

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 29.05.2023 12:43:55

Уникальный программный ключ:

ca953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов»**

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:

09.03.03 Прикладная информатика

(код и наименование направления подготовки/специальности)

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

Прикладная информатика

(наименование (профиль/специализация) ОП ВО)

2023 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель преподавания дисциплины «Линейный и нелинейный регрессионный анализ больших данных» состоит в изучении теоретических основ анализа больших данных, включая базовые элементы статистического программирования и интеллектуального анализа больших наборов данных.

Задачи дисциплины – научить производить расчеты с применением технологий анализа больших данных и решать широкий спектр 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): УК-1; ОПК-1; ОПК-6; ПК-7

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	Способность 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.
ОПК-1	Способен применять естественнонаучные и общетехнические знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности	ОПК-1.1. Знает основы математики, физики, вычислительной техники и программирования.
		ОПК-1.2. Умеет решать стандартные профессиональные задачи с применением естественнонаучных и обще-инженерных знаний, методов математического анализа и моделирования
		ОПК-1.3. Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности
ОПК-10	Способен понимать принципы работы современных информационных технологий и	ОПК-10.1 Знает базовые принципы цифровых технологий и методов, необходимых в профессиональной деятельности в области прикладной информатики для: изучения и моделирования объектов профессиональной

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	использовать их для решения задач профессиональной деятельности	<p>деятельности, анализа данных, представления информации и пр.</p> <p>ОПК-10.2 Умеет применять необходимые в профессиональной деятельности цифровые технологии и методы в области прикладной информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.</p> <p>ОПК-10.3 Владеет необходимыми в профессиональной деятельности технологиями и методами в области прикладной информатики для: изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации и пр.</p>
ПК-7	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	<p>ПК-7.1. Знает основы научно-исследовательской деятельности в области информационных технологий, знает основные методы решения прикладных задач, современные методы информационных технологий</p> <p>ПК-7.2. Умеет применять полученные знания для решения стандартных задач в области информационных технологий и в собственной научно-исследовательской деятельности</p> <p>ПК-7.3. Имеет практический опыт научно-исследовательской деятельности в области информационных технологий</p> <p>ПК-7.4. Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет базовыми навыками подготовки научных обзоров и (или) публикаций, рефератов и библиографий по тематике проводимых исследований на русском и иностранном языке</p>

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Линейный и нелинейный регрессионный анализ больших данных» относится к части, формируемой участниками образовательных отношений блока Б1 ОП ВО.

В рамках ОП ВО 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 ¹
УК-1	Способность осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	Философия Социальные и этические вопросы информационных технологий Интеллектуальные системы Теоретические основы информатики Имитационное моделирование Разработка информационно-аналитических систем Математическое моделирование Компьютерный практикум по моделированию	Технологии интеллектуального анализа данных и прогнозирование Кибербезопасность предприятия Основы проектирования сетей и систем телекоммуникаций Моделирование сложно структурированных систем Компьютерный практикум по статистическому анализу данных Компьютерный практикум по интеллектуальным системам Обработка больших данных с использованием машинного обучения Анализ больших данных при моделировании сложно-структурированных систем Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы) Технологическая (проектно-технологическая) практика Научно-исследовательская работа Преддипломная практика
ОПК-1	Способность применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в	Математический анализ Линейная алгебра Дискретная математика и математическая логика Дифференциальные и разностные уравнения Теория вероятностей и математическая статистика Теория конечных графов Основы	Имитационное моделирование Моделирование сложно структурированных систем Разработка информационно-аналитических систем Математическое моделирование Технологии интеллектуального анализа данных и прогнозирование

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики
	профессиональной деятельности;	программирования Технология программирования Python и его приложения Вычислительные системы, сети и телекоммуникации Концепции современного естествознания Интеллектуальные системы Теоретические основы информатики Структуры данных и парадигмы программирования Компьютерный практикум по моделированию Компьютерный практикум по информационным технологиям	Кибербезопасность предприятия Основы проектирования сетей и систем телекоммуникаций Параллельное программирование Прикладной анализ данных с использованием языка Python Компьютерный практикум по статистическому анализу данных Компьютерный практикум по интеллектуальным системам Обработка больших данных с использованием машинного обучения Анализ больших данных при моделировании сложно-структурированных систем Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы) Технологическая (проектно-технологическая) практика Научно-исследовательская работа Преддипломная практика
ОПК-10	Способность понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	Основы программирования Технология программирования Python и его приложения Вычислительные системы, сети и телекоммуникации Архитектура компьютеров и операционные системы Интеллектуальные системы Теоретические основы информатики	Системы управления базами данных Технологии интеллектуального анализа данных и прогнозирование Кибербезопасность предприятия Основы проектирования сетей и систем телекоммуникаций Параллельное программирование Прикладной анализ данных с использованием языка Python

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики
		<p>Основы Web-технологий Основы администрирования операционных систем Логическое программирование Структуры данных и парадигмы программирования Имитационное моделирование Моделирование сложно структурированных систем Разработка информационно-аналитических систем Математическое моделирование Управление ИТ-сервисами и контентом Компьютерный практикум по моделированию Компьютерный практикум по информационным технологиям</p>	<p>Компьютерный практикум по статистическому анализу данных Компьютерный практикум по интеллектуальным системам 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</p>
ПК-7	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	-	<p>Имитационное моделирование Математическое моделирование Технологии интеллектуального анализа данных и прогнозирование Основы проектирования сетей и систем телекоммуникаций Компьютерный практикум по статистическому анализу данных Компьютерный практикум по интеллектуальным системам Обработка больших данных с использованием</p>

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики
			машинного обучения Анализ больших данных при моделировании сложно-структурированных систем Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы) Технологическая (проектно-технологическая) практика Научно-исследовательская работа Преддипломная практика

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Линейный и нелинейный регрессионный анализ больших данных» составляет 3 зачетных единиц.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)
		5
<i>Контактная работа, ак.ч.</i>	36	36
Лекции (ЛК)	18	18
Лабораторные работы (ЛР)	-	-
Практические/семинарские занятия (СЗ)	18	18
<i>Самостоятельная работа обучающихся, ак.ч.</i>	72	72
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>		
Общая трудоемкость дисциплины	ак.ч.	108
	зач.ед.	3

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Наименование разделов дисциплины	Содержание темы	Вид учебной работы
Раздел 1. Введение в большие данные.	Тема 1.1. Жизненный цикл аналитики данных	ЛК,СЗ
	Тема 1.2. Высокопроизводительные вычисления	ЛК,СЗ
	Тема 1.3. Масштабирования и многоуровневое хранение данных	ЛК,СЗ
Раздел 2. Основной аппарат комбинаторики и	Тема 2.1 Основные понятия математической статистики: статистические оценки, их свойств	ЛК,СЗ

Наименование разделов дисциплины	Содержание темы	Вид учебной работы
математической статистики	Тема 2.2. Проверка гипотез. Наблюдённый уровень значимости	ЛК,СЗ
Раздел 3. Детектирование выбросов и аномалий. Очистка данных и технологии регуляризации.	Тема 3.1. Детектирование выбросов и аномалий. Что такое выбросы, типы выбросов. Методы обнаружения выбросов .Поиск аномалий. Цензурирование выборок	ЛК,СЗ
	Тема 3.2. Отсев объектов-выбросов, удаление выбросов. Очистка данных и технологии регуляризации. Основные виды регуляризации. Метод редукции размерности. Методы отбора признаков.	ЛК,СЗ
Раздел 4. Нейронные сети. Генетические алгоритмы.	Тема 4.1 Нейронные сети: общая архитектура. Многослойные сети.	ЛК,СЗ
	Тема 4.2. Обратное распространение ошибки. Стохастический градиентный спуск. Генетические алгоритмы.	ЛК,СЗ
Раздел 5. Выделение особенностей (Feature detection); нормализация данных. Нечеткие множества. Байесовы сети.	Тема 5.1. Извлечение признаков / выделение особенностей (Feature detection). Преобразования признаков. Нормализация данных. Методы нормализации данных. Нормализация по методу минимакс	ЛК,СЗ
	Тема 5.2Нечёткие множества. Байесовы сети. Задачи байесовского вывода. Методика построения нечеткой байесовой сети	ЛК,СЗ
Раздел 6. Сложные методы аналитики	Тема 6.1. Множественная и нелинейная регрессии	ЛК,СЗ
	Тема 6.2 Классификация. Логит и пробит модели. Деревья.Случайный лес	ЛК,СЗ
	Тема 6.3. Кластеризация .	ЛК,СЗ
	Тема 6.4. Визуализация данных и результатов анализа	ЛК,СЗ

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	Компьютер/ноутбук с доступом сети Интернет и

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
	специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	электронно-образовательной среде Университета, браузер, ПО для просмотра PDF, MS Teams.
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве числа обучающихся), доской (экраном) и техническими средствами мультимедиа презентаций.	ОС Windows или Linux. Пакет Gretl, Пакет PSPP Дополнительное ПО: офисный пакет MS Office или LibreOffice
Для самостоятельной работы обучающихся	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

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 : 194.86. http://lib.rudn.ru/MegaPro/UserEntry?Action=Rudn_FindDoc&id=495562&idb=0
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
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
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

- ЭБС «Лань» <http://e.lanbook.com/>

- 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/модуля²:

1. Курс лекций по дисциплине «Линейный и нелинейный регрессионный анализ больших данных».
2. Методические указания к самостоятельным работам по дисциплине «Линейный и нелинейный регрессионный анализ больших данных».

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

Оценочные материалы и балльно-рейтинговая система³ 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.

РАЗРАБОТЧИКИ:

Доцент кафедры прикладной информатики и теории вероятностей Должность, БУП	 Подпись	Д.А. Пяткина Фамилия И.О.
РУКОВОДИТЕЛЬ БУП: Зав. кафедрой прикладной информатики и теории вероятностей Наименование БУП	 Подпись	К.Е. Самуйлов Фамилия И.О.
РУКОВОДИТЕЛЬ ОП ВО: Зав. кафедрой информационных технологий Должность, БУП	 Подпись	Ю.Н. Орлов Фамилия И.О.

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

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