

УДК 004.8

4.1.1. Информационные системы и технологии

**АСК-АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА
ИСХОД ТЕННИСНЫХ МАТЧЕЙ**

Яковенко Анна Юрьевна

Студент группы ПИ2201

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

В условиях растущей популярности тенниса и развития индустрии спортивных прогнозов актуальной задачей является повышение точности прогнозирования исходов матчей. Однако этому могут препятствовать множество взаимосвязанных факторов различной природы: рейтинг игроков, тип покрытия корта, историческая статистика личных встреч, психологическое состояние и другие. В ряде научных работ обосновывается перспективность применения методов искусственного интеллекта для анализа эмпирических данных в спортивной аналитике. Для анализа таких данных традиционно используются статистические методы и алгоритмы машинного обучения. Эти исследования показали, что учет множества факторов позволяет повысить точность прогнозов, хотя различные подходы демонстрируют разную эффективность в зависимости от предметной области. В данной статье для анализа эмпирических данных в области теннисной аналитики предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример на основе данных о теннисных матчах. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исследований в области спортивной аналитики, 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», ТЕННИС, ПРОГНОЗИРОВАНИЕ, СПОРТИВНАЯ АНАЛИТИКА

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

UDC 004.8 UDC 004.8

4.1.1. Information systems and technologies

**ASK-ANALYSIS OF FACTORS AFFECTING
THE OUTCOME OF TENNIS MATCHES**

Anna Yuryevna Yakovenko

Student, Group PI2201

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

In the context of the growing popularity of tennis and the development of the sports forecasting industry, improving the accuracy of match outcome predictions has become a relevant task. However, this is complicated by numerous interrelated factors of various natures: player rankings, court surface type, historical head-to-head statistics, psychological state, and others. A number of scientific studies substantiate the prospects of applying artificial intelligence methods for analyzing empirical data in sports analytics. Traditionally, statistical methods and machine learning algorithms have been used to analyze such data. These studies have shown that accounting for multiple factors can improve prediction accuracy, although different approaches demonstrate varying effectiveness depending on the subject area. In this article, for the analysis of empirical data in the field of tennis analytics, we propose using Automated System-Cognitive Analysis (ASC-analysis) and its software toolkit – the intelligent system "Eidos". A detailed numerical example based on tennis match data is provided. This example contains many diverse visual tabular and graphical output forms and can be used for training in the application of ASC-analysis and the Eidos system for scientific research in sports analytics, for developing practical recommendations, and for substantiating scientific provisions on the mechanisms of causal relationships in this subject area.

Keywords: ASC-ANALYSIS, AUTOMATED SYSTEM-COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS", TENNIS, FORECASTING, SPORTS ANALYTICS

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.....	4
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.....	26
3.3.1. Универсальная постановка задачи, не зависящая от предметной области	26
3.3.2. Конкретное решение задачи в данной работе	33
3.4. Задача-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	37
3.4.1. Универсальная постановка задачи, не зависящая от предметной области	37
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».....	43
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»	45
3.7. Задача-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ.....	48
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	48
3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области	48
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»	51
3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области	51
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	55
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы).....	55
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области	55
3.8.1.2. Конкретное решение задачи в данной работе	55
3.8.2. Кластерно-конструктивный анализ классов.....	58

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области	58
3.8.2.2. Конкретное решение задачи в данной работе	59
3.8.3. Кластерно-конструктивный анализ значений описательных шкал	63
3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области	63
3.8.3.2. Конкретное решение задачи в данной работе	64
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны.....	68
3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области	68
3.8.4.2. Конкретное решение задачи в данной работе	69
3.8.5. Нелокальная нейронная сеть.....	70
3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области	70
3.8.5.2. Конкретное решение задачи в данной работе	71
3.8.6. 3D-интегральные когнитивные карты Ошибка! Закладка не определена.	
3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области	Ошибка! Закладка не определена.
3.8.6.2. Конкретное решение задачи в данной работе	Ошибка! Закладка не определена.
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	72
3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области	72
3.8.7.2. Конкретное решение задачи в данной работе	73
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	74
3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области	74
3.8.8.2. Конкретное решение задачи в данной работе	74
3.8.9. Когнитивные функции	77
3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области	77
3.8.9.2. Конкретное решение задачи в данной работе	77
3.8.10. Значимость описательных шкал и их градаций.....	80
3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области	80
3.8.10.2. Конкретное решение задачи в данной работе	81
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)	88
5. CONCLUSIONS (ВЫВОДЫ)	89
REFERENCES (ЛИТЕРАТУРА).....	90

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

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

Спортивные события, в частности теннисные матчи, представляют собой сложные динамические системы, исход которых зависит от множества взаимосвязанных факторов: рейтинга игроков, покрытия корта, даты проведения турнира, исторической статистики встреч и других параметров.

Прогнозирование исходов спортивных соревнований является ярко выраженным эмерджентным (системным) свойством, зависящим от многих факторов, формализуемых как в числовом, так и в категориальном виде.

В ряде научных работ обосновывается перспективность применения интеллектуальных систем для анализа эмпирических данных в спортивной аналитике. Для анализа таких данных традиционно используются статистические методы, машинное обучение и нейросетевые подходы.

В данной статье для анализа эмпирических данных в области спортивной аналитики предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий — интеллектуальную систему «Эйдос». Приводится подробный численный пример на основе данных о теннисных матчах. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исследований в области спортивной аналитики, 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) — теннисный матч.

Предмет исследования — выявление причинно-следственных зависимостей исхода теннисного матча от совокупности факторов: рейтинга игроков, типа покрытия, даты проведения турнира и других параметров.

1.3. Проблема, решаемая в работе и ее актуальность

Спецификой данной задачи является то, что независимые переменные (факторы, влияющие на объект моделирования) имеют

разнородную природу: часть из них формализуется в виде числовых переменных (рейтинг, счет), часть — в виде лингвистических (категориальных) переменных (тип покрытия, название турнира). При этом зависимая переменная (победитель матча) измеряется в номинальной 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 в разных типах шкал и разных единицах измерения.

Решение в данной работе проблемы сопоставимости при выявлении причинно-следственных зависимостей исхода теннисного матча от совокупности факторов 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 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

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)

интеллекта, 360 публикации в изданиях, входящих в перечень ВАК РФ и приравненных им (по данным [РИНЦ](#)), 17 публикаций в изданиях, входящих в ядро РИНЦ, 6 статей в журналах, входящих в [WoS](#), 7 публикаций в журналах, входящих в [Скопус](#)³ [5, 6, 7].

Три монографии включены в фонды библиотеки конгресса США⁴.

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

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

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

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

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

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

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/>

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-ml-dotnet-work>, <http://chat.openai.com/>, <https://poe.com/>, <https://neural-university.ru/>, <https://dzen.ru/a/ZCKZRKvrlEMBWO8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://poe.com/Aidos-X>, <https://rudalle.ru/>, <https://bard.google.com/>, <https://chatbot.theb.ai/>, <https://problembo.com/ru/services> (может пригодиться - почта на 10 минут: <https://10minutemail.net/>), <https://poe.com/GPT-3.5-Turbo-Instruct>, <https://www.seaart.ai/home>, <https://ui.chatai.com/>.

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

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

- является одной из первых и наиболее популярных отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий

искусственного интеллекта и программирования: есть акт внедрения системы «Эйдос» 1987 года (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);

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

- имеет «нулевой порог входа»:

- содержит большое количество интеллектуальных локальных (т.е. поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 392, соответственно: http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf, http://lc.kubagro.ru/Presentation_LutsenkoEV.pdf);

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

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

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

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



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

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

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

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

- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) предлагает без какой-либо предварительной обработки осмыслить те данные, которые есть, и, тем самым, преобразовать их в информацию, а

затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для принятия решений и управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

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

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

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

В разработке системы «Эйдос» были следующие этапы:

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

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

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

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

4-й этап, «эра MS Windows-10»: 2020-2021 годы. С 13.12.2020 года по настоящее время система «Эйдос» развивается на языке [Аляска-2.0](#) + [Экспресс++](#). Библиотека 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».

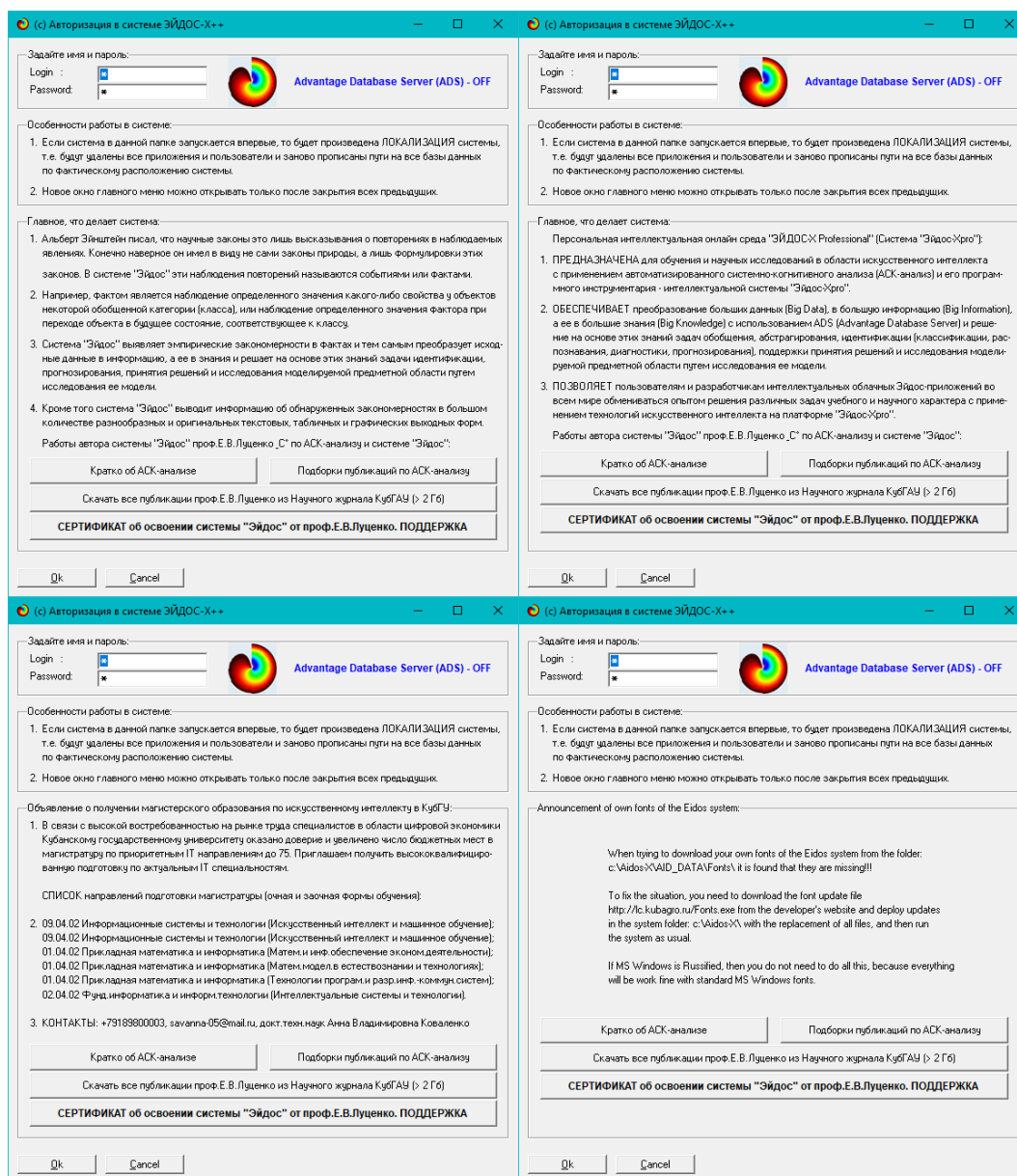
[Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения⁷](#)

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 видеогрaмма DOS-версии системы «Эйдос» (до 2012 года)⁸



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

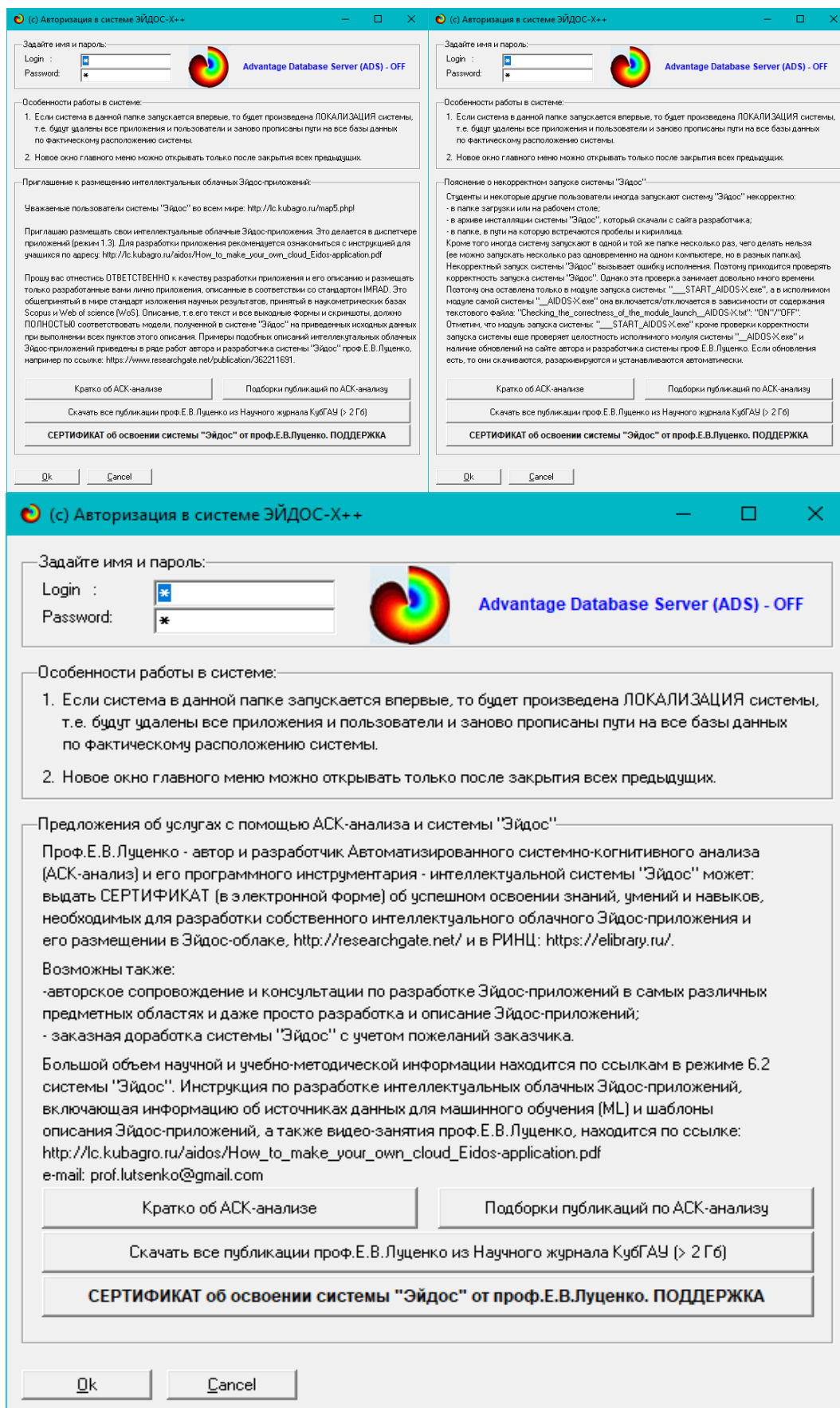


Рисунок 2. Титульные видеогаммы текущей версии системы «Эйдос»

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

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

Задача-1. Когнитивная структуризация предметной области.

Задача-2. Формализация предметной области.

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

Задача-4. Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

Задача-6. Системная идентификация и прогнозирование.

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

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

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

8.2) кластерно-конструктивный анализ классов;

8.3) кластерно-конструктивный анализ значений описательных шкал;

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

8.5) нелокальная нейронная сеть;

8.6) 3d-интегральные когнитивные карты;

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

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

8.9) когнитивные функции;

8.10) значимость описательных шкал и их градаций;

8.11) степень детерминированности классов и классификационных шкал).

Для данной работы особое значение имеет решение подзадачи 8.1, т.к. она позволяет детально исследовать влияние каждого значения каждого фактора на вероятность победы конкретного теннисиста в матче.

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**

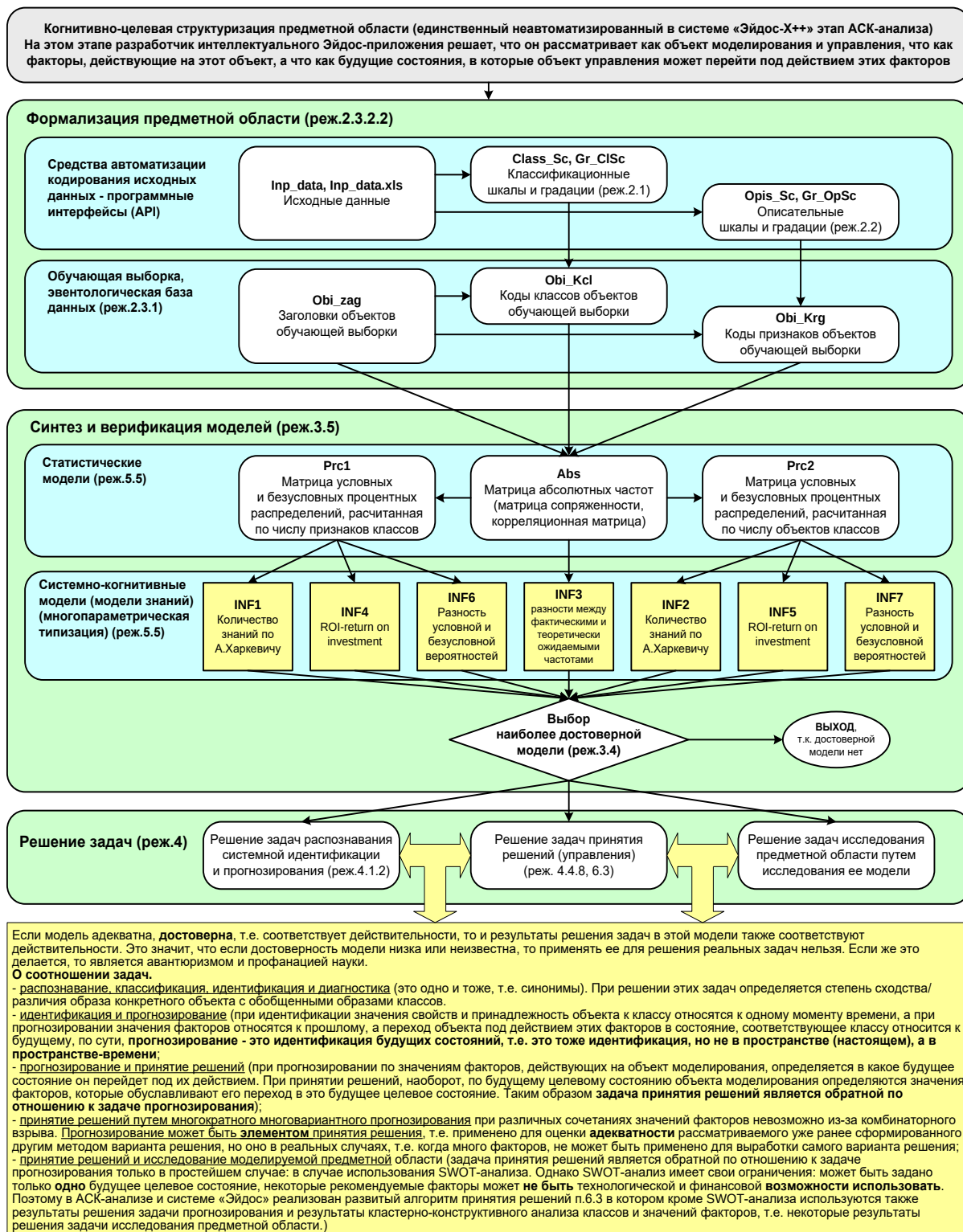


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос» (этапы АСК-анализа)

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 теннисный матч, в качестве *факторов*: ТУРНИР, ДАТА, ПОКРЫТИЕ, ИГРОК_1, ИГРОК_2, РЕЙТИНГ_1, РЕЙТИНГ_2, КОЭФ_1, КОЭФ_2 (таблица 1), а в качестве *результатов* действия этих факторов: ПОБЕДИТЕЛЬ (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	ТУРНИР
2	ДАТА
3	ПОКРЫТИЕ
4	ИГРОК_1
5	ИГРОК_2
6	РЕЙТИНГ_1
7	РЕЙТИНГ_2
8	КОЭФ_1
9	КОЭФ_2

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

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	ПОБЕДИТЕЛЬ

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Inp_data_исх.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

Датасет сформирован на основе данных теннисных матчей с платформы Kaggle (<https://www.kaggle.com/competitions>) и содержит следующие атрибуты:

Таблица 3 – Исходные данные

	A	B	C	D	E	F	G	H	I	J	K
1	Tournament	Date	Surface	Player_1	Player_2	Rank_1	Rank_2	Odd_1	Odd_2	Score	Winner
2	Australian Hardcourt Championships	2000-01-03	Hard	Dosedel S.	Ljubicic I.	63	77	-1	-1	6-4 6-2	Dosedel S.
3	Australian Hardcourt Championships	2000-01-03	Hard	Clement A.	Enqvist T.	56	5	-1	-1	3-6 3-6	Enqvist T.
4	Australian Hardcourt Championships	2000-01-03	Hard	Escude N.	Baccanello P.	40	655	-1	-1	6-7 7-5 6-3	Escude N.
5	Australian Hardcourt Championships	2000-01-03	Hard	Knippschild J.	Federer R.	87	65	-1	-1	1-6 4-6	Federer R.
6	Australian Hardcourt Championships	2000-01-03	Hard	Fromberg R.	Woodbridge T.	81	198	-1	-1	7-6 5-7 6-4	Fromberg R.
7	Australian Hardcourt Championships	2000-01-03	Hard	Arthurs W.	Gambill J.M.	105	58	-1	-1	6-3 6-7 4-6	Gambill J.M.
8	Australian Hardcourt Championships	2000-01-03	Hard	Grosjean S.	Ilie A.	26	51	-1	-1	6-2 6-1	Grosjean S.
9	Australian Hardcourt Championships	2000-01-03	Hard	Balcells J.	Henman T.	218	11	-1	-1	4-6 6-7	Henman T.
10	Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Woodforde M.	24	129	-1	-1	6-3 2-6 6-1	Hewitt L.
11	Australian Hardcourt Championships	2000-01-03	Hard	Tebbutt M.	Lisnard J.	351	134	-1	-1	6-7 7-6 4-6	Lisnard J.
12	Australian Hardcourt Championships	2000-01-03	Hard	Martin A.	Spadea V.	55	21	-1	-1	6-2 6-2	Martin A.
13	Australian Hardcourt Championships	2000-01-03	Hard	Draper S.	Norman M.	155	15	-1	-1	4-6 4-6	Norman M.
14	Australian Hardcourt Championships	2000-01-03	Hard	Petrovic D.	Huet S.	208	122	-1	-1	3-6 6-3 6-3	Petrovic D.
15	Australian Hardcourt Championships	2000-01-03	Hard	Pozzi G.	Sekulov J.	79	142	-1	-1	4-6 6-2 2-6	Sekulov J.
16	Australian Hardcourt Championships	2000-01-03	Hard	Stoltenberg J.	Hrbaty D.	84	20	-1	-1	6-3 6-1	Stoltenberg J.
17	Australian Hardcourt Championships	2000-01-03	Hard	Goldstein P.	Vinck C.	88	227	-1	-1	6-3 3-6 4-6	Vinck C.
18	Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Federer R.	5	65	-1	-1	7-6 6-4	Enqvist T.
19	Australian Hardcourt Championships	2000-01-03	Hard	Dosedel S.	Escude N.	63	40	-1	-1	6-7 4-6	Escude N.
20	Australian Hardcourt Championships	2000-01-03	Hard	Grosjean S.	Gambill J.M.	26	58	-1	-1	3-6 7-6 6-2	Grosjean S.
21	Australian Hardcourt Championships	2000-01-03	Hard	Lisnard J.	Henman T.	134	11	-1	-1	6-3 4-6 2-6	Henman T.
22	Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Petrovic D.	24	208	-1	-1	6-0 6-2	Hewitt L.
23	Australian Hardcourt Championships	2000-01-03	Hard	Sekulov J.	Martin A.	142	55	-1	-1	6-7 6-7	Martin A.
24	Australian Hardcourt Championships	2000-01-03	Hard	Norman M.	Fromberg R.	15	81	-1	-1	7-5 6-4	Norman M.
25	Australian Hardcourt Championships	2000-01-03	Hard	Vinck C.	Stoltenberg J.	227	84	-1	-1	6-3 4-6 4-6	Stoltenberg J.
26	Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Grosjean S.	5	26	-1	-1	7-6 7-5	Enqvist T.
27	Australian Hardcourt Championships	2000-01-03	Hard	Henman T.	Escude N.	11	40	-1	-1	3-6 1-6	Escude N.
28	Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Stoltenberg J.	24	84	-1	-1	6-2 6-2	Hewitt L.
29	Australian Hardcourt Championships	2000-01-03	Hard	Martin A.	Norman M.	55	15	-1	-1	4-6 3-6	Norman M.
30	Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Norman M.	5	15	-1	-1	3-6 6-2 6-2	Enqvist T.

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

Таблица 3 – Excel-таблица исходных данных в стандарте системы «Эйдос»

A	B	C	D	E	F	G	H	I	J	K
Tournament	Date	Surface	Player_1	Player_2	Rank_1	Rank_2	Odd_1	Odd_2	Score	Winner
Australian Hardcourt Championships	2000-01-03	Hard	Dosedel S.	Ljubicic I.	83	77	-1	-1	6-4 6-2	Dosedel S.
Australian Hardcourt Championships	2000-01-03	Hard	Clement A.	Enqvist T.	56	5	-1	-1	3-6 3-6	Enqvist T.
Australian Hardcourt Championships	2000-01-03	Hard	Escude N.	Baccanello P.	40	655	-1	-1	6-7 7-5 6-3	Escude N.
Australian Hardcourt Championships	2000-01-03	Hard	Knippschild J.	Federer R.	87	65	-1	-1	1-6 4-6	Federer R.
Australian Hardcourt Championships	2000-01-03	Hard	Fromberg R.	Woodbridge T.	81	198	-1	-1	7-6 5-7 6-4	Fromberg R.
Australian Hardcourt Championships	2000-01-03	Hard	Arthurs W.	Gambill J.M.	105	58	-1	-1	6-3 6-7 4-6	Gambill J.M.
Australian Hardcourt Championships	2000-01-03	Hard	Grosjean S.	Ilie A.	26	51	-1	-1	6-2 6-1	Grosjean S.
Australian Hardcourt Championships	2000-01-03	Hard	Balcells J.	Henman T.	219	11	-1	-1	4-6 6-7	Henman T.
Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Woodforde M.	24	129	-1	-1	6-3 2-6 6-1	Hewitt L.
Australian Hardcourt Championships	2000-01-03	Hard	Tebbutt M.	Lisnard J.	351	134	-1	-1	6-7 7-6 4-6	Lisnard J.
Australian Hardcourt Championships	2000-01-03	Hard	Martin A.	Spadea V.	55	21	-1	-1	6-2 6-2	Martin A.
Australian Hardcourt Championships	2000-01-03	Hard	Draper S.	Norman M.	155	15	-1	-1	4-6 4-6	Norman M.
Australian Hardcourt Championships	2000-01-03	Hard	Petrovic D.	Huet S.	208	122	-1	-1	3-6 6-3 6-3	Petrovic D.
Australian Hardcourt Championships	2000-01-03	Hard	Pozzi G.	Sekulov J.	79	142	-1	-1	4-6 6-2 2-6	Sekulov J.
Australian Hardcourt Championships	2000-01-03	Hard	Stollenberg J.	Hrbaty D.	84	20	-1	-1	6-3 6-1	Stollenberg J.
Australian Hardcourt Championships	2000-01-03	Hard	Goldstein P.	Vinck C.	88	227	-1	-1	6-3 3-6 4-6	Vinck C.
Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Federer R.	5	65	-1	-1	7-6 6-4	Enqvist T.
Australian Hardcourt Championships	2000-01-03	Hard	Dosedel S.	Escude N.	63	40	-1	-1	6-7 4-6	Escude N.
Australian Hardcourt Championships	2000-01-03	Hard	Grosjean S.	Gambill J.M.	26	58	-1	-1	3-6 7-6 6-2	Grosjean S.
Australian Hardcourt Championships	2000-01-03	Hard	Lisnard J.	Henman T.	134	11	-1	-1	6-3 4-6 2-6	Henman T.
Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Petrovic D.	24	208	-1	-1	6-0 6-2	Hewitt L.
Australian Hardcourt Championships	2000-01-03	Hard	Sekulov J.	Martin A.	142	55	-1	-1	6-7 6-7	Martin A.
Australian Hardcourt Championships	2000-01-03	Hard	Norman M.	Fromberg R.	15	81	-1	-1	7-5 6-4	Norman M.
Australian Hardcourt Championships	2000-01-03	Hard	Vinck C.	Stollenberg J.	227	84	-1	-1	6-3 4-6 4-6	Stollenberg J.
Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Grosjean S.	5	26	-1	-1	7-6 7-5	Enqvist T.
Australian Hardcourt Championships	2000-01-03	Hard	Henman T.	Escude N.	11	40	-1	-1	3-6 1-6	Escude N.
Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Stollenberg J.	24	84	-1	-1	6-2 6-2	Hewitt L.
Australian Hardcourt Championships	2000-01-03	Hard	Martin A.	Norman M.	55	15	-1	-1	4-6 3-6	Norman M.
Australian Hardcourt Championships	2000-01-03	Hard	Enqvist T.	Norman M.	5	15	-1	-1	3-6 6-2 6-2	Enqvist T.
Australian Hardcourt Championships	2000-01-03	Hard	Escude N.	Hewitt L.	40	24	-1	-1	6-3 1-6 4-6	Hewitt L.
Australian Hardcourt Championships	2000-01-03	Hard	Hewitt L.	Enqvist T.	24	5	-1	-1	3-6 6-3 6-2	Hewitt L.
Gold Flake Open	2000-01-03	Hard	Kumar S.	Agenor R.	1198	94	-1	-1	4-6 4-6	Agenor R.
Gold Flake Open	2000-01-03	Hard	Behrend T.	Popp A.	116	124	-1	-1	6-4 4-6 6-1	Behrend T.
Gold Flake Open	2000-01-03	Hard	Rochus C.	Damm M.	120	104	-1	-1	3-6 0-6	Damm M.
Gold Flake Open	2000-01-03	Hard	Delgado J.	Black B.	162	67	-1	-1	6-2 7-6	Delgado J.
Gold Flake Open	2000-01-03	Hard	Spottl M.	Golmard J.	200	38	-1	-1	3-6 4-6	Golmard J.
Gold Flake Open	2000-01-03	Hard	Hantschk M.	Teleman L.	128	89	-1	-1	7-6 6-3	Hantschk M.
Gold Flake Open	2000-01-03	Hard	Krosiak J.	Jonsson F.	103	127	-1	-1	3-6 3-6	Jonsson F.
Gold Flake Open	2000-01-03	Hard	Ketola T.	Ogorodov O.	163	135	-1	-1	7-5 4-6 6-4	Ketola T.
Gold Flake Open	2000-01-03	Hard	Zib T.	Manta L.	115	117	-1	-1	5-7 2-6	Manta L.
Gold Flake Open	2000-01-03	Hard	Paes L.	Boutter J.	145	137	-1	-1	6-4 7-6	Paes L.
Gold Flake Open	2000-01-03	Hard	Ram A.	Pioline C.	414	14	-1	-1	3-6 1-6	Pioline C.
Gold Flake Open	2000-01-03	Hard	Sanguinetti D.	Voinea A.	121	100	-1	-1	7-6 6-4	Sanguinetti D.
Gold Flake Open	2000-01-03	Hard	Kildemv P.	Stanoutchev O.	342	111	-1	-1	4-6 6-7	Stanoutchev O.

Таблица имеет следующую структуру: – каждая строка описывает одно наблюдение (теннисный матч) с определенным сочетанием значений факторов (турнир, игроки, рейтинги, коэффициенты) и определенным результатом (победитель); – 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), а с другой стороны результатами действия этих факторов, т.е. победителем матча, выраженным в классификационной шкале (желтый фон). Такая структура описания наблюдений в технологиях искусственного интеллекта называется «онтологией» и модели представлений знаний Марвина Мински (1975) называется «фрейм-экземпляр»;

– 1-я колонка (A) – не является шкалой и содержит номер наблюдения или другую идентифицирующую информацию;

– колонки со 2-й по 10-ю (B-K) – это описательные шкалы, формализующие факторы, действующие на объект моделирования. Эти шкалы имеют числовой и текстовый тип и их градациями являются лингвистические и числовые переменные (название турнира, дата, тип покрытия, имена игроков, рейтинги, коэффициенты, счет матча);

– колонка 11-я (L) – это классификационная шкала, описывающая результат действия факторов (победитель матча). В общем случае в исходных данных может быть значительно больше классификационных шкал. В системе «Эйдос» существует ограничение на суммарное

количество градаций классификационных шкал: их должно быть не более 2032;

– при вводе данных в систему «Эйдос» нули и пробелы в исходных данных могут рассматриваться как значащие или как отсутствие данных. В данной работе использован вариант, при котором они считаются отсутствием данных.

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 3.

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).

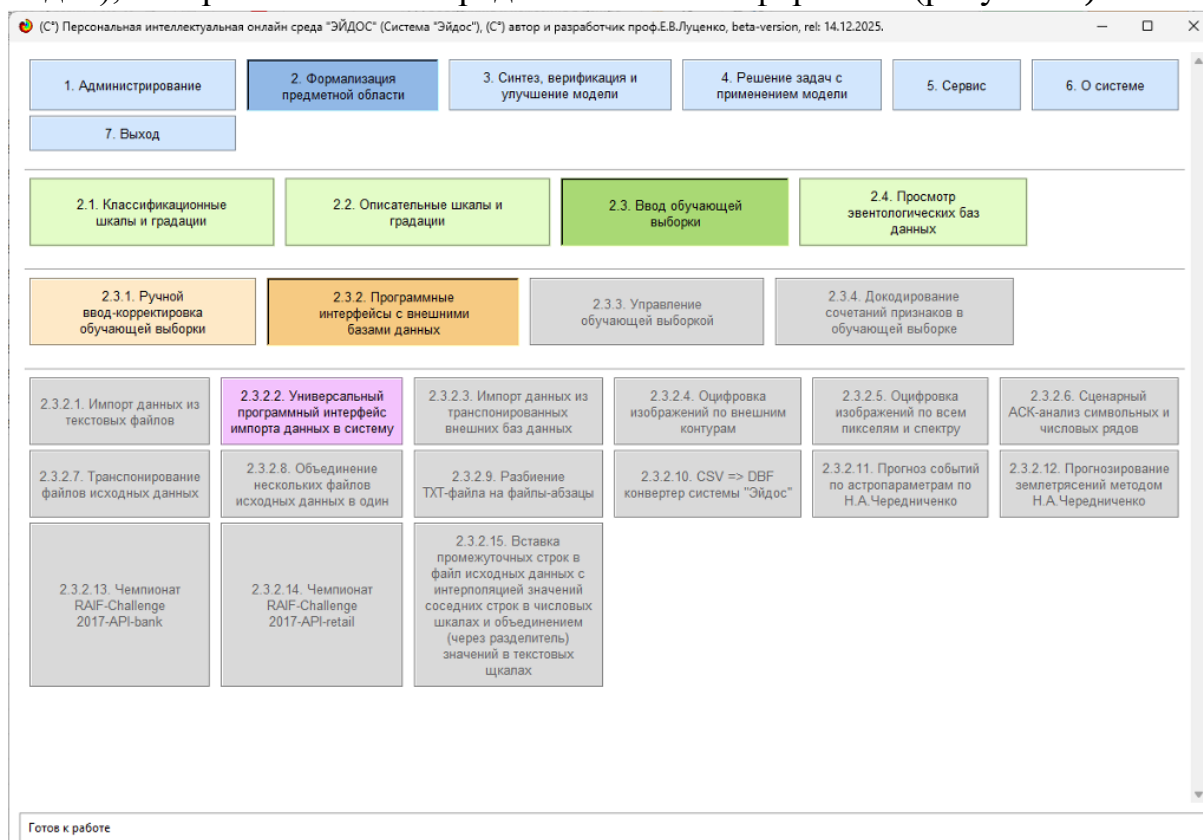


Рисунок 4. Автоматизированные программные интерфейсы (API) системы «Эйдос»

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

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

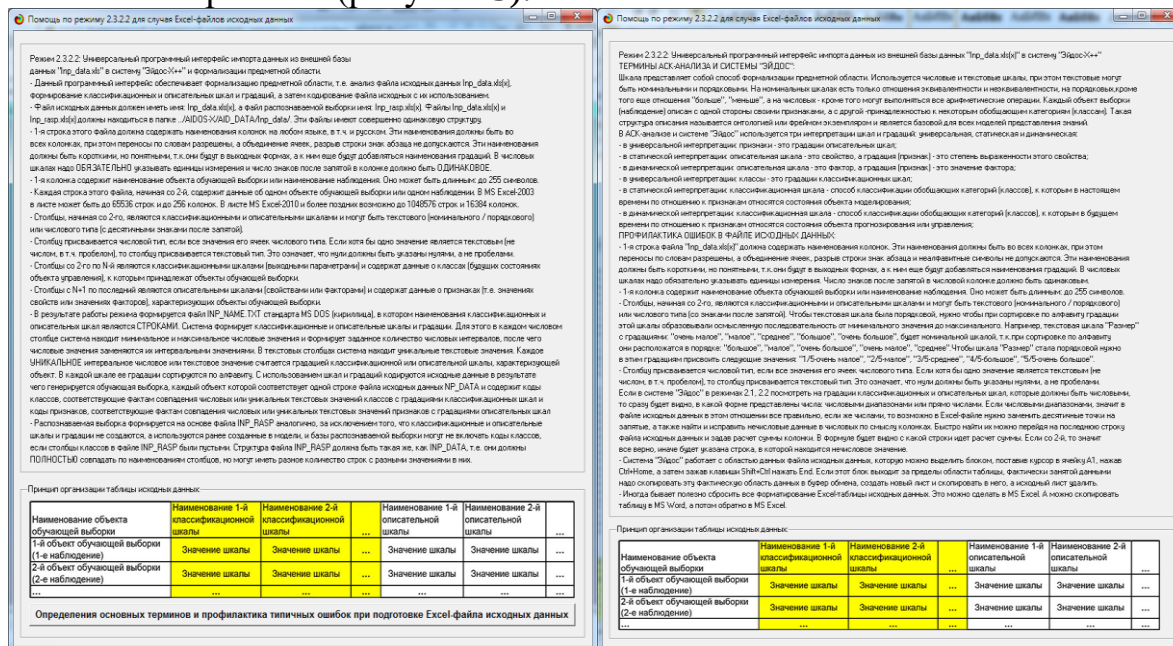


Рисунок 5. Хелпы API-2.3.2.2 системы «Эйдос»

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

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

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

Задайте тип файла исходных данных: "Inp_data":

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

Задайте параметры:

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

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец: Конечный столбец:

Задайте диапазон столбцов описательных шкал:

Начальный столбец: Конечный столбец:

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 ACK-анализа ☐ Применить сценарный метод ACK-анализа
☒ Применить спец.интерпретацию текстовых полей классов ☒ Применить спец.интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

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

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

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

2.3.2.2. Параметры классификационных и описательных шкал и градаций

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [260 классов x 2122 признаков]

Тип	Количество	Суммарное количество	Среднее количество	Количество	Суммарное	Среднее
Числовые	0	0	0.00	5	25	5.00
Текстовые	1	260	260.00	4	2097	524.25
ВСЕГО:	1	260	260.00	9	2122	235.78

Задайте количество числовых диапазонов (интервалов, градаций) в шкале: В классификационных шкалах: 5 В описательных шкалах: 5

ПАРАМЕТРЫ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ

ПАРАМЕТРЫ ШКАЛ И ГРАДАЦИЙ С АДАПТИВНЫМИ ГРАНИЦАМИ И ПРИМЕРНО РАВНЫМ КОЛИЧЕСТВОМ НАБЛЮДЕНИЙ ПО ГРАДАЦИЯМ с коррекцией ошибки округления числа наблюдений по интервалу градации при переходе к следующей градации

КЛАССИФИКАЦИОННАЯ ШКАЛА: код: [1], наим.: "WINNER", тип/число градаций в шкале: "равные интервалы"/260

1 Наим.градации: 1/260-Acasuso J., факт.число наблюдений на градации: 7
 2 Наим.градации: 2/260-Agassi A., факт.число наблюдений на градации: 66
 3 Наим.градации: 3/260-Agenor R., факт.число наблюдений на градации: 6
 4 Наим.градации: 4/260-Alami K., факт.число наблюдений на градации: 25
 5 Наим.градации: 5/260-Allegro Y., факт.число наблюдений на градации: 1
 6 Наим.градации: 6/260-Alonso J., факт.число наблюдений на градации: 4
 7 Наим.градации: 7/260-Alvarez E., факт.число наблюдений на градации: 5
 8 Наим.градации: 8/260-Andersen J.F., факт.число наблюдений на градации: 1
 9 Наим.градации: 9/260-Araci H., факт.число наблюдений на градации: 47
 10 Наим.градации: 10/260-Atmado H., факт.число наблюдений на градации: 1
 11 Наим.градации: 11/260-Archus W., факт.число наблюдений на градации: 29
 12 Наим.градации: 12/260-Bachelot J.F., факт.число наблюдений на градации: 1
 13 Наим.градации: 13/260-Balcells J., факт.число наблюдений на градации: 25
 14 Наим.градации: 14/260-Bastl G., факт.число наблюдений на градации: 21
 15 Наим.градации: 15/260-Behrend T., факт.число наблюдений на градации: 9
 16 Наим.градации: 16/260-Berasategui A., факт.число наблюдений на градации: 4
 17 Наим.градации: 17/260-Bjorkman J., факт.число наблюдений на градации: 43
 18 Наим.градации: 18/260-Black B., факт.число наблюдений на градации: 32
 19 Наим.градации: 19/260-Black W., факт.число наблюдений на градации: 10
 20 Наим.градации: 20/260-Blake J., факт.число наблюдений на градации: 1
 21 Наим.градации: 21/260-Blanco G., факт.число наблюдений на градации: 23

Пересчитать шкалы и градации Записать TXT Выйти на создание модели

Операция завершена

1/7 - Конвертация и ввод файла исходных данных 100%
 2/7 - Создание нового интеллектуального приложения 100%
 3/7 - Анализ файла исходных данных 100%
 4/7 - Формирование классификационных шкал и градаций 0%
 5/7 - Формирование описательных шкал и градаций 100%
 6/7 - Кодирование обучающей выборки 100%
 7/7 - Формирование базы событий 100%

Общий прогресс

Начало: 16.03.2026 10:48:32 Окончание: 16.03.2026 10:49:30
 Прошло: 00:00:58 Осталось: 00:00:00

Сообщения

[10:48:46 INFO]: 5. формирование описательных шкал и градаций
 [10:48:54 INFO]: Сформировано 10 описательных шкал и 2382 градаций.
 [10:48:55 INFO]: 6. Кодирование обучающей выборки
 [10:49:24 INFO]: Таблицы obi_zag, obi_kol, obi_kpr успешно заполнены для 4585 объектов.
 [10:49:24 INFO]: 7. формирование базы событий
 [10:49:30 INFO]: База событий eventsko успешно заполнена. Записано 4585 объектов.
 [10:49:30 SUCCESS]: Операция успешно завершена.

Записать лог в CSV Выход

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.

Таблица 4 – Классификационные шкалы и градации (полностью)

KOD_CLS	NAME_CLS
1	WINNER-1/4585-Acasuso J.
2	WINNER-2/4585-Agassi A.
3	WINNER-3/4585-Agenor R.
...	...
N	WINNER-N/4585-[Имя победителя]

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

Таблица 5 – Описательные шкалы и градации (полностью)

KOD_ATR	NAME_ATR
1	TOURNAMENT-1/2382-Australian Hardcourt Championships
...	...
N	TOURNAMENT-N/2382-[Название турнира]
...	...
N	SURFACE-1/4-Carpet
...	...
N	SURFACE-4/4-Hard
...	...
N	PLAYER_1-1/327-Acasuso J.
...	...
N	PLAYER_1-N/327-[Имя игрока]

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

Таблица 6 – Обучающая выборка (полностью)

NAME_OBJ	N_TOURNAMENT	N_DATE	N_SURFACE	...	N_WINNER
1	1	1	4	...	2
2	1	1	4	...	2
3	1	1	4	...	3
...	...	...	...	...	...
N	N	N	N	...	N

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

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

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

3.3.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

Суть математической модели АСК-анализа состоит в следующем.

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

Таблица 7 – 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 8 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

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

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

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

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

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

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

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

В таблице 10 приведены формулы:

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

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

Таблица 9– 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)	$N_{ij} \text{ – фактическая частота;}$ $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^M \sum_{j=1}^W N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N} \text{ – теоретическая частота.}$	
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	$I_{ij} = \Psi \times \log_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \log_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \log_2 \frac{N_{ij} N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.	---	$I_{ij} = \Psi \times \log_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \log_2 \frac{N_{ij} N}{N_i N_j}$
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	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{N_j} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу	---	$I_{ij} = \frac{N_{ij}}{N_j} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
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	---	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$

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 10 – Матрица системно-когнитивной модели

	Классы					Значимость фактора
	1	...	j	...	W	

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

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

Таблица 11– 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 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

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

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 12 – 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 базовые модели для синтеза и верификации

Базовые статистические модели:

☒ 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: [Помощь](#)

☒ 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

Базовые 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 без усл. вероятностей; вероятности из PRC1

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

Какие модели создавать?

☒ Создавать только базовые модели

☐ Создавать модели 2-го уровня

☐ Создавать модели 3-го уровня

[Помощь](#)

Базовые модели - это модели: Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7 полученные расчетным путем по формулам, приведенным в хелпе режима 3.5 на основе модели Abs.

Модель Abs называется матрицей абсолютных частот и содержит абсолютное количество встреч сочетаний "признак x класс", посчитанное по всей выборке.

Модель Abs еще называют "матрицей сопряженности" или "корреляционной матрицей". Формы по достоверности моделей формируются в режиме 3.4.

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) до

☐ Вообще не менять распознаваемую выборку

Удалять скопированные объекты:

☒ Не удалять

☐ Удалять

[Пояснение по алгоритму верификации](#)

[Подробнее](#)

Выполнить:

☒ Синтез и верификацию

☐ Только верификацию

☐ Только синтез

Задайте процессор

☒ CPU

☐ GPU

Задайте алгоритм

☐ Классика - дольше

☒ Упрощенно-быстрее

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

Расчетный размер БД результатов распознавания Rasp.dbf равен 14194 байт, т.е.: 0.0006610 % от МАХ-возможного, (от 2Гб)

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

[Помощь по синтезу моделей](#) [Помощь по верификации моделей](#)

Старт Cancel

3.5. Синтез и верификация моделей. Процесс завершен

Подготовительные шаги

1. Очистка предыдущих результатов верификации

2. Копирование обучающей выборки в распознаваемую

Стадия синтеза модели

№	Наименование модели	прогресс-бар	%
3	1. ABS - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ВСТРЕЧ СОЧЕТАНИЙ: "КЛАСС-ПРИЗНАК" У ОБЪЕКТОВ ОБУЧ		100%
4	2. PRC1 - ЧАСТНЫЙ КРИТЕРИЙ: УСЛ.ВЕРОЯТНОСТЬ I-ГО ПРИЗНАКА СРЕДИ ПРИЗНАКОВ ОБЪЕКТОВ J-ГО		100%
5	3. PRC2 - ЧАСТНЫЙ КРИТЕРИЙ: УСЛОВНАЯ ВЕРОЯТНОСТЬ I-ГО ПРИЗНАКА У ОБЪЕКТОВ J-ГО КЛАССА		100%
6	4. INF1 - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ЗНАНИЙ ПО А.ХАРКЕВИЧУ; ВЕРОЯТНОСТИ ИЗ PRC1		100%
7	5. INF2 - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ЗНАНИЙ ПО А.ХАРКЕВИЧУ; ВЕРОЯТНОСТИ ИЗ PRC2		100%
8	6. INF3 - ЧАСТНЫЙ КРИТЕРИЙ: ХИ-КВАДРАТ; РАЗНОСТИ МЕЖДУ ФАКТИЧЕСКИМИ И ОЖИДАЕМЫМИ АБС.ЧАСТОТАМИ		100%
9	7. INF4 - ЧАСТНЫЙ КРИТЕРИЙ: ROI (RETURN ON INVESTMENT); ВЕРОЯТНОСТИ ИЗ PRC1		100%
10	8. INF5 - ЧАСТНЫЙ КРИТЕРИЙ: ROI (RETURN ON INVESTMENT); ВЕРОЯТНОСТИ ИЗ PRC2		100%
11	9. INF6 - ЧАСТНЫЙ КРИТЕРИЙ: РАЗН.УСЛ.И БЕЗУСЛ.ВЕРОЯТНОСТЕЙ; ВЕРОЯТНОСТИ ИЗ PRC1		100%
12	10. INF7 - ЧАСТНЫЙ КРИТЕРИЙ: РАЗН.УСЛ.И БЕЗУСЛ.ВЕРОЯТНОСТЕЙ; ВЕРОЯТНОСТИ ИЗ PRC2		100%

Общий прогресс

100%

Начало: 24.02.2026 18:03:51

Прошло: 00:00:17

Окончание: 24.02.2026 18:04:09

Осталось: 00:00:00

Сообщения

```

[18:04:08 SUCCESS]: Агрегированные матрицы для модели 'INF7' сохранены.
[18:04:08 INFO]: Накопление статистики DostRasp для модели 'INF7'...
[18:04:08 INFO]: Расчет метрик VerModClIt для модели 'INF7' (пропущено).
[18:04:08 INFO]: Финальная обработка и сортировка таблиц dost_modelc...
[18:04:08 SUCCESS]: Таблица dost_modelc финализирована и отсортирована.
[18:04:09 INFO]: ЭТАП 8: Формирование итоговой БД DostRasp...
[18:04:09 INFO]: Формирование итоговой таблицы DostRasp...
[18:04:09 SUCCESS]: Таблица DostRasp успешно создана.

```

[Сохранить лог](#) [Выход](#)

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:

[illegible]

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

5.5 Просмотр основных баз данных всех моделей

Задайте модель для просмотра:

Abs

Prc1

Prc2

Inr1

Inr2

Inr3

Inr4

Inr5

Inr7

Количество индексов после запятой: 3 (макс.: 17)

Помощь

Модель: "PrcZ"

Описательные шкалы и градации (примечание)

Код градиента	1 - Winner - Alamo J.	2 - Winner - Agassi A.	3 - Winner - Agner R.	4 - Winner - Alami K.	5 - Winner - Allegro Y.	6 - Winner - Alonso J.	7 - Winner - Alvarez E.	8 - Winner - Andersen J.F.	9 - Winner - Araki H.	10 - Winner - Armando H.	11 - Winner - Arteta W.	12 - Winner - Backfield J.F.	13 - Winner - Balcells J.	14 - Winner - Baró G.	15 - Winner - Beland T.	16 - Winner - Benardjeur A.	17 - Winner - Spelman J.	18 - Winner - Black B.	19 - Winner - Black W.	20 - Winner - Blair J.	21 - Winner - Blanco G.	22 - Winner - Bonier J.	23 - Winner - Bouer J.	24 - Winner - Brausch K.	25 - Winner - Brunen P.	26 - Winner - Engelen S.
1	Date = /-5-[94685750000000000000000000000000, 956413440000000000000000000000]	19 697	66 667	40 000		25 000	20 000		27 660		3 448		12 000	23 810	44 444	25 000	11 628	28 125	40 000		4 348	9 091				3 704
2	Date = /-5-(95641344000000000000000000000000, 965962800000000000000000000000)]	13 636	16 667	28 000		50 000	80 000		14 894		13 793		4 000	9 524	33 333	75 000	25 581	31 250			21 739	4 545	100 000	100 000	100 000	29 630
3	Date = /-5-[96596280000000000000000000000000, 975525120000000000000000000000)]	19 697		16 000		25 000			23 404		27 586	100 000	28 000	38 095			32 558	21 875	40 000		21 739	27 273				25 926
4	Date = /-5-[97552512000000000000000000000000, 985089960000000000000000000000)]	71 429	31 818	16 667	4 000				10 638	100 000	24 138		32 000	28 571			16 279	6 250	10 000		21 739	40 909				14 815
5	Date = /-5-[98508996000000000000000000000000, 994636800000000000000000000000)]	28 571	15 152		12 000	100 000		100 000	23 404		31 034		24 000		22 222		13 953	12 500	10 000	100 000	30 435	18 182				25 926
6	Surface - Carpet	4 545							10 638			100 000		9 524			4 651					22 727				3 704
7	Surface - Clay	85 714	13 636	50 000	52 000	100 000	100 000	100 000	36 170	100 000	17 241		56 000	9 524	55 556	75 000	4 651	3 125		100 000	86 957	27 273			100 000	85 185
8	Surface - Grass	15 152		8 000	100 000				10 638		27 586						30 233	40 625	10 000			4 545	100 000	100 000		
9	Surface - Hard	14 286	66 667	50 000	40 000				42 553		55 172		44 000	80 952	44 444	25 000	60 465	56 250	90 000		13 043	45 455				11 111
10	Player_1 - Accasuo J.	14 286											4 000													
11	Player_1 - Agassi A.	46 970																								
12	Player_1 - Agener R.		50 000																							
13	Player_1 - Ai-Alani S.K.															25 000										
14	Player_1 - Alami K.			44 000					2 128																	
15	Player_1 - Allegro Y.				100 000																					
16	Player_1 - Alonso J.							2 128																		
17	Player_1 - Alvarez E.						60 000																			3 704
18	Player_1 - Andersen J.F.							100 000																		
19	Player_1 - Araki H.	14 286							51 064																	
20	Player_1 - Arquez P.																									

Экспорт в CSV

Экспорт всех моделей в MS Excel

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

[illegible]

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

5.5. Прогноз основных баз данных всех моделей (Не отвечать)

Задайте модель для просмотра:

Аба

Prc1

Prc2

Inr1

Inr2

Inr3

Inr4

Inr5

Inr7

Количество знаков после запятой: 3 (макс: 17)

Получить

Модель: "Inr3"

Описательные шкалы и градации (примечание)

Код признака

1 - Winner - Adriano J.

2 - Winner - Agassi A.

3 - Winner - Agner R.

4 - Winner - Alan K.

5 - Winner - Allego Y.

6 - Winner - Alonso J.

7 - Winner - Alvarez E.

8 - Winner - Andersen J.F.

9 - Winner - Andre H.

10 - Winner - Armando M.

11 - Winner - Arthur W.

12 - Winner - Baines J.P.

13 - Winner - Batista J.

14 - Winner - Best G.

15 - Winner - Bernard T.

16 - Winner - Bernasconi A.

17 - Winner - Brugnara J.

18 - Winner - Buck B.

19 - Winner - Buck W.

20 - Winner - Buar J.

21 - Winner - Banca G.

22 - Winner - Bauer J.

23 - Winner - Bauer J.

24 - Winner - Brauch N.

25 - Winner - Bremer F.

26 - Winner - Burgard S.

27 - Winner - Byun D.

1	Date - 1/5-[945857600000000000 000000, 956413440000000000 000000]	-1.472	-0.877	2.738	4.744	-0.210	0.159	-0.051	-0.210	3.118	-0.210	-0.097	-0.210	-2.256	0.585	2.108	0.159	-0.041	2.272	1.897	-0.210	-0.836	-0.626	-0.210	-0.421	-0.841	-0.677	-1.472
2	Date - 2/5-[956413440000000000 000000, 965962800000000000 000000]	-1.669	-4.733	-0.430	1.040	-0.238	1.046	2.808	-0.238	4.204	-0.238	-2.913	-0.238	-4.960	-3.006	0.855	2.046	0.749	2.372	-2.384	-0.238	-0.483	-4.244	0.762	1.523	3.046	1.564	2.331
3	Date - 3/5-[965962800000000000 000000, 975251200000000000 000000]	-1.247	1.239	-1.069	-0.455	-0.178	0.287	-0.891	-0.178	2.625	-0.178	2.832	0.822	2.545	4.258	-1.604	-0.713	6.338	1.298	2.218	-0.178	0.902	2.080	-0.178	-0.356	-0.713	2.189	-1.247
4	Date - 4/5-[975251200000000000 000000, 985089960000000000 000000]	3.719	8.923	-0.098	-3.575	-0.183	-0.732	-0.915	-0.183	-3.600	0.817	1.693	-0.183	3.425	2.157	-1.647	-0.732	-0.868	-3.856	-0.830	-0.183	0.791	4.974	-0.183	-0.366	-0.732	-0.941	0.719
5	Date - 5/5-[985089960000000000 000000, 994636800000000000 000000]	0.669	-2.552	-1.141	-1.755	0.810	-0.761	-0.951	0.810	2.061	-0.190	3.485	-0.190	1.245	-3.994	0.288	-0.761	-2.178	-2.086	-0.902	0.810	2.626	-0.184	-0.190	-0.380	-0.761	1.865	-0.331
6	Surface - Carpet	-0.249	0.654	-0.213	-0.889	-0.036	-0.142	-0.178	-0.036	3.329	-0.036	-1.031	0.964	-0.889	1.253	-0.320	-0.142	0.471	-1.138	-0.356	-0.036	-0.818	4.218	-0.036	-0.071	-0.142	0.040	-0.249
7	Surface - Clay	3.446	-15.082	0.811	3.878	-0.365	2.540	3.176	0.635	-0.050	-0.635	-5.582	-0.365	4.878	-6.663	1.716	1.540	-13.690	-10.676	-3.649	0.635	11.608	-2.027	-0.365	-0.730	2.540	13.148	-2.564
8	Surface - Grass	-0.861	1.881	-0.738	-1.075	0.877	-0.452	-0.615	-0.123	-0.781	0.123	-0.433	-0.123	-3.075	-5.683	-1.107	-0.492	7.711	9.064	-0.230	-0.123	-8.829	-1.706	0.877	1.754	-0.492	-3.321	4.139
9	Surface - Hard	-2.336	12.547	0.141	-1.914	-0.477	-1.906	-2.383	-0.477	-2.398	-0.477	2.180	-0.477	-9.914	6.992	-0.289	-0.906	5.508	2.750	4.234	-0.477	-7.961	-0.484	-0.477	-9.953	-1.906	-9.867	-1.336
10	Player_1 - Acasuso J.	0.992	-0.072	-0.027	-0.001	-0.004	-0.005	-0.001	-0.051	-0.001	-0.032	-0.001	0.973	-0.023	-0.010	-0.004	-0.047	-0.035	-0.011	-0.001	-0.025	-0.024	-0.001	-0.002	-0.004	-0.029	-0.008	
11	Player_1 - Agassi A.	-0.067	0.367	-0.058	-0.240	0.010	-0.038	-0.048	-0.010	-0.451	-0.010	-0.278	-0.010	-0.240	-0.202	-0.086	-0.338	-0.413	-0.307	-0.095	-0.010	-0.221	-0.010	-0.019	-0.038	-0.259	-0.067	
12	Player_1 - Agner R.	-0.018	-0.173	2.984	-0.065	-0.003	-0.010	-0.013	-0.003	-0.123	-0.003	-0.076	-0.003	-0.065	-0.055	-0.024	-0.010	-0.113	-0.084	-0.026	-0.003	-0.060	-0.058	-0.003	-0.005	-0.010	-0.071	-0.010
13	Player_1 - Al-Atawi S.K.	-0.003	-0.029	-0.003	-0.011	-0.000	-0.002	-0.002	-0.000	-0.621	-0.000	-0.013	-0.000	-0.011	-0.009	-0.004	0.998	-0.019	-0.014	-0.004	-0.000	-0.010	-0.010	-0.000	-0.001	-0.002	-0.012	-0.003
14	Player_1 - Alami K.	-0.050	-0.475	-0.043	10.820	-0.007	-0.029	-0.036	-0.007	0.026	-0.007	-0.209	-0.007	-0.180	-0.151	-0.065	-0.029	-0.309	-0.230	-0.072	-0.007	-0.166	-0.158	-0.007	-0.014	-0.029	-0.194	-0.050
15	Player_1 - Allego Y.	-0.002	-0.014	-0.001	-0.005	1.000	-0.001	-0.001	-0.000	-0.010	-0.000	-0.005	-0.000	-0.005	-0.005	-0.002	-0.001	-0.009	-0.007	-0.002	-0.000	-0.005	-0.005	-0.000	-0.000	-0.001	-0.005	-0.002
16	Player_1 - Alonso J.	-0.003	-0.029	-0.003	-0.011	-0.000	-0.002	-0.002	-0.000	0.979	-0.000	-0.013	-0.000	-0.011	-0.009	-0.004	-0.002	-0.019	-0.014	-0.004	-0.000	-0.010	-0.010	-0.000	-0.001	-0.002	-0.012	-0.003
17	Player_1 - Alvarez E.	-0.011	-0.101	-0.009	-0.038	-0.002	-0.006	2.992	-0.002	-0.072	-0.002	-0.044	-0.002	-0.038	-0.032	-0.014	-0.006	-0.066	-0.049	-0.015	-0.002	-0.035	-0.034	-0.002	-0.003	-0.006	0.959	-0.011
18	Player_1 - Andersen J.F.	-0.003	-0.029	-0.003	-0.011	-0.000	-0.002	-0.002	1.000	-0.001	-0.009	-0.003	-0.000	-0.011	-0.009	-0.004	-0.002	-0.019	-0.014	-0.004	-0.000	-0.010	-0.010	-0.000	-0.001	-0.002	-0.012	-0.003
19	Player_1 - Arazi H.	0.936	-0.005	-0.055	-0.229	-0.009	-0.037	-0.046	-0.009	23.659	-0.009	-0.265	-0.009	-0.219	-0.192	-0.082	-0.037	-0.394	-0.293	-0.002	-0.009	-0.211	-0.202	-0.009	-0.018	-0.037	-0.247	-0.064
20	Player_1 - Arquez P.	-0.002	-0.014	-0.001	-0.005	-0.000	-0.001	-0.001	-0.000	-0.010	-0.000	-0.006	-0.000	-0.005	-0.005	-0.002	-0.001	-0.009	-0.007	-0.002	-0.000	-0.005	-0.005	-0.000	-0.000	-0.001	-0.006	-0.002

Экспорт в CSV

Экспорт всех моделей в MS Excel

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 критерием достоверности L1-мерой проф. Е.В.Луценко наиболее достоверными являются СК-модели INF4 и INF5 с интегральным критерием: «Сумма знаний»: L1=0.798 при максимуме 1 (рисунок 12). Модели INF4 и INF5 имеют одинаковые значения достоверности во всех трех мерах. INF4 будем использовать для решения поставленных в работе задач.

НУЖЕН СКРИН

3.4. Обобщенная форма доств.моделей при разн.крит. Текущая модель: "INF1"												
Наименование модели и частного критерия	Интегральный критерий	ложно-отрицательный (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	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 сомнений: "Клас...	Корреляция абс.частот с обр...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
1. ABS - частный критерий: количество встреч сомнений: "Клас...	Сумма абс.частот по признакам...		0.333	1.000	0.500	11.778		16.667	0.414	1.000	0.586	
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн.частот с о...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Сумма усл.отн.частот по при...		0.333	1.000	0.500	12.292		17.708	0.410	1.000	0.581	
3. PRC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн.частот с о...	4	0.750	0.750	0.750	7.161	9.159	1.998	0.782	1.000	0.878	
3. PRC2 - частный критерий: условная вероятность его признака...	Сумма усл.отн.частот по при...		0.333	1.000	0.500	12.292		17.708	0.410	1.000	0.581	
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	4	0.462	0.750	0.571	7.102	9.102	2.000	0.780	1.000	0.877	
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	6	0.714	0.625	0.667	2.237	9.828	0.353	0.096	0.864	0.959	0.909
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.102	9.102	2.000	0.000	0.780	1.000	0.877
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	6	0.714	0.625	0.667	2.237	9.828	0.353	0.096	0.864	0.959	0.909
6. INF3 - частный критерий: Хинкварт, разности между фактич...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
6. INF3 - частный критерий: Хинкварт, разности между фактич...	Сумма знаний	4	0.750	0.750	0.750	6.667	7.333	0.667	0.909	1.000	0.952	
7. INF4 - частный критерий: ROI (Return On Investment); веро...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
7. INF4 - частный критерий: ROI (Return On Investment); веро...	Сумма знаний	4	0.750	0.750	0.750	6.111	6.778	0.667	0.902	1.000	0.948	
8. INF5 - частный критерий: ROI (Return On Investment); веро...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
8. INF5 - частный критерий: ROI (Return On Investment); веро...	Сумма знаний	4	0.750	0.750	0.750	6.111	6.778	0.667	0.902	1.000	0.948	
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Сумма знаний	4	0.750	0.750	0.750	6.111	6.778	0.667	0.902	1.000	0.948	
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Семантический резонанс зна...	4	0.750	0.750	0.750	7.162	9.162	2.000	0.782	1.000	0.877	
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Сумма знаний	4	0.750	0.750	0.750	6.111	6.778	0.667	0.902	1.000	0.948	
Помощь по мерам достоверности Помощь по частотным распределениям TRTN,FBFN (TF-FP)/(TN-FN) (T-F)/(T+F)*100 Задать интервал сглаживания												

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 INF4.

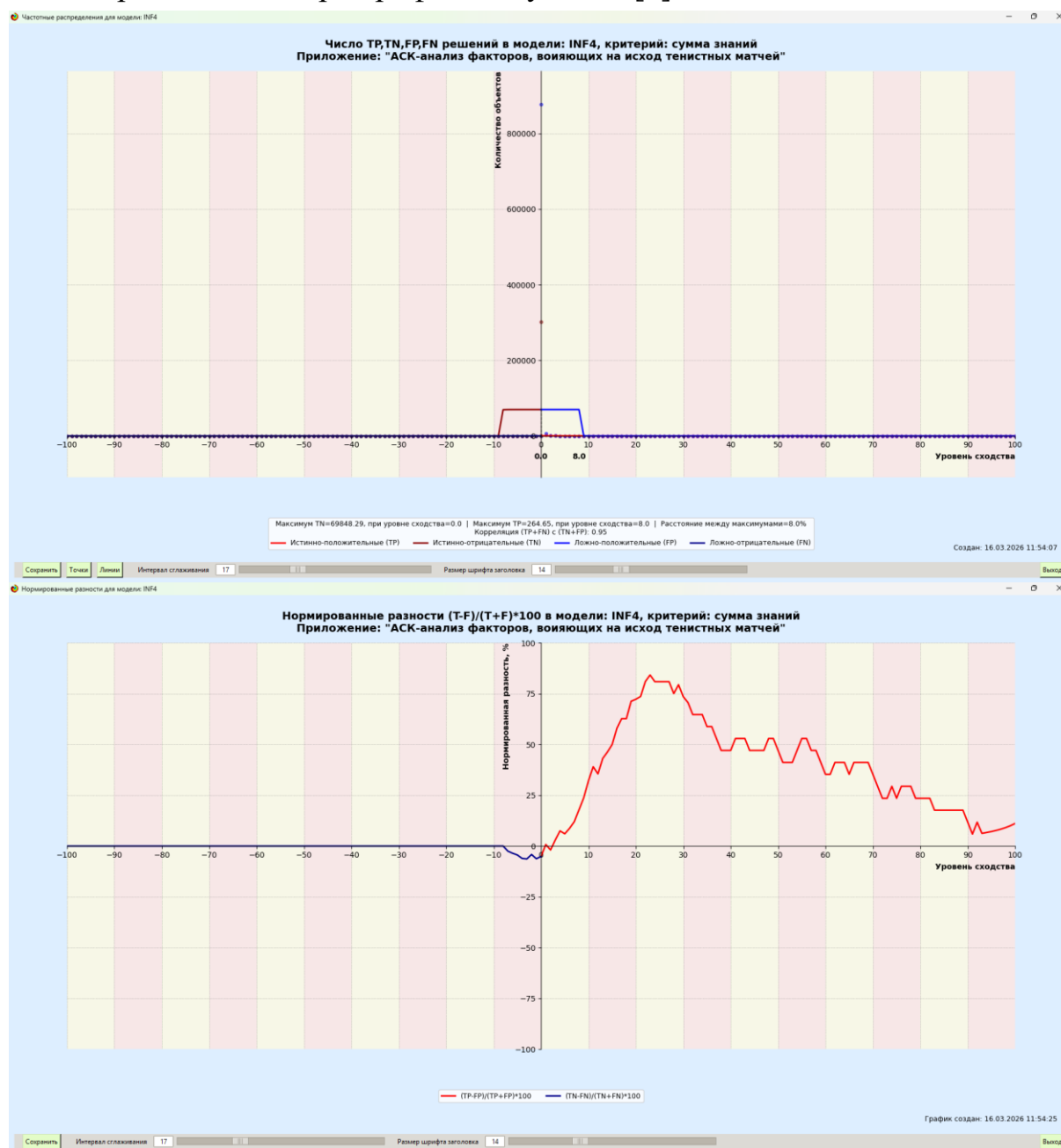


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4 [9]

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

— максимум истинно-положительных решений (TP=264.65) достигается при уровне сходства 8.0%, а максимум истинно-отрицательных решений (TN=69648.29) — при уровне сходства 0.0%.

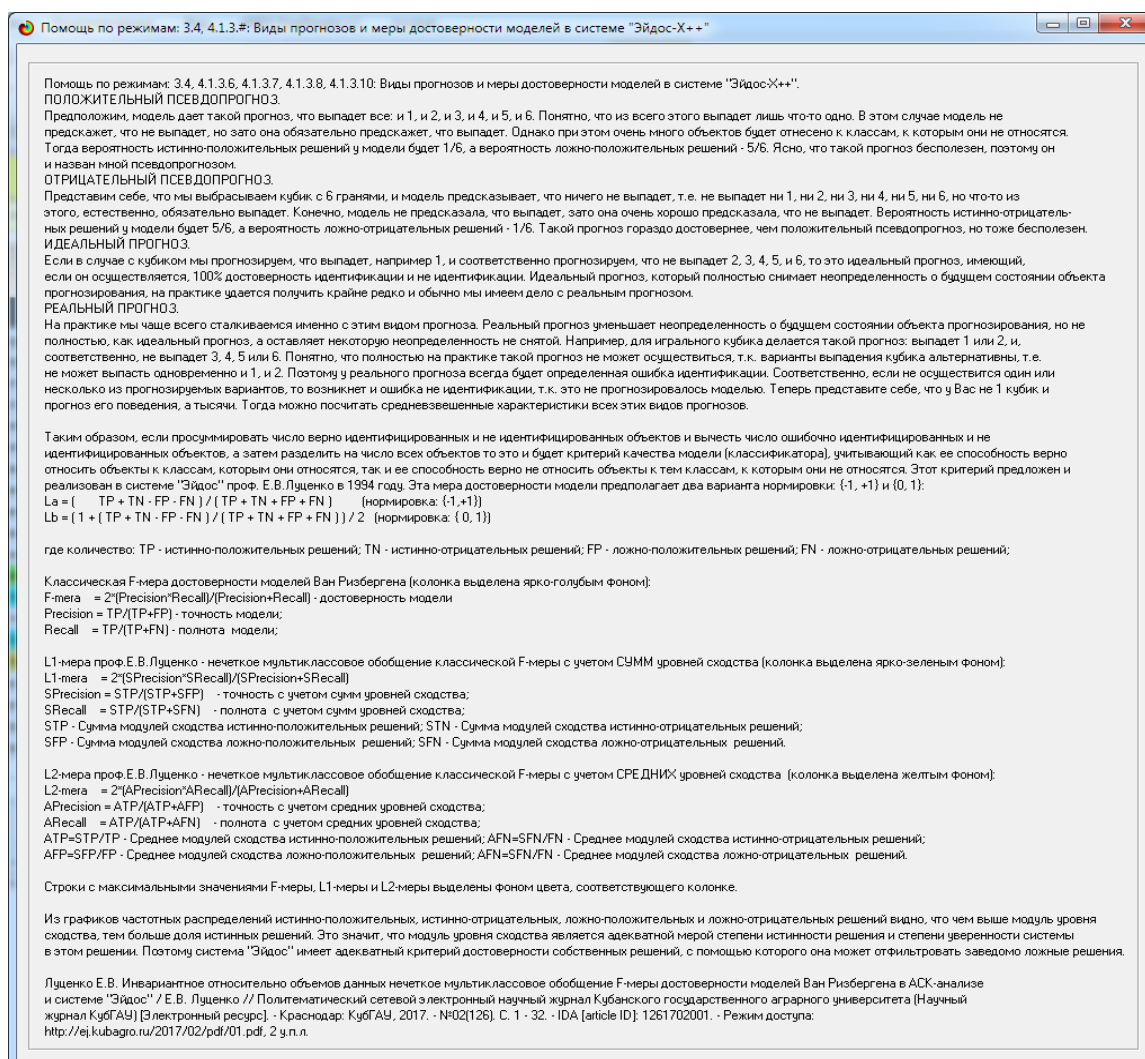
Расстояние между максимумами составляет 8.0%, что свидетельствует о четкой дифференциации решений по уровням сходства;

– при уровнях сходства меньше 10% в данной задаче преобладают ложные положительные решения, а при более высоких уровнях сходства (выше 15-20%) – истинные положительные решения. На графике нормированных разностей видно, что при уровнях сходства выше 25-30% доля истинных решений значительно превышает долю ложных решений (нормированная разность достигает 75-80%);

– корреляция между $(TP+FN)$ и $(TN+FP)$ составляет 0.95, что указывает на высокую согласованность распределений решений;

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

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



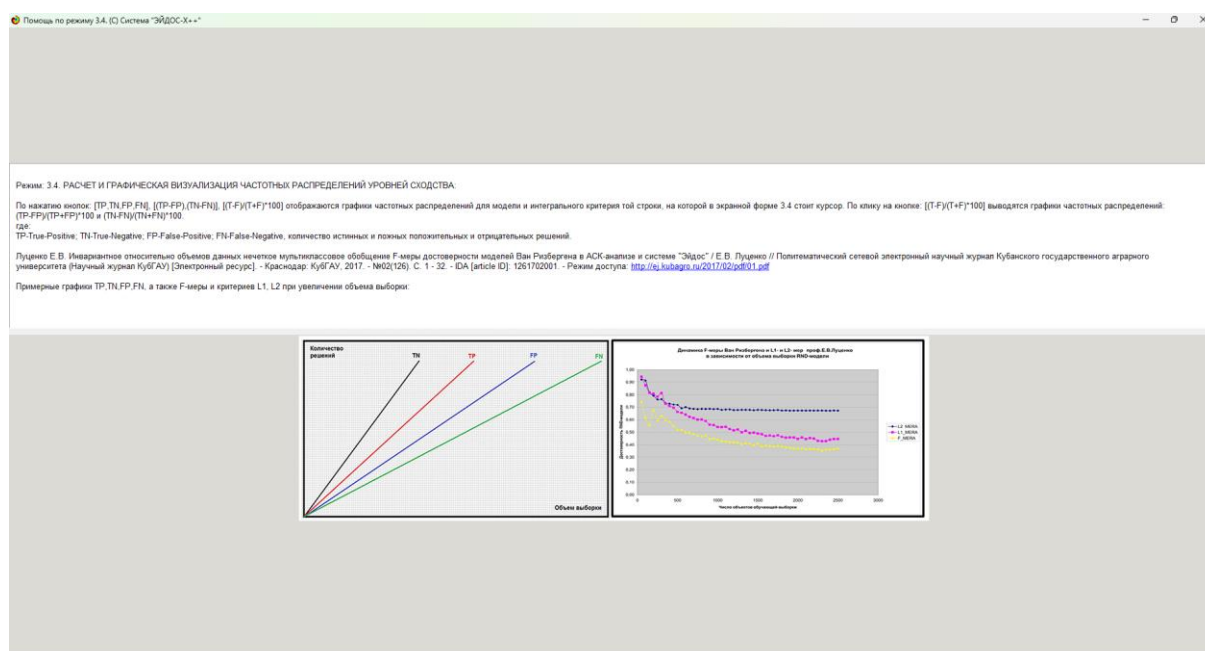


Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

3.5.1. Универсальная постановка задачи, не зависящая от предметной области

Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

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

3.5.2. Конкретное решение задачи в данной работе

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

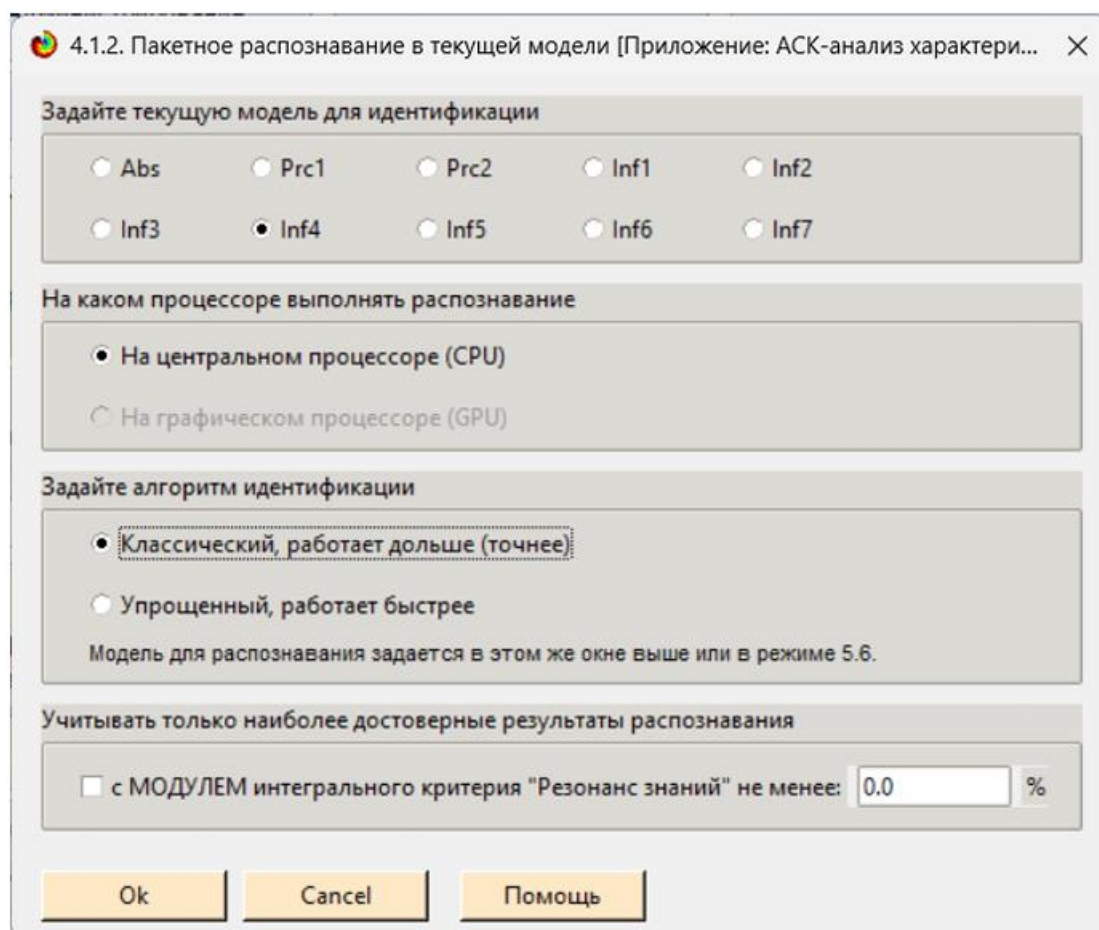


Рисунок 15. Задание СК-модели INF4 в качестве текущей

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 класса 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 ,

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

т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

3.6.1.2. Интегральный критерий «Семантический резонанс знаний»

Интегральный критерий «Семантический резонанс знаний»

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

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

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

где:

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

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

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

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

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

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

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

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

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

3.6.1.3. Важные математические свойства интегральных критериев

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

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

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

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

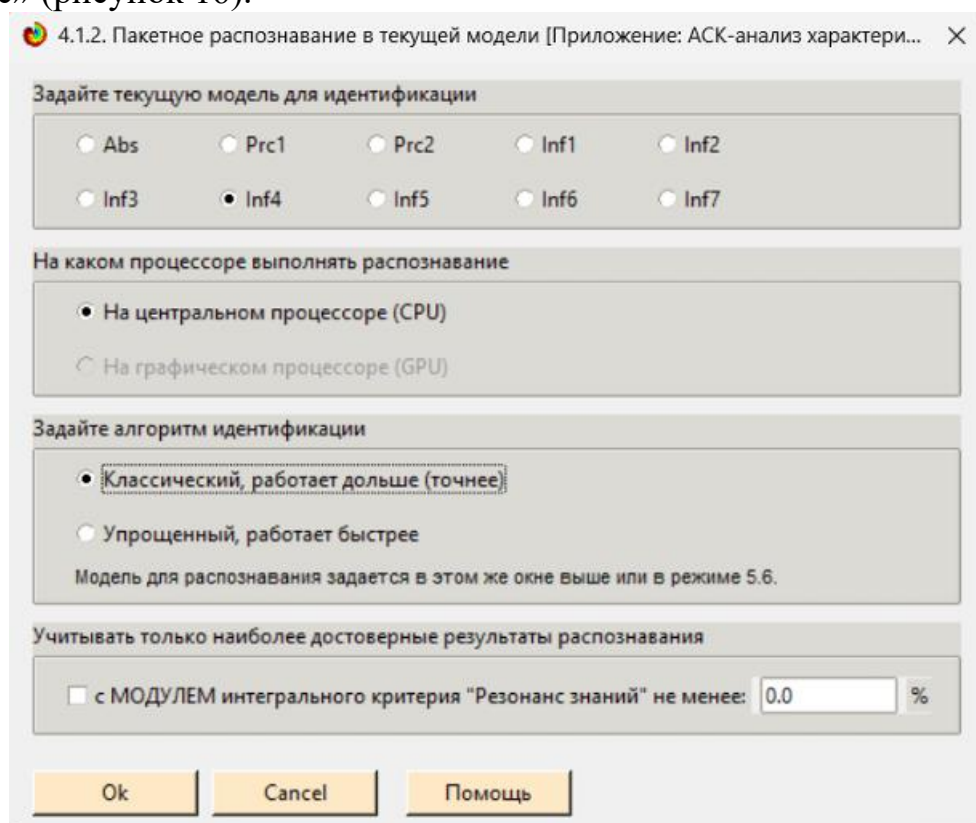
В-четвертых, кроме того значение интегрального критерия сходства представляет собой адекватную самооценку **степени уверенности** системы в положительном или отрицательном решении о принадлежности/непринадлежности объекта к классу или **риска ошибки** при таком решении.

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

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

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

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 освещены 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).



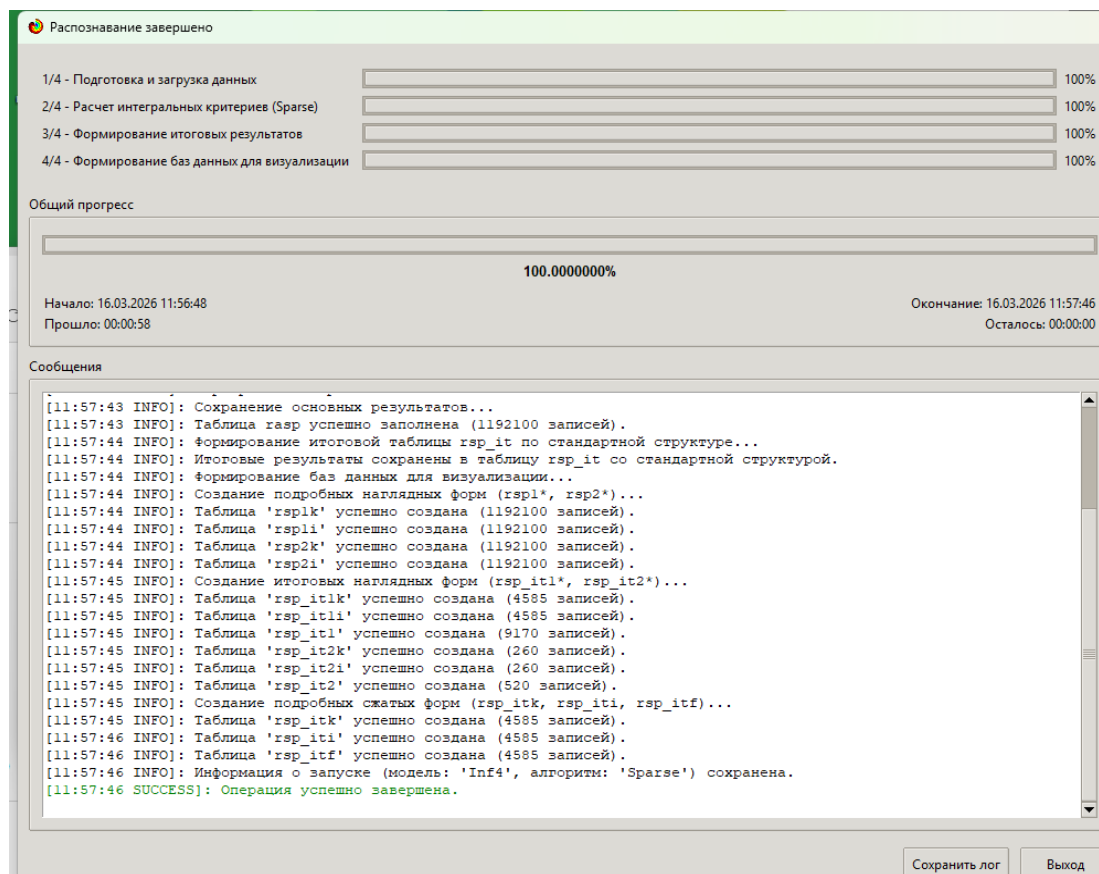


Рисунок 16. Экранные формы режима 4.1.2 идентификации и прогнозирования

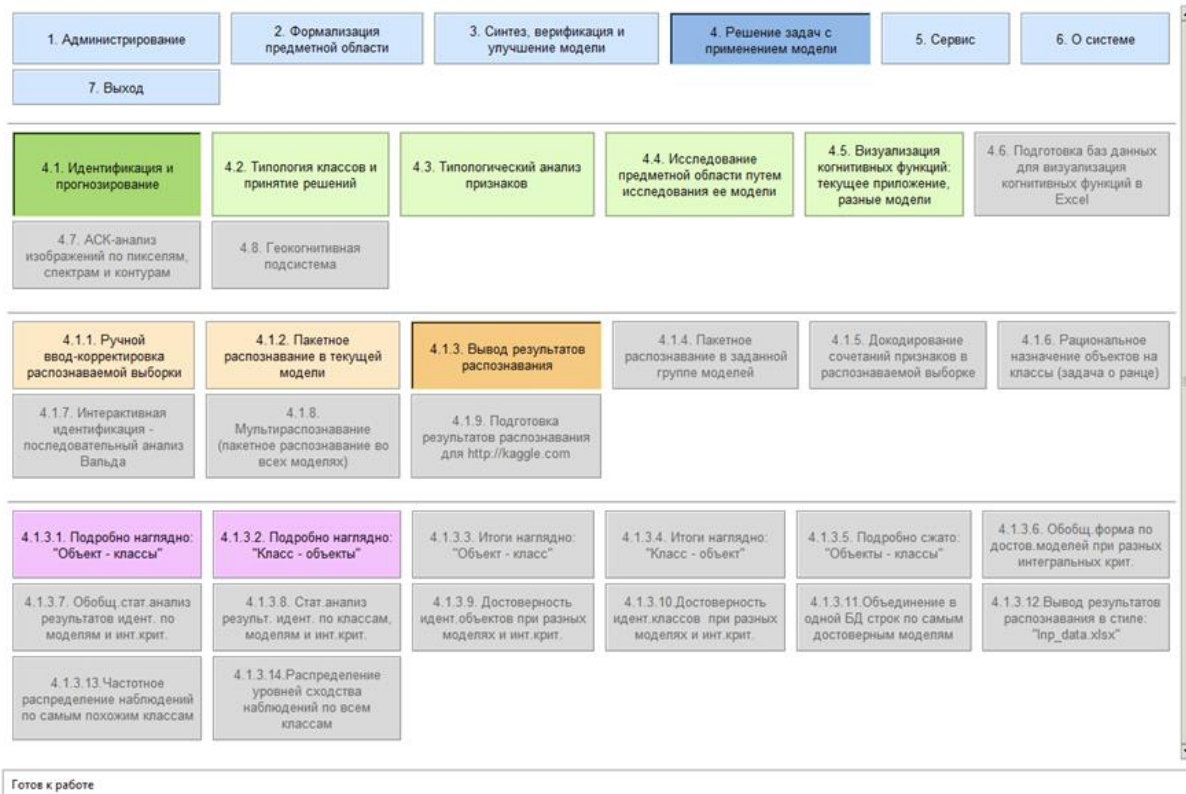


Рисунок 17. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

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

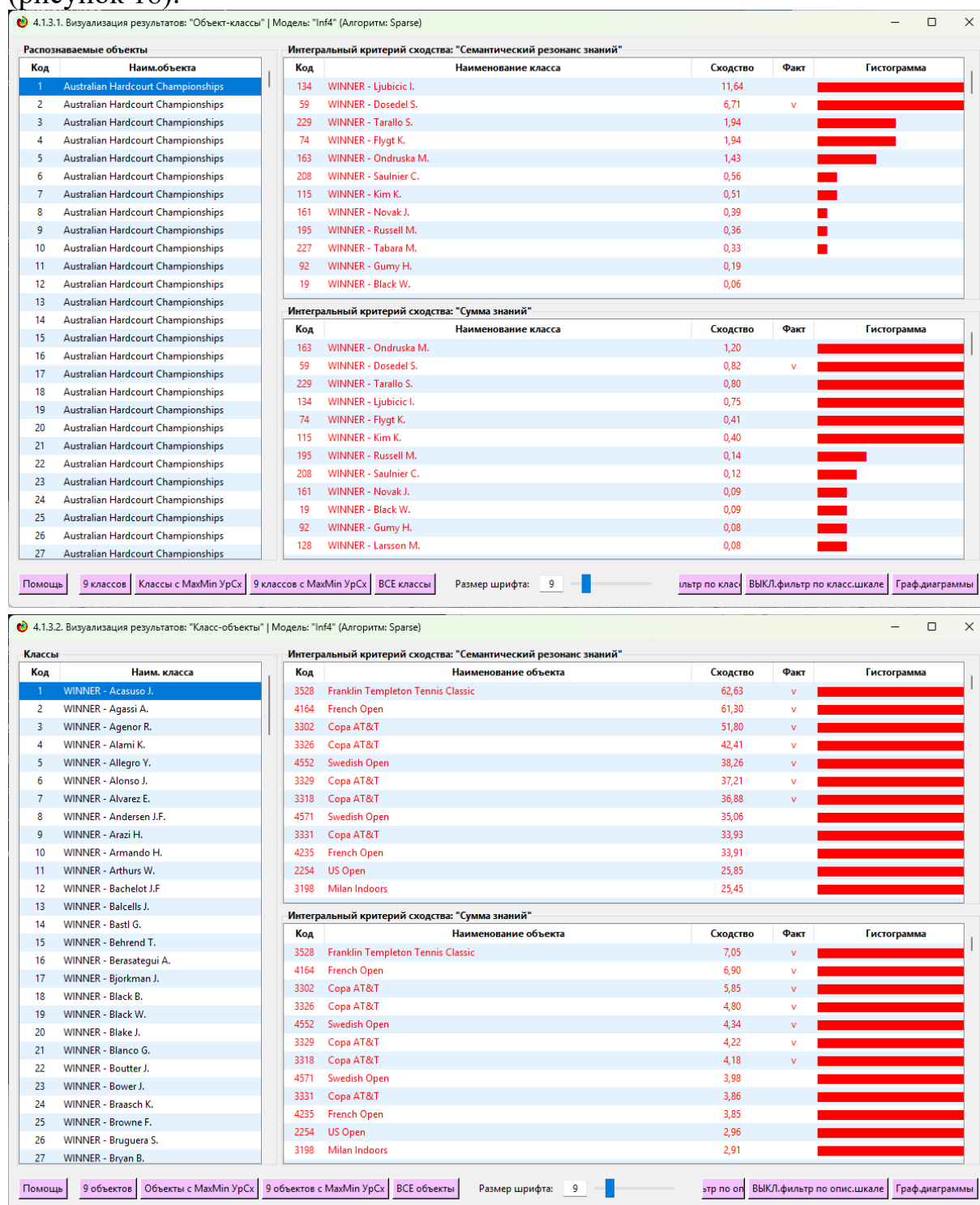


Рисунок 18. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

3.7. Задача-7. Поддержка принятия решений

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

3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

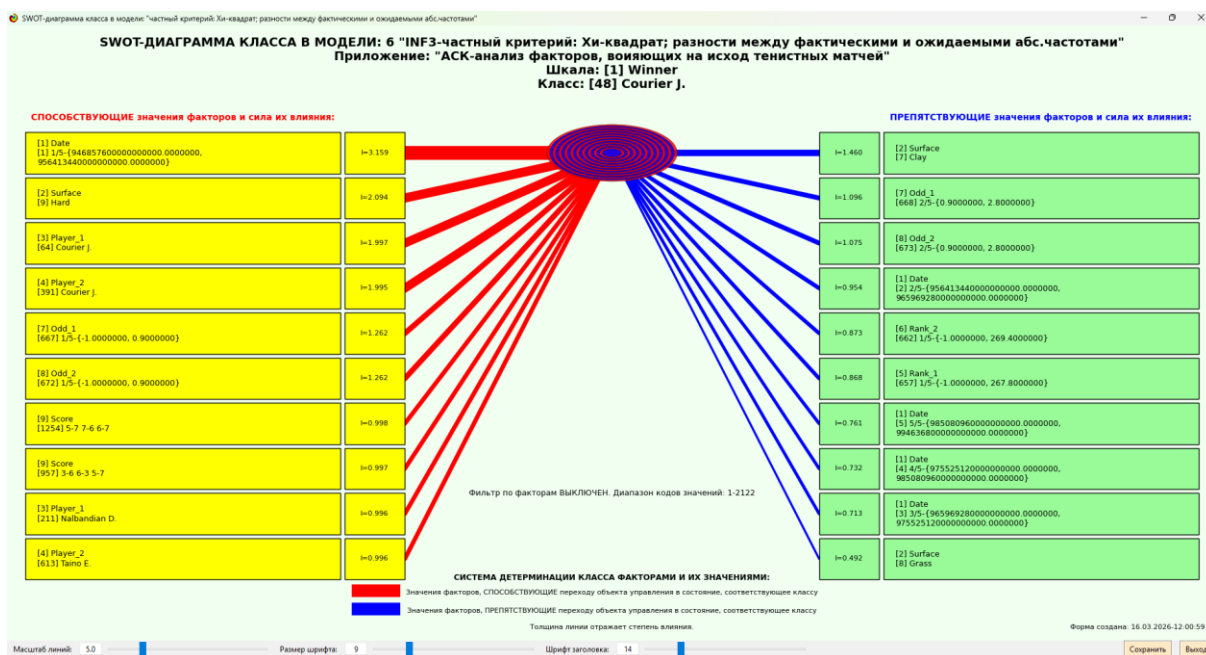
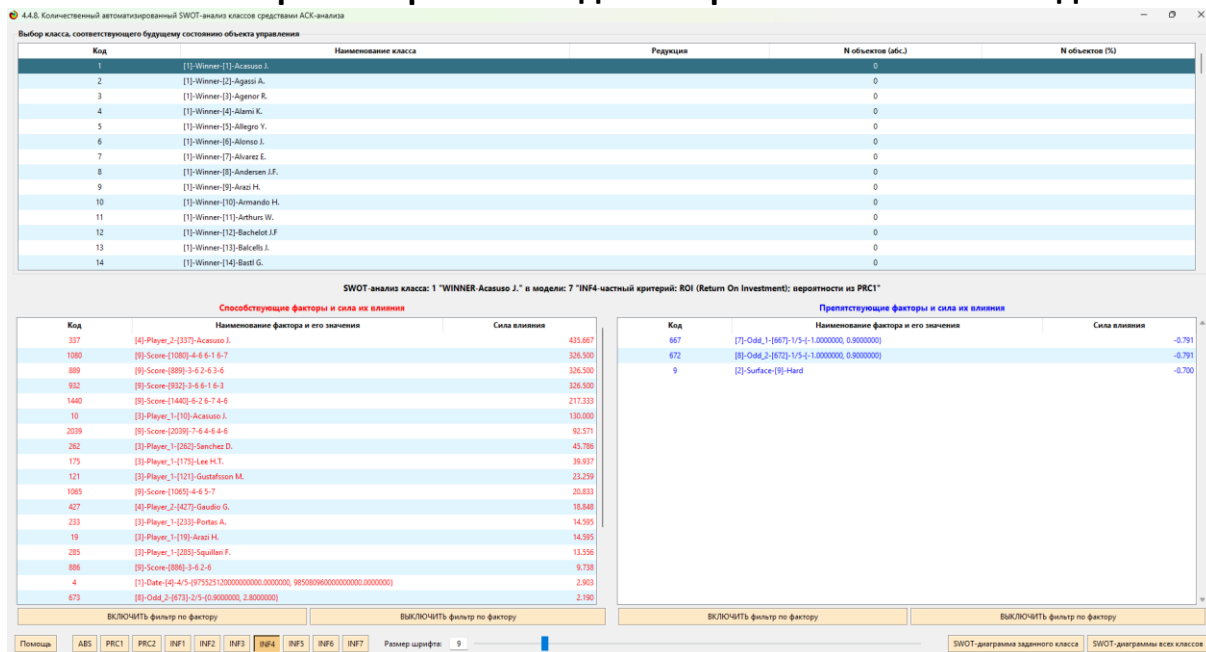
Графические выходные формы, приведенные на рисунках 19, интуитивно понятны и не требуют особых комментариев. Отметим лишь, что на SWOT-диаграммах наглядно показаны знак и сила влияния каждого значения фактора на переход объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне.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

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

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

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

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



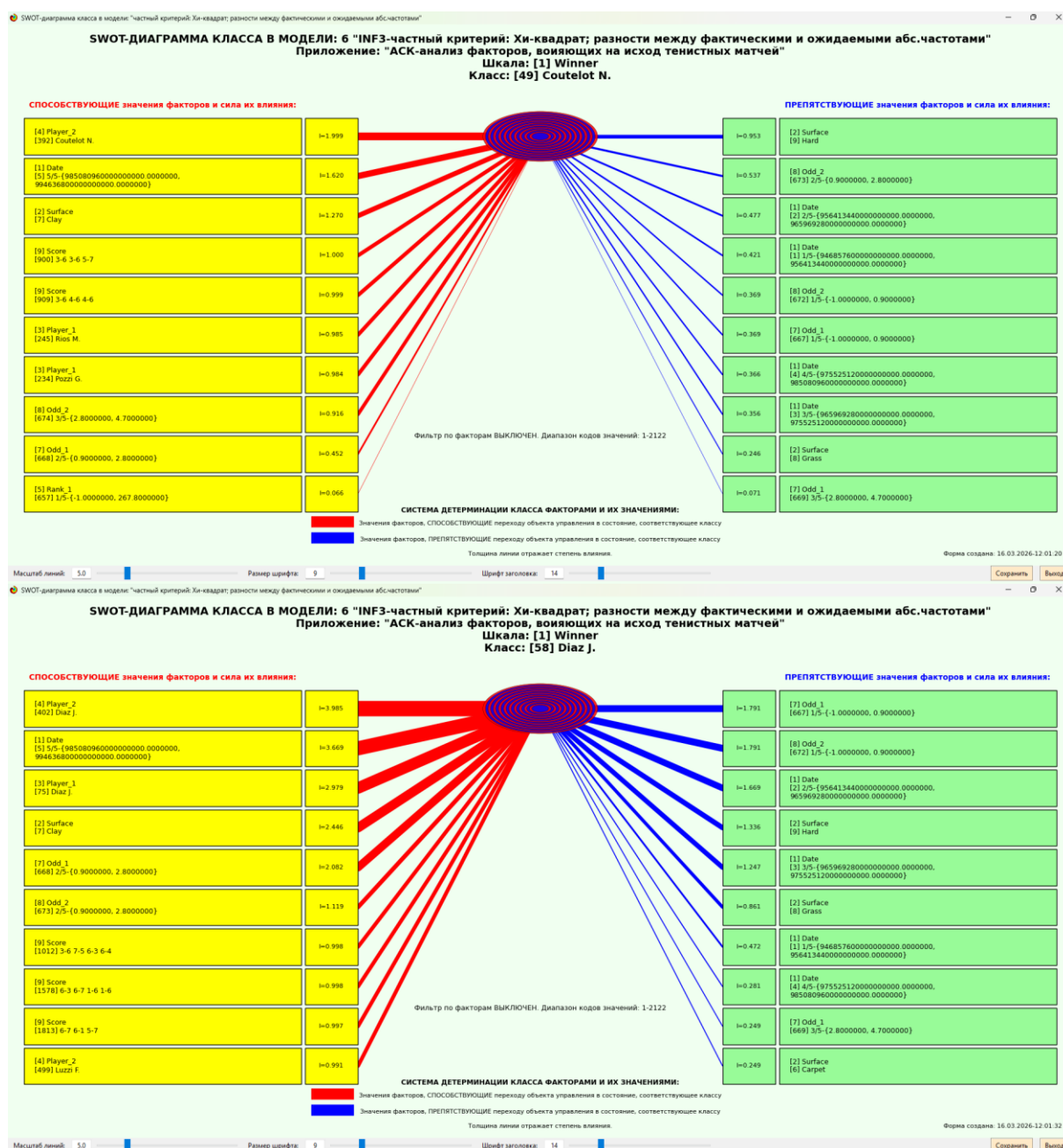


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

Из приведенных выходных форм SWOT-анализа четко видно, какие факторы влияют на победу конкретных теннисистов. Например, для победы Courier J. характерны матчи на покрытии Hard с определенными коэффициентами, тогда как покрытие Clay и Grass препятствуют его победе. Для Courtelat N. благоприятным фактором является покрытие Clay, а препятствующим – покрытие Hard. Для Diaz J. также характерны победы на покрытии Clay, в то время как покрытие Hard, Grass и Carpet препятствуют его победе. Таким образом, тип покрытия корта является одним из наиболее значимых факторов, влияющих на исход матча для различных игроков.

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

3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

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

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

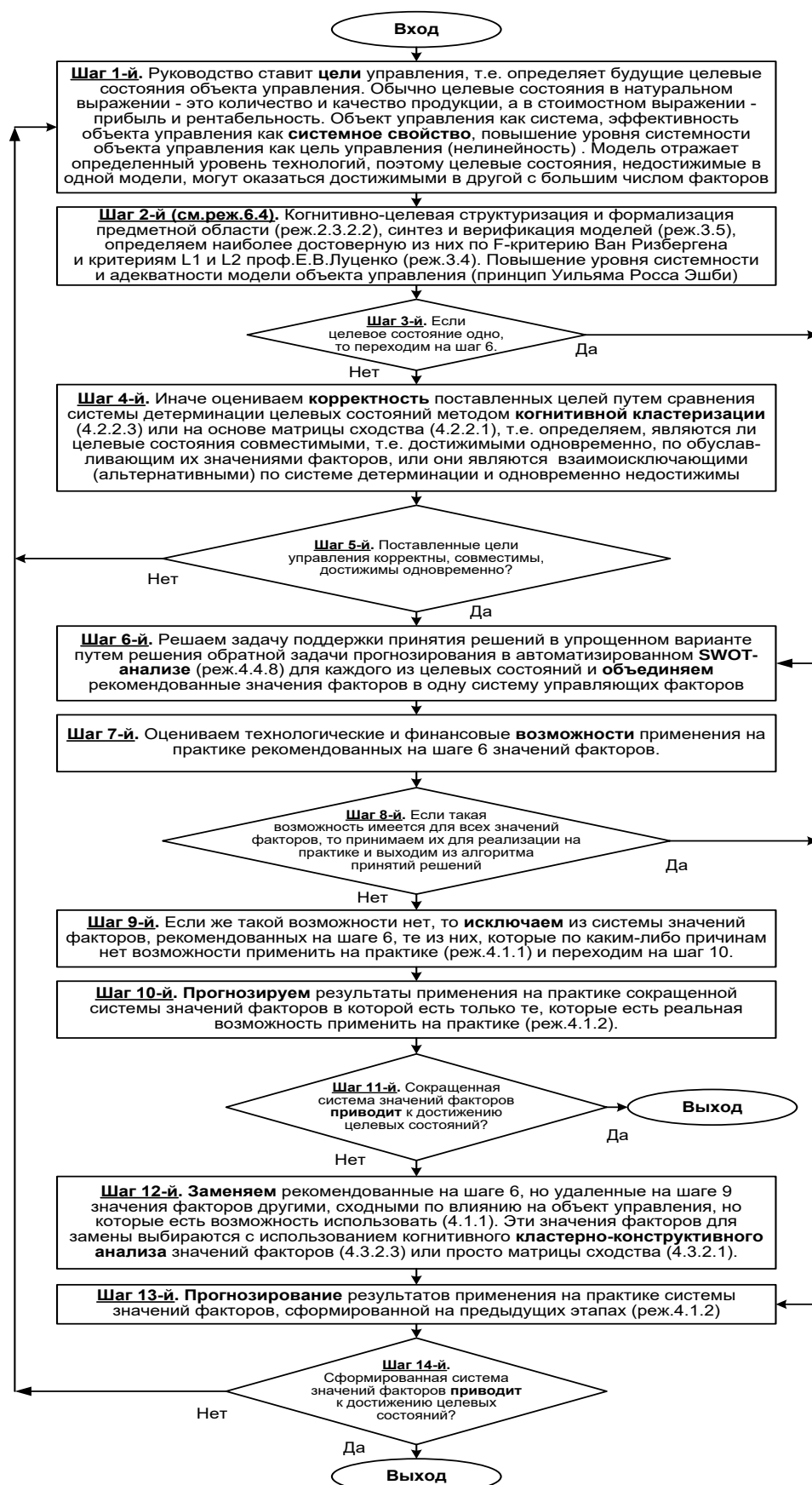


Рисунок 20. Разработанный алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

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

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

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

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

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

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

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

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

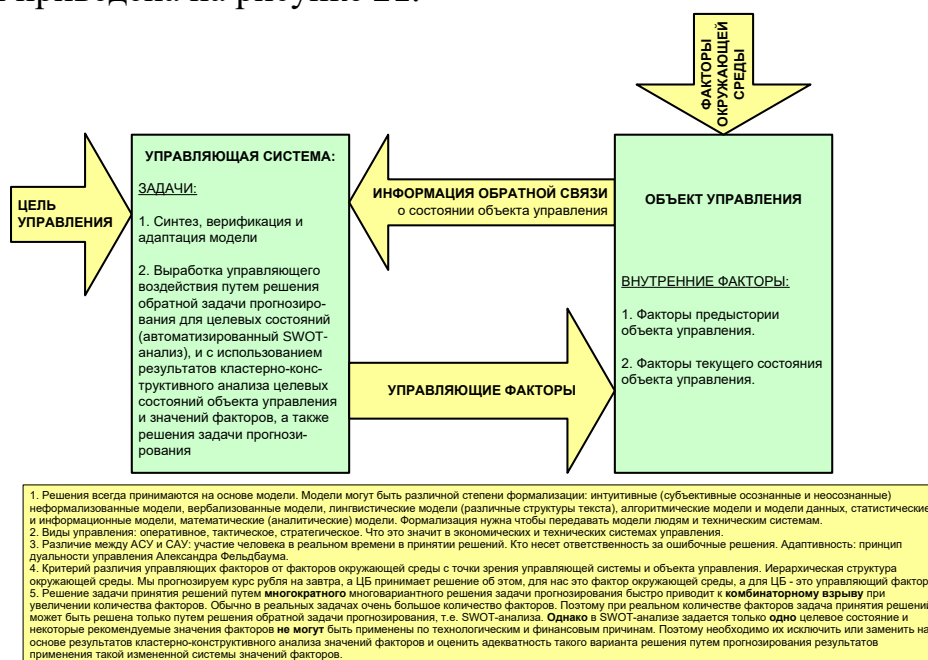


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

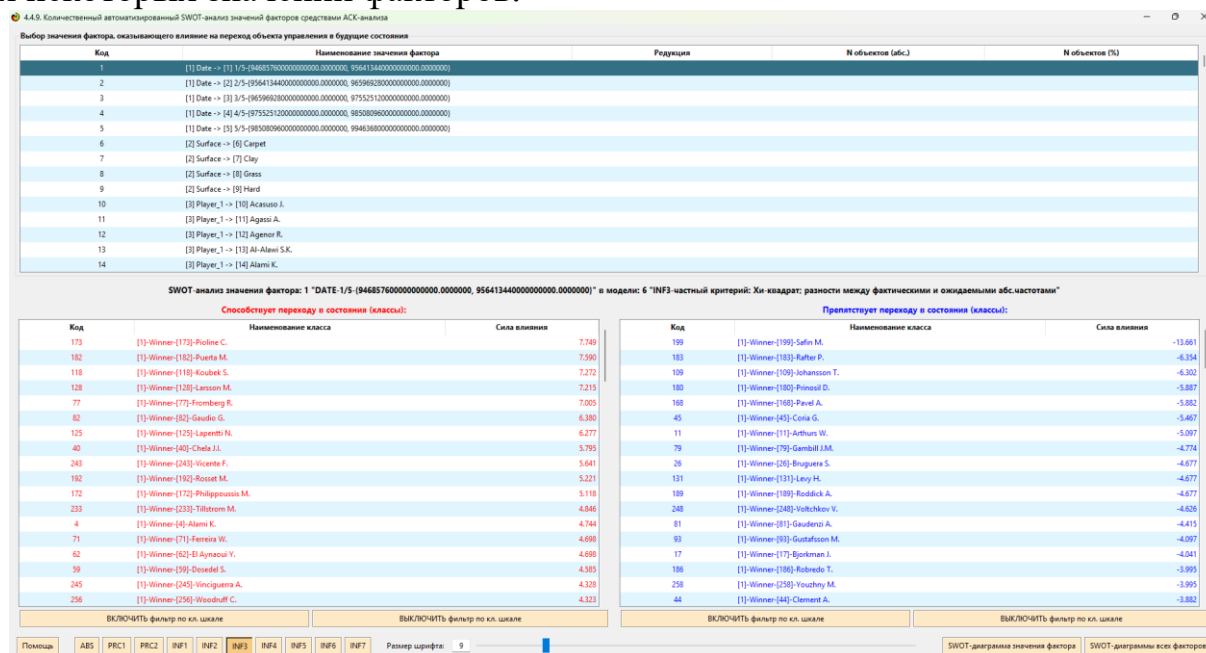
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)

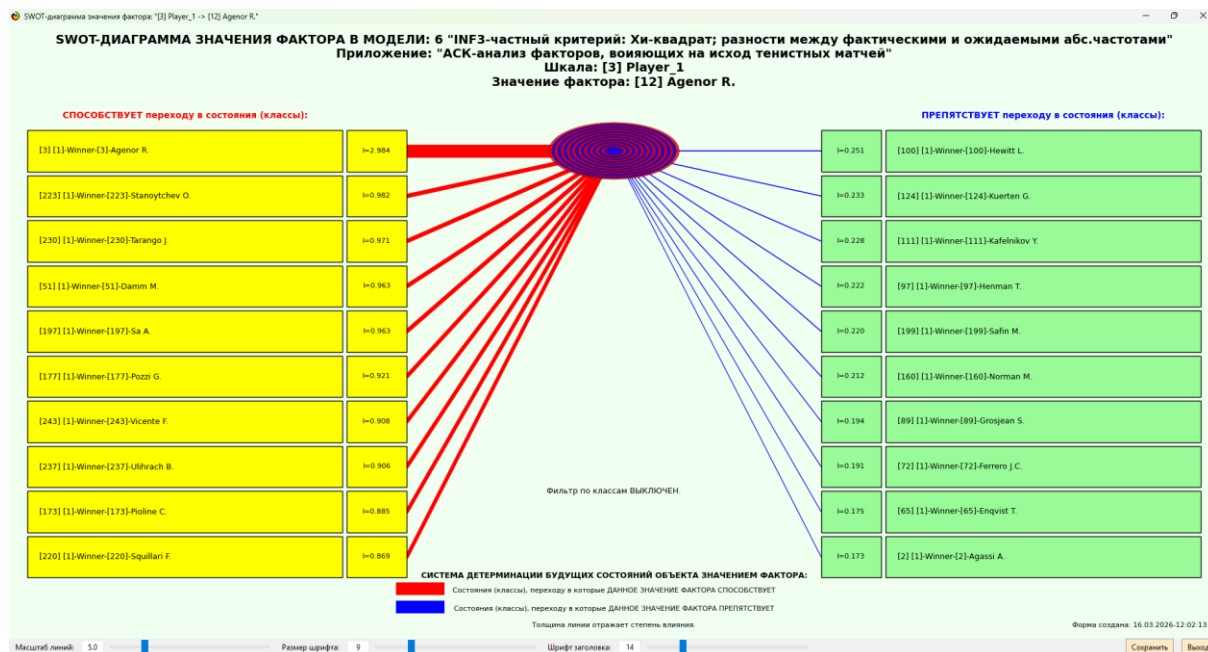
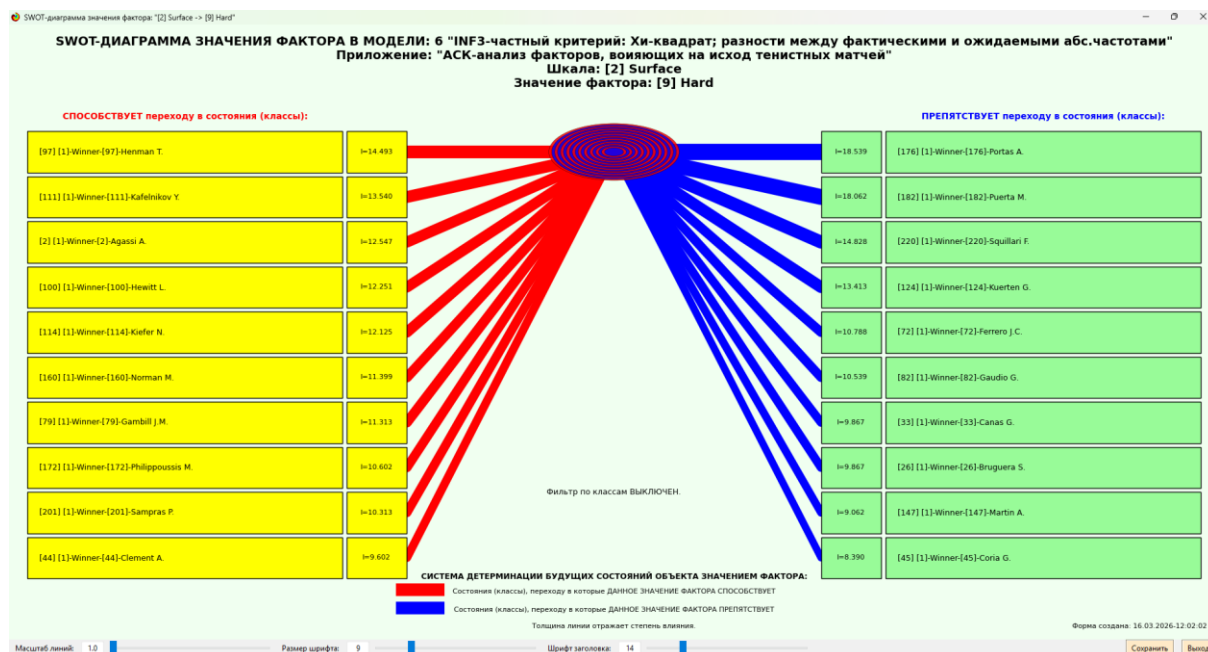
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области

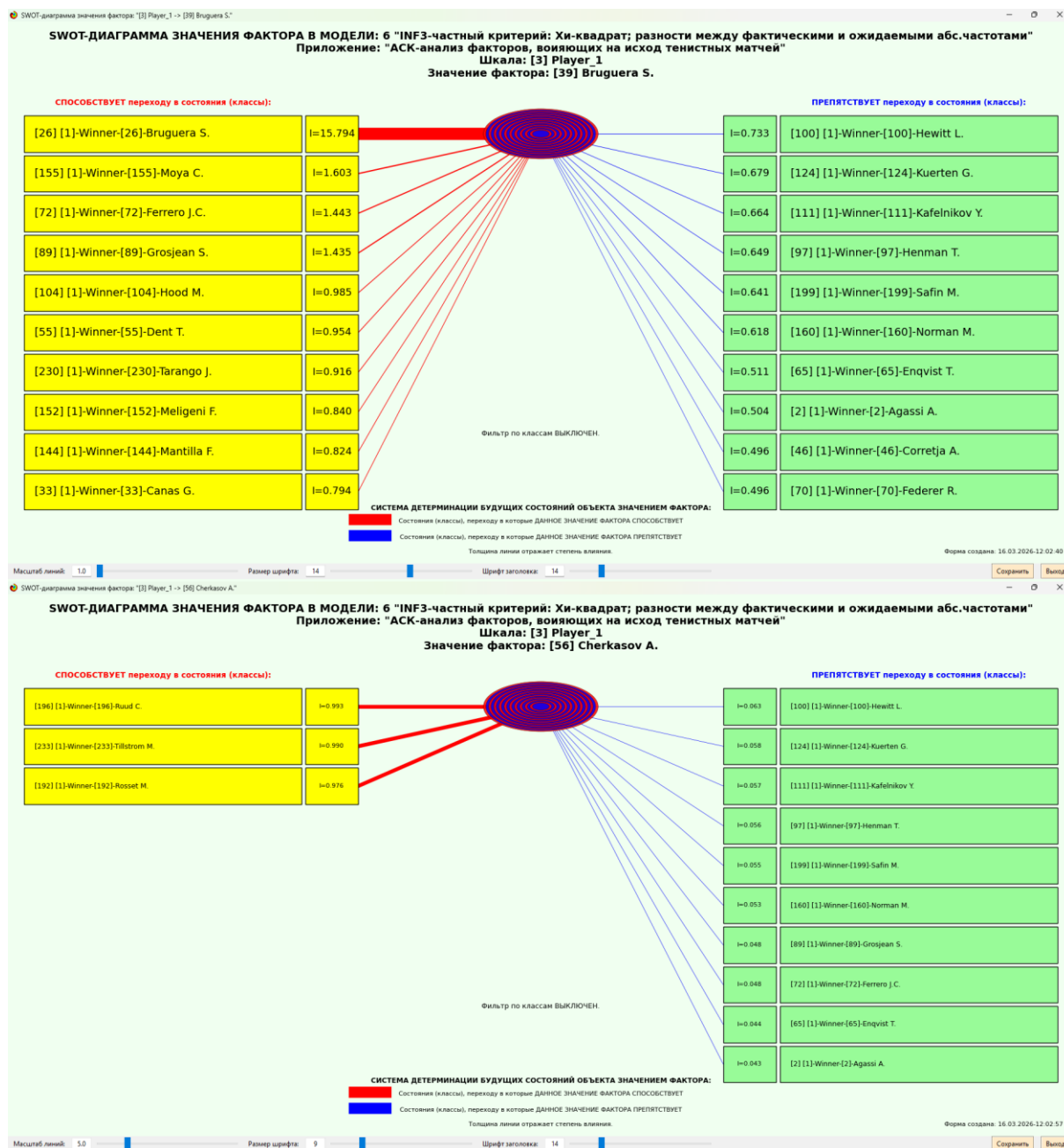
Инвертированные SWOT-диаграммы (предложены автором в работе [10]), отражают силу и направление влияния конкретной градации описательной шкалы на переход объекта моделирования в состояния, соответствующие градациям классификационных шкал (классы). Это и есть **смысл** (семантический потенциал) этой градации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

3.8.1.2. Конкретное решение задачи в данной работе

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







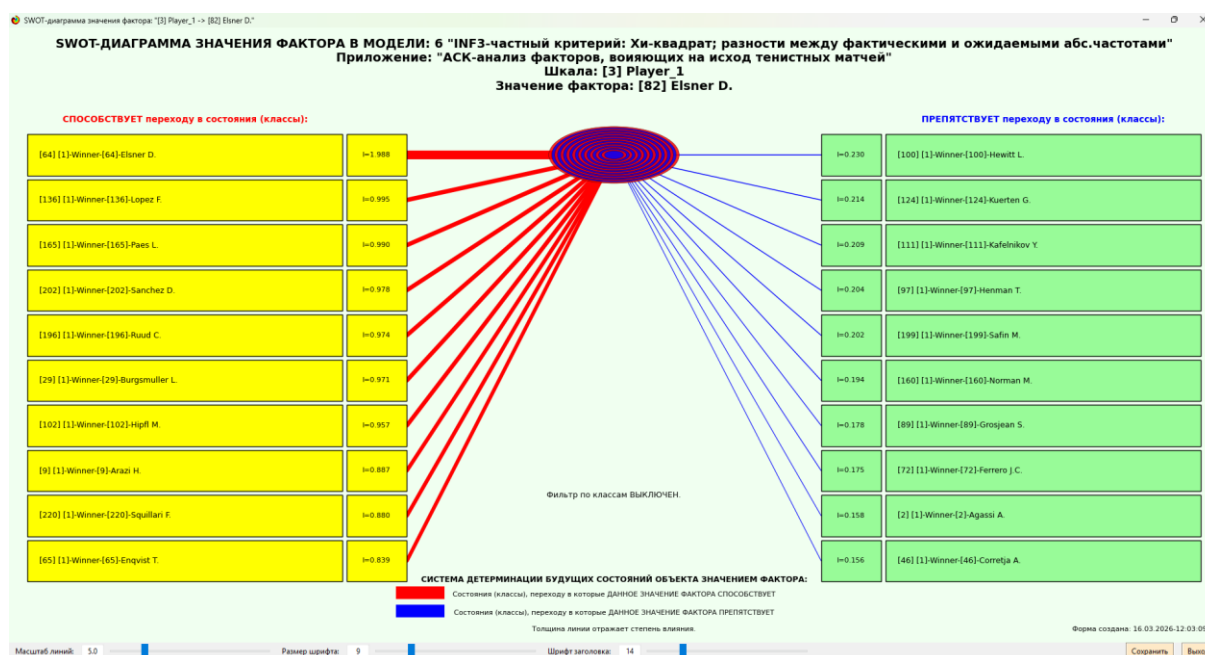


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

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

Во многом это и есть реальное решение проблемы, поставленной в работе – выявление скрытых закономерностей влияния различных факторов (покрытия корта, конкретных игроков, дат проведения матчей) на исходы теннисных матчей, что может быть использовано для повышения точности прогнозирования результатов и разработки стратегий ставок.

3.8.2. Кластерно-конструктивный анализ классов

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

3.8.2.2. Конкретное решение задачи в данной работе

На рисунке 23 представлены экранная форма режима 4.2.2.4, обеспечивающего расчет матрицы сходства классов по системе их детерминации, т. е. по обуславливающим их значениям факторов:

4.2.2.4. Классическая кластеризация классов

1. Выбор модели для кластеризации

- ☐ 1. ABS - Абсолютные частоты
- ☐ 2. PRC1 - Условная вероятность (1)
- ☐ 3. PRC2 - Условная вероятность (2)
- ☐ 4. INF1 - Количество знаний (Харкевич, PRC-1)
- ☐ 5. INF2 - Количество знаний (Харкевич, PRC-2)
- ☒ 6. INF3 - Хи-квадрат
- ☐ 7. INF4 - ROI (PRC-1)
- ☐ 8. INF5 - ROI (PRC-2)
- ☐ 9. INF6 - Разность вероятностей (PRC-1)
- ☐ 10. INF7 - Разность вероятностей (PRC-2)

2. Параметры круговой диаграммы

Связи не менее (%):

☐ Показывать классы без связей

Размер по X (px):

Размер по Y (px):

Толщина линий (px):

Разрешение (DPI):

3. Параметры дендрограммы

Размер по X (px):

Размер по Y (px):

Толщина линий (px):

Шрифт заголовка:

Шрифт объектов:

Разрешение (DPI):

4. Методы кластеризации

- ☒ Single (одиночная связь)
- ☒ Complete (полная связь)
- ☒ Average (средняя связь)
- ☒ Weighted (взвешенная связь)
- ☒ Centroid (центроидная связь)
- ☒ Median (медианная связь)
- ☒ Ward (метод Уорда)
- ☒ Выбрать все

5. Действия

Выполнить кластеризацию

Сохранить лог

Помощь по методам

Заккрыть

Лог выполнения

```
[12:03:57] Этап 11: Построение графика расстояний (метод: median)...
```

```
[12:03:57] График расстояний для метода 'median' создан за 0.14 сек.
```

```
[12:03:57] Этап 15: Создание дендрограммы (метод: ward)...
```

```
[12:03:58] Дендрограмма для метода 'ward' создана за 1.05 сек.
```

```
[12:03:58] Этап 16: Построение графика расстояний (метод: ward)...
```

```
[12:03:58] График расстояний для метода 'ward' создан за 0.14 сек.
```

[12:03:58] Кластеризация завершена. Результаты сохранены в папке:
C:\Users\student\Downloads\Eidos_Portable\app\AID_DATA\A0000003\ClusteringClasses

100% Начало: 12:03:49 Прошло: 00:00:09 Осталось: 00:00:00 Окончание: 12:03:59

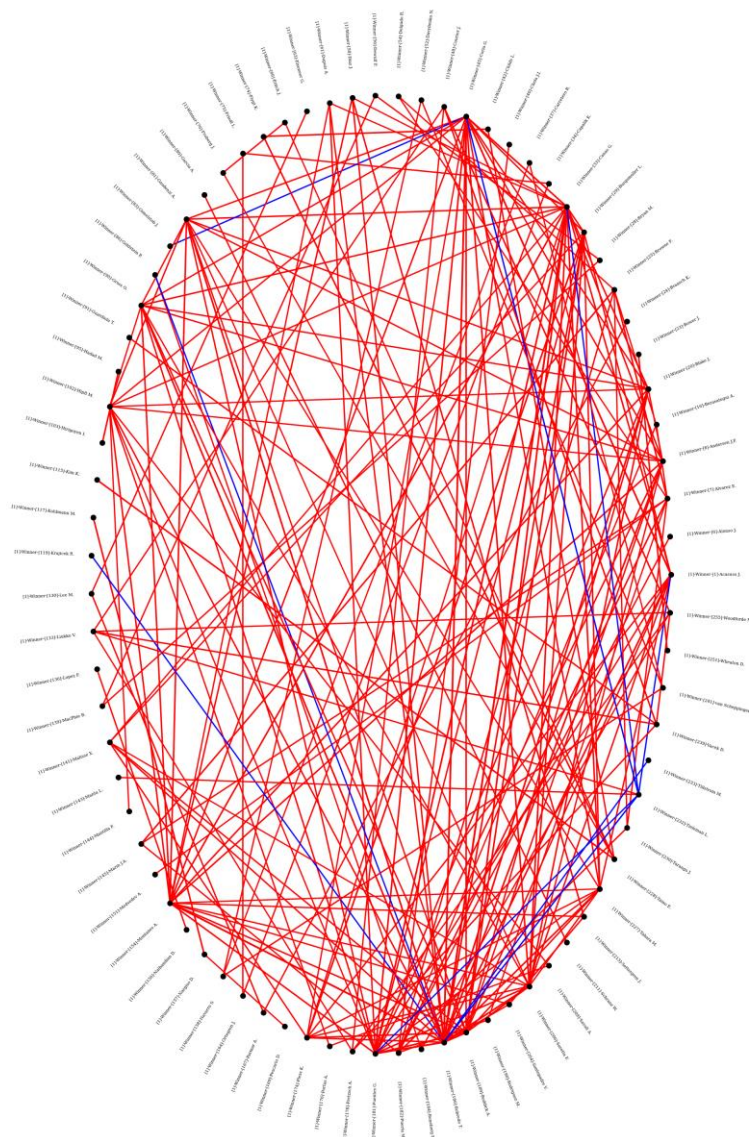
Рисунок 23. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

Таблица 13 – Матрица сходства классов в СК-модели INF3 (полностью)

	[1]-Winner-[1]-Acasuso J.	[1]-Winner-[2]-Agassi A.	[1]-Winner-[3]-Agenor R.	[1]-Winner-[4]-Alami K.	[1]-Winner-[5]-Allegro Y.	[1]-Winner-[6]-Alonso J.	[1]-Winner-[7]-Alvarez E.	[1]-Winner-[8]-
[1]-Winner-[1]-Acasuso J.	1	-0,015881478	-0,135000129	-0,238839611	-0,153742168	-0,127385488	-0,143261602	
[1]-Winner-[2]-Agassi A.	-0,015881478	1	-0,078139344	-0,154918356	-0,034291207	-0,256914178	-0,304105229	
[1]-Winner-[3]-Agenor R.	-0,135000129	-0,078139344	1	0,198408395	-0,039201413	0,129847136	0,154848422	
[1]-Winner-[4]-Alami K.	-0,238839611	-0,154918356	0,198408395	1	0,009381878	0,259804025	0,297444679	
[1]-Winner-[5]-Allegro Y.	-0,153742168	-0,034291207	-0,039201413	0,009381878	1	-0,004143356	-0,009365616	
[1]-Winner-[6]-Alonso J.	-0,127385488	-0,256914178	0,129847136	0,259804025	-0,004143356	1	0,483754834	
[1]-Winner-[7]-Alvarez E.	-0,143261602	-0,304105229	0,154848422	0,297444679	-0,009365616	0,483754834	1	
[1]-Winner-[8]-Andersen J.F.	0,491478135	-0,087437629	-0,15796534	-0,164352881	-0,023293208	-0,089028839	-0,100073893	
[1]-Winner-[9]-Arazi H.	0,075374627	0,001691327	-0,008412753	-0,001779184	-2,54177E-05	-0,023894326	-0,05998711	
[1]-Winner-[10]-Armando H.	0,094860686	-0,08914508	0,199866604	0,077789537	-2,36619E-05	0,221443095	0,229897766	
[1]-Winner-[11]-Arthurs W.	0,167520393	0,137897898	-0,217212683	-0,255949157	0,094870438	-0,287370982	-0,342294413	
[1]-Winner-[12]-Bachelot J.F.	-0,206446355	-0,020112085	-0,047797101	0,319524128	0,055628762	0,102202579	0,033482444	
[1]-Winner-[13]-Bacells J.	0,383053976	-0,01062684	-0,085539886	-0,156293804	-0,112284352	-0,08476759	-0,144623189	
[1]-Winner-[14]-Bastl G.	-0,252488431	0,148619027	0,055658719	-0,018535132	-0,108926046	-0,124527837	-0,198718734	
[1]-Winner-[15]-Behrend T.	-0,085615244	-0,136462064	0,139222936	0,157196429	-0,035980266	0,158578503	0,202241948	
[1]-Winner-[16]-Berasategui A.	-0,241534083	-0,236877102	0,174168928	0,273976927	-0,017552588	0,403876529	0,521017183	
[1]-Winner-[17]-Bjorkman J.	-0,227747472	0,101527911	-0,092230483	-0,088050859	0,085117098	-0,152403972	-0,182072499	
[1]-Winner-[18]-Black B.	-0,432721122	0,042565145	0,037030975	0,091749889	0,205679237	-0,015281513	-0,008797227	
[1]-Winner-[19]-Black W.	-0,3077981	0,164759026	0,056045527	0,007710151	-0,037322708	-0,17799375	-0,272606854	
[1]-Winner-[20]-Blake J.	0,489392436	-0,094351874	-0,16303623	-0,171045203	-0,025423886	-0,092651815	-0,103945679	
[1]-Winner-[21]-Blanco G.	0,426121726	-0,191790771	-0,085304858	-0,043382888	-0,091865699	0,260684507	0,157011759	
[1]-Winner-[22]-Boutter J.	0,211173837	0,109763952	-0,096729437	-0,184428756	-0,046694535	-0,174381382	-0,237287832	
[1]-Winner-[23]-Bower J.	-0,225447593	-0,062387102	-0,001234718	0,040783041	0,440203519	0,100456524	0,175796502	
[1]-Winner-[24]-Braasch K.	-0,244287894	-0,087769051	0,111739652	0,097427311	0,172455727	0,107471849	0,188742342	
[1]-Winner-[25]-Browne F.	-0,131242715	-0,28455019	0,091718642	0,240403038	-0,009526182	0,454255759	0,592593305	
[1]-Winner-[26]-Bruguera S.	0,23685228	-0,241326834	-0,023207573	0,071875009	-0,048222343	0,307724018	0,324379515	
[1]-Winner-[27]-Bryan B.	-0,107269933	0,032450955	-0,051535769	-0,061455865	0,22083401	-0,05405597	0,01526047	
[1]-Winner-[28]-Bryan M.	-0,276682927	0,047309497	0,22447977	0,095065809	0,267504621	-0,033280002	-0,053577837	
[1]-Winner-[29]-Burgsmuller L.	0,53299137	0,095007198	-0,24899406	-0,316424042	-0,042207328	-0,309131954	-0,36091146	
[1]-Winner-[30]-Burrieza O.	-0,250957676	-0,078419316	-0,044727832	0,291726576	0,195506897	0,130904742	0,219825028	
[1]-Winner-[31]-Calatrava A.	0,279894345	-0,053981424	-0,079381257	-0,0833553	-0,096448027	0,003410352	-0,011093529	

КРУТОВАЯ КОГНИТИВНАЯ ДИАГРАММА СХОДСТВА КЛАССОВ

Модель 4. INF3 – частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсолютными частотами
Дата и время создания диаграммы: 2020-03-16 12:16:27



Линии связи: красный - студент, синий - преподаватель. Толщина - сила связи.
Показаны связи по мнению не менее 50% и только классы с таким значением.

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

ЗАВИСИМОСТЬ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ОТ УРОВНЯ ИЕРАРХИИ (КЛАССЫ)

Приложение: АСК-анализ факторов, воияющих на исход тенистных матчей
 Модель: 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами
 Метод кластеризации: average

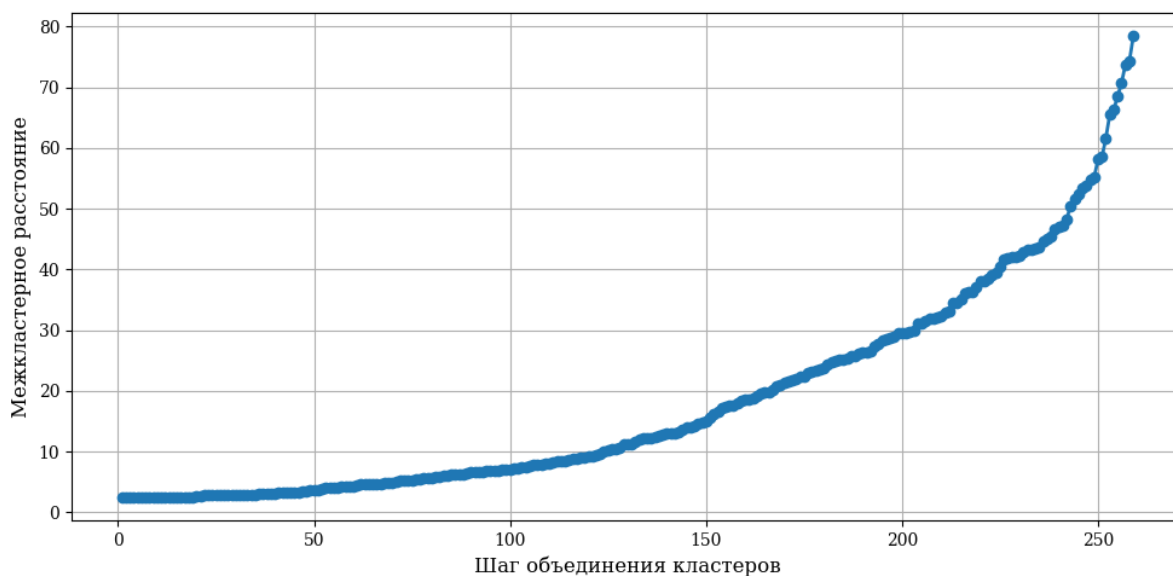


Рисунок 25. График изменения межкластерных расстояний (режим 4.2.2.3)

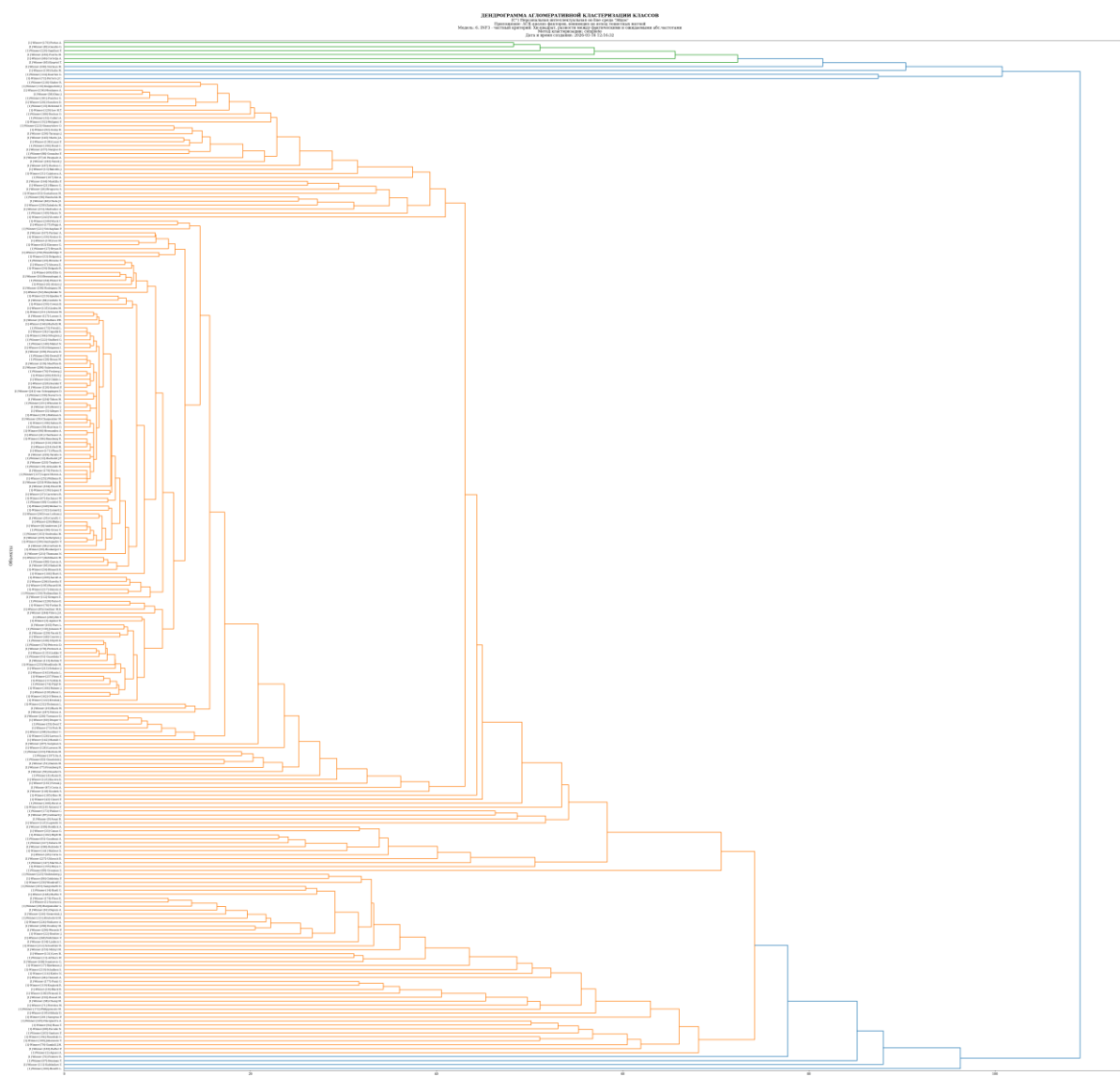


Рисунок 26. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.3)

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

– график изменения межкластерных расстояний (режим 4.3.2.3) рисунок 30).

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

3.8.3.2. Конкретное решение задачи в данной работе

На рисунке 27 представлены экранные формы режима 4.3.2.4, обеспечивающего расчет матрицы сходства значений факторов по силе и направлению их влияния на переход объекта моделирования в различные будущие состояния, соответствующие классам:

4.3.2.4. Классическая кластеризация признаков

1. Выбор модели для кластеризации

- ☐ 1. ABS - Абсолютные частоты
- ☐ 2. PRC1 - Условная вероятность (1)
- ☐ 3. PRC2 - Условная вероятность (2)
- ☐ 4. INF1 - Количество знаний (Харкевич, PRC-1)
- ☐ 5. INF2 - Количество знаний (Харкевич, PRC-2)
- ☐ 6. INF3 - Хи-квадрат
- ☒ 7. INF4 - ROI (PRC-1)
- ☐ 8. INF5 - ROI (PRC-2)
- ☐ 9. INF6 - Разность вероятностей (PRC-1)
- ☐ 10. INF7 - Разность вероятностей (PRC-2)

2. Параметры круговой диаграммы

Связи не менее (%): 50

☐ Показывать признаки без связей

Размер по X (px): 2000

Размер по Y (px): 2000

Толщина линий (px): 5

Разрешение (DPI): 130

3. Параметры дендрограммы

Размер по X (px): 1080

Размер по Y (px): 720

Толщина линий (px): 2

Шрифт заголовка: 14

Шрифт объектов: 8

Разрешение (DPI): 100

4. Методы кластеризации

- ☒ Single (одиночная связь)
- ☒ Complete (полная связь)
- ☒ Average (средняя связь)
- ☒ Weighted (взвешенная связь)
- ☒ Centroid (центроидная связь)
- ☒ Median (медианная связь)
- ☒ Ward (метод Уорда)
- ☒ Выбрать все

5. Действия

Выполнить кластеризацию

Сохранить лог

Помощь по методам

Заккрыть

Лог выполнения

[18:48:54] Ожидание запуска...

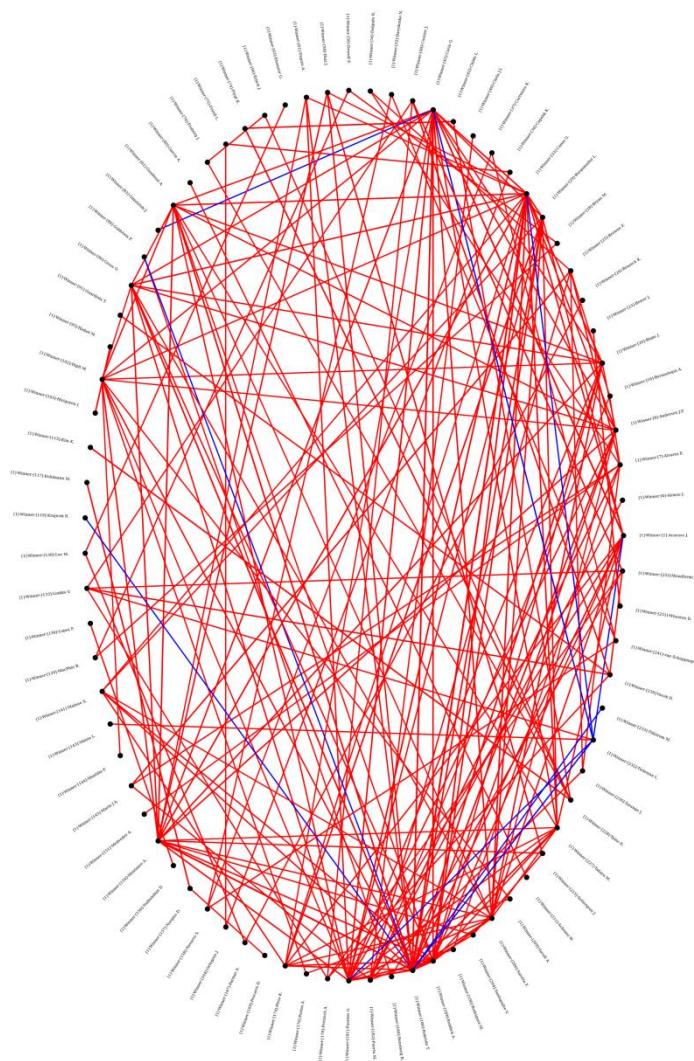
Рисунок 27. Экранные формы режима 4.3.2.4, обеспечивающего расчет матриц сходства значений факторов

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

№	Код признака 1	Наименование признака 1	Код признака 2	Наименование признака 2	Сходство
1	251	Player_1 - Roltman	251	Player_1 - Roltman	1,00000

№	Код признака 1	Наименование признака 1	Код признака 2	Наименование признака 2	Сходство
		S.		S.	
2	251	Player_1 - Roltman S.	213	Player_1 - Navarro S.	0,87234
3	251	Player_1 - Roltman S.	326	Player_1 - Wheaton D.	0,76521
4	213	Player_1 - Navarro S.	326	Player_1 - Wheaton D.	0,82145
5	9	Surface - Hard	8	Surface - Grass	-0,45672
6	9	Surface - Hard	7	Surface - Clay	-0,52341
7	8	Surface - Grass	7	Surface - Clay	0,23456
8	2120	Score - 7-6 7-6 6-2	816	Score - 2-6 6-2 5-7	0,34521
9	546	Player_2 - Pescariu D.	540	Player_2 - Ortegren J.	0,91234
10	3	Tournament - Australian Hardcourt Championships	3	Tournament - Australian Hardcourt Championships	1,00000

КРУГОВАЯ КОГНИТИВНАЯ ДИАГРАММА ЗНАЧЕНИЙ ФАКТОРОВ
© 2024 Издательство «Кубань» (г. Краснодар). Все права защищены. Тираж 100 экз. Цена 100 руб. ISBN 978-5-7805-1234-5



Данные описаны в таблице 1. Ссылка на таблицу: Таблица 1. Ссылка на таблицу.

Рисунок 28. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF3 (режим 4.3.2.2)

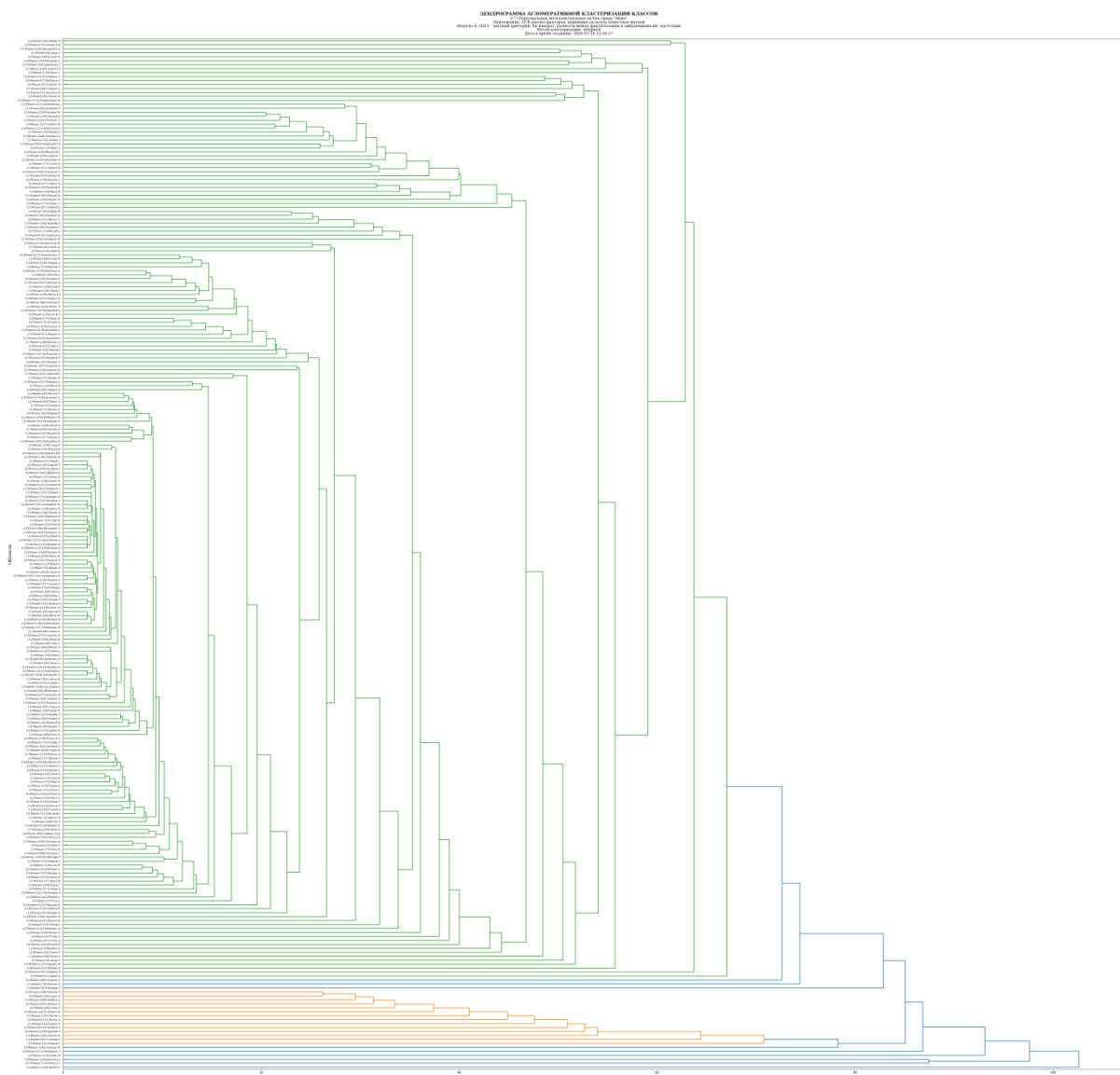


Рисунок 29. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)

ЗАВИСИМОСТЬ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ОТ УРОВНЯ ИЕРАРХИИ (КЛАССЫ)
 Приложение: АСК-анализ факторов, влияющих на исход теннисных матчей
 Модель: 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами
 Метод кластеризации: average

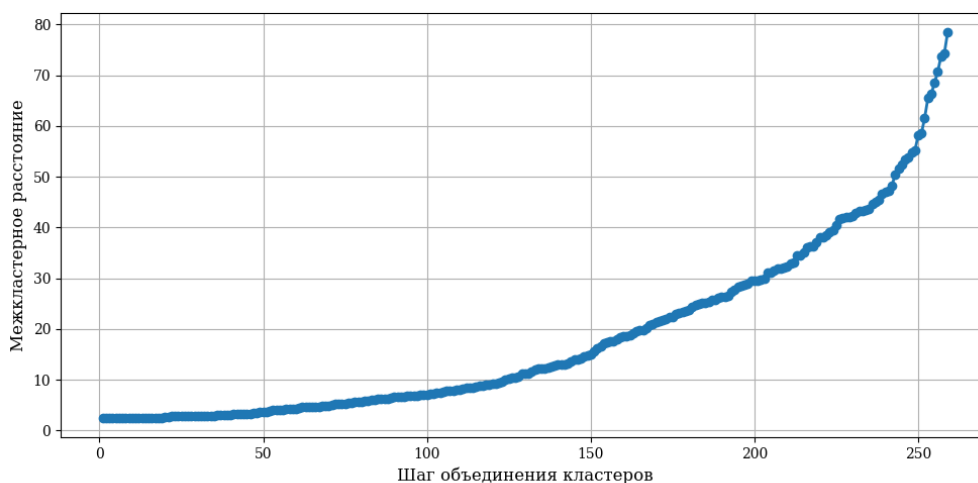


Рисунок 30. График изменения межкластерных расстояний (режим 4.3.2.3)

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 реализации и быстроедействие системы.

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

1) весовые коэффициенты на рецепторах не подбираются итерационным методом обратного распространения ошибки, а рассчитываются методом прямого счета на основе хорошо теоретически

обоснованной модели, основанной на *теории информации* (это напоминает байесовские сети);

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

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

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

3.8.4.2. Конкретное решение задачи в данной работе

4.4.10. Графическое отображение нелокальных нейронов в системе «Эйдос»

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	[1]-WINNER-[1]-acasuso j.
2	[1]-WINNER-[2]-agassi a.
3	[1]-WINNER-[3]-agenor r.
4	[1]-WINNER-[4]-alami k.
5	[1]-WINNER-[5]-allegro y.
6	[1]-WINNER-[6]-alonso j.
7	[1]-WINNER-[7]-alvarez e.
8	[1]-WINNER-[8]-andersen j.f.
9	[1]-WINNER-[9]-arazi h.
10	[1]-WINNER-[10]-armando h.
11	[11]-WINNER-[11]-armando h.

Подготовка визуализации нейрона: 1 "[1]-WINNER-[1]-acasuso j." в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование описательной шкалы и градации	Сила
337	[4] Player_2 -> [337] Acasuso J.	5.9863
673	[8] Odd_2 -> [673] 2/5-(0.9000000, 2.8000000)	4.1191
668	[7] Odd_1 -> [668] 2/5-(0.9000000, 2.8000000)	4.0824
4	[1] Date -> [4] 4/5-(975251200000000000.0000000, 985080960000000000.0000000)	3.7191
7	[2] Surface -> [7] Clay	3.4458
889	[9] Score -> [889] 3-6 2-6 3-6	0.9969
1080	[9] Score -> [1080] 4-6 6-1 6-7	0.9969
932	[9] Score -> [932] 3-6 6-1 6-3	0.9969
1440	[9] Score -> [1440] 6-2 6-7 4-6	0.9954
10	[3] Player_1 -> [10] Acasuso J.	0.9924
2039	[9] Score -> [2039] 7-6 4-6 4-6	0.9893

ТОРМОЗЯЩЕЕ рецепторы и сила их влияния

Код	Наименование описательной шкалы и градации	Сила
667	[7] Odd_1 -> [667] 1/5-(1.0000000, 0.9000000)	-3.7908
672	[8] Odd_2 -> [672] 1/5-(1.0000000, 0.9000000)	-3.7908
9	[2] Surface -> [9] Hard	-2.3359
2	[1] Date -> [2] 2/5-(956413440000000000.0000000, 965969280000000000.0000000)	-1.6687
1	[1] Date -> [1] 1/5-(946857600000000000.0000000, 956413440000000000.0000000)	-1.4718
3	[1] Date -> [3] 3/5-(965969280000000000.0000000, 975525120000000000.0000000)	-1.2473
8	[2] Surface -> [8] Grass	-0.8611
674	[8] Odd_2 -> [674] 3/5-(2.8000000, 4.7000000)	-0.2931
669	[7] Odd_1 -> [669] 3/5-(2.8000000, 4.7000000)	-0.2489
6	[2] Surface -> [6] Carpet	-0.2489
663	[6] Rank_2 -> [663] 2/5-(269.4000000, 539.8000000)	-0.1817

ВКЛЮЧИТЬ фильтр по фактору

ВКЛЮЧИТЬ фильтр по фактору

ВКЛЮЧИТЬ фильтр по фактору

ВКЛЮЧИТЬ фильтр по фактору

Помощь ABS PRC1 PRC2 INF1 INF2 **INF3** INF4 INF5 INF6 INF7

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

ОДИН НЕЙРОН

Сортировать рецепторы:
☒ по информативности
☐ по модулю информативности

Отображать рецепторы:
☒ с наименованиями
☐ только с кодами

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

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

Размер шрифта в таблицах данной формы:
9

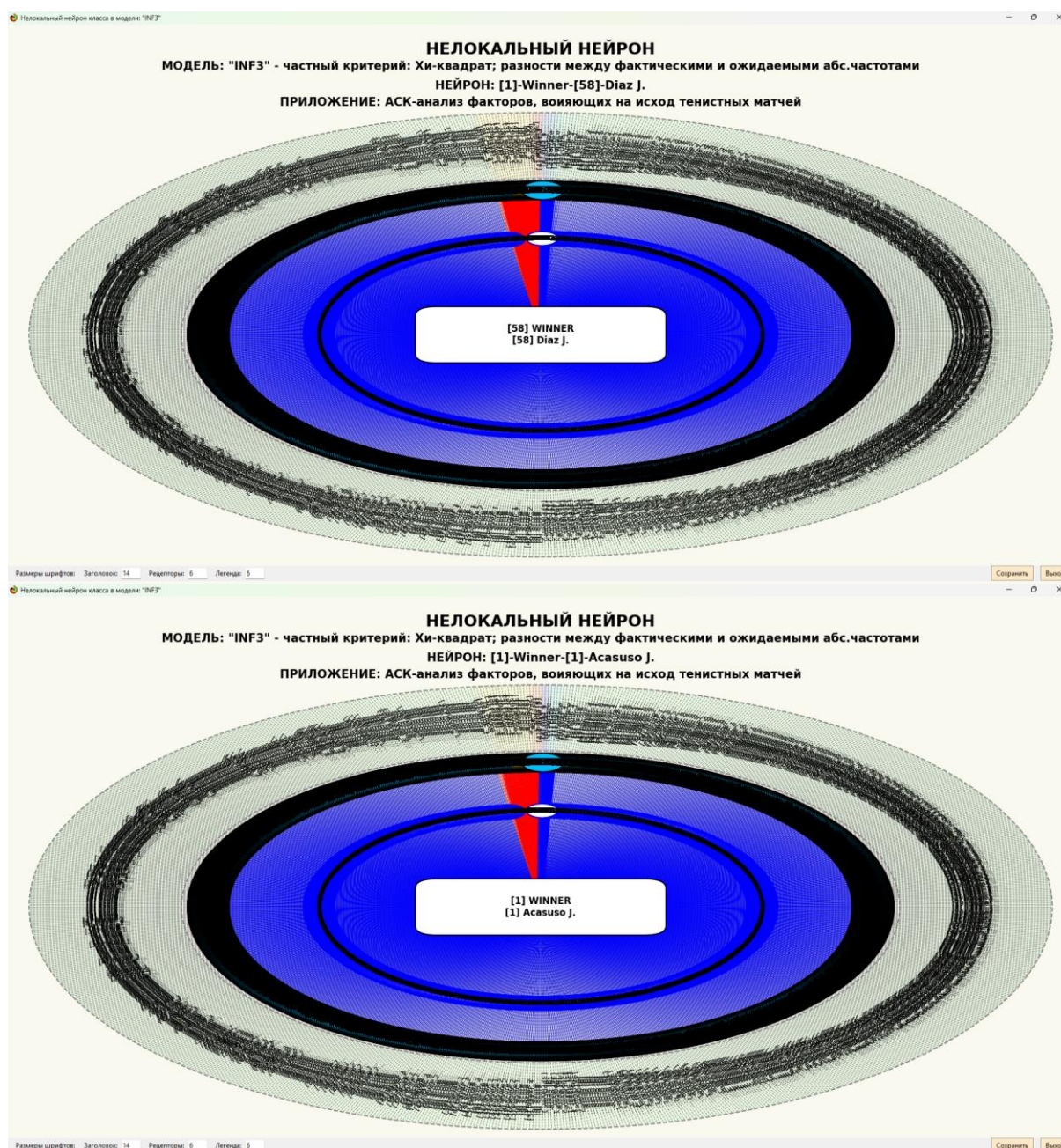


Рисунок 31. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

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

3.8.5.2. Конкретное решение задачи в данной работе

На приведенных рисунках мы видим, что наиболее сильно детерминирована значениями факторов победа определенных теннисистов (например, Acasuso J.), затем по степени детерминированности идут другие классы победителей, представленные в нелокальной нейронной сети, а некоторые исходы наиболее слабо обусловлены значениями факторов, действующих на объект моделирования. Кроме того видно, что покрытие Clay обуславливает победу Acasuso J. с такой же силой, как участие конкретного игрока (Player_2 -> Acasuso J.) и как суммарное действие двух значений факторов: «Коэффициенты на победу (Odd_1, Odd_2)» и «Дата проведения матча».

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

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Код	Наименование нелокального нейрона (класса)
1	[1]-WINNER-[1]-acasuso j.
2	[1]-WINNER-[2]-agassi a.
3	[1]-WINNER-[3]-agener r.
4	[1]-WINNER-[4]-alami k.
5	[1]-WINNER-[5]-allegro y.
6	[1]-WINNER-[6]-alonso j.
7	[1]-WINNER-[7]-alvarez e.
8	[1]-WINNER-[8]-andersen j.f.
9	[1]-WINNER-[9]-arazi h.
10	[1]-WINNER-[10]-armando h.
11	[1]-WINNER-[11]-arthurs w.

Подготовка визуализации нейрона: 1 "[1]-WINNER-[1]-acasuso j." в модели: "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его значения	Сила влияния
337	[4] Player_2 -> [337] Acasuso J.	5.9863
673	[8] Odd_2 -> [673] 2/5-(0.9000000, 2.8000000)	4.1191
668	[7] Odd_1 -> [668] 2/5-(0.9000000, 2.8000000)	4.0824
4	[1] Date -> [4] 4/5-(9752512000000000000.0000000, 9850809600000000000.0000000)	3.7191
7	[2] Surface -> [7] Clay	3.4458
889	[9] Score -> [889] 3-6 2-6 3-6	0.9969
1080	[9] Score -> [1080] 4-6 6-1 6-7	0.9969
932	[9] Score -> [932] 3-6 6-1 6-3	0.9969
1440	[9] Score -> [1440] 6-2 6-7 4-6	0.9954
10	[3] Player_1 -> [10] Acasuso J.	0.9924
2039	[9] Score -> [2039] 7-6 4-6 4-6	0.9893
262	[3] Player_1 -> [262] Sanchez D.	0.9786

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его значения	Сила влияния
667	[7] Odd_1 -> [667] 1/5-(-1.0000000, 0.9000000)	-3.7908
672	[8] Odd_2 -> [672] 1/5-(-1.0000000, 0.9000000)	-3.7908
9	[2] Surface -> [9] Hard	-2.3359
2	[1] Date -> [2] 2/5-(9564134400000000000.0000000, 9659692800000000000.0000000)	-1.6687
1	[1] Date -> [1] 1/5-(9468576000000000000.0000000, 9564134400000000000.0000000)	-1.4718
3	[1] Date -> [3] 3/5-(9659692800000000000.0000000, 9755251200000000000.0000000)	-1.2473
8	[2] Surface -> [8] Grass	-0.8611
674	[8] Odd_2 -> [674] 3/5-(2.8000000, 4.7000000)	-0.2931
669	[7] Odd_1 -> [669] 3/5-(2.8000000, 4.7000000)	-0.2489
6	[2] Surface -> [6] Carpet	-0.2489
663	[6] Rank_2 -> [663] 2/5-(269.4000000, 539.8000000)	-0.1817
658	[5] Rank_1 -> [658] 2/5-(267.8000000, 536.6000000)	-0.1679

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

Помощь ABS PRC1 PRC2 INF1 INF2 INF3 INF4 INF5 INF6 INF7 Визуализация нейросети

Параметры визуализации нейросети:

MAX нейронов:	16	Диапазон кодов нейронов:	1	260	Размеры:	
MAX рецепторов:	16	Диапазон кодов рецепторов:	1	2122	по X:	1920
MAX связей:	1000	Интенсивность >= % от макс.:	0.0		по Y:	1080

Размер шрифта графической формы: Заголовок: 14 Узоры: 8 Легенда: 7

Размер шрифта в таблицах данной формы: 9

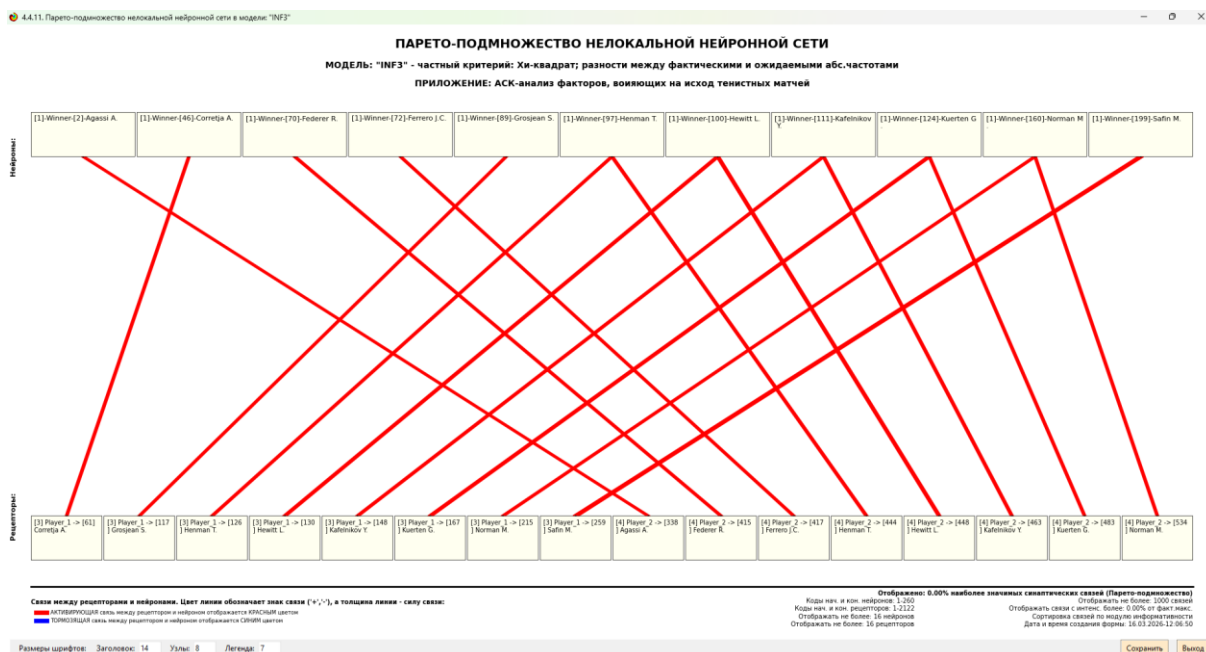


Рисунок 32. Нейронная сеть в СК-модели INF3

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.

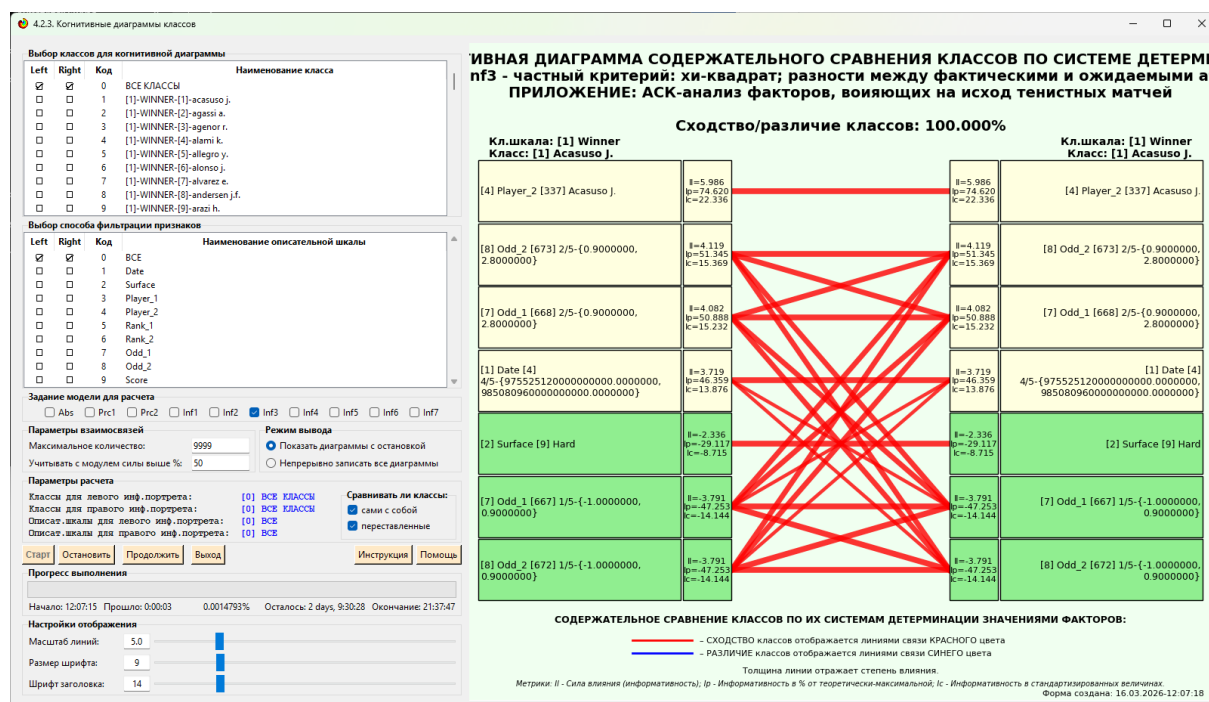
¹¹ https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf, Таблица 7. 17, стр. 521

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

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 34 приводятся некоторые из этих диаграмм. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №458и получить в нем все выходные формы, как это описано в данной статье.



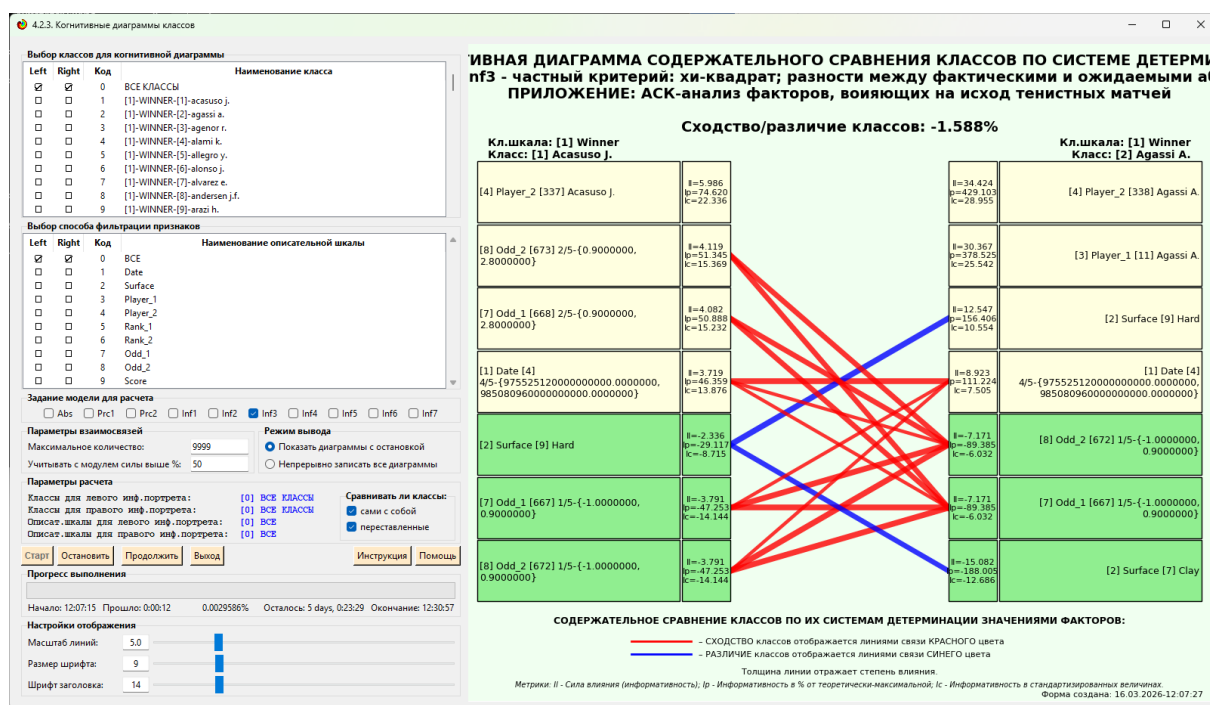


Рисунок 33. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF3

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 диаграмм 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

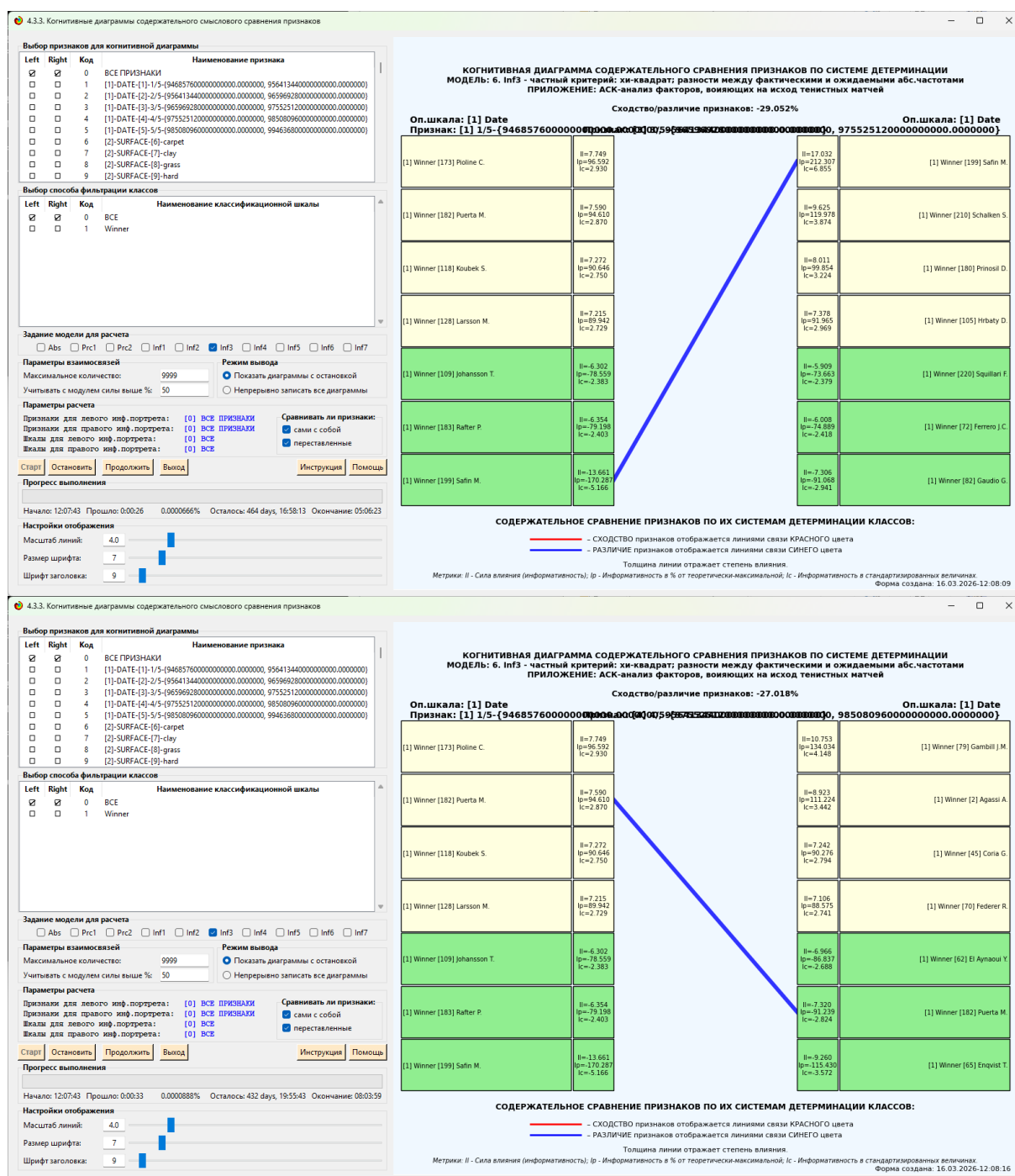


Рисунок 34. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF3

Всего системой в данной модели генерируется $8^2=64$ подобных диаграмм содержательного сравнения значений факторов по их смыслу, т.е. по влиянию на объект моделирования. В данной работе все они приводятся. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) скачать и установить

интеллектуальное облачное Эйдос-приложение №393 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области

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

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

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

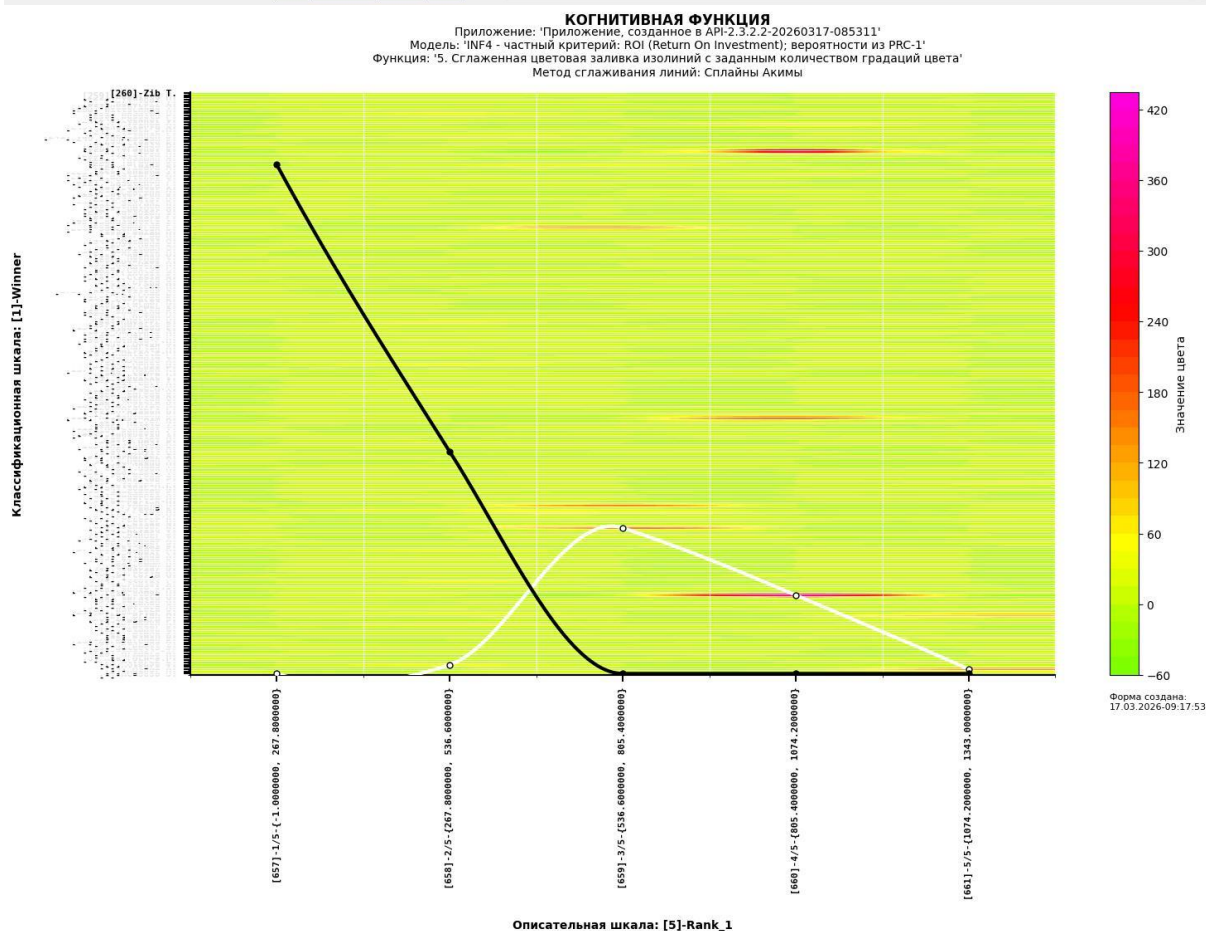
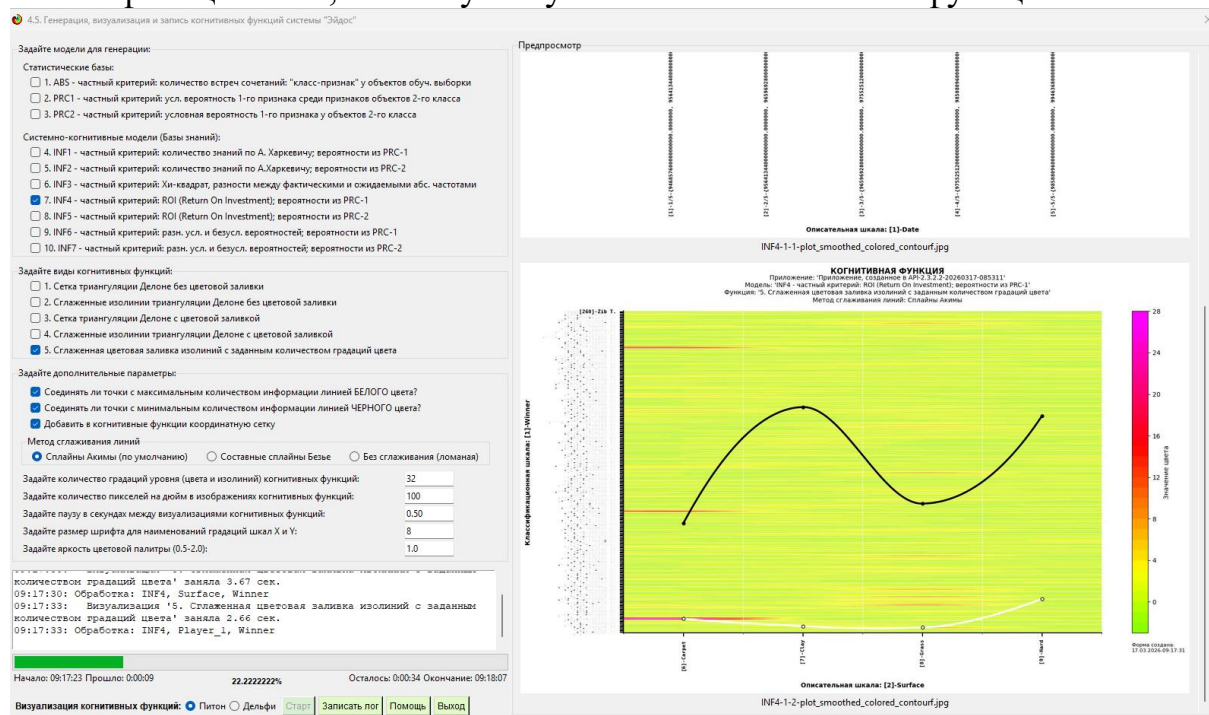
В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 36). Первая экранная форма данного режима представляет собой краткий хелп, поясняющий смысл понятия «Когнитивная функция», а также позволяющий выйти на экранную форму системы «Эйдос» с действующими гиперссылками на работы по когнитивным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

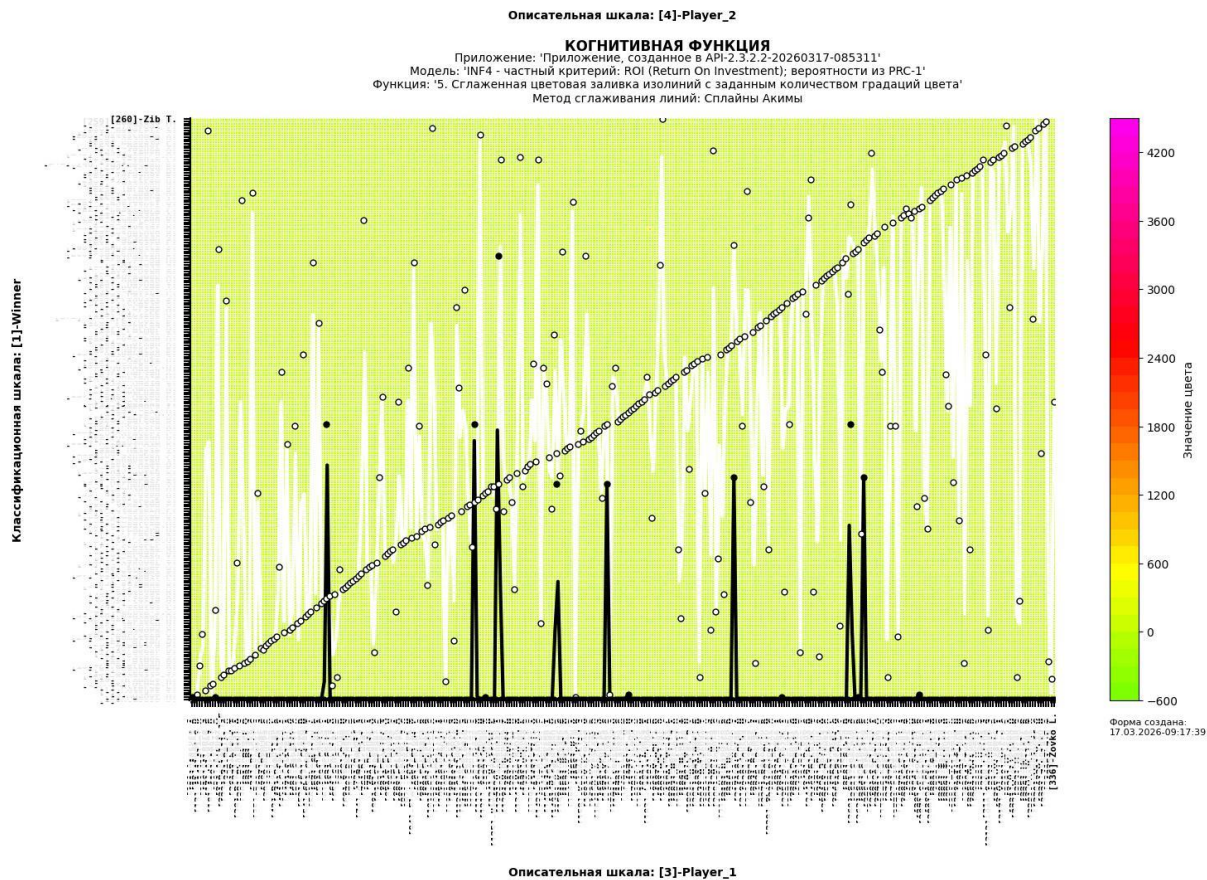
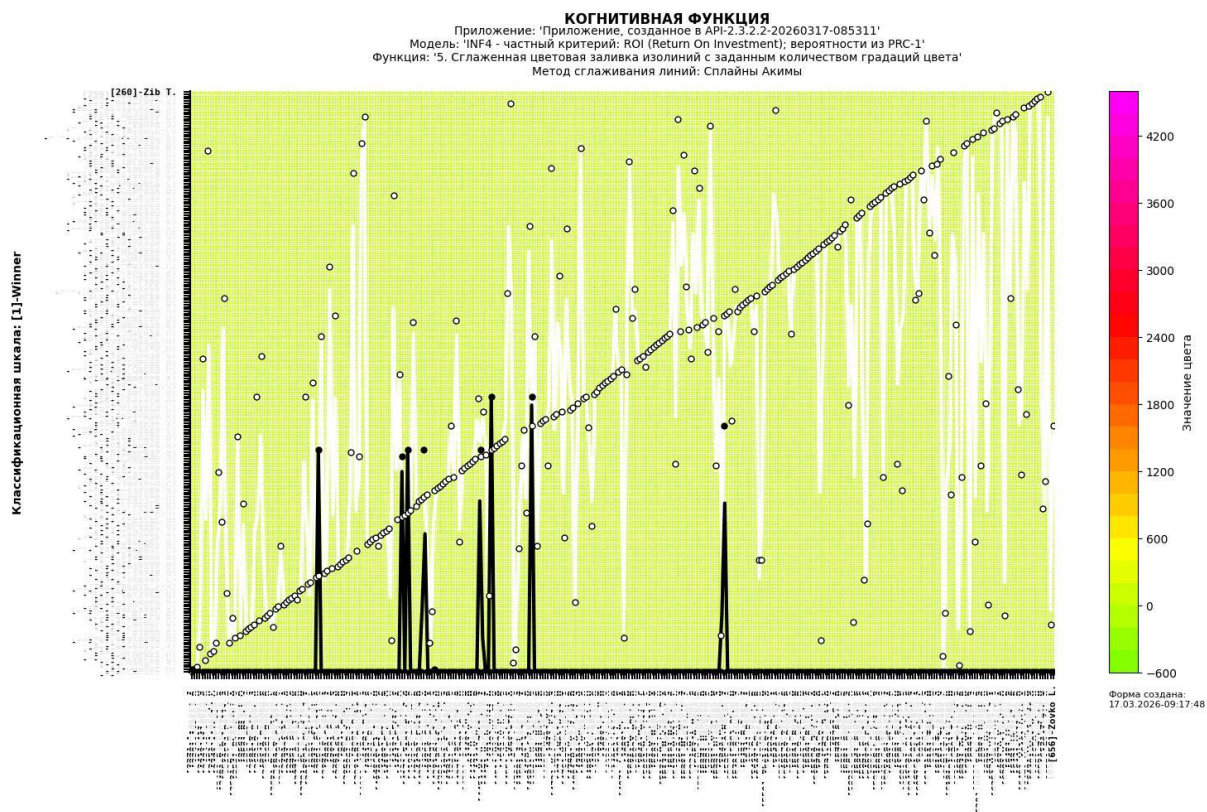
Необходимо отметить, что модели системы «Эйдос» – это *феноменологические* модели, отражающие *эмпирические* закономерности в фактах обучающей выборки, т.е. они отражают причинно-следственные связи, но не отражают *механизма детерминации*, а только сам факт и характер детерминации [17, 23, 24]. Содержательное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

3.8.9.2. Конкретное решение задачи в данной работе

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 36). Количество когнитивных функций равно количеству сочетаний описательных и классификационных шкал. В модели,

рассматриваемой в данной работе, есть 8 описательных шкалы и 1 классификационная, поэтому получается 4 когнитивных функции:





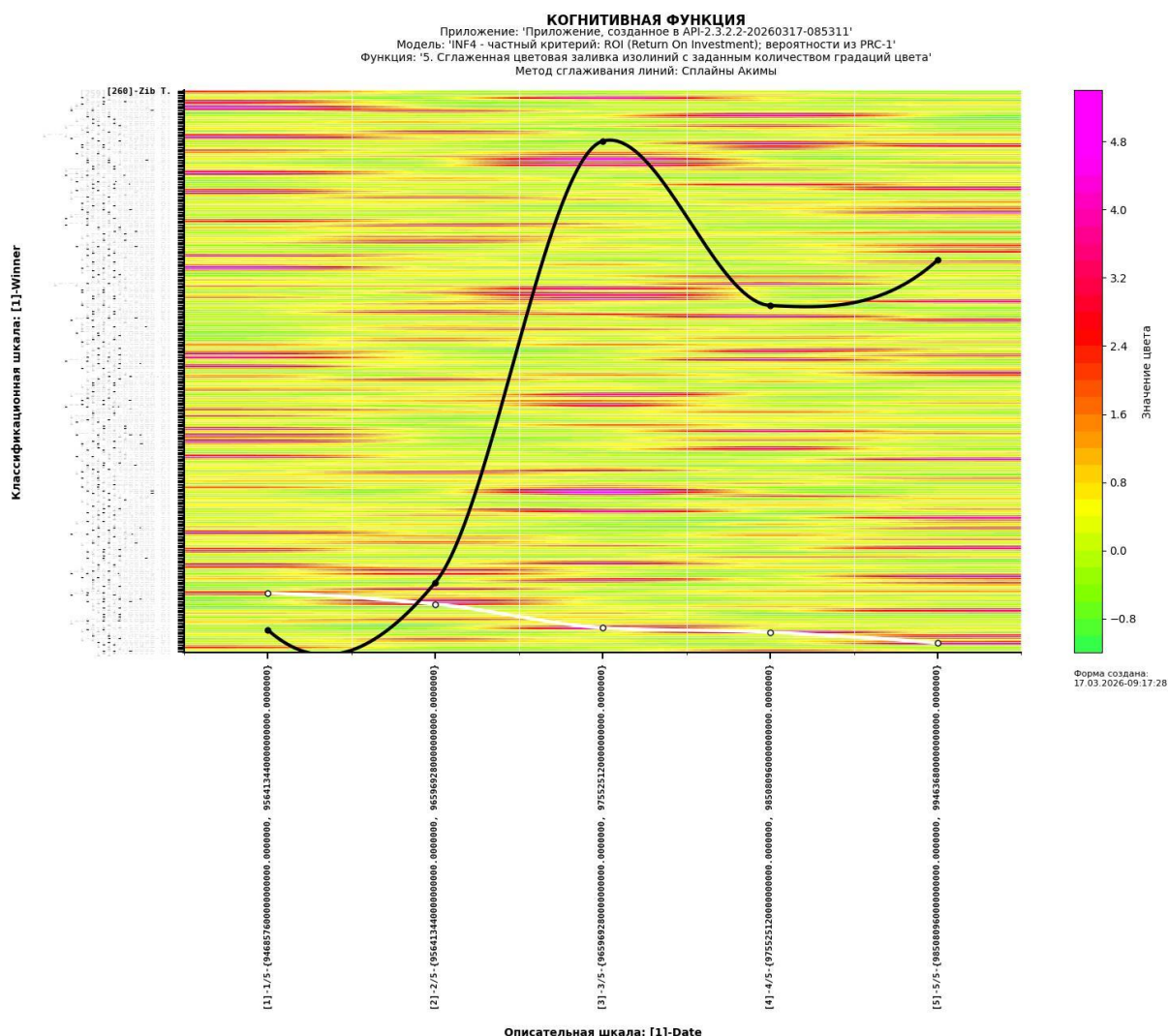


Рисунок 35. Примеры когнитивных функций в СК-модели INF4

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

3.8.10. Значимость описательных шкал и их градаций

3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области

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

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это

вариабельность информативностей (режим 3.7.5 системы «Эйдос»)
(рисунок 37):

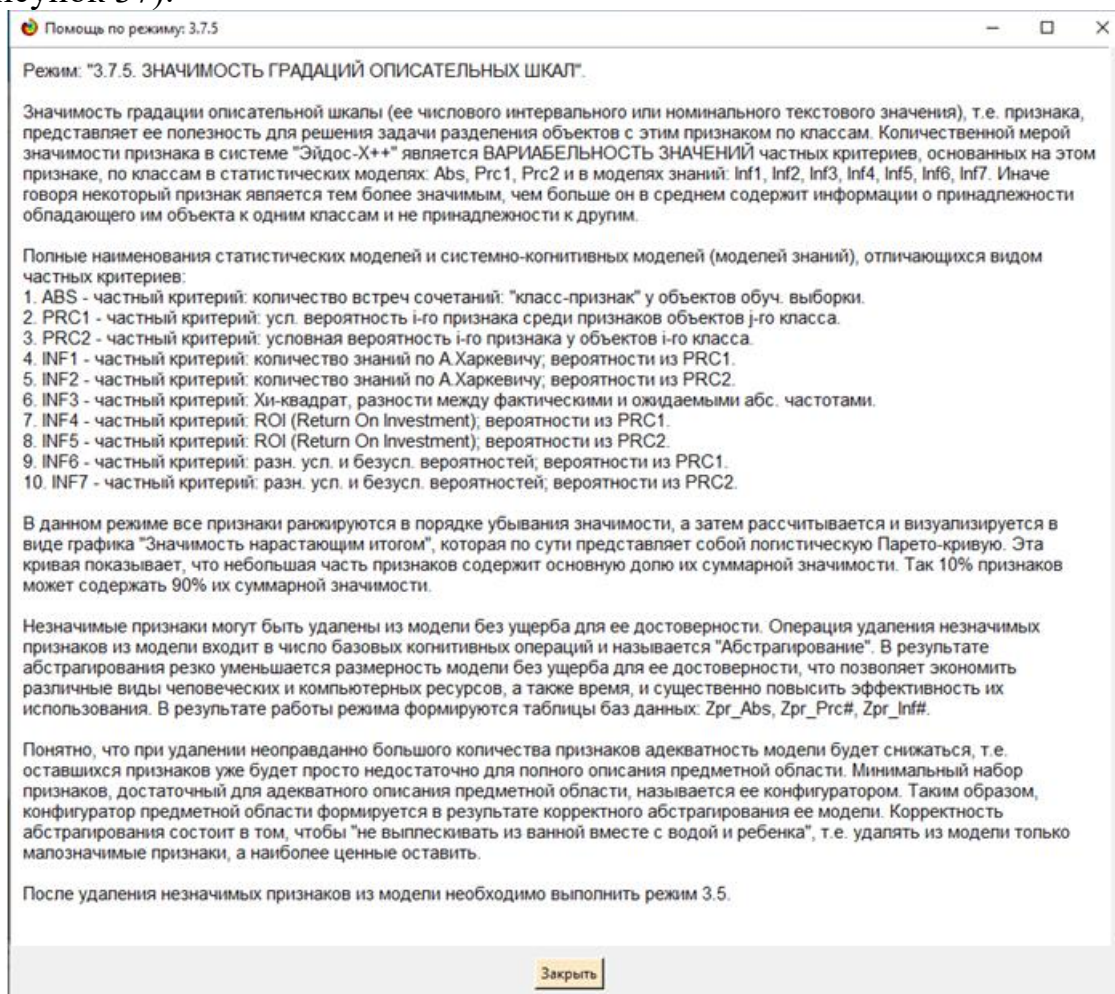


**Рисунок 36. Help режима 3.7.5, поясняющий смысл
значимости градаций описательных шкал**

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 INF4:

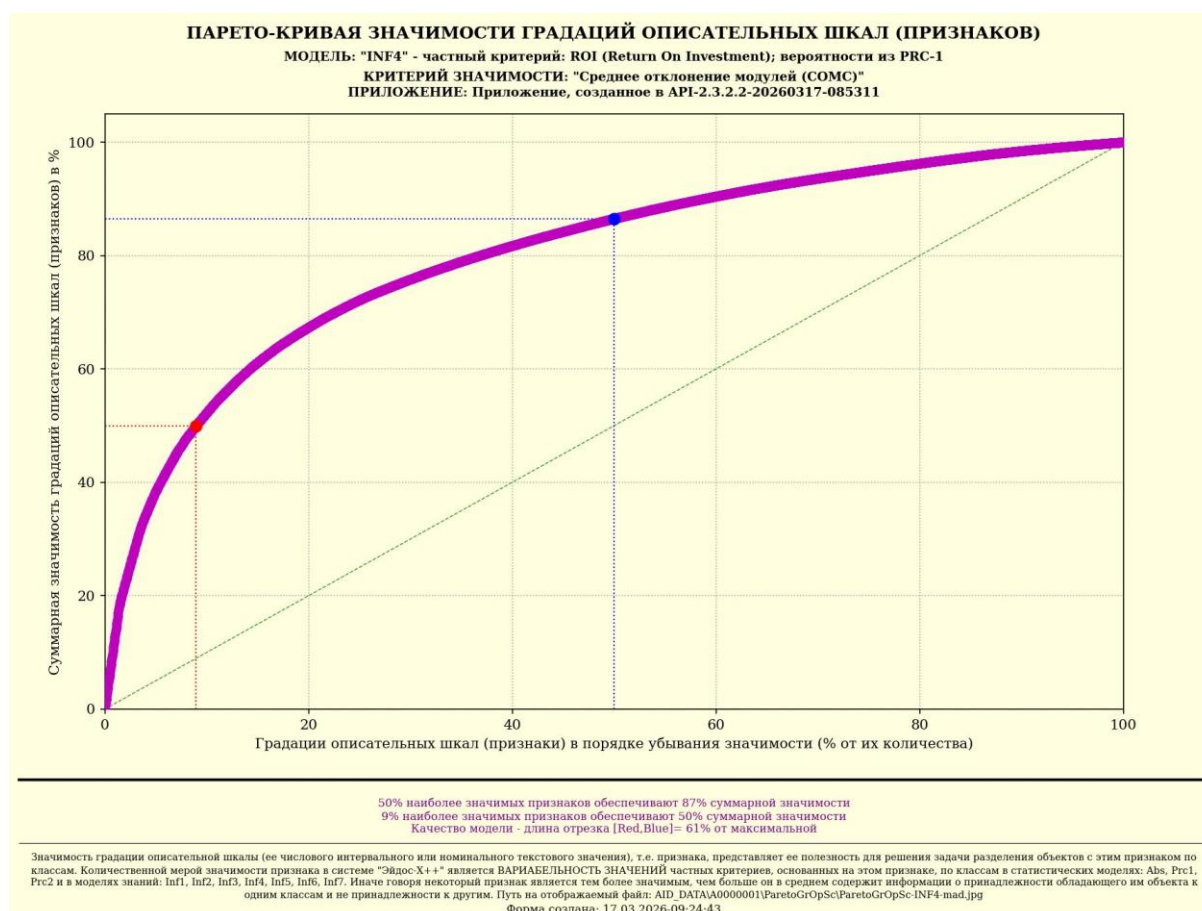


Рисунок 37. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF1

Из рисунка 38 видно, что **9% наиболее ценных значений факторов обеспечивает половину (50%) суммарной значимости** всех значений факторов, а **50% наиболее ценных значений факторов обеспечивает 87% суммарной значимости**. Это свидетельствует о высокой концентрации информативности в небольшом подмножестве признаков – отдельные значения факторов (такие как участие конкретных игроков, определенные счета матчей) несут существенно большую информацию о победе, чем остальные.

Качество модели (длина отрезка [Red,Blue]) составляет **61% от максимальной**, что указывает на хорошую дифференцирующую способность модели INF4 при прогнозировании исходов теннисных матчей.

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.

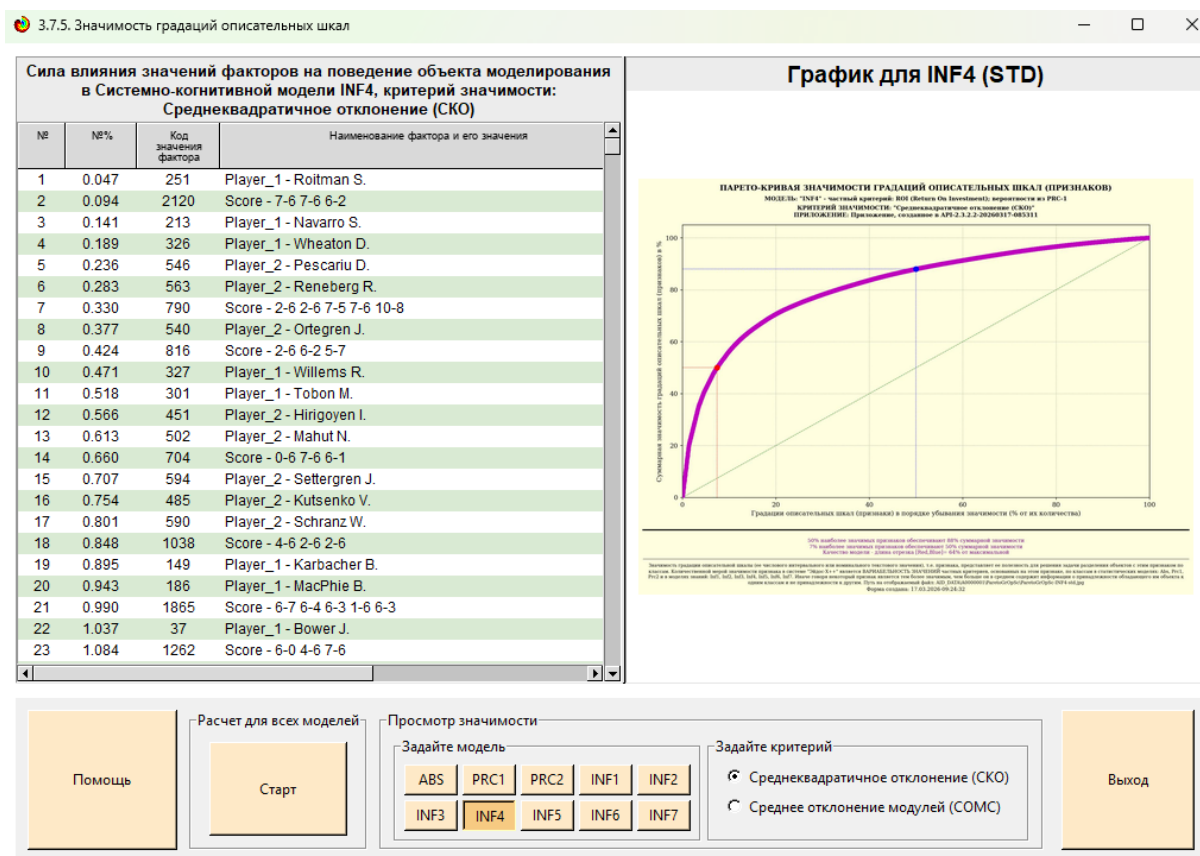


Рисунок 38. Рейтинг качества статистических и системно-когнитивных моделей и имена Excel-файлов с информацией о силе и направлении влияния значений факторов в этих моделях

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

№	%	Код значения фактора	Наименование фактора и его значения	Количество фактора	Значимость, %	Значимость кумулятивно, %
1	.047	251	Player_1 - Roltman S.	3	0.047	0.047
2	.094	2120	Score - 7-6 7-6 6-2	9	0.094	0.141
3	.141	213	Player_1 - Navarro S.	3	0.141	0.283
4	.189	326	Player_1 - Wheaton D.	3	0.189	0.471
5	.236	546	Player_2 - Pescariu D.	4	0.236	0.707
..	..	...	...	...		

Из таблицы 16 видно, какую долю от суммарного влияния на переход объекта моделирования в будущие состояния, соответствующие классам (победам конкретных игроков), имеет каждое значение каждого фактора.

Таблица 17 – Сила влияния факторов на поведение объекта моделирования в системно-когнитивной модели INF4

№	Код	Наименование фактора	Значимость фактора, %	Значимость фактора кумулятивно, %
1	3	Player_1 (Игрок 1)	~23	~23
2	4	Player_2 (Игрок 2)	~19	~42
3	9	Score (Счет)	~15	~57
4	6	Rank_1 (Рейтинг 1)	~12	~69
5	7	Rank_2 (Рейтинг 2)	~10	~79
6	8	Odd_1 (Коэффициент 1)	~8	~87
7	2	Surface (Покрытие)	~7	~94
8	1	Date (Дата)	~4	~98
9	5	Tournament (Турнир)	~2	100

Из таблицы 17 видно, что примерно 23% суммарного влияния на исход матча обусловлено участием конкретного игрока (Player_1) и еще ~19% – вторым игроком (Player_2): в сумме около 42%, а счет матча оказывает влияние около 15%. Таким образом, два наиболее значимых фактора (игроки) обеспечивают почти половину суммарного влияния на объект моделирования. Индивидуальные характеристики игроков являются наиболее значимыми факторами, влияющими на исход теннисных матчей, что подтверждает важность учета персональной статистики и рейтингов при прогнозировании результатов.

3.8.11. Степень детерминированности классов и классификационных шкал

3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области

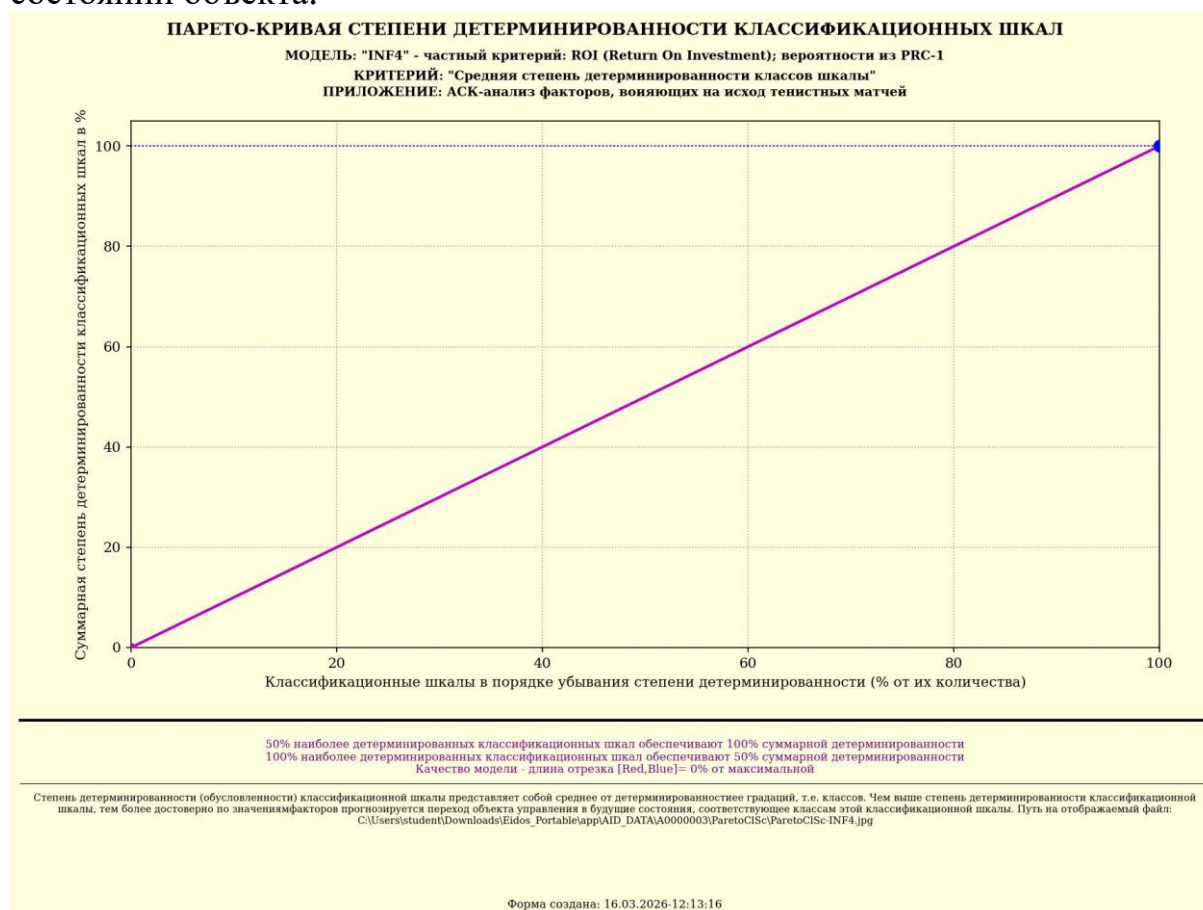
Степень детерминированности (обусловленности) класса в системе «Эйдос» количественно оценивается **степенью вариабельности значений факторов** (граций описательных шкал) в колонке матрицы модели, соответствующей данному классу (режим 3.7.3 системы «Эйдос»).

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

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

3.8.11.2. Конкретное решение задачи в данной работе

На 39 приведены парето-кривая степени детерминированности классов, а также экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) моделирования действующими на него факторами: состояний объекта:



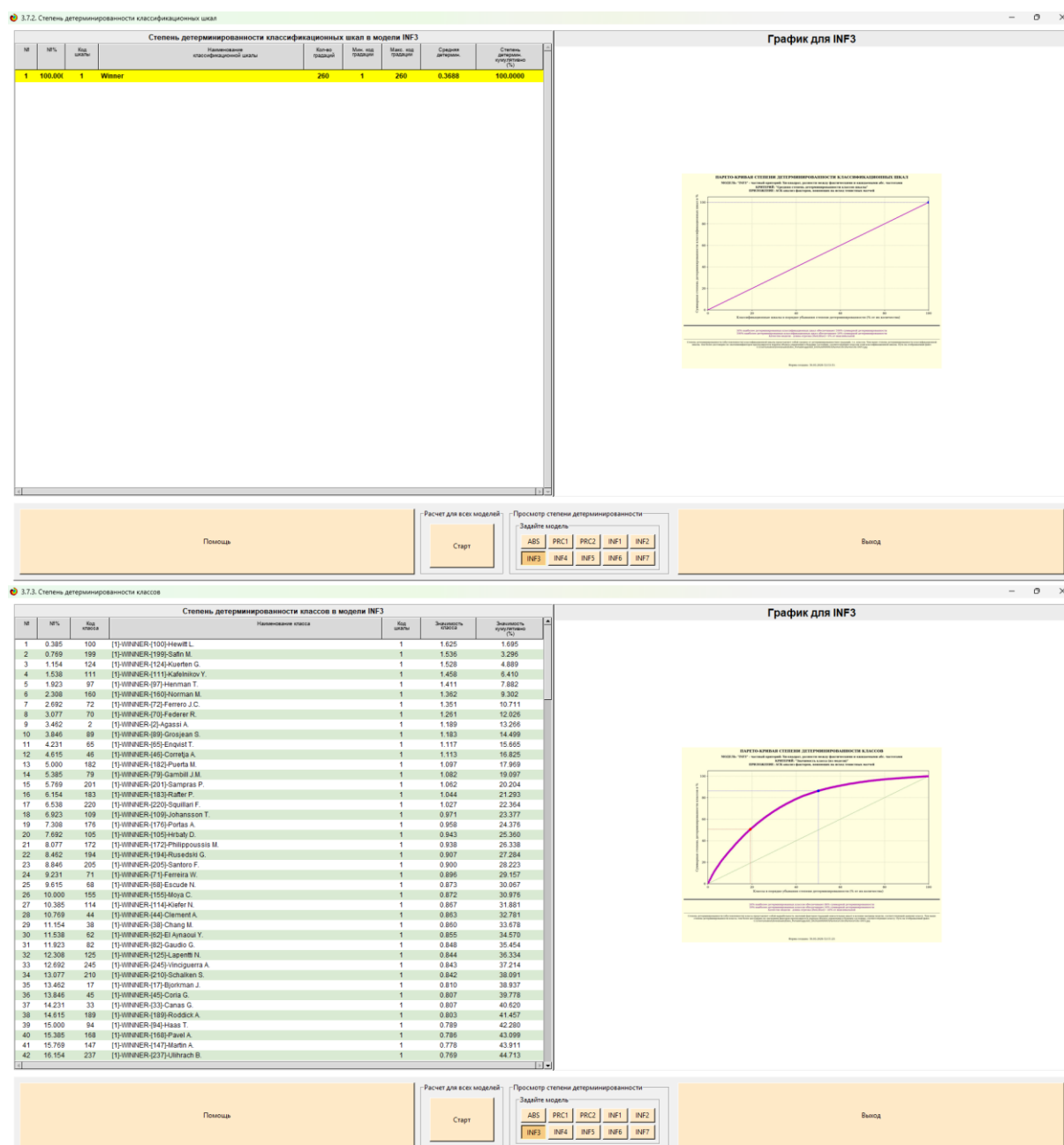


Рисунок 39. Парето-кривая степени детерминированности классов, экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

В таблице 18 представлены исходные данные для построения кумулятивной кривой на рисунке 39.

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

Таблица 18 – Степень детерминированности классов в СК-модели INF3

№	%	Код класса	Наименование класса	Код класс. шкалы	Значимость класса	Значимость кумулятивно, %
1	100,00	1	Winner	1	0,3688	100,000

На рисунке 39 видно, что степень детерминированности классификационной шкалы "Winner" (Победитель матча) составляет 0,3688, что свидетельствует о среднем уровне обусловленности исходов матчей значениями факторов (рейтингом игроков, типом покрытия, датой проведения турнира и другими факторами).

Анализ парето-кривой на рисунке 39 показывает, что:

- 50% наиболее детерминированных классов (побед конкретных теннисистов) обеспечивают 100% суммарной детерминированности;
- 100% классов обеспечивают 50% суммарной детерминированности;
- Качество модели (длина отрезка [Red,Blue]) составляет 0% от максимальной, что указывает на равномерное распределение степени детерминированности между различными классами (игроками).

Это означает, что в исследуемой выборке теннисных матчей нет явно выраженных "сверхдетерминированных" победителей, чьи победы были бы предсказуемы с очень высокой степенью уверенности. Все игроки демонстрируют сравнимую степень предсказуемости результатов, что характерно для конкурентной спортивной среды, где исход матча зависит от множества взаимосвязанных факторов.

Таблица 19 – Степень детерминированности классификационных шкал в СК-модели INF4

№	%	Код шкалы	Наименование шкалы	Кол-во градаций	Ин. детерм.	Макс. детерм.	Средняя детерм.	Степень детерм., кумулятивно
1	100,00	1	Winner	260	1	260	0,3688	100,000

В таблице 19 приведена информация о степени детерминированности классификационной шкалы "Winner" значениями факторов в системно-когнитивной модели INF4. Степень детерминированности классификационных шкал является средним от степени детерминированности их градаций (классов).

Поскольку в данном приложении одна классификационная шкала ("Winner" - Победитель матча), то рейтинг шкал по силе детерминированности состоит из одной шкалы и поэтому дополнительные таблицы здесь не приводятся.

Полученные результаты свидетельствуют о том, что прогнозирование победителей теннисных матчей на основе АСК-анализа

является перспективной задачей, но требует учета большого количества взаимосвязанных факторов различной природы для повышения точности прогнозов.

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 исследований в области спортивной аналитики, на исходных данных которых они основаны. С другой стороны, применение АСК-анализа и системы «Эйдос» весьма существенно расширяет возможности решения задач прогнозирования исходов теннисных матчей, принятия решений и исследования моделируемой предметной области по сравнению с традиционными статистическими методами и алгоритмами машинного обучения. Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований в области спортивной аналитики.

Достижением данной работы является:

Возможность корректного построения сопоставимых системно-когнитивных моделей предметной области на основе исходных данных, содержащих как лингвистические переменные (название турнира, имена игроков, тип покрытия), так и числовые переменные в различных единицах измерения (рейтинги, коэффициенты, даты).

Возможность применения системно-когнитивных моделей для решения задач прогнозирования победителя теннисного матча, поддержки принятия решений (например, при формировании аналитических прогнозов) и исследования моделируемой предметной области путем анализа выявленных причинно-следственных зависимостей.

В качестве перспективы продолжения исследований можно было бы рекомендовать существенно увеличить объем исходных данных (включить матчи за несколько сезонов), количество исследуемых факторов (например, статистику личных встреч, физическую форму игроков, погодные условия во время матча), а также количество классификационных шкал и их градаций для более детального описания возможных исходов (не только победитель, но и счет по сетам, длительность матча, количество эйсов и т.д.).

Например, можно было бы исследовать в создаваемых системно-когнитивных моделях не только индивидуальные характеристики игроков,

но и внешние факторы: тип турнира (Большой шлем, Мастерс, челленджер), географическое расположение, высоту над уровнем моря, влияние зрителей.

Рекомендуется ввести классификационные шкалы, отражающие влияние исследуемых факторов на исход матча не только в качественном выражении (победа/поражение), но и в количественном (разница в счете, вероятность победы в процентах, ожидаемое количество геймов), что позволит повысить точность и практическую ценность прогнозов.

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

У желающих есть все возможности для изучения данной работы и для дальнейших исследований с применением АСК-анализа и системы «Эйдос» на своем компьютере. Для этого надо скачать систему с сайта разработчика по ссылке на странице: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос-приложение №393. По различным аспектам применения данной технологии есть большое количество видеозанятий (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 растущей популярности тенниса и развития индустрии спортивных прогнозов актуальной задачей является повышение точности прогнозирования исходов матчей. Однако этому могут препятствовать множество взаимосвязанных факторов различной природы: рейтинг игроков, тип покрытия, историческая статистика, психологическое состояние и другие. В ряде научных работ обосновывается перспективность применения методов искусственного интеллекта для анализа эмпирических данных в спортивной аналитике. Для анализа таких данных традиционно используются статистические методы и алгоритмы машинного обучения. Эти исследования показали, что учет множества факторов позволяет повысить точность прогнозов, хотя различные подходы демонстрируют разную эффективность в зависимости от предметной области.

В данной статье для анализа эмпирических данных в области теннисной аналитики предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример на основе данных о теннисных матчах. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения

применению АСК-анализа и системы «Эйдос» для научных исследований в области спортивной аналитики, 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 (название турнира, имена игроков, тип покрытия), так и числовыми переменными, измеряемыми в различных единицах измерения (рейтинги, коэффициенты, даты). 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 о вероятности того или иного исхода матча.

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 спортивной аналитики [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 OOS CAB.
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHC.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZJXZZ.
4. Работы проф.Е.В.Луценко по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики http://lc.kubagro.ru/aidos/Work_on_emergence.htm
5. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.

6. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
7. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.
8. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
10. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-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) / 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 ампеелографии:
http://lc.kubagro.ru/aidos/Works_on_cognitive_agronomy.htm

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

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

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

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

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

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

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

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

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

49. Луценко Е.В. Системы искусственного интеллекта как системы автоматизации процесса научного познания и удвоение номенклатуры научных специальностей путем применения этих систем для исследований в различных направлениях науки / Е.В. Луценко, Н.С. Головин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2024. –

№01(195). С. 74 – 111. – IDA [article ID]: 1952401009. – Режим доступа:
<http://ej.kubagro.ru/2024/01/pdf/09.pdf>, 2,375 у.п.л.