

School of Arts, Humanities and Social Science

Module title and code: Machine Learning **CMP020L015S**

Title of coursework(s): Literature Survey and Presentation

Learning outcomes:	LO1: Apply simple statistical learning algorithms such as Naïve Bayesian Classifier to a classification task and measure the classifier's accuracy. LO2: Evaluate the performance of a simple learning systems on a real-world dataset. LO3: Compare and contrast decision making techniques, such as decision trees, neural networks, and Long Short-Term Memory (LSTM) networks. LO4: Evaluate the ethical concerns of applying machine learning techniques to a real-world dataset.
Assessment weighting	40%
Maximum mark	100%
Submission details (e.g. submission link)	https://moodle.roehampton.ac.uk/mod/assign/view.php?id=2092414
Word limit (if applicable)	N/A
Date set	2/Feb/2025
Deadline	Submission date for report: 14 March 2025 by 23:59 GMT online via Moodle.
Feedback and marks	Written feedback and Rubrics will be given. Date:1 April 2025
Assessment setter's name	Kuo-Ming Chao

Academic Misconduct:

"Academic integrity and honesty are fundamental to the academic work you produce at the University of Roehampton. You are expected to complete coursework which is your own and which is referenced appropriately. The university has in place measures to detect academic dishonesty in all its forms. If you are found to be cheating or attempting to gain an unfair advantage over other students in any way, this is considered academic misconduct, and you will be penalised accordingly."

Further details about "Student Code of Conduct" and "Disciplinary Regulations" can be found at:

<https://www.roehampton.ac.uk/corporate-information/policies/>

Grading advice: While fully functioning programs and its documentation that meet all the requirements are the goal of the test, the programs and the rationale for their design, implementation, testing and discussion will be assessed for efficiency, simplicity, creativity, and good style.

Coursework Task

Imagine that you work for an AI company, you are asked to develop a machine learning product that can classify images. You will design a neural network for this image classification task. To complete this task, you are suggested to follow the instructions below.

1. Find an appropriate public image dataset. The dataset should contain images (or pixels) with their associated labels/classes. You can use any image dataset (e.g., animal images, flower images, etc.) except for the digits image dataset in lab solutions. You can look for the dataset on any public data website and here is just an example: <https://paperswithcode.com/datasets?task=image-classification>
2. Import the data to your program. Define the input and output. Split the data into training and testing datasets.
3. Design a neural network for this image classification task. Your neural network should at least have input layer, hidden layer, and output layer. Please feel free to add other layers and design your own neural network structure.
4. Build the neural network using training dataset and evaluate it using testing dataset. Present the evaluation result (e.g., accuracy).
5. Explain the training process and the model structure in a simple clear way. Please feel free to use plot or flowchart.
6. Q & A section for general understanding: Answer two types of questions: one type about the Python code used in your work, and the other type about content mentioned in lectures and seminars. Please note you will be marked as NOT PASS if you fail to answer both questions.
7. [Optional] Preprocess the data to analyse the data set and to address data imbalance or quality if there is any.
8. [Optional] Use other pre-trained models to improve the performance
9. [Optional] Develop any other additional feature you want.

Submission Requirement

You should submit the codes with your name and ID number. The file should contain one page of documentation to analyse and discuss the work and the explanation of your implementation (comment your code).

You will also need to present your work in the following two ways:

- a. Present to a tutor during lab session within two weeks after the due date.
- b. Record your screen while running the codes and include the video link **WITHIN** your submission. The video should be less than 10 minutes and highlight the key features you implemented.