

Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Национальный исследовательский университет «МЭИ»

Направление подготовки/специальность: 27.03.04 Управление в технических системах

Наименование образовательной программы: Управление и информатика в технических системах

Уровень образования: высшее образование - бакалавриат

Форма обучения: Очная

Рабочая программа дисциплины
ПРИКЛАДНОЙ АНАЛИЗ ДАННЫХ

Блок:	Блок 1 «Дисциплины (модули)»
Часть образовательной программы:	Часть, формируемая участниками образовательных отношений
№ дисциплины по учебному плану:	Б4.Ч.01
Трудоемкость в зачетных единицах:	6 семестр - 2;
Часов (всего) по учебному плану:	72 часа
Лекции	не предусмотрено учебным планом
Практические занятия	6 семестр - 14 часов;
Лабораторные работы	не предусмотрено учебным планом
Консультации	проводится в рамках часов аудиторных занятий
Самостоятельная работа	6 семестр - 57,7 часа;
в том числе на КП/КР	не предусмотрено учебным планом
Иная контактная работа	проводится в рамках часов аудиторных занятий
включая:	
Лабораторная работа	
Промежуточная аттестация:	
Зачет	6 семестр - 0,3 часа;

Москва 2023

ПРОГРАММУ СОСТАВИЛ:

Преподаватель

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Толчеев В.О.
	Идентификатор	Rfbd680da-TolcheevVO-692f9924

В.О. Толчеев

СОГЛАСОВАНО:

Руководитель
образовательной программы

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Сидорова Е.Ю.
	Идентификатор	R0dee6ce9-SidorovaYY-923dc6a8

Е.Ю. Сидорова

Заведующий выпускающей
кафедрой

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Бобряков А.В.
	Идентификатор	R2c90f415-BobriakovAV-70dec1fa

А.В. Бобряков

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины: углубленное изучение разделов интеллектуального анализа данных (кластеризация, визуализация и классификация) и информационного поиска, получение навыков по решению конкретных прикладных задач

Задачи дисциплины

- освоение комплексной методики поиска, обработки и анализа разнотипных (фактографических и мультимедийных) данных;
- получение практических навыков самостоятельного решения задач по поиску, скачиванию, хранению, обработке и анализу данных.

Формируемые у обучающегося **компетенции** и запланированные **результаты обучения** по дисциплине, соотнесенные с **индикаторами достижения компетенций**:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Запланированные результаты обучения
ПК-1 Способен разрабатывать и применять технологии сбора, обработки и анализа разнотипных данных для расчета и проектирования систем и средств автоматизации и управления	ИД-1 _{ПК-1} Может формировать выборки и подготавливать данные для проведения анализа	знать: - основы информационного поиска и анализа научной информации, способы повышения релевантности и пертинентности, основные электронные ресурсы в области Data Science. уметь: - формировать выборки из электронных научных библиотек, коллекций фактографических и мультимедийных данных.
ПК-1 Способен разрабатывать и применять технологии сбора, обработки и анализа разнотипных данных для расчета и проектирования систем и средств автоматизации и управления	ИД-4 _{ПК-1} Использует стандартное программное обеспечение и специализированные библиотеки для обработки и анализа данных	знать: - методы обработки, визуализации и анализа разнотипных многомерных данных. уметь: - применять методы интеллектуального анализа, реализованные в стандартных библиотеках Python и общедоступных программных средствах, для решения прикладных задач.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ВО

Дисциплина относится к факультативным дисциплинам основной профессиональной образовательной программе Управление и информатика в технических системах (далее – ОПОП), направления подготовки 27.03.04 Управление в технических системах, уровень образования: высшее образование - бакалавриат.

Базируется на уровне среднего общего образования.

Результаты обучения, полученные при освоении дисциплины, необходимы при выполнении выпускной квалификационной работы.

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 2 зачетных единицы, 72 часа.

№ п/п	Разделы/темы дисциплины/формы промежуточной аттестации	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по видам учебной работы										Содержание самостоятельной работы/ методические указания	
				Контактная работа							СР				
				Лек	Лаб	Пр	Консультация		ИКР		ПА	Работа в семестре	Подготовка к аттестации /контроль		
КПР	ГК	ИККП	ТК												
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	Электронные информационные ресурсы в области науки и техники	40	6	-	-	8	-	-	-	-	-	32	-	<p><u>Подготовка к контрольной работе:</u> Подготовка к контрольной работе №1. «Электронные информационные ресурсы и основы информационного поиска». Для выполнения заданий контрольной работы необходимо изучить лекционный материал по темам 1 и 2.</p> <p><u>Изучение материалов литературных источников:</u> [1], стр. 23-39, 159-167, 167-171, 179-182, 189-198 [4], стр. 9-10, 162-164, 443-449</p>	
1.1	Информационные ресурсы со свободным доступом	20		-	-	4	-	-	-	-	-	16	-		
1.2	Информационный поиск и рекомендательные системы	20		-	-	4	-	-	-	-	-	16	-		
2	Методика поиска, обработки и анализа разнотипных данных	31.7		-	-	6	-	-	-	-	-	25.7	-		<p><u>Подготовка к контрольной работе:</u> Подготовка к выполнению контрольной работы №2: «Методы обработки, визуализации и анализа данных». Для выполнения заданий контрольной работы необходимо изучить лекционный материал по темам 1 и 2.</p> <p><u>Изучение материалов литературных источников:</u> [1], стр. 353-363, 379-393 [2], стр. 13-22, 24-25, 50-52 [3], стр. 41-59 [4], стр. 18-20, 110-112, 301-304, 435-436, 463-467</p>
2.1	Выбор подходящих моделей и методов анализа в зависимости от решаемой задачи	12		-	-	3	-	-	-	-	-	9	-		
2.2	Программные средства анализа данных	19.7		-	-	3	-	-	-	-	-	16.7	-		
	Зачет	0.3		-	-	-	-	-	-	-	0.3	-	-		

	Всего за семестр	72.0		-	-	14	-	-	-	-	0.3	57.7	-	
	Итого за семестр	72.0		-	-	14	-	-	-	-	0.3	57.7	-	

Примечание: Лек – лекции; Лаб – лабораторные работы; Пр – практические занятия; КПр – аудиторные консультации по курсовым проектам/работам; ИККП – индивидуальные консультации по курсовым проектам/работам; ГК- групповые консультации по разделам дисциплины; СР – самостоятельная работа студента; ИКР – иная контактная работа; ТК – текущий контроль; ПА – промежуточная аттестация

3.2 Краткое содержание разделов

1. Электронные информационные ресурсы в области науки и техники

1.1. Информационные ресурсы со свободным доступом

Формирование выборок из коллекций (фактографических и мультимедийных) данных. Информационная потребность пользователя. Полнотекстовые и библиографические документы. Цели и задачи наукометрического анализа. Наукометрические показатели.

1.2. Информационный поиск и рекомендательные системы

Информационный поиск и рекомендательные системы. Полнота и точность поиска, способы их повышения. Релевантность и пертинентность. Обратная связь по релевантности. Виды рекомендательных систем (РС). Области применения РС и примеры использования. РС в научно-образовательной области.

2. Методика поиска, обработки и анализа разнотипных данных

2.1. Выбор подходящих моделей и методов анализа в зависимости от решаемой задачи

Различия задач регрессии, кластеризации и классификации. Способы выявления дубликатов (почти дубликатов). Изучение принципов (статистического) моделирования данных в случае малых выборок и несбалансированных классов. Сэмплирование.

2.2. Программные средства анализа данных

Обзор современных программных средств анализа данных. Использование свободно распространяемых специализированных библиотек Python.

3.3. Темы практических занятий

1. Выбор моделей и методов для решения прикладных задач;
2. Изучение выборки с помощью методов кластеризации и визуализации;
3. Формирование выборок из коллекций данных;
4. Научные рекомендательные системы;
5. Виды рекомендательных систем;
6. Введение в информационный поиск;
7. Электронные информационные ресурсы в области науки и техники. Наукометрический анализ.

3.4. Темы лабораторных работ

не предусмотрено

3.5 Консультации

Групповые консультации по разделам дисциплины (ГК)

1. Консультации направлены на разбор наиболее сложных аспектов прикладного анализа данных и ответы на вопросы студентов.
2. Консультации направлены на разбор наиболее сложных аспектов прикладного анализа данных и ответы на вопросы студентов.

3.6 Тематика курсовых проектов/курсовых работ

Курсовой проект/ работа не предусмотрены

3.7. Соответствие разделов дисциплины и формируемых в них компетенций

Запланированные результаты обучения по дисциплине (в соответствии с разделом 1)	Коды индикаторов	Номер раздела дисциплины (в соответствии с п.3.1)		Оценочное средство (тип и наименование)
		1	2	
Знать:				
основы информационного поиска и анализа научной информации, способы повышения релевантности и пертинентности, основные электронные ресурсы в области Data Science	ИД-1ПК-1	+		Лабораторная работа/Проведение контрольной работы №2
методы обработки, визуализации и анализа разнотипных многомерных данных	ИД-4ПК-1		+	Лабораторная работа/Проведение контрольной работы №4
Уметь:				
формировать выборки из электронных научных библиотек, коллекций фактографических и мультимедийных данных	ИД-1ПК-1	+		Лабораторная работа/Проведение контрольной работы №1
применять методы интеллектуального анализа, реализованные в стандартных библиотеках Python и общедоступных программных средствах, для решения прикладных задач	ИД-4ПК-1	+	+	Лабораторная работа/Проведение контрольной работы №3

4. КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫЕ ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ДИСЦИПЛИНЕ)

4.1. Текущий контроль успеваемости

6 семестр

Форма реализации: Защита задания

1. Проведение контрольной работы №1 (Лабораторная работа)
2. Проведение контрольной работы №2 (Лабораторная работа)
3. Проведение контрольной работы №3 (Лабораторная работа)
4. Проведение контрольной работы №4 (Лабораторная работа)

Балльно-рейтинговая структура дисциплины является приложением А.

4.2 Промежуточная аттестация по дисциплине

Зачет (Семестр №6)

Положением о балльно-рейтинговой системе для студентов НИУ «МЭИ» на основании семестровой и аттестационной составляющих.

В диплом выставляется оценка за 6 семестр.

Примечание: Оценочные материалы по дисциплине приведены в фонде оценочных материалов ОПОП.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1 Печатные и электронные издания:

1. Маннинг, К. Д. Введение в информационный поиск : пер. с англ. / К. Д. Маннинг, П. Рагхаван, Х. Шютце . – М. : Вильямс, 2011 . – 528 с. - ISBN 978-5-8459-1623-5 .;
2. Толчеев, В. О. Современные методы обработки и анализа текстовой информации : учебное пособие по курсу "Интеллектуальные информационные системы" по специальности "Управление и информатика в технических системах" / В. О. Толчеев, Моск. энерг. ин-т (МЭИ ТУ) . – М. : Изд-во МЭИ, 2006 . – 76 с. - ISBN 5-7046-1285-7 .;
3. Толчеев, В. О. Основы теории классификации многомерных наблюдений : учебное пособие по курсу "Идентификация и диагностика систем" по специальности "Управление и информатика в технических системах " / В. О. Толчеев, Нац. исслед. ун-т "МЭИ" . – М. : Издательский дом МЭИ, 2012 . – 124 с. - ISBN 978-5-383-00690-0 .
<http://elibr.mpei.ru/elibr/view.php?id=3463>;
4. А. И. Орлов- "Прикладная статистика", Издательство: "Интернет-Университет Информационных Технологий (ИНТУИТ)", Москва, 2009 - (846 с.)
<https://biblioclub.ru/index.php?page=book&id=234537>.

5.2 Лицензионное и свободно распространяемое программное обеспечение:

1. Office / Российский пакет офисных программ;
2. Windows / Операционная система семейства Linux;
3. Python.

5.3 Интернет-ресурсы, включая профессиональные базы данных и информационно-справочные системы:

1. ЭБС Лань - <https://e.lanbook.com/>
2. ЭБС "Университетская библиотека онлайн" - http://biblioclub.ru/index.php?page=main_ub_red
3. Научная электронная библиотека - <https://elibrary.ru/>
4. База данных Web of Science - <http://webofscience.com/>
5. База данных Scopus - <http://www.scopus.com>
6. Национальная электронная библиотека - <https://rusneb.ru/>
7. ЭБС "Консультант студента" - <http://www.studentlibrary.ru/>
8. Электронная библиотека МЭИ (ЭБ МЭИ) - <http://elib.mpei.ru/login.php>
9. Портал открытых данных Российской Федерации - <https://data.gov.ru>
10. База открытых данных Министерства труда и социальной защиты РФ - <https://rosmintrud.ru/opendata>
11. База открытых данных профессиональных стандартов Министерства труда и социальной защиты РФ - <http://profstandart.rosmintrud.ru/obshchiy-informatsionnyy-blok/natsionalnyy-reestr-professionalnykh-standartov/>
12. База открытых данных Министерства экономического развития РФ - <http://www.economy.gov.ru>
13. База открытых данных Росфинмониторинга - <http://www.fedsfm.ru/opendata>
14. Электронная открытая база данных "Polpred.com Обзор СМИ" - <https://www.polpred.com>
15. Национальный портал онлайн обучения «Открытое образование» - <https://openedu.ru>
16. Официальный сайт Федерального агентства по техническому регулированию и метрологии - <http://protect.gost.ru/>
17. Открытая университетская информационная система «РОССИЯ» - <https://uisrussia.msu.ru>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Тип помещения	Номер аудитории, наименование	Оснащение
Учебные аудитории для проведения практических занятий, КР и КП	М-307, Учебная аудитория	стол преподавателя, стол учебный, стул, доска меловая, компьютерная сеть с выходом в Интернет, мультимедийный проектор, экран
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Учебные аудитории для проведения промежуточной аттестации	М-307, Учебная аудитория	стол преподавателя, стол учебный, стул, доска меловая, компьютерная сеть с выходом в Интернет, мультимедийный проектор, экран
	Ж-120, Машинный зал ИВЦ	сервер, кондиционер
Помещения для самостоятельной работы	НТБ-201, Компьютерный читальный зал	стол компьютерный, стул, стол письменный, вешалка для одежды, компьютерная сеть с выходом в Интернет, компьютер персональный, принтер, кондиционер
Помещения для консультирования	М-304а/2, Учебная лаборатория моделирования систем и анализа данных	кресло рабочее, стол преподавателя, стол учебный, стул, шкаф для документов, шкаф для одежды, компьютерная сеть с выходом в

		Интернет, компьютер персональный
Помещения для хранения оборудования и учебного инвентаря	М-309, Кладовая	стол, стул, шкаф для хранения инвентаря
	М-301/1, Кладовая	стул

БАЛЛЬНО-РЕЙТИНГОВАЯ СТРУКТУРА ДИСЦИПЛИНЫ

Прикладной анализ данных

(название дисциплины)

6 семестр**Перечень контрольных мероприятий текущего контроля успеваемости по дисциплине:**

КМ-1 Проведение контрольной работы №1 (Лабораторная работа)

КМ-2 Проведение контрольной работы №2 (Лабораторная работа)

КМ-3 Проведение контрольной работы №3 (Лабораторная работа)

КМ-4 Проведение контрольной работы №4 (Лабораторная работа)

Вид промежуточной аттестации – Зачет.

Номер раздела	Раздел дисциплины	Индекс КМ:	КМ-1	КМ-2	КМ-3	КМ-4
		Неделя КМ:	4	6	12	14
1	Электронные информационные ресурсы в области науки и техники					
1.1	Информационные ресурсы со свободным доступом		+	+	+	
1.2	Информационный поиск и рекомендательные системы		+	+	+	
2	Методика поиска, обработки и анализа разнотипных данных					
2.1	Выбор подходящих моделей и методов анализа в зависимости от решаемой задачи				+	+
2.2	Программные средства анализа данных				+	+
Вес КМ, %:			20	20	20	40