UNSW Course Outline

BINF9020 Computational Bioinformatics - 2024

Published on the 25 Aug 2024

General Course Information

Course Code : BINF9020
Year : 2024
Term : Term 3
Teaching Period : T3
Is a multi-term course? : No
Faculty : Faculty of Engineering
Academic Unit : School of Computer Science and Engineering
Delivery Mode : In Person
Delivery Format : Standard
Delivery Location : Kensington
Campus : Sydney
Study Level : Postgraduate
Units of Credit : 6

Useful Links

Handbook Class Timetable

Course Details & Outcomes

Course Description

This course explores the computational basis of bioinformatics and complements and extends understanding of bioinformatics tools and resources acquired in other courses, with a focus on the analysis of complex biological datasets and the application of mathematical and

computational methods to problems in modern life science. Example domains include sequence analysis, gene expression and function, networks of interaction, and systems modelling.

Specific topics include: Algorithms and representations in DNA and protein sequence analysis: string matching and alignment, tree building methods, hidden Markov models and other probabilistic representations, genome assembly and mapping. Computational representations in systems biology including Boolean and Bayesian networks. Optimisation and machine learning approaches used in bioinformatics.

Course Aims

The course aims to reinforce students' understanding of core bioinformatics engineering principles, as well as building their capabilities in core bioinformatics engineering practice. In particular, the course will improve students' ability to apply computational methods of analysis and modelling in modern biology, building on the use of bioinformatics tools acquired in prerequisite courses, and relating to research and industrial contexts. Development of software skills in bioinformatics complements strengthening of algorithmic knowledge combined with practical applications to biological data.

This course assumes that students have a background in bioinformatics and programming skills. It is part of the Bioinformatics major in the Master of Information Technology program.

Relationship to Other Courses

Assumed knowledge acquired in the courses BINF2010 or COMP2041, or equivalent.

Course Learning Outcomes

Course Learning Outcomes
CLO1 : Reproduce, explain and apply the major methods of pairwise and multiple sequence analysis in bioinformatics
CLO2 : Define key computational and statistical concepts in bioinformatics approaches to pairwise and multiple sequence analysis.
CLO3 : Define Hidden Markov models, describe their associated algorithms, and discuss their application to biological sequence analysis.
CLO4 : Define and apply machine learning methods of clustering and classification in analysis of genome-scale data.
CLO5 : Express and apply key concepts from graph theory to model biological networks.
CLO6 : Use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology.
CLO7 : Research and present on a range of open-source software tools and techniques available to bioinformatics software developers.

Course Learning Outcomes	Assessment Item
CLO1 : Reproduce, explain and apply the major methods of pairwise and multiple sequence analysis in bioinformatics	• Programming assignment in 3 parts: sequence alignment
CLO2 : Define key computational and statistical concepts in bioinformatics approaches to pairwise and multiple sequence analysis.	• Programming assignment in 3 parts: sequence alignment
CLO3 : Define Hidden Markov models, describe their associated algorithms, and discuss their application to biological sequence analysis.	• Final Exam
CLO4 : Define and apply machine learning methods of clustering and classification in analysis of genome-scale data.	• Homework: systems biology
CLO5 : Express and apply key concepts from graph theory to model biological networks.	• Homework: systems biology
CLO6 : Use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology.	• Final Exam • Homework: systems biology
CLO7 : Research and present on a range of open-source software tools and techniques available to bioinformatics software developers.	• Group presentation • Final Exam

Learning and Teaching Technologies

Moodle - Learning Management System | Microsoft Teams

Assessments

Assessment Structure

Assessment Item	Weight	Relevant Dates
Programming assignment in 3 parts: sequence alignment Assessment Format: Individual	20%	
Group presentation Assessment Format: Group	16%	
Homework: systems biology Assessment Format: Individual	14%	
Final Exam Assessment Format: Individual	50%	

Assessment Details

Programming assignment in 3 parts: sequence alignment

Assessment Overview

A programming assignment focusing on the dynamic programming alignment algorithm.

Two weeks is allocated for development of each programming component, and one week for collecting data and report preparation.

Marking is with respect to a rubric and feedback will be provided with the online assessment.

Course Learning Outcomes

- CLO1 : Reproduce, explain and apply the major methods of pairwise and multiple sequence analysis in bioinformatics
- CLO2 : Define key computational and statistical concepts in bioinformatics approaches to pairwise and multiple sequence analysis.

Assignment submission Turnitin type

Not Applicable

Generative AI Permission Level

Not Applicable

Generative AI is not considered to be of assistance to you in completing this assessment. If you do use generative AI in completing this assessment, you should attribute its use.
For more information on Generative AI and permitted use please see here.

Group presentation

Assessment Overview

Open-source bioinformatics software environments

Research and present on a selected topic. Marks will include a peer-assessed component.

Presentation time of approximately 30 minutes per team, report of about 5 pages, depending on team size.

Marking is with respect to a rubric and feedback will be provided with the online assessment.

Course Learning Outcomes

- CLO7 : Research and present on a range of open-source software tools and techniques

available to bioinformatics software developers.

Detailed Assessment Description

Specification for presentation and report is released on Moodle.

Assessment Length

Team slides plus written report of approximately 1-2 pages per member

Assignment submission Turnitin type

Not Applicable

Generative AI Permission Level

Not Applicable

Generative AI is not considered to be of assistance to you in completing this assessment. If you do use generative AI in completing this assessment, you should attribute its use.
For more information on Generative AI and permitted use please see here.

Homework: systems biology

Assessment Overview

A two-part self-directed data analysis and programming task on methods from lectures.

One week for each part of the homework is allocated for completing exercises and answering questions.

Marking is with respect to a rubric and feedback will be provided with the online assessment.

Course Learning Outcomes

- CLO4 : Define and apply machine learning methods of clustering and classification in analysis of genome-scale data.
- CLO5 : Express and apply key concepts from graph theory to model biological networks.
- CLO6 : Use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology.

Detailed Assessment Description

Specification for each part is released on Moodle.

Assessment Length

Program code plus short report

Assignment submission Turnitin type

Not Applicable

Generative AI Permission Level

Not Applicable

Generative AI is not considered to be of assistance to you in completing this assessment. If you do use generative AI in completing this assessment, you should attribute its use.
For more information on Generative AI and permitted use please see here.

Final Exam

Assessment Overview

A formal exam in the exam period covering the lecture content of the course.

A two hour exam comprising written and multiple choice answers based on problem solving.

Marking is with respect to a rubric.

Course Learning Outcomes

- CLO3 : Define Hidden Markov models, describe their associated algorithms, and discuss their application to biological sequence analysis.
- CLO6 : Use and explain selected engineering approaches to the representation, execution and evaluation of models of biological processes and their application in systems biology.
- CLO7 : Research and present on a range of open-source software tools and techniques available to bioinformatics software developers.

Assessment Length

Two hour final exam

Assignment submission Turnitin type

Not Applicable

Generative AI Permission Level

No Assistance

This assessment is designed for you to complete without the use of any generative AI. You are not permitted to use any generative AI tools, software or service to search for or generate information or answers.
For more information on Generative AI and permitted use please see here.

General Assessment Information

Grading Basis

Standard

Requirements to pass course

Course mark is the sum of marks for assessed course components with a mark of 50 or greater required to pass.

Course Schedule

Teaching Week/Module	Activity Type	Content
Week 1 : 9 September - 15 September	Lecture	Substitution matrices
	Lecture	Global alignment
Week 2 : 16 September - 22 September	Lecture	Local alignment and sequence search (part 1)
	Lecture	Local alignment and sequence search (part 2)
	Tutorial	Substitution scoring
	Assessment	Assignment part 1
Week 3 : 23 September - 29 September	Lecture	Multiple alignment and phylogeny (part 1)
	Lecture	Multiple alignment and phylogeny (part 2)
	Tutorial	Sequence alignment
Week 4 : 30 September - 6 October	Lecture	Family representation and HMMs (part 2)
	Tutorial	Sequence search and sequence assembly
	Lecture	Family representation and HMMs (part 1)
	Assessment	Assignment part 2
Week 5 : 7 October - 13 October	Lecture	Public holiday
	Lecture	Sequence assembly and genome informatics
	Tutorial	Public holiday, rescheduled to week 4
	Assessment	Assignment part 3
Week 6 : 14 October - 20 October	Other	Flexibility week
Week 7 : 21 October - 27 October	Lecture	Gene expression and functional genomics (part 1)
	Lecture	Gene expression and functional genomics (part 2)
	Laboratory	Expression data analysis in Python (HW part 1)
Week 8 : 28 October - 3 November	Lecture	Biological network analysis (part 1)
	Lecture	Biological network analysis (part 2)
	Laboratory	Network analysis in python (HW part 2)
	Assessment	HW due
Week 9 : 4 November - 10 November	Lecture	Systems biology (part 1)
	Lecture	Systems biology (part 2)
	Assessment	Open source software (presentations)
Week 10 : 11 November - 17 November	Lecture	Systems biology (part 3)
	Lecture	Systems biology (part 4)

Attendance Requirements

Students are strongly encouraged to attend all classes and review lecture recordings.

General Schedule Information

Lectures in Term 3 will be **In Person** and will be held on **Monday** (Lecture 1) and on **Wednesdays** (Lecture 2). Each week has a three hour scheduled tutorial/lab times on **Monday**.

Please note: we will use **only** the following tutorial/lab times:

Weeks 2-4: two 1 hour tutorials and one 2 hour tutorial

Weeks 7-8: a 2 hour lab

Week 9: 3 hour group presentation slot

However, remember times may be subject to change!

All schedule details available from the Moodle course page.

Course Resources

Prescribed Resources

There is no required textbook for this course. However, please see the recommended resources to access introductory or deeper background reading and for further study on some of the course topics.

Recommended Resources

- "Bioinformatics Algorithms: An Active Learning Approach (Volumes I and II)" by Compeau and Pevzner, Active Learning Press (2015)
- "Biological Sequence Analysis – Probabilistic models of proteins and nucleic acids" by Durbin, Eddy, Krogh and Mitchison, Cambridge University Press (1998)
- "Introduction to Systems Biology – Design Principles of Biological Circuits (2nd Edition)" by Uri Alon, CRC Press, 2019.
- "Systems Biology – A Textbook (2nd Edition)" by Edda Klipp *et al.*, Wiley-VCH, 2016.
- "Networks (2nd Edition)" by Mark Newman, Oxford University Press, 2018.
- "Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry and Engineering (2nd Edition)" by Steven Strogatz, CRC Press, 2015.
- "Computational Modeling of Gene Regulatory Networks – A Primer" by Hamid Bolouri, Imperial College Press, 2008.

Lecture slides, discussion forums, announcements and specifications for assessed material will be made available on the course Moodle website (accessible through myUNSW).

Course Evaluation and Development

This course will be evaluated through the online MyExperience process at the end of session. Individual lecturers may also distribute surveys on their own teaching. Feedback from these surveys is taken seriously and you are encouraged to respond.

Staff Details

Position	Name	Email	Location	Phone	Availability	Equitable Learning Services Contact	Primary Contact
Convenor	Sara Ballouz					No	Yes
Lecturer	Raymond Louie					No	No
	Bruno Gaeta					No	No

Other Useful Information

Academic Information

I. Special consideration and supplementary assessment

If you have experienced an illness or misadventure beyond your control that will interfere with your assessment performance, you are eligible to apply for Special Consideration prior to, or within 3 working days of, submitting an assessment or sitting an exam.

Please note that UNSW has a Fit to Sit rule, which means that if you sit an exam, you are declaring yourself fit enough to do so and cannot later apply for Special Consideration.

For details of applying for Special Consideration and conditions for the award of supplementary assessment, please see the information on UNSW's Special Consideration page.

II. Administrative matters and links

All students are expected to read and be familiar with UNSW guidelines and polices. In particular, students should be familiar with the following:

- Attendance
- UNSW Email Address
- Special Consideration
- Exams
- Approved Calculators

- Academic Honesty and Plagiarism
- Equitable Learning Services

III. Equity and diversity

Those students who have a disability that requires some adjustment in their teaching or learning environment are encouraged to discuss their study needs with the course convener prior to, or at the commencement of, their course, or with the Equity Officer (Disability) in the Equitable Learning Services. Issues to be discussed may include access to materials, signers or note-takers, the provision of services and additional exam and assessment arrangements. Early notification is essential to enable any necessary adjustments to be made.

IV. Professional Outcomes and Program Design

Students are able to review the relevant professional outcomes and program designs for their streams by going to the following link: https://www.unsw.edu.au/engineering/student-life/student-resources/program-design.

Note: This course outline sets out the description of classes at the date the Course Outline is published. The nature of classes may change during the Term after the Course Outline is published. Moodle or your primary learning management system (LMS) should be consulted for the up-to-date class descriptions. If there is any inconsistency in the description of activities between the University timetable and the Course Outline/Moodle/LMS, the description in the Course Outline/Moodle/LMS applies.

Academic Honesty and Plagarism

UNSW has an ongoing commitment to fostering a culture of learning informed by academic integrity. All UNSW students have a responsibility to adhere to this principle of academic integrity. Plagiarism undermines academic integrity and is not tolerated at UNSW. *Plagiarism at UNSW is defined as using the words or ideas of others and passing them off as your own.*

Plagiarism is a type of intellectual theft. It can take many forms, from deliberate cheating to accidentally copying from a source without acknowledgement. UNSW has produced a website with a wealth of resources to support students to understand and avoid plagiarism, visit: student.unsw.edu.au/plagiarism. The Learning Centre assists students with understanding academic integrity and how not to plagiarise. They also hold workshops and can help students one-on-one.

You are also reminded that careful time management is an important part of study and one of the identified causes of plagiarism is poor time management. Students should allow sufficient time for research, drafting and the proper referencing of sources in preparing all assessment tasks.

Repeated plagiarism (even in first year), plagiarism after first year, or serious instances, may also be investigated under the Student Misconduct Procedures. The penalties under the procedures can include a reduction in marks, failing a course or for the most serious matters (like plagiarism in an honours thesis or contract cheating) even suspension from the university. The Student Misconduct Procedures are available here:

www.gs.unsw.edu.au/policy/documents/studentmisconductprocedures.pdf

Submission of Assessment Tasks

Work submitted late without an approved extension by the course coordinator or delegated authority is subject to a late penalty of five percent (5%) of the maximum mark possible for that assessment item, per calendar day.

The late penalty is applied per calendar day (including weekends and public holidays) that the assessment is overdue. There is no pro-rata of the late penalty for submissions made part way through a day. This is for all assessments where a penalty applies.

Work submitted after five days (120 hours) will not be accepted and a mark of zero will be awarded for that assessment item.

For some assessment items, a late penalty may not be appropriate. These will be clearly indicated in the course outline, and such assessments will receive a mark of zero if not completed by the specified date. Examples include:

- Weekly online tests or laboratory work worth a small proportion of the subject mark;
- Exams, peer feedback and team evaluation surveys;
- Online quizzes where answers are released to students on completion;
- Professional assessment tasks, where the intention is to create an authentic assessment that has an absolute submission date; and,
- Pass/Fail assessment tasks.

Faculty-specific Information

Engineering Student Support Services – The Nucleus - enrolment, progression checks, clash

requests, course issues or program-related queries

Engineering Industrial Training – Industrial training questions

UNSW Study Abroad – study abroad student enquiries (for inbound students)

UNSW Exchange – student exchange enquiries (for inbound students)

UNSW Future Students – potential student enquiries e.g. admissions, fees, programs, credit transfer

Phone

(+61 2) 9385 8500 – Nucleus Student Hub

(+61 2) 9385 7661 – Engineering Industrial Training

(+61 2) 9385 3179 – UNSW Study Abroad and UNSW Exchange (for inbound students)

School Contact Information

CSE Help! - on the Ground Floor of K17

- For assistance with coursework assessments.

The Nucleus Student Hub - https://nucleus.unsw.edu.au/en/contact-us

- Course enrolment queries.

Grievance Officer - grievance-officer@cse.unsw.edu.au

- If the course convenor gives an inadequate response to a query or when the courses convenor does not respond to a query about assessment.

Student Reps - stureps@cse.unsw.edu.au

- If some aspect of a course needs urgent improvement. (e.g. Nobody responding to forum queries, cannot understand the lecturer)

You should **never** contact any of the following people directly:

- Vice Chancellor

- Pro-vice Chancellor Education (PVCE)

- Head of School

- CSE administrative staff

- CSE teaching support staff

They will simply bounce the email to one of the above, thereby creating an unnecessary level of indirection and a delay in the response.