

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 22.05.2024 14:35:09

Уникальный программный ключ:

sa953a01204891083f939673078ef1a989dae18a

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:

02.03.01 МАТЕМАТИКА И КОМПЬЮТЕРНЫЕ НАУКИ

(код и наименование направления подготовки/специальности)

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

МАТЕМАТИКА И КОМПЬЮТЕРНЫЕ НАУКИ

(наименование (профиль/специализация) ОП ВО)

2024 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Анализ больших данных» входит в программу бакалавриата «Математика и компьютерные науки» по направлению 02.03.01 «Математика и компьютерные науки» и изучается в 7 семестре 4 курса. Дисциплину реализует Кафедра теории вероятностей и кибербезопасности. 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 консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и	ОПК-1.1 Обладает базовыми знаниями, полученными в области математических и (или) естественных наук; ОПК-1.2 Умеет использовать базовые знания, полученные в области математических и (или) естественных наук, в профессиональной деятельности; ОПК-1.3 Имеет навыки выбора методов решения задач профессиональной деятельности на основе теоретических знаний;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности	
ОПК-4	Способен находить, анализировать, реализовывать программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	ОПК-4.1 Знает базовые основы современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности; ОПК-4.2 Умеет использовать математический аппарат в профессиональной деятельности; ОПК-4.3 Имеет практический опыт применения современного математического аппарата, связанного с проектированием, разработкой, реализацией и оценкой качества программных продуктов и программных комплексов в различных областях человеческой деятельности;
ПК-1	Способен разрабатывать и отлаживать программный код	ПК-1.1 Знает основы программирования; современные 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	Обработка данных и визуализация; Основы машинного обучения и нейронные сети; Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);	Технологическая (проектно-технологическая) практика; Преддипломная практика; Научно-исследовательская работа;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных		
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	Философия; Математическое моделирование; Вычислительные системы, сети и телекоммуникации; Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);	Преддипломная практика; Научно-исследовательская работа;
ОПК-1	Способен консультировать и использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений, дискретной математики и математической логики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики в профессиональной деятельности	Математический анализ; Алгебра и аналитическая геометрия; Дискретная математика и математическая логика; Физика; Теория вероятностей и математическая статистика; Теория конечных графов; Дифференциальные уравнения; Вычислительные методы; Математическое моделирование; Имитационное моделирование; Марковские процессы; Теоретическая механика; Дифференциальная геометрия и топология; Функциональный анализ; Компьютерная геометрия; Компьютерная алгебра; Основы машинного обучения и нейронные сети; Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);	
ОПК-4	Способен находить, анализировать, реализовывать программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	Основы программирования; Технология программирования; Вычислительные методы; Математическое моделирование; Имитационное моделирование; Алгоритмы машинной графики и обработки изображений; Компьютерная алгебра; Основы машинного обучения и нейронные сети; Компьютерная геометрия;	
ПК-1	Способен разрабатывать и отлаживать программный код	Алгоритмы машинной графики и обработки изображений; <i>Компьютерный практикум по моделированию</i> **;	Преддипломная практика; Научно-исследовательская работа; Технологическая (проектно-

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
		<p><i>Компьютерный практикум по информационным технологиям</i> **;</p> <p>Обработка данных и визуализация;</p> <p>Компьютерная алгебра;</p> <p>Основы машинного обучения и нейронные сети;</p> <p>Компьютерная геометрия;</p> <p>Методы машинного обучения;</p> <p>Машинное обучение в телекоммуникациях;</p> <p>Архитектура компьютеров и операционные системы;</p> <p>Основы программирования;</p> <p>Технология программирования;</p> <p>Введение в программирование для мобильных платформ;</p>	технологическая) практика;

* - 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
Общая трудоемкость дисциплины	ак.ч.	108	108
	зач.ед.	3	3

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 BGData – 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

- 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.elsevierscience.ru/products/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 теории
вероятностей и
кибербезопасности

Должность, БУП

Подпись

Зарядов И.С.

Фамилия И.О.

РУКОВОДИТЕЛЬ БУП:

Заведующий кафедрой теории
вероятностей и
кибербезопасности

Должность БУП

Подпись

Самуйлов К.Е.

Фамилия И.О.

РУКОВОДИТЕЛЬ ОП ВО:

Заведующий кафедрой теории
вероятностей и
кибербезопасности

Должность, БУП

Подпись

Самуйлов К.Е.

Фамилия И.О.