

Документ подписан простой электронной подписью
 Информация о владельце:
 ФИО: Стрельникова Светлана Евгеньевна
 Должность: ректор
 Дата подписания: 28.09.2026 16:46:35
 Уникальный программный ключ:
 ed7e7ef871ef28d870011250a92f17921078cc18

**Автономная некоммерческая образовательная организация высшего образования
 «Контур Университет»**



Технологии обработки больших данных рабочая программа дисциплины (модуля)

Учебный план	09.03.03_ИИС_2026.plx 09.03.03 Прикладная информатика		
Квалификация	бакалавр		
Форма обучения	очная		
Общая трудоемкость	3 ЗЕТ		
Часов по учебному плану	108	Виды контроля в семестрах:	
в том числе:		зачеты 6	
аудиторные занятия	64		
самостоятельная работа	42		
часов на контроль	2		

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	6 (3.2)		Итого	
Неделя	17 2/6			
Вид занятий	уп	рп	уп	рп
Лекции	16	16	16	16
Практические	48	48	48	48
Итого ауд.	64	64	64	64
Контактная работа	64	64	64	64
Сам. работа	42	42	42	42
Часы на контроль	2	2	2	2
Итого	108	108	108	108

Рабочая программа дисциплины

Технологии обработки больших данных

разработана в соответствии с ФГОС ВО:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.03 Прикладная информатика (приказ Минобрнауки России от 19.09.2017 г. № 922)

составлена на основании учебного плана:

09.03.03 Прикладная информатика

Одобрена Ученым советом вуза

Протокол от 15.09.2026 № 05.

Одобрена учебно-методической комиссией

Протокол от 10.09.2026 г. № 01

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
1.1	Цель освоения дисциплины:
1.2	Формирование у обучающихся системы знаний о современных технологиях, методах и инструментах сбора, хранения, обработки и анализа больших данных, а также развитие практических навыков применения распределенных вычислительных систем и платформ для решения задач анализа данных в профессиональной деятельности, включая бизнес-аналитику, прогнозирование и поддержку управленческих решений.
1.3	Задачи изучения дисциплины:
1.4	Ознакомить с концепцией больших данных, их характеристиками (объем, скорость, разнообразие), архитектурными подходами и современными платформами для работы с большими данными (Hadoop, Spark, Kafka, Flink);
1.5	Сформировать знания о методах и технологиях сбора, хранения и предобработки больших данных: распределенные файловые системы (HDFS), NoSQL-базы данных (MongoDB, Cassandra), системы потоковой обработки данных;
1.6	Развить навыки использования инструментов для обработки и анализа больших данных: MapReduce, Apache Spark (PySpark), SQL-движки для больших данных (Hive, Presto/Trino), библиотеки машинного обучения для больших данных (MLlib);
1.7	Сформировать умения проектировать и реализовывать ETL-процессы для больших данных, строить хранилища данных (Data Lake, Data Warehouse) и витрины данных для поддержки принятия управленческих решений;
1.8	Обеспечить понимание методов анализа больших данных: статистический анализ, машинное обучение на больших данных, обработка потоковых данных в реальном времени и визуализация результатов.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	
Цикл (раздел) ОП:	Б1.В
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Базы данных и SQL
2.1.2	Программирование
2.1.3	Математический анализ
2.1.4	Теория вероятности и математическая статистика
2.1.5	Введение в профессиональную деятельность
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Интеллектуальные информационные системы
2.2.2	Производственная практика (эксплуатационная)

3. ФОРМИРУЕМЫЕ КОМПЕТЕНЦИИ И ИНДИКАТОРЫ ИХ ДОСТИЖЕНИЯ	
ПК-4: Способен применять методы и программные средства сбора, обработки и анализа информации для информационно-аналитической поддержки принятия управленческих решений	
ПК-4.2: Проектирует процедуры сбора данных для решения конкретных управленческих задач (сбор показателей эффективности подразделений, оценка удовлетворённости клиентов) с использованием методов опросов, анкетирования, интервьюирования, веб-аналитики, парсинга и API государственных систем	
Знать: методы проектирования процедур сбора данных из гетерогенных источников ; инструменты и методы сбора данных (веб-аналитика, парсинг, API, потоковые данные) ; подходы к оценке удовлетворённости клиентов и сбору показателей эффективности подразделений Уметь: проектировать процедуры сбора данных для решения конкретных управленческих задач ; выбирать и применять методы сбора данных в зависимости от задачи; использовать веб-аналитику, парсинг и API для сбора данных Владеть: навыками проектирования процедур сбора данных из разнородных источников ; методами сбора и обработки данных для управленческих задач; техниками работы с API и веб-аналитикой	
ПК-4.3: Подготавливает данные из учётных систем и внешних источников к анализу с использованием методов ETL/ELT-процессов, обработки пропусков и выбросов, нормализации и стандартизации, формируя единое аналитическое хранилище (витрину данных)	
Знать: методы ETL/ELT-процессов для подготовки данных к анализу; подходы к обработке пропусков, выбросов, нормализации и стандартизации данных; принципы формирования единого аналитического хранилища (витрины данных) Уметь: подготавливать данные из учётных систем и внешних источников к анализу;	

применять методы ETL/ELT-процессов для обработки данных; формировать единое аналитическое хранилище для поддержки принятия решений Владеть: навыками ETL/ELT-процессов для подготовки данных к анализу; методами обработки пропусков и выбросов; техниками нормализации и стандартизации данных
ПК-4.5: Применяет методы прикладной статистики (регрессионный анализ, кластеризация, прогнозирование временных рядов, факторный анализ, сценарное моделирование) для выявления закономерностей в бизнес-данных и формулирования выводов для руководителей
Знать: методы прикладной статистики для анализа бизнес-данных; подходы к выявлению закономерностей в больших данных; методы прогнозирования и факторного анализа для обоснования управленческих решений Уметь: применять методы прикладной статистики для выявления закономерностей в бизнес-данных; интерпретировать результаты статистического анализа и формулировать выводы для руководителей; использовать методы прогнозирования и факторного анализа для поддержки принятия решений Владеть: навыками применения регрессионного анализа и кластеризации для анализа больших данных; методами прогнозирования временных рядов; техниками интерпретации результатов анализа для руководства

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)				
Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Примечание
	Раздел 1.			
1.1	Введение в большие данные: понятие, характеристики (объём, скорость, разнообразие, достоверность, ценность) /Лек/	6	1	
1.2	Введение в большие данные: понятие, характеристики (объём, скорость, разнообразие, достоверность, ценность) /Пр/	6	3	
1.3	Введение в большие данные: понятие, характеристики (объём, скорость, разнообразие, достоверность, ценность) /Ср/	6	2	
1.4	Архитектуры систем обработки больших данных: Lambda-архитектура и Карра-архитектура /Лек/	6	1	
1.5	Архитектуры систем обработки больших данных: Lambda-архитектура и Карра-архитектура /Пр/	6	3	
1.6	Архитектуры систем обработки больших данных: Lambda-архитектура и Карра-архитектура /Ср/	6	2	
1.7	Методы сбора данных из гетерогенных источников (API, веб-скрапинг, лог-файлы, сенсорные данные) /Лек/	6	1	
1.8	Методы сбора данных из гетерогенных источников (API, веб-скрапинг, лог-файлы, сенсорные данные) /Пр/	6	3	
1.9	Методы сбора данных из гетерогенных источников (API, веб-скрапинг, лог-файлы, сенсорные данные) /Ср/	6	2	
1.10	Распределённые файловые системы: архитектура и принципы работы (HDFS) /Лек/	6	1	
1.11	Распределённые файловые системы: архитектура и принципы работы (HDFS) /Пр/	6	3	
1.12	Распределённые файловые системы: архитектура и принципы работы (HDFS) /Ср/	6	2	
1.13	NoSQL-базы данных: классификация, применение, сравнение с реляционными СУБД /Лек/	6	1	
1.14	NoSQL-базы данных: классификация, применение, сравнение с реляционными СУБД /Пр/	6	3	
1.15	NoSQL-базы данных: классификация, применение, сравнение с реляционными СУБД /Ср/	6	2	
1.16	ETL и ELT-процессы: этапы, методы, применение для больших данных /Лек/	6	1	
1.17	ETL и ELT-процессы: этапы, методы, применение для больших данных /Пр/	6	3	

1.18	ETL и ELT-процессы: этапы, методы, применение для больших данных /Ср/	6	2	
1.19	Предобработка данных: обработка пропусков, выбросов, нормализация и стандартизация /Лек/	6	1	
1.20	Предобработка данных: обработка пропусков, выбросов, нормализация и стандартизация /Пр/	6	3	
1.21	Предобработка данных: обработка пропусков, выбросов, нормализация и стандартизация /Ср/	6	3	
1.22	Модель программирования MapReduce: принципы параллельной обработки данных /Лек/	6	1	
1.23	Модель программирования MapReduce: принципы параллельной обработки данных /Пр/	6	3	
1.24	Модель программирования MapReduce: принципы параллельной обработки данных /Ср/	6	3	
1.25	Apache Spark: архитектура, компоненты, работа с DataFrames и Dataset /Лек/	6	1	
1.26	Apache Spark: архитектура, компоненты, работа с DataFrames и Dataset /Пр/	6	3	
1.27	Apache Spark: архитектура, компоненты, работа с DataFrames и Dataset /Ср/	6	3	
1.28	SQL-движки для больших данных: Hive, Presto/Trino /Лек/	6	1	
1.29	SQL-движки для больших данных: Hive, Presto/Trino /Пр/	6	3	
1.30	SQL-движки для больших данных: Hive, Presto/Trino /Ср/	6	3	
1.31	Обработка потоковых данных в реальном времени: Spark Streaming /Лек/	6	1	
1.32	Обработка потоковых данных в реальном времени: Spark Streaming /Пр/	6	3	
1.33	Обработка потоковых данных в реальном времени: Spark Streaming /Ср/	6	3	
1.34	Библиотека Spark MLlib для машинного обучения на больших данных /Лек/	6	1	
1.35	Библиотека Spark MLlib для машинного обучения на больших данных /Пр/	6	3	
1.36	Библиотека Spark MLlib для машинного обучения на больших данных /Ср/	6	3	
1.37	Методы прикладной статистики для анализа больших данных (регрессия, кластеризация, прогнозирование временных рядов) /Лек/	6	1	
1.38	Методы прикладной статистики для анализа больших данных (регрессия, кластеризация, прогнозирование временных рядов) /Пр/	6	3	
1.39	Методы прикладной статистики для анализа больших данных (регрессия, кластеризация, прогнозирование временных рядов) /Ср/	6	3	
1.40	Визуализация больших данных: принципы и инструменты /Лек/	6	1	
1.41	Визуализация больших данных: принципы и инструменты /Пр/	6	5	
1.42	Визуализация больших данных: принципы и инструменты /Ср/	6	5	
1.43	Формирование витрин данных и Data Lake для поддержки управленческих решений /Лек/	6	2	
1.44	Формирование витрин данных и Data Lake для поддержки управленческих решений /Пр/	6	4	
1.45	Формирование витрин данных и Data Lake для поддержки управленческих решений /Ср/	6	4	
1.46	/Зачёт/	6	2	

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

5.1. Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации

Темы для обсуждения на практических занятиях:

Практическое задание № 1 «Анализ характеристик больших данных» (к теме 1.1)

Тема: Введение в большие данные: понятие, характеристики (объём, скорость, разнообразие, достоверность, ценность).

Цель: сформировать понимание концепции больших данных и их характеристик.

Задания:

1. Изучите концепцию больших данных. Дайте определение больших данных.
2. Охарактеризуйте каждую из характеристик больших данных (объём, скорость, разнообразие, достоверность, ценность).

Приведите примеры.

3. Для 3 предложенных бизнес-задач (прогнозирование продаж, анализ поведения клиентов в интернет-магазине, мониторинг работы оборудования в реальном времени) определите, какие характеристики больших данных являются ключевыми.

4. Подготовьте презентацию (5–7 слайдов) о применении больших данных в одной из отраслей (финансы, ритейл, производство, логистика, телеком).

Форма отчётности: Аналитическая записка, презентация.

Практическое задание № 2 «Сравнительный анализ архитектур Lambda и Карра» (к теме 1.2)

Тема: Архитектуры систем обработки больших данных: Lambda-архитектура и Карра-архитектура.

Цель: освоить сравнительный анализ архитектурных подходов к обработке больших данных.

Задания:

1. Изучите Lambda-архитектуру и Карра-архитектуру. Заполните сравнительную таблицу по критериям: компоненты, преимущества, недостатки, области применения.

2. Для каждого из 3 сценариев (обработка исторических данных, аналитика в реальном времени, гибридный сценарий) определите, какая архитектура наиболее подходит. Обоснуйте выбор.

3. Разработайте схему архитектуры для гипотетической Big Data-системы для интернет-магазина. Опишите компоненты и потоки данных.

Форма отчётности: Сравнительная таблица, схема архитектуры, обоснование выбора.

Практическое задание № 3 «Сбор данных из гетерогенных источников» (к теме 1.3)

Тема: Методы сбора данных из гетерогенных источников (API, веб-скрапинг, лог-файлы, сенсорные данные).

Цель: освоить методы сбора данных из гетерогенных источников.

Задания:

1. Изучите методы сбора данных: API, веб-скрапинг, работа с лог-файлами, сенсорные данные. Опишите каждый метод, его преимущества и ограничения.

2. Используя доступный API (например, для получения данных о погоде, курсах валют), получите данные и сохраните их в формате JSON.

3. Выполните веб-скрапинг для сбора данных с веб-страницы (с соблюдением этических норм и правил robots.txt).

4. Проведите анализ лог-файла (предложенного преподавателем): определите структуру, извлеките ключевую информацию.

5. Сравните полученные данные из разных источников. Оцените качество данных и предложите способы их интеграции.

Форма отчётности: Код сбора данных, описание методов, анализ результатов.

Практическое задание № 4 «Работа с распределённой файловой системой» (к теме 1.4)

Тема: Распределённые файловые системы: архитектура и принципы работы (HDFS).

Цель: освоить работу с распределённой файловой системой.

Задания:

1. Изучите архитектуру HDFS: NameNode, DataNode, блоки данных, репликация. Опишите принципы работы.

2. Установите и настройте среду для работы с HDFS (локальная установка или использование облачной среды).

3. Выполните основные операции с HDFS: создание каталогов, загрузка файлов, чтение, копирование, удаление.

4. Сравните HDFS с традиционной файловой системой по критериям: масштабируемость, отказоустойчивость, производительность.

5. Опишите, как HDFS используется для построения Data Lake.

Форма отчётности: Отчёт по работе с HDFS, сравнительный анализ.

Практическое задание № 5 «Выбор и работа с NoSQL-базой данных» (к теме 1.5)

Тема: NoSQL-базы данных: классификация, применение, сравнение с реляционными СУБД.

Цель: освоить выбор и работу с NoSQL-базами данных для больших данных.

Задания:

1. Изучите классификацию NoSQL-баз данных (документно-ориентированные, столбцовые, ключ-значение, графовые).

Заполните сравнительную таблицу.

2. Изучите теорему CAP. Объясните, как она влияет на выбор баз данных для больших данных.

3. Установите одну из NoSQL-баз данных (например, MongoDB или Cassandra). Выполните базовые операции: создание коллекции, вставка, чтение, обновление, удаление данных.

4. Сравните работу с NoSQL-базой и реляционной СУБД на примере одной и той же задачи.

5. Предложите, какая NoSQL-база подходит для заданного сценария (например, хранение логов, управление профилями пользователей, граф связей). Обоснуйте выбор.

Форма отчётности: Сравнительная таблица, отчёт по работе с NoSQL, обоснование выбора.

Практическое задание № 6 «Разработка ETL-процесса» (к теме 1.6)

Тема: ETL и ELT-процессы: этапы, методы, применение для больших данных.

Цель: освоить разработку ETL-процессов для больших данных.

Задания:

1. Изучите этапы ETL-процесса: извлечение, трансформация, загрузка. Опишите методы трансформации данных.
2. Для заданного набора данных (CSV-файлы, JSON, данные из API) разработайте ETL-процесс:
 - о Извлечение данных из источника.
 - о Трансформация: очистка, нормализация, обогащение.
 - о Загрузка в целевое хранилище (базу данных или файловую систему).
3. Реализуйте ETL-процесс с использованием Python или Spark.
4. Оцените качество данных до и после трансформации.
5. Предложите, как ETL-процесс может быть использован для построения витрины данных.

Форма отчётности: Код ETL-процесса, описание этапов, анализ качества данных.

Практическое задание № 7 «Предобработка данных для анализа» (к теме 1.7)

Тема: Предобработка данных: обработка пропусков, выбросов, нормализация и стандартизация.

Цель: освоить методы предобработки данных для анализа.

Задания:

1. Изучите методы обработки пропусков (удаление, заполнение средним/медианой, интерполяция) и выбросов (методы IQR, Z-score).
2. Изучите методы нормализации (Min-Max) и стандартизации (Z-score). В чём их отличие?
3. Для заданного набора данных (содержащего пропуски и выбросы) выполните предобработку:
 - о Определите пропуски и выбросы.
 - о Обработайте пропуски.
 - о Обработайте выбросы.
 - о Выполните нормализацию и стандартизацию.
4. Сравните результаты применения разных методов предобработки.
5. Подготовьте отчёт с описанием выполненных шагов и обоснованием выбора методов.

Форма отчётности: Отчёт по предобработке данных, код.

Практическое задание № 8 «Реализация MapReduce-задачи» (к теме 1.8)

Тема: Модель программирования MapReduce: принципы параллельной обработки данных.

Цель: освоить модель программирования MapReduce для параллельной обработки данных.

Задания:

1. Изучите модель MapReduce: фаза Map, фаза Reduce, принципы параллельной обработки.
2. Для заданной задачи (например, подсчёт частоты слов в тексте, обработка логов, агрегация данных) реализуйте MapReduce-задачу.
3. Запустите задачу на локальной установке Hadoop или в облачной среде.
4. Проанализируйте результаты выполнения: время, объём обработанных данных.
5. Сравните производительность MapReduce с другими подходами. Опишите ограничения MapReduce.

Форма отчётности: Код MapReduce-задачи, отчёт о выполнении, анализ производительности.

Практическое задание № 9 «Разработка приложения на Apache Spark» (к теме 1.9)

Тема: Apache Spark: архитектура, компоненты, работа с DataFrames и Dataset.

Цель: освоить разработку приложений на Apache Spark для обработки больших данных.

Задания:

1. Изучите архитектуру Apache Spark: драйвер, исполнители, кластерный менеджер. Опишите компоненты Spark.
2. Установите и настройте Spark (локально или в облачной среде).
3. Загрузите набор данных и выполните с ним операции:
 - о Создание DataFrame.
 - о Фильтрация, группировка, агрегация.
 - о Применение пользовательских функций.
4. Выполните анализ данных с использованием Spark SQL.
5. Оптимизируйте выполнение запросов: используйте кэширование, партиционирование.
6. Сравните производительность Spark с MapReduce на примере одной задачи.

Форма отчётности: Код Spark-приложения, отчёт о выполнении, анализ производительности.

Практическое задание № 10 «Работа с SQL-движками для больших данных» (к теме 1.10)

Тема: SQL-движки для больших данных: Hive, Presto/Trino.

Цель: освоить работу с SQL-движками для больших данных.

Задания:

1. Изучите архитектуру и принципы работы Hive и Presto/Trino. Заполните сравнительную таблицу.
2. Установите и настройте Hive. Создайте внешнюю таблицу на данных в HDFS.
3. Выполните SQL-запросы к данным: выборка, агрегация, соединение таблиц.
4. Сравните производительность Hive и Presto/Trino на одном наборе данных.
5. Предложите сценарии использования каждого движка в зависимости от задач.

Форма отчётности: Сравнительная таблица, отчёт по выполнению запросов, анализ производительности.

Практическое задание № 11 «Обработка потоковых данных в реальном времени» (к теме 1.11)

Тема: Обработка потоковых данных в реальном времени: Spark Streaming.

Цель: освоить обработку потоковых данных с использованием Spark Streaming.

Задания:

1. Изучите концепцию потоковой обработки данных. В чём отличие потоковой обработки от пакетной?
2. Настройте среду для Spark Streaming.
3. Реализуйте приложение для обработки потоковых данных (например, анализ логов в реальном времени, подсчёт событий за окно).

4. Проанализируйте результаты: как изменяются показатели во времени.

5. Опишите, как потоковая обработка может применяться для бизнес-задач (мониторинг, оповещения, аналитика в реальном времени).

Форма отчётности: Код приложения, отчёт о результатах, описание бизнес-применения.

Практическое задание № 12 «Применение машинного обучения на больших данных» (к теме 1.12)

Тема: Библиотека Spark MLlib для машинного обучения на больших данных.

Цель: освоить применение машинного обучения на больших данных с использованием Spark MLlib.

Задания:

1. Изучите библиотеку Spark MLlib: алгоритмы классификации, регрессии, кластеризации.
2. Загрузите большой набор данных (например, данные о клиентах, транзакциях).
3. Выполните предобработку данных с использованием Spark.
4. Постройте модель машинного обучения (классификация, регрессия или кластеризация).
5. Оцените качество модели с использованием метрик (accuracy, precision, recall, RMSE).
6. Сравните производительность модели на малых и больших данных.
7. Интерпретируйте результаты модели для бизнеса.

Форма отчётности: Код модели, оценка качества, интерпретация результатов.

Практическое задание № 13 «Статистический анализ больших данных» (к теме 1.13)

Тема: Методы прикладной статистики для анализа больших данных (регрессия, кластеризация, прогнозирование временных рядов).

Цель: освоить методы прикладной статистики для анализа больших данных.

Задания:

1. Изучите методы регрессионного анализа, кластеризации и прогнозирования временных рядов.
2. Загрузите большой набор данных (например, данные о продажах, клиентские данные).
3. Примените регрессионный анализ для выявления зависимостей между переменными.
4. Примените кластеризацию для сегментации данных.
5. Выполните прогнозирование временных рядов (например, прогноз продаж).
6. Интерпретируйте полученные результаты. Сформулируйте выводы для руководителей.

Форма отчётности: Отчёт по статистическому анализу, интерпретация результатов.

Практическое задание № 14 «Визуализация больших данных» (к теме 1.14)

Тема: Визуализация больших данных: принципы и инструменты.

Цель: освоить визуализацию результатов анализа больших данных.

Задания:

1. Изучите принципы эффективной визуализации данных. Какие типы графиков используются для разных типов данных?
2. Выберите инструмент для визуализации данных (например, Tableau, Power BI, Matplotlib, Plotly).
3. На основе результатов анализа (из предыдущих работ) создайте дашборд для визуализации ключевых показателей.
4. Добавьте интерактивные элементы (фильтры, срезы).
5. Подготовьте презентацию с визуализацией результатов для руководства.

Форма отчётности: Дашборд, презентация.

Практическое задание № 15 «Построение витрины данных и Data Lake» (к теме 1.15)

Тема: Формирование витрин данных и Data Lake для поддержки управленческих решений.

Цель: освоить формирование витрин данных и Data Lake для поддержки управленческих решений.

Задания:

1. Изучите концепции Data Lake и витрин данных. В чём их отличие?
2. Разработайте архитектуру Data Lake для гипотетической компании (интернет-магазин, производственное предприятие).
3. Опишите источники данных, процесс загрузки, структуру хранения.
4. Спроектируйте витрину данных для конкретной управленческой задачи (анализ продаж, клиентская аналитика, финансовый анализ).
5. Реализуйте ETL-процесс для загрузки данных в витрину (с использованием Python или Spark).
6. Подготовьте отчёт с описанием архитектуры и результатов загрузки данных.

Форма отчётности: Архитектура Data Lake и витрины данных, код ETL-процесса, отчёт.

Примерные вопросы к зачету:

1. Понятие больших данных. Характеристики больших данных (объём, скорость, разнообразие, достоверность, ценность).
2. Эволюция систем хранения и обработки данных: от реляционных баз данных к распределённым системам.
3. Классификация задач обработки больших данных: пакетная обработка, потоковая обработка, интерактивный анализ.
4. Lambda-архитектура: компоненты и принципы работы.
5. Карра-архитектура: компоненты и принципы работы.
6. Сравнительный анализ Lambda-архитектуры и Карра-архитектуры.
7. Требования к Big Data-системам: масштабируемость, отказоустойчивость, производительность.
8. Методы сбора данных из гетерогенных источников (API, веб-скрапинг, лог-файлы, сенсорные данные).
9. Проектирование процедур сбора данных для решения управленческих задач.
10. Оценка качества данных и методы очистки.
11. Архитектура распределённой файловой системы HDFS (NameNode, DataNode, блоки данных, репликация).

12. Основные операции с HDFS: запись, чтение, удаление данных.
13. Сравнение HDFS с другими распределёнными файловыми системами.
14. Использование распределённого хранения для построения Data Lake.
15. Понятие NoSQL. Отличие NoSQL от реляционных СУБД.
16. Теорема CAP и её влияние на выбор баз данных для больших данных.
17. Классификация NoSQL-баз данных: документо-ориентированные, столбцовые, ключ-значение, графовые.
18. Применение NoSQL-баз данных для работы с большими данными.
19. Выбор типа NoSQL-базы в зависимости от задачи.
20. Понятие ETL и ELT-процессов. Этапы ETL-процесса.
21. Обработка пропусков и выбросов в данных.
22. Нормализация и стандартизация данных.
23. Формирование витрин данных для поддержки управленческих решений.
24. Модель программирования MapReduce: этапы Map и Reduce.
25. Экосистема Hadoop: компоненты и их назначение (HDFS, YARN, MapReduce).
26. Ограничения MapReduce и переход к более современным инструментам.
27. Архитектура Apache Spark: драйвер, исполнители, кластерный менеджер.
28. Основные компоненты Spark: Spark Core, Spark SQL, Spark Streaming, MLlib, GraphX.
29. Работа с Spark DataFrame и Dataset.
30. Оптимизация Spark-приложений.
31. Понятие SQL-on-Hadoop. Apache Hive: метастор, HiveQL, интеграция с Hadoop.
32. Presto/Trino: распределённый SQL-движок для интерактивных запросов.
33. Сравнение производительности и областей применения SQL-движков для больших данных.
34. Понятие потоковой обработки данных. Отличие потоковой обработки от пакетной.
35. Apache Spark Streaming: микропакетная обработка.
36. Обработка данных в реальном времени для бизнес-задач.
37. Библиотека Spark MLlib для машинного обучения на больших данных.
38. Алгоритмы классификации, регрессии, кластеризации в распределённой среде.
39. Методы прикладной статистики для анализа больших данных: регрессионный анализ, кластеризация, прогнозирование временных рядов.
40. Принципы визуализации больших данных. Инструменты визуализации данных.
41. Создание дашбордов для мониторинга ключевых показателей.
42. Представление результатов анализа для различных категорий пользователей.
43. Интерпретация результатов статистического анализа для руководителей.
44. Data Lake как подход к хранению больших данных.
45. Витрины данных: назначение, структура, методы построения.
46. Проектирование ETL-процессов для больших данных.
47. Интеграция данных из учётных систем и внешних источников.
48. Оценка эффективности систем обработки больших данных.
49. Применение больших данных в управлении бизнесом (бизнес-аналитика, прогнозирование, поддержка управленческих решений).
50. Перспективы развития технологий обработки больших данных.

5.2. Темы письменных работ

не предусмотрены

5.3. Фонд оценочных средств

Тестовые задания:

I. ЗАДАНИЯ С ОДНИМ ВЕРНЫМ ВАРИАНТОМ ОТВЕТА (№ 1–10)

Инструкция: прочитайте текст, выберите правильный вариант ответа и запишите только букву выбранного варианта ответа (А, Б, В или Г).

ПК-4.2. Вопрос № 1. Что представляет собой процесс проектирования процедур сбора данных для решения управленческих задач?

- А) Случайный выбор источников данных без обоснования
- Б) Определение целей, источников, методов и форматов сбора данных для конкретной задачи
- В) Автоматическое копирование всех данных из внешних систем
- Г) Сбор данных только из внутренней отчётности организации

Ответ:

ПК-4.2. Вопрос № 2. Какой метод сбора данных используется для анализа поведения пользователей на сайте?

- А) Интервьюирование
- Б) Опросы и анкетирование
- В) Веб-аналитика
- Г) Парсинг государственных систем

Ответ:

ПК-4.2. Вопрос № 3. Для сбора данных о показателях эффективности подразделений наиболее целесообразно использовать:

- А) Только данные из открытых источников
- Б) Данные из учётных систем организации (ERP, CRM, 1C)

В) Субъективное мнение руководителей подразделений

Г) Анализ социальных сетей

Ответ:

ПК-4.2. Вопрос № 4. Оценка удовлетворённости клиентов с использованием стандартизированных вопросников относится к методу:

А) Интервьюирования

Б) Анкетирования

В) Веб-аналитики

Г) Парсинга данных

Ответ:

ПК-4.3. Вопрос № 5. Что является основным результатом ETL-процесса?

А) Визуализация данных в виде дашбордов

Б) Формирование единого аналитического хранилища (витрины данных)

В) Сбор данных из внешних источников

Г) Проведение статистического анализа

Ответ:

ПК-4.3. Вопрос № 6. Какой метод обработки данных применяется для устранения пропусков в наборе данных?

А) Нормализация данных

Б) Заполнение пропусков средним значением, медианой или интерполяция

В) Удаление дубликатов

Г) Стандартизация данных

Ответ:

ПК-4.3. Вопрос № 7. Что такое нормализация данных (Min-Max scaling)?

А) Преобразование данных к стандартному нормальному распределению

Б) Преобразование данных к диапазону [0, 1]

В) Удаление выбросов из данных

Г) Сортировка данных по возрастанию

Ответ:

ПК-4.5. Вопрос № 8. Для выявления скрытых групп клиентов с похожими характеристиками используется метод:

А) Регрессионный анализ

Б) Кластеризация

В) Прогнозирование временных рядов

Г) Сценарное моделирование

Ответ:

ПК-4.5. Вопрос № 9. Какой метод используется для прогнозирования будущих значений на основе исторических данных?

А) Регрессионный анализ

Б) Кластеризация

В) Прогнозирование временных рядов

Г) Факторный анализ

Ответ:

ПК-4.5. Вопрос № 10. Для оценки влияния нескольких независимых переменных на зависимую переменную используется:

А) Регрессионный анализ

Б) Кластеризация

В) Факторный анализ

Г) Сценарное моделирование

Ответ:

II. ЗАДАНИЯ С ТРЕМЯ ВЕРНЫМИ ВАРИАНТАМИ ОТВЕТА (№ 11–20)

Инструкция: прочитайте текст, выберите три правильных варианта ответа и запишите только буквы выбранных вариантов ответов (А, Б, В, Г или Д).

ПК-4.2. Вопрос № 11. Выберите три метода сбора данных для оценки удовлетворённости клиентов:

А) Анкетирование

Б) Интервьюирование

В) Веб-аналитика (анализ поведения на сайте)

Г) Парсинг государственных реестров

Д) Сбор данных с сенсоров

Ответ:

ПК-4.2. Вопрос № 12. Выберите три источника данных для сбора показателей эффективности подразделений:

А) Учётные системы организации (ERP, CRM, 1C)

Б) Данные внешних государственных систем

В) Внутренние отчёты подразделений
Г) Социальные сети без привязки к задачам
Д) Случайные данные из интернета
Ответ:

ПК-4.2. Вопрос № 13. Выберите три инструмента для сбора данных из внешних источников:

А) API государственных систем
Б) Веб-парсинг (скрапинг)
В) Веб-аналитика (Google Analytics, Яндекс.Метрика)
Г) Текстовый редактор
Д) Электронная почта
Ответ:

ПК-4.3. Вопрос № 14. Выберите три этапа ETL/ELT-процесса:

А) Извлечение (Extract)
Б) Трансформация (Transform)
В) Загрузка (Load)
Г) Валидация (Validate)
Д) Визуализация (Visualize)
Ответ:

ПК-4.3. Вопрос № 15. Выберите три метода обработки пропусков в данных:

А) Удаление строк с пропусками
Б) Заполнение средним значением
В) Интерполяция
Г) Замена на ноль
Д) Шифрование данных
Ответ:

ПК-4.3. Вопрос № 16. Выберите три метода обработки данных перед загрузкой в витрину:

А) Обработка пропусков
Б) Удаление выбросов
В) Нормализация и стандартизация
Г) Сжатие данных
Д) Сортировка
Ответ:

ПК-4.5. Вопрос № 17. Выберите три метода прикладной статистики для анализа бизнес-данных:

А) Регрессионный анализ
Б) Кластеризация
В) Прогнозирование временных рядов
Г) Блокчейн
Д) Майнинг криптовалют
Ответ:

ПК-4.5. Вопрос № 18. Выберите три задачи, решаемые с помощью кластеризации:

А) Сегментация клиентов
Б) Группировка товаров по схожим характеристикам
В) Выявление групп риска
Г) Прогнозирование цен на акции
Д) Расчёт налогов
Ответ:

ПК-4.5. Вопрос № 19. Выберите три области применения регрессионного анализа:

А) Прогнозирование продаж
Б) Выявление факторов, влияющих на удовлетворённость клиентов
В) Оценка влияния рекламных затрат на выручку
Г) Классификация изображений
Д) Создание веб-сайта
Ответ:

ПК-4.5. Вопрос № 20. Выберите три метода прогнозирования временных рядов:

А) ARIMA
Б) Экспоненциальное сглаживание
В) Скользящее среднее
Г) Линейная регрессия
Д) Кластеризация
Ответ:

III. ЗАДАНИЯ НА УСТАНОВЛЕНИЕ СООТВЕТСТВИЯ (№ 21–25)

Инструкция: к каждому элементу из левого столбца подберите ОДИН соответствующий элемент из правого столбца. Ответ запишите в виде сочетания цифр и букв (например: 1-А, 2-Б, 3-В). Каждая буква используется только один раз.

ПК-4.2. Вопрос № 21. Установите соответствие между методом сбора данных и его описанием:

Метод сбора данных Описание

1. Анкетирование А. Программный интерфейс для получения данных из внешней системы
2. Интервьюирование Б. Стандартизированный сбор ответов от группы респондентов
3. API В. Живая беседа с респондентом по заранее подготовленным вопросам

Ответ: 1 __, 2 __, 3 __

ПК-4.2. Вопрос № 22. Установите соответствие между инструментом сбора данных и его применением:

Инструмент Применение

1. Веб-аналитика А. Автоматический сбор данных с веб-страниц
2. Парсинг Б. Анализ поведения пользователей на сайте
3. API государственных систем В. Получение официальных данных из государственных реестров

Ответ: 1 __, 2 __, 3 __

ПК-4.3. Вопрос № 23. Установите соответствие между методом обработки данных и его описанием:

Метод обработки Описание

1. Нормализация А. Преобразование данных к стандартному распределению со средним 0 и дисперсией 1
2. Стандартизация Б. Приведение данных к диапазону [0, 1]
3. Обработка выбросов В. Выявление и устранение anomalously отличающихся значений

Ответ: 1 __, 2 __, 3 __

ПК-4.5. Вопрос № 24. Установите соответствие между методом статистического анализа и его описанием:

Метод Описание

1. Регрессионный анализ А. Выявление скрытых групп объектов со схожими характеристиками
2. Кластеризация Б. Прогнозирование будущих значений на основе исторических данных
3. Прогнозирование временных рядов В. Оценка влияния независимых переменных на зависимую переменную

Ответ: 1 __, 2 __, 3 __

ПК-4.5. Вопрос № 25. Установите соответствие между задачей управления и методом анализа данных:

Задача управления Метод анализа

1. Прогноз продаж на следующий месяц А. Кластеризация
2. Сегментация клиентской базы Б. Прогнозирование временных рядов
3. Оценка влияния цен на спрос В. Регрессионный анализ

Ответ: 1 __, 2 __, 3 __

IV. ЗАДАНИЯ НА УСТАНОВЛЕНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ (№ 26–30)

Инструкция: прочитайте текст, установите правильную последовательность действий и запишите аргументы, обосновывающие выбранный порядок.

ПК-4.2. Вопрос № 26. Установите правильную последовательность проектирования процедуры сбора данных для управленческой задачи:

- А) Определить методы и инструменты сбора данных
- Б) Сформулировать цель и задачи сбора данных
- В) Определить источники данных
- Г) Разработать план сбора данных

Ответ:

ПК-4.2. Вопрос № 27. Установите правильную последовательность сбора данных для оценки удовлетворённости клиентов:

- А) Провести анкетирование клиентов
- Б) Разработать анкету с вопросами
- В) Определить целевую аудиторию
- Г) Проанализировать результаты опроса

Ответ:

ПК-4.3. Вопрос № 28. Установите правильную последовательность этапов ETL-процесса при подготовке данных к анализу:

- А) Загрузка данных в витрину данных
- Б) Трансформация данных (очистка, нормализация, стандартизация)
- В) Извлечение данных из источников

Ответ:

ПК-4.3. Вопрос № 29. Установите правильную последовательность обработки данных перед загрузкой в аналитическое хранилище:

- А) Обработать пропуски (удаление или заполнение)
- Б) Удалить выбросы (аномальные значения)
- В) Выполнить нормализацию или стандартизацию

Г) Проверить качество данных

Ответ:

ПК-4.5. Вопрос № 30. Установите правильную последовательность применения статистического анализа для поддержки управленческого решения:

А) Сформулировать выводы для руководителей и подготовить рекомендации

Б) Выбрать метод статистического анализа (регрессия, кластеризация и т.д.)

В) Подготовить данные для анализа (очистка, нормализация)

Г) Применить выбранный метод и интерпретировать результаты

Ответ:

5.4. Перечень видов оценочных средств и критерии оценивания

Контроль результатов освоения дисциплины является формой управления качеством образования.

Контроль результатов освоения дисциплины осуществляется в виде текущего контроля и промежуточной аттестации (зачета).

Система текущего контроля включает:

1. контроль работы обучающихся на лекционных и практических занятиях;
2. контроль участия в совместной работе группы;
3. контроль выполнения индивидуальных заданий;
4. контроль выполнения обучающимися заданий для самостоятельной работы.

Промежуточная аттестация осуществляется в виде зачета, который проводится в устной либо письменной форме.

Задачи промежуточной аттестации:

1. определить общий уровень освоения дисциплины в целом;
2. определить уровень основных знаний по каждой теме курса;
3. объективизировать результат контроля, минимизировать возможную субъективность преподавателя.

Критерии и рубрики оценивания Элемент контроля/Критерий/Индикаторы/Уровни (0–5)/Комментарий к уровню «5»

Практическое задание № 1 «Анализ характеристик больших данных» (к теме 1.1)

Элемент контроля: Понимание концепции больших данных и их характеристик / ПК-4.2, ПК-4.5 / 0 / 3 / 5 / Дано четкое определение больших данных; каждая характеристика описана с примерами; для 3 бизнес-задач обоснованно определены ключевые характеристики; презентация структурирована, содержит примеры применения в выбранной отрасли; выводы аргументированы.

Практическое задание № 2 «Сравнительный анализ архитектур Lambda и Kappa» (к теме 1.2)

Элемент контроля: Освоение сравнительного анализа архитектурных подходов к обработке больших данных / ПК-4.3 / 0 / 3 / 5 / Сравнительная таблица заполнена по всем критериям; для каждого сценария обоснованно выбрана архитектура; схема Big Data-системы содержит все необходимые компоненты; описаны потоки данных; выводы аргументированы.

Практическое задание № 3 «Сбор данных из гетерогенных источников» (к теме 1.3)

Элемент контроля: Освоение методов сбора данных из гетерогенных источников / ПК-4.2, ПК-4.3 / 0 / 3 / 5 / Описаны все методы сбора с преимуществами и ограничениями; данные получены через API и сохранены корректно; веб-скрапинг выполнен с соблюдением правил; проведён анализ лог-файла; выполнено сравнение и оценка качества данных с предложением способов интеграции; код документирован.

Практическое задание № 4 «Работа с распределённой файловой системой» (к теме 1.4)

Элемент контроля: Освоение работы с распределённой файловой системой / ПК-4.3 / 0 / 3 / 5 / Архитектура HDFS описана полно; выполнены основные операции с HDFS; сравнительный анализ с традиционной файловой системой содержит обоснованные выводы; описание использования HDFS для Data Lake структурировано; отчёт содержит все необходимые разделы.

Практическое задание № 5 «Выбор и работа с NoSQL-базой данных» (к теме 1.5)

Элемент контроля: Освоение выбора и работы с NoSQL-базами данных / ПК-4.3, ПК-4.5 / 0 / 3 / 5 / Классификация NoSQL заполнена с обоснованием; теорема CAP объяснена; выполнены базовые операции в NoSQL-базе; проведён сравнительный анализ с реляционной СУБД; для заданного сценария обоснованно выбрана NoSQL-база; отчёт структурирован.

Практическое задание № 6 «Разработка ETL-процесса» (к теме 1.6)

Элемент контроля: Освоение разработки ETL-процессов для больших данных / ПК-4.3 / 0 / 3 / 5 / ETL-процесс реализован с выполнением всех этапов; применены методы трансформации; оценка качества данных выполнена до и после трансформации; предложено использование ETL для витрины данных; код документирован; выводы обоснованы.

Практическое задание № 7 «Предобработка данных для анализа» (к теме 1.7)

Элемент контроля: Освоение методов предобработки данных / ПК-4.3 / 0 / 3 / 5 / Изучены методы обработки пропусков и выбросов; выполнена предобработка с обоснованием выбора методов; нормализация и стандартизация применены корректно; сравнение методов выполнено с выводами; отчёт структурирован.

Практическое задание № 8 «Реализация MapReduce-задачи» (к теме 1.8)

Элемент контроля: Освоение модели программирования MapReduce для параллельной обработки данных / ПК-4.3 / 0 / 3 / 5 / MapReduce-задача реализована корректно; запущена на Hadoop; проведён анализ результатов; сравнение производительности с другими подходами выполнено; ограничения MapReduce описаны; код документирован.

Практическое задание № 9 «Разработка приложения на Apache Spark» (к теме 1.9)

Элемент контроля: Освоение разработки приложений на Apache Spark / ПК-4.3, ПК-4.5 / 0 / 3 / 5 / Архитектура Spark изучена; приложение разработано с выполнением операций над DataFrame; Spark SQL применён; выполнена оптимизация запросов; сравнение с MapReduce выполнено; код документирован; выводы обоснованы.

Практическое задание № 10 «Работа с SQL-движками для больших данных» (к теме 1.10)

Элемент контроля: Освоение работы с SQL-движками для больших данных / ПК-4.3 / 0 / 3 / 5 / Сравнительная таблица Hive

и Presto/Trino заполнена; SQL-запросы выполнены; сравнительный анализ производительности проведён; предложены сценарии использования; выводы аргументированы.
Практическое задание № 11 «Обработка потоковых данных в реальном времени» (к теме 1.11)
Элемент контроля: Освоение обработки потоковых данных с использованием Spark Streaming / ПК-4.2, ПК-4.5 / 0 / 3 / 5 / Концепция потоковой обработки изучена; приложение реализовано; проведён анализ результатов с выводами о динамике показателей; описано бизнес-применение потоковой обработки; код документирован.
Практическое задание № 12 «Применение машинного обучения на больших данных» (к теме 1.12)
Элемент контроля: Применение машинного обучения на больших данных с использованием Spark MLlib / ПК-4.3, ПК-4.5 / 0 / 3 / 5 / Предобработка данных выполнена с использованием Spark; модель построена и обучена; качество модели оценено по метрикам; проведён сравнительный анализ производительности на малых и больших данных; результаты интерпретированы для бизнеса; код документирован.
Практическое задание № 13 «Статистический анализ больших данных» (к теме 1.13)
Элемент контроля: Освоение методов прикладной статистики для анализа больших данных / ПК-4.5 / 0 / 3 / 5 / Методы изучены; выполнены регрессионный анализ, кластеризация и прогнозирование; результаты интерпретированы; сформулированы выводы для руководителей; отчёт структурирован; выводы аргументированы.
Практическое задание № 14 «Визуализация больших данных» (к теме 1.14)
Элемент контроля: Освоение визуализации результатов анализа больших данных / ПК-4.5 / 0 / 3 / 5 / Принципы визуализации изучены; создан дашборд для визуализации ключевых показателей; добавлены интерактивные элементы; презентация структурирована; выводы аргументированы.
Практическое задание № 15 «Построение витрины данных и Data Lake» (к теме 1.15)
Элемент контроля: Освоение формирования витрин данных и Data Lake для поддержки управленческих решений / ПК-4.2, ПК-4.3, ПК-4.5 / 0 / 3 / 5 / Архитектура Data Lake и витрины данных разработана; источники данных и процесс загрузки описаны; ETL-процесс реализован; отчёт структурирован; выводы обоснованы; код документирован.
Тест
Корректность ответов на вопросы: $\geq 90\%$ правильных ответов — 5; /75–89% — 4; /60–74% — 3; /< 60% — 2.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

- Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2024. — 121 с. — (Высшее образование). — ISBN 978-5-534-09837-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/539330>.
- Маркин, А. В. Системы графовых баз данных. Neo4j : учебник для вузов / А. В. Маркин. — Москва : Издательство Юрайт, 2025. — 303 с. — (Высшее образование). — ISBN 978-5-534-13996-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/567732>.
- Миркин, Б. Г. Базовые методы анализа данных : учебник и практикум для вузов / Б. Г. Миркин. — 3-е изд., пер. и доп. — Москва : Издательство Юрайт, 2025. — 297 с. — (Высшее образование). — ISBN 978-5-534-19709-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/560414>.
- Мхитарян, В. С. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2024. — 448 с. — (Высшее образование). — ISBN 978-5-534-19964-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/557384>.
- Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебное пособие для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2024. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/545036>.
- Стефанова, Н. Л. Методы математической обработки данных : учебник и практикум для вузов / Н. Л. Стефанова [и др.] ; под общей редакцией Н. Л. Стефановой. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 317 с. — (Высшее образование). — ISBN 978-5-534-18254-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/534612>.
- Нестеров, С. А. Базы данных : учебник и практикум для вузов / С. А. Нестеров. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 258 с. — (Высшее образование). — ISBN 978-5-534-18107-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/536687>.
- Мойзес, Б. Б. Статистические методы контроля качества и обработка экспериментальных данных : учебное пособие для вузов / Б. Б. Мойзес, И. В. Плотнокова, Л. А. Редько. — 2-е изд. — Москва : Издательство Юрайт, 2024. — 118 с. — (Высшее образование). — ISBN 978-5-534-11906-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/542709>.

6.2. Электронные учебные издания и электронные образовательные ресурсы

Э1	Сетевые ресурсы Государственного автономного учреждения культуры Свердловской области «Свердловская областная универсальная научная библиотека им. В.Г. Белинского». Режим доступа: https://book.uraic.ru/el_library/resursy_dostup
Э2	Образовательная платформа Юрайт. Режим доступа: https://urait.ru/
Э3	Научная библиотека Cyberleninka. Режим доступа: https://cyberleninka.ru/

6.3.1 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства

6.3.1.1	Операционная система Astra Linux
6.3.1.2	Офисная система LibreOffice

6.3.1.3	Платформа для анализа больших данных с визуальным интерфейсом и поддержкой машинного обучения (KNIME, лицензия GPL); СУБД с массово-параллельной обработкой для крупномасштабной аналитики (Greenplum Community Edition, бесплатная версия); библиотека аналитических алгоритмов с открытым исходным кодом (MADlib, лицензия Apache 2.0); среда для статистического анализа данных (R, лицензия GNU GPL); средства для визуализации данных (matplotlib, seaborn, лицензия BSD); инструменты для работы с распределёнными вычислениями и кластерными системами; офисный пакет для оформления результатов анализа
6.3.1.4	Контур.Класс
6.3.1.5	Контур.Толк
6.3.1.6	ИС:Университет ПРОФ
6.3.1.7	Среда электронного обучения 3LK (русский Moodle)

7. МТО (оборудование и технические средства обучения)	
7.1	Учебные аудитории для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенные оборудованием и техническими средствами обучения:
7.2	Адрес (местоположение) учебных аудиторий Перечень оборудования и технических средств обучения
7.3	620014, Свердловская область, г.о. город Екатеринбург, г Екатеринбург, пр-кт Ленина, стр. 20а, этаж № 5, кабинет № 3 площадью 66,2 м² в помещении 501 Стол – 13 шт.
7.4	Стул – 25 шт.
7.5	ПК (с возможностью подключения к сети «Интернет») – 1 шт.
7.6	Телевизор (75”) – 1 шт.
7.7	Беспроводная система презентаций – 1 шт.
7.8	Спикерфон – 1 шт.
7.9	ПК (ноутбук) – 12 шт.
7.10	620014, Свердловская область, г.о. город Екатеринбург, г Екатеринбург, пр-кт Ленина, стр. 20а, этаж № 5, кабинет № 19 площадью 31,7 м² в помещении 501 Стол – 7 шт.
7.11	Стул – 13 шт.
7.12	ПК (с возможностью подключения к сети «Интернет») – 1 шт.
7.13	Магнитно-маркерная доска – 1 шт.
7.14	Телевизор (75”) – 1 шт.
7.15	620014, Свердловская область, г.о. город Екатеринбург, г Екатеринбург, пр-кт Ленина, стр. 20а, этаж № 5, кабинет № 21 площадью 35,6 м² в помещении 501 Стол – 7 шт.
7.16	Стул – 13 шт.
7.17	ПК (с возможностью подключения к сети «Интернет») – 1 шт.
7.18	Магнитно-маркерная доска – 1 шт.
7.19	Телевизор (75”) – 1 шт.
7.20	
7.21	Помещение для самостоятельной работы обучающихся, оснащенное компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду университета:
7.22	Адрес (местоположение) помещения для самостоятельной работы обучающихся Перечень компьютерной техники и оборудования
7.23	620014, Свердловская область, г.о. город Екатеринбург, г Екатеринбург, пр-кт Ленина, стр. 20а, этаж № 5, кабинет № 5 площадью 44,3 м² в помещении 501 ПК (ноутбук) – 9 шт.
7.24	Стол – 11 шт.
7.25	Стул – 11 шт.

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для успешного освоения дисциплины рекомендуется ознакомиться с рабочей программой, целями и задачами дисциплины, её связями с другими дисциплинами образовательной программы. Также рекомендуется изучить основную и дополнительную литературу, перечень заданий. Перед лекцией просматривать конспект предыдущей лекции для более глубокого восприятия материала. При подготовке к практическим занятиям изучить основную литературу, ознакомиться с дополнительной и методической литературой, учесть рекомендации преподавателя. Важно также опираться на конспекты лекций. В ходе занятий вести конспектирование учебного материала, обращать внимание на категории, формулировки, раскрывающие содержание тех или иных понятий, явлений и процессов, научные выводы и практические рекомендации по их применению. При пропуске занятия самостоятельно изучить материал и предоставить преподавателю отчёт по пропущенной теме в часы индивидуальных консультаций. При подготовке к текущей аттестации тщательно изучать материал по блокам тем, акцентировать внимание на определениях, терминах, содержании понятий. Для этого рекомендуется использовать основную литературу, а также конспекты лекций, составленные в ходе изучения всей дисциплины. При подготовке к промежуточной аттестации в дополнение к изучению конспектов лекций, словарей, учебных пособий, целесообразно пользоваться учебной литературой, рекомендованной в программе, и электронными ресурсами. При подготовке к зачёту или экзамену организовать учёбу так, чтобы все виды работ и заданий, предусмотренные рабочей программой, были выполнены в срок. Основное в подготовке — это повторение всего материала учебной дисциплины.