

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 29.05.2024 11:28:25

Уникальный программный ключ:

ca953a01204891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Российский университет дружбы народов имени Патриса Лумумбы»**

**Факультет искусственного интеллекта**

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **ПРИКЛАДНАЯ СТАТИСТИКА И АНАЛИЗ ДАННЫХ**

(наименование дисциплины/модуля)

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

### **02.04.02 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):**

### **УПРАВЛЕНИЕ ДАННЫМИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ**

(наименование (профиль/специализация) ОП ВО)

**2024 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Прикладная статистика и анализ данных» входит в программу магистратуры «Управление данными и искусственный интеллект» по направлению 02.04.02 «Фундаментальная информатика и информационные технологии» и изучается в 1 семестре 1 курса. Дисциплину реализует NOT DEFINED. Дисциплина состоит из 7 разделов и 23 тем и направлена на изучение методов сбора, обработки и анализа данных с использованием статистических инструментов. Она позволяет студентам освоить практические навыки работы с данными, проводить их анализ и интерпретацию результатов, а также принимать обоснованные решения на основе полученных данных.

Целью освоения дисциплины является формирование у студентов знаний и навыков в области прикладной статистики и анализа данных. Студенты должны освоить методы сбора, обработки и анализа данных, научиться использовать статистические инструменты для принятия обоснованных решений на основе полученных результатов.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Прикладная статистика и анализ 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 проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 Знает принципы сбора, отбора и обобщения информации; УК-1.2 Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности;
УК-2	Способен управлять проектом на всех этапах его жизненного цикла	УК-2.2 Умеет определять круг задач в рамках избранных видов профессиональной деятельности, планировать собственную деятельность исходя из имеющихся ресурсов; соотносить главное и второстепенное, решать поставленные задачи в рамках избранных видов профессиональной деятельности; УК-2.3 Имеет практический опыт применения нормативной базы и решения задач в области избранных видов профессиональной деятельности;
ОПК-1	Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий	ОПК-1.1 Обладает фундаментальными знаниями в области математических и естественных наук, теории коммуникаций; ОПК-1.2 Умеет осуществлять первичный сбор и анализ материала, интерпретировать различные математические объекты; ОПК-1.3 Имеет практический опыт работы с решением математических задач и применяет его в профессиональной деятельности;
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования	ОПК-3.1 Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей;
ПК-1	Способен разрабатывать и применять алгоритмы	ПК-1.1 Знает существующие алгоритмы интеллектуальной обработки данных;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	интеллектуальной обработки данных для решения задач профессиональной деятельности	ПК-1.2 Умеет модифицировать алгоритмы интеллектуальной обработки данных;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Прикладная статистика и анализ 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 «Прикладная статистика и анализ 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-2	Способен управлять проектом на всех этапах его жизненного цикла		Технологическая (проектно-технологическая) практика (учебная); Преддипломная практика; Технологическая (проектно-технологическая) практика (производственная); Обучение с подкреплением; Управление проектами и Agile; Методы оптимизации;
УК-1	Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий		Алгоритмы и структуры данных; Методы к решению задач Information Extraction; Преддипломная практика;
ОПК-1	Способен находить, формулировать и решать актуальные проблемы прикладной математики, фундаментальной информатики и информационных технологий		Обучение с подкреплением; Машинный перевод; Вопросно-ответные системы; Методы к решению задач Information Extraction; Алгоритмы и структуры данных;
ОПК-3	Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования		Преддипломная практика; Методы глубокого обучения для обработки естественного языка; Вопросно-ответные системы; Методы к решению задач Information Extraction; Методы оптимизации; Практикум по

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
			программированию;
ПК-1	Способен разрабатывать и применять алгоритмы интеллектуальной обработки данных для решения задач профессиональной деятельности		Алгоритмы и структуры данных; Методы оптимизации; Методы глубокого обучения для обработки естественного языка; <i>Современные устройства центров обработки больших данных**;</i> <i>Современные инструментальные средства разработки ПО для искусственного интеллекта**;</i> Методы к решению задач Information Extraction; Вопросно-ответные системы; Технологическая (проектно-технологическая) практика (учебная);

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

\*\* - элективные дисциплины /практики

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

Общая трудоемкость дисциплины «Прикладная статистика и анализ данных» составляет «4» 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)
			1
<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>	81		81
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>144</b>	<b>144</b>
	<b>зач.ед.</b>	<b>4</b>	<b>4</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	Распределение вероятностей случайных величин	1.1	Предварительные замечания вероятностного и статистического характера.	ЛК, СЗ
		1.2	Дискретные распределения. Непрерывные распределения.	ЛК, СЗ
		1.3	Предварительные замечания об используемом математическом аппарате	ЛК, СЗ
		1.4	Многомерное нормальное распределение. Распределение линейной комбинации нормально распределенных величин; независимость величин; частные распределения. Условные распределения и множественный коэффициент корреляции. Характеристическая функция; моменты.	ЛК, СЗ
Раздел 2	Оценки параметров распределений вероятностей	2.1	Общие замечания. Оценки типа максимального правдоподобия (М- оценки). Линейные комбинации порядковых статистик (L-оценки). Оценки, получаемые в ранговых критериях (R-оценки). Оценки метода минимального расстояния (R-оценки). Байесовское оценивание.	ЛК, СЗ
		2.2	Робастность в статистике. Общие основы. Зачем нужны робастные процедуры. Качественная робастность. Количественная робастность. Инфинитезимальные аспекты. Общая идея вычисления робастных оценок. Робастность М-оценок. Параметрическая процедура отбраковки аномальных наблюдений.	ЛК, СЗ
		2.3	Исследование свойств оценок параметров распределений вероятностей по эмпирическим данным	ЛК, СЗ
		2.4	Байесовское оценивание	ЛК, СЗ
		2.5	Экспериментальное исследование робастности оценок	ЛК, СЗ
Раздел 3	Методы анализа законов распределения вероятностей случайных величин	3.1	Общие критерии согласия	ЛК, СЗ
		3.2	Специальные критерии согласия	ЛК, СЗ
		3.3	Критерии проверки отклонения распределения от нормального закона	ЛК, СЗ
		3.4	Специальные критерии согласия	ЛК, СЗ
		3.5	Подбор кривых распределения вероятностей по экспериментальным данным	ЛК, СЗ
Раздел 4	Проверка гипотез о значении параметров распределения	4.1	Сравнение параметров распределения.	ЛК, СЗ
		4.2	Непараметрические критерии однородности статистических данных. Критерии тренда и случайности. Толерантные пределы.	ЛК, СЗ
Раздел 5	Методы исследования связей между случайными величинами	5.1	Корреляционный анализ	ЛК, СЗ
		5.2	Дисперсионный анализ	ЛК, СЗ
		5.3	Регрессионный анализ	ЛК, СЗ
		5.4	Алгоритмы восстановления регрессии. МНК. Непараметрическая регрессия. Многомерная	ЛК, СЗ

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
			линейная регрессия. Нелинейные обобщения линейной регрессии. Логистическая регрессия	
Раздел 6	Статистические методы анализа многомерных данных	6.1	Методы снижения размерности. Методы факторного анализа. Многомерное шкалирование	ЛК, СЗ
		6.2	Многомерное шкалирование.	ЛК, СЗ
Раздел 7	Оценивание плотности и функции вероятности	7.1	Оценивание плотности и функции вероятности. Параметрические оценки плотности. Гистограммы и ядерные оценки плотности. Свойства ядерной оценки одномерной плотности Выбор ширины окна. Референтные эвристические правила. Подстановка. Кросс-валидация на основе наименьших квадратов. Кросс-валидация на основе правдоподобия. Бутстраповские методы. Частотная и ядерная оценка функции вероятности. Ядерное оценивание плотности в случае смешанных дискретных и непрерывных данных Построение доверительных интервалов Проклятие размерности. Оценивание условной плотности. Ядерное оценивание условной плотности. Ядерное оценивание функции условного распределения. Ядерное оценивание условной квантили. Модели бинарного выбора и счетных данных.	ЛК, СЗ

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/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.	Учебная аудитория для проведения занятий лекционного типа, практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Комплект специализированной мебели; технические средства: Моноблок

		<p>Lenovo Idea Centre, выход в интернет, Проектор EPSON, Интерактивный дисплей SMART MX SERIES SBID-MX065-v2. Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/Office 365, Teams)</p>
Семинарская	<p>Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.</p>	<p>Компьютерный класс для проведения занятий лекционного типа, практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Комплект специализированной мебели; технические средства: Моноблок MSI (no) - 13 Моноблок Lenovo c560 – 3, Моноблок hp pro one – 1, Интерактивная система Smart Unifi45 – 1, Проектор Notevision – 1, Проекционный экран Sactus – 1, Компьютерная гарнитура -17, Портативная камера для документов – 1. Имеется выход в интернет. Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/Office 365, Teams,).</p>
Для самостоятельной работы	<p>Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.</p>	<p>Компьютерный класс для самостоятельной работы обучающихся. Комплект специализированной мебели (11 посадочных мест); технические средства: Моноблок Lenovo Idea Centre, проектор ViewSonic p9d6253, имеется выход в</p>



		интернет, Проекционный экран Cactus, гарнитуры Plantronics Audio 655 Dsp Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/Office 365, Teams)
--	--	---

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Основная литература:

1. Михайлов, Геннадий Алексеевич. Численное статистическое моделирование. Методы Монте-Карло : учебное пособие для студентов вузов, обучающихся по направлению подготовки "Прикладная математика" / Г. А. Михайлов, А. В. Войтишек. - Москва : Академия, 2006 (Саратов : Саратовский полиграфкомбинат). - 366, [1] с.; 22 см. - (Университетский учебник. Серия "Прикладная математика и информатика"); ISBN 5-7695-2739-0 (В пер.)
2. Магнус, Ян Р. Матричное дифференциальное исчисление с приложениями к статистике и эконометрике : Пер. с англ. / Я.Р. Магнус, Х. Нейдеккер. - Перераб. изд. - Москва : Физматлит, 2002. - 495 с.; 24 см.; ISBN 5-9221-0262-1
3. Кендэл, Морис Джордж. Теория распределений / М. Дж. Кендэл, А. Стьюарт ; перевод с англ. В. В. Сазонова, А. Н. Ширяева ; под ред. А. Н. Колмогорова. - Москва : Наука, 1966. - 587 с. : черт.; 22 см.
4. Дэйвисон, М.Л. Многомерное шкалирование : Методы нагляд. представления данных / М. Дэйвисон; Пер. с англ. В. С. Каменского; [Предисл. С. А. Айвазяна, В. С. Каменского]. - Москва : Финансы и статистика, 1988. - 253,[1] с. : ил.; 21 см. - (Мат.-стат. методы за рубежом); ISBN 5-279-00276-3 (В пер.) : 2 р. 10 к.

### Дополнительная литература:

1. Горькова К. А., Абрамов Ю. Ш. Факторный анализ (метод главных компонент): учебное пособие Ленинград: Ленинградский финансово-экономический институт, 1981
2. Леман Э. Л., Прохоров А. В. Проверка статистических гипотез: перевод с английского Москва: Наука. Главная редакция физико-математической литературы [Физматлит], 1979

### 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 «Прикладная статистика и анализ 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 «Прикладная статистика и анализ данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

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

**РАЗРАБОТЧИК:**

<hr/>	<hr/>	<hr/>
<i>Должность, БУП</i>	<i>Подпись</i>	Андреева Марина Андреевна <i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ БУП:**

<hr/>	<hr/>	<hr/>
Заведующий кафедрой <i>Должность БУП</i>	<i>Подпись</i>	<i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ ОП ВО:**

<hr/>	<hr/>	<hr/>
<i>Должность, БУП</i>	<i>Подпись</i>	<i>Фамилия И.О.</i>