

CIIC 5110 - Bioinformatics Algorithms

1. General Information:

Alpha-numeric codification: CIIC 5110
Course Title: Bioinformatics Algorithms
Number of credits: 3
Contact Period: 3 hours of lecture per week

2. Course Description:

English: Discussion of algorithms for processing genomic and proteomic data. Discussion of heuristic randomized and non-heuristic algorithmic solutions for sequence comparison, gene finding and gene expression state determination problems, among others. Discussion of algorithmic design principles and their impacts on time and space complexity and the quality of results.

Spanish: Discusión de algoritmos para procesar datos genómicos y proteómicos. Discusión de soluciones algorítmicas heurísticas, aleatorizadas y no heurísticas para comparación de secuencias, hallazgos genéticos y problemas de determinación del estado de expresión genética, entre otros. Discusión de principios de diseño algorítmico y sus impactos en la complejidad del tiempo y el espacio y la calidad de los resultados.

3. Pre/Co-requisites and other requirements:

Prerequisites: (ICOM 4038 or CIIC 4025) and ININ 4010 or authorization from department chair.

4. Course Objectives:

With this course, the student will be exposed to fundamental concepts and algorithms, including their design and analysis, in the bioinformatics field. It is aimed to contribute to the development of important skills in this modern field.

5. Instructional Strategies:

conference discussion computation laboratory
seminar with formal presentation seminar without formal presentation workshop
art workshop practice trip thesis special problems tutoring
research other, please specify:

6. Minimum or Required Resources Available:

Students will use the Departmental computer laboratories to complete course projects.

7. Course time frame and thematic outline

Outline	Contact Hours
Introduction: DNA information, transcription and translation, central dogma of molecular biology. Nucleotide and protein databases.	1.5
String matching methods and the sequence alignment problem: Similarity and homology. Knuth-Morris-Pratt and Suffix-trees exact matching algorithms. Formulation of the gapped alignment problem.	5
Dynamic Programming Sequence Alignment: Scoring schemes, Substitution matrices and gap penalty mappings. Algorithms of Needleman-Wunsh and Smith-Waterman. Boltzmann distributions and the Probalign algorithm.	1.5
Heuristic Sequence Alignment: BLAST/FASTA and the statistics of local alignment. BLAST's E-value and P-value.	3
Statistical Methods: Hidden Markov models methods: Forward, Backward, Viterbi and Baum-Welch algorithms. Applications to CpG islands identification and gene finding.	3
Multiple Sequence Alignment: Problem definition, Levenshtein distance and scoring schemes, Aligners for proteins, Spliced alignment methods.	4.5
Motif Finding: Profiling, Exhaustive search, Greedy approaches, Gibbs sampling and random projections.	4.5
Genome Assembly: Next generation sequencing data. Quality, coverage and depth of a set of reads. De Bruijn graphs and <i>de novo</i> assembly. Mapping assembly, Bowtie2, BLAT.	7
Phylogenetic Analysis: Clustering, neighbor joining, parsimony and maximum likelihood methods.	3
Comparative Genomics: Genome rearrangements.	3
Gene Expression: Binary quantization. Shou-Wang-Daugherty mixture model method. Shmulevich algorithm, Multiscale Binarization , BASC.	6
Exams	3
Total hours: (equivalent to contact period)	45

8. Grading System

Quantifiable (letters) Not Quantifiable

9. Evaluation Strategies

	Quantity	Percent

<input checked="" type="checkbox"/> Exams	3	60%
<input checked="" type="checkbox"/> Final Exam	1	20%
<input type="checkbox"/> Short Quizzes		
<input type="checkbox"/> Oral Reports		
<input type="checkbox"/> Monographies		
<input type="checkbox"/> Portfolio		
<input checked="" type="checkbox"/> Projects	variable	20%
<input type="checkbox"/> Journals		
<input type="checkbox"/> Other, specify:		
TOTAL:		100%

10. Bibliography:

Textbook:

- Neil C. Jones, Pavel Pevsner, *An Introduction to Bioinformatics Algorithms*, 1st Ed., MIT Press, 2004.
- Phillip Compeau, and Pavel Pevzner, *Bioinformatics Algorithms: An Active Learning Approach*, 2nd Ed. Vol. 1, Active Learning Publishers, 2015

References:

- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, 3rd Ed., MIT Press, 2009.
- Bernhard Scholkopf, Koji Tsuda, Jean-Philippe Vert, *Kernel Methods in Computational Biology*, 1st Ed. MIT Press, 2004.
- <http://nihlibrary.nih.org> *Bioinformatics Support Program*, National Institutes of Health.

11. Course Outcomes

After successfully completing this course, the student will be able to:	Program Student Outcomes Impacted
1. Interpret/Explain the theoretical principles behind most common Bioinformatics data analysis methods.	2
2. Explain specific methodologies and approaches for the computational solution of problems that do not have a deterministic mathematical description	1
3. Analyse, adapt or design new methods for solving Bioinformatics problems.	1
4. Judge the influence of choices of parameters or algorithmic paradigms in the results.	7

12. According to Law 51

Students will identify themselves with the Institution and the instructor of the course for purposes of assessment (exams) accommodations. For more information please call the Student with Disabilities Office which is part of the Dean of Students office (Office #4) at (787)265-3862 or (787)832-4040 extensions 3250 or 3258.

13. Academic Integrity

-The University of Puerto Rico promotes the highest standards of academic and scientific integrity. Article 6.2 of the UPR Students General Bylaws (Board of Trustees Certification 13, 2009-2010) states that academic dishonesty includes, but is not limited to: fraudulent actions; obtaining grades or academic degrees by false or fraudulent simulations; copying the whole or part of the academic work of another person; plagiarizing totally or partially the work of another person; copying all or part of another person answers to the questions of an oral or written exam by taking or getting someone else to take the exam on his/her behalf; as well as enabling and facilitating another person to perform the aforementioned behavior. Any of these behaviors will be subject to disciplinary action in accordance with the disciplinary procedure laid down in the UPR Students General Bylaws.—