

Федеральное государственное бюджетное образовательное учреждение высшего образования «Тамбовский государственный университет имени Г.Р. Державина»  
Институт математики, физики и информационных технологий  
Кафедра функционального анализа

УТВЕРЖДАЮ:  
Директор института



Н. Л. Королева  
«21» июня 2023 г.

## **РАБОЧАЯ ПРОГРАММА**

по дисциплине Б1.В.ДВ.04.1 Математические методы анализа данных

Направление подготовки/специальность: 01.04.01 - Математика

Профиль/направленность/специализация: Обработка больших данных и интеллектуальные системы поддержки принятия решений

Уровень высшего образования: магистратура

Квалификация: Магистр

год набора: 2023

**Автор программы:**

Кандидат физико-математических наук, Переславцева Оксана Николаевна

Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.04.01 - Математика (уровень магистратуры) (приказ Министерства образования и науки РФ от «10» января 2018 г. № 12).

Рабочая программа принята на заседании Кафедры функционального анализа «14» июня 2023 г. Протокол № 9

Рассмотрена и одобрена на заседании Ученого совета Института математики, физики и информационных технологий, Протокол от «21» июня 2023 г. № 3.

## СОДЕРЖАНИЕ

1. Цели и задачи дисциплины.....	4
2. Место дисциплины в структуре ОП Магистратуры.....	6
3. Объем и содержание дисциплины.....	6
4. Контроль знаний обучающихся и типовые оценочные средства.....	13
5. Методические указания для обучающихся по освоению дисциплины (модуля).....	19
6. Учебно-методическое и информационное обеспечение дисциплины.....	21
7. Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы.....	21

## 1. Цели и задачи дисциплины

### 1.1 Цель дисциплины – формирование компетенций:

ПК-4 Способен к применению методом математического и алгоритмического моделирования при решении теоретических и прикладных задач

ПК-7 Способен к использованию баз данных и информационных систем при реализации организационно-управленческих функций

### 1.2 Типы задач профессиональной деятельности, к которым готовятся обучающиеся в рамках освоения дисциплины:

- научно-исследовательский
- организационно-управленческий
- проектно-технологический

### 1.3 Дисциплина ориентирована на подготовку обучающихся к профессиональной деятельности в сфере: 40 Сквозные виды профессиональной деятельности в промышленности (в сфере научно-исследовательских и опытноконструкторских разработок)

### 1.4 В результате освоения дисциплины у обучающихся должны быть сформированы:

Обобщенные трудовые функции / трудовые функции / трудовые или профессиональные действия (при наличии профстандарта)	Код и наименование компетенции ФГОС ВО, необходимой для формирования трудового или профессионального действия	Индикаторы достижения компетенций
	ПК-4 Способен к применению методом математического и алгоритмического моделирования при решении теоретических и прикладных задач	Применяет в своей деятельности адекватные математические теории для создания новых и модификации существующих алгоритмов современной компьютерной математики
	ПК-7 Способен к использованию баз данных и информационных систем при реализации организационно-управленческих функций	Обладает знаниями в области реляционной алгебры и умеет применять ее методы на практике

### 1.5 Согласование междисциплинарных связей дисциплин, обеспечивающих освоение компетенций:

ПК-4 Способен к применению методом математического и алгоритмического моделирования при решении теоретических и прикладных задач

№ п/п	Наименование дисциплин, определяющих междисциплинарные связи	Форма обучения			
		Очная (семестр)			
		1	2	3	4
1	Базы данных и системы управления базами данных	+			
2	Избранные вопросы алгебры			+	

3	Научно-исследовательская работа				+
4	Научно-педагогическая практика			+	
5	Онлайн-курс "Анализ данных на практике"		+		
6	Онлайн-курс "Прикладной и статистический анализ"			+	
7	Онлайн-курс "Суперкомпьютеры и параллельная обработка данных"			+	
8	Онлайн-курс "Хранение и обработка данных"			+	
9	Основы вычислимости и теория сложности		+		
10	Программирование на языках высокого уровня		+		
11	Разработка информационных систем и программных продуктов на основе больших данных			+	
12	Теория вычислительных процессов и структур			+	
13	Технология разработки программного обеспечения		+		

ПК-7 Способен к использованию баз данных и информационных систем при реализации организационно-управленческих функций

№ п/п	Наименование дисциплин, определяющих междисциплинарные связи	Форма обучения			
		Очная (семестр)			
		1	2	3	4
1	Базы данных и системы управления базами данных	+			
2	Научно-исследовательская работа				+

3	Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы)		+		
4	Онлайн-курс "Анализ данных на практике"		+		
5	Онлайн-курс "Хранение и обработка данных"			+	
6	Основы вычислимости и теория сложности		+		
7	Параллельные и распределенные вычисления	+			
8	Преддипломная практика				+
9	Разработка информационных систем и программных продуктов на основе больших данных			+	

## 2. Место дисциплины в структуре ОП магистратуры:

Дисциплина «Математические методы анализа данных» относится к части, формируемой участниками образовательных отношений, учебного плана ОП по направлению подготовки 01.04.01 - Математика.

Дисциплина «Математические методы анализа данных» изучается в 3 семестре.

## 3. Объем и содержание дисциплины

3.1. Объем дисциплины: 2 з.е.

Очная: 2 з.е.

Вид учебной работы	Очная (всего часов)
<b>Общая трудоёмкость дисциплины</b>	<b>72</b>
Контактная работа	16
Лекции (Лекции)	8
Практические (Практ. раб.)	8
Самостоятельная работа (СР)	56
Зачет	-

## 3.2. Содержание курса:

№ темы	Название раздела/темы	Вид учебной работы, час.			Формы текущего контроля
		Лек ции	Пра кт. раб.	СР	
		О	О	О	

3 семестр					
1	Построение математических моделей. Принятие решений по моделям	2	2	20	Лабораторная работа; Контрольная работа
2	Дискретное преобразование Фурье. Интегральные прямые и обратные преобразования Фурье и Лапласа.	2	2	18	Лабораторная работа
3	Интеллектуальные методы анализа данных. Понятие кластерного анализа. Статистические методы Data Mining	4	4	18	Лабораторная работа; Контрольная работа

### **Тема 1. Построение математических моделей. Принятие решений по моделям (ПК-4)**

#### **Лекция.**

Математическая постановка задачи принятия решений. Математические модели принятия решений. Модель транспортной задачи. Вероятностно-статистические модели.

#### **Практическое занятие.**

### Лабораторное занятие.

#### Задачи.

1. Составить математическую модель транспортной задачи, исходные данные которой приведены в таблице:

	20	30	40
40	3	5	7
50	4	6	10

2. Необходимо найти решение .

26	30	17	10	16	4
30	37	26	9	23	6
13	4	32	3	1	10
3	1	5	14	24	10
7	7	7	7	2	

3. На трех базах A1, A2, A3 находится однородный груз в количестве  $a_1$ ,  $a_2$ ,  $a_3$  т. Этот груз необходимо развести пяти потребителям B1, B2, B3, B4, B5, потребности которых в данном грузе составляют  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$ ,  $b_5$  т. Соответственно. Стоимость перевозок пропорциональна расстоянию и количеству перевозимого груза. Матрица тарифов и значения  $a_j$ ,  $b_j$  приведены в таблице. Требуется спланировать перевозки так, чтобы их общая стоимость была минимальной. Стоимость доставки единицы груза из каждого пункта отправления в соответствующие пункты назначения задана матрицей тарифов

	1	2	3	4	5	Запасы
1	15	8	9	11	12	100
2	4	10	7	5	8	150
3	6	3	4	15	20	250
Потребности	100	40	140	60	160	

#### Задания для самостоятельной работы.

1. В пунктах А и В находятся соответственно 150 т. и 90 т. горючего. Пунктам 1, 2, 3 требуется соответственно 60, 70, 110 т. Горючего. Стоимость перевозки 1 т. Горючего из пункта А в пункты 1, 2, 3 равна 60, 10, 40 тыс. руб. за 1 т. соответственно, а из пункта В в пункты 1, 2, 3 – 120, 20, 80 тыс. руб. за 1 т. соответственно. Составьте план перевозок горючего, минимизирующий общую сумму транспортных расходов.

2. В угольном бассейне добывается уголь, который хранится на трех складах в количестве 120, 60, 100 ед. соответственно. Добытый уголь доставляется четырем энергетическим установкам в количестве 70, 90, 50, и 70 ед. Стоимость доставки 1 ед. угля из каждого склада соответствующим энергетическим установкам задана матрицей . Определить оптимальный план доставки угля энергетическим установкам, обеспечивающий суммарные минимальные затраты.

3. Три завода выпускают комбайны, которые отправляются потребителям. Первый завод поставляет 50 комбайнов, второй – 40 комбайнов, третий – 70 комбайнов. Каждому из потребителей требуется соответственно 30, 50, 40 и 40 комбайнов. Стоимость перевозки одной единицы техники от поставщика потребителю задана матрицей стоимостей . Составьте оптимальный план, обеспечивающий общую минимальную стоимость перевозки комбайнов.

4. На двух складах А и В находится по 90 т. горючего. Перевозка одной тонны горючего со склада А в пункты 1, 2, 3 соответственно стоит 1, 3 и 5 д.е., а перевозка одной тонны со склада В в те же пункты – соответственно 2, 5 и 4 д.е. В каждый пункт надо доставить по одинаковому количеству тонн горючего. Составить такой план перевозки горючего, при котором транспортные расходы будут наименьшими.



5. В резерве трех железнодорожных станций А, В, С находятся соответственно 60, 80, 100 вагонов. Составить оптимальный план перегона этих вагонов к четырем пунктам погрузки хлеба, если пункту 1 необходимо 40 вагонов, пункту 2 – 60 вагонов, пункту 3 – 80 вагонов и пункту 4 – 60 вагонов. Стоимости перегонов одного вагона со станции А в указанные пункты соответственно равны 1, 2, 3, 4 д.е., со станции В – 4, 3, 2 и 1 д.е., со станции С – 1, 2, 2, 1 д.е.

6. Завод имеет три цеха А, В, С и четыре склада. 1, 2, 3, и 4. Цех А производит 30 тыс. шт. изделий, цех В – 40 тыс. шт., цех С – 20 тыс. шт. Пропускная способность складов за то же время характеризуется следующими показателями: склад 1 – 20 тыс. шт., склад 2 – 30 тыс. шт., склад 3 – 30 тыс. шт., склад 4 – 10 тыс. шт. Стоимости перевозки 1 тыс. шт. изделий из цеха А в склады 1, 2, 3, 4 соответственно равны 2, 3, 2, 4 д.е., из цеха В – 3, 2, 5, 1 д.е., из цеха С – 4, 3, 2, 6 д.е. Составить такой план перевозки изделий, при котором расходы на перевозку 90 тыс. шт. изделий были бы минимальными.

7. На трех автобазах имеются автобусы в количестве 35, 45, 50 шт. соответственно для обслуживания четырех маршрутов. Для перевозки пассажиров каждому из маршрутов требуется автобусов в количестве 40, 25, 35 и 30 шт. соответственно. Расходы по эксплуатации каждой транспортной единицы заданы матрицей.

Распределить имеющиеся транспортные средства (автобусы) по маршрутам таким образом, чтобы общие расходы были минимальными.

8. Три завода выпускают грузовые автомобили, которые отправляются четырем потребителям. Первый завод поставляет 90 платформ грузовиков, второй – 30 платформ, третий – 40 платформ. Требуется поставить платформы следующим потребителям: первому – 70 шт., второму – 30 шт., третьему – 20 шт., четвертому – 40 шт. Стоимость перевозки одной платформы от поставщика до потребителя указана в следующей таблице (д.е.):

Поставщики

Потребители

I

II

III

IV

1

2

3

18

10

16

20

20

22

14

40

10

10

30

20

Составьте оптимальный план доставки грузовых автомобилей, обеспечивающий минимальные расходы.

9. На складах А, В, С находится сортовое зерно 100, 150, 250 т., которое нужно доставить в четыре пункта. Пункту 1 необходимо поставить 50 т., пункту 2 – 100 т., пункту 3 – 200 т., пункту 4 – 150 т. сортового зерна. Стоимость доставки 1 т. зерна со склада А в указанные пункты соответственно равна (д. е.) 80, 30, 50, 20; со склада В – 40, 10, 60, 70; со склада С – 10, 90, 40, 30. Составьте оптимальный план перевозки зерна из условия минимума стоимости перевозки.

10. Груз, находящийся на трех складах и требующий для перевозки 60, 80, 106 автомашин соответственно, необходимо перевезти в четыре магазина, Первому магазину требуется 44 машины груза, второму – 70, третьему – 50 и четвертому – 82 машины. Стоимость пробега одной автомашины за 1 км составляет 10 д.е. Расстояния от складов до магазинов указаны в таблице:

Склады

Машины

1

2

3

4

1

2

3

18

2

12

17

7

18

6

10

2

8

41

22

Составьте оптимальный по стоимости план перевозки груза от складов до магазинов.

## **Тема 2. Дискретное преобразование Фурье. Интегральные прямые и обратные преобразования Фурье и Лапласа. (ПК-4)**

### **Лекция.**

Дискретное преобразование Фурье. Алгоритм быстрого преобразования Фурье. Обратное преобразование Фурье. Интегральное преобразование Фурье.

### **Практическое занятие.**

**Лабораторное занятие. Применение дискретного преобразования для умножения многочленов, для умножения больших чисел.**

- 1 1. Написать программу, реализующую алгоритм быстрого преобразования Фурье.
- 2 2. Разработать алгоритм умножения многочленов с помощью быстрого преобразования Фурье.
- 3 3. Разработать алгоритм умножения больших чисел с помощью быстрого преобразования Фурье.

### **Задания для самостоятельной работы.**

Составить конспект лекции на тему «Интегральное прямое и обратное преобразование Лапласа».

## **Тема 3. Интеллектуальные методы анализа данных. Понятие кластерного анализа. Статистические методы Data Mining (ПК-7)**

### **Лекция.**

Понятие интеллектуального анализа данных (DM). Требования, предъявляемые к новым знаниям. Задачи DataMining. Стадии DataMining. Модели DataMining: классификация, кластеризация, поиск ассоциативных правил, итоговые модели, регрессионные модели и т. д. Сферы применения технологии интеллектуального анализа данных.

### **Практическое занятие.**

**Лабораторное занятие.**

Методика обнаружения нового знания в хранилищах данных (KDD). Характеристика этапов KDD. Природа задач таксономии. Алгоритмы таксономии класса FOREL. Динамическая таксономия. Таксономия с «суперцелью». Кластерный анализ. Задача кластерного анализа.

### **Задания для самостоятельной работы.**

Тема: «Программирование методов кластеризации данных»

В работе необходимо изучить методы кластеризации, предложить модификацию выбранного метода и получить навыки в создании приложений для решения практической задачи анализа данных на основе методов кластеризации.

Исходные данные: определены в задании к лабораторной работе.

Результаты должны быть представлены в виде текстового отчета и работающего ПО.

Отчет должен содержать

1. Титульный лист.
2. Цель, задание и требования.
3. Описание объектов исследования, входных данных (атрибутов объектов) и исследовательских вопросов.
4. Выбор метода кластеризации.
5. Формальную постановку задачи кластеризации, применительно к прикладной области, при этом требуется предложить улучшение выбранного метода.
6. Архитектуру разработанного ПО (IDEF0 (как есть/как должно быть), UML диаграммы) и используемые технологии.
7. Результаты кластеризации выбранным и модифицированным методами и их сравнительная оценка.
8. Заключение и выводы.
9. Список литературы и источников.

Требования к ПО.

Программа должна обеспечивать ввод исходных данных, выполнение кластеризации выбранным методом, кластеризации модифицированным методом, вывода результатов кластеризации, то есть полученных кластеров в табличной форме, строки – наименование кластеров, столбцы – математические характеристики. Для каждого кластера – мощность (количество объектов), центр, среднее внутрикластерное расстояние.

Содержательные требования к работе

- 1) Сформировать структуру данных для анализа объекта исследования и заполнить ее.
- 2) Разработать методику и программу для сегментации объекта исследования.
- 3) Провести сегментацию объекта исследования по количественным признакам на основе выбранного метода кластеризации внутри каждого кластера.
- 4) Провести пространственную (по странам, регионам, городам, предметным областям, корпорациям...) сегментацию объекта исследования.
- 5) Провести временную сегментацию на основе тенденций «рост», «падение», «стабильность».
- 6) Сформулировать выводы и объяснить результаты.

Варианты заданий:

1. Кластеризация рынка технологий Big Data;
2. Кластеризация языков программирования;
3. Кластеризация рынка IT профессий;
4. Кластеризация рынка рекламных технологий (ТВ, радио, интернет);
5. Кластеризация рынка IT продуктов;
6. Кластеризация рынка IT технологий разработки ПО;
7. Кластеризация рынка e-learning в области IT;
8. Кластеризация рынка распределенных систем;
9. Кластеризация рынка IoT;
10. Кластеризация регионов по индексам развития информационного общества;

11. Кластеризация средств визуального моделирования;
12. Кластеризация технических текстов при разработке ПО;
13. Кластеризация программных проектов по метрикам качества;
14. Кластеризация абитуриентов ИТ-направлений вузов;
15. Кластеризация покупателей по типу поведения;
16. Кластеризация ИТ-разработчиков по типу поведения;
17. Кластеризация пользователей ИТ-продуктов;
18. Кластеризация отзывов на ИТ-продукт.

#### 4. Контроль знаний обучающихся и типовые оценочные средства

##### 4.1. Распределение баллов:

##### 3 семестр

- посещаемость – 10 баллов
- текущий контроль – 70 баллов
- контрольные срезы – 2 среза по 10 баллов каждый
- премиальные баллы – 20 баллов

##### Распределение баллов по заданиям:

№ те мы	Название темы / вид учебной работы	Формы текущего контроля / срезы	Мах. кол-во баллов	Методика проведения занятия и оценки
1.	Построение математических моделей. Принятие решений по моделям	Лабораторная работа	25	17-18 баллов – работа выполнена полностью, метод выбран наиболее рационально, студент грамотно отвечает на все поставленные преподавателем вопросы; 11-16 баллов – работа выполнена полностью, метод выбран рационально, студент грамотно отвечает на все поставленные преподавателем вопросы, имеются недочеты в оформлении; 6-10 баллов – работа выполнена почти полностью, задача решена верно, но при этом использован не самый рациональный метод (или студент неуверенно отвечает на поставленные вопросы, ошибается); 1-5 баллов – студент может правильно решить задачу только с помощью наводящих вопросов преподавателя, но в процессе работы осваивает нужный учебный материал; 0 баллов – студент не может привести решение задачи и не может ответить на наводящие вопросы преподавателя и обнаруживает полную неподготовленность по изучаемой тематике.
		Контрольная работа(контрольный срез)	10	10 баллов, если студент правильно выполнил все задания 5 баллов, если студент правильно выполнил 51% заданий Менее 50% - 0 баллов.

2.	Дискретное преобразование Фурье. Интегральные прямые и обратные преобразования Фурье и Лапласа.	Лабораторная работа	25	11-16 баллов – работа выполнена полностью, метод выбран наиболее рационально, студент грамотно отвечает на все поставленные преподавателем вопросы; 7-10 баллов – работа выполнена полностью, метод выбран рационально, студент грамотно отвечает на все поставленные преподавателем вопросы, имеются недочеты в оформлении; 5-6 баллов – работа выполнена почти полностью, задача решена верно, но при этом использован не самый рациональный метод (или студент неуверенно отвечает на поставленные вопросы, ошибается); 0-4 баллов – студент может правильно решить задачу только с помощью наводящих вопросов преподавателя, но в процессе работы осваивает нужный учебный материал; 0 баллов – студент не может привести решение задачи и не может ответить на наводящие вопросы преподавателя и обнаруживает полную неподготовленность по изучаемой тематике.
3.	Интеллектуальные методы анализа данных. Понятие кластерного анализа. Статистические методы Data Mining	Лабораторная работа	20	11-16 баллов – работа выполнена полностью, метод выбран наиболее рационально, студент грамотно отвечает на все поставленные преподавателем вопросы; 7-10 баллов – работа выполнена полностью, метод выбран рационально, студент грамотно отвечает на все поставленные преподавателем вопросы, имеются недочеты в оформлении; 5-6 баллов – работа выполнена почти полностью, задача решена верно, но при этом использован не самый рациональный метод (или студент неуверенно отвечает на поставленные вопросы, ошибается); 0-4 баллов – студент может правильно решить задачу только с помощью наводящих вопросов преподавателя, но в процессе работы осваивает нужный учебный материал; 0 баллов – студент не может привести решение задачи и не может ответить на наводящие вопросы преподавателя и обнаруживает полную неподготовленность по изучаемой тематике.
		Контрольная работа(контрольный срез)	10	10 баллов, если студент правильно выполнил все задания 5 баллов, если студент правильно выполнил 51% заданий Менее 50% - 0 баллов.
4.	Посещаемость		10	10 баллов за отсутствие пропусков занятий без уважительной причины и выполнение в срок всех заданий
5.	Премиальные баллы		20	20 баллов за участие в студенческих научных конференциях и олимпиадах
6.	Индивидуальные задания, с помощью которых можно набрать дополнительные баллы		30	30 баллов за выполнение дополнительных индивидуальных заданий
7.	Итого за семестр		100	

Итоговая оценка по зачету выставляется в 100-балльной шкале и в традиционной четырехбалльной шкале. Перевод 100-балльной рейтинговой оценки по дисциплине в традиционную четырехбалльную осуществляется следующим образом:

100-балльная система	Традиционная система
50 - 100 баллов	Зачтено
0 - 49 баллов	Не зачтено

#### 4.2 Типовые оценочные средства текущего контроля

## Контрольная работа

Тема 1. Построение математических моделей. Принятие решений по моделям

Задания контрольной работы

Решить гладкие конечномерные задачи на безусловный и условный экстремум:

1.  $x^2 - xy + y^2 - 2x + y \rightarrow \text{extr};$
2.  $e^{xy} \rightarrow \text{extr}, x + y = 1;$
3.  $xyz \rightarrow \text{extr}, x^2 + y^2 + z^2 = 1, x + y + z = 0;$
4.  $xyz \rightarrow \text{extr}, x^2 + y^2 + z^2 \leq 1;$
5.  $e^{x_1 - x_2} - x_1 - x_2 \rightarrow \text{extr}, x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0.$

Тема 3. Интеллектуальные методы анализа данных. Понятие кластерного анализа. Статистические методы Data Mining

Задания контрольной работы

- 1 Понятие интеллектуального анализа данных (DM). Требования, предъявляемые к новым знаниям. Задачи DataMining. Стадии DataMining.
- 2 Модели DataMining: классификация, кластеризация, поиск ассоциативных правил, итоговые модели, регрессионные модели и т. д.
- 3 Сферы применения технологии интеллектуального анализа данных.
- 4 Методика обнаружения нового знания в хранилищах данных (KDD). Характеристика этапов KDD.
- 5 Природа задач таксономии. Алгоритмы таксономии класса FOREL. Динамическая таксономия. Таксономия с «суперцелью».
- 6 Кластерный анализ. Задача кластерного анализа.

## Лабораторная работа

Тема 1. Построение математических моделей. Принятие решений по моделям

### Лабораторная работа № 1.

1. Завод имеет три цеха А, В, С и четыре склада 1, 2, 3, и 4. Цех А производит 30 тыс. шт. изделий, цех В – 40 тыс. шт., цех С – 20 тыс. шт. Пропускная способность складов за то же время характеризуется следующими показателями: склад 1 – 20 тыс. шт., склад 2 – 30 тыс. шт., склад 3 – 30 тыс. шт., склад 4 – 10 тыс. шт. Стоимости перевозки 1 тыс. шт. изделий из цеха А в склады 1, 2, 3, 4 соответственно равны 2, 3, 2, 4 д.е., из цеха В – 3, 2, 5, 1 д.е., из цеха С – 4, 3, 2, 6 д.е. Составить такой план перевозки изделий, при котором расходы на перевозку 90 тыс. шт. изделий были бы минимальными.

2. На трех автобазах имеются автобусы в количестве 35, 45, 50 шт. соответственно для обслуживания четырех маршрутов. Для перевозки пассажиров каждому из маршрутов требуется автобусов в количестве 40, 25, 35 и 30 шт. соответственно. Расходы по эксплуатации

$$\begin{pmatrix} 10 & 8 & 12 & 7 \\ 9 & 8 & 11 & 12 \\ 5 & 7 & 10 & 9 \end{pmatrix}$$

каждой транспортной единицы заданы матрицей

Распределить имеющиеся транспортные средства (автобусы) по маршрутам таким образом, чтобы общие расходы были минимальными.

## Тема 2. Дискретное преобразование Фурье. Интегральное прямые и обратные преобразования Фурье и Лапласа.

### Лабораторная работа № 2.

1. Вычислить дискретное преобразование Фурье для многочлена  $f(x) = x^4 - 5x^3 + 3x^2 - 4x + 1$
2. С помощью программы, реализованной для алгоритма быстрого преобразования Фурье, вычислите произведение двух многочленов степени 50.
3. С помощью программы, реализованной для алгоритма быстрого преобразования Фурье, вычислите произведение двух чисел, составленных из 100 цифр.

## Тема 3. Интеллектуальные методы анализа данных. Понятие кластерного анализа. Статистические методы Data Mining

### Задания для лабораторной работы

#### Лабораторная работа № 3. «Программирование методов кластеризации данных»

В работе необходимо изучить методы кластеризации, предложить модификацию выбранного метода и получить навыки в создании приложений для решения практической задачи анализа данных на основе методов кластеризации.

Исходные данные: определены в задании к лабораторной работе.

Результаты должны быть представлены в виде текстового отчета и работающего ПО.

Отчет должен содержать

1. Титульный лист.
2. Цель, задание и требования.
3. Описание объектов исследования, входных данных (атрибутов объектов) и исследовательских вопросов.
4. Выбор метода кластеризации.
5. Формальную постановку задачи кластеризации, применительно к прикладной области, при этом требуется предложить улучшение выбранного метода.
6. Архитектуру разработанного ПО (IDEF0 (как есть/как должно быть), UML диаграммы) и используемые технологии.
7. Результаты кластеризации выбранным и модифицированным методами и их сравнительная оценка.
8. Заключение и выводы.
9. Список литературы и источников.

#### Требования к ПО.

Программа должна обеспечивать ввод исходных данных, выполнение кластеризации выбранным методом, кластеризации модифицированным методом, вывода результатов кластеризации, то есть полученных кластеров в табличной форме, строки – наименование кластеров, столбцы – математические характеристики. Для каждого кластера – мощность (количество объектов), центр, среднее внутрикластерное расстояние.

#### Содержательные требования к работе

- 1) Сформировать структуру данных для анализа объекта исследования и заполнить ее.
- 2) Разработать методику и программу для сегментации объекта исследования.
- 3) Провести сегментацию объекта исследования по количественным признакам на основе выбранного метода кластеризации внутри каждого кластера.
- 4) Провести пространственную (по странам, регионам, городам, предметным областям, корпорациям...) сегментацию объекта исследования.
- 5) Провести временную сегментацию на основе тенденций «рост», «падение», «стабильность».
- 6) Сформулировать выводы и объяснить результаты.

#### Варианты заданий:



1. Кластеризация рынка технологий Big Data;
2. Кластеризация языков программирования;
3. Кластеризация рынка IT профессий;
4. Кластеризация рынка рекламных технологий (ТВ, радио, интернет);
5. Кластеризация рынка IT продуктов;
6. Кластеризация рынка IT технологий разработки ПО;
7. Кластеризация рынка e-learning в области IT;
8. Кластеризация рынка распределенных систем;
9. Кластеризация рынка IoT;
10. Кластеризация регионов по индексам развития информационного общества;
11. Кластеризация средств визуального моделирования;
12. Кластеризация технических текстов при разработке ПО;
13. Кластеризация программных проектов по метрикам качества;
14. Кластеризация абитуриентов IT-направлений вузов;
15. Кластеризация покупателей по типу поведения;
16. Кластеризация IT-разработчиков по типу поведения;
17. Кластеризация пользователей IT-продуктов;
18. Кластеризация отзывов на IT-продукт.

4.3 Промежуточная аттестация по дисциплине проводится в форме зачета

#### **Типовые вопросы зачета (ПК-4, ПК-7)**

##### **Типовые вопросы экзамена**

1. Математическая постановка задачи принятия решений. Математические модели принятия решений.
2. Модель транспортной задачи
3. Вероятностно-статистические модели.
4. Дискретное преобразование Фурье.
5. Алгоритм быстрого преобразования Фурье.
6. Обратное преобразование Фурье.
7. Интегральное преобразование Фурье.
8. Применение дискретного преобразования для умножения многочленов.
9. Применение дискретного преобразования для умножения больших чисел.
10. Интегральное прямое и обратное преобразование Лапласа.
11. Понятие интеллектуального анализа данных (DM). Требования, предъявляемые к новым знаниям. Задачи DataMining. Стадии DataMining.
12. Модели DataMining: классификация, кластеризация, поиск ассоциативных правил, итоговые модели, регрессионные модели и т. д.
13. Сферы применения технологии интеллектуального анализа данных.
14. Методика обнаружения нового знания в хранилищах данных (KDD). Характеристика этапов KDD.
15. Природа задач таксономии. Алгоритмы таксономии класса FOREL. Динамическая таксономия. Таксономия с «суперцелью».
16. Кластерный анализ. Задача кластерного анализа.

#### **Типовые задания для зачета (ПК-4, ПК-7)**

##### **Типовые задания к лабораторным работам.**

Лабораторная работа № 1.

1. Завод имеет три цеха А, В, С и четыре склада. 1, 2, 3, и 4. Цех А производит 30 тыс. шт. изделий, цех В – 40 тыс. шт., цех С – 20 тыс. шт. Пропускная способность складов за то же время характеризуется следующими показателями: склад 1 – 20 тыс. шт., склад 2 – 30 тыс. шт., склад 3 – 30 тыс. шт., склад 4 – 10 тыс. шт. Стоимости перевозки 1 тыс. шт. изделий из цеха А в склады 1, 2, 3, 4 соответственно равны 2, 3, 2, 4 д.е., из цеха В – 3, 2, 5, 1 д.е., из цеха С – 4, 3, 2, 6 д.е. Составить такой план перевозки изделий, при котором расходы на перевозку 90 тыс. шт. изделий были бы минимальными.

2. На трех автобазах имеются автобусы в количестве 35, 45, 50 шт. соответственно для обслуживания четырех маршрутов. Для перевозки пассажиров каждому из маршрутов требуется автобусов в количестве 40, 25, 35 и 30 шт. соответственно. Расходы по эксплуатации каждой транспортной единицы заданы матрицей.

Распределить имеющиеся транспортные средства (автобусы) по маршрутам таким образом, чтобы общие расходы были минимальными.

#### Лабораторная работа № 2.

1. Вычислить дискретное преобразование Фурье для многочлена

$$f(x) = x^4 - 5x^3 + 3x^2 - 4x + 1.$$

2. С помощью программы, реализованной для алгоритма быстрого преобразования Фурье, вычислите произведение двух многочленов степени 50.

3. С помощью программы, реализованной для алгоритма быстрого преобразования Фурье, вычислите произведение двух чисел, составленных из 100 цифр.

#### Лабораторная работа № 3. «Программирование методов кластеризации данных»

В работе необходимо изучить методы кластеризации, предложить модификацию выбранного метода и получить навыки в создании приложений для решения практической задачи анализа данных на основе методов кластеризации.

Исходные данные: определены в задании к лабораторной работе.

Результаты должны быть представлены в виде текстового отчета и работающего ПО.

Отчет должен содержать

1. Титульный лист.
2. Цель, задание и требования.
3. Описание объектов исследования, входных данных (атрибутов объектов) и исследовательских вопросов.
4. Выбор метода кластеризации.
5. Формальную постановку задачи кластеризации, применительно к прикладной области, при этом требуется предложить улучшение выбранного метода.
6. Архитектуру разработанного ПО (IDEF0 (как есть/как должно быть), UML диаграммы) и используемые технологии.
7. Результаты кластеризации выбранным и модифицированным методами и их сравнительная оценка.
8. Заключение и выводы.
9. Список литературы и источников.

#### Требования к ПО.

Программа должна обеспечивать ввод исходных данных, выполнение кластеризации выбранным методом, кластеризации модифицированным методом, вывода результатов кластеризации, то есть полученных кластеров в табличной форме, строки – наименование кластеров, столбцы – математические характеристики. Для каждого кластера – мощность (количество объектов), центр, среднее внутрикластерное расстояние.

#### Содержательные требования к работе

- 1) Сформировать структуру данных для анализа объекта исследования и заполнить ее.
- 2) Разработать методику и программу для сегментации объекта исследования.

- 3) Провести сегментацию объекта исследования по количественным признакам на основе выбранного метода кластеризации внутри каждого кластера.
- 4) Провести пространственную (по странам, регионам, городам, предметным областям, корпорациям...) сегментацию объекта исследования.
- 5) Провести временную сегментацию на основе тенденций «рост», «падение», «стабильность».
- 6) Сформулировать выводы и объяснить результаты.

Варианты заданий:

1. Кластеризация рынка технологий Big Data;
2. Кластеризация языков программирования;
3. Кластеризация рынка IT профессий;
4. Кластеризация рынка рекламных технологий (ТВ, радио, интернет);
5. Кластеризация рынка IT продуктов;
6. Кластеризация рынка IT технологий разработки ПО;
7. Кластеризация рынка e-learning в области ИТ;
8. Кластеризация рынка распределенных систем;
9. Кластеризация рынка IoT;
10. Кластеризация регионов по индексам развития информационного общества;
11. Кластеризация средств визуального моделирования;
12. Кластеризация технических текстов при разработке ПО;
13. Кластеризация программных проектов по метрикам качества;
14. Кластеризация абитуриентов ИТ-направлений вузов;
15. Кластеризация покупателей по типу поведения;
16. Кластеризация ИТ-разработчиков по типу поведения;
17. Кластеризация пользователей ИТ-продуктов;
18. Кластеризация отзывов на ИТ-продукт.

#### 4.4. Шкала оценивания промежуточной аттестации

Оценка	Компетенции	Дескрипторы (уровни) – основные признаки освоения (показатели достижения результата)
«зачтено» (50 - 100 баллов)	ПК-4	Успешно применяет в своей деятельности адекватные математические теории для создания новых и модификации существующих алгоритмов современной компьютерн
	ПК-7	Обладает достаточными знаниями в области реляционной алгебры и умеет применять ее методы на практике
«не зачтено» (0 - 49 баллов)	ПК-4	Не умеет применять в своей деятельности адекватные математические теории для создания новых и модификации существующих алгоритмов современной компьютерн
	ПК-7	Не обладает достаточными знаниями в области реляционной алгебры и не умеет применять ее методы на практике

## 5. Методические указания для обучающихся по освоению дисциплины (модуля)

### 5.1 Методические указания по организации самостоятельной работы обучающихся:

Приступая к изучению дисциплины, в первую очередь обучающимся необходимо ознакомиться содержанием рабочей программы дисциплины (РПД), которая определяет содержание, объем, а также порядок изучения и преподавания учебной дисциплины, ее раздела, части.

Для самостоятельной работы важное значение имеют разделы «Объем и содержание дисциплины», «Учебно-методическое и информационное обеспечение дисциплины» и «Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы».

В разделе «Объем и содержание дисциплины» указываются все разделы и темы изучаемой дисциплины, а также виды занятий и планируемый объем в академических часах.

В разделе «Учебно-методическое и информационное обеспечение дисциплины» указана рекомендуемая основная и дополнительная литература.

В разделе «Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы» содержится перечень профессиональных баз данных и информационных справочных систем, необходимых для освоения дисциплины.

## 5.2 Рекомендации обучающимся по работе с теоретическими материалами по дисциплине

При изучении и проработке теоретического материала необходимо:

- просмотреть еще раз презентацию лекции в системе MOODLe, повторить законспектированный на лекционном занятии материал и дополнить его с учетом рекомендованной дополнительной литературы;
- при самостоятельном изучении теоретической темы сделать конспект, используя рекомендованные в РПД источники, профессиональные базы данных и информационные справочные системы;
- ответить на вопросы для самостоятельной работы, по теме представленные в пункте 3.2 РПД.
- при подготовке к текущему контролю использовать материалы фонда оценочных средств (ФОС).

## 5.3 Рекомендации по работе с научной и учебной литературой

Работа с основной и дополнительной литературой является главной формой самостоятельной работы и необходима при подготовке к устному опросу на семинарских занятиях, к дебатам, тестированию, экзамену. Она включает проработку лекционного материала и рекомендованных источников и литературы по тематике лекций.

Конспект лекции должен содержать реферативную запись основных вопросов лекции, в том числе с опорой на размещенные в системе MOODLe презентации, основных источников и литературы по темам, выводы по каждому вопросу. Конспект может быть выполнен в рамках распечатки выдачи презентаций лекций или в отдельной тетради по предмету. Он должен быть аккуратным, хорошо читаемым, не содержать не относящуюся к теме информацию или рисунки.

Конспекты научной литературы при самостоятельной подготовке к занятиям должны содержать ответы на каждый поставленный в теме вопрос, иметь ссылку на источник информации с обязательным указанием автора, названия и года издания используемой научной литературы. Конспект может быть опорным (содержать лишь основные ключевые позиции), но при этом позволяющим дать полный ответ по вопросу, может быть подробным. Объем конспекта определяется самим студентом.

В процессе работы с основной и дополнительной литературой студент может:

- делать записи по ходу чтения в виде простого или развернутого плана (создавать перечень основных вопросов, рассмотренных в источнике);
- составлять тезисы (цитирование наиболее важных мест статьи или монографии, короткое изложение основных мыслей автора);
- готовить аннотации (краткое обобщение основных вопросов работы);
- создавать конспекты (развернутые тезисы).

## 5.4. Рекомендации по подготовке к отдельным заданиям текущего контроля

Собеседование предполагает организацию беседы преподавателя со студентами по вопросам практического занятия с целью более обстоятельного выявления их знаний по определенному разделу, теме, проблеме и т.п. Все члены группы могут участвовать в обсуждении, добавлять информацию, дискутировать, задавать вопросы и т.д.

Устный опрос может применяться в различных формах: фронтальный, индивидуальный, комбинированный. Основные качества устного ответа подлежащего оценке:

- правильность ответа по содержанию;
- полнота и глубина ответа;
- сознательность ответа;
- логика изложения материала;
- рациональность использованных приемов и способов решения поставленной учебной задачи;
- своевременность и эффективность использования наглядных пособий и технических средств при ответе;

- использование дополнительного материала;
- рациональность использования времени, отведенного на задание.

Устный опрос может сопровождаться презентацией, которая подготавливается по одному из вопросов практического занятия. При выступлении с презентацией необходимо обращать внимание на такие моменты как:

- содержание презентации: актуальность темы, полнота ее раскрытия, смысловое содержание, соответствие заявленной темы содержанию, соответствие методическим требованиям (цели, ссылки на ресурсы, соответствие содержания и литературы), практическая направленность, соответствие содержания заявленной форме, адекватность использования технических средств учебным задачам, последовательность и логичность презентуемого материала;
- оформление презентации: объем (оптимальное количество), дизайн (читаемость, наличие и соответствие графики и анимации, звуковое оформление, структурирование информации, соответствие заявленным требованиям), оригинальность оформления, эстетика, использование возможности программной среды, соответствие стандартам оформления;
- личностные качества: ораторские способности, соблюдение регламента, эмоциональность, умение ответить на вопросы, систематизированные, глубокие и полные знания по всем разделам программы;
- содержание выступления: логичность изложения материала, раскрытие темы, доступность изложения, эффективность применения средств ИКТ, способы и условия достижения результативности и эффективности для выполнения задач своей профессиональной или учебной деятельности, доказательность принимаемых решений, умение аргументировать свои заключения, выводы.

## **6. Учебно-методическое и информационное обеспечение дисциплины**

### **6.1 Основная литература:**

1. Пальмов, С. В. Интеллектуальный анализ данных : учебное пособие. - Весь срок охраны авторского права; Интеллектуальный анализ данных. - Самара: Поволжский государственный университет телекоммуникаций и информатики, 2017. - 127 с. - Текст : электронный // IPR BOOKS [сайт]. - URL: <http://www.iprbookshop.ru/75376.html>

### **6.2 Дополнительная литература:**

1. Мыльников Л. А., Краузе Б., Кютц М., Баде К., Шмидт И. А. Интеллектуальный анализ данных в управлении производственными системами (подходы и методы) : монография. - Москва: Библио-Глобус, 2017. - 334 с. - Текст : электронный // ЭБС «Университетская библиотека онлайн» [сайт]. - URL: <http://biblioclub.ru/index.php?page=book&id=499006>
2. Нестеров, С. А. Интеллектуальный анализ данных средствами MS SQL Server 2008. - 2022-03-31; Интеллектуальный анализ данных средствами MS SQL Server 2008. - Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. - 303 с. - Текст : электронный // IPR BOOKS [сайт]. - URL: <http://www.iprbookshop.ru/62813.html>
3. Воронова, Л. И., Воронов, В. И. Big Data. Методы и средства анализа : учебное пособие. - 2022-04-04; Big Data. Методы и средства анализа. - Москва: Московский технический университет связи и информатики, 2016. - 33 с. - Текст : электронный // IPR BOOKS [сайт]. - URL: <http://www.iprbookshop.ru/61463.html>

### **6.3 Иные источники:**

1. СКА MahtPartner - <http://mathpar.cloud.unihub.ru/>

## **7. Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы**

Для проведения занятий по дисциплине необходимо следующее материально-техническое обеспечение: учебные аудитории для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, помещения для самостоятельной работы.

Учебные аудитории и помещения для самостоятельной работы укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Помещения для самостоятельной работы укомплектованы компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Для проведения занятий лекционного типа используются наборы демонстрационного оборудования, обеспечивающие тематические иллюстрации (проектор, ноутбук, экран/ интерактивная доска).

Лицензионное и свободно распространяемое программное обеспечение:

LibreOffice

Microsoft Office Профессиональный плюс 2007

Microsoft Windows 10

Операционная система "Альт Образование"

Профессиональные базы данных и информационные справочные системы:

1. Платформа Springer Link. – URL: <https://link.springer.com>
2. Научная электронная библиотека eLIBRARY.ru. – URL: <https://elibrary.ru>
3. Президентская библиотека имени Б.Н. Ельцина. – URL: <https://www.prilib.ru>
4. Российская государственная библиотека. – URL: <https://www.rsl.ru>
5. Российская национальная библиотека. – URL: <http://nlr.ru>
6. Тамбовская областная универсальная научная библиотека им. А.С. Пушкина. – URL: <http://www.tambovlib.ru>
7. Университетская библиотека онлайн: электронно-библиотечная система. – URL: <https://biblioclub.ru>
8. Федеральное хранилище «Единая коллекция цифровых образовательных ресурсов». – URL: <http://school-collection.edu.ru>
9. ЭБС «Университетская библиотека онлайн» . – URL: <http://www.biblioclub.ru>

### **Электронная информационно-образовательная среда**

[https://auth.tsutmb.ru/authorize?response\\_type=code&client\\_id=moodle&state=xyz](https://auth.tsutmb.ru/authorize?response_type=code&client_id=moodle&state=xyz)

Взаимодействие преподавателя и студента в процессе обучения осуществляется посредством мультимедийных, гипертекстовых, сетевых, телекоммуникационных технологий, используемых в электронной информационно-образовательной среде университета.