



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное учреждение
высшего образования

«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ФГБОУ ВО «ИГУ»

Факультет бизнес-коммуникаций и информатики

Кафедра естественнонаучных дисциплин

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине Б1.В.09 Машинное обучение с подкреплением

направление подготовки 09.04.03 Прикладная информатика

направленность (профиль) Большие данные и интеллектуальные системы

Одобрено
УМК факультета бизнес-коммуникаций
и информатики

Разработан в соответствии с ФГОС ВО

с учетом требований проф. стандарта

Председатель УМК

М.Г. Синчурина

ФИО, должность, ученая степень, звание

подпись, печать

Разработчики:

(подпись)

доцент

(занимаемая должность)

О.И. Бернгардт

(инициалы, фамилия)

Цель фонда оценочных средств. Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Б1.В.09 Машинное обучение с подкреплением». Перечень видов оценочных средств соответствует рабочей программе дисциплины.

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля (в следующих формах: тест, устный опрос, практическое задание, доклад-/презентация) и промежуточной аттестации в форме вопросов и заданий к зачету с оценкой.

Структура и содержание заданий – задания разработаны в соответствии с рабочей программой дисциплины «Б1.В.09 Машинное обучение с подкреплением».

1. Паспорт фонда оценочных средств

Компетенция	Индикаторы компетенций	Результаты обучения
<p>ПК-3 Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения отраслевых задач, проводить аналитические исследования с применением технологий больших данных</p>	ПК-3.1	<p>Знание основных концепций и принципов машинного обучения. Понимание различных типов алгоритмов машинного обучения, их принципов работы и областей применения. Знание основных методов предобработки данных и подготовки признаков для обучения моделей машинного обучения. Знание принципов оценки производительности и качества моделей машинного обучения</p>
	ПК-3.2	<p>Умение выбирать подходящие методы и алгоритмы машинного обучения для конкретной отраслевой задачи. Навыки разработки и настройки моделей машинного обучения с учетом требований отраслевых задач. Умение анализировать и интерпретировать результаты работы моделей машинного обучения. Умение оптимизировать и настраивать модели машинного обучения для достижения требуемой производительности и качества решений</p>
	ПК-3.3	<p>Владение языками программирования, используемыми в машинном обучении . Владение основными библиотеками и инструментами для разработки и применения моделей машинного обучения. Владение техниками и методами работы с большими объемами данных для обучения моделей машинного обучения. Владения навыками работы с различными типами задач и отраслями, что позволяет эффективно адаптировать методы и алгоритмы машинного обучения под конкретные требования</p>

2. Показатели и критерии оценивания компетенций, шкалы оценивания

2.1. Показатели и критерии оценивания компетенций

№ п/п	Раздел, тема	Код индикатора компетенции	Наименование ОС	
			ТК	ПА
1	Введение	ПК-3.1	Тест, УО	Тест
2	Нейронные сети и методы машинного обучения	ПК-3.2, ПК-3.1, ПК-3.3	Тест, УО, Пз	Тест
3	Базовые понятия. Марковский процесс принятия решений. Динамическое программирование	ПК-3.1, ПК-3.2, ПК-3.3	Тест, УО, Пз	Тест
4	Аппроксимация в алгоритмах обучения с подкреплением.	ПК-3.1	Тест, Пз	Тест
5	Градиент стратегии	ПК-3.1, ПК-3.2, ПК-3.3	Тест, УО, Пз	Тест
6	Алгоритм Актора-Критика	ПК-3.3, ПК-3.1, ПК-3.2	Тест, УО, Пз	Тест
7	Оптимизация градиента стратегии	ПК-3.1, ПК-3.2, ПК-3.3	Тест, УО, Пз	Тест
8	Обучение с подкреплением на основе модели мира	ПК-3.1, ПК-3.2, ПК-3.3	Тест, УО, Пз	Тест
9	Многорукие бандиты. Жадные алгоритмы	ПК-3.2, ПК-3.1	Тест, УО, Пз	Тест
10	Мотивация и ускорение исследования среды	ПК-3.1, ПК-3.2	Тест, Пз	Тест
11	Интеграция планирования в среде, алгоритм MCTS	ПК-3.1, ПК-3.2	Тест, Пз	Тест
12	Обучение на основе демонстраций. Подражательное обучение	ПК-3.3, ПК-3.1, ПК-3.2	Тест, УО, Пз	Тест
13	Автономное обучение с подкреплением и трансформеры принятия решений	ПК-3.2, ПК-3.1	Тест, УО, Пз	Тест
14	Перспективы и актуальные проблемы в обучении с подкреплением	ПК-3.1, ПК-3.2	Тест, УО, Д	Тест

2.2. Критерии оценивания результатов обучения для текущего контроля успеваемости и промежуточной аттестации

Оценочное средство	Критерии оценивания	Шкала оценивания
Тест	Студентом даны правильные ответы на 91-100% заданий	Отлично
	Студентом даны правильные ответы на 81-90% заданий	Хорошо
	Студентом даны правильные ответы на 71-80% заданий	Удовлетворительно
	Студентом даны правильные ответы менее чем на 70% заданий	Неудовлетворительно

Оценочное средство	Критерии оценивания	Шкала оценивания
Устный опрос	<p>Ответ соответствует поставленной теме и содержит ответы на поставленные задачи, имеет четкую структуру, логически сопоставляемую с поставленными вопросами. Ответ демонстрирует способность анализировать и обобщать информацию, опираясь на знания, полученные в ходе изучения темы, а также демонстрировать самостоятельность автора в решении поставленных задач. Ответ содержит качественную речь и аргументацию, которая убедительно подтверждает выводы и ответы на поставленные вопросы</p>	Отлично
	<p>Ответ должен быть направлен на ответ на поставленные вопросы и соответствовать поставленной теме, иметь логическую цепочку рассуждений и четко демонстрировать связь между поставленными вопросами. Ответ выдержан в четкой форме, быть грамотно и без ошибок озвучен, выделены ключевые термины. Ответ должен демонстрировать способность анализировать и критически оценивать информацию, выбирая ключевые аспекты и выделяя главные выводы</p>	Хорошо
	<p>Ответ должен соответствовать поставленной теме и содержать ответы на поставленные вопросы, должен содержать существенную информацию, ясно передавать ответы и идеи. Ответ должен содержать достаточное количество аргументов и примеров, связанных с темой работы и позволяющих изложить свою точку зрения. Ответ должен быть грамотно сформулирован</p>	Удовлетворительно
	<p>Ответ не соответствует поставленной теме или не содержит ответов на поставленные задачи, содержит недостаточно аргументации и примеров, которые подтверждают высказанные в ответе идеи и выводы. Ответ не соответствует логической цепочке рассуждений и не выполняет требования логической последовательности высказывания, затрудняющей понимание ответа. Ответ содержит грубые ошибки, что затрудняет понимание высказывания</p>	Неудовлетворительно

Оценочное средство	Критерии оценивания	Шкала оценивания
Практическое задание	Задание выполнено верно. Выбран оптимальный путь решения. Присутствует развернутое описание алгоритма решения	Отлично
	Задание выполнено верно. Допущены негрубые логические ошибки при описании алгоритма решения. Отсутствуют пояснения к решению задания	Хорошо
	Ход решения задания верный, но допущены ошибки приведшие к неправильному ответу	Удовлетворительно
	В работе получен неверный ответ, связанный с грубыми ошибками допущенными в ходе решения, либо решение отсутствует полностью	Неудовлетворительно
Доклад/презентация	Обучающийся демонстрирует исчерпывающее знание материала и последовательно, четко и логично излагает материал, хорошо ориентируется в материале, не затрудняется с ответом на сопутствующие вопросы	Отлично
	Обучающийся демонстрирует знание материала, не допускает существенных неточностей. Ответы на дополнительные вопросы в целом верные, но содержащие отдельные пробелы	Хорошо
	Обучающийся демонстрирует знания только основного материала, но не знает деталей, допускает неточности как в докладе, так и в ответах на вопросы	Удовлетворительно
	Обучающийся не знает значительной части программного материала, плохо ориентируется в материале, допускает существенные ошибки, выступает неуверенно, с большими затруднениями	Неудовлетворительно

2.3. Оценочные средства для текущего контроля (примеры)

2.3.1. Материалы для компьютерного тестирования обучающихся

Общие критерии оценивания

Процент правильных ответов	Оценка
91% – 100%	5 (отлично)
81% – 90%	4 (хорошо)
71% – 80%	3 (удовлетворительно)
Менее 70%	2 (неудовлетворительно)

Соответствие вопросов теста индикаторам формируемых и оцениваемых компетенций

№ вопроса в тесте	Код индикатора компетенции
1	ПК-3.1
2	ПК-3.1
3	ПК-3.1
4	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2
5	ПК-3.3
6	ПК-3.2
7	ПК-3.3
8	ПК-3.1
9	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2
10	ПК-3.1
11	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2, ПК-3.3
12	ПК-3.3
13	ПК-3.2
14	ПК-3.1
15	ПК-3.1
16	ПК-3.1
17	ПК-3.2

Ключ ответов

№ вопроса в тесте	Номер ответа (или ответ, или соответствие)
1	b
2	b, c
3	b
4	b
5	a, d
6	c
7	d
8	a, b
9	c
10	c
11	b
12	a
13	a
14	b, c
15	b
16	a
17	b

Перечень тестовых вопросов

№ 1. Задание с единичным выбором. Выберите один правильный ответ.

На вход нейрона подается два числа, сколько чисел будет на выходе нейрона

- a. 2
- b. 1
- c. 4
- d. 3
- e. больше 4

№ 2. Задание с множественным выбором. Выберите 2 правильных ответа.

Обучение с подкреплением требует...

- a. нейронной сети
- b. агента
- c. окружения
- d. механизмов наследования
- e. функции потерь
- f. графа соответствий

№ 3. Задание с единичным выбором. Выберите один правильный ответ.

Какие виды мотивации обычно разделяют в машинном обучении

- a. верная и неверная
- b. внутренняя и внешняя

№ 4. Задание с единичным выбором. Выберите один правильный ответ.

Уравнение Белмана позволяет...

- a. найти оптимальную стратегию
- b. рекурсивно определить оптимальные функции ценности
- c. найти минимальный доход

№ 5. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верное утверждение

- a. Метод актор-критик способен обучаться на длинных (непрерывных) задачах, а REINFORCE - нет.
- b. Метод актор-критик сходится медленнее, чем REINFORCE.
- c. Метод актор-критик не способен обучаться на длинных (непрерывных) задачах, а REINFORCE - способен.
- d. Метод актор-критик сходится быстрее, чем REINFORCE.

№ 6. Задание с единичным выбором. Выберите один правильный ответ.

Алгоритм DDPG основан на

- a. Методе Актора-Критика
- b. Q-обучении
- c. Всем перечисленном

№ 7. Задание с единичным выбором. Выберите один правильный ответ.

Выберите неверное утверждение

- a. Алгоритмы подражательного обучения могут быть активными и пассивными

- b. Цель подражательного обучения - перенять стратегию экспертов
- c. Алгоритмы подражательного обучения используют опыт экспертов
- d. При подражательном обучении ученик может превзойти эксперта

№ 8. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верные ответы

- a. стратегия бывает детерминированной и стохастической
- b. стратегия определяет, какое действие предпринять в данной ситуации.
- c. стратегия бывает верной и неверной
- d. стратегия бывает доходной и бездоходной

№ 9. Задание с единичным выбором. Выберите один правильный ответ.

Уравнение Белмана позволяет...

- a. найти минимальный доход
- b. найти оптимальную стратегию
- c. рекурсивно определить оптимальные функции ценности

№ 10. Задание с единичным выбором. Выберите один правильный ответ.

Алгоритм MCTS обычно используют

- a. В играх с бесконечно большим количеством вариантов ходов
- b. В играх с малым количеством вариантов ходов
- c. В играх с большим количеством вариантов ходов

№ 11. Задание с единичным выбором. Выберите один правильный ответ.

основа методов, основанных на модели...

- a. создать модель среды
- b. предсказать переходы состояний и вознаграждения для каждой пары состояние–действие, не взаимодействуя с реальной средой
- c. оптимизировать стратегию методом градиента стратегии

№ 12. Задание с единичным выбором. Выберите один правильный ответ.

Для чего используется метод градиентного спуска в машинном обучении?

- a. Для поиска минимума функции потерь нейронной сети
- b. Для поиска оптимального размера обучающей выборки

№ 13. Задание с единичным выбором. Выберите один правильный ответ.

Многорукий бандит это...

- a. задача оптимизации выигрыша при возможности осознанного выбора противника в игре
- b. нейронная сеть с множественными выходами, обычно классификационная
- c. алгоритм для оптимизации поисковой выдачи

№ 14. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верные ответы

- a. стратегия бывает верной и неверной
- b. стратегия бывает детерминированной и стохастической
- c. стратегия определяет, какое действие предпринять в данной ситуации.
- d. стратегия бывает доходной и бездоходной

№ 15. Задание с единичным выбором. Выберите один правильный ответ.

Какую задачу решают, когда обучают нейронную сеть?

- a. максимизацию функции активации
- b. минимизацию функции потерь

№ 16. Задание с единичным выбором. Выберите один правильный ответ.

Какие методы аппроксимации функций используют в RL?

- a. Все перечисленное
- b. Нейронными сетями
- c. Регрессионными деревьями
- d. RBF-сетями
- e. Линейная и полиномиальная

№ 17. Задание с единичным выбором. Выберите один правильный ответ.

Что такое скорость обучения?

- a. Количество эпох обучения за единицу времени
- b. Коэффициент пропорциональности при градиенте в градиентных методах обучения
- c. Количество эпох, необходимое для получения оптимального результата

2.3.2. Практические задания для оценки компетенции «ПК-3.2»

№ 1. Архитектуры и методы обучения нейронных сетей.

Решить задачи по созданию нейронных сетей на классических датасетах прогноза и классификации

№ 2. Марковский процесс принятия решений.

Построить решение задачи, использующей марковский процесс принятия решений

№ 3. Марковский процесс принятия решений.

Построить решение задачи, использующей марковский процесс принятия решений

№ 4. Реализовать решение задачи на основе алгоритма REINFORCE.

Реализовать решение задачи на основе алгоритма REINFORCE

№ 5. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 6. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 7. Реализовать решение задачи на основе алгоритма REINFORCE.

Реализовать решение задачи на основе алгоритма REINFORCE

№ 8. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 9. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 10. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 11. Многорукые бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 12. Многорукые бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 13. Многорукие бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 14. MuJoCo: Ant2.

Написать алгоритм для максимально быстрого передвижения робота Ant2 (продолжение работы в классе)

№ 15. Реализовать MCTS для простой настольной игры.

Реализовать MCTS для игры в крестики-нолики

№ 16. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

№ 17. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

№ 18. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

№ 19. Обучение автономного воплощенного агента.

Обучение автономного воплощенного агента (продолжение классной работы)

2.3.3. Практические задания для оценки компетенции «ПК-3.3»

№ 20. Марковский процесс принятия решений.

Построить решение задачи, использующей марковский процесс принятия решений

№ 21. Марковский процесс принятия решений.

Построить решение задачи, использующей марковский процесс принятия решений

№ 22. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 23. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 24. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 25. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 26. Решение игры FrozenLake.

Решение игры FrozenLake (openai-gym) методами динамического программирования

№ 27. Многорукие бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 28. Многорукие бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 29. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

№ 30. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

№ 31. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

2.3.4. Практические задания для оценки компетенции «ПК-3.1»

№ 32. Методы аппроксимации.

Рассмотреть методы аппроксимации заданной функции - линейную, полиномиальную, нейронной сетью, RBF-сетью, регрессионным деревом и их ансамблем.

(продолжение классной работы)

№ 33. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 34. Метод актора-критика.

Решить задачу посадки корабля на Луну методом актора-критика

№ 35. Многорукие бандиты.

Решить задачу многорукого бандита методом эpsilon-жадного агента

№ 36. Решение задачи Flappy Bird.

Решение задачи Flappy Bird методами подражательного обучения в среде PyGame LearningEnvironment (PLE)

2.3.5. Доклады/презентации для оценки компетенции «ПК-3.1»

№ 37. Перспективы и актуальные проблемы в обучении с подкреплением.

Перспективы и актуальные проблемы в обучении с подкреплением

3. Промежуточная аттестация

3.1. Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и опыта деятельности

Зачет является заключительным этапом процесса формирования компетенций обучающегося при изучении дисциплины и имеет целью проверку и оценку знаний обучающегося по теории, и применению полученных знаний, умений и навыков при решении практических задач.

Зачет проводится по расписанию, сформированному учебно-методическим управлением, в сроки, предусмотренные календарным учебным графиком. Зачет принимается преподавателем, ведущим лекционные занятия.

Зачет проводится только при предъявлении обучающимся зачетной книжки и при условии выполнения всех контрольных мероприятий, предусмотренных учебным планом и рабочей программой дисциплины. Обучающимся на зачете представляется право выбрать один из билетов. Время подготовки к ответу составляет 30 минут. По истечении установленного времени обучающийся должен ответить на вопросы билета. Результаты зачета оцениваются по четырехбалльной системе и заносятся в зачетно-экзаменационную ведомость и зачетную книжку. В зачетную книжку заносятся только положительные оценки. Подписанный преподавателем экземпляр ведомости сдается не позднее следующего дня в деканат.

В случае неявки обучающегося на зачет в зачетно-экзаменационную ведомость делается отметка «не явка». Обучающиеся, не прошедшие промежуточную аттестацию по

дисциплине, должны ликвидировать академическую задолженность в установленном локальными нормативными актами порядке.

3.2. Вопросы к зачету с оценкой

№	Вопрос	Код компетенции
1.	В каких задачах можно применять обучение с подкреплением? Привести примеры.	ПК-3.3
2.	Какие архитектуры и методы обучения нейронных сетей вы знаете?	ПК-3.1
3.	Описать, в каких задачах и как можно использовать марковский процесс принятия решений	ПК-3.1, ПК-3.1, ПК-3.1
4.	Что такое метод градиентного подъема, где его можно использовать?	ПК-3.1, ПК-3.2, ПК-3.1
5.	Что такое метод градиентного подъема, где его можно использовать?	ПК-3.1, ПК-3.2, ПК-3.1
6.	В каких задачах можно применять методы многоруких бандитов?	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2
7.	Дилемма разведки-эксплуатации, когда возникает, какие методы решения существуют	ПК-3.3
8.	В чем принцип работы случайного поиска по дереву (Monte Carlo Tree Search)	ПК-3.3
9.	В чем принцип обучения на основе демонстраций?	ПК-3.1
10.	Какие актуальные проблемы в обучении с подкреплением вам известны?	ПК-3.3

3.3. Тематика курсовых работ

По данной дисциплине выполнение курсовых проектов (работ) не предусматривается.

3.4. Материалы для компьютерного тестирования обучающихся

Общие критерии оценивания

Процент правильных ответов	Оценка
91% – 100%	5 (отлично)
81% – 90%	4 (хорошо)
71% – 80%	3 (удовлетворительно)
Менее 70%	2 (неудовлетворительно)

Соответствие вопросов теста индикаторам формируемых и оцениваемых компетенций

№ вопроса в тесте	Код индикатора компетенции
1	ПК-3.1
2	ПК-3.1
3	ПК-3.1
4	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2
5	ПК-3.3

№ вопроса в тесте	Код индикатора компетенции
6	ПК-3.2
7	ПК-3.3
8	ПК-3.1
9	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2
10	ПК-3.1
11	ПК-3.1, ПК-3.2, ПК-3.1, ПК-3.2, ПК-3.3
12	ПК-3.3
13	ПК-3.2
14	ПК-3.1
15	ПК-3.1
16	ПК-3.1
17	ПК-3.2

Ключ ответов

№ вопроса в тесте	Номер ответа (или ответ, или соответствие)
1	b
2	b, c
3	b
4	b
5	a, d
6	c
7	d
8	a, b
9	c
10	c
11	b
12	a
13	a
14	b, c
15	b
16	a
17	b

Перечень тестовых вопросов

№ 1. Задание с единичным выбором. Выберите один правильный ответ.

На вход нейрона подается два числа, сколько чисел будет на выходе нейрона

- a. 2
- b. 1
- c. 4
- d. 3

е. больше 4

№ 2. Задание с множественным выбором. Выберите 2 правильных ответа.

Обучение с подкреплением требует...

- a. нейронной сети
- b. агента
- c. окружения
- d. механизмов наследования
- e. функции потерь
- f. графа соответствий

№ 3. Задание с единственным выбором. Выберите один правильный ответ.

Какие виды мотивации обычно разделяют в машинном обучении

- a. верная и неверная
- b. внутренняя и внешняя

№ 4. Задание с единственным выбором. Выберите один правильный ответ.

Уравнение Белмана позволяет...

- a. найти оптимальную стратегию
- b. рекурсивно определить оптимальные функции ценности
- c. найти минимальный доход

№ 5. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верное утверждение

- a. Метод актер-критик способен обучаться на длинных (непрерывных) задачах, а REINFORCE - нет.
- b. Метод актер-критик сходится медленнее, чем REINFORCE.
- c. Метод актер-критик не способен обучаться на длинных (непрерывных) задачах, а REINFORCE - способен.
- d. Метод актер-критик сходится быстрее, чем REINFORCE.

№ 6. Задание с единственным выбором. Выберите один правильный ответ.

Алгоритм DDPG основан на

- a. Методе Актора-Критика
- b. Q-обучении
- c. Всем перечисленном

№ 7. Задание с единственным выбором. Выберите один правильный ответ.

Выберите неверное утверждение

- a. Алгоритмы подражательного обучения могут быть активными и пассивными
- b. Цель подражательного обучения - перенять стратегию экспертов
- c. Алгоритмы подражательного обучения используют опыт экспертов
- d. При подражательном обучении ученик может превзойти эксперта

№ 8. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верные ответы

- a. стратегия бывает детерминированной и стохастической
- b. стратегия определяет, какое действие предпринять в данной ситуации.

- c. стратегия бывает верной и неверной
- d. стратегия бывает доходной и бездоходной

№ 9. Задание с единственным выбором. Выберите один правильный ответ.

Уравнение Белмана позволяет...

- a. найти минимальный доход
- b. найти оптимальную стратегию
- c. рекурсивно определить оптимальные функции ценности

№ 10. Задание с единственным выбором. Выберите один правильный ответ.

Алгоритм MCTS обычно используют

- a. В играх с бесконечно большим количеством вариантов ходов
- b. В играх с малым количеством вариантов ходов
- c. В играх с большим количеством вариантов ходов

№ 11. Задание с единственным выбором. Выберите один правильный ответ.

основа методов, основанных на модели...

- a. создать модель среды
- b. предсказать переходы состояний и вознаграждения для каждой пары состояние–действие, не взаимодействуя с реальной средой
- c. оптимизировать стратегию методом градиента стратегии

№ 12. Задание с единственным выбором. Выберите один правильный ответ.

Для чего используется метод градиентного спуска в машинном обучении?

- a. Для поиска минимума функции потерь нейронной сети
- b. Для поиска оптимального размера обучающей выборки

№ 13. Задание с единственным выбором. Выберите один правильный ответ.

Многоармий бандит это...

- a. задача оптимизации выигрыша при возможности осознанного выбора противника в игре
- b. нейронная сеть с множественными выходами, обычно классификационная
- c. алгоритм для оптимизации поисковой выдачи

№ 14. Задание с множественным выбором. Выберите 2 правильных ответа.

Выберите верные ответы

- a. стратегия бывает верной и неверной
- b. стратегия бывает детерминированной и стохастической
- c. стратегия определяет, какое действие предпринять в данной ситуации.
- d. стратегия бывает доходной и бездоходной

№ 15. Задание с единственным выбором. Выберите один правильный ответ.

Какую задачу решают, когда обучают нейронную сеть?

- a. максимизацию функции активации
- b. минимизацию функции потерь

№ 16. Задание с единственным выбором. Выберите один правильный ответ.

Какие методы аппроксимации функций используют в RL?

- a. Все перечисленное

- b. Нейронными сетями
- c. Регрессионными деревьями
- d. RBF-сетями
- e. Линейная и полиномиальная

№ 17. Задание с единственным выбором. Выберите один правильный ответ.

Что такое скорость обучения?

- a. Количество эпох обучения за единицу времени
- b. Коэффициент пропорциональности при градиенте в градиентных методах обучения
- c. Количество эпох, необходимое для получения оптимального результата