

Методы машинного обучения – список вопросов для экзаменационных билетов.

1. Термин Машинное обучение. Задачи, решаемые с помощью машинного обучения.
2. Обучение по прецедентам (индуктивное обучение). Формальная постановка задачи.
3. Обучение по прецедентам (индуктивное обучение). Общий порядок действий.
4. Типы машинного обучения.
5. Свойства алгоритма обучения. Обобщающая способность (generalization ability). Проблема недообучения и переобучения.
6. Свойства алгоритма обучения. Обобщающая способность (generalization ability). Эмпирические оценки обобщающей способности
7. Исходные данные. Признаковое описание объекта. Типы признаков (Атрибутов). Предобработка.
8. Исходные данные. Возможные проблемы исходных данных. Предобработка.
9. Случайная природа входных данных. Вероятностные характеристики выборки. Основные распределения значений признаков.
10. Статистические гипотезы и их проверка. Статистическая значимость.
11. Статистические гипотезы и их проверка. Ошибки первого и второго рода.
12. Методика проверки статистических гипотез. P-value. Доверительный интервал (Confidence interval).
13. Проверка гипотезы о математическом ожидании. Критерий согласия Пирсона Хи-квадрат (Chi-square test).
14. Корреляционный анализ. Отбор признаков. Ограничения.
15. Регрессионный анализ (regression analysis). Цели и задачи. Математическое определение регрессии
16. Регрессионный анализ. Линейная и Нелинейная регрессия.
17. Регрессионный анализ. Оценки качества регрессии.
18. Обучение с учителем (supervised learning). Постановка задачи классификации, методы классификации, основные метрики качества классификации.
19. Классификация - Байесовский подход. Цепочка расчетов и формула Байеса
20. Модель наивного байесовского классификатора
21. Логистическая регрессия. Подбор параметров: метод Ньютона.
22. Логистическая регрессия. Метод градиентного спуска.
23. Логистическая регрессия и представление в виде однослойной нейронной сети
24. Линейный дискриминантный анализ и линейный дискриминант Фишера.
25. Нейросетевой подход к решению задач классификации.
26. Принципы построения искусственных нейронных сетей. Нейрон, функция активации, обучение.
27. Многослойный персептрон. Выбор числа слоев, числа нейронов в скрытом слое. Методы обучения.
28. Нейронные сети основанные на базисных радиальных функциях.
29. Деревья решений. Сферы применения, решаемые задачи. Алгоритмы обучения.
30. Деревья решений. Проблемы основных этапов построения.
31. Деревья решений. Алгоритмы построения. Выбор атрибута разбиения.
32. Деревья решений. Выбор атрибута разбиения. Split-Info и Gain-Ratio.
33. Ансамбли классификаторов. Типы Ансамблей. Теорема Кондорсе о присяжных.
34. Ансамбли классификаторов. Бэггинг. Random Forest.

35. Ансамбли классификаторов. Бустинг (boosting). Градиентный бустинг над решающими деревьями.
36. Обучение с подкреплением (reinforcement learning). Постановка задачи. Exploration vs Exploitation.
37. Марковские процессы. Постановка задачи обучения с подкреплением. Подход к решению.
38. Задача о многоруком бандите (The multi-armed bandit problem). Стратегии.
39. Марковские процессы. Агент, среда, обратная связь, состояние, функции ценности состояния (Value function), качества действия (Q-function);
40. Эволюционный алгоритм (Evolutionary algorithm). Основные этапы. Преимущества и недостатки.
41. Генетический алгоритм. Основные этапы. Целевая функция приспособленности (fitness function)
42. Алгоритм роя частиц. Инициализация, выполнение.
43. Эволюционные алгоритмы. Алгоритм пчелиной колонии.
44. Эволюционные алгоритмы. Муравьиный алгоритм.
45. Кластеризация. Цели, формальная постановка задачи, принципиальная неоднозначность.
46. Кластеризация. Основные этапы. Меры расстояний.
47. Кластеризация. Меры качества кластеризации.
48. Кластеризация. Типы алгоритмов кластеризации.
49. Кластеризация. Метод k-средних (k-means). Достоинства и недостатки.
50. Кластеризация. Алгоритмы семейства FOREL. Достоинства и недостатки.
51. Кластеризация. Алгоритм DBSCAN. Достоинства и недостатки.
52. Самоорганизующиеся карты Кохонена. Решаемые задачи. Обучение сети.
53. Рекуррентные нейронные сети. Принципы функционирования на примере нейронная сеть Хопфилда.
54. Рекуррентные нейронные сети. Двухнаправленная ассоциативная память.
55. Рекуррентные нейронные сети. Проблема долговременных зависимостей.
56. Рекуррентные нейронные сети. Долгая краткосрочная память - Long short-term memory (LSTM)
57. Сверточные нейронные сети. Принципы функционирования применительно к обработке изображений.
58. Ключевые этапы реализации глубоких сетей.
59. Подходы к извлечению правил из обученных нейронных сетей в задачах классификации.
60. Визуализация данных и результатов работы построенной модели.