Vor- und Zuname →	**Max Mustermann (Student sample)**	
Matrikelnummer →	**Student's Copy**	
Studienkennzahl →	**Student's' Copy**	
Semesteranzahl →	**Student's Copy**	
Prüfungsdatum →	**1.Prüfungstermin 28.05.2019** Vers. A	⬅

Welcome to the final exam of the lecture "Machine Learning for Health Informatics"!

This exam consists of a sum of questions in different question blocks. You can reach a maximum of 100 credit points, which will be used for calculation of your final grade:

Fail	Poor	Average	Good	Excellent!
Nicht genügend	Genügend	Befriedigend	Gut	Sehr gut
5	4	3	2	1
0-50	51-69	70-79	80-89	90-100

Before you start, please answer some general questions.

I feel well at the moment.

Yes, I fully agree *No, I fully disagree*

①----------②----------③----------④----------⑤

Medical Informatics is a difficult subject.

Yes, I fully agree *No, I fully disagree*

①----------②----------③----------④----------⑤

Medical Informatics is very interesting.

Yes, I fully agree *No, I fully disagree*

①----------②----------③----------④----------⑤

Good luck! ☺ ...

A) Yes/No decision question block

Please check the following sentences and decide if the sentence is true = YES; or false = NO; for each correct answer you will be awarded 2 credit points.

01	Medical Decision Making is a search task in arbitrarily high dimensions with the additional problem of limited time.	❒ Yes ❒ No	2 total
02	Causability is the property of a person, while explainability is the property of an AI-System, important is the mapping of both	❒ Yes ❒ No	2 total
03	Highly structured data contains low information entropy, i.e. H = 0 if there is no uncertainty, if everything is in order; consequently, H can be useful to look for regularities in biomedical data.	❒ Yes ❒ No	2 total
04	Humans are much better able to perform transfer learning than current AI Systems.	❒ Yes ❒ No	2 total
05	The prediction accuracy is much more important than the possibility of providing causal explanations on demand.	❒ Yes ❒ No	2 total
06	The Bayesian concept of Prior Probability avoids that the hypotheses are too specific, while the likelihood of a hypothesis ensures that the definition is not too broad.	❒ Yes ❒ No	2 total
07	If we have little data and/or deal with NP-hard problems we still need the human-in-the-loop.	❒ Yes ❒ No	2 total
08	A function f: X → Y between two topological spaces (X, T_X) and (Y, T_Y) is called a homeomorphism if f is bijective, continuous, and the inverse function f^{-1} is also continuous.	❒ Yes ❒ No	2 total
09	In a rule based expert system the certainty factor CF of an element is calculated by: CF[h] = MB[h] – MD[h]; CF is negative, if more evidence is given for a hypothesis, otherwise CF is positive.	❒ Yes ❒ No	2 total
10	Biomarkers are measured molecules which indicate the presence of an abnormal condition within a patient, and can be a gene (e.g., SNP), protein (e.g., prostate-specific antigen), or metabolite.	❒ Yes ❒ No	2 total

Sum of Question Block A (max. **20 points) – filled out by teacher**		

End of Question Block A – please proceed to the next block

B) Multiple choice question block (MCQ)

Note: The following questions are composed of two parts: the stem, which identifies the question or problem and a set of alternatives which can contain 0, 1, 2, 3 or 4 correct answers to the question, along with a number of distractors that might be plausible – but are incorrect.

Please **select the correct answers** by ticking ☒ - and do not forget that every question can have 0, 1, 2, 3 or 4 correct answers. Each question will be awarded 4 points ***only if everything is correct.***

01	To reach a level of usable intelligence we need to … ❒ a) ... collect large amounts of top-quality data. ❒ b) ... disentangle underlying explanatory factors of the data. ❒ c) ... extract knowledge and generalize. ❒ d) ... understand the data in the context of a problem domain.	4 total
02	Automatic Machine Learning fails when … ❒ a) ... having large amounts of top-quality data. ❒ b) ... dealing with rare events. ❒ c) ... dealing with NP-hard problems. ❒ d) ... confronted with Salt-and-Pepper-Noise.	4 total
03	Causality is …. ❒ a) … not so important, because correlation is easier to measure. ❒ b) … important, because it allows to answer "why" questions. ❒ c) … the relationship that you observe between two or more variables ❒ d) … aka Causation.	4 total
04	Abductive Reasoning … ❒ a) … is used to get best explanation from incomplete set of preconditions. ❒ b) … seeks to find the simplest and most likely explanation. ❒ c) … solutions are aka "best available" or "most likely. ❒ d) ... is used often in diagnostic expert systems.	4 total
04	Standardized medical data … ❒ a) … is the majority of all data in the hospital. ❒ b) … is the basis for accurate communication. ❒ c) … contains tags or markers to separate structure and content. ❒ d) ... ensures that information is interpreted by all subsequent medical professionals with the same understanding.	4 total
06	Information retrieval models in the health care domain, which are following the concept of "reasoning with uncertainty" apply … ❒ … Probability theories. ❒ … Graph theories. ❒ … Algebra calculus. ❒ … Fuzzy set theories.	4 total
07	MCMC is important because … ❒ a) ... it uses the concept of randomness to solve problems with uncertainty. ❒ b) ... it allows to solve multidimensional integrals – otherwise intractable. ❒ c) ... it can be used for systems with many dof. ❒ d) ... it allows to calculate characteristics of high-dimensional distributions.	4 total

08	The ethics commission at a standard medical research institution checks ❒ a) … the output of a study in terms of citations. ❒ b) … the social value. ❒ c) … fair selection of study population (inclusion – exclusion criteria). ❒ d) ... risk-benefit ratio and the minimization of risks.	4 total
09	A determination of more/less, a transformation $x \mapsto f(x)$ in $\mathbb{R}$ and the basic statistics median and percentiles are applicable on … ❒ a) ... nominal data structures. ❒ b) ... ordinal data structures. ❒ c) ... interval data structures. ❒ d) ... rational data structures.	4 total
10	Superintelligence … ❒ a) ... is inevitable ❒ b) ... is impossible ❒ c) … we do not know it ❒ d) ... does not need a body – it just needs an Internet connection	4 total

Sum of Question Block B (max. 40 points) – filled out by teacher	

End of Question Block B – please proceed to the next block

C) **Free recall blocks** – please follow the instructions below. At each question you will be assigned the credit points indicated if your option is correct (partial points may be given).

01	Probability. 0,8% of the patients have cancer. Cancer is deteced with a probability of 98% and the chance that a negative test-result is correct is 97 %. a) Write down the equation for the posterior probability b) Calculate **all** probabilities c) Calculate P(Cancer | Test=true) Solutions:	4 each 16 total

02	Decisions: a) Explain in detail the two types of decision and b) provide two examples for each of them:	8 total

03	Shannon-Entropy: a) Explain in your own words what entropy is and why it is important for us b) Write down the equation for the Shannon-Entropy c) Which entropy is higher: A) 123456789 or B) 12344679 d) Calculate the Entropy for 1100101 (b=2)	2 each 8 total

04	Problem Solving: Human vs. Computers In which problem solving tasks do humans perform better and in which problem solving tasks do computers perform better? Provide three different examples of each and explain them and define what "better" means!	4 each 8 total

Sum of Question Block C (max. 40 points) – filled out by teacher	

That's it! Thank you very much for taking the exam!

My best wishes for your further studies, Andreas Holzinger