Автономная некоммерческая образовательная организация
высшего образования
«Воронежский экономико-правовой институт»
(АНОО ВО «ВЭПИ»)

АНОО ВО "ВОРОНЕЖСКИЙ ЭКОНОМИКО-ПРАВОВОЙ ИНСТИТУТ", АНОО ВО "ВЭПИ", ВЭПИ

Подписано цифровой подписью: АНОО ВО "ВОРОНЕЖСКИЙ ЭКОНОМИКО-ПРАВОВОЙ ИНСТИТУТ", АНОО ВО "ВЭПИ", ВЭПИ
Дата: 2024.08.31 17:54:23 +03'00'

УТВЕРЖДАЮ
Проректор
по учебно-методической работе
А.Ю. Жильников
«__» _____ 2023 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Б1.О.36 Системы искусственного интеллекта
(наименование дисциплины (модуля))

37.03.01 Психология
(код и наименование направления подготовки)

Направленность (профиль) Психолог в сфере образования
(наименование направленности (профиля))

Квалификация выпускника Бакалавр
(наименование квалификации)

Форма обучения очная, очно-заочная
(очная, очно-заочная, заочная)

Рекомендован к использованию филиалами АНОО ВО «ВЭПИ»

Воронеж 2023

Фонд оценочных средств по дисциплине (модулю) рассмотрен и одобрен на заседании кафедры прикладной информатики

Протокол от «01» сентября 2023 г. № 1

Фонд оценочных средств по дисциплине (модулю) согласован со следующими представителями работодателей или их объединений, направление деятельности которых соответствует области профессиональной деятельности, к которой готовятся обучающиеся:

Директор Муниципального бюджетного образовательного учреждения «Средняя общеобразовательная школа №102»
Фактор А.М.
(должность, наименование организации, фамилия, инициалы, подпись, дата, печать)

Директор Автономной некоммерческой профессиональной образовательной организации «Региональный экономико-правовой колледж»
Чернусских Ю.Л.
(должность, наименование организации, фамилия, инициалы, подпись, дата, печать)

Заведующий кафедрой М.С. Агафонова

Разработчики:

профессор М.С. Агафонова

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОП ВО

Целью проведения дисциплины Б1.О.36 «Системы искусственного интеллекта» является достижение следующих результатов обучения:

Код компетенции	Наименование компетенции
ОПК-9	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности

В формировании данных компетенций также участвуют следующие дисциплины (модули), практики образовательной программы (по семестрам (курсам) их изучения):

- для очной формы обучения:

Наименование дисциплин (модулей), практик	Этапы формирования компетенций по семестрам изучения							
	1 сем.	2 сем.	3 сем.	4 сем.	5 сем.	6 сем.	7 сем.	8 сем.
Информационные технологии в психологии		ОПК-9						
Математические методы в психологии			ОПК-9					
Учебная практика (учебно-ознакомительная практика)				ОПК-9				
Производственная практика (преддипломная практика)								ОПК-9

- для очно-заочной формы обучения:

Наименование дисциплин (модулей), практик	Этапы формирования компетенций по семестрам изучения									
	1 сем.	2 сем.	3 сем.	4 сем.	5 сем.	6 сем.	7 сем.	8 сем.	9 сем.	Сем. А
Информационные технологии в психологии		ОПК-9								
Математические методы в психологии				ОПК-9						
Учебная практика (учебно-ознакомительная практика)						ОПК-9				
Производственная практика (преддипломная практика)										ОПК-9

Этап дисциплины (модуля) Б1.О.36 «Системы искусственного интеллекта» соответствует:

- для очной формы обучения – 5 семестру;
- для очно-заочной формы обучения – 5 семестру.

2. Показатели и критерии оценивания компетенций на различных этапах их формирования, шкалы оценивания

Показателями оценивания компетенций являются следующие результаты обучения:

Код компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине (модулю)
ОПК-9.	ИОПК-9.1. Понимает принципы работы современных информационных технологий	Знать: принципы работы современных информационных технологий искусственного интеллекта в профессиональной деятельности Уметь: применять принципы работы современных информационных технологий искусственного интеллекта для решения задач профессиональной деятельности Владеть: навыками применения принципов работы современных информационных технологий искусственного интеллекта для решения задач профессиональной деятельности
	ИОПК-9.2. Применяет принципы работы современных информационных технологий для решения задач профессиональной деятельности	

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины (модуля):

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
1	Тема 1. Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными.	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			информационной безопасности при решении задач профессиональной деятельности.		
2	Тема 2. Основные задачи систем искусственного интеллекта.	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено
3	Тема 3. Классификация на примере алгоритма k-ближайших соседей (kNN)	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной без-	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			опасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.		
4	Тема 4. Регрессия. Метрики оценки регрессии: MSE, MAE, R2 – коэффициент детерминации.	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено
5	Тема 5. Кластеризация k-means, k-means++, DBSCAN, агломеративная кластеризация Метрики оценки кластеризации	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информацион-	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			ной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.		
6	Тема 6. Алгоритмы, основанные на применении решающих деревьев	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			информационной безопасности при решении задач профессиональной деятельности.		
7	Тема 7. Метод опорных векторов. Прямая и обратная задача. Определение опорных векторов. Ядерный трюк.	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено
8	Тема 8. Наивный байесовский классификатор	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.		
9	Тема 9. Методы безградиентной оптимизации: случайный поиск, hill climb, отжиг, генетический алгоритм	ОПК-9 ИОПК-9.1 ИОПК-9.2	знать: − современные информационные технологии и программные средства, в том числе отечественного производства при решении задач информационной безопасности; уметь: − правильно выбирать и применять современные информационные технологии и программные средства, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности; Владеть: − навыками выбора и применения современных информационных технологий и программных средств, в том числе отечественного производства для обеспечения информационной безопасности при решении задач профессиональной деятельности.	Устный опрос, доклад, типовые задания, лабораторная работа	- зачтено - не зачтено
ИТОГО			Форма контроля	Оценочные средства промежуточной аттестации	Шкала оценивания

№ п/п	Наименование раздела дисциплины, темы (модуля)	Код компетенции, код индикатора достижения компетенции	Критерии оценивания	Оценочные средства текущего контроля успеваемости	Шкала оценивания
			Зачет	ответ на билет	- зачтено - не зачтено

Критерии оценивания результатов обучения для текущего контроля успеваемости и промежуточной аттестации по дисциплине (модулю)

1. Критерий оценивания устного ответа:

Зачтено – хорошее знание основных терминов и понятий курса, последовательное изложение материала курса, умение формулировать некоторые обобщения по теме вопросов, достаточно полные ответы на вопросы, умение использовать фундаментальные понятия из базовых дисциплин при ответе.

Не зачтено – не выполнены требования, соответствующие оценке «зачтено».

2. Критерии оценивания доклада:

Зачтено – содержание основано на глубоком и всестороннем знании темы, изученной литературы, изложено логично, аргументировано и в полном объеме, основные понятия, выводы и обобщения сформулированы убедительно и доказательно, возможны недостатки в систематизации или в обобщении материала, неточности в выводах, основные категории применяются для изложения материала.

Не зачтено – не выполнены требования, соответствующие оценке «зачтено».

3. Критерии оценивания тестирования:

Оценка «отлично» – 86 % – 100 % правильных ответов.

Оценка «хорошо» – 70 % – 85 % правильных ответов.

Оценка «удовлетворительно» – 51 % – 69 % правильных ответов.

Оценка «неудовлетворительно» – 50 % и менее правильных ответов.

4. Критерии оценивания решения ситуационных задач:

Зачтено – ответ на вопрос задачи дан правильный, объяснение хода её решения подробное, последовательное, грамотное, с теоретическими обоснованиями или решение подробное, но недостаточно логичное, с единичными ошибками в деталях, некоторыми затруднениями в теоретическом обосновании, или ответ на вопрос задачи дан правильный, объяснение хода её решения недостаточно полное, непоследовательное, с ошибками, слабым теоретическим обоснованием.

Не зачтено – не выполнены требования, соответствующие оценке «зачтено».

5. Критерии оценивания ответа на зачете:

Зачтено – выставляется обучающемуся, если он продемонстрировал знание основного теоретического содержания дисциплин учебного плана образовательной программы высшего образования, умение показать уровень сформированности практических профессиональных умений и навыков, способность четко и аргументировано отвечать на дополнительные вопросы.

Не зачтено – не выполнены требования, соответствующие оценке «зачтено».

3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

1 ЭТАП – Текущий контроль освоения дисциплины

3.1. «Вопросы для устного опроса»:

1. Что включает в себя понятие "экономическая безопасность" и как она связана с искусственным интеллектом?
2. Какие области экономической безопасности наиболее подвержены угрозам, связанным с ИИ?
3. Какие типы систем искусственного интеллекта применяются для повышения экономической безопасности?
4. Какие методы машинного обучения чаще всего применяются для выявления финансовых мошенничеств?
5. Как использование ИИ помогает в управлении рисками в экономике?
6. Какие данные являются наиболее важными для работы систем ИИ в области экономической безопасности?
7. Как регулируется использование ИИ в экономической безопасности на государственном уровне?
8. Какие угрозы безопасности могут возникнуть при использовании систем ИИ в экономике?
9. В чем состоит роль искусственного интеллекта в прогнозировании финансовых рисков?
10. Какие моральные и этические проблемы могут возникнуть при использовании ИИ в экономической безопасности?
11. Какие есть способы защиты данных, используемых в системах ИИ для экономической безопасности?
12. Как применение искусственного интеллекта влияет на скорость и эффективность обработки финансовых данных?

13. В каких случаях системы ИИ могут ошибочно выявить угрозу безопасности?

14. Как искусственный интеллект может помочь в анализе крупных массивов данных для выявления финансовых преступлений?

15. Какие примеры успешного применения ИИ в борьбе с экономическими преступлениями вы знаете?

16. Какие преимущества и недостатки использования систем ИИ для обеспечения экономической безопасности?

17. Какие существуют методы противодействия кибератакам с применением систем ИИ?

18. Как системы ИИ могут повысить прозрачность финансовых операций в экономике?

19. Какие ограничения накладываются на системы ИИ при использовании их в экономике?

20. Каким образом можно улучшить технологии ИИ для более эффективного обеспечения экономической безопасности?

21. Методы машинного обучения для работы с табличными данными.

22. Задачи классификации в контексте работы с табличными данными.

23. Задачи регрессии в контексте работы с табличными данными.

24. Основные задачи искусственного интеллекта.

25. Применение искусственного интеллекта.

26. Работа алгоритма k-ближайших соседей (kNN) в задачах классификации.

27. Принципы работы алгоритма k-ближайших соседей (kNN) в задачах классификации.

28. Выбор оптимального значения параметра k при использовании алгоритма kNN.

29. Качество модели регрессии в предоставлении метрик MSE, MAE и R2.

30. Интерпретация значений коэффициента детерминации R2.

31. Система глубокого обучения (Deep learning).

32. Система классического машинного обучения (Classical Machine Learning).

33. Задачи глубокого обучения**.**

34. Особенности глубокого обучения.

35. Типы нейронных сетей для анализа изображения.

36. Обучение нейронных сетей на изображениях.

37. Сверточные слои.

38. Метод решающих деревьев.

39. Принцип работы алгоритмов, основанных на решающих деревьях.

40. Преимущества алгоритма решающих деревьев.

41. Ограничения алгоритмов решающих деревьев.

42. Метод опорных векторов.

43. Зависимость классификации данных от выбора ядра SVM.

44. Работа наивного байесовского классификатора.
45. Сферы применения наивного байесовского классификатора.
46. Особенности применения наивного байесовского классификатора.

3.2. Примерный перечень тем докладов:

1. Основные этапы развития искусственного интеллекта: история и современность.
2. Практическое применение методов машинного обучения для работы с табличными данными.
3. Этические аспекты использования искусственного интеллекта в современном мире.
4. Роль и значение систем искусственного интеллекта в сфере здравоохранения.
5. Интеллектуальный анализ данных в бизнесе: преимущества и вызовы.
6. Перспективы развития искусственного интеллекта в области финансового анализа.
7. Принципы и методы классификации данных на примере алгоритма k-ближайших соседей.
8. Регрессионный анализ: понятие, цели и применение в реальном мире.
9. Сравнительный анализ метрик оценки регрессии: MSE, MAE, R2.
10. Возможности использования систем глубокого обучения для анализа текстовых данных.
11. Применение нейронных сетей для распознавания речи и улучшения систем машинного перевода.
12. Алгоритмы компьютерного зрения: от основ до современных технологий.
13. Кластерный анализ в контексте изучения изображений с использованием нейронных сетей.
14. Обзор методов обучения моделей на основе решающих деревьев.
15. Эффективность прогнозирования с помощью решающих деревьев: практические аспекты.
16. Применение метода опорных векторов для задачи анализа тональности текста.
17. Адаптация метода опорных векторов для задачи детекции объектов на изображениях.
18. Роль и значение наивного байесовского классификатора в обработке естественного языка.
19. Выявление аномалий и их классификация с использованием наивного байесовского метода.
20. Практическое применение наивного байесовского классификатора для анализа потоков данных.

21. Обзор методов ансамблирования при решении задач искусственного интеллекта.

22. Эффективная комбинация методов машинного обучения с алгоритмами глубокого обучения.

23. Интерпретация результатов работы нейронных сетей при анализе изображений.

24. Оценка качества моделей машинного обучения на основе алгоритмов решающих деревьев.

25. Прогнозирование временных рядов с использованием методов регрессии и искусственного интеллекта.

26. Автоматизация процесса выбора параметров модели на примере метода опорных векторов.

27. Оценка вероятности событий и классификации данных с помощью метода наивного байесовского классификатора.

28. Сравнение эффективности алгоритмов машинного обучения при работе с различными типами данных.

29. Сети обратного распространения ошибки: теория и применение в задачах классификации.

30. Оптимизация процесса обучения нейронных сетей для достижения наилучших результатов.

Задания закрытого типа (Тестовые задания)

Номер вопроса и проверка сформированной компетенции

Общие критерии оценивания

№ вопроса	Код компетенции	№ вопроса	Код компетенции
1	ОПК-9. ИОПК-9.1. ИОПК-9.2.	19	ОПК-9. ИОПК-9.1. ИОПК-9.2.
2	ОПК-9. ИОПК-9.1. ИОПК-9.2.	20	ОПК-9. ИОПК-9.1. ИОПК-9.2.
3	ОПК-9. ИОПК-9.1. ИОПК-9.2.	21	ОПК-9. ИОПК-9.1. ИОПК-9.2.
4	ОПК-9. ИОПК-9.1. ИОПК-9.2.	22	ОПК-9. ИОПК-9.1. ИОПК-9.2.
5	ОПК-9. ИОПК-9.1. ИОПК-9.2.	23	ОПК-9. ИОПК-9.1. ИОПК-9.2.
6	ОПК-9. ИОПК-9.1. ИОПК-9.2.	24	ОПК-9. ИОПК-9.1. ИОПК-9.2.
7	ОПК-9. ИОПК-9.1. ИОПК-9.2.	25	ОПК-9. ИОПК-9.1. ИОПК-9.2.
8	ОПК-9. ИОПК-9.1. ИОПК-9.2.	26	ОПК-9. ИОПК-9.1. ИОПК-9.2.

№ вопроса	Код компетенции	№ вопроса	Код компетенции
9	ОПК-9. ИОПК-9.1. ИОПК-9.2.	27	ОПК-9. ИОПК-9.1. ИОПК-9.2.
10	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
11	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
12	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
13	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
14	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
15	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
16	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
17	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
18	ОПК-9. ИОПК-9.1. ИОПК-9.2.		

Ключ ответов

№ вопроса	Верный ответ	№ вопроса	Верный ответ
1	2	19	1
2	3	20	1
3	2	21	3
4	1	22	2
5	3	23	2
6	3	24	1
7	4	25	3
8	2	26	3
9	1	27	2
10	2	28	
11	2	29	
12	3	30	
13	2	31	
14	3	32	
15	3	33	
16	2	34	
17	3	35	
18	1	36	

Тема 1. Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными

Задание № 1

Какая из перечисленных моделей машинного обучения используется для классификации?

1. Линейная регрессия
2. Логистическая регрессия
3. Кластеризация K-средних
4. Главные компоненты анализа (PCA)

Задание № 2

Какой из нижеуказанных методов обучения относится к методу обучения без учителя?

1. Метод опорных векторов (SVM)
2. Дерево решений
3. Кластеризация K-средних
4. Наивный байесовский классификатор

Задание № 3

Какой метод используется для уменьшения размерности данных в машинном обучении?

1. Регрессия Риджа
2. Анализ главных компонент (PCA)
3. Линейная регрессия
4. Градиентный бустинг

Тема 2. Основные задачи систем искусственного интеллекта

Задание № 4

Какая из следующих задач относится к основным задачам искусственного интеллекта?

1. Классификация данных
2. Создание анимаций
3. Производство аппаратного обеспечения
4. Расчет электрических цепей

Задание № 5

Что из перечисленного является задачей обработки естественного языка в системах искусственного интеллекта?

1. Прогнозирование временных рядов
2. Распознавание образов
3. Перевод текста с одного языка на другой

4. Оптимизация маршрутов

Задание № 6
Какую задачу решает алгоритм машинного обучения, определяющий спам в электронной почте?

1. Регрессия
2. Кластеризация
3. Классификация
4. Генерация текстов

Тема 3. Классификация на примере алгоритма k-ближайших соседей (kNN).

Задание № 7
Какой параметр определяет количество соседей, учитываемых алгоритмом k-ближайших соседей?

1. Порог расстояния
2. Коэффициент регуляризации
3. Значение k
4. Максимальная глубина дерева

Задание № 8
Какое из следующих утверждений верно для алгоритма k-ближайших соседей?

1. Алгоритм kNN всегда строит модель перед классификацией
2. Алгоритм kNN является методом классификации и не требует обучения на этапе тренировки
3. Алгоритм kNN использует градиентный спуск для оптимизации
4. Алгоритм kNN работает только с бинарной классификацией

Задание № 9
Какой из перечисленных методов обычно используется для измерения расстояния между точками в алгоритме k-ближайших соседей?

1. Евклидово расстояние
2. Метод главных компонент
3. Логистическая регрессия
4. Метод наименьших квадратов

Тема 4 Регрессия. Метрики оценки регрессии: MSE, MAE, R^2 – коэффициент детерминации

Задание № 10

Какую задачу решает регрессия в машинном обучении?

1. Определение класса объекта
2. Прогнозирование непрерывного значения
3. Разделение данных на кластеры
4. Оптимизация маршрутов

Задание № 11

Какая из перечисленных метрик оценки регрессии измеряет среднюю абсолютную ошибку модели?

1. MSE (среднеквадратичная ошибка).
2. MAE (средняя абсолютная ошибка).
3. R^2 (коэффициент детерминации).
4. Логистическая функция потерь

Задание № 12

Какое значение R^2 (коэффициент детерминации. показывает, что модель идеально объясняет вариативность данных?

1. $R^2 = -1$
2. $R^2 = 0$
3. $R^2 = 1$
4. $R^2 > 1$

Тема 5. Кластеризация: k-means, k-means++, DBSCAN, агломеративная кластеризация. Метрики оценки кластеризации

Задание № 13

Какой из методов кластеризации использует плотность данных для определения кластеров?

1. k-means
2. DBSCAN
3. Агломеративная кластеризация
4. k-means++

Задание № 14

Какую проблему решает метод инициализации k-means++ в алгоритме k-means?

1. Определение плотности точек для кластеров
2. Оптимизацию количества кластеров
3. Снижение зависимости от начального выбора центров кластеров

4. Использование иерархического подхода для формирования кластеров

Задание № 15

Какая из перечисленных метрик может использоваться для оценки качества кластеризации?

1. Коэффициент детерминации (R^2)
2. Среднеквадратичная ошибка (MSE)
3. Индекс силуэта
4. Кросс-энтропия

Тема 6. Алгоритмы, основанные на применении решающих деревьев

Задание № 16

Какой из перечисленных критериев используется для построения решающих деревьев, чтобы разделить данные на подмножества?

1. Коэффициент детерминации (R^2)
2. Энтропия
3. Линейная регрессия
4. Средняя абсолютная ошибка (MAE)

Задание № 17

Какой из алгоритмов является ансамблем решающих деревьев?

1. Линейная регрессия
2. Метод k-ближайших соседей (kNN)
3. Случайный лес
4. Метод главных компонент (PCA)

Задание № 18

Что происходит при увеличении глубины решающего дерева?

1. Увеличивается точность, но возрастает риск переобучения
2. Снижается точность модели
3. Снижается вероятность переобучения
4. Увеличивается скорость обучения

Тема 7. Метод опорных векторов (SVM). Прямая и обратная задача. Определение опорных векторов. Ядерный трюк.

Задание № 19

Какую задачу решает метод опорных векторов (SVM)?

1. Классификация и регрессия

2. Оптимизация маршрутов
3. Кластеризация
4. Прогнозирование временных рядов

Задание № 20

Что такое гиперплоскость в контексте метода опорных векторов?

1. Линия, разделяющая данные на два класса
2. Множество ближайших точек, определяющих границу
3. Линия, определяющая точку перегиба функции потерь
4. Линия, которая минимизирует ошибку

Задание № 21

Какой метод используется для преобразования нелинейно разделимых данных в линейно разделимые в SVM?

1. Линейное ядро
2. Ядро Рамана
3. Ядро Гаусса (RBF.
4. Логистическая функция

Тема 8. Наивный байесовский классификатор

Задание № 22

Что является основным предположением наивного байесовского классификатора?

1. Все признаки зависимы друг от друга
2. Все признаки независимы друг от друга, условно на основе класса
2. Классы равновероятны
3. Признаки имеют нормальное распределение

Задание № 23

Какую роль играет теорема Байеса в наивном байесовском классификаторе?

1. Для оценки зависимости между признаками
2. Для вычисления вероятности классов на основе признаков
3. Для вычисления средней ошибки
4. Для оптимизации градиентного спуска

Задание № 24

Какие типы распределений часто используются в наивном байесовском классификаторе для различных типов признаков?

1. Биномиальное распределение для категориальных признаков и нормальное

распределение для числовых
2. Нормальное распределение для категориальных признаков и экспоненциальное для числовых
3. Геометрическое распределение для всех типов признаков
4. Равномерное распределение для всех признаков

Тема 9. Методы безградиентной оптимизации: случайный поиск, hill climb, отжиг, генетический алгоритм

Задание № 25

Какой из следующих методов безградиентной оптимизации является методом локального поиска, который на каждом шаге выбирает ближайшее улучшение решения?

1. Случайный поиск
2. Генетический алгоритм
3. Hill climb
4. Метод отжига

Задание № 26

Какой метод безградиентной оптимизации использует принцип «температуры» и позволяет избегать попадания в локальные минимумы путем случайных колебаний на этапе отжига?

1. Случайный поиск
2. Генетический алгоритм
3. Отжиг (Simulated Annealing.
4. Hill climb

Задание № 27

Какой из методов оптимизации использует популяцию решений и работает на основе принципов эволюции, таких как отбор, скрещивание и мутация?

1. Случайный поиск
2. Генетический алгоритм
3. Hill climb
3. Метод отжига

Задания открытого типа (типовые задания, ситуационные задачи)

Номер вопроса и проверка сформированной компетенции

№ вопроса	Код компетенции	№ вопроса	Код компетенции
1	ОПК-9. ИОПК-9.1. ИОПК-9.2.	19	ОПК-9. ИОПК-9.1. ИОПК-9.2.

№ вопроса	Код компетенции	№ вопроса	Код компетенции
2	ОПК-9. ИОПК-9.1. ИОПК-9.2.	20	ОПК-9. ИОПК-9.1. ИОПК-9.2.
3	ОПК-9. ИОПК-9.1. ИОПК-9.2.	21	ОПК-9. ИОПК-9.1. ИОПК-9.2.
4	ОПК-9. ИОПК-9.1. ИОПК-9.2.	22	ОПК-9. ИОПК-9.1. ИОПК-9.2.
5	ОПК-9. ИОПК-9.1. ИОПК-9.2.	23	ОПК-9. ИОПК-9.1. ИОПК-9.2.
6	ОПК-9. ИОПК-9.1. ИОПК-9.2.	24	ОПК-9. ИОПК-9.1. ИОПК-9.2.
7	ОПК-9. ИОПК-9.1. ИОПК-9.2.	25	ОПК-9. ИОПК-9.1. ИОПК-9.2.
8	ОПК-9. ИОПК-9.1. ИОПК-9.2.	26	ОПК-9. ИОПК-9.1. ИОПК-9.2.
9	ОПК-9. ИОПК-9.1. ИОПК-9.2.	27	ОПК-9. ИОПК-9.1. ИОПК-9.2.
10	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
11	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
12	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
13	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
14	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
15	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
16	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
17	ОПК-9. ИОПК-9.1. ИОПК-9.2.		
18	ОПК-9. ИОПК-9.1. ИОПК-9.2.		

Ключ ответов к заданиям открытого типа

№ вопроса	Верный ответ
1	Классификация — это задача предсказания категориальных меток, где выходной результат принадлежит одному из заранее определенных классов. Пример: определение, является ли электронное письмо спамом или нет. Регрессия — это задача предсказания непрерывных значений. Пример: предсказание стоимости дома на основе его характеристик.

№ вопроса	Верный ответ
2	1. Очистка данных — удаление или заполнение пропущенных значений, устранение выбросов. 2. Масштабирование и нормализация — приведение данных к единому масштабу для улучшения качества модели. 3. Кодирование категориальных переменных — преобразование категориальных данных в числовой формат, например, с помощью one-hot кодирования. 4. Разделение на тренировочную и тестовую выборки — выделение части данных для проверки качества модели.
3	Метод ближайших соседей (k-Nearest Neighbors, kNN) классифицирует объект на основе его ближайших соседей в многомерном пространстве признаков. Основной параметр — значение k, которое определяет, сколько ближайших соседей учитывать. Также важна метрика расстояния, например, евклидово расстояние. Метод может быть менее эффективным на больших выборках из-за высокой вычислительной сложности и на данных с большим количеством признаков, так как эффект «проклятия размерности» снижает качество модели.
4	Основные задачи ИИ включают: 1. Классификация — определение принадлежности объекта к одному из классов. Применяется, например, в медицинской диагностике для классификации заболеваний. 2. Регрессия — предсказание числовых значений. Используется для прогнозирования, например, цен на акции. 3. Кластеризация — группировка объектов на основе схожести. Применяется в маркетинге для сегментации клиентов. 4. Оптимизация — нахождение оптимальных решений для заданных условий, полезна в логистике для оптимизации маршрутов. 5. Генерация — создание новых данных, например, в творческих приложениях или для генерации изображений.
5	Задача классификации в ИИ — это определение класса, к которому относится новый объект, на основе обучающей выборки. Например, в сфере медицины классификация может использоваться для анализа медицинских изображений и выявления признаков заболеваний, таких как рак. Это полезно для автоматизации диагностики и ускорения процесса выявления заболеваний.
6	Задача оптимизации в ИИ заключается в нахождении лучшего решения для определенной проблемы при заданных ограничениях. Пример — оптимизация маршрутов доставки в логистике, где цель — минимизировать затраты времени и топлива. Это позволяет компаниям улучшить логистику, сократить расходы и уменьшить воздействие на окружающую среду, что полезно как для бизнеса, так и для общества.
7	Алгоритм k-ближайших соседей (kNN) классифицирует объект, основываясь на классах его k ближайших соседей в пространстве признаков. После выбора значения k, алгоритм находит k ближайших точек к новому объекту и относит его к классу, который встречается чаще всего среди этих соседей. Параметры, влияющие на точность, включают: 1. k (количество соседей):маленькие значения k могут привести к переобучению, а слишком большие значения — к недообучению. 2. Метрика расстояния:часто используется евклидово расстояние, но в зависимости от задачи могут применяться и другие метрики (например, манхэттенское расстояние). 3. Масштабирование данных:алгоритм чувствителен к масштабу признаков, поэтому перед использованием kNN важно нормализовать или стандартизировать данные.
8	При k=5 алгоритм найдет 5 ближайших соседей к новому объекту и присвоит ему тот класс, который наиболее часто встречается среди этих соседей. - Если k=1, алгоритм будет классифицировать объект на основе единственного ближайшего соседа, что может привести к переобучению, так как случайный шум может сильно влиять на выбор класса. - Если k=100 (т.е. все объекты в выборке), алгоритм будет учитывать все объекты, и результат будет зависеть от большинства класса в выборке. Такой подход может привести к сильному недообучению, так как индивидуальные особенности соседей не будут учтены, и алгоритм фактически проигнорирует близость к конкретным соседям.
9	Нормализация данных важна, потому что признаки (рост и вес) могут иметь разные масштабы. Например, если рост измеряется в метрах (диапазон от 1 до 2), а вес в килограммах (диапазон от 30 до 150), то вес будет доминировать при вычислении расстояния, так как его значения больше. Без нормализации алгоритм будет уделять больше внимания весу, что может привести к неправильной классификации. Нормализация приводит признаки к одному масштабу (например, от 0 до 1), что позволяет каждому признаку вносить равный вклад в расчет расстояния.
10	Задача регрессии заключается в предсказании непрерывного числового значения на основе имеющихся данных. Примером задачи регрессии может быть прогнозирование цены недвижимости на основе таких данных, как площадь, количество комнат, местоположение и возраст зда-

№ вопроса	Верный ответ
	ния. Обычно для решения задачи регрессии используются числовые и категориальные признаки (после кодирования), которые могут влиять на целевое числовое значение.
11	Средняя квадратичная ошибка (MSE) измеряет среднее значение квадратической ошибки между предсказанными и реальными значениями. MSE рассчитывается как: $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, где y_i — реальное значение, а $\hat{y}_i$ — предсказанное значение. Эта метрика чувствительна к выбросам, так как возводит ошибки в квадрат, что приводит к значительному увеличению значения MSE при наличии больших ошибок. Чтобы снизить влияние группы, можно использовать минимальную абсолютную ошибку (MAE), которая не приводит к ошибкам в квадрате и, следовательно, меньшему влиянию основных отклонений.
12	Коэффициент детерминации P^2 показывает, какую долю вариации целевой переменной объясняет модель. Значение P^2 = 0.85 означает, что модель объясняет 85% изменчивости данных, что указывает на высокое качество предсказаний. Чем ближе значение P^2 к 1, тем лучше модель объясняет данные. Если P^2 близко к 0, это значит, что модель объясняет мало вариации данных и имеет низкое качество.
13	Для определения принадлежности точки P(3, 4) к кластеру, необходимо вычислить расстояние от точки P до каждого из центров кластеров. Используем евклидово расстояние: • Расстояние до C1: [d(P, C1) = \sqrt{(3-1)^2 + (4-2)^2} = \sqrt{4 + 4} = \sqrt{8} \approx 2.83] • Расстояние до C2: [d(P, C2) = \sqrt{(3-5)^2 + (4-6)^2} = \sqrt{4 + 4} = \sqrt{8} \approx 2.83] • Расстояние до C3: [d(P, C3) = \sqrt{(3-9)^2 + (4-10)^2} = \sqrt{36 + 36} = \sqrt{72} \approx 8.49] Поскольку расстояния до C1 и C2 одинаковы, точка P может быть отнесена к любому из этих кластеров. Однако, в случае равенства расстояний, обычно выбирается кластер с меньшим номером, то есть P будет принадлежать к кластеру C1.
14	DBSCAN формирует кластеры на основе плотности точек. Если в плотной области (где находятся 10 точек) найдется хотя бы 3 точки в радиусе ε=1, то они будут отнесены к одному кластеру. Поскольку 10 точек образуют плотную область, они будут объединены в один кластер. 5 оставшихся точек, находящихся далеко от остальных, не будут соответствовать критериям для формирования кластера и будут считаться шумом. Таким образом, будет найдено 1 кластер и 5 шумовых точек.
15	Для оценки качества кластеров, полученных с помощью агломеративной кластеризации, можно использовать различные метрики. Одной из наиболее распространенных является коэффициент силуэта. Он измеряет, насколько хорошо каждая точка соответствует своему кластеру по сравнению с другими кластерами. Коэффициент силуэта варьируется от -1 до 1, где значения близкие к 1 указывают на хорошую кластеризацию, а значения близкие к -1 — на плохую. Таким образом, для оценки качества кластеров A и B можно рассчитать коэффициент силуэта для каждой точки и проанализировать полученные значения.
16	Алгоритм решающих деревьев строит модель в виде дерева, где каждый узел соответствует проверке условия на одном из признаков, а лист — предсказанию класса (для классификации) или значения (для регрессии). На каждом шаге алгоритм разделяет данные на основе признака, который наилучшим образом разделяет данные для максимальной чистоты узлов. Параметры, которые влияют на глубину дерева, включают максимальную глубину (`max_depth`), минимальное количество выборок в узле (`min_samples_split`) и минимальное количество выборок в листе (`min_samples_leaf`). Чем глубже дерево, тем более сложную структуру оно может моделировать, однако чрезмерно глубокие деревья склонны к переобучению, так как слишком точно подгоняются под обучающие данные.
17	Бустинг — это метод ансамблевого обучения, который создает несколько слабых моделей (чаще всего малых деревьев), последовательно обучая каждую на ошибках предыдущих. В отличие от одного решающего дерева, бустинг объединяет множество деревьев, чтобы повысить точность и устойчивость модели. Каждый последующий алгоритм в бустинге фокусируется на ошибках, допущенных предыдущими деревьями, что позволяет улучшить качество модели. Примером алгоритма, использующего бустинг, является Gradient Boosting(градиентный бустинг), а также его улучшенные версии, такие как XGBoostи LightGBM, которые особенно популярны для решения задач с табличными данными.
18	Для оценки качества модели можно использовать метрики, такие как точность (accuracy), полнота (recall) и F1-мера. Точность поможет понять, сколько из всех предсказаний были верными, полнота покажет, насколько хорошо модель находит положительные примеры, а F1-мера даст сбалансированное представление о точности и полноте, что особенно важно в задачах с несба-

№ вопроса	Верный ответ
	лансированными классами.
19	Метод опорных векторов (SVM) — это алгоритм машинного обучения, используемый для классификации и регрессии. Основной принцип работы SVM заключается в нахождении гиперплоскости, которая максимально разделяет классы данных в многомерном пространстве. Гиперплоскость — это обобщение понятия "прямой" в более высоких измерениях. SVM стремится найти такую гиперплоскость, которая обеспечивает максимальный зазор (margin) между ближайшими точками разных классов, называемыми опорными векторами. Это позволяет SVM быть устойчивым к шуму и переобучению.
20	Ядровая функция в контексте SVM — это функция, которая позволяет преобразовать данные в более высокое измерение, где они могут быть линейно разделимы. Это особенно полезно, когда данные не поддаются линейной классификации в исходном пространстве. Ядровые функции позволяют SVM эффективно работать с нелинейными данными, не вычисляя координаты в высоком измерении напрямую. Примеры различных типов ядровых функций включают линейное ядро, полиномиальное ядро, радиальное базисное ядро (RBF) и сигмоидное ядро.
21	Метод опорных векторов (SVM) находит широкое применение в различных областях, таких как распознавание образов, биоинформатика, текстовая классификация и финансовый анализ. Например, SVM может использоваться для классификации изображений, где он помогает различать объекты на фотографиях. В биоинформатике SVM применяется для предсказания структуры белков. В текстовой классификации SVM может эффективно разделять спам и не-спам сообщения. Преимущества использования SVM включают его способность работать с высокоразмерными данными, устойчивость к переобучению и возможность использования различных ядровых функций для обработки нелинейных зависимостей.
22	Используем формулу наивного байесовского классификатора: P(Spam | Скидка) = (P(Скидка | Spam) P(Spam)) / P(Скидка) Где: • P(Скидка | Spam) = 0.8 • P(Spam) = 0.7 • P(Скидка) = P(Скидка | Spam) P(Spam) + P(Скидка | Не Spam) P(Не Spam) = 0.8 0.7 + 0.1 0.3 = 0.56 + 0.03 = 0.59 Теперь подставим значения: P(Spam | Скидка) = (0.8 0.7) / 0.59 ≈ 0.949 Таким образом, вероятность того, что письмо является спамом, составляет примерно 94.9%.
23	Используем ту же формулу: P(Политическая | Выборы) = (P(Выборы | Политическая) P(Политическая)) / P(Выборы) Где: • P(Выборы | Политическая) = 0.9 • P(Политическая) = 0.6 • P(Выборы) = P(Выборы | Политическая) P(Политическая) + P(Выборы | Экономическая) P(Экономическая) = 0.9 0.6 + 0.2 0.4 = 0.54 + 0.08 = 0.62 Теперь подставим значения: P(Политическая | Выборы) = (0.9 0.6) / 0.62 ≈ 0.868 Таким образом, вероятность того, что новость является политической, составляет примерно 86.8%.
24	Используем формулу: P(Положительный | Отлично) = (P(Отлично | Положительный) P(Положительный)) / P(Отлично) Где: • P(Отлично | Положительный) = 0.85 • P(Положительный) = 0.75 • P(Отлично) = P(Отлично | Положительный) P(Положительный) + P(Отлично | Отрицательный) P(Отрицательный) = 0.85 0.75 + 0.15 0.25 = 0.6375 + 0.0375 = 0.675 Теперь подставим значения: P(Положительный | Отлично) = (0.85 0.75) / 0.675 ≈ 0.944 Таким образом, вероятность того, что отзыв положительный, составляет примерно 94.4%.
25	1. Генерируем 1000 случайных значений (x) в интервале ([0, 6]). 2. Вычисляем значение функции (f(x)) для каждого из сгенерированных значений. 3. Находим минимальное значение функции и соответствующее ему значение (x). import numpy as np def f(x): return (x - 3)2 + 2 random_points = np.random.uniform(0, 6, 1000) function_values = f(random_points)

№ вопроса	Верный ответ
	min_index = np.argmin(function_values) min_x = random_points[min_index] min_value = function_values[min_index] print(f"Минимум функции достигается в x = {min_x}, f(x) = {min_value}")

26	1. Начинаем с начальной точки (x = 0).
	2. На каждом шаге проверяем значения функции в точках (x + \Delta x) и (x - \Delta x).
	3. Если значение функции увеличивается, перемещаемся в сторону увеличения, иначе останавливаемся.
	def f(x):
	return -x2 + 4x
	
	x = 0
	delta_x = 0.1
	step = True
	
	while step:
	step = False
	if f(x + delta_x) > f(x):
	x += delta_x
	step = True
	elif f(x - delta_x) > f(x):
	x -= delta_x
	step = True
	
	print(f"Максимум функции достигается в x = {x}, f(x) = {f(x)}")

27	1. Инициализируем популяцию случайными значениями (x) в интервале ([0, 4]).
	2. На каждом поколении выбираем лучших особей, скрещиваем их и мутируем.
	3. Повторяем процесс на протяжении 20 поколений.
	import random
	
	def f(x):
	return -abs(x - 2) + 3
	population = [random.uniform(0, 4) for _ in range(10)]
	for generation in range(20):
	# Оценка приспособленности
	population.sort(key=f, reverse=True)
	next_generation = population[:5] # выбираем лучших
	# Скрещивание
	while len(next_generation) < 10:
	parent1, parent2 = random.sample(next_generation[:5], 2)
	child = (parent1 + parent2) / 2 + random.uniform(-0.1, 0.1) # мутация
	next_generation.append(child)
	population = next_generation
	best_x = population[0]
	print(f"Максимум функции достигается в x = {best_x}, f(x) = {f(best_x)}")

Тема 1. Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными

Задание № 1

Объясните, чем отличаются задачи классификации и регрессии в машинном обучении? Приведите примеры использования для каждой из задач.

Задание № 2

Какие основные этапы включены в процесс предобработки табличных данных перед использованием в моделях машинного обучения? Охарактеризуйте каждый этап.

Задание № 3
Опишите, как работает метод ближайших соседей (k-Nearest Neighbors) и какие параметры влияют на его производительность. В каких случаях он может быть менее эффективным?

Тема 2. Основные задачи систем искусственного интеллекта

Задание № 4
Перечислите и кратко опишите основные задачи, которые решают системы искусственного интеллекта (ИИ). Какую пользу они приносят в различных сферах?

Задание № 5
Что такое задача классификации в системах искусственного интеллекта? Приведите пример, где задача классификации может применяться, и объясните, как она может быть полезной.

Задание № 6
Объясните, что такое задача оптимизации в искусственном интеллекте. Приведите пример задачи оптимизации и охарактеризуйте, как она может улучшить процессы в реальной жизни.

Тема 3. Классификация на примере алгоритма k-ближайших соседей (kNN).

Задание № 7
Опишите, как работает алгоритм k-ближайших соседей (kNN) для задачи классификации. Какие параметры влияют на его поведение и точность?

Задание № 8
Рассмотрите ситуацию: у нас есть обучающая выборка с 100 объектами, разделёнными на два класса. Мы применяем kNN с k=5 для классификации нового объекта. Объясните, как будет происходить классификация и что произойдет, если число соседей k выбрать равным 1 или 100.

Задание № 9
В данных имеется два признака: рост и вес. Вы хотите использовать kNN для классификации новых объектов на основании этих признаков. Объясните, почему важно выполнить нормализацию данных перед применением kNN, и приведите пример того, как отсутствие нормализации может повлиять на результаты классификации.

Тема 4. Регрессия. Метрики оценки регрессии: MSE, MAE, R^2 – коэффициент детерминации

Задание № 10

Объясните, что представляет собой задача регрессии в машинном обучении. Приведите пример задачи регрессии и укажите, какие типы данных обычно используются для её решения.

Задание № 11

Опишите метрику средней квадратичной ошибки (MSE). Почему эта метрика считается чувствительной к выбросам? Какую альтернативную метрику можно использовать, чтобы снизить влияние выбросов?

Задание № 12

Что показывает коэффициент детерминации P^2? Если модель имеет P^2= 0.85, что это означает в контексте качества модели? Какое значение P^2 указывает на то, что модель объясняет мало вариации данных?

Тема 5. Кластеризация: k-means, k-means++, DBSCAN, агломеративная кластеризация. Метрики оценки кластеризации

Задание № 13

У вас есть набор данных, состоящий из 10 точек в двумерном пространстве. Вы применили алгоритм k-means с k=3. После первой итерации центры кластеров оказались в следующих координатах: C1(1, 2), C2(5, 6), C3(9, 10). Определите, к какому кластеру будет принадлежать точка P(3, 4).

Задание № 14

У вас есть набор данных, состоящий из точек, расположенных в двумерном пространстве. Вы применили алгоритм DBSCAN с параметрами ε=1 и minPts=3. Определите, сколько кластеров будет найдено, если в вашем наборе данных есть 15 точек, из которых 10 образуют плотную область, а 5 находятся далеко от остальных.

Задание № 15

Вы применили агломеративную кластеризацию к набору данных, состоящему из 8 точек. В процессе кластеризации вы использовали метод "среднего связывания". После первого объединения у вас образовались два кластера: A (точки 1, 2, 3) и B (точки 4, 5, 6). Какой метод оценки кластеризации можно использовать для оценки качества полученных кластеров?

Тема 6. Алгоритмы, основанные на применении решающих деревьев

Задание № 16

Опишите принцип работы алгоритма решающих деревьев. Какие параметры влияют на глубину дерева и как глубина может повлиять на качество модели?

Задание № 17

Объясните, что такое "бустинг" на основе решающих деревьев и как он отличается от одного решающего дерева. Приведите пример алгоритма, использующего бустинг.

Задание № 18

Банк хочет оценить кредитоспособность своих клиентов на основе таких факторов, как доход, возраст, наличие задолженности и кредитная история. Вам необходимо построить модель, используя решающие деревья. Какие метрики вы будете использовать для оценки качества модели и почему?

Тема 7. Метод опорных векторов (SVM). Прямая и обратная задача. Определение опорных векторов. Ядерный трюк.

Задание № 19

Опишите основные принципы работы метода опорных векторов (SVM). Включите в ответ объяснение, что такое гиперплоскость и как SVM использует ее для классификации данных.

Задание № 20

Что такое ядровая функция в контексте SVM и как она помогает в решении задач классификации? Приведите примеры различных типов ядровых функций.

Задание № 21

Опишите несколько реальных приложений метода опорных векторов (SVM) в различных областях. Каковы преимущества использования SVM в этих приложениях?

Тема 8. Наивный байесовский классификатор

Задание № 22

У вас есть набор писем, которые были помечены как "спам" или "не спам". Известно, что 70% писем являются спамом, а 30% — не спамом. В одном из писем содержится слово "скидка". Из анализа данных известно, что 80% спам-писем содержат слово "скидка", и только 10% не спам-писем содержат это слово. Какова вероятность того, что данное письмо является спамом?

Задание № 23

Предположим, у вас есть классификатор новостей, который определяет, является ли новость "политической" или "экономической". Известно, что 60%

новостей — политические, а 40% — экономические. В одной из новостей упоминается слово "выборы". Из анализа данных известно, что 90% политических новостей содержат слово "выборы", и только 20% экономических новостей содержат это слово. Какова вероятность того, что новость является политической?

Задание № 24

Вы анализируете отзывы о продукте и хотите классифицировать их как "положительные" или "отрицательные". Известно, что 75% отзывов положительные, а 25% — отрицательные. В одном из отзывов встречается слово "отлично". Из анализа данных известно, что 85% положительных отзывов содержат слово "отлично", и только 15% отрицательных отзывов содержат это слово. Какова вероятность того, что отзыв положительный?

Тема 9. Методы безградиентной оптимизации: случайный поиск, hill climb, отжиг, генетический алгоритм

Задание № 25

Реализуйте алгоритм случайного поиска для нахождения минимума функции (f(x) = (x - 3)^2 + 2) на интервале ([0, 6]). Используйте 1000 случайных точек для поиска минимума.

Задание № 26

Реализуйте метод "hill climb" для нахождения максимума функции (f(x) = -x^2 + 4x) на интервале ([0, 4]). Начальная точка - (x = 0). Используйте шаг изменения (\Delta x = 0.1).

Задание № 27

Реализуйте генетический алгоритм для нахождения максимума функции (f(x) = -|x - 2| + 3) на интервале ([0, 4]). Используйте популяцию из 10 особей и 20 поколений.

2 ЭТАП – Промежуточная аттестация по итогам освоения дисциплины

3.3. Вопросы для проведения зачета:

1. Что такое искусственный интеллект и какие его основные компоненты?
2. Какие существуют типы искусственного интеллекта? Приведите примеры для каждого типа.
3. Объясните, что такое машинное обучение и как оно связано с искусственным интеллектом.
4. Назовите и опишите основные алгоритмы машинного обучения. В чем их отличия?

5. Что такое нейронные сети и как они работают? Приведите примеры их применения.

6. Каковы основные задачи обработки естественного языка (NLP)? Какие технологии используются для их решения?

7. Какие этические вопросы возникают при использовании систем искусственного интеллекта? Как можно их решить?

8. Приведите примеры применения искусственного интеллекта в различных отраслях, таких как медицина, финансы и транспорт.

9. Как работают системы рекомендаций? Какие алгоритмы используются для их создания?

10. Каковы ваши прогнозы относительно будущего искусственного интеллекта? Какие технологии могут стать ключевыми в ближайшие годы?

11. Концепции искусственного интеллекта.

12. Методы машинного обучения.

13. Методы машинного обучения при анализе табличных данных.

14. Процесс предварительной обработки данных перед применением методов машинного обучения.

15. Ключевые задачи искусственного интеллекта в современном мире.

16. Проблемы применения систем искусственного интеллекта.

17. Работа алгоритма k-ближайших соседей (kNN) в задаче классификации.

18. Выбор оптимального значения параметра k при использовании алгоритма kNN.

19. Преимущества и ограничения применения алгоритма k-ближайших соседей по сравнению с другими методами классификации.

20. Задача регрессии и ее типы.

21. Применение метрик MSE, MAE, R2 для оценки качества модели регрессии.

22. Интерпретация больших значений MSE и MAE при оценке модели регрессии.

23. Отличия глубокого обучения от классических методов машинного обучения.

24. Особенности решения проблем с помощью систем глубокого обучения.

25. Типы нейронных сетей для обработки изображений.

26. Особенности применения сверточных нейронных сетей в обработке изображений.

27. Работа алгоритмов, основанных на решающих деревьях.

28. Выбор критериев разделения при построении решающего дерева.

29. Основной принцип работы метода опорных векторов (SVM).

30. Выбор ядра при использовании метода опорных векторов.

31. Отличие наивного байесовского классификатора от других алгоритмов классификации.

32. Условия выбора машинного обучения для конкретных задач.

33. Отличие надзорного и ненадзорного обучения в машинном обучении.

34. Учет проблемы переобучения при построении модели машинного обучения.

35. Виды методов ансамблирования моделей.

36. Важность проведения кросс-валидации при оценке качества модели машинного обучения.

37. Аспекты безопасности при разработке систем на основе искусственного интеллекта.

38. Эффективность различных моделей машинного обучения при решении одной и той же задачи.

39. Роль функции потерь в процессе обучения моделей машинного обучения.

40. Методы извлечения признаков из данных перед применением алгоритмов машинного обучения.

Комплекс заданий

Общие критерии оценивания

№ п/п	Процент правильных ответов	Оценка
1	86 % – 100 %	5 («отлично»)
2	70 % – 85 %	4 («хорошо)
3	51 % – 69 %	3 (удовлетворительно)
4	50 % и менее	2 (неудовлетворительно)

Задания закрытого типа

Номер вопроса и проверка сформированной компетенции

№ вопроса	Код компетенции	№ вопроса	Код компетенции
1	ОПК-9. ИОПК-9.1. ИОПК-9.2.	19	ОПК-9. ИОПК-9.1. ИОПК-9.2.
2	ОПК-9. ИОПК-9.1. ИОПК-9.2.	20	ОПК-9. ИОПК-9.1. ИОПК-9.2.
3	ОПК-9. ИОПК-9.1. ИОПК-9.2.	21	ОПК-9. ИОПК-9.1. ИОПК-9.2.
4	ОПК-9. ИОПК-9.1. ИОПК-9.2.	22	ОПК-9. ИОПК-9.1. ИОПК-9.2.
5	ОПК-9. ИОПК-9.1. ИОПК-9.2.	23	ОПК-9. ИОПК-9.1. ИОПК-9.2.
6	ОПК-9. ИОПК-9.1. ИОПК-9.2.	24	ОПК-9. ИОПК-9.1. ИОПК-9.2.
7	ОПК-9.	25	ОПК-9.

№ вопроса	Код компетенции	№ вопроса	Код компетенции
	ИОПК-9.1. ИОПК-9.2.		ИОПК-9.1. ИОПК-9.2.
8	ОПК-9. ИОПК-9.1. ИОПК-9.2.	26	ОПК-9. ИОПК-9.1. ИОПК-9.2.
9	ОПК-9. ИОПК-9.1. ИОПК-9.2.	27	ОПК-9. ИОПК-9.1. ИОПК-9.2.
10	ОПК-9. ИОПК-9.1. ИОПК-9.2.	28	ОПК-9. ИОПК-9.1. ИОПК-9.2.
11	ОПК-9. ИОПК-9.1. ИОПК-9.2.	29	ОПК-9. ИОПК-9.1. ИОПК-9.2.
12	ОПК-9. ИОПК-9.1. ИОПК-9.2.	30	ОПК-9. ИОПК-9.1. ИОПК-9.2.
13	ОПК-9. ИОПК-9.1. ИОПК-9.2.	31	ОПК-9. ИОПК-9.1. ИОПК-9.2.
14	ОПК-9. ИОПК-9.1. ИОПК-9.2.	32	ОПК-9. ИОПК-9.1. ИОПК-9.2.
15	ОПК-9. ИОПК-9.1. ИОПК-9.2.	33	ОПК-9. ИОПК-9.1. ИОПК-9.2.
16	ОПК-9. ИОПК-9.1. ИОПК-9.2.	34	ОПК-9. ИОПК-9.1. ИОПК-9.2.
17	ОПК-9. ИОПК-9.1. ИОПК-9.2.	35	ОПК-9. ИОПК-9.1. ИОПК-9.2.
18	ОПК-9. ИОПК-9.1. ИОПК-9.2.	36	ОПК-9. ИОПК-9.1. ИОПК-9.2.

Ключ ответов

№ вопроса	Верный ответ	№ вопроса	Верный ответ
1	3	19	2
2	2	20	1
3	4	21	3
4	2	22	4
5	3	23	3
6	4	24	4
7	2	25	3
8	3	26	4
9	2	27	4
10	2	28	4
11	2	29	2
12	2	30	1
13	3	31	2
14	2	32	2
15	1	33	1
16	1	34	3
17	1	35	1
18	1	36	2

Задание №1.
Что означает термин "гиперпараметр" в контексте алгоритмов машинного обучения?

Ответ:
1. Параметр модели, определяемый в процессе обучения;
2. Независимая переменная, влияющая на вывод модели;
3. Параметр, настраиваемый вне процесса обучения, влияющий на процесс обучения и результаты;
4. Результат классификации модели на тестовой выборке.

Задание №2.
Что представляет собой метод "градиентного бустинга"?

Ответ:
1. Алгоритм для автоматической генерации новых признаков из исходных;
2. Метод оптимизации функции потерь путем последовательного добавления слабых моделей;
3. Модель для генерации случайного леса на основе исходных данных;
4. Алгоритм для уменьшения размерности данных в процессе обучения.

Задание №3.
Какой тип задач в искусственном интеллекте связан с созданием систем, способных воспринимать и понимать человеческую речь?

Ответ:
1. Генерация естественного языка;
2. Автоматическое управление процессами производства;
3. Обнаружение аномалий в данных;
4. Распознавание речи.

Задание №4.
Что представляет собой задача "симуляции человеческого мышления" в рамках искусственного интеллекта?

Ответ:
1. Оценка эмоциональных реакций на изображения;
2. Моделирование когнитивных процессов и принятия решений;
3. Автоматизированная генерация искусственных текстов;
4. Разработка алгоритмов для оптимизации производственных процессов.

Задание № 5.
Как влияет параметр k на работу алгоритма k-ближайших соседей?

Ответ:
1. Увеличение k ухудшает точность классификации;
2. Уменьшение k увеличивает склонность модели к переобучению;
3. При увеличении k модель становится менее чувствительной к шуму;
4. Выбор k не влияет на работу алгоритма.

Задание № 6.
Что такое "взвешенная" версия алгоритма kNN?

Ответ:
1. Модификация алгоритма, при которой ближайшие соседи учитываются с разными весами в зависимости от удаленности;
2. Применение метода случайного леса для классификации;
3. Адаптация алгоритма для работы с текстовыми данными;
4. Модель, где каждый сосед вкладывает в прогноз с весом, зависящим от расстояния до целевого объекта.

Задание № 7.
Каким образом определяется "оптимальное" значение параметра k в алгоритме k-ближайших соседей?

Ответ:
1. Выборка параметра k всегда зависит от типа задачи и характера данных;
2. Выбор оптимального k выполняется на основе кросс-валидации;
3. Оптимальное значение k рассчитывается аналитически на основе функции потерь;
4. Использование k равного длине обучающей выборки для достижения наилучшего результата.

Задание № 8.
Что происходит при использовании алгоритма k-ближайших соседей в случае несбалансированных классов?

Ответ:
1. Алгоритм kNN не работает с несбалансированными данными;
2. Необходимо провести масштабирование признаков для улучшения работы алгоритма;

3. Рекомендуется применение взвешенной версии kNN или использование метода SMOTE для балансировки данных;
4. Несбалансированные классы не влияют на результат работы алгоритма kNN.

Задание № 9.
Что означает коэффициент детерминации R2 в контексте регрессии?

Ответ:
1. Процент точности предсказаний модели;
2. Доля объяснённой дисперсии в зависимой переменной моделью;
3. Коэффициент корреляции между признаками;
4. Среднее значение входных признаков.

Задание № 10.
Какая из метрик, MSE, MAE или R2, более чувствительна к выбросам в данных?

Ответ:
1. MSE;
2. MAE;
3. R2;
4. Все метрики одинаково чувствительны к выбросам.

Задание № 11.
Чем отличаются метрики MSE и MAE?

Ответ:
1. MSE учитывает только абсолютные разности, MAE - квадраты разностей;
2. MAE сильнее штрафует за большие ошибки, чем MSE;
3. MSE усредняет абсолютные значения, MAE - их квадраты;
4. MAE нормализует ошибки в отличие от MSE.

Задание № 12.
Чем отличается сверточная нейронная сеть (CNN) от обычной полносвязанной нейронной сети?

Ответ:
1. CNN обычно используется для обработки текстовых данных, а полносвязанная сеть для изображений;

2. CNN имеет специализированные слои для обработки пространственных признаков, типично для изображений;
3. Полносвязанная сеть требует меньше вычислительных ресурсов для обучения;
4. CNN используется только для задач классификации данных.

Задание №13.

Какой метод активации чаще всего используется в последних слоях сверточных нейронных сетей для задачи классификации изображений?

Ответ:
1. ReLU;
2. Сигмоидная функция активации;
3. Softmax;
4. Tanh.

Задание №14.

Что такое "обучение на основе передаточного обучения" (transfer learning) в контексте работы с изображениями?

Ответ:
1. Процесс обучения модели с нуля на новом наборе данных;
2. Использование заранее обученных моделей для извлечения признаков из изображений и их дальнейшего обучения на новых задачах;
3. Техника увеличения датасета для обучения через аугментацию изображений;
4. Обучение модели на высокоразмерных изображениях для улучшения ее обобщающей способности.

Задание №15.

Чем отличается алгоритм ID3 от C4.5 в контексте построения решающих деревьев?

Ответ:
1. ID3 использует энтропию Гини для выбора лучшего признака, C4.5 - информационный прирост;
2. ID3 неприменим для задач классификации, C4.5 только для регрессии;
3. ID3 допускает только бинарные признаки, C4.5 работает с категориальными и числовыми данными;
4. ID3 работает только с дискретными данными, C4.5 — с непрерывными.

Задание №16.
Какие методы обработки пропущенных значений часто используются при построении решающих деревьев?

Ответ:
1. Замена отсутствующих значений средним или медианой для числовых признаков, наиболее частым значением для категориальных;
2. Исключение объектов с пропущенными значениями из обучающего набора данных;
3. Замена пропусков нулями для всех признаков;
4. Использование метода k-ближайших соседей для заполнения пропущенных значений.

Задание №17.
Что такое "жесткий зазор" в SVM?

Ответ:
1. Ограничение на ширину разделяющего гиперплоскость между классами, позволяющее до некоторой ошибки классифицировать объекты;
2. Максимально возможное разделение объектов на два класса без ошибок;
3. Зазор между признаками объектов, определяющий их сходство или различие;
4. Уровень шума в данных, который модель способна игнорировать при построении разделяющей границы.

Задание №18.
Каким образом SVM обрабатывает нелинейные зависимости между данными?

Ответ:
1. Использованием функций ядра, которые отображают данные в пространство более высокой размерности, где они становятся линейно разделимыми;
2. Другим комплексным методом машинного обучения, не имеющим отношения к линейной алгебре;
3. Преобразованием данных с помощью метода кластеризации для повышения их обобщающей способности;
4. Увеличением количества обучающих объектов для нахождения новых зависимостей.

Задание №19.
Какие предположения делает наивный байесовский классификатор о данных?

Ответ:
1. Предполагает независимость между всеми парами признаков и равномерное распределение классов;
2. Условная независимость каждого признака от других при условии известного класса;
3. Предполагает линейную зависимость между признаками и классами;
4. Основывается на предположении о нормальном распределении каждого класса в выборке.

Задание №20.

Как выбор используемой модели вероятностей влияет на работу наивного байесовского классификатора?

Ответ:
1. Модель мультиномиального распределения работает лучше для категориальных данных, где встречаются часто повторяющиеся значения;
2. Гауссовская модель используется для непрерывных данных и предполагает нормальное распределение;
3. Бернуллиевская модель хорошо подходит для бинарных данных, где значения признаков равны 0 или 1;
4. Каждая модель вероятностей подходит для определенного типа данных, и выбор зависит от вида признаков в задаче.

Задание № 21.

Какие факторы следует учитывать при выборе алгоритма машинного обучения для конкретной задачи?

Ответ:
1. Скорость обучения модели и уровень шума в данных;
2. Точность классификации на обучающей выборке;
3. Сложность модели, объем данных и вид задачи;
4. Уровень интерпретации результатов визуализации.

Задание № 22.

Что представляет собой метод "линейная регрессия" в машинном обучении?

Ответ:
1. Метод для построения линейных моделей на основе данных;
2. Техника для решения сложных оптимизационных задач;
3. Алгоритм для автоматического извлечения признаков из текстовых данных;
4. Модель для прогнозирования числовых значений на основе линейной зависимости.

Задание № 23.

Какие методы машинного обучения подходят для работы с категориальными признаками в табличных данных?

Ответ:
1. Методы деревьев решений и случайного леса;
2. Логистическая регрессия и метод ближайших соседей;
3. Кодирование категориальных признаков и использование градиентного бустинга;
4. Методы кластеризации и метод опорных векторов.

Задание № 24.

Какая задача в области искусственного интеллекта связана с созданием систем, способных самостоятельно общаться и взаимодействовать с людьми?

Ответ:
1. Классификация больших объемов данных;
2. Нейронные сети глубокого обучения;
3. Обработка естественного языка;
4. Разработка чат-ботов.

Задание № 25.

Что подразумевает задача "робототехники" в контексте искусственного интеллекта?

Ответ:
1. Создание системы распознавания лиц;
2. Разработка алгоритмов обучения без учителя;
3. Проектирование систем управления и обучения роботов;
4. Анализ данных о клиентах для персонализации рекомендаций.

Задание № 26.

Какую задачу в сфере искусственного интеллекта можно связать с разработкой систем, способных распознавать объекты на изображениях?

Ответ:
1. Метод главных компонент;
2. Построение рекомендательных систем;
3. Сегментация изображений;
4. Обработка изображений и компьютерное зрение.

Задание № 27.
Что означает задача "обучения с подкреплением" в области искусственного интеллекта?

Ответ:
1. Программирование виртуальных агентов для ведения диалогов;
2. Настройка параметров нейронных сетей;
3. Обучение алгоритмов на основе опыта и обратной связи;
4. Создание игровых симуляторов для обучения нейронных сетей.

Задание № 28.
Что представляет собой метод "прагматического учителя" в контексте алгоритма k-ближайших соседей?

Ответ:
1. Автоматическое обучение путем наблюдения за поведением других агентов;
2. Метод оценки важности признаков в данных;
3. Техника взвешивания при сравнении ближайших соседей;
4. Подбор оптимального значения k для повышения точности классификации.

Задание № 29.
Что такое "пространство признаков" в контексте применения алгоритма k-ближайших соседей?

Ответ:
1. Область, в которой находятся целевые классы для классификации;
2. Множество всех возможных значений признаков, используемых для описания объектов;
3. Размер окрестности, в которой происходит поиск ближайших соседей;
4. Количество признаков, необходимых для корректной работы алгоритма kNN.

Задание № 30.
В чем сущность использования R2 для оценки моделей регрессии?

Ответ:
1. R2 показывает, насколько качественно модель предсказывает значения по сравнению с простым средним;
2. R2 определяет средние значения показателей в данных;

3. R2 выгодно используется для преобразования данных перед обучением модели;
4. R2 отражает точность предсказания модели на тренировочной выборке.

Задание № 31.
Как влияет увеличение значения MSE на качество модели регрессии?

Ответ:
1. Чем больше MSE, тем лучше качество модели;
2. При увеличении MSE качество модели ухудшается;
3. MSE не влияет на качество модели регрессии;
4. MSE нельзя интерпретировать как метрику качества модели.

Задание № 32.
Какую роль играют функции активации в глубоком обучении?

Ответ:
1. Функции активации преобразуют выходы нейронов в слоях нейронной сети в вероятности классов;
2. Функции активации определяют нелинейности и сложности, которые сеть может моделировать;
3. Функции активации используются для уплотнения данных перед входом в нейронную сеть;
4. Функции активации отвечают за обратное распространение ошибки в сети.

Задание №33.
Что представляет собой метод "Transfer Learning" в глубоком обучении?

Ответ:
1. Передача знаний и весов предобученной модели для обучения новой модели на небольшом наборе данных;
2. Алгоритм для регуляризации нейронных сетей;
3. Техника для автоматического выбора архитектуры нейронной сети;
4. Метод для инициализации весов модели перед началом обучения.

Задание №34.
Что представляют собой глубокие нейронные сети с архитектурой "U-Net"?

Ответ:
1. Сети, специализированные на генерации изображений из текстовых описаний;
2. Модели для обработки видео и последовательных данных;

3. Нейронные сети, применяемые в задачах сегментации изображений, где присутствуют кодировщик и декодировщик;
4. Алгоритмы для анализа эмоциональных состояний людей на изображениях.

Задание №35.
Что такое "подстройка" (pruning) деревьев решений?

Ответ:
1. Процесс удаления недействительных ветвей и узлов дерева для улучшения обобщающей способности модели;
2. Увеличение глубины дерева после завершения обучения для увеличения точности прогнозирования;
3. Процесс увеличения количества листьев в дереве для повышения детализации предсказаний;
4. Использование метода случайного леса для улучшения качества классификации.

Задание №36.
Какая метрика часто используется для выбора лучшего разбиения в узле решающего дерева?

Ответ:
1. Коэффициент корреляции Пирсона;
2. Коэффициент Джини;
3. Коэффициент детерминации R2;
4. Средняя абсолютная ошибка (MAE).

Задания открытого типа

Номер вопроса и проверка сформированной компетенции

№ вопроса	Код компетенции	№ вопроса	Код компетенции
1	ОПК-9. ИОПК-9.1. ИОПК-9.2.	19	ОПК-9. ИОПК-9.1. ИОПК-9.2.
2	ОПК-9. ИОПК-9.1. ИОПК-9.2.	20	ОПК-9. ИОПК-9.1. ИОПК-9.2.
3	ОПК-9. ИОПК-9.1. ИОПК-9.2.	21	ОПК-9. ИОПК-9.1. ИОПК-9.2.
4	ОПК-9. ИОПК-9.1. ИОПК-9.2.	22	ОПК-9. ИОПК-9.1. ИОПК-9.2.
5	ОПК-9. ИОПК-9.1. ИОПК-9.2.	23	ОПК-9. ИОПК-9.1. ИОПК-9.2.

№ вопроса	Код компетенции	№ вопроса	Код компетенции
6	ОПК-9. ИОПК-9.1. ИОПК-9.2.	24	ОПК-9. ИОПК-9.1. ИОПК-9.2.
7	ОПК-9. ИОПК-9.1. ИОПК-9.2.	25	ОПК-9. ИОПК-9.1. ИОПК-9.2.
8	ОПК-9. ИОПК-9.1. ИОПК-9.2.	26	ОПК-9. ИОПК-9.1. ИОПК-9.2.
9	ОПК-9. ИОПК-9.1. ИОПК-9.2.	27	ОПК-9. ИОПК-9.1. ИОПК-9.2.
10	ОПК-9. ИОПК-9.1. ИОПК-9.2.	28	ОПК-9. ИОПК-9.1. ИОПК-9.2.
11	ОПК-9. ИОПК-9.1. ИОПК-9.2.	29	ОПК-9. ИОПК-9.1. ИОПК-9.2.
12	ОПК-9. ИОПК-9.1. ИОПК-9.2.	30	ОПК-9. ИОПК-9.1. ИОПК-9.2.
13	ОПК-9. ИОПК-9.1. ИОПК-9.2.	31	ОПК-9. ИОПК-9.1. ИОПК-9.2.
14	ОПК-9. ИОПК-9.1. ИОПК-9.2.	32	ОПК-9. ИОПК-9.1. ИОПК-9.2.
15	ОПК-9. ИОПК-9.1. ИОПК-9.2.	33	ОПК-9. ИОПК-9.1. ИОПК-9.2.
16	ОПК-9. ИОПК-9.1. ИОПК-9.2.	34	ОПК-9. ИОПК-9.1. ИОПК-9.2.
17	ОПК-9. ИОПК-9.1. ИОПК-9.2.	35	ОПК-9. ИОПК-9.1. ИОПК-9.2.
18	ОПК-9. ИОПК-9.1. ИОПК-9.2.	36	ОПК-9. ИОПК-9.1. ИОПК-9.2.

Ключ ответов к заданиям открытого типа

№ вопроса	Верный ответ
1	Искусственный интеллект (ИИ) — это область компьютерных наук, занимающаяся созданием систем, способных выполнять задачи, требующие интеллекта, такие как восприятие, рассуждение, обучение и принятие решений. Основные направления развития ИИ включают машинное обучение, обработку естественного языка, компьютерное зрение и робототехнику.
2	Контролируемое обучение подразумевает наличие размеченных данных, где алгоритм обучается на входных данных с известными выходами (например, классификация). Неконтролируемое обучение работает с неразмеченными данными, где алгоритм ищет скрытые структуры (например, кластеризация). Примером контролируемого обучения является задача распознавания изображений, а неконтролируемого — сегментация клиентов на основе их покупательского поведения.
3	Многослойная нейронная сеть состоит из входного слоя, одного или нескольких скрытых слоев и выходного слоя. Каждый нейрон в слое связан с нейронами следующего слоя. Функция активации определяет, будет ли нейрон активирован, основываясь на взвешенной сумме входных сигналов. Она добавляет нелинейность в модель, позволяя сети обучаться сложным паттернам.
4	Основные задачи NLP включают анализ тональности, машинный перевод, извлечение информации и генерацию текста. Например, анализ тональности может использоваться для определения настроения в отзывах, а машинный перевод — для перевода текста с одного языка на другой.

№ вопроса	Верный ответ
5	Этические проблемы включают предвзятость алгоритмов, нарушение конфиденциальности и отсутствие прозрачности в принятии решений. Решения могут включать разработку этических стандартов, использование открытых данных для обучения и создание механизмов для объяснения решений ИИ.
6	Метод градиентного спуска — это итеративный алгоритм, используемый для минимизации функции потерь. Он обновляет веса нейронной сети, двигаясь в направлении отрицательного градиента функции потерь. Это позволяет находить оптимальные параметры модели, улучшая ее предсказательную способность.
7	ИИ применяется в здравоохранении (диагностика заболеваний), финансах (анализ рисков), производстве (предиктивное обслуживание) и маркетинге (персонализированные рекомендации). ИИ изменяет эти отрасли, повышая эффективность, снижая затраты и улучшая качество обслуживания.
8	Генеративные модели — это модели, которые могут генерировать новые данные, похожие на обучающие. Примеры включают генеративные состязательные сети (GAN) и вариационные автокодировщики (VAE). Они используются для создания изображений, музыки и текстов.
9	Обучение с подкреплением — это метод, при котором агент обучается принимать решения, получая награды или штрафы за свои действия. В отличие от контролируемого и неконтролируемого обучения, здесь агент взаимодействует с окружающей средой и учится на основе своих действий и их последствий.
10	Будущее ИИ обещает значительные достижения в области автономных систем, улучшенной обработки данных и взаимодействия человека с машиной. Ожидается развитие технологий, таких как квантовые вычисления, улучшенные алгоритмы машинного обучения и более совершенные системы взаимодействия на основе естественного языка.
11	1. Генерируем случайные значения (x) в диапазоне [0, 6]. 2. Вычисляем значение функции для каждого (x). 3. Находим минимальное значение функции и соответствующее значение (x).
12	1. Начинаем с произвольного значения (x) в диапазоне. 2. Находим значение функции и проверяем соседние значения (x). 3. Если соседнее значение дает большее значение функции, перемещаемся к нему. 4. Повторяем, пока не достигнем максимума.
13	1. Начинаем с случайного значения (x). 2. Устанавливаем начальную температуру и уменьшаем ее в процессе. 3. На каждом шаге генерируем новое значение (x') и вычисляем (f(x')). 4. Принимаем новое значение с определенной вероятностью, даже если оно хуже.
14	1. Генерируем начальную популяцию случайных значений (x). 2. Оцениваем каждое значение по функции. 3. Отбираем лучшие значения для скрещивания. 4. Применяем мутацию и создаем новую популяцию. 5. Повторяем процесс до достижения максимума.
15	1. Загружаем набор данных MNIST. 2. Создаем модель нейронной сети с несколькими слоями. 3. Обучаем модель на обучающем наборе данных. 4. Оцениваем точность на тестовом наборе данных.
16	1. Собираем набор данных с текстами и метками. 2. Преобразуем текст в векторы (например, с помощью TF-IDF или word embeddings). 3. Создаем и обучаем нейронную сеть для классификации. 4. Оцениваем точность модели на тестовом наборе.
17	1. Определите диапазоны для гиперпараметров (например, C и gamma). 2. Генерируйте случайные комбинации гиперпараметров. 3. Обучайте модель SVM и оценивайте ее точность. 4. Выберите наилучшие гиперпараметры на основе результатов.
18	1. Инициализация: Выберите начальную точку ((x_0, y_0)). 2. Оценка соседей: Для текущей точки ((x, y)) вычислите значения функции для соседних точек. Соседние точки можно получить, добавляя или вычитая небольшое значение (\epsilon) к (x) и (y). 3. Выбор лучшего соседа: Если значение функции в одной из соседних точек меньше, чем в текущей, переместитесь в эту соседнюю точку. 4. Повторение: Повторяйте шаги 2 и 3, пока не достигнете точки, в которой все соседние точки имеют большее значение функции, чем текущая.
19	1. Создайте начальную популяцию маршрутов. 2. Оценивайте каждый маршрут по длине.

№ вопроса	Верный ответ
	3. Отбирайте лучшие маршруты для скрещивания. 4. Применяйте мутацию и создавайте новые маршруты. 5. Повторяйте процесс до нахождения оптимального маршрута.
20	1. Соберите набор данных с размеченными изображениями. 2. Создайте модель нейронной сети (например, U-Net). 3. Обучите модель на размеченных данных. 4. Оцените качество сегментации на тестовых изображениях.
21	В обучении с подкреплением агент — это сущность, которая принимает решения и обучается взаимодействовать с окружающей средой. Среда представляет собой все, что окружает агента и может влиять на его действия, также она предоставляет агенту информацию о текущем состоянии и награду. Основная цель агента — выработать стратегию, которая будет максимизировать ожидаемую суммарную награду.
22	Агент достигает оптимальной стратегии, когда значения Q-функции стабилизируются и не изменяются значительно с каждым новым обновлением. Это означает, что агент выработал устойчивую стратегию, при которой он максимизирует свои долгосрочные награды.
23	Если все Q-значения равны, агент может использовать случайный выбор действий (например, метод \(\epsilon\)-жадности). В этом методе агент с вероятностью \(\epsilon\) выбирает случайное действие, а с вероятностью \(1 - \epsilon\) — действие с наибольшим Q-значением. Это позволяет агенту исследовать среду, чтобы найти лучшие действия.
24	Параметр \(\alpha\) определяет, насколько быстро агент обновляет свои Q-значения. Высокое значение \(\alpha\) ускоряет обучение, но может сделать оценку нестабильной. Параметр \(\gamma\) отвечает за учет будущих наград: чем выше \(\gamma\), тем больше агент ориентирован на долгосрочные награды, а не на краткосрочные выгоды. Оптимальное сочетание этих параметров важно для успешного обучения.
25	Q-обучение — это метод, при котором агент обучается, обновляя Q-значения для каждой пары "состояние-действие" на основе текущего вознаграждения и максимального ожидаемого значения будущего состояния. Основные шаги: 1. Агент находится в состоянии \(s\) и выбирает действие \(a\). 2. Переходит в новое состояние \(s'\) и получает награду \(r\). 3. Обновляет значение Q-функции: \(Q(s, a) = Q(s, a) + \alpha [r + \gamma \max_a Q(s', a) - Q(s, a)]\), где \(\alpha\) — скорость обучения, а \(\gamma\) — коэффициент дисконтирования. 4. Повторяет процесс до тех пор, пока не выработает оптимальную стратегию.
26	Состояние — это информация, которая характеризует текущее положение агента в среде. Действие — это выбор агента, как реагировать на текущее состояние, чтобы перейти в другое состояние. Награда — числовое значение, которое агент получает за выполнение действия в конкретном состоянии. Цель агента — выбирать действия, которые приведут к максимальной суммарной награде в долгосрочной перспективе.
27	Функция ценности состояния \(V(s)\) оценивает, насколько полезно находиться в состоянии \(s\), учитывая долгосрочные награды, которые агент может получить, начиная из этого состояния и следуя своей текущей стратегии. Она помогает агенту определить, какие состояния являются более предпочтительными для достижения максимальной награды.
28	Q-функция, или функция качества действия \(Q(s, a)\), оценивает "качество" пары "состояние-действие" \(s, a\), то есть ожидаемую награду, если агент выполнит действие \(a\) в состоянии \(s\) и затем будет следовать своей стратегии. Q-функция позволяет агенту сравнивать различные действия и выбирать то, которое ведет к наибольшей награде в долгосрочной перспективе.
29	Функция ценности состояния \(V(s)\) — это обобщение Q-функции, представляющее максимум ожидаемой награды в состоянии \(s\). Q-функция же уточняет выбор, показывая, какое действие \(a\) выбрать в состоянии \(s\). Используя обе функции, агент может находить оптимальную стратегию, учитывая как ценность состояния, так и качество действий.
30	Агент может оптимизировать свою стратегию, выбирая действия, которые максимизируют Q-функцию или функцию ценности состояния. В процессе обучения агент обновляет свои оценки ценности и качества, приближаясь к оптимальной стратегии. Это помогает ему делать более обоснованные действия, ведущие к наилучшим возможным наградам.
31	Алгоритм DQN включает следующие этапы: 1. Инициализация сети: Используется нейронная сеть для приближения Q-функции. 2. Опытное хранилище (Replay Buffer): Агент сохраняет свои прошлые переходы \((s, a, r, s')\) для повторного обучения. Это устраняет корреляцию между последовательными наблюдениями и увеличивает стабильность обучения. 3. Целевая сеть (Target Network): Параллельно основной сети создается копия (целевая сеть), которая обновляется каждые несколько шагов. Это позволяет уменьшить нестабильность Q-обновлений.

№ вопроса	Верный ответ
	4. Обучение: Обновление сети на мини-батчах из хранилища опыта с использованием метода градиентного спуска для минимизации ошибки Беллмана. Основные улучшения: - Double DQN— разделяет выбор действия и оценку качества действия. - Dueling DQN— использует отдельные сети для оценки ценности состояния и преимущества действия. Эти изменения помогают снизить переобучение и повысить стабильность алгоритма.
32	Алгоритм REINFORCE — это метод градиента стратегии, который обновляет параметры \(\theta\) политики агента \(\pi_\theta\) для максимизации ожидаемой награды. Основные этапы: 1. Генерация эпизодов: Агент взаимодействует со средой, формируя эпизоды на основе текущей политики. 2. Оценка награды: После завершения эпизода агент подсчитывает награду \(G_t\) для каждого шага \(t\). 3. Обновление градиента: Градиент функции стоимости J рассчитывается как \(\nabla J(\theta) = \mathbb{E}[\nabla_\theta \log \pi_\theta(a|s) G_t]\), и параметры обновляются в направлении этого градиента. Плюсы: - Простой метод, не требует Q-функции. Минусы: - Высокая дисперсия оценок, что делает обучение медленным и нестабильным.
33	В алгоритме Actor-Critic: - Актор (Actor)обновляет политику \(\pi_\theta(a|s)\) для выбора действий. - Критик (Critic)оценивает действие актера, используя функцию ценности \(V(s)\), и вычисляет разность \(A(s, a) = Q(s, a) - V(s)\) (Advantage). Основное отличиеот REINFORCE заключается в том, что в Actor-Critic есть критик, который уменьшает дисперсию оценок награды, что ускоряет обучение. Актор обновляется с учетом оценки критика, что снижает нестабильность и дисперсию обновлений.
34	В A2C агент использует Advantage (выгоду) для обновления параметров актера: 1. Выгода (Advantage): A2C оценивает разницу между фактической наградой и прогнозируемой ценностью \(A(s, a) = Q(s, a) - V(s)\), что уменьшает дисперсию градиента. 2. Параллельные агенты: A2C запускает несколько агентов параллельно, что повышает эффективность обучения и позволяет быстрее находить глобальный минимум. Использование Advantage позволяет агенту более точно оценивать качество действий и увеличивает стабильность по сравнению с базовым Actor-Critic методом.
35	PPO — это метод, который ограничивает изменения политики с использованием ограничений на величину обновления. Основные принципы: 1. **Вырезка** : PPO ограничивает изменение политикиπ_θ, добавляя срезку к ощущениям вероятностей$\frac{\pi_\theta(a\|c)}{\pi_{\theta_{старый}}(a\|c)}$. Это позволяет избежать больших скачков при обновлении параметров и повысить стабильность. 2. **Функция потерь** : Использует "обрезанную" функцию потерь. мин$(r(\theta)A, clip(r(\theta), 1-\epsilon, 1+\epsilon)A)$, что стабилизирует обучение, позволяет контролировать изменения в защите. PPO улучшает A2C, так как предотвращает большие отклонения в параметрах, что снижает вероятность падения производительности.
36	DDPG (Deep Deterministic Policy Gradient) — это метод глубокого обучения с подкреплением для задач с непрерывным пространством действий. Компоненты: 1. Актор-Критик: В DDPG используются сети актор и критик, где актор выбирает действие, а критик оценивает его качество. 2. Целевые сети: Для стабильности используются целевые сети для актора и критика, которые обновляются медленно для снижения шумов. 3. Шумовой процесс: Для исследования DDPG добавляет шумы к действиям (например, процесс Орнштейна-Уленбека) для случайности в действиях. DDPG позволяет решать задачи с непрерывным пространством действий, так как использует непрерывные функции для выбора действий и обучения, в отличие от DQN, который ориентирован на дискретные действия.

Задание № 1

Дайте определение искусственного интеллекта. Какие основные направления его развития вы можете выделить?

Задание № 2

Опишите различия между контролируемым и неконтролируемым обучением. Приведите примеры.

Задание № 3

Объясните, как работает многослойная нейронная сеть. Какова роль функции активации?

Задание № 4

Какие основные задачи решает обработка естественного языка (NLP)? Приведите примеры.

Задание № 5

Какие этические проблемы могут возникнуть при использовании ИИ? Как их можно решить?

Задание № 6

Опишите метод градиентного спуска. Как он используется в обучении нейронных сетей?

Задание № 7

Приведите примеры применения ИИ в различных отраслях. Как ИИ изменяет эти отрасли?

Задание № 8

Что такое генеративные модели? Приведите примеры их использования.

Задание № 9

Объясните, что такое обучение с подкреплением. Как оно отличается от других типов обучения?

Задание № 10

Как вы видите будущее искусственного интеллекта? Какие технологии могут появиться в ближайшие годы?

Задание № 11

Найдите минимум функции (f(x) = (x - 3)^2 + 4) с помощью случайного поиска. Ограничьте область поиска от 0 до 6.

Задание № 12

Используя метод hill climb, найдите максимум функции (f(x) = -x^2 + 4x) в диапазоне от 0 до 4.

Задание № 13

Примените метод отжига для нахождения минимума функции (f(x) = x^4 - 3x^3 + 2) в диапазоне от -2 до 3.

Задание № 14

Используя генетический алгоритм, найдите максимум функции (f(x) = x \sin(10\pi x) + 1) на интервале [0, 1].

Задание № 15

Обучите нейронную сеть для классификации изображений из набора данных MNIST.

Задание № 16

Реализуйте простую модель для классификации текстов на основе нейронной сети.

Задание №17

Используйте случайный поиск для оптимизации гиперпараметров модели SVM.

Задание №18

Найдите минимум функции (f(x, y) = (x-1)^2 + (y-2)^2) с помощью метода hill climb.

Задание №19

Реализуйте генетический алгоритм для решения задачи коммивояжера.

Задание №20

Реализуйте нейронную сеть для сегментации изображений.

Задание №21

Определите, что такое агент и среда в контексте обучения с подкреплением. Какова роль агента и среды в этой модели?

Задание №22

Как агент определяет, что достиг оптимальной стратегии в Q-обучении? Что является индикатором того, что обучение завершено?

Задание №23

Представьте, что агент находится в состоянии, в котором Q-значения для всех действий равны. Какой стратегии может следовать агент, чтобы выбрать действие?

Задание №24

Какое влияние на Q-обучение оказывают параметры \(\alpha\) (скорость обучения) и \(\gamma\) (коэффициент дисконтирования)?

Задание №25

Как работает алгоритм Q-обучения? Опишите шаги, которые выполняет агент в этом алгоритме.

Задание №26

Объясните понятия состояния, действия и награды. Как они связаны друг с другом в обучении с подкреплением?

Задание №27

Что такое функция ценности состояния (Value Function) и какую роль она играет в обучении с подкреплением?

Задание №28

Опишите функцию качества действия (Q-функцию). Какое значение имеет Q-функция в процессе обучения с подкреплением?

Задание №29

Как связаны между собой функция ценности состояния и Q-функция? Почему обе эти функции необходимы для обучения агента?

Задание №30

Опишите, как агент может оптимизировать свою стратегию с помощью максимизации функций ценности и качества. Как это влияет на его обучение?

Задание №31

Реализация Deep Q-Networks (DQN). Опишите основные этапы алгоритма Deep Q-Networks. Какие основные различия между DQN и классическим Q-обучением? Как можно улучшить стабильность и производительность DQN?

Задание №32

Описать алгоритм REINFORCE для обучения агента с подкреплением. Какие ключевые особенности и ограничения этого метода?

Задание №33

Алгоритм Actor-Critic. Поясните, как работает метод Actor-Critic. Чем он отличается от REINFORCE? Какие задачи решает использование "критика"?

Задание №34

A2C (Advantage Actor-Critic).

Как работает алгоритм A2C? Почему использование Advantage помогает улучшить обучение? Какие преимущества A2C имеет перед обычным Actor-Critic?

Задание №35

Объясните основные принципы PPO и его улучшения по сравнению с A2C.

Какое значение имеет "клиппинг" в PPO?

Задание №36
Описать алгоритм DDPG и его компоненты. Почему DDPG эффективен для задач с непрерывным пространством действий?

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

Промежуточная аттестация обучающихся обеспечивает оценивание промежуточных и окончательных результатов обучения по дисциплине (модулю).

Преподаватель доводит до сведения обучающихся на первом учебном занятии перечень вопросов, выносимых на промежуточную аттестацию, и критерии оценивания знаний, умений и навыков.

Зачеты проводятся в день последнего в данном семестре занятия по соответствующей дисциплине (модулю) в соответствии с рабочей программой.

Зачеты служат формой проверки освоения учебного материала дисциплины (модуля).

Результаты сдачи зачета оцениваются по шкале: «зачтено», «не зачтено».

В целях поощрения обучающихся за систематическую активную работу на учебных занятиях и на основании успешного прохождения текущего контроля допускается выставление зачетной оценки без процедуры сдачи зачета.

Зачеты принимаются преподавателями, ведущим занятия по дисциплине (модулю).

Зачет проводится по билетам. Преподаватель имеет право задавать дополнительные вопросы в соответствии с рабочей программой дисциплины (модуля).

Во время зачета с разрешения преподавателя обучающийся может пользоваться справочниками, таблицами, инструкциями и другими материалами.

Обучающийся при подготовке ответа ведет необходимые записи, которые предъявляет преподавателю. При отказе обучающегося от ответа в зачетно-экзаменационную ведомость проставляется оценка «не зачтено».

Результаты зачета заносятся в зачетно-экзаменационную ведомость. Если обучающийся не явился на зачет, в ведомости напротив фамилии обучающегося делается запись «не явился». Неявка на зачет без уважительной причины приравнивается к оценке «не зачтено».

В зачетную книжку выставляется соответствующая оценка, полученная обучающимся. Оценка «не зачтено» в зачетную книжку не ставится. Заполнение зачетной книжки до внесения соответствующей оценки в ведомость не разрешается.

Прием зачета у обучающегося прекращается при нарушении им дисциплины, использовании неразрешенных материалов и средств мобильной связи. В этом случае обучающемуся в зачетно-экзаменационную ведомость проставляется оценка «не зачтено».

В случае несогласия обучающегося с оценкой, выставленной на зачете, он имеет право подать апелляцию.