

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

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

Факультет заочного обучения

Кафедра компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ заработной платы специалистов в области науки о
данных»

Выполнил студент группы: ИТ32441 Герасимова Анна Валерьевна

Допущен к защите _____

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

(подпись, расшифровка подписи)



Защищен _____

(дата)

Оценка _____

Краснодар

2026

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

Факультет заочного обучения

РЕЦЕНЗИЯ

на курсовую работу

Студента Герасимовой Анны Валерьевны курса 2

заочной формы обучения группы ИТ32441

Направление подготовки 09.04.02 «Информационные системы и технологии»

Наименование темы «АСК-анализ заработной платы специалистов в области науки о данных»

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	ОТЛИЧНО
2	Уровень и корректность использования в работе различных методов исследований	ОТЛИЧНО
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	ОТЛИЧНО
4	Ясность, четкость, последовательность и обоснованность изложения	ОТЛИЧНО
5	Применение современных технологий обработки информации	ОТЛИЧНО
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	ОТЛИЧНО
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	ОТЛИЧНО
8	Ответы на вопросы при защите	ОТЛИЧНО

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____

Рецензент _____ (Е. В. Луценко) «08» апреля 2026 г.

РЕФЕРАТ

Курсовая работа содержит: 85 страниц, 26 рисунков, 12 таблиц, 24 литературных источника.

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

Целью работы является проведение автоматизированного системно-когнитивного анализа покупателей магазина.

Для достижения цели требуется проанализировать методы создания обобщенных представлений классов и решения задач идентификации объектов с применением методов принятия решений и изучения моделируемой области путем анализа модели.

Автоматизированный системно-когнитивный анализ заработной платы специалистов в области науки о данных

Герасимова Анна Валерьевна
студент факультета заочного
обучения, группы ИТЗ2441
GerasimovaAV_1@yandex.ru

Кубанский государственный аграрный университет имени И.Т.Трубилина, Краснодар, Россия

Целью данной работы является изучение заработной платы специалистов в области науки о данных. Достижение данной цели представляет большой личный интерес. Для меня это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», а также обеспечить некоторую научную ценность в сфере изучения востребованности специалистов в области науки о данных. Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

Ключевые слова:

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, АСК-АНАЛИЗ, СИСТЕМА «ЭЙДОС».

Automated System-Cognitive Analysis of the Salaries of Data Science Specialists

Gerasimova Anna Valerievna
Student of the Correspondence Department, Group
ITZ2441
GerasimovaAV_1@yandex.ru

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

The purpose of this work is to study the salaries of data science specialists. Achieving this goal is of great personal interest. For me, this will allow me to gain knowledge in working with the universal cognitive analytical system "Eidos-X++", as well as provide some scientific value in the field of studying the demand for specialists in data science. To achieve this goal, I use Automated System-Cognitive Analysis (ASC-analysis) and its software toolkit, the intelligent system "Eidos".

Key words: AUTOMATED SYSTEM-COGNITIVE ANALYSIS, ASC-ANALYSIS, "EIDOS" SYSTEM.

ОГЛАВЛЕНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	6
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	6
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ	8
1.4. ЦЕЛЬ РАБОТЫ	10
2. METHODS (МЕТОДЫ)	10
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	10
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	11
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ	11
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	14
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	24
3. RESULTS (РЕЗУЛЬТАТЫ)	27
3.1. Задача-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	27
3.2. Задача-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	29
3.3. Задача-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	38
3.4. Задача-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	53
3.5. Задача-5. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	56
3.5.1. <i>Интегральный критерий «Сумма знаний»</i>	57
3.5.2. <i>Интегральный критерий «Семантический резонанс знаний»</i>	58
3.5.3. <i>Важные математические свойства интегральных критериев</i>	60
3.5.4. <i>Решение задачи идентификации и прогнозирования в системе «Эйдос»</i>	61
3.6. Задача-6. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ	65
3.6.1. <i>Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ</i>	65
3.6.2. <i>Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе</i>	67
3.6.3. <i>АСК-анализа и системы «Эйдос»</i>	67
3.7. Задача-7. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	72
3.7.1. <i>Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)</i>	72
3.7.2. <i>Модель знаний системы «Эйдос» и нелокальные нейроны</i>	73
3.8.5. <i>Нелокальная нейронная сеть</i>	75
3.8.10. <i>Значимость описательных шкал и их градаций</i>	77
3.8.11. <i>Степень детерминированности классов и классификационных шкал</i>	78
4. DISCUSSION (ОБСУЖДЕНИЕ)	80
5. CONCLUSIONS (ВЫВОДЫ)	82
REFERENCES (ЛИТЕРАТУРА)	83

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Технологии искусственного интеллекта стремительно проникают во все сферы человеческой деятельности. От финансов и тяжелой промышленности до здравоохранения, управления персоналом и автоматизированной журналистики — алгоритмы машинного обучения и интеллектуального анализа данных становятся неотъемлемой частью современного мира. В этом контексте особое место занимает сфера управления человеческими ресурсами, где ключевым вызовом является объективная оценка стоимости специалиста. Особенно остро эта проблема стоит в области науки о данных (Data Science) — динамично развивающейся индустрии, где дисбаланс спроса и предложения, а также широкий спектр компетенций создают высокую неопределенность в вопросах оплаты труда.

Традиционные методы анализа заработных плат часто опираются на усредненные показатели или линейные зависимости, что не позволяет в полной мере учесть сложные, нелинейные взаимодействия между различными факторами: уровнем технических навыков (hard skills), soft skills, опытом работы, образованием, владением конкретными инструментами (Python, SQL, TensorFlow) и специализацией (Computer Vision, NLP, анализ данных). Для выявления этих глубинных закономерностей необходим инструмент, способный работать с разнородными данными и выявлять причинно-следственные связи в многомерных системах.

Одним из эффективных подходов к решению подобных задач является автоматизированный системно-когнитивный анализ (АСК-анализ). Данный метод позволяет сегментировать исследуемую аудиторию (в данном случае — специалистов) на основе множества факторов, таких

как уровень квалификации, стек технологий, опыт работы и географическое положение, выявляя их влияние на итоговый уровень дохода. В отличие от классических статистических методов, АСК-анализ позволяет строить когнитивные модели, отражающие реально существующие взаимосвязи в исследуемой предметной области.

Целью данной работы является проведение АСК-анализа факторов, влияющих на уровень заработной платы специалистов в области науки о данных, с целью выявления ключевых закономерностей и построения интеллектуальных моделей, отражающих причинно-следственные взаимосвязи между характеристиками специалиста и его доходом.

Для достижения поставленной цели в работе применяется технология автоматизированного системно-когнитивного анализа, реализованная в интеллектуальной системе «AidosX++». Данный инструментарий позволяет структурировать исходные данные, построить когнитивные модели и визуализировать полученные зависимости для последующей интерпретации.

В рамках исследования предполагается решить следующие задачи:

- провести сбор и структурирование данных о специалистах в области Data Science, включающих информацию об их профессиональных навыках, опыте, образовании и уровне дохода;
- выполнить формализацию предметной области и подготовку данных для загрузки в интеллектуальную систему «AidosX++»;
- осуществить АСК-анализ для выявления основных факторов (причин), оказывающих наибольшее влияние на уровень заработной платы (следствие);
- построить когнитивные модели и проинтерпретировать выявленные зависимости, сформулировав выводы о структуре предпочтений работодателей.

Результаты исследования могут быть полезны как для специалистов по управлению персоналом (HR) с целью разработки прозрачных систем мотивации и удержания сотрудников, так и для начинающих специалистов (Junior-уровень) для планирования вектора профессионального развития и обоснованного позиционирования на рынке труда.

1.2. Объект и предмет исследования

Объект исследования (моделирования) – рынок труда специалистов в области науки о данных.

Предмет исследования – количественные и качественные характеристики специалистов (опыт, навыки, образование), определяющие уровень их заработной платы.

1.3. Проблема, решаемая в работе и ее актуальность

Современный рынок труда в сфере науки о данных (Data Science) характеризуется высокой динамикой, дефицитом квалифицированных кадров и значительным разбросом в уровнях заработных плат. Компании, стремясь привлечь и удержать талантливых специалистов, сталкиваются с необходимостью построения объективных и прозрачных систем оплаты труда. Однако многообразие факторов, влияющих на доход специалиста, — от стека технологий и глубины экспертизы до наличия soft skills и опыта работы в конкретных индустриях — создает существенные сложности при принятии кадровых решений.

Проблема заключается в отсутствии системного подхода к анализу взаимосвязей между характеристиками специалиста и его заработной платой. Часто работодатели ориентируются на усредненные рыночные показатели или субъективные оценки, что приводит к неэффективному расходованию бюджета на персонал, несправедливой дифференциации оплаты и, как следствие, снижению мотивации сотрудников. Для

начинающих специалистов (junior-уровня) эта проблема усугубляется недостатком информации о том, какие именно компетенции и личные качества в наибольшей степени влияют на стартовый доход и карьерный рост.

АСК-анализ (автоматизированный системно-когнитивный анализ) позволяет решить данную проблему за счет выявления многомерных, нелинейных зависимостей между набором признаков специалиста (образование, опыт, навыки, сертификации, знание инструментов) и его уровнем дохода. В отличие от классических статистических методов, АСК-анализ дает возможность строить когнитивные модели, отражающие реальные причинно-следственные связи, и количественно оценивать вклад каждого фактора в итоговый результат.

Актуальность данного исследования обусловлена необходимостью создания инструментов для обоснованного управления вознаграждением в сфере Data Science. Результаты АСК-анализа могут быть использованы:

- HR-специалистами и руководителями — для разработки прозрачных грейдов и систем мотивации, соответствующих реальной ценности компетенций;
- соискателями и начинающими специалистами — для осознанного планирования профессионального развития и понимания того, какие навыки позволят максимизировать доход;
- образовательными учреждениями — для актуализации учебных программ под реальные запросы рынка.

Таким образом, решение поставленной проблемы позволит повысить эффективность взаимодействия на рынке труда, снизить неопределенность при определении заработной платы и способствовать более справедливому распределению кадровых ресурсов в высокотехнологичной отрасли.

1.4. Цель работы

Целью работы является изучение и анализ факторов, определяющих уровень заработной платы специалистов в области науки о данных, с использованием АСК-анализа для выявления влияния таких характеристик, как профессиональные навыки, опыт работы, уровень образования и владение инструментарием, на величину дохода.

Для достижения поставленной цели в работе предполагается решить ряд задач, связанных с выбором и применением методов анализа данных, обработкой информации о специалистах и интерпретацией полученных результатов. Конкретные задачи и этапы исследования будут сформулированы после обоснования методики, применяемой для решения проблемы.

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 здесь 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.

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)

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

АСК-анализ включает:

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

Более подробно математический метод АСК-анализа описан в работе [3] и ряде других [5]. Около половины из 688 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано более 40 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 33 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

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

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

наукам и в 8 кандидатских диссертациях по экономическим, техническим, психологическими и медицинским наукам, еще несколько докторских и кандидатских диссертаций с применением АСК-анализа находятся в стадии выхода на защиту.

Автор является основателем междисциплинарной научной школы: «Автоматизированный системно-когнитивный анализ»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;
- Автоматизированный системно-когнитивный анализ текстовых данных;
- Спектральный и контурный автоматизированный системно-когнитивный анализ изображений;
- Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов.

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [5] и страничка в РесечГейт [6] и РИНЦ [7], на которых можно получить более

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

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

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.

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

2.4. Система «Эйдос» – инструментарий АСК-анализа

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

Полезные нейросети и приложения для разных сфер:

Для дизайнеров: SiteKick - нейросеть для создания лендингов; AdCreative - делает рекламные креативы, плакаты; Looka - логотипы по текстовому описанию; Watermarkremover - поможет удалить вотермарки;

Booth ai - создает стоковые фотки по описанию; PatternedAI - паттерны по текстовому описанию; Нама - вырезать лишние элементы с фото или картинки; RoomGPT - «примеряет» новый ремонт на вашу квартиру, помогает выбрать дизайн;

Для фотографов: ; Pallete fm - раскрашивает черно-белые изображения; Relight - меняет светотень на фотографиях; Photoroom - вырезать элементы из фото, поменять фон; LeiaPix - сделает из 2D-фотки 3D.; Nostalgia Photo - улучшает качество старых фото; pfrmaker - генератор аватарок для соцсетей; Picsart - заменяет или удаляет ненужные элементы на фото;

Для тех, кто монтирует видео:; CapCut - удобный редактор, доступен в браузере. Есть цветокорр, разные эффекты; vidyo ai - нарезать видео на короткие фрагменты; Reface - изменить лицо человека на видео; Runwayml - самые разные инструменты для монтажа; Colourlab AI - нейросеть для цветокоррекции; Topaz Video AI - сильно улучшит качество видео, уберет шум и трясущийся экран; Luma AI - сделает 3D изображение из серии фото; Simplified - анимация картинки; SpiritMe - твоя цифровая копия в сети;

Для звукарей и музыкантов; ; Mubert - создаёт музыку любого жанра; Beatoven - ИИ-композитор музыки для видео; Clip audio - подберет музыку для любого видоса; Fadr - порежет трек на отдельные дорожки инструментов и вокала; Adobe Enhance - чистит запись от шумов. Бесплатно; Elevenlabs - мощнейший синтезатор, подделает любой голос; The MetaVoice - меняйте свой голос на один из восьми пресетов; Cleanvoice - уберет из вашей разговорной записи мусор; ;

Для айтишников; ; CodePal - пишет код с нуля, исправляет ошибки, оценивает готовый код; Codesnippets - создает код по текстовому запросу; Buildt AI - поисковик для VSCode, найдет готовый код в инете; Code GPT - плагин-генератор кода для VSCode; Autobackend - автоматический бэкэнд;

Adrenaline - ищет и помогает чинить ошибки в коде; Tabnine - дописывает код, если у тебя не получается; ;

Для школьников и студентов; ; Consensus - база научных статей; ExamCram - превратит сложные учебные материалы в карточки и тесты для самопроверки; MathGPT - решает задачи по математике; editGPT - исправляет ошибки в английском ; Yip - то же самое, но в вебе и с поддержкой Википедии; ChatBA - делает презентации за тебя; YouTube Summary with ChatGPT - конвертирует видео или лекции в текст; Explain Me Like I'm Five - объясняет сложные научные термины простым языком;

Для тех, кто ищет работу: InterviewGPT AI - задает каверзные вопросы и помогает готовиться к собеседованию; Resume Worded - улучшает резюме; kickresume - сделает крутое резюме и напишет мотивационное письмо; Cover Letter AI - написать сопроводительный текст к резюме; ;

Для тех, кому не помог Гугл: Chord - напишет реферат в ответ на запрос в строке; Lexii ai - бот, который умеет ссылаться на источники; Perplexity - нейросеть-поисковик в виде расширения для браузера; Nuclia - поиск по облаку или серверу; Phind - умеет искать код, поможет айтишникам; ;

Для отдыха и развлечения: RadioGPT - радио, где музыку генерируют нейронки; EndlessVN - бесконечная визуальная новелла; Natural Language Playlist - подберет плейлист на 7 часов специально для тебя; Movie Deep Search - найдет фильм по запросу; FashionAdvisor AI - советы от нейро-стилиста; Hello History - с помощью нее пообщаешься с историческим персонажем; Cool Gift Ideas - выберет подарок для человека по его описанию; Endel - нейро-музыка, которая помогает засыпать; PlaylistAI - соберет плейлист в Apple и Spotify по тексту или картинке.; Tattoos AI - делает эскизы для татуировок.

И все не смотря на настоящую революцию в области искусственного интеллекта и связанный с ней бум Универсальная когнитивная аналитическая система «Эйдос-X++» 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*);

- находится в полном открытом бесплатном доступе (http://lc.kubagro.ru/aidos/_Aidos-X.htm), причем с актуальными исходными текстами (http://lc.kubagro.ru/_AidosALL.txt): открытая лицензия: CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>), и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора и разработчика системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 32 свидетельства РосПатента РФ);

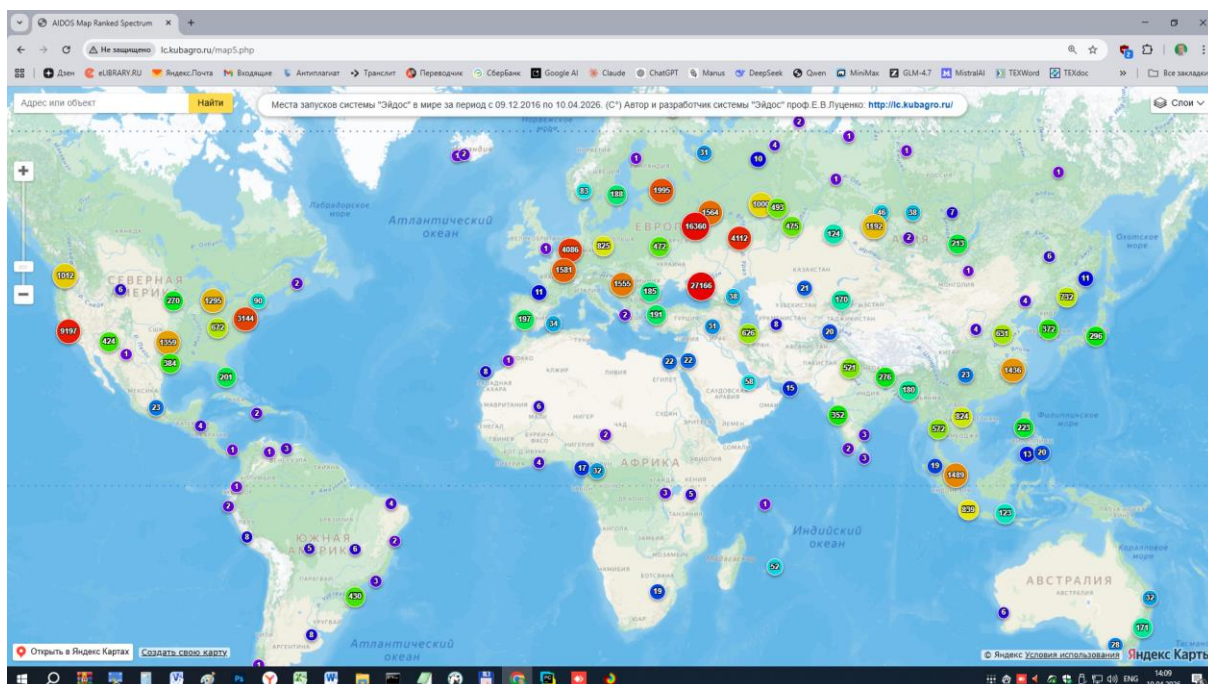
- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта:

«имеет нулевой порог входа» (есть акт внедрения системы «Эйдос» 1987 года) (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);

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

- имеет «нулевой порог входа», содержит большое количество локальных (поставляемых с инсталляцией) и облачных интеллектуальных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 390, 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);

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

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

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

достижения целей (т.е. для управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

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

В 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 внедрения).

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»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке xBase++eXpress++Advantage Database Server (ADS), обеспечивающей обработку больших данных, информации и

знаний (Big Data, Big Information, Big Knowledge), а также на языке Питон (Python). Практически все новые режимы системы «Эйдос» и новые реализации старых режимов будут осуществляться на языке Питон.

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

Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения⁷

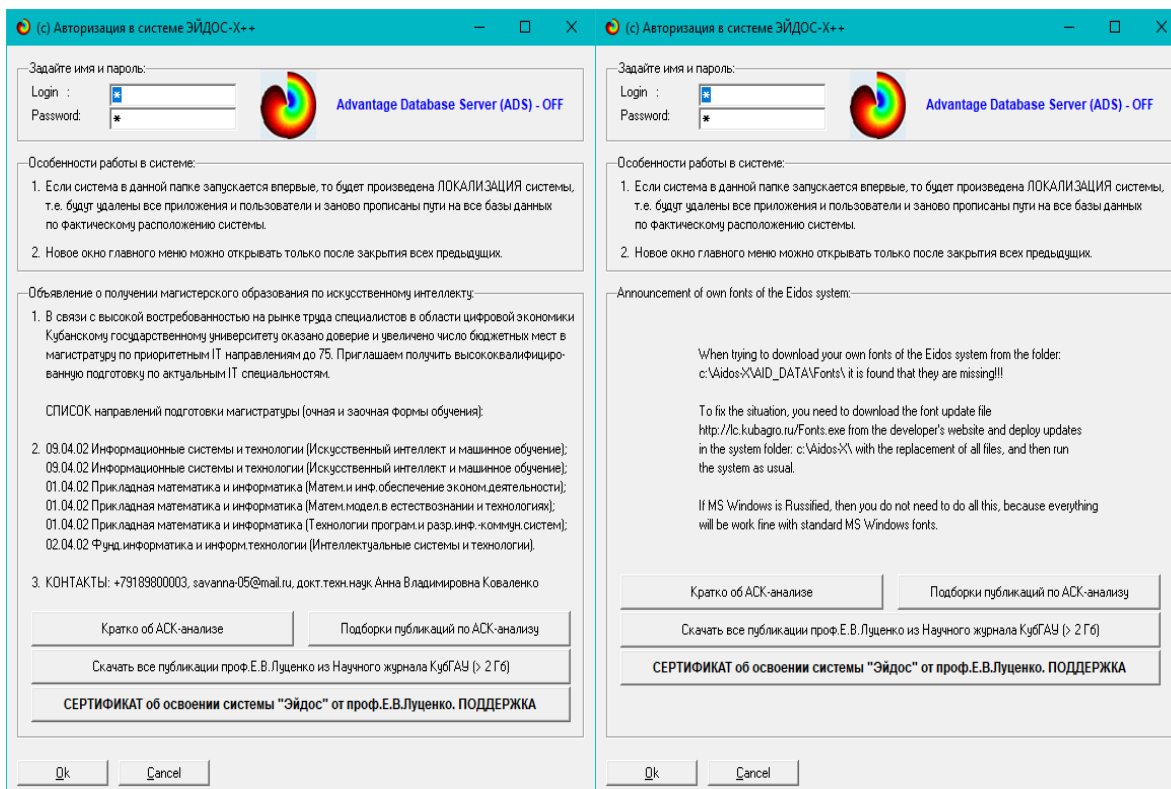
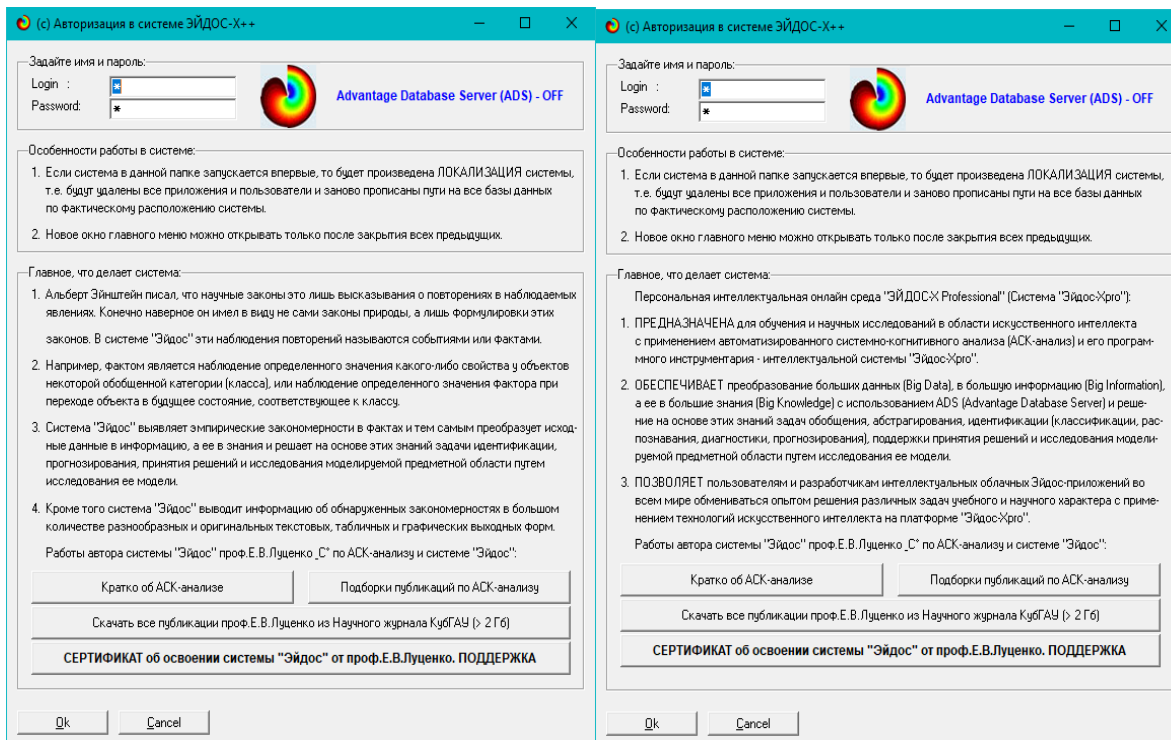
На рисунке 1 приведена титульная видеограмма DOS-версии системы «Эйдос», а на рисунке 2 – текущей версии системы «Эйдос»:



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

⁷ [http://lc.kubagro.ru/aidos/How to make your own cloud Eidos-application.pdf](http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf)

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



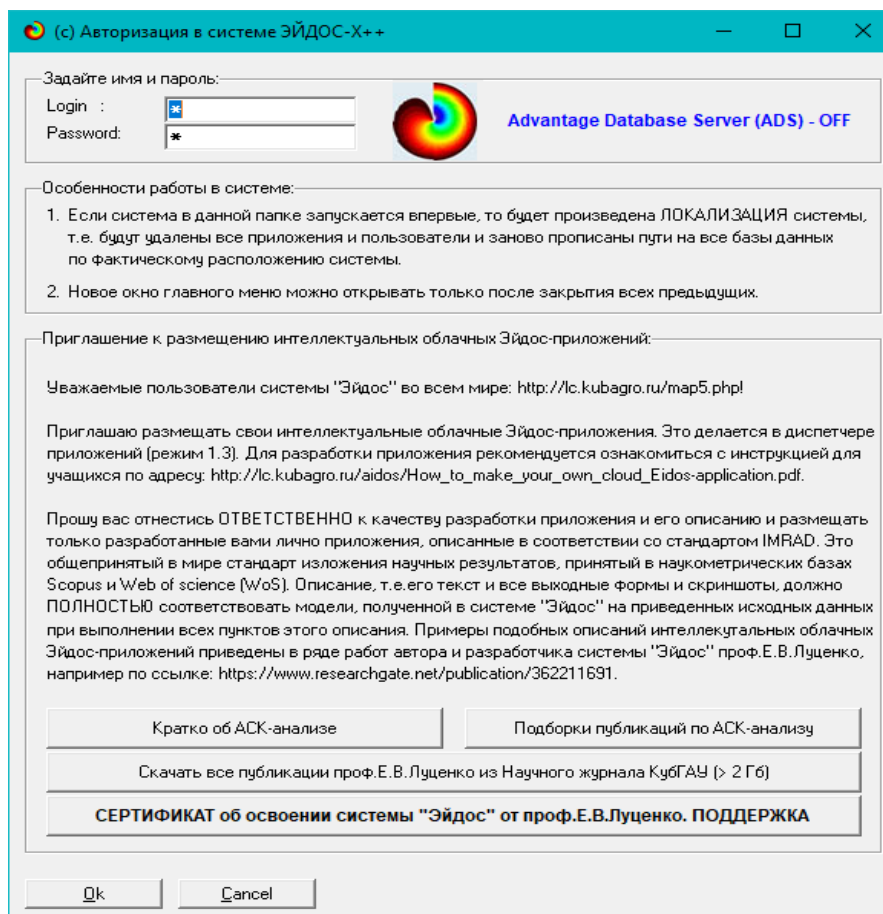


Рисунок 2. Титульные видеogramмы текущей версии системы «Эйдос»

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, т.к. она позволяет детально исследовать влияние каждого значения каждого фактора на уровень заработной платы специалистов в области науки о данных.

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

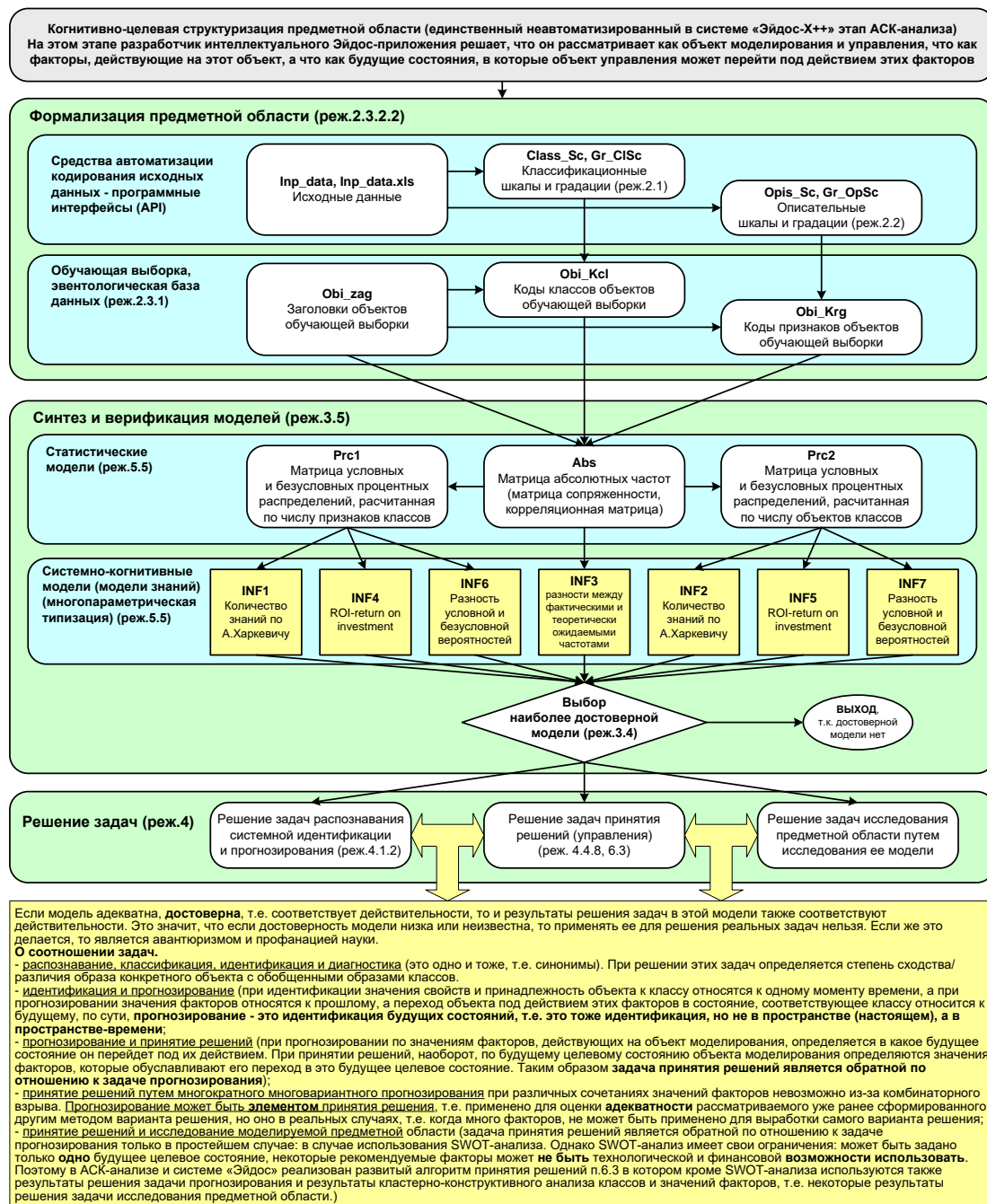


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе

«Эйдос»

3. RESULTS (РЕЗУЛЬТАТЫ)

3.1. Задача-1. Когнитивная структуризация предметной области.

Две интерпретации классификационных и описательных шкал и градаций

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

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

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

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

Статичная интерпретация и терминология:

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

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

Динамичная интерпретация и терминология:

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

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

Обобщающая терминология:

– классификационные шкалы и градации;

– описательные шкалы и градации.

3.1.2. Конкретное решение задачи в данной работе

Исходные данные были взяты по ссылке <https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries>.

Данные содержат информацию о заработных платах специалистов в области науки о данных, включая год выплаты, уровень опыта, тип занятости, должность, зарплату, страну расположения компании и размер компании.

В данной работе в качестве объекта моделирования выступают специалисты в области Data Science, в качестве факторов: год выплаты зарплаты (*work_year*), уровень опыта (*experience_level*), тип занятости (*employment_type*), должность (*job_title*), страна расположения компании (*company_location*) и размер компании (*company_size*) (таблица 1), а в качестве результатов действия этих факторов – уровень заработной платы в долларах США (*salary_in_usd*) (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

1	1
2	2
3	3
4	4
5	5
6	6
7	7

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

Таблица 2 – Классификационные шкалы (результаты действия факторов)

1	1
2	2

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

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

3.2. Задача-2. Формализация предметной области

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 почти везде, где человек применяет естественный интеллект.

3.2.1. Конкретное решение задачи в данной работе

Таблица 1 – Исходные данные по заработным платам специалистов в области науки о данных

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		_work_year	experience_level	employment_type	job_title	salary	salary_currency	salary_in_usd	employee_residence	remote_ratio	company_location	company_size							
2	0,2020	MI	FT	Data Scientist	70000	EUR	79833	DE	0	DE	L								
3	1,2020	SE	FT	Machine Learning Scientist	260000	USD	260000	JP	0	JP	S								
4	2,2020	SE	FT	Big Data Engineer	85000	GBP	109024	GB	50	GB	M								
5	3,2020	MI	FT	Product Data Analyst	20000	USD	20000	HN	0	HN	S								
6	4,2020	SE	FT	Machine Learning Engineer	150000	USD	150000	US	50	US	L								
7	5,2020	EN	FT	Data Analyst	72000	USD	72000	US	100	US	L								
8	6,2020	SE	FT	Lead Data Scientist	190000	USD	190000	US	100	US	S								
9	7,2020	MI	FT	Data Scientist	11000000	HUF	35735	HU	50	HU	L								
10	8,2020	MI	FT	Business Data Analyst	135000	USD	135000	US	100	US	L								
11	9,2020	SE	FT	Lead Data Engineer	125000	USD	125000	NZ	50	NZ	S								
12	10,2020	EN	FT	Data Scientist	45000	EUR	51321	FR	0	FR	S								
13	11,2020	MI	FT	Data Scientist	3000000	INR	40481	IN	0	IN	L								
14	12,2020	EN	FT	Data Scientist	35000	EUR	39916	FR	0	FR	M								
15	13,2020	MI	FT	Lead Data Analyst	87000	USD	87000	US	100	US	L								
16	14,2020	MI	FT	Data Analyst	85000	USD	85000	US	100	US	L								
17	15,2020	MI	FT	Data Analyst	8000	USD	8000	PK	50	PK	L								
18	16,2020	EN	FT	Data Engineer	4450000	JPY	41689	JP	100	JP	S								
19	17,2020	SE	FT	Big Data Engineer	100000	EUR	114047	PL	100	GB	S								
20	18,2020	EN	FT	Data Science Consultant	423000	INR	5707	IN	50	IN	M								
21	19,2020	MI	FT	Lead Data Engineer	56000	USD	56000	PT	100	US	M								

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

Таблица 2 – Таблица исходных данных в стандарте системы «Эйдос» (фрагмент)

id	experience_level	job_title	company_location	work_year	employment_type	company_size	salary_in_usd
0	MI	Data Scientist	DE	2020	FT	L	79833
1	SE	Machine Learning Scientist	JP	2020	FT	S	260000
2	SE	Big Data Engineer	GR	2020	FT	M	109024
3	MI	Product Data Analyst	HN	2020	FT	S	20000
4	SE	Machine Learning Engineer	US	2020	FT	L	150000
5	EN	Data Analyst	US	2020	FT	L	72000
6	SE	Lead Data Scientist	FR	2020	FT	S	190000
7	MI	Data Scientist	HU	2020	FT	L	35735
8	MI	Business Data Analyst	US	2020	FT	L	135000
9	SE	Lead Data Engineer	NZ	2020	FT	S	125000
10	EN	Data Scientist	FR	2020	FT	S	51321
11	MI	Data Scientist	IN	2020	FT	L	40481
12	EN	Data Scientist	DE	2020	FT	M	39916
13	MI	Lead Data Analyst	JP	2020	FT	L	87000
14	MI	Data Analyst	GR	2020	FT	L	85000
15	MI	Data Analyst	HN	2020	FT	L	8000
16	EN	Data Engineer	US	2020	FT	S	41689
17	SE	Big Data Engineer	US	2020	FT	S	114047
18	EN	Data Science Consultant	FR	2020	FT	M	5707
19	MI	Lead Data Engineer	HU	2020	FT	M	56000
20	MI	Machine Learning Engineer	US	2020	FT	M	43331
21	MI	Product Data Analyst	NZ	2020	FT	L	6072

Таблица 4 имеет следующую структуру:

- каждая строка описывает одного специалиста, всего их 607;
- каждое наблюдение описывается одновременно двумя способами: с одной стороны, значениями факторов, действующих на объект моделирования (лингвистические и числовые переменные, градации описательных шкал), а с другой стороны – результатами действия этих

факторов, выраженными в числовой шкале. Такая структура описания наблюдений в технологиях искусственного интеллекта называется «онтологией» и в модели представлений знаний Марвина Мински (1975) называется «фрейм-экземпляр»;

- 1-я колонка – не является шкалой и содержит номер наблюдения (индекс) для идентификации записи;

- колонка 8 (*salary_in_usd*) – это классификационная шкала числового типа, описывающая результат действия факторов – уровень заработной платы в долларах США. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в натуральном и стоимостном выражении: например, уровень удовлетворённости, карьерный рост и так далее. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

- колонки со 2-й по 7-ую (*work_year, experience_level, employment_type, job_title, company_location, company_size*) – это описательные шкалы, формализующие факторы, действующие на объект моделирования. Эти шкалы имеют числовой и текстовый тип, и их градациями являются лингвистические и числовые переменные;

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

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

типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

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

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

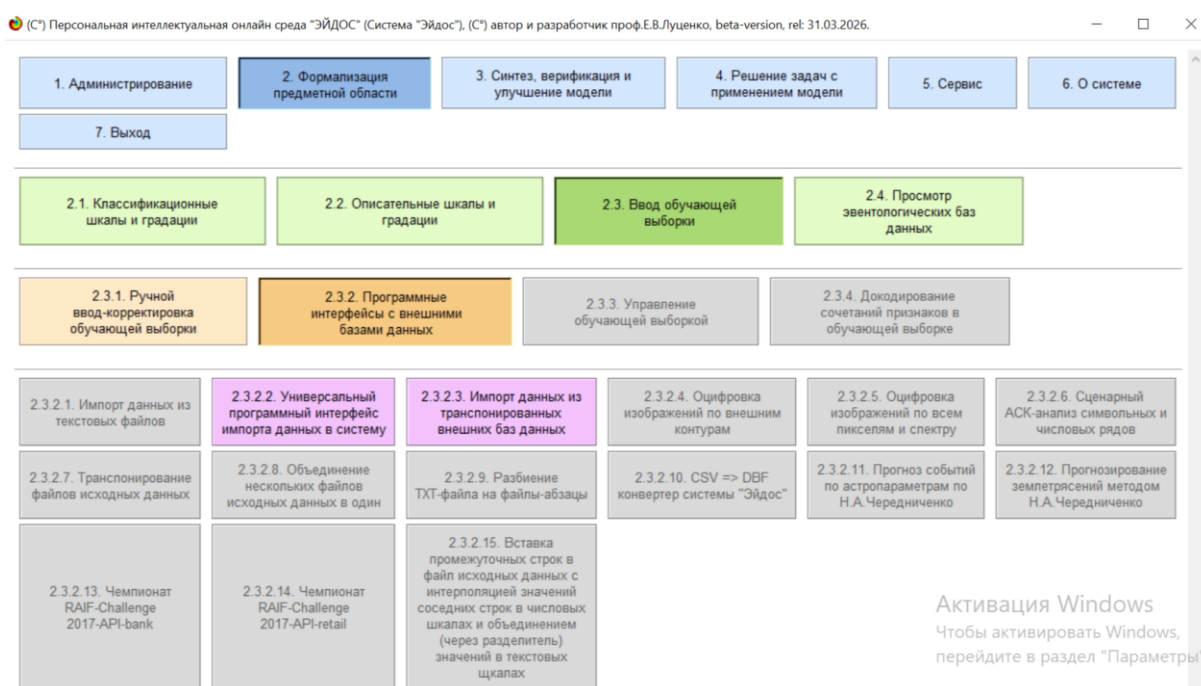


Рисунок 4. Автоматизированные программные интерфейсы (API) системы «Эйдос»

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 7):

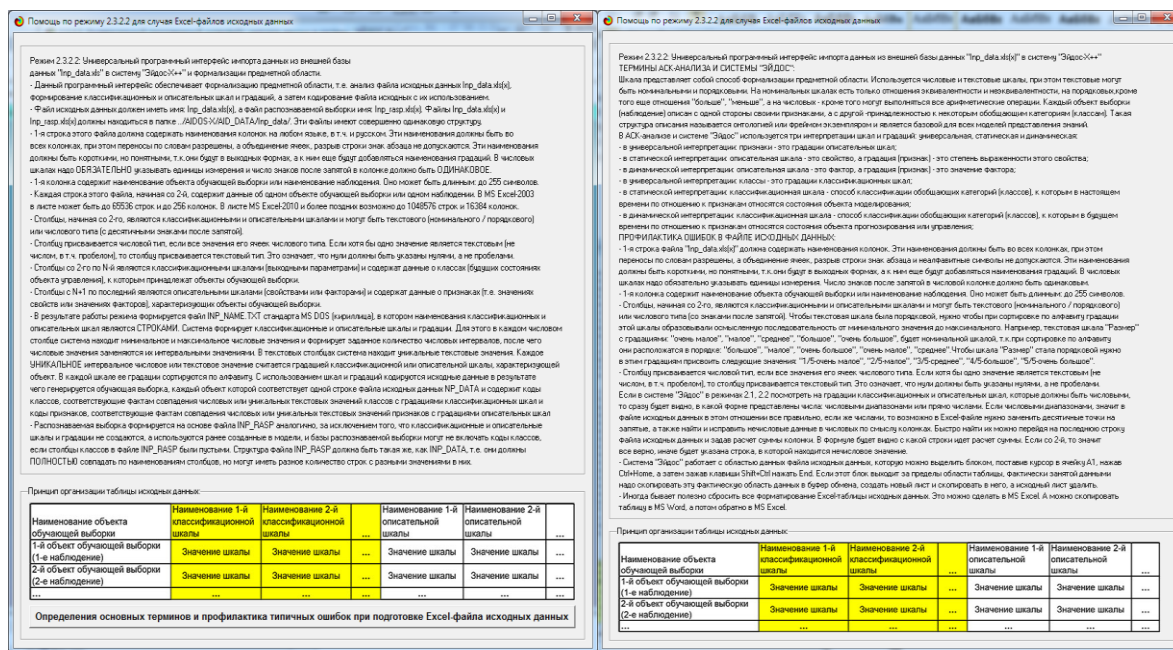


Рисунок 5. Хелпы API-2.3.2.2 системы «Эйдос»

Экранные формы управления API-2.3.2.2 системы «Эйдос» с *реальными параметрами*, использованными в данной работе, приведены на рисунках 8.

На 5, 6, 7 приведены классификационные и описательные шкалы и градации, а также обучающая выборка, сформированные API-2.3.2.2 при параметрах, показанных на рисунках 8.

Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал и между шкалами, как числовыми, так и текстовыми. Математическая модель АСК-анализа позволяет корректно преодолеть несбалансированность данных путем перехода от

абсолютных частот к относительным и к количественным мерам знаний в системно-когнитивных моделях (мы увидим это ниже).

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": <input type="radio"/> XLS - MS Excel-2003 Стандарт XLS-файла <input checked="" type="radio"/> XLSX - MS Excel-2007(2010) Стандарт XLS-файла <input type="radio"/> DBF - DBASE IV (DBF/NTX) Стандарт DBF-файла <input type="radio"/> CSV => DBF конвертер Стандарт CSV-файла		Задайте параметры: <input checked="" type="radio"/> Нули и пробелы считать ОТСУТСТВИЕМ данных <input type="radio"/> Нули и пробелы считать ЗНАЧЕНИЯМИ данных <input type="checkbox"/> Создавать БД средних по классам "Inp_davr.dbf"? <div style="border: 1px solid black; background-color: #e0e0ff; padding: 2px; text-align: center;">Требования к файлу исходных данных</div>	
Задайте диапазон столбцов классификационных шкал: Начальный столбец: <input type="text" value="8"/> Конечный столбец: <input type="text" value="8"/>		Задайте диапазон столбцов описательных шкал: Начальный столбец: <input type="text" value="4"/> Конечный столбец: <input type="text" value="7"/>	
Задайте режим: <input checked="" type="radio"/> Формализации предметной области (на основе "Inp_data") <input type="radio"/> Генерации распознаваемой выборки (на основе "Inp_rasp")		Задайте способ выбора размера интервалов: <input checked="" type="radio"/> Равные интервалы с разным числом наблюдений <input type="radio"/> Разные интервалы с равным числом наблюдений	
Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp_data": <input checked="" type="radio"/> Не применять сценарный метод АСК-анализа <input type="radio"/> Применить сценарный метод АСК-анализа <input type="checkbox"/> Применить спец.интерпретацию текстовых полей классов <input type="checkbox"/> 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: <input checked="" type="radio"/> Только интервальные числовые значения (например: "1/3-(59873.0000000, 178545.6666667)") <input type="radio"/> Только наименования интервальных числовых значений (например: "Минимальное") <input type="radio"/> И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-(59873.0000000, 178545.6666667)")			

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [3 классов x 44 признака]

Тип	Количество	Суммарное количество	Среднее количество	Количество	Суммарное	Среднее
Числовые	1	3	3.00	1	2	2.00
Текстовые	0	0	0.00	3	42	14.00
ВСЕГО:	1	3	3.00	4	44	11.00

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

ПАРАМЕТРЫ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ

ПАРАМЕТРЫ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ

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

КЛАССИФИКАЦИОННАЯ ШКАЛА: код: [1], наим.: "SALARY_IN_USD", тип/число градаций в шкале: "равные интервалы"/3

- Наим.градации: 1/3-{ 2859.0000000, 201906.0000000 }, размер интервала=199047.0000000, расч./факт.число наблюдений на градацию: 202/554
- Наим.градации: 2/3-{ 201906.0000000, 400953.0000000 }, размер интервала=199047.0000000, расч./факт.число наблюдений на градацию: 202/46
- Наим.градации: 3/3-{ 400953.0000000, 600000.0000000 }, размер интервала=199047.0000000, расч./факт.число наблюдений на градацию: 202/7

ОПИСАТЕЛЬНАЯ ШКАЛА: код: [1], наим.: "COMPANY_LOCATION", тип шкалы/число градаций в шкале: "равные интервалы"/35

- Наим.градации: 1/35-AE, факт.число наблюдений на градации: 2
- Наим.градации: 2/35-AI, факт.число наблюдений на градации: 1
- Наим.градации: 3/35-AE, факт.число наблюдений на градации: 1
- Наим.градации: 4/35-BR, факт.число наблюдений на градации: 1
- Наим.градации: 5/35-CA, факт.число наблюдений на градации: 24
- Наим.градации: 6/35-CN, факт.число наблюдений на градации: 6
- Наим.градации: 7/35-CL, факт.число наблюдений на градации: 1
- Наим.градации: 8/35-CO, факт.число наблюдений на градации: 1
- Наим.градации: 9/35-CZ, факт.число наблюдений на градации: 2
- Наим.градации: 10/35-DE, факт.число наблюдений на градации: 15
- Наим.градации: 11/35-DK, факт.число наблюдений на градации: 1
- Наим.градации: 12/35-DZ, факт.число наблюдений на градации: 2
- Наим.градации: 13/35-ES, факт.число наблюдений на градации: 8
- Наим.градации: 14/35-FR, факт.число наблюдений на градации: 37

1/7 - Конвертация и ввод файла исходных данных	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
2/7 - Создание нового интеллектуального приложения	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
3/7 - Анализ файла исходных данных	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
4/7 - Формирование классификационных шкал и градаций	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
5/7 - Формирование описательных шкал и градаций	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
6/7 - Кодирование обучающей выборки	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%
7/7 - Формирование базы событий	<div style="width: 100%; height: 10px; background-color: green;"></div>	100%

Общий прогресс

Начало: 01.04.2026 22:20:19 Окончание: 01.04.2026 22:20:52
 Прошло: 00:00:32 Осталось: 00:00:00

Сообщения

```
[22:20:42 INFO]: 5. Формирование описательных шкал и градаций
[22:20:43 INFO]: Сформировано 4 описательных шкал и 44 градаций.

[22:20:43 INFO]: 6. Кодирование обучающей выборки
[22:20:50 INFO]: Таблицы obi_zag, obi_kol, obi_kpr успешно заполнены для 607 объектов.

[22:20:50 INFO]: 7. Формирование базы событий
[22:20:51 INFO]: База событий eventsko успешно заполнена. Записано 607 объектов.

[22:20:52 SUCCESS]: Операция успешно завершена.
```

[Записать лог в CSV](#) [Выход](#)

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

Таблица 3 – Классификационные шкалы и градации

Классификационные шкалы и градации (Приложение: Приложение, созданное в API-2.3.2.2-20260401-222020)

Код	Наименование классификационной шкалы	Код	Наименование градации классификационной шкалы
	salary_in_usd	1	1/3-{2859.0000000, 201906.0000000}
		2	2/3-{201906.0000000, 400953.0000000}
		3	3/3-{400953.0000000, 600000.0000000}

Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".

Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Удаление и перекодирование | Справочник классов | Графики будущих сценариев | Выход

Таблица 4 – Описательные шкалы и градации

2.2. Описательные шкалы и градации (Приложение: Приложение, созданное в API-2.3.2.2-20260401-222020)

Код	Наименование описательной шкалы	Код	Наименование градации описательной шкалы
1	company_location	42	L
2	work_year	43	M
3	employment_type	44	S
4	company_size		

Активация Windows
Чтобы активировать Windows, перейдите в раздел "Параметры".

Помощь | Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Перекодировать | Очистить | Справочник признаков | Графики прошлых сценариев | Выход

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 [1-4]. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко, акцентируя внимание лишь на математической взаимосвязи коэффициента возврата инвестиций (ROI) с мерой χ -квадрат Карла Пирсона и с семантической мерой целесообразности информации Александра Харкевича.

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

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике [1-7] и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых (нелинейных) данных, представленных в

различных типах шкал (дихотомических, номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот (таблица 7).

Таблица 7 – Матрица абсолютных частот (статистическая модель ABS)

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

На основе таблицы 7 рассчитываются матрицы условных и безусловных процентных распределений (таблица 8).

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

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

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

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

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

На практике часто встречается существенная несбалансированность данных, под которой понимается сильно отличающееся количество наблюдений объектов обучающейся выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 7) было бы очень неразумно 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 7), а матрицы условных и безусловных процентных распределений (таблица 8), 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 7 и 8 с использованием частных критериев, знаний приведенных таблице 9, рассчитываются матрицы 7 системно-когнитивных моделей (таблица 10).

В таблице 10 приведены формулы:

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

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

Таблица 9– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	Через относительные частоты	Через абсолютные частоты
<p>ABS, матрица абсолютных частот, N_{ij} - фактическое число встреч i-го признака у объектов j-го класса;</p> <p>\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</p>	<p>N_{ij} – фактическая частота;</p> $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота.}$	

всей выборке (таблица 7)		
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	---	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу		
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_i N_j}$
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 10 – Матрица системно-когнитивной модели

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

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 10), применяемых в настоящее время в системе «Эйдос» равно 7 определяется тем, что они получаются путем всех возможных вариантов сравнения фактических и теоретических абсолютных частот, условных и безусловных относительных частот путем вычитания и путем деления, и при этом N_j рассматривается как суммарное количество или признаков, или объектов обучающей выборки в j -м классе, а **нормировка к нулю** (для аддитивных интегральных критериев), если нет связи между наличием признака и принадлежностью объекта к классу, осуществляется либо логарифмированием, либо вычитанием единицы (таблица 11).

**Таблица 5– 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^2 -квадрат Карла Пирсона	---	---
Сравнение условных и безусловных относительных частот	Путем деления	---	INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
	Путем вычитания	INF6, INF7	---	---

Обратим особое внимание на то, что сравнение фактических и теоретических абсолютных частот путем деления приводит при нормировках к нулю (что нужно для применения аддитивных интегральных критериев) путем взятия логарифма и путем вычитания 1 к *тем же самым* моделям, что и сравнение условных и безусловных относительных частот путем деления с теми же самыми способами нормировки. Таким образом, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и провести нормировку к 0 путем взятия логарифма и путем вычитания 1, то получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Других же системно-когнитивных моделей, рассчитываемых на основе приведенных статистических моделей просто нет. Это и есть конфигуратор

статистических и когнитивных моделей в смысле В.А.Лефевра. *Под конфигуратором В.А.Лефевр понимал минимальный полный набор понятийных шкал или конструктов, т.е. понятий, достаточный для адекватного описания предметной области* [4]⁹. Необходимо отметить, что все эти модели рассчитываются в интеллектуальной системе «Эйдос».

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

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

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

Вероятность рассматривается как предел, к которому стремится относительная частота (отношение количества благоприятных исходов к числу испытаний) при неограниченном увеличении количества испытаний. Ясно, что вероятность – это математическая абстракция, которая никогда не встречается на практике (также как и другие математические и физические абстракции, типа математической точки, материальной точки, бесконечно малой и т.п.). На практике встречается только относительная

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

частота. Но она может быть весьма близкой к вероятности. Например, при 480 наблюдений различие между относительной частотой и вероятностью (погрешность) составляет около 5%, при 1250 наблюдениях – около 2.5%, при 10000 наблюдениях – 1%.

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

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

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

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 12):

Таблица 12 – Уточнение терминологии АСК-анализа

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

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

получающиеся в результате сравнения фактических и теоретических абсолютных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1 **тождественно совпадают** с моделями, получающимися путем сравнения условных и безусловных относительных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1. Это и есть конфигурактор статистических и когнитивных моделей в смысле В.А.Лефевра, содержащий минимальное количество моделей, позволяющих полно описать моделируемую предметную область.

Показательно, 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):

3.5. Синтез и верификация моделей

Задайте базовые модели для синтеза и верификации

Базовые статистические модели:

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

Задайте источник данных для расчета модели ABS:

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

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

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

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

Базовые системно-когнитивные модели (базы знаний):

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

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

6. INF3 - частный критерий: Хи-квадрат; разности между фактическими и ожидаемыми абс. частотами

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

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

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

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

Какие модели создавать?

Создавать только базовые модели

Создавать модели 2-го уровня

Создавать модели 3-го уровня

[Помощь](#)

Базовые модели - это модели: Prc1,Prc2,Inf1,Inf2,Inf3, Inf4,Inf5,Inf6,Inf7 полученные расчетным путем по формулам, приведенным в хелпе режима 5.5 на основе модели Abs.

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

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

Параметры копирования обучающей выборки в распознаваемую (бутстрепный подход)

Какие объекты обуч.выборки копировать:

Копировать всю обучающую выборку

Копировать только текущий объект

Копировать каждый N-й объект

Копировать N случайных объектов

Копировать объекты от N1 до N2 (fastest) до

Вообще не менять распознаваемую выборку

Удалять скопированные объекты:

Не удалять

Удалять

[Пояснение по алгоритму верификации](#)

[Подробнее](#)

Выполнить:

Синтез и верификацию

Только верификацию

Только синтез

Задайте процессор:

CPU

GPU

Задайте алгоритм:

Классика - дольше

Упрощенно-быстрее

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

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

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

3.5. Синтез и верификация моделей. Процесс завершен

Подготовительные шаги

- Очистка предыдущих результатов верификации 100%
- Копирование обучающей выборки в распознаваемую 100%

Стадия синтеза модели			Стадия верификации модели			
№	Наименование модели	прогресс-бар	%	№	прогресс-бар	%
3	1. ABS - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ВСТРЕЧ СОЧЕТАНИЙ: "КЛАСС-ПРИЗНАК" У ОБЪЕКТОВ ОБУЧ	<div style="width: 100%;"></div>	100%	13	<div style="width: 100%;"></div>	100%
4	2. PRC1 - ЧАСТНЫЙ КРИТЕРИЙ: УСЛ.ВЕРОЯТНОСТЬ I-ГО ПРИЗНАКА СРЕДИ ПРИЗНАКОВ ОБЪЕКТОВ J-ГО	<div style="width: 100%;"></div>	100%	14	<div style="width: 100%;"></div>	100%
5	3. PRC2 - ЧАСТНЫЙ КРИТЕРИЙ: УСЛОВНАЯ ВЕРОЯТНОСТЬ I-ГО ПРИЗНАКА У ОБЪЕКТОВ J-ГО КЛАССА	<div style="width: 100%;"></div>	100%	15	<div style="width: 100%;"></div>	100%
6	4. INF1 - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ЗНАНИЙ ПО А.ХАРКЕВИЧУ; ВЕРоятности ИЗ PRC1	<div style="width: 100%;"></div>	100%	16	<div style="width: 100%;"></div>	100%
7	5. INF2 - ЧАСТНЫЙ КРИТЕРИЙ: КОЛИЧЕСТВО ЗНАНИЙ ПО А.ХАРКЕВИЧУ; ВЕРоятности ИЗ PRC2	<div style="width: 100%;"></div>	100%	17	<div style="width: 100%;"></div>	100%
8	6. INF3 - ЧАСТНЫЙ КРИТЕРИЙ: ХИ-КВАДРАТ; РАЗНОСТИ МЕЖДУ ФАКТИЧЕСКИМИ И ОЖИДАЕМЫМИ АБС.ЧАСТОТАМИ	<div style="width: 100%;"></div>	100%	18	<div style="width: 100%;"></div>	100%
9	7. INF4 - ЧАСТНЫЙ КРИТЕРИЙ: ROI (RETURN ON INVESTMENT); ВЕРоятности ИЗ PRC1	<div style="width: 100%;"></div>	100%	19	<div style="width: 100%;"></div>	100%
10	8. INF5 - ЧАСТНЫЙ КРИТЕРИЙ: ROI (RETURN ON INVESTMENT); ВЕРоятности ИЗ PRC2	<div style="width: 100%;"></div>	100%	20	<div style="width: 100%;"></div>	100%
11	9. INF6 - ЧАСТНЫЙ КРИТЕРИЙ: РАЗН.УСЛИ БЕЗУСЛ.ВЕРоятностей; ВЕРоятности ИЗ PRC1	<div style="width: 0%;"></div>	0%	21	<div style="width: 0%;"></div>	0%
12	10. INF7 - ЧАСТНЫЙ КРИТЕРИЙ: РАЗН.УСЛИ БЕЗУСЛ.ВЕРоятностей; ВЕРоятности ИЗ PRC2	<div style="width: 100%;"></div>	100%	22	<div style="width: 100%;"></div>	100%

Общий прогресс

Начало: 01.04.2026 22:25:32 Окончание: 01.04.2026 22:26:06

Прошло: 00:00:34 Осталось: 00:00:00

Сообщения

```
[22:26:05 SUCCESS]: Агрегированные метрики для модели 'INF7' сохранены.
[22:26:05 INFO]: Накопление статистики DostRasp для модели 'INF7'...
[22:26:05 INFO]: Расчет метрик VerModClsIt для модели 'INF7' (пропущено).

[22:26:05 INFO]: Финальная обработка и сортировка таблиц dost_models...
[22:26:05 SUCCESS]: Таблица dost_modcls finalized и отсортирована.

[22:26:05 INFO]: ШАГ 8: формирование итоговой ЕД DostRasp...
[22:26:05 INFO]: формирование итоговой таблицы DostRasp...
[22:26:06 SUCCESS]: Таблица DostRasp успешно создана.

[22:26:06 INFO]: ШАГ 10: Создание форм 'Достоверность идент. классов'...
[22:26:06 INFO]: Создание форм 'Достоверность идент. классов' (пропущено).

[22:26:06 SUCCESS]: Операция успешно завершена.
```

Рисунок 7. Экранные формы режима синтеза и верификации моделей

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

5.5. Просмотр основных баз данных всех моделей

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

Модель: "Abs"

Код признака	Описательные шкалы и градации (признаки)	1 - salary_in_usd - (13/2859.000000, 20196.000000)			2 - salary_in_usd - (23/20196.000000, 40953.000000)			3 - salary_in_usd - (33/40953.000000, 60000.000000)		
		Сумма	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего					
1	company_location - AE	2.000			2.000	0.667	1.155	0.889		
2	company_location - AT	1.000			1.000	0.333	0.577	0.444		
3	company_location - BE	1.000			1.000	0.333	0.577	0.444		
4	company_location - BR	1.000			1.000	0.333	0.577	0.444		
5	company_location - CA	23.000	1.000		24.000	8.000	13.000	10.000		
6	company_location - CH	6.000			6.000	2.000	3.464	2.667		
7	company_location - CL	1.000			1.000	0.333	0.577	0.444		
8	company_location - CO	1.000			1.000	0.333	0.577	0.444		
9	company_location - CZ	2.000			2.000	0.667	1.155	0.889		
10	company_location - DE	15.000			15.000	5.000	8.660	6.667		
11	company_location - DK	1.000			1.000	0.333	0.577	0.444		
12	company_location - DZ	2.000			2.000	0.667	1.155	0.889		
13	company_location - ES	8.000			8.000	2.667	4.619	3.556		
14	company_location - FR	35.000	2.000		37.000	12.333	19.655	15.111		
15	company_location - GB	45.000			45.000	15.000	25.981	20.000		
16	company_location - GR	15.000			15.000	5.000	8.660	6.667		
17	company_location - HN	14.000			14.000	4.667	8.083	6.222		

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

5.5. Просмотр основных баз данных всех моделей

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

Модель: "Prc2"

Код признака	Описательные шкалы и градации (признаки)	1 - salary_in_usd - 1/3-(2859.0000000, 201906.0000000)			Безусловная вероятность	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	company_location - AE	0.361			0.329	0.120	0.208	0.160
2	company_location - AT	0.181			0.165	0.060	0.104	0.080
3	company_location - BE	0.181			0.165	0.060	0.104	0.080
4	company_location - BR	0.181			0.165	0.060	0.104	0.080
5	company_location - CA	4.152	2.174		3.954	2.109	2.077	1.406
6	company_location - CH	1.083			0.988	0.361	0.625	0.481
7	company_location - CL	0.181			0.165	0.060	0.104	0.080
8	company_location - CO	0.181			0.165	0.060	0.104	0.080
9	company_location - CZ	0.361			0.329	0.120	0.208	0.160
10	company_location - DE	2.708			2.471	0.903	1.563	1.203
11	company_location - DK	0.181			0.165	0.060	0.104	0.080
12	company_location - DZ	0.361			0.329	0.120	0.208	0.160
13	company_location - ES	1.444			1.318	0.481	0.834	0.642
14	company_location - FR	6.318	4.348		6.096	3.555	3.233	2.370
15	company_location - GB	8.123			7.414	2.708	4.690	3.610
16	company_location - GR	2.708			2.471	0.903	1.563	1.203
17	company_location - HN	2.527			2.306	0.842	1.459	1.123

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

5.5. Просмотр основных баз данных всех моделей

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

Модель: "Inf1"

Код признака	Описательные шкалы и градации (признаки)	1 - salary_in_usd - 1/3-(2859.0000000, 201906.0000000)			Сумма	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
1	company_location - AE	0.019			0.019	0.006	0.011	0.008
2	company_location - AT	0.019			0.019	0.006	0.011	0.008
3	company_location - BE	0.019			0.019	0.006	0.011	0.008
4	company_location - BR	0.019			0.019	0.006	0.011	0.008
5	company_location - CA	0.010	-0.122		-0.112	-0.037	0.073	0.056
6	company_location - CH	0.019			0.019	0.006	0.011	0.008
7	company_location - CL	0.019			0.019	0.006	0.011	0.008
8	company_location - CO	0.019			0.019	0.006	0.011	0.008
9	company_location - CZ	0.019			0.019	0.006	0.011	0.008
10	company_location - DE	0.019			0.019	0.006	0.011	0.008
11	company_location - DK	0.019			0.019	0.006	0.011	0.008
12	company_location - DZ	0.019			0.019	0.006	0.011	0.008
13	company_location - ES	0.019			0.019	0.006	0.011	0.008
14	company_location - FR	0.007	-0.069		-0.061	-0.020	0.042	0.032
15	company_location - GB	0.019			0.019	0.006	0.011	0.008
16	company_location - GR	0.019			0.019	0.006	0.011	0.008
17	company_location - HN	0.019			0.019	0.006	0.011	0.008

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

5.5. Просмотр основных баз данных всех моделей

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

Модель: "Inf3"

Код признака	Описательные шкалы и градации (признаки)	1 - salary_in_usd - 1/3-(2859.0000000, 201906.0000000)			Сумма	Среднее	Средн. квадр. отклонение	Средний модуль отклонения от среднего
		0.175	-0.152	-0.023				
1	company_location - AE	0.175	-0.152	-0.023	-0.000	-0.000	0.164	0.116
2	company_location - AT	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
3	company_location - BE	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
4	company_location - BR	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
5	company_location - CA	1.096	-0.819	-0.277	0.000	0.000	0.987	0.730
6	company_location - CH	0.524	-0.455	-0.069	0.000	0.000	0.493	0.349
7	company_location - CL	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
8	company_location - CO	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
9	company_location - CZ	0.175	-0.152	-0.023	-0.000	-0.000	0.164	0.116
10	company_location - DE	1.310	-1.137	-0.173	-0.000	-0.000	1.232	0.873
11	company_location - DK	0.087	-0.076	-0.012	-0.000	-0.000	0.082	0.058
12	company_location - DZ	0.175	-0.152	-0.023	-0.000	-0.000	0.164	0.116
13	company_location - ES	0.699	-0.606	-0.092	-0.000	-0.000	0.657	0.466
14	company_location - FR	1.231	-0.804	-0.427	-0.000	-0.000	1.082	0.820
15	company_location - GB	3.929	-3.410	-0.519	-0.000	-0.000	3.697	2.619
16	company_location - GR	1.310	-1.137	-0.173	-0.000	-0.000	1.232	0.873
17	company_location - HN	1.222	-1.061	-0.161	0.000	0.000	1.150	0.815

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 [9].

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

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

В соответствии с критерием достоверности L1-мерой проф.Е.В.Луценко наиболее достоверной является СК-модель ABS с интегральным критерием: «Сумма абс.частот по признакам объекта»: $L1=0.909$ (рисунок 12). *Эту модель и будем использовать для решения поставленных в работе задач.*

3.4. Обобщенная форма по достов.моделям при разн.инт.крит.([Приложение: Приложение, созданное в API-2.3.2.2-20260401-222020])

№	Наименование модели и частного критерия	Интегральный критерий	Сумма модулей уровней сходства ложноположит. решений (SFP)	Сумма модулей уровней сходства ложноотрицат. решений (SFN)	S. Точность модели	S. Полнота модели	L1-мера проф. Е.В.Луценко	Сред. мод. уров. сход. исти. пооп. реше (AT)
1	ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч. выборки	корреляция абс. частот с обр. объекта	808.669	0.000	0.358	1.000	0.528	0.7
2	ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч. выборки	сумма абс. частот по признакам объекта	95.120	0.000	0.832	1.000	0.909	0.7
3	PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса	корреляция усл. отн. частот с обр. объекта	808.669	0.000	0.358	1.000	0.528	0.7
4	PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса	сумма усл. отн. частот по признакам объекта	849.887	0.000	0.336	1.000	0.504	0.7
5	PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса	корреляция усл. отн. частот с обр. объекта	808.669	0.000	0.358	1.000	0.528	0.7
6	PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса	сумма усл. отн. частот по признакам объекта	849.887	0.000	0.336	1.000	0.504	0.7
7	INF1 - частный критерий: количество знаний по А.Харкевичу, вероятности из PRC1	семантический резонанс знаний	78.891	79.156	0.116	0.115	0.115	0.11
8	INF1 - частный критерий: количество знаний по А.Харкевичу, вероятности из PRC1	сумма знаний	99.023	58.629	0.120	0.188	0.147	0.0
9	INF2 - частный критерий: количество знаний по А.Харкевичу, вероятности из PRC2	семантический резонанс знаний	78.891	79.156	0.116	0.115	0.115	0.11
10	INF2 - частный критерий: количество знаний по А.Харкевичу, вероятности из PRC2	сумма знаний	99.023	58.629	0.120	0.188	0.147	0.0
11	INF3 - частный критерий: Хи-квадрат: разности между фактическими и ожидаемыми абс. частотами	семантический резонанс знаний	343.696	271.469	0.174	0.211	0.191	0.3
12	INF3 - частный критерий: Хи-квадрат: разности между фактическими и ожидаемыми абс. частотами	сумма знаний	277.268	234.661	0.204	0.232	0.217	0.3
13	INF4 - частный критерий: ROI (Return On Investment), вероятности из PRC1	семантический резонанс знаний	14.322	45.212	0.269	0.105	0.151	0.11
14	INF4 - частный критерий: ROI (Return On Investment), вероятности из PRC1	сумма знаний	42.875	25.537	0.123	0.190	0.149	0.0
15	INF5 - частный критерий: ROI (Return On Investment), вероятности из PRC2	семантический резонанс знаний	14.322	45.212	0.269	0.105	0.151	0.11
16	INF5 - частный критерий: ROI (Return On Investment), вероятности из PRC2	сумма знаний	42.875	25.537	0.123	0.190	0.149	0.0
17	INF7 - частный критерий: разн. усл. и без усл. вероятностей, вероятности из PRC2	семантический резонанс знаний	336.923	265.062	0.176	0.213	0.192	0.3
18	INF7 - частный критерий: разн. усл. и без усл. вероятностей, вероятности из PRC2	сумма знаний	232.946	164.564	0.073	0.101	0.085	0.01

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 ABS.

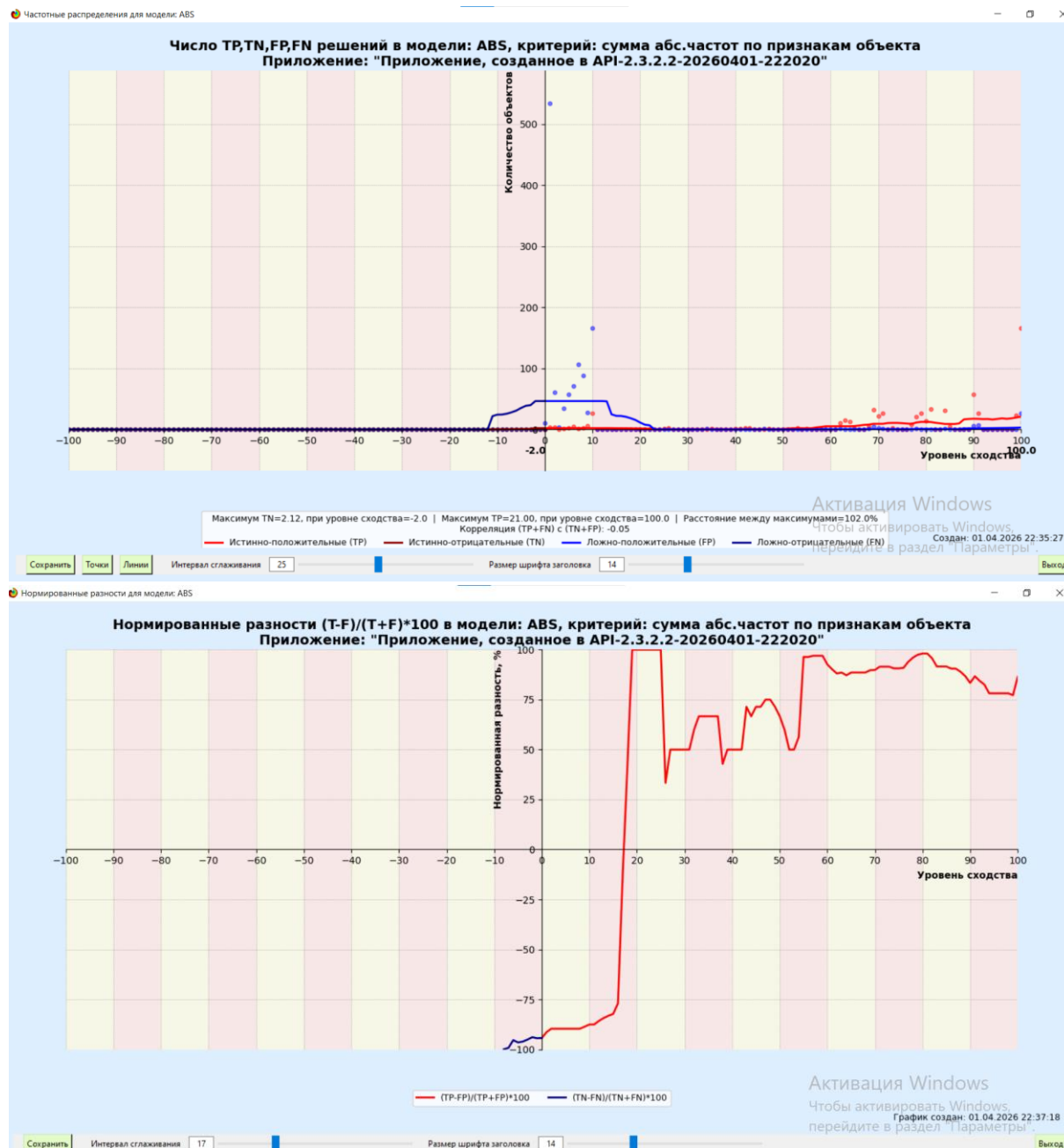


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели ABS

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

Помощь по режиму 3.4: Виды прогнозов и меры достоверности моделей в системе "Эйдос-Х++".

Представим себе, что мы выбрасываем кубик с 6 гранями...
Тогда возможны следующие варианты прогнозов:

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.

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

ИДЕАЛЬНЫЙ ПРОГНОЗ.

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

РЕАЛЬНЫЙ ПРОГНОЗ.

На практике мы чаще всего сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой.

Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у Вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов.

Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно относить объекты к классам, 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}:

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

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

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

Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. 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 i (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.

3.5.1. 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:

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

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

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

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

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

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

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

3.5.2. 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:

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

Поэтому по своей сути он также является скалярным произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применяя сплайнов, в частности линейной интерполяции:

$I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}}$, Это позволяет предложить неограниченное

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

3.5.3. 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 **неортономрированных** пространствах, которые, как правило, и встречаются на практике и в которых применение Евклидова расстояния (теоремы Пифагора) является некорректным.

Во-вторых, данный интегральный критерий являются **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»:

3.5.4. 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):

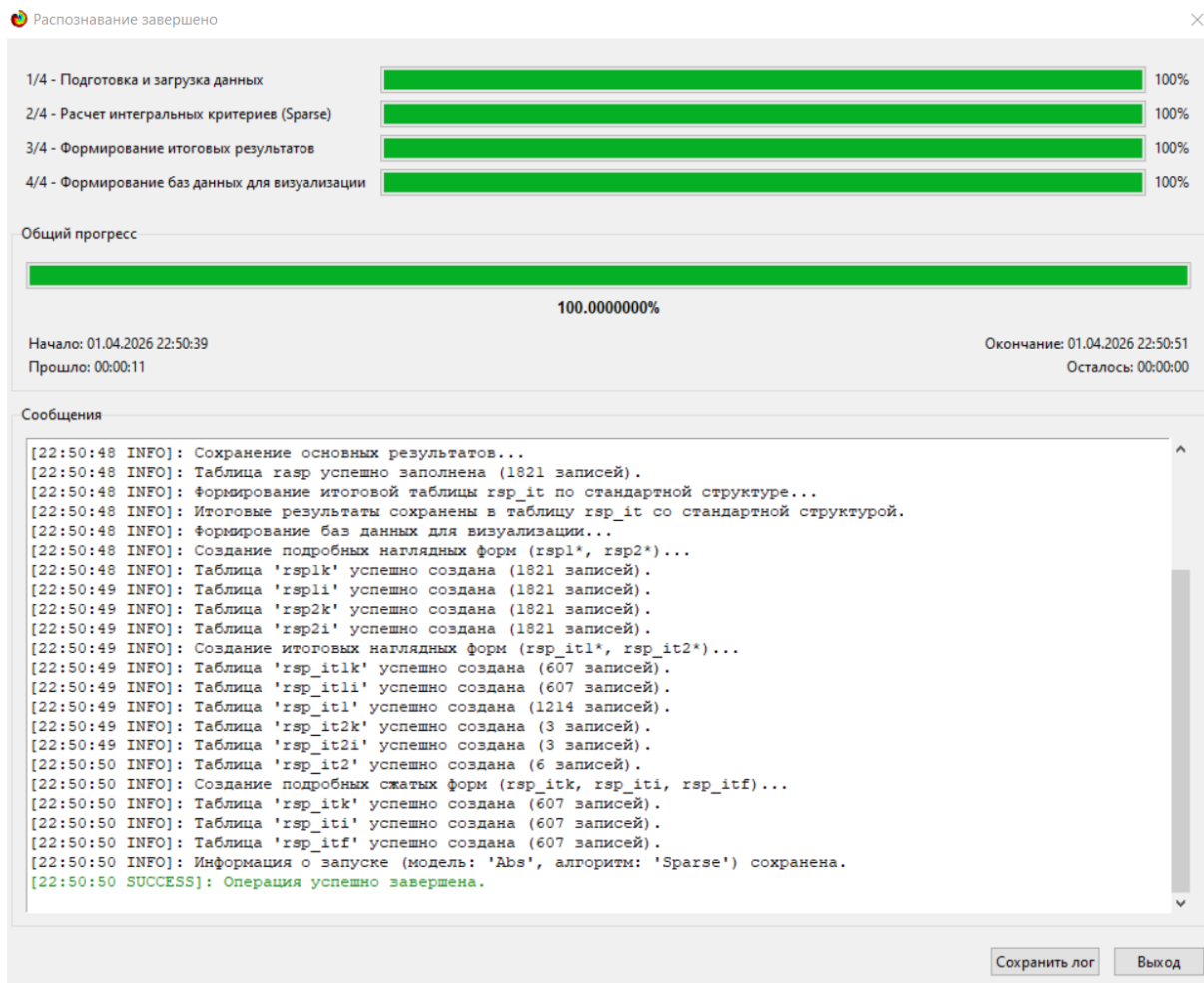


Рисунок 15. Экранные формы режима 4.1.2 идентификации и прогнозирования

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

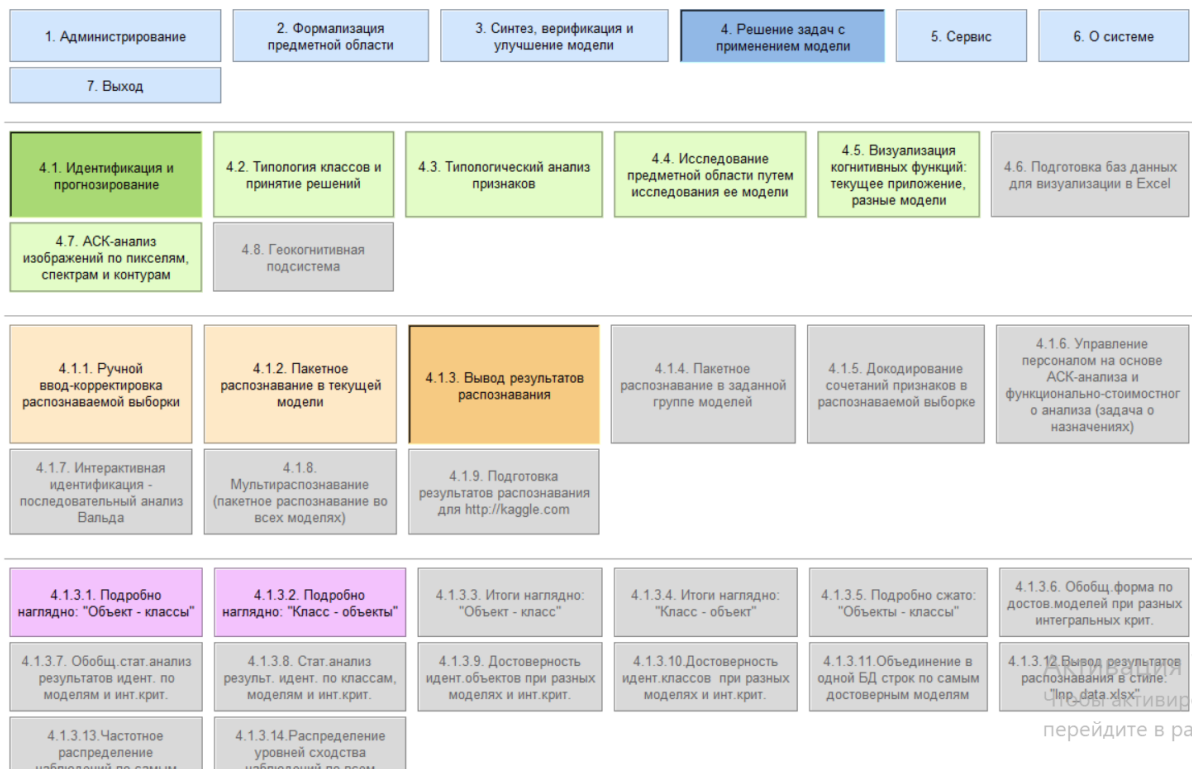


Рисунок 16. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

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

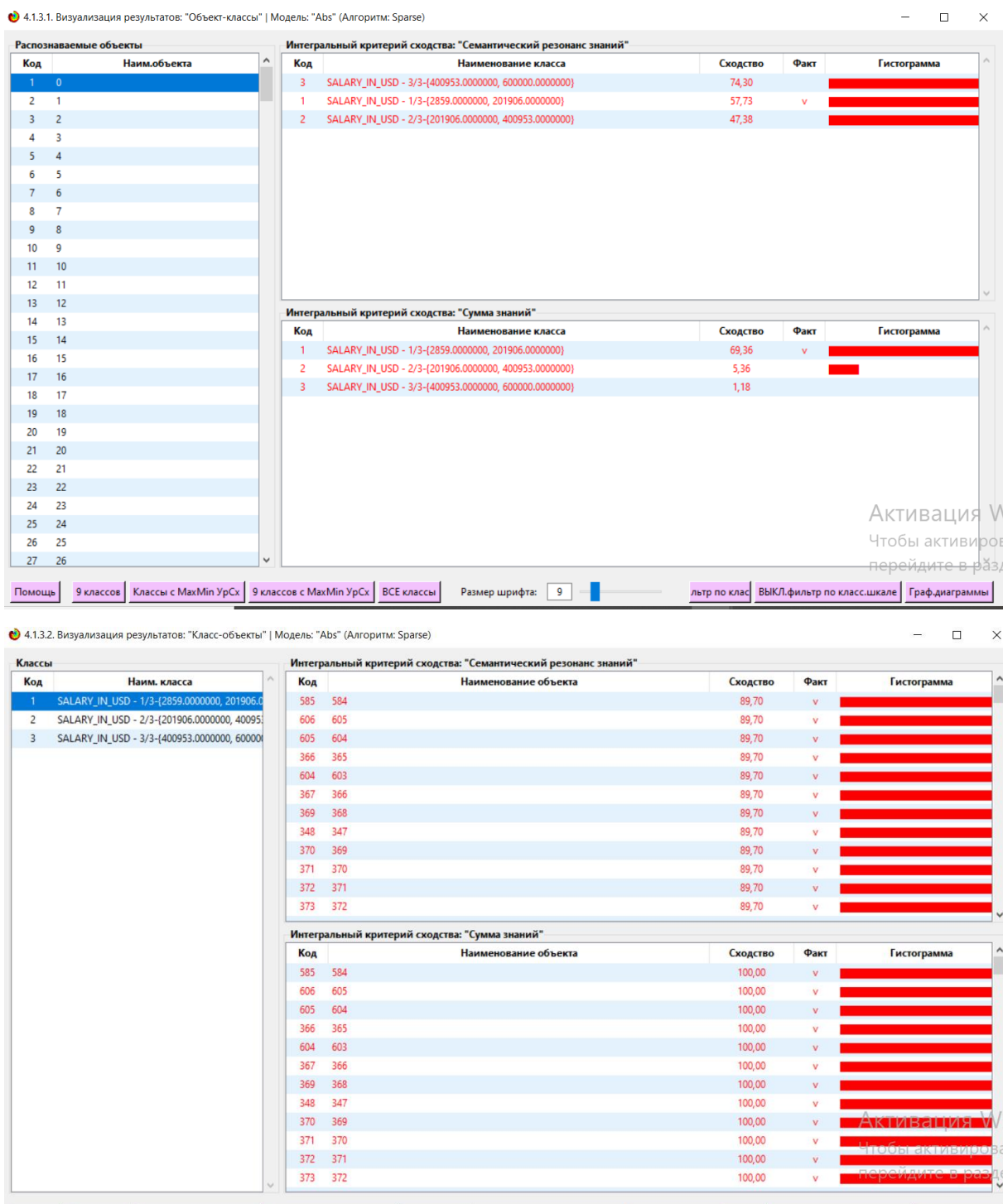


Рисунок 17. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

3.6. Задача-6. Поддержка принятия решений

3.6.1. 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») [10] (рисунки 19).

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

Графические выходные формы, приведенные на рисунках 19, интуитивно понятны и не требуют особых комментариев. Отметим лишь, что на SWOT-диаграммах наглядно показаны знак и сила влияния каждого значения фактора на переход объекта моделирования в состояние,

соответствующее классу, выбранному в верхнем окне. Знак показан цветом, а сила влияния – толщиной линии.

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

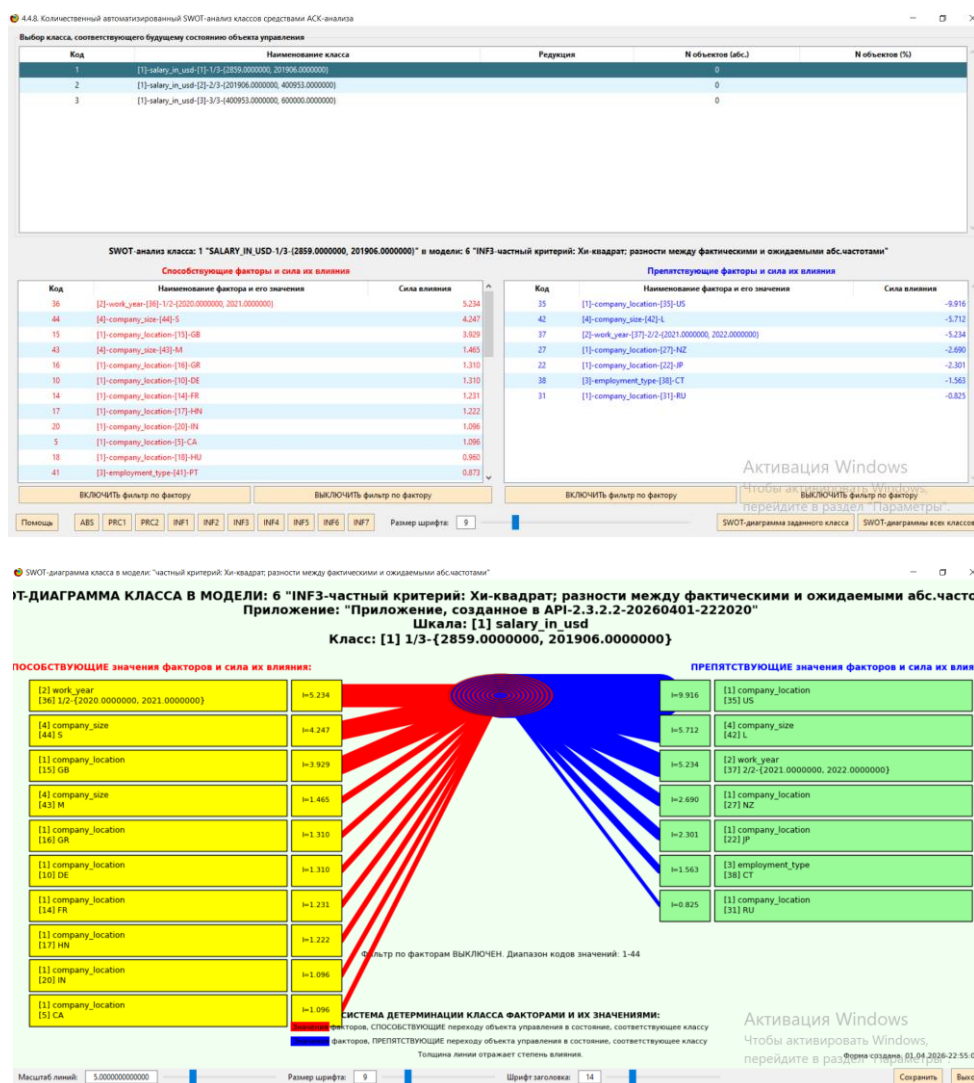


Рисунок 18. Примеры экранных формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

3.6.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе

3.6.3. 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. Этот алгоритм описан в работах [11, 12, 13] и в ряде других работ.

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

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

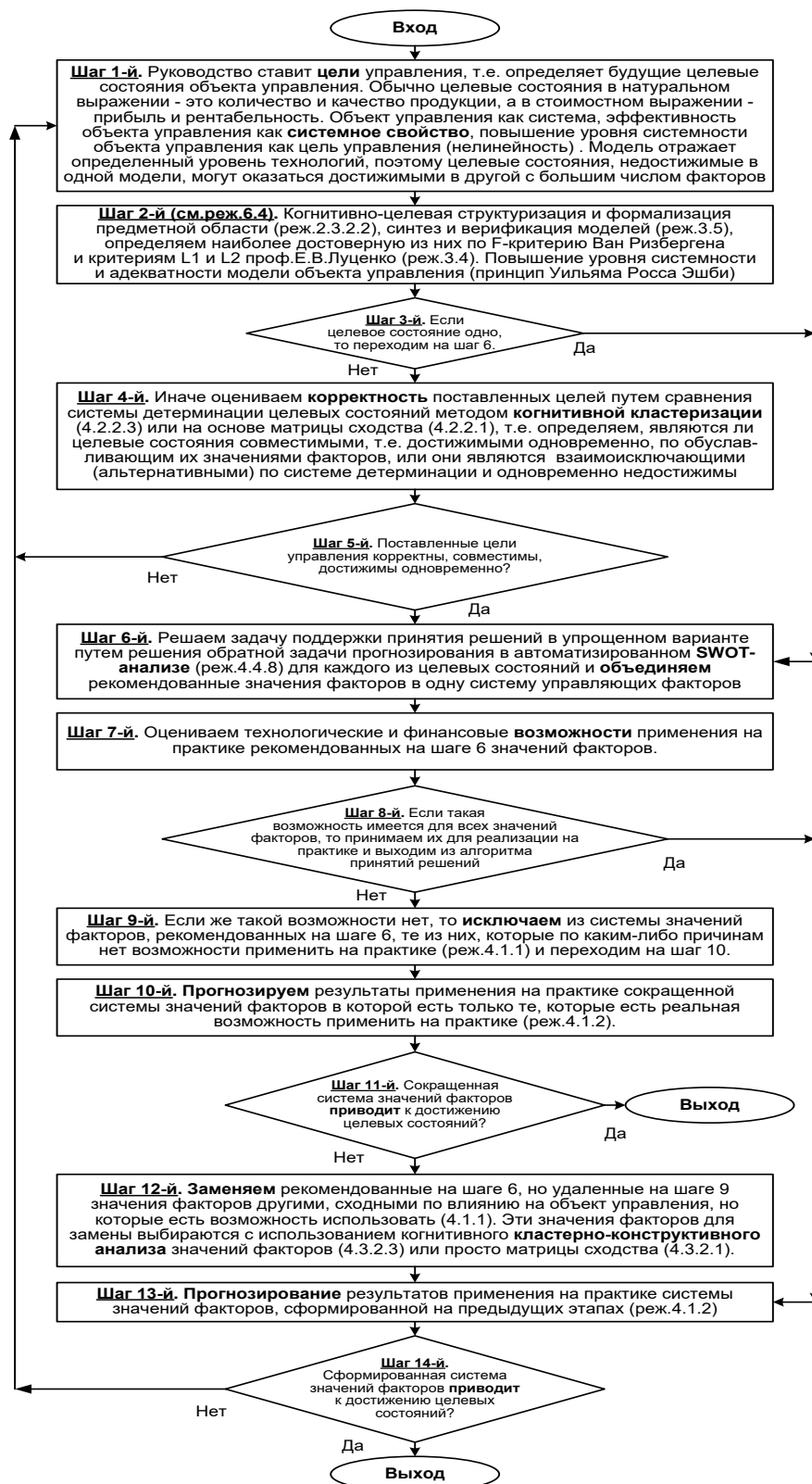


Рисунок 19. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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:

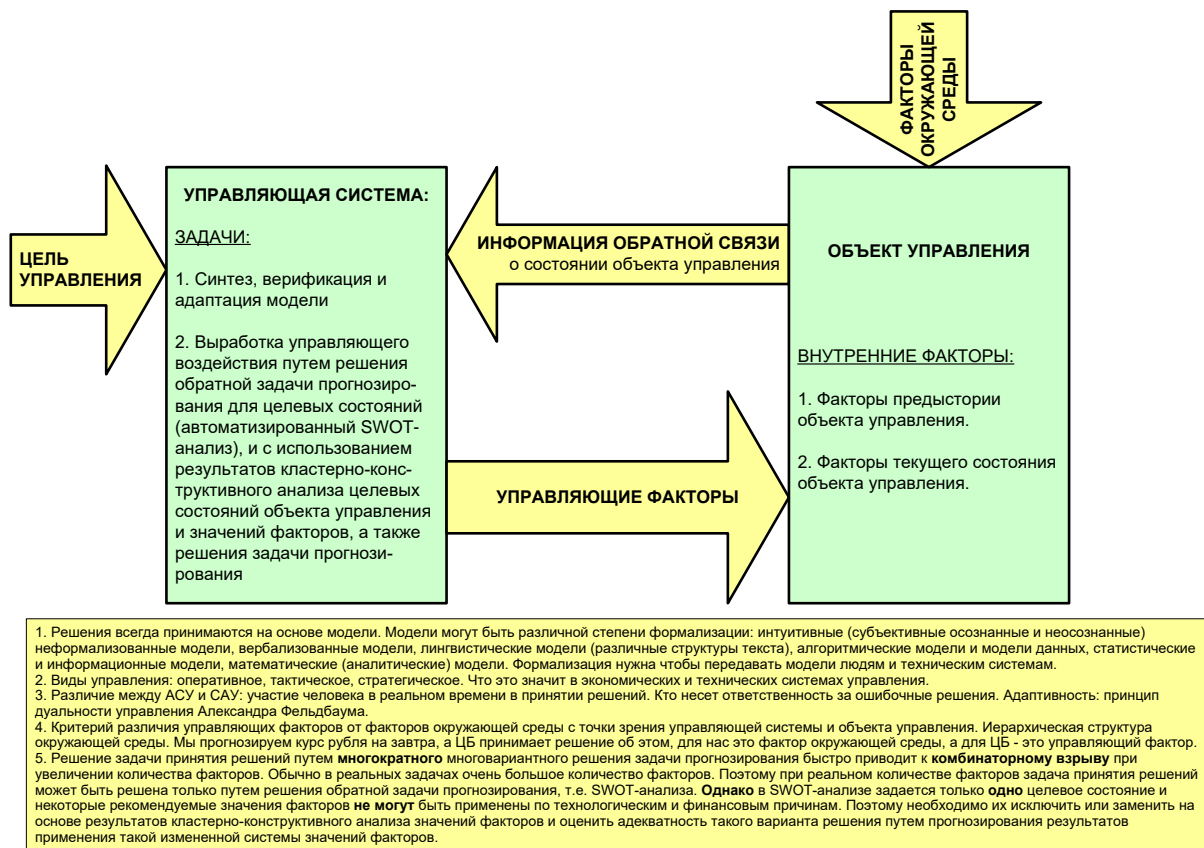


Рисунок 20. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

Как мы видим в развитом алгоритме принятия решений широко используются результаты 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 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. Содержательное объяснение этих эмпирических

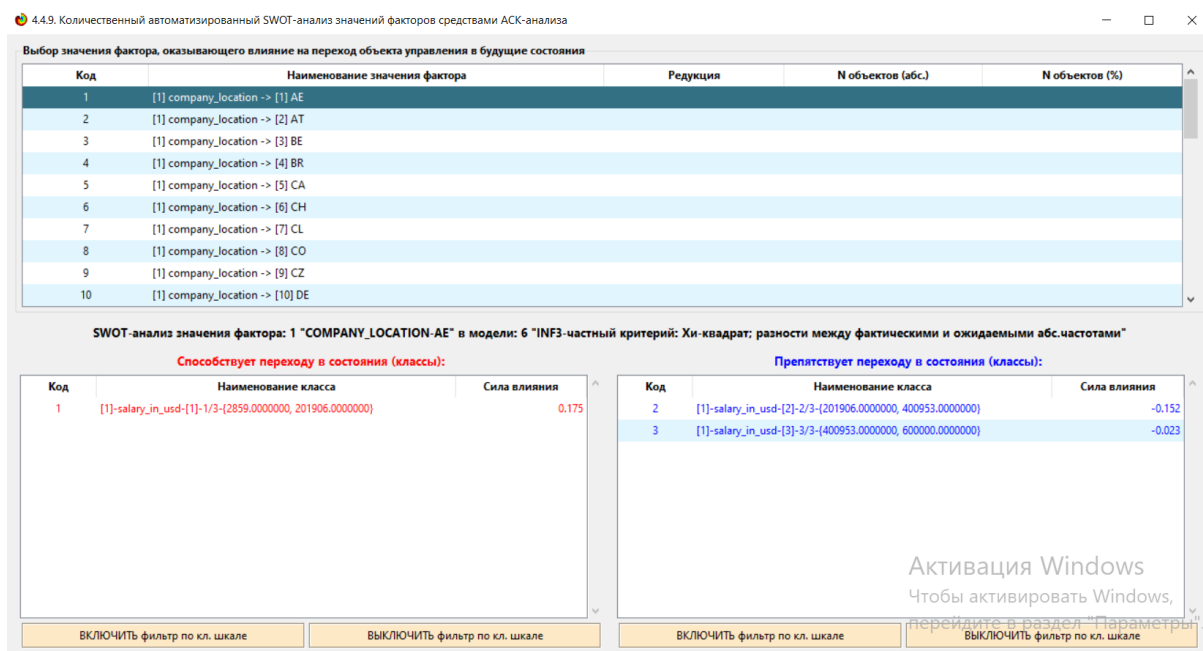
закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [12, 17].

3.7. Задача-7. Исследование объекта моделирования путем исследования его модели

3.7.1. 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:



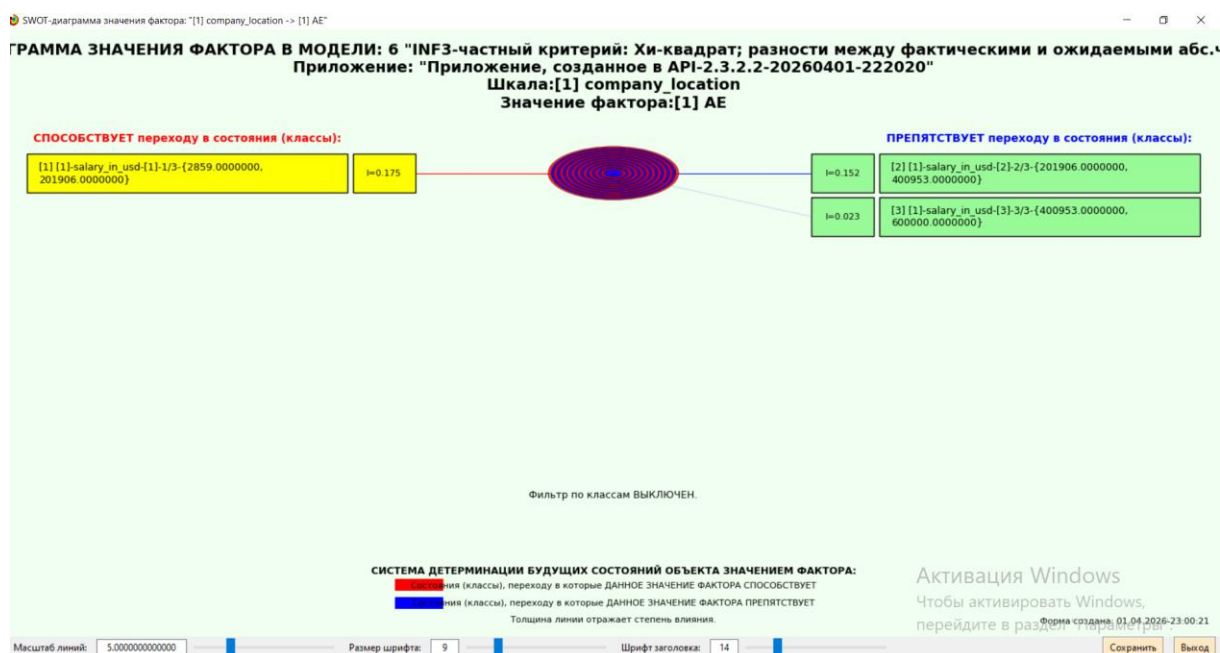


Рисунок 21. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

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. Во многом это и есть решение проблемы, поставленной в работе.

3.7.2. 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.

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

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

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

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

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

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 нелокальной нейронной сети в системе «Эйдос»

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Код	Наименование нелокального нейрона (класса)
1	[1]-SALARY_IN_USD-[1]-1/3-(2859.0000000, 201906.0000000)
2	[1]-SALARY_IN_USD-[2]-2/3-(201906.0000000, 400953.0000000)
3	[1]-SALARY_IN_USD-[3]-3/3-(400953.0000000, 600000.0000000)

Подготовка визуализации нейрона: 1 "[1]-SALARY_IN_USD-[1]-1/3-(2859.0000000, 201906.0000000)" в модели: "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния			ТОРМОЗЯЩИЕ рецепторы и сила их влияния		
Код	Наименование фактора и его значения	Сила влияния	Код	Наименование фактора и его значения	Сила влияния
36	[2] work_year -> [36] 1/2-(2020.0000000, 2021.0000000)	5.2339	35	[1] company_location -> [35] US	-9.9160
44	[4] company_size -> [44] S	4.2471	42	[4] company_size -> [42] L	-5.7117
15	[1] company_location -> [15] GB	3.9292	37	[2] work_year -> [37] 2/2-(2021.0000000, 2022.0000000)	-5.2339
43	[4] company_size -> [43] M	1.4646	27	[1] company_location -> [27] NZ	-2.6903
16	[1] company_location -> [16] GR	1.3097	22	[1] company_location -> [22] JP	-2.3015
10	[1] company_location -> [10] DE	1.3097	38	[3] employment_type -> [38] CT	-1.5634
14	[1] company_location -> [14] FR	1.2306	31	[1] company_location -> [31] RU	-0.8254
17	[1] company_location -> [17] HN	1.2224			
20	[1] company_location -> [20] IN	1.0956			
5	[1] company_location -> [5] CA	1.0956			
18	[1] company_location -> [18] HU	0.9605			
41	[3] employment_type -> [41] PT	0.8731			

Параметры визуализации нейросети:

MAX нейрона:	16	Диапазон кодов нейронов:	1	3	Размеры:	
MAX рецепторов:	16	Диапазон кодов рецепторов:	1	44	по X:	1536
MAX связей:	1000	Инициализация: % от макс:	0.0		по Y:	864

Активация Windows. Чтобы активировать Windows, перейдите в Заголовки, Узоры, Значения, Легенда.

Размер шрифта графической формы: []
 Размер шрифта в таблицах данной формы: []

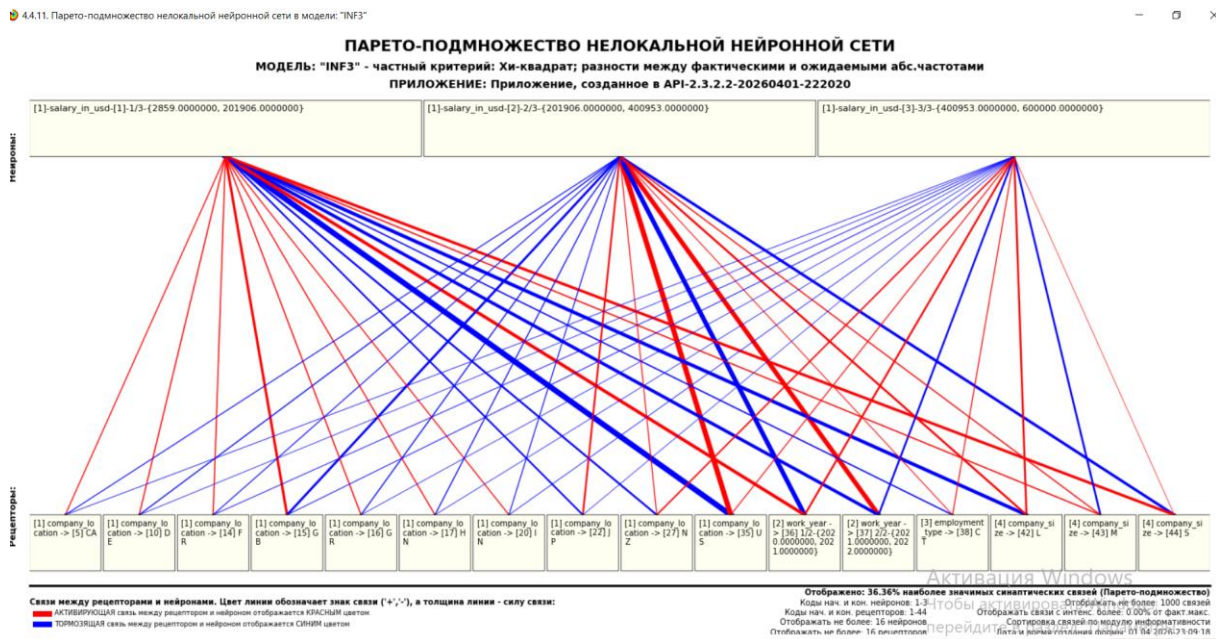


Рисунок 24. Нейронная сеть в СК-модели INF3

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 [6].

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

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 ABS:

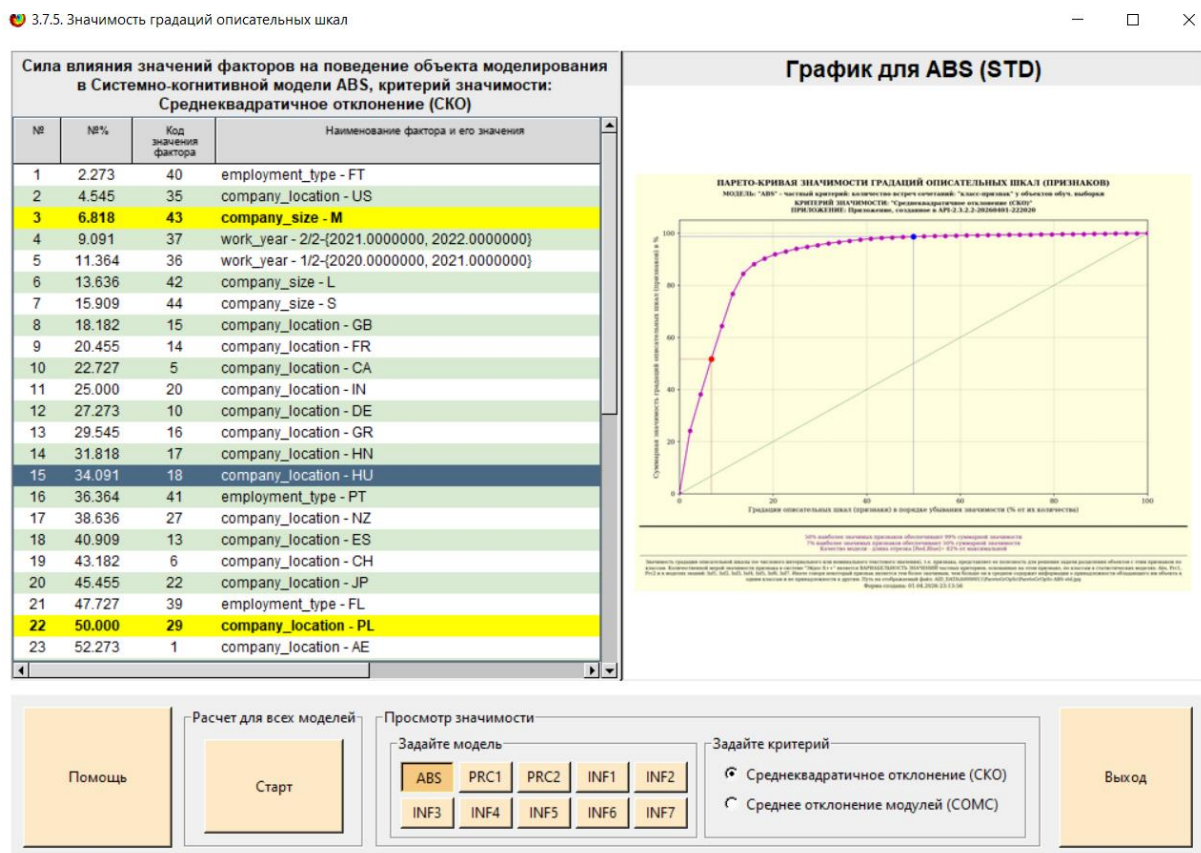


Рисунок 25 . Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели ABS

В таблице 16 представлены исходные данные для построения кумулятивной кривой на рисунке 37. Из таблицы 16 видно, какую долю от суммарного влияния на переход объекта моделирования в будущее состояния, соответствующие классам, имеет каждое значение каждого фактора.

3.8.11. Степень детерминированности классов и классификационных шкал

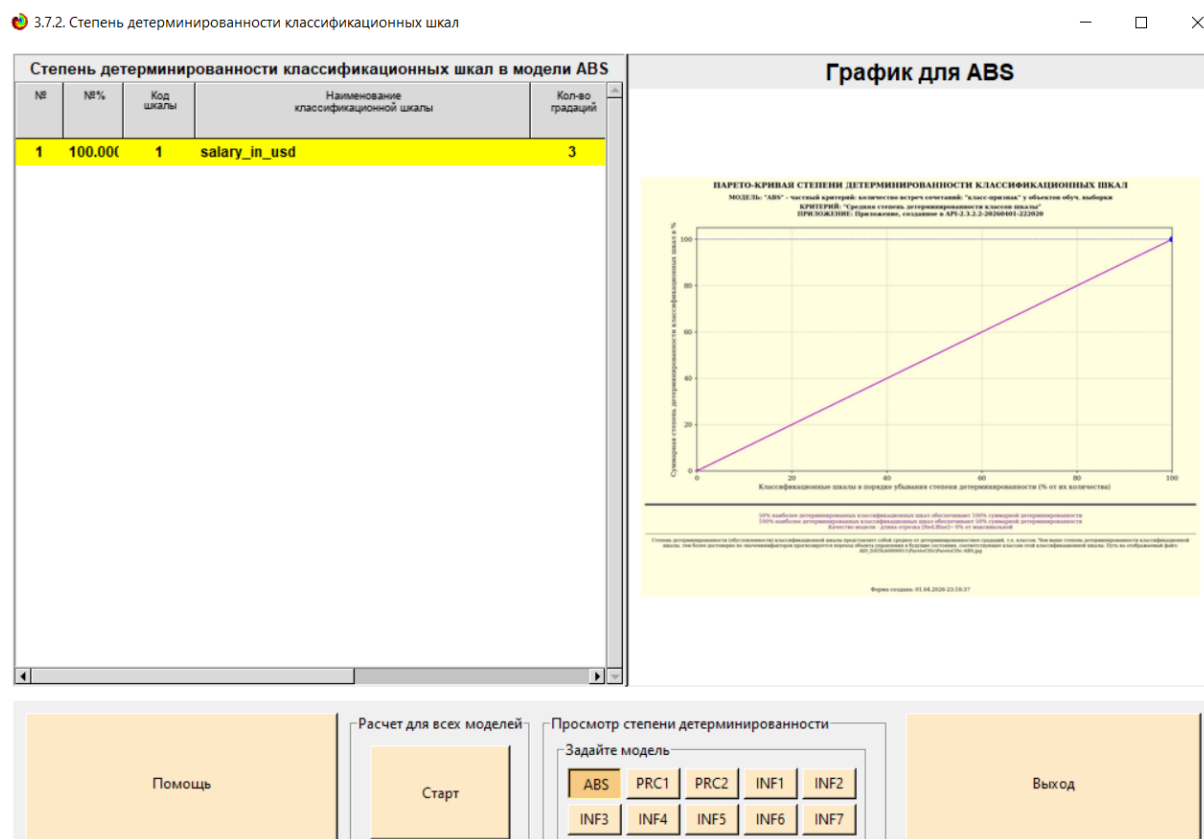
Степень детерминированности (обусловленности) класса в системе «Эйдос» количественно оценивается *степенью варибельности значений*

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 рисунке 38 приведены экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



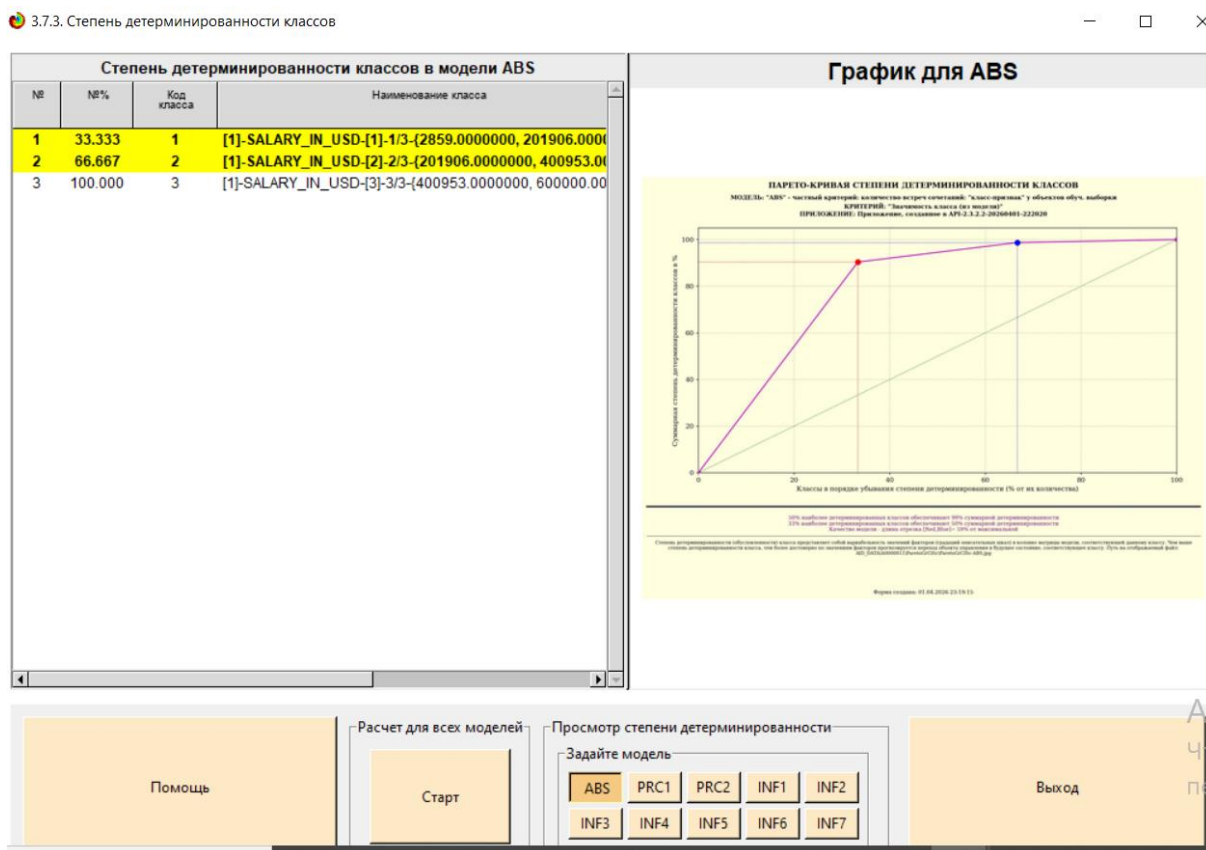


Рисунок 26. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

4. DISCUSSION (ОБСУЖДЕНИЕ)

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

Анализ полученных результатов, проведенный в данной работе, полностью согласуется с результатами работы, на исходных данных которой они основаны. С другой стороны, применение АСК-анализа и системы «Эйдос» весьма существенно расширяет возможности решения задач прогнозирования, принятия решений и исследования моделируемой предметной области, по сравнению с методами, применяемыми в работе.

Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований.

Достижением данной работы является:

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

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

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

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

Для выполнения анализа надо скачать систему с сайта разработчика по ссылке на странице: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос приложение. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Эйдос» решение поставленных задач является вполне эффективным, что позволяет обоснованно утверждать, что цель работы достигнута, поставленная проблема решена.

В результате проделанной работы, с помощью системы «Эйдос» были созданы статистические и системно-когнитивные модели, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по различным характеристикам покупателя в магазине, изучено влияние параметров людей из выборки на эти классы, и, на основе этого, решены задачи идентификации, классификации и исследования моделируемой предметной области путем исследования ее модели

REFERENCES (ЛИТЕРАТУРА)

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

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

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

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

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

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

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

16. Луценко, Е. В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе "Эйдос") / Е. В. Луценко, В. Е. Коржаков // Политематический сетевой

электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 71. – С. 27-74. – EDN OIGYBB.

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

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

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

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

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

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

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

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