

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Петровская Анна Викторовна

Должность: Директор

Дата подписания: 12.09.2025 09:38:42

Уникальный программный ключ:

798bda6555fbdebe827768f6f1710bd17a9070c31fdc1b6a6ac5a1f10c8c5199

*Приложение 3 к основной профессиональной образовательной программе  
по направлению подготовки 38.03.01 Экономика  
направленность (профиль) программы  
«Экономика и управление на предприятии (организации)»*

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное бюджетное образовательное учреждение высшего образования  
«Российский экономический университет имени Г. В. Плеханова»  
Краснодарский филиал РЭУ им. Г.В. Плеханова

Факультет экономики, менеджмента и торговли

Кафедра экономики и цифровых технологий

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **Б1.О.ДЭ.02.01 ОСНОВЫ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ**

**Направления подготовки - 38.03.01 Экономика**

**Направленность (профиль) программы «Экономика и управление  
на предприятии (организации)»**

**Уровень высшего образования - Бакалавриат**

**Год начала подготовки 2025**

**Краснодар – 2024 г.**

Составитель:

к.т.н., доцент кафедры экономики и цифровых технологий Р.Н. Фролов

Рабочая программа одобрена на заседании кафедры экономики и цифровых технологий, протокол № 9 от 14.03.2024 г.

# СОДЕРЖАНИЕ

<b>I. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ РАЗДЕЛ.....</b>	<b>4</b>
Цель и задачи освоения дисциплины .....	4
Место дисциплины в структуре образовательной программы .....	4
Объем дисциплины и виды учебной работы.....	4
Перечень планируемых результатов обучения по дисциплине .....	5
<b>II. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....</b>	<b>7</b>
<b>III. УЧЕБНО-МЕТОДИЧЕСКОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ</b> <b>.....</b>	<b>12</b>
РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА .....	12
ПЕРЕЧЕНЬ ИНФОРМАЦИОННО-СПРАВОЧНЫХ СИСТЕМ.....	13
ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ .....	13
ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	13
ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ.....	13
МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ .....	13
<b>IV. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ .....</b>	<b>14</b>
<b>V. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ И УМЕНИЙ, ХАРАКТЕРИЗУЮЩИХ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ .....</b>	<b>14</b>
<b>VI. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ .....</b>	<b>15</b>
<b>АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ.....</b>	<b>25</b>

# І. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ РАЗДЕЛ

## Цель и задачи освоения дисциплины

**Цель дисциплины** заключается в решении теоретико-методологических знаний и закрепление профессиональных навыков в области анализа больших данных (Big Data) с учетом зарубежного и российского опыта, а также формирование современного мышления, приобретение знаний и компетенций, позволяющих повышать эффективность принятия решений за счет грамотного сбора, структурирования и применения современных техник анализа больших объемов разнородных данных

### Задачи дисциплины:

- изучить типовые задачи, для решения которых применяется машинное обучение: анализ трендов, социальных сетей, распознавание графических, видео- и аудио-образов, текста, прогнозирование действий;
- изучить методы обработки и анализа данных: регрессия, нейронные сети; графы;
- изучить основные инструменты анализа данных, в том числе методы интеллектуального анализа данных и машинного обучения;
- изучить основные принципы поиска, сбора, очистки, хранения, анализа и обработки статистических данных с целью их визуализации;
- изучить основы обработки экономических и финансовых данных с использованием информационных технологий для решения профессиональных задач на основе функционирования систем искусственного интеллекта.

## 2.Содержание дисциплины:

### Место дисциплины в структуре образовательной программы

Дисциплина «Основы работы с Большими данными», относится к обязательной части учебного плана.

## Объем дисциплины и виды учебной работы

Таблица 1

Показатели объема дисциплины *	Всего часов по формам обучения	
	очная	очно-заочная
Объем дисциплины в зачетных единицах	3 ЗЕТ	
Объем дисциплины в акад. часах	108	
Промежуточная аттестация: форма	зачет	зачет
<b>Контактная работа обучающихся с преподавателем (Контакт. часы), всего:</b>	<b>30</b>	<b>14</b>
1. Контактная работа на проведение занятий лекционного и семинарского типов, всего	<b>28</b>	<b>12</b>

часов, в том числе:		
• лекции	<b>12</b>	<b>6</b>
• практические занятия	<b>16</b>	<b>6</b>
• лабораторные занятия	-	-
в том числе практическая подготовка	-	-
2. Индивидуальные консультации (ИК)	-	-
3. Контактная работа по промежуточной аттестации (Катт)	<b>2</b>	<b>2</b>
4. Консультация перед экзаменом (КЭ)		
5. Контактная работа по промежуточной аттестации в период экз. сессии / сессии заочников (Каттэк)		
<b>Самостоятельная работа (СР), всего:</b>	<b>78</b>	<b>94</b>
в том числе:		
• самостоятельная работа в период экз. сессии (СРэк)		
• самостоятельная работа в семестре(СРс)	<b>78</b>	<b>94</b>

## Перечень планируемых результатов обучения по дисциплине

Таблица 2

<b>Формируемые компетенции (код и наименование компетенции)</b>	<b>Индикаторы достижения компетенций (код и наименование индикатора)</b>	<b>Результаты обучения (знания, умения)</b>
<b>УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</b>	<b>УК-1.1.</b> Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>УК-1.1. 3-1. Знает</b> основные методы критического анализа и основы системного подхода как общенаучного метода.
		<b>УК-1.1. У-1. Умеет</b> анализировать задачу, используя основы критического анализа и системного подхода.
		<b>УК-1.1. У-2. Умеет</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.
<b>ОПК-2. Способен осуществлять сбор, обработку и статистический анализ данных, необходимых для решения поставленных экономических задач.</b>	<b>ОПК-2.1.</b> Использует основные методы, средства получения, представления, хранения и обработки статистических данных.	<b>ОПК-2.1. 3-1. Знает</b> методы поиска и систематизации информации об экономических процессах и явлениях.
		<b>ОПК-2.1. У-1. Умеет</b> работать с национальными и международными базами данных с целью поиска информации, необходимой для решения поставленных экономических задач.

		<p><b>ОПК-2.1.У-2. Умеет</b> рассчитывать экономические и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов на основе типовых методик и действующей нормативно-правовой базы</p>
		<p><b>ОПК-2.1.У-3. Умеет</b> представить наглядную визуализацию данных.</p>
<p><b>ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</b></p>	<p>ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы</p>	<p><b>ОПК-6.1. 3-1. Знать:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p>
		<p><b>ОПК-6.1. У-1. Уметь:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности</p>
	<p>ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач</p>	<p><b>ОПК-6.2. 3-1. Знать:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p>
		<p><b>ОПК-6.2.У-1. Умеет</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p>

## II. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### Этапы формирования и критерии оценивания сформированности компетенций для обучающихся очной формы обучения

Таблица 3. 1

№ п/п	Наименование раздела, темы дисциплины	Трудоемкость, академические часы					Индикаторы достижения компетенций	Результаты обучения (знания, умения)	Учебные задания для аудиторных занятий	Текущий контроль	Задания для творческого рейтинга (по теме(-ам)/ разделу или по всему курсу в целом)	
		Лекции	Практические занятия	Лабораторные занятия	Практическая подготовка	Самостоятельная работа/ КЭ, Катгэк, Катг						Всего
<b>Семестр 4</b>												
Раздел 1. Большие данные: современные подходы к обработке и хранению												
1.	<b>Тема 1. Основы больших данных</b> Понятие и характеристики больших данных: современные подходы к обработке и хранению. Терминология, история появления. Технические сложности работы с большими данными. Влияние больших данных в решении аналитических и исследовательских задач профессиональной деятельности. Источники больших данных. Платформа больших данных. Наука о данных. Процесс изучения данных. Методы исследования больших данных.	2	4			18	24	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2. У-1.	Гр.д.	-	Д
2.	<b>Тема 2. Обработка больших данных</b> Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представление наборов данных.	4	4			18	26	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2. У-1.	Гр.д.	К.	Д.

									ОПК-6.2.У-1			
<b>Раздел 2. Программное обеспечение в области анализа больших данных</b>												
3.	<b>Тема 3. Анализ и извлечение данных</b> Аналитические платформы: классификация и особенности применения. Языки визуального моделирования. Реляционные базы данных. Распределенные СУБД. Технологии GOOGLE BIGTABLE. Полнотекстовый поиск. Параллельные запросы. Технология поиска и интеграция. Процесс обработки отчетов. Экосистемы больших данных и его компоненты. Облачные решения BigData. Интеграция данных с использованием технологий BigData; построение и загрузка озер и хранилищ данных; обеспечение качества данных.	2	4			21	27	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2.У-1	Гр.д.	К.	Д
4.	<b>Тема 4. Технология Data Mining</b> Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining. Подготовка данных к анализу. Методика извлечения знаний. Актуальность технологий Data Mining как средств обработки больших объемов информации. Ассоциативные правила. Поиск и достоверность ассоциативных правил. Алгоритм генерации ассоциативных правил. Определение кластеризации. Цели кластеризации в Data Mining. Применение классификации и регрессии. Статистические методы. Основные понятия нейронных сетей. Определение и структура дерева решений. Выбор атрибута и разбиение в узле.	4	4			21	29	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2.У-1	Гр.д.	К.р., Т	Д
	<i>Контактная работа по промежуточной аттестации (Камт)</i>	-	-	-	-	-/2	2					
	<i>Самостоятельная работа в период экз. сессии (СРэк)</i>	-	-	-	-	-	-					
	<b>Итого</b>	<b>12</b>	<b>16</b>	-	-	<b>78/2</b>	<b>108</b>	<b>х</b>	<b>х</b>	<b>х</b>	<b>х</b>	<b>х</b>

**Этапы формирования и критерии оценивания сформированности компетенций  
для обучающихся очно - заочной формы обучения**

Таблица 3.2

№ п/п	Наименование раздела, темы дисциплины	Трудоемкость*, академические часы					Индикаторы достижения компетенций	Результаты обучения** (знания, умения)	Учебные задания для аудиторных занятий	Текущий контроль	Задания для творческого рейтинга (по теме(-ам))/ разделу или по всему курсу в целом)	
		Лекции	Практические занятия	Лабораторные занятия	Практическая подготовка	Самостоятельная работа/ КЭ, Каттэк, Катт						Всего
<b>Семестр 4</b>												
Раздел 1. Большие данные: современные подходы к обработке и хранению												
1.	<b>Тема 1. Основы больших данных</b> Понятие и характеристики больших данных: современные подходы к обработке и хранению. Терминология, история появления. Технические сложности работы с большими данными. Влияние больших данных. Роль и место больших данных в решении аналитических и исследовательских задач профессиональной деятельности. Источники больших данных. Платформа больших данных. Наука о данных. Процесс изучения данных. Методы исследования больших данных.	1	1			22	24	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2.У-1	Гр.д.	-	-
2.	<b>Тема 2. Обработка больших данных</b> Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представление наборов данных.	1	1			24	26	УК-1.1. ОПК-2.1.. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1.	Гр.д.	К.	-

									ОПК-6.2. 3-1 ОПК-6.2.У-1			
<b>Раздел 2. Программное обеспечение в области анализа больших данных</b>												
3.	<b>Тема 3. Анализ и извлечение данных</b> Аналитические платформы: классификация и особенности применения. Языки визуального моделирования. Реляционные базы данных. Распределенные СУБД. Технологии GOOGLE BIGTABLE. Полнотекстовый поиск. Параллельные запросы. Технология поиска и интеграция. Процесс обработки отчетов. Экосистемы больших данных и его компоненты. Облачные решения BigData. Интеграция данных с использованием технологий BigData; построение и загрузка озер и хранилищ данных; обеспечение качества данных.	2	2			23	27	УК-1.2. ОПК-2.1. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2.У-1	Гр.д.	К.	-
4.	<b>Тема 4. Технология KDD и Data Mining</b> Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining. Подготовка данных к анализу. Методика извлечения знаний. Актуальность технологий Data Mining как средств обработки больших объемов информации. Ассоциативные правила. Поиск и достоверность ассоциативных правил. Алгоритм генерации ассоциативных правил. Определение кластеризации. Цели кластеризации в Data Mining. Применение классификации и регрессии. Статистические методы. Основные понятия нейронных сетей. Определение и структура дерева решений. Выбор атрибута и разбиение в узле.	2	2			25	29	УК-1.1. ОПК-2.1. ОПК-6.1 ОПК-6.2.	УК-1.1. 3-1 УК-1.1. У-1. УК-1.1. У-2. ОПК-2.1. 3-1. ОПК-2.1. У-1 ОПК-2.1. У-2. ОПК-2.1. У-3. ОПК-6.1. 3-1. ОПК-6.1. У-1. ОПК-6.2. 3-1 ОПК-6.2.У-1	Гр.д.	К.р., Т	-
	<i>Контактная работа по промежуточной аттестации (Катт)</i>	-	-	-	-	-/2	2					
	<i>Самостоятельная работа в период экз. сессии (СРэк)</i>	-	-	-	-	-	-					
	<b>Итого</b>	<b>6</b>	<b>6</b>	-	-	<b>94/2</b>	<b>108</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>

**Формы учебных заданий на аудиторных занятиях:**

*Групповая дискуссия (Гр.д.)*

**Формы текущего контроля:**

*Контрольная работа (К.р.)*

*Тест (Т.)*

*Кейс (К.)*

**Формы заданий для творческого рейтинга:**

*Доклад (Д)*

### III. УЧЕБНО-МЕТОДИЧЕСКОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

#### РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

##### Основная литература

1. Грибанов, Ю. И. Цифровая трансформация бизнеса: учебное пособие / Ю. И. Грибанов, М. Н. Руденко. - 2-е изд. - Москва : Дашков и К, 2025. - 213 с. - ISBN 978-5-394-04192-1. - Текст: электронный. - Режим доступа <https://znanium.com/read?id=371213>
2. Ильин, В. В. Цифровая экономика: практическая реализация : методическое пособие / В. В. Ильин. - Москва : Агентство электронных изданий «Интермедиатор», 2024. - 202 с. - ISBN 978-5-91349-074-2. - Текст: электронный. - Режим доступа: <https://znanium.com/read?id=358892>
3. Эзрахи, А. Виртуальная конкуренция: посулы и опасности алгоритмической экономики : учебник / Ариэль Эзрахи, Морис Стаки ; пер. с англ. А. А. Резвова. - Москва : Дело (РАНХиГС), 2023. - 384 с. - (Академическая книга). - ISBN 978-5-85006-341-2. - Текст : электронный. - URL: <https://znanium.com/read?id=393902>

##### Дополнительная литература:

1. Салдана, Т. Почему цифровая трансформация не дает результата и что делать, чтобы всё заработало: практическое руководство / Т. Салдана. - Москва : Альпина Паблицер, 2024. - 334 с. - ISBN 978-5-9614-3859-8. - Текст: электронный. - Режим доступа URL: <https://znanium.com/read?id=387180>
2. Дейтел, П. Python: Искусственный интеллект, большие данные и облачные вычисления : практическое руководство / П. Дейтел, Х. Дейтел. - Санкт-Петербург: Питер, 2023. - 864 с. - (Серия «Для профессионалов»). - ISBN 978-5-4461-1432-0. - Текст : электронный. - URL: <https://znanium.com/read?id=378508>
3. Братко, А. Г. Искусственный разум, правовая система и функции государства: монография / А.Г. Братко. — Москва: ИНФРА-М, 2022. — 282 с. — (Научная мысль). — DOI 10.12737/1064996. - ISBN 978-5-16-015890-7. - Текст: электронный. - URL: <https://znanium.com/read?id=390092>

##### Нормативные правовые документы:

1. Федеральный закон от 31 июля 2020 г. № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации» [Электрон.ресурс]. — Режим доступа [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_358738/](http://www.consultant.ru/document/cons_doc_LAW_358738/)
2. "Стратегия развития информационного общества в Российской Федерации на 2017-2030 годы" [Электрон.ресурс]. — Режим доступа [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_216363/](http://www.consultant.ru/document/cons_doc_LAW_216363/)

## **ПЕРЕЧЕНЬ ИНФОРМАЦИОННО-СПРАВОЧНЫХ СИСТЕМ**

1. <http://www.consultant.ru> -Справочно-правовая система Консультант Плюс;
2. <http://www.garant.ru>- Справочно-правовая система Гарант.

## **ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ**

1. <http://www.iep.ru/ru/publikacii/categories.html> Федеральный образовательный портал. Экономика. Социология. Менеджмент
2. <https://rosmintrud.ru/opendata> - База открытых данных Минтруда России
3. <http://www.fedsfm.ru/opendata> - База открытых данных Росфинмониторинга
4. <https://polpred.com/> Электронная база данных "Polpred.com Обзор СМИ"

## **ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ**

1. <https://digital.gov.ru/ru/> - информационный ресурс Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации
2. <http://citforum.ru/>-«Сервер информационных технологий» - on-line библиотека информационных материалов по компьютерным технологиям.
3. <http://www.intuit.ru/>-Образовательный портал дистанционного обучения.
4. [www.coursera.org](http://www.coursera.org/)-Платформа для бесплатных онлайн-лекций (проект по публикации образовательных материалов в интернете, в виде набора бесплатных онлайн-курсов).

## **ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ**

Операционная система Windows 10, Microsoft Office Professional Plus: 2019 год (MS Word, MS Excel, MS Power Point, MS Access)

Антивирус Dr.Web Desktop Security Suite Комплексная защита

Браузер Google Chrome

Adobe Premiere

Power DVD

MediaPlayerClassic

## **МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

Дисциплина «Основы работы с Большими данными» обеспечена:

для проведения занятий лекционного типа:

- учебной аудиторией, оборудованной учебной мебелью, мультимедийными средствами обучения для демонстрации лекций-презентаций;

для проведения занятий семинарского типа (практические занятия);

- компьютерным классом;

- помещением для самостоятельной работы, оснащенным компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа электронной информационно-образовательной среде университета.

#### **IV. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

- Методические рекомендации по организации и выполнению внеаудиторной самостоятельной работы.
- Методические указания по выполнению практических работ.

#### **V. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ И УМЕНИЙ, ХАРАКТЕРИЗУЮЩИХ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ**

Результаты текущего контроля и промежуточной аттестации формируют рейтинговую оценку работы обучающегося. Распределение баллов при формировании рейтинговой оценки работы обучающегося осуществляется в соответствии с «Положением о рейтинговой системе оценки успеваемости и качества знаний студентов в процессе освоения дисциплины «Основы работы с Большими данными» в федеральном государственном бюджетном образовательном учреждении высшего образования «Российский экономический университет имени Г.В. Плеханова».

Таблица 4

<b>Виды работ</b>	<b>Максимальное количество баллов</b>
Выполнение учебных заданий на аудиторных занятиях	20
Текущий контроль	20
Творческий рейтинг	20
Промежуточная аттестация ( <i>зачет</i> )	40
<b>ИТОГО</b>	<b>100</b>

В соответствии с Положением о рейтинговой системе оценки успеваемости и качества знаний обучающихся «преподаватель кафедры, непосредственно ведущий занятия со студенческой группой, обязан проинформировать группу о распределении рейтинговых баллов по всем видам работ на первом занятии учебного модуля (семестра), количестве модулей по учебной дисциплине, сроках и формах контроля их освоения, форме промежуточной аттестации, снижении баллов за несвоевременное выполнение выданных заданий. Обучающиеся в течение учебного модуля (семестра) получают информацию о текущем количестве набранных по дисциплине баллов через личный кабинет студента».

## **VI. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ**

Оценочные материалы по дисциплине разработаны в соответствии с Положением об оценочных материалах в федеральном государственном бюджетном образовательном учреждении высшего образования «Российский экономический университет имени Г.В. Плеханова», протокол №1 от 26.09.2022г.

### ***Тематика курсовых работ/проектов***

«Курсовая работа/проект по дисциплине «Основы работы с Большими данными» учебным планом не предусмотрена.

### ***Перечень вопросов к зачету:***

1. Какие данные называют «большими»?
2. Почему большие данные нецелесообразно хранить в обычной реляционной базе данных?
3. Какой момент считается точкой отсчета науки о больших данных?
4. Укажите порядки объемов данных, обычно называемых «большими».
5. Приведите пример неочевидной полезной закономерности, полученной с помощью методов анализа больших данных.
6. Приведите примеры источников больших данных в области охраны общественного порядка.
7. Как анализ больших данных может быть использован для решения задач биржевой торговли?
8. Как анализ больших данных используется в современной медицине?
9. Приведите пример геоинформационных систем, работа которых существенно зависит от эффективности алгоритмов анализа больших данных?
10. Что представляет из себя интернет вещей и как он связан с методами анализа больших данных?
11. Создание «прозрачной» информации посредством анализа экономических данных большого объема.
12. Методы анализа больших данных в задачах принятия математически обоснованных управленческих решений.
13. Узкое сегментирование клиентов с учетом персональных пожеланий на основе технологий обработки больших данных.
14. Увеличение скорости в принятии решений за счет сложной аналитики больших данных в режиме реального времени.
15. Аналитика больших данных в задачах развития товаров и услуг следующего поколения

16. Каким образом осуществляется оцифровка текста на естественном языке?
17. Какие программные продукты могут использоваться для семантического анализа текстов на естественных языках?
18. Перечислите основные принципы работы современных систем машинного перевода текстов на естественных языках.
19. Перечислите основные базы знаний Wolfram Data Platform.
20. Что представляет из себя формат CDF?
21. Что такое канонический идентификатор WDF?
22. Что такое хеширование?
23. Дайте определение Big Data и охарактеризуйте основные признаки больших данных (4V).
24. В чем заключаются основные отличия традиционных и больших данных?
25. Опишите архитектуру и принципы работы распределенных систем хранения данных.
26. Каковы основные проблемы при работе с большими данными?
27. Опишите архитектуру Hadoop и назначение его основных компонентов.
28. В чем отличие Hadoop от Spark? Когда целесообразно использовать каждую из этих технологий?
29. Объясните принцип работы MapReduce. Приведите примеры задач, где он применяется.
30. Какие существуют типы NoSQL баз данных? Приведите примеры и случаи их использования.
31. Опишите особенности работы с потоковыми данными. Какие инструменты для этого используются?
32. Какие методы предварительной обработки больших данных вы знаете?
33. Опишите основные этапы работы с большими данными (ETL-процесс).
34. Какие алгоритмы машинного обучения наиболее эффективны для работы с большими данными?
35. Как осуществляется визуализация больших данных? Назовите популярные инструменты.
36. Приведите примеры практического применения Big Data в различных отраслях.
37. Каковы основные проблемы безопасности при работе с большими данными?
38. Опишите этические аспекты сбора и использования больших данных.
39. Каковы современные тенденции развития технологий Big Data?
40. В чем особенности работы с данными интернета вещей (IoT)?
41. Сравните подходы Data Lake и Data Warehouse. Когда какой подход предпочтительнее?

42. В чем преимущества облачных платформ для работы с большими данными по сравнению с локальными решениями?
43. Опишите кейсы успешного внедрения Big Data-решений в бизнесе.
44. Сущность больших данных и перспективы их использования в различных сферах.
45. Условия и возможности использования больших данных в различных сферах.
46. Примеры использования больших данных в зарубежной практике.
47. Примеры использования больших данных на национальном уровне.
48. Роль цифровой информации в современных условиях.
49. Источники получения больших данных.
50. Виды и формы данных.
51. Технологии обработки больших данных: NoSQL
52. Примеры и инструменты для визуализации.
53. Data Mining. Постановка основных задач.
54. Машинное обучение.
55. Бизнес-решения с помощью алгоритмов Data Mining.
56. Экосистема Hadoop в анализе больших данных.
57. Вычислительная модель MapReduce.
58. В чем суть HDFS?
59. Цели кластеризации.
60. Платформы больших данных

### **Практические задания к зачету:**

#### **Задание №1:**

1. По индивидуальному заданию (скачать интересующий датасет <https://www.kaggle.com/datasets?search=retail>) провести быструю аналитику в Excel / Google-таблицах.
2. Найдите закономерности в данных с помощью автоматически формируемых диаграмм.
3. Создайте сводные таблицы на основе своих данных.
4. Визуализируйте подмножества данных, выбирая только нужные столбцы или ячейки.
5. Результат работы (отчет в формате docx/pdf – в документе должны быть представлены описание датасета, ссылки на выполненное задание, скриншоты с пояснениями и выводы).

#### Задание №2:

По индивидуальному заданию (скачать интересующий датасет <https://www.kaggle.com/datasets?search=retail>) провести быструю аналитику в Yandex DataLens .

1. Найдите закономерности в данных. Сделайте выводы. Создайте виджеты (чарты, индикаторы, селекторы, дашборды) на основе своих данных.
2. Настройте публичный доступ к дашборду.

Результат работы (отчет в формате docx – в документе должны быть представлены описание датасета, ссылки на выполненное задание, скриншоты с пояснениями, выводы и заключение).

#### Задание №3:

1. Представить данные в виде многомерных кубов (OLAP-кубов) в low-code платформе анализа данных.
2. Произвести разведочный анализ, исследование данных.
3. Сформировать аналитическую отчетность.

#### Задание №4:

По индивидуальному заданию (скачать интересующий датасет <https://www.kaggle.com/datasets?search=retail> выполнить в Power BI ряд действий:

1. Произведите расширение модели данных использованием таблиц календаря.
2. Создайте выражения KPI с использованием анализа данных (DAX)/языка M Power Query
3. Используйте для моделирования параметр (What If)
4. Разработайте интерактивные отчеты.

#### Задание №5:

По индивидуальному заданию (скачать интересующий датасет: <https://www.kaggle.com/datasets?search=retail>) выполнить в Tableau ряд действий:

1. Постройте дашборды.
2. Выявите инсайты.
3. Оформите истории (Story).
4. Опубликуйте в своем аккаунте в Tableau Public.

#### Задание №6:

Используя методологию DataMining, решите задачу поиск аномалий в данных в объекте исследования «Финансовая компания».

#### Задание №7:

Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Интернет-магазин».

#### Задание №8:

Используя методологию DataMining, решите задачу регрессии данных в объекте исследования «Объект розничной торговли».

#### Задание №9:

Используя методологию DataMining, решите задачу поиск аномалий в данных в объекте исследования «Предприятие оптовой торговли».

Задание №10:

Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Коммерческое предприятие».

Задание №11:

Используя методологию DataMining, решите задачу корреляционного анализа в объекте исследования «Коммерческая фирма».

Задание №12:

Используя методологию DataMining, решите задачу кластеризации данных в объекте исследования «Макроэкономические показатели РФ».

Задание №13:

Используя методологию DataMining, решите задачу бинарной классификации данных в объекте исследования «Агентство недвижимости».

Задание №14:

Используя методологию DataMining, решите задачу нейросетевого моделирования в объекте исследования «Маркетплейс».

Задание №15:

Используя методологию DataMining, решите задачу регрессии данных в объекте исследования «Маркетплейс».

Задание №16:

Используя методологию DataMining, решите задачу поиска ассоциативных правил в объекте исследования «Маркетплейс».

Задание №17:

Используя методологию DataMining, решите задачу кластеризации транзакций в объекте исследования «Маркетплейс».

**Тестовые задания к зачету:**

1. **На проверку и очистку данных как правило тратится до ...% времени проекта.**
2. **Синонимы понятия "наука о данных":**
3. **В Hadoop центральное место занимают:**
4. **Big Data – это ... представление фактов, понятий или инструкций в форме, приемлемой для интерпретации, или обработки;**
  - 1) *комплексный набор методов обработки структурированных и неструктурированных данных колоссальных объемов;*
  - 2) *колоссальный объем данных, собранных человечеством;*
  - 3) *класс в Java, предназначенный для хранения данных от 100 Гб.*

**5. Объём накопленных человечеством цифровых данных на 2012 год измеряется:**

- 1) петабайтами;
- 2) зеттабайтами;
- 3) экзабайтами;
- 4) йоттабайтами.

**6. Связь между ... такая же, как между сырой нефтью и нефтеперерабатывающим заводом.**

Ответ: big data и data science

**7. Наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных**

Ответ:

искусственный интеллект.  
artificial intelligence.

**8. Apache Hadoop - инфраструктура, упрощающая работу с компьютерными кластерами. Hadoop пытается достичь следующих целей (отметьте правильные утверждения)**

*Масштабируемость данные и их обработка распределяются в компьютерных кластерах (горизонтальное масштабирование).*

Портируемость - возможность установки на всех видах устройств и операционных систем.

**9. Преобразование данных ... .**

Ответ:

гарантирует, что данные находятся в подходящем формате для использования в ваших моделях

**10. Фаза моделирования состоит из шагов, расположенных в следующей последовательности:**

Ответ:

1) Планирование показателей и выбор модели. 2) Обучение модели. 3) Проверка адекватности модели и выбор. 4) Применение обученной модели к новым данным.

**11. Очистка данных, интеграция данных, преобразование данных – это подфазы фазы ... .**

Ответ:

Подготовки данных

**12. Фаза подготовки данных состоит из ... подфаз.**

Ответ:

3

**13. Примеры машинных данных:**

Ответ:

записи детализации звонков

журналы веб-серверов

журналы сетевых событий и телеметрии

**14. удаляет некорректные значения из источника данных и устраняет расхождения между источниками.**

Ответ:

Очистка данных

**15. Информация, автоматически генерируемая компьютером, процессом, приложением или устройством без вмешательства человека - ... данные.**

Ответ:

Машинные

**16. Отсутствующие значения можно исправить, используя ... .**

Ответ:

удаление выборки или значения

**17. Искусственный интеллект - ...**

Ответ:

свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека.

**18. Этот важный принцип обеспечивает целостность данных. Запись, вставленная в базу данных, ни при каких условиях не должна конфликтовать с заранее установленными правилами (например, в ней не может отсутствовать обязательное поле или поле не может содержать числовую информацию вместо текстовой) - ... .**

Ответ:

Согласованность

**19. Перечислите ошибки, указывающие на ложные значения в одном наборе данных:**

Ответ:

Лишние пропуски  
Отсутствующие значения  
Ошибки при вводе данных  
Выбросы  
Невозможные значения

**20. Первая проверка данных в процессе data science происходит на стадии:**

Ответ:

сбора данных

**21. Раздел информатики, изучающий проблемы анализа, обработки и представления данных в цифровой форме - ...**

Ответ:

data science.  
datalogy.  
наука о данных.

**22. Изолированность (Isolation) - ...**

Ответ:

когда в базе данных что-то изменяется, ничего не может происходить точно с одними и теми же данными точно в один момент. Вместо этого действия выполняются последовательно с другими изменениями.

**23. Интеграция данных ... .**

Ответ:

расширяет информацию посредством объединения информации из нескольких источников

**24. NoSQL – это**

Ответ:

«не только» SQL

**25. Преобразование данных – это:**

Ответ:

Экстраполяция данных  
Агрегирование данных

Создание вспомогательных переменных  
Сокращение количества переменных  
Производные метрики

**26. Машинные данные ... .**

Ответ:

есть информация, автоматически генерируемая компьютером, процессом, приложением или устройством без вмешательства человека

**27. Вклад специалиста по данным в составление проектного задания:**

Ответ:

Предназначение и контекст проекта.  
Предварительное описание методики анализа.  
Чётко сформулированная цель исследований.  
Предъявляемые результаты и критерий успеха.  
Необходимые ресурсы.  
Календарный план.  
Доказательство практической реализуемости проекта.

**28. Укажите правильное сочетание для очистки данных (обработка отсутствующих значений):**

Ответ:

Присваивание статического значения (например, 0 или среднего арифметического) - простота и предотвращение потери информации от других переменных, но возможно формирование ложных оценок на основе модели.  
Моделирование значения (независимое) - незначительно влияет на модель, но может потребовать слишком большой уверенности в модели и может создать искусственные зависимости между переменными.

**29. Машинное обучение - ...**

Ответ:

класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач.

**30. Выбросы можно исправить, используя ... .**

Ответ:

проверку значения, и в случае ошибочности - интерпретировать как отсутствующее значение (удаление или вставка)

**31. Структурированные; неструктурированные; на естественном языке; машинные; графовые; аудио, видео и графика; потоковые. Это всё категории**

... .

Ответ:

Данных

**32. Процесс построения большинства моделей состоит из следующих шагов (укажите этапы):**

Ответ:

- 1)Выбор метода моделирования и переменных для включения в модель.
- 2)Выполнение модели.
- 3)Диагностика и сравнение моделей.

**33. расширяет информацию посредством объединения информации из нескольких источников.**

Ответ:

Интеграция данных

**34. Под обобщающим термином «...» принято понимать любые наборы данных, достаточно большие и сложные для того, чтобы их можно было обработать традиционными средствами работы с данными.**

Ответ:

большие данные

**35. Наука о данных - ...**

Ответ:

раздел информатики, изучающий проблемы анализа, обработки и представления данных в цифровой форме.

**36. Отличие распределённой файловой системы от «обычной»:**

Ответ:

распределённая работает на нескольких серверах сразу

**37. Характеристики больших данных часто называются «триа ...».**

Ответ: V

**38. Статистическая взаимосвязь двух или более случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми). При этом изменения значений одной или нескольких из этих величин сопутствуют систематическому изменению значений другой или других величин - ...**

Ответ: Корреляция

**39. Униграмма содержит ... слово(а)**

Ответ: 1

**40. Атомарность (Atomicity) - это...**

Ответ:

принцип «всё или ничего». Если блок данных включается в базу, то он либо включается полностью, либо не включается вообще. Например, если в середине операции записи произойдет сбой питания, в базу данных не будет занесена половина данных; они не будут записаны.

**41. Наука о данных объединяет методы:**

Ответ:

по обработке данных в условиях больших объёмов и высокого уровня параллелизма  
интеллектуального анализа данных  
статистические

**42. Разбиение системы на более мелкие структурные компоненты и разнесение их по отдельным физическим машинам (или их группам), и (или) увеличение количества серверов, параллельно выполняющих одну и ту же функцию - ...**

Ответ: горизонтальное масштабирование.

**43. Математическая структура для моделирования попарных отношений между объектами –**

Ответ: Граф

**44. Лишние пропуски можно исправить, используя ... .**

Ответ:

строковые функции

**45. SPARQL – язык запросов, применяемый для работы с ...**

Ответ: графовыми базами данных

**46. Экосистему машинного обучения Python можно условно разделить на пакеты следующих типов:**

Ответ:

Оптимизация кода

Большие данные

Данные помещаются в памяти

**47. Очистка данных – это:**

Ответ:

Отсутствующие значения

Ошибки при вводе данных

Выбросы

Пробелы, опечатки

Физически невозможные данные

Отклонения от свода правил

**48. Класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач - ... .**

Ответ:

машинное обучение

machine learning

**49. Долгосрочность (Durability) - ...**

Ответ:

если данные внесены в базу данных, то они должны находиться в базе данных постоянно. Данные могут уничтожаться при физическом повреждении жёстких дисков, но не при отключении электропитания и программных сбоях.

**50. Характеристики ... часто называются «три V»: объем (volume), разнообразие (variety) и скорость (velocity). Часто эти характеристики дополняются «четвертым V» - достоверностью (veracity):**

Ответ:

больших данных

**51. расширяет информацию посредством объединения информации из нескольких источников.**

Ответ: Интеграция данных

**52. SQL - ...**

Ответ:

основное средство управления и обращения с запросами к данным, хранящимся в реляционных базах данных

**53. Процесс data science обычно состоит из шести шагов:**

Ответ:

- 1) Назначение цели исследования.
- 2) Сбор данных.
- 3) Подготовка данных.
- 4) Исследование данных.
- 5) Моделирование данных.
- 6) Отображение и автоматизация

**54. Разные уровни агрегирования можно исправить, используя ... .**

Ответ:

Переход на единый уровень измерений посредством агрегирования или экстраполяции

**55. Ошибки при вводе данных можно исправить, используя ... .**

Ответ:

ручное переопределение

**56. Большие данные**

Ответ:

обозначение структурированных и неструктурированных данных огромных объёмов и значительного многообразия, эффективно обрабатываемых горизонтально масштабируемыми программными инструментами.

**57. Графовые структуры данных используют ..., ... и ... для представления и хранения графических данных.**

Ответ:

Узлы

Свойства

Ребра

**58. ... развивались на базе статистики и традиционного управления данными, но сейчас считаются разными дисциплинами.**

Ответ:

Big data и data science

**59. Big data- ...**

Ответ:

обозначение структурированных и неструктурированных данных огромных объёмов и значительного многообразия, эффективно обрабатываемых горизонтально масштабируемыми программными инструментами.

**60. Эффективность практического использования технологий больших данных ... .**

Ответ:

определяется предметной областью и людьми, реализующими проект

**61. ... данные зависят от модели данных и хранятся в фиксированном поле внутри записи. Соответственно, такие данные часто бывает удобно хранить в таблицах, в базах данных или файлах Excel.**

Ответ: Структурированные

**62. Невозможные значения можно исправить, используя ... .**

Ответ:

ручное переопределение

**63. ... направлено(а) на достижение более глубокого понимания данных.**

Ответ:

Исследование данных

**64. Отклонения от свода правил можно исправить, используя ... .**

Ответ:

Сопоставление по ключу или ручное переопределение

**65. Неструктурированные данные ... .**

Ответ:

трудно подогнать под конкретную модель данных, потому что их содержимое зависит от контекста или имеет переменный характер

**66. Комбинированные данные – это:**

Ответ:

Слияние / объединение наборов данных

Создание представлений

Операторы множеств

**67. Разные единицы измерения можно исправить, используя ... .**

Ответ: Пересчет

**68. ... данные естественным образом подходят для представления социальных сетей, а их структура позволяет вычислять такие специфические метрики, как влияние участников и кратчайший путь между двумя людьми.**

Ответ: Графовые

**69. Характеристики больших данных часто называются «тремя V»: ... . Часто эти характеристики дополняются «четвертым V» - ...:**

Ответ:

скорость (velocity)

объем (volume)

разнообразие (variety)

**70. ... - это расширение статистики, способное справляться с огромными объёмами данных.**

Ответ:

Data science

### ***Типовые тестовые задания:***

1. Какие науки включены в Data Mining?

- а. статистика, базы данных, искусственный интеллект;
- б. информатика, базы данных, статистика;
- в. искусственный интеллект, базы данных, базы знаний;
- г. информатика, базы данных, хранилища данных.

2. Какая подсистема СППР отвечает за интеллектуальный анализ данных?

- а. OLTP;
- б. хранилище данных;
- в. SQL;
- г. OLAP;
- д. Data Mining.

3. Репозиторий – это...

- а. словарь терминов;
- б. хранилище метаданных;
- в. каталог с файлами.

4. Если для реализации многомерной модели используют многомерные базы данных, то способ реализации гиперкуба называется

- а. MOLAP;
- б. ROLAP;

- в. HОLAP;
- г. DOLAP;
- д. JOLAP.

- 5 Задача классификации сводится к ...
- а. нахождению частых зависимостей между объектами или событиями;
  - б. определению класса объекта по его характеристикам;
  - в. определению по известным характеристикам объекта значения некоторого его параметра;
  - г. поиску независимых групп и их характеристик во всем множестве анализируемых данных.

### ***Типовые кейсы***

1. Осуществить поиск информации по заданной теме в сети Интернет. Сохранить данные. Загрузить данные в СУБД Access и распределить по таблицам, выстроив связи между объектами. Сделать отчет на основе полученной информации.

2. Представление исходных данных в программе R-Studio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах Excel и R-Studio: подсчет описательных статистик, графическое представление данных. Группировка данных, обнаружение значимых корреляций, зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы).

### ***Примеры тем групповых дискуссий:***

- 1. Методики анализа больших данных
- 2. Аналитический инструментарий
- 3. Рынки решений для управления большими данными
- 4. Теория и практика больших данных в отраслях
- 5. Как использовать аналитические данные для разработки качественных ИТ-сервисов
- 6. Большие данные (big data) в информационных технологиях.

### ***Примеры типовых заданий для контрольной работы:***

Вариант 1. Поиск и определение Big Data, хранение больших данных

- 1. Требуется осуществить поиск источников информации в сети Интернет: открытые и закрытые источники данных. Рассмотреть портал открытых данных РФ.

2. Определить можно ли классифицировать найденные данные как «Большие данные». Сохранить данные.

3. Описать модель организации хранилища данных. Загрузить данные в Excel и СУБД Access, организовать связь между загруженными таблицами с помощью SQL-запросов. Осуществить выгрузку информации для просмотра описательных статистик. Сделать отчет.

Вариант 2. Сформулируйте общую модель хранения больших данных на примере конкретной отрасли (на примере Краснодара и Краснодарского края), как систему, состоящую из связанных друг с другом компонентов.

1. Описать и проанализировать текущее состояние предприятия/организации по применению ИТ на основе имеющейся информации (в описании указать откуда взята информация: документация организации, информация с сайта организации, собранная информация из Интернет по различным источникам). Составить дерево решений.

2. На основе предложенных компонентов информационных технологий, составить поэтапный план внедрения в каждое структурное подразделение в описании будущее состояние

3. Указать какие существуют ограничения по стратегическому развитию (внутренние, внешние).

### ***Тематика докладов:***

1. Большие данные как дальнейшее развитие информационной экономики
2. Цифровая экономика и цифровая трансформация, взаимодействие с большими данными
3. Распределенные вычисления и хранилища данных (облачное хранение)
4. Проблема создания и размещения дата-центров
5. Большие данные и принятие решений.
6. Искусственный интеллект
7. Понятие bigdata. Новые подходы к накоплению и обработке данных в экономике и финансах на микро- и макроуровнях.
8. Открытые данные компьютерных поисковых систем и социальных сетей. GoogleTrends. YandexWorstat.
9. Цифровая трансформация модели потребления.
10. Тенденции в области цифрового контента.
11. Применение методов машинного обучения к большим данным
12. Применение методов Data Mining к большим данным
13. Применение нейронных сетей к анализу больших данных.
14. Поиск ассоциативных правил в больших данных.
15. Программные средства анализа «больших данных»

## Структура зачетного задания

<i>Наименование оценочного материала</i>	<i>Максимальное количество баллов</i>
<i>Вопрос 1.</i>	<i>10</i>
<i>Вопрос 2.</i>	<i>10</i>
<i>Практическое задание</i>	<i>15</i>
<i>Тестовое задание</i>	<i>5</i>

## Показатели и критерии оценивания планируемых результатов освоения компетенций и результатов обучения, шкала оценивания

Таблица 5

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень освоения компетенций
85 – 100 баллов	«зачтено»	УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>Знает верно и в полном объеме</b> основные методы критического анализа и основы системного подхода как общенаучного метода.	Продвинутый
				<b>Умеет верно и в полном объеме</b> анализировать задачу, используя основы критического анализа и системного подхода.	
				<b>Умеет верно и в полном объеме</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.	
		ОПК-2. Способен осуществлять сбор, обработку и статистический анализ данных, необходимых для решения поставленных экономических задач.	ОПК-2.1. Использует основные методы, средства получения, представления, хранения и обработки статистических данных.	<b>Знает верно и в полном объеме</b> методы поиска и систематизации информации об экономических процессах и явлениях.	
				<b>Умеет верно и в полном объеме</b> работать с национальными и международными базами данных с целью поиска информации, необходимой для решения поставленных	

				экономических задач. <b>Умеет верно и в полном объеме</b> рассчитывать экономические и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов на основе типовых методик и действующей нормативно-правовой базы	
				<b>Умеет верно и в полном объеме</b> представить наглядную визуализацию данных.	
		<b>ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</b>	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<b>Знает верно и в полном объеме:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий <b>Умеет верно и в полном объеме:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности	
			ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач	<b>Знает верно и в полном объеме:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий <b>Умеет верно и в полном объеме:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
70 – 84 баллов	«зачтено»	<b>УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач</b>	<b>УК-1.1.</b> Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>Знает с незначительными замечаниями</b> основные методы критического анализа и основы системного подхода как общенаучного метода. <b>Умеет с незначительными замечаниями</b> анализировать задачу, используя основы	<b>Повышенный</b>

				критического анализа и системного подхода.	
				<b>Умеет с незначительными замечаниями</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.	
		<b>ОПК-2. Способен осуществлять сбор, обработку и анализ данных, необходимых для решения поставленных управленческих задач, с использованием современного инструментария и интеллектуальных информационно-аналитических систем</b>	ОПК-2.1. Определяет источники информации и осуществляет их поиск на основе поставленных целей для решения профессиональных задач	<b>Знает с незначительными замечаниями:</b> методы сбора информации, способы и вид ее представления, применяя современное программное обеспечение	
				<b>Умеет с незначительными замечаниями :</b> использовать современный инструментарий и интеллектуальные информационно-аналитические системы	
		<b>ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</b>	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<b>Знает с незначительными замечаниями:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
				<b>Умеет с незначительными замечаниями:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности	
			ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач	<b>Знает с незначительными замечаниями:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
				<b>Умеет с незначительными замечаниями:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	

50 – 69 баллов	«зачтено»	УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<p><b>Знает на базовом уровне с ошибками</b> основные методы критического анализа и основы системного подхода как общенаучного метода.</p> <p><b>Умеет с незначительными замечаниями</b> анализировать задачу, используя основы критического анализа и системного подхода.</p> <p><b>Умеет с незначительными замечаниями</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.</p>	Базовый
		ОПК-2. Способен осуществлять сбор, обработку и анализ данных, необходимых для решения поставленных управленческих задач, с использованием современного инструментария и интеллектуальных информационно-аналитических систем	ОПК-2.1. Определяет источники информации и осуществляет их поиск на основе поставленных целей для решения профессиональных задач	<p><b>Знает на базовом уровне, с ошибками:</b> методы сбора информации, способы и вид ее представления, применяя современное программное обеспечение</p> <p><b>Умеет на базовом уровне, с ошибками:</b> использовать современный инструментарий и интеллектуальные информационно-аналитические системы</p>	
		ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные цифровые информационные технологии, основываясь на принципах их работы	<p><b>Знает на базовом уровне, с ошибками:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий</p> <p><b>Умеет на базовом уровне, с ошибками:</b> использовать современные цифровые информационные технологии для решения задач профессиональной деятельности</p>	

			ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач	<b>Знает на базовом уровне, с ошибками:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий <b>Умеет на базовом уровне, с ошибками:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий	
менее 50 баллов	«не зачтено»	УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи	<b>Не знает на базовом уровне</b> основные методы критического анализа и основы системного подхода как общенаучного метода.	Компетенции не сформированы
				<b>Не умеет на базовом уровне</b> анализировать задачу, используя основы критического анализа и системного подхода.	
				<b>Не умеет с на базовом уровне</b> осуществлять поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации.	
		ОПК-5. Способен использовать при решении профессиональных задач современные информационные технологии и программные средства, включая управление крупными массивами данных и их интеллектуальный анализ	ОПК-5.2. Применяет современные информационные технологии и системы для постановки и решения задач управления, включая управление крупными массивами данных и их интеллектуальный анализ	<b>Не знает на базовом уровне: особенности ИСПОЛЬЗОВАНИЯ</b> современных информационных технологий и систем для постановки и решения задач управления, включая управление крупными массивами данных и их интеллектуальный анализ <b>Не умеет на базовом уровне:</b> решать задачи управления на основе использования современных информационных технологий и систем	
		ОПК-6. Способен понимать принципы работы современных информационных технологий и использовать их	ОПК-6.1. Использует соответствующие содержанию профессиональных задач современные	<b>Не знает на базовом уровне:</b> характеристики соответствующих содержанию профессиональных задач современных цифровых информационных технологий <b>Не умеет на базовом уровне:</b>	

		для решения задач профессиональной деятельности	цифровые информационные технологии, основываясь на принципах их работы	использовать современные цифровые информационные технологии для решения задач профессиональной деятельности	
	ОПК-6.2. Понимает принципы работы современных цифровых информационных технологий, соответствующих содержанию профессиональных задач		<b>Не знает на базовом уровне:</b> принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий		
			<b>Не умеет на базовом уровне:</b> применять принципы работы соответствующих содержанию профессиональных задач современных цифровых информационных технологий		

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**  
**Федеральное государственное бюджетное образовательное учреждение высшего образования**  
**«Российский экономический университет имени Г.В. Плеханова»**  
**Краснодарский филиал РЭУ им. Г. В. Плеханова**

Факультет экономики, менеджмента и торговли

Кафедра экономики и цифровых технологий

**АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ**  
**Б1.О.ДЭ.02.01 ОСНОВЫ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ**

Направление подготовки **38.03.01 ЭКОНОМИКА**

Направленность (профиль) программы  
**ЭКОНОМИКА И УПРАВЛЕНИЕ НА ПРЕДПРИЯТИИ**  
**(ОРГАНИЗАЦИИ)**

Уровень высшего образования **Бакалавриат**

## 1. Цель и задачи дисциплины:

**Цель дисциплины** заключается в решении теоретико-методологических знаний и закрепление профессиональных навыков в области анализа больших данных (Big Data) с учетом зарубежного и российского опыта, а также формирование современного мышления, приобретение знаний и компетенций, позволяющих повышать эффективность принятия решений за счет грамотного сбора, структурирования и применения современных техник анализа больших объемов разнородных данных

### Задачи дисциплины:

- изучить типовые задачи, для решения которых применяется машинное обучение: анализ трендов, социальных сетей, распознавание графических, видео- и аудио-образов, текста, прогнозирование действий;
- изучить методы обработки и анализа данных: регрессия, нейронные сети; графы;
- изучить основные инструменты анализа данных, в том числе методы интеллектуального анализа данных и машинного обучения;
- изучить основные принципы поиска, сбора, очистки, хранения, анализа и обработки статистических данных с целью их визуализации;
- изучить основы обработки экономических и финансовых данных с использованием информационных технологий для решения профессиональных задач на основе функционирования систем искусственного интеллекта.

## 2. Содержание дисциплины:

№ п/п	Наименование разделов / тем дисциплины
<b><i>Раздел 1. Большие данные: современные подходы к обработке и хранению</i></b>	
1.	Тема 1. Основы больших данных
2.	Тема 2. Обработка больших данных
<b><i>Раздел 2. Программное обеспечение в области анализа больших данных</i></b>	
3.	Тема 3. Анализ и извлечение данных
4.	Тема 4. Технология Data Mining
<b>Трудоемкость дисциплины составляет 3 з.е. / 108 часа.</b>	

**Форма контроля – зачет.**

**Составитель:**

к.т.н., доцент кафедры экономики и цифровых технологий Р.Н. Фролов