REIT6811 Research Methods (Semester 1, 2025)
Assignment 1

Aim: The aim of this assignment is for you to research sources, work effectively with a range of reference formats and metrics relating to researchers and publications. This assignment builds on the databases and practices you have learned in lectures and tutorials. However, you will need to search bibliometrics beyond those given in class. Use complete sentences with clear descriptions and justifications. For all quality measures, state the source, date of access, and provide a correctly formatted reference.

Submission: Submission is via both Gradescope and Turnitin via Blackboard.
Ensure your submissions are text-based using a standard font (e.g. Times Roman, Arial, Calibri, or LaTeX default fonts). Your report should be in .pdf format and named according to the format `a1-REIT6811-[SID].pdf`, where SID is your student ID.

Referencing: Use IEEE format except where stated. If using concepts or text from generative AI, ensure you cite their usage. Further notes on referencing and plagiarism are provided at the end of this Assignment, including Appendix A on how to cite Generative AI. Include all your references in a reference manager as the last question requires an Appendix with an image of your reference manager or a copy of your Bibtex file.

Marking: Answer all 9 questions in this assignment with a total worth of 45 marks. Assignment 1 contributes 20% to your final grade. In the following questions, answers to some sections may be easy to look up online. However, marks will only be given if the answer is both accurate and appropriately referenced to an authoritative source. Explain your logic for why those sources are appropriate.

Full marks for Question 1 will only be given if Questions 2-9 are consistent with your Q1 answer by demonstrating high quality referencing in both in-text citations and in the Reference section. As seen in the class demonstrations, sometimes you may need to adjust how a reference is automatically formatted using a reference manager. For this assignment, if you do need to manually edit your in-text citations or references, you can add an explanation of the issue at the end of your answer to that question. The important pedagogic lessons in question 1 are (1) to use effective tools; (2) but also be aware of their limitations and how you are working with them; and (3) communicate your reasoning to your reader if an explanation is needed.

Due Date: 4pm, Monday 7 April 2025 [AEST]

Question 1. Referencing managers (2 marks)

(a) What word processor and reference manager are you using for this assignment? (see also Question 9 where you need to demonstrate the references in your reference manager) (1 mark)

(b) Will you use cite-as-you-write when preparing this assignment? Why or why not? (1 mark)

Question 2. ArXiv Pre-prints (12 marks)

Many research communities have ways to circulate manuscripts before they are peer reviewed. These include un-refereed sites such as arXiv.org, mailing lists, google groups, or direct sharing of manuscripts between individual researchers. Peer review is an independent assessment of a manuscript by experts in the field. This is typically through a journal or refereed conference.

For this question, you need to do your own background research by finding and reading the arXiv.org "About arXiv" website and searching for other sources of useful information. Cite all your sources and use complete and correctly formatted references (in IEEE format, including the dates of access if required). Collate all your references in one list at the end of the assignment.

(a) What role does arXiv.org play in disseminating scholarly research and which disciplines deposit articles in arXiv.org? Cite your source(s). (<100 words, 1 mark)

(b) What are the key differences between papers in an e-print archive like arXiv, and in an online conference proceedings repository like AAAI Conference Proceedings? Cite your source(s). (<150 words, 1 mark)

(c) What are the key differences between the types of articles (conference publications, journal articles, editorial letters, pre-prints etc.) included in citation databases like Google Scholar and Web of Science? Cite your source(s). (<150 words, 1 mark)

(d) arXiv papers are labeled with an identifier. For example, `2006.10204` is the arXiv identifier for the following URL: `https://arxiv.org/abs/2006.10204`. Citing your source(s), explain what the numbers refer to in the latest version of the identifier. (<50 words, 1 mark)

The following question parts (e) – (h) refer to the article: **Cao, Z., et al. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. arXiv.org** It was first deposited in arXiv.org and then published later in a high-profile journal or conference.

(e) Find the preprint and the publication online and obtain the pdfs for each version. Enter the two versions of the article in your reference manager. Note that the reference above is incomplete and not in a standard format. Ensure that you find any missing details, and correctly format the article title (using sentence capitalization) and the journal or conference title (using title capitalization).

Provide an in-text citation for each version of the publication here and ensure that the full citation is correctly formatted in the reference list at the end of the assignment. (<50 words, 1 mark)

(f) In your own words, summarise what the paper is about in less than 50 words, based on the abstract. Use appropriate paraphrasing techniques, and specify which paraphrasing methods you have used from the following: (3 marks)

- ☐ Synonyms
- ☐ Changing voice e.g. passive to active OR changing word order
- ☐ Changing a clause to a phrase
- ☐ Changing parts of speech
- ☐ Changing conjunctions / linkers

(g) Fill in the table below, reporting exactly how many times the article by Cao et al. has been cited on Google Scholar and Web of Science on the date(s) you accessed them. (2 marks)

Article version	Number of Google Scholar citations	Number of Web of Science citations	Date(s) of access
Peer-reviewed			

(h) For this question you need to find the list of papers that cited Cao et al. ordered using the "Sort by relevance" option on the left hand side of Google Scholar. Determine where the ten most relevant citing papers were published or if they were preprints. Fill in the table below with the publication name. For example, if the first paper in the citing list was published in the journal Bioinformatics, then you would add that name to the 'Journals' column.

Using the information in your table, in one sentence, summarise the types of publications that the first ten references to Cao et al. in Google Scholar are found. Does this support your answer in 2(c) fully, partially, or not at all? (<100 words, 2 marks)

	Journals	Refereed Conferences	Preprints and unrefereed conferences	Date of access
Title of publication of citing articles				
Total Count				

Question 3. Referencing and Plagiarism **(6 marks)**

Below is a passage in an (imaginary) deliverable that a student has submitted to answer the question: **"Why does Simplicity Not Imply Accuracy?"**, and further below are excerpts from one of the original sources they used.

Note that the original source utilised ACM style referencing, while the student used an APA style in-text citation. In your answer, use IEEE style, where necessary.

Student's answer:

> Simplifying things doesn't necessarily make them accurate. For example, support vector machines can effectively have an infinite number of parameters without overfitting. In contrast, a simple function like sign(sin(ax)) can effectively distinguish between an infinite number of points on the x-axis, despite having only one parameter (Domingos et al., 2012).

Original source:
Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78-87. [excerpts from p. 86].

Occam's razor famously states that entities should not be multiplied beyond necessity. In machine learning, this is often taken to mean that, given two classifiers with the same training error, the simpler of the two will likely have the lowest test error. (. . .)

Another counterexample is support vector machines, which can effectively have an infinite number of parameters without overfitting. Conversely, the function sign(sin(ax)) can discriminate an arbitrarily large, arbitrarily labeled set of points on the x axis, even though it has only one parameter [23].

Excerpt from the original source's references
23. Vapnik, V.N. *The Nature of Statistical Learning Theory*. Springer, NY, 1995.

Answer the following two questions about this.

(a) Identify any problems with non-originality, plagiarism, and/or referencing in the student's answer. List the problems, being very specific about where the problems are. (4 marks)

(b) State, using examples, what the student should have done to avoid the problems. (2 marks)

Question 4. Quality metrics for journals and journal articles (5 marks)

Question 4 and Question 5 refer to the two articles cited below. When answering questions, remember to include the databases you used and the dates of access, e.g., h-index = 42 from [database] accessed on [date].

Article 1

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

Article 2

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, 128(2), 336–359. https://doi.org/10.1007/s11263-019-01228-7

(a) The two citations above are in different formats, IEEE and APA style respectively. What are key differences between these two formats in terms of their in-text citations, and the ordering of the reference list? Show an example of an in-text citation for Article 1 using IEEE style and APA style formatting. (2 marks)

(b) Determine and report the h5-Index of the last author for Article 2 based on Scopus, Google Scholar and WoS. Is the h5-Index of the author the same or different across these three databases? Give a suitable explanation for your answer. (2 marks)

(c) Choose one metric suitable for assessing the quality or impact of Article 2. Justify your choice and report its value. (1 mark)

Question 5. Teams of authors (3 marks)

The authorship list can reveal important information about a paper. Fill in the table below answering the following questions for the published version of Article 2:

i) Who is the corresponding author?

ii) What were the academic or industry affiliations of the authors at the time of publication?

iii) What were their positions (e.g. professor, MPhil student, PhD student)?

Author order	Name	Affiliation	Academic or professional position	Corresponding author?
Author #1				
Author #2				
Author #3				
Author #4				
Author #5				
Author #6				

Question 6. Quality metrics for conferences and conference papers (6 marks)

Imagine that your supervisor has asked you to publish in a conference and is considering either "International Joint Conference on Neural Networks" or "IEEE Conference on Computer Vision and Pattern Recognition".

(a) What are two conference-level metrics that will help to determine the quality of the above-mentioned conferences? Justify your choice and report their value. If a metric you chose is not available for both conferences, choose another metric or justify what the absence of that metric means for the quality of the conference. (3 marks)

(b) What are the ANZSRC Fields of Research (FoR) codes in the CORE database used for the above-mentioned conferences from 2023? Report all the values in the table below and provide an accurate citation to the database used.

(3 marks)

Conference	Metric 1	Metric 2	ANZSRC FoR codes
International Joint Conference on Neural Networks			
IEEE Conference on Computer Vision and Pattern Recognition			

Question 7. Research Ethics (6 marks)

(a) Imagine you are collaborating with a team on a machine learning project using medical records to develop predictive models. Discuss the ethical considerations surrounding data collection, ownership, confidentiality, and sharing that must be addressed.

(4 marks)

(b) When evaluating the developed predictive model during your thesis, you find that some of the data points don't fit the expected trend, and you get significantly better performance by removing those data points. Can these data points be removed for publication, and why or why not? What are appropriate steps to do in this situation? (2 marks)

Question 8. Conflicts of Interest (2 marks)

Dr. Williams is conducting research on renewable energy technologies and has recently been awarded a research grant from a major oil company. The company has publicly stated its interest in influencing policy on energy regulations. Dr. Williams is now preparing a paper for a high-profile journal that analyses the environmental impact of renewable energy sources.

(a) Identify the conflict of interest (COI) that may arise from Dr. Williams' funding source and any potential ethical concerns related to this funding arrangement for Dr. Williams' research outcomes. (<100 words, 1 mark)

(b) How can Dr. Williams manage this conflict of interest to maintain research integrity? (<100 words, 1 mark)

Question 9. Demonstrating your use of a reference manager (3 marks)

Add a Reference section in IEEE format which includes all the references in your assignment. Note that you need to cite answers using concepts or text provided by generative AI just as you would any other source (see Academic Misconduct and Appendix A for details of the format requirements).

Add an Appendix to demonstrate the list of articles in your reference manager using either a screen image ("screenshot"), or a copy of your Bibtex file.

Academic Misconduct

The University defines Academic Misconduct as involving "a range of unethical behaviours that are designed to give a student an unfair and unearned advantage over their peers." UQ takes Academic Misconduct very seriously and any suspected cases will be investigated through the University's standard policy (`https://ppl.app.uq.edu.au/content/3.60.04-student-integrity-and-misconduct`). If you are found guilty, you may be expelled from the University with no award.

All sources of information should be attributed, and references should use IEEE format except where stated. Use quotation marks when citing specific sentences. Do not plagiarise. In other words, do not directly copy your answers from other sources, such as lecture notes, material written by others, material written by you for other purposes or courses, or any other source. See university and course material on how to avoid plagiarism. Acknowledge the source of all ideas that you use in your work, using methods discussed in the lectures and tutorials.

Generative AI: If you have utilised Generative AI tools such as ChatGPT, you must clearly cite any use of generative AI in each instance. To reference your use of AI:

- See: https://guides.library.uq.edu.au/referencing/chatgpt-and-generative-ai-tools
- Describe how the AI was used in the reference, e.g.:
  "EMI Instrumentation Presentation" prompt; initial slide outline for presentation on EMI shielding in MRI ChatGPT, Apr 2023 version, OpenAI, 2 Feb. 2024, `chat.openai.com/chat`.
- If referencing specific text, do not copy the generated material – just like a journal article or textbook you should write the generated material in your own words, unless quoting (which you should always do sparingly).

Failure to reference use of generative AI tools constitutes student misconduct under the Student Code of Conduct.

It is the responsibility of the student to ensure that you understand what constitutes Academic Misconduct and to ensure that you do not break the rules. If you are unclear about what is required, please ask.

It is also the responsibility of the student to take reasonable precautions to guard against unauthorised access by others to his/her work, however stored in whatever format, both before and after assessment.

All submitted files will be subject to electronic plagiarism detection and misconduct proceedings will be instituted against students where plagiarism or collusion is suspected. If you collude to develop your report, you will be caught.

For more information, please consult the following University web pages:

- Information regarding Academic Integrity and Misconduct:
  - https://my.uq.edu.au/information-and-services/manage-my-program/student-integrity-and-conduct/academic-integrity-and-student-conduct
  - http://ppl.app.uq.edu.au/content/3.60.04-student-integrity-and-misconduct
- Information on Student Services:
  - https://www.uq.edu.au/student-services/

Late submission

Students should not leave assignment preparation until the last minute and must plan their workloads to meet advertised or notified deadlines. It is your responsibility to manage your time effectively.

Assessment submissions received after the due time (or any approved extended deadline) will be subject to a late penalty of 10% per 24 hours of the maximum possible mark for the assessment item.

In the event of exceptional circumstances, you may submit a request for an extension. You can find guidelines on acceptable reasons for an extension here https://my.uq.edu.au/information-and-services/manage-my-program/exams-and-assessment/applying-extension.

All requests for extension must be submitted on the UQ Application for Extension of Assessment form at least 48 hours prior to the submission deadline.

If you submit your assignment before the deadlines, then Turnitin does not allow any more submissions after the deadline. If you have not submitted before the deadline, then Turnitin will allow just one submission after the deadline. If you have been granted an extension and have trouble submitting, email the course account reit6811@eecs.uq.edu.au for help. Ensure you include the required information in your email (see the Ed Discussion board notice from the first week of semester).

Appendix A: Generative AI Usage

Please use the following Table A1 and statement to acknowledge your use of generative AI in this assignment. Include this table at the end of your report or slides, near your references section. The first two rows (in orange) are example only, and should be removed before use.

I acknowledge the use of generative AI tools in completing this assignment. Details of which tools were used and how they were used are provided in the table below, along with appropriate in-text and full references. I take responsibility for critically evaluating and integrating the AI-generated content, and ensuring it adheres to academic integrity standards.

AI Model Used and date	Table A1. Generative AI usage (tick all that apply)								
	Language Translation	Grammar/Style	Planning/Drafting	Research/ Background Information	Content Creation text	Content Creation visual	Content Creation code	Feedback	Other (provide details)
ChatGPT4o 30/03/25		✓	✓	✓				✓	
Midjourney 30/03/25						✓			