


МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ
Заведующий кафедрой
МО ЭВМ
 **Абрамов Г. В.**
27.03.2025г

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.В.10 Облачные технологии хранения

1. Код и наименование направления подготовки/специальности:

38.04.05 Бизнес информатика

2. Профиль подготовки/специализация:

Информационная бизнес аналитика

3. Квалификация (степень) выпускника: магистр

4. Форма обучения: очная

5. Кафедра, отвечающая за реализацию дисциплины: МО ЭВМ

6. Составители программы: к.ф.-м.н., доц. Трофименко Елена Владимировна

7. Рекомендована: НМС факультета ПММ, протокол №5 от 17.03.2025.

8. Учебный год: 2025/2026

Семестр(ы): 2

9. Цели и задачи учебной дисциплины:

Цели:

- формирование способности к созданию информационных ресурсов глобальных сетей;
- знакомство с основами облачных вычислений, архитектурой и компонентами, с распределенной файловой системой Hadoop и технологией MapReduce.
- знакомство с фреймворком Spark его возможностями и структурой.

Задачи:

- формирование и развитие навыков определения и создания информационных ресурсов глобальных сетей, образовательного контента, средств тестирования систем.
- освоение принципов поиска, хранения, обработки и анализа информации;
- знакомство с форматами представления различных видов данных в облачных технологиях.

10. Место учебной дисциплины в структуре ООП:

Блок 1. Вариативная часть.

Требования к входным знаниям, умениям и компетенциям:

- базовые знания по языкам программирования python;
- знания в области баз данных, язык запросов SQL.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:

Код	Название компетенции	Код(ы)	Индикатор(ы)	Планируемые результаты обучения
ПК-3	Способен проводить обработку и анализ больших данных на базе современных языков программирования и пакетов прикладных программ моделирования	ПК-3.2	Разрабатывает и совершенствует методы анализа массовых количественных и нечисловых данных на базе современных языков программирования и технологий управления данными	Знать – принципы поиска, хранения, обработки и анализа информации; – форматы представления различных видов информации. Уметь – исследовать возможные пути решения поставленной задачи; – осуществлять поиск, хранение, обработку и анализ информации из различных источников и баз данных; – разрабатывать наборы тестовых заданий для проверки корректности разработанного программного кода. Владеть – навыками представления информации в требуемом формате с использованием информационных, компьютерных и сетевых технологий.
ПК-6	Способен управлять электронным предприятием и подразделениями электронного бизнеса	ПК-6.1	Планирует процессы управления жизненным циклом ИТ-инфраструктуры электронного предприятия и организует их	Знать – содержание ИТ-инфраструктуры предприятия и жизненный цикл ее составляющих, особенности и тенденции развития; – основные и вспомогательные процессы управления жизненным циклом ИТ-инфраструктуры предприятия и имеющиеся проблемы.

			исполнение	Уметь – применять методы исследования для выработки организационно-управленческих и проектных решений, с применением облачных технологий. Владеть – современными методами управления жизненным циклом ИТ-инфраструктуры предприятия;
--	--	--	------------	---

12. Объем дисциплины в зачетных единицах/час.(в соответствии с учебным планом) — 2/72.

Форма промежуточной аттестации(зачет/экзамен) экзамен

13. Трудоемкость по видам учебной работы

Вид учебной работы		Трудоемкость			
		Всего	По семестрам		
			2 семестр	№	...
Аудиторные занятия		32	32		
в том числе:	лекции	16	16		
	практические				
	лабораторные	16	16		
Самостоятельная работа		40	40		
в том числе: курсовая работа (проект)					
Форма промежуточной аттестации экзамен					
Итого:		72	72		

13.1. Содержание дисциплины:

№ п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК *
		1. Лекции	
1.1	Введение в основы облачных вычислений	Возникновение и понятие облачных вычислений. Архитектура и компоненты облачных вычислений. Типы облаков.	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517
1.2	Основы распределенной системы данных Hadoop и технологии MapReduce	Структура кластера Hadoop и его элементы. Понятие пара ключ-значение. Знакомство с MapReduce.	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517
1.3	Изучаем фреймворк Spark	Унифицированный стек Spark. Установка и настройка. Программирование операций с RDD. Работа с парами ключ- значение. Загрузка и сохранение данных. Работа с DataFrame	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517
		2. Лабораторные занятия	
2.1	Введение в основы облачных вычислений Основы распределенной системы данных Hadoop и технологии MapReduce	Возникновение и понятие облачных вычислений. Архитектура и компоненты облачных вычислений. Типы облаков. Унифицированный стек PySpark. Установка и настройка.	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517
2.1	Изучаем фреймворк Spark	Структура кластера Hadoop и его элементы. Понятие пара ключ-значение. Программирование операций с RDD. Работа с парами ключ- значение. Загрузка и сохранение данных, работа с файлами.	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517
2.3	Изучаем работу с частично и полностью структурированными данными	Создание RDD, DataFrame применение action и transformation. Загрузка и сохранение данных	Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517

13.2. Темы (разделы) дисциплины и виды занятий:

№ п/п	Наименование раздела дисциплины	Виды занятий (часов)				
		Лекции	Практические	Лабораторные	Самостоятельная работа	Всего
1	Введение в основы облачных вычислений	2			4	6
2	Основы распределенной системы данных Hadoop и технологии MapReduce	4			18	22
3	Изучаем фреймворк Spark. Работа с RDD. Создание и работа с о структурированным и данными	10		16	18	44
	Итого	16		16	40	72

14. Методические указания для обучающихся по освоению дисциплины

Работа с конспектами лекций, выполнение лабораторных заданий, заданий текущей и промежуточной аттестаций.

При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей по работе на LMS-платформе Moodle, своевременно подключаться к online-занятиям, соблюдать рекомендации по организации самостоятельной работы.

15. Учебно-методическое и информационное обеспечение дисциплины

(список оформляется в соответствии с требованиями ГОС и ФГОС, используется общая сквозная нумерация для всех видов литературы)

а) основная литература:

№ п/п	Источник
1	<i>Jules S. Damji, Larning Spark. Lightning - Fast Data Analytic / Jules S. Damji Brooke Wenig, Tathagata Das & Denny Lee. – Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472</i>
2	<i>Захария М. Изучаем Spark. Молниеносный анализ данных./ М. Захария, П. Венделл, Э. Конвински, Х. Карау. — М.: ДМК Пресс, 2015. — 304 с. — ISBN: 978-5-97060-323-9</i>

б) дополнительная литература:

№ п/п	Источник
3	<i>Лэм Ч. Hadoop в действии/Чак Лэм.— М.: ДМК Пресс, 2012.— 424 с.: — ISBN 978-5-97060-156-3, 978-5-94074-785-</i>

в) информационные электронно-образовательные ресурсы:

№ п/п	Источник
4	<i>HYPERLINK "https://docs.microsoft.com/ru-ru/azure/hdinsight/spark/apache-spark-overview" "https://docs.microsoft.com/ru-ru/azure/hdinsight/spark/apache-spark-overview Apache Spark в Azure</i>
5	<i>HYPERLINK "https://spark.apache.org/downloads.html" "https://spark.apache.org/downloads.html официальный сайт</i>
6	<i>Курс Б1.В.10 Облачные технологии хранения https://edu.vsu.ru/course/view.php?id=31517</i>

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	https://edu.vsu.ru - Образовательный портал «Электронный университет ВГУ»- Электронный ресурс <i>Основы облачных вычислений</i>
2	Трофименко Е.В. Введение в облачные вычисления. Основные понятия (Учебное пособие). Воронежский государственный университет. – Воронеж: Издательский дом ВГУ, 2020. – 58с.

17. Информационные технологии, используемые для реализации учебной дисциплины, включая программное обеспечение и информационно-справочные системы (при необходимости)

Для реализации учебного процесса используется бесплатная полнофункциональная интегрированная среда разработки IntelliJ IDEA 2015 и фреймворк Spark. Дисциплина реализуется с применением электронного обучения и дистанционных образовательных технологий. Для организации занятий рекомендован онлайн-курс «Б1.В.10 Облачные технологии хранения <https://edu.vsu.ru/course/view.php?id=31517>», размещенный на платформе Электронного университета ВГУ (LMS moodle), а также Интернет-ресурсы, приведенные в п.15в.

18. Материально-техническое обеспечение дисциплины:

Учебная аудитория для проведения лекций специализированная мебель, компьютер (ноутбук), мультимедийное оборудование (проектор, экран, средства звуковоспроизведения).

Учебная аудитория для практических занятий: специализированная мебель, персональные компьютеры в количестве, обеспечивающем возможность индивидуальной работы, компьютер преподавателя, мультимедийное оборудование (проектор, экран).

ОС Windows 8 (10), интернет-браузер (Google Chrome, Mozilla Firefox), ПО Adobe Reader, пакет стандартных офисных приложений для работы с документами, таблицами (MS Office, МойОфис, LibreOffice)

Учебная аудитория для проведения лабораторных работ: специализированная мебель, персональные компьютеры для индивидуальной работы с возможностью подключения к сети «Интернет», мультимедийное оборудование (проектор, экран, средства звуковоспроизведения).

ОС Windows 8 (10), интернет-браузер (Google Chrome, Mozilla Firefox), ПО Adobe Reader, пакет стандартных офисных приложений для работы с документами, таблицами (MS Office, МойОфис, LibreOffice), допускается демоверсия или виртуальный аналог ПО

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
1.	Основы распределенной системы данных Hadoop и технологии MapReduce	ПК-3	ПК-3.2	Задания для лабораторных работ
2.	Основы распределенной системы данных Hadoop и	ПК-6	ПК-6.1	Задания для лабораторных работ

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
	<i>технологии MapReduce</i>			
3.	<i>Изучаем фреймворк Spark</i>	ПК-3	ОПК-3.2	Задания для лабораторных работ
Промежуточная аттестация форма контроля - зачет				КИМ

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1 Текущий контроль успеваемости

Тестовые задания

Для оценивания результатов обучения используются следующие показатели (ЗУНы из 19.1):

1. Знание структуры экосистемы Hadoop. Знать правила работы и функционирования Data center.
2. Умение связывать теорию с практикой;
3. Умение использовать основные приемы и методы программирования для построения алгоритмов задач по организации облачных вычислений.
4. Умение производить вычисления на языке Spark, Python или Java
5. Умение организовывать коллекции из потока данных и производить с ними различные вычисления и выборки.

Перечень практических заданий

1. Создать текстовый файл (50-100 строк минимум) в качестве имени файла использовать свою фамилию.

Файл содержит текст на английском языке набор слов, из них обязательно есть повторяющиеся и пустые строки.

Над файлом произвести следующие действия:

1. найти заданное слово
2. посчитать количество вхождений заданного слова;
3. разбить текст на слова и удалить пустые строки.
4. Создать наборы RDD на основе массивов (целочисленных и ассоциативных) и применить к ним Transformation reduce и map.

2. Файл содержит данные по успеваемости студентов. Нужно ответить на следующие вопросы:

1. Сколько человек сдало тест?
 - 1.1 Сколько человек в возрасте до 20 лет сдают тест?
 - 1.2 Сколько человек, которым исполнилось 20 лет, сдают экзамен?
 - 1.3 Сколько человек старше 20 лет сдают экзамен?
2. Всего несколько мальчиков сдают экзамен?
 - 2.1 Сколько девушек сдают экзамен?
3. Сколько человек сдают экзамен в 12 классе?

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенции в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций. Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в формах: устного опроса; защиты лабораторных работ, выполнения контрольных работ. Промежуточная аттестация проводится в соответствии с Положением о промежуточной аттестации обучающихся по программам высшего образования. Промежуточная аттестация по итогам освоения дисциплины проводится в форме зачета с оценкой и экзамена. Для получения положительной итоговой оценки необходимо выполнение всех лабораторных и контрольных работ. При оценивании используется следующая шкала:

5 баллов ставится, если обучающийся демонстрирует полное соответствие знаний, умений, навыков приведенным в таблицах показателям, свободно оперирует приобретенными знаниями, умениями, применяет их при решении практических задач;
4 балла ставится, если обучающийся демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателям, но допускает незначительные ошибки, неточности, испытывает затруднения при решении практических задач;
3 балла ставится, если обучающийся демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателям, допускает значительные ошибки при решении практических задач;
2 балла ставится, если обучающийся демонстрирует явное несоответствие знаний, умений, навыков приведенным в таблицах показателям.

20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств: тесты с выбором ответа (ПК -3.2, ПК-6.1)

Примеры теста:

DataFrame это?

- ☐ распределенная коллекция изменяемых данных
- ☐ неизменяемая коллекция данных
- ☐ распределенный набор данных, имеющий структуру, определяется схемой

Правильный ответ: распределенный набор данных, имеющий структуру, определяется схемой

Преобразование where производит с DataFrame следующее...

- ☐ Фильтрацию строк на основе заданных логических условий, возвращая старый DataFrame
- ☐ Поиск строк на основе заданных логических условий, возвращая новый DataFrame
- ☐ Фильтрацию строк на основе заданных логических условий, возвращая новый DataFrame

Правильный ответ: Поиск строк на основе заданных логических условий, возвращая новый DataFrame

Что служит результатом применения действий к RDD?

- ☐ массив данных
- ☐ конкретное значение
- ☐ строковое значение
- ☐ новый набор RDD

Правильный ответ: новый набор RDD

За что отвечает следующий фрагмент кода:

```
new_rdd = my_inner_RDD.map(lambda x:x+2)
```

- ☐ получение нового набора RDD путем увеличения каждого элемента изначального RDD на 2
- ☐ расширение начального RDD в два раза путем добавления случайных чисел
- ☐ добавление двух элементов из начального набора RDD
- ☐ получение нового набора RDD, путем добавления к имеющимся элементам 2-х элементов, полученных случайным образом

Правильный ответ: получение нового набора RDD путем увеличения каждого элемента изначального RDD на 2

Что происходит с RDD в момент вызова нового действия?

- ☐ данные группируются, и вычисления ведутся над каждой группой отдельно
- ☐ вычисление всего набора с самого начала
- ☐ вычисление всего набора, начиная от самого последнего элемента
- ☐ создается новый набор RDD

Правильный ответ: создается новый набор RDD

Операция count, применяемая к DataFrame это?

- ☐ - это действие возвращает количество строк в DataFrame.
- ☐ - это действие возвращает количество переменных в Array.
- ☐ - это трансформация возвращает количество строк в DataFrame.

Правильный ответ: это действие возвращает количество строк в DataFrame

Описание технологии проведения

Текущая аттестация проводится на занятии одновременно во всей учебной группе в виде теста в электронной образовательной среде «Электронный университет ВГУ», адрес курса — <https://edu.vsu.ru/course/> Тест составляется из материалов ФОСа, формируется системой автоматически путём добавления случайных вопросов, количество которых соответствует образцу билета. Большая часть вопросов проверяется автоматически, проверки преподавателем с ручным оцениванием требуют только отдельные вопросы, представленные в форме эссе. Ограничение по времени на каждую попытку — 20 минут, количество попыток — 1, выставление окончательной оценки — по высшему баллу.»

Требования к выполнению заданий (или шкалы и критерии оценивания)

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Задание выполнено: 21-20 правильных ответов	<i>Базовый уровень</i>	<i>зачтено</i>
Задание выполнено: 10-19	-	Не зачтено

Задания раздела 20.2 рекомендуются к использованию при проведении диагностических работ с целью оценки остаточных знаний по результатам освоения данной дисциплины