

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ИНГУШСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ХИМИКО-БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ**

СОГЛАСОВАНО

Руководитель образовательной программы

_____ профессор Саламов А.М.

« 22 » _____ апреля _____ 2026 г.

УТВЕРЖДАЮ

И.о. декана химико-биологического факультета

_____ Б.А.Кодзоева

« 29 » _____ апреля _____ 2026 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

**Б1.О.25 «Искусственный интеллект в хемоинформатике: от молекулярных
репрезентаций до генерации новых соединений»**

Направление подготовки

04.03.01. Химия

Направленность (профиль)

«Аналитическая химия и химическая экспертиза»

Квалификация выпускника

Бакалавр

Форма обучения

Очная

Магас, 2026

1. ЦЕЛЬ И ЗАДАЧИ ДИСЦИПЛИНЫ

Цель: сформировать у студентов системное понимание принципов применения технологий искусственного интеллекта (ИИ) для решения задач хемоинформатики, молекулярного моделирования и молекулярного дизайна; обеспечить освоение инструментальных средств ИИ для прогнозирования свойств веществ, генерации новых соединений и анализа белок-лигандных взаимодействий.

Задачи:

- Изучить принципы представления химических структур в виде данных (SMILES, SELFIES, молекулярные графы, дескрипторы), основы хемоинформатики и нанохимтехнологий.
- Освоить методы машинного обучения для предсказания свойств химических соединений (QSAR/QSPR), построения и валидации моделей, тюнинга гиперпараметров.
- Освоить применение генеративных моделей (VAE, GAN), методов молекулярной фильтрации (ADME, PAINS) и виртуального скрининга для молекулярного дизайна.
- Изучить графовые нейронные сети (GCNN, MPNN, SchNet, ALIGNN) и их применение для предсказания свойств молекул и анализа белок-лигандных взаимодействий.
- Сформировать навыки работы с молекулярными базами (ChEMBL, PubChem, QM9), PDB и инструментами AlphaFold.
- Развить навыки программирования на Python (RDKit, Pandas, Scikit-Learn, PyTorch, PyTorch Geometric, Biopython, PyMOL) для решения задач хемоинформатики.
- Развить навыки анализа этических и социальных рисков применения ИИ в химической промышленности.

2. СВЯЗЬ ДИСЦИПЛИНЫ С ПРЕДШЕСТВУЮЩИМИ ДИСЦИПЛИНАМИ И СРОКИ ИХ ИЗУЧЕНИЯ

Таблица 2.1.

Код дисциплины	Дисциплины, предшествующие дисциплине «ИИ в хемоинформатике»	Семестр
Б1.О.10	Общая и неорганическая химия	1–2
Б1.О.12	Органическая химия	2–3
Б1.О.14	Физическая химия	3–4
Б1.О.17	Информатика и основы программирования	2
Б1.О.19	Математический анализ и статистика	2–3
Б1.О.22	Строение вещества и квантовая химия	4

Для успешного освоения дисциплины обучающимся необходимо обладать: базовыми знаниями в области химии, химической технологии и принципов работы химических производств; базовыми знаниями в области цифровых технологий и информатики, включая

навыки программирования на уровне написания скриптов; пониманием основ машинного обучения; способностью к аналитическому мышлению и самостоятельному обучению.

3. РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО по данному направлению:

Таблица 3.1.

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.1. Понимает принципы современной химической технологии, основы нанохимтехнологий, основы молекулярного моделирования.	Знать: Принципы современной химической технологии; основы нанохимтехнологий; принципы молекулярного моделирования; фундаментальные принципы применения технологий ИИ для решения задач химической промышленности и хемоинформатики; теоретические основы применения методов ИИ для прогнозирования свойств веществ, молекулярного дизайна и анализа данных в химической технологии. Уметь: Применять принципы молекулярного моделирования для решения задач хемоинформатики; использовать методы ИИ для прогнозирования свойств веществ и молекулярного дизайна; выбирать и применять библиотеки Python (RDKit, Pandas) для загрузки молекул, визуализации и вычисления дескрипторов; применять методы машинного обучения для предсказания свойств химических соединений. Владеть: Навыками применения технологий ИИ в хемоинформатике; методами молекулярного моделирования; инструментами Python для обработки научных данных и машинного обучения; навыками

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
			работы с молекулярными базами (ChEMBL, PubChem, QM9).
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.2. Оценивает и прогнозирует перспективность объектов своей профессиональной деятельности для химико-технологических производств.	Знать: Методы оценки и прогнозирования перспективности объектов для химико-технологических производств; принципы построения моделей QSAR/QSPR; связь между химической структурой и макроскопическими свойствами; типы моделей (линейные, нелинейные, ансамблевые); методы тюнинга гиперпараметров; принципы работы генеративных моделей (VAE, GAN); принципы и цели молекулярной фильтрации (ADME, PAINS). Уметь: Оценивать и прогнозировать перспективность химических соединений для химико-технологических производств; прогнозировать свойства веществ на основе их химической структуры; применять методы предобработки данных для ML; применять алгоритмы кластеризации и понижения размерности; применять генеративные модели для создания новых соединений с заданными свойствами; использовать фильтры ADME и PAINS. Владеть: Навыками оценки и прогнозирования перспективности объектов; методами построения и валидации моделей QSAR/QSPR; навыками применения градиентного бустинга (XGBoost, LightGBM); навыками применения генеративных моделей (VAE, GAN) для молекулярного дизайна;

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
			методами виртуального скрининга и высокопроизводительного скрининга.
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-5.3. Знает основные тенденции развития современных информационных технологий, основы информационной безопасности; методы применения информации из различных источников для решения профессиональных задач.	Знать: Основные тенденции развития современных информационных технологий; основы информационной безопасности; методы применения информации из различных источников для решения профессиональных задач; принцип работы графовых нейронных сетей (GNN); архитектуры GCNN, MPNN, SchNet, ALIGNN; принципы Ligand-based и Structure-based drug design; основы работы с белковыми структурами (PDB); принципы работы AlphaFold. Уметь: Применять информацию из различных источников для решения профессиональных задач; использовать готовую реализацию GNN (PyTorch Geometric) для предсказания свойств молекул; сравнивать производительность GNN с классическими ML-моделями; использовать Python-библиотеки (Biopython, PyMOL) для загрузки и визуализации белковых структур; применять AlphaFold; анализировать белок-лигандные взаимодействия. Владеть: Навыками работы с современными информационными технологиями; методами обеспечения информационной безопасности; навыками применения GNN для задач хемоинформатики; навыками работы с PDB; навыками использования AlphaFold; методами анализа белков-

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
			лигандных взаимодействий.
ПК-13	Способен использовать знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей, инженеров, технологов	ПК-13.1. Понимает принципы построения и основные методы систем искусственного интеллекта и применяет их для решения задач профессиональной деятельности.	Знать: Принципы построения и основные методы систем искусственного интеллекта; методы машинного обучения (обучение с учителем, без учителя, с подкреплением); принципы работы градиентного бустинга (XGBoost, LightGBM); принципы работы генеративных моделей (VAE, GAN); принципы работы графовых нейронных сетей (GNN); методы предобработки данных для ML; основные метрики оценки качества моделей. Уметь: Применять методы ИИ для решения задач профессиональной деятельности; применять методы машинного обучения для прогнозирования свойств веществ; применять градиентный бустинг для задач хемоинформатики; применять генеративные модели для генерации новых соединений; применять GNN для предсказания свойств молекул; применять методы предобработки данных; оценивать качество моделей. Владеть: Навыками применения методов ИИ для решения профессиональных задач; инструментами Python (Scikit-Learn, PyTorch, PyTorch Geometric) для реализации методов ИИ; навыками построения и валидации моделей машинного обучения; навыками применения генеративных моделей; навыками применения GNN.
ПК-13	Способен использовать	ПК-13.2. Исследует направления	Знать: Направления применения систем ИИ для различных

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
	знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей, инженеров, технологов	применения систем искусственного интеллекта для различных предметных областей.	предметных областей; особенности применения ИИ в химической промышленности и хемоинформатике; методы анализа данных в химической технологии; принципы построения моделей QSAR/QSPR; методы виртуального скрининга; принципы Ligand-based и Structure-based drug design. Уметь: Исследовать направления применения систем ИИ для различных предметных областей; анализировать особенности применения ИИ в химической промышленности; применять методы ИИ для анализа данных в химической технологии; строить модели QSAR/QSPR; проводить виртуальный скрининг; применять подходы Ligand-based и Structure-based drug design. Владеть: Навыками исследования направлений применения ИИ; методами анализа данных в химической технологии; навыками построения моделей QSAR/QSPR; навыками проведения виртуального скрининга; навыками применения подходов Ligand-based и Structure-based drug design.
ПК-13	Способен использовать знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей,	ПК-13.3. Выбирает комплексы методов и инструментальных средств искусственного интеллекта для решения профессиональных задач в зависимости от особенностей предметной области.	Знать: Комплексы методов и инструментальных средств ИИ; особенности выбора методов ИИ в зависимости от предметной области; критерии выбора инструментальных средств ИИ; особенности применения ИИ в химической промышленности; методы оценки эффективности ИИ-решений. Уметь: Выбирать комплексы методов и инструментальных

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
	инженеров, технологгов		средств ИИ для решения профессиональных задач; обосновывать выбор методов ИИ в зависимости от особенностей предметной области; выбирать инструментальные средства ИИ; оценивать эффективность ИИ-решений. Владеть: Навыками выбора комплексов методов и инструментальных средств ИИ; навыками обоснования выбора методов ИИ; навыками применения инструментальных средств ИИ; навыками оценки эффективности ИИ-решений.
ПК-14 (MnI-1)	Способен осуществлять трудовые функции, обусловленные профессиональной ролью, в ОПД «Обрабатывающая промышленность»	ПК-14.1. Применяет современные методы и технологии ИИ для решения актуальных задач промышленности.	Знать: Современные методы и технологии ИИ, применяемые для решения актуальных задач промышленности. Уметь: Реализовывать end-to-end проект по предсказанию свойств вещества, демонстрируя применение ИИ для конкретной технологической задачи. Владеть: Навыками применения современных методов и технологий ИИ в промышленности.
ПК-15 (SS-1)	Способен применять системное, критическое и креативное мышление в области искусственного интеллекта	ПК-15.1. Определяет ценностные предпосылки, когнитивные искажения, культурно-обусловленные предвзятости в данных, алгоритмах, постановке задач для ИИ.	Знать: Что качество обучающей выборки существенно определяет этико-социальные аспекты функционирования ИИ. Уметь: Выявить очевидные несоответствия между задачей для ИИ и обучающей выборкой. Владеть: Навыками критического анализа данных и моделей.
ПК-15 (SS-1)	Способен применять системное, критическое и	ПК-15.2. Применяет методики работы с этическими и социальными рисками,	Знать: Основные этические принципы (отсутствие дискриминации, справедливость, человекоориентированность,

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
	креативное мышление в области искусственного интеллекта	возникающими на разных стадиях жизненного цикла ИИ.	ответственность, безопасность, прозрачность, автономия человека). Уметь: Осознавать, что ИИ-системы могут порождать этические проблемы. Владеть: Навыками выявления и оценки этических проблем при работе с ИИ.
ПК-16 (BD-1)	Способен осуществлять поиск, сбор, очистку и предварительный анализ данных	ПК-16.1. Обосновывает способы и варианты применения методов предварительного анализа данных в задачах ИИ, включая их математическое (алгоритмическое) преобразование и адаптацию к специфике задачи.	Знать: Основные способы представления молекул (SMILES, молекулярные графы, дескрипторы) и их пригодность для разных задач ИИ. Уметь: Обосновать выбор конкретного представления (например, SMILES для генеративных моделей, графы для GNN, дескрипторы для QSAR). Владеть: Навыками анализа данных и подготовки их к применению современных методов ИИ.
ПК-16 (BD-1)	Способен осуществлять поиск, сбор, очистку и предварительный анализ данных	ПК-16.2. Применяет методы анализа данных для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.	Знать: Методы заполнения пропусков в данных и удаления выбросов в табличных данных. Уметь: Формулировать и проверять простейшие разведочные гипотезы о данных. Владеть: Навыками разведочного анализа данных.
ПК-16 (BD-1)	Способен осуществлять поиск, сбор, очистку и предварительный анализ данных	ПК-16.3. Применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных.	Знать: Существующие библиотеки, реализующие методы понижения размерности. Уметь: Применять метод главных компонент (PCA) для визуализации многомерного химического пространства в 2D/3D и интерпретировать полученные кластеры. Владеть: Навыками визуализации многомерных данных.
ПК-16 (BD-1)	Способен	ПК-16.4. Отбирает	Знать: Методы отбора признаков.

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
	осуществлять поиск, сбор, очистку и предварительный анализ данных	признаки данных, значимые для исследования.	Уметь: Проводить отбор признаков с использованием метода на основе важности признаков из модели градиентного бустинга. Владеть: Навыками отбора значимых признаков.
ПК-17 (ML-2)	Способен применять фундаментальные принципы и методы машинного обучения, включая подготовку данных, оценку качества моделей и работу с признаками	ПК-17.1. Различает основные типы задач машинного обучения и применяет на практике принципы их решения.	Знать: Разницу между задачами регрессии и классификации. Уметь: Выбирать соответствующие метрики качества. Владеть: Навыками решения базовых задач с использованием готовых инструментов и библиотек (Scikit-Learn).
ПК-17 (ML-2)	Способен применять фундаментальные принципы и методы машинного обучения, включая подготовку данных, оценку качества моделей и работу с признаками	ПК-17.2. Применяет методы предварительной обработки данных и работы с признаками.	Знать: Методы предобработки данных для ML: стандартизацию числовых признаков, кодирование категориальных переменных. Уметь: Проводить предобработку данных для ML. Владеть: Навыками работы с признаками.
ПК-17 (ML-2)	Способен применять фундаментальные принципы и методы машинного обучения, включая подготовку данных, оценку качества моделей и работу с признаками	ПК-17.3. Решает проблемы несбалансированных данных и оценивает качество моделей.	Знать: Понятие дисбаланса классов и его влияние на метрики. Уметь: Применять стратегию балансировки весов классов при обучении модели. Владеть: Навыками оценки качества моделей.
ПК-18 (ML-4)	Способен применять методы классического	ПК-18.1. Применяет алгоритмы кластеризации и	Знать: Стандартные алгоритмы кластеризации и понижения размерности.

Код компетенции	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:
	машинного обучения без учителя	понижения размерности для решения практических задач.	Уметь: Применять алгоритм k-средних для кластеризации библиотеки соединений по их дескрипторам и анализировать представительность кластеров. Владеть: Навыками применения библиотек для обучения без учителя (Scikit-Learn).
ПК-19 (PL-1)	Способен применять язык программирования Python для решения задач в области ИИ	ПК-19.1. Осуществляет выбор инструментов разработки на Python, приемлемых для создания прикладной системы обработки научных данных, машинного обучения и визуализации с заданными требованиями.	Знать: Основные библиотеки для научных вычислений (NumPy, SciPy, Pandas); основные библиотеки для визуализации данных (Matplotlib, Seaborn). Уметь: Применять основные библиотеки для научных вычислений и визуализации. Владеть: Навыками разработки на Python для решения задач в области ИИ.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

4.1. Структура дисциплины (модуля)

Общая трудоемкость дисциплины составляет 4 зачетные единицы, 144 часа. Контактная работа — 54 часа, из них: лекции — 18 часов, практические (семинарские) занятия — 18 часов, лабораторные работы — 18 часов. Самостоятельная работа — 90 часов. Дисциплина изучается в 5 семестре. Форма промежуточной аттестации — зачёт.

Таблица 4.1.

№ п/п	Наименование раздела дисциплины	Трудоемкость, акад. часы	Лекции, час.	Практические (семинарские) занятия, час.	Лабораторные работы, час.	Самостоятельная работа, час.	Контроль, час.
1.	Введение в хемоинформатику и представления молекул	36	6	4	6	20	—
1.1.	Графовое и	12	2	2	2	6	—

№ п/п	Наименование раздела дисциплины	Трудоемкость, акад. часы	Лекции, час.	Практические (семинарские) занятия, час.	Лабораторные работы, час.	Самостоятельная работа, час.	Контроль, час.
	линейное представление молекул (SMILES, SELFIES, дескрипторы, молекулярные графы)						
1.2.	Методы конвертации 2D/3D-структур	10	2	—	2	6	—
1.3.	Работа с молекулярными базами (ChEMBL, PubChem, QM9)	14	2	2	2	8	—
2.	Машинное обучение и предсказание свойств веществ	32	4	4	4	20	—
2.1.	Модели «структура-свойство», предобработка и валидация данных	16	2	2	2	10	—
2.2.	Методы тюнинга моделей, градиентный бустинг	16	2	2	2	10	—
3.	AI-инструменты для	32	4	4	4	20	—

№ п/п	Наименование раздела дисциплины	Трудоемкость, акад. часы	Лекции, час.	Практические (семинарские) занятия, час.	Лабораторные работы, час.	Самостоятельная работа, час.	Контроль, час.
	молекулярного дизайна						
3.1.	Молекулярная фильтрация (ADME, PAINS)	14	2	2	2	8	—
3.2.	Скрининг и генерация новых соединений	9	1	1	1	6	—
3.3.	Генеративные модели и вариационные автоэнкодеры	9	1	1	1	6	—
4.	Графовые нейросети в хемоинформатике	25	2	4	4	15	—
4.1.	GCNN, MPNN, SchNet, ALIGNN и их применение в предсказании свойств	13	1	2	2	8	—
4.2.	Ligand-based и structure-based drug design с применением GNN и трансформеров	12	1	2	2	7	—
5.	Работа с	19	2	2	—	15	—

№ п/п	Наименование раздела дисциплины	Трудоемкость, акад. часы	Лекции, час.	Практические (семинарские) занятия, час.	Лабораторные работы, час.	Самостоятельная работа, час.	Контроль, час.
	белковыми структурами и AI-инструменты						
5.1.	Сбор данных о белках (PDB), моделирование белок-лигандных взаимодействий	10	1	1	—	8	—
5.2.	Введение в AlphaFold как инструмент ИИ-предсказания третичной структуры белков	9	1	1	—	7	—
	ИТОГО	144	18	18	18	90	—

4.2. Содержание дисциплины (модуля)

В разделе 4.2 программы учебной дисциплины приводятся краткие аннотации структурных единиц материала дисциплины. Содержание дисциплины структурируется по разделам, темам или модулям и раскрывается в аннотациях рабочей программы с достаточной полнотой, чтобы обучающиеся могли изучать материал самостоятельно, опираясь на программу.

Таблица 4.2.

№ раздела	Наименование раздела дисциплины	Содержание дисциплины	Формируемые компетенции
1.	Введение в	1.1. Графовое и линейное представление молекул	ОПК-5, ПК-13,

№ раздела	Наименование раздела дисциплины	Содержание дисциплины	Формируемые компетенции
	хемоинформатику и представления молекул	(SMILES, SELFIES, дескрипторы, молекулярные графы). Определение химической структуры: пространственное расположение атомов в молекуле и типы связей между ними. Элементный состав соединения и взаимное расположение атомов. Понятие топологии молекулы и ее топологические характеристики (связность, циклы). Валентность атомов в химических соединениях. Различные способы представления молекул в химии (структурные формулы, химические формулы, тривиальные названия) и их ограничения. Особенности и требования к представлениям молекул в хемоинформатике. Линейные представления молекул: основы SMILES и SELFIES, их синтаксис и преимущества для кодирования молекулярной информации. Молекулы как графы: представление атомов как узлов и связей как ребер. SMILES и SELFIES как методы кодирования информации графа. Молекулярные дескрипторы: битовые строки (структурные ключи, отпечатки пальцев, хэшированные отпечатки) как сжатая информация о молекулярной структуре. Матричное представление молекулярных графов (например, матрица смежности) и их виды. Представление молекул с использованием координат атомов (3D-представление) как основа для более сложных анализов. 1.2. Методы конвертации 2D/3D-структур. Принципы конвертации химической структуры в линейные представления (SMILES, SELFIES). Методы преобразования между различными формами представления молекул (линейные, графовые, матричные, координатные). Основы и алгоритмы конвертации из 2D- в 3D-структуры молекул и обратно. 1.3. Работа с молекулярными базами (ChEMBL, PubChem, QM9). Обзор основных молекулярных баз данных: ChEMBL, PubChem, QM9. Типы информации, содержащиеся в этих базах (структуры, биологическая активность, физико-химические свойства). Методы поиска, извлечения и фильтрации данных из молекулярных баз для задач хемоинформатики и	ПК-14, ПК-15, ПК-16, ПК-19

№ раздела	Наименование раздела дисциплины	Содержание дисциплины	Формируемые компетенции
		машинного обучения.	
2.	Машинное обучение и предсказание свойств веществ	2.1. Модели «структура-свойство», предобработка и валидация данных. Концепция датасетов структура-свойство в хемоинформатике. Формирование таких датасетов из молекулярных баз данных (ChEMBL, PubChem, QM9): извлечение структур, дескрипторов и целевых свойств. Предварительная обработка данных для ML: очистка, нормализация, обработка пропусков и выбросов (включая «Предобработка данных. Общие принципы построения моделей»)). Предсказание свойств материалов/соединений на основе их химической структуры. Основные идеи и принципы построения моделей QSAR/QSPR. Связь между химической структурой и макроскопическими свойствами. Роль молекулярных дескрипторов: определение, генерация и использование для кодирования структурной информации. Типы моделей, используемых в QSAR/QSPR (линейные, нелинейные, ансамблевые). Примеры применения в разработке лекарств и материаловедении. 2.2. Методы тюнинга моделей, градиентный бустинг. Введение в ансамблевые методы машинного обучения. Алгоритм машинного обучения: основные принципы градиентного бустинга. Последовательное обучение: минимизация ошибки с использованием градиентного спуска. Преимущества градиентного бустинга (например, XGBoost, LightGBM) в хемоинформатике. Концепция тюнинга гиперпараметров моделей для оптимизации производительности (grid search, random search, Bayesian optimization).	ОПК-5, ПК-13, ПК-14, ПК-16, ПК-17, ПК-18, ПК-19
3.	AI-инструменты для молекулярного дизайна	3.1. Молекулярная фильтрация (ADME, PAINS). Нежелательные фрагменты, механизмы побочных действий. Параметры ADME, поиск подфрагментов PAINS. Виды поверхностей. Молекулярные формы. Нежелательные фрагменты, механизмы побочных действий, параметры ADME, поиск подфрагментов PAINS. Дескрипторы могут быть использованы для	ОПК-5, ПК-13, ПК-14, ПК-15, ПК-18, ПК-19

№ раздела	Наименование раздела дисциплины	Содержание дисциплины	Формируемые компетенции
		предсказания ADME свойств. Физико-химические дескрипторы, необходимые для оценки свойств, связанных с ADME. Поиск подфрагментов PAINS требует специализированных алгоритмов и баз данных. Параметры ADME. Почему важны параметры ADME. Безопасность. Разработка лекарств. Основные параметры ADME. Поиск подфрагментов PAINS (Pan Assay Interference Compounds). Почему поиск подфрагментов PAINS важен. Улучшение качества скрининга. Повышение эффективности. 3.2. Скрининг и генерация новых соединений. Использование для виртуального скрининга и для высокопроизводительного скрининга. Разбросанные и сфокусированные библиотеки. Типы библиотек, используемых в скрининге. 3.3. Генеративные модели и вариационные автоэнкодеры. Генерация структур, новые молекулярные структуры с желаемыми свойствами. RECAP. Fragmenster. Кластеризация молекул: может использоваться для анализа результатов скрининга. Отбор молекул без кластеризации: метод выбора молекул для скрининга. Машинное обучение в FBS.	
4.	Графовые нейросети в хемоинформатике	4.1. GCNN, MPNN, SchNet, ALIGNN и их применение в предсказании свойств. Применение молекулярного моделирования (применяются в молекулярном докинге). Модели «структура-свойство» на основе GCNN, генеративные модели G-SchNet. Дизайн библиотек данных. 4.2. Ligand-based и structure-based drug design с применением GNN и трансформеров. Применение молекулярного моделирования (используются для дизайна молекул и докинга). Генерация новых молекул. Учет свойств. Анализ графов. Симметричная передача сообщений. Вариационный автоэнкодер (VAE). Преимущества G-SchNet.	ОПК-5, ПК-13, ПК-17, ПК-14, ПК-19
5.	Работа с белковыми структурами и AI-инструменты	5.1. Сбор данных о белках (PDB), моделирование белок-лигандных взаимодействий. Обнаружение активных участков связывания, моделирование белок-лигандных взаимодействий методом	ОПК-5, ПК-13, ПК-19, ПК-16, ПК-15

№ раздела	Наименование раздела дисциплины	Содержание дисциплины	Формируемые компетенции
		молекулярной динамики. Графовые нейросети и трансформеры для Structure-based drug design (SBDD). 5.2. Введение в AlphaFold как инструмент ИИ-предсказания третичной структуры белков. Применение молекулярного моделирования (предсказание структуры белков для дальнейшего моделирования). AlphaFold: Система искусственного интеллекта (ИИ). Глубокое обучение. Прогнозирование структуры белков. Архитектура модели. Обучение модели. Точность предсказаний. Влияние на научные исследования.	

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При подготовке обучающихся используются следующие основные формы проведения учебных занятий:

- интерактивные лекции;
- лекции-пресс-конференции;
- тренинги и семинары по развитию профессиональных навыков;
- групповые, научные дискуссии, дебаты;
- проектная технология;
- кейс-технология;
- технология развития критического мышления;
- информационно-коммуникационная технология.

Таблица 5.1.

№	Семестр	Тема программы дисциплины	Применяемые технологии	Кол-во ауд. часов
1.	5	Графовое и линейное представление молекул (SMILES, SELFIES, дескрипторы, молекулярные графы)	Интерактивная лекция, практическое занятие, лабораторная работа	6
2.	5	Методы конвертации 2D/3D-структур	Лекция с презентацией, лабораторная работа	4
3.	5	Работа с молекулярными базами (ChEMBL, PubChem, QM9)	Лекция с презентацией, практическое занятие, лабораторная работа	6

№	Семестр	Тема программы дисциплины	Применяемые технологии	Кол-во ауд. часов
4.	5	Модели «структура-свойство», предобработка и валидация данных	Интерактивная лекция, практическое занятие, лабораторная работа	6
5.	5	Методы тюнинга моделей, градиентный бустинг	Лекция с презентацией, практическое занятие, лабораторная работа	6
6.	5	Молекулярная фильтрация (ADME, PAINS)	Интерактивная лекция, практическое занятие, лабораторная работа	6
7.	5	Скрининг и генерация новых соединений	Кейс-технология, практическое занятие, лабораторная работа	3
8.	5	Генеративные модели и вариационные автоэнкодеры	Лекция-пресс-конференция, практическое занятие, лабораторная работа	3
9.	5	GCNN, MPNN, SchNet, ALIGNN и их применение в предсказании свойств	Интерактивная лекция, практическое занятие, лабораторная работа	5
10.	5	Ligand-based и structure-based drug design с применением GNN и трансформеров	Кейс-технология, практическое занятие, лабораторная работа	5
11.	5	Сбор данных о белках (PDB), моделирование белок-лигандных взаимодействий	Интерактивная лекция, практическое занятие	2
12.	5	Введение в AlphaFold как инструмент ИИ-предсказания третичной структуры белков	Лекция с презентацией, практическое занятие	2
		ИТОГО		54

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ. ОЦЕНОЧНЫЕ СРЕДСТВА

6.1. План самостоятельной работы студентов

Таблица 6.1.

№ п/п	Наименование раздела (темы) дисциплины	Вид самостоятельной работы	Трудоемкость (в акад. часах)
1.	Введение в хемоинформатику и представления молекул	Подготовка реферата по истории хемоинформатики. Анализ методов представления молекул. Решение задач на конвертацию 2D/3D-структур.	20
2.	Машинное обучение и предсказание свойств веществ	Решение задач на построение моделей QSAR/QSPR. Подготовка презентации по методам тюнинга гиперпараметров.	20
3.	AI-инструменты для молекулярного дизайна	Анализ научной статьи по генеративным моделям в хемоинформатике. Решение кейсов по молекулярной фильтрации.	20
4.	Графовые нейросети в хемоинформатике	Подготовка обзора по современным архитектурам GNN для предсказания свойств молекул. Решение кейсов по SBDD.	15
5.	Работа с белковыми структурами и AI-инструменты	Анализ научной статьи по AlphaFold. Подготовка обзора по методам анализа белок-лигандных взаимодействий.	15
	ИТОГО		90

6.2. Методические указания по организации самостоятельной работы студентов

Учебным планом направления подготовки по дисциплине «Искусственный интеллект в хемоинформатике: от молекулярных репрезентаций до генерации новых соединений» предусматривается самостоятельная работа студента, которая выполняется следующими видами самостоятельной работы: написание контрольной работы по дисциплине, сдача коллоквиума, выполнение проектной работы.

6.2.1. Методические рекомендации по выполнению контрольной работы

Контрольная работа — одна из форм самостоятельной работы студентов, направленная на проверку усвоения теоретического материала и умения применять знания при решении практических задач. Данные рекомендации помогут правильно организовать выполнение и оформление работы в соответствии с требованиями дисциплины.

Цель контрольной работы: закрепление и систематизация знаний по теме, развитие навыков самостоятельной работы с литературой, формирование умения анализировать, обобщать и излагать материал, подготовка к зачету.

Требования к содержанию и структуре:

- Работа выполняется индивидуально.
- Объем: зависит от задания (обычно 10–15 страниц).
- Оформление: в соответствии с ГОСТ.
- Структура: титульный лист, содержание, введение (актуальность темы, цели и задачи работы), основная часть, практическая часть, заключение, список литературы, приложения (если требуются).

6.2.2. Методические рекомендации по подготовке и сдаче коллоквиума

Коллоквиум — форма промежуточного контроля знаний, направленная на проверку понимания теоретического материала, умения анализировать и применять полученные знания.

Цель коллоквиума: проверить уровень усвоения материала, оценить способность студента анализировать и систематизировать информацию, развить навыки устного изложения и аргументации. Предполагаемый объем ответа не должен быть большим (примерно 1,5–2 минуты), чтобы преподаватель мог успеть опросить всех студентов.

От студента требуется:

- владение изученным в ходе учебного процесса материалом, относящимся к рассматриваемой проблеме;
- наличие собственного мнения по обсуждаемым вопросам и умение его аргументировать.

6.3. Материалы для проведения текущего и промежуточного контроля знаний студентов

Контроль освоения компетенций

Таблица 6.3.

№ п/п	Вид контроля	Контролируемые темы (разделы)	Компетенции, компоненты которых контролируются
1.	Коллоквиум	Введение в хемоинформатику и представления молекул. Машинное обучение и предсказание свойств веществ. AI-инструменты для молекулярного дизайна. Графовые нейросети в хемоинформатике. Работа с белковыми структурами и AI-инструменты.	ОПК-5, ПК-13, ПК-14, ПК-15, ПК-16, ПК-17, ПК-18, ПК-19

№ п/п	Вид контроля	Контролируемые темы (разделы)	Компетенции, компоненты которых контролируются
2.	Зачёт (защита проектной работы)	Все разделы дисциплины.	ОПК-5, ПК-13, ПК-14, ПК-15, ПК-16, ПК-17, ПК-18, ПК-19

Примеры оценочных средств для проведения промежуточной аттестации:

Раздел 1. Введение в хемоинформатику и представления молекул (тестирование):

1. Какой из перечисленных способов представления молекул является линейной нотацией?

а) Молекулярный граф.

б) SMILES.

в) Матрица смежности.

г) 3D-координаты.

2. Что представляют собой узлы и рёбра в молекулярном графе?

а) Узлы — связи, рёбра — атомы.

б) Узлы — атомы, рёбра — связи.

в) Узлы — функциональные группы, рёбра — атомы.

г) Узлы — электроны, рёбра — орбитали.

3. Какая молекулярная база данных специализируется на биологической активности соединений?

а) QM9.

б) PubChem.

в) ChEMBL.

г) PDB.

4. Какую информацию кодируют молекулярные дескрипторы (фингерпринты)?

а) Только молекулярную массу.

б) Наличие определённых структурных фрагментов в молекуле.

в) Только тип гибридизации атомов.

г) Температуру плавления.

5. Какая библиотека Python чаще всего используется для загрузки и обработки молекулярных структур?

- а) NumPy.
- б) RDKit.
- в) Matplotlib.
- г) Requests.

Раздел 2. Машинное обучение и предсказание свойств веществ (тестирование):

1. Что такое QSAR-модель?

- а) Модель, предсказывающая квантовые свойства молекул.
- б) Модель, устанавливающая количественную связь между структурой молекулы и её свойством.
- в) Модель для генерации новых молекул.
- г) Модель для кластеризации молекул.

2. Какой метод машинного обучения относится к ансамблевым?

- а) Линейная регрессия.
- б) Логистическая регрессия.
- в) Градиентный бустинг.
- г) Метод главных компонент.

3. Какая метрика используется для оценки качества регрессионной модели?

- а) Accuracy.
- б) F1-score.
- в) RMSE.
- г) ROC-AUC.

4. Что означает "область применимости" (Applicability Domain) QSAR-модели?

- а) Область, в которой модель обучалась и в которой её предсказания надёжны.
- б) Область, в которой модель не работает.
- в) Область, в которой модель можно применять без ограничений.
- г) Область, в которой модель не требует данных.

5. Какие гиперпараметры градиентного бустинга влияют на переобучение?

- а) learning_rate и n_estimators.
- б) Только random_state.
- в) Только verbose.

г) Только p_jobs.

Раздел 3. AI-инструменты для молекулярного дизайна (тестирование):

1. Что означает аббревиатура ADME?

а) Adsorption, Distribution, Metabolism, Excretion.

б) Absorption, Distribution, Metabolism, Excretion.

в) Analysis, Design, Modeling, Evaluation.

г) Association, Distribution, Modeling, Estimation.

2. Что такое PAINS?

а) Соединения, обладающие высокой биологической активностью.

б) Соединения, склонные к ложноположительным результатам в скрининге.

в) Соединения, запрещённые к применению.

г) Соединения с высокой молекулярной массой.

3. Какая метрика оценивает "drug-likeness" молекулы?

а) LogP.

б) QED.

в) TPSA.

г) MW.

4. Что является основной задачей генеративных моделей в хемоинформатике?

а) Предсказание свойств существующих молекул.

б) Генерация новых молекул с заданными свойствами.

в) Классификация молекул.

г) Кластеризация молекул.

5. Какая архитектура лежит в основе вариационного автоэнкодера (VAE)?

а) Только полносвязные слои.

б) Энкодер-декодер с латентным пространством.

в) Рекуррентные слои.

г) Свёрточные слои.

Раздел 4. Графовые нейросети в хемоинформатике (тестирование):

1. В чём ключевое преимущество графовых нейронных сетей для задач хемоинформатики?

а) Они работают только с изображениями.

б) Они напрямую работают с топологией молекулы без ручного вычисления дескрипторов.

в) Они требуют меньше данных.

г) Они не требуют обучения.

2. Что такое "передача сообщений" (message passing) в GNN?

а) Обмен информацией между узлами графа.

б) Передача данных между слоями.

в) Передача градиентов.

г) Передача весов.

3. Какая архитектура GNN использует симметричную передачу сообщений?

а) GCNN.

б) MPNN.

в) SchNet.

г) ALIGNN.

4. В чём преимущество SchNet и ALIGNN для работы с 3D-структурами молекул?

а) Они работают с 2D-графами.

б) Они учитывают пространственное расположение атомов.

в) Они не требуют данных.

г) Они работают только с SMILES.

5. Какой подход к дизайну лекарств основан на структуре белка-мишени?

а) Ligand-based drug design.

б) Structure-based drug design.

в) Fragment-based drug design.

г) Phenotypic drug design.

Раздел 5. Работа с белковыми структурами и AI-инструменты (тестирование):

1. Что такое PDB?

а) База данных белковых структур.

б) База данных молекул.

в) База данных реакций.

г) База данных спектров.

2. Что предсказывает AlphaFold?

а) Биологическую активность молекул.

б) Третичную структуру белков.

в) Температуру плавления.

г) Растворимость.

3. Какая метрика используется для оценки достоверности предсказанной AlphaFold структуры?

а) pLDDT.

б) RMSD.

в) TM-score.

г) GDT.

4. Какие взаимодействия важны для связывания лиганда с белком?

а) Водородные связи.

б) Гидрофобные взаимодействия.

в) Электростатические взаимодействия.

г) Все перечисленные.

5. Какая библиотека Python используется для загрузки и анализа белковых структур?

а) NumPy.

б) Biopython.

в) Pandas.

г) Matplotlib.

6.4. Итоговая аттестация по дисциплине

Итоговая аттестация по дисциплине проводится в форме зачёта. Зачёт включает собеседование по вопросам и защиту проектной работы.

Цель проектной работы: демонстрация готовности слушателя к практическому применению технологий искусственного интеллекта для решения профессиональных задач в химической промышленности, подтверждающей освоение заявленных компетенций программы.

Примерный перечень тем проектных работ:

1. Разработка и валидация генеративной модели для дизайна перспективных ингибиторов киназ.
2. Создание robust QSPR-модели для предсказания критических свойств полимеров.
3. Сравнительный анализ архитектур графовых нейросетей для точного предсказания энергии сольватации.

Критерии оценивания сформированности компетенций

Оценивание уровня учебных достижений обучающихся по дисциплине осуществляется в виде промежуточной аттестации по разделам дисциплины и итоговой аттестации по дисциплине. Итоговая аттестация по дисциплине осуществляется в форме зачёта (защиты проектной работы), при этом проводится оценка компетенций, сформированных по дисциплине.

Знания, умения и навыки обучающихся по дисциплине определяются в следующей форме: «зачтено» (А), «не зачтено» (В).

Таблица 6.4.

Индикаторы компетенций	Критерии оценивания компетенций	
	А «Зачтено»	В «Не зачтено»
ОПК-5.1	Знает принципы представления химических структур в виде данных (SMILES, молекулярные графы, дескрипторы). Знает назначение и возможности молекулярных баз данных (PubChem, ChEMBL). Умеет выбирать и применять библиотеки Python (RDKit, Pandas) для загрузки молекулярных структур, их визуализации и вычисления дескрипторов.	Не может объяснить базовые принципы кодирования молекул или назначение молекулярных БД. Выполненный скрипт не работает или выполняет не все требуемые операции.
ОПК-5.2	Знает методы оценки и прогнозирования перспективности объектов, принципы построения моделей QSAR/QSPR. Умеет применять методы предобработки данных, алгоритмы кластеризации и понижения размерности, генеративные модели для создания новых соединений. Владеет навыками применения градиентного бустинга и генеративных моделей.	Не может сформулировать концепцию QSAR или объяснить принципы работы генеративных моделей. Не применяет методы предобработки данных.

Индикаторы компетенций	Критерии оценивания компетенций	
ОПК-5.3	Знает основные тенденции развития современных ИТ, основы информационной безопасности, принципы работы GNN, архитектуры GCNN, MPNN, SchNet, ALIGNN. Умеет использовать готовую реализацию GNN (PyTorch Geometric), применять AlphaFold, анализировать белок-лигандные взаимодействия.	Не может объяснить базовый механизм GNN и их преимущество. Не справляется с загрузкой или визуализацией белковых структур.
ПК-13.1	Понимает принципы построения и основные методы систем ИИ. Знает методы машинного обучения, принципы работы градиентного бустинга, генеративных моделей, GNN. Умеет применять методы ИИ для решения задач профессиональной деятельности.	Не понимает принципы построения систем ИИ. Не может применять методы машинного обучения для прогнозирования свойств веществ.
ПК-13.2	Знает направления применения систем ИИ для различных предметных областей, особенности применения ИИ в химической промышленности. Умеет строить модели QSAR/QSPR, проводить виртуальный скрининг.	Не знает направления применения систем ИИ. Не умеет строить модели QSAR/QSPR.
ПК-13.3	Знает комплексы методов и инструментальных средств ИИ. Умеет выбирать методы ИИ в зависимости от предметной области, обосновывать выбор, оценивать эффективность ИИ-решений.	Не может выбрать методы ИИ для решения профессиональных задач. Не умеет оценивать эффективность ИИ-решений.
ПК-14.1	Применяет современные методы и технологии ИИ для решения актуальных задач промышленности. Реализует end-to-end проект по предсказанию свойств вещества.	Не применяет современные методы и технологии ИИ для решения актуальных задач промышленности. Проект не сдан или не работает.
ПК-15.1	Понимает, что качество обучающей выборки существенно определяет этико-социальные аспекты функционирования ИИ. Может выявить очевидные несоответствия между задачей для ИИ и обучающей выборкой.	Не понимает, что качество обучающей выборки существенно определяет этико-социальные аспекты функционирования ИИ.

Индикаторы компетенций	Критерии оценивания компетенций	
ПК-15.2	Осознает, что ИИ-системы могут порождать этические проблемы. Знает основные этические принципы.	Не осознает, что ИИ-системы могут порождать этические проблемы. Не знает основные этические принципы.
ПК-16.1	Обосновывает выбор конкретного представления (SMILES для генеративных моделей, графы для GNN, дескрипторы для QSAR). Корректно объясняет, почему для прогноза свойства используются определённые представления.	Не может связать выбор представления с типом решаемой задачи. Выполняет слепое копирование.
ПК-16.2	Применяет методы заполнения пропусков и удаления выбросов. Проводит базовый анализ датасета, формулирует обоснованные гипотезы.	Не применяет методы заполнения пропусков и удаления выбросов. Анализ не проводит или гипотезы не сформулированы.
ПК-16.3	Использует существующие библиотеки, реализующие методы понижения размерности. Применяет PCA для визуализации многомерного химического пространства.	Не использует существующие библиотеки, реализующие методы понижения размерности.
ПК-16.4	Использует существующие библиотеки, реализующие методы отбора признаков. Проводит отбор признаков на основе важности из модели градиентного бустинга.	Не использует существующие библиотеки, реализующие методы отбора признаков.
ПК-17.1	Без ошибок определяет тип задачи (регрессия/классификация) по её описанию и выбирает корректные метрики для оценки модели.	Путает типы задач и метрики.
ПК-17.2	В скрипте подготовки данных корректно применяет StandardScaler и OneHotEncoder (или аналоги) до обучения модели.	Предобработка не проводится или приводит к ошибке при обучении модели.
ПК-17.3	На несбалансированном датасете применяет метод балансировки и сравнивает F1-score до и после, давая корректную интерпретацию.	Не распознаёт проблему дисбаланса или применяет метод неверно.
ПК-18.1	Получает кластеры, визуализирует их через PCA и даёт содержательное описание кластеров.	Кластеризация не проведена или интерпретация отсутствует.

Индикаторы компетенций	Критерии оценивания компетенций	
ПК-19.1	Знает и применяет основные библиотеки для научных вычислений (NumPy, SciPy, Pandas). Знает и применяет основные библиотеки для визуализации данных (Matplotlib, Seaborn).	Не знает и не применяет основные библиотеки для научных вычислений и визуализации данных.

Методические указания для обучающихся по освоению дисциплины

Таблица 6.5.

Вид учебных занятий / деятельности	Деятельность обучающегося
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения лекции, выводы, формулировки, обобщения; пометить важные мысли, выделять ключевые слова, термины. Обозначить вопросы, термины или другой материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе.
Практическое (семинарское) занятие	При подготовке к семинарскому (практическому) занятию необходимо проработать материалы лекций, основной и дополнительной литературы по заданной теме. Сформировать собственное мнение по выносимой на обсуждение тематике. Во время занятия активно участвовать в обсуждении вопросов.
Устный/письменный опрос	Отвечать, максимально полно, логично и структурировано, на поставленный вопрос. Основная цель — показать всю глубину знаний по конкретной теме или её части.
Подготовка к итоговой аттестации по дисциплине	При подготовке к итоговой аттестации необходимо проработать вопросы по темам, которые рекомендуются для самостоятельной подготовки. При возникновении затруднений следует ориентироваться на конспекты лекций, семинаров, рекомендуемую литературу.
Самостоятельная работа	Самостоятельная работа состоит из: 1) чтение учебной, справочной, научной литературы; 2) повторение материала лекций; 3) изучение документации к AI-сервисам, API и программным фреймворкам; 4) просмотр и анализ записей лекций, вебинаров. Практическая и проектная деятельность: разработка прототипов, работа с API, проектирование и создание цифровых артефактов.
Дискуссия	Публичное обсуждение спорного вопроса, проблемы. Каждая сторона должна оппонировать мнение собеседника, аргументируя свою позицию.

Вид учебных занятий / деятельности	Деятельность обучающегося
Тестирование (устное/письменное)	При подготовке к тестированию необходимо проработать материалы лекций, семинаров, основной и дополнительной литературы по заданной теме. Основная цель — показать уровень сформированности знаний.
Индивидуальная работа	При выполнении индивидуальной работы необходимо взять задание у преподавателя, ознакомиться с требованиями к выполнению работы, изучить поставленную проблему, найти решение проблемы. Оформить результаты работы.
Выполнение домашних заданий и групповых проектов	Для выполнения домашних заданий и групповых проектов необходимо получить формулировку задания от преподавателя и убедиться в понимании задания. Проработать материалы лекций, основной и дополнительной литературы по заданной теме.

Методические рекомендации для преподавателей

1. Организация практикума

- Проверьте доступность программного обеспечения и данных для слушателей.
- Используйте сквозные кейсы, чтобы показать связь между разделами.
- Поощряйте использование открытых образовательных datasets.

2. Оценивание

- Опирайтесь на матрицу компетенций и индикаторы.
- Учитывайте не только техническую реализацию, но и педагогическую целесообразность решений.

3. Ресурсы

- Рекомендуйте слушателям платформы для развёртывания моделей.
- Организуйте консультации по интеграции с LMS.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

7.1. Учебная литература:

а) основная литература:

1. Суханов, А. Я. Интеллектуальные системы: учебно-методическое пособие по лабораторным и практическим занятиям, самостоятельной и индивидуальной работе студентов / А. Я. Суханов. — Томск: Томский государственный университет систем управления и радиоэлектроники, 2023. — 147 с. — Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprbookshop.ru/152795.html>.

2. Пенькова, Т. Г. Модели и методы искусственного интеллекта: учебное пособие / Т. Г. Пенькова, Ю. В. Вайнштейн. — Красноярск: Сибирский федеральный университет, 2019. — 116 с. — ISBN 978-5-7638-4043-8. — Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprbookshop.ru/100056.html>.

3. Маккинли, Уэс. Python и анализ данных / Уэс Маккинли; перевод А. Слинкина. — 2-е изд. — Саратов: Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprbookshop.ru/88752.html>.

4. Рик, Гаско. Простой Python просто с нуля / Гаско Рик. — Москва: СОЛОН-Пресс, 2019. — 256 с. — ISBN 978-5-91359-334-4. — Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. — URL: <https://www.iprbookshop.ru/94940.html>.

5. Шевченко, Т. Г. Молекулярный дизайн биологически активных соединений в контексте создания лекарственных средств: учебно-методическое пособие / Т. Г. Шевченко; составители О. В. Тиньков, В. В. Люленова. — Тирасполь: ПГУ им. Т.Г. Шевченко, 2021. — 110 с. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/505658>.

6. Якубик, Д. Г. Химическая информатика: учебное пособие / Д. Г. Якубик. — Кемерово: КемГУ, 2021. — 79 с. — ISBN 978-5-8353-2734-8. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/173539>.

7. Газенаур, Е. Г. Компьютерные технологии в науке и образовании: информационные и коммуникационные технологии: учебное пособие / Е. Г. Газенаур, Л. В. Кузьмина, Н. В. Газенаур. — Кемерово: КемГУ, 2022. — 160 с. — ISBN 978-5-8353-2964-9. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/332318>.

б) дополнительная литература:

1. Диченко, Я. В. Компьютерное моделирование строения и реакционной способности молекул: монография / Я. В. Диченко. — Минск: Белорусская наука, 2023. — 137 с. — ISBN 978-985-08-2970-2. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/345584>.

2. Звягинцев, Н. В. Информационная система обработки и хранения данных о кинетике химических реакций / Н. В. Звягинцев, В. А. Биллиг // Программные продукты и системы. — 2018. — № 4. — С. 808–813. — ISSN 2311-2735. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/journal/issue/309488>.

3. Цирельсон, В. Г. Квантовая химия. Молекулы, молекулярные системы и твердые тела: учебное пособие / В. Г. Цирельсон. — 5-е эл. изд. — Москва: Лаборатория знаний, 2025. — 522 с. — ISBN 978-5-93208-518-9. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/495278>.

4. Гуреев, М. А. Молекулярный докинг и его верификация в контексте виртуального скрининга: учебно-методическое пособие / М. А. Гуреев, В. В. Кадочников, Ю. Б. Порозов.

— Санкт-Петербург: НИУ ИТМО, 2018. — 50 с. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/136513>.

5. Нейн, Ю. И. Компьютерное представление химической информации: учебное пособие / Ю. И. Нейн, М. Н. Иванцова. — Екатеринбург: УрФУ, 2020. — 142 с. — ISBN 978-5-7996-3018-8. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/361901>.

6. Глубокое обучение в биологии и медицине / Б. Рамсундар, П. Истман, П. Уолтерс, В. Панде; перевод с английского В. С. Яценкова. — Москва: ДМК Пресс, 2020. — 202 с. — ISBN 978-5-97060-791-6. — Текст: электронный // Лань: электронно-библиотечная система. — URL: <https://e.lanbook.com/book/131725>.

7.2. Интернет-ресурсы и профессиональные базы данных:

Образовательная организация, реализующая дисциплину, должна обеспечить обучающимся доступ (удаленный доступ) к профессиональным базам данных и информационным справочным системам.

Таблица 7.1.

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
Электронно-библиотечные системы:			
1.	http://www.iprbookshop.ru	Русскоязычная ЭБС на платформе «IPR Smart». Учебники и учебные пособия для университетов издательства «IPRbooks».	Индивидуальный неограниченный доступ
2.	http://e.lanbook.com	Русскоязычная ЭБС на платформе «Лань». Учебники и учебные пособия для университетов издательства «Лань».	Индивидуальный неограниченный доступ
3.	https://elibrary.ru	Научная электронная библиотека	Доступ свободный
Профессиональные базы данных и свободно распространяемое ПО:			
4.	https://www.python.org/downloads/	Сайт Python — для освоения и развития языка программирования Python. В каталоге пакетов Python (PyPI) представлены тысячи сторонних модулей для Python.	Индивидуальный неограниченный доступ
5.	https://docs.python.org/3/	Официальная документация Python	Индивидуальный неограниченный доступ

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
6.	https://numpy.org/	Фундаментальный пакет для научных вычислений на Python	Индивидуальный неограниченный доступ
7.	https://pandas.pydata.org/	Инструмент с открытым исходным кодом для анализа и обработки данных на Python	Индивидуальный неограниченный доступ
8.	https://scikit-learn.org/stable/	Scikit-learn — библиотека для реализации классических алгоритмов машинного обучения. Используется для построения диагностических и прогностических моделей	Индивидуальный неограниченный доступ
9.	https://matplotlib.org/	Matplotlib — комплексная библиотека для создания статических, анимированных и интерактивных визуализаций на Python	Индивидуальный неограниченный доступ
10.	https://seaborn.pydata.org/	Seaborn — библиотека визуализации данных Python, основанная на matplotlib	Индивидуальный неограниченный доступ
11.	https://colab.research.google.com/	Colab — облачная среда для обработки данных в области искусственного интеллекта	Индивидуальный неограниченный доступ
12.	https://jupyter.org/	Программное обеспечение, открытые стандарты и веб-сервисы для интерактивных вычислений на всех языках программирования	Индивидуальный неограниченный доступ
13.	https://www.rdkit.org/	RDKit: программное обеспечение для хемоинформатики с открытым исходным кодом	Индивидуальный неограниченный доступ
14.	https://pytorch-geometric.readthedocs.io/en/latest/	PyG (PyTorch Geometric) — библиотека, построенная на PyTorch, позволяет писать и обучать графовые нейронные сети (GNN) для широкого спектра приложений	Индивидуальный неограниченный доступ
15.	https://pytorch.org/	PyTorch — фреймворк для глубокого обучения. Применяется для построения моделей анализа данных и сложных прогнозных моделей	Индивидуальный неограниченный доступ

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
16.	https://www.tensorflow.org/	TensorFlow — фреймворк для глубокого обучения. Применяется для построения моделей анализа данных и сложных прогнозных моделей	Индивидуальный неограниченный доступ
17.	https://opencv.org/	OpenCV — проект с открытым исходным кодом, содержащий более 2500 алгоритмов и поддерживаемый некоммерческой организацией Open Source Vision Foundation	Индивидуальный неограниченный доступ
18.	https://labelstud.io/	Label Studio — инструмент для разметки данных. Используется слушателями для создания размеченных датасетов в рамках проектной работы и практических кейсов	Индивидуальный неограниченный доступ
19.	https://huggingface.co/docs/transformers/index	Hugging Face Transformers — библиотека для работы с архитектурой трансформеров. Используется для обработки SMILES-последовательностей и молекулярных языковых моделей	Индивидуальный неограниченный доступ
20.	https://www.kaggle.com/	Kaggle.com — открытая площадка, на которой собрано огромное количество информации по Data Science и машинному обучению	Индивидуальный неограниченный доступ
21.	https://github.com/git/git	Система контроля версий для ведения проектной документации, хранения кода и организации командной работы над проектами	Индивидуальный неограниченный доступ
22.	https://www.microsoft.com/ru-ru/microsoft-365/excel	Программа для первичной работы с табличными данными, создания структуры баз данных и выполнения простого анализа	Индивидуальный неограниченный доступ
23.	https://www.ebi.ac.uk/chembl/	Explore ChEMBL — база данных биологически активных молекул с аннотированными свойствами. Используется для формирования датасетов «структура–свойство»	Индивидуальный неограниченный доступ
24.	https://www.rcsb.org/	RCSB (RCSB PDB) — база данных трёхмерных структур белков и нуклеиновых кислот. Используется в модуле по работе с белковыми структурами	Индивидуальный неограниченный доступ

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
25.	https://pytorch-geometric.readthedocs.io/en/2.5.0/generated/torch_geometric_datasets.QM9.html	Датасет квантово-химических свойств для ~134 тыс. малых органических молекул. Применяется для обучения моделей предсказания свойств	Индивидуальный неограниченный доступ
26.	https://alphafold.ebi.ac.uk/	База предсказанных структур белков, созданных с помощью AlphaFold. Источник данных для моделирования белок-лигандных взаимодействий	Индивидуальный неограниченный доступ
27.	https://www.docker.com/	Платформа для контейнеризации приложений. Используется для создания воспроизводимых сред выполнения расчётов и развёртывания моделей	Индивидуальный неограниченный доступ
28.	https://streamlit.io/	Фреймворк для создания веб-интерфейсов. Применяется для разработки прототипов веб-сервисов прогнозирования свойств молекул	Индивидуальный неограниченный доступ
29.	https://telemost.yandex.ru/	Платформа для проведения онлайн-лекций, семинаров, консультаций и защиты проектов в режиме реального времени	Индивидуальный неограниченный доступ
30.	https://physionet.org/	Ресурс для исследования сложных физиологических сигналов	Индивидуальный неограниченный доступ
Информационные справочные системы:			
31.	https://minobrnauki.gov.ru/	Официальный сайт Министерства науки и высшего образования Российской Федерации	Доступ свободный
32.	https://www.edu.ru/	Федеральный портал «Российское образование»	Доступ свободный
33.	https://www.sciencedirect.com	База данных ScienceDirect	Доступ свободный
34.	https://edu.garant.ru/go/	Справочно-правовая система «Гарант», сетевая версия, доступна со всех компьютеров в корпоративной сети ИнГУ	Индивидуальный неограниченный доступ

7.3. Программное обеспечение

Университет обеспечен необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства.

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронной информационно-образовательной среде университета из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети «Интернет» как на территории университета, так и вне её.

Таблица 7.2.

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Windows 7, Windows 10	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Офисные приложения:		
МойОфис	отечественное	свободно распространяемое
Microsoft Office 2007, 2010	зарубежное	лицензионное
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
WEEK	отечественное	свободно распространяемое
Shtab	отечественное	свободно распространяемое
Антивирусная защита:		
Dr. Web 13	отечественное	лицензионное
Среды разработки и библиотеки:		
OpenIDE	отечественное	свободно распространяемое
TensorFlow	зарубежное	свободно распространяемое
PyTorch	зарубежное	свободно распространяемое
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Label Studio	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colab	зарубежное	свободно распространяемое
Streamlit	зарубежное	свободно распространяемое
Scikit-learn	зарубежное	свободно распространяемое
Matplotlib	зарубежное	свободно распространяемое
RDKit	зарубежное	свободно распространяемое

Наименование ПО	Производство	Лицензионное / свободно распространяемое
PyTorch Geometric (PyG)	зарубежное	свободно распространяемое
Biopython	зарубежное	свободно распространяемое
PyMOL	зарубежное	свободно распространяемое
Сервисы и службы:		
Docker	зарубежное	свободно распространяемое

7.4. Материально-техническое обеспечение дисциплины

Материально-техническая база университета позволяет обеспечивать качественное проведение теоретических и практических занятий.

Перечень необходимых технических средств обучения, используемых в учебном процессе для освоения дисциплины:

- компьютерное и мультимедийное оборудование;
- видео- и аудиовизуальные средства обучения и др.

Рабочая программа дисциплины составлена в соответствии с ФГОС ВО — бакалавриат по направлению подготовки 04.03.01 Химия, утвержденным приказом Министерства образования и науки Российской Федерации от 17 июля 2017 г. № 671, в действующей редакции, а также с учетом содержания и структуры ОПОП направления подготовки 04.03.01 Химия Ингушского государственного университета. В части библиографического описания использован ГОСТ Р 7.0.100-2018.

Программу составил(а): старший преподаватель Осмиев Р.М.

Рабочая программа утверждена на заседании кафедры химии

Протокол заседания № 8 от « 22 » апреля 2026 г.

Рабочая программа одобрена учебно-методическим советом

химико-биологического факультета

Протокол заседания № 6 от « 28 » апреля 2026 г.

Сведения о переутверждении программы на очередной учебный год и регистрации изменений

Учебный год	Решение кафедры (№ протокола, дата)	Внесенные изменения	Подпись зав. кафедрой