

МИНОБРНАУКИ РОССИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Рабочая программа дисциплины (модуля)
ТЕХНОЛОГИИ BIG DATA

Направление и направленность (профиль)
09.04.03 Прикладная информатика. Искусственный интеллект и машинное обучение в
управлении и принятии решений

Год набора на ОПОП
2022

Форма обучения
очная

Владивосток 2023

Рабочая программа дисциплины (модуля) «Технологии Big Data» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика (утв. приказом Минобрнауки России от 19.09.2017г. №916) и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 06.04.2021 г. N245).

Составитель(и):

Шахгельдян К.И., доктор технических наук, профессор, Кафедра информационных технологий и систем, carina.shahgeldyan@vvsu.ru

Утверждена на заседании кафедры информационных технологий и систем от 31.05.2023 , протокол № 9

СОГЛАСОВАНО:

Заведующий кафедрой (разработчика)

Кийкова Е.В.

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ	
Сертификат	1575633692
Номер транзакции	0000000000BBE2F0
Владелец	Кийкова Е.В.

1 Цель, планируемые результаты обучения по дисциплине (модулю)

Целью дисциплины является знакомство с технологиями больших данных и получение начальных навыков развертывания инфраструктуры Big data.

Задачи освоения дисциплины:

1. Познакомиться с основными открытыми технологиями больших данных
2. Получить начальные навыки развертывания и сопровождения инфраструктуры Big Data

Планируемыми результатами обучения по дисциплине (модулю), являются знания, умения, навыки. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы, представлен в таблице 1.

Таблица 1 – Компетенции, формируемые в результате изучения дисциплины (модуля)

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине		
			Код результата	Формулировка результата	
09.04.03 «Прикладная информатика» (М-ПИ)	ОПК-2 : Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ОПК-2.1к : Решает профессиональные задачи, используя современные интеллектуальные технологии	РД1	Знание	Базовые технологии больших данных
		ОПК-2.2к : Обосновывает выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач	РД2	Умение	Выбор подходящих технологий
	ОПК-5 : Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем	ОПК-5.1к : Разрабатывает и исследует программно-аппаратные комплексы автоматизированных информационных систем	РД3	Знание	Архитектуры инфраструктуры хранения и обработки больших данных
			РД4	Умение	Построение инфраструктуры больших данных
			РД5	Навык	Базовая настройка компонентов инфраструктуры больших данных

2 Место дисциплины (модуля) в структуре ОПОП

3. Объем дисциплины (модуля)

Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися (по видам учебных занятий) и на самостоятельную работу, приведен в таблице 2.

Таблица 2 – Общая трудоемкость дисциплины

Название ОПОП ВО	Форма обучения	Часть УП	Семестр (ОФО) или курс (ЗФО, ОЗФО)	Трудо-емкость	Объем контактной работы (час)					СРС	Форма аттес-тации	
					(З.Е.)	Всего	Аудиторная					Внеауди-торная
				лек.			прак.	лаб.	ПА			КСР
09.04.03 Прикладная информатика	ОФО	М01.Б	4	4	33	8	24	0	1	0	111	Э

4 Структура и содержание дисциплины (модуля)

4.1 Структура дисциплины (модуля) для ОФО

Тематический план, отражающий содержание дисциплины (перечень разделов и тем), структурированное по видам учебных занятий с указанием их объемов в соответствии с учебным планом, приведен в таблице 3.1

Таблица 3.1 – Разделы дисциплины (модуля), виды учебной деятельности и формы текущего контроля для ОФО

№	Название темы	Код ре-зультата обучения	Кол-во часов, отведенное на				Форма текущего контроля
			Лек	Практ	Лаб	СРС	
1	Генерация данных	РД1	2	4	0	15	Практическая работа
2	Инфраструктура больших данных	РД2, РД4, РД5	2	4	0	30	Практическая работа
3	MapReduce	РД3, РД4	2	4	0	30	Практическая работа
4	Технологии Big Data	РД5	2	12	0	36	Практическая работа
Итого по таблице			8	24	0	111	

4.2 Содержание разделов и тем дисциплины (модуля) для ОФО

Тема 1 Генерация данных.

Содержание темы: Источники больших данных: поисковые машины, социальные сети, банковские транзакции, телеком, биоинформатика, Интернет вещей. Характеристики больших данных. Применение больших данных в жизни человека.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: Подготовка к промежуточной аттестации.

Тема 2 Инфраструктура больших данных.

Содержание темы: Распределенные архитектуры больших данных. Параллельные и распределенные вычисления. Grid. Распределенные файловые системы. Файловая система Google. Основные компоненты распределенной файловой системы. Восстановление данных в распределенной файловой системе.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: Подготовка к промежуточной аттестации.

Тема 3 MapReduce.

Содержание темы: Модель распределенных вычислений в вычислительных кластерах.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: Подготовка к промежуточной аттестации.

Тема 4 Технологии Big Data.

Содержание темы: Технологии распределенных вычислений и хранения данных: Hadoop – технология распределенных вычисления на основе модели Map Reduce. HDFS – технология распределенной файловой системы Hadoop. Решение на базе Hadoop: Cloudera. Технология параллельной обработки данных Apache nifi. Диспетчеризация сообщений Apache Kafka. Системы хранения AsterixDB, HP Vertica, Impala, Neo4j, Redis, SparkSQL.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: Подготовка к промежуточной аттестации.

5 Методические указания для обучающихся по изучению и реализации дисциплины (модуля)

5.1 Методические рекомендации обучающимся по изучению дисциплины и по обеспечению самостоятельной работы

Методические рекомендации по организации самостоятельной работы

В ходе изучения дисциплины студенты должны посещать аудиторские занятия (лекции, практические занятия, консультации). Особое место в овладении частью тем данной дисциплины отводится самостоятельной работе, при этом во время аудиторных занятий могут быть рассмотрены и проработаны наиболее важные и трудные вопросы по той или иной теме дисциплины, а применение уже освоенных навыков в смежных технологиях вынесены на самостоятельное обучение.

В соответствии с учебным планом направления подготовки процесс изучения дисциплины предусматривает проведение лекций, практических занятий, консультаций, а также самостоятельную работу студентов.

Ниже перечислены предназначенные для самостоятельного изучения студентами те вопросы, которые во время проведения аудиторных занятий изучаются недостаточно или изучение которых носит обзорный характер.

Перечень и тематика самостоятельных работ студентов по дисциплине

1. Использование базы MongoDB
2. Инструменты интеграции Splunk и Datameer

5.2 Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости обучающимся из числа лиц с ограниченными возможностями здоровья и инвалидов (по заявлению обучающегося) предоставляется учебная информация в доступных формах с учетом их индивидуальных психофизических особенностей:

- для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; индивидуальные задания, консультации и др.

6 Фонд оценочных средств для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине (модулю)

В соответствии с требованиями ФГОС ВО для аттестации обучающихся на соответствие их персональных достижений планируемым результатам обучения по дисциплине (модулю) созданы фонды оценочных средств. Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 1.

7 Учебно-методическое и информационное обеспечение дисциплины (модуля)

0.1 Основная литература

0.2 Дополнительная литература

0.3 Ресурсы информационно-телекоммуникационной сети "Интернет", включая профессиональные базы данных и информационно-справочные системы (при необходимости):

Отсутствуют

8 Материально-техническое обеспечение дисциплины (модуля) и перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

МИНОБРНАУКИ РОССИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Фонд оценочных средств
для проведения текущего контроля
и промежуточной аттестации по дисциплине (модулю)

ТЕХНОЛОГИИ BIG DATA

Направление и направленность (профиль)

09.04.03 Прикладная информатика. Искусственный интеллект и машинное обучение в управлении и принятии решений

Год набора на ОПОП
2022

Форма обучения
очная

Владивосток 2023

1 Перечень формируемых компетенций

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции
09.04.03 «Прикладная информатика» (М-ПИ)	ОПК-2 : Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ОПК-2.1к : Решает профессиональные задачи, используя современные интеллектуальные технологии
		ОПК-2.2к : Обосновывает выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач
	ОПК-5 : Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем	ОПК-5.1к : Разрабатывает и исследует программно-аппаратные комплексы автоматизированных информационных систем

Компетенция считается сформированной на данном этапе в случае, если полученные результаты обучения по дисциплине оценены положительно (диапазон критериев оценивания результатов обучения «зачтено», «удовлетворительно», «хорошо», «отлично»). В случае отсутствия положительной оценки компетенция на данном этапе считается несформированной.

2 Показатели оценивания планируемых результатов обучения

Компетенция ОПК-2 «Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач»

Таблица 2.1 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код результата	Тип результата	Результат	
ОПК-2.1к : Решает профессиональные задачи, используя современные интеллектуальные технологии	РД1	Знание	Базовые технологии больших данных	Сформировавшееся знание о базовых технологиях больших данных
ОПК-2.2к : Обосновывает выбор современных интеллектуальных технологий и программной среды при разработке оригинальных программных средств для решения профессиональных задач	РД2	Умение	Выбор подходящих технологий	Сформировавшееся умение выбирать подходящую технологию

Компетенция ОПК-5 «Способен разрабатывать и модернизировать программное и

аппаратное обеспечение информационных и автоматизированных систем»

Таблица 2.2 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код результата	Тип результата	Результат	
ОПК-5.1к : Разрабатывает и исследует программно-аппаратные комплексы автоматизированных информационных систем	РД3	Знание	Архитектуры инфраструктур хранения и обработки больших данных	Сформированное знание об архитектурах инфраструктур хранения и обработки больших данных
	РД4	Умение	Построение инфраструктуры больших данных	Сформированное умение строить инфраструктуру больших данных
	РД5	Навык	Базовая настройка компонентов инфраструктуры больших данных	Сформированное владение навыками базовой настройки компонентов инфраструктуры больших данных

Таблица заполняется в соответствии с разделом 1 Рабочей программы дисциплины (модуля).

3 Перечень оценочных средств

Таблица 3 – Перечень оценочных средств по дисциплине (модулю)

Контролируемые планируемые результаты обучения	Контролируемые темы дисциплины	Наименование оценочного средства и представление его в ФОС		
		Текущий контроль	Промежуточная аттестация	
Очная форма обучения				
РД1	Знание : Базовые технологии больших данных	1.1. Генерация данных	Практическая работа	Список вопросов
РД2	Умение : Выбор подходящих технологий	1.2. Инфраструктура больших данных	Практическая работа	Список вопросов
РД3	Знание : Архитектуры и инфраструктуры хранения и обработки больших данных	1.3. MapReduce	Практическая работа	Список вопросов
РД4	Умение : Построение инфраструктуры больших данных	1.2. Инфраструктура больших данных	Практическая работа	Список вопросов
		1.3. MapReduce	Практическая работа	Список вопросов
РД5	Навык : Базовая настройка компонентов инфраструктуры больших данных	1.2. Инфраструктура больших данных	Практическая работа	Список вопросов
		1.4. Технологии Big Data	Практическая работа	Список вопросов

4 Описание процедуры оценивания

Качество сформированности компетенций на данном этапе оценивается по результатам текущих и промежуточных аттестаций при помощи количественной оценки, выраженной в баллах. Максимальная сумма баллов по дисциплине (модулю) равна 100 баллам.

Вид учебной деятельности	Оценочное средство		
	Отчёт по практическим работам	Вопросы к экзамену	Итого
Лекция		10	10
Промежуточная аттестация		20	20
Практические занятия	60		60
Самостоятельная работа		10	10
Итого	60	40	100

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей.

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика качества сформированности компетенции
от 91 до 100	«зачтено» / «отлично»	Студент демонстрирует сформированность дисциплинарных компетенций, обнаруживает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями и умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«зачтено» / «хорошо»	Студент демонстрирует сформированность дисциплинарных компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«зачтено» / «удовлетворительно»	Студент демонстрирует сформированность дисциплинарных компетенций: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по некоторым дисциплинарным компетенциям, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено» / «неудовлетворительно»	У студента не сформированы дисциплинарные компетенции, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено» / «неудовлетворительно»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

5 Примерные оценочные средства

5.1 Примеры заданий для выполнения практических работ

Тема 1. Установка сервера

Провести предварительную настройку операционной системы. Установить Java версии 6x или старше. Создать отдельной учетной записи для запуска Hadoop. Настроить доступ по ssh для управления узлами кластера hadoop. Установить Apache Hadoop.

Тема 2. Настройка Apache Hadoop

Тема 3. Интеграция и обработка на платформе Hadoop

Краткие методические указания

После выполнения каждой практической работы студент должен представить отчет о ее выполнении, а также ответить на сопутствующие вопросы по теме.

Шкала оценки

№	Баллы	Описание
5	45-60	Студент демонстрирует умения на итоговом уровне: умеет свободно выполнять задания, предусмотренные программой, свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.
4	28-44	Студент демонстрирует умения на среднем уровне: освоил основные умения, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.
3	19-27	Студент демонстрирует умения и навыки на базовом уровне: допускаются значительные ошибки, проявляется отсутствие отдельных умений, навыков по дисциплинарной компетенции, испытываются значительные затруднения при оперировании умениями и при их переносе на новые ситуации.
2	10-18	Студент демонстрирует умения и навыки на уровне ниже базового: проявляется недостаточность умений и навыков.
1	0-9	Студентом проявляется полное или практически полное отсутствие умений и навыков, но присутствует на занятии и пытается выполнить задание.

5.2 Экзаменационные вопросы

1. Источники больших данных, особенности больших данных
2. Инфраструктура больших данных. Обязательные элементы
3. Технология распределенных вычисления MapReduce
4. Инфраструктура больших данных Hadoop
5. Технологии хранения больших данных
6. Технологии извлечения и обработки больших данных
7. Распределенное хранение

Краткие методические указания

Экзамен проводится после выполнения всех практических работ. Преподаватель на итоговой аттестации задает студенту (письменно или устно) вопрос.

Шкала оценки

№	Баллы	Описание
5	35-40	Студент ответил на заданные вопросы. Студент демонстрирует знания на высоком уровне. Демонстрирует способность
4	29-34	Студент демонстрирует знания на среднем уровне. Демонстрирует способность
3	22-28	Студент демонстрирует знания на базовом уровне.
2	10-21	Студент демонстрирует знания на уровне ниже базового.
1	0-9	Студентом проявляется полное или практически полное отсутствие знаний, но присутствовал на занятии и пытался ответить на вопросы.