} \frac{1}{|x_2|+1} & \frac{-\text{sign}(x_2)x_1}{(|x_2|+1)^2} \\ 0 & 1 \end{bmatrix}$$

We assume a Gaussian base distribution $p_1(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. We observed one point $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

We propose to stack the transformations f_1, f_2 to transform the base distribution p_1 in the distribution p_2 with normalizing flows. Compute the likelihood for $\boldsymbol{x}$ under the transformed distribution p_2 if the order of transformations is f_1 followed by f_2.

Hint: You might use the density of the unit variate Gaussian $p = \mathcal{N}(0,1)$ *at the following points:* $p(1/2) = 0.3521$, $p(1/3) = 0.3774$, $p(1/9) = 0.3965$, $p(5) = 1.4867e^{-06}$, $p(8) = 5.0523e^{-15}$, $p(10) = 7.6946e^{-23}$

0 1 2 3 4

Problem 4 (Version B) **(4 credits)**

We consider two transformations $f_1(\boldsymbol{z}) = \begin{bmatrix} z_1 \\ z_2^{1/3} \end{bmatrix}$ and $f_2(\boldsymbol{z}) = \begin{bmatrix} z_1(|z_2| + 1) \\ z_2 \end{bmatrix}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$.

The respective inverse transformation are $f_1^{-1}(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ x_2^3 \end{bmatrix}$ and $f_2^{-1}(\boldsymbol{x}) = \begin{bmatrix} \frac{x_1}{|x_2|+1} \\ x_2 \end{bmatrix}$.

The respective Jacobians are

$$J_{f_1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3}z_2^{-\frac{2}{3}} \end{bmatrix} \qquad J_{f_2} = \begin{bmatrix} |z_2| + 1 & \mathrm{sign}(z_2)z_1 \\ 0 & 1 \end{bmatrix}$$

$$J_{f_1^{-1}} = \begin{bmatrix} 1 & 0 \\ 0 & 3x_2^2 \end{bmatrix} \qquad J_{f_2^{-1}} = \begin{bmatrix} \frac{1}{|x_2|+1} & \frac{-\mathrm{sign}(x_2)x_1}{(|x_2|+1)^2} \\ 0 & 1 \end{bmatrix}$$

We assume a Gaussian base distribution $p_1(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. We observed one point $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

We propose to stack the transformations f_1, f_2 to transform the base distribution p_1 in the distribution p_2 with normalizing flows. Compute the likelihood for $\boldsymbol{x}$ under the transformed distribution p_2 if the order of transformations is f_2 followed by f_1.

Hint: You might use the density of the unit variate Gaussian $p = \mathcal{N}(0, 1)$ *at the following points:* $p(1/2) = 0.3521$, $p(1/3) = 0.3774$, $p(1/9) = 0.3965$, $p(5) = 1.4867e^{-06}$, $p(8) = 5.0523e^{-15}$, $p(10) = 7.6946e^{-23}$

Additional space for solutions–clearly mark the (sub)problem your answers are related to and strike out invalid solutions.