

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ  
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования «Кабардино-Балкарский государственный университет  
им. Х.М. Бербекова» (КБГУ)

ИНСТИТУТ ИНФОРМАТИКИ, ЭЛЕКТРОНИКИ И РОБОТОТЕХНИКИ

КАФЕДРА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

СОГЛАСОВАНО

УТВЕРЖДАЮ

Руководитель образовательной  
программы \_\_\_\_\_ Т.Ю.Хаширова

Директор института ИЭиР  
\_\_\_\_\_ Н.В. Черкесова

« \_\_\_\_ » \_\_\_\_\_ 2021 г.

« \_\_\_\_ » \_\_\_\_\_ 2021 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Методы обработки больших данных

Направление подготовки

**09.03.01 – Информатика и вычислительная техника**

Профиль

Интеллектуальные системы обработки информации и  
управления

Квалификация (степень) выпускника

Бакалавр

Форма обучения

Очная форма

Нальчик 2021

Рабочая программа дисциплины «Методы обработки больших данных» /сост. Хаширова Т.Ю. – Нальчик: КБГУ, 2021. 23 с.

Рабочая программа предназначена для преподавания дисциплины части, формируемой участниками образовательных отношений, дисциплины по выбору Б1.В.ДВ.01.01 «Методы обработки больших данных» студентам очной формы обучения, по направлению подготовки 09.03.01 – Информатика и вычислительная техника, для профиля: Интеллектуальные системы обработки информации и управления, в 5 семестре, 3 курса.

Рабочая программа дисциплины составлена с учетом Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.03.01 Информатика и вычислительная техника, утвержденного приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 г. №929 (зарегистрировано в Минюсте России 10 октября 2017 г. № 48489).

## СОДЕРЖАНИЕ

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	4
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО .....	4
3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	4
4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.....	6
5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО И РУБЕЖНОГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ...	10
6. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ, УМЕНИЙ, НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ .....	17
7. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	18
8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ....	23
9. ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ .....	25
ПРИЛОЖЕНИЕ .....	26

## 1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью курса является

- знакомство студентов с проблемами и технологиями работы с данными огромного
- размера (Big Data), формирование представления о возможностях и ограничениях
- основных методов обработки, анализа больших данных;
- изучение студентами современного состояния теории Big Data;
- приобретение знаний о технологиях подготовки, хранения, обработки и анализа больших
- данных;
- применение статистических и математических методов для анализа больших объемов информации;
- получение знаний о технологиях и методах анализа Big Data и результатов.

Задачи дисциплины:

- изучение основных методов хранения и анализа больших данных, в том числе в сфере корпоративных финансов, мировых рынков, построения моделей и прогнозирования;
- формирование практического навыка по работе с кластером хранения и обработки Big Data на примере Cloudera Hadoop, выбору методов и технических решений в зависимости от типа решаемой задачи обработки данных и их объема;
- умение использовать инструментарий анализа данных и их визуализации.

## 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

Дисциплина Б1.О.11.05 «Методы обработки больших данных» относится к части, формируемой участниками образовательных отношений, дисциплины по выбору Б1.В.ДВ.01.01 «Методы обработки больших данных» студентам очной формы обучения, по направлению подготовки 09.03.01 – Информатика и вычислительная техника, для профиля: Интеллектуальные системы обработки информации и управления, предназначена для преподавания студентам очной формы обучения в 5 семестре, 3 курса.

Дисциплина «Методы обработки больших данных» занимает важное место в современном образовании и базируется на знаниях, полученных студентами в процессе изучения дисциплин: Дискретный анализ, Теоретические основы информатики, Алгоритмические основы информатики, Программирование и основы алгоритмизации, Языки и методы программирования.

В результате освоения дисциплины обучающийся сможет частично продемонстрировать следующие обобщенные трудовые функции (ОТФ):

- Разработка требований и проектирование программного обеспечения (профессиональный стандарт 06.001 – «Программист», код D, уровень квалификации – 6).
- Концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности (профессиональный стандарт 06.022 – «Системный аналитик», код С, уровень квалификации – 6).

## 3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В совокупности с другими дисциплинами профиля «Интеллектуальные системы обработки информации и управления (ИСОИУ)» дисциплина «Методы обработки больших данных» направлена на формирование следующих компетенций в соответствии с ФГОС ВО и ОПОП ВО по направлению подготовки 09.03.01 – Информатика и вычислительная техника

(ИиВТ) (уровень бакалавриата):

ПКС-2.1; ПКС-3.1; ПКС-2.2; ПКС-3.2; ПКС-3.3; ПКС-2.3

а) профессиональных (ПКС):

- способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности (ПКС-2);
- способен разрабатывать графический дизайн интерфейса, проектировать пользовательские интерфейсы по готовому образцу или концепции интерфейса. (ПКС-3).

**Коды и наименования индикаторов достижения компетенции:**

ПКС-2.1 Способен освоить методы планирования проектных работ; методы классического системного анализа; теорию управления бизнес-процессами; шаблоны оформления бизнестребований; методы концептуального проектирования; методы оценки качества программных систем.

ПКС-2.2 Способен разрабатывать технико-экономическое обоснование; разрабатывать техническое задание на систему; разрабатывать требования к подсистемам системы и осуществлять контроль их качества; организовать оценку соответствия требованиям существующих систем и их аналогов; выполнять сопровождение приемочных испытаний и ввод в эксплуатацию системы; обрабатывать запросы на изменение требований к системе..

ПКС-2.3 Способен применить навыки составления графика контрольных мероприятий; приемы разработки бизнестребований к системе; определять ключевые свойства и ограничения системы; навыками выделения подсистем системы.

ПКС-3.1 Способен освоить технические требования к интерфейсной графике; стандарты, регламентирующие требования к эргономике взаимодействия человек – система; технологии алгоритмической визуализации данных; основы верстки с использованием языков разметки; основы верстки с использованием языков описания стилей; основы программирования с использованием сценарных языков.

ПКС-3.2 Способен разрабатывать и оформлять проектную документацию на интерфейс; создавать интерактивные прототипы интерфейса; создавать графические документы в программах подготовки растровых и векторных изображений.

ПКС-3.3 Способен применить навыки верстки; навыки работы с программами редактирования табличных данных.

В результате освоения дисциплины обучающийся должен:

Знать:

- особенности использования внешних и внутренних каналов связи при передаче больших объемов данных;
- состав и назначение модулей online/offline обработки, подсистем индексирования и хранения данных;
- фазы и действия, выполняемые программным каркасом Hadoop при исполнении MapReduce-программы;
- состав кластера Apache Spark и Apache Kafka;
- состав кластера Apache Cassandra;
- средства обеспечения согласованности в высокопроизводительных системах хранения данных;
- виды алгоритмов машинного обучения, пригодные для использования в offline-модулях,
- виды алгоритмов машинного обучения, пригодные для использования в online-модулях.

Уметь:

- пользоваться основными метриками для оценки производительности BigData-систем;
- применять стандартные средства разработки для организации процессов потоковой обработки больших объемов данных;

- применять технологию контейнеризации OpenShift для развертывания программных решений в облаке;
  - оценивать необходимую производительность узлов сети исходя из прогнозируемой нагрузки;
  - проектировать отказоустойчивые высокопроизводительные модули индексации данных, предназначенные для поддержки конкретных алгоритмов BigData-аналитики;
  - применять гибридные вычислительные технологии в задачах машинного обучения.
- Владеть:
- методами декомпозиции BigData-задач;
  - способами организации внешних сетевых интерфейсов BigData-систем и построения модулей для online/offline обработки;
  - принципами построения интеграционного кластерного интерфейса для организации взаимодействия распределенных систем;
  - навыками управления жизненным циклом Hadoop-кластера;
  - способами развертывания MapReduce-программ в Hadoop-окружении;
  - математическим аппаратом теории вероятностей и теории массового обслуживания для построения моделей потоков данных в BigData-системах;
  - принципами трансформации потока входящей информации в поток объектов хранения Big Data-системы;
  - способами повышения производительности алгоритмов машинного обучения с применением гибридных вычислительных технологий.

#### 4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

В таблице 1 приводится описание содержания дисциплины, структурированное по разделам, с указанием по каждому разделу формы текущего контроля: защита лабораторной работы (ЛР), коллоквиум (К), рубежный контроль (РК), тестирование (Т).

Таблица 1

##### Содержание разделов дисциплины

№	Наименование раздела	Содержание раздела	Код контролируемой компетенции (или ее части)	Форма текущего контроля
1.	Архитектура Big Data систем. Часть 1.	Общие вопросы организации Big Data-систем. Проблемы организации Big Data-систем. Классификация характера задач для выбора подходящей реализации. Методы декомпозиции Big Data-задач. Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.	ПКС-2	К, Т
2.	Архитектура Big Data систем. Часть 2.	Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем	ПКС-2	К, ЛР, Т

		online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.		
3.	Инфраструктура Big Data систем	Системы построения виртуализированных сетевых инфраструктур. Системы легковесной контейнеризации. Вопросы виртуализации сетевых функций и построения программно-определяемых сетей. Назначение технологий SDN/NFV. Механизмы работы систем управления виртуализированными контейнерами. Технология контейнеризации OpenShift для развертывания программных решений в облаке. Принципы использования технологий OpenStack и OpenShift для организации сетей виртуальных машин и систем управления контейнерами.	ПКС-2	К, Т
4.	Пакетная распределенная обработка больших объемов данных	Технология организации пакетной обработки больших объемов данных MapReduce. Фазы и действия, выполняемые программным каркасом Hadoop при исполнении MapReduce-программы. Разработка Java-приложения, использующие Hadoop для выполнения пакетной обработки данных по таймеру. Управление жизненным циклом Hadoop-кластера. Способы развертывания MapReduce-программ в Hadoop-окружении.	ПКС-2, ПКС-3	К, Т
5.	Оценка производительности Big Data-систем	Элементы теории массового обслуживания для оценки производительности распределенных вычислительных Big Data-систем. Основы теории массового обслуживания для расчета интенсивности поступления запросов на каждый узел сети. Оценка необходимой производительности узлов сети исходя из прогнозируемой нагрузки. Математический аппарат теории вероятностей и теории массового обслуживания для построения моделей потоков данных в Big Data-системах.	ПКС-2, ПКС-3	К, ЛР, Т
6.	Асинхронная обработка больших объемов данных	Принципы работы с технологией Apache Spark для выполнения асинхронных вычислительных операций и системы очередей для управления асинхронными процессами в Big Data-системах. Состав кластера Apache Spark и Apache Kafka. Организация загрузки/выгрузки информации в системах Apache Spark и Apache Kafka. Аминистрирование систем Apache Spark и Apache Kafka, способами организации асинхронного взаимодействия нескольких вычислительных задач.	ПКС-2, ПКС-3	К, ЛР, Т
7.	Хранение больших объемов данных	Высокопроизводительные NoSQL-системы. Предпосылки, типы и характеристика. Состав и характеристики высокопроизводительных файловых систем на примере GFS, HDFS и NFS v4.1. Состав кластера Apache Cassandra. Средства обеспечения согласованности в высокопроизводительных системах хранения данных. Определение необходимого типа системы	ПКС-2, ПКС-3	К, ЛР, Т

		хранения и схему упаковки данных в зависимости от задачи. Принципы трансформации потока входящей информации в поток объектов хранения Big Data-системы		
8.	Построение индексов	Модули индексации данных для Big Data-систем. Назначение модулей индексации данных и требования к ним. Проектирование отказоустойчивых высокопроизводительных модулей индексации данных, предназначенные для поддержки конкретных алгоритмов Big Data-аналитики. Принципы построения и оценки производительности подсистем индексации данных	ПКС-2, ПКС-3	
9.	Библиотеки машинного обучения для Big Data-систем	Алгоритмы машинного обучения, предназначенные для пакетной и поточной обработки. Виды алгоритмов машинного обучения, пригодные для использования в offline-модулях, Виды алгоритмов машинного обучения, пригодные для использования в online-модулях. Применение гибридных вычислительных технологий в задачах машинного обучения. Способы повышения производительности алгоритмов машинного обучения с применением гибридных вычислительных технологий.	ПКС-3	

Таблица 2

**Структура дисциплины «Модели и методы представления знаний»**

Вид работы	Трудоемкость	Всего
	5 семестр	
<b>Общая трудоемкость</b>	<b>108</b>	<b>144</b>
<b>Контактная работа:</b>	<b>51</b>	<b>60</b>
<i>Лекции (Л)</i>	17	34
<i>Практические занятия (ПЗ)</i>	-	-
<i>Лабораторные работы (ЛР)</i>	34	51
<b>Самостоятельная работа:</b>	<b>48</b>	<b>50</b>
Курсовой проект (КП), курсовая работа	—	—
Расчетно-графическое задание (РГЗ)	—	—
Реферат (Р)	—	—
Эссе (Э)	—	—
Самостоятельное изучение разделов	48	50
Контрольная работа (К)	—	—
Самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам, рубежному контролю и т.д.),	—	—
Переаттестация	—	—
Подготовка и сдача экзамена	<b>9</b>	<b>9</b>
<b>Вид итогового контроля (зачет, экзамен)</b>	<b>Зачет</b>	<b>Зачет с</b>



Таблица 3

## Лекционные занятия

№	Наименование раздела	Темы лекций
1.	Архитектура Big Data систем. Часть 1.	Общие вопросы организации Big Data-систем. Проблемы организации Big Data-систем. Классификация характера задач для выбора подходящей реализации. Методы декомпозиции Big Data-задач. Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.
2.	Архитектура Big Data систем. Часть 2.	Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.
3.	Инфраструктура Big Data систем	Системы построения виртуализированных сетевых инфраструктур. Системы легковесной контейнеризации. Вопросы виртуализации сетевых функций и построения программно-определяемых сетей. Назначение технологий SDN/NFV. Механизмы работы систем управления виртуализированными контейнерами. Технология контейнеризации OpenShift для развертывания программных решений в облаке. Принципы использования технологий OpenStack и OpenShift для организации сетей виртуальных машин и систем управления контейнерами.
4.	Пакетная распределенная обработка больших объемов данных	Технология организации пакетной обработки больших объемов данных MapReduce. Фазы и действия, выполняемые программным каркасом Hadoop при исполнении MapReduce-программы. Разработка Java-приложения, использующие Hadoop для выполнения пакетной обработки данных по таймеру. Управление жизненным циклом Hadoop-кластера. Способы развертывания MapReduce-программ в Hadoop-окружении.
5.	Оценка производительности Big Data-систем	Элементы теории массового обслуживания для оценки производительности распределенных вычислительных Big Data-систем. Основы теории массового обслуживания для расчета интенсивности поступления запросов на каждый узел сети. Оценка необходимой производительности узлов сети исходя из прогнозируемой нагрузки. Математический аппарат теории вероятностей и теории массового обслуживания для построения моделей потоков данных в Big Data-системах.
6.	Асинхронная обработка больших объемов данных	Принципы работы с технологией Apache Spark для выполнения асинхронных вычислительных операций и системы очередей для управления асинхронными процессами в Big Data-системах. Состав кластера Apache Spark и Apache Kafka. Организация загрузки/выгрузки информации в системах Apache Spark и Apache Kafka. Аминистрирование систем Apache Spark и Apache Kafka, способами организации асинхронного взаимодействия нескольких вычислительных задач.
7.	Хранение больших объемов данных	Высокопроизводительные NoSQL-системы. Предпосылки, типы и характеристика. Состав и характеристики высокопроизводительных файловых систем на примере GFS, HDFS и NFS v4.1. Состав кластера Apache Cassandra. Средства обеспечения согласованности в высокопроизводительных системах хранения данных. Определение необходимого типа системы хранения и схему упаковки данных в зависимости от задачи. Принципы трансформации потока входящей информации в поток объектов хранения Big Data-системы
8.	Построение индексов	Модули индексации данных для Big Data-систем. Назначение модулей индексации данных и требования к ним. Проектирование отказоустойчивых высокопроизводительных модулей индексации данных, предназначенные для поддержки конкретных алго-

		ритмов Big Data-аналитики. Принципы построения и оценки производительности подсистем индексации данных
9.	Библиотеки машинного обучения для Big Data-систем	Алгоритмы машинного обучения, предназначенные для пакетной и поточной обработки. Виды алгоритмов машинного обучения, пригодные для использования в offline-модулях, Виды алгоритмов машинного обучения, пригодные для использования в online-модулях. Применение гибридных вычислительных технологий в задачах машинного обучения. Способы повышения производительности алгоритмов машинного обучения с применением гибридных вычислительных технологий.

Таблица 4.

Практические занятия не предусмотрены

Таблица 5

#### Лабораторные работы

№	Наименование тем
1.	Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.
2.	Стандартные средства разработки для организации процессов поточной обработки больших объемов данных.
3.	Назначение технологий SDN/NFV.
4.	Разработка Java-приложения, использующие Hadoop для выполнения пакетной обработки данных по таймеру.
5.	Математический аппарат теории вероятностей и теории массового обслуживания для построения моделей потоков данных в Big Data-системах.
6.	Состав кластера Apache Spark и Apache Kafka.
7.	Определение необходимого типа системы хранения и схему упаковки данных в зависимости от задачи.
8.	Проектирование отказоустойчивых высокопроизводительных модулей индексации данных, предназначенные для поддержки конкретных алгоритмов Big Data-аналитики.
9.	Алгоритмы машинного обучения, предназначенные для пакетной и поточной обработки.
10.	Виды алгоритмов машинного обучения, пригодные для использования в online-модулях.

## 5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО И РУБЕЖНОГО КОНТРОЛЯ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Формы контроля текущих, рубежных и промежуточных знаний студентов по дисциплине определяются в соответствии с учебным планом образовательной программы и в соответствии с действующим Положением о балльно-рейтинговой системе оценки успеваемости студентов КБГУ.

От обучающихся требуется посещение занятий, выполнение лабораторных работ, знакомство с рекомендованной литературой.

При аттестации обучающихся оценивается качество работы на занятиях (умение вести дискуссию, способность четко и ёмко формулировать свои мысли), уровень подготовки к самостоятельной деятельности, качество выполнения заданий (презентаций, докладов, выполнение лабораторных работ и др.).

Конечными результатами освоения программы дисциплины являются сформированные когнитивные дескрипторы «знать», «уметь», «владеть», расписанные по отдельным компетенциям. Формирование этих дескрипторов происходит в течение всего семестра по этапам в рамках различного вида занятий и самостоятельной работы.

В ходе изучения дисциплины предусматриваются текущий, рубежный контроль и промежуточная аттестация. Формы контроля текущих, рубежных и промежуточных знаний студентов по дисциплине определяются в соответствии с учебным планом образова-

тельной программы и в соответствии с действующим Положением о балльно-рейтинговой системе оценки успеваемости студентов КБГУ.

От обучающихся требуется посещение занятий, выполнение лабораторных работ, знакомство с рекомендованной литературой.

При аттестации обучающихся оценивается качество работы на занятиях (умение вести дискуссию, способность четко и ёмко формулировать свои мысли), уровень подготовки к самостоятельной деятельности, качество выполнения заданий (презентаций, докладов, выполнение лабораторных работ и др.).

**5.1. Оценочные материалы для текущего контроля.** Цель текущего контроля – оценка результатов работы в семестре и обеспечение своевременной обратной связи, для коррекции обучения, активизации самостоятельной работы обучающегося. Объектом текущего контроля являются конкретизированные результаты обучения (учебные достижения) по дисциплине

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины «Методы обработки больших данных», оценка качества подготовки на основании выполненных заданий ведется преподавателем (с обсуждением результатов), баллы начисляются в зависимости от сложности задания.

#### **Критерии формирования оценок (оценивания) устного опроса**

Устный опрос является одним из основных способов учёта знаний обучающегося по дисциплине «Методы обработки больших данных». Развёрнутый ответ должен представлять собой связное, логически последовательное сообщение на заданную тему, показывать его умение применять определения.

В результате устного опроса знания, обучающегося оцениваются по следующей шкале:

<b>3 балла</b>	<b>2 балла</b>	<b>1 балл</b>	<b>0 баллов</b>
ставится, если обучающийся: 1) полно излагает изученный материал, даёт правильное экономическое понятие; 2) обнаруживает понимание материала, может обосновать свои суждения, применить знания на практике, привести необходимые примеры не только по учебнику, но и самостоятельно составленные; 3) излагает материал последовательно и правильно с точки зрения норм литературного языка.	ставится, если обучающийся даёт ответ, удовлетворяющий тем же требованиям, что и для балла «1», но допускает 1-2 ошибки, которые сам же исправляет, и 1-2 недочёта в последовательности и языковом оформлении излагаемого.	ставится, если обучающийся обнаруживает знание и понимание основных положений данной темы, но: 1) излагает материал неполно и допускает неточности в определении понятий; 2) не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры; 3) излагает материал непоследовательно и допускает ошибки в языковом оформлении излагаемого.	ставится, если обучающийся обнаруживает незнание большей части соответствующего раздела изучаемого материала, допускает ошибки в формулировке.

Баллы «1», «2», «3» могут ставиться не только за единовременный ответ, но и за рассредоточенный во времени, т.е. за сумму ответов, данных на протяжении занятия. начисляются в зависимости от сложности задания.

## **5.2. Оценочные материалы для самостоятельной работы обучающегося ( типовые задачи) (при наличии)**

Рабочая программа предусматривает проведение лекционных, лабораторных занятий, а также самостоятельную работу обучающихся. В ФГБОУ ВО «Кабардино-Балкарский государственный университет» действует балльно-рейтинговая система оценки учебных достижений, обучающихся по образовательным программам, реализуемым на основании федеральных государственных образовательных стандартов. Балльно-рейтинговая система оценки знаний является одной из составляющих системы управления качеством образовательной деятельности в университете.

### **Перечень вопросов по дисциплине для самостоятельного изучения**

Способы организации внешних сетевых интерфейсов BigData-систем и построения модулей для online/offline обработки.

Принципы построения интеграционного кластерного интерфейса для организации взаимодействия распределенных систем.

Принципы использования технологий OpenStack и OpenShift для организации сетей виртуальных машин и систем управления контейнерами.

Управления жизненным циклом Hadoop-кластера. Способами развертывания MapReduce-программ в Hadoop-окружении.

Основы теории массового обслуживания для расчета интенсивности поступления запросов на каждый узел сети. Оценка необходимой производительности узлов сети исходя из прогнозируемой нагрузки.

Администрирование систем Apache Spark и Apache Kafka, способами организации асинхронного взаимодействия нескольких вычислительных задач.

Принципы трансформации потока входящей информации в поток объектов хранения Big Data-системы.

Библиотеки машинного обучения для Big Data-систем. Алгоритмы машинного обучения, предназначенные для пакетной и поточной обработки.

### **Критерии формирования оценок по заданиям для самостоятельной работы студента**

«Отлично» (5 баллов) – обучающийся показал глубокие знания материала по поставленным вопросам, грамотно, логично его излагает, структурировал и детализировал информацию, избегая простого повторения информации из текста, информация представлена в переработанном виде. Свободно использует необходимые формулы при решении задач;

«Хорошо» (4 балла) – обучающийся твердо знает материал, грамотно его излагает, не допускает существенных неточностей в процессе решения задач;

«Удовлетворительно» (3 балла) – обучающийся имеет знания основного материала по поставленным вопросам, но не усвоил его деталей, допускает отдельные неточности при решении задач;

«Неудовлетворительно» (менее 3 баллов) – обучающийся допускает грубые ошибки в ответе на поставленные вопросы и при решении задач

## **5.2. Оценочные материалы для рубежного контроля**

Рубежный контроль осуществляется по более или менее самостоятельным разделам – учебным модулям курса и проводится по окончании изучения материала модуля в заранее установленное время. Рубежный контроль проводится с целью определения качества усвоения материала учебного модуля в целом. В течение семестра проводится три таких контрольных мероприятия по графику.

В качестве форм рубежного контроля используется тестирование (письменное или компьютерное), проведение коллоквиума.

Рубежный контроль освоения студентом дисциплины осуществляется в рамках балльно-рейтинговой системы. Распределение баллов в соответствии с действующим Положением о балльно-рейтинговой системе оценки успеваемости студентов КБГУ приведено в таблице 7.

Таблица 7

Распределение баллов в соответствии с действующим Положением о балльно-рейтинговой системе

№ рейтинговой точки	Коллоквиум	Лаб.практикум	Посещаемость	Тестирование	Итого
1	7	8	3	5	23
2	7	8	3	5	23
3	7	8	4	5	24

Таблица 8

#### Критерии оценки

Вид мероприятия	Критерии оценки	Баллы
Коллоквиум (устный опрос по теме)	- ясность, четкость и доказательность изложения ответов на вопросы; - владение специальными терминами; - системность знаний по тематике	0-21 балл
Лабораторное занятие	- понимание цели и задач работы - выполнение заданий и обработка результатов - отчет и защита лабораторной работы	0-24 балла
Компьютерное тестирование по разделам дисциплины	Результаты тестирования (Количество баллов = $5 \cdot \varphi$ , $\varphi$ - доля правильно отвеченных тестов по теме).	0-15 баллов
Посещение занятий	При более 3 пропусках без уважительной причины занятий аннулируются баллы	0-10 баллов
Экзамен	ясность, четкость и доказательность изложения ответов на вопросы; - владение специальными терминами; - системность знаний по тематике дисциплины в целом	0-30 баллов
Итоговая оценка		0-100 баллов

#### 5.2.1. Оценочные материалы для проведения коллоквиума

Перечень вопросов, выносимых на рейтинговый контроль первой точки 5 семестра (контролируемые компетенции ПКС-2, ПКС-3)

##### Архитектура Big Data систем. Часть 1.

Общие вопросы организации Big Data-систем. Проблемы организации Big Data-систем. Классификация характера задач для выбора подходящей реализации. Методы декомпозиции Big Data-задач. Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.

Перечень вопросов, выносимых на рейтинговый контроль второй точки 8 семестра (контролируемые компетенции ПКС-2, ПКС-5)

Архитектура Big Data систем. Часть 2.

Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных. Основные метрики для оценки производительности Big Data-систем. Способы организации внешних сетевых интерфейсов Big Data-систем и построения модулей для online/offline обработки.

Инфраструктура Big Data систем.

Системы построения виртуализированных сетевых инфраструктур. Системы легкой контейнеризации. Вопросы виртуализации сетевых функций и построения программно определяемых сетей. Назначение технологий SDN/NFV. Механизмы работы систем управления виртуализированными контейнерами. Технология контейнеризации OpenShift для развертывания программных решений в облаке. Принципы использования технологий OpenStack и OpenShift для организации сетей виртуальных машин и систем управления контейнерами.

### **5.2.2. Оценочные материалы: тестирование**

Тестирование обучающихся проводится в онлайн-режиме согласно расписанию, в ЭИОС open.kbsu.ru

### **Примерные тестовые задания для РТ 1 (контролируемая компетенция ПКС-2) Полный перечень тестовых заданий представлен в ЭОИС**

1. Что понимается под представлением знаний?

- а) это кодирование информации, на каком – либо формальном языке;
- б) знания, представленные в программе на языке C ++;
- в) знания, представленные в учебниках по математике;
- г) моделирование знаний специалистов – экспертов.

1. Выбор уровня детализации модели это

- +: стратификация
- : верификация
- : идентификация

3. Задача выбора числовых значений неопределенных числовых параметров это

- : стратификация
- : верификация
- +: идентификация

4. Задача проверки адекватности модели при найденных значениях параметров это

- : стратификация
- +: верификация
- : идентификация

### **Примерные тестовые задания для РТ 2 (контролируемая компетенция ОПК-5) Полный перечень тестовых заданий представлен в ЭОИС**

1. Какие определения, представленные ниже не являются моделями представления знаний?

- a) продукционные модели;
- b) фреймы;
- c) компьютерные модели;
- d) семантические сети;
- e) формально - логические модели.

2. Что представляют собой семантическая сеть?

- a) сетевой график, вершины которого - сроки выполнения работ;
- b) это нейронная сеть, состоящая из нейронов;
- c) ориентированный граф, вершины которого - понятия, а дуги – отношения между ними.

3. Какой из основных типов отношений семантической сети, представленных ниже, может быть названа как АКО (A - Kind – Of)?

- a) это;
- b) элемент класса;
- c) имеет частью;
- d) принадлежит;
- e) функциональная связь.

4. Чем отличаются семантические сети и фреймы?

- a) Элемент модели состоит из множества незаполненных значений некоторых атрибутов, именуемых «слотами»;
- b) наследование по АКО- связям;
- c) элемент модели – структура, используемая для обозначения объектов и понятий.

**Примерные тестовые задания для РТ 3 (контролируемая компетенция ОПК-9)**

**Полный перечень тестовых заданий представлен в ЭОИС**

1. Что объединяет семантические сети и фреймы?

- a) организация процедуры вывода;
- b) наследование свойств;
- c) множества незаполненных значений некоторых атрибутов, именуемых «слотами»;
- d) структуры, использующихся для обозначения объектов и понятий.

2. Какие из выражений, представленных ниже, являются структурной частью фрейма?

- a) значение N- го слота;
- b) шаблон;
- c) примитивные типы данных.

3. На каком формализме НЕ основаны логические модели?

- a) исчисление высказываний;
- b) пропозициональная логика;
- c) силлогизмы Аристотеля;
- d) правильно построенные формулы;
- e) нечёткие системы (fuzzy set).

**Критерии формирования оценок по тестовым заданиям**

Выполнение тестирования оценивается согласно проценту правильных ответов. Макси-

мально возможное количество баллов за тестирование – 5.

### 5.3. Оценочные материалы для промежуточной аттестации

Целью промежуточных аттестаций по дисциплине является оценка качества освоения дисциплины обучающимися.

Промежуточная аттестация предназначена для объективного подтверждения и оценивания достигнутых результатов обучения после завершения изучения дисциплины. Осуществляется в конце семестра и представляет собой итоговую оценку знаний по дисциплине «Методы обработки больших данных» в виде проведения зачета в 5 семестре.

Промежуточная аттестация может проводиться в устной, письменной форме, и в форме тестирования. На промежуточную аттестацию отводится до 30 баллов.

Вопросы, выносимые на зачет (контролируемые компетенции ПКС-2, ПКС-3)

1. Общие вопросы организации Big Data-систем. Проблемы организации Big Data-систем.
2. Классификация характера задач для выбора подходящей реализации. Методы декомпозиции Big Data-задач.
3. Метрики производительности Big Data-систем, состав и назначение подсистем внутренних/внешних обменов, подсистем online/offline аналитики.
4. Особенности использования внешних и внутренних каналов связи при передаче больших объемов данных.
5. Способы интеграции подсистем online/offline обработки, подсистем индексирования и хранения данных.
6. Введение в проблематику построения систем мониторинга вычислительных ресурсов. Состав и назначение (на примере системы мониторинга вычислительных ресурсов) модулей online/offline обработки, подсистем индексирования и хранения данных.
7. Системы построения виртуализированных сетевых инфраструктур. Системы легковесной контейнеризации. Вопросы виртуализации сетевых функций и построения программно определяемых сетей.
8. Назначение технологий SDN/NFV.
9. Технология организации пакетной обработки больших объемов данных MapReduce.
10. Фазы и действия, выполняемые программным каркасом Hadoop при исполнении MapReduce-программы.
11. Разработка Java-приложения, использующие Hadoop для выполнения пакетной обработки данных по таймеру.
12. Элементы теории массового обслуживания для оценки производительности распределенных вычислительных Big Data-систем
13. Основы теории массового обслуживания для расчета интенсивности поступления запросов на каждый узел сети.
14. Оценка необходимой производительности узлов сети исходя из прогнозируемой нагрузки.
15. Принципы работы с технологией Apache Spark для выполнения асинхронных вычислительных операций и системы очередей для управления асинхронными процессами в BigData-системах.
16. Состав кластера Apache Spark и Apache Kafka.
17. Высокопроизводительные NoSQL-системы. Предпосылки, типы и характеристика. Состав и характеристики высокопроизводительных файловых систем на примере GFS, HDFS и NFS v4.1.



18. Состав кластера Apache Cassandra. Средства обеспечения согласованности в высокопроизводительных системах хранения данных.
19. Модули индексации данных для Big Data-систем. Назначение модулей индексации данных и требования к ним.
20. Проектирование отказоустойчивых высокопроизводительных модулей индексации данных, предназначенные для поддержки конкретных алгоритмов.
21. Принципы построения и оценки производительности подсистем индексации данных.
22. Алгоритмы машинного обучения, предназначенные для пакетной и поточной обработки.
23. Виды алгоритмов машинного обучения, пригодные для использования в offline-модулях, Виды алгоритмов машинного обучения, пригодные для использования в online-модулях.
24. Применение гибридных вычислительных технологий в задачах машинного обучения.
25. Способы повышения производительности алгоритмов машинного обучения с применением гибридных вычислительных технологий.

### **Критерии формирования оценок по промежуточной аттестации**

«Отлично» получают обучающиеся, которые свободно ориентируются в материале и отвечают без затруднений. Обучающийся способен к выполнению сложных заданий, постановке целей и выборе путей их реализации. Работа выполнена полностью без ошибок, решено 100% задач;

«Хорошо» получают обучающиеся, которые относительно полно ориентируются в материале, отвечают без затруднений, допускают незначительное количество ошибок. Обучающийся способен к выполнению сложных заданий. Работа выполнена полностью, но имеются не более одной негрубой ошибки и одного недочета, не более трех недочетов. Допускаются незначительные неточности при решении задач, решено 70% задач;

«Удовлетворительно» получают обучающиеся, у которых недостаточно высок уровень владения материалом. В процессе ответа на экзамене допускаются ошибки и затруднения при изложении материала. Обучающийся правильно выполнил не менее 2/3 всей работы или допустил не более одной грубой ошибки и двух недочетов, не более одной грубой и одной негрубой ошибки, не более трех негрубых ошибок, одной негрубой. Обучающийся затрудняется с правильной оценкой предложенной задачи, дает неполный ответ, решено 55% задач;

«Неудовлетворительно» получают обучающиеся, которые допускают значительные ошибки. Обучающийся имеет лишь начальную степень ориентации в материале. В работе число ошибок и недочетов превысило норму для оценки 3 или правильно выполнено менее 2/3 всей работы. Обучающийся дает неверную оценку ситуации, решено менее 50% задач.

## **6. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ОПРЕДЕЛЯЮЩИЕ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ЗНАНИЙ, УМЕНИЙ, НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ**

Максимальная сумма (100 баллов), набираемая студентом по дисциплине включает две составляющие:

- первая составляющая – оценка регулярности, своевременности и качества выполнения студентом учебной работы по изучению дисциплины в течение периода изучения дисциплины (семестра, или нескольких семестров) (сумма – не более 70 баллов). Баллы, характеризующие успеваемость студента по дисциплине, набираются им в течение всего периода обучения за изучение отдельных тем и выполнение отдельных видов работ.

- вторая составляющая – оценка знаний студента по результатам промежуточной аттестации (не более 30 –баллов).

Критерием оценки уровня сформированности компетенций в рамках учебной дисциплины «Модели и методы представления знаний» в 4 семестре является зачет с оценкой.

Общий балл текущего и рубежного контроля складывается из следующих составляющих (Приложение).

Целью промежуточных аттестаций по дисциплине является оценка качества освоения дисциплины обучающимися.

Таблица 9. Результаты освоения формирования, подлежащие проверке

Компетенция по ФГОС	Основные показатели освоения (показатели достижения результата)	Код показателя освоения
ПКС-2. Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности	<b>ИД-1<sub>ПКС-2</sub> Знать:</b> языки, утилиты и среды программирования, и средства пакетного выполнения процедур	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы (раздел 5)
	<b>ИД-2<sub>ПКС-2</sub> Уметь:</b> использовать выбранную среду программирования для разработки процедур проверки работоспособности программного обеспечения на выбранном языке программирования.	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы
	<b>ИД-3<sub>ПКС-2</sub> Владеть:</b> навыками разработки процедуры измерения требуемых характеристик программного обеспечения.	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы (раздел 5)
ПКС-3. Способен разрабатывать графический дизайн интерфейса, проектировать пользовательские интерфейсы по готовому образцу или концепции интерфейса.	<b>ИД-1<sub>ПКС-3</sub> Знать:</b> основные виды диагностических данных и способы их представления	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы (раздел 5)
	<b>ИД-2<sub>ПКС-3</sub> Уметь:</b> писать программный код процедур проверки работоспособности программного обеспечения на выбранном языке программирования	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы
	<b>ИД-3<sub>ПКС-3</sub> Владеть:</b> навыками разработки процедуры сбора диагностических данных	Типовые оценочные материалы для устного опроса, типовые тестовые задания, лабораторные работы (раздел 5)

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 7.1. Нормативно-правовая база

1. Приказ Минобрнауки России от 05.04.2017 №301 "Об утверждении Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры" (Зарегистрировано в Минюсте России 14.07.2017 № 47415).
2. Федеральный государственный образовательный стандарт высшего образования по направлению подготовки 02.04.02 – «Фундаментальная информатика и информационные технологии» (уровень магистратуры), утвержденный приказом Министерства образования и науки Российской Федерации от 23 августа 2017 г. N 811 (Зарегистрировано в Минюсте РФ 13 сентября 2017 г. Регистрационный N 48168) (с изм. и доп.,

вступ. в силу с 01.09.2021).

3. Федеральный закон "Об образовании в Российской Федерации" от 29.12.2012 N 273-ФЗ [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_140174/](http://www.consultant.ru/document/cons_doc_LAW_140174/)

## 7.2. Основная литература

1. Железнов М.М. Методы и технологии обработки больших данных : учебно-методическое пособие / Железнов М.М.. — Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/101802.html>
2. Воронов В.И. Data Mining - технологии обработки больших данных : учебное пособие / Воронов В.И., Воронова Л.И., Усачев В.А.. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/81324.html>
3. Глотова М.Ю. ИКТ и математические методы обработки данных : учебное пособие / Глотова М.Ю., Самохвалова Е.А.. — Москва : Московский педагогический государственный университет, 2019. — 244 с. — ISBN 978-5-4263-0767-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/94642.html>.

## 7.3. Дополнительная литература

1. Дубровский С.А. Методы обработки и анализа экспериментальных данных : учебное пособие / Дубровский С.А., Дудина В.А., Садыева Я.В.. — Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2015. — 62 с. — ISBN 978-5-88247-719-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/55640.html>
2. Деревнин Д.А. Статистическая обработка экспериментальных данных : учебно-методическое пособие / Деревнин Д.А., Ситников В.Н.. — Тюмень : Тюменский индустриальный университет, 2019. — 50 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/101430.html>
3. Батищев Р.В. Структуры и алгоритмы обработки данных. Часть 1 : учебное пособие / Батищев Р.В.. — Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2014. — 90 с. — ISBN 5-88247-716-6. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/55658.html>

## 7.4. Периодические издания (газета, вестник, бюллетень, журнал)

1. Журнал «Моделирование, оптимизация и информационные технологии»

## 7.5. Интернет-ресурсы

1. Библиотека по информационным технологиям (<http://www.itexpert.ru/rus/biblio/cobit>)
2. Электронная библиотека образовательных и просветительских изданий (<http://www.iqlib.ru/>)

## 7.6. Современные профессиональные базы данных

1. База данных Science Index (РИНЦ) <http://elibrary.ru>
2. Национальная электронная библиотека РГБ <https://нэб.рф>
3. Крупнейшая единая база данных, содержащая аннотации и информацию о цитируемости рецензируемой научной литературы, со встроенными инструментами отслеживания, анализа и визуализации данных. [www.scopus.com](http://www.scopus.com)
4. Самая полная математическая база данных, охватывающая материалы с конца 19 века. zbMath содержит документы, журналы и книги по математике, статистике, информатике, а также машиностроению, физике, естественным наукам и др. [www.zbmath.org](http://www.zbmath.org) (доступ открытый)

## 7.7. Методические указания по проведению различных учебных занятий и другим видам самостоятельной работы

### Методические рекомендации по изучению дисциплины для обучающихся

Приступая к изучению дисциплины, обучающемуся необходимо ознакомиться с тематическим планом занятий, списком рекомендованной учебной литературы. Следует уяснить последовательность выполнения индивидуальных учебных заданий, занести в свою рабочую тетрадь темы и сроки проведения семинаров, написания учебных и творческих работ. При изучении дисциплины обучающиеся выполняют следующие задания: изучают рекомендованную учебную и научную литературу; пишут контрольные работы, готовят доклады и сообщения к практическим занятиям; выполняют самостоятельные творческие работы, участвуют в выполнении практических заданий. Уровень и глубина усвоения дисциплины зависят от активной и систематической работы на лекциях, изучения рекомендованной литературы, выполнения контрольных письменных заданий

Курс изучается на лекциях, лабораторных занятиях, при самостоятельной и индивидуальной работе обучающихся. Обучающийся для полного освоения материала должен не пропускать занятия и активно участвовать в учебном процессе. Лекции включают все темы и основные вопросы теории и практики. Для максимальной эффективности изучения необходимо постоянно вести конспект лекций, знать рекомендуемую преподавателем литературу, позволяющую дополнить знания и лучше подготовиться к лабораторным занятиям.

В соответствии с учебным планом на каждую тему выделено необходимое количество часов практических занятий, которые проводятся в соответствии с вопросами, рекомендованными к изучению по определенным темам. Обучающиеся должны регулярно готовиться к лабораторным занятиям и участвовать в обсуждении вопросов. При подготовке к занятиям следует руководствоваться конспектом лекций и рекомендованной литературой. Тематический план дисциплины, учебно-методические материалы, а также список рекомендованной литературы приведены в рабочей программе

### Методические рекомендации при работе над конспектом во время проведения лекции

В процессе лекционных занятий целесообразно конспектировать учебный материал. Для этого используются общие и утвердившиеся в практике правила, и приемы конспектирования лекций.

Конспектирование лекций ведется в специально отведенной для этого тетради, каждый лист которой должен иметь поля, на которых делаются пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений.

Целесообразно записывать тему и план лекций, рекомендуемую литературу к теме. Записи разделов лекции должны иметь заголовки, подзаголовки, красные строки. Для вы-

деления разделов, выводов, определений, основных идей можно использовать цветные карандаши и фломастеры.

Названные в лекции ссылки на первоисточники надо пометить на полях, чтобы при самостоятельной работе найти и вписать их. В конспекте дословно записываются определения понятий, категорий и законов. Остальное должно быть записано своими словами.

Каждому обучающемуся необходимо выработать и использовать допустимые сокращения наиболее распространенных терминов и понятий.

### **Методические рекомендации по подготовке к лабораторным занятиям**

Лабораторные занятия – составная часть учебного процесса, групповая форма занятий при активном участии обучающихся. Лабораторные занятия способствуют углубленному изучению наиболее сложных проблем науки и служат основной формой подведения итогов самостоятельной работы обучающихся. Целью лабораторных занятий является углубление и закрепление теоретических знаний, полученных обучающимися на лекциях и в процессе самостоятельного изучения учебного материала, а, следовательно, формирование у них определенных умений и навыков.

В ходе подготовки к лабораторному занятию необходимо прочесть конспект лекции, изучить основную литературу, ознакомиться с дополнительной литературой, выполнить выданные преподавателем задания. При этом учесть рекомендации преподавателя и требования программы. Дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы.

Желательно при подготовке к лабораторным занятиям по дисциплине одновременно использовать несколько источников, раскрывающих заданные вопросы.

### **Методические рекомендации по организации самостоятельной работы**

Самостоятельная работа обучающихся – способ активного, целенаправленного приобретения обучающимися новых для него знаний и умений без непосредственного участия в этом процесса преподавателей. Повышение роли самостоятельной работы обучающихся при проведении различных видов учебных занятий предполагает:

- оптимизацию методов обучения, внедрение в учебный процесс новых технологий обучения, повышающих производительность труда преподавателя, активное использование информационных технологий, позволяющих обучающемуся в удобное для него время осваивать учебный материал;
- широкое внедрение компьютеризированного тестирования;
- совершенствование методики проведения практик и научно-исследовательской работы обучающихся, поскольку именно эти виды учебной работы в первую очередь готовят обучающихся к самостоятельному выполнению профессиональных задач;
- модернизацию системы курсового и дипломного проектирования, которая должна повышать роль обучающихся в подборе материала, поиске путей решения задач.

Самостоятельная работа приводит обучающихся к получению нового знания, упорядочению и углублению имеющихся знаний, формированию у него профессиональных навыков и умений. Самостоятельная работа выполняет ряд функций:

- развивающую;
- информационно-обучающую;
- ориентирующую и стимулирующую;
- воспитывающую;
- исследовательскую.

В рамках курса выполняются следующие виды самостоятельной работы:

- Проработка учебного материала (по конспектам, учебной и научной литературе);

- Выполнение разноуровневых заданий;
- Работа с тестами и вопросами для самопроверки;
- Выполнение итоговой контрольной работы.

Обучающимся рекомендуется с самого начала освоения курса работать с литературой и предлагаемыми заданиями в форме подготовки к очередному аудиторному занятию. При этом актуализируются имеющиеся знания, а также создается база для усвоения нового материала, возникают вопросы, ответы на которые обучающийся получает в аудитории.

Необходимо отметить, что некоторые задания для самостоятельной работы по курсу имеют определенную специфику. При освоении курса обучающийся может пользоваться библиотекой вуза, которая в полной мере обеспечена соответствующей литературой. Значительную помощь в подготовке к очередному занятию может оказать имеющийся в учебно-методическом комплексе краткий конспект лекций. Он же может использоваться и для закрепления полученного в аудитории материала. Самостоятельная работа обучающихся предусмотрена учебным планом и выполняется в обязательном порядке. Задания предложены по каждой изучаемой теме и могут готовиться индивидуально или в группе. По необходимости обучающийся может обращаться за консультацией к преподавателю. Выполнение заданий контролируется и оценивается преподавателем.

Для успешной организации самостоятельной работы все активнее применяются разнообразные образовательные ресурсы в сети Интернет: системы тестирования по различным областям, виртуальные лекции, лаборатории, при этом пользователю достаточно иметь компьютер и подключение к Интернету для того, чтобы связаться с преподавателем, решать вычислительные задачи и получать знания. Использование сетей усиливает роль самостоятельной работы обучающихся и позволяет кардинальным образом изменить методику преподавания.

Обучающийся может получать все задания и методические указания через сервер, что дает ему возможность привести в соответствие личные возможности с необходимыми для выполнения работ трудозатратами. Обучающийся имеет возможность выполнять работу дома или в аудитории. Большое воспитательное и образовательное значение в самостоятельном учебном труде обучающийся имеет самоконтроль. Самоконтроль возбуждает и поддерживает внимание и интерес, повышает активность памяти и мышления, позволяет обучающемуся своевременно обнаружить и устранить допущенные ошибки и недостатки, объективно определить уровень своих знаний, практических умений. Самое доступное и простое средство самоконтроля с применением информационно-коммуникационных технологий - это ряд тестов «on-line», которые позволяют в режиме реального времени определить свой уровень владения предметным материалом, выявить свои ошибки и получить рекомендации по самосовершенствованию.

### **Методические рекомендации по работе с литературой**

Всю литературу можно разделить на учебники и учебные пособия, оригинальные научные монографические источники, научные публикации в периодической печати. Из них можно выделить литературу основную (рекомендуемую), дополнительную и литературу для углубленного изучения дисциплины.

Изучение дисциплины следует начинать с учебника, поскольку учебник – это книга, в которой изложены основы научных знаний по определенному предмету в соответствии с целями и задачами обучения, установленными программой.

При работе с литературой необходимо учитывать, что имеются различные виды чтения, и каждый из них используется на определенных этапах освоения материала.

Предварительное чтение направлено на выявление в тексте незнакомых терминов и поиск их значения в справочной литературе. В частности, при чтении указанной литературы необходимо подробнейшим образом анализировать понятия.

Сквозное чтение предполагает прочтение материала от начала до конца. Сквозное чтение литературы из приведенного списка дает возможность обучающимся сформировать свод основных понятий из изучаемой области и свободно владеть ими.

Выборочное – наоборот, имеет целью поиск и отбор материала. В рамках данного курса выборочное чтение, как способ освоения содержания курса, должно использоваться при подготовке к практическим занятиям по соответствующим разделам.

Аналитическое чтение – это критический разбор текста с последующим его конспектированием. Освоение указанных понятий будет наиболее эффективным в том случае, если при чтении текстов обучающийся будет задавать к этим текстам вопросы. Часть из этих вопросов сформулирована в ФОС в перечне вопросов для собеседования. Перечень этих вопросов ограничен, поэтому важно не только содержание вопросов, но сам принцип освоения литературы с помощью вопросов к текстам.

Целью изучающего чтения является глубокое и всестороннее понимание учебной информации. Есть несколько приемов изучающего чтения: чтение по алгоритму предполагает разбиение информации на блоки: название; автор; источник; основная идея текста; фактический материал; анализ текста путем сопоставления имеющихся точек зрения по рассматриваемым вопросам; новизна.

Прием постановки вопросов к тексту имеет следующий алгоритм:

- медленно прочитать текст, стараясь понять смысл изложенного;
- выделить ключевые слова в тексте;
- постараться понять основные идеи, подтекст и общий замысел автора.

Прием тезирования заключается в формулировании тезисов в виде положений, утверждений, выводов.

К этому можно добавить и иные приемы: прием реферирования, прием комментирования.

Важной составляющей любого солидного научного издания является список литературы, на которую ссылается автор. При возникновении интереса к какой-то обсуждаемой в тексте проблеме всегда есть возможность обратиться к списку относящейся к ней литературы. В этом случае вся проблема как бы разбивается на составляющие части, каждая из которых может изучаться отдельно от других. При этом важно не терять из вида общий контекст и не погружаться чрезмерно в детали, потому что таким образом можно не увидеть главного.

Подготовка к экзамену должна проводиться на основе лекционного материала, материала практических занятий с обязательным обращением к основным учебникам по курсу. Это позволит исключить ошибки в понимании материала, облегчит его осмысление, прокомментирует материал многочисленными примерами.

## **8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **8.1. Требования к материально-техническому обеспечению**

Для проведения лекционных занятий с компьютерной поддержкой требуется наличие аудитории с проекционным оборудованием, также при изучении дисциплины «Методы обработки больших данных» предполагается использование интерактивной доски.

Во время самостоятельной работы студенты используют компьютерные классы института информатики, электроники и компьютерных технологий, электронные читальные залы КБГУ и домашние компьютеры.

При проведении занятий лекционного типа используются:

лицензионное программное обеспечение:

- Продукты Microsoft (Desktop Education ALNG LicSaPk OLVS Academic Edition Enterprise) подписка (Open Value Subscription);
- Антивирусное программное обеспечение Kaspersky Endpoint Security Стандартный Russian Edition;
- AltLinux (Альт Образование 8);

свободно распространяемые программы:

- Academic MarthCAD License - математическое программное обеспечение, которое позволяет выполнять, анализировать важнейшие инженерные расчеты и обмениваться ими;
- WinZip для Windows - программ для сжатия и распаковки файлов;
- Adobe Reader для Windows – программа для чтения PDF файлов;
- Far Manager - консольный файловый менеджер для операционных систем семейства Microsoft Windows.

## **8.2. Особенности реализации дисциплины для инвалидов и лиц с ограниченными возможностями здоровья**

Для студентов с ограниченными возможностями здоровья созданы специальные условия для получения образования. В целях доступности получения высшего образования по образовательным программам инвалидами и лицами с ограниченными возможностями здоровья университетом обеспечивается (аудитория для самостоятельной работы и коллективного пользования специальными техническими средствами для обучения инвалидов и лиц с ОВЗ в КБГУ, аудитория № 145 Главный корпус КБГУ):

1. Альтернативная версия официального сайта в сети «Интернет» для слабовидящих;

2. Для инвалидов с нарушениями зрения (слабовидящие, слепые):

- присутствие ассистента, оказывающего обучающемуся необходимую помощь, дублирование вслух справочной информации о расписании учебных занятий; наличие средств для усиления остаточного зрения, брайлевской компьютерной техники, видеоувеличителей, программ невидимого доступа к информации, программ-синтезаторов речи и других технических средств приема-передачи учебной информации в доступных формах для студентов с нарушениями зрения;
- задания для выполнения на экзамене зачитываются ассистентом;
- письменные задания выполняются на бумаге, надиктовываются ассистенту обучающимся;

3. Для инвалидов и лиц с ограниченными возможностями здоровья по слуху (слабослышащие, глухие):

- на зачете/экзамене присутствует ассистент, оказывающий студенту необходимую техническую помощь с учетом индивидуальных особенностей (он помогает занять рабочее место, передвигаться, прочитать и оформить задание, в том числе записывая под диктовку);
- зачет/экзамен проводится в письменной форме;

4. Для инвалидов и лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата, созданы материально-технические условия, обеспечивающие возможность беспрепятственного доступа обучающихся в учебные помещения, объекты питания, туалетные и другие помещения университета, а также пребывания в указанных помещениях (наличие расширенных дверных проемов, поручней и других приспособлений).

- письменные задания выполняются на компьютере со специализированным программным обеспечением или надиктовываются ассистенту;
- по желанию студента экзамен проводится в устной форме.

Обучающиеся из числа лиц с ограниченными возможностями здоровья обеспечены электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.



## 9. ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Рабочая программа: Методы обработки больших данных

одобрена на 2021/2022 учебный год. Протокол № \_\_\_\_\_ заседания кафедры от

« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

В рабочую программу внесены следующие изменения:

1. В части раздела «Учебно-методическое и информационное обеспечение дисциплины»

---

---

---

---

---

---

Разработчик программы \_\_\_\_\_

Зав. кафедрой \_\_\_\_\_

## ПРИЛОЖЕНИЕ

### Распределение баллов текущего и рубежного контроля

№п/п	Вид контроля	Сумма баллов			
		Общая сумма	1-я точка	2-я точка	3-я точка
1	Посещение занятий	до 10 баллов	до 3 б.	до 3б.	до 4б.
2	Текущий контроль:	до 30 баллов	до 10 б.	до 10 б.	до 10 б.
3	Рубежный контроль (тестирование и коллоквиум)	до 30 баллов	до 10 б.	до 10 б.	до 10 б.
4	Итого сумма текущего и рубежного контроля	до 70 баллов	до 23б	до 23 б	до 24 б