

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Сыров Игорь Анатольевич  
Должность: Директор  
Дата подписания: 30.10.2023 10:59:38  
Уникальный программный ключ:  
b683afe664d7e9f64175886cf9626a196149ad36

СТЕРЛИТАМАКСКИЙ ФИЛИАЛ  
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО  
УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ  
«УФИМСКИЙ УНИВЕРСИТЕТ НАУКИ И ТЕХНОЛОГИЙ»

Факультет  
Кафедра

*Математики и информационных технологий*  
*Математического моделирования*

**Аннотация рабочей программы дисциплины (модуля)**

дисциплина ***Б1.В.07 Методы интеллектуального анализа данных***

часть, формируемая участниками образовательных отношений

Направление

***02.03.03 Математическое обеспечение и администрирование информационных систем***

код наименование направления

Программа

***Сетевое программирование и администрирование информационных систем***

Форма обучения

***Очная***

Для поступивших на обучение в  
***2023 г.***

Стерлитамак 2023

**1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с установленными в образовательной программе индикаторами достижения компетенций**

Формируемая компетенция (с указанием кода)	Код и наименование индикатора достижения компетенции	Результаты обучения по дисциплине (модулю)
ПК-1. Способен использовать основные методы и средства автоматизации, связанные с разработкой, сопровождением и администрированием программных продуктов и информационных систем	ПК-1.1. Знание	Обучающийся должен знать: основные понятия, методы и технологии в области автоматизированной обработки, визуализации и хранения больших объемов данных
	ПК-1.2. Умение	Обучающийся должен уметь: применять методы обработки больших объемов данных и использовать программные средства с технологией Big Data при решении практических задач;
	ПК-1.3. Владение	Обучающийся должен владеть: понятиями и методами обработки больших объемов данных и машинного обучения с применением современных программных продуктов; навыками применения современных методов разработки технического, информационного и алгоритмического обеспечения Big Data.

**2. Цели и место дисциплины (модуля) в структуре образовательной программы**

Цели изучения дисциплины:

Цели изучения дисциплины «Методы интеллектуального анализа данных» формирование научного представления о современных методах интеллектуального анализа данных; приобретение знаний о принципах и алгоритмах, лежащих в основе современных интеллектуальных систем анализа данных; овладение технологиями и методиками сбора, предварительной подготовки и анализа экспериментальных данных; приобретение практических навыков работы с конкретными программными средствами интеллектуального анализа данных. Дисциплина рассматривает наиболее

распространенные методы и алгоритмы интеллектуального анализа экспериментальных данных.

Для достижения целей решаются следующие задачи изучения дисциплины: изучение основных методов интеллектуального анализа данных; изучение основных терминов в области интеллектуальных информационных технологий и анализа данных; изучение методик выбора алгоритмов и методов интеллектуального анализа данных при решении исследовательских задач; изучение эволюции методов интеллектуального анализа данных, сопоставление достоинств и недостатков различных методов; формирование представления о научных основах существующих методов анализа данных; изучение методики совершенствования систем информационно-документационного обеспечения управления с использованием интеллектуальных систем и алгоритмов; Практическая компонента умение осуществлять сбор и систематизацию экспериментальных данных в электронной форме; умение проводить предварительную подготовку данных для анализа; умение подобрать подходящие алгоритмы и методы интеллектуального анализа данных исходя из поставленной задачи и характеристик выборки данных; приобретение опыта решения практических задач в области документооборота с использованием конкретных программных средств.

Дисциплина изучается на 4 курсе в 7, 8 семестрах

**3. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся**

Общая трудоемкость (объем) дисциплины составляет 9 зач. ед., 324 акад. ч.

Объем дисциплины	Всего часов
	Очная форма обучения
Общая трудоемкость дисциплины	324
Учебных часов на контактную работу с преподавателем:	
лекций	28
практических (семинарских)	42
лабораторных	42
другие формы контактной работы (ФКР)	2,4
Учебных часов на контроль (включая часы подготовки):	69,6
экзамен	
Учебных часов на самостоятельную работу обучающихся (СР)	140

Формы контроля	Семестры
экзамен	7, 8

**4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий**

**4.1. Разделы дисциплины и трудоемкость по видам учебных занятий (в академических часах)**

№ п/п	Наименование раздела / темы дисциплины	Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость (в часах)
-------	--	---

		Контактная работа с преподавателем			СР
		Лек	Пр/Сем	Лаб	
<b>1</b>	<b>Основы интеллектуального анализа данных</b>	<b>10</b>	<b>8</b>	<b>8</b>	<b>12</b>
1.1	Интеллектуальный анализ данных: базовые понятия	2	0	0	4
1.2	Матрица объект-объект и признак-признаки	4	4	4	4
1.3	Основные задачи анализа и интерпретации данных .	4	4	4	4
<b>2</b>	<b>Классификация данных с использованием детерминированных моделей</b>	<b>16</b>	<b>8</b>	<b>2</b>	<b>16</b>
2.1	Линейные дискриминантные функции. Линейная разделимость.	4	2	2	4
2.2	Нелинейные дискриминантные функции.	4	2	0	4
2.3	Процедуры обучения с коррекцией ошибок	4	2	0	4
2.4	Перцептронный критерий.	4	2	0	4
<b>3</b>	<b>Классификация данных на основе статистических моделей</b>	<b>0</b>	<b>12</b>	<b>10</b>	<b>20</b>
3.1	Байесовская дискриминантная функция.	0	4	4	0
3.2	Дискриминантная функция Фишера.	0	6	2	0
3.3	Множественный дискриминантный анализ.	0	2	2	10
3.4	Обучение для статистических дискриминантных функций.	0	0	2	10
<b>4</b>	<b>Кластер-анализ</b>	<b>2</b>	<b>2</b>	<b>4</b>	<b>20</b>
4.1	Основные типы задач кластер-анализа.	0	0	2	10
4.2	Метод К- внутригрупповых средних.	2	2	2	10
<b>5</b>	<b>Методы снижения размерностей данных</b>	<b>0</b>	<b>6</b>	<b>10</b>	<b>36</b>
5.1	Метод главных компонент. Корреляционная матрица и ее основные свойства.	0	0	4	6
5.2	Модели факторного анализа.	0	2	2	10
5.3	Многомерное шкалирование.	0	2	2	10
5.4	Многомерные развертки	0	2	2	10
<b>6</b>	<b>Системы DATA MINING в задачах анализа и интерпретации данных</b>	<b>0</b>	<b>6</b>	<b>8</b>	<b>36</b>
6.1	DATA MINING - системы извлечения новых знаний из данных.	0	2	2	12
6.2	DATA MINING -предметно-статистические пакеты, нейронные сети.	0	2	4	12
6.3	Современные пакеты прикладных программ для решения задач обработки экспериментальных данных .	0	2	2	12
	<b>Итого</b>	<b>28</b>	<b>42</b>	<b>42</b>	<b>140</b>

## 4.2. Содержание дисциплины, структурированное по разделам (темам)

Курс лекционных занятий

№	Наименование раздела / темы дисциплины	Содержание
<b>1</b>	<b>Основы интеллектуального анализа данных</b>	
1.1	Интеллектуальный анализ данных: базовые понятия	1. Изучение приемов работы с системой STATISTICA. Матрица данных. Структура матрицы данных и задачи обработки.
1.2	Матрица объект-объект и признак-признаки	Расстояние и близость. Измерение признаков. Отношения и их представление. Основные проблемы измерений. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных .
1.3	Основные задачи анализа и интерпретации данных .	понимание исследуемой ситуации целиком (выявление тенденций, в том числе негативных отклонений от плана, прогнозирование и получение рекомендации).
<b>2</b>	<b>Классификация данных с использованием детерминированных моделей</b>	
2.1	Линейные дискриминантные функции. Линейная разделимость.	Решающие поверхности и дискриминантные функции. Линейные дискриминантные функции. Линейная разделимость. Кусочно-линейные дискриминантные функции.
2.2	Нелинейные дискриминантные функции.	Нелинейные дискриминантные функции. Фи- машины. Потенциальные функции как дискриминантные функции.
2.3	Процедуры обучения с коррекцией ошибок	Пространство весов. Процедуры обучения с коррекцией ошибок: правило с фиксированным приращением, правило абсолютной коррекции, частично корректирующее правило. Обобщенные градиентные методы.
2.4	Перцептронный критерий.	Перцептронный критерий. Процедуры обучения на основе минимальной среднеквадратичной ошибки: псевдоинверсный метод, метод Хо-Кашпа.
<b>4</b>	<b>Кластер-анализ</b>	
4.2	Метод К-внутригрупповых средних.	Обучение для статистических дискриминантных функций. Оценки максимального правдоподобия, байесовские оценки. Непараметрическое оценивание. Парзенковские окна, метод непараметрического оценивания на основе К-ближайшего соседства.

Курс практических/семинарских занятий

№	Наименование раздела / темы дисциплины	Содержание
<b>1</b>	<b>Основы интеллектуального анализа данных</b>	
1.2	Матрица объект-объект и признак-признаки	Примерные задачи о матрице объект-объект и признак-признак, о расстоянии и близость. Измерение признаков. Отношения и их представление. Основные проблемы измерений. Основные типы шкал. Проблема адекватности. Основные задачи анализа и интерпретации данных

1.3	Основные задачи анализа и интерпретации данных .	Обработка информации после ее получения и сбора; сбор информации, структуризация информации, выявление закономерностей, анализ, прогнозирование и получение рекомендаций.
<b>2</b>	<b>Классификация данных с использованием детерминированных моделей</b>	
2.1	Линейные дискриминантные функции. Линейная разделимость.	Решающие поверхности и дискриминантные функции. Линейные дискриминантные функции. Линейная разделимость. Кусочно-линейные дискриминантные функции.
2.2	Нелинейные дискриминантные функции.	Нелинейные дискриминантные функции. Фи-машины. Потенциальные функции как дискриминантные функции.
2.3	Процедуры обучения с коррекцией ошибок	Пространство весов. Процедуры обучения с коррекцией ошибок: правило с фиксированным приращением, правило абсолютной коррекции, частично корректирующее правило. Обобщенные градиентные методы.
2.4	Персептронный критерий.	Персептронный критерий. Процедуры обучения на основе минимальной среднеквадратичной ошибки: псевдоинверсный метод, метод Хо-Кашпа.
<b>3</b>	<b>Классификация данных на основе статистических моделей</b>	
3.1	Байесовская дискриминантная функция.	Функция потерь. Байесовская дискриминантная функция. Принятие решение по максимуму правдоподобия.
3.2	Дискриминантная функция Фишера.	Оптимальная дискриминантная функция для нормально распределенных образов. Дискриминантная функция Фишера.
3.3	Множественный дискриминантный анализ.	Множественный дискриминантный анализ. Пошаговый дискриминантный анализ. Ошибки классификации. Примеры построения статистических дискриминантных функций для различных статистических моделей данных.
<b>4</b>	<b>Кластер-анализ</b>	
4.2	Метод К- внутригрупповых средних.	Метод К- внутригрупповых средних. Использование методов теории графов в задачах кластеризации. Кластеризация на основе анализа плотностей вероятностей.
<b>5</b>	<b>Методы снижения размерностей данных</b>	
5.2	Модели факторного анализа.	Множественный дискриминантный анализ. Пошаговый дискриминантный анализ. Ошибки классификации. Примеры построения статистических дискриминантных функций для различных статистических моделей данных.
5.3	Многомерное шкалирование.	Модели факторного анализа. Оценка факторных нагрузок методом максимального правдоподобия и центроидным методом. Вращение факторов и их интерпретация. Использование кластеризации признаков для снижения размерности
5.4	Многомерные развертки	Многомерное шкалирование. Метрический и неметрический подход к многомерному шкалированию. Методы ортогонального проектирования. Нелинейные методы многомерного

		шкалирования. Многомерное шкалирование неметрических данных. Многомерные развертки.
<b>6</b>	<b>Системы DATA MINING в задачах анализа и интерпретации данных</b>	
6.1	DATA MINING - системы извлечения новых знаний из данных.	Предметно-ориентированные аналитические системы, статистические пакеты, нейронные сети, деревья решений, обнаружение логических закономерностей, генетические алгоритмы, системы визуализации многомерных данных.
6.2	DATA MINING -предметно-статистические пакеты, нейронные сети.	Тредметно-ориентированные аналитические системы, статистические пакеты, нейронные сети, деревья решений, обнаружение логических закономерностей, генетические алгоритмы, системы визуализации многомерных данных.
6.3	Современные пакеты прикладных программ для решения задач обработки экспериментальных данных .	Табличные процессоры и базы данных в задачах обработки данных. Виды статистических пакетов. Требования к статистическим пакетам общего назначения. Общая характеристика пакетов "STATGRAFICS Plus", "STATISTICA", SAS, SPSS . Комплексные системы класса DATA MINING для обработки данных - "PolyAnalist", Intelligent Miner

#### Курс лабораторных занятий

№	Наименование раздела / темы дисциплины	Содержание
<b>1</b>	<b>Основы интеллектуального анализа данных</b>	
1.2	Матрица объект-объект и признак-признаки	Работа-структура матрицы данных и задачи обработки. . Измерение признаков. Основные задачи анализа и интерпретации данных .
1.3	Основные задачи анализа и интерпретации данных .	Работа-Сбор, анализ и интерпретация данных.
<b>2</b>	<b>Классификация данных с использованием детерминированных моделей</b>	
2.1	Линейные дискриминантные функции. Линейная делимость.	Работа-Линейные функции, процедуры обучения с коррекцией ошибок.
<b>3</b>	<b>Классификация данных на основе статистических моделей</b>	
3.1	Байесовская дискриминантная функция.	Работа--решение по методу максимуму правдоподобия. Оценки максимального правдоподобия, байесовские оценки.
3.2	Дискриминантная функция Фишера.	Обучение для статистических дискриминантных функций.
3.3	Множественный дискриминантный анализ.	Проведение пошагового дискриминантного анализа.
3.4	Обучение для статистических дискриминантных функций.	Работа-- построения статистических дискриминантных функций для различных статистических моделей данных.
<b>4</b>	<b>Кластер-анализ</b>	
4.1	Основные типы задач кластер-анализа.	Кластерные методы, основанные на евклидовой метрике. Иерархическая кластеризация.
4.2	Метод К- внутригрупповых средних.	Использование методов теории графов в

		задачах кластеризации.
<b>5</b>	<b>Методы снижения размерностей данных</b>	
5.1	Метод главных компонент. Корреляционная матрица и ее основные свойства.	Проведение анализа матриц исходных данных. Геометрическая интерпретация главных компонент на плоскости.
5.2	Модели факторного анализа.	Оценка факторных нагрузок методом максимального правдоподобия и центроидным методом.
5.3	Многомерное шкалирование.	Проведение анализа матриц исходных данных. Метод главных компонент.
5.4	Многомерные развертки	Применение подходов к многомерному шкалированию. Снижения размерности.
<b>6</b>	<b>Системы DATA MINING в задачах анализа и интерпретации данных</b>	
6.1	DATA MINING - системы извлечения новых знаний из данных.	Знакомство с пакетом программ STATISTICA и подготовка статистических данных
6.2	DATA MINING -предметно-статистические пакеты, нейронные сети.	Изучение дискриминантного анализа с помощью пакета программ STATISTICA.
6.3	Современные пакеты прикладных программ для решения задач обработки экспериментальных данных	Изучение дискриминантного анализа с помощью пакета программ STATISTICA. Исследование методов кластер-анализа