Вопросы к экзамену

1. Цели и задачи обучения с подкреплением. Отличия от других видов машинного обучения
2. Основные понятия обучения с подкреплением. Схема взаимодействия агента со средой.
3. Работа со средой OpenAI Gym. Основные объекты и методы.
4. Задача с тележкой и стержнем. Характеристики. Возможные решения.
5. Постановка задачи многорукого бандита. Особенности задачи. Методы решения.
6. Эпсилон жадные стратегии. Алгоритм. Способы улучшения.
7. Алгоритм верхней доверительной границы.
8. Градиентный метод бандита.
9. Марковский процесс принятия решений в обучении с подкреплением. Основные понятия.
10. Определение дохода в обучении с подкреплением. Эпизодические и бесконечные задачи. Обесценивание.
11. Функция ценности. Фундаментальное свойство.
12. Оптимальная стратегия. Определение. Оптимальная ценность дейтсвий.
13. Уравнение Беллмана для ценности состояний и действий. Способы решений.
14. Итеративное оценивание стратегий.
15. Теорема об улучшении стратегии.
16. Алгоритм итераций по стратегиям.
17. Алгоритм итерации по ценности.
18. Метод Монте-Карло первого посещения. Алгоритм.
19. Метод Монте-Карло с исследовательскими стартами
20. Метод Монте-Карло без исследовательских стартов
21. Методы с разделенной стратегией
22. Метод Монте-Карло с разделенной стратегией для оценки ценности состояний
23. Метод Монте-Карло с разделенной стратегией для поиска оптимальной стратегии
24. Обучение на основе временных различий. TD(0).
25. SARSA.
26. Q-обучение.
27. Определение СППР. Требования к критериям
28. Методы многокритериального анализа альтернатив
29. Инструментальные средства решения задач оптимизации

При ответе на вопрос нужно привести пример задачи, где используется данная теория.

Если в вопросе есть описание алгоритма или метода, то нужно указать его достоинства, недостатки, область применения.