

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 01.06.2023 11:16:49
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

**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

МАШИННОЕ ОБУЧЕНИЕ И АНАЛИЗ БОЛЬШИХ ДАННЫХ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

27.04.04 УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

(код и наименование направления подготовки/специальности)

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

КОСМИЧЕСКАЯ ИНЖЕНЕРИЯ

(наименование (профиль/специализация) ОП ВО)

2023 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Машинное обучение и анализ больших данных» входит в программу магистратуры «Космическая инженерия» по направлению 27.04.04 «Управление в технических системах» и изучается в 3 семестре 2 курса. 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 «Машинное обучение и анализ больших 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 критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	УК-1.1 Анализирует задачу, выделяя ее базовые составляющие;; УК-1.2 Определяет и ранжирует информацию, требуемую для решения поставленной задачи;; УК-1.3 Осуществляет поиск информации для решения поставленной задачи по различным типам запросов;; УК-1.4 Предлагает варианты решения задачи, анализирует возможные последствия их использования;; УК-1.5 Анализирует пути решения проблем мировоззренческого, нравственного и личностного характера на основе использования основных философских идей и категорий в их историческом развитии и социально-культурном контексте.;
ПК-1	Способен формулировать цели, задачи научных исследований в области управления аэрокосмическими системами, выбирать методы и средства решения профессиональных задач	ПК-1.1 Знает методы и средства решения задач научных исследований в области систем искусственного интеллекта и робототехнических систем; ПК-1.2 Умеет формулировать цель и задачи научных исследований в профессиональной области; ПК-1.3 Владеет приемами для формулировки цели и задач научных исследований, умеет выбирать методы и средства решения задач профессиональной деятельности;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов в области управления аэрокосмическими	ПК-2.1 Знает современные теоретические и экспериментальные методы, применяемые для разработки математических моделей исследуемых объектов и процессов профессиональной деятельности; ПК-2.2 Умеет определять эффективность применяемых методов для разработки математических моделей исследуемых объектов и процессов;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	системами	ПК-2.3 Владеет современными теоретическими и экспериментальными методами для разработки математических моделей объектов и процессов профессиональной деятельности по направлению подготовки;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Машинное обучение и анализ больших данных» относится к части, формируемой участниками образовательных отношений 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 «Машинное обучение и 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*
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	Applied Problems of Mathematical Modeling; Numerical Methods for Solving Mathematical Modeling Problems; Research work / Научно-исследовательская работа;	Undergraduate practice / Преддипломная практика;
ПК-1	Способен формулировать цели, задачи научных исследований в области управления аэрокосмическими системами, выбирать методы и средства решения профессиональных задач	Research work / Научно-исследовательская работа; Dynamics and Control of Space Systems;	Undergraduate practice / Преддипломная практика;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов в области управления аэрокосмическими системами	Research work / Научно-исследовательская работа; Applied Problems of Mathematical Modeling; Contemporary Problems of Control Theory; Dynamics and Control of Space Systems; Applied Mechanics and Engineering;	Undergraduate practice / Преддипломная практика;

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

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Машинное обучение и анализ больших 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)
			3
<i>Контактная работа, ак.ч.</i>	72		72
Лекции (ЛК)	36		36
Лабораторные работы (ЛР)	36		36
Практические/семинарские занятия (СЗ)	0		0
<i>Самостоятельная работа обучающихся, ак.ч.</i>	45		45
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
Общая трудоемкость дисциплины	ак.ч.	144	144
	зач.ед.	4	4

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

- 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 «Машинное обучение и анализ больших 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 «Машинное обучение и анализ больших данных» 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.