



Машинное обучение и рекомендательные системы.ТИ

- 1) Какая из перечисленных задач относится к задачам машинного обучения с учителем?
- 2) Какой алгоритм относится к метрическим классификаторам?
- 3) Какая метрика расстояния вычисляет сумму абсолютных разностей координат между двумя точками?
- 4) Какой алгоритм кластеризации требует предварительного задания количества кластеров?
- 5) Какой из перечисленных алгоритмов кластеризации лучше всего подходит для обнаружения кластеров произвольной формы?
- 6) Какой критерий используется для выбора лучшего признака при построении деревьев решений в алгоритме ID3?
- 7) Какой критерий используется для выбора лучшего признака при построении деревьев решений в алгоритме CART?
- 8) Что такое "Kernel Trick" в контексте метода опорных векторов (SVM)?
- 9) Что такое переобучение (overtraining)?
- 10) Что показывает ROC-кривая?
- 11) Что такое AUC (Area Under the Curve)?
- 12) Какое из утверждений верно относительно лесов решающих деревьев (Random Forest)?
- 13) Что такое опорные векторы в методе опорных векторов (SVM)?
- 14) Какая из перечисленных метрик наиболее чувствительна к несбалансированности классов?
- 15) Какой из алгоритмов кластеризации на основе плотности требует указания минимального количества точек в окрестности для определения "основной" точки?
- 16) Какой из перечисленных алгоритмов чаще всего используется для уменьшения размерности данных?



- 17) Что такое “энтропия” в контексте деревьев решений?
- 18) Какой из следующих алгоритмов классификации наиболее подвержен влиянию выбросов?
- 19) Что характеризует “зазор” (margin) в методе опорных векторов (SVM)?
- 20) Какая из следующих характеристик является преимуществом использования ансамблевых методов (например, Random Forest)?
- 21) Какая функция активации чаще всего используется в выходном слое логистической регрессии?
- 22) Какой метод оптимизации используется для обучения логистической регрессии и нейронных сетей?
- 23) Что такое “обратное распространение ошибки” (backpropagation)?
- 24) Какая операция используется в сверточных нейронных сетях (CNN) для обнаружения локальных признаков в изображениях?
- 25) Какова цель операции пулинга в CNN?
- 26) Какой тип регрессии используется, когда зависимость между переменными не является линейной?
- 27) Что такое смещение (bias) в контексте машинного обучения?
- 28) Что такое дисперсия (variance) в контексте машинного обучения?
- 29) Какова основная цель регуляризации в машинном обучении?
- 30) Какой метод объединяет прогнозы нескольких базовых моделей для получения более точного и устойчивого результата?
- 31) Какой из следующих ансамблевых методов обучает базовые модели последовательно, где каждая последующая модель фокусируется на исправлении ошибок предыдущих?
- 32) Какой алгоритм стохастического поиска имитирует процесс отжига металла?
- 33) Какой алгоритм стохастического поиска использует принципы естественного отбора?
- 34) Что такое “функция активации” в нейронной сети?





- 35) Что такое “скорость обучения” (learning rate) в градиентном спуске?
- 36) Какая из следующих функций активации помогает решить проблему “исчезающего градиента”?
- 37) Какой метод помогает уменьшить дисперсию модели, обучая несколько моделей на разных подвыборках данных?
- 38) В каком из следующих алгоритмов каждая модель в ансамбле получает вес в зависимости от её производительности?
- 39) Какой алгоритм стохастического поиска может принимать ухудшающие решения с определенной вероятностью, чтобы избежать застревания в локальных минимумах?
- 40) Какой из методов не относится к коллаборативной фильтрации?
- 41) В каком типе коллаборативной фильтрации рекомендации формируются на основе предпочтений похожих пользователей?
- 42) Какая мера обычно используется для вычисления сходства между пользователями в user-based CF?
- 43) Какой тип матрицы часто используется для представления взаимодействий пользователей с объектами в рекомендательных системах?
- 44) В каком алгоритме используется FP-дерево для поиска частых наборов объектов?
- 45) Какая мера определяет, как часто набор объектов X и Y встречается вместе в транзакциях?
- 46) Какая мера определяет, насколько вероятно, что правило X \rightarrow Y является верным?
- 47) Какая мера определяет, насколько связь между X и Y сильнее, чем если бы они были независимы?
- 48) Какой из алгоритмов предназначен для поиска ассоциативных правил?
- 49) Какое из понятий означает минимальный набор правил, из которого можно вывести все остальные правила с использованием логических операций?
- 50) Что такое “холодный старт” в контексте рекомендательных систем?
- 51) Какая мера характеризует, насколько правило сохраняется при небольших изменениях в данных?





- 52) Какой из подходов к коллаборативной фильтрации лучше подходит для ситуации, когда количество объектов значительно меньше, чем количество пользователей?
- 53) Какое из следующих утверждений верно относительно FP-дерева?
- 54) Что такое “замкнутое множество” (closed itemset)?
- 55) В какой из рекомендательных систем используются знания о характеристиках самих объектов (например, жанр фильма, описание товара)?
- 56) Что такое “лифт” (Lift) равный 1?
- 57) Какая из мер интересности наиболее чувствительна к большим наборам данных?
- 58) Что из перечисленного является преимуществом коллаборативной фильтрации по сравнению с контентно-ориентированными системами?
- 59) Какой метод используется для выявления латентных факторов путем разложения матрицы оценок?
- 60) Какой алгоритм матричной факторизации учитывает как явные оценки, так и неявные взаимодействия?
- 61) Какой алгоритм матричной факторизации учитывает временную динамику предпочтений пользователей и популярности объектов?
- 62) Какой метод оптимизации используется для минимизации ошибки между предсказанными и фактическими оценками в матричной факторизации и обновляет параметры на каждой итерации на основе случайно выбранного образца данных?
- 63) Какой метод оптимизации для матричной факторизации итеративно обновляет латентные факторы пользователей и объектов, фиксируя одну группу факторов и оптимизируя другую?
- 64) Какой метод обобщает линейные модели и модели матричной факторизации и позволяет учитывать взаимодействия между признаками высокой размерности?
- 65) Для чего может использоваться поиск минимального разреза в рекомендательных системах?
- 66) Какой подход к рекламе показывает рекламные объявления пользователям на основе тематики просматриваемой ими веб-страницы?





- 67) Какой метод используется для сравнения двух версий веб-страницы или приложения с целью определения, какая из них лучше работает?
- 68) Что означает встраивание дополнительной информации в модели рекомендательных систем?
- 69) Какой метод объединяет прогнозы нескольких рекомендательных алгоритмов для повышения точности и устойчивости системы?
- 70) В чем преимущество использования глубокого обучения для рекомендательных систем?
- 71) Какое утверждение верно относительно PureSVD?
- 72) В каком из перечисленных случаев особенно полезны факторизационные машины?
- 73) Какой метод позволяет учитывать контекст пользователя (время, местоположение) при формировании рекомендаций?
- 74) Что из перечисленного не является методом оценки качества рекомендательных систем?
- 75) Что является преимуществом использования ансамблей рекомендательных алгоритмов?
- 76) Что подразумевается под латентными факторами в моделях матричной факторизации?

