

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 29.05.2023 14:33:12  
Уникальный программный ключ:  
ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования  
«Российский университет дружбы народов имени Патриса Лумумбы»  
Факультет физико-математических и естественных наук  
(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)**

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **АНАЛИЗ БОЛЬШИХ ДАННЫХ**

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки/специальности:**

### **02.03.02 ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

(код и наименование направления подготовки/специальности)

**Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):**

### **ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

(наименование (профиль/специализация) ОП ВО)

**2023 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Анализ больших данных» входит в программу бакалавриата «Фундаментальная информатика и информационные технологии» по направлению 02.03.02 «Фундаментальная информатика и информационные технологии» и изучается в 7 семестре 4 курса. Дисциплину реализует Кафедра прикладной информатики и теории вероятностей. Дисциплина состоит из 4 разделов и 17 тем и направлена на изучение решения прикладных задач, связанных с анализом данных в статистическом пакете R.

Целью освоения дисциплины является приобретение знаний и навыков работы с большими данными в различных прикладных пакетах, в особенности в пакете R.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Анализ больших данных» направлено на формирование у обучающихся следующих компетенций (части компетенций):

*Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)*

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Знает принципы сбора, отбора и обобщения информации, методики системного подхода для решения профессиональных задач; УК-1.2 Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности; УК-1.3 Владеет навыками научного поиска и практической работы с информационными источниками; методами принятия решений;
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-12.1 Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных;
ОПК-1	Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности	ОПК-1.1 Знает основные положения и концепции в области математических и естественных наук; знает основную терминологию; ОПК-1.2 Умеет осуществлять первичный сбор и анализ материала, интерпретировать различные математические объекты; ОПК-1.3 Имеет практический опыт работы с решением стандартных математических задач и применяет его в профессиональной деятельности;
ПК-1	Способен разрабатывать и	ПК-1.1 Знает основы программирования; современные

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	отлаживать программный код	объектно-ориентированные языки программирования; современные структурные языки программирования; языки современных бизнес-приложений; ПК-1.2 Умеет кодировать на языках программирования; тестировать результаты кодирования; ПК-1.3 Владеет навыками разработки кода информационной системы; навыками верификации кода информационной системы;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Анализ больших данных» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Анализ больших данных».

*Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины*

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-12	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	Обработка данных и визуализация; Основы машинного обучения и нейронные сети; Обработка больших данных с использованием машинного обучения; Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);	Технологическая (проектно-технологическая) практика; Научно-исследовательская работа; Преддипломная практика;
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	История России; Философия; Математическое моделирование; Вычислительные системы, сети и телекоммуникации; Алгоритмы машинной графики и обработки изображений; Стохастический анализ беспроводных сетей; Интеллектуальные системы; Программная инженерия;	Технологическая (проектно-технологическая) практика; Научно-исследовательская работа; Преддипломная практика;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
		Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);	
ОПК-1	Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности	<p>Математический анализ;  Аналитическая геометрия;  Дискретная математика и математическая логика;  <i>Компьютерный практикум по моделированию</i> **;  Физика;  Теория вероятностей и математическая статистика;  Теория конечных графов;  Вычислительные методы;  Математическое моделирование;  Имитационное моделирование;  Технологии искусственного интеллекта;  Линейный и нелинейный регрессионный анализ больших данных;  Обработка больших данных с использованием машинного обучения;  Анализ приоритетного доступа в мультисервисных сетях;  Методы искусственного интеллекта;  Методы машинного обучения;  Машинное обучение в телекоммуникациях;  Стохастический анализ беспроводных сетей;  Алгебра;  Дифференциальные уравнения;  Прикладные задачи теории массового обслуживания;  <i>Компьютерный практикум по информационным технологиям</i> **;  Компьютерная алгебра;  Основы машинного обучения и нейронные сети;  Компьютерная геометрия;  Марковские процессы;  Методы машинного обучения для анализа временных рядов и панельных данных;  Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);</p>	Научно-исследовательская работа; Преддипломная практика;
ПК-1	Способен разрабатывать и отлаживать программный код	<p>Основы программирования;  Технология программирования;  <i>Компьютерный практикум по моделированию</i> **;  Технологии искусственного</p>	Научно-исследовательская работа; Технологическая (проектно-технологическая) практика; Преддипломная практика;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
		<p>интеллекта;  Методы машинного обучения для анализа временных рядов и панельных данных;  Методы искусственного интеллекта;  Основы формальных методов описания бизнес-процессов;  Управление проектами разработки информационных систем;  Введение в управление инфокоммуникациями;  Разработка информационно-аналитических систем;  Методы машинного обучения;  Алгоритмы машинной графики и обработки изображений;  Основы машинного обучения и нейронные сети;  Теория автоматов и формальных языков;  Стохастический анализ беспроводных сетей;  Интеллектуальные системы;  Программная инженерия;  Архитектура компьютеров и операционные системы;  <i>Компьютерный практикум по информационным технологиям**</i>;  Компьютерная алгебра;  Обработка данных и визуализация;  Компьютерная геометрия;  Машинное обучение в телекоммуникациях;  Обработка больших данных с использованием машинного обучения;</p>	

\* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

\*\* - элективные дисциплины /практики

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Анализ больших данных» составляет «3» зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			7
<i>Контактная работа, ак.ч.</i>	54		54
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	36		36
<i>Самостоятельная работа обучающихся, ак.ч.</i>	54		54
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	0		0
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>108</b>	<b>108</b>
	<b>зач.ед.</b>	<b>3</b>	<b>3</b>

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Большие данные: основные понятия и задачи	1.1	Определение больших данных. Области, в которых используется анализ больших данных	ЛК
		1.2	Методы и инструменты анализа больших данных	ЛК
Раздел 2	Введение в статистический пакет R	2.1	Типы переменных. Операции над различными переменными. Простейшая математика	ЛК, СЗ
		2.2	Операторы цикла и условия. Создание функций	ЛК, СЗ
		2.3	Классы данных	ЛК, СЗ
		2.4	Ввод и вывод данных	ЛК, СЗ
		2.5	Графические функции	ЛК, СЗ
Раздел 3	Анализ данных в пакете R	3.1	Базовые вероятностные распределения	ЛК, СЗ
		3.2	Анализ категориальных данных	ЛК, СЗ
		3.3	Статистические числовые характеристики	ЛК, СЗ
		3.4	Графический анализ числовых данных	ЛК, СЗ
		3.5	Оценки неизвестных параметров	ЛК, СЗ
		3.6	Проверка статистических гипотез на случай одной выборки	ЛК, СЗ
		3.7	Проверка статистических гипотез на случай двух и более выборок	ЛК, СЗ
Раздел 4	Пакеты и функции в R, используемые в анализе больших данных	4.1	Пакет bigmemory – создание, хранение, доступ и обработка сверхбольших матриц	ЛК, СЗ
		4.2	Пакет BGDData – анализ данных, связанных с геномными исследованиями	ЛК, СЗ
		4.3	Пакеты bigSurvSGD и bigstatsr	ЛК, СЗ

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – семинарские занятия.

## 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет R
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет R

	презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	Компьютер/ноутбук с доступом к сети Интернет и электронно-образовательной среде Университета, браузер, программное обеспечение для просмотра pdf-файлов, MS Teams, пакет R. Дополнительное программное обеспечение: MS Office или LibreOffice

\* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Роберт Кабаков. R в действии = R in Action. — ДМК-Пресс, 2014. — 588 с. — ISBN 978-5-947060-077-1

2. Хэдли Уикем, Гарретт Гроулмунд. Язык R в задачах науки о данных: импорт, подготовка, обработка, визуализация и моделирование данных = R for Data Science: Visualize, Model, Transform, Tidy, and Import Data. — Вильямс, 2017. — 592 с. — ISBN 978-5-9909446-8-8, 978-1-491-91039-9.

3. Норман Мэтлофф[en]. Искусство программирования на R. Погружение в большие данные. = The Art of R Programming: A Tour of Statistical Software Design.. — Питер, 2019. — 416 с. — ISBN 978-5-4461-1101-5.

*Дополнительная литература:*

1. Шипунов А. Б. R — объектно-ориентированная статистическая среда.  
<http://ashipunov.info/shipunov/software/r/r-ru.htm>

2. R for Data Science. <https://r4ds.had.co.nz/index.html>

3. The Comprehensive R Archive Network. <https://cran.r-project.org/index.html>

*Ресурсы информационно-телекоммуникационной сети «Интернет»:*

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров

- Электронно-библиотечная система РУДН – ЭБС РУДН

<http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>

- ЭБС Юрайт <http://www.biblio-online.ru>

- ЭБС «Консультант студента» [www.studentlibrary.ru](http://www.studentlibrary.ru)

- ЭБС «Троицкий мост»

2. Базы данных и поисковые системы

- электронный фонд правовой и нормативно-технической документации

<http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS

<http://www.elsevier.com/locate/elsevier/scopus/>

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля\*:*

1. Курс лекций по дисциплине «Анализ больших данных».



\* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

## **8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ**

Оценочные материалы и балльно-рейтинговая система\* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Анализ больших данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

\* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

**РАЗРАБОТЧИК:**

Доцент кафедры прикладной  
информатики и теории  
вероятностей

*Должность, БУП*



*Подпись*

Зарядов Иван Сергеевич

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ БУП:**

Заведующий кафедрой  
прикладной информатики и  
теории вероятностей

*Должность, БУП*



*Подпись*

Самуйлов Константин

Евгеньевич

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ ОП ВО:**

Заведующий кафедрой  
прикладной информатики и  
теории вероятностей

*Должность, БУП*



*Подпись*

Самуйлов Константин

Евгеньевич

*Фамилия И.О.*