

CM502650: Machine Learning 2

Assignment: Group Project

March, 2021

(This assignment comprises 40% of the overall unit assessment.)

1 Your task

The objective of this project assignment is to assess your capability of applying deep learning models to solve a real-life problem. You are free to select a suitable problem you like and also find proper data to work on. You can use any deep learning method to solve the problem; you are not limited to the models that we explicitly covered in the unit.

If you want to have some inspirations of project ideas and relevant dataset, I would suggest to visit the Kaggle website (<https://www.kaggle.com/datasets>), which was considered as the airbnb for data scientists. Note that the difficulty of the problem and datasets will be reflected in your marks.

2 Groups

You are encouraged to work in groups (maximum 4 students), because being a good team player is one of the transferable skills you will need in your future career. You are free to form your group on your own. Please sign up a group in the Moodle by **8pm, March 25th, 2021**.

All members of a group are supposed to contribute equally, therefore will get the same mark. However, if there are credible, state the reasons, I will then differentiate the marking of group members. Also, you can choose to work individually, should you wish to do so.

3 Submission

You need submit a zip file which compresses the following 2 files by **8pm, April 30th, 2021**. Note that each group only need have one submission.

1. A Google Colab notebook (.ipynb) with supporting programming code (with comments).

2. A report (.pdf). The main body of the report should contain less than 2000 words.

The paper report must include:

- A **cover page** consists of the following:
 - A form to specify how much contribution each member made, based upon your agreed input. If each member took equal contribution, this would be all 100% (which is the general expectation for this assignment). More or less than 100% means higher or lower contribution for the project.

ID	Contribution
ID1	100%
ID2	100%
ID3	100%
ID4	100%

[Optional] Reasons for different weights:

Do all the members agree with the above contributions? [Yes/No]

- A shared link which your trained models (.h5) are located at. Note that you must use "`model.save_model(MODEL_FILE)`" to save the deep learning models, instead of using "`save_weights()`". No need for sharing the non-deep learning models.
 - A link where your dataset is. Failure to offering the models and the dataset will cause not able to testing your code, thus leading to lower marks.
- The **main body** of the report: What is the real-life problem you are trying to solve? Why is it interesting (i.e., what is the motivation for this project)? What data do you use? What deep learning model(s) do you apply? How do you train and evaluate it? How well have you solved this problem? You should also discuss at least one alternative method (could be non-deep learning or deep learning methods) that are applicable, along with the relative merits of your approach. Finally, any in-depth thinking or discussion about deep learning is expected.

- **References and Appendix.**[Optional]

Note that do not include your names anywhere in your code or in your report.

4 Marking Criteria

Your report and your code will be assessed jointly, using the following criteria. During the marking process, the following criteria could be slightly adjusted.

Criterion	Description	Mks
Problem	The problem is interesting, challenging and motivated.	10
Solution	A good solution (a proper deep learning model) to the problem is proposed and explained.	15
Dataset	Proper dataset is used and properly preprocessed.	8
Model	The model is properly trained and evaluated. A good performance is achieved.	32
Comparison	At least one alternative method is compared to demonstrate the effectiveness of the chosen solution.	10
Understanding	Demonstrate the depth of understanding of deep learning.	15
Writing	The document is well-organised, explains things clearly.	5
Code	The code (with comments) in the notebook is neat and clear.	5

5 Late submissions

The university policy will be followed on late submissions. If a piece of work is submitted after the submission date, the maximum possible mark will be 40% of the full mark. If work is submitted more than five days after the submission date, student will receive zero marks. If you need an extension, please contact your Director of Studies.

6 Plagiarism

Do not plagiarise. Plagiarism is a serious academic offence. For details on what it is and how to avoid it, please visit <http://www.bath.ac.uk/library/help/infoguides/plagiarism.html>

7 Any questions about the assignment

All questions about this assignment will be answered in the **two LOIL sessions**, commencing from **Week 7 (18 March)**. As usual, I, together with one or two tutors, will be in the Thursday's session. All tutors will help you with the projects in the Friday's session. Alternatively, you may use the Moodle Forum to ask questions, then every one can benefit from the answers.