UNSW Course Outline

BINF2010 Introduction to Bioinformatics - 2024

Published on the 25 Aug 2024

General Course Information

Course Code : BINF2010
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 : Multimodal
Delivery Format : Standard
Delivery Location : Kensington
Campus : Sydney
Study Level : Undergraduate
Units of Credit : 6

Useful Links

Handbook Class Timetable

Course Details & Outcomes

Course Description

The course surveys the major areas of bioinformatics, exploring the history of bioinformatics in relation to advances in computing hardware and software; the biological problems currently being addressed using bioinformatics; and future applications of bioinformatics. Major topics

include genomics; genome sequencing projects; proteomics; structural genomics; systems biology; phylogeny; medical informatics; and commercial applications of bioinformatics. The general nature of the data, computational problems and the approaches employed will be discussed in each case. Bioinformatics will be discussed both as a scientific discipline and as an engineering discipline. The course will also explore the role of bioinformatics in the biotechnology and pharmaceutical industries and ethical issues associated with biological data. Lectures are supplemented by practical exposure to public and commercial bioinformatics web sites and to commonly used bioinformatics software.

Course Aims

Bioinformatics is a relatively new discipline joining two very different fields, and it is not always clear what it entails. This course aims to provide clear foundational knowledge and context for students in a bioinformatics-related degree as well as students who aim to apply bioinformatics as part of a biology degree, through an exploration of the field and its impact on modern biology, together with its industrial and technical context.

BINF2010 is the first bioinformatics course in the bioinformatics programs and is also available as an elective or in some cases a general education course for students taking a biology-related major.

After taking the course, students should have a clear idea of how biology, computing and engineering courses tie into bioinformatics and provide context and tools for its development and application.

Course Learning Outcomes

Course Learning Outcomes
CLO1 : Define bioinformatics and provide examples of common uses in analysing genome, protein and expression data, and in modelling biological systems
CLO2 : Describe some common application areas of bioinformatics and the techniques used therein (genome annotation, rational drug design, medical genomics etc) and describe and explain some common bioinformatics algorithms and common data types used in these applications
CLO3 : Apply common bioinformatics procedures using commonly available software and websites, including: retrieving relevant sequences and structures from databases; identifying ORFs in a DNA sequence; identifying the function of an unknown sequence by similarity searching; identifying the function of an unknown sequence by pattern searches or based on physicochemical properties; creating a multiple sequence alignment and a phylogenetic tree;
CLO4 : Construct simple shell scripts in UNIX to perform file management tasks and launch programs.
CLO5 : Explain the impact of bioinformatics on modern biology and relevant ethical and social issues
CLO6 : Discuss some common challenges in the engineering of computer systems for bioinformatics
CLO7 : Use R to visualise complex data and analyse transcriptomics data

Course Learning Outcomes	Assessment Item
CLO1 : Define bioinformatics and provide examples of common uses in analysing genome, protein and expression data, and in modelling biological systems	• Midterm exam • Final Exam
CLO2 : Describe some common application areas of bioinformatics and the techniques used therein (genome annotation, rational drug design, medical genomics etc) and describe and explain some common bioinformatics algorithms and common data types used in these applications	• Midterm exam • Final Exam
CLO3 : Apply common bioinformatics procedures using commonly available software and websites, including: retrieving relevant sequences and structures from databases; identifying ORFs in a DNA sequence; identifying the function of an unknown sequence by similarity searching; identifying the function of an unknown sequence by pattern searches or based on physicochemical properties; creating a multiple sequence alignment and a phylogenetic tree;	• Laboratory and tutorial work (Quizzes) • Scripting assignment
CLO4 : Construct simple shell scripts in UNIX to perform file management tasks and launch programs.	• Laboratory and tutorial work (Quizzes) • Scripting assignment
CLO5 : Explain the impact of bioinformatics on modern biology and relevant ethical and social issues	• Midterm exam
CLO6 : Discuss some common challenges in the engineering of computer systems for bioinformatics	• Final Exam
CLO7 : Use R to visualise complex data and analyse transcriptomics data	• Laboratory and tutorial work (Quizzes)

Learning and Teaching Technologies

Moodle - Learning Management System | Microsoft Teams

Assessments

Assessment Structure

Assessment Item	Weight	Relevant Dates
Midterm exam Assessment Format: Individual	20%	Due Date: Week 5: 07 October - 13 October
Laboratory and tutorial work (Quizzes) Assessment Format: Individual	25%	Start Date: Not Applicable Due Date: Week 3: 23 September - 29 September, Week 4: 30 September - 06 October, Week 5: 07 October - 13 October, Week 9: 04 November - 10 November, Week 10: 11 November - 17 November, Week 11: 18 November - 24 November
Final Exam Assessment Format: Individual	40%	
Scripting assignment Assessment Format: Individual	15%	Due Date: Week 9: 04 November - 10 November

Assessment Details

Midterm exam

Assessment Overview

An exam in the middle of the term, covering the work done so far. The exam is delivered online on the Learning Management System. Questions include short essays as well as problems that require you to apply the algorithms demonstrated in the course. Essay questions are marked by the lecturers that set the questions. Problems are marked automatically by the quiz engine of the LMS. Marks and feedback are provided to the students through the LMS.

Course Learning Outcomes

- CLO1 : Define bioinformatics and provide examples of common uses in analysing genome, protein and expression data, and in modelling biological systems
- CLO2 : Describe some common application areas of bioinformatics and the techniques used therein (genome annotation, rational drug design, medical genomics etc) and describe and explain some common bioinformatics algorithms and common data types used in these applications
- CLO5 : Explain the impact of bioinformatics on modern biology and relevant ethical and social issues

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.

Laboratory and tutorial work (Quizzes)

Assessment Overview

6 quizzes assessing the work done in the labs and in the tutorial. Each quiz is due the week after the relevant lab or tutorial. Quizzes are available on the Learning Management System and are marked automatically. After the quiz has closed, marks and feedback are made available to the students through the LMS quiz engine.

Course Learning Outcomes

- CLO3 : Apply common bioinformatics procedures using commonly available software and websites, including: retrieving relevant sequences and structures from databases; identifying ORFs in a DNA sequence; identifying the function of an unknown sequence by similarity searching; identifying the function of an unknown sequence by pattern searches or based on physicochemical properties; creating a multiple sequence alignment and a phylogenetic tree;
- CLO4 : Construct simple shell scripts in UNIX to perform file management tasks and launch programs.
- CLO7 : Use R to visualise complex data and analyse transcriptomics data

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

An approx 2-hour exam in the final exam period, delivered online (Learning Management System). The exam consists of MCQs covering the entire course, plus some essay questions covering the major course modules that were not covered in the midterm exam. MCQs are marked automatically by the quiz engine, and essay questions are marked manually by the

lecturers who taught the corresponding content, based on relevance, accuracy, completeness and clarity of the answers.

Course Learning Outcomes

- CLO1 : Define bioinformatics and provide examples of common uses in analysing genome, protein and expression data, and in modelling biological systems
- CLO2 : Describe some common application areas of bioinformatics and the techniques used therein (genome annotation, rational drug design, medical genomics etc) and describe and explain some common bioinformatics algorithms and common data types used in these applications
- CLO6 : Discuss some common challenges in the engineering of computer systems for bioinformatics

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.

Scripting assignment

Assessment Overview

An assignment requiring students to write a script automating a "pipeline" of bioinformatics tasks. The assignment may involve shell, Python, Perl or R scripting. Students submissions are evaluated using a number of test cases as well as manual inspection of the submitted code, with the results provided to the students.

Course Learning Outcomes

- CLO3 : Apply common bioinformatics procedures using commonly available software and websites, including: retrieving relevant sequences and structures from databases; identifying ORFs in a DNA sequence; identifying the function of an unknown sequence by similarity searching; identifying the function of an unknown sequence by pattern searches or based on physicochemical properties; creating a multiple sequence alignment and a phylogenetic tree;
- CLO4 : Construct simple shell scripts in UNIX to perform file management tasks and launch programs.

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.

General Assessment Information

Lab quizzes, exams and assignments are all submitted on Moodle

Grading Basis

Standard

Course Schedule

Teaching Week/Module	Activity Type	Content
Week 1 : 9 September - 15 September	Lecture	Introduction to bioinformatics. Sequence data, evolution and alignment
	Lecture	Sequence databases, RESTful APIs
	Tutorial	Setting up your environment(s)
	Laboratory	Introduction to UNIX command line
Week 2 : 16 September - 22 September	Lecture	Sequence alignment by dynamic programming
	Lecture	Database searching and multiple sequence alignment
	Tutorial	Introduction to Python
	Laboratory	Basic sequence analysis using EMBOSS
	Assessment	Online quiz
Week 3 : 23 September - 29 September	Lecture	Genome informatics: applications and assembly algorithms, annotation and genome databases
	Lecture	Phylogeny inference and UPGMA algorithm
	Tutorial	Data viz with Python
	Laboratory	Phylogeny and UPGMA
	Assessment	Online quiz
Week 4 : 30 September - 6 October	Lecture	Structural bioinformatics
	Lecture	Case study: characterising a new protein
	Tutorial	Functions and data wrangling
	Laboratory	Structural bioinformatics
	Assessment	Online quiz
Week 5 : 7 October - 13 October	Lecture	Public holiday
	Lecture	Distributed and application-specific hardware in bioinformatics
	Tutorial	Introduction to R
	Laboratory	TBD
	Assessment	Midterm exam
Week 6 : 14 October - 20 October	Other	Flexibility week
Week 7 : 21 October - 27 October	Lecture	Transcriptomics
	Lecture	Transcriptomics
	Tutorial	Data viz with R
	Laboratory	Advanced UNIX and shell scripting
Week 8 : 28 October - 3 November	Lecture	Experimental design
	Lecture	TBD
	Tutorial	Functions and data wrangling (R)
	Laboratory	Transcriptome analysis using R
	Assessment	Online quiz
Week 9 : 4 November - 10 November	Lecture	Proteomics
	Lecture	Proteomics
	Tutorial	Shiny!
	Laboratory	Proteomics
	Assessment	Scripting assignment
	Assessment	Online quiz
Week 10 : 11 November - 17 November	Lecture	TBD Case Study: Genomics
	Lecture	Case study: medical application of human genomics and transcriptomics
	Tutorial	Recap
	Laboratory	Workflow management using Galaxy
	Assessment	Online quiz

Attendance Requirements

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

Course Resources

Recommended Resources

There is no required textbook for this course.

However, Bruno recommends “Digital Code of Life – How Bioinformatics is Revolutionizing Science, Medicine and Business” by Glyn Moody. "This is quite an entertaining book that discusses the history and context of bioinformatics. This book does not cover the same materials as the course, but provides background information and context on most sections of the course, in a very readable format."

Some resources and links:

http://www.bioinformatics.org/ and their wiki

Australian Bioinformatics and Computational Biology Society (ABACBS)

Australian Computational Biology and Bioinformatics Student Society (COMBINE)

Australian BioCommons

Bioplatforms Australia

Melbourne Bioinformatics

Statistical Bioinformatics Seminars (USyd)

International Society for Computational Biology (ISCB)

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	Bruno Gaeta					No	No
	Peter Humburg					No	No
	Hasindu Gamaarachchi					No	No
	Fabio Luciani					No	No
	Eleni Giannoulatou					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.