San José State University
Computer Science Department
CS 223 Bioinformatics, Sec 01, Fall 2018

Course Information

Instructor: Leonard P. Wesley
Department: Computer Science
College of Science, San Jose State University.
Fall Semester, 2018

Course and Contact Information

Instructor:	Leonard Wesley
Office Location:	MH 212
Telephone:	408.924.5287
Email:	Leonard.Wesley@sjsu.edu
Office Hours:	Thursdays 2:00PM – 4:00PM On 9/20, 10/25, and 11/29 Office hours will be from 2:30PM – 4:30PM
Class Days/Time:	Tuesdays and Thursdays 6:00PM – 7:15PM
Classroom:	DH 450
Prerequisites:	Prerequisite: CS 123A or CS 155.

Course Description

The course investigates the main algorithms for solving computational problems in bioinformatics. Methods will include Hidden Markov Models for gene prediction and protein profiling, and Genetic Algorithms for biological sequence analysis and structure prediction. Students will be given programming projects.

Learning Outcomes

Upon successful completion of this course, students will:

1. SLO-1 Have a basic understanding of molecular biology and the biology central dogma, the transcription and translation process, regulation of transcription and translation, genetic crossover and translocation, genes and alleles, genotype and phenotype, basics of epigenetics,
2. SLO-2 Know to build and implement Genetic algorithms, HMMs, hierarchical, k-means, and Knn clustering solutions to biological problems using Python.
3. SLO-3 Know how the CRISPR-Cas system works, find CRISPR arrays, and find genomic repeats.
4. SLO-4 Have overview knowledge of Next Generation Sequencing (NGS) and genome assembly.
5. SLO-6 How the microarray technology works and how to carry out microarray analysis.

Required Texts/Readings

All required text, publications, reference material, and so forth will be provided to the class.

Other Readings

Developing Bioinformatics Computer Skills, Cynthia Gibas and Per Jambeck, O'Reilly & associates. (A good book for beginners)

Introduction to Computational Biology: Maps, Sequences and Genomes, Michael S. Waterman, CRC Press. (A statistical oriented view of bioinformatics)

Bioinfromatics: A Practical Guide to the Analysis of Genes and Proteins, Andreas D. Baxecvanis and B.F. Francis Ouellette, John Wiley & Sons 2nd Ed. (Includes contributions from several authors providing a wide perspective)

Computational Resources

Students are required to make sure that they have access to sufficient UNIX, Windows, or Mac based computational resources (e.g., computers and software) to carryout assignments in the course. An attempt to offer the course in a classroom with sufficient computation resources will be made by the department to support classroom instruction and demonstrations. However, students should be prepared to bring their portable laptops to class.

Course Requirements and Assignments

Students will be assigned a video and or related multi-media or electronic copies of Python programming, programming in general, or literature related to developing computational solutions on a weekly basis. Students will be expected and required to read the assigned material and complete all homework or programming tasks prior to the indicated next class meeting.

In class instruction will consist of a short quiz at the start of elected classes to test comprehension of assigned material. Then the class will be divided into groups of 2-3 students to work on in-class programming exercises during indicated classes.

Course Logistics

Students should expect to spend approximately nine (9) hours per week (on average) completing the assigned course work. This includes viewing videos, homework, in-class lecture and in-class exercise time. The amount of time that a student actually spends depends on their individual skills and the time allocated to the course. The nine (9) hours per week estimate is based on the previous experiences of the instructor and students. So please plan and schedule accordingly.

Previously, students have asked for special exception to policies and procedures for this course. An example includes asking the instructor for extra assignments or work to help improve a student's grade. Even if such a request is reasonable in the view of the instructor, no exception will be given to a student unless it can be made available to the entire class, AND does not constitute significant extra work on the part of students, instructors, graders and so forth. Students should have no concern that other students will receive special exceptions that will not be available to them to pursue.

NOTE: University policy F69-24 at http://www.sjsu.edu/senate/docs/F69-24.pdf states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading." However, attendance will be required in order to complete and submit many in-class exercises, quizzes, and exams.

Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the syllabus.

Quizzes and Exams:

There will be three quizzes, one midterm and a final exam all of which will count toward the final grad as specified in the “Grades” section below. During quizzes and exams, communication with other individuals via any means is strictly prohibited without the express permission of the instructor. Violations will be met with the full impact of SJSU’s academic integrity policy and procedures.

Projects:

Several life science related project topics will be described near the start of the course. Projects will involve applying the skills and knowledge learned in the course to the project. Teams of 2-3 students will be formed to work on a selected project topic. Teams will be required to submit a project proposal before starting on a project, and submit a project report along with working code at the end of the course. Individual student scores on a project will be determined by the content and quality of the contribution of each student toward the project. The score on the course project and project presentation will count toward the final grad (percentage wise) as specified in the “Grades” section below.

Reading, Homework, Programming, In-Class Exercises, Participation Assignments

Graded reading, homework, programming, and class participation and brief course feedback assignments will be given almost weekly, and will count toward the final grade. There will be 4 In-class Exercise sessions. These will typically involve forming teams of 2-3 students that work on assigned programming exercises in the classroom. They provide an opportunity to get started on homework programming assignments that are to be submitted on a designated due date. Participation is mandatory, and scores will count toward final grade.

Tentative course calendar of assignment due dates & exam dates:

(Please note that course calendar below, and its content is "subject to change with fair notice")

Week and Class Mtg #	Tue	Thur	Module # & Name	TOPIC	Assignment See Canvas For Module & Weekly Assignment Details and Due Dates
Week 1 Class Mtgs 1 & 2	8/21	8/23	#1 Biology Basics	8/21: Intro To Course: -Topics, learning objectives, course logistics, Instructor background - Greensheet 8/23: - Intro to molecular biology, DNA, RNA, and the central dogma. - DNA Replication	Learning Module #1 Week #1 Project Team Formation By Instructor
Week 2 Class Mtgs 3 & 4	8/28	8/30	#1 Biology Basics	8/28: - Transcription, Translation 8/30: - Regulation of transcription and translation	Learning Module #1 Week #2 August 31st Last Day To Drop Classes Project Teams Considering Possible Projects
Week 3 Class Mtgs 5 & 6	9/4	9/6	#1 Biology Basics	9/4 - Genes, alleles, genotype, phenotype, genetic crossover, translocation, types of mutations 9/6: - **In-Class Exercise 1** Topics Covered 8/30 – 9/4	Learning Module #1 Week 3 Project Proposals Due 9/4

4	9/11	9/13	#1 Biology Basics	9/11: - Proteins and their structure - Function from structure 9/13: - Epigenetics, regulation, diseases	Learning Module #1 Week 4
5	9/18	9/20	#2 Genetic Algorithms	9/18: - **Quiz 1 (~35 mins): Covers Topics Week 1 thru Week 4** - Genetic Algorithm 9/20: - Genetic Algorithm (cont.)	Learning Module #2 Week 5
6	9/25	9/27	#2 Genetic Algorithms & HMMs	9/25: - Application of Genetic Algorithms 9/27: - HMMs	Learning Module #2 Week 6
7	10/2	10/4	#2 HMMs	10/2: - **In-Class Exercise 2** Topics Covered 9/11 – 9/27 10/4: - **Midterm (Full period): Covers Topics Week 1 thru Week 6**	Learning Module #2 Week 7
8	10/9	10/11	#2 HMMs	10/9: - HMMs (cont.) 10/11: - Applications of HMMs to gene finding	Learning Module #2 Week 8

9	10/16	10/18	#2 Hierarchical Clustering ,K-means, KNN	10/16: - **Quiz 2 (~35 mins): Covers Topics Week 7 thru Week 8** - Hierarchical Clustering 10/18: - Hierarchical Clustering (cont.) - K-means Clustering - KNN	Module #2 Week 9
10	10/23	10/25	#3 CRISPR	10/23: - CRISPR-Cas 10/25: - CRISPER-Cas (cont.)	Module #3 Week 10
11	10/30	11/1	#3 CRISPR	10/30: - **In-Class Exercise 3** Topics Covered 10/2 – 10/25 11/1 - CRISPER-Cas (cont.)	Module #3 Week 11
12	11/6*	11/8	#4 NGS	11/6: - NGS *(* NOTE: Instructor will be on travel. This lecture will be a reading assignment. No class meeting)* 11/8: - **Quiz 3 (~35 mins): Covers Topics Week 9 thru Week 11** - NGS (cont)	Module #4 Week 12
13	11/13	11/15	#4 NGS	11/13: - NGS (cont) 11/15: - **In-Class Exercise 4** (Work on Team Projects, Q&A)	Module #4 Week 13

14	11/20	11/22	#5 Microarray Technologies	11/20: - Microarray Technologies 11/22: - **Thanksgiving Holiday**	Module #5 Week 14
15	11/27	11/29	#5 Microarray Technologies	12/6: - Microarray Technologies (cont) 12/8: - Microarray Technologies (cont) - Review For Final Exam	Module #5 Week 15
			Final Project Code and Project Report Due To Canvas December 13, 2018 By 11:59PM **Final Exam** December 13, 2018, DH 450, 5:15PM to 7:30PM		

SCHEDULE FOOTNOTES:
NONE AS OF AUGUST 2018

Grades *

WRITTEN HOMEWORK (10 at 10 points each)	100 pts
QUIZZES (3 at 50pts each)	150 pts
MIDTERM	100 pts
IN-CLASS EXERCISES (4 at 50pts each)	150 pts
WEEKLY COURSE FEEDBACK (14 at 5pts each)	70 pts
PROGRAMMING ASSIGNMENTS (5 @ 40pts each)	200 pts
FINAL EXAM	200 pts
FINAL PROJECT REPORT & CODE	200 pts
Total Course Points	= 1,170 pts Total

* The total points for each category might change depending on the number of project teams and assignments. The instructor reserves the right to adjust, with sufficient advanced notice, the above point distributionby ± 5 pts. Such adjustments might be based on the difficulty or simplicity of assignments or quizzes or exams.

Grading Percentage Breakdown (NOTE: Ranges might change if point totals change)

Percentage of Total Pts	Pts	Letter Grade
96.66% and above	≥ 1,131	A+
93.33% - 96.65%	≥ 1,092- 1,130	A
90% - 93.32%	≥ 1,053 – 1,091	A-
86.66% - 89.99%	≥ 1,013 – 1,052	B+
83.33% - 86.65%	≥975 – 1,012	B
80% - 83.32%	≥ 936 – 974	B-
76.66% - 79.99%	≥ 897 – 935	C+
73.33% - 76.65%	≥ 858 - 896	C
70% - 73.32%	≥ 819 – 857	C-
66.66% - 69.99%	≥ 778– 818	D+
63.33% - 66.65%	≥ 741 – 777	D
60% - 63.32%	≥ 702 - 740	D-
Below 60%	**< 702**	**F**

HOW TO CALCULATE/ESTIMATE YOUR GRADE

If students would like to calculate their numeric grade percentage, the formula is as follows:

Numeric CS 223 Grade Percentage =

$$\frac{Total\ points\ from\ assignments}{Total\ course\ points} x\ 100\%$$

There is no guarantee that grades will be curved. If so, it will be done at the end of the semester. The instructor is already aware that graduate students need to maintain an overall GPA of B or better. Just because a student NEEDS a particular grade doesn't mean that the instructor will automatically GIVE the student that grade. Students must EARN a passing grade based on submitted and evaluated course work.

Extra credit options, if available

There are no extra credit assignments in this course except for completing designated "Advanced" assignments. However, homework assignments and exams might contain extra credit options.

Penalty for late or missed work

Late assignments will receive a 25% deduction for every 24hr period the submission is late. There will be partial credit for assignments.

Receiving An Incomplete (I) Grade

Receiving a grade of Incomplete (I) is not automatic. Students must complete at least 80% of course assignments by the end of the semester to be eligible to receive a grade of incomplete. Students must also provide documentation to support the reason for the request to receive an Incomplete grade. The instructor has the final decision to give an Incomplete grade. If the instructor agrees to give a student an Incomplete grade, the instructor will enter the remaining work to be completed as part of the PeopleSoft grade submission process.

Grade Change Policy

It is a university policy that course grade changes must be made within one semester from the end of the course. Requests for exceptions to this policy must be accompanied with a documented and compelling reason.

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/. Make sure to review these policies and resources