



Statistical Methods in Bioinformatics

The University of Toledo
Bioinformatics, Proteomics and Genomics, COMLS
BIPG 520/720 (45720/45059)

Instructor:	Sadik A. Khuder. Ph.D.	Term:	Fall 2020
Email:	sadik.khuder@utoledo.edu	Course Website:	Blackboard Learn
Office Hours:	By arrangement	Class Location:	Online
Office Location:	Room 12 RHC	Class Day/Time:	Mondays / 2 – 5 pm
Office Phone:	(419) 383-4089	Lab Location:	127 HEB
Instructor Phone:	(419) 897-9258	Lab Day/Time:	10/26, 11/02, 11/09/ 1-3 pm
		Credit Hours:	3

SPECIAL COURSE EXPECTATIONS DURING COVID-19 ATTENDANCE

The University of Toledo has a missed class policy. It is important that students and instructors discuss attendance requirements for the course. Students must perform a daily health assessment, based on based on [CDC guidelines](#), before coming to campus each day, which included taking their temperature. Students who are symptomatic/sick should not come to class and should contact the Main Campus Health Center at 419-530-3451. *Absences due to COVID-19 quarantine or isolation requirements are considered excused absences.* Students should notify their instructors and these absences may not require written notice.

FACE COVERINGS

All students must wear face coverings while on campus, except while eating, alone in an enclosed space, or outdoors practicing social distancing. NO students will be permitted in class without a face covering. If you have a medical reason that prevents you from wearing a face covering due to a health condition deemed high-risk for COVID-19 by the Centers for Disease Control and Prevention (CDC), you should submit a request for an accommodation through the Student Disability Services Office (SDS) by completing the [online application](#). Students will need to provide documentation that verifies their health condition or disability and supports the need for accommodations. If a student is already affiliated with SDS and would like to request additional accommodations due to the impact of COVID-19, should contact their accessibility specialist to discuss their specific needs.

SOCIAL DISTANCING

Students should practice social distancing inside and outside the classroom please follow signage and pay attention to the seating arrangements. Do not remove stickers or tape from seats and/or tables, this is there to provide guidance on the appropriate classroom capacity based on the recommended 6 feet of social distancing between individuals. Please be conscious of your personal space and respectful of others. Also be cognizant of how you enter and exit the room; always try to maintain at least 6 feet of distance between yourself and others.

DESKS AND WORK SPACES

Students will need to sanitize their desks and/or work space before class with the University provided sanitizing spray and paper towels their desks.

SPECIAL NOTES

It's important to note that based on the unpredictability of the COVID-19 virus things can change at any time so please be patience and understanding as we move through the semester. I also ask that you keep me informed of concerns you may have about class, completing course work/assignments timely and/or health concerns related to COVID.



CATALOG/COURSE DESCRIPTION*

This course introduces students to statistical methods commonly used in bioinformatics. Students will learn to use statistical programs and related bioinformatics resources locally and on the Internet. Lectures and lab discussion will emphasize on the statistical models and methods underlying the computational tools. The course will focus on the application of the newer statistical methods and the reasoning behind these applications. More emphasis will be placed on the analysis of genomic and proteomic experiments and students will learn statistical techniques to handle RNA-seq, microRNA, microarray, methylation, and proteomic data.

COURSE OVERVIEW/ TEACHING METHODOLOGY

A variety of teaching methods will be used, including in class lectures, exercises, quizzes, project, and online instructional lectures and videos.

STUDENT LEARNING OUTCOMES*

After completion of the course, students should be able to:

1. State the fundamental concepts of statistics in bioinformatics.
2. Describe the formulation of stochastic models for genomic and proteomic data.
3. Apply statistical techniques to analyze RNA-seq & single cell RNA-seq data.
4. Apply statistical techniques to analyze microarray, microRNA, and methylated data
5. Interpret the results of the analysis of RNA-seq, microarray, microRNA, and methylated data.
6. Download and analyze data from GEO and ArrayExpress data repositories.
7. Use statistical tests commonly employed in bioinformatics.
8. Recognize modern statistical methods and software to solve complex problems in bioinformatics.
9. Interpret the statistical results as reported in the bioinformatics literature.

PREREQUISITES AND COREQUISITES*

None

TEXTS AND ANCILLARY MATERIALS*

There is no required text for this course. All the required materials will be available on the course web site. Readings will consist of original literature, review articles, and R based books.

TECHNOLOGY REQUIREMENTS

None

ACADEMIC POLICIES*

[Graduate Policies: http://www.utoledo.edu/policies/academic/graduate/](http://www.utoledo.edu/policies/academic/graduate/)

COURSE EXPECTATIONS

Attendance for the lectures on Blackboard Collaborate is mandatory. Instructions for assignments, exams, quizzes, lecture and other course announcements will be posted on the Course News. A copy of each announcement will be sent to your email. All email communications to me should be sent to my utoledo email address, sadik.khuder@utoledo.edu. Assignments, quizzes and exams must be the original work of each student. Neither plagiarism nor cheating will be tolerated and will be addressed in accordance with the applicable policy of the University of Toledo.



OVERVIEW OF COURSE GRADE ASSIGNMENT*

Assignments 30%

There are 12 assignments in this course. These assignments are intended to improve skills in bioinformatics and statistical analysis of genomic and proteomic data. The lowest assignment score will be dropped from the calculation of your final grade.

Quizzes 15%

There are 7 quizzes in this course. Each quiz will cover the material presented in that particular week. Few questions from previous quizzes may also be included. Further instructions will be supplied for each particular quiz.

Project 20%

A project will explore a topic of the course in greater depth. A written report, along with a power point presentation, is required. A project that addresses a question through data analysis with a written report and summary of conclusions would be sufficient. More detailed descriptions of the project along with suggested topics will be posted to the course webpage at the appropriate times during the term.

Final examination 35%

The final exam will be a take-home exam. The exam will cover all the materials presented in the course.

Midterm Grading*

Midterm grades will be calculated based on assignments 1-7 and quizzes 1-3.

Final Grading*

≥ 90 A, 87-89 A-, 84-86 B+, 80-83 B, 77-79 B-, 74-76 C+, 70-73 C, < 70 D, < 50 F.

UNIVERSITY POLICIES

Policy Statement on Non-Discrimination on the Basis of Disability (ADA)* The University is an equal opportunity educational institution. Please read [The University's Policy Statement on Nondiscrimination on the Basis of Disability Americans with Disability Act Compliance](#). Students can find this policy along with other university policies listed by audience on the [University Policy webpage](#) (<http://www.utoledo.edu/policies/audience.html/#students>).

Academic Accommodations

The University of Toledo embraces the inclusion of students with disabilities. We are committed to ensuring equal opportunity and seamless access for full participation in all courses. For students who have an accommodations memo from Student Disability Services, I invite you to correspond with me as soon as possible so that we can communicate confidentially about implementing accommodations in this course. For students who have not established affiliation with Student Disability Services and are experiencing disability access barriers or are interested in a referral to healthcare resources for a potential disability or would like information regarding eligibility for academic accommodations, please contact the [Student Disability Services Office](#) (<http://www.utoledo.edu/offices/student-disability-services/>) by phone: 419.530.4981 or email at StudentDisability@utoledo.edu.



ACADEMIC AND SUPPORT SERVICES*

Please follow this link to view a comprehensive list of [Student Academic and Support Services](http://www.utoledo.edu/studentaffairs/departments.html) (http://www.utoledo.edu/studentaffairs/departments.html) available to you as a student

SAFETY AND HEALTH SERVICES FOR UT STUDENTS*

Please use the following link to view a comprehensive list [Campus Health and Safety Services](#) available to you as a student.

INCLUSIVE CLASSROOM STATEMENT

In this class, we will work together to develop a learning community that is inclusive and respectful. Our diversity may be reflected by differences in race, culture, age, religion, sexual orientation, gender identity/expression, socioeconomic background, and a myriad of other social identities and life experiences. We will encourage and appreciate expressions of different ideas, opinions, and beliefs so that conversations and interactions that could potentially be divisive turn, instead, into opportunities for intellectual and personal development.

COURSE SCHEDULE

Week*	Date	Topic	Assignments***	Quiz**
1	8/17	Review of probability, and statistics Genome sequence of SARS-Coronavirus	1	1
2	8/24	Modeling DNA	2	2
3	8/31	Genomic data bases	3	
4	9/08	Differential gene expression	4	
5	9/14	Analyzing RNA-seq data I	5	
6	9/21	Analyzing RNA-seq data II	6	3
7	9/28	Analyzing Single-Cell RNA data	7	
8	10/05	Analyzing MicroRNA data	8	
9	10/12	Analyzing Methylation data	9	4
10	10/19	Evolutionary models Phylogenetic tree estimation	10	
11	10/26	Protein structure analysis Analyzing proteomic data	11	5
12	11/02	Markov Chains Hidden Markov model	12	
13	11/09	Cluster Analysis		6
14	11/16	Discriminant analysis ANOVA Models	Project submission	7
15	11/23	Applied Bioinformatics	Final Exam	

*Modules are posted online on the indicated dates

**Quizzes are posted on Mondays at 6 am and closed on Sundays at 11:30 pm.

***Assignments are posted on Mondays at 6 am and due on Sundays at 11:59 pm.