

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ястребов Олег Александрович  
Должность: Ректор  
Дата подписания: 06.06.2023 12:14:33  
Уникальный программный ключ:  
ca953a01204891083f939673078ef1a989dae18a

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

**Инженерная академия**

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

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **АНАЛИЗ ДАННЫХ И МАШИННОЕ ОБУЧЕНИЕ**

(наименование дисциплины/модуля)

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

### **27.03.04 УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ**

(код и наименование направления подготовки/специальности)

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

### **DATA ENGINEERING, ПРОГРАММИРОВАНИЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ**

(наименование (профиль/специализация) ОП ВО)

**2023 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Анализ данных и машинное обучение» входит в программу бакалавриата «Data Engineering, программирование и компьютерное моделирование интеллектуальных систем» по направлению 27.03.04 «Управление в технических системах» и изучается в 6, 7 семестрах 3, 4 курсов. Дисциплину реализует Департамент механики и процессов управления. Дисциплина состоит из 5 разделов и 22 тем и направлена на изучение фундаментальных основ аппарата комбинаторики и мат. статистики, регрессионного анализа и сжатия данных, детектированных выбросов и аномалий, очистки данных и технологий регуляризации, технологий кластеризации и классификации, нейронных сетей, генетических алгоритмов, выделения особенностей (feature detection), нормализации данных, нечетких множеств, байесовых сетей; разбор основных методов решения типовых задач и знакомство с областью их применения в профессиональной деятельности.

Целью освоения дисциплины является формирование фундаментальных знаний и навыков применения методов решения задач, необходимых для профессиональной деятельности (например, последующая разработка стратегий продвижения продуктов и услуг в цифровом пространстве), повышение общего уровня цифровой грамотности студентов,

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Анализ данных и машинное обучение» 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-6	Способен разрабатывать и использовать алгоритмы и программы, современные информационные технологии, методы и средства контроля, диагностики и управления, пригодные для практического применения в сфере своей профессиональной деятельности	ОПК-6.1 Знает основные алгоритмы и программы, современные информационные технологии, методы и средства контроля, диагностики и управления, пригодные для практического применения в сфере своей профессиональной деятельности; ОПК-6.2 Умеет применять алгоритмы и программы, современные информационные технологии, методы и средства контроля, диагностики и управления, пригодные для практического применения в сфере своей профессиональной деятельности; ОПК-6.3 Уверенно владеет алгоритмами и и программы, современными информационными технологиями, методами и средствами контроля, диагностикой и управлением, пригодными для практического применения в сфере своей профессиональной деятельности;
ПК-3	Способен реализовать корректную обработку данных, эффективный обмен данными и проведение базовой разведки больших сложных наборов данных	ПК-3.1 Знать методы обработки данных, технологии и языки манипулирования данными; ПК-3.2 Уметь применять технологии и языки манипулирования большими сложными наборами данных; ПК-3.3 Владеть технологиями и языками манипулирования и обработки данными;
ПК-6	Способен участвовать в подготовке технико-экономического обоснования проектов создания систем и средств автоматизации и управления техническими	ПК-6.1 Знает способы реализации основных технологических процессов; закономерности построения автоматизированных и автоматических производственных процессов; способы подготовки технико-экономического обоснования проектов создания систем и средств автоматизации и управления;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	проектами гражданской тематики	ПК-6.2 Умеет рационально выбирать различные варианты средств автоматизации, в том числе и вспомогательных, проектировать системы автоматизации с использованием микропроцессорной техники; выполнять работы по автоматизации технологических процессов и производств их обеспечению средствами автоматизации и управления; использовать современные методы и средства автоматизации, контроля, диагностики, испытаний;; ПК-6.3 Владеет способами реализации основных технологических процессов; навыками к практическому освоению и совершенствованию систем автоматизации производственных и технологических процессов, контроля, диагностики, испытаний, управления процессами; навыками разработок обобщенных вариантов решения проблем, связанных с автоматизацией производств, выборе на основе анализа вариантов оптимального, прогнозировании последствий решения;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Анализ данных и машинное обучение» 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 «Анализ данных и машинное обучение».

*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-6	Способен разрабатывать и использовать алгоритмы и программы, современные информационные технологии, методы и средства контроля, диагностики и управления, пригодные для практического применения в сфере своей профессиональной деятельности	Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы); Теория автоматического управления; Информатика и программирование; Механика космического полета; Алгоритмы и структуры данных;	Преддипломная практика;
ПК-3	Способен реализовать корректную обработку данных, эффективный обмен данными и проведение базовой разведки больших сложных наборов данных	Анализ геоинформационных данных; Алгоритмы и структуры данных;	Преддипломная практика;
ПК-6	Способен участвовать в подготовке технико-экономического обоснования проектов создания систем и средств	Теория автоматического управления;	Преддипломная практика;

<b>Шифр</b>	<b>Наименование компетенции</b>	<b>Предшествующие дисциплины/модули, практики*</b>	<b>Последующие дисциплины/модули, практики*</b>
	автоматизации и управления техническими проектами гражданской тематики		

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

\*\* - элективные дисциплины /практики

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

Общая трудоемкость дисциплины «Анализ данных и машинное обучение» составляет «7» зачетных единиц.

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)	
			6	7
<i>Контактная работа, ак.ч.</i>	144		72	72
Лекции (ЛК)	72		36	36
Лабораторные работы (ЛР)	72		36	36
Практические/семинарские занятия (СЗ)	0		0	0
<i>Самостоятельная работа обучающихся, ак.ч.</i>	90		54	36
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	18		18	0
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>252</b>	<b>144</b>	<b>108</b>
	<b>зач.ед.</b>	<b>7</b>	<b>4</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 машинное обучение и обработку данных. Программный инструментарий Data Mining и Machine Learning.	1.1	Введение в машинное обучение и обработку данных. Постановка основных классов задач в машинном обучении.	ЛК, ЛР
		1.2	Регрессия и классификация; кластеризация, снижение размерности	ЛК, ЛР
		1.3	Обработка текстов; обработка изображений	ЛК, ЛР
Раздел 2	Основной аппарат комбинаторики и мат. статистики. Регрессионный анализ и сжатие данных.	2.1	Основной аппарат комбинаторики и мат. статистики. Основные понятия математической статистики	ЛК, ЛР
		2.2	Статистические оценки, их свойства, проверка гипотез. Регрессионный анализ и сжатие данных.	ЛК, ЛР
		2.3	Задача регрессии. Минимизация квадрата отклонения. Регрессионная функция: условное мат. ожидание	ЛК, ЛР
		2.4	Линейная регрессия и метод k ближайших соседей. Переобучение и недообучение	ЛК, ЛР
		2.5	Разложение ошибки на шум, смещение и разброс	ЛК, ЛР
Раздел 3	Детектирование выбросов и аномалий. Очистка данных и технологии регуляризации.	3.1	Детектирование выбросов и аномалий. Что такое выбросы, типы выбросов	ЛК, ЛР
		3.2	Методы обнаружения выбросов. Поиск аномалий	ЛК, ЛР
		3.3	Цензурирование выборки. Отсев объектов-выбросов, удаление выбросов	ЛК, ЛР
		3.4	Очистка данных и технологии регуляризации. Основные виды регуляризации	ЛК, ЛР
		3.5	Метод редукции размерности. Методы отбора признаков	ЛК, ЛР
Раздел 4	Технологии кластеризации и классификации. Нейронные сети. Генетические алгоритмы.	4.1	Технологии кластеризации и классификации. K-means. EM-алгоритм	ЛК, ЛР
		4.2	Другие методы кластеризации. Задачи классификации. Байесовский классификатор	ЛК, ЛР
		4.3	Линейные методы для классификации. Логистическая регрессия, максимизация правдоподобия	ЛК, ЛР
		4.4	Нейронные сети: общая архитектура. Многослойные сети. Обратное распространение ошибки	ЛК, ЛР
		4.5	Стохастический градиентный спуск. Генетические алгоритмы	ЛК, ЛР
Раздел 5	Выделение особенностей (Feature detection); нормализация данных. Нечеткие множества. Байесовы сети.	5.1	Извлечение признаков / выделение особенностей (Feature detection)	ЛК, ЛР
		5.2	Преобразования признаков. Нормализация данных. Методы нормализации данных	ЛК, ЛР
		5.3	Нормализация по методу минимакса. Нормализация по Z-показателю. Десятичное масштабирование	ЛК, ЛР
		5.4	Нечеткие множества. Байесовы сети. Задачи байесовского вывода. Методика построения нечеткой байесовой сети	ЛК, ЛР

\* - 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.	
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 15 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
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.	

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

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Основная литература:

1. James, G. et al. An introduction to statistical learning. – Springer, 2013. – 426 pp
2. Trevor Hastie, Robert Tibshirani, et al., The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, 2017
3. Вьюгин, В. В. Математические основы машинного обучения и прогнозирования : учебное пособие / В. В. Вьюгин. — Москва : МЦНМО, 2014. — 304 с.

### Дополнительная литература:

1. Bruce, P. C., & Bruce, A. (2017). Practical Statistics for Data Scientists : 50 Essential Concepts (Vol. First edition). Sebastopol, CA: O'Reilly Media
2. Molnar, C. (2018). iml: An R package for Interpretable Machine Learning
3. Explainable and interpretable models in computer vision and machine learning. (2018)
4. Комбинаторика и теория вероятностей, учебное пособие, 99 с., Райгородский, А. М., 2013

### 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 «Анализ данных и машинное обучение».

\* - 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 данных и машинное обучение» представлены в Приложении к настоящей Рабочей программе дисциплины.

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



**РАЗРАБОТЧИК:**

Доцент

*Должность, БУП*



*Подпись*

Салтыкова Ольга  
Александровна

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ БУП:**

Директор ДМПУ

*Должность БУП*



*Подпись*

Разумный Юрий  
Николаевич

*Фамилия И.О.*

**РУКОВОДИТЕЛЬ ОП ВО:**

Профессор

*Должность, БУП*



*Подпись*

Разумный Юрий  
Николаевич

*Фамилия И.О.*