Queen Mary
University of London

Main Examination period 2022 – January – Semester A

# MTH793: Advanced Machine Learning

**You should attempt ALL questions. Marks available are shown next to the questions.**

**In completing this assessment:**

- **You may use books and notes.**
- **You may use calculators and computers, but you must show your working for any calculations you do.**
- **You may use the Internet as a resource, but not to ask for the solution to an exam question or to copy any solution you find.**
- **You must not seek or obtain help from anyone else.**

All work should be **handwritten** and should **include your student number**.

The exam is available for a period of **24 hours**. Upon accessing the exam, you will have **4 hours** in which to complete and submit this assessment.

When you have finished:

- scan your work, convert it to a **single PDF file**, and submit this file using the tool below the link to the exam;
- e-mail a copy to **maths@qmul.ac.uk** with your student number and the module code in the subject line;
- with your e-mail, include a photograph of the first page of your work together with either yourself or your student ID card.

Please try to upload your work well before the end of the submission window, in case you experience computer problems. **Only one attempt is allowed – once you have submitted your work, it is final.**

**Examiners: N. Otter, M. Poplavskyi**

 

The set of real numbers is denoted by $\mathbb{R}$. All computations should be done by hand where possible, with marks being awarded for intermediate steps in order to discourage computational aids.

**Question 1 [37 marks].**

(a) Perform two steps of the $K$-means algorithm for the data points

$$x_1 = (0,1)^t \,, x_2 = (1,0)^t \,, x_3 = (1,4)^t \,, x_4 = (-1,-3)^t \,, x_5 = (-2,-2)^t$$

and for $K = 2$. Let the centroids be

$$\mu_1^0 = (0,d)^t \text{ and } \mu_2^0 = (-d,0)\,,$$

where $d$ is the maximum between the fifth digit of your student ID and one. [9]

(b) Compute the rank of the following matrix: [2]

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ \sqrt{2} & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

(c) Find a matrix $B$ with rank 1 that is the best approximation of matrix $A$ with respect to the Frobenius norm. [10]

(d) Complete the following matrix by choosing non-zero values for the missing entries so that it has **minimal** rank. Here $d$ is the maximum between the eighth digit of your student ID and 1. [8]

$$A = \begin{bmatrix} ? & -3 & 3 & d \\ ? & ? & 6 & ? \\ -2 & ? & ? & ? \end{bmatrix}$$

(e) Complete the following matrix by choosing non-zero values for the missing entries so that it has **maximal** rank. Here $d$ is the maximum between the eighth digit of your student ID and 1. [8]

$$A = \begin{bmatrix} 1 & 0 & 4 & d \\ 2 & 1 & 8 & ? \\ 0 & 0 & 1 & ? \end{bmatrix}$$

**Question 2 [22 marks].**

(a) Rewrite the function $f(x) = \frac{x^2-1}{3x+2}$ as a function of the form $g(a) + \epsilon h(a,b)$ where $x = a + \epsilon b$ with $\epsilon^2 = 0$ . [8]

(b) Compute the derivative $f'(a)$ by using your result from Question 2(a). [6]

(c) Consider the function $f\colon \mathbb{R}^n \to \mathbb{R}^n$, with $f(x) = (x^t x)x$. Rewrite $f(x)$ as a function of the form $g(a) + \epsilon h(a, b)$, where $x = a + \epsilon b$ is a vector of dual numbers. Specify both $g$ and $h$, and compute the partial derivates $\frac{\partial}{\partial_i} f(a_1, \ldots, a_n)$ for all $i$. [**8**]

**Question 3 [8 marks].** Determine the parameters of a one-layer perceptron with two inputs that models the IMPLY function $f$, namely

| $x_1$ | $x_2$ | $f(x_1, x_2)$ |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 1 | 1 |

.

[**8**]

**Question 4 [33 marks].**

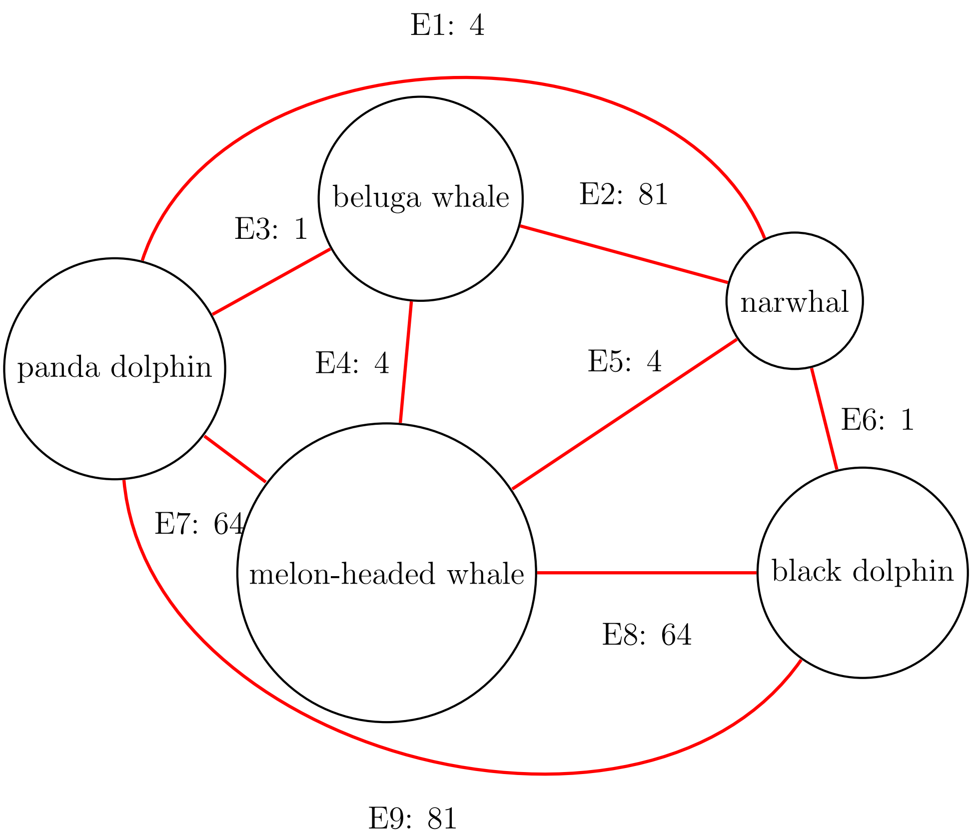


Figure 1: A weighted undirected graph.

(a) Write down the incidence matrix of the weighted undirected graph in Figure 1. Use the definition of an incidence matrix from the lecture notes and order the columns of the incidence matrix alphabetically according to the vertex name and the rows according to the edge numbering $(E_1, E_2, E_3, \dots)$. **[7]**

(b) Compute the graph Laplacian for the incidence matrix in Question 4(a). **[8]**

(c) We want to use the graph from Question 4(a) to determine whether a node in the graph belongs to the class "dolphin" or the class "whale". Suppose that we are in a semi-supervised setting, where the node "narwhal" is already labelled $v_{\text{narwhal}} = 1$ (class "whale") and the node "black dolphin" is labelled as $v_{\text{black dolphin}} = 0$ (class "dolphin"). Determine the remaining labels with the same procedure as discussed in the lecture notes and classify each node. **[9]**

(d) Now suppose that we are in an unsupervised setting, and we do not know any node label. Use the second eigenvector of the graph Laplacian to divide the nodes of the graph into two classes. **[9]**

---

**End of Paper.**