

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе
и образовательным инновациям

 О.И. Тирошенко

«05» июля 2024 г.

Регистрационный № УД -1474/м.

АНАЛИЗ И ОБРАБОТКА ДАННЫХ В RUTNOM

Учебная программа учреждения образования
по учебной дисциплине для специальности

7-06-0531-01 Химия

профилизация специальности:

Хемоинформатика

2024 г.

Учебная программа составлена на основе образовательного стандарта ОСВО 7-06-0531-01-2023 и учебного плана БГУ №М44с-5.5-149/уч. от 29.03.2024.

СОСТАВИТЕЛИ:

С.В. ШОЛТАНЮК, старший преподаватель кафедры компьютерных технологий и систем факультета прикладной математики и информатики Белорусского государственного университета.

РЕЦЕНЗЕНТЫ:

Н.В. МАЛАЩЁНОК, руководитель проектов разработки ООО «АСД Лабс», кандидат химических наук, доцент.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой компьютерных технологий и систем БГУ (протокол № 15 от 25.06.2024);

Научно-методическим советом БГУ (протокол № 9 от 28.06.2024)

Заведующий кафедрой _____



В.В. Казачёнок

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Цель и задачи учебной дисциплины

Настоящая дисциплина знакомит магистрантов с основами программирования на языке Python, широко используемом в хемоинформатике, в т.ч. при работе с различными данными, в первую очередь о строении и свойствах неорганических и органических химических соединений, а также с основами анализа данных и некоторыми методами, применяющимися при анализе химических данных. В рамках настоящей дисциплины предлагается изучение как основ самого языка, так и базовые примеры его использования в химии и хемоинформатике, например, моделирование и визуализация химических соединений, а также работа с химическими базами данных.

Цель учебной дисциплины «Анализ и обработка данных в Python» – ознакомление магистрантов с языком программирования Python, формированием у них умений и навыков разработки на языке Python, в частности для решения базовых задач химии и хемоинформатики.

Задачи учебной дисциплины:

1. Ознакомить студентов с основными особенностями языка Python, его областями применения.
2. Изучить основные приёмы работы с данными, базами данных (в том числе химическими).
3. Приобрести знания, умения и навыки разработки программного кода и приложений на языке Python.

Место учебной дисциплины. В системе подготовки специалиста с углубленным высшим образованием учебная дисциплина относится к модулю «Специальные программные средства» компонента учреждения образования.

Учебная программа составлена с учетом межпредметных **связей** и программ по следующим дисциплинам:

- «Компьютерное моделирование строения и реакционной способности молекул»;
- «Методы визуализации в анализе данных средствами языка R»;
- «Основы молекулярной биологии»;
- дисциплины модуля «Медицинская химия»: «Хеометрика» и «Современные методы прогнозирования свойств веществ».

Требования к компетенциям

Освоение учебной дисциплины «Анализ и обработка данных в Python» должно обеспечить формирование следующей *специализированной* компетенции:

СК. Строить, анализировать и тестировать алгоритмы и программы решения типовых задач обработки информации с использованием структурного, объектно-ориентированного и иных парадигм программирования. Применять полученные навыки в рамках решения химических задач.

В результате освоения учебной дисциплины магистрант должен:

знать:

- базовые принципы организации языка программирования Python;
- основные принципы представления данных, работы с данными и базами данных в Python;
- основные принципы организации вычислений в языке Python;
- основные области применения языка Python (в том числе в химии и хемоинформатике);

уметь:

- устанавливать и эффективно управлять средствами разработки на языке Python;
- разрабатывать и отлаживать программные коды на языке Python;
- разрабатывать пользовательские приложения на языке Python;

владеть:

- базовым набором инструментов для разработки приложений на языке Python.

Структура учебной дисциплины

Дисциплина изучается в 1 и во 2 семестре. Количество часов, отведённых на изучение учебной дисциплины «Анализ и обработка данных в Python», их распределение по разным видам занятий, а также трудоёмкость дисциплины и формы промежуточной аттестации представлены в таблице ниже.

Специальность (профилизация)		7-06-0531-01 Химия (Хемоинформатика)	
Форма получения углубленного высшего образования		Очная форма	
Курс		1	
Семестр		I	II
Всего часов по дисциплине	180	90	90
Всего аудиторных часов по дисциплине	96	56	40
Лекции	48	28	20
Лабораторные занятия	44	26 (ДОТ)	18
Управляемая самостоятельная работа	4	2 (ДОТ)	2
Трудоёмкость учебной дисциплины	6	3 з.е.	3 з.е.
Форма промежуточной аттестации		зачёт	экзамен

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Раздел 1. Основы Python

Тема 1.1. Основы языка программирования Python

Общие сведения о языке программирования Python. Основные возможности и преимущества Python. Сравнение языков Python и R. Установка языка Python. Среды разработки и дистрибутивы для Python. IPython – интерактивная вычислительная среда для Python. Версии языка Python (Python 2 и Python 3), их сравнение.

Тема 1.2. Базовые операции Python

Синтаксис Python. Базовые типы данных в Python. Консольный ввод и вывод, функции `input` и `print`. Целые и вещественные числа. Условный оператор `if`. Операторы циклов. Ключевые слова и встроенные идентификаторы Python. Основы работы со строками. Функции и методы строк. Форматирование строк, метод `format`. Регулярные выражения, модуль `re`.

Тема 1.3. Стандартные структуры данных Python

Списки (`list`). Функции и методы списков. Массивы. Индексы и срезы. Кортежи (`tuple`). Операции с кортежами, упаковка и распаковка переменных. Словари (`dict`) и работа с ними. Методы словарей. Множества (`set` и `frozenset`). Операции с множествами. Генераторы списков и словарей. Изменяемые (`mutable`) и неизменяемые (`immutable`) объекты в Python.

Тема 1.4. Работа с файлами в Python

Чтение из файла и запись в файл. Оператор `with`. Файловые форматы CSV и JSON. Работа с файловой системой с использованием Python, модуль `os`.

Тема 1.5. Модульное программирование и объектно-ориентированное программирование (ООП)

Определение пользовательских функций в Python. Аргументы функции: позиционные и именованные аргументы, аргументы со значением по умолчанию. Упаковка и распаковка аргументов функции. Генераторы в Python, ключевое слово `yield`. Основы ООП. Инкапсуляция, наследование, полиморфизм. Классы в Python. Обработка исключений. Иерархия исключений. Модули и пакеты. Основные пакеты, используемые в Python. Комментарии и документация. Области видимости функций, модулей и пакетов.

Раздел 2. Анализ и обработка данных в Python

Тема 2.1. Векторные и матричные операции, научные вычисления в Python

Модуль `numpy`. Массивы (`array`), их сходства со списками `list` и отличия от них. Базовые операции с массивами, их сходство с операциями над векторами и матрицами в языке R. Понятие оси массива (`axis`).

Тема 2.2. Работа с данными в Python

Модуль `pandas`. Работа с рядами (`Series`) и таблицами данных (`DataFrame`) при помощи модуля `pandas`. Сортировка, фильтрация и агрегация данных в таблице. Сходство рядов и таблиц данных с аналогичными объектами языка R. Их сходство со стандартными типами данных Python. Их сходство с объектами `NumPy`.

Тема 2.3. Графические возможности Python

Модуль `matplotlib`. Особенности работы с модулем `matplotlib` в интерактивной среде `Jupyter Notebook` и других средах разработки. Графический интерфейс `pyplot`. Построение диаграмм и графиков при помощи `pyplot`. Оформление диаграмм и графиков (легенды, текст и т.д.). Графическое представление табличных данных.

Тема 2.4. Анализ данных в хемоинформатике

Этапы анализа данных. Подготовка данных (обработка пропущенных и аномальных значений, фильтрация и нормализация). Основы анализа данных. Корреляционный и регрессионный анализ. Методы классификации и кластеризации. Методы снижения размерности PCA (метод главных компонент) и PLS (метод частичных наименьших квадратов). Модули `scipy` и `scikit-learn`.

Раздел 3. Основы химического моделирования в Python

Тема 3.1. Моделирование химических веществ в Python

Моделирование атомов, ионов, молекул и прочих способов представления химических веществ. Визуализация молекул в неорганической и органической химии. SMILES – спецификация описания молекул химических веществ. Моделирование химических реакций.

Тема 3.2. Работа с моделями веществ в Python

Поиск подструктур в молекулах веществ. Работа со свойствами химических веществ. Определение молекулярного подобия веществ.

Тема 3.3. Пакеты Python для специальных приложений хемоинформатики

Предсказание свойств химических веществ. Пакеты Python для квантовой химии. Пакеты для обработки и анализа спектральных данных. Основные химические базы данных и доступ к ним посредством специальных пакетов Python.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Очная форма получения высшего образования
с применением дистанционных образовательных технологий (ДОТ)

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСП	Формы контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	Основы Python							
1.1	Основы языка программирования Python	2			2 (ДОТ)			Устный опрос
1.2	Базовые операции Python	6			6 (ДОТ)			Устный опрос
1.3	Стандартные структуры данных Python	4			4 (ДОТ)			Проверка программного кода лабораторных работ, защита отчета по лабораторной работе
1.4	Работа с файлами в Python	2			2 (ДОТ)			Устный опрос
1.5	Модульное программирование и объектно-ориентированное программирование (ООП)	2			2 (ДОТ)			Устный опрос
2	Анализ и обработка данных в Python							
2.1	Векторные и матричные операции, научные вычисления в Python	4			4 (ДОТ)			Устный опрос

2.2	Работа с данными в Python	4			4 (ДОТ)			Проверка программного кода лабораторных работ, защита отчета по лабораторной работе
2.3	Графические возможности Python	4			2 (ДОТ)		2 (ДОТ)	Проверка программного кода лабораторных работ, защита отчета по лабораторной работе
2.4	Анализ данных в хемоинформатике	8			8			Проверка программного кода лабораторных работ, защита отчета по лабораторной работе
3	Основы химического моделирования в Python							
3.1	Моделирование химических веществ в Python	6			6			Проверка программного кода лабораторных работ, защита отчета по лабораторной работе
3.2	Работа с моделями веществ в Python	4			2		2	Открытое задание
3.3	Пакеты Python для специальных приложений хемоинформатики	2			2			Устный опрос
ИТОГО		48			44		4	зачёт, экзамен

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Бизли, Д. Python. Исчерпывающее руководство = Python Distilled / Д. Бизли ; [пер. с англ. Е. Матвеев]. – Санкт-Петербург ; Москва ; Минск : Питер, 2023. – 366 с.
2. Косицын, Д.Ю. Язык программирования Python : учеб.-метод. пособие [Электронный ресурс] / Д. Ю. Косицин. – Минск : БГУ, 2019. – 136 с. – URL : <http://elib.bsu.by/handle/123456789/237193>.
3. Мэттиз, Э. Изучаем Python = Python Crash Course : программирование игр, визуализация данных, веб-приложения / Э. Мэттиз ; [перевел с англ. Е. Матвеев]. – 3-е изд. – Санкт-Петербург [и др.] : Питер, 2020. – 511 с.

Перечень дополнительной литературы

1. Hill, C. Python for Chemists / C. Hill. – Cambridge, UK : Cambridge University Press, 2023. – 560 p.
2. Interactive Data Visualization with Python / A. Belorkar [et al.]. – 2nd ed. – Birmingham, UK : Packt Publishing Ltd, 2020. – 362 p.
3. Kanagasabapathy, M. Python for Chemistry : An introduction to Python algorithms, Simulations, and Programing for Chemistry / M. Kanagasabapathy. – London : BPB Publications, 2023. – 368 p.
4. Маккини, У. Python и анализ данных : Первичная обработка данных с применением pandas, NumPy и Jupyter / У. Маккини – 3-е изд.: пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2023. – 536 с.: ил.

Интернет-ресурсы

1. 3.12.5 Documentation [Electronic resource]. – Mode of access: <https://docs.python.org/3/>. – Date of access: 15.05.2024.
2. The Python Tutorial — Python 3.12.5 documentation [Electronic resource]. – Mode of access: <https://docs.python.org/3/tutorial/index.html>. – Date of access: 15.05.2024.
3. Cheminformatics - Chemistry LibreTexts [Electronic resource]. – Mode of access: https://chem.libretexts.org/Courses/Intercollegiate_Courses/Cheminformatics. – Date of access: 15.05.2024.
Fall 2019 Cheminformatics OLCC, a cheminformatics course sponsored by the ACS Division of Chemical Education's [Committee on Computers in Chemical Education](#).

4. Matthes, E. Python Crash Course - Third Edition [Electronic resource] / E. Matthes // a Github repository. – Mode of access: https://ehmatthes.github.io/pcc_3e/. – Date of access: 15.05.2024.
Online resources for *Python Crash Course*, 3rd edition, from No Starch Press.
5. McKinney, W. Python for Data Analysis, 3E [Electronic resource]. / W. McKinney. – Mode of access: <https://wesmckinney.com/book/>. – Date of access: 15.05.2024.
An open access HTML version of the 3rd edition of *Python for Data Analysis*
6. Mentel, L. Awesome Python Chemistry / Lukasz Mentel // a GitHub repository. – Mode of access: <https://github.com/lmmentel/awesome-python-chemistry>. – Date of access: 15.05.2024.
A curated list of awesome Python frameworks, libraries, software and resources related to Chemistry. Inspired by <https://awesome-python.com/>.
7. Самоучитель Python | Python 3 для начинающих и чайников [Электронный ресурс]. – Режим доступа: <https://pythonworld.ru/samouchitel-python>. – Дата доступа: 15.05.2024.
8. Matplotlib documentation — Matplotlib 3.9.2 documentation [Electronic resource]. – Mode of access: <https://matplotlib.org/stable/index.html>. – Date of access: 15.05.2024.
9. NumPy Documentation [Electronic resource]. – Mode of access: <https://numpy.org/doc/>. – Date of access: 15.05.2024.
10. Numpy and Scipy Documentation — Numpy and Scipy documentation [Electronic resource]. – Mode of access: <https://docs.scipy.org/doc/>. – Date of access: 15.05.2024.
11. Pandas documentation — pandas 2.2.2 documentation [Electronic resource]. – Mode of access: <https://pandas.pydata.org/docs/index.html>. – Date of access: 15.05.2024.
12. PubChemPy documentation — PubChemPy 1.0.4 documentation [Electronic resource]. – Mode of access: <http://pubchempy.readthedocs.io/>. – Date of access: 15.05.2024.
13. Getting Started with the RDKit in Python — The RDKit 2024.03.5 documentation [Electronic resource]. – Mode of access: <https://www.rdkit.org/docs/GettingStartedInPython.html>. – Date of access: 15.05.2024.
RDKit is a collection of cheminformatics and machine-learning software written in C++ and Python.

Перечень рекомендуемых средств диагностики и методика формирования итоговой отметки

Объектом диагностики компетенций магистрантов являются знания, умения, полученные ими в результате изучения учебной дисциплины. Выявление учебных достижений магистрантов осуществляется с помощью мероприятий текущей и промежуточной аттестации.

Текущая аттестация магистрантов проходит в следующих формах:

- устная: устные опросы, проводимые в целях первичного мониторинга усвоения материала магистрантами и оцениваемые исходя из полноты и последовательности ответа, понимания основных понятий, методов и алгоритмов, изложенных на лекционных или лабораторных занятиях.

- техническая: лабораторные работы, выполняемые на компьютере. Они оцениваются исходя из читаемости и оптимизированности программного кода, а также путём тестирования программного кода при работе на различных примерах.

- устно-письменная: устная и/или письменная (в виде отчёта) защита лабораторных работ, оцениваемая на основе полноты и последовательности ответа (отчёта), полноты раскрытия содержания выполненного задания, понимания работы алгоритмов и методов, использованных при выполнении задания; а также открытое (эвристическое) задание, направленное на закрепление и применение полученных знаний к практическим задачам химии и хемоинформатики.

Формами промежуточной аттестации по дисциплине «Анализ и обработка данных в Python» предусмотрены **зачёт** в 1 семестре и **экзамен** во 2 семестре.

При формировании итоговой отметки используется рейтинговая система оценки знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая система предусматривает использование весовых коэффициентов в ходе проведения контрольных мероприятий текущей аттестации.

Примерные весовые коэффициенты, определяющие вклад текущей аттестации в отметку при прохождении промежуточной аттестации:

Формирование отметки за текущую аттестацию:

- устные опросы – 10%,
- проверка программного кода лабораторных работ – 50%,
- защита отчётов по лабораторным работам – 20%.
- выполнение открытого (эвристического) задания – 20%.

Итоговая отметка по дисциплине рассчитывается на основе отметки текущей аттестации (рейтинговой системы оценки знаний) – 40% и экзаменационной отметки – 60%.

Зачёт по дисциплине проходит в устной и/или письменной форме в виде защиты отчётов по лабораторным работам 1 семестра. В случае успешной защиты отчётов по всем лабораторным работам допускается получение зачёта без проведения дополнительного опроса. При этом явка обучающегося на зачёт является обязательной.

В случае успешной защиты эвристического задания допускается определение результатов промежуточной аттестации во 2 семестре без проведения дополнительного опроса на экзамене. При этом явка обучающегося на экзамен является обязательной.

Примерный перечень заданий для управляемой самостоятельной работы студентов

Тема 2.3. Графические возможности Python (2 ч. ДОТ)

Примерный перечень вопросов и заданий:

1. Построение диаграмм и графиков при помощи пакета `matplotlib`. Оформление диаграмм и графиков (легенды, текст и т.д.). Графическое представление табличных данных.

2. На основе данных, загруженных из файла или пакета (например `sklearn` или `statsmodels`), построить диаграмму рассеяния и/или график динамики этих данных).

Форма контроля – проверка программного кода лабораторных работ, защита отчета по лабораторной работе.

Тема 3.2. Работа с моделями веществ в Python (2 ч.)

Примерный перечень вопросов и заданий:

1. Поиск подструктур в молекулах веществ. Работа со свойствами химических веществ. Определение молекулярного подобия веществ.

2. Выбрать объект для исследования в рамках задания и найти соответствующие данные (например, согласно теме магистерской диссертации или дипломной работы).

3. Осуществить предобработку данных, представить их в подходящем формате.

4. Осуществить анализ данных, их динамику и взаимосвязь между ними, используя различные пакеты и модули языка Python.

Форма контроля – открытое задание.

Примерная тематика лабораторных занятий

Занятие № 1. Установка языка Python. Среды разработки и дистрибутивы для Python. IPython – интерактивная вычислительная среда для Python.

Занятие № 2. Синтаксис Python. Базовые типы данных в Python. Консольный ввод и вывод, функции `input` и `print`. Целые и вещественные числа.

Занятие № 3. Условный оператор `if`. Операторы циклов. Ключевые слова и встроенные идентификаторы Python.

Занятие № 4. Основы работы со строками. Функции и методы строк. Форматирование строк, метод `format`. Регулярные выражения, модуль `re`.

Занятия №5-6. Списки (`list`). Функции и методы списков. Массивы. Индексы и срезы. Кортежи (`tuple`). Операции с кортежами, упаковка и распаковка переменных. Словари (`dict`) и работа с ними. Методы словарей. Множества (`set` и `frozenset`). Операции с множествами. Генераторы списков и словарей.

Занятие №7. Чтение из файла и запись в файл. Оператор `with`. Файловые форматы CSV и JSON. Работа с файловой системой с использованием Python, модуль `os`.

Занятие №8. Основы ООП. Инкапсуляция, наследование, полиморфизм. Классы в Python. Обработка исключений. Иерархия исключений.

Занятия №9-10. Модуль `numpy`. Массивы (`array`), их сходства со списками `list` и отличия от них. Базовые операции с массивами, их сходство с операциями над векторами и матрицами в языке R. Понятие оси массива (`axis`).

Занятия №11-12. Модуль `pandas`. Работа с рядами (`Series`) и таблицами данных (`DataFrame`) при помощи модуля `pandas`. Сортировка, фильтрация и агрегация данных в таблице. Сходство рядов и таблиц данных с аналогичными объектами языка R. Их сходство со стандартными типами данных Python. Их сходство с объектами NumPy.

Занятие №13. Модуль `matplotlib`. Особенности работы с модулем `matplotlib` в интерактивной среде Jupyter Notebook и других средах разработки. Графический интерфейс `pyplot`.

Занятия №14-17. Этапы анализа данных. Подготовка данных (обработка пропущенных и аномальных значений, фильтрация и нормализация). Основы

анализа данных. Корреляционный и регрессионный анализ. Методы классификации и кластеризации. Методы снижения размерности PCA (метод главных компонент) и PLS (метод частичных наименьших квадратов). Модули `scipy` и `scikit-learn`.

Занятия №18-20. Моделирование атомов, ионов, молекул и прочих способов представления химических веществ. Визуализация молекул в неорганической и органической химии. SMILES – спецификация описания молекул химических веществ. Моделирование химических реакций.

Занятие №21. Поиск подструктур в молекулах веществ. Работа со свойствами химических веществ. Определение молекулярного подобия веществ.

Занятие №22. Предсказание свойств химических веществ. Пакеты Python для квантовой химии. Пакеты для обработки и анализа спектральных данных. Основные химические базы данных и доступ к ним посредством специальных пакетов Python.

Описание инновационных подходов и методов к преподаванию учебной дисциплины

При организации образовательного процесса используются *эвристический* и *практико-ориентированный подходы*.

Эвристический подход предполагает:

- осуществление студентами лично-значимых открытий окружающего мира;
- демонстрацию многообразия решений большинства профессиональных задач и жизненных проблем;
- творческую самореализацию обучающихся в процессе создания образовательных продуктов;
- индивидуализацию обучения через возможность самостоятельно ставить цели, осуществлять рефлексию собственной образовательной деятельности.

Практико-ориентированный подход предполагает:

- освоение содержания образования через решения практических задач;
- приобретение навыков эффективного выполнения разных видов профессиональной деятельности;
- ориентацию на генерирование идей, реализацию групповых студенческих проектов, развитие предпринимательской культуры;
- использованию процедур, способов оценивания, фиксирующих сформированность профессиональных компетенций.

Методические рекомендации по организации самостоятельной работы обучающихся

Самостоятельная работа с целью изучения материала учебной дисциплины предполагает работу с рекомендованной учебной литературой и Интернет-ресурсами. Теоретические сведения закрепляются выполнением лабораторных заданий, при выполнении которых следует руководствоваться методическими разработками, размещенными в электронной библиотеке университета и на образовательном портале. Кроме того, с учётом учебной и научной деятельности магистрантов формируется открытое (эвристическое) задание, направленное на закрепление и применение полученных знаний к практическим задачам химии и хемоинформатики (например, поставленных в рамках научно-исследовательской работы магистрантов при написании магистерских диссертаций). Могут быть предложены дополнительные задания (тесты, задания для самостоятельного выполнения) для самооценки и более глубокого усвоения полученного материала.

Примерный перечень вопросов к экзамену

1. Общие сведения о языке программирования Python. Основные возможности и преимущества Python.
2. Различия между версиями Python 2 и Python 3.
3. Основные особенности синтаксиса Python.
4. Консольный ввод и вывод, функции `input` и `print`.
5. Основные типы данных Python.
6. Условный оператор `if`.
7. Операторы циклов.
8. Работа со строками. Функции и методы строк.
9. Форматирование строк, метод `format`.
10. Регулярные выражения. Модуль `re`.
11. Стандартные структуры данных Python. Списки (`list`). Функции и методы списков.
12. Стандартные структуры данных Python. Кортежи (`tuple`). Операции с кортежами, упаковка и распаковка переменных.
13. Стандартные структуры данных Python. Словари (`dict`). Методы словарей.
14. Стандартные структуры данных Python. Множества (`set` и `frozenset`).
15. Стандартные структуры данных Python. Генераторы списков и словарей.
16. Работа с файлами (текстовые форматы, CSV, JSON).
17. Определение пользовательских функций в Python.

18. Основы объектно-ориентированного программирования. Классы в Python.
19. Обработка исключений. Иерархия исключений.
20. Модули и пакеты. Основные пакеты, используемые в Python. Комментарии и документация. Области видимости функций, модулей и пакетов.
21. Модуль `numpy`. Работа с массивами (`array`). Сходства массивов со списками `list` и отличия от них. Базовые операции с массивами.
22. Обработка данных в Python. Модуль `pandas`. Основные операции с рядами (`Series`).
23. Обработка данных в Python. Модуль `pandas`. Основные операции с таблицами данных (`DataFrame`).
24. Графика в Python. Построение диаграмм и графиков при помощи пакета `matplotlib`. Оформление диаграмм и графиков (легенды, текст и т.д.).
25. Графика в Python. Графическое представление табличных данных.
26. Основы анализа данных. Этапы анализа данных. Подготовка данных.
27. Корреляционный анализ с использованием Python.
28. Регрессионный анализ с использованием Python.
29. Методы классификации и кластеризации с использованием Python.
30. Методы снижения размерности с использованием Python.
31. Основные пакеты Python для моделирования химического строения веществ.
32. Моделирование химических реакций в Python.
33. Работа с моделями веществ в Python.
34. Пакеты Python для специальных приложений хемоинформатики.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
1. Компьютерное моделирование строения и реакционной способности молекул	Кафедра неорганической химии	Изменений не требуется	
2. Методы визуализации в анализе данных средствами языка R	Кафедра теории вероятностей и математической статистики	Изменений не требуется	
3. Современная неорганическая химия	Кафедра неорганической химии	Изменений не требуется	
4. Современная органическая химия	Кафедра органической химии	Изменений не требуется	
5. Современные методы прогнозирования свойств веществ	Кафедра неорганической химии	Изменений не требуется	
6. Хемометрика	Кафедра физической химии и электрохимии	Изменений не требуется	

Заведующий кафедрой неорганической химии

Доктор химических наук, профессор,
член-корреспондент НАН Беларуси

«21» июня 2024 г

Д.В. Свиридов

Заведующий кафедрой теории вероятностей и математической статистики

Доктор физико-математических наук, профессор

«25» июня 2024 г

А.Ю. Харин

Заведующий кафедрой органической химии

Кандидат химических наук, доцент

«20» июня 2024 г

И.П. Антоневиц

Заведующий кафедрой физической химии и электрохимии

Доктор химических наук, профессор

«24» июня 2024 г

Е.А. Стрельцов

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ
ПО ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 202_ г.)

Заведующий кафедрой

УТВЕРЖДАЮ
Декан факультета
