Note: This is a draft of the coursework assignment, just to be used as a guide. The final version will be available in Learning Central in the coming weeks, and may slightly change.

----------DRAFT----------

Cardiff School of Computer Science and Informatics

Coursework Assessment Pro-forma

Module Code: CMT307
Module Title: Applied Machine Learning
Lecturer: Jose Camacho-Collados, Yukun Lai
Assessment Title: Coursework 2
Assessment Number: 2
Date Set: Thursday, February 6th
Submission Date and Time: Friday, April 24th at 9:30am (group report) + Tuesday, April 28th at 9:30am (individual essay)
Return Date: Tuesday, May 19th.

This assignment is worth **50%** of the total marks available for this module. If coursework is submitted late (and where there are no extenuating circumstances):

1 If the assessment is submitted no later than 24 hours after the deadline, the mark for the assessment will be capped at the minimum pass mark;
2 If the assessment is submitted more than 24 hours after the deadline, a mark of 0 will be given for the assessment.

This will apply to any of the three parts to be submitted as part of this assignment.

Your individual submission must include the official Coursework Submission Cover sheet, which can be found here:

https://docs.cs.cf.ac.uk/downloads/coursework/Coversheet.pdf

Submission Instructions

This coursework consists of a group project divided into three parts with different weights:

- Part (1) consists of a **group report** on a specific machine learning project. The final deliverable consists of a single PDF file and a zip file with the code. The deliverable

includes a zip file with the code, and a written summary (up to 4500 words) describing solutions, design choices, evaluation and a reflection on the main challenges faced during development and insights gained throughout the process.

- Part (2) consists of an **individual reflective essay** (up to 1500 words) where students reflect on the main insights gained as part of the group project.

- Part (3) consists of a **group presentation** where groups present their project in front of the module instructors and students.

Description		Type	Name
Part 1	**Compulsory**	One PDF (.pdf) file	groupreport_[group number].pdf
Part 1	**Compulsory**	One ZIP (.zip) file containing the Python code	groupcode_[group_number].zip
Cover sheet	**Compulsory**	One PDF (.pdf) file	[student number].pdf
Part 2	**Compulsory**	One PDF (.pdf) file for the individual essay	individualessay_[student number].pdf
Part 3	**Compulsory**	One PDF (.pdf) file for the presentation slides.	groupslides_[group_number].pdf

Submission Instructions

Part 1: A nominated team member should email Dr Jose Camacho Collados (camachocolladosj@cardiff.ac.uk) and Dr Yukun Lai (LaiY4@cardiff.ac.uk) your group report as a single PDF document by 9:30am on Friday, April 24th, adding in CC all group members. Prior to handing in make sure all documentation has been collected. A single zip archive with additional supporting material, such as sources, may also be submitted if appropriate along with the code zip file. Any code submitted will be run in Python 3 (Linux) and must be submitted as stipulated in the instructions. All team members must have seen and agreed to the final version of the submission. Make sure the report clearly mentions your group number, a list of all members of the group (with full name and student id as on learning central), the project title, and the name of supervisor on the title page of your report.

Part 2: The cover sheet will be submitted together with the individual report. Both the cover sheet and individual report will be submitted in Learning Central by the individual report deadline (i.e. Tuesday, April 28th at 9:30am).

Part 3: The slides for the presentation will be sent to Dr Jose Camacho Collados (camachocolladosj@cardiff.ac.uk) and Dr Yukun Lai (LaiY4@cardiff.ac.uk) the day before the presentation takes place (date will be communicated at a later stage).

Any deviation from the submission instructions above (including the number and types of files submitted) will result in a mark of zero for the assessment or question part.

Staff reserve the right to invite students to a meeting to discuss coursework submissions

Assignment

In this coursework, students demonstrate their familiarity with the topics covered in the module via a group project.

Marks will be awarded to the individual student based on the quality of the group report and their contribution, the individual report and the presentation. Extenuating circumstances submitted for the spring term project period will be considered pro-rata for the contribution weighting.

Your submission must include the official Coursework Submission Cover sheet, which can be found here:

https://docs.cs.cf.ac.uk/downloads/coursework/Coversheet.pdf

Any deviation from the submission instructions above (including the number and types of files submitted) may result in a mark of zero for the assessment or question part.

Staff reserve the right to invite students to a meeting to discuss coursework submissions.

Part 1: Group report

In Part 1, students will be allocated in groups to design a machine learning project in one specific topic. The list of all topics along with their descriptions is available in the following link:
https://docs.google.com/document/d/1hRyzJnEOvSsNmajB13Veua0_hw8NVGK5Ba4D5ACFUP4/edit?usp=sharing

Each group will be given a specific dataset and a supervisor. The task of each group will consist of developing a whole machine learning pipeline that attempts at solving the task. Throughout the course the groups will have several milestones and should present their progress to their supervisor each session. Finally, the group will write a report summarizing the steps followed and the main insights gained as part of the process.

As part of the group decisions, each student will be allocated to one of the following tasks:

- Descriptive analysis of the dataset + Error analysis
- Preprocessing + Literature review
- Implementation + Results

Each of these tasks will have a minimum of two students involved (except in exceptional cases when this is not possible), who will work together in the specific task and as part of the group. The structure of the report will be decided by the group members. In the following link, students can find some guidelines to write the report, including some of the common sections that groups may want to include in their report:

https://docs.google.com/document/d/172LYApfCINTWez-nHw-pb2FUE-a8S_CMvpU1x9g3lWY/edit?usp=sharing

Note: These are just guidelines and students are not forced to follow this structure. New sections may be added or adjusted if necessary.

Each student will also be involved in all group activities/tasks and will be responsible with the well functioning and coordination of the team members.

Deliverables

The deliverables for this part includes **a report of no more than 4500 words** and a **zip file with the Python code**. The code should contain three specific parts:

(1) Code to get the statistics used to complement the descriptive analysis of the dataset.
(2) Code to train one of the best performing models in the training set and evaluate it in the test set. This code should also include all steps for preprocessing the original dataset, if it were necessary.
(3) A README file explaining how to run the code for each of the two parts.

The code will not be marked separately and will only be used as a complement to assess specific parts of the report.

Assessment

The final mark for this part (70% of the total marks) will result from the following items:

- Descriptive analysis of the dataset + Error analysis (14%)
- Preprocessing + Literature review (14%)
- Implementation + Results (14%)
- Student's own allocated task from one of the three above (14%)
- Group report as a whole, including its coherence and structure (14%)

Note: In addition to the specific individual task assigned, all marks may be weighted by the individual contribution in the project.

All main criteria carry equal weight as indicated above for your total mark and will be evaluated on the following scale:

Excellent (70-100%): rigorous, methodical, analytic, content meets all requirements of the work, very few errors/omissions.

Good (60-69%): competent, reasoned, coherent, content very sound, few errors/omissions.

Fair (50-59%): satisfactory, relevant, content meets many of the required elements, some errors/omissions.

Fail (1-49%): not passable, evident weaknesses, gaps in content, evident errors/omissions.

None (0%): indicates that the topic has not at all been covered.

Part 2: Individual reflection essay

In Part 2, students are asked to write a reflective essay about their group projects. The **individual essay** must discuss your contribution to the group report and to the overall group work. You must show that you contributed to the group work, which will be determined via the individual report and the contribution monitoring, conducted by the supervisor, if it were necessary. Discuss what tasks you have performed and provide evidence of your work (you may refer to the group report for the actual work/results). Discuss how you approached these tasks and how you interacted with other members, both in sharing your results and in organising the team's activities. Consider how well your existing skills were utilised and what new skills you have learned. Then reflect on your overall performance and role in the team and suggest what went well and what changes you will be making to improve (1) your performance in particular, and (2) the performance and results of methods and analyses performed as part of the group. You may also reflect on how your perspective and approach changed over time and adapted to improve your work.

The individual report must have **no more than 1500 words**. It does not have to be exhaustive, but should contain good examples of what you have done and discuss key aspects. This part weighs 20% of the total marks.

Part 3: Group presentations

In Part 3, students are asked to present their projects. Specific guidelines and dates for the presentations will be available in the following link: https://docs.google.com/document/d/1byDngcv7Sa76ItkwdgrxytBwaX9E3RypKBCR9YSNxjo/edit?usp=sharing

This information will be made available to the students at least two weeks before their scheduled presentation. The presentation weighs 10% of the total marks.

Learning Outcomes Assessed

This coursework covers the six LOs listed in the module description.

Criteria for assessment

Criteria for each individual part is provided separately. The final mark will be obtained from a weighted sum of the three parts: Part 1 - 70%; Part 2 - 20%; Part 3 - 10%.

The grade range is divided in:
Distinction (70-100%)
Merit (60-69%)
Pass (50-59%)
Fail (0-50)

Feedback

Feedback on your coursework will address the given criteria. Feedback and marks will be returned by May 19th via Learning Central. There will be opportunity for individual and group feedback during an agreed time.