MSc in Computer Science (Data Analytics, Artificial Intelligence, Artificial Intelligence - Online)

NUI Galway School of Computer Science

January 28, 2021

1 Introduction

The capstone project is a major component of the degree, worth 30 of 90 ECTS. That represents about 600 hours of work.

2 Learning Outcomes

- Identify/research a Data Analytics (DA) / Artificial Intelligence (AI) problem
- Identify, describe, and synthesize the state-of-the-art approaches to the problem
- Devise and implement a solution to the problem which may have some novelty
- Carry out experimental evaluation on the solution
- Write up the problem, state-of-the-art, implementation, experimental evaluation, results, and conclusion as a thesis.

The project is in the area of DA/AI. It is a research project, not software development. The outcome is a document (the *thesis*) describing your research – not (usually) a viable commercial product.

3 Supervisor Allocation Process

Each supervisor will have a budget of projects, related to their workload (e.g., project supervision on other MSc programs). Thus, not all students will necessarily be allocated to the supervisor they may prefer. Individual supervisors cannot make commitments to supervise particular students in advance of the final decision by faculty. Allocation of students to supervisors and projects is carried out by faculty to maximize fit and fairness.

4 Supervision

Every student has an academic supervisor, one of the members of NUI Galway School of CS faculty. In some cases, there may be a co-supervisor from the same or another School. Students may consult http://www.it.nuigalway.ie/people/ and https://datascienceinstitute.ie/people/ for information on supervisors' research interests.

Students can expect regular meetings with their supervisor in the early stages – once a week or once a fortnight is common. The frequency may decrease when the project is "up and running". Supervisors will typically take annual leave at some point in the project also. In addition

to face-to-face meetings, email, phone, Zoom meetings, etc. may be used depending on both parties' preferences.

Supervisors will expect students to be prepared for meetings, including by finding and reading relevant material and being prepared to summarise it; by working on ideas discussed at previous meetings and being prepared to summarise it; and bringing well-formed questions. Students should take notes in meetings. Some students may, e.g., photograph whiteboards after meetings.

Students in full-time employment in a related industry may carry out their project in conjunction with their employer if appropriate, subject to NUI Galway approval. In these cases, there may be a mentor from the employer. The ideal is to work on a project which is useful to the employer, and thus good for the student's career, as well as fulfilling the requirements of the degree. Students may wish to consider how to demonstrate value to the employer when scoping and writing the project, e.g., choosing certain figures or sections of the text as most appropriate for the employer. There are no extra expectations or workload on the student or the supervisor (relative to full-time).

5 Project Timeline

5.1 MSc DA / MSc AI

• Last day of February: Students submit a 1-page *Project Definition Document (PDD)* on Blackboard informally describing their proposed project, relevance, possible datasets and availability. A statement of any preliminary reading or other work already underway can also be included. A template for the PDD is provided on Blackboard. The PDD needs to be agreed with the project supervisor before submission.

• Last day of June: Students submit a formal *Project Proposal* document on Blackboard and additionally by email to their supervisor. It should describe a research problem to be solved in some detail, including background, motivation, and proposed research methods. It should incorporate a high-quality Literature Review. For further information and guidelines see Sect. 6.

• Last day of August: Students submit their final thesis document on Blackboard and additionally by email to the supervisor. A thesis template is provided on Blackboard where generic guidelines are given. Individual supervisors should be consulted with respect to any divergence with these guidelines that they might wish to specify. Observe the strict upper limit of 13,000 for the number of words. For further information and guidelines see Sect. 7.

5.2 MSc AI Online

• Semester 2, Sunday at end of Week 3: students provide up to 1 page to Programme Director informally describing their proposed project, employer relevance, possible datasets and availability, preferred supervisors. Students who choose not to propose their own projects should still provide as much detail as possible on preferred topics/supervisors, and may make contact with the Programme Director and possible supervisors in advance. Students should then include a very brief summary of any discussion with the preferred supervisor, e.g., including that supervisor's comments on project viability. A statement of any preliminary reading or other work already underway can also be included.

- End of Semester 2: Literature Review deliverable as part of Research Skills module.

- Last day of June: Students submit a formal *Project Proposal* document on <u>Blackboard</u> and additionally by email to their supervisor. It should describe a research problem to be solved in some detail, including background, motivation, and proposed research methods. It should incorporate a <u>high-quality Literature Review</u>. For further information and guidelines see Sect. 6.

- Last day of August, Year 2: Students submit their final thesis document on <u>Blackboard</u> and additionally by email to the supervisor. A thesis template is provided on Blackboard where generic guidelines are given. Individual supervisors should be consulted with respect to any divergence with these guidelines that they might wish to specify. Observe the strict <u>upper limit</u> of 13,000 for the number of words. For further information and guidelines see Sect. 7.

6   Project Proposal

The formal Project Proposal should describe a research problem to be solved in some detail, including background, motivation, and proposed research methods. It should incorporate a <u>high-quality Literature Review</u>. (For students who have taken the Research Skills module, the Literature Review produced earlier may be re-used in the proposal in part or in whole.)

This document should normally be between 11 and 24 pages long (plus references). Supervisors will give feedback verbally or on Blackboard.

The template for the final thesis (see 7.2) should also be used for the Project Proposal (but normally only a part of the template can already be completed at this time).

Normally, faculty approve the Project Proposal. However, faculty may reject a low-quality proposal and require a complete redesign or a new topic.

7   Final Thesis Document

The document should follow a typical scientific structure (see template). Guideline page counts are given in the template. It should clearly distinguish between your primary and secondary research, i.e., between what has previously been accomplished, and what has been accomplished by you. The number of words in the PDF thesis document (not in the LaTeX source) should be between 8,000 and 13,000 (not including title page, Declaration, Acknowledgements, List of Figures, List of Tables, Table of Contents and Bibliography). <u>13,000 is a strict upper limit – longer theses may be rejected without review.</u>

7.1   Writing and Style

The goal of the document is to convince a skeptical, well-informed reader, and to document work

done.

The document should be written for other researchers. It should assume the reader knows at least undergrad-level material. It should be written in clear, simple English, with short sentences and little or no "poetic" language. Graphics and tables should be clean, with no frills or decoration. The best way to develop an appropriate style of writing and formatting a paper is by reading many high-quality research papers. Examples of good-quality theses from previous years are also available in Blackboard.

A draft which has not been edited will rarely achieve quality. "The best writing is rewriting." - E B White. "My pencils outlast their erasers." - Vladimir Nabokov. Supervisors will give feedback on writing. The best way to receive useful feedback is to edit your own drafts first, then request feedback on specific sections and/or on overall structure – the earlier, the better.

The document should be focused on relevant material, with no padding and little repetition. (But it is appropriate to have some repetition between the abstract, introduction, and conclusions; and some repetition where main points in the introduction are expanded elsewhere.)

7.2   Thesis Template

We provide a thesis template in Latex format. The .zip is downloadable from Blackboard. It contains instructions for use. Latex is recommended, but students who prefer other software are free to use it. They need then mimic the thesis template format. The Project Proposal and final thesis should both adhere to this format and should both be submitted in PDF format. So, the proposal can be seen as a first (incomplete) draft of the thesis. Also see Sect. 10 (Submission).

8   Grading

The document will be evaluated by two internal supervisors. It will then be discussed at a full School meeting. Grading is overseen by an external examiner. A grading guidelines document will be released in due course.

9   Plagiarism

"All work submitted by students for assessment purposes is accepted on the understanding that it is their own work and written in their own words except where explicitly referenced using the accepted norms and formats of the appropriate academic discipline."

See http://www.nuigalway.ie/plagiarism/

Each thesis will be automatically screened to detect plagiarism, but this is just one of the methods we use. Students who try to "trick" automated plagiarism detectors are guilty of attempting to subvert the University's examination procedure. This can be detected by other methods.

9.1   Accidental Plagiarism

"Whilst some cases of plagiarism can arise through poor academic practice with no deliberate intent to cheat, this still constitutes a breach of acceptable practice and requires to be appropriately investigated and acted upon."

In other words, accidental plagiarism is plagiarism, and is subject to disciplinary action.

9.2 Examples of Plagiarism

• Use of material created or provided by another person or agency, such as an "essay mill";
• Copying the work of another student or individual, with or without that person's consent;
• Submission of a student's own work for credit in more than one course;
• Copying from a website, book, journal or other document, without proper citation;
• Use of material from a website, book, journal or other document, without an appropriate reference to the exact source;
• Use of a direct quotation, not included in quotation marks and properly attributed to the original source;
• Close paraphrasing of the work of another person, such as, for example, where only superficial changes are made to the original sentence structure.

10 Thesis Submission

The last day of August is the hard deadline (no extension possible!) for submission of your thesis. Upload your final thesis in PDF format on Blackboard by that date, and send it additionally by email to your supervisor. Also, send a draft in PDF format to the supervisor one week in advance. The submission must also include a link to a Web location where the data and code can be found, except where agreed otherwise in advance with the supervisor, e.g., because of sensitive data.

11 Publication

For faculty, publication is a core goal of research. A thesis which we evaluate as publishable receives a high grade. Students can optionally demonstrate publishability by extracting a submission-ready research paper from their thesis. This would typically by 8-15 pages in a much more compressed and focused style, targeted at a specific journal or conference. Students should discuss the possibility with their supervisor before engaging in this work.

For some projects, it may be appropriate to provide Latex source, code and data to the supervisor to allow for further experimentation and writing either by the supervisor or by a future student – standing on the shoulders of giants.

12 Intellectual Property

The copyright for documents, code and results produced by the student remains with the student. (Students in employment may need to refer to their employer and/or contract for clarification on copyright.)

Students should discuss IP issues with their supervisor if the supervisor is contributing ideas, data, or any other IP to the project.

Our strong preference is for projects involving open, anonymised or synthetic data such that the student is entitled to publish their work in the thesis and in scientific venues. Note that such publication does not necessarily require the data to be made public.

Where company data is used, workarounds may be needed to preserve sensitive data. In some cases anonymised or synthesized data can be used to avoid exposing students or super- visors to sensitive data. If a sponsoring company wishes, they can request that the students not include any sensitive information in the project outputs, so long as this doesn't prevent the students from describing their findings. For example, the students could describe their findings in detail while omitting mention of specific customer names or monetary amounts. If appropriate, the supervisor, the students and the employer would agree broadly at the outset what types of data can be included, and agree a timetable by which the students will provide representative drafts of documents or slides, and dates by which the partner will respond with confirmation that they can be published.

NUI Galway faculty cannot in general sign Non-Disclosure Agreements for MSc projects.

Normally, thesis documents are considered as published material, and may be (e.g.) copied to other faculty and future students, or deposited in the Library, or on a web page. Students (especially those in employment, carrying out projects in conjunction with their employer) may request an embargo. This means that we undertake not to deposit the document (and associated code, data, presentation slides, etc.) in the Library, or on our websites, or to allow their contents to become known outside School of CS faculty.