

# CS3244 Exam 1: Part 2

## 28 Sep 2020

Please do not turn to the next page until you are told to do so by your proctor.

- This exam part is worth **25** marks out of a **75** mark total for all three parts.
- This exam part is estimated to take you about **25** minutes to complete.
- This exam part has a total of **7** questions.
- This exam part contains only multiple choice questions (MCQs) and multiple response questions (MRQ). Please key these into the assessment system.
- You can visit <http://www.comp.nus.edu.sg/~cs3244/2010/e1.part2.html> to reach the entry form for this survey if you lose your browser window.
- Do remember that you will need to key in and submit your answers to the according assessment system as designated by your proctor or by exam central.

1. (MRQ with 4 options; 3 marks) Bias–Variance. Mark all statements that are true.
  - (a) A model with high variance will tend to have low test error over different samples of training data.
  - (b) An  $\mathcal{H}$  with high complexity will tend to exhibit high variance.
  - (c) High variance models tend to have low bias.
  - (d) An  $\mathcal{H}$  with low bias will generally improve over one with high bias, given sufficient training data.
2. (MCQ; 2 marks) Which of the following statements is true of **Logistic Regression**?
  - (a) The sigmoid function is differentiable everywhere except at  $x = 0$ .
  - (b) The output interval of logistic regression will be  $[-1, 1]$ .
  - (c) When using gradient descent for the training of logistic regression at Step  $t$  to update  $\theta$ , we update in same direction as  $\nabla J_{train}(\theta(t))$ .
  - (d) It is used as a linear classifier by setting a cutoff value to the output of the sigmoid function.
3. (MCQ; 2 marks) In lecture, we have seen the Bias–Variance decomposition for squared loss; i.e.,  $J(f(x), h(x)) = (f(x) - h(x))^2$ . Let us now define the error as  $J(f(x) - h(x)) = (|f(x) - h(x)|^p)$ . Given this, choose the correct Bias–Variance decomposition.
  - (a)  $Bias^p + Variance^{p/2}$ .
  - (b)  $Bias^{p/2} + Variance^p$ .
  - (c)  $Bias^p + Variance^p$ .
  - (d)  $Bias^{p/2} + Variance^{p/2}$ .
  - (e) None of the above.
4. (MRQ with 5 options; 3 marks) Which set of features can result in zero training loss on the following training examples when using a linear model?

- (a)  $x_1^2, x_2^2, -1$ .
- (b)  $x_1^2, x_2^2, 1$ .
- (c)  $x_1, x_2, 1$ .
- (d)  $x_1, x_2, x_1^2, x_2^2, 1$ .
- (e)  $x_1, x_2, x_1^2, x_2^2, -1$ .

[Questions 5–7] Consider the instance space consisting of integer points in the range  $\{0, 1, \dots, 10\}$  in the  $x, y$  plane and the set of hypotheses  $\mathcal{H}$  consisting of rectangles with sides parallel to the  $x$ -axis and  $y$ -axis; i.e.,  $h \in \mathcal{H}$  is of the form  $a \leq x \leq b, c \leq y \leq d$ , where  $a, b, c$ , and  $d$  is also restricted to the range  $\{0, 1, \dots, 10\}$ , and an  $h(x, y) = 1$  if lies within or touches the rectangle defined by  $h$ , and 0 otherwise.

5. (MCQ; 7 marks) Given the input below, what does **Candidate-Elimination algorithm** output for  $G$ ?

Example	Input Instance	Target Concept
1	(6,4)	1
2	(8,7)	0
3	(4,7)	1
4	(2,2)	0

- (a)  $\{(3, 10, 0, 6), (0, 10, 3, 6), (0, 7, 3, 10), (3, 7, 0, 10)\}$   
 (b)  $\{(2, 10, 0, 6), (0, 10, 2, 6), (0, 7, 2, 10), (2, 7, 0, 10)\}$   
 (c)  $\{(0, 7, 3, 10), (3, 7, 0, 10)\}$   
 (d)  $\{(0, 7, 2, 10), (2, 7, 0, 10)\}$
6. (MCQ; 4 marks) Given an arbitrary target concept  $f$  of the same form as  $h$  (e.g.,  $3 \leq x \leq 5, 2 \leq y \leq 9$ ), what is the fewest number of training examples you must provide so that Candidate-Elimination learns it perfectly?
- (a) 6  
 (b) 2  
 (c) 4  
 (d) 5  
 (e) 8
7. (MCQ; 4 marks) Suppose we are given several input instances and obtain the  $S$  and  $G$  boundary sets after running Candidate-Elimination:

$$S: \{(6 \leq x \leq 8, 3 \leq y \leq 5)\}$$

$$G: \{(2 \leq x \leq 9, 2 \leq y \leq 7), (3 \leq x \leq 9, 2 \leq y \leq 9)\}$$

Which new instance below will reduce the number of hypotheses remaining in the version space?

- (a) (6, 5)  
 (b) (4, 6)  
 (c) (2, 9)  
 (d) (7, 4)  
 (e) None of the above. It will also depend on the instance's label.

**This marks the end of this part of the exam.  
 These is no additional material beyond this point.**