



## SCHEMA DELL'INSEGNAMENTO (SI)

### "TECNICHE CHEMOINFORMATICHE NELL'ANALISI CHIMICA"

SSD CHIM/08

DENOMINAZIONE DEL CORSO DI STUDIO: CONTROLLO DI QUALITÀ

ANNO ACCADEMICO: **2021-2022**

#### INFORMAZIONI GENERALI - DOCENTE

DOCENTE: **FEDERICA MORACA**

TELEFONO: **081-678551**

EMAIL: **federica.moraca@unina.it**

#### INFORMAZIONI GENERALI - ATTIVITÀ

INSEGNAMENTO INTEGRATO (EVENTUALE):

MODULO (EVENTUALE):

CANALE (EVENTUALE):

ANNO DI CORSO: II/III

SEMESTRE: II

CFU: 6

## INSEGNAMENTI PROPEDEUTICI (se previsti dal Regolamento del CdS)

Nessuno

## EVENTUALI PREREQUISITI

Nozioni acquisite con lo studio della Chimica Generale, Chimica Organica, Informatica

## OBIETTIVI FORMATIVI

Il Corso si propone di fornire allo studente le conoscenze di base sulle moderne metodologie Chemoinformatiche per gestire ed estrarre le informazioni chimiche contenute in diversi database di molecole organiche (sostanze chimiche e/o farmaci) ed il loro utilizzo per l'analisi chimica. Tale obiettivo sarà raggiunto portando lo studente ad apprendere la rappresentazione molecolare 2D, a leggere i formati elettronici più comuni che definiscono le caratteristiche delle strutture chimiche (es. SMILES, MOL, InChI, SDF), a riconoscere e calcolare i principali descrittori molecolari, predire i parametri qualitativi chimico-fisici (es. pKa, lipofilia, solubilità in acqua) e farmacocinetici/tossicologici (ADMET) di sostanze chimiche e/o farmaci mediante l'utilizzo di appositi tool online. Il Corso sarà completato da esercitazioni pratiche in laboratorio chemoinformatico, utilizzando la piattaforma open source KNIME Analytics Platform, e riguarderanno l'estrazione e l'analisi di dati, da database molecolari di farmaci e/o sostanze chimiche, al fine di identificare composti reattivi/tossici e contaminanti, nonché di effettuare ricerche di similarità e sottostruttura. La parte pratica è finalizzata ad applicare in autonomia le conoscenze acquisite nella parte teorica, attraverso l'elaborazione di flussi di lavoro (workflow) sulla medesima piattaforma, liberamente accessibile e scaricabile.

## RISULTATI DI APPRENDIMENTO ATTESI (DESCRITTORI DI DUBLINO)

### Conoscenza e capacità di comprensione

Apprendimento dei principi di base delle tecniche chemoinformatiche. Saper ricercare, leggere e gestire le informazioni chimiche contenute in diverse banche dati molecolari tra cui: descrittori molecolari, caratteristiche chimico-fisiche e farmacocinetiche/tossicologiche.

### Capacità di applicare conoscenza e comprensione

Leggere e modificare i formati elettronici più comuni per definire strutture chimiche. Capacità di saper costruire in autonomia dei flussi di lavoro in grado di manipolare database molecolari e saper analizzare ed interpretare i dati finali.

## PROGRAMMA

**PARTE GENERALE.** *Introduzione alla normativa europea REACH:* registrazione, valutazione, autorizzazione e restrizione delle sostanze chimiche. *Introduzione alla Chemoinformatica:* concetto di banche dati (database), tipologie di banche dati usati in chimica per ricerche ed analisi di strutture di interesse (es. per similarità, sottostruttura): pubblici (PubChem, ChEMBL), commerciali (CAS) e chimici-farmaceutici (ZINC, ChemSpider, DrugBank); analisi di dati (Data mining). *Rappresentazione delle strutture molecolari:* lineare (SMILES e InChI) e bidimensionale con richiamo alle diverse tipologie di files (MOL; SDF; PDB). *I descrittori molecolari:* definizione e classificazione (topologici, geometrici, frammentali, di superficie molecolare). *Ricerca per similarità:* algoritmi di fingerprint 2D e ricerca per sottostruttura. *Predittori dei parametri qualitativi chimico-fisici e dei profili farmacocinetici/tossicologici:* pKa, lipofilia, solubilità in acqua, profilo ADMET (assorbimento-distribuzione-metabolismo-eliminazione-tossicità), tool predittivi online (es. SwissADME). *Introduzione alla piattaforma KNIME Analytics:* Interfaccia grafica, nodi, metanodi e componenti, KNIME Hub repository, costruzione dei processi di lavoro (workflow).

**CASI DI STUDIO.** I casi di studio prevedono esercitazioni guidate sull'utilizzo della piattaforma open source KNIME Analytics ([www.knime.com](http://www.knime.com)), che permette di assemblare dei nodi per la processazione di dati, in flussi di lavoro

(workflow), allo scopo di facilitare e migliorare l'analisi dei dati raccolti. Durante le esercitazioni guidate, lo studente apprenderà come costruire i workflow raggiungendo i seguenti obiettivi:

- Importare un insieme di dati (xls, csv, txt, SDF) (nodi di lettura);
- Manipolare database di molecole organiche attraverso l'utilizzo di connettori, nodi di concatenazione, nodi di filtraggio e nodi di raggruppamento;
- Esportare i dati combinati in un file excel o in un file DB (nodi di scrittura);
- Analisi dei dati.

## MATERIALE DIDATTICO

- **Libro di testo consigliato:**

M. Mabilia, M. Bacilieri, A. Bassan, L. Broccardo, E. Fioravanzo, S. Moro, L. Sartori, M. Stocchero  
Chemoinformatica: Soluzioni e strumenti per scienze e tecnologie biomediche. Springer Biomed 2012.

- **Slides e materiale didattico distribuito dalla docente a lezione e disponibile sul sito web**

- **Lecture di approfondimento**

Chen, W.L. (2006) Chemoinformatics: Past, Present, and Future. *J. Chem. Inf. Model.* 46:2230-2255.

Gasteiger, J. (2003) Handbook of chemoinformatics: from data to knowledge. Volumes 1–4. Wiley-VCH, Weinheim.

N.B. -> Lo studente può utilizzare altri testi in possesso ed attinenti al corso previa consultazione con la docente.

## MODALITÀ DI SVOLGIMENTO DELL'INSEGNAMENTO

La prima parte del Corso verrà erogata attraverso lezioni frontali, i cui argomenti verranno trattati con l'ausilio di diapositive PowerPoint. La seconda parte si svolgerà nel laboratorio chemo-informatico e prevede esercitazioni pratiche sulla piattaforma KNIME Analytics. Tali esercitazioni verranno effettuate dagli studenti sotto la supervisione della docente e avranno come obiettivo la costruzione di specifici flussi di lavoro.

## VERIFICA DI APPRENDIMENTO E CRITERI DI VALUTAZIONE

### a) Modalità di esame:

La verifica dell'apprendimento avviene attraverso un esame finale, che accerti sia l'acquisizione delle conoscenze teoriche, sia l'abilità nell'utilizzo della piattaforma KNIME.

L'esame si articola in prova	
scritta e orale	
solo scritta	
solo orale	x
discussione di elaborato progettuale	
altro	

In caso di prova scritta i quesiti sono (*)	A risposta multipla	
	A risposta libera	
	Esercizi numerici	

(\*) È possibile rispondere a più opzioni

### b) Modalità di valutazione:

Il voto finale d'esame è espresso in trentesimi da 18/30 a 30/30 e lode e tiene conto di:

- 1) Acquisizione di un linguaggio scientifico corretto e rigoroso adeguato alla disciplina;

- 2) Capacità di definire le caratteristiche di una molecola in base al valore dei predittori chimico-fisici;
- 3) Capacità di creare flussi di lavoro specifici con KNIME per una determinata analisi di dati;
- 4) Capacità di interpretare i dati finali.

L'attribuzione del voto avviene secondo i criteri riportati in Tabella:

<b>Voto</b>	<b>Descrittori</b>
< 18 <i>insufficiente</i>	<i>Conoscenze frammentarie e superficiali dei contenuti, errori nell'applicare i concetti, esposizione carente</i>
18 - 20	<i>Conoscenze dei contenuti sufficienti ma generali, esposizione semplice, incertezze nell'applicazione di concetti teorici</i>
21 - 23	<i>Conoscenze dei contenuti appropriate ma non approfondite, capacità di applicare i concetti teorici, capacità di presentare i contenuti in modo semplice</i>
24 - 25	<i>Conoscenze dei contenuti appropriate ed ampie, discreta capacità di applicazione delle conoscenze, capacità di presentare i contenuti in modo articolato.</i>
26 - 27	<i>Conoscenze dei contenuti precise e complete, buona capacità di applicare le conoscenze, capacità di analisi, esposizione chiara e corretta</i>
28 - 29	<i>Conoscenze dei contenuti ampie, complete ed approfondite, buona applicazione dei contenuti, buona capacità di analisi e di sintesi, esposizione sicura e corretta,</i>
30 30 e lode	<i>Conoscenze dei contenuti molto ampie, complete ed approfondite, capacità ben consolidata di applicare i contenuti, ottima capacità di analisi, di sintesi e di collegamenti interdisciplinari, padronanza di esposizione</i>