

Учреждение образования
«Гродненский государственный университет имени Янки Купаль»

УТВЕРЖДАЮ

Ректор учреждения образования

«Гродненский государственный
университет имени Янки Купаль»

И.Ф. Китурко



2023 г.

Регистрационный № УД-2023 ФЭУ д/о-7 /уч.

Технологии интеллектуального анализа данных

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности**

7-06-0311-01

Экономика

Профилизация Экономика и организация производства продукции и услуг

Учебная программа составлена на основе образовательного стандарта высшего образования (ОСВО 7-06-0311-01-2023)

СОСТАВИТЕЛИ:

Н.В. Гринь, доцент, кандидат экономических наук

РЕЦЕНЗЕНТЫ:

А.В. Карканица, заведующий кафедрой современных технологий программирования, кандидат технических наук, доцент

А. А. Денисковец, доцент кафедры технической механики и математики ГрГАУ, кандидат физ.-мат. наук, доцент

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой экономики и управления на предприятии

(протокол № 08 от 29 мая 2023 г.);

Научно-методическим советом учреждения образования «Гродненский государственный университет имени Янки Купалы»

(протокол № 5 от 08 июня 2023 г.)

1. ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

1.1. Цели и задачи учебной дисциплины

Стремительное развитие информационных технологий, в частности, прогресс в методах сбора, хранения и обработки данных позволил многим организациям собирать огромные массивы данных, которые необходимо анализировать. Объемы этих данных настолько велики, что возможностей экспертов уже не хватает. На сегодняшний день интенсивно развивается направление, связанное с интеллектуализацией методов обработки и анализа данных. Интеллектуальные системы анализа данных (ИСАД) призваны минимизировать усилия лица, принимающего решения в процессе анализа данных, а также в настройке алгоритмов анализа. Многие ИСАД позволяют не только решать классические задачи принятия решения, но и способны выявлять причинно-следственные связи, скрытые закономерности в системе, подвергаемой анализу.

Интеллектуальный анализ данных (ИАД) – собирательное название, используемое для обозначения совокупности методов обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

Учебная программа по дисциплине «Технологии интеллектуального анализа данных» разработана для магистрантов факультета экономики и управления. Она предусматривает требования к содержанию тем лекционных и лабораторных занятий.

Целью учебной дисциплины является формирование представления о типах задач, возникающих в области интеллектуального анализа данных и методах их решения, которые помогут обучающимся успешно осуществлять анализ данных для решения экономических, управленческих, научно-исследовательских задач.

Задачи дисциплины – ознакомление с основными понятиями и типами закономерностей, выявляемыми с помощью ИАД, изучение методов и технологий ИАД, обучение использованию программных средств на основе технологий интеллектуального анализа данных для решения экономических, управленческих, научно-исследовательских задач.

Целями освоения дисциплины "Интеллектуальные информационные системы" является формирование у студентов комплекса теоретических знаний и методологических основ в области проведения анализа данных с использованием интеллектуальных систем обработки информации, а также практических навыков, необходимых для внедрения и практического использования таких систем.

Задачи дисциплины:

1. Получить представление о современных подходах и процедурах, выполняемых с использованием интеллектуальных систем.
2. Познакомиться с основными методами интеллектуального анализа данных.

3. Изучить основные технологии построения алгоритмов анализа данных с использованием информационных систем на основе технологий прикладного искусственного интеллекта.

4. Выработать навыки использования современных информационных технологий и программных средств, реализующих алгоритмы интеллектуального анализа данных.

1.2. Место учебной дисциплины в системе подготовки специалиста, связи с другими дисциплинами учебного плана

Учебная программа составлена с учетом межпредметных связей и программ по учебным дисциплинам «Высшая математика», «Компьютерные информационные технологии», «Эконометрика и экономико-математические методы и модели».

1.3. Требования к компетентности (согласно образовательному стандарту специальности)

Освоение учебной дисциплины «Интеллектуальный анализ данных» будет способствовать формированию у магистрантов следующих компетенций:

УПК-5. Быть способным осуществлять анализ данных для решения экономических, управленческих, научно-исследовательских задач.

В результате изучения учебной дисциплины студент должен:

– знать:

- базовые понятия и принципы ИАД;
- основные алгоритмы методов и моделей ИАД и подходы к их применению;
- этапы процесса и технологии ИАД, сферы применения и стандарты ИАД.

– уметь:

- производить расчеты с применением аппарата ИАД;
- применять методы ИАД в среде табличного процессора "MsExcel" для решения экономических, управленческих, научно-исследовательских задач.;
- творчески и эффективно использовать полученные знания в профессиональной деятельности.

– владеть навыками:

- применения алгоритмов и программных систем анализа данных;
- работы со средствами автоматизации интеллектуального анализа и обработки данных.

1.4. Форма получения высшего образования дневная - 2.0 г.

1.5. Распределение общих и аудиторных часов по семестрам, видам занятий

Семестр	Общее количество часов	Количество аудиторных часов (всего)	Количество часов лекций	Количество часов практических, семинарских занятий	Количество часов лабораторных занятий	Самостоятельная работа (в том числе внеаудиторная СР, КСР)	Формы текущей аттестации
3	100	42	18	24		66	зачет

2. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Номер	Название раздела	Содержание в соответствии с типовой учебной программой (стандартом)
1	Введение в интеллектуальный анализ данных	Основные понятия интеллектуального анализа данных. Интеллектуальный анализ данных (ИАД). Постановка и классификация задач ИАД. Данные. Понятие данных. Типы шкал наборов данных. Типы наборов данных. Задача классификации и регрессии. Задача поиска ассоциативных правил. Задача кластеризации. Примеры практического применения дисциплины ИАД: медицина, биоинформатика, геология, банковское дело, торговля, другие области применения.
2	Процесс интеллектуального анализа данных	Основные этапы процесса интеллектуального анализа данных. Анализ предметной области. Постановка задачи. Сбор и подготовка данных. Этапы процесса интеллектуального анализа данных, связанные с построением, проверкой, оценкой, выбором и коррекцией моделей. Анализ данных. Интерпретация и прогнозирование.
3	Статистические методы	Методы предварительного анализа данных. Описательная статистика. Характеристики центральной тенденции. Характеристики вариации. Очистка данных. Нормировка и стандартизация данных. Анализ выбросов и аномальных значений. Обработка пропущенных значений. Корреляционный анализ. Коэффициент корреляции Пирсона. Ранговая корреляция. Критерий Спирмена. Множественная и частная корреляции. Регрессионный анализ. Дисперсионный анализ.
4	Power Query	Основные понятия и функции Power Query. Процесс подключения к источникам данных и загрузки данных в Power Query. Основные методы трансформации данных, включая фильтрацию, сортировку, объединение таблиц и другие операции. Создание запросов в Power Query. Возможности Power Query для очистки и преобразования данных.
5	Визуализация как этап анализа данных	Понятие визуализации данных, базовые принципы. Исследовательская и презентационная задачи визуализации. Графики и диаграммы. Инфографика и схемы. Визуализация данных в программе MsExcel.

6	Бизнес аналитика и информационные панели (дашборды)	Бизнес аналитика и информационные панели (дашборды). Построение сводных таблиц и диаграмм для вычисления, сведения и анализа данных в программе MsExcel. Построение дашборда в программе MsExcel.
7	Метод анализа иерархий	Классификация задач. Участники процесса принятия решений. Альтернативы. Критерии. Типы задач принятия решений. Предварительная структуризация задач принятия решений. Преимущества и недостатки метода. Иерархическое представление. Шкала отношений. Матрицы парных сравнений. Собственные векторы и значения матриц.
8	Методы классификации	Деревья решений. Анализ данных с использованием деревьев решений. Анализ признаков объектов. Критерии и функции качества разбиения. Индекс Джини. Энтропия. Ошибка классификации. Остановка построения дерева. Сокращение дерева. Извлечение правил из деревьев. Алгоритмы построения деревьев решений.
9	Кластерный анализ	Основы кластерного анализа. Математические характеристики кластера. Критерии качества кластеризации. Выделение значимых кластеров. Иерархические агломеративные и дивизимные методы иерархического кластерного анализа. Неиерархические итерационные методы кластерного анализа. Алгоритм k-средних.

3. ТРЕБОВАНИЯ К КУРСОВОЙ РАБОТЕ (ПРОЕКТУ)

3.1. Цель курсовой работы (проекта) по дисциплине

Курсовая работа (проект) не предусмотрена учебным планом.

3.2. Объем задания

4. УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

№	Семес тр	Название раздела, темы	Количество аудиторных часов						Сам. раб. студе нта	Средст ва обучени я	Литерат ура	Формы контро ля знаний
			Лекции		Практ заняти я		Лаб. заняти я					
			Ауд .	УС Р	Ауд .	УС Р	Ауд .	УС Р				
1.		Введение в интеллектуальный анализ данных	2					2				
1.1.	3	Интеллектуальный анализ данных (ИАД). Основные понятия ИАД. Постановка и классификация задач ИАД. Данные. Понятие данных. Типы шкал наборов данных. Типы наборов данных. Основные задачи и области применения ИАД.	2					2	Доска. Проектор	1, 2, 4, 7, 8, 9, 10, 11	Опрос	
2.		Процесс интеллектуального анализа данных	2					2				
2.1.	3	Основные этапы процесса интеллектуального анализа данных. Анализ предметной области. Постановка задачи. Сбор и подготовка данных. Этапы процесса интеллектуального анализа данных, связанные с построением, проверкой, оценкой, выбором и коррекцией моделей. Анализ данных. Интерпретация и прогнозирование.	2					2	Доска. Проектор	1, 2, 4, 7, 8, 9, 10, 11	Опрос	

3.		Статистические методы	6				8		16			
3.1.	3	Методы предварительного анализа данных. Описательная статистика. Характеристики центральной тенденции. Характеристики вариации. Очистка данных. Нормировка и стандартизация данных.		2					2	Образовательный портал	1, 2, 4, 7, 8, 9, 10, 11	Опрос
3.2.	3	Анализ выбросов и аномальных значений. Анализ пропущенных значений.	2						2	Доска. Проектор	1, 2, 4, 7, 8, 9, 10, 11	Опрос
3.3	3	Корреляционный анализ. Коэффициент корреляции Пирсона. Ранговая корреляция. Критерий Спирмена. Множественная и частная корреляции. Регрессионный анализ. Дисперсионный анализ.		2					2	Образовательный портал	1, 2, 4, 7, 8, 9, 10, 11	Опрос
3.4	3	Методы предварительного анализа данных. Описательная статистика. Характеристики центральной тенденции. Характеристики вариации.					2		2	Учебная лаборатория	1, 2, 4, 7, 8, 9, 10, 11	Индивидуальное задание
3.5	3	Корреляционный анализ. Коэффициент корреляции Пирсона. Ранговая корреляция. Критерий Спирмена. Множественная и частная корреляции. Регрессионный анализ. Дисперсионный анализ.					2		2	Учебная лаборатория	1, 2, 4, 7, 8, 9, 10, 11	Индивидуальное задание
3.6	3	Анализ выбросов и аномальных значений.					2		3	Учебная лаборатория	4, 7, 8, 9	Индивидуальное задание
3.7	3	Анализ пропущенных значений.					2		3	Учебная	4, 7, 8,	Индивидуал

									лаборатория	9	ьное задание
4	3	Power Query	2				4	12			
4.1	3	Основные понятия и функции Power Query. Процесс подключения к источникам данных и загрузки данных в Power Query. Основные методы трансформации данных, включая фильтрацию, сортировку, объединение таблиц и другие операции. Создание запросов в Power Query. Возможности Power Query для очистки и преобразования данных.	2					4	Доска. Проектор	1, 2, 4, 7, 8, 9, 10, 11	Опрос
4.2	3	Основные понятия и функции Power Query. Процесс подключения к источникам данных и загрузки данных в Power Query. Основные методы трансформации данных, включая фильтрацию, сортировку, объединение таблиц и другие операции.					2	4	Учебная лаборатория	1, 2, 4, 7, 8, 9, 10, 11	Индивидуал ьное задание
4.3	3	Создание запросов в Power Query. Возможности Power Query для очистки и преобразования данных.					2	4	Образователь ный портал	1, 2, 4, 7, 8, 9, 10, 11	Тест
5.	3	Визуализация как этап анализа данных	2				2	8			
5.1.	3	Понятие визуализации данных, базовые принципы. Исследовательская и презентационная задачи визуализации. Графики и диаграммы. Инфографика и схемы.	2					4	Доска. Проектор	5, 6, 12, 13, 14, 15, 16	Опрос

		Визуализация данных в программе MsExcel.										
5.2.	3	Понятие визуализации данных, базовые принципы. Исследовательская и презентационная задачи визуализации. Графики и диаграммы. Инфографика и схемы. Визуализация данных в программе MsExcel.				2		4	Учебная лаборатория	5,6, 12, 13, 14, 15, 16	Индивидуальное задание	
6.	3	Бизнес аналитика и информационные панели (дашборды)				4		8				
5.1.	3	Бизнес аналитика и информационные панели (дашборды). Построение сводных таблиц и диаграмм для вычисления, сведения и анализа данных в программе MsExcel.				2		4	Учебная лаборатория	5,6, 12, 13, 14, 15, 16	Индивидуальное задание	
5.2.	3	Построение дашборда в программе MsExcel.				2		4	Учебная лаборатория	5,6, 12, 13, 14, 15, 16	Индивидуальное задание	
7.	3	Метод анализа иерархий	2			4		4				
7.1.	3	Классификация задач. Участники процесса принятия решений. Альтернативы. Критерии. Типы задач принятия решений. Предварительная структуризация задач принятия решений. Преимущества и недостатки метода. Иерархическое представление. Шкала отношений. Матрицы парных	2						Доска. Проектор	1, 2, 4, 7, 8, 9, 10, 11	Опрос	

		сравнений. Собственные векторы и значения матриц.										
7.2.	3	Классификация задач. Участники процесса принятия решений. Альтернативы. Критерии. Типы задач принятия решений. Предварительная структуризация задач принятия решений. Преимущества и недостатки метода. Иерархическое представление. Шкала отношений. Матрицы парных сравнений. Собственные векторы и значения матриц.					4		4	Учебная лаборатория	1, 2, 4, 7, 8, 9, 10, 11	Индивидуальное задание
8.	3	Методы классификации					2		4			
8.1.	3	Деревья решений. Анализ данных с использованием деревьев решений. Анализ признаков объектов. Критерии и функции качества разбиения. Индекс Джини. Энтропия. Ошибка классификации. Остановка построения дерева. Сокращение дерева. Извлечение правил из деревьев. Алгоритмы построения деревьев решений.					2		4	Образовательный портал	1, 2, 4, 7, 8, 9, 10, 11	Тест
9.	3	Кластерный анализ					2		4			
9.1.	3	Неиерархические итерационные методы кластерного анализа. Алгоритм k-средних.					2		4	Учебная лаборатория	1, 2, 4, 7, 8, 9, 10, 11	Индивидуальное задание
		ИТОГО:	14	4			20	4				
			18				24					

5. ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

5.1. Перечень рекомендуемой литературы

Основная литература:

1. Рафалович, Владимир Data Mining, или интеллектуальный анализ данных для занятых : практический курс / В. Рафалович. – М. : СмартБук, 2014. – 96 с.
2. Анализ данных : учебник для академического бакалавриата : учебник для студ. вузов, обуч. по экономическим напр. и спец. / Высшая школа экономики. Национальный исслед. ун-т ; под ред. В. С. Мхитаряна. – М. : Юрайт, 2017. – 490 с. – (Бакалавр. Академический курс).
3. Хакимзянов, Г. С Анализ данных и подготовка публикаций: программный инструментарий математика. Ч.3: учеб. пособие /Новосибирский гос. ун-т / Г. С. Хакимзянов, Л. Б. Чубаров. – Новосибирск : Изд-во Новосибирского ун-та, 1999. – 63 с.
4. Гринчук, А. Интеллектуальный анализ данных в телекоммуникациях: пример анализа лояльности клиентов / А. Гринчук // It Бел. – 2011. – № 1-2. – С. 34-36.
5. Кабаков, Роберт И. R в действии. Анализ и визуализация данных в программе R / Р. И. Кабаков ; пер. с англ. П. А. Волковой. – М. : ДМК Пресс, 2016. – 587 с.
6. Мастицкий, С. Э. Статистический анализ и визуализация данных с помощью R / С. Э. Мастицкий, В. К. Шитиков. – М. : ДМК Пресс, 2015. – 495 с.

Дополнительная литература:

7. Тирни, Б Наука о данных. Базовый курс / Брендан Тирни – Альпина Паблишер., 2021. – 222 с.
8. Дж., Морроу Как вытащить из данных максимум. Навыки аналитики для неспециалистов / Морроу Джордан – Альпина Паблишер., 2021. – 256 с.
9. Дюк, В. Data Mining: учебный курс / В. Дюк, А. Самойленко. –СПб., 2001. 368 с.
10. Яцков, Н. Н. Интеллектуальный анализ данных : пособие / Н. Н. Яцков. – Минск : БГУ, 2014. – 151 с.
11. Анализ данных и процессов: учеб. пособие / А. А. Барсегян [и др.]. –3-е изд., перераб. и доп. – СПб., 2009. – 512 с.
12. Каир, А Функциональное искусство: введение в информационную графику и визуализацию / Альберто Каир — Издательство: New Riders, 2012. — 384 с.
13. Тафти, Э Визуальные пояснения: образы и количества, доказательства и описание / Эдвард Тафти —Издательство:Graphics Press USA, 1997. — 151 с.
14. Нейтан Яу. Искусство визуализации в бизнесе. –М.: Манн, Иванов и Фербер, 2013. —352 с.

15. Желязны, Д. Говори на языке диаграмм : пособие по визуальным коммуникациям / Джин Желязны ; пер. с англ. [А. Мучника и Ю.Корнилович] — 5-е изд. — М. : Манн, Иванов и Фербер, 2012. — 304 с.
16. Фью, С. Дизайн информационной панели: эффективная визуальная передача данных / Фью Стивен — Издательство: O'Reilly Media, 2006. — 224 с.

5.2. Перечни заданий и контрольных мероприятий управляемой самостоятельной работы студентов по учебной дисциплине

Материалы для самостоятельной работы обучающихся и методические рекомендации по её выполнению расположены на Образовательном портале ГрГУ в личном кабинете магистранта <https://edu.grsu.by/>

5.2.1 Варианты тестовых заданий для проверки знаний магистрантов на лабораторных занятиях

1. Что включает в себя ИАД?

- а) Статистический анализ данных
- б) Машинное обучение
- в) Обработку естественного языка
- г) Все вышеперечисленное

2. Какие методы входят в ИАД?

- а) Кластеризация
- б) Классификация
- в) Регрессия
- г) Все вышеперечисленное

3. Что такое данные?

- а) Факты, цифры, статистика, тексты и т.д.
- б) Только числовые значения
- в) Только текстовые значения

4. Какие типы шкал наборов данных не существуют?

- а) Номинальная
- б) Порядковая
- в) Интервальная
- г) Ординальная
- д) Кумулятивная

5. Какую задачу решает классификация?

- а) Распределение объектов по заранее заданным классам
- б) Поиск взаимосвязей между признаками
- в) Поиск аномалий в данных

6. Какую задачу решает регрессия?
- а) Распределение объектов по заранее заданным классам
 - б) Поиск взаимосвязей между признаками
 - в) Предсказание числовых значений на основе имеющихся данных
7. Какую задачу решает поиск ассоциативных правил?
- а) Распределение объектов по заранее заданным классам
 - б) Поиск взаимосвязей между признаками
 - в) Поиск аномалий в данных
8. Какую задачу решает кластеризация?
- а) Распределение объектов по заранее заданным классам
 - б) Поиск взаимосвязей между признаками
 - в) Группировка объектов на основе их сходства
9. Какой коэффициент используется для измерения силы линейной связи между двумя переменными?
- а) Коэффициент корреляции Пирсона
 - б) Ранговая корреляция
 - с) Критерий Спирмена
10. Какой тип графика наиболее удобен для сравнения значений различных категорий?
- а) Столбчатая диаграмма
 - б) Круговая диаграмма
 - с) Линейный график
 - д) Точечная диаграмма
11. Какой инструмент программы MsExcel используется для создания дашборда?
- а) Формулы и функции
 - б) Панель инструментов
 - с) Параметры страницы
 - д) Сводные таблицы
12. Что такое дашборд в программе MsExcel?
- а) Крупномасштабная диаграмма
 - б) Интерактивная панель управления для отображения данных
 - с) Сложный алгоритм обработки данных
 - д) Режим просмотра данных без возможности редактирования

В тестовых заданиях предусмотрен 1 правильный ответ.

Критерии оценки результатов тестирования, 10-балльная шкала
10 баллов = 100% (правильных ответов)

96% ≤ 9 баллов ≤ 99%
91% ≤ 8 баллов ≤ 95%
86% ≤ 7 баллов ≤ 90%
80% ≤ 6 баллов ≤ 85%
70% ≤ 5 баллов ≤ 79%
60% ≤ 4 балла ≤ 69%
50% ≤ 3 балла ≤ 59%
40% ≤ 2 балла ≤ 49%
1 балл ≤ 40 %

Студент, выполнивший 60 % и более заданий, получает зачет по данной дисциплине.

5.2.2 Внеаудиторная самостоятельная работа

В рамках теоретической подготовки к практическим занятиям студент должен самостоятельно подготовить ответы на следующие вопросы по темам дисциплины:

Тема 1. Статистические методы

1. Что такое методы предварительного анализа данных и почему они важны при работе с данными?
2. Какие основные методы описательной статистики используются для анализа данных?
3. Какие характеристики центральной тенденции помогают определить типичное значение в наборе данных?
4. Какие характеристики вариации используются для измерения разброса значений в данных?
5. Какие методы можно применить для очистки данных от выбросов и ошибок?
6. Что такое нормировка данных и какие преимущества она может иметь при анализе данных?
7. Какие методы можно использовать для обработки пропущенных значений в данных?
8. Какие статистические показатели и графики могут помочь визуализировать распределение данных?
9. Какие методы можно применить для исследования связи между переменными в наборе данных?

Тема 2. Корреляционный и регрессионный анализ

1. Что такое корреляционный анализ и какая его основная цель?
2. Что представляет собой коэффициент корреляции Пирсона и как он интерпретируется?
3. Какая роль ранговой корреляции в анализе данных и в чем ее отличие от коэффициента корреляции Пирсона?

4. Что такое критерий Спирмена и как он применяется в корреляционном анализе?
5. В чем заключается множественная корреляция и какие преимущества она имеет перед простой корреляцией?
6. Как можно провести частную корреляцию и какие выводы о взаимосвязи переменных можно сделать на основе результатов?
7. Что представляет собой регрессионный анализ и какие задачи он решает?
8. Какие виды регрессии существуют и в каких случаях они применяются?

Тема 3. Создание запросов в Power Query. Возможности Power Query для очистки и преобразования данных.

1. Что такое Power Query и какую роль он играет в анализе данных?
2. Какие возможности предоставляет Power Query для импорта данных из различных источников?
3. Какие инструменты в Power Query позволяют очищать данные от ошибок, дубликатов и пропущенных значений?
4. Как можно использовать Power Query для объединения нескольких таблиц или файлов данных?
5. Какие операции преобразования данных можно выполнять с помощью Power Query?
6. Как можно применять фильтры и сортировку данных в Power Query?
7. Какие методы и инструменты в Power Query помогают обрабатывать текстовые данные?
8. Какие функции Power Query позволяют работать с датами и временем?
9. Какие способы сохранения и использования созданных запросов в Power Query существуют, например, для автоматического обновления данных или создания отчетов?

Тема 4. Методы классификации

1. Что такое дерево решений и какой основной принцип лежит в его основе?
2. Какие виды задач можно решать с помощью деревьев решений в анализе данных?
3. Каким образом происходит анализ признаков объектов при построении дерева решений?
4. Какие критерии и функции качества разбиения используются для определения наилучшего разделения в дереве решений?
5. Что такое индекс Джини и как он используется для измерения неопределенности в дереве решений?
6. Как энтропия применяется в контексте деревьев решений и какая роль она играет в принятии решений?
7. Что представляет собой ошибка классификации и как она учитывается при построении дерева решений?

8. Какие методы и критерии можно использовать для остановки построения дерева решений и избежания переобучения?
9. 9. Что такое сокращение дерева?
10. Как можно извлечь правила из построенного дерева решений и как они могут быть интерпретированы в контексте анализа данных?

5.3. Перечень средств диагностики результатов учебной деятельности

1. Тестирование знаний магистрантов по темам дисциплины;
2. Выполнение индивидуальных практических заданий;
3. Опрос;
4. Зачет.

Вопросы к зачету

1. Основные понятия интеллектуального анализа данных. Понятие данных. Типы шкал наборов данных. Типы наборов данных.
2. Задача классификации и регрессии. Задача поиска ассоциативных правил. Задача кластеризации.
3. Примеры практического применения дисциплины ИАД: медицина, биоинформатика, геология, банковское дело, торговля, другие области применения.
4. Основные этапы процесса интеллектуального анализа данных.
5. Методы предварительного анализа данных. Описательная статистика.
6. Характеристики центральной тенденции. Характеристики вариации.
7. Очистка данных. Нормировка и стандартизация данных.
8. Анализ выбросов и аномальных значений.
9. Обработка пропущенных значений.
10. Корреляционный анализ. Коэффициент корреляции Пирсона. Ранговая корреляция. Критерий Спирмена. Множественная и частная корреляции.
11. Регрессионный анализ.
12. Основные понятия и функции Power Query. Процесс подключения к источникам данных и загрузки данных в Power Query.
12. Основные методы трансформации данных, включая фильтрацию, сортировку, объединение таблиц и другие операции в Power Quer. Создание запросов. Возможности Power Query для очистки и преобразования данных.
13. Понятие визуализации данных, базовые принципы. Исследовательская и презентационная задачи визуализации.
14. Графики и диаграммы. Инфографика и схемы. Визуализация данных в программе MsExcel.
15. Бизнес аналитика и информационные панели (дашборды). Построение сводных таблиц и диаграмм для вычисления, сведения и анализа данных в программе MsExcel.
16. Построение дашборда в программе MsExcel.
17. Метода анализа иерархий.
18. Методы классификации. Деревья решений. Анализ данных с использованием деревьев решений. Анализ признаков объектов. Критерии и функции качества разбиения.

19. Индекс Джини. Энтропия. Ошибка классификации. Остановка построения дерева. Сокращение дерева. Извлечение правил из деревьев. Алгоритмы построения деревьев решений.

20. Основы кластерного анализа. Математические характеристики кластера. Критерии качества кластеризации. Выделение значимых кластеров. Иерархические агломеративные и дивизимные методы иерархического кластерного анализа. Неиерархические итерационные методы кластерного анализа.

21. Алгоритм k-средних.

Академическая оценка по дисциплине «Технологии интеллектуального анализа данных», которая фиксируется в зачетной книжке студента, определяется на основе суммы баллов итогового семестрового рейтинга. Итоговая оценка формируется в соответствии со шкалой, приведенной ниже в таблице:

Шкала оценивания результатов обучения по дисциплине

Количество баллов итогового семестрового рейтинга	Оценка
85–100	«Отлично»
75–84	«Хорошо»
60–74	«Удовлетворительно»
Менее 60	«Неудовлетворительно»

Характеристика оценки знаний в соответствии с четырех балльной шкалой

Оценка	Студент обнаруживает ...	Применение
«Отлично»	Всестороннее, систематическое и глубокое знание учебно-программного материала, умение свободно выполнять задания, предусмотренные программой, усвоил основную и знаком с дополнительной литературой, рекомендованной программой.	Как правило, выставляется студентам, усвоившим взаимосвязь основных понятий дисциплины и их значение для приобретаемой профессии, проявившим творческие способности в понимании, изложении и использовании учебного материала.
«Хорошо»	Полное знание учебного материала, успешно выполняет предусмотренные в программе задания, усвоил основную литературу, рекомендованную в программе. Студент допускает ошибки, которые исправляет самостоятельно в процессе работы (ответа).	Как правило, выставляется студентам, показавшим систематический характер знаний по дисциплине и способным к их самостоятельному пополнению и обновлению в ходе дальнейшей учебной работы и профессиональной деятельности.
«Удовлетворительно»	Знание основного учебного материала в объеме, необходимом для дальнейшей учебы и	Выставляется студентам, допустившим погрешности в ответе на экзамене и при выполнении

	<p>предстоящей работы по профессии, справляется с выполнением заданий, предусмотренных программой, знаком с основной литературой, рекомендованной программой. Студент допускает ошибки, которые не может исправить самостоятельно в процессе работы (ответа).</p>	<p>экзаменационных заданий, но обладающим необходимыми знаниями для их устранения под руководством преподавателя.</p>
«Неудовлетворительно»	<p>Пробелы в знаниях основного учебного материала, допускает принципиальные ошибки в выполнении предусмотренных программой заданий.</p>	<p>Ставится студентам, которые не могут продолжить обучение или приступить к профессиональной деятельности по окончании вуза без дополнительных занятий по соответствию</p>

5.4 Дополнительные критерии оценки результатов учебной деятельности обучающихся в учреждениях высшего образования по десятибалльной шкале

10 баллов	<ul style="list-style-type: none"> – систематизированные, глубокие и полные знания по всем разделам учебной программы учреждения высшего образования по учебной дисциплине, а также по основным вопросам, выходящим за ее пределы; – точное использование научной терминологии (в том числе на иностранном языке), грамотное, логически правильное изложение ответа на вопросы; – безупречное владение инструментарием учебной дисциплины, умение его эффективно использовать в постановке и решении научных и профессиональных задач; – выраженная способность самостоятельно и творчески решать сложные проблемы в нестандартной ситуации; – полное и глубокое усвоение основной, дополнительной литературы, по изучаемой учебной дисциплине; – умение свободно ориентироваться в теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им аналитическую оценку, использовать научные достижения других дисциплин; – творческая самостоятельная работа на практических, лабораторных занятиях, активное творческое участие в групповых обсуждениях высокий уровень культуры исполнения заданий
9 баллов	<ul style="list-style-type: none"> – систематизированные, глубокие и полные знания по всем разделам учебной программы учреждения высшего образования по учебной дисциплине; – точное использование научной терминологии (в том числе на иностранном языке), грамотное, логически правильное изложение ответа на вопросы; – владение инструментарием учебной дисциплины, умение его эффективно

	<p>использовать в постановке и решении научных и профессиональных задач;</p> <ul style="list-style-type: none"> – способность самостоятельно и творчески решать сложные проблемы в нестандартной ситуации в рамках учебной программы учреждения высшего образования по учебной дисциплине; – полное усвоение основной и дополнительной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – умение ориентироваться в теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им аналитическую оценку; – систематическая, активная самостоятельная работа на практических, лабораторных занятиях, творческое участие в групповых обсуждениях, высокий уровень культуры исполнения заданий
8 баллов	<ul style="list-style-type: none"> – систематизированные, глубокие и полные знания по всем разделам учебной программы учреждения высшего образования по учебной дисциплине в объеме учебной программы учреждения высшего образования по учебной дисциплине; – использование научной терминологии (в том числе на иностранном языке), грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы и обобщения; – владение инструментарием учебной дисциплины (методами комплексного анализа, техникой информационных технологий), умение его использовать в постановке и решении научных и профессиональных задач; – способность самостоятельно решать сложные проблемы в рамках учебной программы учреждения высшего образования по учебной дисциплине; – усвоение основной и дополнительной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – умение ориентироваться в теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им аналитическую оценку; – активная самостоятельная работа на практических, лабораторных занятиях, систематическое участие в групповых обсуждениях, высокий уровень культуры исполнения заданий
7 баллов	<ul style="list-style-type: none"> – систематизированные, глубокие и полные знания по всем разделам учебной программы учреждения высшего образования по учебной дисциплине; – использование научной терминологии (в том числе на иностранном языке), грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы и обобщения; – владение инструментарием учебной дисциплины, умение его использовать в постановке и решении научных и профессиональных задач; – свободное владение типовыми решениями в рамках учебной программы учреждения высшего образования по учебной дисциплине; – усвоение основной и дополнительной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – умение ориентироваться в основных теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им аналитическую оценку; – самостоятельная работа на практических, лабораторных занятиях, участие в групповых обсуждениях, (высокий) уровень культуры исполнения заданий
6 баллов	<ul style="list-style-type: none"> – достаточно полные и систематизированные знания в объеме учебной программы учреждения высшего образования по учебной дисциплине; – использование необходимой научной терминологии, грамотное, логически правильное изложение ответа на вопросы, умение делать обобщения и

	<p>обоснованные выводы;</p> <ul style="list-style-type: none"> – владение инструментарием учебной дисциплины, умение его использовать в решении учебных и профессиональных задач; – способность самостоятельно применять типовые решения в рамках, учебной программы учреждения высшего образования по учебной дисциплине; – усвоение основной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – умение ориентироваться в базовых теориях, концепциях и направлениях по изучаемой дисциплине и давать им сравнительную оценку; – активная самостоятельная работа на практических, лабораторных занятиях, периодическое участие в групповых обсуждениях, высокий уровень культуры исполнения заданий
5 баллов	<ul style="list-style-type: none"> – достаточные знания в объеме учебной программы учреждения высшего образования по учебной дисциплине; – использование научной терминологии, грамотное, логически правильное изложение ответа на вопросы, умение делать выводы; – владение инструментарием учебной дисциплины, умение его использовать в решении учебных и профессиональных задач; – способность самостоятельно применять типовые решения в рамках учебной программы учреждения высшего образования по учебной дисциплине; – усвоение основной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – умение ориентироваться в базовых теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им сравнительную оценку; – самостоятельная работа на практических, лабораторных занятиях, фрагментарное участие в групповых обсуждениях, достаточный уровень культуры исполнения заданий
4 балла	<ul style="list-style-type: none"> – достаточный объем знаний в рамках образовательного стандарта высшего образования; – усвоение основной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – использование научной терминологии, логическое изложение ответа на вопросы, умение делать выводы без существенных ошибок; – владение инструментарием учебной дисциплины, умение его использовать в решении стандартных (типовых) задач; – умение под руководством преподавателя решать стандартные (типовые) задачи; – умение ориентироваться в основных теориях, концепциях и направлениях по изучаемой учебной дисциплине и давать им оценку; – работа под руководством преподавателя на практических, лабораторных занятиях, допустимый уровень культуры исполнения заданий

<p>3 балла, не зачтено</p>	<ul style="list-style-type: none"> – недостаточно полный объем знаний в рамках образовательного стандарта высшего образования; – знание части основной литературы, рекомендованной учебной программой учреждения высшего образования по учебной дисциплине; – использование научной терминологии, изложение ответа на вопросы с существенными, логическими ошибками; – слабое владение инструментарием учебной дисциплины, некомпетентность в решении стандартных (типовых) задач; – неумение ориентироваться в основных теориях, концепциях и направлениях изучаемой учебной дисциплины; – пассивность на практических и лабораторных занятиях, низкий уровень культуры исполнения заданий
<p>2 балла, не зачтено</p>	<ul style="list-style-type: none"> – фрагментарные знания в рамках образовательного стандарта высшего образования; – знания отдельных литературных источников, рекомендованных учебной программой учреждения высшего образования по учебной дисциплине; – неумение использовать научную терминологию учебной дисциплины, наличие в ответе грубых, логических ошибок; – пассивность на практических и лабораторных занятиях, низкий уровень культуры исполнения заданий
<p>1 балл, не зачтено</p>	<ul style="list-style-type: none"> – отсутствие знаний и (компетенций) в рамках образовательного, стандарта высшего образования, отказ от ответа, неявка на аттестацию без уважительной причины

6. ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
			Согласование не требуется. Протокол № 08 от 29 мая 2023 г.

7. ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ

на ____ / ____ учебный год

№№ пп	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
_____ (протокол № ____ от _____ 20__ г.)
(название кафедры)

Заведующий кафедрой
к.э.н., доцент



И.И. Бычек

УТВЕРЖДАЮ
Декан факультета
к.э.н., доцент



М.Е. Карпицкая