

**Программа государственной итоговой аттестации по
направлению подготовки “Информатика и вычислительная
техника”, 2023/24 учебный год**

*Примечание: используется сквозная нумерация разделов после программы от
кафедры ДМ.*

4. Алгоритмы и структуры данных

| | |
|-------|--|
| 4.1. | Динамический массив. Амортизационный анализ. Учетная оценка времени добавления элемента в динамический массив (с удвоением заполненного буфера). |
| 4.2. | Связные списки. Стек, очередь, дек и их реализации. |
| 4.3. | Быстрая сортировка (QuickSort). Поиск порядковой статистики методом “Разделяй и властвуй” (QuickSelect). |
| 4.4. | Сортировка слиянием (MergeSort). Поразрядные сортировки. |
| 4.5. | Двоичная куча и сортировка кучей (HeapSort). Слияние k отсортированных массивов с помощью кучи. |
| 4.6. | Хеш-таблица, полиномиальная хеш-функция. |
| 4.7. | Динамическое программирование: общая идея, линейная динамика, матричная, динамика на отрезках. |
| 4.8. | RMQ. Sparse table. Дерево отрезков. |
| 4.9. | LCA: сведение к RMQ и метод двоичного подъёма. |
| 4.10. | Двоичное дерево поиска. Обходы в глубину и в ширину. Поиск ключа, наивные вставка и удаление ключа. АВЛ-дерево. |
| 4.11. | Декартово дерево. Декартово дерево по неявному ключу. |
| 4.12. | Минимальное остовное дерево: алгоритмы Прима и Крускала. |
| 4.13. | Обход графа в глубину, ширину. |
| 4.14. | Поиск кратчайших путей в графе: алгоритмы Дейкстры, Форда-Беллмана, Флойда-Уоршелла. |
| 4.15. | Поиск сильно-связных компонент в графе. |
| 4.16. | Мосты и точки сочленения в графе. |
| 4.17. | Нахождение подстроки в строке: префикс-функция, алгоритм Кнута-Морриса-Пратта. |
| 4.18. | Стандартные контейнеры: vector, deque, queue, priority_queue, set, map, итераторы, компараторы. |
| 4.19. | Бор. Алгоритм Ахо-Корасик. |
| 4.20. | Вычисление выпуклой оболочки множества точек в 2D. |
| 4.21. | Планиметрия: примитивы точки, прямой, окружности. Построение прямой. Пересечения (прямых, окружностей). |
| 4.22. | Проверка принадлежности точки многоугольнику. |

5. Машинное обучение

| | |
|-------|---|
| 5.1. | Постановка задачи обучения с учителем (supervised learning), постановка задачи обучения без учителя (unsupervised learning). Постановка задачи классификации и метрики качества классификации. Наивный Байесовский классификатор. |
| 5.2. | Постановка задачи регрессии и метрики качества регрессии. Линейная регрессия. Теорема Гаусса-Маркова (формулировка). Проблема мультиколлинеарных признаков. L1 и L2 регуляризация, их влияние на веса признаков. |
| 5.3. | Проблема несбалансированных классов. Работа с категориальными признаками и пропущенными значениями. Mean encoding. Примеры простых алгоритмов, решающих стандартные задачи: kNN, наивный байесовский классификатор. |
| 5.4. | Логистическая регрессия. Понятия отступа (Margin). Эквивалентность решений полученных методом максимального правдоподобия и минимизации логистической функции потерь. Логистическая функция потерь, кросс-энтропия. |
| 5.5. | Процедура построения решающего дерева, критерии информативности: энтропийный, Джини. Бустинг (принцип построения ансамбля). |
| 5.6. | Bias-Variance decomposition (декларативно). Процедура bootstrap, алгоритм bagging. Random Forest. |
| 5.7. | Градиентный бустинг. Принцип построения. Какие алгоритмы могут использоваться в качестве базовых алгоритмов. |
| 5.8. | Задача снижения размерности: алгоритм PCA. Связь PCA и SVD. Может ли PCA давать не единственное решение? |
| 5.9. | Проблема переобучения. Пример, причины возникновения (идейно). Кросс-валидация. Мотивация использования train, val и test выборок. Понятие параметров и гиперпараметров моделей. Процедура подбора гиперпараметров. Понятие регуляризации (в общем случае). Способы регуляризации различных моделей (линейные модели, деревья, ансамбли, нейронные сети). |
| 5.10. | Метод градиентного спуска. Метод обратного распространения ошибки (backpropagation). Функции активации (Sigmoid, tanh, ReLU), их свойства и проблемы. Функции потерь в задаче многоклассовой классификации и регрессии. Методы регуляризации нейронных сетей: Dropout, Batch normalization, data augmentation. |
| 5.11. | Методы регуляризации нейронных сетей: Dropout, Batch normalization, data augmentation. Слабые стороны стохастического градиентного спуска. Способы доработки: Momentum, Nesterov momentum, RMSprop, Adam. Минусы данных подходов. |
| 5.12. | Рекуррентные нейронные сети (RNN) для упорядоченных данных. Основные принципы работы. Проблема затухающего градиента и ее возможные решения. Рекуррентные блоки: наивный (Vanilla RNN), LSTM, GRU, мотивация их использования. Функции активации в рекуррентных блоках. |
| 5.13. | Методы работы с изображениями. Почему линейные слои не получили широкого применения в задаче компьютерного зрения? Сверточные слои в нейронной сети. Мотивация их использования в задачах анализа изображений и сигналов. |
| 5.14. | Одномерные и двумерные свертки (Conv1d и Conv2d). Max & average pooling. Генеративные нейронные сети: VAE, GAN. Архитектура и сфера применения. |

6. Формальные языки и трансляции [опционально для ИВТ]

| | |
|------|---|
| 6.1. | Недетерминированные конечные автоматы (НКА). Различные варианты определений. |
| 6.2. | Детерминированные конечные автоматы (ДКА). Эквивалентность ДКА и НКА. |
| 6.3. | Свойства класса автоматных языков. Замкнутость относительно булевых операций. Минимальный ДКА |
| 6.4. | Регулярные выражения. Теорема Клини о совпадении классов регулярных и автоматных языков. Регулярный автомат, выводимость в регулярном автомате. Алгоритм построения регулярного выражения по регулярному автомату. |
| 6.5. | Теорема Майхилла-Нероуда. Лемма о разрастании для автоматных языков. Примеры неавтоматных языков. |

7. Операционные системы, параллельные и распределенные вычисления

| | |
|------|--|
| 7.1. | Операционные системы и их компоненты. Ядро операционных систем. Системные вызовы и их отличия от обычных библиотечных функций. Способы реализации системных вызовов (прерывания, sysenter, syscall). |
| 7.2. | Целочисленная арифметика в представлении компьютера. Знаковые и беззнаковые значения, способы представления отрицательных значений. Целочисленное переполнение и его контроль. Длинная целочисленная арифметика. |
| 7.3. | Вещественная арифметика. Представления с фиксированной и плавающей точкой. Стандарт IEEE754. Специальные вещественные значения, определенные стандартом IEEE754 и операции над ними. |
| 7.4. | Процессы и потоки. Сходства и различия между ними. Реализация многозадачности и алгоритмы планирования задач в операционных системах. |
| 7.5. | Проблема многопоточной синхронизации. Атомарные переменные и объекты блокировки. Свободные от блокировок (lock-free) структуры данных и их реализация. |
| 7.6. | Интерфейс передачи сообщений (MPI) Существующие реализации, задачи MPI как среды программирования. Жизненный цикл MPI программы. Создание и завершение процессов. Организация потока ввода-вывода. |
| 7.7. | Понятие ускорения и масштабируемости параллельных программ. Вертикальная и горизонтальная масштабируемость. Закон Амдала. Оценка эффективности параллельных программ. Ярусно-параллельная форма программы. |
| 7.8. | Распределенные файловые системы. Роли элементов системы, обеспечение отказоустойчивости. Алгоритмы чтения и записи в распределенных файловых системах. Репликация данных. |
| 7.9. | Модель вычислений MapReduce. Пары ключ-значение в реализациях MapReduce. |

| | |
|-------|--|
| | Основные стадии вычислений и дополнительные элементы модели. |
| 7.10. | Соединение данных (операция Join) в модели MapReduce. Модель вычислений и оптимизации данных. |
| 7.11. | Итеративные вычисления на больших объемах данных. Модель ленивых вычислений и структура хранения данных в реализации Spark (RDD). Кэширование результатов вычислений и итеративные вычисления. |
| 7.12. | Распределенные диспетчеры сообщений. Репликация и реализация отказоустойчивости. Семантики доставки сообщений. |
| 7.13. | Степени изоляции транзакций. Принципы атомарности, согласованности, изолированности и устойчивости (правила ACID). Применение (commit) и отмена транзакций. |
| 7.14. | Теорема Фишера-Линч-Патерсона (FLP-теорема) без доказательства, CAP-теорема и их применение. Распределенные системы хранения конфигураций. |
| 7.15. | Иерархическая и сетевая модели данных: свойства, сходства и различия, достоинства и недостатки. |
| 7.16. | Реляционная модель данных. Основные понятия реляционной модели данных. Реляционная алгебра. Операции реляционной алгебры. |
| 7.17. | Нормальные формы. Перечислить все. Определения нормальных форм (до нормальной формы Бойса-Кодда включительно). Потенциальные, первичные, внешние ключи. Ссылочная целостность базы данных. |