

EXAMINATION

Course code: ITI43210	Course: Machine Learning
Date: 4 May to 6 May 2015	Take home exam Duration from 09:00 a.m. 4 May 2015 to 14:00 p.m. 6 May 2015
Permitted sources: This is a take home exam. Please take note of the regulations cited in the Declaration (to be handed in with the final answer paper)	Lecturer: Lars Vidar Magnusson
The examination: The examination papers consist of 4 pages including this page. Please check that the examination papers are complete before you start answering the questions.	
Date of announcement of the examination results: The examination results will be made available on the Studentweb no later than two workdays after the announcement of the examination results (www.hiof.no/studentweb).	

Each student must hand in a complete solution to the exam as well as complete project reports even if he or she has been cooperating with other students. Those students who have been working together must state who they have been working with and what is to be regarded as joint work.

1 The projects (65%)

Attach project reports both on paper and on a CD or DVD disc. The reports should be given both on paper and on .pdf format. The disc should also contain all files that are needed for runs with C5.0, neural nets, automatic programming and other machine learning software that has been used. The outputs from the runs should also be on the disc.

Two complete copies of the reports both on paper and on disc should be handed in on 6. May by 2.00 pm (14.00).

2 The exam (35%)

Clearly explain how each answer has been produced and give good motivations for all calculations.

1. A bank has decided to use machine learning to determine if a person should be granted a loan. Assume that the database of the bank contains the following data, where the last column shows how good a customer has been at paying back his loans.

Age	Loan for house	Unemployed	History
39	yes	yes	bad
35	yes	no	good
19	yes	yes	bad
25	yes	no	good
34	no	no	good
26	no	no	good
22	no	no	bad
25	no	yes	bad
18	yes	no	good
20	no	no	bad

- (a) Use entropy heuristics as in the ID3 algorithm to construct a decision tree that perfectly classifies this data set. Clearly explain all calculations.
- (b) Convert the decision tree to rules.
- (c) How are the following new customers classified by your tree?

Age	Loan for house	Unemployed	History
28	yes	yes	?
23	yes	no	?
27	no	no	?

2. A survey has been conducted in Manchester to determine whether a person considers themselves as United or a City supporter. The following data set describes whether a person interviewed at coordinate (x, y) is either red or blue.

(x, y)	Class
(1,6)	red
(1,9)	red
(2,5)	red
(3,4)	blue
(3,8)	red
(6,6)	red
(6,8)	red
(7,3)	blue
(8,5)	blue
(9,1)	blue
(10,5)	blue
(10,9)	red

- (a) Construct a neural net with only one perceptron node that perfectly differentiates between red and blue supporters. Describe how you find the weights in the neural net. You are not supposed to give a general algorithm for training of neural nets.
- (b) After the initial survey, it was found two additional city supporters at coordinates $(5, 10)$ and $(6, 10)$. Construct a new neural net that gives perfect classification for both the data above and these new points.
3. Assume that a function f is supposed to find the smallest of five integers. The type of f may be given as

```
fun f( ( I1, I2, I3, I4, I5 ) : int * int * int * int * int ) : int = ?
```

- (a) Create a machine learning data set suitable for automatically learning a definition of f .
- (b) Show all the steps needed to produce a definition of f using automatic programming.

4. Assume that the bank in the example above no longer wishes to use age discrimination and that the dataset then becomes as follows.

Loan for house	Unemployed	History
yes	yes	bad
yes	no	good
yes	yes	bad
yes	no	good
no	no	good
no	no	good
no	no	bad
no	yes	bad
yes	no	good
no	no	bad

Show how naive Bayes method would classify each of the following three new potential clients. Remember to clearly explain all of your calculations.

Loan for house	Unemployed	History
yes	yes	?
yes	no	?
no	no	?