

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ
Директор по образовательной
деятельности


С.Т. Князев
«07» сентября 2023



РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1164094	Разработка систем анализа больших данных

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа Прикладной искусственный интеллект	Код ОП 09.03.03
Направление подготовки Прикладная информатика	Код направления и уровня подготовки 09.03.03

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Коломыцева Анна Олеговна	к.э.н. доцент	доцент	информационных технологий и систем управления
2	Павлов Марк Владимирович	без ученой степени, без ученого звания	ассистент	информационных технологий и систем управления

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ Разработка систем анализа больших данных

1.1. Аннотация содержания модуля

Целью освоения модуля «Разработка систем анализа больших данных» является формирование у студентов понимания архитектур информационных систем обработки больших данных. Подробно рассматривается концепция MapReduce, архитектура Hadoop. В результате освоения модуля студент понимает основные принципы секционирования, сегментирования и перемешивания. В результате слушатель может оптимизировать ациклические графы вычислений, основанные на многоузловой обработке, понимает разницу между узкими и широкими преобразованиями. В модуле дается основа микросервисной архитектуры на примере кластеров Kubernetes.

1.2. Структура и объем модуля

Таблица 1

№	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Разработка систем анализа больших данных	3
ИТОГО по модулю:		3

1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	1. Подсистемы хранения и извлечения данных; 2. Пакетная и потоковая обработка данных
Постреквизиты и кореквизиты модуля	Не предусмотрены

1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения
1	2	3	4
Разработка систем анализа больших данных	ПК-2. Способен разрабатывать и тестировать программные компоненты решения задач в системах	ПК-2.1. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. 3-1. Знает современные языки программирования, библиотеки и программные платформы для функционального, логического, объектно-ориентированного программирования приложений

	искусственного интеллекта		систем искусственного интеллекта (Python, R, C++, C#) ПК-2.1. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)
Разработка систем анализа больших данных	ПК-7. Способен разрабатывать системы анализа больших данных	ПК-7.1. Разрабатывает программные компоненты извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных	ПК-7.1 З-1. Знает общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных ПК-7.1 З-2. Знает принципы работы экосистемы Hadoop, фреймворка SPARK ПК-7.1 З-3. Знает устройство интерфейсов между реляционными SQL-хранилищами данных и нереляционными NoSQL-хранилищами данных ПК-7.1 З-4. Знает предметно-ориентированные языки ПК-7.1 У-1. Умеет настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных ПК-7.1 У-2. Умеет разрабатывать программное обеспечение для очистки и валидации наборов больших данных ПК-7.1 У-3. Умеет выполнять потоковую обработку данных (data streaming, event processing) ПК-7.1 У-4. Умеет использовать шины данных (Apache Kafka) ПК-7.1 У-5. Умеет использовать языки запросов, в том числе нереляционных, для поддержки различных типов данных (например, XML, RDF, JSON, мультимедиа) и операций с

			большими данными (например, матричные операции)
		ПК-7.2. Разрабатывает программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных	<p>ПК-7.2. З-1. Знает принципы и методы анализа больших данных, включая спецификации и стандартизацию метаданных</p> <p>ПК-7.2. З-2. Знает устройство и принципы работы систем обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL)</p> <p>ПК-7.2. З-3. Знает архитектуру и принципы работы промышленных решений, созданных на основе искусственного интеллекта</p> <p>ПК-7.2. З-4. Знает методы и технологии машинного обучения на больших данных</p> <p>ПК-7.2. У-1. Умеет разрабатывать программное обеспечение для анализа больших данных</p> <p>ПК-7.2. У-2. Умеет разрабатывать программные и технические средства визуализации больших данных и результатов их анализа</p> <p>ПК-7.2. У-3. Умеет использовать системы обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL, процессы и инструменты)</p> <p>ПК-7.2. У-4. Умеет использовать технологии науки о данных и больших данных в разработке для решения практических задач промышленности</p> <p>ПК-7.2. У-5. Умеет описывать и управлять качеством и достоверностью больших данных</p>

1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной форме.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Разработка систем анализа больших данных

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Коломыцева Анна Олеговна	к.э.н. доцент	доцент	информационных технологий и систем управления
2	Павлов Марк Владимирович	без ученой степени, без ученого звания	ассистент	информационных технологий и систем управления

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Коломыцева Анна Олеговна, кандидат экономических наук, доцент, информационных технологий и систем управления;
- Павлов Марк Владимирович, ассистент, информационных технологий и систем управления.

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Базовый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания; Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Введение в обработку больших данных	Определение больших данных. Типичные решаемые задачи с помощью больших данных. Концепция MapReduce.
2	Введение в Hadoop	Hadoop как набор инструментов для разработки систем анализа больших данных. Архитектура Hadoop. Понятие распределенной файловой системы. HDFS (Hadoop Distributed File System).
3	Основы MapReduce в Hadoop	Парадигма и ключевые функции MapReduce. Этапы процесса обработки данных с помощью MapReduce. Использование YARN. Понятия и функции ResourceManager и NodeManager. Конфигурирование MapReduce.
4	Применение MapReduce для решения практических задач	Структура MapReduce программы. Сцепление задач MapReduce. Соединение данных из разных источников.
5	Работа с графами в Hadoop	Реализация графов в MapReduce. Особенности решения основных задач на графах в MapReduce.
6	Введение в Pig и Hive	Процедурный язык для выполнения запросов к данным Pig. Сравнение Pig Latin с SQL. Особенности выполнения Pig-скриптов. NoSQL-хранилище класса SQL-on-Hadoop Hive. Основы HiveQL.
7	Использование ZooKeeper в Hadoop	Необходимость сервиса для координации распределенных процессов. Архитектура ZooKeeper. Клиенты и сессии. Основы ZooKeeper Atomic Broadcast.
8	Использование NoSQL баз данных	Понятие NoSQL баз данных. Введение в HBase. Архитектура HBase. Работа с HBase в Hadoop.

9	Использование Kubernetes и Docker для разработки Hadoop приложений	Основные понятия Kubernetes и Docker. Использование Docker-образов для разворачивания Hadoop приложений. Декларативное и императивное создание Hadoop кластеров.
---	--	--

1.3. Направление, виды воспитательной деятельности и используемые технологии

Таблица 1.2

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения
Профессиональное воспитание	профориентационная деятельность	Технология формирования уверенности и готовности к самостоятельной успешной профессиональной деятельности Технология самостоятельной работы	ПК-2. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. У-1. Умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации.

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Разработка систем анализа больших данных

Печатные издания

1. Bernard Marr Big Data in Practice. How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results / Bernard Marr. 2020. – 952 с.
2. Nitin Kumar. Big Data Using Hadoop and Hive. David Pallai. Mercury Learning and information. Dulles, 2021. – 201с.
3. Сеймур Митч. Kafka Streams и ksql: данные в реальном времени. – СПб.: Питер, 2023. – 432с.: ил. – (Серия «Библиотека программиста»).

4. Билл Фрэнкс. Революция в аналитике. Как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики. – М.: Альпина Диджитал, 2020. – 430с.
5. Дейтел Пол, Дейтел Харви. Python. Искусственный интеллект, большие данные и облачные вычисления. – СПб.: Питер, 2020. – 864с.: ил. – (Серия «Для профессионалов»).
6. Rathinaraja Jeyaraj. Big Data with Hadoop MapReduce: A Classroom Approach. – Apple Academic, 2020. – 426с.
7. Алан Бьюли Изучаем SQL / Алан Бьюли. - М.: Символ-плюс, 2022. – 851 с.
8. Esayas Chakiso Aloto. MongoDB Administration: The Comprehensive Guide to Administering MongoDB Database. 2020. – 304с.
9. Gerardus Blokdyk. Hbase A Complete Guide. – Sebastopol: O'Reilly Media, 2021. – 244с.
10. Джон Гудсон Практическое руководство по доступу к данным / Джон Гудсон, Роб Стюард. - М.: БХВ-Петербург, 2022. – 304 с.

Профессиональные базы данных, информационно-справочные системы

- 1) Единое окно доступа к образовательным ресурсам. Раздел Информатика и информационные технологии - http://window.edu.ru/catalog/p_rubr=2.2.75.6
- 2) Зональная научная библиотека УрФУ <http://lib.urfu.ru>
- 3) Научная электронная библиотека Elibrary.ru <https://www.elibrary.ru/>
- 4) Электронная библиотечная сеть "Лань" <http://e.lanbook.com/>
- 5) Портал информационно-образовательных ресурсов УрФУ <http://study.urfu.ru/>

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

- 1) Государственная публичная научно-техническая библиотека <http://www.gpntb.ru>
- 2) Список библиотек, доступных в Интернет и входящих в проект «Либнет» <http://www.valley.ru/-nicr/listrum.htm>
- 3) Российская национальная библиотека <http://www.rsl.ru>
- 4) Свободная энциклопедия Википедия <https://ru.wikipedia.org/>

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Разработка систем анализа больших данных

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Подключение к сети Интернет</p>	<p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>Microsoft Windows 8.1 Pro 64-bit RUS OLP NL Acdmc</p>
2	Лабораторные занятия	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Доска аудиторная</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>SQL Svr Standard Core ALNG LicSAPk MVL 2Lic CoreLic EES</p> <p>Microsoft Windows 8.1 Pro 64-bit RUS OLP NL Acdmc</p> <p>СУБД Microsoft SQL Server 2012 или более новая</p>
3	Самостоятельная работа студентов	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>SQL Svr Standard Core ALNG LicSAPk MVL 2Lic CoreLic EES</p> <p>Microsoft Windows 8.1 Pro 64-bit RUS OLP NL Acdmc</p> <p>СУБД Microsoft SQL Server 2012 или более новая</p>
4	Текущий контроль и промежуточная аттестация	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	<p>Office 365 EDUA3 ShrdSvr ALNG SubsVL MVL PerUsr B Faculty EES</p> <p>Microsoft Windows 8.1 Pro 64-bit RUS OLP NL Acdmc</p>

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
ПО ДИСЦИПЛИНЕ**
Разработка систем анализа больших данных

Код модуля
1164094(1)

Модуль
Разработка систем анализа больших данных

Екатеринбург

Оценочные материалы составлены автором(ами):

№ п/п	Фамилия, имя, отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Коломыцева Анна Олеговна	к.э.н., доцент	доцент	кафедра информационных технологий и систем управления
2	Павлов Марк Владимирович	без ученой степени, без ученого звания	ассистент	кафедра информационных технологий и систем управления

Авторы:

- Коломыцева Анна Олеговна, доцент, информационных технологий и систем управления
- Павлов Марк Владимирович, ассистент, информационных технологий и систем управления

1. СТРУКТУРА И ОБЪЕМ ДИСЦИПЛИНЫ Разработка систем анализа больших данных

1.	Объем дисциплины в зачетных единицах	3	
2.	Виды аудиторных занятий	Лекции Лабораторные занятия	
3.	Промежуточная аттестация	Зачет	
4.	Текущая аттестация	Домашняя работа	2

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (ИНДИКАТОРЫ) ПО ДИСЦИПЛИНЕ МОДУЛЯ Разработка систем анализа больших данных

Индикатор – это признак / сигнал/ маркер, который показывает, на каком уровне обучающийся должен освоить результаты обучения и их предъявление должно подтвердить факт освоения предметного содержания данной дисциплины, указанного в табл. 1.3 РПМ-РПД.

Таблица 1

Код и наименование компетенции	Индикаторы достижения компетенции	Планируемые результаты обучения	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3	4
ПК-2. Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта	ПК-2.1. Разрабатывает приложения систем искусственного интеллекта	ПК-2.1. 3-1. Знает современные языки программирования, библиотеки и программные платформы для функционального, логического, объектно-ориентированного программирования приложений систем искусственного интеллекта (Python, R, C++, C#) ПК-2.1. У-1. Умеет разрабатывать	Домашняя работа №1 Домашняя работа №2 Зачет Лабораторные занятия Лекции

		программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#)	
ПК-7. Способен разрабатывать системы анализа больших данных	ПК-7.1. Разрабатывает программные компоненты извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных	<p>ПК-7.1 З-1. Знает общедоступные репозитории и специализированные библиотеки, содержащие наборы больших данных</p> <p>ПК-7.1 З-2. Знает принципы работы экосистемы Hadoop, фреймворка SPARK</p> <p>ПК-7.1 З-3. Знает устройство интерфейсов между реляционными SQL-хранилищами данных и нереляционными NoSQL-хранилищами данных</p> <p>ПК-7.1 З-4. Знает предметно-ориентированные языки</p> <p>ПК-7.1 У-1. Умеет настраивать и оптимизировать конфигурацию программного и аппаратного обеспечения с целью интеграции больших данных</p> <p>ПК-7.1 У-2. Умеет разрабатывать программное обеспечение для очистки и валидации наборов больших данных</p> <p>ПК-7.1 У-3. Умеет выполнять потоковую обработку данных (data streaming, event processing)</p> <p>ПК-7.1 У-4. Умеет использовать шины</p>	<p>Домашняя работа №1</p> <p>Домашняя работа №2</p> <p>Зачет</p> <p>Лабораторные занятия</p> <p>Лекции</p>

		<p>данных (Apache Kafka) ПК-7.1 У-5. Умеет использовать языки запросов, в том числе нереляционных, для поддержки различных типов данных (например, XML, RDF, JSON, мультимедиа) и операций с большими данными (например, матричные операции)</p>	
	<p>ПК-7.2. Разрабатывает программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных</p>	<p>ПК-7.2. 3-1. Знает принципы и методы анализа больших данных, включая спецификации и стандартизацию метаданных ПК-7.2. 3-2. Знает устройство и принципы работы систем обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL) ПК-7.2. 3-3. Знает архитектуру и принципы работы промышленных решений, созданных на основе искусственного интеллекта ПК-7.2. 3-4. Знает методы и технологии машинного обучения на больших данных ПК-7.2. У-1. Умеет разрабатывать программное обеспечение для анализа больших данных ПК-7.2. У-2. Умеет разрабатывать программные и технические средства визуализации больших данных и результатов их анализа ПК-7.2. У-3. Умеет использовать системы обработки и анализа больших массивов данных (SQL, NoSQL, Hadoop, ETL, процессы и</p>	<p>Домашняя работа №1 Домашняя работа №2 Зачет Лабораторные занятия Лекции</p>

		инструменты) ПК-7.2. У-4. Умеет использовать технологии науки о данных и больших данных в разработке для решения практических задач промышленности ПК-7.2. У-5. Умеет описывать и управлять качеством и достоверностью больших данных	
--	--	---	--

3. ПРОЦЕДУРЫ КОНТРОЛЯ И ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ В РАМКАХ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ МОДУЛЯ В БАЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЕ (ТЕХНОЛОГИЧЕСКАЯ КАРТА БРС)

3.1. Процедуры текущей и промежуточной аттестации по дисциплине

1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.6		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>домашняя работа №1</i>	7,7	50
<i>домашняя работа №2</i>	7,14	50
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 0.5		
Промежуточная аттестация по лекциям – зачет		
Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.5		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – не предусмотрено		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям – не предусмотрено		
Промежуточная аттестация по практическим/семинарским занятиям – нет		
Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям – не предусмотрено		
3. Лабораторные занятия: коэффициент значимости совокупных результатов лабораторных занятий – 0.4		
Текущая аттестация на лабораторных занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>отчет по лабораторным работам</i>	7,16	100

Весовой коэффициент значимости результатов текущей аттестации по лабораторным занятиям -1		
Промежуточная аттестация по лабораторным занятиям –нет		
Весовой коэффициент значимости результатов промежуточной аттестации по лабораторным занятиям – не предусмотрено		
4. Онлайн-занятия: коэффициент значимости совокупных результатов онлайн-занятий –не предусмотрено		
Текущая аттестация на онлайн-занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент значимости результатов текущей аттестации по онлайн-занятиям -не предусмотрено		
Промежуточная аттестация по онлайн-занятиям –нет		
Весовой коэффициент значимости результатов промежуточной аттестации по онлайн-занятиям – не предусмотрено		

3.2. Процедуры текущей и промежуточной аттестации курсовой работы/проекта

Текущая аттестация выполнения курсовой работы/проекта	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент текущей аттестации выполнения курсовой работы/проекта– не предусмотрено		
Весовой коэффициент промежуточной аттестации выполнения курсовой работы/проекта– защиты – не предусмотрено		

4. КРИТЕРИИ И УРОВНИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ МОДУЛЯ

4.1. В рамках БРС применяются утвержденные на кафедре/институте критерии (признаки) оценивания достижений студентов по дисциплине модуля (табл. 4) в рамках контрольно-оценочных мероприятий на соответствие указанным в табл.1 результатам обучения (индикаторам).

Таблица 4

Критерии оценивания учебных достижений обучающихся

Результаты обучения	Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам
Знания	Студент демонстрирует знания и понимание в области изучения на уровне указанных индикаторов и необходимые для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Умения	Студент может применять свои знания и понимание в контекстах, представленных в оценочных заданиях, демонстрирует освоение умений на уровне указанных индикаторов и необходимых для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Опыт /владение	Студент демонстрирует опыт в области изучения на уровне указанных индикаторов.
Другие результаты	Студент демонстрирует ответственность в освоении результатов

	<p>обучения на уровне запланированных индикаторов. Студент способен выносить суждения, делать оценки и формулировать выводы в области изучения. Студент может сообщать преподавателю и коллегам своего уровня собственное понимание и умения в области изучения.</p>
--	--

4.2 Для оценивания уровня выполнения критериев (уровня достижений обучающихся при проведении контрольно-оценочных мероприятий по дисциплине модуля) используется универсальная шкала (табл. 5).

Таблица 5

Шкала оценивания достижения результатов обучения (индикаторов) по уровням

Характеристика уровней достижения результатов обучения (индикаторов)				
№ п/п	Содержание уровня выполнения критерия оценивания результатов обучения (выполненное оценочное задание)	Шкала оценивания		
		Традиционная характеристика уровня		Качественная характеристика уровня
1.	Результаты обучения (индикаторы) достигнуты в полном объеме, замечаний нет	Отлично (80-100 баллов)	Зачтено	Высокий (В)
2.	Результаты обучения (индикаторы) в целом достигнуты, имеются замечания, которые не требуют обязательного устранения	Хорошо (60-79 баллов)		Средний (С)
3.	Результаты обучения (индикаторы) достигнуты не в полной мере, есть замечания	Удовлетворительно (40-59 баллов)		Пороговый (П)
4.	Освоение результатов обучения не соответствует индикаторам, имеются существенные ошибки и замечания, требуется доработка	Неудовлетворительно (менее 40 баллов)	Не зачтено	Недостаточный (Н)
5.	Результат обучения не достигнут, задание не выполнено	Недостаточно свидетельств для оценивания		Нет результата

5. СОДЕРЖАНИЕ КОНТРОЛЬНО-ОЦЕНОЧНЫХ МЕРОПРИЯТИЙ ПО ДИСЦИПЛИНЕ МОДУЛЯ

5.1. Описание аудиторных контрольно-оценочных мероприятий по дисциплине модуля

5.1.1. Лекции

Самостоятельное изучение теоретического материала по темам/разделам лекций в соответствии с содержанием дисциплины (п. 1.2. РПД)

5.1.2. Лабораторные занятия

Примерный перечень тем

1. Изучение вариантов запуска Hadoop
 2. Написание простых MapReduce-программ.
 - 2.1 Поиск максимума, минимума среди значений
 - 2.2 Подсчет (агрегация) значений
 - 2.3 Вычисление произведения скалярных векторов
 3. Продвинутое использование возможностей MapReduce
 - 3.1 Фильтрация данных
 - 3.2 Умножение вектора на матрицу
 - 3.3 Комбинирование нескольких задач
 4. Изучение возможностей конфигурирования
 5. Использование возможностей оптимизации
 6. Написание и выполнение Pig-скриптов
 7. Написание и выполнение запросов на языке HiveQL
 8. Работа с ZooKeeper.
 9. Основы практической работы с HBase
 10. Основы практической работы с Kubernetes и Docker.
- LMS-платформа – не предусмотрена

5.2. Описание внеаудиторных контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

Разноуровневое (дифференцированное) обучение.

Базовый

5.2.1. Домашняя работа №1

Примерные перечень тем

1. MapReduce

Примерные задания

1. Установить и запустить демонстрационный пример Hadoop.
2. Реализация и выполнение простых MapReduce, согласно варианту по предоставленным данным.
3. Самостоятельная реализация нескольких MapReduce по собственным данным.
4. Реализация и выполнение составных MapReduce, согласно варианту по предоставленным данным.
5. Самостоятельная реализация составных MapReduce по собственным данным.

LMS-платформа – не предусмотрена

5.2.2. Домашняя работа 2.

Примерный перечень тем

1. Pig и Hive

Примерные задания

1. Написание скриптов Pig, согласно варианту по предоставленным данным

2. Самостоятельное написание скриптов Pig по собственным данным
3. Написание запросов HiveQL, согласно варианту по предоставленным данным
4. Самостоятельное написание запросов HiveQL по собственным данным

LMS-платформа – не предусмотрена

5.3. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля

5.3.1. Зачет

Список примерных вопросов

1. Определение больших данных. Решаемые задачи.
2. Концепция MapReduce.
3. Что такое Hadoop. Архитектура Hadoop.
4. Описание Hadoop Distributed File System.
5. Работа с файлами в системе HDFS.
6. Парадигма и ключевые функции MapReduce в Hadoop
7. Основные этапы обработки данных с помощью MapReduce в Hadoop
8. Использование YARN. Понятия и функции ResourceManager и NodeManager
9. Конфигурирование MapReduce. Основные параметры, на что влияют.
10. Структура простейшей MapReduce программы.
11. Укажите способы сцепления задач MapReduce.
12. Особенности реализации и решения задач на графах в MapReduce.
13. Процедурный язык для выполнения запросов к данным Pig.
14. NoSQL-хранилище класса SQL-on-Hadoop Hive
15. Сравнительная характеристика Pig и Hive.
16. ZooKeeper как сервис для координации распределенных процессов
17. Архитектура ZooKeeper. Клиенты и сессии
18. Понятие NoSQL баз данных. HBase.
19. Основные понятия Kubernetes.
20. Основные понятия Docker.

5.4 Содержание контрольно-оценочных мероприятий по направлениям воспитательной деятельности

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения	Контрольно-оценочные мероприятия
Профессиональное воспитание	профорекционная деятельность	Технология самостоятельной работы Технология анализа образовательных задач	ПК-2	ПК-2.1. У-1	Домашняя работа №1 Домашняя работа №2 Зачет Лабораторные занятия Лекции