

Corso integrato di Management di Laboratorio
Modulo di **BIOINFORMATICA**
(5 CFU - 40 ore)

Dipartimenti di Area Medica - Università di Foggia
CdL in Scienze e Tecnologie Biomolecolari
a.a. 2019/2020

Prof. Crescenzo Gallo
CRESCENZIO.GALLO@UNIFG.IT

Contesto formativo

L'insegnamento dell'informatica è stato inserito ufficialmente nella Facoltà di Medicina e Chirurgia con il DPR 95 del 1986 (nuova Tab. XVIII) nel quale venivano indicati, tra gli altri, i seguenti obiettivi: *“lo studente deve essere capace di applicare il metodo sperimentale allo studio dei fenomeni della vita dimostrando di conoscere e saper utilizzare i principi fondamentali della fisica, statistica, matematica, informatica...”*.

Successivamente, con la revisione della Tab. XVIII, è stato eliminato il riferimento esplicito all'informatica, ma si è concessa maggiore autonomia alle singole Facoltà nel definire l'ordinamento degli Studi e si sono introdotte le **Attività Didattiche Opzionali** tra le quali vi è libertà di scelta da parte degli studenti.

Contesto formativo

Tale trattazione si può riassumere su tre livelli:

attività formative di base, in riferimento ai sistemi di elaborazione delle informazioni;

attività affini ed integrative, in cui esplicitamente parla di informatica, sistemi di elaborazione delle informazioni, bioingegneria elettronica e bioinformatica;

attività “altre”, tornando a riferirsi esplicitamente al laboratorio di informatica e sottolineando la necessità che gli studenti non solo conoscano ma acquisiscano abilità informatiche.

Obiettivi del Corso

- Il corso si propone l'obiettivo di introdurre gli allievi all'analisi e modellizzazione dei dati di natura biomedica secondo l'approccio “data mining” e fornire competenze per l'elaborazione dei dati mediante strumenti software open-source.
- Gli allievi dovranno acquisire la capacità di individuare i dati rilevanti in un studio, impostare i modelli di elaborazione, implementare e testare gli algoritmi di Machine Learning/Data Mining utilizzando tool software open-source.

Programma del Corso

- **Unità didattica 1:** La raccolta e la presentazione dei dati. L'approccio "tradizionale" all'analisi dei dati. Analisi dei dati ed estrazione della conoscenza secondo il modello concettuale del data mining. Problemi, modelli ed algoritmi di Machine Learning per la classificazione, supervisionata e non supervisionata. Metodi di valutazione della performance predittiva dei modelli.
- **Unità didattica 2:** Il tool Open Source di analisi dei dati Orange per la Bioinformatica: caratteristiche generali, download, installazione, configurazione. L'interfaccia applicativa interattiva di Orange: i Widget e il Workflow per l'implementazione di modelli di data mining.
- **Unità didattica 3:** Uso del linguaggio R per la Bioinformatica. Caratteristiche generali, strutture dati fondamentali, funzioni, pacchetti. Import/export di dati. R e database relazionali. Grafici con R. Uso di R per il machine learning.
- **Unità didattica 4:** Applicazioni della Bioinformatica -- Parte I. Accesso alle banche dati pubbliche di sequenze biologiche. Analisi statistica di sequenza biologiche, Allineamenti a coppie (pairwise), multipli (MSA) e alberi filogenetici. Gene-finding computazionale.
- **Unità didattica 5:** Applicazioni della Bioinformatica -- Parte II. Modelli nascosti di Markov (HMM). Estrazione e visualizzazione delle caratteristiche strutturali delle proteine. Analisi di dati da microarray. Analisi dei dati di spettrometria di massa (MS).

Avvertenze

- Ai fini dell'esame occorre conoscere il materiale illustrato a lezione.
- Durante lo svolgimento del corso sarà eventualmente reso disponibile ulteriore materiale didattico, “scaricabile” dal sito Web del corso.
- L'esame consiste in un test scritto preliminare, cui seguirà un orale di chiarimento e integrazione di quanto emerso dalla prova scritta.
- Per l'iscrizione all'esame si seguano le modalità abituali sulla piattaforma [ESSE3](#).
- Per qualsiasi esigenza si consiglia di utilizzare principalmente la posta elettronica, comunicando con il docente all'indirizzo crescenzo.gallo@unifg.it

Testi consigliati

- Dispense e lucidi del corso.
- M. Helmer Citterich, F. Ferré, G. Pavesi, G. Pesole, C. Romualdi (2018). *Fondamenti di Bioinformatica*. Zanichelli.
- C. Vercellis (2006). *Business intelligence: modelli matematici e sistemi per le decisioni*. McGraw-Hill.
- F. Azuaje, J. Dopazo (2005). *Data Analysis and Visualization in Genomics and Proteomics*. John Wiley & Sons.
- D. Talia, P. Trunfio, F. Marozzo (2016). *Data Analysis in the Cloud: Models, Techniques and Applications*. Elsevier.
- I.H. Witten, E. Frank, M.A. Hall (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier.
- S. Mitra, T. Acharya (2003). *Data Mining: Multimedia, Soft Computing, and Bioinformatics*. Wiley-Interscience.

Calendario delle lezioni (II semestre)

Per il calendario delle lezioni e gli esami consultare la **homepage del corso**.

+ |  <http://www.crescenziogallo.it/unifg/>

Home-page
del docente



Università degli Studi di Foggia

Prof. Crescenzo Gallo



Sezione
didattica

[Didattica](#) | [Tutorato](#) | [Curriculum vitae](#) | [StudentCard](#)
[Link utili](#) | [App iUnifg](#)

Qualifica

Ricercatore universitario a tempo indeterminato (SSD ING-INF/05)
Professore aggregato di "Informatica Medica" e "Sistemi di Elaborazione delle Informazioni"

Strutture di appartenenza

[Università degli Studi di Foggia](#) | [Dip.to di Medicina Clinica e Sperimentale](#)



Follow me
on ResearchGate



Dip.to di Medicina Clinica e Sperimentale

Studio n.9 - Ricevimento: **su appuntamento e-mail** - Tel. fisso: **+39 0881-588067** (interno: 25867)

Dip.to di Economia

Studio n.7 - Ricevimento: **su appuntamento e-mail** - Tel. fisso: **+39 0881-781801** (interno: 22801)

VoIP: **+39 0881-1896124** - Fax: **+39 178-2240797**

E-mail: **crescenzo.gallo@unifg.it** - PEC: **info@pec.crescenziogallo.it**



Dipartimento di Medicina Clinica e Sperimentale

Prof. Crescenzo Gallo



Attività didattica

Home-page
del corso

a.a. 2019/2020

Dipartimenti di Area Medica

CdL in Scienze e Tecnologie Biomolecolari
Bioinformatica - III anno (II semestre, 5 CFU, 40 ore)