

**Министерство науки и высшего образования РФ**  
**Федеральное государственное бюджетное образовательное учреждение**  
**высшего образования**  
**«Национальный исследовательский университет «МЭИ»**

---

Направление подготовки/специальность: 09.04.01 Информатика и вычислительная техника

Наименование образовательной программы: Цифровые технологии

Уровень образования: высшее образование - магистратура

Форма обучения: Очная

**Рабочая программа дисциплины**  
**ТЕХНОЛОГИЯ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ**

<b>Блок:</b>	<b>Блок 1 «Дисциплины (модули)»</b>
<b>Часть образовательной программы:</b>	<b>Часть, формируемая участниками образовательных отношений</b>
<b>№ дисциплины по учебному плану:</b>	<b>Б1.Ч.03</b>
<b>Трудоемкость в зачетных единицах:</b>	<b>2 семестр - 4;</b>
<b>Часов (всего) по учебному плану:</b>	<b>144 часа</b>
<b>Лекции</b>	<b>2 семестр - 16 часов;</b>
<b>Практические занятия</b>	<b>не предусмотрено учебным планом</b>
<b>Лабораторные работы</b>	<b>2 семестр - 32 часа;</b>
<b>Консультации</b>	<b>2 семестр - 2 часа;</b>
<b>Самостоятельная работа</b>	<b>2 семестр - 93,5 часа;</b>
<b>в том числе на КП/КР</b>	<b>не предусмотрено учебным планом</b>
<b>Иная контактная работа</b>	<b>проводится в рамках часов аудиторных занятий</b>
<b>включая:</b>	
<b>Лабораторная работа</b>	
<b>Промежуточная аттестация:</b>	
<b>Экзамен</b>	<b>2 семестр - 0,5 часа;</b>

**Москва 2022**

## ПРОГРАММУ СОСТАВИЛ:

Преподаватель

(должность)

	<b>Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»</b>	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Зейн А.Н.
	Идентификатор	R54353a8f-ZeynAIN-7d1f3849

(подпись)

А.Н. Зейн

(расшифровка  
подписи)

## СОГЛАСОВАНО:

Руководитель  
образовательной программы

(должность, ученая степень, ученое звание)

	<b>Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»</b>	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Вишняков С.В.
	Идентификатор	R35b26072-VishniakovSV-02810d9

(подпись)

С.В. Вишняков

(расшифровка подписи)

Заведующий выпускающей  
кафедры

(должность, ученая степень, ученое звание)

	<b>Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»</b>	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Вишняков С.В.
	Идентификатор	R35b26072-VishniakovSV-02810d9

(подпись)

С.В. Вишняков

(расшифровка подписи)

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель освоения дисциплины:** Цель освоения дисциплины является изучение методов решения задач обработки больших по объему, быстро изменяющихся и плохо структурированных данных.

### Задачи дисциплины

- ознакомление с технологиями загрузки, обработки и хранения структурированных и не структурированных данных;
- изучение методов обработки динамических данных;
- изучение методов анализа, интерпретации и применение больших данных для решения задач кластеризации и классификации;
- приобретение навыков преобразования сложных неструктурированных данных для загрузки в реляционную базу.

Формируемые у обучающегося **компетенции** и запланированные **результаты обучения** по дисциплине, соотнесенные с **индикаторами достижения компетенций**:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Запланированные результаты обучения
ПК-3 Способен осуществлять проектирование вычислительных комплексов и систем, включая разработку аппаратного, программного обеспечения, системную интеграцию, ввод в эксплуатацию	ИД-2 <sub>ПК-3</sub> Демонстрирует знание теории баз данных, включая перспективные технологии обработки больших данных	знать: - ограничения реляционной модели при работе с большими данными; - основные методы, обработки, анализа и интерпретации данных в системах Big Data и принцип миграции данных в гетерогенных системах.  уметь: - получать большие данные из открытых и доступных источников, их обрабатывать и модифицировать для работы в реляционной/не реляционной БД; - использовать существующие алгоритмы для классификации объектов и раскрывать не очевидные закономерности.

## 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ВО

Дисциплина относится к основной профессиональной образовательной программе Цифровые технологии (далее – ОПОП), направления подготовки 09.04.01 Информатика и вычислительная техника, уровень образования: высшее образование - магистратура.

Базируется на уровне среднего общего образования.

Результаты обучения, полученные при освоении дисциплины, необходимы при выполнении выпускной квалификационной работы.

### 3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

#### 3.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 4 зачетных единицы, 144 часа.

№ п/п	Разделы/темы дисциплины/формы промежуточной аттестации	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по видам учебной работы										Содержание самостоятельной работы/ методические указания		
				Контактная работа							СР					
				Лек	Лаб	Пр	Консультация		ИКР		ПА	Работа в семестре	Подготовка к аттестации /контроль			
КПР	ГК	ИККП	ТК													
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
1	Данные: статические и динамические; Структурированные и мало-структурированные данные. Недостатки реляционных баз данных.	4	2	2	-	-	-	-	-	-	-	2	-	<p><b><u>Подготовка к лабораторной работе:</u></b> студент должен понимать, что такое реляционная база данных. Как определить родительскую и дочернюю таблицу. какие ограничения в реляционной базе данных. <b><u>Изучение материалов литературных источников:</u></b> [1], стр.20-28 [4], стр.5-12</p>		
1.1	Данные: статические и динамические; Структурированные и мало-структурированные данные.	2		1	-	-	-	-	-	-	-	-	1		-	
1.2	Недостатки реляционных баз данных	2		1	-	-	-	-	-	-	-	-	1		-	
2	Источники больших данных. DOM-модель интернет страницы.	6		2	-	-	-	-	-	-	-	-	4		-	<p><b><u>Подготовка к лабораторной работе:</u></b> Студент проводит анализ доступных источников больших данных (сайты, порталы, открытые ари: ари.hh.ru как пример) <b><u>Изучение материалов литературных источников:</u></b> [2], стр. 169-183</p>
2.1	Источники больших данных.	3		1	-	-	-	-	-	-	-	-	2		-	
2.2	DOM-модель интернет страницы	3		1	-	-	-	-	-	-	-	-	2		-	
3	Копирование больших объемов данных. XML и JSON.	22	2	8	-	-	-	-	-	-	-	12	-	<p><b><u>Подготовка к лабораторной работе:</u></b> Студент изучает формат XML и JSON. Различает коллекции и документы. <b><u>Изучение материалов литературных</u></b></p>		
3.1	Копирование больших	11	1	4	-	-	-	-	-	-	-	6	-			

	объемов данных.												<b><u>источников:</u></b>
3.2	XML и JSON	11	1	4	-	-	-	-	-	-	6	-	[8], стр.73-91
4	Критерий Тьюки. Принцип Бонферрони. MapReduce.	6	2	-	-	-	-	-	-	-	4	-	<b><u>Подготовка к лабораторной работе:</u></b> Студент изучает принципы реализации задачи mapreduce.
4.1	Критерий Тьюки. Принцип Бонферрони.	3	1	-	-	-	-	-	-	-	2	-	<b><u>Изучение материалов литературных источников:</u></b>
4.2	MapReduce.	3	1	-	-	-	-	-	-	-	2	-	[6], стр.1-52
5	Работа с базой данных MongoDB.	42	2	16	-	-	-	-	-	-	24	-	<b><u>Подготовка к лабораторной работе:</u></b> Студент изучает документацию СУБД mongoDB.
5.1	MongoDB. Установка и работа в диалоговом режиме и в интерактивном режиме.	42	2	16	-	-	-	-	-	-	24	-	<b><u>Изучение материалов литературных источников:</u></b> [7], стр.10-52
6	Перенос данных из MongoDB в реляционную базу данных.	5	1	-	-	-	-	-	-	-	4	-	<b><u>Подготовка к лабораторной работе:</u></b> Студент знакомится с различными утилитами, командами и программами для копирования и восстановления баз данных в MongoDB.
6.1	Перенос данных между MongoDB и реляционной базы данных.	5	1	-	-	-	-	-	-	-	4	-	<b><u>Изучение материалов литературных источников:</u></b> [5], стр. 211-247
7	Алгоритмы классификации	10	2	4	-	-	-	-	-	-	4	-	<b><u>Самостоятельное изучение теоретического материала:</u></b> Студент делает обзор существующих методов классификации.
7.1	Алгоритмы классификации	10	2	4	-	-	-	-	-	-	4	-	<b><u>Изучение материалов литературных источников:</u></b> [3], стр.24-73
8	Алгоритмы кластеризации	10	2	4	-	-	-	-	-	-	4	-	<b><u>Самостоятельное изучение теоретического материала:</u></b> Студент делает обзор существующих методов кластеризации
8.1	Алгоритмы кластеризации	10	2	4	-	-	-	-	-	-	4	-	<b><u>Изучение материалов литературных источников:</u></b> [3], стр.81-105
9	Некоторые типовые и перспективные задачи	3	1	-	-	-	-	-	-	-	2	-	<b><u>Подготовка к лабораторной работе:</u></b> студент подготавливает план работы в

	обработки больших данных.												рамках лабораторной работы: выбор компонентов, методов и решений в среде knime.
9.1	Некоторые типовые и перспективные задачи обработки больших данных	3	1	-	-	-	-	-	-	-	2	-	<b><u>Изучение материалов литературных источников:</u></b> [2], стр.12-79 [3], стр. 128-140
	Экзамен	36.0	-	-	-	-	2	-	-	0.5	-	33.5	
	<b>Всего за семестр</b>	<b>144.0</b>	<b>16</b>	<b>32</b>	<b>-</b>	<b>-</b>	<b>2</b>	<b>-</b>	<b>-</b>	<b>0.5</b>	<b>60</b>	<b>33.5</b>	
	<b>Итого за семестр</b>	<b>144.0</b>	<b>16</b>	<b>32</b>	<b>-</b>	<b>2</b>	<b>-</b>	<b>-</b>	<b>0.5</b>	<b>60</b>	<b>33.5</b>	<b>93.5</b>	

**Примечание:** Лек – лекции; Лаб – лабораторные работы; Пр – практические занятия; КПП – аудиторные консультации по курсовым проектам/работам; ИККП – индивидуальные консультации по курсовым проектам/работам; ГК- групповые консультации по разделам дисциплины; СР – самостоятельная работа студента; ИКР – иная контактная работа; ТК – текущий контроль; ПА – промежуточная аттестация

## 3.2 Краткое содержание разделов

### 1. Данные: статические и динамические; Структурированные и мало-структурированные данные. Недостатки реляционных баз данных.

1.1. Данные: статические и динамические; Структурированные и мало-структурированные данные.

Статические данные. Источники статических данных, примеры. Структурированные данные, малоструктурированные данные, не структурированные данные, примеры..

1.2. Недостатки реляционных баз данных

Изолированность, надежность и точность данных в реляционных БД: повышение качества данных за счет производительности..

### 2. Источники больших данных. DOM-модель интернет страницы.

2.1. Источники больших данных.

Источники больших данных. Открытые API, интернет сайты..

2.2. DOM-модель интернет страницы

DOM-модель интернет страницы. Парсинг интернет сайтов..

### 3. Копирование больших объемов данных. XML и JSON.

3.1. Копирование больших объемов данных.

Копирование больших объемов с помощью скриптов Питона..

3.2. XML и JSON

Примеры использование XML файлов для хранения больших данных. Примеры использование JSON файлов для хранения больших данных.

### 4. Критерий Тьюки. Принцип Бонферрони. MapReduce.

4.1. Критерий Тьюки. Принцип Бонферрони.

Критерий Тьюки. Определение. решение задачи.. Принцип Бонферрони. Дополнения и ограничения.

4.2. MapReduce.

MapReduce. Реализация, применение, пример реализации в MongoDB..

### 5. Работа с базой данных MongoDB.

5.1. MongoDB. Установка и работа в диалоговом режиме и в интерактивном режиме.

Работа в интерактивном режима в интерфейсе TROBOT.. Установка MongoDB. Работа в режиме диалогового окна..

### 6. Перенос данных из MongoDB в реляционную базу данных.

6.1. Перенос данных между MongoDB и реляционной базы данных.

Перенос данных от MongoDB к MSSQL. Перенос данных от MSSQL к MongoDB.

## 7. Алгоритмы классификации

### 7.1. Алгоритмы классификации

Основные алгоритмы классификации.. применение элементов аналитической системы Knime.

## 8. Алгоритмы кластеризации

### 8.1. Алгоритмы кластеризации

Основные алгоритмы кластерного анализа.. Применение аналитической системы Knime для кластеризации..

## 9. Некоторые типовые и перспективные задачи обработки больших данных.

### 9.1. Некоторые типовые и перспективные задачи обработки больших данных

Персональная задача: выбор тематики и методов кластеризации/классификации..

## **3.3. Темы практических занятий**

не предусмотрено

## **3.4. Темы лабораторных работ**

1. Ограничения реляционной модели данных;
2. MongoDB (часть1) - установка MongoDB, создание документов и коллекций.;
3. MongoDB (часть2) - Сложные запросы для реализации Map() и Reduce а также MapReduce(). Введение в работе с аналитической системы Knime.;
4. Кластеризация и классификация в Knime.

## **3.5 Консультации**

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

Курсовой проект/ работа не предусмотрены

### 3.7. Соответствие разделов дисциплины и формируемых в них компетенций

Запланированные результаты обучения по дисциплине (в соответствии с разделом 1)	Коды индикаторов	Номер раздела дисциплины (в соответствии с п.3.1)									Оценочное средство (тип и наименование)	
		1	2	3	4	5	6	7	8	9		
<b>Знать:</b>												
основные методы, обработки, анализа и интерпретации данных в системах Big Data и принцип миграции данных в гетерогенных системах	ИД-2ПК-3			+	+							Лабораторная работа/Лабораторная работа 1 (MongoDB часть1) - установка MongoDB, создание документов и коллекций.
ограничения реляционной модели при работе с большими данными	ИД-2ПК-3	+	+									Лабораторная работа/Ограничения реляционной модели данных
<b>Уметь:</b>												
использовать существующие алгоритмы для классификации объектов и раскрывать не очевидные закономерности	ИД-2ПК-3					+	+					Лабораторная работа/Лабораторные работы 2 и 3 - Сложные запросы для реализации Map() и Reduce а также MapReduce(). Введение в работе с аналитической системы Ktime.
получать большие данные из открытых и доступных источников, их обрабатывать и модифицировать для работы в реляционной/не реляционной БД	ИД-2ПК-3								+	+	+	Лабораторная работа/Лабораторная работа 4 - Кластеризация и классификация в Ktime

#### **4. КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫЕ ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ДИСЦИПЛИНЕ)**

##### **4.1. Текущий контроль успеваемости**

###### **2 семестр**

Форма реализации: Выполнение задания

1. Ограничения реляционной модели данных (Лабораторная работа)

Форма реализации: Защита задания

1. Лабораторная работа 1 (MongoDB часть1) - установка MongoDB, создание документов и коллекций. (Лабораторная работа)
2. Лабораторные работы 2 и 3 - Сложные запросы для реализации Map() и Reduce а также MapReduce(). Введение в работе с аналитической системы Kpime. (Лабораторная работа)

Форма реализации: Проверка задания

1. Лабораторная работа 4 - Кластеризация и классификация в Kpime (Лабораторная работа)

Балльно-рейтинговая структура дисциплины является приложением А.

##### **4.2 Промежуточная аттестация по дисциплине**

*Экзамен (Семестр №2)*

учитывается работа студента в течении всего семестра.

В диплом выставляется оценка за 2 семестр.

**Примечание:** Оценочные материалы по дисциплине приведены в фонде оценочных материалов ОПОП.

#### **5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

##### **5.1 Печатные и электронные издания:**

1. Борисова, С. В. Базы данных. Часть 2. Лабораторные работы № 5-8 : методические указания по курсу "Базы данных" по направлению 09.03.01 "Информатика и вычислительная техника" / С. В. Борисова, А. Н. Зейн, Нац. исслед. ун-т "МЭИ" (НИУ"МЭИ") . – М. : Изд-во МЭИ, 2019 . – 28 с.

[http://elib.mpei.ru/action.php?kt\\_path\\_info=ktcore.SecViewPlugin.actions.document&fDocumentId=10965;](http://elib.mpei.ru/action.php?kt_path_info=ktcore.SecViewPlugin.actions.document&fDocumentId=10965)

2. Мартишин, С. А. Базы данных. Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala : учебное пособие для направлений "Информатика и вычислительная техника", "Информационные системы и технологии", "Программная инженерия", "Прикладная математика и информатика", "Прикладная информатика", "Управление в технических системах", "Бизнес-информатика", "Педагогическое образование"(профиль "Информатика") / С. А. Мартишин, В. Л. Симонов, М. В. Храпченко . – Москва : ИНФРА-М, 2020 . – 235 с. – (Высшее образование - Бакалавриат) . - ISBN 978-5-16-015133-5 .;

3. Чураков, Е. П. Введение в многомерные статистические методы : учебное пособие / Е. П. Чураков . – Санкт-Петербург : Лань, 2021 . – 148 с. – (Учебники для вузов. Специальная литература) . - ISBN 978-5-8114-2149-7 .;
4. Бородин, Г. А. Язык разметки XML в СУБД : учебное пособие по курсу "Проектирование баз данных" по направлению "Информатика и вычислительная техника" / Г. А. Бородин, И. Н. Андреева, Нац. исслед. ун-т "МЭИ" . – М. : Изд-во МЭИ, 2016 . – 48 с. - ISBN 978-5-7046-1739-6 .  
[http://elib.mpei.ru/action.php?kt\\_path\\_info=ktcore.SecViewPlugin.actions.document&fDocumentId=8185;](http://elib.mpei.ru/action.php?kt_path_info=ktcore.SecViewPlugin.actions.document&fDocumentId=8185)
5. Мартишин, С. А. Базы данных. Практическое применение СУБД SQL- и NoSQL типа для проектирования информационных систем : учебное пособие для среднего профессионального образования по специальностям УГС 09.02.00 "Информатика и вычислительная техника" / С. А. Мартишин, В. Л. Симонов, М. В. Храпченко . – М. : Форум : ИНФРА-М, 2019 . – 368 с. – (Среднее профессиональное образование) . - ISBN 978-5-8199-0785-6 .;
6. Адлер Ю. П., Черных Е. А.- "Статистическое управление процессами. «Большие данные»", Издательство: "МИСИС", Москва, 2016 - (52 с.)  
[https://e.lanbook.com/book/93652;](https://e.lanbook.com/book/93652)
7. Кайл Б.- "MongoDB в действии", Издательство: "ДМК Пресс", Москва, 2012 - (394 с.)  
[http://e.lanbook.com/books/element.php?pl1\\_cid=25&pl1\\_id=4156;](http://e.lanbook.com/books/element.php?pl1_cid=25&pl1_id=4156)
8. Макшанов А. В., Журавлев А. Е., Тындыкаръ Л. Н.- "Большие данные. Big Data", (2-е изд., стер.), Издательство: "Лань", Санкт-Петербург, 2022 - (188 с.)  
[https://e.lanbook.com/book/198599.](https://e.lanbook.com/book/198599)

## 5.2 Лицензионное и свободно распространяемое программное обеспечение:

1. СДО "Прометей";
2. Office / Российский пакет офисных программ;
3. Windows / Операционная система семейства Linux;
4. Майнд Видеоконференции;
5. Python.

## 5.3 Интернет-ресурсы, включая профессиональные базы данных и информационно-справочные системы:

1. ЭБС Лань - <https://e.lanbook.com/>
2. ЭБС "Университетская библиотека онлайн" - [http://biblioclub.ru/index.php?page=main\\_ub\\_red](http://biblioclub.ru/index.php?page=main_ub_red)
3. Научная электронная библиотека - <https://elibrary.ru/>
4. База данных Web of Science - <http://webofscience.com/>
5. База данных Scopus - <http://www.scopus.com>
6. Электронная библиотека МЭИ (ЭБ МЭИ) - <http://elib.mpei.ru/login.php>
7. База открытых данных Министерства труда и социальной защиты РФ - <https://rosmintrud.ru/opendata>
8. База открытых данных профессиональных стандартов Министерства труда и социальной защиты РФ - <http://profstandart.rosmintrud.ru/obshchiy-informatsionnyy-blok/natsionalnyy-reestr-professionalnykh-standartov/>
9. База открытых данных Министерства экономического развития РФ - <http://www.economy.gov.ru>

## 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Тип помещения	Номер аудитории, наименование	Оснащение
---------------	-------------------------------	-----------

Учебные аудитории для проведения лекционных занятий и текущего контроля	Е-505, Учебная аудитория каф. "ВМСС"	парта, стол преподавателя, стул, мультимедийный проектор, экран, доска маркерная, компьютер персональный
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Учебные аудитории для проведения практических занятий, КР и КП	Е-505, Учебная аудитория каф. "ВМСС"	парта, стол преподавателя, стул, мультимедийный проектор, экран, доска маркерная, компьютер персональный
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Учебные аудитории для проведения лабораторных занятий	Е-522/3, Компьютерный класс №1	стол преподавателя, стол компьютерный, стул, доска маркерная, компьютер персональный
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Учебные аудитории для проведения промежуточной аттестации	Е-505, Учебная аудитория каф. "ВМСС"	парта, стол преподавателя, стул, мультимедийный проектор, экран, доска маркерная, компьютер персональный
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Помещения для самостоятельной работы	Е-522/3, Компьютерный класс №1	стол преподавателя, стол компьютерный, стул, доска маркерная, компьютер персональный
	Е-522/4, Компьютерный класс №2	стол преподавателя, стол компьютерный, стул, мультимедийный проектор, экран, доска маркерная, компьютер персональный
	Е-522/6, Компьютерный класс №3	стол преподавателя, стол компьютерный, стул, доска маркерная, компьютер персональный
	Е-522/9, Компьютерный класс №4	стол преподавателя, стол компьютерный, стул, мультимедийный проектор, экран, доска маркерная, компьютер персональный
Помещения для консультирования	Е-402, Кабинет сотрудников "ВМСС"	
	Е-504а, Кабинет сотрудников	
Помещения для хранения оборудования и учебного инвентаря	Е-403, Склад	стол для работы с документами, шкаф, шкаф для документов

## БАЛЛЬНО-РЕЙТИНГОВАЯ СТРУКТУРА ДИСЦИПЛИНЫ

### Технология обработки больших данных

(название дисциплины)

#### 2 семестр

**Перечень контрольных мероприятий текущего контроля успеваемости по дисциплине:**

- КМ-1 Ограничения реляционной модели данных (Лабораторная работа)
- КМ-2 Лабораторная работа 1 (MongoDB часть1) - установка MongoDB, создание документов и коллекций. (Лабораторная работа)
- КМ-3 Лабораторные работы 2 и 3 - Сложные запросы для реализации Map() и Reduce а также MapReduce(). Введение в работе с аналитической системы Knime. (Лабораторная работа)
- КМ-4 Лабораторная работа 4 - Кластеризация и классификация в Knime (Лабораторная работа)

**Вид промежуточной аттестации – Экзамен.**

Номер раздела	Раздел дисциплины	Индекс КМ:	КМ-1	КМ-2	КМ-3	КМ-4
		Неделя КМ:	4	8	12	16
1	Данные: статические и динамические; Структурированные и мало-структурированные данные. Недостатки реляционных баз данных.					
1.1	Данные: статические и динамические; Структурированные и мало-структурированные данные.		+			
1.2	Недостатки реляционных баз данных		+			
2	Источники больших данных. DOM-модель интернет страницы.					
2.1	Источники больших данных.		+			
2.2	DOM-модель интернет страницы		+			
3	Копирование больших объемов данных. XML и JSON.					
3.1	Копирование больших объемов данных.			+		
3.2	XML и JSON			+		
4	Критерий Тьюки. Принцип Бонферрони. MapReduce.					
4.1	Критерий Тьюки. Принцип Бонферрони.			+		
4.2	MapReduce.			+		
5	Работа с базой данных MongoDB.					

5.1	MongoDB. Установка и работа в диалоговом режиме и в интерактивном режиме.			+	
6	Перенос данных из MongoDB в реляционную базу данных.				
6.1	Перенос данных между MongoDB и реляционной базы данных.			+	
7	Алгоритмы классификации				
7.1	Алгоритмы классификации				+
8	Алгоритмы кластеризации				
8.1	Алгоритмы кластеризации				+
9	Некоторые типовые и перспективные задачи обработки больших данных.				
9.1	Некоторые типовые и перспективные задачи обработки больших данных				+
Вес КМ, %:		20	20	30	30