

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования «Новосибирский национальный исследовательский
государственный университет» (Новосибирский государственный университет, НГУ)

Факультет информационных технологий

СОГЛАСОВАНО

Декан ФИТ НГУ

 М.М. Лаврентьев

«25» апреля 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Приложения ИИ и Data Science в задачах биологии

Направление подготовки: 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Направленность (профиль): Искусственный интеллект и Data Science

Форма обучения: очная

Год обучения: 1, семестр: 2

№	Вид деятельности	Семестр
		2
1	Лекции, час.	16
2	Практические занятия, час.	32
3	Лабораторные занятия, час.	
4	Занятий в контактной форме без учета промежуточной аттестации, час, из них	48
5	в электронной форме, час.	
6	из них аудиторных занятий, час.	48
7	из них в активной и интерактивной форме, час.	48
8	консультаций, час.	
9	Самостоятельная работа, час.	58
10	в том числе на выполнение письменных работ, час	40
11	Форма аттестации (экзамен, зачет, дифференцированный зачет), час	ДЗ 2
12	Всего зачетных единиц ¹	3

Новосибирск 2023

¹ С учетом выделенных часов на промежуточную аттестацию

Рабочая программа дисциплины составлена на основании федерального государственного образовательного стандарта (ФГОС) высшего образования - магистратура по направлению подготовки 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА.

Федеральный государственный образовательный стандарт (ФГОС) высшего образования - магистратура по направлению подготовки 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА введен в действие приказом Минобрнауки от 19.09.2017 № 918.

Место дисциплины в структуре учебного плана: Блок 1 Дисциплины (модули); часть, формируемая участниками образовательных отношений, дисциплина по выбору

Рабочая программа дисциплины утверждена решением Ученого совета факультета информационных технологий от 24.04.2023, протокол №91.

Программу разработал:

Старший преподаватель кафедры систем информатики ФИТ
кандидат биологических наук



Д.В.Антонец

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук



М.М. Лаврентьев

Ответственный за образовательную программу:

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук



М.М. Лаврентьев

Аннотация к рабочей программе дисциплины «Приложения ИИ и Data Science в задачах биологии»

Дисциплина «**Приложения ИИ и Data Science в задачах биологии**» реализуется в рамках образовательной программы высшего образования – программы магистратуры 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА, направленность (профиль): ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И DATA SCIENCE по очной форме обучения на русском языке.

Место в образовательной программе:

Дисциплина «**Приложения ИИ и Data Science в задачах биологии**» реализуется во 2 семестре в рамках части, формируемой участниками образовательных отношений, дисциплин (модулей) Блока 1 и является дисциплиной по выбору.

Данный курс является базовым для работы в рамках практики и выполнения выпускной квалификационной работы.

Дисциплина «**Приложения ИИ и Data Science в задачах биологии**» направлена на формирование компетенций:

Способен разрабатывать программные решения на основе аналитики больших данных (ПКС-1), в части следующих индикаторов достижения компетенции:

ПКС-1.1 Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности

ПКС-1.2 Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными

Перечень основных разделов дисциплины:

Преподавание дисциплины предусматривает проведение следующих видов учебных занятий: лекции, практические занятия, самостоятельную работу студента.

Содержание дисциплины охватывает круг вопросов, связанных с применением искусственного интеллекта и Data Science при решении актуальных задач медицины.

Основные темы:

1. Введение. Что такое биоинформатика. Разнообразие биологических данных, разнообразие форматов хранения биологических данных.
2. Анализ изображений. Выделение объектов, сегментация, классификация, подсчет объектов и т.п. Анализ медицинских изображений, поиск патологии. Примеры.
3. Анализ табличных данных. Поиск маркерных признаков. Построение функциональной зависимости - задачи регрессии. Примеры.
4. Задачи снижения размерности. PCA, PLS, tSNE, UMAP, автокодирующие нейронные сети. Примеры.
5. Задачи кластеризации. Метод k-средних, иерархическая кластеризация, DBSCAN, нечеткая кластеризация. Метод k-ближайших соседей. Примеры.
6. Задачи, связанные с анализом биологических последовательностей. Важность выбора схемы способа параметризации данных. Предсказание структурно-функциональных свойств. Языковые модели. Примеры.

7. Задачи хемоинформатики. Примеры. Классические подходы на основе таблиц дескрипторов. Языковые модели.
8. Анализ результатов секвенирования нового поколения. Использование методов машинного обучения для интерпретации сигнала. Предсказание эффектов мутаций.
9. Работа с большими данными в биологии. Геномика, транскриптомика, транскриптомика единичных клеток. Форматы данных.

Общий объем дисциплины – 3 зачетные единицы (108 часов)

Правила аттестации по дисциплине.

Текущий контроль по дисциплине «Приложения ИИ и Data Science в задачах биологии» осуществляется на практических занятиях на основании оценки за портфолио (оценка за задания по темам занятий). По результатам защиты портфолио выставляется оценка «зачтено» или «не зачтено».

Промежуточная аттестация по дисциплине «Приложения ИИ и Data Science в задачах биологии» проводится по завершению периода ее освоения (семестра). Промежуточная аттестация по дисциплине включает 2 этапа:

- 1) портфолио (оценка за реферат и устные ответы на занятиях);
- 2) дифзачет.

Оценка «зачтено» за портфолио является необходимым условием для прохождения промежуточной аттестации. Результаты промежуточной (итоговой по дисциплине) аттестации определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Учебно-методическое обеспечение дисциплины.

Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора

1. Внешние требования к дисциплине

Таблица 1.1

Компетенция ПКС-1 Способен разрабатывать программные решения на основе аналитики больших данных, в части следующих индикаторов достижения компетенции:
ПКС-1.1 Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности
ПКС-1.2 Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными

2. Требования к результатам освоения дисциплины

Таблица 2.1

Результаты изучения дисциплины по уровням освоения (иметь представление, знать, уметь, владеть)	Формы организации занятий		
	Лекции	Практики / семинары	Самостояте льная работа
ПКС-1.1 Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности			
1. Знать основы инструментария и технологий больших данных для решения задач биологии	+	+	+
ПКС-1.2 Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными			
2. Уметь применять основы инструментария и технологий больших данных для решения задач биологии	+	+	+

3. Содержание и структура учебной дисциплины

Таблица 3.1

Темы лекций	Активные формы, час.	Часы	Ссылки на результаты обучения
Семестр: 2			
Введение. Что такое биоинформатика. Разнообразие биологических данных, разнообразие форматов хранения биологических данных. Анализ изображений. Выделение объектов, сегментация, классификация, подсчет объектов и т.п. Анализ медицинских изображений, поиск патологии. Примеры.	4	4	1, 2
Анализ табличных данных. Поиск маркерных признаков. Построение функциональной зависимости - задачи регрессии. Примеры. Задачи снижения размерности. PCA, PLS, tSNE, UMAP, автокодирующие нейронные сети. Примеры. Задачи кластеризации. Метод k-средних, иерархическая кластеризация, DBSCAN, нечеткая кластеризация. Метод k-ближайших соседей. Примеры. Задачи, связанные с анализом биологических последовательностей. Важность выбора схемы способа параметризации данных. Предсказание структурно-функциональных свойств. Языковые модели. Примеры.	4	4	1, 2

Задачи хемоинформатики. Примеры. Классические подходы на основе таблиц дескрипторов. Языковые модели. Анализ результатов секвенирования нового поколения. Использование методов машинного обучения для интерпретации сигнала. Предсказание эффектов мутаций. Работа с большими данными в биологии. Геномика, транскриптомика, транскриптомика единичных клеток. Форматы данных.	8	8	1, 2
Итого	16	16	

Таблица 3.2

Темы практических занятий	Активные формы, час.	Часы	Ссылки на результаты обучения	Учебная деятельность
Семестр: 2				
Введение. Что такое биоинформатика. Разнообразие биологических данных, разнообразие форматов хранения биологических данных.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Анализ изображений. Выделение объектов, сегментация, классификация, подсчет объектов и т.п. Анализ медицинских изображений, поиск патологии. Примеры.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Анализ табличных данных. Поиск маркерных признаков. Построение функциональной зависимости - задачи регрессии. Примеры.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Задачи снижения размерности. PCA, PLS, tSNE, UMAP, автокодирующие нейронные сети. Примеры.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Задачи кластеризации. Метод k-средних, иерархическая кластеризация, DBSCAN, нечеткая кластеризация. Метод k-ближайших соседей. Примеры.	2	2	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по

				дисциплине
Задачи, связанные с анализом биологических последовательностей. Важность выбора схемы способа параметризации данных. Предсказание структурно-функциональных свойств. Языковые модели. Примеры.	2	2	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Задачи хемоинформатики. Примеры. Классические подходы на основе таблиц дескрипторов. Языковые модели.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Анализ результатов секвенирования нового поколения. Использование методов машинного обучения для интерпретации сигнала. Предсказание эффектов мутаций.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Работа с большими данными в биологии. Геномика, транскриптомика, транскриптомика единичных клеток. Форматы данных.	4	4	1, 2	Разбор теоретической темы, представленной на лекции, выполнение заданий, подготовка к контролю текущих знаний по дисциплине
Итого	32	32		

4. Самостоятельная работа студентов

Таблица 4.1

№	Виды самостоятельной работы	Ссылки на результаты обучения	Часы на выполнении	Часы на консультации
Семестр: 2				
1	изучение разделов дисциплины по учебной литературе, в том числе вопросов, не освещаемых на лекциях Изучение предлагаемых теоретических разделов в соответствии с настоящей Программой. Учебно-методические материалы по дисциплине «Приложения ИИ и Data Science в задачах биологии» выложены на странице курса в сети Интернет	1, 2	4	
2	Подготовка к практическим занятиям, к текущему контролю знаний Выполнение заданий	1, 2	40	

3	Подготовка к дифзачету	1. 2	14	
	Повторение теоретического материала по вопросам, совпадающим с темами лекций			
	Итого		58	0

5. Образовательные технологии

В ходе реализации учебного процесса по дисциплине применяются лекционные, лабораторные и практические занятия, а также применяются следующие интерактивные формы обучения (таблица 5.1).

Таблица 5.1

1	Лекция в форме дискуссии	ПКС-1.1,1.2
Формируемые умения: Знать основы инструментария и технологий больших данных для решения задач биологии Уметь применять основы инструментария и технологий больших данных для решения задач биологии		
Краткое описание применения: Представляется теория, проблематика вопросов, связанных с применением ИИ и Data Science для решения типовых задач, возникающих в области биологии; основные подходы к их решению		
2	Портфолио	ПКС-1.1,1.2
Формируемые умения: Знать основы инструментария и технологий больших данных для решения задач биологии Уметь применять основы инструментария и технологий больших данных для решения задач биологии		
Краткое описание применения: студенты ведут портфолио, которое является основой для проведения аттестации по дисциплине		

Для организации и контроля самостоятельной работы студентов, а также проведения консультаций применяются информационно-коммуникационные технологии (таблица 5.2).

Таблица 5.2

Информирование	Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора
Консультирование	Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора
Контроль	Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора
Размещение учебных материалов	Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора

6. Правила аттестации по учебной дисциплине

Текущий контроль по дисциплине «Приложения ИИ и Data Science в задачах биологии» осуществляется на практических занятиях на основании оценки за портфолио. По результатам защиты портфолио выставляется оценка «зачтено» или «не зачтено».

Промежуточная аттестация по дисциплине «Приложения ИИ и Data Science в задачах биологии» проводится по завершению периода ее освоения (семестра). Промежуточная аттестация по дисциплине включает 2 этапа:

- 1) портфолио (задания по темам занятий);
- 2) дифзачет.

Оценка «зачтено» за портфолио является необходимым условием для прохождения промежуточной аттестации. Результаты промежуточной (итоговой по дисциплине) аттестации определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

В таблице 6.1 представлено соответствие форм аттестации заявляемым требованиям к результатам освоения дисциплины.

Таблица 6.1

Коды компетенций ФГОС	Результаты обучения	Формы аттестации	
		портфолио	дифзачет
ПКС-1	ПКС-1.1 Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности	+	+
ПКС-1	ПКС-1.2 Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными	+	+

Требования к структуре и содержанию портфолио, оценочные средства, а также критерии оценки сформированности компетенций и освоения дисциплины в целом, представлены в Фонде оценочных средств, являющемся приложением 1 к настоящей рабочей программе дисциплины.

7. Литература

1. Володченкова, Л.А. Биоинформатика : учебное пособие : [16+] / Л.А. Володченкова ; Министерство образования и науки РФ, Омский государственный университет им. Ф. М. Достоевского. – Омск : Омский государственный университет им. Ф.М. Достоевского, 2018. – 44 с. : ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=563147> – Библиогр. в кн. – ISBN 978-5-7779-2214-4. – Текст : электронный.

Дополнительная литература (в т.ч. учебная)

5. Пак, И.В. Введение в биотехнологию : учебное пособие : [16+] / И.В. Пак, О.В. Трофимов, О.А. Величко ; Тюменский государственный университет. – 3-е изд., перераб. и доп. – Тюмень : Тюменский государственный университет, 2018. – 160 с. : ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=567615> – Библиогр.: с. 144. – ISBN 978-5-400-01454-3. – Текст : электронный.

Интернет-ресурсы

Таблица 7.1

№ п/п	Наименование Интернет-ресурса	Краткое описание
1.	http://www.nsu.ru/xmlui/	Электронная библиотека НГУ
2.	http://www.spsl.nsc.ru	Портал ГПНТБ СО РАН
3.	http://libra.nsu.ru	НГУ. Научная электронная библиотека
4.	http://e.lanbook.com/	Электронно-библиотечная система «Лань»

8. Учебно-методическое и программное обеспечение дисциплины

8.1. Учебно-методическое обеспечение

Электронный учебный ресурс на Google Classroom. Ресурс создается для каждого нового набора

8.2. Программное обеспечение

Для обеспечения реализации дисциплины используется стандартный комплект программного обеспечения (ПО), включающий регулярно обновляемое лицензионное ПО Windows и MS Office.

Специализированное программное обеспечение не требуется.

9. Профессиональные базы данных и информационные справочные системы

1. Полнотекстовые журналы Springer Journals за 1997-2015 г., электронные книги (2005-2016 гг.), коллекция научных биомедицинских и биологических протоколов SpringerProtocols, коллекция научных материалов в области физических наук и инжиниринга SpringerMaterials, реферативная БД по чистой и прикладной математике zbMATH.

2. Электронная библиотека диссертаций Российской государственной библиотеки (ЭБД РГБ)

3. Электронные ресурсы Web of Science Core Collection (Thomson Reuters Scientific LLC.), Journal Citation Reports + ESI

4. БД Scopus (Elsevier)

10. Материально-техническое обеспечение

Таблица 10.1

№	Наименование	Назначение
1	Презентационное оборудование (мультимедиа-проектор, экран, компьютер для управления)	Для проведения лекционных и практических занятий
2	Компьютерный класс (с выходом в Internet)	Для организации самостоятельной работы обучающихся

Материально-техническое обеспечение образовательного процесса по дисциплине для обучающихся из числа лиц с ограниченными возможностями здоровья осуществляется согласно «Порядку организации и осуществления образовательной деятельности по образовательным программам для инвалидов и лиц с ограниченными возможностями здоровья в Новосибирском государственном университете».

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования «Новосибирский национальный исследовательский
государственный университет» (Новосибирский государственный университет, НГУ)

Факультет информационных технологий

СОГЛАСОВАНО

Декан ФИТ НГУ

 М.М. Лаврентьев

«25» апреля 2023 г.

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ
по дисциплине Приложения ИИ и Data Science в задачах биологии**

Направление подготовки: 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Направленность (профиль): Искусственный интеллект и Data Science

Квалификация: магистр

Форма обучения: очная

Год обучения: 1, семестр 2

Форма аттестации	Семестр
Дифференцированный зачет	2

Новосибирск 2023

Фонд оценочных средств промежуточной аттестации по дисциплине является **Приложением 1** к рабочей программе дисциплины «Приложения ИИ и Data Science в задачах биологии», реализуемой в рамках образовательной программы высшего образования – программы магистратуры 09.04.01 Информатика и вычислительная техника. Искусственный интеллект и Data Science.

Фонд оценочных средств промежуточной аттестации по дисциплине утвержден решением ученого совета факультета информационных технологий, протокол №91 от 24.04.2023.

Разработчик:

Старший преподаватель кафедры систем информатики ФИТ,
кандидат биологических наук



Д.В.Антонец

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук



М.М. Лаврентьев

Ответственный за образовательную программу:

Заведующий кафедрой систем информатики ФИТ,
доктор физико-математических наук



М.М. Лаврентьев

1. Содержание и порядок проведения промежуточной аттестации по дисциплине

1.1. Общая характеристика содержания промежуточной аттестации

Промежуточная аттестация по дисциплине «Приложения ИИ и Data Science в задачах биологии» проводится по завершению периода освоения образовательной программы (семестра) для оценки сформированности компетенций в части следующих индикаторов достижения компетенции (таблица П1.1).

Таблица П1.1

Код	Компетенции, формируемые в рамках дисциплины «Приложения ИИ и Data Science в задачах биологии»	Семестр 2	
		Портфолио	Дифзачет
	ПКС-1 Способен разрабатывать программные решения на основе аналитики больших данных		
ПКС-1.1	Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности	+	+
ПКС-1.2	Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными	+	+

Промежуточная аттестация по дисциплине включает 2 этапа: портфолио и дифзачет.

Тематика вопросов соответствует разделам (темам) дисциплины:

Введение. Что такое биоинформатика. Разнообразие биологических данных, разнообразие форматов хранения биологических данных.

Анализ изображений. Выделение объектов, сегментация, классификация, подсчет объектов и т.п. Анализ медицинских изображений, поиск патологии. Примеры.

Анализ табличных данных. Поиск маркерных признаков. Построение функциональной зависимости - задачи регрессии. Примеры.

Задачи снижения размерности. PCA, PLS, tSNE, UMAP, автокодирующие нейронные сети. Примеры.

Задачи кластеризации. Метод k-средних, иерархическая кластеризация, DBSCAN, нечеткая кластеризация. Метод k-ближайших соседей. Примеры.

Задачи, связанные с анализом биологических последовательностей. Важность выбора схемы способа параметризации данных. Предсказание структурно-функциональных свойств. Языковые модели. Примеры.

Задачи хеоминформатики. Примеры. Классические подходы на основе таблиц дескрипторов. Языковые модели.

Анализ результатов секвенирования нового поколения. Использование методов машинного обучения для интерпретации сигнала. Предсказание эффектов мутаций.

Работа с большими данными в биологии. Геномика, транскриптомика, транскриптомика единичных клеток. Форматы данных.

1.2. Порядок проведения промежуточной аттестации по дисциплине

Промежуточная аттестация проводится в форме дифзачета. Необходимым условием для прохождения промежуточной аттестации является оценка «зачтено» за портфолио. Оценка «зачтено» за портфолио выставляется при условии выполнения и защиты работы.

Дифзачет проводится в устной форме, в аудитории, студентам разрешено пользоваться бумагой для записей и авторучкой. Справочной, учебной и другой литературой пользоваться не разрешается. Использование электронных устройств (телефоны, любые виды компьютеров, т.д.) запрещено.

2. Требования к структуре и содержанию фонда оценочных средств промежуточной аттестации по дисциплине

Перечень оценочных средств, применяемых на каждом этапе проведения промежуточной аттестации по дисциплине, представлен в таблице П1.2.

Таблица П1.2

№ п/п	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
Этап 1 - портфолио			
1.	Портфолио	Целевая подборка работ студента, раскрывающая его индивидуальные образовательные достижения в одной или нескольких учебных дисциплинах.	Структура портфолио
Этап 2 – дифзачет			
2	Собеседование	Средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т.п.	Вопросы по темам/разделам дисциплины

2.1. Требования к структуре и содержанию оценочных средств аттестации

2.1.1 Требования к структуре и содержанию портфолио

Портфолио включает выполнение заданий по темам практических занятий.

По результатам портфолио выставляется оценка «зачтено» или «не зачтено». Оценка «зачтено» является необходимым условием для прохождения промежуточной аттестации.

2.1.2 Перечень вопросов дифзачета

1. Разнообразие биологических данных, разнообразие форматов хранения биологических данных.
2. Анализ изображений. Выделение объектов, сегментация, классификация, подсчет объектов и т.п. Анализ медицинских изображений, поиск патологии. Примеры.
3. Анализ табличных данных. Поиск маркерных признаков. Построение функциональной зависимости - задачи регрессии. Примеры.
4. Задачи снижения размерности. PCA, PLS, tSNE, UMAP, автокодирующие нейронные сети. Примеры.
5. Задачи кластеризации. Метод k-средних, иерархическая кластеризация, DBSCAN, нечеткая кластеризация. Метод k-ближайших соседей. Примеры.

6. Задачи, связанные с анализом биологических последовательностей. Важность выбора схемы способа параметризации данных. Предсказание структурно-функциональных свойств. Языковые модели. Примеры.
7. Задачи хемоинформатики. Примеры. Классические подходы на основе таблиц дескрипторов. Языковые модели.
8. Анализ результатов секвенирования нового поколения. Использование методов машинного обучения для интерпретации сигнала. Предсказание эффектов мутаций.
9. Работа с большими данными в биологии. Геномика, транскриптомика, транскриптомика единичных клеток. Форматы данных.

Набор вопросов для дифзачета формируется и утверждается в установленном порядке в начале учебного года при наличии контингента обучающихся, завершающих освоение дисциплины «Приложения ИИ и Data Science в задачах биологии» в текущем учебном году.

3. Критерии оценки сформированности компетенций в рамках промежуточной аттестации по дисциплине

Таблица П1.5

Шифр компетенций	Структурные элементы оценочных средств	Показатель сформированности	Не сформирован	Пороговый уровень	Базовый уровень	Продвинутый уровень
ПКС-1	Портфолио (этап 1), Дифзачет (этап 2)	ПКС-1.1 Разрабатывает программные решения на основе аналитики больших данных в области профессиональной деятельности	Не знает основы инструментария и технологий больших данных для решения задач биологии	Демонстрирует слабое знание основ инструментария и технологий больших данных для решения задач биологии	Допускает незначительные погрешности, в целом, знает основы инструментария и технологий больших данных для решения задач биологии	Знает глубоко и уверенно основы инструментария и технологий больших данных для решения задач биологии
ПКС-1	Портфолио (этап 1), Дифзачет (этап 2)	ПКС-1.2 Разрабатывает новые и адаптирует существующие методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными	Не умеет применять основы инструментария и технологий больших данных для решения задач биологии	Слабо умеет применять основы инструментария и технологий больших данных для решения задач биологии	Допускает незначительные ошибки, умеет применять основы инструментария и технологий больших данных для решения задач биологии	Демонстрирует высокий уровень умения применять основы инструментария и технологий больших данных для решения задач биологии

4. Критерии выставления оценок по результатам промежуточной аттестации по дисциплине

Результаты промежуточной аттестации в 2 семестре определяются оценками «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Оценки «отлично», «хорошо», «удовлетворительно» означают успешное прохождение промежуточной аттестации.

Оценка «отлично» соответствует продвинутому уровню сформированности компетенции.

Оценка «хорошо» соответствует базовому уровню сформированности компетенции.

Оценка «удовлетворительно» соответствует пороговому уровню сформированности компетенции.

Оценка «неудовлетворительно» выставляется, если хотя бы одна компетенция не сформирована.

Итоговой оценкой результатов промежуточной аттестации выставляется оценка за дифзачет.