



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное учреждение
высшего образования
**«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ИГУ»)**

Институт математики и информационных технологий
Кафедра алгебраических и информационных систем



Рабочая программа дисциплины (модуля)

Б1.В.01 Анализ данных научных экспериментов

Направление подготовки информационные технологии	02.04.02	Фундаментальная информатика и	и
Направленность (профиль) подготовки		Анализ данных научных исследований и	
Квалификация выпускника	магистр		
Форма обучения	очная		

1. Цели и задачи дисциплины:

Цель

Целью дисциплины Б1.В.01 **Анализ данных научных экспериментов** является знакомство студентов с задачами обработки данных экспериментов, выработка навыков применения математических методов и алгоритмов для обработки, анализа и интерпретации экспериментальных данных.

Задачи:

Познакомить студентов с актуальными задачами научных экспериментов; на примерах опубликованных исследований продемонстрировать как за счёт большого объема экспериментальных данных может быть получено новое знание; дать представление об основных математических подходах, используемых при анализе.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина (модуль) Б1.В.01 **Анализ данных научных экспериментов** относится к дисциплинам по выбору.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы знания, умения и навыки, формируемые дисциплинами, включенными в программу бакалавриата: высшая математика, математический анализ, дискретная математика, линейная алгебра, теория вероятностей. В программе магистратуры предшествующей дисциплиной является «Прикладная статистика».

2.3. Перечень последующих учебных дисциплин, для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной: все виды практик, подготовка и защита ВКР.

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций (элементов следующих компетенций) в соответствии с ФГОС ВО по соответствующему направлению подготовки.

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
<i>ПК-3. Способен формулировать задачи, анализировать и применять способы и методы научных исследований, проводить информационный поиск и использовать информационные ресурсы для решения научно-исследовательских задач, формулировать и</i>	ИДК ПК3.1 Умеет выделять проблемы, относящиеся к области естественных наук	Знать: основные понятия и методы, относящиеся к области вычислительной биологии, необходимые для задач профессиональной деятельности. Уметь: выделять проблемы, относящиеся к области естественных наук, фундаментальной информатике и информационным технологиям Владеть: навыками применения современного математического аппарата для обработки и анализа
	ИДК ПК3.2 Умеет решать актуальные проблемы, относящиеся к области естественных наук	
	ИДК ПК3.3 Способен формулировать проблемы, относящиеся к области естественных наук	
	ИДК ПК3.4	

представлять научные результаты в форме презентаций и публикаций	Способен проводить анализ математических моделей, выбирать оптимальные	экспериментальных данных (в том числе больших).
	ИДК ПК3.4 Способен создавать инновационные методы решения прикладных задач профессиональной деятельности	

4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 4 зачетных единицы, 144 часа,

Форма промежуточной аттестации: зачет

4.1 Содержание дисциплины, структурированное по темам, с указанием видов учебных занятий и отведенного на них количества академических часов

№ п/п	Раздел дисциплины/темы	С е м е с тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости; Форма промежуточной аттестации (по семестрам)
			Контактная работа преподавателя с обучающимися			Самостоятельная работа	
			Лекции	Лабораторные работы	Консультации		
1	Основные инструменты анализа данных	2	4	4		28	тест
	Библиотеки numpy, pandas		2	2		14	
	Библиотеки scipy, matplotlib		2	2		14	
2	Язык программирования R	2	4	4		28	тест
	Синтаксис и основные отличия языка		2	2		14	
	Статистический анализ и визуализация		2	2		14	
3	Хранение и управление данными	2	4	4		28	тест
	Облачные платформы для больших данных		2	2		14	
	Инструменты распределённой обработки		2	2		14	
4	Практикум по анализу данных реальных экспериментов (на основе публикаций)	2	4	4		28	тест
	Часть 1. Естественные науки		2	2		14	
	Часть 2. Экономика, языки, социальные науки		2	2		14	
Итого часов			16	16		112	

4.2 План внеаудиторной самостоятельной работы обучающихся по дисциплине

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Затраты времени (час.)		
1	Основные инструменты анализа данных	<i>УИЛ</i>	1-я четверть курса	28	тест	
	Библиотеки numpy, pandas	<i>УИЛ</i>		14		
	Библиотеки scipy, matplotlib	<i>УИЛ</i>		14		
	Язык программирования R	<i>УИЛ</i>	2-я четверть курса	28	тест	
	Синтаксис и основные отличия языка	<i>УИЛ</i>		14		
	Статистический анализ и визуализация	<i>УИЛ</i>		14		
	Хранение и управление данными	<i>УИЛ</i>	3-я четверть курса	28	тест	
	Облачные платформы для больших данных	<i>УИЛ</i>		14		
	Инструменты распределённой обработки	<i>УИЛ</i>		14		
	Практикум по анализу данных реальных экспериментов (на основе публикаций)	<i>УИЛ</i>	4-я четверть курса	28	тест	
	Часть 1. Естественные науки	<i>УИЛ</i>		14		
	Часть 2. Экономика, языки, социальные науки	<i>УИЛ</i>		14		
Общая трудоёмкость самостоятельной работы по дисциплине (час)				112		
Бюджет времени самостоятельной работы, предусмотренный учебным планом для данной дисциплины (час)						

Виды самостоятельной работы:

Р – написание реферата,

Д – подготовка доклада,

У – выполнение упражнений,

Э – написание эссе,

Пт – выполнение проекта,

К – кейс-задание,

Пф – портфолио,

И – информационный поиск,

Прз – презентация,

Л – изучение литературы,

4.3. Методические указания по организации самостоятельной работы студентов

Методические указания по организации самостоятельной работы расположены в ЭОС Educa

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

а) основная литература

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2021. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/469022>

2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>

3. Кузьмин, В. И. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Гадзаов. — 2-е изд., перераб. и доп. — Москва : РТУ МИРЭА, 2020. — 155 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171433>

4. Дюк, В. А. Логический анализ данных : учебное пособие / В. А. Дюк. — Санкт-Петербург : Лань, 2020. — 80 с. — ISBN 978-5-8114-4180-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126935>

Дополнительная литература

1. Павлов, Л. А. Структуры и алгоритмы обработки данных : учебник для вузов / Л. А. Павлов, Н. В. Первова. — 3-е изд., стер. — Санкт-Петербург : Лань, 2021. — 256 с. — ISBN 978-5-8114-7259-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/156929>

2. Карангин, В. П. Обработка экспериментальных данных : учебное пособие / В. П. Карангин, С. Ф. Елецкая. — Омск : ОмГТУ, 2018. — 48 с. — ISBN 978-5-8149-2603-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/149107>

3. Стефанова, И. А. Обработка данных и компьютерное моделирование : учебное пособие / И. А. Стефанова. — Санкт-Петербург : Лань, 2020. — 112 с. — ISBN 978-5-8114-4010-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126939>

4. Васильев, А. Н. Числовые расчеты в Excel : справочник / А. Н. Васильев. — Санкт-Петербург : Лань, 2021. — 608 с. — ISBN 978-5-8114-1580-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/168874>

в) периодические издания

1. CODATA Data Science Journal <https://datascience.codata.org/> CODATA Data Science Journal - это рецензируемый электронный журнал с открытым доступом, публикующий статьи по управлению, распространению, использованию и повторному использованию исследовательских данных и баз данных во всех областях исследований, включая науку, технологии, гуманитарные науки и искусство.

2. Научный рецензируемый журнал Моделирование и анализ информационных систем <https://www.mais-journal.ru/jour>

3. Научный рецензируемый журнал “PeerJ Computer Science” <https://peerj.com/computer-science/>

4. Научный рецензируемый журнал «Машинное обучение и анализ данных» <https://www.elibrary.ru/contents.asp?titleid=32828>

г) список авторских методических разработок:

нет

д) базы данных, информационно-справочные и поисковые системы

1. Единое окно доступа к образовательным ресурсам. Полнотекстовая электронная библиотека учебных и учебно-методических материалов (федеральный ресурс). <http://www.window.edu.ru>.
2. Лекторий МФТИ <https://mipt.lectoriy.ru/>
3. Московский центр непрерывного математического образования, МЦНМО. Материалы (полные тексты) свободно распространяемых книг по математике. <http://www.mcsme.ru/free-books>.
4. База знаний и набор вычислительных алгоритмов. <http://www.wolframalpha.com>.
5. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru/>
6. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
7. ЭОС ИГУ <https://educa.isu.ru/>
8. Онлайн-курсы от ведущих вузов и компаний страны <https://welcome.stepik.org/ru>
9. Образовательный онлайн-проект <https://www.coursera.org/>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Учебно-лабораторное оборудование:

Для проведения лекционных занятий необходима аудитория с презентационным оборудованием, для проведения практических занятий необходима аудитория на 15-30 рабочих мест (в зависимости от численности учебной группы), оборудованная доской, презентационной техникой.

6.2. Программное обеспечение:

LaTeX – с использованием сборки TeXLive (или возможность выхода на онлайн-ресурс Overleaf), pdf-view'er, произвольный пакет или библиотеки для математических вычислений.

6.3. Технические и электронные средства:

ЭОС EDUCA, презентационное оборудование, персональный компьютер с возможностью просмотра презентаций.

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации данного курса используются следующие образовательные технологии: технологии традиционного обучения, игровые технологии, технологии проблемного обучения, технологии обучения в сотрудничестве, интерактивные технологии, технологии дистанционного обучения, активные педагогические технологии.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ


Оценочные средства (ОС):

- 8.1. Оценочные средства для входного контроля – тест в ЭОС Educa.
- 8.2. Оценочные средства текущего контроля – тесты в ЭОС Educa в соответствии с п. 4.1.
- 8.3. Оценочные средства для промежуточной аттестации (в форме зачета). Необходимым условием получения зачета является выполнение всех тестовых заданий.

Примерные вопросы для зачета.

- 1) Какие преимущества даёт совместная работа в онлайн-сервисе Google Colab?
- 2) Перечислите основные библиотеки языка программирования Python, используемые для анализа данных
- 3) С помощью каких библиотек языка программирования Python можно проводить статистический анализ экспериментальных данных?
- 4) Назовите и сравните основные подходы машинного обучения для классификации
- 5) В чём преимущество языка R в обработке данных?
- 6) Перечислите популярные «облачные» сервисы для хранения и обработки экспериментальных данных
- 7) Продемонстрируйте возможности библиотек визуализации при построении графиков и диаграмм по реальным данным
- 8) Какие подходы используются для кластеризации данных?
- 9) С помощью каких методов можно сгенерировать данные с необходимыми статистическими характеристиками?
- 10) Проведите линейно-регрессионный анализ предложенного набора данных.

Разработчики:

 Доцент кафедры АиИС ИМИТ ИГУ Петрушин И.С.
(подпись) (занимаемая должность) (инициалы, фамилия)

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 02.04.02 «Фундаментальная информатика и информационные технологии» (уровень магистратуры), утвержденный приказом Министерства образования и науки Российской Федерации от «23» августа 2017 г. № 811, зарегистрированный в Минюсте России «13» сентября 2017 г. № 48168 с изменениями и дополнениями от 26 ноября 2020 г., 8 февраля 2021 г.

Программа рассмотрена на заседании кафедры Алгебраических и информационных систем ИМИТ ИГУ «24» марта 2022 г.

Протокол № 9 Зав. кафедрой  Пантелеев В.И.