

Список вопросов к экзамену

1 Структуры данных

1. О-нотация. Приведите примеры алгоритмов над структурами данных с константным и квадратичным по числу элементов временем работы.
2. О-нотация. Приведите примеры алгоритмов над структурами данных с линейным и логарифмическим по числу элементов временем работы.
3. О-нотация. Приведите примеры, в которых выбор асимптотически более медленного с точки зрения О-нотации алгоритма предпочтителен.
4. Сортировка. Приведите примеры методов сортировки.
5. Деревья. Определения: двоичное дерево, глубина дерева, сбалансированное дерево. Какова глубина сбалансированного двоичного дерева?
6. Деревья. Определения: дерево поиска, двоичное дерево поиска, минимальное и максимальное время поиска значения.
7. Сбалансированное двоичное дерево поиска, для чего используется, время поиска значения.
8. Сбалансированное двоичное дерево поиска, пример построения из упорядоченного массива.
9. k-мерное двоичное дерево поиска. Время поиска значения, примеры использования.
10. Хэш-функции. Какими свойствами должна обладать хорошая хэш-функция. Область применения.
11. Хэш-таблица. Алгоритм построения. Скорость доступа к элементу. Скорость добавления нового элемента.
12. Хэш-таблица. Требование к памяти. Парадокс дней рождения и коллизии.
13. Хэш-таблица. Разрешение коллизий цепочкой. Влияние коллизий на скорость работы.
14. Структуры данных: массив, дерево поиска. Примеры использования.
15. Структуры данных: дерево, хеш-таблица. Примеры использования.

2 Оптимизация

1. Задачи оптимизации. Классификация. Примеры.
2. Методы оптимизации. Классификация. Примеры.
3. Сформулируйте необходимые условия локального минимума $f(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^N$.
4. Сформулируйте достаточные условия локального минимума $f(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^N$.
5. Оптимизация с ограничениями. Примеры барьерных методов.
6. Оптимизация с ограничениями. Примеры проекционных методов.

3 Статистика

1. Формулировка задачи линейной регрессии. Идея и постановка задачи определения параметров линии регрессии методом наименьших квадратов.
2. Формулировка задачи линейной регрессии. Идея и постановка задачи определения параметров линии регрессии методом максимального правдоподобия.
3. Сравните метод максимального правдоподобия и метод наименьших квадратов. В каких случаях они применяются, как выбрать тот или иной метод.
4. Метод максимального правдоподобия. Покажите как методом максимального правдоподобия получить параметр пуассоновского распределения по набору измерений.
5. Метод максимального правдоподобия.
6. Проверка статистических гипотез. Нулевая и альтернативная гипотеза. Ошибки первого и второго рода.
7. Проверка статистических гипотез. Идея t -критерия Стьюдента.
8. Проверка статистических гипотез. Идея χ^2 -критерия Пирсона.
9. Проверка статистических гипотез. Идея критерия Колмогорова.
10. Стохастические процессы. Как автокорреляционная функция позволяет их изучать.
11. Стохастические линейные процессы. Идея аппроксимации моделью авторегрессии.
12. Стохастические линейные процессы. Идея аппроксимации моделью скользящего среднего.

4 Машинное обучение

1. Основная идея подхода к решению задач методами машинного обучения. Фундаментальное ограничение машинного обучения.
2. Классификация методов машинного обучения. Приведите примеры задач каждого класса.
3. Задача классификации с учителем. Примеры.
4. Функция потерь в задачах машинного обучения.
5. Понятие переобучения. Тестовая выборка.
6. Идея метода k ближайших соседей (kNN).
7. Гиперпараметры и валидационная выборка.
8. Задача регрессии с учителем. Примеры. Популярные алгоритмы решения.
9. Методы машинного обучения без учителя. Приведите примеры задач.
10. Идея метода нейронных сетей.
11. Идея методов random forest и gradient boosting.
12. Приведите примеры физических задач для которых подходит и не подходит машинное обучение.