МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ГЕОДЕЗИИ И КАРТОГРАФИИ
(МИИГАиК)

Картографический факультет
Кафедра высшей математики

**ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ**

Сертификат: 5387FD2AF5B17FED816FDE9A1D4973B1
Владелец: Степанченко Алексей Леонидович
Действителен: с 21.07.2025 по 14.10.2026

«УТВЕРЖДАЮ»
Проректор
А. Л. Степанченко
24 марта 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «Б1.В.05 ОСНОВЫ АНАЛИЗА ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ»

Уровень высшего образования: магистратура

Направление подготовки: 05.04.03 Картография и геоинформатика

Направленность (профиль) подготовки: Картография и геоинформатика

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2025

Срок получения образования: 2 года

Объем: в зачетных единицах: 3 з.е.
в академических часах: 108 ак.ч.

Москва, 2025

Разработчики:
Доцент кафедры Высшей математики, кандидат
физико-математических наук Волковая Т. А.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 05.04.03 Картография и геоинформатика, утвержденного приказом Минобрнауки от 07.08.2020 № 893

Согласование и утверждение

№	Подразделение	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	Картографический факультет	Председатель ученого совета факультета	Загребин Г. И.	Рассмотрено	20.12.2024, № 4/24

Содержание

1. Цель и задачи освоения дисциплины (модуля)

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

3. Место дисциплины в структуре ОП

4. Объем дисциплины и виды учебной работы

5. Содержание дисциплины
   5.1. Разделы, темы дисциплины и виды занятий
   5.2. Содержание разделов, тем дисциплины и формы текущего контроля

6. Оценочные материалы текущего контроля

7. Оценочные материалы промежуточной аттестации

8. Порядок проведения промежуточной аттестации

9. Материально-техническое и учебно-методическое обеспечение дисциплины
   9.1. Перечень основной и дополнительной учебной литературы
   9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся
   9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине
   9.4. Оборудование и технические средства обучения

10. Методические указания по освоению дисциплины (модуля)

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - формирование компетенций, позволяющих составить представление о методах машинного обучения, современных методов восстановления зависимостей по эмпирическим данным.

Задачи изучения дисциплины:

- изучить теоретические основы машинного обучения и алгоритмы, обучаемые на данных;
- научиться выполнять исследовательские циклы анализа данных: выдвижение гипотез, визуализацию, выбор подходящих алгоритмов, обучение моделей, оценку качества и интерпретацию полученных результатов;
- сформировать навыки решения прикладных задач на реальных данных с помощью алгоритмов машинного обучения.

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

ОПК-1 Способен использовать философские концепции и основы методологии научного познания при изучении различных уровней организации материи, пространства и времени при решении задач профессиональной деятельности

ОПК-1.1 Способен решать производственные и (или) исследовательские задачи производственной деятельности на основе фундаментальных знаний в области геодезии и картографии

Знать:

ОПК-1.1/Зн3 Методологические и теоретические основы и инструменты машинного обучения, используемые для решения научных и практических задач

Уметь:

ОПК-1.1/Ум3 Использовать базовые методы машинного обучения для решение практических и (или) научно-исследовательских задач

Владеть:

ОПК-1.1/Нв3 Способностью к применению методов математического и алгоритмического моделирования при решении теоретических и прикладных задач

3. Место дисциплины в структуре ОП

Дисциплина (модуль) Б1.В.05 «Основы анализа данных и машинного обучения» относится к формируемой участниками образовательных отношений части образовательной программы и изучается в семестре(ах): 2.

В процессе изучения дисциплины студент готовится к решению типов задач профессиональной деятельности, предусмотренных ФГОС ВО и образовательной программой.

4. Объем дисциплины (модуля) и виды учебной работы

Период	доемкость сы)	доемкость ЗЕТ)	ые занятия сы)	ие занятия сы)	ьная работа сы)	ая аттестация сы)

обучения	Общая тру (ча	Общая тру (ЗЕ	Лекционни (ча	Практическ (ча	Самостоятел (ча	Промежуточн (ча
2 семестр	108	3	28	28	52	Зачет
Всего	108	3	28	28	52	

5. Содержание дисциплины (модуля)

5.1. Разделы, темы дисциплины и виды занятий

(часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа	Планируемые результаты обучения, соотнесенные с результатами освоения программы
Раздел 1. Основы анализа данных и машинного обучения	**108**	**28**	**28**	**52**	ОПК-1.1
Тема 1.1. Основы машинного обучения. Обучение с учителем и без. Предобработка данных.	32	8	8	16	
Тема 1.2. Обучение с учителем (задача регрессии, задача классификации). Обучение без учителя (задача кластеризации).	40	12	12	16	
Тема 1.3. Нейронные сети	36	8	8	20	
Итого	**108**	**28**	**28**	**52**	

5.2. Содержание разделов, тем дисциплин и формы текущего контроля

Раздел 1. Основы анализа данных и машинного обучения

Тема 1.1. Основы машинного обучения. Обучение с учителем и без. Предобработка данных.

Постановка задачи МО в общем виде. Решаемые задачи: кластеризация, классификация, регрессия. Главные вопросы: задача; метрики качества; валидация; переобучение; обнаружение и обработка выбросов; метрические методы и гипотеза компактности; функции потерь; обучение с учителем и без. Признаки. Типы признаков. Понятие функционала качества. Вероятностная постановка задачи. Оценка обобщающей способности. Проблема переобучения. Критерии оценки качества работы алгоритмов машинного обучения.

Текущий контроль

Вид (форма) контроля
Письменный опрос

Тема 1.2. Обучение с учителем (задача регрессии, задача классификации). Обучение без учителя (задача кластеризации).

Линейная регрессия. Градиентные методы. Линейный классификатор. Логистическая регрессия. Метод опорных векторов. Метод K ближайших соседей. Решающее дерево. Отбор признаков для задач классификации и регрессии. Функции потерь. Виды метрик качества. Ансамблирование (бэггинг, бустинг, стэкинг). Случайный лес. Модели градиентного бустинга (catboost, XGB и др.). Методы кластеризации (K-средних, K-медиан, DBSCAN др.). Разбор основных алгоритмов МО с точки зрения механики их работы и границ применимости.

Текущий контроль

Вид (форма) контроля
Письменный опрос

Тема 1.3. Нейронные сети

Модели искусственного нейрона. Функции активации. Нейрон с векторным входом. Искусственные нейронные сети. Архитектура искусственных нейронных сетей. Набор средств для создания, инициализации, обучения, моделирования и визуализации сети. Методы и алгоритмы обучения искусственных нейронных сетей.

Текущий контроль

Вид (форма) контроля
Письменный опрос

6. Порядок проведения промежуточной аттестации

Порядок проведения промежуточной аттестации регламентируется локальными нормативными актами Университета:

- Положением о формах, периодичности, порядке текущего контроля успеваемости и промежуточной аттестации обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.6 «Организация и проведение промежуточной аттестации».
- Положением о порядке осуществления образовательного процесса, проведения текущего контроля успеваемости и промежуточной аттестации обучающихся с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.8 «Проведение промежуточной аттестации с применением ЭО и ДОТ».

Промежуточная аттестация: Зачет, 2 семестр.

7. Оценочные материалы текущего контроля

Раздел 1. Основы анализа данных и машинного обучения

Контролируемые ИДК: ОПК-1.1

Тема 1.1. Основы машинного обучения. Обучение с учителем и без. Предобработка данных.

Форма контроля/оценочное средство: Письменный опрос

Вопросы/Задания:

1. Назовите четыре основные задачи предобработки данных

2. Опишите задачу очистки данных
3. Опишите задачу интеграции данных
4. Опишите задачу редукции данных
5. Приведите примеры преобразования данных
6. Какой признак является количественным?
7. Какой признак является качественным?
8. Что называют целевой переменной (целевым признаком)?
9. Что называют переобучением?
10. В чем суть кросс-валидации?

Тема 1.2. Обучение с учителем (задача регрессии, задача классификации). Обучение без учителя (задача кластеризации).

Форма контроля/оценочное средство: Письменный опрос

Вопросы/Задания:

1. Сформулируйте постановку задачи обучения с учителем.
2. Сформулируйте постановку задачи кластеризации.
3. Сформулируйте постановку задачи классификации
4. Сформулируйте постановку задачи бинарной классификации
5. Сформулируйте постановку задачи регрессии
6. Сформулируйте постановку задачи линейной регрессии
7. Дайте описание алгоритма градиентного спуска
8. В решении каких задачи используется логистическая регрессия?
9. Дайте формальное описание метода случайных подпространств
10. Опишите алгоритм градиентного бустинга

Тема 1.3. Нейронные сети

Форма контроля/оценочное средство: Письменный опрос

Вопросы/Задания:

1. По какой формуле определяется состояние нейрона?
2. Что вычисляет функция активации нейрона?
3. Приведите примеры известных вам функций активации
4. Охарактеризуйте линейный порог активации
5. Охарактеризуйте пороговую функцию активации
6. Охарактеризуйте сигмоидальную функцию активации
7. Дайте классификация нейронных сетей по характеру обучения
8. Дайте классификация нейронных сетей по типу настройки весов
9. Что понимают под архитектурой нейронной сети?
10. Приведите примеры структур многослойных сетей

8. Оценочные материалы промежуточной аттестации

2 семестр, Зачет

Контролируемые ИДК: ОПК-1.1

Вопросы/Задания:

1. Дана выборка: -33, -28, 4, 9, 14, 19, 24, 30, 40, 50. Определите медиану.

А) 14
б) 19
в) 16, 5
г) 12, 9

2. Дана выборка: -33, -28, 4, 9, 14, 19, 24, 30, 40, 50. Определите среднее значение признака.

А) 14
б) 19
в) 16, 5
г) 12, 9

3. Дана выборка: -33, -28, 4, 9, 14, 19, 24, 30, 40, 50. Определите квантиль уровня 0, 25.

А) 4
б) 9
в) 5, 25
г) 12, 9

4. Дана выборка: -33, -28, 4, 9, 14, 19, 24, 30, 40, 50. Определите квантиль уровня 0, 75.

А) 24
б) 30
в) 28, 5
г) 12, 9

5. Для выборки 22, 18, 15, 24, 31, 25, 28, 1, 61 объема n=9 указать выбросы, если они есть.

А) 1, 61
б) 31, 61
в) только 61
г) выбросов нет

6. Класс …
Заменяет каждое уникальное значение признака на целое неотрицательное число, тем самым задаётся порядок этих значений класс …

А) LabelEncoder
б) OneHotEncoder
в) MinMaxScaler
г) StandardScaler

7. Класс … преобразует один категориальный признак в несколько количественных признаков, каждый из которых несёт информацию о наличии одного из уникальных значений преобразуемого категориального признака. В случае присутствия данного уникального значения соответствующий количественный признак будет иметь значение единица, и ноль — в противном случае.

А) LabelEncoder
б) OneHotEncoder
в) MinMaxScaler
г) StandardScaler

8. Линейное преобразование признака, которое переводит наименьшее значение данного признака в некоторое значение a (по умолчанию 0), а наибольшее – в значение b (по умолчанию 1), осуществляет класс ...

А) LabelEncoder
б) OneHotEncoder
в) MinMaxScaler
г) StandardScaler

9. Класс … стандартизирует данные, то есть осуществляет линейное преобразование данных так, чтобы преобразованные данные имели нулевое среднее и единичное среднеквадратичное отклонение.

А) LabelEncoder
б) OneHotEncoder
в) MinMaxScaler
г) StandardScaler

10. Модель обучается на размеченных данных, где каждому входному примеру соответствует выходной результат. Это

А) обучение без учителя
б) обучение с учителем
в) обучение с подкреплением

11. К какой из задач машинного обучения относится следующая задача: определение рыночной цены недвижимости на основании характеристик (площадь, количество комнат, удаленность от метро и др.)

А) задача регрессии
б) задача классификации
в) задача кластеризации

12. К какой из задач машинного обучения относится следующая задача: пометка электронных писем как «спам» и «не спам»

А) задача регрессии
б) задача классификации
в) задача кластеризации

13. К какой из задач машинного обучения относится следующая задача: выделение групп людей на основе графа связей в социальных сетях

А) задача регрессии
б) задача классификации
в) задача кластеризации

14. Accuracy (общая доля верных предсказаний) вычисляется по формуле:

а) $\frac{TP+TN}{TP+TN+FP+FN}$
б) $\frac{TP}{TP+FP}$
в) $\frac{TP}{TP+FN}$
г) $\frac{TP}{TP+TN}$

15. Precision (среди предсказанных положительными, сколько действительно положительные) вычисляется по формуле:

а) $\frac{TP+TN}{TP+TN+FP+FN}$
б) $\frac{TP}{TP+FP}$
в) $\frac{TP}{TP+FN}$
г) $\frac{TP}{TP+TN}$

16. Recall (среди фактически положительных, сколько предсказаны как положительные) вычисляется по формуле:

а) $\frac{TP+TN}{TP+TN+FP+FN}$
б) $\frac{TP}{TP+FP}$
в) $\frac{TP}{TP+FN}$
г) $\frac{TP}{TP+TN}$

17. Укажите функцию потерь для гребневой (Ridge) регрессии

а) $Loss = MSE + C\sum_{k=1}^{n} a_k^2$

б) $Loss = MSE + C\sum_{k=1}^{n}|a_k|$
в) $Loss = MSE + C_1\sum_{k=1}^{n}|a_k| + C_2\sum_{k=1}^{n}a_k^2$
г) $Loss = MSE$

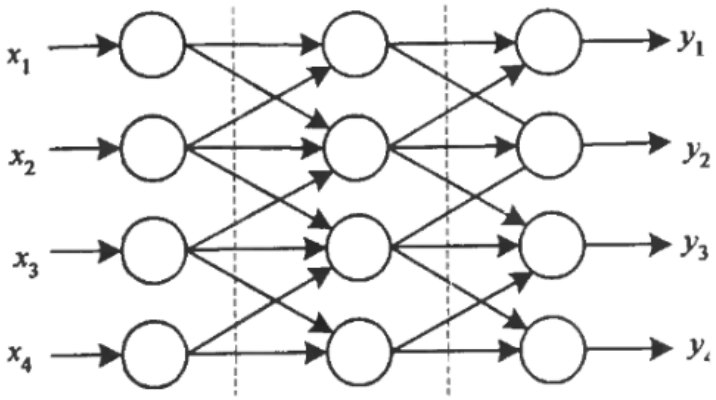
18. Укажите функцию потерь для Lasso регрессии

а) $Loss = MSE + C\sum_{k=1}^{n}a_k^2$
б) $Loss = MSE + C\sum_{k=1}^{n}|a_k|$
в) $Loss = MSE + C_1\sum_{k=1}^{n}|a_k| + C_2\sum_{k=1}^{n}a_k^2$
г) $Loss = MSE$

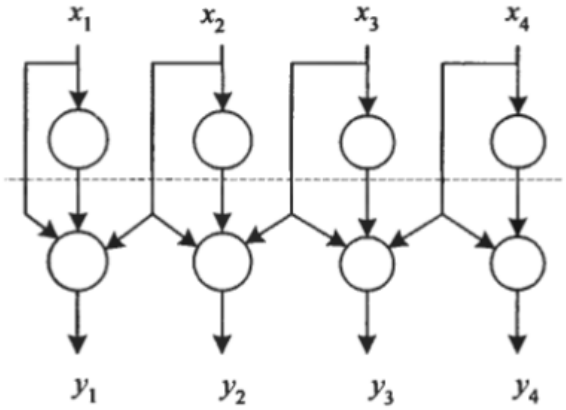
19. Укажите функцию потерь для Elastic-Net

а) $Loss = MSE + C\sum_{k=1}^{n}a_k^2$
б) $Loss = MSE + C\sum_{k=1}^{n}|a_k|$
в) $Loss = MSE + C_1\sum_{k=1}^{n}|a_k| + C_2\sum_{k=1}^{n}a_k^2$
г) $Loss = MSE$

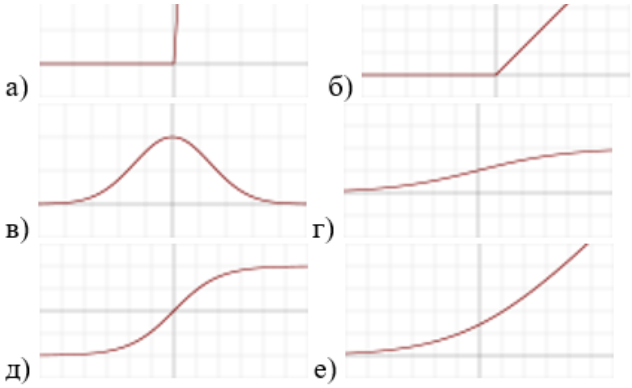
20. На рисунке изображена
А) сеть с прямыми связями
б) сеть с перекрестными связями
в) сеть с обратными связями
г) сеть с латеральными связями



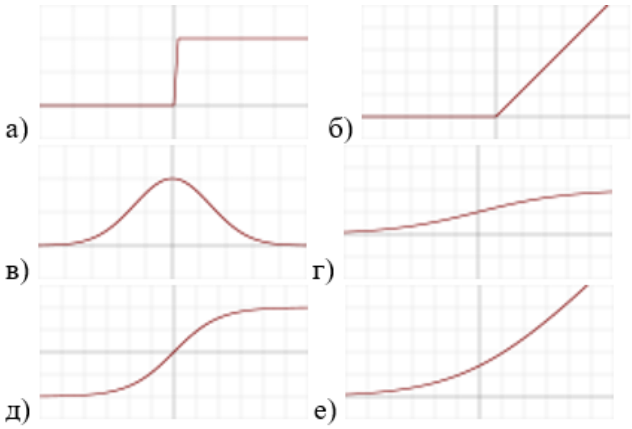
21. На рисунке изображена
А) сеть с прямыми связями
б) сеть с перекрестными связями
в) сеть с обратными связями
г) сеть с латеральными связями



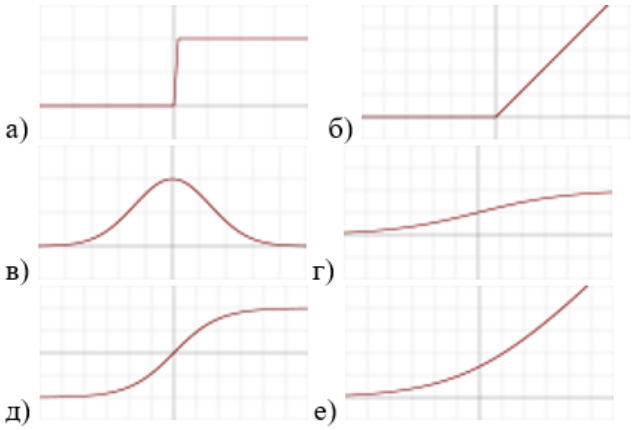
22. Выберите пороговую функцию активации (единичная ступенька)



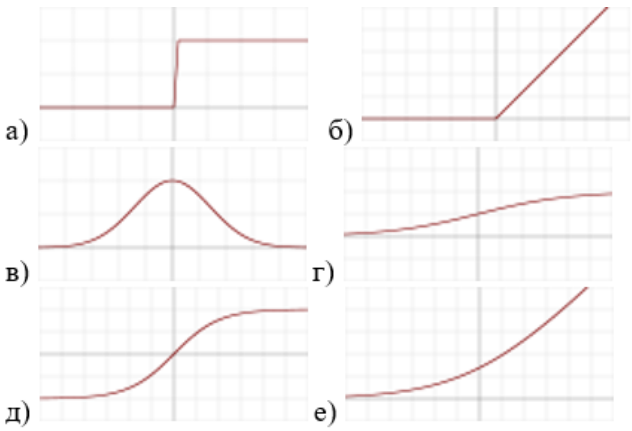
23. Выберите функцию активации линейный выпрямитель (ReLU)



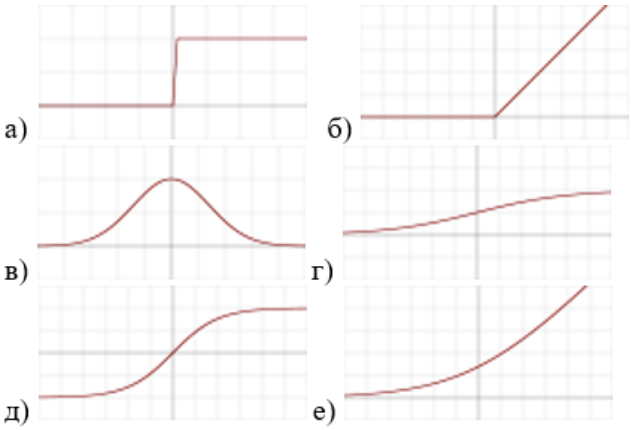
24. Выберите Гауссову функцию активации



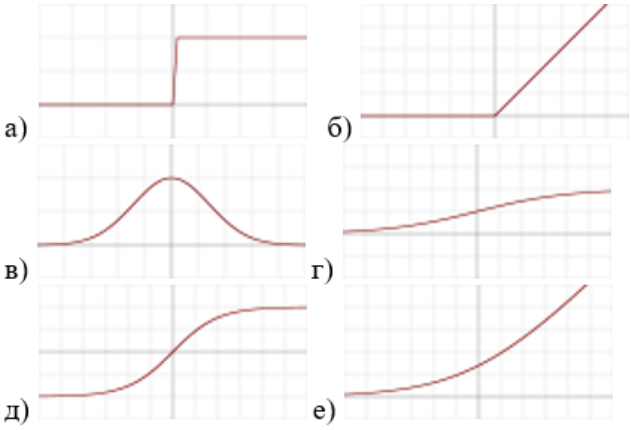
25. Выберите логистическую функцию активации (сигмоида или гладкая ступенька)



26. Выберите функцию активации гиперболический тангенс

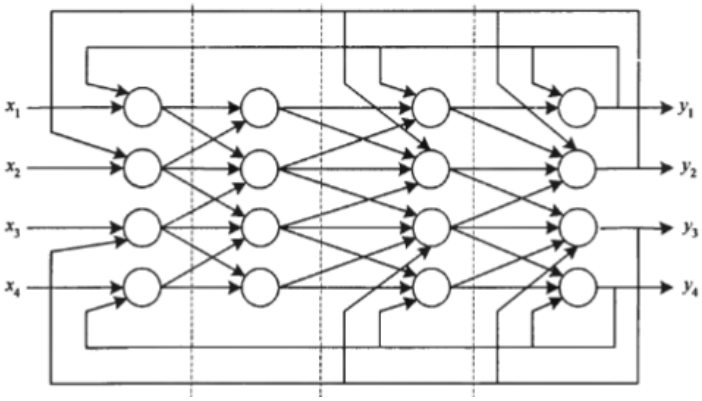


27. Выберите пороговую функцию SoftPlus



28. На рисунке изображена

А) сеть с прямыми связями
б) сеть с перекрестными связями
в) сеть с обратными связями
г) сеть с латеральными связями



9. Материально-техническое и учебно-методическое обеспечение дисциплины

9.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Запечников С. В. Основы интеллектуального анализа данных и машинного обучения: конспект лекций: учебное пособие / С. В. Запечников. - Основы интеллектуального анализа данных и машинного обучения: конспект лекций - Москва: Национальный исследовательский ядерный университет «МИФИ», 2022. - 136 c. - 978-5-7262-2856-3. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/132687.html - Режим доступа: по подписке

2. Медведев М. Ю. Neural networks fundamentals in mobile robot control systems: учебное пособие / М. Ю. Медведев, А. Е. Кульченко. - Neural networks fundamentals in mobile robot control systems - Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2020. - 144 c. - 978-5-9275-3587-3. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/107940.html - Режим доступа: по подписке

3. Ростовцев В. С. Искусственные нейронные сети: учебник для вузов / В. С. Ростовцев. - 5-е изд., стер. - Санкт-Петербург: Лань, 2025. - 216 с. - 978-5-507-50568-5. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/447392 - Режим доступа: по подписке

Дополнительная литература

1. Митяков Е. С. Искусственный интеллект и машинное обучение: учебное пособие для вузов / Е. С. Митяков, А. Г. Шмелева, А. И. Ладынин. - Санкт-Петербург: Лань, 2025. - 252 с. - 978-5-507-51465-6. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/450827 - Режим доступа: по подписке

2. Баланов А. Н. Машинное обучение и искусственный интеллект: учебное пособие для вузов / А. Н. Баланов. - Санкт-Петербург: Лань, 2024. - 172 с. - 978-5-507-49194-0. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/414920 - Режим доступа: по подписке

9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. https://www.iprbookshop.ru/ - Онлайн библиотека

9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Перечень программного обеспечения
(обновление производится по мере появления новых версий программы)

Не используется.

Перечень информационно-справочных систем
(обновление выполняется еженедельно)

Не используется.

9.4. Оборудование и технические средства обучения

Смешанный тип

Интерактивная панель - 1 шт.
Меловая доска - 4 шт.
Автоматизированное рабочее место (АРМ) - 2 шт.
Проектор - 2 шт.

10. Методические указания по освоению дисциплины (модуля)

Текущий контроль представляет собой проверку усвоения учебного материала теоретического и практического характера, регулярно осуществляемую на протяжении семестра.
К основным формам текущего контроля (текущей аттестации) относятся собеседование, тестирование.
Промежуточная аттестация осуществляется в конце семестра, завершает изучение данной дисциплины и помогает оценить более крупные совокупности знаний и умений, в некоторых случаях – даже формирование определенных компетенций.
Форма проведения промежуточной аттестации: зачет.
Оценивание знаний, умений, навыков и (или) опыта деятельности носит комплексный, системный характер и учитывает, как место дисциплины в структуре ОПОП, так и содержательные и смысловые внутренние связи. Связи формируемых компетенций с разделами дисциплины обеспечивают возможность реализации для текущего контроля, промежуточной аттестации по дисциплине и итогового контроля наиболее подходящих оценочных средств.

Рекомендуемая литература:
1. Ефимов, А. И. Основы машинного обучения : учебное пособие / А. И. Ефимов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2025. — 277 с. — ISBN 978-5-9275-5105-7. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/157821.html (дата обращения: 01.03.2026). — Режим доступа: для авторизир. пользователей
2. Медведев, М. Ю. Введение в методы машинного обучения с подкреплением : учебное пособие / М. Ю. Медведев. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2025. — 217 с. — ISBN 978-5-9275-5006-7. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/157814.html (дата обращения: 01.03.2026). — Режим доступа: для авторизир. пользователей
3. Запечников, С. В. Основы интеллектуального анализа данных и машинного обучения: конспект лекций : учебное пособие / С. В. Запечников. — Москва : Национальный исследовательский ядерный университет «МИФИ», 2022. — 136 с. — ISBN 978-5-7262-2856-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/132687.html (дата обращения: 01.03.2026). — Режим доступа: для авторизир. пользователей
4. Пылов, П. А. Основы работы с моделями машинного и глубокого обучения : учебное пособие / П. А. Пылов, Р. В. Майтак, А. В. Дягилева. — Москва, Вологда : Инфра-Инженерия, 2023. — 256 с. — ISBN 978-5-9729-1547-7. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: https://www.iprbookshop.ru/133369.html (дата обращения: 01.03.2026). — Режим доступа: для авторизир. пользователей

Описание возможностей изучения дисциплины лицами с ОВЗ и инвалидами

Инвалиды и лица с ограниченными возможностями здоровья могут допускаться на аттестационные испытания в сопровождении ассистентов-сопровождающих. Во время процедуры оценивания обучающиеся могут пользоваться рабочей программой дисциплины, а также с разрешения преподавателя справочной и нормативной литературой.