

УДК 004.8

1.2.1. Искусственный интеллект и машинное обучение

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ ВЫЖИВШИХ ПРИ КРУШЕНИИ «ТИТАНИКА» ПАССАЖИРОВ

Ефимиади Леонид Константинович

Студент группы ПИ2342

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

В условиях продолжающегося интереса к историческим событиям, особенно к таким трагедиям, как крушение «Титаника» 1912 года, важно углубить понимание факторов, повлиявших на выживаемость пассажиров. В ряде исследований, проведенных специалистами в области статистики, обосновывается необходимость анализа выживаемости на основе таких переменных, как пол, возраст, социальный класс и наличие или отсутствие родственников на борту. Для анализа эмпирических данных использованы традиционные методы. Эти исследования показали, что пол и социальный класс значительно влияли на шансы выживания, в то время как возраст оказывал как положительное, так и отрицательное влияние в зависимости от прочих факторов. В данной статье для анализа эмпирических данных в области статистики и исторических исследований предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исторических исследований, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей в данной предметной области.

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

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

UDC 004.8 UDC 004.8

1.2.1. Artificial Intelligence and Machine Learning

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF TITANIC SURVIVORS

Efimiadi Leonid Konstantinovich

Student of group PI2342

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

With the continuing interest in historical events, especially tragedies such as the sinking of the Titanic in 1912, it is important to gain a better understanding of the factors that influenced the survival of passengers. Several studies conducted by statisticians have argued for the need to analyze survival based on variables such as gender, age, social class, and the presence or absence of relatives on board. Traditional methods were used to analyze the empirical data. These studies showed that gender and social class significantly affected the chances of survival, while age had both a positive and a negative effect depending on other factors. This paper proposes to use Automated System-Cognitive Analysis (ASC-analysis) and its software toolkit, the Eidos intelligent system, to analyze empirical data in the field of statistics and historical research. A detailed numerical example is given. This example contains many different visual tabular and graphical output forms and can be used to teach the use of ASC analysis and the Eidos system for scientific historical research, to develop practical recommendations and substantiate scientific provisions on the mechanisms of action of cause-and-effect relationships in this subject area.

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

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ).....	4
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	4
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	4
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	5
1.4. ЦЕЛЬ РАБОТЫ.....	5
2. METHODS (МЕТОДЫ).....	5
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	5
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	6
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	6
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА.....	8
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	15
3. RESULTS (РЕЗУЛЬТАТЫ).....	17
3.1. Задача-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	17
3.1.1. Универсальная постановка задачи, не зависящая от предметной области	17
3.1.2. Конкретное решение задачи в данной работе	18
3.2. Задача-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	18
3.2.1. Универсальная постановка задачи, не зависящая от предметной области	18
3.2.2. Конкретное решение задачи в данной работе	19
3.3. Задача-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ.....	25
3.3.1. Универсальная постановка задачи, не зависящая от предметной области	25
3.3.2. Конкретное решение задачи в данной работе	33
3.4. Задача-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	36
3.4.1. Универсальная постановка задачи, не зависящая от предметной области	36
3.4.2. Конкретное решение задачи в данной работе	37
3.5. Задача-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	40
3.5.1. Универсальная постановка задачи, не зависящая от предметной области	40
3.5.2. Конкретное решение задачи в данной работе	40
3.6. Задача-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	41
3.6.1. Универсальная постановка задачи, не зависящая от предметной области	41
3.6.1.1. Интегральный критерий «Сумма знаний»	42
3.6.1.2. Интегральный критерий «Семантический резонанс знаний».....	42
3.6.1.3. Важные математические свойства интегральных критериев.....	44
3.6.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»	44
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
3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области	47
3.7.1.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос».....	49
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	50
3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области	50
3.7.2.2. Конкретное решение задачи управления в данной работе в системе «Эйдос»	54
3.8. Задача-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	54
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы).....	54
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области	54
3.8.1.2. Конкретное решение задачи в данной работе	54
3.8.2. Кластерно-конструктивный анализ классов.....	59

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области	59
3.8.2.2. Конкретное решение задачи в данной работе	59
3.8.3. Кластерно-конструктивный анализ значений описательных шкал	61
3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области	61
3.8.3.2. Конкретное решение задачи в данной работе	61
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны.....	64
3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области	64
3.8.4.2. Конкретное решение задачи в данной работе	65
3.8.5. Нелокальная нейронная сеть.....	66
3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области	66
3.8.5.2. Конкретное решение задачи в данной работе	66
3.8.6. 3D-интегральные когнитивные карты	68
3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области	68
3.8.6.2. Конкретное решение задачи в данной работе	68
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	69
3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области	69
3.8.7.2. Конкретное решение задачи в данной работе	69
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	71
3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области	71
3.8.8.2. Конкретное решение задачи в данной работе	71
3.8.9. Когнитивные функции	75
3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области	75
3.8.9.2. Конкретное решение задачи в данной работе	76
3.8.10. Значимость описательных шкал и их градаций.....	79
3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области	79
3.8.10.2. Конкретное решение задачи в данной работе	80
3.8.11. Степень детерминированности классов и классификационных шкал	84
3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области	84
3.8.11.2. Конкретное решение задачи в данной работе	85
4. DISCUSSION (ОБСУЖДЕНИЕ)	87
5. CONCLUSIONS (ВЫВОДЫ)	88
REFERENCES (ЛИТЕРАТУРА)	89

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Данная работа является продолжением серии работ автора по применению Автоматизированного системно-когнитивного анализа (АСК-анализ) для исследования влияния различных факторов на эмерджентные свойства систем в целом [1-4].

Исторические катастрофы, такие, как крушение «Титаника», представляют собой сложные системы, в которых переплетаются различные социальные, психологические и экономические факторы. Важно рассматривать их не только как непосредственные инциденты, но и учитывать контекст, в котором они происходили.

В условиях анализа выживаемости пассажиров «Титаника» критически важно учитывать множество факторов, таких, как пол, возраст, социальный класс и наличие или отсутствие родственников на борту. Эти факторы образуют эмерджентные свойства исследуемой системы и оказывают влияние на шансы выживания.

В ряде исследований, проведенных специалистами в области статистики, обосновывается необходимость анализа выживаемости на основе таких переменных, как пол, возраст, социальный класс и наличие или отсутствие родственников на борту. Для анализа эмпирических данных использованы традиционные методы.

Эти исследования показали, что пол и социальный класс значительно влияли на шансы выживания, в то время как возраст оказывал как положительное, так и отрицательное влияние в зависимости от прочих факторов.

В данной статье для анализа эмпирических данных в области статистики и исторических исследований предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исторических исследований, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей в данной предметной области.

1.2. Объект и предмет исследования

Объект исследования (моделирования) – выживаемость пассажиров «Титаника».

Предмет исследования – выявление причинно-следственных зависимостей фактора выживаемости пассажира от его личных качеств и характеристик.

1.3. Проблема, решаемая в работе и ее актуальность

Спецификой данной задачи является то, что независимые переменные (факторы, влияющие на объект моделирования), имеют очень низкую степень формализации, т.к. формализуются в виде *лингвистических* (категориальных) переменных. При этом зависимая переменная измеряется в числовой шкале.

Таким образом, в работе решается **проблема** построения гибридной модели, включающей как номинальные (текстовые), так и числовые шкалы и обеспечивающей *сопоставимость* обработки данных разных типов, представленных (формализуемых) в разных типах шкал и разных единицах измерения.

Решение в данной работе *проблемы сопоставимости* при выявлении причинно-следственных зависимости фактора выживаемости пассажира от его личных качеств и характеристик, делает данную работу **актуальной**.

1.4. Цель работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели обеспечивается решением ряда *задач* и подзадач, которые являются *этапами* достижения цели. Конкретная формулировка этих задач зависит от метода решения проблемы, поэтому обоснованно мы сформулируем их в конце раздела, т.е. после обоснованного выбора метода решения проблемы и его краткого описания.

2. METHODS (МЕТОДЫ)

2.1. Обоснование требований к методу решения проблемы

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

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

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

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

программный инструментарий, находящийся в полном открытом бесплатном доступе.

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

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

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

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

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

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

АСК-анализ включает:

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

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

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

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

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

Три монографии включены в фонды библиотеки конгресса США⁴.

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

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

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

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

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

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

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

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

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

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

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

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

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

2.4. Система «Эйдос» – инструментарий АСК-анализа

Конечно, на системе «Эйдос» как говорят «Свет клином не сошелся». Существует много очень достойных систем искусственного интеллекта. Чтобы лично убедиться в этом достаточно самостоятельно осуществить поиск в Internet, просто посмотреть файлы: [NCKR-1](#), [NCKR-2](#), [NCKR-3](#), [NCKR-4](#) или пройти по ссылкам: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/how-does-mldotnet-work>, <http://chat.openai.com/>, <https://poe.com/>, <https://neural-university.ru/>, <https://dzen.ru/a/ZCKZRKvrlEMBWOк8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://poe.com/Aidos-X>, <https://rudalle.ru/>, <https://bard.google.com/>, <https://chatbot.theb.ai/>, <https://problembo.com/ru/services> (может пригодиться - почта на 10 минут: <https://10minutemail.net/>), <https://poe.com/GPT-3.5-Turbo-Instruct>, <https://www.seaart.ai/home>, <https://ui.chatai.com/>.

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

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

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

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

- имеет «нулевой порог входа»:

- содержит большое количество интеллектуальных локальных (т.е. поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более **420**, соответственно: http://lc.kubagro.ru/Source_data_applications/WebAppls.htm)

(http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf,

http://lc.kubagro.ru/Presentation_LutsenkoEV.pdf);

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

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

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

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



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

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

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

- хорошо имитирует человеческий стиль мышления и является инструментом познания: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции, если эти эксперты уже есть, а если их еще нет, то она все равно дает верные результаты познания, что будет признано будущими экспертами, когда они появятся;

- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ)

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

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

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

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

В разработке системы «Эйдос» были следующие этапы:

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

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBM-совместимых персональных компьютеров система "Эйдос" впервые реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены свидетельства РосПатента, первые в Краснодарском крае и, возможно, в России на системы искусственного интеллекта (слева приведена титульная видеграмма финальной DOS-версии системы «Эйдос-12.5», июнь 2012 года). С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

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

работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

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

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

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

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

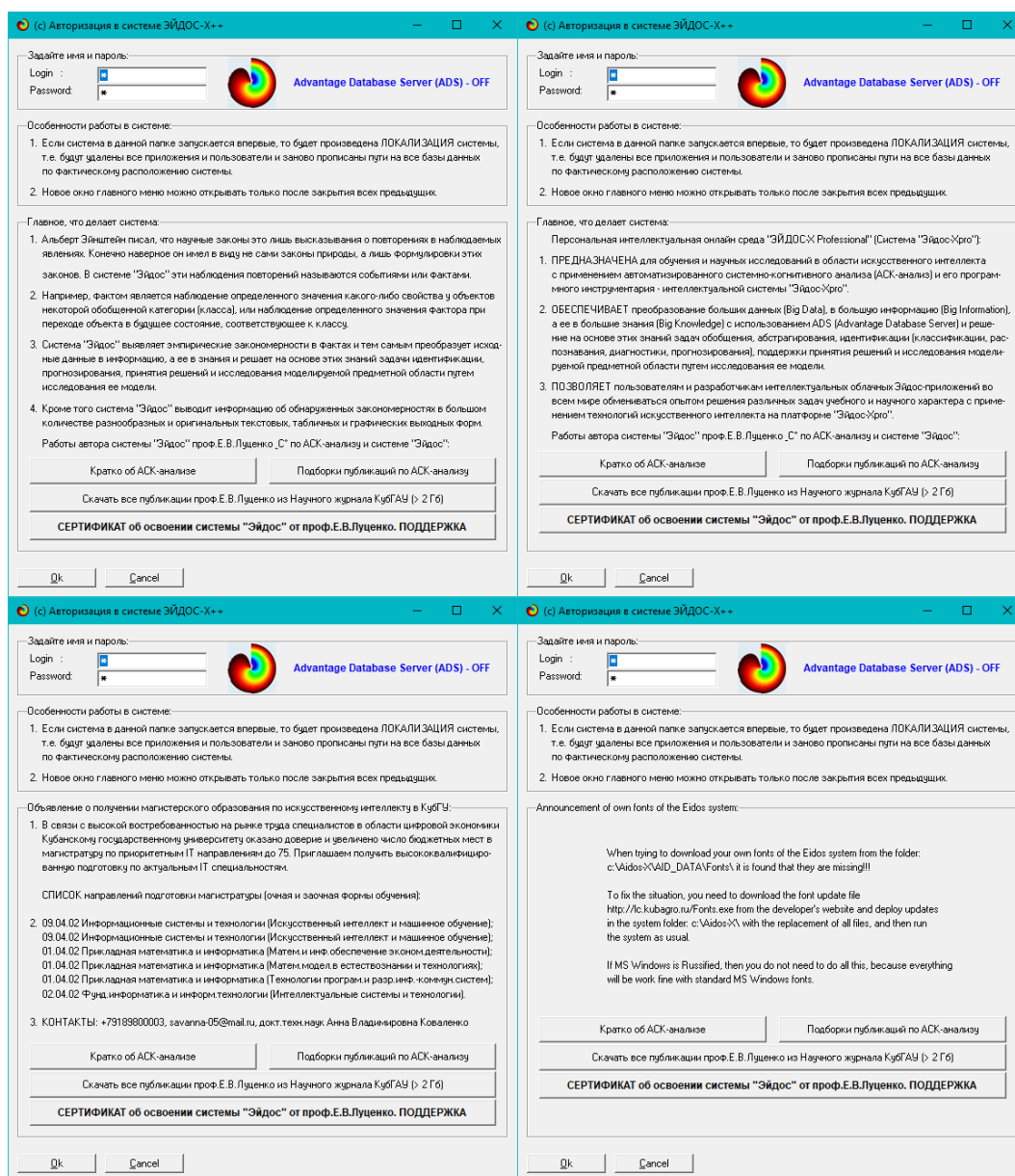
[Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения⁷](#)

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

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



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



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

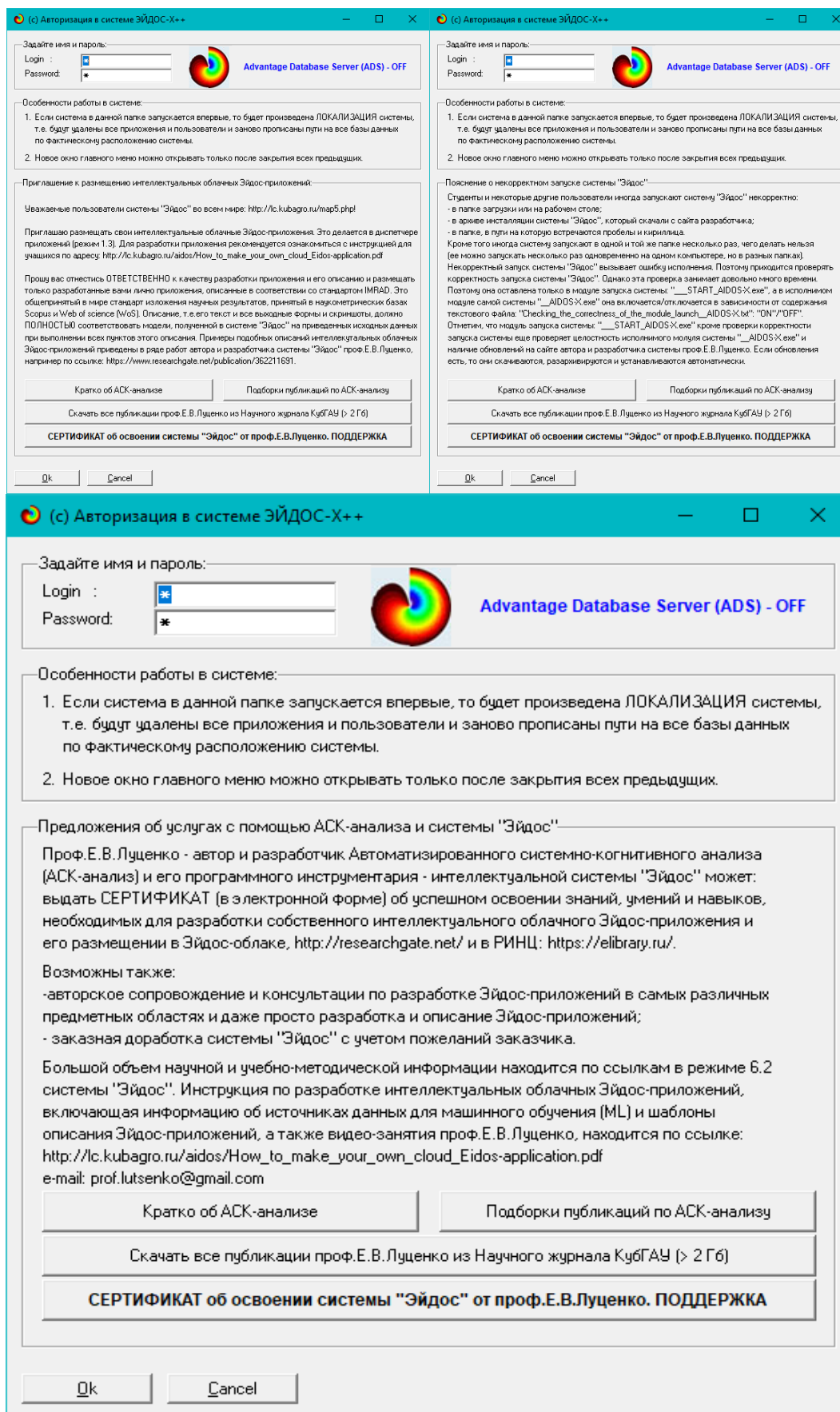


Рисунок 2. Титульные видеограммы текущей версии системы «Эйдос»

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

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

Задача-1. Когнитивная структуризация предметной области.

Задача-2. Формализация предметной области.

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

Многопараметрическая типизация и частные критерии знаний.

Задача-4. Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

Задача-6. Системная идентификация и прогнозирование.

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

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

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

8.2) кластерно-конструктивный анализ классов;

8.3) кластерно-конструктивный анализ значений описательных шкал;

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

8.5) нелокальная нейронная сеть;

8.6) 3d-интегральные когнитивные карты;

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

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

8.9) когнитивные функции;

8.10) значимость описательных шкал и их градаций;

8.11) степень детерминированности классов и классификационных шкал).

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

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

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

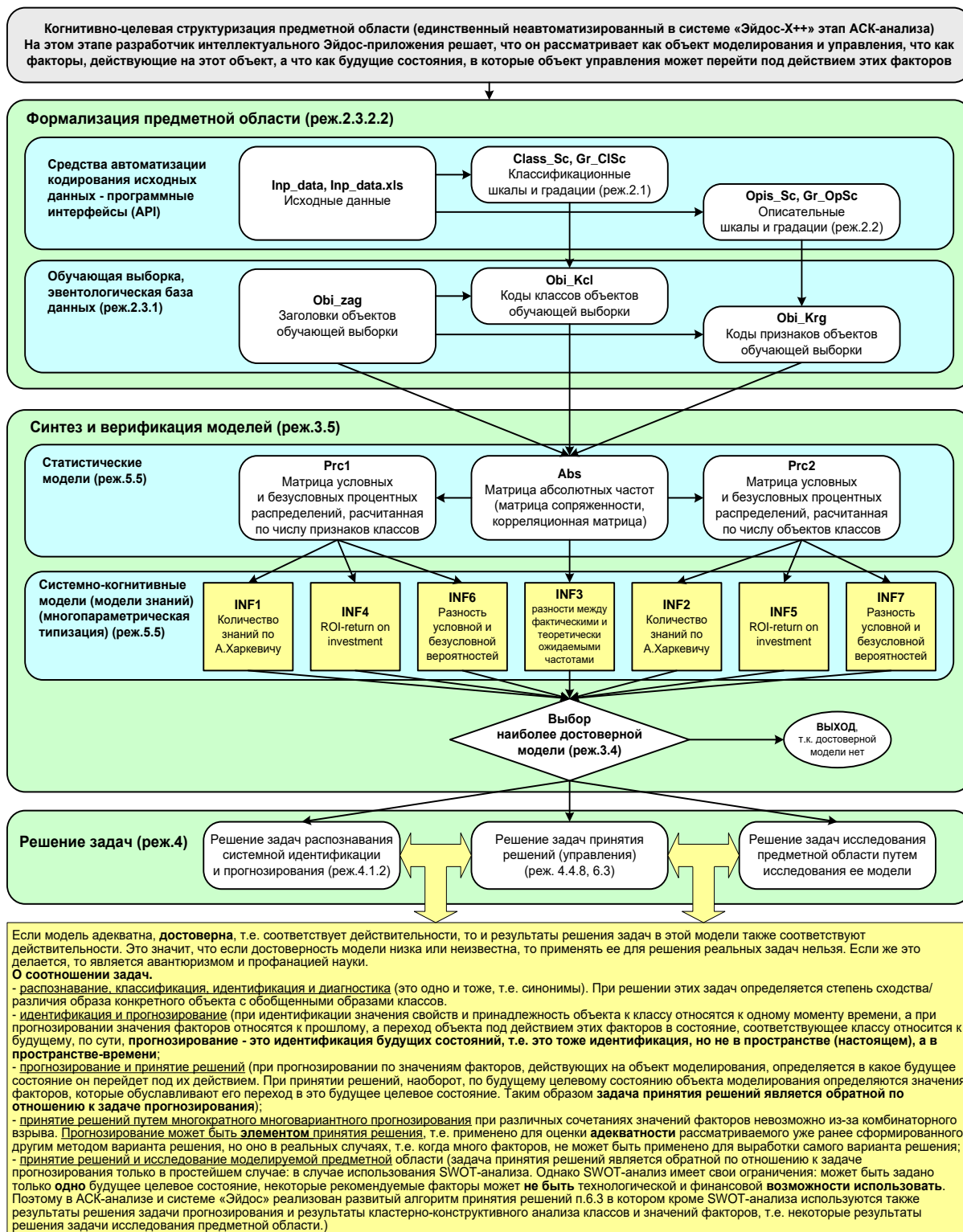


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос» (этапы АСК-анализа)

3. RESULTS (РЕЗУЛЬТАТЫ)

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

3.1.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

Статичная интерпретация и терминология:

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

Динамичная интерпретация и терминология:

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

Обобщающая терминология:

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

3.1.2. Конкретное решение задачи в данной работе

В данной работе в качестве *объекта моделирования* выступает признак выживаемости пассажира, в качестве *факторов*: класс, пол, возраст, наличие родственников на борту (таблица 1), а в качестве *результатов* действия этих факторов анализируется, выжил пассажир или нет (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	КЛАСС
2	ПОЛ
3	ВОЗРАСТ
4	РОДСТВЕННИКИ НА БОРТУ

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Opis_Sc.xlsx

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	ВЫЖИЛ

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Class_Sc.xlsx

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

3.2. Задача-2. Формализация предметной области

3.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

Этим обеспечивается возможность комфортного для пользователя применения системы «Эйдос» для проведения научных исследований и решения практических задач в самых различных предметных областях и

направлениях науки, практически везде, где человек применяет естественный интеллект [48, 49].

3.2.2. Конкретное решение задачи в данной работе

В качестве *источника исходных данных* в данной работе используем портал Kaggle.

Таблица 3 – Исходные данные по влиянию различных факторов на выживаемость пассажиров «Титаника» (фрагмент)

Name	Survived	Pclass	Sex	Age	SibSp or Parch
Wilkes, Mrs. James	yes	Third	female	47	yes
Hirvonen, Mrs. Alexander	yes	Third	female	22	yes
Caldwell, Mr. Albert Francis	no	Second	male	26	yes
Davies, Mr. John Samuel	no	Third	male	21	yes
Snyder, Mrs. John Pillsbury	yes	First	female	23	yes
Howard, Mr. Benjamin	no	Second	male	63	yes
Chaffee, Mrs. Herbert Fuller	yes	First	female	47	yes
del Carlo, Mrs. Sebastiano	yes	Second	female	24	yes
Ilmakangas, Miss. Ida Livija	yes	Third	female	27	yes

Источник: <https://www.kaggle.com/datasets/brendan45774/test-file/data>

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

Таблица 4 – Excel-таблица исходных данных в стандарте системы «Эйдос»

	A	B	C	D	E	F
1	ФИО пассажира	Выжил	Класс	Пол	Возраст	Родственники на борту
2	Wilkes, Mrs. James (Ellen Needs)	да	Третий	жен	47	да
3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	да	Третий	жен	22	да
4	Caldwell, Mr. Albert Francis	нет	Второй	муж	26	да
5	Davies, Mr. John Samuel	нет	Третий	муж	21	да
6	Snyder, Mrs. John Pillsbury (Nelle Stevenson)	да	Первый	жен	23	да
7	Howard, Mr. Benjamin	нет	Второй	муж	63	да
8	Chaffee, Mrs. Herbert Fuller (Carrie Constance Toogood)	да	Первый	жен	47	да
9	del Carlo, Mrs. Sebastiano (Argenia Genovesi)	да	Второй	жен	24	да
10	Ilmakangas, Miss. Ida Livija	да	Третий	жен	27	да
11	Rothschild, Mr. Martin	нет	Первый	муж	55	да
12	Olsen, Master. Artur Karl	нет	Третий	муж	9	да
13	Williams, Mr. Richard Norris II	нет	Первый	муж	21	да
14	Ryerson, Mrs. Arthur Larned (Emily Maria Borie)	да	Первый	жен	48	да
15	Robins, Mr. Alexander A	нет	Третий	муж	50	да
16	Ostby, Miss. Helene Ragnhild	да	Первый	жен	22	да
17	Samaan, Mr. Elias	нет	Третий	муж	18	да
18	Louch, Mr. Charles Alexander	нет	Второй	муж	50	да
19	Jefferys, Mr. Clifford Thomas	нет	Второй	муж	24	да
20	Dean, Mrs. Bertram (Eva Georgetta Light)	да	Третий	жен	33	да
21	Johnston, Mrs. Andrew G (Elizabeth Lily" Watson)"	да	Третий	жен	7	да
22	Mock, Mr. Philipp Edmund	нет	Первый	муж	30	да
23	Karun, Mr. Franz	нет	Третий	муж	39	да
24	Kimball, Mrs. Edwin Nelson Jr (Gertrude Parsons)	да	Первый	жен	45	да
25	Coutts, Mrs. William (Winnie Minnie" Treanor)"	да	Третий	жен	36	да
26	Smith, Mr. Lucien Philip	нет	Первый	муж	24	да
27	Hocking, Miss. Ellen Nellie""	да	Второй	жен	20	да
28	Fortune, Miss. Ethel Flora	да	Первый	жен	28	да
29	Rice, Master. Albert	нет	Третий	муж	10	да
30	Davison, Mr. Thomas Henry	нет	Третий	муж	32	да

Примечание: Разработка автора

Таблица 4 имеет следующую структуру:

- каждая строка описывает одно наблюдение с определенным сочетанием значений факторов и признаком выживаемости пассажира, всего 418 наблюдений;

- каждое *наблюдение* описывается одновременно *двумя* способами: с одной стороны значениями факторов, действующих на объект моделирования (лингвистические и числовые переменные, градации описательных шкал, бесцветный фон в таблице 4), а с другой стороны результатами действия этих факторов, т.е. признаком выживаемости пассажира, выраженными в текстовых и числовых шкалах в разных единицах измерения (желтый фон). Такая структура описания наблюдений в технологиях искусственного интеллекта называется «*онтологией*» и модели представлений знаний Марвина Мински (1975) называется «*фрейм-экземпляр*»;

- 1-я колонка – не является шкалой и содержит ФИО пассажира;

- колонки со 2-й по 2-ю – это классификационные шкалы – это шкалы *текстового* и *числового* типа описывающие *результаты* действия факторов в различных единицах измерения (таблица 4), в данном случае признак выживаемости пассажира. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в *натуральном* и *стоимостном* выражении [8]: например *количество* и *качество* продукции, *прибыль* и *рентабельность*. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

- колонки с 3-й по 6-ю – это описательные шкалы, формализующие факторы, действующие на объект моделирования (таблица 4). Эти шкалы имеют числовой и текстовый тип и их градациями являются лингвистические и числовые переменные;

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

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

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

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

2.3.2. Программные интерфейсы с внешними базами данных

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

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

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

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

Резюме 2.3.2.2: Универсальный программный интерфейс импорта данных из внешней базы данных "Inp_data.xls" в систему "Эйдос" и форматизация предметной области.

Данный программный интерфейс обеспечивает форматизацию предметной области, т.е. анализ файла исходных данных Inp_data.xls, формирование классификационных и описательных шкал и графиков, а затем кодирование файла исходных с их использованием.

*Файл исходных данных должен иметь имя Inp_data.xls и файл расширенной выборки имя Inp_sel.xls. Файлы Inp_data.xls и Inp_sel.xls должны находиться в папке _AIDOS\XND_DATA\Inp_data. Эти файлы имеют совершенно стандартную структуру:

- 1-я строка этого файла должна содержать наименования колонок на любом языке, в т.ч. и русском. Эти наименования должны быть во всех колонках, при этом переводы по словам разрешены, а обозначение ячеек, разрыв строки знак абзаца не допускается. Эти наименования должны быть короткими, но понятными, т.к. они будут в вышеразрешенной форме, а к ним еще будет добавляться наименование графика. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения и число колонок после запятой в колонке должно быть ОДНАКОМБЛОК.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным, до 255 символов.
- Каждая строка этого файла, начиная со 2-й, содержит данные об одном объекте обучающей выборки или одном наблюдении. В MS Excel 2003 в ящике может быть до 65536 строк и до 256 колонок. В ящике MS Excel 2010 и более позднее возможно до 1048576 строк и 16384 колонок.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (именованного / пареваемого) или числового типа (двухзначные значения после запятой).
- Столбцы присваиваются числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбцы присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Столбцы со 2-го по N-й являются классификационными шкалами (выпадающим параметром) и содержат данные о классах (базисном состоянии объекта управления), в котором размещает объект обучающей выборки.
- Столбцы с N+1 по последний являются описательными шкалами (свойствами или факторами) и содержат данные о признаках (т.е. значениях свойства или значения фактора), характеризующие объект обучающей выборки.
- В результате работы режима формируется файл INF_NAME.TXT с расширением MS EXCEL (именованный), в котором наименования классификационных и описательных шкал являются СТРОКАМИ. Система формирует классификационные и описательные шкалы и графики. Для этого в каждом числовом столбце система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего числовые значения заменяются их интервальными значениями. В текстовых столбцах система находит уникальные текстовые значения. Каждое УНИКАЛЬНОЕ интервальное числовое или текстовое значение считается градивной классификационной или описательной шкалой, характеризующей объект. В каждой шкале ее градивы сортируются по алфавиту. С использованием шкал и графиков кодируются исходные данные в результате чего генерируется обучающая выборка, каждый объект которой соответствует одной строке файла исходных данных INF_DATA и содержит коды классов, соответствующие фактам сведениям классификационным или уникальным текстовым значениям классов с градивной классификационной шкалы и коды признаков, соответствующие фактам сведениям числовым или уникальным текстовым значениям признаков с градивной описательной шкалы.
- Разположенные в столбце на основе файла INF_RASP анализируются, за исключением того, что классификационные и описательные шкалы и графики не создаются, а используются ранее созданные в ящике, и базы расположенной выборки могут не включать коды классов, если столбцы классов в файле INF_RASP были пустыми. Структура файла INF_RASP должна быть также же, как INF_DATA, т.е. они должны ПОЛНОСТЬЮ совпадать по наименованиям столбцов, но могут иметь разное количество строк с разными значениями в ящике.

Принцип организации таблицы исходных данных:

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

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

Принцип организации таблицы исходных данных:

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

Рисунок 5. Желпы API-2.3.2.2 системы «Эйдос»

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

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

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

Задайте тип файла исходных данных: "Inp_data":

XLS - MS Excel-2003 Стандарт XLS-файла

XLSX - MS Excel-2007(2010) Стандарт XLS-файла

DBF - DBASE IV (DBF/NTX) Стандарт DBF-файла

CSV - CSV => DBF конвертер Стандарт CSV-файла

Задайте параметры:

Нули и пробелы считать ОТСУТСТВИЕМ данных

Нули и пробелы считать ЗНАЧЕНИЯМИ данных

Создавать БД средних по классам "Inp_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

Формализации предметной области (на основе "Inp_data")

Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

Равные интервалы с разным числом наблюдений

Разные интервалы с равным числом наблюдений

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

Не применять сценарный метод АСК-анализа Применить сценарный метод АСК-анализа

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

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

Интерпретация TXT-полей признаков:

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

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

Только интервальные числовые значения (например: "1/3-(59873.0000000, 178545.6666667)")

Только наименования интервальных числовых значений (например: "Минимальное")

И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-(59873.0000000, 178545.6666667)")

Ok Cancel

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

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (адаптивные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [2 классов x 19 признаков]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	0	0	0,00	1	12	12,00
Текстовые	1	2	2,00	3	7	2,33
ВСЕГО:	1	2	2,00	4	19	4,75

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

В описательных шкалах:

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

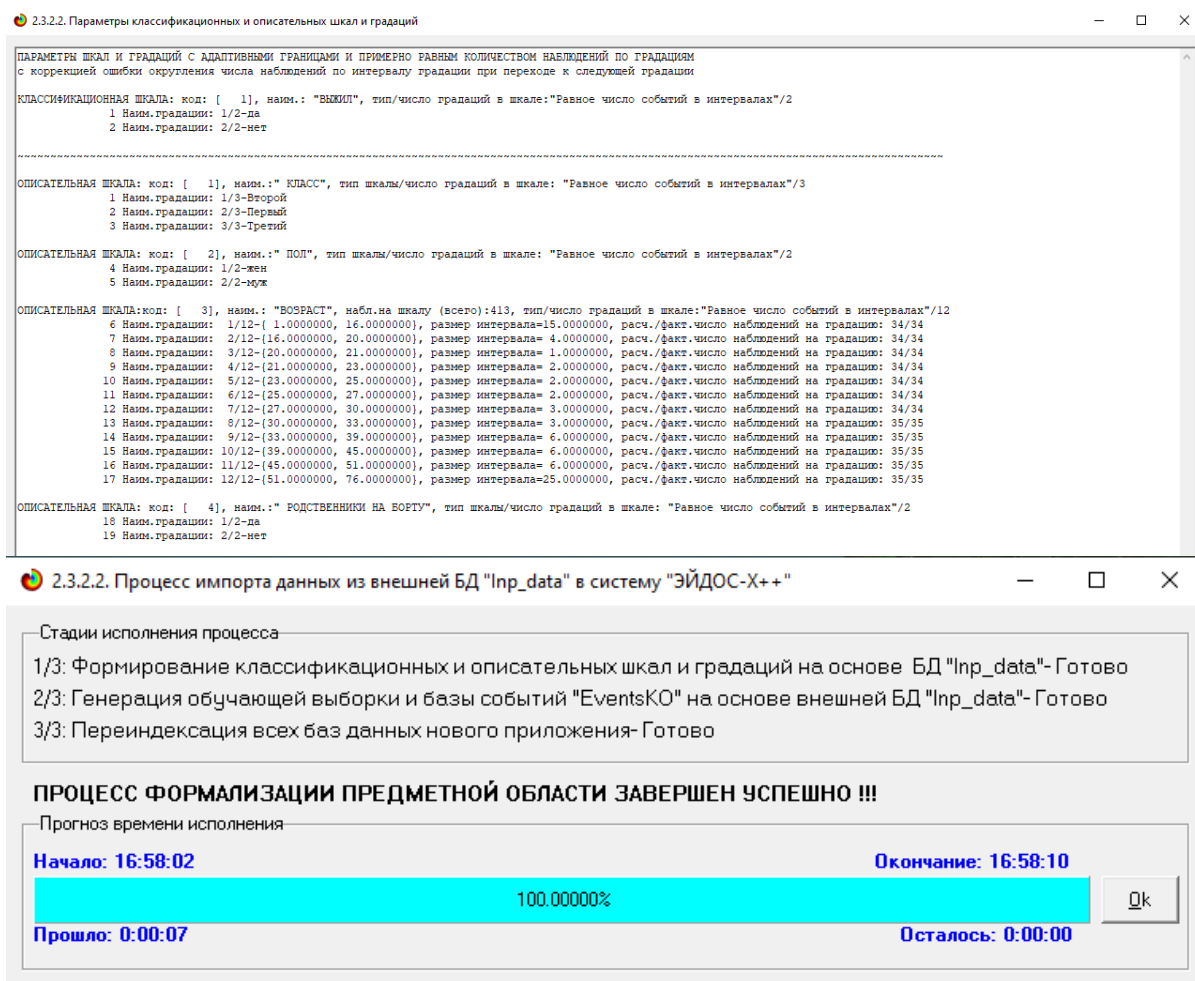


Рисунок 6. Экранные формы управления API-2.3.2.2 системы «Эйдос»

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

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

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

Таблица 5 – Классификационные шкалы и градации (полностью)

KOD_CLS	NAME_CLS
1	ВЫЖИЛ-1/2-да
2	ВЫЖИЛ-2/2-нет

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Classes.xlsx

Таблица 6 – Описательные шкалы и градации (полностью)

KOD_ATR	NAME_ATR
1	КЛАСС-1/3-Второй
2	КЛАСС-2/3-Первый
3	КЛАСС-3/3-Третий
4	ПОЛ-1/2-жен
5	ПОЛ-2/2-муж
6	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}
7	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}
8	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}
9	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}
10	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}
11	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}
12	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}
13	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}
14	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}
15	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}
16	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}
17	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}
18	РОДСТВЕННИКИ НА БОРТУ-1/2-да
19	РОДСТВЕННИКИ НА БОРТУ-2/2-нет

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Attributes.xlsx

Таблица 7 – Обучающая выборка (фрагмент)

NAME_OBJ	N2	N3	N4	N5	N6
Wilkes, Mrs. James (Ellen Needs)	1	3	4	16	18
Hirvonen, Mrs. Alexander (Helga E Lindqvist)	1	3	4	9	18
Caldwell, Mr. Albert Francis	2	1	5	11	18
Davies, Mr. John Samuel	2	3	5	8	18
Snyder, Mrs. John Pillsbury (Nelle Stevenson)	1	2	4	9	18
Howard, Mr. Benjamin	2	1	5	17	18
Chaffee, Mrs. Herbert Fuller (Carrie Constance Toogood)	1	2	4	16	18
del Carlo, Mrs.	1	1	4	10	18

Sebastiano (Argenia Genovesi)					
Ilmakangas, Miss. Ida Livija	1	3	4	11	18
Rothschild, Mr. Martin	2	2	5	17	18
Olsen, Master. Artur Karl	2	3	5	6	18
Williams, Mr. Richard Norris II	2	2	5	8	18
Ryerson, Mrs. Arthur Larned (Emily Maria Borie)	1	2	4	16	18
Robins, Mr. Alexander A	2	3	5	16	18
Ostby, Miss. Helene Ragnhild	1	2	4	9	18
Samaan, Mr. Elias	2	3	5	7	18
Louch, Mr. Charles Alexander	2	1	5	16	18
Jefferys, Mr. Clifford Thomas	2	1	5	10	18
Dean, Mrs. Bertram (Eva Georgetta Light)	1	3	4	13	18

Источник: c:\Aidos-X\AID_DATA\A0000004\System\EventsKO.xlsx

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

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

3.3.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

Суть математической модели АСК-анализа состоит в следующем.

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

Таблица 8 – Матрица абсолютных частот (статистическая модель ABS)

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

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

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

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

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

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

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

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

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

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

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

В таблице 10 приведены формулы:

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

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

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

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	Через относительные частоты	Через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке (таблица 7)	$N_{ij} - \text{фактическая частота,}$ $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота.}$	
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	---	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij} N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предьявлен объект j -го класса, то у него будет обнаружен i -й признак.	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij} N}{N_i N_j}$
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	---	$I_{ij} = N_{ij} - \bar{N}_{ij} = N_{ij} - \frac{N_i N_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$

Обозначения к таблице:*i* – значение прошлого параметра;*j* – значение будущего параметра;*N_{ij}* – количество встреч *j*-го значения будущего параметра при *i*-м значении прошлого параметра;*M* – суммарное число значений всех прошлых параметров;*W* – суммарное число значений всех будущих параметров.*N_i* – количество встреч *i*-м значения прошлого параметра по всей выборке;*N_j* – количество встреч *j*-го значения будущего параметра по всей выборке;*N* – количество встреч *j*-го значения будущего параметра при *i*-м значении прошлого параметра по всей выборке.*I_{ij}* – частный критерий знаний: количество знаний в факте наблюдения *i*-го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее *j*-му значению будущего параметра;*Ψ* – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;*P_i* – безусловная относительная частота встречи *i*-го значения прошлого параметра в обучающей выборке;*P_{ij}* – условная относительная частота встречи *i*-го значения прошлого параметра при *j*-м значении будущего параметра.**Таблица 11 – Матрица системно-когнитивной модели**

		Классы					Значимость фактора
		1	...	<i>j</i>	...	<i>W</i>	
Значения факторов	1	<i>I₁₁</i>		<i>I_{1j}</i>		<i>I_{1W}</i>	$\sigma_{1\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{1j} - \bar{I}_1)^2}$
	...						
	<i>i</i>	<i>I_{i1}</i>		<i>I_{ij}</i>		<i>I_{iW}</i>	$\sigma_{i\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
	...						
<i>M</i>	<i>I_{M1}</i>		<i>I_{Mj}</i>		<i>I_{MW}</i>	$\sigma_{M\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{Mj} - \bar{I}_M)^2}$	
Степень редукции класса	$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$	

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

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

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

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

Когда мы сравниваем фактические и теоретические абсолютные частоты путем вычитания у нас получается частный критерий знаний: «хи-квадрат» (СК-модель INF3), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2) или «коэффициент возврата

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

инвестиций ROI» - Return On Investment (СК-модели INF4, INF5) в зависимости от способа нормировки.

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

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

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

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

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

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

Таблица 13 – Уточнение терминологии АСК-анализа

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

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

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

частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1. Это и есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра, содержащий минимальное количество моделей, позволяющих полно описать моделируемую предметную область.

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

3.3.2. Конкретное решение задачи в данной работе

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

3.5. Синтез и верификация моделей

— □ ×

Задайте модели для синтеза и верификации

Статистические базы

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

Задайте источник данных для расчета модели ABS:

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

Задайте значение фона в матрице абсолютных частот:

2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса

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

Системно-когнитивные модели (базы знаний):

4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1

5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2

6. INF3 - частный критерий: χ -квадрат, разности между фактическими и ожидаемыми абс. частотами

7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1

8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2

9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1

10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

Текущая модель

1. ABS

2. PRC1

3. PRC2

4. INF1

5. INF2

6. INF3

7. INF4

8. INF5

9. INF6

10. INF7

Параметры копирования обучающей выборки в распознаваемую (бутстрепный подход):

Какие объекты обуч.выборки копировать:

Копировать всю обучающую выборку

Копировать только текущий объект

Копировать каждый N-й объект

Копировать N случайных объектов

Копировать объекты от N1 до N2 [fastest]

Вообще не менять распознаваемую выборку

Удалять из обуч.выборки скопированные объекты:

Не удалять

Удалять

Пояснение по алгоритму верификации

Подробнее

Измеряется внутренняя достоверн. модели

Выполнить:

Синтез и верификацию

Только верификацию

Только синтез (xBase++)

Только синтез (Python)

Задайте процессор:

CPU GPU

Задайте алгоритм:

Классика - дольше

Упрощенно-быстрее

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

Расчетный размер БД результатов распознавания Rasp.dbf равен 41934 байт, т.е.: 0.0019555 % от MAX-возможного, (от 2Гб)

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

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

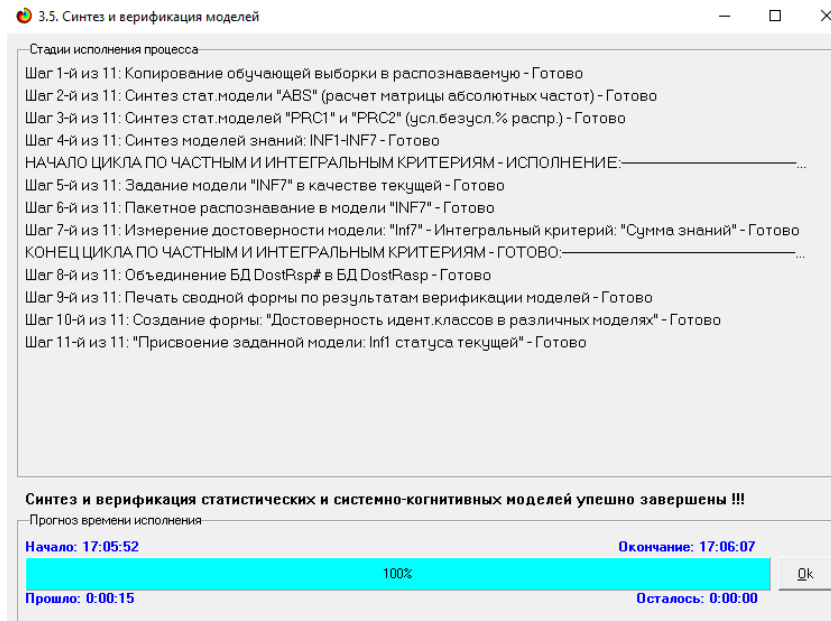


Рисунок 7. Экранные формы режима синтеза и верификации моделей

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

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

Код признака	Наименование описательной шкалы и градации	1. ВЫЖИЛ 1/2 ДА	2. ВЫЖИЛ 2/2 НЕТ	Сумма	Среднее	Средн. квадр.
1.0	КЛАСС-1/3-Второй	30.0	63.0	93.0	46.50	23.33
2.0	КЛАСС-2/3-Первый	50.0	57.0	107.0	53.50	4.95
3.0	КЛАСС-3/3-Третий	72.0	146.0	218.0	109.00	52.33
4.0	ПОЛ-1/2-жен	152.0		152.0	76.00	107.48
5.0	ПОЛ-2/2-муж		266.0	266.0	133.00	188.09
6.0	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}	15.0	20.0	35.0	17.50	3.54
7.0	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}	20.0	25.0	45.0	22.50	3.54
8.0	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}	7.0	17.0	24.0	12.00	7.07
9.0	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}	16.0	21.0	37.0	18.50	3.54
10.0	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}	8.0	27.0	35.0	17.50	13.44
11.0	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}	10.0	20.0	30.0	15.00	7.07
12.0	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}	13.0	27.0	40.0	20.00	9.90
13.0	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}	10.0	20.0	30.0	15.00	7.07
14.0	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}	15.0	22.0	37.0	18.50	4.95
15.0	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}	9.0	29.0	38.0	19.00	14.14
16.0	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}	11.0	16.0	27.0	13.50	3.54
17.0	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}	16.0	19.0	35.0	17.50	2.12
18.0	РОДСТВЕННИКИ НА БОРТУ-1/2-да	84.0	81.0	165.0	82.50	2.12
19.0	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	68.0	185.0	253.0	126.50	82.73
	Сумма числа признаков	606.0	1061.0	1667.0		
	Среднее	31.9	55.8		43.87	
	Среднеквадратичное отклонение	38.1	69.4			56.55
	Сумма числа объектов обуч.выборки	152.0	266.0	418.0		

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

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

Код признака	Наименование описательной шкалы и градации	1. ВЫЖИЛ 1/2 ДА	2. ВЫЖИЛ 2/2 НЕТ	Безусл. вероятн.	Среднее	Средн. квадр.
1.0	КЛАСС-1/3-Второй	19.737	23.684	22.249	21.711	2.914
2.0	КЛАСС-2/3-Первый	32.895	21.429	25.598	27.162	8.232
3.0	КЛАСС-3/3-Третий	47.368	54.887	52.153	51.128	5.440
4.0	ПОЛ-1/2-жен	100.000		36.364	50.000	70.836
5.0	ПОЛ-2/2-муж		100.000	63.636	50.000	70.836
6.0	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}	9.868	7.519	8.373	8.694	1.783
7.0	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}	13.158	9.398	10.766	11.278	2.781
8.0	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}	4.605	6.391	5.742	5.498	1.382
9.0	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}	10.526	7.895	8.852	9.211	1.982
10.0	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}	5.263	10.150	8.373	7.707	3.579
11.0	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}	6.579	7.519	7.177	7.049	0.780
12.0	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}	8.553	10.150	9.569	9.352	1.249
13.0	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}	6.579	7.519	7.177	7.049	0.780
14.0	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}	9.868	8.271	8.852	9.070	1.249
15.0	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}	5.921	10.902	9.091	8.412	3.645
16.0	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}	7.237	6.015	6.459	6.626	0.982
17.0	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}	10.526	7.143	8.373	8.835	2.515
18.0	РОДСТВЕННИКИ НА БОРТУ-1/2-да	55.263	30.451	39.474	42.857	17.670
19.0	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	44.737	69.549	60.526	57.143	17.670
	Сумма	398.684	398.872	797.556		
	Среднее	20.983	20.993		20.988	
	Среднеквадратичное отклонение	25.052	26.110			25.232
	Сумма числа объектов обуч.выборки	152.000	266.000	418.000		

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

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"

Код признака	Наименование описательной шкалы и градации	1. ВЫЖИЛ 1/2 ДА	2. ВЫЖИЛ 2/2 НЕТ	Сумма	Среднее	Средн. квадр.
1.0	КЛАСС-1/3-Второй	-0.016	0.008	-0.008	-0.004	0.017
2.0	КЛАСС-2/3-Первый	0.034	-0.024	0.010	0.005	0.041
3.0	КЛАСС-3/3-Третий	-0.013	0.007	-0.006	-0.003	0.014
4.0	ПОЛ-1/2-жен	0.136		0.136	0.068	0.096
5.0	ПОЛ-2/2-муж		0.061	0.061	0.030	0.043
6.0	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}	0.022	-0.015	0.008	0.004	0.026
7.0	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}	0.027	-0.018	0.009	0.004	0.032
8.0	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}	-0.030	0.014	-0.015	-0.008	0.031
9.0	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}	0.023	-0.015	0.008	0.004	0.027
10.0	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}	-0.063	0.026	-0.037	-0.018	0.063
11.0	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}	-0.012	0.006	-0.005	-0.003	0.013
12.0	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}	-0.015	0.008	-0.007	-0.004	0.016
13.0	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}	-0.012	0.006	-0.005	-0.003	0.013
14.0	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}	0.015	-0.009	0.006	0.003	0.017
15.0	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}	-0.058	0.024	-0.033	-0.017	0.058
16.0	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}	0.015	-0.010	0.006	0.003	0.018
17.0	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}	0.031	-0.021	0.009	0.005	0.037
18.0	РОДСТВЕННИКИ НА БОРТУ-1/2-да	0.045	-0.035	0.010	0.005	0.057
19.0	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	-0.041	0.019	-0.022	-0.011	0.042
	Сумма	0.091	0.033	0.124		
	Среднее	0.005	0.002		0.003	
	Среднеквадратичное отклонение	0.044	0.022			0.035
	Сумма числа объектов обуч.выборки	152.000	266.000	418.000		

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

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

Код признака	Наименование описательной шкалы и градации	1. ВЫЖИЛ 1/2 ДА	2. ВЫЖИЛ 2/2 НЕТ	Сумма	Среднее	Средн. квадр.
1.0	КЛАСС-1/3-Второй	-3.808	3.808			5.385
2.0	КЛАСС-2/3-Первый	11.103	-11.103			15.701
3.0	КЛАСС-3/3-Третий	-7.249	7.249			10.252
4.0	ПОЛ-1/2-жен	96.744	-96.744			136.816
5.0	ПОЛ-2/2-муж	-96.698	96.698			136.752
6.0	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}	2.277	-2.277			3.220
7.0	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}	3.641	-3.641			5.150
8.0	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}	-1.725	1.725			2.439
9.0	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}	2.549	-2.549			3.606
10.0	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}	-4.723	4.723			6.680
11.0	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}	-0.906	0.906			1.281
12.0	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}	-1.541	1.541			2.179
13.0	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}	-0.906	0.906			1.281
14.0	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}	1.549	-1.549			2.191
15.0	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}	-4.814	4.814			6.808
16.0	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}	1.185	-1.185			1.676
17.0	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}	3.277	-3.277			4.634
18.0	РОДСТВЕННИКИ НА БОРТУ-1/2-да	24.018	-24.018			33.967
19.0	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	-23.972	23.972			33.902
	Сумма					
	Среднее					
	Среднеквадратичное отклонение	33.453	33.453			32.998
	Сумма числа объектов обуч. выборки	152.000	266.000	418.000		

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

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

3.4. Задача-4. Верификация моделей

3.4.1. Универсальная постановка задачи, не зависящая от предметной области

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных и ложных положительных и отрицательных решений по F-мере Ван 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 [9].

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

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

3.4.2. Конкретное решение задачи в данной работе

В соответствии с критериями достоверности F-мерой Ван Ризбергена и L1 и L2 -мерами проф.Е.В.Луценко наиболее достоверными являются СК-модели INF3 (хи-квадрат К.Пирсона) с интегральными критериями: «Сумма знаний» и «Резонанс знаний»: все эти меры равны 1 при максимуме 1 (рисунок 12). *Эту СК-модель INF3 модель с интегральным критерием «Сумма знаний» и будем использовать для решения поставленных в работе задач.*

3.4. Обобщ.форма по достов.моделей при разнит.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Юлнота модели	F-мера Ван Ризбергена	Сумма модул. уровней сход. истинно-полож. решений (STP)	Сумма модул. уровней сход. истинно-отриц. решений (STN)	Сумма модул. уровней сход. ложно-полож. решений (SFP)	Сумма модул. уровней сход. ложно-отрицат. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	1.000	0.735	291.633	9.997	35.151		0.892	1.000	0.943
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс. частот по признак...	1.000	0.667	305.823		121.179		0.716	1.000	0.835
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн. частот с о...	1.000	0.735	291.633	9.997	35.151		0.892	1.000	0.943
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн. частот по приз...	1.000	0.667	360.292		166.671		0.684	1.000	0.812
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн. частот с о...	1.000	0.735	291.635	9.998	35.151		0.892	1.000	0.943
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн. частот по приз...	1.000	0.667	360.232		166.621		0.684	1.000	0.812
4. INF1 - частный критерий: количество знаний по А\Харкевичу: в...	Семантический резонанс зна...	0.964	0.862	224.457	104.394	18.707	2.913	0.923	0.987	0.954
4. INF1 - частный критерий: количество знаний по А\Харкевичу: в...	Сумма знаний	0.964	0.835	162.388	60.180	21.299	0.937	0.884	0.994	0.936
5. INF2 - частный критерий: количество знаний по А\Харкевичу: в...	Семантический резонанс зна...	0.964	0.862	224.470	104.385	18.709	2.912	0.923	0.987	0.954
5. INF2 - частный критерий: количество знаний по А\Харкевичу: в...	Сумма знаний	0.964	0.835	162.489	60.281	21.286	0.932	0.884	0.994	0.936
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	1.000	1.000	316.561	316.561			1.000	1.000	1.000
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Сумма знаний	1.000	1.000	325.167	325.167			1.000	1.000	1.000
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.974	0.896	254.637	101.469	9.867	0.789	0.963	0.997	0.980
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	1.000	0.846	170.960	33.903	18.091		0.904	1.000	0.950
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.974	0.896	254.648	101.467	9.868	0.789	0.963	0.997	0.980
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	1.000	0.846	171.065	33.965	18.084		0.904	1.000	0.950
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	1.000	0.903	294.079	107.205	7.141		0.976	1.000	0.988
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	1.000	0.849	233.562	46.666	21.141		0.917	1.000	0.957
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	1.000	0.903	294.139	107.244	7.148		0.976	1.000	0.988
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	1.000	0.849	233.726	46.743	21.144		0.917	1.000	0.957

Рисунок 12. Экранные формы режима измерения достоверности моделей 3.4

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



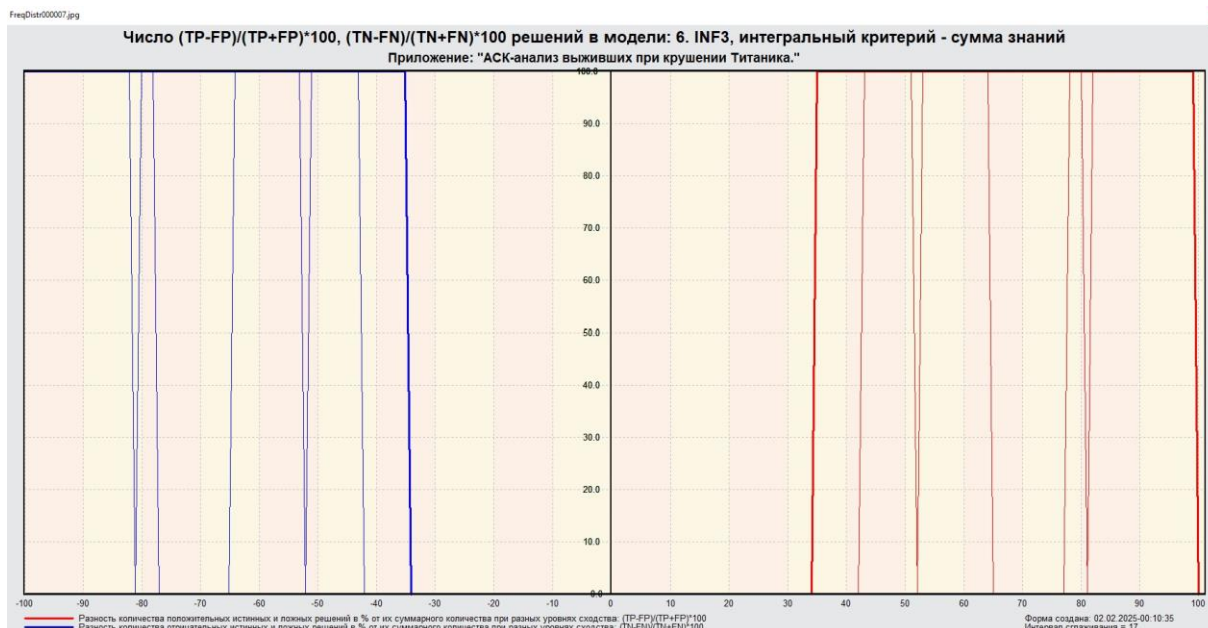


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF3 [9]

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

- отрицательные ложные решения в данной задаче вообще не встречаются не встречаются, а в общем случае такие решения практически не встречаются за исключением очень небольшого количества случаев с очень низкими уровнями различия;

- при уровнях сходства меньше 20% в данной задаче преобладают ложные положительные решения, а при более высоких уровнях сходства – истинные положительные решения. В общем случае при уровнях сходства выше 60% ложных положительных решений практически вообще нет;

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

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

Помощь по режимам: 3.4, 4.1.3.6, 4.1.3.7, 4.1.3.8, 4.1.3.10: Виды прогнозов и меры достоверности моделей в системе "Эйдос-X++"

ПОЛОЖИТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Предположим, модель дает такой прогноз, что выпадет все: и 1, и 2, и 3, и 4, и 5, и 6. Понятно, что из всего этого выпадет лишь что-то одно. В этом случае модель не предскажет, что не выпадет, но зато она обязательно предскажет, что выпадет. Однако при этом очень много объектов будет отнесено к классам, к которым они не относятся. Тогда вероятность истинно-положительных решений у модели будет 1/6, а вероятность ложно-положительных решений - 5/6. Ясно, что такой прогноз бесполезен, поэтому он и назван мной псевдопрогнозом.

ОТРИЦАТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Представим себе, что мы выбрасываем кубик с 6 гранями, и модель предсказывает, что ничего не выпадет, т.е. не выпадет ни 1, ни 2, ни 3, ни 4, ни 5, ни 6, но что-то из этого, естественно, обязательно выпадет. Конечно, модель не предсказала, что выпадет, зато она очень хорошо предсказала, что не выпадет. Вероятность истинно-отрицательных решений у модели будет 5/6, а вероятность ложно-отрицательных решений - 1/6. Такой прогноз гораздо достовернее, чем положительный псевдопрогноз, но тоже бесполезен.

ИДЕАЛЬНЫЙ ПРОГНОЗ.
Если в случае с кубиком мы прогнозируем, что выпадет, например 1, и соответственно прогнозируем, что не выпадет 2, 3, 4, 5, и 6, то это идеальный прогноз, имеющий, если он осуществляется, 100% достоверность идентификации и не идентификации. Идеальный прогноз, который полностью снимает неопределенность о будущем состоянии объекта прогнозирования, на практике удается получить крайне редко и обычно мы имеем дело с реальным прогнозом.

РЕАЛЬНЫЙ ПРОГНОЗ.
На практике мы чаще всего сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой. Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у Вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов.

Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно относить объекты к классам, к которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Этот критерий предложен и реализован в системе "Эйдос" проф. Е.В. Луценко в 1994 году. Эта мера достоверности модели предполагает два варианта нормировки: {-1,+1} и {0,1}:

$$L_a = \frac{TP + TN - FP - FN}{TP + TN + FP + FN} \quad (\text{нормировка: } \{-1,+1\})$$

$$L_b = \frac{1 + (TP + TN - FP - FN)}{(TP + TN + FP + FN)} / 2 \quad (\text{нормировка: } \{0,1\})$$

где количество: TP - истинно-положительных решений; TN - истинно-отрицательных решений; FP - ложно-положительных решений; FN - ложно-отрицательных решений;

Классическая F-мера достоверности моделей Ван Ризбергера (колонка выделена ярко-голубым фоном):
 $F\text{-мера} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$
Precision = TP / (TP + FP) - точность модели;
Recall = TP / (TP + FN) - полнота модели;

L1-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом SUMM уровней сходства (колонка выделена ярко-зеленым фоном):
 $L1\text{-мера} = 2 * (S * \text{Precision} * S * \text{Recall}) / (S * \text{Precision} + S * \text{Recall})$
S Precision = STP / (STP + SFP) - точность с учетом сумм уровней сходства;
S Recall = STP / (STP + SFN) - полнота с учетом сумм уровней сходства;
STP - Сумма модулей сходства истинно-положительных решений; STN - Сумма модулей сходства истинно-отрицательных решений;
SFP - Сумма модулей сходства ложно-положительных решений; SFN - Сумма модулей сходства ложно-отрицательных решений.

L2-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СРЕДНИХ уровней сходства (колонка выделена желтым фоном):
 $L2\text{-мера} = 2 * (AP * \text{Precision} * A * \text{Recall}) / (AP * \text{Precision} + A * \text{Recall})$
APrecision = ATP / (ATP + AFP) - точность с учетом средних уровней сходства;
ARecall = ATP / (ATP + AFN) - полнота с учетом средних уровней сходства;
ATP = STP / TP - Среднее модулей сходства истинно-положительных решений; AFN = SFN / FN - Среднее модулей сходства истинно-отрицательных решений;
AFP = SFP / FP - Среднее модулей сходства ложно-положительных решений; AFN = SFN / FN - Среднее модулей сходства ложно-отрицательных решений.

Строки с максимальными значениями F-меры, L1-меры и L2-меры выделены фоном цвета, соответствующего колонке.

Из графиков частотных распределений истинно-положительных, истинно-отрицательных, ложно-положительных и ложно-отрицательных решений видно, что чем выше модуль уровня сходства, тем больше доля истинных решений. Это значит, что модуль уровня сходства является адекватной мерой степени истинности решения и степени уверенности системы в этом решении. Поэтому система "Эйдос" имеет адекватный критерий достоверности собственных решений, с помощью которого она может отфильтровать заведомо ложные решения.

Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета [Научный журнал КубГАУ] [Электронный ресурс] - Краснодар: КубГАУ, 2017. - №02(126). С. 1 - 32. - IDA [article ID]: 1261702001. - Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 у.п.л.

Помощь по режиму 3.4. (С) Система "ЭЙДОС-X++"

Режим: 4.1.3.11. РАСЧЕТ И ГРАФИЧЕСКАЯ ВИЗУАЛИЗАЦИЯ ЧАСТОТНЫХ РАСПРЕДЕЛЕНИЙ УРОВНЕЙ СХОДСТВА:

По нажатию кнопок: [TP;TN;FP;FN], [(TP;FP);(TN;FN)], [(T-F)/(T+F)*100] отображаются графики частотных распределений для модели и интегрального критерия той строки, на которой в экранной форме 3.4 стоит курсор. По клику на кнопки: [(T-F)/(T+F)*100] выводятся графики частотных распределений: (TP-FP)/(TP+FP)*100 и (TN-FN)/(TN+FN)*100, где:
TP-True-Positive; TN-True-Negative; FP-False Positive; FN-False-Negative, количество истинных и ложных положительных и отрицательных решений.

Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета [Научный журнал КубГАУ] [Электронный ресурс] - Краснодар: КубГАУ, 2017. - №02(126). С. 1 - 32. - IDA [article ID]: 1261702001. - Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 у.п.л.

Примерные графики TP,TN,FP,FN, а также F-меры и критериев L1, L2 при увеличении объема выборки:

The figure consists of two side-by-side graphs. The left graph, titled 'Количество решений' (Number of solutions), plots TP, TN, FP, and FN against 'Объем выборки' (Sample size). All four metrics increase linearly with sample size. The right graph, titled 'Динамика F-меры Ван Ризбергера и L1- и L2- мер проф.Е.В.Луценко в зависимости от объема выборки RND-модели' (Dynamics of F-мера, L1- and L2- measures of Prof. E.V. Lutsenko depending on the volume of RND-model selection), plots 'Достоверность, %' (Reliability, %) against 'Число объектов обучающей выборки' (Number of objects in the training sample). It shows three curves: L2-MERA (blue), L1-MERA (magenta), and F-MERA (yellow). All three curves start at 1.00 for small sample sizes and decrease as the sample size increases, eventually stabilizing around 0.65-0.70.

Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

3.5.1. Универсальная постановка задачи, не зависящая от предметной области

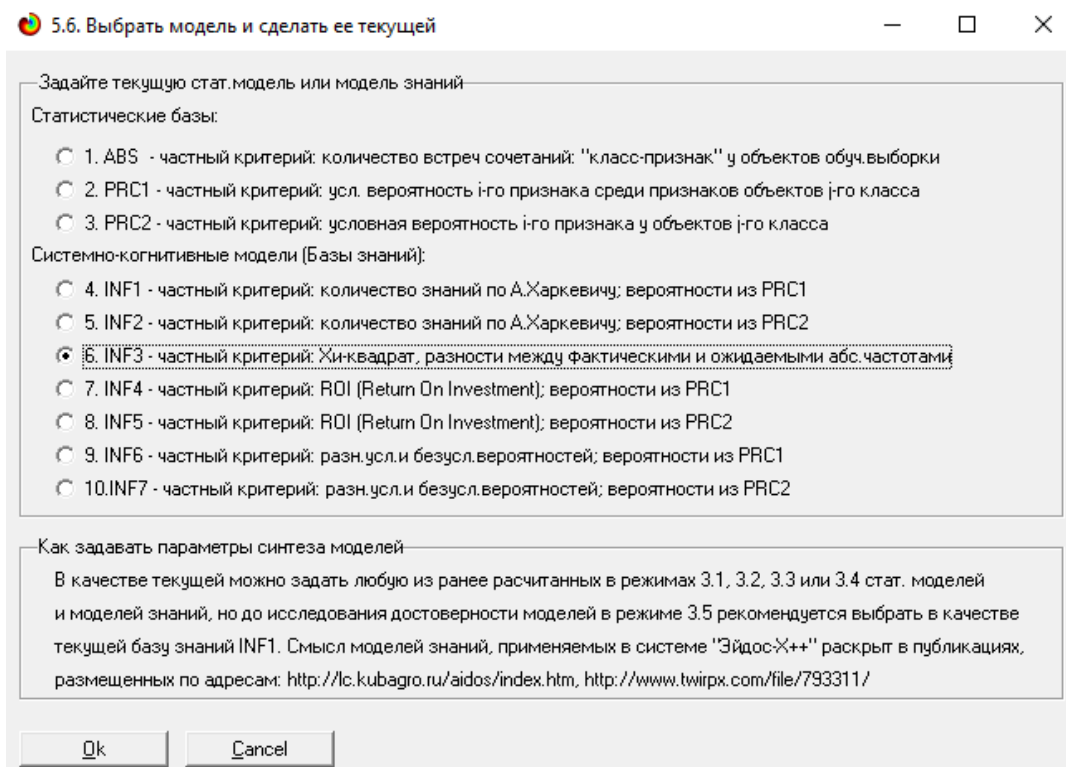
Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

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

3.5.2. Конкретное решение задачи в данной работе

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



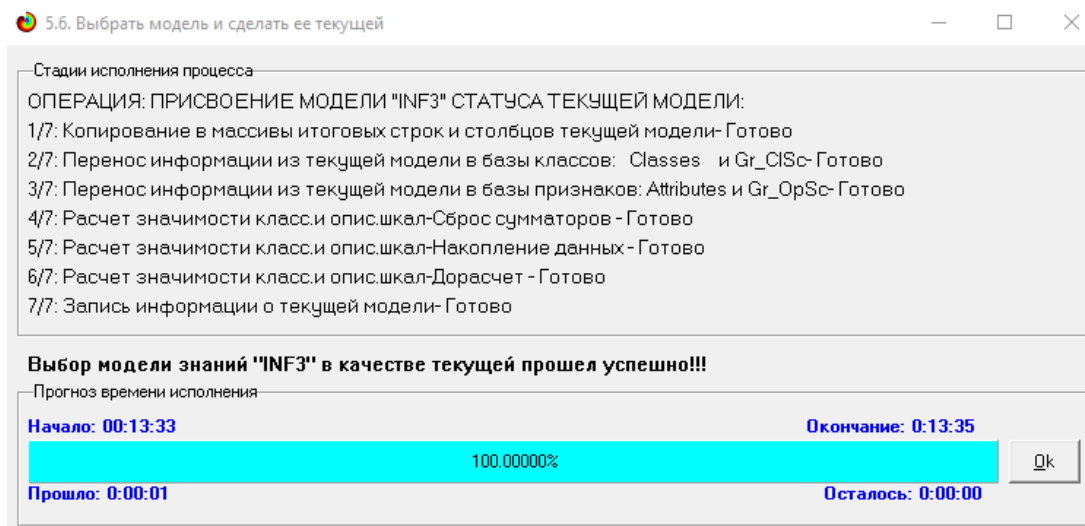


Рисунок 15. Задание СК-модели INF3 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

3.6.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

Сравнение осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что

корректны¹⁰ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1.1. Интегральный критерий «Сумма знаний»

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

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

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

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

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

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

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

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

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

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

3.6.1.2. Интегральный критерий «Семантический резонанс знаний»

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

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

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

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

где:

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

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

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

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

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

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

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

Поэтому по своей сути он также является скалярным

произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применения вейвлетов и сплайнов, в частности линейной интерполяции: $I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, \quad L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}}$, Это позволяет предложить

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

3.6.1.3. Важные математические свойства интегральных критериев

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

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

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

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

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

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

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

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

В АСК-анализе разработаны, а в системе «Эйдос» реализованы развитые методы прогнозирования, основанные на сценарном методе АСК-анализа или сценарном АСК-анализе [11]. Но в задачи данной работы не входит их подробное рассмотрение, тем более, что они подробно

освящены и на теоретическом уровне, и с детальными численными примерами в работах [4-7, 11] и в ряде других [48].

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

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

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

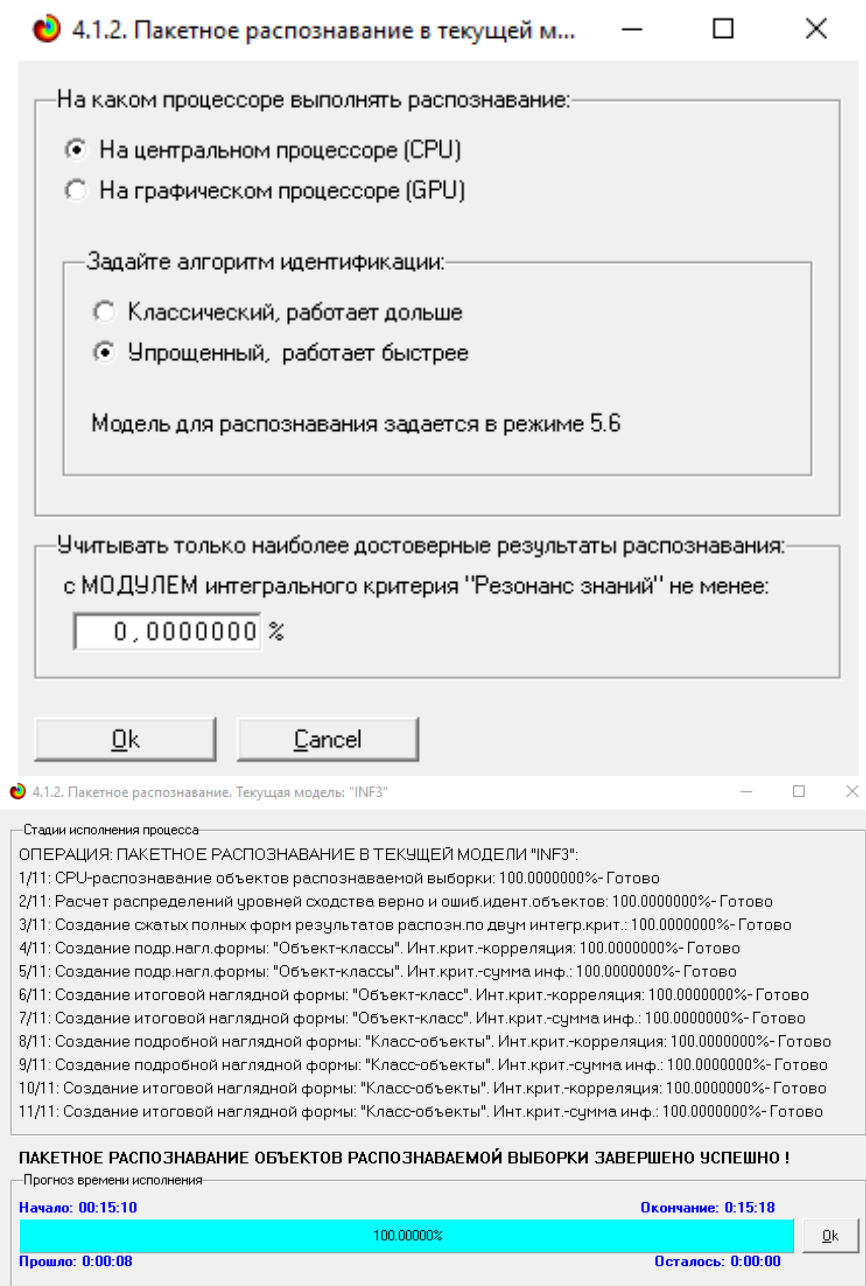


Рисунок 16. Экранные формы режима 4.1.2 идентификации и прогнозирования

- 4.1.3.1. Подробно наглядно: "Объект - классы"
 - 4.1.3.2. Подробно наглядно: "Класс - объекты"
 - 4.1.3.3. Итоги наглядно: "Объект - класс"
 - 4.1.3.4. Итоги наглядно: "Класс - объект"
 - 4.1.3.5. Подробно сжато: "Объекты - классы"
-
- 4.1.3.6. Обобщ. форма по достов. моделей при разных интегральных крит.
 - 4.1.3.7. Обобщ. стат. анализ результатов идент. по моделям и инт. крит.
 - 4.1.3.8. Стат. анализ результ. идент. по классам, моделям и инт. крит.
 - 4.1.3.9. Достоверность идент. объектов при разных моделях и инт. крит.
 - 4.1.3.10. Достоверность идент. классов при разных моделях и инт. крит.
 - 4.1.3.11. Объединение в одной БД строк по самым достоверным моделям
 - 4.1.3.12. Вывод результатов распознавания в стиле: "Inp_data.xlsx"
 - 4.1.3.13. Частотное распределение наблюдений по самым похожим классам
 - 4.1.3.14. Распределение уровней сходства наблюдений по всем классам

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

4.1.3.1. Визуализация результатов распознавания в отношении: "Объект-классы". Текущая модель: "INF3"

Расознаваемые объекты		Интегральный критерий сходства: "Семантический резонанс знаний"				
Код	Наим. объекта	Код	Наименование класса	Сходство	Ф...	Сходство
1	Wilkes, Mrs. James (Ellen Nee...	1	ВЫЖИЛ-1/2-да	82,23...	v	
2	Hirvonen, Mrs. Alexander (Hel...	2	ВЫЖИЛ-2/2-нет	-82,23...		
3	Caldwell, Mr. Albert Francis					
4	Davies, Mr. John Samuel					
5	Snyder, Mrs. John Pillsbury (N...					
6	Howard, Mr. Benjamin					
7	Chalfee, Mrs. Herbert Fuller (C...					
8	del Carlo, Mrs. Sebastiano (Arg...					
9	Ilmakangas, Miss. Ida Livija					
10	Rothschild, Mr. Martin					
11	Olsen, Master. Artur Karl					
12	Williams, Mr. Richard Norris II					
13	Ryerson, Mrs. Arthur Larned (E...					
14	Robins, Mr. Alexander A					
15	Ostby, Miss. Helene Ragnhild					
16	Samaan, Mr. Elias					
17	Louch, Mr. Charles Alexander					
18	Jefferys, Mr. Clifford Thomas					
19	Dean, Mrs. Bertram (Eva Geor...					
20	Johnston, Mrs. Andrew G (Eliz...					
21	Mock, Mr. Philipp Edmund					
22	Karun, Mr. Franz					
23	Kimball, Mrs. Edwin Nelson Jr (...)					

Интегральный критерий сходства: "Сумма знаний"	
Код	Наименование класса
1	ВЫЖИЛ-1/2-да
2	ВЫЖИЛ-2/2-нет

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

4.1.3.2. Визуализация результатов распознавания в отношении: "Класс-объекты". Текущая модель: "INF3"

Классы	
Код	Наим. класса
1	ВЫЖИЛ-1/2-да
2	ВЫЖИЛ-2/2-нет

Интегральный критерий сходства: "Семантический резонанс знаний"				
Код	Наименование объекта	Сходство	Ф...	Сходство
102	West, Miss. Barbara J	100,00...	v	[Progress Bar]
72	Marvin, Mrs. Daniel Warner (Mary Graham...	97,15...	v	[Progress Bar]
156	Smith, Mrs. Lucien Philip (Mary Eloise Hug...	97,15...	v	[Progress Bar]
137	Dean, Miss. Elizabeth Gladys Millvina"	97,05...	v	[Progress Bar]
31	Fortune, Mrs. Mark (Mary McDougald)	96,89...	v	[Progress Bar]
32	Cornell, Mrs. Robert Clifford (Malvina Hele...	96,89...	v	[Progress Bar]
39	Cavendish, Mrs. Tyrell William (Julia Floren...	96,89...	v	[Progress Bar]
43	Straus, Mrs. Isidor (Rosalie Ida Blun)	96,89...	v	[Progress Bar]
70	Compton, Mrs. Alexander Taylor (Mary Eliz...	96,89...	v	[Progress Bar]
120	Crosby, Mrs. Edward Gifford (Catherine Eli...	96,89...	v	[Progress Bar]

Интегральный критерий сходства: "Сумма знаний"				
Код	Наименование объекта	Сходство	Ф...	Сходство
72	Marvin, Mrs. Daniel Warner (Mary Graham...	100,00...	v	[Progress Bar]
156	Smith, Mrs. Lucien Philip (Mary Eloise Hug...	100,00...	v	[Progress Bar]
31	Fortune, Mrs. Mark (Mary McDougald)	99,73...	v	[Progress Bar]
32	Cornell, Mrs. Robert Clifford (Malvina Hele...	99,73...	v	[Progress Bar]
39	Cavendish, Mrs. Tyrell William (Julia Floren...	99,73...	v	[Progress Bar]
43	Straus, Mrs. Isidor (Rosalie Ida Blun)	99,73...	v	[Progress Bar]
70	Compton, Mrs. Alexander Taylor (Mary Eliz...	99,73...	v	[Progress Bar]
120	Crosby, Mrs. Edward Gifford (Catherine Eli...	99,73...	v	[Progress Bar]
133	Cardeza, Mrs. James Warburton Martinez (...)	99,73...	v	[Progress Bar]
138	Brown, Mrs. John Murray (Caroline Lane L...	99,73...	v	[Progress Bar]

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

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

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

3.7. Задача-7. Поддержка принятия решений

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

3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

Таким образом, задача принятия решений является *обратной* по отношению к задаче прогнозирования. Но это так только в простейшем

случае: в случае использования SWOT-анализа (режим 4.4.8 системы «Эйдос») [10] (рисунки 19).

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

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

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

Необходимо подчеркнуть, что в системе «Эйдос» SWOT-диаграммы формируются автоматически на основе статистических и системно-когнитивных моделей, созданных непосредственно на основе эмпирических данных, а не как обычно не формализуемым экспертным путем на основе интуиции, опыта и профессиональной компетенции, т.е. практически «на глазок», а в некоторых случаях и вообще «от фонаря».

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

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

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	ВЫЖИЛ-1/2-да	33,4529026	606	36,3636364
2	ВЫЖИЛ-2/2-нет	33,4529026	1061	63,6363636

SWOT-анализ класса: 1 "ВЫЖИЛ-1/2-да" в модели: 6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	ПОЛ-1/2-жен	96,744
18	РОДСТВЕННИКИ НА БОРТУ-1/2-да	24,018
2	КЛАСС-2/3-Первый	11,103
7	ВОЗРАСТ-2/12-{16,0000000, 20,0000000}	3,641
17	ВОЗРАСТ-12/12-{51,0000000, 76,0000000}	3,277
9	ВОЗРАСТ-4/12-{21,0000000, 23,0000000}	2,549
6	ВОЗРАСТ-1/12-{1,0000000, 16,0000000}	2,277
14	ВОЗРАСТ-9/12-{33,0000000, 39,0000000}	1,549
16	ВОЗРАСТ-11/12-{45,0000000, 51,0000000}	1,185

Препятствующие факторы и сила их влияния

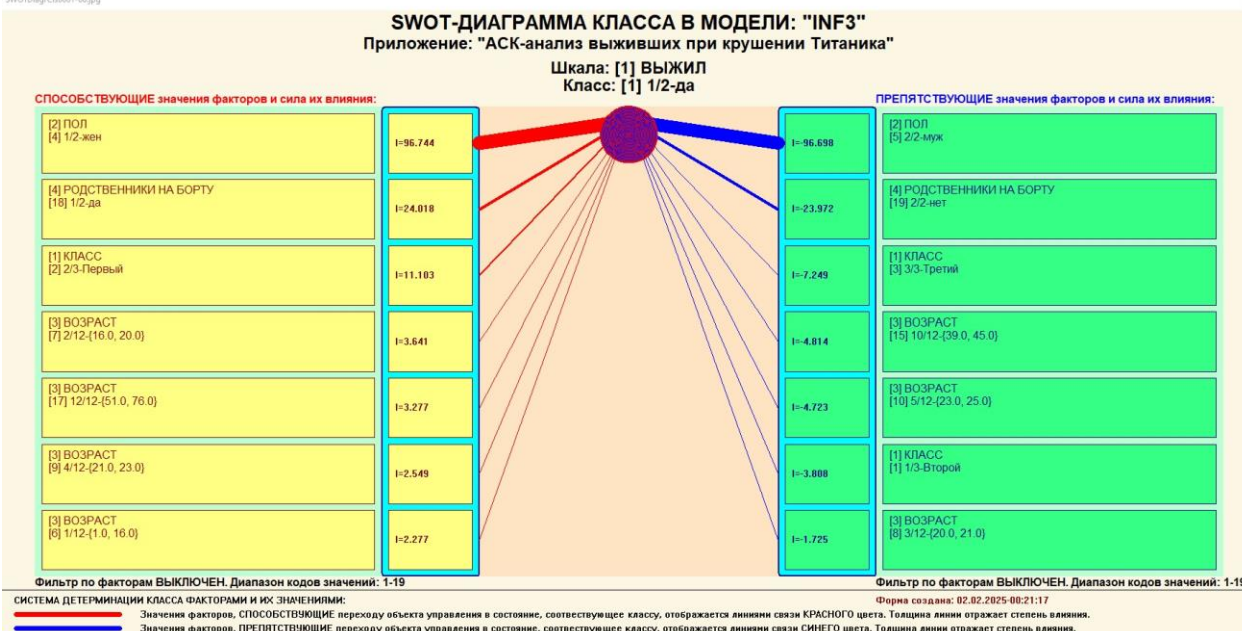
Код	Наименование фактора и его интервального значения	Сила влияния
5	ПОЛ-2/2-муж	-96,698
19	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	-23,972
3	КЛАСС-3/3-Третий	-7,249
15	ВОЗРАСТ-10/12-{39,0000000, 45,0000000}	-4,814
10	ВОЗРАСТ-5/12-{23,0000000, 25,0000000}	-4,723
1	КЛАСС-1/3-Второй	-3,808
8	ВОЗРАСТ-3/12-{20,0000000, 21,0000000}	-1,725
12	ВОЗРАСТ-7/12-{27,0000000, 30,0000000}	-1,541
13	ВОЗРАСТ-8/12-{30,0000000, 33,0000000}	-0,906
11	ВОЗРАСТ-6/12-{25,0000000, 27,0000000}	-0,906

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

SWOT-диаграмма

SWOTdiag\Sk0001-06.jpg



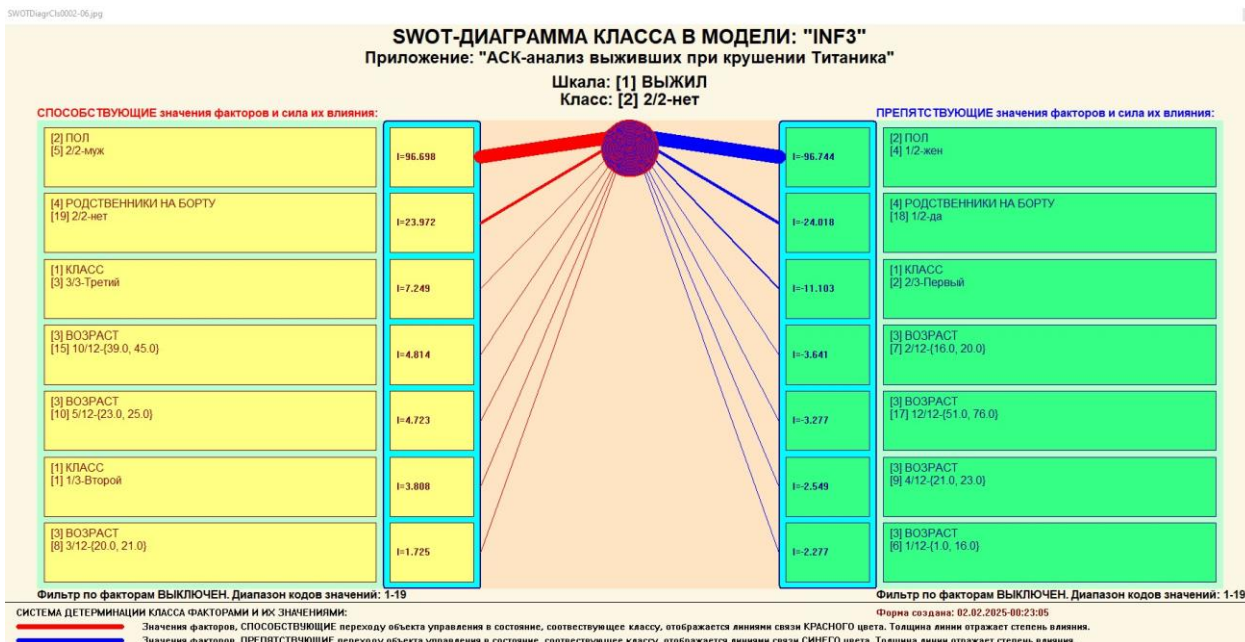


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

Из приведенных выходных форм SWOT-анализа четко видно, что выживаемость пассажиров «Титаника» зависела от ряда ключевых факторов. Способствующие выживанию стали такие факторы, как принадлежность к женскому полу, наличие родственников на борту и путешествие в первом классе. В то же время, пассажиры мужского пола, путешествовавшие в одиночку в третьем классе, были подвержены наибольшим рискам, которые препятствовали их выживаемости.

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

3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

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

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

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

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

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

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

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

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

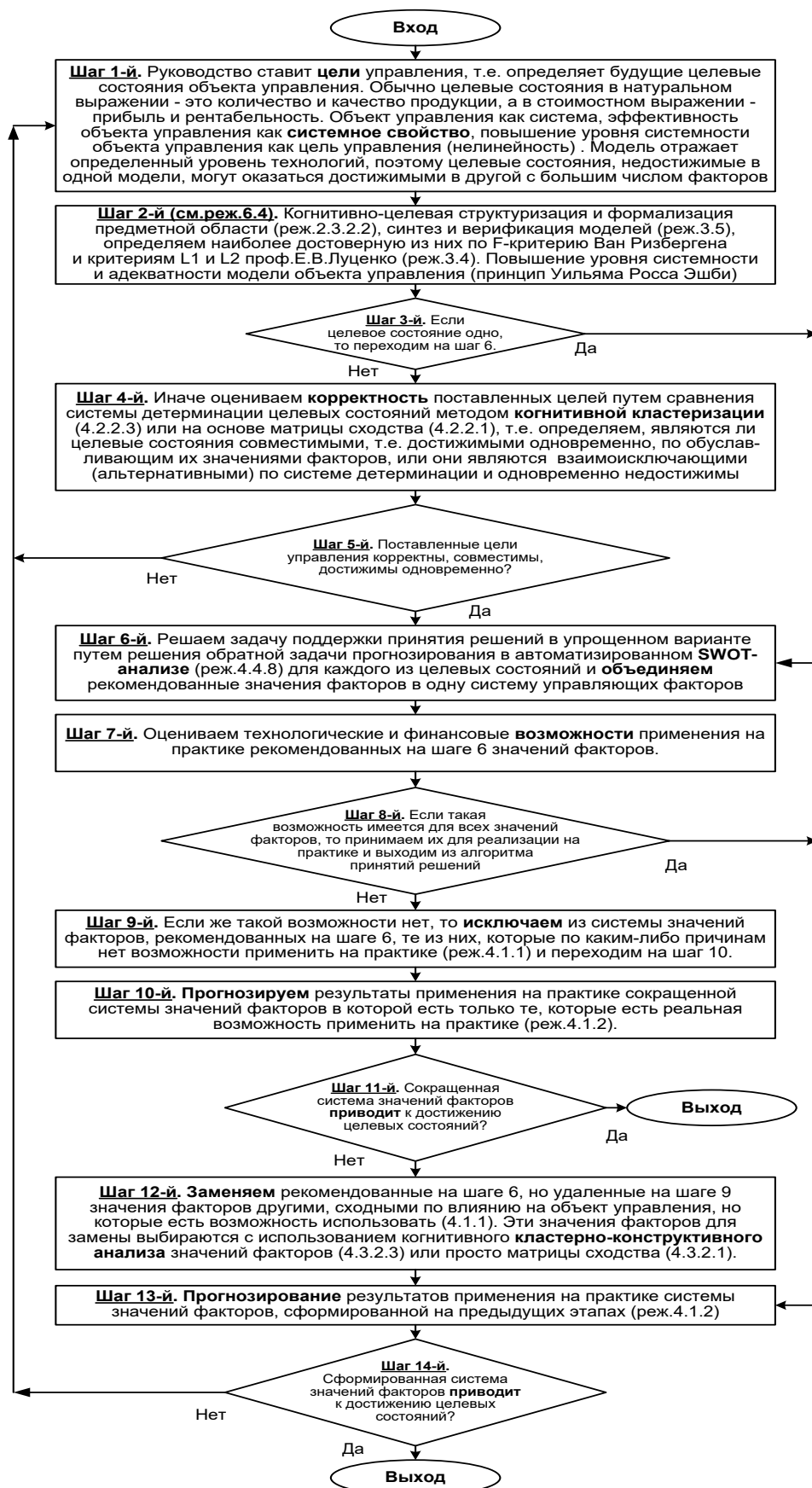


Рисунок 20. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

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

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

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

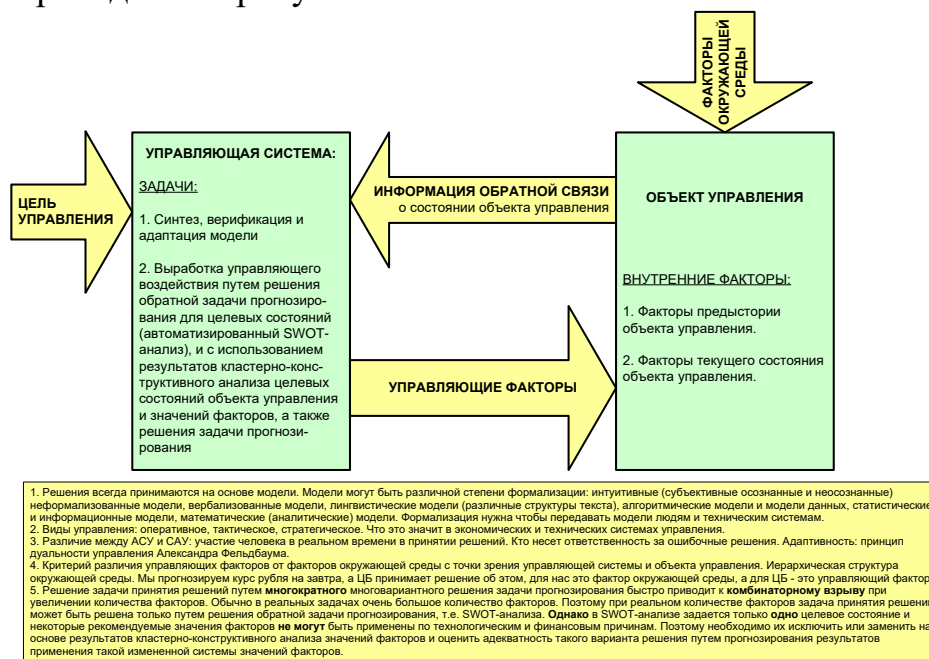


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

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

3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области

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

3.8.1.2. Конкретное решение задачи в данной работе

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

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

Код	Наименование значения фактора
1	КЛАСС-1/3-Второй
2	КЛАСС-2/3-Первый
3	КЛАСС-3/3-Третий
4	ПОЛ-1/2-жен
5	ПОЛ-2/2-муж
6	ВОЗРАСТ-1/12-(1.00000000, 16.00000000)

SWOT-анализ значения фактора: 1 "КЛАСС-1/3-Второй" в модели: 6 "INF3"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влияния
2	ВЫЖИЛ-2/2-нет	3.808

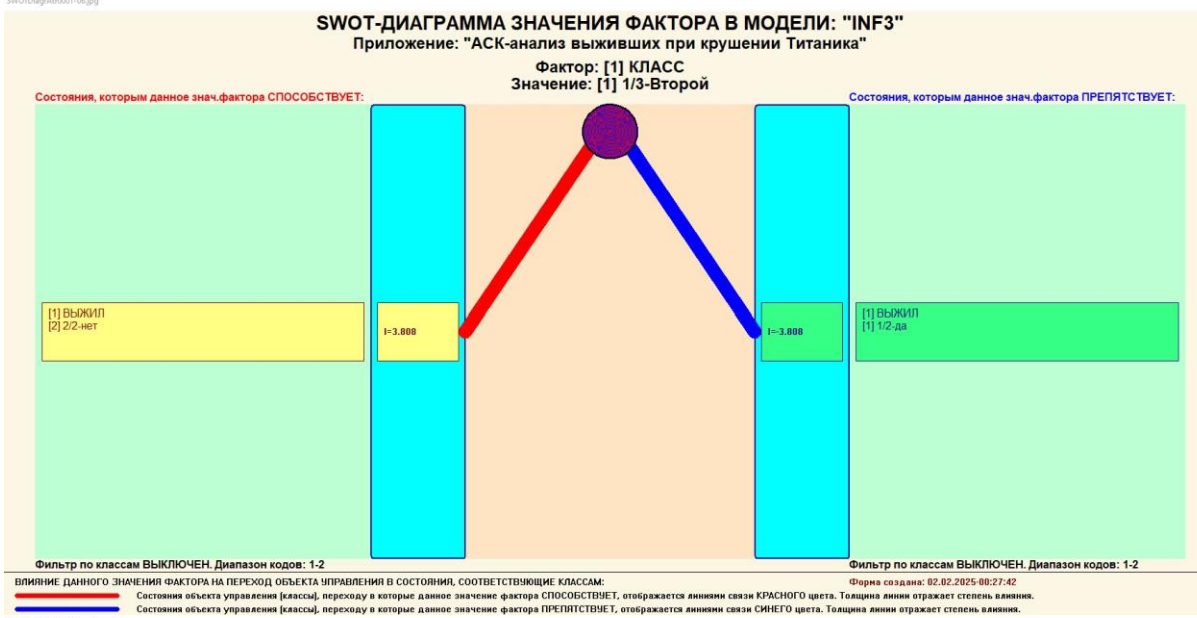
ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
1	ВЫЖИЛ-1/2-да	-3.808

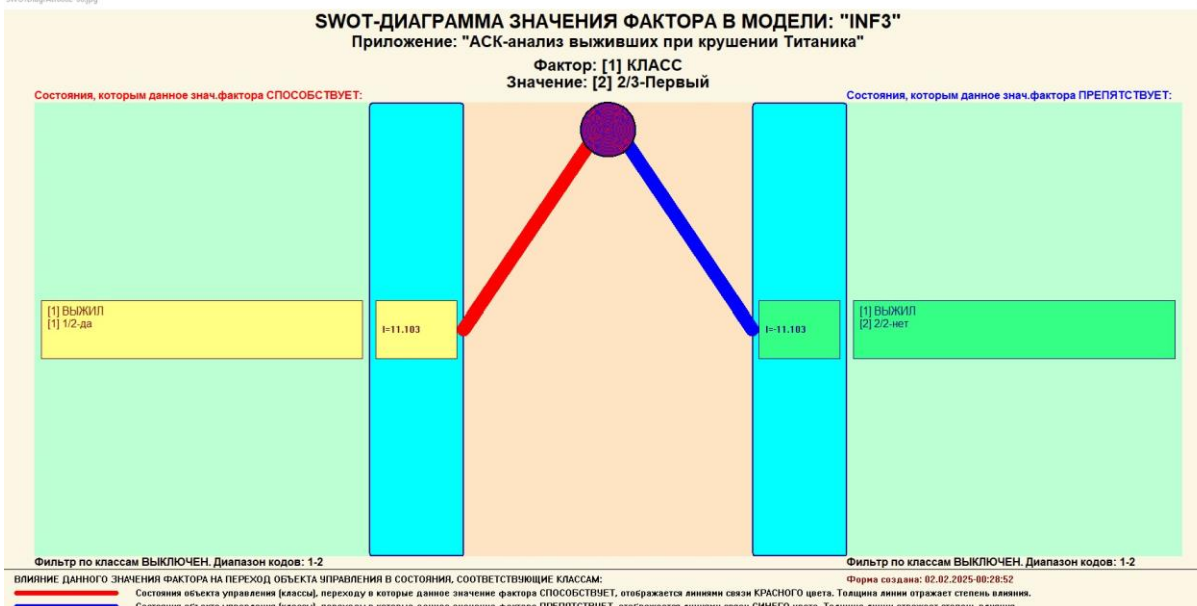
ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале

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

SWOTdiagAm001-06.jpg



SWOTdiagAm001-06.jpg



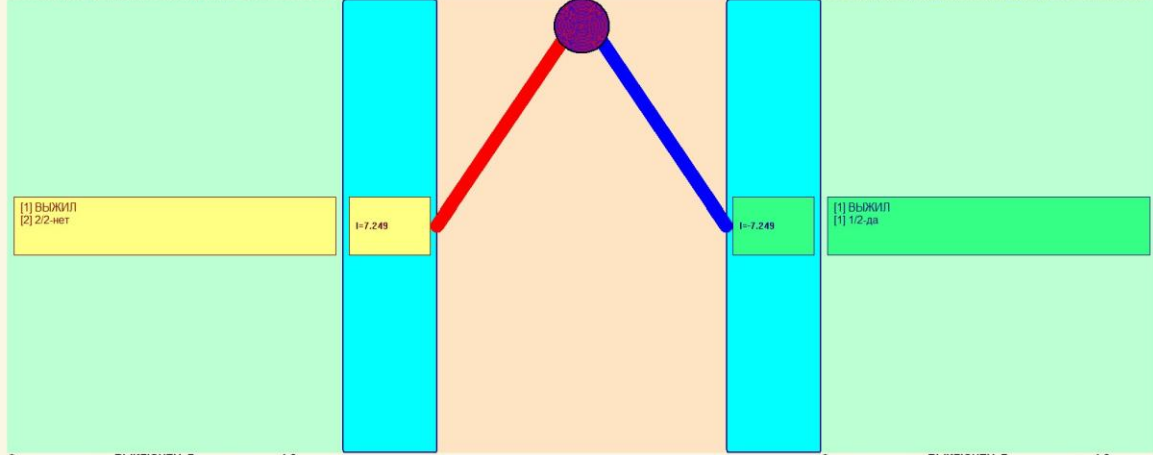
SWOTdiag4a0003-06.jpg

SWOT-ДИАГРАММА ЗНАЧЕНИЯ ФАКТОРА В МОДЕЛИ: "INF3"
 Приложение: "АСК-анализ выживших при крушении Титаника"

Фактор: [1] КЛАСС
 Значение: [3] 3/3-Третий

Состояния, которым данное знач. фактора СПОСОБСТВУЕТ:

Состояния, которым данное знач. фактора ПРЕПЯТСТВУЕТ:



Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

ВЛИЯНИЕ ДАННОГО ЗНАЧЕНИЯ ФАКТОРА НА ПЕРЕХОД ОБЪЕКТА УПРАВЛЕНИЯ В СОСТОЯНИЯ, СООТВЕТСТВУЮЩИЕ КЛАССАМ:

Форма создана: 02.02.2025 00:29:21

- Состояния объекта управления (классы), переходу в которые данное значение фактора СПОСОБСТВУЕТ, отображается линией связи КРАСНОГО цвета. Толщина линии отражает степень влияния.
- Состояния объекта управления (классы), переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ, отображается линией связи СИНЕГО цвета. Толщина линии отражает степень влияния.

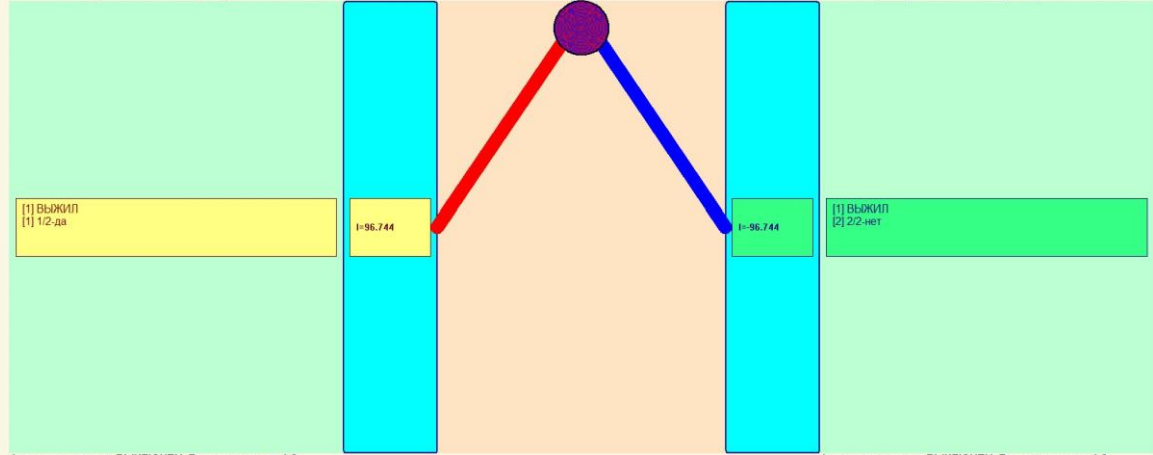
SWOTdiag4a0004-06.jpg

SWOT-ДИАГРАММА ЗНАЧЕНИЯ ФАКТОРА В МОДЕЛИ: "INF3"
 Приложение: "АСК-анализ выживших при крушении Титаника"

Фактор: [2] ПОЛ
 Значение: [4] 1/2-жен

Состояния, которым данное знач. фактора СПОСОБСТВУЕТ:

Состояния, которым данное знач. фактора ПРЕПЯТСТВУЕТ:



Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

ВЛИЯНИЕ ДАННОГО ЗНАЧЕНИЯ ФАКТОРА НА ПЕРЕХОД ОБЪЕКТА УПРАВЛЕНИЯ В СОСТОЯНИЯ, СООТВЕТСТВУЮЩИЕ КЛАССАМ:

Форма создана: 02.02.2025 00:29:41

- Состояния объекта управления (классы), переходу в которые данное значение фактора СПОСОБСТВУЕТ, отображается линией связи КРАСНОГО цвета. Толщина линии отражает степень влияния.
- Состояния объекта управления (классы), переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ, отображается линией связи СИНЕГО цвета. Толщина линии отражает степень влияния.

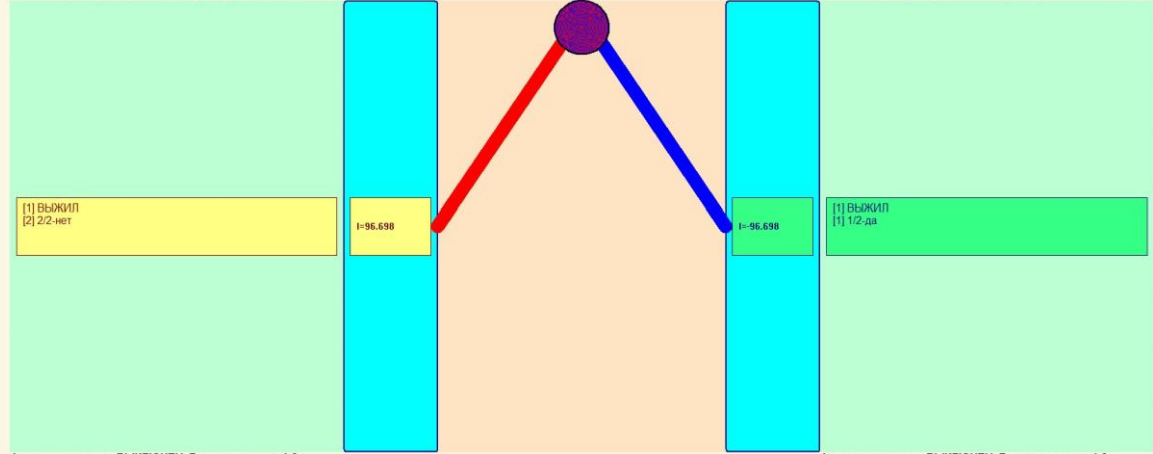
SWOTdiag4a0005-06.jpg

SWOT-ДИАГРАММА ЗНАЧЕНИЯ ФАКТОРА В МОДЕЛИ: "INF3"
 Приложение: "АСК-анализ выживших при крушении Титаника"

Фактор: [2] ПОЛ
 Значение: [5] 2/2-муж

Состояния, которым данное знач. фактора СПОСОБСТВУЕТ:

Состояния, которым данное знач. фактора ПРЕПЯТСТВУЕТ:



Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

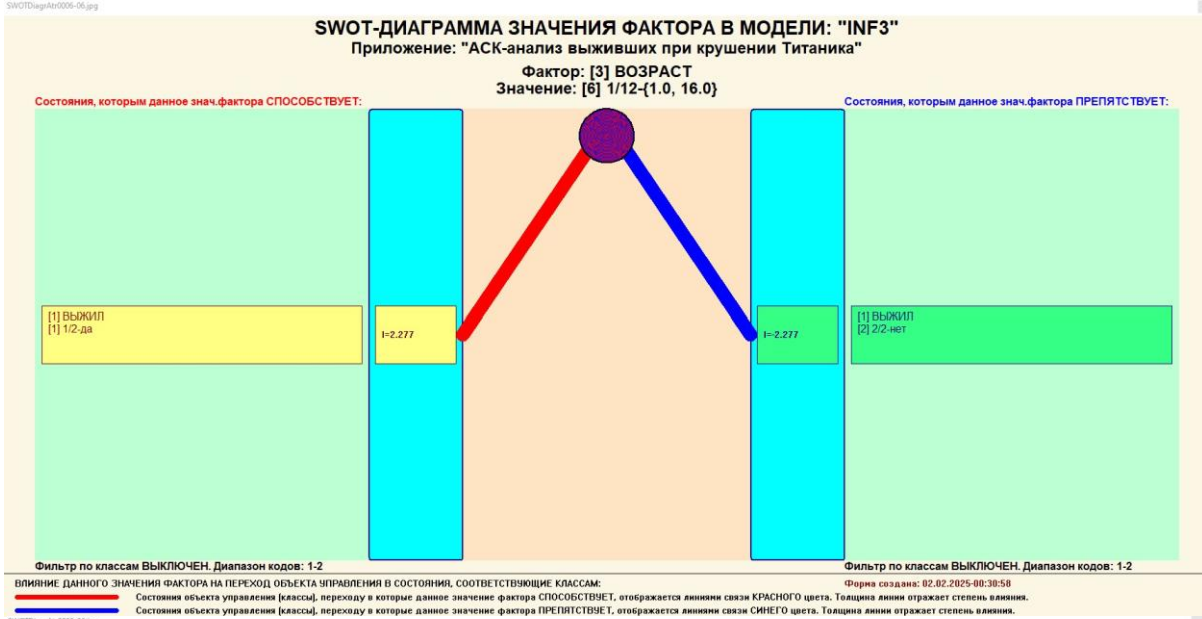
Фильтр по классам ВКЛЮЧЕН. Диапазон кодов: 1-2

ВЛИЯНИЕ ДАННОГО ЗНАЧЕНИЯ ФАКТОРА НА ПЕРЕХОД ОБЪЕКТА УПРАВЛЕНИЯ В СОСТОЯНИЯ, СООТВЕТСТВУЮЩИЕ КЛАССАМ:

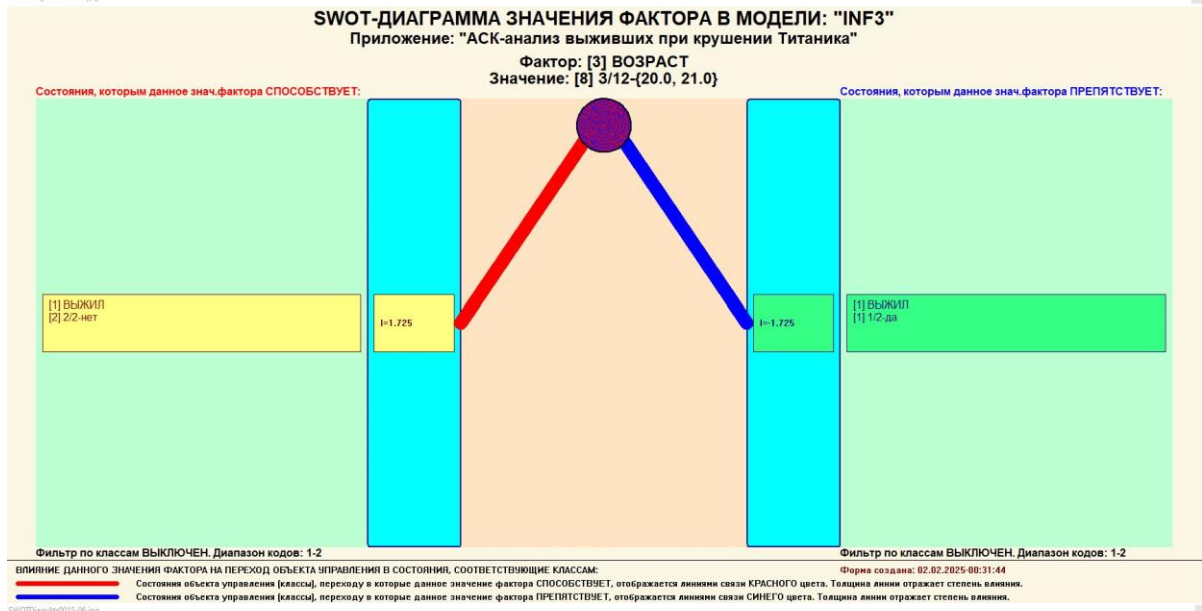
Форма создана: 02.02.2025 00:30:15

- Состояния объекта управления (классы), переходу в которые данное значение фактора СПОСОБСТВУЕТ, отображается линией связи КРАСНОГО цвета. Толщина линии отражает степень влияния.
- Состояния объекта управления (классы), переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ, отображается линией связи СИНЕГО цвета. Толщина линии отражает степень влияния.

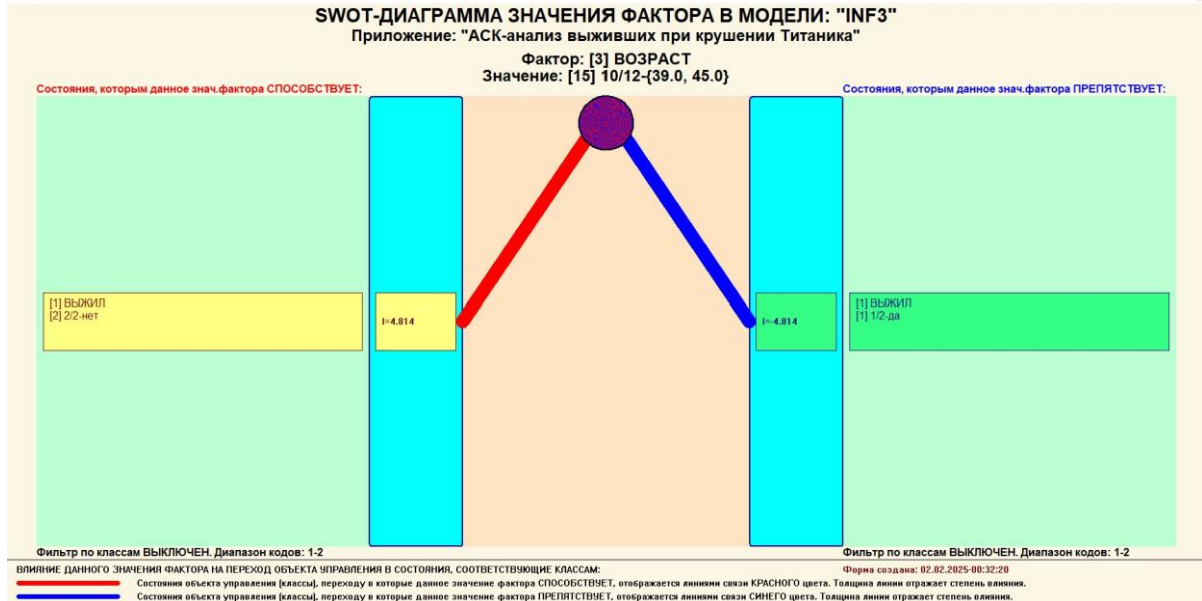
SWOTdiagAn006-06.jpg



SWOTdiagAn008-06.jpg



SWOTdiagAn015-06.jpg



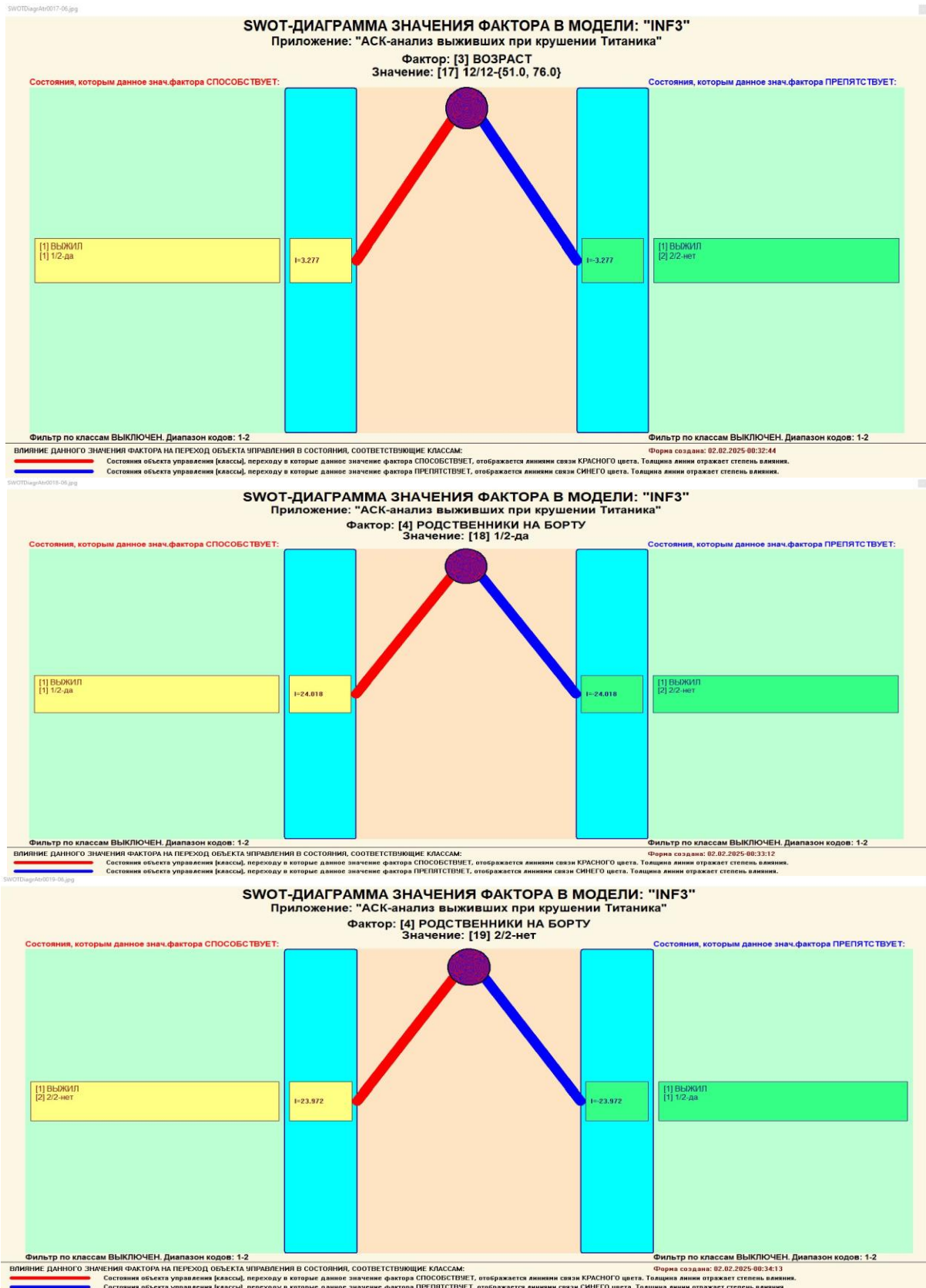


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 22 инвертированные SWOT-диаграммы исчерпывающим образом отражают силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в состояния, соответствующие различным классам.

Из приведенных выходных форм SWOT-анализа четко видно, что выживаемость пассажиров «Титаника» зависела от ряда ключевых факторов. Способствующие выживанию стали такие факторы, как принадлежность к женскому полу, наличие родственников на борту и путешествие в первом классе. В то же время, пассажиры мужского пола, путешествовавшие в одиночку в третьем классе, были подвержены наибольшим рискам, которые препятствовали их выживаемости. Во многом это и есть реальное решение проблемы, поставленной в работе.

3.8.2. Кластерно-конструктивный анализ классов

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

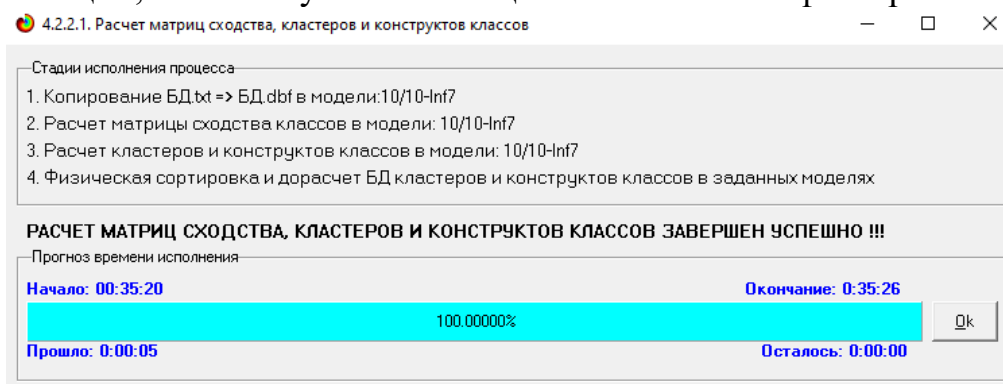
- круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2) (рисунок 24);

- агломеративные дендрограммы, полученные в результате **когнитивной (истинной) кластеризации классов** (предложена автором в 2011 году в работе [16]) (режим 4.2.2.4) (рисунок 25).

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

3.8.2.2. Конкретное решение задачи в данной работе

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



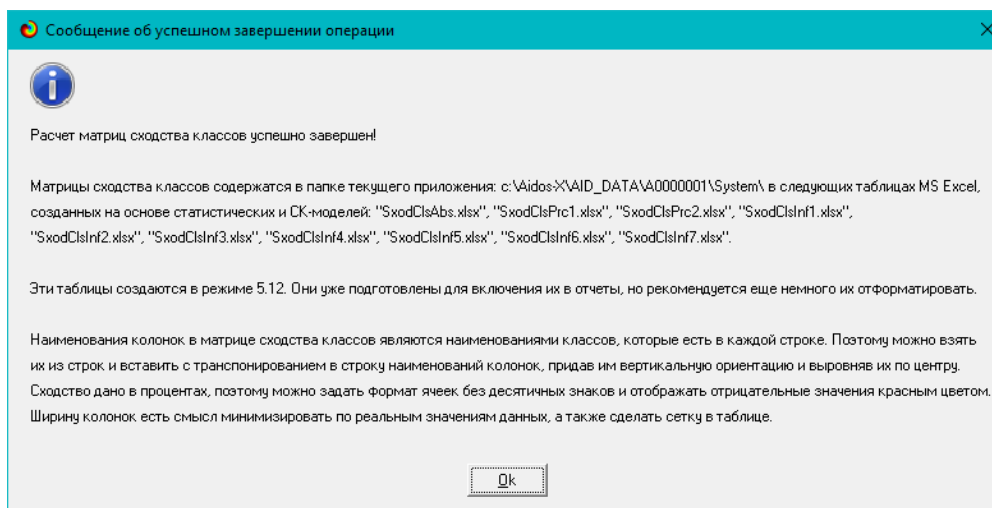


Рисунок 23. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

Таблица 14 – Матрица сходства классов в СК-модели INF3 (полностью)

	A	B	C	D
1	KOD_CLS	NAME_CLS	ВЫЖИЛ-1/2-да	ВЫЖИЛ-2/2-нет
2	1	ВЫЖИЛ-1/2-да	100,000000	-100,000000
3	2	ВЫЖИЛ-2/2-нет	-100,000000	100,000000



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

ДЕНДРОГРАММА АГЛОМЕРАТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ
 (С) Персональная интеллектуальная on-line среда 'Эйдос'
 Приложение: АСК-анализ выживших при крушении Титаника
 Модель: 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами
 Метод кластеризации: complete
 Дата и время создания дендрограммы: 2025-02-02 17:33:10.977764

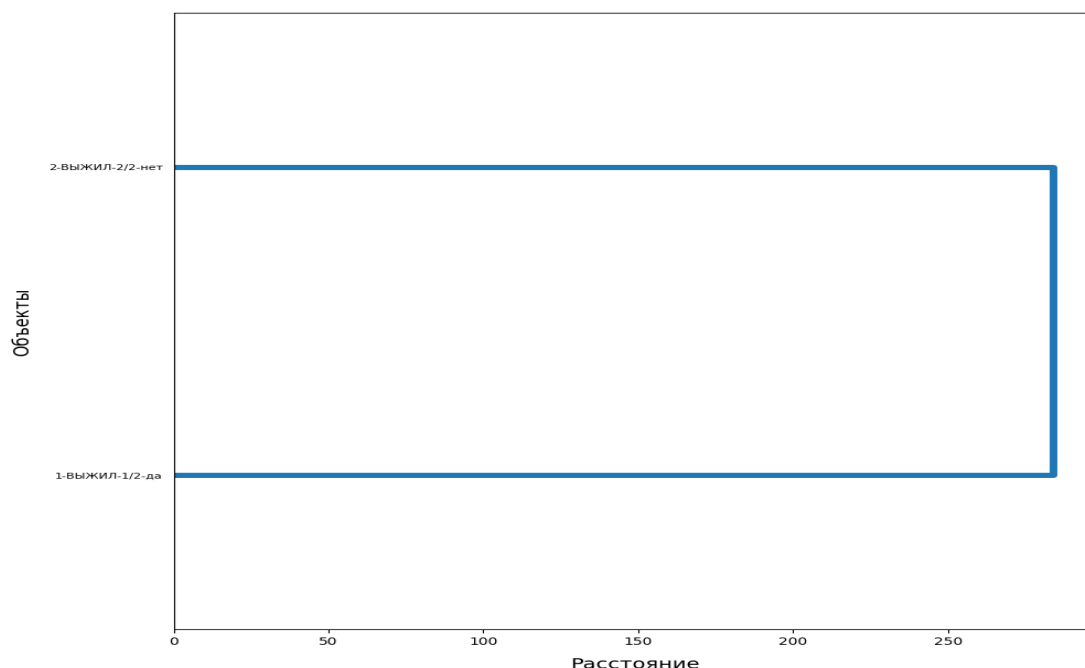


Рисунок 25. Дендрограмма когнитивной кластеризации классов (режим 4.2.2.4)

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 26) рассчитывается матрица сходства признаков (таблица 15) по их смыслу и на основе этой матрицы рассчитывается и выводится четыре основных формы:

- круговая 2d-когнитивная диаграмма признаков (режим 4.3.2.2) рисунок 27);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации признаков* (предложена автором в 2011 году в работе [16]) (режим 4.3.2.3) рисунок 28);
- график изменения межкластерных расстояний (режим 4.3.2.3) рисунок 29).

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

3.8.3.2. Конкретное решение задачи в данной работе

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

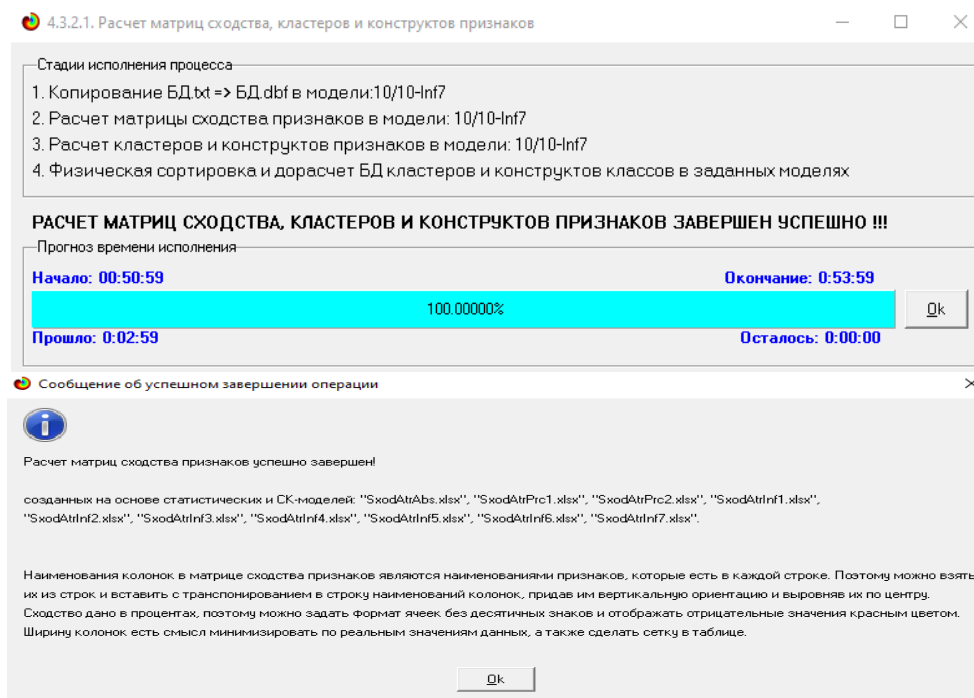


Рисунок 26. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

Таблица 15 – Матрица сходства значений факторов в СК-модели INF3 (фрагмент)

A	B	C	D	E	F	G	H	I
KOD_ATR	KOD_OPSC	NAME_ATR	КЛАСС-1/3-Второй	КЛАСС-2/3-Первый	КЛАСС-3/3-Третий	ПОЛ-1/2-жен	ПОЛ-2/2-муж	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}
1	1	КЛАСС-1/3-Второй	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
2	1	КЛАСС-2/3-Первый	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
3	1	КЛАСС-3/3-Третий	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
4	2	ПОЛ-1/2-жен	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
5	2	ПОЛ-2/2-муж	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
6	3	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
7	3	ВОЗРАСТ-2/12-{16.0000000, 20.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
8	3	ВОЗРАСТ-3/12-{20.0000000, 21.0000000}	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
9	3	ВОЗРАСТ-4/12-{21.0000000, 23.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
10	3	ВОЗРАСТ-5/12-{23.0000000, 25.0000000}	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
11	3	ВОЗРАСТ-6/12-{25.0000000, 27.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
12	3	ВОЗРАСТ-7/12-{27.0000000, 30.0000000}	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
13	3	ВОЗРАСТ-8/12-{30.0000000, 33.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
14	3	ВОЗРАСТ-9/12-{33.0000000, 39.0000000}	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
15	3	ВОЗРАСТ-10/12-{39.0000000, 45.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
16	3	ВОЗРАСТ-11/12-{45.0000000, 51.0000000}	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000
17	3	ВОЗРАСТ-12/12-{51.0000000, 76.0000000}	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
18	4	РОДСТВЕННИКИ НА БОРТУ-1/2-да	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000
19	4	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	100,0000000	-100,0000000	100,0000000	-100,0000000	100,0000000	-100,0000000

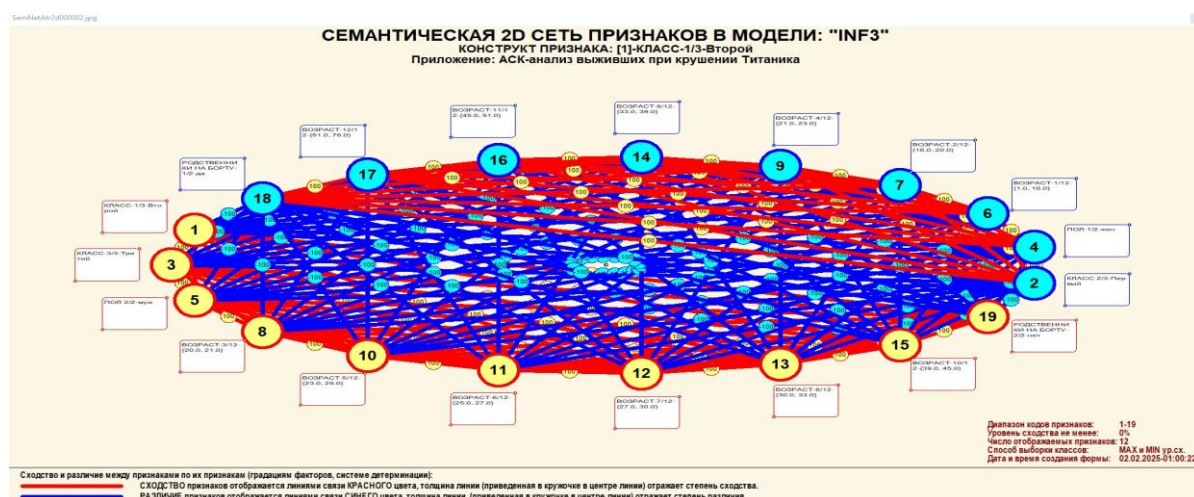


Рисунок 27. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF3 (режим 4.3.2.2)

ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF3"
 Приложение: "АСК-анализ выживших при крушении Титаника"

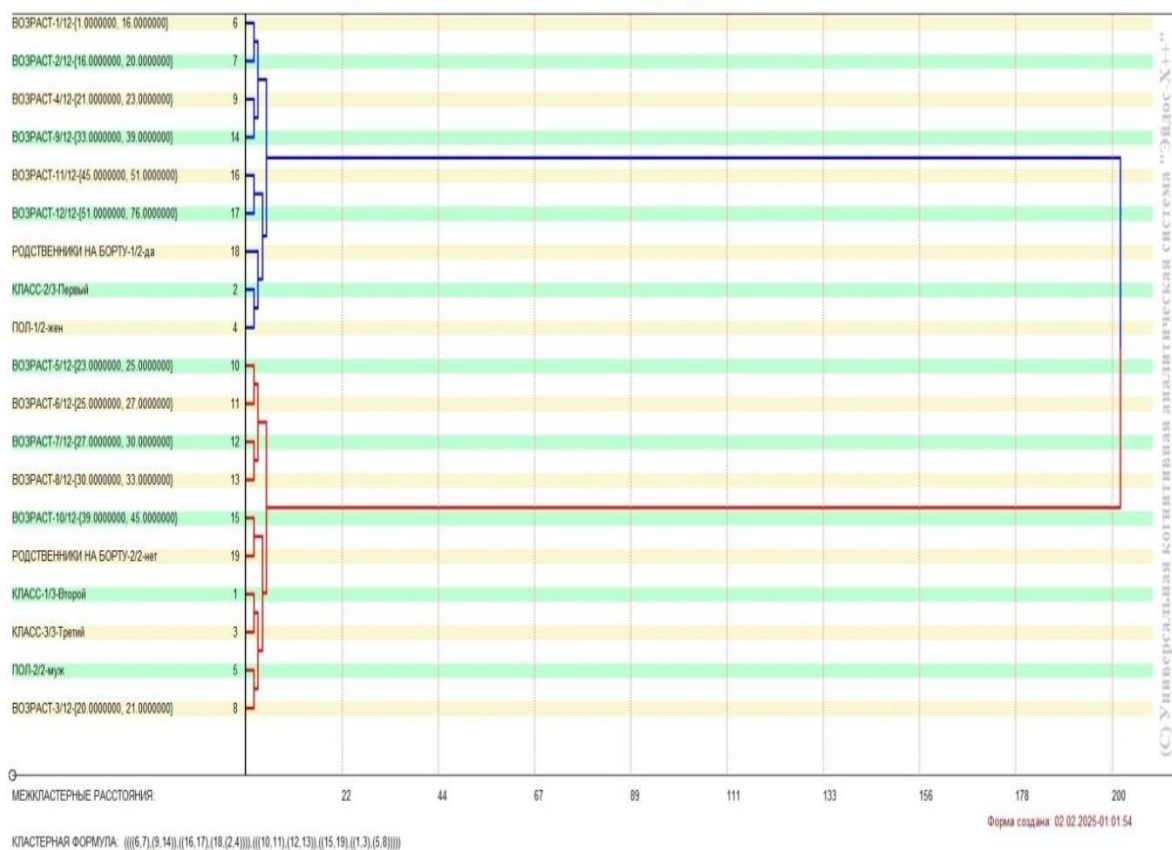


Рисунок 28. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)
ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF3"
 Приложение: "АСК-анализ выживших при крушении Титаника"

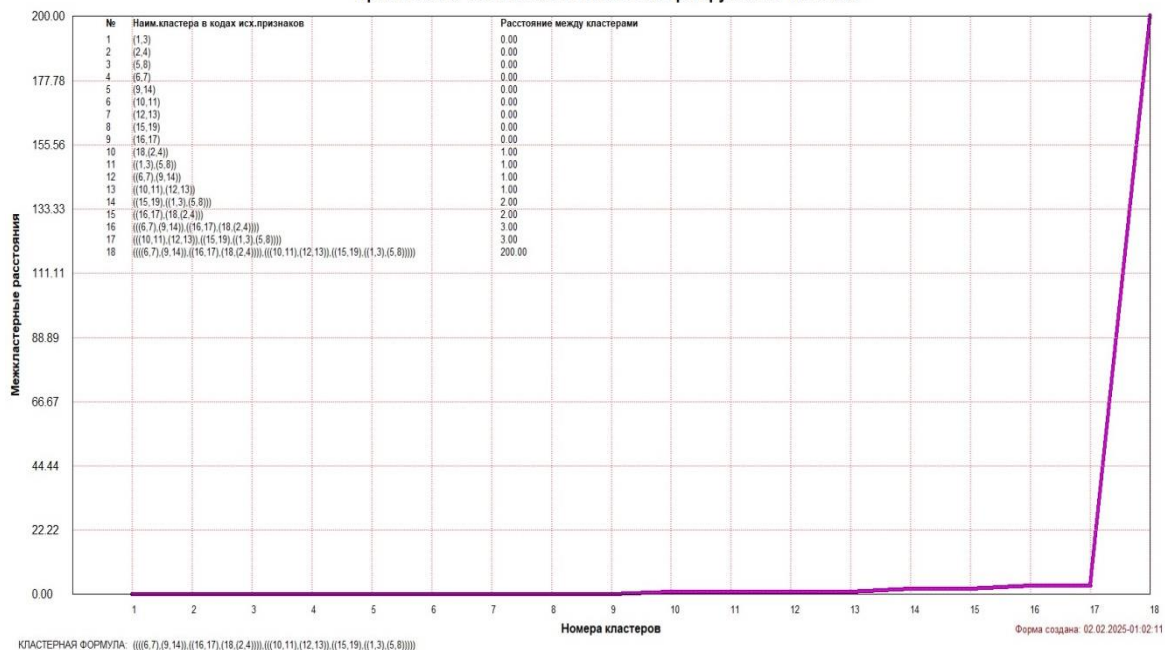


Рисунок 29. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

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

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

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

В системе «Эйдос» нелокальные нейроны визуализируются (режим 4.4.10 системы «Эйдос») в виде специальных графических форм, на которых сила и направление влияния рецепторов нейрона на степень его активации/торможения, которые отображаются соответственно в форме цвета и толщины дендрита (рисунок 30). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.4.2. Конкретное решение задачи в данной работе

4.4.10. Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	ВЫЖИЛ-1/2-да
2	ВЫЖИЛ-2/2-нет

Подготовка визуализации нейрона: 1 "ВЫЖИЛ-1/2-да" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	ПОЛ-1/2-жен	96.744
18	РОДСТВЕННИКИ НА БОРТУ-1/2-да	24.018
2	КЛАСС-2/3-Первый	11.103
7	ВОЗРАСТ-2/12-(16.00000000, 20.00000000)	3.641
17	ВОЗРАСТ-12/12-(51.00000000, 76.00000000)	3.277
9	ВОЗРАСТ-4/12-(21.00000000, 23.00000000)	2.549
6	ВОЗРАСТ-1/12-(1.00000000, 16.00000000)	2.277
14	ВОЗРАСТ-9/12-(33.00000000, 39.00000000)	1.549
16	ВОЗРАСТ-11/12-(45.00000000, 51.00000000)	1.185

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
5	ПОЛ-2/2-муж	-96.698
19	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	-23.972
3	КЛАСС-3/3-Третий	-7.249
15	ВОЗРАСТ-10/12-(39.00000000, 45.00000000)	-4.814
10	ВОЗРАСТ-5/12-(23.00000000, 25.00000000)	-4.723
1	КЛАСС-1/3-Второй	-3.808
8	ВОЗРАСТ-3/12-(20.00000000, 21.00000000)	-1.725
12	ВОЗРАСТ-7/12-(27.00000000, 30.00000000)	-1.541
13	ВОЗРАСТ-8/12-(30.00000000, 33.00000000)	-0.906

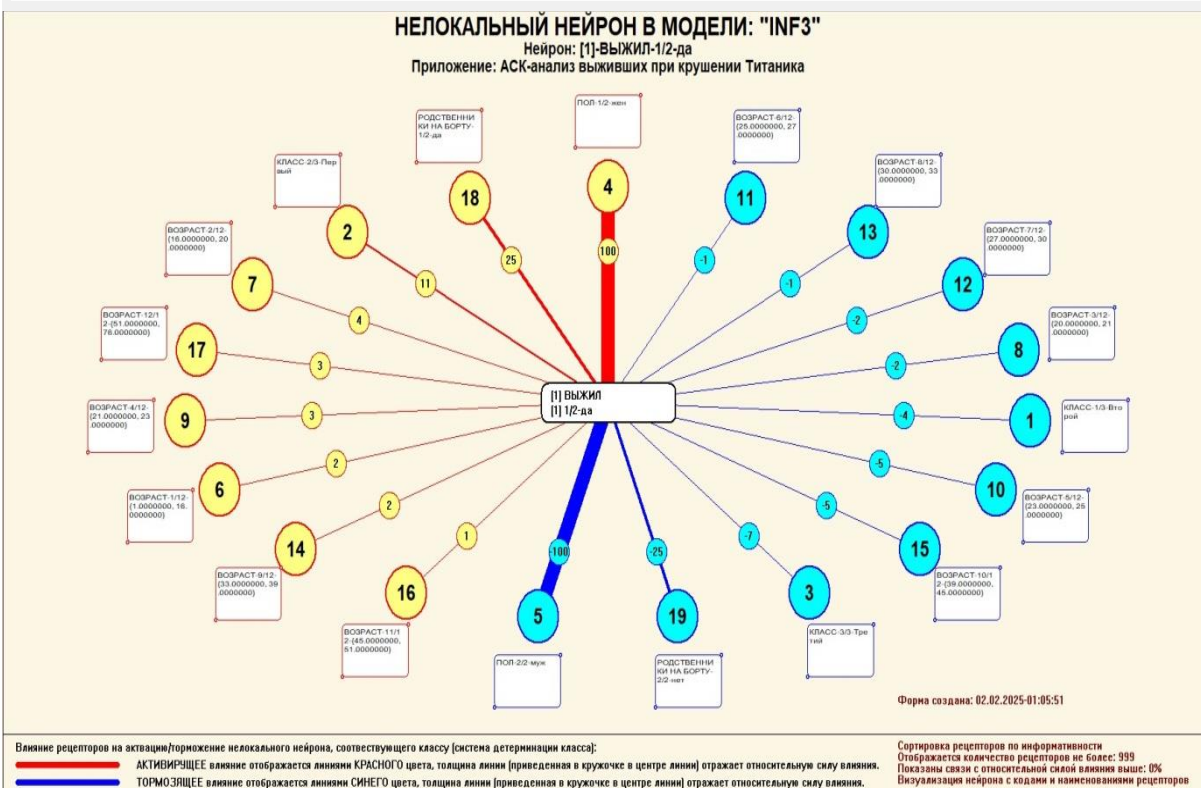
ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

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

Сортировать рецепторы:
 по информативности
 по модулю информативности

Отображать рецепторы:
 с наименованиями
 только с кодами



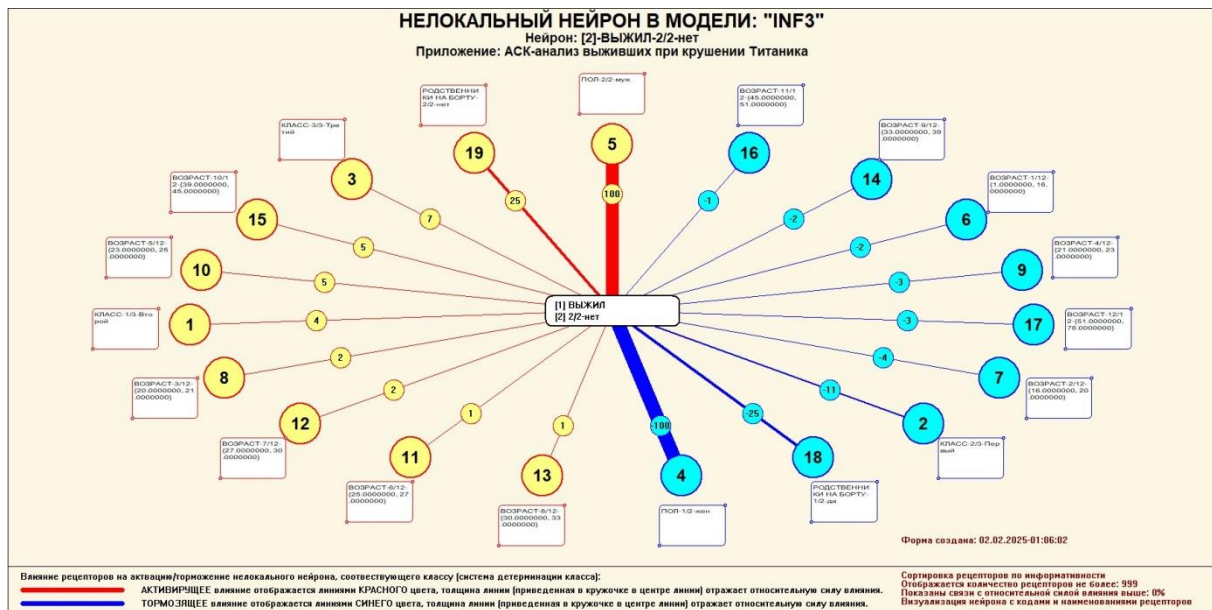


Рисунок 30. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

Нейроны на изображении слоя нейронной сети расположены слева направо в порядке убывания модуля суммарной силы их детерминации рецепторами, т.е. слева находятся результаты, наиболее жестко обусловленные действующими на них значениями факторов, а справа – менее жестко обусловленные (рисунок 31). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.5.2. Конкретное решение задачи в данной работе

На рисунке 31 мы видим, что выживаемость пассажиров «Титаника» зависела от ряда ключевых факторов. Способствующие выживанию стали такие факторы, как принадлежность к женскому полу, наличие родственников на борту и путешествие в первом классе. В то же время, пассажиры мужского пола, путешествовавшие в одиночку в третьем классе, были подвержены наибольшему риску, которые препятствовали их выживаемости.

Выбор нелокальных нейронов (классов) для визуализации в нейросети

№	Код	Наименование нелокального нейрона (класса)
1	1	ВЫЖИЛ-1/2-да
2	2	ВЫЖИЛ-2/2-нет

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

Подготовка визуализации нейрона: 1 "ВЫЖИЛ-1/2-да" в модели: 6 "INF3"
АКТИВИРУЮЩИЕ рецепторы и сила их влияния **ТОРМОЗЯЩИЕ рецепторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
4	ПОЛ-1/2-жен	96.744
18	РОДСТВЕННИКИ НА БОРТУ-1/2-да	24.018
2	КЛАСС-2/3-Первый	11.103
7	ВОЗРАСТ-2/12-{16.00000000, 20.00000000}	3.641
17	ВОЗРАСТ-12/12-{51.00000000, 76.00000000}	3.277
9	ВОЗРАСТ-4/12-{21.00000000, 23.00000000}	2.549
6	ВОЗРАСТ-1/12-{1.00000000, 16.00000000}	2.277
14	ВОЗРАСТ-9/12-{33.00000000, 39.00000000}	1.549
16	ВОЗРАСТ-11/12-{45.00000000, 51.00000000}	1.185

Код	Наименование фактора и его интервального значения	Сила влияния
5	ПОЛ-2/2-муж	-96.698
19	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	-23.972
3	КЛАСС-3/3-Третий	-7.249
15	ВОЗРАСТ-10/12-{39.00000000, 45.00000000}	-4.814
10	ВОЗРАСТ-5/12-{23.00000000, 25.00000000}	-4.723
1	КЛАСС-1/3-Второй	-3.808
8	ВОЗРАСТ-3/12-{20.00000000, 21.00000000}	-1.725
12	ВОЗРАСТ-7/12-{27.00000000, 30.00000000}	-1.541
13	ВОЗРАСТ-8/12-{30.00000000, 33.00000000}	-0.906
11	ВОЗРАСТ-6/12-{25.00000000, 27.00000000}	-0.906

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

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

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

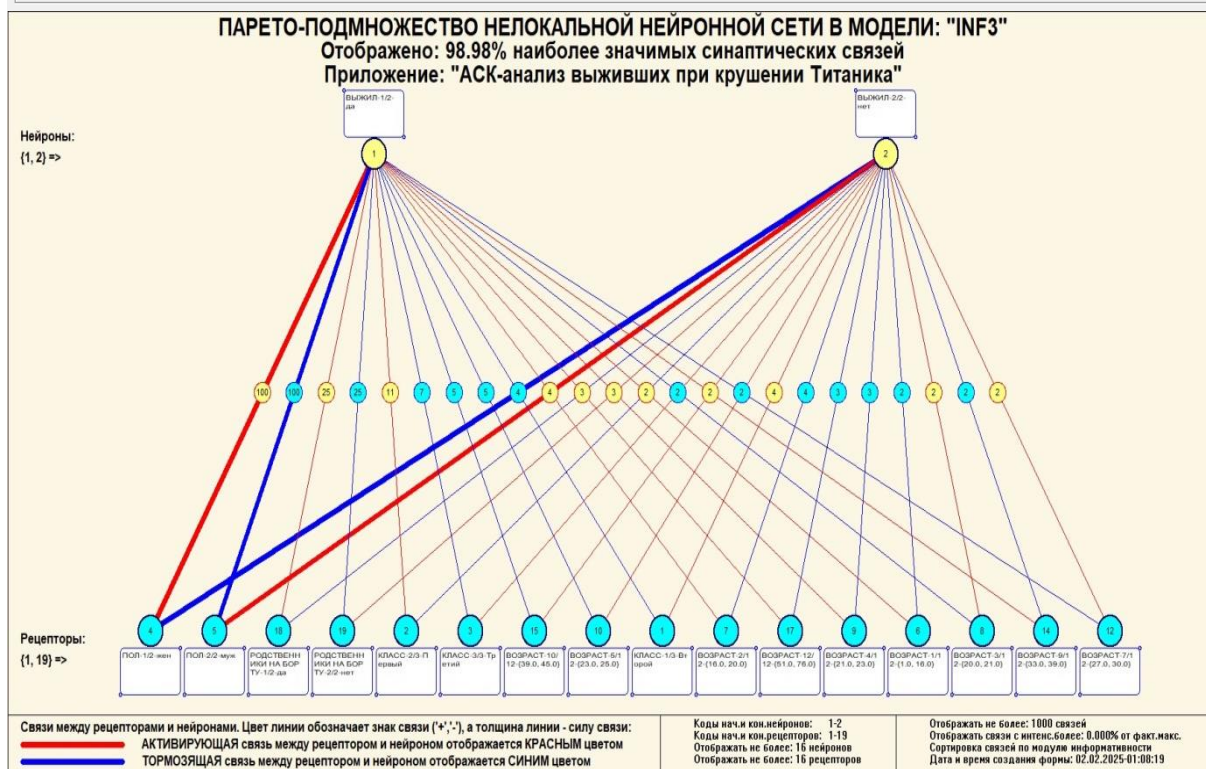


Рисунок 31. Нейронная сеть в СК-модели INF3

3.8.6. 3D-интегральные когнитивные карты

3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 24) вверху и когнитивной диаграммы значений факторов (рисунок 27) внизу и соединяющего их одного слоя нейронной сети (рисунок 31) (режим 4.4.12 системы «Эйдос») (рисунок 32).

3.8.6.2. Конкретное решение задачи в данной работе

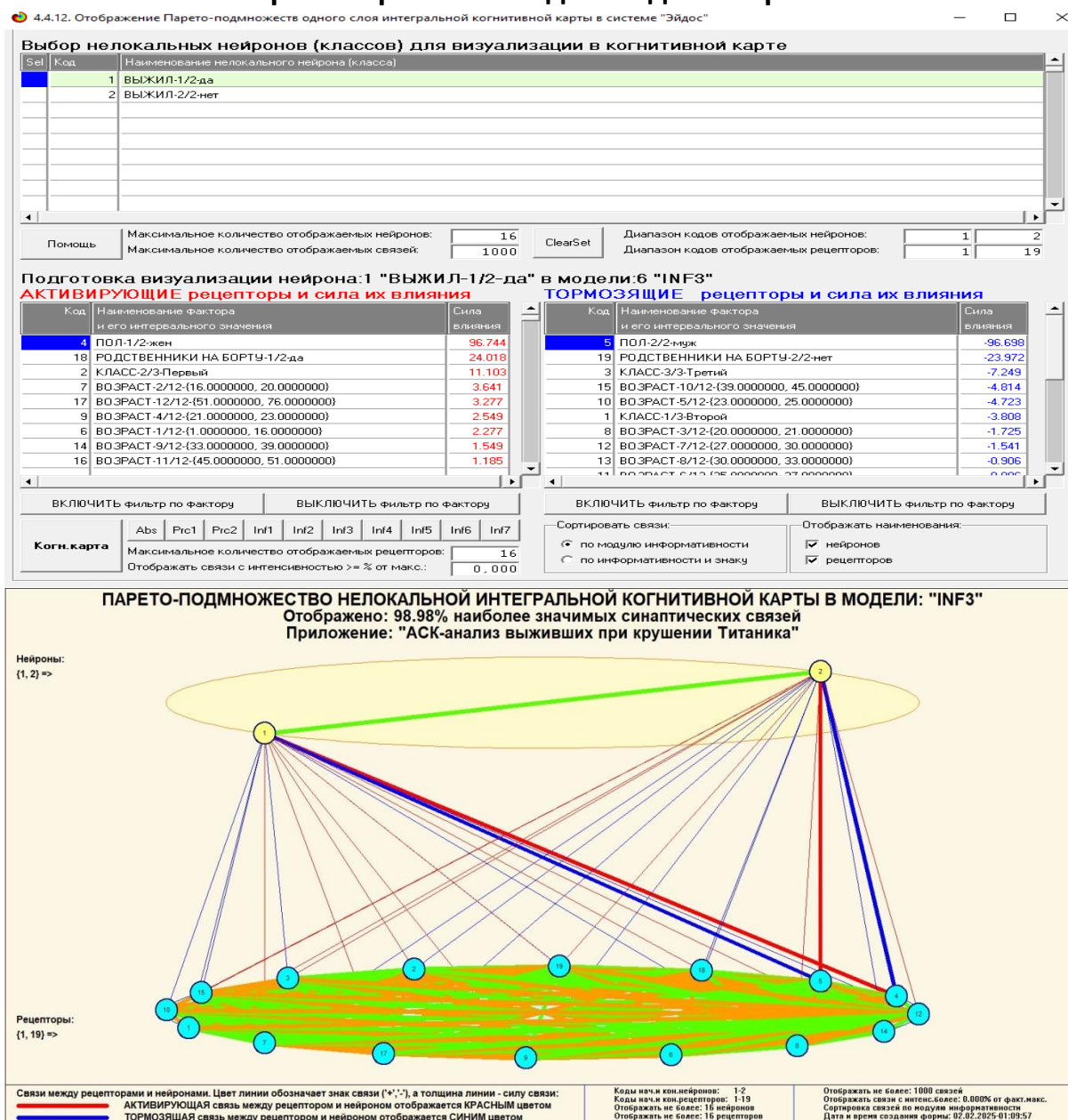


Рисунок 32. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

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

3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

Пример опосредованных правдоподобных рассуждений.

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

3.8.7.2. Конкретное решение задачи в данной работе

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 34. Всего системой в данной модели генерируется 9 форм содержательного сравнения классов. Так как каждый из 3 классов сравнивается со всеми остальными, в т.ч. с собой, то всего получается $3^2=9$ подобных диаграмм. На рисунках 33 приводятся некоторые из этих диаграмм. Пользователь при желании всегда

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

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

может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №393 и получить в нем все выходные формы, как это описано в данной статье.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

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

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	ВЫЖИЛ-1/2-да
2	ВЫЖИЛ-2/2-нет

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

Выбор способа фильтрации признаков в информационных портретах когнитивной диаграммы

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

Код	Наименование описательной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛ	1	19
1	КЛАСС	1	3
2	ПОЛ	4	5
3	ВОЗРАСТ	6	17
4	РОДСТВЕННИКИ НА БОРТУ	18	19

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

Задайте модели, в которых проводить расчеты когнитивных диаграмм:

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Класс для левого инф. портрета: [0] ВСЕ КЛАССЫ
Класс для правого инф. портрета: [0] ВСЕ КЛАССЫ
Описат. шкала для левого инф. портрета: [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛ
Описат. шкала для правого инф. портрета: [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛ
Модели, заданные для расчета: Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

КОГНИТИВНАЯ ДИАГРАММА КЛАССОВ В МОДЕЛИ: "INF3"
Приложение: "АСК-анализ выживших при крушении Титаника"
Сход./разл. классов: 100.000%

Кл.шкала: [1] ВЫЖИЛ Класс: [1] 1/2-да Сход./разл. классов: 100.000% Кл.шкала: [1] ВЫЖИЛ Класс: [1] 1/2-да

Наименования признаков:

[2] ПОЛ [4] 1/2-жен	Ив=96.744 Ip=9674.385 Ic=0.707	[2] ПОЛ [4] 1/2-жен	Ив=96.744 Ip=9674.385 Ic=0.707
[4] РОДСТВЕННИКИ НА БОРТУ [18] 1/2-да	Ив=24.018 Ip=2401.800 Ic=0.707	[4] РОДСТВЕННИКИ НА БОРТУ [18] 1/2-да	Ив=24.018 Ip=2401.800 Ic=0.707
[1] КЛАСС [2] 2/3-Первый	Ив=11.103 Ip=1110.258 Ic=0.707	[1] КЛАСС [2] 2/3-Первый	Ив=11.103 Ip=1110.258 Ic=0.707
[3] ВОЗРАСТ [15] 10/12-(39.0, 45.0)	Ив=4.814 Ip=481.404 Ic=0.707	[3] ВОЗРАСТ [15] 10/12-(39.0, 45.0)	Ив=4.814 Ip=481.404 Ic=0.707
[1] КЛАСС [3] 3/3-Третий	Ив=7.249 Ip=724.895 Ic=0.707	[1] КЛАСС [3] 3/3-Третий	Ив=7.249 Ip=724.895 Ic=0.707
[4] РОДСТВЕННИКИ НА БОРТУ [19] 2/2-нет	Ив=23.972 Ip=2397.241 Ic=0.707	[4] РОДСТВЕННИКИ НА БОРТУ [19] 2/2-нет	Ив=23.972 Ip=2397.241 Ic=0.707
[2] ПОЛ [5] 2/2-муж	Ив=96.839 Ip=9683.826 Ic=0.707	[2] ПОЛ [5] 2/2-муж	Ив=96.839 Ip=9683.826 Ic=0.707

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛ 1-19 Фильтр по оп.шкале: [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛ 1-19

Сходство и различие между классами по их признакам с учетом сходств/различия между признаками (градациями факторов, системе детерминации):
— СХОДСТВО классов отображается линиями связи КРАСНОГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень сходства.
— РАЗЛИЧИЕ классов отображается линиями связи СИНЕГО цвета, толщина линии (приведенная в кружочке в центре линии) отражает степень различия.

Форма создана: 02.02.2025-01:11:44. Показано количество связей <=999

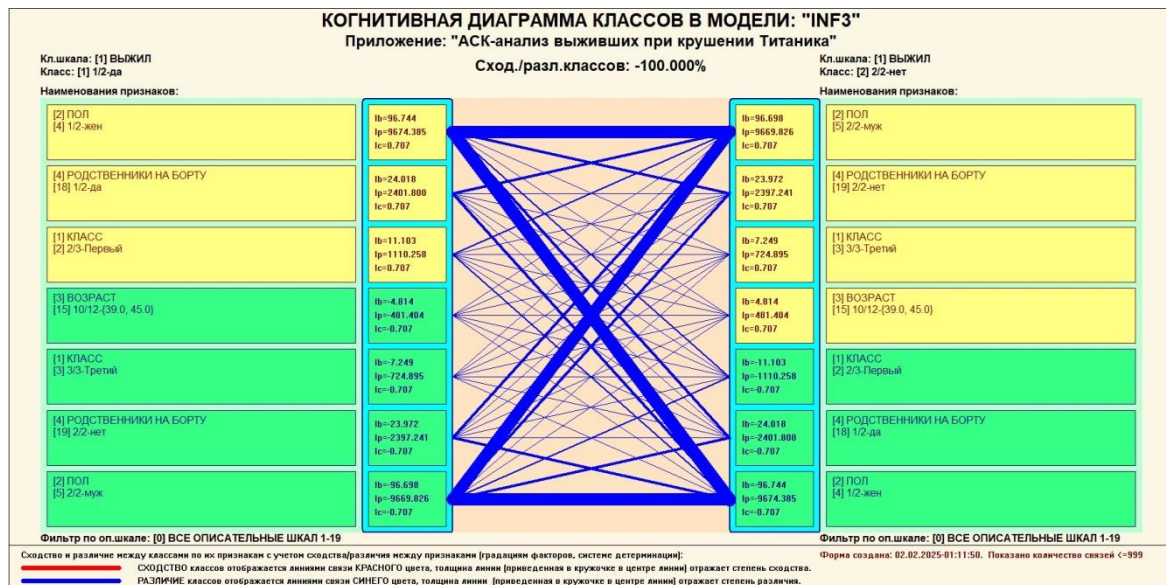


Рисунок 33. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF3

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

3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

3.8.8.2. Конкретное решение задачи в данной работе

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта моделирования в будущие состояния, соответствующие классам, приведены ниже на рисунках 34:

4.3.3. Когнитивные диаграммы признаков. Задание параметров генерации выходных форм

Выбор признаков для когнитивной диаграммы

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

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	КЛАСС-1/3-Второй
2	КЛАСС-2/3-Первый
3	КЛАСС-3/3-Третий
4	ПОЛ-1/2-жен
5	ПОЛ-2/2-муж
6	ВОЗРАСТ-1/12-{1.0000000, 16.0000000}

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

Выбор способа фильтрации классов в информационных портретах когнитивной диаграммы

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

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

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

Задайте модели, в которых проводить расчеты когнитивных диаграмм:

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

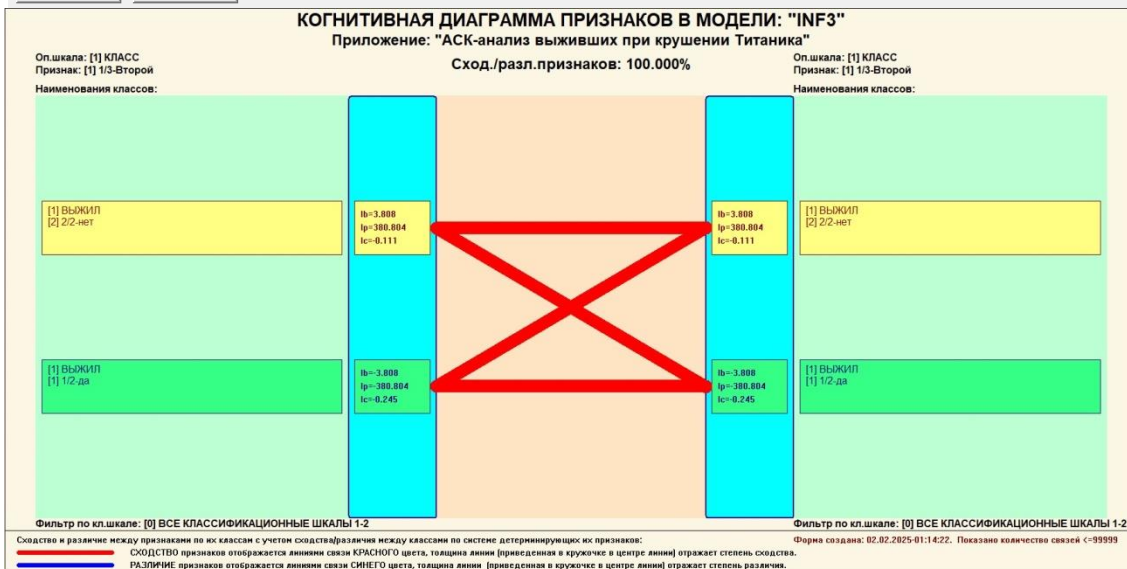
В диалоге заданы следующие параметры расчета когнитивных диаграмм:

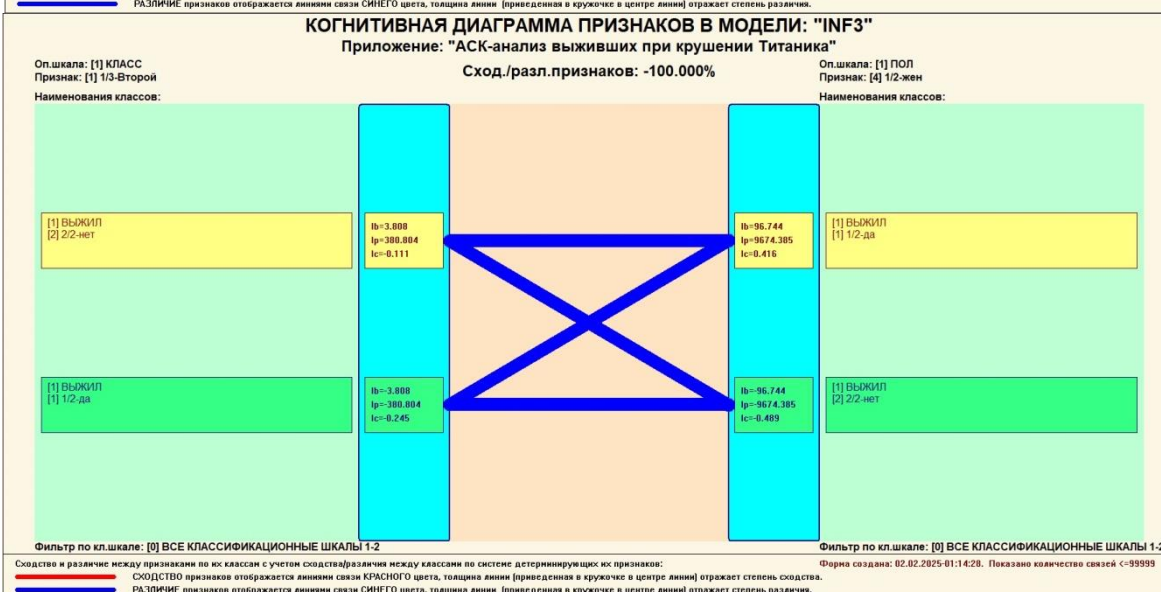
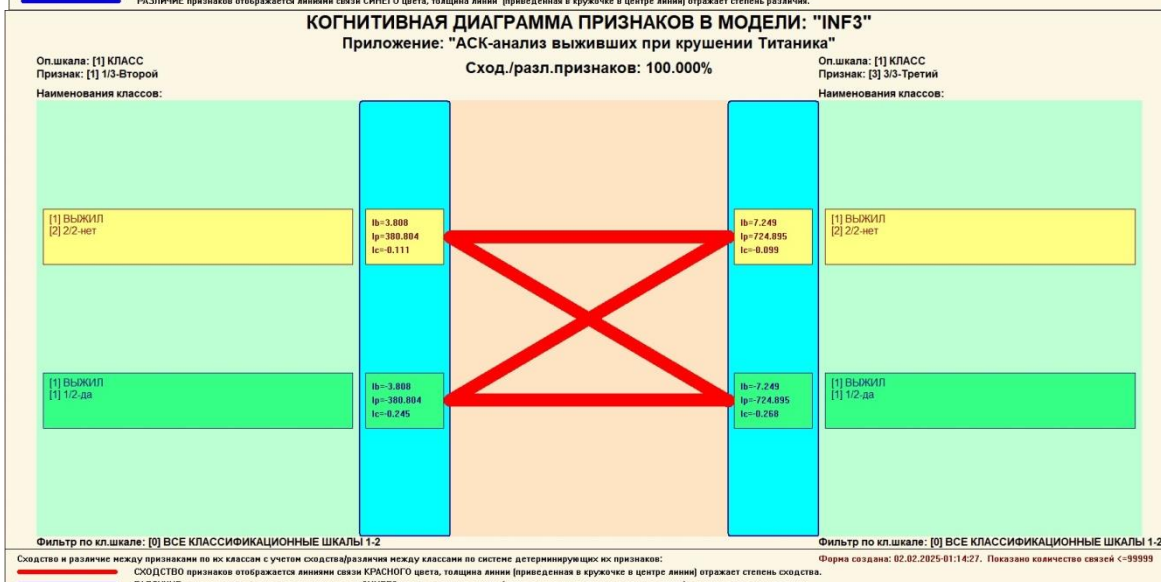
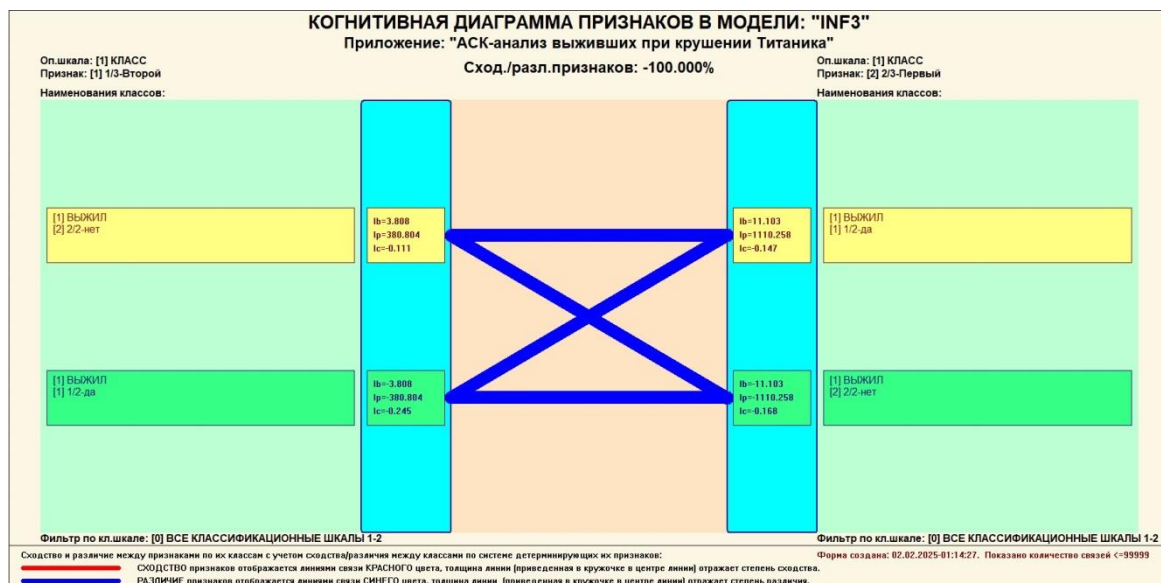
Признак для левого инф. портрета: [0] ВСЕ ПРИЗНАКИ
Признак для правого инф. портрета: [0] ВСЕ ПРИЗНАКИ
Классиф. шкала для левого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Классиф. шкала для правого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Модели, заданные для расчета: Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7

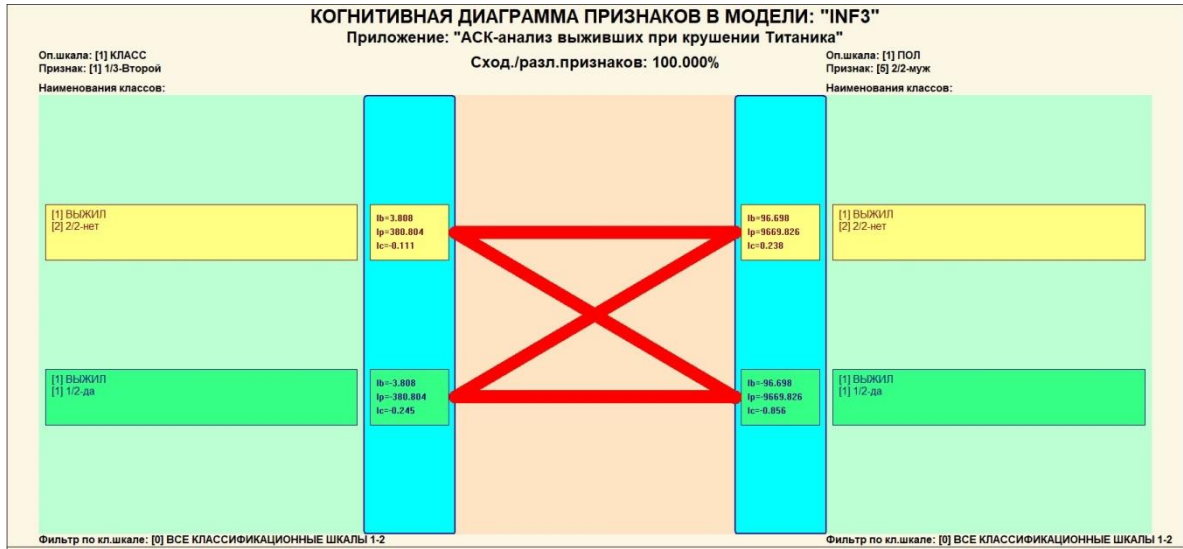
Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

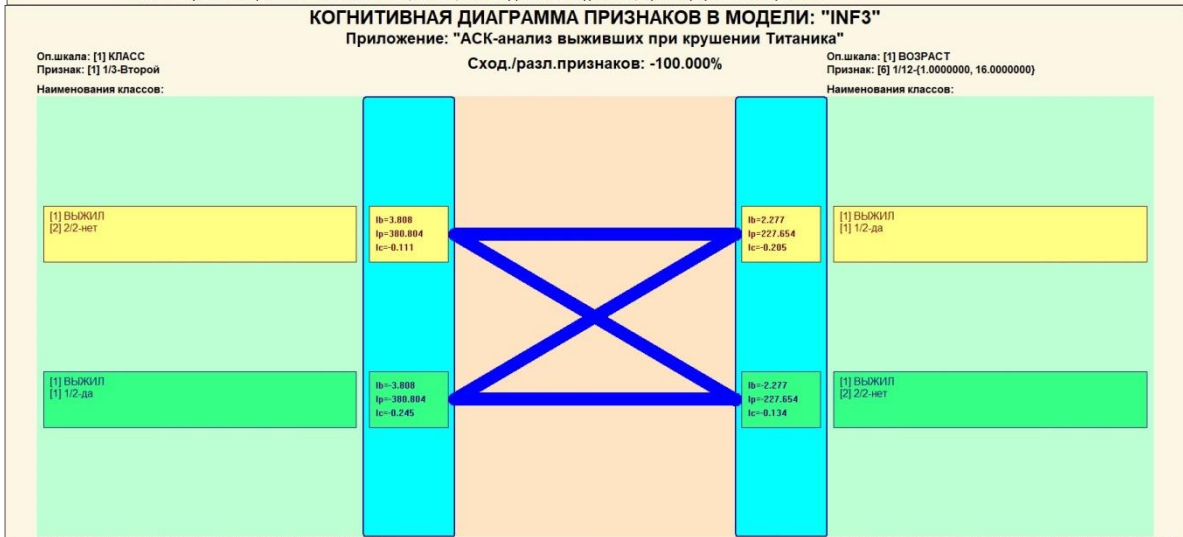
Ok Cancel







Сходство и различие между признаками по их классам с учетом сходства/различия между классами по системе детерминирующих их признаков:
 СХОДСТВО признаков отображается линиями связи КРАСНОГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень сходства.
 РАЗЛИЧИЕ признаков отображается линиями связи СИНЕГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень различия.



Сходство и различие между признаками по их классам с учетом сходства/различия между классами по системе детерминирующих их признаков:
 СХОДСТВО признаков отображается линиями связи КРАСНОГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень сходства.
 РАЗЛИЧИЕ признаков отображается линиями связи СИНЕГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень различия.



Сходство и различие между признаками по их классам с учетом сходства/различия между классами по системе детерминирующих их признаков:
 СХОДСТВО признаков отображается линиями связи КРАСНОГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень сходства.
 РАЗЛИЧИЕ признаков отображается линиями связи СИНЕГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень различия.

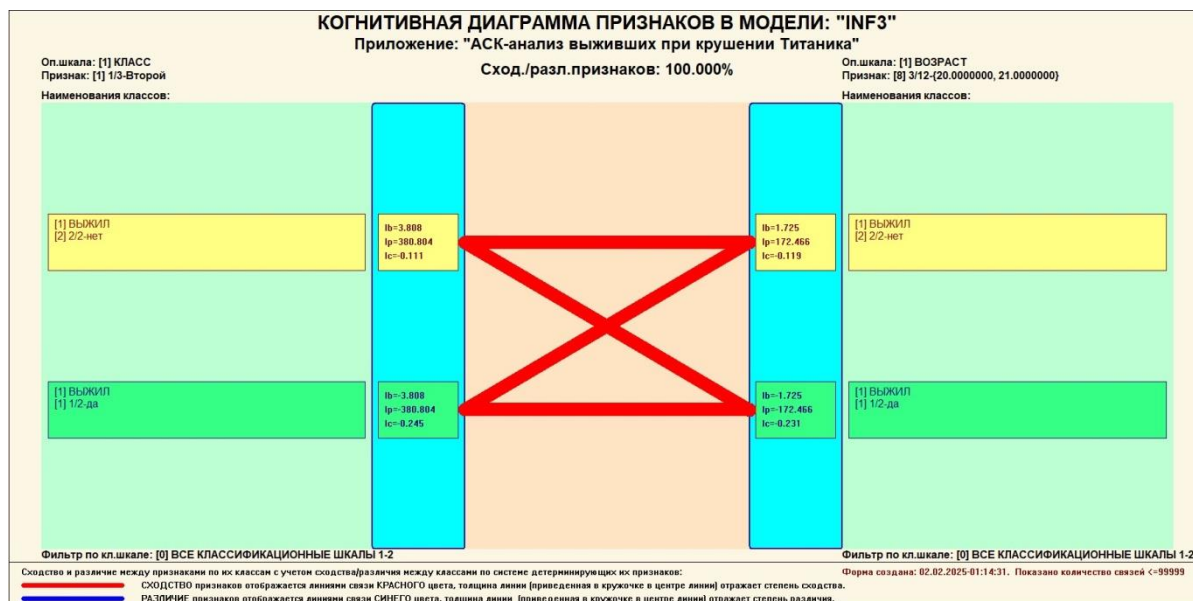


Рисунок 34. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF3

Всего системой в данной модели генерируется 361 подобных диаграмм содержательного сравнения значений факторов по их смыслу, т.е. по влиянию на объект моделирования. В данной работе все они не приводятся. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №421 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

значению аргумента соответствуют все значения функции, но соответствуют в разной степени, причем как положительной, так и отрицательной, которая отображается цветом.

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

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 35). Первая экранная форма данного режима представляет собой краткий хелп, поясняющий смысл понятия «Когнитивная функция», а также позволяющий выйти на экранную форму системы «Эйдос» с действующими гиперссылками на работы по когнитивным функциям, а также страницы сайта автора со списком этих работ и работ автора по выявлению, представлению и использованию знаний, логике и методологии научного познания.

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

3.8.9.2. Конкретное решение задачи в данной работе

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 35). Количество когнитивных функций равно количеству сочетаний описательных и классификационных шкал. В модели, рассматриваемой в данной работе, есть 4 описательных шкалы и 1 классификационная, поэтому получается 4 когнитивных функции:

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

— □ ×

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

Статистические базы:

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

Системно-когнитивные модели (Базы знаний):

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

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

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

— Задайте дополнительные параметры визуализации когнитивных функций:

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

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

Задайте количество пикселей на дюйм в изображениях когнитивных функций:

Задайте паузу в секундах между визуализациями когнитивных функций:

Задайте размер шрифта для наименований градаций шкал X и Y:

Визуализация когнитивных функций new **Визуализация когнитивных функций old**

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

4.5. Визуализация когнитивных функций

— □ ×

— Что такое когнитивная функция:

Визуализация прямых, обратных, позитивных, негативных, полностью и частично редуцированных когнитивных функций

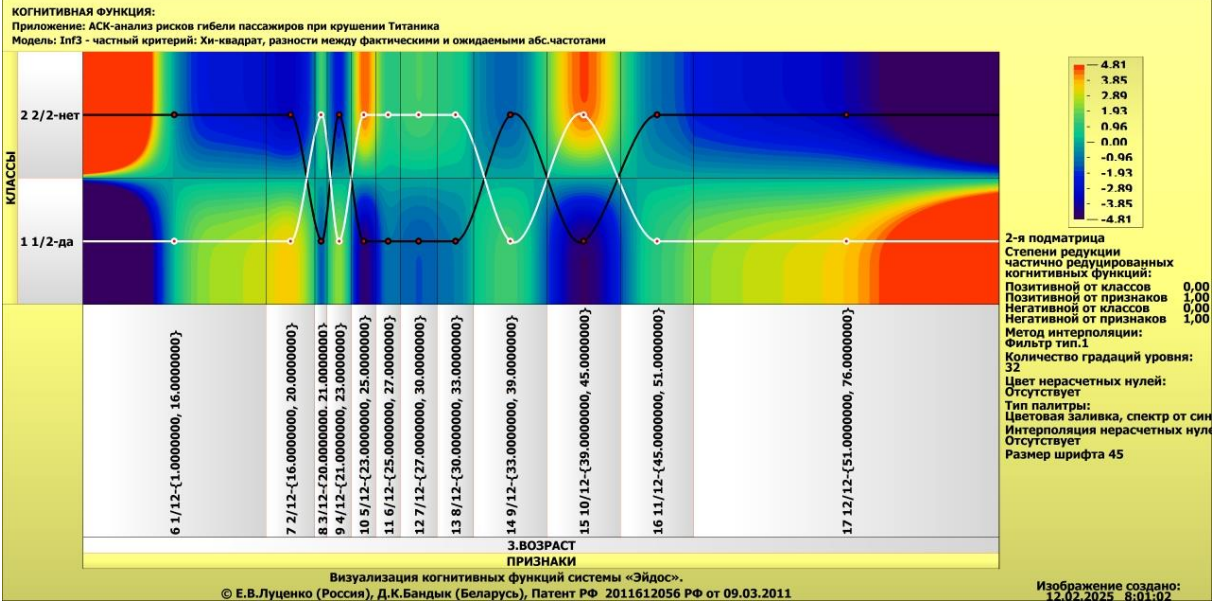
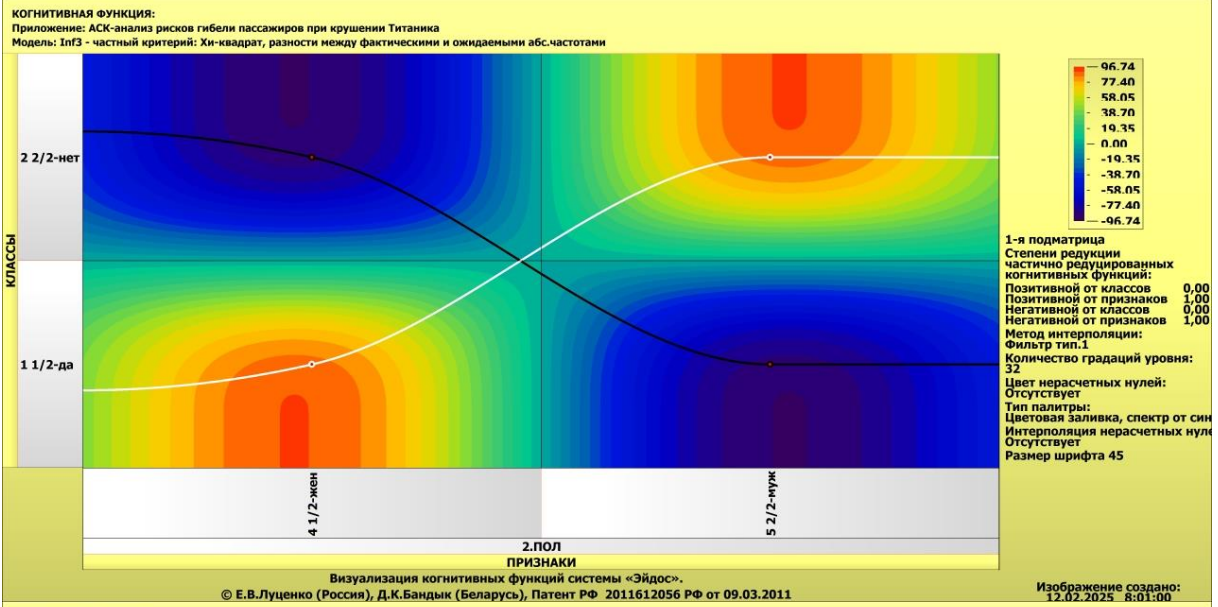
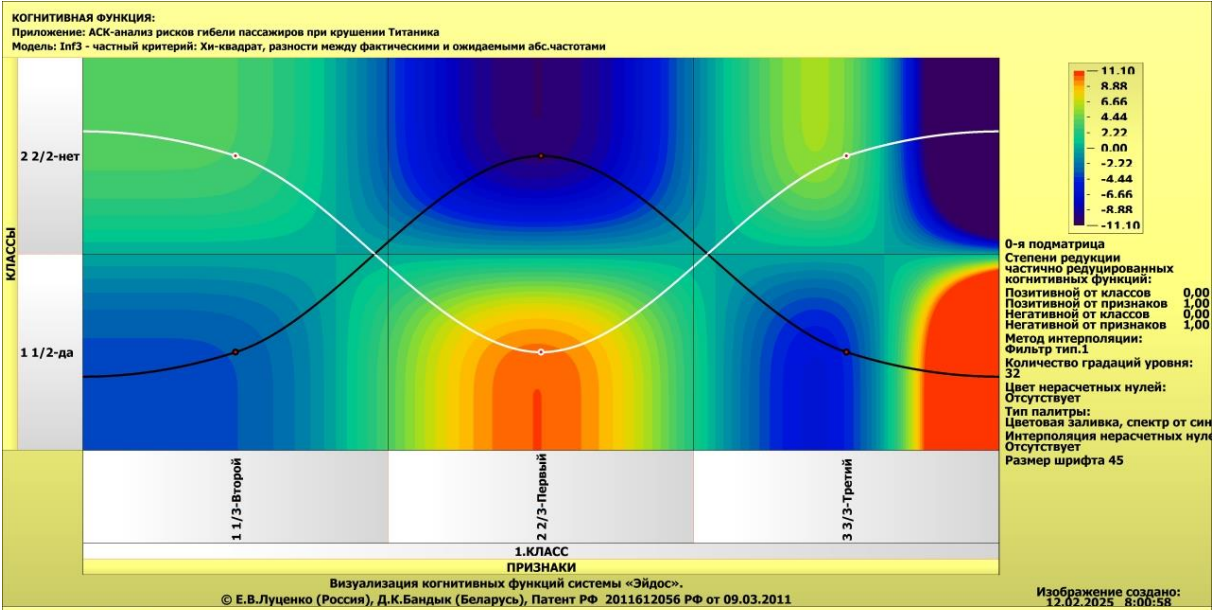
Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам. Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации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

Луценко Е.В. Метод визуализации когнитивных функций - новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета [Научный журнал КубГАУ] [Электронный ресурс]. - Краснодар: КубГАУ, 2011. - №03(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077. - 2,688 з.л. - Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>

— Задайте нужный режим:

Визуализации когнитивных функций Литератур.ссылки на работы по когнитивным функциям

Литератур.ссылки на работы по когнитивным функциям Литератур.ссылки на работы по управлению знаниями



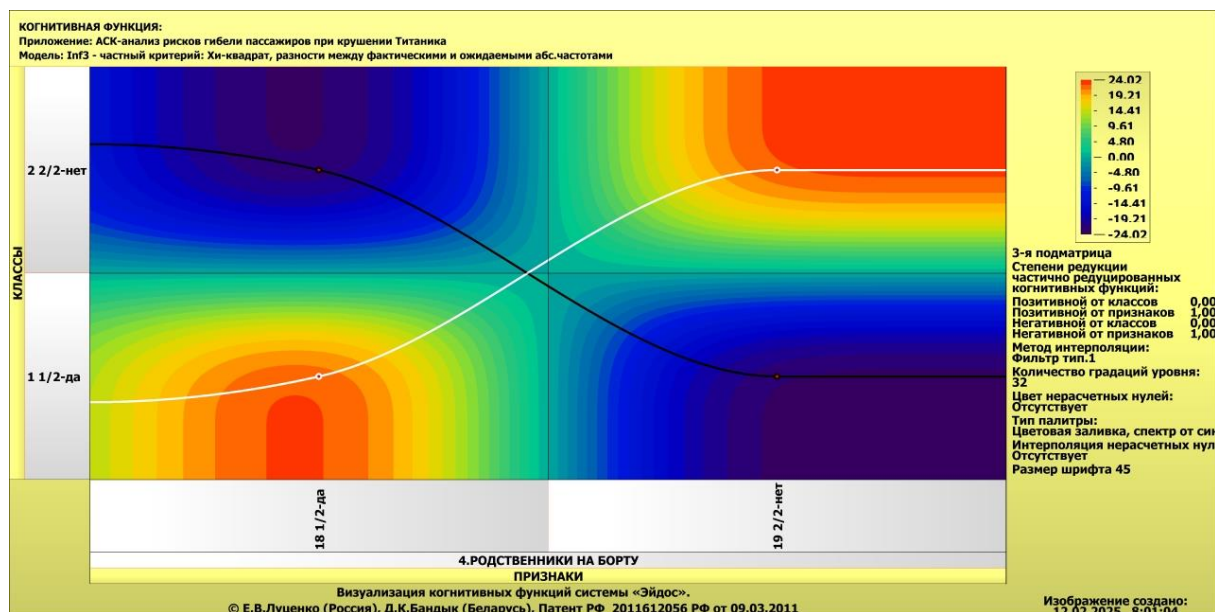


Рисунок 35. Примеры когнитивных функций в СК-модели INF3

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

3.8.10. Значимость описательных шкал и их градаций

3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области

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

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это вариабельность информативностей (режим 3.7.5 системы «Эйдос») (рисунок 36):

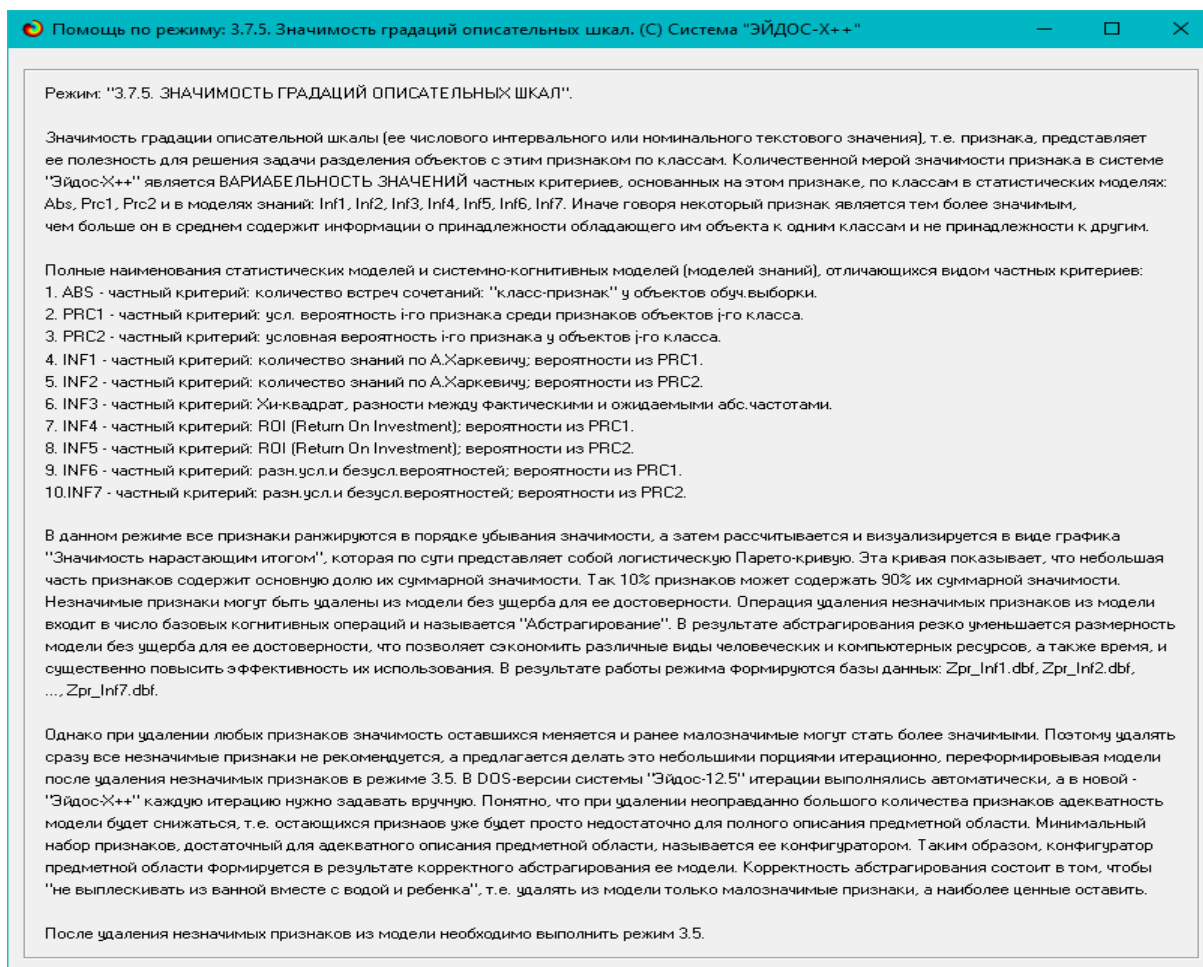


Рисунок 36. Help режима 3.7.5, поясняющий смысл значимости градаций описательных шкал

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

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

3.8.10.2. Конкретное решение задачи в данной работе

На рисунке 37 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели PRC1:

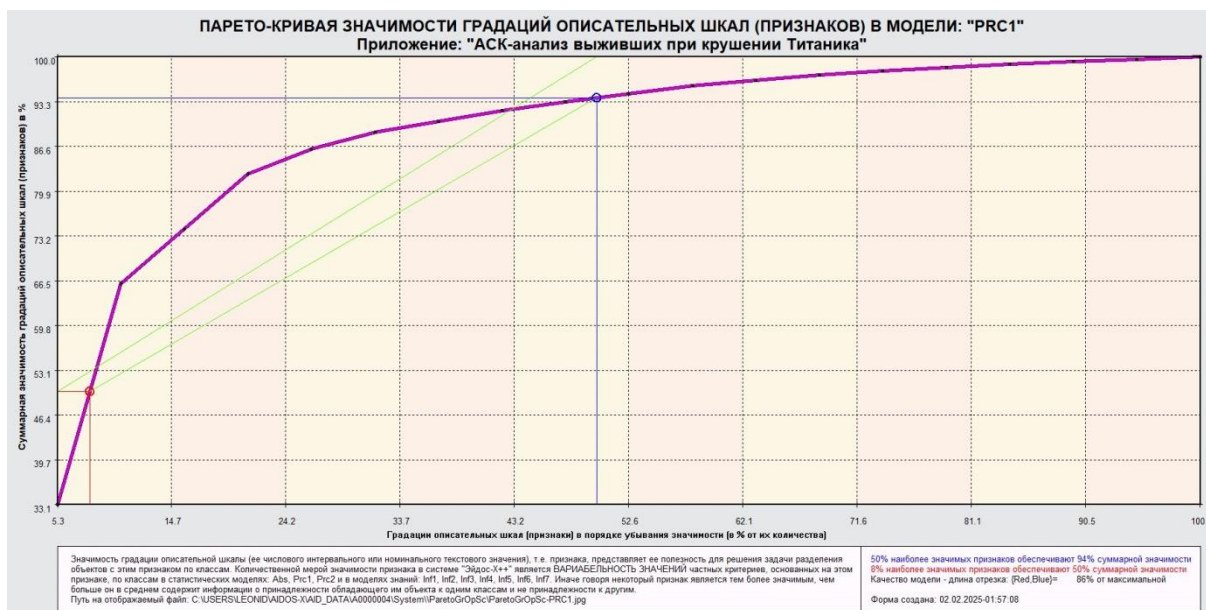


Рисунок 37. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели PRC1

Из рисунка 37 видно, что примерно 8% наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 92% суммарного влияния. На рисунке 38 система «Эйдос» привела рейтинг качества статистических и системно-когнитивных моделей. Кроме того на этом рисунке приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях.

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

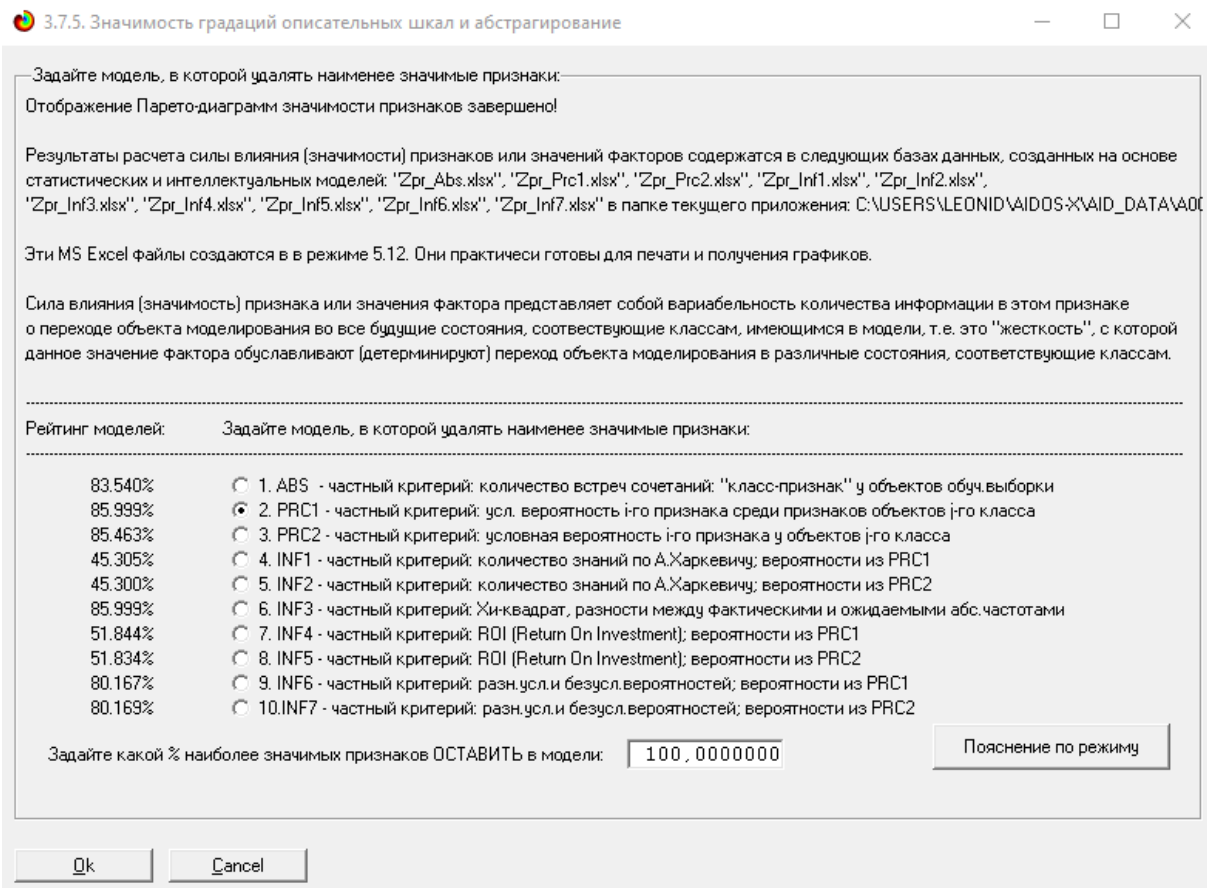


Рисунок 38. Рейтинг качества статистических и системно-когнитивных моделей и имена Excel-файлов с информацией о силе и направлении влияния значений факторов в этих моделях

Таблица 16 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели PRC1

NUM	№%	Код значения фактора	Наименование фактора и его значения	Кода фактора	Значимость, %	Значимость кумулятивно, %
1	5,2631579	4	ПОЛ-1/2-жен	2	33,0539080	33,0539080
2	10,5263158	5	ПОЛ-2/2-муж	2	33,0383311	66,0922391
3	15,7894737	18	РОДСТВЕННИКИ НА БОРТУ-1/2-да	4	8,2060887	74,2983278
4	21,0526316	19	РОДСТВЕННИКИ НА БОРТУ-2/2-нет	4	8,1905122	82,4888400
5	26,3157895	2	КЛАСС-2/3-Первый	1	3,7933537	86,2821937
6	31,5789474	3	КЛАСС-3/3-Третий	1	2,4767067	88,7589004
7	36,8421053	15	ВОЗРАСТ-10/12- {39.0000000, 45.0000000}	3	1,6447840	90,4036844
8	42,1052632	10	ВОЗРАСТ-5/12- {23.0000000, 25.0000000}	3	1,6138355	92,0175198
9	47,3684211	1	КЛАСС-1/3-Второй	1	1,3010704	93,3185902
10	52,6315789	7	ВОЗРАСТ-2/12- {16.0000000, 20.0000000}	3	1,2440921	94,5626823

11	57,8947368	17	ВОЗРАСТ-12/12- {51.0000000, 76.0000000}	3	1,1194779	95,6821602
12	63,1578947	9	ВОЗРАСТ-4/12- {21.0000000, 23.0000000}	3	0,8710694	96,5532296
13	68,4210526	6	ВОЗРАСТ-1/12- {1.0000000, 16.0000000}	3	0,7778137	97,3310433
14	73,6842105	8	ВОЗРАСТ-3/12- {20.0000000, 21.0000000}	3	0,5892528	97,9202961
15	78,9473684	14	ВОЗРАСТ-9/12- {33.0000000, 39.0000000}	3	0,5294052	98,4497014
16	84,2105263	12	ВОЗРАСТ-7/12- {27.0000000, 30.0000000}	3	0,5265360	98,9762373
17	89,4736842	16	ВОЗРАСТ-11/12- {45.0000000, 51.0000000}	3	0,4047910	99,3810283
18	94,7368421	11	ВОЗРАСТ-6/12- {25.0000000, 27.0000000}	3	0,3094858	99,6905142
19	100,000000 0	13	ВОЗРАСТ-8/12- {30.0000000, 33.0000000}	3	0,3094858	100,0000000

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Zpr_Prc1.xlsx

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

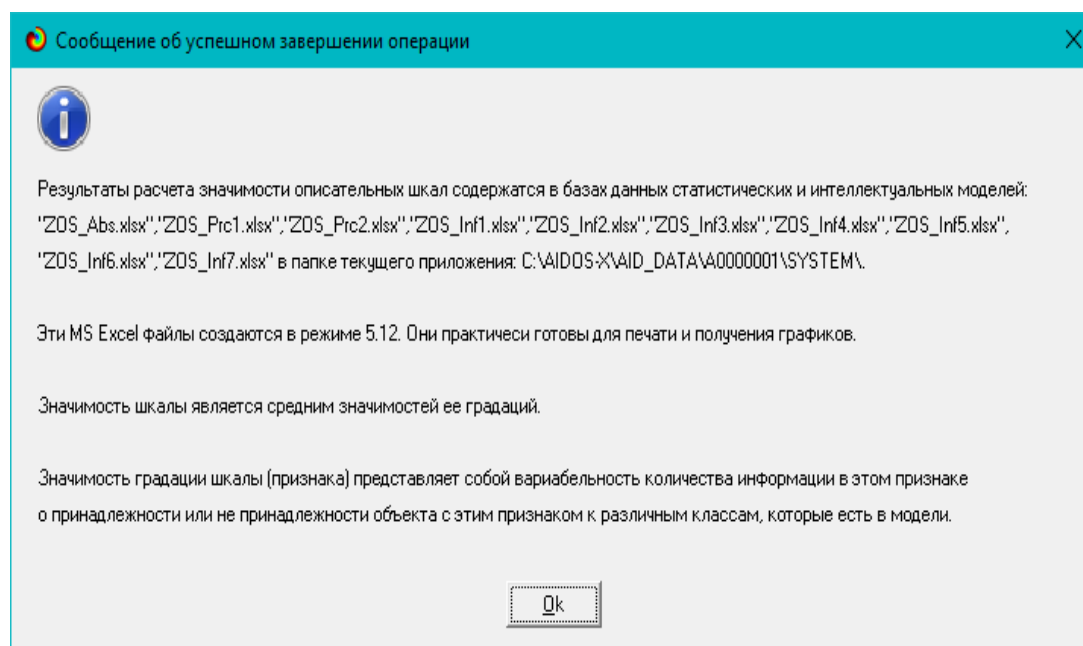


Рисунок 39. Имена Excel-файлов с информацией о силе влияния факторов в статистических и системно-когнитивных моделях

В таблице 17 приведена информация о силе влияния факторов на переход объекта моделирования в различные будущие состояния, соответствующие классам, в системно-когнитивной модели PRC1.

Таблица 17 – Сила влияния факторов на поведение объекта моделирования в системно-когнитивной модели PRC1

№	Код	Наименование фактора	Значимость фактора, %	Значимость фактора кумулятивно, %
1	2	ПОЛ	74,1003098	74,1003098
2	4	РОДСТВЕННИКИ НА БОРТУ	18,3832962	92,4836060
3	1	КЛАСС	5,6589915	98,1425975
4	3	ВОЗРАСТ	1,8574025	100,0000000

Источник: c:\Aidos-X\AID_DATA\A0000004\System\ZOS_Prc1.xlsx

Из таблицы 17 видно, что примерно 74% суммарного влияния на поведение объекта моделирования обусловлено полом пассажира, а 18% наличием родственников на борту: в сумме около 92%. Класс и возраст оказывают сравнительно меньшее влияние: в сумме около 8%. Таким образом, два наиболее значимых фактора, т.е. 50% от всех факторов, обеспечивают примерно 92% суммарного влияния на объект моделирования, а 50% фактора дают суммарно 8% влияния. *Пол пассажира является одним из двух факторов, наиболее сильно влияющих на выживаемость пассажира при крушении «Титаника».*

3.8.11. Степень детерминированности классов и классификационных шкал

3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области

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

3.8.11.2. Конкретное решение задачи в данной работе

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



Сообщение об успешном завершении операции



Отображение Парето-диаграмм степени сформированности классов завершено!

Результаты расчета степени детерминированности (значимости) классов содержатся в следующих базах данных, созданных на основе статистических и интеллектуальных моделей: 'Zkl_Abs.xlsx', 'Zkl_Prc1.xlsx', 'Zkl_Prc2.xlsx', 'Zkl_Inf1.xlsx', 'Zkl_Inf2.xlsx', 'Zkl_Inf3.xlsx', 'Zkl_Inf4.xlsx', 'Zkl_Inf5.xlsx', 'Zkl_Inf6.xlsx', 'Zkl_Inf7.xlsx' в папке текущего приложения: c:\Aidos\X\AID_DATA\A0000001\System\.

Эти MS Excel файлы создаются в режиме 5.12. Они практически готовы для печати и получения графиков.

Степень детерминированности класса представляет собой вариативность количества информации в всех признаках модели о принадлежности или не принадлежности объекта с этим признаком к данному классу, т.е. это "жесткость", с которой значения факторов обуславливают (детерминируют) переход объекта моделирования в состояние, соответствующее классу.

Ok

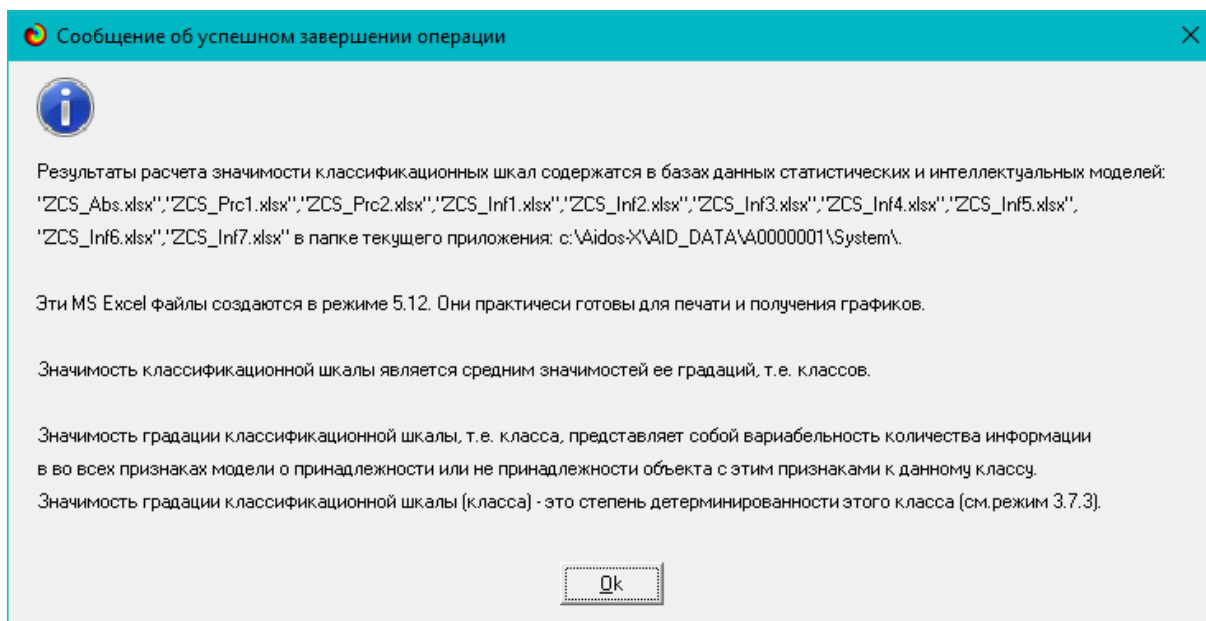


Рисунок 40. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

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

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

Например, 50% наиболее жестко детерминированных классов суммарно обеспечивают примерно 63% степень детерминированности, а 50% суммарной детерминированности обеспечивают 37% наиболее жестко детерминированных классов.

Таблица 18 – Степень детерминированности классов в СК-модели PRC1

№	Код	Наименование класса	Значимость, %	Значимость кумулятивно, %
1	2	ВЫЖИЛ-2/2-нет	63,6363636	63,6363636
2	1	ВЫЖИЛ-1/2-да	36,3636364	100,0000000

Источник: c:\Aidos-X\AID_DATA\A0000004\System\Zkl_Prc1.xlsx

В таблице 18 приведена информация о степени детерминированности классов значениями факторов в системно-когнитивной модели PRC1.

Степень детерминированности классификационных шкал представлена в таблицах, наименования которых приведены на третьем рисунке 40. Но поскольку в данном приложении одна классификационная шкала, то рейтинг шкал по силе детерминированности состоит из одной шкалы и поэтому эти таблицы здесь не приводятся.

4. DISCUSSION (ОБСУЖДЕНИЕ)

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

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

Достижением данной работы является:

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

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

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

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

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

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

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

Для этого надо скачать систему с сайта разработчика по ссылке на странице: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос-приложение №421. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

В условиях продолжающегося интереса к историческим событиям, особенно к таким трагедиям, как крушение «Титаника» 1912 года, важно углубить понимание факторов, повлиявших на выживаемость пассажиров. В ряде исследований, проведенных специалистами в области статистики, обосновывается необходимость анализа выживаемости на основе таких переменных, как пол, возраст, социальный класс и наличие или отсутствие родственников на борту. Для анализа эмпирических данных использованы традиционные методы. Эти исследования показали, что пол и социальный класс значительно влияли на шансы выживания, в то время как возраст оказывал как положительное, так и отрицательное влияние в зависимости от прочих факторов. В данной статье для анализа эмпирических данных в области статистики и исторических исследований предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исторических исследований, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей в данной предметной области.

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

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

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

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

REFERENCES (ЛИТЕРАТУРА)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

37. Работы по виртуальной реальности: http://lc.kubagro.ru/aidos/Virtual_reality_publications.htm

38. Работы по когнитивной ветеринарии: http://lc.kubagro.ru/aidos/Publications_on_cognitive_veterinary_medicine.htm

39. Работы по когнитивной агрономии и когнитивной ампелографии: http://lc.kubagro.ru/aidos/Works_on_cognitive_agronomy.htm

40. Работы по тематике, связанной с АПК: http://lc.kubagro.ru/aidos/Work_with_agricultural.htm

41. Работы по наукометрии: http://lc.kubagro.ru/aidos/Works_on_scientometrics.htm

42. Работы о высших формах сознания, перспективах человека, технологии и общества: http://lc.kubagro.ru/aidos/Works_on_higher_forms_of_consciousness.htm

43. Работы по разработке и применению профиограмм и тестов (психологических, профориентационных, медицинских и ветеринарных): http://lc.kubagro.ru/aidos/Work_on_the_development_and_application_tests.htm

44. Работы по сценарному автоматизированному системно-когнитивному анализу (сценарный АСК-анализ): http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm

45. MVP-проект «Внедрение технологий АСК-анализа и системы «Эйдос» для решения задач АПК»: <http://lc.kubagro.ru/aidos/MVP-projects.htm>

46. Кратко об АСК-анализе и системе «Эйдос»: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

47. Ссылки на видео-занятия и проф.Е.В.Луценко в Пермском национальном университете: <https://bigbluebutton.pstu.ru/b/w3y-2ir-ukd-bqn> (2021), <https://bigbluebutton.pstu.ru/b/3kc-n8a-gon-tjz> (2022), в Кубанском государственном университете и Кубанском государственном аграрном университете: <https://disk.yandex.ru/d/knISAD5qzV83Ng?w=1>

48. Луценко, Е. В. Революция начала XXI века в искусственном интеллекте: глубинные механизмы и перспективы / Е. В. Луценко, Н. С. Головин. – Краснодар: Кубанский государственный аграрный университет им. И.Т. Трубилина, 2024. – 394 с. – DOI 10.13140/RG.2.2.17056.56321. – EDN OMIPII.

49. Луценко Е.В. Системы искусственного интеллекта как системы автоматизации процесса научного познания и удвоение номенклатуры научных специальностей путем применения этих систем для исследований в различных направлениях науки / Е.В. Луценко, Н.С. Головин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2024. – №01(195). С. 74 – 111. – IDA [article ID]: 1952401009. – Режим доступа: <http://ej.kubagro.ru/2024/01/pdf/09.pdf>, 2,375 у.п.л.