

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2013 2014

BEng Honours Degree in Computing Part III
MEng Honours Degree in Electronic and Information Engineering Part IV
BEng Honours Degree in Mathematics and Computer Science Part III
MEng Honours Degree in Mathematics and Computer Science Part III
MEng Honours Degrees in Computing Part III
MSc in Advanced Computing
MSc in Computing Science (Specialist)
for Internal Students of the Imperial College of Science, Technology and Medicine

*This paper is also taken for the relevant examinations for the
Associateship of the City and Guilds of London Institute*

PAPER C395

MACHINE LEARNING

2013, 10:00
Duration: 90 minutes

Answer TWO questions

Paper contains 3 questions
Calculators required

Section A (Use a separate answer book for this Section)

- 1a Given the target function output representation

$$o_d = w_0 + w_1x_1 + w_1x_1^3 + \dots + w_nx_n + w_nx_n^3$$

what is the Least Mean Squares (LMS) training rule used for and how is it defined?

- 1b Derive the gradient descent training rule assuming that the target function representation is:

$$o_d = w_0 + w_1x_1 + w_1x_1^3 + \dots + w_nx_n + w_nx_n^3.$$

Define explicitly the cost/error function E , assuming that a set of training examples D is provided, where each training example $d \in D$ is associated with the target output t_d .

- 1c Prove that the LMS training rule performs a gradient descent to minimize the cost/error function E defined in 1b.
- 1d Consider the instance space consisting of integer points in the x, y plane, where $0 \leq x, y \leq 10$, and the set of hypothesis consisting of rectangles (i.e., being of the form $(a \leq x \leq b, c \leq y \leq d)$, where $0 \leq a, b, c, d \leq 10$).
- What is the smallest number of training examples one needs to provide so that the CANDIDATE-ELIMINATION algorithm perfectly learns a particular target concept (e.g., $(2 \leq x \leq 4, 6 \leq y \leq 9)$)? Explain your answer in a clear manner (i.e., explain when can we say that the target concept is exactly learned in the case of the CANDIDATE-ELIMINATION algorithm and what is the optimal query strategy).

The four parts carry, respectively, 10%, 25%, 40%, 25% of the marks.

2a Consider the following set of positive (+) and negative (-) training examples:

	techno	house	D&B	venue	DJ	price	Enjoy Rave
1	deep	acid	180bpm	Fabric	AF	15£	+
2	deep	acid	140bpm	Fabric	AF	15£	+
3	tekkno	no	140bpm	Fabric	AF	20£	-
4	deep	acid	140bpm	Fabric	Beyer	20£	+
5	deep	acid	180bpm	Egg	AF	15£	-

Show the decision tree that would be learned by ID3 assuming that it is given the above-listed training examples. Write out the intermediate calculations.

- 2b Apply a distance-weighted k-Nearest Neighbour algorithm, $k=3$, to classify the instance <tekkno, acid, 140bpm, Fabric, AF, 15£>, assuming that the above-listed examples are already known. Write out the algorithm, the distance function, the weight function, and the intermediate results.
- 2c Which types of knowledge can we distinguish in Case-Based Reasoning (CBR)? Provide a short explanation of each of the types.
- 2d Explain the workings of the Roulette Wheel method and Two-point Crossover, which are typically used in Genetic algorithms.

The four parts carry, respectively, 30%, 40%, 15%, 15% of the marks.

3a Let

C be $\text{member}(U, \text{cons}(V, W)) \leftarrow \text{member}(U, W)$,

D be

$\text{member}(U1, \text{cons}(V1, \text{cons}(W1, Z1))) \leftarrow \text{member}(U1, Z1)$

- i) Is it the case that $C \models D$? Explain your answer.
- ii) Is it the case that $C \succeq D$? Explain your answer based on the definition of \succeq .
- iii) What is the lgg of C and D ? Explain your answer in terms of compatible literals.

b Let

B_1 be $p(X) \leftarrow q(X)$

B_2 be $q(a) \leftarrow$

$B = B_1 \wedge B_2$ be the background knowledge

$E = r(a) \leftarrow$ be an example.

- i) Letting H stand for the hypothesis, state the condition in Inductive Logic Programming which H , B and E must satisfy.
- ii) For B and E above, what is \perp (the most specific hypothesis)? Explain.

c Consider the following three statements.

C: Paolo lives in Rome

D: Paolo lives in Italy

E: Paolo lives in Europe

In each case below explain your answer.

- i) Represent statements **C,D,E** as definite clauses.
- ii) Define background knowledge clauses **B1,B2** which allow the clauses for **C,D,E** to be related according to their generality.

- iii) What are the generality relations between the clauses for **C,D,E** given **B**?

The three parts carry, respectively, 40%, 20%, and 40% of the marks.