



MODELLO SCHEDA INSEGNAMENTO

Corso di L/LM/LMCU	Scienze e Tecnologie Genetiche
Denominazione insegnamento:	Bioinformatica
Numero di Crediti:	6
Anno:	II
Semestre:	II
Docente Titolare:	Luigi Cerulo
Dottorandi/assegnisti di ricerca che svolgono attività didattica a supporto del corso:	Teresa Noviello
Orario di ricevimento:	Mercoledì dalle 9 alle 11
Indirizzo:	via Port'Arsa 11

PRESENTAZIONE DEL CORSO:

Negli ultimi anni la bioinformatica gioca un ruolo fondamentale nell'avanzamento della ricerca biologica. Il corso di bioinformatica si propone di fornire allo studente i concetti teorici e gli strumenti pratici di base per trattare ed elaborare i dati biologici e per risolvere tipici problemi di genomica funzionale ed analisi di sequenza attraverso la progettazione di un algoritmo implementato con un linguaggio di programmazione su calcolatore.

GLI OBIETTIVI FORMATIVI

- Conoscere alcune importanti codifiche delle informazioni biologiche (database biologici, formati fasta, ecc.).
- Capacità di comprendere i problemi risolvibili mediante un algoritmo.
- Conoscere il linguaggio di programmazione R.
- Capacità di codificare un algoritmo mediante il linguaggio R ed eseguirlo su un calcolatore.
- Conoscere ed applicare con efficacia i principali algoritmi usati per le analisi di sequenza e la genomica funzionale.

PREREQUISITI RICHIESTI

Conoscenze logico-matematiche e informatiche di base

FREQUENZA DELLE LEZIONI

Data la natura degli argomenti trattati, la frequenza del corso è fortemente consigliata. Il corso prevede delle sessioni di esercitazioni durante le quali gli studenti possono fare esperienza diretta delle tecniche di programmazione che difficilmente sono reperibili sui libri di testo.

CONTENUTI DEL CORSO

1. Database biologici
 - 1.1. Codifica delle informazioni biologiche (i principali formati)
 - 1.2. Database biologici
 - 1.2.1. Nucleotide databanks: EMBL, NCBI, DDBJ
 - 1.2.2. UCSC
 - 1.2.3. Ensembl (BioMart)
 - 1.2.4. PDB, UniProt
2. Richiami di programmazione in R
 - 2.1. Concetto di algoritmo
 - 2.2. Linguaggio R
 - 2.2.1. Tipi di dati (numeric, stringa, logic)
 - 2.2.2. Variabili ed espressioni numeriche
 - 2.2.3. Modello di Memoria e istruzione di assegnazione
 - 2.2.4. Struttura di un programma, il modello input-elaborazione-output (istruzioni di input/output)
 - 2.3. Istruzione di controllo IF
 - 2.4. Istruzione iterativa WHILE
 - 2.5. Strutture dati elementari
 - 2.5.1. Vettore
 - 2.5.2. Matrice
 - 2.5.3. Tabella
3. Analisi di Sequenze
 - 3.1. Algoritmi di allineamento

- 3.1.1. Needleman-Wunsch
- 3.1.2. Smith-Waterman
- 3.1.3. Matrici di sostituzione: i modelli PAM e BLOSUM
- 3.1.4. Algoritmi di allineamento approssimati (BLAST e FASTA)
- 3.2. Identificazione di segnali
 - 3.2.1. Siti di legame sui promotori: modello PWM
 - 3.2.2. Uso della libreria Biostrings per l'analisi dei promotori
 - 3.2.3. Modelli markoviani per la ricerca di segnali più complessi
4. Genomica funzionale
 - 4.1. Analisi di geni differenzialmente espressi
 - 4.1.1. Metodi di normalizzazione
 - 4.2. Algoritmi di clustering dei dati
 - 4.2.1. Clustering gerarchico
 - 4.2.2. K-means
 - 4.3. Cenni su analisi di dati NGS

METODI DIDATTICI

- Lezioni frontali tramite lucidi
- Esercitazioni pratiche tramite il calcolatore del docente
- Sessioni di esercitazioni svolte dagli studenti e guidate dal docente

TESTI DI RIFERIMENTO

- Libro di testo: "Understandig Bioinformatics" di Jeremy O. Baum, Marketa J. Zvelebil (Garland Science)
- Dispense del docente
- Lucidi del docente

ESAME DI PROFITTO

L'esame di profitto consiste in una prova scritta ed una prova orale successiva se si supera la prova scritta.

La prova scritta consiste in 4 o 5 domande a difficoltà crescente e ha la durata di 2 ore. Ogni domanda può essere risposta mediante l'implementazione di un algoritmo. La correttezza dell'algoritmo e i commenti forniti saranno valutati con un punteggio in trentesimi. Il superamento della soglia di 18/30 consente l'accesso alla prova orale successiva.

La prova orale consiste in tre/quattro domande che vertono a verificare le conoscenze teoriche del programma svolto durante il corso. Il voto della prova orale sarà stabilito valutando la conoscenza degli argomenti in particolare degli algoritmi bioinformatici trattati, la capacità di ragionamento logico nella risoluzione dei problemi, la proprietà di linguaggio tecnico e la capacità espressiva complessiva dello studente.

Il voto finale dell'esame sarà calcolato facendo la media dei voti ottenuti nella prova scritta e orale.

CALENDARIO ESAMI

Rinvio al link

PRENOTAZIONE ESAMI

Rinvio al link

SYLLABUS

Argomenti	Ore	Riferimenti bibliografici	Tipologia di lezione
Database biologici	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Richiami di programmazione in R	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Eserciziario del docente	Esercitazione al calcolatore
Istruzione di controllo IF	2	Dispense del docente, Lucidi	Lezione frontale
Esercitazioni	2	Eserciziario del docente	Esercitazione al calcolatore
Istruzione interattiva WHILE	2	Dispense del docente, Lucidi	Lezione frontale
Esercitazioni	2	Eserciziario del docente	Esercitazione al calcolatore
Strutture dati elementari (vettori e matrici)	2	Dispense del docente, Lucidi	Lezione frontale
Esercitazioni	2	Eserciziario del docente	Esercitazione al calcolatore
Schemi tipici di elaborazione per vettori e matrici	2	Dispense del docente, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Algoritmi di allineamento	3	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Matrici di sostituzione (PAM e BLOSUM)	3	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore

Algoritmi di allineamento approssimati (BLAST e FASTA)	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Identificazione dei segnali: modello PWM	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Identificazione dei segnali complessi: modelli markoviani	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Genomica funzionale: analisi di geni differenzialmente espressi	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore
Algoritmi di clustering dei dati	2	Libro di testo, Lucidi	Lezione frontale
Esercitazioni	2	Esercizi forniti dal docente	Esercitazione al calcolatore