

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»**

Факультет прикладной информатики

Кафедра Компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ факторов, влияющих на уровень зарплаты по
характеристикам вакансий»

Выполнил студент группы: ИТ2341 Сальков Алексей Алексеевич

Допущен к защите _____

Руководитель проекта д.э.н., к.т.н., профессор Луценко Е.В. (_____)

(подпись, расшифровка подписи)



Защищен _____

(дата)

Оценка отлично

Краснодар
2025

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ
УНИВЕРСИТЕТ ИМЕНИ И. Т. ТРУБИЛИНА»

Факультет прикладной информатики

**РЕЦЕНЗИЯ
на курсовую работу**

Студента Салькова Алексея Алексеевич курса
2 очной формы обучения группы ИТ2341
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ факторов, влияющих на уровень зарплаты
по характеристикам вакансий»

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)
1	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
2	Уровень и корректность использования в работе различных методов исследований	5
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
4	Ясность, четкость, последовательность и обоснованность изложения	5
5	Применение современных технологий обработки информации	5
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	5
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	5
8	Ответы на вопросы при защите	5

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите отлично

Рецензент  (Е. В. Луценко)

«06» февраля 2025 г.

РЕФЕРАТ

Курсовая работа содержит: 87 страниц, 39 рисунков, 19 таблиц, 20 литературных источников.

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

Целью работы является проведение автоматизированного системно-когнитивного анализа факторов, влияющих на уровень зарплаты по характеристикам вакансий.

Для достижения поставленной цели необходимо провести анализ методов формирования обобщенных моделей факторов, влияющих на уровень зарплаты, и решения задач классификации вакансий по их характеристикам. Это включает идентификацию конкретных зависимостей между параметрами вакансий и уровнями заработной платы, а также исследование моделируемой предметной области рынка труда через анализ и интерпретацию полученной модели с использованием выбранных инструментальных средств.

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА УРОВЕНЬ ЗАРПЛАТЫ ПО ХАРАКТЕРИСТИКАМ ВАКАНСИЙ

Сальков Алексей Алексеевич
студент факультета ПИ, группы
ИТ2341
salkow12345@gmail.com

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF FACTORS AFFECTING SALARY LEVEL BASED ON JOB CHARACTERISTICS

Salkov Alexey Alekseevich
student of the faculty of PI, group
IT2341
salkow12345@gmail.com

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

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

В условиях современного рынка труда важно понимать факторы, влияющие на уровень заработной платы, чтобы оптимизировать процессы найма и повысить конкурентоспособность работодателей. Однако на уровень зарплат могут влиять различные характеристики вакансий, такие как требуемый опыт, уровень квалификации, регион и отрасль. В данной статье предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос» для анализа эмпирических данных, связанных с уровнем зарплат. Приводится подробный численный пример, который включает разнообразные табличные и графические формы представления данных. Этот пример может быть использован для обучения применению АСК-анализа и системы «Эйдос» в исследованиях рынка труда, разработки практических рекомендаций и обоснования научных выводов о механизмах влияния различных факторов на уровень заработной платы.

Ключевые слова: АСК-АНАЛИЗ, АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «ЭЙДОС», УРОВЕНЬ ЗАРПЛАТЫ, ХАРАКТЕРИСТИКИ ВАКАНСИЙ

In the context of the modern labor market, it is important to understand the factors affecting salary levels to optimize hiring processes and increase the competitiveness of employers. However, salary levels can be influenced by various job characteristics, such as required experience, skill level, region, and industry. This article proposes the use of Automated System-Cognitive Analysis (ASC-analysis) and its software tool, the intelligent Eidos system, to analyze empirical data related to salary levels. A detailed numerical example is provided, including various tabular and graphical data representations. This example can be used to teach the application of ASC-analysis and the Eidos system in labor market research, develop practical recommendations, and substantiate scientific conclusions about the mechanisms of influence of various factors on salary levels.

Keywords: ASC-ANALYSIS, AUTOMATED SYSTEM-COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS", SALARY LEVELS, JOB CHARACTERISTICS

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	6
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	6
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	6
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ	6
1.4. ЦЕЛЬ РАБОТЫ.....	7
2. METHODS (МЕТОДЫ)	7
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	7
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	8
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	8
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	10
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	16
3. RESULTS (РЕЗУЛЬТАТЫ)	19
3.1. ЗАДАЧА-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	19
3.2. ЗАДАЧА-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ.....	20
3.3. ЗАДАЧА-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ.....	30
3.4. ЗАДАЧА-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	40
3.5. ЗАДАЧА-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	44
3.6. ЗАДАЧА-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ.....	44
3.6.1. Интегральный критерий «Сумма знаний».....	45
3.6.2. Интегральный критерий «Семантический резонанс знаний».....	46
3.6.3. Важные математические свойства интегральных критериев.....	47
3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос».....	48
3.7. ЗАДАЧА-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ	51
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ	51
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос».....	53
3.8. ЗАДАЧА-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ ...	57
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)	57
3.8.2. Кластерно-конструктивный анализ классов	58
3.8.3. Кластерно-конструктивный анализ значений описательных шкал.....	60
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны.....	63
3.8.5. Нелокальная нейронная сеть	65
3.8.6. 3D-интегральные когнитивные карты	67
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	68
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	71
3.8.9. Когнитивные функции	73
3.8.10. Значимость описательных шкал и их градаций	75
3.8.11. Степень детерминированности классов и классификационных шкал.....	78
4. DISCUSSION (ОБСУЖДЕНИЕ)	80
5. CONCLUSIONS (ВЫВОДЫ)	81
REFERENCES (ЛИТЕРАТУРА)	82

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Уровень заработной платы на рынке труда зависит от множества факторов, включая характеристики вакансий, требования к кандидатам и условия работы. Эффективное определение зарплаты требует учета различных социальных и экономических аспектов, таких как квалификация, опыт, обязанности, а также расположение вакансии и особенности отрасли. Одним из методов анализа этих факторов является АСК-анализ, который помогает выявить связь между характеристиками вакансий и уровнем заработной платы, а также определить, какие из этих факторов наиболее значимы для работодателей и работников.

Цель данного исследования заключается в анализе факторов, влияющих на уровень заработной платы, с фокусом на характеристиках вакансий. Для этого будут собраны и систематизированы данные о вакансиях, включая информацию о должностных обязанностях, требованиях к квалификации, опыте, а также условиях труда. На основе этих данных будет проведен анализ взаимосвязей между параметрами вакансий и уровнем зарплаты. В результате исследования будут разработаны рекомендации для работодателей, направленные на оптимизацию процесса установления зарплаты, что позволит более эффективно привлекать и удерживать квалифицированных специалистов.

1.2. Объект и предмет исследования

Объектом исследования является рынок труда, а именно вакансии, предлагаемые работодателями в различных отраслях экономики, и факторы, влияющие на определение уровня заработной платы для этих вакансий. Рынок труда представляет собой совокупность всех предложений по трудовой занятости, включая требования к кандидатам, условия работы и размер вознаграждения.

Предметом исследования являются конкретные характеристики вакансий, такие как квалификационные требования, опыт работы, должностные обязанности, географическое расположение, а также социальные и экономические факторы, влияющие на установление уровня заработной платы. Основное внимание уделяется анализу того, как эти параметры соотносятся с величиной оплаты труда и как различные факторы могут быть использованы для оптимизации процесса формирования заработных плат в зависимости от типа вакансии и условий рынка труда.

1.3. Проблема, решаемая в работе и ее актуальность

Современный рынок труда характеризуется высокой конкуренцией и быстрыми изменениями в экономической и социальной сферах, что

оказывает влияние на установление уровней заработной платы. Одна из ключевых проблем заключается в сложности определения объективных критериев, которые должны быть использованы для справедливого и конкурентоспособного определения зарплаты на различных вакансиях. Многие работодатели сталкиваются с трудностями при установлении оптимальной величины заработной платы, что может привести к недооценке или переоценке труда сотрудников, снижению привлекательности вакансий и, как следствие, недостаточной эффективности привлечения и удержания квалифицированных кадров.

Актуальность данного исследования обусловлена необходимостью разработки эффективных методов оценки и установления уровня заработной платы на основе характеристик вакансий. В условиях динамично меняющегося рынка труда важно иметь точные и обоснованные данные о влиянии различных факторов на размер вознаграждения, чтобы работодатели могли правильно реагировать на запросы соискателей, а также справедливо и конкурентоспособно формировать зарплаты, что способствует улучшению кадровой политики и повышению производительности труда.

Таким образом, решение данной проблемы поможет компаниям и организациям разработать более точные подходы к установлению уровня оплаты труда, что повысит эффективность управления персоналом и конкурентоспособность на рынке труда.

1.4. Цель работы

Целью работы является анализ факторов, влияющих на уровень заработной платы по характеристикам вакансий, с целью разработки рекомендаций для оптимизации процесса установления зарплаты.

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»).

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)

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 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].

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/>

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.

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/ZCKZRKvrlEMBWOк8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://poe.com/Aidos-X>, <https://rudalle.ru/>, <https://bard.google.com/>, <https://chatbot.theb.ai/>, <https://problembo.com/ru/services> (может пригодиться - почта на 10 минут: <https://10minutemail.net/>), <https://poe.com/GPT-3.5-Turbo-Instruct>, <https://www.seaart.ai/home>, <https://ui.chatai.com/>.

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

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

идентификации, прогнозирования, принятия решений и исследования предметной области путем исследования ее модели (автоматические системы работают без такого участия человека);

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

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

- имеет «нулевой порог входа», содержит большое количество интеллектуальных локальных (т.е. поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 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);

- 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 видеogramма финальной DOS-версии системы «Эйдос-12.5», июнь 2012 года). С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

3-й этап, «эра MS Windows xp, 8, 7»: 2012-2020 годы. С июня 2012 по 14.12.2020 система «Эйдос» развивалась на языке [Аляска-1.9](#) + [Экспресс++](#) + библиотека для работы с Internet xb2net. Система «Эйдос-X1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

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

5-й этап, «эра Больших данных, информации и знаний»: 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-приложения](#)⁷

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



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

⁷ <http://lc.kubagro.ru/aidos/How to make your own cloud Eidos-application.pdf>

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

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Главное, что делает система:

- Альберт Эйнштейн писал, что научные законы это лишь высказывания о повторениях в наблюдаемых явлениях. Конечно наверное он имел в виду не сами законы природы, а лишь формулировки этих законов. В системе "Эйдос" эти наблюдения повторений называются событиями или фактами.
- Например, фактом является наблюдение определенного значения какого-либо свойства у объектов некоторой обобщенной категории (класса), или наблюдение определенного значения фактора при переходе объекта в будущее состояние, соответствующее к классу.
- Система "Эйдос" выявляет эмпирические закономерности в фактах и тем самым преобразует исходные данные в информацию, а ее в знания и решает на основе этих знаний задачи идентификации, прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования ее модели.
- Кроме того система "Эйдос" выводит информацию об обнаруженных закономерностях в большом количестве разнообразных и оригинальных текстовых, табличных и графических выходных форм.

Работы автора системы "Эйдос" проф.Е.В.Луценко, С" по АСК-анализу и системе "Эйдос".

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Главное, что делает система:

Персональная интеллектуальная онлайн среда "ЭЙДОС-X Professional" (Система "Эйдос-Xpro").

- ПРЕДНАЗНАЧЕНА для обучения и научных исследований в области искусственного интеллекта с применением автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария - интеллектуальной системы "Эйдос-Xpro".
- ОБЕСПЕЧИВАЕТ преобразование больших данных (Big Data) в большую информацию (Big Information), а ее в большие знания (Big Knowledge) с использованием ADS (Advantage Database Server) и решение на основе этих знаний задач обобщения, абстрагирования, идентификации (классификации), распознавания, диагностики, прогнозирования, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели.
- ПОЗВОЛЯЕТ пользователям и разработчикам интеллектуальных облачных Эйдос-приложений во всем мире обмениваться опытом решения различных задач учебного и научного характера с применением технологий искусственного интеллекта на платформе "Эйдос-Xpro".

Работы автора системы "Эйдос" проф.Е.В.Луценко, С" по АСК-анализу и системе "Эйдос".

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Объявление о получении магистерского образования по искусственному интеллекту в КубГУ:

- В связи с высокой востребованностью на рынке труда специалистов в области цифровой экономики Кубанскому государственному университету оказано доверие и увеличено число бюджетных мест в магистратуру по приоритетным ИТ направлениям до 75. Приглашаем получать высококвалифицированное подготовку по актуальным ИТ специальностям.

СПИСОК направлений подготовки магистратуры (очная и заочная формы обучения):

- 09.04.02 Информационные системы и технологии (Искусственный интеллект и машинное обучение); 09.04.02 Информационные системы и технологии (Искусственный интеллект и машинное обучение); 01.04.02 Прикладная математика и информатика (Матем. и инф. обеспечение экономики); 01.04.02 Прикладная математика и информатика (Матем. модел. в естественных и технических); 01.04.02 Прикладная математика и информатика (Технологии программ. и разрач. инф. систем); 02.04.02 Фунд. информатика и информ. технологии (Интеллектуальные системы и технологии).

3. КОНТАКТЫ: +79189800003, zavanna-05@mail.ru, докт. техн. наук. Анна Владимировна Коваленко

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Announcement of own fonts of the Eidos system:

When trying to download your own fonts of the Eidos system from the folder: c:\Eidos\VAID_DATA\Fonts\ it is found that they are missing!!!

To fix the situation, you need to download the font update file <http://ic.kubagro.ru/Fonts.exe> from the developer's website and deploy updates in the system folder: c:\Eidos\X\ with the replacement of all files, and then run the system as usual.

If MS Windows is Russified, then you do not need to do all this, because everything will be work fine with standard MS Windows fonts.

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Приглашение к размещению интеллектуальных облачных Эйдос-приложений:

Уважаемые пользователи системы "Эйдос" во всем мире: [http://ic.kubagro.ru/map5.php!](http://ic.kubagro.ru/map5.php)

Приглашаю размещать свои интеллектуальные облачные Эйдос-приложения. Это делается в диспетчере приложений (Режим 1.3). Для разработки приложения рекомендуется ознакомиться с инструкцией для учащихся по адресу: http://ic.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf

Прошу вас отнестись ОТВЕТСТВЕННО к качеству разработки приложения и его описанию и разрешать только разработанные вами лично приложения, описанные в соответствии со стандартом IMRAD. Это общепринятый в мире стандарт изложения научных результатов, принятый в наукометрических базах Scopus и Web of science (WoS). Описание, т.е. его текст и все выходные формы и скриншоты, должно ПОЛНОСТЬЮ соответствовать модели, полученной в системе "Эйдос" на приведенных исходных данных при выполнении всех пунктов этого описания. Примеры подобных описаний интеллектуальных облачных Эйдос-приложений приведены в ряде работ автора и разработчика системы "Эйдос" проф.Е.В.Луценко, например по ссылке: <https://www.researchgate.net/publication/362211691>.

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

(c) Авторизация в системе ЭЙДОС-X++

Задайте имя и пароль:

Login :

Password :

Advantage Database Server (ADS) - OFF

Особенности работы в системе:

- Если система в данной папке запускается впервые, то будет произведена ЛОКАЛИЗАЦИЯ системы, т.е. будут удалены все приложения и пользователи и заново прописаны пути на все базы данных по фактическому расположению системы.
- Новое окно главного меню можно открывать только после закрытия всех предыдущих.

Пояснение о некорректном запуске системы "Эйдос"

Студенты и некоторые другие пользователи иногда запускают систему "Эйдос" некорректно: - в папке загрузки или на рабочем столе; - в архиве инсталляции системы "Эйдос", который скачали с сайта разработчика; - в папке, в пути на которую встречается проблема и кириллица.

Кроме того иногда систему запускают в одной и той же папке несколько раз, чего делать нельзя (ее можно запускать несколько раз одновременно на одном компьютере, но в разных папках). Некорректный запуск системы "Эйдос" вызывает ошибку исполнения. Поэтому приходится проверять корректность запуска системы "Эйдос". Однако эта проверка занимает довольно много времени. Поэтому она оставлена только в модуле запуска системы: "___START_AIDOSX.exe", а в исполнимом модуле самой системы "___AIDOSX.exe" она включается/отключается в зависимости от содержания текстового файла: "Checking_the_correctness_of_the_module_launch_AIDOSX.txt"; "DN"/"OFF". Отметим, что модуль запуска системы: "___START_AIDOSX.exe" кроме проверки корректности запуска системы еще проверяет целостность исполнимого модуля системы "___AIDOSX.exe" и наличие обновлений на сайте автора и разработчика системы проф.Е.В.Луценко. Если обновления есть, то они скачиваются, разархивируются и устанавливаются автоматически.

Кратко об АСК-анализе Подборки публикаций по АСК-анализу

Скачать все публикации проф.Е.В.Луценко из Научного журнала КубГАУ (> 2 Гб)

СЕРТИФИКАТ об освоении системы "Эйдос" от проф.Е.В.Луценко. ПОДДЕРЖКА

Ok Cancel

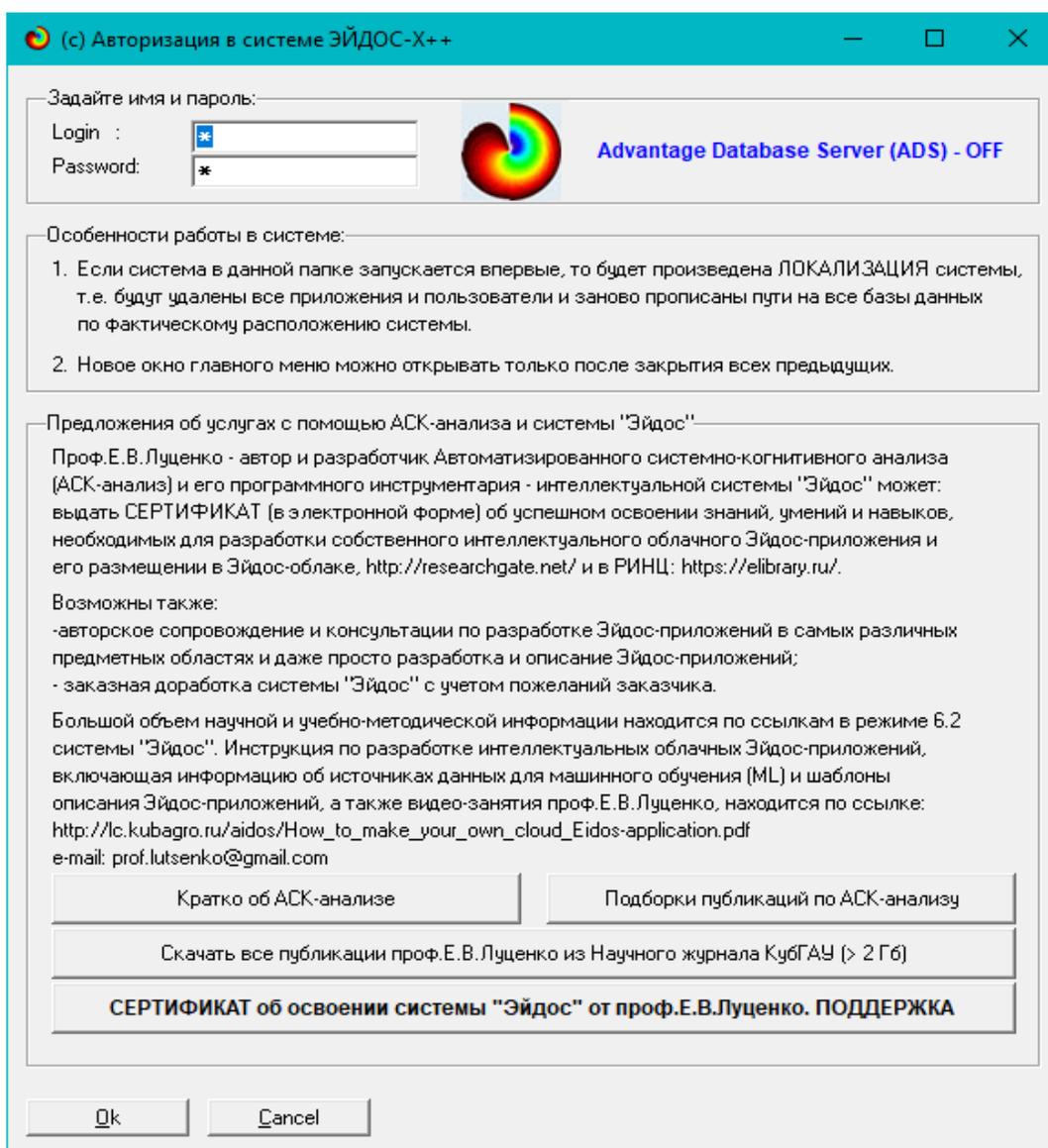


Рисунок 2. Титульные видеogramмы текущей версии системы «Эйдос»

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, связанной с анализом факторов, влияющих на уровень заработной платы, так как это позволяет детально исследовать влияние различных характеристик вакансий на установление размера вознаграждения, включая роль квалификации, опыта и обязанностей сотрудников.

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**

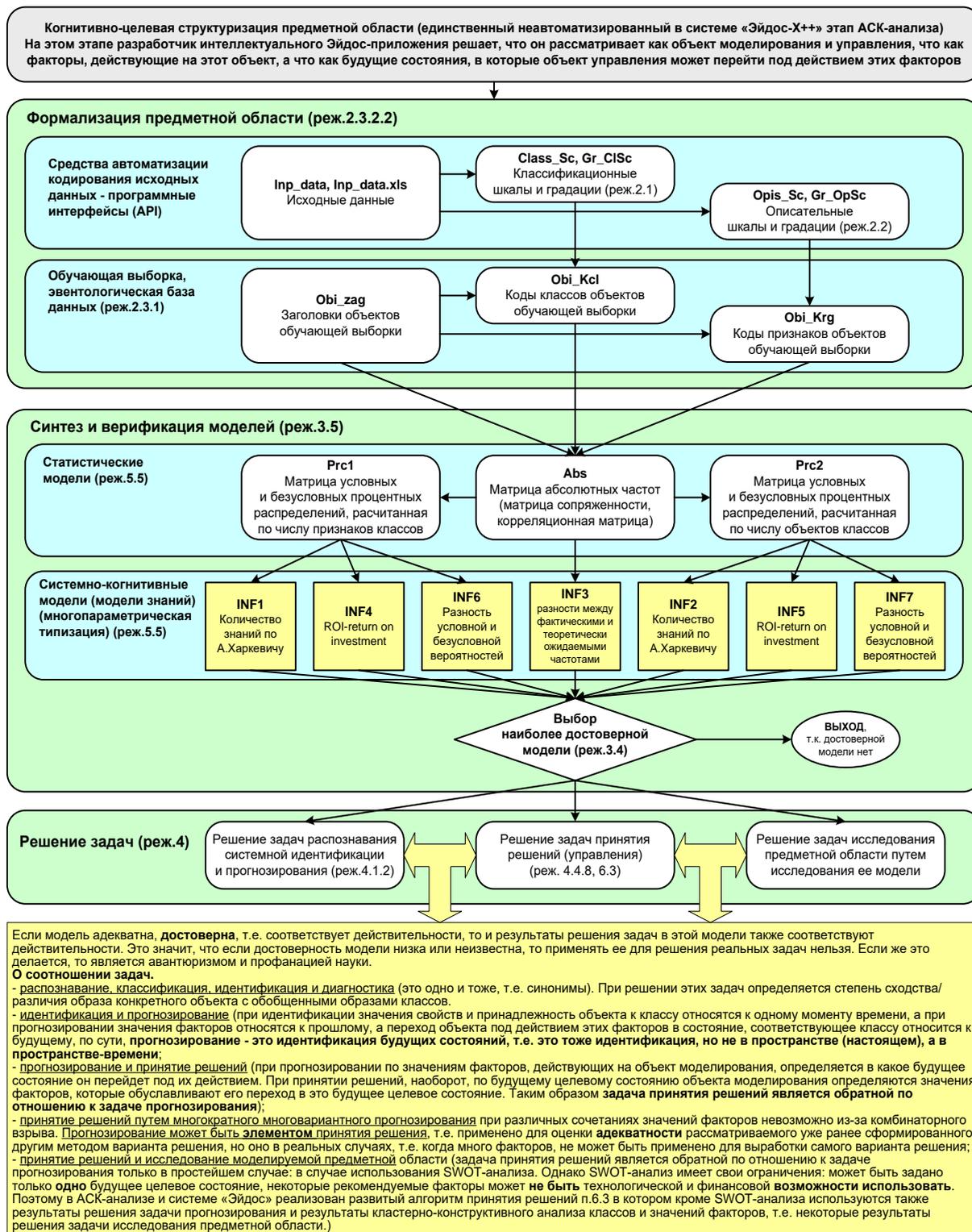


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос» (этапы АСК-анализа)

3. RESULTS (РЕЗУЛЬТАТЫ)

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

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

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

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

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

Статичная интерпретация и терминология:

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

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

Динамичная интерпретация и терминология:

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

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

Обобщающая терминология:

– классификационные шкалы и градации;

– описательные шкалы и градации.

В данной работе в качестве *объекта моделирования* выступает уровень заработной платы, в качестве *факторов*: образование, опыт работы, место расположения, должность, возраст и пол (таблица 1). В качестве результата действия этих факторов рассматривается размер зарплаты (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	Education
2	Experience
3	Location
4	Job_Title
5	Age
6	Gender

Источник: c:\Aidos-X\AID_DATA\A0000027\System\Opis_Sc.xlsx

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	Salary

Источник: c:\Aidos-X\AID_DATA\A0000027\System\Class_Sc.xlsx

3.2. Задача-2. Формализация предметной области

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

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

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

В качестве источника исходных данных в данной работе используем Excel-таблицы (см. таблицы 3 и 4):

Таблица 3 – Исходные данные по влиянию различных факторов, влияющих на уровень зарплаты по характеристикам вакансий

№	Education	Experience	Location	Job_Title	Age	Gender	Salary
1	High School	8	Urban	Manager	63	Male	84620
2	PhD	11	Suburban	Director	59	Male	142591
3	Bachelor	28	Suburban	Manager	61	Female	97800
4	High School	29	Rural	Director	45	Male	96835
5	PhD	25	Urban	Analyst	26	Female	132158
6	PhD	19	Rural	Director	27	Female	156313
7	PhD	4	Rural	Director	60	Female	130568
8	PhD	13	Suburban	Director	49	Female	148708
9	Bachelor	20	Urban	Engineer	25	Female	95945
10	PhD	14	Urban	Analyst	58	Female	133339
11	Bachelor	23	Rural	Manager	23	Female	101164
12	Master	1	Urban	Director	27	Female	124252
13	High School	11	Urban	Manager	63	Male	55066
14	PhD	4	Rural	Manager	43	Male	124973
15	Master	7	Suburban	Director	44	Female	126139
16	High School	28	Suburban	Engineer	49	Male	81007
17	High School	4	Suburban	Analyst	37	Male	60694
18	High School	26	Rural	Director	63	Male	103386
19	Master	7	Urban	Director	53	Female	140042
20	Bachelor	20	Rural	Analyst	34	Male	86048
21	Master	3	Rural	Engineer	62	Female	73862
22	PhD	19	Suburban	Analyst	62	Male	106145
23	PhD	19	Urban	Director	36	Female	142982
24	Master	7	Suburban	Analyst	21	Female	96621
25	High School	21	Urban	Manager	62	Female	89427
26	Bachelor	12	Urban	Manager	20	Male	104762
27	Bachelor	23	Suburban	Manager	35	Female	101748
28	Bachelor	1	Suburban	Engineer	36	Male	68635

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

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

№	Education	Experience	Location	Job_Title	Age	Gender	Salary
1	High School	8	Urban	Manager	63	Male	84620
2	PhD	11	Suburban	Director	59	Male	142591
3	Bachelor	28	Suburban	Manager	61	Female	97800
4	High School	29	Rural	Director	45	Male	96835
5	PhD	25	Urban	Analyst	26	Female	132158
6	PhD	19	Rural	Director	27	Female	156313
7	PhD	4	Rural	Director	60	Female	130568
8	PhD	13	Suburban	Director	49	Female	148708
9	Bachelor	20	Urban	Engineer	25	Female	95945
10	PhD	14	Urban	Analyst	58	Female	133339
11	Bachelor	23	Rural	Manager	23	Female	101164
12	Master	1	Urban	Director	27	Female	124252
13	High School	11	Urban	Manager	63	Male	55066
14	PhD	4	Rural	Manager	43	Male	124973
15	Master	7	Suburban	Director	44	Female	126139
16	High School	28	Suburban	Engineer	49	Male	81007
17	High School	4	Suburban	Analyst	37	Male	60694
18	High School	26	Rural	Director	63	Male	103386
19	Master	7	Urban	Director	53	Female	140042
20	Bachelor	20	Rural	Analyst	34	Male	86048
21	Master	3	Rural	Engineer	62	Female	73862
22	PhD	19	Suburban	Analyst	62	Male	106145
23	PhD	19	Urban	Director	36	Female	142982
24	Master	7	Suburban	Analyst	21	Female	96621
25	High School	21	Urban	Manager	62	Female	89427
26	Bachelor	12	Urban	Manager	20	Male	104762
27	Bachelor	23	Suburban	Manager	35	Female	101748
28	Bachelor	1	Suburban	Engineer	36	Male	68635

Таблица 4 имеет следующую структуру:

- каждая строка описывает одного работника;
- каждое наблюдение описывается одновременно двумя способами: с одной стороны значениями факторов, действующих на объект моделирования (лингвистические и числовые переменные, градации описательных шкал, бесцветный фон в таблице 4), а с другой стороны результатами действия этих факторов, т.е. зарплата, выраженная в числовых шкалах (желтый фон). Такая структура описания наблюдений в технологиях искусственного интеллекта называется «онтологией» и

модели представлений знаний Марвина Мински (1975) называется «фрейм-экземпляр»;

- 1-я колонка – не является шкалой и содержит порядковый номер;

- колонка 8 – это классификационные шкалы – это шкалы **текстового** и **числового** типа описывающие **результаты** действия факторов в различных единицах измерения (таблица 1), в данном случае образование, опыт, местоположение, должность, возраст, пол. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в натуральном и стоимостном выражении [8]: например, возраст и опыт человека, зарплата. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

- колонки с 2-й по 7-ю – это описательные шкалы, формализующие факторы, действующие на объект моделирования (таблица 4). Эти шкалы имеют числовой и текстовый тип и их градациями являются лингвистические и числовые переменные;

- при вводе данных в систему «Эйдос» нули и пробелы в исходных данных могут рассматриваться как значащие или как отсутствие данных. 1-й вариант и будет использован в данной работе.

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

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

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

2.3.2. Программные интерфейсы с внешними базами данных

- 2.3.2.1. Импорт данных из текстовых файлов
- 2.3.2.2. Универсальный программный интерфейс импорта данных в систему
- 2.3.2.3. Импорт данных из транспонированных внешних баз данных
- 2.3.2.4. Оцифровка изображений по внешним контурам
- 2.3.2.5. Оцифровка изображений по всем пикселям и спектру
- 2.3.2.6. Сценарный АСК-анализ символьных и числовых рядов

- 2.3.2.7. Транспонирование файлов исходных данных
- 2.3.2.8. Объединение нескольких файлов исходных данных в один
- 2.3.2.9. Разбиение TXT-файла на файлы-абзацы
- 2.3.2.10. CSV => DBF конвертер системы "Эйдос"

- 2.3.2.11. Прогноз событий по астропараметрам по Н.А.Чередниченко
- 2.3.2.12. Прогнозирование землетрясений методом Н.А.Чередниченко
- 2.3.2.13. Чемпионат RAIF-Challenge 2017-API-bank
- 2.3.2.14. Чемпионат RAIF-Challenge 2017-API-retail

- 2.3.2.15. Вставка промежуточных строк в файл исходных данных Inp_data

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

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

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

Помощь по режиму 2.3.2.2 для случая Excel-файлов исходных данных

Регистр 2.3.2.2. Универсальный программный интерфейс импорта данных из внешней базы данных "Inp_data.xls" в систему "Эйдос" и формализация прачной области.

Данный программный интерфейс обеспечивает формализацию прачной области, т.е. анализ файла исходных данных (Inp_data.xls), формирование классификационных и описательных шкал и графидов, а затем и оцифровку файла исходных данных с их использованием.

*Файл исходных данных должен иметь имя Inp_data.xls, а файл распознающей выборки имя Inp_pars.xls(Файлы Inp_data.xls и Inp_pars.xls должны находиться в папке J:\EIDOS\RAW_DATA\Inp_data. Эти файлы имеют совершенно одинаковую структуру.

- 1-я строка этого файла должна содержать наименования колонок на латинском языке, в т.ч. и русским. Эти наименования должны быть во всех колонках, при этом переводы по словам разрешены, а объединение имен, разрыв строки знак абзаца не допускается. Эти наименования должны быть короткими, но понятными, т.е. они будут в выводе в форме, а не как код. Не добавляйте наименования графидов. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения и число знаков после запятой в колонке должно быть ODMAXSIGOES.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длиной до 255 символов.
- Каждая строка этого файла, начиная со 2-ой, содержит данные об одном объекте обучающей выборки или одном наблюдении. В MS Excel 2003 в акте может быть до 65536 строк и до 256 колонок. В акте MS Excel 2010 и более подача колонок до 1048576 строк и 16384 колонок.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (именованного / порядкового) или числового типа (с десятичными знаками после запятой).
- Столбцы представляются числовой тип, если все значения его имен числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбец представляется текстовым тип. Это означает, что нули должны быть удалены нулями, а не пробелами.
- Столбцы со 2-го по N-й являются классификационными шкалами (выделяем и параметризм) и содержат данные о классах (Будем считать состояние объекта управляемым), к которым принадлежат объекты обучающей выборки.
- Столбцы с (N+1) по последние являются описательными или факторными (содержат данные о признаках (т.е. значениях свойств или значениях факторов), характеризующих объекты обучающей выборки.

В результате работы режима формируется файл Inp_NAME.TXT стандарта MS DDS (интерпретация), в котором наименования классификационных и описательных шкал являются СПРАЖИМЫ. Система формирует классификационные и описательные шкалы и графиды. Для этого в каждом числовом столбце система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего столбцы значения заменяются на интервальные значения. В текстовых столбцах система находит уникальные текстовые значения. Кладов ШИКАРДЮС интерпретация числового или текстового значения считается графидом классификационной или описательной шкалы, характеризующий объект. В каждой шкале не графиды сортируются по алфавиту. С использованием шкал и графидов кодируются исходные данные в результате чего генерируется обучающая выборка, каждый объект которой соответствует одной строке файла исходных данных. NP_DATA и содержит коды классов, соответствующие фактам сопоставления числовых или уникальных текстовых значений классов с графидными классификационными шкалами и коды признаков, соответствующие фактам сопоставления числовых или уникальных текстовых значений признаков с графидными описательными шкалами. Распознающая выборка формируется на основе файла Inp_PARS аналогично, за исключением того, что классификационные и описательные шкалы и графиды не создаются, и используются ранее созданные в файле Inp_NAME.TXT, а файлы распознающей выборки могут не включать коды классов, если столбцы классов в файле Inp_PARS были пустыми. Структура файла Inp_PARS должна быть такой же, как Inp_DATA, т.е. описатели ПОЛНОСТЬЮ совпадают по наименованиям столбцов, но могут иметь разное количество строк с разными значениями в них.

Принцип организации таблицы исходных данных:

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

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

Принцип организации таблицы исходных данных:

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

Рисунок 5. 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-Х++"

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)
 DBF - DBASE IV (DBF/NTX) Стандарт DBF-файла
 CSV - CSV => DBF конвертер Стандарт CSV-файла

Задайте параметры:

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

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:
 Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:
 Конечный столбец описательных шкал:

Задайте режим:

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

Задайте способ выбора размера интервалов:

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

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

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

Ok Cancel

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

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [2 классов x 23 признаков]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	1	2	2,00	2	10	5,00
Текстовые	0	0	0,00	4	13	3,25
ВСЕГО:	1	2	2,00	6	23	3,83

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

В классификационных шкалах: В описательных шкалах:

Пересчитать шкалы и градации Выйти на создание модели

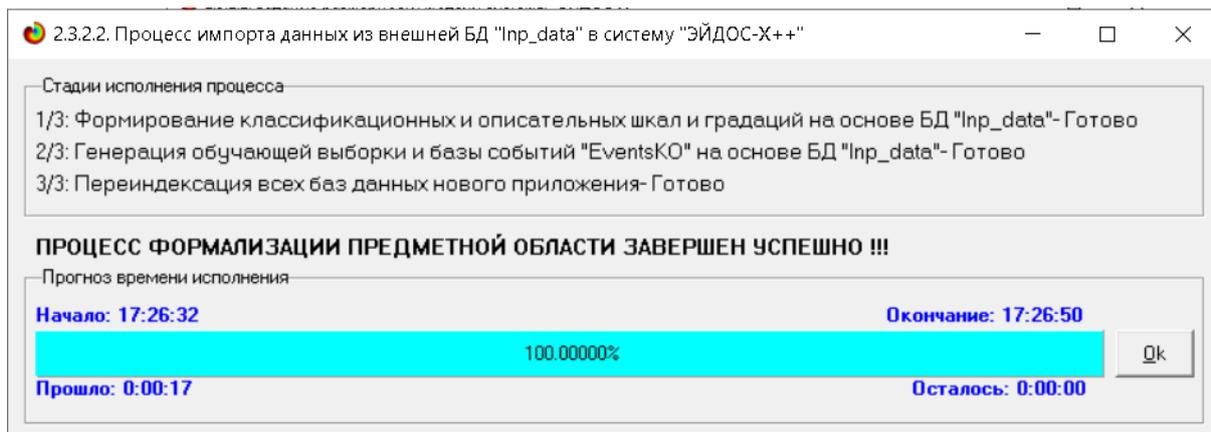


Рисунок 6. Экранные формы управления API-2.3.2.2 системы «Эйдос»

Таблица 5 – Классификационные шкалы и градации

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы
1	SALARY	1	Малое
		2	Большое

Таблица 6 – Описательные шкалы и градации

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	EDUCATION	17	Очень малое
2	EXPERIENCE	18	Малое
3	LOCATION	19	Среднее
4	JOB_TITLE	20	Большое
5	AGE	21	Очень большое
6	GENDER		

Таблица 7 – Обучающая выборка

№	Education	Experience	Location	Job_Title	Age	Gender	Salary
1	High School	8	Urban	Manager	63	Male	84620
2	PhD	11	Suburban	Director	59	Male	142591
3	Bachelor	28	Suburban	Manager	61	Female	97800
4	High School	29	Rural	Director	45	Male	96835
5	PhD	25	Urban	Analyst	26	Female	132158
6	PhD	19	Rural	Director	27	Female	156313
7	PhD	4	Rural	Director	60	Female	130568
8	PhD	13	Suburban	Director	49	Female	148708
9	Bachelor	20	Urban	Engineer	25	Female	95945
10	PhD	14	Urban	Analyst	58	Female	133339
11	Bachelor	23	Rural	Manager	23	Female	101164
12	Master	1	Urban	Director	27	Female	124252
13	High School	11	Urban	Manager	63	Male	55066
14	PhD	4	Rural	Manager	43	Male	124973
15	Master	7	Suburban	Director	44	Female	126139
16	High School	28	Suburban	Engineer	49	Male	81007
17	High School	4	Suburban	Analyst	37	Male	60694

18	High School	26	Rural	Director	63	Male	103386
19	Master	7	Urban	Director	53	Female	140042
20	Bachelor	20	Rural	Analyst	34	Male	86048
21	Master	3	Rural	Engineer	62	Female	73862
22	PhD	19	Suburban	Analyst	62	Male	106145
23	PhD	19	Urban	Director	36	Female	142982
24	Master	7	Suburban	Analyst	21	Female	96621
25	High School	21	Urban	Manager	62	Female	89427
26	Bachelor	12	Urban	Manager	20	Male	104762
27	Bachelor	23	Suburban	Manager	35	Female	101748
28	Bachelor	1	Suburban	Engineer	36	Male	68635
29	Bachelor	5	Rural	Manager	28	Male	90550
30	High School	14	Urban	Analyst	61	Female	75559
31	Bachelor	10	Urban	Engineer	40	Female	87433
32	High School	3	Urban	Analyst	63	Female	57615
33	PhD	2	Urban	Engineer	45	Male	126043
34	High School	13	Urban	Analyst	20	Male	73043
35	PhD	16	Rural	Director	59	Male	134560
36	Bachelor	23	Urban	Engineer	22	Female	95804
37	Master	2	Rural	Director	26	Female	99695
38	PhD	24	Suburban	Director	23	Female	153931
39	PhD	13	Rural	Director	50	Male	151123
40	High School	6	Rural	Director	33	Female	81463
41	Master	7	Rural	Director	40	Male	106019
42	PhD	2	Suburban	Director	31	Female	142123
43	High School	11	Urban	Analyst	31	Male	66858
44	Bachelor	7	Rural	Manager	25	Female	82061
45	PhD	7	Suburban	Engineer	36	Female	107935
46	Bachelor	7	Rural	Analyst	33	Male	70252
47	PhD	7	Rural	Analyst	28	Female	118799
48	PhD	12	Suburban	Analyst	47	Female	100369
49	Master	19	Rural	Analyst	26	Male	100039
50	PhD	22	Rural	Manager	64	Male	164543
51	High School	19	Urban	Manager	62	Male	95726
52	Bachelor	19	Suburban	Engineer	24	Male	96598
53	Bachelor	20	Rural	Engineer	60	Female	82628
54	Bachelor	3	Suburban	Analyst	61	Female	66217
55	PhD	13	Suburban	Engineer	35	Female	127167
56	High School	10	Urban	Director	20	Female	106029
57	PhD	20	Urban	Analyst	59	Male	139656
58	Master	9	Urban	Engineer	57	Male	100278
59	High School	6	Rural	Engineer	26	Female	67768
60	PhD	21	Rural	Director	49	Male	148074
61	PhD	22	Rural	Analyst	45	Female	143973
62	Master	12	Rural	Manager	32	Female	106334

63	PhD	2	Rural	Analyst	20	Female	96967
64	Master	12	Rural	Manager	48	Female	113746
65	PhD	16	Suburban	Engineer	27	Female	140823
66	High School	13	Suburban	Manager	21	Female	88374
67	Master	16	Urban	Analyst	64	Female	112592
68	High School	13	Rural	Engineer	27	Male	53119
69	High School	22	Suburban	Manager	60	Female	81947
70	High School	25	Urban	Manager	46	Female	74817
71	Bachelor	19	Urban	Analyst	60	Male	84869
72	Bachelor	26	Urban	Engineer	42	Male	104215
73	Master	14	Urban	Engineer	34	Male	117890
74	High School	29	Suburban	Manager	51	Male	90605
75	High School	13	Suburban	Manager	41	Female	77430
76	Bachelor	10	Suburban	Manager	27	Male	91038
77	PhD	22	Urban	Director	56	Male	169510
78	High School	7	Urban	Engineer	36	Male	58009
79	Bachelor	5	Rural	Manager	26	Female	93375
80	Master	17	Rural	Manager	62	Female	136609
81	Master	9	Urban	Engineer	40	Female	113728
82	PhD	27	Suburban	Director	58	Male	156729
83	High School	6	Urban	Engineer	24	Female	60637
84	Bachelor	16	Rural	Director	36	Female	118098
85	Bachelor	6	Rural	Analyst	54	Male	59734
86	PhD	6	Suburban	Engineer	34	Male	111462
87	Bachelor	28	Urban	Engineer	59	Male	96356
88	Bachelor	18	Rural	Analyst	35	Male	70215
89	PhD	1	Suburban	Engineer	44	Male	105552
90	Master	17	Rural	Analyst	34	Male	109899
91	PhD	17	Suburban	Director	30	Male	163023
92	PhD	1	Rural	Engineer	24	Male	104029
93	Master	11	Rural	Analyst	40	Female	98073
94	Master	13	Urban	Manager	61	Female	130592
95	PhD	1	Suburban	Director	60	Male	137006
96	High School	27	Suburban	Director	31	Male	92451
97	Master	8	Rural	Manager	49	Female	93296
98	PhD	23	Urban	Manager	23	Male	163104
99	Bachelor	20	Rural	Analyst	64	Male	89962
100	High School	6	Rural	Manager	46	Male	71770
101	Bachelor	5	Rural	Engineer	47	Male	63924
102	Master	20	Suburban	Analyst	22	Female	112350
103	High School	5	Urban	Director	22	Male	91290
104	PhD	6	Rural	Engineer	43	Female	119140
105	High School	29	Rural	Engineer	48	Male	71550
106	Master	13	Rural	Director	38	Female	133123
107	High School	27	Suburban	Engineer	21	Female	66724

108	PhD	18	Rural	Director	24	Male	157287
109	PhD	20	Suburban	Director	56	Female	165811
110	High School	13	Suburban	Director	37	Female	103723
111	PhD	7	Suburban	Director	44	Female	138575
112	High School	8	Suburban	Analyst	45	Female	48006
113	High School	21	Suburban	Engineer	63	Female	83939
114	High School	17	Urban	Manager	32	Male	80008
115	High School	1	Suburban	Analyst	54	Male	51440
116	Master	24	Suburban	Manager	41	Female	125825
117	PhD	16	Rural	Analyst	29	Female	111979
118	High School	11	Urban	Engineer	56	Female	74210
119	PhD	29	Urban	Director	63	Male	155080
120	Master	4	Rural	Manager	61	Male	109922
121	PhD	4	Suburban	Analyst	26	Male	104331
122	PhD	17	Urban	Engineer	37	Female	119683
123	Bachelor	17	Suburban	Engineer	51	Female	91499
124	Bachelor	10	Rural	Analyst	64	Male	79316
125	Bachelor	21	Urban	Director	26	Male	126769
126	High School	21	Suburban	Analyst	36	Female	79184
127	Bachelor	6	Urban	Analyst	50	Male	57464
128	Bachelor	25	Urban	Analyst	48	Male	89949
129	Bachelor	25	Suburban	Director	31	Male	121067
130	PhD	17	Suburban	Manager	52	Female	157190
131	High School	18	Urban	Analyst	58	Male	80837
132	PhD	22	Rural	Engineer	45	Male	151013
133	Bachelor	26	Urban	Manager	58	Female	110751
134	Master	10	Rural	Engineer	49	Male	102744
135	High School	13	Urban	Engineer	22	Female	62000
136	Bachelor	16	Suburban	Analyst	39	Female	81956
137	Master	7	Rural	Manager	32	Female	99647
138	High School	15	Rural	Analyst	38	Female	53660
139	Master	2	Urban	Engineer	20	Male	98066

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

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. 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 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).

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

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.

Таблица 5 – 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				$N_{\Sigma j}$			$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

Таблица 6 – 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				$P_{\Sigma j}$			

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*.

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

Таблица 7– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	Через относительные частоты	Через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке (таблица 7)	N_{ij} – фактическая частота; $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N}$ – теоретическая частота.	
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	...	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу	...	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij} N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.		
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	...	$I_{ij} = N_{ij} - \bar{N}_{ij} = N_{ij} - \frac{N_i N_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		

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 8 – 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		$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Таблица 9– 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 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 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

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

информацию о наблюдениях объекта моделирования, представленную в различных типах измерительных шкал и различных единицах измерения.

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

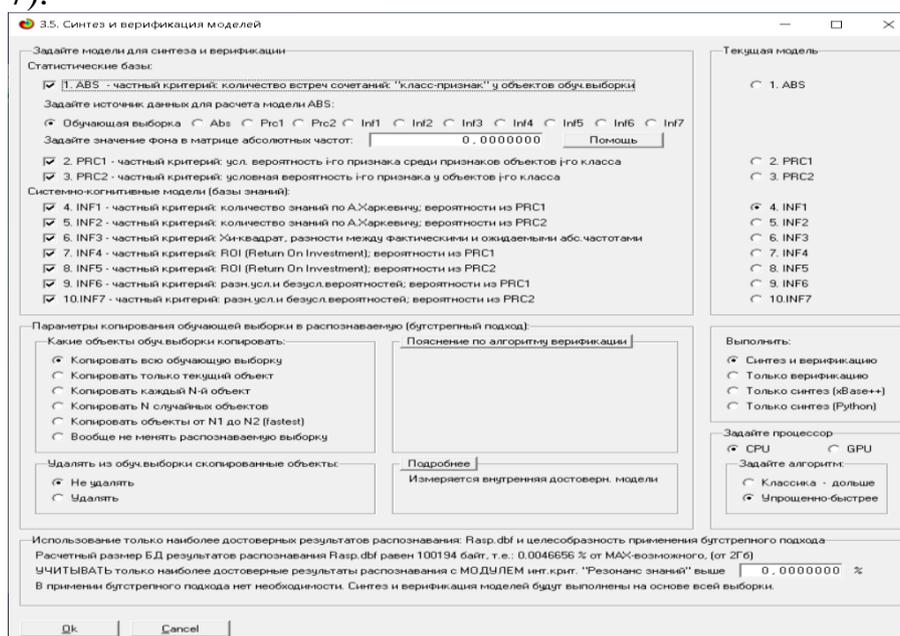
Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности 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».*

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



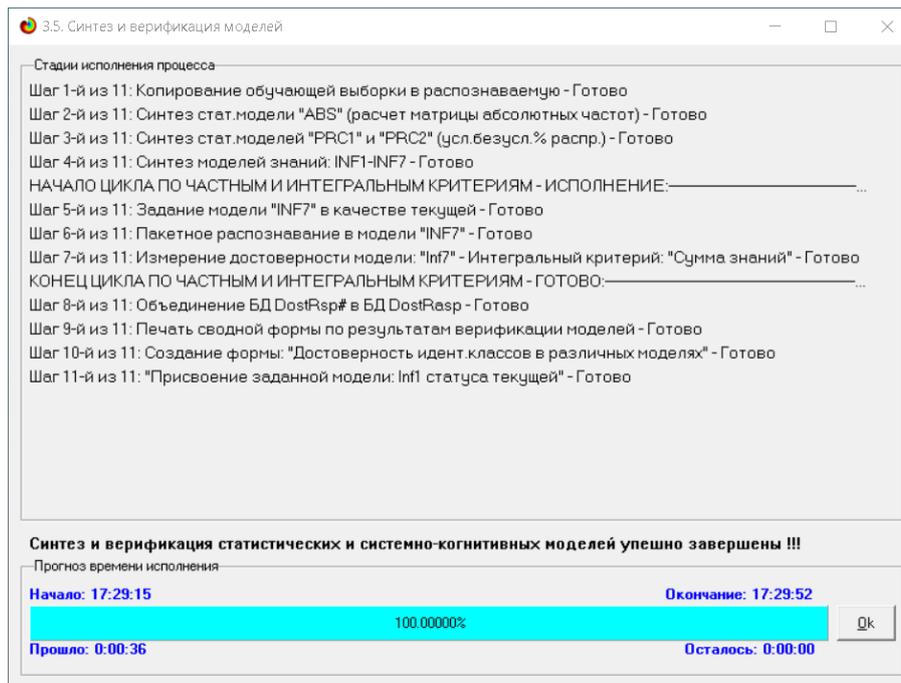


Рисунок 7. Экранные формы режима синтеза и верификации моделей

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

Код признака	Наименование описательной шкалы и градации	1. SALARY МАЛОЕ	2. SALARY БОЛЬШОЕ	Сумма	Среднее	Средн. квадр. откл.
1.0	EDUCATION-Bachelor	212.0	41.0	253.0	126.50	120.92
2.0	EDUCATION-High School	254.0	1.0	255.0	127.50	178.90
3.0	EDUCATION-Master	117.0	124.0	241.0	120.50	4.95
4.0	EDUCATION-PhD	34.0	217.0	251.0	125.50	129.40
5.0	EXPERIENCE-Очень малое	171.0	56.0	227.0	113.50	81.32
6.0	EXPERIENCE-Малое	133.0	52.0	185.0	92.50	57.28
7.0	EXPERIENCE-Среднее	113.0	67.0	180.0	90.00	32.53
8.0	EXPERIENCE-Большое	113.0	98.0	211.0	105.50	10.61
9.0	EXPERIENCE-Очень большое	87.0	110.0	197.0	98.50	16.26
10.0	LOCATION-Rural	251.0	94.0	345.0	172.50	111.02
11.0	LOCATION-Suburban	202.0	143.0	345.0	172.50	41.72
12.0	LOCATION-Urban	164.0	146.0	310.0	155.00	12.73
13.0	JOB_TITLE-Analyst	197.0	58.0	255.0	127.50	98.29
14.0	JOB_TITLE-Director	110.0	165.0	275.0	137.50	38.89
15.0	JOB_TITLE-Engineer	166.0	63.0	229.0	114.50	72.83
16.0	JOB_TITLE-Manager	144.0	97.0	241.0	120.50	33.23
17.0	AGE-Очень малое	128.0	96.0	224.0	112.00	22.63
18.0	AGE-Малое	103.0	64.0	167.0	83.50	27.58
19.0	AGE-Среднее	119.0	73.0	192.0	96.00	32.53
20.0	AGE-Большое	113.0	65.0	178.0	89.00	33.94
21.0	AGE-Очень большое	154.0	85.0	239.0	119.50	48.79
22.0	GENDER-Female	304.0	180.0	484.0	242.00	87.68
23.0	GENDER-Male	313.0	203.0	516.0	258.00	77.78
	Сумма числа признаков	3702.0	2298.0	6000.0		
	Среднее	161.0	99.9		130.43	
	Среднеквадратичное отклонение	69.5	54.3			68.98
	Сумма числа объектов обучающей выборки	617.0	383.0	1000.0		

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

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

Код признака	Наименование описательной шкалы и градации	1. SALARY МАЛОЕ	2. SALARY БОЛЬШОЕ	Безусл. вероятн.	Среднее	Средн. квадр. откл.
1.0	EDUCATION-Bachelor	34.360	10.705	25.300	22.532	16.810
2.0	EDUCATION-High School	41.167	0.261	25.500	20.714	29.008
3.0	EDUCATION-Master	18.963	32.376	24.100	25.669	9.568
4.0	EDUCATION-PhD	5.511	56.658	25.100	31.084	36.250
5.0	EXPERIENCE-Очень малое	27.715	14.621	22.700	21.168	9.341
6.0	EXPERIENCE-Малое	21.556	13.577	18.500	17.566	5.725
7.0	EXPERIENCE-Среднее	18.314	17.493	18.000	17.904	0.659
8.0	EXPERIENCE-Большое	18.314	25.587	21.100	21.951	5.225
9.0	EXPERIENCE-Очень большое	14.100	28.721	19.700	21.411	10.421
10.0	LOCATION-Rural	40.681	24.543	34.500	32.612	11.494
11.0	LOCATION-Suburban	32.739	37.337	34.500	35.038	3.333
12.0	LOCATION-Urban	26.580	38.120	31.000	32.350	8.243
13.0	JOB_TITLE-Analyst	31.929	15.144	25.500	23.536	11.952
14.0	JOB_TITLE-Director	17.828	43.081	27.500	30.455	17.940
15.0	JOB_TITLE-Engineer	26.904	16.449	22.900	21.677	7.476
16.0	JOB_TITLE-Manager	23.339	25.326	24.100	24.333	1.486
17.0	AGE-Очень малое	20.746	25.065	22.400	22.905	3.137
18.0	AGE-Малое	16.694	16.710	16.700	16.702	0.046
19.0	AGE-Среднее	19.287	19.060	19.200	19.173	0.229
20.0	AGE-Большое	18.314	16.971	17.800	17.643	1.030
21.0	AGE-Очень большое	24.959	22.193	23.900	23.576	2.038
22.0	GENDER-Female	49.271	46.997	48.400	48.134	1.689
23.0	GENDER-Male	50.729	53.003	51.600	51.866	1.689
	Сумма	600.000	600.000	1200.000		
	Среднее	26.087	26.087		26.087	
	Среднеквадратичное отклонение	11.275	14.179			12.663
	Сумма числа объектов обуч.выборки	617.000	383.000	1000.000		

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

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

Код признака	Наименование описательной шкалы и градации	1. SALARY МАЛОЕ	2. SALARY БОЛЬШОЕ	Сумма	Среднее	Средн. квадр. откл.
1.0	EDUCATION-Bachelor	0.035	-0.099	-0.064	-0.032	0.095
2.0	EDUCATION-High School	0.055	-0.527	-0.472	-0.236	0.411
3.0	EDUCATION-Master	-0.028	0.034	0.006	0.003	0.043
4.0	EDUCATION-PhD	-0.174	0.094	-0.081	-0.040	0.189
5.0	EXPERIENCE-Очень малое	0.023	-0.051	-0.028	-0.014	0.052
6.0	EXPERIENCE-Малое	0.018	-0.036	-0.018	-0.009	0.038
7.0	EXPERIENCE-Среднее	0.002	-0.003	-0.001	-0.001	0.004
8.0	EXPERIENCE-Большое	-0.016	0.022	0.006	0.003	0.027
9.0	EXPERIENCE-Очень большое	-0.038	0.043	0.005	0.002	0.058
10.0	LOCATION-Rural	0.019	-0.039	-0.020	-0.010	0.041
11.0	LOCATION-Suburban	-0.006	0.009	0.003	0.002	0.011
12.0	LOCATION-Urban	-0.018	0.024	0.006	0.003	0.029
13.0	JOB_TITLE-Analyst	0.026	-0.060	-0.034	-0.017	0.061
14.0	JOB_TITLE-Director	-0.050	0.052	0.002	0.001	0.072
15.0	JOB_TITLE-Engineer	0.019	-0.038	-0.020	-0.010	0.040
16.0	JOB_TITLE-Manager	-0.004	0.006	0.002	0.001	0.007
17.0	AGE-Очень малое	-0.009	0.013	0.004	0.002	0.015
18.0	AGE-Малое	0.000	0.000	0.000	0.000	0.000
19.0	AGE-Среднее	0.001	-0.001	0.000	0.000	0.001
20.0	AGE-Большое	0.003	-0.005	-0.002	-0.001	0.006
21.0	AGE-Очень большое	0.005	-0.009	-0.004	-0.002	0.010
22.0	GENDER-Female	0.002	-0.003	-0.001	-0.001	0.004
23.0	GENDER-Male	-0.002	0.003	0.001	0.001	0.004
	Сумма	-0.138	-0.571	-0.709		
	Среднее	-0.006	-0.025		-0.015	
	Среднеквадратичное отклонение	0.043	0.117			0.088
	Сумма числа объектов обуч.выборки	617.000	383.000	1000.000		

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

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

Код признака	Наименование описательной шкалы и градации	1. SALARY МАЛОЕ	2. SALARY БОЛЬШОЕ	Сумма	Среднее	Средн. квадрат. откл.
1.0	EDUCATION-Bachelor	55.899	-55.899			79.053
2.0	EDUCATION-High School	96.665	-96.665			136.705
3.0	EDUCATION-Master	-31.697	31.697			44.826
4.0	EDUCATION-PhD	-120.867	120.867			170.932
5.0	EXPERIENCE-Очень малое	30.941	-30.941			43.757
6.0	EXPERIENCE-Малое	18.855	-18.855			26.665
7.0	EXPERIENCE-Среднее	1.940	-1.940			2.744
8.0	EXPERIENCE-Большое	-17.187	17.187			24.306
9.0	EXPERIENCE-Очень большое	-34.549	34.549			48.860
10.0	LOCATION-Rural	38.135	-38.135			53.931
11.0	LOCATION-Suburban	-10.865	10.865			15.365
12.0	LOCATION-Urban	-27.270	27.270			38.566
13.0	JOB_TITLE-Analyst	39.665	-39.665			56.095
14.0	JOB_TITLE-Director	-59.675	59.675			84.393
15.0	JOB_TITLE-Engineer	24.707	-24.707			34.941
16.0	JOB_TITLE-Manager	-4.697	4.697			6.643
17.0	AGE-Очень малое	-10.208	10.208			14.436
18.0	AGE-Малое	-0.039	0.039			0.055
19.0	AGE-Среднее	0.536	-0.536			0.758
20.0	AGE-Большое	3.174	-3.174			4.489
21.0	AGE-Очень большое	6.537	-6.537			9.245
22.0	GENDER-Female	5.372	-5.372			7.597
23.0	GENDER-Male	-5.372	5.372			7.597
	Сумма					
	Среднее					
	Среднеквадратичное отклонение	42.204	42.204			41.732
	Сумма числа объектов обуч. выборки	617.000	383.000	1000.000		

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

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

3.4. Задача-4. Верификация моделей

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

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

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

В соответствии с критерием достоверности L1-мерой проф.Е.В.Луценко наиболее достоверной является СК-модель INF4 и INF5 (хи-квадрат К.Пирсона) с интегральным критерием: «Сумма знаний»: **L1=0.948** при максимуме 1 (рисунок 12). Мы будем использовать модель **INF4** для решения поставленных в работе задач.

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

Наименование модели и частного критерия	Интегральный критерий	Полнота модели	Федера Ван Риббергена	Сумма модул. уровней сход. истинно-поло. 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	Средний модуль истинно-полож. решений	Средний модуль истинно-отриц. реш.
1. ABS - частный критерий: количество встреч сометаний "класс"	Корреляция абс.частот с обр...	4	0.991	0.788	438.718	68.387	80.774	0.478	0.845	0.999	0.915	0.443
1. ABS - частный критерий: количество встреч сометаний "класс"	Сумма абс.частот по признак...	0	1.000	0.667	743.818		555.561	0.572	1.000	0.728	0.744	
2. PFC1 - частный критерий: усл. вероятность i-го признака сред.	Корреляция усл.отн.частот с о...	4	0.991	0.788	438.718	68.387	80.774	0.478	0.845	0.999	0.915	0.443
2. PFC1 - частный критерий: усл. вероятность i-го признака сред.	Сумма усл.отн.частот по призна...	0	1.000	0.667	779.485		643.846	0.548	1.000	0.708	0.779	
3. PFC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	4	0.991	0.788	438.701	68.386	80.772	0.478	0.845	0.999	0.915	0.443
3. PFC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по призна...	0	1.000	0.667	779.485		643.846	0.548	1.000	0.708	0.779	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0	0.937	0.821	288.338	304.137	42.357	15.104	0.872	0.950	0.909	0.308
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	2	0.795	0.854	77.417	328.548	2.596	14.562	0.968	0.842	0.900	0.097
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0	0.937	0.821	288.338	304.136	42.357	15.104	0.872	0.950	0.909	0.308
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	2	0.795	0.854	77.417	328.548	2.596	14.562	0.968	0.842	0.900	0.097
6. INF3 - частный критерий: Миквадар, разности между факт...	Семантический резонанс зна...	5	0.875	0.875	321.882	321.882	17.870	17.870	0.947	0.947	0.947	0.368
6. INF3 - частный критерий: Миквадар, разности между факт...	Сумма знаний	5	0.875	0.875	321.882	321.882	17.870	17.870	0.947	0.947	0.947	0.368
7. INF4 - частный критерий: ROI (Return On Investment); вероят...	Семантический резонанс зна...	4	0.874	0.874	321.577	321.577	17.737	17.737	0.948	0.948	0.948	0.368
7. INF4 - частный критерий: ROI (Return On Investment); вероят...	Сумма знаний	4	0.874	0.874	258.261	262.936	16.711	12.036	0.939	0.955	0.947	0.295
8. INF5 - частный критерий: ROI (Return On Investment); вероят...	Семантический резонанс зна...	4	0.874	0.874	321.577	321.577	17.737	17.737	0.948	0.948	0.948	0.368
8. INF5 - частный критерий: ROI (Return On Investment); вероят...	Сумма знаний	4	0.874	0.874	258.261	262.936	16.711	12.036	0.939	0.955	0.947	0.295
9. INF6 - частный критерий: разн усл и безуслов. вероятностей; вер...	Семантический резонанс зна...	5	0.875	0.875	321.882	321.882	17.870	17.870	0.947	0.947	0.947	0.368
9. INF6 - частный критерий: разн усл и безуслов. вероятностей; вер...	Сумма знаний	5	0.875	0.875	258.377	263.312	16.807	12.156	0.939	0.955	0.947	0.295
10. INF7 - частный критерий: разн усл и безуслов. вероятностей; ве...	Семантический резонанс зна...	5	0.875	0.875	321.882	321.882	17.870	17.870	0.947	0.947	0.947	0.368
10. INF7 - частный критерий: разн усл и безуслов. вероятностей; ве...	Сумма знаний	5	0.875	0.875	258.377	263.312	16.807	12.156	0.939	0.955	0.947	0.295

Помощь по мерам достоверности | Помощь по частотным распределениям | TR|TN|FP|FN | (TP-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 СК-модели INF4.

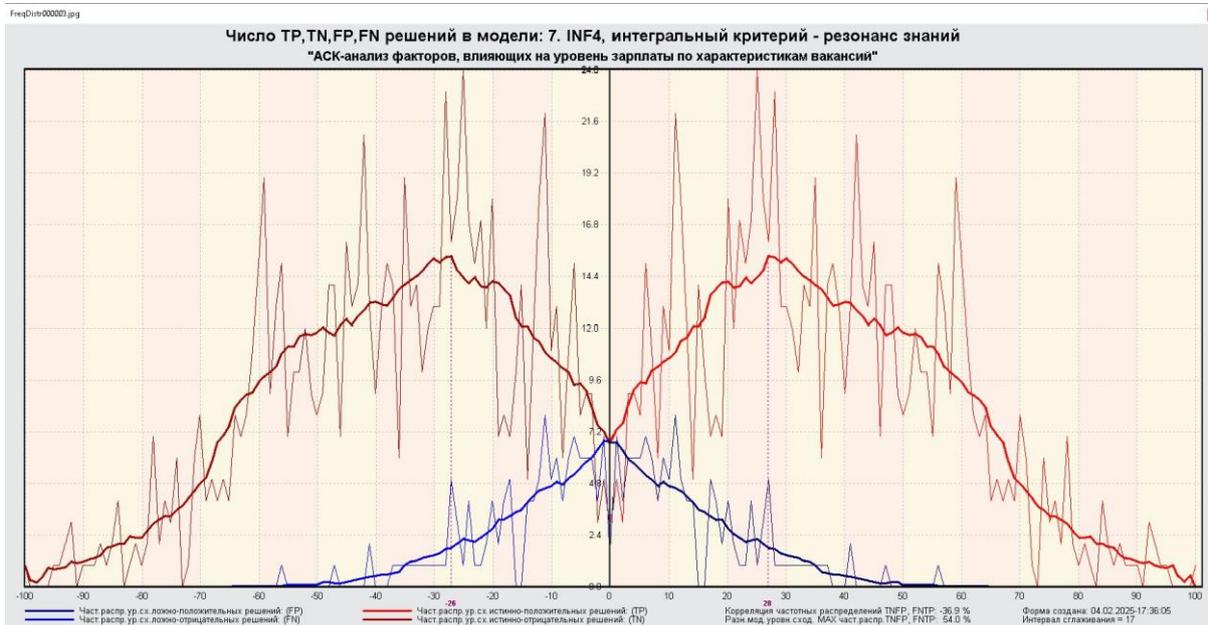




Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4

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

- отсутствие отрицательных ложных решений: В задаче анализа факторов, влияющих на уровень зарплаты, отрицательные ложные решения (FN) практически не встречаются. Это указывает на высокую точность модели в выявлении незначительного влияния различных характеристик вакансий на уровень зарплаты. В общем случае такие ошибки возможны только в очень редких случаях с минимальными различиями между параметрами;

- ложные положительные и истинные положительные решения: При уровнях сходства менее 25% преобладают ложные положительные решения (FP), что может быть связано с недостаточной информативностью данных или сложностью выявления значимых факторов при малых различиях. Однако при более высоких уровнях сходства (выше 72%) модель демонстрирует преимущественно истинно положительные решения (TP). Это свидетельствует о том, что система «Эйдос» эффективно распознает значимые факторы, влияющие на уровень зарплаты, при наличии достаточного количества схожих данных;

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

На рисунках 14 приведены экранные формы хелпов режима 3.4, в которых подробно объясняется смысл этого режима. Эти формы

приводятся в данной работе вместо более детального описания данного режима.

Помощь по режимам: 3.4, 4.1.3.#: Виды прогнозов и меры достоверности моделей в системе "Эйдос-Х++"

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

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

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

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

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

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

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

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

Классическая F-мера достоверности моделей Ван Ризбергера (колонка выделена ярко-голубым фоном):

$$F\text{-мера} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$
 - достоверность модели
Precision = TP/(TP+FP) - точность модели;
Recall = TP/(TP+FN) - полнота модели;

L1-мера проф.Е.В.Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СУММ уровней сходства (колонка выделена ярко-зеленым фоном):

$$L1\text{-мера} = 2 * (SPrecision * SRecall) / (SPrecision + SRecall)$$

SPrecision = STP/(STP+SFP) - точность с учетом сумм уровней сходства;
SRecall = STN/(STN+SFN) - полнота с учетом сумм уровней сходства;
STP - Сумма модулей сходства истинно-положительных решений; STN - Сумма модулей сходства истинно-отрицательных решений;
SFP - Сумма модулей сходства ложно-положительных решений; SFN - Сумма модулей сходства ложно-отрицательных решений.

L2-мера проф.Е.В.Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СРЕДНИХ уровней сходства (колонка выделена желтым фоном):

$$L2\text{-мера} = 2 * (APrecision * ARecall) / (APrecision + ARecall)$$

APrecision = ATP/(ATP+AFP) - точность с учетом средних уровней сходства;
ARecall = ATN/(ATN+AFN) - полнота с учетом средних уровней сходства;
ATP=STP/TP - Среднее модулей сходства истинно-положительных решений; AFN=SFN/FN - Среднее модулей сходства истинно-отрицательных решений;
AFP=SFP/FP - Среднее модулей сходства ложно-положительных решений; AFN=SFN/FN - Среднее модулей сходства ложно-отрицательных решений.

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

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

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

Помощь по режиму 3.4. (C) Система "ЭЙДОС-Х++"

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

По нажатию кнопка: [TP,TN,FP,FN], [(TP-FP)/(TN-FN)], [(T-F)/(T+F)*100] отображаются графики частотных распределений для модели и интегрального критерия той строки, на которой в экранной форме 3.4 стоит курсор. По клику на кнопку: [(T-F)/(T+F)*100] выводятся графики частотных распределений: (TP-FP)/(TP+FP)*100 и (TN-FN)/(TN+FN)*100.

где:
TP-True-Positive; TN-True-Negative; FP-False Positive; FN-False-Negative, количество истинных и ложных положительных и отрицательных решений.

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

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

The left graph, titled 'Количество решений' (Number of solutions), plots TP, TN, FP, and FN against 'Объем выборки' (Sample size). All four metrics increase linearly with sample size. The right graph, titled 'Динамика F-меры Ван Ризбергера и L1- и L2- мер проф.Е.В.Луценко в зависимости от объема выборки RND-модели' (Dynamics of F-MER, L1-MER, and L2-MER metrics depending on the volume of RND-model samples), plots 'Достоверность, %' (Reliability, %) against 'Число объектов обучающей выборки' (Number of objects in the training sample). The F-MER metric (yellow line) starts at approximately 0.9 and decreases towards 0.4. The L1-MER metric (purple line) starts at approximately 0.9 and decreases towards 0.5. The L2-MER metric (blue line) starts at approximately 0.9 and decreases towards 0.6.

Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

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

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (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.

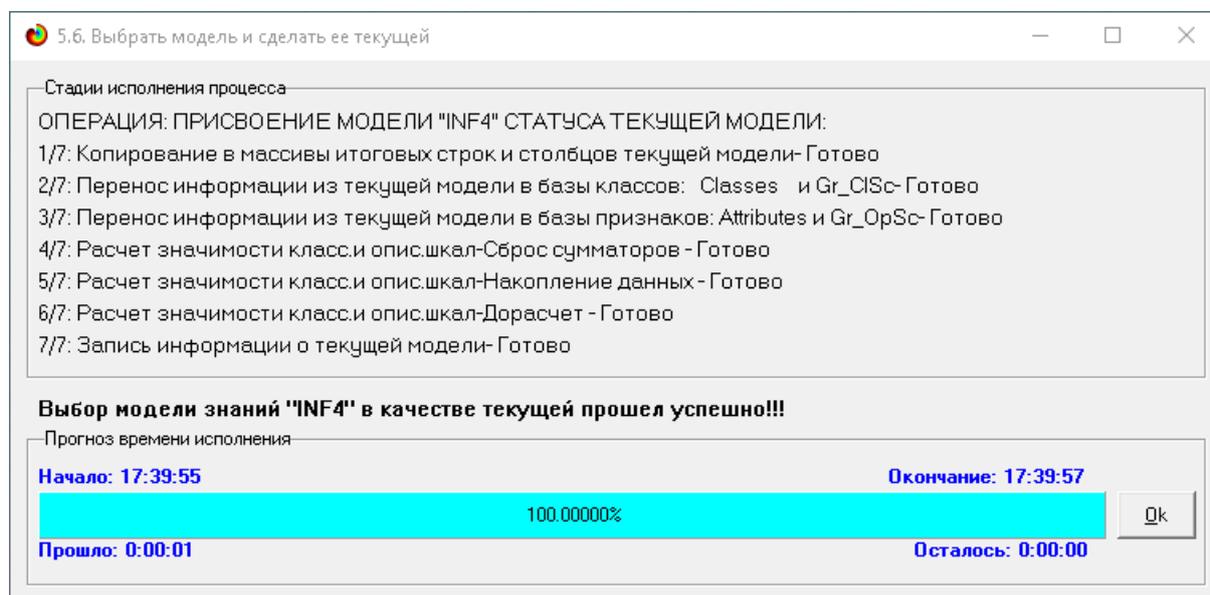


Рисунок 15. Задание СК-модели INF4 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

При решении задачи идентификации каждый объект распознаваемой выборки сравнивается по всем своим признакам с каждым из обобщенных образов классов. Смысл решения задачи идентификации заключается в том, что при определении принадлежности конкретного объекта к обобщенному образу классу об этом конкретном объекте *по аналогии*

становится известно все, что известно об объектах этого класса, по крайней мере, самое существенное о них, т.е. чем они отличаются от объектов других классов.

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

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

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

3.6.1. 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;

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

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

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

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

3.6.2. Интегральный критерий «Семантический резонанс знаний»

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

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

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

где:

M – количество градаций описательных шкал (признаков); \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.

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

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

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

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

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

Произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применения вейвлетов и сплайнов, в частности линейной интерполяции:

$$I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, \quad L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}},$$

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

3.6.3. 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 **неортометризованных** 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.

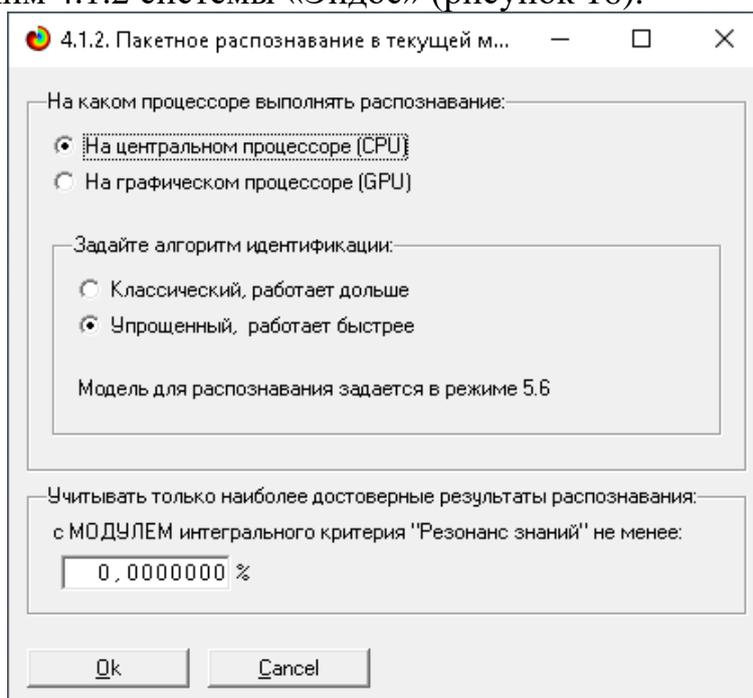
В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется **вес** каждого обобщенного образа класса в образе объекта, что подробнее описано в работах [2, 3, 4].

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

3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос»

В АСК-анализе разработаны, а в системе «Эйдос» реализованы развитые методы прогнозирования, основанные на сценарном методе АСК-анализа или сценарном АСК-анализе. 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] и в ряде других.

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



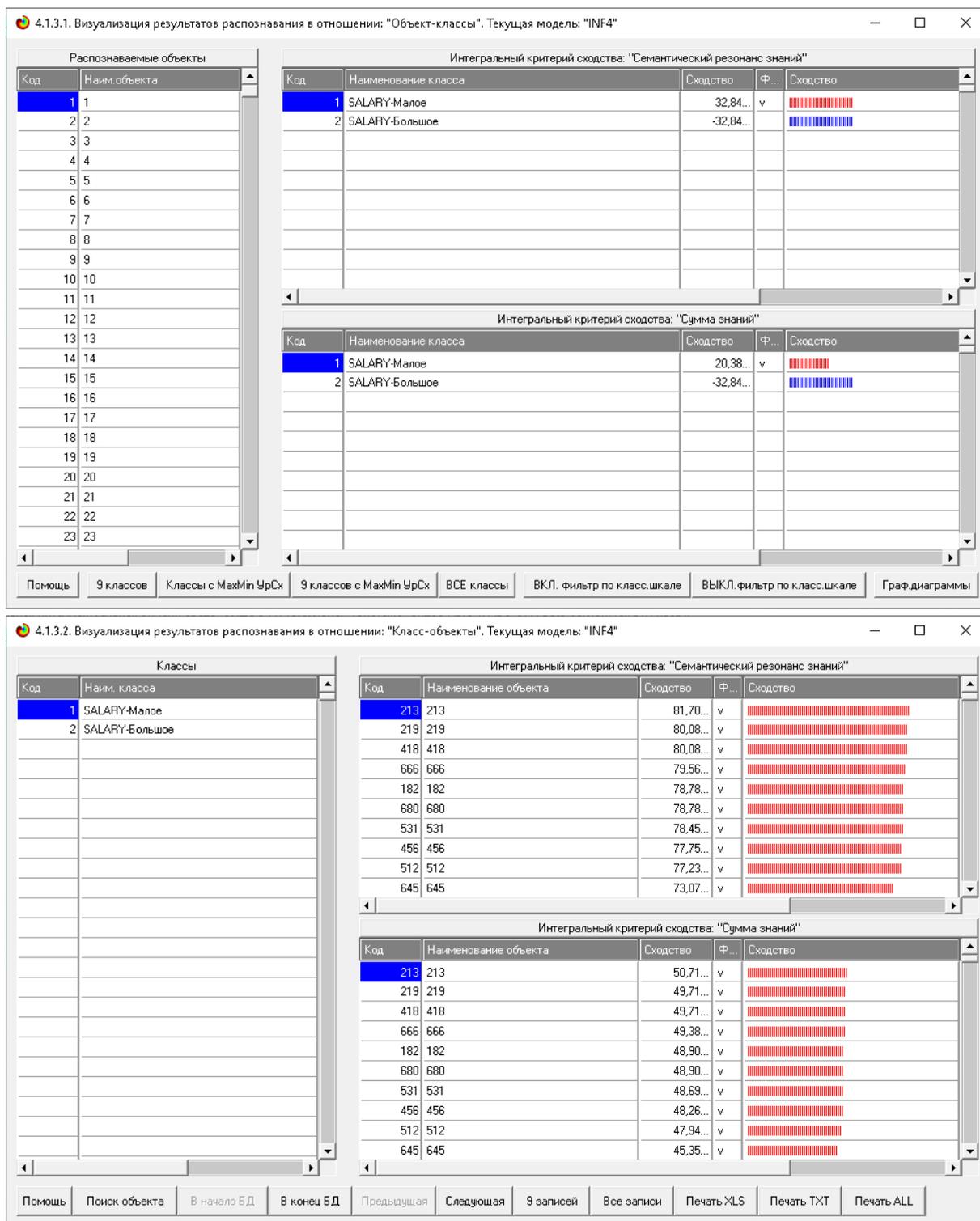


Рисунок 18. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

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») (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.

Важно отметить, что в системе «Эйдос» SWOT-диаграммы создаются автоматически, основываясь на статистических и системно-когнитивных моделях, которые строятся непосредственно на основе реальных эмпирических данных. В отличие от традиционного подхода, при котором SWOT-анализ часто проводится экспертами интуитивно, на основе личного опыта и профессиональной компетенции – то есть

практически "на глазок", а иногда даже "от фонаря", система «Эйдос» обеспечивает формализованный и объективный процесс анализа.

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

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	SALARY-Малое	0,2711601	3702	61,7000000
2	SALARY-Большое	0,4368298	2298	38,3000000

SWOT-анализ класса: 1 "SALARY-Малое" в модели: 7 "INF4"

Способствующие факторы и сила их влияния

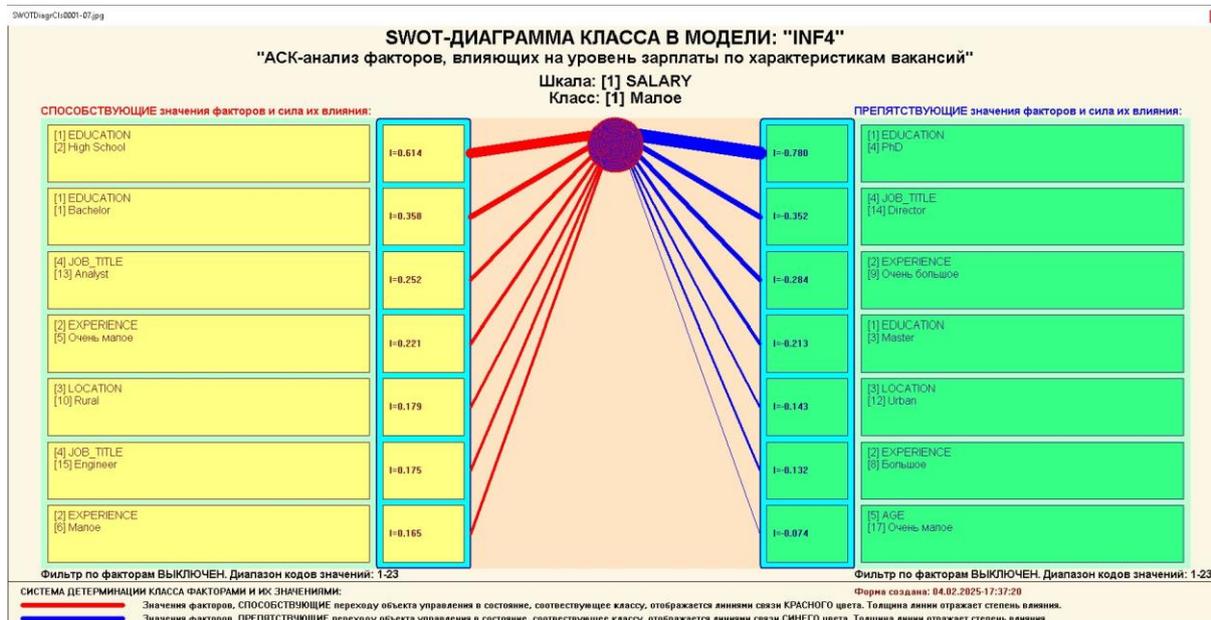
Код	Наименование фактора и его интервального значения	Сила влияния
2	EDUCATION-High School	0.614
1	EDUCATION-Bachelor	0.358
13	JOB_TITLE-Analyst	0.252
5	EXPERIENCE-Очень малое	0.221
10	LOCATION-Rural	0.179
15	JOB_TITLE-Engineer	0.175
6	EXPERIENCE-Малое	0.165
21	AGE-Очень большое	0.044
20	AGE-Большое	0.029
22	GENDER-Female	0.018
7	EXPERIENCE-Среднее	0.017
19	AGE-Среднее	0.005

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	EDUCATION-PhD	-0.780
14	JOB_TITLE-Director	-0.352
9	EXPERIENCE-Очень большое	-0.284
3	EDUCATION-Master	-0.213
12	LOCATION-Urban	-0.143
8	EXPERIENCE-Большое	-0.132
17	AGE-Очень малое	-0.074
11	LOCATION-Suburban	-0.051
16	JOB_TITLE-Manager	-0.032
23	GENDER-Male	-0.017
18	AGE-Малое	0.000

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

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



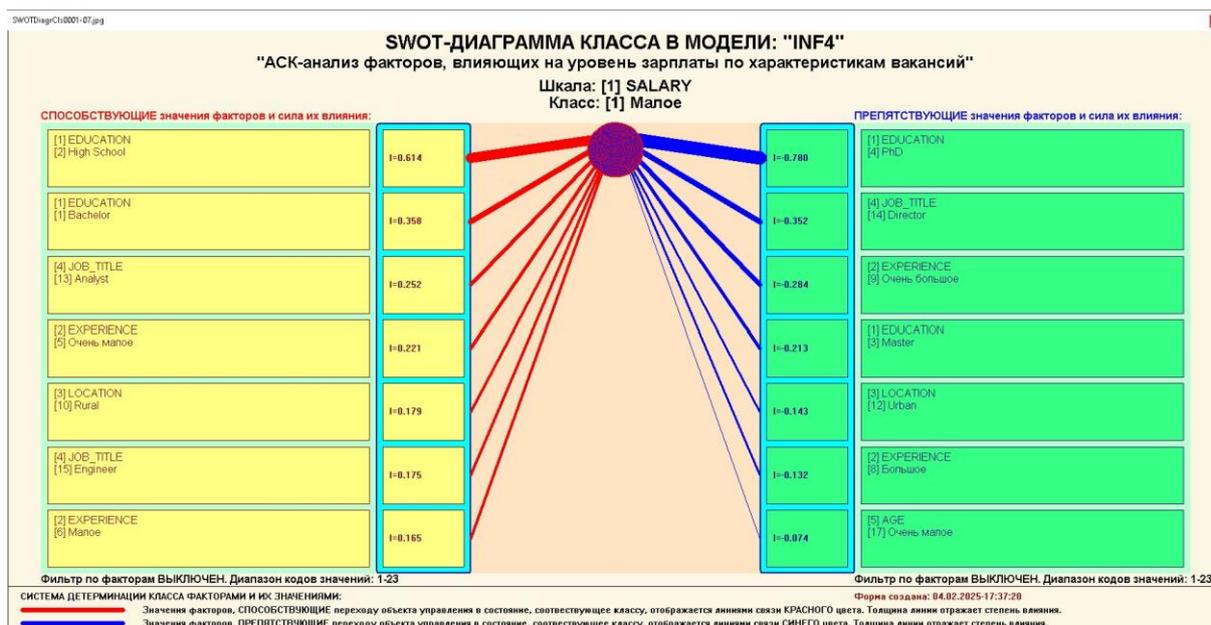


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

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

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

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

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.

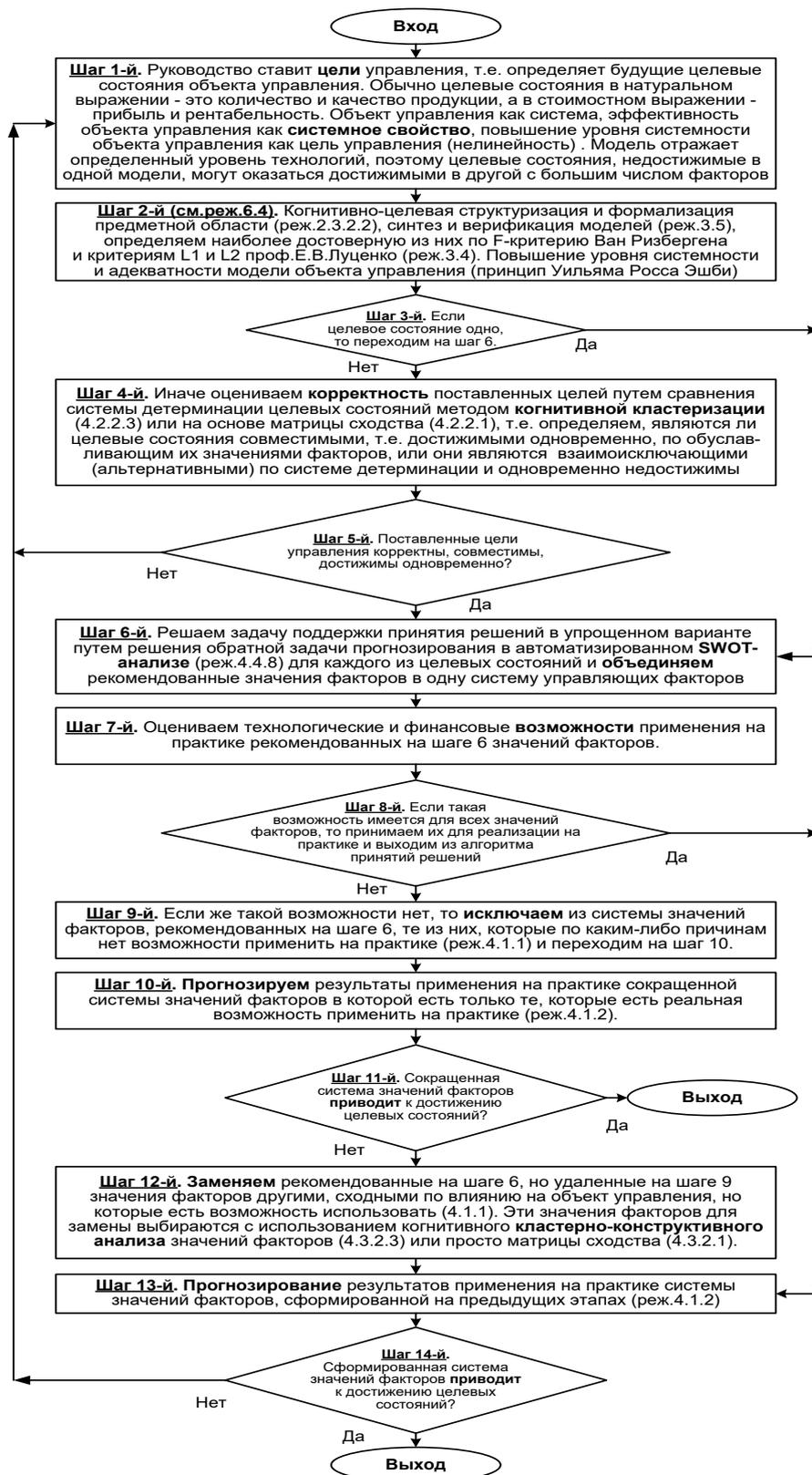


Рисунок 20. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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

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

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

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

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

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

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

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:

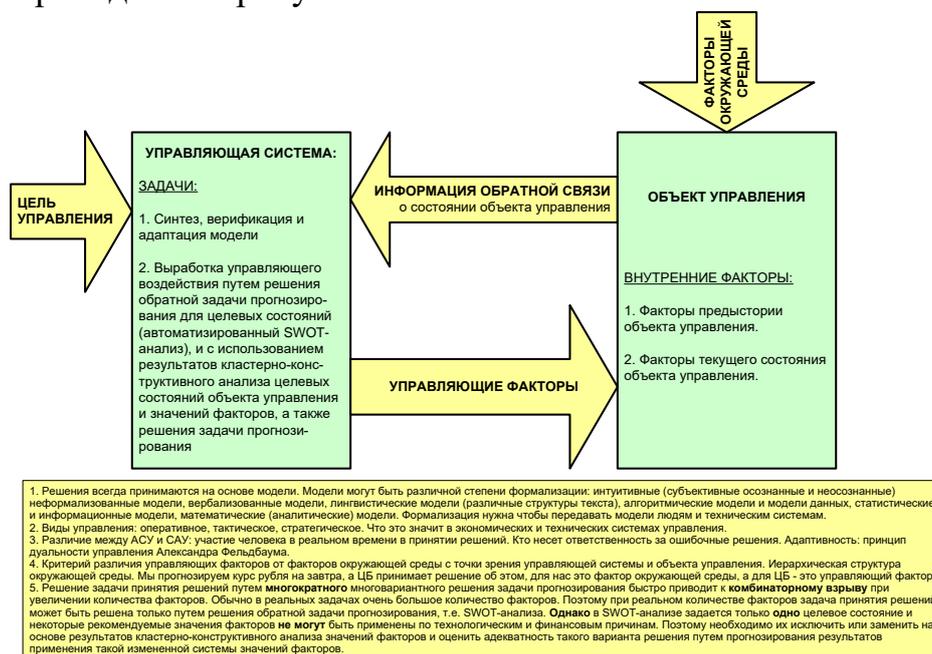


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

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.

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.

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

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

Инвертированные SWOT-диаграмм (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».

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

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами АСК-анализа в системе "Эйдос"

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

Код	Наименование значения фактора
1	EDUCATION-Bachelor
2	EDUCATION-High School
3	EDUCATION-Master
4	EDUCATION-PhD
5	EXPERIENCE-Очень малое
6	EXPERIENCE-Малое

SWOT-анализ значения фактора:1 "EDUCATION-Bachelor" в модели:6 "INF3"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влияния
1	SALARY-Малое	55.899

ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
2	SALARY-Большое	-55.899

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

Помощь Abs Prc1 Prc2 Inf1 Inf2 **Inf3** Inf4 Inf5 Inf6 Inf7

SWOT-диаграмма

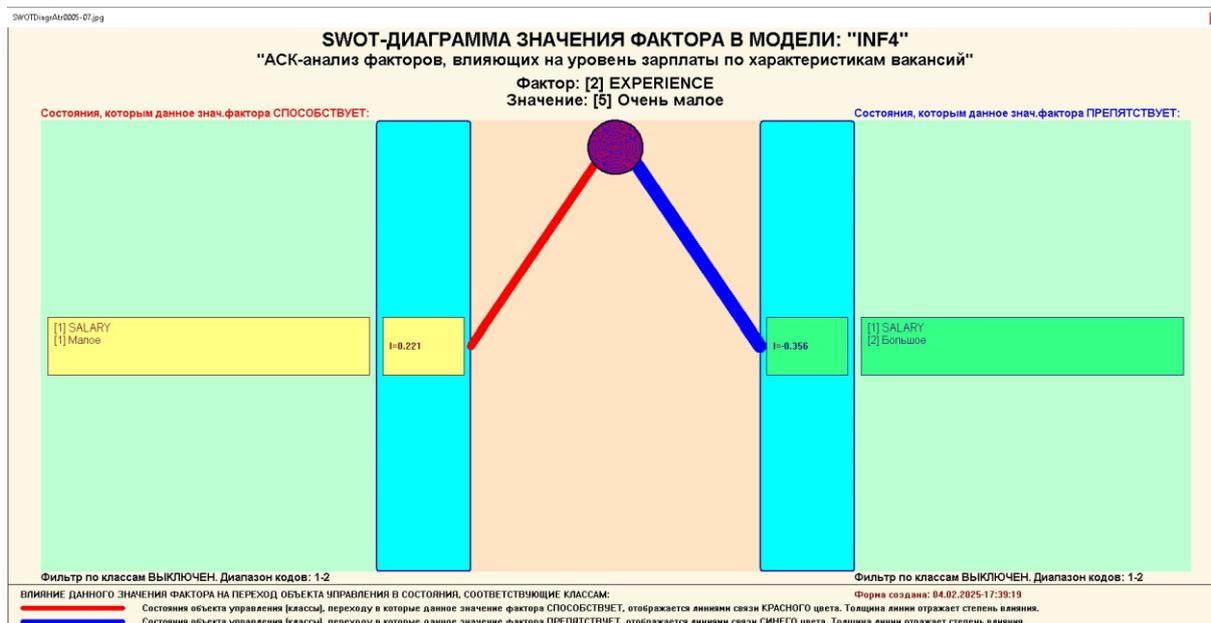
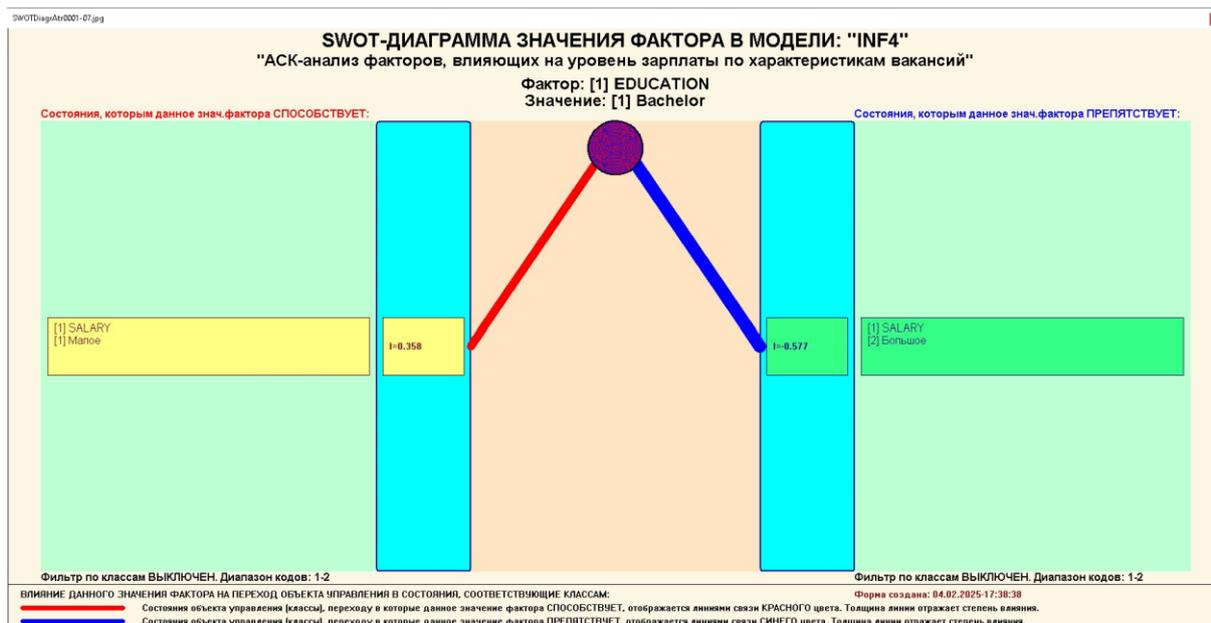


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

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

3.8.2. Кластерно-конструктивный анализ классов

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.

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

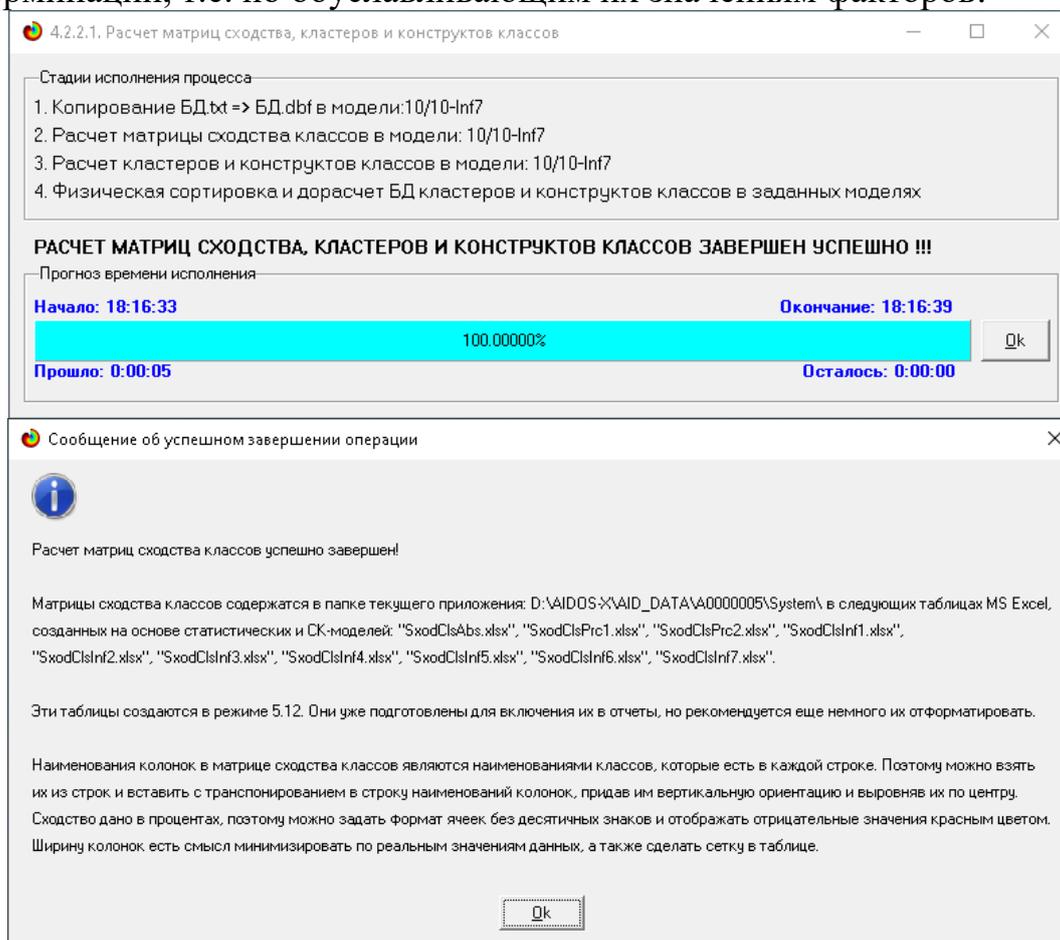


Рисунок 23. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

Таблица 11 – Матрица сходства классов в СК-модели INF4 (полностью)

KOD_CLS	KOD_CLSC	NAME_CLS	N1	N2
1		1 SALARY-Малое	100,0000000	-100,0000000
2		1 SALARY-Большо	-100,0000000	100,0000000



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

3.8.3.Кластерно-конструктивный анализ значений описательных шкал

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) (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.

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

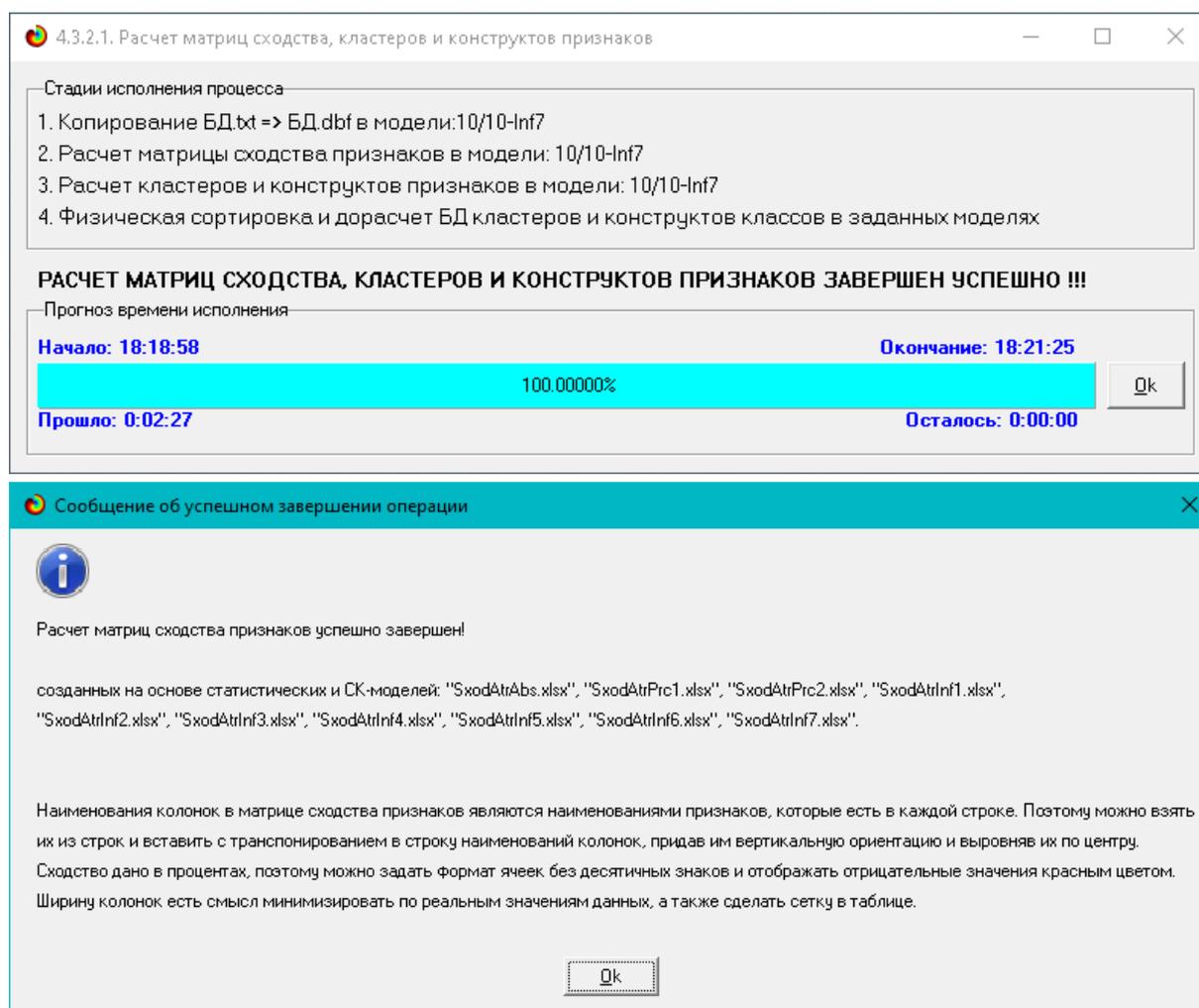


Рисунок 25. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

Таблица 12 – Матрица сходства значений факторов в СК-модели INF4 (фрагмент)

KOD_ATR	KOD_OPSC	NAME_ATR	N1	N2	N3
1	1	EDUCATION-Bachelor	100,0000000	100,0000000	-100,0000000
2	1	EDUCATION-High School	100,0000000	100,0000000	-100,0000000
3	1	EDUCATION-Master	-100,0000000	-100,0000000	100,0000000
4	1	EDUCATION-PhD	-100,0000000	-100,0000000	100,0000000
5	2	EXPERIENCE-Очень малое	100,0000000	100,0000000	-100,0000000
6	2	EXPERIENCE-Малое	100,0000000	100,0000000	-100,0000000
7	2	EXPERIENCE-Среднее	100,0000000	100,0000000	-100,0000000
8	2	EXPERIENCE-Большое	-100,0000000	-100,0000000	100,0000000
9	2	EXPERIENCE-Очень большое	-100,0000000	-100,0000000	100,0000000
10	3	LOCATION-Rural	100,0000000	100,0000000	-100,0000000
11	3	LOCATION-Suburban	-100,0000000	-100,0000000	100,0000000
12	3	LOCATION-Urban	-100,0000000	-100,0000000	100,0000000
13	4	JOB_TITLE-Analyst	100,0000000	100,0000000	-100,0000000
14	4	JOB_TITLE-Director	-100,0000000	-100,0000000	100,0000000
15	4	JOB_TITLE-Engineer	100,0000000	100,0000000	-100,0000000
16	4	JOB_TITLE-Manager	-100,0000000	-100,0000000	100,0000000
17	5	AGE-Очень малое	-100,0000000	-100,0000000	100,0000000
18	5	AGE-Малое	-100,0000000	-100,0000000	100,0000000
19	5	AGE-Среднее	100,0000000	100,0000000	-100,0000000
20	5	AGE-Большое	100,0000000	100,0000000	-100,0000000
21	5	AGE-Очень большое	100,0000000	100,0000000	-100,0000000
22	6	GENDER-Female	100,0000000	100,0000000	-100,0000000
23	6	GENDER-Male	-100,0000000	-100,0000000	100,0000000

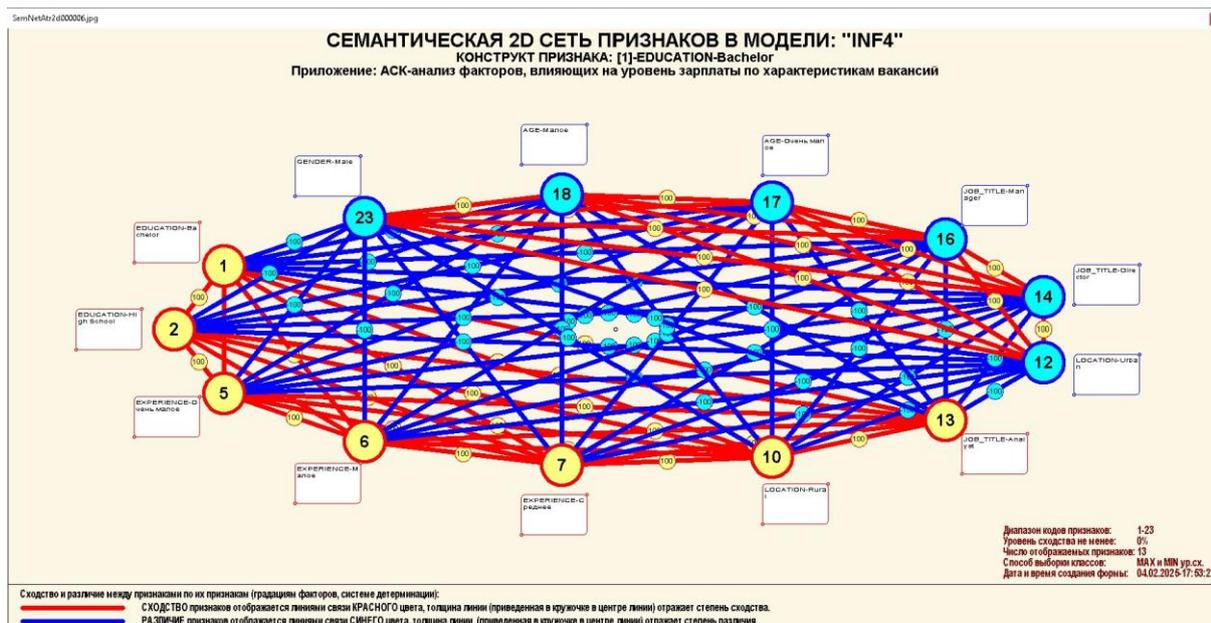


Рисунок 26. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF4 (режим 4.3.2.2)

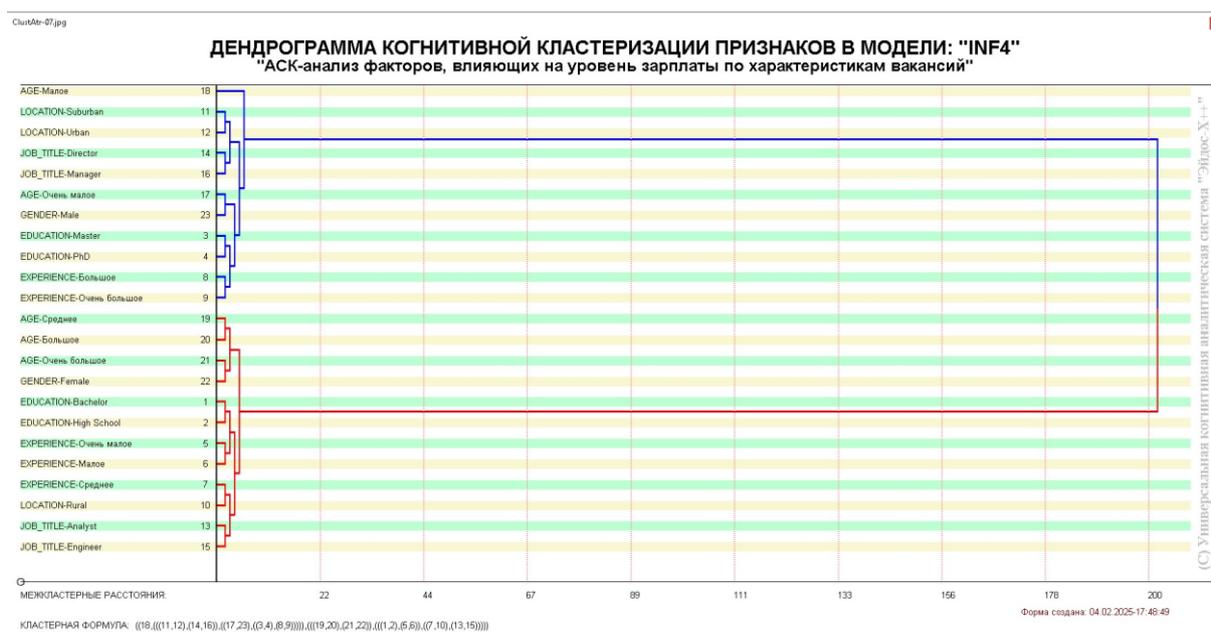


Рисунок 27. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)

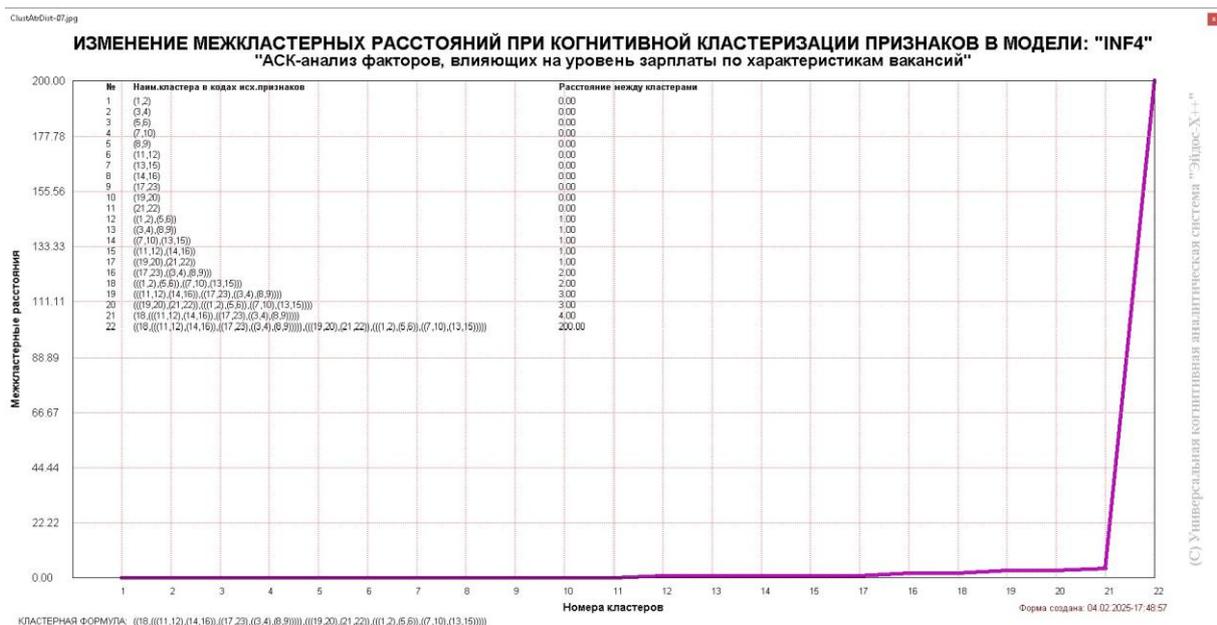


Рисунок 28. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

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

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

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

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

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

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

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

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

4.4.10. Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	SALARY-Малое
2	SALARY-Большое

Подготовка визуализации нейрона: 1 "SALARY-Малое" в модели: 7 "INF4"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
2	EDUCATION-High School	0.614
1	EDUCATION-Bachelor	0.358
13	JOB_TITLE-Analyst	0.252
5	EXPERIENCE-Очень малое	0.221
10	LOCATION-Rural	0.179
15	JOB_TITLE-Engineer	0.175
6	EXPERIENCE-Малое	0.165
21	AGE-Очень большое	0.044
20	AGE-Большое	0.029
22	СЕМЬЯ-Большая	0.017

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	EDUCATION-PhD	-0.780
14	JOB_TITLE-Director	-0.352
9	EXPERIENCE-Очень большое	-0.284
3	EDUCATION-Master	-0.213
12	LOCATION-Urban	-0.143
8	EXPERIENCE-Большое	-0.132
17	AGE-Очень малое	-0.074
11	LOCATION-Suburban	-0.051
16	JOB_TITLE-Manager	-0.032
23	СЕМЬЯ-Малая	0.017

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

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

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

Сортировать рецепторы: по информативности по модулю информативности

Отображать рецепторы: с наименованиями только с кодами

4.4.10. Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	SALARY-Малое
2	SALARY-Большое

Подготовка визуализации нейрона: 1 "SALARY-Малое" в модели: 7 "INF4"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
2	EDUCATION-High School	0.614
1	EDUCATION-Bachelor	0.358
13	JOB_TITLE-Analyst	0.252
5	EXPERIENCE-Очень малое	0.221
10	LOCATION-Rural	0.179
15	JOB_TITLE-Engineer	0.175
6	EXPERIENCE-Малое	0.165
21	AGE-Очень большое	0.044
20	AGE-Большое	0.029
22	СЕМЬЯ-Большая	0.017

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
4	EDUCATION-PhD	-0.780
14	JOB_TITLE-Director	-0.352
9	EXPERIENCE-Очень большое	-0.284
3	EDUCATION-Master	-0.213
12	LOCATION-Urban	-0.143
8	EXPERIENCE-Большое	-0.132
17	AGE-Очень малое	-0.074
11	LOCATION-Suburban	-0.051
16	JOB_TITLE-Manager	-0.032
23	СЕМЬЯ-Малая	0.017

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

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

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

Сортировать рецепторы: по информативности по модулю информативности

Отображать рецепторы: с наименованиями только с кодами

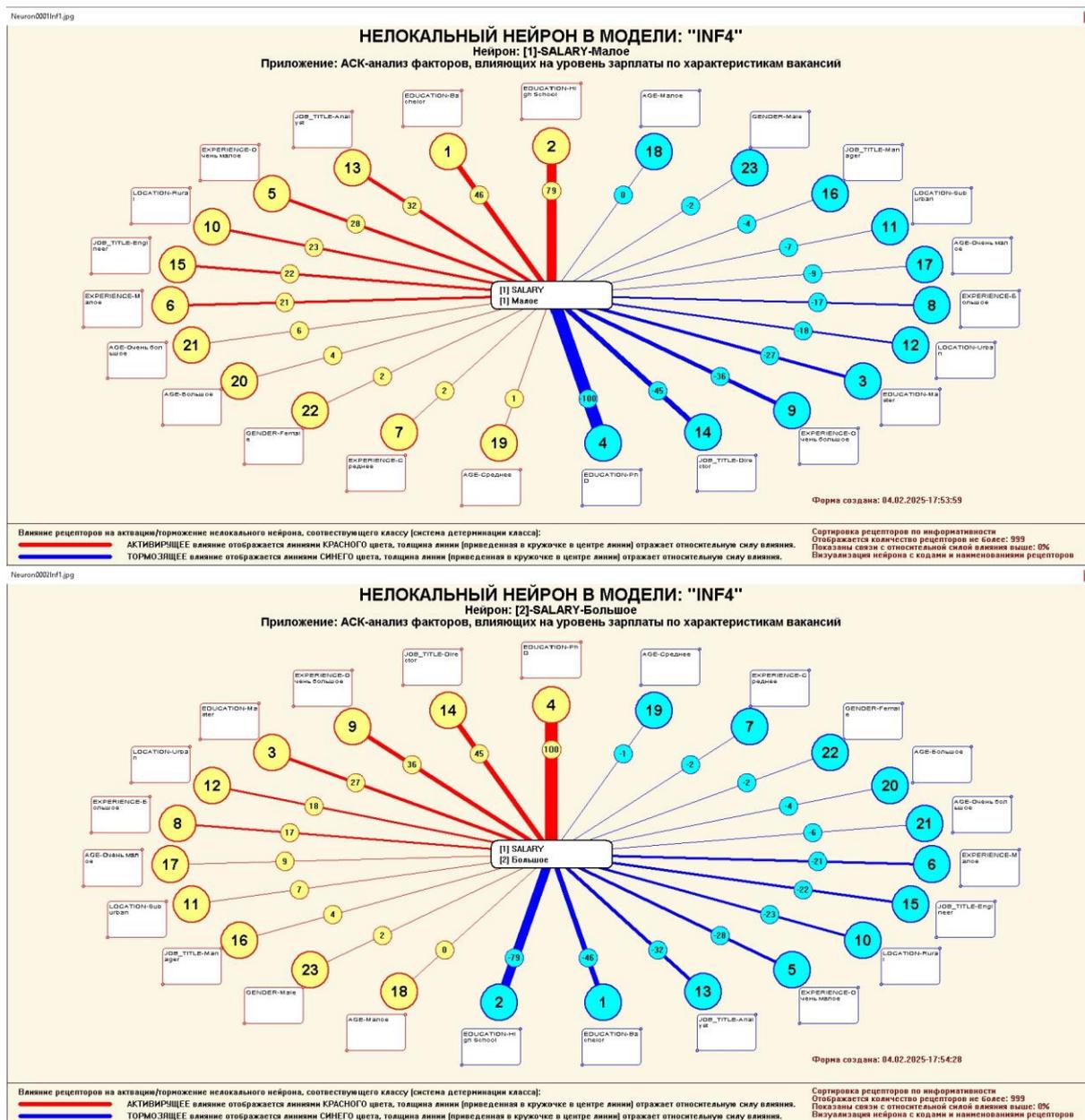


Рисунок 29. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

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

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

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.

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

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	SALARY-Малое
<input type="checkbox"/>	2	SALARY-Большое

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

Подготовка визуализации нейрона: 1 "SALARY-Малое" в модели: 7 "INF4"

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

Код	Наименование фактора и его интервального значения	Сила влияния
2	EDUCATION-High School	0.614
1	EDUCATION-Bachelor	0.358
13	JOB_TITLE-Analyst	0.252
5	EXPERIENCE-Очень малое	0.221
10	LOCATION-Rural	0.179
15	JOB_TITLE-Engineer	0.175
6	EXPERIENCE-Малое	0.165
21	AGE-Очень большое	0.044
20	AGE-Большое	0.029
22	AGE-Малое	0.019

Код	Наименование фактора и его интервального значения	Сила влияния
4	EDUCATION-PhD	-0.780
14	JOB_TITLE-Director	-0.352
9	EXPERIENCE-Очень большое	-0.284
3	EDUCATION-Master	-0.213
12	LOCATION-Urban	-0.143
8	EXPERIENCE-Большое	-0.132
17	AGE-Очень малое	-0.074
11	LOCATION-Suburban	-0.051
16	JOB_TITLE-Manager	-0.032
23	AGE-Очень малое	0.017

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

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

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

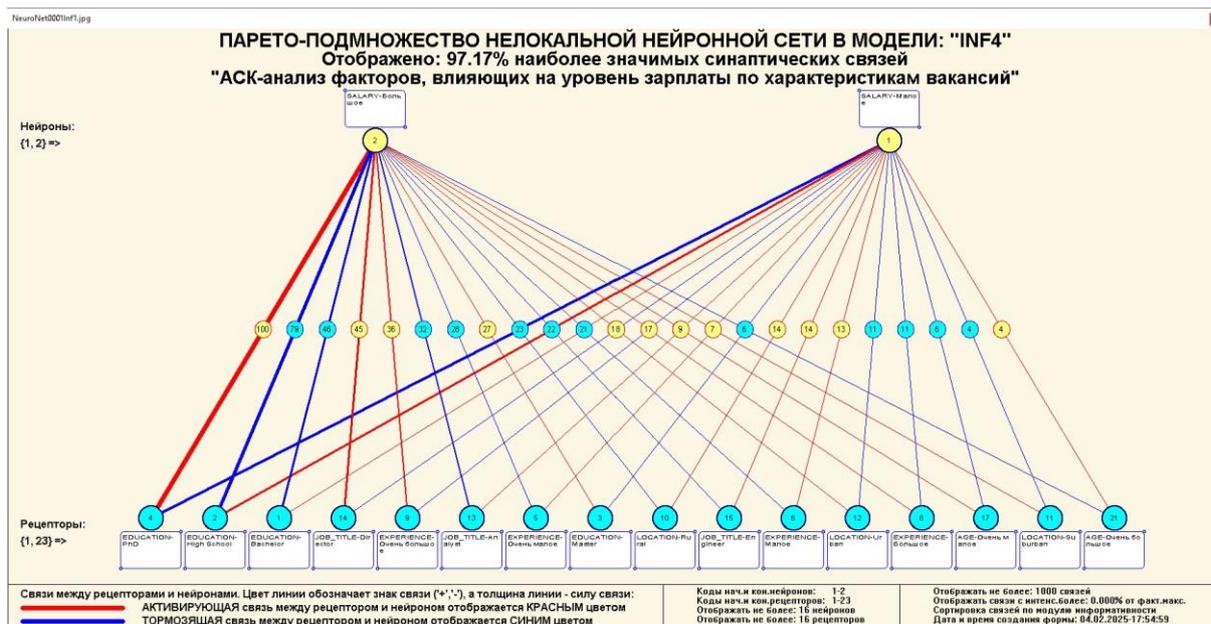


Рисунок 30. Нейронная сеть в СК-модели INF4

3.8.6. 3D-интегральные когнитивные карты

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 24) вверху и когнитивной диаграммы значений факторов (рисунок 28) внизу и соединяющего их одного слоя нейронной сети (рисунок 32) (режим 4.4.12 системы «Эйдос») (рисунок 33).

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

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	SALARY-Малое
<input type="checkbox"/>	2	SALARY-Большое

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

Подготовка визуализации нейрона:1 "SALARY-Малое" в модели:6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
<input checked="" type="checkbox"/>	2 EDUCATION-High School	96.665
<input type="checkbox"/>	1 EDUCATION-Bachelor	55.899
<input type="checkbox"/>	5 EXPERIENCE-Малое	53.032
<input type="checkbox"/>	10 JOB_TITLE-Analyst	39.665
<input type="checkbox"/>	7 LOCATION-Rural	38.135
<input type="checkbox"/>	12 JOB_TITLE-Engineer	24.707
<input type="checkbox"/>	15 AGE-Большое	6.649
<input type="checkbox"/>	16 GENDER-Female	5.372

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
<input checked="" type="checkbox"/>	4 EDUCATION-PhD	-120.867
<input type="checkbox"/>	11 JOB_TITLE-Director	-59.675
<input type="checkbox"/>	6 EXPERIENCE-Большое	-53.032
<input type="checkbox"/>	3 EDUCATION-Master	-31.697
<input type="checkbox"/>	9 LOCATION-Urban	-27.270
<input type="checkbox"/>	8 LOCATION-Suburban	-10.865
<input type="checkbox"/>	14 AGE-Малое	-6.649
<input type="checkbox"/>	17 GENDER-Male	-5.372
<input type="checkbox"/>	13 JOB_TITLE-Manager	-4.697

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

Когн. карта

Сортировать связи:
 по модулю информативности
 по информативности и знаку

Отображать наименования:
 нейронов
 рецепторов

Максимальное количество отображаемых рецепторов:
 Отображать связи с интенсивностью >= % от макс.:

ПАРЕТО-ПОДМНОЖЕСТВО НЕЛОКАЛЬНОЙ ИНТЕГРАЛЬНОЙ КОГНИТИВНОЙ КАРТЫ В МОДЕЛИ: "INF4"
 Отображено: 97.17% наиболее значимых синаптических связей
 "АСК-анализ факторов, влияющих на уровень зарплаты по характеристикам вакансий"

Нейроны:
 {1, 2} =>

Рецепторы:
 {1, 23} =>

Связи между рецепторами и нейронами. Цвет линии обозначает знак связи (*, /, \), а толщина линии - силу связи:
 АКТИВИРУЮЩАЯ связь между рецептором и нейроном отображается КРАСНЫМ цветом
 ТОРМОЗЯЩАЯ связь между рецептором и нейроном отображается СИНИМ цветом

Коды нач. и кон. нейронов: 1-2
 Коды нач. и кон. рецепторов: 1-23
 Отображать не более: 16 нейронов
 Отображать не более: 16 рецепторов

Отображать не более: 1000 связей
 Отображать связи с интенсивностью >= 0,000% от факт. макс.
 Сортировка связей по модулю информативности
 Дата и время создания формы: 04.02.2025 17:55:46

Рисунок 31. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

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

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

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

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

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых, может быть, одним из первых писал Дьердь Пойа. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году в работе на странице 521¹¹. Позже об этом писалось в работе¹² и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

Пример опосредованных правдоподобных рассуждений.

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

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 9409 форм содержательного сравнения классов. Так как каждый из 97 классов сравнивается со всеми остальными, в т.ч. с собой, то всего получается $97 \cdot 96 = 9312$ подобных диаграмм. Естественно, все они не приводятся. Но 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-

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.

приложение №417 и получить в нем все выходные формы, как это описано в данной статье.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

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

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	SALARY-Малое
2	SALARY-Большое

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

Выбор способа фильтрации признаков в информационных портретах когнитивной диаграммы

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

Код	Наименование описательной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ ОПИСАТ	1	17
1	EDUCATION	1	4
2	EXPERIENCE	5	6
3	LOCATION	7	9
4	JOB_TITLE	10	13
5	AGE	14	15

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

Задать модели, в которых проводить расчеты когнитивных диаграмм

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задать max количество отображаемых связей: [Помощь](#)

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Класс для левого инф.портрета: [0] ВСЕ КЛАССЫ
Класс для правого инф.портрета: [0] ВСЕ КЛАССЫ
Описат.шкала для левого инф.портрета: [0] ВСЕ ОПИСАТ
Описат.шкала для правого инф.портрета: [0] ВСЕ ОПИСАТ
Модели, заданные для расчета: Inf4

Задать режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

КОГНИТИВНАЯ ДИАГРАММА КЛАССОВ В МОДЕЛИ: "INF4"
"АСК-анализ факторов, влияющих на уровень зарплаты по характеристикам вакансий"

Кл.шкала: [1] SALARY
Класс: [1] Малое

Сход./разл.классов: 100.000%

Кл.шкала: [1] SALARY
Класс: [1] Малое

Наименования признаков:

[1] EDUCATION [2] High School	Ib=0.614 Ip=61.439 Ic=0.707	[1] EDUCATION [2] High School
[1] EDUCATION [1] Bachelor	Ib=0.358 Ip=35.810 Ic=0.707	[1] EDUCATION [1] Bachelor
[4] JOB_TITLE [13] Analyst	Ib=0.252 Ip=25.211 Ic=0.707	[4] JOB_TITLE [13] Analyst
[2] EXPERIENCE [5] Очень малое	Ib=0.221 Ip=22.091 Ic=0.707	[2] EXPERIENCE [5] Очень малое
[2] EXPERIENCE [9] Очень большое	Ib=0.284 Ip=28.424 Ic=0.707	[2] EXPERIENCE [9] Очень большое
[4] JOB_TITLE [14] Director	Ib=0.352 Ip=35.170 Ic=0.707	[4] JOB_TITLE [14] Director
[1] EDUCATION [4] PhD	Ib=0.780 Ip=78.046 Ic=0.707	[1] EDUCATION [4] PhD

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТ 1-23

Фильтр по оп.шкале: [0] ВСЕ ОПИСАТ 1-23

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

Форма создана: 04.02.2025-17:56:19. Показано количество связей <=999

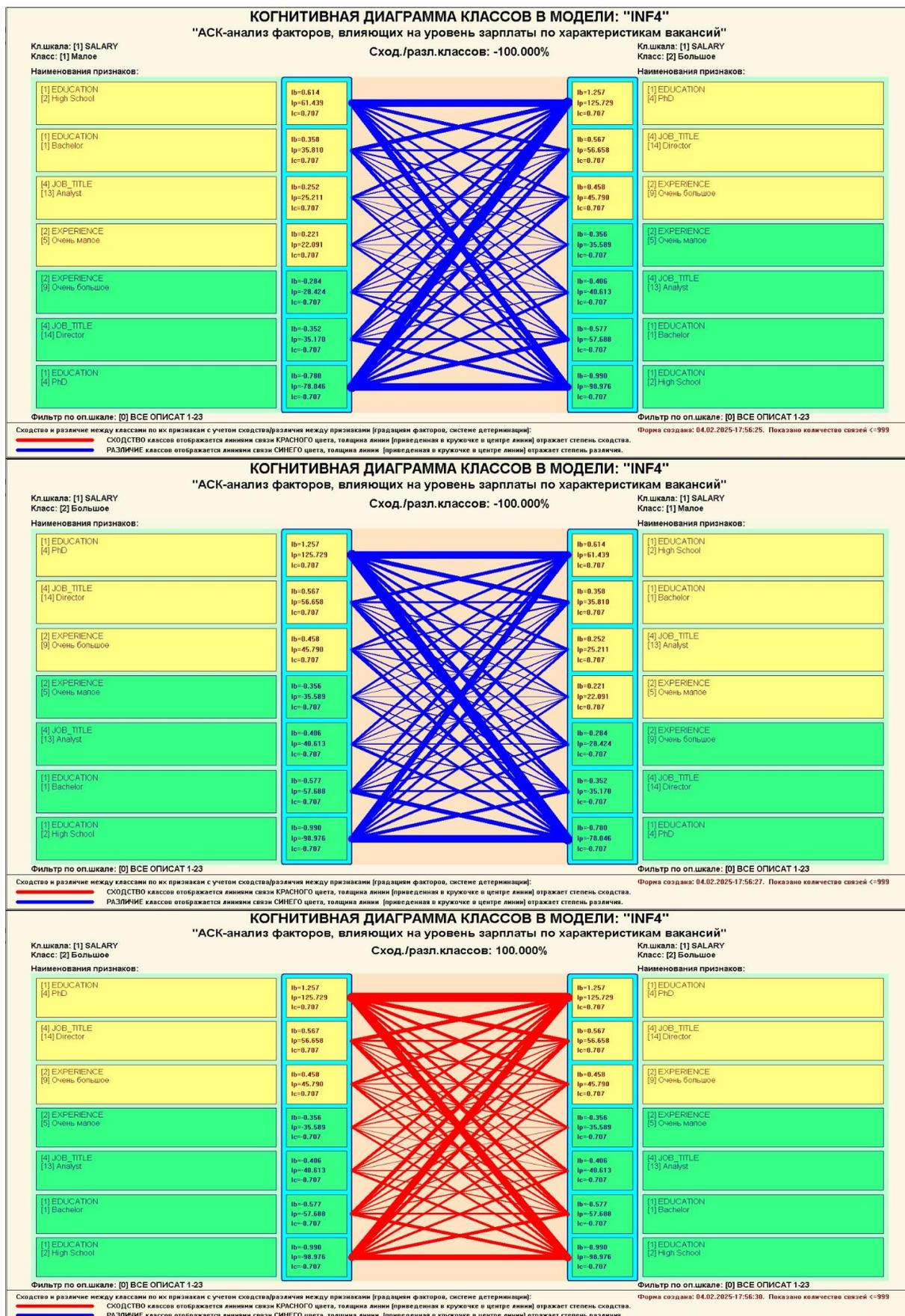


Рисунок 32. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF4

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

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

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

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».

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

4.3.3. Когнитивные диаграммы признаков. Задание параметров генерации выходных форм

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

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	EDUCATION-Bachelor
2	EDUCATION-High School
3	EDUCATION-Master
4	EDUCATION-PhD
5	EXPERIENCE-Малое
6	EXPERIENCE-Большое

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

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

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

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

Задайте модели, в которых проводить расчеты когнитивных диаграмм:
 Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

В диалоге заданы следующие параметры расчета когнитивных диаграмм:
Признак для левого инф.портрета: [0] ВСЕ ПРИЗНАКИ
Признак для правого инф.портрета: [0] ВСЕ ПРИЗНАКИ
Классиф.шкала для левого инф.портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Классиф.шкала для правого инф.портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
Модели, заданные для расчета: Inf4

Задайте режим вывода когнитивных диаграмм:
 Показать все диаграммы с остановкой
 Записать все диаграммы без показа

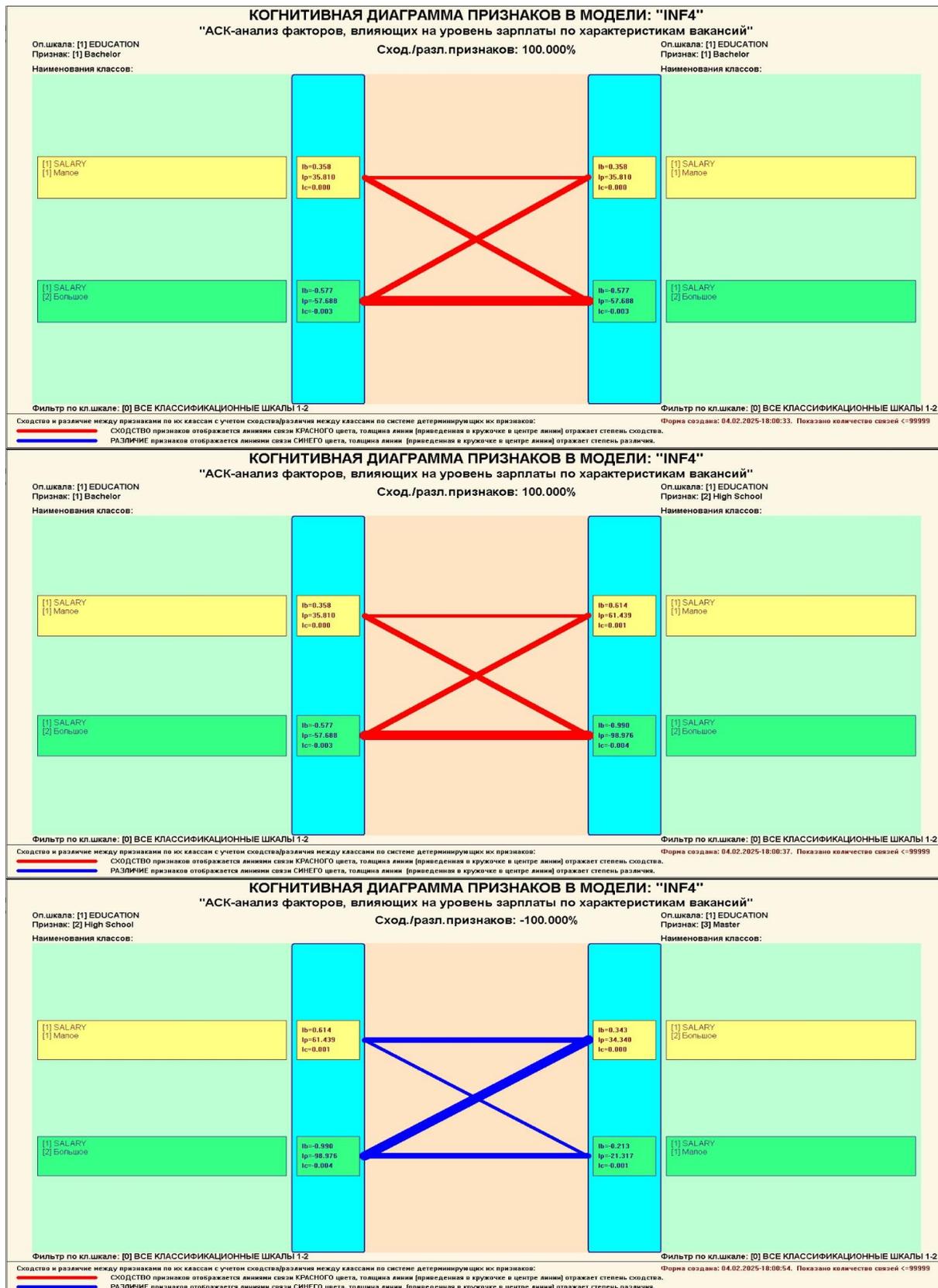


Рисунок 35. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF4

Всего системой в данной модели генерируется $75^2=5625$ подобных диаграмм содержательного сравнения значений факторов по их смыслу, т.е. по влиянию на объект моделирования. Естественно, все они в данной работе не приводятся. Но 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 №417 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

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

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

Когнитивные функции являются одним из наиболее мощных и наглядных средств когнитивной графики, имеющих 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. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах.

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

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

Статистические базы:

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

Системно-когнитивные модели (Базы знаний):

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

Задайте виды когнитивных функций для генерации, визуализации и записи:

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

Задайте дополнительные параметры визуализации когнитивных функций:

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

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

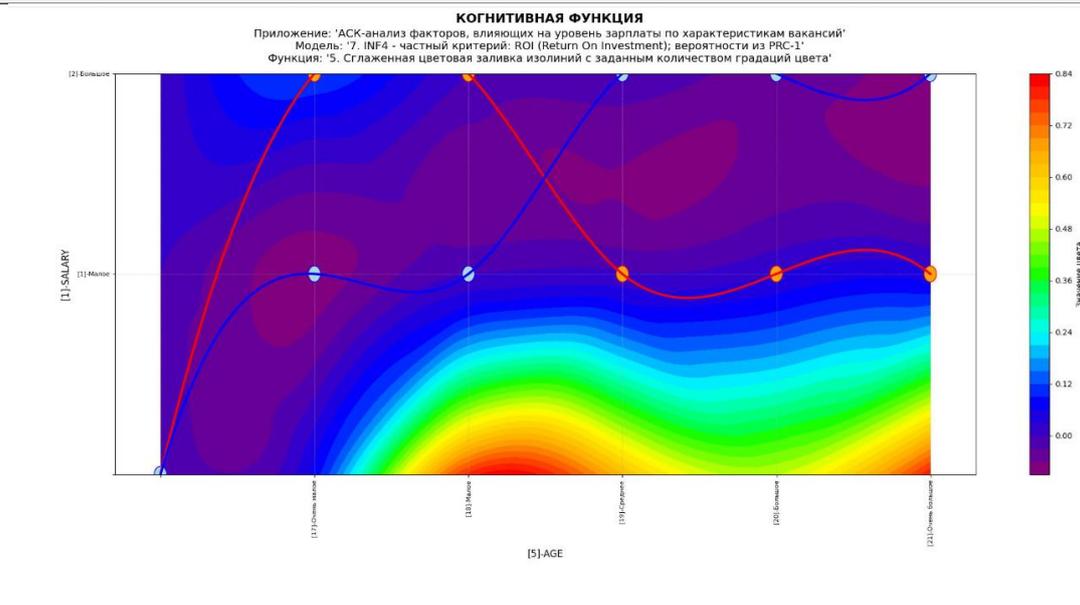
Задайте количество пикселей на дюйм в изображениях когнитивных функций:

Задайте паузу в секундах между визуализациями когнитивных функций:

Задайте размер шрифта для наименований градаций шкал X и Y:

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

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



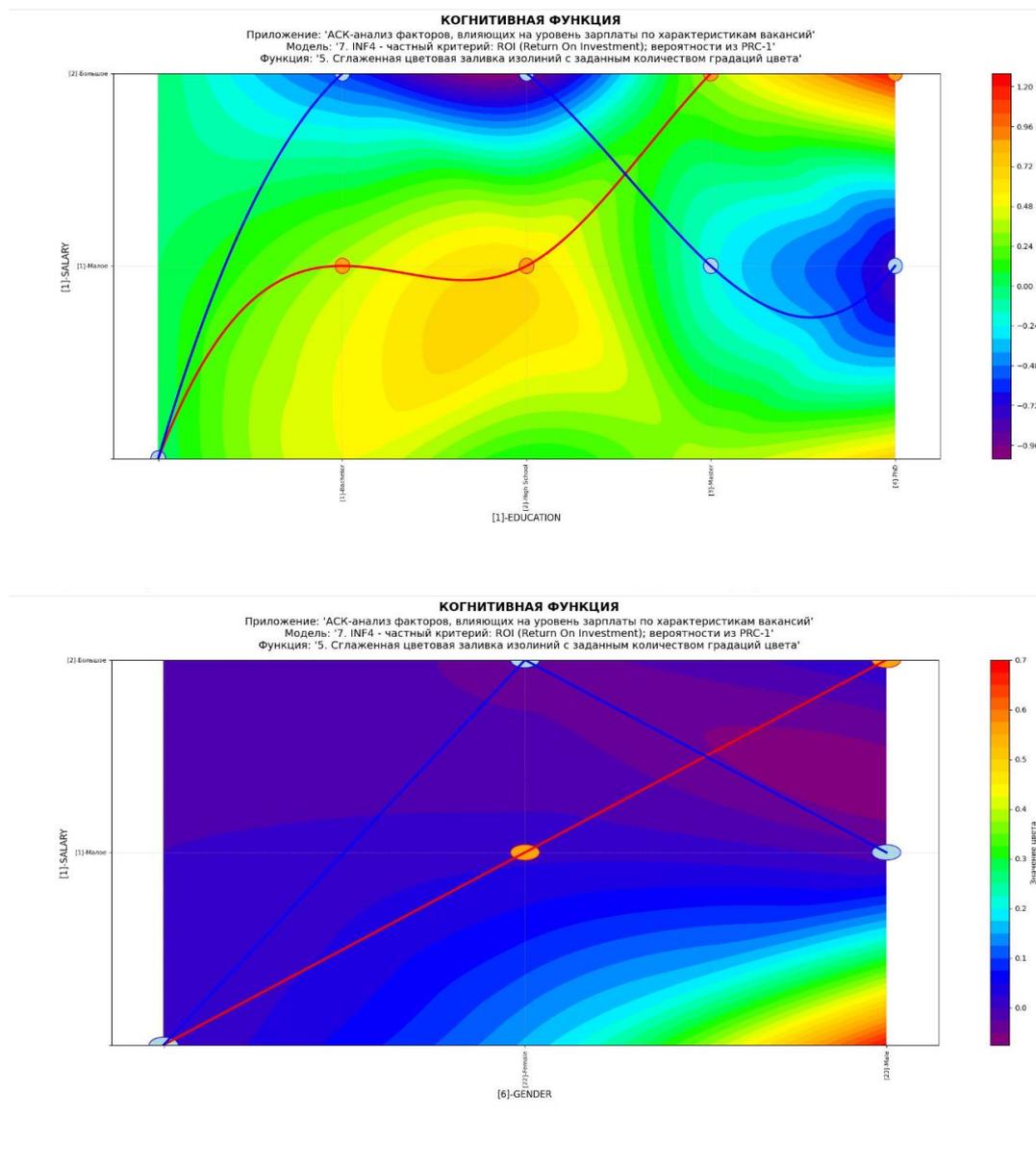


Рисунок 33. Примеры когнитивных функций в СК-модели INF4

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

3.8.10. Значимость описательных шкал и их градаций

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

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это вариабельность информативностей (режим 3.7.5 системы «Эйдос») (рисунок 38):

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

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

На рисунке 37 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4:

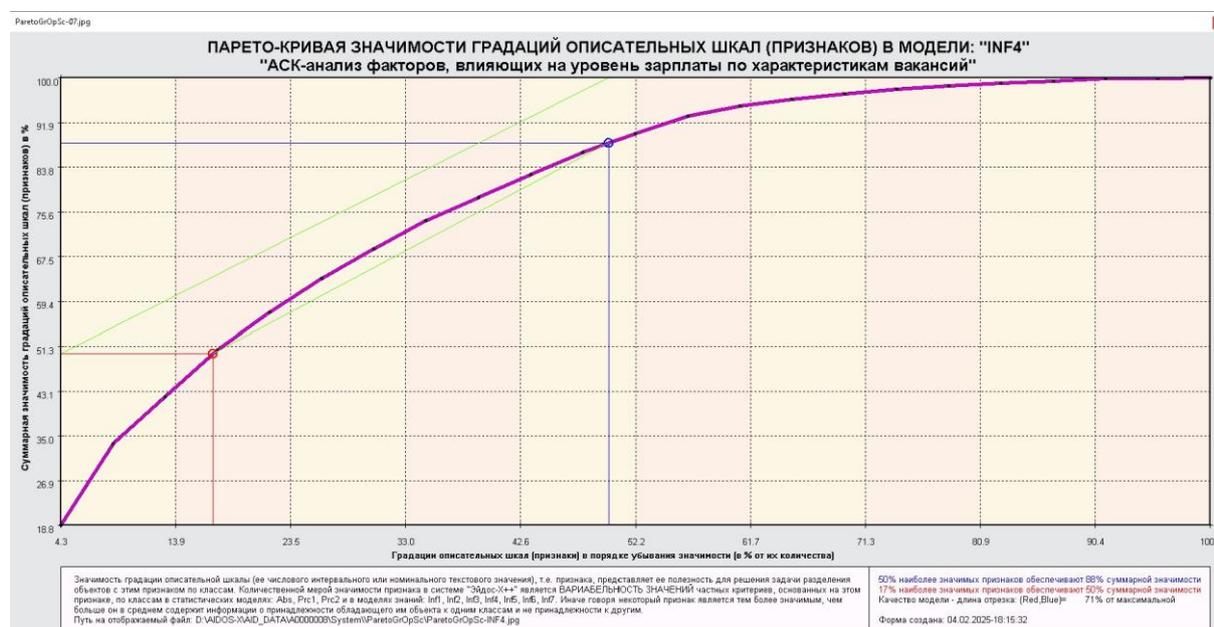


Рисунок 37. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4

Из рисунка 37 видно, что 17% наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 50% суммарного влияния. На рисунке 38 система «Эйдос» привела рейтинг качества статистических и системно-когнитивных моделей. Кроме того на этом рисунке приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях.

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

Таблица 13 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели INF4

NUM	NUM_PRC	KOD_ATR	NAME_ATR	KOD_OPSC	ZNACH_ATR	ZN_ATRNIT	ZNACH_PRC	ZN_PRCNIT
1	4,3478261		4 EDUCATION-PhD	1	1,4409042	1,4409042	18,7798968	18,7798968
2	8,6956522		2 EDUCATION-High School	1	1,1343058	2,5752100	14,7838738	33,5637706
3	13,0434783		1 EDUCATION-Bachelor	1	0,6611265	3,2363365	8,6167335	42,1805040
4	17,3913043		14 JOB_TITLE-Director	4	0,6493231	3,8856596	8,4628949	50,6433989
5	21,7391304		9 EXPERIENCE-Очень большо	2	0,5247716	4,4104312	6,8395640	57,4829629
6	26,0869565		13 JOB_TITLE-Analyst	4	0,4654450	4,8758762	6,0663360	63,5492989
7	30,4347826		5 EXPERIENCE-Очень малое	2	0,4078586	5,2837348	5,3157888	68,8650877
8	34,7826087		3 EDUCATION-Master	1	0,3935520	5,6772868	5,1293250	73,9944127
9	39,1304348		10 LOCATION-Rural	3	0,3307544	6,0080412	4,3108581	78,3052708
10	43,4782609		15 JOB_TITLE-Engineer	4	0,3228387	6,3308799	4,2076895	82,5129603
11	47,8260870		6 EXPERIENCE-Малое	2	0,3049691	6,6358490	3,9747876	86,4877479
12	52,1739130		12 LOCATION-Urban	3	0,2632234	6,8990724	3,4306988	89,9184467
13	56,5217391		8 EXPERIENCE-Большое	2	0,2437355	7,1428079	3,1767050	93,0951517
14	60,8695652		17 AGE-Очень малое	5	0,1363621	7,2791700	1,7772633	94,8724150
15	65,2173913		11 LOCATION-Suburban	3	0,0942348	7,3734048	1,2282009	96,1006158
16	69,5652174		21 AGE-Очень большое	5	0,0818430	7,4552478	1,0666935	97,1673093
17	73,9130435		16 JOB_TITLE-Manager	4	0,0583182	7,5135660	0,7600851	97,9273944
18	78,2608696		20 AGE-Большое	5	0,0533565	7,5669225	0,6954172	98,6228116
19	82,6086957		22 GENDER-Female	6	0,0332117	7,6001342	0,4328617	99,0556733
20	86,9565217		7 EXPERIENCE-Среднее	2	0,0322500	7,6323842	0,4203275	99,4760008
21	91,3043478		23 GENDER-Male	6	0,0311521	7,6635363	0,4060181	99,8820190
22	95,6521739		19 AGE-Среднее	5	0,0083534	7,6718897	0,1088733	99,9908923
23	100,0000000		18 AGE-Малое	5	0,0006988	7,6725885	0,0091077	100,0000000

На экранной форме рисунка 38 приведены имена Excel-файлов с информацией о силе влияния факторов в разных моделях:

3.7.5. Значимость градаций описательных шкал и абстрагирование

— □ ×

— Задайте модель, в которой удалять наименее значимые признаки:
 Отображение Парето-диаграмм значимости признаков завершено!

Результаты расчета силы влияния (значимости) признаков или значений факторов содержатся в следующих базах данных, созданных на основе статистических и интеллектуальных моделей: "Zpr_Abs.xlsx", "Zpr_Prc1.xlsx", "Zpr_Prc2.xlsx", "Zpr_Inf1.xlsx", "Zpr_Inf2.xlsx", "Zpr_Inf3.xlsx", "Zpr_Inf4.xlsx", "Zpr_Inf5.xlsx", "Zpr_Inf6.xlsx", "Zpr_Inf7.xlsx" в папке текущего приложения: D:\AIDOS\X\AID_DATA\A0000005\System\.

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

Сила влияния (значимость) признака или значения фактора представляет собой вариабельность количества информации в этом признаке о переходе объекта моделирования во все будущие состояния, соответствующие классам, имеющимся в модели, т.е. это "жесткость", с которой данное значение фактора обуславливают (детерминируют) переход объекта моделирования в различные состояния, соответствующие классам.

Рейтинг моделей: Задайте модель, в которой удалять наименее значимые признаки:

48.142% 1. ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч.выборки
 61.294% 2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса
 60.912% 3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса
 79.134% 4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1
 79.134% 5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2
 61.294% 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами
 66.164% 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1
 66.164% 8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
 61.294% 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1
 61.294% 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

Задайте какой % наиболее значимых признаков ОСТАВИТЬ в модели:

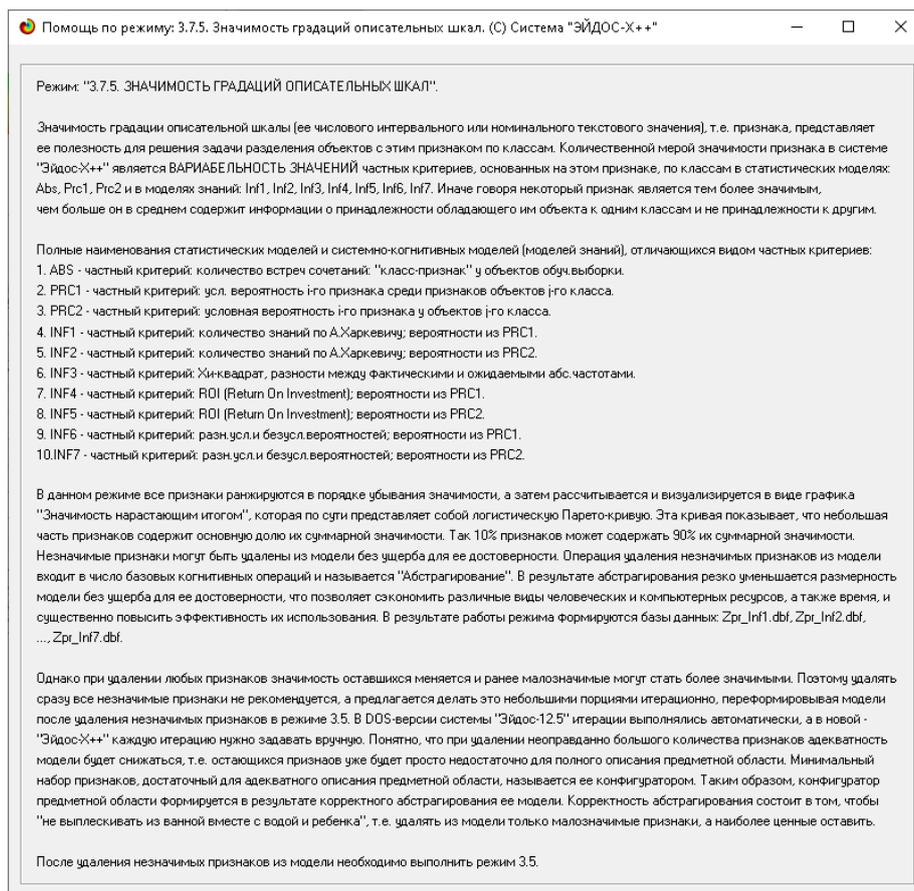


Рисунок 38. Имена Excel-файлов с информацией о силе влияния факторов в статистических и системно-когнитивных моделях

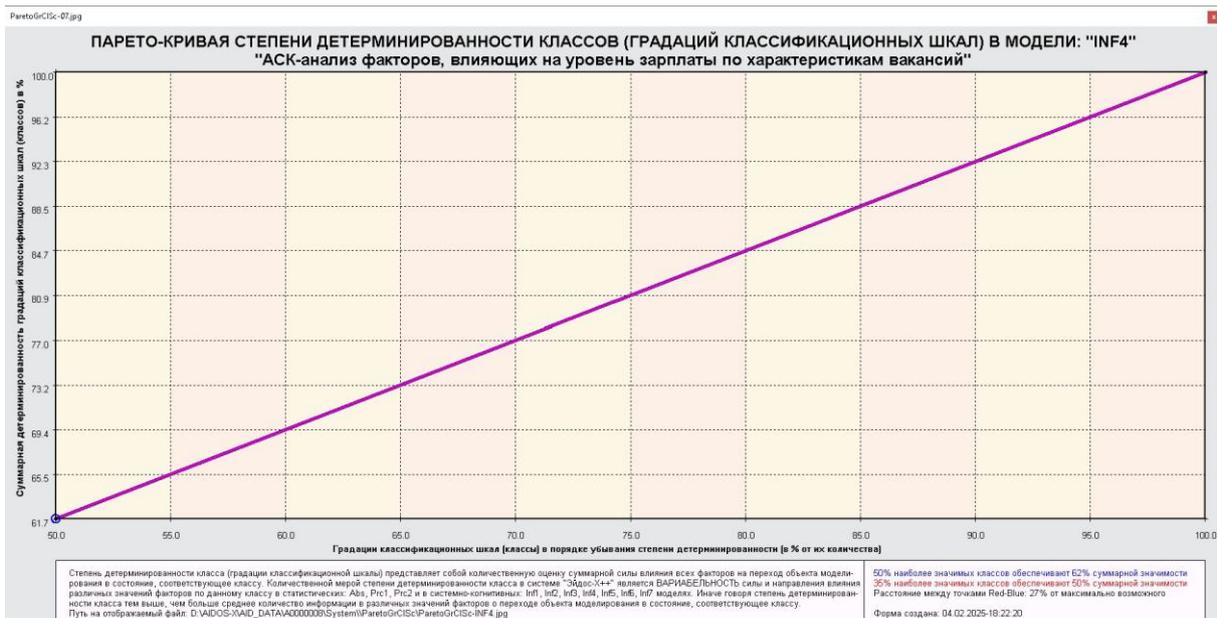
3.8.11. Степень детерминированности классов и классификационных шкал

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

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

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

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



Сообщение об успешном завершении операции

Отображение Парето-диаграмм степени сформированности классов завершено!

Результаты расчета степени детерминированности (значимости) классов содержатся в следующих базах данных, созданных на основе статистических и интеллектуальных моделей: "Zkl_Abs.xlsx", "Zkl_Prc1.xlsx", "Zkl_Prc2.xlsx", "Zkl_Inf1.xlsx", "Zkl_Inf2.xlsx", "Zkl_Inf3.xlsx", "Zkl_Inf4.xlsx", "Zkl_Inf5.xlsx", "Zkl_Inf6.xlsx", "Zkl_Inf7.xlsx" в папке текущего приложения: D:\AID\OS\X\AID_DATA\A0000005\System\.

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

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

Ok

Сообщение об успешном завершении операции

Результаты расчета значимости классификационных шкал содержатся в базах данных статистических и интеллектуальных моделей: "ZCS_Abs.xlsx", "ZCS_Prc1.xlsx", "ZCS_Prc2.xlsx", "ZCS_Inf1.xlsx", "ZCS_Inf2.xlsx", "ZCS_Inf3.xlsx", "ZCS_Inf4.xlsx", "ZCS_Inf5.xlsx", "ZCS_Inf6.xlsx", "ZCS_Inf7.xlsx" в папке текущего приложения: D:\AID\OS\X\AID_DATA\A0000005\System\.

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

Значимость классификационной шкалы является средней значимостью ее градаций, т.е. классов.

Значимость градации классификационной шкалы, т.е. класса, представляет собой вариабельность количества информации в во всех признаках модели о принадлежности или не принадлежности объекта с этим признаками к данному классу.

Значимость градации классификационной шкалы (класса) - это степень детерминированности этого класса (см.режим 3.7.3).

Ok

Рисунок39. 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, соответствующие классам, довольно существенно отличается друг от друга.

Например, 50% наиболее жестко детерминированных классов суммарно обеспечивают примерно 62% степень детерминированности, а 50% суммарной детерминированности обеспечивают 35% наиболее жестко детерминированных классов.

Таблица 14 – Степень детерминированности классов в СК-модели INF4

NUM	NUM_PRC	KOD_CLS	NAME_CLS	KOD_CLSC	ZNACH_CLS	ZN_CLSNIT	ZNACH_PRC	ZN_PRCNIT
1	50,0000000		1 SALARY-Малое		1 617,0000000	617,0000000	61,7000000	61,7000000
2	100,0000000		2 SALARY-Большо		1 383,0000000	1000,0000000	38,3000000	100,0000000

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

Таблица 19 – Степень детерминированности классификационных шкал в системно-когнитивной модели INF4 (фрагмент)

NUM	NUM_PRC	KOD_CLSC	NAME_CLSC	N_GRCLSC	KODGR_MIN
1	100,0000000		1 SALARY		2 1

4. DISCUSSION (ОБСУЖДЕНИЕ)

Результаты, полученные в данной работе, можно считать успешным решением поставленной задачи и достижением целей исследования. Эти результаты были достигнуты благодаря использованию Автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного обеспечения – интеллектуальной системы «Эйдос».

Проведенный анализ результатов полностью соответствует выводам работы, на основе данных которой они были получены. Однако применение АСК-анализа и системы «Эйдос» значительно расширяет возможности решения задач прогнозирования, принятия решений и исследования моделируемой области по сравнению с методами, использованными в работе. Это дает основание рекомендовать использование АСК-анализа и системы «Эйдос» для дальнейших исследований.

Достижения данной работы включают:

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

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

В качестве перспективы дальнейших исследований предлагается увеличить объем исходных данных, количество исследуемых факторов и классификационных шкал, а также их градаций для более точного описания будущих состояний объекта моделирования. Рекомендуется также ввести классификационные шкалы, отражающие влияние факторов не только в натуральном выражении (например, количество и качество продукции), но и в стоимостном (прибыль и рентабельность).

Ценность и перспективность таких исследований очевидна и подтверждается работами в этой области].

Желающие могут изучить данную работу и провести собственные исследования с использованием АСК-анализа и системы «Эйдос». Для этого необходимо скачать систему с сайта разработчика по ссылке: http://lc.kubagro.ru/aidos/_Aidos-X.htm, установить интеллектуальное облачное Эйдос-приложение через диспетчера приложений (режим 1.3). Также доступно большое количество видеоуроков (около 300) по различным аспектам применения технологии, которые можно найти по ссылке: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

В условиях современного рынка труда важно понимать факторы, определяющие уровень зарплат по характеристикам вакансий. Традиционные методы анализа часто не справляются с выявлением сложных взаимосвязей между параметрами вакансий и уровнем оплаты.

В статье предложен подход на основе Автоматизированного системно-когнитивного анализа (АСК-анализа) и системы «Эйдос». Приведен численный пример с таблицами и графиками для обучения работе с АСК-анализом. Особенность задачи – использование как категориальных, так и числовых данных, решаемая через метризацию шкал.

Результаты могут быть использованы для практических рекомендаций участникам рынка труда и научного объяснения механизмов формирования зарплат. Материалы подходят для лабораторных работ и исследований в области экономики труда с применением технологий искусственного интеллекта.

REFERENCES (ЛИТЕРАТУРА)

1. Луценко, Е. В. Исследование влияния подсистем различных уровней иерархии на эмерджентные свойства системы в целом с применением АСК-анализа и интеллектуальной системы "Эйдос" (микроструктура системы как фактор управления ее макросвойствами) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2012. – № 75. – С. 321-363. – EDN OOSCAV.
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHC.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZJXZZ.
4. Работы проф.Е.В.Луценко по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики http://lc.kubagro.ru/aidos/Work_on_emergence.htm
5. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.
6. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
7. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.
8. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
10. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-Х++" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2014. – № 101. – С. 1367-1409. – EDN SZVWRV.
11. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. – 288 с. – DOI 10.13140/RG.2.2.22981.37608. – EDN ZQLITW.
12. Орлов, А. И. Анализ данных, информации и знаний в системной нечеткой интервальной математике / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2022. – 405 с. – ISBN 978-5-907550-62-9. – DOI 10.13140/RG.2.2.15688.44802. – EDN OQULUW.
13. Луценко, Е. В. Автоматизация функционально-стоимостного анализа и метода "Директ-костинг" на основе АСК-анализа и системы "Эйдос" (автоматизация управления натуральной и финансовой эффективностью затрат без содержательных технологических и финансово-экономических расчетов на основе информационных и

когнитивных технологий и теории управления)1 / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.

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

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

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

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

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

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

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