

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 29.06.2022 16:21:56  
Уникальный программный ключ:  
ca953a0120d891083f939673078ef1a989dae18a

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

**Инженерная академия**

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

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Анализ и классификация данных**

(наименование дисциплины)

Рекомендовано МССН для направления подготовки

**27.03.05 Инноватика**

(код и наименование направления подготовки)

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

**Управление инновациями в отраслях промышленности**

(наименование (направленность/профиль) ОП ВО)

Форма обучения: **очная**

2022 г.

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Анализ и классификация данных» является изучение базовых алгоритмов анализа и интерпретации табличных данных, формирование практических навыков работы с современными пакетами прикладных программ для решения задач анализа и интерпретации данных реализации систем цифровой обработки сигналов.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

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

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

Код компетенции	Наименование компетенции	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. Рассматривает возможные варианты решения задачи, оценивая Анализ и классификация данных их достоинства и недостатки

## 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Анализ и классификация данных» относится к части, формируемой участниками образовательных отношений дисциплина по выбору блока Б1.В.ДВ.08.01.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Анализ и классификация данных»

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

Код компетенции	Наименование компетенции	Предшествующие дисциплины/ модули, практики*	Последующие дисциплины/модули, практики*
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	Математика Численные методы и методы оптимизации в технике Введение в управление инновационными процессами Управление инновационной деятельностью в промышленности Основы инженерной экономики и менеджмента Управление инновациями на различных этапах жизненного цикла Логистика Дискретная математика в инженерных приложениях Планирование и контроллинг инновационных предприятий	Ознакомительная практика Проектная практика Организационно-управленческая практика Преддипломная практика Подготовка к сдаче и сдача государственного экзамена Выполнение, подготовка к процедуре защиты и защита выпускной квалификационной работы

	Численные методы и методы оптимизации в технике Маркетинг Природоохранная деятельность инновационного предприятия Основы информационной безопасности Основы применения данных дистанционного зондирования Земли и геоинформационных систем Теория автоматического управления Управление инновационными проектами Технологии виртуальной и дополненной реальности Системы управления базами данных Экономическая безопасность инновационного предприятия	
--	--	--

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

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

Общая трудоемкость дисциплины составляет 3 зачетных единиц.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО

Вид учебной работы	Всего часов	Семестр
		7
<i>Контактная работа, ак.ч.</i>	54	54
Лекции (ЛК)	18	18
Лабораторные работы (ЛР)	36	36
Практические/семинарские занятия (СЗ)		
<i>Самостоятельная работа обучающегося, ак.ч.</i>	54	54
<i>Контроль (экзамен), ак.ч.</i>		
<b>Общая трудоемкость дисциплины</b>	ак.ч.	108
	зач.ед.	3

#### 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины по видам учебной работы

Наименование раздела дисциплины	Содержание раздела (темы)	Виды учебной работы
<b>Раздел 1</b> Основные понятия дисциплины	Тема 1.1. Ведение в анализ данных. Проблема обработки данных. Матрица данных. Тема 1.2. Гипотезы компактности и скрытых факторов. Структура матрицы данных и задачи обработки. Тема 1.3. Матрица объект-объект и признак-признак. Расстояние и близость. Измерение признаков.. Тема 1.4. Отношения и их представление. Основные проблемы измерений. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных.	Л Л, СР СР СР
<b>Раздел 2</b>	Тема 2.1. Решающие поверхности и дискриминантные функции.	СР

Классификация данных с использованием детерминированных моделей.	<p>Линейные дискриминантные функции классификатор по минимуму расстояния.</p> <p>Тема 2.2. Линейная разделимость. Кусочно линейные дискриминантные функции.</p> <p>Тема 2.3. Нелинейные дискриминантные функции. Фи машины. Потенциальные функции как дискриминантные функции. Пространство весов. Процедуры обучения с коррекцией ошибок: правило с фиксированным приращением, правило абсолютной коррекции, частично корректирующее правило.</p> <p>Тема 2.4. Обобщенные градиентные методы</p>	<p>Л, СР</p> <p>Л, СР</p> <p>СР</p> <p>СР</p>
<b>Раздел 3</b> Классификация данных на основе статистических моделей.	<p>Тема 3.1. Функция потерь. Байесовская дискриминантная функция. Принятие решения по максимуму правдоподобия.</p> <p>Тема 3.2. Оптимальная дискриминантная функция для нормально распределенных образов. Дискриминантная функция Фишера.</p> <p>Тема 3.3. Множественный дискриминантный анализ. Пошаговый дискриминантный анализ. Ошибки классификации.</p> <p>Тема 3.4. Примеры построения статистических дискриминантных функций для различных статистических нескольких моделей данных.</p> <p>Тема 3.5. Обучение для статистических дискриминантных функций. Оценки максимального правдоподобия, байесовские оценки. Непараметрическое оценивание.</p> <p>Тема 3.6. Парзеновские окна, метод непараметрического оценивания на основе Кближайшего соседства.</p>	<p>Л, СР</p> <p>Л, СР</p> <p>СР</p> <p>Л, СР</p> <p>СР</p> <p>СР</p>
<b>Раздел 4</b> Кластер-анализ	<p>Тема 4.1. Основные типы задач кластер-анализа.</p> <p>Тема 4.2. Меры подобия и функции расстояния .</p> <p>Тема 4.3. Выбор критерия кластеризации.</p> <p>Тема 4.4. Кластерные методы, основанные на евклидовой метрике. Иерархическая кластеризация.</p> <p>Тема 4.5. Метод К-внутригрупповых средних</p> <p>Тема 4.6. Использование методов теории графов в задачах кластеризации. Кластеризация на основе анализа плотностей вероятностей.</p>	<p>Л, СР</p> <p>Л, СР</p> <p>Л, СР</p> <p>СР</p> <p>СР</p> <p>СР</p>
<b>Раздел 5</b> Методы снижения размерностей данных	<p>Тема 5.1. Анализ матриц исходных данных. Метод главных компонент. Корреляционная матрица и ее основные свойства.</p> <p>Тема 5.2. Собственные векторы и собственные числа корреляционной матрицы. Приведение корреляционной матрицы к диагональной форме.</p> <p>Тема 5.3. Геометрическая интерпретация главных компонент на плоскости. Модели факторного анализа. Оценка факторных нагрузок методом максимального правдоподобия и центроидным методом.</p> <p>Тема 5.4. Вращение факторов и их интерпретация. Использование кластеризации признаков для снижения размерности.</p> <p>Тема 5.5. Многомерное шкалирование (МИ). Метрический и неметрический подход к МИ. Методы ортогонального проектирования. Нелинейные методы МИ</p> <p>Тема 5.6. Многомерное шкалирование неметрических данных. Многомерные развертки.</p>	<p>Л, СР</p> <p>Л, СР</p> <p>СР</p> <p>СР</p> <p>СР</p> <p>СР</p> <p>СР</p>
<b>Раздел 6</b> Методы прогнозирования временных рядов	<p>Тема 6.1. Классификация методов прогнозирования. Оценивание трендов.</p> <p>Тема 6.2. Методы скользящего среднего.</p> <p>Тема 6.3. Экспоненциальное сглаживание. Регрессионный анализ и прогнозирование</p> <p>Тема 6.4. Линейные параметрические модели временных рядов.</p> <p>Тема 6.5. Методы оценивания моделей авторегрессии, скользящего среднего и смешанных моделей.</p> <p>Тема 6.6. Сезонные модели. Прогнозирование на основе параметрических моделей. Прогнозирование с использованием нейронных сетей.</p>	<p>Л, СР</p> <p>Л, СР</p> <p>СР</p> <p>СР</p> <p>СР</p> <p>СР</p>
<b>Раздел 7</b> Системы DATA MINING. в задачах анализа и интерпретации данных	<p>Тема 7.1. Понятие об интеллектуальных системах анализа и интерпретации данных.</p> <p>Тема 7.2. DATA MINING - системы извлечения новых знаний из данных.</p> <p>Тема 7.3. Типы систем DATA MINING -предметноориентированные аналитические системы, статистические пакеты, нейронные сети, деревья решений, обнаружение логических закономерностей, генетические алгоритмы, системы визуализации многомерных данных</p>	<p>Л, СР</p> <p>Л, СР</p> <p>СР</p>
<b>Раздел 8</b>	Тема 8.1. Табличные процессоры и базы данных в задачах обработки	Л, СР

Современные пакеты прикладных программ для решения задач обработки экспериментальных данных	данных. Виды статистических пакетов. Тема 8.2. Требования к статистическим пакетам общего назначения. Общая характеристика пакетов "STATGRAFICS Plus", "STATISTICA", SAS, SPSS . Комплексные системы класса DATA MINING для обработки данных - "PolyAnalist", Intelligent Miner.	Л,СР
---	---	------

## 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	
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием	
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	

аудитория для самостоятельной работы обучающихся указывается обязательно

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

*Основная литература:*

- 1) Балдин К.В., Уткин В.Б. Информационные системы в экономике: Учебник. - Дашков и К, 2015. - 395 с. // <http://www.knigafund.ru/books/16943> 1. Нестеров С.А. Интеллектуальный анализ данных средствами MS SQLServer 2008 Национальный Открытый Университет «ИНТУИТ» 2016 г. 338 страниц.
- 2) Сирота А.А. Методы и алгоритмы анализа данных и их моделирование в Matlab. Учебное пособие, Санкт-Петербург, БХВ-Петербург, 2016 г. – 376 с.

*Дополнительная литература:*

- 1) Тюрин Ю.Н. Анализ данных на компьютере: учебное пособие. – 4-е изд. – М.: МД ФОРУЧ, 2013. – 368 с. (гриф) 2. Кудинов Ю.И., Пашенко Ф.Ф. Основы современной информатики: учебное пособие. - С.Пб. :Изд-во «Лань», 2009. – 256 с. (гриф) 3. Гуда А.Н. Информатика и программирование: компьютерный практикум. - Ростов/н/Д.: Наука-Спектр; М.: Изд-во "Дашков и К", 2010. - 240 с. (гриф)

*Ресурсы информационно-телекоммуникационной сети «Интернет»:*

1) Электронно-библиотечная система (ЭБС) РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров:

– 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 «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/модуля\*:*

1) Курс лекций по дисциплине «Анализ и классификация 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 и классификация данных» представлены в Приложении к настоящей Рабочей программе дисциплины.

\* Ом и БРС формируются на основании требований соответствующего локального нормативного акта

### **Разработчик:**

Доцент департамента механики и процессов управления,  
к.ф.-м.н., доцент



О.А. Салтыкова

### **Руководитель базового учебного подразделения:**

Директор департамента механики и процессов управления,  
Д-р.т.н., профессор



Ю.Н. Разумный

### **Руководитель программы:**

Доцент департамента инновационного менеджмента  
в отраслях промышленности, к.э.н., доцент



Ю.А. Назарова