

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет компьютерных технологий и прикладной математики

УТВЕРЖДАЮ

Проректор по учебной работе,
качеству образования – первый
проректор

Хагуров Т.А.

подпись

«30» мая 2025

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

К.М.02.04«Машинное обучение»

Направление подготовки 02.03.03 Математическое обеспечение и
администрирование информационных систем

Направленность (профиль) Технологии разработки программных систем

Форма обучения очная

Квалификация бакалавр

Краснодар 2025

Рабочая программа дисциплины «Машинное обучение» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем.

Программу составил(и):

Харченко Анна Владимировна, доцент, канд. пед наук
И.О. Фамилия, должность, ученая степень, ученое звание


подпись

Рабочая программа дисциплины «Машинное обучение» утверждена на заседании кафедры информационных технологий протокол №15 от «14» мая 2025г.

Заведующий кафедрой (разработчика)
В. В. Подколзин


подпись

Рабочая программа обсуждена на заседании кафедры информационных технологий протокол №15 от «14» мая 2025г.

Заведующий кафедрой (выпускающей)
В. В. Подколзин


подпись

Утверждена на заседании учебно-методической комиссии факультета компьютерных технологий и прикладной математики протокол №4 от «23» мая 2025 г.

Председатель УМК факультета
А. В. Коваленко


подпись

Рецензенты:

Бегларян М. Е., Проректор по учебной работе, Краснодарский кооперативный институт (филиал) АНО ВО Центросоюза РФ «Российский университет кооперации»

Рубцов Сергей Евгеньевич, кандидат физико-математических наук, доцент кафедры математического моделирования ФГБОУ ВО «КубГУ»

1 Цели и задачи изучения дисциплины (модуля)

1.1 Цель освоения дисциплины

Целью освоения дисциплины является формирование у студентов фундаментальных теоретических знаний и практических компетенций в области машинного обучения, включая овладение современными методами, алгоритмами и инструментами анализа данных. Курс направлен на подготовку специалистов, способных разрабатывать, реализовывать и оптимизировать математические модели для решения прикладных задач, а также интерпретировать и визуализировать результаты их работы.

1.2 Задачи дисциплины

- освоение ключевых концепций, областей применения и ограничений методов машинного обучения;
- изучение методологических основ выбора, настройки и оценки алгоритмов в зависимости от типа решаемой задачи;
- формирование навыков предобработки данных, включая очистку, нормализацию и преобразование признаков;
- развитие умений применять методы машинного обучения для решения задач классификации, регрессии и кластеризации;
- приобретение опыта визуализации данных и интерпретации результатов работы моделей;
- освоение инструментов оценки качества моделей и методов предотвращения переобучения;
- практическое применение библиотек машинного обучения для реализации и тестирования алгоритмов.

1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Машинное обучение» относится к «К.М.Комплексные модули» Блока 1 «Дисциплины (модули)» учебного плана.

1.4 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

ПК-2 **Способен выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности в области моделирования и анализа сложных естественных и искусственных систем, определять структуру программного обеспечения, методы и средства его проектирования на основе требований с учетом существующих ограничений**

ИД-1.ПК-2 *Разрабатывает концепцию и архитектуру программной системы, ее функциональные возможности и логику работы, делает выбор средств проектирования и реализации на основе требований с учетом существующих ограничений*

Знать *Методологии и технологии проектирования и использования баз данных
Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения
Методы и средства проектирования программного обеспечения
Инструменты и методы проектирования и дизайна ИС
Инструменты и методы верификации структуры программного кода*

Методы и средства планирования и организации исследований и разработок

Уметь *Вырабатывать варианты реализации требований
Проводить оценку и обоснование рекомендуемых решений
Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов*

Владеть *Согласование требований к программному обеспечению с заинтересованными сторонами
Разработка, изменение и согласование архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения
Разработка структуры программного кода ИС
Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач*

ИД-2.ПК-2 ***Способен использовать знания о базовых принципах организации и основных этапах проектирования ИС***

Знать *Методологии и технологии проектирования и использования баз данных
Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения
Методы и средства проектирования программного обеспечения
Методы и средства проектирования программных интерфейсов
Инструменты и методы проектирования и дизайна ИС
Инструменты и методы верификации структуры программного кода
Цели и задачи проводимых исследований и разработок
Методы и средства планирования и организации исследований и разработок*

Уметь *Проводить оценку и обоснование рекомендуемых решений
Использовать существующие типовые решения и шаблоны проектирования программного обеспечения
Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов*

Владеть *Оценка времени и трудоемкости реализации требований к программному обеспечению
Согласование требований к программному обеспечению с заинтересованными сторонами
Разработка, изменение и согласование архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения
Проектирование программных интерфейсов
Разработка структуры программного кода ИС
Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач*

ИД-3.ПК-2 ***Использует методологии разработки программного обеспечения и технологии программирования, методологии и технологии проектирования и использования баз данных, методы и средства***

проектирования программных интерфейсов, принципы построения архитектуры программного обеспечения

Знать	Возможности существующей программно-технической архитектуры Методологии и технологии проектирования и использования баз данных Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения Методы и средства проектирования программного обеспечения Методы и средства проектирования баз данных Методы и средства проектирования программных интерфейсов Инструменты и методы проектирования и дизайна ИС Инструменты и методы верификации структуры программного кода Методы и средства планирования и организации исследований и разработок
Уметь	Вырабатывать варианты реализации требований Использовать существующие типовые решения и шаблоны проектирования программного обеспечения Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов
Владеть	Разработка, изменение и согласование архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения Проектирование программных интерфейсов Разработка структуры программного кода ИС Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач

ПК-3 **Способен применять современные информационные технологии при проектировании, реализации, оценке качества и анализа эффективности математических моделей и(или) программного обеспечения для решения задач в различных предметных областях**

ИД-1.ПК-3 **Использует современные решения и технологии проектирования при разработке программного обеспечения**

Знать	Возможности современных и перспективных средств разработки программных продуктов, технических средств Методологии разработки программного обеспечения и технологии программирования Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения Методы и средства проектирования программных интерфейсов Языки программирования и работы с базами данных Инструменты и методы проектирования и дизайна ИС Современные объектно-ориентированные языки программирования Современные структурные языки программирования Методы анализа и обобщения отечественного и международного опыта в соответствующей области исследований
Уметь	Вырабатывать варианты реализации требований Кодировать на языках программирования Верифицировать структуру программного кода

Владеть	Разработка, изменение и согласование архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения Проектирование программных интерфейсов Разработка структуры программного кода ИС Верификация структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС Устранение обнаруженных несоответствий
ИД-2.ПК-3	Использует современные языки и системы программирования, технологии проектирования программного обеспечения
Знать	Возможности современных и перспективных средств разработки программных продуктов, технических средств Методологии разработки программного обеспечения и технологии программирования Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения Методы и средства проектирования программных интерфейсов Языки программирования и работы с базами данных Инструменты и методы проектирования и дизайна ИС Современные объектно-ориентированные языки программирования Современные структурные языки программирования Методы анализа и обобщения отечественного и международного опыта в соответствующей области исследований
Уметь	Вырабатывать варианты реализации требований Кодировать на языках программирования Верифицировать структуру программного кода
Владеть	Разработка, изменение и согласование архитектуры программного обеспечения с системным аналитиком и архитектором программного обеспечения Проектирование программных интерфейсов Верификация структуры программного кода ИС относительно архитектуры ИС и требований заказчика к ИС Устранение обнаруженных несоответствий
ПК-4	Способен использовать знания современных программных средств, тенденции развития функций и архитектур проблемно-ориентированных программных систем и комплексов в профессиональной деятельности
ИД-1.ПК-4	Проводит классификацию и осуществляет выбор современных инструментальных средств разработки прикладного программного обеспечения вычислительных средств и систем различного функционального назначения, с учетом тенденций развития функций и архитектур в соответствующих проблемно-ориентированных системах и комплексах
Знать	Возможности современных и перспективных средств разработки программных продуктов, технических средств Современные структурные языки программирования
Уметь	Проводить анализ исполнения требований

*Вырабатывать варианты реализации требований
Проводить оценку и обоснование рекомендуемых решений*

Владеть *Сбор, обработка, анализ и обобщение передового отечественного и международного опыта в соответствующей области исследований
Внедрение результатов исследований и разработок в соответствии с установленными полномочиями*

ИД-2.ПК-4 *Реализует приемы работы с современными инструментальными средствами, поддерживающими создание программных проблемно-ориентированных продуктов*

Знать *Возможности современных и перспективных средств разработки программных продуктов, технических средств
Современные структурные языки программирования*

Уметь

Владеть *Устранение обнаруженных несоответствий
Внедрение результатов исследований и разработок в соответствии с установленными полномочиями*

ПК-6 *Способен использовать современные методы разработки программных систем и реализации конкретных алгоритмов математических моделей на базе языков программирования, баз данных и пакетов прикладных программ*

ИД-1.ПК-6 *Использует современные инструментальные средства разработки баз данных, прикладного программного обеспечения и систем различного функционального назначения*

Знать *Возможности современных и перспективных средств разработки программных продуктов, технических средств
Методологии и технологии проектирования и использования баз данных
Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения
Методы и средства проектирования баз данных
Методы и средства проектирования программных интерфейсов
Языки программирования и работы с базами данных
Инструменты и методы проектирования и дизайна ИС
Основы современных систем управления базами данных
Современные объектно-ориентированные языки программирования
Методы анализа и обобщения отечественного и международного опыта в соответствующей области исследований*

Уметь *Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов
Кодировать на языках программирования*

Владеть *Проектирование баз данных
Проектирование программных интерфейсов
Устранение обнаруженных несоответствий
Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач*

ИД-2.ПК-6 *Демонстрирует знания методов, технологий и средств разработки программных систем и реализации конкретных алгоритмов математических моделей на базе языков программирования, баз данных и пакетов прикладных программ*

Знать *Возможности современных и перспективных средств разработки программных продуктов, технических средств
Методологии и технологии проектирования и использования баз данных
Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения
Методы и средства проектирования баз данных
Методы и средства проектирования программных интерфейсов
Языки программирования и работы с базами данных
Инструменты и методы проектирования и дизайна ИС
Основы современных систем управления базами данных
Современные объектно-ориентированные языки программирования
Методы анализа и обобщения отечественного и международного опыта в соответствующей области исследований*

Уметь *Использовать существующие типовые решения и шаблоны проектирования программного обеспечения
Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов
Кодировать на языках программирования*

Владеть *Анализ возможностей реализации требований к программному обеспечению
Оценка и согласование сроков выполнения поставленных задач
Проектирование баз данных
Оценка и согласование сроков выполнения поставленных задач
Устранение обнаруженных несоответствий
Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач*

ИД-3.ПК-6 *Применяет современные приемы работы с инструментальными средствами, поддерживающими создание программных продуктов и программных комплексов на базе языков программирования, баз данных и пакетов прикладных программ*

Знать *Возможности современных и перспективных средств разработки программных продуктов, технических средств
Методологии и технологии проектирования и использования баз данных
Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения
Методы и средства проектирования баз данных
Методы и средства проектирования программных интерфейсов
Языки программирования и работы с базами данных
Инструменты и методы проектирования и дизайна ИС
Основы современных систем управления базами данных
Современные объектно-ориентированные языки программирования
Методы анализа и обобщения отечественного и международного опыта в соответствующей области исследований*

<i>Уметь</i>	<i>Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов Кодировать на языках программирования</i>
<i>Владеть</i>	<i>Проектирование баз данных Устранение обнаруженных несоответствий Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач</i>
ПК-7	Способность использовать знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей образовательных организаций высшего образования, инженеров, технологов
<i>ИД-1.ПК-7</i>	<i>Использует современные инструментальные средства и методы искусственного интеллекта при разработке баз данных, прикладного программного обеспечения и систем различного функционального назначения</i>
<i>Знать</i>	<i>Основы алгоритмизации и программирования Стандарты оформления кода для используемых языков программирования Тестирование ПО, ориентированное на дефекты Техники тестирования ПО, ориентированные на код Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке компьютерного программного обеспечения</i>
<i>Уметь</i>	<i>Использовать существующие типовые решения и шаблоны проектирования компьютерного программного обеспечения Применять методы и средства проектирования компьютерного программного обеспечения, структур данных, баз данных, программных интерфейсов</i>
<i>Владеть</i>	<i>Проектирование баз данных Проектирование структур данных Разработка, изменение архитектуры компьютерного программного обеспечения и ее согласование с системным аналитиком и архитектором программного обеспечения</i>
<i>ИД-2.ПК-7</i>	<i>Использует современные инструментальные средства и методы искусственного интеллекта для сбора, анализа и представления информации</i>
<i>Знать</i>	<i>Тестирование ПО, базирующееся на надежности инженерного процесса Техники тестирования ПО, базирующиеся на интуиции и опыте инженера Техники тестирования ПО, базирующиеся на природе приложения Техники тестирования ПО, базирующиеся на спецификации Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке компьютерного программного обеспечения</i>
<i>Уметь</i>	<i>Составлять отчет о проведении тестирования ПО по разработанным тестовым случаям</i>

Владеть *Проектирование баз данных*
Проектирование программных интерфейсов
Проектирование структур данных
Разработка, изменение архитектуры компьютерного программного обеспечения и ее согласование с системным аналитиком и архитектором программного обеспечения

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 2 зач. ед. (72 часов), их распределение по видам работ представлено в таблице

Вид учебной работы	Всего часов	Семестры (часы)					
		7					
Контактная работа, в том числе:	52,2	52,2					
Аудиторные занятия (всего):	50	50					
Занятия лекционного типа	16	16					
Лабораторные занятия	34	34					
Занятия семинарского типа (семинары, практические занятия)							
Иная контактная работа:	2,2	2,2					
Контроль самостоятельной работы (КСР)	2	2					
Промежуточная аттестация (ИКР)	0,2	0,2					
Самостоятельная работа, в том числе:	19,8	19,8					
<i>Проработка учебного (теоретического) материала</i>	6	6					
<i>Выполнение индивидуальных заданий (подготовка сообщений, презентаций)</i>	8	8					
Подготовка к текущему контролю	5,8	5,8					
Контроль:							
Подготовка к экзамену							
Общая трудоемкость	час.	72	72				
	в том числе контактная работа	52,2	52,2				
	зач. ед	2	2				

2.2 Структура дисциплины

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.
 Разделы (темы) дисциплины, изучаемые в 7 семестре

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа СРС
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Введение в машинное обучение	10	2		6	2
2.	Линейные модели регрессии	10	2		6	2
3.	Логистическая регрессия	8	2		4	2

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
4.	Классификация	10	2		6	2
5.	Кластеризация	8	2		4	4
6.	Деревья решений	9,8	2		4	3,8
7.	Ансамблевые методы	12	4		4	4
ИТОГО по разделам дисциплины		69,8	16		34	19,8
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0,2				
Подготовка к текущему контролю						
Общая трудоемкость по дисциплине		72				

Примечание: Л – лекции, ПЗ – практические занятия/семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов (тем) дисциплины

2.3.1 Занятия лекционного типа

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
1.	Введение в машинное обучение	Введение в предметную область. Примеры использования методов машинного обучения для решения прикладных задач. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и анализа данных. NumPy, SciPy, pandas. Знакомство с различными методами предобработки данных, описательными статистиками и основными способами визуализации данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработка данных. Работа с пропущенными значениями.	К, Т
2.	Линейные модели регрессии	Постановка задачи регрессии. Линейный регрессионный анализ. Отбор признаков, коллинеарность, влияющие наблюдения, анализ остатков. Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества.	К, Т
3.	Логистическая регрессия	Разделение данных на обучающие и тестовые. Нормировка данных. Определение переобученности модели. Критерии оценки качества полученных моделей.	К, Т
4.	Классификация	Постановка задачи классификации, обзор основных методов ее решения. Бинарная и многоклассовая классификация. Логистическая регрессия. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой).	К, Т
5.	Кластеризация	Обучение на неразмеченных данных. Кластеризация. Иерархическая кластеризация. Метод К-средних, DBSCAN и др. Обзор методов	К, Т

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
		кластеризации, реализованных в библиотеке sklearn..	
6.	Деревья решений	Структура деревьев решений. Виды разделяющих функций. Обучения дерева решений. Алгоритм Random Forest.	К, Т
7.	Ансамблевые методы	Ансамбли алгоритмов машинного обучения. Агрегирование моделей	К, Т
8.	Ансамблевые методы	. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг.	К, Т

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.2 Занятия семинарского типа

Не предусмотрены

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.3 Лабораторные занятия

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
1.	Введение в машинное обучение	Примеры использования методов машинного обучения для практических задач. Краткий обзор синтаксиса языка Python. Встроенные операции и функции, типы и структуры данных.	РЗ
2.	Введение в машинное обучение	Библиотеки NumPy и SciPy. Матрицы. Разреженные матрицы. Индексирование, срезы. Объединение массивов. Библиотека pandas. Запросы к таблицам: выборка строк/столбцов по заданным критериям. Модификация элементов таблицы. Добавление строк/столбцов.	РЗ
3.	Введение в машинное обучение	Группировка и агрегирование. Объединение таблиц (различные виды join). Многомерные данные: мультииндексы. Операции stack-unstack. Построение сводных таблиц (pivottables).	РЗ
4.	Линейные модели регрессии	Описательные статистики. Обзор библиотек matplotlib, seaborn, bokeh. Базовые типы визуализации данных. Знакомство с библиотекой scikit-learn (sklearn).	РЗ
5.	Линейные модели регрессии	Предобработка данных. Метод главных компонент. Работа с пропущенными значениями	РЗ
6.	Линейные модели регрессии	Полиномиальная регрессия. Смещение и дисперсия. Гребневая регрессия.	РЗ
7.	Логистическая регрессия	Целевая функция логистической регрессии. Регуляризация логистической регрессии.	РЗ
8.	Логистическая регрессия	Определения целевой метрики качества. Сравнение различных метрик качества моделей. Работа с несбалансированными наборами	РЗ
9.	Классификация	Реализация регрессионных и классификационных моделей с помощью sklearn.	РЗ
10.	Классификация	Работа с синтетическими данными. Реализация метода градиентного спуска.	РЗ
11.	Классификация	Реализация моделей на основе метода k-ближайших соседей метода опорных векторов	РЗ
12.	Деревья решений	Правила и анализ ROC кривой. Точность, полнота). Анализ с помощью ROC кривой. Алгоритм построения деревьев решений.	РЗ

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
13.	Деревья решений	. Критерий информационного выигрыша и критерий Джини. Леса решающих деревьев.	РЗ
14.	Кластеризация	Алгоритмы кластеризации с фиксированным количеством кластеров.	РЗ
15.	Кластеризация	. Алгоритмы кластеризации по плотности. Иерархическая кластеризация.	РЗ
16.	Ансамблевые методы	Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Блендинг и стеккинг.	РЗ
17.	Ансамблевые методы	Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Методы отбора признаков. Оптимизация гиперпараметров	РЗ

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.4 Примерная тематика курсовых работ (проектов)

Не предусмотрены

.....

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Изучение теоретического материала	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой информационных технологий, протокол №1 от 30.08.2019
2	Решение задач	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой информационных технологий, протокол №1 от 30.08.2019

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии

В соответствии с требованиями ФГОС в программа дисциплины предусматривает использование в учебном процессе следующих образовательные технологии: чтение лекций с использованием мультимедийных технологий; метод малых групп, разбор практических задач и кейсов.

При обучении используются следующие образовательные технологии:

- Технология коммуникативного обучения – направлена на формирование коммуникативной компетентности студентов, которая является базовой, необходимой для адаптации к современным условиям межкультурной коммуникации.

- Технология разноуровневого (дифференцированного) обучения – предполагает осуществление познавательной деятельности студентов с учётом их индивидуальных способностей, возможностей и интересов, поощряя их реализовывать свой творческий потенциал. Создание и использование диагностических тестов является неотъемлемой частью данной технологии.

- Технология модульного обучения – предусматривает деление содержания дисциплины на достаточно автономные разделы (модули), интегрированные в общий курс.

- Информационно-коммуникационные технологии (ИКТ) – расширяют рамки образовательного процесса, повышая его практическую направленность, способствуют интенсификации самостоятельной работы учащихся и повышению познавательной активности. В рамках ИКТ выделяются 2 вида технологий:

- Технология использования компьютерных программ – позволяет эффективно дополнить процесс обучения языку на всех уровнях.

- Интернет-технологии – предоставляют широкие возможности для поиска информации, разработки научных проектов, ведения научных исследований.

- Технология индивидуализации обучения – помогает реализовывать личностно-ориентированный подход, учитывая индивидуальные особенности и потребности учащихся.

- Проектная технология – ориентирована на моделирование социального взаимодействия учащихся с целью решения задачи, которая определяется в рамках профессиональной подготовки, выделяя ту или иную предметную область.

- Технология обучения в сотрудничестве – реализует идею взаимного обучения, осуществляя как индивидуальную, так и коллективную ответственность за решение учебных задач.

- Игровая технология – позволяет развивать навыки рассмотрения ряда возможных способов решения проблем, активизируя мышление студентов и раскрывая личностный потенциал каждого учащегося.

- Технология развития критического мышления – способствует формированию разносторонней личности, способной критически относиться к информации, умению отбирать информацию для решения поставленной задачи.

Комплексное использование в учебном процессе всех вышеперечисленных технологий стимулируют личностную, интеллектуальную активность, развивают познавательные процессы, способствуют формированию компетенций, которыми должен обладать будущий специалист.

Основные виды интерактивных образовательных технологий включают в себя:

- работа в малых группах (команде) - совместная деятельность студентов в группе под руководством лидера, направленная на решение общей задачи путём творческого сложения результатов индивидуальной работы членов команды с делением полномочий и ответственности;

– проектная технология - индивидуальная или коллективная деятельность по отбору, распределению и систематизации материала по определенной теме, в результате которой составляется проект;

– анализ конкретных ситуаций - анализ реальных проблемных ситуаций, имевших место в соответствующей области профессиональной деятельности, и поиск вариантов лучших решений;

– развитие критического мышления – образовательная деятельность, направленная на развитие у студентов разумного, рефлексивного мышления, способного выдвинуть новые идеи и увидеть новые возможности.

Подход разбора конкретных задач и ситуаций широко используется как преподавателем, так и студентами во время лекций, лабораторных занятий и анализа результатов самостоятельной работы. Это обусловлено тем, что при исследовании и решении каждой конкретной задачи имеется, как правило, несколько методов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

Семестр	Вид занятия	Используемые интерактивные образовательные технологии	количество интерактивных часов
7	ЛР	Практические занятия в режимах взаимодействия «преподаватель – студент» и «студент – студент»	4
Итого			4

Примечание: Л – лекции, ПЗ – практические занятия/семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

Темы, задания и вопросы для самостоятельной работы призваны сформировать навыки поиска информации, умения самостоятельно расширять и углублять знания, полученные в ходе лекционных и практических занятий.

Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4. Оценочные и методические материалы

4.1 Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Машинное обучение».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме тестовых заданий, разноуровневых заданий, и **промежуточной аттестации** в форме вопросов и заданий к зачету.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на зачете;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Контролируемые разделы (темы) дисциплины*	Код контролируемой компетенции (или ее части)	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация
1	Введение в машинное обучение	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 1-5	Теоретический вопрос 1-4

2	Линейные модели регрессии	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 6-8 Типовые контрольные задания 1-2	Теоретический вопрос 5-7
3	Логистическая регрессия	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Тестовые вопросы 9-11 Типовые контрольные задания 3	Теоретический вопрос 8
4	Классификация	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 12-14 Типовые контрольные задания 4-6	Теоретический вопрос 9-13
5	Деревья решений	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 15 Типовые контрольные задания 7	Теоретический вопрос 15-17

6	Кластеризация	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 16 -18 Типовые контрольные задания 8-9	Теоретический вопрос 14
7	Ансамблевые методы	ИД-1.ПК-3 ИД-2.ПК-3 ИД-1.ПК-4 ИД-2.ПК-4 ИД-1.ПК-2 ИД-2.ПК-2 ИД-3.ПК-2 ИД-1.ПК-6 ИД-2.ПК-6 ИД-3.ПК-6 ИД-1.ПК-7 ИД-2.ПК-7	Типовые тестовые задания 19-21 Типовые контрольные задания 10	Теоретический вопрос 18-20

Показатели, критерии и шкала оценки сформированных компетенций

Соответствие **пороговому уровню** освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **удовлетворительно /зачтено**):

ПК-2 **Способен выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности в области моделирования и анализа сложных естественных и искусственных систем, определять структуру программного обеспечения, методы и средства его проектирования на основе требований с учетом существующих ограничений**

ИД-1.ПК-2 *Разрабатывает концепцию и архитектуру программной системы, ее функциональные возможности и логику работы, делает выбор средств проектирования и реализации на основе требований с учетом существующих ограничений*

Знать *Методологии и технологии проектирования и использования баз данных в контексте ML
Принципы построения архитектуры ML-систем и виды архитектур ПО
Методы и средства проектирования ML-решений
Инструменты и методы проектирования информационных систем для анализа данных
Методы верификации структуры программного кода в ML-проектах
Методы планирования и организации исследований в области машинного обучения*

Уметь	Разрабатывать варианты реализации требований к ML-моделям Оценивать и обосновывать выбор алгоритмов и архитектур Применять методы проектирования ПО, структур данных и программных интерфейсов для ML
Владеть	Согласованием требований к ML-системам с заинтересованными сторонами Разработкой и модификацией архитектуры ML-решений совместно с аналитиками и архитекторами Проектированием структуры кода для ML-приложений Решением аналитических задач с выбором оптимальных методов машинного обучения
ИД-2.ПК-2	Способен использовать знания о базовых принципах организации и основных этапах проектирования ИС
Знать	Методологии проектирования баз данных для ML Архитектурные подходы к построению ML-систем Методы проектирования программных интерфейсов для ML-моделей Инструменты верификации кода в ML-разработке Методы организации исследований в области ML
Уметь	Оценивать и обосновывать решения в ML-проектах Использовать типовые шаблоны проектирования для ML
Владеть	Применять методы проектирования структур данных и API для ML Оценкой трудоемкости реализации ML-решений Согласованием требований к ML-ПО Проектированием API для ML-сервисов Решением аналитических задач с применением ML-методов
ИД-3.ПК-2	Использует методологии разработки программного обеспечения и технологии программирования, методологии и технологии проектирования и использования баз данных, методы и средства проектирования программных интерфейсов, принципы построения архитектуры программного обеспечения
Знать	Возможности программно-технической архитектуры для ML Методологии проектирования баз данных для анализа данных Принципы построения ML-архитектур Методы проектирования API для ML-сервисов
Уметь	Разрабатывать варианты реализации ML-моделей Использовать типовые решения в ML-проектах Применять методы проектирования баз данных и API для ML
Владеть	Разработкой и согласованием архитектуры ML-систем Проектированием API для ML-приложений Решением задач аналитического характера с применением ML
ПК-3	Способен применять современные информационные технологии при проектировании, реализации, оценке качества и анализа эффективности математических моделей и(или) программного обеспечения для решения задач в различных предметных областях
ИД-1.ПК-3	Использует современные решения и технологии проектирования при разработке программного обеспечения
Знать	Современные средства разработки ML-продуктов Методологии разработки ML-ПО

Уметь	Принципы построения архитектуры ML-систем
	Языки программирования для ML (Python, R и др.)
Владеть	Разрабатывать варианты реализации ML-моделей
	Кодировать алгоритмы машинного обучения
Владеть	Верифицировать структуру кода ML-приложений
	Проектированием архитектуры ML-систем
Владеть	Разработкой API для ML-сервисов
	Верификацией кода относительно требований
ИД-2.ПК-3	
Использует современные языки и системы программирования, технологии проектирования программного обеспечения	
Знать	Возможности современных ML-ориентированных языков (Python, R, Julia) и сред разработки (Jupyter, Colab, VS Code с ML-расширениями)
	ML-фреймворки (TensorFlow, PyTorch, Keras) и их архитектурные особенности
Уметь	Методологии разработки ML-совместимого ПО (CRISP-DM, MLOps)
	Принципы построения микросервисной архитектуры для ML-решений
Владеть	Выбирать язык/фреймворк под конкретную ML-задачу (например, PyTorch для NLP, TensorFlow для CV)
	Оптимизировать производительность кода (ускорение вычислений с помощью CUDA, векторизации)
Владеть	Разработкой архитектуры ML-приложений (например: сбор данных → предобработка → обучение → сервис)
	Рефакторингом ML-кода для улучшения поддерживаемости
Владеть	Документированием ML-решений (описание архитектуры, метрик, ограничений моделей)
ПК-4	
Способен использовать знания современных программных средств, тенденции развития функций и архитектур проблемно-ориентированных программных систем и комплексов в профессиональной деятельности	
ИД-1.ПК-4	
Проводит классификацию и осуществляет выбор современных инструментальных средств разработки прикладного программного обеспечения вычислительных средств и систем различного функционального назначения, с учетом тенденций развития функций и архитектур в соответствующих проблемно-ориентированных систем и комплексов	
Знать	Возможности современных ML-фреймворков (TensorFlow, PyTorch, Scikit-learn)
	Структурные языки программирования (Python, R), применяемые в ML
Уметь	Анализировать соответствие ML-алгоритмов требованиям задачи
	Разрабатывать варианты реализации ML-моделей
Владеть	Оценивать и обосновывать выбор инструментов
	Анализом передового опыта в ML-разработке
Владеть	Внедрением результатов исследований в ML-проекты
ИД-2.ПК-4	
Реализует приемы работы с современными инструментальными средствами, поддерживающими создание программных проблемно-ориентированных продуктов	
Знать	Современные библиотеки для обработки данных (Pandas, NumPy)
	Языки программирования для ML-разработки

<i>Уметь</i>	<i>Работать с ML-инструментами (Jupyter Notebook, MLflow)</i>
<i>Владеть</i>	<i>Устранением несоответствий в ML-коде</i>
	<i>Внедрением ML-решений в производственные среды</i>

ПК-6 **Способен использовать современные методы разработки программных систем и реализации конкретных алгоритмов математических моделей на базе языков программирования, баз данных и пакетов прикладных программ**

ИД-1.ПК-6 *Использует современные инструментальные средства разработки баз данных, прикладного программного обеспечения и систем различного функционального назначения*

Знать *Современные ML-фреймворки и СУБД для данных*
Шаблоны проектирования для ML (Feature Store, Pipeline)
Методы проектирования API для ML-сервисов

Уметь *Проектировать структуры данных для ML*
Кодировать алгоритмы обучения

Владеть *Проектированием ML-интерфейсов*
Устранением ошибок в ML-моделях

ИД-2.ПК-6 *Демонстрирует знания методов, технологий и средств разработки программных систем и реализации конкретных алгоритмов математических моделей на базе языков программирования, баз данных и пакетов прикладных программ*

Знать *Типовые ML-решения (ансамбли, нейросети)*
Методы проектирования векторных баз данных

Уметь *Использовать шаблоны проектирования для ML*
Применять методы оптимизации моделей

Владеть *Оценкой времени реализации ML-моделей*
Проектированием ML-конвейеров

ИД-3.ПК-6 *Применяет современные приемы работы с инструментальными средствами, поддерживающими создание программных продуктов и программных комплексов на базе языков программирования, баз данных и пакетов прикладных программ*

Знать *MLOps-инструменты (DVC, Kubeflow)*
Методы деплоя ML-моделей

Уметь *Разрабатывать воспроизводимые ML-эксперименты*

Владеть *Устранением переобучения и смещений в данных*

ПК-7 **Способность использовать знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей образовательных организаций высшего образования, инженеров, технологов**

ИД-1.ПК-7 *Использует современные инструментальные средства и методы искусственного интеллекта при разработке баз данных, прикладного программного обеспечения и систем различного функционального назначения*

Знать *Алгоритмы ИИ (обучение с подкреплением, NLP)*
Стандарты тестирования ML-моделей

Уметь *Применять типовые архитектуры нейросетей*

Владеть *Проектированием дата-пайплайнов*

Согласованием архитектуры ML-систем

ИД-2.ПК-7 *Использует современные инструментальные средства и методы искусственного интеллекта для сбора, анализа и представления информации*

Знать *Техники тестирования ML-моделей (A/B-тесты)
Инструменты визуализации (TensorBoard, Weights & Biases)*

Уметь *Составлять отчеты по экспериментам*

Владеть *Проектированием REST API для ML
Оптимизацией гиперпараметров моделей*

Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Типовые тестовые задания

1) Метеоролог Вася хочет построить модель, прогнозирующую температуру воздуха на завтра. К какому типу относится данная задача?

- Кластеризация
- Классификация
- Ранжирование
- Регрессия

2) Доктор Петя хочет построить модель, прогнозирующую возможность развития осложнений у пациента после операции - все ли будет благополучно в течение нескольких следующих месяцев или нет. К какому типу относится данная задача?

- Ранжирование
- Регрессия
- Кластеризация
- Классификация

3) Астроном Витя хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача?

- Классификация
- Регрессия
- Кластеризация
- Ранжирование

4) В задаче какого типа в обучающей выборке для объектов НЕТ ответов?

- Регрессия
- Классификация
- Кластеризация
- Во всех этих задачах в обучающей выборке есть ответы

5) Выберите все верные утверждения:

- Модель машинного обучения, по сути, является отображением пространства ответов в пространство объектов
- Функционал ошибки показывает, насколько плохое качество имеют данные, используемые для решения задачи
- Процесс обучения модели заключается в минимизации функционала ошибки
- Элементами обучающей выборки являются объекты, характеристики которых являются значениями признаков

6) Рассмотрим признак "Образовательная программа" при анализе данных по студентам университета. Этот признак может принимать три значения: "Экономика", "Математика", "Философия". Воспользуемся one-hot кодированием и заменим этот признак на три бинарных, которые будут соответствовать категориям в том порядке, в котором они перечислены выше. Как будет закодирован признак со значением "Философия"?

- (0, 0, 1)
- (0, 1, 0)
- (1, 0, 0)

7) Чему будет равен корень из среднеквадратичной ошибки для набора из 3 наблюдений, где отклонение предсказания линейной регрессии от реальных значений равны: -1, 2, 2?

- 0
- 3
- 2

8) Предположим, что мы строим модель предсказания роста по возрасту и весу человека. Модель с какими коэффициентами вероятнее всего переобучилась?

- (возраст) + 0.5 * (вес)
- 1402325.3 * (возраст) + -1404370.5 (вес)
- (возраст) + 0.33 (вес)

9) Чем стохастический градиентный спуск (SGD) лучше обычного градиентного спуска? Выберите все подходящие ответы.

- Один шаг в SGD быстрее, поэтому в целом этот метод может быстрее выдать решение.
- В SGD гарантируется, что на каждой итерации уменьшается ошибка модели.
- В SGD гарантируется, что будет найден глобальный минимум.
- Один шаг в SGD точнее, чем в обычном градиентном спуске, поэтому требуется меньше шагов для получения решения.

10) Заполните пропуск: «Для выпуклой функции указывает сторону наискорейшего убывания»

- Изохрона
- Линия уровня
- Антиградиент
- Градиент

11) Что может являться критерием останова в градиентном спуске?

- Норма градиента на текущем шаге
- Норма разницы весов на соседних шагах алгоритма
- Сумма элементов вектора градиента
- Число шагов алгоритма

12) Рассмотрим два объекта: у первого отступ линейного классификатора равен -10, у второго -1000.

Выберите верные утверждения про то, как соотносятся эти два объекта:

- Первый объект находится дальше от разделяющей поверхности, чем второй
- Второй объект находится дальше от разделяющей поверхности, чем первый -
- Модель больше уверена в своём ответе на первом объекте, чем на втором
- Модель больше уверена в своём ответе на втором объекте, чем на первом -

13) Рассмотрим пользователя социальной сети как объект в задаче машинного обучения. Что из перечисленного является задачей классификации?

- Предсказание заработной платы пользователя
- Предсказание пола пользователя
- Предсказание профессии пользователя
- Предсказание, какой пост пользователь сделает следующим

14) Чем задача классификации с пересекающимися классами (П) отличается от задачи классификации с непересекающимися классами (Н)?

- В задаче П один объект может относиться к нескольким классам одновременно, а в задаче Н один объект относится ровно к одному классу
- В задаче П один объект может относиться либо к одному классу, либо ни к какому классу, а в задаче Н один объект относится ровно к одному классу
- В задаче П один признак может использоваться для предсказания нескольких классов, а в задаче Н один признак используется для предсказания только одного класса
- В задаче П для объектов обучающей выборки известны метки классов, а в задаче Н - нет

15) Выберите, какой из приведенных алгоритмов предсказания стоимости квартиры является примером решающего дерева:

- Если площадь больше 100 кв. метров, то 4 млн руб., иначе если этаж первый, то 2.3 млн. рублей, иначе 3.15 млн. рублей
- 400 тыс. рублей. за каждый квадратный метр и 27 тыс. руб. за каждый этаж
- Использовать стоимость квартиры, наиболее похожей на текущую, из продаваемых ранее
- Продавать за 3 млн. рублей

16) Метод главных компонент... (один правильный ответ)

- Строит новые признаки как сложные нелинейные функции от исходных признаков
- Строит новые признаки как линейные функции от исходных признаков
- Отбирает самые важные признаки из исходных
- Предсказывает класс объекта на основе главных компонент

17) Выберите верные утверждения про задачу кластеризации:

- В задаче кластеризации сложно придумать универсальный критерий качества, говорящий, хорошая или плохая получилась кластеризация, что усложняет решение этой задачи
- На практике данные, как правило, не разделяются на четкие кластеры, что усложняет решение задачи кластеризации
- Метод k-Means - самый эффективный метод кластеризации, способный находить качественную кластеризацию практически любых данных
- Задача кластеризации предполагает выделение самых важных признаков и удаление всех остальных признаков

18) Какие преимущества дает отбор признаков?

- Чем меньше признаков, тем более сложные зависимости может восстановить алгоритм
- Чем меньше признаков, тем быстрее работает алгоритм
- Отбор признаков уменьшает объем данных и затраты на их хранение
- Отбор неинформативных признаков может повысить качество работы алгоритма

19) Ансамблевые методы обычно делают более качественные предсказания, чем отдельные алгоритмы (у ансамблей ниже ошибка на тестовой выборке). А какой показатель всегда ухудшается при ансамблировании?

- Ошибка на обучающей выборке
- Скорость выполнения предсказаний
- Количество данных, необходимое для обучения
- Число используемых алгоритмом признаков

20) Алгоритм бэггинг подразумевает выбор случайных частей данных, обучение алгоритма (например, решающего дерева) на каждой части и составление ансамбля из обученных алгоритмов. Для чего нужно обучать алгоритмы на разных частях данных?

- Для того, чтобы обученные алгоритмы выполняли несовпадающие (различные) предсказания
- Для того, чтобы ускорить обучение
- Для того, чтобы выделить важные для обучения объекты
- Без этого невозможно обучить решающее дерево

21) Вы обучили три модели для предсказания суммы кредита, которая потребуется клиенту, и выполняете предсказание для нового клиента: линейная модель предсказывает 50 тысяч, решающее дерево - 30 тысяч, метод ближайшего соседа - 20 тысяч. Какое предсказание выполнит ансамбль трех алгоритмов, если вес линейной модели 0.4, вес решающего дерева 0.4, а вес метода ближайшего соседа - 0.2?

- 33.3 тысяч
- 30 тысяч
- 36 тысяч
- 38 тысяч

Типовые контрольные задания

1. Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета noisysine – степеней от 2 до 5, для датасета hydrodynamics – степени 2) без регуляризации.
2. Реализуйте алгоритм гребневой регрессии и найдите оптимальный параметр регуляризации.
3. Реализуйте алгоритм логистической регрессии со стохастическим градиентным спуском, обучите его на датасете spambase_old (train) и проверьте на датасете spambase_new (val). Получите ROC кривые для вариантов без нормировки и с нормировкой признаков.
4. Реализуйте алгоритм kNN классификации по k ближайшим соседям, используя простое евклидовое расстояние.
5. Примените метод SVM (например, из библиотеки sklearn) для датасета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).
6. Реализуйте мультиклассовую классификацию с softmax в качестве решающей функции и кросс-энтропией в качестве функции потерь и обучите на подготовленном датасете mnist.
7. Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10).
8. Реализуйте алгоритм k-means для кластеризации на 2-4 кластера.
9. Реализуйте алгоритм DBSCAN, найдите параметры для кластеризации на 4 кластера.

10. Примените ансамбевые методы для решения задачи классификации на подготовленном датасете mnist.

Зачетно-экзаменационные материалы для промежуточной аттестации (экзамен/зачет)

Вопросы для подготовки к зачету

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
3. Смещение и дисперсия (bias and variance). Понятие средней гипотезы.
4. Ошибка внутри и вне выборки. Ошибка обобщения. Неравенство Хёффдинга. Валидация и кросс-валидация.
5. Линейная регрессия.
6. Полиномиальная регрессия.
7. Гребневая регрессия.
8. Логистическая регрессия. Градиентный спуск.
9. Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). ЕМ алгоритм.
10. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
11. Пороговые условия. Эффективность по Парето.
12. Precision-Recall и ROC кривые. AUC.
13. Метрические классификаторы. kNN. WkNN. Отбор эталонов. DROP5. Kdtree.
14. Кластеризация. kMeans, MeanShift, DBSCAN, Affinity Propagation.
15. Деревья решений. Информационный выигрыш, критерий Джини.
16. Регуляризация деревьев. Небрежные решающие деревья.
17. Бустинг деревьев решений.
18. Ансамблевые методы. Soft and Hard Voting. Bagging.
19. Случайные леса. AdaBoost.
20. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.

4.2 Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Методические рекомендации, определяющие процедуры оценивания тестов:

Тест проводится онлайн в системе Moodle или Google Docs и ограничен по времени. На сдачу теста дается две попытки. Тест считается успешно пройденным если студент правильно ответил на 70% вопросов.

Методические рекомендации, определяющие процедуры оценивания выполнения контрольных заданий:

Задание считается выполненным при выполнении следующих условий:

- предоставлен исходный код в среде PyCharm, Google Collab
- продемонстрирована работоспособность программы
- студент понимает исходный код и отвечает на вопросы по его организации.

Методические рекомендации к сдаче экзамена

Процедура промежуточной аттестации проходит в соответствии с Положением о текущем контроле и промежуточной аттестации обучающихся ФГБОУ ВО «КубГУ».

Текущий контроль успеваемости студентов проводится в целях совершенствования и непрерывного контроля качества образовательного процесса, проверки усвоения учебного материала, активизации самостоятельной работы студентов, стимулирования их учебной работы, обеспечения эффективности образовательного процесса, предупреждения рисков отчисления студентов.

Текущий контроль знаний студентов осуществляется постоянно в течение всего семестра.

Виды текущего контроля: устный (письменный) опрос на занятиях; проверка выполнения домашних заданий; проведение контрольных работ; оценка активности студента на занятии.

При отсутствии зачетной книжки у студента экзаменатор не имеет права принимать у него зачет/экзамен. Такой студент считается не явившимся на зачет/экзамен. В исключительных случаях, на основании распоряжения декана преподаватель может допустить студента к зачету/экзамену при наличии документа, удостоверяющего личность.

В целях объективного оценивания знаний во время проведения зачетов и экзаменов не допускается наличие у студентов посторонних предметов и технических устройств.

Студенту, использующему в ходе зачета/экзамена неразрешенные источники и средства получения информации, выставляется неудовлетворительная оценка, и он удаляется из аудитории.

Во время экзамена студенты могут пользоваться утвержденной рабочей программой учебной дисциплины, которая должна быть в наличии на экзамене, а также с разрешения экзаменатора справочной литературой и другими пособиями.

Студенты, нарушающие правила поведения при проведении зачетов и экзаменов, могут быть незамедлительно удалены из аудитории, к ним могут быть применены меры дисциплинарного воздействия.

На зачете/экзамене могут присутствовать ректор, проректор по учебной работе, декан факультета, заведующий кафедрой, которая обеспечивает учебный процесс по данной дисциплине. Присутствие на экзаменах и зачетах посторонних лиц без разрешения ректора или проректора по учебной работе не допускается.

После прослушивания лекции рекомендуется выполнить упражнения, приводимые в лекции для самостоятельной работы, а также выполнить на компьютере с использованием среды Python задачи, приводимые в лекции в качестве примеров.

При самостоятельной работе студентов необходимо изучить литературу, приведенную в перечнях, для осмысления вводимых понятий, анализа предложенных подходов и методов разработки параллельных программ. Разрабатывая решение новой задачи, студент должен уметь выбрать методы решения задачи с учетом целевой аппаратной платформы, проводить отладку и профилирование программы на языке Python.

В качестве систем программирования для решения задач и изучения методов и алгоритмов, приведенных в лекциях, рекомендуется использовать на практических занятиях и при самостоятельной работе стандартную реализацию языка Python в связке со средой разработки PyCharm, Google Collab.

Методические рекомендации, определяющие процедуры оценивания на зачете:

Критерии оценивания и шкала оценки

Оценка	
Не зачтено	Зачтено
• если студент правильно решил менее 70 % задач и/или не имеет представление как решать остальные задачи • не знает значительной части теоретического материала, допускает существенные ошибки.	• если студент правильно решил 70 % задач, имеет представление как решать остальные задачи • на теоретические вопросы дан развернутый ответ. Материал изложен в целом последовательно. Имеются логичные и аргументированные выводы

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на зачете;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

5. Перечень учебной литературы, информационных ресурсов и технологий

5.1 Основная литература:

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262> (дата обращения: 08.06.2025)

2. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2022. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/469022> (дата обращения: 08.06.2025)

3. Келлехер, Д. Наука о данных: базовый курс : [16+] / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров ; пер. с англ. М. Белоголовского. — Москва : Альпина Паблишер, 2020. — 224 с. : схем., табл. — Режим доступа: по подписке. — URL: <https://biblioclub.ru/index.php?page=book&id=598235> (дата обращения: 08.06.2025). — ISBN 978-5-9614-3170-4. — Текст : электронный.

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечных системах «Лань» и «Юрайт».

5.2 Дополнительная литература:

1. Маккинли У. (пер. с англ. Слинкин А.А.), Python и анализ данных // Издательство “ДМК Пресс”, 2015, 482 с.

2. Мюллер А., Гвидо С. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными // Вильямс, 2017, 480 с.

3. Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python : [16+] / В. М. Волкова, М. А. Семенова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск : Новосибирский государственный технический университет, 2017. — 74 с. : ил., табл. — Режим доступа: по подписке. — URL: <https://biblioclub.ru/index.php?page=book&id=576496> (дата обращения: 08.06.2025). — Библиогр.: с. 48. — ISBN 978-5-7782-3183-2. — Текст : электронный.

5.3. Периодические издания:

1. Базы данных компании «Ист Вью» <http://dlib.eastview.com>

2. Электронная библиотека GREBENNIKON.RU <https://grebennikon.ru/>

5.4. Интернет-ресурсы, в том числе современные профессиональные базы данных и информационные справочные системы

Электронно-библиотечные системы (ЭБС):

1. ЭБС «ЮРАЙТ» <https://urait.ru/>

2. ЭБС «УНИВЕРСИТЕТСКАЯ БИБЛИОТЕКА ОНЛАЙН» <http://www.biblioclub.ru/>

3. ЭБС «BOOK.ru» <https://www.book.ru>

4. ЭБС «ZNANIUM.COM» www.znanium.com

5. ЭБС «ЛАНЬ» <https://e.lanbook.com>

Профессиональные базы данных

1. Scopus <http://www.scopus.com/>

2. ScienceDirect <https://www.sciencedirect.com/>

3. Журналы издательства Wiley <https://onlinelibrary.wiley.com/>

4. Научная электронная библиотека (НЭБ) <http://www.elibrary.ru/>

5. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН <http://archive.neicon.ru>

6. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) <https://rusneb.ru/>

7. Президентская библиотека им. Б.Н. Ельцина <https://www.prilib.ru/>

8. База данных CSD Кембриджского центра кристаллографических данных (CCDC) <https://www.ccdc.cam.ac.uk/structures/>

9. Springer Journals: <https://link.springer.com/>

10. Springer Journals Archive: <https://link.springer.com/>
11. Nature Journals: <https://www.nature.com/>
12. Springer Nature **Protocols and Methods**:
<https://experiments.springernature.com/sources/springer-protocols>
13. Springer Materials: <http://materials.springer.com/>
14. Nano Database: <https://nano.nature.com/>
15. Springer eBooks (i.e. 2020 eBook collections): <https://link.springer.com/>
16. "Лекториум ТВ" <http://www.lektorium.tv/>
17. Университетская информационная система РОССИЯ <http://uisrussia.msu.ru>

Информационные справочные системы

1. Консультант Плюс - справочная правовая система (доступ по локальной сети с компьютеров библиотеки)

Ресурсы свободного доступа

1. КиберЛенинка <http://cyberleninka.ru/>;
2. Американская патентная база данных <http://www.uspto.gov/patft/>
3. Министерство науки и высшего образования Российской Федерации
<https://www.minobrnauki.gov.ru/>;
4. Федеральный портал "Российское образование" <http://www.edu.ru/>;
5. Информационная система "Единое окно доступа к образовательным ресурсам"
<http://window.edu.ru/>;
6. Единая коллекция цифровых образовательных ресурсов <http://school-collection.edu.ru/> .
7. Проект Государственного института русского языка имени А.С. Пушкина
"Образование на русском" <https://pushkininstitute.ru/>;
8. Справочно-информационный портал "Русский язык" <http://gramota.ru/>;
9. Служба тематических толковых словарей <http://www.glossary.ru/>;
10. Словари и энциклопедии <http://dic.academic.ru/>;
11. Образовательный портал "Учеба" <http://www.ucheba.com/>;
12. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--p1ai/voprosy_i_otvety

Собственные электронные образовательные и информационные ресурсы КубГУ

1. Электронный каталог Научной библиотеки КубГУ <http://megapro.kubsu.ru/MegaPro/Web>
2. Электронная библиотека трудов ученых КубГУ
<http://megapro.kubsu.ru/MegaPro/UserEntry?Action=ToDb&idb=6>
3. Среда модульного динамического обучения <http://moodle.kubsu.ru>
4. База учебных планов, учебно-методических комплексов, публикаций и конференций
<http://infoneeds.kubsu.ru/>
5. Библиотека информационных ресурсов кафедры информационных образовательных технологий <http://mschool.kubsu.ru;>
6. Электронный архив документов КубГУ <http://docspace.kubsu.ru/>
7. Электронные образовательные ресурсы кафедры информационных систем и технологий в образовании КубГУ и научно-методического журнала "ШКОЛЬНЫЕ ГОДЫ"
<http://icdau.kubsu.ru/>

5.5 Перечень информационно-коммуникационных технологий

- Проверка домашних заданий и консультирование посредством электронной почты.
- Использование электронных презентаций при проведении лекционных занятий

- Система MOODLE
- Проверка домашних заданий и консультирование посредством ЭОИС КубГУ

5.6 Перечень лицензионного и свободно распространяемого программного обеспечения

OpenOffice
 Компилятор C++
 Oracle VirtualBox 6
 VMware Workstation 16
 Putty 0.76 или Kitty 0.76
 FileZilla 3.57.0
 WinSCP 5.19
 Advanced port scanner 2.5
 Python 3 (3.7 И 3.9)
 numpy 1.22.0
 opencv 4.5.5
 Keras 2.7.0
 Tensor flow 2.7.0
 matplotlib 3.5.1
 PyCharm 2021
 Cuda Toolkit 11.6
 Фреймворк Django
 Firefox, любая версия
 Putty, любая версия
 Visual Studio Code, версия 1.52+
 Eclipse PHP Development Tools, версия 2020-06+
 Плагин Remote System Explorer (RSE) для Eclipse PDT
 JetBrains PHP Storm
 GIT
 Java Version 8 Update 311
 Clojure 1.10.3.1029.ps1
 SWI Prolog 8.4
 IntelliJ Idea IDE 2021
 Mozilla Firefox 96
 Google Chrome 97
 GitHub Desktop 2.9
 PHP Storm 2021
 FileZilla 3.57.0
 Putty 0.76

6. Методические указания для обучающихся по освоению дисциплины (модуля)

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

7. Материально-техническое обеспечение по дисциплине (модулю)

№	Вид работ	Наименование учебной аудитории, ее оснащенность оборудованием и техническими средствами обучения
1.	Лекционные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
2.	Лабораторные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, проектором, программным обеспечением
3.	Групповые (индивидуальные) консультации	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
4.	Текущий контроль, промежуточная аттестация	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
5.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.

Примечание: Конкретизация аудиторий и их оснащение определяется ОПОП.