

EXAM – METHODS IN AI RESEARCH

6 NOVEMBER 2019, 9:00-12:00, EDUC THEATRON/MEGARON
MSC ARTIFICIAL INTELLIGENCE

Instructions – please read carefully before you open the exam!

- Please do not open the exam until the lecturers have given the start signal.
- If you have any questions, please raise your hand.
- You have 3 hours to complete the exam.
- There are 3 parts, each with subquestions.
- For the exam you can get a maximum of 90 points. For each question the amount of points is indicated.
- Use the separate sheets with text blocks after each question to write down your answers. If needed there are some additional answer sheets at the back of the exam, make sure to clearly indicate the question if you write on these sheets.
- Please write your name and student number on EACH PAGE.
- Please DO NOT REMOVE THE STAPLE from the exam.
- Please provide your answers in English.
- It is not allowed to have any written materials besides this exam and the official pages for notes on your table.
- Electronic equipment such as mobile phones, smartwatches etc. must be stowed away during the exam.
- Calculators are not allowed. If the final answer contains a fraction or a square root and you don't know the exact numerical answer, it is fine to leave the fraction or square root in the answer.

Once you are done

- Check that you have answered all questions.
- Check that each page has your name and student number on it.
- Walk up to the attendants who will collect your exam. Have your student ID ready.

Information about grades, resit, and exam inspection

- The final grade is calculated as follows: $(\text{points}+10)/10$.
- The grades will be published on Blackboard within two to three weeks.
- This exam contributes 40% of your final grade.
- To pass the course you need a minimum unrounded grade of at least 5.0 on this exam and the team assignment, and average unrounded grade of at least 5.5.
- After the publication of the grades, we will arrange for an exam inspection moment as soon as possible.
- The resit exam will be on Monday 6 January 2020, 17:00-20:00.

F 4
D 13
C 8
R 1

Name: *Otto Matthes*

Student number: *6324363*

PART 1 – MACHINE LEARNING AND LANGUAGE (26 OF 36 PTS)

1. Machine Learning Experimentation (14pt)

You want to use a decision tree for your classification problem. To prevent overfitting, you want to set a maximum depth to which your tree can grow.

- a. How would you determine the value for the maximum depth? (4 pt)

You are building a fraud detection system for banking transactions. Based on the output of your classifier (fraudulent, non-fraudulent), flagged transactions will be manually reviewed by bank employees and customers will be contacted if necessary. The table below shows a confusion matrix with the predicted and true labels.

		True labels	
		Fraudulent	non-fraudulent
Pred. labels	fraudulent	5	5
	non-fraudulent	10	45

- b. Provide the precision and recall for the 'fraudulent' category, as well as the overall accuracy. (6 pt)
c. Which metric is more important for this task, precision or recall? Please explain your answer. (4 pt)

2. Logistic Regression (8pt)

You are developing a classifier for a task with two categories (0 or 1). Your logistic regression classifier has learned the weights shown below:

Feature	Weight
Bias term	-0.1
Feature 1	1
Feature 2	-0.5
Feature 3	0.5

Weights of the logistic regression classifier

Feature	Value
Feature 1	2
Feature 2	2
Feature 3	-1

Test instance

Your classifier outputs the label 1 when $P(Y = 1|X) \geq 0.5$. Which label would be predicted by your classifier for the test instance? Please explain your answer (8 pt).

3. Text similarity (4pt)

We have the following two sentences:

- the cat chases the dog
- the cat sits on the chair

What is the cosine similarity between these two sentences when each sentence is represented by a vector with the frequency counts of the words? (4 pt)

Please write your answers in the text box on the next page

- ① a) First, separate data to train/dev/test set. Then look at error rates for train and dev sets which should be similar. Adjust max depth accordingly.
- b) Precision : ~~$\frac{5}{15} = \frac{1}{3}$~~ $\frac{5}{10} = \frac{1}{2}$ $\frac{5}{5} = 1$ ✗
- Recall : ~~$\frac{5}{15} = \frac{1}{3}$~~ $\frac{5}{10} = 0.5$ ✗
- Accuracy : ~~$\frac{5}{15} = \frac{1}{3}$~~ $\frac{5}{15} = \frac{1}{3}$ $\frac{5}{15} = \frac{1}{3}$ ✗
- c) Recall as it represents how well did we predict fraudulent classifier^{labels} in terms of all true fraudulent classifier^{labels}. In short, how many fraudulent labels were we able to "recall" from all of them. (High value in identifying all fraudulent labels)

3

0

2

0

- ② ~~$1 \cdot 2 = 2$~~ Label predicted would be 0.
- ~~$-0.5 \cdot 2 = -1$~~ This is due to not all features having the value $X \geq 0.5$.

- ③ Let sentence one be vector one (v_1) and sentence two be vector two (v_2), then

$$v_1 [2, 1, 1, 0, 1, 0] \times$$

$$v_2 [2, 1, 0, 1, 0, 1]$$

$$\text{length}(v_1) = \sqrt{2^2 + 1^2 + 1^2 + 1^2} = \sqrt{7}$$

$$\text{length}(v_2) = \sqrt{2^2 + 1^2 + 1^2 + 1^2} = \sqrt{7}$$

$$\text{Dot product} = 2 \cdot 2 + 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 = 5$$

$$\text{cosine similarity} = \frac{5}{\sqrt{7} \cdot \sqrt{7}} = \frac{5}{7} \times$$

2

Name: Otto Matthes

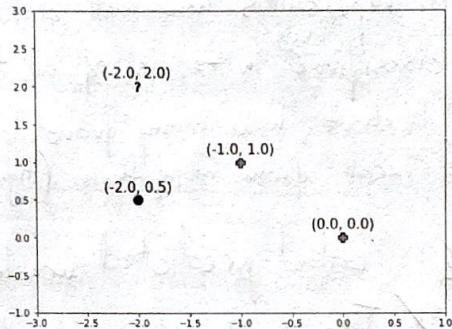
Student number: 6324363

PART 1 (CONT'D) - MACHINE LEARNING AND LANGUAGE (10 OF 36 PTS)

4. Nearest neighbors (6pt)

The figure below shows you a dataset with three training points (labeled with • and +) and a test point (indicated with a ?).

- What would a 1NN classifier predict for the test point when using the Manhattan distance? Also provide the distance to the closest point. (3 pt)
- What would a 1NN classifier predict when using the Euclidian distance? Also provide the distance to the closest point. (3 pt)



5. Speech acts (4 pt)

- What type of speech act (assertive/commissive/directive/ expressive/declaration) is "I'm really sorry!" according to Searle's Speech Act taxonomy? (2 pt)
- What type of speech act (assertive/commissive/directive/ expressive/declaration) is "Call me at seven." according to Searle's Speech Act taxonomy? (2 pt)

Please write your answers in the text box on the next page

④ a) Let the test point be a vector (v_{test}), then

$v_{test} [-2, 2]$ and the closest point to be a vector (v_{close}),
then $v_{close} [-2, 0, 5]$.

$$\cancel{L_1(v_{test})} = \cancel{| -2 | + | 2 |} = \cancel{4}$$

$$L_1 = |-2 - 2| + |2 - 0,5| = 5,5 \times$$

Answer. Test point would be labelled ~~+~~ with \bullet . 8

2

b) Let the test point be a vector (v_{test}), then

$v_{test} [-2, 2]$ and the closest point to be a vector (v_{close}),

$$\text{then } v_{close} \cancel{[-1; 1]}.$$

$$\cancel{L_2(v_{test})} = \cancel{\sqrt{-2^2 + 2^2}} = \cancel{\sqrt{8}}$$

$$L_2 = \sqrt{(-2-2)^2 + (2-1)^2} = \sqrt{5} \quad +$$

Answer. Test point would be labeled as $+$. 8

4

⑤ a) expressive

+

b) directive

Name: Otto Mattern

Student number: 6324363

PART 2 – COGNITIVE MODELING AND EXPERIMENTATION (18 PTS)

1. Cognitive Modeling (8 pt)

Cognitive models can be developed at different 'levels of abstraction' or "bands of cognition" (see also Anderson, 2002; lecture slides).

Consider this case study: Jane wants to improve the design of a website to identify which items take long to find and which items are confusing (e.g., lead to errors). For this effort she also wants to develop a model. To develop the model and compare it to human data, she has access to behavioral data of 10 users that browse the website: eye fixation times, mouse clicks, and key presses.

Explain which of Newell's bands is most appropriate to model human behavior for this situation.

2. Experimentation (10 pt)

Please fill in the blank or make the right choice for the following 5 subquestions about the topic of experimentation. No motivation is needed, just the correct word/answer for each subquestion.

(A) If all participants experience all conditions of an experiment, this is a _____ - subjects experiment.

(B) The variable that an experimenter manipulates in an experiment is the dependent / independent variable. (choose the correct one)

(C) A _____ occurs when an experimenter wants to manipulate just one factor to answer a specific question (e.g., "who is the best lecturer: Floris, Dong, Rosalie, or Chris"), but accidentally also varies other factors (e.g., time of day, topic of the lecture) at the same time.

(D) Imagine there are two experimental conditions: "coffee" and "tea". In an experiment, one group of participants first drinks "coffee" and then "tea", the other group of participants first drinks "tea" and then "coffee". We call this method to control for order effects: _____

(E) _____ validity is about making sure that you are "measuring what you think you are measuring". For example, you violate this type of validity when you invent your own unvalidated questionnaire for immersion in games to test immersion level of a game.

Please write your answers in the text box on the next page

① In this case, the cognitive band would be appropriate.

Even though eye fixation times might open the door to be able to look at biological band, most meaningful information should reveal itself in ~~less~~ the scope of seconds (rather than milliseconds or minutes). Couple the time-scale with ^(user cognition) the measurements available and the cognitive band should allow for proper analysis and possibly improvements of the website through model creation.

8

② A) ~~over~~

B) ~~dependent~~

c) ~~error in data collection~~

0

D) ~~algorithmic theory~~

E) ~~internal~~

Name: Otto Waffas

Student number: 6324363

PART 3 - REASONING & ARGUMENTATION (24 OF 36 PTS)

1. Description Logic (8pt)

In the DL sentence in (a) below, Master Degree stands for *has a master's degree*.

- a. Which class of human beings is described by the following sentence? (4 pt)

$$(\forall \text{parent. Master Degree}) \sqcap (\forall \text{grandparent. } \neg \text{Master Degree})$$

- b. Translate the sentence given in (a) into FOL. (4 pt)

2. First-order Logic (12pt)

Consider a FOL language that contains exactly one function symbol, f , and one predicate symbol S . Both f and S are unary. The language is used to formalize a computer program that on a given input produces a certain output that is either safe or not. $f(x)$ denotes the output on input x and $S(y)$ denotes that y is safe. In the following questions, you may assume that all quantifiers range over all possible inputs.

- b. Formalize the following statement in FOL: *For any input, if the output is different from the input, then that output is safe.* (4 pt)

In the following questions, let T be the sentence $\forall x(f(x) \neq x \vee S(x)) \wedge \exists z S(z) \wedge \neg \exists x \exists y(x \neq y)$

- c. Suppose model $M = (D, I)$ consists of one element, meaning that D consists of one element, say
a. Explain why $M \models T$ implies that $M \models f(a) = a$ and $M \models S(a)$. Please use no more than 3 lines for your answer. (4 pt)
d. Explain why all models $M = (D, I)$ of T consist of one element. Please use no more than 3 lines for your answer. (4 pt)

3. Human reasoning (4pt)

Recall that the Syllogism Task uses the following abbreviations:

$$\begin{array}{ll} Axy & \text{all } x \text{ are } y \\ Exy & \text{no } x \text{ is } y \end{array}$$

$$\begin{array}{ll} Ixy & \text{some } x \text{ are } y \\ Oxy & \text{some } x \text{ are not } y \end{array}$$

In general, humans tend to conclude Ayx from Axy . The following syllogisms (a and b) are stated to be true by many subjects. One of them is true according to FOL, the other is not. Indicate the false one and argue why humans may come to the opposite conclusion. Please use no more than 3 lines for your answer. (4 pt)

$$(a) \frac{Oxy \ Azy}{Oxz}$$

$$(b) \frac{Ixy \ Azy}{Oxz}$$

Please write your answers in the text box on the next page

① a) The class of human beings described is
~~all people whose parents with a Master's Degreee, whose parents (the grandparents) do not have a Master's Degna~~

1

~~b) $\forall x (\exists y \rightarrow y) \wedge \exists (\forall y \rightarrow \exists y)$~~

0

② b) $\forall x (\exists y \rightarrow x = f(y))$

0

c) As there are no more elements than 1, so the other variables next to x drop out.

0

d) As x will always conjunction.

0

③ Look at the back.

a) is false as FOL is undecidable. People might think that if all z are y and if some x are not y, then surely the x might hold values without z. Euler circles representation, think we cannot say this as it's



a) some x are not y
 all z are y
 some x are not z

b) some x are y
 all z are y
 some x are not z

Answer: b) is false, as we can not conclude that some x are not z. People might want to say "all z are y & some x are y, thus some x are not y" while all z are y.

Name: Oto Matas

Student number: 6324363

Name:

PART 3 (CONT'D) - REASONING & ARGUMENTATION (12 OF 36 PTS)

4. Arguments and attacks (6 pt)

Take the knowledge base $\mathcal{K} = \{\text{Bird}, \text{Penguin}\}$ and the rule-base $\mathcal{R}_d = \{r_1: \text{Bird} \Rightarrow \text{Flies}, r_2: \text{Penguin} \Rightarrow \neg\text{Flies}, r_3: \text{Penguin} \Rightarrow \neg r_1\}$. With this knowledge and rule base, the following arguments can be built in ASPIC+.

$A_1: \text{Bird}$

$A_2: A_1 \Rightarrow \text{Flies}$

$B_1: \text{Penguin}$

$B_2: B_1 \Rightarrow \neg\text{Flies}$

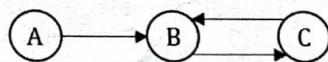
$C: B_1 \Rightarrow \neg r_1$

Indicate which of these arguments attack each other according to ASPIC+, and what the type of each attack is (rebut/undercut/undermine).

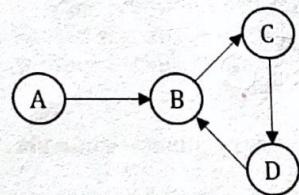
5. Extensions (6 pt)

What are the preferred extensions of the following argumentation frameworks (nodes are arguments and arrows are attacks)?

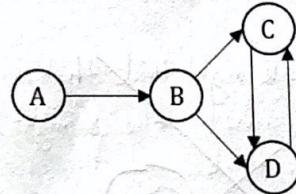
a.



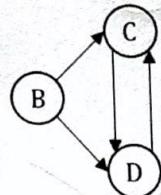
b.



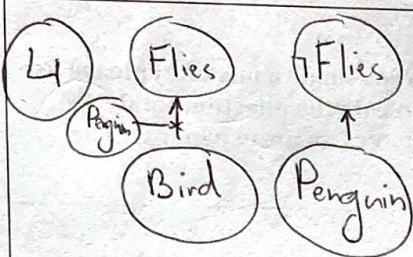
c.



d.



Please write your answers in the text box on the next page



Class Penguin attacks the argument "if it's a bird, it flies".
Penguins do not fly while being a bird so the argument attacked, is attacked by rebuttal.

- ⑤ a) $\{A, C\}$ ✓
- b) $\{\cancel{A}, D\}$ ✗
- c) $\{A, C\} \{A, D\}$ ✓
- d) $\{B, C\} \{B, D\}$ ✗

END OF THE EXAM

Extra pages: if you do not have enough space (e.g. because you made a mistake) please use the following pages to write down your answer. Clearly indicate the questions and subquestion. Please do not use these pages for your own notes (separate paper will be provided for this).

Part 3

3. Human reasoning

$$\text{a) } \frac{\exists x y A_{xy}}{\exists x_2}$$

$$\text{b) } \frac{\forall x y A_{xy}}{\forall x_2}$$

Answer

FOL is undecidable. People might think that in case all x are y and some x are y , then surely there must be some x that are $\underline{\text{not}}$ y (and inherently $\underline{\text{not}}$ y).

b) is ~~false~~. (While some x are y , also all x can be y).

and thus..

3