

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Луковникова Елена Ивановна
Должность: Проректор по учебной работе
Дата подписания: 16.11.2021 12:45:34
Уникальный программный ключ:
890f5aae3463de1924cbcf76ac5d7ab89e9fe7d3

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ

БРАТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ"



УТВЕРЖДАЮ

Проректор по учебной работе

Е.И. Луковникова

Е.И.Луковникова

31 мая

20*21* г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.05 Введение в анализ больших данных

Закреплена за кафедрой **Информатики, математики и физики**

Учебный план b090302_21_ИСиТ.plx

Направление: 09.03.02 Информационные системы и технологии

Квалификация **Бакалавр**

Форма обучения **очная**

Общая трудоемкость **4 ЗЕТ**

Виды контроля в семестрах:

Контрольная работа 7, Зачет 7

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	17			
Неделя	уп	рп	уп	рп
Вид занятий				
Лекции	34	34	34	34
Лабораторные	51	51	51	51
В том числе инт.	16	16	16	16
Итого ауд.	85	85	85	85
Контактная работа	85	85	85	85
Сам. работа	59	59	59	59
Итого	144	144	144	144

Программу составил(и):
б.с., ст.пр., Васильева Лариса Васильевна 

Рабочая программа дисциплины

Введение в анализ больших данных

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.02 Информационные системы и технологии (приказ Минобрнауки России от 19.09.2017 г. № 926)

составлена на основании учебного плана:


Направление: 09.03.02 Информационные системы и технологии
утвержденного приказом ректора от 01.03.2021 протокол № 80.

Рабочая программа одобрена на заседании кафедры

Информатики, математики и физики

Протокол от 16 апреля 2021 г. № 9

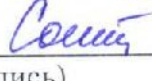
Срок действия программы: 2021-2025 уч.г.

Зав. кафедрой Горохов Д.Б. 

Председатель МКФ

старший преподаватель Латушкина С.В. 18 20 апреля 2021 г. 

Ответственный за реализацию ОПОП  Д.Б. Горохов
(подпись) (ФИО)

Директор библиотеки  Т.Ф. Сотник
(подпись) (ФИО)

№ регистрации 221
(методический отдел)

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	формирование профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных.
-----	--

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП

Цикл (раздел) ООП:	Б1.В.05
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Технологии обработки информации
2.1.2	Базы данных
2.1.3	Математическое моделирование
2.1.4	Программирование
2.1.5	Алгоритмы и структуры данных
2.1.6	Теория вероятностей и математическая статистика
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Производственная (преддипломная) практика
2.2.2	Выполнение и защита выпускной квалификационной работы

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-2: Способность разрабатывать прототип информационной системы на базе типового решения и кодировать на языках программирования

Индикатор 1	ПК-2.2. Разрабатывает код информационной системы (базы данных информационной системы), используя современные языки и технологии программирования
-------------	--

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	особенности анализа больших данных.
3.2	Уметь:
3.2.1	применять технологии анализа больших данных.
3.3	Владеть:
3.3.1	навыками разработки кода информационной системы с помощью технологий анализа больших данных.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Вид занятия	Наименование разделов и тем	Семестр / Курс	Часов	Компетенции	Литература	Инте ракт.	Примечание
	Раздел	Раздел 1. Введение в большие данные						
1.1	Лек	Современные подходы к хранению, обработке и использованию данных. Предпосылки формирования тренда больших данных. Тенденции и перспективы рынка больших данных и их инструментов. Возможности для бизнеса. Большие данные (Big Data): понятие, сущность, ключевые признаки. Жизненный цикл больших данных. Экосистема больших данных.	7	6	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4	2	ПК-2.2 Лекция-визуализация
1.2	Ср	Подготовка к лекциям	7	2	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4	0	ПК-2.2
1.3	Зачёт	Подготовка и сдача зачета	7	2	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4	0	ПК-2.2

	Раздел	Раздел 2. Технологии хранения и анализа больших данных						
2.1	Лек	Архитектура системы обработки больших данных. Масштабирование и многоуровневое хранение данных. Парадигма NoSQL. Классификация NoSQL баз данных. Базы данных NewSQL и полихранилища. Высокопроизводительные вычисления. Параллельные алгоритмы для работы с данными. Вычислительная парадигма MapReduce. Визуализация данных и результатов анализа. Сложные методы аналитики. Программные платформы и системы для больших данных. Проект Apache Hadoop и его экосистема. Инструментарий Apache Hadoop.	7	28	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4	6	ПК-2.2 Лекция-визуализация
2.2	Лаб	MySQL для больших данных	7	17	ПК-2	Л1.1 Л1.2Л2.1 Л2.3 Л2.4 Э1	0	ПК-2.2
2.3	Лаб	Основы Hadoop	7	34	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4 Э2	8	ПК-2.2 Работа в малых группах
2.4	Ср	Подготовка к лекциям, к лабораторным работам	7	24	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4 Э2	0	ПК-2.2
2.5	Контр.раб.	Выполнение контрольной работы, подготовка отчета, подготовка к защите и защита работы	7	29	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4 Э2	0	ПК-2.2
2.6	Зачёт	Подготовка и сдача зачета	7	2	ПК-2	Л1.1 Л1.2Л2.1 Л2.2 Л2.4	0	ПК-2.2

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

Традиционная (репродуктивная) технология (преподаватель знакомит обучающихся с порядком выполнения задания, наблюдает за выполнением и при необходимости корректирует работу обучающихся)

Образовательные технологии с использованием активных методов обучения (лекция – беседа, лекция – дискуссия, проблемная лекция, лекция-визуализация, лекция с заранее запланированными ошибками, лекция – пресс-конференция, лекция с разбором конкретных ситуаций, лекция-консультация, занятия с применением затрудняющих условий, методы группового решения творческих задач, метод развивающейся кооперации)

Технология компьютерного обучения(использование в учебном процессе компьютерных технологий и предоставляемых ими возможностей (электронные библиотеки, онлайн тесты, практические задания и т.д.))

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Контрольные вопросы и задания

I. Задания для лабораторных работ.

ЛР 1. MySQL для больших данных.

Изучить возможности MySQL для обработки больших данных.

ЛР 2. Основы Hadoop.

1. Выполнить установку Hadoop кластера.
2. Освоить базовые операции обслуживания кластера и файловые операции HDFS.
3. Приобрести навыки управление ресурсами и запуском задач, использования основных инструментов и компонент Apache Hadoop экосистемы.

II. Отчеты по лабораторным работам.

6.2. Темы письменных работ

Целью контрольной работы является закрепление знаний об организации больших данных и практических навыков применения современных технологий их обработки и анализа.

Тема контрольной работы определяется вариантом индивидуального задания, который устанавливает предметную область получения данных, задачу обработки данных:

1. Социальные медиа.
2. Политика.
3. Наука и исследование.
4. Энергетика.
5. Здравоохранение.
6. Бизнес.
7. Транспорт.

Конечным результатом контрольной работы является отчет (документ текстового редактора), отражающий основные этапы решения поставленной задачи: получение данных, выбор инструментария для обработки, обработка и анализ полученных результатов.

Рекомендуемый объем отчета – 15-20 страниц.

Выполненная работа в печатном и электронном вариантах сдается на проверку преподавателю. Преподаватель принимает решение о допуске работы к защите. При наличии значительных ошибок и замечаний, работа возвращается обучающемуся на доработку.

Выдача задания, прием выполненных работ и их защита производится в соответствии с календарным учебным графиком.

6.3. Фонд оценочных средств

Вопросы к зачету.

Раздел 1. Введение в большие данные.

1. Роль цифровой информации в 21 веке. Предпосылки формирования тренда больших данных.
2. Тенденции и перспективы рынка больших данных и их инструментов. Возможности для бизнеса.
3. Большие данные (Big Data): понятие, сущность, ключевые признаки.
4. Жизненный цикл больших данных.
5. Экосистема больших данных.

Раздел 2. Технологии хранения и анализа больших данных.

1. Архитектура системы обработки больших данных.
2. Масштабирование и многоуровневое хранение данных.
3. Парадигма NoSQL. Классификация NoSQL баз данных.
4. Базы данных NewSQL и полихранилища.
5. Высокопроизводительные вычисления. Вычислительная парадигма MapReduce.
6. Визуализация данных и результатов анализа.
7. Сложные методы аналитики.
8. Программные платформы и системы для больших данных.
9. Проект Apache Hadoop и его экосистема. Инструментарий Apache Hadoop.

6.4. Перечень видов оценочных средств

Задания для лабораторных работ.

Отчеты по лабораторным работам.

Отчет по контрольной работе.

Вопросы к зачету.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Кол-во	Эл. адрес
ЛП.1		Базы данных в высокопроизводительных информационных системах: учебное пособие	Ставрополь: Северо-Кавказский Федеральный университет (СКФУ), 2016	1	http://biblioclub.ru/index.php?page=book&id=466799

	Авторы,	Заглавие	Издательство,	Кол-во	Эл. адрес
Л1. 2	Лесковец Ю., Раджара ман А., Ульман Дж.	Анализ больших наборов данных: Учебное пособие	Москва: ДМК Пресс, 2016	1	https://ecat.brstu.ru/catalog/Ресурсы%20свободного%20доступа/Лесковец%20Ю.Анализ%20больших%20наборов%20данных.2016.pdf

7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Кол-во	Эл. адрес
Л2. 1	Д. Келлехер, Б. Тирни	Наука о данных: базовый курс: учебное пособие	Москва : Альпина Паблишер, 2020	1	http://biblioclub.ru/index.php?page=book&id=598235
Л2. 2	Силен Д., Мейсман А., Али М.	Основы Data Science и Big Data. Python и наука о данных: учебное пособие	Санкт- Петербург: Пите р, 2017	1	https://ecat.brstu.ru/catalog/Ресурсы%20свободного%20доступа/Силен%20Д.Основы%20Data%20Science%20и%20Big%20Data.%20Python%20и%20наука%20о%20данных.2017.pdf
Л2. 3	Чаллавала Ш. и др.	MySQL 8 для больших данных: учебное издание	Москва: ДМК Пресс, 2018	1	https://ecat.brstu.ru/catalog/Ресурсы%20свободного%20доступа/Чаллавала%20Ш.MySQL%208%20для%20больших%20данных.2018.pdf
Л2. 4	Радченко И.А., Никола ев И.Н.	Технологии и инфраструктура Big Data: учебное пособие	Санкт- Петербург: Унив ерситет ИТМО, 2018	1	https://ecat.brstu.ru/catalog/Ресурсы%20свободного%20доступа/Радченко%20И.А.%20Технологии%20и%20инфраструктура%20Big%20Data.2018.pdf

7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Официальный сайт MySQL	https://www.mysql.com/
Э2	Официальный сайт проекта Apache Hadoop	https://hadoop.apache.org/

7.3.1 Перечень программного обеспечения

7.3.1.1	Microsoft Office Professional Plus 2010 Russian Academic OPEN 1 license No Level
7.3.1.2	Adobe Reader
7.3.1.3	LibreOffice
7.3.1.4	Chrome
7.3.1.5	Anaconda
7.3.1.6	MySQL Community Edition
7.3.1.7	Python IDLE
7.3.1.8	MySQL Server

7.3.2 Перечень информационных справочных систем

7.3.2.1	Национальная электронная библиотека НЭБ
7.3.2.2	Университетская информационная система РОССИЯ (УИС РОССИЯ)
7.3.2.3	
7.3.2.4	Научная электронная библиотека eLIBRARY.RU
7.3.2.5	Электронная библиотека БрГУ
7.3.2.6	Электронный каталог библиотеки БрГУ
7.3.2.7	«Университетская библиотека online»

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

3125	Дисплейный класс	Учебная мебель Комплект серверного оборудования для построения технической архитектуры комплекса терминальных решений в составе терминального сервера, терминальных рабочих мест и периферии в составе: терминальный сервер Dell PowerEdge RX740XD, монитор Samsung SM493 19", 15 тонких клиентов SmartClient Mini PC (Intel CPU J1900 1.99GHzx4, 4GB), монитор Forgame Liquid Crystal Display MK27FC 27" 1800R 1920x1080 144 Hz, вебкамера Logitech C920 PRO), МФУ Canon i-Sensys MF 421dw, доска интерактивная сенсорная Smart Board SB480.
------	------------------	--

3125	Дисплейный класс	Учебная мебель Комплект серверного оборудования для построения технической архитектуры комплекса терминальных решений в составе терминального сервера, терминальных рабочих мест и периферии в составе: терминальный сервер Dell PowerEdge RX740XD, монитор Samsung SM493 19", 15 тонких клиентов SmartClient Mini PC (Intel CPU J1900 1.99GHzx4, 4GB), монитор Forgame Liquid Crystal Display MK27FC 27" 1800R 1920x1080 144 Hz, вебкамера Logitech C920 PRO), МФУ Canon i-Sensys MF 421dw, доска интерактивная сенсорная Smart Board SB480.
2201	читальный зал №1	Учебная мебель Оборудование 10- ПК i5-2500/H67/4Gb (монитор TFT19 Samsung); принтер HP Laser Jet P2055D
3125	Дисплейный класс	Учебная мебель Комплект серверного оборудования для построения технической архитектуры комплекса терминальных решений в составе терминального сервера, терминальных рабочих мест и периферии в составе: терминальный сервер Dell PowerEdge RX740XD, монитор Samsung SM493 19", 15 тонких клиентов SmartClient Mini PC (Intel CPU J1900 1.99GHzx4, 4GB), монитор Forgame Liquid Crystal Display MK27FC 27" 1800R 1920x1080 144 Hz, вебкамера Logitech C920 PRO), МФУ Canon i-Sensys MF 421dw, доска интерактивная сенсорная Smart Board SB480.

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Учебным планом предусмотрены следующие виды учебных занятий: лекции, лабораторные работы, контрольная работа, самостоятельная работа обучающегося, зачет.

Лекции.

Написание конспекта лекций: краткое, последовательное изложение основных положений, формулировок, выводов, обобщений; техническое оформление записей (подчеркивание, выделение ключевых слов и терминов). Активная работа на лекции.

Лабораторные работы.

Выполнение заданий с использованием методических указаний по выполнению лабораторных работ, оформление отчетов, защита лабораторных работ.

Самостоятельная работа обучающихся.

- Подготовка к лекциям: систематическая работа с конспектом лекций (чтение записей; проверка терминов с помощью энциклопедий, словарей и справочников; обозначение вопросов, материал, которых вызывает трудности; попытка найти ответ в рекомендуемых источниках; подготовка вопросов преподавателю, если не удается самостоятельно разобраться в материале).

- Подготовка к лабораторным работам: проработка материалов по теме лабораторной работы с использованием рекомендуемой литературы, конспекта лекций, ресурсов информационно-телекоммуникационной сети Интернет; оформление отчетов по лабораторным работам; подготовка к защите лабораторных работ.

- Выполнение контрольной работы: выполнение задания с использованием рекомендуемых источников; оформление отчета; подготовка к защите контрольной работы.

- Подготовка к зачету: проработка вопросов к зачету с использованием конспекта лекций, рекомендуемых источников; подготовка вопросов преподавателю, если не удается самостоятельно разобраться в материале.