Machine Learning in Science Research Project (PHYS4037)

Module Handbook 2021-22

Convenor: Dr Simon Dye (simon.dye@nottingham.ac.uk)

Module Aims

The overall aim of this module is to engage students in an original scientific problem, chosen to ensure exposure to a real research environment and to investigate new, previously unexplored avenues of enquiry into the application of machine learning in science. Projects run through the entire summer term until submission of a project report at the end of term. Students work together in teams of two or three on the project but submit their own independently written reports.

Objectives

On completion of the module, students should be able to:

(a) focus a range of knowledge and skills on the solution of a real research problem
(b) describe and present technically sophisticated material in a clear, logical fashion
(c) manage time to meet strict deadlines
(d) gather information using online resources and the library where necessary
(e) work in a group and to adapt to unfamiliar environments
(f) produce a detailed but concise final report to a tight deadline
(g) follow safety procedures and keep appropriate records

Introduction

The Machine Learning in Science Research Project module is an important element of the MSc training programme, accounting for 60 of the 180 available course credits. It provides an opportunity for practising the taught professional machine learning skills by applying the knowledge acquired during the earlier stages of the course and to considerably increase students' research experience. Projects are supervised by a researcher belonging to one of the many research groups within the Faculty of Science or its industrial partners. All will involve real problems in the sense that they will contain a significant element of 'discovery' and innovation. Students will be investigating new pieces of science that have not been studied before.

As is the nature of research, there will not necessarily be a well-defined end. Even where there is, it may not be possible to provide a full solution in the time available. Nevertheless, credit will be given for an intelligent approach to the project, an appreciation of the methods used and, naturally, a comprehensive and informative final report.

Project Topics

Students are encouraged to propose their own project topics, although a list of pre-prepared topics will be made available on Moodle near the start of spring term.

- *Pre-prepared topics* – Student teams considering a pre-prepared project topic should discuss in detail each project of interest with the proposing supervisor before deciding their project choices. Those students opting for a pre-prepared project should email the module convenor (simon.dye@nottingham.ac.uk) with their **top five** preferences by **Friday 8th April 2022.** Allocated projects will be announced on **Monday April 18th 2022**.
- *Proposed topics* – Student teams proposing their own project topics should develop an idea for a project topic and then identify a member of staff within the Faculty of Science (i.e. the schools of: Biosciences, Chemistry, Computer Science, Mathematical Sciences, Pharmacy, Physics and Astronomy, Psychology) who is prepared to supervise the project. A brief project outline should then be written to justify the choice of topic. This should be no more than 1 side of A4 paper and no more than 500 words. The outline should contain a description of the science problem being addressed, the methods used (including specifics of how machine learning will be applied), the source and availability of any external datasets required and an indication of a contingency plan in case the main project goals cannot be reached in the time available. Project outlines should be emailed with the name and email address of the proposed supervisor to the project convenor by **Friday 8th April 2022.** These will be assessed by the module convenor to ensure that they meet a threshold for each of the following eligibility criteria: 1) viability of completion within the given time period; 2) originality of proposal; 3) suitability of proposed methods; 4) availability of any supporting data required. If necessary, a period of arbitration **lasting no more than 1 week** will ensue between the module convenor, supervisor and student team until the criteria are met. It is therefore in the interests of students to ensure that a well thought-out proposal is submitted in the first instance. Proposed projects must be formally accepted by **Friday 15th April 2022**. Those students who have failed to meet this deadline will have a project assigned to them at random from the list of pre-prepared projects. Student teams who are having difficulty identifying a supervisor should contact the module convenor for assistance well before the topic submission deadline.

Project Teams

Students will be randomly assigned one or two partners unless they request to work with specific individuals. Such requests must be emailed with the project choices/proposals to the module convenor by **Friday 8th April 2022.**

Project Supervisor

Students should arrange to meet with their supervisor before 2nd May 2022. This will allow any preparatory work, reading of any recommended material and giving consideration to how the project may develop to be done well before full-time project work begins. Student teams should then **meet their supervisor a second time in the week commencing 2nd May 2022** and agree upon a time slot for regular project meetings. If the project is with a research group in the faculty of science then the project supervisor is the person with whom specific, specialised aspects of the project should be discussed on a weekly basis. If the project is with an external organisation, then meetings will probably be less frequent, but students will also be assigned to a member of the academic staff who has expertise in the project area and who will be available locally for discussions.

The role of the project supervisor is to guide students in the conduct of the research project. However, at this more advanced masters level, we will be assessing the performance of students as professional scientists, including their originality, organisation and effort. Students must be prepared to take responsibility for making their own project decisions, informed as necessary by discussions with experienced researchers; students should not expect decisions to be made for them.

Project supervisors participate in the marking of the final project report and will provide more general feedback to the module convenor on the conduct, progress and balance of effort of the project by the student team.

Module Supervisor

In addition to a project supervisor, each project is assigned a module supervisor who is one of the MLiS course leaders (Prof Juan Garrahan, Dr Adam Moss, Dr Steven Bamford, Dr Simon Dye). This staff member will be responsible for monitoring the general progress of the project and ensuring its smooth running. They will take a keen interest in the project, but they may not necessarily have the specialised knowledge of the project supervisor. The module supervisor will be available for consultation throughout the project to facilitate the smooth and ongoing progress of the work. In order to ensure marking standards are consistent across all projects, the course leaders will be responsible for moderating the project report marks awarded by the project supervisor and producing the final project mark.

Commercial sensitivity and Intellectual Property

A final but important point is regarding commercial sensitivity. Any projects that involve commercially active companies or collaborations between academic schools and commercially active companies must treat any sensitive material being handled in the strictest confidence. Leaking of material to a third party, even inadvertently, may lead to serious consequences including, in extreme cases, prosecution. Where appropriate, students are bound by a Disclosure Agreement between the University and the sponsoring company or organisation. Also, each student is responsible for their own behaviour and should be aware that they are representing the University.

Timetable and Assessment

The period up until the 9th April is dedicated to finalising research topics and assigning these to a supervisor. Students are only expected to start the hands-on research element of their project from **Monday May 16th 2022 onwards**). This process, particularly planning and background reading, may begin earlier once a project has been formally accepted. If students have examinations in the summer exam period then full-time project work can be delayed, although it is strongly recommend that part-time work on the project commences. It is expected that approximately 37.5 hours per week will be spent per student on the project once in full-time mode.

Progress report: In order to ensure that projects are on-track, student teams must submit a brief progress report mid-way through the project. This report is also intended to directly benefit students, to serve as an indicator of what remains to be done in the time left. The report should be word processed and approximately one side of A4 in length. It should contain a statement of the topic under investigation, a brief summary of progress so far and a plan for the rest of the project. One progress report for each project must be emailed to the **project supervisor AND the module convenor** during

the week ending **Friday 1st July 2022**. The project supervisor will give feedback on the progress report at the next project meeting.

Final report submission: Project reports must be electronically submitted as a PDF file to the MLiS project Moodle page by **1pm on August 19th 2022**. Alongside this, students should also submit a zipped file of all code developed during the project. **Reports submitted late without good cause will receive the standard university penalty of 5 marks per working day from the 100 total marks available**. If extra time is required due to extenuating circumstances, then a standard university extenuating circumstances form must be submitted before the project deadline.

Timetable Summary

Date	***Activity***	***Notes***
08/4/2022	Deadline for submission of project proposal or, if choosing from pre-selected list, project topic, to module convenor.	Email to the module convenor. For proposals, include supervisor details.
08/4/2022	Deadline for project partner requests (otherwise assignment is random).	Both parties must email the module convenor.
18/4/2022	Project allocations announced	
02/5/2022	Deadline for completing first meeting with project supervisor.	
16/5/2022	Deadline for completing second meeting with project supervisor.	Arrange weekly meeting slot, discuss first steps.
16/5/2022	Full-time project work begins in week commencing 16th May 2022.	Students can start earlier once project approved and supervisor met.
01/7/2022	Deadline for submission of progress report.	Email to project supervisor **and** module convenor during week ending 1st July 2022.
19/8/2022	Deadline for submission of project reports.	Submit PDF and zipped code via the Moodle MLiS project page. Penalty of 5 marks per day for late submission.

Preparation of the Final Report

The report should comprise around 10,000 words with a hard upper-limit of 12,000 words including the abstract, main text, figures and bibliography. It should span no more than 50 sides of A4 paper size and should be written in a 10pt or 11pt font. Text used in figures and diagrams should be at least the size of the main body text.

The final report is a valuable document recording an intense period of original work and so every effort should be made to ensure it has a professional presentation, is well-structured and has many citations to previous related research studies. The report should be written in the style of a scientific research paper organised into sections and

sub-sections as required. Appendices should be used sparingly and only where relevant; these will not be counted in the word-limit.

The structure of the report will depend on the nature of the work conducted and so there is a degree of flexibility in how this will be presented, but it should contain the following key elements:

- *Abstract* – The report should begin with a single paragraph containing an overview of the aim of the report, the methods used, the results found and a discussion of the significance of the results.
- *Introduction* – The first section should introduce the background science with appropriate references to already published scientific literature including books and journal papers. This referencing can take a number of forms; for example, references to original material where experimental techniques are described, mathematical expressions are derived or computational techniques have been developed. Additionally, it will encompass a 'literature review' outlining the current state of knowledge in the published literature; what papers, past and recent, have been published that are directly relevant to the project? This should demonstrate that students have read widely and have a good acquaintance with the scientific literature. The introduction should conclude with a brief overview of how the project will further existing studies. It is typical for introductions to account for 10-15% of the report word count.
- *Methods used* – An account of how the work was conducted is an essential part of any report or scientific research paper as this allows others to repeat or build upon your work. This account should include any scientific, mathematical and computational techniques used and any external datasets that have been used. It could be written as a single section with sub-sections or separate sections as is appropriate. A good methodology strikes a sensible compromise between containing enough detail to allow another researcher to repeat what has been done to a good level of fidelity with being concise and not containing unnecessary detail.
- *Results* – A separate section should be dedicated to describing the outcomes of the work. Results should be included in the most concise but clear way possible. Plots and graphs are typically a more efficient way of presenting results but sometimes tables are a more viable option. Large tables of raw data or several plots showing minor variations in experimental set up should be avoided; plotting several lines onto one graph can be a nice way of keeping results compact so long as lines are discernible (i.e. not too many and using obviously different line styles). A good results section will also include an analysis of any uncertainties involved.
- *Discussion* – The discussion section should provide an interpretation of the results in a scientific context. Here, results should be compared and contrasted with other studies and there should be a discussion about any limitations of the work carried out and how this might be improved with further research.
- *Conclusion/Summary* – Often combined with the discussion section, the conclusion should summarise the findings and conclusions of the paper. This should not contain any new material and should be brief. A researcher wanting a quick idea about the work carried out will typically read the abstract and the conclusion/summary section.
- *References* - The sources of scientific statements, derivations, mathematical expressions, external datasets, software libraries etc, must always be cited in the main text body of the report and fully referenced in the bibliography section situated at the end of the report (but before the appendices which come last). In a project of this nature, references to peer reviewed journal articles and papers as well as books should be used as far as possible and always in preference to

web sites; it is not acceptable to refer to web-sites as a primary source of information apart from when necessary, such as referring to software repositories or online datasets. Furthermore, it is important that students identify published material from their own work; citations will help achieve this (see further comments regarding plagiarism below). It is not always necessary to include the full title in the citation, but it can be helpful to the reader if it is included. Different journals have slightly different citation formats; students should adopt a particular format consistently throughout the report.

A style guide entitled '*Guidance on Writing and Formatting Reports and other Written Physics Articles*' has been produced by Dr Mellor. This gives guidance on preparing written work in all years of study in the School of Physics & Astronomy. Students should read this document and adopt the guidance in preparing their submitted written work for this project module. A copy is available alongside this handbook on the MLiS project Moodle page. Students should also gain an appreciation of the standards required in writing a report by reading work published in peer-reviewed journals as part of their research.

Important note: the project supervisor may be approached for general guidance on the structure and content of the report but it is not reasonable to expect them to read a draft of the report and comment in detail.

Professional Conduct, Teamwork and Plagiarism

Professional conduct, both in carrying out the project work and in composing written work is expected at all times.

The University takes plagiarism, or any other form of cheating, very seriously and anyone suspected of committing an offence will be reported to the University Academic Offences Committee. Students are encouraged to refer to the section on "Academic Offences" in the University Regulations. Particularly important: a) all experimental and computational results must be original; fabrication of results is a very serious offence indeed; b) it is essential to refer correctly to original sources of information; all verbatim and all derived material should be adequately referenced; c) students claiming another person's work as their own is a very serious professional offence for all scientists. All reports will be run through the Turnitin software for detecting plagiarism.

It is important not to confuse plagiarism and teamwork particularly since students work in teams. The teamwork in the conduct of the project is a very important aspect of the training. However, the report is written and assessed individually and so it should represent the work of the individual student alone.

Marking Standards

Guides to the criteria applied when marking the report are given below and form the basis on which marks will be allocated. Items marked in italics are particularly important.

Presentation

Credit will be given for:

- *A clear logical argument.*
- *Accurate grammatical constructions.*
- Clear and well-chosen illustrations.
- Adequate captions for all figures which allow them to be understood independently and to be referenced from the text.
- *A sensible partition into sections, including a general introduction and a discussion.*
- Choice of clear print format and layout of text.
- *Use of simple, clear English with the minimum of jargon.* All acronyms must be defined.
- Comprehensive list of references which are cited from the text.
- Sensible use of appendices where appropriate for non-essential mathematical detail etc.
- Clear and informative abstract.
- Indication of future work.
- Matching to allocated word/time limits.

Science

Credit will be given for:

- *A demonstrable understanding of the fundamental science involved.*
- An appreciation of the nature of the problem.
- A realistic approach to the problem.
- *Imagination and flexibility of approach.*
- Evidence for independence of thought and action.
- *Evidence for industry and hard work.*
- Sensible division of labour within the team and evidence of planning.
- Ability to work in a team.
- Decisiveness.
- *Good experimental/theoretical technique.*
- *Critical analysis of theoretical and experimental results.*
- An awareness of safety issues.

The following tables summarise some of the attributes that characterise the different grades. Feedback for the various components will be provided in written form.

Effort and organisation

90	Detailed feedback will be provided
80	The project shows evidence of an outstanding level of industry and organisation, with time and effort well allocated and a very high volume of work produced.
70	The project shows evidence of a very good level of industry and organisation, with time and effort well allocated and a high volume of work produced.
60	The project shows evidence of a satisfactory level of industry and organisation. A reasonable allocation of time and effort has led to a satisfactory amount of work done.
50	The project shows evidence of insufficient industry and organisation. Time and effort have been poorly allocated.
40	The project shows evidence of an insufficient level of industry and organisation. Very little thought has been given to allocation of time and effort.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Initiative and independence

90	Detailed feedback will be provided
80	You have shown outstanding initiative and independence in conducting the project and demonstrated creativity by exploring novel directions.
70	You have demonstrated very good initiative and independence, at times probing new avenues.
60	You have demonstrated a satisfactory level of initiative and independence in conducting the project and overcoming difficulties.
50	You have demonstrated little initiative and independence in conducting the project.
40	You have demonstrated insufficient initiative and independence in conducting the project.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Quality of outputs and analysis

90	Detailed feedback will be provided
80	The work constitutes an outstandingly professional piece of research. The outputs and analysis are of outstandingly high quality, at a standard suitable for publication.
70	The work constitutes a competent and professional piece of research, and the outputs and analysis are of very good quality. They could be of a standard suitable for publication, subject to confirmation by research staff in the group.
60	Most aspects of the work meet the required professional standards. The research outputs and analysis are of satisfactory quality, and constitute useful results.
50	Aspects of the work fall short of the required professional standards. The research outputs and analysis are of poor quality, but perhaps constitute some useful results.
40	The work falls well short of the required professional standards. The research outputs and analysis are of inadequate quality.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Understanding shown of theory, method and techniques

90	Detailed feedback will be provided
80	Your writing displays an outstanding level of understanding of the theory, method and techniques relevant to the project.
70	Your writing displays a very good understanding of the theory, method and techniques relevant to the project.
60	Your writing displays a satisfactory level of understanding of the theory, method and techniques relevant to the project.
50	Your writing displays a poor understanding of the theory, method and techniques relevant to the project.
40	Your writing displays insufficient understanding of the theory, method and techniques relevant to the project.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Discussion and conclusions

90	Detailed feedback will be provided
80	The discussion is outstanding, making full use of the research outputs and relevant literature and leading to highly developed, insightful conclusions that are set within the context of the research area.
70	The discussion is very good, making use of the research outputs and relevant literature and leading to insightful conclusions.
60	The discussion is satisfactory, and some progress has been made in developing conclusions.
50	The discussion in the project report is poor, with only elementary progress towards relevant conclusions.
40	The discussion and conclusions in the project report are inadequate.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Language, figures and structure

90	Detailed feedback will be provided
80	An outstandingly well-written and well-structured report, which conveys its message efficiently and concisely. The figures are outstandingly clear and informative. The spelling, grammar and formatting are near perfect.
70	A very well-written report, which has very good structure and is easy to read. The figures are clear and informative. The spelling, grammar and formatting are very good.
60	The writing in this report is satisfactory, but with some errors in spelling and/or grammar. The logical structure is not always ideal, restricting the efficiency with which the message is conveyed. The figures are mostly clear and informative.
50	A poorly written report with frequent errors in spelling and/or grammar. It is poorly organised with little logical flow or clear structure. The use of figures is poor.
40	The writing in this report is inadequate. It is insufficiently well organised and has little logical structure. The figures are inadequate.
30	Detailed feedback will be provided
20	Detailed feedback will be provided
10	Detailed feedback will be provided
0	Detailed feedback will be provided

Project Health and Safety

Although it is not anticipated that projects will involve any laboratory work nor working with machinery or direct hazards, all students must individually sign and submit a risk assessment form prior to starting project work. Projects will not be given formal approval until a risk assessment form has been signed and submitted by all members of the project team.

A risk assessment form specific to projects that only involve use of a computer (i.e. no laboratory work) has been prepared and put onto Moodle. This form covers the use of display screen equipment. Students must sign this form and hand it back to the module convenor. Any projects that involve work that goes beyond the use of standard display screen equipment (i.e. anything other than use of a computer or use of a chair and desk) must consult the module convenor for further advice prior to starting the project.

A copy of the UK Health and Safety Executive's display screen equipment checklist is available on Moodle alongside the risk assessment form. This will be covered in an introductory talk given to students in the spring semester; all students must ensure they are familiar with the recommendations on the checklist.

More detailed and specific information (which is regularly updated) can be found in the School of Physics and Astronomy's health and safety online policy document: http://www.nottingham.ac.uk/~ppzphy/safety/Policy.htm.