

**Комплект оценочных материалов по дисциплине  
«Системы обработки данных»**

**Задания закрытого типа**

**Задания закрытого типа на выбор правильного ответа**

1. Выберите один правильный ответ

Что такое система обработки данных?

- А) Программное обеспечение для создания графиков
- Б) Метод шифрования информации
- В) Устройство для печати документов
- Г) Комплекс аппаратных и программных средств для сбора, хранения и анализа данных

Правильный ответ: Г

Компетенции (индикаторы): ПК-1

2. Выберите один правильный ответ

Какой язык программирования чаще всего используется для работы с большими данными?

- А) Java
- Б) Python
- В) SQL
- Г) HTML

Правильный ответ: Б

Компетенции (индикаторы): ПК-1

3. Выберите один правильный ответ

Что такое NoSQL база данных?

- А) База данных, которая не использует SQL для запросов
- Б) База данных, которая работает только с текстовыми данными
- В) База данных, которая не поддерживает транзакции
- Г) База данных, которая работает только в облаке

Правильный ответ: А

Компетенции (индикаторы): ПК-1

4. Выберите один правильный ответ

Какой из перечисленных форматов данных является структурированным?

- А) JSON
- Б) XML
- В) CSV
- Г) Все перечисленные

Правильный ответ: Г

Компетенции (индикаторы): ПК-1

5. Выберите один правильный ответ

Что такое ETL?

- А) Метод шифрования данных
- Б) Процесс извлечения, преобразования и загрузки данных
- В) Язык программирования для анализа данных
- Г) Инструмент для визуализации данных

Правильный ответ: Б

Компетенции (индикаторы): ПК-1

6. Выберите один правильный ответ

Какая из перечисленных баз данных является реляционной?

- А) MongoDB
- Б) Cassandra
- В) PostgreSQL
- Г) Redis

Правильный ответ: В

Компетенции (индикаторы): ПК-1

7. Выберите один правильный ответ

Какой язык используется для написания запросов в реляционных базах данных?

- А) Python
- Б) SQL
- В) Java
- Г) HTML

Правильный ответ: Б

Компетенции (индикаторы): ПК-1

8. Выберите один правильный ответ

Какой инструмент используется для потоковой обработки данных?

- А) Apache Kafka
- Б) Tableau
- В) MySQL
- Г) Hadoop

Правильный ответ: а

Компетенции (индикаторы): ПК-1

### **Задания закрытого типа на установление соответствия**

1. Установите соответствие между типами данных и их примерами:

Тип данных

Пример

- |                             |               |
|-----------------------------|---------------|
| 1) Структурированные данные | A) Видеофайлы |
|-----------------------------|---------------|

- 2) Неструктурированные данные      Б) Таблицы в реляционной базе данных  
 3) Полуструктурированные данные    В) JSON-файлы  
 4) Двоичные данные                      Г) Изображения в формате PNG  
 Правильный ответ:

1	2	3	4
Б	А	В	Г

Компетенции (индикаторы): ПК-1

2. Установите соответствие между языками программирования и их использованием в обработке данных:

- | Язык программирования | Использование в обработке данных                              |
|-----------------------|---|
| 1) Python             | А) Анализ данных и статистика                                 |
| 2) SQL                | Б) Работа с базами данных и запросы                           |
| 3) R                  | В) Разработка приложений для обработки данных                 |
| 4) Java               | Г) Универсальный язык для анализа данных и машинного обучения |

Правильный ответ:

1	2	3	4
Г	Б	А	В

Компетенции (индикаторы): ПК-1

3. Установите соответствие между форматами данных и их характеристиками:

- | Формат данных | Характеристика  |
|---------------|---|
| 1) JSON       | А) Текстовый формат, разделенный запятыми                                     |
| 2) XML        | Б) Формат для хранения столбчатых данных, оптимизированный для аналитики      |
| 3) CSV        | В) Формат на основе тегов, используемый для структурированных данных          |
| 4) Parquet    | Г) Легковесный формат для обмена данными, основанный на синтаксисе JavaScript |

Правильный ответ:

1	2	3	4
Г	В	А	Б

## Компетенции (индикаторы): ПК-1

4. Установите соответствие между типами баз данных и их описаниями:

База данных	Описание
1) Реляционная база данных	А) Хранение данных в виде документов (например, JSON)
2) NoSQL база данных	Б) Использование таблиц и связей между ними
3) Графовая база данных	В) Оптимизирована для работы с данными, представленными в виде графов
4) Документоориентированная база данных	Г) Не использует реляционную модель, подходит для неструктурированных данных

Правильный ответ:

1	2	3	4
Б	Г	В	А

## Компетенции (индикаторы): ПК-1

5. Установите соответствие между инструментами и их назначением:

Инструмент	Назначение
1) Apache Hadoop	А) Визуализация данных
2) Apache Kafka	Б) Распределенная обработка больших данных
3) Tableau	В) Реляционная база данных
4) PostgreSQL	Г) Потоковая обработка данных

Правильный ответ:

1	2	3	4
Б	Г	А	В

## Компетенции (индикаторы): ПК-1

6 Установите соответствие между понятиями и их определениями:

Понятие	Определение
1) Big Data	А) Хранилище данных в сыром виде, без строгой структуры
2) Data Mining	Б) Технологии анализа больших объемов данных для поиска закономерностей
3) Data Warehouse	В) Централизованное хранилище данных для аналитики
4) Data Lake	Г) Обработка огромных объемов данных, которые невозможно обработать традиционными методами

Правильный ответ:

1	2	3	4
Г	Б	В	А

Компетенции (индикаторы): ПК-1

7. Установите соответствие между технологиями и их применением:

Технология	Применение
1) Apache Spark	А) Поиск и анализ текстовых данных в реальном времени
2) Apache Flink	Б) Обработка потоковых данных
3) Elasticsearch	В) Хранение документоориентированных данных
4) MongoDB	Г) Высокопроизводительная обработка больших данных

Правильный ответ:

1	2	3	4
Г	Б	А	В

Компетенции (индикаторы): ПК-1

8. Установите соответствие между этапами машинного обучения и их описанием:

Этап машинного обучения	Описание
1) Сбор данных	А) Очистка и преобразование данных
2) Предобработка данных	Б) Проверка точности модели на тестовых данных
3) Обучение модели	В) Настройка параметров модели на обучающих данных
4) Оценка модели	Г) Получение данных из различных источников

Правильный ответ:

1	2	3	4
Г	А	В	Б

Компетенции (индикаторы): ПК-1

**Задания закрытого типа на установление правильной последовательности**

1. Установите правильную последовательность этапов работы с базой данных:

- А) Подключение к базе данных
- Б) Выполнение SQL-запроса

- В) Закрытие соединения
  - Г) Получение результатов запроса
- Правильный ответ: А, Б, Г, В
- Компетенции (индикаторы): ПК-1

2. Установите правильную последовательность этапов анализа данных:

- А) Сбор данных
  - Б) Построение моделей
  - В) Очистка данных
  - Г) Визуализация данных
- Правильный ответ: А, В, Б, В
- Компетенции (индикаторы): ПК-1

3. Установите правильную последовательность этапов работы с большими данными:

- А) Хранение данных
  - Б) Сбор данных
  - В) Анализ данных
  - Г) Обработка данных
- Правильный ответ: Б, А, Г, В
- Компетенции (индикаторы): ПК-1

4. Установите правильную последовательность этапов работы с машинным обучением:

- А) Подготовка данных
  - Б) Обучение модели
  - В) Выбор алгоритма
  - Г) Оценка модели
- Правильный ответ: А, В, Б, Г
- Компетенции (индикаторы): ПК-1

5. Установите правильную последовательность этапов разработки Data Pipeline:

- А) а) Проектирование архитектуры
  - Б) Тестирование и отладка
  - В) Реализация ETL-процессов
  - Г) Мониторинг и оптимизация
- Правильный ответ: А, В, Б, Г
- Компетенции (индикаторы): ПК-1

6. Установите правильную последовательность этапов работы с Apache Kafka

- А) Создание топика (topic)
- Б) Настройка продюсера (producer)

- В) Настройка консьюмера (consumer)
  - Г) Отправка и получение сообщений
- Правильный ответ: А, Б, В, Г  
Компетенции (индикаторы): ПК-1

7. Установите правильную последовательность этапов работы с Hadoop:

- А) Загрузка данных в HDFS
  - Б) Запуск MapReduce-задачи
  - В) Получение результатов обработки
  - Г) Настройка кластера Hadoop
- Правильный ответ: Г, А, Б, В  
Компетенции (индикаторы): ПК-1

8. Установите правильную последовательность этапов работы с Data Warehouse:

- А) Проектирование схемы данных (например, звезда или снежинка)
  - Б) Загрузка данных из источников
  - В) Оптимизация запросов
  - Г) Создание отчетов и визуализация
- Правильный ответ: А, Б, В, Г  
Компетенции (индикаторы): ПК-1

## **Задания открытого типа**

### **Задания открытого типа на дополнение**

1. Напишите пропущенное слово (словосочетание).

Язык запросов, используемый для работы с реляционными базами данных, называется \_\_\_\_\_.

Правильный ответ: SQL.

Компетенции (индикаторы): ПК-1

2. Напишите пропущенное слово (словосочетание).

Технология распределенной обработки больших данных, разработанная Apache, называется \_\_\_\_\_.

Правильный ответ: Hadoop.

Компетенции (индикаторы): ПК-1

3. Напишите пропущенное слово (словосочетание).

Формат данных, который использует теги для структурирования информации, называется \_\_\_\_\_.

Правильный ответ: XML.

Компетенции (индикаторы): ПК-1

4. Напишите пропущенное слово (словосочетание).

Система управления базами данных, которая не использует реляционную модель, называется \_\_\_\_\_.

Правильный ответ: NoSQL.

Компетенции (индикаторы): ПК-1

5. Напишите пропущенное слово (словосочетание).

Архитектура, которая позволяет хранить данные в сыром виде без строгой структуры, называется \_\_\_\_\_.

Правильный ответ: Data Lake.

Компетенции (индикаторы): ПК-1

6. Напишите пропущенное слово (словосочетание).

Процесс извлечения, преобразования и загрузки данных называется \_\_\_\_\_.

Правильный ответ: ETL.

Компетенции (индикаторы): ПК-1

7. Напишите пропущенное слово (словосочетание).

Процесс поиска закономерностей в больших объемах данных называется \_\_\_\_\_.

Правильный ответ: Data Mining.

Компетенции (индикаторы): ПК-1

8. Напишите пропущенное слово (словосочетание).

Инструмент для визуализации данных, который позволяет создавать интерактивные дашборды, называется \_\_\_\_\_.

Правильный ответ: Tableau.

Компетенции (индикаторы): ПК-1

9. Напишите пропущенное слово (словосочетание).

Технология потоковой обработки данных, разработанная Apache, называется \_\_\_\_\_.

Правильный ответ: Apache Kafka.

Компетенции (индикаторы): ПК-1

10. Напишите пропущенное слово (словосочетание).

Компонент Hadoop, отвечающий за распределенное хранение данных, называется \_\_\_\_\_.

Правильный ответ: HDFS.

Компетенции (индикаторы): ПК-1

**Задания открытого типа с кратким свободным ответом**

1. Напишите пропущенное слово (словосочетание)

Какие основные этапы включает процесс ETL?

Правильный ответ: Извлечение, преобразование, загрузка/ Extract, Transform, Load.

Компетенции (индикаторы): ПК-1

2. Дайте ответ на вопрос.

Назовите пример реляционной базы данных.

Правильный ответ: PostgreSQL / MySQL/ Oracle

Компетенции (индикаторы): ПК-1

3. Дайте ответ на вопрос

Какие компоненты входят в экосистему Apache Hadoop?

Правильный ответ: HDFS, MapReduce, YARN, Hive, HBase.

Компетенции (индикаторы): ПК-1

4. Дайте ответ на вопрос

Формат данных который используются для хранения полустроктурированных данных?

Правильный ответ: JSON / XML / YAML.

Компетенции (индикаторы): ПК-1

5. Дайте ответ на вопрос

Инструмент для визуализации данных?

Правильный ответ: Tableau / Power BI / Matplotlib, Seaborn / D3.js

Компетенции (индикаторы): ПК-1

6. Дайте ответ на вопрос

Какие преимущества предоставляет использование Parquet для хранения данных?

Правильный ответ: Эффективное хранение столбчатых данных, оптимизация для аналитики, сжатие данных

Компетенции (индикаторы): ПК-1

7. Дайте ответ на вопрос.

Назовите пример NoSQL баз данных.

Правильный ответ: MongoDB / Cassandra / Redis

Компетенции (индикаторы): ПК-1

8. Дайте ответ на вопрос

Какие этапы включает процесс машинного обучения?

Правильный ответ: Сбор данных, предобработка, выбор модели, обучение, оценка, внедрение.

Компетенции (индикаторы): ПК-1

9. Дайте ответ на вопрос

Какие задачи решает Data Mining?

Правильный ответ: Классификация, кластеризация, прогнозирование, поиск закономерностей.

Компетенции (индикаторы): ПК-1

10. Дайте ответ на вопрос

Для чего он используется Data Warehouse

Правильный ответ: Аналитика, отчетность.

Компетенции (индикаторы): ПК-1

### **Задания открытого типа с развернутым ответом**

1. Упростить и разложить на множители полином:

$$P(x,y,z) = x^2y + xy^2 + x^2z + xz^2 + y^2z + yz^2 + 2xyz.$$

Привести расширенное решение.

Время выполнения – 60 мин.

Ожидаемый результат:

1. Группировка членов:

Попробуем сгруппировать члены полинома так, чтобы можно было вынести общие множители. Заметим, что каждый член содержит две переменные из трёх. Сгруппируем их следующим образом:

$$P(x,y,z) = (x^2y + xy^2) + (x^2z + xz^2) + (y^2z + yz^2) + 2xyz.$$

2. Вынесение общих множителей:

В каждой группе вынесем общие множители:

$$P(x,y,z) = xy(x+y) + xz(x+z) + yz(y+z) + 2xyz.$$

3. Дополнительное упрощение:

Теперь заметим, что  $2xyz$  можно распределить между группами.

Добавим  $xxyz$  к каждой группе:

$$P(x,y,z) = xy(x+y) + xyz + xz(x+z) + xyz + yz(y+z) + xyz.$$

4. Факторизация:

Теперь в каждой группе можно вынести общие множители:

$$P(x,y,z) = xy(x+y+z) + xz(x+y+z) + yz(x+y+z).$$

Теперь видно, что  $(x+y+z)(x+y+z)$  является общим множителем:

$$P(x,y,z) = (x+y+z)(xy + xz + yz).$$

5. Итоговый результат:

Полином  $P(x,y,z)$  разложен на множители:

$$P(x,y,z) = (x+y+z)(xy + xz + yz)$$

Проверка

Чтобы убедиться в правильности результата, раскроем скобки:

$$(x+y+z)(xy + xz + yz) = x^2y + x^2z + xy^2 + xz^2 + y^2z + yz^2 + 2xyz.$$

Это совпадает с исходным полиномом, значит, преобразование выполнено верно.

Правильный ответ:  $P(x,y,z) = (x+y+z)(xy+xz+yz)$ .

Компетенции (индикаторы): ПК-1

2. Каков будет результат применения операций PUSH(S,4), PUSH(S,1), PUSH(S,3), POP(S), PUSH(S,6), PUSH(S,4), PUSH(S,1), PUSH(S,3), POP(S), POP(S) к пустому стеку, хранящемуся в массиве S[1..6]? К тому же стеку, уже хранящему 2 элемента?

Привести расширенное решение.

Время выполнения – 60 мин.

Ожидаемый результат:

Рассмотрим пошагово выполнение операций для двух случаев: когда стек пуст и когда в стеке уже есть 2 элемента.

1. Стек пуст (начальное состояние:  $S = []$ )

Размер стека: 6 ( $S[1..6]$ )

Операции:

PUSH(S, 4): Добавляем 4 в стек.

Стек: [4]

PUSH(S, 1): Добавляем 1 в стек.

Стек: [4, 1]

PUSH(S, 3): Добавляем 3 в стек.

Стек: [4, 1, 3]

POP(S): Удаляем верхний элемент (3).

Стек: [4, 1]

PUSH(S, 6): Добавляем 6 в стек.

Стек: [4, 1, 6]

PUSH(S, 4): Добавляем 4 в стек.

Стек: [4, 1, 6, 4]

PUSH(S, 1): Добавляем 1 в стек.

Стек: [4, 1, 6, 4, 1]

PUSH(S, 3): Добавляем 3 в стек.

Стек: [4, 1, 6, 4, 1, 3]

POP(S): Удаляем верхний элемент (3).

Стек: [4, 1, 6, 4, 1]

POP(S): Удаляем верхний элемент (1).

Стек: [4, 1, 6, 4]

Итоговое состояние стека:

$S = [4, 1, 6, 4]$

2. В стеке уже есть 2 элемента (начальное состояние:  $S = [a, b]$ )

Размер стека: 6 ( $S[1..6]$ )

Начальное состояние:

Стек: [a, b]

Операции:

PUSH(S, 4): Добавляем 4 в стек.

Стек: [a, b, 4]

PUSH(S, 1): Добавляем 1 в стек.  
Стек: [a, b, 4, 1]  
PUSH(S, 3): Добавляем 3 в стек.  
Стек: [a, b, 4, 1, 3]  
POP(S): Удаляем верхний элемент (3).  
Стек: [a, b, 4, 1]  
PUSH(S, 6): Добавляем 6 в стек.  
Стек: [a, b, 4, 1, 6]  
PUSH(S, 4): Добавляем 4 в стек.  
Стек: [a, b, 4, 1, 6, 4]  
PUSH(S, 1): Пытаемся добавить 1, но стек уже заполнен (6 элементов).  
Ошибка: Переполнение стека.  
Стек остается: [a, b, 4, 1, 6, 4]  
Итоговое состояние стека:  
S=[a,b,4,1,6,4]  
Правильный ответ: S=[4,1,6,4], . S=[a,b,4,1,6,4]  
Компетенции (индикаторы): ПК-1

3. Определить вероятность попадания случайной величины x в интервал от 10 до 16, если известно, что  $m_x = 13$ , а оценка дисперсии  $D = S^2 = 1$ .

Привести расширенное решение.

Время выполнения – 60 мин.

Ожидаемый результат:

Для решения задачи используем свойства нормального распределения, так как заданы математическое ожидание ( $m_x = 13$ ) и дисперсия ( $D = S^2 = 1$ ). Случайная величина x распределена по нормальному закону:  $x \sim N(m_x, \sigma^2)$ , где  $\sigma = \sqrt{D} = 1$

Дано:

Математическое ожидание:

$$m_x = 13$$

$$D = S^2 = 1$$

Интервал: от 10 до 16.

Нормализуем случайную величину:

Перейдем к стандартному нормальному распределению с помощью замены:

$$z = \frac{x - m_x}{\sigma}$$

$$\text{где } \sigma = \sqrt{D} = 1$$

Тогда границы интервала:

$$z_1 = \frac{10 - 13}{1} = -3$$

$$z_2 = \frac{16 - 13}{1} = 3$$

Найдем вероятность:

Вероятность попадания  $x$  в интервал от 10 до 16 равна вероятности попадания  $z$  в интервал от  $-3$  до  $3$

$$P(10 \leq x \leq 16) = P(-3 \leq z \leq 3)$$

Используем таблицу стандартного нормального распределения или калькулятор для нахождения вероятностей:

$$P(z \leq 3) \approx 0,99865$$

$$P(z \leq -3) \approx 0,00135$$

Тогда:

$$P(-3 \leq z \leq 3) = P(z \leq 3) - P(z \leq -3) = 0,99865 - 0,00135 = 0,9973$$

$$P(-3 \leq z \leq 3) = P(z \leq 3) - P(z \leq -3) = 0,99865 - 0,00135 = 0,9973.$$

Ответ:

Вероятность попадания случайной величины  $x$  в интервал от 10 до 16 равна: 0,9973

Правильный ответ: 0,9973.

Компетенции (индикаторы): ПК-1

4. Определить оценку математического ожидания, а также смещенную и несмещенную оценки дисперсии  $D_x$ , если в 3 опытах выпали цифры 7, 8 и 9.

Привести расширенное решение.

Время выполнения – 60 мин.

Ожидаемый результат:

1. Оценка математического ожидания:

Математическое ожидание  $m_x$  оценивается как среднее арифметическое значений:

$$m_x = \frac{x_1 + x_2 + x_3}{3} = \frac{7 + 8 + 9}{3} = \frac{24}{3} = 8$$

2. Смещенная оценка дисперсии:

Смещенная оценка дисперсии  $D_x$  вычисляется по формуле:

$$D_x = \frac{(x_1 - m_x)^2 + (x_2 - m_x)^2 + (x_3 - m_x)^2}{3}$$

Подставим значения:

$$(x_1 - m_x)^2 = (7 - 8)^2 = 1$$

$$(x_2 - m_x)^2 = (8 - 8)^2 = 0$$

$$(x_3 - m_x)^2 = (9 - 8)^2 = 1$$

Тогда:

$$D_x = \frac{1 + 0 + 1}{3} = \frac{2}{3} \approx 0,6667$$

3. Несмещенная оценка дисперсии:

Несмещенная оценка дисперсии  $S^2$  вычисляется по формуле:

$$S^2 = \frac{(x_1 - m_x)^2 + (x_2 - m_x)^2 + (x_3 - m_x)^2}{3 - 1}$$

Подставим значения:

$$S^2 = \frac{1 + 0 + 1}{2} = \frac{2}{2} = 1$$

Правильный ответ: 8, 0,6667, 1  
Компетенции (индикаторы): ПК-1

## Экспертное заключение

Представленный комплект оценочных материалов по дисциплине «Системы обработки данных» соответствует требованиям ФГОС ВО.

Предлагаемые оценочные материалы адекватны целям и задачам реализации основной профессиональной образовательной программы по направлению подготовки 09.04.04 Программная инженерия.

Виды оценочных средств, включённые в представленный фонд, отвечают основным принципам формирования ФОС.

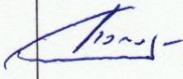
Разработанные и представленные для экспертизы оценочные материалы рекомендуются к использованию в процессе подготовки обучающихся по указанному направлению.

Председатель учебно-методической  
комиссии института компьютерных  
систем и информационных технологий



Ветрова Н.Н.

## Лист изменений и дополнений

№ п/п	Виды дополнений и изменений	Дата и номер протокола заседания кафедры (кафедр), на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)
1.	Дополнен комплектом оценочных материалов	протокол заседания кафедры компьютерных систем и сетей № <u>8</u> от <u>10.03.2025</u>	 С.В. Попов