

УДК 004.8
09.03.03. Прикладная информатика

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ ВЛИЯНИЯ ПОКАЗАТЕЛЕЙ НА ОДОБРЕНИЕ БАНКОВСКОГО КРЕДИТА

Чувилев Егор Станиславович
Студент факультета ПИ, группы ПИ2203
egor23ru6@mail.ru

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

В условиях цифровизации банковского сектора и ужесточения требований к управлению кредитными рисками важно повышать точность прогнозирования одобрения кредитов физическим лицам. Однако этому могут препятствовать сложные нелинейные взаимосвязи между характеристиками заемщика и вероятностью дефолта, которые традиционные статистические методы не всегда способны выявить. В ряде научных работ обосновывается перспективность использования методов машинного обучения для повышения качества скоринговых моделей. Для анализа эмпирических данных использованы традиционные методы. В данной статье для анализа эмпирических данных в области банковского скоринга предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исследований в области кредитного скоринга, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей в данной предметной области.

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

UDC 004.8 UDC 004.8
09.03.03. Applied computer science

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF THE IMPACT OF INDICATORS ON BANK LOAN APPROVAL

Chuvilov Egor Stanislavovich
Student of the faculty of PI, group PI2203
egor23ru6@mail.ru

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

In the context of digitalization of the banking sector and stricter requirements for credit risk management, it is important to increase the accuracy of forecasting credit approvals for individuals. However, this may be hindered by complex non-linear relationships between the characteristics of the borrower and the probability of default, which traditional statistical methods are not always able to identify. A number of scientific papers substantiate the prospects of using machine learning methods to improve the quality of scoring models. Traditional methods have been used to analyze empirical data. In this article, it is proposed to use Automated System Cognitive Analysis (ASK analysis) and its software toolkit, the intelligent Eidos system, to analyze empirical data in the field of banking scoring. A detailed numerical example is given. This example contains many different visual tabular and graphical output forms and can be used to teach the use of ASK analysis and the Eidos system for scientific research in the field of credit scoring, to develop practical recommendations and substantiate scientific provisions on the mechanisms of causal relationships in this subject area.

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

<http://dx.doi.org/10.21515/1990-4665-197-009>

<http://ej.kubagro.ru/2024/03/pdf/09.pdf>

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ).....	4
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	4
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	5
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	5
1.4. ЦЕЛЬ РАБОТЫ.....	5
2. METHODS (МЕТОДЫ).....	5
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	5
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	6
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	6
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА.....	8
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	15
3. RESULTS (РЕЗУЛЬТАТЫ)	17
3.1. Задача-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	17
3.1.1. Универсальная постановка задачи, не зависящая от предметной области	17
3.1.2. Конкретное решение задачи в данной работе	18
3.2. Задача-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	19
3.2.1. Универсальная постановка задачи, не зависящая от предметной области	19
3.2.2. Конкретное решение задачи в данной работе	20
3.3. Задача-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ.....	27
3.3.1. Универсальная постановка задачи, не зависящая от предметной области	27
3.3.2. Конкретное решение задачи в данной работе	35
3.4. Задача-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	40
3.4.1. Универсальная постановка задачи, не зависящая от предметной области	40
3.4.2. Конкретное решение задачи в данной работе	40
3.5. Задача-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	43
3.5.1. Универсальная постановка задачи, не зависящая от предметной области	43
3.5.2. Конкретное решение задачи в данной работе	43
3.6. Задача-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	45
3.6.1. Универсальная постановка задачи, не зависящая от предметной области	45
3.6.1.1. Интегральный критерий «Сумма знаний»	46
3.6.1.2. Интегральный критерий «Семантический резонанс знаний».....	47
3.6.1.3. Важные математические свойства интегральных критериев.....	48
3.6.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»	49
3.7. Задача-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ.....	53
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	53
3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области	53
3.7.1.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос».....	54
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	56
3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области	56
3.7.2.2. Конкретное решение задачи управления в данной работе в системе «Эйдос»	60
3.8. Задача-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	61
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы).....	61
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области	61
3.8.1.2. Конкретное решение задачи в данной работе	61
3.8.2. Кластерно-конструктивный анализ классов.....	66

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области	66
3.8.2.2. Конкретное решение задачи в данной работе	66
3.8.3. Кластерно-конструктивный анализ значений описательных шкал	66
3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области	66
3.8.3.2. Конкретное решение задачи в данной работе	66
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны.....	67
3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области	67
3.8.4.2. Конкретное решение задачи в данной работе	68
3.8.5. Нелокальная нейронная сеть.....	69
3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области	69
3.8.5.2. Конкретное решение задачи в данной работе	69
3.8.6. 3D-интегральные когнитивные карты Ошибка! Закладка не определена.	
3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области	Ошибка! Закладка не определена.
3.8.6.2. Конкретное решение задачи в данной работе	Ошибка! Закладка не определена.
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	70
3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области	70
3.8.7.2. Конкретное решение задачи в данной работе	71
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	72
3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области	72
3.8.8.2. Конкретное решение задачи в данной работе	73
3.8.9. Когнитивные функции	74
3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области	74
3.8.9.2. Конкретное решение задачи в данной работе	75
3.8.10. Значимость описательных шкал и их градаций.....	85
3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области	85
3.8.10.2. Конкретное решение задачи в данной работе	86
3.8.11. Степень детерминированности классов и классификационных шкал	90
3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области	90
3.8.11.2. Конкретное решение задачи в данной работе	91
4. DISCUSSION (ОБСУЖДЕНИЕ)	93
5. CONCLUSIONS (ВЫВОДЫ)	94
REFERENCES (ЛИТЕРАТУРА).....	95

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Предметная область исследования относится к сфере банковского кредитования, а именно — к процессу принятия решений о выдаче кредитов физическим лицам. В современной банковской практике ключевую роль играет оценка кредитоспособности заемщика, которая позволяет прогнозировать вероятность своевременного возврата заемных средств. Эта оценка реализуется через систему кредитного скоринга — совокупность статистических и математических моделей, которые на основе множества факторов формируют итоговое решение: одобрить кредит или отказать в выдаче. Скоринговые модели используются банками для автоматизации процессов, снижения влияния человеческого фактора и обеспечения единообразия подходов к оценке клиентов.

Основными источниками данных для скоринга являются анкетные данные заемщика (возраст, пол, трудовой стаж), его финансовые показатели (уровень дохода, запрашиваемая сумма кредита, соотношение платежа к доходу), а также параметры, отражающие кредитную историю и благонадежность (продолжительность кредитной истории, кредитный рейтинг). В условиях цифровизации банковского сектора объемы таких данных постоянно растут, что открывает возможности для применения методов машинного обучения. В отличие от классических статистических подходов (например, логистической регрессии), современные алгоритмы позволяют выявлять сложные нелинейные зависимости между характеристиками заемщика и его склонностью к дефолту.

Ключевой особенностью данной предметной области является необходимость балансировки между двумя противоположными рисками. С одной стороны, излишне либеральная политика одобрения кредитов ведет к росту просроченной задолженности и прямым финансовым потерям банка. С другой стороны, чрезмерно строгий подход снижает доступность кредитных продуктов для надежных заемщиков, что ограничивает прибыльность банка и ухудшает клиентский опыт. Таким образом, задача построения точной и интерпретируемой модели одобрения кредитов имеет как экономическое, так и управленческое значение, находясь на стыке финансового анализа, статистики и методов искусственного интеллекта.

В рамках данного исследования предметная область рассматривается через призму анализа исторических данных о заявках на кредиты. На основе набора признаков, характеризующих заемщиков и условия кредитования, решается задача бинарной классификации — прогнозирование целевой переменной `loan_status`, отражающей факт одобрения кредита. Результаты такого анализа могут быть использованы для совершенствования скоринговых систем, выявления ключевых факторов, влияющих на решение, и снижения уровня кредитного риска.

1.2. Объект и предмет исследования

Объект исследования (моделирования) – Процесс кредитования физических лиц в коммерческом банке.

Предмет исследования – Взаимосвязь между характеристиками заемщика (возраст, доход, стаж, кредитная история и др.) и вероятностью одобрения кредита.

1.3. Проблема, решаемая в работе и ее актуальность

Проблемой является высокий уровень кредитного риска из-за недостаточно точной оценки платежеспособности заемщиков традиционными методами, что приводит к росту просроченной задолженности или необоснованным отказам клиентам.

В условиях цифровизации банковского сектора и ужесточения требований регуляторов, внедрение точных моделей прогнозирования дефолта (статуса кредита) позволяет снизить финансовые потери банка, автоматизировать принятие решений и повысить доступность кредитных продуктов для надежных заемщиков. Это является описанием **актуальности** проекта

1.4. Цель работы

Целью работы является решение поставленной проблемы.

Разработать модель прогнозирования статуса одобрения кредита (*loan_status*) на основе анализа социально-демографических, финансовых и кредитно-исторических данных заемщика для минимизации кредитных рисков.

2. METHODS (МЕТОДЫ)

2.1. Обоснование требований к методу решения проблемы

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

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

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

3. Метод должен не теоретически, а реально на практике решать поставленную проблему, а значит, он должен иметь поддерживающий его

программный инструментарий, находящийся в полном открытом бесплатном доступе.

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

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

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

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

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

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

АСК-анализ включает:

- теоретические основы, в частности базовую формализуемую когнитивную концепцию;
- математическую модель, основанную на системном обобщении теории информации (СТИ);
- методику численных расчетов (структуры баз данных и алгоритмы их обработки);
- программный инструментарий, в качестве которого в настоящее время выступает универсальная когнитивная аналитическая система «Эйдос» (интеллектуальная система «Эйдос»).

Более подробно математический метод АСК-анализа описан в работе [3] и ряде других [5]. Около половины из 706 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент

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

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

написания данной работы автором опубликовано 45 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получено 34 патента РФ на системы искусственного интеллекта, 360 публикации в изданиях, входящих в перечень ВАК РФ и приравненных им (по данным [РИНЦ](#)), 17 публикаций в изданиях, входящих в ядро РИНЦ, 6 статей в журналах, входящих в [WoS](#), 7 публикаций в журналах, входящих в [Скопус](#)³ [5, 6, 7].

Три монографии включены в фонды библиотеки конгресса США⁴.

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

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

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

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

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных [1-47];
- Автоматизированный системно-когнитивный анализ текстовых данных [32];
- Спектральный и контурный автоматизированный системно-когнитивный анализ изображений [31];
- Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов [44].

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [5] и страничка в РесечГейт [6] и РИНЦ [7], на которых можно получить более полную информацию о методе АСК-анализа и системе «Эйдос». Краткая

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

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

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

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

информация об АСК-анализе и системе «Эйдос» есть в материале: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf.

На сайте автора размещены тематические подборки публикаций по применению АСК-анализа и системы «Эйдос» в различных предметных областях [26-47].

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

2.4. Система «Эйдос» – инструментарий АСК-анализа

Конечно, на системе «Эйдос» как говорят «Свет клином не сошелся». Существует много очень достойных систем искусственного интеллекта. Чтобы лично убедиться в этом достаточно самостоятельно осуществить поиск в Internet, просто посмотреть файлы: [NCKR-1](#), [NCKR-2](#), [NCKR-3](#), [NCKR-4](#) или пройти по ссылкам: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/how-does-mldotnet-work>, <http://chat.openai.com/>, <https://poe.com/>, <https://neural-university.ru/>, <https://dzen.ru/a/ZCKZRKvrlEMBWO8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://poe.com/Aidos-X>, <https://rudalle.ru/>, <https://bard.google.com/>, <https://chatbot.theb.ai/>, <https://problembo.com/ru/services> (может пригодиться - почта на 10 минут: <https://10minutemail.net/>), <https://poe.com/GPT-3.5-Turbo-Instruct>, <https://www.seaart.ai/home>, <https://ui.chatai.com/>.

И все же Универсальная когнитивная аналитическая система «Эйдос-Х++» отличается от большинства из этих систем, по крайней мере, некоторыми из следующих своих параметров:

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

- является одной из первых и наиболее популярных отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта и программирования: 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 интеллектуальных локальных (т.е. поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 392, соответственно: http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf, http://lc.kubagro.ru/Presentation_LutsenkoEV.pdf);

- находится в полном открытом бесплатном доступе (http://lc.kubagro.ru/aidos/_Aidos-X.htm), причем с актуальными исходными текстами (http://lc.kubagro.ru/_AidosALL.txt): открытая лицензия: [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) (<https://creativecommons.org/licenses/by-sa/4.0/>), и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора и разработчика системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 34 свидетельства РосПатента РФ);

- является «интерпретатором интеллектуальных моделей», т.е. с одной стороны является инструментальной оболочкой, позволяющей без какого-либо программирования создавать интеллектуальные приложения на основе [конфигуратора статистических и системно-когнитивных моделей](#), а с другой стороны является run-time системой или средой исполнения, обеспечивающей эксплуатацию этих интеллектуальных приложений в адаптивном режиме.

- чтобы самостоятельно освоить систему Эйдос достаточно скачать со страницы: <http://lc.kubagro.ru/aidos/index.htm> и установить полную версию систему, а затем в режиме 1.3 скачать и установить из Эйдос-облака одно из интеллектуальных облачных Эйдос-приложений (http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) и выполнять его, следуя описанию приложения. Обычно это файл readme.pdf в папке: c:\Aidos-X\AID_DATA\Inp_data. Для изучения лучше выбирать самые новые приложения, автором которых является проф.Е.В.Луценко. Кроме того на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf есть более 300 полутора-часовых видео-занятий (на русском языке) и много других учебных материалов и примеров описания интеллектуальных-Эйдос-приложений.

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



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

- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение 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, т.е. поддерживать язык OpenGL);

- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач идентификации, прогнозирования, 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 и является инструментом познания: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции, если эти эксперты уже есть, а если их еще нет, то она все равно дает верные результаты познания, что будет признано будущими экспертами, когда они появятся;

- 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 те данные, которые есть, и, тем самым, преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для принятия решений и управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

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

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

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

В разработке системы «Эйдос» были следующие этапы:

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

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 видеограмма 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 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).

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»: 2022 год. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке xBase++eXpress++Advantage Database Server (ADS), обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge).

6-й этап, «эра Больших данных, информации и знаний»: с 2023 года по настоящее время. С 2023 развитие системы «Эйдос» будет осуществляться на языках Питон (Python), а также [Аляска-2.0](#) + [Экспресс++](#).

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

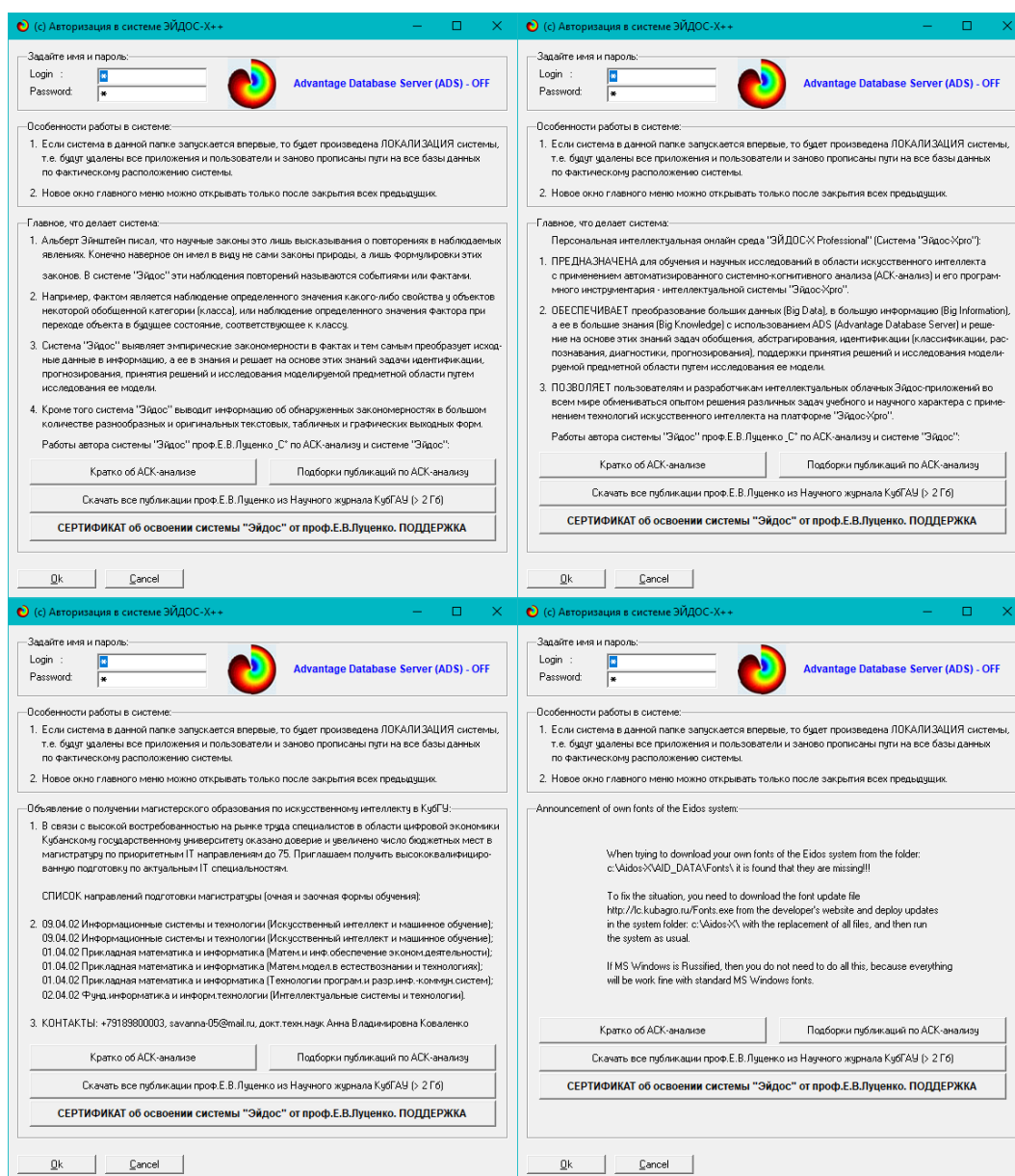
[Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения⁷](#)

На рисунке 1 приведена титульная видеограмма DOS-версии системы «Эйдос», а на рисунке 2 – титульные видеограммы текущей версии системы «Эйдос» (их в настоящее время 7 и они меняются по очереди при каждом запуске системы):

⁷ http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf



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



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

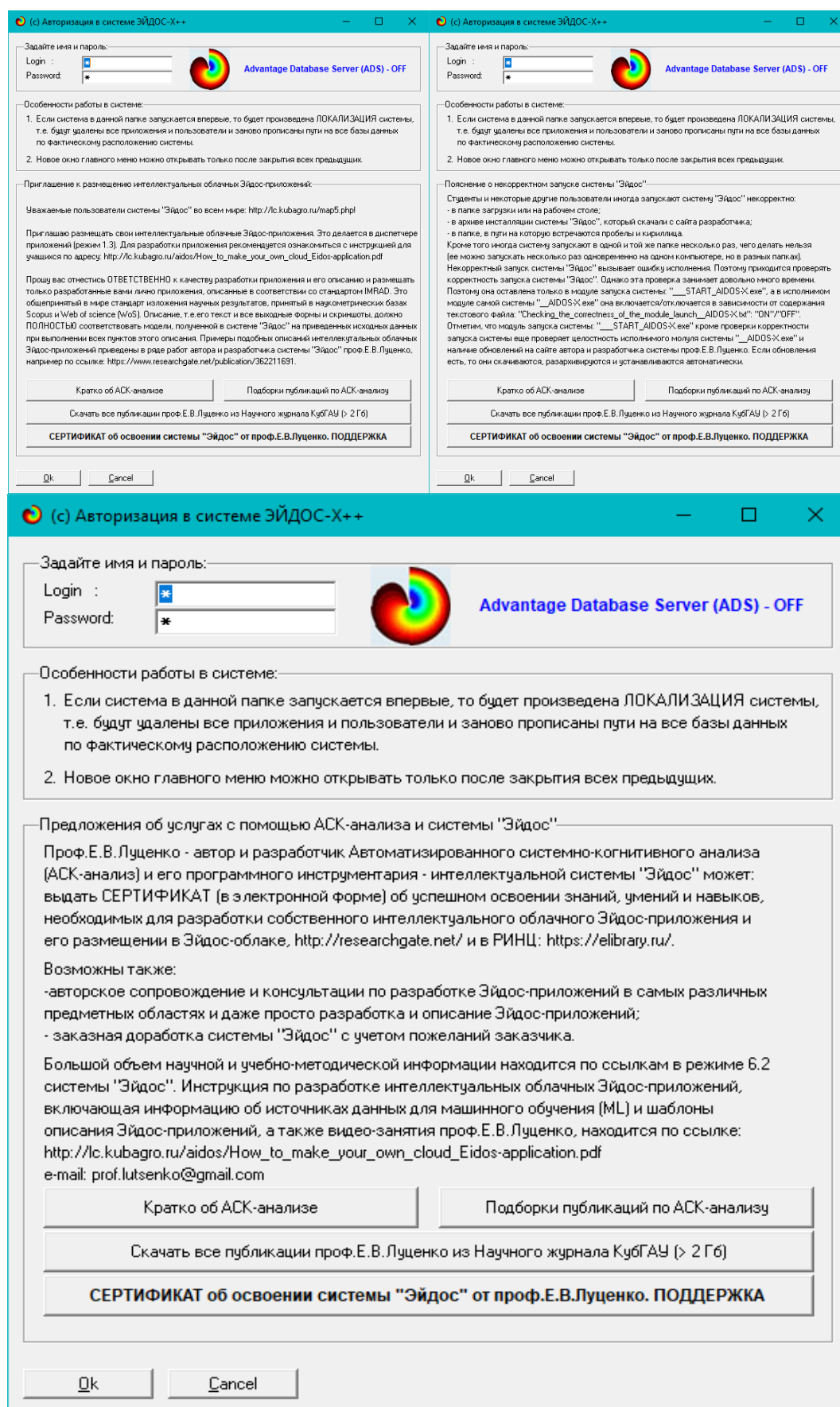


Рисунок 2. Титульные видеогаммы текущей версии системы «Эйдос»

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

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

Задача-1. Когнитивная структуризация предметной области.

Задача-2. Формализация предметной области.

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

Многопараметрическая типизация и частные критерии знаний.

Задача-4. Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

Задача-6. Системная идентификация и прогнозирование.

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

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

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

8.2) кластерно-конструктивный анализ классов;

8.3) кластерно-конструктивный анализ значений описательных шкал;

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

8.5) нелокальная нейронная сеть;

8.6) 3d-интегральные когнитивные карты;

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

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

8.9) когнитивные функции;

8.10) значимость описательных шкал и их градаций;

8.11) степень детерминированности классов и классификационных шкал).

Для данной работы особое значение имеет решение подзадачи 8.1, т.к. она позволяет детально исследовать влияние каждого значения каждого фактора на устойчивость сливы к температурным стрессорам, в т.ч. от обработки деревьев борной кислотой.

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

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

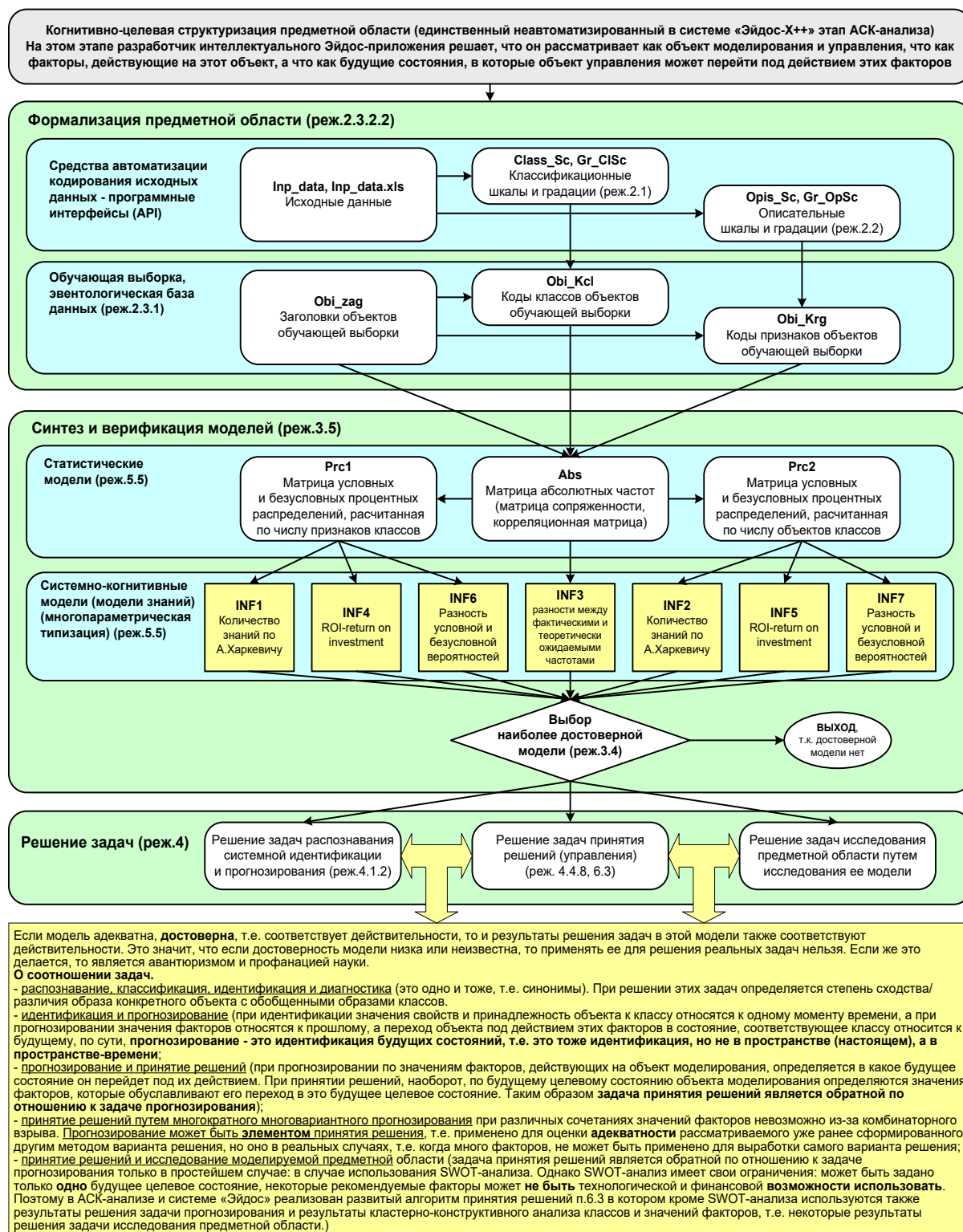


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос» (этапы АСК-анализа)

3. RESULTS (РЕЗУЛЬТАТЫ)

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

3.1.1. Универсальная постановка задачи, не зависящая от предметной области

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.

3.1.2. Конкретное решение задачи в данной работе

В данной работе в качестве **объекта моделирования** выступает процесс принятия решения о выдаче кредита физическому лицу. В качестве **факторов**, влияющих на этот процесс, рассматриваются: возраст заемщика, уровень дохода, трудовой стаж, сумма запрашиваемого кредита, процентная ставка, отношение суммы кредита к доходу, продолжительность кредитной истории, кредитный рейтинг, а также пол заемщика. (Скриншот 1) В качестве **результата** действия этих факторов выступает бинарный показатель — статус одобрения кредита, который отражает, был ли кредит выдан заемщику или отказано в выдаче. (Скриншот 2)

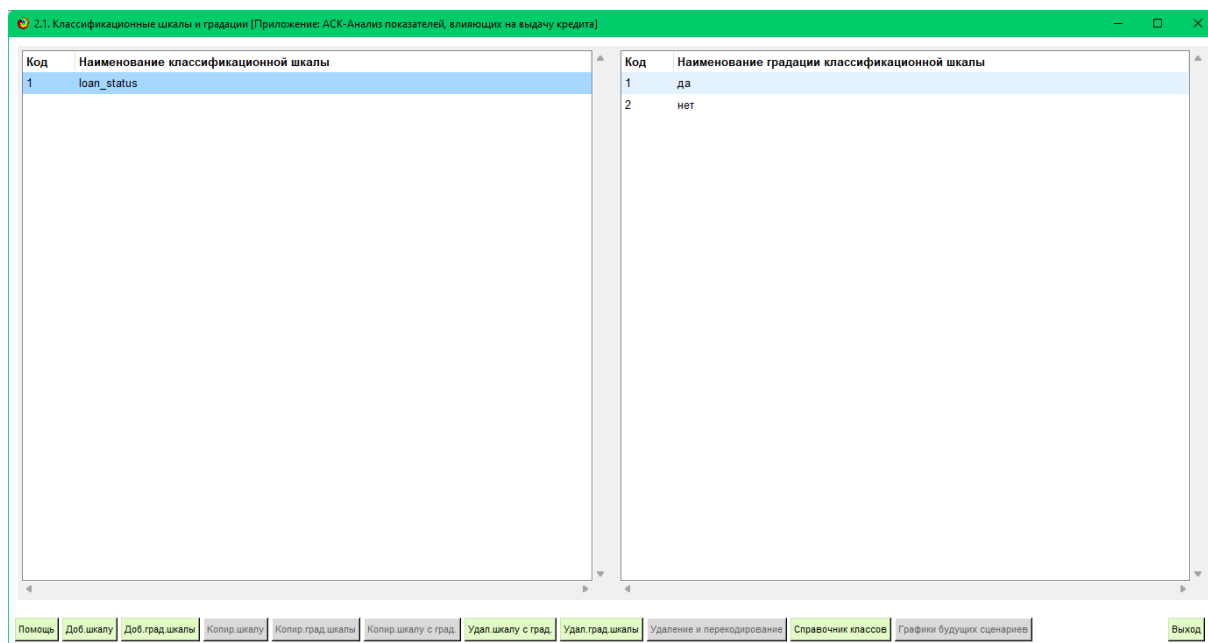
Скриншот 1 – Описательные шкалы (факторы)

Код	Наименование описательной шкалы
1	person_age
2	person_income
3	person_emp_exp
4	loan_amnt
5	loan_int_rate
6	loan_percent_income
7	cb_person_cred_hist_length
8	credit_score
9	person_gender_male

Код	Наименование градации описательной шкалы
1	1/5-[20.0000000, 44.8000000]
2	2/5-[44.8000000, 69.6000000]
3	3/5-[69.6000000, 94.4000000]
4	4/5-[94.4000000, 119.2000000]
5	5/5-[119.2000000, 144.0000000]

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Opis_Sc.xlsx

Скриншот 2 – Классификационные шкалы (результаты действия факторов)



Источник: c:\Aidos-X\AID_DATA\A0000001\System\Class_Sc.xlsx

Для формирования.xlsx-файлов, приведенных в таблицах 1 и 2, необходимо выполнить в системе "Эйдос" режим 5.12.

3.2. Задача-2. Формализация предметной области

3.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

Этим обеспечивается возможность комфортного для пользователя применения системы «Эйдос» для проведения научных исследований и решения практических задач в самых различных предметных областях и направлениях науки, практически везде, где человек применяет естественный интеллект [48, 49].

3.2.2. Конкретное решение задачи в данной работе

В качестве *источника исходных данных* в данной работе используем найденную нами таблицу данных [1]. Таблица найдена на сайте Kaggle.com.(Таблица 1)

Таблица 1 – Исходные данные по влиянию различных факторов на получения одобрения кредита в банке

id	person_age	person_income	person_emp_exp	loan_amnt	loan_int_rate	loan_percent_income	cb_person_cred_hist_length	credit_score	person_gender_male	loan_status
1	22	71948	0	35000	16,02	0,49	3	561	нет	да
2	21	12282	0	1000	11,14	0,08	2	504	нет	нет
3	25	12438	3	5500	12,87	0,44	3	635	нет	да
4	23	79753	0	35000	15,23	0,44	2	675	нет	да
5	24	66135	1	35000	14,27	0,53	4	586	да	да
6	21	12951	0	2500	7,14	0,19	2	532	нет	да
7	26	93471	1	35000	12,42	0,37	3	701	нет	да
8	24	95550	5	35000	11,11	0,37	4	585	нет	да
9	24	100684	3	35000	8,9	0,35	2	544	нет	да
10	21	12739	0	1600	14,74	0,13	3	640	нет	да
11	22	102985	0	35000	10,37	0,34	4	621	нет	да
12	21	13113	0	4500	8,63	0,34	2	651	нет	да
13	23	114860	3	35000	7,9	0,3	2	573	да	да
14	26	130713	0	35000	18,39	0,27	4	708	да	да
15	23	138998	0	35000	7,9	0,25	4	583	нет	нет
16	23	600891	5	30000	10,65	0,05	3	670	нет	нет
17	23	144943	0	35000	7,9	0,24	4	663	да	нет
18	23	111369	0	35000	20	0,31	4	694	нет	да
19	23	136628	0	35000	18,25	0,26	4	709	да	да
20	24	14283	1	1750	10,99	0,12	2	679	нет	да
21	25	195718	0	35000	7,49	0,18	4	684	да	нет
22	25	165792	4	34800	16,77	0,21	2	662	да	нет
23	22	79255	0	34000	17,58	0,43	4	691	нет	да
24	24	13866	0	1500	7,29	0,11	3	600	нет	нет
25	22	97420	1	33950	14,54	0,35	4	691	да	да

Источник: [25]: <https://www.kaggle.com/datasets/mytalkwithyou/bank-loan-approval-dataset>

Таблица 4 имеет следующую структуру:

- каждая строка описывает одно наблюдение с определенным сочетанием значений факторов и определенной жизнеспособностью пыльцы сливы, всего 10 наблюдений;

- каждое **наблюдение** описывается одновременно **двумя** способами: с одной стороны значениями факторов, действующих на объект моделирования (лингвистические и числовые переменные, градации описательных шкал, бесцветный фон в таблице 1), а с другой стороны результатами действия этих факторов, т.е. одобрен кредит или нет, выраженными в текстовых и числовых шкалах в разных единицах измерения (желтый фон). Такая структура описания наблюдений в технологиях искусственного интеллекта называется **«онтологией»** и модели представлений знаний Марвина Мински (1975) называется **«фрейм-экземпляр»**;

- 1-я колонка – не является шкалой и содержит номер наблюдения или другую идентифицирующую информацию о том, откуда взято описание этого наблюдения;

- колонки со 11-й по 11-ю – это классификационные шкалы – это шкалы **текстового** и **числового** типа описывающие **результаты** действия факторов в различных единицах измерения (таблица 1), в данном случае, одобрен кредит или нет. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в **натуральном** и **стоимостном** выражении [8]: например **количество** и **качество** продукции, **прибыль** и **рентабельность**. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

- колонки с 2-й по 10-ю – это описательные шкалы, формализующие факторы, действующие на объект моделирования (таблица 1). Эти шкалы имеют числовой и текстовый тип и их градациями являются лингвистические и числовые переменные;

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

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

Таким образом, система «Эйдос» не предъявляет чрезмерно жестких практически невыполнимых требований к исходным данным, а

обрабатывает те данные, которые реально есть, например подобные представленным в таблице 1.

В системе «Эйдос» есть 6 основных автоматизированных программных интерфейсов (API), обеспечивающих ввод в систему и интеллектуальную обработку числовых, текстовых и графических данных, представленных в виде таблиц и файлов. Возможна обработка и других видов данных (например, данных землетрясений, ЭЭГ, ЭКГ, аудио и видео), которые возможно представить в этих форматах (рисунок 4).

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

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

Для ввода исходных данных, представленных в таблице 1, в систему «Эйдос», используется один из ее автоматизированных программных интерфейсов (API), а именно универсальный автоматизированный программный интерфейс ввода данных из файлов MS Excel (API-2.3.2.2).

Требования API-2.3.2.2 к исходным данным подробно описаны в хелпах этого режима (рисунки 5):

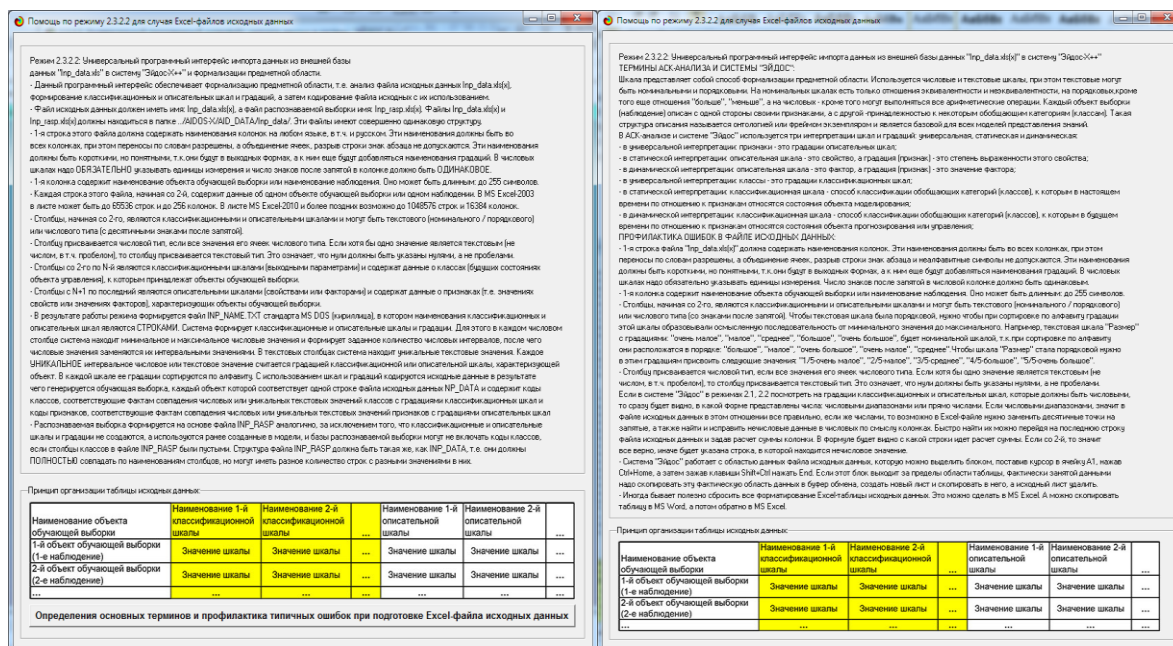


Рисунок 5. Хелпы API-2.3.2.2 системы «Эйдос»

Экранные формы управления API-2.3.2.2 системы «Эйдос» с реальными параметрами, использованными в данной работе, приведены на рисунках 6:

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

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

Задайте тип файла исходных данных: "Inp_data":

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

Задайте параметры:

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

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец: 11 Конечный столбец: 11

Задайте диапазон столбцов описательных шкал:

Начальный столбец: 2 Конечный столбец: 10

Задайте режим:

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

Задайте способ выбора размера интервалов:

☒ Равные интервалы с разным числом наблюдений
☐ Разные интервалы с равным числом наблюдений

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

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

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:
 Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:
 Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

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

Старт Выход

2.3.2.2. Задание размерности модели системы "ЭЙДОС-X++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (адаптивные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [3 классов x 8 признаков]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	1	3	3,00	0	0	0,00
Текстовые	0	0	0,00	4	8	2,00
ВСЕГО:	1	3	3,00	4	8	2,00

Задайте количество числовых диапазонов (интервалов, градаций) в шкале:

В классификационных шкалах: 3

Пересчитать шкалы и градации Параметры числ.шкал и градаций Выйти на создание модели

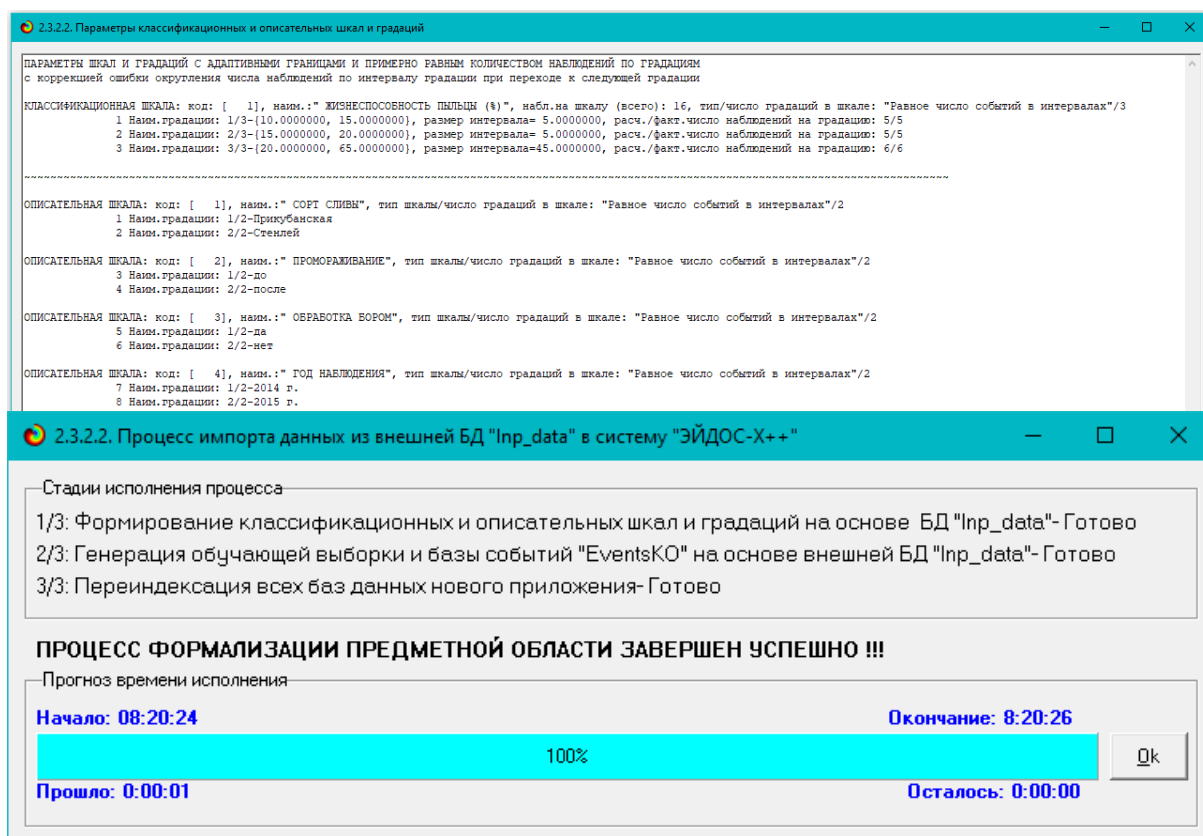


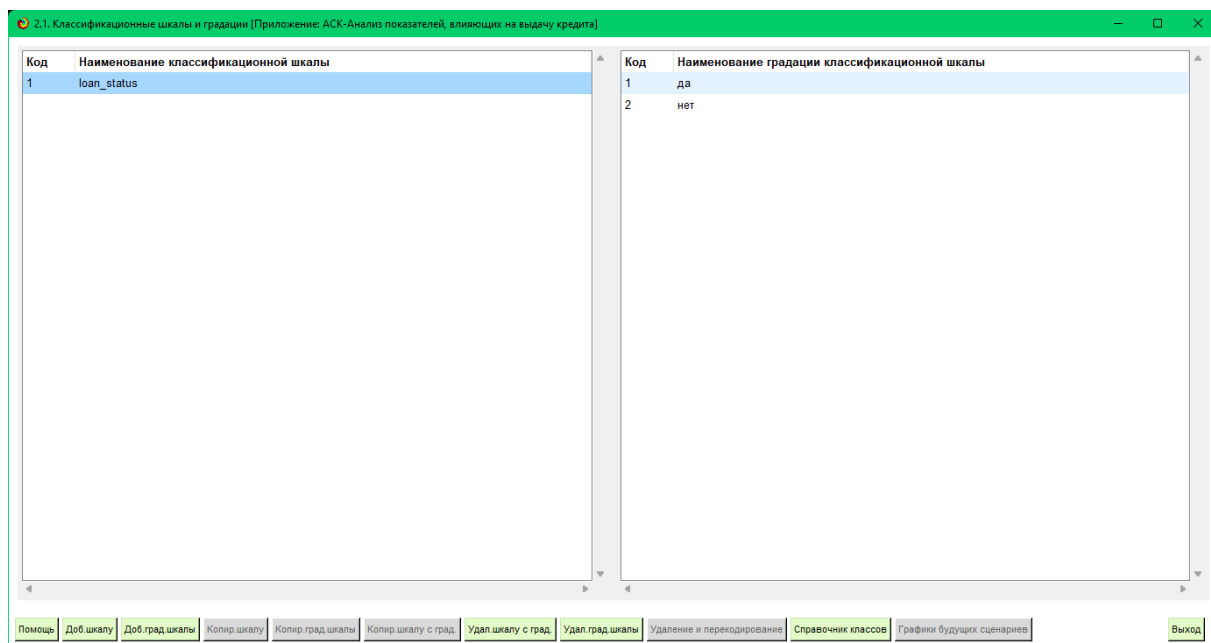
Рисунок 6. Экранные формы управления API-2.3.2.2 системы «Эйдос»

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

Отметим, что суммарное количество градаций на 2-м рисунке 6 и в таблице 6 может не совпадать, если в некоторых описательных шкалах есть градации «Пробел» или нули, которые в соответствии с параметрами на 1-м рисунке 6 рассматриваются не как значения, а как *отсутствие данных*.

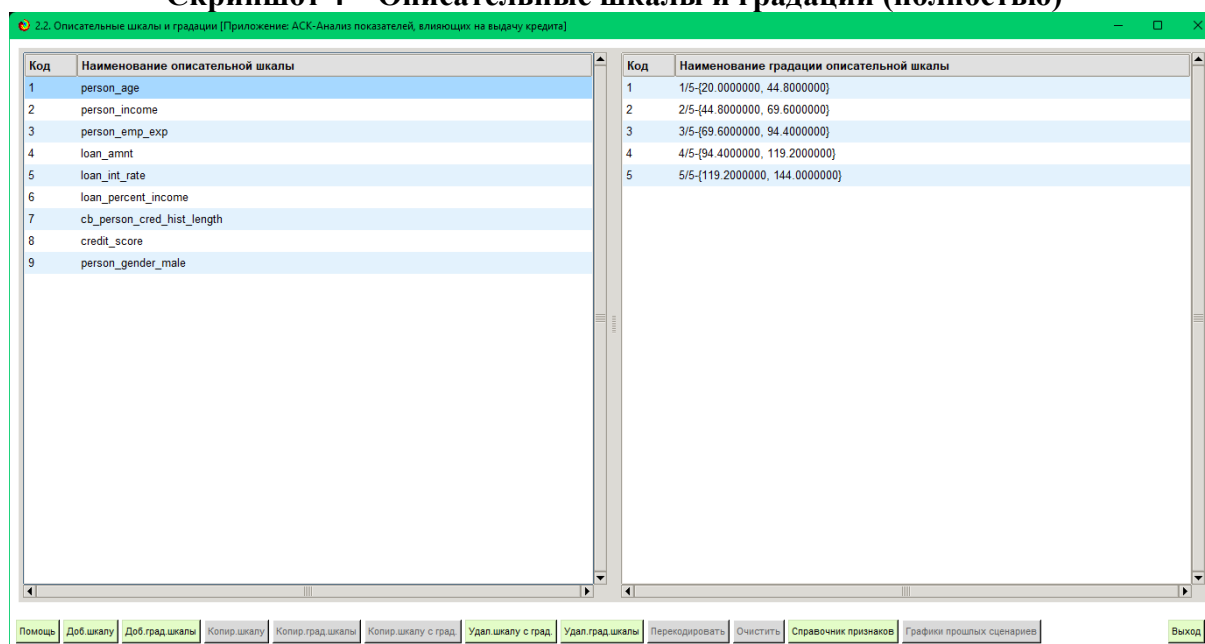
Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал и между шкалами, как числовыми, так и текстовыми. Математическая модель АСК-анализа позволяет корректно преодолеть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

Скриншот 3 – Классификационные шкалы и градации (полностью)



Источник: c:\Aidos-X\AID_DATA\A0000001\System\Classes.xlsx

Скриншот 4 – Описательные шкалы и градации (полностью)



Источник: c:\Aidos-X\AID_DATA\A0000001\System\Attributes.xlsx

Скриншот 5 – Обучающая выборка (полностью)

2.3.1. Ручной ввод-корректировка обучающей выборки [Приложение: АСК-Анализ показателей, влияющих на выдачу кредита]

Код объекта	Наименование объекта обучающей выборки	Дата	Время
1	1	17.03.2026	16:24:21
2	2	17.03.2026	16:24:21
3	3	17.03.2026	16:24:21
4	4	17.03.2026	16:24:21
5	5	17.03.2026	16:24:21
6	6	17.03.2026	16:24:21
7	7	17.03.2026	16:24:21
8	8	17.03.2026	16:24:21
9	9	17.03.2026	16:24:21
10	10	17.03.2026	16:24:21
11	11	17.03.2026	16:24:21
12	12	17.03.2026	16:24:21

Коды классов (градаций классификационных шкал)				Коды признаков (градаций описательных шкал)						
1	0	0	0	1	6	0	16	22	28	31
				38	42	0	0	0	0	0

Шкала: Градация:

Помощь Склонировать обуч. выб в расп. Добавить объект Добавить классы Добавить признаки Удалить объект Удалить классы Удалить признаки Очистить БД Выход

Источник: c:\Aidos-X\AID_DATA\A0000001\System\EventsKO.xlsx

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

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

3.3.1. Универсальная постановка задачи, не зависящая от предметной области

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

Отметим, что модели системы «Эйдос» основаны на матрице абсолютных частот, отражающей число встреч градаций описательных

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

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

Суть математической модели АСК-анализа состоит в следующем.

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

Таблица 3 – Матрица абсолютных частот (статистическая модель ABS)

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

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

Таблица 4 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

		Классы					Безусловная вероятность признака
		1	...	j	...	w	
Значения факторов	1	P_{11}		P_{1j}		P_{1w}	
	...						
	i	P_{i1}		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		P_{iw}	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	m	P_{m1}		P_{mj}		P_{mw}	
Безусловная вероятность класса				$P_{\Sigma j}$			

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

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

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

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

Этот переход полностью снимает проблему **несбалансированности** данных, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 8), а матрицы условных и безусловных процентных распределений (таблица 9), а также матрицы системно-когнитивных моделей, рассчитываемые на основе матрица абсолютных частот и матрицы условных и безусловных процентных распределений. Этот

подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (дихотомических, номинальных, порядковых и числовых) и в разных единицах измерения [8]. В системе «Эйдос» этот подход применяется всегда при решении любых задач.

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

В таблице 10 приведены формулы:

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

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

Таблица 5– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

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

INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij}N - N_iN_j}{N_jN}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		

Обозначения к таблице:

i – значение прошлого параметра;

j – значение будущего параметра;

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

M – суммарное число значений всех прошлых параметров;

W – суммарное число значений всех будущих параметров.

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

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

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

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

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

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

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

Таблица 6 – Матрица системно-когнитивной модели

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

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

осуществляется либо логарифмированием, либо вычитанием единицы (таблица 12).

Таблица 7– Конфигуратор системно-когнитивных моделей АСК-анализа и интеллектуальной системы «Эйдос»

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

Обратим особое внимание на то, что сравнение фактических и теоретических абсолютных частот путем деления приводит при нормировках к нулю (что нужно для применения аддитивных интегральных критериев) путем взятия логарифма и путем вычитания 1 к *тем же самым* моделям, что и сравнение условных и безусловных относительных частот путем деления с теми же самыми способами нормировки. Таким образом, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и провести нормировку к 0 путем взятия логарифма и путем вычитания 1, то получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Других же системно-когнитивных моделей, рассчитываемых на основе приведенных статистических моделей просто нет. Это и есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра. *Под конфигуратором В.А.Лефевр понимал минимальный полный набор понятийных шкал или конструктов, т.е. понятий, достаточный для адекватного описания предметной области* [4]⁹. Необходимо отметить, что все эти модели рассчитываются в интеллектуальной системе «Эйдос».

Когда мы сравниваем фактические и теоретические абсолютные частоты путем вычитания у нас получается частный критерий знаний: «хи-квадрат» (СК-модель INF3), когда же мы сравниваем их путем деления, то

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

у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2) или «коэффициент возврата инвестиций ROI» - Return On Investment (СК-модели INF4, INF5) в зависимости от способа нормировки.

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

Таким образом, мы видим, что все частные критерии знаний тесно взаимосвязаны друг с другом. Особенно интересна связь знаменитого критерия хи-квадрат К.Пирсона с замечательной мерой количества информации А.Харкевича и с известным в экономике коэффициентом ROI.

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

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

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

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

Таблица 8 – Уточнение терминологии АСК-анализа

№	Традиционные термины	Новый	Формула
---	----------------------	-------	---------

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

Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 13):

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

моделей в смысле В.А.Левеева, содержащий минимальное количество моделей, позволяющих полно описать моделируемую предметную область.

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

3.3.2. Конкретное решение задачи в данной работе

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

3.5. Синтез и верификация моделей

Задайте модели для синтеза и верификации

Статистические базы:

- ☒ 1. ABS - частный критерий: количество встреч сочетаний "класс-признак" у объектов обучающей выборки

Задайте источник данных для расчета модели ABS:

☒ Обучающая выборка ☐ Abs ☐ Prc1 ☐ Prc2 ☐ Inf1 ☐ Inf2 ☐ Inf3 ☐ Inf4 ☐ Inf5 ☐ Inf6 ☐ Inf7

Задайте значение фона в матрице абсолютных частот: [Помощь](#)

- ☒ 2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов i-го класса
- ☒ 3. PRC2 - частный критерий: условная вероятность i-го признака у объектов i-го класса

Системно-когнитивные модели (базы знаний):

- ☒ 4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1
- ☒ 5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2
- ☒ 6. INF3 - частный критерий: χ^2 -квадрат, разности между фактическими и ожидаемыми абс. частотами
- ☒ 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1
- ☒ 8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
- ☒ 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1
- ☒ 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

Параметры копирования обучающей выборки в распознаваемую (бутстрепный подход):

Какие объекты обучающей выборки копировать:

- ☒ Копировать всю обучающую выборку
- ☐ Копировать только текущий объект
- ☐ Копировать каждый N-й объект
- ☐ Копировать N случайных объектов
- ☐ Копировать объекты от N1 до N2 (fastest)
- ☐ Вообще не менять распознаваемую выборку

Удалять из обучающей выборки скопированные объекты:

- ☒ Не удалять
- ☐ Удалять

Пояснение по алгоритму верификации

[Подробнее](#)

Измеряется внутренняя достоверн. модели

Для каждой заданной модели выполнить:

- ☒ Синтез и верификацию
- ☐ Только верификацию
- ☐ Только синтез

Задайте процессор:

- ☒ CPU ☐ GPU

Задайте алгоритм:

- ☐ Классика - дольше
- ☒ Упрощенно-быстрее

Использование только наиболее достоверных результатов распознавания: Rasp.dbf и целесообразность применения бутстрепного подхода

Расчетный размер БД результатов распознавания Rasp.dbf равен 2594 байт, т.е.: 0.0001208 % от MAX-возможного, (от 2Gb)

УЧИТЫВАТЬ только наиболее достоверные результаты распознавания с МОДУЛЕМ инт. крит. "Резонанс знаний" выше %

В применении бутстрепного подхода нет необходимости. Синтез и верификация моделей будут выполнены на основе всей выборки.

[Ok](#) [Cancel](#)

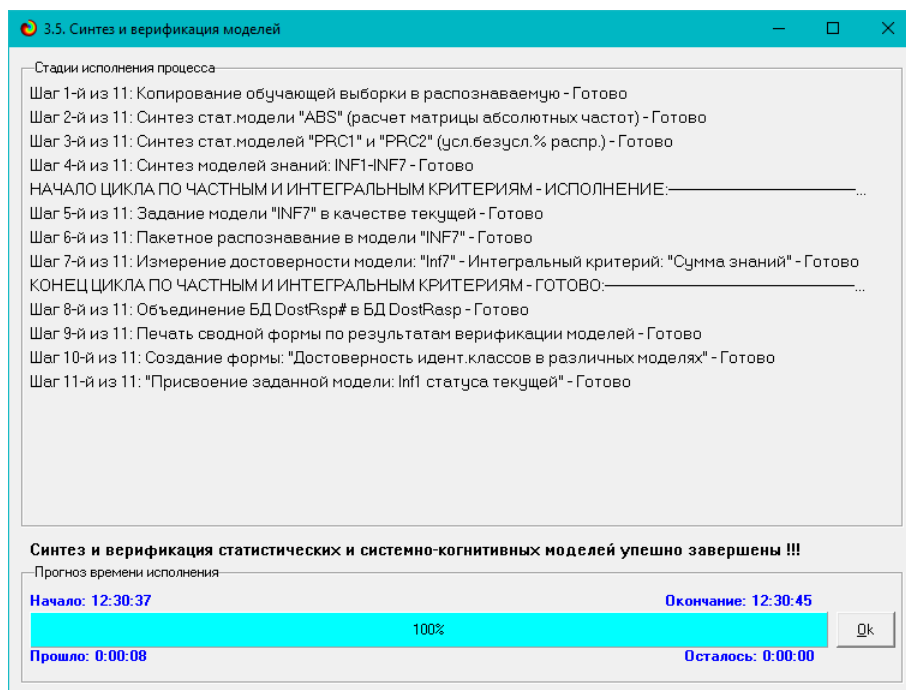


Рисунок 7. Экранные формы режима синтеза и верификации моделей

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

3.5. Просмотр основных баз данных всех моделей

Задайте модель для просмотра: Abs PRC1 PRC2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Количество знаков после запятой: 3 (макс: 17) Помощь

Модель: "Abs"

Код признака	Описательные шкалы и градации (признаки)	1 - loan_amnt - abs	2 - loan_amnt - int	Сумма	Среднее	Средн. макс. отклонение	Средней максим. отклонение от среднего
1	person_age - 1/5-[20 0000000, 44 8000000]	2662 000	8021 000	10683 000	5341 500	3789 385	2679 500
2	person_age - 2/5-[44 8000000, 69 6000000]						
3	person_age - 3/5-[69 6000000, 94 4000000]						
4	person_age - 4/5-[94 4000000, 119 2000000]						
5	person_age - 5/5-[119 2000000, 144 0000000]		4 000	4 000	2 000	2 828	2 000
6	person_income - 1/5-[12231 0000000, 129963 0000000]	2618 000	7680 000	10298 000	5149 000	3579 375	2531 000
7	person_income - 2/5-[129963 0000000, 247695 0000000]	39 000	302 000	341 000	170 500	185 969	131 500
8	person_income - 3/5-[247695 0000000, 365427 0000000]	5 000	41 000	46 000	23 000	25 456	18 800
9	person_income - 4/5-[365427 0000000, 483159 0000000]	1 000	1 000	0 500	0 707	0 500	
10	person_income - 5/5-[483159 0000000, 600891 0000000]	1 000	1 000	0 500	0 707	0 500	
11	person_emp_exp - 1/5-[1 8000000, 25 8000000]	1511 000	4714 000	6225 000	3112 500	2264 863	1601 500
12	person_emp_exp - 2/5-[25 8000000, 50 6000000]						
13	person_emp_exp - 3/5-[50 6000000, 75 4000000]						
14	person_emp_exp - 4/5-[75 4000000, 100 2000000]	1 000	1 000	0 500	0 707	0 500	
15	person_emp_exp - 5/5-[100 2000000, 125 0000000]	3 000	3 000	1 500	2 121	1 500	
16	loan_amnt - 1/5-[500 0000000, 7400 0000000]	1261 000	4737 000	5998 000	2999 000	2457 903	1738 000
17	loan_amnt - 2/5-[7400 0000000, 14300 0000000]	752 000	2428 000	3180 000	1590 000	1185 111	838 000
18	loan_amnt - 3/5-[14300 0000000, 21200 0000000]	410 000	681 000	1091 000	545 500	191 626	135 500
19	loan_amnt - 4/5-[21200 0000000, 28100 0000000]	207 000	156 000	363 000	181 500	36 062	25 500
20	loan_amnt - 5/5-[28100 0000000, 35000 0000000]	32 000	23 000	55 000	27 500	6 364	4 500
21	loan_int_rate - 1/5-[5 4200000, 8 3360000]	302 000	2364 000	2666 000	1333 000	1458 054	1031 000
22	loan_int_rate - 2/5-[8 3360000, 11 2520000]	696 000	2813 000	3509 000	1754 500	1496 945	1058 500
23	loan_int_rate - 3/5-[11 2520000, 14 1680000]	676 000	2091 000	2767 000	1383 500	1000 556	707 500

Экспорт в CSV Экспорт всех моделей в MS Excel Выход

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

5.5. Просмотр основных баз данных всех моделей

Задайте модель для просмотра: Abs Prc1 **Prc2** Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Количество знаков после запятой: 3 (макс: 17) Помощь

Модель: "Prc2"

Код признака	Описательные шкалы и градации (признаки)	1 - loan_status - да	2 - loan_status - нет	Безусловная вероятность	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	person_age - 1/5-{20.0000000, 44.8000000}	100.000	99.950	99.963	99.975	0.035	0.025
2	person_age - 2/5-{44.8000000, 69.6000000}						
3	person_age - 3/5-{69.6000000, 94.4000000}						
4	person_age - 4/5-{94.4000000, 119.2000000}						
5	person_age - 5/5-{119.2000000, 144.0000000}		0.050	0.037	0.025	0.035	0.025
6	person_income - 1/5-{12231.0000000, 129963.0000000}	98.347	95.701	96.360	97.024	1.871	1.323
7	person_income - 2/5-{129963.0000000, 247695.0000000}	1.465	3.763	3.191	2.614	1.625	1.149
8	person_income - 3/5-{247695.0000000, 365427.0000000}	0.188	0.511	0.430	0.349	0.228	0.162
9	person_income - 4/5-{365427.0000000, 483159.0000000}		0.012	0.009	0.006	0.009	0.006
10	person_income - 5/5-{483159.0000000, 600891.0000000}		0.012	0.009	0.006	0.009	0.006
11	person_emp_exp - 1/5-{1.0000000, 25.8000000}	56.762	58.741	58.248	57.752	1.400	0.990
12	person_emp_exp - 2/5-{25.8000000, 50.6000000}						
13	person_emp_exp - 3/5-{50.6000000, 75.4000000}						
14	person_emp_exp - 4/5-{75.4000000, 100.2000000}		0.012	0.009	0.006	0.009	0.006
15	person_emp_exp - 5/5-{100.2000000, 125.0000000}		0.037	0.028	0.019	0.026	0.019
16	loan_amnt - 1/5-{500.0000000, 7400.0000000}	47.370	59.028	56.124	53.199	8.243	5.829
17	loan_amnt - 2/5-{7400.0000000, 14300.0000000}	28.249	30.255	29.756	29.252	1.418	1.003
18	loan_amnt - 3/5-{14300.0000000, 21200.0000000}	15.402	8.486	10.209	11.944	4.890	3.458
19	loan_amnt - 4/5-{21200.0000000, 28100.0000000}	7.776	1.944	3.397	4.860	4.124	2.916
20	loan_amnt - 5/5-{28100.0000000, 35000.0000000}	1.202	0.287	0.515	0.744	0.647	0.458
21	loan_int_rate - 1/5-{5.4200000, 8.3360000}	11.345	29.458	24.946	20.401	12.808	9.057
22	loan_int_rate - 2/5-{8.3360000, 11.2520000}	26.146	35.053	32.834	30.599	6.298	4.454
23	loan_int_rate - 3/5-{11.2520000, 14.1680000}	25.394	26.056	25.891	25.725	0.468	0.331
24	loan_int_rate - 4/5-{14.1680000, 17.0840000}	30.654	8.561	14.064	19.607	15.622	11.046

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

5.5. Просмотр основных баз данных всех моделей

Задайте модель для просмотра: Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Количество знаков после запятой: 3 (макс: 17) Помощь

Модель: "Prc2"

Код признака	Описательные шкалы и градации (признаки)	1 - loan_status - да	2 - loan_status - нет	Безусловная вероятность	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	person_age - 1/5-{20.0000000, 44.8000000}	100.000	99.950	99.963	99.975	0.035	0.025
2	person_age - 2/5-{44.8000000, 69.6000000}						
3	person_age - 3/5-{69.6000000, 94.4000000}						
4	person_age - 4/5-{94.4000000, 119.2000000}						
5	person_age - 5/5-{119.2000000, 144.0000000}		0.050	0.037	0.025	0.035	0.025
6	person_income - 1/5-{12231.0000000, 129963.0000000}	98.347	95.701	96.360	97.024	1.871	1.323
7	person_income - 2/5-{129963.0000000, 247695.0000000}	1.465	3.763	3.191	2.614	1.625	1.149
8	person_income - 3/5-{247695.0000000, 365427.0000000}	0.188	0.511	0.430	0.349	0.228	0.162
9	person_income - 4/5-{365427.0000000, 483159.0000000}		0.012	0.009	0.006	0.009	0.006
10	person_income - 5/5-{483159.0000000, 600891.0000000}		0.012	0.009	0.006	0.009	0.006
11	person_emp_exp - 1/5-{1.0000000, 25.8000000}	56.762	58.741	58.248	57.752	1.400	0.990
12	person_emp_exp - 2/5-{25.8000000, 50.6000000}						
13	person_emp_exp - 3/5-{50.6000000, 75.4000000}						
14	person_emp_exp - 4/5-{75.4000000, 100.2000000}		0.012	0.009	0.006	0.009	0.006
15	person_emp_exp - 5/5-{100.2000000, 125.0000000}		0.037	0.028	0.019	0.026	0.019
16	loan_amnt - 1/5-{500.0000000, 7400.0000000}	47.370	59.028	56.124	53.199	8.243	5.829
17	loan_amnt - 2/5-{7400.0000000, 14300.0000000}	28.249	30.255	29.756	29.252	1.418	1.003
18	loan_amnt - 3/5-{14300.0000000, 21200.0000000}	15.402	8.486	10.209	11.944	4.890	3.458
19	loan_amnt - 4/5-{21200.0000000, 28100.0000000}	7.776	1.944	3.397	4.860	4.124	2.916
20	loan_amnt - 5/5-{28100.0000000, 35000.0000000}	1.202	0.287	0.515	0.744	0.647	0.458
21	loan_int_rate - 1/5-{5.4200000, 8.3360000}	11.345	29.458	24.946	20.401	12.808	9.057
22	loan_int_rate - 2/5-{8.3360000, 11.2520000}	26.146	35.053	32.834	30.599	6.298	4.454
23	loan_int_rate - 3/5-{11.2520000, 14.1680000}	25.394	26.056	25.891	25.725	0.468	0.331
24	loan_int_rate - 4/5-{14.1680000, 17.0840000}	30.654	8.561	14.064	19.607	15.622	11.046

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

5.5. Просмотр основных баз данных всех моделей

Задайте модель для просмотра: Abs Prc1 Prc2 Inf1 Inf2 **Inf3** Inf4 Inf5 Inf6 Inf7 Количество знаков после запятой: 3 (макс: 17) Помощь

Модель: "Inf3"

Код признака	Описательные шкалы и градации (признаки)	1 - loan_status - да	2 - loan_status - нет	Сумма	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	person_age - 1/5-{20.0000000, 44.8000000}	5.692	-5.692			8.050	5.692
2	person_age - 2/5-{44.8000000, 69.6000000}						
3	person_age - 3/5-{69.6000000, 94.4000000}						
4	person_age - 4/5-{94.4000000, 119.2000000}						
5	person_age - 5/5-{119.2000000, 144.0000000}	-0.995	0.995	0.000	0.000	1.407	0.995
6	person_income - 1/5-{12231.0000000, 129963.0000000}	57.422	-57.422			81.207	57.422
7	person_income - 2/5-{129963.0000000, 247695.0000000}	-45.789	45.789	0.000	0.000	64.755	45.789
8	person_income - 3/5-{247695.0000000, 365427.0000000}	-6.438	6.438	0.000	0.000	9.104	6.438
9	person_income - 4/5-{365427.0000000, 483159.0000000}	-0.249	0.249	0.000	0.000	0.352	0.249
10	person_income - 5/5-{483159.0000000, 600891.0000000}	-0.249	0.249	0.000	0.000	0.352	0.249
11	person_emp_exp - 1/5-{1.0000000, 25.8000000}	-36.835	36.835			52.092	36.835
12	person_emp_exp - 2/5-{25.8000000, 50.6000000}						
13	person_emp_exp - 3/5-{50.6000000, 75.4000000}						
14	person_emp_exp - 4/5-{75.4000000, 100.2000000}	-0.249	0.249	0.000	0.000	0.352	0.249
15	person_emp_exp - 5/5-{100.2000000, 125.0000000}	-0.746	0.746	-0.000	-0.000	1.055	0.746
16	loan_amnt - 1/5-{500.0000000, 7400.0000000}	-230.391	230.391	0.000	0.000	325.823	230.391
17	loan_amnt - 2/5-{7400.0000000, 14300.0000000}	-38.701	38.701	-0.000	-0.000	54.732	38.701
18	loan_amnt - 3/5-{14300.0000000, 21200.0000000}	138.725	-138.725			196.187	138.725
19	loan_amnt - 4/5-{21200.0000000, 28100.0000000}	116.741	-116.741			165.096	116.741
20	loan_amnt - 5/5-{28100.0000000, 35000.0000000}	18.324	-18.324			25.915	18.324
21	loan_int_rate - 1/5-{5.4200000, 8.3360000}	-360.896	360.896			510.384	360.896
22	loan_int_rate - 2/5-{8.3360000, 11.2520000}	-176.506	176.506			249.618	176.506
23	loan_int_rate - 3/5-{11.2520000, 14.1680000}	-12.009	12.009	-0.000	-0.000	16.984	12.009
24	loan_int_rate - 4/5-{14.1680000, 17.0840000}	442.282	-442.282	-0.000	-0.000	625.481	442.282

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

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

3.4. Задача-4. Верификация моделей

3.4.1. Универсальная постановка задачи, не зависящая от предметной области

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных и ложных положительных и отрицательных решений по F-мере Ван Ризбергера, а также по критериям L1- L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры [9].

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

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

3.4.2. Конкретное решение задачи в данной работе

В соответствии с критерием достоверности L1-мерой проф.Е.В.Луценко наиболее достоверной является СК-модель INF3 (хи-квадрат К.Пирсона) с интегральным критерием: «Сумма знаний»: **L1=0.952** при максимуме 1 (рисунок 12). *Эту модель и будем использовать для решения поставленных в работе задач.*

3.4. Обобщенная форма для расчета крит. (Примечание: СК-модели показаны, выделены на выдел. крит.)

№	Наименование модели и критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-полож. решений (TP)	Число истинно-отриц. решений (TN)	Число ложно-полож. решений (FP)	Число ложно-отриц. решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модулей уровней складств истинно-положит. решений (STP)	Сумма модулей уровней складств истинно-отрицат. решений (STN)
1	ABS - частный критерий: количество встреч сочетаний "класс-признак" у объектов обуч. выборки	корреляция abs частот с обр. объекта	10687	10687	0	10687	0	0.500	1.000	0.667	8803.629	0.000
2	ABS - частный критерий: количество встреч сочетаний "класс-признак" у объектов обуч. выборки	сумма abs частот по признакам объекта	10687	10687	0	10687	0	0.500	1.000	0.667	7596.865	0.000
3	PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса	корреляция усл. отн. частот с обр. объекта	10687	10687	0	10687	0	0.500	1.000	0.667	8803.629	0.000
4	PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса	сумма усл. отн. частот по признакам объекта	10687	10687	0	10687	0	0.500	1.000	0.667	9026.198	0.000
5	PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса	корреляция усл. отн. частот с обр. объекта	10687	10687	0	10687	0	0.500	1.000	0.667	8803.629	0.000
6	PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса	сумма усл. отн. частот по признакам объекта	10687	10687	0	10687	0	0.500	1.000	0.667	9021.181	0.000
7	INF1 - частный критерий: количество знаний по А.Харавану; вероятности из PRC1	семантический резонанс знаний	10687	8604	8448	2239	2083	0.794	0.805	0.799	2758.764	3481.361
8	INF1 - частный критерий: количество знаний по А.Харавану; вероятности из PRC1	сумма знаний	10687	7733	8116	2571	2894	0.752	0.729	0.740	902.271	2186.457
9	INF2 - частный критерий: количество знаний по А.Харавану; вероятности из PRC2	семантический резонанс знаний	10687	8596	8448	2239	2091	0.793	0.804	0.799	2747.891	3486.408
10	INF2 - частный критерий: количество знаний по А.Харавану; вероятности из PRC2	сумма знаний	10687	7820	8131	2556	2867	0.754	0.732	0.743	907.590	2224.140
11	INF3 - частный критерий: Хи-квадрат; разности между фактическими и ожидаемыми abs частотами	семантический резонанс знаний	10687	7884	7884	2803	2803	0.738	0.738	0.738	4090.477	4090.477
12	INF3 - частный критерий: Хи-квадрат; разности между фактическими и ожидаемыми abs частотами	сумма знаний	10687	7884	7884	2803	2803	0.738	0.738	0.738	4205.423	4205.423
13	INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1	семантический резонанс знаний	10687	7950	8985	1702	2737	0.824	0.744	0.782	2142.625	4492.136
14	INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1	сумма знаний	10687	7893	7889	2798	2794	0.738	0.739	0.738	774.450	928.699
15	INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2	семантический резонанс знаний	10687	7947	8987	1700	2740	0.824	0.744	0.782	2146.945	4502.658
16	INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2	сумма знаний	10687	7913	7909	2778	2774	0.740	0.740	0.740	777.633	946.784
17	INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1	семантический резонанс знаний	10687	7884	7884	2803	2803	0.738	0.738	0.738	4088.446	4091.469
18	INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1	сумма знаний	10687	7884	7884	2803	2803	0.738	0.738	0.738	1679.028	3918.111
19	INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2	семантический резонанс знаний	10687	7946	7948	2739	2741	0.744	0.744	0.744	4121.654	4124.679
20	INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2	сумма знаний	10687	7978	7978	2709	2709	0.747	0.747	0.747	1674.803	3966.092

Помощь по мерам Помощь по распрд. TP,TN,FP,FN (1-F)(1+T)*100

Рисунок 12. Экранные формы режима измерения достоверности моделей 3.4

На рисунках 13 приведены частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко [9] СК-модели INF3.

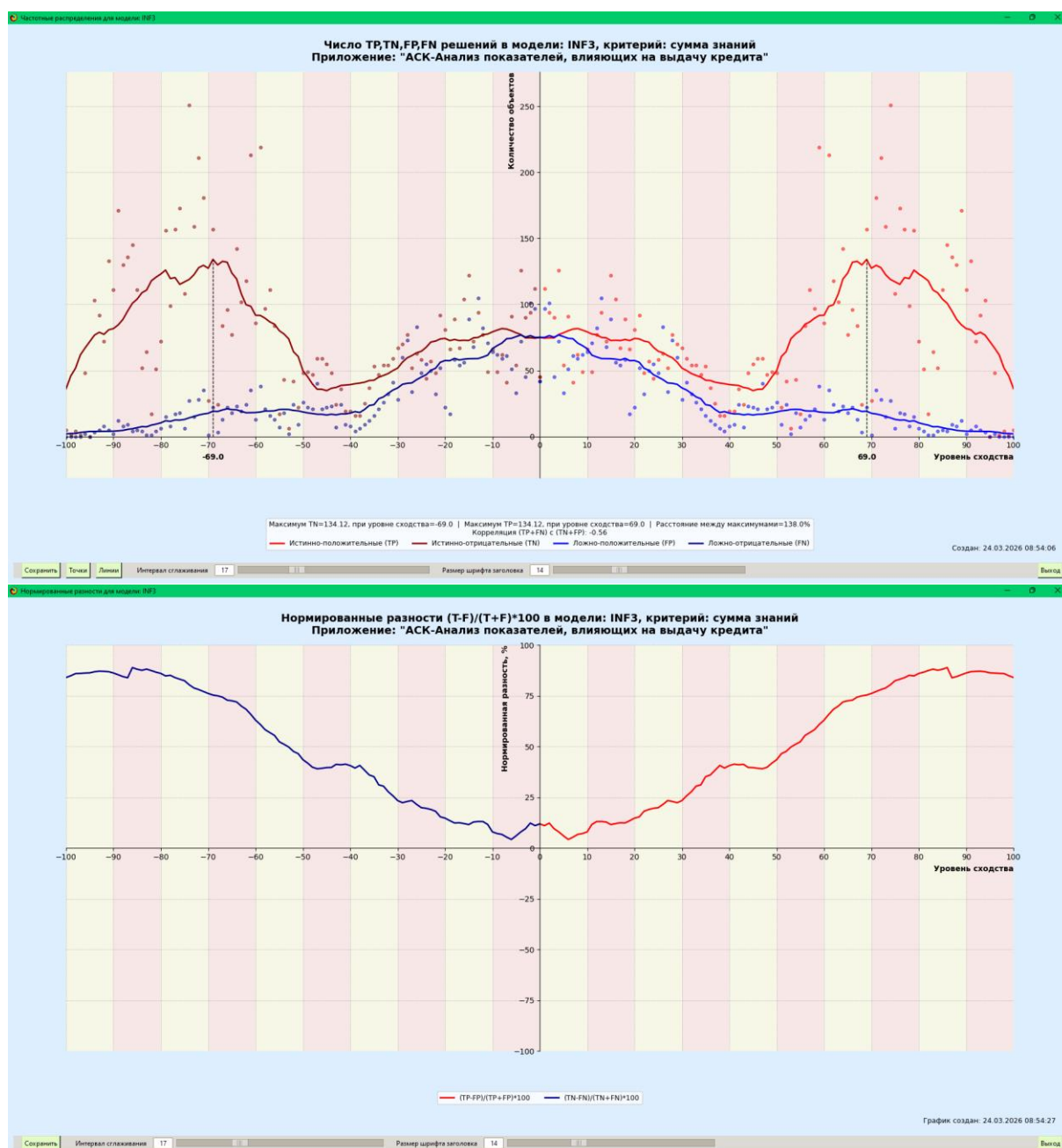


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF3 [9]

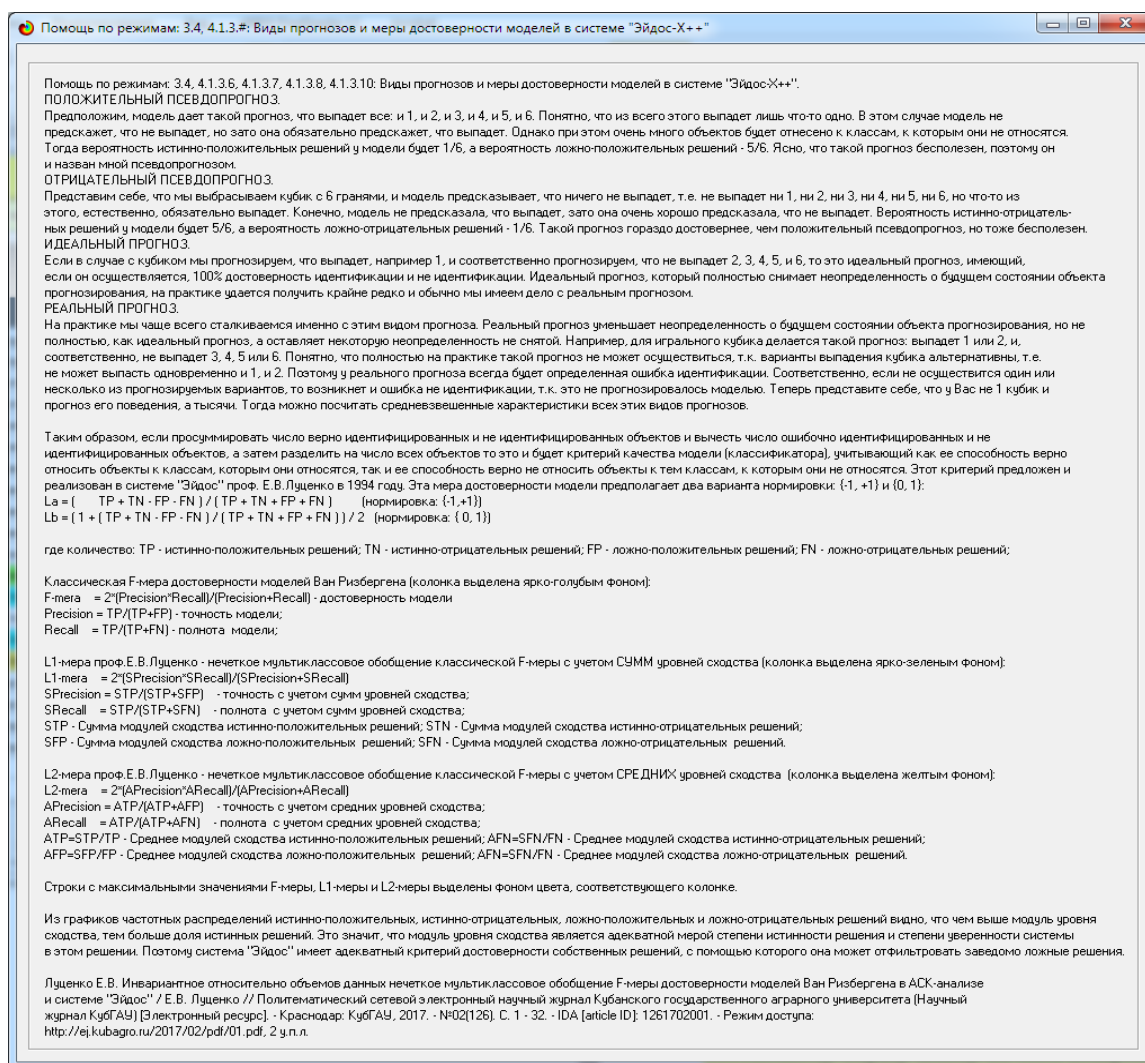
Из этих частотных распределений видно, что в наиболее достоверной по критерию достоверности L1-мерой проф.Е.В.Луценко СК-модели INF3:

- отрицательные ложные решения в данной задаче вообще не встречаются, а в общем случае такие решения практически не встречаются за исключением очень небольшого количества случаев с очень низкими уровнями различия;
- при уровнях сходства меньше 20% в данной задаче преобладают ложные положительные решения, а при более высоких уровнях сходства —

истинные положительные решения. В общем случае при уровнях сходства выше 60% ложных положительных решений практически вообще нет;

– **чем выше уровень сходства, тем больше доля истинных решений. Поэтому уровень сходства является адекватной внутренней мерой системы «Эйдос», так сказать адекватной мерой самооценки или аудита степени достоверности решений и уровня риска ошибочного решения.**

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



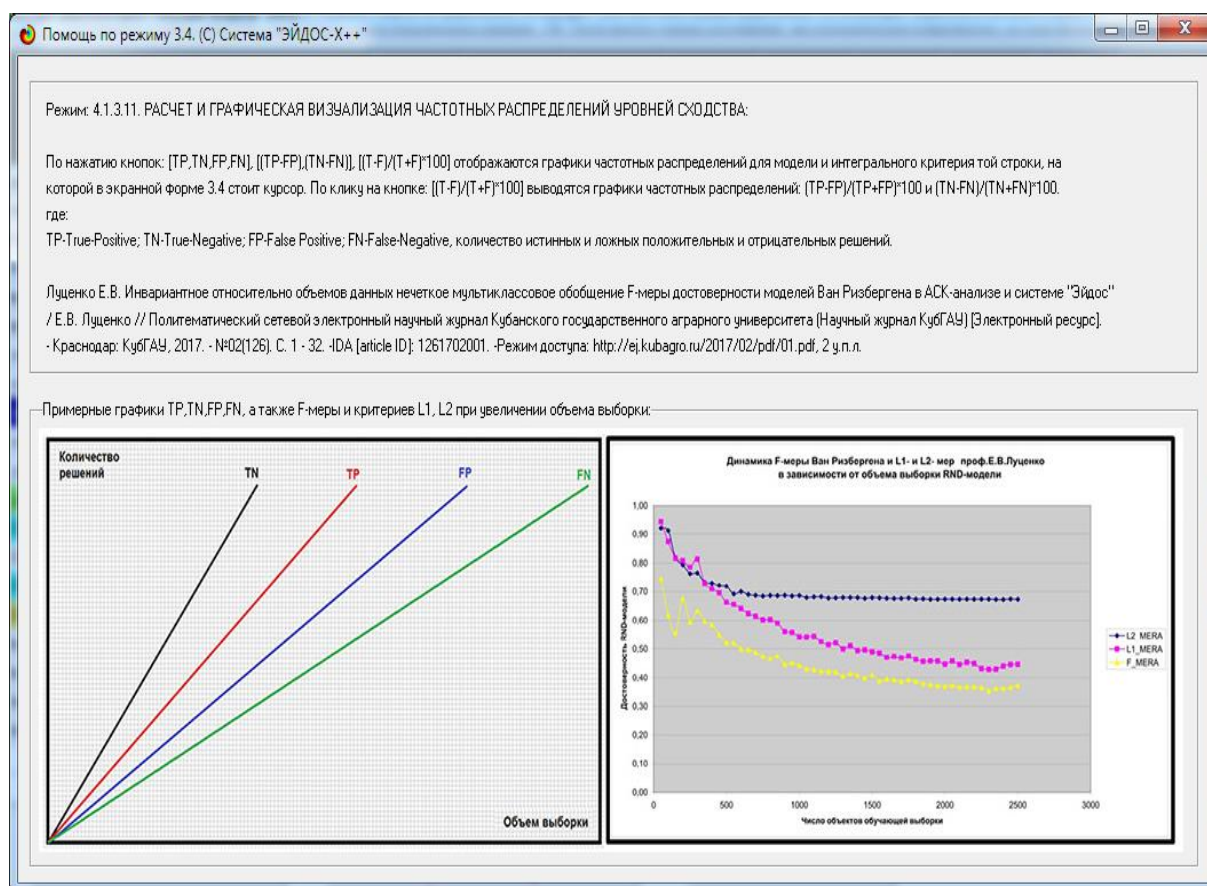


Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

3.5.1. Универсальная постановка задачи, не зависящая от предметной области

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.

3.5.2. Конкретное решение задачи в данной работе

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (рисунки 15). Это необходимо делать лишь для решения задачи идентификации и прогнозирования (в режиме 4.1.2), которая требует и

потребляет наибольшие вычислительные ресурсы и поэтому решается только для модели, заданной в качестве текущей. Все остальные расчеты проводятся в системе «Эйдос» сразу во всех моделях.

The dialog box contains the following sections and controls:

- Задайте текущую модель для идентификации**
 - Radio buttons: Abs, Prc1, Prc2, Inf1, Inf2 (top row); Inf3 (selected), Inf4, Inf5, Inf6, Inf7 (bottom row).
- На каком процессоре выполнять распознавание**
 - Radio buttons: На центральном процессоре (CPU) (selected), На графическом процессоре (GPU).
- Задайте алгоритм идентификации**
 - Radio buttons: Классический, работает дольше (точнее); Упрощенный, работает быстрее (selected).
 - Text: Модель для распознавания задается в этом же окне выше или в режиме 5.6.
- Учитывать только наиболее достоверные результаты распознавания**
 - Checkbox: с МОДУЛЕМ интегрального критерия "Резонанс знаний" не менее: 0.0 %.
- Buttons:** Ok, Cancel, Помощь.

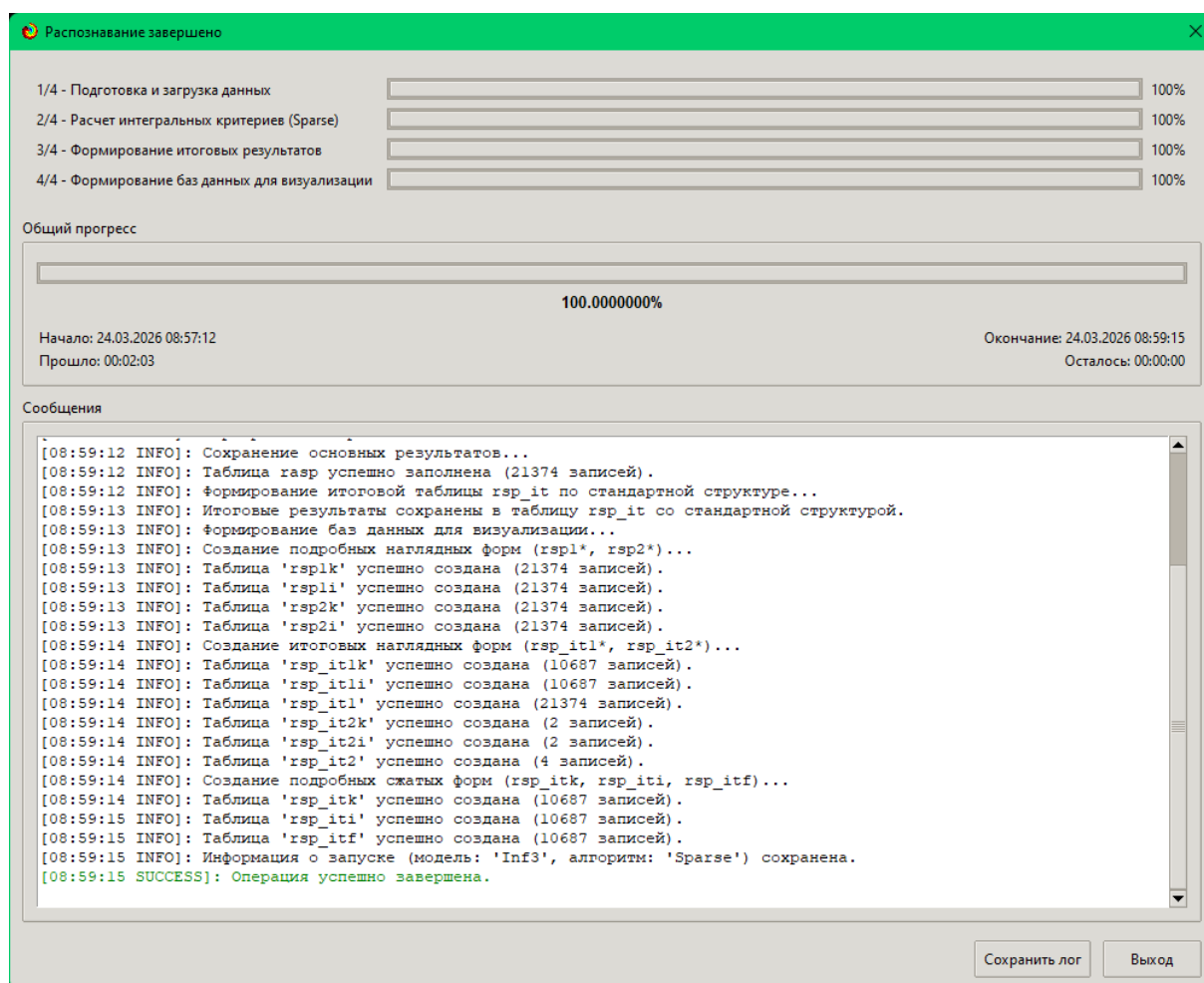


Рисунок 15. Задание СК-модели INF3 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

3.6.1. Универсальная постановка задачи, не зависящая от предметной области

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 (рисунок 3).

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

Сравнение осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что корректны¹⁰ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1.1. Интегральный критерий «Сумма знаний»

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

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

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

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

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

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

¹⁰ В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

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).

3.6.1.2. Интегральный критерий «Семантический резонанс знаний»

Интегральный критерий «Семантический резонанс знаний»

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

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

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

где:

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

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

$\bar{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса; $\bar{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта (состояния или явления), включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив-локатор), т.е.:

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

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

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

$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}$. Поэтому по своей сути он также является скалярным

произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применения вейвлетов и 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».

3.6.1.3. Важные математические свойства интегральных критериев

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

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

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

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

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

В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям

классов I_{ij} , т.е. определяется вес каждого обобщенного образа класса в образе объекта, что подробнее описано в работах [2, 3, 4].

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

3.6.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»

В АСК-анализе разработаны, а в системе «Эйдос» реализованы развитые методы прогнозирования, основанные на сценарном методе АСК-анализа или сценарном АСК-анализе [11]. Но в задачи данной работы не входит их подробное рассмотрение, тем более, что они подробно описаны и на теоретическом уровне, и с детальными численными примерами в работах [4-7, 11] и в ряде других [48].

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

По результатам решения задачи идентификации и прогнозирования в системе «Эйдос» выдается большое количество разнообразных выходных форм, которых в настоящее время 14 (рисунок 17).

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

4.1.2. Пакетное распознавание в текущей модели [Приложение: АСК-Анализ показател... X

Задайте текущую модель для идентификации

☐ Abs ☐ Prc1 ☐ Prc2 ☐ Inf1 ☐ Inf2
☒ Inf3 ☐ Inf4 ☐ Inf5 ☐ Inf6 ☐ Inf7

На каком процессоре выполнять распознавание

☒ На центральном процессоре (CPU)
☐ На графическом процессоре (GPU)

Задайте алгоритм идентификации

☐ Классический, работает дольше (точнее)
☒ Упрощенный, работает быстрее

Модель для распознавания задается в этом же окне выше или в режиме 5.6.

Учитывать только наиболее достоверные результаты распознавания

☐ с МОДУЛЕМ интегрального критерия "Резонанс знаний" не менее: %

Ok Cancel Помощь

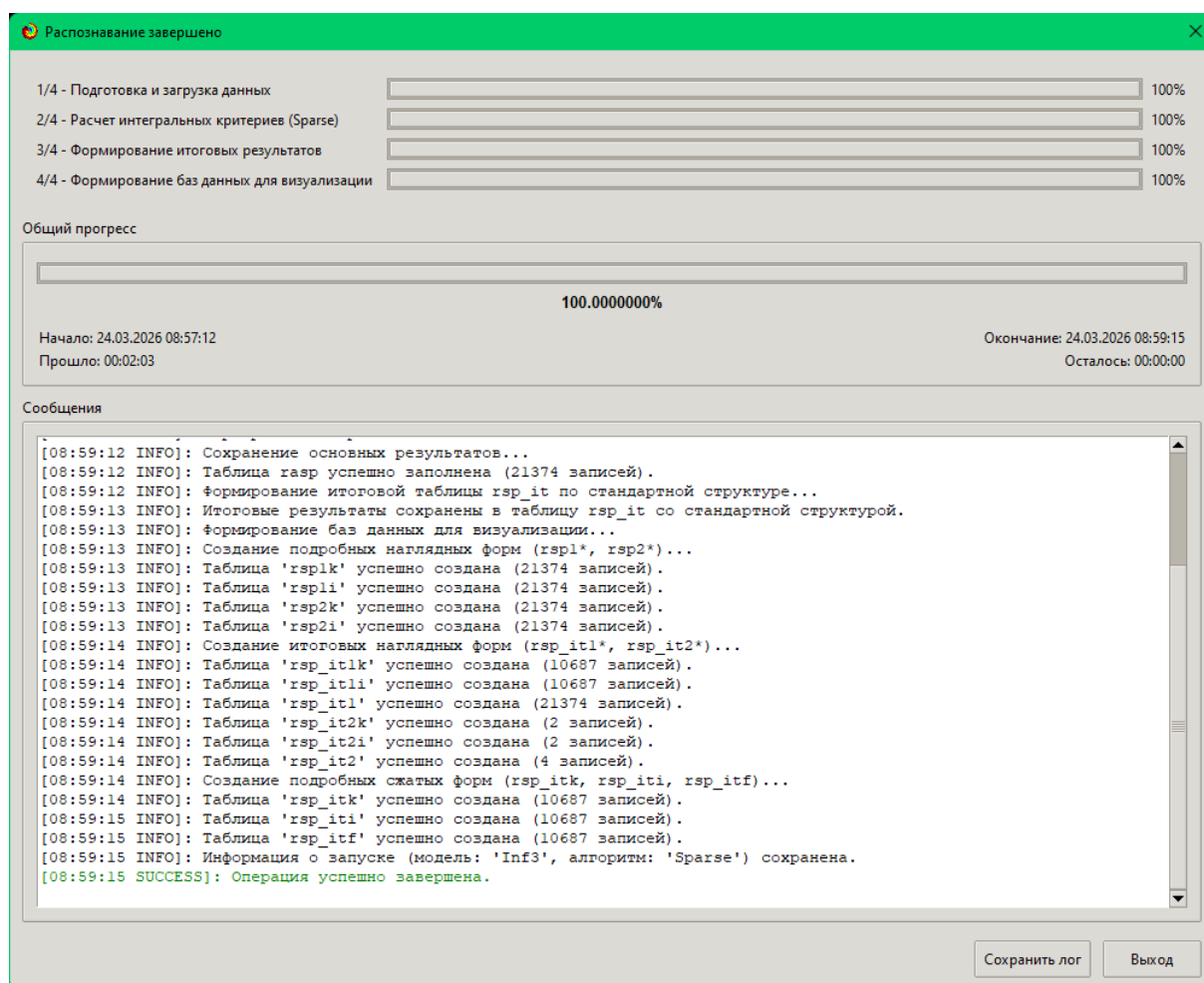


Рисунок 16. Экранные формы режима 4.1.2 идентификации и прогнозирования

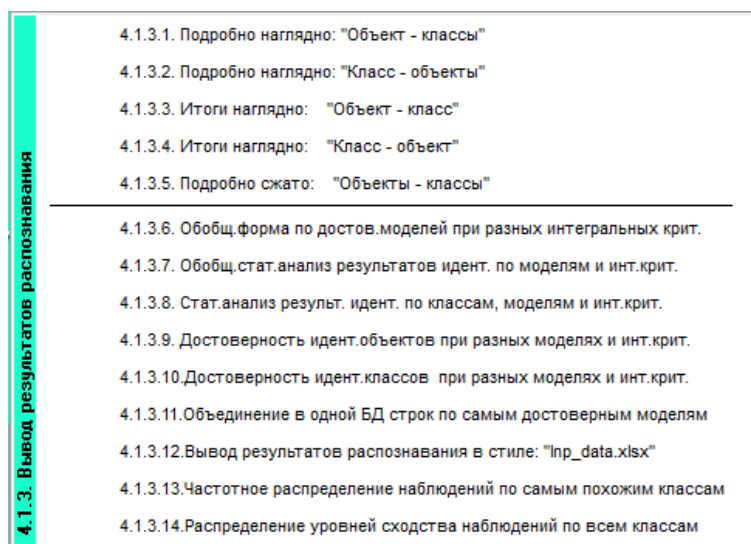


Рисунок 17. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

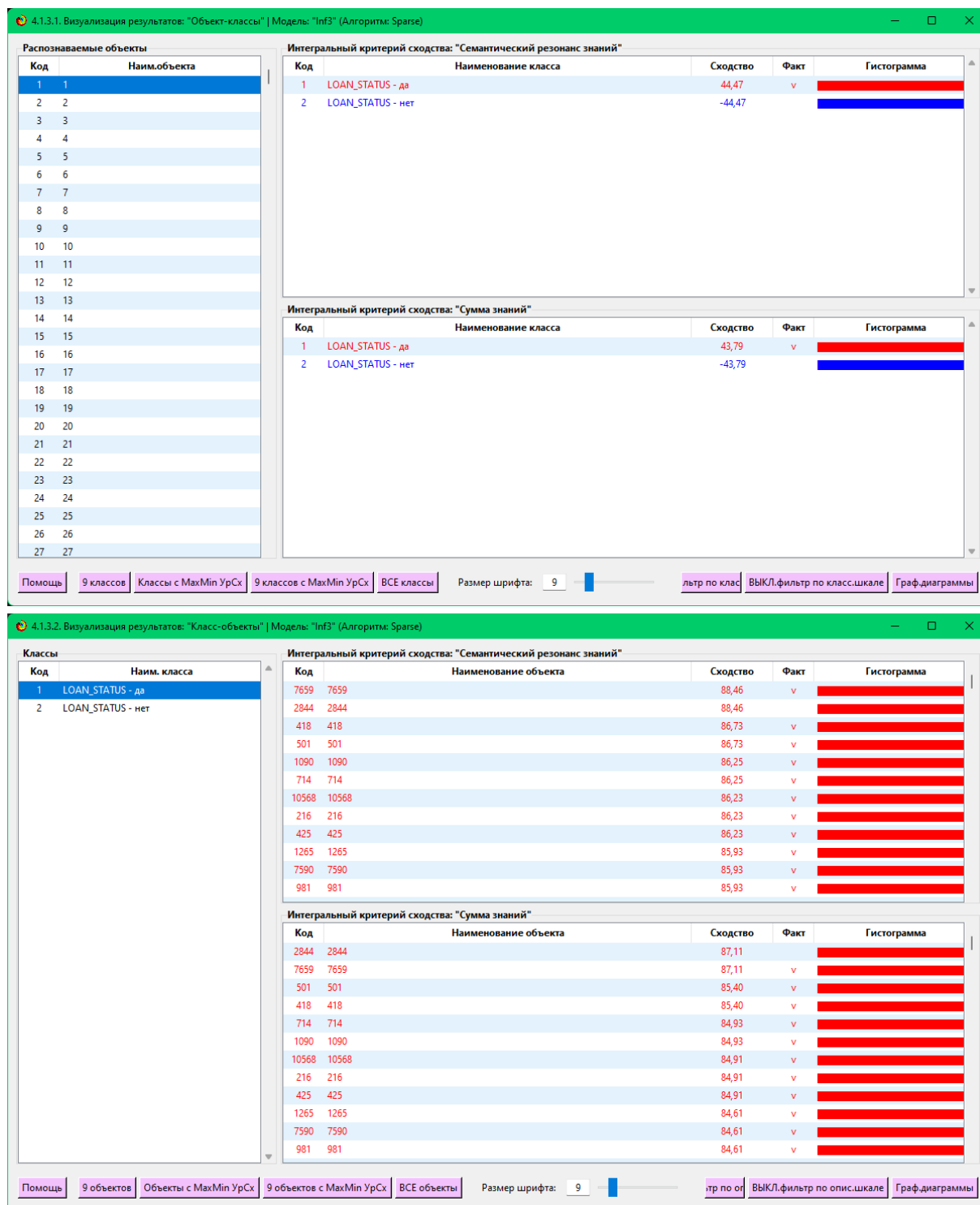


Рисунок 18. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

3.7. Задача-7. Поддержка принятия решений

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

3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области

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

- при прогнозировании по значениям факторов, действующих на объект моделирования, определяется в какое будущее состояние он перейдет под их действием;
- при принятии решений, наоборот, по будущему целевому состоянию объекта моделирования определяются значения факторов, которые обуславливают его переход в это будущее целевое состояние.

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

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

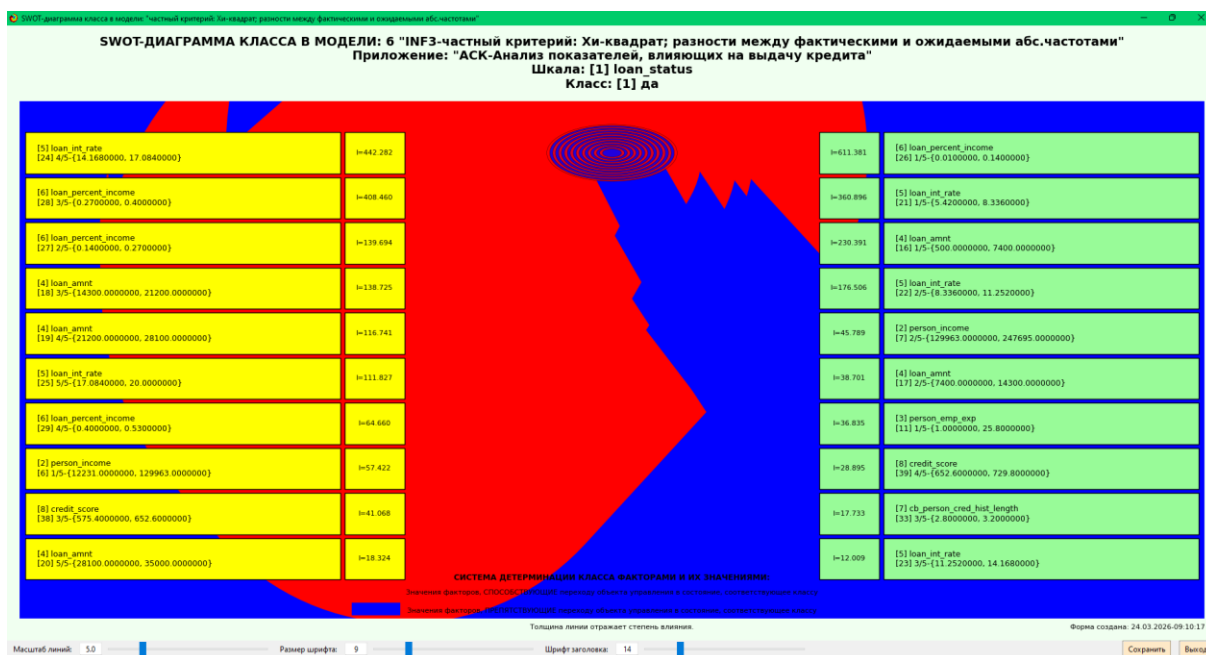
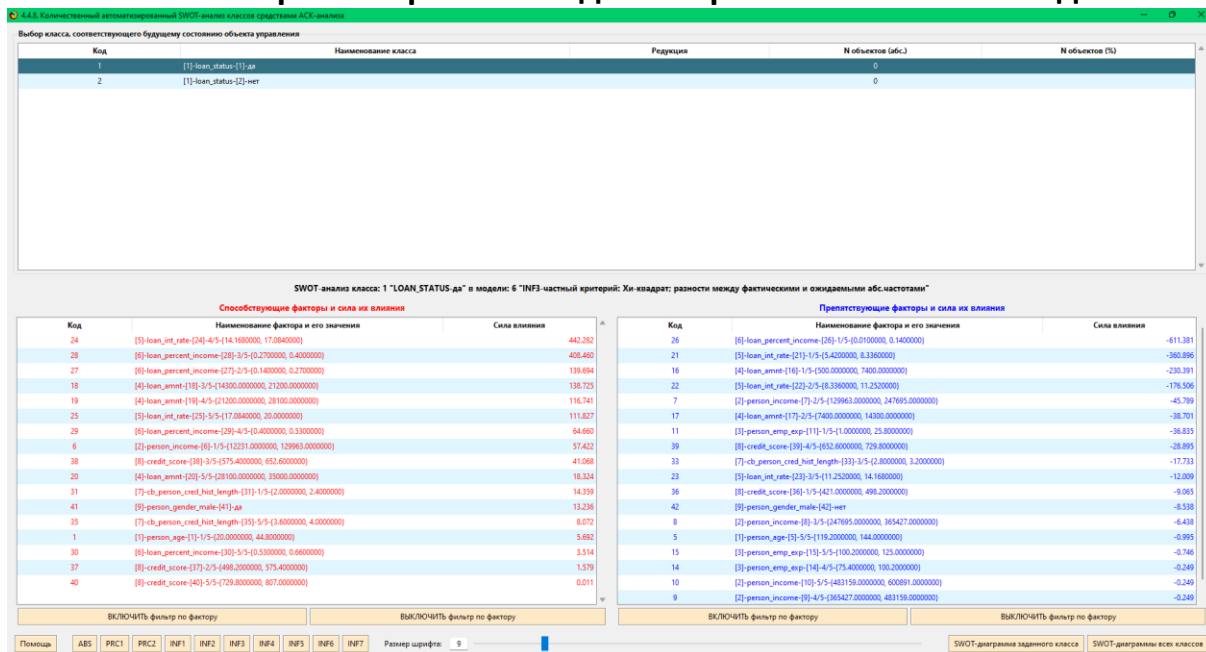
Графические выходные формы, приведенные на рисунках 19, интуитивно понятны и не требуют особых комментариев. Отметим лишь, что на SWOT-диаграммах наглядно показаны знак и сила влияния каждого значения фактора на переход объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне.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

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

Необходимо подчеркнуть, что в системе «Эйдос» SWOT-диаграммы формируются автоматически на основе статистических и системно-когнитивных моделей, созданных непосредственно на основе эмпирических данных, а не как обычно не формализуемым экспертным

путем на основе интуиции, опыта и профессиональной компетенции, т.е. практически «на глазок», а в некоторых случаях и вообще «от фонаря».

3.7.1.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»



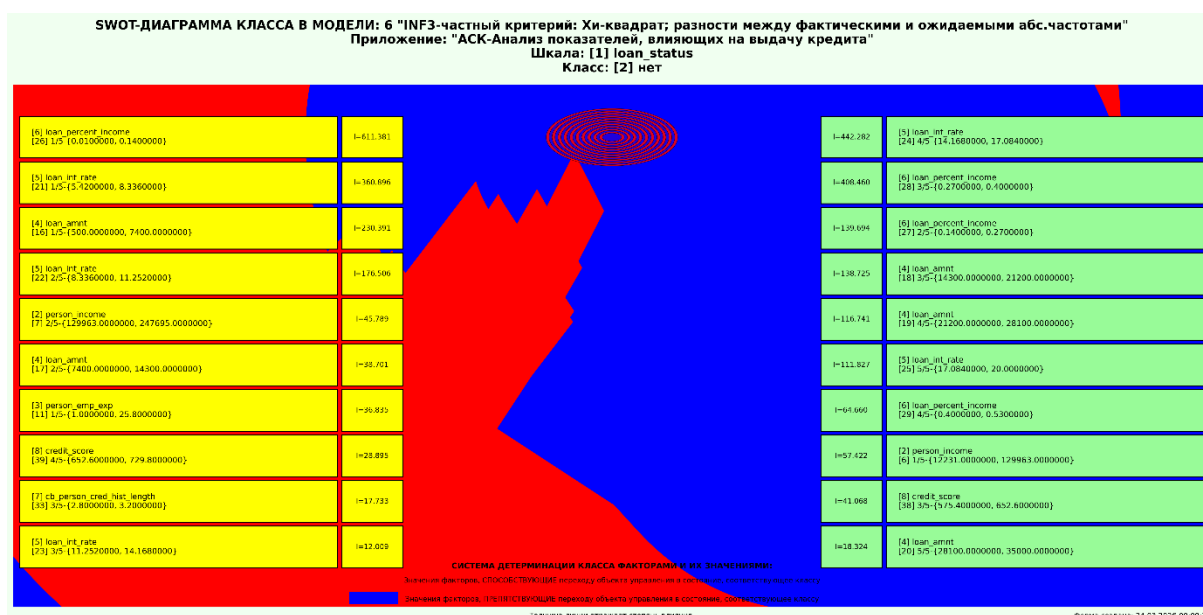


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

На основе приведенных выходных форм SWOT-анализа (критерий Хи-квадрат, разности между фактическими и ожидаемыми абсолютными частотами) для класса «нет» (кредит не одобрен) четко видно, что высокая доля отказа в выдаче кредита наиболее сильно детерминируется следующими факторами:

Наибольшее влияние оказывает показатель «loan_percent_income» (отношение суммы кредита к доходу) в интервале значений от 0,01 до 0,14 (первый квинтиль). Значение Хи-квадрат = 611,381. Это означает, что заемщики, запрашивающие сумму кредита, которая составляет менее 14% от их годового дохода, получают отказ значительно чаще, чем ожидалось при случайном распределении. Данный фактор вносит максимальный вклад в детерминацию отказа.

Вторым по значимости фактором является «loan_int_rate» (процентная ставка) в интервале от 5,42% до 8,336% (первый квинтиль). Значение Хи-квадрат = 360,896. Заемщики с наиболее низкими процентными ставками также демонстрируют повышенную частоту отказов, что может свидетельствовать о том, что низкая ставка предлагается высокорисковым клиентам в рамках компенсаторных механизмов либо же такие заявки отклоняются по иным причинам.

Третий ключевой фактор — «loan_amnt» (сумма кредита) в минимальном диапазоне от 500 до 740 единиц (первый квинтиль). Значение Хи-квадрат = 230,391. Наименьшие суммы кредита также ассоциированы с повышенной частотой отказов.

В совокупности указанные факторы формируют устойчивый портрет заемщика, которому наиболее часто отказывают: это лица,

запрашивающие небольшие суммы кредита под низкий процент, при этом сумма кредита составляет незначительную долю от их дохода. Парадоксальность ситуации заключается в том, что формально такие параметры свидетельствуют о низкой кредитной нагрузке, однако на практике они могут отражать недостаточную кредитную историю, низкий кредитный рейтинг либо принадлежность к определенным категориям заемщиков с высоким риском дефолта.

Кроме того, система детерминации класса факторами из зависимых показывает, что наибольшая стоимость требований (наборы объектов управления) формируется также за счет:

`loan_int_rate` в интервале 14,168–17,084% (четвертый квинтиль) со значением $\chi^2 = 442,282$ — здесь уже высокие ставки детерминируют отказы;

`loan_percent_income` в интервале 0,27–0,40 (третий квинтиль) со значением 408,460 — средняя доля кредита к доходу также значимо влияет на отказы.

Таким образом, анализ показывает, что детерминация отказа в выдаче кредита не сводится к какому-то одному фактору, а представляет собой сложную комбинацию: как минимальные, так и высокие значения по ряду финансовых показателей могут быть сигналами повышенного риска. Наиболее весомый вклад вносит сочетание низкого отношения кредита к доходу, низкой процентной ставки и минимальной суммы кредита, что в сумме формирует устойчивый профиль заемщика с высокой вероятностью отказа.

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

3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

Поэтому в АСК-анализе и системе «Эйдос» реализован развитый алгоритм принятия решений (режим 6.3) в котором кроме SWOT-анализа используются также результаты решения задачи прогнозирования и

результаты кластерно-конструктивного анализа классов и значений факторов, т.е. некоторые результаты решения задачи исследования предметной области. Этот алгоритм описан в работах [11, 12, 13] и в ряде других работ.

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

Шаг 1-й. Руководство ставит **цели** управления, т.е. определяет будущие целевые состояния объекта управления. Обычно целевые состояния в натуральном выражении – это количество и качество продукции, а в стоимостном выражении – прибыль и рентабельность. Объект управления как система, эффективность объекта управления как **системное свойство**, повышение уровня системности объекта управления как цель управления (нелинейность). Модель отражает определенный уровень технологий, поэтому целевые состояния, недостижимые в одной модели, могут оказаться достижимыми в другой с большим числом факторов [11, 12, 13, 14, 15, 16].

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

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

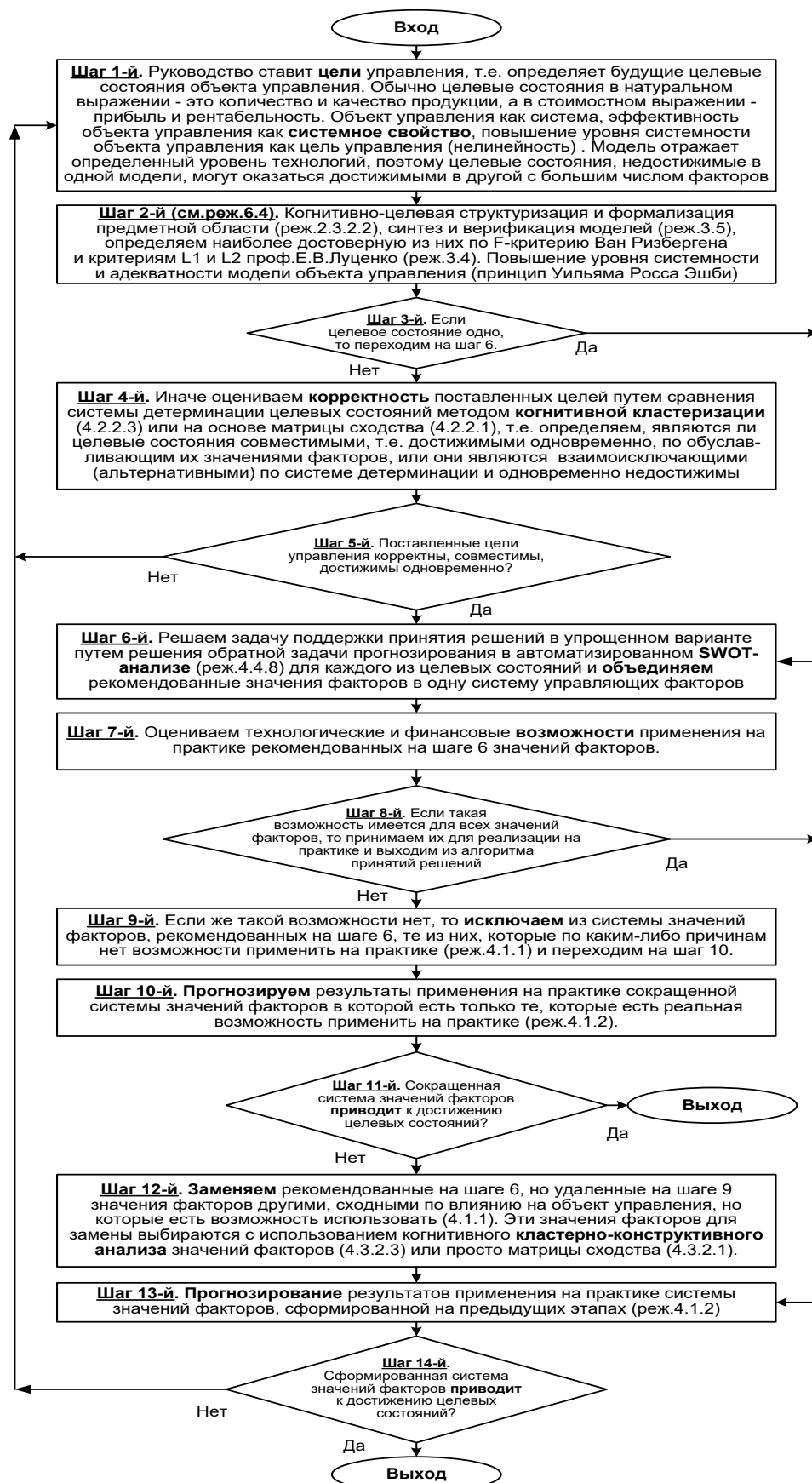


Рисунок 20. Разработанный алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

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

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

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

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

Шаг 12-й. **Заменяем** рекомендованные на шаге 6, но удаленные на шаге 9 значения факторов другими, сходными по влиянию на объект управления, но которые есть возможность использовать (4.1.1). Эти значения факторов для замены выбираются с использованием когнитивного **кластерно-конструктивного анализа** значений факторов (4.3.2.3) или просто матрицы сходства (4.3.2.1) [16].

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

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

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

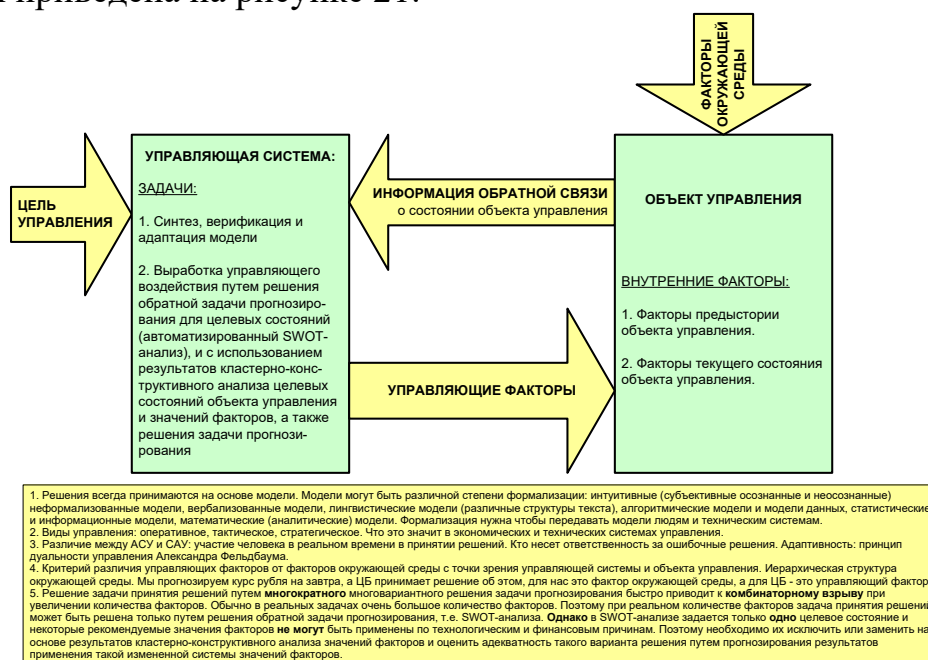


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

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

3.7.2.2. Конкретное решение задачи управления в данной работе в системе «Эйдос»

Необходимо отметить, что модели системы Эйдос – это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного *механизма* детерминации, а только сам факт и характер детерминации (силу и направление влияния факторов). *Содержательное* объяснение конкретных механизмов этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [12, 17].

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

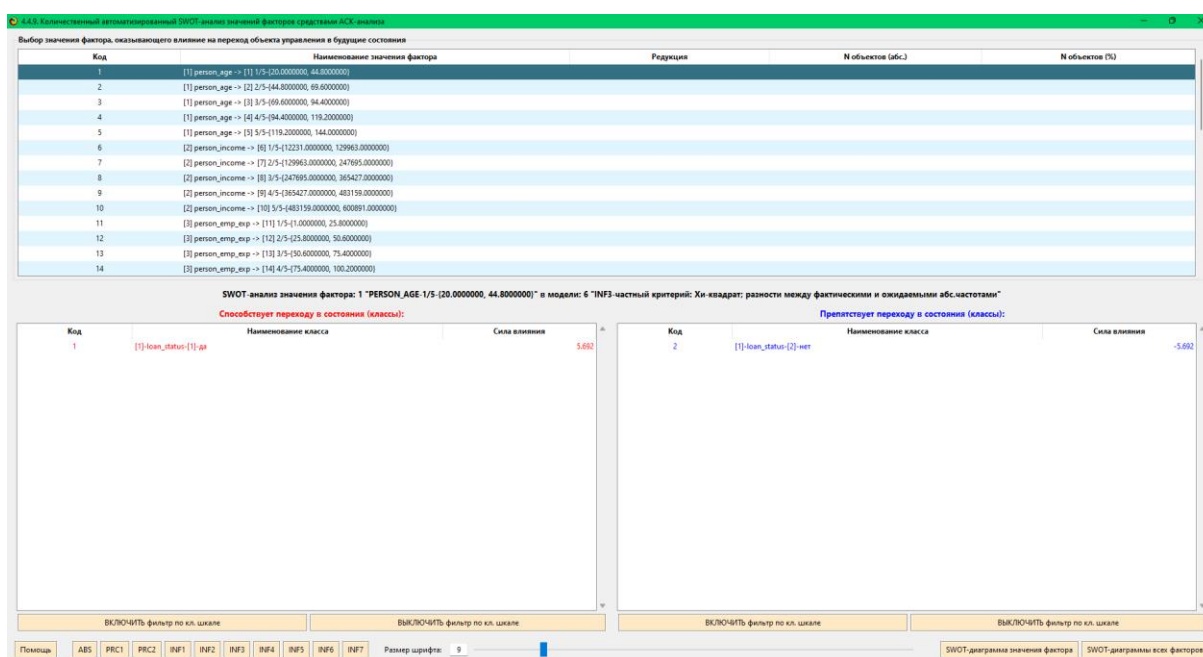
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)

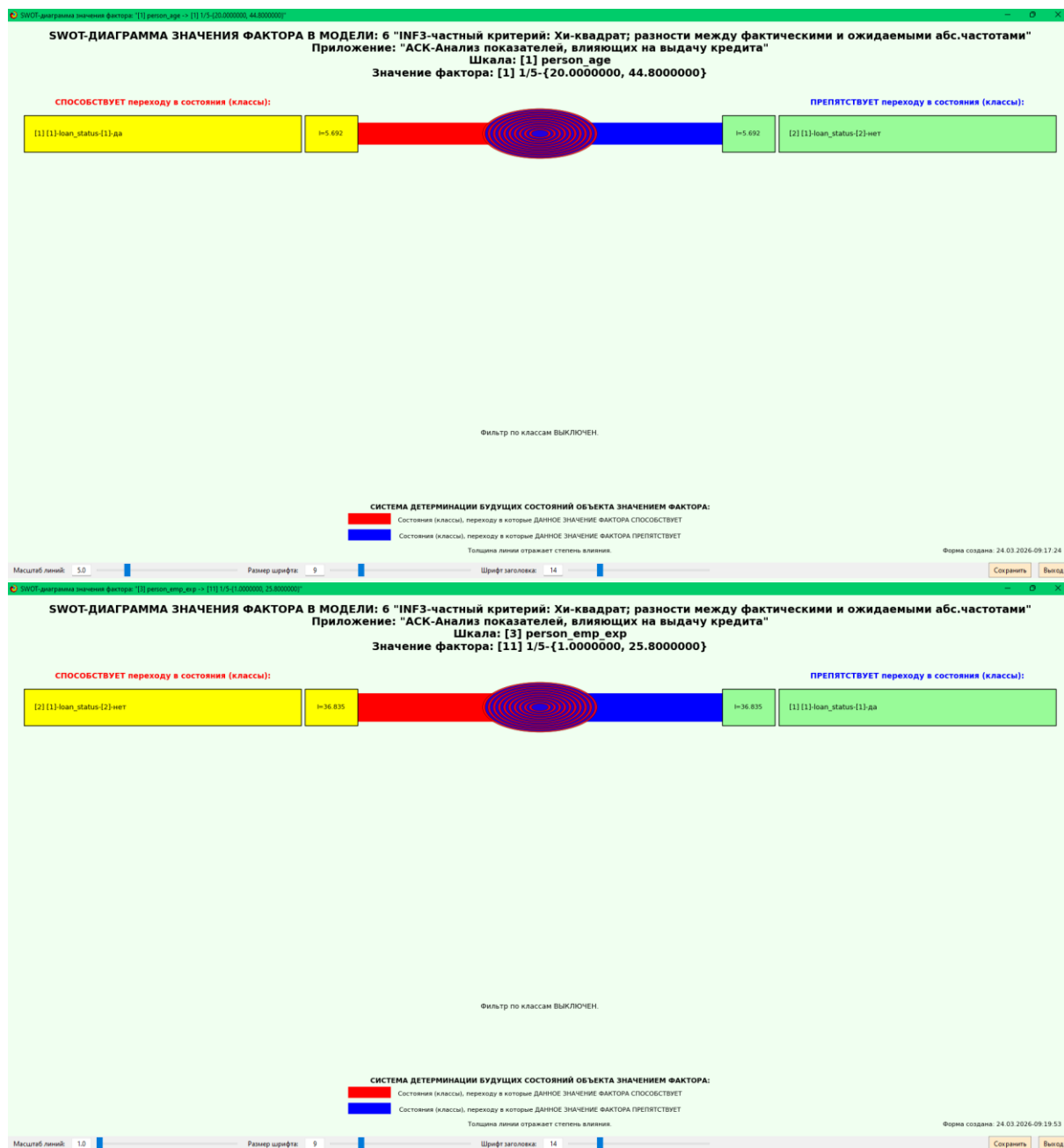
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области

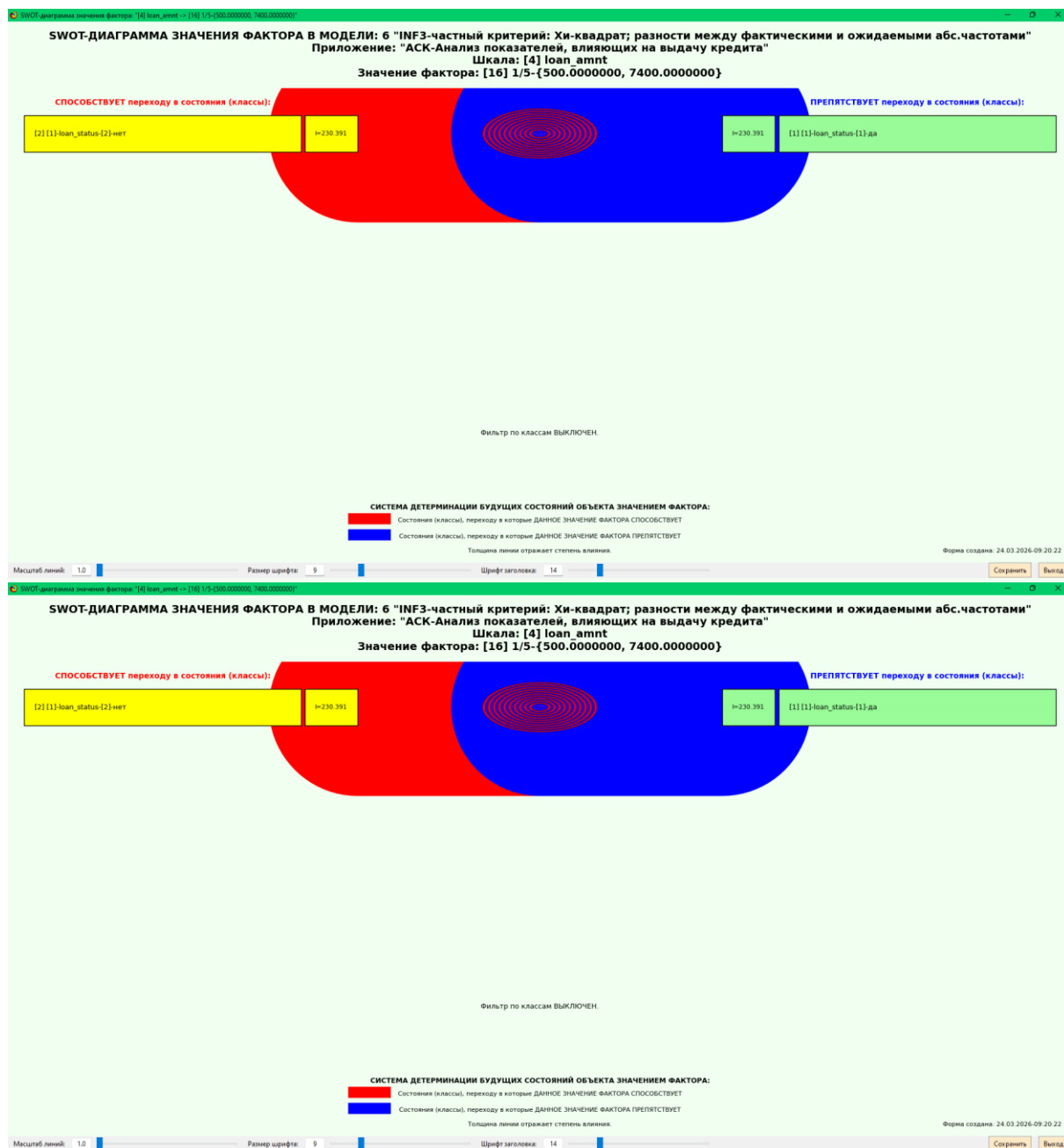
Инвертированные SWOT-диаграмм (предложены автором в работе [10]), отражают силу и направление влияния конкретной градации описательной шкалы на переход объекта моделирования в состояния, соответствующие градациям классификационных шкал (классы). Это и есть *смысл* (семантический потенциал) этой градации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

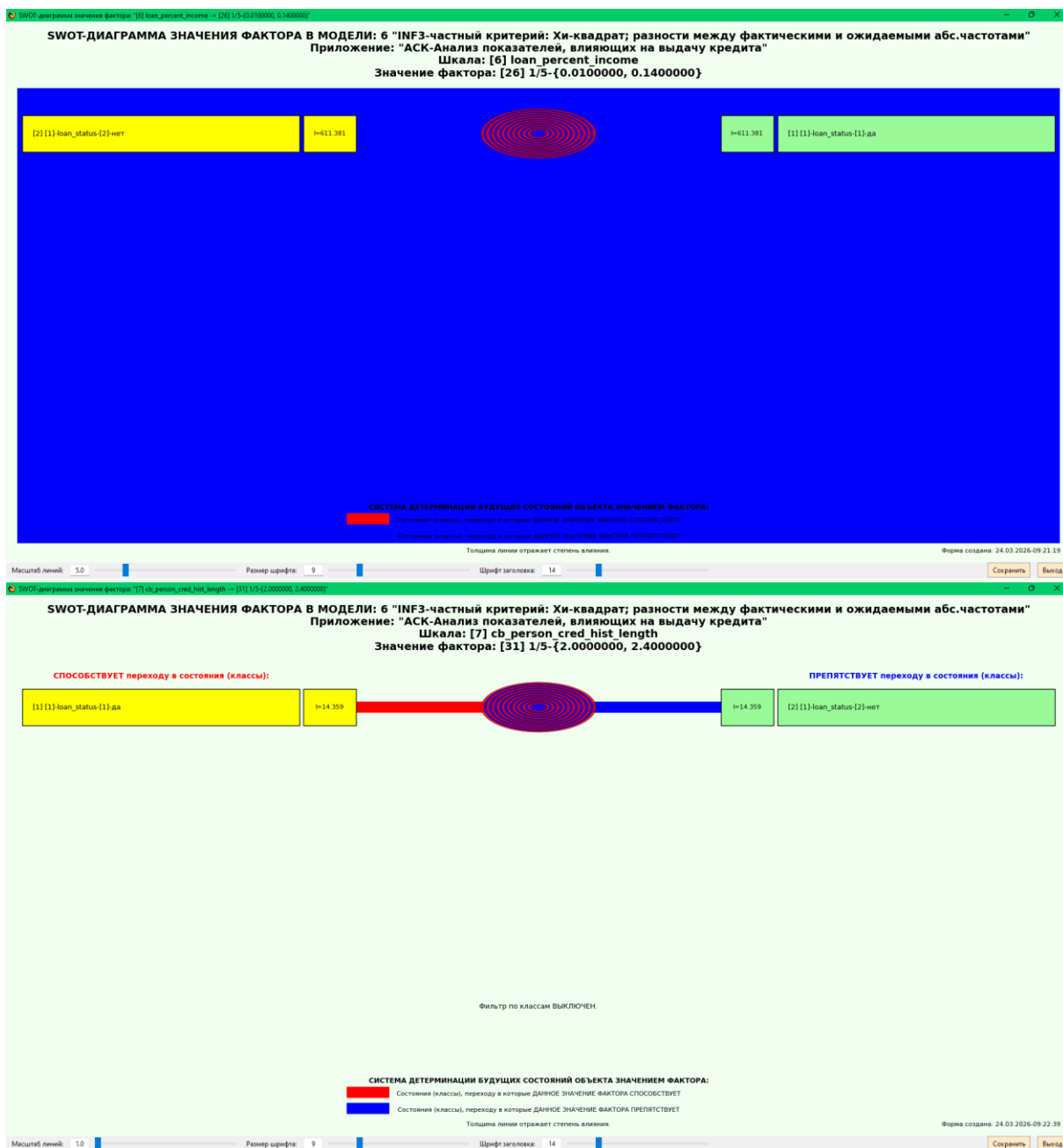
3.8.1.2. Конкретное решение задачи в данной работе

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









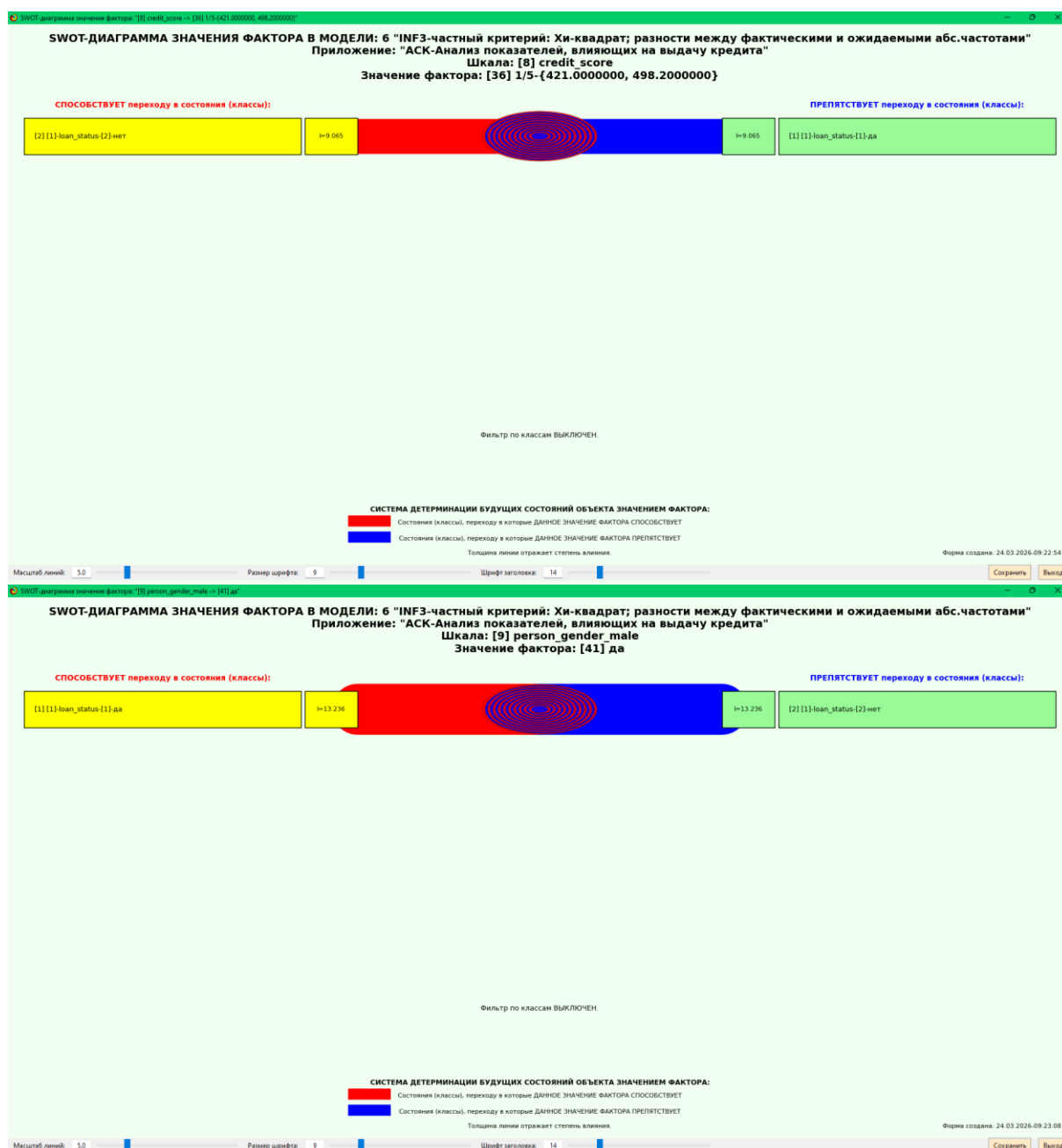


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 22 инвертированные SWOT-диаграммы исчерпывающим образом отражают силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в состояния, соответствующие различным классам. Из-за большого кол-ва факторов, все их втсавлять не имеет смысла, я всталвил 10 факторов из 47

Из приведенных выходных форм SWOT-анализа видно, что большее влияние оказывает "Заработок в год" и "Кредитная история"о многом это и есть реальное решение проблемы, поставленной в работе.

3.8.2. Кластерно-конструктивный анализ классов

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

3.8.2.2. Конкретное решение задачи в данной работе

У меня всего 2 класса, поэтому не имеет смысла делать когнитивную 2d-диаграмму классов.

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

3.8.3.2. Конкретное решение задачи в данной работе

Поскольку у меня всего 2 класса, то делать дендрограммный анализ классов не имеет смысла

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области

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 реализации и быстродействие системы.

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

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

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

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

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

3.8.4.2. Конкретное решение задачи в данной работе

4.4.10. Графическое отображение нелокальных нейронов в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	[1]-LOAN_STATUS-[1]-да
2	[1]-LOAN_STATUS-[2]-нет

Подготовка визуализации нейрона: 1 "[1]-LOAN_STATUS-[1]-да" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование описательной шкалы и градации	Сила
24	[5] loan_int_rate -> [24] 4/5-(14.1680000, 17.0840000)	442.2819
28	[6] loan_percent_income -> [28] 3/5-(0.2700000, 0.4000000)	408.4599
27	[6] loan_percent_income -> [27] 2/5-(0.1400000, 0.2700000)	139.6937
18	[4] loan_amnt -> [18] 3/5-(14300.0000000, 21200.0000000)	138.7249
19	[4] loan_amnt -> [19] 4/5-(21200.0000000, 28100.0000000)	116.7407
25	[5] loan_int_rate -> [25] 5/5-(17.0840000, 20.0000000)	111.8272
29	[6] loan_percent_income -> [29] 4/5-(0.4000000, 0.5300000)	64.6595
6	[2] person_income -> [6] 1/5-(12231.0000000, 129963.0000000)	57.4217
38	[8] credit_score -> [38] 3/5-(575.4000000, 652.6000000)	41.0677
20	[4] loan_amnt -> [20] 5/5-(28100.0000000, 35000.0000000)	18.3244
31	[7] cb_person_cred_hist_length -> [31] 1/5-(2.0000000, 2.4000000)	14.3587

ТОРМОЗЯЩЕЕ рецепторы и сила их влияния

Код	Наименование описательной шкалы и градации	Сила
26	[6] loan_percent_income -> [26] 1/5-(0.0100000, 0.1400000)	-611.3805
21	[5] loan_int_rate -> [21] 1/5-(5.4200000, 8.3360000)	-360.8959
16	[4] loan_amnt -> [16] 1/5-(500.0000000, 7400.0000000)	-230.3914
22	[5] loan_int_rate -> [22] 2/5-(8.3360000, 11.2520000)	-176.5062
7	[2] person_income -> [7] 2/5-(129963.0000000, 247695.0000000)	-45.7890
17	[4] loan_amnt -> [17] 2/5-(7400.0000000, 14300.0000000)	-38.7010
11	[3] person_emp_exp -> [11] 1/5-(1.0000000, 25.8000000)	-36.8345
39	[8] credit_score -> [39] 4/5-(652.6000000, 729.8000000)	-28.8954
33	[7] cb_person_cred_hist_length -> [33] 3/5-(2.8000000, 3.2000000)	-17.7333
23	[5] loan_int_rate -> [23] 3/5-(11.2520000, 14.1680000)	-12.0093
36	[8] credit_score -> [36] 1/5-(421.0000000, 498.2000000)	-9.0648

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

Помощь ABS PRC1 PRC2 INF1 INF2 INF3 INF4 INF5 INF6 INF7

СТАРТ Максимальное количество отображаемых рецепторов: 999 Минимальный вес. коэфф. отображаемых рецепторов: 0.000 ОДИН НЕЙРОН

Сортировать рецепторы: ☒ по информативности ☐ по модулю информативности

Отображать рецепторы: ☒ с наименованиями ☐ только с кодами

Размер изображения: по оси X: 1920 по оси Y: 1080

Размер шрифта графической формы: Заголовок: 14 Рецепторы: 6 Легенда: 6

Размер шрифта в таблицах данной формы: 9



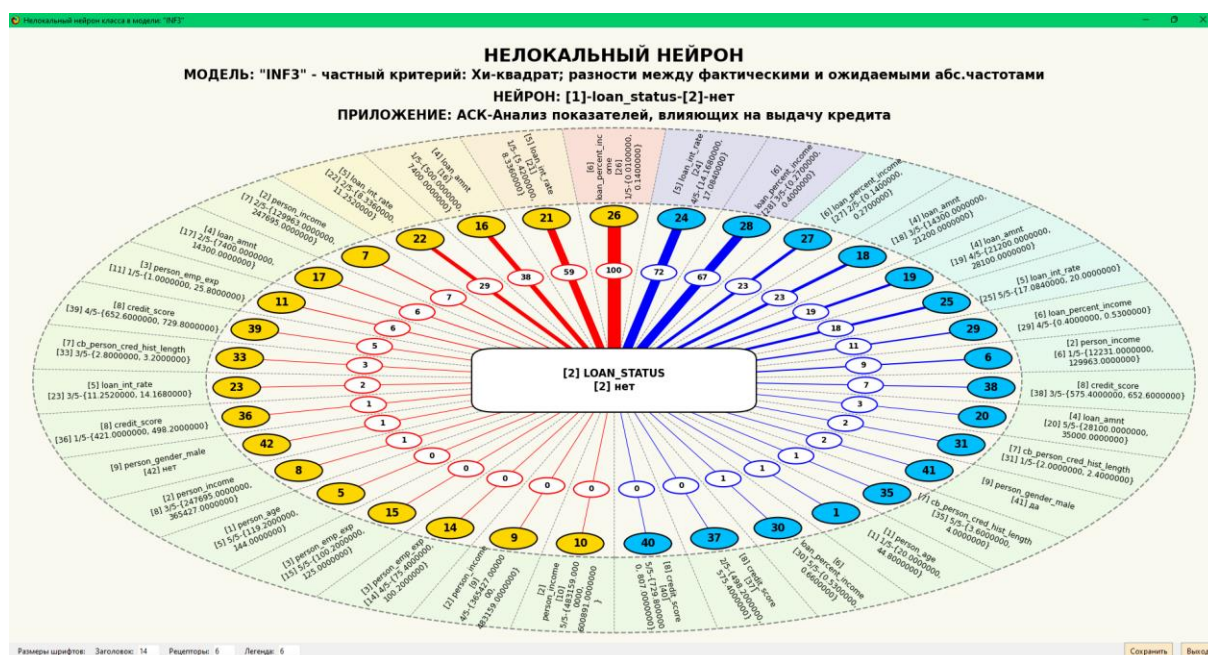


Рисунок 23. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

Нейроны на изображении слоя нейронной сети расположены слева направо в порядке убывания модуля суммарной силы их детерминации рецепторами, т.е. слева находятся результаты, наиболее жестко обусловленные действующими на них значениями факторов, а справа – менее жестко обусловленные (рисунок 32). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.5.2. Конкретное решение задачи в данной работе

На рисунке 32 мы видим, что наиболее сильно детерминирована значениями факторов одобрения кредита в банке, затем по степени детерминированности идет годовой доход, а средняя наиболее слабо обусловлена значениями факторов, действующими на объект моделирования.

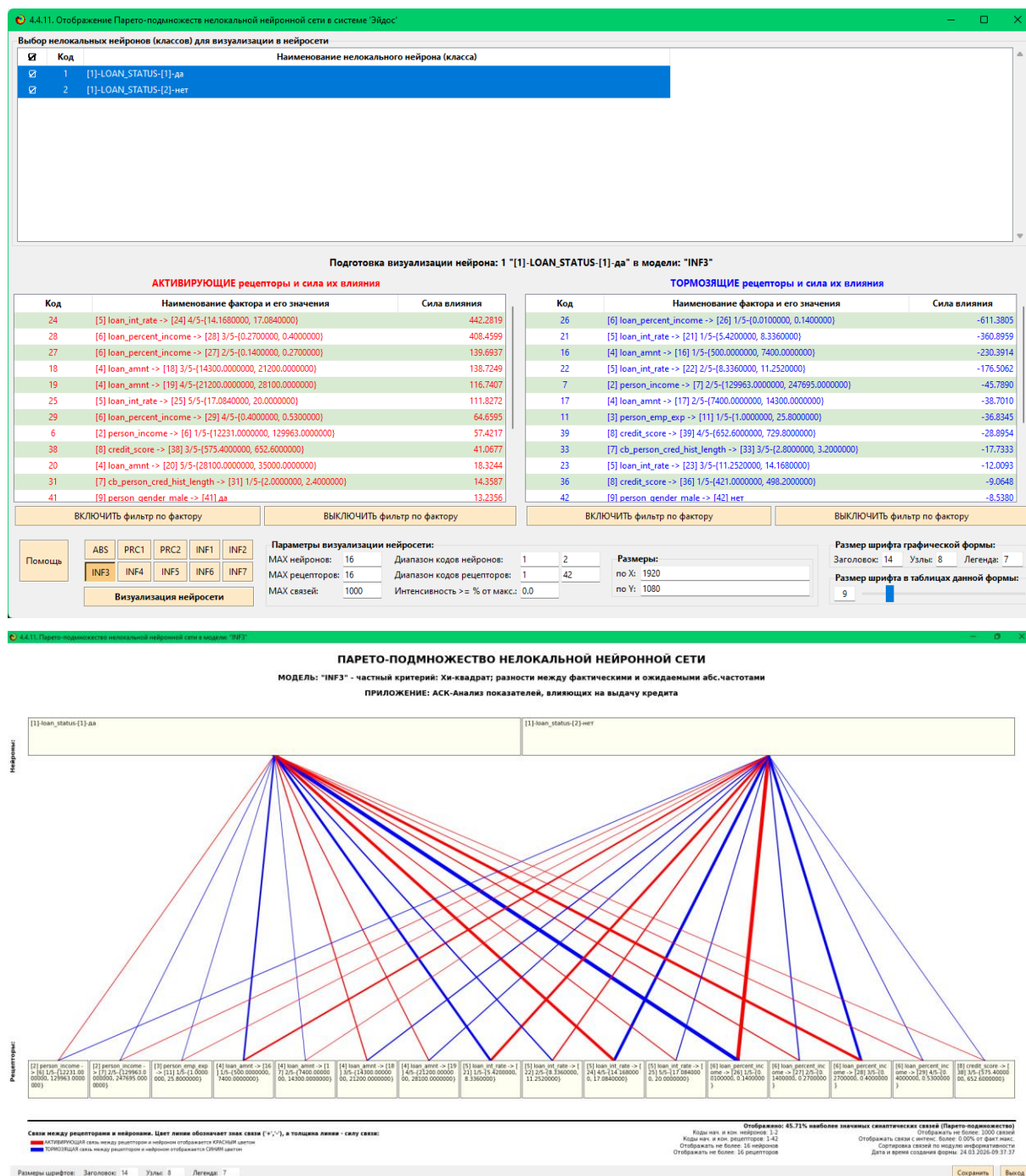


Рисунок 24. Нейронная сеть в СК-модели INF3

3.8.6. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения)

3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых может быть одним из первых писал Дьердь Пойа [19, 20]. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году в работе [2] на странице 521¹¹. Позже об этом писалось в работе [3]¹² и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

Пример опосредованных правдоподобных рассуждений.

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

3.8.6.2. Конкретное решение задачи в данной работе

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 34. Всего системой в данной модели генерируется 9 форм содержательного сравнения классов. Так как каждый из 3 классов сравнивается со всеми остальными, в т.ч. с собой, то всего получается $3^2=9$ подобных диаграмм. На рисунках 34 приводятся некоторые из этих диаграмм. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №393 и получить в нем все выходные формы, как это описано в данной статье.

¹¹ https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf, Таблица 7. 17, стр. 521

¹² <http://ej.kubagro.ru/2013/07/pdf/15.pdf>, стр.44.

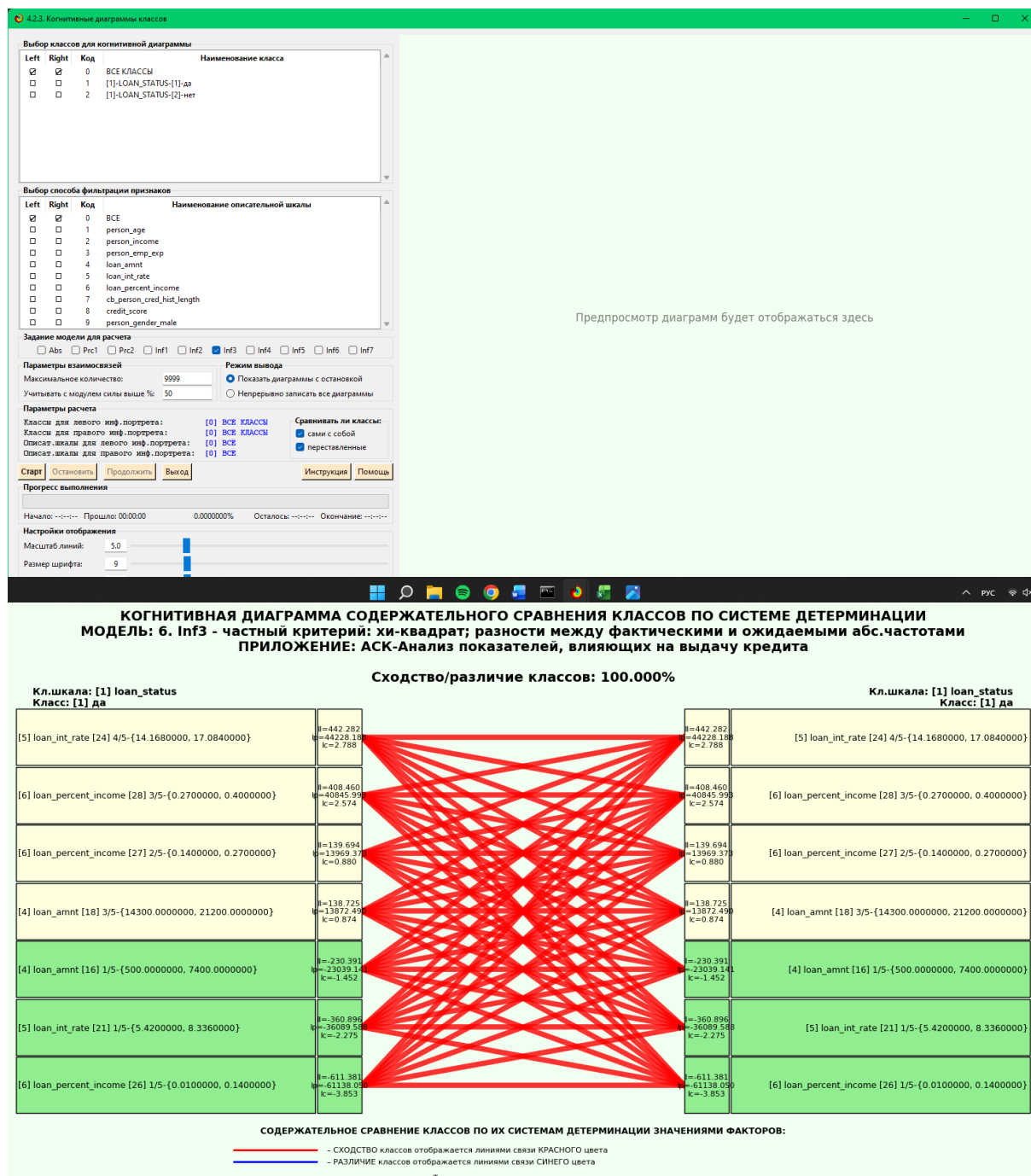


Рисунок 25. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF3

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

3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области

Из 2d-когнитивных диаграммах сравнения значений факторов по их влиянию на объект моделирования, т.е. на его переходы в состояния,

соответствующие классам вполне понятно, насколько сходны или отличаются любые два значения факторов по их смыслу.

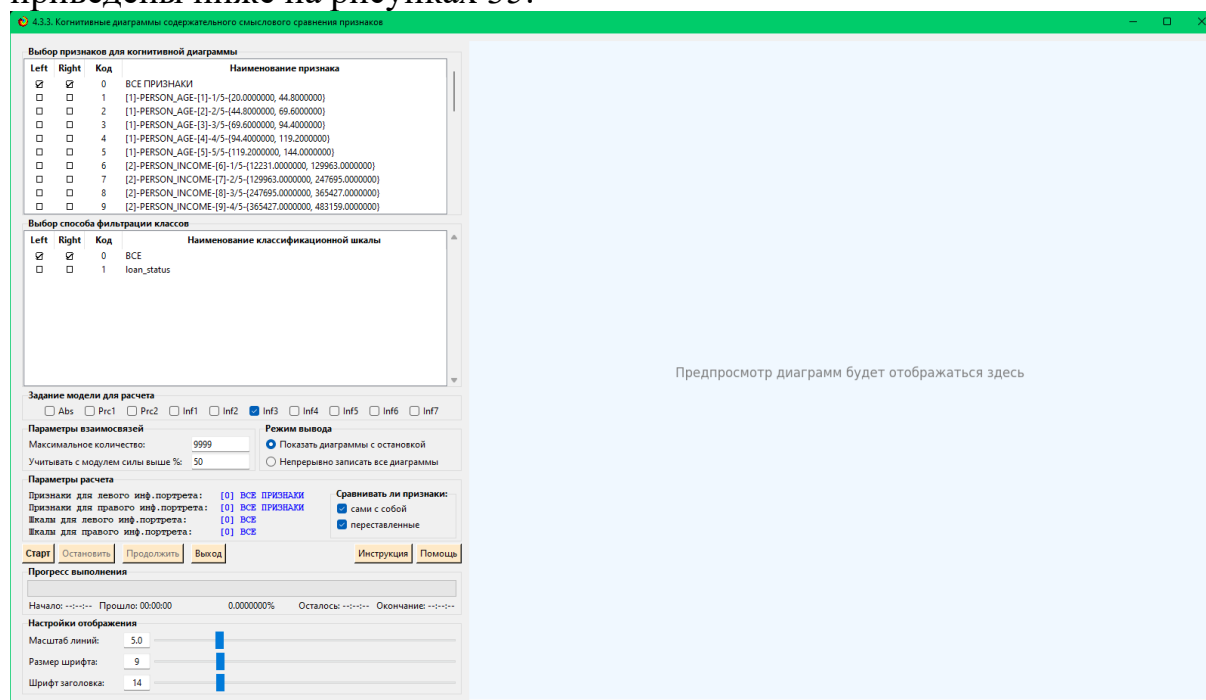
Напомним, что смысл событий, согласно концепции смысла Шенка-Абельсона, используемой в АСК-анализе, состоит в знании причин и последствий этих событий [21].

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

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

3.8.7.2. Конкретное решение задачи в данной работе

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



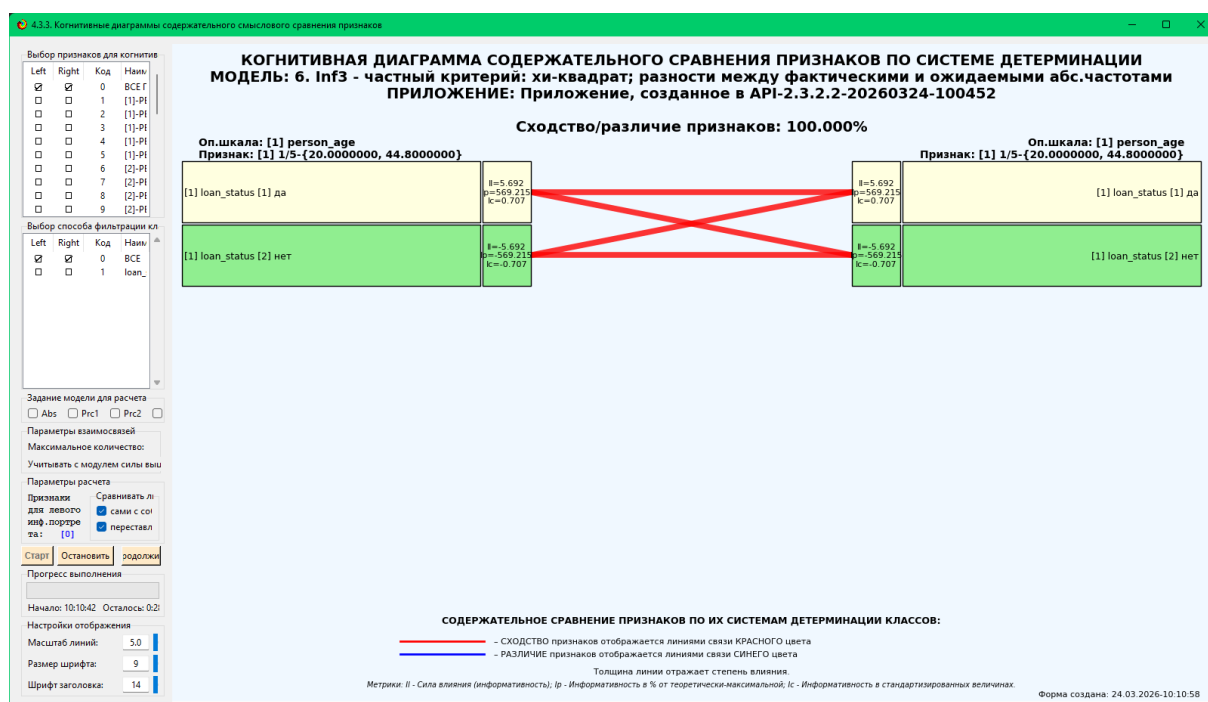


Рисунок 26. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF3

Всего системой в данной модели генерируется $8^2=64$ подобных диаграмм содержательного сравнения значений факторов по их смыслу, т.е. по влиянию на объект моделирования. В данной работе все они приводятся. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №393 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

соответствуют в разной степени, причем как положительной, так и отрицательной, которая отображается цветом.

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

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

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

3.8.9.2. Конкретное решение задачи в данной работе

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 36). Количество когнитивных функций равно количеству сочетаний описательных и классификационных шкал. В модели, рассматриваемой в данной работе, есть 9 описательных шкалы и 1 классификационная, поэтому получается 9 когнитивных функции:

4.5. Визуализация когнитивных функций

—Что такое когнитивная функция:

Визуализация прямых, обратных, позитивных, негативных, полностью и частично редуцированных когнитивных функций. Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам. Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека. Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации; негативные, отражающие чему препятствуют система детерминации; средневзвешенные, отражающие совокупное влияние всех значений факторов на поведение объекта (причем в качестве весов наблюдений используется количество информации в значении аргумента о значениях функции) различной степенью редукции или степенью детерминации, которая отражает в графической форме (в форме полосы) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала. Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию. Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой "Эйдос". Необходимо отметить, что на вид функций влияния математической моделью АСК-анализа не накладывается никаких ограничений, в частности, они могут быть и не дифференцируемыми.

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(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077. - 2,688 у.п.л. - Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/282.pdf>

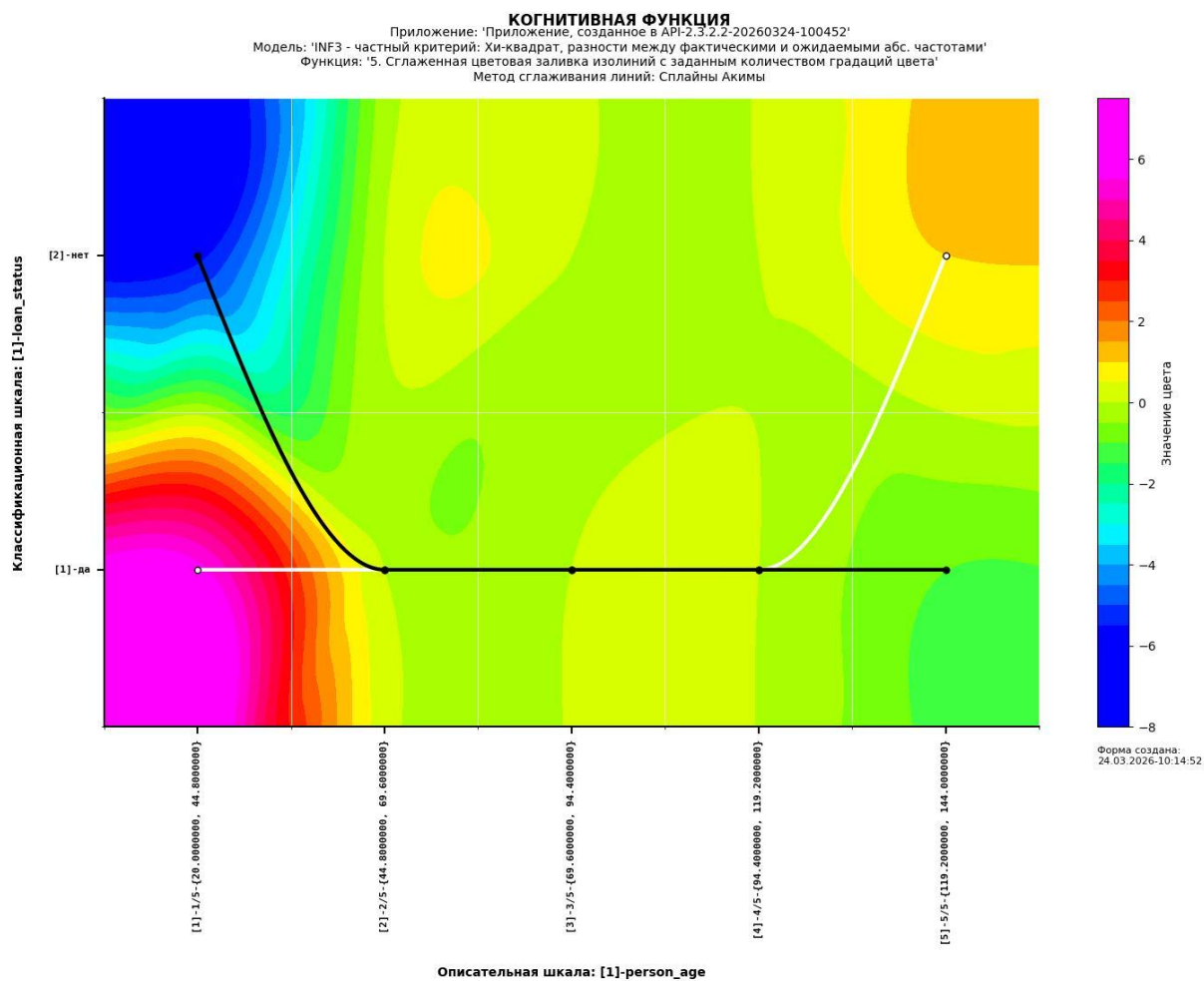
— Задайте нужный режим:

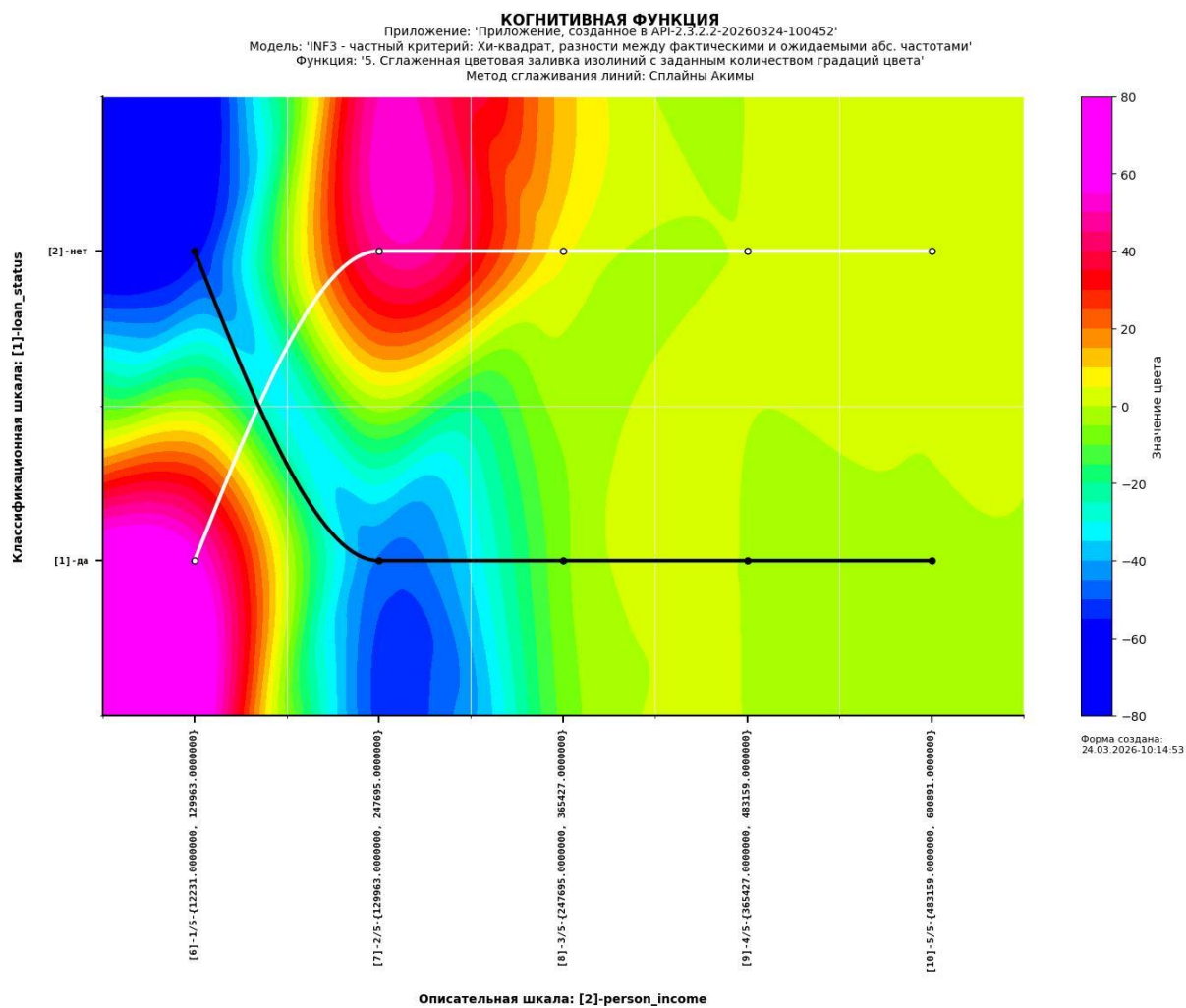
Визуализации когнитивных функций

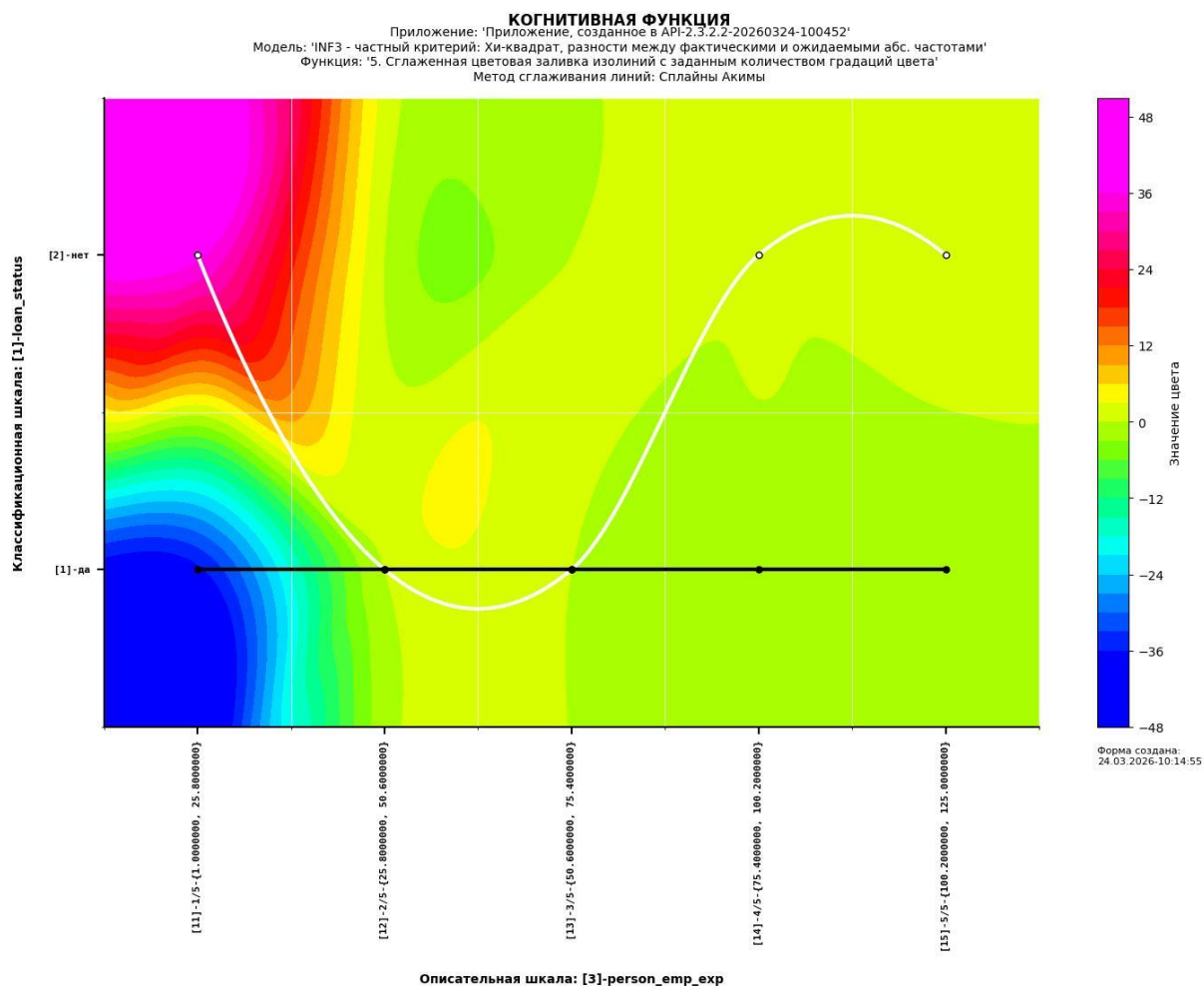
Литератур.ссылки на работы по когнитивным функциям

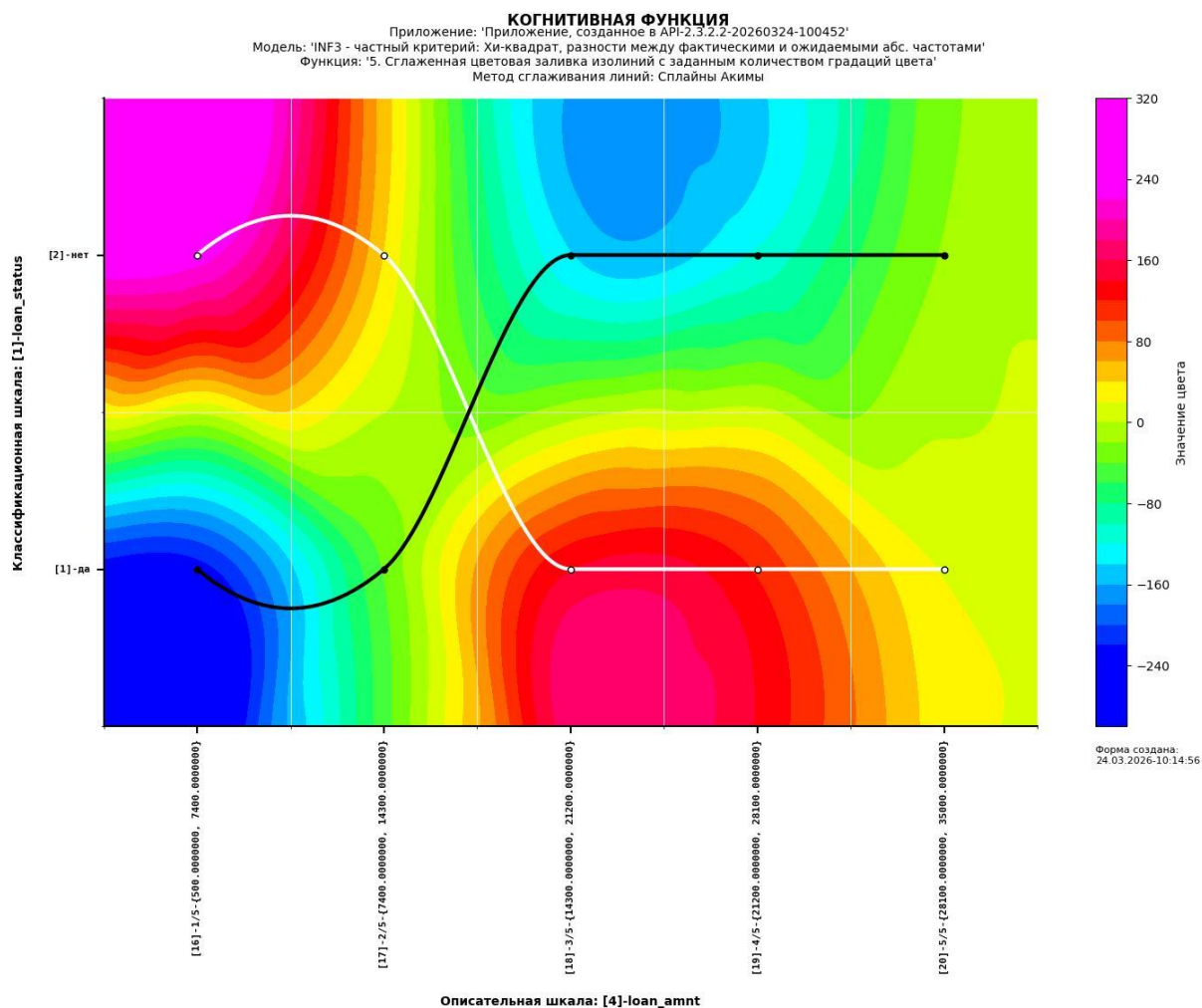
Литератур.ссылки на работы по когнитивным фнкциям

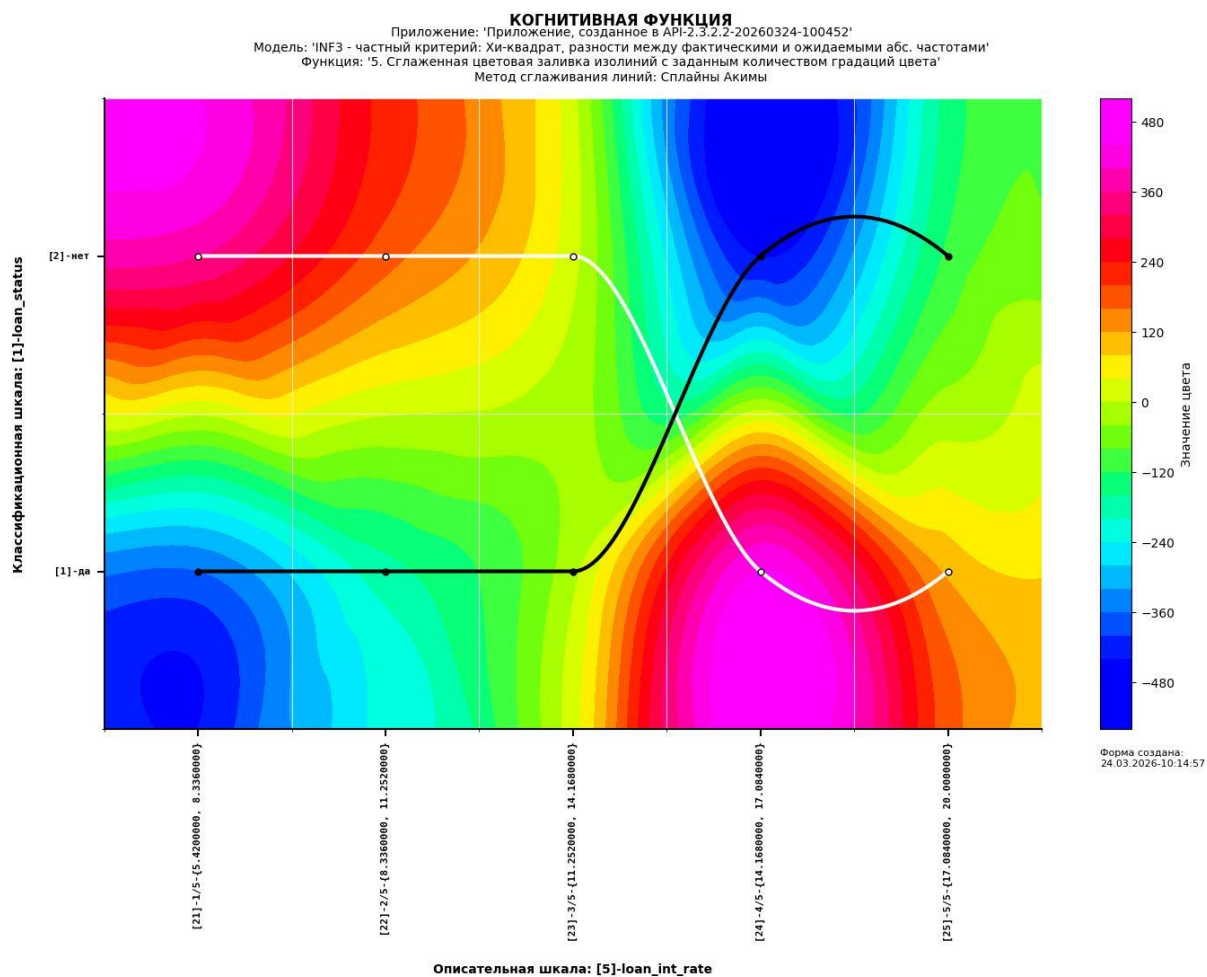
Литератур.ссылки на работы по управлению знаниями

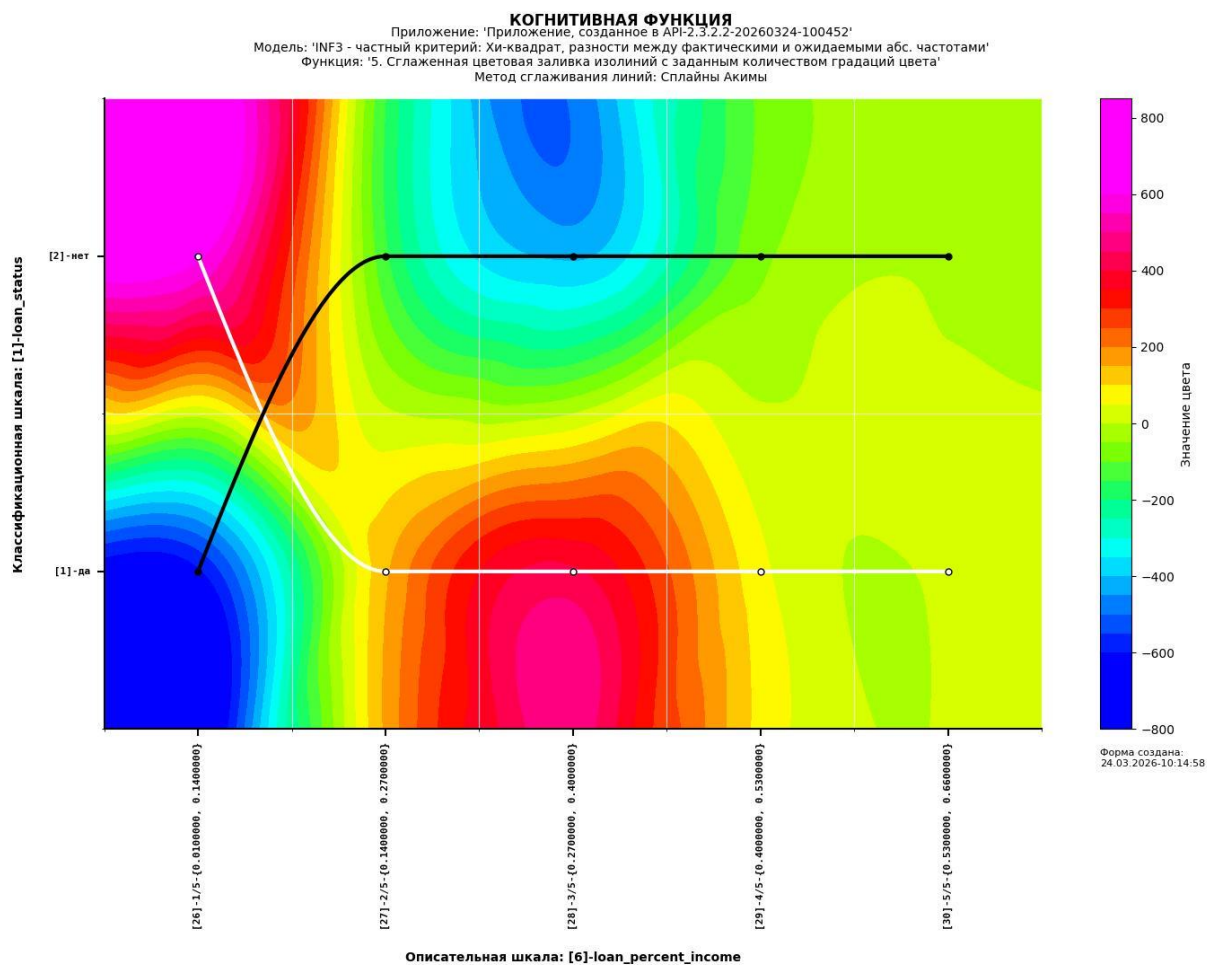


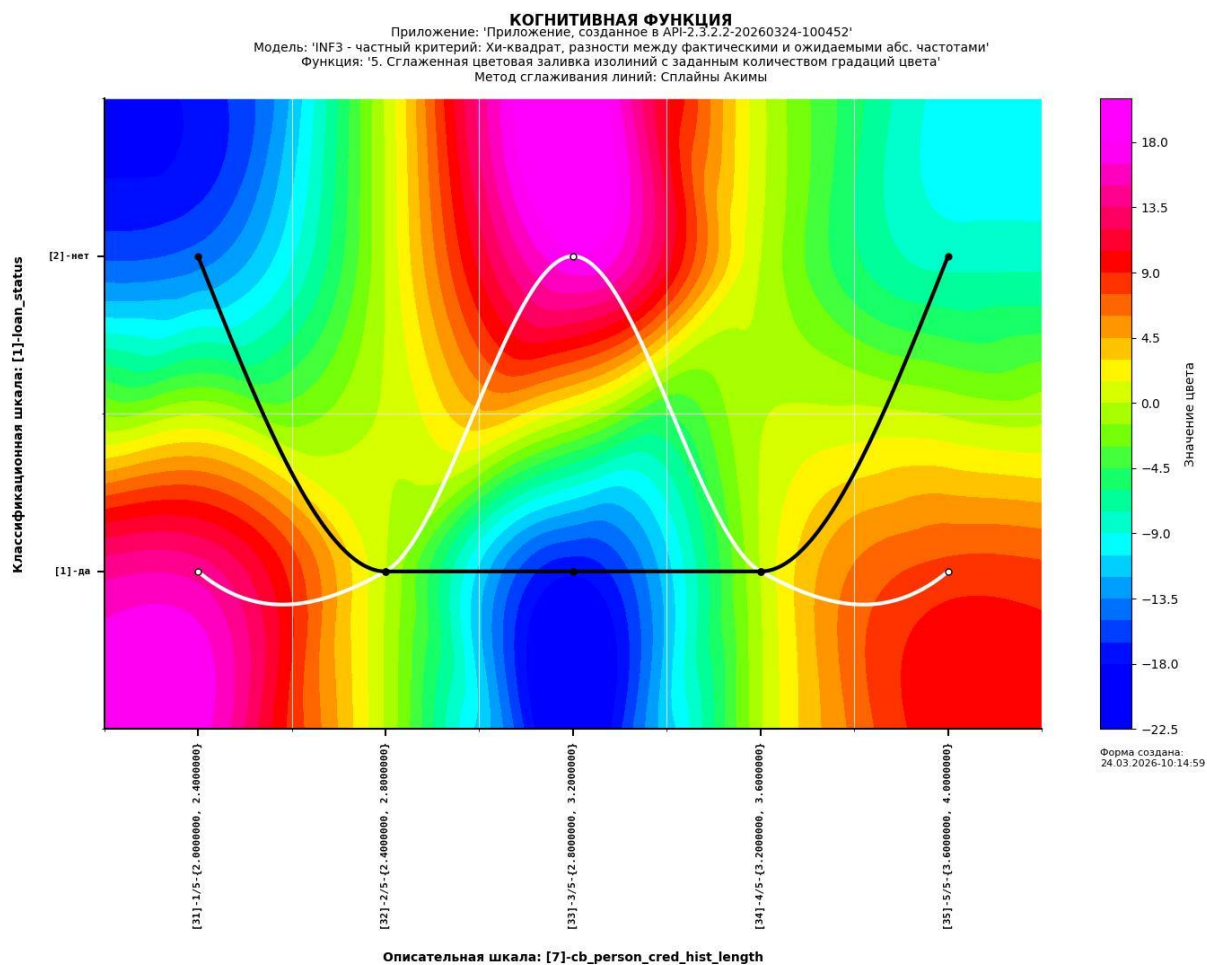












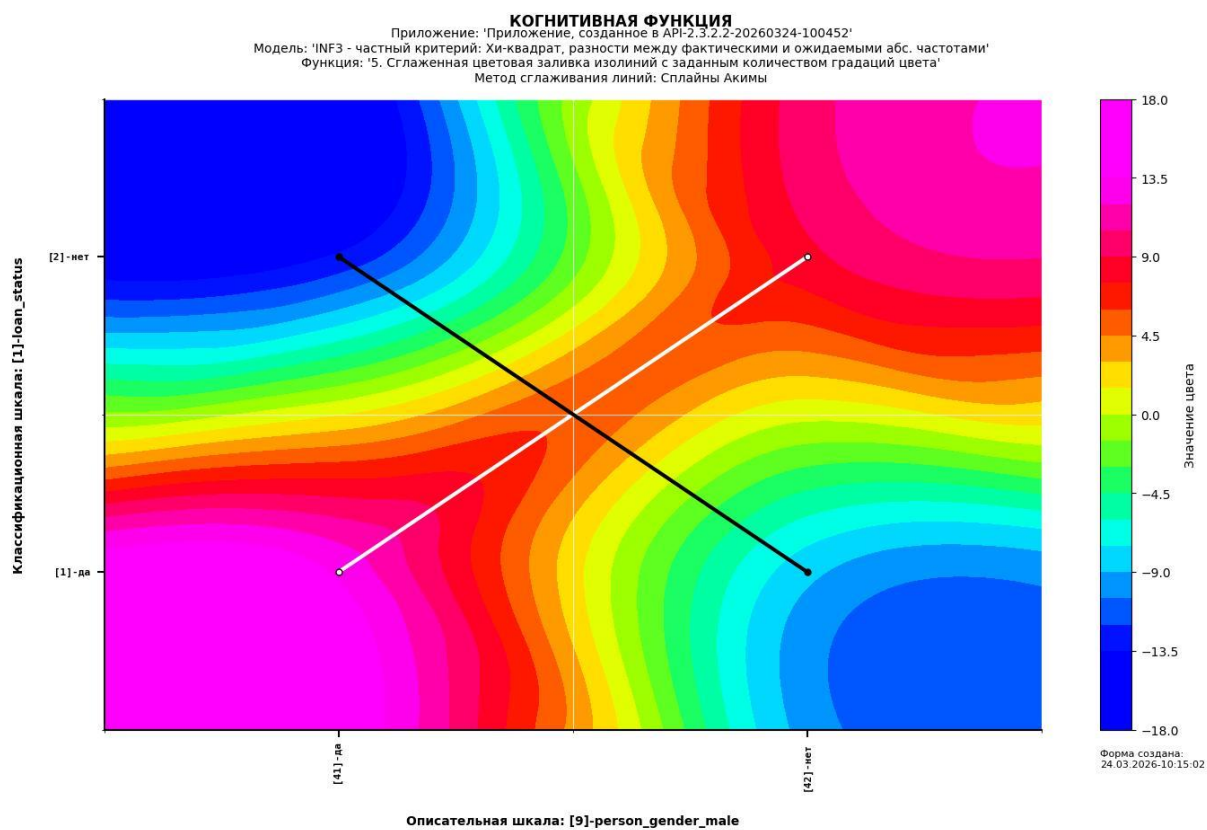
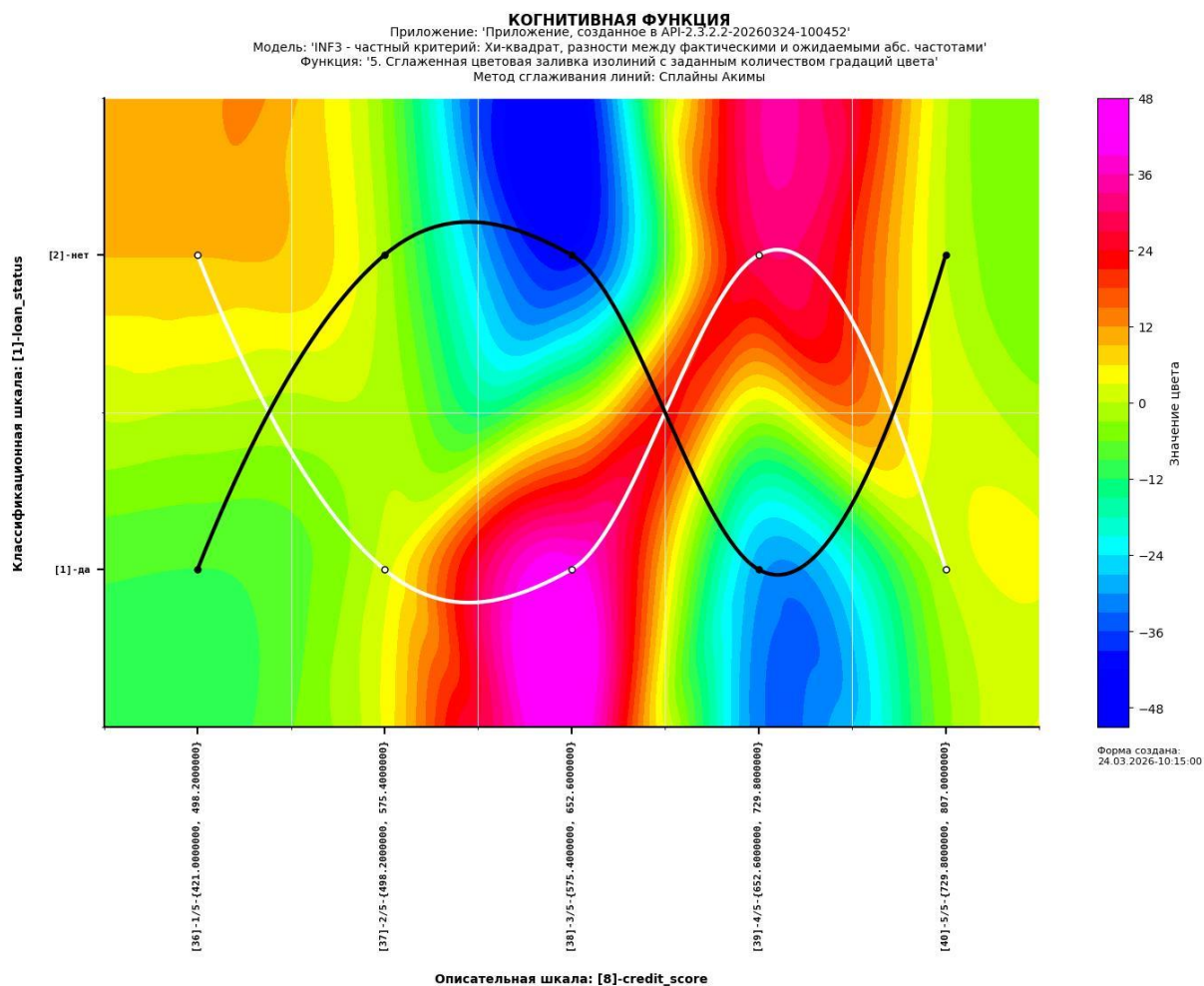


Рисунок 27. Примеры когнитивных функций в СК-модели INF3

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

3.8.10. Значимость описательных шкал и их градаций

3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области

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

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это вариабельность информативностей (режим 3.7.5 системы «Эйдос») (рисунок 37):

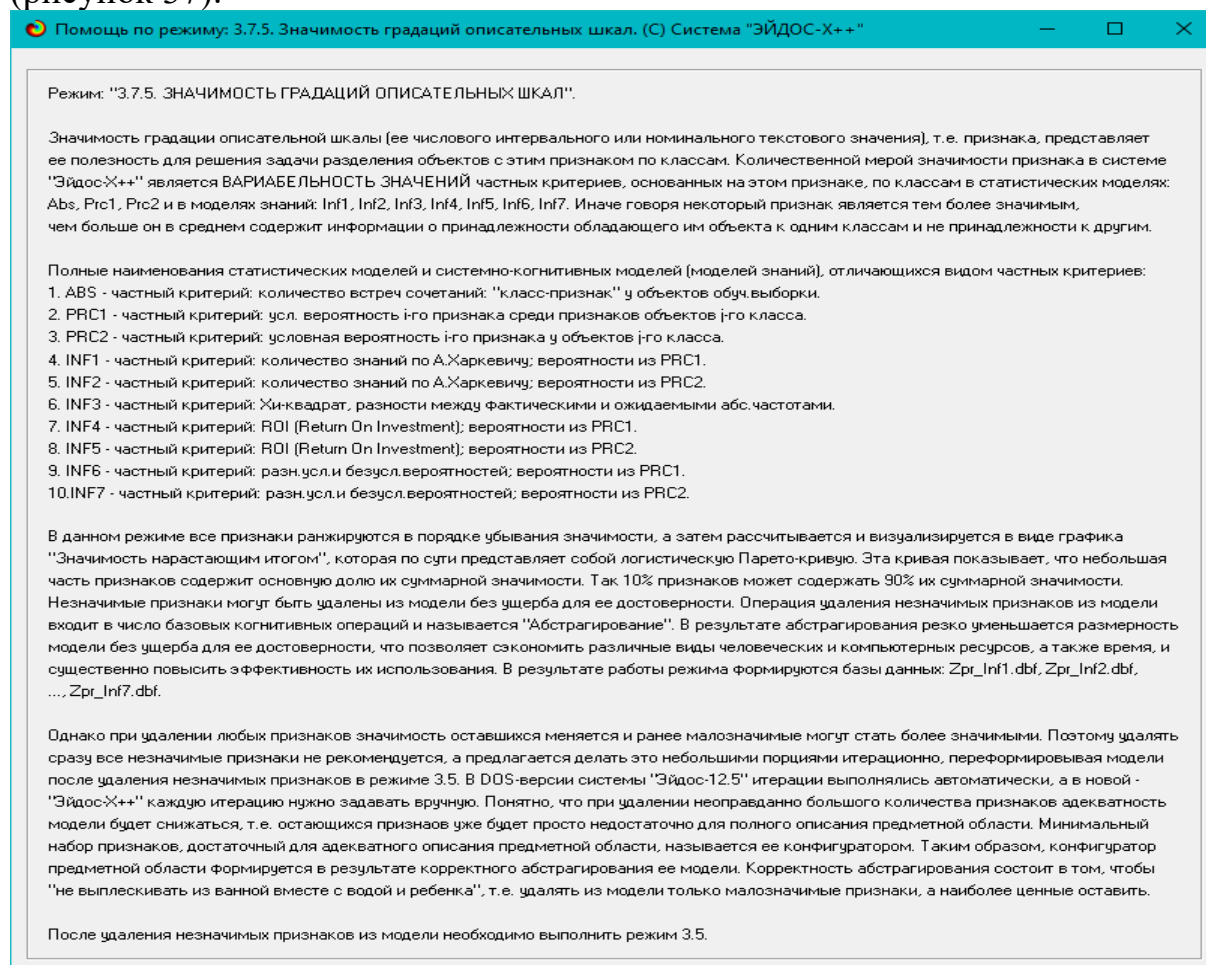


Рисунок 28. Help режима 3.7.5, поясняющий смысл значимости градаций описательных шкал

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.

3.8.10.2. Конкретное решение задачи в данной работе

На рисунке 38 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF1:

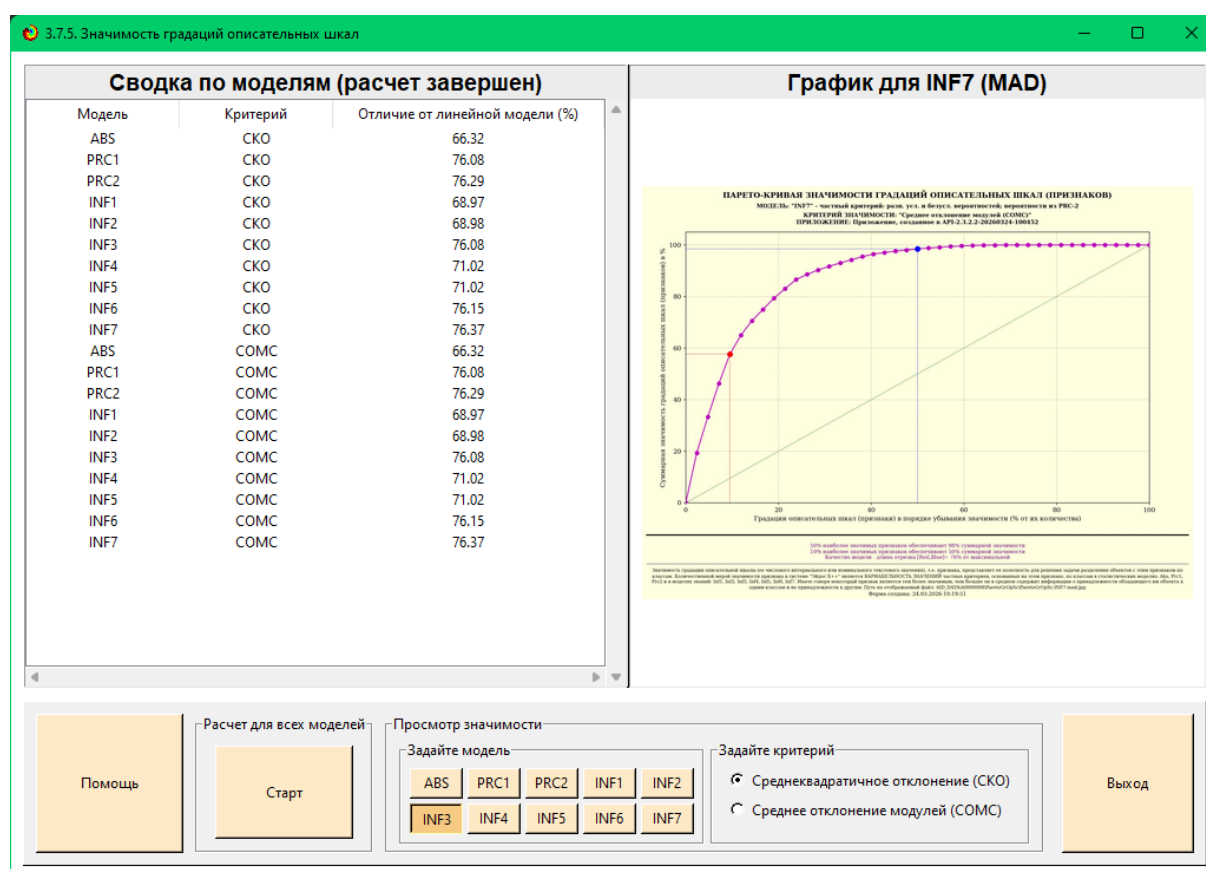


Рисунок 29. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF1

Из рисунка 38 видно, что 43% наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 58% суммарного влияния. На рисунке 39 система «Эйдос» привела рейтинг качества статистических и системно-когнитивных моделей. Кроме того на этом рисунке приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях.

В таблице 16 представлены исходные данные для построения кумулятивной кривой на рисунке 38. Из таблицы 16 видно, какую долю от суммарного влияния на переход объекта моделирования в будущее состояния, соответствующие классам, имеет каждое значение каждого фактора.

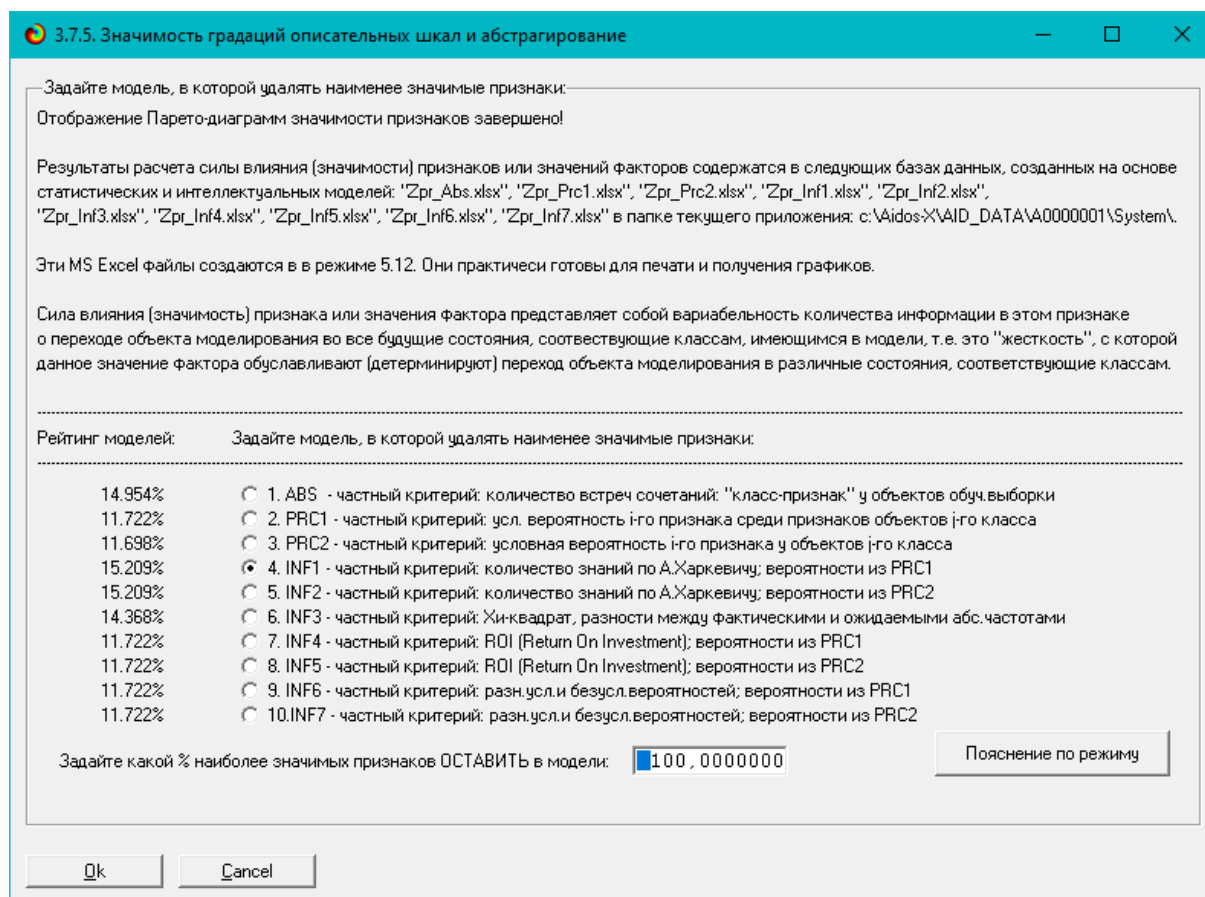


Рисунок 30. Рейтинг качества статистических и системно-когнитивных моделей и имена Excel-файлов с информацией о силе и направлении влияния значений факторов в этих моделях

Рисунок 40 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели INF1

num	num_prc	kod_atr	name_atr	kod_opsc	znach_atr	zn_atrnit	znach_prc	zn_prcnit
1	2,380952	26	loan_perc	6	864,6226	864,6226	19,27838	19,27838
2	4,761905	24	loan_int_r	5	625,481	1490,104	13,94627	33,22465
3	7,142857	28	loan_perc	6	577,6496	2067,753	12,87978	46,10443
4	9,52381	21	loan_int_r	5	510,3839	2578,137	11,37996	57,48439
5	11,90476	16	loan_amn	4	325,8227	2903,96	7,264826	64,74922
6	14,28571	22	loan_int_r	5	249,6175	3153,577	5,56569	70,31491
7	16,66667	27	loan_perc	6	197,5568	3351,134	4,404898	74,7198
8	19,04762	18	loan_amn	4	196,1866	3547,321	4,374348	79,09415
9	21,42857	19	loan_amn	4	165,0963	3712,417	3,681132	82,77528
10	23,80952	25	loan_int_r	5	158,1475	3870,564	3,526194	86,30148
11	26,19048	29	loan_perc	6	91,44237	3962,007	2,038879	88,34036
12	28,57143	6	person_in	2	81,20652	4043,213	1,810651	90,15101
13	30,95238	7	person_in	2	64,75544	4107,969	1,443844	91,59485
14	33,33333	38	credit_scc	8	58,07855	4166,047	1,29497	92,88982
15	35,71429	17	loan_amn	4	54,7315	4220,779	1,220341	94,11016
16	38,09524	11	person_er	3	52,0919	4272,871	1,161486	95,27165
17	40,47619	39	credit_scc	8	40,86426	4313,735	0,911145	96,18279
18	42,85714	20	loan_amn	4	25,91455	4339,65	0,577813	96,76061
19	45,2381	33	cb_persor	7	25,07862	4364,728	0,559175	97,31978
20	47,61905	31	cb_persor	7	20,30622	4385,034	0,452765	97,77255
21	50	41	person_ge	9	18,71792	4403,752	0,417351	98,1899
22	52,38095	23	loan_int_r	5	16,98378	4420,736	0,378685	98,56858
23	54,7619	36	credit_scc	8	12,81955	4433,556	0,285836	98,85442
24	57,14286	42	person_ge	9	12,07456	4445,63	0,269225	99,12364
25	59,52381	35	cb_persor	7	11,41576	4457,046	0,254536	99,37818
26	61,90476	8	person_in	2	9,104443	4466,15	0,203001	99,58118
27	64,28571	1	person_ag	1	8,049919	4474,2	0,179488	99,76067
28	66,66667	30	loan_perc	6	4,968866	4479,169	0,11079	99,87146
29	69,04762	37	credit_scc	8	2,233318	4481,403	0,049796	99,92125
30	71,42857	5	person_ag	1	1,406566	4482,809	0,031362	99,95262
31	73,80952	15	person_er	3	1,054925	4483,864	0,023522	99,97614
32	76,19048	9	person_in	2	0,351642	4484,216	0,007841	99,98398
33	78,57143	14	person_er	3	0,351642	4484,567	0,007841	99,99182
34	80,95238	10	person_in	2	0,351642	4484,919	0,007841	99,99966
35	83,33333	40	credit_scc	8	0,015295	4484,934	0,000341	100
36	85,71429	4	person_ag	1	0	4484,934	0	100

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Zpr_Inf3.xlsx

На экранной форме рисунка 41 приведены имена Excel-файлов с информацией о силе влияния факторов в разных моделях.

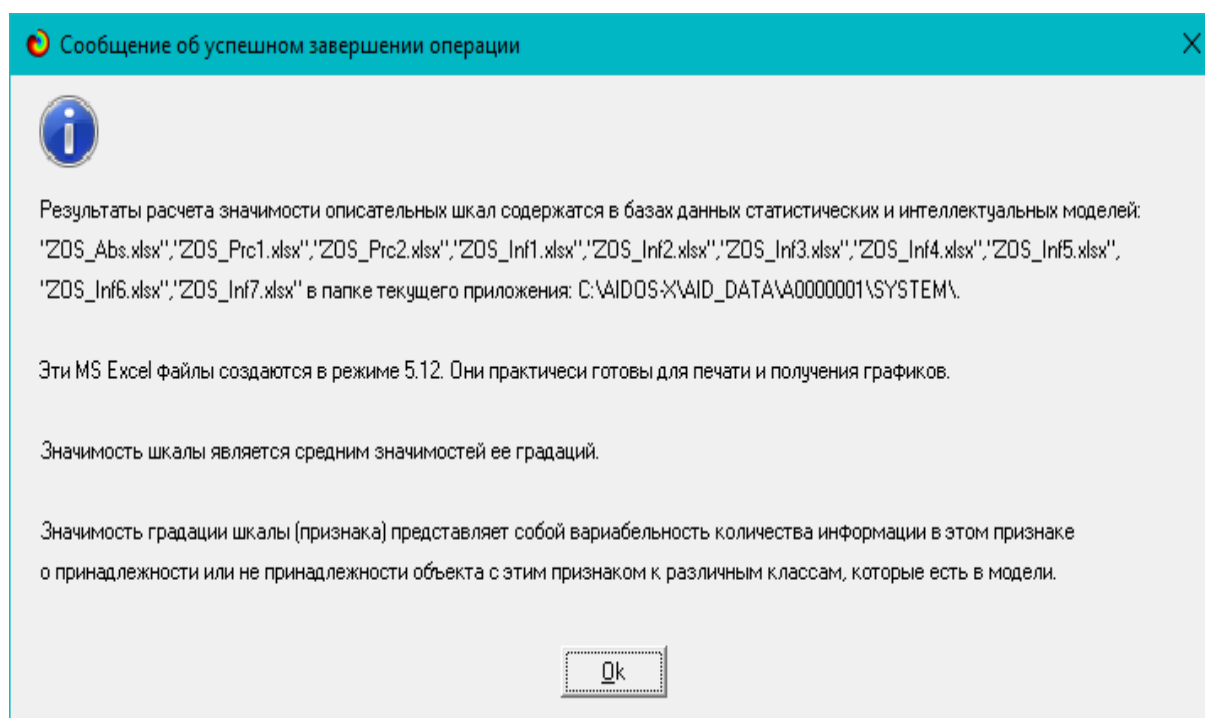


Рисунок 41. Имена Excel-файлов с информацией о силе влияния факторов в статистических и системно-когнитивных моделях

На рисунке 42 приведена информация о силе влияния факторов на переход объекта моделирования в различные будущие состояния, соответствующие классам, в системно-когнитивной модели INF1.

Таблица 42 – Сила влияния факторов на поведение объекта моделирования в системно-когнитивной модели INF1

5.5. Просмотр основных баз данных всех моделей

Задайте модель для просмотра: Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Количество знаков после запятой: 3 (макс: 17) Помощь

Модель: "Inf1"

Код признака	Описательные шкалы и градации (признаки)	1 - loan_status - да	2 - loan_status - нет	Сумма	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	person_age - 1/5-{20.0000000, 44.8000000}	0.000	-0.000	0.000	0.000	0.000	0.000
2	person_age - 2/5-{44.8000000, 69.6000000}						
3	person_age - 3/5-{69.6000000, 94.4000000}						
4	person_age - 4/5-{94.4000000, 119.2000000}						
5	person_age - 5/5-{119.2000000, 144.0000000}		0.025	0.025	0.013	0.018	0.013
6	person_income - 1/5-{12231.0000000, 129963.0000000}	0.002	-0.001	0.001	0.001	0.002	0.001
7	person_income - 2/5-{129963.0000000, 247695.0000000}	-0.068	0.014	-0.054	-0.027	0.058	0.041
8	person_income - 3/5-{247695.0000000, 365427.0000000}	-0.072	0.015	-0.057	-0.029	0.062	0.044
9	person_income - 4/5-{365427.0000000, 483159.0000000}		0.025	0.025	0.013	0.018	0.013
10	person_income - 5/5-{483159.0000000, 600891.0000000}		0.025	0.025	0.013	0.018	0.013
11	person_emp_exp - 1/5-{1.0000000, 25.8000000}	-0.002	0.001	-0.001	-0.001	0.002	0.001
12	person_emp_exp - 2/5-{25.8000000, 50.6000000}						
13	person_emp_exp - 3/5-{50.6000000, 75.4000000}						
14	person_emp_exp - 4/5-{75.4000000, 100.2000000}		0.025	0.025	0.013	0.018	0.013
15	person_emp_exp - 5/5-{100.2000000, 125.0000000}		0.025	0.025	0.013	0.018	0.013
16	loan_amnt - 1/5-{500.0000000, 7400.0000000}	-0.015	0.004	-0.010	-0.005	0.013	0.010
17	loan_amnt - 2/5-{7400.0000000, 14300.0000000}	-0.004	0.001	-0.003	-0.001	0.004	0.003
18	loan_amnt - 3/5-{14300.0000000, 21200.0000000}	0.036	-0.016	0.020	0.010	0.037	0.026
19	loan_amnt - 4/5-{21200.0000000, 28100.0000000}	0.073	-0.049	0.024	0.012	0.086	0.061
20	loan_amnt - 5/5-{28100.0000000, 35000.0000000}	0.074	-0.051	0.023	0.012	0.089	0.063
21	loan_int_rate - 1/5-{5.4200000, 8.3360000}	-0.069	0.014	-0.054	-0.027	0.059	0.042
22	loan_int_rate - 2/5-{8.3360000, 11.2520000}	-0.020	0.006	-0.014	-0.007	0.018	0.013
23	loan_int_rate - 3/5-{11.2520000, 14.1680000}	-0.002	0.001	-0.001	-0.001	0.001	0.001
24	loan_int_rate - 4/5-{14.1680000, 17.0840000}	0.068	-0.043	0.025	0.012	0.079	0.056

Источник: c:\Aidos-X\AID_DATA\A0000001\System\ZOS_Inf1.xlsx

3.8.11. Степень детерминированности классов и классификационных шкал

3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области

Степень детерминированности (обусловленности) класса в системе «Эйдос» количественно оценивается *степенью варибельности значений*

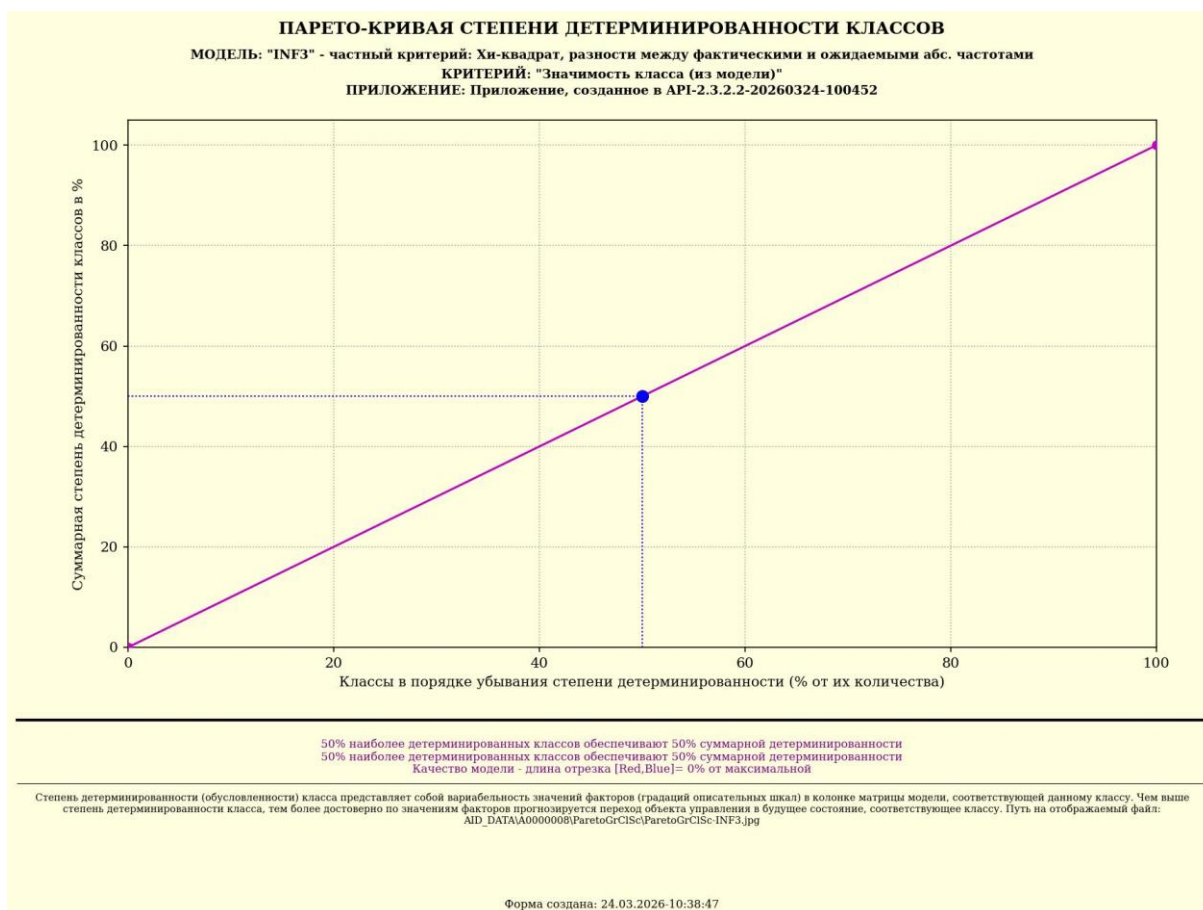
факторов (граций описательных шкал) в колонке матрицы модели, соответствующей данному классу (режим 3.7.3 системы «Эйдос»).

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 граций, т.е. классов (режим 3.7.2 системы «Эйдос»).

3.8.11.2. Конкретное решение задачи в данной работе

На рисунках 42 приведены экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



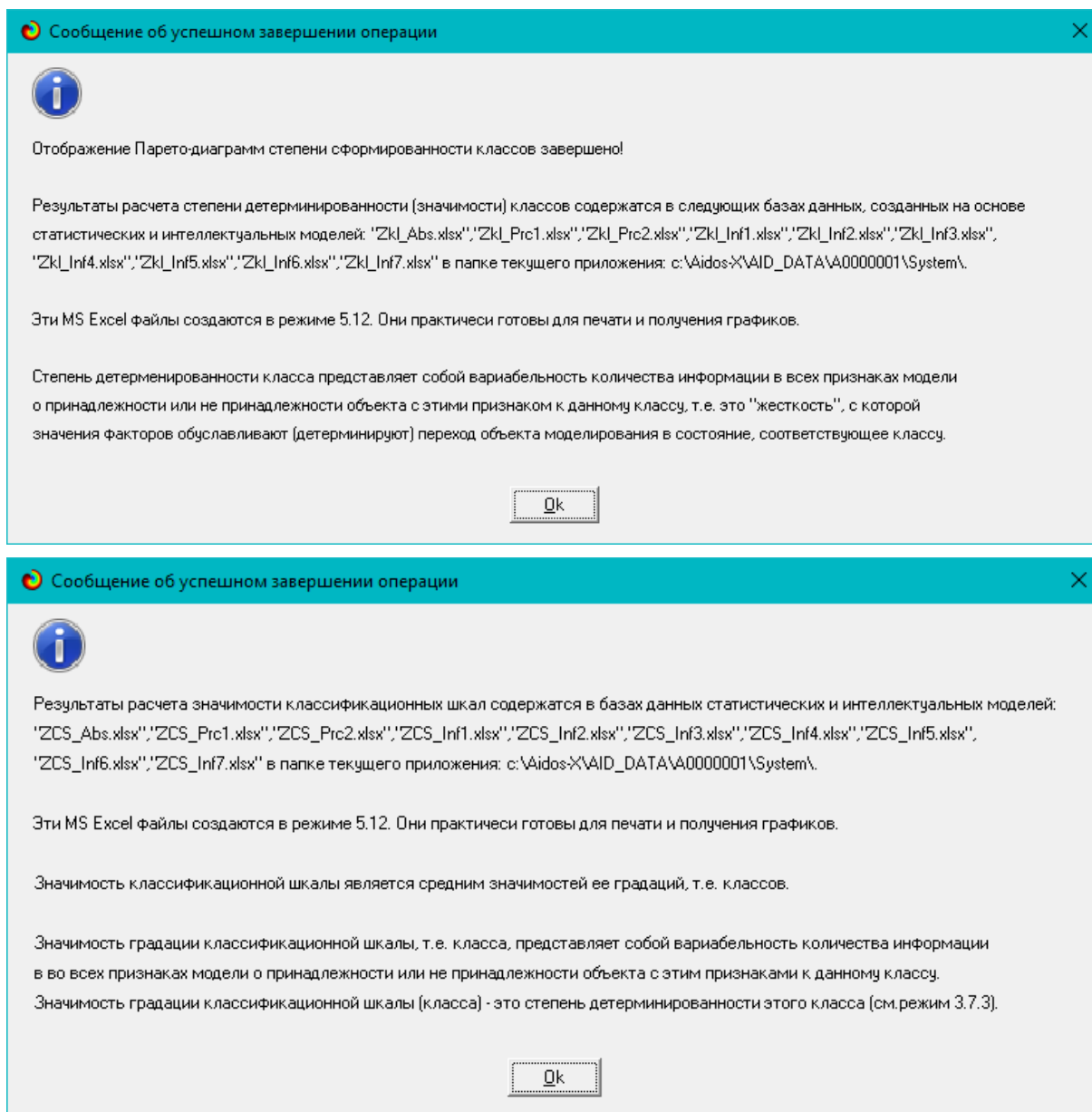


Рисунок 31. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

В таблице 18 представлены исходные данные для построения кумулятивной кривой на рисунке 42.

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

Например, 50% наиболее жестко детерминированных классов суммарно обеспечивают примерно 66% степень детерминированности, а 50% суммарной детерминированности обеспечивают 37% наиболее жестко детерминированных классов.

Степень детерминированности классификационных шкал представлена в таблицах, наименования которых приведены на третьем рисунке 42. Но поскольку в данном приложении одна классификационная шкала, то рейтинг шкал по силе детерминированности состоит из одной шкалы и поэтому эти таблицы здесь не приводятся.

4. DISCUSSION (ОБСУЖДЕНИЕ)

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

Анализ полученных результатов, проведенный в данной работе, полностью согласуется с результатами работы [25], на исходных данных которой они основаны. С другой стороны применение АСК-анализа и системы «Эйдос» весьма существенно расширяет возможности решения задач прогнозирования, принятия решений и исследования моделируемой предметной области, по сравнению с методами, применяемыми в работе [25]. Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований.

Достижением данной работы является:

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

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

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

Например, можно было бы исследовать в создаваемых системно-когнитивных моделях, и технологические, и природно-климатические факторы.

Рекомендуется ввести классификационные шкалы, отражающие влияние исследуемых факторов на объект моделирования не только в натуральном выражении (количество и качество различных видов продукции), но и в стоимостном выражении (прибыль и рентабельность, как общая по предприятию, так и в разрезе по видам продукции).

Перспективность и ценность результатов подобных исследований и разработок для теории и практики не вызывает особых сомнений, что подтверждается работами в этой области [1-49].

У желающих есть все возможности для изучения данной работы и для дальнейших исследований с применением АСК-анализа и системы «Эйдос» на своем компьютере.

Для этого надо скачать систему с сайта разработчика по ссылке на странице: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос-приложение №393. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

В условиях цифровизации банковского сектора и ужесточения требований к управлению кредитными рисками важно повышать точность прогнозирования одобрения кредитов физическим лицам. Однако этому могут препятствовать сложные нелинейные взаимосвязи между характеристиками заемщика и вероятностью дефолта, которые традиционные статистические методы не всегда способны выявить. В ряде научных работ обосновывается перспективность использования методов машинного обучения для повышения качества скоринговых моделей. Для анализа эмпирических данных использованы традиционные методы. Эти исследования показали, что учет таких факторов, как кредитный рейтинг, продолжительность кредитной истории и отношение суммы кредита к доходу, способствует повышению точности прогнозирования, хотя различные категории заемщиков реагируют на эти факторы по-разному. В данной статье для анализа эмпирических данных в области банковского скоринга предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исследований в области кредитного скоринга, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей в данной предметной области.

Спецификой данной задачи является то, что независимые переменные являются как лингвистическими (категориальными) переменными, так и числовыми переменными, измеряемыми в различных единицах измерения. Поэтому для решения данной задачи применяется АСК-анализ, обеспечивающий построение гибридных моделей,

включающих как номинальные (текстовые), так и числовые шкалы, причем в различных единицах измерения.

Сопоставимость обработки данных разных типов, представленных в разных типах шкал и разных единицах измерения обеспечивается путем метризации номинальных шкал, т.е. повышения их степени формализации до уровня числовых шкал [8]. Это достигается путем вычисления количества информации, содержащегося в градациях номинальных шкал и получении той или иной урожайности.

В работе приводится краткое описание АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос».

Работа может быть основой для лабораторных работ и научных исследований по применению систем искусственного интеллекта, в частности лингвистического АСК-анализа для решения задач в области *когнитивной агрономии* [48, 49].

REFERENCES (ЛИТЕРАТУРА)

1. Луценко, Е. В. Исследование влияния подсистем различных уровней иерархии на эмерджентные свойства системы в целом с применением АСК-анализа и интеллектуальной системы "Эйдос" (микроструктура системы как фактор управления ее макросвойствами) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2012. – № 75. – С. 321-363. – EDN OOSCAV.
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHC.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZJXZZ.
4. Работы проф.Е.В.Луценко по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики http://lc.kubagro.ru/aidos/Work_on_emergence.htm
5. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.
6. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
7. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.
8. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.

10. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-Х++" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2014. – № 101. – С. 1367-1409. – EDN SZVWRV.

11. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. – 288 с. – DOI 10.13140/RG.2.2.22981.37608. – EDN ZQLITW.

12. Орлов, А. И. Анализ данных, информации и знаний в системной нечеткой интервальной математике / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2022. – 405 с. – ISBN 978-5-907550-62-9. – DOI 10.13140/RG.2.2.15688.44802. – EDN OQULUW.

13. Луценко, Е. В. Автоматизация функционально-стоимостного анализа и метода "Директ-костинг" на основе АСК-анализа и системы "Эйдос" (автоматизация управления натуральной и финансовой эффективностью затрат без содержательных технологических и финансово-экономических расчетов на основе информационных и когнитивных технологий и теории управления)¹ / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.

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

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

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

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

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

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

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

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

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

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

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

25. Влияние сроков применения борной кислоты на генеративную деятельность косточковых культур / Т. Н. Дорошенко, Л. Г. Рязанова, Н. В. Захарчук, Д. В. Максимцов // Плодоводство и виноградарство Юга России. – 2016. – № 41(5). – С. 121-130. – EDN WKBFHT, <https://journalkubansad.ru/pdf/16/05/12.pdf>

26. Монографии по АСК-анализу: http://lc.kubagro.ru/aidos/_Aidos-X.htm#_Toc128746370

27. Некоторые учебники и учебные пособия проф.Е.В.Луценко: http://lc.kubagro.ru/aidos/_Aidos-X.htm#_Toc128746372.

28. Свидетельства Роспатента на систему «Эйдос» и ее подсистемы: http://lc.kubagro.ru/aidos/_Aidos-X.htm#_Toc128746371.

29. Тематические подборки публикаций по применению АСК-анализа и системы «Эйдос» в различных предметных областях: http://lc.kubagro.ru/aidos/_Aidos-X.htm

30. Работы по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики: http://lc.kubagro.ru/aidos/Work_on_emergence.htm .

31. Работы по АСК-анализу изображений: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_images.htm

32. Работы по АСК-анализу текстов: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_texts.htm

33. Работы по когнитивным функциям: http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm

34. Работы по выявлению, представлению и использованию знаний, логике и методологии научного познания: http://lc.kubagro.ru/aidos/Work_on_identification_presentation_and_use_of_knowledge.htm

35. Работы по экологии, климатологии и изучению влияния космической среды на различные глобальные процессы на Земле: http://lc.kubagro.ru/aidos/Work_on_the_study_of_the_influence_of_the_space_environment_on_various_processes_on_Earth.htm

36. Работы по современным информационно-коммуникационным технологиям в научно-исследовательской деятельности и образовании: http://lc.kubagro.ru/aidos/Information_and_communication_technologies_in_research_activities_and_education.htm

37. Работы по виртуальной реальности:
http://lc.kubagro.ru/aidos/Virtual_reality_publications.htm
38. Работы по когнитивной ветеринарии:
http://lc.kubagro.ru/aidos/Publications_on_cognitive_veterinary_medicine.htm
39. Работы по когнитивной агрономии и когнитивной ампелографии:
http://lc.kubagro.ru/aidos/Works_on_cognitive_agronomy.htm
40. Работы по тематике, связанной с АПК:
http://lc.kubagro.ru/aidos/Work_with_agricultural.htm
41. Работы по наукометрии: http://lc.kubagro.ru/aidos/Works_on_scientometrics.htm
42. Работы о высших формах сознания, перспективах человека, технологии и общества: http://lc.kubagro.ru/aidos/Works_on_higher_forms_of_consciousness.htm
43. Работы по разработке и применению профессиограмм и тестов (психологических, профориентационных, медицинских и ветеринарных):
http://lc.kubagro.ru/aidos/Work_on_the_development_and_application_tests.htm
44. Работы по сценарному автоматизированному системно-когнитивному анализу (сценарный АСК-анализ): http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm
45. MVP-проект «Внедрение технологий АСК-анализа и системы «Эйдос» для решения задач АПК»: <http://lc.kubagro.ru/aidos/MVP-projects.htm>
46. Кратко об АСК-анализе и системе «Эйдос»:
http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf
47. Ссылки на видео-занятия и проф.Е.В.Луценко в Пермском национальном университете: <https://bigbluebutton.pstu.ru/b/w3y-2ir-ukd-bqn> (2021), <https://bigbluebutton.pstu.ru/b/3kc-n8a-gon-tjz> (2022), в Кубанском государственном университете и Кубанском государственном аграрном университете: <https://disk.yandex.ru/d/knISAD5qzV83Ng?w=1>
48. Луценко, Е. В. Революция начала XXI века в искусственном интеллекте: глубинные механизмы и перспективы / Е. В. Луценко, Н. С. Головин. – Краснодар: Кубанский государственный аграрный университет им. И.Т. Трубилина, 2024. – 394 с. – DOI 10.13140/RG.2.2.17056.56321. – EDN OMIPIL.
49. Луценко Е.В. Системы искусственного интеллекта как системы автоматизации процесса научного познания и удвоение номенклатуры научных специальностей путем применения этих систем для исследований в различных направлениях науки / Е.В. Луценко, Н.С. Головин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2024. – №01(195). С. 74 – 111. – IDA [article ID]: 1952401009. – Режим доступа: <http://ej.kubagro.ru/2024/01/pdf/09.pdf>, 2,375 у.п.л.