University of Lincoln
School of Computer Science

CMP9137M Assessment Item 1 Briefing 2023-2024

Module Code and Title: CMP9137M Advanced Machine Learning

Contribution to Final Module Mark: 50%

Description of Assessment Task and Purpose:
This assessment comprises two tasks on machine learning as follows.

Task 1
You are required to use Machine Learning to tackle the problem of "Image-Text Matching". Given an image and a textual description, the task is to predict whether they match or not. This implies a binary classification task, where there is match when the image and text correspond to each other—no match otherwise, see examples in the table below. The data for this task will be generated by the delivery team from the Flickr8k or Flicker30k datasets. This data was originally proposed by (Hodosh et al, 2013, JAIR) and (Young et al, 2014, TACL) to investigate methods at the intersection between machine learning, computer vision, and natural language processing.

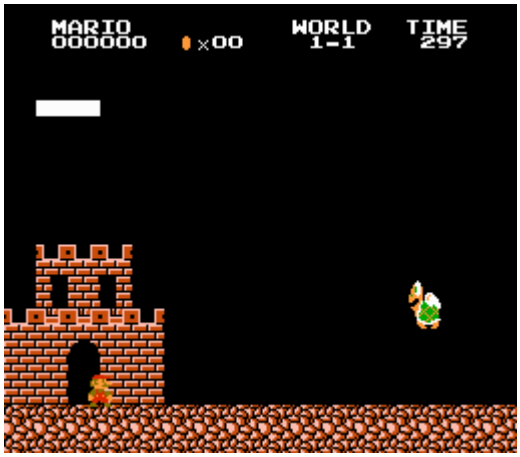
Image	Text	Label
	A young child is walking on a stone paved street with a metal pole and a man behind him.	match
	A little boy skateboarder is doing a trick on his board while another young skateboarder watches.	no-match
	Two workers at a grocery store observing something out of camera range.	match
	A man is riding a tiny motorcycle through a busy market.	no-match

Your task consists of training and evaluating at least three machine learning classifiers for matching images and sentences. To do that, you will use a unique dataset (different and potentially smaller than the originally proposed due to compute requirements) which will be available via Blackboard. Since the dataset will contain different data splits (training, validation, and test examples), you are reminded that the *test data* **should not** be involved during training. It should only be used for testing your trained classifiers, and it should be used only to report the performance of the predictions of your proposed models.

You are expected to explore a range of machine learning classifiers inspired by the materials presented and discussed within the module and beyond (i.e., from reading related literature). You will investigate and compare their performance and critique them to justify your recommendations. Your comparisons should include metrics such as Balanced Classification Accuracy, Precision-Recall, F-measure, and any other metrics that you consider relevant. In this assignment you are free to train any classifier, to do any pre- and/or post-processing of the data, and to implement your own algorithm(s) whenever possible instead of only using libraries. You are allowed to use publicly available libraries to train your models; for example, you can use any deep learning frameworks and tools (e.g. Keras, Tensorflow, Pytorch, ChatGPT). But you will need to clearly mention your resources, acknowledge appropriately, and compare between classifiers and their results in your own report.

Task 2

You are required to use Machine Learning to tackle the problem of "Game Learning". Your goal in this task is to train Deep Reinforcement Learning (DRL) agents that receive image-inputs from a game simulator, and that output game actions to play the game autonomously. The following simulator, but only version **SuperMarioBros2-v1**, will be used to play the game: https://github.com/Kautenja/gym-super-mario-bros



You are required to use your knowledge acquired in the module regarding DRL agents, and knowledge acquired from additional recommended readings. This will be useful to investigate the performance of those agents, and to compare and criticise them so you can recommend your best agent. You are expected to evaluate your agents using metrics such as Avg. Reward, Avg. Game Score, Avg. Steps Per Episode, and Training and Test Times – and any others that you consider relevant.

You are expected to train at least three different agents, which can differ in their state representation (CNN, Transformer, CNN-Transformer) and/or different learning algorithms or training methodologies. Once you have decided the agents that you want to report, you should train them with multiple seeds and average their results—to reduce the potential noise (due to randomness) in the performance of your models. If you report learning curves, they should be based on those average results instead of using a single seed (run). You are expected to justify your choices in terms of architectures, hyperparameters and algorithms.

In this assignment, you are free to train any DRL agent, in any programming language, to pre-process the data, and to implement your own solutions whenever possible. While you are free to use libraries such as PFRL, StableBaselines or Pearl (among others), you should mention your resources used, acknowledge appropriately, and compare between agents in your report.

Please read the Criterion Reference Grid for details on how your work will be graded.

Learning Outcomes Assessed:

[L03] Use machine learning software to solve complex real-world problems in an application domain of interest.

Knowledge & Skills Assessed:

- *Subject Specific Knowledge, Skills and Understanding*: e.g., literature searching, referencing, project planning, techniques, and skills subject-specific knowledge.
- *Professional Graduate Skills*: e.g., independence and personal responsibility, adaptability, verbal communication, written communication, creativity, critical thinking, IT skills, problem solving, effective time management, working under pressure to meet deadlines.
- *Emotional intelligence*: e.g., self-awareness, self-management, motivation, resilience, self-confidence.

Assessment Submission Instructions:

You must make an electronic submission of your work in PDF format (not MS Word) by using the assessment link on Blackboard for this component. You must attend the lectures and workshops for further details, guidance, and clarifications regarding these instructions.

DO NOT include this briefing document with your submission.

The deadline for submission of this work is included in the School Submission dates on Blackboard.

Date for Return of Feedback:

Please see the School's assessment dates spreadsheet.

Format for Assessment:

Your submission describing the solutions to the machine learning problems above should include a concise report of 4 pages including references using the IEEE template provided via Blackboard, which must be submitted as a PDF file on Blackboard. Whenever possible, you should cite previous works from the related literature to justify your arguments or choices. Your 4-page report should not exceed that number of pages. Whilst you can distribute your content equally per task, you are free to elaborate more on task 1 instead of task 2 or the other way round if needed. But please include all your references at the end of your report, and do not exceed the maximum number of pages in total (4).

The software implemented (source code) to solve the targeted problems should be submitted as a ZIP file in the assignment support documentation on Blackboard. Please note that your ZIP file should also include a **video of up to 3-minutes** summarising your achievements on solving the tasks above. You can use the tools of your choice to generate your video, but it should be in MP4 format – example tools to use include MS Teams and VLC Media Player.

Feedback Format:

Written and numerical feedback will be provided via Blackboard, and additional feedback can be provided upon request in a meeting or via email.

Additional Information for Completion of Assessment:

This assessment is an individually assessed component. Your work must be presented according to the Lincoln School of Computer Science guidelines for the presentation of assessed written work. Please make sure you have a clear understanding of the grading principles for this component as detailed in the accompanying Criterion Reference Grid. You are expected to take the following into account:

- Your submission should be a PDF file generated by one of the provided templates (in MS Word or Latex), which will be available via Blackboard under the materials of assessment item 1. This PDF file should be a concise report of maximum **4 pages** including references. If you fail to stick to the length requirement or omitting submitting the source code of your solutions for both tasks, your submission of this assessment **will not be marked**.
- Do not submit the dataset provided via Blackboard.
- Do not submit all your models to avoid uploading and downloading really large files. Only submit the best model per task in addition to your source code.
- Do not submit Jupyter notebook, do submit *.py files and resource/config files.
- Make sure that you submit **your own work** and not somebody else's writing or results. Failure to do so will incur plagiarism or collusion, which will the reported to the School for investigation of potential academic misconduct—see more on this below.

If you are unsure about any aspect of this assessment component, please seek advice with a member of the delivery team.

Assessment Support Information:

Assignment support will be provided during the workshop sessions and surgery hours.

Important Information on Dishonesty, Plagiarism and AI Tools:

University of Lincoln Regulations define plagiarism as 'the passing off of another person's thoughts, ideas, writings or images as one's own...Examples of plagiarism include the unacknowledged use of another person's material whether in original or summary form. Plagiarism also includes the copying of another student's work'. Plagiarism is a serious offence and is treated by the University as a form of academic dishonesty.

Please note, if you use AI tools in the production of assessment work **where it is not permitted**, then it will be classed as an academic offence and treated by the University as a form of academic dishonesty.

Students are directed to the University Regulations for details of the procedures and penalties involved.

For further information, see www.plagiarism.org