

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ

"БРАТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ"

УТВЕРЖДАЮ

Проректор по учебной работе

_____ Е.И.Луковникова

_____ 05 июня _____ 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.ДВ.02.01 Большие данные

Закреплена за кафедрой **Базовая кафедра менеджмента и информационных технологий**

Учебный план b090303_23_ПИЭ.plx

Направление: 09.03.03 Прикладная информатика

Квалификация **Бакалавр**

Форма обучения **очная**

Общая трудоемкость **3 ЗЕТ**

Виды контроля в семестрах:

Зачет 8

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	8 (4.2)		Итого	
Неделя	10			
Вид занятий	уп	рп	уп	рп
Лекции	30	30	30	30
Лабораторные	40	40	40	40
В том числе инт.	14	14	14	14
В том числе в форме практ.подготовки	40	40	40	40
Итого ауд.	70	70	70	70
Контактная работа	70	70	70	70
Сам. работа	38	38	38	38
Итого	108	108	108	108

Программу составил(и):
к.т.н., дек., Патрусова А.М. _____

Рабочая программа дисциплины

Большие данные

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.03 Прикладная информатика (приказ Минобрнауки России от 19.09.2017 г. № 922)

составлена на основании учебного плана:

Направление: 09.03.03 Прикладная информатика
утвержденного приказом ректора от 17.02.2023 № 72.

Рабочая программа одобрена на заседании кафедры

Базовая кафедра менеджмента и информационных технологий

Протокол от 27 апреля 2023 г. № 14

Срок действия программы: 2023-2027 уч.г.

Зав. кафедрой Вахрушева М.Ю.

Председатель МКФ

доцент, к.э.н., Грудистова Е.Г. 11.05.2023 г. № 9 _____

Ответственный за реализацию ОПОП _____ Вахрушева М.Ю.
(подпись)

Директор библиотеки _____ Сотник Т.Ф.
(подпись)

№ регистрации _____ 52 _____
(методический отдел)

Визирование РПД для исполнения в очередном учебном году

Председатель МКФ

_____ 2024 г.

Рабочая программа пересмотрена, обсуждена и одобрена для
исполнения в 2024-2025 учебном году на заседании кафедры
Базовая кафедра менеджмента и информационных технологий

Внесены изменения/дополнения (Приложение _____)

Протокол от _____ 2024 г. № __

Зав. кафедрой _____

Визирование РПД для исполнения в очередном учебном году

Председатель МКФ

_____ 2025 г.

Рабочая программа пересмотрена, обсуждена и одобрена для
исполнения в 2025-2026 учебном году на заседании кафедры
Базовая кафедра менеджмента и информационных технологий

Внесены изменения/дополнения (Приложение _____)

Протокол от _____ 2025 г. № __

Зав. кафедрой _____

Визирование РПД для исполнения в очередном учебном году

Председатель МКФ

_____ 2026 г.

Рабочая программа пересмотрена, обсуждена и одобрена для
исполнения в 2026-2027 учебном году на заседании кафедры
Базовая кафедра менеджмента и информационных технологий

Внесены изменения/дополнения (Приложение _____)

Протокол от _____ 2026 г. № __

Зав. кафедрой _____

Визирование РПД для исполнения в очередном учебном году

Председатель МКФ

_____ 2027 г.

Рабочая программа пересмотрена, обсуждена и одобрена для
исполнения в 2027-2028 учебном году на заседании кафедры
Базовая кафедра менеджмента и информационных технологий

Внесены изменения/дополнения (Приложение _____)

Протокол от _____ 2027 г. № __

Зав. кафедрой _____

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	Овладение основами теоретических и практических знаний в области больших массивов данных, методов их обработки, анализа и эффективного применения в экономике.
-----	--

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП

Цикл (раздел) ООП:	Б1.В.ДВ.02.01
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Интеллектуальные информационные системы
2.1.2	Основы интеллектуального анализа данных
2.1.3	Статистика и анализ данных
2.1.4	Базы данных
2.1.5	Теория систем и системный анализ
2.1.6	Цифровая экономика
2.1.7	Теория вероятностей и математическая статистика
2.1.8	Финансовая грамотность
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Выполнение и защита выпускной квалификационной работы
2.2.2	Производственная (преддипломная) практика

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)**ПК-4: Способен разрабатывать базы данных ИС**

Индикатор 1	ПК-4.1. Владеет принципами и методами разработки баз данных; основами современных систем управления базами данных
Индикатор 2	ПК-4.2. Способен выбирать инструментальные средства и технологии проектирования базы данных ИС
Индикатор 3	ПК-4.3. Способен разрабатывать структуру баз данных ИС

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	основы современных систем управления базами данных; инструментальные средства и технологии проектирования базы данных ИС; структуру баз данных ИС
3.2	Уметь:
3.2.1	разрабатывать базы данных; выбирать инструментальные средства и технологии проектирования базы данных ИС; разрабатывать структуру баз данных ИС
3.3	Владеть:
3.3.1	принципами и методами разработки баз данных; навыками применения инструментальных средств и технологий проектирования базы данных ИС; навыками создания баз данных ИС

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. Большие данные (Big Data): современные подходы к обработке и хранению						
1.1	Лек	Общее понятие и признаки Big Data	8	2	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.2	Лек	Формы представления данных. Типы и виды данных	8	2	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.3	Лек	Подходы к обработке и хранению Big Data	8	2	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3

1.4	Лаб	Поиск и определение Big Data	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.5	Лаб	Извлечение и визуализация данных	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.6	Лаб	Хранение больших данных	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.7	Ср	Подготовка к лабораторным работам	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
1.8	Зачёт	Подготовка к зачёту	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
	Раздел	Раздел 2. Инструментарий анализа больших данных						
2.1	Лек	Программное обеспечение в области анализа больших данных	8	10	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	8	Технология проектного обучения ПК-4.1, ПК-4.2, ПК-4.3
2.2	Лаб	Программное обеспечение в области анализа больших данных	8	10	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	6	Технология проектного обучения ПК-4.1, ПК-4.2, ПК-4.3
2.3	Ср	Подготовка к лабораторным работам	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
2.4	Зачёт	Подготовка к зачёту	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
	Раздел	Раздел 3. Технологии KDD и Data Mining						
3.1	Лек	Технологии KDD	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.2	Лек	Технологии Data Mining	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.3	Лек	Методология Big Data	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.4	Лаб	Прогнозирование с помощью линейной регрессии	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.5	Лаб	Алгоритмы кластеризации больших данных	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3

3.6	Лаб	Поиск ассоциативных правил	8	4	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.7	Ср	Подготовка к лабораторным работам	8	10	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3
3.8	Зачёт	Подготовка к зачёту	8	6	ПК-4	Л1.1 Л1.2 Л1.3 Л1.4 Л1.5Л2.1 Л2.2	0	ПК-4.1, ПК-4.2, ПК-4.3

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

Технология проектного обучения (приобретение знаний, умений и личного опыта по созданию и реализации проектов)

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Контрольные вопросы и задания

Вопросы для текущего контроля освоения дисциплины:

Раздел 1. Большие данные (Big Data): современные подходы к обработке и хранению.

1. Большие данные (Big Data): сущность, основные понятия, история появления.
2. Большие данные (Big Data): технические сложности работы с большими данными.
3. Большие данные (Big Data): признаки больших данных.
4. Большие данные (Big Data): роль и место больших данных в решении аналитических и исследовательских задач профессиональной деятельности.
5. Общая схема анализа данных.
6. Извлечение и визуализация данных.
7. Этапы моделирования.
8. Процесс построения моделей.
9. Представление данных. Типы и виды данных.
10. Представление наборов данных.

Раздел 2. Инструментарий анализа больших данных.

Интерактивная форма проведения лекционных занятий с применением технологии проектного обучения с целью приобретения знаний, умений и личного опыта по созданию и реализации проектов предполагает просмотр видеозаписи вебинара во время лекционных занятий на тему "Программное обеспечение в области анализа больших данных" для поиска и группового обсуждения ответов на следующие вопросы (задания).

1. Аналитические платформы: классификация и особенности применения.
2. Языки визуального моделирования.
3. Язык программирования Python.
4. Реляционные базы данных.
5. Параллельные базы данных.
6. Распределенные файловые системы.
7. ETL процесс по обработке отчетов.

Интерактивная форма проведения лабораторных работ "Программное обеспечение в области анализа больших данных" с применением технологии проектного обучения с целью приобретения знаний, умений и личного опыта по применению программного обеспечения в области анализа данных предполагает изучение соответствующего программного обеспечения на конкретном примере.

Тема 3. Технологии KDD и Data Mining.

1. Понятие и сущность технологии Data Mining.
2. Подготовка данных к анализу.
3. Методика извлечения знаний.
4. Причины распространения KDD и Data Mining.
5. Актуальность технологий Data Mining как средств обработки больших объемов информации.
6. Ассоциативные правила. Поддержка и достоверность ассоциативного правила.

7. Поиск ассоциативных правил.
8. Кластеризация. Цели кластеризации в Data Mining.
9. Примеры кластеризации в различных областях.
10. Применение классификации и регрессии.
11. Обзор методов классификации и регрессии.
12. Статистические методы.
13. Методы, основанные на обучении, разнообразии подходов.

6.2. Темы письменных работ

Не предусмотрено учебным планом

6.3. Фонд оценочных средств

Вопросы к зачету:

Раздел 1. Большие данные (Big Data): современные подходы к обработке и хранению.

- 1.1. Большие данные (Big Data): сущность, основные понятия, история появления.
- 1.2. Большие данные (Big Data): технические сложности работы с большими данными.
- 1.3. Большие данные (Big Data): признаки больших данных.
- 1.4. Большие данные (Big Data): роль и место больших данных в решении аналитических и исследовательских задач профессиональной деятельности.
- 1.5. Общая схема анализа данных.
- 1.6. Извлечение и визуализация данных.
- 1.7. Этапы моделирования.
- 1.8. Процесс построения моделей.
- 1.9. Представление данных. Типы и виды данных.
- 1.10. Представление наборов данных.

Раздел 2. Инструментарий анализа больших данных.

- 2.1. Аналитические платформы: классификация и особенности применения.
- 2.2. Языки визуального моделирования.
- 2.3. Язык программирования Python.
- 2.4. Реляционные базы данных.
- 2.5. Параллельные базы данных.
- 2.6. Распределенные файловые системы.
- 2.7. ETL процесс по обработке отчетов.

Тема 3. Технологии KDD и Data Mining.

- 3.1. Понятие и сущность технологии Data Mining.
- 3.2. Подготовка данных к анализу.
- 3.3. Методика извлечения знаний.
- 3.4. Причины распространения KDD и Data Mining.
- 3.5. Актуальность технологий Data Mining как средств обработки больших объемов информации.
- 3.6. Ассоциативные правила. Поддержка и достоверность ассоциативного правила.
- 3.7. Поиск ассоциативных правил.
- 3.8. Кластеризация. Цели кластеризации в Data Mining.
- 3.9. Примеры кластеризации в различных областях.
- 3.10. Применение классификации и регрессии.
- 3.11. Обзор методов классификации и регрессии.
- 3.12. Статистические методы.
- 3.13. Методы, основанные на обучении, разнообразии подходов.

6.4. Перечень видов оценочных средств

Вопросы для текущего контроля освоения дисциплины, 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
--	---------	----------	---------------	--------	-----------

	Авторы,	Заглавие	Издательство,	Кол-во	Эл. адрес
Л1.1	Волкова В. М., Семенова М. А., Четвертакова Е. С., Вожов С. С.	Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python: учебное пособие	Новосибирск: Новосибирский государственный технический университет, 2017	1	http://biblioclub.ru/index.php?page=book&id=576496
Л1.2	Каган Е. С.	Прикладной статистический анализ данных: учебное пособие	Кемерово: Кемеровский государственный университет, 2018	1	http://biblioclub.ru/index.php?page=book&id=573550
Л1.3	Жуковский О. И.	Информационные технологии и анализ данных: учебное пособие	Томск: Эль Контент, 2014	1	http://biblioclub.ru/index.php?page=book&id=480500
Л1.4	Силен Д., Мейсман А., Али М.	Основы Data Science и Big Data. Python и наука о данных: учебное пособие	Санкт-Петербург: Питер, 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
Л1.5	Радченко И.А., Николаев И.Н.	Технологии и инфраструктура Big Data: учебное пособие	Санкт-Петербург: Университет ИТМО, 2018	1	https://ecat.brstu.ru/catalog/Ресурсы%20свободного%20доступа/Радченко%20И.А.%20Технологии%20и%20инфраструктура%20Big%20Data.2018.pdf

7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Кол-во	Эл. адрес
Л2.1	Боровиков В.П., Ивченко Г.И.	Прогнозирование в системе STATISTICA в среде Windows: Основы теории и интенсивная практика на компьютере: Учеб. пособие для вузов	Москва: Финансы и статистика, 2006	10	
Л2.2	Бродовская Е. В., Домбровская А. Ю.	Большие данные в исследовании политических процессов: учебное пособие	Москва: Московский педагогический государственный университет (МПГУ), 2018	1	http://biblioclub.ru/index.php?page=book&id=563578

7.3.1 Перечень программного обеспечения

7.3.1.1	Microsoft Windows Professional 7 Russian Upgrade Academic OPEN No Level
7.3.1.2	Microsoft Office 2007 Russian Academic OPEN No Level
7.3.1.3	LibreOffice
7.3.1.4	Chrome
7.3.1.5	Python
7.3.1.6	R

7.3.2 Перечень информационных справочных систем

7.3.2.1	Университетская информационная система РОССИЯ (УИС РОССИЯ)
7.3.2.2	Национальная электронная библиотека НЭБ
7.3.2.3	Научная электронная библиотека eLIBRARY.RU
7.3.2.4	Информационная система "Единое окно доступа к образовательным ресурсам"
7.3.2.5	Электронная библиотека БрГУ
7.3.2.6	Электронный каталог библиотеки БрГУ
7.3.2.7	«Университетская библиотека online»
7.3.2.8	Издательство "Лань" электронно-библиотечная система

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Аудитория	Назначение	Оснащение аудитории	Вид занятия
3217	Учебная аудитория (мультимедийный класс)	Основное оборудование: - интерактивная доска SMART Board 680i2/Unifl, - интерактивный планшет Wacom PL-720,	Лек

		- колонки Microlab Solo-7C, - ноутбук Samsung R610<NP-R610-FS08>, - телевизор плазменный Samsung 63 PS-63A756T1M. Дополнительно: - маркерная доска – 1 шт. Учебная мебель: - комплект мебели (посадочных мест) – 42 шт.; - комплект мебели (посадочных мест) для преподавателя – 1 шт.	
3234	Учебная аудитория (дисплейный класс)	Основное оборудование: - ПК AMD 3.9 GHz 4GbDVD 19 KbMs (13 шт.). Дополнительно: - маркерная доска – 1 шт. Учебная мебель: - комплект мебели (посадочных мест/АРМ) – 24/12 шт.; - комплект мебели (посадочных мест) для преподавателя – 1 шт.; - комплект мебели (посадочных мест/АРМ) для оператора – 1/1 шт.	Ср
3236	Учебная аудитория (дисплейный класс)	Основное оборудование: - системный блок AMD A10-7800 Radeon R7 (12 шт.), - системный блок для слабовидящих пользователей AMD A10-7850K (1 шт.), - монитор Philips233 V5QHABP (13 шт.). Дополнительно: - маркерная доска – 1 шт. Учебная мебель: - комплект мебели (посадочных мест/АРМ) – 26/12 шт.; - комплект мебели (посадочных мест) для преподавателя – 1 шт.; - комплект мебели (посадочных мест/АРМ) для оператора – 1/1 шт.	Зачёт
2201	читальный зал №1	Комплект мебели (посадочных мест) Стеллажи Комплект мебели (посадочных мест) для библиотекаря Выставочные шкафы ПК i5-2500/H67/4Gb (монитор TFT19 Samsung) (10шт.); принтер HP Laser Jet P2055D (1шт.)	Ср

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

Дисциплина «Большие данные» направлена на овладение основами теоретических и практических знаний в области больших массивов данных, методов их обработки, анализа и эффективного применения в экономике.

Изучение дисциплины «Большие данные» предусматривает лекции; самостоятельную работу обучающихся; зачёт.

Помимо освоения основных разделов дисциплины необходимо овладеть навыками и умениями сбора и применения больших данных в конкретных ситуациях.

В процессе изучения дисциплины на первом этапе рекомендуется обратить внимание на понятийно-категориальный аппарат дисциплины. Овладение ключевыми понятиями является важным этапом в освоении содержания современных методов и способов работы с большими данными.

При подготовке к сдаче зачёта рекомендуется особое внимание уделить вопросам, связанным с признаками и принципами работы с большими данными, основными способами анализа и перспективами развития больших данных.

Самостоятельную работу по изучению дисциплины необходимо начинать с проработки конспекта лекций, обобщения, систематизации, углубления и конкретизации полученных теоретических знаний с использованием основной и дополнительной литературы, а также рекомендуемых ресурсов информационно-телекоммуникационной сети «Интернет».

В процессе консультации с преподавателем необходимо уточнять вопросы, термины, материал, вызвавший трудности при самостоятельной работе.

Важнейшим элементом в получении знаний по дисциплине является работа с литературой. Прежде всего, необходимо воспользоваться списком рекомендуемой по данной дисциплине литературы. Дополнительные сведения по изучаемым темам можно найти в периодической печати и информационно-телекоммуникационной сети «Интернет».

Предусмотрено проведение аудиторных занятий в виде лекций в сочетании с внеаудиторной работой.