



Аналіз біологічних послідовностей
Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 – Інформаційні технології</i>
Спеціальність	<i>122 Комп'ютерні науки</i>
Освітня програма	<i>Комп'ютерні технології в біології та медицині</i>
Статус дисципліни	Вибіркова
Форма навчання	Денна
Рік підготовки, семестр	3курс, весняний семестр
Обсяг дисципліни	Загальна кількість 120 год.
Семестровий контроль/ контрольні заходи	Залік/МКР, РГР
Розклад занять	<i>Згідно розкладу лекції 2год/тиждень, практика – 2 год. тиждень</i>
Мова викладання	Українська
Інформація про керівника курсу / викладачів	Лектор: ст. викладач, Кисляк Сергій Володимирович, bmk-ksv-fbmi@111.kpi.ua Практичні: ст. викладач, Кисляк Сергій Володимирович, bmk-ksv-fbmi@111.kpi.ua
Розміщення курсу	https://classroom.google.com/

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчання та результати навчання

Актуальність. Стрімкий розвиток природничих наук в 21 сторіччі відбувається в рамках шостого технологічного укладу, який базується на досягненнях в області сучасних інформаційних технологій, біотехнологій, молекулярної біології та генної інженерії, а також систем штучного інтелекту. В таких умовах, велику увагу вчені приділяють розвитку наук про життя. Становлення та розвиток біоінформатики та суміжних з нею наук (наприклад, хемоінформатики та хемогеноміки) призводить до зміщення вектору розвитку сучасної медицини в напрямку персоналізованого принципу. Відповідно до такого підходу, кількісний та якісний склад кодуєчих ділянок геному людини може бути визначальним на етапі пошуку ефективної стратегії лікування, з застосуванням сучасних лікарських препаратів. Серед великої кількості хімічних сполук (більше 10 млн.), які були синтезовані (або потенційно можуть бути синтезовані) тільки більше 1000 є лікарськими препаратами. З біля 25 тисяч досліджених генів людини ідентифіковано 800, що є мішенями для лікарських препаратів. З урахуванням надзвичайно складного, багатостадійного процесу розробки нових препаратів медичного призначення існує необхідність в оптимізації даного процесу. Така можливість досягається за допомогою комп'ютерного дизайну лікарських препаратів (CADD).

Метою дисципліни є формування у студентів навичок роботи з популярними програмами та сервісами, що можуть бути застосовані для вирішення різноманітних задач обчислювальної молекулярної біології, таких як: проведення ресеквенування біологічних послідовностей; вивчення молекулярних основ генетичних захворювань; побудови філогенетичних дерев; проектування лікарських препаратів, ідентифікації генів. При цьому студенти отримують детальну інформацію щодо застосованих алгоритмів з можливістю їх подальшої реалізації та можливого удосконалення (оптимізації).

Інтегральна компетентність (ОП введено в дію Наказом ректора НОН/75/2022 від 15.02.2022 р.)

-**ІК**- Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми у галузі комп'ютерних наук або у процесі навчання, що передбачає застосування теорій та методів комп'ютерних наук, інформаційних технологій і характеризується комплексністю та невизначеністю умов. Згідно з вимогами освітньо-професійної програми здобувачі вищої освіти повинні засвоїти компетентності:

- **ЗК 3** - Знання та розуміння предметної області та розуміння професійної діяльності;
- **ЗК 11** - Здатність приймати обґрунтовані рішення;

Спеціальні (фахові) компетентності (ОП введено в дію Наказом ректора НОН/75/2022 від 15.02.2022 р.)

-**ФК19** - Здатність аналізувати масиви медико-біологічних даних сигналів, за допомогою машинних алгоритмів та статистичних методів, включаючи традиційні методи секвенування ДНК та конструювання сигнальних мереж за даними ДНК-мікрочіпів;

Програмні результати навчання (ОП введено в дію Наказом ректора НОН/75/2022 від 15.02.2022 р.)

ПР 22 - Використовувати методи біоінформатики для вирівнювання послідовностей, пошуку генів, збірки геномів, вирівнювання структур білків передбачення структур білків, передбачення експресії генів та білок-білкової взаємодії та реконструювання процесу еволюції;

Згідно з вимогами програми навчальної дисципліни студенти після засвоєння кредитного модуля мають продемонструвати такі результати навчання:

знання :

- основних концепцій аналізу біологічних текстів;
- методів оцінки якості даних секвенування;
- алгоритмів асемблювання геномів та оцінки якості отриманої збірки геному;
- молекулярної організації геномів та методів (алгоритмів) їх порівняльного аналізу;
- особливостей проведення ресеквенування геномів прокариот;
- методів розрахунку генетичних дистанцій;
- методів побудови філогенетичних дерев;
- алгоритмів пошуку генів,
- основних принципів та підходів щодо створення сучасних лікарських препаратів

уміння:

- використовувати програмні пакети та веб-сервіси при проведенні ресеквенування біологічних послідовностей;
- отримувати множинне вирівнювання, визначати еволюційну відстань між послідовностями та отримувати ультраметричні та адитивні філогенетичні дерева, проводити бутстреп аналіз;
- використовувати програмні пакети та веб-сервіси з метою побудови філогенетичних дерев для отримання інформації щодо динаміки зараження на коронавірус (Sars-Cov-2) по різних країнам світу;
- отримання навичок роботи з сучасними молекулярно-біологічними БД;
- працювати з геномними браузерями;
- проводити аналіз геномів еукаріот та визначати мутації, що призводять до генетичних захворювань;
- використовувати технології автоматизованого проектування лікарських препаратів in silico;
- здійснювати пошук та проводити аналіз просторових структур білків за допомогою бази даних RCSB Protein Data Bank;
- працювати з спеціалізованими програмними продуктами, що використовують для молекулярного моделювання, 3D візуалізації та конвертації текстових файлів.
- визначати просторову структуру білка-мішені за допомогою методу моделювання за гомологією та оцінювати якість отриманої моделі;
- моделювати взаємодії ліганд-рецептор та порівнювати результати докінг взаємодії білка-мішені з вже існуючими лігандами до даного рецептору;

- проводити аналіз властивостей лігандів та робити висновки щодо їх хімічної модифікації;
- передбачити білки-рецептори, на які буде впливати ліганд, що може бути лікарським препаратом;
- за допомогою фармакофорного моделювання знаходити нові ліганди для біологічних мішеней;
- визначати біологічну активність (або властивість) лігандів за допомогою методу QSAR ;
- використовувати технології автоматизованого проектування лікарських препаратів *in silico*;

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі)

Дисципліна відноситься до вибірових дисциплін циклу професійної підготовки і базується на знаннях з дисциплін: «Дискретна математика», «Основи інформатики та програмування», «Основи біології та медицини», «Теорія вибору альтернатив», «Проектування та аналіз обчислювальних алгоритмів», «Основи молекулярної біології та біоінформатики»

Дисципліна забезпечує наступні навчальні дисципліни та кредитні модулі: «Основи геноміки та протеоміки»

3. Зміст навчальної дисципліни

Вступна частина.

Розділ 1 Молекулярна еволюція та філогенетичний аналіз

Тема 2.1 Алгоритми парного та множинного вирівнювання біологічних послідовностей

Тема 2.2 Філогенетичний аналіз біологічних послідовностей

Тема 2.3 Еволюція нуклеотидних та амінокислотних послідовностей

Тема 2.4 Визначення генетичних дистанцій для нуклеотидних послідовностей

Тема 2.5 Дистанційні методи побудови ультра метричних філогенетичних дерев. Метод UPGMA, трансформованої дистанції.

Тема 2.5 Дистанційні методи побудови адитивних філогенетичних дерев. Метод приєднання сусідів, мінімуму еволюції.

Тема 2.6 Визначення довжин гілок філогенетичного дерева

Тема 2.7 Методи аналізу дискретних ознак

Тема 2.8 Переваги та недоліки методів побудови філогенетичних дерев. Бутстреп аналіз

Розділ 2 Моделювання просторової структури білків по гомології

Тема 2.1 Особливості просторової організації білків. База даних просторових структур Protein Data Bank

Тема 2.2 Формати даних SMILES, SMARTS , .pdb

Тема 2.3 Пошук білків-гомологів. Ідентифікація консервативних та варіабельних ділянок. Моделювання бічних ланцюгів.

Тема 2.4 Оптимізація білкової моделі методами молекулярної динаміки.

Тема 2.5 Методи оцінки якості білкової моделі.

Розділ 3 Автоматизоване проектування лікарських препаратів на основі структури біологічної мішені

Тема 3.1 Підготовка бібліотеки хімічних сполук. Міжнародний хімічний ідентифікатор IUPAC (InChI). Формати mol, mol2

Тема 3.2 Ідентифікація сайту зв'язування ліганду

Тема 3.3 Молекулярний докінг

Тема 3.4 Методи аналізу конформаційного простору ліганду

Тема 3.5 Оцінка енергії зв'язування комплексу ліганд - біологічна мішень

Тема 3.6 Фільтрація результатів віртуального скринінгу

Тема 3.7 Методи оцінки якості результатів молекулярного моделювання.

Тема 3.8 Особливості застосування та обмеження молекулярного докінгу

Розділ 4 Автоматизоване проектування лікарських препаратів на основі структури ліганду

Тема 4.1 Фармакофорний пошук на основі структури ліганду. Молекулярні дескриптори лігандів.

Тема 4.2 Метод QSAR (Quantative structure – activity relationship)

4. Навчальні матеріали та ресурси

Базова література:

1. Paul G. Higgs and Teresa K. Attwood. *Bioinformatics and molecular evolution*. – Blackwell publishing, 2005, ISBN (Pbk) 1-4051-0683-2
2. Nei, M. and Kumar S. (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.- 2004.- 418с.
3. Durbin, R., Eddy, S., Krogh, A. and Mitchison, G. *Biological Sequence Analysis Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge.
4. Кисляк С.В., Основи молекулярної біології та біоінформатики [Електронний ресурс] : навч. посіб. для студ. спеціальності 122 «Комп'ютерні науки та інформаційні технології спеціалізації «Інформаційні технології в біології та медицині» / С.В.Кисляк, Є.А.Настенко; КПІ ім. Ігоря Сікорського. – Електронні текстові данні (1 файл, 2957 Кбайт). – Київ : КПІ ім. Ігоря Сікорського, 2018. – 95 с.
5. Кисляк С.В., Голуб Н.Б., Дуган О.М., Аверьянова О.А. Моделювання молекулярної взаємодії [Електронний ресурс] : підручник для студ. спеціальності 162 «Біотехнології та біоінженерія» / С.В.Кисляк; КПІ ім. Ігоря Сікорського. – Електронні текстові данні (1 файл 27167 Кбайт). – Київ: КПІ ім. Ігоря Сікорського, 2023 – 203с.

Інформаційні ресурси

1. <https://ru.coursera.org/>
2. <http://ugene.net/ru/>
3. <https://www.ncbi.nlm.nih.gov/>
4. <https://www.megasoftware.net/docs>
5. <https://www.ncbi.nlm.nih.gov/>
6. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
7. <https://www.rcsb.org/>
8. <https://daylight.com/dayhtml/doc/theory/theory.smiles.html>
9. <https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>
10. <https://swissmodel.expasy.org/>
11. <https://saves.mbi.ucla.edu/>
12. <https://www.cgl.ucsf.edu/chimera/>
13. <https://pubchem.ncbi.nlm.nih.gov/>
14. <http://www.chemspider.com/>
15. <https://chemapps.stolaf.edu/jmol/jsmol/hackamol.htm>
16. http://openbabel.org/wiki/Main_Page
17. <https://www.click2drug.org/>

5. Методика опанування навчальної дисципліни (освітнього компонента)

Інформація (за розділами, темами) про всі навчальні заняття

Назви розділів і тем	Програмні результати навчання	Основні завдання	
		Контрольний захід	Термін виконання
Розділ 1 Молекулярна еволюція та філогенетичний аналіз			
Тема 2.1 Алгоритми парного та множинного вирівнювання біологічних послідовностей	ПР 22	Комп'ютерний практикум 1,2,3	1-3-й тиждень
Тема 2.2 Еволюція нуклеотидних та амінокислотних послідовностей	ПР 22	Комп'ютерний практикум 4	4-й тиждень
Тема 2.3 Філогенетичний аналіз біологічних послідовностей	ПР 22	Комп'ютерний практикум 5	4-й тиждень
Тема 2.4 Визначення генетичних дистанцій для нуклеотидних послідовностей	ПР 22	Комп'ютерний практикум 5,6	5-6-й тиждень
Тема 2.5 Дистанційні методи побудови ультра метричних філогенетичних дерев. Метод UPGMA, трансформованої дистанції.	ПР 22	Комп'ютерний практикум 7,8	7-й тиждень
Тема 2.6 Визначення довжин гілок філогенетичного дерева	ПР 22	Комп'ютерний практикум 9	8-й тиждень
Тема 2.7 Методи аналізу дискретних ознак	ПР 22	Опитування	9-й тиждень
Тема 2.8 Переваги та недоліки методів побудови філогенетичних дерев. Бутстреп аналіз	ПР 22	Опитування	9-й тиждень
Розділ 2 Моделювання просторової структури білків по гомології			
Тема 2.1 Особливості просторової організації білків. База даних просторових структур Protein Data Bank Моделювання бічних ланцюгів	ПР 22	Опитування	10-й тиждень
Тема 2.2 Формати даних SMILES, SMARTS , .pdb	ПР 22	Опитування	10-й тиждень
Тема 2.3 Пошук білків-гомологів. Ідентифікація консервативних та варіабельних ділянок.	ПР 22	Комп'ютерний практикум 10	11-й тиждень
Тема 2.4 Оптимізація білкової моделі методами молекулярної динаміки	ПР 22	Опитування	11-й тиждень
Тема 2.5 Методи оцінки якості білкової моделі.	ПР 22	Комп'ютерний практикум 11	12-й тиждень
Розділ 3 Автоматизоване проєктування лікарських препаратів на основі структури біологічної мішені			
Тема 3.1 Підготовка бібліотеки хімічних сполук. Міжнародний хімічний ідентифікатор IUPAC (InChI). Формати mol, mol2	ПР 22	Опитування	12-й тиждень

Тема 3.2 Ідентифікація сайту зв'язування ліганду	ПР 22	Опитування	12-й тиждень
Тема 3.3 Молекулярний докінг	ПР 22	Комп'ютерний практикум 12	13-й тиждень
Тема 3.4 Методи аналізу конформаційного простору ліганду	ПР 22	Опитування	13-й тиждень
Тема 3.5 Оцінка енергії зв'язування комплексу ліганд - біологічна мішень	ПР 22	Комп'ютерний практикум 13	14-й тиждень
Тема 3.6 Фільтрація результатів віртуального скринінгу	ПР 22	Комп'ютерний практикум 14	15-й тиждень
Тема 3.7 Методи оцінки якості результатів молекулярного моделювання	ПР 22	Опитування	15-й тиждень
Тема 3.8 Особливості застосування та обмеження молекулярного докінгу	ПР 22	Комп'ютерний практикум 15	16-й тиждень
Розділ 4 Автоматизоване проєктування лікарських препаратів на основі структури ліганду			
Тема 4.1 Фармокофорний пошук на основі структури ліганду. Молекулярні дескриптори лігандів.	ПР 22	Комп'ютерний практикум 16	17-й тиждень
Тема 4.2 Метод QSAR (Quantative structure – activity relationship)	ПР 22	Опитування	17-й тиждень

Рекомендації щодо засвоєння навчальних занять (у формі деталізованого опису кожного заняття та запланованої роботи):

№	Теми лекційних занять	Кількість годин
1	Молекулярні основи еволюції. Історія становлення молекулярної філогенетики. Еволюційне дерево життя, мутації, види мутацій, транзиції і трансверсії, механізми еволюції, гіпотеза молекулярного годинника, нейтральної еволюції. Література: 1,2,3	2
2	Філогенетичні дерева. Види філогенетичних дерев, структура дерева, адитивні та ультраметричні філогенетичні дерева Література: 1,2,3	2
3	Множинне вирівнювання біологічних послідовностей. Множинне вирівнювання за допомогою динамічного програмування, евристичні прогресивні алгоритми множинного вирівнювання, алгоритм Clustal, Muscle. Алгоритми парного вирівнювання, алгоритми побудови направляючого дерева, визначення оптимального множинного вирівнювання мінімізацією інформаційної ентропії. Література: 3,4	2
4	Еволюція нуклеотидних послідовностей. Визначення еволюційних дистанцій, p - дистанція, модель Джукса-Кантора, модель Кимури, модель Таджими-Неї, гамма дистанції, визначення математичної моделі Література:2	2
5	Еволюція амінокислотних послідовностей. p – дистанція, корекція p-дистанції, Пуассон скорегована дистанція, дистанція с поправкой Кимури, визначення дистанцій між амінокислотними послідовностями за допомогою матриць заміни амінокислот PAM, BLOSUM	2

	Література:2	
6	Дистанційні методи побудови ультра метричних філогенетичних дерев. Метод. UPGMA, метод трансформованої дистанції. Визначення довжин гілок та особливості побудови ультраметричного філогенетичного дерева. Проведення тесту на ультраметричність Література:1,2,3	2
7	Дистанційні методи побудови адитивних філогенетичних дерев. Метод приєднання сусідів, метод мінімуму еволюції. Особливості побудови адитивних філогенетичних дерев. Проведення тесту на адитивність. Метод найменших квадратів. Визначення довжин гілок філогенетичного дерева. Метод Фітча-Марголіаша. Визначення довжин гілок ультраметричних та адитивних філогенетичних дерев. Література:1,2,3	2
8	Методи аналізу дискретних ознак Метод максимальної економії, максимальної правдоподібності. Переваги та недоліки методів побудови філогенетичних дерев. Бутстреп аналіз. Визначення достовірності вузлів філогенетичного дерева, методи тестування філогенії. Література: 2	2
9	Молекулярна організація рибонуклеїнових кислот. Транспортні РНК; структура тРНК; основні елементи вторинної структури РНК. Методи передбачення вторинної структури РНК. Алгоритми передбачення вторинної структури РНК, алгоритм Нуссинів, алгоритм Зукера. Література:3	2
10	Особливості просторової організації білків. Хімічна будова білків. Амінокислоти, пептидний зв'язок і поліпептидний ланцюг. Вторинна структура. α -Спіраль; β -Структура. Спіраль ₃₁₀ . β -Поворот; глобулярна структура. Роль вторинної структури в утворенні глобули, стабілізація глобули, структурна класифікація глобулярних білків. Структура мембранних білків. формування торсійних кутів ϕ та ψ на рівні амінокислотних залишків білкових молекул. Література: 5	2
11	Особливості збереженні інформації про структуру ліганду та білкової мішені. Формати даних SMILES, SMARTS та .pdb. Міжнародний хімічний ідентифікатор IUPAC (InChI). Формати mol, mol2. Література: 5	2
12	Ефективні алгоритми пошуку білків-гомологів. Алгоритм BLAST. Особливості ідентифікації консервативних та варіабельних ділянок. Моделювання бічних ланцюгів. Гомологічне моделювання білків за допомогою веб-сервісу SwissModel Оцінка якості моделі. Стереохімічні параметри. Якість упаковки. Достовірність укладки. Література: 5	2
13	Основи молекулярної динаміки біополімерів	2
14	Основні етапи та особливості проведення віртуального скринінгу за умови наявної просторової структури біологічної мішені. Отримання та підготовка структури мішені; Ідентифікація сайту зв'язування ліганду; Підготовка бібліотеки хімічних сполук; Молекулярний докінг; Література: 5	2
15	Особливості проведення віртуального скринінгу за умови наявної просторової структури біологічної мішені. Розрахунок вільної енергії зв'язування та фільтрація отриманих результатів віртуального скринінгу. Оцінка якості результатів молекулярного моделювання. Література: 5	2
16	Автоматизоване проектування лікарських препаратів на основі структури ліганду. Фармокофорний пошук на основі структури ліганда. Література: 5	2

17	Автоматизоване проектування лікарських препаратів на основі структури ліганду. Метод QSAR. Література: 5	2
18	Етапи розробки сучасних лікарських препаратів. Огляд програмного забезпечення, що може бути застосоване на різних етапах <i>in silico</i> пошуку та створення лікарських препаратів. Література: 5	2
Разом		36

Практичні заняття

Основні завдання циклу практичних занять:

- працювати з програмами та веб-сервісами для вирішення задач молекулярної біології;
- орієнтуватись в базових алгоритмах геноміки, молекулярної філогенетики та алгоритмах передбачення структури РНК та пошуку генів.

№	Тематика комп'ютерних практикумів	Кіль-ть годин
1	Топологія, формати запису та різновиди філогенетичних дерев Кладограма, ультра метричне та адитивне філогенетичне дерево, нетривіальні гілки дерева, Newick формат. Література:1,2,3	2
2	Визначення швидкості синонімічних та не синонімічних замін нуклеотидних послідовностей. Транзиції, трансверсії, синонімічні та несинонімічні заміни, метод Нея-Годжборі. Література:1,2,3	2
3	Модифіковані алгоритми парного вирівнювання Алгоритм вирівнювання, що враховує повтори. Література:1,2,3	2
4	Алгоритм вирівнювання з афінною штрафною функцією. Трьохрівнева матриця динамічного програмування, алгоритм глобального та локального вирівнювання з використанням афінної штрафної функції. Література:1,2,3	2
5	Філогенетичний аналіз біологічних послідовностей. Множинне вирівнювання біологічних послідовностей. Розрахунок еволюційних дистанцій між амінокислотними послідовностями. Література:1,2,3	2
6	Ультраметричне філогенетичне дерево. Метод UPGMA. Ультраметричне філогенетичне дерево. Метод UPGMA. Метод трансформованої дистанції. Метод трансформованої дистанції. Література:1,2,3	2
7	Визначення довжини гілок адитивного філогенетичного дерева. Метод Фітча-Марголіаша. .Визначення довжини гілок адитивного філогенетичного дерева. Метод приєднання сусідів. Література:1,2,3	2
8	Оцінка надійності вузлів філогенетичного дерева. Бутстреп аналіз	2
9	Передбачення вторинної структури РНК. Алгоритм Нуссінов Алгоритм мінімізації енергії Цукера. Література:1,2,3	2
10	Моделюванні білків по гомології. Гомологічне моделювання білків за допомогою веб-сервісу SwissModel. Література: 5	2
11	Оцінка якості просторової моделі білкової молекули. Література: 1	2
12	Підготовка бібліотеки хімічних сполук. Фільтрація бібліотеки лігандів відповідно до правил п'яти Ліпінського. Література: 5	2
13	Молекулярний докінг лігандів за допомогою веб-сервісу Swiss-dock Література: 5	2
14	Молекулярний докінг лігандів за допомогою програмного забезпечення UCSF Chimera та Autodock vina. Література: 5	2
15	Молекулярний докінг лігандів за допомогою програмного забезпечення UCSF Chimera та Autodock vina. Література: 5	2

16	Проведення аналізу результатів порівняння просторових структур та їх послідовностей, за допомогою класичних алгоритмів біоінформатики. Література: 5	2
17	Пошук нових лігандів для біологічних мішеней. Фармакофорний пошук за допомогою веб-сервісів PharmaGist та Pharmit. Література: 5	2
18	Залікове заняття	2
Разом		36

6. Самостійна робота студента

Самостійна робота студента по дисципліні включає підготовку до аудиторних занять (18 годин), модульної контрольної (4 години), підготовка до заліку (6 годин), РГР (10) год.) та самостійне вивчення певних тем (10 годин), перелік яких наводиться нижче.

№	Назви тем і питань, що виносяться на самостійне опрацювання, та посилання на навчальну літературу	Кількість годин СРС
1.	Розділ 3 Методи структурної біоінформатики. Основи молекулярної динаміки біополімерів	10

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Правила відвідування занять. Відвідування лекцій, практичних занять, а також відсутність на них, не оцінюється. Студентам рекомендується відвідувати заняття, оскільки на них викладається теоретичний матеріал та розвиваються навички, необхідні для формування компетентностей, визначених стандартом освіти. Система оцінювання орієнтована на отримання балів за активність студента, а також виконання завдань, які здатні розвинути практичні уміння та навички. За об'єктивних причин (наприклад, хвороба, працевлаштування, міжнародне стажування тощо) навчання може відбуватися у додатковий час за погодженням із керівником курсу. Захист індивідуальних завдань проводиться на практичних заняттях.

Політика щодо дедлайнів та перескладання. Роботи, які здаються із порушенням термінів без поважних причин, оцінюються з пониженням рейтингової оцінки. Викладачем визначаються дедлайни виконання розрахункової роботи та захистів 16 практичних робіт. Політика та принципи академічної доброчесності визначені у розділі 3 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше можна ознайомитись за посиланням: <https://kpi.ua/code>.

Правила поведінки на заняттях. Норми етичної поведінки студентів і працівників визначені у розділі 2 Кодексу честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського». Детальніше за посиланням: <https://kpi.ua/code>. Використання додаткових джерел інформації під час оцінювання знань заборонено (у т.ч. мобільних пристроїв).

Дистанційне навчання

Дистанційне навчання відбувається через Платформу дистанційного навчання «Сікорський». Дистанційне навчання через проходження додаткових он-лайн курсів за певною тематикою допускається за умови погодження зі студентами. У разі, якщо невелика кількість студентів має бажання пройти он-лайн курс за певною тематикою, вивчення матеріалу за допомогою таких курсів допускається, але студенти повинні виконати всі завдання, які передбачені у навчальній дисципліні. Список курсів пропонується викладачем після виявлення бажання студентами (оскільки банк доступних курсів поновлюється майже щомісяця). Студент надає документ, що підтверджує проходження дистанційного курсу (у разі проходження повного курсу) або надає виконані практичні завдання з дистанційного курсу та за умови проходження усної співбесіди з викладачем за пройденими темами може отримати оцінки за контрольні заходи, які передбачені за вивченими

темами (експрес-контрольні / тестові завдання, практичні роботи). Виконання практичних робіт, а також виконання розрахунково-графічної роботи, здійснюється під час самостійної роботи студентів у дистанційному режимі (з можливістю консультування з викладачем через електронну пошту, соціальні мережі).

8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Поточний контроль: виконання практичних робіт (48 балів), виконання розрахункової роботи (25 балів), МКР (27 балів). Докладніша інформація щодо поточного контролю та критеріїв оцінювання наведена в PCO з дисципліни (додаток 1)

Календарний контроль: проводиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Семестровий контроль: залік.

Умови допуску до семестрового контролю: зарахування усіх практичних робіт, написання МКР, виконання та захист РГР та семестровий рейтинг $R_c > 40$. Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<i>Кількість балів</i>	<i>Оцінка</i>
100-95	Відмінно
94-85	Дуже добре
84-75	Добре
74-65	Задовільно
64-60	Достатньо
Менше 60	Незадовільно
Не виконані умови допуску	Не допущено

Рейтинг студента складається з балів, що він отримує за:

- 1) виконання завдань на 16 практичних заняттях;
- 2) 1 модульну контрольну роботу;
- 3) виконання та захист РР.

Система рейтингових (вагових) балів занять і рейтингових оцінок за видами контролю

№	Вид контролю	Бал	Кількість	Сума балів
1	Завдання на практичному занятті	3	16	48
	ваговий бал			
	якість виконання та захист ¹	0-3		
2	Модульна контрольна робота	25	1	25
	ваговий бал			
	якість виконання та захист ²	0-27		
3	РГР	27	1	27
	Якість виконання ³	0-27		
	Всього			100

¹ – Якість виконання завдання на практичному занятті:

- завдання виконане повністю, звіт надано своєчасно - 3 бали;

- робота містить несуттєві помилки або не повністю виконане завдання, звіт надано своєчасно - 2 бали;

- робота не зарахована, звіт надано не своєчасно з суттєвими помилками - 0 балів.

²- Якість виконання контрольної роботи:

- повна розкрита відповідь - 23 - 25 балів;

- помилка в одному завданні або неповна відповідь в двох завданнях - 18 - 23 балів;

- помилка в одному завданні або неповна відповідь на усі завдання - 13- 18 балів;

- помилка в двох завданнях - 8 - 13 балів;

- робота не зарахована - 0 балів.

³- Якість виконання РР та захист:

- захист роботи (захист включає знання з лекційного матеріалу) - 7 балів;

- правильно виконана робота - 16 - 20 балів;

- робота виконана з помилками в одному завданні - 12 - 16 балів;

- робота виконана з помилками у двох завданнях - 6 - 12 балів;

- робота не зарахована - 0 балів.

Штрафні бали:

№		
1	Несвоєчасна здача виконаного завдання з практичного заняття	-1 б

Сума як штрафних балів не має перевищувати $0,1 R_C = 60 \text{ балів} \times 0,1 = 6 \text{ балів}$

Розрахунок шкали (R) рейтингу

Сума вагових балів контрольних заходів протягом семестру складає:

$$R_C = 48 + 25 + 27 = 100 \text{ балів}$$

Форма атестації передбачена у вигляді заліку з сумарною кількістю балів 100.

Умови допуску до заліку: зарахування РГР, відпрацювання та захист 16 комп'ютерних практикумів, написання МКР не менше ніж на «задовільно», а також стартовий рейтинг (r_C) не менше 40% від R_C , тобто 40 балів. Студенти, які набрали протягом семестру рейтинг з кредитного модуля менше $0,6 R$, зобов'язані виконувати залікову контрольну роботу. Студенти, які набрали протягом семестру необхідну кількість балів ($RD \geq 0,6 R$), мають можливості:

- отримати залікову оцінку (залік) так званим «автоматом» відповідно до набраного рейтингу;
- виконувати залікову контрольну роботу з метою підвищення оцінки;
- у разі отримання оцінки, більшої ніж «автоматом» з рейтингу, студент отримує оцінку за результатами залікової контрольної роботи;
- у разі отримання оцінки меншої, ніж «автоматом» з рейтингу, кафедра застосовує жорстку РСО – попередній рейтинг студента з дисципліни скасовується і він отримує оцінку тільки за результатами залікової контрольної роботи.

Для отримання студентом відповідних оцінок (ECTS та традиційних) його рейтингова оцінка **RD** розраховується згідно з таблицею:

$RD = r_C + r_E$	Оцінка ECTS	Традиційна оцінка
95...100	A	відмінно
85...94	B	добре
75...84	C	
65...74	D	задовільно
64...60	E	
$RD \leq 60$	Fx	не задовільно
$r_C < 30$ або не виконані інші умови допуску до заліку	F	не допущений

9. Додаткова інформація з дисципліни (освітнього компонента)

Перелік питань для підготовки до модульної контрольної роботи та заліку наведено у Додатку 1. Дистанційне навчання через проходження додаткових онлайн-курсів за певною тематикою допускається, за умови погодження зі студентами. У разі, якщо невелика кількість студентів має бажання пройти онлайн-курс за певною тематикою, вивчення матеріалу за допомогою таких курсів допускається, але студенти повинні виконати всі завдання, що передбачені програмою навчальної дисципліни. Список курсів пропонується викладачем після виявлення бажання студентами, оскільки банк доступних курсів поновлюється майже щомісяця. Студент надає документ, що підтверджує проходження дистанційного курсу (у разі проходження повного курсу), або надає виконані практичні завдання з дистанційного курсу та, за умови проходження усної співбесіди з викладачем за пройденими темами, може отримати оцінки за контрольні заходи, що передбачені за вивченими темами.

Робочу програму навчальної дисципліни (силабус):

Складено: старшим викладачем кафедри біомедичної кібернетики Кисляком Сергієм Володимировичем

Ухвалено кафедрою біомедичної кібернетики (протокол № 18 від 24.06.2024)

Погоджено Методичною комісією ФБМІ (протокол № 9 від 26.06. 2024)

Перелік питань для підготовки до модульної контрольної роботи та заліку

1. Мутації. Основні види точкових мутацій. Транзиції та трансверсії. Не синонімічні та синонімічні заміни.
2. Особливості еволюції нуклеотидних послідовностей (первинні, вторинні та зворотні заміни, характеристика консервативних ділянок, конвергенція та дивергенція)
3. Концепція молекулярного годинника та нейтральна теорія еволюції.
4. Характеристика основних етапів філогенетичного аналізу біологічних послідовностей.
5. Множинне вирівнювання біологічних послідовностей. Алгоритм прогресивного множинного вирівнювання. Алгоритм Clustal в модифікації W. Оцінка якості вирівняних послідовностей. Ітеративні алгоритми множинного вирівнювання біологічних послідовностей.
6. Опис консенсусного рядка програми у множинному вирівнюванні. Дати визначення: домен, мотив, патерн та сайт білкової послідовності. Правила створення запитів у БД Prosite.
7. Спостережені, істинні та розрахункові еволюційні дистанції між послідовностями. P-дистанція. Обмеження та недоліки використання P-дистанції.
8. Визначення генетичних відстаней між нуклеотидними послідовностями. Модель Джукса-Кантора, Кімури, Таджими-Неї.
9. Визначення еволюційних дистанцій між амінокислотними послідовностями.
10. Філогенетичні дерева. Топологія, види дерев, ультраметричні та адитивні філогенетичні дерева.
11. Методи побудови філогенетичних дерев. Загальна характеристика дистанційних методів побудови дерев філогенезу.
12. Метод неваженого попарного групування із середнім арифметичним (UPGMA).
13. Метод трансформованої дистанції.
14. Метод приєднання сусідів (Neighbor-Joining).
15. Метод Фітча – Марголіаша
16. Метод максимальної економії (Maximum parsimony)
17. Статистична оцінка достовірності вузлів філогенетичного дерева. Бутстреп-аналіз.
18. Основні елементи вторинної структури РНК. Алгоритми передбачення вторинної структури РНК. Алгоритм Нуссінов. Алгоритм Зукера.
19. Особливості запису інформації з урахуванням хімічної структури та іонного складу молекул в лінійній нотації Smiles та Smarts.
20. Застосування веб-сервісу Open Babel з метою перетворення форматів даних. Для заданої структурної формули молекули отримати лінійний запис у форматі Smiles.
21. Особливості запису молекулярно-біологічної інформації у форматі Fasta. Для відповідної послідовності у Fasta форматі знайти інформацію про просторову структуру.

22. Основи молекулярного моделювання білків. Принципи структурної організації білків. Первинний, вторинний, третинний та четвертинний рівень просторової організації білків. Конформаційні властивості білків. Торсійні кути бокових ланцюгів N-C α (кут ϕ) та C-C α (кут ψ). Карти Рамачандрана.
23. Моделювання білків по гомології. Для білкової біологічної послідовності знайти гомолога з кращою роздільною здатністю.
24. База даних просторових структур Protein Data Bank (PDB). Особливості запису молекулярно-біологічної інформації у форматі .pdb.
25. Процедура оцінки якості отриманої просторової моделі. Критерії оцінки якості моделі (стереохімічна точність, якість упаковки, достовірність укладки). Еталонні значення стереохімічних параметрів, як індикатор якості отриманої моделі. Програма PROCHECK.
26. Віртуальний скринінг - як базовий метод пошуку нових лікарських препаратів. Зменшення кількості хімічних сполук – кандидатів, при пошуку в сучасних базах даних лігандів. Правило п'яти Ліпінського.
27. Особливості моделювання ліганд рецепторної (докінг) взаємодії. Відбір результатів докінга. Веб-сервіс моделювання докінг взаємодії SwissDock.
28. Створення моделі фармакофора. Характеристика основних фармакофорних ознак, що необхідно врахувати при побудові моделі. Веб -сервіс Pharma Gist.
29. Пошук нових лікарських препаратів на основі моделі фармакофора. Інструмент ZINCPharmer.
30. Особливості опису рядків текстового файлу з атомними координатами у форматі .pdb. Параметри B-value та occupancy.
31. Формати даних *smiles* та *smarts* з лінійною нотацією. Основні вимоги до текстових файлів з лінійною нотацією у форматі *smiles*.
32. Які переваги має формат *smarts* у порівнянні з *smiles*. Логічні оператори формату *smarts*?
33. Основні етапи проведення гомологічного моделювання. Які алгоритми лежать в основі отримання інформації про просторову структуру білкової молекули відповідно до методу моделювання по гомології?
34. Особливості алгоритмів парного вирівнювання (глобальне з афінним штрафом) методом динамічного програмування.
35. Особливості пошуку білка – моделі, що може бути використаний для гомологічного моделювання. Яке значення E-value буде відповідати для знахідки, що потенційно буде гомологом до послідовності-запиту. Який відсоток ідентичності послідовностей може вказувати на їх гомологію?
36. Оптимізація отриманої моделі. Основи молекулярної динаміки біополімерів. Особливості визначення загальної E_{total} потенційної енергії молекулярної системи.

37. Загальна характеристика базових сил, що враховуються при розрахунках повної E_{total} потенційної енергії молекулярної системи.
38. Особливості просторової організації білків. Первинна, вторинна, третинна та четвертинна структури.
39. Основні критерії визначення якості отриманої моделі білка. Стереохімічні параметри. Якість упаковки. Достовірність укладки.
40. Особливості формування торсійних кутів ϕ та ψ на рівні амінокислотних залишків білкових молекул.
41. Характеристика статистичних параметрів, що використовують для оцінки якості отриманої моделі сервісу SwissModel. Параметри QSQE (Quaternary Structure Quality Estimation), GMQE (Global Model Quality Estimate), QMEANZ-score.
42. Особливості організації бази даних UniProt.
43. Загальна характеристика основних методів CADD.
44. Особливості проведення CADD за умови відомої просторової структури біологічної мішені.
45. Особливості проведення CADD за умови відомої просторової структури ліганду.
46. Загальна схема робочого процесу автоматизованого проєктування лікарських препаратів.
47. Основні етапи та особливості проведення віртуального скринінгу за умови відомої просторової структури біологічної мішені.
48. Загальна характеристика бібліотек хімічних сполук, які використовуються для проведення віртуального скринінгу.
49. Особливості фільтрації бібліотеки лігандів. Правило п'яти Ліпінського.
50. Формати даних, що використовуються для збереження інформації про просторові структури лігандів. Формати *mol* та *mol2*.
51. Загальна характеристика та особливості проведення конформаційного пошуку лігандів. Алгоритми систематичного та стахостичного пошуку.
52. Оцінка енергії зв'язування ліганд – біологічна мішень. Оціночні функції на основі силового поля, на основі знань. Емпіричні оціночні функції.
53. Оцінка якості результатів молекулярного моделювання.
54. Алгоритми парного вирівнювання методом динамічного програмування. Алгоритм Ніделмана-Вунша з афінним штрафом. Матриці замін амінокислот PAM та BLOSUM.
55. Алгоритми множинного вирівнювання. Многочисленне динамічне програмування. Прогресивний алгоритм множинного вирівнювання ClustalW. Ітераційні та стахостичні алгоритми множинного вирівнювання. Алгоритм Muscle.
56. Особливості стабілізації просторової структури білкових молекул. Основні елементи вторинної структури білків.
57. Особливості створення запитів в Бд Prosite. Мова паттернів Бд Prosite.
58. Особливості застосування та обмеження молекулярного докінгу.

59. Вплив молекул води на результати молекулярного докінгу.
60. Особливості проведення молекулярного докінгу в активний центр з кофактором.
61. Вплив ковалентних зв'язків на результати молекулярного докінгу.
62. Особливості проведення молекулярного докінгу за допомогою програмного забезпечення UCSF Chimera
63. Особливості проведення віртуального скринінгу за умов відсутньої просторової структури білка – мішені.
64. Основи фармакофорного моделювання. Фармакофорний пошук. Основні дескриптори фармакофору.
65. Класифікація молекулярних дескрипторів. Особливості топологічних дескрипторів.
66. Загальна характеристика основних етапів розробки нових лікарських препаратів. Можливості CADD з точки зору оптимізації процесу розробки лікарських препаратів.
67. Базові метрики, що використовують для визначення подібності/відмінності між молекулами, Коефіцієнти Танімото, Дайса, Евклідова відстань та відстань Хеммінга.
68. Алгоритм Вагнера-Фішера визначення редакційної відстані між біологічними послідовностями .
69. Основи QSAR моделювання. Базові дескриптори методу QSAR.
70. Особливості валідації лінійної моделі QSAR. Тест та крос-валідація. Коефіцієнти детермінації для тест та крос-валідації. Лінійний коефіцієнт кореляції.