

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ  
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ  
ИМЕНИ И.Т. ТРУБИЛИНА»

Факультет прикладной информатики

Кафедра компьютерных технологий и систем


КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ зарплат разработчиков ПО в США»

Выполнил студент группы: ИТ2141 Кутумов Кирилл Сергеевич

Допущен к защите \_\_\_\_\_

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

Защищен \_\_\_\_\_

Оценка \_\_\_\_\_

Краснодар

2023

## РЕФЕРАТ

Курсовая работа содержит: 33 страниц, 35 рисунков, 10 литературных источников.

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

Цель данной курсовой работы заключается в создании интеллектуальных моделей, которые в полной мере отражают существующие причинно-следственные взаимосвязи между определенными фактами внешнего мира и уровнем заработных плат разработчиков ПО в США.

Для достижения поставленной цели применяется Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная система «Эйдос».

В данной курсовой работе необходимо проанализировать методы формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами, принятия решений и исследования моделируемой предметной области путем исследования ее модели.

# СОДЕРЖАНИЕ

_Тос127832295ВВЕДЕНИЕ .....	4
1 СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ .....	5
1.1 Описание решения .....	5
1.2 Поиск и объединение исходных данных .....	7
1.3 Ввод выборки в систему Aidos-X .....	8
1.4 Синтез и верификация моделей .....	10
1.5 Виды моделей Aidos-X .....	11
1.6 Результаты верификации моделей .....	12
2 РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ .....	17
2.1 Решение задач идентификации .....	17
2.2 SWOT и PEST матрицы и диаграммы .....	20
2.3 Исследование моделируемой предметной области путем исследования ее модели .....	21
2.3.1 Когнитивные диаграммы классов .....	22
2.3.2 Агломеративная когнитивная кластеризация классов .....	22
2.3.3 Когнитивные диаграммы значений факторов .....	24
2.3.4 Агломеративная когнитивная кластеризация значений факторов ..	24
2.3.5 Нелокальные нейроны и нелокальные нейронные сети .....	26
2.3.6 3D-интегральные когнитивные карты .....	27
2.3.7 Когнитивные функции .....	28
ЗАКЛЮЧЕНИЕ .....	30
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ .....	31

## ВВЕДЕНИЕ

Технологии искусственного интеллекта быстро развиваются и применяются в самых разных областях, таких как финансы, тяжелая промышленность, здравоохранение, управление персоналом, музыка, журналистика, обслуживание клиентов и многое другое. Искусственный интеллект используется в алгоритмической торговле, маркетинговых исследованиях, управлении финансовым портфелем, опасной работе с роботами, медицинской диагностике, уходе за пожилыми людьми и пациентами, подборе персонала и управлении им, автоматизированной журналистике, поддержке клиентов и во многих других отраслях.

Данная курсовая работа посвящена решению задачи анализа, а именно анализу заработных плат разработчиков ПО в США по некоторым статистическим характеристикам внешнего мира. Выявление зависимостей между наблюдаемыми данными и уровнем зарплат позволит облегчить процесс формирования новых вакантных позиций и их условий для компаний.

Из всего вышесказанного можно выделить цель работы. Целью данной работы является разработка интеллектуальных моделей, отражающих реально существующие причинно-следственные взаимосвязи между некоторыми характеристиками реального мира и зарплатой.

Для достижения поставленной цели в данной работе будет применена технология автоматизированного системно-когнитивного анализа (АСК-анализ), который будет осуществлён с помощью интеллектуальной системы «AidosX++».

# 1 СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

## 1.1 Описание решения

В качестве метода исследования, решения проблемы и достижения цели было решено использовать новый метод искусственного интеллекта: Автоматизированный системно-когнитивный анализ или АСК-анализ.

Главным фактором выбора АСК-анализа является то, что он включает в себя не только теорию, но и метод количественного выявления в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных данных крайне большой размерности числовой и не числовой природы, которые являются измеряемыми в различных типах шкал и единицах измерения.

Универсальная когнитивная аналитическая система Aidos-X в настоящее время выступает очень важным, развитым и доступным программным инструментарием для реализации АСК-анализа.

Преимущества данной системы:

- универсальность;
- содержит большое количество локальных и облачных учебных и научных приложений (в настоящее время их 31 и 364, соответственно);
- мультиязычная поддержка интерфейса (больше 50 языков);
- поддерживает online-среду накопления знаний и широко используется во всем мире;
- 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;
- доступность, то есть данная система находится в полном открытом бесплатном доступе, причем с актуальными исходными текстам;

- одна из первых отечественных систем искусственного интеллекта персонального уровня, т.е. она не требует от пользователя специальной подготовки в области технологий искусственного интеллекта (акт внедрения системы Aidos-X от 1987 года);
- 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;
- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение 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 (развитая когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе);
- хорошо имитирует человеческий стиль мышления: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции.

В данной работе будет реализовано решение поставленной задачи в четыре этапа в соответствии с методологией АСК-анализа:

1. Поиск и объединение исходных данных в единую таблицу и сохранение в файле MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы Aidos-X.
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

## 1.2 Поиск и объединение исходных данных

Входные данные, необходимые для обучения модели, были взяты из открытых источников. Итоговый файл содержит следующие данные:

- City – город наблюдения;
- Number of Software Developer Jobs – количество вакансий;
- Median Home Price – средняя цена дома;
- Cost of Living Plus Rent avg – средняя стоимость жизни с арендой;
- Local Purchasing Power avg – средняя покупательская способность;
- Mean Software Developer Salary (unadjusted) – средняя зарплата;

Фрагмент исходных данных в виде таблицы MS Excel представлен на рисунке 1.

	A	B	C	D	E	F
	City	Number of Software Developer Jobs	Median Home Price	Cost of Living Plus Rent avg	Local Purchasing Power avg	Mean Software Developer Salary (unadjusted)
1	Columbus, OH	13430	192000	2856.5	9335.4	108500
2	Seattle, WA	65760	491600	4091.5	8971.3	131167
3	Charlotte, NC	12800	208500	3221.1	8939.8	107046
4	Colorado Springs, CO	5780	296500	3094.5	8493.1	111670
5	Dayton, OH	4240	124100	2586	4887.7	99338
6	Greensboro, NC	1560	136000	2888	5721.9	99612
7	San Francisco, CA	59550	946700	5290.7	8216.7	142101
8	Oakland, CA	59550	946700	4754.9	6894.3	142101
9	San Antonio, TX	5410	194100	2688.9	8494.9	104688
10	Fort Worth, TX	46130	243900	2960	6949.2	111113
11	Dallas, TX	46130	243900	3128.6	10530.6	111113
12	Durham, NC	6800	243700	2999.9	6464.2	105018
13	Atlanta, GA	38110	219600	3334	8063	106334
14	Cincinnati, OH	8620	169900	2775.1	7723.6	98112
15	Austin, TX	21160	311300	3275.7	9762.4	109535
16	Raleigh, NC	12560	267200	2922.2	8892.9	103584
17	Houston, TX	20400	205600	2852.3	10673.7	109503
18	Baltimore, MD	16200	267800	3078.7	8399.9	115006
19	Las Vegas, NV	3900	279900	2848.1	7702.6	104495
20	Tucson, AZ	4480	206200	130097.8	5164.1	101919
21	Salt Lake City, UT	9600	366000	2854.9	8676.4	105334
22	Richmond, VA	6390	229600	2814.5	8079.7	101153
23	Lexington, KY	1140	178400	130097.8	6318.6	95986
24	Denver, CO	28750	407900	3460.6	8020.4	111624
25	Santa Clara, CA	78730	1193600	4720.8	9589.6	137397
26	San Jose, CA	78730	1193600	4083.6	9711.8	137397
27	Portland, OR	15420	397400	3446.9	6669.7	106108
28	Vancouver, WA	15420	397400	2983.6	7780.4	106108
29	Wichita, KS	1220	132900	2241.4	7357.7	92595
30	Phoenix, AZ	20440	266600	3015.1	8388.2	100804
31	Mesa, AZ	20440	266600	2588.6	6381.5	100804
32	Santa Barbara, CA	2060	602600	4582.1	5791.6	112648
33	Los Angeles, CA	53190	649500	4043.7	7782.2	119662
34	Long Beach, CA	53190	649500	3356.6	6789.4	119662
35	Boston, MA	54140	465000	4312.1	7140.5	114227
36	Orlando, FL	11050	238900	3029.9	7609.5	100358
37	Tampa, FL	12160	213800	3257.8	6064.4	100614
38	Birmingham, AL	2310	146000	2893.8	5240.6	90219
39	Syracuse, NY	2140	135000	2794.5	4840.1	98320
40	Washington, DC	63760	407700	4146.1	7517.5	119806
41	Sacramento, CA	9980	411500	3418	6268.7	102801
42	Philadelphia, PA	26370	232800	3430.6	7133.7	103352
43	Des Moines, IA	4060	188800	2674.2	6685.7	93386
44	Omaha, NE	5220	172200	2842.6	6642.6	92326

Рисунок 1 – Фрагмент обучающей выборки

### 1.3 Ввод выборки в систему Aidos-X

Затем обучающая выборка в формате xls, импортирована в систему Aidos-X: скопирована в папку Inp\_data и переименована в Inp\_data.

Далее был использован универсальный программный интерфейс импорта данных в систему Aidos-X, а именно режим 2.3.2.2 (Рисунок 2).

В форме задания размерности модели системы указываем диапазоны столбцов и желаемое количество градаций описательных и классификационных шкал (Рисунок 3).

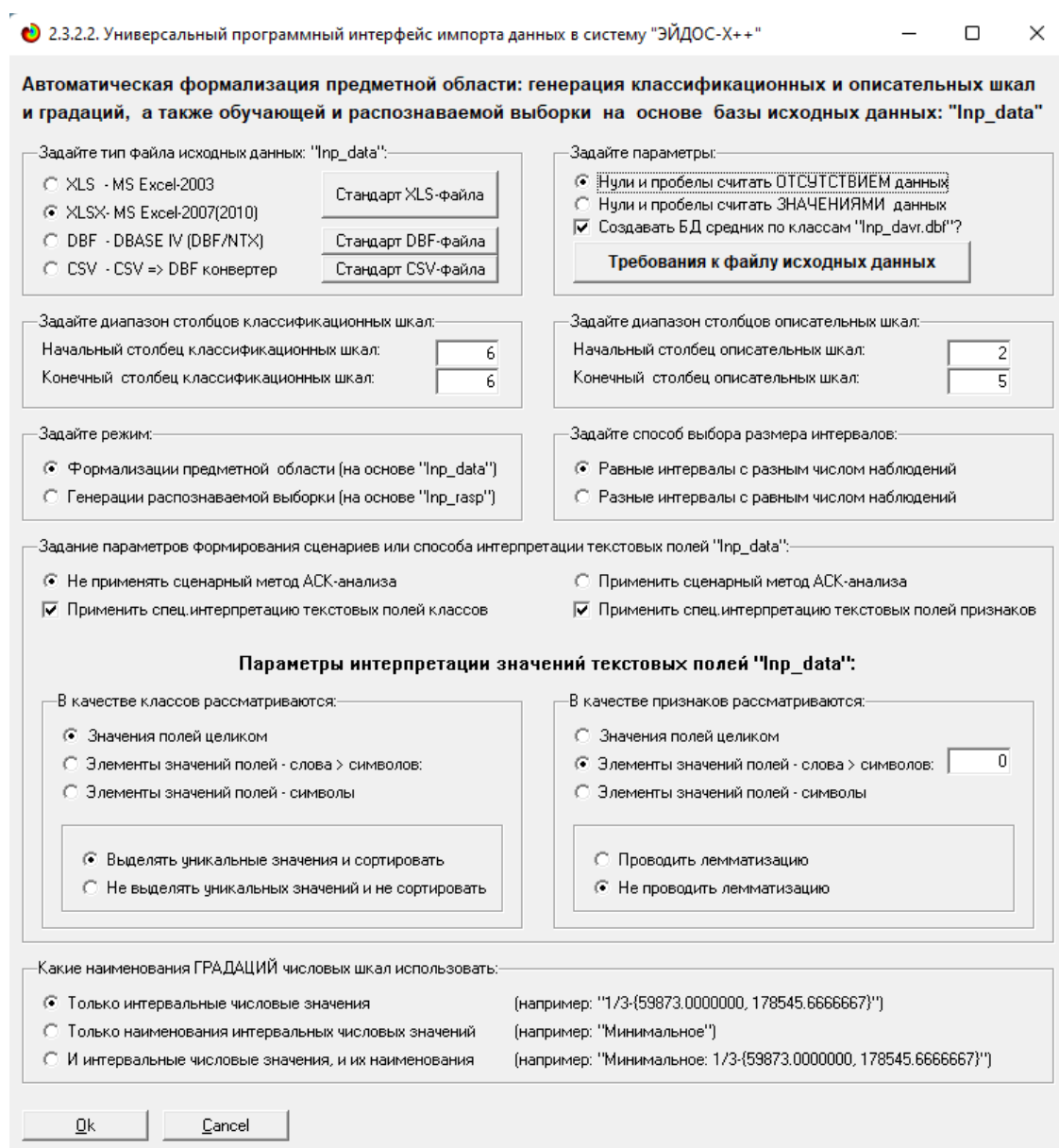


Рисунок 2 – Интерфейс импорта данных



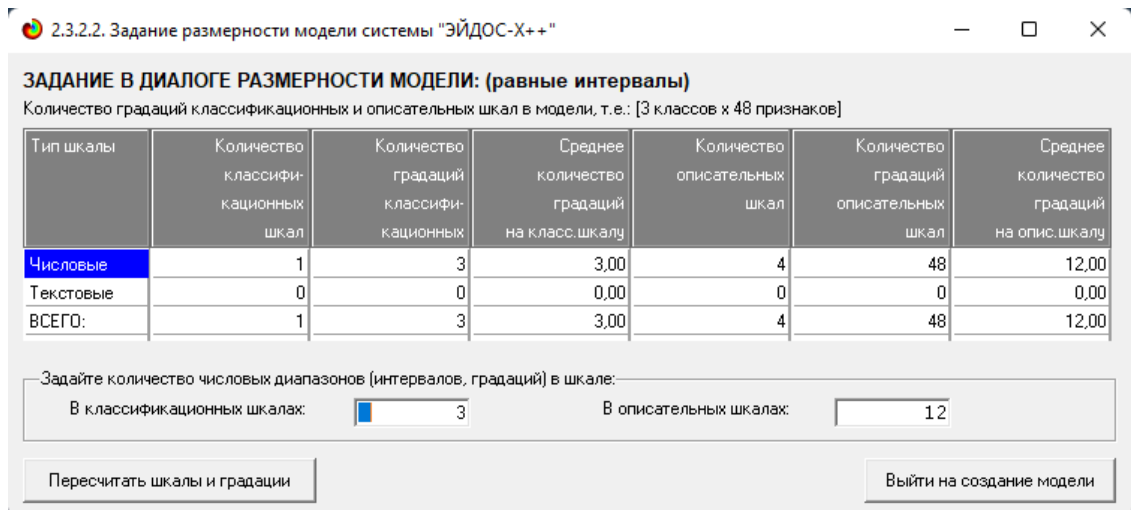


Рисунок 3 – Задание размерности модели

Процесс импорта данных из внешней БД "Inp\_data.xls" в систему Aidos-X представлен на рисунке 4.

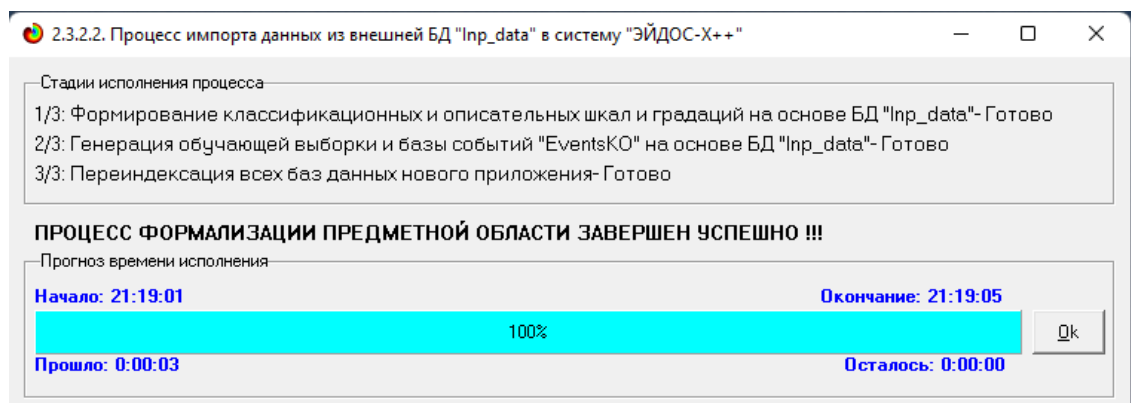


Рисунок 4 – Процесс импорта данных

Формирование классификационных и описательные шкал, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных, осуществляется автоматически после импорта. Тем самым этап формализации предметной области выполняется полностью автоматизировано.

## 1.4 Синтез и верификация моделей

Следующим шагом требуется запустить режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей (рисунок 5).

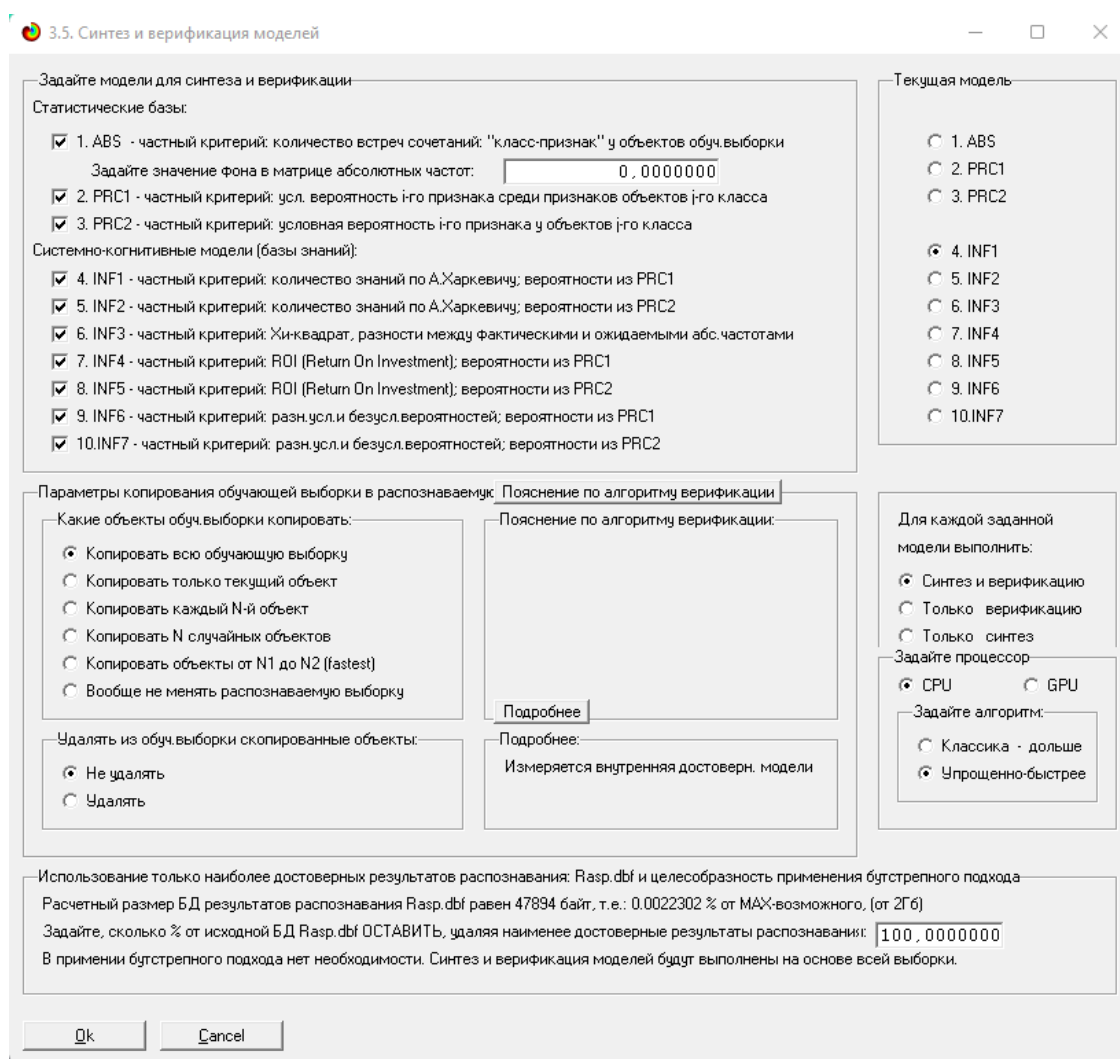


Рисунок 5 – Выбор моделей для синтеза и верификации

В данном режиме имеется много различных методов верификации моделей. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 6.

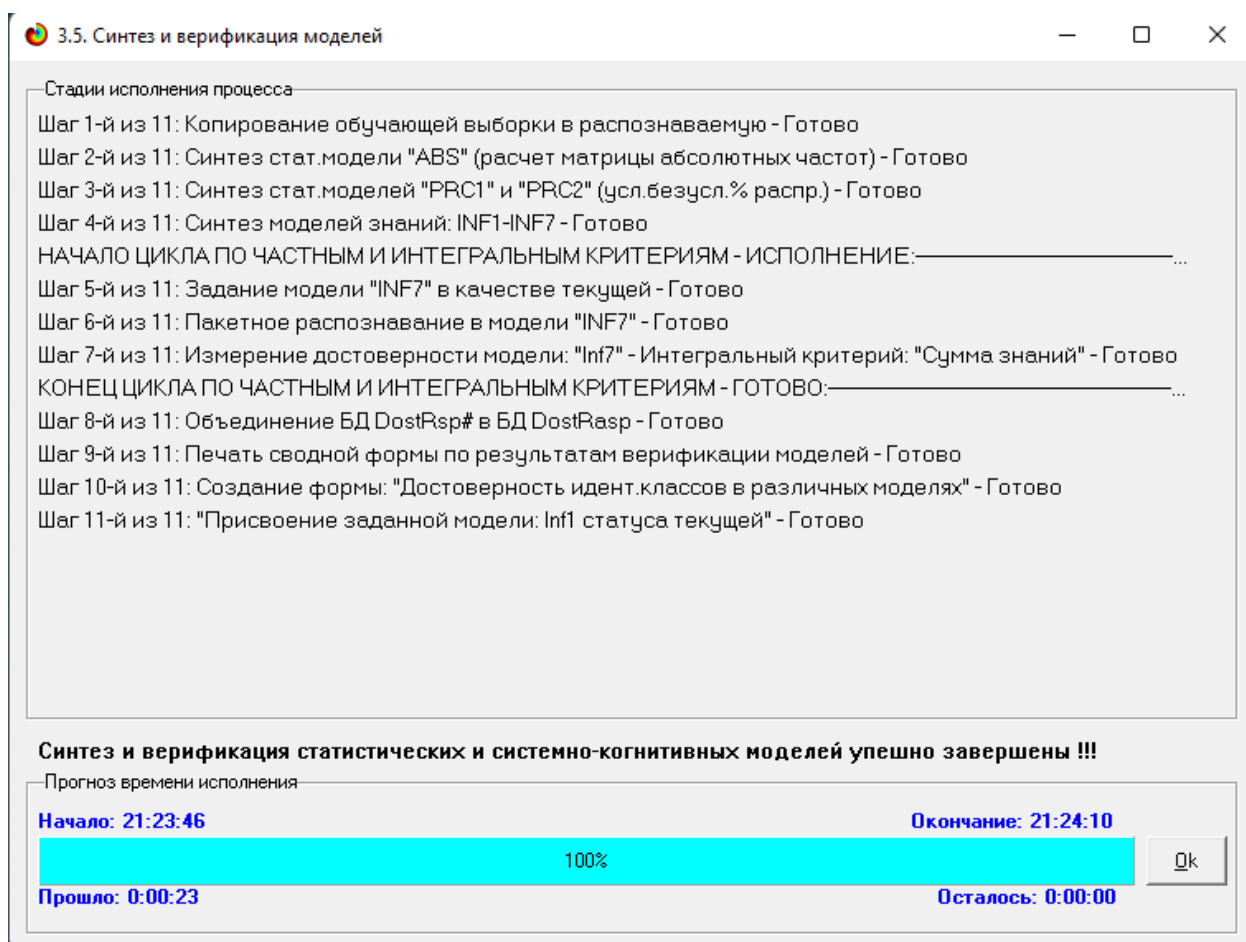


Рисунок 6 – Процесс исполнения режима 3.5

Стоит отметить, что синтез и верификация всех 10 моделей на данной задаче заняли 23 секунды. Далее перейдем непосредственно к выбору наиболее достоверной модели.

### 1.5 Виды моделей Aidos-X

На рисунках 7 и 8 приведены фрагменты созданных статистических и системно-когнитивных моделей (СК-модели).

Рассмотрим решение задачи идентификации на примере нескольких моделей, в которых рассчитано количество информации по А. Харкевичу, которое было получено по принадлежности идентифицируемого объекта к каждому из классов, если известно, что у этого объекта есть некоторый признак.

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

Код признака	Наименование описательной шкалы и градации	1. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 1/3 (80314.0000000, 100909.6666667)	2. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 2/3 (100909.666666, 121505.3333333)	3. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 3/3 (121505.33333, 142101.0000000)	Сумма	Среднее	Средн. квадрат. откл.
1.0	NUMBER OF SOFTWARE DEVELOPER JOBS-1/12-(1120.0000000, 9247.5000000)	32.0	7.0		39.0	13.00	16.82
2.0	NUMBER OF SOFTWARE DEVELOPER JOBS-2/12-(9247.5000000, 17375.0000000)	6.0	8.0		14.0	4.67	4.16
3.0	NUMBER OF SOFTWARE DEVELOPER JOBS-3/12-(17375.0000000, 25502.5000000...)	2.0	4.0		6.0	2.00	2.00
4.0	NUMBER OF SOFTWARE DEVELOPER JOBS-4/12-(25502.5000000, 33630.0000000...)	1.0	2.0		3.0	1.00	1.00
5.0	NUMBER OF SOFTWARE DEVELOPER JOBS-5/12-(33630.0000000, 41757.5000000...)		1.0		1.0	0.33	0.58
6.0	NUMBER OF SOFTWARE DEVELOPER JOBS-6/12-(41757.5000000, 49885.0000000...)		3.0		3.0	1.00	1.73
7.0	NUMBER OF SOFTWARE DEVELOPER JOBS-7/12-(49885.0000000, 58012.5000000...)		3.0		3.0	1.00	1.73
8.0	NUMBER OF SOFTWARE DEVELOPER JOBS-8/12-(58012.5000000, 66140.0000000...)		1.0	3.0	4.0	1.33	1.53
9.0	NUMBER OF SOFTWARE DEVELOPER JOBS-9/12-(66140.0000000, 74267.5000000...)						
10.0	NUMBER OF SOFTWARE DEVELOPER JOBS-10/12-(74267.5000000, 82395.0000000...)			2.0	2.0	0.67	1.15
11.0	NUMBER OF SOFTWARE DEVELOPER JOBS-11/12-(82395.0000000, 90522.5000000...)						
12.0	NUMBER OF SOFTWARE DEVELOPER JOBS-12/12-(90522.5000000, 98650.0000000...)		2.0		2.0	0.67	1.15
13.0	MEDIAN HOME PRICE-1/12-(124100.0000000, 213225.0000000)	23.0	5.0		28.0	9.33	12.10
14.0	MEDIAN HOME PRICE-2/12-(213225.0000000, 302350.0000000)	16.0	12.0		28.0	9.33	8.33
15.0	MEDIAN HOME PRICE-3/12-(302350.0000000, 391475.0000000)	1.0	2.0		3.0	1.00	1.00
16.0	MEDIAN HOME PRICE-4/12-(391475.0000000, 480600.0000000)		8.0		8.0	2.67	4.62
17.0	MEDIAN HOME PRICE-5/12-(480600.0000000, 569725.0000000)			1.0	1.0	0.33	0.58
18.0	MEDIAN HOME PRICE-6/12-(569725.0000000, 658850.0000000)		4.0		4.0	1.33	2.31
19.0	MEDIAN HOME PRICE-7/12-(658850.0000000, 747975.0000000)	1.0			1.0	0.33	0.58
20.0	MEDIAN HOME PRICE-8/12-(747975.0000000, 837100.0000000)						
21.0	MEDIAN HOME PRICE-9/12-(837100.0000000, 926225.0000000)						

Рисунок 7 – Фрагмент матрицы абсолютных частот

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

Код признака	Наименование описательной шкалы и градации	1. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 1/3 (80314.0000000, 100909.6666667)	2. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 2/3 (100909.6666667, 121505.3333333)	3. MEAN SOFTWARE DEVELOPER SALARY (UNADJUSTED) 3/3 (121505.3333333, 142101.0000000)	Сумма	Среднее	Средн. квадрат. откл.
1.0	NUMBER OF SOFTWARE DEVELOPER JOBS-1/12-(1120.0000000, 9247.5000000)	11.234	-8.701	-2.532			10.206
2.0	NUMBER OF SOFTWARE DEVELOPER JOBS-2/12-(9247.5000000, 17375.0000000)	-1.455	2.364	-0.909			2.065
3.0	NUMBER OF SOFTWARE DEVELOPER JOBS-3/12-(17375.0000000, 25502.5000000...)	-1.195	1.584	-0.390			1.430
4.0	NUMBER OF SOFTWARE DEVELOPER JOBS-4/12-(25502.5000000, 33630.0000000...)	-0.597	0.792	-0.195			0.715
5.0	NUMBER OF SOFTWARE DEVELOPER JOBS-5/12-(33630.0000000, 41757.5000000...)	-0.532	0.597	-0.065			0.568
6.0	NUMBER OF SOFTWARE DEVELOPER JOBS-6/12-(41757.5000000, 49885.0000000...)	-1.597	1.792	-0.195			1.703
7.0	NUMBER OF SOFTWARE DEVELOPER JOBS-7/12-(49885.0000000, 58012.5000000...)	-1.597	1.792	-0.195			1.703
8.0	NUMBER OF SOFTWARE DEVELOPER JOBS-8/12-(58012.5000000, 66140.0000000...)	-2.130	-0.610	2.740			2.492
9.0	NUMBER OF SOFTWARE DEVELOPER JOBS-9/12-(66140.0000000, 74267.5000000...)						
10.0	NUMBER OF SOFTWARE DEVELOPER JOBS-10/12-(74267.5000000, 82395.0000000...)	-1.065	-0.805	1.870			1.625
11.0	NUMBER OF SOFTWARE DEVELOPER JOBS-11/12-(82395.0000000, 90522.5000000...)						
12.0	NUMBER OF SOFTWARE DEVELOPER JOBS-12/12-(90522.5000000, 98650.0000000...)	-1.065	1.195	-0.130			1.135
13.0	MEDIAN HOME PRICE-1/12-(124100.0000000, 213225.0000000)	8.091	-6.273	-1.818			7.352
14.0	MEDIAN HOME PRICE-2/12-(213225.0000000, 302350.0000000)	1.091	0.727	-1.818			1.585
15.0	MEDIAN HOME PRICE-3/12-(302350.0000000, 391475.0000000)	-0.597	0.792	-0.195			0.715
16.0	MEDIAN HOME PRICE-4/12-(391475.0000000, 480600.0000000)	-4.260	4.779	-0.519			4.542
17.0	MEDIAN HOME PRICE-5/12-(480600.0000000, 569725.0000000)	-0.532	-0.403	0.935			0.812
18.0	MEDIAN HOME PRICE-6/12-(569725.0000000, 658850.0000000)	-2.130	2.390	-0.260			2.271
19.0	MEDIAN HOME PRICE-7/12-(658850.0000000, 747975.0000000)	0.468	-0.403	-0.065			0.439
20.0	MEDIAN HOME PRICE-8/12-(747975.0000000, 837100.0000000)						
21.0	MEDIAN HOME PRICE-9/12-(837100.0000000, 926225.0000000)						

Рисунок 8 – Фрагмент модели INF3

## 1.6 Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями, были отражены на рисунке 9.

3.4. Обобщенная форма по доств.моделям при разн.инт.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Всего признаков объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Фигура Ван Ризбергера	Средняя мод. уровн. спец. решений (SP)	Средняя мод. уровн. спец. истинно-отриц. решений (STN)	Средняя мод. уровн. довл. по решению
1. ABS - частный критерий: количество встреч соизвещаний "1.клас...	Корреляция абс частот с обр...	77	77	2	152	0	0.336	1.000	0.503	59.895	0.316	73
1. ABS - частный критерий: количество встреч соизвещаний "1.клас...	Сумма абс частот по признак...	77	77	2	152	0	0.336	1.000	0.503	50.097		44
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн частот с о...	77	77	2	152	0	0.336	1.000	0.503	59.895	0.316	73
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн частот по при...	77	77	2	152	0	0.336	1.000	0.503	50.192		77
3. PR2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн частот с о...	77	77	2	152	0	0.336	1.000	0.503	59.894	0.316	73
3. PR2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн частот по при...	77	77	2	152	0	0.336	1.000	0.503	50.192		77
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	77	64	132	22	13	0.744	0.831	0.785	28.227	47.369	3
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	77	64	67	87	13	0.424	0.831	0.561	12.753	9.960	3
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	77	64	132	22	13	0.744	0.831	0.785	28.227	47.369	3
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	77	64	67	87	13	0.424	0.831	0.561	12.753	9.960	3
6. INF3 - частный критерий: Минкватрат, разности между факти...	Семантический резонанс зна...	77	66	134	20	11	0.767	0.857	0.810	35.227	56.031	6
6. INF3 - частный критерий: Минкватрат, разности между факти...	Сумма знаний	77	66	134	20	11	0.767	0.857	0.810	32.536	34.771	5
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	77	62	141	13	15	0.827	0.805	0.816	30.050	42.163	3
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Сумма знаний	77	70	61	93	7	0.429	0.909	0.583	6.988	1.431	1
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	77	62	141	13	15	0.827	0.805	0.816	30.050	42.163	3
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Сумма знаний	77	70	61	93	7	0.429	0.909	0.583	6.988	1.431	1
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Семантический резонанс зна...	77	63	136	18	14	0.778	0.818	0.797	34.239	36.504	6
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Сумма знаний	77	66	59	95	11	0.410	0.857	0.555	22.059	14.721	7
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Семантический резонанс зна...	77	63	136	18	14	0.778	0.818	0.797	34.239	36.504	6
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Сумма знаний	77	66	59	95	11	0.410	0.857	0.555	22.059	14.721	7

Рисунок 9 – Оценка достоверности моделей (F-критерий)

3.4. Обобщенная форма по доств.моделям при разн.инт.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	S Полнота модели	L1-нр проф. Е.В. Луценко	Средний модуль уровн. спец. истинно-полож. решений	Средний модуль уровн. спец. истинно-отриц. решений	Средний модуль уровн. спец. ложно-полож. решений	Средний модуль уровн. спец. ложно-отриц. решений	A Точность модели AРесон = АТР/АТР <sub>4</sub>	A Полнота модели AРесон = АТР/АТР <sub>4</sub>	L2-нр проф. Е.В. Луценко	Процент правльн. идентифиц.	Процент правльн. не идрт	
1. ABS - частный критерий: количество встреч соизвещаний "1.клас...	Корреляция абс частот с обр...	10	1.000	0.620	0.778	0.158	0.482	0.617	1.000	0.763	100.000	0	
1. ABS - частный критерий: количество встреч соизвещаний "1.клас...	Сумма абс частот по признак...	10	1.000	0.693	0.651	0.292	0.292	0.690	1.000	0.817	100.000	0	
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн частот с о...	10	1.000	0.620	0.778	0.158	0.482	0.617	1.000	0.763	100.000	0	
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн частот по при...	10	1.000	0.600	0.756	0.158	0.510	0.597	1.000	0.748	100.000	0	
3. PR2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн частот с о...	10	1.000	0.620	0.778	0.158	0.482	0.617	1.000	0.763	100.000	0	
3. PR2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн частот по при...	10	1.000	0.600	0.756	0.158	0.510	0.597	1.000	0.748	100.000	0	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	18	0.853	0.865	0.441	0.359	0.179	0.373	0.711	0.542	0.615	83.117	75
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	12	0.902	0.826	0.199	0.149	0.046	0.106	0.813	0.652	0.724	83.117	70
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	18	0.853	0.865	0.441	0.359	0.179	0.373	0.711	0.542	0.615	83.117	75
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	12	0.902	0.826	0.199	0.149	0.046	0.106	0.813	0.652	0.724	83.117	70
6. INF3 - частный критерий: Минкватрат, разности между факти...	Семантический резонанс зна...	19	0.898	0.868	0.534	0.418	0.313	0.404	0.631	0.569	0.598	85.714	79
6. INF3 - частный критерий: Минкватрат, разности между факти...	Сумма знаний	15	0.897	0.870	0.493	0.259	0.298	0.339	0.623	0.593	0.608	85.714	79
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	17	0.894	0.890	0.465	0.299	0.294	0.238	0.622	0.671	0.646	80.519	82
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Сумма знаний	15	0.979	0.884	0.100	0.023	0.018	0.021	0.846	0.827	0.836	90.909	64
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	17	0.894	0.890	0.465	0.299	0.294	0.238	0.622	0.671	0.646	80.519	82
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Сумма знаний	15	0.979	0.884	0.100	0.023	0.018	0.021	0.846	0.827	0.836	90.909	64
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Семантический резонанс зна...	18	0.878	0.863	0.543	0.268	0.340	0.338	0.615	0.616	0.616	81.818	78
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей; вер...	Сумма знаний	17	0.895	0.808	0.334	0.250	0.083	0.235	0.801	0.587	0.678	85.714	61
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Семантический резонанс зна...	18	0.878	0.863	0.543	0.268	0.340	0.338	0.615	0.616	0.616	81.818	78
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей; ве...	Сумма знаний	17	0.895	0.808	0.334	0.250	0.083	0.235	0.801	0.587	0.678	85.714	61

Рисунок 10 – Оценка достоверности моделей (L1 и L2 критерии)

Из рисунка 9-10 можно сделать вывод о том, что в данном интеллектуальном приложении по F-критерию Ван Ризбергера наиболее достоверными моделями являются INF4 и INF5 с интегральным критерием «Семантический резонанс знаний» (F=0,816 максимуме 1,000), по критерию L1 профессора Е.В. Луценко наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием «Семантический резонанс знаний» (L1=0,890 при максимуме 1,000), а по критерию L2 профессора Е.В. Луценко наиболее достоверной является СК-модель INF4 и INF5 с интегральным критерием «Сумма знаний» (L2=0,836 при максимуме 1,000). Для оценки достоверности моделей в АСК- анализе и системе Aidos-X используется F-

мера Ван Ризбергена и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Е.В. Луценко (рисунок 11).

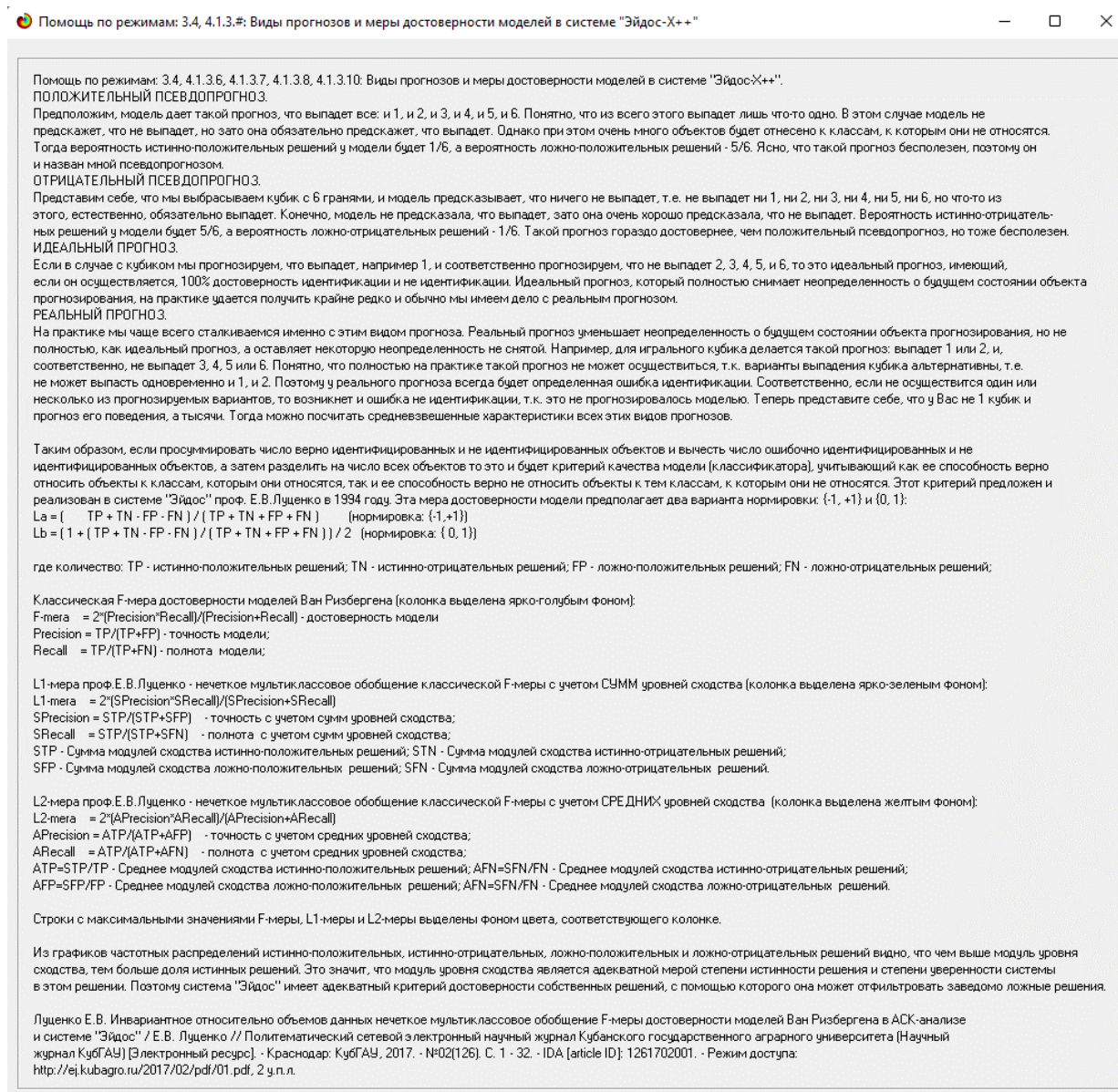


Рисунок 11 – Виды прогнозов и принципы определения

Также необходимо обратить внимание на то, что статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую.

На рисунках 12-15 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений в СК-моделях по данным обучающей выборки.

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

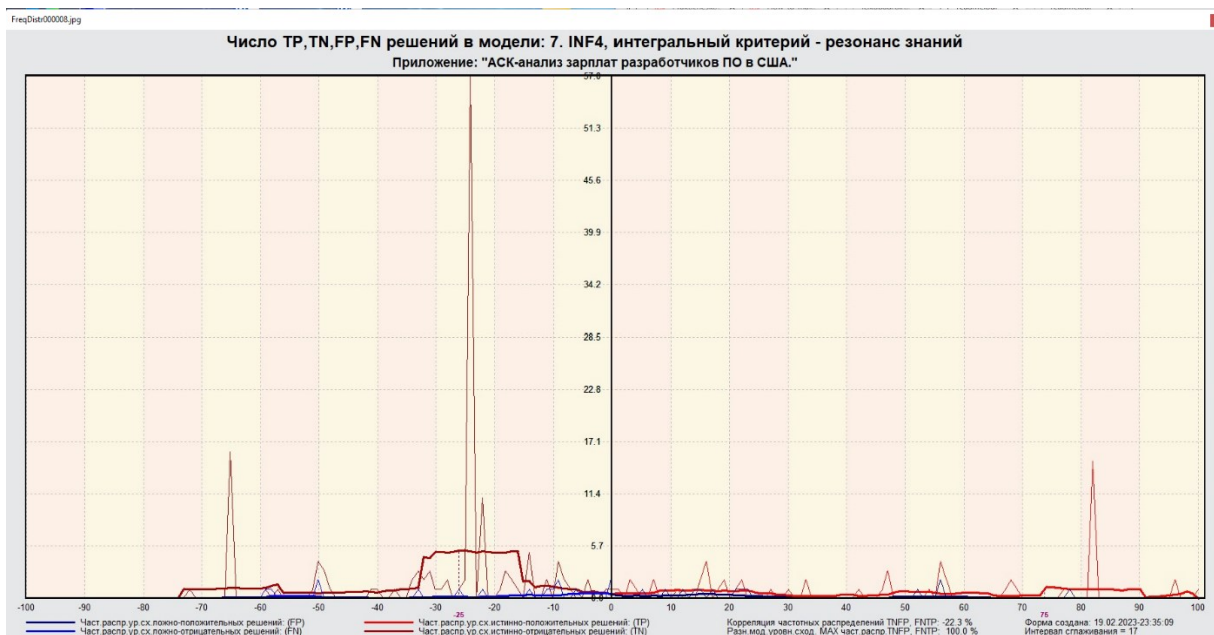


Рисунок 12 – Частотные распределения модели INF4 (резонанс знаний)

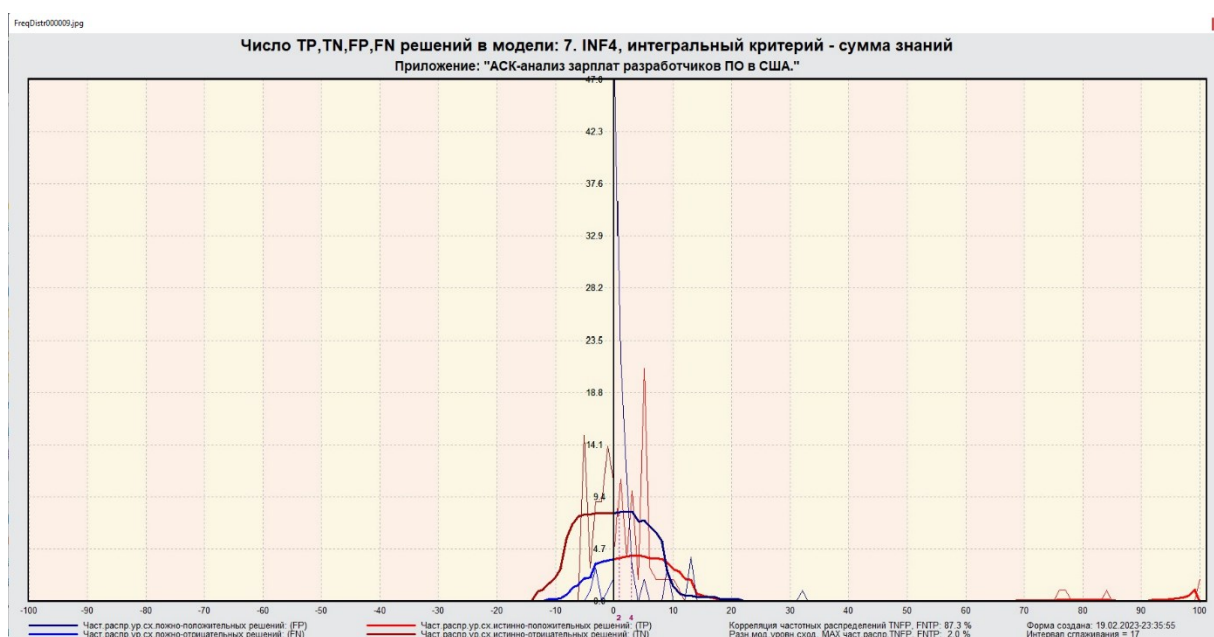


Рисунок 13 – Частотные распределения модели INF4 (сумма знаний)

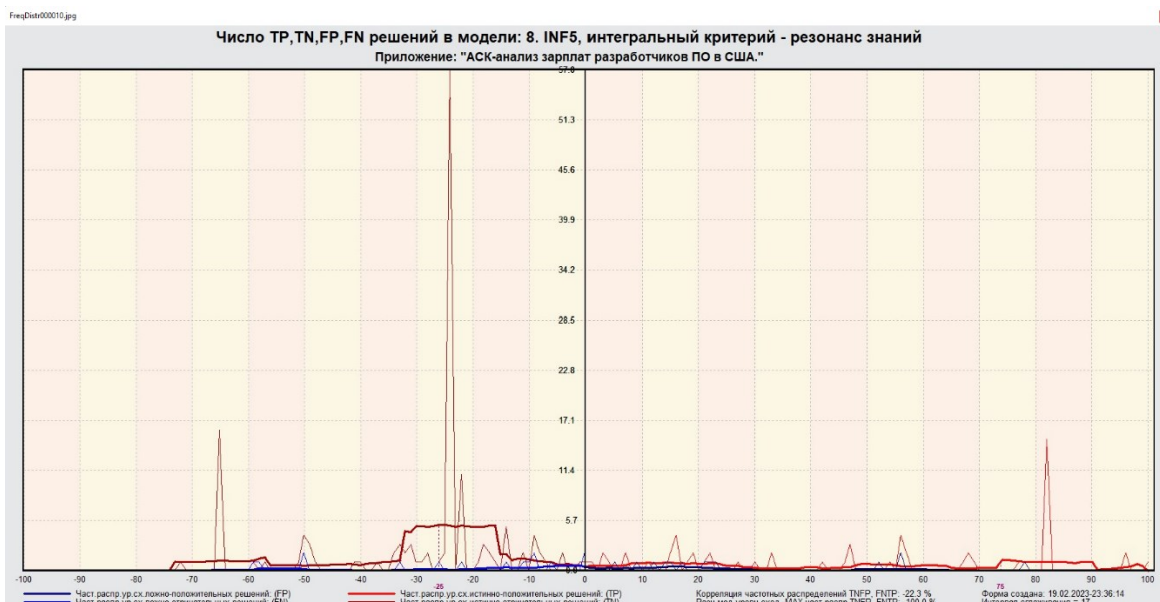


Рисунок 14 – Частотные распределения модели INF5 (резонанс знаний)

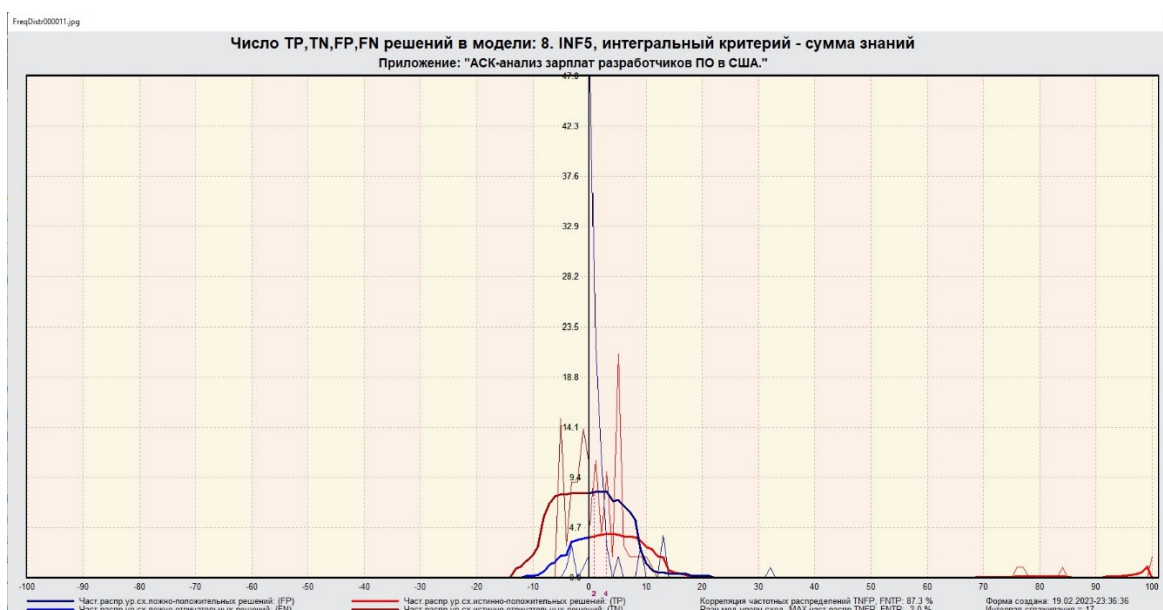


Рисунок 15 – Частотные распределения модели INF5 (сумма знаний)

Видно, что модели INF4 и INF5 на основе интегрального критерия «резонанс знаний» при проценте схожести более 70% практически не дают ложноотрицательных решений. При этом модели INF4 и INF5 на основе интегрального критерия «сумма знаний» при проценте схожести меньше примерно 15% дают меньше ложноположительных, чем истинно-положительных решений. Из этого можно сделать вывод о том, что система лучше определяет принадлежность объекта к классу, чем непринадлежность.



## 2 РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

### 2.1 Решение задач идентификации

Решим задачу системной идентификации, т.е. определим к каким респонденты к каким категориям относятся на основе обучающей выборки в наиболее достоверной СК-модели INF4 (режим 5.6) (рисунок 16, 17).

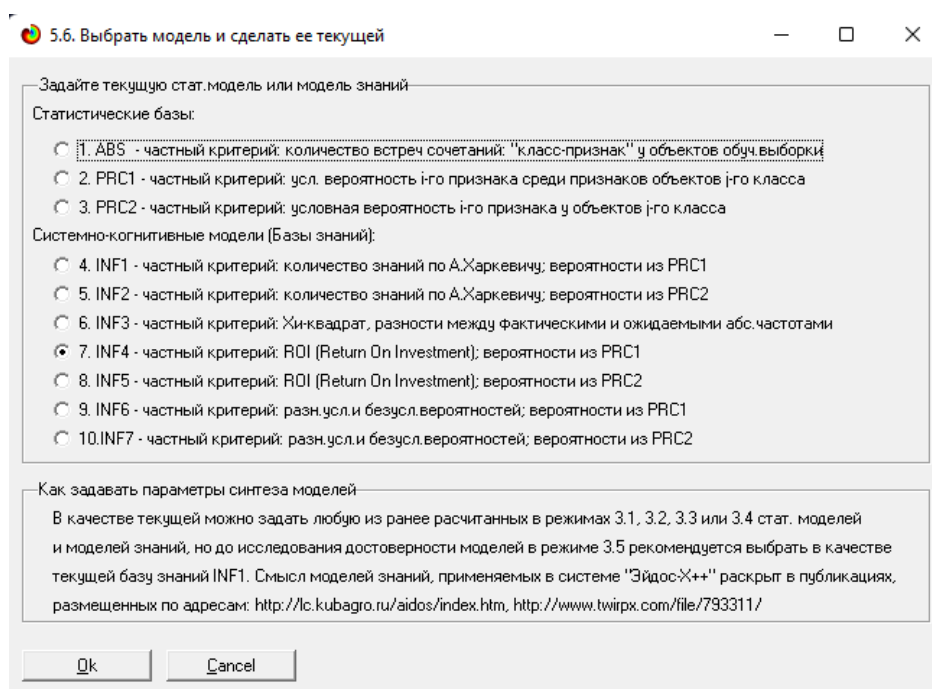


Рисунок 16 – Интерфейс задания модели

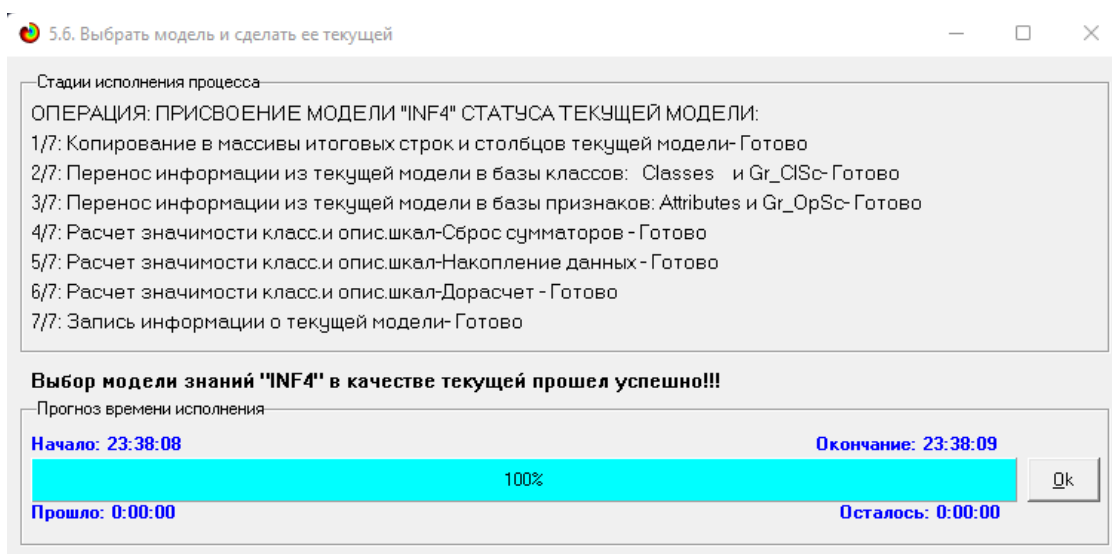


Рисунок 17 – Процесс загрузки модели

Затем следует провести пакетное распознавание в режиме 4.1.2. (рисунок 18).

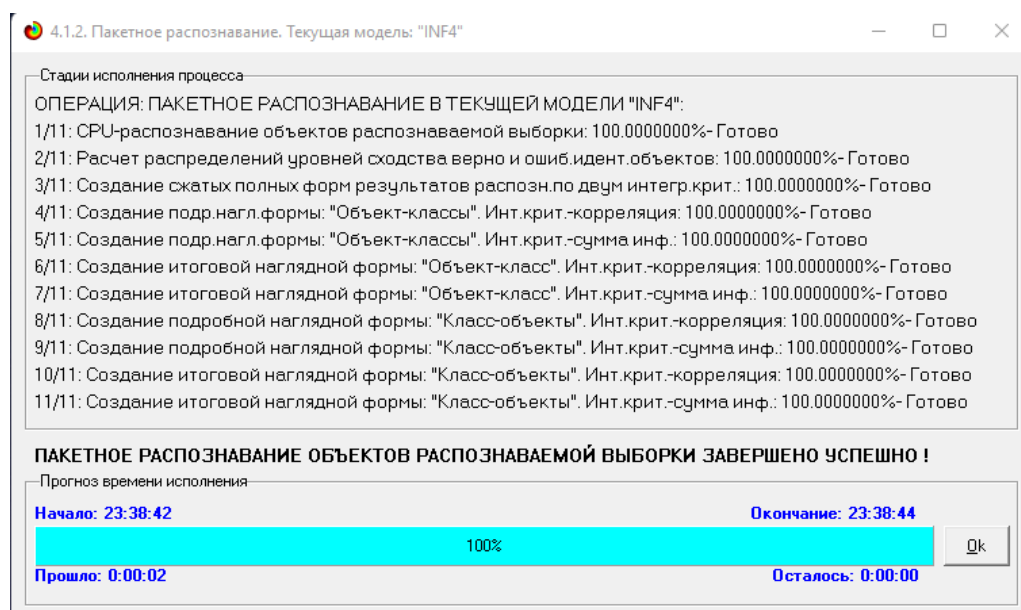


Рисунок 18 – Процесс пакетного распознавания модели

В результате пакетного распознавания в текущей модели создается ряд баз данных, которые визуализируются в выходных экранных формах, отражающих результаты решения задачи идентификации и прогнозирования.

Режим 4.1.3 системы Aidos-X обеспечивает отображение результатов идентификации и прогнозирования в различных формах:

- Подробно наглядно: «Объект – классы».
- Подробно наглядно: «Класс – объекты».
- Итоги наглядно: «Объект – классы».
- Итоги наглядно: «Класс – объекты».
- Подробно сжато: «Объект – классы».
- Обобщенная форма по достоверности моделей при разных интегральных критериях.
- Обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям.

– Статистический анализ результатов идентификации по классам, моделям и интегральным критериям.

На рисунке 19 и 20 приведен пример прогноза высокой и низкой достоверности частоты объектов и классов в модели INF4 на основе наблюдения предыстории их развития. Визуализация результатов распознавания представлена в соотношении: «Объект-классы» и «Класс-объекты».

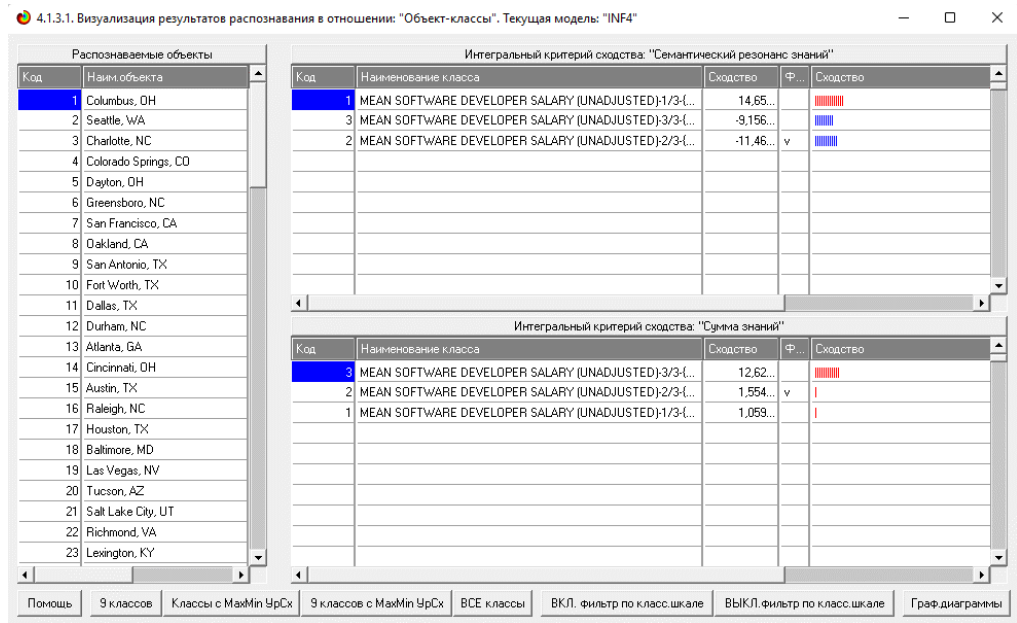


Рисунок 19 – Визуализация «Объект-классы»

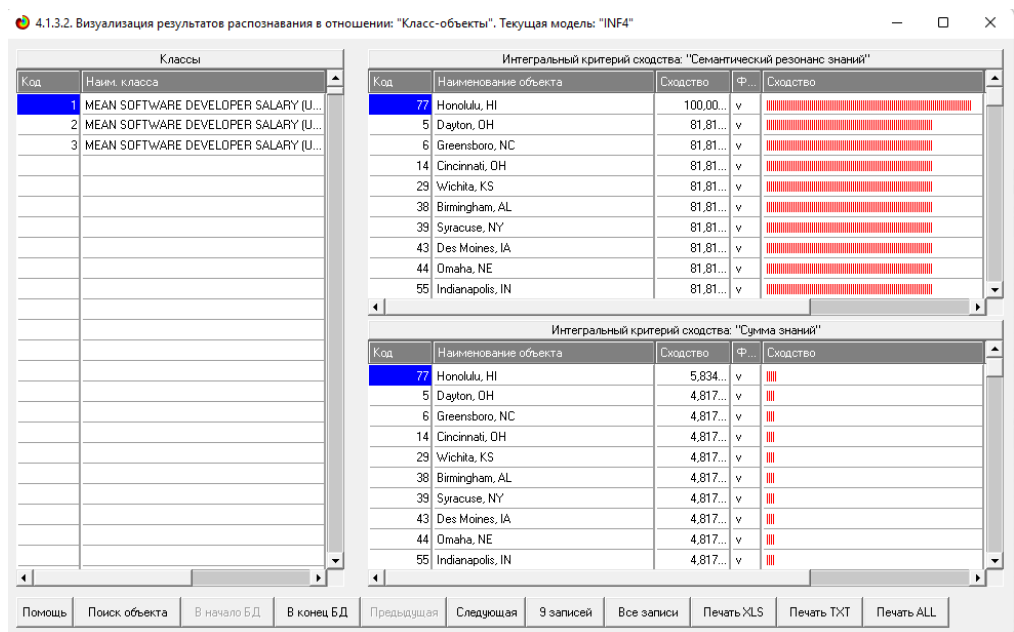


Рисунок 20 – Визуализация «Класс-объекты»

## 2.2 SWOT и PEST матрицы и диаграммы

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 решаемой задаче, SWOT-анализ показывает, как различные критерии влияют на уровень заработной платы (рисунок 21).

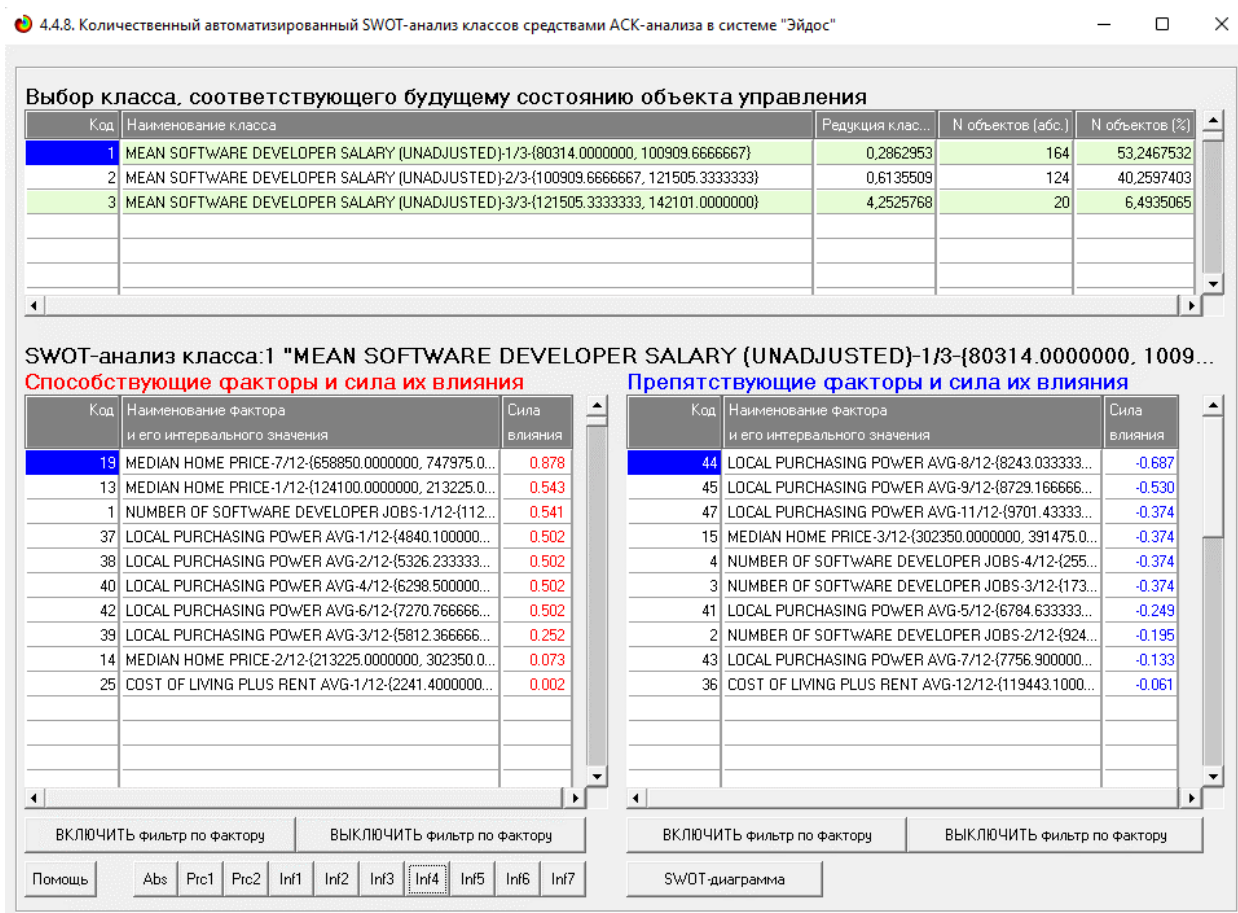


Рисунок 21 – SWOT-матрица

На рисунке 22 приведен пример табличной выходной формы количественного автоматизированного SWOT-анализа и PEST-анализа средствами системы Aidos-X.

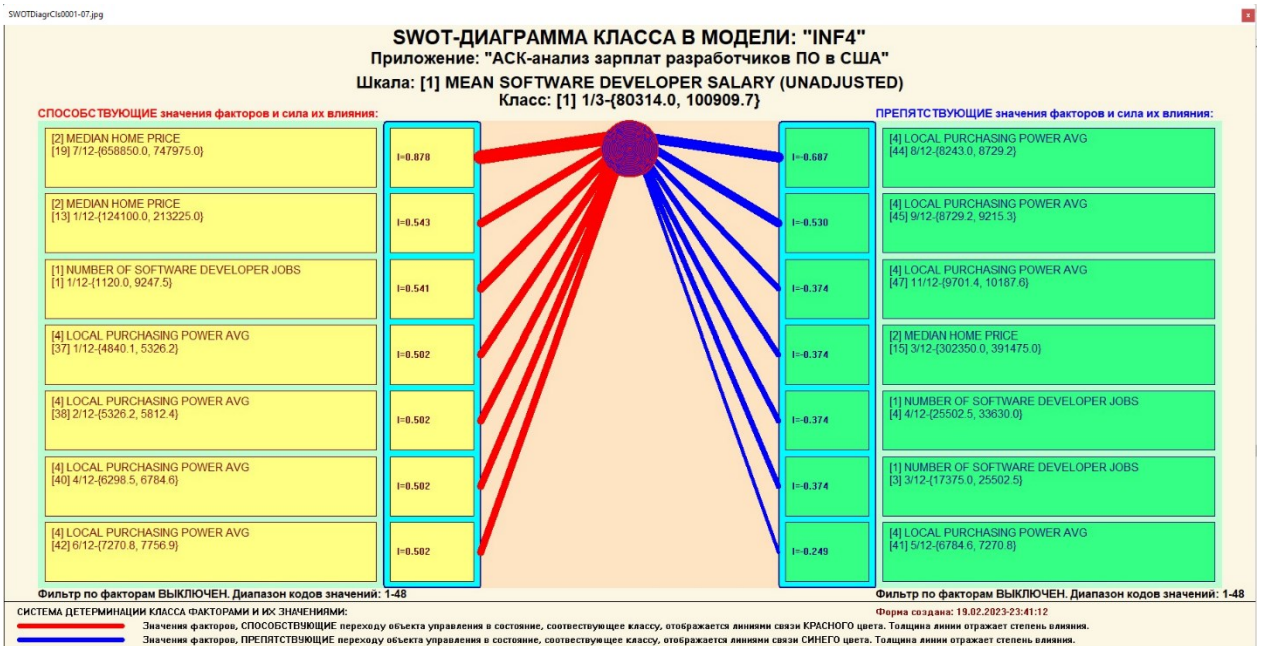


Рисунок 22 – SWOT-диаграмма

В левой стороне представлены факторы характерные для класса, а справа – нехарактерные.

### 2.3 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 «Aidos-X» 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.

### 2.3.1 Когнитивные диаграммы классов

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

Отметим также, что на когнитивной диаграмме, приведенной на рисунке 23, показаны количественные оценки сходства/различия показателей. 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.

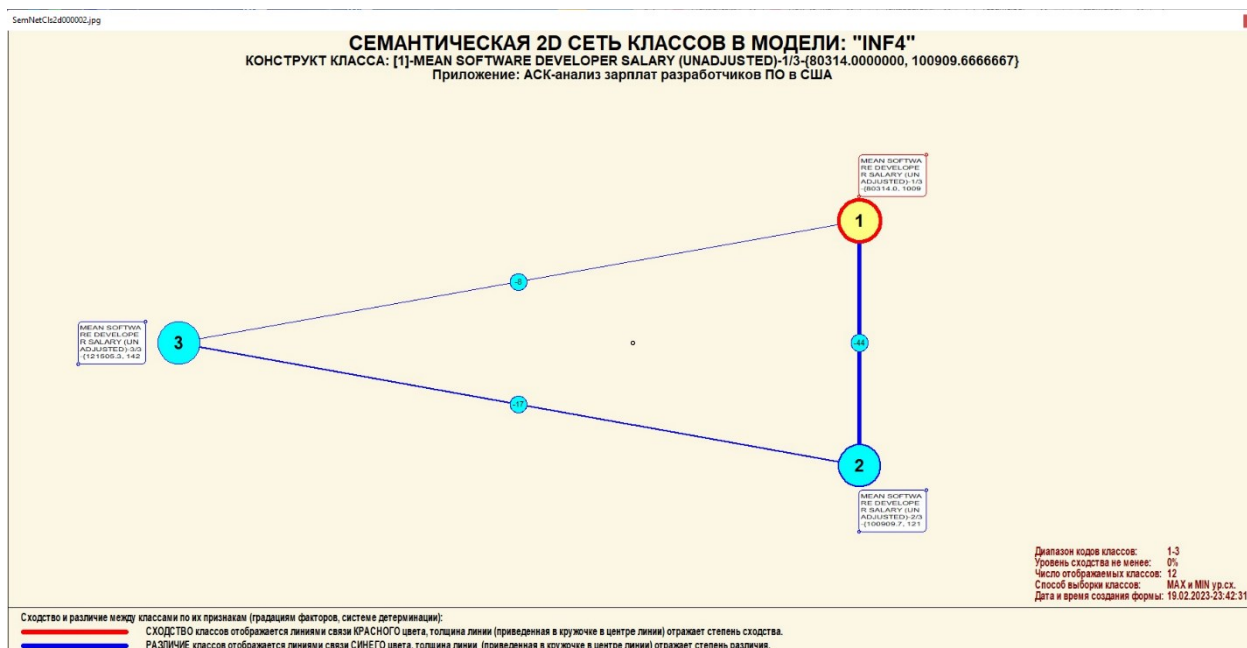


Рисунок 23 – Когнитивная диаграмма классов

### 2.3.2 Агломеративная когнитивная кластеризация классов

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме, когнитивных диаграмм, но и в форме агломеративных дендрограмм с графиком изменения межкластерного расстояния, полученных в результате когнитивной кластеризации (рисунок 24 и 25).

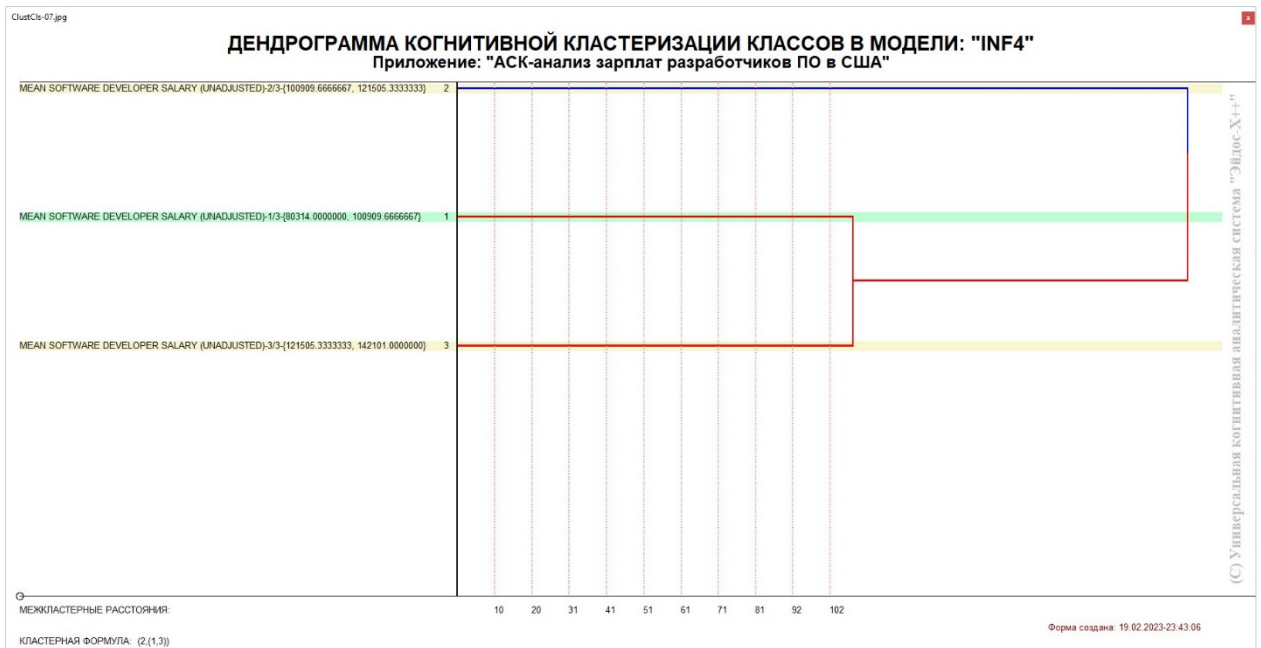


Рисунок 24 – Дендрограмма когнитивной кластеризации классов

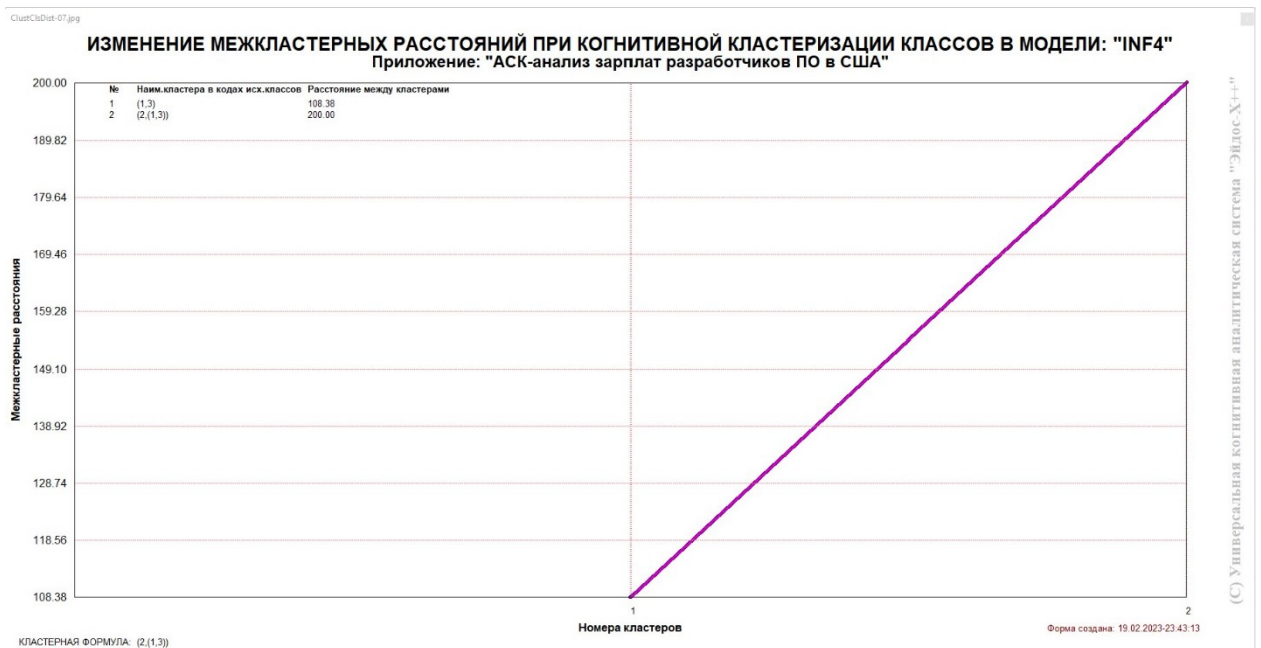


Рисунок 25 – График изменения межкластерных расстояний при когнитивной кластеризации классов

Из рисунка 24 мы видим, что некоторые характеристики сходны по системе детерминирующих их факторов, и, следовательно, могут наблюдаться одновременно, а другие по этой системе обуславливающих

факторов сильно отличаются, и, следовательно, являются взаимоисключающими, т.е. являются альтернативными.

### 2.3.3 Когнитивные диаграммы значений факторов

Подобные диаграммы отражают зависимость параметров. Эти диаграммы мы получаем в режимах 4.3.2.1 и 4.3.2.2 (рисунок 26).

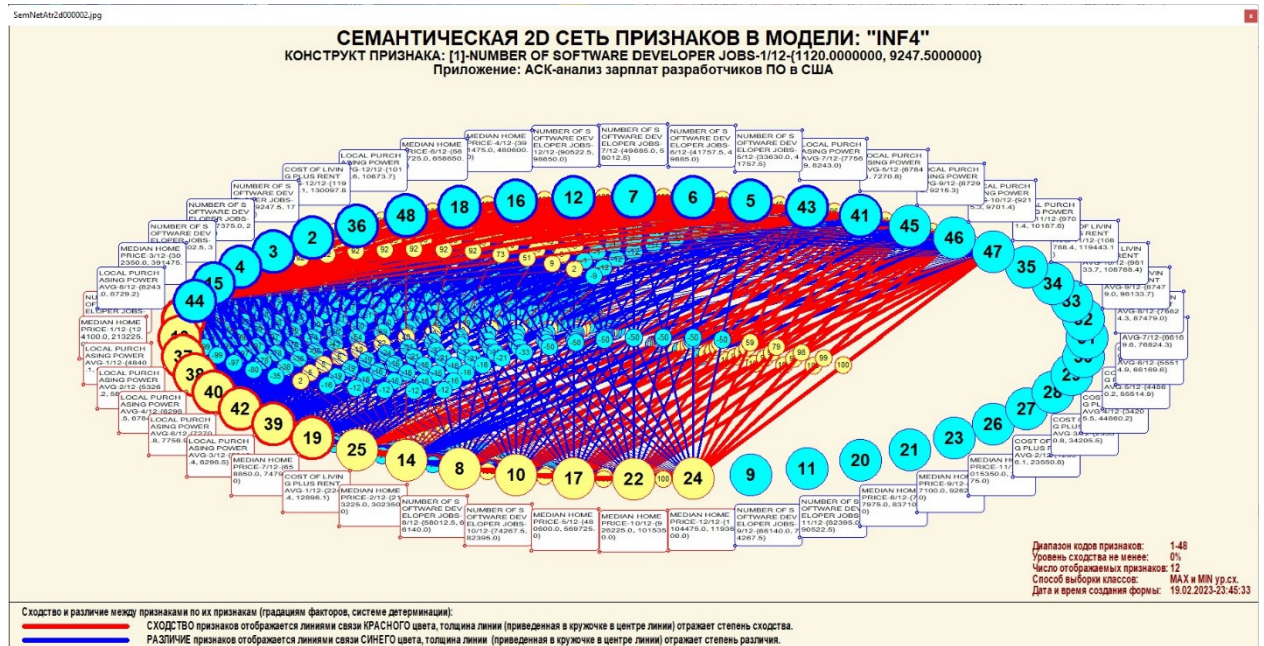


Рисунок 26 – График изменения межкластерных расстояний

### 2.3.4 Агломеративная когнитивная кластеризация значений факторов

На рисунках 27-28 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.



### ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF4" Приложение: "АСК-анализ зарплат разработчиков ПО в США"

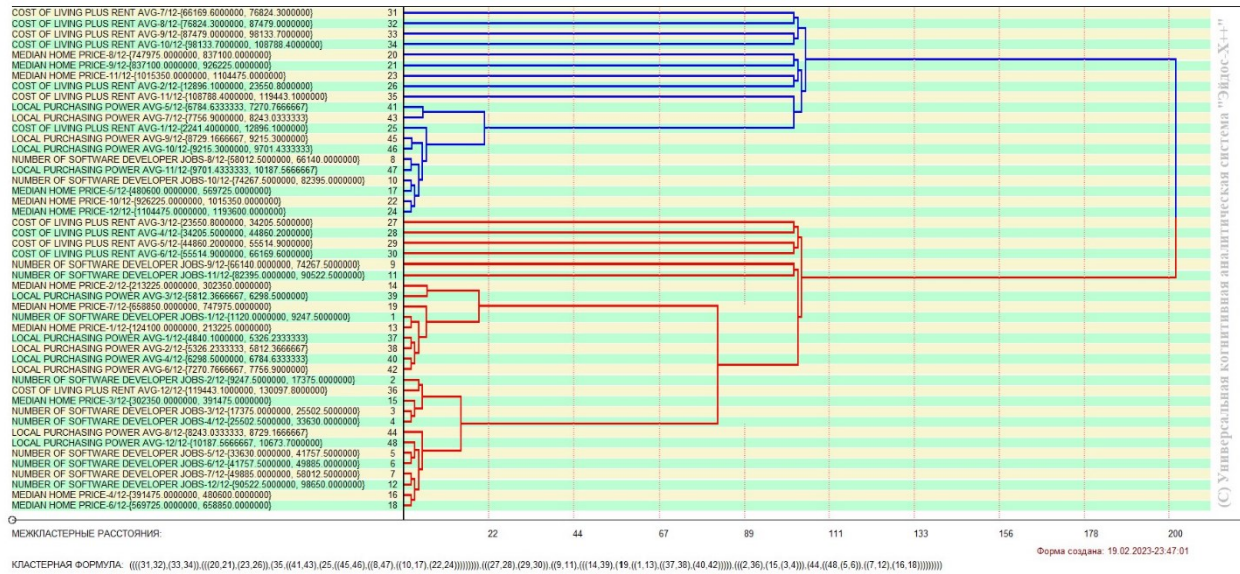


Рисунок 27 – Дендрограмма когнитивной кластеризации признаков

### ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF4" Приложение: "АСК-анализ зарплат разработчиков ПО в США"

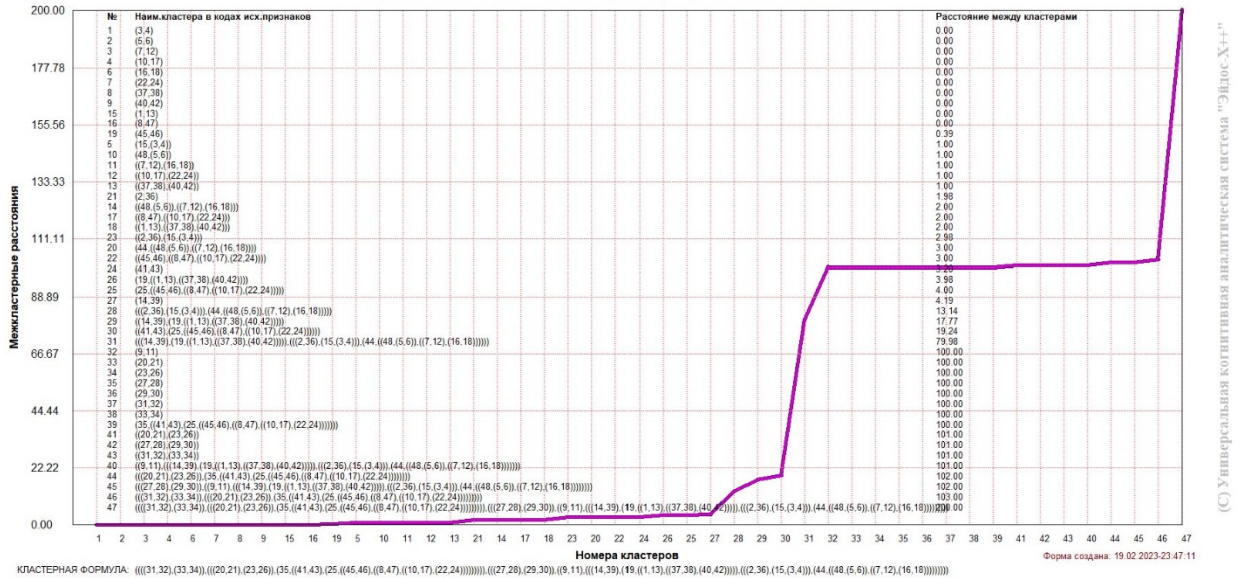


Рисунок 28 – График изменения межкластерных расстояний при когнитивной кластеризации признаков

Из дендрограммы мы видим, что все значения факторов образуют 2 четко выраженных кластера, объединенных в полюса конструкта (показаны синими и красными цветами). Хорошо видна группировка признаков по их смыслу.

### 2.3.5 Нелокальные нейроны и нелокальные нейронные сети

На рисунке 29 приведён пример нелокального нейрона, а на рисунке 30 – фрагмент одного слоя нелокальной нейронной сети, отражающий 87% наиболее значимых синаптических связей.

В приведенном фрагменте слоя нейронной сети нейроны соответствуют уровню заработной платы, а рецепторы – параметрам, влияющим на ее размер. 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.



Рисунок 29 – Пример нелокального нейрона

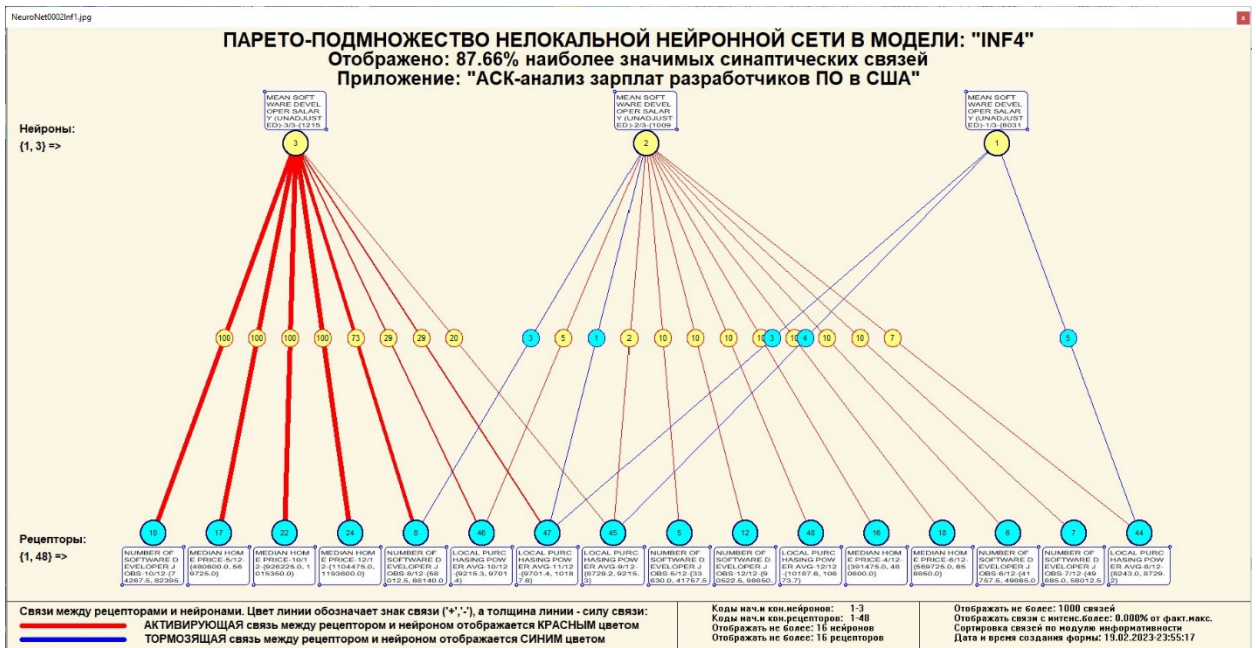


Рисунок 30 – Один слой нелокальной нейронной сети

### 2.3.6 3D-интегральные когнитивные карты

На рисунке 31 приведен фрагмент 3D-интегральной когнитивной карты, отражающий фрагмент СК-модели INF4.

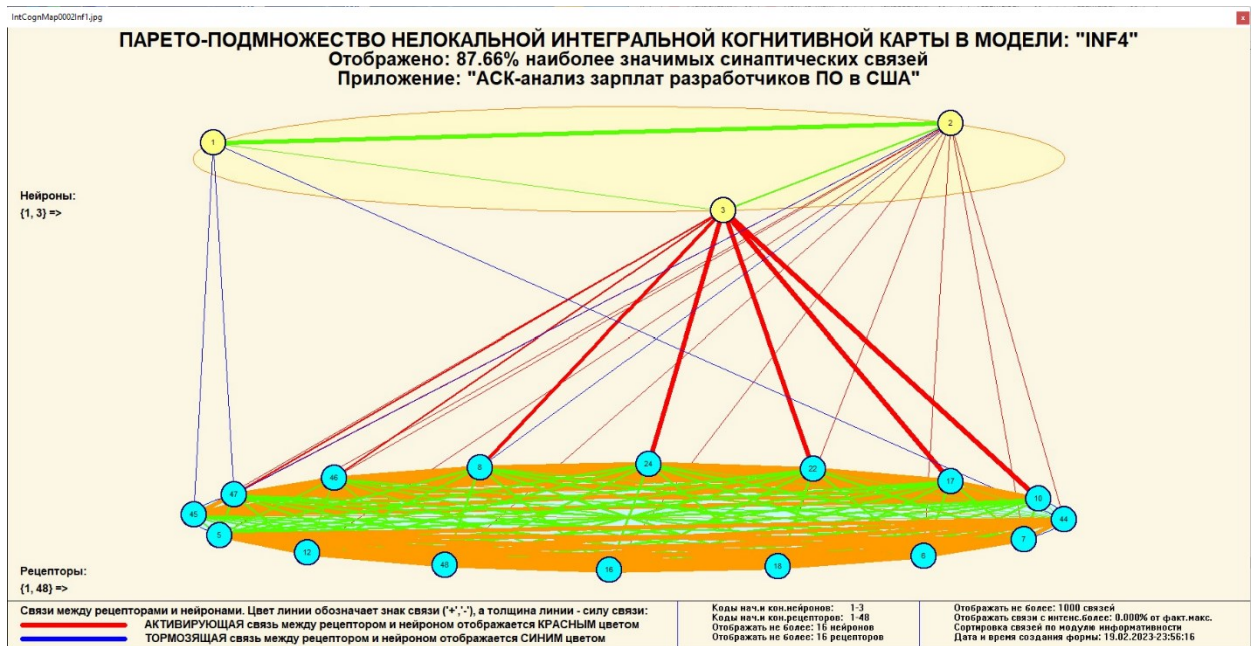


Рисунок 31 – 3D-интегральная когнитивная карта

### 2.3.7 Когнитивные функции

Такие диаграммы отражают то, как влияют градации различных факторов на прогноз. Эти диаграммы мы получаем в режиме 4.5 «Визуализация когнитивных функций» и некоторые из них представлены на рисунках 32, 33, 34, 35.

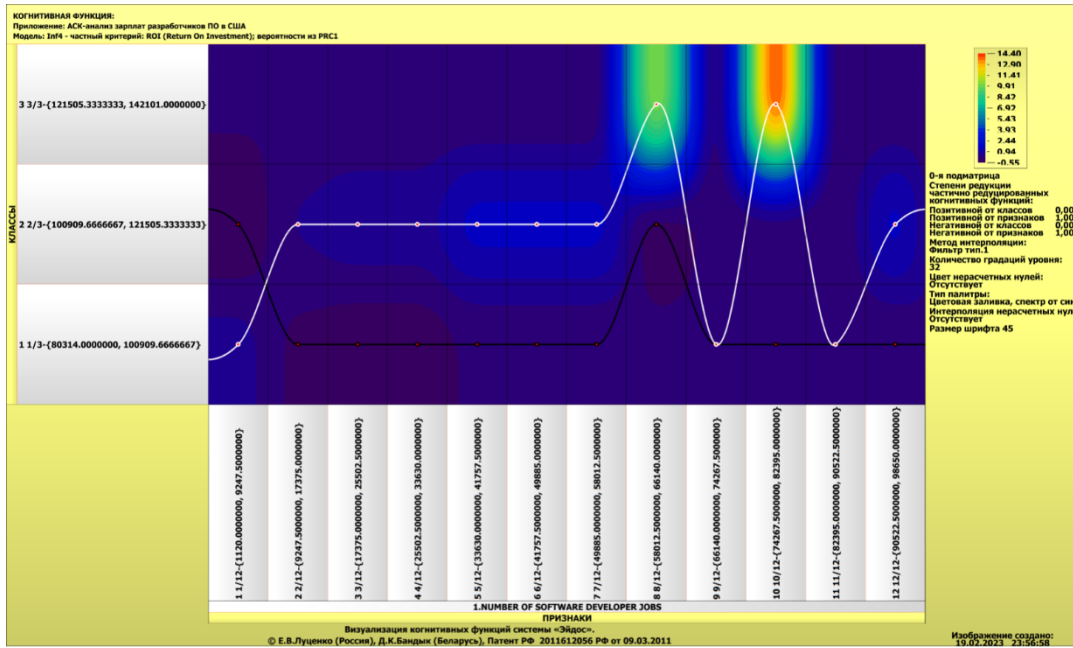


Рисунок 32 – Пример визуализации когнитивной функции по признаку 1

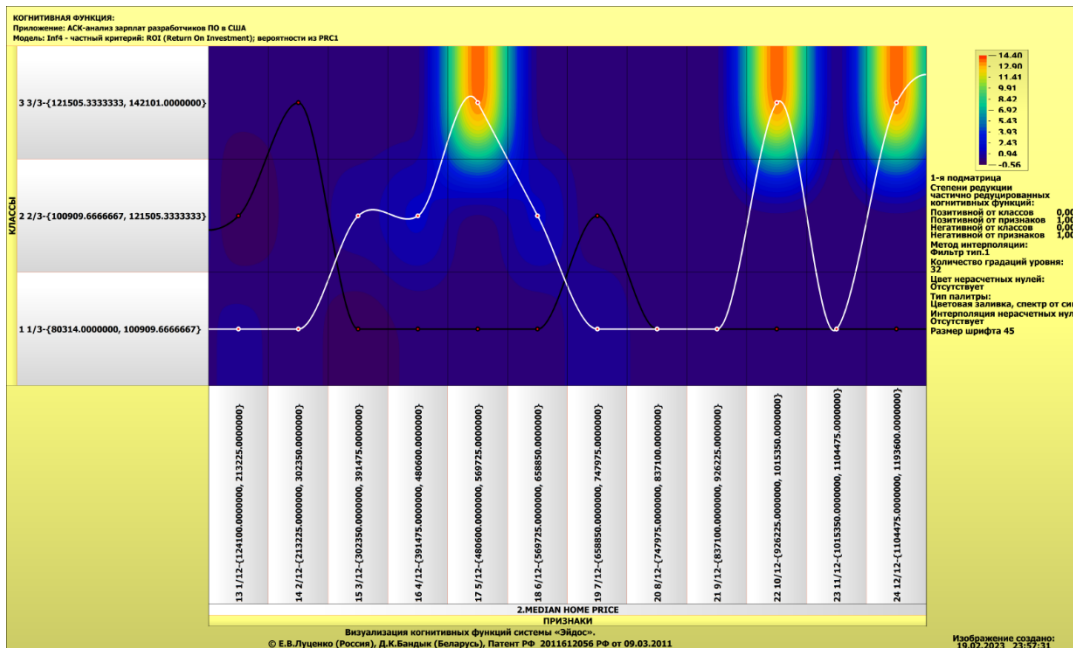


Рисунок 33 – Пример визуализации когнитивной функции по признаку 2

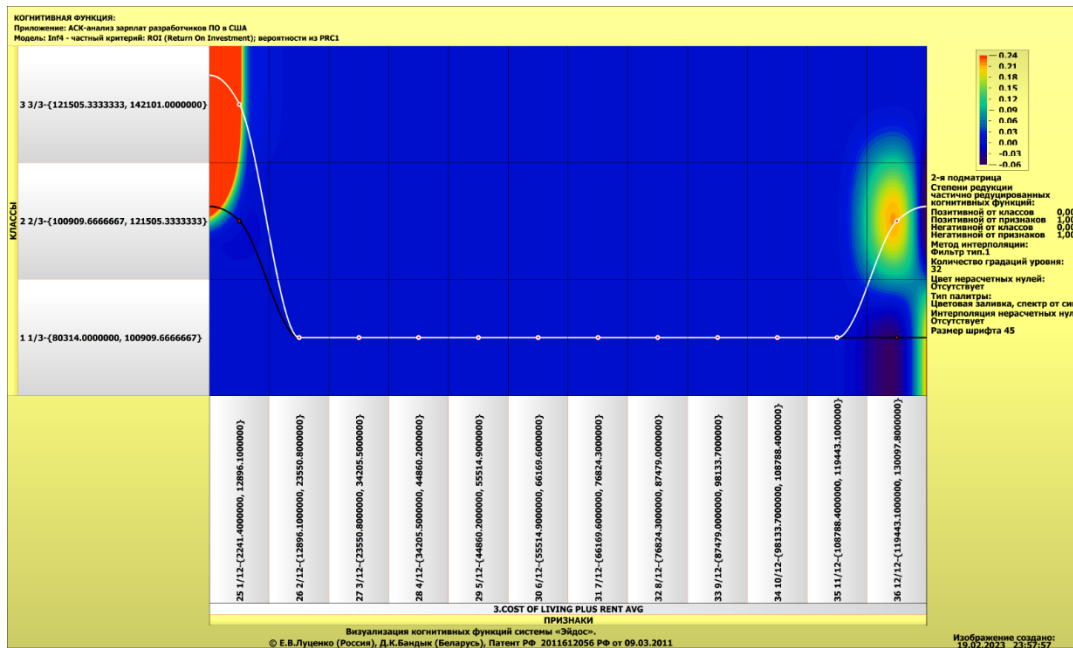


Рисунок 34 – Пример визуализации когнитивной функции по признаку 3

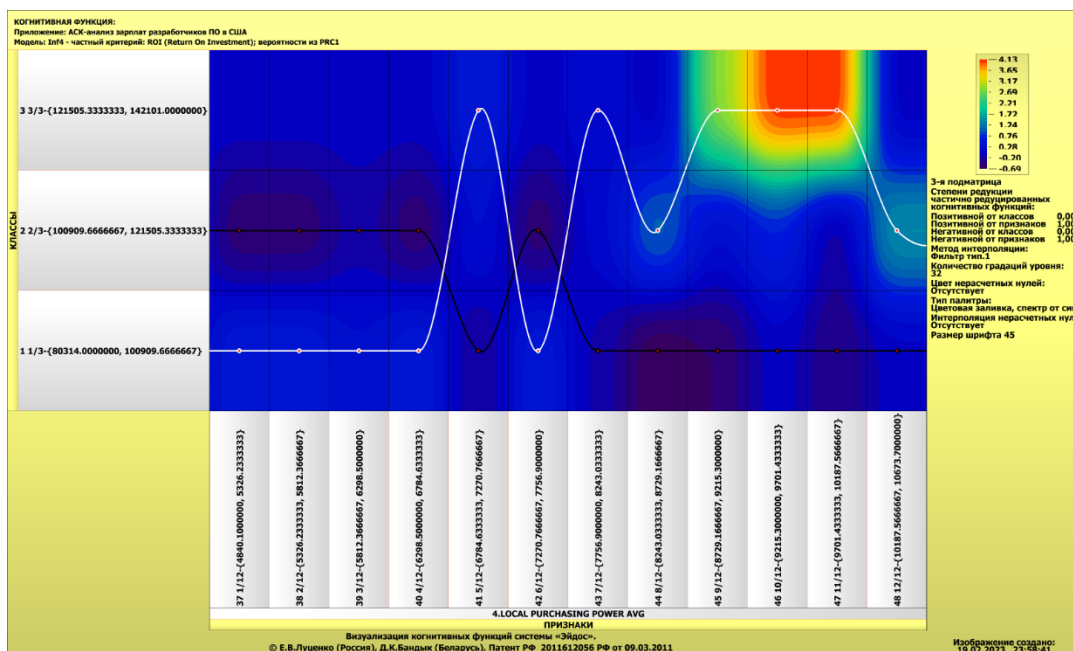


Рисунок 35 – Пример визуализации когнитивной функции по признаку 4

Зависимости между значениями параметров внешнего мира и показателями заработной платы в основном имеют довольно сложный характер, в чем можно наглядно убедиться из приведенных когнитивных функций.

## ЗАКЛЮЧЕНИЕ

Цель работы достигнута, поставленная проблема решена, что показывает анализ результатов численного эксперимента предложенных и реализованных в системе «Aidos-X», решение поставленных задач является вполне эффективным.

При помощи системы «Aidos-X» результатом реализованной работы являются созданные статистические и системно-когнитивные модели, в которых на основе эмпирических данных были сформированы обобщенные образы классов, показывающие предполагаемый уровень заработной платы, и на основе этого, были достигнуты такие задачи, как идентификация, классификация и исследование моделируемой предметной области путем исследования ее модели.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос» / Е.В. Луценко, Е.К. Печурина, А.Э. Сергеев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №06(160). С. 95 – 114. – IDA [article ID]: 1602006009. – Режим доступа: <http://ej.kubagro.ru/2020/06/pdf/09.pdf>, 1,25 у.п.л.

2. 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.

3. Луценко Е.В., Подсистема агломеративной когнитивной кластеризации классов системы «Эйдос» ("Эйдос-кластер"). Пат. № 2012610135 РФ. Заяв. № 2011617962 РФ 26.10.2011. Опубл. От 10.01.2012. – Режим доступа: <http://lc.kubagro.ru/aidos/2012610135.jpg>, 3,125 у.п.л.

4. 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.

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

6. 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.

7. Луценко Е.В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №01(001). С. 79 – 91. – IDA [article ID]: 0010301011. – Режим доступа: <http://ej.kubagro.ru/2003/01/pdf/11.pdf>, 0,812 у.п.л.

8. Луценко Е.В. Метод визуализации когнитивных функций – новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного



аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №03(067). С. 240 – 282. – Шифр Информрегистра: 0421100012\0077, IDA [article ID]: 0671103018. – Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>, 2,688 у.п.л.

9. 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.

10. Луценко Е.В. Детальный численный пример сценарного Автоматизированного системно-когнитивного анализа в интеллектуальной системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2020. – №08(162). С. 273 – 355. – IDA [article ID]: 1622008020. – Режим доступа: <http://ej.kubagro.ru/2020/08/pdf/20.pdf>, 5,188 у.п.л.