

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**

**Data analysis and classification / 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**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Data analysis and classification / 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 «Data analysis and classification / 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 их достоинства и недостатки

## 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Data analysis and classification / Анализ и классификация данных» относится к части, формируемой участниками образовательных отношений дисциплина по выбору блока Б1.В.ДВ.08.02.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Data analysis and classification / 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	Преддипломная практика Подготовка к сдаче и сдача государственного экзамена Выполнение, подготовка к процедуре защиты и защита выпускной квалификационной работы

\* - 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
Контактная работа, ак.ч.	54	54
Лекции (ЛК)	18	18
Лабораторные работы (ЛР)	36	36
Практические/семинарские занятия (СЗ)		
Самостоятельная работа обучающегося, ак.ч.	54	54
Контроль (экзамен), ак.ч.		
Общая трудоемкость дисциплины	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> Basic concepts of the discipline/ Основные понятия дисциплины	Тема 1.1. Conducting data analysis. Data processing problem. Data matrix/ Ведение в анализ данных. Проблема обработки данных. Матрица данных. Тема 1.2. Hypotheses of compactness and hidden factors. Data matrix structure and processing tasks/ Гипотезы компактности и скрытых факторов. Структура матрицы данных и задачи обработки. Тема 1.3. Object-object matrix and feature-attribute. Distance and proximity. Measurement of features/ Матрица объект-объект и признак-признак. Расстояние и близость. Измерение признаков.. Тема 1.4. Relationships and their presentation. Basic measurement problems. The main types of scales. Adequacy problem. The main tasks of data analysis and interpretation/ Отношения и их представление. Основные проблемы измерений. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных.	Л Л, СР СР СР СР
<b>Раздел 2</b> Data classification using deterministic models/ Классификация данных с использованием детерминированных моделей.	Тема 2.1. Decision surfaces and discriminant functions. Linear discriminant functions classifier by minimum distance. Linear separability/ Решающие поверхности и дискриминантные функции. Линейные дискриминантные функции классификатор по минимуму расстояния. Тема 2.2. Piecewise linear discriminant functions. Nonlinear discriminant functions/ Линейная разделимость. Кусочно линейные дискриминантные функции. Тема 2.3. Fimashins. Potential functions as discriminant functions. Weights space. Error correction learning procedures: fixed increment rule, absolute correction rule, partially correcting rule./ 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. Тема 2.4. Generalized gradient methods/ Обобщенные градиентные методы	СР Л, СР Л, СР СР СР
<b>Раздел 3</b> Data classification based on statistical models/ Классификация данных на основе статистических моделей.	Тема 3.1. Loss function. Bayesian discriminant function. Making a decision with maximum likelihood/ Функция потерь. Байесовская дискриминантная функция. Принятие решение по максимуму правдоподобия. Тема 3.2. Optimal discriminant function for normally distributed images. Fisher's discriminant function/ Оптимальная дискриминантная функция для нормально распределенных образов. Дискриминантная функция Фишера.	Л, СР Л, СР

	<p>Тема 3.3. Multiple Discriminant Analysis. Stepwise discriminant analysis. Classification errors./ Множественный дискриминантный анализ. Пошаговый дискриминантный анализ. Ошибки классификации.</p> <p>Тема 3.4. Examples of constructing statistical discriminant functions for various statistical several data models/ Примеры построения статистических дискриминантных функций для различных статистических нескольких моделей данных.</p> <p>Тема 3.5. Training for statistical discriminant functions. Maximum likelihood estimates, Bayesian estimates. Nonparametric estimation/ Обучение для статистических дискриминантных функций. Оценки максимального правдоподобия, байесовские оценки. Непараметрическое оценивание.</p> <p>Тема 3.6. Parzen windows, a nonparametric estimation method based on the K nearest neighbor/ Парзеновские окна, метод непараметрического оценивания на основе Кближайшего соседства.</p>	<p>CP</p> <p>Л, CP</p> <p>CP</p> <p>CP</p> <p>Л,CP</p>
<p><b>Раздел 4</b> Cluster analysis/ Кластер-анализ</p>	<p>Тема 4.1. The main types of cluster analysis tasks./ Основные типы задач кластер-анализа.</p> <p>Тема 4.2. Similarity measures and distance functions./ Меры подобия и функции расстояния .</p> <p>Тема 4.3. Selection of the clustering criterion./ Выбор критерия кластеризации.</p> <p>Тема 4.4. Cluster methods based on the Euclidean metric. Hierarchical clustering./ Кластерные методы, основанные на евклидовой метрике. Иерархическая кластеризация.</p> <p>Тема 4.5. Method of K-within-group means. / Метод К-внутригрупповых средних</p> <p>Тема 4.6. The use of graph theory methods in clustering problems. Clustering based on the analysis of probability densities./ Использование методов теории графов в задачах кластеризации. Кластеризация на основе анализа плотностей вероятностей.</p>	<p>Л, CP</p> <p>Л,CP</p> <p>Л, CP</p> <p>CP</p> <p>CP</p> <p>CP</p>
<p><b>Раздел 5</b> Data Dimension Reduction Techniques/ Методы снижения размерностей данных</p>	<p>Тема 5.1. Analysis of input data matrices. Principal component method. Correlation matrix and its main properties/ Анализ матриц исходных данных. Метод главных компонент. Корреляционная матрица и ее основные свойства.</p> <p>Тема 5.2. Eigenvectors and eigenvalues of the correlation matrix. Bringing the correlation matrix to a diagonal form./ Собственные векторы и собственные числа корреляционной матрицы. Приведение корреляционной матрицы к диагональной форме.</p> <p>Тема 5.3. Geometric interpretation of the principal components in the plane. Factor analysis models. Estimation of factor loadings by maximum likelihood and centroid methods/ Геометрическая интерпретация главных компонент на плоскости. Модели факторного анализа. Оценка факторных нагрузок методом максимального правдоподобия и центроидным методом.</p> <p>Тема 5.4. Rotation of factors and their interpretation. Using feature clustering for dimensionality reduction./ Вращение факторов и их интерпретация. Использование кластеризации признаков для снижения размерности.</p> <p>Тема 5.5. Multidimensional scaling (MI). Metric and non-metric approach to MI. Orthogonal design methods. Nonlinear MI methods./ Многомерное шкалирование (МИ). Метрический и неметрический подход к МИ. Методы ортогонального проектирования. Нелинейные методы МИ</p> <p>Тема 5.6. Multidimensional scaling of non-metric data. Multidimensional sweeps./ Многомерное шкалирование неметрических данных. Многомерные развертки.</p>	<p>Л, CP</p> <p>Л, CP</p> <p>CP</p> <p>CP</p> <p>CP</p> <p>Л, CP</p> <p>Л,CP</p> <p>Л,CP</p>
<p><b>Раздел 6</b> Time series forecasting methods/ Методы прогнозирования временных рядов</p>	<p>Тема 6.1. Classification of forecasting methods. Assessing trends./ Классификация методов прогнозирования. Оценивание трендов.</p> <p>Тема 6.2. Moving average methods./ Методы скользящего среднего.</p> <p>Тема 6.3. Exponential smoothing. Regression analysis and forecasting./ Экспоненциальное сглаживание. Регрессионный анализ и прогнозирование</p> <p>Тема 6.4. Linear parametric time series models./ Линейные параметрические модели временных рядов.</p> <p>Тема 6.5. Estimation methods for autoregressive models, moving average and mixed models. / Методы оценивания моделей авторегрессии, скользящего</p>	<p>Л,CP</p> <p>Л,CP</p> <p>CP</p> <p>CP</p>

	<p>среднего и смешанных моделей.</p> <p>Тема 6.6. Seasonal models. Forecasting based on parametric models. Forecasting using neural networks./ Сезонные модели. Прогнозирование на основе параметрических моделей. Прогнозирование с использованием нейронных сетей.</p>	<p>СР</p> <p>Л, СР</p>
<p><b>Раздел 7</b></p> <p>DATA MINING systems. in the tasks of data analysis and interpretation/ Системы DATA MINING. в задачах анализа и интерпретации данных</p>	<p>Тема 7.1. The concept of intelligent systems for data analysis and interpretation. / Понятие об интеллектуальных системах анализа и интерпретации данных.</p> <p>Тема 7.2. DATA MINING - systems for extracting new knowledge from data./ DATA MINING - системы извлечения новых знаний из данных.</p> <p>Тема 7.3. Types of DATA MINING systems - subject-oriented analytical systems, statistical packages, neural networks, decision trees, detection of logical patterns, genetic algorithms, multidimensional data visualization systems/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>Л, СР</p>
<p><b>Раздел 8</b></p> <p>Modern software packages for solving problems of experimental data processing/ Современные пакеты прикладных программ для решения задач обработки экспериментальных данных</p>	<p>Тема 8.1. Table processors and databases in data processing tasks. Types of statistical packages./ Табличные процессоры и базы данных в задачах обработки данных. Виды статистических пакетов.</p> <p>Тема 8.2. Requirements for general-purpose statistical packages. General characteristics of the packages "STATGRAFICS Plus", "STATISTICA", SAS, SPSS. Integrated systems of the DATA MINING class for data processing - "PolyAnalist", Intelligent Miner./ 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.</p>	<p>Л,СР</p> <p>Л,СР</p>

## 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 техническим средствами мультимедиа презентаций	
Семинарская	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническим средствами мультимедиа презентаций Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели 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 «Data analysis and classification / 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 «Data analysis and classification / 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