

Submission Guidelines

1. Overview

Your submission will consist of three of four files that need to be uploaded to MOLE. These are

- system.py (mandatory)
- model.json.gz (mandatory)
- requirements.txt (optional)
- report.md (mandatory)

No other files will be accepted. The files listed above are detailed below.

2. The files to submit

system.py

This file stores your python code. All the python code that you have written must be in this file. It must work correctly when called from the train.py and evaluate.py provided.

model.json.gz

This is the file that is generated when you call train.py in conjunction with your own system.py. As stated in the assignment instructions it must not exceed the size limit of 3 MB.

requirements.txt

If you have developed your system outside of CoCalc or if you have installed any modules using pip then I will need to know which modules and module versions you have used. The easiest way to tell me this is to include a requirements file. This is generated by typing 'pip freeze' in your python environment and saving the output into a file called requirements.txt. (If you have tested your code in CoCalc then there is no need to include this file.)

report.md

Your report must be a concisely written text file called 'report.md' strictly following the format below. Cut and paste the text below and then replace the text in square brackets with your answers. The word limits are strict.

OCR assignment report

Feature Extraction (Max 200 Words)

[Describe and justify the design of your feature extraction. Note, feature extraction includes all the steps that you have taken to produce the 10 dimensional feature vector starting from the initial image of the page.]

Classifier (Max 200 Words)

[Describe and justify the design of your classifier and any associated classifier training stage.]

Error Correction (Max 200 Words)

[Describe and justify the design of any post classification error correction that you may have attempted.]

Performance

The percentage errors (to 1 decimal place) for the development data are as follows:

- Page 1: [Insert percentage here, e.g. 98.1%]
- Page 2: [Insert percentage here, e.g. 98.0%]
- Page 3: [Insert percentage here, e.g. 83.3%]
- Page 4: [Insert percentage here, e.g. 58.1%]
- Page 5: [Insert percentage here, e.g. 38.7%]
- Page 6: [Insert percentage here, e.g. 28.9%]

Other information (Optional, Max 100 words)

[Optional: highlight any significant aspects of your system that are NOT covered in the sections above]