

Recupero di dati ed elaborazione di segnali e
immagini per bioinformatica

*MODULO: Riconoscimento e Recupero
dell'informazione per Bioinformatica*

Manuele Bicego

Corso di Laurea in Bioinformatica
Dipartimento di Informatica - Università di Verona

Il docente

Manuele Bicego

Dipartimento di informatica

Ufficio: Ca' Vignal 2 – Primo Piano – Stanza 1.48b

Telefono: 045 8027072

e-mail: manuele.bicego@univr.it

Ricevimento:

⇒ (Per il momento): su appuntamento concordato
via e-mail

Il corso

Modulo da 6 CFU: 4 CFU teoria, 2 CFU laboratorio

Orario:

Lunedì	13.30 – 16.30	Aula T.03
Martedì	14.30 – 17.30	Lab Alfa

NOTA: Il Laboratorio comincerà il 12 ottobre

Modalità di erogazione

- ♦ Modalità **principale**: in presenza
 - ♦ E' necessario utilizzare l'apposita applicazione "UNIVR Lezioni" per la prenotazione del posto in aula (<https://www.univr.it/it/app-univr-lezioni>).
- ♦ **Altre modalità**: per chi non può assistere in presenza, sarà possibile assistere alla lezione in streaming tramite il software Zoom (sicuramente voce e slides – per il video dipende dalle strumentazioni dell'aula)
 - ♦ nella piattaforma e-learning (moodle) troverete i link zoom e il materiale

Modalità di erogazione

- ♦ **Altre modalità:** Le lezioni verranno **registrate** e rese disponibili agli studenti tramite la pagina moodle del corso.
 - ♦ Saranno disponibili per tutta la durata del semestre, fino alla conclusione della prima sessione di esami che segue il completamento del corso.

Inviatemi una mail se intendete seguire il corso ma non avete i diritti per accedere allo spazio moodle

Pre-requisiti

Pre-requisiti per le attività in aula:

- ♦ Conoscenze di base di Probabilità, Statistica, Analisi

Pre-requisiti per le attività in laboratorio:

- ♦ Minima capacità di programmare
- ♦ Verranno rivisti i fondamenti di Matlab

Punto di vista

- ♦ **Titolo del modulo:** Riconoscimento e Recupero dell'informazione per bioinformatica
 - ⇒ è un titolo molto generico!
- ♦ In questo modulo: studio delle tecniche di "Pattern Recognition" per estrarre informazioni (da dati biologici)
 - ⇒ spesso alla base di programmi largamente utilizzati (ad esempio BLAST, Phylip, HMMER)

Obiettivi formativi

- ♦ Fornire le basi delle metodologie di Pattern Recognition
 - ⇒ Capire cos'è la pattern recognition
 - ⇒ Capire la differenza tra le diverse tipologie di problemi risolvibili con tecniche di pattern recognition
 - ⇒ Capire come creare un sistema automatico di pattern recognition
 - ⇒ Capire come validare i risultati ottenuti
 - ⇒ Vedere esempi di applicazione di tecniche di Pattern Recognition a problemi di bioinformatica
- ♦ L'attenzione è rivolta principalmente alla descrizione delle metodologie piuttosto che ai dettagli dei programmi applicativi (già visti /da vedere in altri corsi).

Programma (in generale)

Il corso si compone di due parti

- ♦ **Teoria:** in questa parte verranno presentate le diverse metodologie di Pattern Recognition, le motivazioni che portano al loro studio, e i problemi connessi al loro utilizzo.
 - ♦ Verranno inoltre analizzati alcuni problemi bioinformatici che sono classicamente risolti con metodologie di pattern recognition
- ♦ **Laboratorio:** verranno implementati in matlab semplici algoritmi di pattern recognition

Materiale didattico

Materiale didattico:

- ♦ lucidi del corso
- ♦ appunti presi a lezione (per lezioni alla lavagna)
- ♦ libri suggeriti
- ♦ articoli
- ♦ internet in generale.

Testi

- ♦ R. Duda, P. Hart, D. Stork *Pattern Classification*. Wiley, 2001 (2nd edition).
- ♦ P. Baldi, S. Brunak, *Bioinformatics, The Machine Learning Approach*. MIT Press, 2001
- ♦ G. Gan, C. Ma, J, Wu: *Data Clustering: Theory, Algorithms and Applications*, ASA-SIAM Series on Statistics and Applied Probability, 2007
- ♦ A.K. Jain and R.C. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, 1988.
 - ⇒ Disponibile on line
http://homepages.inf.ed.ac.uk/rbf/BOOKS/JAIN/Clustering_Jain_Dubes.pdf
- ♦ S. Theodoridis, K. Koutroumbas: *Pattern Recognition*, Second edition, Academic press, 2003

Altri testi consigliati

- ♦ N. Cristianini, M.W. Hahn: *Introduction to Computational Genomics*, Cambridge University Press, 2007
- ♦ W.J. Ewens, G.R. Grant: *Statistical Methods in Bioinformatics*, Springer 2001
- ♦ C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- ♦ W.J. Ewens, G.R. Grant, *Statistical Methods in Bioinformatics*. Springer, 2001
- ♦ E. Keedwell, A. Narayanan, *Intelligent Bioinformatics*. Wiley, 2005
- ♦ M. Berthold, D.J. Hand, *Intelligent Data Analysis*. Springer, 2003 (2nd edition).

Modalità d'esame

*NOTA: Vale per **tutto il corso** di Recupero di dati ed elaborazione di segnali e immagini per bioinformatica*

- ♦ PARTE 1 (15 punti): modulo di Riconoscimento e Recupero dell'informazione per Bioinformatica
- ♦ PARTE 2 (15 punti): modulo Segnali e Immagini 1

Modalità d'esame

*NOTA: Vale per **tutto il corso** di Recupero di dati ed elaborazione di segnali e immagini per bioinformatica*

- ♦ Il voto totale è la somma dei due scritti
- ♦ Occorre totalizzare almeno 9 punti in ognuno delle due parti
- ♦ Non è necessario passare le due parti nello stesso appello/sessione
- ♦ I voti dei moduli singoli rimangono validi per l'intero anno accademico (fino febbraio 2023)

Modalità d'esame

Dettagli sulla parte di Riconoscimento e Recupero dell'informazione per Bioinformatica

ESAME **SCRITTO** (se possibile)

- ⇒ Tre domande a risposta aperta su argomenti del corso
- ⇒ Esempio: “come funziona l’algoritmo K-means?”

Modalità d'esame

Dettagli sullo scritto di Riconoscimento e Recupero dell'informazione per Bioinformatica

- ⇒ Una domanda di “comprensione di codice” (relativo alla parte di laboratorio)
 - ♦ Esempio: “Il seguente codice implementa il calcolo della media? Perché?”

```
% x è un vettore di numeri  
m = 0;  
for i = 1:N  
    m = m+x(i);  
end  
m = m/(N-1);
```

Seguite il lab!!

NOTE FINALI

- ♦ Tutte le informazioni, il materiale didattico e gli aggiornamenti saranno disponibili sulla pagina moodle del corso
- ♦ Gli avvisi verranno pubblicati anche sulla pagina pubblica del corso
- ♦ Non esitate a contattarmi per qualsiasi dubbio o domanda