

**Федеральное государственное бюджетное учреждение науки
Институт машиноведения им. А.А. Благонравова
Российской академии наук
(ИМАШ РАН)**

УТВЕРЖДАЮ

Заместитель директора по вопросам
подготовки кадров высшей
квалификации ИМАШ РАН

к.ф-м.н. Жданов Э.Р.

« 01 » сентября 2026 года
м.п.



**Рабочая программа дисциплины
«Наука о данных для медицины и биотехнологий»**

Направление подготовки:	02.04.03 Математическое обеспечение и администрирование информационных систем
Профиль подготовки:	Искусственный интеллект и интеллектуальная обработка данных
Квалификация:	Магистр
Форма обучения	Очная

Москва, 2026 г.

1. Аннотация к дисциплине

1.1.	Составлена в соответствии с ФГОС ВО - магистратура по направлению подготовки 02.04.03 Математическое обеспечение и администрирование информационных систем, утвержденным приказом Минобрнауки от 23.08.2017г. № 812.
1.2.	Дисциплина направлена на формирование у обучающихся системных представлений о применении методов искусственного интеллекта, машинного обучения и анализа больших данных для решения прикладных задач в сфере здравоохранения, фармацевтики и биотехнологического производства.
1.3.	В рамках дисциплины рассматриваются особенности медицинских данных, методы их предобработки, алгоритмы построения прогностических моделей и биоинформатического анализа. Особое внимание уделяется практическому применению полученных знаний для разработки интеллектуальных систем поддержки принятия врачебных решений и оптимизации биотехнологических процессов.

2. Место дисциплины в структуре ООП

2.1.	Цикл (раздел) ООП: ФТД.2
2.2.	1 курс, 2 семестр

3. Цель освоения дисциплин

3.1.	Формирование у обучающихся необходимых компетенций для успешного освоения образовательной программы
------	---

4. Задачи освоения дисциплины

4.1.	<ul style="list-style-type: none">- сформировать знания о специфике медицинских и биологических данных (структура, источники, проблемы качества и конфиденциальности);- сформировать знания о ключевых методах машинного обучения (классификация, регрессия, кластеризация) и их применимости в задачах диагностики, прогнозирования заболеваний и анализа геномных данных;- сформировать знания об архитектуре информационных систем в медицине (МИС, PACS) и стандартах обмена медицинскими данными (HL7, DICOM);- сформировать знания о этических и правовых аспектах обработки персональных медицинских данных;- сформировать навыки предобработки и очистки разнородных биомедицинских данных;- сформировать навыки визуализации медицинских данных для выявления скрытых закономерностей;- сформировать навыки работы с библиотеками машинного обучения (например, Scikit-learn) для решения прикладных задач на языке Python;- сформировать умение формулировать постановку задачи анализа данных в медицинской или биотехнологической области;- сформировать умение выбирать адекватный математический аппарат и алгоритм для решения поставленной задачи;- сформировать умение интерпретировать результаты моделирования и оценивать их клиническую или технологическую значимость.
------	---

5. Перечень планируемых результатов обучения, соотнесенных с планируемыми результатами освоения образовательной программы

ПК-1	Создание информационных технологий нового поколения
ПК-1.2	Управление разработкой продуктов, услуг и решений на основе больших данных
ПК-1.3	Разработка и внедрение новых методов и технологий исследования больших данных

Код индикатора компетенции	Планируемые результаты обучения по дисциплине
----------------------------	---

ПК-1.2	<p>Знать:</p> <ul style="list-style-type: none"> - жизненный цикл разработки продуктов на основе анализа медицинских данных методологии управления проектами в области Data Science (CRISP-DM); - требования к качеству и безопасности медицинских информационных систем. <p>Уметь:</p> <ul style="list-style-type: none"> - декомпозировать сложную бизнес-задачу медицинского учреждения на подзадачи анализа данных; формировать техническое задание для команды разработчиков на сбор и обработку данных; оценивать риски и ресурсы при реализации проектов с использованием больших данных. <p>Иметь практический опыт:</p> <ul style="list-style-type: none"> - участия в планировании проекта по созданию прототипа системы поддержки принятия врачебных решений; управления требованиями к данным на всех этапах их жизненного цикла в рамках учебного проекта.
ПК-1.3	<p>Знать:</p> <ul style="list-style-type: none"> - современные методы машинного обучения и глубокого обучения, применяемые в биомедицине; алгоритмы оптимизации гиперпараметров моделей; - подходы к интерпретации результатов работы сложных моделей ("Explainable AI"). <p>Уметь:</p> <ul style="list-style-type: none"> - адаптировать существующие алгоритмы машинного обучения под специфику биомедицинских задач; разрабатывать пайплайны обработки данных от их загрузки до получения финального прогноза; сравнивать эффективность различных моделей и выбирать оптимальную для конкретной задачи. <p>Иметь практический опыт:</p> <ul style="list-style-type: none"> - самостоятельной разработки и тестирования нового метода или модификации существующего алгоритма для решения прикладной задачи; внедрения разработанного решения в виде программного модуля или скрипта и оценки его производительности на реальных или синтетических данных.

6. Тематический план

Наименование темы/раздела	Виды занятий и трудоемкость (в часах)					Формы контроля Форма ПА
	Всего	Л	ПЗ	ЛЗ	СР	
Тема 1. Введение в науку о данных в медицине и биотехнологиях	9	1	1	-	7	Отчет по практикуму/25
Тема 2. Математические основы и методы машинного обучения	9	1	1	-	7	Отчет по практикуму/25
Тема 3. Анализ изображений и сигналов в медицине	9	1	1	-	7	Отчет по практикуму/25
Тема 4. Биоинформатика и оптимизация биотехнологических процессов	9	1	1	-	7	Отчет по практикуму/25
Всего часов	36	4	4	-	28	100 ТКУ
Контроль, час	-					Зачет
Объем дисциплины (в академических часах)	36					
Объем дисциплины (в зачетных единицах)	1					

7. Содержание дисциплины

Наименование темы/раздела	Краткое содержание	Лекции (часов)
Тема 1. Введение в науку о данных в медицине и биотехнологиях	Обзор предметной области. Источники данных: электронные медицинские карты (ЭМК), данные с носимых устройств, геномные и протеомные данные. Проблемы качества данных (пропуски, шум).	1

	Этические нормы и законодательство (152-ФЗ). Анализ структуры реального датасета ЭМК. Выявление пропусков и аномалий. Визуализация распределения показателей здоровья.	
Тема 2. Математические основы и методы машинного обучения	Линейные модели для классификации (логистическая регрессия) и регрессии. Методы обучения без учителя: кластеризация (K-means) для сегментации пациентов. Оценка качества моделей в медицине (ROC-AUC, чувствительность, специфичность). Проблема переобучения. Построение модели для прогнозирования риска развития заболевания на основе анамнеза пациента. Оценка метрик качества модели на отложенной выборке.	1
Тема 3. Анализ изображений и сигналов в медицине	Основы компьютерного зрения: работа с медицинскими изображениями (рентген, МРТ). Сверточные нейронные сети (CNN) для задач детекции патологий. Анализ временных рядов: обработка сигналов ЭКГ, ЭЭГ. Разработка прототипа системы для автоматической детекции патологий на рентгеновских снимках с использованием предобученной нейронной сети.	1
Тема 4. Биоинформатика и оптимизация биотехнологических процессов	Введение в биоинформатику: анализ геномных последовательностей. Применение методов анализа данных для оптимизации процессов биосинтеза (например, в ферментере). Моделирование кинетики роста клеточных культур. Анализ экспрессии генов с целью поиска биомаркеров заболевания. Построение модели для предсказания выхода целевого продукта в биотехнологическом процессе.	1

8. Самостоятельная работа студентов

Наименование темы/раздела	Виды самостоятельной работы	Трудоемкость в часах
Тема 1. Введение в науку о данных в медицине и биотехнологиях	Работа с литературой, ЭБС, источниками в сети Internet, подготовка заданий для сдачи текущего контроля успеваемости	7
Тема 2. Математические основы и методы машинного обучения	Работа с литературой, ЭБС, источниками в сети Internet, подготовка заданий для сдачи текущего контроля успеваемости	7
Тема 3. Анализ изображений и сигналов в медицине	Работа с литературой, ЭБС, источниками в сети Internet, подготовка заданий для сдачи текущего контроля успеваемости	7
Тема 4. Биоинформатика и оптимизация биотехнологических процессов	Работа с литературой, ЭБС, источниками в сети Internet, подготовка заданий для сдачи текущего контроля успеваемости	7

9. Тематика практических и лабораторных занятий

Наименование темы/раздела	Вид занятия	Краткое содержание занятия	Объем часов
Тема 1. Введение в науку о данных в медицине и биотехнологиях	лабораторный практикум	Выполнение заданий в лабораторных условиях. Первичный анализ и визуализация данных электронных медицинских карт (ЭМК)	1
Тема 2. Математические основы и методы машинного обучения	лабораторный практикум	Выполнение заданий в лабораторных условиях. Построение и оценка модели для прогнозирования риска заболевания	1

Тема 3. Анализ изображений и сигналов в медицине	лабораторный практикум	Выполнение заданий в лабораторных условиях. Детекция патологий на медицинских изображениях с помощью нейронных сетей	1
Тема 4. Биоинформатика и оптимизация биотехнологических процессов	лабораторный практикум	Выполнение заданий в лабораторных условиях. Поиск биомаркеров по данным экспрессии генов и предсказание выхода продукта	1

10. Фонд оценочных средств

12.1. Контрольные вопросы и задания для проведения текущего контроля
В приложении
12.2. Фонд оценочных средств для проведения промежуточной аттестации
В приложении
12.3. Порядок начисления баллов и критерии оценивания
В приложении

11. Учебно-методическое и информационное обеспечение дисциплины

9.1. Основная литература			
Авторы	Заглавие	Издательство, год	Эл. адрес
	Биотехнология: учебник и практикум для вузов / под редакцией Н. В. Загоскиной, Л. В. Назаренко. — 4-е изд., испр. и доп.	Москва: Издательство Юрайт, 2025	https://urait.ru/bcode/567470
Вадутов, О. С.	Электроника. Математические основы обработки сигналов: учебник и практикум для вузов	Москва: Издательство Юрайт, 2024	https://urait.ru/bcode/537281
Новиков, Ф. А.	Символический искусственный интеллект: математические основы представления знаний: учебник для вузов	Москва: Издательство Юрайт, 2025	https://urait.ru/bcode/584114
Стефанов, В. Е.	Биоинформатика: учебник для вузов / В. Е. Стефанов, А. А. Тулуб, Г. Р. Мавропуло-Столяренко.	Москва: Издательство Юрайт, 2026	https://urait.ru/bcode/583663
9.2. Дополнительная литература			
Авторы	Заглавие	Издательство, год	Эл. адрес
	Анализ данных: учебник для вузов / под редакцией В. С. Мхитаряна	Москва: Издательство Юрайт, 2026	https://urait.ru/bcode/583032
Миркин, Б. Г.	Базовые методы анализа данных: учебник и практикум для вузов / Б. Г. Миркин. - 3-е изд., перераб. и доп.	Москва: Издательство Юрайт, 2026	https://urait.ru/bcode/583143

12. Материально-техническое обеспечение дисциплины

Учебные помещения	кабинеты оснащены мебелью и оборудованием, соответствующими санитарным нормам и требованиям техники безопасности.
Информационно-технологическая инфраструктура	- серверы и хостинг: наличие сервера или облачного хранилища для размещения учебных ресурсов и платформ электронного обучения; - интернет-канал: стабильный широкополосный доступ к интернету для бесперебойной передачи больших объемов данных.
Организационное обеспечение	- технические средства обучения: компьютеры, ноутбуки, планшеты и мобильные устройства;

	<ul style="list-style-type: none"> - программы для видеоконференцсвязи для проведения лекций, консультаций и практических занятий в режиме реального времени; - антивирусное ПО и защита данных для предотвращения утечек конфиденциальной информации и защиты пользователей от киберугроз.
Условия взаимодействия участников образовательного процесса	<ul style="list-style-type: none"> - дистанционная поддержка преподавателей: помощь техническими специалистами настройки оборудования и программного обеспечения. - поддержка слушателя: организация технической поддержки, включая инструкции по установке и настройке необходимых приложений, устранение возникающих проблем.