

Список вопросов к экзамену

1 Структуры данных

1. О-нотация. Приведите примеры алгоритмов над структурами данных с константным, логарифмическим и квадратичным по числу элементов временем работы.
2. О-нотация. Выбор алгоритма для конкретной задачи. Приведите примеры, в которых выбор более медленного с точки зрения О-нотации алгоритма предпочтителен.
3. Сортировка. Приведите примеры методов сортировки, один из них разберите подробно и представьте вывод для времени его работы.
4. Деревья. Основные определения: дерево, бинарное дерево, бинарное дерево поиска, глубина дерева, сбалансированное дерево. Какова глубина сбалансированного дерева?
5. Сбалансированное дерево поиска. Для чего используется, время поиска значения. Как получить его из упорядоченного массива.
6. Хэш-функции. Какими свойствами должна обладать хорошая хэш-функция. Область применения.
7. Хэш-таблица. Алгоритм построения. Скорость доступа к элементу. Скорость добавления нового элемента.
8. Хэш-таблица. Требование к памяти. Парадокс дней рождений. Разрешение коллизий цепочкой. Влияние коллизий на скорость работы
9. Структуры данных: массив, дерево, хеш-таблица. Примеры использования.

2 Оптимизация

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

3 Статистика

1. Формулировка задачи линейной регрессии. Сформулируйте задачу двух-параметрической линейной регрессии ($\hat{y} = a + b x$) и решите её методом наименьших квадратов.
2. Формулировка задачи линейной регрессии. Сформулируйте задачу двух-параметрической линейной регрессии ($\hat{y} = a + b x$) и решите её методом максимального правдоподобия, полагая распределение нормальным.
3. Сформулируйте и решите методом наименьших квадратов задачу однопараметрической линейной регрессии ($\hat{y} = b x$). Найдите математическое ожидание и дисперсию параметра модели.
4. Сравните метод максимального правдоподобия и метод наименьших квадратов. В каких случаях они применяются, как выбрать тот или иной метод.
5. Метод максимального правдоподобия как метод сравнения статистических гипотез. Покажите как методом максимального правдоподобия получить параметр пуассоновского распределения по набору измерений.
6. Проверка статистических гипотез. Идея t -критерия Стьюдента и χ^2 -критерия Пирсона.
7. Проверка статистических гипотез. Идея критерия Колмогорова.

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. Тестовая подвыборка и функция потерь.
4. Идея метода k ближайших соседей (kNN).
5. Гиперпараметры и валидационная выборка.
6. Задача регрессии с учителем. Примеры. Популярные алгоритмы решения.
7. Методы машинного обучения без учителя. Приведите примеры задач.
8. Идея метода нейронных сетей.

9. Идея методов random forest и gradient boosting.
10. Приведите примеры физических задач для которых подходит и не подходит машинное обучение.