

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

Направление подготовки/специальность: 09.03.01 Информатика и вычислительная техника

Наименование образовательной программы: Технологии разработки программного обеспечения

Уровень образования: высшее образование - бакалавриат

Форма обучения: Заочная

Рабочая программа дисциплины
АНАЛИЗ ДАННЫХ

Блок:	Блок 1 «Дисциплины (модули)»
Часть образовательной программы:	Часть, формируемая участниками образовательных отношений
№ дисциплины по учебному плану:	Б1.Ч.01.12
Трудоемкость в зачетных единицах:	6 семестр - 3;
Часов (всего) по учебному плану:	108 часов
Лекции	6 семестр - 4 часа;
Практические занятия	6 семестр - 4 часа;
Лабораторные работы	не предусмотрено учебным планом
Консультации	6 семестр - 2 часа;
Самостоятельная работа	6 семестр - 96,8 часа;
в том числе на КП/КР	не предусмотрено учебным планом
Иная контактная работа	6 семестр - 0,9 часа;
включая: Тестирование Контрольная работа	
Промежуточная аттестация:	
Зачет с оценкой	6 семестр - 0,3 часа;

Москва 2023

ПРОГРАММУ СОСТАВИЛ:

Преподаватель

(должность)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Раскатова М.В.
	Идентификатор	R6bc62db2-RaskatovaMV-ead4381

(подпись)

М.В. Раскатова

(расшифровка подписи)

СОГЛАСОВАНО:

Руководитель
образовательной программы

(должность, ученая степень, ученое звание)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Вишняков С.В.
	Идентификатор	R35b26072-VishniakovSV-02810d9

(подпись)

С.В. Вишняков

(расшифровка подписи)

Заведующий выпускающей
кафедры

(должность, ученая степень, ученое звание)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Вишняков С.В.
	Идентификатор	R35b26072-VishniakovSV-02810d9

(подпись)

С.В. Вишняков

(расшифровка подписи)

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины: Формирование знаний об анализе данных, о базовых инструментах анализа данных

Задачи дисциплины

- Использование современных компьютерных технологий в качестве средства получения необходимых данных;
- Владение методами и алгоритмами обработки и анализа данных;
- Нарботка опыта практического применения методов анализа данных для решения практических задач реального сектора экономики;
- Представление результатов анализа 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-2 способен осуществлять интеграцию программных модулей и компонентов и проверки работоспособности выпусков программного продукта	ИД-2ПК-2 Подключение программного продукта к компонентам внешней среды	знать: - Основные понятия анализа данных. уметь: - Верно интерпретировать результаты анализа данных.
ПК-2 способен осуществлять интеграцию программных модулей и компонентов и проверки работоспособности выпусков программного продукта	ИД-4ПК-2 Внесение изменений в процедуры сборки модулей и компонентов компьютерного программного обеспечения, развертывания компьютерного программного обеспечения, миграции и преобразования данных	знать: - Основные методы анализа данных. уметь: - Идентифицировать тип исходных данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ВО

Дисциплина относится к основной профессиональной образовательной программе Технологии разработки программного обеспечения (далее – ОПОП), направления подготовки 09.03.01 Информатика и вычислительная техника, 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 3 зачетных единицы, 108 часов.

№ 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	Введение. Основные понятия	23.60	6	1.0	-	1.0	-	0.50	-	0.30	-	20.8	-	<p><u>Подготовка к текущему контролю:</u> Работа ориентирована на изучение основных понятий, используемых при работе с данными</p> <p><u>Самостоятельное изучение теоретического материала:</u> Работа ориентирована на изучение литературных источников, конспектирование материалов</p> <p><u>Подготовка к текущему контролю:</u> Работа ориентирована на изучение принципов работы Data Mining</p> <p><u>Самостоятельное изучение теоретического материала:</u> Работа ориентирована на изучение целей и задач Data Mining</p> <p><u>Подготовка к текущему контролю:</u> Работа ориентирована на выполнение анализа данных при помощи метода визуализации</p> <p><u>Самостоятельное изучение теоретического материала:</u> Работа ориентирована на изучение различных методов анализа данных</p>
1.1	Основные понятия. Задачи анализа данных	10.80		0.5	-	0.5	-	0.25	-	0.15	-	9.4	-	
1.2	Данные	12.80		0.5	-	0.5	-	0.25	-	0.15	-	11.4	-	
2	Data Mining	22.60		1.0	-	1.0	-	0.50	-	0.30	-	19.8	-	
2.1	Методы и стадии Data Mining.	10.80		0.5	-	0.5	-	0.25	-	0.15	-	9.4	-	
2.2	Задачи Data Mining. Классификация и кластеризация	11.80		0.5	-	0.5	-	0.25	-	0.15	-	10.4	-	
3	Методы анализа данных	43.80		2.0	-	2.0	-	1.00	-	0.30	-	38.5	-	
3.1	Метод деревьев решений	10.70		0.5	-	0.5	-	0.25	-	0.05	-	9.4	-	
3.2	Метод прогнозирования	10.70		0.5	-	0.5	-	0.25	-	0.05	-	9.4	-	
3.3	Метод кластерного анализа	10.75		0.5	-	0.5	-	0.25	-	0.1	-	9.4	-	
3.4	Метод визуализации	11.65	0.5	-	0.5	-	0.25	-	0.1	-	10.3	-		
	Зачет с оценкой	18.0	-	-	-	-	-	-	-	-	0.3	-	17.7	
	Всего за семестр	108.000	4.0	-	4.0	-	2.00	-	0.900	0.3	79.1	17.7		
	Итого за семестр	108.000	4.0	-	4.0	-	2.00	-	0.900	0.3	96.8			

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. Введение. Основные понятия

1.1. Основные понятия. Задачи анализа данных

Постановка задачи для получения конечного результата на основе исходных данных.

1.2. Данные

Данные должны быть представлены в форме, пригодной для хранения, передачи и обработки.

2. Data Mining

2.1. Методы и стадии Data Mining.

Особенность Data Mining - сочетание математического инструментария (формализованные методы) и последних достижений в сфере информационных технологий (методы неформального анализа).

2.2. Задачи Data Mining. Классификация и кластеризация

Задачи Data Mining включают в себя анализ предметной области, постановку задачи, подготовку данных, очистку данных..

3. Методы анализа данных

3.1. Метод деревьев решений

Способ представления правил в иерархической, последовательной структуре. Основа структуры - ответы "ДА" и "НЕТ" на ряд вопросов..

3.2. Метод прогнозирования

Метод сводится к поиску ассоциативных правил.

3.3. Метод кластерного анализа

Результаты кластеризации зависят от критериев разбиения совокупности исходных данных.

3.4. Метод визуализации

Визуализация - инструментарий, позволяющий увидеть конечный результат вычислений.

3.3. Темы практических занятий

1. Прогнозирование;
2. Функции предсказания и сглаживания;
3. Data Mining;
4. Построение графиков в трехмерном пространстве;
5. Визуализация;
6. Понятия анализа данных.

3.4. Темы лабораторных работ

не предусмотрено

3.5 Консультации

Групповые консультации по разделам дисциплины (ГК)

1. Рассмотрение основных понятий и терминов, применяемых при анализе данных.
Формулирование задач анализа данных
2. Рассмотрение особенностей Data Mining
3. Рассмотрение особенностей методов, используемых для анализа данных

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	3	
Знать:					
Основные понятия анализа данных	ИД-2ПК-2	+			Тестирование/Понятия анализа данных
Основные методы анализа данных	ИД-4ПК-2		+		Тестирование/Data Mining
Уметь:					
Верно интерпретировать результаты анализа данных	ИД-2ПК-2			+	Контрольная работа/Визуализация
Идентифицировать тип исходных данных	ИД-4ПК-2			+	Контрольная работа/Визуализация

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: Компьютерное задание

1. Понятия анализа данных (Тестирование)
2. Data Mining (Тестирование)

Форма реализации: Письменная работа

1. Визуализация (Контрольная работа)

Балльно-рейтинговая структура дисциплины является приложением А.

4.2 Промежуточная аттестация по дисциплине

Зачет с оценкой (Семестр №6)

Оценка определяется по совокупности результатов текущего контроля успеваемости в соответствии с Положением о балльно-рейтинговой системе для студентов НИУ «МЭИ»

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. А. И. Китайгородский- "Теория структурного анализа", Издательство: "Издательство Академии Наук СССР", Москва, 1957 - (286 с.)
<https://biblioclub.ru/index.php?page=book&id=475614>;
2. Александров, В. В. Анализ данных на ЭВМ (на примере системы СИТО) / В. В. Александров, А. И. Алексеев, Н. Д. Горский . – М. : Финансы и статистика, 1990 . – 191 с. – (Математическое обеспечение прикладной статистики) .;
3. Боровиков, В. П. STATISTICA: Искусство анализа данных на компьютере : Для профессионалов / В. П. Боровиков . – СПб. : Питер, 2001 . – 656 с. + Приложение CD-ROM . - ISBN 5-318-00302-8 ..

5.2 Лицензионное и свободно распространяемое программное обеспечение:

1. СДО "Прометей";
2. Office / Российский пакет офисных программ;
3. Windows / Операционная система семейства Linux;
4. Майнд Видеоконференции.

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. Электронная библиотека МЭИ (ЭБ МЭИ) - <http://elib.mpei.ru/login.php>
5. Портал открытых данных Российской Федерации - <https://data.gov.ru>
6. База открытых данных Министерства труда и социальной защиты РФ - <https://rosmintrud.ru/opendata>
7. База открытых данных профессиональных стандартов Министерства труда и социальной защиты РФ - <http://profstandart.rosmintrud.ru/obshchiy-informatsionnyy-blok/natsionalnyy-reestr-professionalnykh-standartov/>
8. База открытых данных Министерства экономического развития РФ - <http://www.economy.gov.ru>
9. База открытых данных Росфинмониторинга - <http://www.fedsfm.ru/opendata>
10. Электронная открытая база данных "Polpred.com Обзор СМИ" - <https://www.polpred.com>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Тип помещения	Номер аудитории, наименование	Оснащение
Учебные аудитории для проведения лекционных занятий и текущего контроля	Ж-417/6, Белая мультимедийная студия	стол компьютерный, доска интерактивная, компьютерная сеть с выходом в Интернет, мультимедийный проектор, компьютер персональный
	Ж-417/7, Световая черная студия	стул, компьютерная сеть с выходом в Интернет, микрофон, мультимедийный проектор, экран, оборудование специализированное, компьютер персональный
Учебные аудитории для проведения практических занятий, КР и КП	Ж-417/1, Компьютерный класс ИДДО	стол преподавателя, стол компьютерный, шкаф для документов, шкаф для одежды, стол письменный, компьютерная сеть с выходом в Интернет, доска маркерная передвижная, компьютер персональный, принтер, кондиционер, стенд информационный
Учебные аудитории для проведения промежуточной аттестации	Ж-417/1, Компьютерный класс ИДДО	стол преподавателя, стол компьютерный, шкаф для документов, шкаф для одежды, стол письменный, компьютерная сеть с выходом в Интернет, доска маркерная передвижная, компьютер персональный, принтер, кондиционер, стенд информационный
Помещения для самостоятельной работы	НТБ-303, Компьютерный читальный зал	стол компьютерный, стул, стол письменный, вешалка для одежды, компьютерная сеть с выходом в Интернет, компьютер персональный, принтер, кондиционер
Помещения для консультирования	Ж-2006, Конференц-зал ИДДО	стол, стул, компьютер персональный, кондиционер
Помещения для хранения оборудования и учебного инвентаря	Ж-417 /2а, Помещение для инвентаря	стеллаж для хранения инвентаря, экран, указка, архивные документы, дипломные и курсовые работы студентов, канцелярский принадлежности, спортивный инвентарь, хозяйственный инвентарь, запасные комплектующие для оборудования

БАЛЛЬНО-РЕЙТИНГОВАЯ СТРУКТУРА ДИСЦИПЛИНЫ

Анализ данных

(название дисциплины)

6 семестр

Перечень контрольных мероприятий текущего контроля успеваемости по дисциплине:

КМ-1 Понятия анализа данных (Тестирование)

КМ-2 Data Mining (Тестирование)

КМ-3 Визуализация (Контрольная работа)

Вид промежуточной аттестации – Зачет с оценкой.

Номер раздела	Раздел дисциплины	Индекс КМ:	КМ-1	КМ-2	КМ-3
		Неделя КМ:	3	9	12
1	Введение. Основные понятия				
1.1	Основные понятия. Задачи анализа данных		+		
1.2	Данные		+		
2	Data Mining				
2.1	Методы и стадии Data Mining.			+	
2.2	Задачи Data Mining. Классификация и кластеризация			+	
3	Методы анализа данных				
3.1	Метод деревьев решений				+
3.2	Метод прогнозирования				+
3.3	Метод кластерного анализа				+
3.4	Метод визуализации				+
Вес КМ, %:			30	35	35