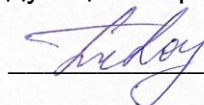


**МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)**

УТВЕРЖДАЮ

Заведующий кафедрой экономической теории
и мировой экономики



д.э.н., проф. Т.Н.Гоголева

12.05.2022 г.

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
Б1.В.ДВ.04.01 Интеллектуальный анализ данных
в экономических исследованиях**

- 1. Код и наименование направления подготовки:** 38.04.01 Экономика
- 2. Профиль подготовки:** Экономика организаций и рынков
- 3. Квалификация выпускника:** магистр
- 4. Форма обучения:** очная
- 5. Кафедра, отвечающая за реализацию дисциплины:** кафедра экономической теории и мировой экономики
- 6. Составители программы:** Азарнова Т.В., д.т.н., профессор кафедры экономической теории и мировой экономики
- 7. Рекомендована:** Научно-методическим советом экономического факультета ВГУ от 21.04.2022 г., протокол №4
- 8. Учебный год:** 2023/2024 **Семестр:** 4

9. Цели и задачи учебной дисциплины:

Целями освоения учебной дисциплины являются:

- формирование у обучающихся представления о типах задач, возникающих в области интеллектуального анализа (DataMining) экономических данных и методах их решения, которые помогут обучающимся выявлять, формализовать и успешно решать практические задачи анализа данных, возникающие в процессе прогнозирования показателей деятельности предприятия, отрасли, региона и экономики в целом;

- выбор методов решения исследовательских задач, проведение исследования в рамках реализации научного проекта с использованием законов естественно-научных дисциплин, применения математического аппарата, методов оптимизации, теории вероятностей, математической статистики, системного анализа для принятия решений в области стратегического и тактического планирования и организации производства.

Задачи учебной дисциплины:

- приобрести знания по методам и моделям Data Mining;
- изучить методы, основанные на не четкой логике;
- получить представление о преимуществах и недостатках использования нейротехнологий в экономических исследованиях;
- получить представление об алгоритмах построения деревьев решений; освоить методы кластеризации экономических данных;
- получить практические навыки в работе с существующими программными пакетами по интеллектуальному анализу данных.

10. Место учебной дисциплины в структуре ООП: дисциплина «Интеллектуальный анализ данных в экономических исследованиях» относится к дисциплинам по выбору вариативной части базового цикла.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:

Код	Название компетенции	Коды	Индикатор(ы)	Планируемые результаты обучения
ПК-3	Способен составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом с использованием современных технических средств и информационных технологий	ПК-3.1	ПК-3.1 Использует современные технические средства и информационные технологии при анализе социально-экономических систем	<i>Знать:</i> -методы обработки результатов исследования с применением интеллектуального анализа данных; -методы обработки знаний, полученных в результате работы с интеллектуальными информационными системами, базирующимися на инструментах интеллектуального анализа данных; <i>Уметь:</i> -применять в исследованиях стандартное и оригинальное программное обеспечение, базирующееся на инструментах интеллектуального анализа данных. <i>Владеть:</i> -навыками работы с современными интеллектуальными информационными системами.
ПК-4	Способен планировать работу, выбирать методы	ПК-4.2 ПК-4.4	ПК-4.2 Анализирует и обрабатывает ин-	<i>Знать:</i> -основные стадии проведения исследований в области проектирования интел-

решения исследовательских задач, проводить исследования в рамках реализации научного проекта адекватно поставленным целям с учетом широкого понимания профессиональной области обучения, в том числе на междисциплинарном уровне	формацию по тематике исследования в выбранной области науки на основании широкого понимания профессиональной области, в том числе на междисциплинарном уровне ПК-4.4 Обрабатывает полученные данные с использованием современных методов анализа информации	лектуальных систем, базирующихся на инструментах интеллектуального анализа данных; - технологию планирования экспериментов в области решения экономических задач с помощью интеллектуальных систем; <i>Уметь:</i> - ставить задачи исследования с применением инструментов интеллектуального анализа данных, -осуществлять эксперименты по решению экономических задач с применением инструментов интеллектуального анализа данных; <i>Владеть:</i> - средствами составления отчетов по результатам проведения эксперимента, направленного на решение экономических задач с применением инструментов интеллектуального анализа данных
--	--	---

12. Объем дисциплины в зачетных единицах/час. — 3 ЗЕТ / 108 час.

Форма промежуточной аттестации — зачет с оценкой

13. Трудоемкость по видам учебной работы:

Вид учебной работы		Трудоемкость	
		Всего	По семестрам 4 семестр
Контактная работа		28	28
в том числе:	лекции	14	14
	практические	14	14
Самостоятельная работа		80	80
Промежуточная аттестация		Зачёт с оценкой	Зачёт с оценкой
Итого:		108	108

13.1 Содержание разделов дисциплины:

№ п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК
1. Лекции			
1	Базовые понятия и основные направления развития искусственного интеллекта (ИИ) и интеллектуальных систем (ИС)	Понятие искусственного интеллекта. Основные направления развития искусственного интеллекта. Данные и знания. Свойства, характеристики знаний. Процедурные и декларативные знания. Классификация знаний по глубине, по жесткости. Формализация знаний. Формальные языки. Языки (модели) представления знаний. Классификация моделей знаний и данных. Определение интеллектуальных систем. Основные источники научно-технической информации по основным направлениям, методам, моделям и инструментальным средствам конструирования интеллектуальных систем. Формализация знаний в интеллектуальных системах (ИС) Классификация ИС. Обобщенная функциональная структура ИС. Основные (базовые) свойства и возможности. Обобщенная типология интеллектуальных систем.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях

2	Основы нечеткого анализа	Нечеткие множества, нечеткие величины, нечеткие числа. Основные типы нечетких чисел. Операции над нечеткими числами. Сравнение нечетких чисел. Понятие нечеткой функции. Примеры прикладных задач с нечеткой информацией. Задача нечеткой кластеризации.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
3	Лингвистическая модель представления информации	Лингвистическая переменная. Лингвистическая шкала. Принцип «нечеткого большинства» и лингвистические кванторы. Лингвистические операторы агрегирования информации. Нечеткие системы: структура, база знаний, механизм вывода.	Интеллектуальный анализ данных в экономических исследованиях
4	Системы нечеткого вывода	Базовая архитектура систем нечеткого вывода. Основные этапы нечеткого вывода. Формирование базы правил систем нечеткого вывода. Фаззификация (Fuzzification). Агрегирование (Aggregation). Активизация (Activation). Аккумуляция (Accumulation). Дефаззификация (Defuzzification). Основные алгоритмы нечеткого вывода. Алгоритм Мамдани (Mamdani). Алгоритм Цукамото (Tsukamoto). Алгоритм Ларсена (Larsen). Алгоритм Сугено (Sugeno). Примеры использования систем нечеткого вывода в задачах управления	ЭУМК Мягкие вычисления; Интеллектуальный анализ данных
5	Нечеткое моделирование в среде MATLAB	Основные элементы системы MATLAB. Основные приемы работы в системе MATLAB. Редактор систем нечеткого вывода FIS. Редактор функций принадлежности. Редактор правил системы нечеткого вывода. Программа просмотра правил системы нечеткого вывода. Программа просмотра поверхности системы нечеткого вывода.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
6	Задачи Data Mining.	Классификация задач Data Mining. Сферы применения Data Mining	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
7	Задача классификации. Методы классификации.	Задача классификации Процесс классификации Методы, применяемые для решения задач классификации Точность классификации: оценка уровня ошибок Оценивание классификационных методов/ Деревья решений. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация. Нейронные сети	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
8	Задача кластеризации. Методы кластерного анализа.	Задача кластеризации Оценка качества кластеризации Процесс кластеризации Применение кластерного анализа Кластерный анализ в маркетинговых исследованиях Практика применения кластерного анализа в маркетинговых исследованиях. Иерархические методы. Итеративные методы.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
9	Задача прогнозирования	Задача прогнозирования Сравнение задач прогнозирования и классификации Прогнозирование и временные ряды Тренд, сезонность и цикл Точность прогноза Виды прогнозов Методы прогнозирования Задача визуализации	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
2. Лабораторные занятия			
1	Базовые понятия и основные направления развития искусственного интеллекта (ИИ) и интеллектуальных систем (ИС)	Понятие искусственного интеллекта. Данные и знания. Свойства, характеристики знаний. Процедурные и декларативные знания. Классификация знаний по глубине, по жесткости. Формализация знаний. Формальные языки. Языки (модели) представления знаний. Классификация моделей знаний и данных. Определение интеллектуальных систем. Формализация знаний в интеллектуальных системах (ИС) Классификация ИС. Обобщенная функциональная структура ИС. Основные (базовые) свойства и возможности. Обобщенная типология интеллектуальных систем.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
2	Основы нечеткого анализа	Нечеткие множества, нечеткие величины, нечеткие числа. Основные типы нечетких чисел. Операции над нечеткими числами. Сравнение нечетких чисел. Понятие нечеткой функции. Примеры прикладных задач с нечеткой информацией. Задача нечеткой кластеризации.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях

3	Лингвистическая модель представления информации	Лингвистическая переменная. Лингвистическая шкала. Принцип «нечеткого большинства» и лингвистические кванторы. Лингвистические операторы агрегирования информации. Нечеткие системы: структура, база знаний, механизм вывода.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
4	Системы нечеткого вывода	Базовая архитектура систем нечеткого вывода. Основные этапы нечеткого вывода. Формирование базы правил систем нечеткого вывода. Фаззификация (Fuzzification). Агрегирование (Aggregation). Активизация (Activation). Аккумуляция (Accumulation). Дефаззификация (Defuzzification). Основные алгоритмы нечеткого вывода. Алгоритм Мамдани (Mamdani). Алгоритм Цукамото (Tsukamoto). Алгоритм Ларсена (Larsen). Алгоритм Сугено(Sugeno). Примеры использования систем нечеткого вывода в задачах управления	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
5	Нечеткое моделирование в среде MATLAB	Основные элементы системы MATLAB. Основные приемы работы в системе MATLAB. Редактор систем нечеткого вывода FIS. Редактор функций принадлежности. Редактор правил системы нечеткого вывода. Программа просмотра правил системы нечеткого вывода. Программа просмотра поверхности системы нечеткого вывода.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
6	Задачи Data Mining.	Классификация задач Data Mining. Сферы применения Data Mining	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
7	Задача классификации. Методы классификации.	Задача классификации Процесс классификации Методы, применяемые для решения задач классификации Точность классификации: оценка уровня ошибок Оценивание классификационных методов/ Деревья решений. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация. Нейронные сети	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
8	Задача кластеризации. Методы кластерного анализа.	Задача кластеризации Оценка качества кластеризации Процесс кластеризации Применение кластерного анализа Кластерный анализ в маркетинговых исследованиях Практика применения кластерного анализа в маркетинговых исследованиях. Иерархические методы. Итеративные методы.	ЭУМК Интеллектуальный анализ данных в экономических исследованиях
9	Задача прогнозирования	Задача прогнозирования Сравнение задач прогнозирования и классификации Прогнозирование и временные ряды Тренд, сезонность и цикл Точность прогноза Виды прогнозов Методы прогнозирования Задача визуализации	ЭУМК Интеллектуальный анализ данных в экономических исследованиях

13.2. Темы (разделы) дисциплины и виды занятий:

№	Наименование темы (раздела) дисциплины	Виды занятий (количество часов)			
		Лекции	Практические	Самостоятельная работа	Всего
1	Базовые понятия и основные направления развития искусственного интеллекта (ИИ) и интеллектуальных систем (ИИС)	1	-	10	11
2	Основы нечеткого анализа	1	-	10	11
3	Лингвистическая модель представления информации	2	2	10	14
4	Системы нечеткого вывода	2	2	10	14
5	Нечеткое моделирование в среде MATLAB	-	2	5	7
6	Задачи Data Mining.	2	2	5	9
7	Задача классификации. Методы классификации.	2	2	10	14
8	Задача кластеризации. Методы кластерного анализа.	2	2	10	14
9	Задача прогнозирования	2	2	10	14
	Итого:	14	14	80	108

14. Методические указания для обучающихся по освоению дисциплины:

Для лучшего усвоения материала студентам рекомендуется домашняя работа с конспектами лекций, презентациями, выполнение практических заданий для самостоятельной работы, выполнение лабораторных работ, использование рекомендованной литературы и методических материалов. В рамках общего объема часов, отведенных для изучения дисциплины, предусматривается выполнение следующих видов самостоятельных работ студентов (СРС): изучение теоретического материала, выполнение в пакете статистического анализа данных заданий по темам, изученным на лекционных и практических занятиях.

При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей, вовремя подключаться к online занятиям, ответственно подходить к заданиям для самостоятельной работы.

15. Перечень основной и дополнительной литературы, интернет-ресурсов, необходимых для освоения дисциплины

а) основная литература

№ п/п	Источник
1	Ясницкий, Л. Н. Интеллектуальные системы : учебник / Л. Н. Ясницкий. — 2-е изд. — Москва : Лаборатория знаний, 2020. — 224 с. — ISBN 978-5-00101-897-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/151510
2	Белозерова, Г. И. Нечеткая логика и нейронные сети : учебное пособие : в 2 частях / Г. И. Белозерова, Д. М. Скуднев, З. А. Кононова. — Липецк : Липецкий ГПУ, [б. г.]. — Часть 1 — 2017. — 64 с. — ISBN 978-5-88526-875-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/111969
3	Целых, А. Н. Современные методы прикладной информатики в задачах анализа данных : учебное пособие / А. Н. Целых, А. А. Целых, Э. М. Котов. — Ростов-на-Дону : ЮФУ, 2021. — 130 с. — ISBN 978-5-9275-3783-9. — Текст : электронный // Лань:электронно-библиотечная система. — URL: https://e.lanbook.com/book/195357
4	Сапрыкин, О. Н. Интеллектуальный анализ данных: учебное пособие / О. Н. Сапрыкин. — Самара : Самарский университет, 2020. — 80 с. — ISBN 978-5-7883-1563-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/188906 .

б) дополнительная литература:

№ п/п	Источник
5	Киселев, В. Ю. Теория нечётных множеств и нечетная логика. Задачи и упражнения : учебное пособие / В. Ю. Киселев, Т. Ф. Калугина. — Иваново : ИГЭУ, 2019. — 72 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/154561
6	Талипов, Н. Г. Технологии интеллектуального анализа данных : учебно-методическое пособие / Н. Г. Талипов. — Казань : КНИТУ-КАИ, 2020. — 184 с. — ISBN 978-5-7579-2489-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/193529
7	Остроух, А. В. Системы искусственного интеллекта : монография / А. В. Остроух, Н. Е. Суркова. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 228 с. — ISBN 978-5-8114-8519-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/176662
8	Макшанов, А. В. Технологии интеллектуального анализа данных : учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/120063

в) информационные электронно-образовательные ресурсы:

№ п/п	Источник
9	www.lib.vsu.ru – Зональная научная библиотека ВГУ
10	edu.vsu.ru
11	ЭБС Лань
12	ЭБС ЮРАЙТ

16. Перечень учебно-методического обеспечения

№ п/п	Источник
1	Интеллектуальные системы: учебное пособие / А. Семенов, Н. Соловьев, Е. Чернопрудова, А. Цыганков; Оренбургский государственный университет. – Оренбург: Оренбургский государственный университет, 2013. – 236 с. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=259148 . – Текст : электронный.

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии, электронное обучение, смешанное обучение.

При реализации дисциплины могут использоваться технологии электронного обучения и дистанционные образовательные технологии на базе портала edu.vsu.ru, а также другие доступные ресурсы сети.

18. Материально-технического обеспечения дисциплины:

Учебные аудитории для проведения учебных (лекционных) занятий, оснащенные оборудованием и техническими средствами обучения: специализированная мебель, проектор, экран для проектора, компьютер с возможностью подключения к сети "Интернет", проводной микрофон, комплект активных громкоговорителей

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенции	Индикаторы достижения компетенции	Оценочные средства
1.	2	3	4	5
1.	Базовые понятия и основные направления развития искусственного интеллекта (ИИ) и интеллектуальных систем (ИИС)	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 1
2.	Основы нечеткого анализа	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 2
3.	Лингвистическая модель представления информации	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 3
4.	Системы нечеткого вывода	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 4
5.	Нечеткое моделирование в среде MATLAB	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 4
6.	Задачи Data Mining.	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Устный опрос

№ п/п	Наименование раздела дисциплины (модуля)	Компетенции	Индикаторы достижения компетенции	Оценочные средства
1.	2	3	4	5
7	Задача классификации. Методы классификации.	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 5
8.	Задача кластеризации. Методы кластерного анализа.	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 6
9	Задача прогнозирования	ПК-3 ПК-4	ПК-3.1. ПК-4.2 ПК-4.4.	Лабораторная работа 7
Промежуточная аттестация форма контроля – зачет с оценкой				Перечень вопросов Практическое задание

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1 Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств:

Лабораторная работа 1 (пример задания)

Задание.

Представить декларативное знание о понятии «Квартира» четырьмя моделями представления знаний:

1. в виде семантической сети.
2. в виде фреймов.
3. в виде логической модели.
4. в виде продукционной модели.

Квартира состоит из:

1. Кухня.
2. Гостиная.
3. Прихожая.
4. Спальня.
5. Детская.
6. Санитарный узел (туалет).
7. Ванная комната.
8. Кладовка.
9. Гардеробная.
10. Комната отдыха (игровая комната).
11. Спортивная комната (тренажерная).
12. Бытовая комната.

Дополнительные задания: компьютерный класс; компьютерный клуб.

Порядок выполнения и результаты.

Студенты по вариантам (вариант - часть квартиры) самостоятельно выполняют задание в рабочей тетради (лучше сначала на черновиках). Используется конспект лекций или рекомендуемая литература. Семантическая сеть должна содержать не менее 20 вершин с разными типами связей. Фреймовая модель должна содержать не менее 6 фреймов, связанных двумя типами связи. В продукции отразить все составляющие. В процессе работы каждый студент предоставляет преподавателю составленные модели. Преподаватель со студентом обсуждают и уточняют модели, при необходимости модели дорабатываются.

ся. Затем обучаемые составляют общую семантическую сеть и сеть фреймов понятия «Квартира» на доске и зарисовывают ее в тетрадь. Результат: четыре модели (согласно варианта) и две общие модели в рабочей тетради.

Лабораторная работа 2 (пример варианта)

Задача 1. Пусть функции принадлежности нечетких множеств A и B заданы в виде

$$\mu_A x = \begin{cases} 1 - e^x, & \text{если } x \leq 0, \\ 1 - e^{-x}, & \text{если } x \geq 0. \end{cases} \quad \mu_B(x) = \frac{1}{1 + 3x^2}.$$

Постройте графики функций принадлежности $A \cup B$, $A \cap B$, $A \setminus B$.

Задача 2. Для нечеткого множества с функцией принадлежности

$$\mu_A x = \begin{cases} \frac{1}{2} + \frac{1}{2} \sin \pi \cdot x + 1,5, & -2 \leq x \leq -1 \\ 1, & -1 \leq x \leq 1, \\ \frac{1}{2} - \frac{1}{2} \sin \pi \cdot x - 1,5, & 1 \leq x \leq 2, \\ 0, & \text{иначе,} \end{cases}$$

найдите линейный индекс нечеткости.

Задача 3. Пусть заданы нечеткие множества

$$A = a/0, b/0.3, c/0.7, d/0.1, e/0, f/0.2,$$

$$B = a/0.4, b/1, c/0.5, d/0.8, e/1, f/0.5.$$

Найдите нечеткое множество, являющееся их средним арифметическим.

Задача 4. Непрерывная, строго убывающая функция $\varphi: 0,1 \rightarrow 0,\infty$, такая что $\varphi 1 = 0$, называется *убывающим* генератором. Непрерывная, строго возрастающая функция $\varphi: 0,1 \rightarrow 0,\infty$, такая что $\varphi 0 = 0$, называется *возрастающим* генератором.

Найти условия, при которых функция $\varphi x = \pm \frac{1}{\sqrt{\alpha\gamma}} \arctg \left(x \sqrt{\frac{\gamma}{\alpha}} \right) + C \left(\frac{\alpha}{\gamma} > 0 \right)$ яв-

ляется возрастающим и/или убывающим генератором. Подтвердите графически Ваши выводы.

Лабораторная работа 3 (пример варианта)

Компания «Трансферт», занимающаяся перепродажей квартир, решила приобрести несколько квартир в новых жилых комплексах города. Для анализа качества жилья были выбраны следующие критерии: *характеристика района (ХР), репутация компании-застройщика (Р), инвестиционная привлекательность (ИП), планировка (П), вид из окна (В), наличие внешней инфраструктуры (ВИ), наличие автономной инфраструктуры (паркинг, служба консьерж, охрана, пассажирский и грузовой лифт, профессиональное управление комплексом) (АИ), скидки для покупателей и выгодные условия кредита (СК), средняя стоимость 1 кв. метра жилья (С)*. Оценки жилых комплексов в шкале $S = \{N, VL, L, M, H, VH, P\}$ представлены в таблице.

<i>Жилой комплекс</i>	<i>ХР</i>	<i>Р</i>	<i>ИП</i>	<i>П</i>	<i>В</i>	<i>ВИ</i>	<i>АИ</i>	<i>СК</i>	<i>С</i>
<i>Три богатыря</i>	VH	H	VH	H	M	VH	VH	VH	VH
<i>Северная корона</i>	M	H	M	M	H	M	L	M	M
<i>Лесная поляна</i>	M	M	M	H	VH	L	H	H	H
<i>Арка</i>	H	M	VH	M	M	L	L	H	H
<i>Петровский пассаж</i>	VH	H	M	VH	H	H	VH	H	VH
<i>Синяя птица</i>	M	M	H	M	H	M	M	M	M
<i>Алые паруса</i>	L	H	M	H	VH	L	L	M	M

Выберите из таблицы два жилых комплекса и по любым трем показателям, наиболее важным с Вашей точки зрения, постройте обобщенные оценки на основе функции порядкового взвешенного агрегирования (OWA) с вектором весов, полученным на основе квантора $Q x = x^2$. Какова стратегия агрегирования? Какой объект является лучшим?

Лабораторная работа 4 (пример работы)

Проектирование нечетких Мамдани (Mamdani), Цукамото (Tsukamoto), (Larsen), Сугено (Sugeno). В качестве примера рассмотрено проектирование нечетких систем Мамдани.

Проектирование нечетких систем Мамдани

В среде *MatLab* присутствуют следующие пять основных средств графического интерфейса пользователя, которые обеспечивают процесс нечеткого моделирования.

Редактор системы нечеткого вывода предполагает ввод и редактирование количества входных и выходных переменных и наименований переменных.

Редактор функции принадлежности используется для определения формы функции принадлежности, ассоциированной с каждым термом лингвистической переменной.

Редактор правил вывода применяется для редактирования списка правил, описывающих поведение моделируемой системы.

Средство просмотра правил вывода используется в целях диагностики и может показывать, например, активность правил или форму влияния отдельных функций принадлежности на результат нечеткого вывода.

Средство просмотра поверхности вывода используется для отображения зависимости выходной переменной от одной или двух входных переменных.

Эти средства связаны между собой динамически, и производимые изменения в одном из них влекут изменения в других.

Рассмотрим основные этапы проектирования нечетких систем Мамдани на примере создания системы нечеткого логического вывода, моделирующей зависимость

$$y = x_1^2 \cdot \sin(x_2 - 1), \quad x_1 \in [-7, 3], \quad x_2 \in [-4.4, 1.7],$$

для представления трехмерного изображения которой используем следующую программу:

```
% Построение графика функции y=x1^2*sin(x2-1)
% в области x1∈[-7,3] и x2∈[-4.4,1.7].
n=15;
x1=-7:10/(n-1):3;
x2=-4.4:6.1/(n-1):1.7;
y=zeros(n,n);
for j=1:n
y(j,:)=x1.^2*sin(x2(j)-1);
end
surf(x1,x2,y)
xlabel('x1')
ylabel('x2')
zlabel('y')
title('Target');
```

В результате выполнения программы получим графическое изображение

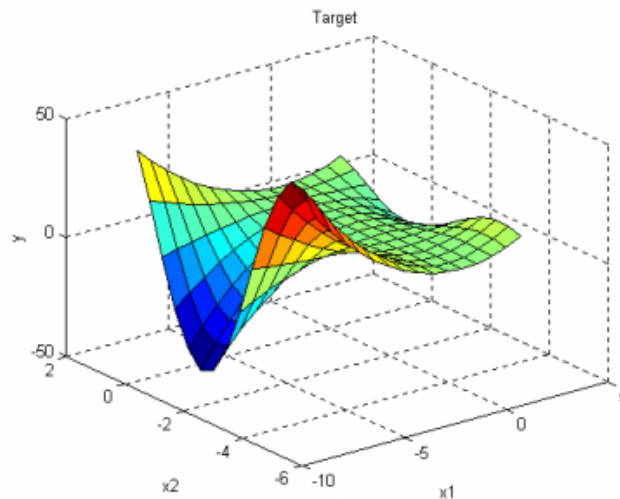


Рис. 1

Проектирование системы нечеткого логического вывода, соответствующей приведенному графику, состоит в выполнении следующей последовательности шагов.

Шаг 1. Для загрузки основного **fis**-редактора введем команду **fuzzy** в командной строке. После этого откроется новое графическое окно, показанное на рис. 2.

Шаг 2. Добавим вторую входную переменную. Для этого в меню **Edit** выбираем команду **Add input**.

Шаг 3. Переименуем первую входную переменную. Для этого сделаем один щелчок левой кнопкой мыши на блоке **input1**, введем новое обозначение **x1** в поле редактирования имени текущей переменной и нажмем **<Enter>**.

Шаг 4. Переименуем вторую входную переменную. Для этого сделаем один щелчок левой кнопкой мыши на блоке **input2**, введем новое обозначение **x2** в поле редактирования имени текущей переменной и нажмем **<Enter>**.

Шаг 5. Переименуем выходную переменную. Для этого сделаем один щелчок левой кнопкой мыши на блоке **output1**, введем новое обозначение **y** в поле редактирования имени текущей переменной и нажмем **<Enter>**.

Шаг 6. Зададим имя системы. Для этого в меню **File** выбираем в подменю **Export** команду **To disk** и вводим имя файла, например, **first**.

Шаг 7. Перейдем в редактор функций принадлежности. Для этого сделаем двойной щелчок левой кнопкой мыши на блоке **x1**.

Шаг 8. Зададим диапазон изменения переменной **x1**. Для этого введем **-7 3** в поле **Range** (рис. 3) и нажмем **<Enter>**.

Шаг 9. Зададим функции принадлежности переменной **x1**. Для лингвистической оценки этой переменной будем использовать 3 термина с треугольными функциями принадлежности. Для этого в меню **Edit** выберем команду **Add MFs...** В результате появится диалоговое окно выбора типа и количества

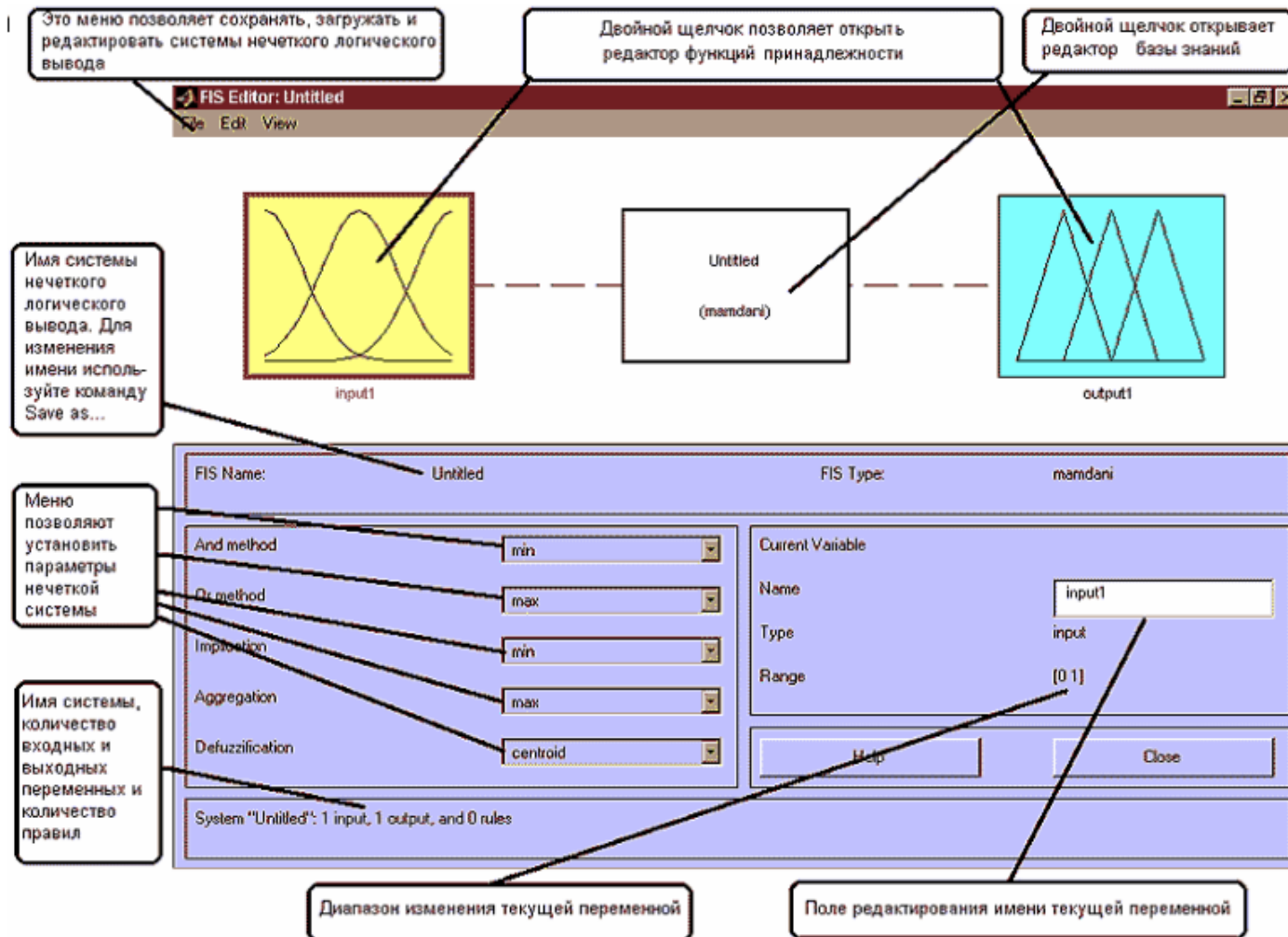


Рис. 2. Окно редактора FIS-Editor

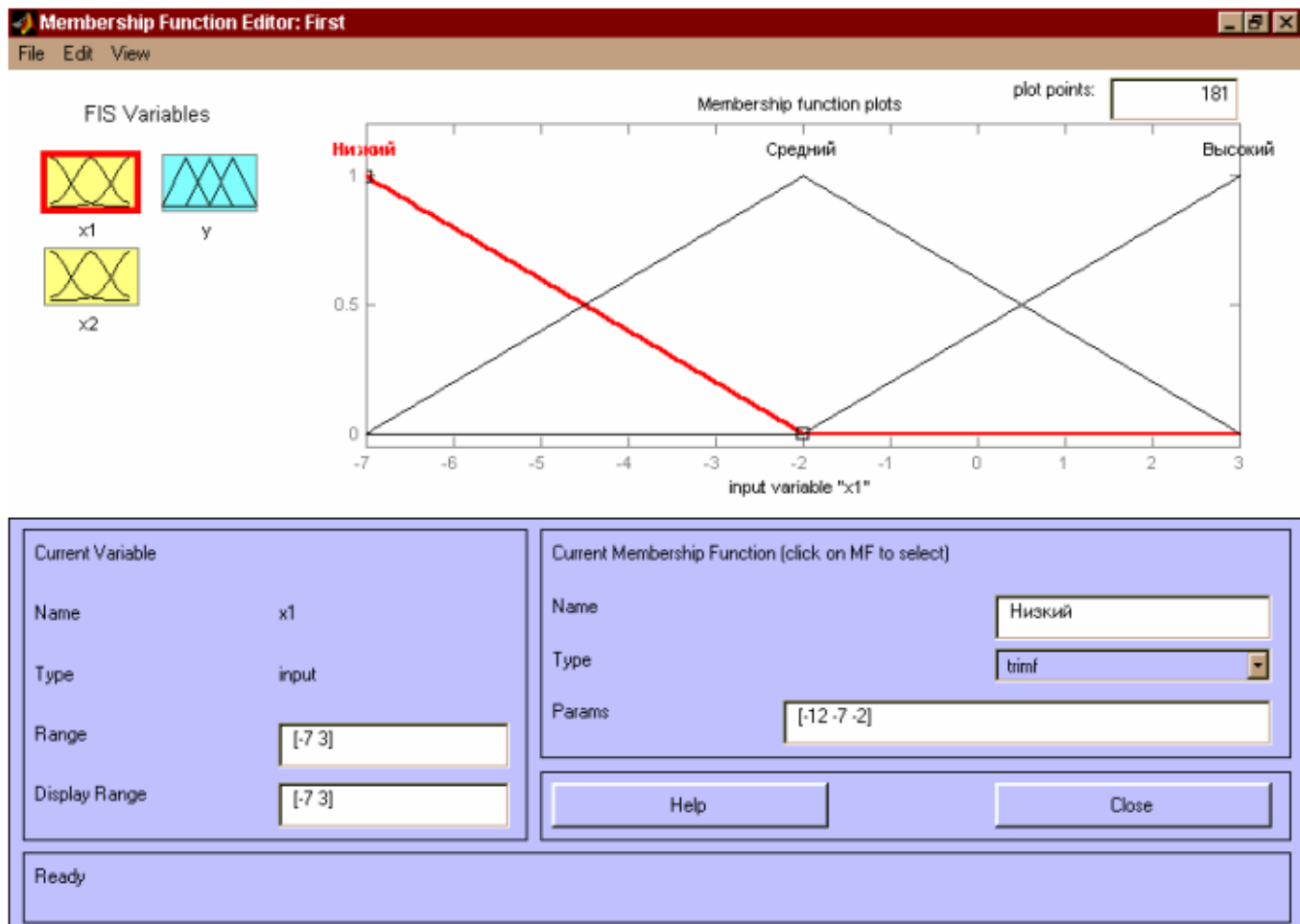


Рис. 3. Функции принадлежности переменной x1

функций принадлежности. По умолчанию это 3 термина с треугольными функциями принадлежности. Поэтому просто нажимаем **<Enter>**.

Шаг 10. Зададим наименования термов переменной **x1**. Для этого делаем один щелчок левой кнопкой мыши по графику первой функции принадлежности (рис. 3). Затем вводим наименование термина, например, *низкий* в поле **Name** и нажимаем **<Enter>**. После делаем один щелчок левой кнопкой мыши по графику второй функции принадлежности и вводим наименование термина, например, *средний* в поле **Name** и нажимаем **<Enter>**. Еще раз делаем один щелчок левой кнопкой мыши по графику третьей функции принадлежности и вводим наименование термина, например, *высокий* в поле **Name** и нажимаем **<Enter>**. В результате получим графическое окно, изображенное на рис. 3.

Шаг 11. Зададим функции принадлежности термов для переменной **x2**. Для лингвистической оценки этой переменной будем использовать 5 термов с гауссовыми функциями принадлежности. Для этого активизируем переменную **x2** с помощью щелчка левой кнопки мыши на блоке **x2**. Зададим диапазон изменения переменной **x2**. Для этого введем **-4.4 1.7** в поле **Range** (рис. 4) и нажимаем **<Enter>**. Затем в меню **Edit** выберем команду **Add MFs...** В появившемся диалоговом окне выбираем тип функции принадлежности **gaussmf** в поле **MF type** и **5** термов в поле **Number of MFs**. После этого нажимаем **<Enter>**.

Шаг 12. По аналогии с *шагом 10* зададим следующие наименования термов переменной **x2**: *низкий, ниже среднего, средний, выше среднего, высокий*. В результате получим графическое окно, изображенное на рис. 4.

Шаг 13. Зададим функции принадлежности переменной **y**. Для лингвистической оценки этой переменной будем использовать 5 термов с треугольными функциями принадлежности. Для этого активизируем переменную **y** с помощью щелчка левой кнопки мыши на блоке **y**. Зададим диапазон изменения переменной **y**. Для этого введем **-50 50** в поле **Range** (рис. 5) и нажимаем **<Enter>**. Затем в меню **Edit** выберем команду **Add MFs...** В появившемся диалоговом окне выбираем **5** термов в поле **Number of MFs**. После этого нажимаем **<Enter>**.

Шаг 14. По аналогии с *шагом 10* зададим следующие наименования термов переменной **y**: *низкий, ниже среднего, средний, выше среднего, высокий*. В результате получим графическое окно, изображенное на рис. 5.

Шаг 15. Перейдем в редактор базы знаний **RuleEditor**. Для этого в меню **Edit** выберем команду **Edit rules...**

Шаг 16. На основе визуального наблюдения за графиком, изображенным на рис. 1, сформулируем следующие девять правил.

1. Если $x1=средний$, то $y=средний$.
2. Если $x1=низкий$ и $x2=низкий$, то $y=высокий$.
3. Если $x1=низкий$ и $x2=высокий$, то $y=высокий$.
4. Если $x1=высокий$ и $x2=высокий$, то $y=выше среднего$.

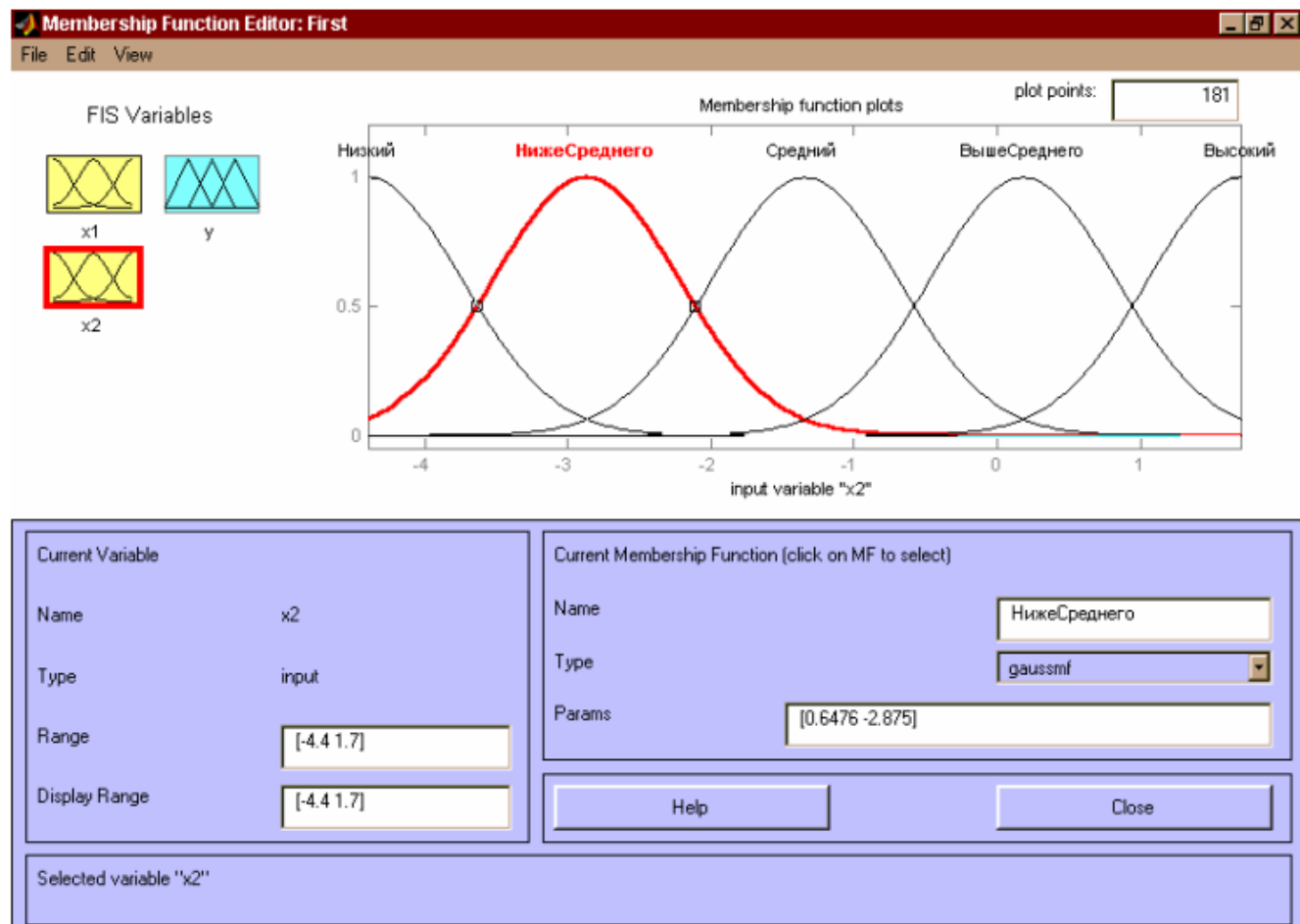


Рис. 4. Функции принадлежности переменной x2

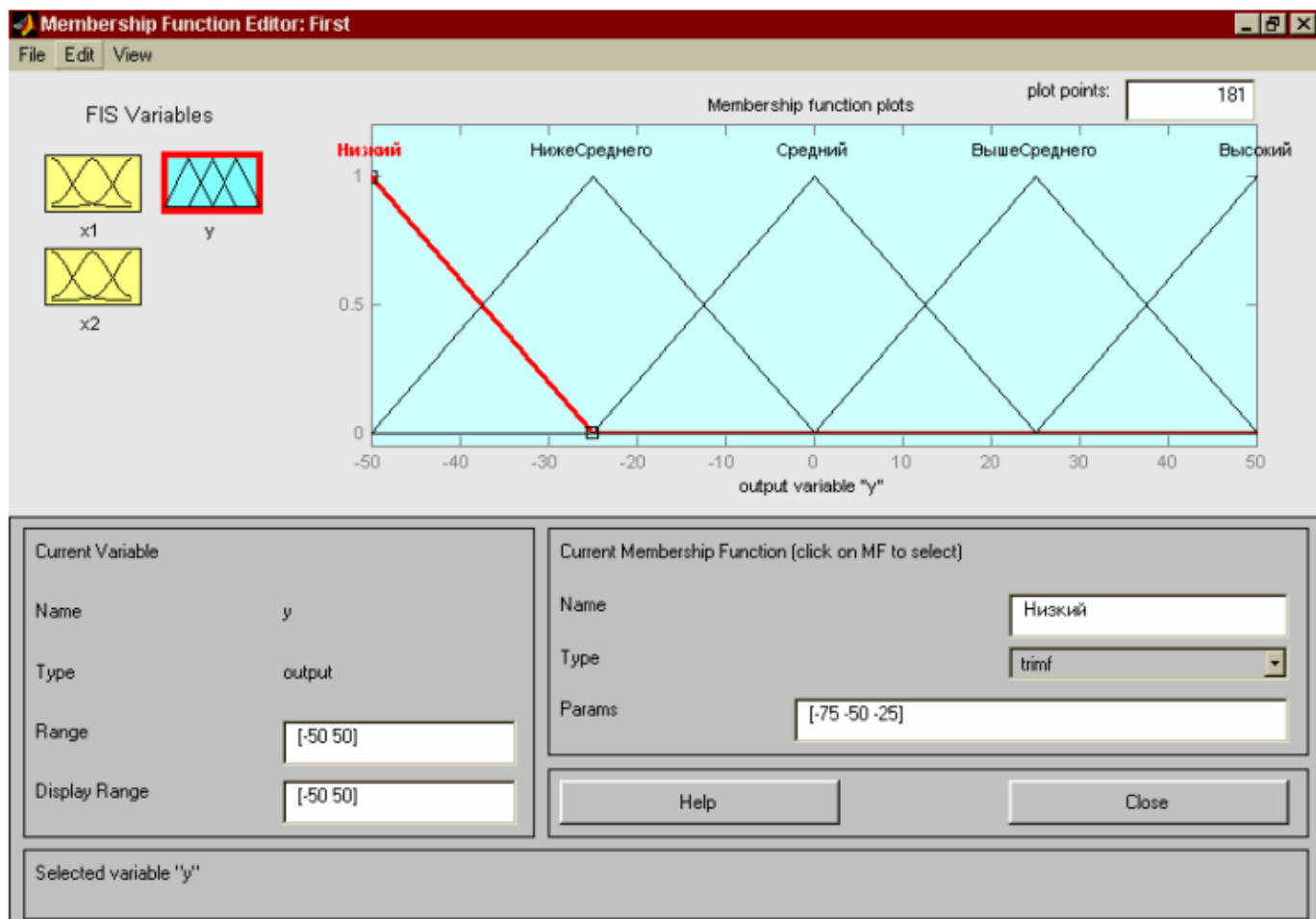


Рис. 5. Функции принадлежности переменной y

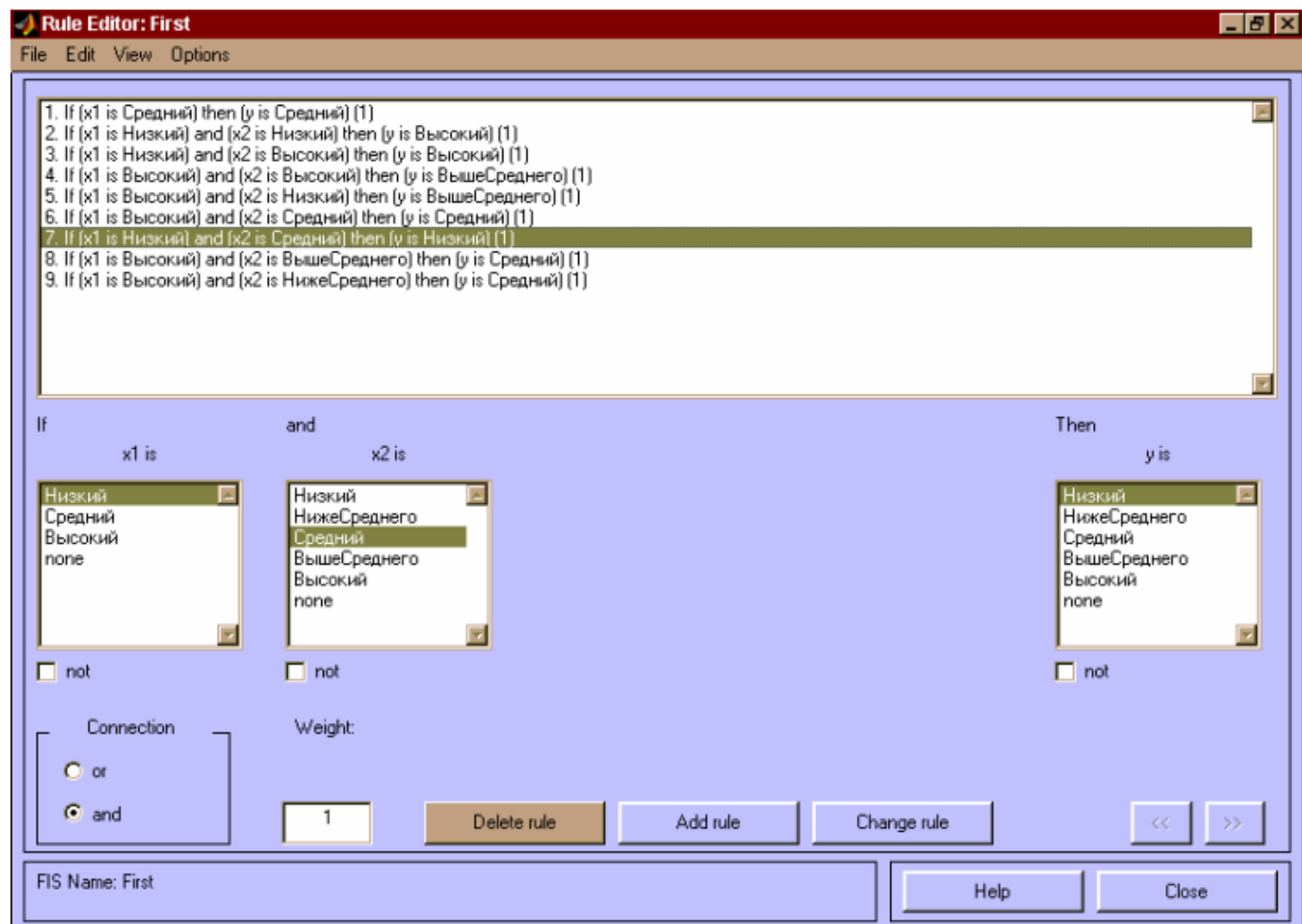


Рис. 6. База знаний в RuleEditor

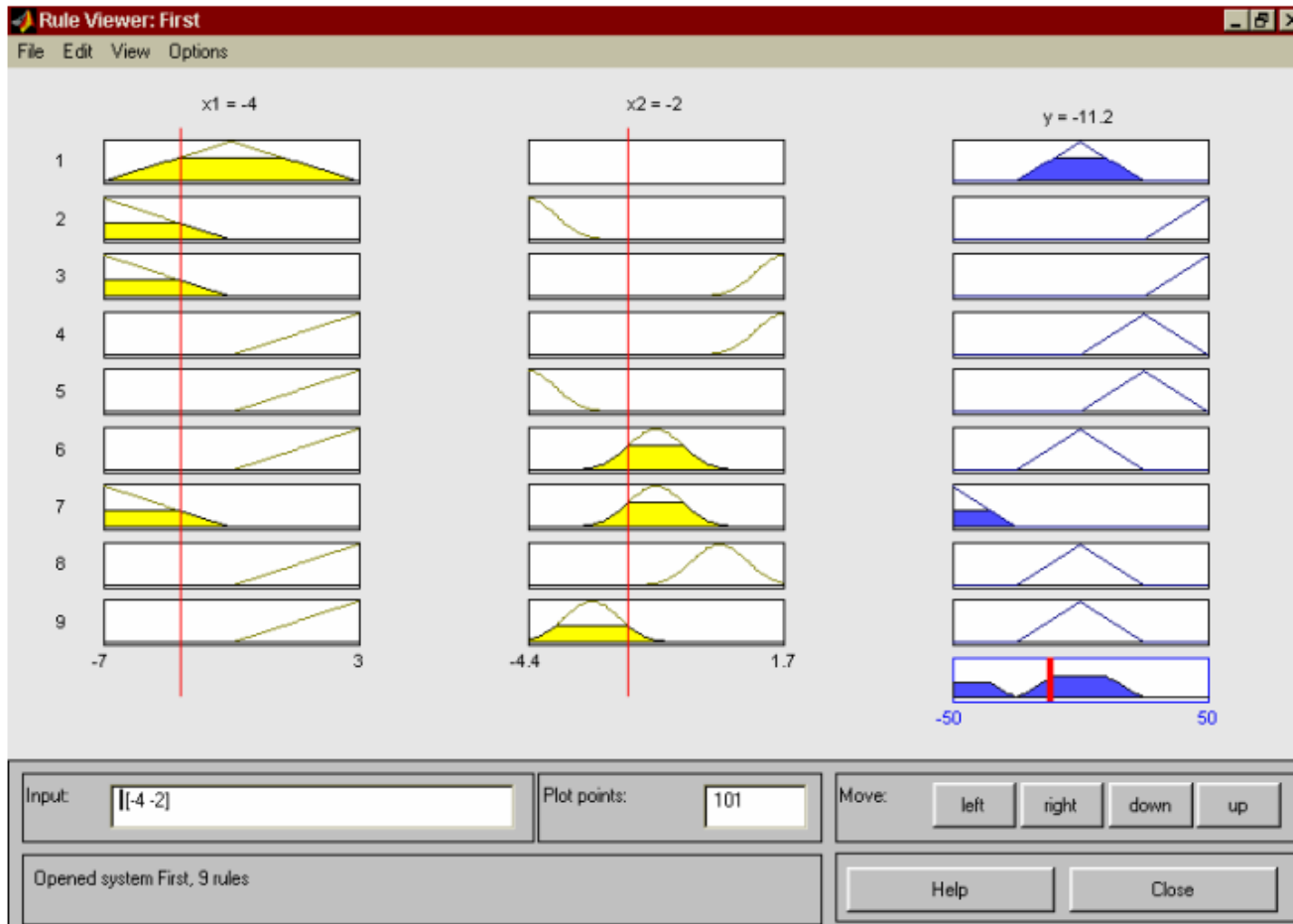


Рис. 7. Визуализация нечеткого логического вывода в RuleViewer

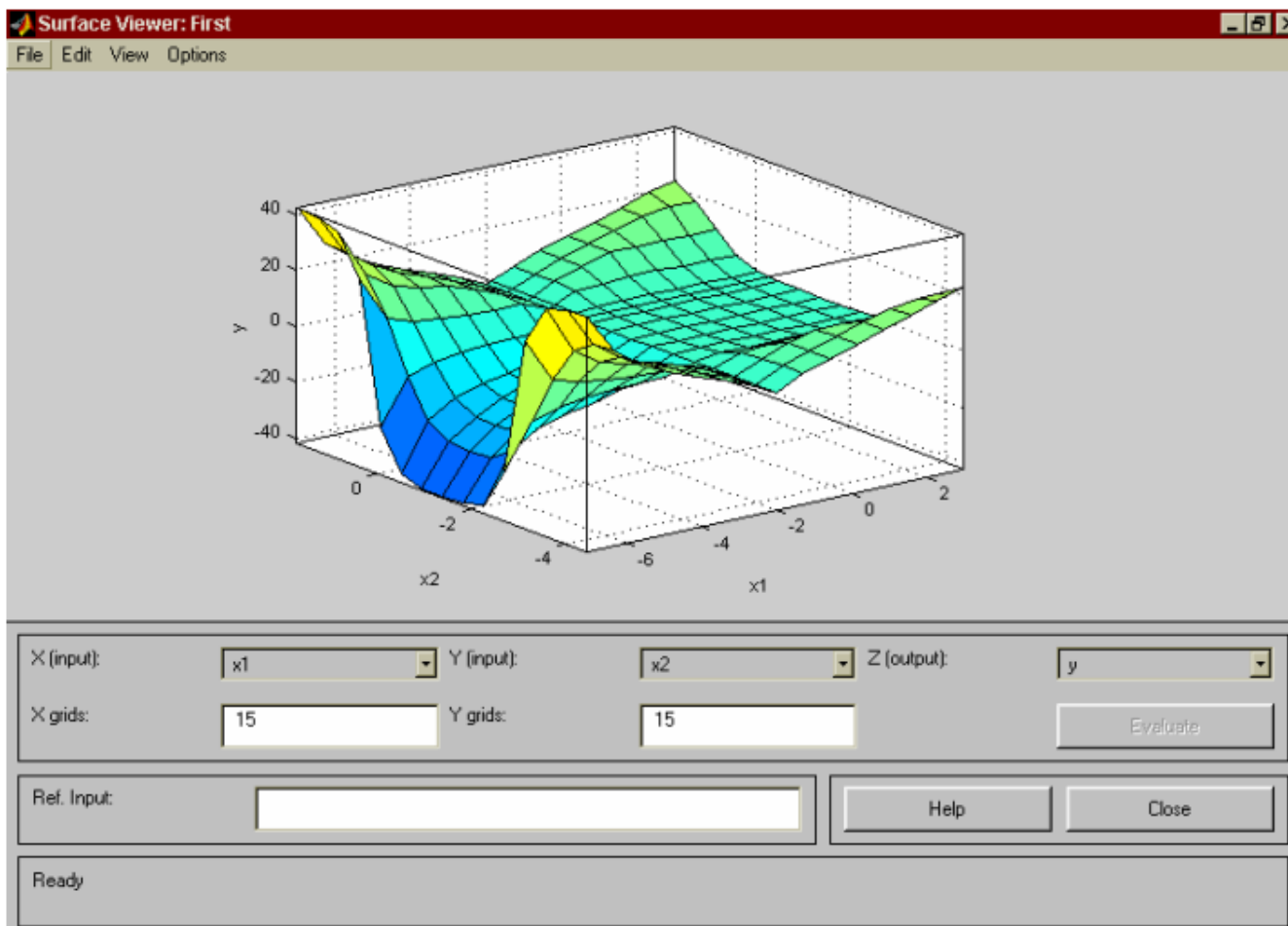


Рис. 8. Поверхность “входы-выход” в окне SurfaceViewer

5. Если x_1 =высокий и x_2 =низкий, то y =выше среднего.
6. Если x_1 =высокий и x_2 =средний, то y =средний.
7. Если x_1 =низкий и x_2 =средний, то y =низкий.
8. Если x_1 =высокий и x_2 =выше среднего, то y =средний.
9. Если x_1 =высокий и x_2 =ниже среднего, то y =средний.

Для ввода правила необходимо выбрать в меню соответствующую комбинацию термов и нажать кнопку **Add rule**. На рис. 6 изображено окно редактора базы знаний после ввода всех девяти правил. Число, приведенное в скобках в конце каждого правила, представляет собой весовой коэффициент соответствующего правила.

Шаг 17. Сохраним созданную систему. Для этого в меню **File** выбираем в подменю **Export** команду **To disk**.

На рис. 7 приведено окно визуализации нечеткого логического вывода. Это окно активизируется командой **View rules...** меню **View**. В поле **Input** указываются значения входных переменных, для которых выполняется логический вывод. На рис. 8 приведена поверхность “входы-выход”, соответствующая синтезированной нечеткой системе. Для вывода этого окна необходимо использовать команду **View surface...** меню **View**. Сравнивая поверхности на рис. 1 и на рис. 8, можно сделать вывод, что нечеткие правила достаточно хорошо описывают сложную нелинейную зависимость.

Лабораторная работа 5

Класстеризация объектов

1. Цель работы

- изучить алгоритмы и методы кластерного и дискриминантного анализов данных на примере решения конкретной задачи ИАД;
- исследовать эффективность использования различных алгоритмов и методов кластерного анализа данных для решения прикладной задачи;
- ознакомиться и получить практические навыки работы с модулями интегрированной статистической системы, реализующими решение задачи классификации объектов.

2. Задание к лабораторной работе

1. Подготовьте файл данных для проведения интеллектуального анализа в соответствии с вариантом.

2. Проведите классификацию объектов, используя алгоритмы и методы кластерного анализа.

2.1. Выполните кластеризацию объектов с помощью иерархических агломеративных методов, используя различные меры расстояния между объектами. В качестве исходных данных для решения задачи возьмите количественные показатели. Для каждого метода необходимо выполнить следующее:

- построить дендрограмму иерархического метода для предложенных данных;
- проанализировать матрицу расстояний и схему объединения, построить график зависимости числа кластеров от коэффициента слияния;

- на основе визуального анализа построенного графика выдвинуть гипотезу о числе кластеров в данных;
- рассчитать оптимальное число кластеров с помощью правила Мойена;
- внести полученные результаты в таблицу

2.2. Выполните кластеризацию объектов на основе использования качественных признаков. Воспользуйтесь одним из предложенных в варианте иерархическим агломеративным методом и процентом несогласия в качестве меры расстояния между объектами. Выполните шаги п. 2.1.

2.3. Проведите кластеризацию объектов, используя метод К-средних и метод карт Кохонена. Выполните следующие шаги:

- задать число кластеров, определенное в пп. 2.1 и 2.2;
- рассчитать евклидово расстояние между кластерами;
- рассчитать значимость влияния каждого из признаков на выделение кластеров на основе F -статистики;
- построить график средних значений признаков по каждому из кластеров;
- определить объекты, относящиеся к каждому кластеру;
- построить диаграмму рассеяния в пространстве трех признаков, которые вносят наибольший вклад в выделение кластеров, на диаграмме выделить полученные кластеры;
- внести полученные результаты в таблицу.

2.4. Сопоставьте полученные результаты кластеризации с заданным разбиением объектов на классы. По результатам анализа заполните таблицу

Таблица

метод	мера расстояния	количество кластеров	количество ошибочно отнесенных к кластеру объектов				достоинства, недостатки метода и меры расстояния
			кластер 1	кластер 2	

2.5. Опишите полученные кластеры в терминах предметной области, дайте каждому кластеру условное название.

2.6. На основе проведенного анализа ответьте на следующие вопросы:

- С помощью какого иерархического метода было наиболее точно определено количество кластеров? Какая при этом использовалась мера расстояния между объектами? Какой метод оказался наименее точным? Объясните полученный результат.
- Какой из иерархических методов дал наиболее точное разбиение объектов на кластеры? Какая при этом использовалась мера расстояния между объектами? Какой метод оказался наименее точным? Объясните полученный результат.
- На основе использования какого типа исходных данных (количественные, качественные) получено более точное разбиение объектов на классы и почему?
- Какие признаки вносят наибольший вклад в разделение объектов на кластеры и почему? Проранжируйте признаки в порядке значимости.
- Какие признаки можно не рассматривать при кластеризации объектов и почему?
- Выделите два кластера объектов наиболее и наименее схожих по совокупности рассматриваемых признаков. Обоснуйте свой выбор.

3. Постройте правило отнесения объектов к одному из выделенных классов (кластеров) на основе методов дискриминантного анализа.

3.1. Постройте функцию классификации объектов, используя разные признаки и их комбинации.

3.2. Рассчитайте матрицу классификации.

3.3. По результатам расчетов п. 3.1-3.2 заполните таблицу

Таблица

признаки, используемые в классификационных функциях	Количество ошибочно отнесенных объектов к классу (кластеру)				% корректно отнесенных объектов
	класс 1	класс 2	

3.4. На основе анализа таблицы сделайте окончательные выводы о структуре классификационных функций (какие признаки необходимо включить в классификационные функции). Обоснуйте выводы.

Лабораторная работа 6

Применение деревьев классификации в решении задач интеллектуального анализа данных (ИАД)

1. Цель работы

- изучить алгоритмы и методы деревьев классификации на примере решения конкретной задачи ИАД;
- исследовать эффективность использования алгоритмов и методов деревьев классификации для решения прикладной задачи;
- ознакомиться и получить практические навыки работы с модулями интегрированной системы, реализующими методы деревьев классификации.

2. Задание к лабораторной работе

Подготовьте исходные данные.

Постройте правило отнесения объектов к одному из заданных классов, используя методы деревьев решений.

1. Постройте деревья решений, используя разные комбинации признаков (показателей), заданных в задаче:

1 вариант построения – все качественные и количественные признаки;

2 вариант построения – количественные признаки;

3 вариант построения - качественные признаки.

При построении дерева решений задайте следующие опции:

- одинаковые цены ошибки классификации и априорные вероятности попадания объекта в тот или иной класс;
- метод полного перебора деревьев с одномерным ветвлением по методу CART;
- меру Джини однородности вершины в качестве критерия согласия;
- правило останова - прямая остановка по методу FACT, доля неклассифицированных объектов – 0,05;
- параметры V – кратной кросс-проверка: V=3.

Оцените сложность и точность построенных деревьев решений, по результатам анализа заполните таблицу.

Таблица.
Показатели качества деревьев решений

номер вар. дерева	признаки	априорные вероятности	правило останова	точность дерева		сложность дерева				
				цена обучения	цена кросс проверки	число листьев	число вершин	максимальн. длина пути	длина внешнего пути	

2. Постройте дерево решений на основе использования качественных признаков. Укажите правило останова «отсечение по ошибке классификации»:

- используйте параметры по умолчанию (*вариант построения 5*): максимальное количество неверно классифицированных объектов – 5; стандартная ошибка правила – 1;

- задайте следующие параметры (*вариант построения дерева б*): максимальное количество неверно классифицированных объектов – 2; стандартная ошибка правила – 0,1.

Остальные опции такие же, как в п. 2.1. Полученные результаты внесите в таблицу

3. Для варианта б. дерева решений постройте:

- графическое представление дерева (в узлах дерева задайте гистограммы);

- матрицу ошибок классификации;

- последовательность оптимально усеченных деревьев и их параметры качества;

- график цены обучения и цены кросс проверки для цепочки последовательно усечаемых деревьев;

- выполните глобальную кросс-проверку (используйте параметры глобальной кросс проверки - по умолчанию), постройте матрицу ошибок классификации по результатам глобальной кросс проверки;

- сделайте вывод об эффективности работы процедуры «автоматического» выбора оптимального варианта дерева.

4. Выберите наилучший по совокупности показателей качества вариант построения дерева решений. Постройте графическое представление дерева для этого варианта.

5. По результатам проведенного исследования ответьте на следующие вопросы.

- На основе каких признаков (качественных, количественных) построено дерево решений, обладающее лучшими показателями точности и сложности?

- Как повлияло изменение априорной вероятности попадания объектов в классы на результаты классификации, на точность и сложность, построенного дерева решений?

- Как повлияло изменение правила останова и параметров правила останова на результаты классификации объектов, на точность и сложность, построенного дерева решений?

6. Проведите кросс проверку выбранного варианта и варианта б. построения дерева решения на тестовой выборке. По результатам анализа заполните таблицу. Оцените прогностическую способность дерева, сделайте выводы.

Таблица.
Оценка прогностической способности деревьев классификации

вариант построения	цена V-кросс проверки	стандартная ошибка цены V-кросс проверки	цена тестовой выборки	стандартная ошибка цены тестовой выборки

7. Сопоставьте результаты классификации объектов тестовой выборки на основе использования методов дискриминантного анализа и методов деревьев решений. Сделайте выводы в терминах решаемой задачи.
8. Сделайте выводы по результатам проведенного исследования в произвольной форме.

Лабораторная работа 7.

Прогнозирование временных рядов

1. Цель лабораторной работы

- изучить методы и алгоритмы прогнозирования временных рядов на примере решения конкретной задачи ИАД;
- исследовать эффективность использования различных методов прогнозирования временных рядов для решения прикладной задачи;
- ознакомиться и получить практические навыки работы с модулями интегрированной статистической системы, реализующими решение задачи прогнозирования временных рядов.

2. Задание к лабораторной работе

2.1. Подготовьте исходные данные для проведения интеллектуального анализа.

2.2. Постройте линейный график временного ряда. На основе визуального анализа графика сделайте предварительные выводы о структуре временного ряда:

- наличие тренда; характер основной тенденции (монотонность; существование вертикальных и/или горизонтальных асимптот; рост (спад) уровней ряда с течением времени); тип функции тренда (линейная, нелинейная);
- наличие сезонной составляющей и характер сезонной составляющей (периодичность; амплитуда колебаний; постоянство (изменчивость) амплитуды колебаний с течением времени).

2.3. Определите структурную модель тренда временного ряда с помощью метода характеристик прироста:

- сгладьте временной ряд, используя пятимесячную скользящую среднюю;
- определите средние приросты;
- определите производные характеристики прироста: \bar{u}_t ; $\bar{u}_t^{(2)}$; \bar{u}_t / \bar{y}_t ; $\log \bar{u}_t$; $\log(\bar{u}_t / \bar{y}_t)$; $\log(\bar{u}_t / \bar{y}_t^2)$;
- постройте линейные графики производных характеристик прироста.

На основе анализа характеристик прироста определите две наиболее вероятные структурные модели тренда.

2.4. Идентифицируйте параметры выбранных структурных моделей тренда. Рассчитайте характеристики точности прогнозных моделей, заполните табл. 2.

2.5. По результатам расчетов (табл. 2) сделайте окончательный вывод относительно вида модели тренда. Постройте график исходного временного ряда с наложенной прогнозной моделью тренда.

2.6. Определите структурную модель сезонной составляющей ряда:

- постройте и проанализируйте периодограмму временного ряда;
- постройте структуру периодической гармонической функции.

2.7. Идентифицируйте параметры сезонной составляющей ряда. Рассчитайте характеристики точности прогнозной модели, содержащей тренд и сезонную составляющую, заполните таблицу. Постройте график исходного временного ряда с наложенной прогнозной моделью.

Таблица .
Характеристики точности прогнозных моделей

Номер варианта модели	Прогнозная модель	Остатки				сумма квадратов отклонений	коэффициент детерминации
		миним.	максим.	средний	СКО		

2.8. Постройте автокорреляционную и частную автокорреляционную функции остатков прогнозной модели, построенной в п. 2.7. Сделайте вывод о наличии (отсутствии) автокорреляции в остатках и необходимости учета авторегрессионной составляющей в прогнозной модели ряда.

2.9. Определите структуру и параметры авторегрессионной составляющей ряда (в случае необходимости). Рассчитайте характеристики точности прогнозной модели, содержащей тренд, сезонную и авторегрессионную составляющие временного ряда.

2.10. Проанализируйте таблицу, выберите окончательный вариант прогнозной модели, обоснуйте свой выбор.

2.11. Для выбранного варианта прогнозной модели постройте гистограмму остатков и проверьте гипотезу о согласии распределения остатков с моделью нормального распределения, постройте автокорреляционную и частную автокорреляционную функции остатков.

2.12. Сделайте выводы об адекватности построенной прогнозной модели данным наблюдения.

2.13. Дайте содержательную интерпретацию полученных результатов. Опишите составляющие прогнозной модели в терминах решаемой задачи.

2.14. На основании построенной модели осуществите прогноз временного ряда на 3 временных интервала вперед. Сделайте выводы в терминах решаемой задачи.

2.15. По результатам проведенного исследования сделайте выводы в свободной форме.

20.2 Промежуточная аттестация

Промежуточная аттестация выставляется на основе выполненных лабораторных работ

Для оценивания результатов обучения на зачете используются следующие показатели:

- 1) знание учебного материала и владение понятийным аппаратом по курсу «Интеллектуальный анализ в экономических исследованиях»;
- 2) знание современных направлений исследований в области искусственного интеллекта;
- 3) владение технологиями планирования экспериментов в области решения прикладных задач с помощью интеллектуальных систем
- 4) владение навыками использования пакета MATLAB для решения задач;
- 5) владение навыками работы с инструментами Data Mining;
- 6) владение навыками написания отчетов по результатам проведенного эксперимента;
- 7) владение навыками интерпретации полученных результатов в терминах прикладной области с целью получения новых знаний и выводов.

Для оценивания результатов обучения на зачете используется 4-балльная шкала: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Соотношение показателей, критериев и шкалы оценивания результатов обучения.

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся в полной мере владеет понятийным аппаратом данной области науки (теоретическими основами дисциплины), сдал все лабораторные работы.	Повышенный уровень	Отлично
Обучающийся владеет понятийным аппаратом данной области науки (теоретическими основами дисциплины), но не сдал одну лабораторную работу.	Базовый уровень	Хорошо
Обучающийся демонстрирует неуверенное владение понятийным аппаратом данной области науки (теоретическими основами дисциплины), не сдал две лабораторные работы.	Пороговый уровень	Удовлетворительно
Обучающийся демонстрирует отрывочные, фрагментарные знания, допускает грубые ошибки, не сдал более двух лабораторных работ.	–	Неудовлетворительно