UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Matematica e Informatica
ACADEMIC YEAR	2023/2024
MASTER'S DEGREE (MSC)	DATA, ALGORITHMS AND MACHINE INTELLIGENCE
SUBJECT	PATTERN DISCOVERY FOR LIFE SCIENCES
TYPE OF EDUCATIONAL ACTIVITY	B
AMBIT	50341-Discipline Informatiche
CODE	22454
SCIENTIFIC SECTOR(S)	INF/01
HEAD PROFESSOR(S)	EPIFANIO CHIARA Ricercatore Univ. di PALERMO
OTHER PROFESSOR(S)	
CREDITS	6
INDIVIDUAL STUDY (Hrs)	102
COURSE ACTIVITY (Hrs)	48
PROPAEDEUTICAL SUBJECTS	
MUTUALIZATION	
YEAR	1
TERM (SEMESTER)	1° semester
ATTENDANCE	Not mandatory
EVALUATION	Out of 30
TEACHER OFFICE HOURS	**EPIFANIO CHIARA** Tuesday 14:30 17:00 Dipartimento di Matematica ed Informatica, via Archirafi 34, Room 104, primo piano/ first floor

DOCENTE: Prof.ssa CHIARA EPIFANIO

PREREQUISITES	Theory of algorithms: algorithmic techniques and complexity of an algorithm.
LEARNING OUTCOMES	Knowledge and Understanding After a brief biological digression, students will be introduced to some fundamental problems in the field of Bioinformatics. They will acquire advanced tools for analyzing bioinformatics problems and will learn to use Bioinformatics specific language. Applying knowledge and understanding Students will acquire the ability to apply the presented algorithms and to evaluate their importance in applications in the field of biology. Making judgements Students will be able to evaluate the implications of the studied results and to address in a critical way the proposed subjects, also with the help of scientific publications. Communication skills Students will be able to present the results of algorithmic studies, even to a non-expert public. Moreover, they will be able to support the importance and highlight the technological implications of the studied theories. Learning skills Students will be able to read scientific publications in the field of bioinformatics. They will also be able to attend, using the knowledge acquired in the course, both second-level masters, and advanced courses, and specialized seminars in bioinformatics.
ASSESSMENT METHODS	The final assessment is built on a written test through which the student can show the results obtained thanks to the course, in terms of knowledge and comprehension, ability in applying knowledge and understanding, autonomy in making judgments, communications skills, learning ability. The exam consists of some closed questions (1 point each), some open questions (max 2.5 points each) and the application of an algorithm on a given input (max 10 points). The assessment is expressed in thirties. In details, the outcome of the assessment can be described as follows. 30 cum laude: excellent knowledge of the topics, excellent properties of language, excellent analytical skills, excellent ability to independently apply the acquired knowledge 30: excellent knowledge of the topics, excellent properties of language, good analytical ability, good ability to independently apply the acquired knowledge 26-29: Good mastery of the subjects, good properties of language, reasonable ability to independently apply the acquired knowledge 24-25: Good knowledge of the main topics, reasonable property of language, with limited ability to independently apply the acquired knowledge 21-23: student knows the main topics of teaching but has not fully mastered; limited property of language, poor ability to independently apply the acquired knowledge 18-20: minimal knowledge of the main topics of teaching and of the technical language, very little ability to independently apply the acquired knowledge Insufficient: student does not have an acceptable knowledge of the contents of the topics covered in the teaching and has no ability to independently apply the acquired knowledge.
EDUCATIONAL OBJECTIVES	The aim of the course is to provide both theoretical and practical tools to solve problems in the field of bioinformatics. Students will also acquire the ability to deepen the acquired knowledge by consulting scientific specialized publications.
TEACHING METHODS	Lectures
SUGGESTED BIBLIOGRAPHY	Dispense del Corso/Lecture notes Libri di Consultazione/Reference Books • Dan Gusfield - Algorithms on strings, trees and sequence, Cambridge University Press, 2011; ISBN 0-521-58519-8. • Neil C. Jones and Pavel A. Pevzner - An Introduction to Bioinformatics Algorithms. The MIT Press, 2004; ISBN 0-262-10106-8

SYLLABUS

Hrs	Frontal teaching
4	Introduction to problems in bioinformatics
5	Similarity between strings. Alignments and dynamic programming: global alignment, local alignment. Substitution matrices: Pam and Blosum
3	Edit distance between two strings; calculation of the edit distance with dynamic programming. Edit Graphs; Weighted edit-distance, alphabet-weighted edit-distance.

SYLLABUS

Hrs	Frontal teaching
3	Heuristic alignments FASTA . BLAST.
4	Multiple sequence alignments Introduction to the multiple alignment problem; algorithms for Multiple alignment
6	Biological databases and similarity search.
5	Molecular evolution Introduction to the problem; molecular mechanisms of evolution processes; orthologous and paralogous Genes; Genetic distances between nucleotide sequences and amino acid sequences.
6	The molecular clock ; Molecular phylogenetics; Methods for the phylogenetic trees construction; UPGMA , Neighbor -joining.
2	PHYLIP Package
5	Markov chains; Markov sources; Hidden Markov models. Fair-Bet Casino Problem. Forward procedures; Viterbi algorithm
2	HMM for a phoneme ; Introduction to profile HMMs
3	Alignment-free methods