

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

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

Факультет прикладной информатики

Кафедра Компьютерных систем и технологий


КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ зависимости сожжённых калорий во время
физических тренировок на основе набора данных ресурса Kaggle

Выполнил студент группы: ИТ2341 Хомицкий Евгений Евгеньевич

Допущен к защите _____

Руководитель проекта д.э.н., к.т.н., профессор Луценко Е.В. ()
(подпись, расшифровка подписи)

Защищен _____
(дата)

Оценка отлично

Краснодар
2025

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

Факультет прикладной информатики

РЕЦЕНЗИЯ
на курсовую работу

Студента Хомицкого Евгения Евгеньевича
курса 2 очной формы обучения группы ИТ2341
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ зависимости сожжённых калорий во
время физических тренировок на основе набора данных ресурса Kaggle

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

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____ отлично

Рецензент _____ (Е. В. Луценко)

«21» февраля 2025 г.

РЕФЕРАТ

Курсовая работа содержит: 89 страниц, 41 рисунок, 16 таблиц, 23 литературных источника.

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

Целью работы является проведение автоматизированного системно-когнитивного анализа зависимости сожжённых калорий во время физических тренировок на основе набора данных ресурса Kaggle

Объект исследования – организм человека.

Предмет исследования – выявление зависимостей между количеством сожженных калорий и 7 биологическими характеристиками тела человека с учетом длительности тренировки.

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

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ ЗАВИСИМОСТИ СОЖЖЁННЫХ КАЛОРИЙ ВО ВРЕМЯ ФИЗИЧЕСКИХ ТРЕНИРОВОК НА ОСНОВЕ НАБОРА ДАННЫХ РЕСУРСА KAGGLE

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

В данной работе используется автоматизированный системно-когнитивный анализ (АСК-анализ) для изучения зависимости между сжиганием калорий в организме человека и семью биологическими характеристиками, учитывая продолжительность тренировки. На основе выявленных зависимостей решаются задачи прогнозирования, принятия решений и исследования предметной области через анализ её системно-когнитивной модели. Особенность задачи заключается в том, что все независимые переменные (факторы) представлены в виде лингвистических (категориальных) переменных. Поэтому для её решения применяется лингвистический АСК-анализ, основанный на методах когнитивной математической лингвистики. При этом зависимые переменные, отражающие результаты влияния факторов, измеряются в числовых шкалах. Таким образом, строится гибридная модель, объединяющая номинальные (текстовые) и числовые данные.

Для обеспечения сопоставимости данных разных типов, представленных в различных шкалах и единицах измерения, используется метризация номинальных шкал. Это означает повышение уровня их формализации до числовых шкал путём вычисления количества информации, содержащейся в градациях номинальных шкал, и её связи с целевыми показателями. В работе кратко описывается методология АСК-анализа и его программная реализация в интеллектуальной системе «Эйдос».

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

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF THE DEPENDENCE OF CALORIES BURNED DURING PHYSICAL TRAINING BASED ON A DATASET FROM THE KAGGLE RESOURCE

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

In this work, an automated system-cognitive analysis (ASC-analysis) is used to study the relationship between calorie burning in the human body and seven biological characteristics, taking into account the duration of training. Based on the identified relationships, tasks of forecasting, decision-making, and exploring the subject area through the analysis of its system-cognitive model are addressed. A distinctive feature of the task is that all independent variables (factors) are represented as linguistic (categorical) variables. Therefore, linguistic ASC-analysis, based on methods of cognitive mathematical linguistics, is applied to solve it. At the same time, dependent variables, reflecting the results of the influence of factors, are measured in numerical scales. Thus, a hybrid model is constructed, combining nominal (textual) and numerical data.

To ensure the comparability of different types of data represented in various scales and units of measurement, the metricization of nominal scales is used. This means increasing the level of their formalization to numerical scales by calculating the amount of information contained in the gradations of nominal scales and its connection with target indicators. The work briefly describes the methodology of ASC-analysis and its software implementation in the intelligent system "Eidos."

Keywords: LINGUISTIC ASK-ANALYSIS, LINGUISTIC AUTOMATED SYSTEMIC COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS",

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	6
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	6
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	7
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ	7
1.4. ЦЕЛЬ РАБОТЫ	8
2. METHODS (МЕТОДЫ)	9
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	9
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	9
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ	9
2.4. СИСТЕМА «ЭЙДОС» - ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	11
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	16
3. RESULTS (РЕЗУЛЬТАТЫ)	18
3.1. ЗАДАЧА-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	18
3.2. ЗАДАЧА-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	19
3.3. ЗАДАЧА-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	26
3.4. ЗАДАЧА-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	35
3.5. ЗАДАЧА-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	38
3.6. ЗАДАЧА-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	40
3.6.1. <i>Интегральный критерий «Сумма знаний»</i>	41
3.6.2. <i>Интегральный критерий «Семантический резонанс знаний»</i>	41
3.6.3. <i>Важные математические свойства интегральных критериев</i>	43
3.6.4. <i>Решение задачи идентификации и прогнозирования в системе «Эйдос»</i>	43
3.7. ЗАДАЧА-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ	46
3.7.1. <i>Упрощенный вариант принятия решений как обратная задача прогнозирования,</i> <i>позитивный и негативный информационные портреты классов, SWOT-анализ</i>	46
3.7.2. <i>Развитый алгоритм принятия решений в адаптивных интеллектуальных системах</i> <i>управления на основе АСК-анализа и системы «Эйдос»</i>	48
3.8. ЗАДАЧА-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ ...	52
3.8.1. <i>Инвертированные SWOT-диаграммы значений описательных шкал (семантические</i> <i>потенциалы)</i>	52
3.8.2. <i>Кластерно-конструктивный анализ классов</i>	54
3.8.3. <i>Кластерно-конструктивный анализ значений описательных шкал</i>	57
3.8.4. <i>Модель знаний системы «Эйдос» и нелокальные нейроны</i>	60
3.8.5. <i>Нелокальная нейронная сеть</i>	62
3.8.6. <i>3D-интегральные когнитивные карты</i>	64
3.8.7. <i>2D-интегральные когнитивные карты содержательного сравнения классов</i> <i>(опосредованные нечеткие правдоподобные рассуждения)</i>	65
3.8.8. <i>2D-интегральные когнитивные карты содержательного сравнения значений факторов</i> <i>(опосредованные нечеткие правдоподобные рассуждения)</i>	68
3.8.9. <i>Когнитивные функции</i>	71
3.8.10. <i>Значимость описательных шкал и их градаций</i>	78
3.8.11. <i>Степень детерминированности классов и классификационных шкал</i>	82
4. DISCUSSION (ОБСУЖДЕНИЕ)	85
5. CONCLUSIONS (ВЫВОДЫ)	86
REFERENCES (ЛИТЕРАТУРА)	87

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Гиппократ, древнегреческий философ и отец медицины, еще в V веке до нашей эры утверждал, что многие болезни возникают из-за образа жизни. Он подчеркивал, что если человек не готов изменить свои привычки, ему ничто не поможет. Гиппократ сравнивал тело с тканью, которую необходимо очищать, а гимнастику — с инструментом для "очистки" организма, подчеркивая важность физической активности для поддержания здоровья.

Сегодня, несмотря на технологический прогресс и доступность современных медицинских услуг, люди часто забывают о важности физической активности. Спорт не только укрепляет здоровье, но и развивает характер, целеустремленность, выносливость и лидерские качества. Даже небольшие, но регулярные физические нагрузки способны значительно улучшить состояние организма. Например, они помогают исправить осанку, что способствует правильному расположению внутренних органов и их здоровой работе. Кроме того, спорт укрепляет кости и связки, снижая риск травм, особенно в пожилом возрасте.

Физическая активность также положительно влияет на дыхательную систему. Регулярные тренировки увеличивают объем легких, улучшают их вентиляцию и предотвращают застойные явления, что снижает риск заболеваний дыхательных путей. Сердечно-сосудистая система также укрепляется: сердце становится более выносливым, улучшается кровообращение, снижается риск инфарктов, инсультов и других сердечных заболеваний.

Одним из ключевых аспектов физической активности является сжигание калорий. Калории — это единицы энергии, которые организм получает из пищи. Если человек потребляет больше калорий, чем тратит, избыток энергии накапливается в виде жира, что может привести к лишнему весу и ожирению. Ожирение, в свою очередь, является фактором риска для многих заболеваний, включая диабет, гипертонию, болезни сердца и опорно-двигательного аппарата. Регулярные физические нагрузки помогают сжигать калории, поддерживая энергетический баланс и предотвращая накопление лишнего жира.

Для точного подсчета сожженных калорий во время тренировок были разработаны специальные модели и формулы. Эти модели учитывают различные факторы, такие как вес человека, возраст, пол, интенсивность и продолжительность тренировки, а также тип физической активности. Например, для ходьбы, бега, плавания или силовых упражнений существуют отдельные алгоритмы расчета. Современные технологии, такие как фитнес-трекеры и умные часы, используют эти модели для предоставления пользователям точных данных о количестве

сожженных калорий. Это помогает людям лучше контролировать свои тренировки и достигать поставленных целей, будь то похудение, поддержание веса или улучшение физической формы.

Процесс сжигания калорий особенно важен для поддержания здорового веса и улучшения общего самочувствия. Во время физической активности организм использует запасы энергии, начиная с глюкозы в крови, затем переходя к гликогену в мышцах и, наконец, к жировым отложениям. Таким образом, спорт способствует не только похудению, но и улучшению метаболизма. Ускоренный обмен веществ помогает организму быстрее перерабатывать пищу, эффективнее использовать питательные вещества и выводить токсины.

Кроме того, спорт улучшает внешний вид. Регулярные тренировки делают фигуру более подтянутой, укрепляют мышцы и улучшают тонус кожи. Физическая активность ускоряет микроциркуляцию крови, что способствует здоровому цвету лица, уменьшению отечности и сокращению количества морщин. Это происходит благодаря снижению уровня гормона стресса кортизола, который подавляет синтез коллагена — белка, отвечающего за упругость кожи.

Таким образом, спорт играет ключевую роль в укреплении здоровья и улучшении качества жизни. Он не только помогает поддерживать физическую форму, но и способствует психическому благополучию, снижая уровень стресса и повышая самооценку. Как говорил Гиппократ, здоровье начинается с образа жизни, и физическая активность — это простой, доступный и эффективный способ его поддержать. Регулярные тренировки, даже умеренные, способны значительно улучшить состояние организма, предотвратить множество заболеваний и помочь человеку чувствовать себя лучше как физически, так и эмоционально. А благодаря современным технологиям и моделям подсчета калорий каждый может точно отслеживать свои достижения и прогресс, делая тренировки еще более эффективными.

1.2. Объект и предмет исследования

Объект исследования – организм человека.

Предмет исследования – зависимостей между количеством сожженных калорий и 7 биологическими характеристиками тела человека, а также учитывается длительность тренировки.

1.3. Проблема, решаемая в работе и ее актуальность

Проблема, которая решается в данном исследовании, заключается в недостаточном понимании ключевых факторов, влияющих на количество сжигаемых калорий во время физической активности, и отсутствии эффективных методов для точного подсчета и оптимизации этого процесса. В современном мире, где малоподвижный образ жизни,

неправильное питание и стресс становятся все более распространенными, вопросы поддержания здоровья и физической формы приобретают особую актуальность. Многие люди стремятся к здоровому образу жизни, занимаются спортом, но сталкиваются с трудностями в достижении желаемых результатов, таких как снижение веса, улучшение физической подготовки или поддержание формы. Это связано с тем, что процесс сжигания калорий является сложным и зависит от множества факторов, которые часто не учитываются в повседневной практике.

Актуальность данной проблемы обусловлена несколькими причинами. Во-первых, рост числа людей с избыточным весом и ожирением становится глобальной проблемой, которая приводит к увеличению случаев сердечно-сосудистых заболеваний, диабета, проблем с опорно-двигательным аппаратом и других хронических заболеваний. Во-вторых, в условиях современного ритма жизни многие люди испытывают дефицит времени и не могут уделять достаточно внимания длительным и интенсивным тренировкам. В-третьих, отсутствие точных и персонализированных методов подсчета сожженных калорий приводит к тому, что люди часто переоценивают или недооценивают эффективность своих тренировок, что снижает мотивацию и препятствует достижению целей.

Исследование направлено на решение этих проблем путем системного анализа факторов, влияющих на сжигание калорий, таких как тип физической активности, интенсивность и продолжительность тренировок, а также индивидуальные характеристики человека (вес, возраст, пол, уровень физической подготовки). Это позволяет разработать более точные модели и алгоритмы, которые помогут людям оптимизировать свои тренировки, учитывая их уникальные особенности и цели.

Таким образом, исследование не только решает актуальную проблему недостаточной эффективности тренировок, но и способствует популяризации здорового образа жизни, повышению мотивации людей к физической активности и улучшению их общего самочувствия. В условиях современного общества, где здоровье становится одной из ключевых ценностей, такие исследования имеют большое значение для улучшения качества жизни и профилактики хронических заболеваний.

1.4. Цель работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели обеспечивается решением ряда *задач* и подзадач, которые являются *этапами* достижения цели. Конкретная формулировка этих задач зависит от метода решения проблемы, поэтому обоснованно мы сформулируем их в конце раздела, т.е. после обоснованного выбора и описания метода 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 очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

2. Иначе говоря, метод решения проблемы не должен предъявлять жестких требований к исходным данным, которые невозможно выполнить, а должен обеспечивать обработку тех данных, которые реально есть.

3. Метод должен 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 программных систем, *одновременно*:

- находящихся в полном открытом бесплатном доступе;
- обеспечивающих сопоставимую обработку числовой и текстовой информации в одной модели, дает следующие результаты; показал, что альтернатив Автоматизированному системно-когнитивному анализу и его программному инструментарю – системе «Эйдос» в настоящее время здесь нет [3].

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 [1].

*Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф.Е.В.Луценко. На тот момент он вообще не встречался в Internet. Сегодня по соответствующему запросу в Яндексe находится 9 миллионов сайтов с этим сочетанием слов*².

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)

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 [2] и ряде других [4]. Около половины из 730 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 50 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 32 патент РФ на системы искусственного интеллекта, 346 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 10 докторских диссертациях по экономическим, техническим, биологическим наукам и в 10 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.

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

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

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

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

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 [6] и страничка в РесечГейт [7], на которых можно получить более полную информацию о методе АСК-анализа и системе «Эйдос». Краткая информация об АСК-анализе и системе «Эйдос» есть в материале: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf.

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 [5]. Сама метризация номинальных шкал достигается *путем* вычисления количества информации, содержащегося в градациях номинальных шкал о получении той или иной урожайности [4]. Для работы с лингвистическими переменными применяется лингвистический АСК-анализ [3].

2.4. Система «Эйдос» - инструментарий АСК-анализа

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

- является универсальной и может быть применена во многих предметных областях, т.к. разработана в универсальной постановке, не зависящей от предметной области (<http://lc.kubagro.ru/aidos/index.htm>).
- Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени при решении задач идентификации, прогнозирования, принятия решений и исследования предметной области (автоматические системы работают без такого участия человека);
- находится в полном открытом бесплатном доступе (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), причем с актуальными исходными текстами (<http://lc.kubagro.ru/AidosALL.txt>): открытая лицензия: **CC BY-SA 4.0** (<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 системы

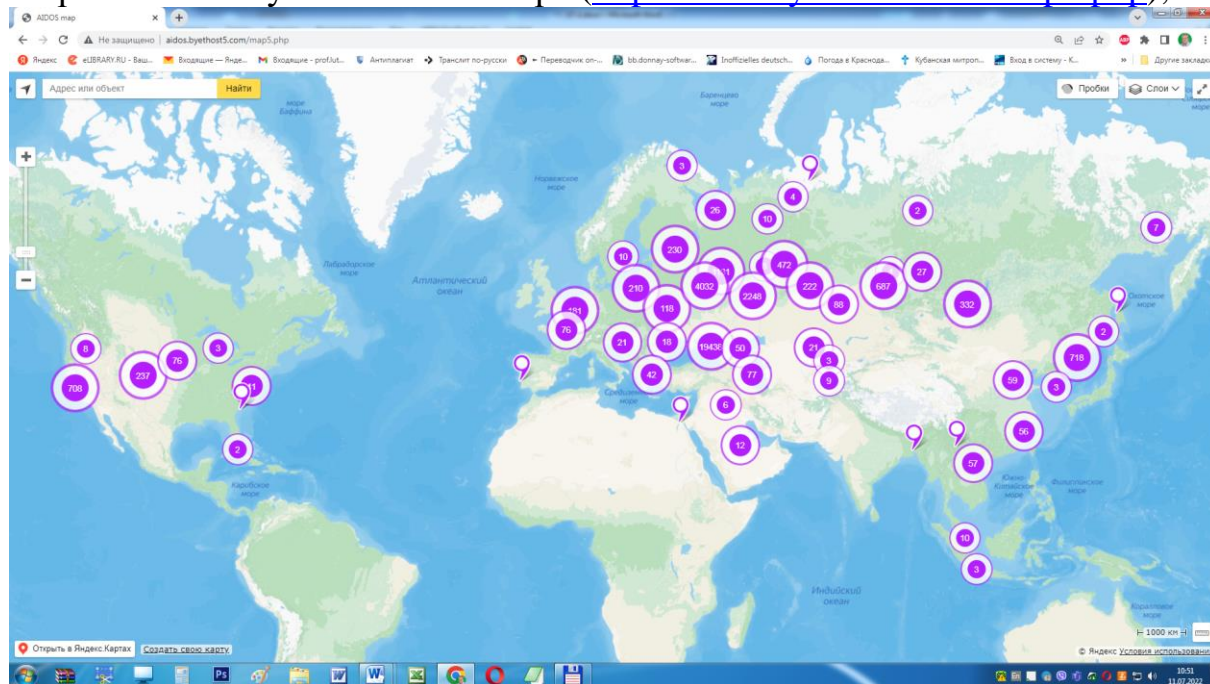
«Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 31 свидетельство РосПатента РФ);

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

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

– имеет «нулевой порог входа», содержит большое количество локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более **435**, соответственно: http://aidos.byethost5.com/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

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



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

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

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

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

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

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

[В 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 законах⁷.

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-2200С). В 1987 году впервые получен [акт внедрения](#) на одну из ранних версий системы «Эйдос», реализованную в среде персональной технологической системы «Вега-М» разработки автора (см. 2-й акт по ссылке: <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>).

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBM-совместимых персональных компьютеров система "Эйдос" впервые реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены [свидетельства РосПатента](#), первые в Краснодарском крае и, возможно, в России на системы искусственного интеллекта (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-Х1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

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

5-й этап, «эра Больших данных, информации и знаний»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке Аляска+Экспресс,

⁷ Ссылка на это краткое описание системы «Эйдос» на английском языке: http://lc.kubagro.ru/aidos/The_Eidos_en.htm

обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge) с использованием ADS (Advantage Database Server), а также на языке C# (Visual Studio | C#).

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



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

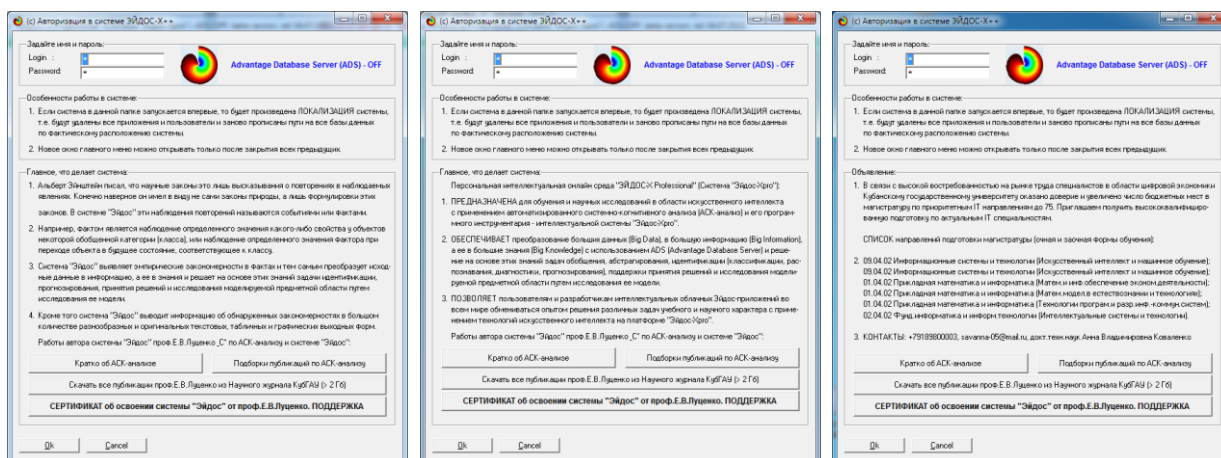


Рисунок 2. Титульные видеोगраммы текущей версии системы «Эйдос»

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

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

Как уже показано выше, для работы с лингвистическими переменными целесообразно применить лингвистический АСК-анализ [3].

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

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:

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

2) кластерно-конструктивный анализ классов;

3) кластерно-конструктивный анализ значений описательных шкал;

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

5) нелокальная нейронная сеть;

6) 3d-интегральные когнитивные карты;

7) 2d-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения);

8) 2d-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения);

9) когнитивные функции;

10) значимость описательных шкал и их градаций;

11) степень детерминированности классов и классификационных шкал).

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**

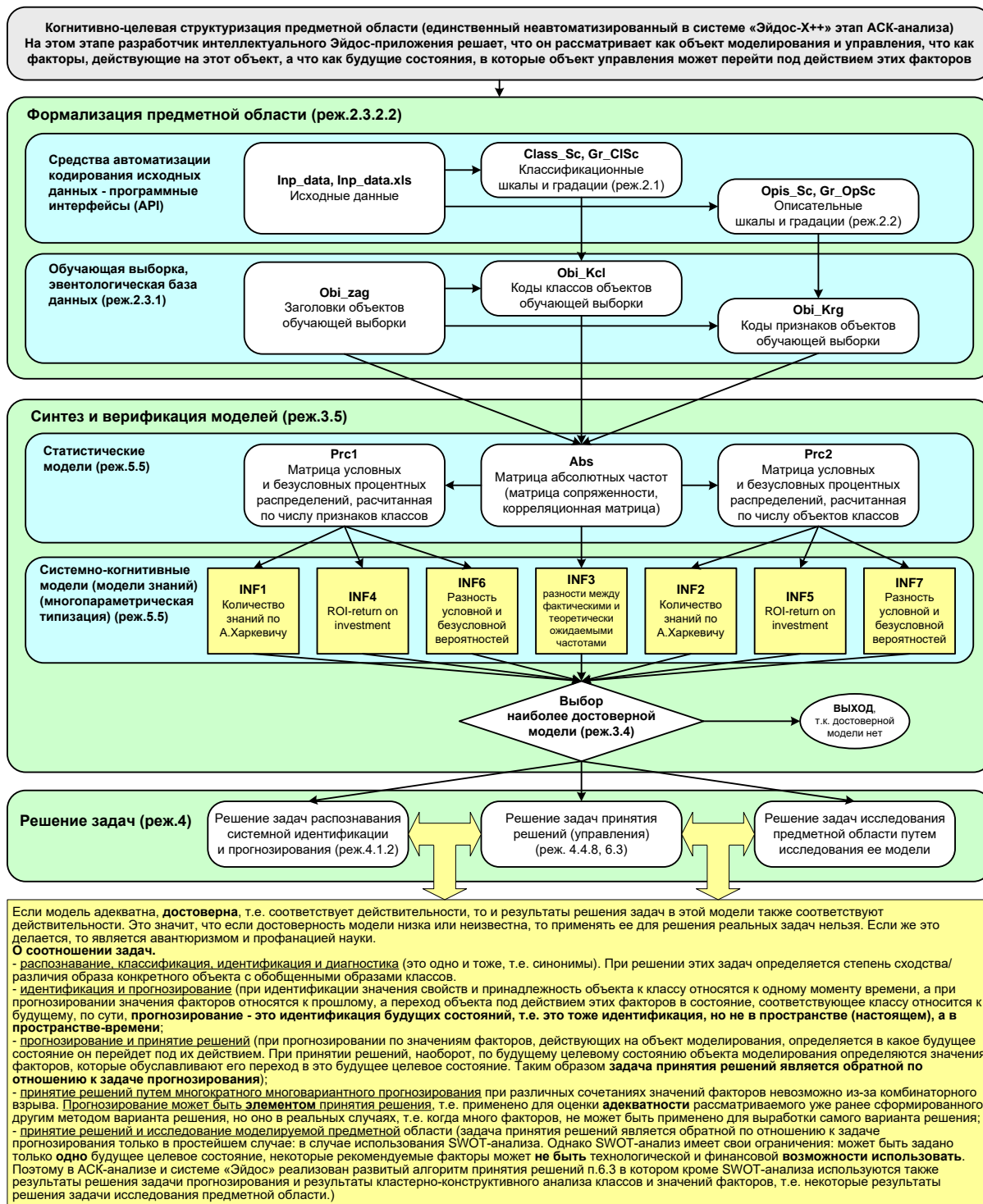


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»

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.

В данной работе в качестве *объекта моделирования* человеческий организм, в качестве *факторов* вид, пол человека, рост, вес, продолжительность тренировки, частота сердцебиения, температура тела (таблица 1), а в качестве *результатов* действия этих факторов количество сожженных калорий (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	GENDER
2	AGE
3	HEIGHT
4	WEIGHT
5	DURATION
6	HEART_RATE
7	BODY_TEMP

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Opis_Sc.dbf

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	CALORIES

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Class_Sc.dbf

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 в самых различных направлениях науки и решения практических задач в самых различных предметных областях, практически почти везде, где человек применяет естественный интеллект.

В качестве источника данных для представленных в этой работе исследований с электронного ресурса kaggle.com был взят набор данных «Calories Burnt Prediction» <https://www.kaggle.com/datasets/ruchikakumbhar/calories-burnt-prediction>.

	A	B	C	D	E	F	G	H	I
1	User_ID	Gender	Age	Height	Weight	Duration	Heart_Rate	Body_Ten	Calories
2	14733363	male	68	190	94	29	105	40,8	231
3	14861698	female	20	166	60	14	94	40,3	66
4	11179863	male	69	179	79	5	88	38,7	26
5	16180408	female	34	179	71	13	100	40,5	71
6	17771927	female	27	154	58	10	81	39,8	35
7	15130815	female	36	151	50	23	96	40,7	123
8	19602372	female	33	158	56	22	95	40,5	112
9	11117088	male	41	175	85	25	100	40,7	143
10	12132339	male	60	186	94	21	97	40,4	134
11	17964668	female	26	146	51	16	90	40,2	72
12	13723164	female	36	177	76	1	74	37,8	3
13	13681290	female	21	157	56	17	100	40	92
14	15566424	male	66	171	79	11	90	40	58
15	12891699	female	32	157	54	18	93	40,4	88
16	13823829	male	53	182	85	2	82	38,1	7
17	17557348	female	39	156	62	28	104	40,8	170
18	12198133	male	39	182	82	4	82	38,6	11
19	15236104	male	46	169	67	11	89	40,2	43
20	11042324	female	27	171	65	4	85	38,6	15
21	16864285	male	50	188	86	14	94	40,2	74
22	11674347	male	67	189	93	8	77	39,2	29
23	19797300	female	31	148	50	8	84	39,5	32
24	14711095	female	33	157	60	3	80	38,7	10
25	14434854	female	20	165	59	29	100	41	155
26	14893804	male	48	182	85	1	80	37,7	3
27	17231597	male	29	176	75	10	83	39,7	24
28	10901446	male	33	173	73	7	78	39,3	12
29	15874362	male	42	190	88	3	83	38,9	10
30	15569252	female	62	159	59	29	106	41,2	195
31	15615743	male	38	171	75	2	81	38,2	5
32	13363046	male	20	183	88	16	97	40,5	69
33	17572853	female	25	160	59	24	102	40,3	136
34	17157339	female	24	165	59	18	91	40,2	80
35	18328111	female	42	165	68	22	93	40,8	107
36	19303479	male	22	182	84	29	114	41	200
37	10699201	female	74	158	59	10	93	39,6	56

Рисунок 4 – Фрагмент набора данных Excel-таблица исходных данных в стандарте системы «Эйдос»

Данные на рисунке 4 имеют следующую структуру:
 – каждая строка описывает одно наблюдение, всего 15000 наблюдений;

В первом столбце указывается номер испытуемого (User_ID) в итоге, исследование проведено на 15000 человек. В сидящих столбцах (со 2 по 8) указываются пол, возраст, особенности тела человека (рост и вес) и

продолжительность тренировки и параметры на которые оказывается влияние физических нагрузок, это частота сердцебиения и температура тела. В 9 столбце указывается количество сожжённых калорий за период физической тренировки.

Пол (Gender): Пол может влиять на метаболизм и количество сожжённых калорий, так как у мужчин и женщин разная мышечная масса и гормональный фон.

Возраст (Age): С возрастом метаболизм замедляется, что может влиять на количество сожжённых калорий.

Рост (Height) и Вес (Weight): Эти параметры напрямую связаны с расчётом индекса массы тела (ИМТ) и общим уровнем энергозатрат. Чем больше вес и рост, тем больше калорий может сжигаться при физической нагрузке.

Продолжительность тренировки (Duration): Чем дольше длится тренировка, тем больше калорий сжигается.

Частота сердцебиения (Heart Rate): Это показатель интенсивности тренировки. Чем выше пульс, тем больше энергии тратит организм.

Температура тела (Body Temperature): Повышение температуры тела может указывать на интенсивность метаболических процессов, что также связано с расходом калорий.

Отметим, что 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 жестких практически невыполнимых требований к исходным данным, а обрабатывает те данные, которые реально есть, например подобные представленным в таблице 5.

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 5).

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

Рисунок 5. Автоматизированные программные интерфейсы (API) системы «Эйдос»

Для ввода исходных данных, представленных в таблице 5, 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 (рисунок 6):

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

В таблицах 3, 4, 5 приведены классификационные и описательные шкалы и градации, а также обучающая выборка, сформированные API-2.3.2.2 при параметрах, показанных на рисунках 7.

Для классификационных шкал на 3-м рисунке 7 приведено также количество наблюдений для каждого интервального значения (градации) и его размер. За счет того, что интервальные значения имеют разные размеры удается преодолеть *несбалансированность данных*, т.к. число наблюдений в каждом интервальном значении некоторой шкалы получается равным с точностью 1 (т.к. число наблюдений – всегда целое число).

Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал.

Помощь по режиму 2.3.2.2 для случая Excel-файлов исходных данных

Режим 2.3.2.2: Универсальный программный интерфейс импорта данных из внешней базы данных "Inp_data.xls" в систему "Эйдос:Х++" и формализации предметной области.

- Данный программный интерфейс обеспечивает формализацию предметной области, т.е. анализ файла исходных данных Inp_data.xls(x), формирование классификационных и описательных шкал и градаций, а затем кодирование файла исходных с их использованием.
- Файл исходных данных должен иметь имя: Inp_data.xls(x), а файл распознаваемой выборки имя: Inp_rasp.xls(x). Файлы Inp_data.xls(x) и Inp_rasp.xls(x) должны находиться в папке: \AIDOS\AID_DATA\Inp_data/. Эти файлы имеют совершенно одинаковую структуру.
- 1-я строка этого файла должна содержать наименования колонок на любом языке, в т.ч. и русском. Эти наименования должны быть во всех колонках, при этом переносы по словам разрешены, а объединение ячеек, разрыв строки знак абзаца не допускаются. Эти наименования должны быть короткими, но понятными, т.к. они будут в выходных формах, а к ним еще будут добавляться наименования градаций. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения и число знаков после запятой в колонке должно быть ОДИНАКОВОЕ.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным: до 255 символов.
- Каждая строка этого файла, начиная со 2-й, содержит данные об одном объекте обучающей выборки или одном наблюдении. В MS Excel-2003 в листе может быть до 65536 строк и до 256 колонок. В листе MS Excel-2010 и более поздних возможно до 1048576 строк и 16384 колонок.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (номинального / порядкового) или числового типа (с десятичными знаками после запятой).
- Столбцу присваивается числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбцу присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Столбцы со 2-го по N-й являются классификационными шкалами (выходными параметрами) и содержат данные о классах (будущих состояниях объекта управления), к которым принадлежат объекты обучающей выборки.
- Столбцы с N+1 по последний являются описательными шкалами (свойствами или факторами) и содержат данные о признаках (т.е. значениях свойства или значениях факторов), характеризующих объекты обучающей выборки.
- В результате работы режима формируется файл INP_NAME.TXT стандарта MS DOS (кириллица), в котором наименования классификационных и описательных шкал являются СТРОКАМИ. Система формирует классификационные и описательные шкалы и градации. Для этого в каждом числовом столбце система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего числовые значения заменяются их интервальными значениями. В текстовых столбцах система находит уникальные текстовые значения. Каждое УНИКАЛЬНОЕ интервальное числовое или текстовое значение считается градацией классификационной или описательной шкалы, характеризующей объект. В каждой шкале ее градации сортируются по алфавиту. С использованием шкал и градаций кодируются исходные данные в результате чего генерируется обучающая выборка, каждый объект которой соответствует одной строке файла исходных данных NP_DATA и содержит коды классов, соответствующие фактам совпадения числовых или уникальных текстовых значений классов с градациями классификационных шкал и коды признаков, соответствующие фактам совпадения числовых или уникальных текстовых значений признаков с градациями описательных шкал
- Расознаваемая выборка формируется на основе файла INP_RASP аналогично, за исключением того, что классификационные и описательные шкалы и градации не создаются, а используются ранее созданные в модели, и базы распознаваемой выборки могут не включать коды классов, если столбцы классов в файле INP_RASP были пустыми. Структура файла INP_RASP должна быть такая же, как INP_DATA, т.е. они должны ПОЛНОСТЬЮ совпадать по наименованиям столбцов, но могут иметь разное количество строк с разными значениями в них.

Принцип организации таблицы исходных данных:

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

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

Помощь по режиму 2.3.2.2 для случая Excel-файлов исходных данных

Режим 2.3.2.2: Универсальный программный интерфейс импорта данных из внешней базы данных "Inp_data.xls(x)" в систему "Эйдос:Х++"

ТЕРМИНЫ АСК-АНАЛИЗА И СИСТЕМЫ "ЭЙДОС":

Шкала представляет собой способ формализации предметной области. Используются числовые и текстовые шкалы, при этом текстовые могут быть номинальными и порядковыми. На номинальных шкалах есть только отношения эквивалентности и неэквивалентности, на порядковых кроме того еще отношения "больше", "меньше", а на числовых - кроме того могут выполняться все арифметические операции. Каждый объект выборки (наблюдение) описан с одной стороны своими признаками, а с другой - принадлежностью к некоторым обобщающим категориям (классам). Такая структура описания называется онтологией или фреймом экземпляром и является базовой для всех моделей представления знаний.

В АСК-анализе и системе "Эйдос" используется три интерпретации шкал и градаций: универсальная, статическая и динамическая:

- в универсальной интерпретации: признаки - это градации описательных шкал;
- в статической интерпретации: описательная шкала - это свойство, а градация (признак) - это степень выраженности этого свойства;
- в динамической интерпретации: описательная шкала - это фактор, а градация (признак) - это значение фактора;
- в универсальной интерпретации: классы - это градации классификационных шкал;
- в статической интерпретации: классификационная шкала - способ классификации обобщающих категорий (классов), к которым в настоящем времени по отношению к признакам относятся состояния объекта моделирования;
- в динамической интерпретации: классификационная шкала - способ классификации обобщающих категорий (классов), к которым в будущем времени по отношению к признакам относятся состояния объекта прогнозирования или управления;

ПРОФИЛАКТИКА ОШИБОК В ФАЙЛЕ ИСХОДНЫХ ДАННЫХ:

- 1-я строка файла "Inp_data.xls(x)" должна содержать наименования колонок. Эти наименования должны быть во всех колонках, при этом переносы по словам разрешены, а объединение ячеек, разрыв строки знак абзаца и неалфавитные символы не допускаются. Эти наименования должны быть короткими, но понятными, т.к. они будут в выходных формах, а к ним еще будут добавляться наименования градаций. В числовых шкалах надо обязательно указывать единицы измерения. Число знаков после запятой в числовой колонке должно быть одинаковым.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным: до 255 символов.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (номинального / порядкового) или числового типа (со знаками после запятой). Чтобы текстовая шкала была порядковой, нужно чтобы при сортировке по алфавиту градации этой шкалы образовывали осмысленную последовательность от минимального значения до максимального. Например, текстовая шкала "Размер" с градациями: "очень малое", "малое", "среднее", "большое", "очень большое", будет номинальной шкалой, т.к. при сортировке по алфавиту они расположатся в порядке: "большое", "малое", "очень большое", "очень малое", "среднее". Чтобы шкала "Размер" стала порядковой нужно в эти градации присвоить следующие значения: "1/5-очень малое", "2/5-малое", "3/5-среднее", "4/5-большое", "5/5-очень большое".
- Столбцу присваивается числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбцу присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Если в системе "Эйдос" в режиме 2.1, 2.2 посмотреть на градации классификационных и описательных шкал, которые должны быть числовыми, то сразу будет видно, в какой форме представлены числа: числовыми диапазонами или прямо числами. Если числовыми диапазонами, значит в файле исходных данных в этом отношении все правильно, если же числами, то возможно в Excel-файле нужно заменить десятичные точки на запятые, а также найти и исправить нечисловые данные в числовых по смыслу колонках. Быстро найти их можно перейдя на последнюю строку файла исходных данных и задав расчет сумм в колонке. В формуле будет видно с какой строки идет расчет сумм. Если со 2-й, то значит все верно, иначе будет указана строка, в которой находится нечисловое значение.
- Система "Эйдос" работает с областью данных файла исходных данных, которую можно выделить блоком, поставив курсор в ячейку A1, нажав Shift+Home, а затем эжжав клавиши Shift+Ctrl нажать End. Если этот блок выходит за пределы области таблицы, фактически занятой данными надо скопировать эту фактическую область данных в буфер обмена, создать новый лист и скопировать в него, а исходный лист удалить.
- Иногда бывает полезно сбросить все форматирование Excel-таблицы исходных данных. Это можно сделать в MS Excel. А можно скопировать таблицу в MS Word, а потом обратно в MS Excel.

Принцип организации таблицы исходных данных:

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

Рисунок 6. Хелпы API-2.3.2.2 системы «Эйдос»

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

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

Задайте тип файла исходных данных: "Inp_data":

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

Задайте параметры:

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

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:
 Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:
 Конечный столбец описательных шкал:

Задайте режим:

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

Задайте способ выбора размера интервалов:

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

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

Не применять сценарный метод АСК-анализа Применить сценарный метод АСК-анализа
 Применить спец. интерпретацию текстовых полей классов Применить спец. интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

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

Интерпретация TXT-полей признаков:

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

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

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

Ok Cancel

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

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)

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

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	1	3	3,00	6	42	7,00
Текстовые	0	0	0,00	1	2	2,00
ВСЕГО:	1	3	3,00	7	44	6,29

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

В классификационных шкалах: В описательных шкалах:

Пересчитать шкалы и градации Выйти на создание модели

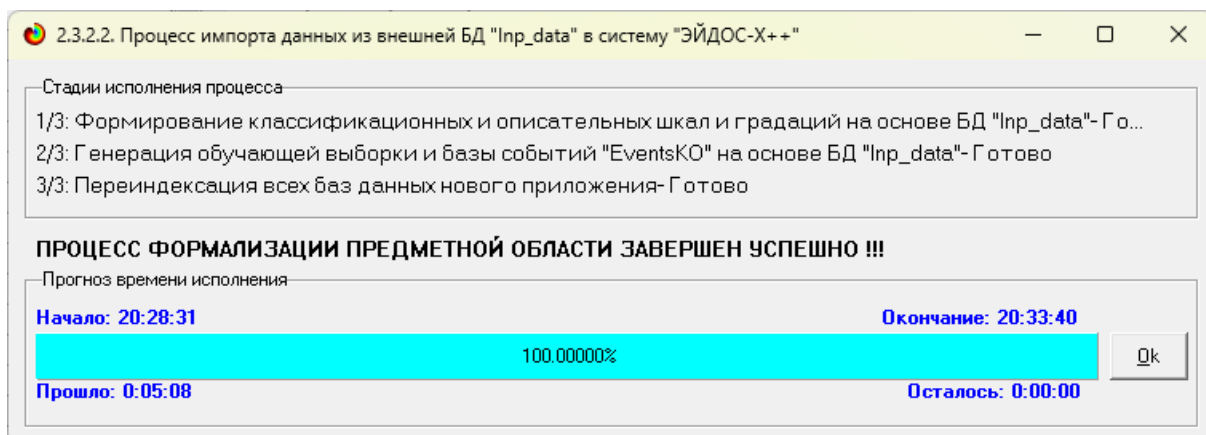


Рисунок7. Экранные формы управления API-2.3.2.2 системы «Эйдос»

Таблица 3 – Классификационные шкалы и градации (числовые шкалы)

KOD_CLS	NAME_CLS
1	CALORIES-1/3-{1.0, 105.3}
2	CALORIES-2/3-{105.3, 209.7}
3	CALORIES-3/3-{209.7, 314.0}

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Classes.dbf

Таблица 4 – Описательные шкалы и градации (лингвистические переменные)

KOD_ATR	NAME_ATR
1	GENDER-female
2	GENDER-male
3	AGE-1/7-{20.0000000, 28.4285714}
4	AGE-2/7-{28.4285714, 36.8571429}
5	AGE-3/7-{36.8571429, 45.2857143}
6	AGE-4/7-{45.2857143, 53.7142857}
7	AGE-5/7-{53.7142857, 62.1428571}
8	AGE-6/7-{62.1428571, 70.5714286}
9	AGE-7/7-{70.5714286, 79.0000000}
10	HEIGHT-1/7-{123.0000000, 137.1428571}
11	HEIGHT-2/7-{137.1428571, 151.2857143}

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Attributes.dbf

Таблица 5 – Обучающая выборка (полностью)

NAME_OBJ	N2	N3	N4	N5	N6	N7	N8	N9
14733363	2	8	14	21	30	35	43	3
14861698	1	3	13	18	27	34	43	1
11179863	2	8	13	20	24	33	40	1
16180408	1	4	13	19	26	34	43	1
17771927	1	3	12	18	26	32	42	1
15130815	1	4	11	18	29	34	43	2

19602372	1	4	12	18	29	34	43	2
11117088	2	5	13	20	29	34	43	2
12132339	2	7	14	21	28	34	43	2
17964668	1	3	11	18	27	33	42	1
13723164	1	4	13	19	24	31	39	1
13681290	1	3	12	18	27	34	42	1
15566424	2	8	13	20	26	33	42	1
12891699	1	4	12	18	28	33	43	1
13823829	2	6	14	20	24	32	39	1
17557348	1	5	12	18	30	35	43	2
12198133	2	5	14	20	24	32	40	1
15236104	2	6	13	19	26	33	42	1

Источник: c:\Aidos-X\AID_DATA\A0000002\System\EventsKO.dbf

Отметим, что в системе «Эйдос» обычно используются базы данных с расширением «dbf». Они открываются в MS Excel или могут быть конвертированы в файлы xls, xlsx с помощью онлайн-сервисов.

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

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

Но для решения всех задач используется не непосредственно сама эта матрица, а матрицы условных и безусловных процентных распределений и системно-когнитивные модели, которые рассчитываются на ее основе и отражают какое количество информации содержится в факте наблюдения определенной градации описательной шкалы о том, что объект моделирования перейдет в состояние, соответствующее определенной градации классификационной шкалы (классу) (рисунок 3).

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике и обеспечивает [1, 2] сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых данных, представленных в различных типах шкал (номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных (см. Help режима 2.3.2.2) рассчитывается матрица абсолютных частот (таблица 6).

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

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

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

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

На практике часто встречается существенная **несбалансированность данных**, под которой понимается сильно отличающееся количество наблюдений объектов обучающейся выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 9) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 10) является весьма обоснованным и логичным.

Таблица 6 – Матрица абсолютных частот (статистическая модель ABS)

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

Таблица 7 – 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				$P_{\Sigma j}$			

Этот переход полностью снимает проблему несбалансированности данных, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 6), а матрицы условных и безусловных процентных распределений (таблица 7) и матрицы системно-когнитивных моделей (СК-модели, таблица 9), в частности матрица информативностей.

Этот подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (номинальных, порядковых и числовых) и в разных единицах измерения [5].

В системе «Эйдос» этот подход применяется *всегда* при решении любых задач.

Затем на основе таблиц 6 и 7 с использованием *частных критериев, знаний* приведенных таблице 8, рассчитываются матрицы семи системно-когнитивных моделей (таблица 9).

В таблице 8 приведены формулы:

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

И это сравнение в таблицах 6 и 7 осуществляется двумя возможными способами: путем **вычитания** и путем **деления**.

Таблица 8 3– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	через относительные частоты	через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке (таблица 7)	$N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $N_{ij} - \text{фактическая частота,}$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота.}$	
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	---	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу	---	
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \text{Log}_2 \frac{N_{ij} N}{N_i N_j}$
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	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$
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.

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 8), 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**.

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 Пирсона с замечательной мерой количества информации А.Харкевича и с известным в экономике коэффициентом ROI.

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

относительной частотой и вероятностью (погрешность) составляет около 5%, при 1250 наблюдениях – около 2.5%, при 10000 наблюдениях – 1%.

Таблица 9 – Матрица системно-когнитивной модели

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

Суть этих методов в том, что вычисляется количество информации в значении фактора о том, что объект моделирования перейдет под его действием в определенное состояние, соответствующее классу. Это позволяет сопоставимо и корректно обрабатывать разнородную информацию о наблюдениях объекта моделирования, представленную в различных типах измерительных шкал и различных единицах измерения [5].

На основе системно-когнитивных моделей, представленных в таблице 12 (отличаются частыми критериями, приведенными в таблице 8), 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 9).

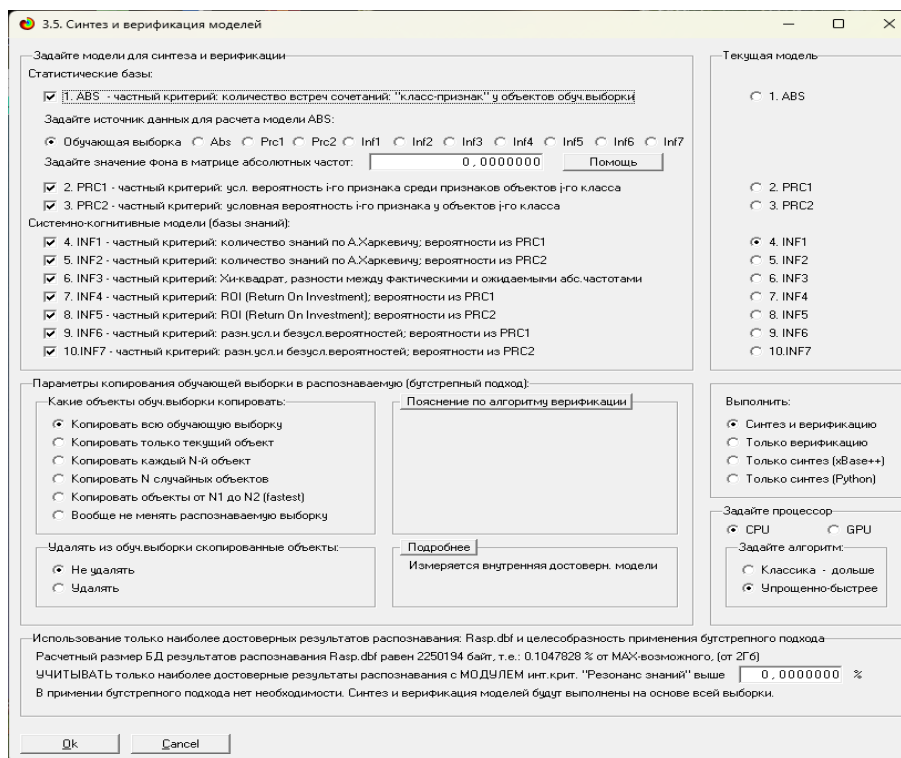
Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта

величина совпадает с **мощностью сигнала**, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 13):

Таблица 104 – 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	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

В системе «Эйдос» синтез моделей производится в режиме 3.5 (рисунок 8):



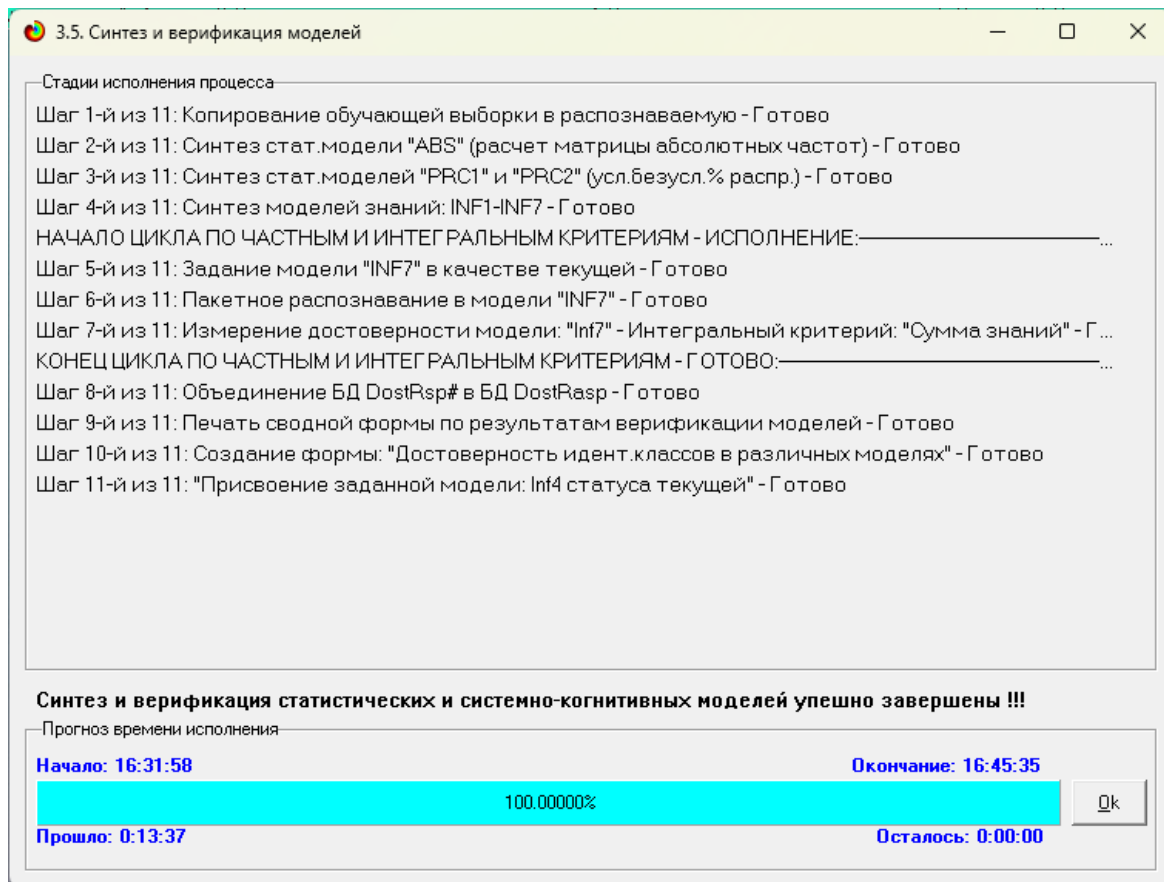


Рисунок 8. Экранные формы режима синтеза и верификации моделей

В результате работы режима 3.5 создано 3 статистических и 7 системно-когнитивных моделей, некоторые из которых приведены на рисунках 9-12:

Код признака	Наименование описательной шкалы и градаций	1. CALORIES 1/3 (1.0, 105.3)	2. CALORIES 2/3 (105.3, 209.7)	3. CALORIES 3/3 (209.7, 314.0)	Сумма	Среднее	Средн. квадрат. откл.
1.0	GENDER:female	4715.0	2728.0	110.0	7553.0	2517.67	2309.69
2.0	GENDER:male	4591.0	2433.0	423.0	7447.0	2482.33	2084.44
3.0	AGE:1/7(20.000000, 28.4285714)	2706.0	1282.0	20.0	4008.0	1336.00	1343.81
4.0	AGE:2/7(28.4285714, 36.8571429)	1728.0	857.0	39.0	2624.0	874.67	844.64
5.0	AGE:3/7(36.8571429, 45.2857143)	1480.0	833.0	68.0	2381.0	793.67	706.82
6.0	AGE:4/7(45.2857143, 53.7142857)	1018.0	650.0	58.0	1726.0	575.33	484.34
7.0	AGE:5/7(53.7142857, 62.1428571)	989.0	611.0	101.0	1701.0	567.00	445.63
8.0	AGE:6/7(62.1428571, 70.5714286)	690.0	483.0	112.0	1285.0	428.33	292.85
9.0	AGE:7/7(70.5714286, 79.0000000)	695.0	445.0	135.0	1275.0	425.00	280.54
10.0	HEIGHT:1/7(123.000000, 137.1428571)	22.0	8.0	1.0	31.0	10.33	10.69
11.0	HEIGHT:2/7(137.1428571, 151.2857143)	438.0	276.0	14.0	728.0	242.67	213.96
12.0	HEIGHT:3/7(151.2857143, 165.4285714)	2180.0	1297.0	60.0	3537.0	1179.00	1064.91
13.0	HEIGHT:4/7(165.4285714, 179.5714286)	3163.0	1750.0	130.0	5043.0	1681.00	1517.68
14.0	HEIGHT:5/7(179.5714286, 193.7142857)	2631.0	1369.0	222.0	4222.0	1407.33	1204.96
15.0	HEIGHT:6/7(193.7142857, 207.8571429)	825.0	439.0	92.0	1356.0	452.00	366.67
16.0	HEIGHT:7/7(207.8571429, 222.0000000)	47.0	22.0	14.0	83.0	27.67	17.21
17.0	WEIGHT:1/7(36.000000, 49.7142857)	218.0	123.0	5.0	346.0	115.33	106.71
18.0	WEIGHT:2/7(49.7142857, 63.4285714)	2226.0	1367.0	65.0	3658.0	1219.33	1088.04
19.0	WEIGHT:3/7(63.4285714, 77.1428571)	2912.0	1544.0	77.0	4533.0	1511.00	1417.79
20.0	WEIGHT:4/7(77.1428571, 90.8571429)	2377.0	1263.0	174.0	3814.0	1271.33	1101.52
21.0	WEIGHT:5/7(90.8571429, 104.5714286)	1375.0	765.0	172.0	2312.0	770.67	601.52
22.0	WEIGHT:6/7(104.5714286, 118.2857143)	186.0	97.0	35.0	318.0	106.00	75.90
23.0	WEIGHT:7/7(118.2857143, 132.0000000)	12.0	2.0	5.0	19.0	6.33	5.13
24.0	DURATION:1/7(1.000000, 5.1428571)	2262.0			2262.0	754.00	1305.97
25.0	DURATION:2/7(5.1428571, 9.2857143)	2073.0			2073.0	691.00	1196.85
26.0	DURATION:3/7(9.2857143, 13.4285714)	2102.0			2102.0	700.67	1213.59
27.0	DURATION:4/7(13.4285714, 17.5714286)	1990.0	148.0		2138.0	712.67	1109.60
28.0	DURATION:5/7(17.5714286, 21.7142857)	813.0	1246.0		2059.0	686.33	632.58
29.0	DURATION:6/7(21.7142857, 25.8571429)	65.0	1939.0	33.0	2037.0	679.00	1091.31
30.0	DURATION:7/7(25.8571429, 30.0000000)	1.0	1828.0	500.0	2329.0	776.33	944.33
31.0	HEART_RATE:1/7(67.000000, 75.7142857)	160.0			160.0	53.33	92.38
32.0	HEART_RATE:2/7(75.7142857, 84.4285714)	1937.0			1937.0	645.67	1118.33
33.0	HEART_RATE:3/7(84.4285714, 93.1428571)	4250.0	70.0		4320.0	1440.00	2433.78
34.0	HEART_RATE:4/7(93.1428571, 101.8571429)	2669.0	1532.0	5.0	4206.0	1402.00	1336.75
35.0	HEART_RATE:5/7(101.8571429, 110.5714286)	290.0	3031.0	212.0	3533.0	1177.67	1605.51
36.0	HEART_RATE:6/7(110.5714286, 119.2857143)		525.0	298.0	823.0	274.33	263.30
37.0	HEART_RATE:7/7(119.2857143, 128.0000000)		3.0	18.0	21.0	7.00	9.64
38.0	BODY_TEMP:1/7(37.100000, 37.2285714)	137.0			137.0	45.67	79.10
39.0	BODY_TEMP:2/7(37.2285714, 38.3571429)	531.0			531.0	177.00	306.57

Рисунок 9. Статистическая модель «ABS», матрица абсолютных частот (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	1 CALORIES 1/3 (1.0 105.3)	2 CALORIES 2/3 (105.3 203.7)	3 CALORIES 3/3 (203.7, 314.0)	Безул. вероятн.	Среднее	Средн. кв.ад. откл.
1.0	GENDER-female	50.666	52.858	20.638	50.353	41.387	18.039
2.0	GENDER-male	49.334	47.142	79.362	49.647	58.613	18.039
3.0	AGE-1/7-(20.000000, 28.4285714)	29.078	24.840	3.752	26.720	19.224	13.601
4.0	AGE-2/7-(28.4285714, 36.8571429)	18.569	16.605	7.317	17.493	14.164	6.046
5.0	AGE-3/7-(36.8571429, 45.2857143)	15.904	16.140	12.758	15.873	14.934	1.924
6.0	AGE-4/7-(45.2857143, 53.7142857)	10.939	12.594	10.882	11.507	11.472	1.008
7.0	AGE-5/7-(53.7142857, 62.1428571)	10.628	11.839	18.949	11.340	13.805	4.531
8.0	AGE-6/7-(62.1428571, 70.5714286)	7.415	9.359	21.013	8.567	12.595	7.390
9.0	AGE-7/7-(70.5714286, 79.0000000)	7.468	8.622	25.328	8.500	13.806	10.031
10.0	HEIGHT-1/7-(123.000000, 137.1428571)	0.236	0.155	0.188	0.207	0.193	0.068
11.0	HEIGHT-2/7-(137.1428571, 151.2857143)	4.707	5.348	2.627	4.853	4.227	1.458
12.0	HEIGHT-3/7-(151.2857143, 165.4285714)	23.426	25.131	11.257	23.580	19.938	7.602
13.0	HEIGHT-4/7-(165.4285714, 179.5714286)	33.989	33.908	24.390	33.620	30.762	5.554
14.0	HEIGHT-5/7-(179.5714286, 193.7142857)	28.272	26.526	41.651	28.147	32.150	8.310
15.0	HEIGHT-6/7-(193.7142857, 207.8571429)	8.865	8.506	17.261	9.040	11.544	4.990
16.0	HEIGHT-7/7-(207.8571429, 222.0000000)	0.505	0.426	2.627	0.553	1.186	1.283
17.0	WEIGHT-1/7-(36.0000000, 48.7142857)	2.343	2.383	0.938	2.307	1.888	0.858
18.0	WEIGHT-2/7-(48.7142857, 63.4285714)	23.920	26.487	12.195	24.387	20.867	7.655
19.0	WEIGHT-3/7-(63.4285714, 77.1428571)	31.292	29.917	14.447	30.220	25.218	9.390
20.0	WEIGHT-4/7-(77.1428571, 90.8571429)	25.543	24.472	32.645	25.427	27.553	4.478
21.0	WEIGHT-5/7-(90.8571429, 104.5714286)	14.775	14.823	32.270	15.413	20.623	10.123
22.0	WEIGHT-6/7-(104.5714286, 118.2857143)	1.999	1.879	6.567	2.120	3.482	2.708
23.0	WEIGHT-7/7-(118.2857143, 132.0000000)	0.129	0.039	0.938	0.127	0.369	0.530
24.0	DURATION-1/7-(1.000000, 5.1428571)	24.307			15.080	8.102	14.069
25.0	DURATION-2/7-(5.1428571, 9.2857143)	22.276			13.820	7.425	12.897
26.0	DURATION-3/7-(9.2857143, 13.4285714)	22.588			14.013	7.529	13.077
27.0	DURATION-4/7-(13.4285714, 17.5714286)	21.384	2.868		14.253	8.084	11.643
28.0	DURATION-5/7-(17.5714286, 21.7142857)	8.736	24.143		13.727	10.960	12.260
29.0	DURATION-6/7-(21.7142857, 25.8571429)	0.698	37.570	6.191	13.580	14.820	19.928
30.0	DURATION-7/7-(25.8571429, 30.0000000)	0.011	35.419	93.809	15.527	43.080	47.402
31.0	HEART_RATE-1/7-(67.0000000, 75.7142857)	1.719			1.067	0.573	1.028
32.0	HEART_RATE-2/7-(75.7142857, 84.4285714)	20.815			12.913	6.938	12.053
33.0	HEART_RATE-3/7-(84.4285714, 93.1428571)	45.669	1.356		28.800	15.675	26.020
34.0	HEART_RATE-4/7-(93.1428571, 101.8571429)	28.680	29.684	0.938	28.040	19.768	16.350
35.0	HEART_RATE-5/7-(101.8571429, 110.5714286)	3.116	58.729	39.775	23.553	33.873	28.308
36.0	HEART_RATE-6/7-(110.5714286, 119.2857143)		10.172	55.910	5.487	22.027	29.816
37.0	HEART_RATE-7/7-(119.2857143, 128.0000000)		0.058	3.377	0.140	1.145	1.969
38.0	BODY_TEMP-1/7-(37.100000, 37.7285714)	1.472			0.913	0.491	0.885
39.0	BODY_TEMP-2/7-(37.7285714, 38.3571429)	5.706			3.540	1.902	3.330

Рисунок -10. Статистическая модель «PRC2», матрица условных и безусловных процентных распределений (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	1 CALORIES 1/3 (1.0 105.3)	2 CALORIES 2/3 (105.3 203.7)	3 CALORIES 3/3 (203.7, 314.0)	Средн.	Среднее	Средн. кв.ад. откл.
1.0	GENDER-female	0.001	0.007	-0.122	-0.115	-0.038	0.073
2.0	GENDER-male	-0.001	-0.007	0.064	0.056	0.019	0.040
3.0	AGE-1/7-(20.000000, 28.4285714)	0.012	-0.010	-0.269	-0.268	-0.089	0.156
4.0	AGE-2/7-(28.4285714, 36.8571429)	0.008	-0.007	-0.119	-0.118	-0.039	0.070
5.0	AGE-3/7-(36.8571429, 45.2857143)	0.000	0.002	-0.030	-0.027	-0.009	0.018
6.0	AGE-4/7-(45.2857143, 53.7142857)	-0.007	0.012	-0.008	-0.002	-0.001	0.011
7.0	AGE-5/7-(53.7142857, 62.1428571)	-0.009	0.006	0.070	0.067	0.022	0.042
8.0	AGE-6/7-(62.1428571, 70.5714286)	-0.020	0.012	0.123	0.115	0.038	0.075
9.0	AGE-7/7-(70.5714286, 79.0000000)	-0.018	0.002	0.150	0.134	0.045	0.092
10.0	HEIGHT-1/7-(123.000000, 137.1428571)	0.018	-0.039	-0.013	-0.034	-0.011	0.029
11.0	HEIGHT-2/7-(137.1428571, 151.2857143)	-0.004	0.013	-0.084	-0.075	-0.025	0.052
12.0	HEIGHT-3/7-(151.2857143, 165.4285714)	-0.001	0.009	-0.101	-0.094	-0.031	0.061
13.0	HEIGHT-4/7-(165.4285714, 179.5714286)	0.001	0.001	-0.044	-0.041	-0.014	0.026
14.0	HEIGHT-5/7-(179.5714286, 193.7142857)	0.001	-0.008	0.054	0.046	0.015	0.033
15.0	HEIGHT-6/7-(193.7142857, 207.8571429)	-0.003	-0.008	0.089	0.078	0.026	0.054
16.0	HEIGHT-7/7-(207.8571429, 222.0000000)	-0.013	-0.036	0.234	0.165	0.055	0.138
17.0	WEIGHT-1/7-(36.0000000, 48.7142857)	0.002	0.004	-0.123	-0.117	-0.039	0.073
18.0	WEIGHT-2/7-(48.7142857, 63.4285714)	-0.003	0.011	-0.095	-0.086	-0.029	0.058
19.0	WEIGHT-3/7-(63.4285714, 77.1428571)	0.005	-0.001	-0.101	-0.098	-0.033	0.059
20.0	WEIGHT-4/7-(77.1428571, 90.8571429)	0.001	-0.005	0.034	0.030	0.010	0.021
21.0	WEIGHT-5/7-(90.8571429, 104.5714286)	-0.006	-0.005	0.101	0.090	0.030	0.062
22.0	WEIGHT-6/7-(104.5714286, 118.2857143)	-0.008	-0.017	0.155	0.130	0.043	0.097
23.0	WEIGHT-7/7-(118.2857143, 132.0000000)	0.002	-0.162	0.274	0.115	0.038	0.221
24.0	DURATION-1/7-(1.000000, 5.1428571)	0.065			0.065	0.022	0.038
25.0	DURATION-2/7-(5.1428571, 9.2857143)	0.065			0.065	0.022	0.038
26.0	DURATION-3/7-(9.2857143, 13.4285714)	0.065			0.065	0.022	0.038
27.0	DURATION-4/7-(13.4285714, 17.5714286)	0.056	-0.220		-0.164	-0.055	0.146
28.0	DURATION-5/7-(17.5714286, 21.7142857)	-0.062	0.077		0.015	0.005	0.070
29.0	DURATION-6/7-(21.7142857, 25.8571429)	-0.407	0.140	-0.108	-0.375	-0.125	0.274
30.0	DURATION-7/7-(25.8571429, 30.0000000)	-0.997	0.113	0.247	-0.638	-0.213	0.683
31.0	HEART_RATE-1/7-(67.0000000, 75.7142857)	0.065			0.065	0.022	0.038
32.0	HEART_RATE-2/7-(75.7142857, 84.4285714)	0.065			0.065	0.022	0.038
33.0	HEART_RATE-3/7-(84.4285714, 93.1428571)	0.063	-0.419		-0.356	-0.119	0.262
34.0	HEART_RATE-4/7-(93.1428571, 101.8571429)	0.003	0.008	-0.466	-0.455	-0.152	0.272
35.0	HEART_RATE-5/7-(101.8571429, 110.5714286)	-0.277	0.125	0.072	-0.800	-0.027	0.219
36.0	HEART_RATE-6/7-(110.5714286, 119.2857143)		0.085	0.318	0.403	0.134	0.165
37.0	HEART_RATE-7/7-(119.2857143, 128.0000000)		-0.120	0.436	0.316	0.105	0.293
38.0	BODY_TEMP-1/7-(37.100000, 37.7285714)	0.065			0.065	0.022	0.038
39.0	BODY_TEMP-2/7-(37.7285714, 38.3571429)	0.065			0.065	0.022	0.038

Рисунок 11. Системно-когнитивная модель «INF1», матрица информативностей (по А.Харкевичу) (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	1. CALORIES 1/3 (1.0, 116.3)	2. CALORIES 2/3 (105.3, 203.7)	3. CALORIES 3/3 (209.7, 314.0)	Сумма	Среднее	Средн. квадрат. откл.
1.0	GENDER:female	0.006	0.050	-0.590	-0.534	-0.178	0.358
2.0	GENDER:male	-0.006	-0.050	0.599	0.542	0.181	0.363
3.0	AGE:1/7:(20.000000, 28.4285714)	0.088	-0.070	-0.860	-0.842	-0.281	0.508
4.0	AGE:2/7:(28.4285714, 36.8571429)	0.061	-0.051	-0.582	-0.571	-0.190	0.344
5.0	AGE:3/7:(36.8571429, 45.2857143)	0.002	0.017	-0.196	-0.178	-0.059	0.119
6.0	AGE:4/7:(45.2857143, 53.7142857)	-0.049	0.095	-0.054	-0.009	-0.003	0.085
7.0	AGE:5/7:(53.7142857, 62.1428571)	-0.063	0.044	0.671	0.652	0.217	0.396
8.0	AGE:6/7:(62.1428571, 70.5714286)	-0.134	0.092	1.453	1.411	0.470	0.858
9.0	AGE:7/7:(70.5714286, 79.0000000)	-0.121	0.014	1.980	1.873	0.624	1.176
10.0	HEIGHT:1/7:(123.000000, 137.1428571)	0.144	-0.250	-0.092	-0.198	-0.066	0.198
11.0	HEIGHT:2/7:(137.1428571, 151.2857143)	-0.030	0.102	-0.459	-0.387	-0.129	0.293
12.0	HEIGHT:3/7:(151.2857143, 165.4285714)	-0.007	0.066	-0.523	-0.463	-0.154	0.321
13.0	HEIGHT:4/7:(165.4285714, 179.5714286)	0.011	0.009	-0.275	-0.255	-0.085	0.164
14.0	HEIGHT:5/7:(179.5714286, 193.7142857)	0.004	-0.058	0.480	0.427	0.142	0.294
15.0	HEIGHT:6/7:(193.7142857, 207.8571429)	-0.019	-0.059	0.909	0.831	0.277	0.548
16.0	HEIGHT:7/7:(207.8571429, 222.0000000)	-0.087	-0.230	3.747	3.430	1.143	2.256
17.0	WEIGHT:1/7:(36.000000, 49.7142857)	0.016	0.033	-0.593	-0.545	-0.182	0.357
18.0	WEIGHT:2/7:(49.7142857, 63.4285714)	-0.019	0.086	-0.500	-0.433	-0.144	0.312
19.0	WEIGHT:3/7:(63.4285714, 77.1428571)	0.035	-0.010	-0.522	-0.497	-0.166	0.310
20.0	WEIGHT:4/7:(77.1428571, 90.8571429)	0.005	-0.038	0.284	0.251	0.084	0.175
21.0	WEIGHT:5/7:(90.8571429, 104.5714286)	-0.041	-0.038	1.094	1.014	0.338	0.654
22.0	WEIGHT:6/7:(104.5714286, 118.2857143)	-0.057	-0.113	2.097	1.927	0.642	1.261
23.0	WEIGHT:7/7:(118.2857143, 132.0000000)	0.018	-0.694	6.406	5.730	1.910	3.910
24.0	DURATION:1/7:(1.000000, 5.1428571)	0.612			0.612	0.204	0.353
25.0	DURATION:2/7:(5.1428571, 9.2857143)	0.612			0.612	0.204	0.353
26.0	DURATION:3/7:(9.2857143, 13.4285714)	0.612			0.612	0.204	0.353
27.0	DURATION:4/7:(13.4285714, 17.5714286)	0.500	-0.799		-0.299	-0.100	0.655
28.0	DURATION:5/7:(17.5714286, 21.7142857)	-0.364	0.759		0.395	0.132	0.573
29.0	DURATION:6/7:(21.7142857, 25.8571429)	-0.949	1.767	-0.544	0.274	0.091	1.465
30.0	DURATION:7/7:(25.8571429, 30.0000000)	-0.999	1.281	5.042	5.324	1.775	3.051
31.0	HEART_RATE:1/7:(67.000000, 75.7142857)	0.612			0.612	0.204	0.353
32.0	HEART_RATE:2/7:(75.7142857, 84.4285714)	0.612			0.612	0.204	0.353
33.0	HEART_RATE:3/7:(84.4285714, 93.1428571)	0.586	-0.953		-0.367	-0.122	0.777
34.0	HEART_RATE:4/7:(93.1428571, 101.8571429)	0.023	0.059	-0.967	-0.885	-0.295	0.582
35.0	HEART_RATE:5/7:(101.8571429, 110.5714286)	-0.868	1.493	0.689	1.314	0.438	1.200
36.0	HEART_RATE:6/7:(110.5714286, 119.2857143)		0.854	9.190	10.044	3.348	5.077
37.0	HEART_RATE:7/7:(119.2857143, 128.0000000)		-0.585	23.122	22.537	7.512	13.522
38.0	BODY_TEMP:1/7:(37.100000, 37.7285714)	0.612			0.612	0.204	0.353
39.0	BODY_TEMP:2/7:(37.7285714, 38.3571429)	0.612			0.612	0.204	0.353

Рисунок 12 – Системно-когнитивная модель «INF4», частный критерий ROI (Return On Investment) (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 Ризбергера, 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 F-мерой Ван Ризбергера наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием: «Семантический резонанс знаний»: F=0.863 (1-й рисунок 13).

В соответствии с критерием достоверности L1-мерой проф. Е.В.Луценко модели СК-модель INF4 и INF5 также, являются наиболее достоверными L1=0.920 (2-й рисунок 14). **Модель INF4 и будем использовать для решения поставленных в работе задач.**

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

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Ф-мера Ван Райсбергера	Средняя модель: уровень сход. истинно-полож. решений (STP)	Средняя модель: уровень сход. истинно-отриц. решений (STN)	Средняя модель: уровень ложно-полож. решений
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	15000	14991	6334	23666	9	0.388	0.999	0.559	8474.584	838.243	5978
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сходна абс частот по призна...	15000	15000	30000	30000	9	0.333	1.000	0.500	10293.801		7083
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отн.частот с о...	15000	14991	6334	23666	9	0.388	0.999	0.559	8474.584	838.243	5978
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Сходна усл.отн.частот по при...	15000	15000	30000	30000	9	0.333	1.000	0.500	8752.084		11926
3. PRC2 - частный критерий: условная вероятность иго признака.	Корреляция усл.отн.частот с о...	15000	14991	6334	23666	9	0.388	0.999	0.559	8474.329	838.239	5978
3. PRC2 - частный критерий: условная вероятность иго признака.	Сходна усл.отн.частот по при...	15000	15000	30000	30000	9	0.333	1.000	0.500	8752.084		11926
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	15000	14526	23287	6713	474	0.684	0.968	0.802	3818.947	8166.894	885
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Сходна знаниям	15000	13041	23490	6510	1959	0.667	0.869	0.755	1704.344	6749.630	746
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	15000	14526	23287	6713	474	0.684	0.968	0.802	3818.948	8166.897	885
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Сходна знаниям	15000	13041	23490	6510	1959	0.667	0.869	0.755	1704.346	6749.632	746
6. INF3 - частный критерий: Хинкварт, разности между факти...	Семантический резонанс зна...	15000	13374	22950	7050	1626	0.655	0.892	0.755	5293.753	7726.144	1668
6. INF3 - частный критерий: Хинкварт, разности между факти...	Сходна знаниям	15000	13374	22950	7050	1626	0.655	0.892	0.755	7093.618	7732.957	1096
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	15000	13101	27749	2251	1899	0.853	0.873	0.863	5793.204	10294.515	620
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Сходна знаниям	15000	13872	19130	10870	1128	0.561	0.925	0.698	1007.858	801.344	869
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	15000	13101	27749	2251	1899	0.853	0.873	0.863	5793.207	10294.517	620
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Сходна знаниям	15000	13872	19130	10870	1128	0.561	0.925	0.698	1007.858	801.344	869
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	15000	13291	22150	7850	1709	0.629	0.886	0.736	5005.063	7520.265	1534
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сходна знаниям	15000	13374	19145	10855	1626	0.552	0.892	0.682	2607.442	3137.441	2023
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	15000	13291	22150	7850	1709	0.629	0.886	0.736	5005.063	7520.265	1534
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сходна знаниям	15000	13374	19145	10855	1626	0.552	0.892	0.682	2607.442	3137.441	2023

Помощь по мерам достоверности | Помощь по частотным распределениям | TR:TN,FP:FN | (TP-FP)/(TN-FN) | (T-F)/(T+F)*100 | Задать интервал сглаживания

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

Наименование модели и частного критерия	Интегральный критерий	из модел. ный сход. ный сход. ный сход.	Средняя модель: уровень сход. истинно-полож. решений (STP)	Средняя модель: уровень сход. истинно-отриц. решений (STN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В.Луценко	Средняя модель: уровень сход. истинно-полож. решений	Средняя модель: уровень сход. истинно-отриц. решений	Средняя модель: уровень ложно-полож. решений	Средняя модель: уровень ложно-отриц. решений	
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	474.584	838.243	5978.946	0.446	0.586	1.000	0.739	0.565	0.132	0.253	0.05
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сходна абс частот по призна...	283.801		7083.178		0.592	1.000	0.744	0.686		0.236	
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отн.частот с о...	474.584	838.243	5978.945	0.446	0.586	1.000	0.739	0.565	0.132	0.253	0.05
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Сходна усл.отн.частот по при...	752.084		11926.601		0.423	1.000	0.595	0.583		0.398	
3. PRC2 - частный критерий: условная вероятность иго признака.	Корреляция усл.отн.частот с о...	474.329	838.239	5978.865	0.446	0.586	1.000	0.739	0.565	0.132	0.253	0.05
3. PRC2 - частный критерий: условная вероятность иго признака.	Сходна усл.отн.частот по при...	752.084		11926.601		0.423	1.000	0.595	0.583		0.398	
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	818.947	8166.894	885.363	55.329	0.812	0.986	0.890	0.263	0.351	0.132	0.11
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Сходна знаниям	704.344	6749.630	746.840	152.039	0.695	0.918	0.791	0.131	0.287	0.115	0.07
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	818.948	8166.897	885.363	55.329	0.812	0.986	0.890	0.263	0.351	0.132	0.11
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Сходна знаниям	704.346	6749.632	746.840	152.039	0.695	0.918	0.791	0.131	0.287	0.115	0.07
6. INF3 - частный критерий: Хинкварт, разности между факти...	Семантический резонанс зна...	293.753	7726.144	1698.504	288.538	0.757	0.948	0.842	0.396	0.337	0.241	0.17
6. INF3 - частный критерий: Хинкварт, разности между факти...	Сходна знаниям	093.618	7732.957	1066.420	427.080	0.869	0.943	0.905	0.530	0.337	0.151	0.26
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	793.204	10294.515	620.823	381.465	0.903	0.938	0.920	0.442	0.371	0.276	0.21
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Сходна знаниям	007.858	801.344	869.674	19.908	0.537	0.981	0.694	0.073	0.042	0.080	0.01
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	793.207	10294.517	620.823	381.465	0.903	0.938	0.920	0.442	0.371	0.276	0.21
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Сходна знаниям	007.858	801.344	869.674	19.908	0.537	0.981	0.694	0.073	0.042	0.080	0.01
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	005.063	7520.265	1534.552	326.313	0.765	0.939	0.843	0.377	0.340	0.195	0.15
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сходна знаниям	607.442	3137.441	2023.689	102.948	0.563	0.962	0.710	0.195	0.164	0.186	0.06
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	005.063	7520.265	1534.552	326.313	0.765	0.939	0.843	0.377	0.340	0.195	0.15
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сходна знаниям	607.442	3137.441	2023.689	102.948	0.563	0.962	0.710	0.195	0.164	0.186	0.06

Помощь по мерам достоверности | Помощь по частотным распределениям | TR:TN,FP:FN | (TP-FP)/(TN-FN) | (T-F)/(T+F)*100 | Задать интервал сглаживания

Рисунок 13. Экранные формы режима измерения достоверности моделей 3.4

На рисунках 14 приведены частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4.

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

- отрицательные ложные решения встречаются на уровне различия до 60%, но истинно отрицательные решения превышают;
- на уровне сходства меньше 8-9% преобладают ложные положительные решения, а при более высоких уровнях сходства – истинные положительные решения. При уровнях сходства выше 80% ложных положительных решений вообще нет;

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

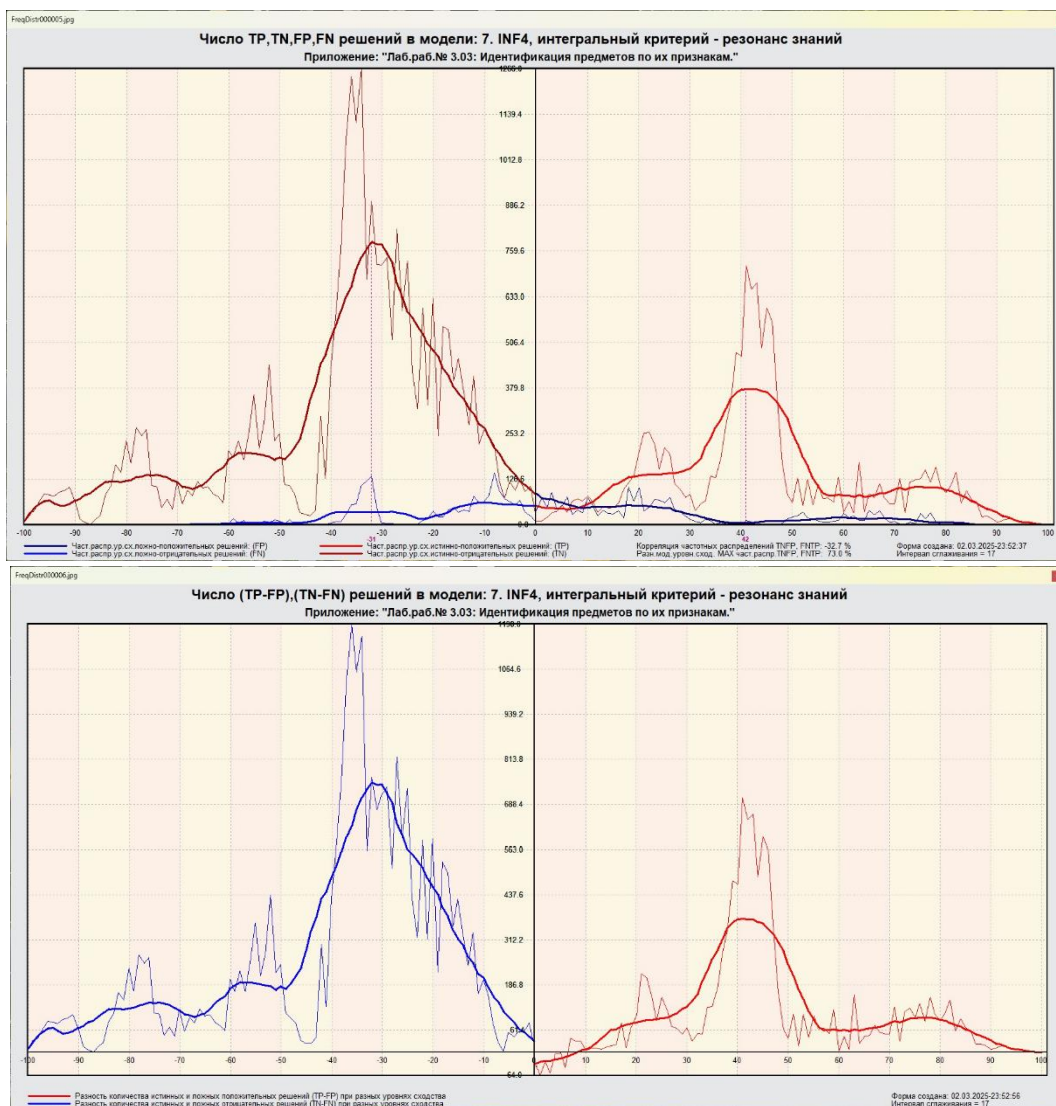


Рисунок 14. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF4

На рисунках 15 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 режимам: 3.4, 4.1.3.6, 4.1.3.7, 4.1.3.8, 4.1.3.10: Виды прогнозов и меры достоверности моделей в системе "Эйдос-Х++"

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Динамика F-меры Ван Ризбергена и L1- и L2- мер проф.Е.В.Луценко в зависимости от объема выборки RND-модели.

Рисунок 45. Экранные формы хелпов режима измерения достоверности моделей

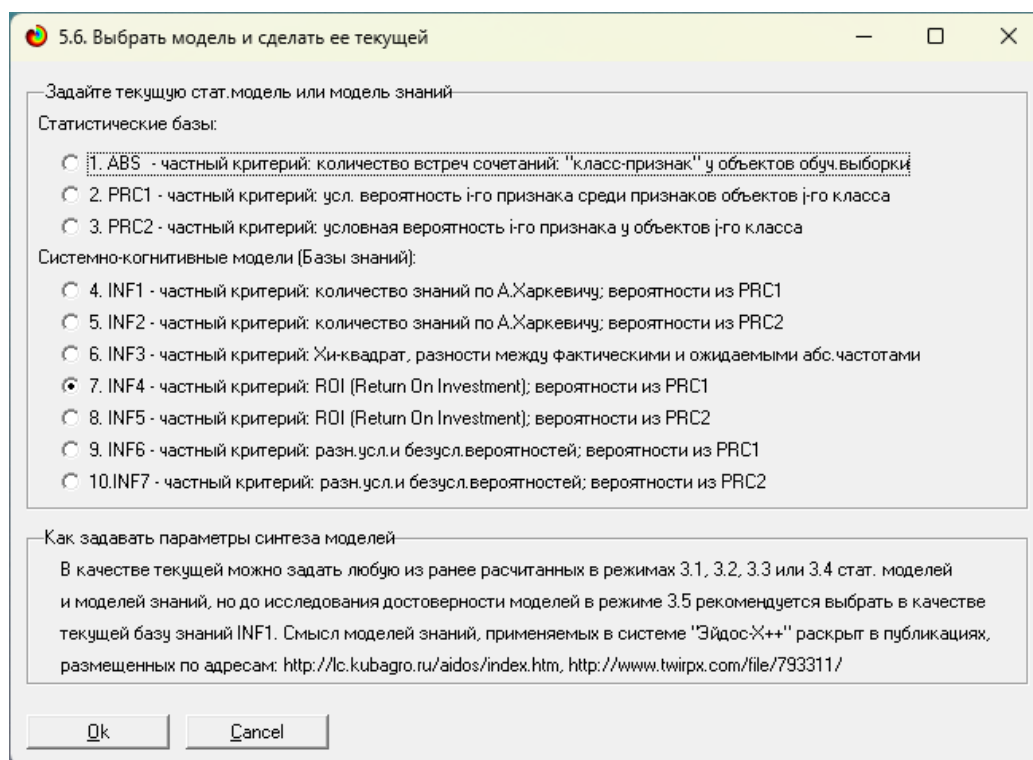
3.5. Задача-5. Выбор наиболее достоверной модели

Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

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

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (рисунки 16). Это необходимо делать лишь для решения задачи идентификации и прогнозирования (в режиме 4.1.2), которая требует и потребляет наибольшие вычислительные ресурсы и поэтому решается только для модели, заданной в качестве текущей. Все остальные расчеты проводятся в системе «Эйдос» сразу во всех моделях.



корректны⁹ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1. Интегральный критерий «Сумма знаний»

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

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

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

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

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

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

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

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

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

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

3.6.2. Интегральный критерий «Семантический резонанс знаний»

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

⁹ В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

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

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

где:

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

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

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

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

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

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

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

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

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

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

3.6.3. Важные математические свойства интегральных критериев

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

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

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

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

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

В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется **вес** каждого обобщенного образа класса в образе объекта, что подробнее описано в монографии [8, 9].

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

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

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

освящены и на теоретическом уровне, и с детальными численными примерами в работах [10, 11] и в ряде других¹⁰.

Поэтому в данной работе рассмотрим стандартный несколько упрощенный вариант решения задачи идентификации и прогнозирования. Запустим режим 4.1.2 системы «Эйдос» (рисунок 17):

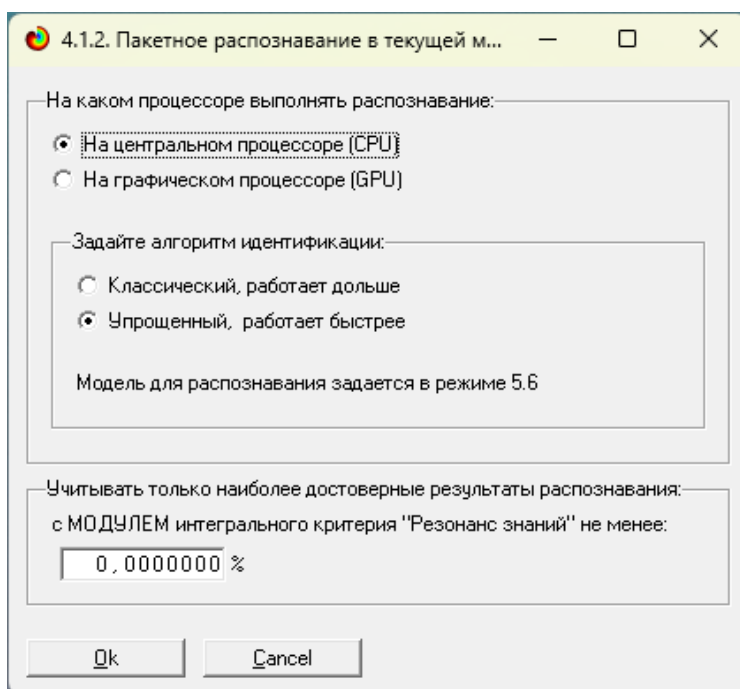


Рисунок 17. Экранные формы режима 4.1.2 идентификации и прогнозирования

По результатам решения задачи идентификации и прогнозирования в системе «Эйдос» выдается большое количество разнообразных выходных форм, которых в настоящее время 12 (рисунок 18):

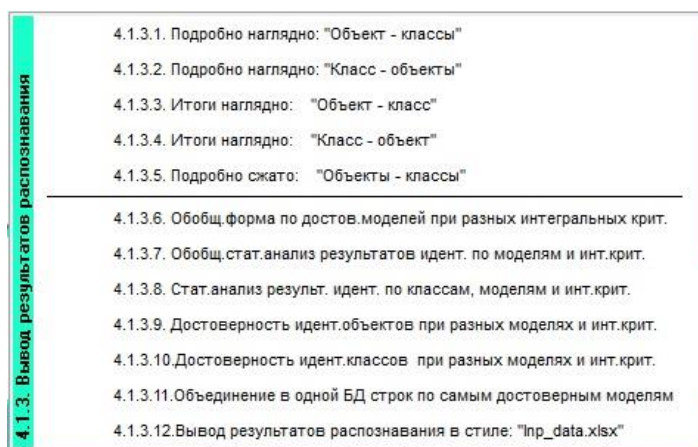


Рисунок 18. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

¹⁰ См., например: http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm

Из этих выходных форм рассмотрим только две: 4.1.3.1 и 4.1.3.2 (рисунок 19):

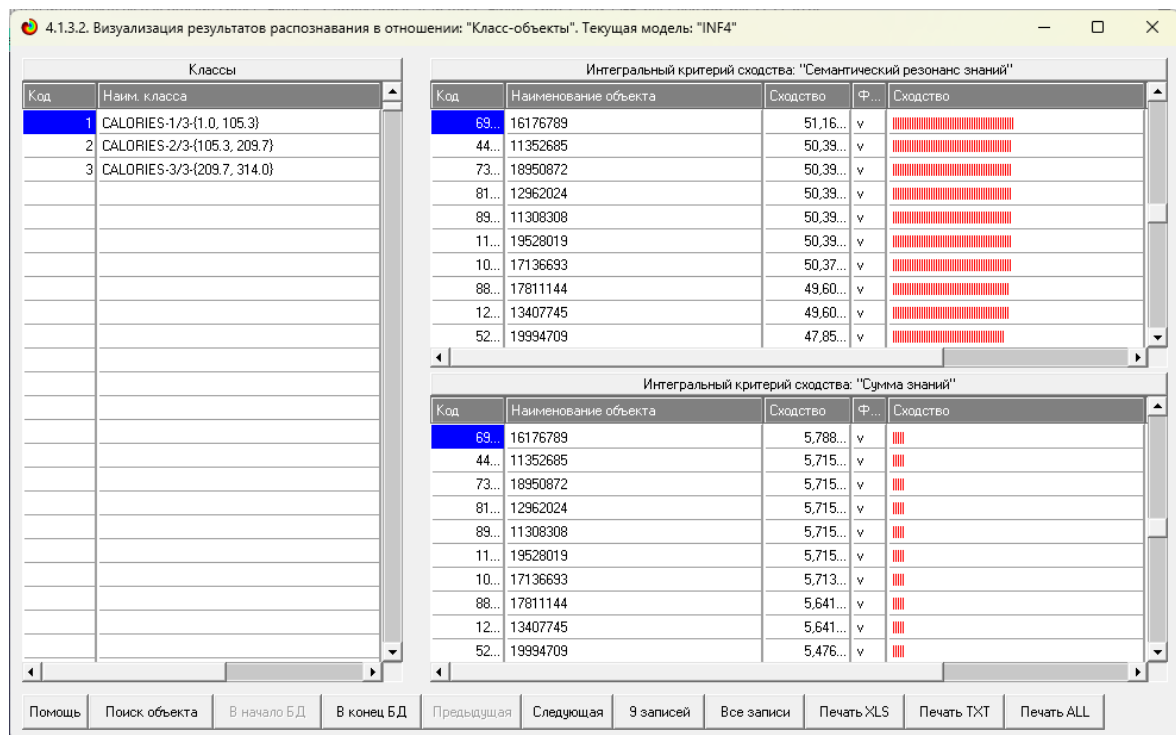
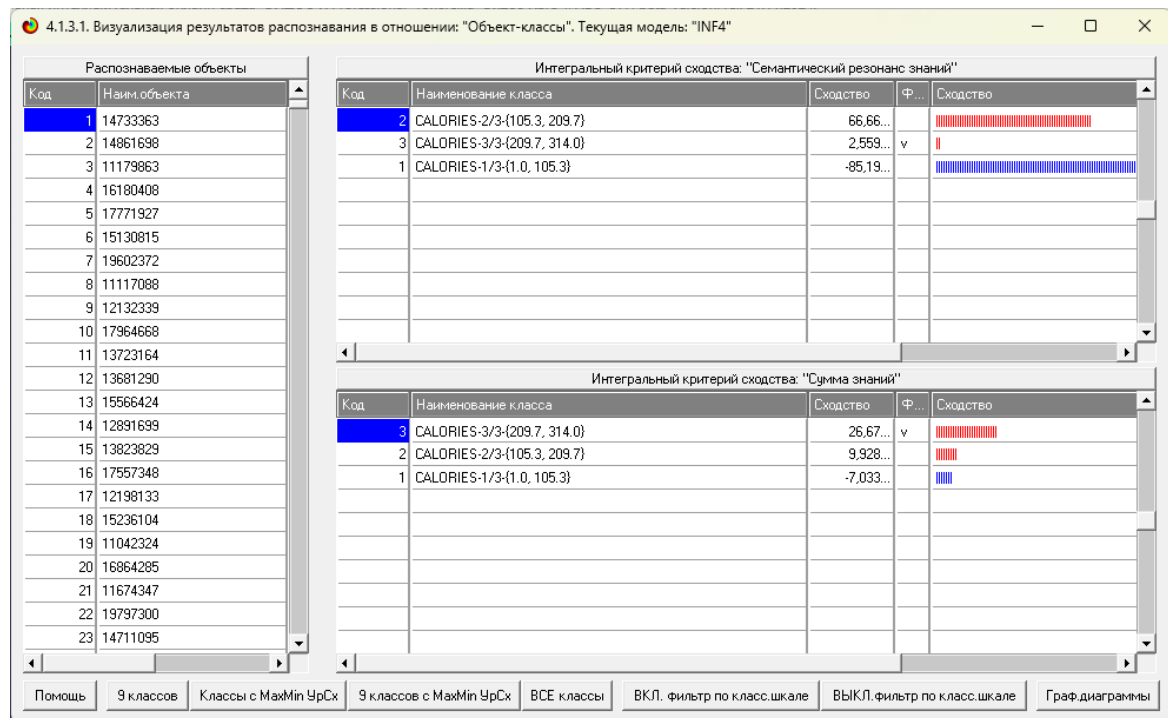


Рисунок 196. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

3.7. Задача-7. Поддержка принятия решений

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

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

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

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

Таким образом, задача принятия решений является обратной по отношению к задаче прогнозирования. Но это так только в простейшем случае: в случае использования SWOT-анализа (режим 4.4.8 системы «Эйдос») [12] (рисунки 20).

Выходные формы, приведенные на рисунках 20, 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.

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

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	CALORIES-1/3-{1.0, 105.3}	0.4299995	65142	62.0400000
2	CALORIES-2/3-{105.3, 209.7}	0.5746561	36127	34.4066667
3	CALORIES-3/3-{209.7, 314.0}	4.2985805	3731	3.5533333

SWOT-анализ класса:1 "CALORIES-1/3-{1.0, 105.3}" в модели:7 "INF4"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
38	WEIGHT-12/12-{124.0000000, 132.0000000}	0.612
65	BODY_TEMP-3/12-{37.8333333, 38.2000000}	0.612
43	DURATION-5/12-{10.6666667, 13.0833333}	0.612
66	BODY_TEMP-4/12-{38.2000000, 38.5666667}	0.612
68	BODY_TEMP-6/12-{38.9333333, 39.3000000}	0.612
39	DURATION-1/12-{1.0000000, 3.4166667}	0.612
40	DURATION-2/12-{3.4166667, 5.8333333}	0.612
41	DURATION-3/12-{5.8333333, 8.2500000}	0.612
42	DURATION-4/12-{8.2500000, 10.6666667}	0.612
53	HEART_RATE-3/12-{77.1666667, 82.2500000}	0.612
64	BODY_TEMP-2/12-{37.4666667, 37.8333333}	0.612
69	BODY_TEMP-7/12-{39.3000000, 39.6666667}	0.612
67	BODY_TEMP-5/12-{38.5666667, 38.9333333}	0.612
52	HEART_RATE-2/12-{72.0833333, 77.1666667}	0.612

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
49	DURATION-11/12-{25.1666667, 27.5833333}	-0.998
59	HEART_RATE-9/12-{107.6666667, 112.7500000}	-0.991
48	DURATION-10/12-{22.7500000, 25.1666667}	-0.975
73	BODY_TEMP-11/12-{40.7666667, 41.1333333}	-0.945
58	HEART_RATE-8/12-{102.5833333, 107.6666667}	-0.871
47	DURATION-9/12-{20.3333333, 22.7500000}	-0.801
72	BODY_TEMP-10/12-{40.4000000, 40.7666667}	-0.642
15	HEIGHT-1/12-{123.0000000, 131.2500000}	-0.463
57	HEART_RATE-7/12-{97.5000000, 102.5833333}	-0.331
46	DURATION-8/12-{17.9166667, 20.3333333}	-0.243
14	AGE-12/12-{74.0833333, 79.0000000}	-0.143
12	AGE-10/12-{64.2500000, 69.1666667}	-0.141
37	WEIGHT-11/12-{116.0000000, 124.0000000}	-0.121
13	AGE-11/12-{69.1666667, 74.0833333}	-0.096

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

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

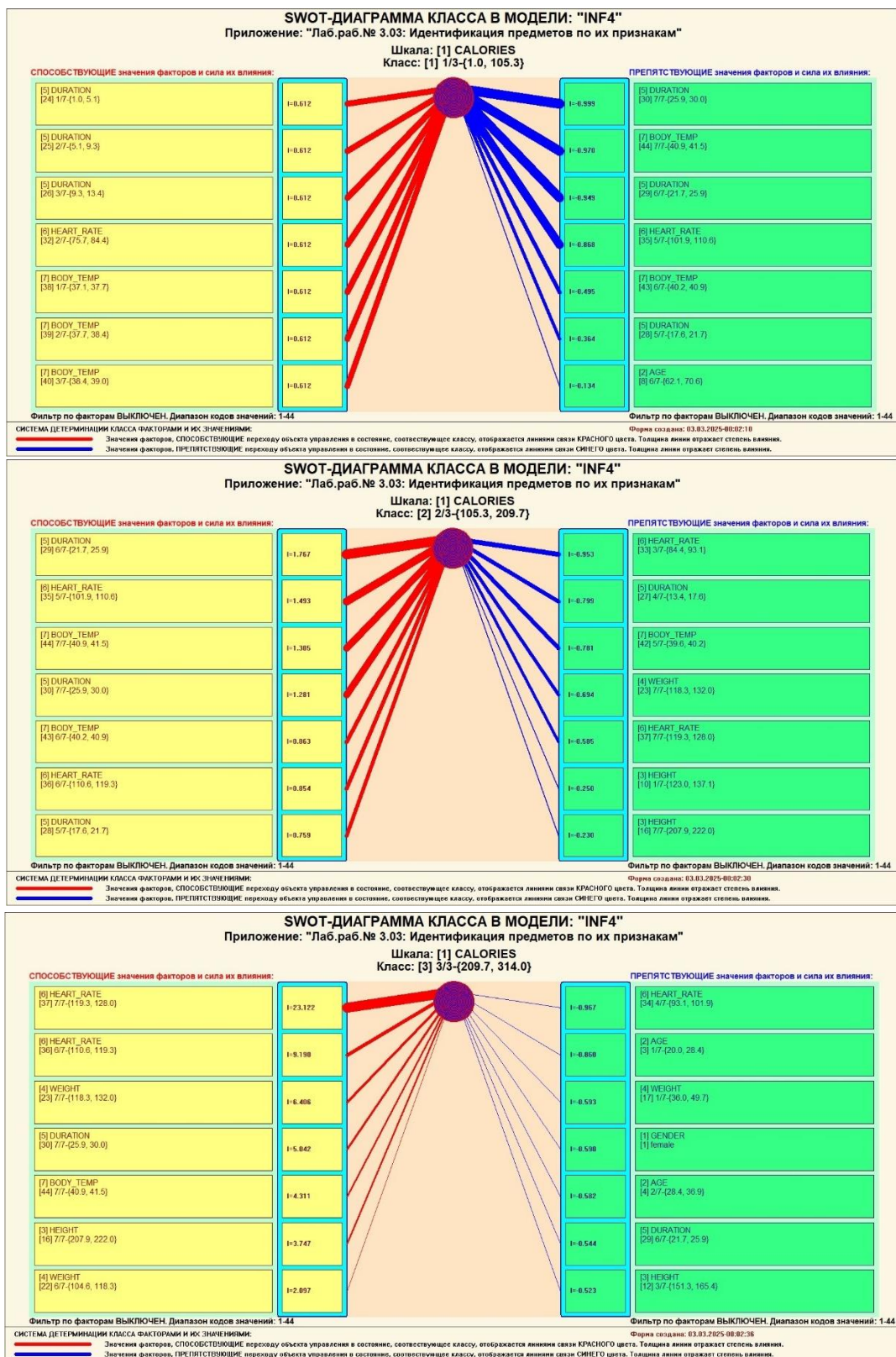


Рисунок 20. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

На первом рисунке 20 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.

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.

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

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

Поэтому в АСК-анализе и системе «Эйдос» реализован развитый алгоритм принятия решений (режим 6.3) в котором кроме SWOT-анализа используются также результаты решения задачи прогнозирования и результаты кластерно-конструктивного анализа классов и значений факторов, т.е. некоторые результаты решения задачи исследования предметной области. Этот алгоритм описан в работах [10, 11, 13] и в ряде других работ.

Приведем этот алгоритм и в данной работе (рисунок 22).

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 [13, 14, 15].

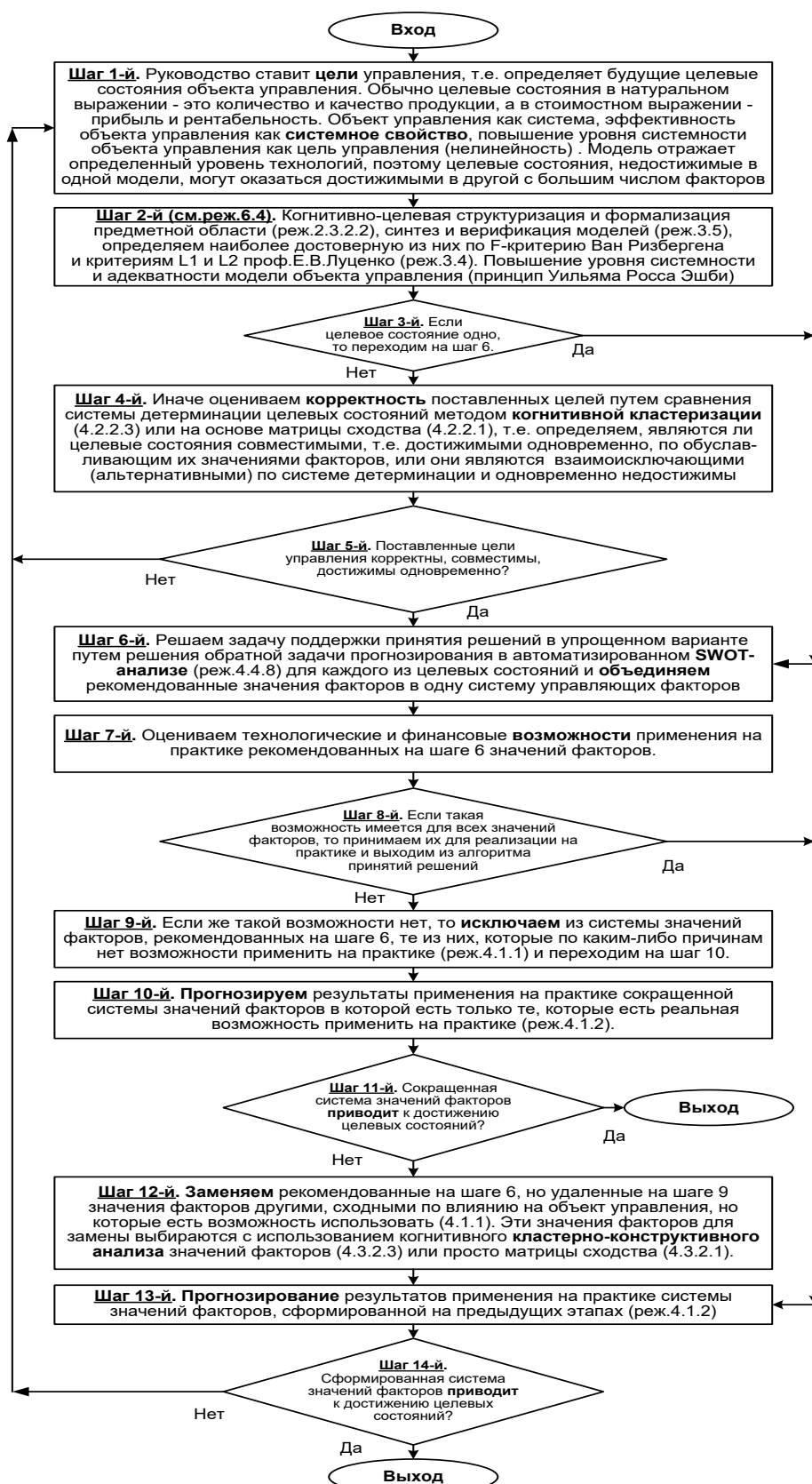


Рисунок 21. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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) [4]. Повышение уровня системности и адекватности модели объекта управления (принцип Уильяма Росса Эшби) [14].

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 [12].

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 22:

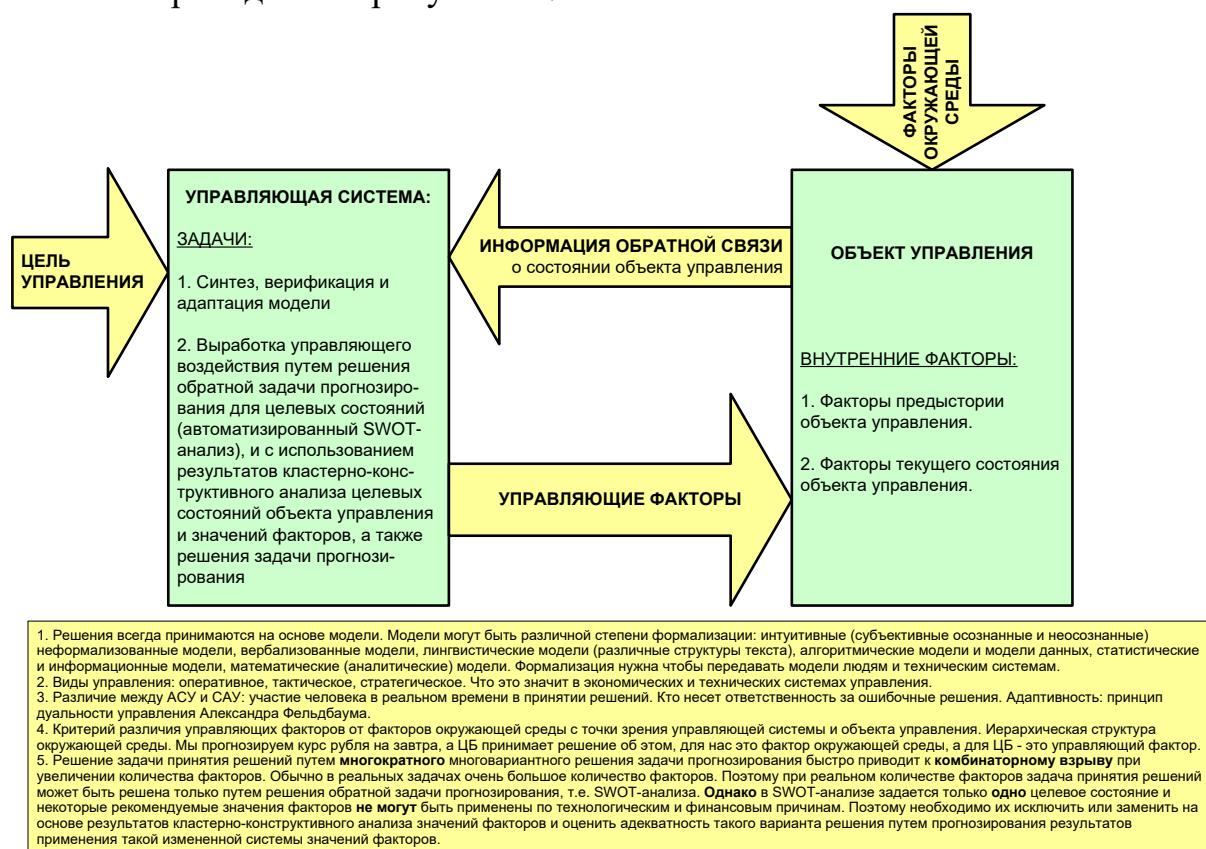


Рисунок 22. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

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 [18, 19].

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

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

Инвертированные SWOT-диаграмм (предложены автором в работе [15]), 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».

Инвертированные SWOT-диаграммы для каждого значения фактора, которые представляют собой лингвистические переменные, приведены ниже на рисунке 23:

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

Код	Наименование значения фактора
1	GENDER-female
2	GENDER-male
3	AGE-1/7-(20.0000000, 28.4285714)
4	AGE-2/7-(28.4285714, 36.8571429)
5	AGE-3/7-(36.8571429, 45.2857143)
6	AGE-4/7-(45.2857143, 53.7142857)

SWOT-анализ значения фактора:1 "GENDER-female" в модели:7 "INF4"

СПОСОБСТВУЕТ:

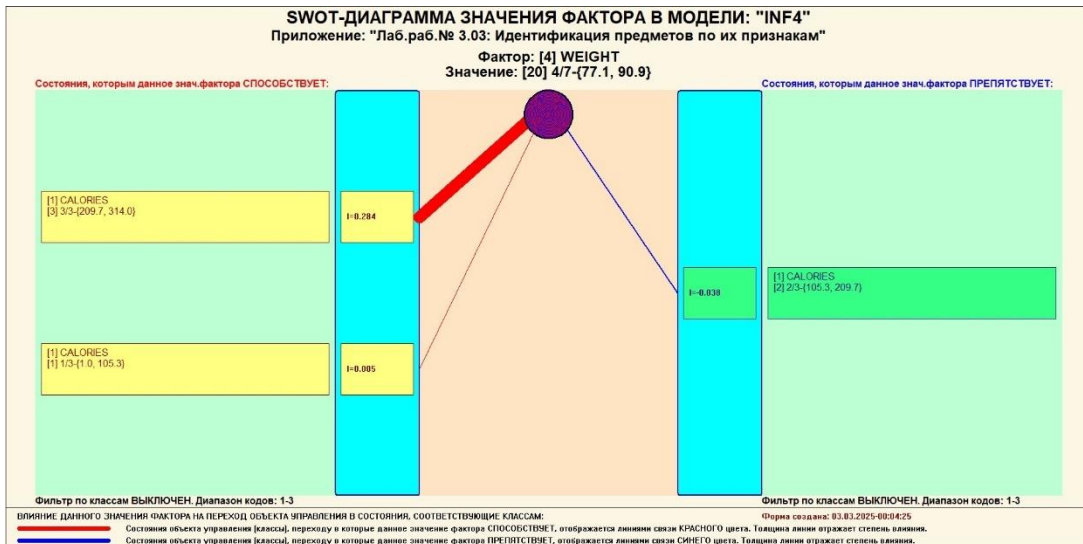
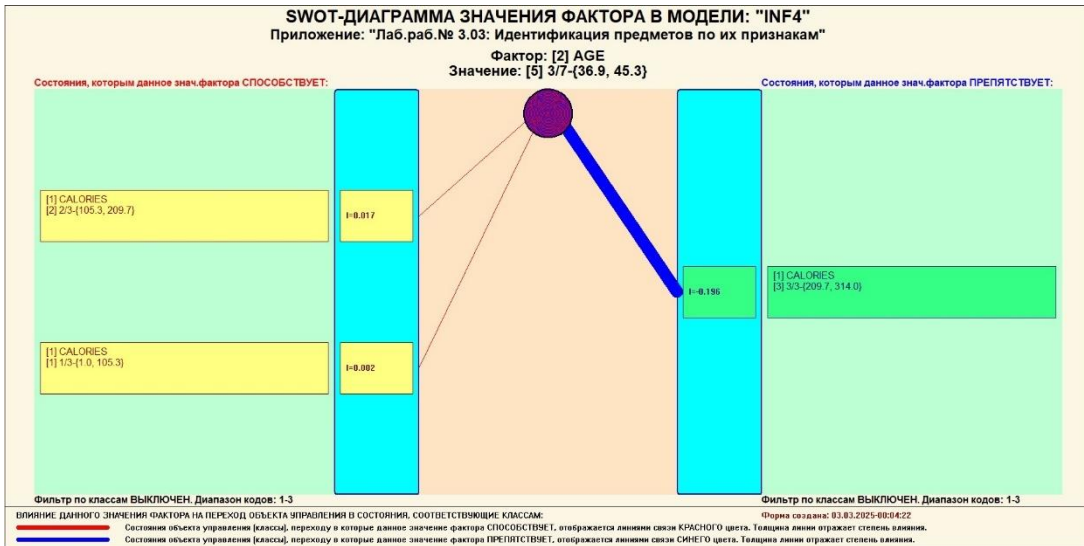
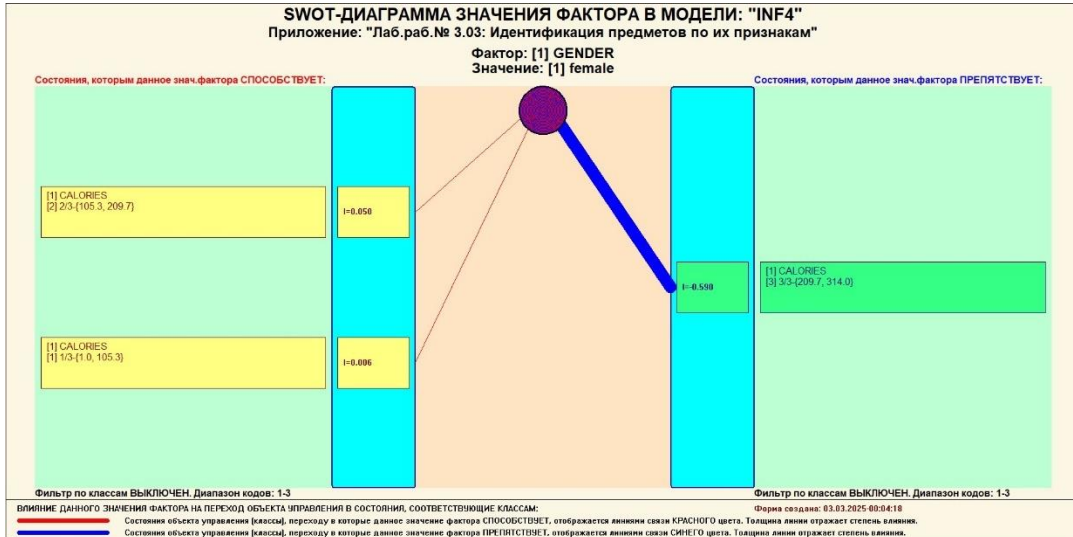
Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влияния
2	CALORIES-2/3-(105.3, 209.7)	0.050
1	CALORIES-1/3-(1.0, 105.3)	0.006

ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
3	CALORIES-3/3-(209.7, 314.0)	-0.590

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

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



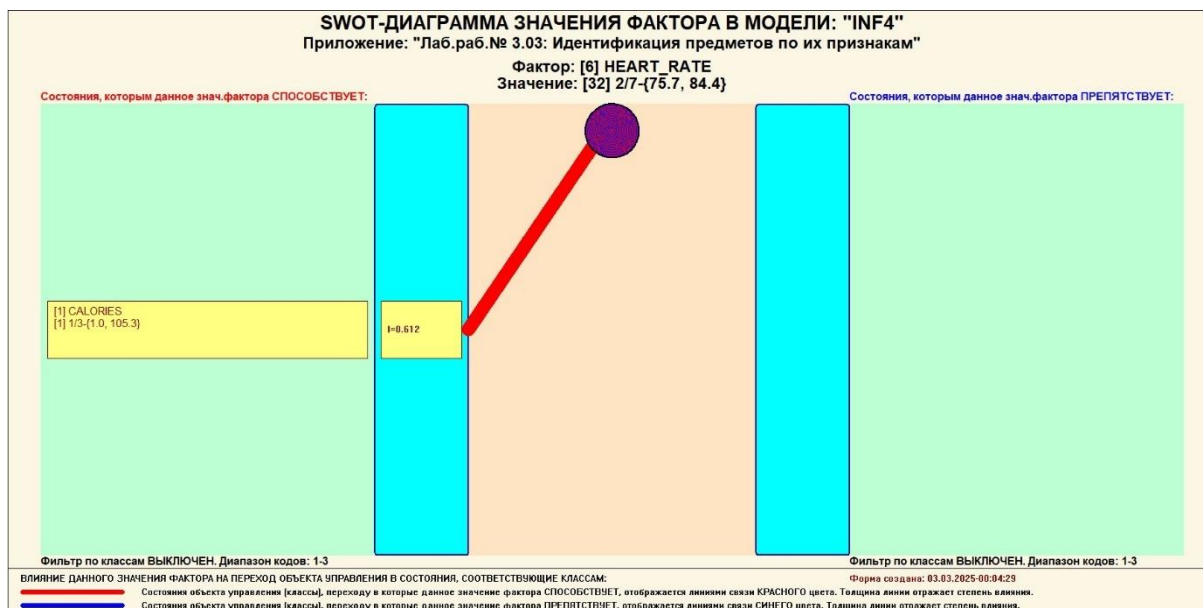


Рисунок 23. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 24 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.

3.8.2. Кластерно-конструктивный анализ классов

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 24) рассчитывается матрица сходства классов (таблица 11) по системе их детерминации и на основе этой матрицы рассчитывается и выводится три основных формы:

- круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2) (рисунок 25);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации классов* (предложена автором в 2011 году в работе [16]) (режим 4.2.2.3) (рисунок 26);
- график изменения межкластерных расстояний (режим 4.2.2.3) (рисунок 27).

Эта матрица сходства (таблица 11) используются и при расчете некоторых других выходных форм.

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

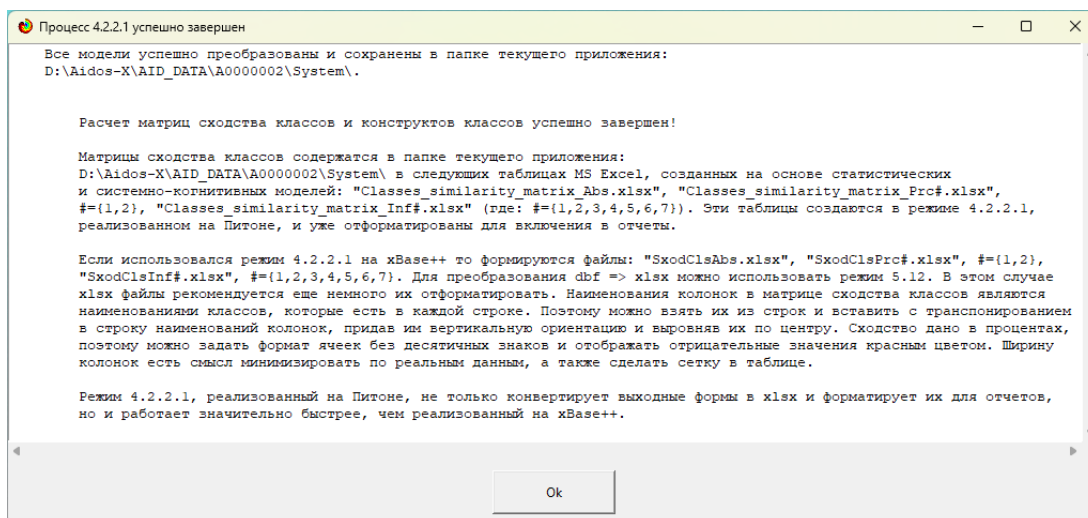
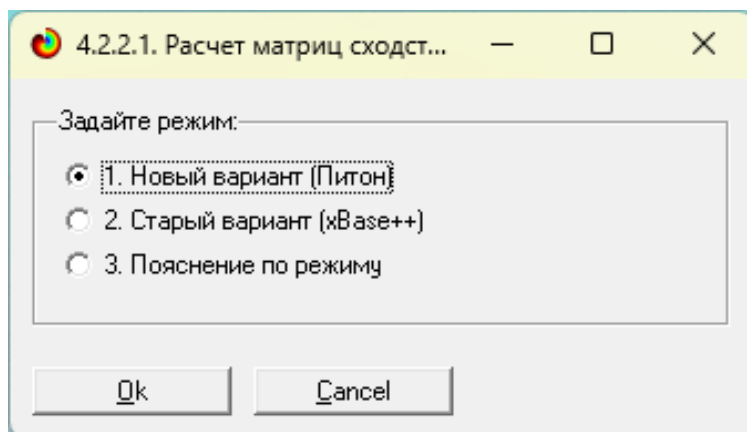


Рисунок 24. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

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

	1-CALORIES-1/3-{1.0, 105.3}	2-CALORIES-2/3-{105.3, 209.7}	3-CALORIES-3/3-{209.7, 314.0}
1-CALORIES-1/3-{1.0, 105.3}	100	-74,569	-18,0035
2-CALORIES-2/3-{105.3, 209.7}	-74,569	100	-4,46969
3-CALORIES-3/3-{209.7, 314.0}	-18,0035	-4,46969	100

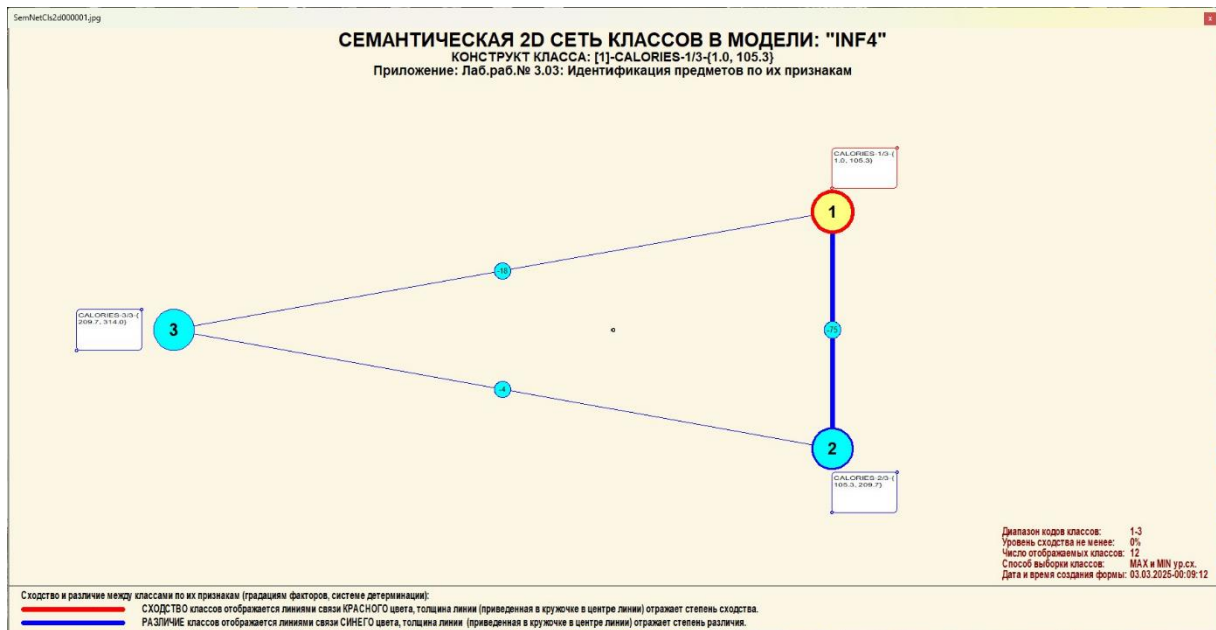


Рисунок 25. Круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2)

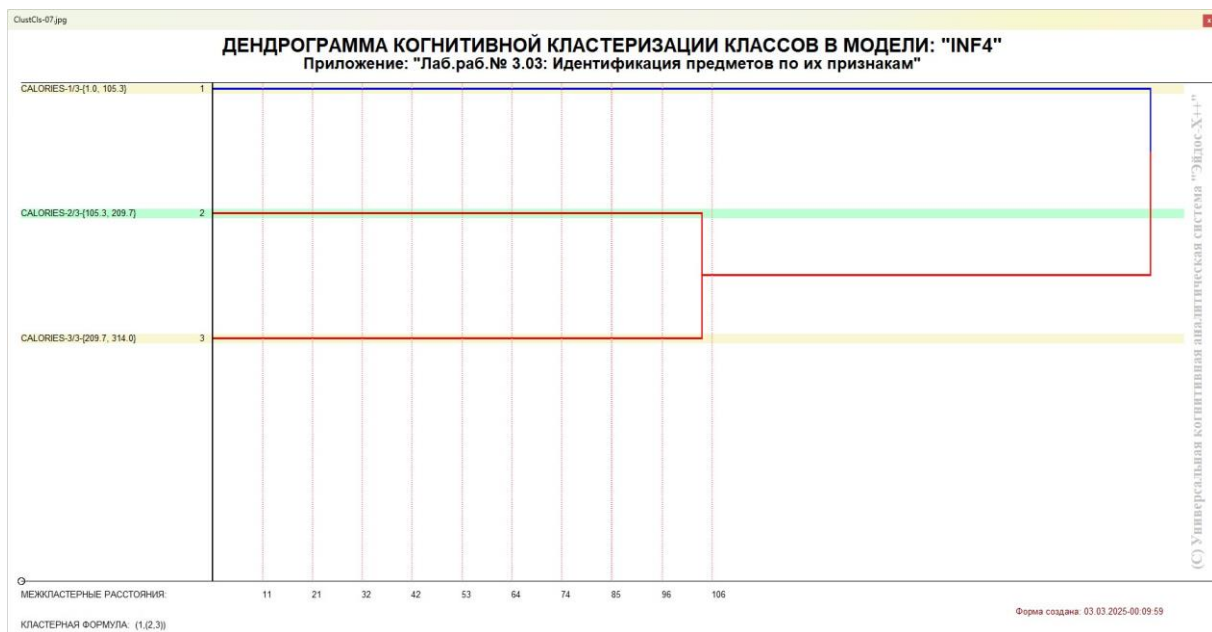


Рисунок 26. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.3)

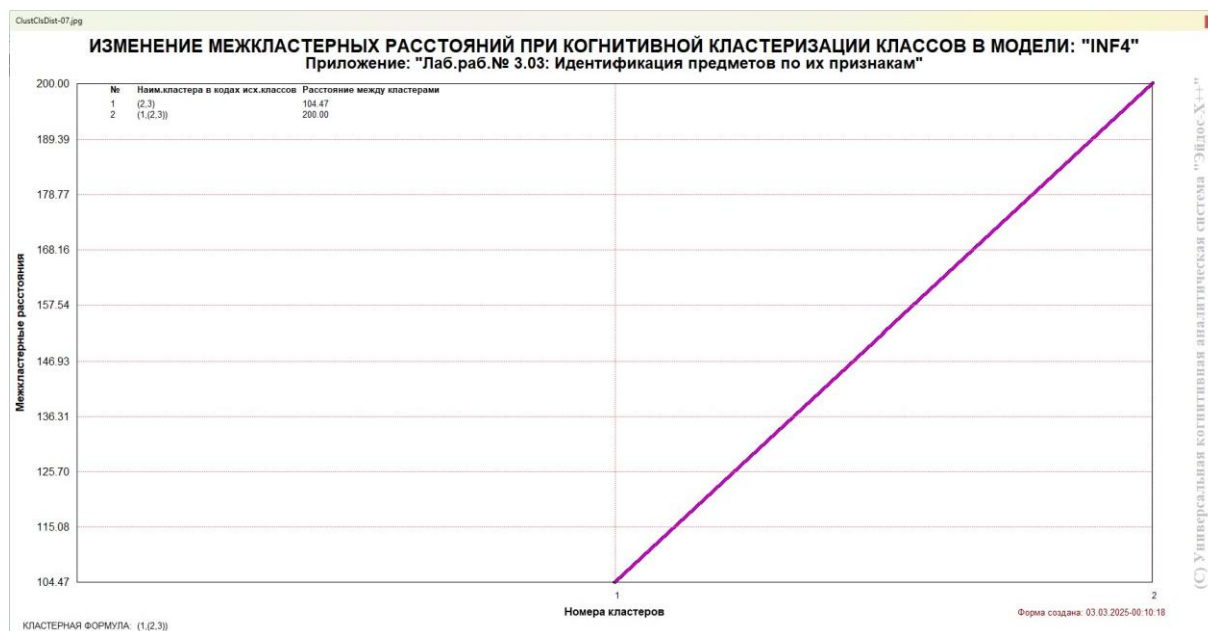


Рисунок 27. График изменения межкластерных расстояний (режим 4.2.2.3)

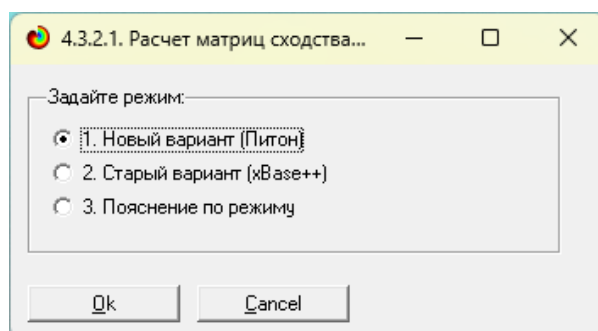
3.8.3. Кластерно-конструктивный анализ значений описательных шкал

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 28) рассчитывается матрица сходства признаков (таблица 12) по их смыслу и на основе этой матрицы рассчитывается и выводится четыре основных формы:

- круговая 2d-когнитивная диаграмма признаков (режим 4.3.2.2) рисунок 29);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации признаков* (предложена автором в 2011 году в работе [16]) (режим 4.3.2.3) рисунок 30);
- график изменения межкластерных расстояний (режим 4.3.2.3) рисунок 30).

Эта матрица сходства (таблица 12) используются и при расчете некоторых других выходных форм.

На рисунке 28 представлены экранные формы режима 4.3.2.1, обеспечивающего расчет матрицы сходства значений факторов по силе и направлению их влияния на переход объекта моделирования в различные будущие состояния, соответствующие классам:



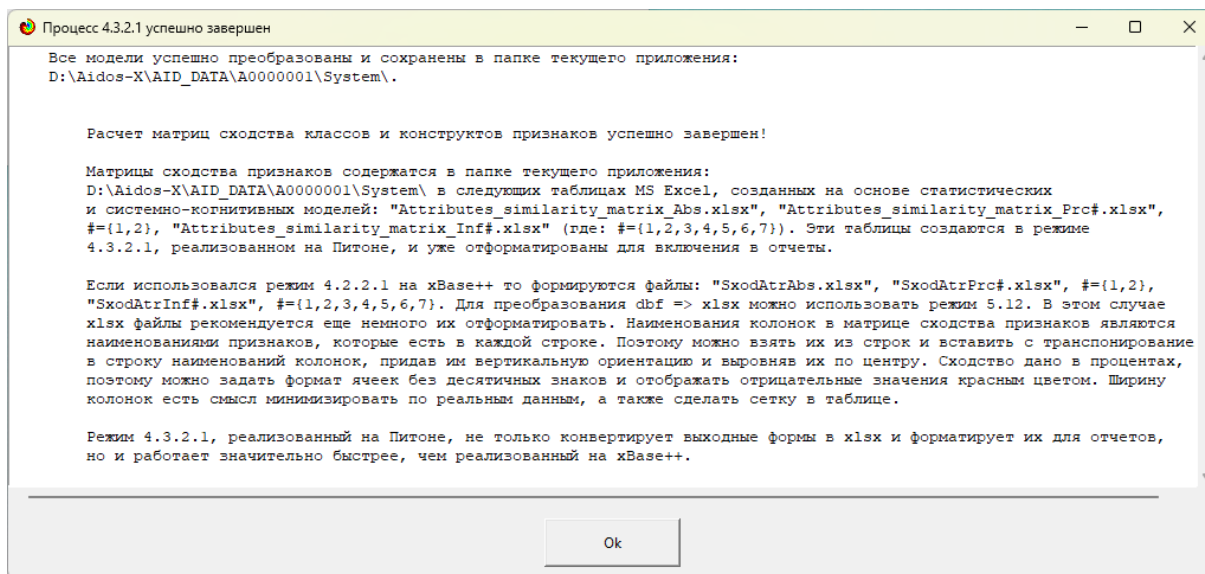


Рисунок 28. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

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

	1 - GENDER-female	2 - GENDER-male	3 - AGE-1/7-{20.0000000, 28.4285714}	4 - AGE-2/7-{28.4285714, 36.8571429}
1 - GENDER-female	100	-100	97,63842	97,47977
2 - GENDER-male	-100	100	-97,6384	-97,4798
3 - AGE-1/7-{20.0000000, 28.4285714}	97,63842	-97,6384	100	99,99739
4 - AGE-2/7-{28.4285714, 36.8571429}	97,47977	-97,4798	99,99739	100

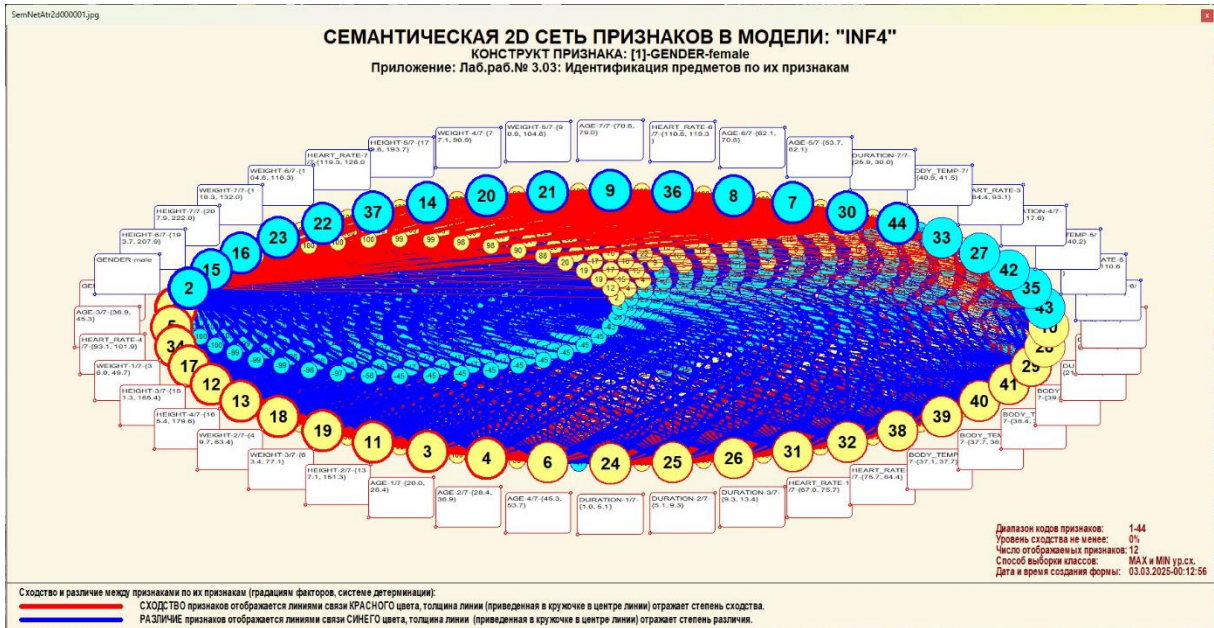


Рисунок 297. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF4 (режим 4.3.2.2)

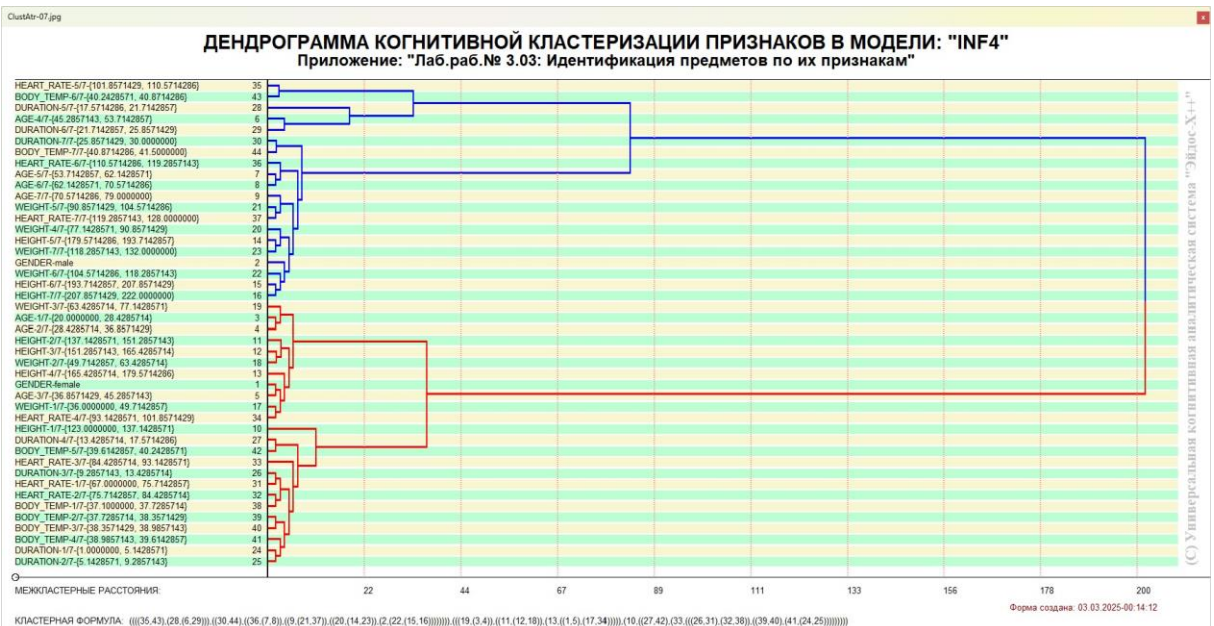


Рисунок 30. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)



Рисунок 30. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

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

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

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

От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что [17]:

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

4.4.10.Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	CALORIES-1/3-{1.0, 105.3}
2	CALORIES-2/3-{105.3, 209.7}
3	CALORIES-3/3-{209.7, 314.0}

Подготовка визуализации нейрона:1 "CALORIES-1/3-{1.0, 105.3}" в модели:7 "INF4"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
24	DURATION-1/7-{1.0000000, 5.1428571}	0.612
25	DURATION-2/7-{5.1428571, 9.2857143}	0.612
26	DURATION-3/7-{9.2857143, 13.4285714}	0.612
32	HEART_RATE-2/7-{75.7142857, 84.4285714}	0.612
38	BODY_TEMP-1/7-{37.1000000, 37.7285714}	0.612
39	BODY_TEMP-2/7-{37.7285714, 38.3571429}	0.612
40	BODY_TEMP-3/7-{38.3571429, 38.9857143}	0.612
41	BODY_TEMP-4/7-{38.9857143, 39.6142857}	0.612
31	HEART_RATE-1/7-{67.0000000, 75.7142857}	0.612

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
30	DURATION-7/7-{25.8571429, 30.0000000}	-0.999
44	BODY_TEMP-7/7-{40.8714286, 41.5000000}	-0.970
29	DURATION-6/7-{21.7142857, 25.8571429}	-0.949
35	HEART_RATE-5/7-{101.8571429, 110.5714286}	-0.868
43	BODY_TEMP-6/7-{40.2428571, 40.8714286}	-0.495
28	DURATION-5/7-{17.5714286, 21.7142857}	-0.364
8	AGE-6/7-{62.1428571, 70.5714286}	-0.134
9	AGE-7/7-{70.5714286, 79.0000000}	-0.121
16	HEIGHT-7/7-{207.8571429, 222.0000000}	-0.087

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

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

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Сортировать рецепторы: Отображать рецепторы:

по информативности с наименованиями
 по модулю информативности только с кодами

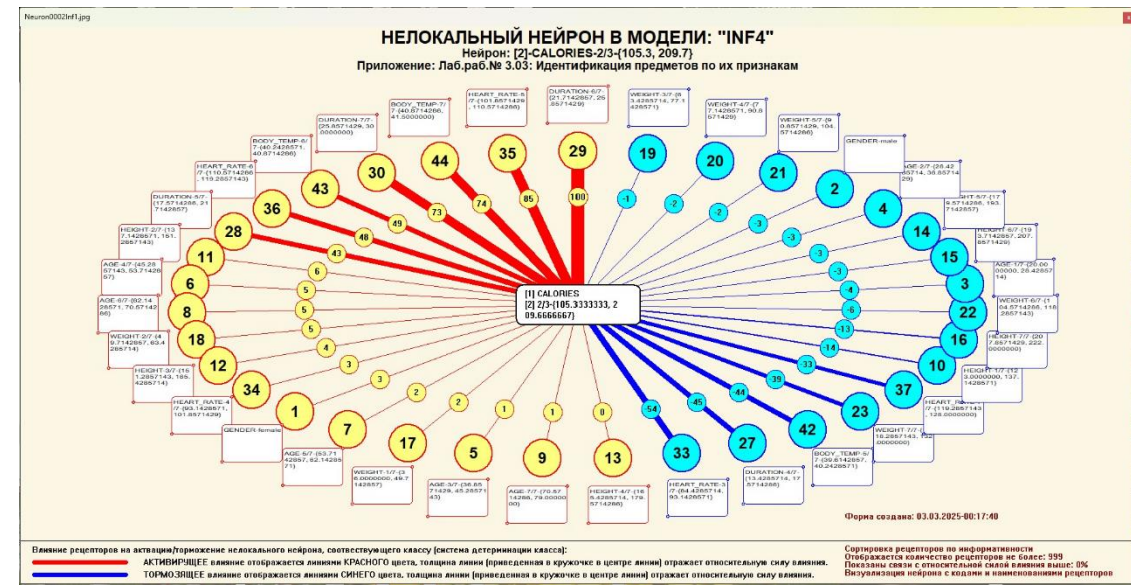
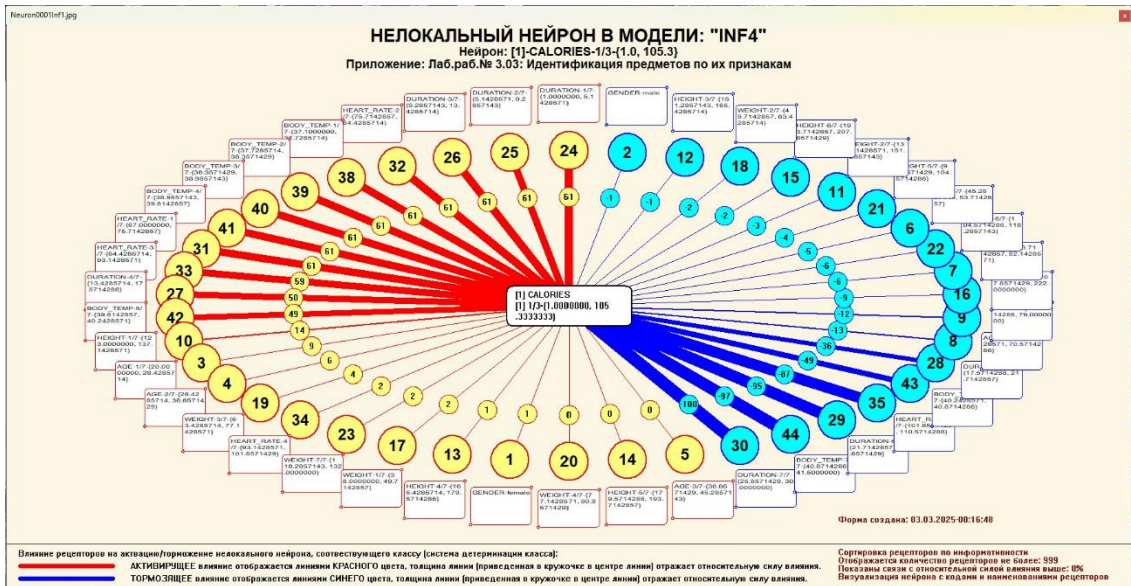


Рисунок 32. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

В системе «Эйдос» есть возможность построения моделей, соответствующих многослойным нейронным сетям [17].

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

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

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	CALORIES-1/3-{1.0, 105.3}
<input type="checkbox"/>	2	CALORIES-2/3-{105.3, 209.7}
<input type="checkbox"/>	3	CALORIES-3/3-{209.7, 314.0}

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

Подготовка визуализации нейрона: 1 "CALORIES-1/3-{1.0, 105.3}" в модели: 7 "INF4"

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

Код	Наименование фактора и его интервального значения	Сила влияния
24	DURATION-1/7-{1.000000, 5.1428571}	0.612
25	DURATION-2/7-{5.1428571, 9.2857143}	0.612
26	DURATION-3/7-{9.2857143, 13.4285714}	0.612
32	HEART_RATE-2/7-{75.7142857, 84.4285714}	0.612
38	BODY_TEMP-1/7-{37.1000000, 37.7285714}	0.612
39	BODY_TEMP-2/7-{37.7285714, 38.3571429}	0.612
40	BODY_TEMP-3/7-{38.3571429, 38.9857143}	0.612
41	BODY_TEMP-4/7-{38.9857143, 39.6142857}	0.612
31	HEART_RATE-1/7-{67.0000000, 75.7142857}	0.612

Код	Наименование фактора и его интервального значения	Сила влияния
30	DURATION-7/7-{25.8571429, 30.0000000}	-0.999
44	BODY_TEMP-7/7-{40.8714286, 41.5000000}	-0.970
29	DURATION-6/7-{21.7142857, 25.8571429}	-0.949
35	HEART_RATE-5/7-{101.8571429, 110.5714286}	-0.868
43	BODY_TEMP-6/7-{40.2428571, 40.8714286}	-0.495
28	DURATION-5/7-{17.5714286, 21.7142857}	-0.364
8	AGE-6/7-{62.1428571, 70.5714286}	-0.134
9	AGE-7/7-{70.5714286, 79.0000000}	-0.121
16	HEIGHT-7/7-{207.8571429, 222.0000000}	-0.087

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

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

Сортировать связи: Отображать наименования:

по модулю информативности нейронов
 по информативности и знаку рецепторов

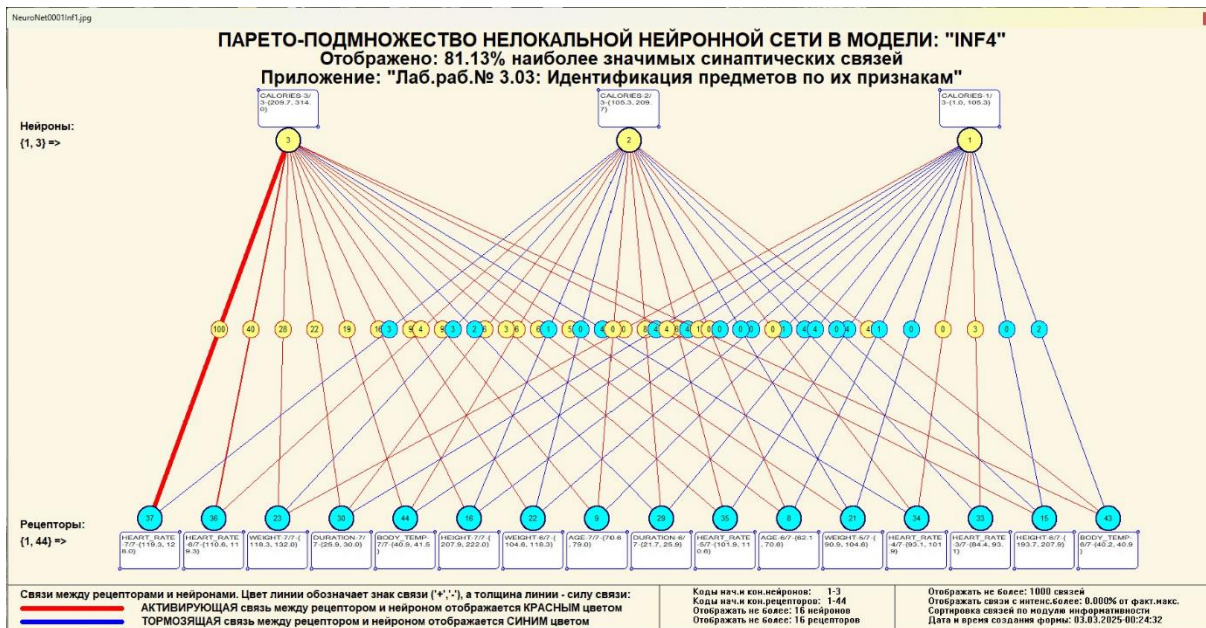


Рисунок 33. Нейронная сеть в СК-модели INF4

3.8.6. 3D-интегральные когнитивные карты

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 25) вверху и когнитивной диаграммы значений факторов (рисунок 29) внизу и соединяющего их одного слоя нейронной сети (рисунок 33) (режим 4.4.12 системы «Эйдос») (рисунок 34):

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

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

Сей	Код	Наименование нелокального нейрона (класса)
1	CALORIES-1/3-(1.0, 105.3)	
2	CALORIES-2/3-(105.3, 209.7)	
3	CALORIES-3/3-(209.7, 314.0)	

Помощь Максимальное количество отображаемых нейронов: 16 Диапазон кодов отображаемых нейронов: 1 3
 Максимальное количество отображаемых связей: 1000 ClearSet Диапазон кодов отображаемых рецепторов: 1 44

Подготовка визуализации нейрона:1 "CALORIES-1/3-[1.0, 105.3]" в модели:7 "INF4"
АКТИВИРУЮЩИЕ рецепторы и сила их влияния **ТОРМОЗЯЩИЕ рецепторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
24	DURATION-1/7-(1.0000000, 5.1428571)	0.612
25	DURATION-2/7-(5.1428571, 9.2857143)	0.612
26	DURATION-3/7-(9.2857143, 13.4285714)	0.612
32	HEART_RATE-2/7-(75.7142857, 84.4285714)	0.612
38	BODY_TEMP-1/7-(37.1000000, 37.7285714)	0.612
39	BODY_TEMP-2/7-(37.7285714, 38.3571429)	0.612
40	BODY_TEMP-3/7-(38.3571429, 38.9857143)	0.612
41	BODY_TEMP-4/7-(38.9857143, 39.6142857)	0.612
31	HEART_RATE-1/7-(67.0000000, 75.7142857)	0.612
33	HEART_RATE-2/7-(75.7142857, 84.4285714)	0.612

Код	Наименование фактора и его интервального значения	Сила влияния
30	DURATION-7/7-(25.8571429, 30.0000000)	-0.999
44	BODY_TEMP-7/7-(40.8714286, 41.5000000)	-0.970
29	DURATION-6/7-(21.7142857, 25.8571429)	-0.949
35	HEART_RATE-5/7-(101.8571429, 110.5714286)	-0.868
43	BODY_TEMP-6/7-(40.2428571, 40.8714286)	-0.495
28	DURATION-5/7-(17.5714286, 21.7142857)	-0.364
8	AGE-6/7-(62.1428571, 70.5714286)	-0.134
9	AGE-7/7-(70.5714286, 79.0000000)	-0.121
16	HEIGHT-7/7-(207.8571429, 222.0000000)	-0.087
7	AGE-5/7-(54.2857143, 62.1428571)	0.000

ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору

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

Сортировать связи:
 по модулю информативности
 по информативности и знаку

Отображать наименования:
 нейронов
 рецепторов

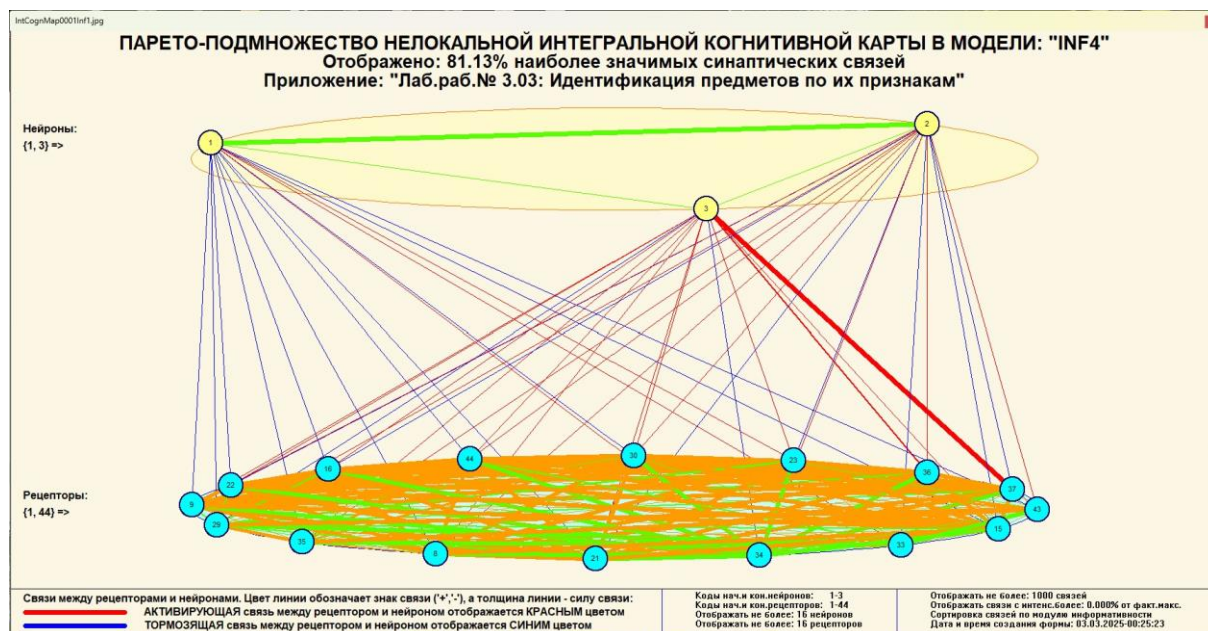


Рисунок 34. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

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

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

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

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

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых может быть одним из первых писал Дьердь Пойа [20]. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году в работе [1] на странице 521¹¹. Позже об этом писалось в работе [2]¹² и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

Пример опосредованных правдоподобных рассуждений.

Допустим нам известно, что один человек имеет голубые глаза, а другой черные волосы. Спрашивается, эти признаки вносят вклад в сходство или в различие этих двух людей? В АСК-анализе и системе

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

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

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

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 35. Всего системой в данной модели генерируется 324 подобных диаграмм, поэтому, естественно, все они не приводятся. Но пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: http://lc.kubagro.ru/aidos/_Aidos-X.htm, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №335 и получить в нем все выходные формы, как описано в данной статье.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

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

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	CALORIES-1/3-(1.0, 105.3)
2	CALORIES-2/3-(105.3, 209.7)
3	CALORIES-3/3-(209.7, 314.0)

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

Выбор способа фильтрации признаков в информационных портретах когнитивной диаграммы

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

Код	Наименование описательной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ ОПИСАТ	1	44
1	GENDER	1	2
2	AGE	3	9
3	HEIGHT	10	16
4	WEIGHT	17	23
5	DURATION	24	30

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

Задать модели, в которых проводить расчеты когнитивных диаграмм:

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

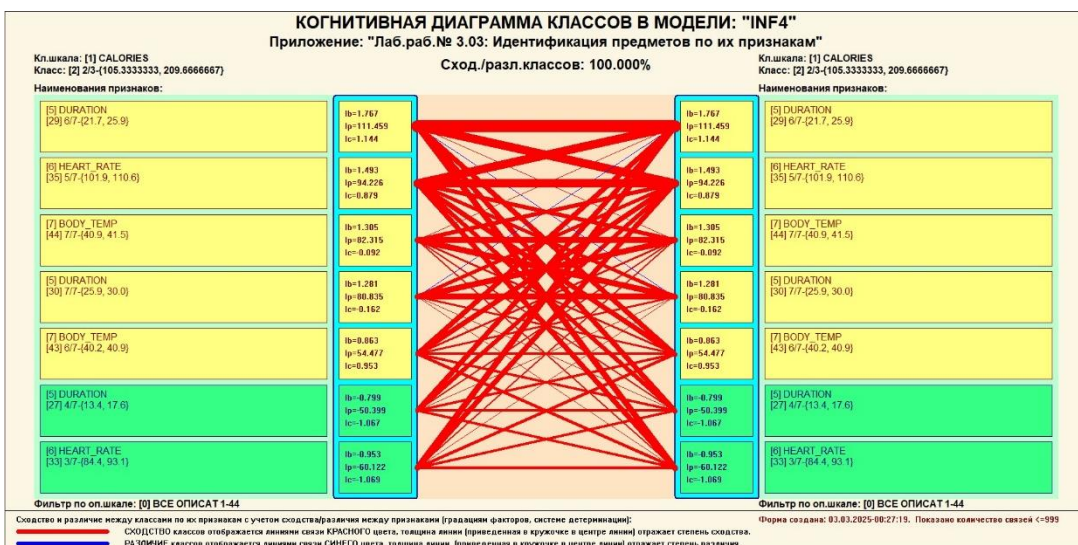
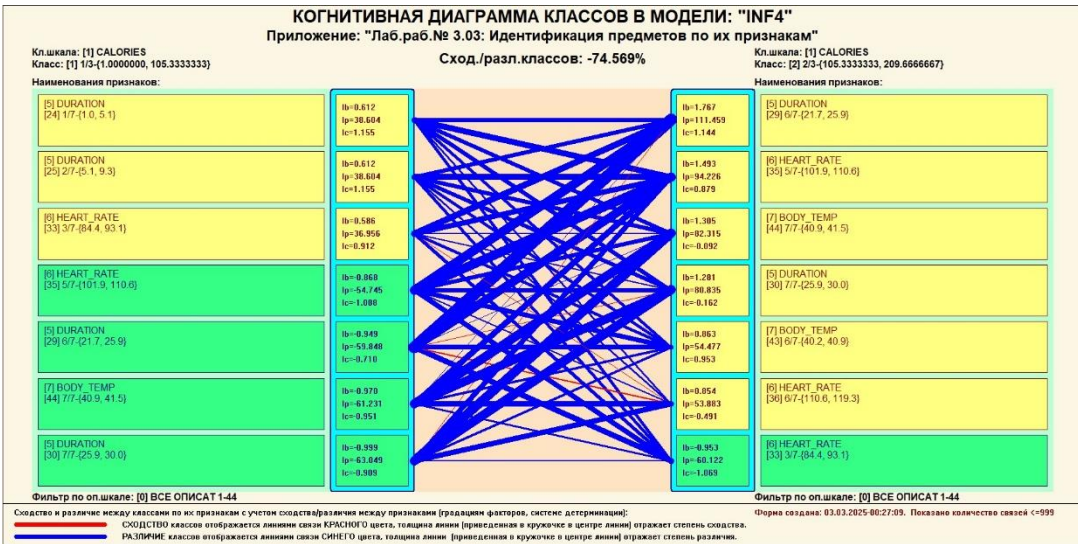
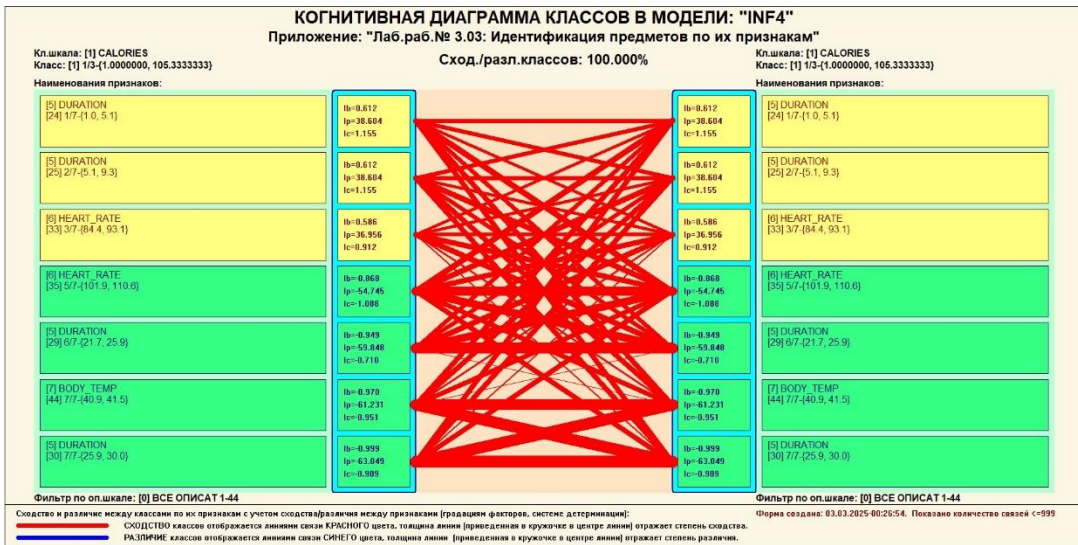
Задать max количество отображаемых связей:

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

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

Задать режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа



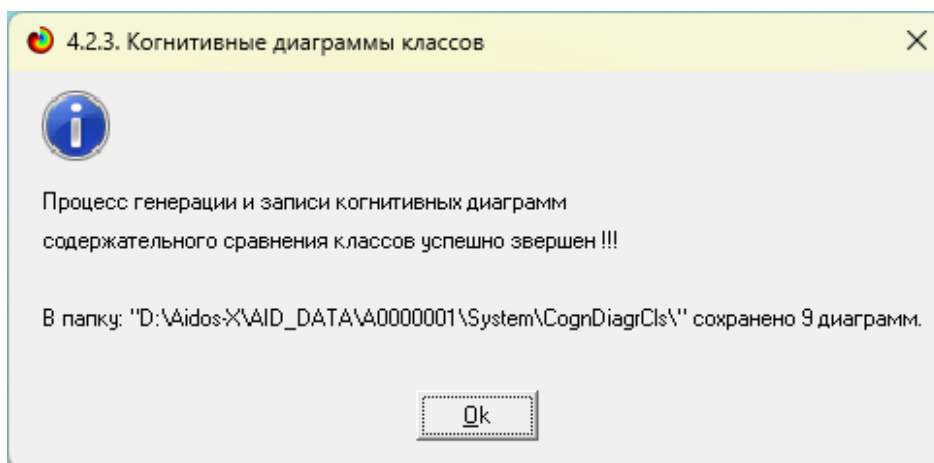
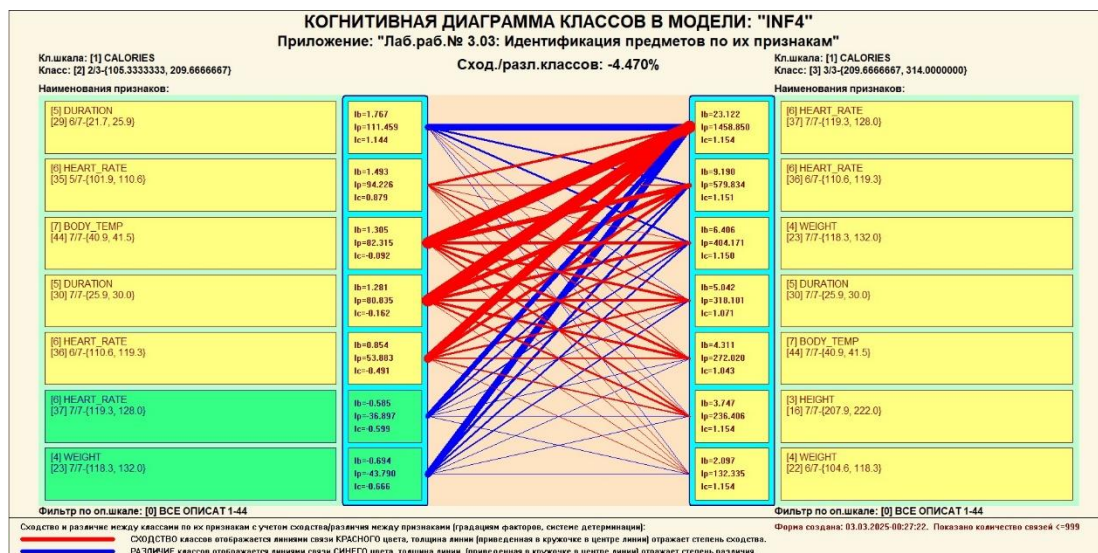


Рисунок 35. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF4

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

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

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

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

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

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта

моделирования в будущие состояния, соответствующие классам, приведены ниже на рисунках 36.

4.3.3. Когнитивные диаграммы признаков. Задание параметров генерации выходных форм

Выбор признаков для когнитивной диаграммы

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

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	GENDER-female
2	GENDER-male
3	AGE-1/7-(20.0000000, 28.4285714)
4	AGE-2/7-(28.4285714, 36.8571429)
5	AGE-3/7-(36.8571429, 45.2857143)
6	AGE-4/7-(45.2857143, 53.7142857)

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

Выбор способа фильтрации классов в информационных портретах когнитивной диаграммы

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

Код	Наименование классификационной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ	1	3
1	CALORIES	1	3

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

Задать модели, в которых проводить расчеты когнитивных диаграмм

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

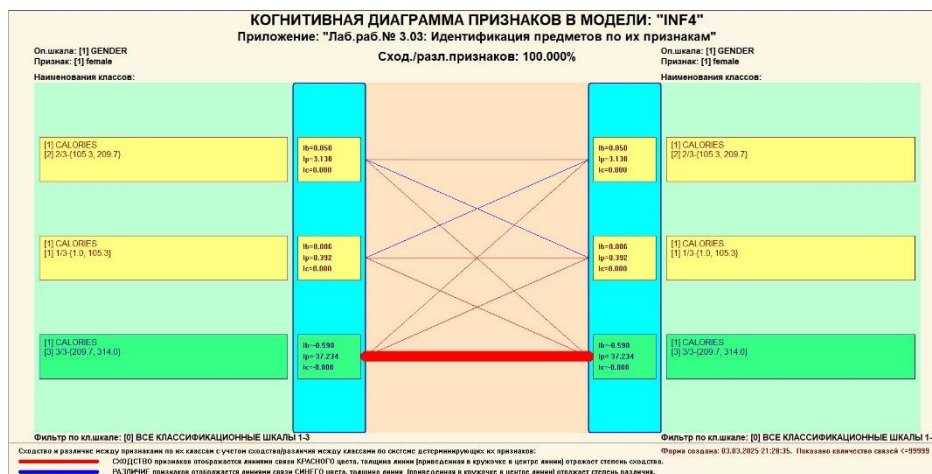
Задать max количество отображаемых связей:

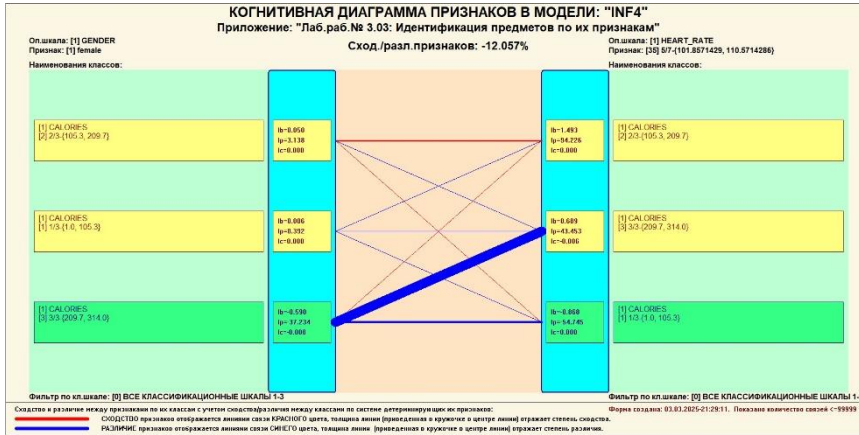
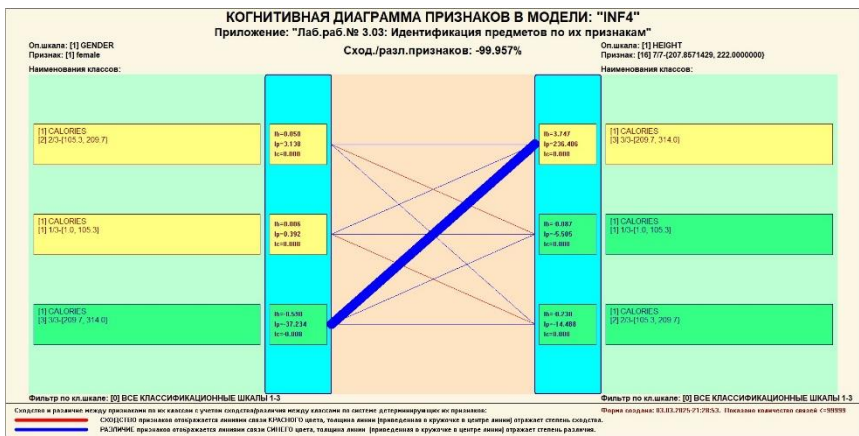
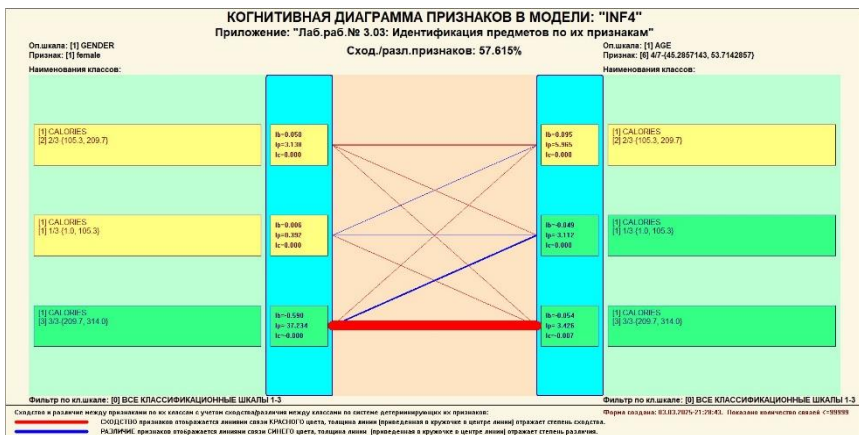
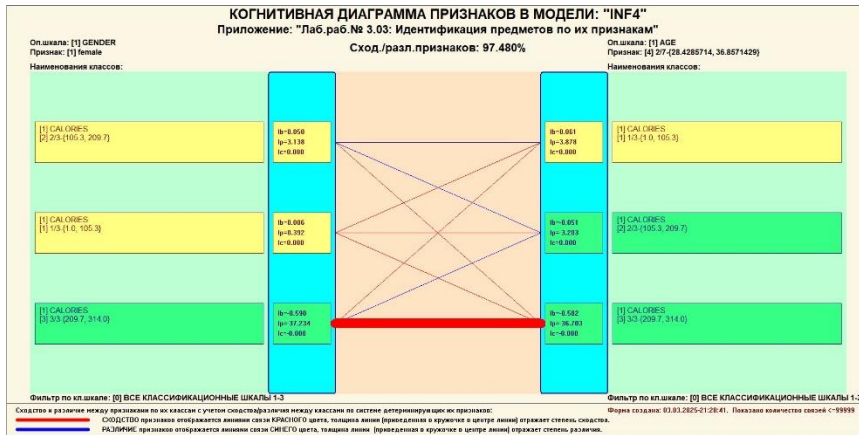
В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Признак для левого инф. портрета: [0] ВСЕ ПРИЗНАКИ
 Признак для правого инф. портрета: [0] ВСЕ ПРИЗНАКИ
 Классиф. шкала для левого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКА...
 Классиф. шкала для правого инф. портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКА...
 Модели, заданные для расчета: Inf4

Задать режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа





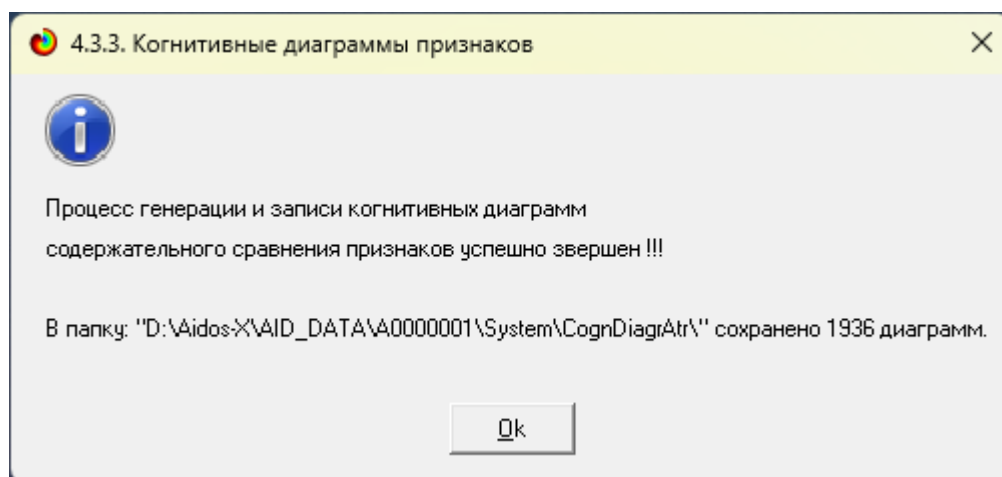


Рисунок 36. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF4

Всего системой в данной модели генерируется 121 подобных диаграмм, поэтому, естественно, все они не приводятся. Но пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №335 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

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

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

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

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 37). Первая экранная форма данного режима представляет собой краткий хелп, поясняющий смысл понятия «Когнитивная функция»,

а также позволяющий выйти на экранную форму системы «Эйдос» с действующими гиперссылками на работы по когнитивным функциям, а также страницы сайта автора со списком этих работ и работ автора по выявлению, представлению и использованию знаний, логике и методологии научного познания.

Необходимо отметить, что модели системы «Эйдос» – это **феноменологические** модели, отражающие **эмпирические** закономерности в фактах обучающей выборки, т.е. они отражают причинно-следственные связи, но не отражают **механизма детерминации**, а только сам факт и характер детерминации [18]. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [19].

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

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

Статистические базы:

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

Системно-когнитивные модели (Базы знаний):

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

Задайте виды когнитивных функций для генерации, визуализации и записи:

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

Задайте дополнительные параметры визуализации когнитивных функций:

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

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

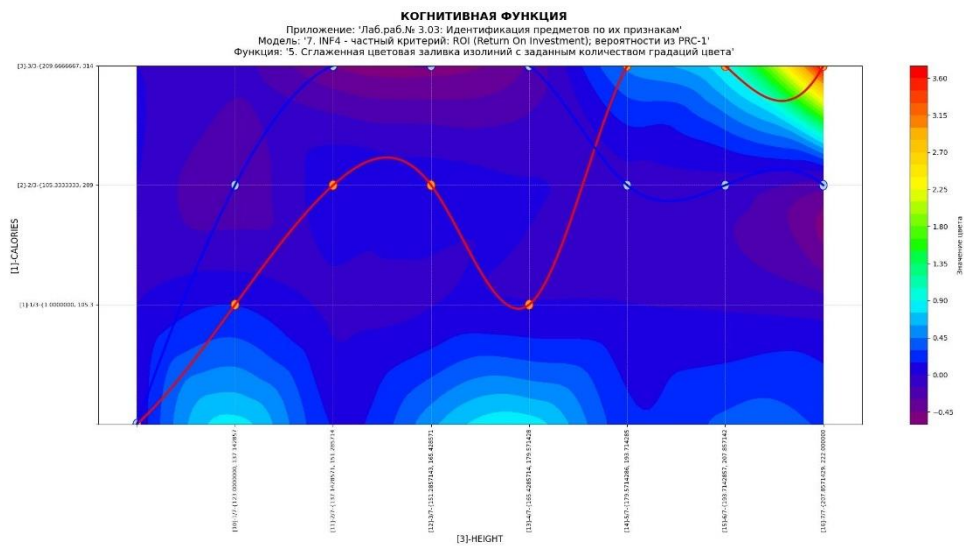
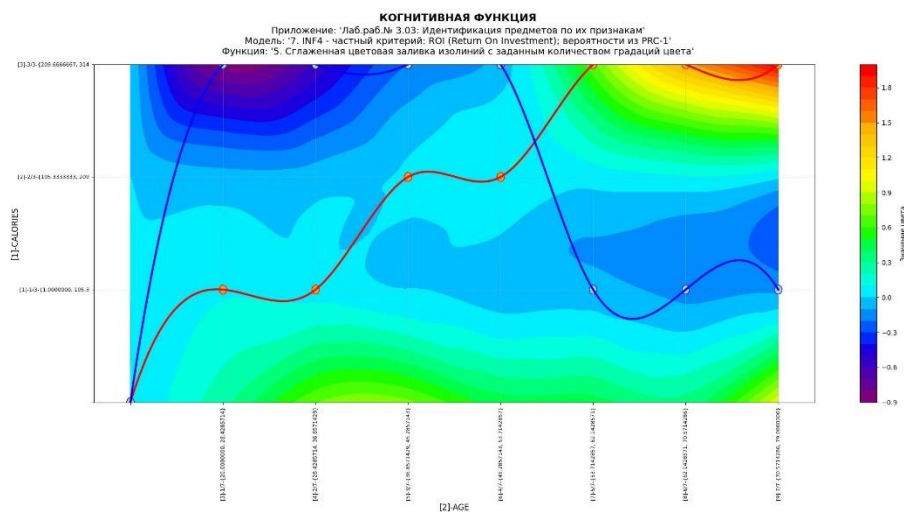
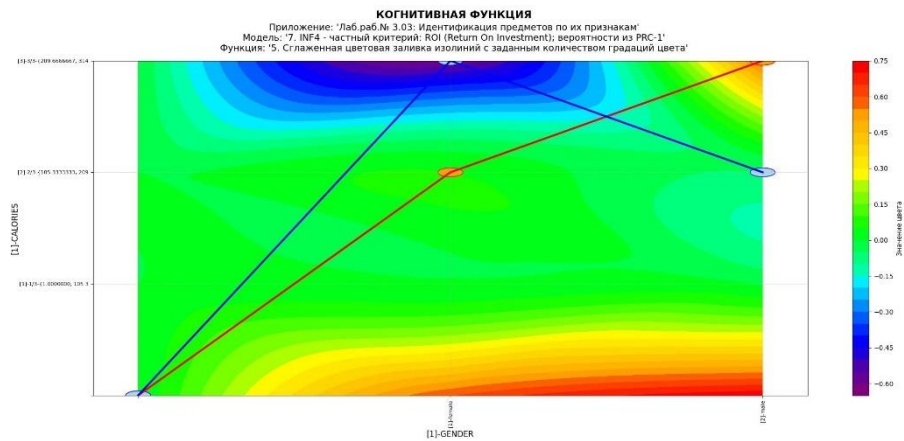
Задайте количество пикселей на дюйм в изображениях когнитивных функций:

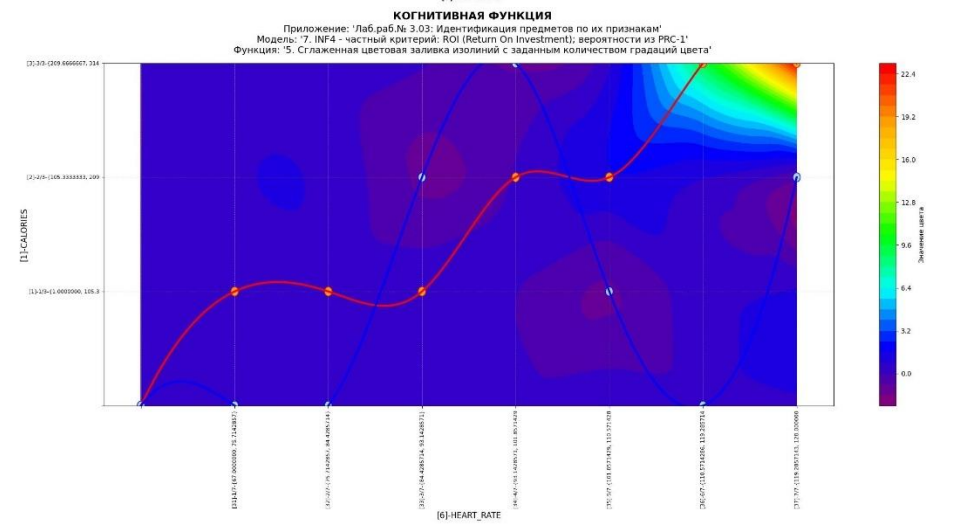
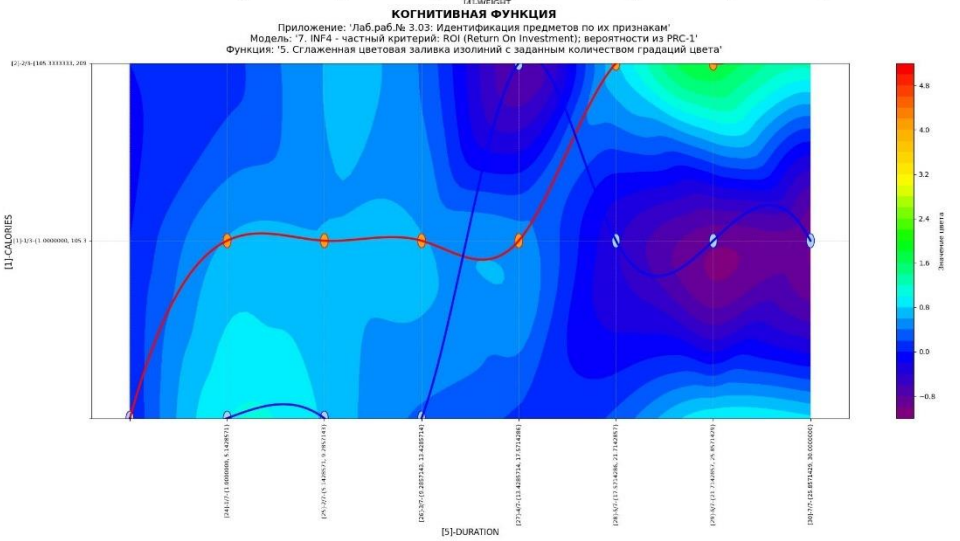
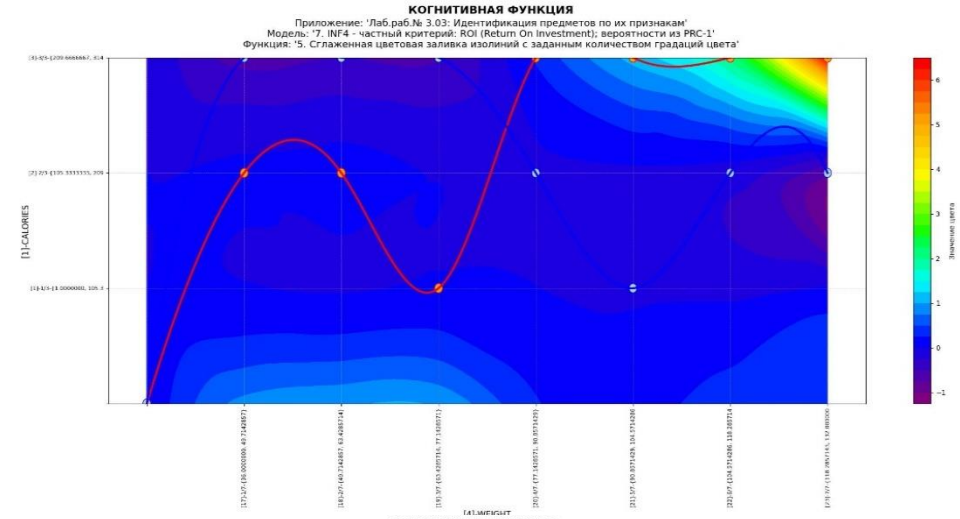
Задайте паузу в секундах между визуализациями когнитивных функций:

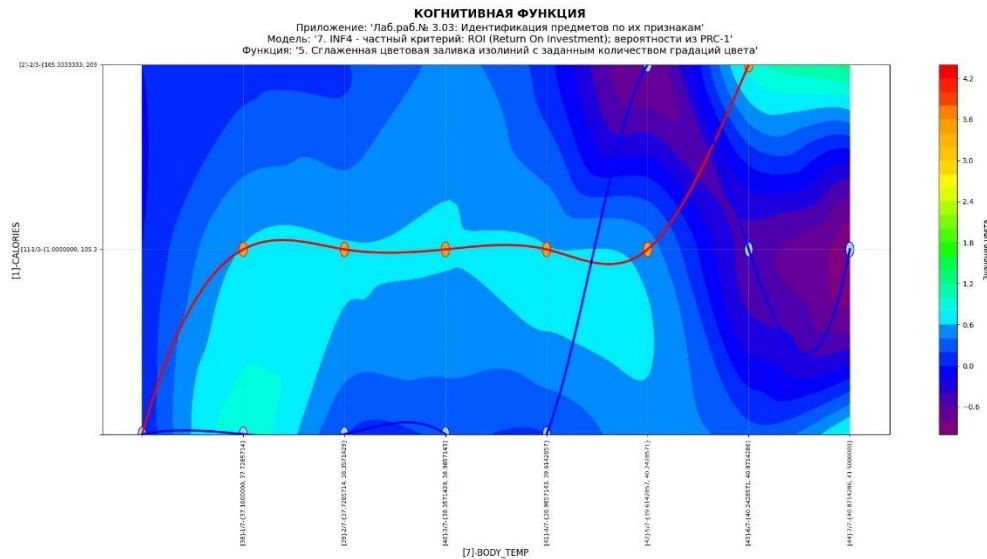
Задайте размер шрифта для наименований градаций шкал X и Y:

Визуализация когнитивных функций new Визуализация когнитивных функций old

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







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

Стадии исполнения процесса

Записана когнитивная функция: "5. Сглаженная цветовая заливка изолиний с заданн"...в модели: "7."

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

Прогноз времени исполнения

Начало: 22:02:15 **Окончание: 22:04:04**

100.00000%

Прошло: 0:01:49 **Осталось: 0:00:00**

Ok

4.5. Визуализация когнитивных функций

Что такое когнитивная функция:

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

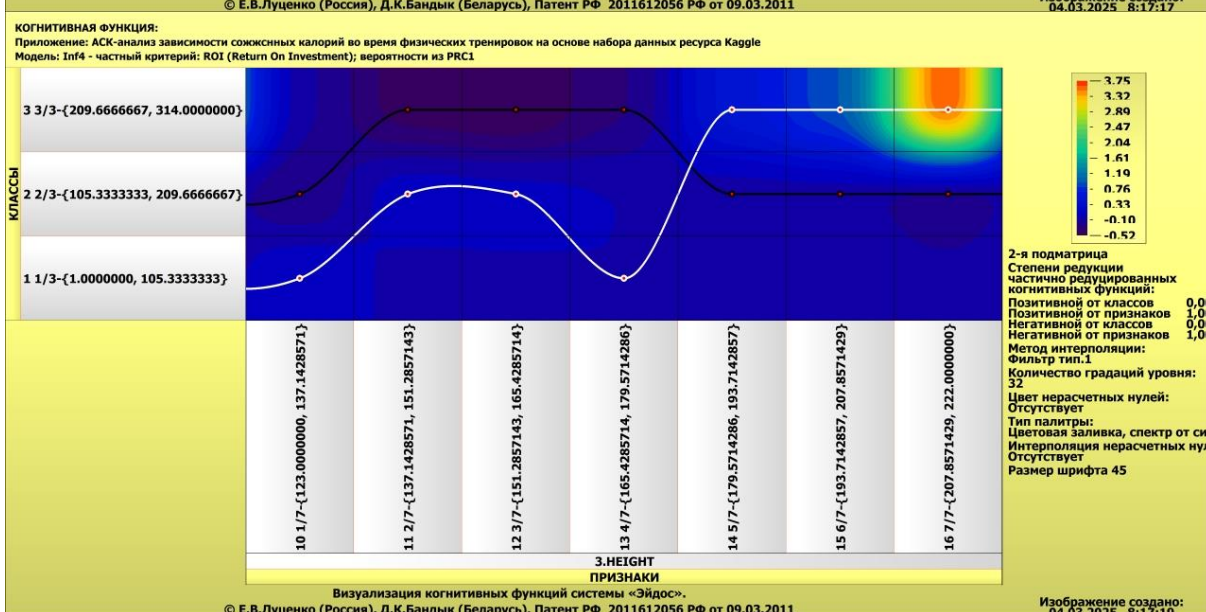
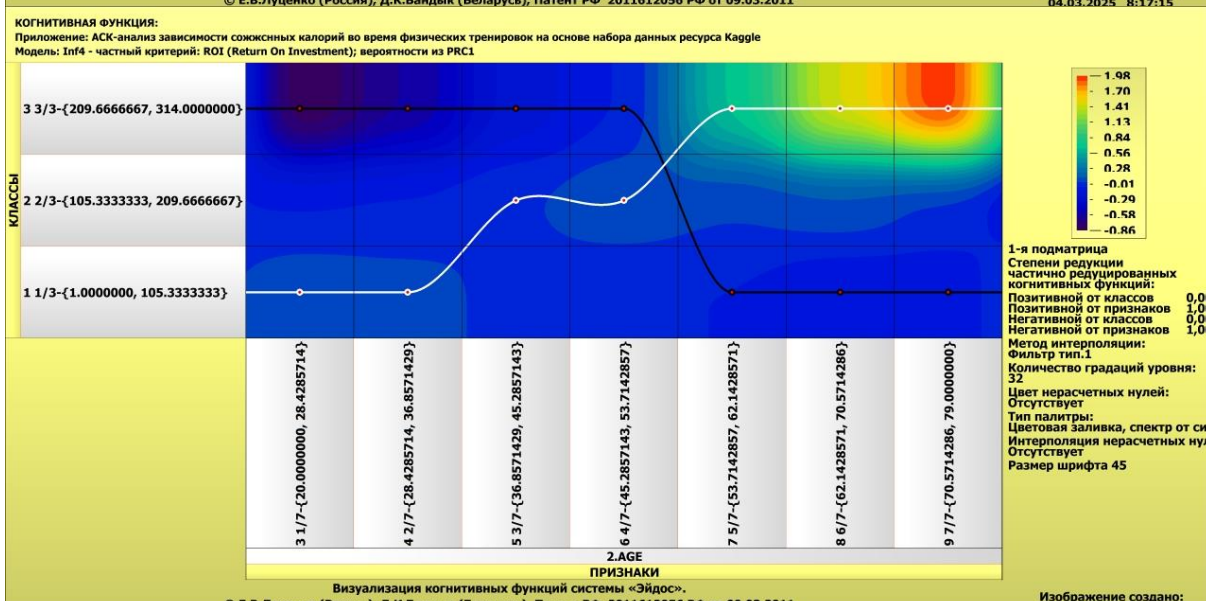
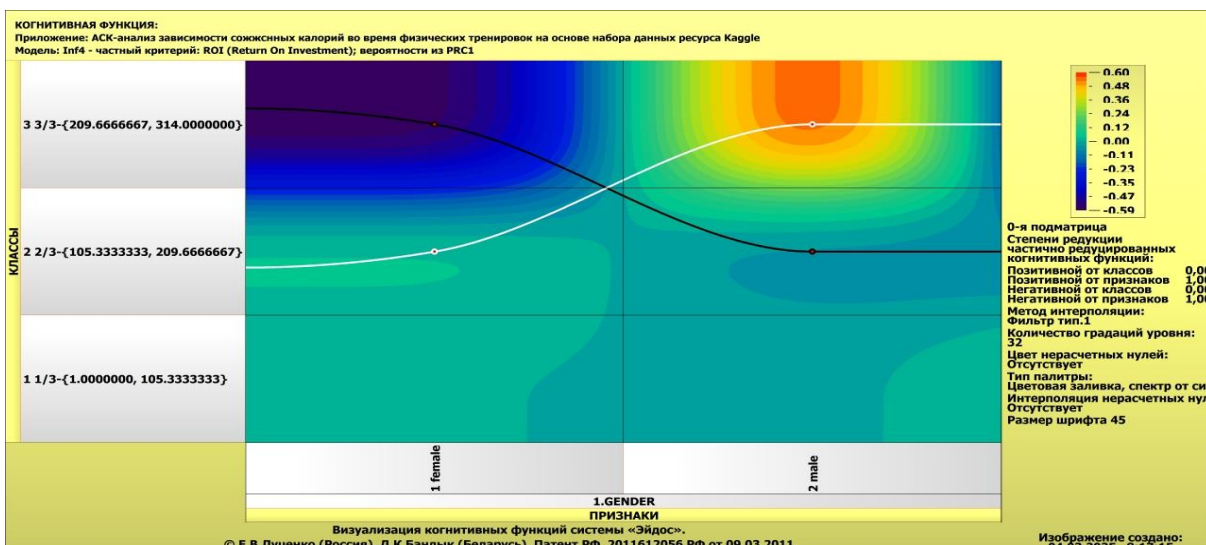
Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам. Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации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

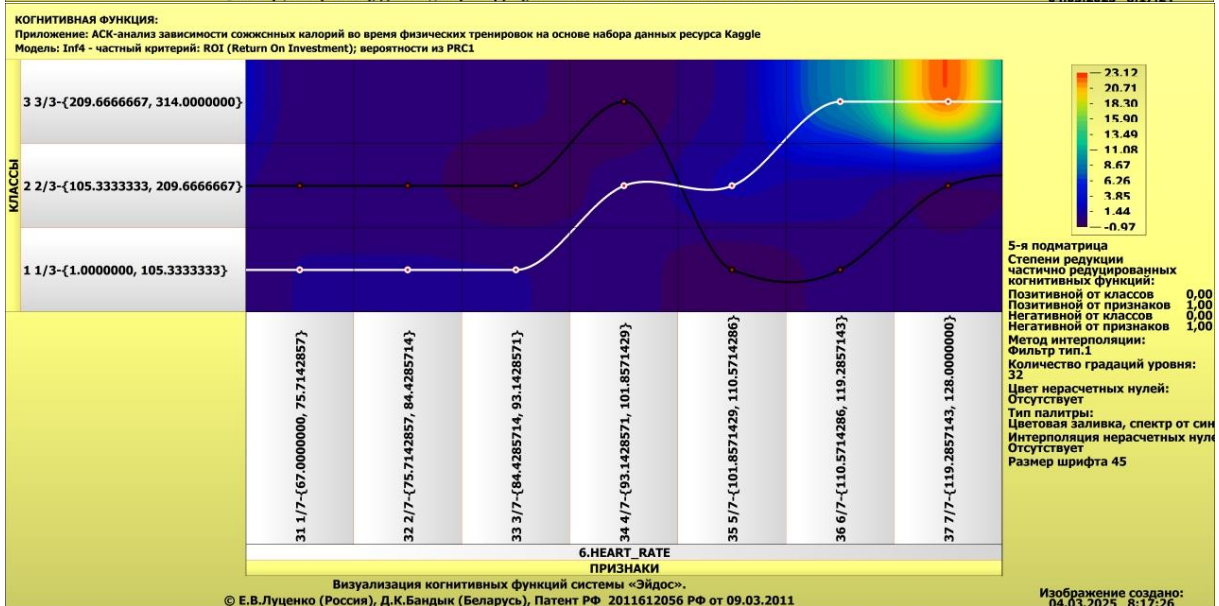
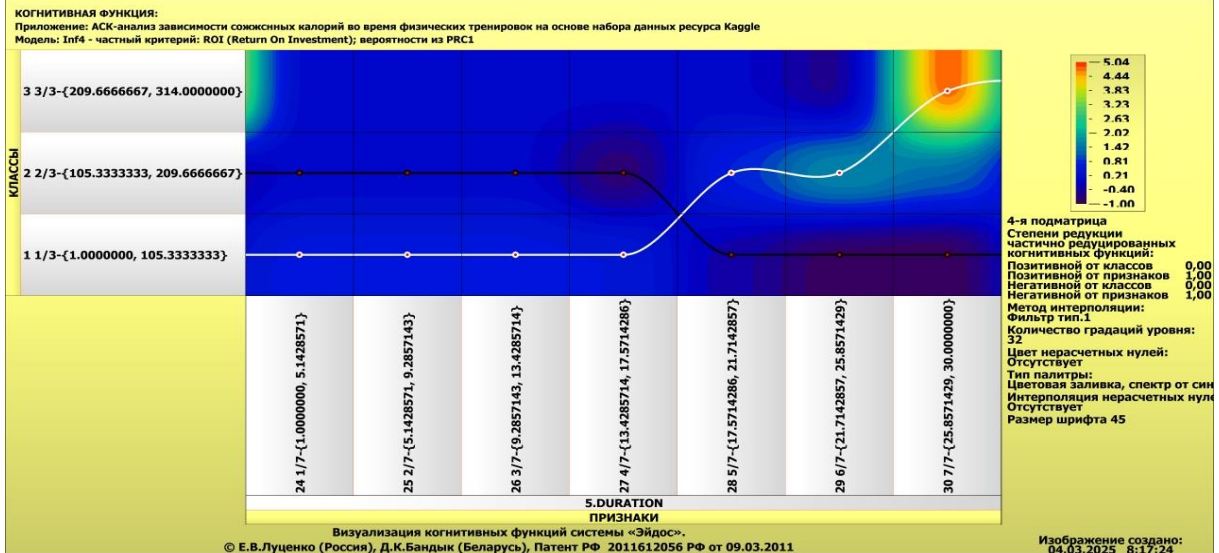
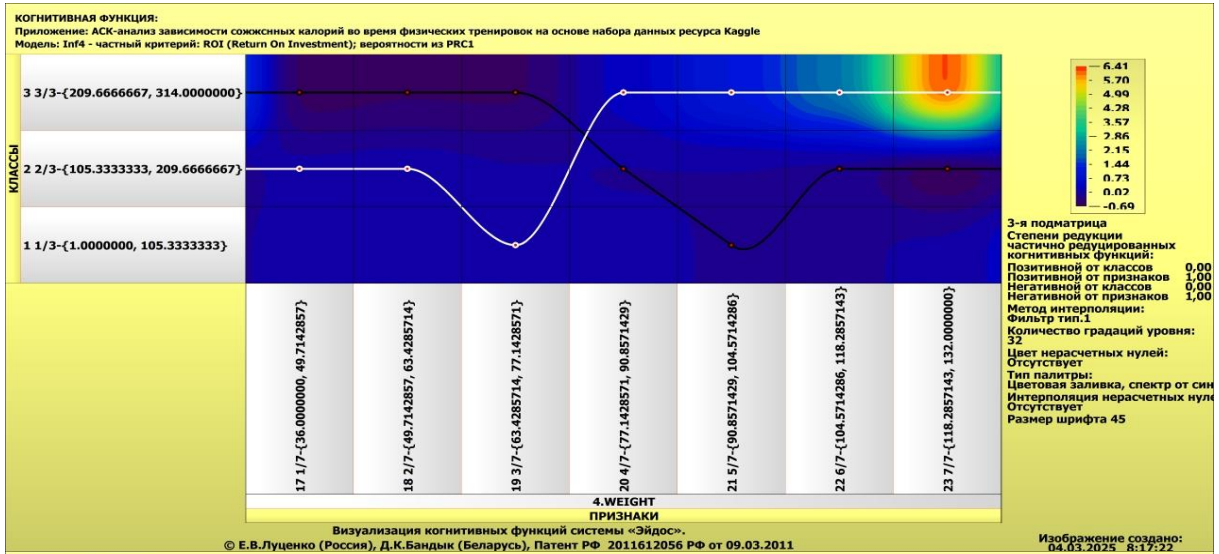
Луценко Е.В. Метод визуализации когнитивных функций - новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Полиматематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный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

Задайте нужный режим:

Визуализации когнитивных функций Литератур.ссылки на работы по когнитивным функциям

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





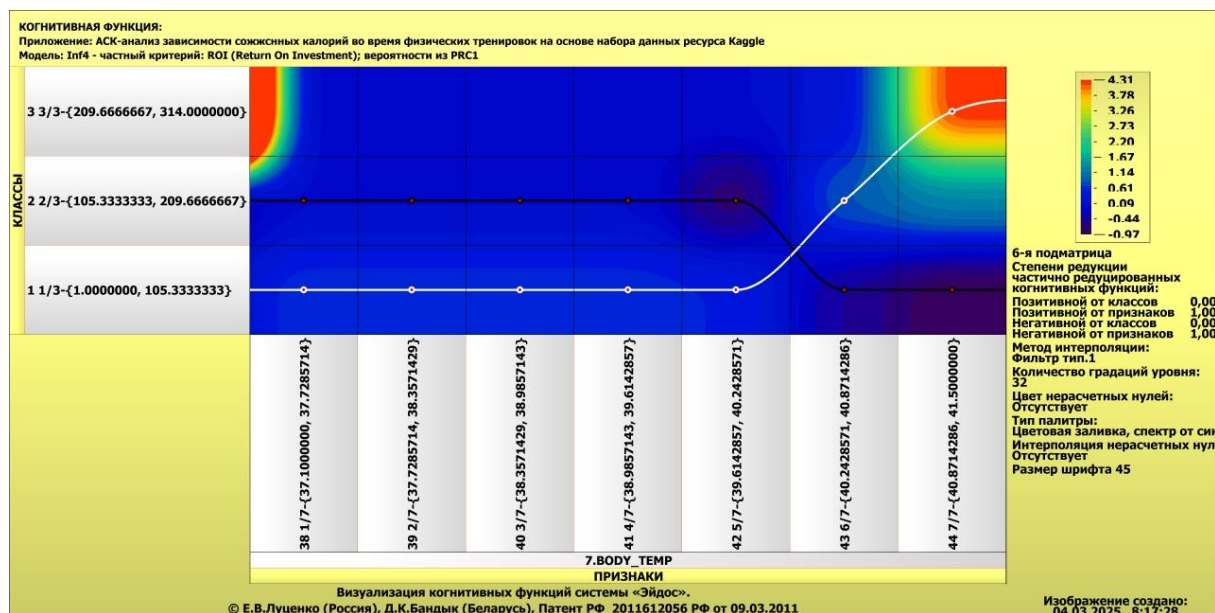


Рисунок 37. Примеры когнитивных функций в СК-модели INF4

Как уже отмечалось содержательное объяснение когнитивных функций на теоретическом уровне познания – это дело специалистов в той предметной области, к которой относится предмет моделирования [19].

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 [5].

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

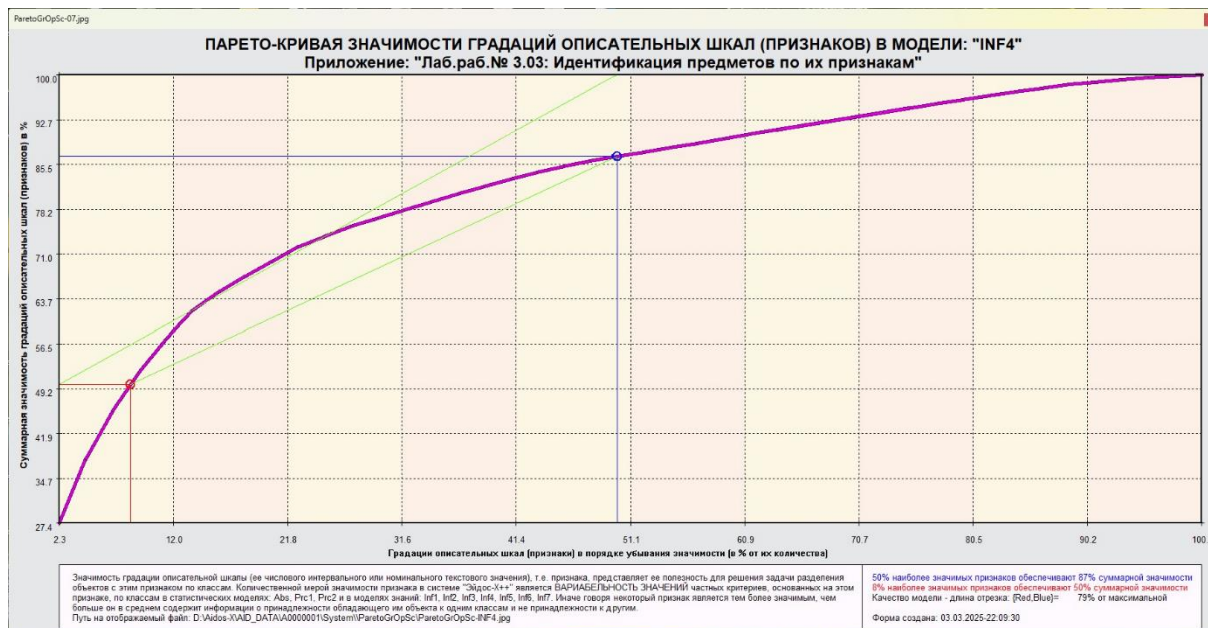


Рисунок 38. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF4

Из рисунка 38 видно, что примерно пятая часть наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 78,5% суммарного влияния.

В таблице 13 представлены исходные данные для построения кумулятивной кривой на рисунке 38. Из таблицы 16 видно, какую долю от суммарного влияния на переход объекта моделирования в будущее состояния, соответствующие классам, имеет каждое значение каждого фактора. Например, видно, что значение фактора: HEART_RATE-7/7-{119.2857143, 128.0} имеет влияние на объект моделирования почти в 161 раз выше, чем значение фактора: AGE-4/7-{45.2857143, 53.7142857}. Это значит, что сила влияния разных значений факторов на объект моделирования довольно существенно отличается друг от друга.

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

№ п./п.	№,%	Код значения фактора	Наименования фактора и его значения	Код фактора	Значимость, %	Значимость кумулятивно, %
1	2,2727273	37	HEART_RATE-7/7- {119.2857143, 128.0000000}	6	27,4215324	27,4215324
2	4,5454545	36	HEART_RATE-6/7- {110.5714286, 119.2857143}	6	10,2968356	37,7183680
3	6,8181818	23	WEIGHT-7/7- {118.2857143, 132.0000000}	4	7,9291641	45,6475321
4	9,0909091	30	DURATION-7/7- {25.8571429, 30.0000000}	5	6,1865754	51,8341075
5	11,3636364	44	BODY_TEMP-7/7- {40.8714286, 41.5000000}	7	5,3728971	57,2070047
1	2,2727273	37	HEART_RATE-7/7- {119.2857143, 128.0000000}	6	27,4215324	27,4215324
42	95,4545455	13	HEIGHT-4/7- {165.4285714, 179.5714286}	3	0,3328858	99,5873379
43	97,7272727	5	AGE-3/7- {36.8571429, 45.2857143}	2	0,2412361	99,8285740
44	100,0000000	6	AGE-4/7- {45.2857143, 53.7142857}	2	0,1714260	100,0000000

На экранной форме рисунка 39 приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях:

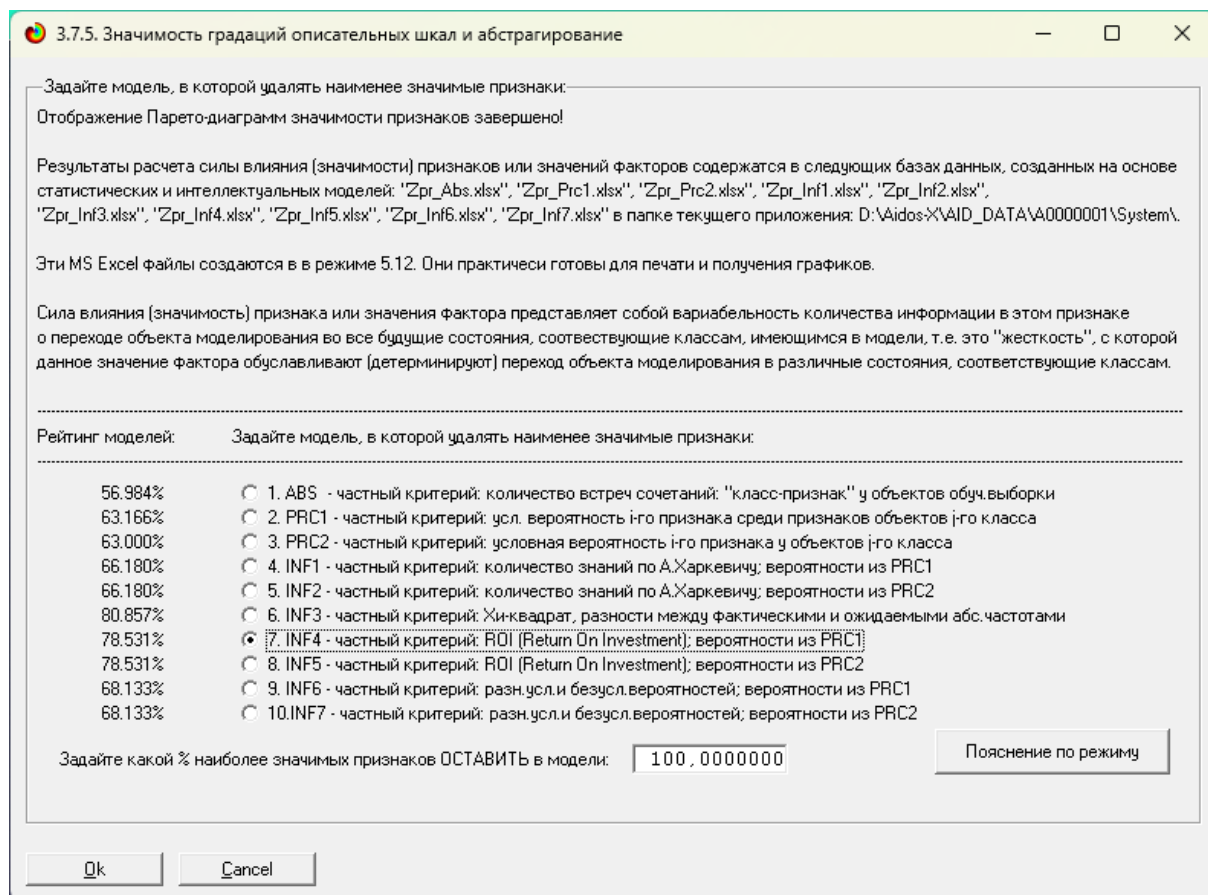


Рисунок 398. Имена Excel-файлов с информацией о силе влияния значений факторов в разных моделях

На экранной форме рисунка 40 приведены имена Excel-файлов с информацией о силе влияния факторов в разных моделях.

В таблице 14 приведена информация о силе влияния факторов на переход объекта моделирования в различные будущие состояния, соответствующие классам, в системно-когнитивной модели INF4. Из таблицы 14 видно, что почти 46% суммарного влияния на поведение объекта моделирования обусловлено частотой сердцебиения, еще 14% влияния оказывает длительность тренировки, а пол человека оказывают сравнительно меньшее влияние: около 5%.

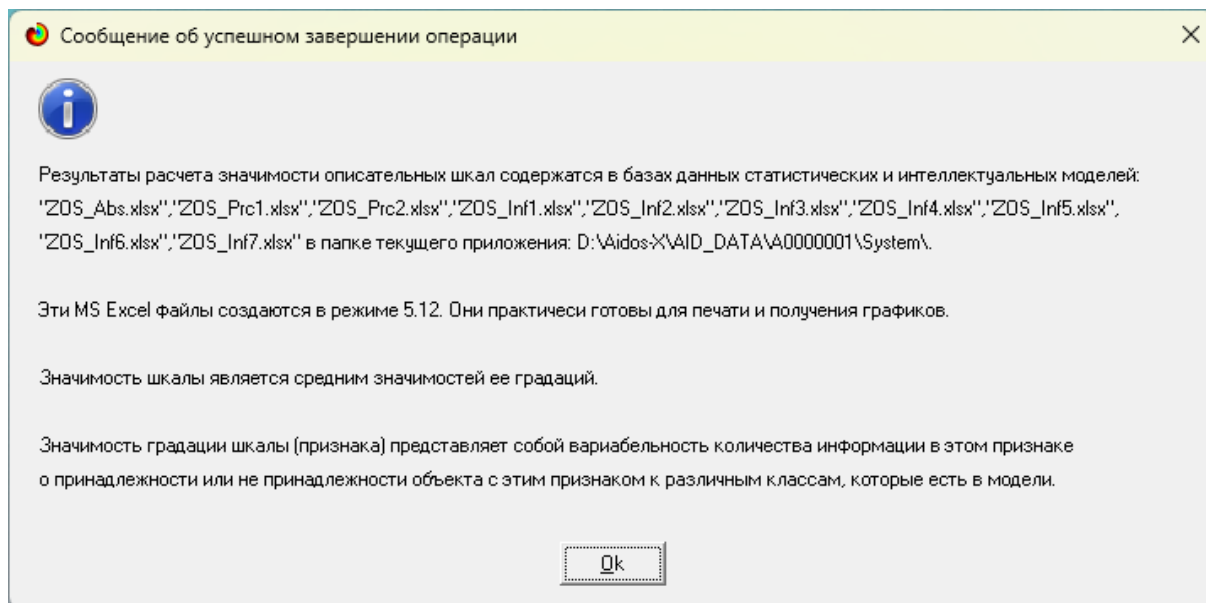


Рисунок 40. имена Excel-файлов с информацией о силе влияния факторов в разных моделях

Таблица 14 – Сила влияния факторов на поведение объекта моделирования в системно-когнитивной модели INF4

№	№,%	Код фактора	Наименование фактора	Сила влияния фактора, %	Сила влияния фактора кумулятивно, %
1	14,2857143	6	HEART_RAT E	42,7784014	42,7784014
2	28,5714286	5	DURATION	13,6532958	56,4316971
3	42,8571429	7	BODY_TEMP	13,3106541	69,7423512
4	57,1428571	4	WEIGHT	10,5350086	80,2773598
5	71,4285714	2	AGE	7,9714427	88,2488025
6	85,7142857	3	HEIGHT	6,8196527	95,0684552
7	100,0000000	1	GENDER	4,9315448	100,0000000

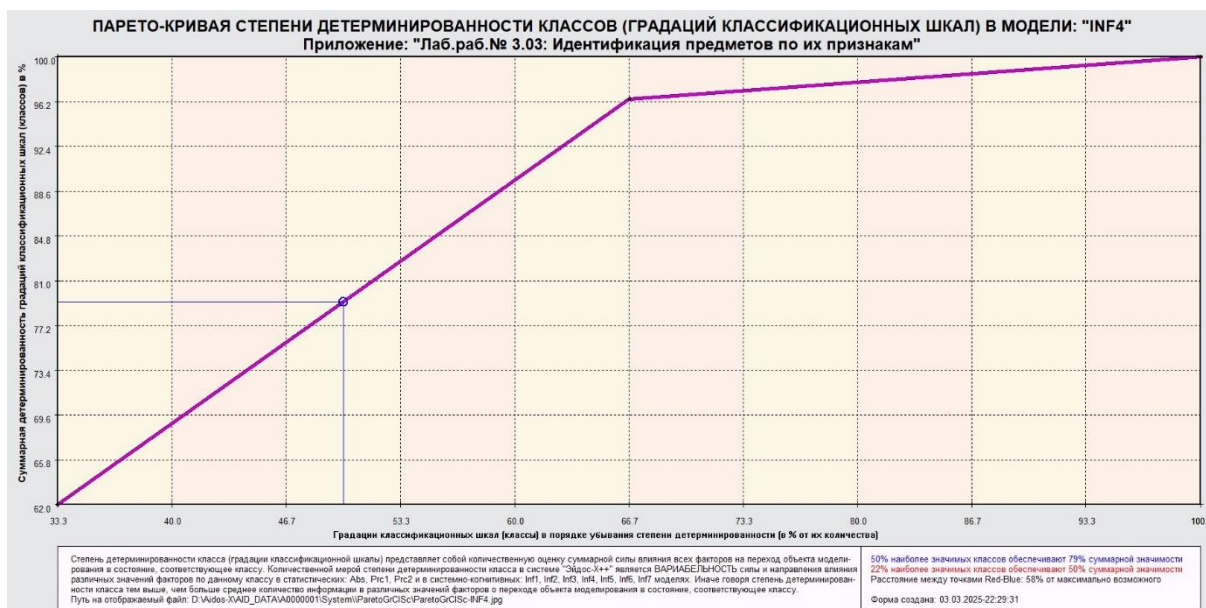
3.8.11. Степень детерминированности классов и классификационных шкал

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

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

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

На рисунках 41 приведены экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



Сообщение об успешном завершении операции

Отображение Парето-диаграмм степени сформированности классов завершено!

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

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

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

Ok

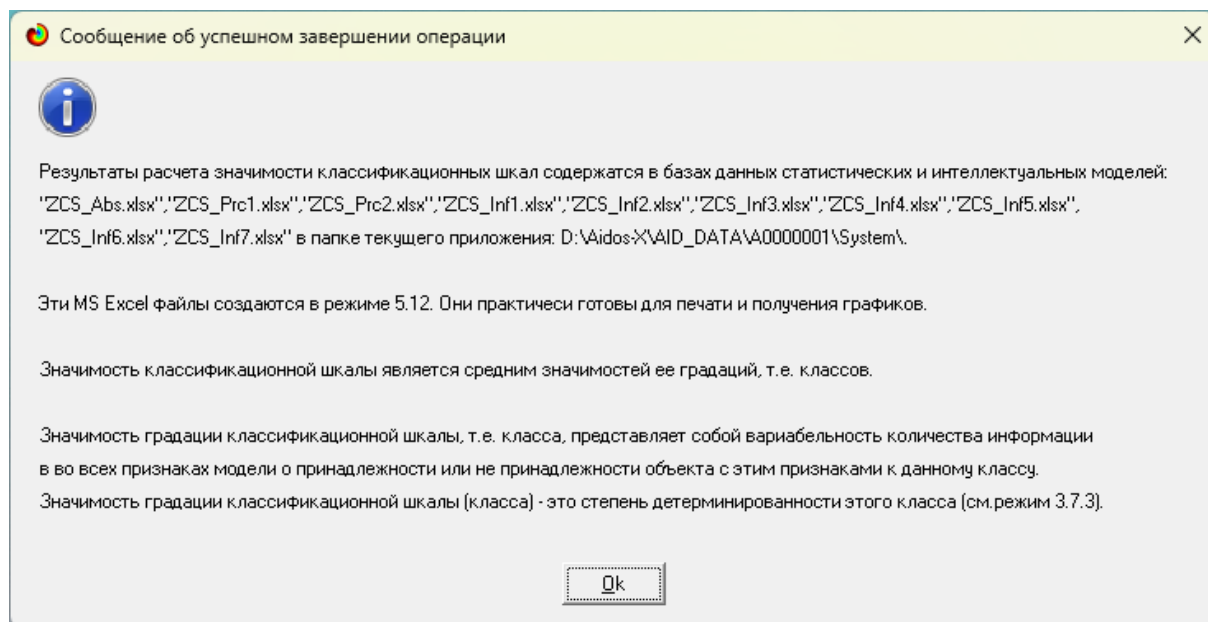


Рисунок 41. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

В таблице 15 представлены исходные данные для построения кумулятивной кривой на рисунке 41. Из таблицы 15 видно, какую долю от суммарной степени детерминированности всех классов имеет каждый класс. Например, видно, что класс: CALORIES-1/3-{1.0, 105.3} 17 раз превышает класс CALORIES-3/3-{209.7, 314.0}. Это значит, что степень обусловленности значениями факторов разных будущих состояний объекта моделирования, соответствующие классам, довольно существенно отличается друг от друга.

Таблица 15 – Степень детерминированности классов в СК-модели INF4

№	№, %	Код класса	Наименование классификационной шкалы и класса	Код классификационной шкалы	Степень детерминированности класс, %	Степень детерминированности класса, кумулятивно, %
1	33,3333333	1	CALORIES-1/3-{1.0, 105.3}	1	62,0400000	62,0400000
2	66,6666667	2	CALORIES-2/3-{105.3, 209.7}	1	34,4066667	96,4466667
3	100,0000000	3	CALORIES-3/3-{209.7, 314.0}	1	3,5533333	100,0000000

По таблице 16, нет возможности продемонстрировать степень детерминированности классификационных шкал, так как так как изначально факторы соответствовали одному классом.

Таблица 16 – Степень детерминированности классификационных шкал в системно-когнитивной модели INF4

№	№, %	Код класса	Наименование классификационной шкалы и класса	Степень детерминированности класс, %	Степень детерминированности класса, кумулятивно, %
1	100,0000000	1	CALORIES	100,0000000	100,0000000

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 лингвистического Автоматизированного системно-когнитивного анализа (лингвистический 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 построения системно-когнитивных моделей предметной области на основе исходных данных, содержащих лингвистические переменные.

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.

Рекомендуется ввести описательные шкалы, не только продолжительность тренировок, но и **о типе тренировки**, разные виды физической активности (бег, плавание, силовые тренировки) могут по-разному влиять на количество сожжённых калорий. Если тип тренировки не указан, модель может быть менее точной. Так следует добавить **данные о питании и уровне метаболизма**: Количество сожжённых калорий может зависеть от диеты и базового уровня метаболизма, которые не учтены в данных.

Перспективность и ценность результатов подобных исследований и разработок для теории и практики не вызывает особых сомнений, так как данные исследования помогут людям при оценке сжигаемых калорий во время занятием физическими нагрузками.

5. CONCLUSIONS (ВЫВОДЫ)

В рамках проведенного исследования был выполнен автоматизированный системно-когнитивный анализ процесса сжигания калорий в организме человека. Основной целью работы стало выявление ключевых факторов, которые влияют на количество сжигаемых калорий во время физической активности. Для этого были рассмотрены различные параметры продолжительность тренировки, а также индивидуальные характеристики человека, включая вес, возраст, пол и уровень физической подготовки. Это позволило определить, как каждый из этих факторов влияет на объем сжигаемых калорий и общую эффективность тренировок.

Цель исследования была успешно достигнута: проведен детальный анализ процесса сжигания калорий, выявлены ключевые характеристики, которые оказывают наибольшее влияние на их количество. В процессе работы были последовательно решены следующие задачи:

1. **Когнитивная структуризация предметной области** – на этом этапе была проведена систематизация знаний о процессе сжигания калорий, выделены основные понятия и взаимосвязи между ними.

2. **Формализация предметной области** – разработана математическая и логическая модель, которая позволила описать процесс сжигания калорий в виде структурированных данных.

3. **Синтез статистических и системно-когнитивных моделей** – на основе собранных данных были созданы модели, позволяющие анализировать и прогнозировать количество сжигаемых калорий в зависимости от различных факторов.

4. **Верификация моделей** – проведена проверка точности и достоверности разработанных моделей на реальных данных.

5. **Выбор наиболее достоверной модели** – из нескольких вариантов была выбрана модель, которая наиболее точно отражает процесс сжигания калорий.

6. **Системная идентификация и прогнозирование** – на основе выбранной модели были проведены расчеты и прогнозы, позволяющие определить, как изменение тех или иных параметров влияет на количество сжигаемых калорий.

7. **Поддержка принятия решений** – разработаны рекомендации, которые могут быть использованы для оптимизации тренировочного процесса и повышения его эффективности.

8. **Анализ объекта моделирования** – проведено исследование модели, что позволило глубже понять механизмы сжигания калорий и их зависимость от различных факторов.

Исследование было направлено на анализ факторов, определяющих количество сжигаемых калорий, что имеет важное практическое значение для оптимизации тренировочного процесса. Полученные результаты позволяют не только лучше понять, как различные параметры влияют на эффективность физической активности, но и разработать индивидуальные программы тренировок, которые будут максимально эффективны для каждого конкретного человека. Например, зная, как интенсивность и продолжительность тренировки влияют на сжигание калорий, можно подобрать оптимальный режим занятий для достижения конкретных целей, будь то похудение, поддержание формы или подготовка к спортивным соревнованиям.

Кроме того, результаты исследования могут быть использованы для дальнейших научных изысканий в области физической активности, спортивной медицины и фитнеса. Они также могут быть полезны для разработки новых методик тренировок, создания специализированных приложений и устройств, которые помогут людям более точно отслеживать свои достижения и прогресс. Например, на основе полученных данных можно разработать алгоритмы для фитнес-трекеров, которые будут учитывать индивидуальные особенности пользователя и предоставлять персонализированные рекомендации по тренировкам.

Таким образом, проведенное исследование вносит значительный вклад в понимание механизмов сжигания калорий и их зависимости от различных факторов. Это открывает новые возможности для оптимизации тренировочного процесса, повышения его эффективности и улучшения качества жизни людей, занимающихся спортом или ведущих активный образ жизни.

REFERENCES (ЛИТЕРАТУРА)

1. 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 OCZFHС.

2. 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.

3. Работы проф.Е.В.Луценко по АСК-анализу текстов, т.е. по когнитивной математической лингвистике: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_texts.htm

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. 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.

6. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.

7. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>

8. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.

9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.

10. Луценко, Е. В. 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.

11. 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.

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

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

14. Луценко, Е. В. Системное обобщение принципа Эшби и повышение уровня системности модели объекта познания как необходимое условие адекватности процесса его познания / Е. В. Луценко // Политематический сетевой электронный научный

журнал Кубанского государственного аграрного университета. – 2020. – № 163. – С. 100-134. – DOI 10.21515/1990-4665-163-009. – EDN SWKGWY.

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

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

17. Луценко, Е. В. 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.

18. 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.

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.