

**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 популярности книг, на основе 50 мировых
бестселлеров»

Выполнил студент группы: ИТ2241 Чернов Владислав Викторович

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 Чернова Владислава Викторовича
курса 2 очной формы обучения группы ИТ2241
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ популярности книг, на основе 50 мировых бестселлеров»

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	
2	Уровень и корректность использования в работе различных методов исследований	
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	
4	Ясность, четкость, последовательность и обоснованность изложения	
5	Применение современных технологий обработки информации	
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	
8	Ответы на вопросы при защите	

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____

Рецензент _____ (Е. В. Луценко)

«19» февраля 2024 г.

РЕФЕРАТ

Курсовая работа содержит: 75 страниц, 45 рисунков, 5 таблиц, 10 литературных источников.

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

Целью работы является проведение автоматизированного системно-когнитивного анализа популярности книг, на основе 50 мировых бестселлеров.

Для достижения поставленной цели необходимо провести анализ методов формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами принятия решений и исследования моделируемой предметной области путем исследования модели.

**АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-
КОГНИТИВНЫЙ АНАЛИЗ ПОПУЛЯРНОСТИ
КНИГ, НА ОСНОВЕ 50 МИРОВЫХ
БЕСТСЕЛЛЕРОВ**

Чернов Владислав Викторович
студент факультета ИТ, группы ИТ2241
Vladi4ka200vlad@gmail.com

*Кубанский Государственный Аграрный университет
имени И.Т.Трубилина, Краснодар, Россия*

Суть данной работы заключается в проведении анализа популярности книг, на основе 50 мировых бестселлеров. На основе знания этих зависимостей решаются задачи прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования ее системно-когнитивной модели. Это достигается путем вычисления количества информации, содержащегося в градациях номинальных шкал и получении того или иного шанса на победы. Приводится краткое описание АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос».

Ключевые слова: АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «ЭЙДОС».

**AUTOMATED SYSTEM-COGNITIVE ANALYSIS
OF BOOK POPULARITY, BASED ON 50 WORLD
BEST SELLERS**

Chernov Vladislav Victorovich
student of faculty of IT, group IT2241
Vladi4ka200vlad@gmail.com

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

The essence of this work is to analyze the popularity of books based on 50 world bestsellers. Based on the knowledge of these dependencies, the problems of prediction, decision-making and research of the modeled subject area are solved by investigating its system-cognitive model. This is achieved by calculating the amount of information contained in the gradations of nominal scales and obtaining this or that chance of wins. A brief description of ASK-analysis and its software toolkit - the intellectual system "Aidos" is given.

Keywords: AUTOMATED SYSTEM-COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS".

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1 ПРЕДМЕТНАЯ ОБЛАСТЬ.....	7
1.1 ОПИСАНИЕ ИСПОЛЬЗУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ.....	7
1.2 ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ.....	7
1.3 ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	7
1.4 ЦЕЛЬ И ЗАДАЧИ РАБОТЫ.....	8
2 МЕТОДЫ	11
2.1 ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ.....	11
2.2 АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ)	13
2.3 СИСТЕМА «ЭЙДОС» - ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	15
3 РЕЗУЛЬТАТЫ	21
3.1 Задача – 1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ.....	21
3.2 Задача – 2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	22
3.3 Задача – 3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	27
3.4 Задача – 4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ.....	38
3.5 Задача – 5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	41
3.6 Задача – 6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	43
3.6.1 Интегральный критерий «Сумма знаний».....	45
3.6.2 Интегральный критерий «Семантический резонанс знаний»	46
3.6.3 Важные математические свойства интегральных критериев	47
3.7 Задача – 7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ.....	49
3.7.1 Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	49
3.7.2 Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	49
3.8 Задача – 8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	50
3.8.1 Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы).....	50
3.8.2 Кластерно-конструктивный анализ классов.....	51
3.8.3 Кластерно-конструктивный анализ значений описательных шкал.....	54
3.8.4 Модель знаний системы «Эйдос» и нелокальные нейроны	56
3.8.5 Нелокальная нейронная сеть.....	59
3.8.6 3D-интегральные когнитивные карты	60
3.8.7 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	62
3.8.8 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	64
3.8.9 Когнитивные функции	66
3.8.10 Значимость описательных шкал и градаций.....	68
3.8.11 Степень детерминированности классов и классификационных шкал	69
4 ОБСУЖДЕНИЕ.....	71
ЗАКЛЮЧЕНИЕ	72
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	73

ВВЕДЕНИЕ

В настоящей работе представлен автоматизированный системно-когнитивный анализ популярности книг с использованием системы «Эйдос». Цель исследования заключается в проведении анализа популярности книг, на основе 50 мировых бестселлеров с целью определения факторов, влияющих на успешность персонажа. Для достижения поставленной цели предполагается провести анализ методов формирования обобщенных образов классов, а также разработать решения для задач идентификации конкретных объектов с учетом классов принятия решений.

В рамках исследования предполагается провести системно-когнитивный анализ, включающий в себя изучение методов формирования обобщенных образов классов в контексте продажи книг. Также планируется провести исследование моделируемой предметной области с использованием системы «Эйдос», фокусируясь на анализе решений, принимаемых в различных ситуациях в игре.

В ходе исследования будет уделено внимание анализу факторов, влияющих на успешность продаж. Предполагается, что полученные результаты не только предоставят более глубокое понимание популярности определенных книг, но и могут иметь практическое применение при написании будущих произведений.

1 ПРЕДМЕТНАЯ ОБЛАСТЬ

1.1 Описание используемой предметной области

Книга – это не просто набор бумажных страниц с текстом, переплетённый в твёрдую или мягкую обложку. Это целый мир, который автор создаёт для чтения, погружения и обдумывания. Книги могут быть самыми разными: учебными, художественными, научными и даже справочными. Они передают знания, идеи и развлечения людям по всему миру. Книга – это способ путешествовать во времени и пространстве, не покидая своего комнатного уголка.

Бестселлер - книга, пользующаяся большим спросом, изданная большим тиражом, а также кинофильм, телеспектакль и т. п., ставшие наиболее популярными.

1.2 Объект и предмет исследования

Объект исследования – выявление зависимостей популярности определенных книг от различных факторов.

Предмет исследования – выявление зависимостей популярности определенных книг, используя данных о мировых бестселлерах.

1.3 Проблема, решаемая в работе и ее актуальность

Актуальность исследования заключается в возможности авторов произведений выявить для себя факторы, влияющие на продажи их произведений.

Настоящая работа нацелена на автоматизированный системно-когнитивный анализ популярности книг, на основе 50 мировых бестселлеров. Поставленная цель предполагает проведение методологического обоснования исследования, охватывающего анализ методов формирования обобщенных образов классов в контексте данной предметной области.

Применение системно-когнитивного подхода позволит осуществить глубокий анализ взаимосвязей между различными параметрами и факторами, влияющими на продажи книг.

Для успешной реализации поставленной цели планируется применение методов математического моделирования и статистического анализа данных, что позволит получить объективные результаты и детальное понимание взаимосвязей в предметной области.

1.4 Цель и задачи работы

Для решения задач, представленных в работе, будет использоваться: Microsoft Excel и система искусственного интеллекта «Эйдос».

Целью данной курсовой работы является, составление АСК-анализа о зависимости заболеваний детей на основе их пола и района проживания.

Поскольку для решения поставленной проблемы используется автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий, в качестве которого в настоящее время выступает интеллектуальная система «Эйдос», то достижение поставленной цели обеспечивается решением следующих задач и подзадач, которые являются этапами достижения цели:

1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций;
2. Формализация предметной области;
3. Синтез статистических и системно-когнитивных моделей. Многопараметрическая типизация и частные критерии знаний;
4. Верификация моделей;
5. Выбор наиболее достоверной модели;
6. Системная идентификации и прогнозирование;
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);

8. Исследование объекта моделирования путем исследования его модели (Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы); Кластерно-конструктивный анализ классов и значений описательных шкал; Модель знаний системы «Эйдос» и нелокальные нейроны; Нелокальная нейронная сеть; 3D-интегральные когнитивные карты; 2D-интегральные когнитивные карты содержательного сравнения классов и факторов (опосредованные нечеткие правдоподобные рассуждения); Когнитивные функции; Значимость описательных шкал и их градаций; Степень детерминированности классов и классификационных шкал).

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

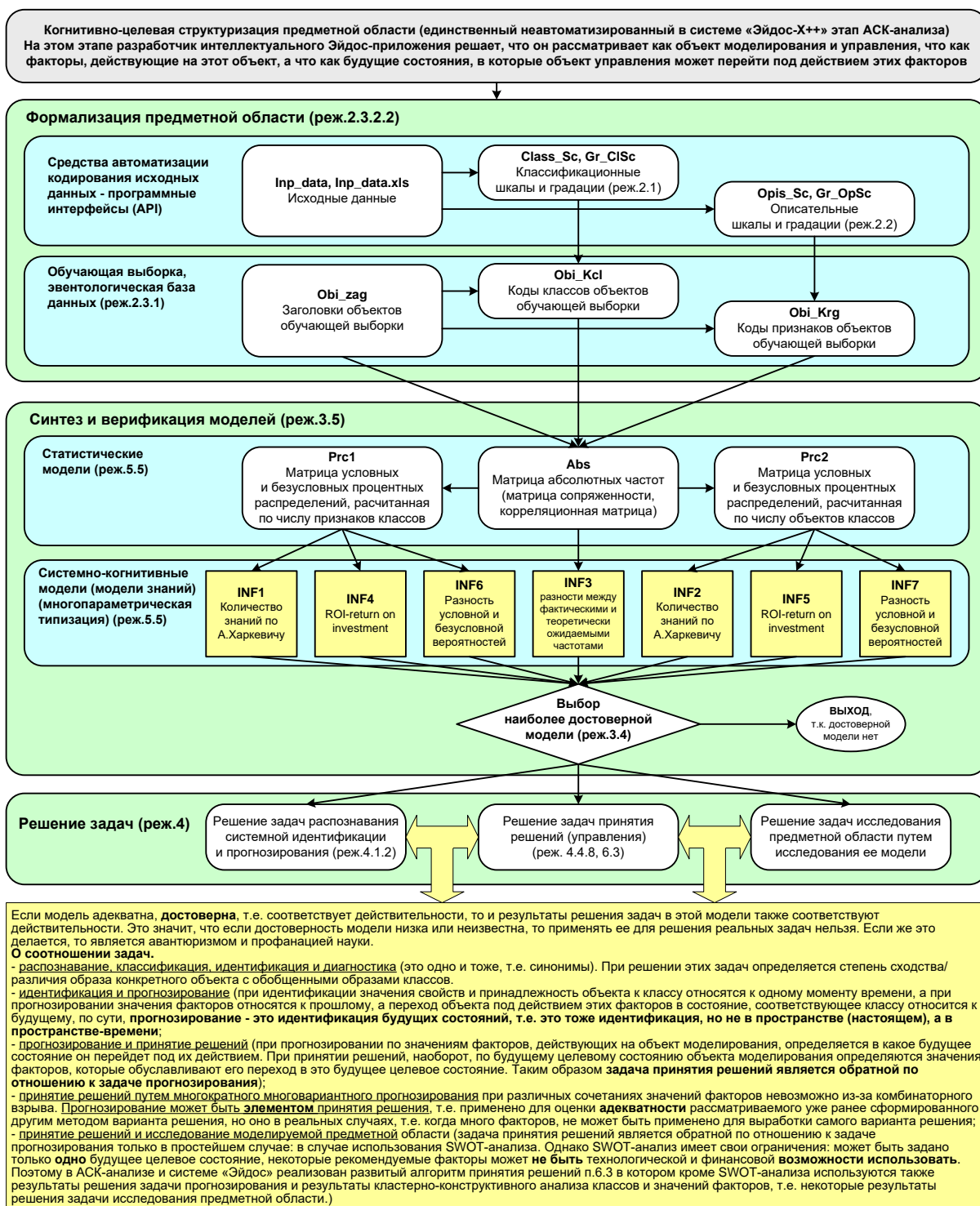


Рисунок 1 – Последовательность преобразования исходных данных в информацию и знаний в системе «Эйдос»

2 МЕТОДЫ

2.1 Обоснование требований к методу решения проблемы

В качестве метода исследования, решения проблемы и достижения цели было решено использовать новый метод искусственного интеллекта автоматизированный системно-когнитивный анализ или АСК-анализ. Основной причиной выбора АСК-анализа является то, что он включает теорию и метод количественного выявления в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал и единицах измерения.

Очень важным является также то, что АСК-анализ имеет свой развитый доступный программный инструментарий, в качестве которого в настоящее время выступает Универсальная когнитивная аналитическая система Aidos-X. Система «Эйдос» выгодно отличается от других интеллектуальных систем следующими параметрами:

- разработана в универсальной постановке, не зависящей от предметной области. Поэтому она является универсальной и может быть применена во многих предметных областях, в которых не требуется автоматического, т. е. без непосредственного участия человека в реальном времени решения задач идентификации, прогнозирования, приятия решений и исследования предметной области;

- находится в полном открытом бесплатном доступе причем с актуальными исходными текстами;

- 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»;

- 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);

— содержит большое количество локальных (поставляемых с инсталляцией) и более 300 учебных и научных интеллектуальных облачных Эйдос-приложений;

— поддерживает онлайн среду накопления знаний и широко используется во всем мире;

— 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 решения 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);

— 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 повторности 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.

2.2 Автоматизированный системно-когнитивный анализ (АСК-анализ)

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

Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф. Е.В.Луценко. На тот момент он вообще не встречался в Internet. Сегодня по соответствующему запросу в Яндексе находится 9 миллионов сайтов с этим сочетанием слов.

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);

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

Около половины из 706 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 45 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 34 патентов РФ на системы искусственного интеллекта, 346 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.

Три монографии включены в фонды библиотеки конгресса США.

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, психологическими 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». 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 РФ). 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 текстовых данных;
- автоматизированный системно-когнитивный анализ текстовых данных;
- спектральный и контурный автоматизированный системно-когнитивный анализ изображений;
- 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 и страничка в ResearchGate, 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.

2.3 Система «Эйдос» - инструментарий АСК-анализа

Существует много систем искусственного интеллекта. Универсальная когнитивная аналитическая система «Эйдос» отличается от них следующими параметрами:

- 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>). Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени при решении задач идентификации, прогнозирования, принятия решений и исследования предметной области (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, и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора системы «Эйдос» проф. Е.В. Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 31 свидетельство 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 учебных и научных Эйдос-приложений (в настоящее время их 31 и более 399, соответственно: http://aidos.byethost5.com/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

— поддерживает онлайн среду накопления знаний и обмена ими, широко используется во всем мире (<http://aidos.byethost5.com/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 решения 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.

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 по ссылке: <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>).

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 xb2net. 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 xb2net 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 Аляска+Экспресс, обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge) с использованием ADS (Advantage Database Server), а также на языке C# (Visual Studio | C#).

На рисунке 2 приведена титульная видеogramма DOS-версии системы «Эйдос», а на рисунке 3 – текущей версии системы «Эйдос».



Рисунок 2 – Титульная видеोगрамма DOS-версии «Эйдос»

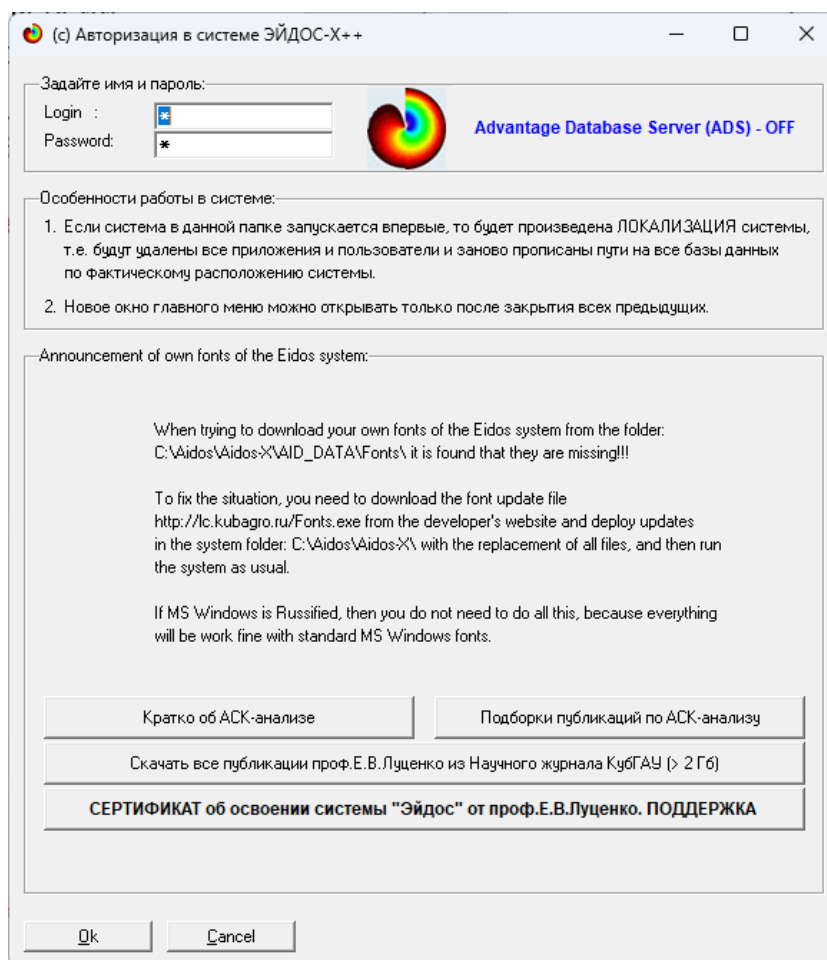


Рисунок 3 – Титульная видеोगрамма текущей версии системы «Эйдос»

3 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.

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 в самых различных направлениях науки и решения практических задач в

самых различных предметных областях, практически почти везде, где человек применяет естественный интеллект.

В качестве источника данных с электронного ресурса kaggle.com был взят набор данных «Best Selling Books», <https://www.kaggle.com/datasets/agarwalyashhh/best-selling-books/data>. Набор данных содержит 10 столбцов, рассмотрим их:

- Books – название книги;
- Sales_in_millions – количество продаж в миллионах;
- Authors – автор произведения;
- Language – язык, на котором написана книга;
- First_Published – год написания произведения;

Для загрузки модели в систему «Эйдос» необходимо конвертировать CSV-файл в файл формата XLSX. Для конвертации был использован встроенный механизм импорта из CSV в Microsoft Excel.

Классифицирующим столбцом было решено выбрать Win Rate, итоговую таблицу можно увидеть на рисунке 4.

Books	Sales in millions	Authors	Language	First Published
A Tale of Two Cities	200	Charles Dickens	English	1859
The Little Prince (Le Petit Prince)	200	Antoine de Saint-Exupéry	French	1943
Harry Potter and the Philosopher's Stone	120	J. K. Rowling	English	1997
And Then There Were None	100	Agatha Christie	English	1939
Dream of the Red Chamber (紅樓夢)	100	Cao Xueqin	Chinese	1791
The Hobbit	100	J. R. R. Tolkien	English	1937
She: A History of Adventure	83	H. Rider Haggard	English	1887
The Da Vinci Code	80	Dan Brown	English	2003
Harry Potter and the Chamber of Secrets	77	J. K. Rowling	English	1998
Harry Potter and the Prisoner of Azkaban	65	J. K. Rowling	English	1999
Harry Potter and the Goblet of Fire	65	J. K. Rowling	English	2000
Harry Potter and the Order of the Phoenix	65	J. K. Rowling	English	2003
Harry Potter and the Half-Blood Prince	65	J. K. Rowling	English	2005
Harry Potter and the Deathly Hallows	63	J. K. Rowling	English	2007
The Alchemist (O Alquimista)	63	Paulo Coelho	Portuguese	1988
The Catcher in the Rye	63	J. D. Salinger	English	1951
The Bridges of Madison County	60	Robert James Waller	English	1992
One Hundred Years of Solitude (Cien años de soledad)	50	Gabriel García Márquez	Spanish	1967
Lolita	50	Vladimir Nabokov	English	1955
Heidi	50	Johanna Spyri	German	1880
The Common Sense Book of Baby and Child Care	50	Benjamin Spock	English	1946
Anne of Green Gables	50	Lucy Maud Montgomery	English	1908

Рисунок 4 – Фрагмент обучающей выборки

Для импорта обучающей выборки в систему AIDOS-X необходимо скопировать ее в папку Inp_data и переименовать в Inp_data.xlsx, после этого можно запустить саму программу и универсальный программный интерфейс

импорта данных в систему (режим 2.3.2.2), результат заполнения которого представлен на рисунке 5.

Рисунок 5 – Интерфейс импорта данных

Следует выделить следующие настройки:

- Тип файла – xlsx;
- Классификационная шкала – 2;
- Описательные шкалы – 3-10.

После импорта данных система просит задать размерности модели системы, оставляем предложенную размерность (рисунок 6).

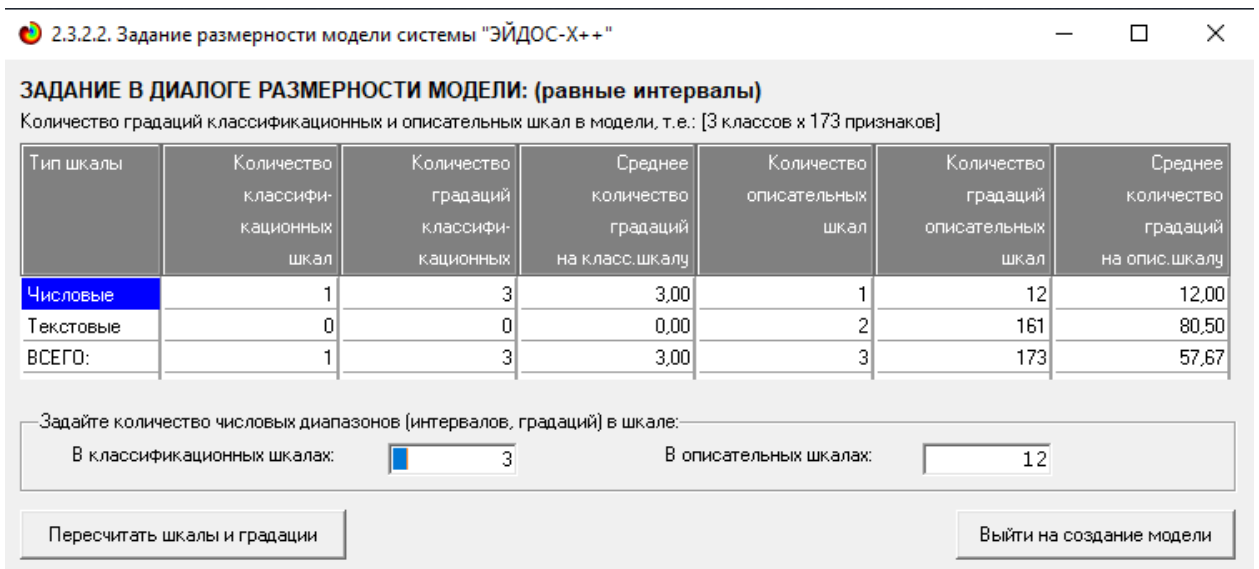


Рисунок 6 – Задание размерностей системы

Процесс импорта данных из внешнего файла в систему представлен на рисунке 7.

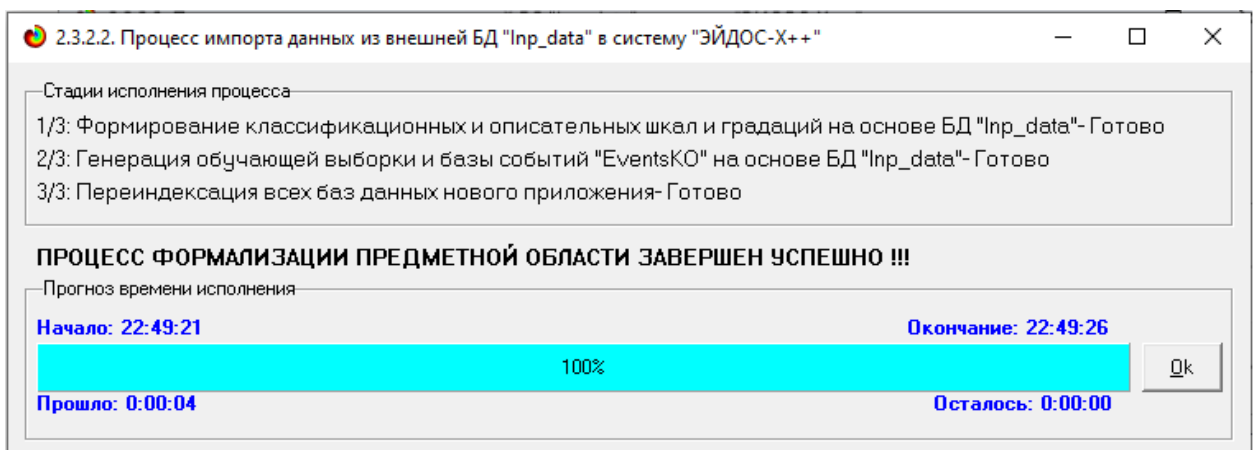


Рисунок 7 – Импорт данных

После загрузки данных, система автоматически нашла классификационные шкалы (рисунок 8) и описательные шкалы (рисунок 9).

2.1. Классификационные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы
1	SALES_IN_MILLIONS	1	1/3 {10.0000000, 73.3333333}
		2	2/3 {73.3333333, 136.6666667}
		3	3/3 {136.6666667, 200.0000000}

Помощь | Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Удаление и перекодирование | Графики будущих сценариев

Рисунок 8 – Классификационные шкалы

2.2. Описательные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	AUTHORS	1	Agatha Christie
2	LANGUAGE	2	Albert Camus
3	FIRST_PUBLISHED	3	Alex Comfort
		4	Alexander Alexandrovich Fadeyev
		5	Alice Sebold
		6	Andrew Morton
		7	Anna Sewell
		8	Anne Frank
		9	Anthony Doerr
		10	Antoine de Saint-Exupery
		11	Ariene Eisenberg and Heidi Mukoif
		12	Astrid Lindgren
		13	Banana Yoshimoto
		14	Beatrix Potter
		15	Benjamin Sprock
		16	Cao Xueqin
		17	Carl Sagan
		18	Carlo Colodi
		19	Carlos Ruiz Zafon
		20	Charles Berlitz
		21	Charles Dickens
		22	Chinua Achebe
		23	Colleen McCullough
		24	Dale Carnegie
		25	Dan Brown
		26	Dante Alighieri
		27	Daphne du Maurier
		--	--

Помощь | Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Перекодировать | Очистить | Графики прошлых сценариев

Рисунок 9 – Описательные шкалы

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

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) моделей осуществляется в режиме 3.5 системы «Эйдос». Математические модели, на основе которых рассчитываются статистические и СК-модели, подробно описаны в ряде монографий и статей автора. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко. Отметим лишь, 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) (рисунок 4).

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

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных (см. Help режима 2.3.2.2) рассчитывается матрица абсолютных частот (таблица 1).

Таблица 1 – Матрица абсолютных частот (статистическая модель ABS)

		Классы					Сумма
		<i>l</i>	...	<i>j</i>	...	<i>w</i>	
Значения факторов	<i>l</i>	N_{11}		N_{1j}		N_{1w}	
	...						
	<i>i</i>	N_{i1}		N_{ij}		N_{iw}	$N_{i\Sigma} = \sum_{j=1}^w N_{ij}$
	...						
	<i>M</i>	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 ее основе рассчитываются матрицы условных и безусловных процентных распределений (таблица 2).

Таблица 2 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

		Классы					Безусловная вероятность признака
		<i>l</i>	...	<i>j</i>	...	<i>w</i>	
Значения факторов	<i>l</i>	P_{11}		P_{1j}		P_{1w}	
	...						
	<i>i</i>	P_{i1}		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		P_{iw}	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	<i>M</i>	P_{M1}		P_{Mj}		P_{MW}	
Безусловная вероятность класса				$P_{\Sigma j}$			

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

1. В качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;
2. 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 объектов обучающейся выборки, относящихся к различным классам. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 1) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) является весьма обоснованным и логичным.

Этот переход полностью снимает проблему несбалансированности данных, т.к. в последующем анализе используется не матрица абсолютных частот, а матрицы условных и безусловных процентных распределений, и матрицы системно-когнитивных моделей (СК-модели, таблица 4), в частности матрица информативностей.

Данный подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (номинальных, порядковых и числовых) и в разных единицах измерения.

В системе «Эйдос» указанный подход используется всегда при решении любых задач.

Затем на основе таблицы 2 с использованием частных критериев, знаний, приведенных таблице 3, рассчитываются матрицы системно-когнитивных моделей (таблица 4).

В таблице 3 приведены формулы:

- для сравнения фактических и теоретических абсолютных частот;

— для сравнения условных и безусловных относительных частот (вероятностей).

Это сравнение в таблице 3 осуществляется двумя возможными способами: путем вычитания и путем деления.

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).

Таблица 3 – 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	$N_i = \sum_{j=1}^M N_{ij}; N_j = \sum_{i=1}^W N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $N_{ij} - \text{фактическая частота};$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота.}$	
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	---	
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	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij} N}{N_i N_j}$
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	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$
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.

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

Таблица 4 – Матрица системно-когнитивной модели

		Классы					Значимость фактора
		<i>l</i>	...	<i>j</i>	...	<i>W</i>	
Значения факторов	<i>l</i>	I_{l1}		I_{lj}		I_{lW}	$\sigma_{l\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^W (I_{lj} - \bar{I}_l)^2}$
	...						
	<i>i</i>	I_{i1}		I_{ij}		I_{iW}	$\sigma_{i\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
	...						
	<i>M</i>	I_{M1}		I_{Mj}		I_{MW}	$\sigma_{M\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^W (I_{Mj} - \bar{I}_M)^2}$
Степень редукции класса		$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

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.

На основе системно-когнитивных моделей, 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 4).

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 5).

Таблица 5 – 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	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Для синтеза и верификации моделей необходимо использовать режим 3.5, после его запуска задается текущая модель (рисунок 10).



Рисунок 10 – Выбор моделей для синтеза и верификации

Режим 3.5 содержит множество различных методов верификации моделей. Стадия процесса выполнена и прогноз времени окончания изображен на форме, приведенной на рисунке 11.

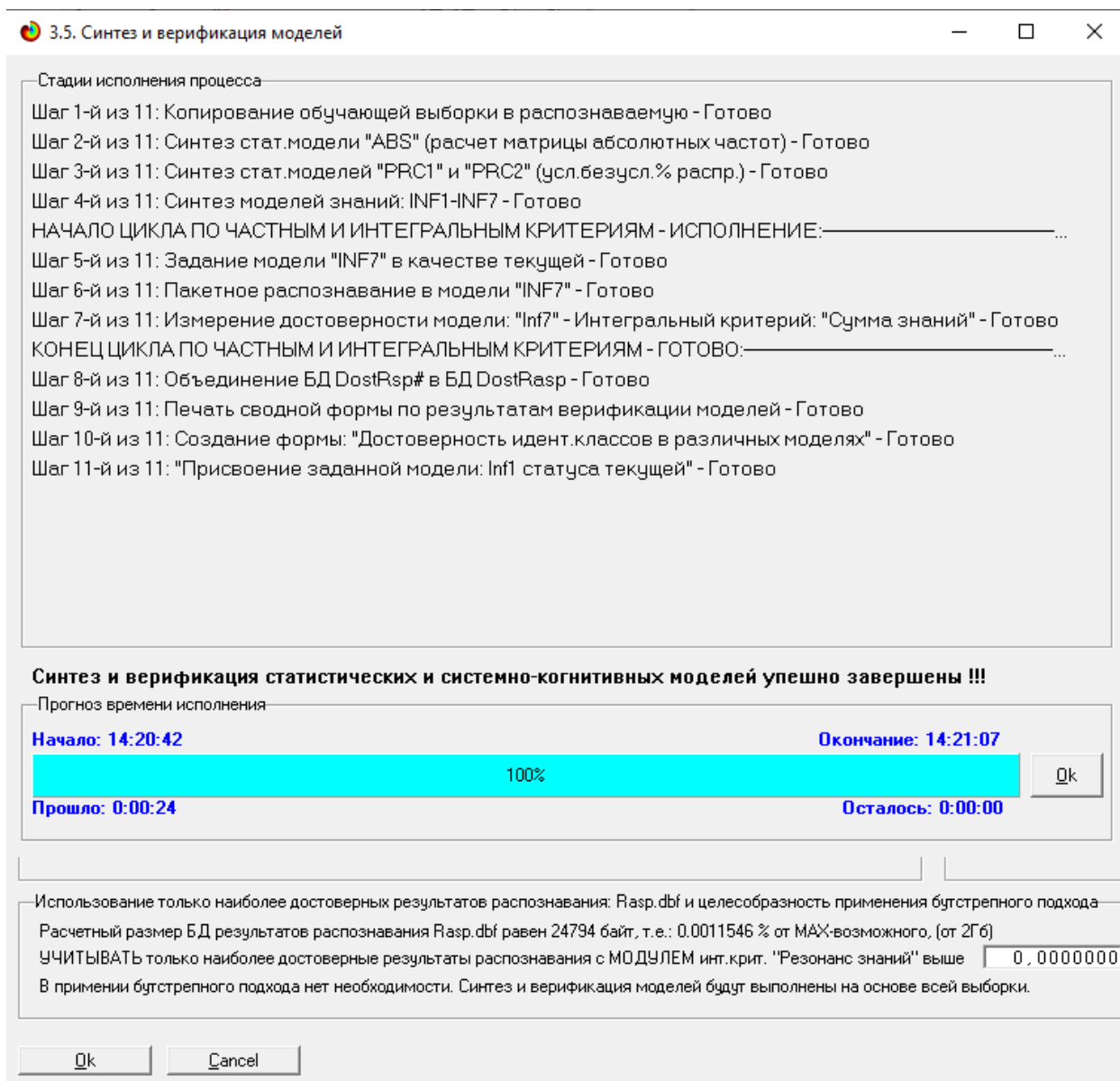


Рисунок 11 – Исполнение режима синтеза моделей

Синтез и верификации всех моделей заняла 24 секунды. Далее можно приступить к выбору наиболее достоверной модели.

Решение задачи идентификации может быть рассмотрено на нескольких моделях, в которых рассчитано количество информации по А. Харкевичу, которое было сгенерировано по принадлежности идентифицируемого объекта к каждому из классов, учитывая признаки объекта.

На рисунках 12 и 13 представлены фрагменты созданных статистических и системно-когнитивных моделей (СК-моделей).

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

Код признака	Наименование описательной шкалы и градации	1. SALES_IN_MILL... 1/3 {10.0, 73.3}	2. SALES_IN_MILL... 2/3 {73.3, 136.7}	3. SALES_IN_MILL... 3/3 {136.7, 200.0}	Сумма	Среднее	Средн. квадрат. откл.
1.0	AUTHORS-Agatha Christie		1.0		1.0	0.33	0.58
2.0	AUTHORS-Albert Camus	2.0			2.0	0.67	1.15
3.0	AUTHORS-Alex Comfort	1.0			1.0	0.33	0.58
4.0	AUTHORS-Alexander Alexandrovich Fadeyev	1.0			1.0	0.33	0.58
5.0	AUTHORS-Alice Sebold	1.0			1.0	0.33	0.58
6.0	AUTHORS-Andrew Morton	1.0			1.0	0.33	0.58
7.0	AUTHORS-Anna Sewell	1.0			1.0	0.33	0.58
8.0	AUTHORS-Anne Frank	1.0			1.0	0.33	0.58
9.0	AUTHORS-Anthony Doerr	1.0			1.0	0.33	0.58
10.0	AUTHORS-Antoine de Saint-Exupery			1.0	1.0	0.33	0.58
11.0	AUTHORS-Arlene Eisenberg and Heidi Murkoff	1.0			1.0	0.33	0.58
12.0	AUTHORS-Astrid Lindgren	1.0			1.0	0.33	0.58
13.0	AUTHORS-Banana Yoshimoto	1.0			1.0	0.33	0.58
14.0	AUTHORS-Beatrix Potter	1.0			1.0	0.33	0.58
15.0	AUTHORS-Benjamin Spock	1.0			1.0	0.33	0.58
16.0	AUTHORS-Cao Xueqin		1.0		1.0	0.33	0.58
17.0	AUTHORS-Carl Sagan	1.0			1.0	0.33	0.58
18.0	AUTHORS-Carlo Collodi	1.0			1.0	0.33	0.58
19.0	AUTHORS-Carlos Ruiz Zafon	1.0			1.0	0.33	0.58
20.0	AUTHORS-Charles Berlitz	1.0			1.0	0.33	0.58
21.0	AUTHORS-Charles Dickens			1.0	1.0	0.33	0.58
22.0	AUTHORS-Chinua Achebe	1.0			1.0	0.33	0.58
23.0	AUTHORS-Colleen McCullough	1.0			1.0	0.33	0.58
24.0	AUTHORS-Dale Carnegie	1.0			1.0	0.33	0.58

Рисунок 12 – Матрица абсолютных частот

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

Код признака	Наименование описательной шкалы и градации	1. SALES_IN_MILL... 1/3 {10.0, 73.3}	2. SALES_IN_MILL... 2/3 {73.3, 136.7}	3. SALES_IN_MILL... 3/3 {136.7, 200.0}	Сумма	Среднее	Средн. квадрат. откл.
1.0	AUTHORS-Agatha Christie	-0.945	0.957	-0.012			0.951
2.0	AUTHORS-Albert Camus	0.110	-0.085	-0.024			0.100
3.0	AUTHORS-Alex Comfort	0.055	-0.043	-0.012			0.050
4.0	AUTHORS-Alexander Alexandrovich Fadeyev	0.055	-0.043	-0.012			0.050
5.0	AUTHORS-Alice Sebold	0.055	-0.043	-0.012			0.050
6.0	AUTHORS-Andrew Morton	0.055	-0.043	-0.012			0.050
7.0	AUTHORS-Anna Sewell	0.055	-0.043	-0.012			0.050
8.0	AUTHORS-Anne Frank	0.055	-0.043	-0.012			0.050
9.0	AUTHORS-Anthony Doerr	0.055	-0.043	-0.012			0.050
10.0	AUTHORS-Antoine de Saint-Exupery	-0.945	-0.043	0.988			0.967
11.0	AUTHORS-Arlene Eisenberg and Heidi Murkoff	0.055	-0.043	-0.012			0.050
12.0	AUTHORS-Astrid Lindgren	0.055	-0.043	-0.012			0.050
13.0	AUTHORS-Banana Yoshimoto	0.055	-0.043	-0.012			0.050
14.0	AUTHORS-Beatrix Potter	0.055	-0.043	-0.012			0.050
15.0	AUTHORS-Benjamin Spock	0.055	-0.043	-0.012			0.050
16.0	AUTHORS-Cao Xueqin	-0.945	0.957	-0.012			0.951
17.0	AUTHORS-Carl Sagan	0.055	-0.043	-0.012			0.050
18.0	AUTHORS-Carlo Collodi	0.055	-0.043	-0.012			0.050
19.0	AUTHORS-Carlos Ruiz Zafon	0.055	-0.043	-0.012			0.050
20.0	AUTHORS-Charles Berlitz	0.055	-0.043	-0.012			0.050
21.0	AUTHORS-Charles Dickens	-0.945	-0.043	0.988			0.967
22.0	AUTHORS-Chinua Achebe	0.055	-0.043	-0.012			0.050
23.0	AUTHORS-Colleen McCullough	0.055	-0.043	-0.012			0.050
24.0	AUTHORS-Dale Carnegie	0.055	-0.043	-0.012			0.050

Рисунок 13 – Модель INF3

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 Ризбергена, а также по критериям L1- L2-мерам проф. Е. В. Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры.

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 F-мерой Ван Ризбергена наиболее достоверной является СК-модель INF1, INF2, INF4 и INF5 с интегральным критерием: «Сумма знаний»: $F=0.684$ (рисунок 14).

В соответствии с критерием достоверности L1-мерой проф. Е. В. Луценко наиболее достоверной является СК-модель ABS с интегральным критерием: «Сумма абсолютных частот по признакам»: $L1=0.954$ (рисунок 15). Эту модель и будем использовать для решения поставленных в работе задач.

С результатами верификации моделей, отличающихся частными критериями, можно ознакомиться в режиме 3.4 системы «Эйдос», они представлены в трех частях на рисунках 14-16.

3.4. Обобщ.форма по достов.моделям при разн.крит. Текущая модель: "ИИТ"

Наименование модели и частного критерия	Интегральный критерий	Всего полученных значений	Число истинных положительных решений (ПР)	Число истинных отрицательных решений (ПН)	Число ложных положительных решений (ЛР)	Число ложных отрицательных решений (ЛН)	Точность модели	Попытки модели	Г-меры Ван Розбергена	Среднее число решений (с.д.с.) истинный (СТР)
1. ABS - частный критерий: количество встреч сигнала "1" на...	Корреляция abs частот с обр.	164	162	23	305	2	0.347	0.988	0.513	122.394
1. ABS - частный критерий: количество встреч сигнала "1" на...	Средн abs частот по границе	164	164	22	306		0.349	1.000	0.517	119.393
2. PRC1 - частный критерий: усл. вероятность того признака сред.	Корреляция усл.отн частот с о.	164	162	23	305	2	0.347	0.988	0.513	122.394
2. PRC1 - частный критерий: усл. вероятность того признака сред.	Средн усл.отн частот по грани	164	164	22	306		0.349	1.000	0.517	117.948
3. PRC2 - частный критерий: условная вероятность того признака сред.	Корреляция усл.отн частот с о.	164	162	23	305	2	0.347	0.988	0.513	122.393
3. PRC2 - частный критерий: условная вероятность того признака сред.	Средн усл.отн частот по грани	164	164	22	306		0.349	1.000	0.517	117.948
4. ИИТ - частный критерий: количество знаков по А.Харкевичу, в.	Семантический резонанс: зна.	164	26	287	41	138	0.388	0.159	0.226	4.878
4. ИИТ - частный критерий: количество знаков по А.Харкевичу, в.	Среднe знаков	164	143	217	111	21	0.563	0.872	0.684	6.161
5. ИИТ2 - частный критерий: количество знаков по А.Харкевичу, в.	Семантический резонанс: зна.	164	26	287	41	138	0.388	0.159	0.226	4.878
5. ИИТ2 - частный критерий: количество знаков по А.Харкевичу, в.	Среднe знаков	164	143	217	111	21	0.563	0.872	0.684	6.161
6. ИИТ3 - частный критерий: Знаки надг. различия между фактн.	Семантический резонанс: зна.	164	115	259	49	49	0.625	0.701	0.661	67.023
6. ИИТ3 - частный критерий: Знаки надг. различия между фактн.	Среднe знаков	164	115	259	49	49	0.625	0.701	0.661	92.216
7. ИИТ4 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс: зна.	164	26	301	27	138	0.491	0.159	0.240	5.078
7. ИИТ4 - частный критерий: ROI (Return On Investment), вероятно.	Среднe знаков	164	143	217	111	21	0.563	0.872	0.684	5.286
8. ИИТ5 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс: зна.	164	26	301	27	138	0.491	0.159	0.240	5.078
8. ИИТ5 - частный критерий: ROI (Return On Investment), вероятно.	Среднe знаков	164	143	217	111	21	0.563	0.872	0.684	5.286
9. ИИТ6 - частный критерий: разн услн безудл вероятностей, вер.	Семантический резонанс: зна.	164	115	248	80	49	0.590	0.701	0.641	62.563
9. ИИТ6 - частный критерий: разн услн безудл вероятностей, вер.	Среднe знаков	164	115	248	80	49	0.590	0.701	0.641	4.373
10. ИИТ7 - частный критерий: разн услн безудл вероятностей, вер.	Семантический резонанс: зна.	164	115	248	80	49	0.590	0.701	0.641	62.563
10. ИИТ7 - частный критерий: разн услн безудл вероятностей, вер.	Среднe знаков	164	115	248	80	49	0.590	0.701	0.641	4.373

Рисунок 14 – Оценка достоверности моделей (первая часть)

3.4. Обобщ.форма по достов.моделям при разн.крит. Текущая модель: "ИИТ"

Наименование модели и частного критерия	Интегральный критерий	Среднe число реш. истинных (СТР)	Среднe число реш. ложных (ЛН)	3-Точность модели	5-Точность модели	Г-мера пров. С.В. Розберга	Среднее число истинных позит. решений (с.д.с.) истинный (СТР)	Среднее число истинных негат. решений (с.д.с.) истинный (ЛН)	Среднее число ложных позит. решений (с.д.с.) ложный (ЛР)	Среднее число ложных негат. решений (с.д.с.) ложный (ЛН)
1. ABS - частный критерий: количество встреч сигнала "1" на...	Корреляция abs частот с обр.	154.754	0.014	0.442	1.000	0.412	0.756	0.830	0.587	0.007
1. ABS - частный критерий: количество встреч сигнала "1" на...	Средн abs частот по границе	11.395		0.912	1.000	0.954	0.732		0.037	0.007
2. PRC1 - частный критерий: усл. вероятность того признака сред.	Корреляция усл.отн частот с о.	154.754	0.014	0.442	1.000	0.412	0.756	0.830	0.587	0.007
2. PRC1 - частный критерий: усл. вероятность того признака сред.	Средн усл.отн частот по грани	164.205		0.418	1.000	0.590	0.719		0.537	0.007
3. PRC2 - частный критерий: условная вероятность того признака сред.	Корреляция усл.отн частот с о.	154.755	0.014	0.442	1.000	0.412	0.756	0.830	0.587	0.007
3. PRC2 - частный критерий: усл. вероятность того признака сред.	Средн усл.отн частот по грани	164.205		0.418	1.000	0.590	0.719		0.537	0.007
4. ИИТ - частный критерий: количество знаков по А.Харкевичу, в.	Семантический резонанс: зна.	6.450	14.154	0.431	0.256	0.321	0.198	0.377	0.157	0.103
4. ИИТ - частный критерий: количество знаков по А.Харкевичу, в.	Среднe знаков	9.822	0.950	0.390	0.918	0.547	0.043	0.642	0.087	0.026
5. ИИТ2 - частный критерий: количество знаков по А.Харкевичу, в.	Семантический резонанс: зна.	6.450	14.154	0.431	0.256	0.321	0.198	0.377	0.157	0.103
5. ИИТ2 - частный критерий: количество знаков по А.Харкевичу, в.	Среднe знаков	9.822	0.950	0.390	0.918	0.547	0.043	0.642	0.087	0.026
6. ИИТ3 - частный критерий: Знаки надг. различия между фактн.	Семантический резонанс: зна.	15.843	8.197	0.809	0.491	0.588	0.583	0.449	0.230	0.167
6. ИИТ3 - частный критерий: Знаки надг. различия между фактн.	Среднe знаков	15.308	11.527	0.850	0.889	0.873	0.802	0.371	0.222	0.226
7. ИИТ4 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс: зна.	2.493	19.604	0.713	0.206	0.318	0.195	0.357	0.076	0.142
7. ИИТ4 - частный критерий: ROI (Return On Investment), вероятно.	Среднe знаков	2.493	0.337	0.569	0.999	0.722	0.023	0.603	0.022	0.002
8. ИИТ5 - частный критерий: ROI (Return On Investment), вероятно.	Семантический резонанс: зна.	2.494	19.604	0.713	0.206	0.319	0.195	0.357	0.076	0.142
8. ИИТ5 - частный критерий: ROI (Return On Investment), вероятно.	Среднe знаков	2.493	0.337	0.569	0.999	0.722	0.023	0.603	0.022	0.002
9. ИИТ6 - частный критерий: разн услн безудл вероятностей, вер.	Семантический резонанс: зна.	14.188	8.970	0.815	0.876	0.884	0.544	0.224	0.177	0.183
9. ИИТ6 - частный критерий: разн услн безудл вероятностей, вер.	Среднe знаков	10.637	0.217	0.291	0.453	0.446	0.038	0.130	0.133	0.004
10. ИИТ7 - частный критерий: разн услн безудл вероятностей, вер.	Семантический резонанс: зна.	14.188	8.970	0.815	0.876	0.884	0.544	0.224	0.177	0.183
10. ИИТ7 - частный критерий: разн услн безудл вероятностей, вер.	Среднe знаков	10.637	0.217	0.291	0.453	0.446	0.038	0.130	0.133	0.004

Рисунок 15 – Оценка достоверности моделей (вторая часть)

Наименование модели и критерий	Интегральный критерий	Число ошибок на тесте	А-Точность модели (RBF)	А-Точность модели (Анализ АТ/ФАП)	Сумма погр. ЭЭ Лямбда	Процент ошибок на тесте	Процент ошибок на валидации	Процент ошибок на тесте	Процент ошибок на валидации	Процент ошибок на тесте	Процент ошибок на валидации	Дата получения результата
1. ABS - частный критерий: количество ошибок (сигналы) "Упас"	Классовые абсолютные частоты с обр.	0.007	0.598	0.393	0.746	98.780	10.687	89.313	1.228	54.734	23.02.20	
1. ABS - частный критерий: количество ошибок (сигналы) "Упас"	Средняя абсолютная частота по признаку		0.951	1.000	0.495	108.090	0.185	99.818		50.893	23.02.20	
2. RBF1 - частный критерий: усл. вероятность (по признаку сред.)	Классовые доли частот с обр.	0.007	0.598	0.393	0.746	98.780	10.687	89.313	1.228	54.734	23.02.20	
2. RBF1 - частный критерий: усл. вероятность (по признаку сред.)	Средняя условная частота по признаку		0.573	1.000	0.728	108.090	0.185	99.818		50.893	23.02.20	
3. RBF2 - частный критерий: условная вероятность (по признаку сред.)	Классовые доли частот с обр.	0.007	0.598	0.393	0.746	98.780	10.687	89.313	1.228	54.734	23.02.20	
3. RBF2 - частный критерий: условная вероятность (по признаку сред.)	Средняя условная частота по признаку		0.573	1.000	0.728	108.090	0.185	99.818		50.893	23.02.20	
4. RBF1 - частный критерий: количество ошибок по А.Характерности	Семантический резонанс: зна.	0.103	0.544	0.647	0.591	15.854	99.318	0.661	84.144	57.586	23.02.20	
5. RBF1 - частный критерий: количество ошибок по А.Характерности	Средняя ошибка	0.026	0.332	0.423	0.432	87.195	97.985	2.035	12.301	32.590	23.02.20	
6. RBF2 - частный критерий: количество ошибок по А.Характерности	Семантический резонанс: зна.	0.103	0.544	0.647	0.591	15.854	99.318	0.661	84.144	57.586	23.02.20	
6. RBF2 - частный критерий: количество ошибок по А.Характерности	Средняя ошибка	0.026	0.332	0.423	0.432	87.195	97.985	2.035	12.301	32.590	23.02.20	
6. RBF3 - частный критерий: (по признаку: разности между частот.)	Семантический резонанс: зна.	0.167	0.717	0.777	0.746	79.122	67.134	32.886	29.878	68.618	23.02.20	
6. RBF3 - частный критерий: (по признаку: разности между частот.)	Средняя ошибка	0.235	0.703	0.773	0.778	79.122	67.134	32.886	29.878	68.618	23.02.20	
7. RBF4 - частный критерий: RD (Rankin On Inequality), вероятн.	Семантический резонанс: зна.	0.142	0.721	0.575	0.642	15.854	99.482	0.518	84.144	57.668	23.02.20	
7. RBF4 - частный критерий: RD (Rankin On Inequality), вероятн.	Средняя ошибка	0.052	0.504	0.325	0.465	87.195	97.985	2.035	12.301	32.590	23.02.20	
8. RBF5 - частный критерий: RD (Rankin On Inequality), вероятн.	Семантический резонанс: зна.	0.142	0.721	0.575	0.642	15.854	99.482	0.518	84.144	57.668	23.02.20	
8. RBF5 - частный критерий: RD (Rankin On Inequality), вероятн.	Средняя ошибка	0.052	0.504	0.325	0.465	87.195	97.985	2.035	12.301	32.590	23.02.20	
9. RBF6 - частный критерий: разн. усл. и безусловных вероятностей, вер.	Семантический резонанс: зна.	0.193	0.754	0.748	0.751	79.122	66.835	33.165	29.878	68.479	23.02.20	
9. RBF6 - частный критерий: разн. усл. и безусловных вероятностей, вер.	Средняя ошибка	0.004	0.222	0.896	0.356	79.122	66.835	33.165	29.878	68.479	23.02.20	
10. RBF7 - частный критерий: разн. усл. и безусловных вероятностей, ва...	Семантический резонанс: зна.	0.183	0.754	0.748	0.751	79.122	66.835	33.165	29.878	68.479	23.02.20	
10. RBF7 - частный критерий: разн. усл. и безусловных вероятностей, ва...	Средняя ошибка	0.004	0.222	0.896	0.356	79.122	66.835	33.165	29.878	68.479	23.02.20	

Рисунок 16 – Оценка достоверности моделей (третья часть)

Важно обратить внимание на то, что статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую.

На рисунке 17 представлены частотные распределения числа истинных и ложных положительных и отрицательных решений в СК-модели ABS по данным обучающей выборки.



Рисунок 17 – Частотные распределение модели ABS

Левое распределение включает только истинно-отрицательные, а правое включает ложно-положительные и истинно-положительные решения. Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу популярности определенных книг и другие задачи.

Для положительных решений от 17% до примерно 36% количество ложных решений больше положительных. Начиная от примерно 36% до примерно 48% количество истинных решений больше. Далее от примерно 88% до 100% количество истинных решений намного больше ложных.

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. 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.

Выберем одну из наиболее достоверных моделей INF4, что показано на рисунках 18 и 19.

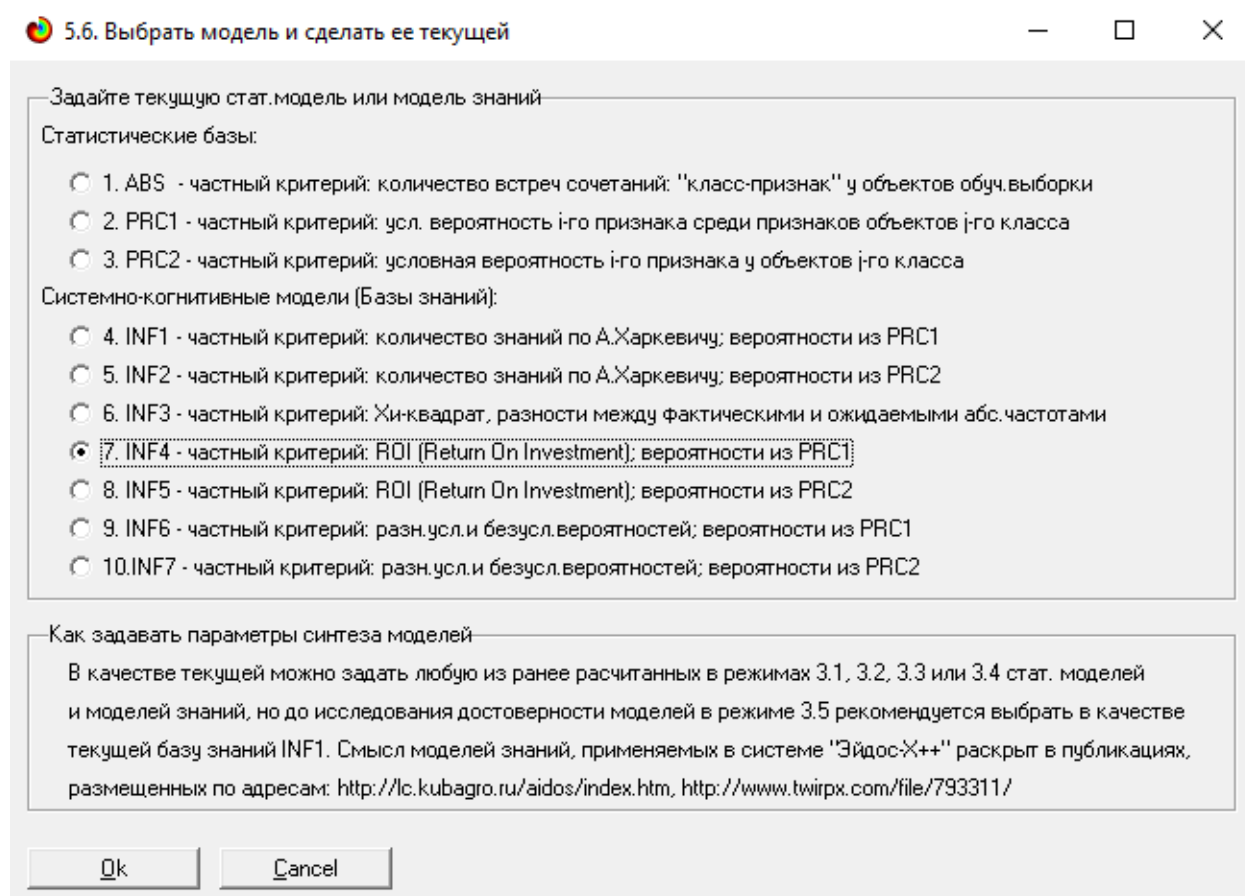


Рисунок 18 – Выбор текущей модели

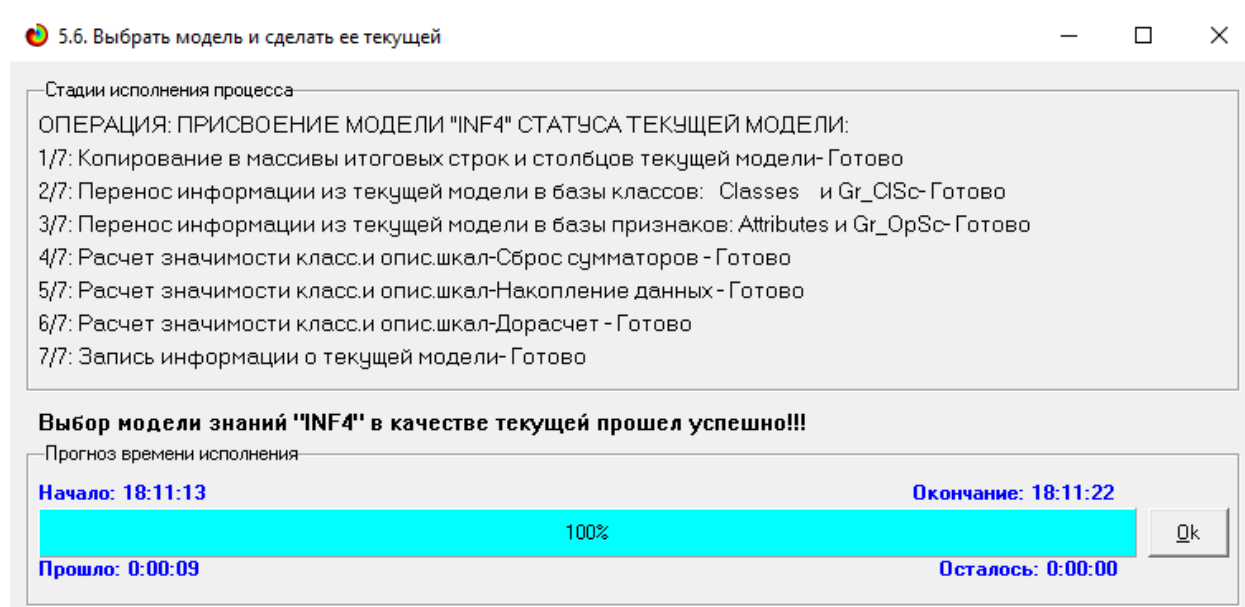


Рисунок 19 – Исполнение присвоения статуса текущей модели

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.

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.

Режиме 4.1.3 системы «Эйдос» обеспечивает отображение результатов идентификации и прогнозирования в различных формах:

- подробно наглядно: «Объект – классы»;
- подробно наглядно: «Класс – объекты»;
- итого наглядно: «Объект – класс»;
- итого наглядно: «Класс – объект»;
- подробно сжато: «Объекты – классы»;
- обобщенная форма по достоверности моделей при разных интегральных критериях;
- обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям;

- статистический анализ результатов идентификации по классам, моделям и интегральным критериям;
- достоверность идентификации объектов при разных моделях и интегральных критериях;
- достоверность идентификации классов при разных моделях и интегральных критериях;
- объединение в одной БД строк по самым достоверным моделям;
- вывод результатов распознавания в стиле «Inp_data.xlsx»;
- частотное распределение наблюдений по самым похожим классам;
- распределение уровней сходства наблюдений по всем классам.

На рисунках 20 и 21 приведены примеры прогноза высокой и низкой достоверности частоты объектов и классов в модели INF4 на основе наблюдения предыстории их развития. Визуализация результатов распознавания представлена в формах: «Объект – классы» и «Класс – объекты».

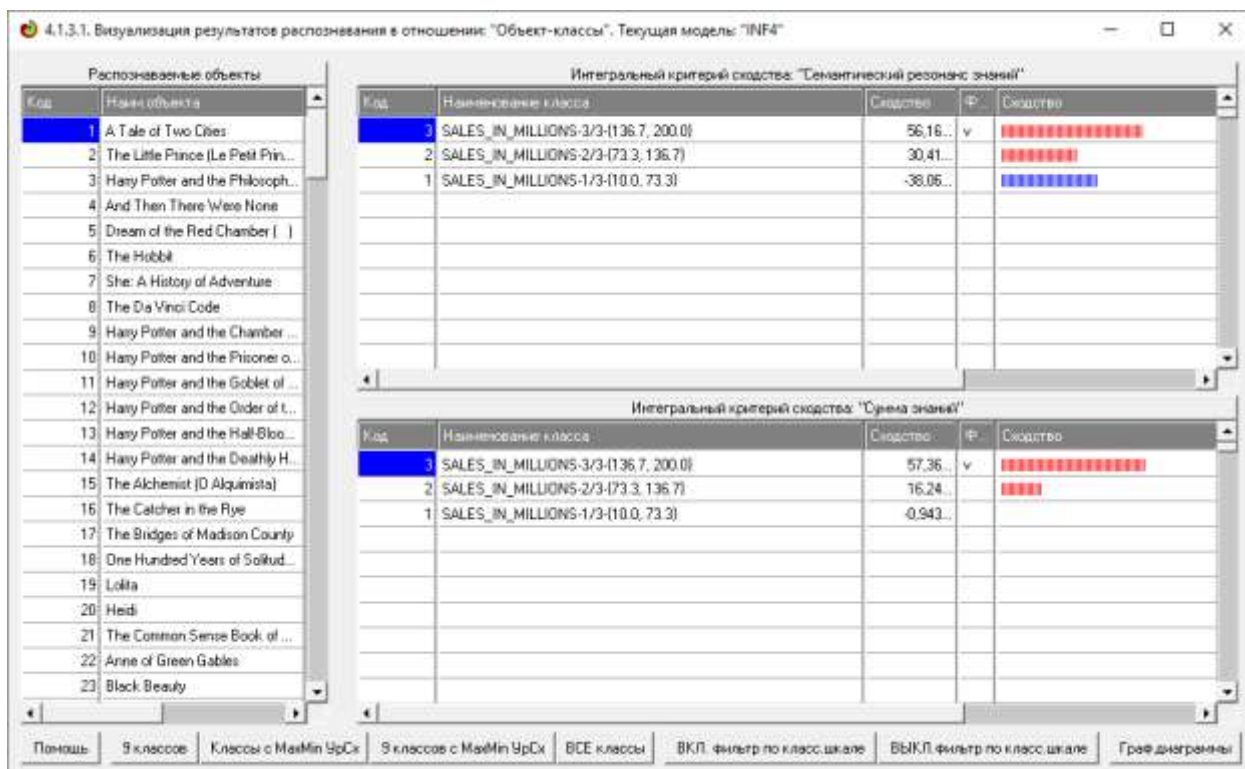


Рисунок 20 – Пример идентификации объектов в модели INF4

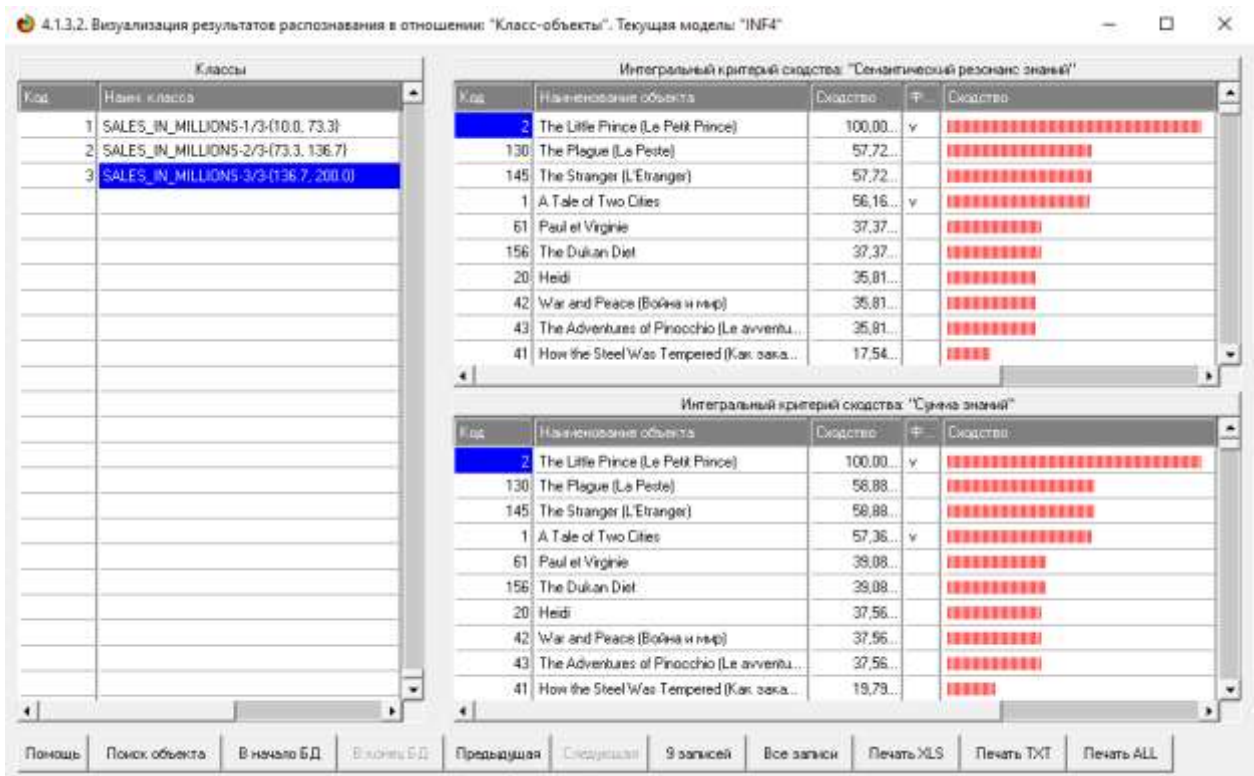


Рисунок 21 – Пример идентификации объектов в модели INF4

После данного этапа можно приступить к выбору наиболее достоверной модели.

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 режиме 5.5:

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

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

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

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);
– средняя информативность по вектору класса;
– среднее по вектору объекта;
– среднеквадратичное отклонение частных критериев знаний вектора класса;
– среднеквадратичное отклонение по вектору распознаваемого объекта.

– вектор состояния j -го класса; – 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: $I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}$, $L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}$. 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: $I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}$, $L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}}$, 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.

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»).

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.

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 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».

Инвертированная SWOT-диаграмма влияния авторства «Agatha Christie» популярность работы изображен на рисунке 23.

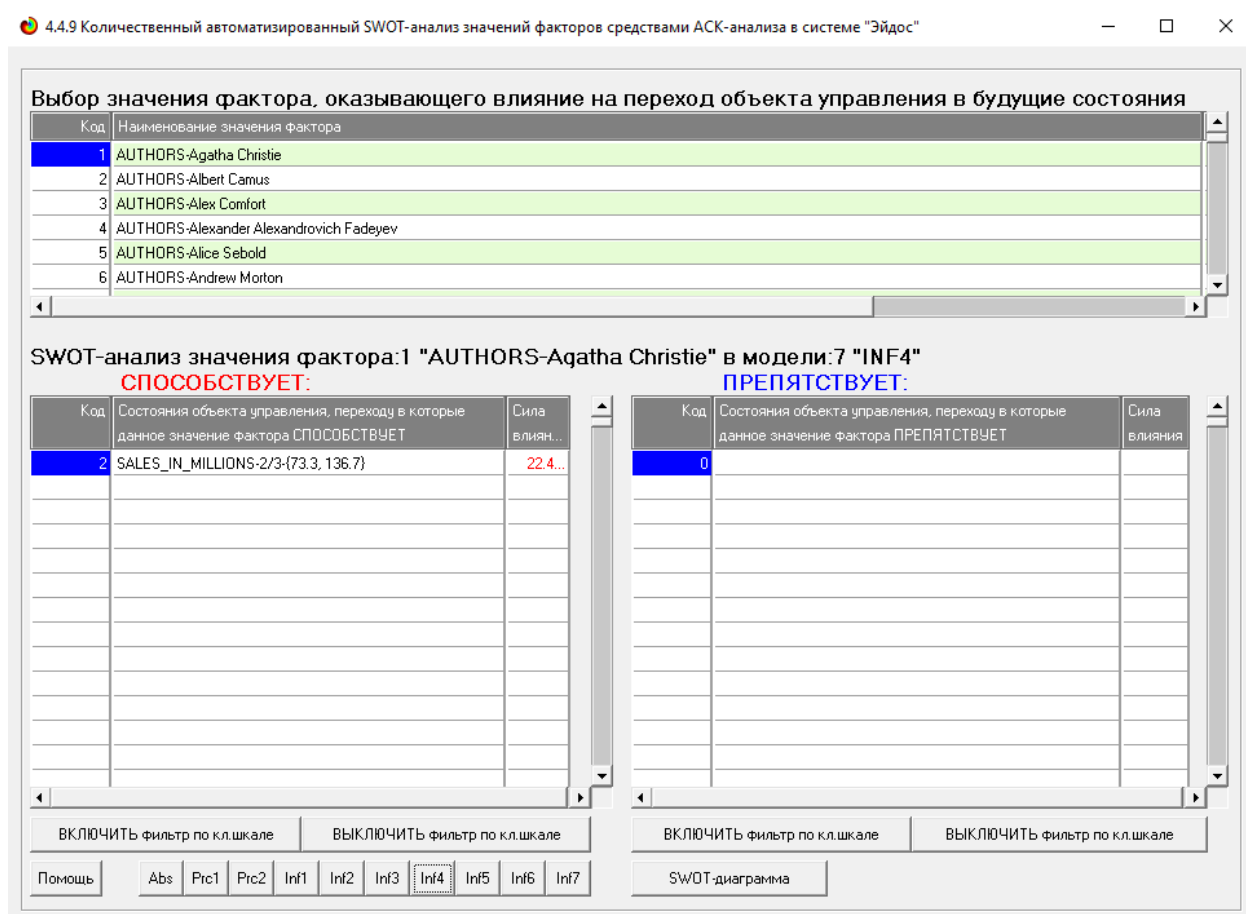


Рисунок 22 – Количественный автоматизированный SWOT-анализ значений факторов

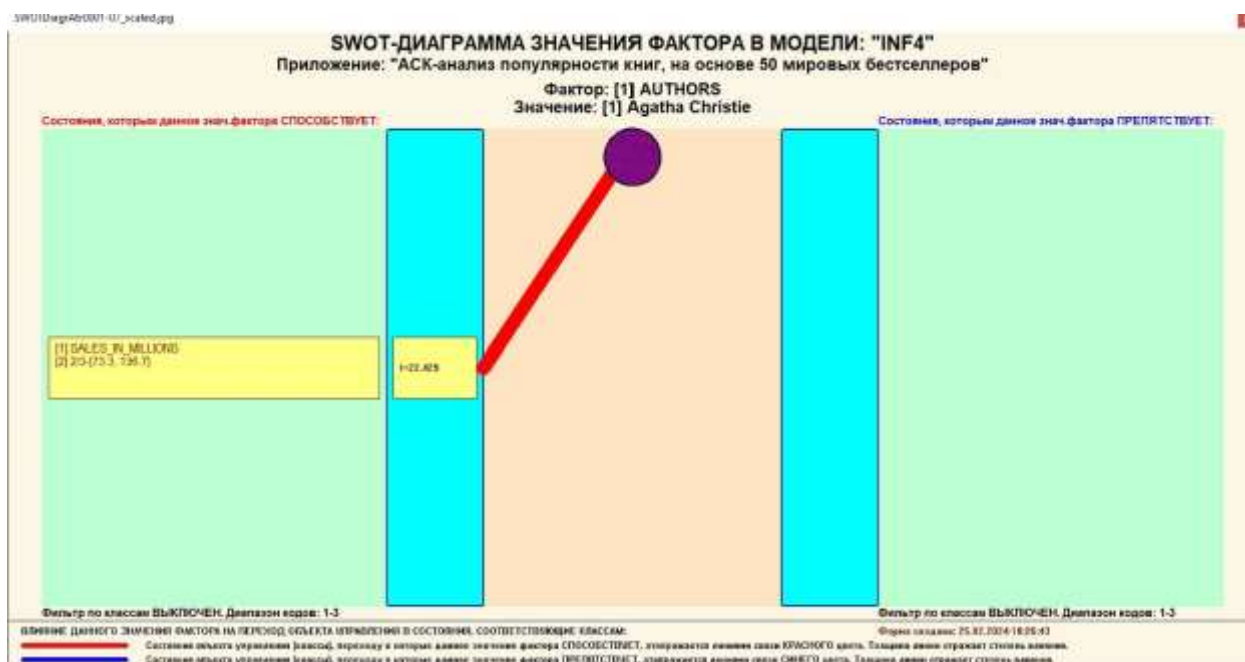


Рисунок 23 – Инвертированная SWOT-диаграмма

3.8.2 Кластерно-конструктивный анализ классов

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 25) рассчитывается матрица сходства классов 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 26);
- агломеративные дендрограммы, полученные в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.3) (рисунок 27);
- график изменения межкластерных расстояний (режим 4.2.2.3) (рисунок 28).

На рисунке 24 представлены экранные формы режима 4.2.2.1, обеспечивающего расчет матрицы сходства классов по системе их детерминации, т.е. по обуславливающим их значениям факторов.

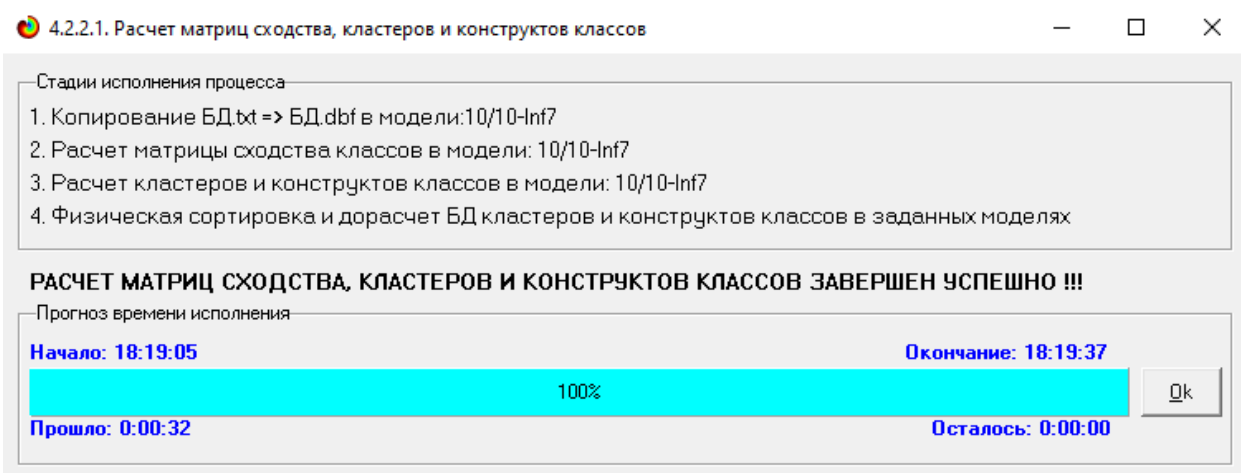


Рисунок 24 – Исполнение расчета матриц сходства, кластеров и конструкторов классов

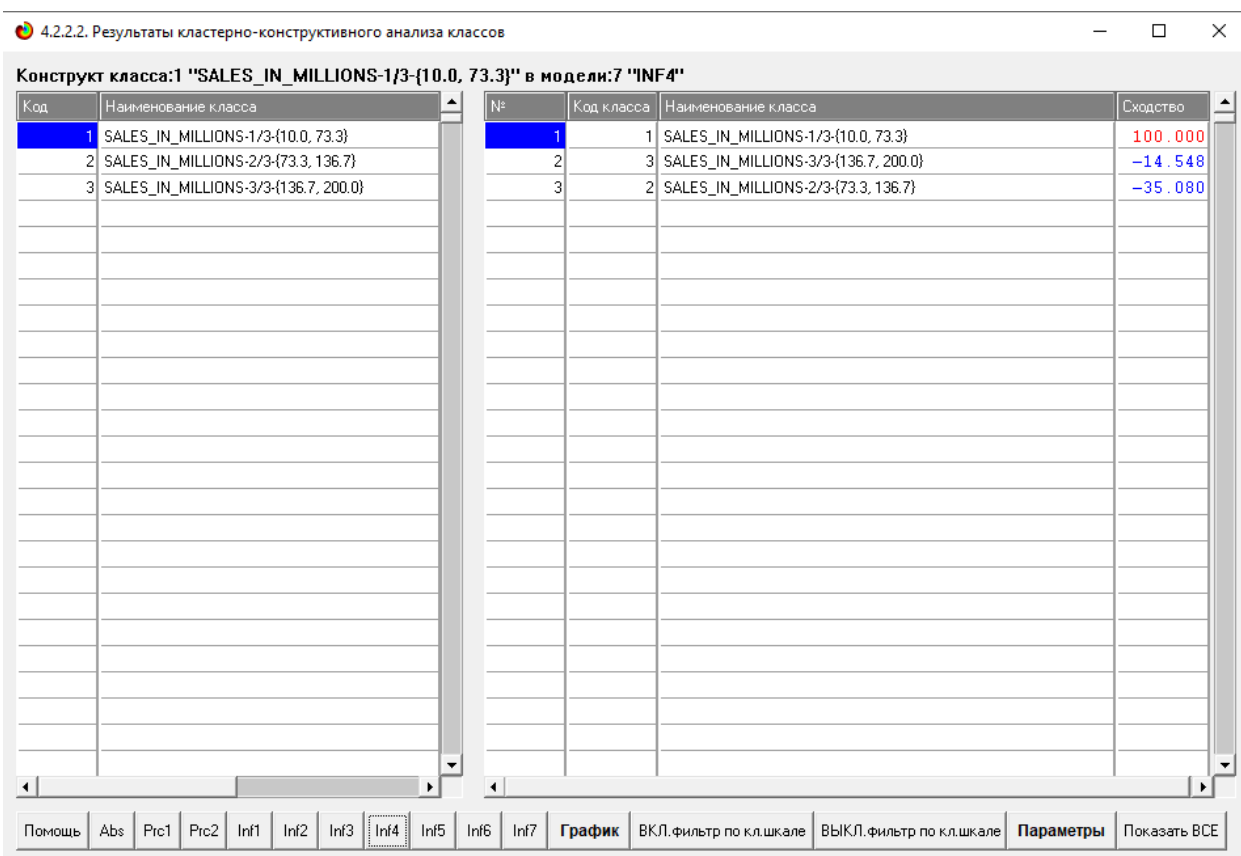


Рисунок 25 – Кластерно-конструктивный анализ модели INF4

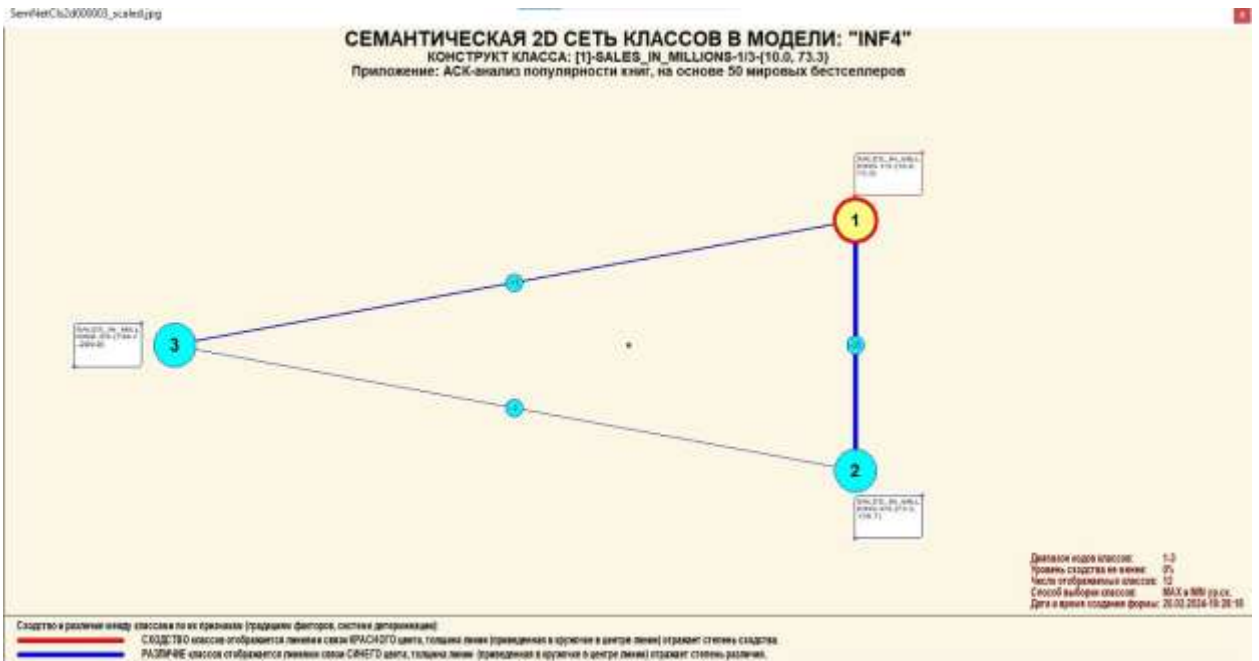


Рисунок 26 – Круговая 2D-когнитивная диаграмма классов

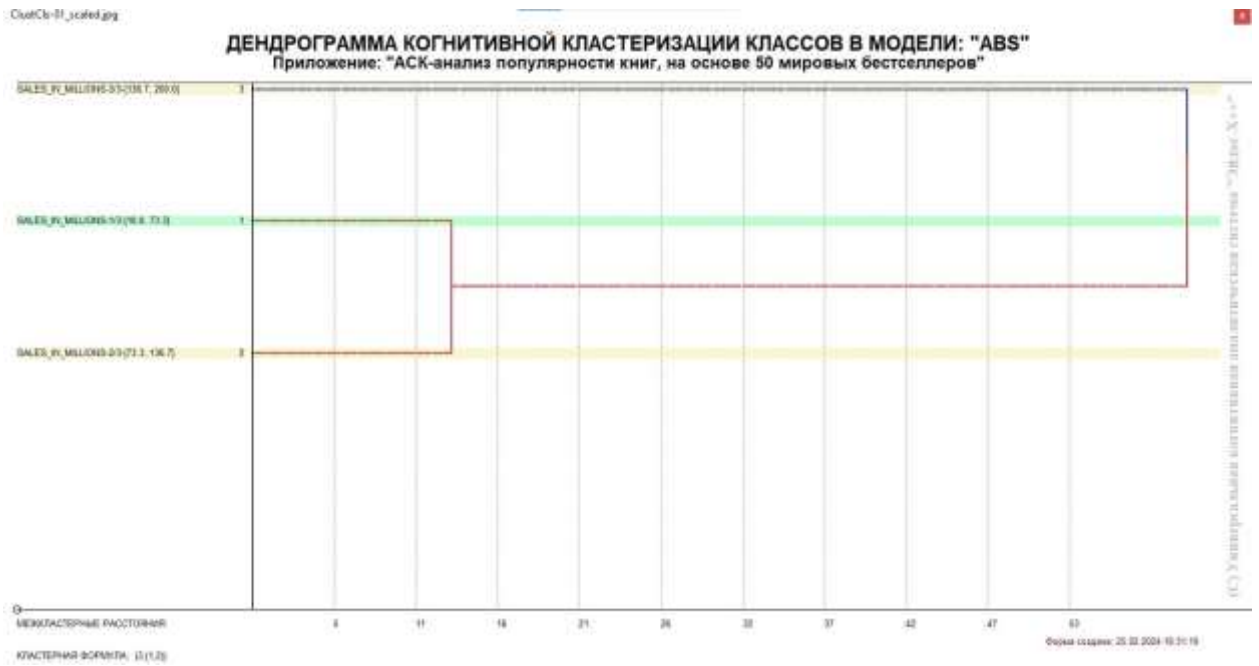


Рисунок 27 – Дендрограмма когнитивной кластеризации классов

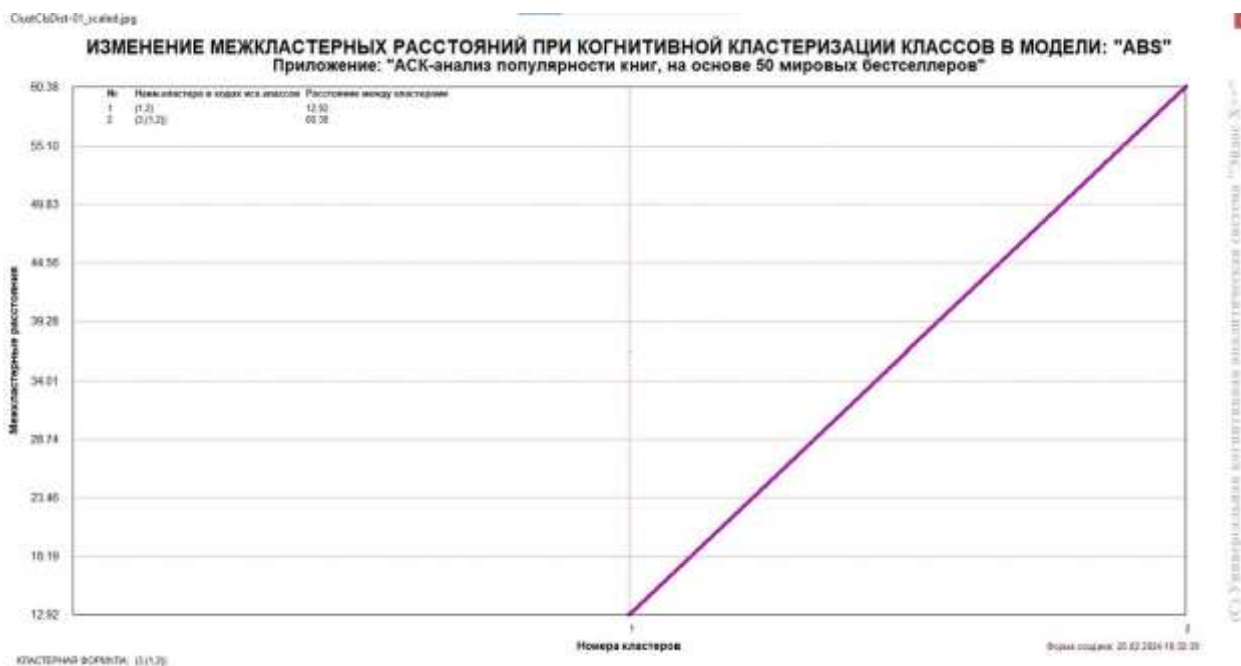


Рисунок 28 – График изменения межкластерных расстояний

3.8.3 Кластерно-конструктивный анализ значений описательных шкал

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 29) рассчитывается матрица сходства признаков по их смыслу и на основе этой матрицы рассчитывается и выводится три основных формы:

- круговая 2D-когнитивная диаграмма признаков (режим 4.3.2.2) рисунок 30);
- агломеративные дендрограммы, полученные в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3) рисунок 31);
- график изменения межкластерных расстояний (режим 4.3.2.3) (рисунок 32).

На рисунке 29 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:

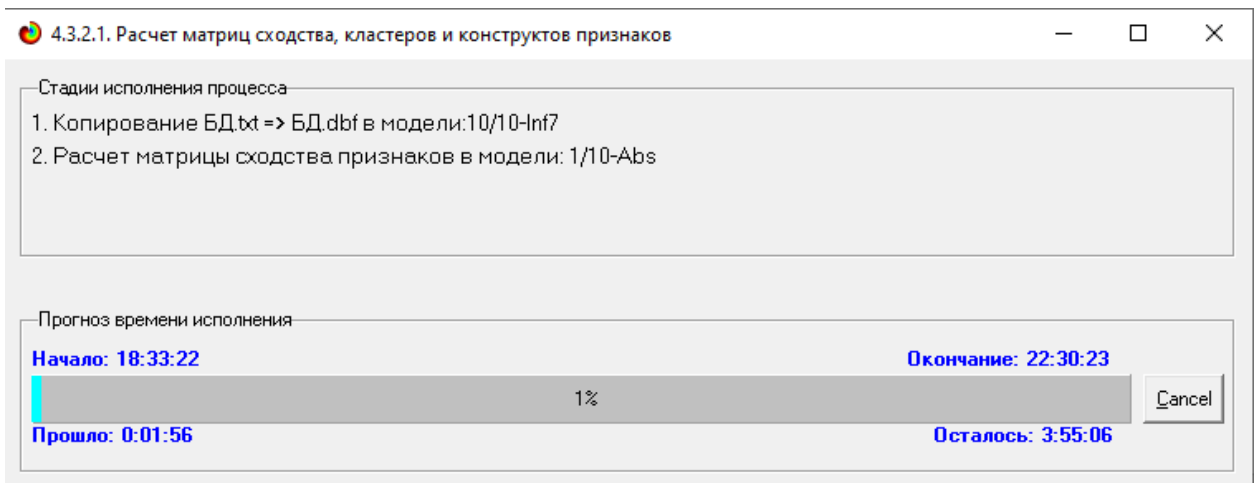


Рисунок 29 – Исполнение расчета матриц



Рисунок 30 – Круговая 2D-когнитивная диаграмма значений факторов в СК-модели ABS

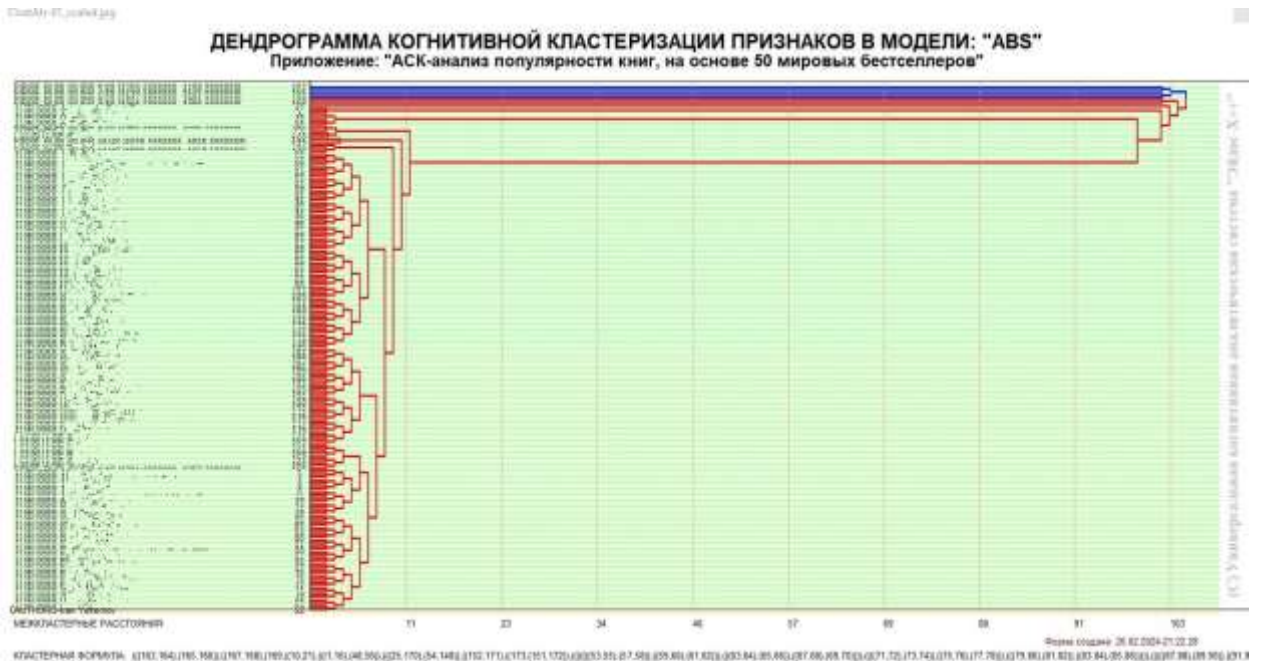


Рисунок 31 – Агломеративная дендрограмма, полученная в результате когнитивной кластеризации признаков

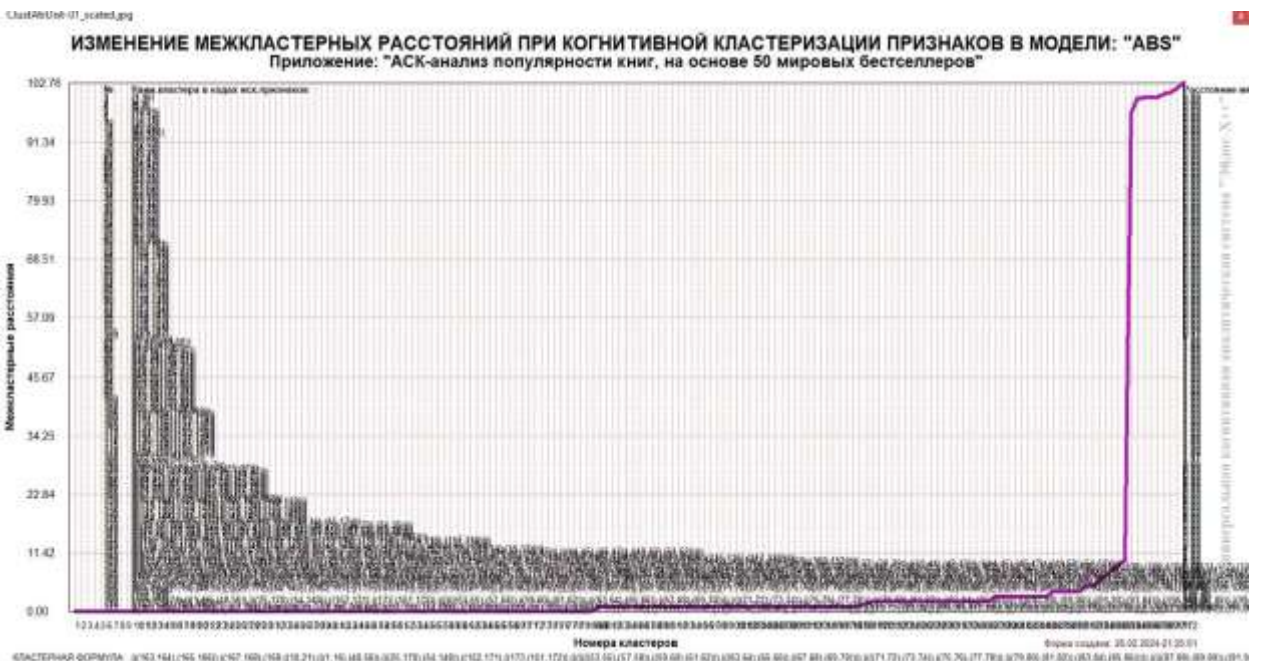


Рисунок 32 – График изменений межкластерных расстояний

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:

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 34). В форме управления визуализацией (рисунок 33) есть возможность задавать фильтры по факторам, которые надо визуализировать.

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	SALES_IN_MILLIONS-1/3-(10.0, 73.3)
2	SALES_IN_MILLIONS-2/3-(73.3, 136.7)
3	SALES_IN_MILLIONS-3/3-(136.7, 200.0)

Подготовка визуализации нейрона: 1 "SALES_IN_MILLIONS-1/3-(10.0, 73.3)" в модели: 7 "INF4"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
3	AUTHORS-Alex Comfort	0.058
4	AUTHORS-Alexander Alexandrovich Fadeyev	0.058
5	AUTHORS-Alice Sebold	0.058
6	AUTHORS-Andrew Morton	0.058
7	AUTHORS-Anna Sewell	0.058
8	AUTHORS-Anne Frank	0.058
9	AUTHORS-Anthony Doerr	0.058
11	AUTHORS-Arlene Eisenberg and Heidi Murkoff	0.058
12	AUTHORS-Astrid Lindgren	0.058

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
170	FIRST_PUBLISHED-9/12-(1780.0000000, 1839.5000000)	-0.295
25	AUTHORS-Dan Brown	-0.295
54	AUTHORS-J. K. Rowling	-0.244
171	FIRST_PUBLISHED-10/12-(1839.5000000, 1899.0000000)	-0.206
148	LANGUAGE-Chinese	-0.206
152	LANGUAGE-French	-0.154
172	FIRST_PUBLISHED-11/12-(1899.0000000, 1958.5000000)	-0.016
151	LANGUAGE-English	-0.002

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 33 – Форма управления визуализацией нейрона

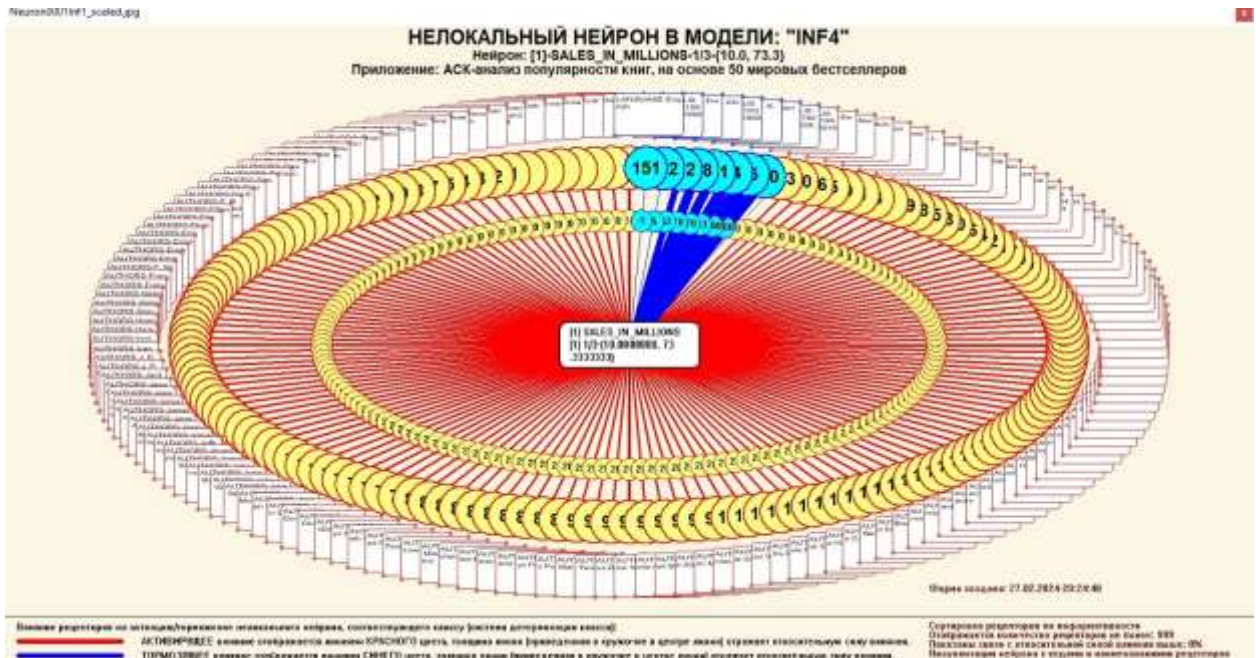


Рисунок 34 – Нелокальный нейрон, соответствующий классу SALES_IN_MILLIONS 1/3

3.8.5 Нелокальная нейронная сеть

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

Есть также возможность визуализации любого одного слоя нелокальной нейронной сети (режим 4.4.11 системы «Эйдос», рисунок 35).

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 36). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

4.4.11. Отображение Парето-подмножеств одного слоя нелокальной нейронной сети в системе "Эйдос"

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	SALES_IN_MILLIONS-1/3-{10.0, 73.3}
<input type="checkbox"/>	2	SALES_IN_MILLIONS-2/3-{73.3, 136.7}
<input checked="" type="checkbox"/>	3	SALES_IN_MILLIONS-3/3-{136.7, 200.0}

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

Подготовка визуализации нейрона: 1 "SALES_IN_MILLIONS-1/3-{10.0, 73.3}" в модели: 7 "INF4"

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

Код	Наименование фактора и его интервального значения	Сила влияния
3	AUTHORS-Alex Comfort	0.058
4	AUTHORS-Alexander Alexandrovich Fadeyev	0.058
5	AUTHORS-Alice Sebold	0.058
6	AUTHORS-Andrew Morton	0.058
7	AUTHORS-Anna Sewell	0.058
8	AUTHORS-Anne Frank	0.058
9	AUTHORS-Anthony Doerr	0.058
11	AUTHORS-Arlene Eisenberg and Heidi Murkoff	0.058
12	AUTHORS-Astrid Lindgren	0.058
13	AUTHORS-Bernard Werber	0.058

Код	Наименование фактора и его интервального значения	Сила влияния
170	FIRST_PUBLISHED-9/12-{1780.0000000, 1839.5000000}	-0.295
25	AUTHORS-Dan Brown	-0.295
54	AUTHORS-J. K. Rowling	-0.244
171	FIRST_PUBLISHED-10/12-{1839.5000000, 1899.0000000}	-0.206
148	LANGUAGE-Chinese	-0.206
152	LANGUAGE-French	-0.154
172	FIRST_PUBLISHED-11/12-{1899.0000000, 1958.5000000}	-0.016
151	LANGUAGE-English	-0.002

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

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

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

Рисунок 35 – Визуализация слоя нелокальной нейронной сети

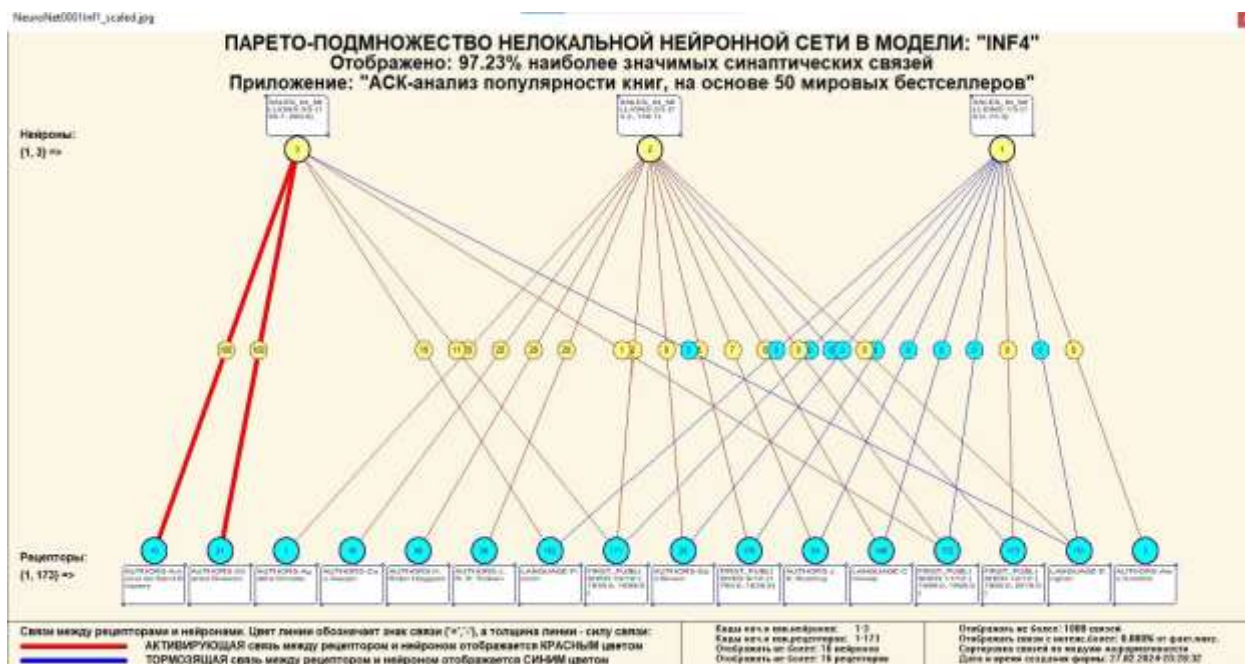


Рисунок 36 – Нейронная сеть в СК-модели INF4

3.8.6 3D-интегральные когнитивные карты

3D-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 26) вверху и когнитивной диаграммы значений факторов (рисунок 30) внизу и соединяющего их одного слоя нейронной сети (рисунок 36) (режим 4.4.12 системы «Эйдос» рисунок 37) (рисунок 38):

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

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

№	Код	Наименование нелокального нейрона (класса)
1	SALES_IN_MILLIONS-1/3-{10.0, 73.3}	
2	SALES_IN_MILLIONS-2/3-{73.3, 136.7}	
3	SALES_IN_MILLIONS-3/3-{136.7, 200.0}	

Помощь Максимальное количество отображаемых нейронов: 16 ClearSet Диапазон кодов отображаемых нейронов: 1 3
 Максимальное количество отображаемых связей: 1000 Диапазон кодов отображаемых рецепторов: 1 173

Подготовка визуализации нейрона:1 "SALES_IN_MILLIONS-1/3-{10.0, 73.3}" в модели:7 "INF4"

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

Код	Наименование фактора и его интервального значения	Сила влияния
3	AUTHORS-Alex Comfort	0.058
4	AUTHORS-Alexander Alexandrovich Fadeyev	0.058
5	AUTHORS-Alice Sebold	0.058
6	AUTHORS-Andrew Morton	0.058
7	AUTHORS-Anna Sewell	0.058
8	AUTHORS-Anne Frank	0.058
9	AUTHORS-Anthony Doerr	0.058
11	AUTHORS-Arlene Eisenberg and Heidi Murkoff	0.058
12	AUTHORS-Astrid Lindgren	0.058
13	AUTHORS-Bernard Yarkovitsky	0.058

Код	Наименование фактора и его интервального значения	Сила влияния
170	FIRST_PUBLISHED-9/12-{1780.0000000, 1839.5000000}	-0.295
25	AUTHORS-Dan Brown	-0.295
54	AUTHORS-J. K. Rowling	-0.244
171	FIRST_PUBLISHED-10/12-{1839.5000000, 1899.0000000}	-0.206
148	LANGUAGE-Chinese	-0.206
152	LANGUAGE-French	-0.154
172	FIRST_PUBLISHED-11/12-{1899.0000000, 1958.5000000}	-0.016
151	LANGUAGE-English	-0.002

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: 16 Отображать связи с интенсивностью >= % от макс.: 0,000

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

Рисунок 37 – Составление слоя интегральной когнитивной карты

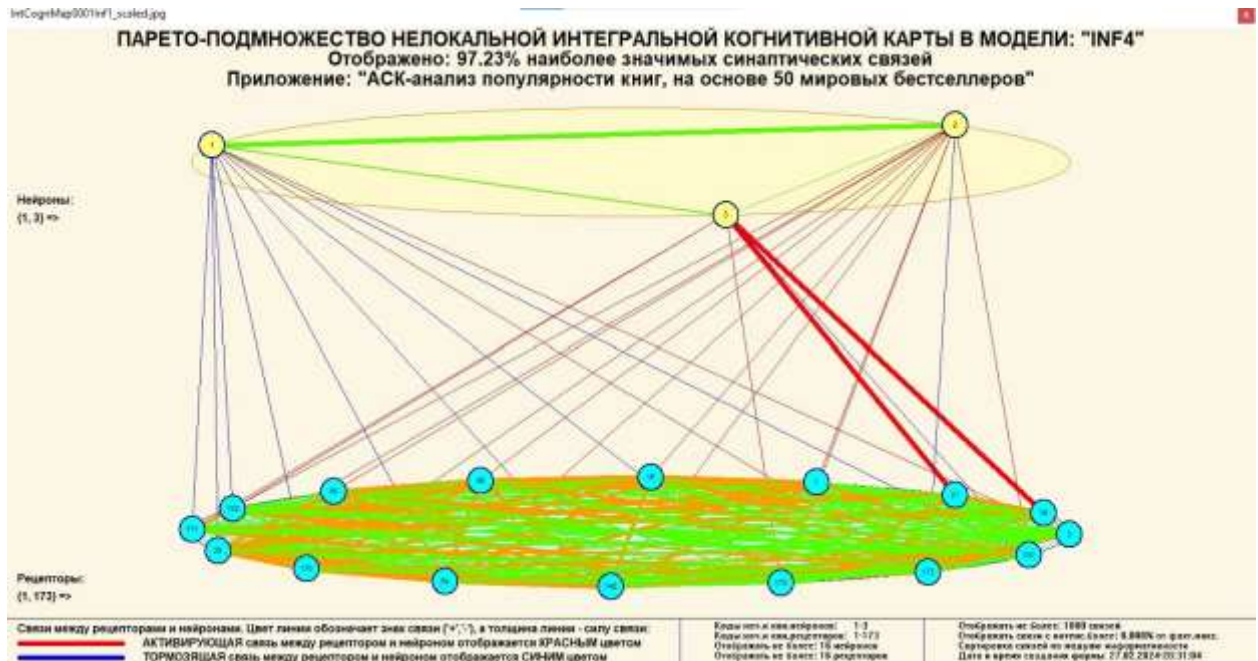


Рисунок 38 – 3D-когнитивная диаграмма классов и признаков

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. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году.

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 (рисунок 39) и 2D-интегральной когнитивной карты содержательного сравнения классов по их системе детерминации приведены ниже на рисунке 40. Всего системой в данной

модели генерируется 324 подобных диаграмм, поэтому, естественно, все они не приводятся.

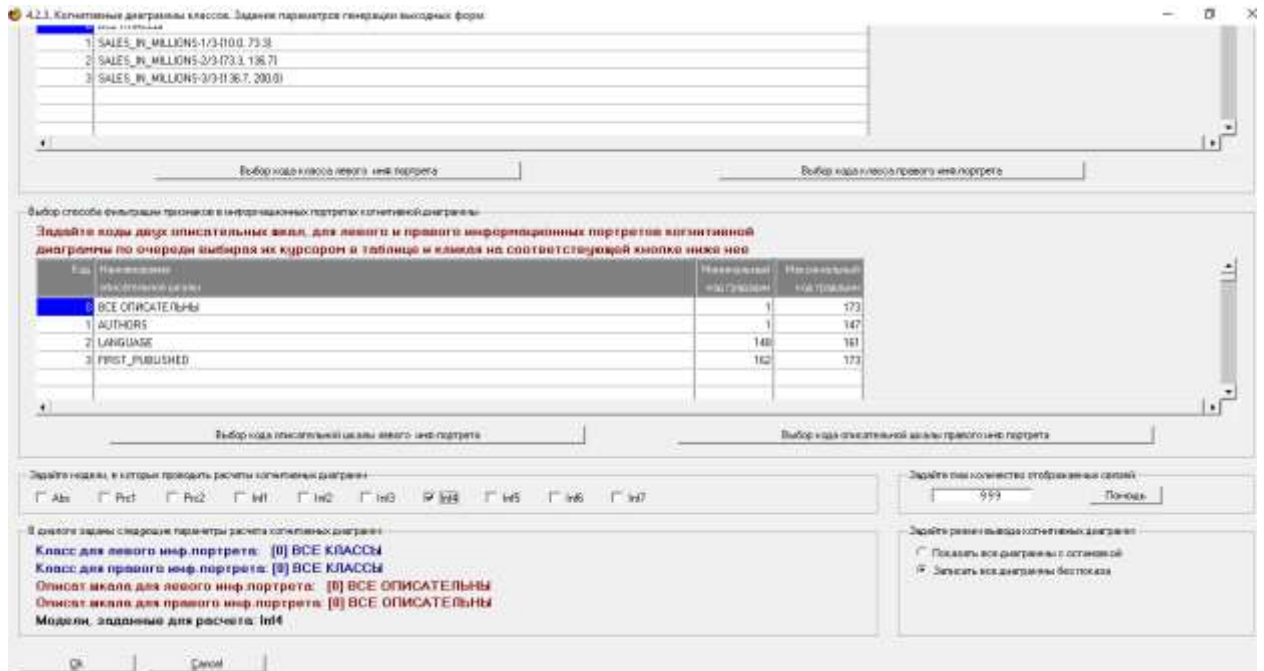


Рисунок 39 – Задание параметров генерации когнитивной диаграммы классов



Рисунок 40 – Пример 2D-интегральной когнитивной карты содержательного сравнения классов по их системе детерминации в СК-модели INF4

3.8.8 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечетки 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, согласно концепции смысла Шенка-Абельсона, используемой в АСК-анализе, состоит в знании причин и последствий.

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

Это видно из когнитивных диаграмм, которые можно получить в режиме 4.3.3 системы «Эйдос» (рисунок 41).

Примеры экранной формы управления и 2D-интегральной когнитивной карты содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта моделирования в будущие состояния, соответствующие классам, приведена ниже на рисунке 42.

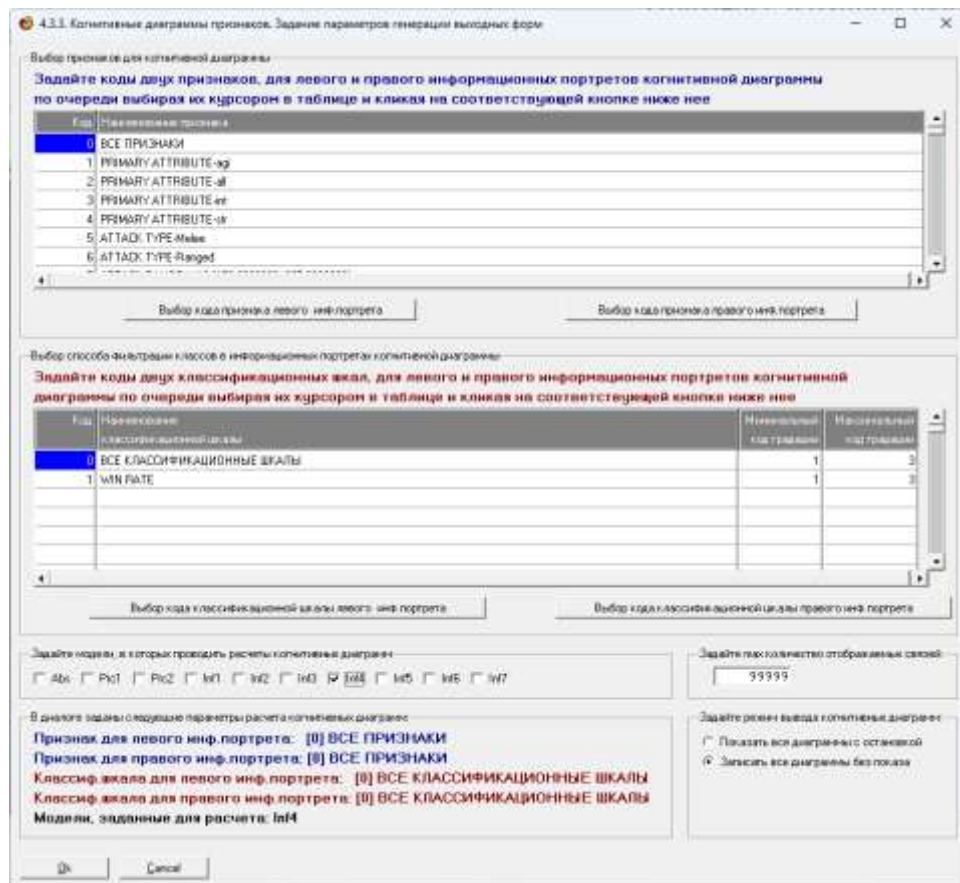


Рисунок 41 – Задание параметров генерации когнитивной диаграммы признаков



Рисунок 42 – Пример интегральной когнитивной карты содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующей классу в СК-модели INF3

Всего системой в данной модели генерируется 156 подобных диаграмм, поэтому, естественно, все они не приводятся.

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.

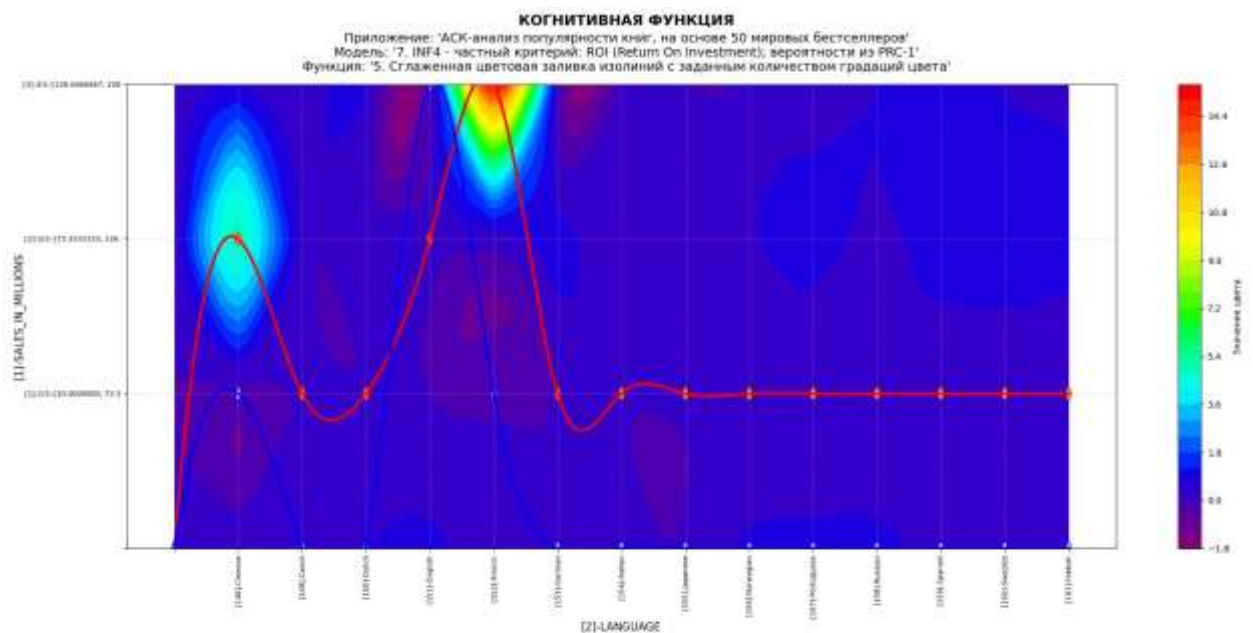
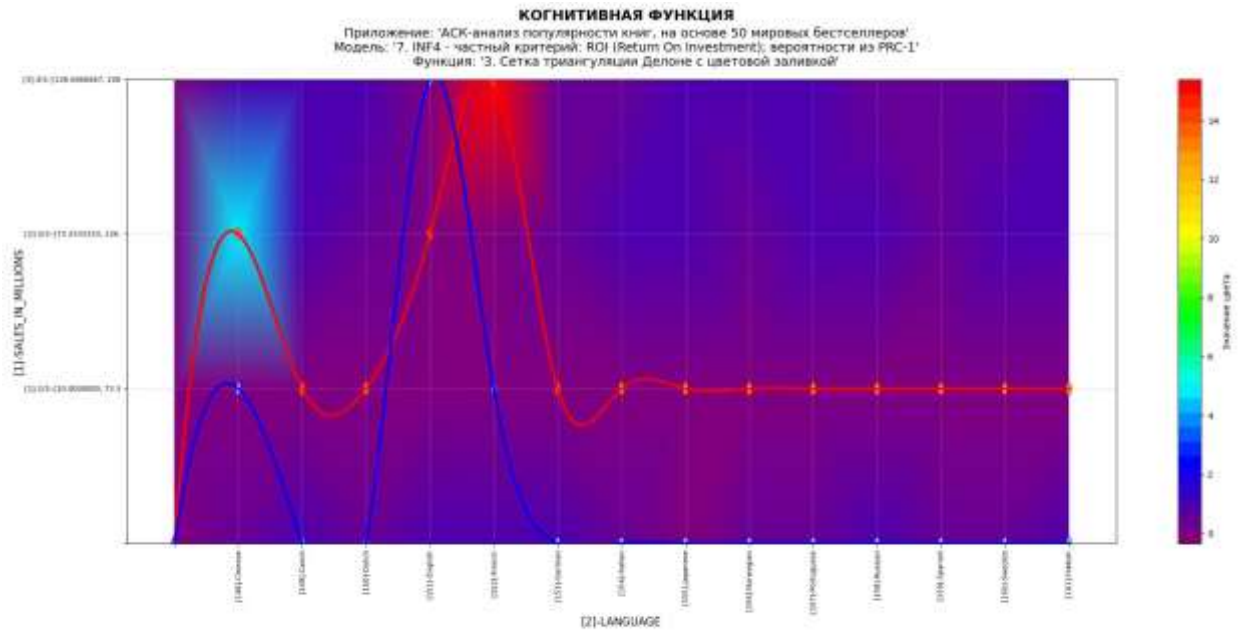
Когнитивные функции отображают, какое количество информации содержится в градациях описательной шкалы о переходе объекта моделирования в состояния, соответствующие градациям классификационной шкалы. При этом в статистических и системно-когнитивных моделях в каждой градации описательной шкалы содержится 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, имеющихся 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 43). 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. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах.



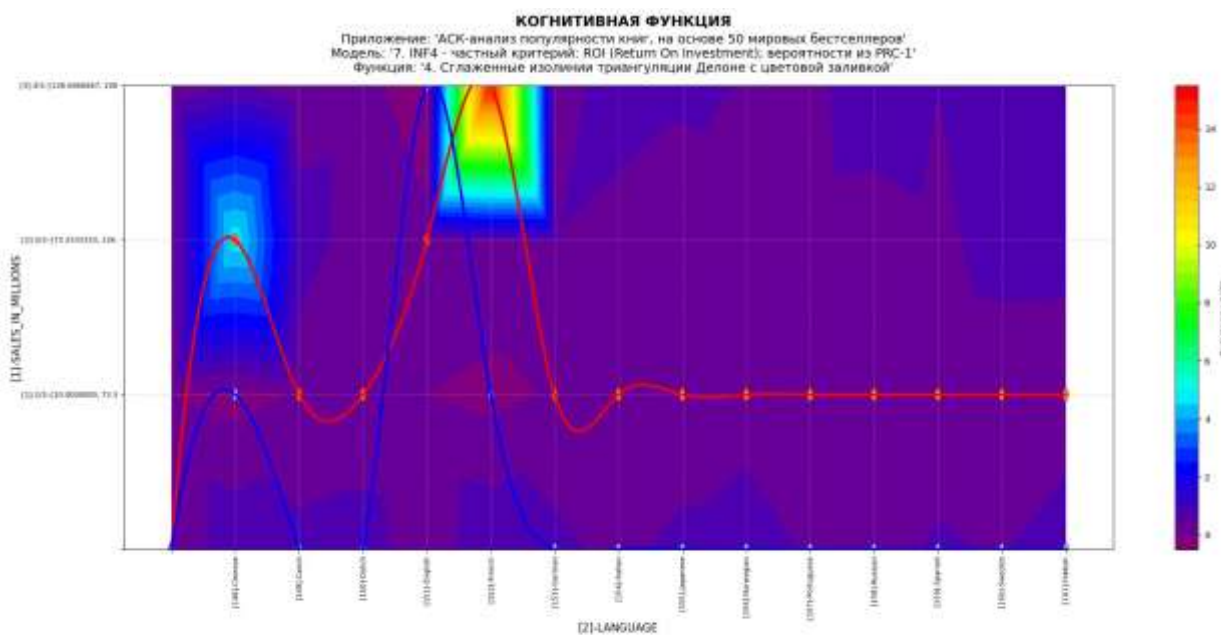


Рисунок 43 – Примеры когнитивных функций в СК-модели INF4

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

3.8.10 Значимость описательных шкал и 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.

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



Рисунок 44 – Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4.

Из рисунка 44 видно, что примерно пятая часть наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 85% суммарного влияния.

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 рисунке 45 приведена экранная форма режима 3.7.3 системы «Эйдос», содержащая информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



Рисунок 45 – Парето-кривая степени детерминированности классов в СК-модели INF4

4 ОБСУЖДЕНИЕ

Сверхзадачей искусственного интеллекта является построение компьютерной интеллектуальной системы, которая обладала бы уровнем эффективности решений неформализованных задач, сравнимым с человеческим или превосходящим его. Самую существенную часть систем искусственного интеллекта составляют экспертные системы. Экспертная система обычно определяется как программа ЭВМ, моделирующая действия эксперта человека при решении задач в узкой предметной области: составление базы знаний и накопления их.

В данной курсовой работе было показано построение модели зависимости популярности книг, на основе 50 мировых бестселлеров системой искусственного интеллекта «Эйдос».

При этом наиболее достоверной в данном приложении оказалась модель ABS, при интегральном критерии «Сумма абсолютных частот по признакам». Точность модели составила 0.954, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 49,4%.

ЗАКЛЮЧЕНИЕ

В заключение, проведенное исследование системно-когнитивного анализа популярности книг с использованием системы «Эйдос», позволило получить важные данные для книгоиздательств. Анализ методов формирования обобщенных образов классов, а также исследование решений для идентификации конкретных объектов с классами принятия решений, дали нам возможность лучше понять факторы, влияющие на успех печатного произведения.

Системно-когнитивный подход, примененный в данном исследовании, предоставил более глубокий анализ того, что интересно читателям. Полученные результаты могут быть использованы для написания новых работ.

Таким образом, результаты данного исследования могут иметь практическое значение для молодых авторов и книгоизданий.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Автоматизированный системно-когнитивный анализ силы и направления влияния морфологических свойств помидоров на количественные, качественные и финансово-экономические результаты их выращивания и степень детерминированности этих результатов в условиях неотапливаемых теплиц Юга России / Е.В. Луценко, Р.А.Гиш, Е.К. Печурина, С.С. Цыгикало // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. №06(150). С. 92 – 142. – IDA [articleID]:1501906015. – Режим доступа: [http://ej.kubagro.ru/2019/06/pdf/15.pdf,3,188у.п.л.](http://ej.kubagro.ru/2019/06/pdf/15.pdf,3,188у.п.л.;);

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): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.;

3. Луценко Е.В. АСК-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: Куб- ГАУ, 2014. – №06(100). С. 1346 – 1395. – IDA [article ID]: 1001406090. – Режим доступа: [http://ej.kubagro.ru/2014/06/pdf/90.pdf, 3,125 у.п.л.](http://ej.kubagro.ru/2014/06/pdf/90.pdf,3,125 у.п.л.;);

4. Луценко Е.В. Количественная оценка степени манипулирования индексом Хирша и его модификация, устойчивая к манипулированию / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №07(121). С. 202 – 234. – IDA [article ID]: 1211607005. – Режим доступа:

<http://ej.kubagro.ru/2016/07/pdf/05.pdf>,

2,062

у.п.л..

<http://dx.doi.org/10.21515/1990-4665-121-0057.;>

5. Луценко Е.В. Метод визуализации когнитивных функций – новый инструмент исследования эмпирических данных большой размерности /Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №03(067). С. 240 – 282. – Шифр Информрегистра: 0421100012\ 0077, IDA [articleID]:0671103018. – Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf> ,2,688у.п.л;

6. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/01.pdf>, 3,75 у.п.л.3.;

7. Луценко Е.В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы «Эйдос» (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е.В. Луценко, Н.В. Андрафанова, Н.В. Потапова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №01(145). С. 31 – 102. – IDA [article ID]: 1451901033. – Режим доступа: <http://ej.kubagro.ru/2019/01/pdf/33.pdf>, 4,5 у.п.л.;

8. 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 (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №01(001). С. 79 – 91. – IDA [article ID]: 0010301011. – Режим доступа: <http://ej.kubagro.ru/2003/01/pdf/11.pdf>, 0,812 у.п.л.;

9. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- X++» Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №09(083). С. 328 – 356. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 у.п.л.;

10. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://lc.kubagro.ru/>, свободный. - Загл. с экрана. Яз.рус.