

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Химический факультет

УТВЕРЖДЕНО:

И.о. декана

А. С. Князев

Оценочные материалы по дисциплине

Хемоинформатика

по направлению подготовки

04.04.01 Химия

Направленность (профиль) подготовки:

Трансляционные химические и биомедицинские технологии

Форма обучения

Очная

Квалификация

Магистр

Год приема

2023

СОГЛАСОВАНО:

Руководитель ОП

И.А. Курзина

Председатель УМК

Л.Н. Мишенина

1. Компетенции и индикаторы их достижения, проверяемые данными оценочными материалами

Целью освоения дисциплины является формирование следующих компетенций:

ПК-1 Способен планировать работу и выбирать адекватные методы решения научно-исследовательских и/или производственных задач в выбранной области химии, химической технологии или смежных с химией науках.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК 1.1 Разрабатывает стратегию научных исследований, составляет общий план и детальные планы отдельных стадий

ИПК 1.2 Выбирает экспериментальные и расчетно-теоретические методы решения поставленной задачи, используя достижения современной химической науки, и исходя из имеющихся, материальных, информационных и временных ресурсов

ИПК 1.3 Использует современное физико-химическое оборудование для получения и интерпретации достоверных результатов исследования в выбранной области химии, химической технологии или смежных с химией науках, применяя взаимодополняющие методы исследования

Проводит поиск, анализирует и обобщает результаты патентного поиска по тематике исследовательской работы

2. Оценочные материалы текущего контроля и критерии оценивания

Элементы текущего контроля:

- тесты;
- практическое задание;
- индивидуальное задание.

Индивидуальное задание (ИПК-1.1, ИПК-1.2, ИПК-1.3)

- Для заданной органической молекулы сгенерировать 2D и 3D представления структуры, сохранить их в виде файлов подходящих форматов, охарактеризовать молекулу набором дескрипторов.

- Для заданной органической молекулы оценить параметры ADME и токсичность с использованием on-line сервисов.

Критерии оценивания:

«отлично» выставляется магистранту, если показаны всесторонние и глубокие знания программного материала;

«хорошо» выставляется магистранту при ответе на поставленные вопросы и задания, в случае если были допущены единичные существенные неточности;

«удовлетворительно» выставляется магистранту, допустившему погрешности в ответе, но обладающему необходимыми знаниями для их устранения под руководством преподавателя;

«неудовлетворительно» выставляется магистранту, показавшему пробелы в знании основного учебно-программного материала, допустившему принципиальные ошибки при ответе на вопрос.

Практическое задание (ИПК-1.1, ИПК-1.2, ИПК-1.3)

1. С использованием базы ChEMBL отберите данные по активности соединений против циклооксигеназы-2 (COX-2). Отберите наиболее надежные данные. Проведите химическую и математическую чистку данных. Отберите 10% набора в качестве внешнего валидирующего набора случайным образом. С помощью программы ISIDA-MLR и различных типов фрагментных дескрипторов постройте несколько QSAR моделей (не менее 10). Для валидации модели используйте 5-кратную кросс-валидацию и внешний

валидирующий набор. Отберите лучшую модель. Используйте эту модель для скрининга данных из базы ZINC. В качестве набора возьмите только «лекарствоподобные соединения» базы. Из числа соединений, показавших наибольшую активность ($IC_{50} < 10^{-7}$ моль/л), отберите 5%, имеющих наименьшую сумму расстояний от соединений, использованных для тренировки модели, в пространстве фрагментных дескрипторов.

2. Провести поиск конформаций и выполнить оптимизацию геометрии заданного химического соединения методом молекулярной механики.

3. Выполнить молекулярный докинг малой органической молекулы в сайт связывания заданной белковой макромолекулы. Дать описание основных взаимодействий лиганд-рецептор в полученном комплексе.

- Для заданного химического соединения выполнить оптимизацию геометрии, найти частоты нормальных колебаний и оценить значения термодинамических параметров с применением квантовохимического метода DFT.

Критерии оценивания:

«отлично» выставляется магистранту, если показаны всесторонние и глубокие знания программного материала;

«хорошо» выставляется магистранту при ответе на поставленные вопросы и задания, в случае если были допущены единичные существенные неточности;

«удовлетворительно» выставляется магистранту, допустившему погрешности в ответе, но обладающему необходимыми знаниями для их устранения под руководством преподавателя;

«неудовлетворительно» выставляется магистранту, показавшему пробелы в знании основного учебно-программного материала, допустившему принципиальные ошибки при ответе на вопрос.

Тесты (ИПК-1.2)

1. Какие из указанных SMILES соответствует молекуле аспирина?

- a. CC(=O)Oc1ccccc1C(O)=O
- b. c1c(C(O)=O)cc(OC(O)C)ccc1
- c. OC(=O)c(cccc1)c1OC(=O)C
- d. c1(C(=O)O)ccccc1OC(=O)C

2. Какие InChI для приведенной молекулы гуанина соответствуют молекуле и являются стандартными?

- a. InChI=1/C5H5N5O/c6-5-9-3-2(4(11)10-5)7-1-8-3/h1H,(H4,6,7,8,9,10,11)/f/h8,10H,6H2
- b. InChI=1S/C5H5N5O/c6-5-9-3-2(4(11)10-5)7-1-8-3/h1H,(H4,6,7,8,9,10,11)
- c. InChI=1S/C6H6N5O2/c6-5-9-3-2(4(11)10-5)7-1-8-3/h1H,(H4,6,7,8,9,10,11)
- d. InChI=1/C5H5N5O/c6-5-9-3-2(4(11)10-5)7-1-8-3/h1H,(H4,6,7,8,9,10,11)/f/h7,9H,6H2

3. Какой SMARTS запроса будет определять выделенную подструктуру в приведённой молекуле? Атомы водорода не принимать во внимание.

- a. N~*~*~N
- b. NcccN
- c. [#7]ccc[#7]
- d. [NH2]aaa[NH2]

4. Какая из приведенных SMILES удовлетворяет приведенной структуре Маркуша?

- a. OCCc1c(C)cccc1
- b. CCCCCCCCC1cc(C(=O)O)ccc1
- c. CCCCc1ccc(C(C)=O)cc1
- d. CCCCc1ccc(COC=O)cc1

5. Какое из приведенных отнесений отмеченных фармакофорных центров 1,2 и 3 является наиболее полным и корректным (один ответ)? Обозначения: N - negative charge, P - positive charge, H - hydrophobe, Ar - aromatic ring, A - H-акцептор, D - H-donor.

- a. 1: N; 2: D; 3: P;
- b. 1: N; 2: A, D; 3: P, D;
- c. 1: N, A; 2: A, D; 3: P, D, Ar;
- d. 1: N, A; 2: A, D, N; 3: P, A, D, Ar;

6. Две структуры задаются указанными ниже битовыми строками. Какой будет индекс схожести Танимото между данными структурами?

Mol 1 1 0 0 0 1 1 1 0 1 1

Mol 2 0 1 0 0 1 0 1 1 0 1

- a. 6/11
- b. 5/8
- c. 3/10
- d. 3/8

7. Какое из приведенных выражений содержит формулу для вычисления индекса схожести Тверского? a - число активных бит в одной молекуле, b - число включенных бит в другой молекуле, c - число бит, которые являются активными в обеих молекулах.

- a. $c/(a+b-c)$
- b. $2c/(a+b)$
- c. $(a+b-2c)/(a+b-c)$
- d. $c/(c+\alpha(a-c)+\beta(b-c))$

8. В какой из приведенных баз данных можно найти информацию, характеризующую прочность связывания данного химического соединения с различными белками?

- a. CAS
- b. PubChem
- c. ChEMBL
- d. ZINC

9. Какую информацию о соединении можно найти в базе ChemSpider?

- a. Химическая структура
- b. Информация об испытании данного соединения на bioassay
- c. Кристаллическая структура молекулы
- d. Индекс LASSO, характеризующий насколько данная молекула подходит для

связывания с активными центрами различных ферментов

10. Какие этапы входят в процесс осуществления поиска по структуре?

- a. Стандартизация соединения
- b. Генерация хэш-кода
- c. Поиск молекулы с помощью скринов
- d. Поиск индекса схожести данного соединения с другими соединениями базы

Критерии оценивания:

Уровень выполнения текущих тестовых заданий оценивается в баллах, которые затем переводятся в оценку. Баллы выставляются следующим образом:

- правильное выполнение задания, где надо выбрать верный ответ – 1 балл.

Оценка соответствует следующей шкале:

| Отметка | Кол-во баллов |
|---------------------|---------------|
| Отлично | 7-10 |
| Хорошо | 5-6 |
| Удовлетворительно | 4 |
| Неудовлетворительно | 3 и менее |

3. Оценочные материалы итогового контроля (промежуточной аттестации) и критерии оценивания

Зачет по курсу «Хемоинформатика» проводится в форме устного опроса студентов, проверяющего освоение компетенций ИПК-1.1, ИПК-1.2, ИПК-1.3. Результаты зачета определяются оценками «зачтено» или «не зачтено».

Вопросы для подготовки к зачету:

1. Основные способы представления объектов в хемоинформатике;
2. Линейные представления;
3. Матричные представления;
4. Битовые представления;
5. Стандартные файлы в хемоинформатике;
6. Хеометрика и хемоинформатика;
7. Биоинформатика и хемоинформатика;
8. Дизайн библиотек данных;
9. Разбросанные и сфокусированные библиотеки;
10. Генерация структур;
11. RECAP;
12. Кластеризация молекул;
13. Иерархические и неиерархические подходы;
14. Отбор молекул без кластеризации;
15. Основные способы валидации моделей;
16. Классификационные модели;
17. Регрессионные модели;
18. Отбор дескрипторов;
19. Навигация в химическом пространстве как способ моделирования;
20. Методы машинного обучения.
21. Молекулярная динамика, ее механические и идейные основы;
22. Физическая природа потенциалов молекулярных взаимодействий и их функциональный вид;
23. Уравнения движения молекулярной системы. Их разностная аппроксимация;
24. Моделирование динамики конденсированных систем. Типы ансамблей. Периодические граничные условия;
25. Алгоритм Верле (составление списка соседей) для вычисления невалентных взаимодействий;
26. Температура. Способы оценки и вычисления. Термостатирование молекулярной системы;
27. Учет растворителя. Явный и неявный учет растворителя;
28. Вычисление давления в малых молекулярных системах. Баростат Берендсена;
29. Моделирование биологических макромолекул. Основы подхода. Назначение моделирования;
30. Общая схема молекулярно-динамического вычислительного эксперимента;
31. Обработка траекторий молекулярной динамики. Временные и пространственные автокорреляционные функции.

Критерии оценивания:

«Зачтено» – студент выполнил индивидуальное задание, тест и практическое задание правильно или с небольшими неточностями. Показал успешное или сопровождающееся отдельными ошибками применение навыков полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала. Ответил на большинство дополнительных вопросов правильно.

«Не зачтено» – студент имеет пробелы по отдельным теоретическим разделам дисциплины, контрольного задания аспирант продемонстрировал частично освоенное

умение и применение полученных навыков при решении профессиональных задач в рамках усвоенного учебного материала. При ответах на дополнительные вопросы было допущено множество неточностей.

Информация о разработчиках

Хлебников Андрей Иванович, д-р хим. наук, профессор, кафедра природных соединений, фармацевтической и медицинской химии, ХФ ТГУ.