Test Pearl 110 — Intelligent Interaction

Pearls of Computer Science (201300070) / Introduction to BIT (201300073)
October 21 2016, 13:45–14:45
Module coordinator: Maurice van Keulen
Instructor: Rieks op den Akker & Mannes Poel

- You may use 1 A4 sheet with your own notes for this test, as well as a *simple* calculator
- Scientific or graphical calculators, laptops, mobile phones, books etc. are not allowed. ***Put those in your bag now (with the sound switched off)!***
- Total number of points: 100

15 points **Question 1**

A bag contains 10 fair dice: of which one is a 4 sided die, one 6 sided die, 2 are 8-sided dice, 2 are 12-sided and 4 are 20-sided dice. They are all fair dice, so that for example the probability of throwing a 3 with a 20 sided die is $1/20$. John randomly draws a die from the bag. John throws with the die he has drawn from the bag. The outcome is 6. This is the data D.

H	$P(H)$	$P(D\|H)$	$P(D\|H).P(H)$	$P(H\|D)$
4				
6				
8				
12				
20				

(a) Copy the above table to your answer form and fill the table (all entries should be filled in). In the table H stands for the possible dice (hypothesis), $P(H)$ is the prior probability, $P(D|H)$ the likelihood, and $P(H|D)$, the normalized posterior of H given the data D.
(b) What is the most likely value of H when we know D (i.e. that the outcome is 6)?

15 points **Question 2**

Consider the following case of a car accident that involved a taxi.

All taxis in town are blue or green. It is known that under dim lighting conditions discrimination between blue and green is 75% reliable; which means that $P(WB|B)$ as well as $P(\overline{WB}|\overline{B})$ are 0.75, where B is the boolean variable for the predicate *"the taxi is blue"* and WB is the boolean variable for the predicate *"the witness says the taxi is blue"*.

Suppose that 8 out of 10 taxis are actually blue. There are two witnesses of which one declares that the taxi was blue, the other declares the taxi was green. Given the declarations of our witnesses what is the probability that the taxi is blue?

20 points **Question 3**

A bag H_1 contains 10 marbles: 1 red, 4 white and 5 blue. Bag H_2 also has 10 marbles, now: 1 red, 3 white and 6 blue. Someone randomly selects a bag and takes 5 times one marbles from this bag. After each take he puts the marble back into the bag before he does the next take.

(a) What is the probability of outcome $D =< 2,1,2 >$ (i.e. 2 red, 1 white and 2 times blue) if he takes from bag H_1?
(b) What is the probability of $D =< 2,1,2 >$ if he takes them from H_2?
(c) What is the most likely bag the marbles were taken from when the outcome is $D =< 2,1,2 >$?

20 points **Question 4** Consider the following dataset with attributes (features) A and B, which both can take values T and F. The class label is given in the last column and is $+$ or $-$.

Id	A	B	*Class label*
1	T	F	+
2	T	T	+
3	T	T	+
4	T	F	+
5	T	T	-
6	F	F	-
7	F	F	-
8	F	F	-
9	T	T	-
10	T	F	-

(a) What is the information gain for attribute A and for attribute B?
A table with values for $-p\log(p)$ can be found at the end of this exam.
(b) Which attribute will be at the top of the decision tree and why?
(c) Compute and draw the complete decision tree.
(d) What is the error rate on the training set for this decision tree?

15 points **Question 5** A certain classifier was tested on a test, resulting in the following confusion matrix:

		Predicted class		
		C_1	C_2	C_3
Actual Class	C_1	120	15	18
	C_2	16	150	10
	C_3	11	3	130

(a) What is the accuracy of this classifier?
(b) What is the definition of recall for class C_3 and what is the recall of this classifier for class C_3?
(c) What is the definition of precision for class C_1 and what is the precision of this classifier for class C_1?

15 points **Question 6** Assume that we are training a linear classifier and that the current linear classifier is given by the line $2+1x_1-2x_2=0$. The next feature point in our training set is given by $x=(-1,1)$.

(a) How will the feature point x be classified, given the current weights $w=(2,1,-2)$ of the linear classifier, 0 or 1?
(b) Assume that the feature point x is misclassified How will the weights of the linear classifier be adapted. Assume a learning rate α of 0.2.
(c) How will x be classified after the above adaptation of the weight vectors w? Is this adaptation a step in the right direction? **Motivate your answer!**

Table for $-p\log_2(p)$

p	$-p\log_2(p)$	p	$-p\log_2(p)$
0	0	1/6	0.43
1	0	2/6	0.53
1/2	0.50	3/6	0.50
1/3	0.53	4/6	0.39
2/3	0.39	5/6	0.22
1/4	0.50	1/7	0.40
2/4	0.50	2/7	0.51
3/4	0.31	3/7	0.52
1/5	0.46	4/7	0.46
2/5	0.53	5/7	0.35
3/5	0.44	6/7	0.19
4/5	0.26		