

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 29.06.2022 15:09:29
Уникальный программный ключ:
ca953a0120d891083f939673078e1a987dae18a

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

Инженерная академия

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Обработка больших данных

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

01.04.02 «Прикладная математика и информатика»

(код и наименование направления подготовки/специальности)

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

Data Science и цифровая трансформация

(наименование (профиль/специализация) ОП ВО)

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью изучения дисциплины «Обработка больших данных» заключается в ознакомлении с базовыми понятиями и основными алгоритмами обработки больших данных, особенностями их применения.

Задачи дисциплины:

- 1) предоставление студентам знаний о методах подготовки к обработке и анализу информации, полученной в виде определенных данных;
- 2) формирование у студентов умений и навыков практической работы по адекватному применению соответствующих методов, техник и процедур при подготовке к обработке, анализу, научной интерпретации информации;
- 3) формирования представления у студентов об использовании результатов исследований в практике.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Big Data Mining / Обработка 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	Способность осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1 - Анализирует задачу, выделяя ее базовые составляющие;
		УК-1.2 - Определяет и ранжирует информацию, требуемую для решения поставленной задачи;
ПК-1	Способен формулировать цели, задачи научных исследований в области прикладной математики и информатики, вычислительной техники и современных технологий программирования, выбирать методы и средства решения задач	ПК-1.1 - Обладает фундаментальными знаниями, полученными в области математических и (или) естественных наук, программирования и информационных технологий
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов, относящихся	ПК-2.1 - Знает современные теоретические и экспериментальные методы разработки математических моделей, инновационные инструментальные средства проектирования и элементы архитектурных решений информационных систем

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	к профессиональной деятельности по направлению подготовки и участвовать в их реализации в виде программных продуктов	

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Big Data Mining / Обработка больших данных» относится к *вариативной* части блока 1 учебного плана.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Big Data Mining / Обработка 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	Способность осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий		Искусственные нейронные сети (Глубокое обучение) Научно-исследовательская работа
ПК-1	Способен формулировать цели, задачи научных исследований в области прикладной математики и информатики, вычислительной техники и современных технологий программирования, выбирать методы и		Искусственные нейронные сети (Глубокое обучение) Научно-исследовательская работа

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики
	средства решения задач		
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов, относящихся к профессиональной деятельности по направлению подготовки и участвовать в их реализации в виде программных продуктов		Искусственные нейронные сети (Глубокое обучение) Научно-исследовательская работа

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины составляет 6 зачетных единиц.

Вид учебной работы	Всего часов	Семестр
		1
Аудиторные занятия (всего)	51	51
В том числе:	-	-
<i>Лекции</i>	34	34
<i>Практические занятия (ПЗ)</i>	17	17
<i>Семинары (С)</i>	17	17
<i>Лабораторные работы (ЛР)</i>	17	17
Самостоятельная работа (всего)	112	112
Общая трудоемкость час	216	216
зач. ед.	6	6

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Наименование раздела (темы) дисциплины	Содержание раздела (темы)	Вид учебной работы
Раздел Введение	1. Тема 1.1. Обзор задач, решаемых алгоритмами машинного обучения	ЛК, ЛР

предмет «Обработка больших данных»		Тема 1.2. Классификация алгоритмов машинного обучения	ЛК, ЛР
Раздел 2. Линеинные модели регрессии.		Тема 2.1. Линейная регрессия	ЛК, ЛР
		Тема 2.2. Линейные модели регрессии	ЛК, ЛР
		Тема 2.3. Базисные функции	ЛК, ЛР
		Тема 2.4. Регуляризация	ЛК, ЛР
Раздел 3. Логистическая регрессия		Тема 3.1. Целевая функция логистической регрессии	ЛК, ЛР
		Тема 3.2. Регуляризация логистической регрессии	ЛК, ЛР
Раздел 4. Кластер-анализ.		Тема 4.1. Основные типы задач кластер-анализа	ЛК, ЛР
		Тема 4.2. Меры подобия и функции расстояния	ЛК, ЛР
		Тема 4.3. Выбор критерия кластеризации	ЛК, ЛР
		Тема 4.4. Кластерные методы, основанные на евклидовой метрике	ЛК, ЛР
		Тема 4.5. Иерархическая кластеризация	ЛК, ЛР
		Тема 4.6. Метод К-внутригрупповых средних	ЛК, ЛР
		Тема 4.7. Использование методов теории графов в задачах кластеризации	ЛК, ЛР
		Тема 4.8. Кластеризация на основе анализа плотностей вероятностей	ЛК, ЛР
Раздел 5. Нейронные сети		Тема 5.1. Структура нейрона	ЛК, ЛР
		Тема 5.2. Структура нейронной сети	ЛК, ЛР
		Тема 5.3. Обучения нейронной сети с помощью алгоритма обратного распространения ошибки	ЛК, ЛР
Раздел 6. Деревья решений		Тема 6.1. Структура деревьев решений	ЛК, ЛР
		Тема 6.2. Виды разделяющих функций	ЛК, ЛР
		Тема 6.3. Обучения дерева решений	ЛК, ЛР
		Тема 6.4. Алгоритм Random Forest	ЛК, ЛР
Раздел 7. Классификация		Тема 7.1. Обзор существующих алгоритмов классификации	ЛК, ЛР
		Тема 7.2. Алгоритм k-means	ЛК, ЛР

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	

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
	техническими средствами мультимедиа презентаций	
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием.	
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели 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 ___ шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
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.	

а) программное обеспечение: Windows, Microsoft Office, SMathStudio, SciLab, Code::Blocks

б) базы данных, информационно-справочные и поисковые системы: www.yandex.ru, www.google.ru, www.mathnet.ru, http://www.math.ru, http://eqworld.ipmnet.ru, http://gen.lib.rus.ec, www.twirpx.com.

а) основная литература

1. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с Microsoft Word - BigData_Radchenko_final.docx (ifmo.ru)
2. Теоретический минимум по Big Data. Анналин Ын, Кеннет Су <https://pps.kaznu.kz/kz/Main/FileShow/2016766/102/446/12258/Ибраев%20Гулама-Гарип%20Алишер%20Ерикжанович/2020/1>
3. ГрасД. Data Science. Наука о данных с нуля: Пер. с англ. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2021. - 416 с.: ил. ISBN 978-5-9775-6731-2 Data-Science-Nauka-o-dannyh-s-nulya_RuLit_Me_652328.pdf

б) дополнительная литература

1. Основы Data Science и Big Data, Python и наука о данных, Силен Д., Мейсман А., Али М., 2017. Основы Data Science и Big Data, Python и наука о данных, Силен Д., Мейсман А., Али М., 2017 (obuchalka.org)

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 «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/>




8.ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ

Материалы для оценки уровня освоения учебного материала дисциплины «Big Data Mining / Обработка больших данных» (оценочные материалы), включающие в себя перечень компетенций с указанием этапов их формирования, описание

показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания, типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы, методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций, разработаны в полном объеме и доступны для обучающихся на странице дисциплины в ТУИС РУДН.

Программа составлена в соответствии с требованиями ОС ВО РУДН.

РАЗРАБОТЧИКИ:

Доцент ДМПУ		Салтыкова О.А.
_____	_____	_____
Должность, БУП	Подпись	Фамилия И.О.
РУКОВОДИТЕЛЬ БУП:		
ДМПУ		Разумный Ю.Н.
_____	_____	_____
Наименование БУП	Подпись	Фамилия И.О.
РУКОВОДИТЕЛЬ ОП ВО:		
Доцент ДМПУ		Салтыкова О.А.
_____	_____	_____
Должность, БУП	Подпись	Фамилия И.О.