

Coursework Brief

Task To Be Done

You are given the ML 2025A3 Coursework dataset (access via [this link](#)) to build a neural network model that can be used for predicting whether or not it will rain (i.e. rain vs no rain) a week in advance, e.g. given data up until or at 17 June 2036, your model can predict whether or not there would be rainfall on 24 June 2036). For more information about the dataset, see the README file included in the dataset folder.

What To Submit

You are to submit 3 separate documents as specified below. Do NOT submit them as a zipped folder. Each document should include ONLY your candidate number (i.e. NOT your name) for identification.

1. (90 marks) **Report** (submit as a pdf file) – The report **MUST** have exactly 4 sections in the given order below. See below for what you are expected to report in each section.

A. (max – 25 marks) **Performance**

- Report the performance of your model. Describe very clearly how it was computed.
- Describe very clearly how data was split to train and evaluate the model, including the number of data instances used in each case.

B. (max – 25 marks) **Model**

- Describe your model very clearly, including its hyperparameters, how it was optimized, and its number of trainable parameters. Provide clear rationale all through.
- List and describe very clearly the steps you took to prevent overfitting for your model. (**NOTE** – 1 mark will be deducted for incorrectly including strategies that do not target overfitting.)

C. (max – 25 marks) **Features & Labels**

- Specify very clearly the input data (features) and output data (label) used, and describe well how each was extracted from the given dataset and selected for use. Provide a rationale for the approach taken at every step, and make sure to make clear the shape of the input data tensor/array as well as the shape of the output data tensor/array for your model.

D. (max – 15 marks) **Preprocessing**

- List and describe very clearly any (other) preprocessing that you did on your data for building your model, with clear details and rationale provided.
2. (5 marks) **Code notebook** (submit as a ipynb file) - Prepare the code used to complete all the tasks above and submit as a single ipynb file that can be run. NOTE – Even if you explore multiple models, only include code for the one model reported for #1B above in your submitted code, as it is only the first seen model implementation that will be assessed.
 3. (5 marks) **Model outputs for the test set(s)** (submit as a csv file) - Prepare the prediction outputs of your model for the test set in a single csv file. The file should have an appropriate header row and at least 2 columns (1 for the true labels & at least 1 for corresponding model predictions).

Other Important Notice

1. Your report must NOT be longer than 10 pages. Also, it must have NOT more than 3 tables and 2 figures. Nothing after p. 10 in the report will be read or marked. Also, tables after the first 3 and figures after the first 2 will be ignored.
2. You are not required to code algorithms from scratch. You can use standard libraries including Scikit-learn, PyTorch, TensorFlow. However, you must only use standard libraries.
3. You are NOT permitted to use or submit someone else's code, output, or report as yours. You are only allowed to use code snippets from the lab materials given by the tutor, from software library documentations as stated in #2 above, or from recommended textbooks.
4. You must NOT use generative artificial intelligence (AI) to generate any materials or content for your assessment submissions.

Note that:

- The baseline position at the University of Sussex is that the use of generative AI material in assessment submissions is prohibited, unless explicitly permitted by the module convenor.
- Students registered with the Disability Advice team and in receipt of reasonable adjustments are still permitted to use other assistive technology as required.
- If in any doubt about what is permissible, students should check with the module convenor.

Link to learning outcomes

This assessment is designed to evaluate how well you have achieved the module's learning outcomes, i.e. are able to

- demonstrate comprehensive understanding of key aspects of machine learning and standard methods;
- show awareness of relevant issues and current challenges in machine learning;
- systematically and creatively build and evaluate machine learning models;

- act autonomously in preparing data appropriately to address a given problem, selecting the most suitable techniques, and communicating valid rationale for choices made.