

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ ИМЕНИ
И.Т. ТРУБИЛИНА»**

Факультет прикладной информатики

Кафедра Компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ покупателей магазина»

Выполнил студент группы: ИТ2342 Лисовин Олег Александрович

\ Допущен к защите

Руководитель проекта д.э.н., к.т.н., профессор Луценко Е.В. (

(подпись, расшифровка подписи)



Защищен

(дата)

Оценка

Краснодар

2024

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ
УНИВЕРСИТЕТ ИМЕНИ И. Т. ТРУБИЛИНА»

Факультет прикладной информатики

РЕЦЕНЗИЯ

на курсовую работу

Студента Лисовина Олега Александровича курса

2 очной формы обучения группы ИТ2342

Направление подготовки 09.04.02 «Информационные системы и технологии»

Наименование темы «АСК-анализ покупателей магазина»

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор

(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)
1	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
2	Уровень и корректность использования в работе различных методов исследований	5
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
4	Ясность, четкость, последовательность и обоснованность изложения	5
5	Применение современных технологий обработки информации	5
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	5
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	5
8	Ответы на вопросы при защите	5

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____

Рецензент  (Е. В. Луценко)

«27» декабря 2024 г.

РЕФЕРАТ

Курсовая работа содержит: 86 страниц, 40 рисунков, 18 таблиц, 24 литературных источника.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, ШКАЛЫ, КЛАССЫ.

Целью работы является проведение автоматизированного системно-когнитивного анализа покупателей магазина.

Для достижения цели требуется проанализировать методы создания обобщенных представлений классов и решения задач идентификации объектов с применением методов принятия решений и изучения моделируемой области путем анализа модели.

Автоматизированный системно-когнитивный анализ покупателей магазина

Лисовин Олег Александрович
студент факультета ПИ, группы
ИТ2342
o.lisowin@yandex.ru

Кубанский государственный аграрный университет имени И.Т.Трубилина, Краснодар, Россия

Целью данной работы является изучение характеристик покупателей магазина. Достижение данной цели представляет большой личный интерес. Для меня это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», а также обеспечить некоторую научную ценности в сфере изучения корреляции покупателей в магазине. Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

Ключевые слова:

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, АСКАНАЛИЗ, СИСТЕМА «ЭЙДОС».

Automated system-cognitive analysis of store customers

Lisovin Oleg Alexandrovich
student of the faculty of PI, group
IT2342
o.lisowin@yandex.ru

Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia

The purpose of this work is to study the characteristics of the store's customers. Achieving this goal is of great personal interest. For me, this will allow me to gain knowledge in working with the universal cognitive analytical system "Eidos-X++", as well as provide some scientific value in the field of studying correlations of customers in the store. To achieve this goal, Automated Cognitive System Analysis (ASC-analysis) and its software toolkit, the intelligent Eidos system, are used.

Key words: AUTOMATED SYSTEM-COGNITIVE ANALYSIS, ASC-ANALYSIS, "EIDOS" SYSTEM.

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	6
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	6
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	6
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	6
1.4. ЦЕЛЬ РАБОТЫ.....	7
2. METHODS (МЕТОДЫ)	7
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	7
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	8
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	8
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА.....	10
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	18
3. RESULTS (РЕЗУЛЬТАТЫ)	20
3.1. Задача-1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций	20
3.2. Задача-2. Формализация предметной области	21
3.3. Задача-3. Синтез статистических и системно-когнитивных моделей. Многопараметрическая типизация и частные критерии знаний	31
3.4. Задача-4. Верификация моделей	41
3.5. Задача-5. Выбор наиболее достоверной модели	44
3.6. Задача-6. Системная идентификация и прогнозирование	46
3.6.1. <i>Интегральный критерий «Сумма знаний»</i>	46
3.6.2. <i>Интегральный критерий «Семантический резонанс знаний»</i>	47
3.6.3. <i>Важные математические свойства интегральных критериев</i>	48
3.6.4. <i>Решение задачи идентификации и прогнозирования в системе «Эйдос»</i>	49
3.7. Задача-7. Поддержка принятия решений	52
3.7.1. <i>Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ</i>	52
3.7.2. <i>Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»</i>	54
3.8. Задача-8. Исследование объекта моделирования путем исследования его модели	58
3.8.1. <i>Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)</i>	58
3.8.2. <i>Кластерно-конструктивный анализ классов</i>	60
3.8.3. <i>Кластерно-конструктивный анализ значений описательных шкал</i>	61
3.8.4. <i>Модель знаний системы «Эйдос» и нелокальные нейроны</i>	64
3.8.5. <i>Нелокальная нейронная сеть</i>	66
3.8.6. <i>3D-интегральные когнитивные карты</i>	68
3.8.7. <i>2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения)</i>	69
3.8.8. <i>2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения)</i>	72
3.8.9. <i>Когнитивные функции</i>	75
3.8.10. <i>Значимость описательных шкал и их градаций</i>	78
3.8.11. <i>Степень детерминированности классов и классификационных шкал</i>	81
4. DISCUSSION (ОБСУЖДЕНИЕ)	83
5. CONCLUSIONS (ВЫВОДЫ)	84
REFERENCES (ЛИТЕРАТУРА)	85

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

В современном мире розничная торговля переживает значительные трансформации, обусловленные развитием технологий и изменением потребительских предпочтений. Для достижения успеха на конкурентном рынке важно не только предлагать качественные товары, но и понимать потребности, интересы и поведение покупателей. Это позволяет более точно формировать ассортимент, разрабатывать стратегии продвижения и предлагать персонализированные решения, что, в свою очередь, способствует повышению лояльности клиентов и увеличению прибыли.

Одним из эффективных подходов к изучению потребительского поведения является АСК-анализ (анализ структуры клиентов), который позволяет сегментировать аудиторию на основе таких факторов, как пол, возраст и доход. Эти параметры оказывают значительное влияние на выбор товаров, предпочтения в ценовой категории и восприятие маркетинговых сообщений.

Целью данного исследования является проведение АСК-анализа покупателей магазина для выявления ключевых закономерностей в поведении клиентов. Основное внимание будет уделено изучению влияния пола, возраста и дохода на структуру покупательской аудитории, а также определению их роли в формировании потребительских предпочтений.

В рамках исследования предполагается:

1. Сбор и обработка данных о клиентах магазина с учетом пола, возраста и уровня дохода.
2. Анализ полученных данных для выявления основных тенденций и закономерностей.
3. Формирование рекомендаций для оптимизации работы магазина, улучшения ассортимента и разработки персонализированных маркетинговых стратегий.

Результаты исследования могут быть полезны для повышения эффективности управления магазином, разработки индивидуальных предложений для различных групп клиентов и формирования устойчивого конкурентного преимущества.

1.2. Объект и предмет исследования

Объект исследования (моделирования) – покупатели в магазине.

Предмет исследования – выявление корреляции в возрасте, поле и доходе покупателей магазина.

1.3. Проблема, решаемая в работе и ее актуальность

Современные магазины сталкиваются с необходимостью глубокого анализа покупательского поведения для повышения эффективности своей деятельности. Учитывая разнообразие клиентов по полу, возрасту и

уровню дохода, возникают трудности в сегментации аудитории и разработке персонализированных предложений.

Проблема заключается в недостаточной систематизации данных о покупателях и недостаточном понимании влияния демографических характеристик на их поведение. В условиях высокой конкуренции магазины, не уделяющие должного внимания анализу своей клиентской базы, рискуют потерять долю рынка и снизить уровень лояльности клиентов.

АСК-анализ, включающий изучение взаимосвязей между полом, возрастом и доходом покупателей, позволяет выявить закономерности, влияющие на их предпочтения. Например, можно определить, какие товарные категории наиболее популярны среди женщин или мужчин, как изменяются предпочтения покупателей в зависимости от возрастной группы, и как доход влияет на выбор продукции.

Актуальность данного исследования обусловлена необходимостью разработки стратегий персонализации для повышения конкурентоспособности магазина. Результаты анализа могут быть использованы для создания более точных маркетинговых кампаний, оптимизации ассортимента и повышения удовлетворенности клиентов.

1.4. Цель работы

Целью работы является изучение и анализ структуры клиентской базы магазина с использованием АСК-анализа для выявления влияния таких факторов, как пол, возраст и доход, на поведение покупателей.

Для достижения поставленной цели в работе предполагается решить ряд задач, связанных с выбором и применением методов анализа данных, обработкой информации о клиентах и интерпретацией полученных результатов. Конкретные задачи и этапы исследования будут сформулированы после обоснования методики, применяемой для решения проблемы.

2. METHODS (МЕТОДЫ)

2.1. Обоснование требований к методу решения проблемы

Из специфики поставленной проблемы сопоставимости обработки в одной модели исходных, представленных в разных типах шкал числовых и текстовых (лингвистических) и в разных единицах измерения, вытекают следующие **требования** к методу решения проблемы:

1. Метод должен обеспечивать устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных (неточных) взаимозависимых (нелинейных) данных большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

2. Метод решения проблемы не должен предъявлять чрезмерно жестких требований к исходным данным, которые практически невозможно выполнить, а должен обеспечивать обработку тех данных, которые реально есть.

3. Метод должен не теоретически, а реально на практике решать поставленную проблему, а значит, он должен иметь поддерживающий его программный инструментарий, находящийся в полном открытом бесплатном доступе.

2.2. Литературный обзор методов решения проблемы, их характеристика и оценка степени соответствия обоснованным требованиям

Поиск в Internet математических методов и реализующих их программных систем, *одновременно* удовлетворяющих всем требованиям, обоснованным в п.2.1 данной работы показал, что альтернатив Автоматизированному системно-когнитивному анализу и его программному инструментарии – системе «Эйдос» в настоящее время здесь практически нет [1-4].

2.3. Автоматизированный системно-когнитивный анализ (АСК-анализ) как метод решения проблемы

Автоматизированный системно-когнитивный анализ (АСК-анализ) предложен *проф.Е.В.Луценко* в 2002 году в ряде статей 1997-2001 годов¹ и фундаментальной монографии [2].

Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф.Е.В.Луценко в 2001 году. На тот момент этот термин вообще не встречался в Internet. Сегодня по соответствующему запросу в Гугле находится около 23000 сайтов с этим сочетанием слов².

***Примечание:** Ниже приведено очень краткое описание АСК-анализа и системе «Эйдос». Это описание может выглядеть как нескромность и самовосхваление. Но автор просит читателей понять его правильно. Это сделано исключительно для тех довольно многочисленных читателей, которые впервые слышат об этом методе и системе.*

АСК-анализ включает:

- теоретические основы, в частности базовую формализуемую когнитивную концепцию;
- математическую модель, основанную на системном обобщении теории информации (СТИ);
- методику численных расчетов (структуры баз данных и алгоритмы их обработки);

¹ <http://lc.kubagro.ru/aidos/Sprab0802.pdf> (см. с публикации № 48).

² [https://yandex.ru/search/?text=Автоматизированный%2Bсистемно-когнитивный%2Bанализ%2B\(АСК-анализ\)&lr=35&clid=2327117-18&win=360](https://yandex.ru/search/?text=Автоматизированный%2Bсистемно-когнитивный%2Bанализ%2B(АСК-анализ)&lr=35&clid=2327117-18&win=360)

– программный инструментарий, в качестве которого в настоящее время выступает универсальная когнитивная аналитическая система «Эйдос» (интеллектуальная система «Эйдос»).

Более подробно математический метод АСК-анализа описан в работе [3] и ряде других [5]. Около половины из 688 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано более 40 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 33 патента РФ на системы искусственного интеллекта, 360 публикации в изданиях, входящих в перечень ВАК РФ и приравненных им (по данным [РИНЦ](#)), 17 публикаций в изданиях, входящих в ядро РИНЦ, 6 статей в журналах, входящих в [WoS](#), 7 публикаций в журналах, входящих в [Скопус](#)³ [5, 6, 7].

Три монографии включены в фонды библиотеки конгресса США⁴.

АСК-анализ и система "Эйдос" были успешно применены в 8 докторских диссертациях по экономическим, техническим, биологическим наукам и в 8 кандидатских диссертациях по экономическим, техническим, психологическим и медицинским наукам, еще несколько докторских и кандидатских диссертаций с применением АСК-анализа находятся в стадии выхода на защиту.

Автор является основателем междисциплинарной научной школы: «Автоматизированный системно-когнитивный анализ»⁵. Научная школа: "Автоматизированный системно-когнитивный анализ" является междисциплинарным научным направлением на пересечении по крайней мере трех научных специальностей (согласно недавно утвержденной новой номенклатуры научных специальностей ВАК РФ⁶). Основные научные специальности, которым соответствует научная школа:

- 5.12.4. Когнитивное моделирование;
- 1.2.1. Искусственный интеллект и машинное обучение;
- 2.3.1. Системный анализ, управление и обработка информации.

Научная школа: "Автоматизированный системно-когнитивный анализ" включает следующие междисциплинарные научные направления:

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных;
- Автоматизированный системно-когнитивный анализ текстовых данных;
- Спектральный и контурный автоматизированный системно-когнитивный анализ изображений;

³ <http://lc.kubagro.ru/aidos/Sprab0802.pdf>

⁴ <https://catalog.loc.gov/vwebv/search?searchArg=Lutsenko+E.V.> (и кликнуть: "Search")

⁵ <https://www.famous-scientists.ru/school/1608>

⁶ <https://www.garant.ru/products/ipo/prime/doc/400450248/>

– Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов.

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [5] и страничка в РесечГейт [6] и РИНЦ [7], на которых можно получить более полную информацию о методе АСК-анализа и системе «Эйдос». Краткая информация об АСК-анализе и системе «Эйдос» есть в материале: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf.

Решение поставленной в работе проблемы сопоставимости в АСК-анализе и системе «Эйдос» обеспечивается *путем* метризации номинальных шкал, т.е. повышения их степени формализации до уровня числовых шкал [8]. Сама метризация номинальных шкал достигается *путем* вычисления количества информации, содержащегося в градациях номинальных шкал о получении той или иной урожайности [8]. Для работы с лингвистическими переменными применяются стандартные возможности АСК-анализа [5].

2.4. Система «Эйдос» – инструментарий АСК-анализа

Конечно, на системе «Эйдос» как говорят «Свет клином не сошелся». Существует много очень достойных систем искусственного интеллекта. Чтобы лично убедиться в этом достаточно самостоятельно осуществить поиск в Internet, просто посмотреть файлы: [NCKR-1](#), [NCKR-2](#), [NCKR-3](#), [NCKR-4](#) или пройти по ссылкам: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/how-does-mldotnet-work>, <http://chat.openai.com/>, <https://poe.com/>, <https://neural-university.ru/>, <https://dzen.ru/a/ZCKZRKvrIEMBWOк8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://rudalle.ru/>, еще очень много отличных нейросетей: <https://problembo.com/ru/services> (и это здесь может пригодиться - почта на 10 минут: <https://10minutemail.net/>).

Полезные нейросети и приложения для разных сфер:

🌀 Для дизайнеров: SiteKick - нейросеть для создания лендингов; AdCreative - делает рекламные креативы, плакаты; Looka - логотипы по текстовому описанию; Watermarkremover - поможет удалить вотермарки; Booth ai - создает стоковые фотки по описанию; PatternedAI - паттерны по текстовому описанию; Nama - вырезать лишние элементы с фото или картинки; RoomGPT - «примеряет» новый ремонт на вашу квартиру, помогает выбрать дизайн;

📷 Для фотографов: ; Pallete fm - раскрашивает черно-белые изображения; Relight - меняет светотень на фотографиях; Photoroom - вырезать элементы из фото, поменять фон; LeiaPix - сделает из 2D-фотки 3D.; Nostalgia Photo - улучшает качество старых фото; pfpmaker - генератор

аватарок для соцсетей; Picsart - заменяет или удаляет ненужные элементы на фото;

🎬 Для тех, кто монтирует видео:; CapCut - удобный редактор, доступен в браузере. Есть цветокорр, разные эффекты; vidyo ai - нарезать видео на короткие фрагменты; Reface - изменить лицо человека на видео; Runwayml - самые разные инструменты для монтажа; Colourlab AI - нейросеть для цветокоррекции; Topaz Video AI - сильно улучшит качество видео, уберет шум и трясущийся экран; Luma AI - сделает 3D изображение из серии фото; Simplified - анимация картинки; SpiritMe - твоя цифровая копия в сети;

🎵 Для звукарей и музыкантов; ; Mubert - создаёт музыку любого жанра; Beatoven - ИИ-композитор музыки для видео; Clip audio - подберет музыку для любого видоса; Fadr - порежет трек на отдельные дорожки инструментов и вокала; Adobe Enhance - чистит запись от шумов. Бесплатно; Elevenlabs - мощнейший синтезатор, подделает любой голос; The MetaVoice - меняйте свой голос на один из восьми пресетов; Cleanvoice - уберет из вашей разговорной записи мусор; ;

📄 Для айтишников; ; CodePal - пишет код с нуля, исправляет ошибки, оценивает готовый код; Codesnippets - создает код по текстовому запросу; Buildt AI - поисковик для VSCode, найдет готовый код в инете; Code GPT - плагин-генератор кода для VSCode; Autobackend - автоматический бэкэнд; Adrenaline - ищет и помогает чинить ошибки в коде; Tabnine - дописывает код, если у тебя не получается; ;

📖 Для школьников и студентов; ; Consensus - база научных статей; ExamCram - превратит сложные учебные материалы в карточки и тесты для самопроверки; MathGPT - решает задачи по математике; editGPT - исправляет ошибки в английском ; Yip - то же самое, но в вебе и с поддержкой Википедии; ChatBA - делает презентации за тебя; YouTube Summary with ChatGPT - конвертирует видео или лекции в текст; Explain Me Like I'm Five - объясняет сложные научные термины простым языком;

✂ Для тех, кто ищет работу:; InterviewGPT AI - задает каверзные вопросы и помогает готовиться к собеседованию; Resume Worded - улучшает резюме; kickresume - сделает крутое резюме и напишет мотивационное письмо; Cover Letter AI - написать сопроводительный текст к резюме; ;

🔍 Для тех, кому не помог Гугл:; Chord - напишет реферат в ответ на запрос в строке; Lexii ai - бот, который умеет ссылаться на источники; Perplexity - нейросеть-поисковик в виде расширения для браузера; Nuclia - поиск по облаку или серверу; Phind - умеет искать код, поможет айтишникам; ;

🎧 Для отдыха и развлечения:; RadioGPT - радио, где музыку генерируют нейронки; EndlessVN - бесконечная визуальная новелла;

Natural Language Playlist - подберет плейлист на 7 часов специально для тебя; Movie Deep Search - найдет фильм по запросу; FashionAdvisor AI - советы от нейро-стилиста; Hello History - с помощью нее пообщаешься с историческим персонажем; Cool Gift Ideas - выберет подарок для человека по его описанию; Endel - нейро-музыка, которая помогает засыпать; PlaylistAI - соберет плейлист в Apple и Spotify по тексту или картинке.; Tattoos AI - делает эскизы для татуировок.

И все не смотря на настоящую революцию в области искусственного интеллекта и связанный с ней бум Универсальная когнитивная аналитическая система «Эйдос-X++» отличается от большинства из этих систем, по крайней мере, некоторыми из следующих своих параметров:

- является универсальной и может быть применена во многих предметных областях, т.к. разработана в универсальной постановке, не зависящей от предметной области (<http://lc.kubagro.ru/aidos/index.htm>) и имеет 6 автоматизированных программных интерфейсов (API) ввода данных из внешних источников данных различных типов: таблиц, текстов и графики. Система «Эйдос» является *автоматизированной* системой, т.е. предполагает непосредственное участие человека в реальном времени в процессе создания моделей и их использования для решения задач идентификации, прогнозирования, принятия решений и исследования предметной области путем исследования ее модели (*автоматические* системы работают без такого участия человека);

- находится в полном открытом бесплатном доступе (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), причем с актуальными исходными текстами (<http://lc.kubagro.ru/AidosALL.txt>): открытая лицензия: CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>), и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора и разработчика системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 32 свидетельства РосПатента РФ);

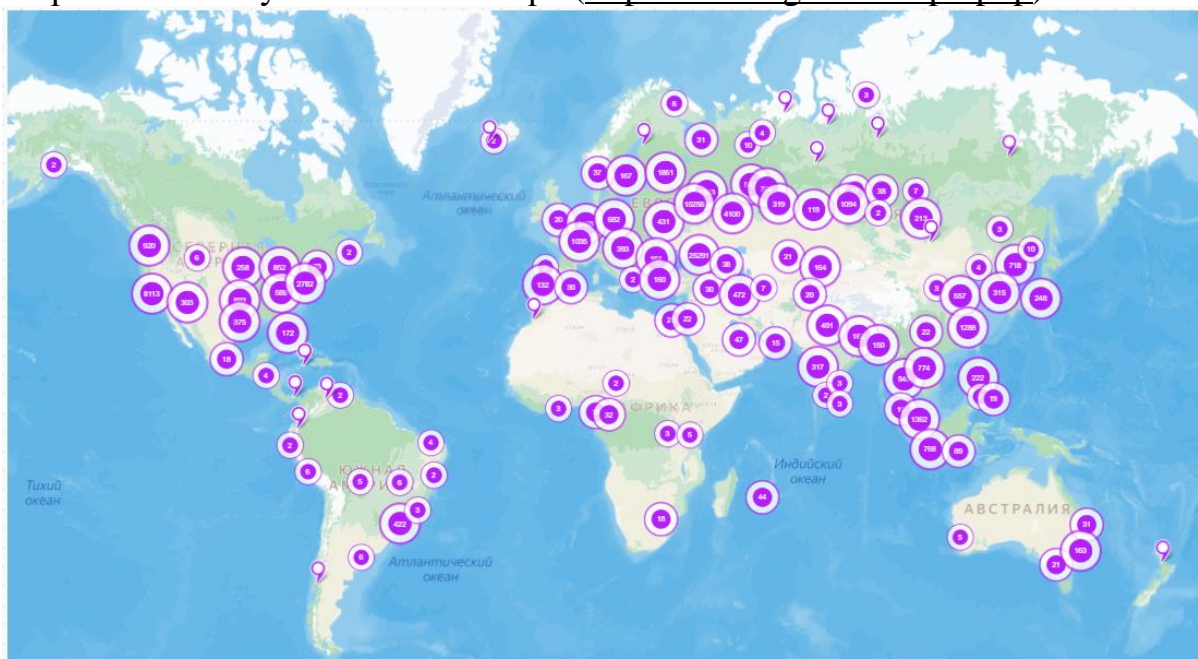
- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта: «имеет нулевой порог входа» (есть акт внедрения системы «Эйдос» 1987 года) (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);

- реально работает, обеспечивает *устойчивое* выявление в *сопоставимой* форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких

требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

- имеет «нулевой порог входа», содержит большое количество локальных (поставляемых с инсталляцией) и облачных интеллектуальных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 390, соответственно: http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf, http://lc.kubagro.ru/Presentation_LutsenkoEV.pdf);

- поддерживает on-line среду накопления знаний и обмена ими, широко используется во всем мире (<http://lc.kubagro.ru/map5.php>):



- обеспечивает мультиязычную поддержку интерфейса на 51 языке. Языковые базы входят в инсталляцию и могут пополняться в автоматическом режиме;

- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA);

- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач классификации, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе: http://lc.kubagro.ru/aidos/aidos18_LLS/aidos18_LLS.pdf);

- хорошо имитирует человеческий стиль мышления: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции;

- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) предлагает без какой-либо предварительной обработки осмыслить эти данные и тем самым преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

В чем сила подхода, реализованного в системе Эйдос? В том, что она реализует подход, эффективность которого не зависит от того, что мы думаем о предметной области и думаем ли вообще. Она формирует модели непосредственно на основе эмпирических данных, а не на основе наших представлений о механизмах реализации закономерностей в этих данных. Именно поэтому Эйдос-модели эффективны даже если наши представления о предметной области ошибочны или вообще отсутствуют.

В этом и слабость этого подхода, реализованного в системе Эйдос. Модели системы Эйдос - это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах.

We are briefly describing a new innovative method of artificial intelligence: Automated system-cognitive analysis (ASC-analysis), which has its own software tools – intelligent system called "Eidos" (open source software).

В разработке системы «Эйдос» были следующие этапы:

1-й этап, «подготовительный»: 1979-1992 годы. Математическая модель системы "Эйдос" разработана в 1979 и впервые прошла экспериментальную апробацию в 1981 году (первый расчет на компьютере на основе модели). С 1981 по 1992 система "Эйдос" неоднократно реализовалась на платформе Wang (на компьютерах Wang-2200C). В 1987 году впервые получен акт внедрения на одну из ранних версий системы «Эйдос», реализованную в среде персональной технологической системы «Вега-М» разработки автора (см.2-й акт внедрения).

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBM-совместимых персональных компьютеров система "Эйдос" впервые

реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены свидетельства РосПатента, первые в Краснодарском крае и, возможно, в России на системы искусственного интеллекта (слева приведена титульная видеogramма финальной DOS-версии системы «Эйдос-12.5», июнь 2012 года). С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

3-й этап, «эра MS Windows xp, 8, 7»: 2012-2020 годы. С июня 2012 по 14.12.2020 система «Эйдос» развивалась на языке Аляска-1.9 + Экспресс++ + библиотека для работы с Internet xh2net. Система «Эйдос-X1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

4-й этап, «эра MS Windows-10»: 2020-2021 годы. С 13.12.2020 года по настоящее время система «Эйдос» развивается на языке Аляска-2.0 + Экспресс++. Библиотека xh2net в ней больше не используется, т.к. все возможности работы с Internet входят в базовые возможности языка программирования.

5-й этап, «эра Больших данных, информации и знаний»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке xBase++eXpress++Advantage Database Server (ADS), обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge), а также на языке Питон (Python). Практически все новые режимы системы «Эйдос» и новые реализации старых режимов будут осуществляться на языке Питон.

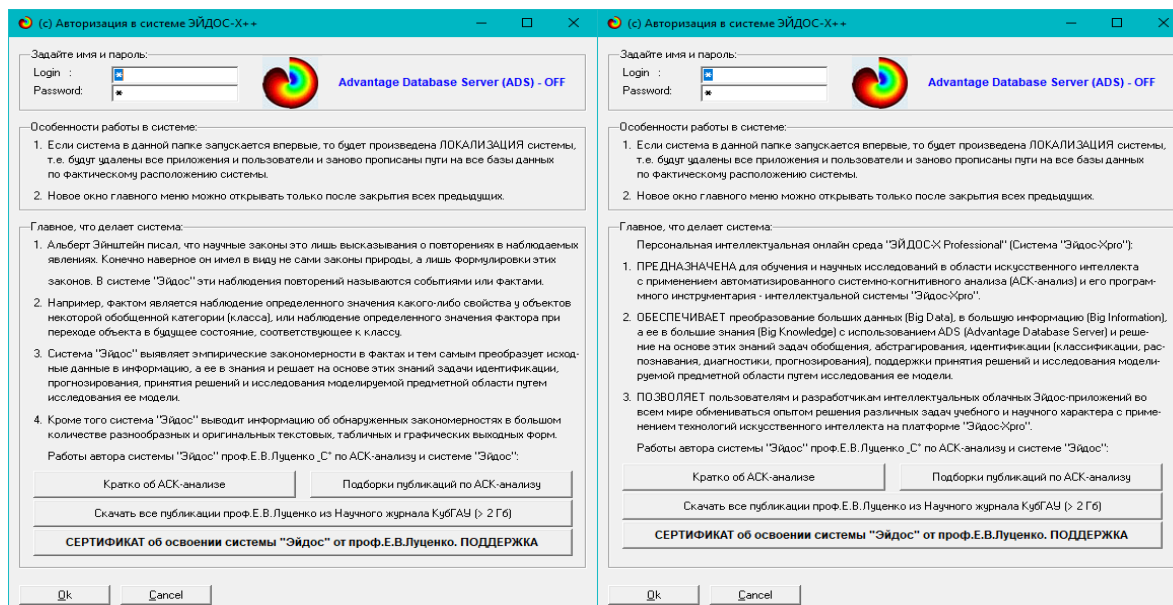
Скачать и запустить систему «Эйдос-X++» (самую новую на текущий момент версию) или обновление системы до текущей версии. Это наиболее полная на данный момент незащищенная от несанкционированного копирования портативная (portable) версия системы (не требующая инсталляции) с полными исходными текстами текущей версии (за исключением ключей доступа к ftp-серверу системы «Эйдос» и ключей лицензионного программного обеспечения), находящаяся в полном открытом бесплатном доступе (около 180 Мб). Обновление имеет объем около 10 Мб. Кредо. Лаборатория в ResearchGate по АСК-анализу и системе «Эйдос».

Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения⁷

На рисунке 1 приведена титульная видеोगрамма DOS-версии системы «Эйдос», а на рисунке 2 – текущей версии системы «Эйдос»:



Рисунок 1. Титульная видеोगрамма DOS-версии системы «Эйдос» (до 2012 года)⁸



⁷ <http://lc.kubagro.ru/aidos/How to make your own cloud Eidos-application.pdf>

⁸ http://lc.kubagro.ru/pic/aidos_titul.jpg

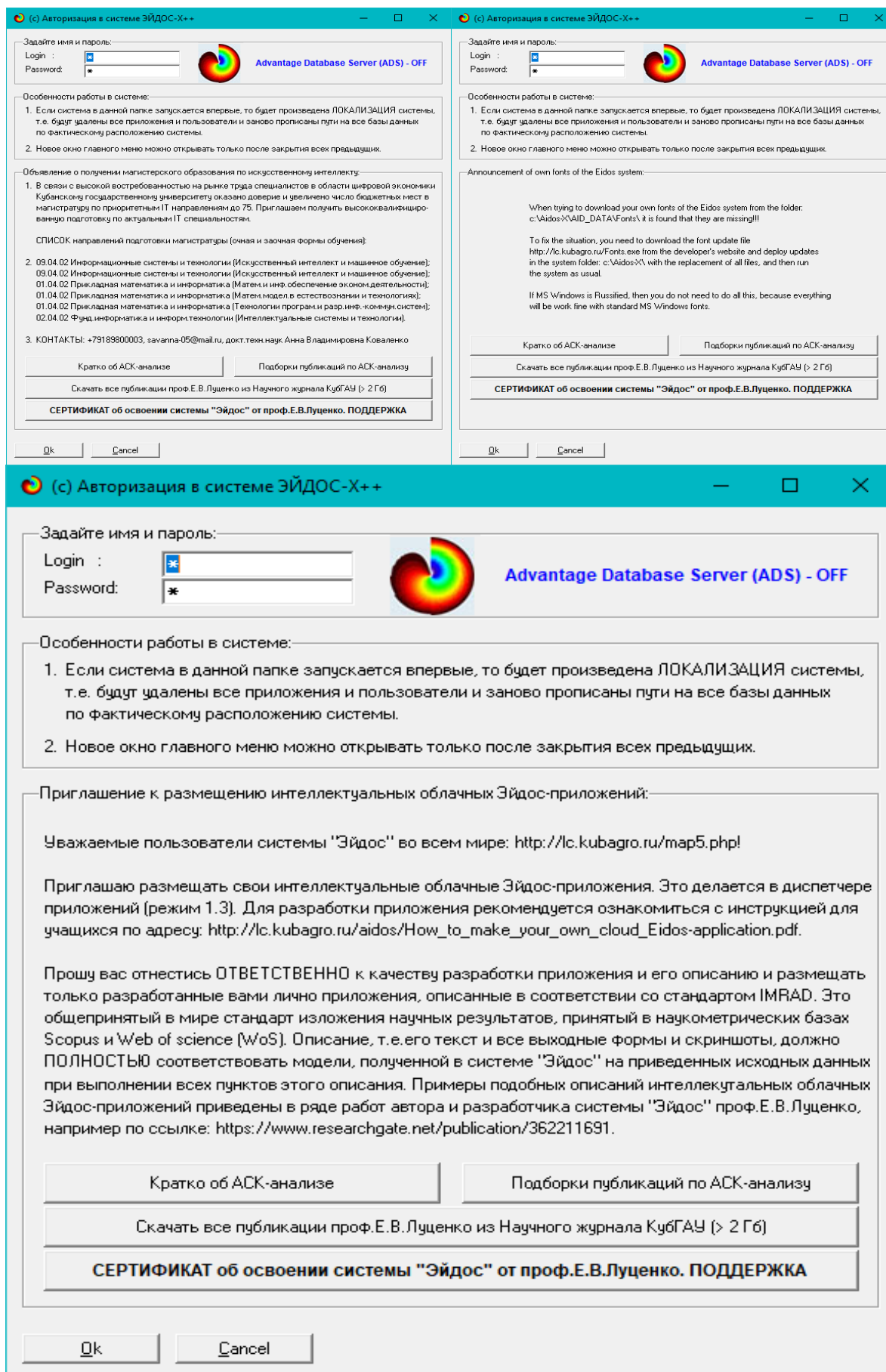


Рисунок 2. Титульные видеограммы текущей версии системы «Эйдос»

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели в АСК-анализе обеспечивается решением следующих *задач* и подзадач, которые получаются в результате декомпозиции цели и являются *этапами* ее достижения:

Задача-1. Когнитивная структуризация предметной области.

Задача-2. Формализация предметной области.

Задача-3. Синтез статистических и системно-когнитивных моделей.

Многопараметрическая типизация и частные критерии знаний.

Задача-4. Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

Задача-6. Системная идентификация и прогнозирование.

Задача-7. Поддержка принятия решений (упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ; развитый алгоритм принятия решений в АСК-анализе).

Задача-8 исследование объекта моделирования путем исследования его модели, *включает ряд подзадач:*

8.1) инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы);

8.2) кластерно-конструктивный анализ классов;

8.3) кластерно-конструктивный анализ значений описательных шкал;

8.4) модель знаний системы «Эйдос» и нелокальные нейроны;

8.5) нелокальная нейронная сеть;

8.6) 3d-интегральные когнитивные карты;

8.7) 2d-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения);

8.8) 2d-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения);

8.9) когнитивные функции;

8.10) значимость описательных шкал и их градаций;

8.11) степень детерминированности классов и классификационных шкал).

Для данной работы особое значение имеет решение подзадачи 8.1, т.к. она позволяет детально исследовать влияние каждого значения каждого фактора на характеристики покупателей.

На рисунке 3 приведена последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»:

**Последовательность обработки данных, информации и знаний в системе «Эйдос»,
повышение уровня системности данных, информации и знаний,
повышение уровня системности моделей**

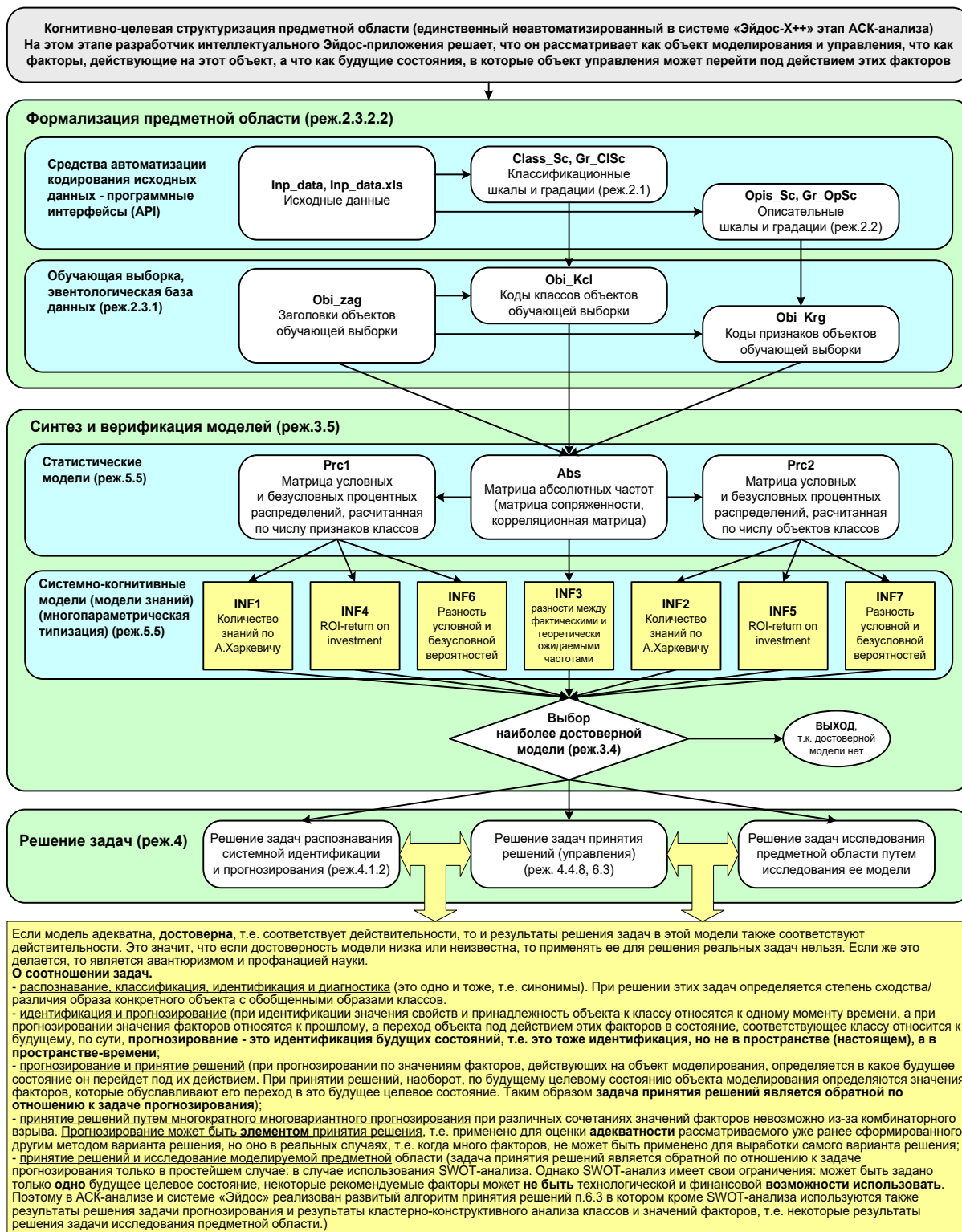


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»

3. RESULTS (РЕЗУЛЬТАТЫ)

3.1. Задача-1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций

На этапе когнитивно-целевой структуризации предметной области мы неформализуемым путем решаем на качественном уровне, что будем рассматривать в качестве объекта моделирования, что в качестве факторов, действующих на моделируемый объект (причин), а что в качестве результатов действия этих факторов (последствий). По сути это постановка решаемой проблемы.

Описательные шкалы служат для формального описания факторов, а классификационные – результатов их действия на объект моделирования. Шкалы могут быть числовые и текстовые. Текстовые шкалы могут быть номинальные и порядковые.

Когнитивная структуризация предметной области является первым и единственным неавтоматизированным в системе «Эйдос» этапом АСК-анализа, т.е. все последующие этапы АСК анализа в ней полностью автоматизированы.

В АСК-анализе и системе «Эйдос» применяется две интерпретации классификационных и описательных шкал и градаций: *статичная и динамичная* и соответствующая терминология (обобщающая, статичная и динамичная). Есть также обобщающая интерпретация и соответствующая ей терминология.

Статичная интерпретация и терминология:

- градации классификационных шкал – это обобщающие категории видов объектов (классы);
- описательные шкалы – свойства объектов, градации описательных шкал – значения свойств (признаки) объектов.

Динамичная интерпретация и терминология:

- градации классификационных шкал – это обобщающие категории будущих состояний объекта моделирования (классы), описывающие результаты действия факторов на объект моделирования в натуральном и стоимостном выражении: например количество и качество продукции, прибыль и рентабельность;
- описательные шкалы – факторы, действующие на объект моделирования, градации описательных шкал – значения факторов, действующие на объект моделирования.

Обобщающая терминология:

- классификационные шкалы и градации;
- описательные шкалы и градации.

В данной работе в качестве *объекта моделирования* выступает покупатель в магазине, в качестве *факторов* пол, возраст, доход (таблица 1), а в качестве *результатов* действия этих факторов является оценка трат (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	Gender
2	Age
3	Annual Income

Источник: c:\Aidos-X\AID_DATA\A0000027\System\Opis_Sc.xlsx

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	Spending Score

Источник: c:\Aidos-X\AID_DATA\A0000027\System\Class_Sc.xlsx

3.2. Задача-2. Формализация предметной области

На этапе формализации предметной области разрабатываются классификационные и описательные шкалы и градации, а затем исходные данные кодируются с их использованием, в результате чего получается обучающая выборка. Обучающая выборка, по сути, представляет собой исходные данные, *нормализованные* с помощью классификационных и описательных шкал и градаций. Этим самым подготавливается все необходимое для выполнения следующего этапа АСК-анализа: синтеза и верификации статистических и системно-когнитивных моделей.

В системе «Эйдос» имеется большое количество разнообразных автоматизированных программных интерфейсов (API), обеспечивающих ввод в систему внешних данных различных типов: текстовых, табличных и графических, а также других, которые могут быть представлены в этом виде, например, аудио или данные электроэнцефалограммы (ЭЦГ) или кардиограммы (ЭКГ).

Этим обеспечивается возможность комфортного для пользователя применения системы «Эйдос» для проведения научных исследований и решения практических задач в самых различных предметных областях и направлениях науки, практически почти везде, где человек применяет естественный интеллект.

В качестве *источника исходных данных* в данной работе используем Excel-таблицы из работы [26]⁹ (см. таблицы 3 и 4):

Таблица 3 – Исходные данные по влиянию характеристик покупателей на оценку их трат

	A	B	C	D	E	F	G	H
1	CustomerID	Genre	Age	Annual Income	Spending Score			
2	0001	Male	19	15	39			
3	0002	Male	21	15	81			
4	0003	Female	20	16	6			
5	0004	Female	23	16	77			
6	0005	Female	31	17	40			
7	0006	Female	22	17	76			
8	0007	Female	35	18	6			
9	0008	Female	23	18	94			
10	0009	Male	64	19	3			
11	0010	Female	30	19	72			
12	0011	Male	67	19	14			
13	0012	Female	35	19	99			
14	0013	Female	58	20	15			
15	0014	Female	24	20	77			
16	0015	Male	37	20	13			
17	0016	Male	22	20	79			
18	0017	Female	35	21	35			
19	0018	Male	20	21	66			
20	0019	Male	52	23	29			
21	0020	Female	35	23	98			
22	0021	Male	35	24	35			
23	0022	Male	25	24	73			
24	0023	Female	46	25	5			
25	0024	Male	31	25	73			
26	0025	Female	54	28	14			
27	0026	Male	29	28	82			
28	0027	Female	45	28	32			
29	0028	Male	35	28	61			
30	0029	Female	40	29	31			
31	0030	Female	23	29	87			
32	0031	Male	60	30	4			
33	0032	Female	21	30	73			
34	0033	Male	53	33	4			
35	0034	Male	18	33	92			
36	0035	Female	49	33	14			
37	0036	Female	21	33	81			
38	0037	Female	42	34	17			
39	0038	Female	30	34	73			
40	0039	Female	26	37	26			

Используя стандартные возможности MS Excel, исходные данные из таблицы 3 представим их в виде, стандартном для системы «Эйдос» (таблица 4):

Таблица 4 – Таблица исходных данных в стандарте системы «Эйдос» (фрагмент)

	A	B	C	D	E	F
1	CustomerID	Genre	Age	Annual Income	Spending Score	
2	0001	Male	19	15	39	
3	0002	Male	21	15	81	
4	0003	Female	20	16	6	
5	0004	Female	23	16	77	
6	0005	Female	31	17	40	
7	0006	Female	22	17	76	
8	0007	Female	35	18	6	
9	0008	Female	23	18	94	
10	0009	Male	64	19	3	
11	0010	Female	30	19	72	
12	0011	Male	67	19	14	
13	0012	Female	35	19	99	
14	0013	Female	58	20	15	
15	0014	Female	24	20	77	
16	0015	Male	37	20	13	
17	0016	Male	22	20	79	
18	0017	Female	35	21	35	
19	0018	Male	20	21	66	
20	0019	Male	52	23	29	
21	0020	Female	35	23	98	
22	0021	Male	35	24	35	
23	0022	Male	25	24	73	
24	0023	Female	46	25	5	
25	0024	Male	31	25	73	
26	0025	Female	54	28	14	
27	0026	Male	29	28	82	
28	0027	Female	45	28	32	
29	0028	Male	35	28	61	

Таблица 4 имеет следующую структуру:

- каждая строка описывает одного покупателя, всего их 200;
- каждое *наблюдение* описывается одновременно двумя способами: с одной стороны, значениями факторов, действующих на объект моделирования (лингвистические и числовые переменные, градации описательных шкал, бесцветный фон), а с другой стороны результатами действия этих факторов, т.е. покупательская способность, выраженная в

числовой шкале (желтый фон). Такая структура описания наблюдений в технологиях искусственного интеллекта называется «онтологией» и модели представлений знаний Марвина Мински (1975) называется «фрейм-экземпляр»;

– 1-я колонка – номер наблюдения (не является шкалой);

– колонка 5 – это классификационная шкала – это шкалы *текстового* и *числового* типа описывающие *результаты* действия факторов, в данном случае пол, возраст, доход (таблица 2). В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в натуральном и стоимостном выражении [8]: например, количество покупок, товаров и так далее. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

– колонки со 2-й по 4-ю – это описательные шкалы, формализующие факторы, действующие на объект моделирования (таблица 1). Эти шкалы имеют числовой и текстовый тип и их градациями являются лингвистические и числовые переменные;

– при вводе данных в систему «Эйдос» нули и пробелы в исходных данных могут рассматриваться как значащие или как отсутствие данных. 2-й вариант и будет использован в данной работе.

Математическая модель системы «Эйдос» обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных (фрагментированных) зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

Таким образом, система «Эйдос» не предъявляет чрезмерно жестких практически невыполнимых требований к исходным данным, а обрабатывает те данные, которые реально есть, например, подобные представленным в таблице 4.

В системе «Эйдос» есть 6 основных автоматизированных программных интерфейсов (API), обеспечивающих ввод в систему и интеллектуальную обработку числовых, текстовых и графических данных, представленных в виде таблиц и файлов. Возможна обработка и других видов данных (например, данных землетрясений, ЭЭГ, ЭКГ, аудио и видео), которые возможно представить в этих форматах (рисунок 6).

2.3.2. Программные интерфейсы с внешними базами данных

- 2.3.2.1. Импорт данных из текстовых файлов
- 2.3.2.2. Универсальный программный интерфейс импорта данных в систему
- 2.3.2.3. Импорт данных из транспонированных внешних баз данных
- 2.3.2.4. Оцифровка изображений по внешним контурам
- 2.3.2.5. Оцифровка изображений по всем пикселям и спектру
- 2.3.2.6. Сценарный АСК-анализ символьных и числовых рядов
- 2.3.2.7. Транспонирование файлов исходных данных
- 2.3.2.8. Объединение нескольких файлов исходных данных в один
- 2.3.2.9. Разбиение TXT-файла на файлы-абзацы
- 2.3.2.10. CSV => DBF конвертер системы "Эйдос"
- 2.3.2.11. Прогноз событий по астропараметрам по Н.А.Чередниченко
- 2.3.2.12. Прогнозирование землетрясений методом Н.А.Чередниченко
- 2.3.2.13. Чемпионат RAIF-Challenge 2017-API-bank
- 2.3.2.14. Чемпионат RAIF-Challenge 2017-API-retail
- 2.3.2.15. Вставка промежуточных строк в файл исходных данных Inp_data

Рисунок 4. Автоматизированные программные интерфейсы (API) системы «Эйдос»

Для ввода исходных данных, представленных в таблице 4, в систему «Эйдос», используется один из ее автоматизированных программных интерфейсов (API), а именно универсальный автоматизированный программный интерфейс ввода данных из файлов MS Excel (API-2.3.2.2).

Требования API-2.3.2.2 к исходным данным подробно описаны в хелпах этого режима (рисунки 7):

Принцип организации таблиц исходных данных

Наименование объекта обучающей выборки	Наименование 1-й описательной шкалы	Наименование 2-й описательной шкалы	Наименование 1-й описательной шкалы	Наименование 2-й описательной шкалы
1-й объект обучающей выборки (1-е наблюдение)	Значение шкалы	Значение шкалы	Значение шкалы	Значение шкалы
2-й объект обучающей выборки (2-е наблюдение)	Значение шкалы	Значение шкалы	Значение шкалы	Значение шкалы
...

Принцип организации таблиц исходных данных

Наименование объекта обучающей выборки	Наименование 1-й описательной шкалы	Наименование 2-й описательной шкалы	Наименование 1-й описательной шкалы	Наименование 2-й описательной шкалы
1-й объект обучающей выборки (1-е наблюдение)	Значение шкалы	Значение шкалы	Значение шкалы	Значение шкалы
2-й объект обучающей выборки (2-е наблюдение)	Значение шкалы	Значение шкалы	Значение шкалы	Значение шкалы
...

Определения основных терминов и профилактика типичных ошибок при подготовке Excel-файла исходных данных

Принцип организации таблиц исходных данных

Термины АСК-АНАЛИЗА И СИСТЕМЫ "ЭЙДОС":
 Шкала представляет собой способ формализации предметной области. Используются числовые и текстовые шкалы, при этом текстовые могут быть номинальными и порядковыми. Но номинальные шкалы есть только отношения эквивалентности и неэквивалентности, на порядковые кроме того еще отношения "больше", "меньше", а на числовые - кроме того могут выполняться все арифметические операции. Каждый объект выборки (наблюдение) описан с одной стороны своими признаками, а с другой - принадлежностью к некоторым обучающим категориям (классам). Такая структура описания позволяет использовать методы машинного обучения и является базовой для всех надежных предельных значений в АСК-анализе и системе "Эйдос". Используются три интерпретации шкал и градаций: универсальная, статистическая и динамическая:
 - в универсальной интерпретации: признаки - это градации описательных шкал;
 - в статистической интерпретации: описательных шкалы - это свойства, а градации признаков - это степени выраженности этого свойства;
 - в динамической интерпретации: описательных шкалы - это факторы, а градации признаков - это значения факторов;
 - в универсальной интерпретации: классы - это градации классификационных шкал;
 - в статистической интерпретации: классификационная шкалы - способ классификации обучающих категорий (классов), в которых в настоящее время по отношению к признакам описано состояние объекта моделирования;
 - в динамической интерпретации: классификационная шкалы - способ классификации обучающих категорий (классов), в которых в будущем времени по отношению к признакам описано состояние объекта прогнозирования или управления.
ПРОФИЛАКТИКА ОШИБОК В ФАЙЛЕ ИСХОДНЫХ ДАННЫХ:
 - 1-я строка файла "Inp_data.xls" должна содержать наименования колонок. Эти наименования должны быть во всех колонках, при этом перенося по словам разрывы, а объединение ячеек, разрыв строки знак абзаца и неактивные символы не допускаются. Эти наименования должны быть короткими, но полными, т.е. они будут в выделенной форме, а не как был бы добавлен наименования градаций. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения и число знаков после запятой в колонке должно быть ОДИНАКОЕ.
 - 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным, до 255 символов.
 - Каждая строка этого файла, начиная со 2-го, содержит данные об одном объекте обучающей выборки или одном наблюдении. В MS Excel 2003 в листе может быть до 65536 строк и до 256 колонок. В листе MS Excel 2010 и более поздних колонок до 1048576 строк и 16384 колонок.
 - Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (номинального / порядкового) или числового типа (с десятичными знаками после запятой).
 - Столбцы присваиваются числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбец присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
 - Столбцы со 2-го по N-й являются классификационными шкалами (выявлены параметрами) и содержат данные о классах (будущая состояние объекта управления), в которых принадлежит объект обучающей выборки.
 - Столбцы с N+1 по последний являются описательными шкалами (свойствами или факторами) и содержат данные о признаках (т.е. значениях свойств или значений факторов), характеризующие объект обучающей выборки.
 - В результате работы режима формируется файл INP_NAME.TXT стандарта MS DOS (информация), в котором наименования классификационных и описательных шкал являются СТРОКАМИ. Система формирует классификационные и описательные шкалы и градации. Два этого в каждой числовой столбце система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего числовые значения заменяются и интервальные значения. В текстовых столбцах система находит уникальные текстовые значения. Каждый УНИКАЛЬНЫЙ интервальный числовой или текстовый значение считается градацией классификационной или описательной шкалы, кодифицирующей объект. В каждой строке ее градаций соответствует по крайней мере одна строка исходных данных. NP_DATA и содержит коды классов, соответствующие фактам сведениям числовым или уникальным текстовым значениям классов с градациями классификационных шкал и коды признаков, соответствующие фактам сведениям числовым или уникальным текстовым значениям признаков с градациями описательных шкал. Расположением файла формируется на основе файла INP_RASP анализатор, за исключением того, что классификационные и описательные шкалы и градации не создаются, а используются ранее созданные в файлах, и базы расположением выборки могут не включать коды классов, если столбцы классов в файле INP_RASP были пустыми. Структура файла INP_RASP должна быть также же, как INP_DATA, т.е. они должны ПОЛНОСТЬЮ соответствовать по наименованиям столбцам, но могут иметь разное количество строк с равными значениями в них.

Рисунок 5. Хелпы API-2.3.2.2 системы «Эйдос»

Экранные формы управления API-2.3.2.2 системы «Эйдос» с реальными параметрами, использованными в данной работе, приведены на рисунках 8.

На 5, 6, 7 приведены классификационные и описательные шкалы и градации, а также обучающая выборка, сформированные API-2.3.2.2 при параметрах, показанных на рисунках 8.

Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал и между шкалами, как числовыми, так и текстовыми. Математическая модель АСК-анализа позволяет корректно преодолеть несбалансированность данных путем перехода от абсолютных частот к относительным и к количественным мерам знаний в системно-когнитивных моделях (мы увидим это ниже).

2.3.2.2. Универсальный программный интерфейс импорта данных в систему "ЭЙДОС-Х++"

Автоматическая формализация предметной области: генерация классификационных и описательных шкал и градаций, а также обучающей и распознаваемой выборки на основе базы исходных данных: "Inp_data"

Задайте тип файла исходных данных: "Inp_data":

- XLS - MS Excel-2003 Стандарт XLS-файла
- XLSX - MS Excel-2007(2010) Стандарт DBF-файла
- DBF - DBASE IV (DBF/NTX) Стандарт CSV-файла
- CSV - CSV => DBF конвертер

Задайте параметры:

- Нули и пробелы считать ОТСУТСТВИЕМ данных
- Нули и пробелы считать ЗНАЧЕНИЯМИ данных
- Создавать БД средних по классам "Inp_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

- Формализации предметной области (на основе "Inp_data")
- Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

- Равные интервалы с разным числом наблюдений
- Разные интервалы с равным числом наблюдений

Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp_data":

- Не применять сценарный метод АСК-анализа
- Применить сценарный метод АСК-анализа
- Применить спец.интерпретацию текстовых полей классов
- Применить спец.интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:

Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

- Только интервальные числовые значения (например: "1/3-{59873.0000000, 178545.6666667}")
- Только наименования интервальных числовых значений (например: "Минимальное")
- И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-{59873.0000000, 178545.6666667}")

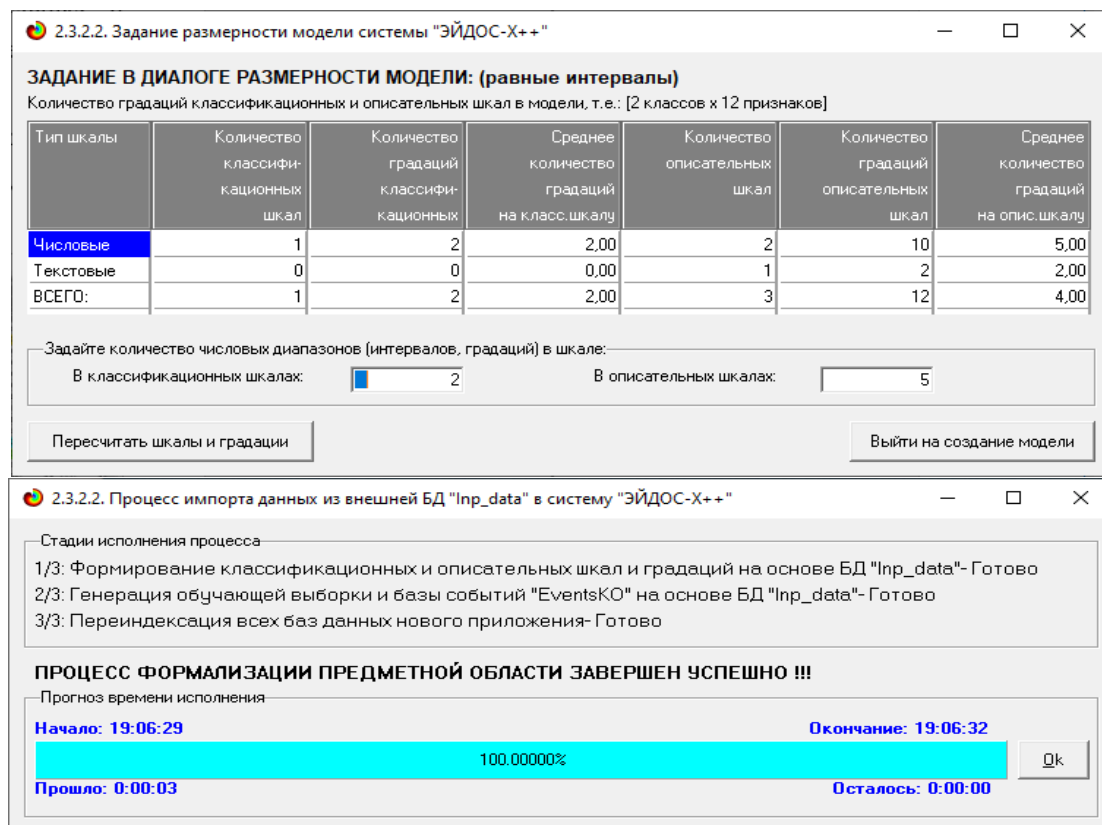


Рисунок 6. Экранные формы управления API-2.3.2.2 системы «Эйдос»

Таблица 5 – Классификационные шкалы и градации

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы
1	SPENDING SCORE	1	Малое
		2	Большое

Таблица 6 – Описательные шкалы и градации

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	GENRE	8	Очень малое
2	AGE	9	Малое
3	ANNUAL INCOME	10	Среднее
		11	Большое
		12	Очень большое

Таблица 7 – Обучающая выборка

CustomerID	Genre	Age	Annual Income	Spending Score
0001	Male	19	15	39
0002	Male	21	15	81
0003	Female	20	16	6
0004	Female	23	16	77
0005	Female	31	17	40
0006	Female	22	17	76
0007	Female	35	18	6
0008	Female	23	18	94
0009	Male	64	19	3
0010	Female	30	19	72
0011	Male	67	19	14
0012	Female	35	19	99
0013	Female	58	20	15
0014	Female	24	20	77
0015	Male	37	20	13
0016	Male	22	20	79
0017	Female	35	21	35
0018	Male	20	21	66
0019	Male	52	23	29
0020	Female	35	23	98
0021	Male	35	24	35
0022	Male	25	24	73
0023	Female	46	25	5
0024	Male	31	25	73
0025	Female	54	28	14
0026	Male	29	28	82
0027	Female	45	28	32
0028	Male	35	28	61
0029	Female	40	29	31
0030	Female	23	29	87
0031	Male	60	30	4
0032	Female	21	30	73
0033	Male	53	33	4
0034	Male	18	33	92
0035	Female	49	33	14
0036	Female	21	33	81
0037	Female	42	34	17
0038	Female	30	34	73
0039	Female	36	37	26
0040	Female	20	37	75
0041	Female	65	38	35
0042	Male	24	38	92
0043	Male	48	39	36
0044	Female	31	39	61
0045	Female	49	39	28
0046	Female	24	39	65
0047	Female	50	40	55
0048	Female	27	40	47
0049	Female	29	40	42
0050	Female	31	40	42
0051	Female	49	42	52
0052	Male	33	42	60

0053	Female	31	43	54
0054	Male	59	43	60
0055	Female	50	43	45
0056	Male	47	43	41
0057	Female	51	44	50
0058	Male	69	44	46
0059	Female	27	46	51
0060	Male	53	46	46
0061	Male	70	46	56
0062	Male	19	46	55
0063	Female	67	47	52
0064	Female	54	47	59
0065	Male	63	48	51
0066	Male	18	48	59
0067	Female	43	48	50
0068	Female	68	48	48
0069	Male	19	48	59
0070	Female	32	48	47
0071	Male	70	49	55
0072	Female	47	49	42
0073	Female	60	50	49
0074	Female	60	50	56
0075	Male	59	54	47
0076	Male	26	54	54
0077	Female	45	54	53
0078	Male	40	54	48
0079	Female	23	54	52
0080	Female	49	54	42
0081	Male	57	54	51
0082	Male	38	54	55
0083	Male	67	54	41
0084	Female	46	54	44
0085	Female	21	54	57
0086	Male	48	54	46
0087	Female	55	57	58
0088	Female	22	57	55
0089	Female	34	58	60
0090	Female	50	58	46
0091	Female	68	59	55
0092	Male	18	59	41
0093	Male	48	60	49
0094	Female	40	60	40
0095	Female	32	60	42
0096	Male	24	60	52
0097	Female	47	60	47
0098	Female	27	60	50
0099	Male	48	61	42
0100	Male	20	61	49
0101	Female	23	62	41
0102	Female	49	62	48
0103	Male	67	62	59
0104	Male	26	62	55
0105	Male	49	62	56
0106	Female	21	62	42
0107	Female	66	63	50

0108	Male	54	63	46
0109	Male	68	63	43
0110	Male	66	63	48
0111	Male	65	63	52
0112	Female	19	63	54
0113	Female	38	64	42
0114	Male	19	64	46
0115	Female	18	65	48
0116	Female	19	65	50
0117	Female	63	65	43
0118	Female	49	65	59
0119	Female	51	67	43
0120	Female	50	67	57
0121	Male	27	67	56
0122	Female	38	67	40
0123	Female	40	69	58
0124	Male	39	69	91
0125	Female	23	70	29
0126	Female	31	70	77
0127	Male	43	71	35
0128	Male	40	71	95
0129	Male	59	71	11
0130	Male	38	71	75
0131	Male	47	71	9
0132	Male	39	71	75
0133	Female	25	72	34
0134	Female	31	72	71
0135	Male	20	73	5
0136	Female	29	73	88
0137	Female	44	73	7
0138	Male	32	73	73
0139	Male	19	74	10
0140	Female	35	74	72
0141	Female	57	75	5
0142	Male	32	75	93
0143	Female	28	76	40
0144	Female	32	76	87
0145	Male	25	77	12
0146	Male	28	77	97
0147	Male	48	77	36
0148	Female	32	77	74
0149	Female	34	78	22
0150	Male	34	78	90
0151	Male	43	78	17
0152	Male	39	78	88
0153	Female	44	78	20
0154	Female	38	78	76
0155	Female	47	78	16
0156	Female	27	78	89
0157	Male	37	78	1
0158	Female	30	78	78
0159	Male	34	78	1
0160	Female	30	78	73
0161	Female	56	79	35
0162	Female	29	79	83

0163	Male	19	81	5
0164	Female	31	81	93
0165	Male	50	85	26
0166	Female	36	85	75
0167	Male	42	86	20
0168	Female	33	86	95
0169	Female	36	87	27
0170	Male	32	87	63
0171	Male	40	87	13
0172	Male	28	87	75
0173	Male	36	87	10
0174	Male	36	87	92
0175	Female	52	88	13
0176	Female	30	88	86
0177	Male	58	88	15
0178	Male	27	88	69
0179	Male	59	93	14
0180	Male	35	93	90
0181	Female	37	97	32
0182	Female	32	97	86
0183	Male	46	98	15
0184	Female	29	98	88
0185	Female	41	99	39
0186	Male	30	99	97
0187	Female	54	101	24
0188	Male	28	101	68
0189	Female	41	103	17
0190	Female	36	103	85
0191	Female	34	103	23
0192	Female	32	103	69
0193	Male	33	113	8
0194	Female	38	113	91
0195	Female	47	120	16
0196	Female	35	120	79
0197	Female	45	126	28
0198	Male	32	126	74
0199	Male	32	137	18
0200	Male	30	137	83

Отметим, что в системе «Эйдос» обычно используются базы данных с расширением «dbf». Они открываются в MS Excel или могут быть конвертированы в файлы xlsm с помощью онлайн-сервисов или в режиме 5.12 (этот режим системы «Эйдос» написан на Питоне).

3.3. Задача-3. Синтез статистических и системно-когнитивных моделей. Многопараметрическая типизация и частные критерии знаний

Математические модели, на основе которых рассчитываются статистические и системно-когнитивные модели автоматизированного системно-когнитивного анализа (АСК-анализ) и интеллектуальной системы «Эйдос», подробно описаны в ряде монографий и статей автора [1-4]. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко, акцентируя внимание лишь на математической взаимосвязи коэффициента возврата инвестиций (ROI) с мерой χ -квадрат Карла Пирсона и с семантической мерой целесообразности информации Александра Харкевича.

Отметим, что модели системы «Эйдос» основаны на матрице абсолютных частот, отражающей число встреч градаций описательных шкал по градациям классификационных шкал (фактов). Но для решения всех задач используется не непосредственно сама эта матрица, а матрицы условных и безусловных процентных распределений и системно-когнитивные модели, которые рассчитываются на ее основе и отражают какое количество информации содержится в факте наблюдения определенной градации описательной шкалы о том, что объект моделирования перейдет в состояние, соответствующее определенной градации классификационной шкалы (классу).

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике [1-7] и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых (нелинейных) данных, представленных в различных типах шкал (дихотомических, номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот (таблица 8).

Таблица 8 – Матрица абсолютных частот (статистическая модель ABS)

		Классы					Сумма
		1	...	j	...	W	
Значения факторов	1	N_{11}		N_{1j}		N_{1W}	
	...						
	i	N_{i1}		N_{ij}		N_{iW}	$N_{i\Sigma} = \sum_{j=1}^W N_{ij}$
	...						
	M	N_{M1}		N_{Mj}		N_{MW}	
Суммарное количество признаков по классу				$N_{\Sigma j} = \sum_{i=1}^M N_{ij}$			$N_{\Sigma\Sigma} = \sum_{i=1}^W \sum_{j=1}^M N_{ij}$
Суммарное количество объектов обучающей выборки по классу				$N_{\Sigma j}$			$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

На основе таблицы 8 рассчитываются матрицы условных и безусловных процентных распределений (таблица 9).

Отметим, что в АСК-анализе и его программном инструментарии интеллектуальной системе «Эйдос» используется два способа расчета матриц условных и безусловных процентных распределений:

Таблица 9 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

		Классы					Безусловная вероятность признака
		1	...	j	...	W	
Значения факторов	1	P_{11}		P_{1j}		P_{1W}	
	...						
	i	P_{i1}		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		P_{iW}	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	M	P_{M1}		P_{Mj}		P_{MW}	
Безусловная вероятность класса				$P_{\Sigma j}$			

1-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;

2-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу.

На практике часто встречается существенная несбалансированность данных, под которой понимается сильно отличающееся количество наблюдений объектов обучающейся выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 8) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 9) является весьма обоснованным и логичным.

Этот переход полностью снимает проблему несбалансированности данных, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 8), а матрицы условных и безусловных процентных распределений (таблица 9), а также матрицы системно-когнитивных моделей, рассчитываемые на основе матрица абсолютных частот и матрицы условных и безусловных процентных распределений. Этот подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (дихотомических, номинальных, порядковых и числовых) и в разных единицах измерения [8]. В системе «Эйдос» этот подход применяется всегда при решении любых задач.

Затем на основе таблиц 8 и 9 с использованием частных критериев, знаний приведенных таблице 10, рассчитываются матрицы 7 системно-когнитивных моделей (таблица 11).

В таблице 10 приведены формулы:

- для сравнения фактических и теоретических абсолютных частот;
- для сравнения условных и безусловных относительных частот («вероятностей»).

И это **сравнение** в таблицах 8 и 9 осуществляется двумя возможными способами: путем **вычитания** и путем **деления**.

Таблица 10– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	Через относительные частоты	Через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке (таблица 7)	N_{ij} – фактическая частота, $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N}$ – теоретическая частота.	
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	...	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу	...	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \log_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \log_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \log_2 \frac{N_{ij} N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.		
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	...	$I_{ij} = N_{ij} - \bar{N}_{ij} = N_{ij} - \frac{N_i N_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		

Обозначения к таблице:

i – значение прошлого параметра;

j – значение будущего параметра;

N_{ij} – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра;

M – суммарное число значений всех прошлых параметров;

W – суммарное число значений всех будущих параметров.

N_i – количество встреч i -м значения прошлого параметра по всей выборке;

N_j – количество встреч j -го значения будущего параметра по всей выборке;

N – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра по всей выборке.

I_{ij} – частный критерий знаний: количество знаний в факте наблюдения i -го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее j -му значению будущего параметра;

Ψ – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;

P_i – безусловная относительная частота встречи i -го значения прошлого параметра в обучающей выборке;

P_{ij} – условная относительная частота встречи i -го значения прошлого параметра при j -м значении будущего параметра.

Таблица 11 – Матрица системно-когнитивной модели

		Классы					Значимость фактора
		1	...	j	...	W	
Значения факторов	1	I_{11}		I_{1j}		I_{1W}	$\sigma_{1\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{1j} - \bar{I}_1)^2}$
	...						
	i	I_{i1}		I_{ij}		I_{iW}	$\sigma_{i\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
	...						
	M	I_{M1}		I_{Mj}		I_{MW}	$\sigma_{M\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{Mj} - \bar{I}_M)^2}$
Степень редукции класса		$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 10), применяемых в настоящее время в системе «Эйдос» равно 7 определяется тем, что они получаются путем всех возможных вариантов сравнения фактических и теоретических абсолютных частот, условных и безусловных относительных частот путем вычитания и путем деления, и при этом N_j рассматривается как суммарное количество или признаков, или объектов обучающей выборки в j -м классе, а **нормировка к нулю** (для аддитивных интегральных критериев), если нет связи между наличием признака и принадлежностью объекта к классу, осуществляется либо логарифмированием, либо вычитанием единицы (таблица 12).

Таблица 12– Конфигуратор системно-когнитивных моделей АСК-анализа и интеллектуальной системы «Эйдос»

	Способ сравнения	Нормировка не требуется	Нормировка к 0 путем взятия логарифма	Нормировка к 0 путем вычитания 1
Сравнение фактических и теоретических абсолютных частот	Путем деления	---	INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
	Путем вычитания	INF3, χ -квадрат Карла Пирсона	---	---
Сравнение условных и безусловных относительных частот	Путем деления	---	INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
	Путем вычитания	INF6, INF7	---	---

Обратим особое внимание на то, что сравнение фактических и теоретических абсолютных частот путем деления приводит при нормировках к нулю (что нужно для применения аддитивных интегральных критериев) путем взятия логарифма и путем вычитания 1 к *тем же самым* моделям, что и сравнение условных и безусловных относительных частот путем деления с теми же самыми способами нормировки. Таким образом, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и провести нормировку к 0 путем взятия логарифма и путем вычитания 1, то получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Других же системно-когнитивных моделей, рассчитываемых на основе приведенных статистических моделей просто нет. Это и есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра. *Под конфигуратором В.А.Лефевр понимал минимальный полный набор понятийных шкал или конструкторов, т.е. понятий, достаточный для адекватного описания предметной области* [4]¹⁰. Необходимо отметить, что все эти модели рассчитываются в интеллектуальной системе «Эйдос».

Когда мы сравниваем фактические и теоретические абсолютные частоты путем вычитания у нас получается частный критерий знаний: «хи-квадрат» (СК-модель INF3), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2) или «коэффициент возврата инвестиций ROI» - Return On Investment (СК-модели INF4, INF5) в зависимости от способа нормировки.

Когда же мы сравниваем условные и безусловные относительные частоты путем вычитания у нас получается частный критерий знаний: «коэффициент взаимосвязи» (СК-модели INF6, INF7), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2).

Таким образом, мы видим, что все частные критерии знаний тесно взаимосвязаны друг с другом. Особенно интересна связь знаменитого

10 См. 1.2.1.2.1.1. Определение понятия конфигулятора, http://lc.kubagro.ru/aidos/aidos06_lec/index.htm

критерия хи-квадрат К.Пирсона с замечательной мерой количества информации А.Харкевича и с известным в экономике коэффициентом ROI.

Вероятность рассматривается как предел, к которому стремится относительная частота (отношение количества благоприятных исходов к числу испытаний) при неограниченном увеличении количества испытаний. Ясно, что вероятность – это математическая абстракция, которая никогда не встречается на практике (также как и другие математические и физические абстракции, типа математической точки, материальной точки, бесконечно малой и т.п.). На практике встречается только относительная частота. Но она может быть весьма близкой к вероятности. Например, при 480 наблюдений различие между относительной частотой и вероятностью (погрешность) составляет около 5%, при 1250 наблюдениях – около 2.5%, при 10000 наблюдениях – 1%.

Суть этих методов в том, что вычисляется количество информации в значении фактора о том, что объект моделирования перейдет под его действием в определенное состояние, соответствующее классу. Это позволяет сопоставимо и корректно обрабатывать разнородную информацию о наблюдениях объекта моделирования, представленную в различных типах измерительных шкал и различных единицах измерения [8].

На основе системно-когнитивных моделей, представленных в таблице 11 (отличаются частыми критериями, приведенными в таблице 10), решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений (обратная задача прогнозирования), а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Отметим, что как значимость значения фактора, степень детерминированности класса и ценность или качество модели в АСК-анализе рассматривается вариабельность значений частных критериев этого значения фактора, класса или модели в целом (таблица 13).

Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте

моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 13):

Таблица 13 – Уточнение терминологии АСК-анализа

№	Традиционные термины (синонимы)	Новый термин	Формула
1	1. Значимость значения фактора (признака). 2. Дифференцирующая мощность значения фактора (признака). 3. Ценность значения фактора (признака) для решения задачи идентификации и других задач	Корень из информационной мощности значения фактора	$\sigma_{i\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
2	1. Степень детерминированности класса. 2. Степень обусловленности класса.	Корень из информационной мощности класса	$\sigma_{\Sigma j} = \sqrt[2]{\frac{1}{M-1} \sum_{i=1}^M (I_{ij} - \bar{I}_j)^2}$
3	1. Качество модели. 2. Ценность модели. 3. Степень сформированности модели. 4. Количественная мера степени выраженности закономерностей в моделируемой предметной области	Корень из информационной мощности модели	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Итак, в разделе раскрывается простая Математическая взаимосвязь меры χ -квадрат Карла Пирсона с коэффициентом возврата инвестиций (ROI) и с семантической мерой целесообразности информации Александра Харкевича. Эта взаимосвязь обнаруживается, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и выполнить нормировку к нулю путем взятия логарифма или вычитания 1. При этом получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Именно 7, а не большее количество системно-когнитивных моделей в итоге получается потому, что модели, получающиеся в результате сравнения фактических и теоретических абсолютных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1 **тождественно совпадают** с моделями, получающимися путем сравнения условных и безусловных относительных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1. Это и есть конфигурактор статистических и когнитивных моделей в смысле В.А.Лефевра, содержащий минимальное количество

моделей, позволяющих полно описать моделируемую предметную область.

Показательно, что модель меры χ -квадрат Карла Пирсона из статистики оказалась математически тесно связанной с коэффициентом возврата инвестиций (ROI), применяемой в экономике в теории управления портфелем инвестиций и с мерой информации Александра Харкевича из семантической теории информации и теории управления знаниями. Все эти модели рассчитываются в интеллектуальной системе «Эйдос».

В системе «Эйдос» синтез моделей производится в режиме 3.5 (рисунок 7):

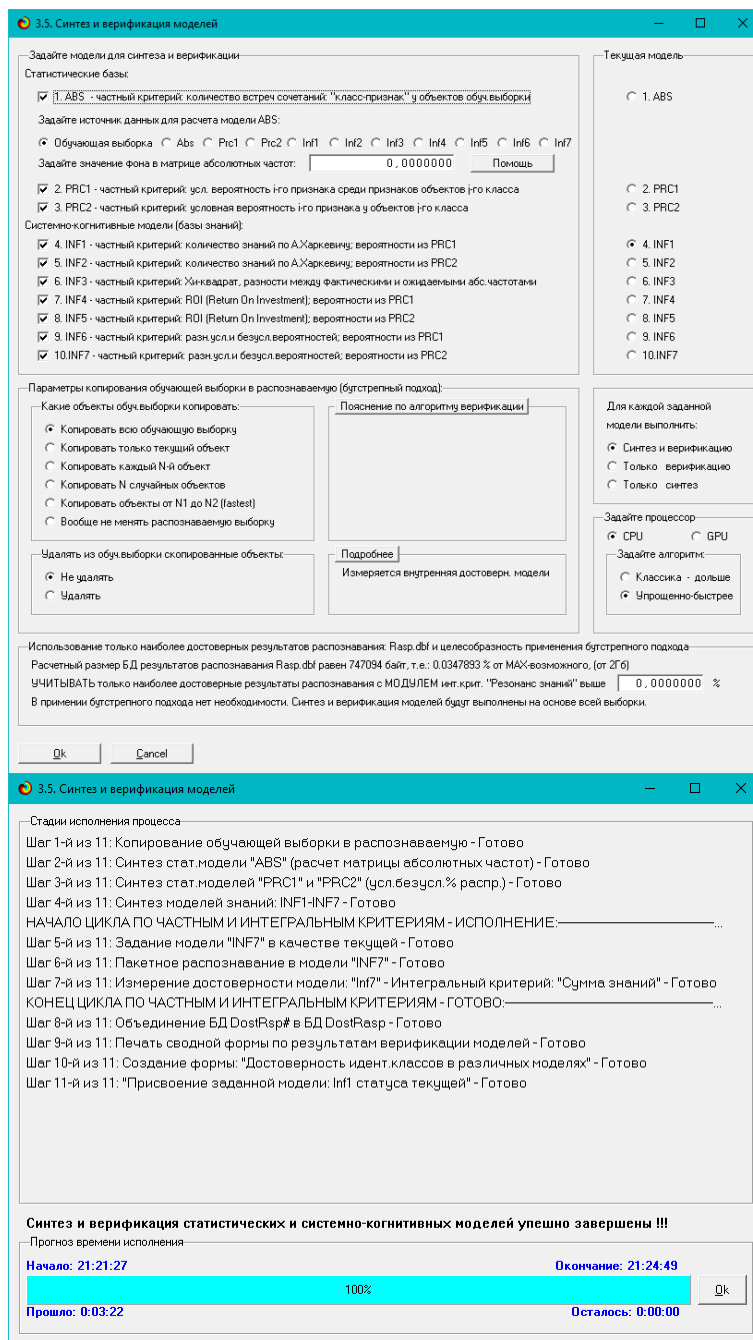


Рисунок 7. Экранные формы режима синтеза и верификации моделей

В результате работы режима 3.5 создано 3 статистических и 7 системно-когнитивных моделей, некоторые из которых приведены на рисунках 8-11:

5.5. Модель: "1. ABS - частный критерий: количество встреч сочетаний: "Класс-признак" у объектов обуч.выборки"

Код признака	Наименование описательной шкалы и градации	1. SPENDING SCORE МАЛОЕ	2. SPENDING SCORE БОЛЬШОЕ	Сумма	Среднее	Средн. квадр. откл.
1.0	GENRE-Female	59.0	53.0	112.0	56.00	4.24
2.0	GENRE-Male	44.0	44.0	88.0	44.00	
3.0	AGE-Очень малое	18.0	32.0	50.0	25.00	9.90
4.0	AGE-Малое	21.0	42.0	63.0	31.50	14.85
5.0	AGE-Среднее	33.0	9.0	42.0	21.00	16.97
6.0	AGE-Большое	19.0	6.0	25.0	12.50	9.19
7.0	AGE-Очень большое	12.0	8.0	20.0	10.00	2.83
8.0	ANNUAL INCOME-Очень малое	23.0	23.0	46.0	23.00	
9.0	ANNUAL INCOME-Малое	35.0	31.0	66.0	33.00	2.83
10.0	ANNUAL INCOME-Среднее	34.0	32.0	66.0	33.00	1.41
11.0	ANNUAL INCOME-Большое	7.0	7.0	14.0	7.00	
12.0	ANNUAL INCOME-Очень большое	4.0	4.0	8.0	4.00	
	Сумма числа признаков	309.0	291.0	600.0		
	Среднее	25.8	24.3		25.00	
	Среднеквадратичное отклонение	15.9	17.2			16.21
	Сумма числа объектов обуч.выборки	103.0	97.0	200.0		

Рисунок 8. Статистическая модель «ABS», матрица абсолютных частот (фрагмент)

5.5. Модель: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"

Код признака	Наименование описательной шкалы и градации	1. SPENDING SCORE МАЛОЕ	2. SPENDING SCORE БОЛЬШОЕ	Безусл. вероятн.	Среднее	Средн. квадр. откл.
1.0	GENRE-Female	57.282	54.639	56.000	55.960	2.028
2.0	GENRE-Male	42.718	45.361	44.000	44.040	2.028
3.0	AGE-Очень малое	17.476	32.990	25.000	25.233	11.135
4.0	AGE-Малое	20.388	43.299	31.500	31.844	16.366
5.0	AGE-Среднее	32.039	9.278	21.000	20.659	16.260
6.0	AGE-Большое	18.447	6.186	12.500	12.316	8.835
7.0	AGE-Очень большое	11.650	8.247	10.000	9.949	2.568
8.0	ANNUAL INCOME-Очень малое	22.330	23.711	23.000	23.021	1.131
9.0	ANNUAL INCOME-Малое	33.981	31.959	33.000	32.970	1.588
10.0	ANNUAL INCOME-Среднее	33.010	32.990	33.000	33.000	0.070
11.0	ANNUAL INCOME-Большое	6.796	7.216	7.000	7.006	0.433
12.0	ANNUAL INCOME-Очень большое	3.883	4.124	4.000	4.004	0.292
	Сумма	300.000	300.000	600.000		
	Среднее	25.000	25.000		25.000	
	Среднеквадратичное отклонение	15.489	17.698			16.250
	Сумма числа объектов обуч.выборки	103.000	97.000	200.000		

Рисунок 9. Статистическая модель «PRC2», матрица условных и безусловных процентных распределений (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. SPENDING SCORE МАЛОЕ	2. SPENDING SCORE БОЛЬШОЕ	Сумма	Среднее	Средн. квадр. откл.
1.0	GENRE-Female	0.004	-0.004	0.000	0.000	0.005
2.0	GENRE-Male	-0.005	0.005	0.000	0.000	0.007
3.0	AGE-Очень малое	-0.056	0.043	-0.013	-0.006	0.070
4.0	AGE-Малое	-0.068	0.050	-0.018	-0.009	0.083
5.0	AGE-Среднее	0.066	-0.128	-0.062	-0.031	0.137
6.0	AGE-Большое	0.061	-0.110	-0.049	-0.025	0.121
7.0	AGE-Очень большое	0.024	-0.030	-0.006	-0.003	0.038
8.0	ANNUAL INCOME-Очень малое	-0.005	0.005	0.000	0.000	0.007
9.0	ANNUAL INCOME-Малое	0.005	-0.005	0.000	0.000	0.007
10.0	ANNUAL INCOME-Среднее	0.000	0.000	0.000	0.000	0.000
11.0	ANNUAL INCOME-Большое	-0.005	0.005	0.000	0.000	0.007
12.0	ANNUAL INCOME-Очень большое	-0.005	0.005	0.000	0.000	0.007
	Сумма	0.016	-0.165	-0.148		
	Среднее	0.001	-0.014		-0.006	
	Среднеквадратичное отклонение	0.039	0.054			0.046
	Сумма числа объектов обуч.выборки	103.000	97.000	200.000		

Рисунок 10. Системно-когнитивная модель «INF1», матрица информанностей (по А.Харкевичу) (фрагмент)

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами"

Код признака	Наименование описательной шкалы и градации	1. SPENDING SCORE МАЛОЕ	2. SPENDING SCORE БОЛЬШОЕ	Сумма	Среднее	Средн. квадр. откл.
1.0	GENRE-Female	1.320	-1.320			1.867
2.0	GENRE-Male	-1.320	1.320			1.867
3.0	AGE-Очень малое	-7.750	7.750			10.960
4.0	AGE-Малое	-11.445	11.445			16.186
5.0	AGE-Среднее	11.370	-11.370			16.080
6.0	AGE-Большое	6.125	-6.125			8.662
7.0	AGE-Очень большое	1.700	-1.700			2.404
8.0	ANNUAL INCOME-Очень малое	-0.690	0.690			0.976
9.0	ANNUAL INCOME-Малое	1.010	-1.010			1.428
10.0	ANNUAL INCOME-Среднее	0.010	-0.010			0.014
11.0	ANNUAL INCOME-Большое	-0.210	0.210			0.297
12.0	ANNUAL INCOME-Очень большое	-0.120	0.120			0.170
	Сумма					
	Среднее					
	Среднеквадратичное отклонение	5.766	5.766			5.640
	Сумма числа объектов обуч.выборки	103.000	97.000	200.000		

Рисунок 11. Системно-когнитивная модель «INF3», матрица Хи-квадрат (по К.Пирсону) (фрагмент)

Полученные модели корректно использовать для решения задач только в том случае, если они достаточно достоверны (адекватны), т.е. верно отражают моделируемую предметную область. Поэтому в следующем разделе оценим достоверность созданных статистических и системно-когнитивных моделей.

3.4. Задача-4. Верификация моделей

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных и ложных положительных и отрицательных решений по F-мере Ван Ризбергера, а также по критериям L1- L2-мерам проф. Е.В.Луценко, которые

предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры [9].

Достоверность моделей можно оценивать и путем решения других задач, например, задач прогнозирования, выработки управляющих решений, исследования объекта моделирования путем исследования его модели. Но это более трудоемко и даже всегда возможно, особенно на экономических и политических моделях.

В режиме 3.4 системы «Эйдос» и ряде других изучается достоверность каждой частной модели в соответствии с этими мерами достоверности.

В соответствии с критерием достоверности L1-мерой проф.Е.В.Луценко наиболее достоверной является СК-модель INF4 с интегральным критерием: «Сумма знаний»: $L1=0.707$ (рисунок 12). *Эту модель и будем использовать для решения поставленных в работе задач.*

Наименование модели и частного критерия	Интегральный критерий	кно-льн. (FF)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма модел... уровенй склад... истинно-поло... решений (STP)	Сумма модел... уровенй склад... истинно-отриц... решений (STN)	Сумма модел... уровенй склад... ложно-полож... решений (SFP)	Сумма модел... уровенй склад... ложно-отрицат... решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В.Луценко	Средний модел... уровенй склад... истинно-полож... истинно-отриц... реш	Средний модел... уровенй склад... истинно-полож... истинно-отриц... реш
1. ABS - частный критерий: количество встреч сочетаний: "Клас...	Корреляция абс частот с обр...	188	3	0.512	0.985	0.674	115.557	1.736	89.826	0.471	0.563	0.996	0.719	0.587	
1. ABS - частный критерий: количество встреч сочетаний: "Клас...	Сумма абс частот по признак...	200		0.500	1.000	0.667	165.701		154.252		0.518	1.000	0.682	0.829	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция услоитн частот с о...	188	3	0.512	0.985	0.674	115.557	1.736	89.826	0.471	0.563	0.996	0.719	0.587	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма услоитн частот по приз...	200		0.500	1.000	0.667	160.770		149.759		0.518	1.000	0.682	0.804	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция услоитн частот с о...	188	3	0.512	0.985	0.674	115.550	1.736	89.820	0.471	0.563	0.996	0.719	0.587	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма услоитн частот по приз...	200		0.500	1.000	0.667	160.770		149.759		0.518	1.000	0.682	0.804	
4. INF1 - частный критерий: количество знаний по А/Харкевичу: в...	Семантический резонанс зна...	74	54	0.664	0.730	0.695	100.484	99.075	45.181	42.756	0.690	0.702	0.696	0.688	
4. INF1 - частный критерий: количество знаний по А/Харкевичу: в...	Сумма знаний	62	62	0.690	0.690	0.690	52.562	83.608	22.098	33.282	0.704	0.612	0.655	0.381	
5. INF2 - частный критерий: количество знаний по А/Харкевичу: в...	Семантический резонанс зна...	74	54	0.664	0.730	0.695	100.483	99.075	45.181	42.756	0.690	0.702	0.696	0.688	
5. INF2 - частный критерий: количество знаний по А/Харкевичу: в...	Сумма знаний	62	62	0.690	0.690	0.690	52.562	83.608	22.098	33.282	0.704	0.612	0.655	0.381	
6. INF3 - частный критерий: Унквadrat, разности между фактич...	Семантический резонанс зна...	65	65	0.675	0.675	0.675	91.425	91.425	38.861	38.861	0.702	0.702	0.702	0.677	
6. INF3 - частный критерий: Унквadrat, разности между фактич...	Сумма знаний	65	65	0.675	0.675	0.675	91.425	91.425	38.861	38.861	0.702	0.702	0.702	0.677	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	65	65	0.675	0.675	0.675	104.196	104.196	45.594	45.594	0.696	0.696	0.696	0.772	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	62	62	0.690	0.690	0.690	89.455	89.805	37.189	36.839	0.706	0.708	0.707	0.648	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	65	65	0.675	0.675	0.675	104.196	104.196	45.594	45.594	0.696	0.696	0.696	0.772	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	62	62	0.690	0.690	0.690	89.455	89.805	37.189	36.839	0.706	0.708	0.707	0.648	
9. INF6 - частный критерий: разн усли и безуслов вероятностей; вер...	Семантический резонанс зна...	65	65	0.675	0.675	0.675	91.425	91.425	38.861	38.861	0.702	0.702	0.702	0.677	
9. INF6 - частный критерий: разн усли и безуслов вероятностей; вер...	Сумма знаний	65	65	0.675	0.675	0.675	89.152	88.373	38.148	37.310	0.700	0.705	0.703	0.660	
10. INF7 - частный критерий: разн усли и безуслов вероятностей; ве...	Семантический резонанс зна...	65	65	0.675	0.675	0.675	91.425	91.425	38.861	38.861	0.702	0.702	0.702	0.677	
10. INF7 - частный критерий: разн усли и безуслов вероятностей; ве...	Сумма знаний	65	65	0.675	0.675	0.675	89.152	88.373	38.148	37.310	0.700	0.705	0.703	0.660	

Рисунок 12. Экранные формы режима измерения достоверности моделей 3.4

На рисунках 13 приведены частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4.

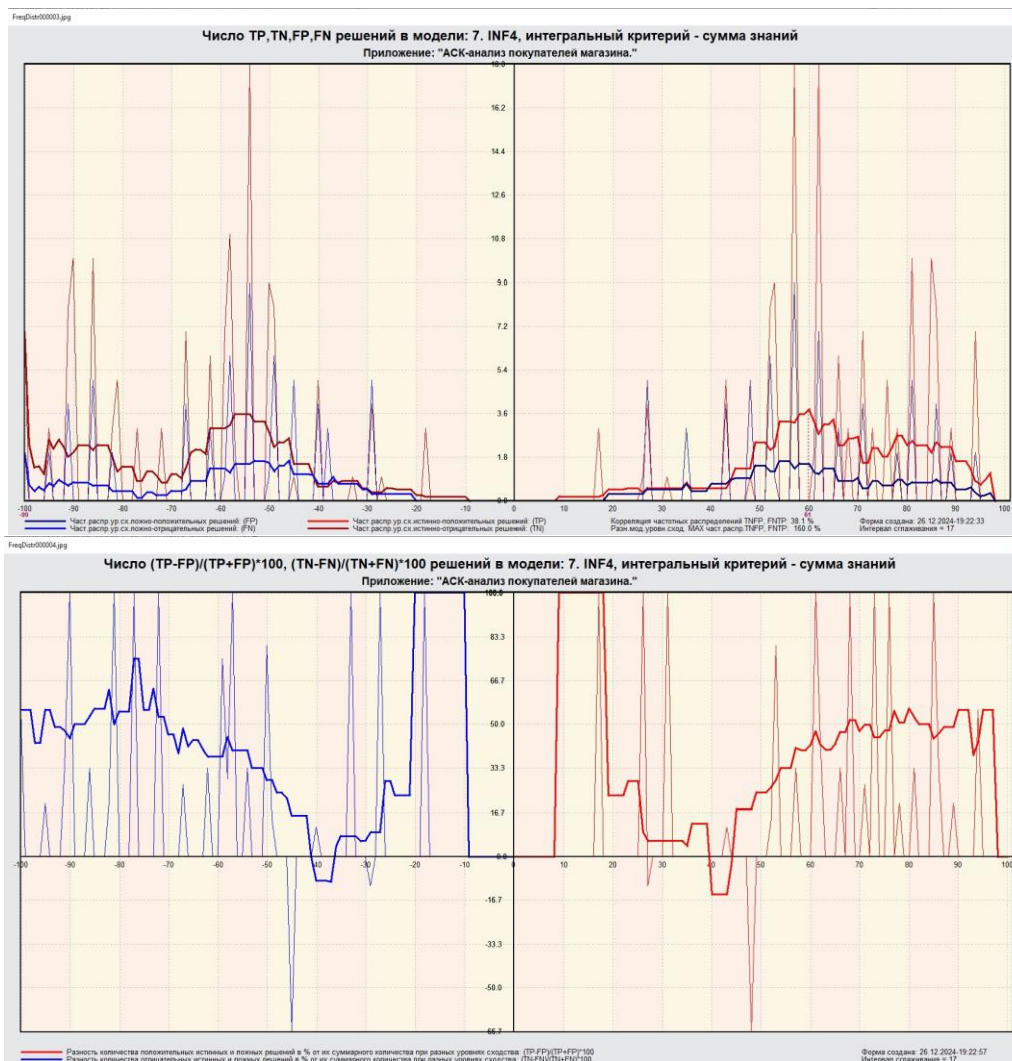


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4

Из этих частотных распределений видно, что в наиболее достоверной по критерию достоверности L1-мерой проф.Е.В.Луценко СК-модели INF4.

На рисунках 14 приведены экранные формы хелпов режима 3.4, в которых подробно объясняется смысл этого режима. Эти формы приводятся в работе вместо более детального описания данного режима.



Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

- идентификация объекта с классом достоверна, т.е. модель относит объекты к классам, к которым они действительно принадлежат;
- прогнозирование достоверно, т.е. действительно наступают те события, которые прогнозируются;
- принятие решений адекватно (достоверно), т.е. после реализации принятых управляющих решений объект управления действительно переходит в целевые будущие состояния;

– исследование достоверно, т.е. полученные в результате исследования модели объекта моделирования выводы могут быть с полным основанием отнесены к объекту моделирования.

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (рисунки 15). Это необходимо делать лишь для решения задачи идентификации и прогнозирования (в режиме 4.1.2), которая требует и потребляет наибольшие вычислительные ресурсы и поэтому решается только для модели, заданной в качестве текущей. Все остальные расчеты проводятся в системе «Эйдос» сразу во всех моделях.

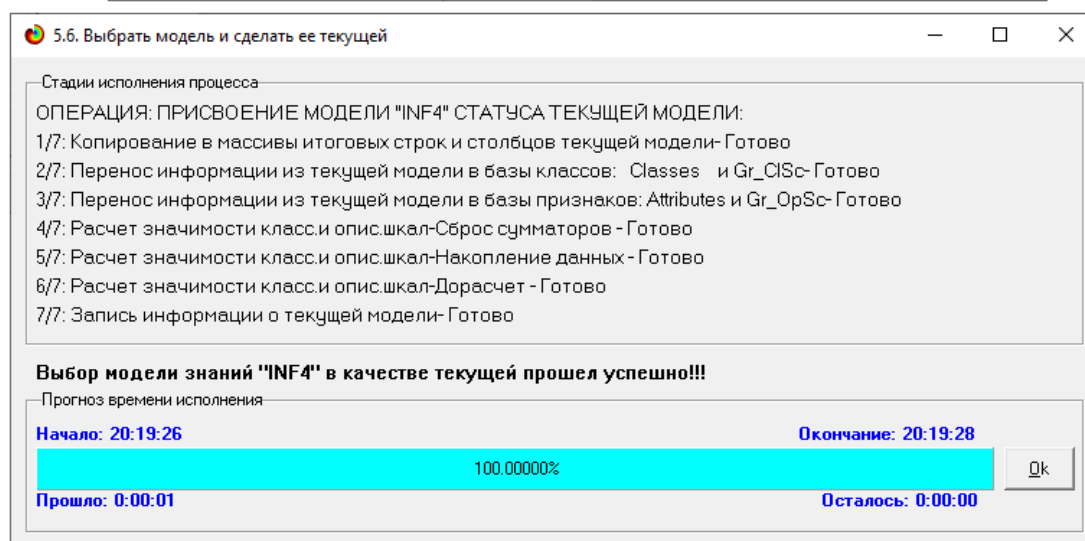
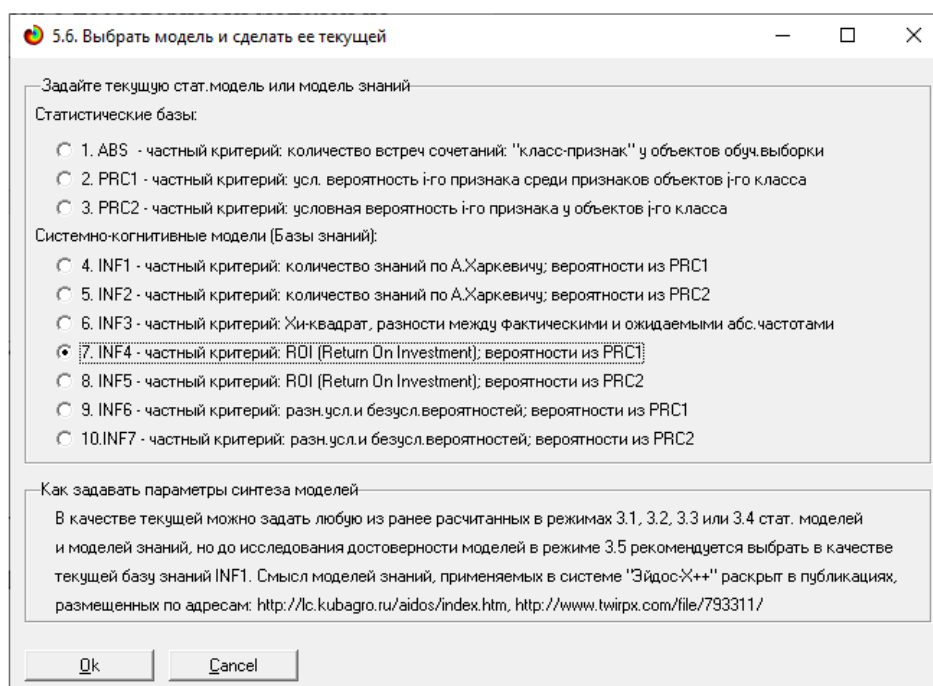


Рисунок 15. Задание СК-модели INF4 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

При решении *задачи идентификации* каждый объект распознаваемой выборки сравнивается по всем своим признакам с каждым из обобщенных образов классов. Смысл решения задачи идентификации заключается в том, что при определении принадлежности конкретного объекта к обобщенному образу классу об этом конкретном объекте *по аналогии становится известно все, что известно об объектах этого класса, по крайней мере, самое существенное о них, т.е. чем они отличаются от объектов других классов.*

Задачи идентификации и прогнозирования взаимосвязаны и мало чем отличаются друг от друга. Главное различие между ними в том, что при идентификации значения свойств и принадлежность объекта к классу относятся к одному моменту времени, а при прогнозировании значения факторов относятся к прошлому, а переход объекта под действием этих факторов в состояние, соответствующее классу относится к будущему i (рисунок 3).

Задача решается в модели, заданной в качестве текущей, т.к. является весьма трудоемкой в вычислительном отношении. Правда с разработкой и реализацией в системе «Эйдос» высокоэффективных алгоритмов распознавания и использованием графического процессора (GPU) для расчетов эта проблема практически снялась.

Сравнение осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что корректны¹¹ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1. Интегральный критерий «Сумма знаний»

Интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 5.5:

$$I_j = (\vec{I}_{ij}, \vec{L}_i).$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид:

¹¹ В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

$$I_j = \sum_{i=1}^M I_{ij} L_i,$$

где: M – количество градаций описательных шкал (признаков);

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса;

$\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i - \text{й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i - \text{й фактор действует с истинностью } n; \\ 0, & \text{если } i - \text{й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

3.6.2. Интегральный критерий «Семантический резонанс знаний»

Интегральный критерий «Семантический резонанс знаний»

представляет собой *нормированное* суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 3.3 и имеет вид:

$$I_j = \frac{1}{\sigma_j \sigma_l M} \sum_{i=1}^M (I_{ij} - \bar{I}_j) (L_i - \bar{L}),$$

где:

M – количество градаций описательных шкал (признаков); \bar{I}_j – средняя информативность по вектору класса; \bar{L} – среднее по вектору объекта;

σ_j – среднее квадратичное отклонение частных критериев знаний вектора класса; σ_l – среднее квадратичное отклонение по вектору распознаваемого объекта.

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса; $\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта (состояния или явления), включающий все виды

факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\bar{L}_i = \begin{cases} 1, & \text{если } i - \text{й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i - \text{й фактор действует с истинностью } n; \\ 0, & \text{если } i - \text{й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

Приведенное выражение для интегрального критерия «Семантический резонанс знаний» получается непосредственно из выражения для критерия «Сумма знаний» после замены координат перемножаемых векторов их стандартизированными значениями:

$$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, \quad L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_i}.$$

Поэтому по своей сути он также является скалярным произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применяя сплайнов, в частности линейной интерполяции:

$$I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, \quad L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}},$$

Это позволяет предложить неограниченное количество других видов интегральных критериев. Но результаты их применения едва ли они будут существенно отличаться от уже существующих, поэтому они не реализованы в системе «Эйдос».

3.6.3. Важные математические свойства интегральных критериев

Данные интегральные критерии обладают очень интересными *математическими свойствами*, которые обеспечивают ему важные достоинства:

Во-первых, интегральный критерий имеет **неметрическую** природу, т.е. он является мерой сходства векторов класса и объекта, но не расстоянием между ними, а косинусом угла между ними, т.е. это межвекторное или информационное расстояние. Поэтому его применение является корректным в **неортонормированных** пространствах, которые, как правило, и встречаются на практике и в которых применение Евклидова расстояния (теоремы Пифагора) является некорректным.

Во-вторых, данный интегральный критерий является **фильтром**, подавляющим белый шум, который всегда присутствует в эмпирических исходных данных и в моделях, созданных на их основе. Это свойство подавлять белый шум проявляется у данного критерия тем ярче, чем больше в модели градаций описательных шкал.

В-третьих, интегральный критерий сходства представляет собой количественную меру сходства/различия конкретного объекта с обобщенным образом класса и имеет тот же смысл, что и **функция принадлежности** элемента множеству в нечеткой логике Лотфи Заде. **Однако** в нечеткой логике эта функция задается исследователем априорно путем выбора из нескольких возможных вариантов, а в АСК-анализе и его программном инструментарии – интеллектуальной системе «Эйдос» она рассчитывается в соответствии с хорошо обоснованной математической моделью непосредственно на основе эмпирических данных.

В-четвертых, кроме того значение интегрального критерия сходства представляет собой адекватную самооценку **степени уверенности** системы в положительном или отрицательном решении о принадлежности/непринадлежности объекта к классу или **риска ошибки** при таком решении.

В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется **вес** каждого обобщенного образа класса в образе объекта, что подробнее описано в работах [2, 3, 4].

На рисунках 17 приведены экранные формы режима идентификации и прогнозирования 4.1.2 системы «Эйдос»:

3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос»

В АСК-анализе разработаны, а в системе «Эйдос» реализованы развитые методы прогнозирования, основанные на сценарном методе АСК-анализа или сценарном АСК-анализе. Но в задачи данной работы не входит их подробное рассмотрение, тем более, что они подробно освещены и на теоретическом уровне, и с детальными численными примерами в работах [4-7] и в ряде других.

Поэтому в данной работе рассмотрим стандартный несколько упрощенный вариант решения задачи идентификации и прогнозирования. Запустим режим 4.1.2 системы «Эйдос» (рисунок 16):

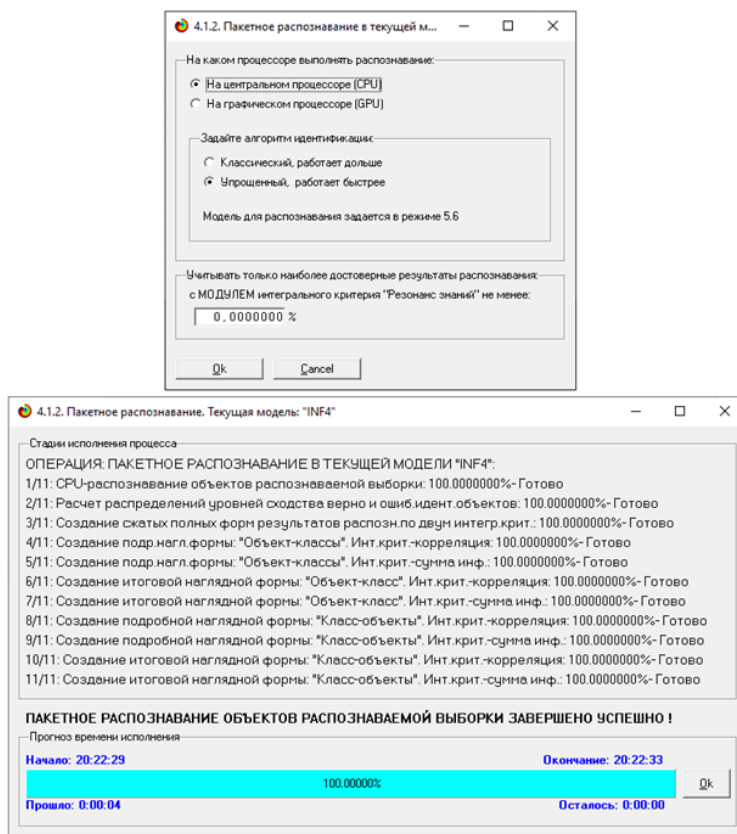


Рисунок 16. Экранные формы режима 4.1.2 идентификации и прогнозирования

По результатам решения задачи идентификации и прогнозирования в системе «Эйдос» выдается большое количество разнообразных выходных форм, которых в настоящее время 12 (рисунок 17):

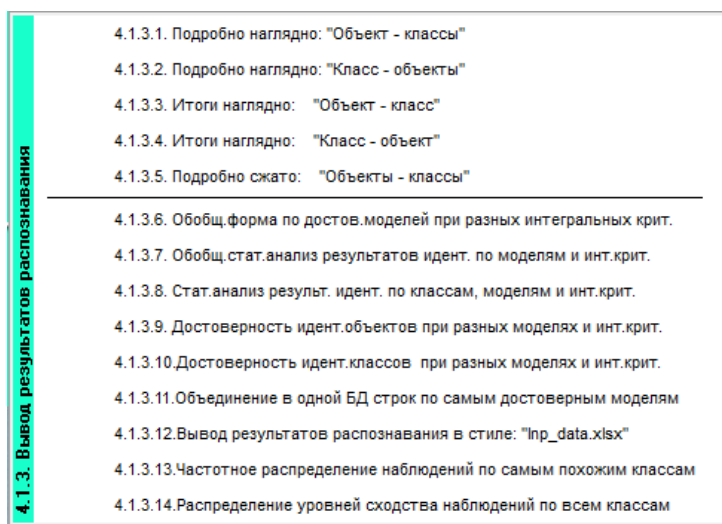


Рисунок 17. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

Из этих выходных форм рассмотрим только две: 4.1.3.1 и 4.1.3.2 (рисунок 18):

4.1.3.1. Визуализация результатов распознавания в отношении: "Объект-классы". Текущая модель: "INF4"

Код	Наим. объекта
1	0001
2	0002
3	0003
4	0004
5	0005
6	0006
7	0007
8	0008
9	0009
10	0010
11	0011
12	0012
13	0013
14	0014
15	0015
16	0016
17	0017
18	0018
19	0019
20	0020
21	0021
22	0022
23	0023

Код	Наименование класса	Сходство	Ф...	Сходство
2	SPENDING SCORE-Большое	71,24...		
1	SPENDING SCORE-Малое	-71,24...	v	

Код	Наименование класса	Сходство	Ф...	Сходство
2	SPENDING SCORE-Большое	71,24...		
1	SPENDING SCORE-Малое	-67,09...	v	

Помощь 9 классов Классы с MaxMin УрСх 9 классов с MaxMin УрСх ВСЕ классы ВКЛ. фильтр по класс. шкале ВЫКЛ. фильтр по класс. шкале Графдиаграммы

4.1.3.2. Визуализация результатов распознавания в отношении: "Класс-объекты". Текущая модель: "INF4"

Код	Наим. класса
1	SPENDING SCORE-Малое
2	SPENDING SCORE-Большое

Код	Наименование объекта	Сходство	Ф...	Сходство
51	0051	100,00...		
67	0067	100,00...	v	
72	0072	100,00...	v	
77	0077	100,00...		
80	0080	100,00...	v	
84	0084	100,00...	v	
94	0094	100,00...	v	
97	0097	100,00...	v	
102	0102	100,00...	v	
118	0118	92,70...		

Код	Наименование объекта	Сходство	Ф...	Сходство
51	0051	94,17...		
67	0067	94,17...	v	
72	0072	94,17...	v	
77	0077	94,17...		
80	0080	94,17...	v	
84	0084	94,17...	v	
94	0094	94,17...	v	
97	0097	94,17...	v	
102	0102	94,17...	v	
118	0118	87,30...		

Помощь Поиск объекта В начало БД В конец БД Предыдущая Следующая 9 записей Все записи Печать XLS Печать TXT Печать ALL

Рисунок 18. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

Эти выходные формы, учитывая сказанное выше об интегральных критериях системы «Эйдос», как говорят «интуитивно понятны» и не требуют особых комментариев.

3.7. Задача-7. Поддержка принятия решений

3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ

Задачи прогнозирования и принятия решений относятся друг к другу как прямая и *обратная* задачи:

– при прогнозировании по значениям факторов, действующих на объект моделирования, определяется в какое будущее состояние он перейдет под их действием;

– при принятии решений, наоборот, по будущему целевому состоянию объекта моделирования определяются значения факторов, которые обуславливают его переход в это будущее целевое состояние.

Таким образом, задача принятия решений является обратной по отношению к задаче прогнозирования. Но это так только в простейшем случае: в случае использования SWOT-анализа (режим 4.4.8 системы «Эйдос») [10] (рисунки 19).

На первом рисунке 19 приведена экранная форма задания в диалоге параметров отображения SWOT-диаграммы. На этой экранной форме в верхнем окне пользователь курсором выбирает исследуемый класс, внизу слева задает модель для исследования, а справа внизу задает отображать ли SWOT-диаграмму. Кроме того, пользователь может включить или выключить фильтры по факторам и посмотреть помощь по режиму. При включении фильтра по фактору, на котором стоит курсор, на экранных формах отображается влияние только значения этого фактора.

Графические выходные формы, приведенные на рисунках 19, интуитивно понятны и не требуют особых комментариев. Отметим лишь, что на SWOT-диаграммах наглядно показаны знак и сила влияния каждого значения фактора на переход объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне. Знак показан цветом, а сила влияния – толщиной линии.

Слева на SWOT-диаграмме приведены значения факторов, способствующих переходу объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне (показаны красным цветом), а справа – препятствующих этому переходу (показаны синим цветом). Сила влияния каждого значения фактора на поведение объекта моделирования показана толщиной линии.

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	SPENDING SCORE-Малое	0,0387133	309	51,5000000
2	SPENDING SCORE-Большое	0,0535418	291	48,5000000

SWOT-анализ класса: 1 "SPENDING SCORE-Малое" в модели: 6 "INF3"

Способствующие факторы и сила их влияния

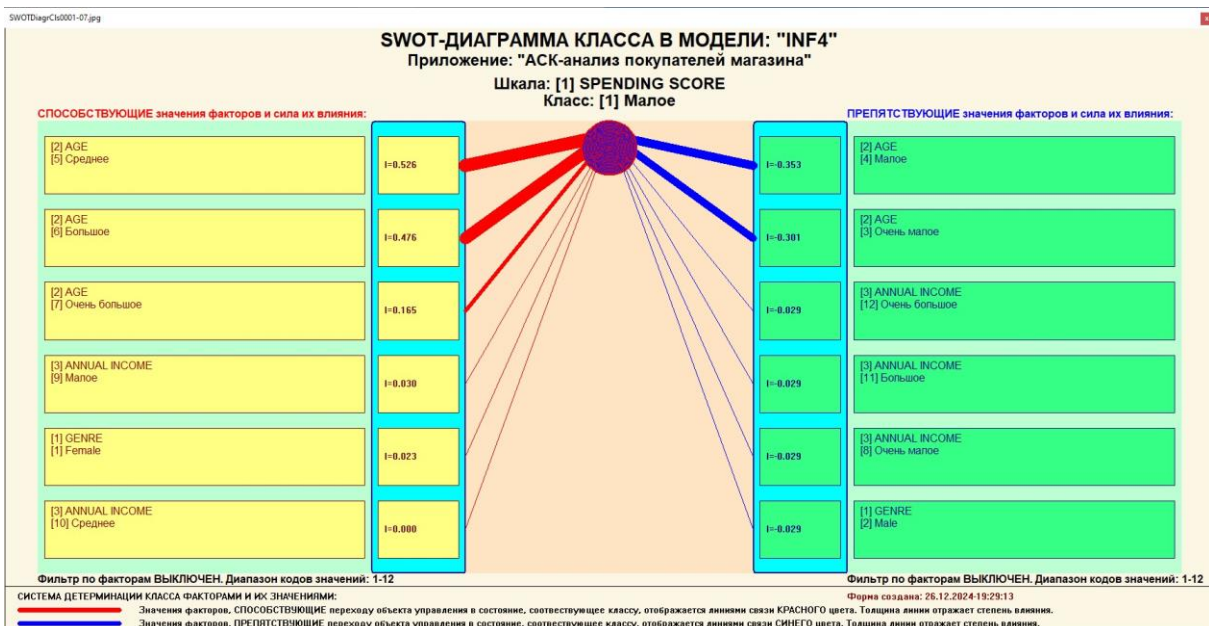
Код	Наименование фактора и его интервального значения	Сила влияния
5	AGE-Среднее	11.370
6	AGE-Большое	6.125
7	AGE-Очень большое	1.700
1	GENRE-Female	1.320
9	ANNUAL INCOME-Малое	1.010
10	ANNUAL INCOME-Среднее	0.010

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	AGE-Малое	-11.445
3	AGE-Очень малое	-7.750
2	GENRE-Male	-1.320
8	ANNUAL INCOME-Очень малое	-0.690
11	ANNUAL INCOME-Большое	-0.210
12	ANNUAL INCOME-Очень большое	-0.120

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма



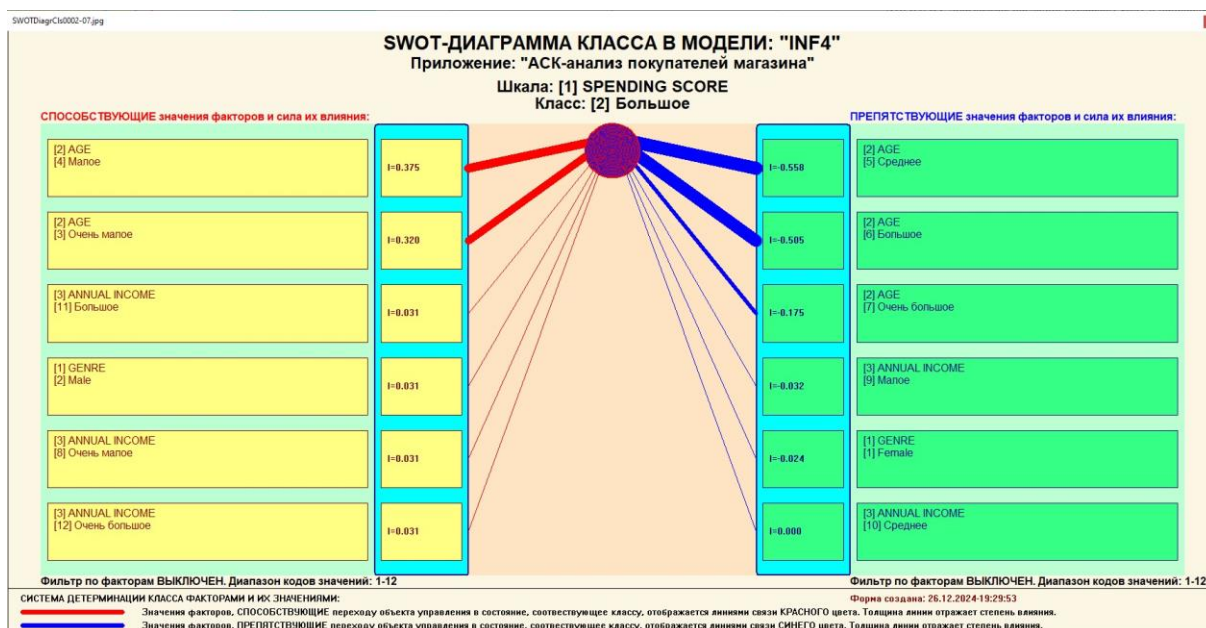


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

Однако SWOT-анализ (режим 4.4.8 системы «Эйдос») имеет свои ограничения: может быть задано только одно будущее целевое состояние, целевые состояния могут быть недостижимыми одновременно (альтернативными) или совместимыми по системе обуславливающих их значений факторов, некоторые рекомендуемые факторы может не быть технологической и финансовой возможности использовать и возможно надо искать им замену, примерно так же влияющую на объект моделирования.

Поэтому в АСК-анализе и системе «Эйдос» реализован развитый алгоритм принятия решений (режим 6.3) в котором кроме SWOT-анализа используются также результаты решения задачи прогнозирования и результаты кластерно-конструктивного анализа классов и значений факторов, т.е. некоторые результаты решения задачи исследования предметной области. Этот алгоритм описан в работах [11, 12, 13] и в ряде других работ.

Приведем этот алгоритм и в данной работе (рисунок 20).

Шаг 1-й. Руководство ставит цели управления, т.е. определяет будущие целевые состояния объекта управления. Обычно целевые состояния в натуральном выражении – это количество и качество продукции, а в стоимостном выражении – прибыль и рентабельность. Объект управления как система, эффективность объекта управления как **системное свойство**, повышение уровня системности объекта управления как цель управления (нелинейность). Модель отражает определенный

уровень технологий, поэтому целевые состояния, недостижимые в одной модели, могут оказаться достижимыми в другой с большим числом факторов [11, 12, 13, 14, 15, 16].

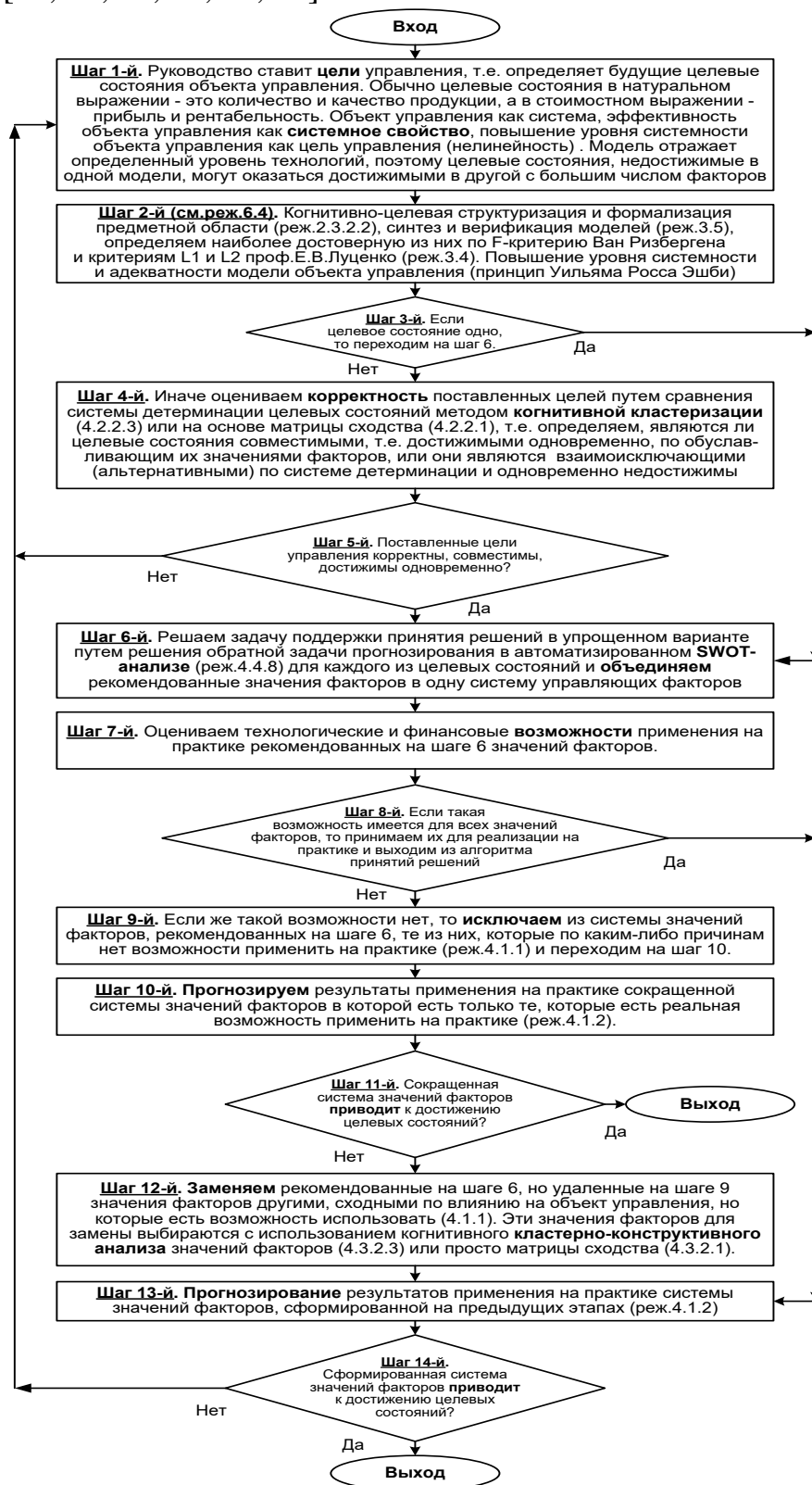


Рисунок 20. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

Шаг 2-й (см.реж.6.4). Когнитивно-целевая структуризация и формализация предметной области (реж.2.3.2.2), синтез и верификация моделей (реж.3.5), определяем наиболее достоверную из них по F-критерию Ван Ризбергена и критериям L1 и L2 проф.Е.В.Луценко (реж.3.4) [9]. Повышение уровня системности и адекватности модели объекта управления (принцип Уильяма Росса Эшби) [14, 15].

Шаг 3-й. Если целевое состояние одно, то переходим на шаг 6, а иначе на шаг 4.

Шаг 4-й. Иначе оцениваем **корректность** поставленных целей путем сравнения системы детерминации целевых состояний методом **когнитивной кластеризации** (4.2.2.3) или на основе матрицы сходства (4.2.2.1), т.е. определяем, являются ли целевые состояния совместимыми, т.е. достижимыми одновременно, по обуславливающим их значениями факторов, или они являются взаимоисключающими (альтернативными) по системе детерминации и одновременно недостижимы.

Шаг 5-й. Поставленные цели управления корректны, совместимы, достижимы одновременно? Если да – переход на шаг 6, иначе на шаг 1.

Шаг 6-й. Решаем задачу поддержки принятия решений в упрощенном варианте путем решения обратной задачи прогнозирования в автоматизированном **SWOT-анализе** (реж.4.4.8) для каждого из целевых состояний и **объединяем** рекомендованные значения факторов в одну систему управляющих факторов [10].

Шаг 7-й. Оцениваем технологические и финансовые **возможности** применения на практике рекомендованных на шаге 6 значений факторов.

Шаг 8-й. Если такая возможность имеется для всех значений факторов, то принимаем их для реализации на практике и переходим на шаг 13 для проверки эффективности принятых решений, а иначе переходим на шаг 9.

Шаг 9-й. Если же такой возможности нет, то **исключаем** из системы значений факторов, рекомендованных на шаге 6, те из них, которые по каким-либо причинам нет возможности применить на практике (реж.4.1.1) и переходим на шаг 10.

Шаг 10-й. **Прогнозируем** результаты применения на практике сокращенной системы значений факторов в которой есть только те, которые есть реальная возможность применить на практике (реж.4.1.2).

Шаг 11-й. Сокращенная система значений факторов **приводит** к достижению целевых состояний? Если да, то выход из алгоритма принятия решений, а иначе переход на шаг 12.

Шаг 12-й. **Заменяем** рекомендованные на шаге 6, но удаленные на шаге 9 значения факторов другими, сходными по влиянию на объект управления, но которые есть возможность использовать (4.1.1). Эти значения факторов для замены выбираются с использованием

когнитивного кластерно-конструктивного анализа значений факторов (4.3.2.3) или просто матрицы сходства (4.3.2.1) [16].

Шаг 13-й. Прогнозирование результатов применения на практике системы значений факторов, сформированной на предыдущих этапах (реж.4.1.2)

Шаг 14-й. Сформированная система значений факторов **приводит** к достижению целевых состояний? Если да – то выход из алгоритма принятия решений, а иначе переход на шаг 1.

Приведенный выше развитый алгоритм принятия решений может быть применен в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос». Принципиальная схема подобной системы приведена на рисунке 21:

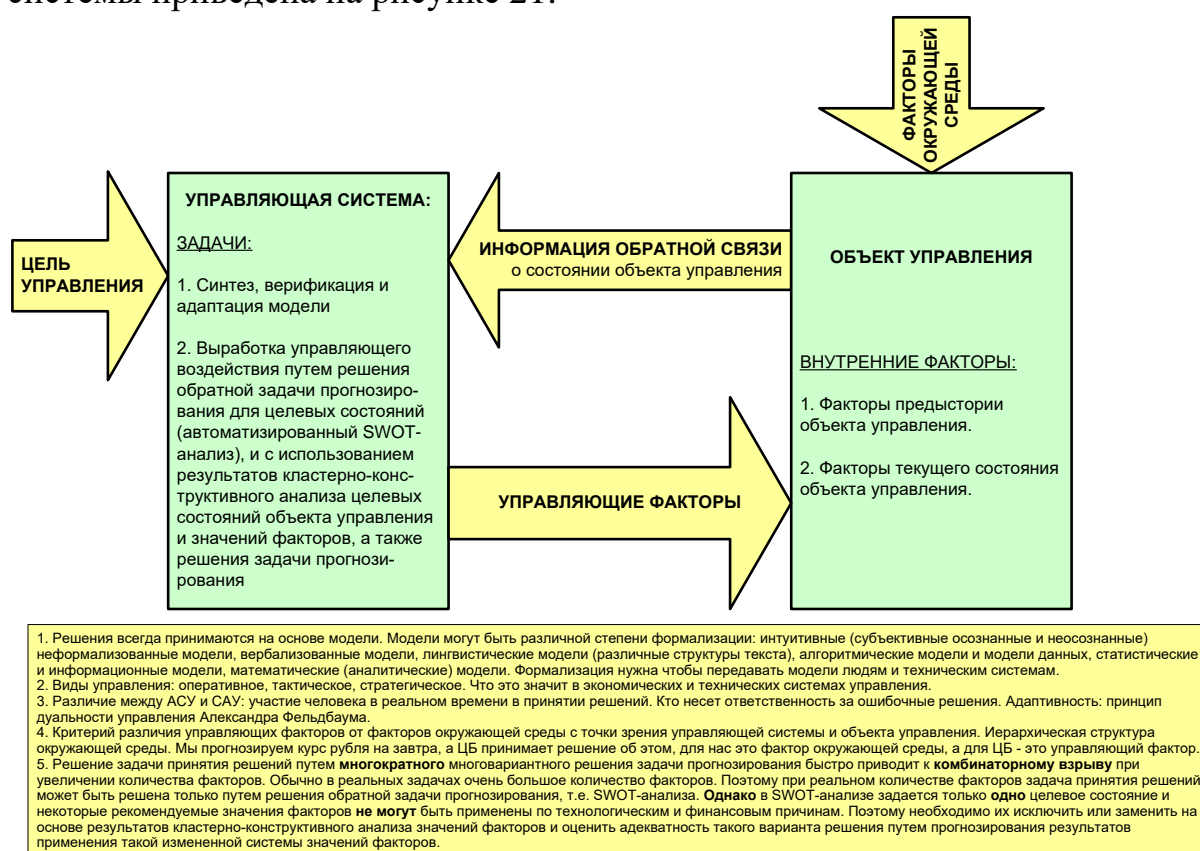


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

Как мы видим в развитом алгоритме принятия решений широко используются результаты решения различных задач: и задачи прогнозирования, и некоторых задач исследования объекта моделирования путем исследования его модели. Необходимо особо отметить, что система «Эйдос» поддерживает решение всех задач, которые необходимо решать в развитом алгоритме принятия решений. Потому ниже кратко рассмотрим решение этих и некоторых других задач.

Необходимо отметить, что модели системы Эйдос – это феноменологические модели, отражающие эмпирические закономерности

в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [12, 17].

3.8. Задача-8. Исследование объекта моделирования путем исследования его модели

3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)

Инвертированные SWOT-диаграмм (предложены автором в работе [10]), отражают силу и направление влияния конкретной градации описательной шкалы на переход объекта моделирования в состояния, соответствующие градациям классификационных шкал (классы). Это и есть *смысл* (семантический потенциал) этой градации описательной шкалы. Инвертированные SWOT-диаграммы выводятся в режиме 4.4.9 системы «Эйдос».

Примеры инвертированных SWOT-диаграмм приведены на рисунках 22 для некоторых значений факторов:

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами АСК-анализа в системе "Эйдос"

Выбор значения фактора, оказывающего влияние на переход объекта управления в будущие состояния

Код	Наименование значения фактора
1	GENRE-Female
2	GENRE-Male
3	AGE-Очень малое
4	AGE-Малое
5	AGE-Среднее
6	AGE-Большое

SWOT-анализ значения фактора:1 "GENRE-Female" в модели:6 "INF3"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влияния
1	SPENDING SCORE-Малое	1.320

ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
2	SPENDING SCORE-Большое	-1.320

ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма

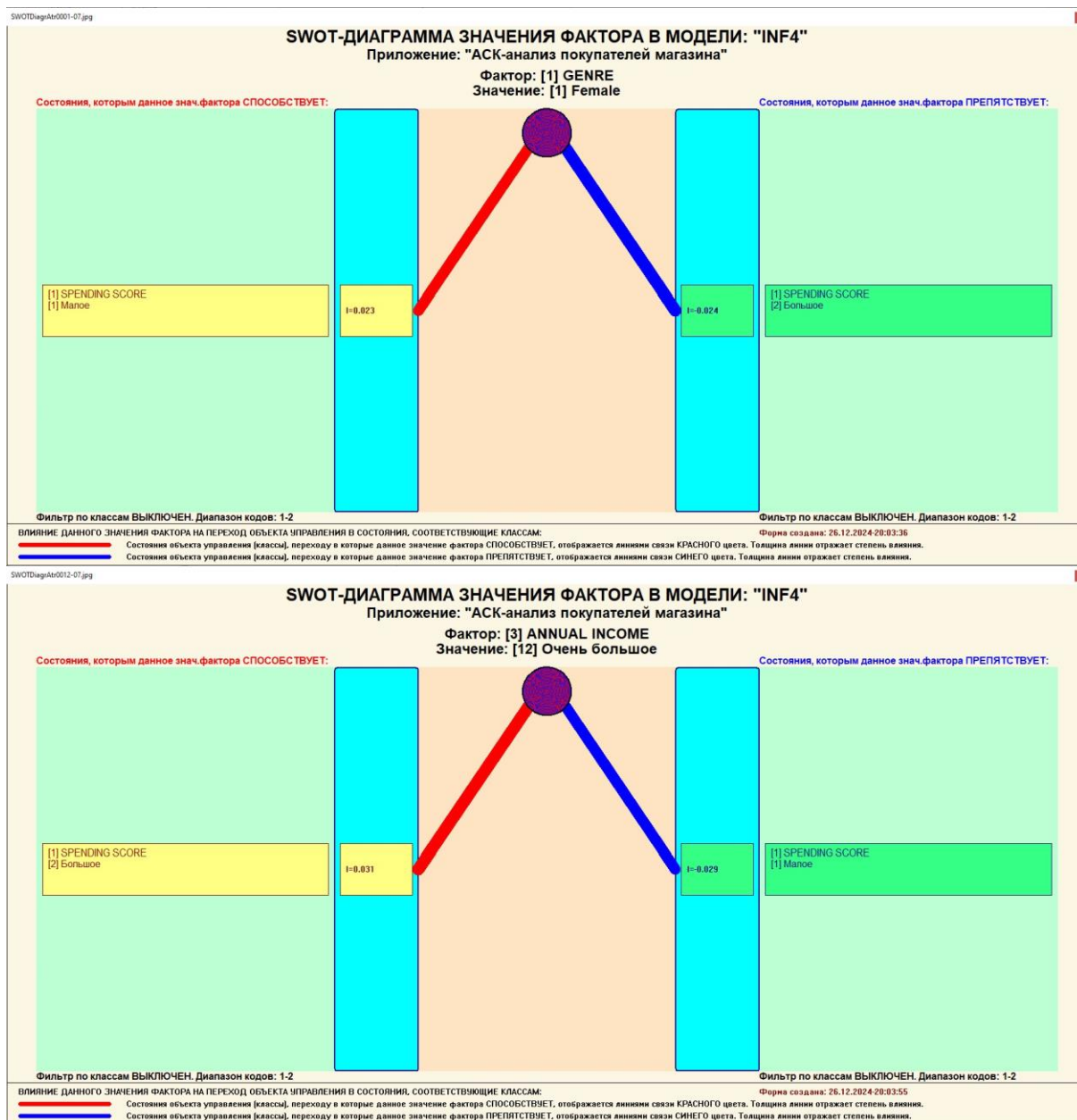


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 22 инвертированные SWOT-диаграммы исчерпывающим образом отражают силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в состояния, соответствующие различным классам. Во многом это и есть решение проблемы, поставленной в работе.

3.8.2. Кластерно-конструктивный анализ классов

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 23) рассчитывается матрица сходства классов (таблица 14) по системе их детерминации и на основе этой матрицы рассчитывается и выводится три основных формы:

- круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2) (рисунок 24);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации классов* (предложена автором в 2011 году в работе [16]) (режим 4.2.2.3) (рисунок 26);
- график изменения межкластерных расстояний (режим 4.2.2.3) (рисунок 25).

Эта матрица сходства (таблица 14) используются и при расчете некоторых других выходных форм.

На рисунке 23 представлены экранные формы режима 4.2.2.1, обеспечивающего расчет матрицы сходства классов по системе их детерминации, т.е. по обуславливающим их значениям факторов:

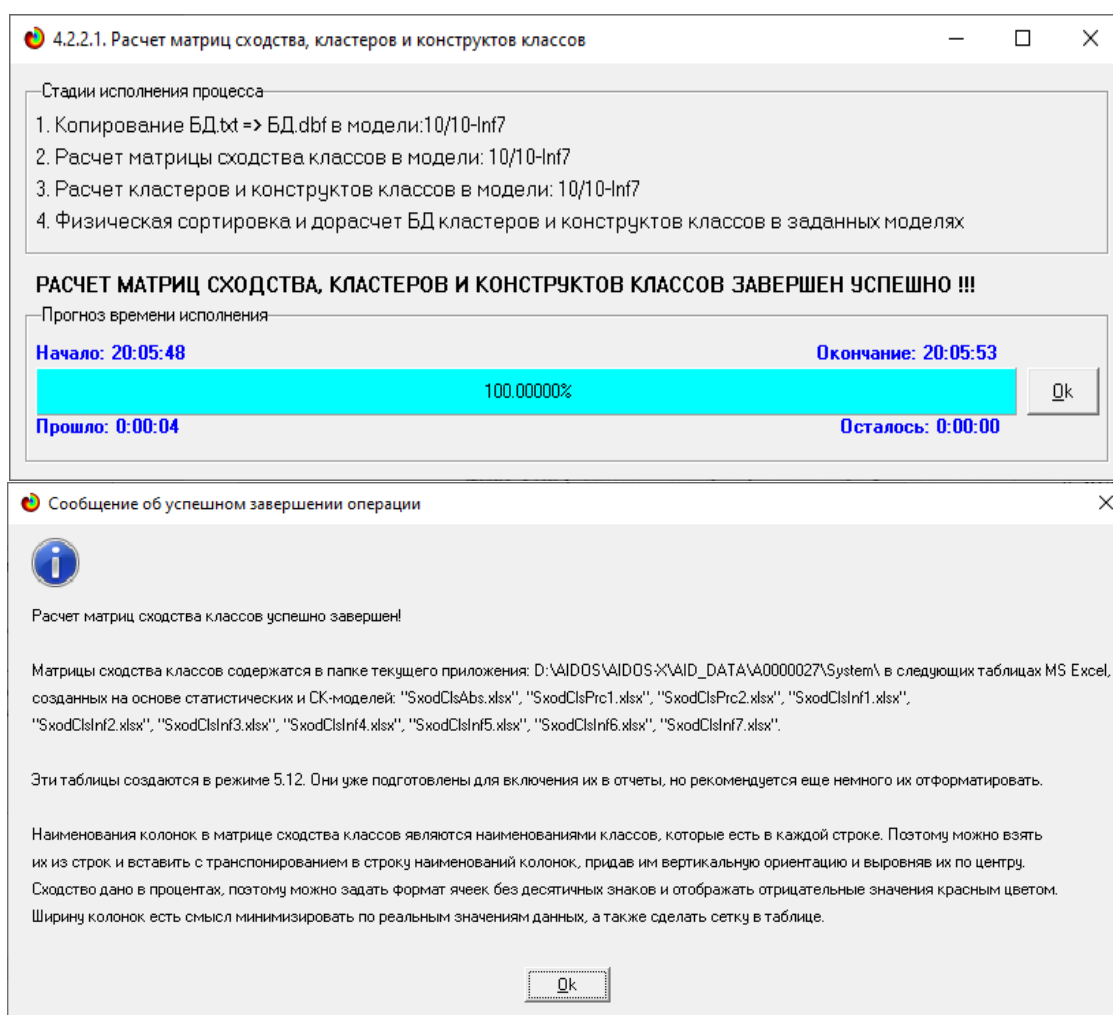


Рисунок 23. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

Таблица 14 – Матрица сходства классов в СК-модели INF4

	A	B	C	D	E	F
	kod_cls	kod_clsc	name_cls	n1	n2	
	1	1	SPENDING SCORE-Малое	100	-46,38274218	
	2	1	SPENDING SCORE-Большое	-46,38274218	100	



Рисунок 24. Круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2)

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 27) рассчитывается матрица сходства признаков (таблица 15) по их смыслу и на основе этой матрицы рассчитывается и выводится четыре основных формы:

- круговая 2d-когнитивная диаграмма признаков (режим 4.3.2.2) рисунок 28);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации признаков* (предложена автором в 2011 году в работе [16]) (режим 4.3.2.3) рисунок 29);
- график изменения межкластерных расстояний (режим 4.3.2.3) рисунок 30).

Эта матрица сходства (таблица 15) используются и при расчете некоторых других выходных форм.

На рисунке 27 представлены экранные формы режима 4.3.2.1, обеспечивающего расчет матрицы сходства значений факторов по силе и направлению их влияния на переход объекта моделирования в различные будущие состояния, соответствующие классам:

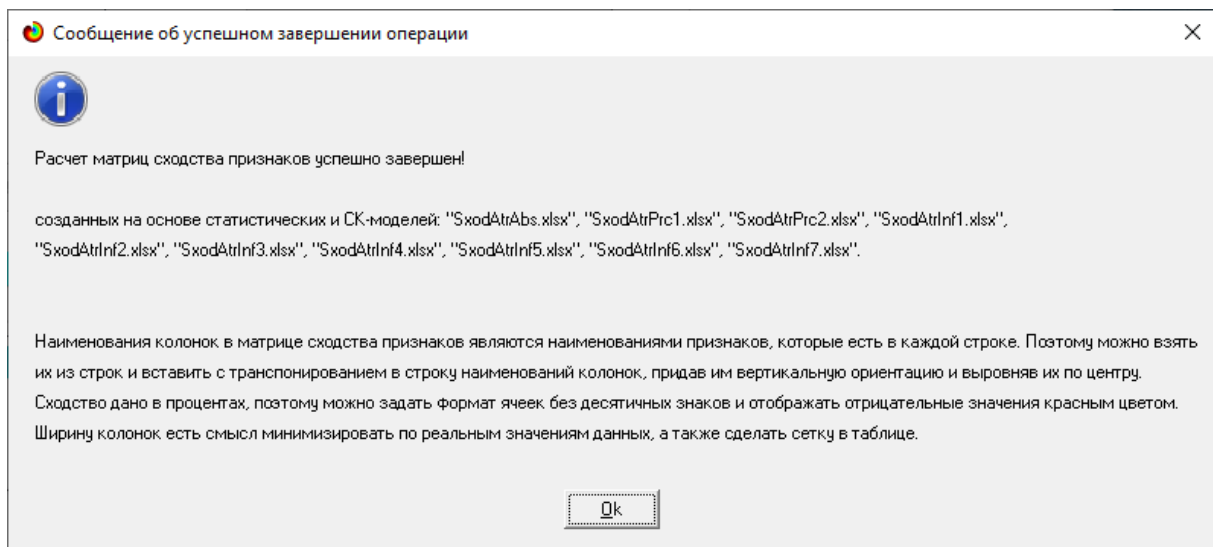
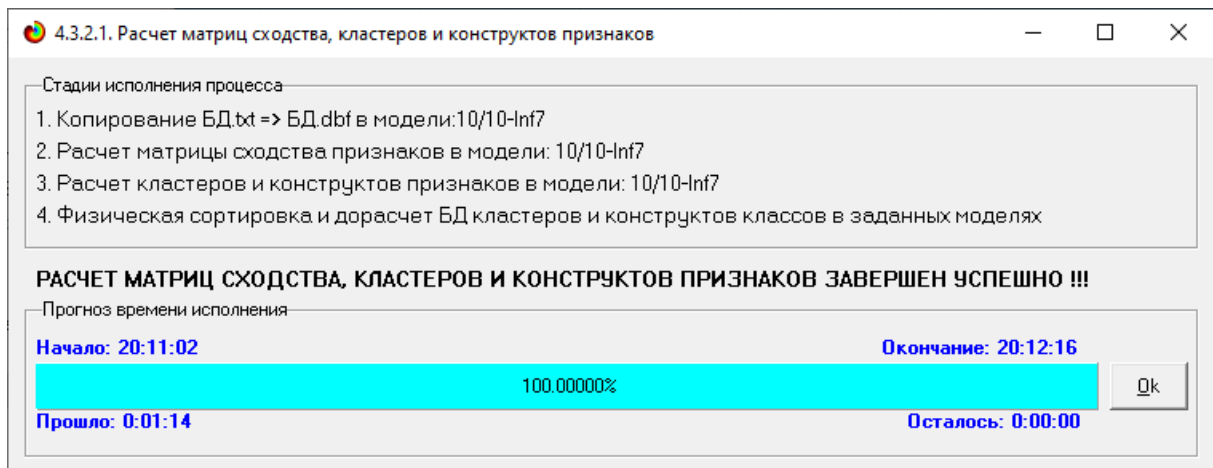


Рисунок 25. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

Таблица 15 – Матрица сходства значений факторов в СК-модели INF4

KOD_ATR	KOD_OPSC	NAME_ATR	N1	N2	N3	N4	N5
1	1	GENRE-Female	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
2	1	GENRE-Male	-100,0000000	100,0000000	100,0000000	100,0000000	-100,
3	2	AGE-Очень малое	-100,0000000	100,0000000	100,0000000	100,0000000	-100,
4	2	AGE-Малое	-100,0000000	100,0000000	100,0000000	100,0000000	-100,
5	2	AGE-Среднее	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
6	2	AGE-Большое	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
7	2	AGE-Очень большое	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
8	3	ANNUAL INCOME-Очень малое	-100,0000000	100,0000000	100,0000000	100,0000000	-100,
9	3	ANNUAL INCOME-Малое	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
10	3	ANNUAL INCOME-Среднее	100,0000000	-100,0000000	-100,0000000	-100,0000000	100,
11	3	ANNUAL INCOME-Большое	-100,0000000	100,0000000	100,0000000	100,0000000	-100,
12	3	ANNUAL INCOME-Очень больш	-100,0000000	100,0000000	100,0000000	100,0000000	-100,

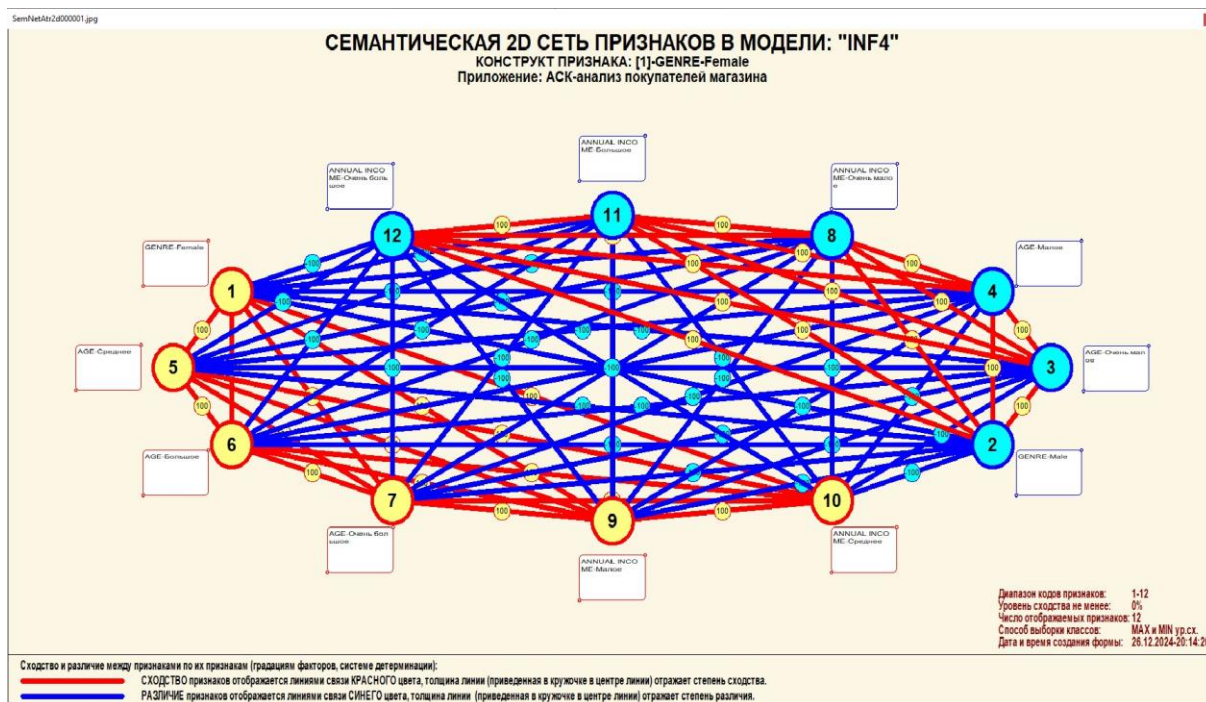


Рисунок 26. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF4 (режим 4.3.2.2)

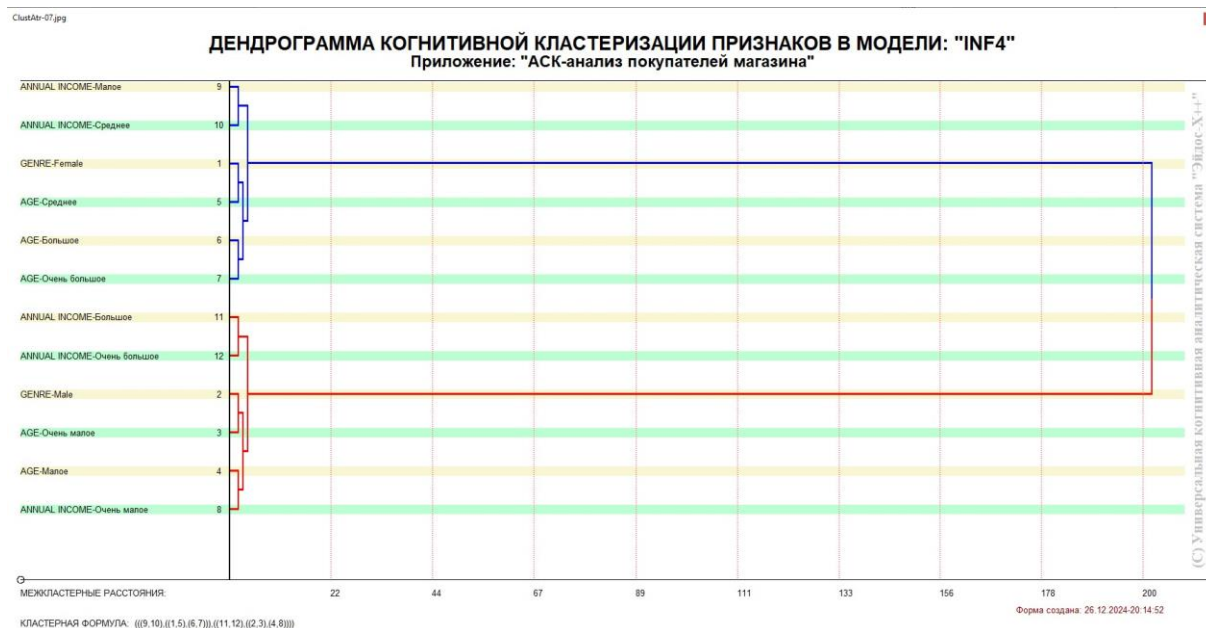


Рисунок 27. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)

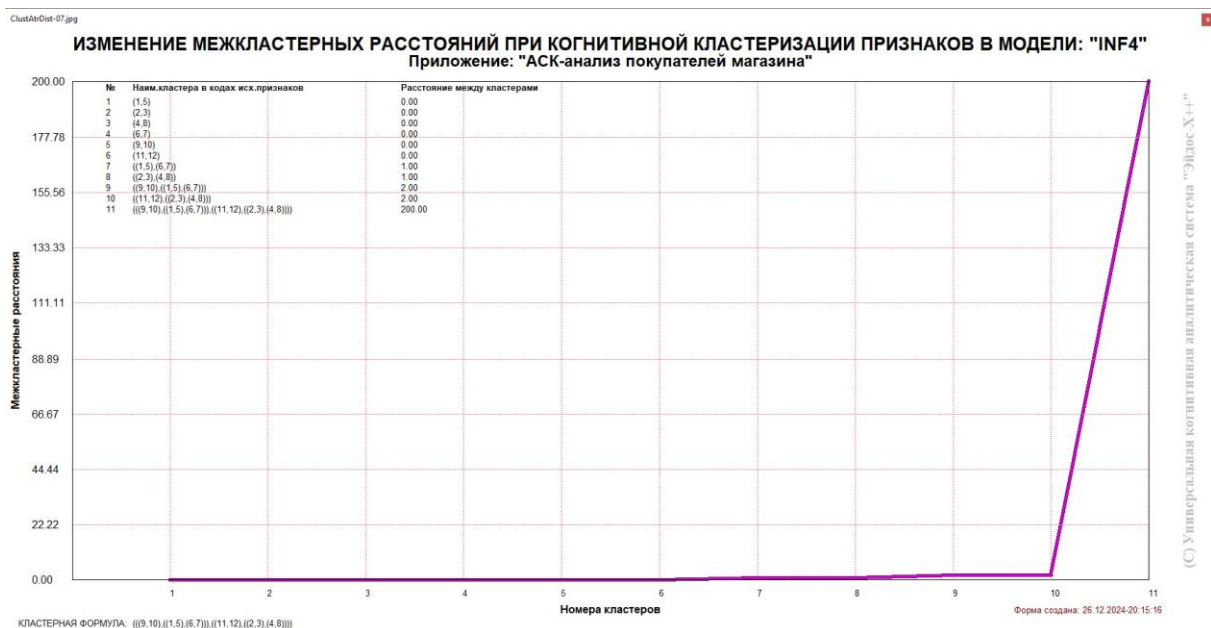


Рисунок 28. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

Модель знаний системы «Эйдос» относится к *нечетким декларативным* гибридным моделям и объединяет в себе некоторые положительные особенности нейросетевой и фреймовой моделей представления знаний.

Классы в этой модели соответствуют нейронам и фреймам, а признаки рецепторам и шпациям (описательные шкалы – слотам).

От фреймовой модели представления знаний модель системы «Эйдос» отличается своей эффективной и простой программной реализацией, полученной за счет того, что разные фреймы отличаются друг от друга не набором слотов и шпаций, а лишь информацией в них. *Поэтому в системе «Эйдос» при увеличении числа фреймов само количество баз данных не увеличивается, а увеличивается лишь их размерность.* Это является очень важным свойством моделей системы «Эйдос», существенно облегчающим и упрощающим программную реализацию.

От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что [18]:

1) весовые коэффициенты на рецепторах не подбираются итерационным методом обратного распространения ошибки, а рассчитываются методом прямого счета на основе хорошо теоретически обоснованной модели, основанной на *теории информации* (это напоминает байесовские сети);

2) весовые коэффициенты имеют хорошо теоретически обоснованную *содержательную интерпретацию*, основанную на теории информации;

3) нейросеть является *нелокальной*, как сейчас говорят «полносвязной».

В системе «Эйдос» нелокальные нейроны визуализируются (режим 4.4.10 системы «Эйдос») в виде специальных графических форм, на которых сила и направление влияния рецепторов нейрона на степень его активации/торможения, которые отображаются соответственно в форме цвета и толщины дендрита (рисунок 31). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

4.4.10.Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	SPENDING SCORE-Малое
2	SPENDING SCORE-Большое

Подготовка визуализации нейрона:1 "SPENDING SCORE-Малое" в модели:6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
5	AGE-Среднее	11.370
6	AGE-Большое	6.125
7	AGE-Очень большое	1.700
1	GENRE-Female	1.320
9	ANNUAL INCOME-Малое	1.010
10	ANNUAL INCOME-Среднее	0.010

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	AGE-Малое	-11.445
3	AGE-Очень малое	-7.750
2	GENRE-Male	-1.320
8	ANNUAL INCOME-Очень малое	-0.690
11	ANNUAL INCOME-Большое	-0.210
12	ANNUAL INCOME-Очень большое	-0.120

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

НЕЙРОН Максимальное количество отображаемых рецепторов: 999
Минимальный вес.коэф.отображаемых рецепторов: 0,000

Сортировать рецепторы:
 по информативности
 по модулю информативности

Отображать рецепторы:
 с наименованиями
 только с кодами

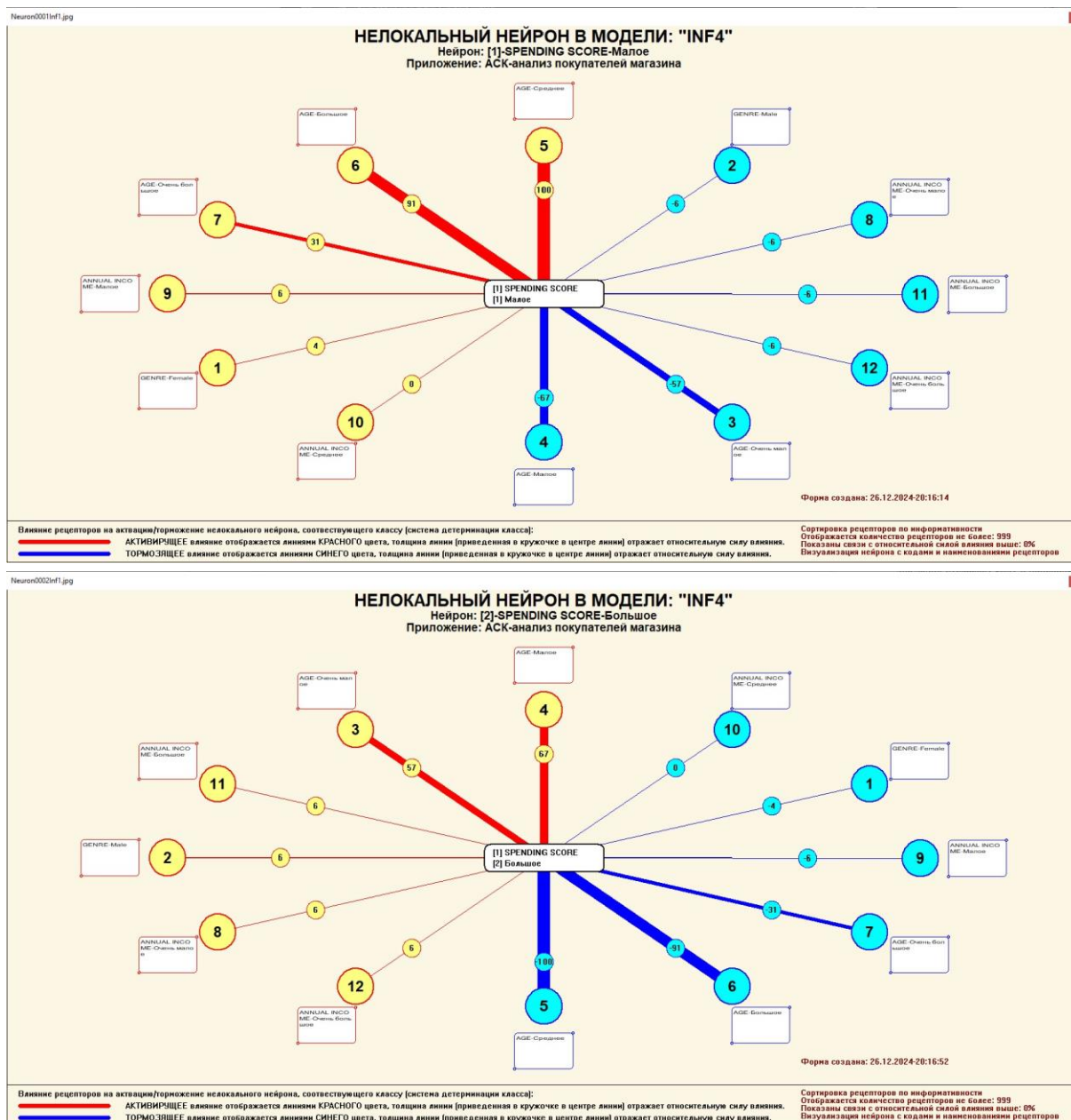


Рисунок 29. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

В системе «Эйдос» есть возможность построения моделей, соответствующих многослойным нейронным сетям [18].

Есть также возможность визуализации любого одного слоя нелокальной нейронной сети (режим 4.4.11 системы «Эйдос»).

Такой слой в наглядной форме отражает силу и направление влияния рецепторов ряда нейрона на степень их активации/торможения в форме цвета и толщины дендритов.

Нейроны на изображении слоя нейронной сети расположены слева направо в порядке убывания модуля суммарной силы их детерминации рецепторами, т.е. слева находятся результаты, наиболее жестко

обусловленные действующими на них значениями факторов, а справа – менее жестко обусловленные (рисунок 32). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

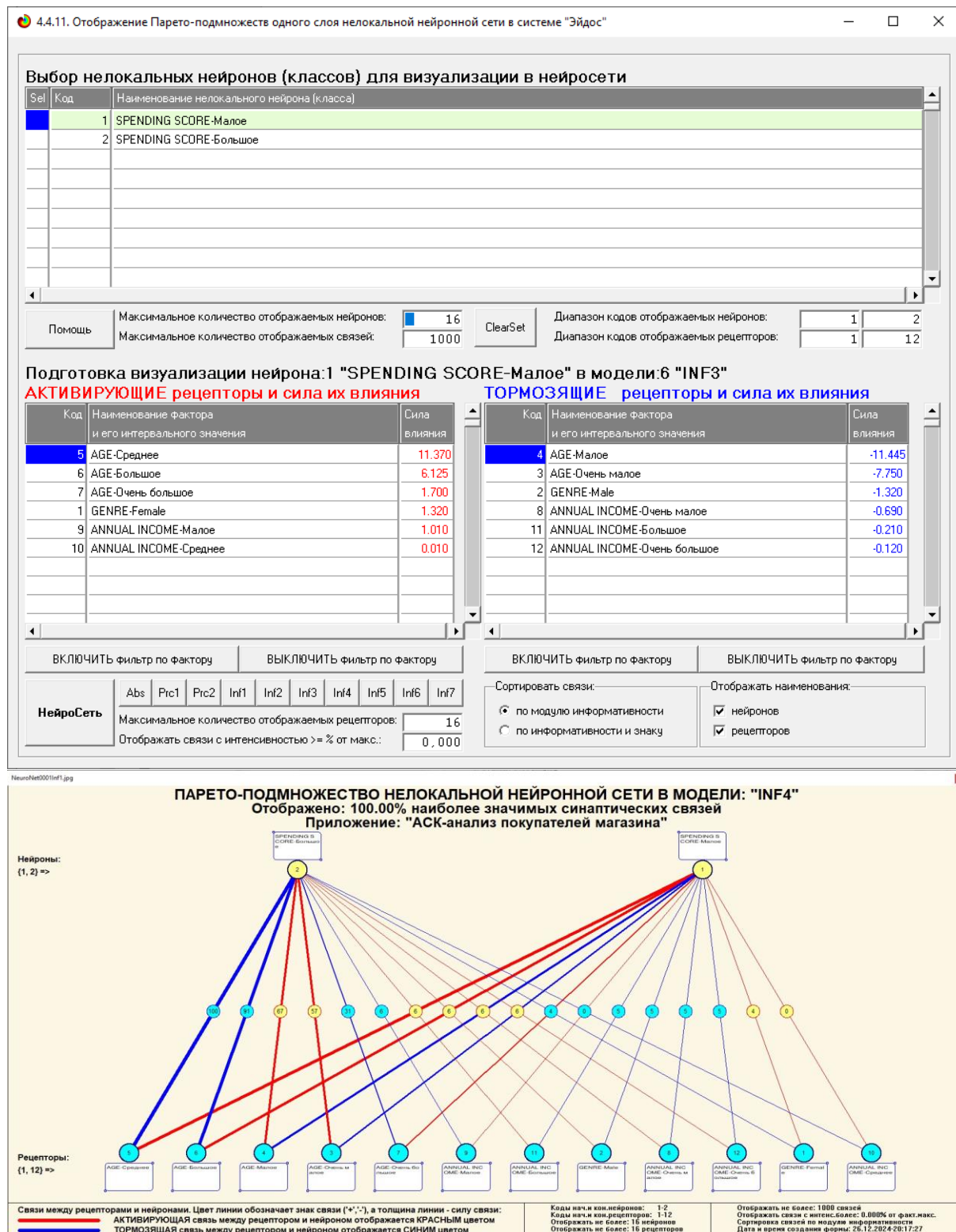


Рисунок 30. Нейронная сеть в СК-модели INF4

3.8.6. 3D-интегральные когнитивные карты

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 24) вверху и когнитивной диаграммы значений факторов (рисунок 28) внизу и соединяющего их одного слоя нейронной сети (рисунок 32) (режим 4.4.12 системы «Эйдос») (рисунок 33):

4.4.12. Отображение Парето-подмножеств одного слоя интегральной когнитивной карты в системе "Эйдос"

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	SPENDING SCORE-Малое
<input type="checkbox"/>	2	SPENDING SCORE-Большое

Помощь Максимальное количество отображаемых нейронов: ClearSet Диапазон кодов отображаемых нейронов: -
 Максимальное количество отображаемых связей: Диапазон кодов отображаемых рецепторов: -

Подготовка визуализации нейрона:1 "SPENDING SCORE-Малое" в модели:6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния **ТОРМОЗЯЩИЕ рецепторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
<input checked="" type="checkbox"/>	5 AGE-Среднее	11.370
<input type="checkbox"/>	6 AGE-Большое	6.125
<input type="checkbox"/>	7 AGE-Очень большое	1.700
<input type="checkbox"/>	1 GENRE-Female	1.320
<input type="checkbox"/>	9 ANNUAL INCOME-Малое	1.010
<input type="checkbox"/>	10 ANNUAL INCOME-Среднее	0.010

Код	Наименование фактора и его интервального значения	Сила влияния
<input checked="" type="checkbox"/>	4 AGE-Малое	-11.445
<input type="checkbox"/>	3 AGE-Очень малое	-7.750
<input type="checkbox"/>	2 GENRE-Male	-1.320
<input type="checkbox"/>	8 ANNUAL INCOME-Очень малое	-0.690
<input type="checkbox"/>	11 ANNUAL INCOME-Большое	-0.210
<input type="checkbox"/>	12 ANNUAL INCOME-Очень большое	-0.120

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Когн. карта Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Максимальное количество отображаемых рецепторов: Отображать связи с интенсивностью >= % от макс.:

Сортировать связи: по модулю информативности по информативности и знаку Отображать наименования: нейронов рецепторов

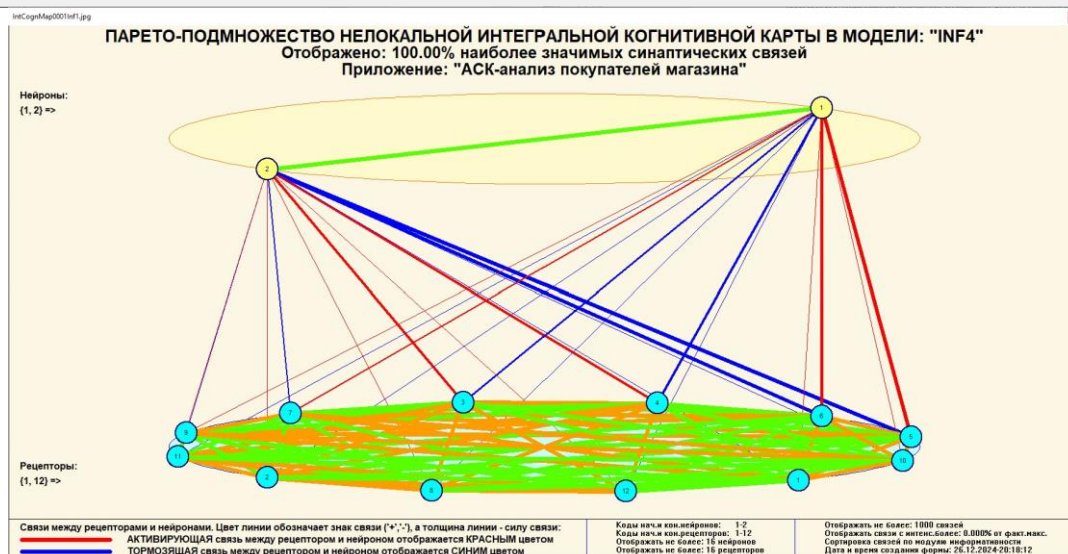


Рисунок 31. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения)

В 2d-когнитивных диаграммах сравнения классов по системе их детерминации видно, насколько сходны или насколько отличаются друг от друга классы по значениям обуславливающих их факторов.

Однако мы не видим из этой диаграммы, чем именно конкретно сходны и чем именно отличаются эти классы по значениям обуславливающих их факторов.

Это мы можем увидеть из когнитивной диаграммы содержательного сравнения классов, которая отображается в режиме 4.2.3 системы «Эйдос».

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых может быть одним из первых писал Дьердь Пойа [19, 20]. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году в работе [2] на странице 521¹². Позже об этом писалось в работе [3]¹³ и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

Пример опосредованных правдоподобных рассуждений.

Допустим нам известно, что один человек имеет голубые глаза, а другой черные волосы. Спрашивается, эти признаки вносят вклад в сходство или в различие этих двух людей? В АСК-анализе и системе «Эйдос» этот вопрос решается так. В модели на основе кластерно-конструктивного анализа классов и значений факторов (признаков) известно, насколько те или иные признаки сходны или отличаются по их влиянию на объект моделирования. Поэтому понятно, что человек с голубыми глазами вероятнее всего блондин, а брюнет, скорее всего, имеет темные глаза. Так что понятно, что эти признаки вносят вклад в различие этих двух людей.

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 34. Всего системой в данной модели генерируется 9409 форм содержательного сравнения классов. Так как каждый из 97 классов сравнивается со всеми остальными, в т.ч. с собой, то всего получается $97^2=9409$ подобных диаграмм. Естественно, все они не приводятся. Но пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-

¹² https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf, Таблица 7. 17, стр. 521

¹³ <http://ej.kubagro.ru/2013/07/pdf/15.pdf>, стр.44.

приложение №391 и получить в нем все выходные формы, как это описано в данной статье.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

Задайте коды двух классов, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	SPENDING SCORE-Малое
2	SPENDING SCORE-Большое

Выбор кода класса левого инф.портрета Выбор кода класса правого инф.портрета

Выбор способа фильтрации признаков в информационных портретах когнитивной диаграммы

Задайте коды двух описательных шкал, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование описательной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ ОПИСАТЕЛЬ	1	12
1	GENRE	1	2
2	AGE	3	7
3	ANNUAL INCOME	8	12

Выбор кода описательной шкалы левого инф.портрета Выбор кода описательной шкалы правого инф.портрета

Задайте модели, в которых проводить расчеты когнитивных диаграмм

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

 [Помощь](#)

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Класс для левого инф.портрета: [0] ВСЕ КЛАССЫ
Класс для правого инф.портрета: [0] ВСЕ КЛАССЫ
Описат.шкала для левого инф.портрета: [0] ВСЕ ОПИСАТЕЛЬ
Описат.шкала для правого инф.портрета: [0] ВСЕ ОПИСАТЕЛЬ
Модели, заданные для расчета: Inf4

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

КОГНИТИВНАЯ ДИАГРАММА КЛАССОВ В МОДЕЛИ: "INF4"

Приложение: "АСК-анализ покупателей магазина"

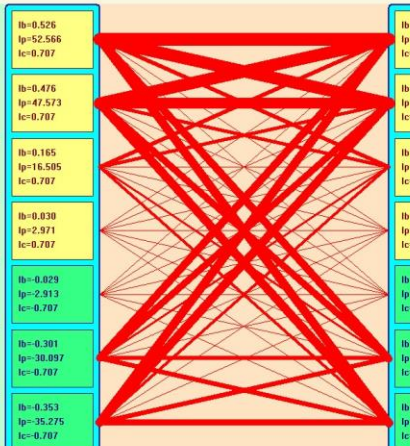
Кл.шкала: [1] SPENDING SCORE
Класс: [1] Малое

Сход./разл.классов: 100.000%

Кл.шкала: [1] SPENDING SCORE
Класс: [1] Малое

Наименования признаков:

[2] AGE [5] Среднее	Ib=-0.526 Ip=52.566 Ic=0.787
[2] AGE [6] Большое	Ib=-0.476 Ip=47.573 Ic=0.787
[2] AGE [7] Очень большое	Ib=-0.165 Ip=16.505 Ic=0.787
[3] ANNUAL INCOME [9] Малое	Ib=-0.830 Ip=2.971 Ic=0.787
[1] GENRE [2] Male	Ib=-0.029 Ip=-2.913 Ic=0.787
[2] AGE [3] Очень малое	Ib=-0.301 Ip=-38.097 Ic=0.787
[2] AGE [4] Малое	Ib=-0.353 Ip=-35.275 Ic=0.787



[2] AGE [5] Среднее	Ib=-0.526 Ip=52.566 Ic=0.787
[2] AGE [6] Большое	Ib=-0.476 Ip=47.573 Ic=0.787
[2] AGE [7] Очень большое	Ib=-0.165 Ip=16.505 Ic=0.787
[3] ANNUAL INCOME [9] Малое	Ib=-0.830 Ip=2.971 Ic=0.787
[1] GENRE [2] Male	Ib=-0.029 Ip=-2.913 Ic=0.787
[2] AGE [3] Очень малое	Ib=-0.301 Ip=-38.097 Ic=0.787
[2] AGE [4] Малое	Ib=-0.353 Ip=-35.275 Ic=0.787

Наименования признаков:

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬ 1-12

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬ 1-12

Сходство и различие между классами по их признакам с учетом сходства/различия между признаками (градация факторов, системы детерминации):

СХОДСТВО классов отображается линиями связи КРАСНОГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень сходства.

РАЗЛИЧИЕ классов отображается линиями связи СИНЕГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень различия.

Формы созданы: 26.12.2024 20:20:04. Показано количество связей < 999

КОГНИТИВНАЯ ДИАГРАММА КЛАССОВ В МОДЕЛИ: "INF4"

Приложение: "АСК-анализ покупателей магазина"

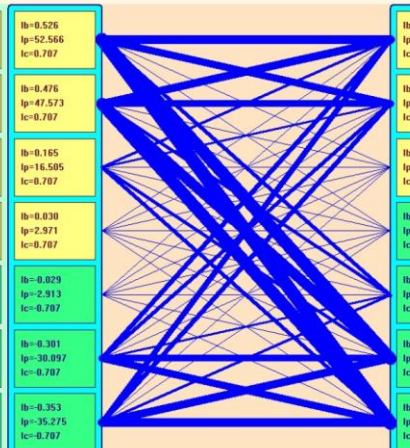
Кл.шкала: [1] SPENDING SCORE
Класс: [1] Малое

Сход./разл.классов: -100.000%

Кл.шкала: [1] SPENDING SCORE
Класс: [2] Большое

Наименования признаков:

[2] AGE [5] Среднее	Ib=-0.526 Ip=52.566 Ic=0.787
[2] AGE [6] Большое	Ib=-0.476 Ip=47.573 Ic=0.787
[2] AGE [7] Очень большое	Ib=-0.165 Ip=16.505 Ic=0.787
[3] ANNUAL INCOME [9] Малое	Ib=-0.830 Ip=2.971 Ic=0.787
[1] GENRE [2] Male	Ib=-0.029 Ip=-2.913 Ic=0.787
[2] AGE [3] Очень малое	Ib=-0.301 Ip=-38.097 Ic=0.787
[2] AGE [4] Малое	Ib=-0.353 Ip=-35.275 Ic=0.787



[2] AGE [4] Малое	Ib=-0.375 Ip=37.457 Ic=0.787
[2] AGE [3] Очень малое	Ib=-0.320 Ip=31.959 Ic=0.787
[1] GENRE [2] Male	Ib=-0.031 Ip=3.093 Ic=0.787
[3] ANNUAL INCOME [9] Малое	Ib=-0.032 Ip=3.155 Ic=0.787
[2] AGE [7] Очень большое	Ib=-0.175 Ip=17.526 Ic=0.787
[2] AGE [8] Большое	Ib=-0.505 Ip=50.515 Ic=0.787
[2] AGE [5] Среднее	Ib=-0.558 Ip=55.817 Ic=0.787

Наименования признаков:

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬ 1-12

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬ 1-12

Сходство и различие между классами по их признакам с учетом сходства/различия между признаками (градация факторов, системы детерминации):

СХОДСТВО классов отображается линиями связи КРАСНОГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень сходства.

РАЗЛИЧИЕ классов отображается линиями связи СИНЕГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень различия.

Формы созданы: 26.12.2024 20:20:06. Показано количество связей < 999

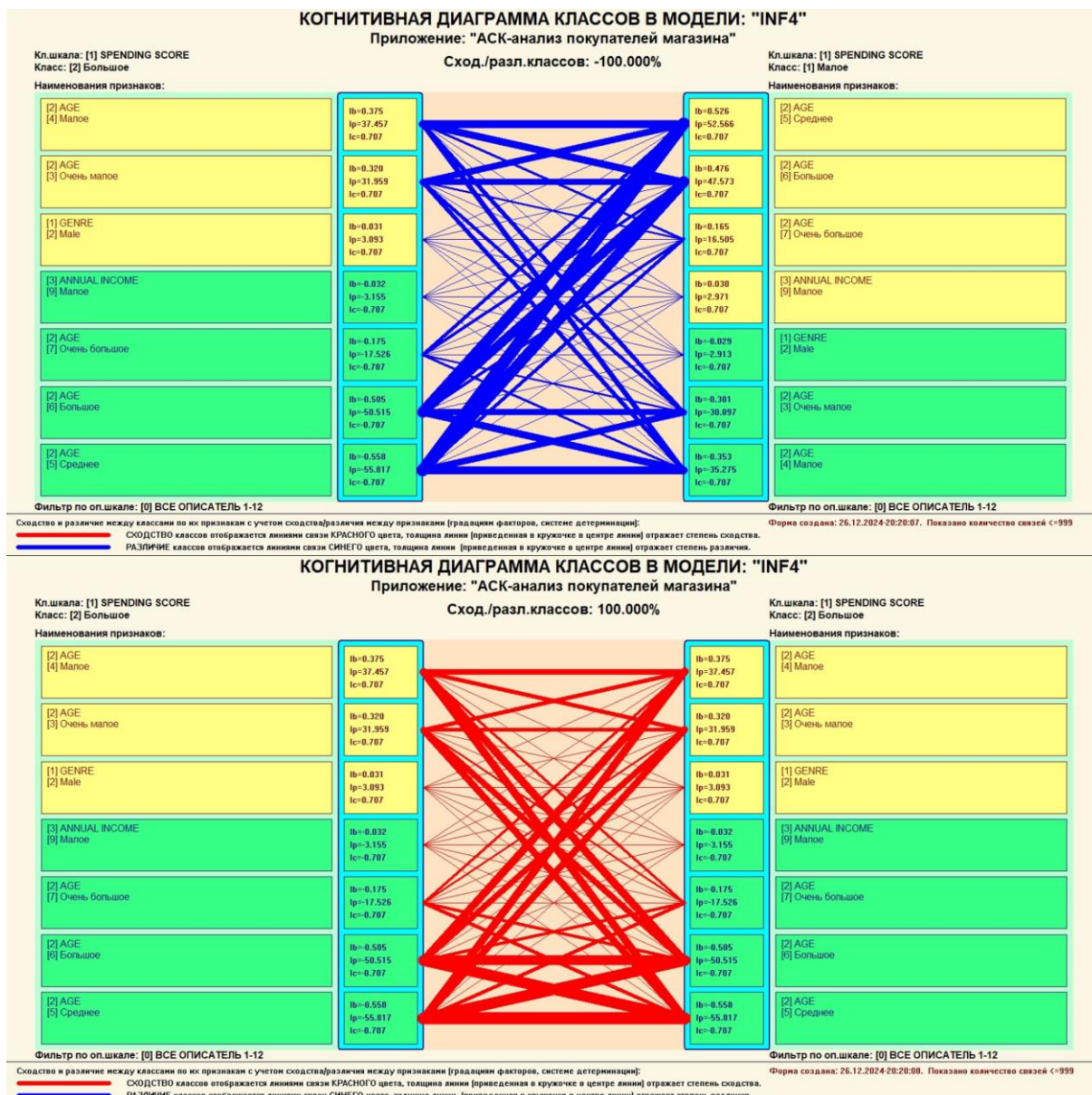


Рисунок 32. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF4

3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения)

Из 2d-когнитивных диаграммах сравнения значений факторов по их влиянию на объект моделирования, т.е. на его переходы в состояния, соответствующие классам вполне понятно, насколько сходны или отличаются любые два значения факторов по их смыслу.

Напомним, что смысл событий, согласно концепции смысла Шенка-Абельсона, используемой в АСК-анализе, состоит в знании причин и последствий этих событий [21].

Однако из этой диаграммы не видно, чем именно конкретно содержательно сходны или отличаются значения факторов по их смыслу.

Это видно из когнитивных диаграмм, которые можно получить в режиме 4.3.3 системы «Эйдос».

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта моделирования в будущие состояния, соответствующие классам, приведены ниже на рисунках 35:

4.3.3. Когнитивные диаграммы признаков. Задание параметров генерации выходных форм

Выбор признаков для когнитивной диаграммы

Задайте коды двух признаков, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	GENRE-Female
2	GENRE-Male
3	AGE-Очень малое
4	AGE-Малое
5	AGE-Среднее
6	AGE-Большое

Выбор кода признака левого инф. портрета Выбор кода признака правого инф. портрета

Выбор способа фильтрации классов в информационных портретах когнитивной диаграммы

Задайте коды двух классификационных шкал, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование классификационной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ	1	2
1	SPENDING SCORE	1	2

Выбор кода классификационной шкалы левого инф. портрета Выбор кода классификационной шкалы правого инф. портрета

Задайте модели, в которых проводить расчеты когнитивных диаграмм:

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Признак для левого инф. портрета: [0] ВСЕ ПРИЗНАКИ
Признак для правого инф. портрета: [0] ВСЕ ПРИЗНАКИ
Классиф. шкала для левого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Классиф. шкала для правого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Модели, заданные для расчета: Inf4

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

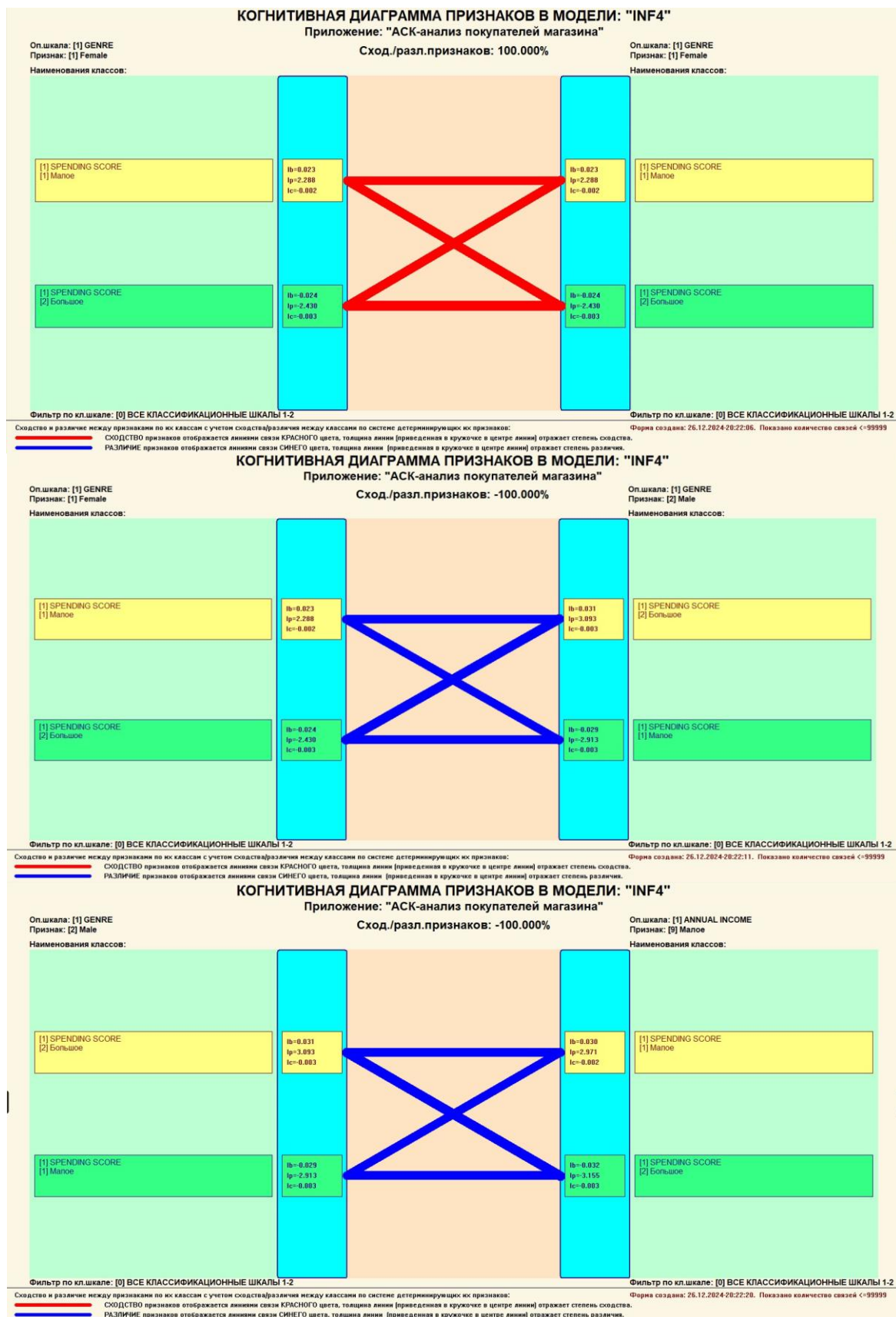


Рисунок 33. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF4

Всего системой в данной модели генерируется $75^2=5625$ подобных диаграмм содержательного сравнения значений факторов по их смыслу, т.е. по влиянию на объект моделирования. Естественно, все они в данной работе не приводятся. Но пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №391 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

Когнитивные функции являются обобщением классического математического понятия функции на основе системной теории информации и предложены Е.В.Луценко в 2005 году [3, 22, 23].

Когнитивные функции отображают, какое количество информации содержится в градациях описательной шкалы о переходе объекта моделирования в состояния, соответствующие градациям классификационной шкалы. При этом в статистических и системно-когнитивных моделях в каждой градации описательной шкалы содержится информация обо всех градациях классификационной шкалы, т.е. *каждому значению аргумента соответствуют все значения функции, но соответствуют в разной степени, причем как положительной, так и отрицательной, которая отображается цветом.*

Когнитивные функции являются одним из наиболее мощных и наглядных средств когнитивной графики, имеющих в системе «Эйдос», позволяющих отобразить силу и направление влияния каждого значения фактора на переход объекта моделирования в каждое из будущих состояний.

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 36). Первая экранная форма данного режима представляет собой краткий хелп, поясняющий смысл понятия «Когнитивная функция», а также позволяющий выйти на экранную форму системы «Эйдос» с действующими гиперссылками на работы по когнитивным функциям, а также страницы сайта автора со списком этих работ и работ автора по выявлению, представлению и использованию знаний, логике и методологии научного познания.

Необходимо отметить, что модели системы «Эйдос» – это *феноменологические* модели, отражающие *эмпирические* закономерности в фактах обучающей выборки, т.е. они отражают причинно-следственные связи, но не отражают *механизма детерминации*, а только сам факт и характер детерминации [17, 23, 24]. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [24].

4.5. Генерация, визуализация и запись когнитивных функций системы "Эйдос"

— □ ×

— Задайте статистические и/или системно-когнитивные модели для генерации когнитивных функций: —

Статистические базы:

- 1. ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч.выборки
- 2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса
- 3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса

Системно-когнитивные модели (Базы знаний):

- 4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1
- 5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2
- 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами
- 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1
- 8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
- 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1
- 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

— Задайте виды когнитивных функций для генерации, визуализации и записи: —

- 1. Сетка триангуляции Делоне без цветовой заливки.
- 2. Сглаженные изолинии триангуляции Делоне без цветовой заливки.
- 3. Сетка триангуляции Делоне с цветовой заливкой.
- 4. Сглаженные изолинии триангуляции Делоне с цветовой заливкой.
- 5. Сглаженная цветочная заливка изолиний с заданным количеством градаций цвета.

— Задайте дополнительные параметры визуализации когнитивных функций: —

- Соединять ли точки с максимальным количеством информации линией КРАСНОГО цвета?
- Соединять ли точки с минимальным количеством информации линией СИНЕГО цвета?

Задайте количество градаций уровня (цвета и изолиний) когнитивных функций:

Задайте количество пикселей на дюйм в изображениях когнитивных функций:

Задайте паузу в секундах между визуализациями когнитивных функций:

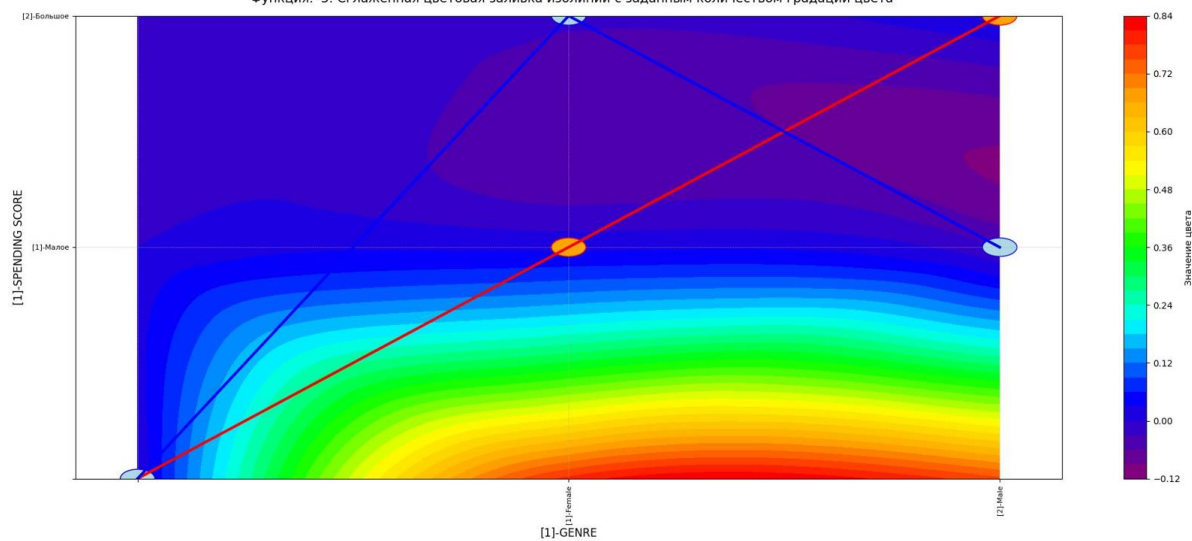
Задайте размер шрифта для наименований градаций шкал X и Y:

Визуализация когнитивных функций new Визуализация когнитивных функций old

Работы по когнитивным функциям-1 Работы по когнитивным функциям-2

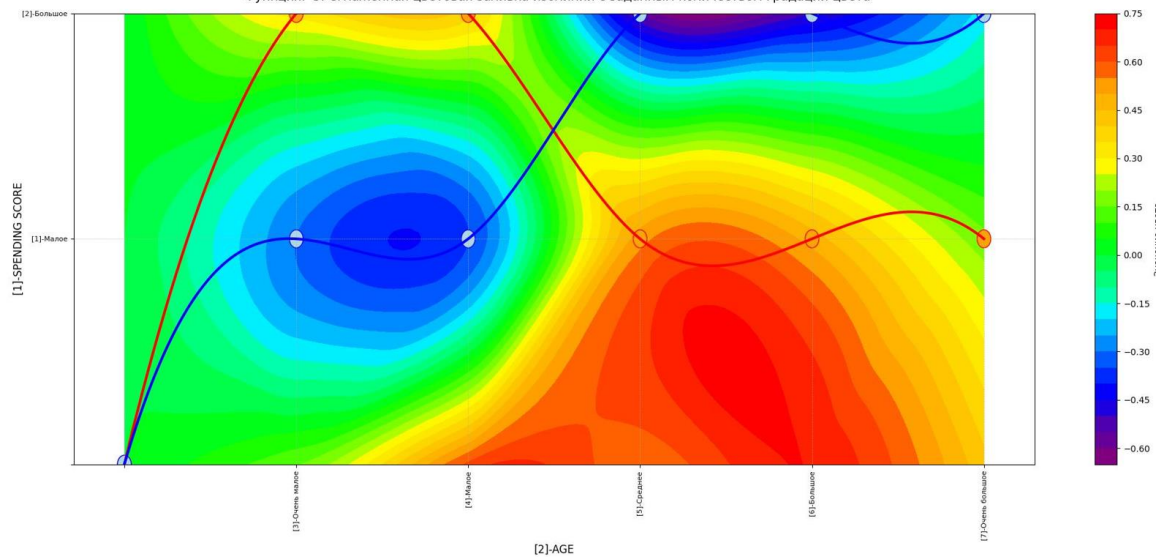
КОГНИТИВНАЯ ФУНКЦИЯ

Приложение: 'АСК-анализ покупателей магазина'
 Модель: '7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC-1'
 Функция: '5. Сглаженная цветовая заливка изолиний с заданным количеством градаций цвета'



КОГНИТИВНАЯ ФУНКЦИЯ

Приложение: 'АСК-анализ покупателей магазина'
 Модель: '7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC-1'
 Функция: '5. Сглаженная цветовая заливка изолиний с заданным количеством градаций цвета'



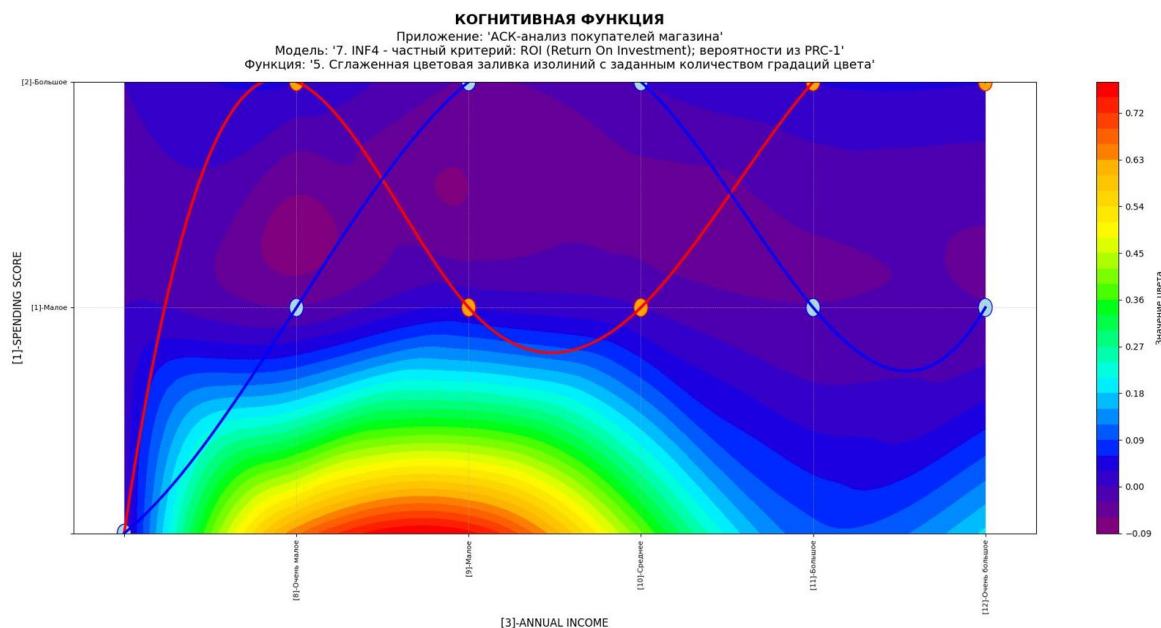


Рисунок 34. Примеры когнитивных функций в СК-модели INF4

Как уже отмечалось содержательное объяснение когнитивных функций на теоретическом уровне познания – это дело специалистов в той предметной области, к которой относится предмет моделирования [24].

3.8.10. Значимость описательных шкал и их градаций

В АСК-анализе все факторы рассматриваются с одной единственной точки зрения: сколько информации содержится в их значениях о переходе объекта моделирования и управления, на который они действуют, в определенное будущее состояние, описываемое классом (градация классификационной шкалы), и при этом сила и направление влияния всех значений факторов на объект измеряется в одних общих для всех факторов единицах измерения: единицах количества информации [6].

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это вариабельность информативностей (режим 3.7.5 системы «Эйдос»).

Значимость всей описательной шкалы является средним от степени значимости ее градаций (режим 3.7.4 системы «Эйдос»).

Если рассортировать все градации факторов (признаки) в порядке убывания селективной силы и получить сумму селективной силы системы значений факторов нарастающим итогом, то получим Парето-кривую.

На рисунке 37 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4:

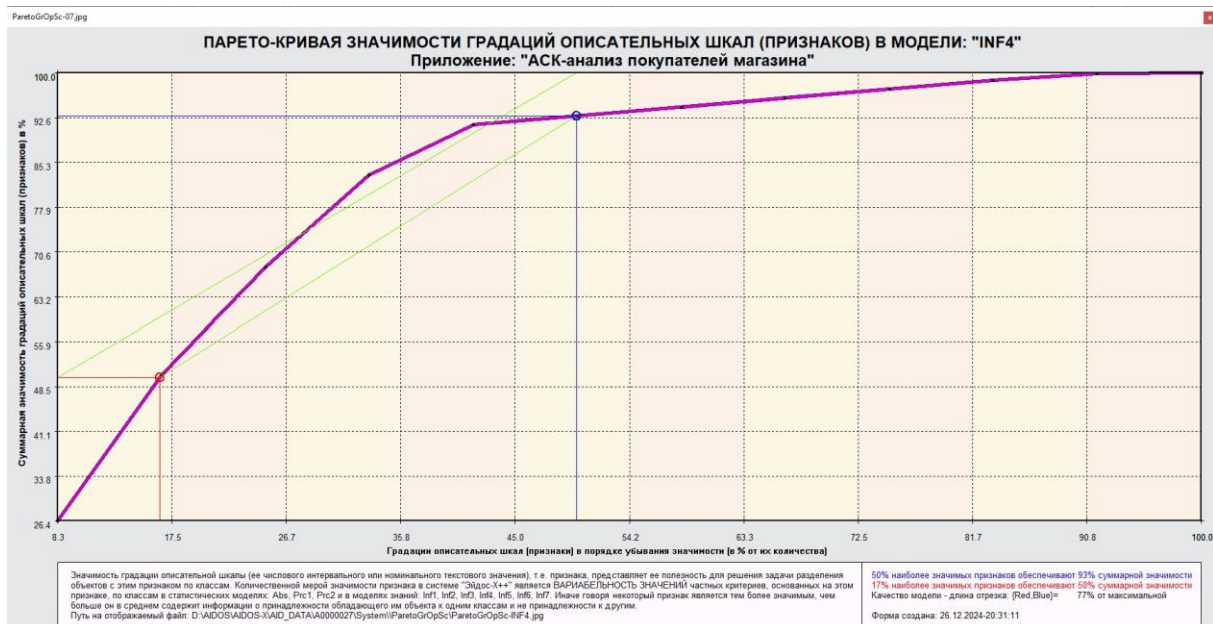


Рисунок 35. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4

Из рисунка 37 видно, что примерно двенадцатая часть наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 93% суммарного влияния.

В таблице 16 представлены исходные данные для построения кумулятивной кривой на рисунке 37. Из таблицы 16 видно, какую долю от суммарного влияния на переход объекта моделирования в будущее состояния, соответствующие классам, имеет каждое значение каждого фактора.

Таблица 16 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели INF4 (фрагмент)

NUM	NUM_PR C	KOD_OPS C	NAME_OPS C	N_GROPS C	KODGR_MI N	KODGR_MA X	ZNACH_O S
1	33,33333	2	AGE	5	3	7	0,53074
2	66,66667	1	GENRE	2	1	2	0,037915
3	100	3	ANNUAL INCOME	5	8	12	0,034229

На экранной форме рисунка 38 приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях:

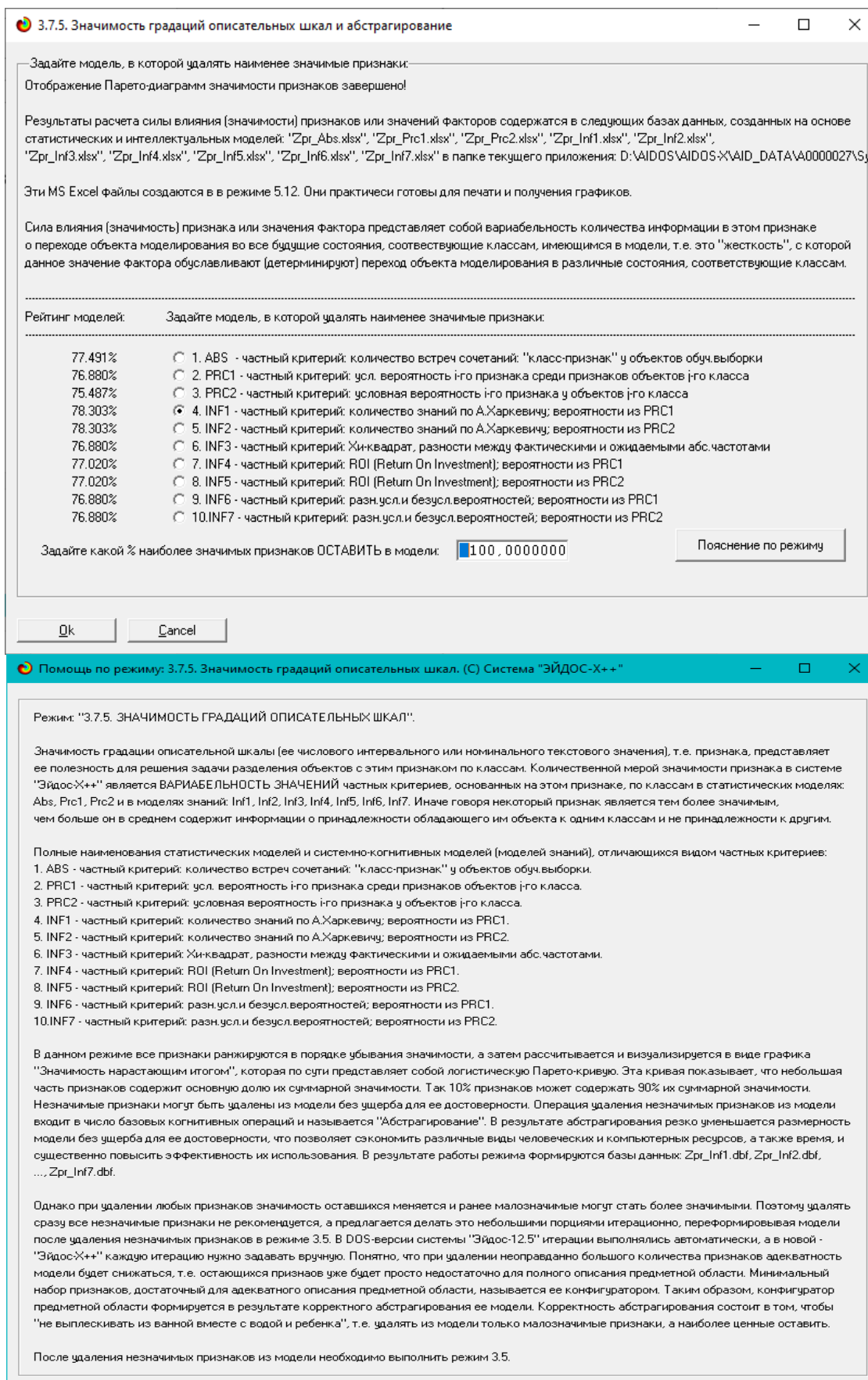


Рисунок 36. Имена Excel-файлов с информацией о силе влияния значений факторов в разных моделях

На экранной форме рисунка 39 приведены имена Excel-файлов с информацией о силе влияния факторов в разных моделях.

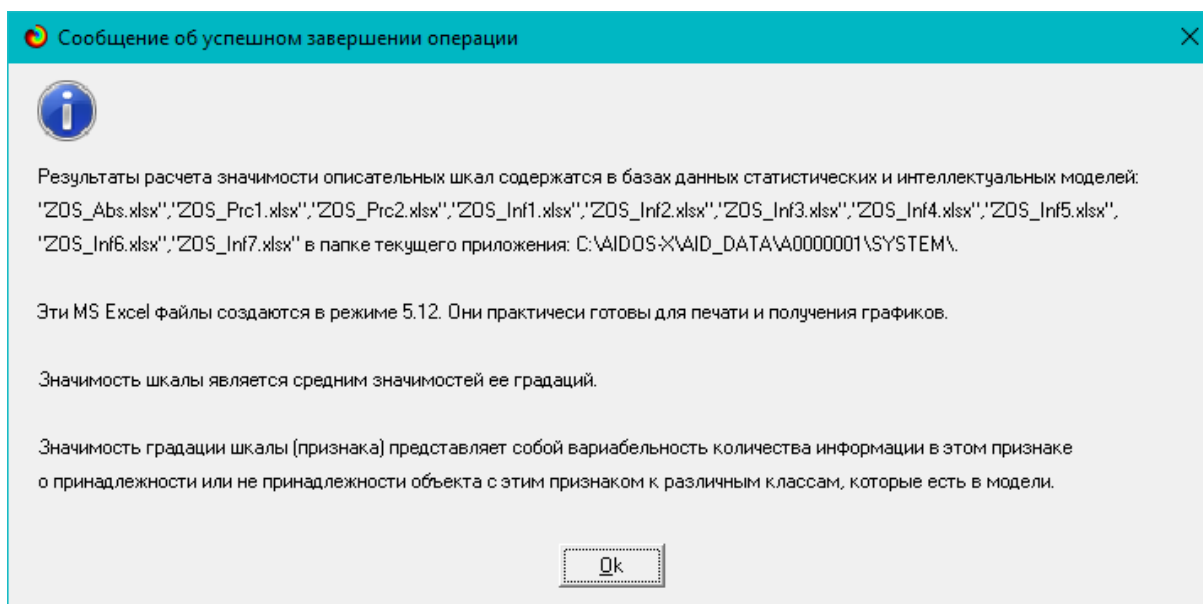


Рисунок 37. имена Excel-файлов с информацией о силе влияния факторов в разных моделях

3.8.11. Степень детерминированности классов и классификационных шкал

Степень детерминированности (обусловленности) класса в системе «Эйдос» количественно оценивается *степенью вариабельности значений факторов* (градаций описательных шкал) в колонке матрицы модели, соответствующей данному классу (режим 3.7.3 системы «Эйдос»).

Чем выше степень детерминированности класса, тем более достоверно он прогнозируется по значениям факторов.

Степень детерминированности (обусловленности) всей классификационной шкалы является средним от степени детерминированности ее градаций, т.е. классов (режим 3.7.2 системы «Эйдос»).

На рисунках 40 приведены экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:

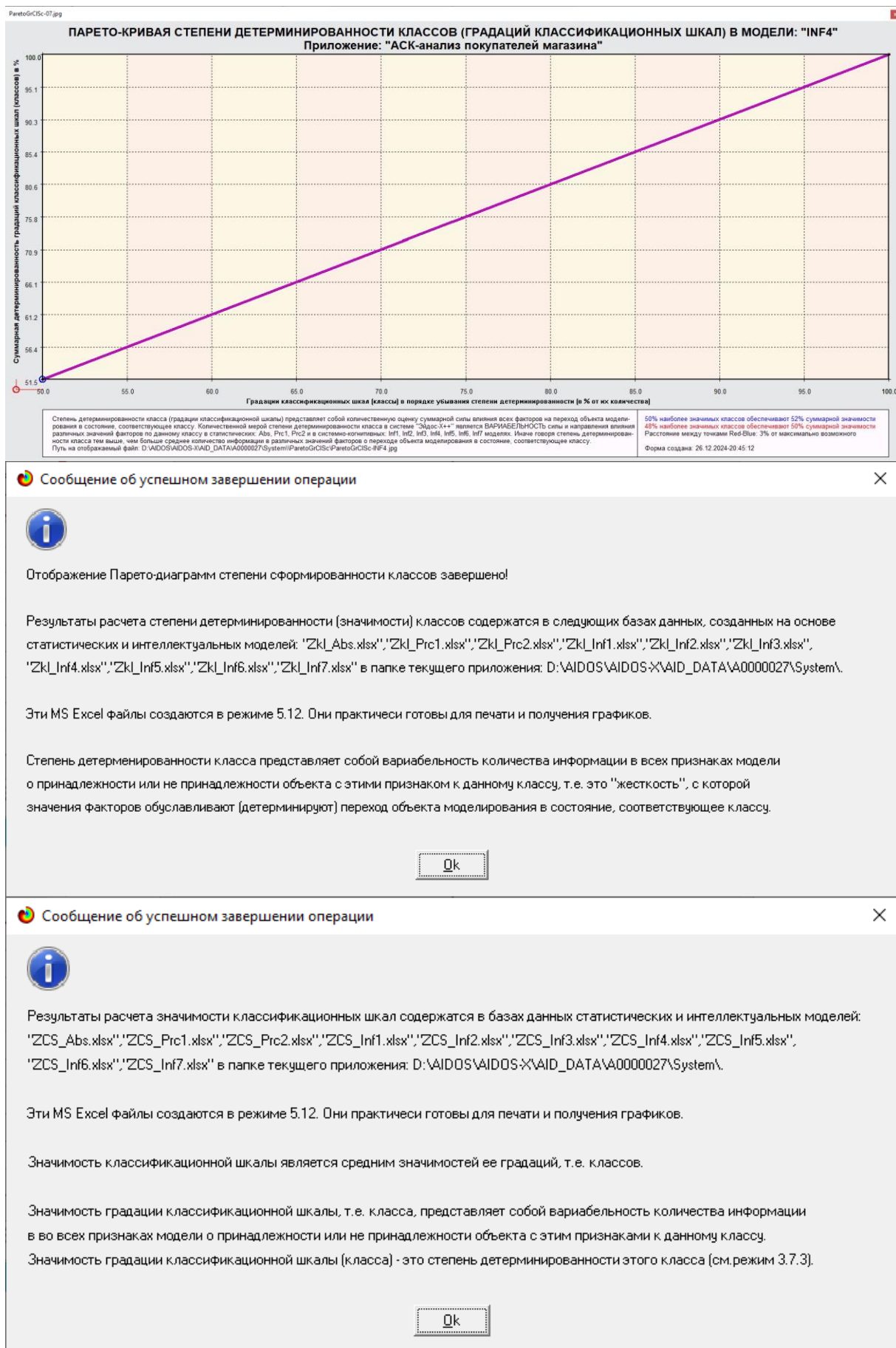


Рисунок 38. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

В таблице 17 представлены исходные данные для построения кумулятивной кривой на рисунке 40.

Из таблицы 18 видно, какую долю от суммарной степени детерминированности всех классов имеет каждый класс. Степень обусловленности значениями факторов разных будущих состояний объекта моделирования, соответствующие классам, довольно существенно отличается друг от друга.

Например, всего лишь 16% наиболее жестко детерминированных классов суммарно имеют примерно 50% степень детерминированности, а 50% наиболее детерминированных классов обеспечивают около 90% суммарной детерминированности всех классов.

Таблица 17 – Степень детерминированности классов в СК-модели INF4 (фрагмент)

NUM	NUM_PRC	KOD_CLS	NAME_CLS	KOD_CLSC	ZNACH_CLS
1	50,0000000	1	SPENDING SCORE- Малое	1	103,0000000
2	100,0000000	2	SPENDING SCORE- Большое	1	97,0000000

В таблице 18 приведена информация о степени детерминированности классов значениями факторов в системно-когнитивной модели INF4. Степень детерминированности классификационных шкал является средним от степени детерминированности их градаций.

Таблица 18 – Степень детерминированности классификационных шкал в системно-когнитивной модели INF4 (фрагмент)

NUM	NUM_PRC	KOD_CLSC	NAME_CLSC	N_GRCLSC	KODGR_MIN
1	100,0000000	1	SPENDING SCORE	2	1

4. DISCUSSION (ОБСУЖДЕНИЕ)

Полученные результаты можно оценить, как успешно решающие сформулированную в работе проблему и обеспечивающие достижение поставленной в работе цели. Эти результаты получены путем применения Автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария – интеллектуальной системы «Эйдос».

Анализ полученных результатов, проведенный в данной работе, полностью согласуется с результатами работы, на исходных данных которой они основаны. С другой стороны, применение АСК-анализа и системы «Эйдос» весьма существенно расширяет возможности решения

задач прогнозирования, принятия решений и исследования моделируемой предметной области, по сравнению с методами, применяемыми в работе. Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований.

Достижением данной работы является:

1. Возможность корректного построения сопоставимых системнокогнитивных моделей предметной области на основе исходных данных, содержащих как лингвистические переменные, так и числовые переменные в различных единицах измерения.

2. Возможность применения системно-когнитивных моделей для решения задач прогнозирования, принятия решений и исследования моделируемой предметной области.

В качестве перспективы продолжения исследований можно было бы рекомендовать существенно увеличить объем исходных данных, количество исследуемых факторов, а также количество классификационных шкал и их градаций (классов) для описания будущих состояний объекта моделирования.

Перспективность и ценность результатов подобных исследований и разработок для теории и практики не вызывает особых сомнений, что подтверждается работами в этой области.

Для выполнения анализа надо скачать систему с сайта разработчика по ссылке на странице: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос приложение. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Эйдос» решение поставленных задач является вполне эффективным, что позволяет обоснованно утверждать, что цель работы достигнута, поставленная проблема решена.

В результате проделанной работы, с помощью системы «Эйдос» были созданы статистические и системно-когнитивные модели, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по различным характеристикам покупателя в магазине, изучено влияние параметров людей из выборки на эти классы, и, на основе этого, решены задачи идентификации, классификации и исследования моделируемой предметной области путем исследования ее модели

REFERENCES (ЛИТЕРАТУРА)

1. Луценко, Е. В. Исследование влияния подсистем различных уровней иерархии на эмерджентные свойства системы в целом с применением АСК-анализа и интеллектуальной системы "Эйдос" (микроструктура системы как фактор управления ее макросвойствами) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2012. – № 75. – С. 321-363. – EDN OOSCAV.
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHC.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZJXZZ.
4. Работы проф.Е.В.Луценко по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики http://lc.kubagro.ru/aidos/Work_on_emergence.htm
5. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.
6. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
7. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.
8. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
10. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-X++" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2014. – № 101. – С. 1367-1409. – EDN SZVWRV.
11. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. – 288 с. – DOI 10.13140/RG.2.2.22981.37608. – EDN ZQLITW.
12. Орлов, А. И. Анализ данных, информации и знаний в системной нечеткой интервальной математике / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2022. – 405 с. – ISBN 978-5-907550-62-9. – DOI 10.13140/RG.2.2.15688.44802. – EDN OQULUW.
13. Луценко, Е. В. Автоматизация функционально-стоимостного анализа и метода "Директ-костинг" на основе АСК-анализа и системы "Эйдос" (автоматизация управления натуральной и финансовой эффективностью затрат без содержательных технологических и финансово-экономических расчетов на основе информационных и

когнитивных технологий и теории управления)1 / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.

14. Луценко, Е. В. Системное обобщение принципа Эшби и повышение уровня системности модели объекта познания как необходимое условие адекватности процесса его познания / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2020. – № 163. – С. 100-134. – DOI 10.21515/1990-4665-163-009. – EDN SWKGWY.

15. Луценко, Е. В. Эффективность объекта управления как его эмерджентное свойство и повышение уровня системности как цель управления / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2021. – № 165. – С. 77-98. – DOI 10.13140/RG.2.2.11887.25761. – EDN UMTAMT.

16. Луценко, Е. В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе "Эйдос") / Е. В. Луценко, В. Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 71. – С. 27-74. – EDN OIGYBB.

17. Луценко, Е. В. Проблемы и перспективы теории и методологии научного познания и автоматизированный системно-когнитивный анализ как автоматизированный метод научного познания, обеспечивающий содержательное феноменологическое моделирование / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 127. – С. 1-60. – DOI 10.21515/1990-4665-127-001. – EDN YLZTMX.

18. Луценко, Е. В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2003. – № 1. – С. 76-88. – EDN JWXLKT.

19. Работы проф.Е.В.Луценко & С^о по выявлению, представлению и использованию знаний, логике и методологии научного познания: http://lc.kubagro.ru/aidos/Works_on_identification_presentation_and_use_of_knowledge.htm

20. Пойа Дьердь. Математика и правдоподобные рассуждения. // под редакцией С.А.Яновской. Пер. с английского И.А.Вайнштейна., М., Наука, 1975 — 464 с., <http://ilib.mccme.ru/djvu/polya/rassuzhdenija.htm>

21. Луценко, Е. В. Системно-когнитивный анализ как развитие концепции смысла Шенка - Абельсона / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2004. – № 5. – С. 14-35. – EDN JWXMKX.

22. Работы проф.Е.В.Луценко & С^о по когнитивным функциям: http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm

23. Луценко, Е. В. Системы представления и приобретения знаний / Е. В. Луценко, В. И. Лойко, В. Н. Лаптев. – Краснодар : Экоинвест, 2018. – 513 с. – ISBN 978-5-94215-415-8. – EDN UZZBLC.

24. Луценко, Е. В. Проблемы и перспективы теории и методологии научного познания и автоматизированный системно-когнитивный анализ как автоматизированный метод научного познания, обеспечивающий содержательное феноменологическое моделирование / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 127. – С. 1-60. – DOI 10.21515/1990-4665-127-001. – EDN YLZTMX.