

РЕФЕРАТ

Курсовая работа содержит: 73 страниц, 44 рисунка, 5 таблиц, 8 литературных источников.

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

Целью работы является проведение автоматизированного системно-когнитивного анализа кредитных историй физлиц.

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

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ КРЕДИТНЫХ ИСТОРИЙ ФИЗ.ЛИЦ

Малая Полина Павловна
студентка факультета заочного обучения, группы ИТз2241
polina-malaya-00@bk.ru

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

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

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

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF CREDIT HISTORIES OF INDIVIDUALS

Malaya Polina Pavlovna
student of faculty of Correspondence Studies, group ITz2241
polina-malaya-00@bk.ru

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

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

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

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 ПРЕДМЕТНАЯ ОБЛАСТЬ.....	6
1.1 ОПИСАНИЕ ИСПОЛЬЗУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ.....	6
1.2 ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ.....	7
1.3 ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	7
1.4 ЦЕЛЬ И ЗАДАЧИ РАБОТЫ.....	7
2 МЕТОДЫ	10
2.1 ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ.....	10
2.2 АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ)	12
2.3 СИСТЕМА «ЭЙДОС» - ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	14
3 РЕЗУЛЬТАТЫ	20
3.1 Задача – 1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ.....	20
3.2 Задача – 2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	21
3.3 Задача – 3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	26
3.4 Задача – 4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ.....	37
3.5 Задача – 5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	40
3.6 Задача – 6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	42
3.6.1 Интегральный критерий «Сумма знаний».....	44
3.6.2 Интегральный критерий «Семантический резонанс знаний»	45
3.6.3 Важные математические свойства интегральных критериев	46
3.7 Задача – 7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ.....	47
3.7.1 Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	47
3.7.2 Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	48
3.8 Задача – 8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	48
3.8.1 Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы).....	48
3.8.2 Кластерно-конструктивный анализ классов.....	50
3.8.3 Кластерно-конструктивный анализ значений описательных шкал.....	53
3.8.4 Модель знаний системы «Эйдос» и нелокальные нейроны	55
3.8.5 Нелокальная нейронная сеть.....	58
3.8.6 3D-интегральные когнитивные карты	59
3.8.7 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	61
3.8.8 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечетки правдоподобные рассуждения).....	63
3.8.9 Когнитивные функции	65
3.8.10 Значимость описательных шкал и градаций.....	67
3.8.11 Степень детерминированности классов и классификационных шкал	69
4 ОБСУЖДЕНИЕ.....	69
ЗАКЛЮЧЕНИЕ	71
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	72

ВВЕДЕНИЕ

В настоящей работе представлен автоматизированный системно-когнитивный анализ кредитных историй с использованием системы «Эйдос». Цель исследования заключается в проведении анализа кредитных историй физических лиц с целью определения факторов, влияющих на принятие решения по выдаче кредита. Для достижения поставленной цели предполагается провести анализ методов формирования обобщенных образов классов, а также разработать решения для задач идентификации конкретных объектов с учетом классов принятия решений.

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

В ходе исследования будет уделено внимание анализу факторов, влияющих на решение по кредитной заявке. Предполагается, что полученные результаты могут иметь практическое применение при дальнейшем принятии решения относительно конкретной кредитной заявки.

1 ПРЕДМЕТНАЯ ОБЛАСТЬ

1.1 Описание используемой предметной области

Под анализом кредитоспособности заемщика понимается оценка банком возможности и целесообразности предоставления заемщику кредита, определение вероятности их своевременного возврата в соответствии с заключаемым кредитным договором.

Оценка кредитоспособности физического лица основывается на соотношении запрашиваемой ссуды и его личного дохода, общей оценке финансового положения и имущества, составе семьи, личностных характеристиках, изучении кредитной истории заемщика. Каждый банк по-своему определяет понятие идеального заемщика, а значит, и параметры оценки у банков могут различаться. Кредитоспособность клиента выражается в его желании и возможности платить за кредит, выражаемые в простой аббревиатуре WAS, где W – willness (желание), A – ability (возможность), S – stability (стабильность).

На сегодняшний день банки используют в своей практической деятельности различные методики оценки кредитоспособности заемщиков, среди которых можно выделить:

- системы оценки кредитоспособности клиентов, основанные на расчете платежеспособности заемщиков, исходя из среднемесячного дохода клиента за последние полгода за вычетом всех обязательных платежей;
- балльные системы оценки кредитоспособности клиентов.

При оценке кредитоспособности заемщиков, исходя из среднемесячного дохода документами, запрашиваемыми банком для подтверждения величины доходов и размера производимых удержаний заемщика и его поручителей, являются: для работающих – справка от предприятия, на котором работает заемщик или его поручитель; для пенсионеров – пенсионное удостоверение и справка из госорганов социальной защиты населения.

1.2 Объект и предмет исследования

Объект исследования – выявление зависимостей вынесения решения по кредитной заявке от различных факторов.

Предмет исследования – выявление зависимостей одобрения кредитов, используя данные о кредитных историях.

1.3 Проблема, решаемая в работе и ее актуальность

Актуальность исследования заключается в возможности кредитных организаций выявить факторы, способствующие и препятствующие одобрению кредитной заявки.

Настоящая работа нацелена на автоматизированный системно-когнитивный анализ кредитных историй физлиц. Поставленная цель предполагает проведение методологического обоснования исследования, охватывающего анализ методов формирования обобщенных образов классов в контексте данной предметной области.

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

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

1.4 Цель и задачи работы

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

Целью данной курсовой работы является, составление АСК-анализа о зависимости кредитных историй физлиц на основе некоторых факторов.

Поскольку для решения поставленной проблемы используется автоматизированный системно-когнитивный анализ (АСК-анализ) и его

программный инструментарий, в качестве которого в настоящее время выступает интеллектуальная система «Эйдос», то достижение поставленной цели обеспечивается решением следующих задач и подзадач, которые являются этапами достижения цели:

1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций;

2. Формализация предметной области;

3. Синтез статистических и системно-когнитивных моделей.

Многопараметрическая типизация и частные критерии знаний;

4. Верификация моделей;

5. Выбор наиболее достоверной модели;

6. Системная идентификация и прогнозирование;

7. Поддержка принятия решений (Упрощенный вариант принятия решений как обратная задача прогнозирования; Позитивный и негативный информационные портреты классов; SWOT-анализ; Развитый алгоритм принятия решений в АСК-анализе);

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

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

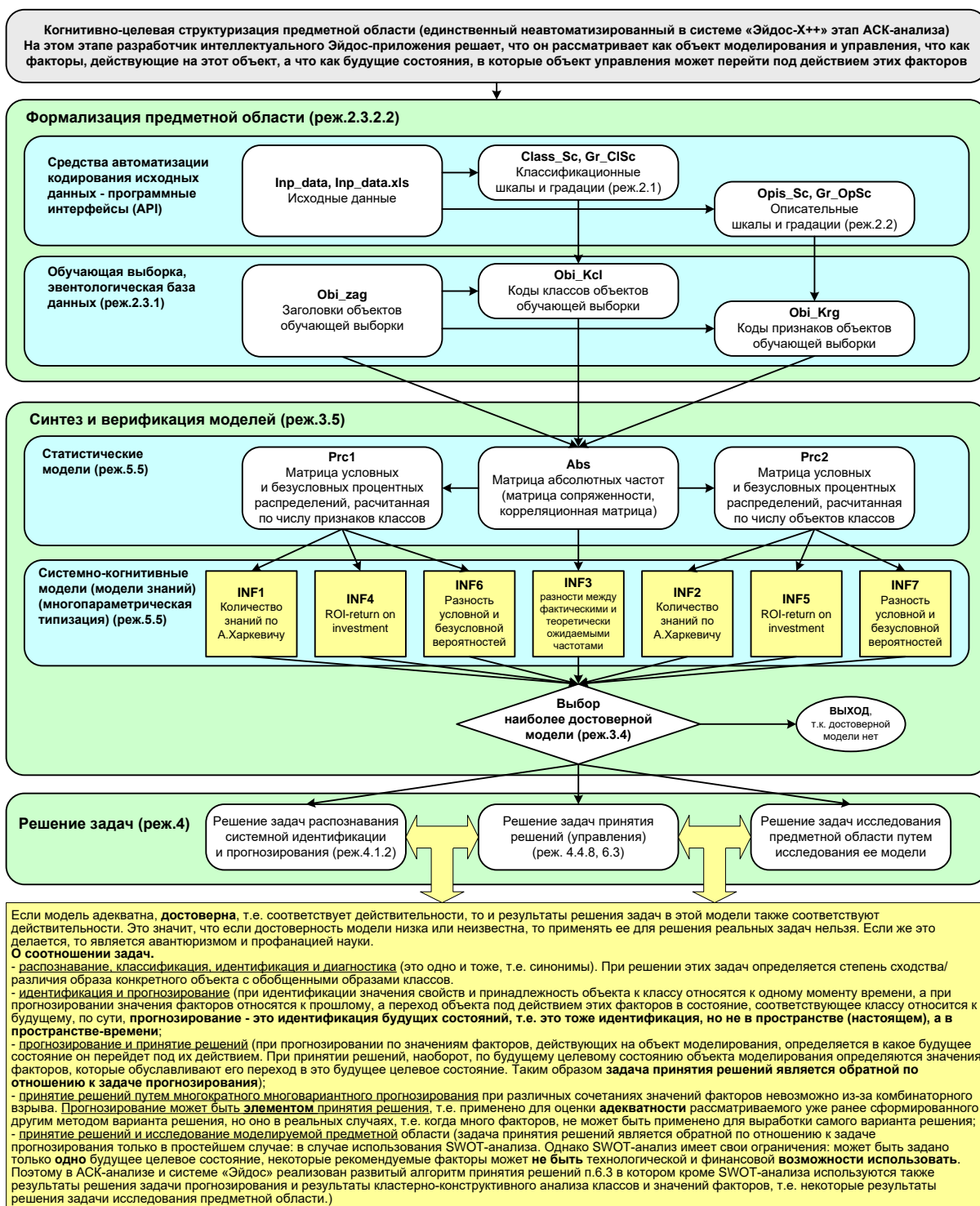


Рисунок 1 – Последовательность преобразования исходных данных в информацию и знаний в системе «Эйдос»

2 МЕТОДЫ

2.1 Обоснование требований к методу решения проблемы

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

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

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

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

— является одной из первых отечественных систем искусственного интеллекта персонального уровня, т. е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта: «имеет нулевой порог входа»;

— обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных

зашумленных взаимосвязанных (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

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

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

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

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

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

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

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

2.2 Автоматизированный системно-когнитивный анализ (АСК-анализ)

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

Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф. Е.В.Луценко. На тот момент он вообще не встречался в Internet. Сегодня по соответствующему запросу в Яндексе находится 19 тысяч сайтов с этим сочетанием слов.

АСК-анализ включает:

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

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

Около половины из опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано более 45 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, 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.

Три монографии включены в фонды библиотеки конгресса США.

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

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

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

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

- автоматизированный системно-когнитивный анализ текстовых данных;
- автоматизированный системно-когнитивный анализ текстовых данных;
- спектральный и контурный автоматизированный системно-когнитивный анализ изображений;
- сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов.

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

2.3 Система «Эйдос» - инструментарий АСК-анализа

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

- является универсальной и может быть применена во многих предметных областях, т.к. разработана в универсальной постановке, не зависящей от предметной области (<http://lc.kubagro.ru/aidos/index.htm>). Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени при решении задач идентификации, прогнозирования, принятия решений и исследования предметной области (автоматические системы работают без такого участия человека);

– находится в полном открытом бесплатном доступе (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), причем с актуальными исходными текстами (<http://lc.kubagro.ru/AidosALL.txt>): открытая лицензия: CC BY-SA 4.0, и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора системы «Эйдос» проф. Е.В. Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 31 свидетельство РосПатента РФ);

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

– реально работает, обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

– имеет «нулевой порог входа», содержит большое количество локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений

(http://aidos.byethost5.com/Source_data_applications/WebAppls.htm,
http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

– поддерживает онлайн среду накопления знаний и обмена ими, широко используется во всем мире (<http://aidos.byethost5.com/map5.php>);

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

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

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

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

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

содержательного эмпирического исследования моделируемой предметной области.

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

В разработке системы «Эйдос» были следующие этапы:

1-й этап, «подготовительный»: 1979-1992 годы. Математическая модель системы "Эйдос" разработана в 1979 и впервые прошла экспериментальную апробацию в 1981 году (первый расчет на компьютере на основе модели). С 1981 по 1992 система "Эйдос" неоднократно реализовалась на платформе Wang (на компьютерах Wang-2200C). В 1987 году впервые получен акт внедрения на одну из ранних версий системы «Эйдос», реализованную в среде персональной технологической системы «Вега-М» разработки автора (см. 2-й акт по ссылке: <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>).

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBM-совместимых персональных компьютеров система "Эйдос" впервые реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены свидетельства РосПатента, первые в Краснодарском крае и, возможно, в России на системы искусственного интеллекта. С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

3-й этап, «эра MS Windows xp, 8, 7»: 2012-2020 годы. С июня 2012 по 14.12.2020 система «Эйдос» развивалась на языке Аляска-1.9 + Экспресс++ + библиотека для работы с Internet xb2net. Система «Эйдос-X1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала

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

4-й этап, «эра MS Windows-10»: 2020-2021 годы. С 13.12.2020 года по настоящее время система «Эйдос» развивается на языке Аляска-2.0 + Экспресс++. Библиотека xb2net в ней больше не используется, т.к. все возможности работы с Internet входят в базовые возможности языка программирования.

5-й этап, «эра Больших данных, информации и знаний»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке Аляска+Экспресс, обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge) с использованием ADS (Advantage Database Server), а также на языке C# (Visual Studio | C#).

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



Рисунок 2 – Титульная видеोगрамма DOS-версии «Эйдос»

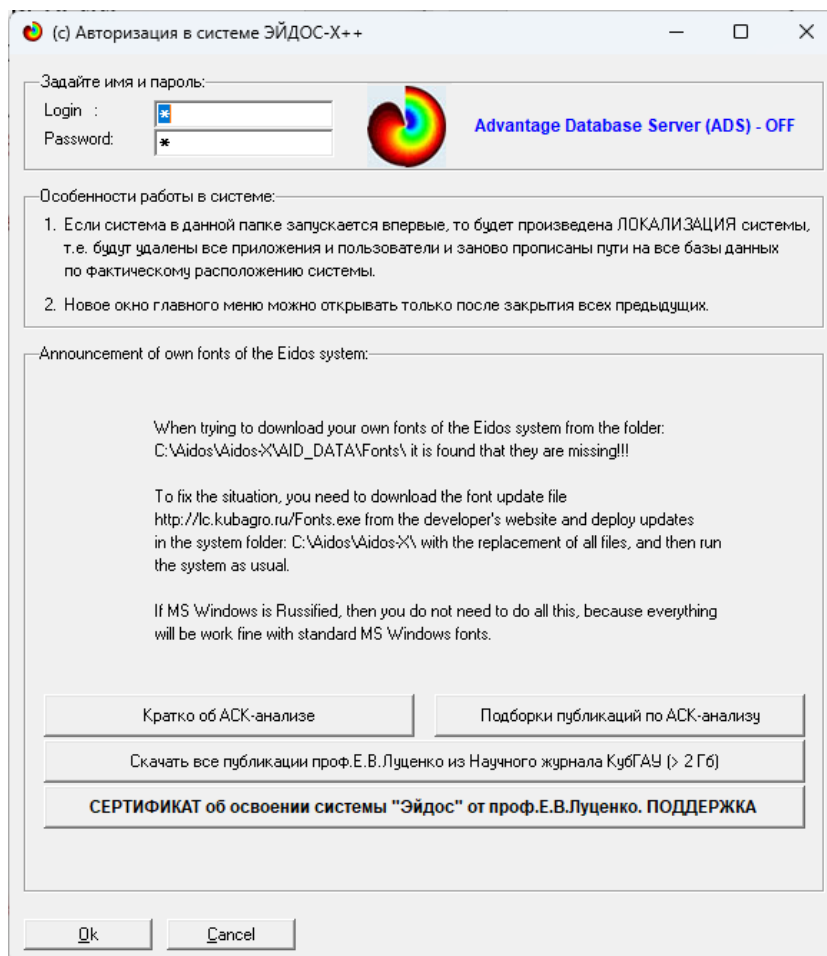


Рисунок 3 – Титульная видеोगрамма текущей версии системы «Эйдос»

3 РЕЗУЛЬТАТЫ

3.1 Задача – 1. Когнитивная структуризация предметной области.

Две интерпретации классификационных и описательных шкал и градаций

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

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

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

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

Статичная интерпретация и терминология:

– градации классификационных шкал – это обобщающие категории видов объектов (классы);

– описательные шкалы – свойства объектов, градации описательных шкал – значения свойств (признаки) объектов.

Динамичная интерпретация и терминология:

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

– описательные шкалы – факторы, действующие на объект моделирования, градации описательных шкал – значения факторов, действующие на объект моделирования.

Обобщающая терминология:

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

3.2 Задача – 2. Формализация предметной области

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

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

В качестве источника данных с электронного ресурса kaggle.com был взят набор данных «Loan Rejection or Approval Status Prediction»,

<https://www.kaggle.com/datasets/bsugiarto9/loan-status-prediction-with-added-nans>. Набор данных содержит 13 столбцов, рассмотрим их:

- Loan_ID – ID заемщика;
- Gender – пол;
- Married – семейное положение;
- Dependents – количество иждивенцев;
- Education – образование;
- Self_Employed – самозанятость;
- ApplicantIncome – заявленный доход заемщика;
- CoapplicantIncome – доход созаемщика (при наличии);
- LoanAmount – сумма кредита;
- Loan_Amount_Term – срок кредита;
- Credit_History – соответствие кредитной истории требованиям;
- Property_Area – область проживания;
- Loan_Status – одобрен кредит или нет.

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
1	LP001003	Male	Yes	1	Graduate	No	4583	1508	128	360	Yes	Rural	N
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0	66	360	Yes	Urban	Y
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358	120	360	Yes	Urban	Y
4	LP001008	Male	No	0	Graduate	No	6000	0	141	360	Yes	Urban	Y
5	LP001013	Male	Yes	0	Not Graduate	No	2333	1516	95	360	Yes	Urban	Y
6	LP001024	Male	Yes	2	Graduate	No	3200	700	70	360	Yes	Urban	Y
7	LP001027	Male	Yes	2	Graduate	No	2500	1840	109	360	Yes	Urban	Y
8	LP001029	Male	No	0	Graduate	No	1851	2840	114	360	Yes	Rural	N
9	LP001030	Male	Yes	2	Graduate	No	1299	1086	17	120	Yes	Urban	Y
10	LP001032	Male	No	0	Graduate	No	4950	0	125	360	Yes	Urban	Y
11	LP001034	Male	No	1	Not Graduate	No	3596	0	100	240	No	Urban	Y
12	LP001036	Female	No	0	Graduate	No	3510	0	76	360	No	Urban	N
13	LP001038	Male	Yes	0	Not Graduate	No	4887	0	133	360	Yes	Rural	N
14	LP001041	Male	Yes	0	Graduate	No	2600	3500	115	360	Yes	Urban	Y
15	LP001043	Male	Yes	0	Not Graduate	No	7660	0	104	360	No	Urban	N
16	LP001047	Male	Yes	0	Not Graduate	No	2600	1911	116	360	No	Semiurban	N
17	LP001050	Male	Yes	2	Not Graduate	No	3365	1917	112	360	No	Rural	N
18	LP001068	Male	Yes	0	Graduate	No	2799	2253	122	360	Yes	Semiurban	Y
19	LP001073	Male	Yes	2	Not Graduate	No	4226	1040	50	360	Yes	Urban	Y
20	LP001086	Male	No	0	Not Graduate	No	1442	0	35	360	Yes	Urban	N
21	LP001087	Female	No	2	Graduate	No	3750	2083	120	360	Yes	Semiurban	Y
22	LP001095	Male	No	0	Graduate	No	3167	0	74	360	Yes	Urban	N
23	LP001097	Male	No	1	Graduate	Yes	4692	0	106	360	Yes	Rural	N
24	LP001098	Male	Yes	0	Graduate	No	3500	1667	114	360	Yes	Semiurban	Y
25	LP001109	Male	Yes	0	Graduate	No	1828	1330	100	360	No	Urban	N
26	LP001112	Female	Yes	0	Graduate	No	3667	1459	144	360	Yes	Semiurban	Y
27	LP001116	Male	No	0	Not Graduate	No	3748	1668	75	360	Yes	Semiurban	Y
28	LP001119	Male	No	0	Graduate	No	3600	0	80	360	Yes	Urban	N
29	LP001120	Male	No	0	Graduate	No	1800	1213	47	360	Yes	Urban	Y
30	LP001123	Male	Yes	0	Graduate	No	2400	0	75	360	No	Urban	Y
31	LP001131	Male	Yes	0	Graduate	No	3941	2336	134	360	Yes	Semiurban	Y

Рисунок 4 – Фрагмент обучающей выборки

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

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

Для импорта обучающей выборки в систему AIDOS-X необходимо скопировать ее в папку Inp_data и переименовать в Inp_data.xlsx, после этого можно запустить саму программу и универсальный программный интерфейс импорта данных в систему (режим 2.3.2.2) (рисунке 5).



Рисунок 5 – Интерфейс импорта данных

Следует выделить следующие настройки:

- Тип файла – xlsx;
- Классификационная шкала – 13;
- Описательные шкалы – 2-12.

После импорта данных система просит задать размерности модели системы, меняем предложенную размерность на 5 (рисунок 6).



Рисунок 6 – Задание размерностей системы

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

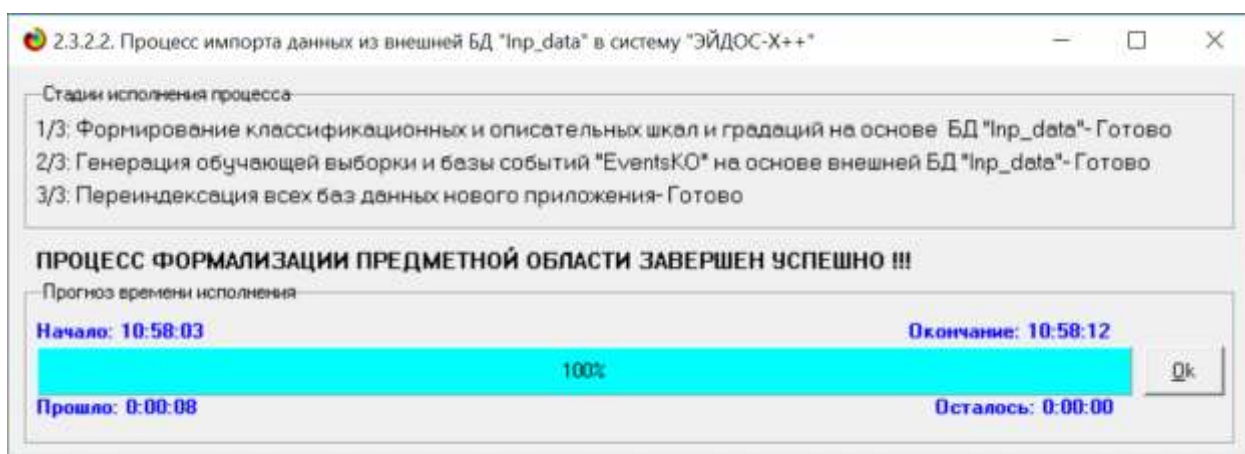


Рисунок 7 – Импорт данных

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

2.1. Классификационные шкалы и градации. Текущая модель: "INF3"

Код шкалы	Наименование классификационной шкалы
1	LOAN_STATUS

Код градации	Наименование градации классификационной шкалы
1	1/2N
2	2/2Y

Панель: Доб. шкалы | Доб. град. шкалы | Копир шкалы | Копир град. шкалы | Копир шкалы с град. | Удал. шкалы с град. | Удал. град. шкалы | Удаление и перекодирование | График будущей сценариев

Рисунок 8 – Классификационные шкалы

2.2. Описательные шкалы и градации. Текущая модель: "INF3"

Код шкалы	Наименование описательной шкалы
1	GENDER
2	MARRIED
3	DEPENDENTS
4	EDUCATION
5	SELF_EMPLOYED
6	APPLICANTINCOME
7	COAPPLICANTINCOME
8	LOANAMOUNT
9	LOAN_AMOUNT_TERM
10	CREDIT_HISTORY
11	PROPERTY_AREA

Код градации	Наименование градации описательной шкалы
5	1/5 (1.0000000, 1.0000000)
6	2/5 (1.0000000, 2.0000000)
7	3/5 (2.0000000, 2.0000000)
8	4/5 (2.0000000, 2.0000000)
9	5/5 (2.0000000, 3.0000000)

Панель: Доб. шкалы | Доб. град. шкалы | Копир шкалы | Копир град. шкалы | Копир шкалы с град. | Удал. шкалы с град. | Удал. град. шкалы | Перекодировать | Очистить | График прошлых сценариев

Рисунок 9 – Описательные шкалы

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

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) моделей осуществляется в режиме 3.5 системы «Эйдос». Математические модели, на основе которых рассчитываются статистические и СК-модели, подробно описаны в ряде монографий и статей автора. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко. Отметим лишь, что модели системы «Эйдос» основаны на матрице абсолютных частот, отражающей число встреч градаций описательных шкал по градациям классификационных шкал (фактов).

Но для решения всех задач используется не непосредственно сама эта матрица, а матрицы условных и безусловных процентных распределений и системно-когнитивные модели, которые рассчитываются на ее основе и отражают какое количество информации содержится в факте наблюдения определенной градации описательной шкалы о том, что объект моделирования перейдет в состояние, соответствующее определенной градации классификационной шкалы (классу) (рисунок 4).

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

Суть математической модели АСК-анализа состоит в следующем.

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

Таблица 1 – Матрица абсолютных частот (статистическая модель ABS)

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

На ее основе рассчитываются матрицы условных и безусловных процентных распределений (таблица 2).

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

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

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

1. В качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;
2. В качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу.

На практике часто встречается существенная несбалансированность данных, под которой понимается сильно отличающееся количество объектов обучающейся выборки, относящихся к различным классам. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 1) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) является весьма обоснованным и логичным.

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

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

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

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

В таблице 3 приведены формулы:

- для сравнения фактических и теоретических абсолютных частот;

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

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

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

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

Таблица 3 – Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	через относительные частоты	через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке	$N_i = \sum_{j=1}^w N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^w \sum_{j=1}^M N_{ij};$ $N_{ij} - \text{фактическая частота,}$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота.}$	
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	---	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество		

объектов обучающей выборки по классу		
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{N_j} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.		
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	---	$I_{ij} = N_{ij} - \bar{N}_{ij} = N_{ij} - \frac{N_i N_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{N_j} - 1 = \frac{N_{ij}N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		

Обозначения к таблице:

- i – значение прошлого параметра;
- j - значение будущего параметра;
- N_{ij} – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра;
- M – суммарное число значений всех прошлых параметров;
- W - суммарное число значений всех будущих параметров.
- N_i – количество встреч i -м значения прошлого параметра по всей выборке;
- N_j – количество встреч j -го значения будущего параметра по всей выборке;

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

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

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

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

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

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

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

Таблица 4 – Матрица системно-когнитивной модели

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

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

На основе системно-когнитивных моделей, решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений (обратная задача прогнозирования), а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Отметим, что как значимость значения фактора, степень детерминированности класса и ценность или качество модели в АСК-анализе рассматривается вариабельность значений частных критериев этого значения фактора, класса или модели в целом (таблица 4).

Численно эта вариабельность может измеряться разными способами, например, средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В

системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 5).

Таблица 5 – Уточнение терминологии АСК-анализа

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

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



Рисунок 10 – Выбор моделей для синтеза и верификации

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

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

Код признака	Наименование описательной шкалы и градации	1. LOAN_STATUS 1/2/N	2. LOAN_STATUS 2/2/Y	Среднее	Среднее	Средн. квадрат. откл.
1.0	GENDER-1/2-Female	28.0	59.0	87.0	43.50	21.92
2.0	GENDER-2/2-Male	82.0	212.0	294.0	147.00	91.92
3.0	MARRIED-1/2-No	52.0	101.0	153.0	76.50	34.65
4.0	MARRIED-2/2-Yes	58.0	170.0	228.0	114.00	79.20
5.0	DEPENDENTS-1/5-(1.0000000, 1.0000000)	17.0	35.0	52.0	26.00	12.73
6.0	DEPENDENTS-2/5-(1.0000000, 2.0000000)	15.0	44.0	59.0	29.50	20.51
7.0	DEPENDENTS-3/5-(2.0000000, 2.0000000)					
8.0	DEPENDENTS-4/5-(2.0000000, 2.0000000)					
9.0	DEPENDENTS-5/5-(2.0000000, 3.0000000)	8.0	20.0	28.0	14.00	6.49
10.0	EDUCATION-1/2-Graduate	73.0	201.0	274.0	137.00	90.51
11.0	EDUCATION-2/2-Not Graduate	32.0	69.0	101.0	50.50	26.16
12.0	SELF_EMPLOYED-1/2-No	99.0	247.0	346.0	173.00	104.65
13.0	SELF_EMPLOYED-2/2-Yes	11.0	24.0	35.0	17.50	9.19
14.0	APPLICANTINCOME-1/5-(150.0000000, 2425.0000000)	22.0	51.0	73.0	36.50	20.51
15.0	APPLICANTINCOME-2/5-(2425.0000000, 3052.0000000)	16.0	58.0	74.0	37.00	29.70
16.0	APPLICANTINCOME-3/5-(3052.0000000, 3598.0000000)	26.0	48.0	74.0	37.00	15.56
17.0	APPLICANTINCOME-4/5-(3598.0000000, 4583.0000000)	18.0	59.0	77.0	38.50	28.99
18.0	APPLICANTINCOME-5/5-(4583.0000000, 9703.0000000)	22.0	49.0	71.0	35.50	19.09
19.0	COAPPLICANTINCOME-1/5-(16.1200008, 1425.0000000)	6.0	32.0	38.0	19.00	18.38
20.0	COAPPLICANTINCOME-2/5-(1425.0000000, 1779.0000000)	9.0	29.0	38.0	19.00	14.14
21.0	COAPPLICANTINCOME-3/5-(1779.0000000, 2166.0000000)	10.0	28.0	38.0	19.00	12.73
22.0	COAPPLICANTINCOME-4/5-(2166.0000000, 2816.0000000)	5.0	33.0	38.0	19.00	19.80
23.0	COAPPLICANTINCOME-5/5-(2816.0000000, 33837.0000000)	10.0	28.0	38.0	19.00	12.73
24.0	LOANAMOUNT-1/5-(9.0000000, 80.0000000)	26.0	53.0	79.0	39.50	19.09
25.0	LOANAMOUNT-2/5-(80.0000000, 101.0000000)	21.0	52.0	73.0	36.50	21.92
26.0	LOANAMOUNT-3/5-(101.0000000, 115.0000000)	25.0	52.0	77.0	38.50	19.09

Рисунок 12 – Матрица абсолютных частот

5.5. Модель: "9. INF6 - частный критерий: разность логарифмов вероятностей; вероятности из PRC1".

Код признака	Наименование описательной шкалы и градации	1. LOAN_STATUS 1/2/N	2. LOAN_STATUS 2/2/Y	Среднее	Среднее	Средн. квадрат. откл.
1.0	GENDER-1/2-Female	0.318	-0.126	0.192	0.096	0.313
2.0	GENDER-2/2-Male	-0.118	0.047	-0.071	-0.036	0.116
3.0	MARRIED-1/2-No	0.819	-0.324	0.496	0.248	0.808
4.0	MARRIED-2/2-Yes	-0.619	0.245	-0.375	-0.187	0.611
5.0	DEPENDENTS-1/5-(1.0000000, 1.0000000)	0.215	-0.085	0.130	0.065	0.212
6.0	DEPENDENTS-2/5-(1.0000000, 2.0000000)	-0.161	0.064	-0.097	-0.049	0.159
7.0	DEPENDENTS-3/5-(2.0000000, 2.0000000)					
8.0	DEPENDENTS-4/5-(2.0000000, 2.0000000)					
9.0	DEPENDENTS-5/5-(2.0000000, 3.0000000)	0.007	-0.003	0.004	0.002	0.007
10.0	EDUCATION-1/2-Graduate	-0.433	0.171	-0.262	-0.131	0.427
11.0	EDUCATION-2/2-Not Graduate	0.321	-0.127	0.194	0.097	0.317
12.0	SELF_EMPLOYED-1/2-No	0.097	-0.038	0.059	0.029	0.096
13.0	SELF_EMPLOYED-2/2-Yes	0.103	-0.041	0.062	0.031	0.101
14.0	APPLICANTINCOME-1/5-(150.0000000, 2425.0000000)	0.126	-0.050	0.076	0.038	0.124
15.0	APPLICANTINCOME-2/5-(2425.0000000, 3052.0000000)	-0.468	0.185	-0.283	-0.141	0.461
16.0	APPLICANTINCOME-3/5-(3052.0000000, 3598.0000000)	0.477	-0.188	0.288	0.144	0.470
17.0	APPLICANTINCOME-4/5-(3598.0000000, 4583.0000000)	-0.359	0.142	-0.217	-0.109	0.354
18.0	APPLICANTINCOME-5/5-(4583.0000000, 9703.0000000)	0.179	-0.071	0.108	0.054	0.177
19.0	COAPPLICANTINCOME-1/5-(16.1200008, 1425.0000000)	-0.449	0.178	-0.272	-0.136	0.443
20.0	COAPPLICANTINCOME-2/5-(1425.0000000, 1779.0000000)	-0.166	0.066	-0.101	-0.050	0.164
21.0	COAPPLICANTINCOME-3/5-(1779.0000000, 2166.0000000)	-0.072	0.028	-0.043	-0.022	0.071
22.0	COAPPLICANTINCOME-4/5-(2166.0000000, 2816.0000000)	-0.544	0.215	-0.329	-0.165	0.537
23.0	COAPPLICANTINCOME-5/5-(2816.0000000, 33837.0000000)	-0.072	0.028	-0.043	-0.022	0.071
24.0	LOANAMOUNT-1/5-(9.0000000, 80.0000000)	0.343	-0.135	0.207	0.104	0.336
25.0	LOANAMOUNT-2/5-(80.0000000, 101.0000000)	0.031	-0.012	0.019	0.009	0.031
26.0	LOANAMOUNT-3/5-(101.0000000, 115.0000000)	0.302	-0.119	0.183	0.091	0.298

Рисунок 13 – Модель INF6

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

3.4 Задача – 4. Верификация моделей

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

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

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

В соответствии с критерием достоверности F-мерой Ван Ризбергена наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием: «Семантический резонанс знаний»: $F=0.819$ (рисунок 14).

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

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

3.4. Обобщенная модель при разном крит. Тестовая модель: "BPT"

Идентификатор модели	Имя модели	Среднее значение критерия	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)
1. AB1	частный критерий: количество стран с рейтингом "Удов."	381	381	381	381	381	381	381	381	381	381	381	381	381	381
2. AB2	частный критерий: количество стран с рейтингом "Удов."	381	381	381	381	381	381	381	381	381	381	381	381	381	381
3. PR1	частный критерий: доля вероятности от 0 до 100%	381	381	381	381	381	381	381	381	381	381	381	381	381	381
4. PR2	частный критерий: условная вероятность от 0 до 100%	381	381	381	381	381	381	381	381	381	381	381	381	381	381
5. PR3	частный критерий: количество стран по А/Среднему в.	381	317	296	85	68	3 799	0.632	0.810	137.838	104.082	19			
6. PR4	частный критерий: количество стран по А/Среднему в.	381	280	319	62	101	3 819	0.738	0.776	48.126	113.772	18			
7. PR5	частный критерий: количество стран по А/Среднему в.	381	316	296	85	65	3 788	0.629	0.808	139.324	109.064	20			
8. PR6	частный критерий: количество стран по А/Среднему в.	381	381	326	55	100	3 836	0.734	0.784	46.529	121.646	8			
9. PR7	частный критерий: Уникальность: разность между факт.	381	387	397	74	74	3 896	0.804	0.899	136.504	166.584	30			
10. PR8	частный критерий: Уникальность: разность между факт.	381	387	397	74	74	3 896	0.804	0.899	136.504	166.584	30			
11. PR9	частный критерий: ROI (Return On Investment) - абсолют.	381	312	312	89	89	3 818	0.819	0.819	138.326	138.326	13			
12. PR10	частный критерий: ROI (Return On Investment) - абсолют.	381	296	296	85	85	3 777	0.777	0.777	49.485	93.485	11			
13. PR11	частный критерий: ROI (Return On Investment) - абсолют.	381	322	312	89	89	3 818	0.819	0.819	138.326	138.326	13			
14. PR12	частный критерий: ROI (Return On Investment) - абсолют.	381	380	380	81	81	3 787	0.787	0.787	53.162	69.312	17			
15. PR13	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			
16. PR14	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			
17. PR15	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			
18. PR16	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			
19. PR17	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			
20. PR18	частный критерий: разн. доли безразличности: вар.	381	387	387	74	74	3 896	0.804	0.896	136.504	166.584	30			

Рисунок 14 – Оценка достоверности моделей (первая часть)

3.4. Обобщенная модель при разном крит. Тестовая модель: "BPT"

Идентификатор модели	Имя модели	Среднее значение критерия	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)	Среднее значение критерия (с погрешностью)
1. AB1	частный критерий: количество стран с рейтингом "Удов."	264.548	264.548	242.844	0.524	1.888	0.488	8.700	0.635						
2. AB2	частный критерий: количество стран с рейтингом "Удов."	264.548	264.548	242.844	0.524	1.888	0.488	8.700	0.635						
3. PR1	частный критерий: доля вероятности от 0 до 100%	264.548	264.548	242.844	0.524	1.888	0.488	8.700	0.635						
4. PR2	частный критерий: условная вероятность от 0 до 100%	264.548	264.548	242.844	0.524	1.888	0.488	8.700	0.635						
5. PR3	частный критерий: количество стран по А/Среднему в.	107.038	106.992	19.962	16.279	0.044	0.049	0.876	0.350	8.358	8.216				
6. PR4	частный критерий: количество стран по А/Среднему в.	48.126	113.772	10.851	20.841	0.027	0.704	0.762	8.172	8.387	8.162				
7. PR5	частный критерий: количество стран по А/Среднему в.	109.324	109.064	20.828	16.700	0.045	0.047	0.876	0.346	8.360	8.236				
8. PR6	частный критерий: количество стран по А/Среднему в.	46.529	121.646	8.971	20.595	0.052	0.693	0.768	8.166	8.373	8.147				
9. PR7	частный критерий: Уникальность: разность между факт.	166.584	166.584	30.276	10.375	0.048	0.048	0.846	0.542	8.542	8.489				
10. PR8	частный критерий: Уникальность: разность между факт.	166.584	166.584	30.276	10.375	0.048	0.048	0.846	0.542	8.542	8.489				
11. PR9	частный критерий: ROI (Return On Investment) - абсолют.	138.326	138.326	27.895	17.876	0.058	0.058	0.856	0.534	8.547	8.515				
12. PR10	частный критерий: ROI (Return On Investment) - абсолют.	49.485	93.485	13.314	8.402	0.035	0.035	0.817	0.164	8.160	8.117				
13. PR11	частный критерий: ROI (Return On Investment) - абсолют.	138.326	138.326	27.895	17.876	0.058	0.058	0.856	0.534	8.547	8.515				
14. PR12	частный критерий: ROI (Return On Investment) - абсолют.	53.162	69.312	12.849	5.970	0.015	0.042	0.828	0.177	8.231	8.149				
15. PR13	частный критерий: разн. доли безразличности: вар.	166.584	166.584	30.276	10.375	0.048	0.048	0.846	0.542	8.542	8.489				
16. PR14	частный критерий: разн. доли безразличности: вар.	89.645	149.669	16.290	24.930	0.032	0.764	0.776	8.263	8.488	8.220				
17. PR15	частный критерий: разн. доли безразличности: вар.	167.968	167.968	30.589	10.589	0.048	0.048	0.846	0.547	8.547	8.413				
18. PR16	частный критерий: разн. доли безразличности: вар.	77.974	159.313	14.965	27.740	0.028	0.728	0.705	8.258	8.526	8.189				

Рисунок 15 – Оценка достоверности моделей (вторая часть)

1.4. Обзор формы по результатам приравненных. Текущая модель: "INF3"

Идентификатор модели	Идентификатор критерия	Число истинных решений	Число ложных решений	Число истинных решений + ложных решений	Число истинных решений / общее число решений	Число ложных решений / общее число решений	Число истинных решений / общее число решений	Число ложных решений / общее число решений	Число истинных решений / общее число решений	Число ложных решений / общее число решений	Дата создания модели	Дата обновления модели	
1. INF3 - интегральный критерий: количество истинных решений	Корреляция инф. частот с обр.	0.524	1.800	2.324	0.226	0.774	100.000	130.800	50.000	29.03.2024	11.00.52		
2. INF3 - интегральный критерий: количество истинных решений	Средняя инф. частот по признакам	0.524	1.800	2.324	0.226	0.774	100.000	130.800	50.000	29.03.2024	11.00.53		
3. INF3 - интегральный критерий: доля вероятности инф. признака в прав.	Корреляция инф. частот с о.	0.524	1.800	2.324	0.226	0.774	100.000	130.800	50.000	29.03.2024	11.00.54		
4. INF3 - интегральный критерий: доля вероятности инф. признака в лев.	Корреляция инф. частот с о.	0.524	1.800	2.324	0.226	0.774	100.000	130.800	50.000	29.03.2024	11.00.55		
5. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.254	0.570	0.824	0.309	0.691	83.292	63.409	36.400	16.798	73.303	29.03.2024	11.00.56
6. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.254	0.570	0.824	0.309	0.691	83.292	63.409	36.400	16.798	73.303	29.03.2024	11.00.56
7. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.257	0.570	0.827	0.307	0.693	82.948	63.409	36.400	17.060	73.170	29.03.2024	11.00.57
8. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.204	0.570	0.774	0.329	0.671	73.763	76.237	24.763	26.247	74.496	29.03.2024	11.00.59
9. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.409	0.570	0.979	0.577	0.423	88.577	62.412	37.508	19.423	71.494	29.03.2024	11.00.58
10. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.399	0.570	0.969	0.577	0.423	88.577	62.412	37.508	19.423	71.494	29.03.2024	11.00.58
11. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.252	0.570	0.822	0.313	0.687	81.898	67.265	31.735	18.110	74.957	29.03.2024	11.00.59
12. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.099	0.511	0.610	0.155	0.845	77.698	75.282	24.718	22.310	76.489	29.03.2024	11.01.00
13. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.254	0.570	0.824	0.309	0.691	81.898	67.265	31.735	18.110	74.957	29.03.2024	11.01.01
14. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.121	0.543	0.664	0.221	0.779	79.748	70.847	29.153	21.240	74.794	29.03.2024	11.01.02
15. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.409	0.570	0.979	0.577	0.423	88.577	62.412	37.508	19.423	71.494	29.03.2024	11.01.02
16. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.337	0.544	0.881	0.616	0.384	88.577	62.412	37.508	19.423	71.494	29.03.2024	11.01.02
17. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.413	0.570	0.983	0.571	0.429	88.577	62.412	37.508	19.423	71.494	29.03.2024	11.01.03
18. INF3 - интегральный критерий: количество знаний по А.Хорошему в.	Семантический разрыв: зна.	0.251	0.577	0.828	0.433	0.567	79.265	69.838	41.362	28.716	69.952	29.03.2024	11.01.03

Рисунок 16 – Оценка достоверности моделей (третья часть)

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

На рисунке 17 представлены частотные распределения числа истинных и ложных положительных и отрицательных решений в СК-модели INF3 по данным обучающей выборки.

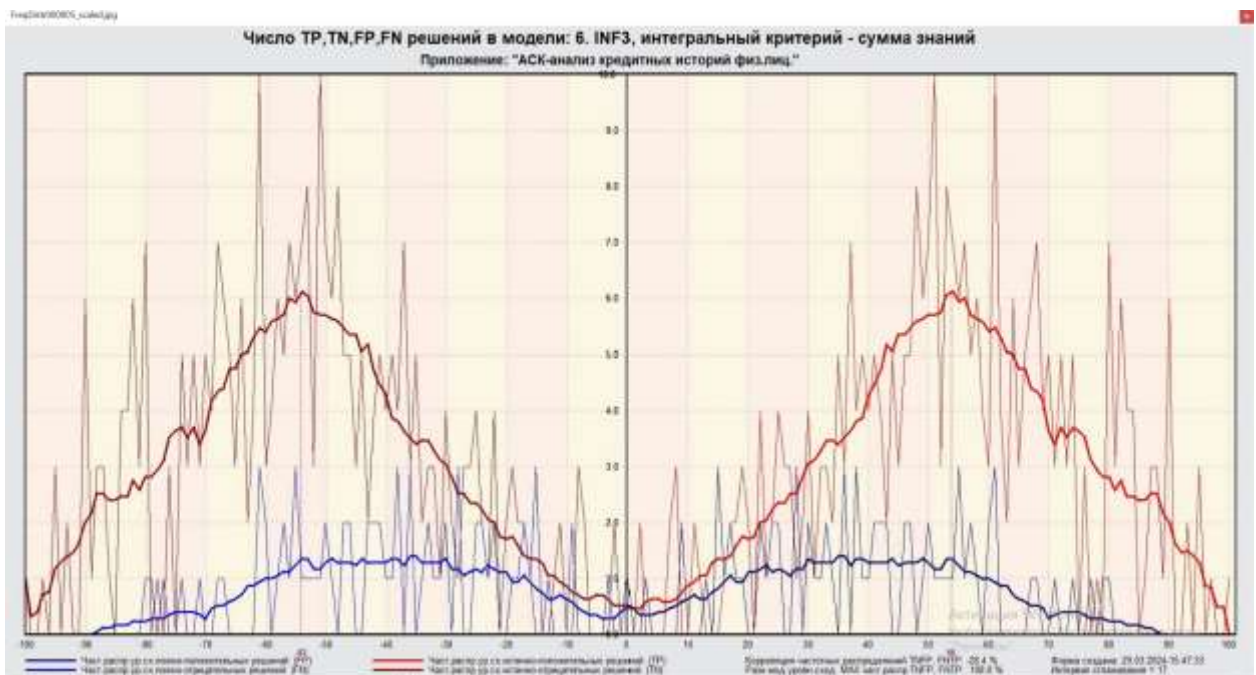


Рисунок 17 – Частотные распределение модели INF3

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

3.5 Задача – 5. Выбор наиболее достоверной модели

Все последующие задачи решаются в наиболее достоверной модели. Причины этого просты. Если модель достоверна, то:

- идентификация объекта с классом достоверна, т.е. модель относит объекты к классам, к которым они действительно принадлежат;
- прогнозирование достоверно, т.е. действительно наступают те события, которые прогнозируются;
- принятие решений адекватно (достоверно), т.е. после реализации принятых управляющих решений объект управления действительно переходит в целевые будущие состояния;
- исследование достоверно, т.е. полученные в результате исследования модели объекта моделирования выводы могут быть с полным основанием отнесены к объекту моделирования.

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

Выберем одну из наиболее достоверных моделей INF3, что показано на рисунках 18 и 19.

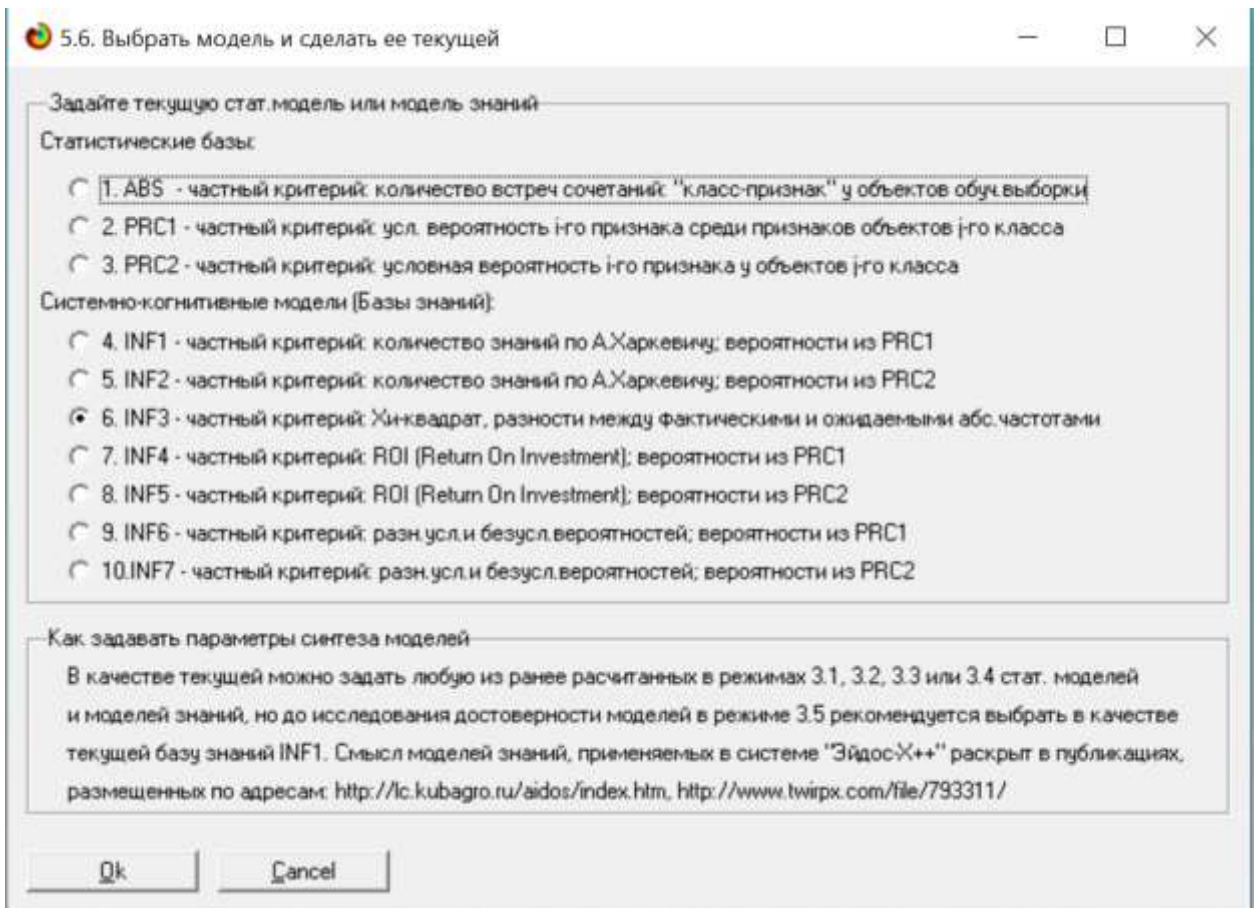


Рисунок 18 – Выбор текущей модели

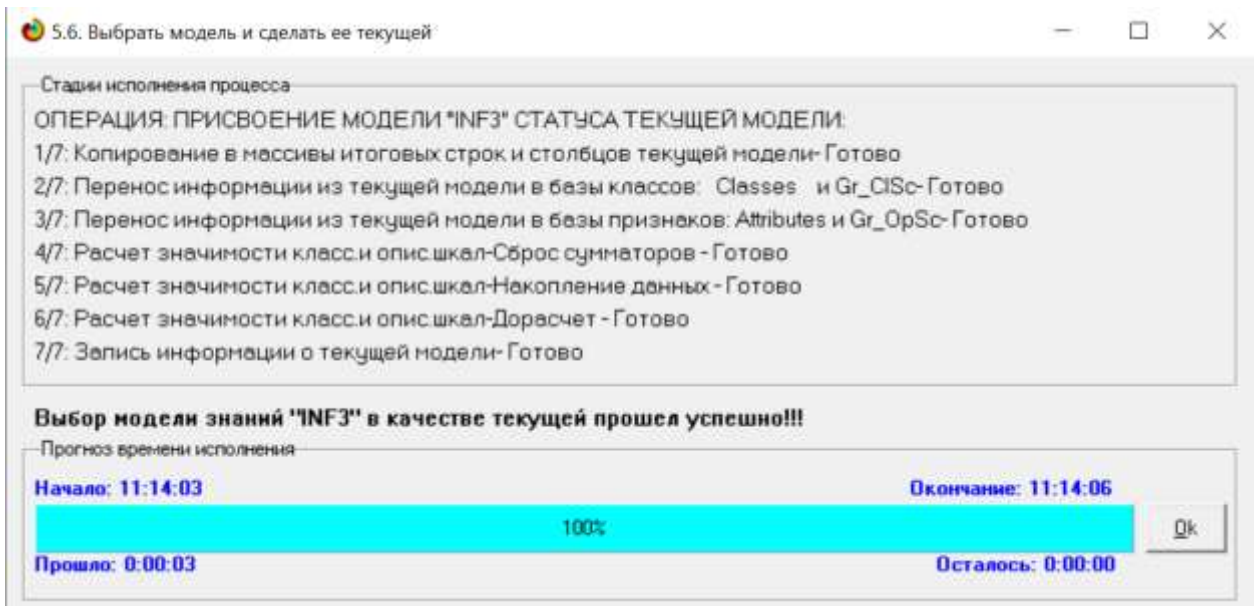


Рисунок 19 – Исполнение присвоения статуса текущей модели

3.6 Задача – 6. Системная идентификация и прогнозирование

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

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

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

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

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

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

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

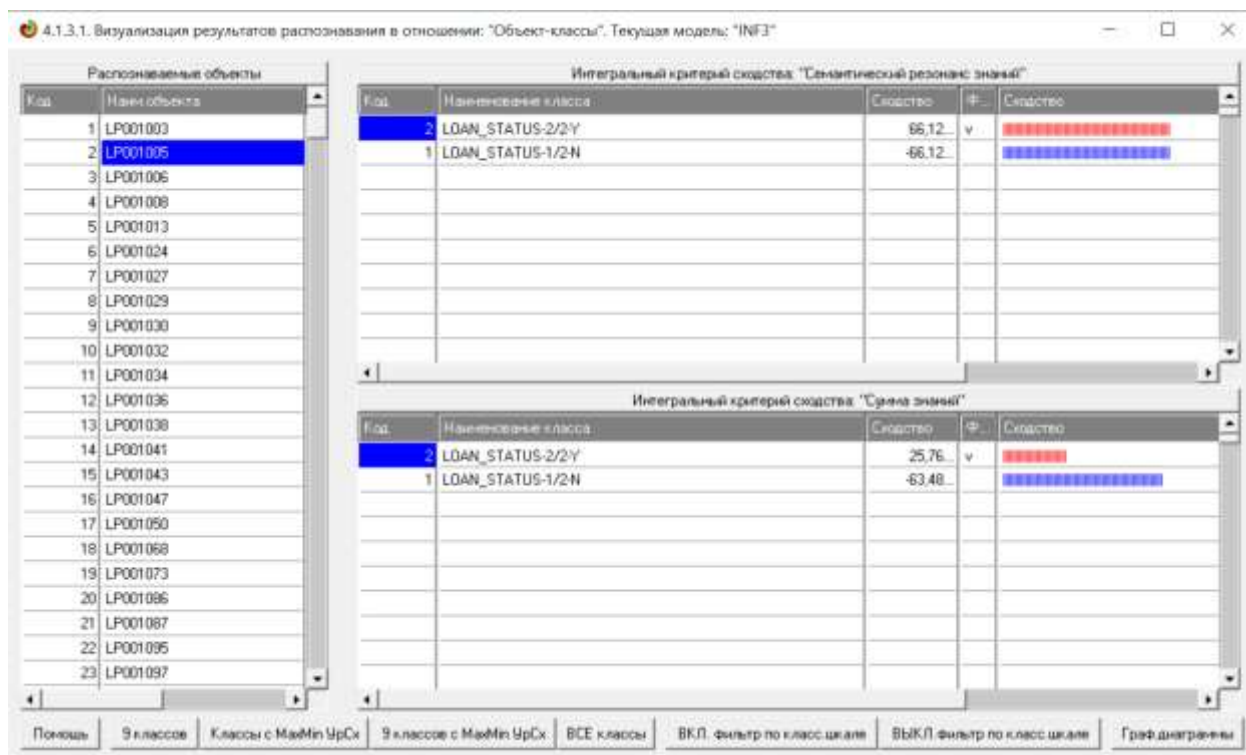


Рисунок 20 – Пример идентификации объектов в модели INF3

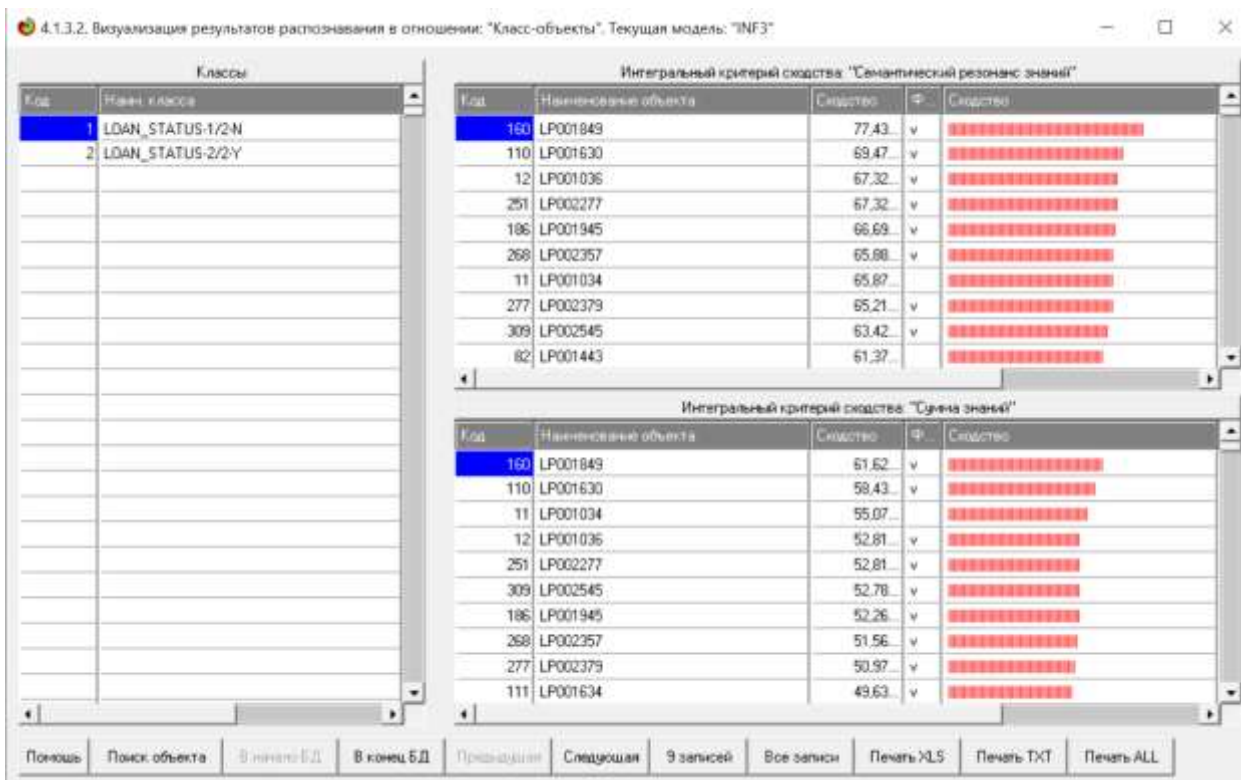


Рисунок 21 – Пример идентификации объектов в модели INF3

После данного этапа можно приступить к выбору наиболее достоверной модели.

3.6.1 Интегральный критерий «Сумма знаний»

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

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help 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 признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

3.6.2 Интегральный критерий «Семантический резонанс знаний»

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

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

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

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: $I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}$, $L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_i}$. Поэтому по своей

сути он также является скалярным произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применения сплайнов, в частности

линейной интерполяции: $I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}$, $L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}}$, Это позволяет

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

3.6.3 Важные математические свойства интегральных критериев

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

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

Во-вторых, данный интегральный критерий является фильтром, подавляющим белый шум, который всегда присутствует в эмпирических

исходных данных и в моделях, созданных на их основе. Это свойство подавлять белый шум проявляется у данного критерия тем ярче, чем больше в модели градаций описательных шкал.

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

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

В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется вес каждого обобщенного образа класса в образе объекта.

3.7 Задача – 7. Поддержка принятия решений

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

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

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

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

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

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

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

Поэтому в АСК-анализе и системе «Эйдос» реализован развитый алгоритм принятия решений (режим 6.3), в котором кроме SWOT-анализа используются также результаты решения задачи прогнозирования и результаты кластерно-конструктивного анализа классов и значений факторов, т.е. некоторые результаты решения задачи исследования предметной области.

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

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

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

Инвертированная SWOT-диаграмма влияния фактора «Пол: мужской» на итоговое решение по кредитной заявке на рисунке 23.

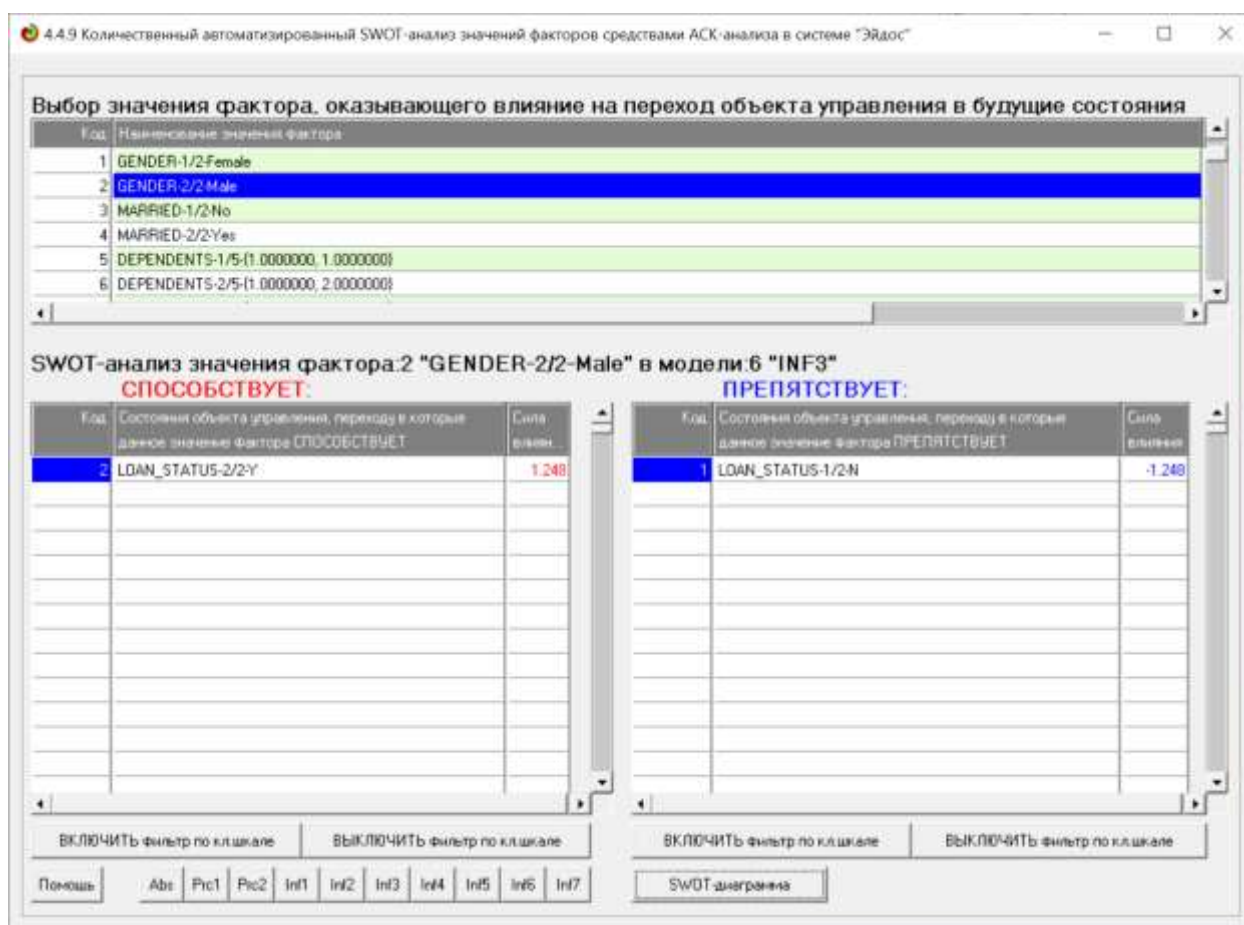


Рисунок 22 – Количественный автоматизированный SWOT-анализ значений факторов

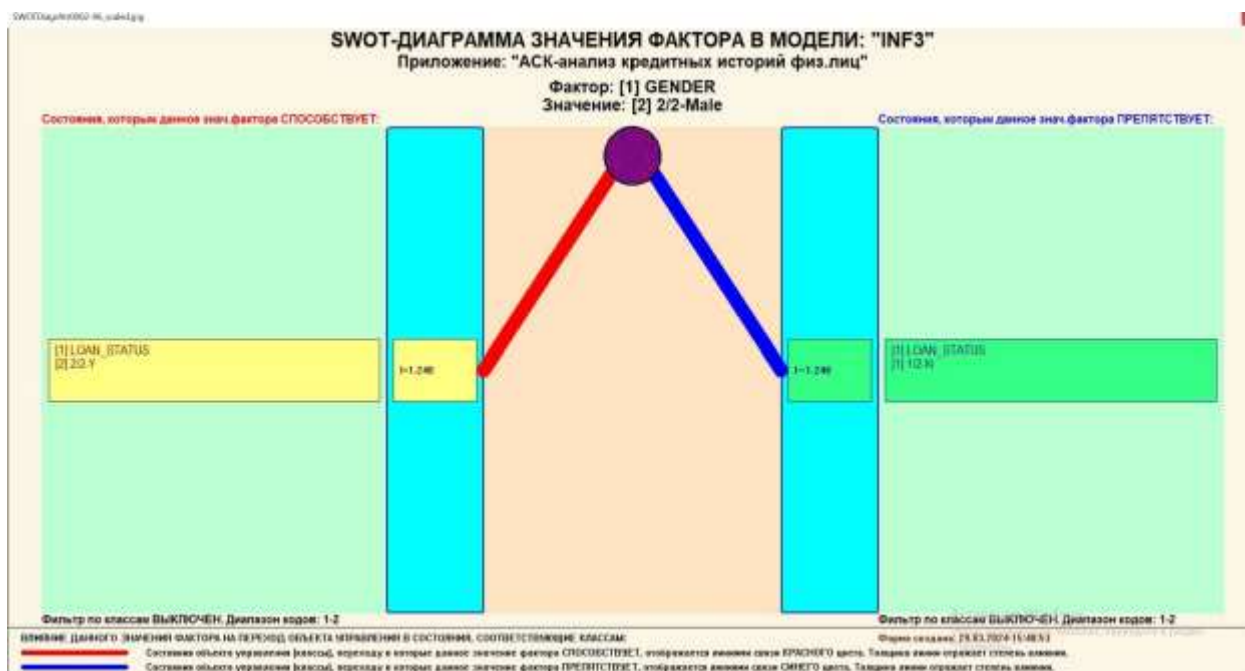


Рисунок 23 – Инвертированная SWOT-диаграмма

3.8.2 Кластерно-конструктивный анализ классов

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 25) рассчитывается матрица сходства классов по системе их детерминации и на основе этой матрицы рассчитывается и выводится три основных формы:

— круговая 2D-когнитивная диаграмма классов (режим 4.2.2.2) (рисунок 26);

— агломеративные дендрограммы, полученные в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.4) (рисунок 27).

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



Рисунок 26 – Семантическая 2D-когнитивная диаграмма классов

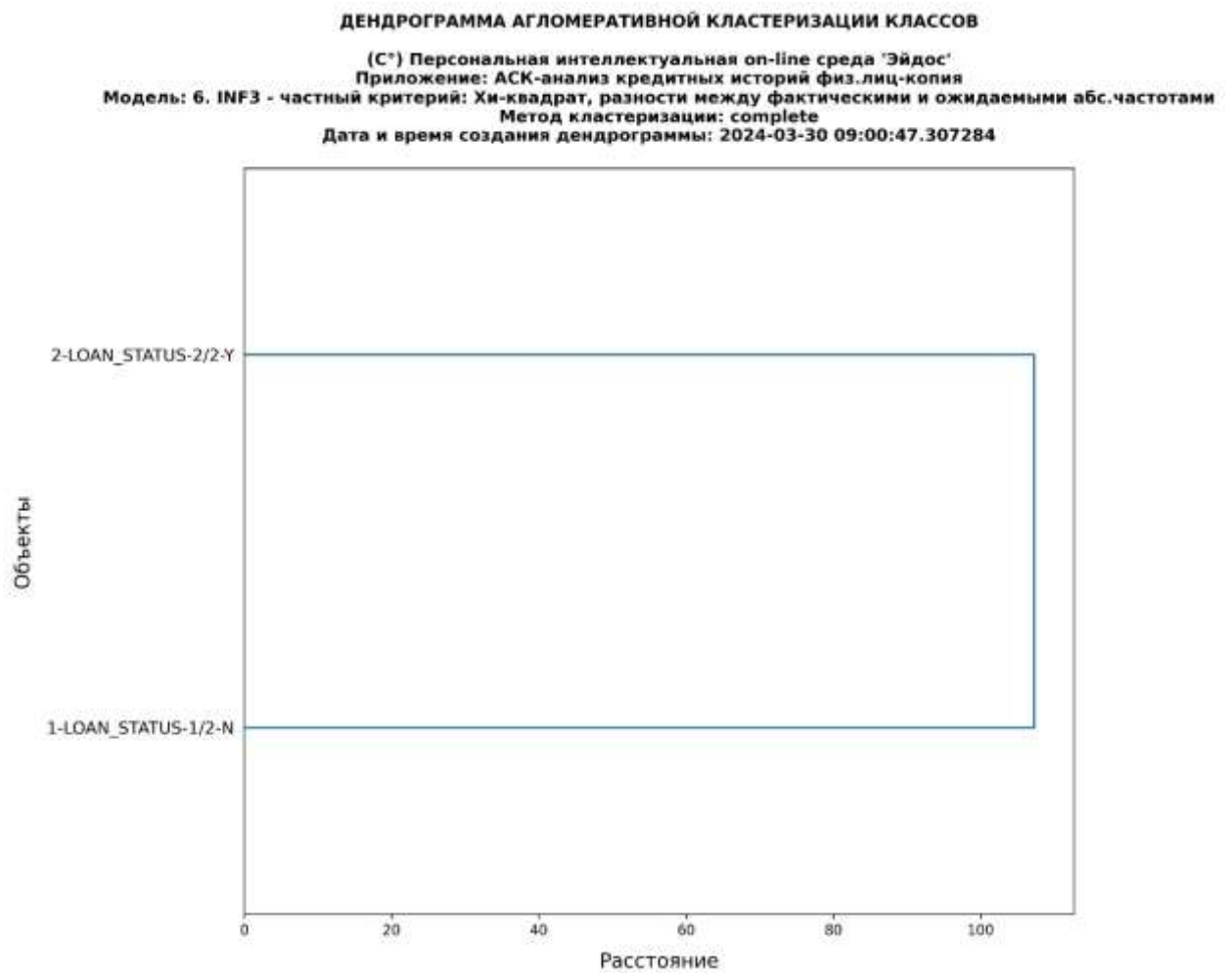


Рисунок 27 – Дендрограмма когнитивной кластеризации классов

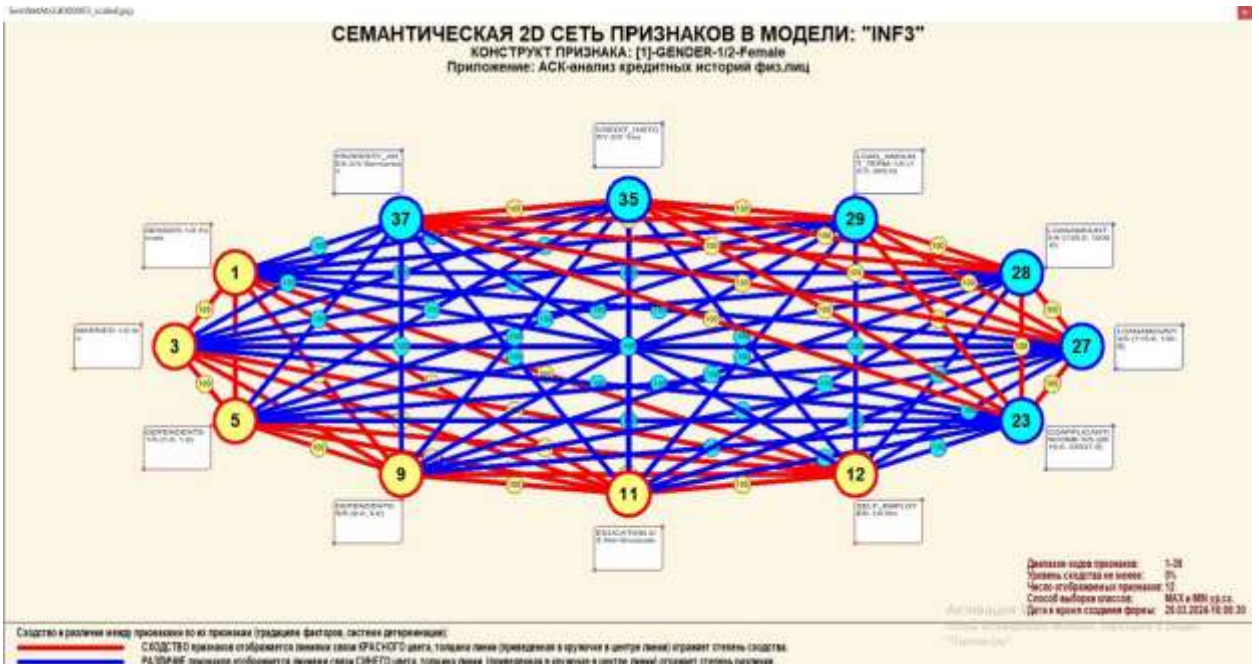


Рисунок 29 – Круговая 2D-когнитивная диаграмма значений факторов в СК-модели INF3



Рисунок 30 – Агломеративная дендрограмма, полученная в результате когнитивной кластеризации признаков

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 33). В форме управления визуализацией (рисунок 32) есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.5 Нелокальная нейронная сеть

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

Есть также возможность визуализации любого одного слоя нелокальной нейронной сети (режим 4.4.11 системы «Эйдос», рисунок 34).

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

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

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Set	Код	Наименование нелокального нейрона (класса)
1	1	LOAN_STATUS-1/2-N
2	2	LOAN_STATUS-2/2-Y

Помощь | Максимальное количество отображаемых нейронов: 16 | ClearSet | Диапазон кодов отображаемых нейронов: 1 - 2
 Максимальное количество отображаемых связей: 1000 | Диапазон кодов отображаемых рецепторов: 1 - 38

Подготовка визуализации нейрона 1 "LOAN_STATUS-1/2-N" в модели 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
34	CREDIT_HISTORY-1/2-No	34.480
36	PROPERTY_AREA-1/3-Rural	10.906
3	MARRIED-1/2-No	8.677
16	APPLICANTINCOME-3/5-{3052.000000, 3598.000000}	5.047
33	LOAN_AMOUNT_TERM-5/5-{360.000000, 480.000000}	3.895
24	LOANAMOUNT-1/5-{9.000000, 80.000000}	3.631
11	EDUCATION-2/2-Not Graduate	3.401
1	GENDER-1/2-Female	3.355
38	PROPERTY_AREA-3/3-Urban	3.322
27	LOANAMOUNT-4/5-{115.000000, 130.000000}	1.707

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
35	CREDIT_HISTORY-2/2-Yes	-32.362
37	PROPERTY_AREA-2/3-Semiurban	-12.190
4	MARRIED-2/2-Yes	-6.559
22	COAPPLICANTINCOME-4/5-{2166.000000, 2616.000000}	-5.760
15	APPLICANTINCOME-2/5-{2425.000000, 3052.000000}	-4.953
19	COAPPLICANTINCOME-1/5-{16.1200008, 1425.000000}	-4.760
10	EDUCATION-1/2-Graduate	-4.584
17	APPLICANTINCOME-4/5-{3598.000000, 4563.000000}	-3.803
27	LOANAMOUNT-4/5-{115.000000, 130.000000}	-3.652
20	LOAN_AMOUNT_TERM-4/5-{360.000000, 480.000000}	-1.707

ВКЛЮЧИТЬ фильтр по фактору | ВЫКЛЮЧИТЬ фильтр по фактору

НейроСеть | Максимальное количество отображаемых рецепторов: 16 | Отображать связи с интенсивностью >= % от макс.: 0.000

Сортировать связи: по модулю информативности | по информативности и знаку

Отображать наименования: нейронов | рецепторов

Рисунок 34 – Визуализация слоя нелокальной нейронной сети

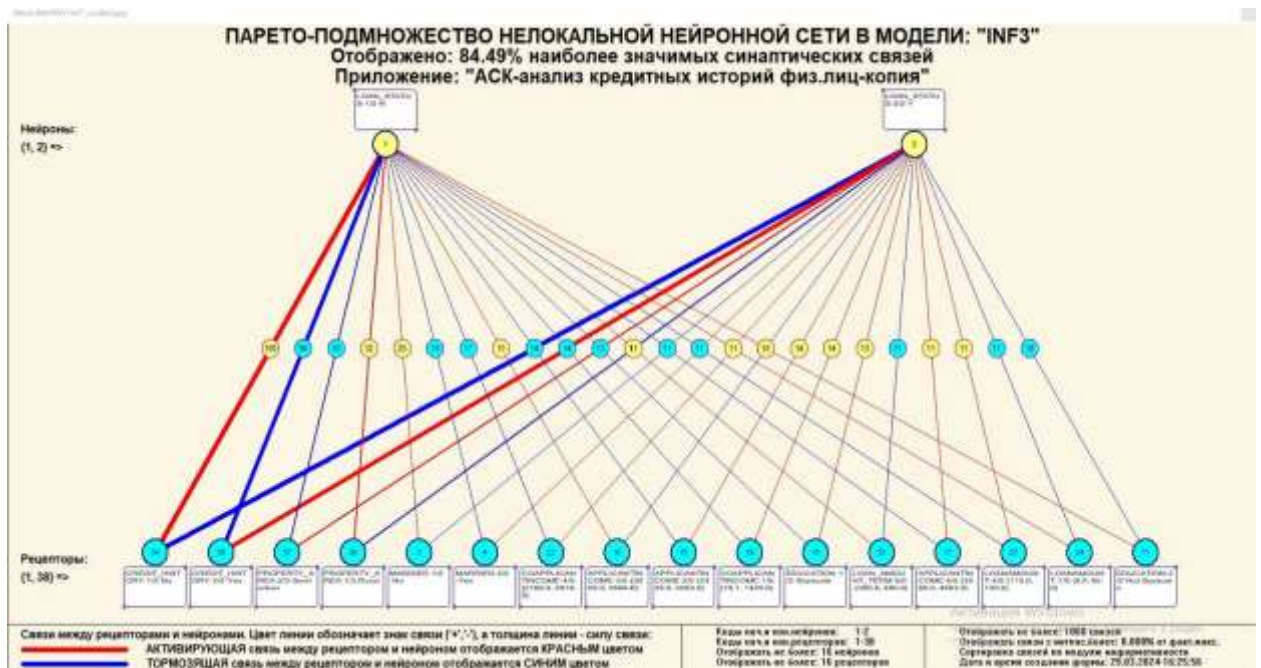


Рисунок 35 – Нейронная сеть в СК-модели INF3

3.8.6 3D-интегральные когнитивные карты

3D-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 26) вверху и когнитивной диаграммы значений факторов (рисунок 29) внизу и соединяющего их одного слоя нейронной сети (рисунок 35) (режим 4.4.12 системы «Эйдос») (рисунок 36) (рисунок 37):

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

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

№	Код	Наименование нелокального нейрона (класса)
1	LOAN_STATUS-1/2-N	
2	LOAN_STATUS-2/2-Y	

Поисковая: Максимальное количество отображаемых нейронов: 16
 Максимальное количество отображаемых связей: 1000 ClearSel
 Диапазон кодов отображаемых нейронов: 1 2
 Диапазон кодов отображаемых рецепторов: 1 38

Подготовка визуализации нейрона: 1 "LOAN_STATUS-1/2-N" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
34	CREDIT_HISTORY-1/2-No	34.490
36	PROPERTY_AREA-1/3-Rural	10.996
3	MARRIED-1/2-No	8.677
16	APPLICANTINCOME-3/5-(3052.0000000, 3598.0000000)	5.047
33	LOAN_AMOUNT_TERM-5/5-(360.0000000, 480.0000000)	3.895
24	LOANAMOUNT-1/5-(9.0000000, 80.0000000)	3.631
11	EDUCATION-2/2-Not Graduate	3.401
1	GENDER-1/2-Female	3.366
38	PROPERTY_AREA-3/3-Urban	3.322
30	LOAN_AMOUNT_TERM-1/5-(9.0000000, 115.0000000)	3.107

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
35	CREDIT_HISTORY-2/2-Yes	-32.362
37	PROPERTY_AREA-2/3-Semiurban	-12.190
4	MARRIED-2/2-Yes	-6.559
22	COAPPLICANTINCOME-4/5-(2166.0000000, 2816.0000000)	-5.760
15	APPLICANTINCOME-2/5-(2425.0000000, 3052.0000000)	-4.953
19	COAPPLICANTINCOME-1/5-(16.1290008, 1425.0000008)	-4.760
10	EDUCATION-1/2-Graduate	-4.584
17	APPLICANTINCOME-4/5-(3598.0000000, 4583.0000000)	-3.803
27	LOANAMOUNT-4/5-(115.0000000, 130.0000000)	-3.652
30	LOAN_AMOUNT_TERM-1/5-(9.0000000, 115.0000000)	-3.307

ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору

Когн. карта

Сортировать связи:
 по модулю информативности
 по информативности и знаку

Отображать наименования:
 нейроны
 рецепторы

Максимальное количество отображаемых рецепторов: 16
 Отображать связи с интенсивностью >= % от макс.: 0,000

Рисунок 36 – Составление слоя интегральной когнитивной карты

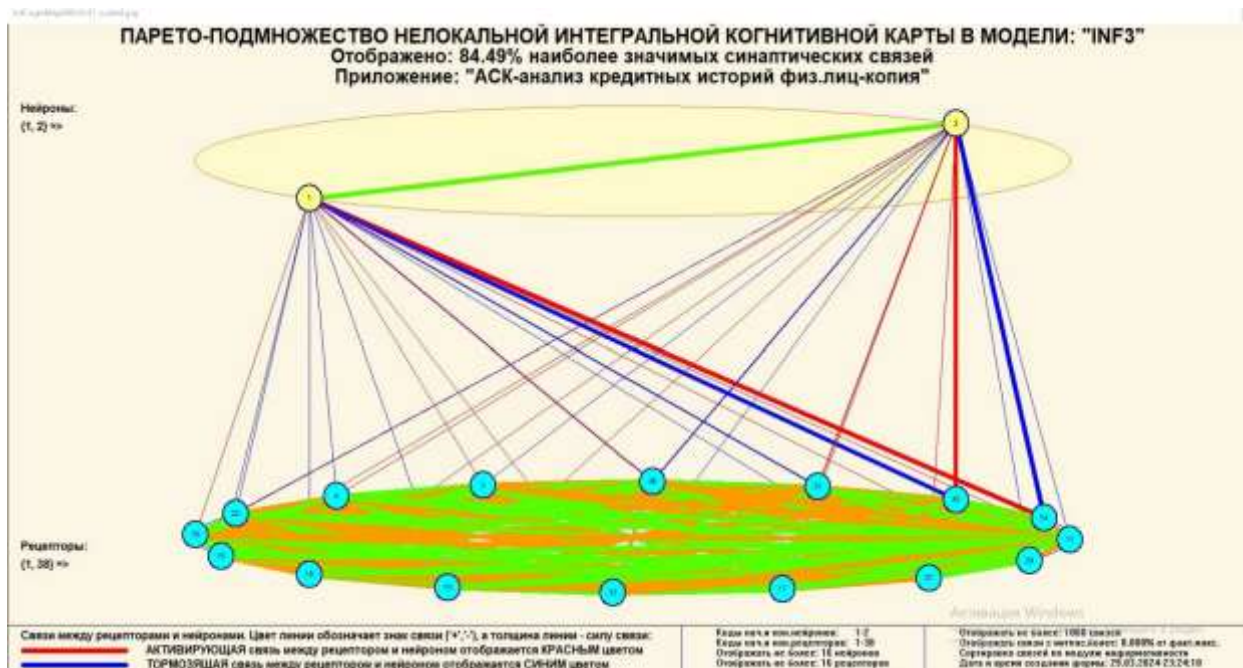


Рисунок 37 – 3D-когнитивная диаграмма классов и признаков

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

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

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

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

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых, может быть, одним из первых писал Дьердь Пойа. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году.

Пример опосредованных правдоподобных рассуждений.

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

Примеры экранной формы управления (рисунок 38) и 2D-интегральной когнитивной карты содержательного сравнения классов по их системе детерминации приведены ниже на рисунке 39. Всего системой в данной модели генерируется 324 подобных диаграмм, поэтому, естественно, все они не приводятся.

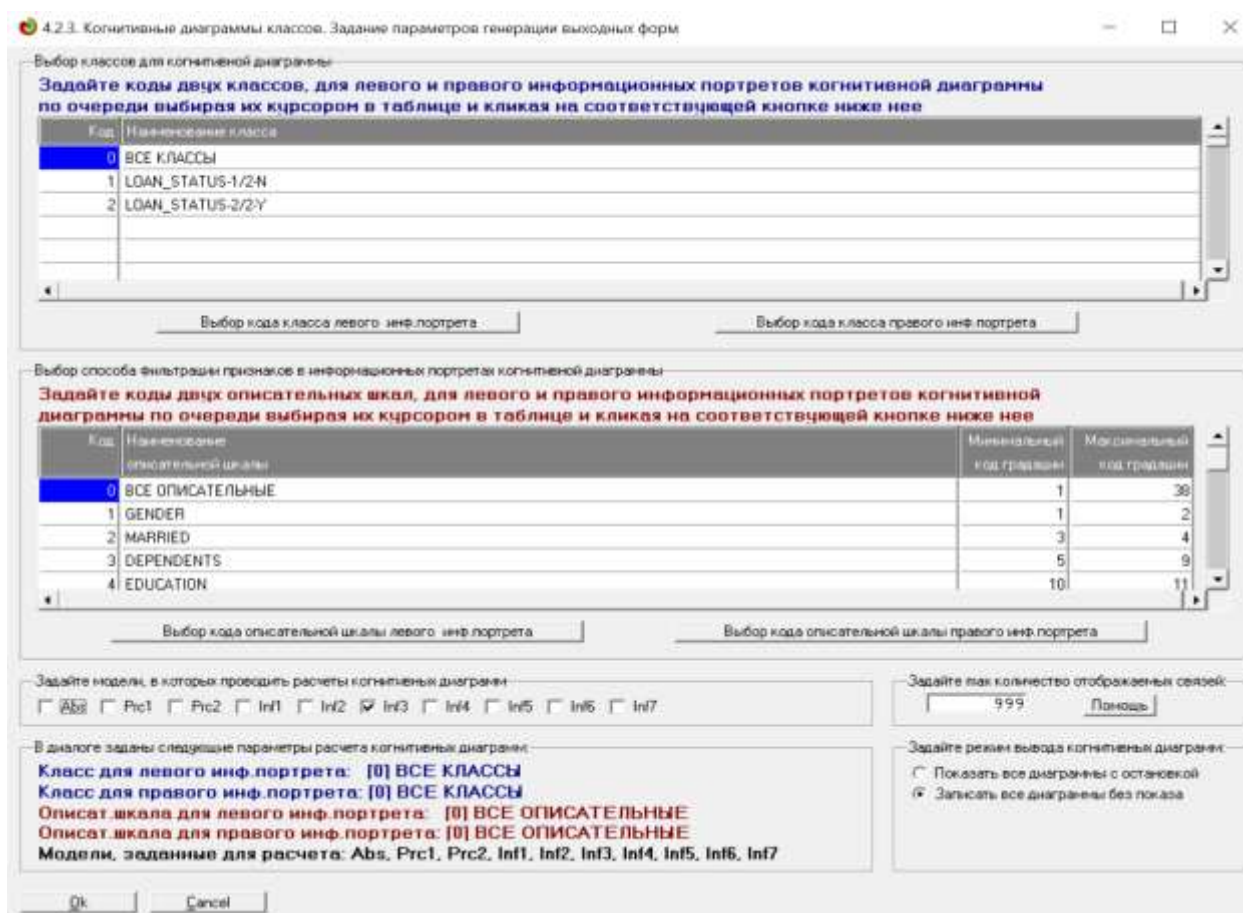


Рисунок 38 – Задание параметров генерации когнитивной диаграммы классов

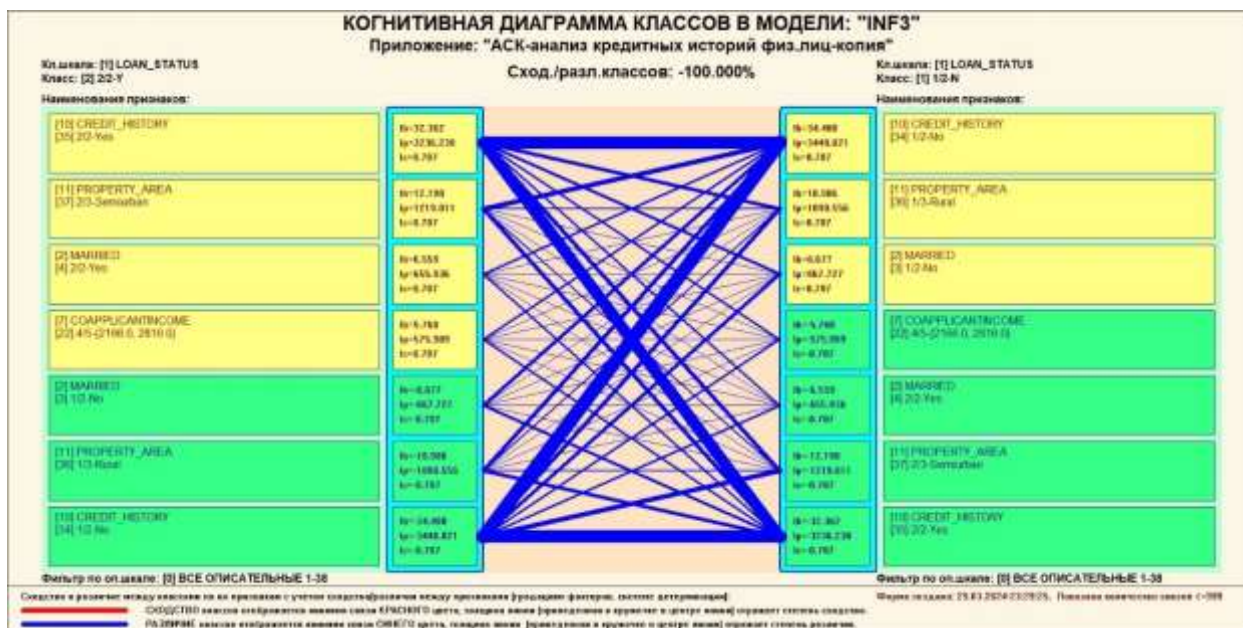


Рисунок 39 – Пример 2D-интегральной когнитивной карты содержательного сравнения классов по их системе детерминации в СК-модели INF3

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

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

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

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

Это видно из когнитивных диаграмм, которые можно получить в режиме 4.3.3 системы «Эйдос» (рисунок 40).

Примеры экранной формы управления и 2D-интегральной когнитивной карты содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта моделирования в будущие состояния, соответствующие классам, приведена ниже на рисунке 41.

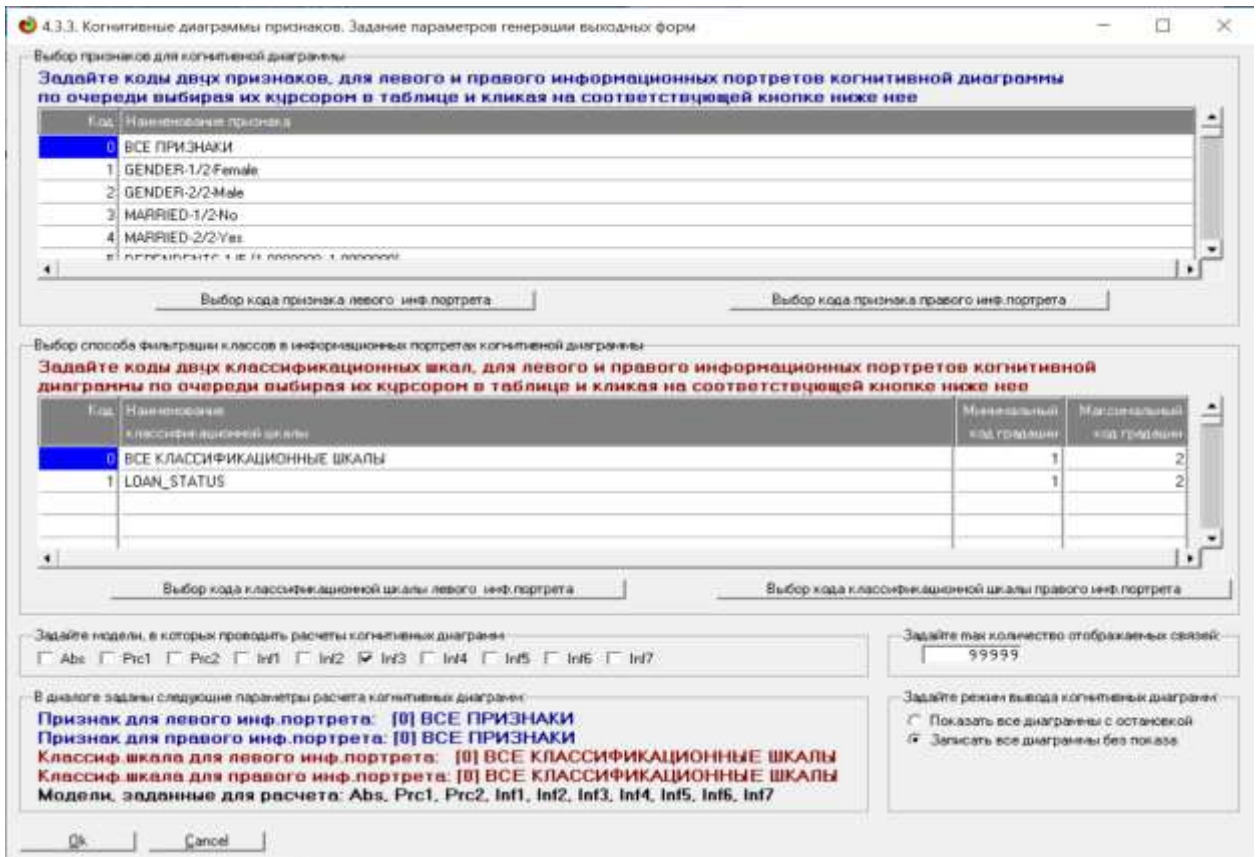


Рисунок 40 – Задание параметров генерации когнитивной диаграммы признаков

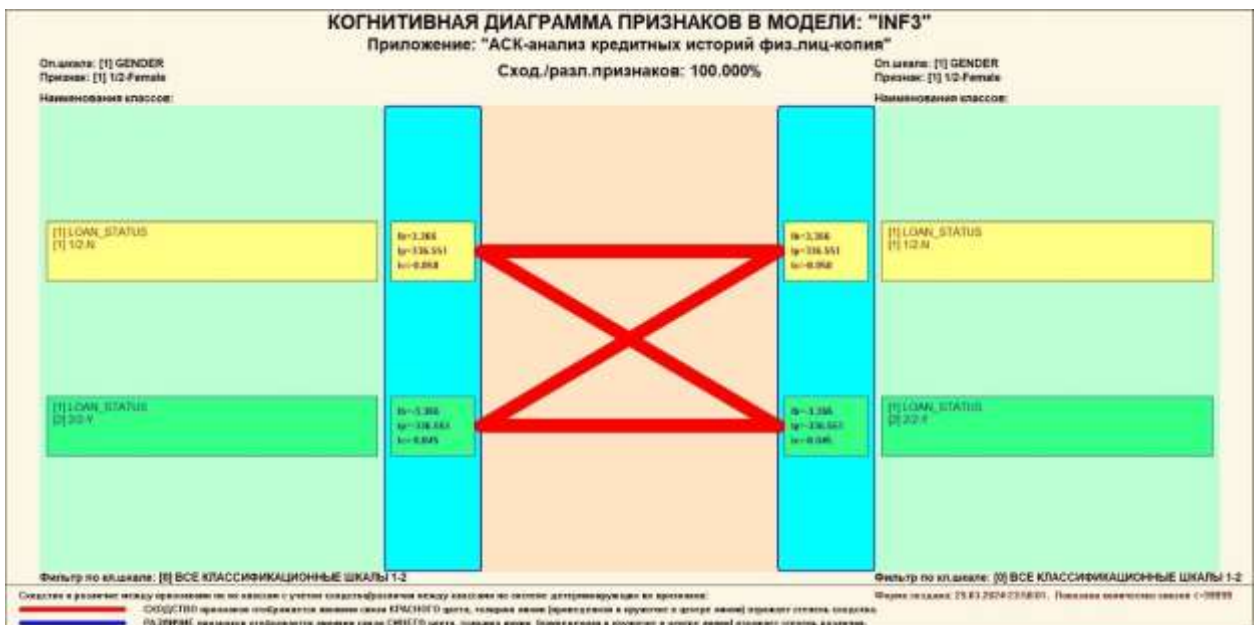


Рисунок 41 – Пример интегральной когнитивной карты содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояние, соответствующей классу в СК-модели INF3

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

3.8.9 Когнитивные функции

Когнитивные функции являются обобщением классического математического понятия функции на основе системной теории информации и предложены Е. В. Луценко в 2005 году.

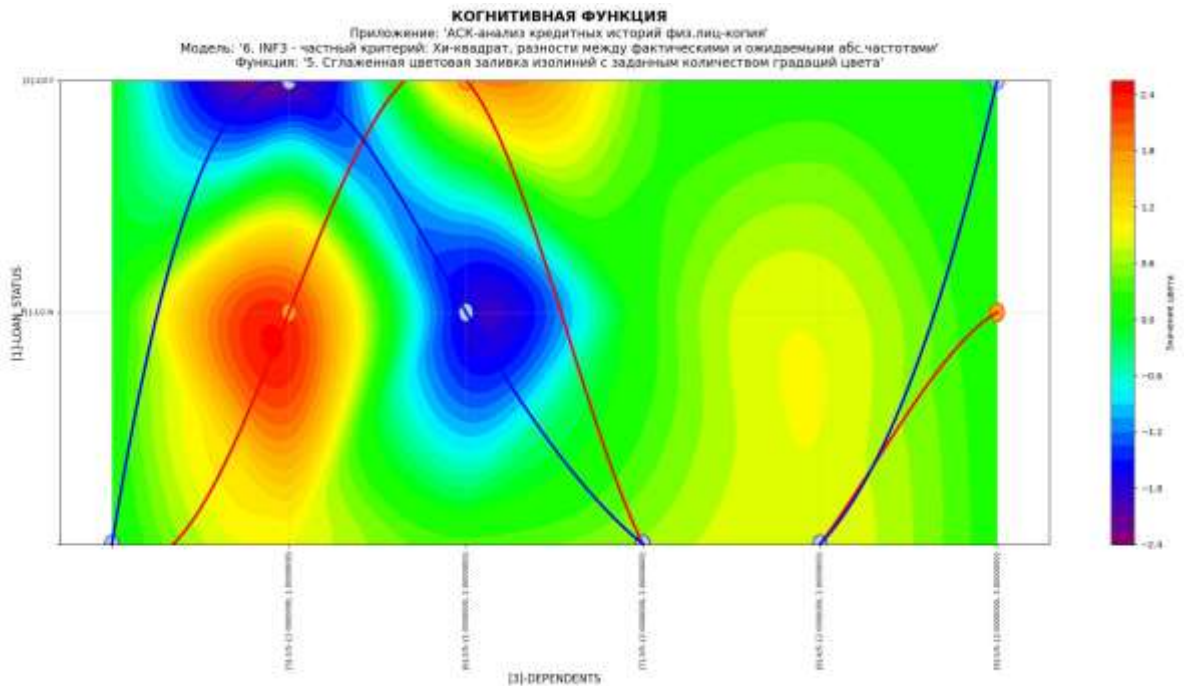
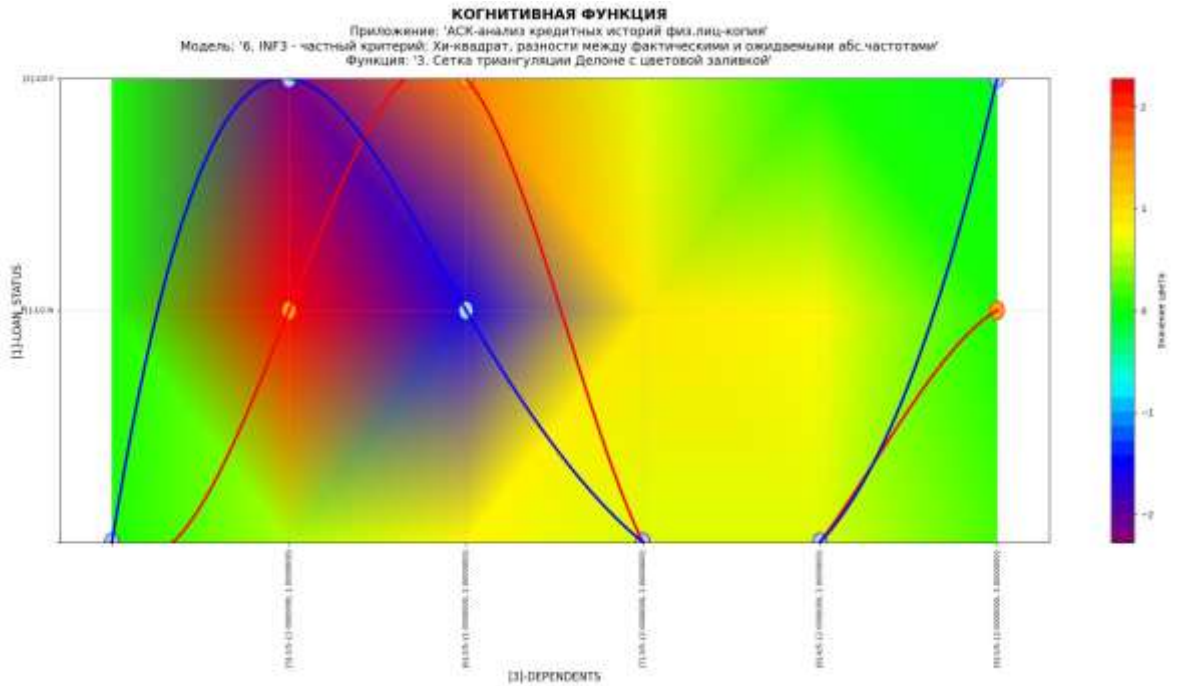
Когнитивные функции отображают, какое количество информации содержится в градациях описательной шкалы о переходе объекта моделирования в состояния, соответствующие градациям классификационной шкалы. При этом в статистических и системно-когнитивных моделях в каждой градации описательной шкалы содержится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

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

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 42). 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, логики и методологии научного познания.

Необходимо отметить, что модели системы «Эйдос» – это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они отражают причинно-следственные

связи, но не отражают механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах.



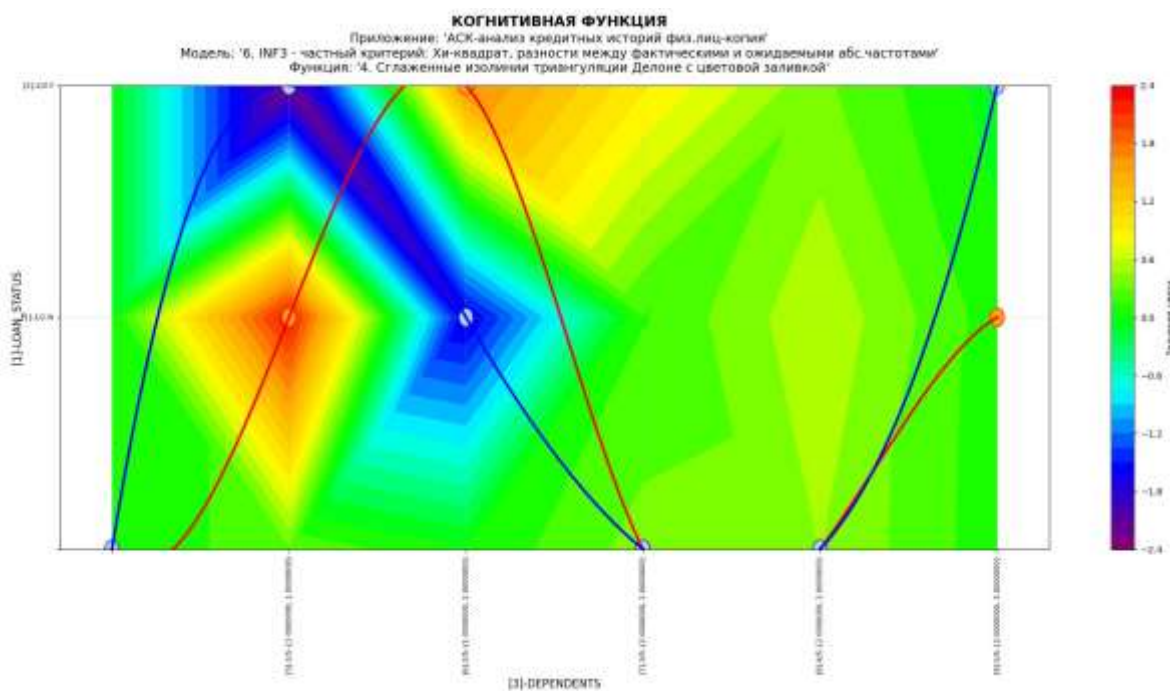


Рисунок 42 – Примеры когнитивных функций в СК-модели INF4

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

3.8.10 Значимость описательных шкал и градаций

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

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

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

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

На рисунке 43 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF3:

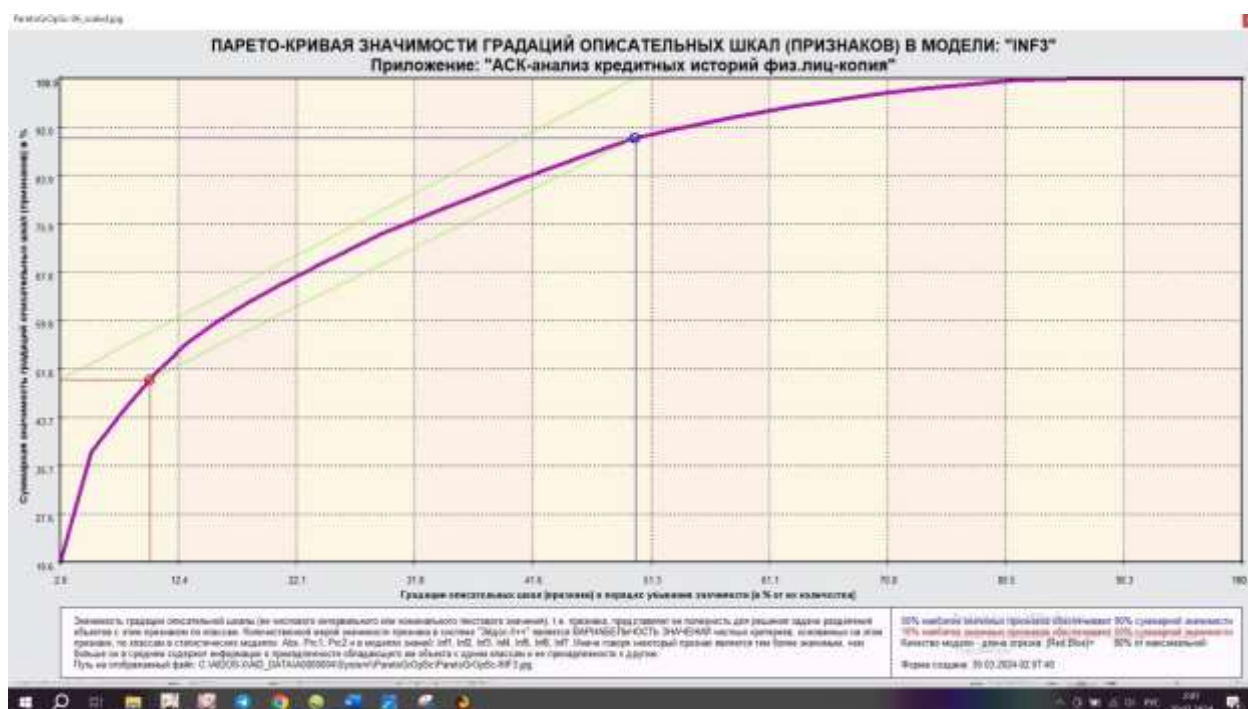


Рисунок 43 – Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF3.

Из рисунка 43 видно, что примерно 10% наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 90% суммарного влияния.

3.8.11 Степень детерминированности классов и классификационных шкал

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

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

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

На рисунке 44 приведена экранная форма режима 3.7.3 системы «Эйдос», содержащая информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



Рисунок 44 – Парето-кривая степени детерминированности классов в СК-модели INF3

4 ОБСУЖДЕНИЕ

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

В данной курсовой работе было показано построение модели зависимости решения по кредитной заявке системой искусственного интеллекта «Эйдос».

При этом наиболее достоверной в данном приложении оказалась модель ABS, при интегральном критерии «Сумма абсолютных частот по признакам».

ЗАКЛЮЧЕНИЕ

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

Системно-когнитивный подход, примененный в данном исследовании, предоставил более глубокий анализ того, на основе чего принимается решение о выдаче кредитов.

Таким образом, результаты данного исследования могут иметь практическое значение как для кредитных организаций, так и для заемщиков-физлиц.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

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

2. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.;

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

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

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

2,062

у.п.л..

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

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

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

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

8. Луценко Е.В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №01(001). С. 79 – 91. – IDA [article ID]: 0010301011. – Режим доступа: <http://ej.kubagro.ru/2003/01/pdf/11.pdf>, 0,812 у.п.л.;

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