

1. К какому способу машинного обучения относится линейная регрессия?
 - Обучение без учителя
 - Обучение с учителем
 - Обучение с подкреплением
 - Обучение без подкрепления
2. Что обычно используется в качестве метрики производительности для линейной регрессии?
 - Средняя абсолютная ошибка
 - Логистическая функция потерь
 - Среднеквадратичная ошибка
 - F1-мера
3. В случае одного регрессора сколько параметров необходимо определить для решения задачи линейной регрессии?
 - 1
 - 2
 - 3
 - 4
4. При градиентном спуске как на каждом шаге изменяются искомые параметры?
 - В сторону увеличения первой производной
 - В сторону уменьшения первой производной
 - В сторону увеличения второй производной
 - В сторону уменьшения второй производной
5. В чем особенность стохастического градиентного спуска?
 - Значения частных производных вычисляются по всему набору данных
 - Значения частных производных заменяются на случайные числа
 - Значения частных производных вычисляются по небольшой группе случайно выбранных элементов данных
 - Значения частных производных вычисляются по одному, случайно выбранному элементу данных
6. В множественной линейной регрессии
 - Один регрессор и одна зависимая переменная
 - Несколько регрессоров и одна зависимая переменная
 - Один регрессор и несколько зависимых переменных
 - Несколько регрессоров и несколько зависимых переменных
7. Какой тип задач машинного обучения решает логистическая регрессия?
 - Регрессия
 - Классификация
 - Кластеризация
 - Всё вышеперечисленное
8. К чему приводит регуляризация?
 - Смещение и разброс уменьшаются
 - Смещение уменьшается, разброс увеличивается
 - Смещение увеличивается, разброс уменьшается
 - Смещение и разброс увеличиваются

9. Что из нижеперечисленного не является типом регуляризации?
- Lasso
 - Logloss
 - Elastic Net
 - Ridge
10. Условная вероятность $P(A|B)$ – это
- Вероятность наступления A при условии, что B уже произошло.
 - Вероятность наступления B при условии, что A уже произошло.
 - Вероятность наступления A при условии, что B не произошло.
 - Вероятность наступления B при условии, что A не произошло.
11. Если события A и B независимы, то чему равна условная вероятность $P(A|B)$?
- $P(A)P(B)$
 - $P(A)$
 - $P(B)$
 - $P(A) / P(B)$
12. Для определения $P(A|B)$ с помощью теоремы Байеса нужно знать
- $P(A)$
 - $P(B)$
 - $P(B|A)$
 - Всё вышеперечисленное.
13. Теорема Байеса позволяет
- Уточнить вероятность гипотезы до проведения эксперимента.
 - Вычислить вероятность успешного проведения эксперимента.
 - Уточнить вероятность гипотезы после проведения эксперимента.
 - Вычислить вероятность неуспешного проведения эксперимента.
14. В чём заключается «наивность» Байесовского классификатора?
- Значения параметров в элементе не могут повторяться.
 - Параметры элементов в наборе данных считаются независимыми.
 - Количество классов равно двум.
 - Порядок следования параметров в элемент определяет их важность.
15. С каким минусом Байесовского классификатора борется сглаживание Лапласа?
- Низкая скорость обучения.
 - Параметры должны быть независимыми.
 - Значения параметров должны быть категориальными.
 - Невозможно работать со значениями параметров, которых не было в обучающей выборке.
16. Что из перечисленного не является плюсом Байесовского классификатора?
- Хорошо работает в случае малого количества элементов.
 - Хорошо работает в случае большого количества параметров.
 - Возможность работы с числовыми значениями параметров.
 - Быстрые обучения и работа.
17. Какие типы задач решают наивный Байесовский классификатор и логистическая регрессия?

- Наивный Байесовский классификатор – задачу классификации, логистическая регрессия – задачу регрессии.
 - Наивный Байесовский классификатор – задачу регрессии, логистическая регрессия – задачу классификации.
 - Оба алгоритма решают задачу регрессии.
 - Оба алгоритма решают задачу классификации.
18. Какие из алгоритмов относятся к генеративному и дискриминативному подходу?
- Наивный Байесовский классификатор и логистическая регрессия оба относятся к генеративному подходу.
 - Наивный Байесовский классификатор и логистическая регрессия оба относятся к дискриминативному подходу.
 - Наивный Байесовский классификатор – генеративный подход, логистическая регрессия – дискриминативный подход.
 - Наивный Байесовский классификатор – дискриминативный подход, логистическая регрессия – генеративный подход.
19. Какие типы задач машинного обучения решаются методом k -ближайших соседей?
- Кластеризация и классификация.
 - Классификация и регрессия.
 - Регрессия и кластеризация.
 - Только кластеризация.
20. Что означает k в методе k -ближайших соседей?
- Количество используемых параметров из набора данных.
 - Размер обучающей выборки.
 - Количество элементов, необходимых для принятия решения.
 - Размерность пространства, в котором производятся вычисления.
21. Что утверждает гипотеза компактности?
- Более близкие объекты чаще относятся к одному и тому же классу, чем к разным.
 - Более близкие объекты чаще относятся к разным классам, чем к одному и тому же.
 - Количество элементов в каждом классе примерно одинаково.
 - Количество классов не превышает заранее выбранное значение.
22. Что из перечисленного не является проблемой метода k -ближайших соседей?
- Алгоритм медленно работает с большими наборами данных.
 - Алгоритм плохо работает с большим числом параметров.
 - Алгоритм плохо интерпретируем.
 - Необходимо заранее выбрать функцию расстояния.
23. Для решения какой проблемы метода k -ближайших соседей применяется кросс-валидация?
- Необходимо заранее выбрать значение k .
 - Алгоритм медленно работает из-за периферийных элементов.
 - Алгоритм плохо работает с несбалансированными данными.
 - Алгоритм совершает ошибки из-за выбросов.
24. Что из перечисленного не относится к преимуществам метода k -ближайших соседей?

- Нет требований к значениям параметров.
- Обучающая выборка может легко дополняться.
- Устойчивость к несбалансированным данным.
- Нет явного процесса обучения.

25. Что происходит с числом кластеров в процессе работы метода k -средних?

- Число кластеров увеличивается от 1 до k .
- Число кластеров уменьшается от k до 1.
- Число кластеров всё время равно k .
- Число кластеров никак не зависит от k .

26. Что из перечисленного не является недостатком метода k -средних?

- Алгоритм плохо распараллеливается.
- Алгоритм останавливается в первом достигнутом локальном минимуме.
- Нужно заранее знать число кластеров k .
- Результат зависит от изначального выбора центроидов.

27. Что применяется для определения наилучшего значения k в методе k -средних?

- Коэффициент силуэта.
- Перебалансировка данных.
- Кросс-валидация.
- Гипотеза компактности.