

CS5063 Assessment 1: Test an AI System

In this assessment, you will create and execute a set of test cases for an AI system. Remember that the job of a tester is not to be fair, it is to try to break the system!

You can choose the AI system you wish to test. **Please email Prof Reiter in advance with the system you would like to test**, he will tell you whether it is appropriate or not.

The AI system you test in Assessment 1 must be different from the AI system you evaluate in assessment 2.

If you are testing the same system as another student, you must do this independently, you cannot collaborate and submit a joint assessment.

This assessment is worth 50% of your overall mark for the module.

Task

Create a set of test cases for an AI system, which include

- *Easy cases*: for example, for face recognition, pictures of faces in good lighting without makeup, looking straight at the camera
- *Hard cases*: for example, for face recognitions, pictures of faces in dim lighting, with heavy makeup, or not directly looking at the camera
- *Challenging cases*: for example, for face recognition, pictures of faces in variable lighting (eg, half the face is shadowed and half is not), with full theatrical makeup, or with a sideways view of the face

You should submit a report which

- Explains the rationale for choosing the test cases
- Gives the results when running the test cases through the system
- Analyses when the AI system works well and when it does not, based on the test cases

The test cases themselves should either be listed in an Appendix or submitted as a separate file, along with your report. For each test case, you must

- Show the input. If the input is audio (eg for voice assistants) or video, you must include an audio/video recording; a transcript on its own is not acceptable.
- Show the expected output. If several outputs are acceptable, show all of them.

Marking

CGS C

30 sensible test cases (10 Easy, 10 Hard, and 10 Challenging) and the results of running these test cases through the AI system.

CGS B

As in CGS C, but also including an excellent justification of the choice of test cases, and high-quality analysis of when the AI system works well and when it does not.

CGS A

As in CGS B, but with 50 test cases (10 Easy, 20 Hard, and 20 Challenging). Some of the Hard and Challenging test cases should look at what happens when there are multiple problems, for example face recognition when there is both dim lighting and heavy makeup.

Submissions

Submissions are due at midnight on Sunday, 1 Nov.

This assignment is part of the formal assessment of the course, and the work done must be that of your own group. You are reminded to look at the section on Cheating and Plagiarism in your student handbook.

A 10% penalty will be applied to late assessments submitted up to 1 working day late, and a 25% penalty will be applied to late assessments handed in between 1 day and 1 week late. No credit will be given to assessments submitted later than 1 week after the first deadline.