

УДК 4.08

**Автоматизированный системно-когнитивный
анализ стоимости недвижимости**

Пигарев Марк Романович

студент факультета ПИ, группы ИТ2003
mark.pigarev@mail.ru

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

Целью данной работы является изучение влияние характеристик недвижимости на стоимость. Достижение данной цели представляет большой личный интерес. Для нас это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-X++», а также получить зачет. 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».

СОДЕРЖАНИЕ

ЗАДАЧА 1: КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	6
ЗАДАЧА 2: ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ И ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ.....	7
ЗАДАЧА 3: СИНТЕЗ И ВЕРИФИКАЦИЯ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ И ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ ИЗ НИХ ДЛЯ РЕШЕНИЯ ЗАДАЧ	12
ЗАДАЧА 4: РЕШЕНИЕ РАЗЛИЧНЫХ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	17
Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)	17
Подзадача 4.2. Поддержка принятия решений (SWOT-анализ)	19
Подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели	22
4.3.1. Когнитивные диаграммы классов.....	22
4.3.2. Агломеративная когнитивная кластеризация классов.....	23
4.3.3. Агломеративная когнитивная кластеризация значений факторов	25
4.3.4. Нелокальные нейроны и нелокальные нейронные сети.....	27
4.3.5. 3d-интегральные когнитивные карты.....	28
4.3.6. Когнитивные функции	28
7. ВЫВОДЫ.....	31
СПИСОК ЛИТЕРАТУРЫ	32

**Automated system-cognitive
Analysis of property price forecasting**

Mark Pigarev

student of the faculty of PI, group IT2003
mark.pigarev@mail.ru

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

The purpose of this work is to study the impact of real estate characteristics on the value . The achievement of this goal is of great personal interest. For us, this will allow us to gain knowledge in working with the universal analytical system "Eidos-X ++", as well as get a credit. To achieve this goal, the Automated System-Cognitive Analysis (ASK-analysis) and its software tools are used - the intelligent system "Eidos".

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

Введение

Целью данной работы является изучение влияние характеристик недвижимости на ее стоимость.

Достижение данной цели представляет большой личный интерес. Для нас это позволит получить знания в работе с универсальной когнитивной аналитической системой «Эйдос-Х++», а также получить зачет.

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-когнитивный анализа (АСК-анализ), 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-Х++» (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 автоматического, т.е. без непосредственного участия человека в реальном времени решения задач идентификации, прогнозирования, принятия решений и исследования 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 227, 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 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;

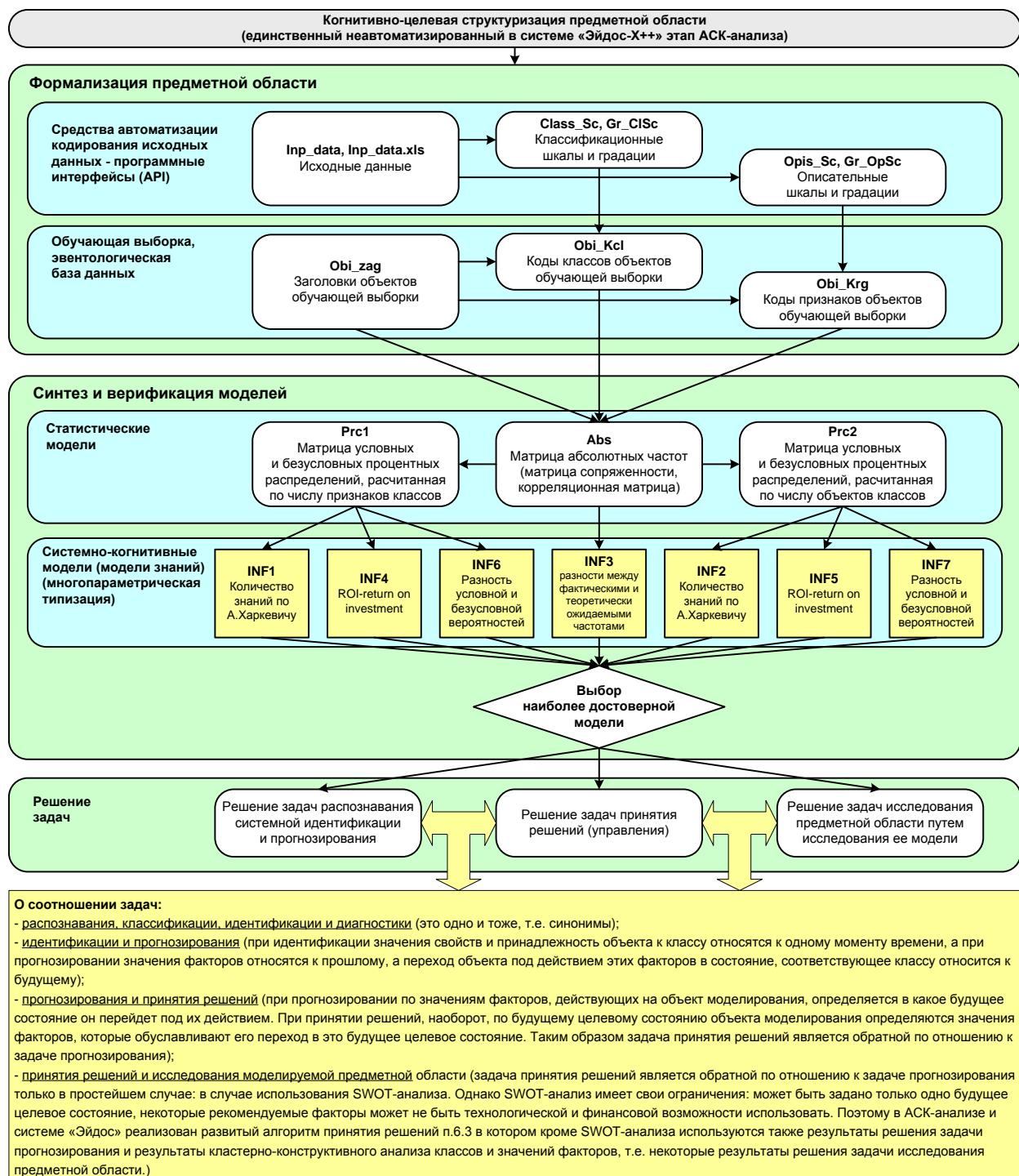
- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) 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 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 неформализуем 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 являются 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 категорию Medv (таблица 1), а в качестве факторов, влияющих на эти результаты – Crim, Zn, Indus, chas, nox, rm, age, dis, rad, tax, ptratio, b, lstat(таблица 2):

Таблица 1 – Классификационная шкала

Код	Наименование
1	Medv

Таблица 2 – Описательные шкалы

Код	Наименование
1	CRIM
2	ZN
3	INDUS
4	CHAS
5	NOX
6	RM
7	AGE
8	DIS
9	RAD
10	TAX
11	PTRATIO
12	B
13	LSTAT

Задача 2: подготовка исходных данных и формализация предметной области

Исходные данные для данной статьи (рисунок 2) подробных таблиц по недвижимости, найденных в интернете.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Obj	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	MEDV
2	1	0,00632	18,0	2,31	0	0,538	6,575	65,2	4,09	1	296	15,3	396,90	4,98	24,0
3	2	0,02731	0,0	7,07	0	0,469	6,421	78,9	4,9671	2	242	17,8	396,90	9,14	21,6
4	3	0,02729	0,0	7,07	0	0,469	7,185	61,1	4,9671	2	242	17,8	392,83	4,03	34,7
5	4	0,03237	0,0	2,18	0	0,458	6,998	45,8	6,0622	3	222	18,7	394,63	2,94	33,4
6	5	0,06905	0,0	2,18	0	0,458	7,147	54,2	6,0622	3	222	18,7	396,90	5,33	36,2
7	6	0,02985	0,0	2,18	0	0,458	6,430	58,7	6,0622	3	222	18,7	394,12	5,21	28,7
8	7	0,08829	12,5	7,87	0	0,524	6,012	66,6	5,5605	5	311	15,2	395,60	12,43	22,9
9	8	0,14455	12,5	7,87	0	0,524	6,172	96,1	5,9505	5	311	15,2	396,90	19,15	27,1
10	9	0,21124	12,5	7,87	0	0,524	5,631	100,0	6,0821	5	311	15,2	386,63	29,93	16,5
11	10	0,17004	12,5	7,87	0	0,524	6,004	85,9	6,5921	5	311	15,2	386,71	17,10	18,9
12	11	0,22489	12,5	7,87	0	0,524	0,000	94,3	6,3467	5	311	15,2	392,52	20,45	15,0
13	12	0,11747	12,5	7,87	0	0,524	6,009	82,9	6,2267	5	311	15,2	396,90	13,27	18,9
14	13	0,09378	12,5	7,87	0	0,524	5,889	39,0	5,4509	5	311	15,2	390,50	15,71	21,7
15	14	0,62976	0,0	8,14	0	0,538	5,949	61,8	4,7075	4	307	21,0	396,90	8,26	20,4
16	15	0,63796	0,0	8,14	0	0,538	6,096	84,5	4,4619	4	307	21,0	380,02	10,26	18,2
17	16	0,62739	0,0	8,14	0	0,538	5,834	56,5	4,4986	4	307	21,0	395,62	8,47	19,9
18	17	1,05393	0,0	8,14	0	0,538	5,935	29,3	4,4986	4	307	21,0	386,85	6,58	23,1
19	18	0,78420	0,0	8,14	0	0,538	5,990	81,7	4,2579	4	307	21,0	386,75	14,67	17,5
20	19	0,80271	0,0	8,14	0	0,538	5,456	36,6	3,7965	4	307	21,0	288,99	11,69	20,2
21	20	0,72580	0,0	8,14	0	0,538	5,727	69,5	3,7965	4	307	21,0	390,95	11,28	18,2
22	21	1,25179	0,0	8,14	0	0,538	5,570	98,1	3,7979	4	307	21,0	376,57	21,02	13,6
23	22	0,85204	0,0	8,14	0	0,538	5,965	89,2	4,0123	4	307	21,0	392,53	13,83	19,6
24	23	1,23247	0,0	8,14	0	0,538	6,142	91,7	3,9769	4	307	21,0	396,90	18,72	15,2
25	24	0,98843	0,0	8,14	0	0,538	5,813	100,0	4,0952	4	307	21,0	394,54	19,88	14,5
26	25	0,75026	0,0	8,14	0	0,538	5,924	94,1	4,3996	4	307	21,0	394,33	16,30	15,6
27	26	0,84054	0,0	8,14	0	0,538	5,599	85,7	4,4546	4	307	21,0	303,42	16,51	13,9
28	27	0,67191	0,0	8,14	0	0,538	5,813	90,3	4,682	4	307	21,0	376,88	14,81	16,6
29	28	0,95577	0,0	8,14	0	0,538	6,047	88,8	4,4534	4	307	21,0	306,38	17,28	14,8
30	29	0,77299	0,0	8,14	0	0,538	6,495	94,4	4,4547	4	307	21,0	387,94	12,80	18,4
31	30	1,00245	0,0	8,14	0	0,538	6,674	87,3	4,239	4	307	21,0	380,23	11,98	21,0
32	31	1,13081	0,0	8,14	0	0,538	5,713	94,1	4,233	4	307	21,0	360,17	22,60	12,7
33	32	1,35472	0,0	8,14	0	0,538	6,072	100,0	4,175	4	307	21,0	376,73	13,04	14,5
34	33	1,38799	0,0	8,14	0	0,538	5,950	82,0	3,99	4	307	21,0	232,60	27,71	13,2
35	34	1,15172	0,0	8,14	0	0,538	5,701	95,0	2,7872	4	307	21,0	252,77	12,25	12,1

Рисунок 2 – Часть исходных данных 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.

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

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

Задайте тип файла исходных данных: "Inp_data":

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

Стандарт XLS-файла
Стандарт DBF-файла
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")

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

- Не применять сценарный метод ACK-анализа
- Применять специал-интерпретацию текстовых полей классов
- Применять сценарный метод ACK-анализа
- Применять специал-интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

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

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

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

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

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

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [5 x 65]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных шкал	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	1	5	5.00	13	65	5.00
Текстовые	0	0	0.00	0	0	0.00
ВСЕГО:	1	5	5.00	13	65	5.00

Задайте число интервалов (градаций) в шкале:

В классификационных шкалах: В описательных шкалах:

Пересчитать шкалы и градации Выйти на создание модели

2.3.2.2. Процесс импорта данных из внешней БД "Inp_data" в систему "ЭЙДОС-Х++"

Стадии исполнения процесса

- 1/3: Формирование классификационных и описательных шкал и градаций на основе БД "Inp_data"- Готово
- 2/3: Генерация обучающей выборки и базы событий "EventsKO" на основе БД "Inp_data"- Готово
- 3/3: Переиндексация всех баз данных нового приложения-Готово

ПРОЦЕСС ФОРМАЛИЗАЦИИ ПРЕДМЕТНОЙ ОБЛАСТИ ЗАВЕРШЕН УСПЕШНО !!!

Прогноз времени исполнения

Начало: 12:42:00	Окончание: 12:42:08
Прошло: 0:00:08	Осталось: 0:00:00

100% Ok

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

Обратим внимание на то, что заданы адаптивные интервалы, учитывающие 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 3 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) 5 (таблица 3) и 13 описательных шкал с суммарным числом градаций 60 (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

Код	Название
1	1/5-{5.0000000, 17.4000000}
2	2/5-{17.4000000, 29.8000000}
3	3/5-{29.8000000, 42.2000000}
4	4/5-{42.2000000, 54.6000000}
5	5/5-{54.6000000, 67.0000000}

Таблица 4 – Описательные шкалы и градации

1	CRIM-1/5-{0.0063200,17.8002960}
2	CRIM-2/5-{17.8002960,35.5942720}
3	CRIM-3/5-{35.5942720,53.3882480}
4	CRIM-4/5-{53.3882480,71.1822240}
5	CRIM-5/5-{71.1822240,88.9762000}
6	ZN-1/5-{12.5000000,30.0000000}
7	ZN-2/5-{30.0000000,47.5000000}
8	ZN-3/5-{47.5000000,65.0000000}
9	ZN-4/5-{65.0000000,82.5000000}
10	ZN-5/5-{82.5000000,100.0000000}
11	INDUS-1/5-{0.4600000,5.9160000}
12	INDUS-2/5-{5.9160000,11.3720000}
13	INDUS-3/5-{11.3720000,16.8280000}
14	INDUS-4/5-{16.8280000,22.2840000}
15	INDUS-5/5-{22.2840000,27.7400000}
16	NOX-1/5-{0.3850000,0.4822000}
17	NOX-2/5-{0.4822000,0.5794000}
18	NOX-3/5-{0.5794000,0.6766000}
19	NOX-4/5-{0.6766000,0.7738000}
20	NOX-5/5-{0.7738000,0.8710000}
21	RM-1/5-{3.5610000,4.6048000}
22	RM-2/5-{4.6048000,5.6486000}
23	RM-3/5-{5.6486000,6.6924000}
24	RM-4/5-{6.6924000,7.7362000}
25	RM-5/5-{7.7362000,8.7800000}
26	AGE-1/5-{2.9000000,22.3200000}
27	AGE-2/5-{22.3200000,41.7400000}
28	AGE-3/5-{41.7400000,61.1600000}
29	AGE-4/5-{61.1600000,80.5800000}
30	AGE-5/5-{80.5800000,100.0000000}
31	DIS-1/5-{1.1296000,3.3289800}
32	DIS-2/5-{3.3289800,5.5283600}
33	DIS-3/5-{5.5283600,7.7277400}

34	DIS-4/5-{7.7277400,9.9271200}
35	DIS-5/5-{9.9271200,12.1265000}
36	RAD-1/5-{1.0000000,5.6000000}
37	RAD-2/5-{5.6000000,10.2000000}
38	RAD-3/5-{10.2000000,14.8000000}
39	RAD-4/5-{14.8000000,19.4000000}
40	RAD-5/5-{19.4000000,24.0000000}
41	TAX-1/5-{187.0000000,291.8000000}
42	TAX-2/5-{291.8000000,396.6000000}
43	TAX-3/5-{396.6000000,501.4000000}
44	TAX-4/5-{501.4000000,606.2000000}
45	TAX-5/5-{606.2000000,711.0000000}
46	PTRATIO-1/5-{12.6000000,14.6800000}
47	PTRATIO-2/5-{14.6800000,16.7600000}
48	PTRATIO-3/5-{16.7600000,18.8400000}
49	PTRATIO-4/5-{18.8400000,20.9200000}
50	PTRATIO-5/5-{20.9200000,23.0000000}
51	B-1/5-{0.3200000,79.6360000}
52	B-2/5-{79.6360000,158.9520000}
53	B-3/5-{158.9520000,238.2680000}
54	B-4/5-{238.2680000,317.5840000}
55	B-5/5-{317.5840000,396.9000000}
56	LSTAT-1/5-{1.7300000,16.5840000}
57	LSTAT-2/5-{16.5840000,31.4380000}
58	LSTAT-3/5-{31.4380000,46.2920000}
59	LSTAT-4/5-{46.2920000,61.1460000}
60	LSTAT-5/5-{61.1460000,76.0000000}

2.3.1. Ручной ввод: корректировка обучающей выборки. Текущая модель: "INF1".

Код объекта	Наименование объекта	Дата	Время
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7	
1	2	0	0	0	1	36	42	47	55	56	0	32

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) моделей осуществляется в режиме 3.5 системы «Эйдос» (рисунок 5).

3.5. Выбор моделей для синтеза и верификации

Задайте модели для синтеза и верификации:

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 Хиквадрат, разности между фактическими и ожидаемыми abs.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 безузл.вероятностей; вероятности из PRC1
- 10.INF7 - частный критерий: разн.усл.и безузл.вероятностей; вероятности из PRC2

Текущая модель:

- 1. ABS
- 2. PRC1
- 3. PRC2
- 4. INF1
- 5. INF2
- 6. INF3
- 7. INF4
- 8. INF5
- 9. INF6
- 10.INF7

Параметры копирования обучающей выборки в распознаваемую:

Какие объекты обуч.выборки копировать:

- Копировать всю обучающую выборку
- Копировать только текущий объект
- Копировать каждый N-й объект
- Копировать N случайных объектов
- Копировать все объекты от N1 до N2
- Вообще не менять распознаваемую выборку

Пояснение по алгоритму верификации:

Удалять из обуч.выборки скопированные объекты:

- Не удалять
- Удалять

Подробнее:

Измеряется внутренняя достоверн. модели

Для каждой заданной модели выполнить:

- Синтез и верификацию
- Только верификацию
- Только синтез

На каком процессоре выполнять расчеты:

- CPU
- GPU

Уменьшение размеров базы данных результатов распознавания: Rasp.dbf
Расчетный размер БД результатов распознавания Rasp.dbf равен 127950 байт, т.е.: 0.0059581 % от MAX-возможного, [от 2Гб]
Задайте, сколько % от исходной БД Rasp.dbf оставить, удаляя наименее достоверные результаты распознавания: 100.0000000

Ok Cancel

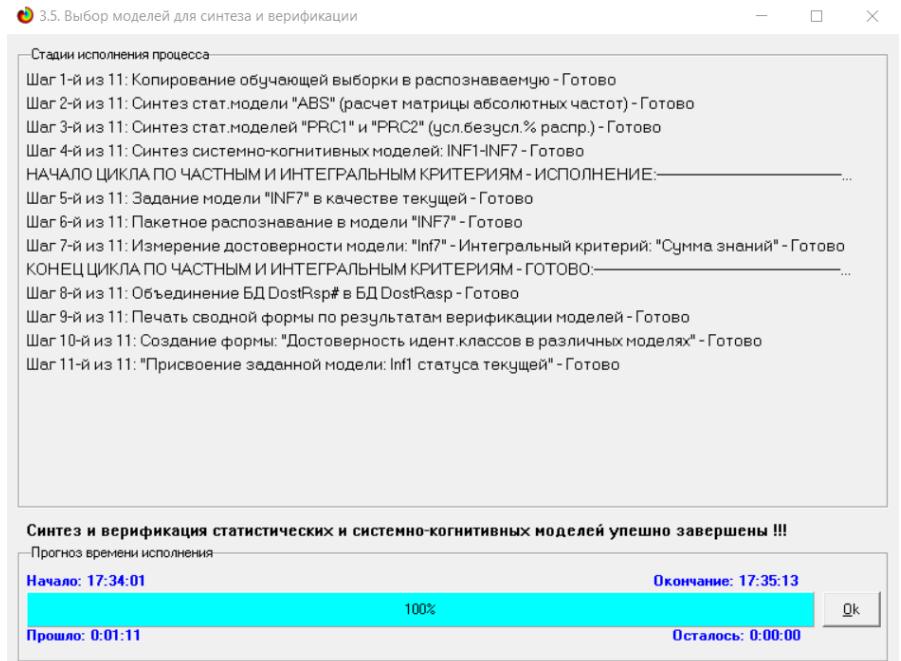


Рисунок 5. Экранные формы режима синтеза и верификации статистических и системно-когнитивных моделей системы «Эйдос»

Обратим внимание на то, что на рисунке 5 в правом нижнем углу окна задана опция: «Расчеты проводить на графическом процессоре (GPU)».

Из рисунка 5 видно, что весь процесс синтеза и верификации моделей занял 1 минуту 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 6, 7:

5.5. Модель: "1. ABS - частный критерий: количество встреч сочетаний: "Класс-признак" у объектов обучавшись"									
Код признака	Наименование описательной шкалы и градации	1. MEDV 1/5 (5.000000, 17.400000)	2. MEDV 2/5 (17.400000, 29.800000)	3. MEDV 3/5 (29.800000, 42.200000)	4. MEDV 4/5 (42.200000, 54.600000)	5. MEDV 5/5 (54.600000, 67.000000)	Сумма	Среднее	Средн квад откл
1	CRIM-1/5-(0.0063200, 17.8002960)	116	284	57	30	1	488	97.60	112.51
2	CRIM-2/5-(17.8002960, 35.5942720)	14	1				15	3.00	6.16
3	CRIM-3/5-(35.5942720, 53.3882480)	5					5	1.00	2.24
4	CRIM4-4/5-(53.3882480, 71.1822240)	1					1	0.20	0.45
5	CRIM5/5-(71.1822240, 88.9762000)	2					2	0.40	0.89
6	ZN-1/5-(12.5000000, 30.0000000)	4	44	10	8	66	13.20	17.64	
7	ZN-2/5-(30.0000000, 47.5000000)	1	9	13		23	4.60	6.02	
8	ZN-3/5-(47.5000000, 65.0000000)		8	2		10	2.00	3.46	
9	ZN-4/5-(65.0000000, 82.9000000)	15	6	2		23	4.60	6.31	
10	ZN-5/5-(82.9000000, 100.0000000)	2	6	4		12	2.40	2.61	
11	INDUS-1/5-(0.4600000, 5.9160000)	3	91	41	16	151	30.20	37.64	
12	INDUS-2/5-(5.9160000, 11.3720000)	21	104	14	4	143	28.60	42.95	
13	INDUS-3/5-(11.3720000, 16.8280000)	2	23	1	1	28	5.60	9.74	
14	INDUS-4/5-(16.8280000, 22.3840000)	106	61	1	9	177	35.40	46.83	
15	INDUS-5/5-(22.3840000, 27.7400000)	6	6			12	2.40	3.29	
16	NOX-1/5-(0.3850000, 0.4822000)	6	104	35	10	155	31.00	42.93	
17	NOX-2/5-(0.4822000, 0.5794000)	19	111	15	8	154	30.80	45.36	
18	NOX-3/5-(0.5794000, 0.6766000)	41	48	7	12	108	21.60	21.48	
19	NOX-4/5-(0.6766000, 0.7738000)	61	17			78	15.60	26.43	
20	NOX-5/5-(0.7738000, 0.8710000)	11	5			16	3.20	4.87	
21	RM-1/5-(3.5610000, 4.5040000)	4	2			6	1.20	1.79	
22	RM-2/5-(4.5040000, 5.6486000)	39	19		1	59	11.80	17.22	
23	RM-3/5-(5.6486000, 6.6924000)	84	232	5	4	326	65.20	99.58	

Рисунок 6. Матрица абсолютных частот (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. MEDV	2. MEDV	3. MEDV	4. MEDV	5. MEDV	Сумма	Среднее	Станд. квадр. откл.
1	CRIM-1/5-0.0056200, 17.8002960	-13.171	11.654	0.647	0.805	0.066	0.000	0.809	
2	CRIM-2/5-17.0002960, 35.5942720	10.930	-7.371	-1.732	-0.897	-0.029	0.000	6.300	
3	CRIM-3/5-05.5942720, 53.3892490	3.677	-2.790	-0.577	-0.299	-0.010		2.331	
4	CRIM-4/5-53.3892490, 71.1622240	0.735	-0.558	-0.115	-0.060	-0.002		0.466	
5	CRIM-5/5-71.1622240, 88.5762000	1.471	-1.116	-0.231	-0.120	-0.004	0.000	0.932	
6	ZN-1/5-12.5000000, 30.0000000	-13.470	7.166	2.378	4.051	-0.126	0.000	7.983	
7	ZN-2/5-30.0000000, 47.5000000	-5.080	-3.836	10.344	-1.376	-0.044		6.113	
8	ZN-3/5-47.5000000, 65.0000000	-2.647	2.419	0.845	-0.598	-0.019		1.866	
9	ZN-4/5-65.0000000, 82.5000000	-6.088	2.164	3.344	0.624	-0.044	0.000	3.651	
10	ZN-5/5-82.5000000, 100.0000000	-3.176	-4.697	4.614	3.282	-0.023	0.000	4.007	
11	INDUS-1/5-0.4600000, 59.1600000	-36.969	6.729	23.563	6.966	-0.289	0.000	22.449	
12	INDUS-2/5-0.5916000, 11.3720000	-16.851	24.194	-2.513	-4.555	-0.274		14.970	
13	INDUS-3/5-11.3720000, 16.8280000	-5.411	7.374	-2.233	-0.675	0.946	0.000	4.743	
14	INDUS-4/5-16.8280000, 22.2840000	59.149	-37.781	-19.440	-1.589	-0.339	0.000	36.423	
15	INDUS-5/5-22.2840000, 27.7400000	2.824	-0.697	-1.386	-0.718	-0.023	0.000	1.650	
16	NOX-1/5-0.3850000, 0.4820000	-35.028	17.497	17.101	0.727	-0.297	0.000	21.367	
17	NOX-2/5-0.4820000, 0.5794000	-21.763	25.055	-2.784	-1.213	0.705	0.000	0.000	16.666
18	NOX-3/5-0.5794000, 0.5760000	12.413	-12.273	-5.472	5.539	-0.207	0.000	9.557	
19	NOX-4/5-0.6760000, 0.7738000	40.354	-26.531	-9.007	-4.666	-0.149	0.000	0.000	24.674
20	NOX-5/5-0.7738000, 0.9710000	6.765	-3.929	-1.848	-0.957	-0.031		4.048	
21	RM-1/5-3.5619000, 4.6548000	2.412	-1.349	-0.693	-0.359	-0.011	0.000	1.436	
22	RM-2/5-4.6040000, 5.6480000	23.383	-13.927	-6.813	-2.530	-0.113	0.000	14.085	
23	RM-3/5-5.6480000, 6.6324000	-2.291	50.064	-32.646	-15.503	0.376	0.000	0.000	30.894

Рисунок 7. 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 обобщенным образом классов и подсчета количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1-L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры. В режиме 3.4 системы «Эйдос» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности (рисунок 8).

Название модели и критерий	Интегральный критерий	очень	удовл.	Полож. модели	Г-факт	Все	Рейтинг	Сумма под. зонами по критериям (ST)	Сумма под. зонами по критериям (ST)	Средн. под. зонами по критериям (ST)	Сумма под. зонами по критериям (ST)	Б-Годность	С-Годность	Л-Годность	Средн. под. зонами по критериям (ST)	Средн. под. зонами по критериям (ST)
1. ABS - частный критерий: количество встреч соединит. "Клас-1"	Корреляция abs. частот с обр.	0,352	0,947	-0,516	239,017	56,693	234,855	1,441	0,504	0,994	0,669	0,454	0,	0,454	0,	0,
1. ABS - частный критерий: количество встреч соединит. "Клас-1"	Сумма abs. частот по признакам	0,270	1,000	-0,425	235,150	385,224	0,390	0,390	1,000	0,551	0,440	0,	0,	0,	0,	0,
2. PRCT - частный критерий: из вероятности его применения	Корреляция abs. частот с обр.	0,352	0,947	-0,516	239,017	56,493	234,888	1,441	0,504	0,994	0,669	0,454	0,	0,454	0,	0,
2. PRCT - частный критерий: из вероятности его применения	Сумма abs. частот по признакам	0,270	1,000	-0,425	232,668	376,423	0,382	0,382	1,000	0,552	0,455	0,	0,	0,	0,	0,
3. PRCT2 - частный критерий: из вероятности его применения	Корреляция abs. частот с обр.	0,352	0,947	-0,516	239,013	56,492	234,882	1,441	0,504	0,994	0,669	0,454	0,	0,454	0,	0,
3. PRCT2 - частный критерий: извероятность его применения	Сумма abs. частот по признакам	0,270	1,000	-0,425	236,900	390,080	0,380	0,380	1,000	0,551	0,449	0,	0,	0,	0,	0,
4. INF1 - частный критерий: количество знаний по Альбому-1	Семантический разноск. ана.	0,417	0,885	-0,567	232,453	244,152	157,070	0,443	0,557	0,941	0,709	0,514	0,	0,	0,	0,
4. INF1 - частный критерий: количество знаний по Альбому-1	Сумма знаний	0,480	0,822	-0,407	109,593	271,856	57,341	15,127	0,496	0,579	0,703	0,259	0,	0,	0,	0,
5. INF2 - частный критерий: количество знаний по Альбому-2	Семантический разноск. ана.	0,417	0,885	-0,567	234,381	246,307	166,713	14,581	0,557	0,941	0,709	0,518	0,	0,	0,	0,
5. INF2 - частный критерий: количество знаний по Альбому-2	Сумма знаний	0,480	0,822	-0,407	109,593	271,856	57,341	15,127	0,496	0,579	0,703	0,259	0,	0,	0,	0,
6. INF3 - частный критерий Ханчадзе: различия между частотами	Семантический разноск. ана.	0,411	0,890	-0,562	203,466	180,451	154,082	9,547	0,568	0,955	0,715	0,448	0,	0,	0,	0,
6. INF3 - частный критерий Ханчадзе: различия между частотами	Сумма знаний	0,411	0,890	-0,562	120,145	97,076	82,527	4,331	0,568	0,946	0,783	0,212	0,	0,	0,	0,
7. INF4 - частный критерий ROI (Return On Investment): вероятность	Семантический разноск. ана.	0,434	0,873	-0,585	196,488	179,945	141,729	10,687	0,581	0,948	0,723	0,441	0,	0,	0,	0,
7. INF4 - частный критерий ROI (Return On Investment): вероятность	Сумма знаний	0,403	0,898	-0,556	137,160	80,171	56,316	4,187	0,616	0,971	0,794	0,302	0,	0,	0,	0,
8. INF5 - частный критерий ROI (Return On Investment): вероятность	Семантический разноск. ана.	0,435	0,875	-0,593	195,992	180,564	142,057	10,724	0,581	0,948	0,723	0,441	0,	0,	0,	0,
8. INF5 - частный критерий ROI (Return On Investment): вероятность	Сумма знаний	0,403	0,898	-0,556	144,184	83,244	60,777	4,333	0,614	0,971	0,795	0,314	0,	0,	0,	0,
9. INF6 - частный критерий разницы и безразличия вероятности	Семантический разноск. ана.	0,417	0,875	-0,564	205,264	155,849	145,514	10,932	0,574	0,948	0,715	0,448	0,	0,	0,	0,
9. INF6 - частный критерий разницы и безразличия вероятности	Сумма знаний	0,398	0,890	-0,589	122,644	70,431	74,727	4,262	0,623	0,966	0,797	0,270	0,	0,	0,	0,
10. INF7 - частный критерий разницы и безразличия вероятности	Семантический разноск. ана.	0,414	0,875	-0,562	200,926	155,580	145,236	10,971	0,574	0,948	0,715	0,449	0,	0,	0,	0,
10. INF7 - частный критерий разница и безразличия вероятности	Сумма знаний	0,397	0,888	-0,549	127,057	73,103	78,025	4,401	0,620	0,967	0,795	0,280	0,	0,	0,	0,

Рисунок 8. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергена и L1-критерию проф. Е.В.Луценко

Из рисунка 8 мы видим, что в данном интеллектуальном приложении по F-критерию Ван Ризбергена наиболее достоверной является СК-модель INF3 с интегральным критерием «Сумма знаний» ($F=0,562$ при максимуме 1,000), что является довольно хорошим результатом, по критерию L1 проф. Е.В.Луценко та же модель является наиболее достоверной ($L1=0,763$ при максимуме 1,000), что тоже является хорошим результатом.

Это подтверждает наличие и адекватное отражение в СК-модели INF3 сильной причинно-следственной зависимости между недвижимостью и ее стоимостью.

На рисунке 9 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам идентификации характеристик недвижимости в СК-модели INF3 по данным обучающей выборки:

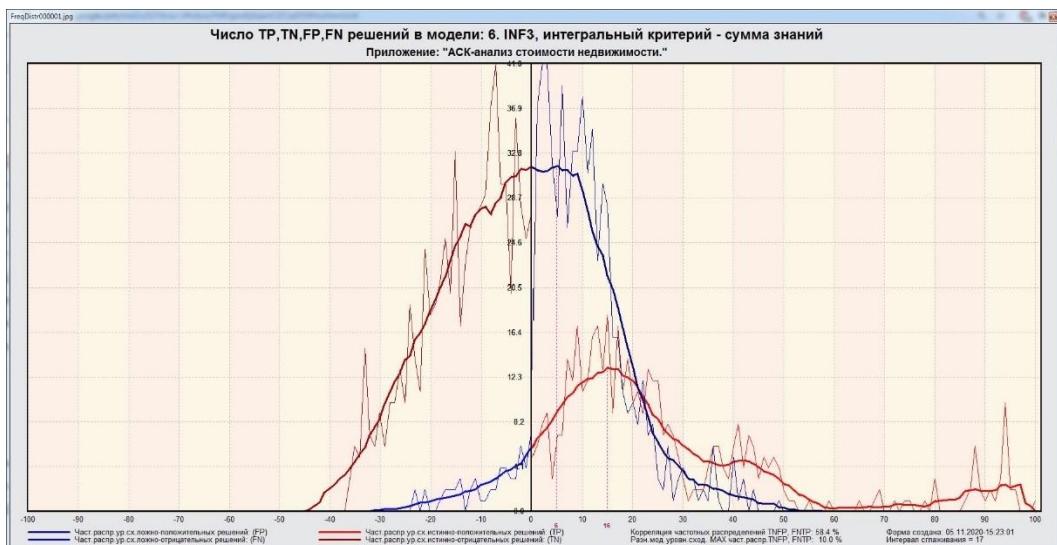


Рисунок 9. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf3

Рисунок 9 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе.

Левое распределение, большее по амплитуде включает истинно-отрицательные и ложно-положительные решения, а правое, меньшее по амплитуде, включает ложные отрицательные и истинно-положительные решения.

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу идентификации недвижимости по ее характеристикам и другие задачи.

Видно, что для отрицательных решений количество истинных решений всегда значительно превосходит количество ложных решений, причем при уровнях различия больше примерно 5% ложные отрицательные решения вообще практически отсутствуют.

Видно также, что для положительных решений картина более сложная и включает 3 диапазона уровней сходства

1) при уровнях сходства от 0% до 30% количество ложных решений больше числа истинных;

2) при уровнях сходства от 30% до примерно 42% есть и истинные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства;

3) при уровнях сходства выше 42% ложные положительные решения не встречаются.

На рисунке 10 приведен Help по режиму 3.4, в котором описаны меры достоверности моделей, применяемые в системе «Эйдос»:

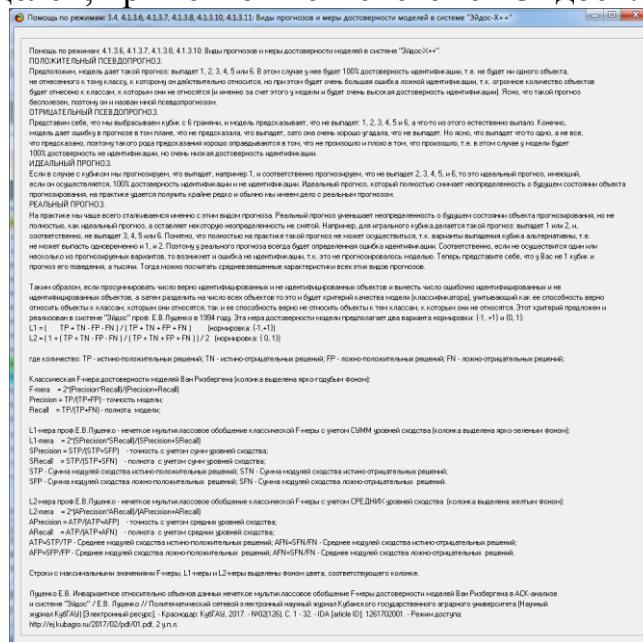


Рисунок 10. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергена и L1- и L2-критериям проф. Е.В.Луценко

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 11):

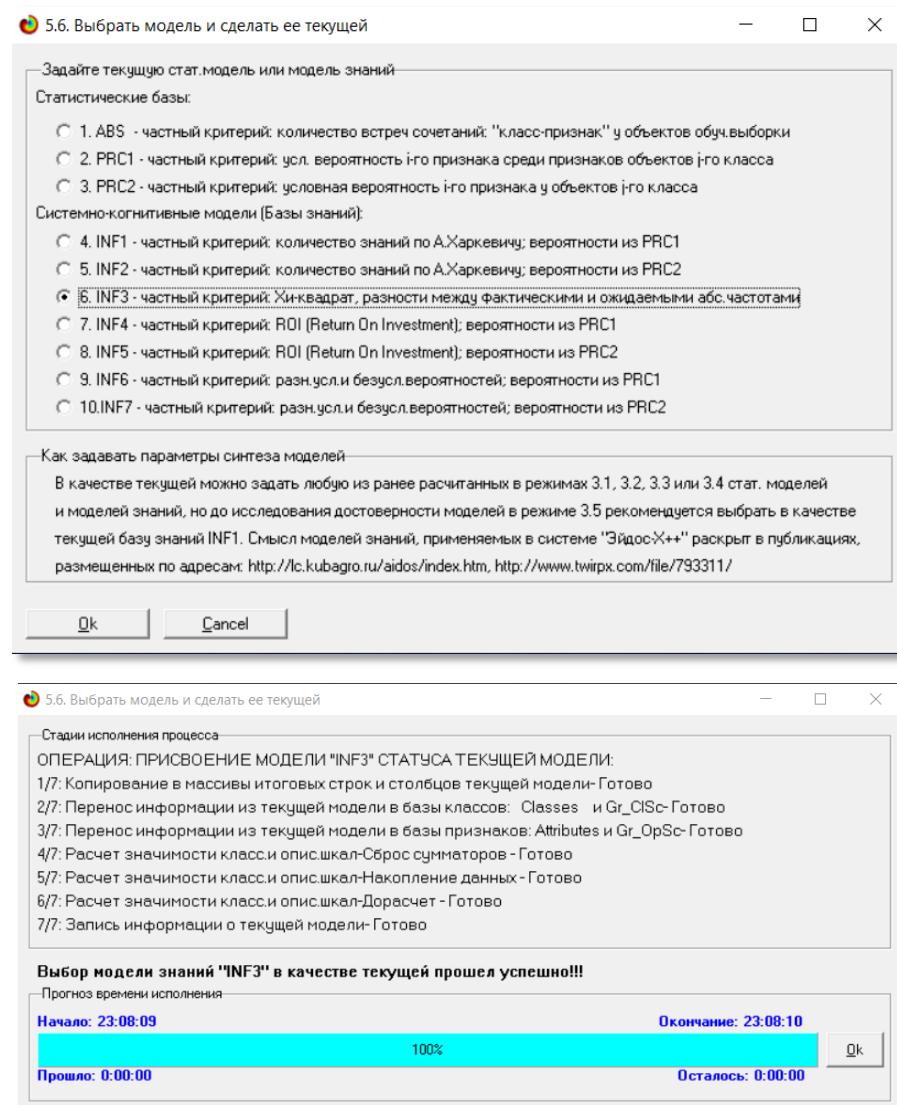


Рисунок 11. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

Задача 4: решение различных задач в наиболее достоверной модели

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

Решим задачу системной идентификации, т.е. определение класса недвижимости на основе обучающей выборки в наиболее достоверной СК-модели INF3 на GPU (рисунок 12).

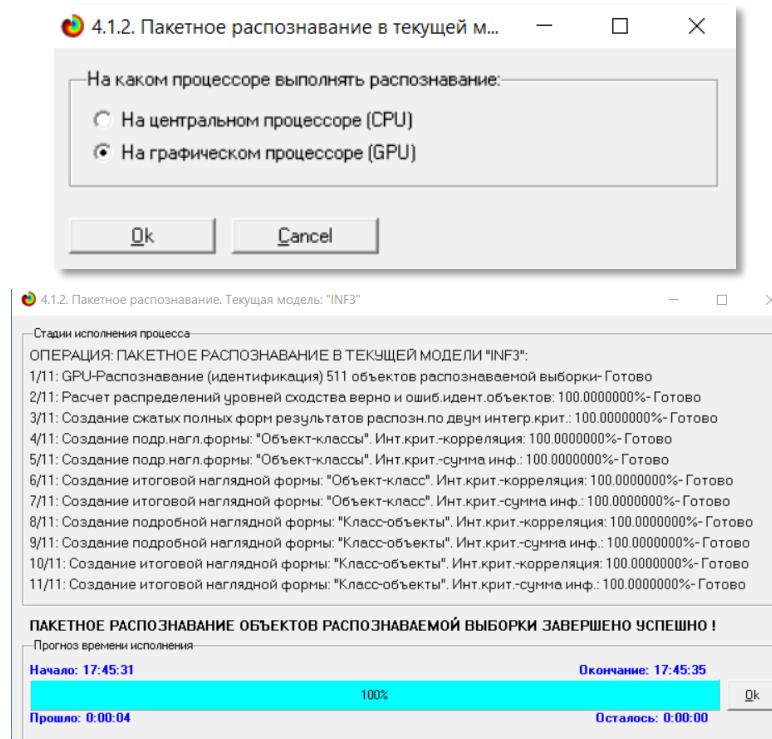
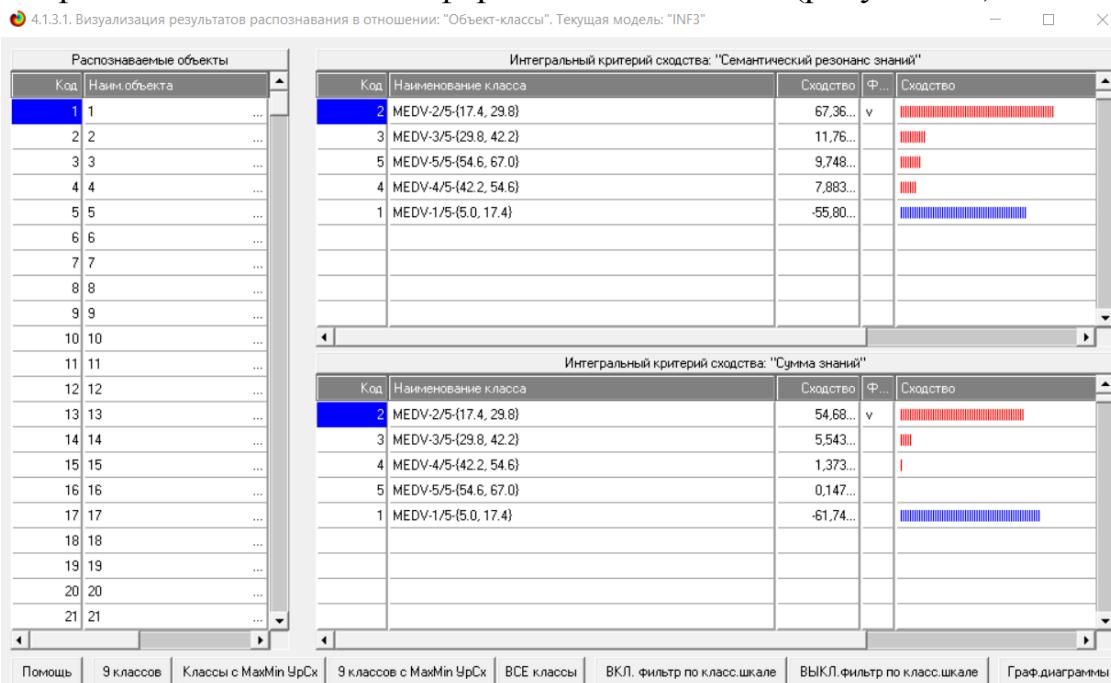


Рисунок 12. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Из рисунка 12 видно, что процесс идентификации занял 4 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 13).



The screenshot shows a software window with three tables. The first table on the left is titled 'Классы' (Classes) and lists five entries: MEDV-1/5-{5.000000, 17.400000}, MEDV-2/5-{17.400000, 29.800000}, MEDV-3/5-{29.800000, 42.200000}, MEDV-4/5-{42.200000, 54.600000}, and MEDV-5/5-{54.600000, 67.000000}. The second table in the middle is titled 'Интегральный критерий сходства: "Семантический резонанс знаний"' (Integral criterion of similarity: "Semantic resonance of knowledge") and lists objects 418, 416, 385, 419, 414, 387, 404, 426, and 430 with similarity scores ranging from 84.72% to 93.42%. The third table at the bottom is titled 'Интегральный критерий сходства: "Сумма знаний"' (Integral criterion of similarity: "Sum of knowledge") and lists the same objects with similarity scores ranging from 90.69% to 100.00%. The bottom of the window contains standard menu items: Помощь (Help), Поиск объекта (Search object), В начало БД (To the beginning of the database), В конец БД (To the end of the database), Предыдущая (Previous), Следующая (Next), 9 записей (9 records), Все записи (All records), Печать XLS (Print XLS), Печать TXT (Print TXT), and Печать ALL (Print ALL).

Рисунок 13. Выходные формы по результатам идентификации недвижимости по ее характеристикам

Символ « \vee » стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка 13 видно, что результаты идентификации являются отличными, естественно при учете информации из рисунка 9 о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 42%, 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 характеристик недвижимости на ее стоимость.

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 14 приведены SWOT-диаграммы наглядно отражающие силу и направление влияния различных значений характеристик недвижимости на ее стоимость.

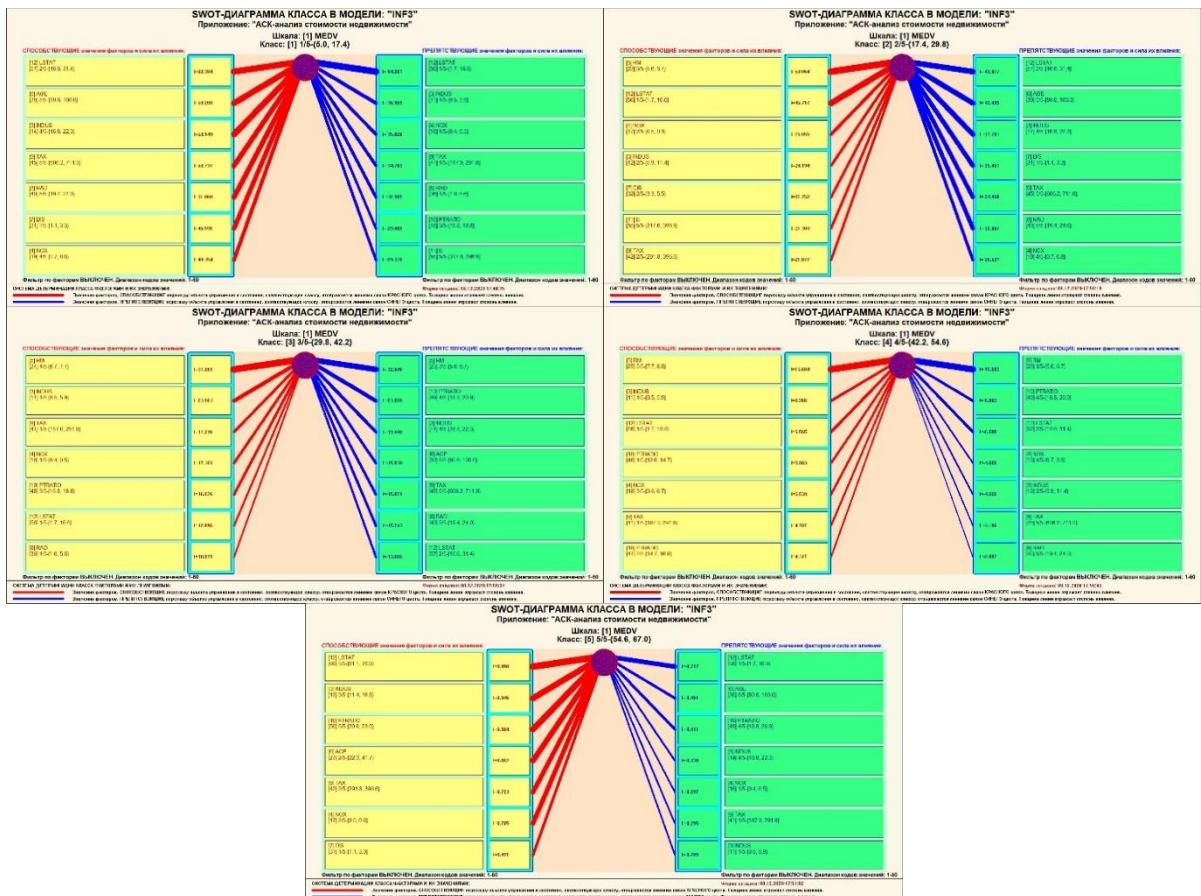


Рисунок 14. SWOT-диаграммы детерминации характеристик недвижимости

Эти SWOT-диаграммы наглядно отражают силу и направление влияния различных значений характеристик недвижимости на ее стоимость.

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.

<p>У Т В Е Р Ж Д А Ъ</p> <p>Заведующий Краснодарским сектором ИСИ АН СССР, к.ф.н. <i>А.Хагуров</i> 1987г.</p>	<p>У Т В Е Р Ж Д А Ъ</p> <p>Директор Северо-Кавказского филиала ВНИЦ "АИУС-агроресурсы", к.э.н. <i>З.М.Трахов</i> 1987г.</p>
--	---

А К Т

Настоящий акт составлен комиссией в составе: Кириченко М.М., Ляшко Г.А., Самсонов Г.А., Коренец В.И., Луценко Е.В. в том, что в соответствии с договором о научно-техническом сотрудничестве между Северо-Кавказским филиалом ВНИЦ "АИУС-агроресурсы" и Краснодарским сектором Института социологических исследований АН СССР Северо-Кавказским филиалом ВНИЦ "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
Руслан М.М. Кириченко
19.05.1987г.

Мл. научный сотрудник
Ляшко Г.А. Ляшко
19.05.1987г.

От СКФ ВНИЦ "АИУС-агроресурсы":

Зав. отделом аэрокосмических и
тематических изысканий №4, к.э.н.
Самсонов Г.А. Самсонов
19.05.1987г.

Главный конструктор проекта
Борисов В.И. Коренец
19.05.1987г.

Главный конструктор проекта
Луценко Е.В. Луценко
19.05.1987г.

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 14, 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 имеющаяся 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.

В заключение отметим, что SWOT-анализ является широко известным и общепризнанным метод 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 неформализуемым 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 уже более 30 лет, но, к сожалению, она сравнительно малоизвестна – это интеллектуальная система «Эйдос».

Подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели

Если модель предметной области достоверна, то исследование модели можно считать исследованием самого моделируемого объекта, т.е. результаты исследования модели корректно относить к самому объекту моделирования, «переносить на него».

В системе «Эйдос» есть довольно много возможностей для такого исследования, но в данной работе из-за ограничений на ее объем мы рассмотрим лишь результаты кластерно-конструктивного анализа классов и признаков (когнитивные диаграммы и дендрограммы), а также нелокальные нейроны, нелокальные нейронные сети, 3d-интегральные когнитивные карты и когнитивные функции.

4.3.1. Когнитивные диаграммы классов

Эти диаграммы отражают сходство/различие классов. Мы получаем их в режимах 4.2.2.1 и 4.2.2.2 (рисунок 16).

Отметим также, что на когнитивной диаграмме, приведенной на рисунке 16, показаны **количественные** оценки сходства/различия недвижимости по связанным с ними значениям их характеристик. Важно, что эти результаты сравнения получены с применением системно-когнитивной модели, созданной *непосредственно на основе эмпирических данных*, а не как традиционно делается на основе экспертных оценок неформализуемым путем на основе опыта, интуиции и профессиональной компетенции.

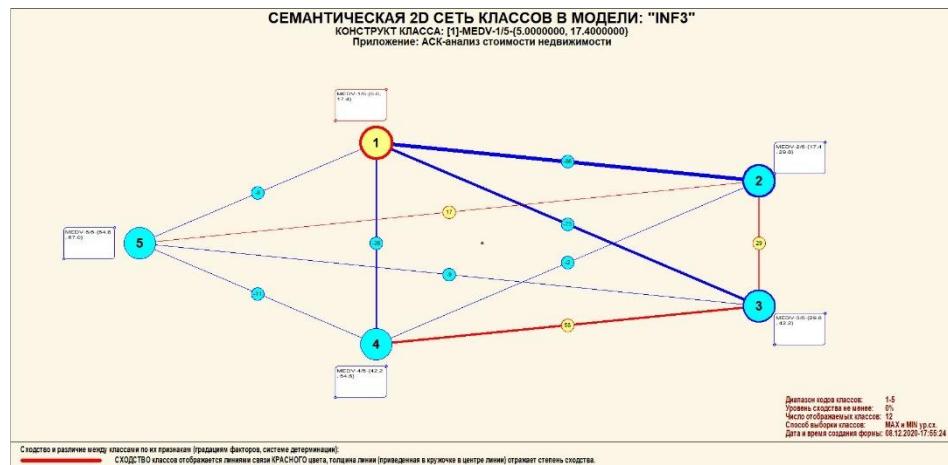


Рисунок 15. Когнитивная диаграмма классов, отражающая сходства/различия недвижимости по связанным с ними значениям их характеристик.

В системе «Эйдос» есть возможность управлять параметрами формирования и вывода изображения, приведенного на рисунке 16. Для этого используется диалоговое окно, приведенное на рисунке 17.

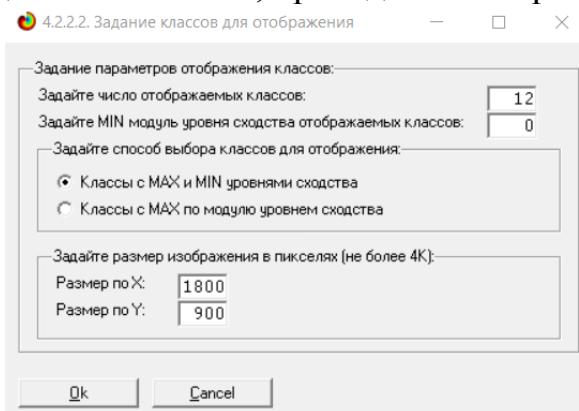


Рисунок 16. Диалоговое окно управления параметрами формирования и вывода изображения когнитивной диаграммы классов

4.3.2. Агломеративная когнитивная кластеризация классов

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме, когнитивных диаграмм, пример которой приведен на рисунке 16, но и в форме агломеративных дендрограмм, полученных в результате *когнитивной кластеризации* (рисунок 18):

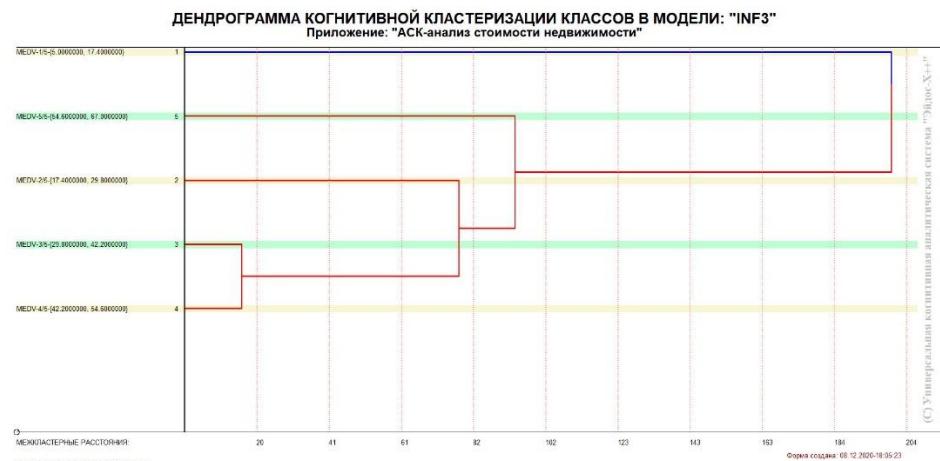


Рисунок 17. Дендрограмма когнитивной агломеративной кластеризации, отражающая сходство/различие недвижимости

Из рисунков 16 и 18 мы видим, что некоторые виды недвижимости сходны по детерминирующей их системе значений характеристик, и, следовательно, корректно ставить задачу их одновременного достижения, а другие по системе значений этих параметров сильно отличаются, и, следовательно, являются взаимоисключающими, т.е. альтернативными и цель их одновременного достижения является некорректной и недостижимой, т.к. для достижения одного из альтернативных результатов необходимы одни значения характеристик, а для достижения другого – совершенно другие, которые не могут наблюдаться одновременно с первыми.

Из дендрограммы когнитивной агломеративной кластеризации классов, приведенной на рисунке 18, мы видим также, что все классы образуют два противоположных кластера, являющихся полюсами конструкта, по системе значений обуславливающих значениям параметров их характеристик.

На рисунке 19 мы видим график изменения межкластерных расстояний:

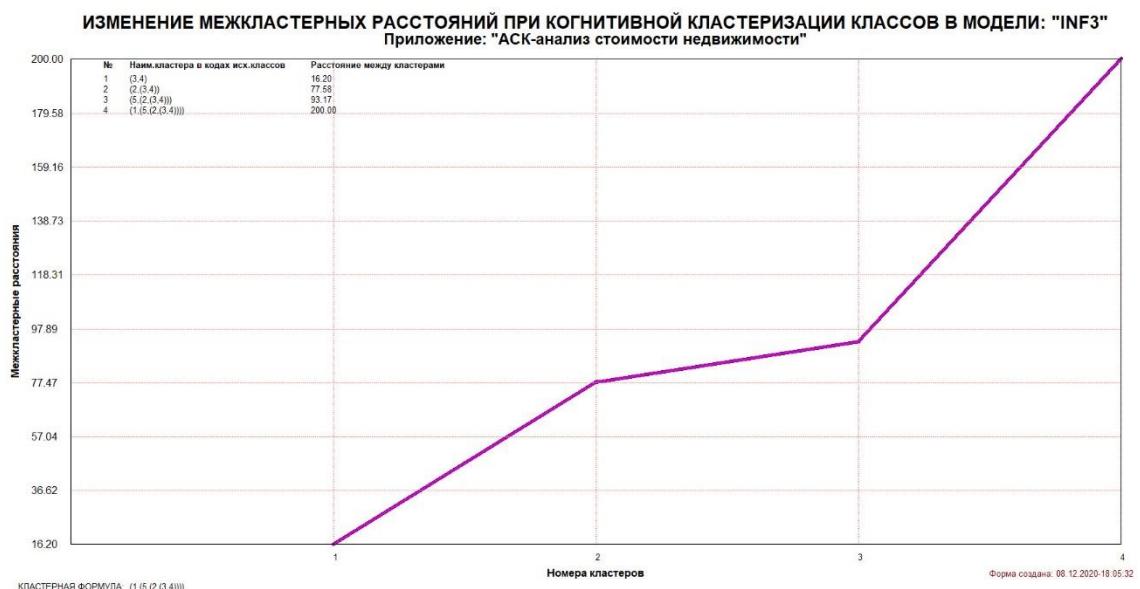


Рисунок 18. График изменения межкластерных расстояний

4.3.3. Агломеративная когнитивная кластеризация значений факторов

На рисунке 22 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 20.

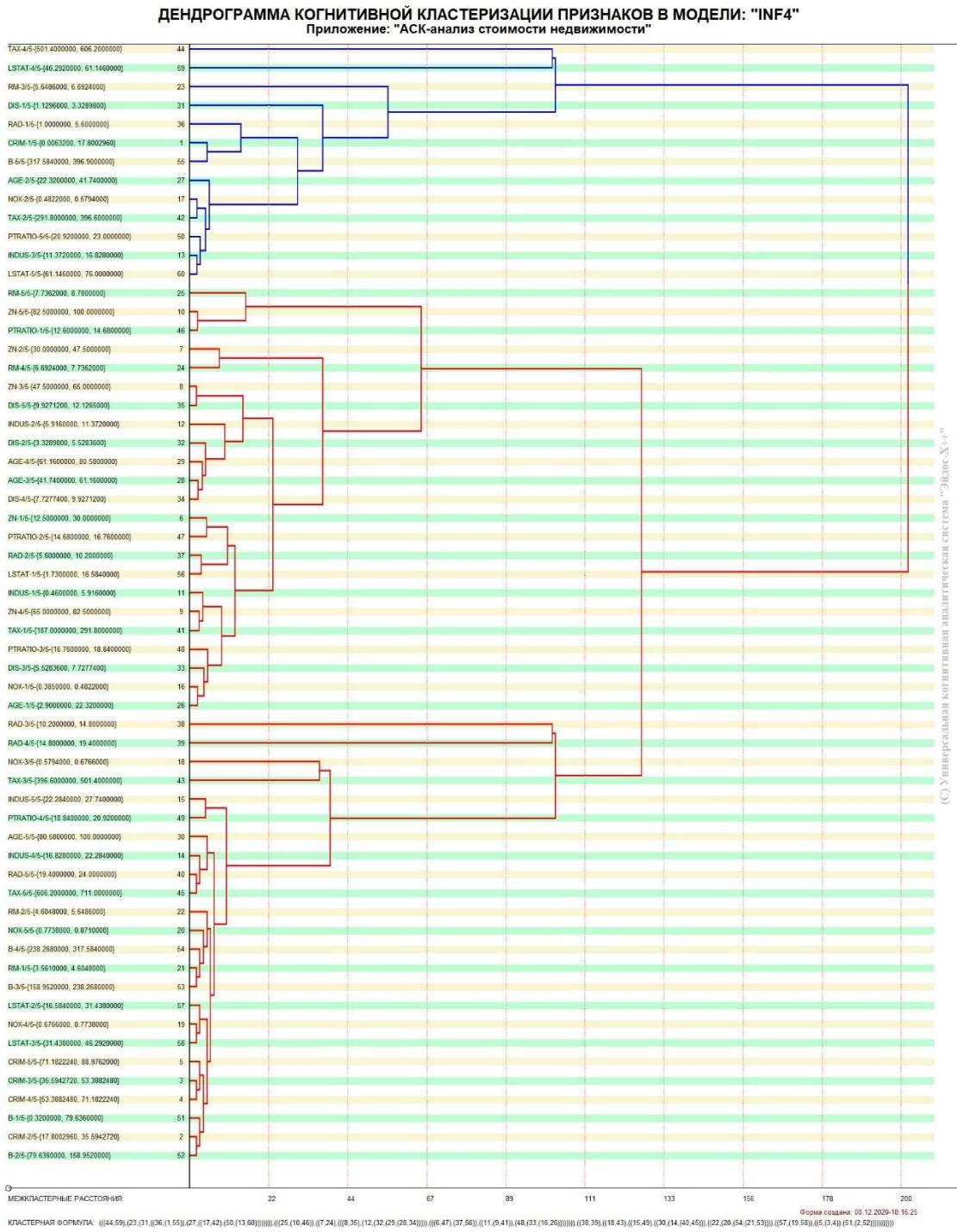


Рисунок 9. Дендрограмма агломеративной когнитивной кластеризации признаков

Из дендрограммы на рисунке 22 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 недвижимости. *Значения факторов на полюсах конструкта факторов (рисунок 22) обуславливают переход объекта моделирования в состояния, соответствующие классам, представленным на полюсах конструкта классов (рисунки 18 и 20).*

На рисунке 23 приведен график межкластерных расстояний значений признаков.

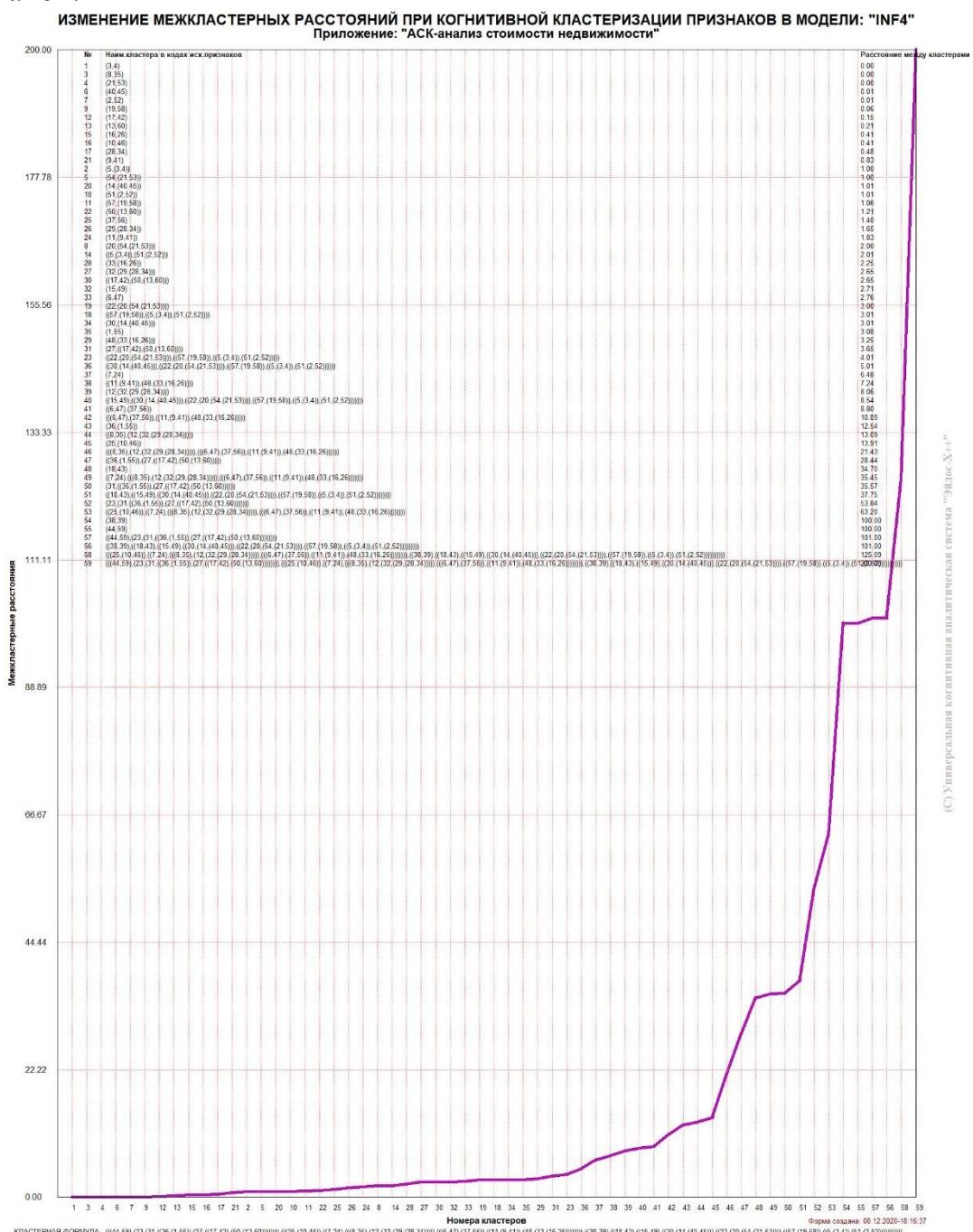


Рисунок 10. График изменения межкластерных расстояний при когнитивной кластеризации значений факторов

4.3.4. Нелокальные нейроны и нелокальные нейронные сети

На рисунке 24 приведён пример нелокального нейрона, а на рисунке 25 – фрагмент одного слоя нелокальной нейронной сети:

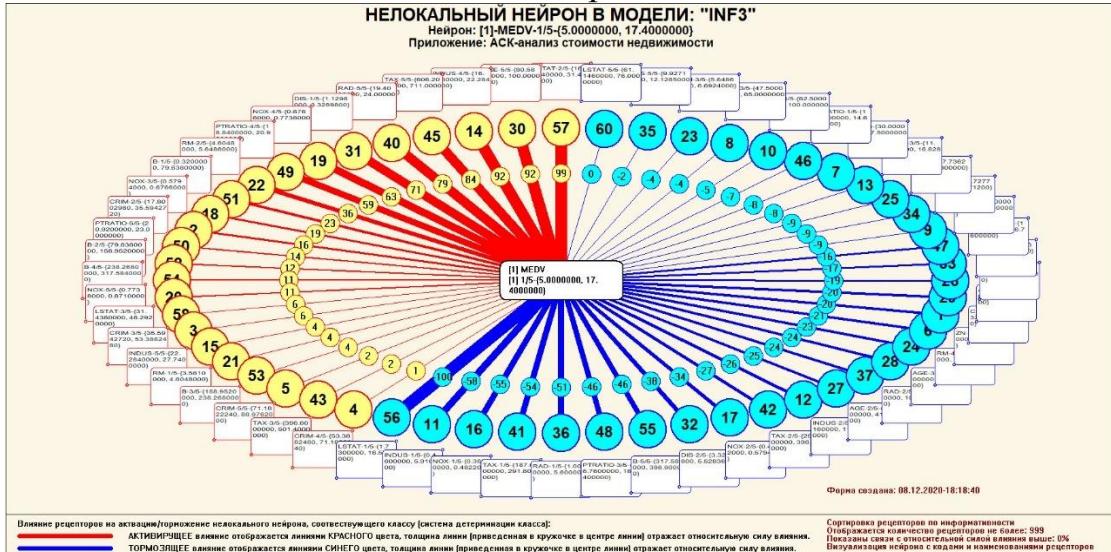


Рисунок 11. Пример нелокального нейрона, отражающего силу и направление влияния значений характеристик недвижимости на ее стоимость

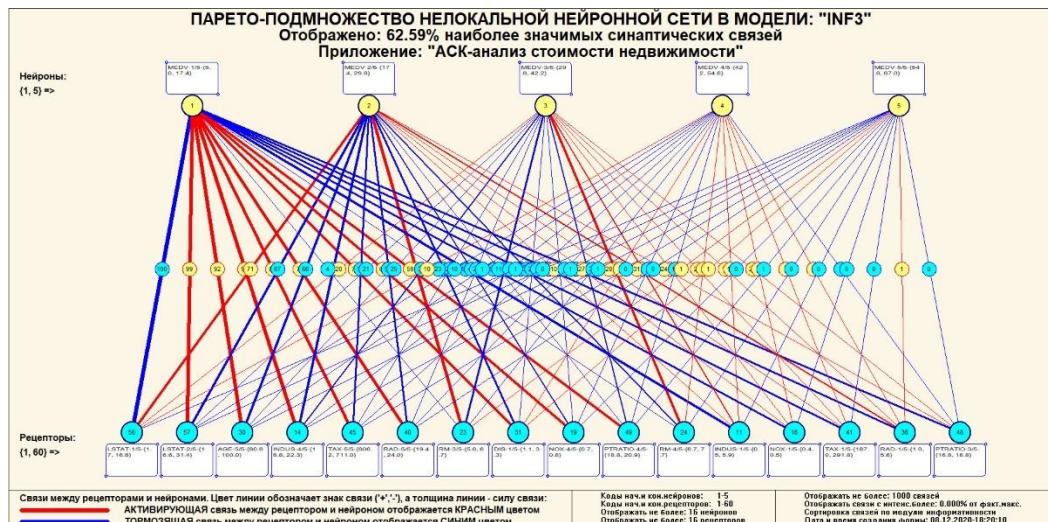


Рисунок 12. Один слой нелокальной нейронной сети, отражающий силу и направление влияния значений характеристик недвижимости на ее стоимость (фрагмент 62% сети)

В приведенном фрагменте слоя нейронной сети нейроны соответствуют недвижимости, а рецепторы – их характеристикам. 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 «полно связной».

4.3.5. 3d-интегральные когнитивные карты

На рисунке 26 приведен фрагмент 3d-интегральной когнитивной карты, отражающий фрагмент около 62% СК-модели Inf3.

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивных диаграмм классов и значений факторов, отображеных соответственно на рисунках 16 и 20, и одного слоя нейронной сети, приведенного на рисунке 25.

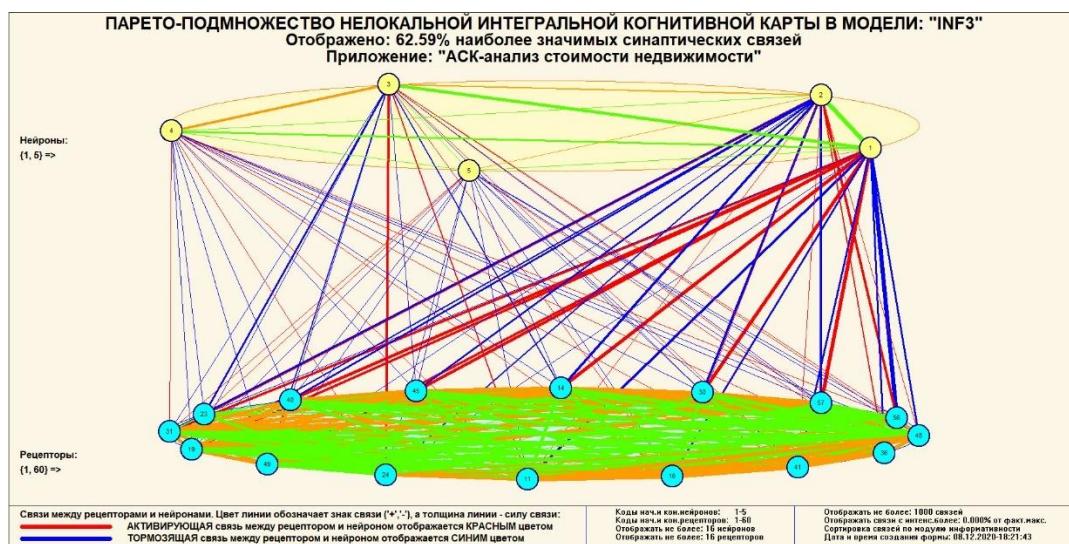


Рисунок 13. 3d-интегральная когнитивная карта в СК-модели Inf3

4.3.6. Когнитивные функции

Вместо описания того, что представляют собой когнитивные функции, приведем help соответствующего режима системы «Эйдос» (рисунок 27).

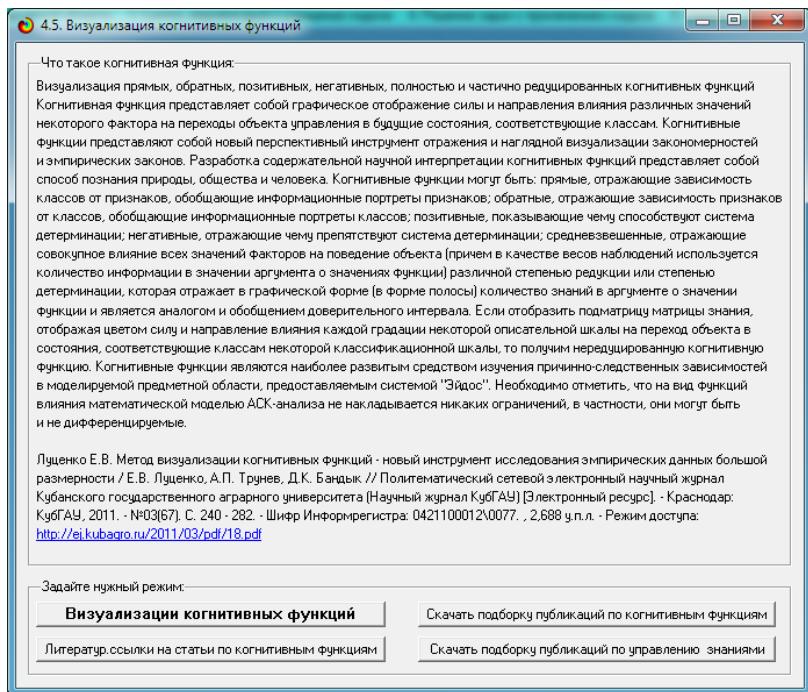


Рисунок 14. Help режима визуализации когнитивных функций

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.

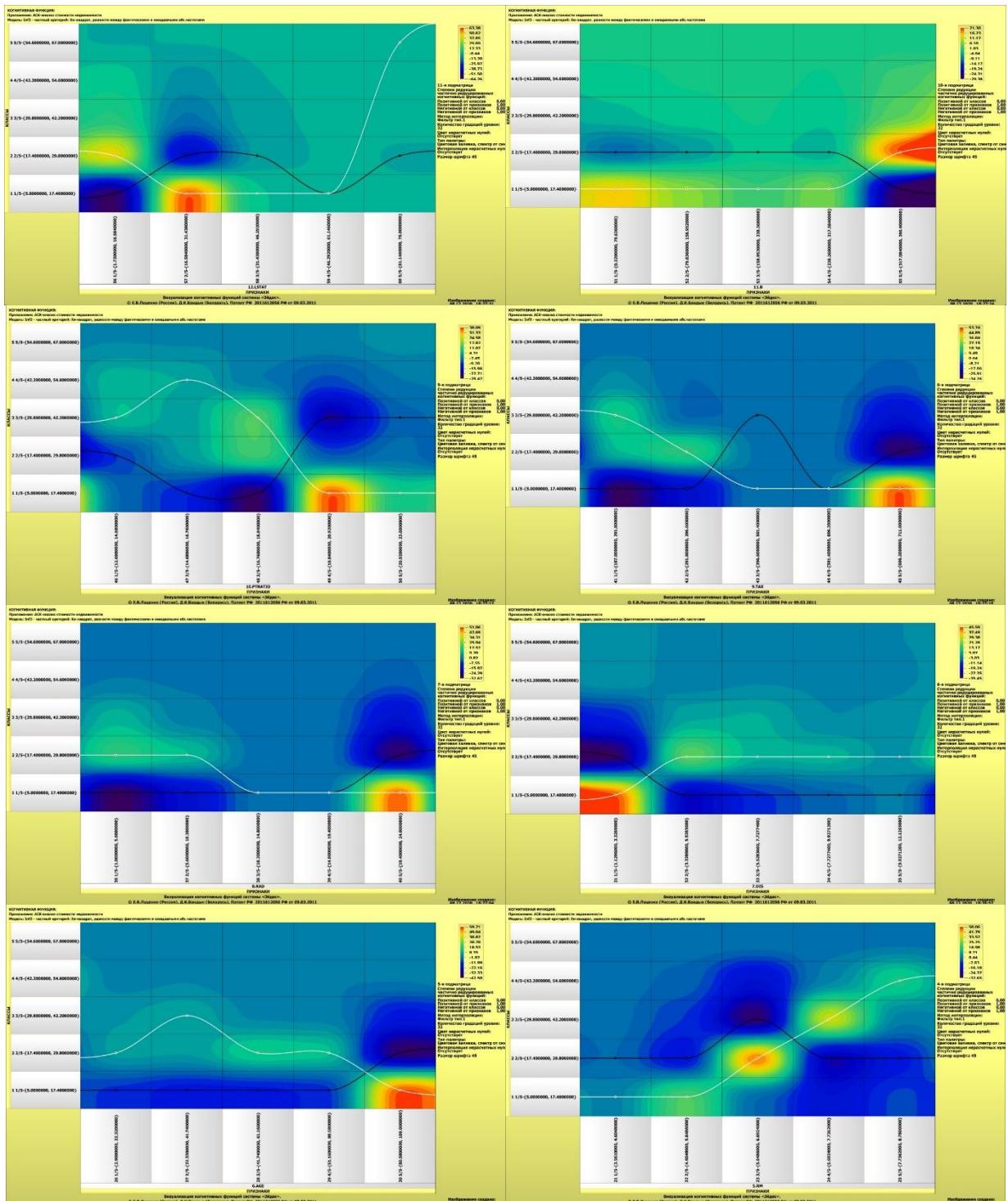
Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют системы детерминации (обозначены белой линией); негативные, отражающие чему препятствуют системы 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 28 приведены когнитивные функции, наглядно отражающие силу и направление влияния значений (т.е. степени выраженности) различных характеристик недвижимости на ее стоимость .



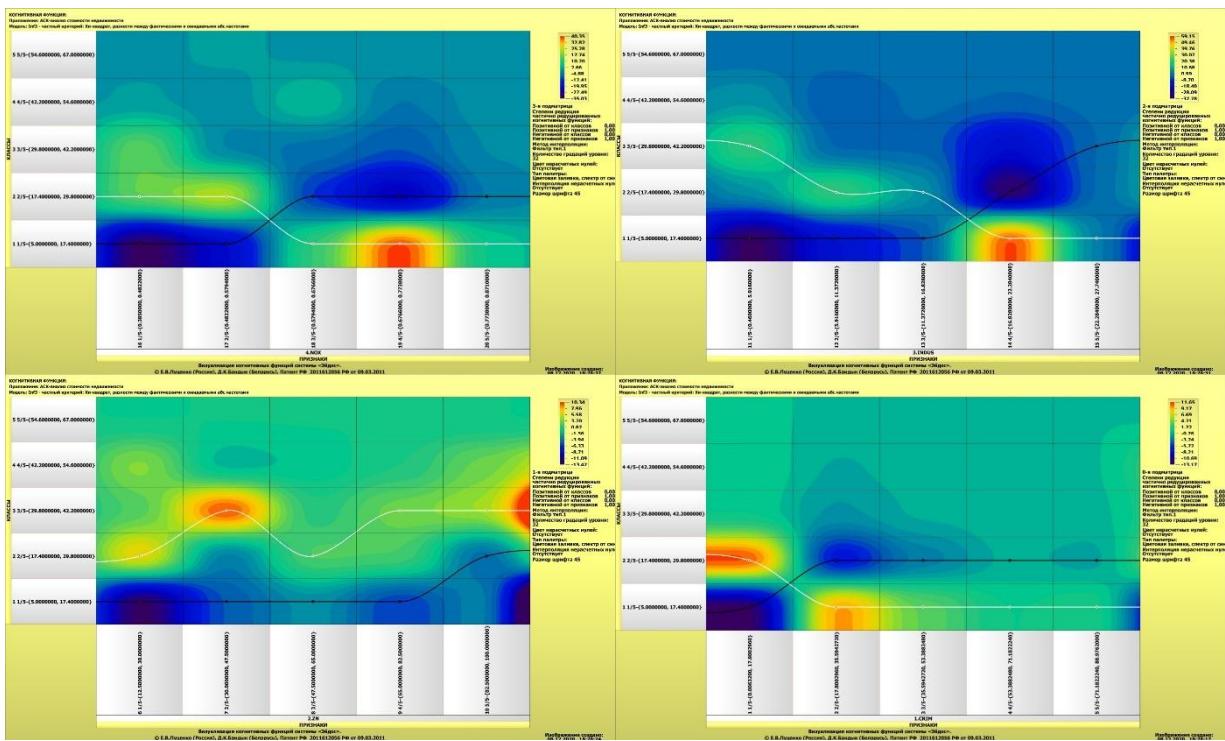


Рисунок 15. Примеры когнитивных функций в СК-модели INF3, отражающих силу и направление влияния значений характеристик недвижимости на ее стоимость

Из когнитивных функций, приведенных на рисунке 28, хорошо видно, что **зависимости между значениями характеристик недвижимости и ее стоимостью в основном имеют довольно предсказуемый характер**.

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 6 системно-когнитивных моделей, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по различным классам недвижимости, изучено влияние характеристик недвижимости на стоимость, и, на основе этого, решены задачи идентификации, классификации и исследования моделируемой предметной области путем исследования ее модели.

Со всеми моделями, созданными в данной статье, можно ознакомиться установив облачное Эйдос-приложение №228 в режиме 1.3 системы «Эйдос».

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 // Политехнический 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](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](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](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 // Политехнический 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 // Политехнический 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. Страница 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 приятие 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.