University of Aberdeen

South China Normal University

Aberdeen Institute of Data Science

& Artificial Intelligence.

BSc in Artificial Intelligence 2024 – 2025

Please read all the information below carefully

Assessment I Briefing Document – Group Assessed (Group of Two)

Course: JC3509 – Machine Learning	***Note: This part of assessment accounts for 30% of your total mark of the course.***

Learning Outcomes

On successful completion of this component a student will have demonstrated competence in the following areas:

- Ability to identify, prepare, & manage appropriate datasets for analysis.
- Ability to appropriately present the results of data analysis.
- Ability to analyse the results of data analyses, and to evaluate the performance of analytic techniques in context.
- Knowledge and understanding of analytic techniques, and ability to appropriately apply them in context, making correct judgements about how this needs to be done.

Information for Plagiarism and Conduct: Your submitted report may be submitted for plagiarism check (e.g., Turnitin). Please refer to the slides available at MyAberdeen for more information about avoiding plagiarism before you start working on the assessment. Please also read the following information provided by the university: https://www.abdn.ac.uk/sls/online-resources/avoiding-plagiarism/

In addition, please familiarise yourselves with the following document "code of practice on student discipline (Academic)": https://tinyurl.com/y92xgkq6

Report Guidance & Requirements

Your report must conform to the below structure and formatting, including the required content as outlined in each section. Each tasks have its percentage of marks allocated. You must supply a written report, along with the corresponding code, containing all distinct sections/subtasks.

Overview

In this assessment, you will work in **groups of two** to write a **5-page maximum** workshop paper. Workshop papers are concise, highly informative research papers that challenge authors to effectively communicate key findings with clarity and efficiency. This assessment is designed to evaluate your ability to implement and analyse Machine Learning (ML) methods, as well as to refine your academic writing skills—an essential attribute for any ML practitioner.

Please read all the information below carefully

You will select an ML task of your choice (e.g., car model classification, crop yield prediction, image generation, text summarisation, etc.) and analyse two distinct methods for solving it using an open-source dataset. Your submission should include a **5-page paper** (references are not included in the page limit) formatted using the CVPR template and all relevant code used for your analysis.

Assessment Requirements

1. **Task Selection**
   You are required to choose a machine learning task and corresponding dataset that aligns with your interests and demonstrates practical relevance. The task should be well-defined, with measurable objectives and real-world applications. It is essential that the dataset you use is open-source, publicly available, and free from restrictive licensing. The paper (discussed in the following) should provide brief discussion of dataset characteristics such as size, structure, and feature types should be included, along with a justification for why the dataset is appropriate for the selected task.

2. **Method Selection**
   To solve the chosen task, you must identify two distinct machine learning. The selected methods should differ significantly in their underlying approach, such as comparing a convolutional neural network (CNN) to a transformer-based model or a support vector machine (SVM) to a neural network. Methods that are too similar, such as ResNet versus VGG, should be avoided. A theoretical explanation of each method should be provided, outlining why it is suitable for the problem at hand.

3. **Workshop Paper (4 Pages, CVPR Template)**
   You must present the findings of your study in a concise, well-structured workshop paper. The paper should effectively communicate the problem definition, chosen methodology, experimental results, and critical analysis. The document should be formatted using double-column formatting provide on blackboard, ensuring that key findings are expressed concisely. You should provide sufficient context for your problem, making it accessible to a broader audience while maintaining technical rigor. Use figures, tables, and concise explanations to present findings in a structured manner. The analysis should be objective, quantitative, and critical, evaluating the advantages and limitations of both methods. The conclusions should be well-supported by evidence and discuss potential future work.

4. **Code Submission**

   In Addition to the written paper, you must also submit your code used to train each method and all evaluation scripts. The code submission must include:

   - Submit all code required to run your experiments.
   - Ensure the code is well-documented and reproducible.
   - Code should be structured logically and include a README file with setup instructions.

Please read all the information below carefully

Workshop Paper Structure:

Your paper should be structured as follows with the percentage weighting of marks given:

Abstract (10%)

- Summarise the problem, selected methods, key findings, and conclusions in a concise yet informative manner.

Introduction (10%)

- Clearly define the problem and its relevance.
- Justify why the problem is worth solving.
- State your research contributions explicitly.

Problem Setting (10%)

- Provide a formal definition of the task.
- Describe the dataset, including key properties (size, features, etc.).
- Discuss any preprocessing or feature engineering applied.

Methodology (20%)

- Provide a concise yet clear explanation of the two selected methods.
- Reference related literature and explain why these methods were chosen.
- Avoid re-writing existing work—focus on how and why the methods apply to your problem.

Experimental Setup (10%)

- Detail the experimental design, including hyperparameter choices, training/validation split, and computational resources used.
- Explain how fairness and reproducibility were ensured.

Quantitative Evaluation (20%)

- Compare the performance of both methods using appropriate quantitative metrics for the problem.
- Use tables, figures, and charts to enhance clarity.
- Include complexity analysis (e.g., training time, memory consumption).
- Provide a critical discussion of strengths, weaknesses, and trade-offs.

Conclusion (10%)

- Summarise the key insights from your study.
- Highlight advantages and limitations of the compared methods.
- Suggest potential improvements and future research directions.

Code Submission (10%)

- The code submission to evidence the written work.

Please read all the information below carefully

Marking Criteria

- Depth and breadth of knowledge.
- Technical details of formalisation, implementation and pseudo-code.
- Communication skills (clear, technical contents and sound reasoning)
- Structure of document.

Submission Instructions

- **Group Formation:** Each group must assign themselves to a group of two via Blackboard (MyAberdeen). Those who do not assign themselves to a group will be required to complete the assessment individually.
- **File Format:** Submit a **ZIP file** containing:
  - A **PDF** of your workshop paper.
  - All **Python code** used for experimentation.
- **Deadline:** 23:59 (Beijing Time), Friday 4th April 2025.
- **Late Submission:** Subject to penalties as per University of Aberdeen regulations.
- **Extensions:** If required, apply through the University of Aberdeen extension request process.

The name of the Zip file should have the form. "JC3509_Assessment_<group_number>.zip"

Any questions pertaining to any aspects of this assessment, please address them to the course coordinator Aiden Durrant, aiden.durrant@abdn.ac.uk