



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

Факультет геоинформатики и информационной безопасности
Кафедра информационно-измерительных систем



**ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ**

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

«УТВЕРЖДАЮ»

Проректор
А. Л. Степанченко
16 февраля 2024 г.

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)
«Б1.О.16 ОСНОВЫ МАШИННОГО ОБУЧЕНИЯ»**

Уровень высшего образования: бакалавриат

Направление подготовки: 10.03.01 Информационная безопасность

Направленность(профиль) подготовки: Организация и технология защиты информации

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Год набора (приема на обучение): 2024

Срок получения образования: 4 года

Объем:
в зачетных единицах: 4 з.е.
в академических часах: 144 ак.ч.

Москва, 2024

Разработчики:

Преподаватель кафедры Информационно-измерительных систем
Жаров С. И.

Фонд оценочных средств по дисциплине (модулю) составлен в соответствии с требованиями ФГОС ВО по направлению подготовки 10.03.01 Информационная безопасность, утвержденного приказом Минобрнауки от 17.11.2020 № 1427

Согласование и утверждение

№	Подразделение	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	факультет Геоинформатики и информационной безопасности	Председатель ученого совета факультета	Матерухин А. В.	Рассмотрено	14.12.2023, № 7-23

1. Планируемые результаты обучения, соотнесенные с планируемыми результатами освоения образовательной программы

ОПК-3 Способен использовать необходимые математические методы для решения задач профессиональной деятельности

ОПК-3.1 Выбирает и использует соответствующие ресурсы, современные методики и оборудование для проведения экспериментальных исследований и измерений

Знать:

ОПК-3.1/Зн1 основные математические методы и статистические подходы, применяемые в машинном обучении для анализа данных и проведения экспериментов

ОПК-3.1/Зн2 современные методики проведения экспериментальных исследований в области машинного обучения, включая методы сбора, обработки и анализа данных

Уметь:

ОПК-3.1/Ум1 выбирать адекватные математические методы и алгоритмы для решения конкретных задач машинного обучения с учетом особенностей данных и целей исследования

Владеть:

ОПК-3.1/Нв1 навыками работы с вычислительными платформами и программными средствами для проведения экспериментов и измерений

ОПК-3.2 Обрабатывает и представляет полученные экспериментальные данные для получения обоснованных выводов

Знать:

ОПК-3.2/Зн1 основные методы обработки экспериментальных данных в машинном обучении: предобработка, очистка, нормализация, масштабирование

ОПК-3.2/Зн2 принципы интерпретации результатов машинного обучения и критерии оценки качества моделей

Уметь:

ОПК-3.2/Ум1 обрабатывать экспериментальные данные с использованием математических и статистических методов для подготовки к анализу

ОПК-3.2/Ум2 применять методы визуализации для наглядного представления данных и результатов машинного обучения

Владеть:

ОПК-3.2/Нв1 навыками работы с инструментами обработки и анализа данных, включая программирование на Python и использование специализированных библиотек

ОПК-3.2/Нв2 методиками визуализации данных для эффективного представления результатов экспериментов и поддержки принятия решений

2. Показатели и критерии оценивания компетенций

Планируемые результаты освоения программы	Уровень овладения	Критерии оценивания (перечень сформированных ИДК/ЗУВ по уровням освоения)
---	-------------------	---

ОПК-3 Способен использовать необходимые математические методы для решения задач профессиональной деятельности	Повышенный	ИДК ОПК-3.1 Выбирает и использует соответствующие ресурсы, современные методики и оборудование для проведения экспериментальных исследований и измерений Знать: ОПК-3.1/Зн1 основные математические методы и статистические подходы, применяемые в машинном обучении для анализа данных и проведения экспериментов ОПК-3.1/Зн2 современные методики проведения экспериментальных исследований в области машинного обучения, включая методы сбора, обработки и анализа данных Уметь: ОПК-3.1/Ум1 выбирать адекватные математические методы и алгоритмы для решения конкретных задач машинного обучения с учетом особенностей данных и целей исследования Владеть: ОПК-3.1/Нв1 навыками работы с вычислительными платформами и программными средствами для проведения экспериментов и измерений
ОПК-3 Способен использовать необходимые математические методы для решения задач профессиональной деятельности	Повышенный	ИДК ОПК-3.2 Обрабатывает и представляет полученные экспериментальные данные для получения обоснованных выводов Знать: ОПК-3.2/Зн1 основные методы обработки экспериментальных данных в машинном обучении: предобработка, очистка, нормализация, масштабирование ОПК-3.2/Зн2 принципы интерпретации результатов машинного обучения и критерии оценки качества моделей Уметь: ОПК-3.2/Ум1 обрабатывать экспериментальные данные с использованием математических и статистических методов для подготовки к анализу ОПК-3.2/Ум2 применять методы визуализации для наглядного представления данных и результатов машинного обучения Владеть: ОПК-3.2/Нв1 навыками работы с инструментами обработки и анализа данных, включая программирование на Python и использование специализированных библиотек ОПК-3.2/Нв2 методиками визуализации данных для эффективного представления результатов экспериментов и поддержки принятия решений

Базовый	ИДК ОПК-3.1 Выбирает и использует соответствующие ресурсы, современные методики и оборудование для проведения экспериментальных исследований и измерений Знать: ОПК-3.1/Зн2 современные методики проведения экспериментальных исследований в области машинного обучения, включая методы сбора, обработки и анализа данных ОПК-3.1/Зн1 основные математические методы и статистические подходы, применяемые в машинном обучении для анализа данных и проведения экспериментов Уметь: ОПК-3.1/Ум1 выбирать адекватные математические методы и алгоритмы для решения конкретных задач машинного обучения с учетом особенностей данных и целей исследования
Базовый	ИДК ОПК-3.2 Обрабатывает и представляет полученные экспериментальные данные для получения обоснованных выводов Знать: ОПК-3.2/Зн2 принципы интерпретации результатов машинного обучения и критерии оценки качества моделей ОПК-3.2/Зн1 основные методы обработки экспериментальных данных в машинном обучении: предобработка, очистка, нормализация, масштабирование Уметь: ОПК-3.2/Ум1 обрабатывать экспериментальные данные с использованием математических и статистических методов для подготовки к анализу ОПК-3.2/Ум2 применять методы визуализации для наглядного представления данных и результатов машинного обучения
Пороговый	ИДК ОПК-3.1 Выбирает и использует соответствующие ресурсы, современные методики и оборудование для проведения экспериментальных исследований и измерений Знать: ОПК-3.1/Зн1 основные математические методы и статистические подходы, применяемые в машинном обучении для анализа данных и проведения экспериментов ОПК-3.1/Зн2 современные методики проведения экспериментальных исследований в области машинного обучения, включая методы сбора, обработки и анализа данных

		ИДК ОПК-3.2 Обрабатывает и представляет полученные экспериментальные данные для получения обоснованных выводов Знать: ОПК-3.2/Зн2 принципы интерпретации результатов машинного обучения и критерии оценки качества моделей ОПК-3.2/Зн1 основные методы обработки экспериментальных данных в машинном обучении: предобработка, очистка, нормализация, масштабирование
--	--	--

3. Шкала оценивания

Оценка	Критерии оценивания (перечень сформированных ИДК/ЗУВ)
Зачтено	ОПК-3.1 Выбирает и использует соответствующие ресурсы, современные методики и оборудование для проведения экспериментальных исследований и измерений ОПК-3.1/Зн1 основные математические методы и статистические подходы, применяемые в машинном обучении для анализа данных и проведения экспериментов ОПК-3.1/Зн2 современные методики проведения экспериментальных исследований в области машинного обучения, включая методы сбора, обработки и анализа данных ОПК-3.2 Обрабатывает и представляет полученные экспериментальные данные для получения обоснованных выводов ОПК-3.2/Зн1 основные методы обработки экспериментальных данных в машинном обучении: предобработка, очистка, нормализация, масштабирование ОПК-3.2/Зн2 принципы интерпретации результатов машинного обучения и критерии оценки качества моделей
Не зачтено	Результаты обучения не сформированы на достаточном уровне

4. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля/Оценочное средство
Текущий контроль	Тестирование
Промежуточная аттестация	Зачет

№ п/п	Наименование раздела	Контролируемые ИДК	Вид контроля/ используемые оценочные материалы	
			Текущий	Промежут. аттестация
1	Знакомство с анализом данных и машинным обучением	ОПК-3.1 ОПК-3.2	Тестирование	Зачет
2	Вероятностная постановка задачи и некоторые методы ее решения	ОПК-3.1 ОПК-3.2	Тестирование	Зачет

3	Метод опорных векторов	ОПК-3.1 ОПК-3.2	Тестирование	Зачет
4	Многомерная линейная регрессия	ОПК-3.1 ОПК-3.2	Тестирование	Зачет
5	Деревья решений	ОПК-3.1 ОПК-3.2	Тестирование	Зачет
6	Кластеризация и визуализация	ОПК-3.1 ОПК-3.2	Тестирование	Зачет
7	Машинное обучение в прикладных задачах	ОПК-3.1 ОПК-3.2	Тестирование	Зачет

5. 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, 4 семестр.

Зачет оценивается по следующей шкале:

- оценка «зачтено» выставляется, если обучающийся имеет глубокие, исчерпывающие знания в объеме освоенной программы; знание основной (обязательной) литературы; дает правильные и уверенные ответы, свидетельствующие о наличии твердых знаний и навыков в использовании технических средств; полное, четкое, грамотное и логически стройное изложение материала; демонстрирует свободное применение теоретических знаний; а также автоматически, когда обучающийся за промежуточную аттестацию имеет оценки «отлично»;
- оценка «не зачтено» выставляется обучающемуся в случае наличия грубых ошибок при изложении ответа на основные вопросы программы, свидетельствующих о неправильном понимании предмета; при решении практических задач показано незнание способов их решения, материал изложен беспорядочно и неуверенно.

6. Оценочные материалы текущего контроля

Раздел 1. Знакомство с анализом данных и машинным обучением

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 1.1. Знакомство с анализом данных и машинным обучением

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Машинное обучение — это:

- 1) метод кодирования сигналов
- 2) раздел науки, изучающий алгоритмы, извлекающие закономерности из данных
- 3) система хранения данных
- 4) способ решения ОДУ

2. Наличие размеченной целевой переменной характерно для:

- 1) обучения без учителя
- 2) обучения с подкреплением
- 3) обучения с учителем
- 4) кластеризации

3. К основным задачам обучения с учителем НЕ относится:

- 1) классификация
- 2) кластеризация
- 3) регрессия
- 4) ранжирование

4. Задача регрессии предполагает:

- 1) отнесение к классу
- 2) предсказание непрерывного значения
- 3) уменьшение размерности
- 4) генерацию текстов

Раздел 2. Вероятностная постановка задачи и некоторые методы ее решения

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 2.1. Вероятностная постановка задачи и некоторые методы ее решения

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какая из метрик предназначена для оценки бинарной классификации?

- 1) Silhouette Score
- 2) F-score
- 3) RMSE
- 4) ARI

2. Переобучение проявляется как:

- 1) высокая ошибка на обучении
- 2) низкая ошибка на обучении и высокая на тесте
- 3) отсутствие выбросов
- 4) равенство bias и variance

3. Какой метод помогает уменьшить переобучение?

- 1) добавление шума
- 2) L2-регуляризация
- 3) увеличение случайных признаков
- 4) уменьшение тестовой выборки

4. Обучающая выборка используется для:

- 1) итоговой оценки
- 2) визуализации
- 3) оптимизации параметров модели
- 4) хранения метаданных

Раздел 3. Метод опорных векторов

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 3.1. Метод опорных векторов

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Тестовая выборка предназначена для:

- 1) оптимизации гиперпараметров
- 2) непредвзятой оценки обобщения

- 3) нормализации признаков
- 4) балансировки классов

2. К базовым алгоритмам классификации относится:

- 1) линейная регрессия
- 2) PCA
- 3) метод опорных векторов (SVM)
- 4) K-Means

3. В методе k-ближайших соседей класс объекта определяется:

- 1) градиентным спуском
- 2) по большинству классов k ближайших точек
- 3) центроидами кластеров
- 4) глубиной дерева

4. Дерево решений формируется путём:

- 1) свёртки признаков
- 2) последовательного разбиения по информативности
- 3) усреднения моделей
- 4) снижения размерности

Раздел 4. Многомерная линейная регрессия

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 4.1. Многомерная линейная регрессия

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Увеличение сложности модели обычно:

- 1) повышает bias и снижает variance
- 2) снижает bias и повышает variance
- 3) не влияет на bias-variance
- 4) уменьшает обе величины

2. Глубина дерева в решающих деревьях является:

- 1) параметром ошибки
- 2) гиперпараметром
- 3) метрикой качества
- 4) функцией потерь

3. Precision измеряет:

- 1) долю верных предсказаний
- 2) долю истинно положительных в предсказанных положительных
- 3) площадь под ROC
- 4) среднюю абсолютную ошибку

4. При сильном дисбалансе классов предпочтительнее использовать:

- 1) Accuracy
- 2) ROC-AUC
- 3) PR-AUC
- 4) MAE

Раздел 5. Деревья решений

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 5.1. Деревья решений

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Линейная регрессия минимизирует:

- 1) энтропию
- 2) сумму квадратов ошибок
- 3) силуэт
- 4) log-loss

2. Признаковое пространство — это:

- 1) набор целевых меток
- 2) N-мерное пространство признаков объектов
- 3) набор гиперпараметров
- 4) матрица ошибок

3. Нормализация признаков необходима, чтобы:

- 1) увеличить размерность
- 2) привести признаки к сопоставимому масштабу
- 3) удалить коллинеарность
- 4) изменить классы

4. Нормализация преобразует признаки к:

- 1) диапазону $[-\infty; \infty]$
- 2) лог-шкале
- 3) нулевому среднему и единичной дисперсии
- 4) бинарному виду

Раздел 6. Кластеризация и визуализация

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 6.1. Кластеризация и визуализация

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Обучение модели — это:

- 1) визуализация данных
- 2) настройка параметров по обучающей выборке
- 3) сжатие датасета
- 4) предсказание классов

2. Бэггинг отличается от бустинга тем, что:

- 1) использует одну модель
- 2) обучает модели параллельно и усредняет их
- 3) фокусируется на ошибках предыдущих моделей
- 4) применяет градиентный спуск

3. Пример ансамблевого метода на деревьях:

- 1) Случайный лес
- 2) K-Means
- 3) LASSO
- 4) t-SNE

4. Искусственная нейронная сеть представляет собой:

- 1) дерево решений
- 2) композицию линейных и нелинейных преобразований
- 3) набор центроидов
- 4) граф кластеров

5. Одна эпоха обучения нейросети — это:

- 1) один шаг оптимизатора
- 2) один полный проход по обучающим данным
- 3) обновление гиперпараметра
- 4) вычисление метрики

Раздел 7. Машинное обучение в прикладных задачах

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Тема 7.1. Машинное обучение в прикладных задачах

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Ошибка обобщения измеряется на:
 - 1) обучающей выборке
 - 2) данных, не участвовавших в обучении
 - 3) валидационной выборке
 - 4) наборе гиперпараметров
2. L2-регуляризация при мультиколлинеарности:
 - 1) увеличивает variance
 - 2) снижает дисперсию коэффициентов
 - 3) удаляет признаки
 - 4) меняет метки классов
3. Размерность данных определяется:
 - 1) числом строк
 - 2) количеством признаков
 - 3) объёмом памяти
 - 4) числом кластеров
4. Случайный лес строится как:
 - 1) одно глубокое дерево
 - 2) ансамбль случайно построенных деревьев
 - 3) K-средних
 - 4) нейронная сеть
5. В градиентном бустинге каждая новая модель аппроксимирует:
 - 1) данные заново
 - 2) метки классов
 - 3) веса предыдущих деревьев
 - 4) градиент функции потерь текущего ансамбля

7. Оценочные материалы промежуточной аттестации

4 семестр, Зачет

Контролируемые ИДК: ОПК-3.1 ОПК-3.2

Вопросы/Задания:

1. Дайте определение машинного обучения.
2. Сформулируйте ключевое отличие обучения с учителем от обучения без учителя.
3. Перечислите два основных типа задач обучения с учителем.
4. Дайте определение задачи регрессии.
5. Назовите три распространённые метрики качества для бинарной классификации.
6. Объясните, что такое переобучение.

7. Укажите хотя бы два способа профилактики переобучения.
8. Определите «обучающую выборку».
9. Для чего предназначена тестовая выборка?
10. Назовите четыре базовых алгоритма классификации.
11. Объясните метод k-ближайших соседей.
12. Как строится дерево решений?
13. Как увеличение сложности модели влияет на смещение (bias) и дисперсию (variance) ошибок?
14. Что называют гиперпараметрами модели?
15. В чём различие метрик Accuracy и Precision?
16. Почему при сильном дисбалансе классов PR-AUC предпочтительнее ROC-AUC?
17. Сформулируйте идею линейной регрессии.
18. Определите «признаковое пространство».
19. Зачем выполняется нормализация признаков?
20. Что такое нормализация?
21. Различите обучение и предсказание.
22. Опишите бэггинг и бустинг как стратегии ансамблирования.
23. Приведите пример ансамблевого метода на деревьях решений.
24. Дайте определение искусственной нейронной сети.
25. Что называют эпохой обучения нейронной сети?
26. Чем отличается ошибка обучения от ошибки обобщения?
27. Как L2-регуляризация помогает при мультиколлинеарности признаков в линейных моделях?
28. Что такое «размерность данных»?
29. Сформулируйте идею алгоритма «случайный лес».
30. Какую величину аппроксимирует каждая новая модель в градиентном бустинге?

