

УТВЕРЖДАЮ
Директор ВШТЭ



Рабочая программа дисциплины

Б1.В.10 Основы хемоинформатики

Учебный план: _____ ФГОС3++b010302БИ-1_23-14.plx

Кафедра: **16** Прикладной математики и информатики

Направление подготовки:
(специальность) 01.03.02 Прикладная математика и информатика

Профиль подготовки:
(специализация) Биоинформатика

Уровень образования: бакалавриат

Форма обучения: очная

План учебного процесса

Семестр (курс для ЗАО)	Контактная работа обучающихся		Сам. работа	Контроль, час.	Трудоё мкость, ЗЕТ	Форма промежуточной аттестации
	Лекции	Практ. занятия				
7	УП	17	34	56,75	0,25	Зачет
	РПД	17	34	56,75	0,25	
Итого	УП	17	34	56,75	0,25	
	РПД	17	34	56,75	0,25	

Рабочая программа дисциплины составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки России от 10.01.2018 г. № 9

Составитель (и):

старший преподаватель

Кушнеров А.И.

От кафедры составителя:

Заведующий кафедрой прикладной математики и информатики

Яковлев В.П.

От выпускающей кафедры:

Заведующий кафедрой

Яковлев В.П.

Методический отдел:

Смирнова В.Г.

1 ВВЕДЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

1.1 Цель дисциплины: формирование у студентов устойчивых представлений о хемоинформатике.

1.2 Задачи дисциплины:

1. ознакомить с основами и методами хемоинформатики;
2. сформировать умения построения SAR/QSAR зависимостей;
3. освоить специализированные программные средства и системы тестирования в области хемоинформатики.

1.3 Требования к предварительной подготовке обучающегося:

Предварительная подготовка предполагает создание основы для формирования компетенций, указанных в п. 2, при изучении дисциплин:

Базы данных

Языки и методы программирования

Компьютерное моделирование

Методы анализа органических соединений

Проектирование и сопровождение биоинформационных систем

Основы химии

Введение в биоинформатику

2 КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-7: Способен проводить тестирования по разработанным тестовым случаям в области биоинформатики

Знать: основные понятия, определения, методы и подходы, используемые в хемоинформатике; способы построения SAR/QSAR зависимостей.

Уметь: создавать собственные базы данных; строить простейшие зависимости SAR/QSAR/QSPR, определять их статистическую значимость.

Владеть: навыками решения химических задач с использованием средств хемоинформатики.

ПК-8: Способен анализировать результаты тестирования в области биоинформатики

Знать: способы представления химических данных, методы осуществления поиска в химических базах данных; основные химические базы данных и методы работы с ними.

Уметь: подготавливать и курировать данные; оперировать данными созданных баз, проводить поиск в базах данных; использовать методы SAR/QSAR/QSPR в создании веществ с заданными свойствами.

Владеть: навыками анализа химических баз данных для решения прикладных задач; основными программами для создания и оперирования базами данных; навыками построения и использования SAR/QSAR/QSPR зависимостей.

3 РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Наименование и содержание разделов, тем и учебных занятий	Семестр (курс для ЗАО)	Контактная работа		СР (часы)	Инновац. формы занятий	Форма текущего контроля
		Лек. (часы)	Пр. (часы)			
Раздел 1. Компьютерное представление химической информации	7					0
Тема 1. Введение в хемоинформатику. Основные понятия хемоинформатики. Квантовая химия. Химическое пространство. Хемометрика. Биоинформатика		2	2	3		
Тема 2. Представление молекул в хемоинформатике. Линейные представления. Представление графов. Трехмерные представления молекул. Форматы данных. Методы конвертации. Представление химических реакций.		2	2	3		
Раздел 2. Химические базы данных						
Тема 3. Общие сведения и поиск в базах данных химической информации Классификация баз данных. Виды поиска, основанного на структуре. Поиск по сходству. Поиск в базах данных химических реакций. Поиск в базах данных трехмерных структур.		1	2	3		
Тема 4. Алгоритмы на графах. Основные определения теории графов. Изоморфизм, автоморфизм и симметрия графов. Алгоритмы прохождения деревьев решений.	1	2	3		0	
Тема 5. Анализ химических баз данных. CAS/SciFinder. Reaxys. BindingDB. Кембриджская структурная база данных (CSD). Базы данных профайлинга химических соединений.	1	2	3		ИЛ	
Раздел 3. Статистические модели хемоинформатики						0

<p>Тема 6. Дескрипторы в хемоинформатике.</p> <p>Классификация. Топологические, трехмерные, фрагментные, фармакофорные дескрипторы. Константы заместителей. Физико-химические дескрипторы. Программы расчета дескрипторов.</p>	1	2	4		
<p>Тема 7. Разработка моделей.</p> <p>Обработка данных. Основы построения моделей. Валидация моделей. Область применимости моделей. Модели "структура-свойство".</p>	1	2	4		
<p>Тема 8. Методы SAR/QSAR/QSPR в создании веществ.</p> <p>Методы 3D-QSAR зависящие от выравнивания молекул. Методы: COMBINE, AFMoC, 4D-QSAR и ND-QSAR.</p>	1	2	4		
<p>Тема 9. Модели различных типов химических объектов.</p> <p>Индивидуальные и смеси химических соединений. Полимеры. Наноматериалы. Программы моделирования.</p>	1	2	4		
<p>Тема 10. Методы машинного обучения в хемоинформатике.</p> <p>Классификация. Особенности и рекомендации по применению методов.</p>	1	2	4		
<p>Раздел 4. Информатика химических реакций</p>					
<p>Тема 11. Математическое описание и кодирование химических реакций.</p> <p>Использование компьютеров для обработки информации о химических реакциях. Функциональный подход Хендриксона к систематизации синтетических реакций. Матричный формализм. Описание реакций с помощью графов.</p>	1	2	4		0

Тема 12. Анализ и прогнозирование характеристик реакций Задача автоматической классификации реакций в базах данных. Картография пространства химических реакций. Прогнозирование характеристик реакций.	1	2	4		
Тема 13. Компьютерные методы анализа химических реакций. Дизайн катализаторов. Методы планирования органического синтеза. Развитие информатики химических реакций.	1	2	4		
Раздел 5. Химическое пространство					
Тема 14. Основы химического пространства. Объекты и отношения сходства. Описательный анализ химического пространства. Библиотеки химических соединений. Фармакофорный анализ.	1	4	4		
Тема 15. Виртуальный скрининг. Воронка виртуального скрининга и ее компоненты. Базы данных для виртуального скрининга. Простейшие фильтры для виртуального скрининга. Ранжирование химических соединений в виртуальном скрининге с использованием 2D и 3D-сходства с активными структурами.	1	4	5,75		0
Итого в семестре (на курсе для ЗАО)	17	34	56,75		
Консультации и промежуточная аттестация (Зачет)	0,25				
Всего контактная работа и СР по дисциплине	51,25		56,75		

4 КУРСОВОЕ ПРОЕКТИРОВАНИЕ

Курсовое проектирование учебным планом не предусмотрено

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1 Описание показателей, критериев и системы оценивания результатов обучения

5.1.1 Показатели оценивания

Код компетенции	Показатели оценивания результатов обучения	Наименование оценочного средства
ПК-7	Излагает основные понятия, определения, методы и подходы, используемые в хемоинформатике; способы построения SAR/QSAR зависимостей. Создает собственные базы данных; строит простейшие	Вопросы устного собеседования Практико-ориентированные задания

	зависимости SAR/QSAR/QSPR, определяет их статистическую значимость. Демонстрирует навыки решения химических задач с использованием средств хемоинформатики.	
ПК-8	Имеет представления о способах представления химических данных, методах осуществления поиска в химических базах данных, а так же об основных химических баз данных и методах работы с ними. Обрабатывает данные; оперирует данными созданных баз, проводит поиск в базах данных; использует методы SAR/QSAR/QSPR в создании веществ с заданными свойствами. Демонстрирует навыки анализа химических баз данных для решения прикладных задач; Активно использует программы для создания и оперирования базами данных; демонстрирует навыки построения и использования SAR/QSAR/QSRP зависимостей.	Вопросы устного собеседования Практико-ориентированные задания

5.1.2 Система и критерии оценивания

Шкала оценивания	Критерии оценивания сформированности компетенций	
	Устное собеседование	Письменная работа
Зачтено	Обучающийся в полном объеме знает материал, грамотно и по существу излагает его, не допускает существенных неточностей в ответе на вопросы, способен правильно применить основные методы и инструменты при решении практических задач, владеет необходимыми навыками и приемами их выполнения.	
Не зачтено	Обучающийся не может изложить значительной части программного материала, допускает существенные ошибки, допускает неточности в формулировках и доказательствах, нарушения в последовательности изложения программного материала; неуверенно, с большими затруднениями выполняет практические задания.	

5.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

5.2.1 Перечень контрольных вопросов

№ п/п	Формулировки вопросов
Семестр 7	
1	Базовая молекулярная модель
2	Представление химических объектов
3	Хемоинформатика и хеометрика
4	Хемоинформатика и биоинформатика
5	Виды представлений
6	Линейные представления
7	Векторное представление
8	Матричное представление
9	Табличное представление
10	Структуры Маркуша
11	Координатные представления
12	Молекулярные поверхности и формы
13	Форматы данных
14	Конвертация между представлениями
15	Представление реакций
16	Структурный поиск в химических базах данных

17	Алгоритмы прохождения деревьев решений
18	Важнейшие химические базы данных
19	Классификация дескрипторов.
20	Предобработка данных
21	Разработка и валидация моделей "структура-свойство"
22	Методы SAR/QSAR/QSPR
23	Модели "структура-свойства" для разных типов химических соединений
24	Методы машинного обучения в хемоинформатике
25	Математическое описание и кодирование химических реакций
26	Анализ и визуализация реакционных данных
27	Прогнозирование характеристик реакций
28	Компьютерный дизайн катализаторов
29	Программное планирование органического синтеза
30	Химическое пространство
31	Описательный анализ химического пространства
32	Библиотеки химических соединений
33	Фармакофорный анализ
34	Виртуальный скрининг

5.2.2 Типовые тестовые задания

Не предусмотрено

5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)

Задание №1. Нарисовать молекулу и пронумеровать по алгоритму Моргана, записать в коде SMILES, SLN и WLN:

```
-ISIS- 10130513072D
9 8 0 0 0 0 0 0 0 0999 V2000
2.7490 1.0140 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0
2.7473 1.9020 0.0000 C 0 0 3 0 0 0 0 0 0 0 0 0
3.4773 2.2688 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0
3.4773 3.0797 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0
2.0216 2.2735 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0
1.3430 1.8280 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0
0.5852 2.1933 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0
1.3759 1.0411 0.0000 N 0 0 0 0 0 0 0 0 0 0 0 0
4.1795 1.8357 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0
2 1 1 0 0 0 0
2 3 1 0 0 0 0
3 4 2 0 0 0 0
2 5 1 0 0 0 0
6 7 2 0 0 0 0
6 8 1 0 0 0 0
5 6 1 0 0 0 0
3 9 1 0 0 0 0
M END
```

Задание №2. Пронумеровать молекулу по алгоритму Моргана и записать в коде SMILES, SLN и WLN:
-S-NH2-OH=O

Задание №3. Нарисовать молекулу и пронумеровать по алгоритму Моргана, записать в коде SMILES, SLN и WLN:

```
N 0.000000 0 0.000000 0 0.000000 0 0 0 0
C 1.469900 1 0.000000 0 0.000000 0 1 0 0
C 1.519900 1 109.470900 1 0.000000 0 2 1 0
O 1.220000 1 119.999900 1 179.999917 1 3 2 1
C 1.540000 1 109.471000 1 240.000500 1 2 1 3
C 1.519900 1 109.471000 1 180.000003 1 5 2 1
O 1.219900 1 120.000000 1 179.999953 1 6 5 2
N 1.320000 1 119.999900 1 359.999953 1 6 5 2
O 1.499900 1 120.000000 1 359.999917 1 3 2 1
H 0.990000 1 119.999900 1 59.999700 1 1 2 3
H 1.010000 1 109.470993 1 292.239097 1 1 2 3
H 1.090000 1 109.471400 1 180.000000 1 2 1 10
H 1.090000 1 109.471000 1 299.999403 1 5 2 1
H 1.089900 1 109.471400 1 59.999703 1 5 2 1
H 1.010000 1 120.000000 1 180.000000 1 8 6 5
H 1.010000 1 120.000000 1 359.999900 1 8 6 5
H 0.960000 1 109.471000 1 180.000000 1 9 3 2
```

5.3 Методические материалы, определяющие процедуры оценивания знаний, умений, владений (навыков и (или) практического опыта деятельности)

5.3.1 Условия допуска обучающегося к промежуточной аттестации и порядок ликвидации академической задолженности

Проведение промежуточной аттестации регламентировано локальным нормативным актом СПбГУПТД «Положение о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся»

5.3.2 Форма проведения промежуточной аттестации по дисциплине

Устная

+

Письменная

Компьютерное тестирование

Иная

5.3.3 Особенности проведения промежуточной аттестации по дисциплине

- Возможность пользоваться справочными таблицами, калькулятором;
- Время на подготовку ответа по билету 15 минут;
- Зачет проводится в компьютерном классе.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Учебная литература

Автор	Заглавие	Издательство	Год издания	Ссылка
6.1.1 Основная учебная литература				
Володченкова, Л. А.	Биоинформатика	Омск: Издательство Омского государственного университета	2018	http://www.iprbookshop.ru/108109.html
Кравцов, А. В., Чеканцев, Н. В., Шарова, Е. С., Гынгазова, М. С., Смышляева, Ю. А., Иванчина, Э. Д.	Проблемно ориентированная информатика химико-технологических процессов	Томск: Томский политехнический университет	2013	http://www.iprbookshop.ru/34700.html
6.1.2 Дополнительная учебная литература				
Т. Л. Луканина, И. С. Михайлова, Л. П. Ардашева, А. Ю. Вахрушев	Химия. (Основы химии для самостоятельного изучения): учеб. пособие	М-во науки и высшего образования РФ, С.-Петербург. гос. ун-т пром. технологий и дизайна, Высш. шк. технологии и энергетики. - Санкт-Петербург : ВШТЭ СПбГУПТД	2020	http://nizrp.narod.ru/metod/kafobshineorgh/1600205290.pdf
Радыгин, В. Ю., Куприянов, Д. Ю.	Базы данных: основы, проектирование, разработка информационных систем, проекты	Москва: Национальный исследовательский ядерный университет «МИФИ»	2020	https://www.iprbookshop.ru/116387.html

6.2 Перечень профессиональных баз данных и информационно-справочных систем

1. Электронно-библиотечная система IPRbooks [Электронный ресурс]. URL: <http://www.iprbookshop.ru/>
2. Электронная библиотека ВШТЭ СПб ГУПТД [Электронный ресурс]. URL: <http://nizrp.narod.ru/>
3. справочная правовая система КонсультантПлюс [Электронный ресурс]. URL: <http://www.consultant.ru/>
4. Электронная библиотека eLIBRARY.RU [Электронный ресурс]. URL: <https://elibrary.ru/>
5. Национальный Открытый университет ИНТУИТ [Электронный ресурс]. URL: <https://www.intuit.ru/>
6. Информационная система «Единое окно доступа к образовательным ресурсам. Раздел. Информатика и информационные технологии» [Электронный ресурс]. URL: http://window.edu.ru/catalog/?p_rubr=2.2.75.6
7. Федеральное государственное автономное учреждение «Государственный научно-исследовательский институт информационных технологий и телекоммуникаций», «Информика» [Электронный ресурс]. URL: <http://www.informika.ru/>

6.3 Перечень лицензионного и свободно распространяемого программного обеспечения

MicrosoftWindows 8
MicrosoftOfficeProfessional 2013
PTC Mathcad 15

6.4 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Аудитория	Оснащение
-----------	-----------

Компьютерный класс	Мультимедийное оборудование, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду
Лекционная аудитория	Мультимедийное оборудование, специализированная мебель, доска