

# Machine Learning

## Coursework 1 Specifications

### General Information

Please read the following information carefully before starting your coursework.

- Submission date: **25 Oct 2024 by 23:59 GMT** online via Moodle.
- Contribution: **30%** of the total mark.
- Academic misconduct: All submissions will be processed through plagiarism tool. If signs of misconduct are found, all students involved will be contacted to discuss further steps. Please see here for information on academic integrity at the university.

<https://portal.roehampton.ac.uk/information/Pages/Academic-Integrity.aspx>.

Our guiding principle is that 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. Please don't do it. It is not worth it.

- 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

To complete this coursework, you will develop a Python program to solve a supervised learning problem. The goals of the program are to import the data, split the data into training and testing set, build a model for prediction and evaluate the model.

- Import the data.

You can access the dataset in section 'Assessment' of the Moodle page, named as 'cw1data.csv'. Please download the data and import it into your program.

- Visualise the data

Develop some codes to visualise the data. You should at least produce one figure of the data.

- Split the data into training and testing set.

The column 'y' will be used as output variable and the remaining columns will be used as input variables. Define the input and output variables and then split them into training and testing datasets.

- Build the model and generate prediction

Develop codes to build regression models and then use the models to generate predictions on the testing dataset. You should at least choose two types of regression models, build the models and use the models to generate predictions.

- Q & A section for general understanding

Answer two general questions: one about the Python code used in your work, and the other about content mentioned in lectures and seminars. Please note you will be marked as NOT PASS if you fail to answer both questions.

- Reproduction of code (optional)

Rerun the designed codes during the presentation to achieve the similar results

- Develop an additional feature (optional)

Develop codes to implement at least one additional feature. For example, you could try to identify the importance of the variables, visualise your predictions results, or perform a cross validation, etc.

### **Submission Requirement**

You should submit the codes with your name and ID number. The file should contain documentation explaining your implementation (comment your code).

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

- Present to a tutor during lab session within two weeks after the due date.
- Record your screen while running the codes and include the video link WITHIN your submission. The video should be less than 3 minutes with the key features you implemented

highlighted and do the Q & A section in the lab or seminar within two weeks after the due date.

## Machine Learning Coursework 2 Specification

### General Information

Please read the following information carefully before starting your coursework.

1. Submission date: 22 Nov 2024 by 23:59 GMT online via Moodle.
2. Contribution: 40% of the total mark.
3. Academic misconduct: All submissions will be processed through plagiarism tool. If signs of misconduct are found, all students involved will be contacted to discuss further steps. Please see [here](https://portal.roehampton.ac.uk/information/Pages/Academic-Integrity.aspx) for information on academic integrity at the university.  
<https://portal.roehampton.ac.uk/information/Pages/Academic-Integrity.aspx>. Our guiding principle is that 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. Please don't do it. It is not worth it.
4. 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 a 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. [Optional] Quickly rerun the designed codes and achieve the similar outcome without error.
7. [Optional] Ask the user to import or select an image as test sample. Present the classification result.
8. [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.



Module title and code: Machine Learning CMP020L015A

Title of coursework(s): Literature Survey and Presentation

Learning outcomes:	<b>LO1: Apply simple statistical learning algorithms such as Naïve Bayesian Classifier to a classification task and measure the classifier's accuracy.</b> <b>LO2: Evaluate the performance of a simple learning systems on a real-world dataset.</b> <b>LO3: Compare and contrast decision making techniques, such as decision trees, neural networks, and Long Short-Term Memory (LSTM) networks.</b> <b>LO4: Evaluate the ethical concerns of applying machine learning techniques to a real-world dataset.</b>
Assessment weighting	30%
Maximum mark	100%
Submission details (e.g. submission link)	<a href="https://moodle.roehampton.ac.uk/mod/assign/view.php?id=2068857">https://moodle.roehampton.ac.uk/mod/assign/view.php?id=2068857</a>
Word limit (if applicable)	1000 words
Date set	27/11/2024
Deadline	<ul style="list-style-type: none"><li>• <b>Submission date for report:</b> 13 Dec 2024 by 23:59 GMT online via Moodle.</li><li>• <b>Presentation date starting:</b> 9 Dec 2024, but no later than 20 Dec 2024. Group presentation timetable will be available later.</li></ul>
Feedback and marks	Written feedback and Rubrics will be given. Date:10 January 2025
Assessment setter's name	Kuo-Ming Chao/Sameena Naaz

**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/>

**Assessment introduction (if applicable):**

Task 1:

You need to form a group with 4 members for group presentation.

Task 2:

Each student in the group must review at least two papers on **AI ethical issues and/or bias** to write a literature survey report. The students in a group cannot have the same papers. In other words, there is no paper overlapped within the group (at least 8 different papers in one group if it has 4 members).

**The word count of the report is up to 1000 words.**

**Note: The paper should be less than 4 years old.**

Task 3:

Present a literature review as groups. Each group with 4 members has 20 minutes for presentation, followed by 5 minutes for questions. The contents of the presentation should be integral as one single coherent report. So, each member needs to present around 5 minutes for part of the group report.

Deliverables (what you will need to submit):

Report in the Moodle and Presentation in the classroom or lab

Additional Information (if required):

Assessment Criteria (Grade Boundaries or Rubric): They are in a separate file.

