

УДК 004.8

Автоматизированный системно-когнитивный анализ и прогнозирования сердечных заболеваний

Малая Полина Павловна
студентка факультета ПИ, группы ИТ1821
polina-malaya-00@bk.ru

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

Целью данной работы прогнозирование сердечных заболеваний. Достижение данной цели представляет большой личный интерес. Это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», что в значительной степени улучшит понимание концепции искусственного интеллекта. Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

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

Automated system-cognitive analysis of heart disease prediction

Malaya Polina Pavlovna
student of the faculty of PI, group IT1821
polina-malaya-00@bk.ru

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

The purpose of this work is to predict heart disease. The achievement of this goal is of great personal interest. This will allow to gain knowledge in working with the universal analytical system "Aidos-X ++", which will greatly improve the understanding of the concept of artificial intelligence. To achieve this goal, the Automated System-Cognitive Analysis (ASK-analysis) and its software tools are used - the intelligent system "Aidos".

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

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	1
ЗАДАЧА 1: КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	5
ЗАДАЧА 2: ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ И ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ.....	6
ЗАДАЧА 3: СИНТЕЗ И ВЕРИФИКАЦИЯ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ И ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ ИЗ НИХ ДЛЯ РЕШЕНИЯ ЗАДАЧ	10
ЗАДАЧА 4: РЕШЕНИЕ РАЗЛИЧНЫХ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	16
Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)	16
Подзадача 4.2. Поддержка принятия решений (SWOT-анализ).....	17
Подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели	21
4.3.1. Когнитивные диаграммы значений факторов	21
4.3.2. Агломеративная когнитивная кластеризация значений факторов	22
4.3.3. Нелокальные нейроны и нелокальные нейронные сети.....	23
4.3.4. 3d-интегральные когнитивные карты.....	24
4.3.5. Когнитивные функции	25
4.3.6. Сила и направление влияния значений признаков заболеваний сердца и сила влияния этих признаков на прогнозирование смерти	28
7. ВЫВОДЫ.....	31
СПИСОК ЛИТЕРАТУРЫ.....	32

Введение

Целью данной работы является предсказание сердечных заболеваний.

Достижение данной цели представляет большой личный интерес. Это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», что в значительной степени улучшит понимание концепции искусственного интеллекта.

АСК-анализ предполагает, что для достижения поставленной цели необходимо решить следующие задачи, которые получаются путем декомпозиции цели и являются этапами ее достижения:

Задача 1: когнитивная структуризация предметной области.

Задача 2: подготовка исходных данных и формализация предметной области.

Задача 3: синтез и верификация статистических и системно-когнитивных моделей и выбор наиболее достоверной модели.

Задача 4: решение различных задач в наиболее достоверной модели:

- подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация);
- подзадача 4.2. Поддержка принятия решений;
- подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели (когнитивные диаграммы классов и значений факторов, агломеративная когнитивная кластеризация классов и значений факторов, нелокальные нейроны и нейронные сети, 3d-интегральные когнитивные карты, когнитивные функции).

Эти задачи, по сути, представляют собой **этапы** автоматизированного системно-когнитивного анализа (АСК-анализа), который поэтому и предлагается применить для их решения.

АСК-анализ представляет собой метод искусственного интеллекта, разработанный проф. Е.В. Луценко в 2002 году для решения широкого класса задач идентификации, прогнозирования, классификации, диагностики, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели. АСК-анализ доведен до **инновационного** уровня благодаря тому, что имеет свой программный инструментарий – универсальную когнитивную аналитическую систему «Эйдос-X++» (система «Эйдос»).

Система «Эйдос» выгодно отличается от других интеллектуальных систем следующими параметрами:

- разработана в универсальной постановке, не зависящей от предметной области. Поэтому она является универсальной и может быть применена во многих предметных областях (<http://lc.kubagro.ru/aidos/index.htm>), в которых не требуется автоматическое (без непосредственного участия человека в реальном времени решения задач идентификации) прогнозирование, принятие решений и исследование предметной области;

- находится в полном открытом бесплатном доступе (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), причем с актуальными исходными текстами (<http://lc.kubagro.ru/AIDOS-X.txt>);

- 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>);

- 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);

- содержит большое количество локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и 234, соответственно) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

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

- обеспечивает мультязычную 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;

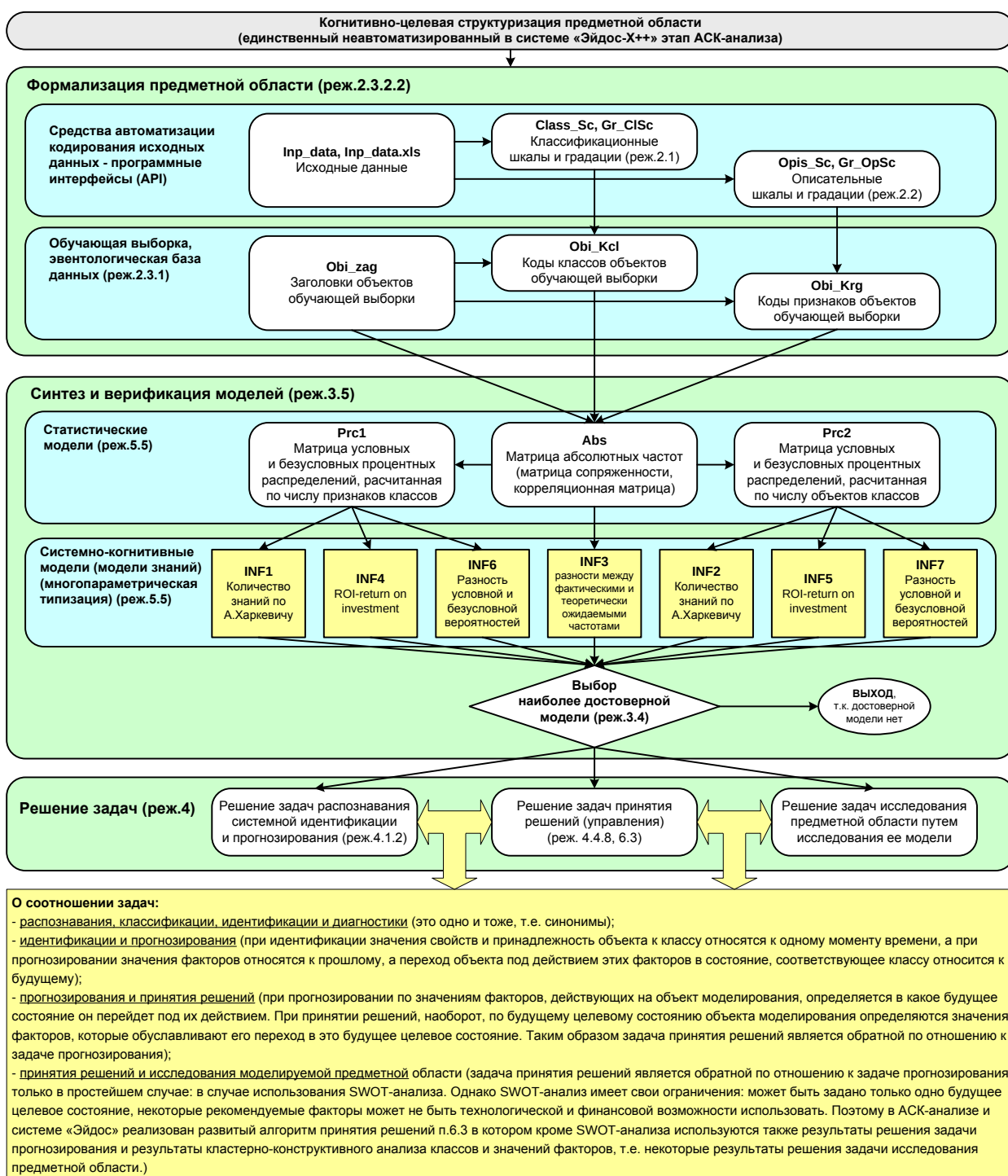
- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) 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 на теоретическом уровне познания в теоретических научных законах.

Всем этим и обусловлен выбор АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос» в качестве метода и инструмента решения поставленной проблемы и достижения цели работы (рисунок 1).

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



**Рисунок 1. Последовательность решения задач
в АСК-анализе и системе «Эйдос»**

Рассмотрим решение поставленных задач в подробном численном примере.

Задача 1: когнитивная структуризация **предметной области**

На этапе когнитивно-целевой структуризации предметной области мы не формализуемым путем решаем на качественном уровне, что будем рассматривать в качестве факторов, действующих на моделируемый объект, (причин), а что – в качестве результатов действия этих факторов (последствий).

При этом необходимо отметить, что системно-когнитивные модели (СК-модели) отражают лишь сам факт наличия зависимостей между значениями факторов и результатами их действия. Но они не отражают причин и механизмов такого влияния.

Это значит:

– во-первых, что содержательная интерпретация СК-моделей – это компетенция специалистов-экспертов, хорошо разбирающихся в данной предметной области. Иногда встречается ситуация, когда и то, что на первый взгляд является причинами, и то, что, казалось бы, является их последствиями, на самом деле является последствиями неких глубинных причин, которых мы не видим и никоим образом непосредственно не отражаем в модели;

– во-вторых, даже если содержательной интерпретации не разработано, то в принципе это не исключает возможности пользоваться ими на практике для достижения заданных результатов и поставленных целей, т.е. для управления.

В данной работе в качестве классификационных шкал выберем факт смерти (таблица 1), а в качестве факторов, влияющих на эти результаты – возраст, уровень фермента КФК в крови, диабет, фракцию выброса, высокое кровяное давление, тромбоциты, уровень сывороточного креатинина в крови, уровень сывороточного натрия в крови, пол, курение, период наблюдения (таблица 2):

Таблица 1 – Классификационная шкала

Код	Наименование
1	DEATH_EVEN

Таблица 2 – Описательные шкалы

Код	Наименование
1	AGE
2	ANAEMIA
3	CREATININE
4	DIABETES
5	EJECTION_F
6	HIGH_BLOOD
7	PLATELETS
8	SERUM_CREA
9	SERUM_SODI
10	SEX
11	SMOKING
12	TIME

Задача 2: подготовка исходных данных и формализация предметной области

Исходные данные для данной статьи (рисунок 2) получены из баз данных на сайте Kaggle.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	RECNUMBER	AGE	ANAEMIA	CREATININE	DIABETES	EJECTION_F	HIGH_BLOOD	PLATELETS	SERUM_CREA	SERUM_SODI	SEX	SMOKING	TIME	DEATH_EVENT
2	1	75,000	0	582	0	20	1	265000,00	1,9	130	1	0	4	да
3	2	65,000	0	7861	0	38	0	263358,03	1,1	136	1	0	6	да
4	3	65,000	0	146	0	20	0	162000,00	1,3	129	1	1	7	да
5	4	50,000	1	111	0	20	0	210000,00	1,9	137	1	0	7	да
6	5	65,000	1	160	1	20	0	327000,00	2,7	116	0	0	8	да
7	6	90,000	1	47	0	40	1	204000,00	2,1	132	1	1	8	да
8	7	75,000	1	246	0	15	0	127000,00	1,2	137	1	0	10	да
9	8	60,000	1	315	1	60	0	454000,00	1,1	131	1	1	10	да
10	9	65,000	0	157	0	65	0	263358,03	1,5	138	0	0	10	да
11	10	80,000	1	123	0	35	1	388000,00	9,4	133	1	1	10	да
12	11	75,000	1	81	0	38	1	368000,00	4,0	131	1	1	10	да
13	12	62,000	0	231	0	25	1	253000,00	0,9	140	1	1	10	да
14	13	45,000	1	981	0	30	0	136000,00	1,1	137	1	0	11	да
15	14	50,000	1	168	0	38	1	276000,00	1,1	137	1	0	11	да
16	15	49,000	1	80	0	30	1	427000,00	1,0	138	0	0	12	нет
17	16	82,000	1	379	0	50	0	47000,00	1,3	136	1	0	13	да
18	17	87,000	1	149	0	38	0	262000,00	0,9	140	1	0	14	да
19	18	45,000	0	582	0	14	0	166000,00	0,8	127	1	0	14	да
20	19	70,000	1	125	0	25	1	237000,00	1,0	140	0	0	15	да
21	20	48,000	1	582	1	55	0	87000,00	1,9	121	0	0	15	да
22	21	65,000	1	52	0	25	1	276000,00	1,3	137	0	0	16	нет
23	22	65,000	1	128	1	30	1	297000,00	1,6	138	0	0	20	да
24	23	68,000	1	220	0	35	1	289000,00	0,9	140	1	1	20	да
25	24	53,000	0	63	1	60	0	368000,00	0,8	135	1	0	22	нет
26	25	75,000	0	582	1	30	1	263358,03	1,8	134	0	0	23	да
27	26	80,000	0	148	1	38	0	149000,00	1,9	144	1	1	23	да
28	27	95,000	1	112	0	40	1	196000,00	1,0	138	0	0	24	да
29	28	70,000	0	122	1	45	1	284000,00	1,3	136	1	1	26	да
30	29	58,000	1	60	0	38	0	153000,00	5,8	134	1	0	26	да
31	30	82,000	0	70	1	30	0	200000,00	1,2	132	1	1	26	да
32	31	94,000	0	582	1	38	1	263358,03	1,8	134	1	0	27	да
33	32	85,000	0	23	0	45	0	360000,00	3,0	132	1	0	28	да
34	33	50,000	1	249	1	35	1	319000,00	1,0	128	0	0	28	да
35	34	50,000	1	159	1	30	0	302000,00	1,2	138	0	0	29	нет
36	35	65,000	0	94	1	50	1	188000,00	1,0	140	1	0	29	да
37	36	69,000	0	582	1	35	0	228000,00	3,5	134	1	0	30	да
38	37	90,000	1	60	1	50	0	226000,00	1,0	134	1	0	30	да
39	38	82,000	1	855	1	50	1	321000,00	1,0	145	0	0	30	да
40	39	89,000	0	2656	1	30	0	305000,00	2,3	137	1	0	30	нет

Рисунок 2 – Исходные данные для ввода в систему «Эйдос»

Затем с параметрами, показанными на рисунке 3, запустим режим 2.3.2.2 системы «Эйдос», представляющий собой автоматизированный программный интерфейс (API) с внешними данными табличного типа.

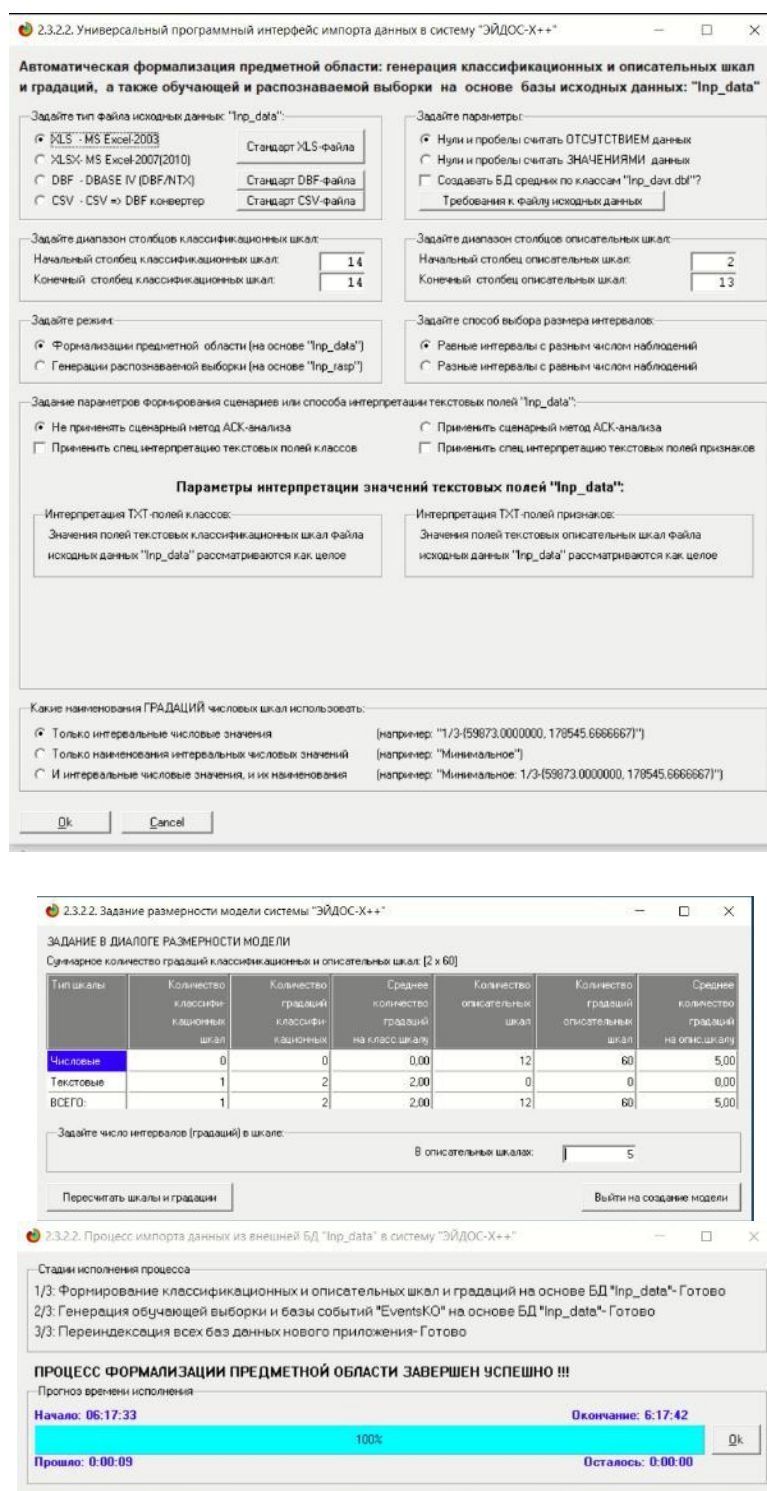


Рисунок 3. Экранные форма программного интерфейса (API) 2.3.2.2 системы «Эйдос» с внешними данными табличного типа

Обратим внимание на то, что заданы адаптивные интервалы, учитывающее неравномерность распределения данных по диапазону значений, что важно при относительно небольшом числе наблюдений. Если бы интервалы были заданы равными по величине, то в них бы учитывалось сильно отличающееся число наблюдений, а в некоторых интервалах их бы могло не оказаться вовсе. В описательных шкалах задано 5 числовых интервальных значения.

На рисунке 4 приведен Help данного режима, в котором объясняется принцип организации таблицы исходных данных для данного режима. Здесь же обратим внимание на то, что в таблице 3 значения параметров могут быть представлены как числовыми, так и текстовыми значениями.



Рисунок 4. Экранная форма HELP программного интерфейса (API) 2.3.2.2

В результате работы режима сформирована 1 классификационная шкала с суммарным количеством градаций (классов) 2 (таблица 3) и 7 описательных шкал с суммарным числом градаций 35 (таблица 4). С использованием классификационных и описательных шкал и градаций исходные данные (рисунок 2) были закодированы и в результате получена обучающая выборка (рисунок 5):

Таблица 3 – Классификационные шкалы и градации
(факт смерти)

Код	Название
1	да
2	нет

Таблица 4 – Описательные шкалы и градации
(признаки)

Код	Название
1	1/5-{40.00000000, 51.00000000}
2	2/5-{51.00000000, 62.00000000}
3	3/5-{62.00000000, 73.00000000}
4	4/5-{73.00000000, 84.00000000}
5	5/5-{84.00000000, 95.00000000}
6	1/5-{23.00000000, 1590.60000000}
7	2/5-{1590.60000000, 3158.20000000}
8	3/5-{3158.20000000, 4725.80000000}
9	4/5-{4725.80000000, 6293.40000000}
10	5/5-{6293.40000000, 7861.00000000}
11	1/5-{14.00000000, 27.20000000}
12	2/5-{27.20000000, 40.40000000}

13	3/5-{40.4000000, 53.6000000}
14	4/5-{53.6000000, 66.8000000}
15	5/5-{66.8000000, 80.0000000}
16	1/5-{25100.0000000, 19080.0000000}
17	2/5-{19080.0000000, 355060.0000000}
18	3/5-{355060.0000000, 520040.0000000}
19	4/5-{520040.0000000, 685020.0000000}
20	5/5-{685020.0000000, 850000.0000000}
21	1/5-{0.5000000, 2.2800000}
22	2/5-{2.2800000, 4.0600000}
23	3/5-{4.0600000, 5.8400000}
24	4/5-{5.8400000, 7.6200000}
25	5/5-{7.6200000, 9.4000000}
26	1/5-{113.0000000, 120.0000000}
27	2/5-{120.0000000, 127.0000000}
28	3/5-{127.0000000, 134.0000000}
29	4/5-{134.0000000, 141.0000000}
30	5/5-{141.0000000, 148.0000000}
31	1/5-{4.0000000, 60.2000000}
32	2/5-{60.2000000, 116.4000000}
33	3/5-{116.4000000, 172.6000000}
34	4/5-{172.6000000, 228.8000000}
35	5/5-{228.8000000, 285.0000000}

№	Наименование objects	14	2	3	4	5	6	7	8	9	10	11	12	13
		DEATH_EVENT	AGE	SEX	CREATININE	DIA...	EJECTION_F	HIGH...	PLATELETS	SERUM_CREA	SERUM_SODI	S...	SM...	TIME
29		1	2		6		12		16	23	28			31
30		1	4		6		12		17	21	28			31
31		1	5		6		12		17	21	28			31
32		1	5		6		13		18	22	28			31
33		1	1		6		12		17	21	28			31
34		2	1		6		12		17	21	29			31
35		1	3		6		13		16	21	29			31
36		1	3		6		12		17	22	28			31
37		1	5		6		13		17	21	28			31
38		1	4		6		13		17	21	30			31
39		2	2		7		12		17	22	29			31
40		1	2		6		12		17	22	30			31
41		1	3		6		11		17	21	28			31
42		1	1		6		12		16	21	29			31
43		1	3		6		13		16	21	29			31
44		2	3		6		13		17	21	28			31
45		1	2		6		14		17	21	30			31
46		1	1		6		12		17	21	29			31
47		1	1		6		11		17	21	28			31
48		1	2		6		12		18	21	29			31
49		1	4		6		11		16	23	28			31
50		1	2		6		12		18	21	29			31
0	1	14	2	3	4	5	6	7	8	9	10	11	12	13

Рисунок 5 – Обучающая выборка (фрагмент)

Обучающая выборка (рисунок 5), по сути, представляет собой нормализованные исходные данные, т.е. таблицу исходных данных (рисунок 2), закодированную с помощью классификационных и описательных шкал и градаций (таблицы 3 и 4). Таким образом, созданы все необходимые и достаточные условия для выполнения следующего этапа АСК-анализа: т.е. для синтеза и верификации моделей.

Задача 3: синтез и верификация статистических и системно-когнитивных моделей и выбор наиболее достоверной из них для решения задач

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) осуществляется в режиме 3.5 системы «Эйдос» (рисунок 6).

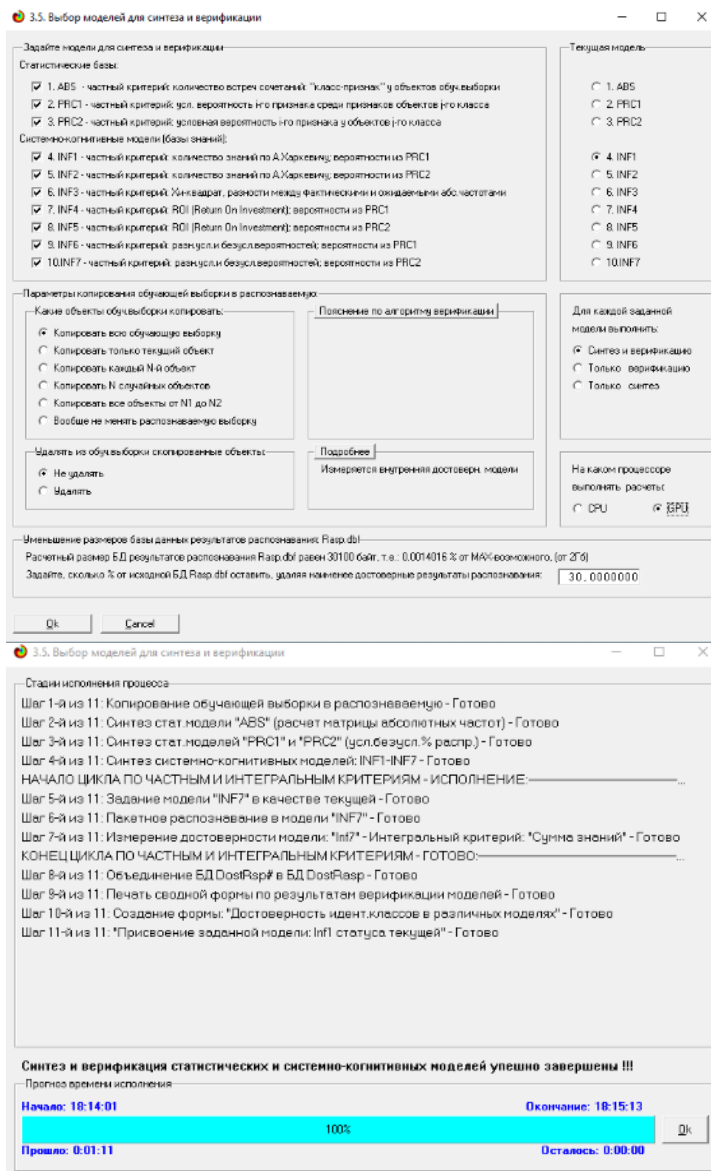


Рисунок 6. Экранные формы режима синтеза и верификации статистических и системно-когнитивных моделей системы «Эйдос»

Обратим внимание на то, что на рисунке 6 в правом нижнем углу окна задана опция: «Расчеты проводить на графическом процессоре (GPU)».

Из рисунка 6 видно, что весь процесс синтеза и верификации моделей занял 1 минуту и 11 секунд. Отметим, что при синтезе и верификации моделей использовался графический процессор (GPU) видеокарты. На центральном процессоре (CPU) выполнение этих операций занимает значительно большее время (на некоторых задачах это происходит в десятки, сотни и даже тысячи раз дольше). Таким образом, неграфические вычисления на графических процессорах видеокарты делает возможной обработку больших объемов исходных данных за разумное время. В процесс синтеза и верификации моделей осуществляется также расчет 10 выходных форм, на что уходит более 99% времени исполнения.

Фрагменты самих созданных статистических и системно-когнитивных моделей (СК-модели) приведены на рисунках 7, 8:

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

Код признака	Наименование описательной шкалы и градаций	1. DEATH_EVEN ДА	2. DEATH_EVEN НЕТ	Сумма	Среднее	Средн. коэфф. откл.
1	AGE-1/5 (40.000000, 51.000000)	20	58	78	39.00	26.87
2	AGE-2/5 (51.000000, 62.000000)	26	69	95	47.50	30.41
3	AGE-3/5 (62.000000, 73.000000)	25	60	85	42.50	24.75
4	AGE-4/5 (73.000000, 84.000000)	15	12	27	13.50	2.12
5	AGE-5/5 (84.000000, 95.000000)	10	4	14	7.00	4.24
6	CREATININE-1/5 (23.000000, 1590.600000)	88	185	273	136.50	68.59
7	CREATININE-2/5 (1590.600000, 3158.200000)	4	15	19	9.50	7.78
8	CREATININE-3/5 (3158.200000, 4725.800000)	1	2	3	1.50	0.71
9	CREATININE-4/5 (4725.800000, 6293.400000)	1	1	2	1.00	
10	CREATININE-5/5 (6293.400000, 7861.000000)	2		2	1.00	1.41
11	EJECTION_F-1/5 (14.000000, 27.200000)	38	21	59	29.50	12.02
12	EJECTION_F-2/5 (27.200000, 40.400000)	39	121	160	80.00	57.98
13	EJECTION_F-3/5 (40.400000, 53.600000)	11	30	41	20.50	13.44
14	EJECTION_F-4/5 (53.600000, 66.800000)	7	30	37	18.50	16.26
15	EJECTION_F-5/5 (66.800000, 80.000000)	1	1	2	1.00	
16	PLATELETS-1/5 (25100.000000, 190080.000000)	20	35	55	27.50	10.61
17	PLATELETS-2/5 (190080.000000, 355060.000000)	62	141	203	101.50	55.86
18	PLATELETS-3/5 (355060.000000, 520040.000000)	13	23	36	18.00	7.07
19	PLATELETS-4/5 (520040.000000, 685020.000000)	1	2	3	1.50	0.71
20	PLATELETS-5/5 (685020.000000, 850000.000000)		2	2	1.00	1.41
21	SERUM_CREA-1/5 (0.500000, 2.280000)	79	192	271	135.50	79.90
22	SERUM_CREA-2/5 (2.280000, 4.060000)	12	9	21	10.50	2.12
23	SERUM_CREA-3/5 (4.060000, 5.840000)	2	1	3	1.50	0.71
24	SERUM_CREA-4/5 (5.840000, 7.620000)	1	1	2	1.00	
25	SERUM_CREA-5/5 (7.620000, 9.400000)	2		2	1.00	1.41
26	SERUM_SODI-1/5 (113.000000, 120.000000)	1	1	2	1.00	

Рисунок 7. Матрица абсолютных частот (фрагмент)

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

Код признака	Наименование описательной шкалы и градаций	1. DEATH_EVEN ДА	2. DEATH_EVEN НЕТ	Сумма	Среднее	Средн. коэфф. откл.
1	AGE-1/5 (40.000000, 51.000000)	-5.043	5.043			7.133
2	AGE-2/5 (51.000000, 62.000000)	-4.502	4.502			6.366
3	AGE-3/5 (62.000000, 73.000000)	-2.291	2.291			3.240
4	AGE-4/5 (73.000000, 84.000000)	6.331	-6.331			8.954
5	AGE-5/5 (84.000000, 95.000000)	5.505	-5.505			7.785
6	CREATININE-1/5 (23.000000, 1590.600000)	0.348	-0.348			0.492
7	CREATININE-2/5 (1590.600000, 3158.200000)	-2.100	2.100			2.970
8	CREATININE-3/5 (3158.200000, 4725.800000)	0.037	-0.037			0.052
9	CREATININE-4/5 (4725.800000, 6293.400000)	0.358	-0.358			0.506
10	CREATININE-5/5 (6293.400000, 7861.000000)	1.358	-1.358			1.920
11	EJECTION_F-1/5 (14.000000, 27.200000)	19.057	-19.057			26.950
12	EJECTION_F-2/5 (27.200000, 40.400000)	-12.371	12.371			17.496
13	EJECTION_F-3/5 (40.400000, 53.600000)	-2.164	2.164			3.060
14	EJECTION_F-4/5 (53.600000, 66.800000)	-4.880	4.880			6.901
15	EJECTION_F-5/5 (66.800000, 80.000000)	0.358	-0.358			0.506
16	PLATELETS-1/5 (25100.000000, 190080.000000)	2.341	-2.341			3.311
17	PLATELETS-2/5 (190080.000000, 355060.000000)	-3.177	3.177			4.493
18	PLATELETS-3/5 (355060.000000, 520040.000000)	1.441	-1.441			2.039
19	PLATELETS-4/5 (520040.000000, 685020.000000)	0.037	-0.037			0.052
20	PLATELETS-5/5 (685020.000000, 850000.000000)	-0.642	0.642			0.908
21	SERUM_CREA-1/5 (0.500000, 2.280000)	-8.010	8.010			11.328
22	SERUM_CREA-2/5 (2.280000, 4.060000)	5.258	-5.258			7.435
23	SERUM_CREA-3/5 (4.060000, 5.840000)	1.037	-1.037			1.466
24	SERUM_CREA-4/5 (5.840000, 7.620000)	0.358	-0.358			0.506
25	SERUM_CREA-5/5 (7.620000, 9.400000)	1.358	-1.358			1.920
26	SERUM_SODI-1/5 (113.000000, 120.000000)	0.358	-0.358			0.506

Рисунок 8. Модель INF3 (фрагмент)

Отметим, что в АСК-анализе и СК-моделях степень выраженности различных свойств объектов наблюдения рассматривается с единственной точки зрения: с точки зрения того, какое *количество информации* содержится в них о том, к каким обобщающим категориям (классам) будут принадлежать или не принадлежать эти объекты. Поэтому не играет никакой роли, в каких единицах измерения измеряются те или иные свойства объектов наблюдения, а также в каких единицах измеряются результаты влияния этих свойств, натуральных, в процентах или стоимостных. Это и есть решение проблемы сопоставимости в АСК-анализе и системе «Эйдос», отличающее их от других интеллектуальных технологий.

Верификация статистических и системно-когнитивных моделей

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по

Рисунок 9. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергена и L1-критерию проф. Е.В.Луценко

Это подтверждает наличие и адекватное отражение в СК-модели INF3 сильной причинно-следственной зависимости между фактом смерти и выявленными признаками.

На рисунке 10 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам идентификации выявленных признаков в СК-модели INF3 по данным обучающей выборки:



Рисунок 10. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf3

Рисунок 10 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе.

Левое распределение включает истинно-отрицательные и ложно-положительные решения, а правое включает ложные отрицательные и истинно-положительные решения.

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу прогнозирования сердечных заболеваний и другие задачи.

Из рисунка 10 видно, что для положительных решений картина включает 2 диапазона уровней сходства:

1) при уровнях сходства от 0% до примерно 50% есть и истинные, и ложные положительные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства, когда как доля отрицательных, наоборот, падает;

3) при уровнях сходства выше 50% ложные положительные решения не встречаются.

Для отрицательных решений картина включает 3 диапазона уровней различия:

1) при уровнях различия от 0% до примерно 23% ложно-отрицательных решений больше, чем истинно-отрицательных;

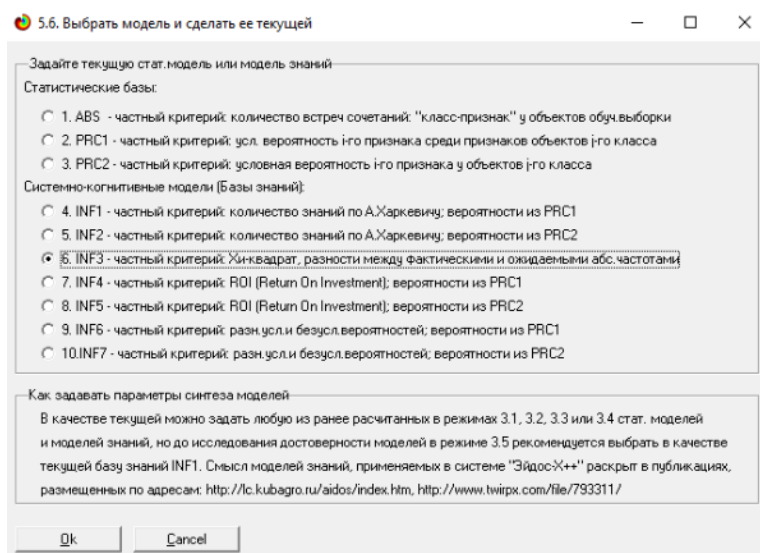
2) при уровнях различия от 23% до примерно 73%, встречаются как истинно-отрицательные, так и ложно-отрицательные, но число истинных решений больше;

3) при уровнях различия больше 73% ложно-отрицательные решения отсутствуют.

На рисунке 11 приведен Help по режиму 3.4, в котором описаны меры достоверности моделей, применяемые в системе «Эйдос»:

Выбор наиболее достоверной модели и присвоение ей статуса текущей

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос» (рисунок 1), присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (рисунок 12):



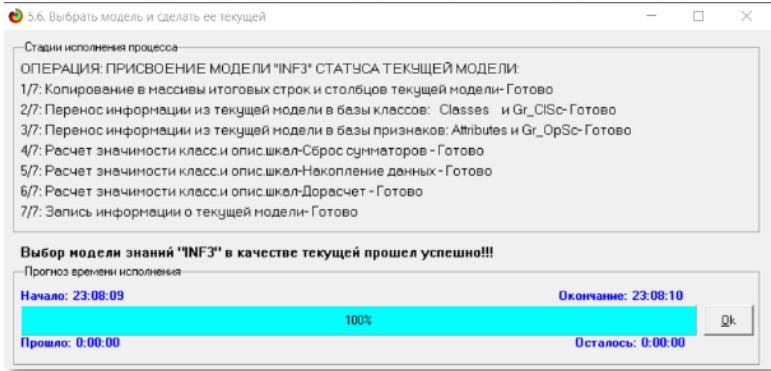


Рисунок 12. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

Задача 4: решение различных задач в наиболее достоверной модели

Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)

Решим задачу системной идентификации, т.е. прогнозирования сердечных заболеваний на основе обучающей выборки в наиболее достоверной СК-модели INF3 на GPU (рисунок 13).

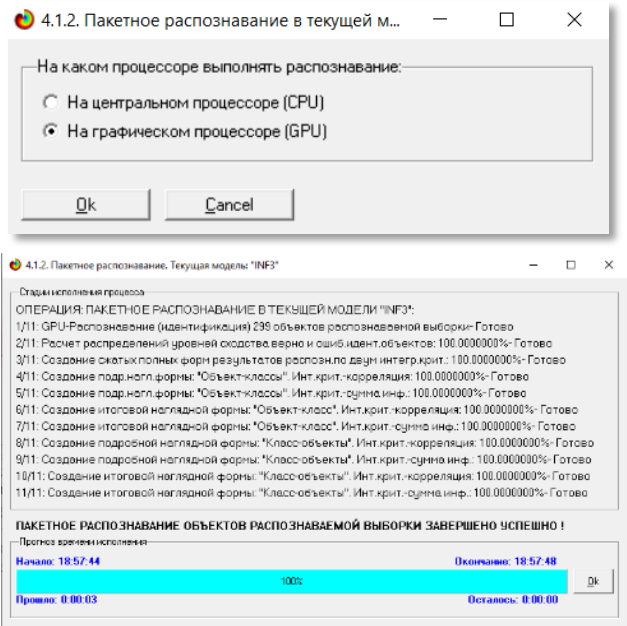


Рисунок 13. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Из рисунка 13 видно, что процесс идентификации занял 3 секунды. Отметим, что 99% этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этого прогнозирования. Эти формы отражают результаты прогнозирования в различных разрезах и обобщениях:

Приведем две из этих 10 форм: 4.1.3.1 и 4.1.3.2 (рисунок 14).

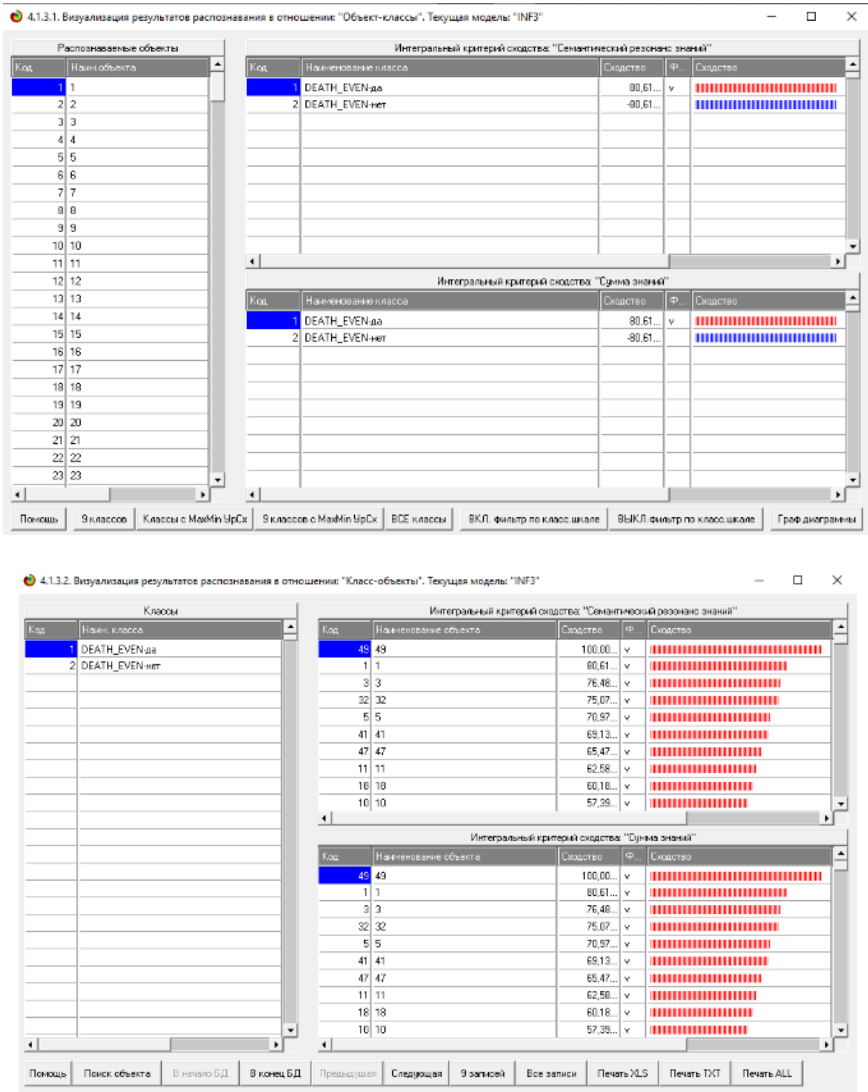


Рисунок 14. Выходные формы по результатам прогнозирования сердечных заболеваний

Символ «✓» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка 14 видно, что результаты идентификации являются отличными, естественно при учете информации из рисунка 10 о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 50%, т.е., по сути, результаты с более низким уровнем сходства надо просто игнорировать.

Подзадача 4.2. Поддержка принятия решений (SWOT-анализ)

При принятии решений определяется сила и направление влияния значений факторов на принадлежность состояний объекта моделирования к тем или иным классам, соответствующим различным будущим состояниям. По сути, это решение задачи SWOT-анализа.

Применительно к задаче, решаемой в данной работе, SWOT-анализ показывает степень влияния различных значений выявленных признаков на факт смерти пациента от сердечных заболеваний.

В системе «Эйдос» в режиме 4.4.8 поддерживается решение этой задачи. При этом ***выявляется система детерминации заданного класса***, т.е. система значений факторов, обуславливающих переход объекта моделирования и управления в состояние, соответствующее данному классу, а также препятствующих этому переходу.

На рисунках 15 приведены SWOT-диаграммы, наглядно отражающие силу и направление влияния различных значений признаков на факт смерти в результате сердечных заболеваний.

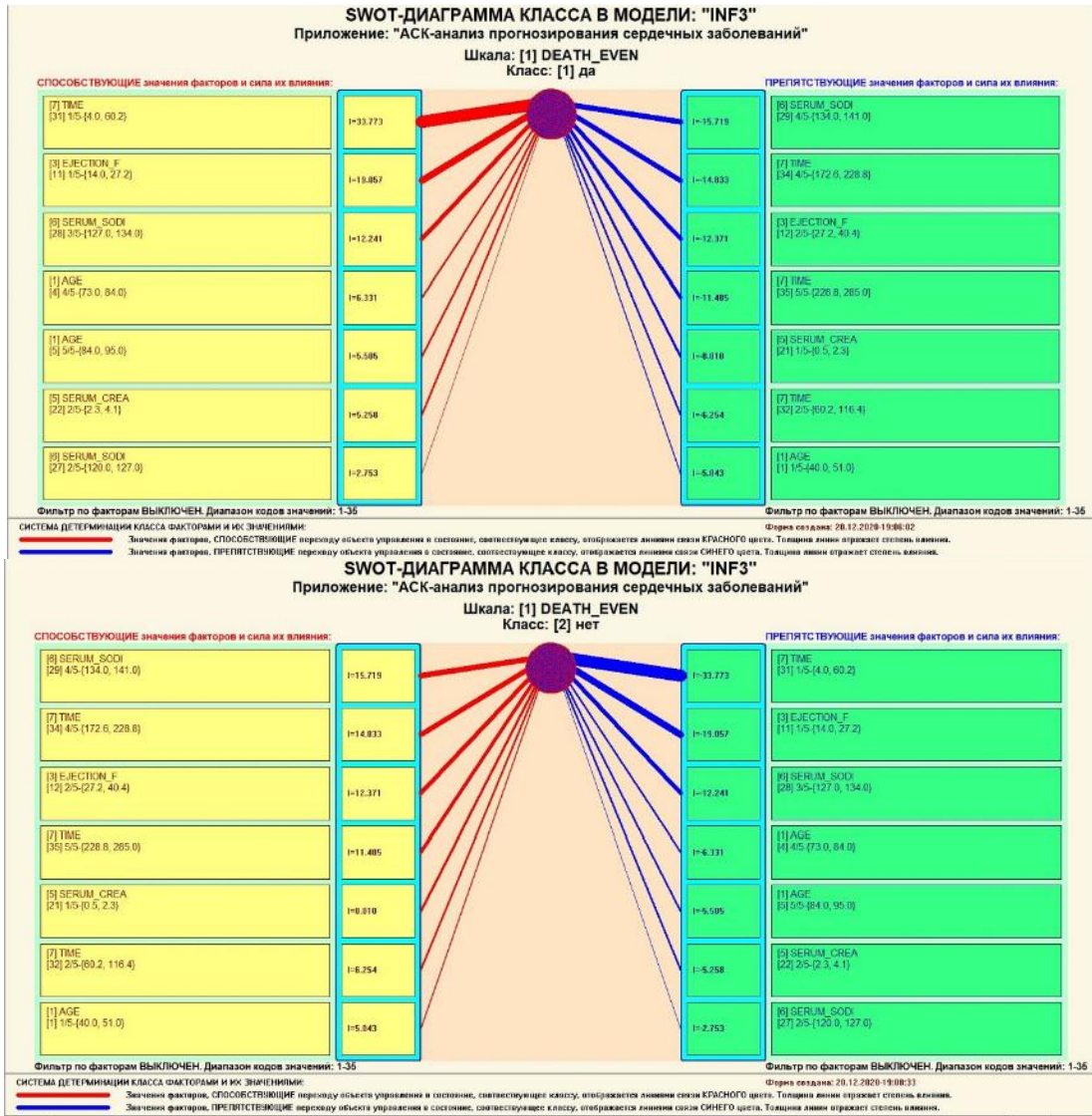


Рисунок 15. SWOT-диаграммы детерминации признаков сердечных заболеваний

Эти SWOT-диаграммы наглядно отражают силу и направление влияния различных значений признаков на смерть пациента в результате сердечных заболеваний.

Отметим также, что система «Эйдос» обеспечивала решение этой задачи ***всегда***, т.е. даже в самых ранних DOS-версиях и в реализациях системы «Эйдос» на других языках и типах компьютеров. Например, первый

акт внедрения системы «Эйдос», где об этом упоминается в явном виде, датируется 1987 годом, а первый подобный расчет относится к 1981 году.

УТВЕРЖДАЮ Заведующий Краснодарским сектором ИСИ АН СССР, к.ф.н. <u>А.А. Хагуров</u> 1987г.	УТВЕРЖДАЮ Директор Северо-Кавказского филиала ВНИИ "АИУС-агроресурсы", к.э.н. <u>Э.М. Трахов</u> 1987г.
А К Т	

Настоящий акт составлен комиссией в составе: Кириченко М.М., Ляшко Г.А., Самсонов Г.А., Коренец В.И., Луценко В.В. в том, что в соответствии с договором о научно-техническом сотрудничестве между Северо-Кавказским филиалом ВНИИ "АИУС-агроресурсы" и Краснодарским сектором Института социологических исследований АН СССР Северо-Кавказским филиалом ВНИИ "АИУС-агроресурсы" выполнены следующие работы:

- осуществлена постановка задачи: "Обработка на ЭВМ социологических анкет Крайагропрома;
- разработаны математическая модель и программное обеспечение подсистемы распознавания образов, позволяющие решать данную задачу в среде персональной технологической системы ВЕГА-М;
- на профессиональной персональной ЭВМ "Искра-226" осуществлены расчеты по задаче в объеме:

Входная информация составила 425 анкет по 9-ти предприятиям.
 Выходная информация - 4 вида выходных форм объемом 90 листов формата А3 и 20 листов формата А4 содержит:

- процентное распределение ответов в разрезе по социальным типам корреспондентов;
- распределение информативностей признаков (в битах) для распознавания социальных типов корреспондентов;
- позитивные и негативные информационные портреты 30-ти социальных типов на языке 212 признаков;
- обобщенная характеристика информативности признаков для выбора такого минимального набора признаков, который содержит максимум информации о распознаваемых объектах (оптимизация анкет).

Работы выполнены на высоком научно-методическом уровне и в срок.

От ИСИ АН СССР: Мл. научный сотрудник <u>Кириченко М.М.</u> 19.05.1987г. Мл. научный сотрудник <u>Ляшко Г.А.</u> 19.05.1987г.	От СКФ ВНИИ "АИУС-агроресурсы": Зав. отделом аэрокосмических и тематических исследований №4, к.э.н. <u>Самсонов Г.А.</u> 19.05.1987г. Главный конструктор проекта <u>Коренец В.И.</u> 19.05.87г. Главный конструктор проекта <u>Луценко В.В.</u> 19.05.87г.
--	---

Но тогда SWOT-диаграммы назывались позитивным и негативным информационными портретами классов.

Информация о системе значений факторов, обуславливающих переход объекта моделирования в различные будущие состояния, соответствующие классам, может быть приведена не только в диаграммах, показанных на рисунках 15, но и во многих других табличных и графических формах, которые в данной работе не приводятся только из-за ограниченности ее объема. В частности, в этих формах может быть выведена значительно более полная информация (в т. ч. вообще вся имеющая в модели). Подобная подробная информация содержится в базах данных, расположенных по пути: \Aidos-X\AID_DATA\A0000001\System\SWOTCls####Inf3.DBF, где: «####» – код класса с ведущими нулями. Эти базы открываются в MS Excel.

На рисунке 16 приведены примеры нескольких инвертированных SWOT-диаграмм, отражающих силу и направление влияния факта смерти от сердечных заболеваний на их признаки.

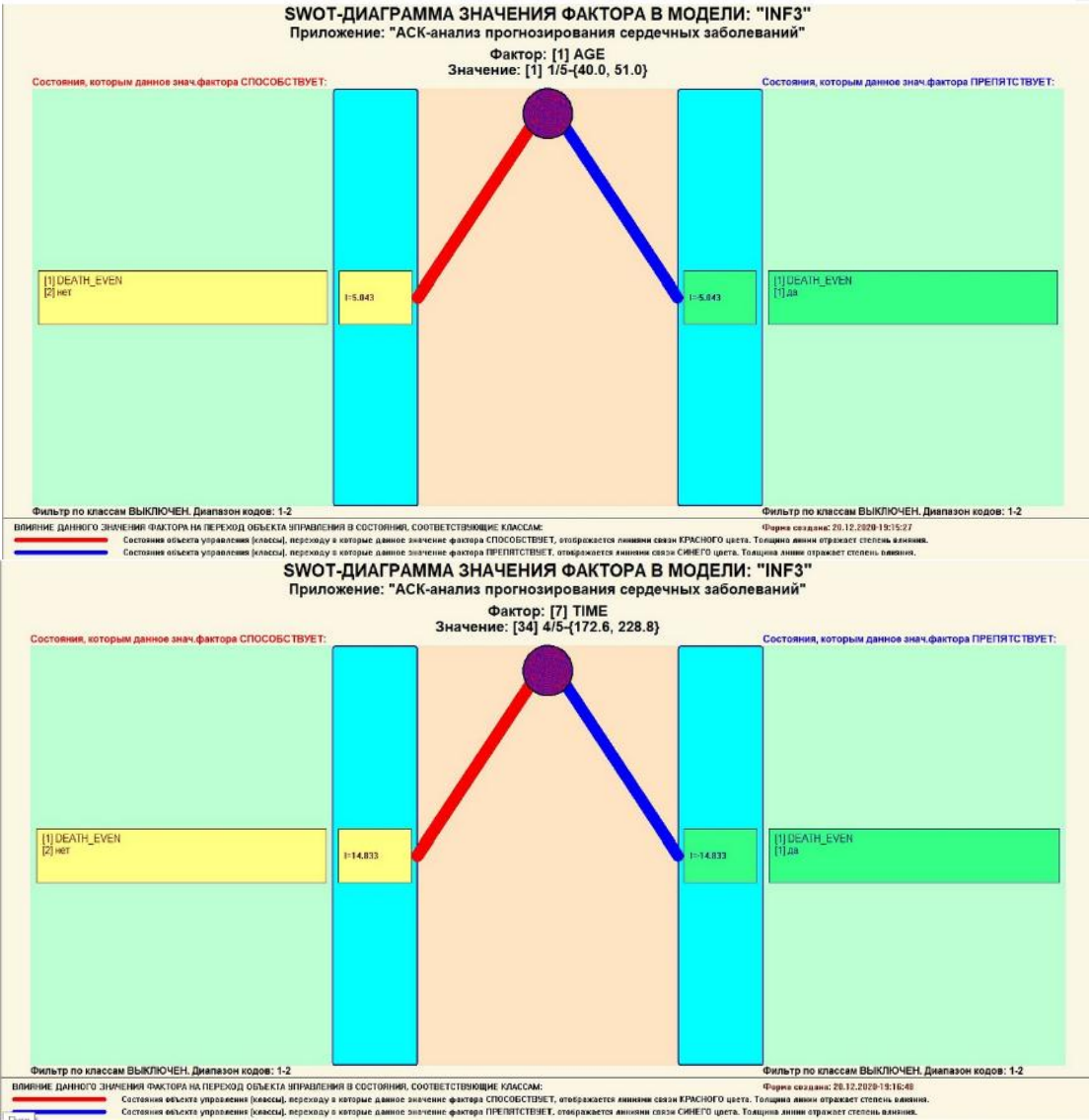


Рисунок 16. Примеры SWOT-диаграмм, отражающих силу и направление влияния факта смерти от сердечных заболеваний на их признаки

В заключение отметим, что SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых чаще всего является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают не формализуемым путем (интуитивно), на основе своего опыта и профессиональной компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей

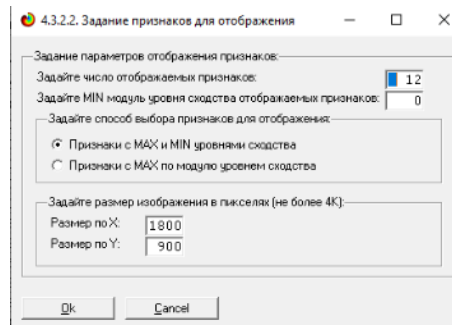


Рисунок 18. Параметры отображения когнитивной диаграммы, приведенной на рисунке 17

4.3.2. Агломеративная когнитивная кластеризация значений факторов

На рисунке 19 приведена агломеративная дендрограмма когнитивной кластеризации значений факторов и график изменения межкластерных расстояний, полученные на основе той же матрицы сходства признаков по их смыслу, что и в когнитивных диаграммах, пример которой приведен на рисунке 17.

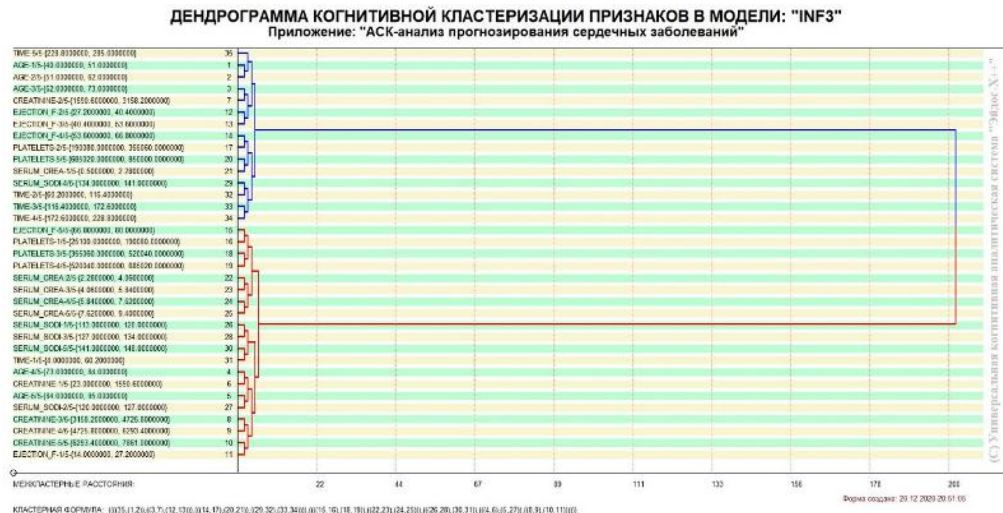


Рисунок 19. Дендрограмма агломеративной когнитивной кластеризации признаков

Из дендрограммы на рисунке 19 мы видим, что все значения факторов образуют 2 четко выраженных кластера, объединенных в полюса конструкта (показаны синими и красным цветами).

Хорошо видна группировка признаков по их смыслу, т.е. по содержащейся в них информации о прогнозе смерти в результате заболеваний сердца. **Значения факторов на полюсах конструкта факторов (рисунок 19) обуславливают переход объекта моделирования в состояния, соответствующие классам, представленным на полюсах конструкта классов.**

На рисунке 20 приведен график межкластерных расстояний значений признаков.



Рисунок 20. График изменения межкластерных расстояний при когнитивной кластеризации значений факторов

4.3.3. Нелокальные нейроны и нелокальные нейронные сети

На рисунке 21 приведён пример нелокального нейрона, а на рисунке 22 - фрагмент одного слоя нелокальной нейронной сети:

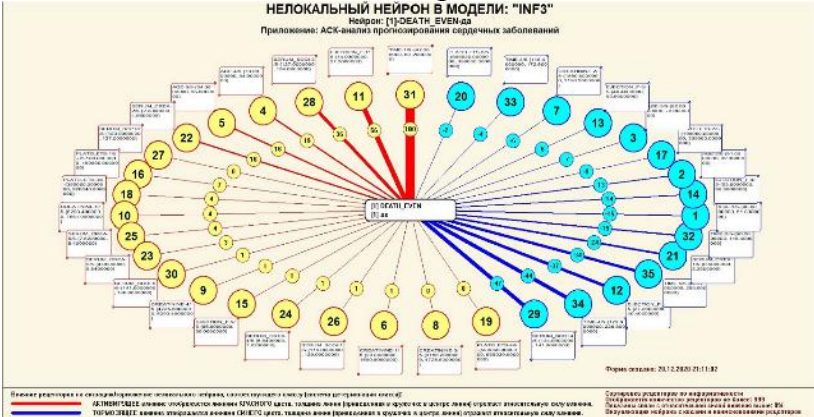


Рисунок 21. Пример нелокального нейрона, отражающего силу и направление влияния значений признаков на прогнозирование заболеваний

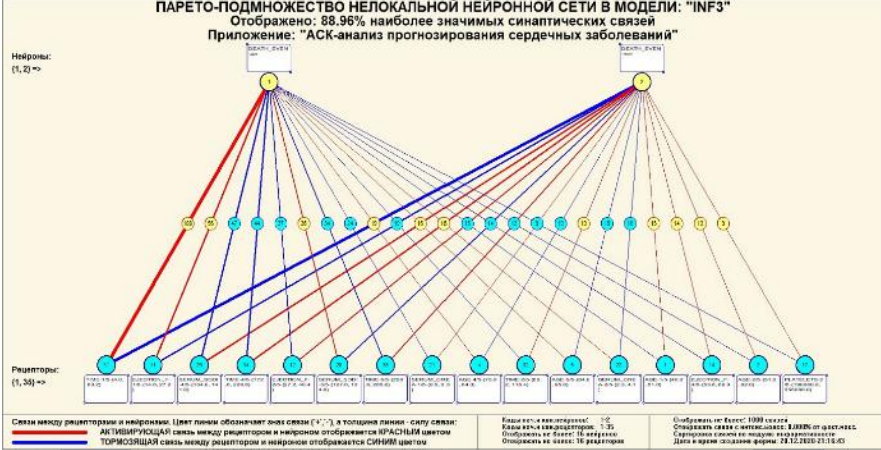


Рисунок 2. Один слой нелокальной нейронной сети, отражающий силу и направление влияния значений признаков на прогноз заболеваний (фрагмент 89% сети)

В приведенном фрагменте слоя нейронной сети нейроны соответствуют классу прогноза летального исхода в результате заболеваний сердца, а рецепторы – их признаков. Нейроны расположены слева направо в порядке убывания силы детерминации, т.е. слева находятся результаты, наиболее жестко обусловленные обуславливающими их значениями факторами, а справа – менее жестко обусловленные.

Модель знаний системы «Эйдос» относится к **нечетким декларативным** гибридным моделям и объединяет в себе некоторые 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.

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

- 1) весовые коэффициенты на рецепторах не подбираются итерационным методом обратного распространения ошибки, а считаются прямым счетом 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».

4.3.4. 3d-интегральные когнитивные карты

На рисунке 23 приведен фрагмент 3d-интегральной когнитивной карты, отражающий фрагмент около 89% СК-модели Inf3.

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивных диаграмм классов и значений факторов, отображенных рисунке 17, и одного слоя нейронной сети, приведенного на рисунке 22.

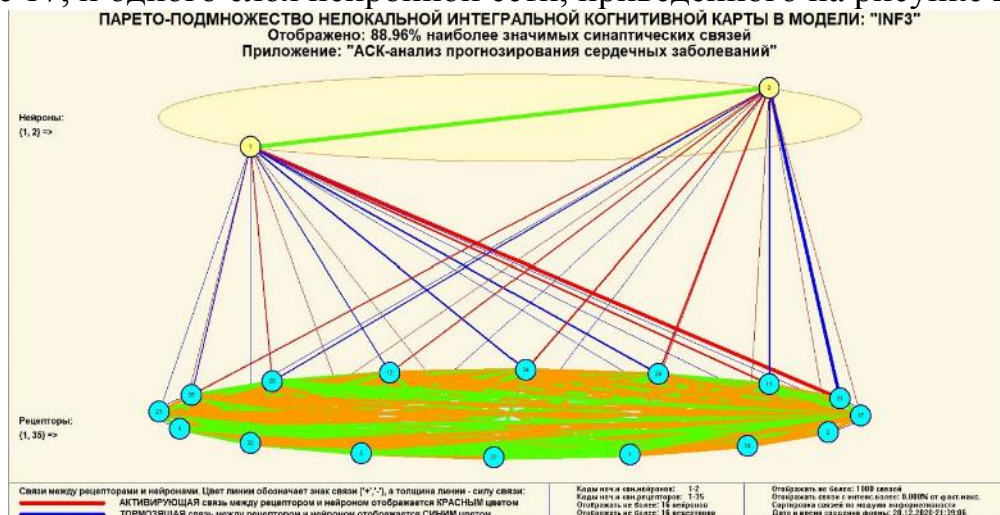


Рисунок 23. 3d-интегральная когнитивная карта в СК-модели Inf3

4.3.5. Когнитивные функции

Вместо описания того, что представляют собой когнитивные функции, приведем help соответствующего режима системы «Эйдос» (рисунок 24).

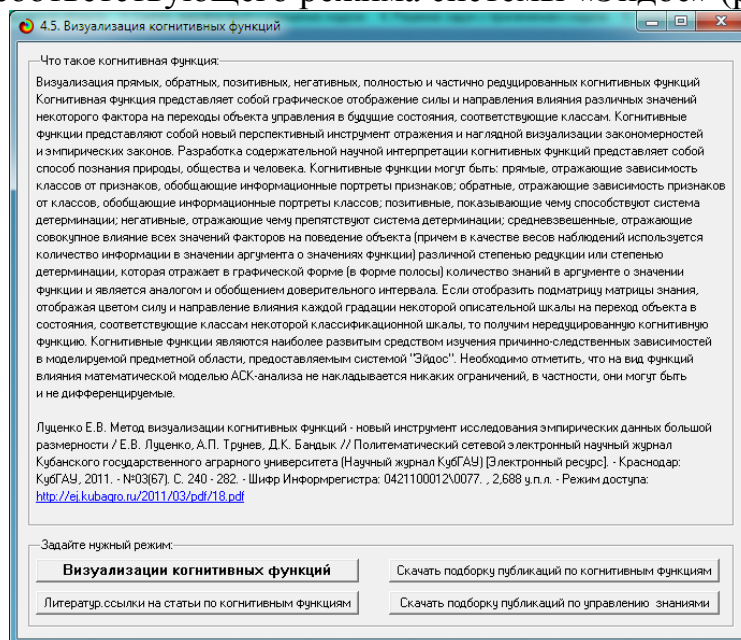


Рисунок 24. Help режима визуализации когнитивных функций

Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора (признаков) на переходы объекта управления в будущие состояния, соответствующие классам. Классы являются градациями классификационных шкал.

Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации эмпирических закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека.

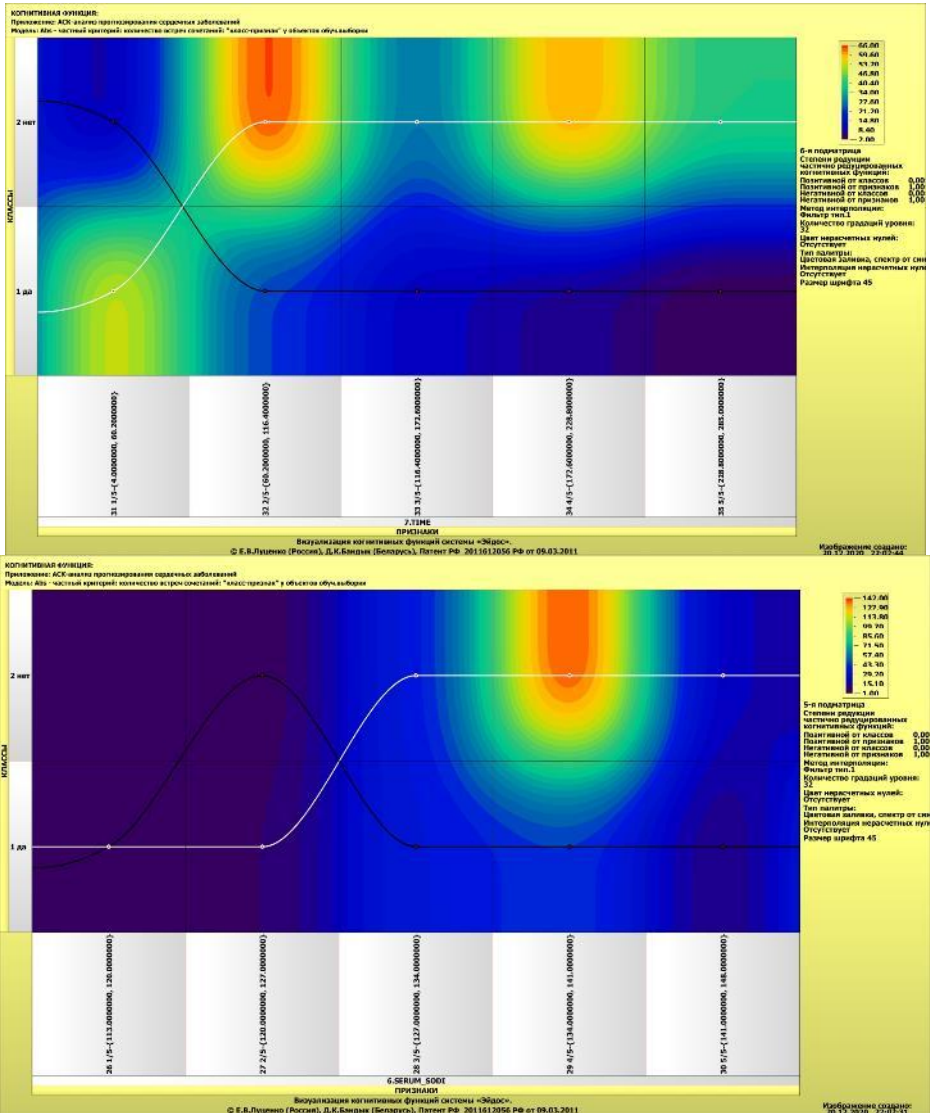
Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации (обозначены белой линией); негативные, отражающие чему препятствуют система детерминации (обозначены черной линией); 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 разной толщины) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала.

Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию.

Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой "Эйдос".

Необходимо отметить, что *на вид функций влияния математической моделью АСК-анализа не накладывается никаких ограничений*, в частности, они могут быть и не дифференцируемые.

На рисунках 25 приведены когнитивные функции, наглядно отражающие силу и направление влияния значений (т.е. степени выраженности) различных признаков сердечных заболеваний на прогнозирование летального исхода.



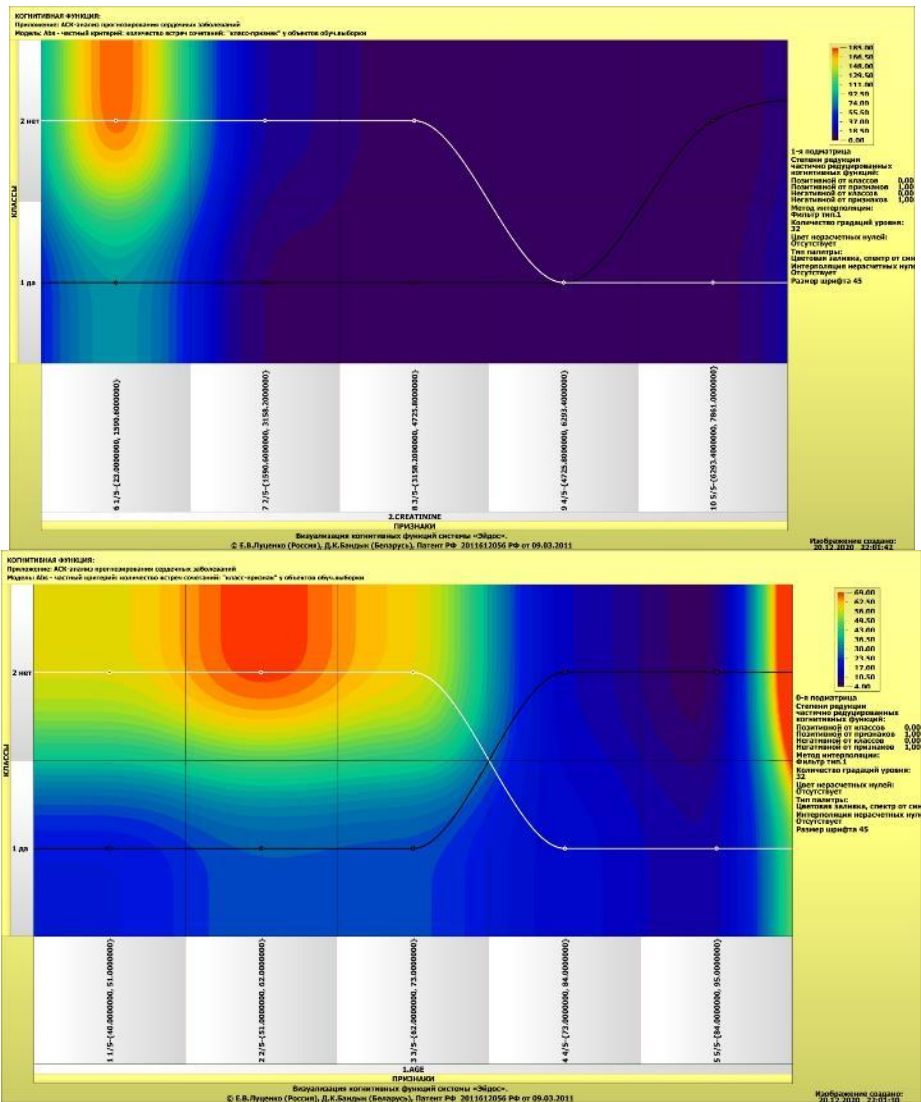


Рисунок 25. Примеры когнитивных функций в СК-модели INF3, отражающих силу и направление влияния значений признаков заболеваний сердца на прогнозирование летального исхода

Из когнитивных функций, приведенных на рисунке 25, хорошо видно, что *зависимости между значениями признаков заболеваний сердца на прогнозирование летального исхода в основном имеют довольно предсказуемый характер.*

4.3.6. Сила и направление влияния значений признаков заболеваний сердца и сила влияния этих признаков на прогнозирование смерти

На рисунках 7, 8 приведены фрагменты некоторых статистических и системно-когнитивных моделей, отражающих моделируемую предметную область.

Строки матриц моделей соответствуют значениям факторов, т.е. значениям признаков сердечных заболеваний (градации описательных шкал).

Колонки матриц моделей соответствуют различным классам, отражающим прогнозирование смерти (градации классификационных шкал).

Числовые значения в ячейках матриц моделей, находящихся на пересечении строк и колонок, отражают направление (знак) и силу влияния конкретного значения признаков заболеваний, на конкретное значение его класса.

Если какое-то значение признака слабо влияет на прогнозирование летального исхода, то в соответствующей строке матрицы модели будут малые по модулю значения разных знаков, если же влияние сильное – то и значения будут большие по модулю разных знаков.

Если значение какого-либо признака способствует получению некоторого определенного прогноза, то в соответствующей этому результату ячейке матрицы модели будут положительные значения, если же понижает – то и значения будут отрицательные.

Из этого следует, что суммарную силу влияния того или иного значения признаков заболевания сердца на прогноз (т.е. ценность данного значения признака для решения задачи квалиметрии и других задач) можно количественно оценивать ***степенью вариабельности значений*** в строке матрицы модели, соответствующего этому значению признака болезни.

Существует много мер вариабельности значений: это и среднее модулей отклонения от среднего, и дисперсия, и среднеквадратичное отклонение, и другие. В АСК-анализе и системе «Эйдос» для этой цели принято использовать среднеквадратичное отклонение. Численно оно равно стандартному отклонению и вычисляется по той же формуле, но мы предпочитаем не использовать термин «стандартное отклонение», т.к. он предполагает нормальность распределения исследуемых последовательностей чисел, а значит и проверку соответствующих статистических гипотез.

Самая правая колонка в матрицах моделей на рисунках 7, 8 содержит количественную оценку вариабельности значений строки модели (среднеквадратичное отклонение), которая и представляет собой ценность значения признаков заболевания, соответствующего строке, для решения задачи квалиметрии и других задач, рассмотренных в работе.

Если рассортировать матрицу модели по этой самой правой колонке в порядке убывания, а потом просуммировать значения в ней нарастающим итогом, то получим логистическую Парето-кривую, отражающую зависимость ценности модели от числа наиболее ценных признаков в ней (рисунки 26, 27).

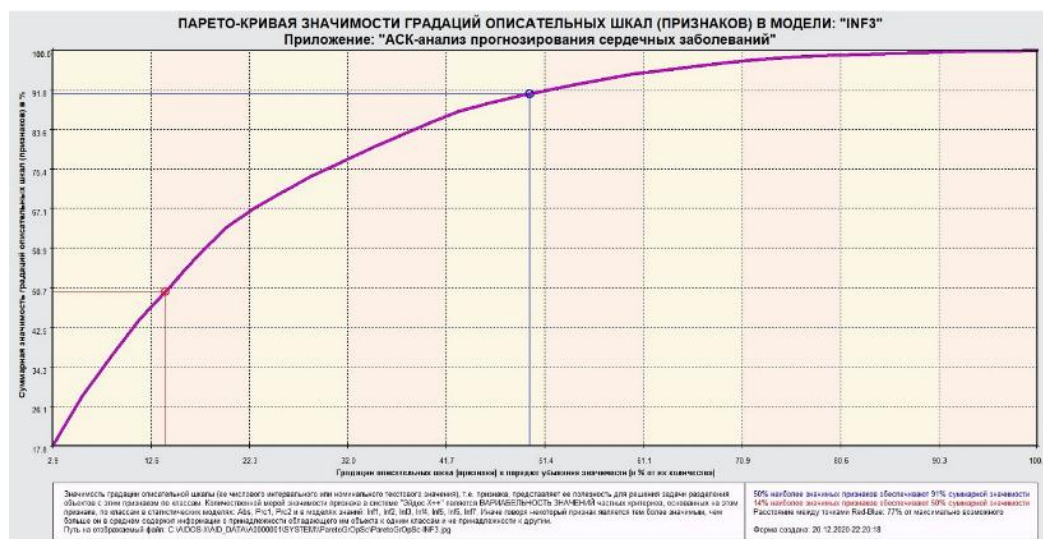


Рисунок 26. Парето-кривая значимости градаций описательных шкал

№	Код	Наименование	Код шкалы	Значимость, %	Значимость, нараст. итогом, %
1	31	TIME-1/5-{4.0000000, 60.2000000}	7	17,8365775	17,8365775
2	11	EJECTION_F-1/5-{14.0000000, 27.2000000}	3	10,0646482	27,9012257
3	29	SERUM_SODI-4/5-{134.0000000, 141.0000000}	6	8,3018335	36,2030592
4	34	TIME-4/5-{172.6000000, 228.8000000}	7	7,8337514	44,0368106
5	12	EJECTION_F-2/5-{27.2000000, 40.4000000}	3	6,5337197	50,5705303
6	28	SERUM_SODI-3/5-{127.0000000, 134.0000000}	6	6,4648320	57,0353623
7	35	TIME-5/5-{228.8000000, 285.0000000}	7	6,0656375	63,1009999
8	21	SERUM_CREA-1/5-{0.5000000, 2.2800000}	5	4,2304024	67,3314023
9	4	AGE-4/5-{73.0000000, 84.0000000}	1	3,3436960	70,6750983
10	32	TIME-2/5-{60.2000000, 116.4000000}	7	3,3030698	73,9781681
11	5	AGE-5/5-{84.0000000, 95.0000000}	1	2,9074081	76,8855762
12	22	SERUM_CREA-2/5-{2.2800000, 4.0600000}	5	2,7766983	79,6622744
13	1	AGE-1/5-{40.0000000, 51.0000000}	1	2,6635522	82,3259266
14	14	EJECTION_F-4/5-{53.6000000, 66.8000000}	3	2,5771010	84,9030277
15	2	AGE-2/5-{51.0000000, 62.0000000}	1	2,3775038	87,2805315
16	17	PLATELETS-2/5-{190080.0000000, 355060.0000000}	4	1,6780302	88,9585616
17	27	SERUM_SODI-2/5-{120.0000000, 127.0000000}	6	1,4537040	90,4122657
18	16	PLATELETS-1/5-{25100.0000000, 190080.0000000}	4	1,2364433	91,6487090
19	3	AGE-3/5-{62.0000000, 73.0000000}	1	1,2099480	92,8586570
20	13	EJECTION_F-3/5-{40.4000000, 53.6000000}	3	1,1428269	94,0014838
21	7	CREATININE-2/5-{1590.6000000, 3158.2000000}	2	1,1092662	95,1107501
22	18	PLATELETS-3/5-{355060.0000000, 520040.0000000}	4	0,7612958	95,8720459
23	10	CREATININE-5/5-{6293.4000000, 7861.0000000}	2	0,7171371	96,5891830
24	25	SERUM_CREA-5/5-{7.6200000, 9.4000000}	5	0,7171371	97,3063200
25	33	TIME-3/5-{116.4000000, 172.6000000}	7	0,6341188	97,9404388
26	23	SERUM_CREA-3/5-{4.0600000, 5.8400000}	5	0,5475677	98,4880666
27	20	PLATELETS-5/5-{685020.0000000, 850000.0000000}	4	0,3391387	98,8271453
28	30	SERUM_SODI-5/5-{141.0000000, 148.0000000}	6	0,1942982	99,0214435
29	9	CREATININE-4/5-{4725.8000000, 6293.4000000}	2	0,1889992	99,2104427
30	15	EJECTION_F-5/5-{66.8000000, 80.0000000}	3	0,1889992	99,3994418
31	24	SERUM_CREA-4/5-{5.8400000, 7.6200000}	5	0,1889992	99,5884410
32	26	SERUM_SODI-1/5-{113.0000000, 120.0000000}	6	0,1889992	99,7774402
33	6	CREATININE-1/5-{23.0000000, 1590.6000000}	2	0,1837002	99,9611404
34	8	CREATININE-3/5-{3158.2000000, 4725.8000000}	2	0,0194298	99,9805702
35	19	PLATELETS-4/5-{520040.0000000, 685020.0000000}	4	0,0194298	100,0000000

Рисунок 27 – Парето-таблица значимости градаций описательных шкал, т.е. сила влияния значений признаков болезни на ее прогноз в СК-модели INF3

Из рисунков 26 и 27 видно, что 50% наиболее ценных для решения задачи прогнозирования летального исхода обуславливают 91% суммарной ценности, а 50% суммарной ценности обеспечиваются 14% наиболее ценных признаков болезни.

Обращаем внимание, что наиболее ценным является наиболее короткий период наблюдения, а наименее ценным – средний уровень тромбоцитов и креатинина.

Из рисунка 27 видно, что наиболее сильное влияние на прогноз летального исхода оказывают значения признаков:

- TIME-1/5-{4.0000000, 60.2000000};
- EJECTION_F-1/5-{14.0000000, 27.2000000};
- SERUM_SODI-4/5-{134.0000000, 141.0000000};
- TIME-4/5-{172.6000000, 228.8000000};
- EJECTION_F-2/5-{27.2000000, 40.4000000}.

А наиболее низкое:

- SERUM_CREA-4/5-{5.8400000, 7.6200000};
- SERUM_SODI-1/5-{113.0000000, 120.0000000};
- CREATININE-1/5-{23.0000000, 1590.6000000};
- CREATININE-3/5-{3158.2000000, 4725.8000000};
- PLATELETS-4/5-{520040.0000000, 685020.0000000}.

Ценность же признаков болезней (всей описательной шкалы или фактора), для решения этих задач можно количественно оценивать, как среднее от ценности значений этого параметра (рисунок 28).

№	Код	Наименование	Значимость, %	Значимость, нараст. итогом, %
1	7	TIME	35,6731550	35,6731550
2	3	EJECTION_F	20,5072950	56,1804500
3	6	SERUM_SODI	16,6036669	72,7841169
4	1	AGE	12,5022080	85,2863249
5	5	SERUM_CREA	8,4608047	93,7471297
6	4	PLATELETS	4,0343378	97,7814675
7	2	CREATININE	2,2185325	100,0000000

Рисунок 28 – Парето-таблица значимости описательных шкал, т.е. сила влияния признаков болезни на прогноз смерти в СК-модели INF3

Из рисунка 28 видно, что наиболее сильное влияние на прогноз смерти оказывают время действия и выброс фракции, а наиболее низкое – уровень креатинина и тромбоцитов.

7. Выводы

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

В результате проделанной работы, с помощью системы «Эйдос» были созданы 3 статистические и 7 системно-когнитивных моделей, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по прогнозу летального исхода, изучено

влияние признаков сердечных болезней на прогнозирование, и, на основе этого, решены задачи идентификации, классификации и исследования моделируемой предметной области путем исследования ее модели.

Список литературы

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): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с. <http://elibrary.ru/item.asp?id=18632909>
2. Луценко Е.В. Универсальный информационный вариационный принцип развития систем / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №07(041). С. 117 – 193. – Шифр Информрегистра: 0420800012\0091, IDA [article ID]: 0410807010. – Режим доступа: <http://ej.kubagro.ru/2008/07/pdf/10.pdf>, 4,812 у.п.л.
3. Lutsenko E. V. System analysis and decision-making (Automated system-cognitive analysis and solving problems of identification, decision-making and research of the simulated subject area): textbook / E. V. Lutsenko. - Krasnodar: ECSC "Eidos", 2020. - 1031 p. // August 2020, DOI: [10.13140/RG.2.2.27247.05289](https://doi.org/10.13140/RG.2.2.27247.05289), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), https://www.researchgate.net/publication/343998862_SYSTEM_ANALYSIS_AND_DECISION_MAKING_Automated_system-cognitive_analysis_and_solving_problems_of_identification_decision-making_and_research_of_the_simulated_subject_area, см. учебный вопрос-2.8.5. Повышение уровня системности объекта управления как цель управления.
4. Lutsenko E.V. On higher forms of consciousness, the prospects of man, technology and society (selected works) // August 2019, DOI: [10.13140/RG.2.2.21336.24320](https://doi.org/10.13140/RG.2.2.21336.24320), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), https://www.researchgate.net/publication/335057548_On_HIGHER_FORMS_of_CONSCIOUSNESS_the_PROSPECTS_of_MAN_TECHNOLOGY_AND_SOCIETY_selected_works
5. Lutsenko E.V. ABOUT THE INTERFACE: "SOUL-COMPUTER» (artificial intelligence: problems and solutions within the system information and functional paradigm of society development) // April 2019, DOI: [10.13140/RG.2.2.23132.85129](https://doi.org/10.13140/RG.2.2.23132.85129), https://www.researchgate.net/publication/332464278_ABOUT_THE_INTERFACE_SOUL-COMPUTER_artificial_intelligence_problems_and_solutions_within_the_system_information_and_functional_paradigm_of_society_development
6. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с. ISBN 978-5-94672-757-0. <http://elibrary.ru/item.asp?id=21358220>
7. Луценко Е.В. Количественные меры возрастания эмерджентности в процессе эволюции систем (в рамках системной теории информации) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2006. – №05(021). С. 355 – 374. – Шифр Информрегистра: 0420600012\0089, IDA [article ID]: 0210605031. – Режим доступа: <http://ej.kubagro.ru/2006/05/pdf/31.pdf>, 1,25 у.п.л.
8. Луценко Е.В. Существование, несуществование и изменение как эмерджентные свойства систем // Квантовая Магия, том 5, вып. 1, стр. 1215-1239, 2008. <http://quantmagic.narod.ru/volumes/VOL512008/p1215.html>
9. 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 (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №03(127). С. 1 – 60. – IDA [article ID]: 1271703001. – Режим доступа: <http://ej.kubagro.ru/2017/03/pdf/01.pdf>, 3,75 у.п.л.

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

11. Страницка Е.В.Луценко: https://www.researchgate.net/profile/Eugene_Lutsenko

12. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 164 – 188. – IDA [article ID]: 0911307012. – Режим доступа: <http://ej.kubagro.ru/2013/07/pdf/12.pdf>, 1,562 у.п.л.

13. ПАМЯТИ ПРОФЕССОРА О.Г. КУКОСЯНА // – Режим доступа: <https://www.elibrary.ru/item.asp?id=20266263>

14. Луценко Е.В. Формирование субъективных (виртуальных) моделей физической и социальной реальности сознанием человека и неоправданное приращение им онтологического статуса (гипостазирование) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №09(113). С. 1 – 32. – IDA [article ID]: 1131509001. – Режим доступа: <http://ej.kubagro.ru/2015/09/pdf/01.pdf>, 2 у.п.л.

15. Луценко Е.В. Принципы и перспективы корректной содержательной интерпретации субъективных (виртуальных) моделей физической и социальной реальности, формируемых сознанием человека / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №01(115). С. 22 – 75. – IDA [article ID]: 1151601003. – Режим доступа: <http://ej.kubagro.ru/2016/01/pdf/03.pdf>, 3,375 у.п.л.

16. Луценко Е.В. Сценарный АСК-анализ как метод разработки на основе эмпирических данных базисных функций и весовых коэффициентов для разложения в ряд функции состояния объекта или ситуации по теореме А.Н.Колмогорова (1957) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №07(161). С. 76 – 120. – IDA [article ID]: 1612007009. – Режим доступа: <http://ej.kubagro.ru/2020/07/pdf/09.pdf>, 2,812 у.п.л.

17. 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 (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 у.п.л.

18. Луценко Е.В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы «Эйдос-Х++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 1367 – 1409. – IDA [article ID]: 1011407090. – Режим доступа: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 у.п.л.

19. 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 (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №07(071). С. 528 – 576. – Шифр Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>, 3,062 у.п.л.