COM2004/3004 - Submission Guidelines

1. Overview

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

- system.py (mandatory)
- model.noisy.json.gz (mandatory)
- model.clean.json.gz (mandatory)
- requirements.txt (optional)
- report.txt (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.noisy.json.gz and model.clean.json.gz

These are the files that are 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.txt

Your report must be concisely written in an ascii text file called ‘report.txt’ 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.]

Board Context (Max 200 Words)

[Describe and justify any steps you have taken to make use of full board context.]

Performance

[Cut and paste the output that is generated when you run evaluate.py]

Other information (Optional, Max 100 words)

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