

**КАЗАНСКИЙ КООПЕРАТИВНЫЙ ИНСТИТУТ (ФИЛИАЛ)
АВТОНОМНОЙ НЕКОММЕРЧЕСКОЙ ОБРАЗОВАТЕЛЬНОЙ
ОРГАНИЗАЦИИ ВЫСШЕГО ОБРАЗОВАНИЯ
ЦЕНТРОСОЮЗА РОССИЙСКОЙ ФЕДЕРАЦИИ
«РОССИЙСКИЙ УНИВЕРСИТЕТ КООПЕРАЦИИ»**

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

МАШИННОЕ ОБУЧЕНИЕ

Специальность: 10.05.01 Компьютерная безопасность

Специализация: «Разработка защищенного программного обеспечения»

Формы обучения: очная

Объем дисциплины:

в зачетных единицах: 4 з.е.

в академических часах: 144 ак.ч.

Форма промежуточной аттестации: экзамен

Оценочные материалы по дисциплине «Машинное обучение»

обсуждены и рекомендованы к утверждению решением кафедры естественных дисциплин, сервиса и туризма от 25 марта 2025 г., протокол № 6

одобрены Научно-методическим советом Казанского кооперативного института (филиала) от 2 апреля 2025 г., протокол № 3

СОДЕРЖАНИЕ

1. Паспорт оценочных материалов по дисциплине
2. Оценочные материалы по дисциплине:
 - 2.1. Оценочные материалы для проведения текущего контроля.
 - 2.2. Оценочные материалы для проведения промежуточной аттестации.

Обновление оценочных материалов дисциплины

Целью оценочных материалов по дисциплине (модулю) (далее – ОМ) является комплексная оценка результатов обучения по дисциплине Машинное обучение и установление уровня сформированности компетенций обучающегося.

ОМ предназначены для проведения текущего контроля успеваемости и промежуточной аттестации.

ОМ включают оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации.

Оценочные материалы разработаны в соответствии с рабочей программой дисциплины Машинное обучение.

1. Паспорт оценочных материалов по дисциплине «Машинное обучение»

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций	Планируемые результаты обучения	Контролируемые разделы, темы дисциплины	Наименование оценочного средства ¹	
				Текущий контроль	Промежуточная аттестация
1	2	3	4	5	6
ОПК-8 Способен применять методы научных исследований при проведении разработок в области обеспечения безопасности компьютерных систем и сетей	ОПК-8.1. Демонстрирует умение применять методику имитационного моделирования и инструментальные средства её поддержки при проведении исследований в области обеспечения безопасности компьютерных систем и сетей	Знать: основы теории и методики имитационного моделирования применительно к исследованиям информационной безопасности, современные инструментальные средства и среды моделирования, используемые в анализе безопасности компьютерных систем и сетей; Уметь: разрабатывать и настраивать модели компьютерных систем и сетей с целью проверки гипотез относительно уровня их защищенности, создавать сценарии тестирования с применением имитационного моделирования для оценки влияния различных факторов риска на безопасность; Владеть: навыками выбора оптимальных методик и инструментариев для решения конкретных исследовательских задач в сфере безопасности компьютерных систем и сетей, умениями самостоятельного проектирования экспериментальных схем	Тема 1. Введение в машинное обучение Тема 2. Регрессия Постановка задачи регрессии Тема 3. Классификация. Постановка задачи классификации Тема 4. Методы обучения с учителем	Реферат (Тема 1) Тест (тема 1) Реферат (Тема 2) Тест (тема 2) Реферат (Тема 3) Тест (тема 3) Реферат (Тема 4) Тест (тема 4)	Тест (1-20 тестового задания)

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций	Планируемые результаты обучения	Контролируемые разделы, темы дисциплины	Наименование оценочного средства ¹	
				Текущий контроль	Промежуточная аттестация
1	2	3	4	5	6
		и проведения сложных вычислительных экспериментов с применением имитации;			
	ОПК-8.2. Демонстрирует умение решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и математического моделирования	Знать: базовые законы природы и фундаментальные научные принципы, лежащие в основе инженерных дисциплин, ключевые математические методы и алгоритмы, применяемые для анализа технических систем и решения профессиональных задач. Уметь: ставить и грамотно формализовать стандартные профессиональные задачи, используя знания естественных наук и математики, выбирать адекватные методы математического анализа и моделирования для решения поставленных задач. Владеть: навыками эффективного применения естественно-научных и общеинженерных знаний в своей профессиональной деятельности, методами постановки и решения стандартных профессиональных задач с привлечением математического аппарата;	Тема 5. Диагностика систем машинного обучения Тема 6. Предварительный анализ и обработка данных Тема 7. Задачи обучения без учителя Тема 8. Практическое использование моделей машинного обучения	Реферат (Тема 5) Тест (Тема 5) Реферат (Тема 6) Тест (тема 6) Реферат (Тема 7) Тест (тема 7) Реферат (Тема 8) Тест (тема 8)	Тест (21-40 тестового задания)

2. Оценочные материалы по дисциплине «Машинное обучение»

2.1 Оценочные материалы для проведения текущего контроля успеваемости

Тематика рефератов

Тема 1. Введение в машинное обучение

1. История развития машинного обучения: от зарождения до современных технологий
2. Основные задачи и направления машинного обучения
3. Сравнение машинного обучения с другими методами анализа данных
4. Применение машинного обучения в различных отраслях (медицина, финансы, маркетинг и т.д.)
5. Этика и машинное обучение: проблемы и вызовы
6. Машинное обучение и искусственный интеллект: взаимосвязь и различия
7. Методы обучения с учителем и без учителя: сравнительный анализ
8. Глубокое обучение: принципы и примеры применения
9. Машинное обучение и большие данные: взаимозависимость и перспективы
10. Проблемы и ограничения машинного обучения: что мешает идеальному алгоритму?
11. Машинное обучение в образовании: перспективы и вызовы
12. Машинное обучение и кибербезопасность: как технологии помогают защищать данные?
13. Машинное обучение в социальных сетях: анализ данных и персонализация контента
14. Машинное обучение и робототехника: интеграция и перспективы
15. Будущее машинного обучения: прогнозы и направления развития

Тема 2. Регрессия. Постановка задачи регрессии.

1. История и развитие методов регрессионного анализа
2. Виды регрессионных моделей: линейная, нелинейная, множественная регрессия
3. Методы оценки параметров регрессионной модели: метод наименьших квадратов и его модификации
4. Регрессия и корреляция: взаимосвязь и различия
5. Применение регрессионного анализа в экономике и финансах
6. Регрессия в социальных науках: прогнозирование и моделирование социальных процессов
7. Регрессия в медицине: прогнозирование заболеваний и анализ медицинских данных
8. Регрессия в маркетинге: анализ потребительского поведения и прогнозирование продаж
9. Регрессия в экологии: моделирование и прогнозирование экологических процессов
10. Проблемы мультиколлинеарности в регрессионном анализе и методы их решения
11. Регрессия с использованием категориальных переменных: методы кодирования и интерпретация результатов
12. Регрессия в условиях гетероскедастичности: методы диагностики и коррекции
13. Регрессия с использованием временных рядов: особенности и методы анализа
14. Регрессия и нейронные сети: интеграция и перспективы
15. Оценка качества регрессионных моделей: критерии и методы

Тема 3. Классификация. Постановка задачи классификации

1. История и развитие методов классификации в машинном обучении
2. Основные задачи и типы классификации: бинарная и многоклассовая классификация
3. Методы классификации: деревья решений, случайный лес, метод k-ближайших соседей
4. Классификация и регрессия: взаимосвязь и различия
5. Применение классификации в медицине: диагностика заболеваний и прогнозирование исходов
6. Классификация в маркетинге: сегментация клиентов и прогнозирование поведения
7. Классификация в финансах: оценка кредитоспособности и выявление мошенничества
8. Классификация в социальных сетях: анализ и категоризация контента
9. Классификация в биоинформатике: анализ геномных данных и классификация белков
10. Проблемы несбалансированных классов в задачах классификации и методы их решения
11. Классификация с использованием категориальных переменных: методы кодирования и интерпретация результатов
12. Классификация в условиях шума и пропущенных данных: методы обработки и улучшения качества данных
13. Классификация с использованием нейронных сетей: архитектура и применение
14. Оценка качества классификаторов: точность, полнота, F-мера и ROC-кривая
15. Классификация и этические аспекты: предвзятость и справедливость алгоритмов

Тема 4. Методы обучения с учителем

1. История и развитие методов обучения с учителем в машинном обучении
2. Сравнение методов обучения с учителем и без учителя: задачи и подходы
3. Линейная регрессия: принципы, методы и примеры применения
4. Логистическая регрессия: принципы, методы и примеры применения
5. Деревья решений: принципы построения и применения в задачах классификации и регрессии
6. Метод k-ближайших соседей: принципы, методы и примеры применения
7. Метод опорных векторов (SVM): принципы, методы и примеры применения
8. Нейронные сети: архитектура, обучение и применение в задачах классификации и регрессии
9. Обучение с учителем в задачах распознавания образов: методы и примеры
10. Обучение с учителем в задачах обработки естественного языка: методы и примеры
11. Проблемы переобучения и недообучения в методах обучения с учителем: диагностика и методы решения
12. Методы регуляризации в обучении с учителем: L1, L2 и их применение
13. Градиентный спуск и его модификации: принципы, методы и примеры применения
14. Оценка качества моделей обучения с учителем: метрики и методы
15. Этика и обучение с учителем: предвзятость данных и справедливость алгоритмов

Тема 5. Диагностика систем машинного обучения.

1. Методы оценки качества моделей машинного обучения: метрики и критерии
2. Диагностика переобучения и недообучения: методы и примеры

3. Анализ ошибок в системах машинного обучения: причины и способы исправления
4. Валидация и тестирование моделей машинного обучения: методы и подходы
5. Интерпретируемость моделей машинного обучения: методы и примеры
6. Диагностика и улучшение моделей на основе анализа данных: методы и примеры
7. Проблемы и методы диагностики систем машинного обучения в реальном времени
8. Диагностика систем машинного обучения в задачах классификации: методы и примеры
9. Диагностика систем машинного обучения в задачах регрессии: методы и примеры
10. Диагностика систем машинного обучения в задачах обработки естественного языка: методы и примеры
11. Диагностика систем машинного обучения в задачах компьютерного зрения: методы и примеры
12. Проблемы и методы диагностики систем машинного обучения в условиях ограниченных данных
13. Диагностика систем машинного обучения в условиях несбалансированных классов: методы и примеры
14. Диагностика систем машинного обучения в условиях шума и пропущенных данных: методы и примеры
15. Этика и диагностика систем машинного обучения: предвзятость и справедливость алгоритмов

Тема 6. Предварительный анализ и обработка данных.

1. Этапы предварительного анализа данных: от сбора до подготовки
2. Методы очистки данных: обработка пропущенных значений и аномалий
3. Методы нормализации и стандартизации данных: принципы и примеры
4. Кодирование категориальных переменных: методы и примеры
5. Методы отбора признаков: фильтрация, оберстка и встроенные методы
6. Методы уменьшения размерности данных: PCA, t-SNE и другие
7. Анализ и обработка временных рядов: методы и примеры
8. Анализ и обработка текстовых данных: методы и примеры
9. Анализ и обработка изображений: методы и примеры
10. Проблемы и методы обработки несбалансированных данных
11. Проблемы и методы обработки данных с шумом и выбросами
12. Методы визуализации данных: принципы и примеры
13. Методы анализа и обработки больших данных: Hadoop, Spark и другие
14. Методы анализа и обработки данных в реальном времени: потоковая обработка
15. Этика и обработка данных: конфиденциальность и защита персональных данных

Тема 7. Задачи обучения без учителя. Задача кластеризации

1. История и развитие методов кластеризации в машинном обучении
2. Основные задачи и типы кластеризации: иерархическая, k-средних и другие
3. Методы кластеризации: k-средних, DBSCAN, иерархическая кластеризация
4. Кластеризация и классификация: взаимосвязь и различия
5. Применение кластеризации в маркетинге: сегментация клиентов и анализ поведения
6. Кластеризация в социальных сетях: анализ и категоризация контента
7. Кластеризация в биоинформатике: анализ геномных данных и классификация белков
8. Кластеризация в медицине: анализ медицинских данных и диагностика заболеваний
9. Кластеризация в финансах: анализ финансовых данных и выявление аномалий
10. Проблемы и методы определения количества кластеров: методы и примеры
11. Кластеризация с использованием категориальных переменных: методы и примеры

12. Кластеризация в условиях шума и пропущенных данных: методы и примеры
13. Кластеризация с использованием нейронных сетей: архитектура и применение
14. Оценка качества кластеризации: метрики и методы
15. Этика и кластеризация: предвзятость и справедливость алгоритмов

Тема 8. Практическое использование моделей машинного обучения.

1. Применение машинного обучения в медицине: диагностика заболеваний и прогнозирование исходов
2. Применение машинного обучения в финансах: оценка кредитоспособности и выявление мошенничества
3. Применение машинного обучения в маркетинге: сегментация клиентов и прогнозирование продаж
4. Применение машинного обучения в социальных сетях: анализ и категоризация контента
5. Применение машинного обучения в биоинформатике: анализ геномных данных и классификация белков
6. Применение машинного обучения в экологии: моделирование и прогнозирование экологических процессов
7. Применение машинного обучения в робототехнике: интеграция и управление роботами
8. Применение машинного обучения в образовании: персонализация обучения и анализ успеваемости
9. Применение машинного обучения в энергетике: прогнозирование энергопотребления и оптимизация ресурсов
10. Применение машинного обучения в сельском хозяйстве: прогнозирование урожайности и управление ресурсами
11. Применение машинного обучения в транспортной логистике: оптимизация маршрутов и управление грузопотоками
12. Применение машинного обучения в безопасности: анализ данных и предотвращение угроз
13. Применение машинного обучения в управлении персоналом: подбор кадров и оценка эффективности
14. Применение машинного обучения в юриспруденции: анализ юридических текстов и прогнозирование судебных решений
15. Этика и практическое использование моделей машинного обучения: предвзятость и справедливость алгоритмов

Критерии оценки выполнения рефератов по учебной дисциплине

Оценка «отлично» выставляется студенту, если тема реферата раскрыта в полной мере и все требования по написанию реферата соблюдены.

Оценка «хорошо» выставляется студенту, если недостаточно раскрыта тема реферата, но все требования по написанию реферата соблюдены.

Оценка «удовлетворительно» выставляется студенту, если недостаточно раскрыта тема реферата, не все требования по написанию реферата соблюдены, присутствуют ошибки и неточности в работе.

Оценка «неудовлетворительно» выставляется студенту, если отсутствует реферат.

Рекомендации по написанию реферата:

Реферат — письменная работа, выполняемая обучающимся в течение определенного срока.

Реферат — краткое точное изложение сущности какого-либо вопроса, темы на основе одной или нескольких книг, монографий или других первоисточников. Реферат должен содержать основные фактические сведения и выводы по рассматриваемому вопросу.

Реферат отвечает на вопрос — что содержится в данной публикации (публикациях). Реферат — не механический пересказ работы, а изложение ее существа.

Кроме реферирования прочитанной литературы, от обучающегося требуется аргументированное изложение собственных мыслей по рассматриваемому вопросу. Тему реферата предлагает преподаватель или сам обучающийся, в последнем случае она должна быть согласованна с преподавателем.

В реферате нужны развернутые аргументы, рассуждения, сравнения. Материал подается не столько в развитии, сколько в форме констатации или описания.

Содержание реферируемого произведения излагается объективно от имени автора. Если в первичном документе главная мысль сформулирована недостаточно четко, в реферате она должна быть конкретизирована и выделена.

Структура реферата:

1. Титульный лист.
2. После титульного листа на отдельной странице следует оглавление (план, содержание), в котором указаны названия всех разделов (пунктов плана) реферата и номера страниц, указывающие начало этих разделов в тексте реферата.
3. После оглавления следует введение. Объем введения составляет 1,5-2 страницы.
4. Основная часть реферата может иметь одну или несколько глав, состоящих из 2-3 параграфов (подпунктов, разделов) и предполагает осмысленное и логичное изложение главных положений и идей, содержащихся в изученной литературе. В тексте обязательны ссылки на первоисточники. В том случае если цитируется или используется чья-либо неординарная мысль, идея, вывод, приводится какой-либо цифрой материал, таблицу - обязательно сделайте ссылку на того автора у кого вы взяли данный материал.
5. Заключение содержит главные выводы, и итоги из текста основной части, в нем отмечается, как выполнены задачи и достигнуты ли цели, сформулированные во введении.
6. Приложение может включать графики, таблицы, расчеты.
7. Библиография (список литературы) здесь указывается реально использованная для написания реферата литература. Список составляется согласно правилам библиографического описания.

Этапы работы над рефератом.

Работу над рефератом можно условно подразделить на три этапа:

1. Подготовительный этап, включающий изучение предмета исследования;
2. Изложение результатов изучения в виде связного текста;
3. Устное сообщение по теме реферата.

1. Подготовительный этап работы.

Формулировка темы. Подготовительная работа над рефератом начинается с формулировки темы. Тема в концентрированном виде выражает содержание будущего текста, фиксируя как предмет исследования, так и его ожидаемый результат. Для того чтобы работа над рефератом была успешной, необходимо, чтобы тема заключала в себе проблему, скрытый вопрос (даже если наука уже давно дала ответ на этот вопрос, студент, только знакомящийся с соответствующей областью знаний, будет вынужден искать ответ заново, что даст толчок к развитию проблемного, исследовательского мышления).

Поиск источников. Грамотно сформулированная тема зафиксировала предмет изучения; задача обучающегося — найти информацию, относящуюся к данному предмету и разрешить поставленную проблему. Выполнение этой задачи начинается с поиска источников. На этом этапе необходимо вспомнить, как работать с энциклопедиями и энциклопедическими словарями.

Работа с источниками. Работу с источниками надо начинать с ознакомительного чтения, т. е. просмотреть текст, выделяя его структурные единицы. При ознакомительном

чтении закладками отмечаются те страницы, которые требуют более внимательного изучения. В зависимости от результатов ознакомительного чтения выбирается дальнейший способ работы с источником. Если для разрешения поставленной задачи требуется изучение некоторых фрагментов текста, то используется метод выборочного чтения. Если в книге нет подробного оглавления, следует обратить внимание на предметные и именные указатели.

Избранные фрагменты или весь текст (если он целиком имеет отношение к теме) требуют вдумчивого, неторопливого чтения с «мысленной проработкой» материала. Такое чтение предполагает выделение: 1) главного в тексте; 2) основных аргументов; 3) выводов. Особое внимание следует обратить на то, вытекает тезис из аргументов или нет. Необходимо также проанализировать, какие из утверждений автора носят проблематичный, гипотетический характер и уловить скрытые вопросы.

Понятно, что умение таким образом работать с текстом приходит далеко не сразу. Наилучший способ научиться выделять главное в тексте, улавливать проблематичный характер утверждений, давать оценку авторской позиции — это сравнительное чтение, в ходе которого студент знакомится с различными мнениями по одному и тому же вопросу, сравнивает весомость и доказательность аргументов сторон и делает вывод о наибольшей убедительности той или иной позиции.

Создание конспектов для написания реферата.

Подготовительный этап работы завершается созданием конспектов, фиксирующих основные тезисы и аргументы. Здесь важно вспомнить, что конспекты пишутся на одной стороне листа, с полями и достаточным для исправления и ремарок межстрочным расстоянием (эти правила соблюдаются для удобства редактирования). Если в конспектах приводятся цитаты, то непременно должно быть дано указание на источник (автор, название, выходные данные, № страницы).

По завершении предварительного этапа можно переходить непосредственно к созданию текста реферата.

2. Создание текста.

Общие требования к тексту.

Текст реферата должен подчиняться определенным требованиям: он должен раскрывать тему, обладать связностью и цельностью.

Раскрытие темы предполагает, что в тексте реферата излагается относящийся к теме материал и предлагаются пути решения содержащейся в теме проблемы; связность текста предполагает смысловую соотносительность отдельных компонентов, а цельность — смысловую законченность текста.

С точки зрения связности все тексты делятся на тексты-констатации и тексты-рассуждения. Тексты-констатации содержат результаты ознакомления с предметом и фиксируют устойчивые и несомненные суждения. В текстах-рассуждениях одни мысли извлекаются из других, некоторые ставятся под сомнение, дается им оценка, выдвигаются различные предположения.

План реферата.

Универсальный план реферата — введение, основной текст и заключение.

Требования к введению.

Во введении аргументируется актуальность исследования, — т. е. выявляется практическое и теоретическое значение данного исследования. Далее констатируется, что сделано в данной области предшественниками; перечисляются положения, которые должны быть обоснованы. Введение может также содержать обзор источников или экспериментальных данных, уточнение исходных понятий и терминов, сведения о методах исследования. Во введении обязательно формулируются цель и задачи реферата.

Объем введения — в среднем около 10% от общего объема реферата.

Основная часть реферата.

Основная часть реферата раскрывает содержание темы. Она наиболее значительна по объему, наиболее значима и ответственна. В ней обосновываются основные тезисы реферата,

приводятся развернутые аргументы, предполагаются гипотезы, касающиеся существа обсуждаемого вопроса. Важно проследить, чтобы основная часть не имела форму монолога. Аргументируя собственную позицию, можно и должно анализировать, и оценивать позиции различных исследователей, с чем-то соглашаться, чему-то возражать, кого-то опровергать. Текст основной части делится на главы, параграфы, пункты. План основной части может быть составлен с использованием различных методов группировки материала: классификации (эмпирические исследования), типологии (теоретические исследования), периодизации (исторические исследования).

Заключение.

Заключение — последняя часть научного текста. В ней краткой и сжатой форме излагаются полученные результаты, представляющие собой ответ на главный вопрос исследования. Здесь же могут намечаться и дальнейшие перспективы развития темы. Небольшое по объему сообщение также не может обойтись без заключительной части - пусть это будут две-три фразы. Но в них должен подводиться итог проделанной работы.

Список использованной литературы.

Реферат любого уровня сложности обязательно сопровождается списком используемой литературы. Названия книг в списке располагают по алфавиту с указанием выходных данных использованных книг.

Требования, предъявляемые к оформлению реферата

Объемы рефератов – от 10-18 машинописных страниц. Работа выполняется на одной стороне листа стандартного формата. По обеим сторонам листа оставляются поля размером 35 мм. слева и 15 мм. справа, рекомендуется шрифт 12-14, интервал - 1,5. Все листы реферата должны быть пронумерованы. Каждый вопрос в тексте должен иметь заголовок в точном соответствии с наименованием в плане-оглавлении. При написании и оформлении реферата следует избегать типичных ошибок, например, таких:

- поверхностное изложение основных теоретических вопросов выбранной темы, когда автор не понимает, какие проблемы в тексте являются главными, а какие второстепенными,

- в некоторых случаях проблемы, рассматриваемые в разделах, не раскрывают основных аспектов выбранной для реферата темы,

- дословное переписывание книг, статей, заимствования рефератов из интернета и т. д.

При проверке реферата преподавателем оцениваются:

1. Знания и умения на уровне требований стандарта конкретной дисциплины: знание фактического материала, усвоение общих представлений, понятий, идей.

2. Характеристика реализации цели и задач исследования (новизна и актуальность поставленных в реферате проблем, правильность формулирования цели, определения задач исследования, правильность выбора методов решения задач и реализации цели; соответствие выводов решаемым задачам, поставленной цели, убедительность выводов).

3. Степень обоснованности аргументов и обобщений (полнота, глубина, всесторонность раскрытия темы, логичность и последовательность изложения материала, корректность аргументации и системы доказательств, характер и достоверность примеров, иллюстративного материала, широта кругозора автора, наличие знаний интегрированного характера, способность к обобщению).

4. Качество и ценность полученных результатов (степень завершенности реферативного исследования, спорность или однозначность выводов).

5. Использование литературных источников.

6. Культура письменного изложения материала.

7. Культура оформления материалов работы.

Типовые тесты

Тест №1

Тема 1. Введение в машинное обучение

Вопрос 1:

Что такое машинное обучение?

- А) Процесс, при котором компьютерные программы самостоятельно обучаются на данных без явного программирования.
- В) Процесс, при котором компьютерные программы обучаются только на заранее заданных правилах.
- С) Процесс, при котором компьютерные программы обучаются исключительно на больших данных.
- Д) Процесс, при котором компьютерные программы обучаются только на малых данных.

Ответ: А

Обоснование: Машинное обучение — это область искусственного интеллекта, которая позволяет компьютерным программам обучаться на данных и улучшать свои результаты без явного программирования. Это означает, что алгоритмы машинного обучения могут автоматически выявлять закономерности и делать прогнозы на основе данных, не требуя ручного программирования для каждого конкретного случая.

Вопрос 2:

Какие основные задачи решает машинное обучение?

- А) Регрессия и классификация.
- В) Кластеризация и регрессия.
- С) Классификация и кластеризация.
- Д) Все вышеперечисленное.

Ответ: D

Обоснование: Машинное обучение решает широкий спектр задач, включая:

Регрессию — предсказание числовых значений (например, прогнозирование цены на недвижимость).

Классификация — отнесение объектов к определенным категориям (например, определение спама в электронной почте).

Кластеризация — группировка объектов на основе их сходства (например, сегментация клиентов по их поведению).

Вопрос 3:

Какие методы обучения используются в машинном обучении?

- А) Обучение с учителем и обучение без учителя.
- В) Обучение с подкреплением и обучение с учителем.
- С) Обучение без учителя и обучение с подкреплением.
- Д) Все вышеперечисленное.

Ответ: D

Обоснование: В машинном обучении используются три основных метода:

Обучение с учителем — когда алгоритм обучается на размеченных данных (например, классификация изображений).

Обучение без учителя — когда алгоритм обучается на неразмеченных данных

(например, кластеризация).

Обучение с подкреплением — когда алгоритм обучается путем взаимодействия с окружающей средой и получения обратной связи (например, обучение роботов).

Вопрос 4:

Какая из следующих областей не относится к машинному обучению?

- A) Глубокое обучение.
- B) Нейронные сети.
- C) Линейная регрессия.
- D) Теория игр.

Ответ: D

Обоснование: Теория игр — это область математики, которая изучает стратегическое взаимодействие между игроками. Она не относится к машинному обучению, в отличие от:

Глубокого обучения — подкатегории машинного обучения, использующей нейронные сети с большим количеством слоев.

Нейронных сетей — математических моделей, вдохновленных работой мозга, которые используются в машинном обучении.

Линейной регрессии — метода машинного обучения для предсказания числовых значений.

Вопрос 5:

Что такое переобучение в машинном обучении?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.
- B) Когда модель недостаточно обучается на тренировочных данных.
- C) Когда модель обучается только на малых данных.
- D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение происходит, когда модель машинного обучения слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Тема 2. Регрессия Постановка задачи регрессии.

Вопрос 1:

Что такое регрессия в машинном обучении?

- A) Задача классификации объектов на несколько категорий.
- B) Задача предсказания числовых значений на основе входных данных.
- C) Задача группировки объектов на основе их сходства.
- D) Задача обучения с подкреплением.

Ответ: B

Обоснование: Регрессия — это задача машинного обучения, в которой цель состоит в предсказании числовых значений (непрерывных величин) на основе входных данных. Например, предсказание цены на недвижимость или прогнозирование температуры на следующий день.

Вопрос 2:

Какие из следующих методов относятся к регрессионным моделям?

- A) Линейная регрессия.
- B) Логистическая регрессия.
- C) Метод k-ближайших соседей.
- D) Деревья решений.

Ответ: A

Обоснование: Линейная регрессия — это классический метод регрессии, используемый для предсказания числовых значений на основе линейной зависимости между входными и выходными переменными. Логистическая регрессия, несмотря на название, используется для классификации, а не для регрессии. Метод k-ближайших соседей и деревья решений могут использоваться как для классификации, так и для регрессии, но в данном контексте правильный ответ — линейная регрессия.

Вопрос 3:

Что такое мультиколлинеарность в регрессионном анализе?

- A) Ситуация, когда одна из независимых переменных сильно коррелирует с зависимой переменной.
- B) Ситуация, когда две или более независимых переменных сильно коррелируют друг с другом.
- C) Ситуация, когда модель переобучается на тренировочных данных.
- D) Ситуация, когда модель недостаточно обучается на тренировочных данных.

Ответ: B

Обоснование: Мультиколлинеарность возникает, когда две или более независимых переменных в регрессионной модели сильно коррелируют друг с другом. Это может привести к проблемам с интерпретацией коэффициентов модели и снижению ее предсказательной способности.

Вопрос 4:

Какая из следующих метрик используется для оценки качества регрессионной модели?

- A) Точность (Accuracy).
- B) F-мера (F-score).
- C) Среднее абсолютное отклонение (Mean Absolute Error, MAE).
- D) Точность (Precision).

Ответ: C

Обоснование: Для оценки качества регрессионных моделей используются метрики, которые измеряют разницу между предсказанными и фактическими значениями. Среднее абсолютное отклонение (MAE) — это одна из таких метрик, которая вычисляет среднее абсолютное значение ошибок предсказания. Точность (Accuracy), F-мера (F-score) и Точность (Precision) используются для оценки качества классификационных моделей.

Вопрос 5:

Что такое переобучение в контексте регрессионного анализа?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо

обобщает на новых данных.

- В) Когда модель недостаточно обучается на тренировочных данных.
- С) Когда модель обучается только на малых данных.
- Д) Когда модель обучается только на больших данных.

Ответ: А

Обоснование: Переобучение в регрессионном анализе происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Тема 3. Классификация. Постановка задачи классификации

Вопрос 1:

Что такое классификация в машинном обучении?

- А) Задача предсказания числовых значений на основе входных данных.
- В) Задача отнесения объектов к определенным категориям.
- С) Задача группировки объектов на основе их сходства.
- Д) Задача обучения с подкреплением.

Ответ: В

Обоснование: Классификация — это задача машинного обучения, в которой цель состоит в отнесении объектов к определенным категориям (классам) на основе входных данных. Например, определение спама в электронной почте или распознавание изображений.

Вопрос 2:

Какие из следующих методов относятся к методам классификации?

- А) Линейная регрессия.
- В) Логистическая регрессия.
- С) Метод k-ближайших соседей.
- Д) Деревья решений.

Ответ: В, С, D

Обоснование: Логистическая регрессия, метод k-ближайших соседей и деревья решений — это методы, которые используются для решения задач классификации. Линейная регрессия, напротив, используется для предсказания числовых значений и не подходит для классификации.

Вопрос 3:

Что такое бинарная классификация?

- А) Классификация объектов на два класса.
- В) Классификация объектов на три и более классов.
- С) Классификация объектов на основе их сходства.
- Д) Классификация объектов с использованием нейронных сетей.

Ответ: А

Обоснование: Бинарная классификация — это тип задачи классификации, в которой объекты относятся к одному из двух классов. Например, определение, является ли электронное письмо спамом или нет.

Вопрос 4:

Какие метрики используются для оценки качества классификаторов?

- A) Точность (Accuracy).
- B) F-мера (F-score).
- C) Точность (Precision).
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для оценки качества классификаторов используются различные метрики, включая:

Точность (Accuracy) — доля правильно классифицированных объектов.

F-мера (F-score) — взвешенное среднее между точностью и полнотой.

Точность (Precision) — доля правильно классифицированных объектов среди тех, которые были отнесены к определенному классу.

Вопрос 5:

Что такое переобучение в контексте классификации?

A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.

B) Когда модель недостаточно обучается на тренировочных данных.

C) Когда модель обучается только на малых данных.

D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение в классификации происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Тема 4. Методы обучения с учителем

Вопрос 1:

Что такое обучение с учителем в машинном обучении?

A) Процесс обучения модели на размеченных данных, где для каждого примера известен правильный ответ.

B) Процесс обучения модели на неразмеченных данных, где правильный ответ неизвестен.

C) Процесс обучения модели путем взаимодействия с окружающей средой и получения обратной связи.

D) Процесс обучения модели на данных без явного программирования.

Ответ: A

Обоснование: Обучение с учителем — это метод машинного обучения, при котором модель обучается на размеченных данных, где для каждого примера известен правильный ответ (целевая переменная). Это позволяет модели учиться предсказывать правильные ответы для новых данных.

Вопрос 2:

Какие из следующих методов относятся к методам обучения с учителем?

- A) Линейная регрессия.
- B) Логистическая регрессия.
- C) Метод k-ближайших соседей.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Все перечисленные методы относятся к методам обучения с учителем:

Линейная регрессия используется для предсказания числовых значений.

Логистическая регрессия используется для классификации.

Метод k-ближайших соседей также применяется для классификации и регрессии.

Вопрос 3:

Что такое переобучение в контексте обучения с учителем?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.
- B) Когда модель недостаточно обучается на тренировочных данных.
- C) Когда модель обучается только на малых данных.
- D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 4:

Какие методы используются для предотвращения переобучения в обучении с учителем?

- A) Регуляризация.
- B) Кросс-валидация.
- C) Ранняя остановка обучения.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для предотвращения переобучения в обучении с учителем используются различные методы:

Регуляризация — добавление штрафа за сложность модели.

Кросс-валидация — разделение данных на несколько частей для оценки модели.

Ранняя остановка обучения — прекращение обучения, когда модель перестает улучшать свои результаты на валидационных данных.

Вопрос 5:

Что такое градиентный спуск в контексте обучения с учителем?

- A) Метод оптимизации, используемый для минимизации ошибки модели.
- B) Метод, используемый для увеличения ошибки модели.
- C) Метод, используемый для обучения без учителя.
- D) Метод, используемый для обучения с подкреплением.

Ответ: A

Обоснование: Градиентный спуск — это метод оптимизации, который используется для минимизации ошибки модели в обучении с учителем. Он работает путем постепенного изменения параметров модели в направлении, противоположном градиенту функции ошибки, что позволяет находить оптимальные значения параметров.

Тема 5. Диагностика систем машинного обучения экранирования статических и динамических полей

Вопрос 1:

Что такое диагностика систем машинного обучения?

- A) Процесс оценки качества и производительности модели на этапе обучения.
- B) Процесс выявления и устранения ошибок в модели после ее развертывания.
- C) Процесс проверки данных на наличие ошибок перед обучением модели.
- D) Процесс оптимизации гиперпараметров модели.

Ответ: B

Обоснование: Диагностика систем машинного обучения — это процесс выявления и устранения ошибок в модели после ее развертывания. Это включает мониторинг производительности модели, выявление аномалий и принятие мер для улучшения ее работы.

Вопрос 2:

Какие из следующих методов используются для диагностики систем машинного обучения?

- A) Мониторинг метрик производительности.
- B) Анализ аномалий в данных.
- C) Валидация модели на новых данных.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для диагностики систем машинного обучения используются различные методы:

Мониторинг метрик производительности — отслеживание таких метрик, как точность, F-мера и другие.

Анализ аномалий в данных — выявление неожиданных или ошибочных данных, которые могут повлиять на модель.

Валидация модели на новых данных — проверка модели на данных, которые не использовались для обучения.

Вопрос 3:

Что такое переобучение в контексте диагностики систем машинного обучения?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.
- B) Когда модель недостаточно обучается на тренировочных данных.
- C) Когда модель обучается только на малых данных.
- D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее

предсказательную способность.

Вопрос 4:

Какие методы используются для диагностики переобучения в системах машинного обучения?

- A) Кросс-валидация.
- B) Мониторинг метрик на валидационных данных.
- C) Анализ распределения ошибок.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для диагностики переобучения используются различные методы:

Кросс-валидация — разделение данных на несколько частей для оценки модели.

Мониторинг метрик на валидационных данных — отслеживание метрик на данных, не использованных для обучения.

Анализ распределения ошибок — изучение ошибок модели для выявления закономерностей.

Вопрос 5:

Что такое дрейф данных в контексте диагностики систем машинного обучения?

- A) Постепенное изменение характеристик данных, на которых обучалась модель.
- B) Постепенное улучшение точности модели.
- C) Постепенное ухудшение точности модели.
- D) Постепенное увеличение объема данных.

Ответ: A

Обоснование: Дрейф данных — это постепенное изменение характеристик данных, на которых обучалась модель. Это может привести к снижению точности модели, так как данные, на которых она обучалась, могут стать неактуальными или измениться со временем.

Тема 6. Предварительный анализ и обработка данных

Вопрос 1:

Что такое предварительный анализ данных?

- A) Процесс сбора данных из различных источников.
- B) Процесс подготовки данных к обучению модели.
- C) Процесс анализа и обработки данных перед обучением модели.
- D) Процесс оценки качества модели после обучения.

Ответ: C

Обоснование: Предварительный анализ данных — это процесс анализа и обработки данных перед обучением модели. Он включает такие этапы, как очистка данных, нормализация, кодирование категориальных переменных и другие.

Вопрос 2:

Какие из следующих методов относятся к этапам предварительной обработки данных?

- A) Очистка данных от пропущенных значений.

- В) Нормализация данных.
 - С) Кодирование категориальных переменных.
 - Д) Все вышеперечисленное.
- Ответ: D

Обоснование: Предварительная обработка данных включает различные методы:

Очистка данных от пропущенных значений — удаление или заполнение пропущенных данных.

Нормализация данных — приведение данных к одному масштабу.

Кодирование категориальных переменных — преобразование категориальных данных в числовые.

Вопрос 3:

Что такое мультиколлинеарность в контексте предварительной обработки данных?

- А) Ситуация, когда одна из независимых переменных сильно коррелирует с зависимой переменной.
 - В) Ситуация, когда две или более независимых переменных сильно коррелируют друг с другом.
 - С) Ситуация, когда модель переобучается на тренировочных данных.
 - Д) Ситуация, когда модель недостаточно обучается на тренировочных данных.
- Ответ: B

Обоснование: Мультиколлинеарность возникает, когда две или более независимых переменных в модели сильно коррелируют друг с другом. Это может привести к проблемам с интерпретацией коэффициентов модели и снижению ее предсказательной способности.

Вопрос 4:

Какие методы используются для нормализации данных?

- А) Стандартизация (z-score).
 - В) Масштабирование в диапазон [0, 1].
 - С) Логарифмическое преобразование.
 - Д) Все вышеперечисленное.
- Ответ: D

Обоснование: Для нормализации данных используются различные методы:

Стандартизация (z-score) — приведение данных к нулевому среднему и единичной дисперсии.

Масштабирование в диапазон [0, 1] — приведение данных к диапазону от 0 до 1.

Логарифмическое преобразование — используется для уменьшения влияния выбросов и нормализации распределения данных.

Вопрос 5:

Что такое выбросы в данных и как с ними бороться?

- А) Выбросы — это данные, которые сильно отличаются от остальных, и их можно игнорировать.
- В) Выбросы — это данные, которые сильно отличаются от остальных, и их можно удалить или скорректировать.
- С) Выбросы — это данные, которые сильно отличаются от остальных, и их всегда нужно удалять.

D) Выбросы — это данные, которые сильно отличаются от остальных, и их всегда нужно корректировать.

Ответ: B

Обоснование: Выбросы — это данные, которые сильно отличаются от остальных. В зависимости от контекста, их можно удалить, скорректировать или оставить, если они представляют собой важные, но редкие события. Решение о том, как поступить с выбросами, зависит от конкретной задачи и данных.

Тема 7. Задачи обучения без учителя

Вопрос 1:

Что такое обучение без учителя в машинном обучении?

A) Процесс обучения модели на размеченных данных, где для каждого примера известен правильный ответ.

B) Процесс обучения модели на неразмеченных данных, где правильный ответ неизвестен.

C) Процесс обучения модели путем взаимодействия с окружающей средой и получения обратной связи.

D) Процесс обучения модели на данных без явного программирования.

Ответ: B

Обоснование: Обучение без учителя — это метод машинного обучения, при котором модель обучается на неразмеченных данных, где правильный ответ неизвестен. Это позволяет модели выявлять скрытые закономерности и структуры в данных.

Вопрос 2:

Какие из следующих методов относятся к задачам обучения без учителя?

A) Кластеризация.

B) Ассоциативные правила.

C) Снижение размерности.

D) Все вышеперечисленное.

Ответ: D

Обоснование: К задачам обучения без учителя относятся:

Кластеризация — группировка объектов на основе их сходства.

Ассоциативные правила — выявление закономерностей и связей между элементами в данных.

Снижение размерности — уменьшение количества признаков в данных без значительной потери информации.

Вопрос 3:

Что такое кластеризация в контексте обучения без учителя?

A) Задача отнесения объектов к определенным категориям.

B) Задача группировки объектов на основе их сходства.

C) Задача предсказания числовых значений на основе входных данных.

D) Задача обучения с подкреплением.

Ответ: B

Обоснование: Кластеризация — это задача обучения без учителя, в которой цель

состоит в группировке объектов на основе их сходства. Это позволяет выявлять скрытые структуры и закономерности в данных.

Вопрос 4:

Какие методы используются для кластеризации?

- A) Метод k-средних.
- B) Иерархическая кластеризация.
- C) DBSCAN.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для кластеризации используются различные методы:

Метод k-средних — алгоритм, который разделяет данные на k кластеров.

Иерархическая кластеризация — метод, который создает иерархию кластеров.

DBSCAN — алгоритм, который находит кластеры на основе плотности данных.

Вопрос 5:

Что такое снижение размерности в контексте обучения без учителя?

- A) Процесс увеличения количества признаков в данных.
 - B) Процесс уменьшения количества признаков в данных без значительной потери информации.
 - C) Процесс удаления всех признаков из данных.
 - D) Процесс добавления новых признаков в данные.
- Ответ: B

Обоснование: Снижение размерности — это процесс уменьшения количества признаков в данных без значительной потери информации. Это позволяет упростить модель и улучшить ее производительность, а также уменьшить вычислительную сложность.

Тема 8. Практическое использование моделей машинного обучения

Вопрос 1:

Какие области применения машинного обучения наиболее распространены?

- A) Медицина и здравоохранение.
- B) Финансы и банковское дело.
- C) Маркетинг и реклама.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Машинное обучение находит применение в самых разных областях, включая:

Медицину и здравоохранение — для диагностики заболеваний и анализа медицинских данных.

Финансы и банковское дело — для оценки кредитоспособности и выявления мошенничества.

Маркетинг и рекламу — для сегментации клиентов и прогнозирования продаж.

Вопрос 2:

Какие задачи машинного обучения используются в медицине?

- A) Диагностика заболеваний.
- B) Прогнозирование исходов лечения.
- C) Анализ медицинских изображений.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: В медицине машинное обучение применяется для:

Диагностики заболеваний — например, распознавание рака на основе медицинских изображений.

Прогнозирования исходов лечения — предсказание вероятности выздоровления или осложнений.

Анализа медицинских изображений — автоматическое выявление патологий на рентгеновских снимках или МРТ.

Вопрос 3:

Какие задачи машинного обучения используются в финансах?

- A) Оценка кредитоспособности.
- B) Выявление мошенничества.
- C) Прогнозирование цен на акции.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: В финансах машинное обучение применяется для:

Оценки кредитоспособности — предсказание вероятности возврата кредита.

Выявления мошенничества — обнаружение подозрительных транзакций.

Прогнозирования цен на акции — предсказание движения рынка на основе исторических данных.

Вопрос 4:

Какие задачи машинного обучения используются в маркетинге?

- A) Сегментация клиентов.
- B) Прогнозирование продаж.
- C) Персонализация рекламы.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: В маркетинге машинное обучение применяется для:

Сегментации клиентов — разделение аудитории на группы с похожими характеристиками.

Прогнозирования продаж — предсказание объемов продаж на основе исторических данных.

Персонализации рекламы — создание индивидуальных предложений для каждого клиента.

Вопрос 5:

Какие задачи машинного обучения используются в социальных сетях?

- A) Анализ и категоризация контента.
- B) Рекомендация контента.
- C) Выявление фейковых новостей.

D) Все вышеперечисленное.

Ответ: D

Обоснование: В социальных сетях машинное обучение применяется для:

Анализа и категоризации контента — определение тематики публикаций и комментариев.

Рекомендации контента — предложение пользователям интересных им постов и видео.

Выявления фейковых новостей — автоматическое обнаружение недостоверной информации.

Критерии оценивания тестовых заданий для текущего контроля

Максимальное количество баллов за решение одного теста составляет 5 баллов, которые переводятся в академические оценки следующим образом:

- оценка «отлично» (5 баллов) выставляется при условии правильного ответа студента не менее чем на 85% контрольных заданий

- оценка «хорошо» (4 балла) выставляется при условии правильного ответа студента не менее чем на 70% контрольных заданий

- оценка «удовлетворительно» (3 балла) выставляется при условии правильного ответа студента не менее чем на 51% контрольных заданий

- оценка «неудовлетворительно» (1-2 балла) выставляется при условии правильного ответа студента менее чем на 50% контрольных заданий

2.2 Оценочные материалы для проведения промежуточной аттестации

Вопрос 1:

Что такое машинное обучение?

A) Процесс, при котором компьютерные программы самостоятельно обучаются на данных без явного программирования.

B) Процесс, при котором компьютерные программы обучаются только на заранее заданных правилах.

C) Процесс, при котором компьютерные программы обучаются исключительно на больших данных.

D) Процесс, при котором компьютерные программы обучаются только на малых данных.

Ответ: A

Обоснование: Машинное обучение — это область искусственного интеллекта, которая позволяет компьютерным программам обучаться на данных и улучшать свои результаты без явного программирования. Это означает, что алгоритмы машинного обучения могут автоматически выявлять закономерности и делать прогнозы на основе данных, не требуя ручного программирования для каждого конкретного случая.

Вопрос 2:

Какие основные задачи решает машинное обучение?

A) Регрессия и классификация.

B) Кластеризация и регрессия.

C) Классификация и кластеризация.

D) Все вышеперечисленное.

Ответ: D

Обоснование: Машинное обучение решает широкий спектр задач, включая:

Регрессию — предсказание числовых значений (например, прогнозирование цены на недвижимость).

Классификация — отнесение объектов к определенным категориям (например, определение спама в электронной почте).

Кластеризация — группировка объектов на основе их сходства (например, сегментация клиентов по их поведению).

Вопрос 3:

Какие методы обучения используются в машинном обучении?

A) Обучение с учителем и обучение без учителя.

B) Обучение с подкреплением и обучение с учителем.

C) Обучение без учителя и обучение с подкреплением.

D) Все вышеперечисленное.

Ответ: D

Обоснование: В машинном обучении используются три основных метода:

Обучение с учителем — когда алгоритм обучается на размеченных данных (например, классификация изображений).

Обучение без учителя — когда алгоритм обучается на неразмеченных данных (например, кластеризация).

Обучение с подкреплением — когда алгоритм обучается путем взаимодействия с окружающей средой и получения обратной связи (например, обучение роботов).

Вопрос 4:

Какая из следующих областей не относится к машинному обучению?

A) Глубокое обучение.

B) Нейронные сети.

C) Линейная регрессия.

D) Теория игр.

Ответ: D

Обоснование: Теория игр — это область математики, которая изучает стратегическое взаимодействие между игроками. Она не относится к машинному обучению, в отличие от:

Глубокого обучения — подкатегории машинного обучения, использующей нейронные сети с большим количеством слоев.

Нейронных сетей — математических моделей, вдохновленных работой мозга, которые используются в машинном обучении.

Линейной регрессии — метода машинного обучения для предсказания числовых значений.

Вопрос 5:

Что такое переобучение в машинном обучении?

A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.

B) Когда модель недостаточно обучается на тренировочных данных.

- С) Когда модель обучается только на малых данных.
 - Д) Когда модель обучается только на больших данных.
- Ответ: А

Обоснование: Переобучение происходит, когда модель машинного обучения слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 6:
Что такое регрессия в машинном обучении?

- А) Задача классификации объектов на несколько категорий.
 - В) Задача предсказания числовых значений на основе входных данных.
 - С) Задача группировки объектов на основе их сходства.
 - Д) Задача обучения с подкреплением.
- Ответ: В

Обоснование: Регрессия — это задача машинного обучения, в которой цель состоит в предсказании числовых значений (непрерывных величин) на основе входных данных. Например, предсказание цены на недвижимость или прогнозирование температуры на следующий день.

Вопрос 7:
Какие из следующих методов относятся к регрессионным моделям?

- А) Линейная регрессия.
 - В) Логистическая регрессия.
 - С) Метод k-ближайших соседей.
 - Д) Деревья решений.
- Ответ: А

Обоснование: Линейная регрессия — это классический метод регрессии, используемый для предсказания числовых значений на основе линейной зависимости между входными и выходными переменными. Логистическая регрессия, несмотря на название, используется для классификации, а не для регрессии. Метод k-ближайших соседей и деревья решений могут использоваться как для классификации, так и для регрессии, но в данном контексте правильный ответ — линейная регрессия.

Вопрос 8:
Что такое мультиколлинеарность в регрессионном анализе?

- А) Ситуация, когда одна из независимых переменных сильно коррелирует с зависимой переменной.
 - В) Ситуация, когда две или более независимых переменных сильно коррелируют друг с другом.
 - С) Ситуация, когда модель переобучается на тренировочных данных.
 - Д) Ситуация, когда модель недостаточно обучается на тренировочных данных.
- Ответ: В

Обоснование: Мультиколлинеарность возникает, когда две или более независимых переменных в регрессионной модели сильно коррелируют друг с другом. Это может

привести к проблемам с интерпретацией коэффициентов модели и снижению ее предсказательной способности.

Вопрос 9:

Какая из следующих метрик используется для оценки качества регрессионной модели?

- A) Точность (Accuracy).
- B) F-мера (F-score).
- C) Среднее абсолютное отклонение (Mean Absolute Error, MAE).
- D) Точность (Precision).

Ответ: C

Обоснование: Для оценки качества регрессионных моделей используются метрики, которые измеряют разницу между предсказанными и фактическими значениями. Среднее абсолютное отклонение (MAE) — это одна из таких метрик, которая вычисляет среднее абсолютное значение ошибок предсказания. Точность (Accuracy), F-мера (F-score) и Точность (Precision) используются для оценки качества классификационных моделей.

Вопрос 10:

Что такое переобучение в контексте регрессионного анализа?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.
- B) Когда модель недостаточно обучается на тренировочных данных.
- C) Когда модель обучается только на малых данных.
- D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение в регрессионном анализе происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 11:

Что такое классификация в машинном обучении?

- A) Задача предсказания числовых значений на основе входных данных.
- B) Задача отнесения объектов к определенным категориям.
- C) Задача группировки объектов на основе их сходства.
- D) Задача обучения с подкреплением.

Ответ: B

Обоснование: Классификация — это задача машинного обучения, в которой цель состоит в отнесении объектов к определенным категориям (классам) на основе входных данных. Например, определение спама в электронной почте или распознавание изображений.

Вопрос 12:

Какие из следующих методов относятся к методам классификации?

- A) Линейная регрессия.
- B) Логистическая регрессия.
- C) Метод k-ближайших соседей.

D) Деревья решений.

Ответ: B, C, D

Обоснование: Логистическая регрессия, метод k-ближайших соседей и деревья решений — это методы, которые используются для решения задач классификации. Линейная регрессия, напротив, используется для предсказания числовых значений и не подходит для классификации.

Вопрос 13:

Что такое бинарная классификация?

A) Классификация объектов на два класса.

B) Классификация объектов на три и более классов.

C) Классификация объектов на основе их сходства.

D) Классификация объектов с использованием нейронных сетей.

Ответ: A

Обоснование: Бинарная классификация — это тип задачи классификации, в которой объекты относятся к одному из двух классов. Например, определение, является ли электронное письмо спамом или нет.

Вопрос 14:

Какие метрики используются для оценки качества классификаторов?

A) Точность (Accuracy).

B) F-мера (F-score).

C) Точность (Precision).

D) Все вышеперечисленное.

Ответ: D

Обоснование: Для оценки качества классификаторов используются различные метрики, включая:

Точность (Accuracy) — доля правильно классифицированных объектов.

F-мера (F-score) — взвешенное среднее между точностью и полнотой.

Точность (Precision) — доля правильно классифицированных объектов среди тех, которые были отнесены к определенному классу.

Вопрос 15:

Что такое переобучение в контексте классификации?

A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.

B) Когда модель недостаточно обучается на тренировочных данных.

C) Когда модель обучается только на малых данных.

D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение в классификации происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 16:

Что такое обучение с учителем в машинном обучении?

- A) Процесс обучения модели на размеченных данных, где для каждого примера известен правильный ответ.
- B) Процесс обучения модели на неразмеченных данных, где правильный ответ неизвестен.
- C) Процесс обучения модели путем взаимодействия с окружающей средой и получения обратной связи.
- D) Процесс обучения модели на данных без явного программирования.

Ответ: A

Обоснование: Обучение с учителем — это метод машинного обучения, при котором модель обучается на размеченных данных, где для каждого примера известен правильный ответ (целевая переменная). Это позволяет модели учиться предсказывать правильные ответы для новых данных.

Вопрос 17:

Какие из следующих методов относятся к методам обучения с учителем?

- A) Линейная регрессия.
- B) Логистическая регрессия.
- C) Метод k-ближайших соседей.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Все перечисленные методы относятся к методам обучения с учителем:

Линейная регрессия используется для предсказания числовых значений.

Логистическая регрессия используется для классификации.

Метод k-ближайших соседей также применяется для классификации и регрессии.

Вопрос 18:

Что такое переобучение в контексте обучения с учителем?

- A) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.
- B) Когда модель недостаточно обучается на тренировочных данных.
- C) Когда модель обучается только на малых данных.
- D) Когда модель обучается только на больших данных.

Ответ: A

Обоснование: Переобучение происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 19:

Какие методы используются для предотвращения переобучения в обучении с учителем?

- A) Регуляризация.
- B) Кросс-валидация.
- C) Ранняя остановка обучения.

D) Все вышеперечисленное.

Ответ: D

Обоснование: Для предотвращения переобучения в обучении с учителем используются различные методы:

Регуляризация — добавление штрафа за сложность модели.

Кросс-валидация — разделение данных на несколько частей для оценки модели.

Ранняя остановка обучения — прекращение обучения, когда модель перестает улучшать свои результаты на валидационных данных.

Вопрос 20:

Что такое градиентный спуск в контексте обучения с учителем?

A) Метод оптимизации, используемый для минимизации ошибки модели.

B) Метод, используемый для увеличения ошибки модели.

C) Метод, используемый для обучения без учителя.

D) Метод, используемый для обучения с подкреплением.

Ответ: A

Обоснование: Градиентный спуск — это метод оптимизации, который используется для минимизации ошибки модели в обучении с учителем. Он работает путем постепенного изменения параметров модели в направлении, противоположном градиенту функции ошибки, что позволяет находить оптимальные значения параметров.

Тема 5. Диагностика систем машинного обучения экранирования статических и динамических полей

Вопрос 21:

Что такое диагностика систем машинного обучения?

A) Процесс оценки качества и производительности модели на этапе обучения.

B) Процесс выявления и устранения ошибок в модели после ее развертывания.

C) Процесс проверки данных на наличие ошибок перед обучением модели.

D) Процесс оптимизации гиперпараметров модели.

Ответ: B

Обоснование: Диагностика систем машинного обучения — это процесс выявления и устранения ошибок в модели после ее развертывания. Это включает мониторинг производительности модели, выявление аномалий и принятие мер для улучшения ее работы.

Вопрос 22:

Какие из следующих методов используются для диагностики систем машинного обучения?

A) Мониторинг метрик производительности.

B) Анализ аномалий в данных.

C) Валидация модели на новых данных.

D) Все вышеперечисленное.

Ответ: D

Обоснование: Для диагностики систем машинного обучения используются различные методы:

Мониторинг метрик производительности — отслеживание таких метрик, как точность, F-мера и другие.

Анализ аномалий в данных — выявление неожиданных или ошибочных данных, которые могут повлиять на модель.

Валидация модели на новых данных — проверка модели на данных, которые не использовались для обучения.

Вопрос 23:

Что такое переобучение в контексте диагностики систем машинного обучения?

А) Когда модель слишком хорошо обучается на тренировочных данных и плохо обобщает на новых данных.

В) Когда модель недостаточно обучается на тренировочных данных.

С) Когда модель обучается только на малых данных.

Д) Когда модель обучается только на больших данных.

Ответ: А

Обоснование: Переобучение происходит, когда модель слишком хорошо подстраивается под тренировочные данные, включая их шум и случайные колебания. В результате модель плохо обобщает на новых, ранее не виденных данных, что снижает ее предсказательную способность.

Вопрос 24:

Какие методы используются для диагностики переобучения в системах машинного обучения?

А) Кросс-валидация.

В) Мониторинг метрик на валидационных данных.

С) Анализ распределения ошибок.

Д) Все вышеперечисленное.

Ответ: D

Обоснование: Для диагностики переобучения используются различные методы:

Кросс-валидация — разделение данных на несколько частей для оценки модели.

Мониторинг метрик на валидационных данных — отслеживание метрик на данных, не использованных для обучения.

Анализ распределения ошибок — изучение ошибок модели для выявления закономерностей.

Вопрос 25:

Что такое дрейф данных в контексте диагностики систем машинного обучения?

А) Постепенное изменение характеристик данных, на которых обучалась модель.

В) Постепенное улучшение точности модели.

С) Постепенное ухудшение точности модели.

Д) Постепенное увеличение объема данных.

Ответ: А

Обоснование: Дрейф данных — это постепенное изменение характеристик данных, на которых обучалась модель. Это может привести к снижению точности модели, так как данные, на которых она обучалась, могут стать неактуальными или измениться со временем.

Вопрос 26:

Что такое предварительный анализ данных?

- A) Процесс сбора данных из различных источников.
- B) Процесс подготовки данных к обучению модели.
- C) Процесс анализа и обработки данных перед обучением модели.
- D) Процесс оценки качества модели после обучения.

Ответ: C

Обоснование: Предварительный анализ данных — это процесс анализа и обработки данных перед обучением модели. Он включает такие этапы, как очистка данных, нормализация, кодирование категориальных переменных и другие.

Вопрос 27:

Какие из следующих методов относятся к этапам предварительной обработки данных?

- A) Очистка данных от пропущенных значений.
- B) Нормализация данных.
- C) Кодирование категориальных переменных.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Предварительная обработка данных включает различные методы:

Очистка данных от пропущенных значений — удаление или заполнение пропущенных данных.

Нормализация данных — приведение данных к одному масштабу.

Кодирование категориальных переменных — преобразование категориальных данных в числовые.

Вопрос 28:

Что такое мультиколлинеарность в контексте предварительной обработки данных?

- A) Ситуация, когда одна из независимых переменных сильно коррелирует с зависимой переменной.
- B) Ситуация, когда две или более независимых переменных сильно коррелируют друг с другом.
- C) Ситуация, когда модель переобучается на тренировочных данных.
- D) Ситуация, когда модель недостаточно обучается на тренировочных данных.

Ответ: B

Обоснование: Мультиколлинеарность возникает, когда две или более независимых переменных в модели сильно коррелируют друг с другом. Это может привести к проблемам с интерпретацией коэффициентов модели и снижению ее предсказательной способности.

Вопрос 29:

Какие методы используются для нормализации данных?

- A) Стандартизация (z-score).
- B) Масштабирование в диапазон $[0, 1]$.
- C) Логарифмическое преобразование.
- D) Все вышеперечисленное.

Ответ: D

Обоснование: Для нормализации данных используются различные методы:

Стандартизация (z-score) — приведение данных к нулевому среднему и единичной дисперсии.

Масштабирование в диапазон [0, 1] — приведение данных к диапазону от 0 до 1.

Логарифмическое преобразование — используется для уменьшения влияния выбросов и нормализации распределения данных.

Вопрос 30:

Что такое выбросы в данных и как с ними бороться?

A) Выбросы — это данные, которые сильно отличаются от остальных, и их можно игнорировать.

B) Выбросы — это данные, которые сильно отличаются от остальных, и их можно удалить или скорректировать.

C) Выбросы — это данные, которые сильно отличаются от остальных, и их всегда нужно удалять.

D) Выбросы — это данные, которые сильно отличаются от остальных, и их всегда нужно корректировать.

Ответ: B

Обоснование: Выбросы — это данные, которые сильно отличаются от остальных. В зависимости от контекста, их можно удалить, скорректировать или оставить, если они представляют собой важные, но редкие события. Решение о том, как поступить с выбросами, зависит от конкретной задачи и данных.

Вопрос 31:

Что такое обучение без учителя в машинном обучении?

A) Процесс обучения модели на размеченных данных, где для каждого примера известен правильный ответ.

B) Процесс обучения модели на неразмеченных данных, где правильный ответ неизвестен.

C) Процесс обучения модели путем взаимодействия с окружающей средой и получения обратной связи.

D) Процесс обучения модели на данных без явного программирования.

Ответ: B

Обоснование: Обучение без учителя — это метод машинного обучения, при котором модель обучается на неразмеченных данных, где правильный ответ неизвестен. Это позволяет модели выявлять скрытые закономерности и структуры в данных.

Вопрос 32:

Какие из следующих методов относятся к задачам обучения без учителя?

A) Кластеризация.

B) Ассоциативные правила.

C) Снижение размерности.

D) Все вышеперечисленное.

Ответ: D

Обоснование: К задачам обучения без учителя относятся:

Кластеризация — группировка объектов на основе их сходства.

Ассоциативные правила — выявление закономерностей и связей между элементами в данных.

Снижение размерности — уменьшение количества признаков в данных без значительной потери информации.

Вопрос 33:

Что такое кластеризация в контексте обучения без учителя?

A) Задача отнесения объектов к определенным категориям.

B) Задача группировки объектов на основе их сходства.

C) Задача предсказания числовых значений на основе входных данных.

D) Задача обучения с подкреплением.

Ответ: B

Обоснование: Кластеризация — это задача обучения без учителя, в которой цель состоит в группировке объектов на основе их сходства. Это позволяет выявлять скрытые структуры и закономерности в данных.

Вопрос 34:

Какие методы используются для кластеризации?

A) Метод k-средних.

B) Иерархическая кластеризация.

C) DBSCAN.

D) Все вышеперечисленное.

Ответ: D

Обоснование: Для кластеризации используются различные методы:

Метод k-средних — алгоритм, который разделяет данные на k кластеров.

Иерархическая кластеризация — метод, который создает иерархию кластеров.

DBSCAN — алгоритм, который находит кластеры на основе плотности данных.

Вопрос 35:

Что такое снижение размерности в контексте обучения без учителя?

A) Процесс увеличения количества признаков в данных.

B) Процесс уменьшения количества признаков в данных без значительной потери информации.

C) Процесс удаления всех признаков из данных.

D) Процесс добавления новых признаков в данные.

Ответ: B

Обоснование: Снижение размерности — это процесс уменьшения количества признаков в данных без значительной потери информации. Это позволяет упростить модель и улучшить ее производительность, а также уменьшить вычислительную сложность.

Вопрос 36:

Какие области применения машинного обучения наиболее распространены?

A) Медицина и здравоохранение.

- В) Финансы и банковское дело.
- С) Маркетинг и реклама.
- Д) Все вышеперечисленное.

Ответ: D

Обоснование: Машинное обучение находит применение в самых разных областях, включая:

Медицину и здравоохранение — для диагностики заболеваний и анализа медицинских данных.

Финансы и банковское дело — для оценки кредитоспособности и выявления мошенничества.

Маркетинг и рекламу — для сегментации клиентов и прогнозирования продаж.

Вопрос 37:

Какие задачи машинного обучения используются в медицине?

- А) Диагностика заболеваний.
- В) Прогнозирование исходов лечения.
- С) Анализ медицинских изображений.
- Д) Все вышеперечисленное.

Ответ: D

Обоснование: В медицине машинное обучение применяется для:

Диагностики заболеваний — например, распознавание рака на основе медицинских изображений.

Прогнозирования исходов лечения — предсказание вероятности выздоровления или осложнений.

Анализа медицинских изображений — автоматическое выявление патологий на рентгеновских снимках или МРТ.

Вопрос 38:

Какие задачи машинного обучения используются в финансах?

- А) Оценка кредитоспособности.
- В) Выявление мошенничества.
- С) Прогнозирование цен на акции.
- Д) Все вышеперечисленное.

Ответ: D

Обоснование: В финансах машинное обучение применяется для:

Оценки кредитоспособности — предсказание вероятности возврата кредита.

Выявления мошенничества — обнаружение подозрительных транзакций.

Прогнозирования цен на акции — предсказание движения рынка на основе исторических данных.

Вопрос 39:

Какие задачи машинного обучения используются в маркетинге?

- А) Сегментация клиентов.
- В) Прогнозирование продаж.
- С) Персонализация рекламы.

D) Все вышеперечисленное.

Ответ: D

Обоснование: В маркетинге машинное обучение применяется для:

Сегментации клиентов — разделение аудитории на группы с похожими характеристиками.
Прогнозирования продаж — предсказание объемов продаж на основе исторических данных.

Персонализации рекламы — создание индивидуальных предложений для каждого клиента.

Вопрос 40:

Какие задачи машинного обучения используются в социальных сетях?

A) Анализ и категоризация контента.

B) Рекомендация контента.

C) Выявление фейковых новостей.

D) Все вышеперечисленное.

Ответ: D

Обоснование: В социальных сетях машинное обучение применяется для:

Анализа и категоризации контента — определение тематики публикаций и комментариев.

Рекомендации контента — предложение пользователям интересных им постов и видео.

Выявления фейковых новостей — автоматическое обнаружение недостоверной информации.

Критерии оценки для проведения зачета по дисциплине (тестирование)

Шкала оценивания результатов промежуточной аттестации и критерии выставления оценок.

Оценка	Уровень сформированности компетенции	Критерии
«Отлично» («зачтено»)	Высокий	Выполнено более 80% заданий предложенного теста
«Хорошо» («зачтено»)	Хороший	Выполнено 60-80% заданий предложенного теста
«Удовлетворительно» («зачтено»)	Достаточный	Выполнено 35-60% заданий предложенного теста
«Неудовлетворительно» («не зачтено»)	Недостаточный	Выполнено менее 35% заданий предложенного теста

Обновление оценочных материалов дисциплины

обсуждено и рекомендовано к утверждению решением кафедры
наименование кафедры _____
от ____ ____ 20__ г., протокол № ____

одобрено Научно-методическим советом _____ от ____ ____ 20__ г.,
протокол № ____