Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение высшего образования
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ
филиал ФГБОУ ВО «РГГМУ» в г. Туапсе

Кафедра «Экономики и управления на предприятии природопользования»

Рабочая программа дисциплины

СОЗДАНИЕ BIG DATA

Основная профессиональная образовательная программа
высшего образования программы бакалавриата по направлению подготовки

09.03.03 «Прикладная информатика»

Направленность (профиль):
Прикладные информационные системы и технологии

Квалификация:
Бакалавр

Форма обучения
заочная

Год поступления 2021

Согласовано
Руководитель ОПОП
«Прикладная информатика»

_____ **Аракелов М.С.**

Утверждаю
Директор филиала ФГБОУ
ВО «РГГМУ» в г. Туапсе _____ **Аракелов М.С.**

Рассмотрена и утверждена на заседании кафедры
15 июня 2021 г., протокол № 4

Руководитель
кафедры _____ **Продолятченко П.А.**

Авторы-разработчики:

Яготинцева _____ **Яготинцева Н.В.**

_____

Туапсе 2021

Рассмотрена и рекомендована к использованию в учебном процессе на 2021/2022 учебный год без изменений*

Протокол заседания кафедры №4 от 15 июня 2021 г

Рассмотрено и рекомендовано к использованию в учебном процессе на _____/_____ учебный год с изменениями (см. лист изменений)**

Протокол заседания кафедры _____ от __.__.20__ №__

*Заполняется при ежегодном пересмотре программы, если в неё не внесены изменения

** Заполняется при ежегодном пересмотре программы, если в неё внесены изменения

1. Цель и задачи освоения дисциплины

Цель освоения дисциплины – изучение концепций Big Data, получение знаний о технологиях и методах анализа Big Data и интерпретации результатов.

Задачи:

- формирование практического навыка по работе с кластером хранения и обработки Big Data на примере Cloudera Hadoop, выбору методов и технических решений в зависимости от типа решаемой задачи обработки данных и их объема. Умение использовать инструментарий анализа данных и их визуализации.
- приобретение навыков управления проектом развития системы информационного обеспечения принятия управленческих решений от сбора данных, их хранения и обработки до представления результатов

2. Место дисциплины в структуре основной профессиональной образовательной программы

Дисциплина «Создание Big data» относится к дисциплинам по выбору, формируемых участниками образовательных отношений изучается на 3 курсе.

Для освоения данной дисциплины, обучающиеся должны иметь базовые знания по дисциплине «Информационные системы и технологии».

Параллельно с дисциплиной идёт изучение дисциплин «Обработка, анализ и хранение данных», «Методы тестирования программного обеспечения», «Информационные технологии 3D-моделирования».

Дисциплина «Создание Big data» является базовой для написания выпускной квалификационной работы.

3. Перечень планируемых результатов обучения

Процесс изучения дисциплины направлен на формирование компетенций ПК-11

Таблица 1

Профессиональные компетенции

Код и наименование профессиональной компетенции	**Код и наименование индикатора достижения профессиональной компетенции**	**Результаты обучения**
ПК-11 Способен проектировать программное обеспечение	ИДПК-11.1 Использует существующие типовые решения и шаблоны проектирования программного обеспечения	**Знать:** существующие типовые решения и шаблоны проектирования программного обеспечения **Уметь:** использовать существующие типовые решения и шаблоны проектирования программного обеспечения **Владеть:** методами и средствами проектирования программного обеспечения
	ИДПК-11.2 Применяет методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов	**Знать:** методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов **Уметь:** применять методы и средства проектирования

		программного обеспечения, структур данных, баз данных, программных интерфейсов **Владеть: методами и средствами** проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов
	ИДПК-11.3 Использует принципы и виды построения архитектуры программного обеспечения	**Знать:** принципы и виды построения архитектуры программного обеспечения **Уметь:** применять принципы и виды построения архитектуры программного обеспечения программного обеспечения **Владеть:** методами и средствами построения архитектуры программного обеспечения

4. Структура и содержание дисциплины

4.1. Объем дисциплины

Объем дисциплины составляет 3 зачетные единицы, 108 академических часа.

Таблица 2

Объем дисциплины по видам учебных занятий в академических часах

Объём дисциплины	заочная форма обучения
Объем дисциплины	**108**
Контактная работа обучающихся с преподавателем (по видам аудиторных учебных занятий) – всего:	**14**
в том числе:	**-**
лекции	**6**
лабораторные занятия	**8**
Самостоятельная работа (далее – СРС) – всего:	**94**
Вид промежуточной аттестации	**зачет**

4.2. Структура дисциплины

Таблица 3

Структура дисциплины для заочной формы обучения

№	Тема дисциплины	Курс	Виды учебной работы, в т.ч. самостоятельная работа студентов, час.	Формы текущего контроля успеваемости	Формируемые компетенции	Индикаторы достижения компетенций

			Лекции	Лабораторные работы	СРС			
1	Источники данных	3	1	1	18	Опрос Сдача лабораторных работ	ПК-11	ИДПК-11.1 ИДПК-11.2 ИДПК-11.3
2	Основные понятия Big Data	3	1	1	19	Опрос Сдача лабораторных работ	ПК-11	ИДПК-11.1 ИДПК-11.2 ИДПК-11.3
3	Технологии работы с Big Data	3	1	2	19	Опрос Сдача лабораторных работ	ПК-11	ИДПК-11.1 ИДПК-11.2 ИДПК-11.3
4	Аналитика данных	3	1	2	19	Опрос Сдача лабораторных работ	ПК-11	ИДПК-11.1 ИДПК-11.2 ИДПК-11.3
5	Управление проектами Big Data	3	2	2	19	Опрос Сдача лабораторных работ	ПК-11	ИДПК-11.1 ИДПК-11.2 ИДПК-11.3
	ИТОГО	**-**	**6**	**8**	**94**	**-**	**-**	**-**

4.3. Содержание разделов дисциплины

Введение

Общий обзор дисциплины, специфика, области применения, терминология.

Тема 1. Источники данных

Обзор открытых источников данных. Организация сбора данных о среде с помощью концепции Интернета вещей. Взаимодействие с облачными хранилищами данных.

Тема 2. Основные понятия Big Data

Краткая история развития концепции. Виды и характеристики Big Data

Тема 3. Технологии работы с Big Data

Шардинг и репликация. Архитектуры хранения. Стек технологий Hadoop, распределенная файловая система HDFS, модель вычислений MapReduce

Тема 4. Аналитика данных

Основные понятия. Виды анализа данных. Методология исследования данных CRISP-DM

Тема 5. Управление проектами Big Data

Анализ сервисов облачных вычислений. Программные комплексы машинного обучения Apache Spark и Vowpal Wabbit. Моделирование и оценка результатов

4.4. Содержание занятий семинарского типа

Таблица 4.

Содержание лабораторных занятий для заочной формы обучения

№ темы дисциплины	Тематика лабораторных занятий	Всего часов
1	Источники данных	1
2	Основные понятия Big Data	1
3	Технологии работы с Big Data	2
4	Аналитика данных	2
6	Управление проектами Big Data	2

5. Перечень учебно-методического обеспечения самостоятельной работы обучающихся по дисциплине

Методические материалы по дисциплине представлены в Методических рекомендациях для обучающихся по освоению дисциплины «Создание Big data».
Попов Н.Н., Александрова Л.В., Абрамов В.М. Инновационные технологии геоинформационного обеспечения управления данными предприятия. Режим доступа: http://elib.rshu.ru/files_books/pdf/rid_04837d21305f4a808ed637c5fda17db0.pdf
Методические рекомендации по организации самостоятельной работы обучающихся.

6. Оценочные средства для текущего контроля успеваемости и промежуточной аттестации по итогам освоения дисциплины

Учет успеваемости обучающегося по дисциплине осуществляется по 100-балльной шкале. Максимальное количество баллов по дисциплине за один семестр – 100:

- максимальное количество баллов за выполнение всех видов текущего контроля -60;
- максимальное количество баллов за посещение лекционных занятий - 10;
- максимальное количество баллов за прохождение промежуточной аттестации - 30;

6.1. Текущий контроль

Текущий контроль проводиться в форме выполнения лабораторных работ.

Типовые задания, методика выполнения и критерии оценивания текущего контроля по разделам дисциплины представлены в Фонде оценочных средств по данной дисциплине.

6.2. Промежуточная аттестация

Форма промежуточной аттестации по дисциплине – **зачет.**

Форма проведения зачета: устно по билетам

Перечень вопросов для подготовки к зачету:

ПК-11

Перечень вопросов для подготовки к зачету

1. Что такое большие данные
2. Характериситка VVV
3. Применение Hadoop
4. Пример открытых источников гидрометеорологических данных
5. Сбора данных о среде
6. Концепция «Интернет вещей»
7. Взаимодействие с облачными хранилищами данных
8. Работа с thingspeak.com
9. Виды и характеристики Big Data
10. Технологии работы с Big Data
11. Шардинг и репликация

12. Архитектуры хранения
13. Стек технологий Hadoop
14. Распределенная файловая система HDFS
15. Модель вычислений MapReduce
16. Аналитика данных
17. Виды анализа данных
18. Методология исследования данных CRISP-DM
19. Инструменты анализа обычных данных (RapidMiner, Weka, Knime)
20. Инструменты анализа Big Data (Hive, Pig)
21. Batch обработка и обработка в реальном времени
22. Анализ сервисов облачных вычислений
23. Программные комплексы машинного обучения Apache Spark и Vowpal Wabbit
24. Моделирование и оценка результатов
25. Создание слоев ГИС и их отображение на веб-портале

6.3. Балльно-рейтинговая система оценивания

Таблица 5

Распределение баллов по видам учебной работы

Вид учебной работы, за которую ставятся баллы	**Баллы**
Посещение лекционных занятий	0-10
Выполнение лабораторных работ	0-60
Промежуточная аттестация	0-30
ИТОГО	**0-100**

Минимальное количество баллов для допуска до промежуточной аттестации составляет 40 баллов при условии выполнения всех видов текущего контроля.

Таблица 6

Балльная шкала итоговой оценки на зачете

Оценка	**Баллы**
Зачтено	40-100
Незачтено	0-39

7. Методические указания для обучающихся по освоению дисциплины

Методические рекомендации ко всем видам аудиторных занятий, а также методические рекомендации по организации самостоятельной работы, в том числе по подготовке к текущему контролю и промежуточной аттестации представлены в Методических рекомендации для обучающихся по освоению дисциплины «Создание Big data».

8. Учебно-методическое и информационное обеспечение дисциплины обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Аппаратно-программные средства геоинформационного обеспечения поддержки решений в рамках рационального природопользования / Н.Н. Попов, Л.В. Александрова, В.М. Абрамов, – СПб.: СпецЛит, 2016. - 51 с. (elib.rshu.ru/files_books/pdf/rid_f982b417571f4e62a275b6c34e00be1c.pdf)
2. Инновационные технологии геоинформационного обеспечения управления данными предприятия / Н.Н. Попов, Л.В. Александрова, В.М. Абрамов, – СПб.: СпецЛит, 2017. - 51 с. (elib.rshu.ru/files_books/pdf/rid_04837d21305f4a808ed637c5fda17db0.pdf)

Дополнительная литература

1. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных [Электронный ресурс]. – Электрон. дан. – М.: ДМК Пресс, 2015. – 400 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=69955. – Загл. с экрана.
2. Коэльо Л.П. Построение систем машинного обучения на языке Python [Электронный ресурс] / Л.П. Коэльо, В. Ричарт. – Электрон. дан. – М.: ДМК Пресс, 2016. – 302 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=82818. – Загл. с экрана.

Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

1. http://www.citforum.ru/database/case/index.shtml. (CASE - технологии. Современные методы и средства проектирования информационных систем).
2. http://books.listsoft.ru/book.asp?cod=123239&rp=1 (List SOFT. Каталог программ).

8.2. Перечень программного обеспечения

1. Операционная система Windows XP, Microsoft Office 2007
2. Программы электронных таблиц Excel
3. Текстовый редактор Word
4. Программа для создания презентаций Power Point
5. Программа распознавания текста FineReader

8.3. Перечень информационных справочных систем

1. Консультант Плюс.

8.4. Электронные библиотечные ресурсы:

1. Электронно-библиотечная система ГидроМетеоОнлайн- http://elib.rshu.ru/
2. Информация электронной библиотечной системы http://znanium.com/
3. Электронный каталог библиотеки РГГМУ http://lib.rshu.ru/jirbis2/index.php?option=com_irbis&view=irbis&Itemid=108
4. Издательство ЮРАИТ https://biblio-online.ru/

8.5. Перечень профессиональных баз данных

1. Научная электронная библиотека eLIBRARY.RU https://elibrary.ru/defaultx.asp
2. Федеральная государственная информационная система Национальная электронная библиотека (НЭБ). https://rusneb.ru/
3. Мультидисциплинарная реферативная и наукометрическая база данных Scopus компании Elsevier https://www.scopus.com/search/form.uri?display=basic#basic
4. Политематическая реферативно-библиографическая и наукометрическая (библиометрическая) база данных Web of Science компании Clarivate Analytics http://apps.webofknowledge.com/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=F4DWwm8nvkgneH3Gu7t&preferencesSaved=

9. Материально-техническое обеспечение дисциплины

Материально-техническое обеспечение программы соответствует действующим санитарно-техническим и противопожарным правилам и нормам и обеспечивает проведение всех видов лекционных, практических занятий и самостоятельной работы бакалавров.

Учебный процесс обеспечен аудиториями, комплектом лицензионного программного обеспечения, доступом к электронно-библиотечным системам.

Учебная аудитория для проведения занятий практического типа - укомплектована специализированной мебелью (ученические столы, стулья, компьютерные столы), компьютерная техника с подключением к информационно-коммуникационной сети

«Интернет» и доступом в электронную информационно-образовательную среду образовательной организации, коммуникационное оборудование, обеспечивающее доступ к сети интернет (проводное соединение и беспроводное соединение по технологии Wi-Fi), доской меловой, мультимедиа проектором, аудиоколонками, учебно-наглядными пособиями, программным обеспечением.

Учебная аудитория для групповых и индивидуальных консультаций– укомплектована специализированной мебелью (ученические столы, стулья), доской меловой, компьютером с доступом в сеть Интернет, мультимедиа проектором, аудиоколонками, учебно-наглядными пособиями.

Учебная аудитория для текущего контроля и промежуточной аттестации– укомплектована специализированной мебелью (ученические столы, стулья), доской меловой, компьютером с доступом в сеть Интернет, мультимедиа проектором, аудиоколонками, учебно-наглядными пособиями.

Помещение для самостоятельной работы укомплектовано специализированной мебелью (ученические столы, стулья, компьютерные столы), компьютерная техника с подключением к информационно-коммуникационной сети «Интернет» и доступом в электронную информационно-образовательную среду образовательной организации, коммуникационное оборудование, обеспечивающее доступ к сети интернет (проводное соединение и беспроводное соединение по технологии Wi-Fi), доской меловой, мультимедиа проектором, аудиоколонками, учебно-наглядными пособиями, программным обеспечением.

10. Особенности освоения дисциплины для инвалидов и лиц с ограниченными возможностями здоровья

Обучение обучающихся с ограниченными возможностями здоровья при необходимости осуществляется на основе адаптированной рабочей программы с использованием специальных методов обучения и дидактических материалов, составленных с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся (обучающегося).

При определении формы проведения занятий с обучающимся-инвалидом учитываются рекомендации, содержащиеся в индивидуальной программе реабилитации инвалида, относительно рекомендованных условий и видов труда.

При необходимости для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья создаются специальные рабочие места с учетом нарушенных функций и ограничений жизнедеятельности.

11. Возможность применения электронного обучения и дистанционных образовательных технологий

Дисциплина может реализовываться с применением электронного обучения и дистанционных образовательных технологий