

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**

### **ЛИНЕЙНЫЙ И НЕЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ 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 «Линейный и нелинейный регрессионный 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 5 семестре 3 курса. Дисциплину реализует Кафедра прикладной информатики и теории вероятностей. Дисциплина состоит из 6 разделов и 15 тем и направлена на изучение теоретических основ анализа больших данных, включая базовые элементы статистического программирования и интеллектуального анализа больших наборов данных

Целью освоения дисциплины является получение навыков проведения расчетов с применением технологий анализа больших данных и решения широкого спектра прикладных задач обработки больших наборов данных.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Линейный и нелинейный регрессионный 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)
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;
ОПК-6	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	ОПК-6.1 Знает базовые принципы цифровых технологий и методов, необходимых в профессиональной деятельности в области фундаментальной информатики и информационных технологий для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.; ОПК-6.2 Умеет применять необходимые в профессиональной деятельности цифровые технологии и методы в области фундаментальной информатики и информационных технологий для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.; ОПК-6.3 Владеет необходимыми в профессиональной деятельности технологиями и методами в области фундаментальной информатики и информационных технологий для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.;
ПК-4	Способен проводить работы по обработке и анализу научно-технической информации и результатов исследований	ПК-4.1 Знает основы научно-исследовательской деятельности в области информационных технологий, основные методы решения прикладных задач, современные методы информационных технологий; принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		ПК-4.2 Умеет применять полученные знания для решения стандартных задач в области информационных технологий и в собственной научно-исследовательской деятельности; ПК-4.3 Владеет базовыми навыками подготовки научных обзоров и (или) публикаций, рефератов и библиографий по тематике проводимых исследований на русском и иностранном языке;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Линейный и нелинейный регрессионный анализ больших данных» относится к блоку по выбору блока образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Линейный и нелинейный регрессионный 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*
ОПК-1	Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности	Математический анализ; Аналитическая геометрия; Дискретная математика и математическая логика; Физика; Теория вероятностей и математическая статистика; Теория конечных графов; Алгебра; Дифференциальные уравнения; Компьютерная алгебра; Основы машинного обучения и нейронные сети; Компьютерная геометрия;	<i>Компьютерный практикум по моделированию**</i> ; Математическое моделирование; Имитационное моделирование; Сети массового обслуживания и их применения; Обработка больших данных с использованием машинного обучения; Анализ приоритетного доступа в мультисервисных сетях; Методы машинного обучения; Интеллектуальные обучающие системы; Анализ больших данных при моделировании сложно-структурированных систем; Моделирование сетей передачи данных; <i>Прикладной анализ данных с использованием языка Python**</i> ; Методы оптимизации и исследование операций; <i>Компьютерный практикум по интеллектуальным системам**</i> ; <i>Компьютерный практикум по статистическому</i>

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
			<p><i>анализу данных**;</i>  Стохастический анализ беспроводных сетей;  <i>Компьютерный практикум по информационным технологиям**;</i>  Анализ больших данных;  Технологии интеллектуального анализа данных и прогнозирование;  Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);  Научно-исследовательская работа;  Преддипломная практика;</p>
ОПК-6	<p>Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности</p>	<p>Интеллектуальные системы;  Обработка данных и визуализация;  Основы машинного обучения и нейронные сети;  Компьютерная геометрия;</p>	<p><i>Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);</i>  <i>Технологическая (проектно-технологическая) практика;</i>  <i>Научно-исследовательская работа;</i>  <i>Преддипломная практика;</i>  <i>Компьютерный практикум по моделированию**;</i>  <i>Математическое моделирование;</i>  <i>Имитационное моделирование;</i>  <i>Обработка больших данных с использованием машинного обучения;</i>  <i>Анализ приоритетного доступа в мультисервисных сетях;</i>  <i>Администрирование локальных сетей;</i>  <i>Введение в управление инфокоммуникациями;</i>  <i>Разработка информационно-аналитических систем;</i>  <i>Технологии интеллектуального анализа данных и прогнозирование;</i>  <i>Методы машинного обучения;</i>  <i>Интеллектуальные обучающие системы;</i>  <i>Анализ больших данных при моделировании сложно-структурированных систем;</i>  <i>Моделирование сетей передачи данных;</i></p>

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
			<p><i>Параллельное программирование**;</i>  <i>Прикладной анализ данных с использованием языка Python**;</i>  <i>Компьютерный практикум по интеллектуальным системам**;</i>  <i>Компьютерный практикум по статистическому анализу данных**;</i>  <i>Стохастический анализ беспроводных сетей;</i>  <i>Компьютерный практикум по информационным технологиям**;</i>  <i>Сети массового обслуживания и их применения;</i></p>
ПК-4	Способен проводить работы по обработке и анализу научно-технической информации и результатов исследований		<p><i>Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы);</i>  <i>Научно-исследовательская работа;</i>  <i>Преддипломная практика;</i>  <i>Сети массового обслуживания и их применения;</i>  <i>Анализ приоритетного доступа в мультисервисных сетях;</i>  <i>Обработка больших данных с использованием машинного обучения;</i>  <i>Анализ больших данных при моделировании сложно-структурированных систем;</i>  <i>Технологии интеллектуального анализа данных и прогнозирование;</i>  <i>Компьютерный практикум по моделированию**;</i>  <i>Интеллектуальные обучающие системы;</i>  <i>Моделирование сетей передачи данных;</i>  <i>Компьютерный практикум по интеллектуальным системам**;</i>  <i>Компьютерный практикум по статистическому анализу данных**;</i>  <i>Компьютерный практикум по информационным технологиям**;</i>  <i>Иностранный язык</i></p>

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/модули, практики*</b>	<b>Последующие дисциплины/модули, практики*</b>
			<i>(дополнительные разделы)**; Русский язык как иностранный (дополнительные разделы)**;</i>

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

\*\* - элективные дисциплины /практики

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

Общая трудоемкость дисциплины «Линейный и нелинейный регрессионный 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)
			5
<i>Контактная работа, ак.ч.</i>	36		36
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	18		18
<i>Самостоятельная работа обучающихся, ак.ч.</i>	72		72
<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	Введение в большие данные.	1.1	Жизненный цикл аналитики данных	ЛК, СЗ
		1.2	Высокопроизводительные вычисления	СЗ
		1.3	Масштабирования и многоуровневое хранение данных	СЗ
Раздел 2	Основной аппарат комбинаторики и математической статистики	2.1	Основные понятия математической статистики: статистические оценки, их свойств	ЛК
		2.2	Проверка гипотез. Наблюдённый уровень значимости	ЛК, СЗ
Раздел 3	Детектирование выбросов и аномалий. Что такое выбросы, типы выбросов. Методы обнаружения выбросов .Поиск аномалий. Цензурирование выборок	3.1	Детектирование выбросов и аномалий. Что такое выбросы, типы выбросов. Методы обнаружения выбросов .Поиск аномалий. Цензурирование выборок	ЛК, СЗ
		3.2	Отсев объектов-выбросов, удаление выбросов. Очистка данных и технологии регуляризации. Основные виды регуляризации. Метод редукции размерности. Методы отбора признаков	ЛК
Раздел 4	Нейронные сети. Генетические алгоритмы.	4.1	Нейронные сети: общая архитектура. Многослойные сети	СЗ
		4.2	Обратное распространение ошибки. Стохастический градиентный спуск. Генетические алгоритмы	ЛК, СЗ
Раздел 5	Выделение особенностей (Feature detection); нормализация данных. Не-четкие множества. Байесовы сети.	5.1	Извлечение признаков / выделение особенностей (Feature detection). Преобразования признаков. Нормализация данных. Методы нормализации данных. Нормализация по методу минимакс	СЗ
		5.2	Нечёткие множества. Байесовы сети. Задачи байесовского вывода. Методика построения нечеткой байесовой сети	ЛК, СЗ
Раздел 6	Сложные методы аналитики	6.1	Множественная и нелинейная регрессии	ЛК
		6.2	Классификация. Логит и пробит модели. Деревья. Случайный лес	СЗ
		6.3	Кластеризация	ЛК, СЗ
		6.4	Визуализация данных и результатов анализа	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
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	Компьютер/ноутбук с доступом сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет Gretl, Пакет PSPP.
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 сети Интернет и электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams. Пакет Gretl, Пакет PSPP.

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

1. Модели с фиктивными переменными и бинарным откликом в пакете Gretl : учебно- методическое пособие для студентов факультета физико-математических и естественных наук / Д.А. Пяткина, С.И. Матюшенко. - Электронные текстовые данные. - Москва : РУДН, 2020. - 40 с. : ил. - ISBN 978-5-209-10433-9
2. Математическое моделирование в экономике и финансах : учебно- методическое пособие для студентов, обучающихся по специальности "Бизнес-информатика" / Д.А. Пяткина, С.И. Матюшенко. - Электронные текстовые данные. - М. : Изд-во РУДН, 2018. - 40 с. : ил. - ISBN 978-5-209-08322-1 : 71.04.  
[http://lib.rudn.ru/MegaPro/UserEntry?Action=Rudn\\_FindDoc&id=468107&idb=0](http://lib.rudn.ru/MegaPro/UserEntry?Action=Rudn_FindDoc&id=468107&idb=0)
3. Эконометрика : учебно-методические указания к лабораторным работам для студентов, обучающихся по специальности "Бизнес-информатика" / Д.А. Пяткина. - Электронные текстовые данные. - М. : Изд-во РУДН, 2017. - 40 с. - ISBN 978-5-209-07659-9 : 70.53. [http://lib.rudn.ru/MegaPro/UserEntry?Action=Rudn\\_FindDoc&id=454906&idb=0](http://lib.rudn.ru/MegaPro/UserEntry?Action=Rudn_FindDoc&id=454906&idb=0)
4. Елисеева, И. И. Эконометрика : учебник для магистров / И. И. Елисеева ; под ред. И. И. Елисеевой. — Москва : Издательство Юрайт, 2014. — 449 с. — (Серия : Бакалавр и магистр. Академический курс). — ISBN 978-5-9916-3202-7. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/376042>
5. Регрессионный анализ в пакете Eviews [Текст/электронный ресурс] : Учебно-методическое пособие / Д.А. Пяткина, С.И. Матюшенко. - Электронные текстовые данные. - М. : Изд-во РУДН, 2013. - 30 с. - ISBN 978-5-209-05291-3 : 61.38
6. Модели с фиктивными переменными и бинарным откликом в пакете Eviews [Текст/электронный ресурс] : Учебно-методическое пособие / Д.А. Пяткина. - Электронные текстовые данные. - М. : Изд-во РУДН, 2015. - 36 с. : ил. - ISBN 978-5-209-06176-2 : 63.33. <http://lib.rudn.ru/ProtectedView/Book/ViewBook/4396>
7. Соловьев Владимир Игоревич.  
Анализ данных в экономике: теория вероятностей, прикладная статистика, обработка и визуализация данных в Microsoft Excel : учебник / В.И. Соловьев. - Москва : КноРус, 2019. - 498 с. - (Бакалавриат). - ISBN 978-5-406-06940-0 : 1169.00.

*Дополнительная литература:*

1. Просветов Г.И. Эконометрика: задачи и решения : Учебно-методическое пособие / Г. И. Просветов. - М. : Изд-во РДЛ, 2012. - 104 с. : ил. - ISBN 5-93840-056-2 : 35.97

2. Носко В.П. Эконометрика. Элементарные методы и введение в регрессионный анализ временных рядов / В. П. Носко. - М. : ИЭПП, 2004. - 501 с. - ISBN 5-93255-141-0 : 70.00

3. Эконометрика [Текст] : Учебное пособие для вузов / С.А. Бородич. - 3-е изд., стереотип. - Минск : Новое знание, 2006. - 408 с. : ил. - (Экономическое образование). - ISBN 985-475-206-2 : 215.05

4. Зехин, В.А. Практикум по многомерным статистическим методам : учебное пособие / В.А. Зехин, В.С. Мхитарян, С.А. Айвазян. - 1-е изд. - Москва : Московский государственный университет экономики, статистики и информатики, 2003. - 76 с. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=90409>

*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.elsevierscience.ru/products/scopus/>

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

1. Курс лекций по дисциплине «Линейный и нелинейный регрессионный 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 «Линейный и нелинейный регрессионный 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*

Пяткина Дарья  
Анатольевна

*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.*