



КГУ

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ЭНЕРГЕТИЧЕСКИЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КГУ»)

УТВЕРЖДАЮ

Директор

Института цифровых технологий и
экономики

_____ Э.И. Беляев
«28» ноября 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.11 Разработка систем анализа больших данных

Направление подготовки 09.03.03 Прикладная информатика

Направленность (профиль) Прикладной искусственный интеллект

Квалификация Бакалавр

Казань, 2023

Программу разработал(и):

Наименование кафедры	Должность, уч.степень, уч.звание	ФИО разработчика
УрФУ ИТСУ	Доцент, к.э.н., доцент	Коломыцева А.О.
УрФУ ИТСУ	Ассисент	Павлов М.В.
КГЭУ ИТИС	Доцент, к.ф.-м.н., доцент	Ишмуратов Р.А.

Согласован е	Наименование подразделения	Дата	№ протокола	Подпись
Одобрена	ИТИС	27.11.2023	11	_____ И.о. зав.каф., к.ф.-м.н., доц. Соловьев С.А.
Согласована	Учебно-методический совет ИЦТЭ	27.11.2023	3	_____ Директор, к.т.н., доц. Беляев Э.И.
Одобрена	Ученый совет ИЦТЭ	28.11.2023	3	_____ Директор, к.т.н., доц. Беляев Э.И.

1. Цель, задачи и планируемые результаты обучения по дисциплине

Целью освоения дисциплины «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 дисциплины студент (обучающийся) понимает основные принципы секционирования, сегментирования и перемешивания данных. Обучающийся 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.

Задачами дисциплины являются: приобретение знаний, умений и навыков в разработке и тестировании систем анализа больших данных.

Компетенции и индикаторы, формируемые у обучающихся:

Код и наименование компетенции	Код и наименование индикатора
ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.	ПК-2.1 – Разрабатывает приложения систем искусственного интеллекта.
	ПК-2.2 – Проводит тестирование систем искусственного интеллекта.
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-7.2 – Разрабатывает программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных.
ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих, информационных), осознавая свою роль и ответственность в проекте.	ПК-10.1 – Решает задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством.
	ПК-10.2 – Определяет возможные ограничения ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.
	ПК-10.3 – Определяет свою роль и ответственность в команде при реализации проекта.

2. Место дисциплины в структуре ОП

Предшествующие дисциплины (модули), практики, НИР, др.:

Математические модели и методы; Дискретная математика; Информационные технологии; Программирование; Операционные системы; Объектно-ориентированное программирование; Базы данных; Проектирование информационных систем; Современные языки программирования; Базовая архитектура программного обеспечения; Теория и практика программной инженерии; Аналитика и визуализация данных; Машинное обучение; Анализ данных и искусственный интеллект; Программирование глубоких нейронных сетей.

Последующие дисциплины (модули), практики, НИР, др.: Подготовка к процедуре защиты и защита выпускной квалификационной работы.

3. Структура и содержание дисциплины

3.1. Структура дисциплины

Для очной формы обучения

Вид учебной работы	Всего ЗЕ	Всего часов	Семестр
			7
ОБЩАЯ ТРУДОЕМКОСТЬ ДИСЦИПЛИНЫ	6	216	216
КОНТАКТНАЯ РАБОТА*	-	50	50
АУДИТОРНАЯ РАБОТА		50	50
Лекции		18	18
Практические (семинарские) занятия		-	-
Лабораторные работы		32	32
САМОСТОЯТЕЛЬНАЯ РАБОТА ОБУЧАЮЩЕГОСЯ		130	130
Проработка учебного материала			
Курсовой проект		32	32
Подготовка к промежуточной аттестации		36	36
Промежуточная аттестация:			Э

3.2. Содержание дисциплины, структурированное по разделам и видам занятий

Разделы дисциплины	Всего часов	Распределение трудоемкости по видам учебной работы				Формы и вид контроля	Индексы индикаторов формируемых компетенций
		лекции	лаб. раб.	пр. зан.	сам. раб.		
1. Введение в обработку больших данных	28	2			26	ТК1	ПК-2.1 З,У,В ПК-2.2 З,У,В ПК-7.1 З,У,В ПК-7.2 З,У,В

							ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
2. Основы Hadoop и MapReduce	30	4	8		18	TK2	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
3. Применение MapReduce для решения практических задач	30	4	8		18	TK3	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
4. Основы Pig и Hive	30	4	8		18	TK4	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
5. Модули Hadoop и их использование для разработки Hadoop приложений	30	4	8		18	TK5	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
Курсовой проект (КП)					32	TK6	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
Экзамен	36				36	OM 1	ПК-2.1 3,У,В ПК-2.2 3,У,В ПК-7.1 3,У,В ПК-7.2 3,У,В ПК-10.1 3,У,В ПК-10.2 3,У,В ПК-10.3 3,У,В
ИТОГО	216	18	32		166		

3.3. Содержание дисциплины

Раздел 1. Введение в обработку больших данных

Определение больших данных. Типичные решаемые задачи с помощью больших данных. Концепция MapReduce.

Раздел 2. Основы Hadoop и MapReduce

Hadoop как набор инструментов для разработки систем анализа больших данных. Архитектура Hadoop. Понятие распределенной файловой системы. HDFS (Hadoop Distributed File System).

Парадигма и ключевые функции MapReduce. Этапы процесса обработки данных с помощью MapReduce. Использование YARN. Понятия и функции ResourceManager и NodeManager. Конфигурирование MapReduce.

Раздел 3. Применение MapReduce для решения практических задач

Применение MapReduce для решения практических задач . Структура MapReduce программы. Сцепление задач MapReduce. Соединение данных из разных источников.

Реализация графов в MapReduce. Особенности решения основных задач на графах в MapReduce.

Раздел 4. Основы 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.

Раздел 5. Модули Hadoop и их использование для разработки Hadoop приложений.

Использование ZooKeeper в Hadoop. Необходимость сервиса для координации распределенных процессов. Архитектура ZooKeeper. Клиенты и сессии. Основы ZooKeeper Atomic Broadcast.

Использование NoSQL баз данных . Понятие NoSQL баз данных. Введение в HBase. Архитектура HBase. Работа с HBase в Hadoop.

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.

3.4. Тематический план практических занятий

Данный вид работы не предусмотрен учебным планом.

3.5. Тематический план лабораторных работ

Лабораторная работа 1

Применение прикладных пакетов для решения задач анализа данных.
(8 час).

Лабораторная работа 2. Основы работы Hadoop.

(8 час).

Лабораторная работа 3. Применение MapReduce для решения практических задач.

(8 час).

Лабораторная работа 4.

Модули Hadoop и их использование для разработки Hadoop приложений.
(8 час).

3.6. Курсовой проект /курсовая работа

Учебным планом предусмотрено выполнение курсового проекта (КП).
Тема КП – «Разработка Hadoop приложения».

4. Оценивание результатов обучения

Оценивание результатов обучения по дисциплине осуществляется в рамках текущего контроля (ТК) и промежуточной аттестации (ПА), проводимых по балльно-рейтинговой системе (БРС).

Шкала оценки результатов обучения по дисциплине:

Код компетенции	Код индикатора компетенции	Запланированные результаты обучения по дисциплине	Уровень сформированности индикатора компетенции			
			Высокий	Средний	Ниже среднего	Низкий
			от 85 до 100	от 70 до 84	от 55 до 69	от 0 до 54
			Шкала оценивания			
			отлично	хорошо	удовлетворительно	неудовлетворительно
ПК-2	ПК-2.1	знать:				
		Основы систем искусственного интеллекта.	Высокий уровень знаний основ систем искусственного интеллекта.	Хороший уровень знаний основ систем искусственного интеллекта.	Низкий (не ниже требуемого) уровень знаний основ систем искусственного интеллекта.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
		Разрабатывать систем искусственного интеллекта (ИИ).	Демонстрирует высокое умение разрабатывать системы ИИ	Демонстрирует достаточно хорошее умение разрабатывать системы ИИ	Частично демонстрирует умение разрабатывать системы ИИ	Не умеет разрабатывать системы ИИ
	владеть:					
	Навыками разработки систем искусственного интеллекта.	Глубоко владеет навыками разработки системы ИИ	Демонстрирует хорошие навыки разработки системы ИИ	Посредственно владеет навыками разработки системы ИИ	Не владеет навыками разработки системы ИИ	
ПК-2.2	знать:					
	Основы тестирования систем искусственного интеллекта.	Высокий уровень знаний основ тестирования систем ИИ.	Хороший уровень знаний основ тестирования систем ИИ.	Низкий (не ниже требуемого) уровень знаний основ тестирования систем ИИ.	Уровень знаний неудовлетворительный (ниже требуемого)	

		уметь:				
		Тестировать системы искусственного интеллекта.	Демонстрирует высокое умение тестировать системы ИИ	Демонстрирует достаточно хорошее умение тестировать системы ИИ.	Частично демонстрирует умение тестировать системы ИИ	Не умеет тестировать системы ИИ
		владеть:				
		Навыками тестирования системы искусственного интеллекта.	Глубоко владеет навыками тестировать системы ИИ	Демонстрирует хорошие навыки тестировать системы ИИ	Посредственно владеет навыками тестировать системы ИИ	Не владеет навыками тестировать системы ИИ
ПК-7	ПК-7.1	знать:				
		Основы программных компонентов извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Высокий уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Хороший уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Низкий (не ниже требуемого) уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
		Разрабатывать программные компоненты извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Демонстрирует высокое умение разрабатывать программные компоненты извлечения, хранения, подготовки больших данных	Демонстрирует достаточно хорошее умение программные компоненты извлечения, хранения, подготовки больших данных.	Частично демонстрирует умение программные компоненты извлечения, хранения, подготовки больших данных.	Не умеет разрабатывать программные компоненты извлечения, хранения, подготовки больших данных.
		владеть:				
		Навыками разработки программных компонентов извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Глубоко владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.	Демонстрирует хорошие навыки разработки программных компонентов извлечения, хранения, подготовки больших данных.	Посредственно владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.	Не владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.
	ПК-7.2	знать:				
		Основы разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и	Высокий уровень знаний основ разработки программных компонентов обработки, удаленной,	Хороший уровень знаний основ тестирования разработки программных компонентов обработки,	Низкий (не ниже требуемого) уровень знаний основ разработки программных компонентов	Уровень знаний неудовлетворительный (ниже требуемого)

		достоверностью больших данных.	распределенной и объединенной аналитики.	удаленной, распределенной и объединенной аналитики.	обработки, удаленной, распределенной и объединенной аналитики	
		уметь:				
		Разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных..	Демонстрирует высокое умение разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики.	Демонстрирует достаточно хорошее умение разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики.	Частично демонстрирует умение разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики.	Не умеет разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики.
		владеть:				
		Навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных.	Глубоко владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Демонстрирует хорошие навыки разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Посредственно владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Не владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.
ПК-10	ПК-10.1	знать:				
		Основные принципы решения задач профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством.	Высокий уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Хороший уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Низкий (не ниже требуемого) уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
			Демонстрирует высокое умение решать задачи профессиональной деятельности в проектном формате для	Демонстрирует достаточно хорошее умение решать задачи профессиональной деятельности в проектном формате для	Частично демонстрирует умение решать задачи профессиональной деятельности в проектном формате для	Не умеет решать задачи профессиональной деятельности в проектном формате для достижения заданной

		достижения заданной цели.	достижения заданной цели.	достижения заданной цели.	цели.
	владеть:				
	Навыками	Глубоко владеет навыками решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Демонстрирует хорошие навыки решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Посредственно владеет навыками решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Не владеет навыками решения задач профессиональной деятельности в проектном формате для достижения заданной цели.
ПК-10.2	знать:				
	Определяет возможные ограничения ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Высокий уровень знаний возможных ограничений ресурсов в проектной деятельности.	Хороший уровень знаний возможных ограничений ресурсов в проектной деятельности.	Низкий (не ниже требуемого) уровень знаний возможных ограничений ресурсов в проектной деятельности.	Уровень знаний неудовлетворительный (ниже требуемого)
	уметь:				
	Определять возможные ограничения ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Демонстрирует высокое умение определять возможные ограничения ресурсов в проектной деятельности.	Демонстрирует достаточно хорошее умение определять возможные ограничения ресурсов в проектной деятельности.	Частично демонстрирует умение определять возможные ограничения ресурсов в проектной деятельности.	Не умеет определять возможные ограничения ресурсов в проектной деятельности.
ПК-10.3	владеть:				
	Навыками определения возможных ограничений ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Глубоко владеет навыками определения возможных ограничений ресурсов в проектной деятельности.	Демонстрирует хорошие навыки определения возможных ограничений ресурсов в проектной деятельности.	Посредственно владеет навыками определения возможных ограничений ресурсов в проектной деятельности.	Не владеет навыками определения возможных ограничений ресурсов в проектной деятельности.
	знать:				
	Свою роль и ответственность в команде при реализации проекта.	Высокий уровень знаний своей роли и ответственности в команде при реализации проекта.	Хороший уровень знаний своей роли и ответственности в команде при реализации проекта.	Низкий (не ниже требуемого) уровень знаний своей роли и ответственности в команде при реализации проекта.	Уровень знаний неудовлетворительный (ниже требуемого)

		уметь:				
	Определять свою роль и ответственность в команде при реализации проекта.	Демонстрирует высокое умение определять свою роль и ответственность в команде при реализации проекта.	Демонстрирует достаточно хорошее умение определять свою роль и ответственность в команде при реализации проекта.	Частично демонстрирует умение определять свою роль и ответственность в команде при реализации проекта.	Не умеет определять свою роль и ответственность в команде при реализации проекта.	
		владеть:				
	Определения своей роли и ответственности в команде при реализации проекта.	Глубоко владеет навыками определения своей роли и ответственности в команде при реализации проекта.	Демонстрирует хорошие навыки определения своей роли и ответственности в команде при реализации проекта.	Посредственно владеет навыками определения своей роли и ответственности в команде при реализации проекта.	Не владеет навыками определения своей роли и ответственности в команде при реализации проекта.	

Оценочные материалы для проведения текущего контроля и промежуточной аттестации приведены в Приложении к рабочей программе дисциплины.

Полный комплект заданий и материалов, необходимых для оценивания результатов обучения по дисциплине, хранится на кафедре разработчика.

5. Учебно-методическое и информационное обеспечение дисциплины

5.1. Учебно-методическое обеспечение

5.1.1. Основная литература

1. Анализ больших данных: учебное пособие / И.Б. Тесленко, В.Е. Крылов, А.М. Губернаторов [и др.]. – Москва : КноРус, 2023. – 295 с. – ISBN 978-5-406-10550-4. – URL: <https://book.ru/book/950469> – Текст: электронный.

2. Макшанов, А. В. Большие данные. Big Data / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. – 4-е изд., стер. – Санкт-Петербург : Лань, 2024. – 188 с. – ISBN 978-5-507-47346-5. – Текст: электронный // Лань : электронно-библиотечная система. – URL: <https://e.lanbook.com/book/362318>.

3. Чернышев, С.А., Алгоритмы и структуры данных на языке GO : учебник / С.А. Чернышев. – Москва : КноРус, 2024. – 353 с. – ISBN 978-5-406-11685-2. – URL: <https://book.ru/book/949702> – Текст: электронный.

4. Лесковец Юре. Анализ больших наборов данных. / Пер. с англ. А.А. Слинкин / Д. Джеффри. – Москва: ДМК Пресс, 2016. – 498 с. – ISBN 978-5-97060-190-7. – URL: <https://ibooks.ru/bookshelf/364297/> – Текст: электронный.

5.1.2.Дополнительная литература

1. Ясницкий Л.Н. Интеллектуальные системы: учебник. – 2-е изд., электрон. – (Учебник для высшей школы) / Л.Н. Ясницкий. – Москва: Лаборатория знаний, 2020. – 224с. – ISBN 978-5-00101-417-1. – URL: <https://ibooks.ru/bookshelf/372666> – Текст: электронный.

2. Чернышев, С. А., Алгоритмы и структуры данных на Python : учебное пособие / С. А. Чернышев. — Москва : КноРус, 2024. – 326 с. – ISBN 978-5-406-11683-8. – URL: <https://book.ru/book/949701> – Текст: электронный.

3. Заботина Н.Н. Проектирование информационных систем / Н.Н. Заботина. – Москва: Инфра-М, 2020. – 331 с. – ISBN 978-5-16-004509-2. – URL: <https://ibooks.ru/bookshelf/361556> – Текст: электронный.

4. Одинцов И. Профессиональное программирование. Системный подход, 2 изд. / И. Одинцов. - Санкт-Петербург : БХВ-Петербург, 2014. – 624 с. – ISBN 5-94157-457-6. – URL: <https://ibooks.ru/bookshelf/18535> – Текст: электронный.

5. Дадян Э.Г. Данные: хранение и обработка / Э.Г. Дадян. – Москва : Инфра-М, 2021. – 205с. – ISBN 978-5-16-016447-2. – URL: <https://ibooks.ru/bookshelf/377999> – Текст: электронный.

5.2. Информационное обеспечение

5.2.1. Электронные и интернет-ресурсы

№ п/п	Наименование электронных и интернет-ресурсов	Ссылка
1	Электронный университет КГЭУ - виртуальная образовательная среда	Среда электронного обучения LMS Moodle
2	ИНТУИТ. Национальный Открытый Университет	https://www.intuit.ru/
3	Гудов, А. М.; Базы данных и системы управления базами данных. Программирование на языке PL/SQL : учебное пособие.; Кемеровский государственный университет, Кемерово; 2010; (Электронное издание)	https://biblioclub.ru/index.php?page=book&id=232497

5.2.2. Профессиональные базы данных / Информационно-справочные системы

№ п/п	Наименование профессиональных баз данных	Адрес	Режим доступа
1.	Российская национальная библиотека	http://nlr.ru/	открытый
2.	Российская государственная библиотека	https://www.rsl.ru/	открытый
3.	Государственная публичная научно-техническая библиотека	http://www.gpntb.ru	открытый
4.	Национальная электронная библиотека (НЭБ)	https://rusneb.ru/	открытый
5.	Техническая библиотека	http://techlibrary.ru	открытый
6.	Образовательный портал	http://www.ucheba.com	открытый

5.2.3. Лицензионное и свободно распространяемое программное обеспечение дисциплины

№ п/п	Наименование программного обеспечения	Описание	Реквизиты подтверждающих документов
1	Windows 7 Профессиональная	Пользовательская операционная система	ЗАО "СофтЛайнТрейд" №2011.25486 от 28.11.2011
2	Windows 7 Профессиональная (Pro)	Пользовательская операционная система	ЗАО "СофтЛайнТрейд" №2011.25486 от 28.11.2011
3	Браузер Chrome	Система поиска информации в сети интернет (включая русскоязычный интернет).	Свободная лицензия Неискл. право. Бессрочно.
4	Браузер Firefox	Свободный веб-браузер	Свободная лицензия Неискл. право. Бессрочно.
5	OpenOffice	Пакет офисных приложений. Одним из первых стал поддерживать новый открытый	Свободная лицензия Неискл. право. Бессрочно.

6	LMS Moodle	Это современное программное обеспечение	Свободная лицензия Неискл. право. Бессрочно.
---	------------	---	--

6. Материально-техническое обеспечение дисциплины

Наименование вида учебной работы	Наименование учебной аудитории, специализированной лаборатории	Перечень необходимого оборудования и технических средств обучения
Лекции	Учебная аудитория для проведения занятий лекционного типа	Специализированная учебная мебель, технические средства обучения, служащие для представления учебной информации большой аудитории (мультимедийный проектор, компьютер (ноутбук), экран), демонстрационное оборудование, учебно-наглядные пособия
Лабораторные работы	Компьютерный класс с выходом в Интернет В-608	доска аудиторная, персональный компьютер (25 шт.), интерактивная доска, мультимедийный проектор
	Компьютерный класс с выходом в Интернет В-600б	Специализированная учебная мебель на 30 посадочных мест, 30 компьютеров, технические средства обучения (мультимедийный проектор, компьютер (ноутбук), экран), видеокамеры, программное обеспечение
Самостоятельная работа	Компьютерный класс с выходом в Интернет В-617	Специализированная учебная мебель на 30 посадочных мест, 25 компьютеров, программное обеспечение
	Читальный зал библиотеки	Специализированная мебель, компьютерная техника с возможностью выхода в Интернет и обеспечением доступа в ЭИОС, экран, мультимедийный проектор, программное обеспечение

7. Особенности организации образовательной деятельности для лиц с ограниченными возможностями здоровья и инвалидов

Лица с ограниченными возможностями здоровья (ОВЗ) и инвалиды имеют возможность беспрепятственно перемещаться из одного учебно-лабораторного корпуса в другой, подняться на все этажи учебно-лабораторных корпусов, заниматься в учебных и иных помещениях с учетом особенностей психофизического развития и состояния здоровья.

Для обучения лиц с ОВЗ и инвалидов, имеющих нарушения опорно-двигательного аппарата, обеспечены условия беспрепятственного доступа во все учебные помещения. Информация о специальных условиях, созданных для обучающихся с ОВЗ и инвалидов, размещена на сайте университета www/kgeu.ru. Имеется возможность оказания технической помощи ассистентом, а также услуг сурдопереводчиков и тифлосурдопереводчиков.

Для адаптации к восприятию лицами с ОВЗ и инвалидами с нарушенным

слухом справочного, учебного материала по дисциплине обеспечиваются следующие условия:

- для лучшей ориентации в аудитории, применяются сигналы оповещения о начале и конце занятия (слово «звонок» пишется на доске);
- внимание слабослышащего обучающегося привлекается педагогом жестом (на плечо кладется рука, осуществляется нерезкое похлопывание);
- разговаривая с обучающимся, педагогический работник смотрит на него, говорит ясно, короткими предложениями, обеспечивая возможность чтения по губам.

Компенсация затруднений речевого и интеллектуального развития слабослышащих обучающихся проводится путем:

- использования схем, диаграмм, рисунков, компьютерных презентаций с гиперссылками, комментирующими отдельные компоненты изображения;
- регулярного применения упражнений на графическое выделение существенных признаков предметов и явлений;
- обеспечения возможности для обучающегося получить адресную консультацию по электронной почте по мере необходимости.

Для адаптации к восприятию лицами с ОВЗ и инвалидами с нарушениями зрения справочного, учебного, просветительского материала, предусмотренного образовательной программой по выбранному направлению подготовки, обеспечиваются следующие условия:

- ведется адаптация официального сайта в сети Интернет с учетом особых потребностей инвалидов по зрению, обеспечивается наличие крупношрифтовой справочной информации о расписании учебных занятий;
- педагогический работник, его собеседник (при необходимости), присутствующие на занятии, представляются обучающимся, при этом каждый раз называется тот, к кому педагогический работник обращается;
- действия, жесты, перемещения педагогического работника коротко и ясно комментируются;
- печатная информация предоставляется крупным шрифтом (от 18 пунктов), тотально озвучивается;
- обеспечивается необходимый уровень освещенности помещений;
- предоставляется возможность использовать компьютеры во время занятий и право записи объяснений на диктофон (по желанию обучающихся).

Форма проведения текущей и промежуточной аттестации для обучающихся с ОВЗ и инвалидов определяется педагогическим работником в соответствии с учебным планом. При необходимости обучающемуся с ОВЗ, инвалиду с учетом их индивидуальных психофизических особенностей дается возможность пройти промежуточную аттестацию устно, письменно на бумаге, письменно на компьютере, в форме тестирования и т.п., либо предоставляется дополнительное время для подготовки ответа.

8. Методические рекомендации для преподавателей по организации воспитательной работы с обучающимися

Методическое обеспечение процесса воспитания обучающихся выступает одним из определяющих факторов высокого качества образования. Преподаватель вуза, демонстрируя высокий профессионализм, эрудицию, четкую гражданскую позицию, самодисциплину, творческий подход в решении профессиональных задач, в ходе образовательного процесса способствует формированию гармоничной личности.

При реализации дисциплины преподаватель может использовать следующие методы воспитательной работы:

- методы формирования сознания личности (беседа, диспут, внушение, инструктаж, контроль, объяснение, пример, самоконтроль, рассказ, совет, убеждение и др.);

- методы организации деятельности и формирования опыта поведения (задание, общественное мнение, педагогическое требование, поручение, приучение, создание воспитывающих ситуаций, тренинг, упражнение, и др.);

- методы мотивации деятельности и поведения (одобрение, поощрение социальной активности, порицание, создание ситуаций успеха, создание ситуаций для эмоционально–нравственных переживаний, соревнование и др.)

При реализации дисциплины преподаватель должен учитывать следующие направления воспитательной деятельности:

Гражданское и патриотическое воспитание:

- формирование у обучающихся целостного мировоззрения, российской идентичности, уважения к своей семье, обществу, государству, принятым в семье и обществе духовно–нравственным и социокультурным ценностям, к национальному, культурному и историческому наследию, формирование стремления к его сохранению и развитию;

- формирование у обучающихся активной гражданской позиции, основанной на традиционных культурных, духовных и нравственных ценностях российского общества, для повышения способности ответственно реализовывать свои конституционные права и обязанности;

- развитие правовой и политической культуры обучающихся, расширение конструктивного участия в принятии решений, затрагивающих их права и интересы, в том числе в различных формах самоорганизации, самоуправления, общественно–значимой деятельности;

- формирование мотивов, нравственных и смысловых установок личности, позволяющих противостоять экстремизму, ксенофобии, дискриминации по социальным, религиозным, расовым, национальным признакам, межэтнической и межконфессиональной нетерпимости, другим негативным социальным явлениям.

Духовно–нравственное воспитание:

- воспитание чувства достоинства, чести и честности, совестливости, уважения к родителям, учителям, людям старшего поколения;

- формирование принципов коллективизма и солидарности, духа милосердия и сострадания, привычки заботиться о людях, находящихся в трудной жизненной ситуации;

- формирование солидарности и чувства социальной ответственности по

отношению к людям с ограниченными возможностями здоровья, преодоление психологических барьеров по отношению к людям с ограниченными возможностями;

- формирование эмоционально насыщенного и духовно возвышенного отношения к миру, способности и умения передавать другим свой эстетический опыт.

Культурно-просветительское воспитание:

– формирование эстетической картины мира;

– формирование уважения к культурным ценностям родного города, края, страны;

– повышение познавательной активности обучающихся.

Научно-образовательное воспитание:

– формирование у обучающихся научного мировоззрения;

– формирование умения получать знания;

– формирование навыков анализа и синтеза информации, в том числе в профессиональной области.

Вносимые изменения и утверждения на новый учебный год

№ п/п	№ раздела внесения изменений	Дата внесения изменений	Содержание изменений	«Согласовано» Зав. каф. реализующей дисциплину	«Согласовано» председатель УМК института (факультета), в состав которого входит выпускающая
1	2	3	4	5	6
1					
2					
3					

*Приложение к рабочей
программе дисциплины*



КГУ

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ЭНЕРГЕТИЧЕСКИЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «КГУ»)**

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
по дисциплине**

Б1.В.11 Разработка систем анализа больших данных

г. Казань, 2023

2. Оценочные материалы текущего контроля и промежуточной аттестации
Шкала оценки результатов обучения по дисциплине:

Код компетенции	Код индикатора компетенции	Запланированные результаты обучения по дисциплине	Уровень сформированности индикатора компетенции			
			Высокий	Средний	Ниже среднего	Низкий
			от 85 до 100	от 70 до 84	от 55 до 69	от 0 до 54
			Шкала оценивания			
			отлично	хорошо	удовлетворительно	неудовлетворительно
ПК-2	ПК-2.1	знать:				
		Основы систем искусственного интеллекта.	Высокий уровень знаний основ систем искусственного интеллекта.	Хороший уровень знаний основ систем искусственного интеллекта.	Низкий (не ниже требуемого) уровень знаний основ систем искусственного интеллекта.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
		Разрабатывать систем искусственного интеллекта (ИИ).	Демонстрирует высокое умение разрабатывать системы ИИ	Демонстрирует достаточно хорошее умение разрабатывать системы ИИ	Частично демонстрирует умение разрабатывать системы ИИ	Не умеет разрабатывать системы ИИ
		владеть:				
	Навыками разработки систем искусственного интеллекта.	Глубоко владеет навыками разработки системы ИИ	Демонстрирует хорошие навыки разработки системы ИИ	Посредственно владеет навыками разработки системы ИИ	Не владеет навыками разработки системы ИИ	
	ПК-2.2	знать:				
		Основы тестирования систем искусственного интеллекта.	Высокий уровень знаний основ тестирования систем ИИ.	Хороший уровень знаний основ тестирования систем ИИ.	Низкий (не ниже требуемого) уровень знаний основ тестирования систем ИИ.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
		Тестировать системы искусственного интеллекта.	Демонстрирует высокое умение тестировать системы ИИ	Демонстрирует достаточно хорошее умение тестировать системы ИИ.	Частично демонстрирует умение тестировать системы ИИ	Не умеет тестировать системы ИИ
владеть:						
Навыками тестирования системы искусственного интеллекта.	Глубоко владеет навыками тестировать системы ИИ	Демонстрирует хорошие навыки тестировать системы ИИ	Посредственно владеет навыками тестировать системы ИИ	Не владеет навыками тестировать системы ИИ		

ПК-7	ПК-7.1	знать:				
		Основы программных компонентов извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Высокий уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Хороший уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Низкий (не ниже требуемого) уровень знаний основ программных компонентов извлечения, хранения, подготовки больших данных.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
		Разрабатывать программные компоненты извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Демонстрирует высокое умение разрабатывать программные компоненты извлечения, хранения, подготовки больших данных	Демонстрирует достаточно хорошее умение программные компоненты извлечения, хранения, подготовки больших данных.	Частично демонстрирует умение программные компоненты извлечения, хранения, подготовки больших данных.	Не умеет разрабатывать программные компоненты извлечения, хранения, подготовки больших данных.
	владеть:					
	Навыками разработки программных компонентов извлечения, хранения, подготовки больших данных с учетом вариантов использования больших данных, определений, словарей и эталонной архитектуры больших данных.	Глубоко владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.	Демонстрирует хорошие навыки разработки программных компонентов извлечения, хранения, подготовки больших данных.	Посредственно владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.	Не владеет навыками разработки программных компонентов извлечения, хранения, подготовки больших данных.	
	ПК-7.2	знать:				
		Основы разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных.	Высокий уровень знаний основ разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Хороший уровень знаний основ тестирования разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Низкий (не ниже требуемого) уровень знаний основ разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
	Разрабатывать программные компоненты обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и	Демонстрирует высокое умение разрабатывать программные компоненты обработки,	Демонстрирует достаточно хорошее умение разрабатывать программные компоненты обработки,	Частично демонстрирует умение разрабатывать программные компоненты обработки,	Не умеет разрабатывать программные компоненты обработки, распределенной	

		управления качеством и достоверностью больших данных..	удаленной, распределенной и объединенной аналитики.	удаленной, распределенной и объединенной аналитики.	удаленной, распределенной и объединенной аналитики.	ой и объединенной аналитики.
		владеть:				
		Навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики, использования результатов анализа, описания и управления качеством и достоверностью больших данных.	Глубоко владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Демонстрирует хорошие навыки разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Посредственно владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.	Не владеет навыками разработки программных компонентов обработки, удаленной, распределенной и объединенной аналитики.
ПК-10	ПК-10.1	знать:				
		Основные принципы решения задач профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством.	Высокий уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Хороший уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Низкий (не ниже требуемого) уровень знаний основных принципов решения задач профессиональной деятельности в проектном формате для достижения заданной цели.	Уровень знаний неудовлетворительный (ниже требуемого)
		уметь:				
			Демонстрирует высокое умение решать задачи профессиональной деятельности в проектном формате для достижения заданной цели.	Демонстрирует достаточно хорошее умение решать задачи профессиональной деятельности в проектном формате для достижения заданной цели.	Частично демонстрирует умение решать задачи профессиональной деятельности в проектном формате для достижения заданной цели.	Не умеет решать задачи профессиональной деятельности в проектном формате для достижения заданной цели.
		владеть:				
		Навыками	Глубоко владеет навыками решения задач профессиональной деятельности в проектном формате для достижения	Демонстрирует хорошие навыки решения задач профессиональной деятельности в проектном формате для достижения	Посредственно владеет навыками решения задач профессиональной деятельности в проектном формате для достижения	Не владеет навыками решения задач профессиональной деятельности в проектном формате для достижения

		заданной цели.	заданной цели.	заданной цели.	цели.
ПК-10.2	знать:				
	Определяет возможные ограничения ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Высокий уровень знаний возможных ограничений ресурсов в проектной деятельности.	Хороший уровень знаний возможных ограничений ресурсов в проектной деятельности.	Низкий (не ниже требуемого) уровень знаний возможных ограничений ресурсов в проектной деятельности.	Уровень знаний неудовлетворительный (ниже требуемого)
	уметь:				
	Определять возможные ограничения ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Демонстрирует высокое умение определять возможные ограничения ресурсов в проектной деятельности.	Демонстрирует достаточно хорошее умение определять возможные ограничения ресурсов в проектной деятельности.	Частично демонстрирует умение определять возможные ограничения ресурсов в проектной деятельности.	Не умеет определять возможные ограничения ресурсов в проектной деятельности.
ПК-10.3	владеть:				
	Навыками определения возможных ограничений ресурсов (временных, финансовых, информационных и человеческих) в проектной деятельности.	Глубоко владеет навыками определения возможных ограничений ресурсов в проектной деятельности.	Демонстрирует хорошие навыки определения возможных ограничений ресурсов в проектной деятельности.	Посредственно владеет навыками определения возможных ограничений ресурсов в проектной деятельности.	Не владеет навыками определения возможных ограничений ресурсов в проектной деятельности.
	знать:				
ПК-10.3	Свою роль и ответственность в команде при реализации проекта.	Высокий уровень знаний своей роли и ответственности в команде при реализации проекта.	Хороший уровень знаний своей роли и ответственности в команде при реализации проекта.	Низкий (не ниже требуемого) уровень знаний своей роли и ответственности в команде при реализации проекта.	Уровень знаний неудовлетворительный (ниже требуемого)
	уметь:				
	Определять свою роль и ответственность в команде при реализации проекта.	Демонстрирует высокое умение определять свою роль и ответственность в команде при реализации проекта.	Демонстрирует достаточно хорошее умение определять свою роль и ответственность в команде при реализации проекта.	Частично демонстрирует умение определять свою роль и ответственность в команде при реализации проекта.	Не умеет определять свою роль и ответственность в команде при реализации проекта.
ПК-10.3	владеть:				
	Определения своей роли и ответственности в команде	Глубоко владеет	Демонстрирует хорошие	Посредственно владеет	Не владеет навыками

		при реализации проекта.	навыками определения своей роли и ответственности в команде при реализации проекта.	навыки определения своей роли и ответственности в команде при реализации проекта.	навыками определения своей роли и ответственности в команде при реализации проекта.	определения своей роли и ответственности в команде при реализации проекта.
--	--	-------------------------	---	---	---	--

Оценка **«отлично»** выставляется за выполнение *расчетных работ в семестре; тестовых заданий; глубокое понимание технологических методов расчета норм расхода материалов, полные и содержательные ответы на вопросы билета (теоретическое и практическое задание);*

Оценка **«хорошо»** выставляется за выполнение *расчетных работ в семестре; тестовых заданий; понимание технологических методов расчета норм расхода материалов, ответы на вопросы билета (теоретическое или практическое задание);*

Оценка **«удовлетворительно»** выставляется за выполнение *расчетных работ в семестре и тестовых заданий;*

Оценка **«неудовлетворительно»** выставляется за слабое и неполное выполнение *расчетных работ в семестре и тестовых заданий.*

3. Перечень оценочных средств

Краткая характеристика оценочных средств, используемых при текущем контроле успеваемости и промежуточной аттестации обучающегося по дисциплине:

Наименование оценочного средства	Краткая характеристика оценочного средства	Описание оценочного средства
Отчет по лабораторной работе (ОЛР)	Выполнение лабораторной работы заканчивается представлением отчета. Результатом выполнения лабораторной работы может быть файл с выполненными заданиями, прикрепленный в электронную среду MOODL или задания, выполненные на лабораторной работе и представленные на проверку преподавателю.	Перечень заданий для защиты лабораторной работы
Тест (Тест)	Система стандартизированных заданий, позволяющая автоматизировать процедуру измерения уровня знаний и умений обучающегося	Комплект тестовых заданий

4. Перечень контрольных заданий или иные материалы, необходимые для оценки знаний, умений и навыков, характеризующих этапы формирования компетенций в процессе освоения дисциплины

Пример задания

Для текущего контроля ТК1:

Проверяемые компетенции:

ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.

ПК-7 – Способен разрабатывать системы анализа больших данных.

ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального

продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих, информационных), осознавая свою роль и ответственность в проекте.

Контрольные вопросы:

1. Какую дату можно считать точкой отсчета BigData?
2. Как можно определить понятие больших данных?
3. Какие составляющие входят в описание больших данных через три “V”?
4. Какие составляющие дополняют определение трех “V”?
5. Какие сферы применения больших данных можно выделить?
6. Какие проблемы могут возникать при работе с большими данными?
7. Что из себя представляет модель MapReduce?
8. Каковы основные этапы MapReduce и их содержание?

Источники по теме для СРС:

1. Большие данные. Подготовка волнорезов. <https://habr.com/ru/articles/290714/>
2. Что такое «Big Data»? <https://habr.com/ru/companies/productstar/articles/503580/>

Для текущего контроля ТК2:

Проверяемые компетенции:

ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.

ПК-7 – Способен разрабатывать системы анализа больших данных.

ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих, информационных), осознавая свою роль и ответственность в проекте.

Лабораторная работа 1. Применение прикладных пакетов для решения задач анализа данных.

Примеры заданий:

Решить предлагаемую задачу по анализу данных с использованием программных средств и инструментов предлагаемого прикладного программного пакета.

Контрольные вопросы:

1. Что такое Hadoop?
2. Какие возможности предоставляет Hadoop?
3. Какие основные компоненты входят в Hadoop?
4. Какой компонент Hadoop непосредственно отвечает за обработку данных?

5. Какие два компонента предоставляют возможность для доступа к данным, хранимым в Hadoop?

6. Каковы основные отличия хранения и обработки данных в Hadoop и в реляционной базе данных?

7. Каковы принципы работы заложенные в распределенной файловой системе Hadoop?

8. Какие возможности предоставляет HDFS?

9. Каковы основные компоненты HDFS?

Для текущего контроля ТК3:

Проверяемые компетенции:

ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.

ПК-7 – Способен разрабатывать системы анализа больших данных.

ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих, информационных), осознавая свою роль и ответственность в проекте.

Лабораторная работа 2. Основы работы Hadoop.

Примеры заданий:

Изучить основы работы Hadoop. Решить предлагаемую задачу по анализу данных с использованием программных средств и инструментов Hadoop.

Контрольные вопросы

1. Что представляет собой MapReduce?

2. Какие особенности MapReduce обуславливают ее применение для обработки больших данных?

3. Какие основные этапы выделяются в модели MapReduce?

4. Какие типы данных используются в Hadoop?

5. Что необходимо сделать, чтобы реализовать пользовательский тип данных в Hadoop?

6. Что необходимо сделать, для определения собственного распределителя? редуктора?

Для текущего контроля ТК4:

Проверяемые компетенции:

ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.

ПК-7 – Способен разрабатывать системы анализа больших данных.

ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих,

информационных), осознавая свою роль и ответственность в проекте.

Лабораторная работа 3. Применение MapReduce для решения практических задач.

Примеры заданий:

Изучить основы работы MapReduce. Решить предлагаемую практическую задачу по анализу данных с использованием программных средств и инструментов MapReduce.

Для текущего контроля ТК5:

Проверяемые компетенции:

ПК-2 – Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта.

ПК-7 – Способен разрабатывать системы анализа больших данных.

ПК-10 – Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов (временных, финансовых, человеческих, информационных), осознавая свою роль и ответственность в проекте.

Лабораторная работа 4. Модули Hadoop и их использование для разработки Hadoop приложений.

Примеры заданий:

Изучить основы работы с модулями Hadoop. Разработать приложение Hadoop с использованием программных средств и инструментов модулей Hadoop.

Контрольные вопросы:

1. Что такое Pig в Hadoop?
2. Чем отличается MapReduce программа от скрипта Pig Latin?
3. Какие основные типы данных существуют в Pig Latin?
4. С какими типами данных можно работать, используя Pig?
5. Какие основные операторы имеются в Pig Latin?
6. Что такое Hive в Hadoop?
- 44
7. Какие особенности имеет Hive, относительно типов данных, с которыми можно работать?
8. Что позволяет, а что нет Hive по сравнению с использованием SQL в реляционных базах данных?

Для промежуточной аттестации:

Вопросы для подготовки к экзамену.

1. История развития понятия и технологии больших данных.

2. Классификация BigData.
3. Основные составляющие описания больших данных (определение три “V”).
4. Дополнительные составляющие определения трех “V”.
5. Области применения больших данных.
6. Особенности и ограничения при работе с большими данными.
7. Основы Hadoop.
8. Возможности Hadoop.
9. Основные компоненты Hadoop.
10. Типы данных, которые используются в Hadoop.
11. Компонент Hadoop, непосредственно ответственный за обработку данных.
12. Компоненты, предоставляющие возможность для доступа к данным, хранимым в Hadoop.
13. Основные отличия хранения и обработки данных в Hadoop и в реляционной базе данных.
14. Принципы работы, заложенные в распределенной файловой системе Hadoop.
15. Реализация пользовательского типа данных в Hadoop.
16. Возможности HDFS.
17. Основные компоненты HDFS.
18. Модель MapReduce. Основные сведения.
19. Основные этапы модели MapReduce.
20. Возможности MapReduce для обработки больших данных.
21. Определения собственного распределителя.
22. Определения собственного редуктора.
23. Основы Pig в Hadoop.
24. Отличие MapReduce программы от скрипта Pig Latin.
25. Основные типы данных Pig Latin.
26. Типы данных Pig.
27. Основные операторы Pig Latin.
28. Основы Hive в Hadoop.
29. Особенности Hive, которые связаны с работой с типами данных.
30. Отличие Hive по сравнению с использованием SQL в реляционных базах данных.