

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 специалистов по данным»

Выполнил студент группы: ИТ2141 Карокозьян Андрей 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: 34 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 специалистов по данным.

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

ВВЕДЕНИЕ .....	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 .....	31
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ .....	32

## 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 специалистов по данным 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 365, 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:

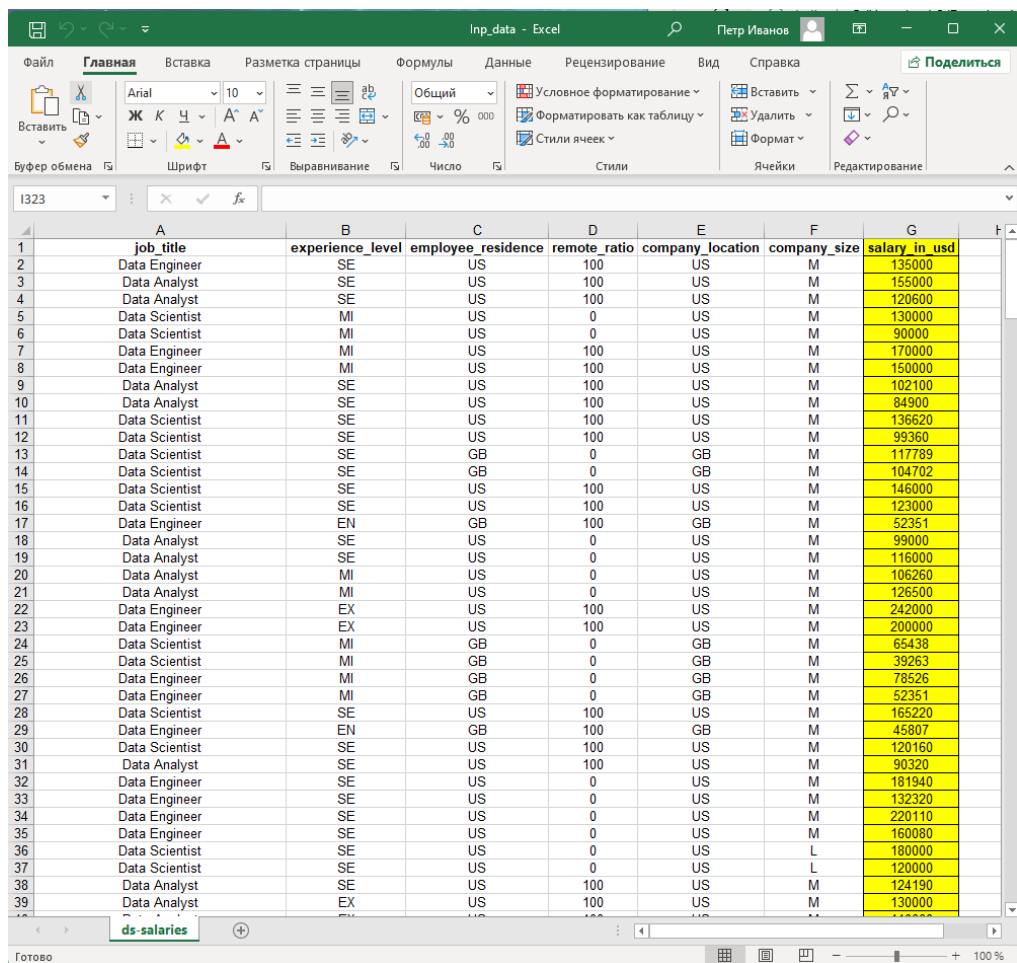
1. Поиск и объединение исходных данных в единую таблицу и сохранение в файле MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы Aidos-X.
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования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

- Job\_title – название вакансии;
- Experience\_level – уровень навыков;
- Employee\_residence – гражданство;
- Remote\_ratio – возможность удаленной работы;
- Company\_location – местоположение компании;
- Company\_size – размер компании;
- Salary\_in\_usd – уровень зарплаты;

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



	A	B	C	D	E	F	G
	job_title	experience_level	employee_residence	remote_ratio	company_location	company_size	salary_in_usd
1	Data Engineer	SE	US	100	US	M	135000
2	Data Analyst	SE	US	100	US	M	155000
3	Data Analyst	SE	US	100	US	M	120600
4	Data Scientist	MI	US	0	US	M	130000
5	Data Scientist	MI	US	0	US	M	90000
6	Data Engineer	MI	US	100	US	M	170000
7	Data Engineer	MI	US	100	US	M	150000
8	Data Analyst	SE	US	100	US	M	102100
9	Data Analyst	SE	US	100	US	M	84900
10	Data Scientist	SE	US	100	US	M	136620
11	Data Scientist	SE	US	100	US	M	99360
12	Data Scientist	SE	GB	0	GB	M	117789
13	Data Scientist	SE	GB	0	GB	M	104702
14	Data Scientist	SE	US	100	US	M	146000
15	Data Scientist	SE	US	100	US	M	123000
16	Data Engineer	EN	GB	100	GB	M	52351
17	Data Analyst	SE	US	0	US	M	99000
18	Data Analyst	SE	US	0	US	M	116000
19	Data Analyst	MI	US	0	US	M	106260
20	Data Analyst	MI	US	0	US	M	126500
21	Data Engineer	EX	US	100	US	M	242000
22	Data Engineer	EX	US	100	US	M	200000
23	Data Scientist	MI	GB	0	GB	M	65438
24	Data Scientist	MI	GB	0	GB	M	39263
25	Data Engineer	MI	GB	0	GB	M	78526
26	Data Engineer	MI	GB	0	GB	M	52351
27	Data Scientist	SE	US	100	US	M	165220
28	Data Engineer	EN	GB	100	GB	M	45807
29	Data Scientist	SE	US	100	US	M	120160
30	Data Analyst	SE	US	100	US	M	90320
31	Data Engineer	SE	US	0	US	M	181940
32	Data Engineer	SE	US	0	US	M	132320
33	Data Engineer	SE	US	0	US	M	220110
34	Data Engineer	SE	US	0	US	M	160080
35	Data Scientist	SE	US	0	US	L	180000
36	Data Scientist	SE	US	0	US	L	120000
37	Data Analyst	SE	US	100	US	M	124190
38	Data Analyst	EX	US	100	US	M	130000
39							

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).

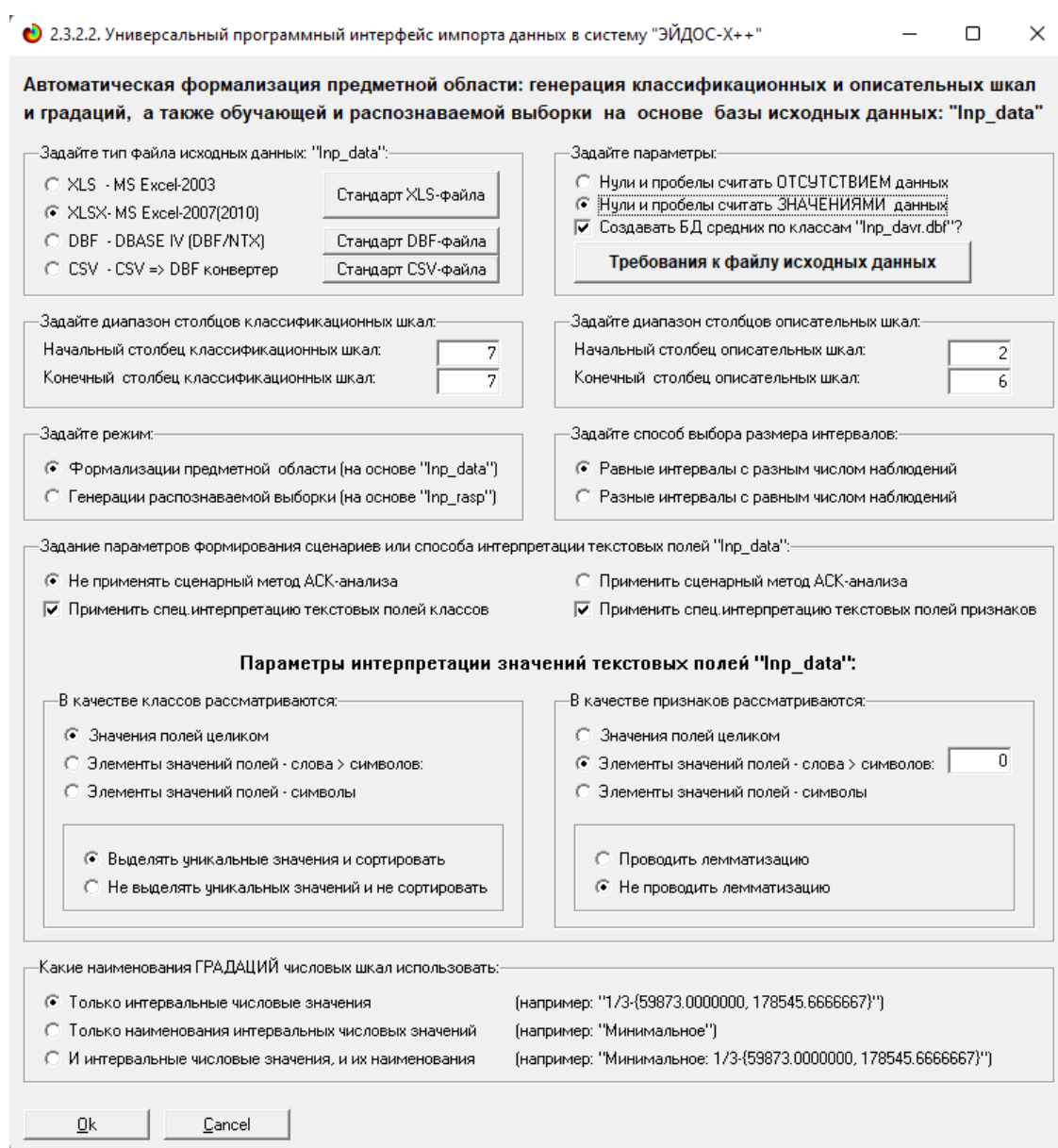


Рисунок 2 – Интерфейс импорта данных



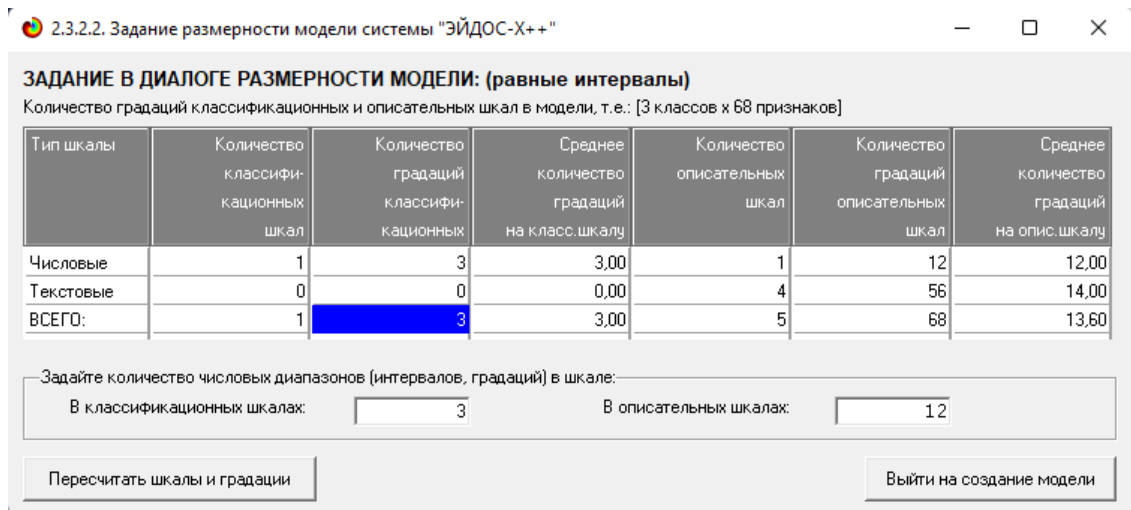


Рисунок 3 – Задание размерности модели

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

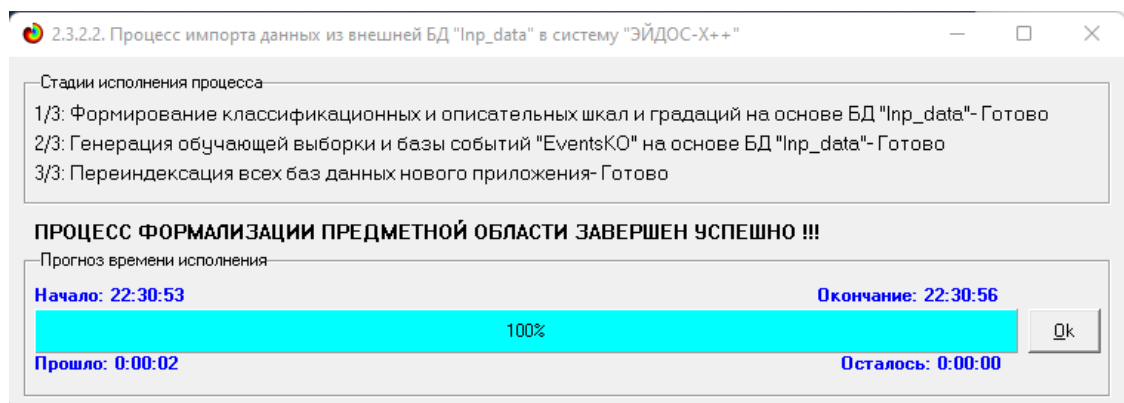


Рисунок 4 – Процесс импорта данных

Формирование классификационных и описательных 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).

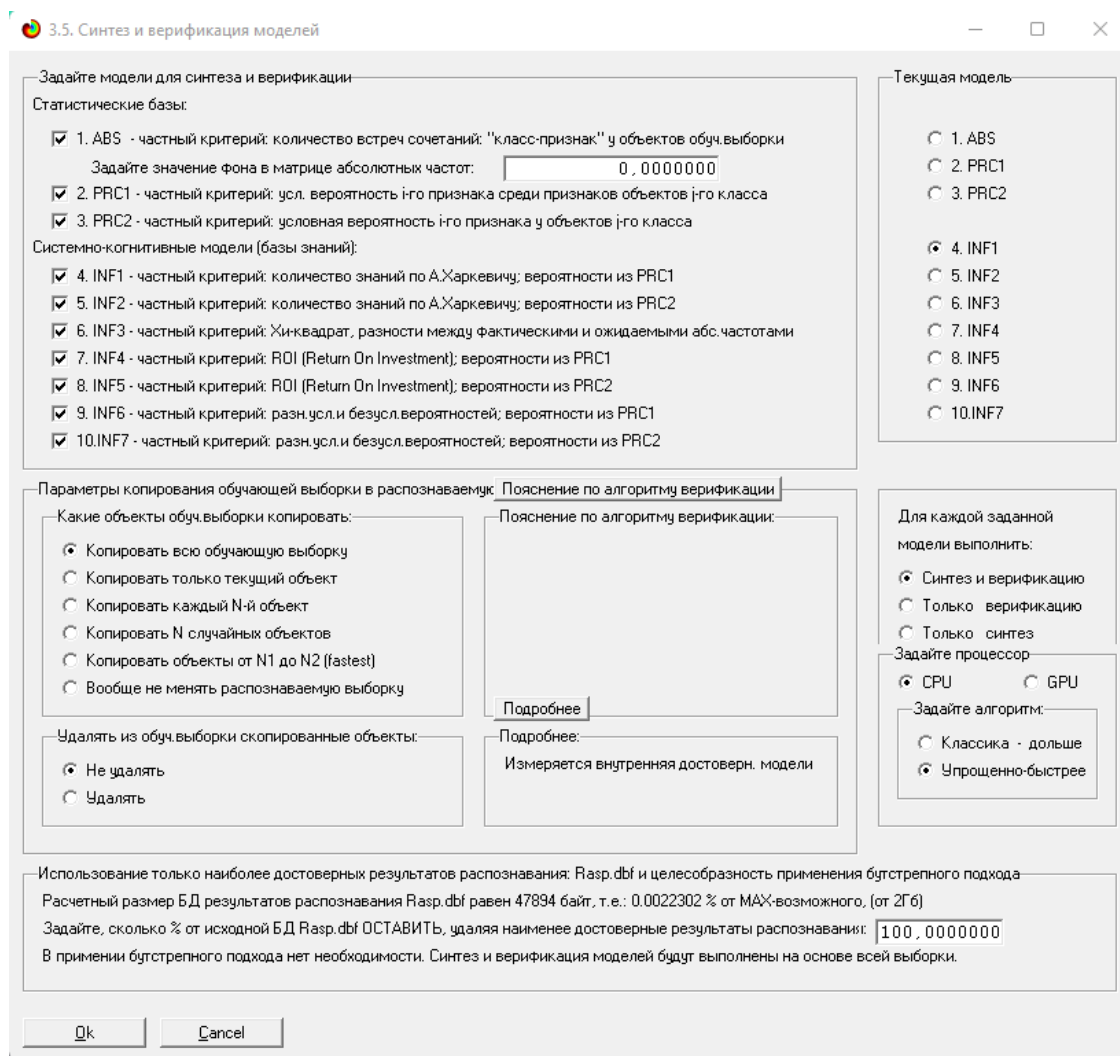


Рисунок 5 – Выбор моделей для синтеза и верификации

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.

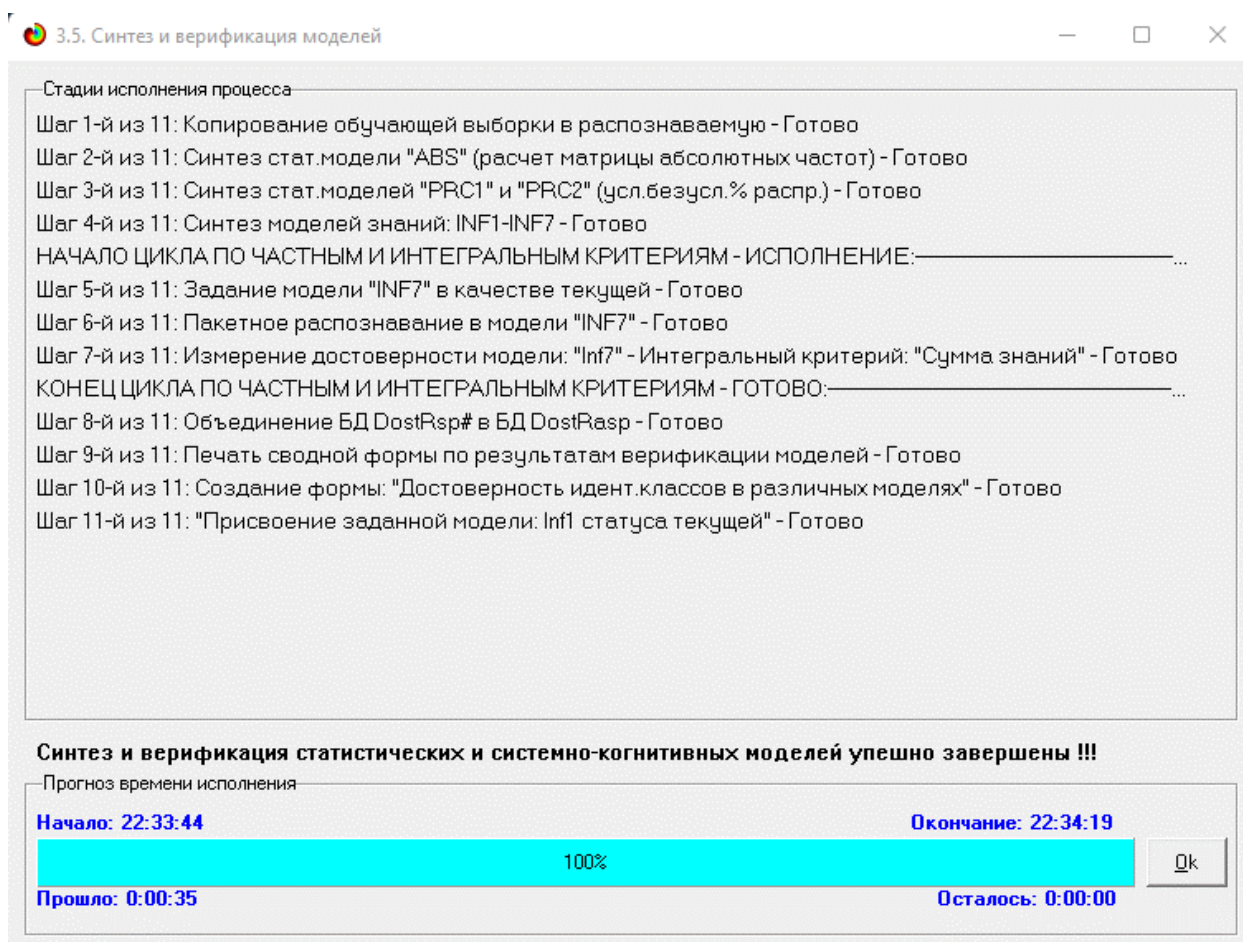


Рисунок 6 – Процесс исполнения режима 3.5

Стоит отметить, что синтез и верификация всех 10 моделей на данной задаче заняли 35 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. SALARY_IN_USD 1/3 (10000.0000000, 141666.6666667)	2. SALARY_IN_USD 2/3 (141666.6666667, 273333.3333333)	3. SALARY_IN_USD 3/3 (273333.3333333, 405000.0000000)	Сумма	Среднее	Средн. квадр. откл.
1.0	EXPERIENCE_LEVEL-EN	20.0	1.0		21.0	7.00	11.27
2.0	EXPERIENCE_LEVEL-EX	5.0	7.0	1.0	13.0	4.33	3.06
3.0	EXPERIENCE_LEVEL-MI	78.0	13.0		91.0	30.33	41.79
4.0	EXPERIENCE_LEVEL-SE	107.0	84.0	2.0	193.0	64.33	55.19
5.0	EMPLOYEE_RESIDENCE-AE	2.0			2.0	0.67	1.15
6.0	EMPLOYEE_RESIDENCE-AR	1.0			1.0	0.33	0.58
7.0	EMPLOYEE_RESIDENCE-AT	1.0			1.0	0.33	0.58
8.0	EMPLOYEE_RESIDENCE-AU	2.0	1.0		3.0	1.00	1.00
9.0	EMPLOYEE_RESIDENCE-BO	1.0			1.0	0.33	0.58
10.0	EMPLOYEE_RESIDENCE-BR	1.0			1.0	0.33	0.58
11.0	EMPLOYEE_RESIDENCE-CA	17.0	2.0		19.0	6.33	9.29
12.0	EMPLOYEE_RESIDENCE-CH	1.0			1.0	0.33	0.58
13.0	EMPLOYEE_RESIDENCE-DE	2.0	1.0		3.0	1.00	1.00
14.0	EMPLOYEE_RESIDENCE-DZ	1.0			1.0	0.33	0.58
15.0	EMPLOYEE_RESIDENCE-EE	1.0			1.0	0.33	0.58
16.0	EMPLOYEE_RESIDENCE-ES	6.0			6.0	2.00	3.46
17.0	EMPLOYEE_RESIDENCE-FR	3.0			3.0	1.00	1.73
18.0	EMPLOYEE_RESIDENCE-GB	28.0	1.0		29.0	9.67	15.89
19.0	EMPLOYEE_RESIDENCE-GR	9.0			9.0	3.00	5.20
20.0	EMPLOYEE_RESIDENCE-IE	1.0			1.0	0.33	0.58
21.0	EMPLOYEE_RESIDENCE-IN	4.0	1.0		5.0	1.67	2.08
22.0	EMPLOYEE_RESIDENCE-JP	1.0			1.0	0.33	0.58
23.0	EMPLOYEE_RESIDENCE-MY		1.0		1.0	0.33	0.58
24.0	EMPLOYEE_RESIDENCE-NL	1.0			1.0	0.33	0.58

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. SALARY_IN_USD 1/3 (10000.0000000, 141666.6666667)	2. SALARY_IN_USD 2/3 (141666.6666667, 273333.3333333)	3. SALARY_IN_USD 3/3 (273333.3333333, 405000.0000000)	Сумма	Среднее	Средн. квадр. откл.
1.0	EXPERIENCE_LEVEL-EN	6.132	-5.934	-0.198			6.035
2.0	EXPERIENCE_LEVEL-EX	-3.585	2.708	0.877			3.237
3.0	EXPERIENCE_LEVEL-MI	17.906	-17.047	-0.858			17.492
4.0	EXPERIENCE_LEVEL-SE	-20.453	20.274	0.179			20.364
5.0	EMPLOYEE_RESIDENCE-AE	0.679	-0.660	-0.019			0.670
6.0	EMPLOYEE_RESIDENCE-AR	0.340	-0.330	-0.009			0.335
7.0	EMPLOYEE_RESIDENCE-AT	0.340	-0.330	-0.009			0.335
8.0	EMPLOYEE_RESIDENCE-AU	0.019	0.009	-0.028			0.025
9.0	EMPLOYEE_RESIDENCE-BO	0.340	-0.330	-0.009			0.335
10.0	EMPLOYEE_RESIDENCE-BR	0.340	-0.330	-0.009			0.335
11.0	EMPLOYEE_RESIDENCE-CA	4.453	-4.274	-0.179			4.366
12.0	EMPLOYEE_RESIDENCE-CH	0.340	-0.330	-0.009			0.335
13.0	EMPLOYEE_RESIDENCE-DE	0.019	0.009	-0.028			0.025
14.0	EMPLOYEE_RESIDENCE-DZ	0.340	-0.330	-0.009			0.335
15.0	EMPLOYEE_RESIDENCE-EE	0.340	-0.330	-0.009			0.335
16.0	EMPLOYEE_RESIDENCE-ES	2.038	-1.981	-0.057			2.010
17.0	EMPLOYEE_RESIDENCE-FR	1.019	-0.991	-0.028			1.005
18.0	EMPLOYEE_RESIDENCE-GB	8.849	-8.575	-0.274			8.715
19.0	EMPLOYEE_RESIDENCE-GR	3.057	-2.972	-0.085			3.015
20.0	EMPLOYEE_RESIDENCE-IE	0.340	-0.330	-0.009			0.335
21.0	EMPLOYEE_RESIDENCE-IN	0.698	-0.651	-0.047			0.676
22.0	EMPLOYEE_RESIDENCE-JP	0.340	-0.330	-0.009			0.335
23.0	EMPLOYEE_RESIDENCE-MY	-0.660	0.670	-0.009			0.665
24.0	EMPLOYEE_RESIDENCE-NL	0.340	-0.330	-0.009			0.335

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. Обобщ.форма по достов.моделей при разн.инт.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	Фигура Ван Ризбергера	Сумма модулей уровней сход. истинно-отриц. решений (STR)	Сумма модулей уровней сход. истинно-отриц. решений (STN)	Сумма модулей уровней сход. ложно-полож. решений
1.ABS - частный критерий: количество встреч сонетовых. "Улас"	Корреляция абс.частот с обр...	318	316	28	608	2	0.342	0.994	0.509	233.910	0.245	416
1.ABS - частный критерий: количество встреч сонетовых. "Улас"	Сумма абс.частот по призна...	318	318	636	636	0.333	1.000	0.500	219.952	0.245	396	
2.PRC1 - частный критерий: усл. вероятность fго признака сред.	Корреляция усл.отн.частот с о...	318	316	28	608	2	0.342	0.994	0.509	233.910	0.245	416
2.PRC1 - частный критерий: усл. вероятность fго признака сред.	Сумма усл.отн.частот по при...	318	318	636	636	0.333	1.000	0.500	209.992	0.245	420	
3.PRC2 - частный критерий: условная вероятность fго признака.	Корреляция усл.отн.частот с о...	318	318	636	636	0.333	1.000	0.500	233.907	0.245	416	
3.PRC2 - частный критерий: условная вероятность fго признака.	Сумма усл.отн.частот по при...	318	318	636	636	0.333	1.000	0.500	208.992	0.245	420	
4.INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	318	107	434	202	0.211	0.346	0.336	0.341	29.330	144.949	62
4.INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	318	196	327	309	0.122	0.388	0.616	0.476	27.178	60.290	39
5.INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	318	107	434	202	0.211	0.346	0.336	0.341	29.330	144.949	62
5.INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	318	196	327	309	0.122	0.388	0.616	0.476	27.178	60.290	39
6.INF3 - частный критерий: Хинкварат, равенности между фактик.	Семантический резонанс зна...	318	171	266	350	0.147	0.328	0.538	0.408	107.292	137.848	169
6.INF3 - частный критерий: Хинкварат, равенности между фактик.	Семантический резонанс зна...	318	171	266	350	0.147	0.328	0.538	0.408	102.078	106.959	106
7.INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	318	107	456	180	0.211	0.373	0.336	0.354	32.917	122.272	58
7.INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	318	209	271	365	0.109	0.364	0.657	0.469	20.455	18.679	48
8.INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	318	107	456	180	0.211	0.373	0.336	0.354	32.917	122.272	58
8.INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	318	209	271	365	0.109	0.364	0.657	0.469	20.455	18.679	48
9.INF6 - частный критерий: разн.усли и безусл.вероятностей, вер...	Семантический резонанс зна...	318	171	240	396	0.147	0.302	0.538	0.386	105.766	123.623	173
9.INF6 - частный критерий: разн.усли и безусл.вероятностей, вер...	Семантический резонанс зна...	318	171	239	397	0.147	0.301	0.538	0.386	54.097	46.813	145
10.INF7 - частный критерий: разн.усли и безусл.вероятностей, ве...	Семантический резонанс зна...	318	171	240	396	0.147	0.302	0.538	0.386	105.766	123.623	173
10.INF7 - частный критерий: разн.усли и безусл.вероятностей, ве...	Семантический резонанс зна...	318	171	239	397	0.147	0.301	0.538	0.386	54.097	46.813	145

Рисунок 9 – Оценка достоверности моделей (F-критерий)

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

Наименование модели и частного критерия	Интегральный критерий	С-Полнота модели	L1-индекс проф. Е.В. Луценко	Средняя модуль уровней сход. истинно-полож. решений	Средний модуль уровней сход. истинноотрицат. решений	Средняя модуль уровней сход. ложно-положит. решений	Средний модуль уровней сход. ложноотрицат. решений	A-Точность модели A-Резонанс = ATR/AATP+	A-Полнота модели A-Ресал = ATR/AATP+	L2-индекс проф. Е.В. Луценко	Процент правильной классификации	
1.ABS - частный критерий: количество встреч сонетовых. "Улас"	Корреляция абс.частот с обр...	0.59	1.000	0.529	0.740	0.009	0.686	0.032	0.519	0.959	0.674	99.371
1.ABS - частный критерий: количество встреч сонетовых. "Улас"	Сумма абс.частот по призна...	0.28	1.000	0.691	0.692	0.310	0.691	1.000	0.691	1.000	0.817	100.000
2.PRC1 - частный критерий: усл. вероятность fго признака сред.	Корреляция усл.отн.частот с о...	0.59	1.000	0.529	0.740	0.009	0.686	0.032	0.519	0.959	0.674	99.371
2.PRC1 - частный критерий: усл. вероятность fго признака сред.	Сумма усл.отн.частот по при...	0.32	1.000	0.499	0.657	0.661	0.661	0.499	1.000	0.665	100.000	
3.PRC2 - частный критерий: условная вероятность fго признака.	Корреляция усл.отн.частот с о...	0.59	1.000	0.529	0.740	0.009	0.686	0.032	0.519	0.959	0.674	99.371
3.PRC2 - частный критерий: условная вероятность fго признака.	Сумма усл.отн.частот по при...	0.32	1.000	0.499	0.657	0.661	0.661	0.499	1.000	0.665	100.000	
4.INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.21	0.217	0.259	0.274	0.334	0.308	0.502	0.471	0.353	0.404	33.648
4.INF1 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.05	0.701	0.513	0.139	0.184	0.129	0.095	0.518	0.593	0.553	61.635
5.INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.21	0.217	0.259	0.274	0.334	0.308	0.502	0.471	0.353	0.404	33.648
5.INF2 - частный критерий: количество знаний по А.Харкевичу, в.	Семантический резонанс зна...	0.05	0.701	0.513	0.139	0.184	0.129	0.095	0.518	0.593	0.553	61.635
6.INF3 - частный критерий: Хинкварат, равенности между фактик.	Семантический резонанс зна...	0.87	0.514	0.442	0.627	0.482	0.485	0.691	0.564	0.476	0.516	53.774
6.INF3 - частный критерий: Хинкварат, равенности между фактик.	Семантический резонанс зна...	0.90	0.501	0.495	0.597	0.374	0.304	0.691	0.662	0.464	0.545	53.774
7.INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.62	0.236	0.286	0.308	0.268	0.323	0.505	0.488	0.379	0.427	33.648
7.INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.97	0.791	0.432	0.098	0.069	0.132	0.049	0.425	0.664	0.518	65.723
8.INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.62	0.236	0.286	0.308	0.268	0.323	0.505	0.488	0.379	0.427	33.648
8.INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.97	0.791	0.432	0.098	0.069	0.132	0.049	0.425	0.664	0.518	65.723
9.INF6 - частный критерий: разн.усли и безусл.вероятностей, вер...	Семантический резонанс зна...	0.79	0.510	0.435	0.619	0.515	0.438	0.692	0.585	0.472	0.523	53.774
9.INF6 - частный критерий: разн.усли и безусл.вероятностей, вер...	Семантический резонанс зна...	0.72	0.657	0.384	0.316	0.196	0.366	0.192	0.464	0.622	0.532	53.774
10.INF7 - частный критерий: разн.усли и безусл.вероятностей, ве...	Семантический резонанс зна...	0.79	0.510	0.435	0.619	0.515	0.438	0.692	0.585	0.472	0.523	53.774
10.INF7 - частный критерий: разн.усли и безусл.вероятностей, ве...	Семантический резонанс зна...	0.72	0.657	0.384	0.316	0.196	0.366	0.192	0.464	0.622	0.532	53.774

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 PRC1 и PRC2 с интегральным критерием «Корреляция условно относительных частот» (F=0,509 максимуме 1,000), по критерию L1 профессора Е.В. Луценко наиболее достоверной является СК-модель ABS с интегральным критерием «Сумма абсолютных частот» (L1=0,691 при максимуме 1,000), а по критерию L2 профессора Е.В. Луценко наиболее достоверной является СК-модель ABS с интегральным критерием «Сумма абсолютных частот» (L2=0,817 при максимуме 1,000). Для оценки достоверности моделей в АСК- анализе и системе Aidos-X используется F-

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

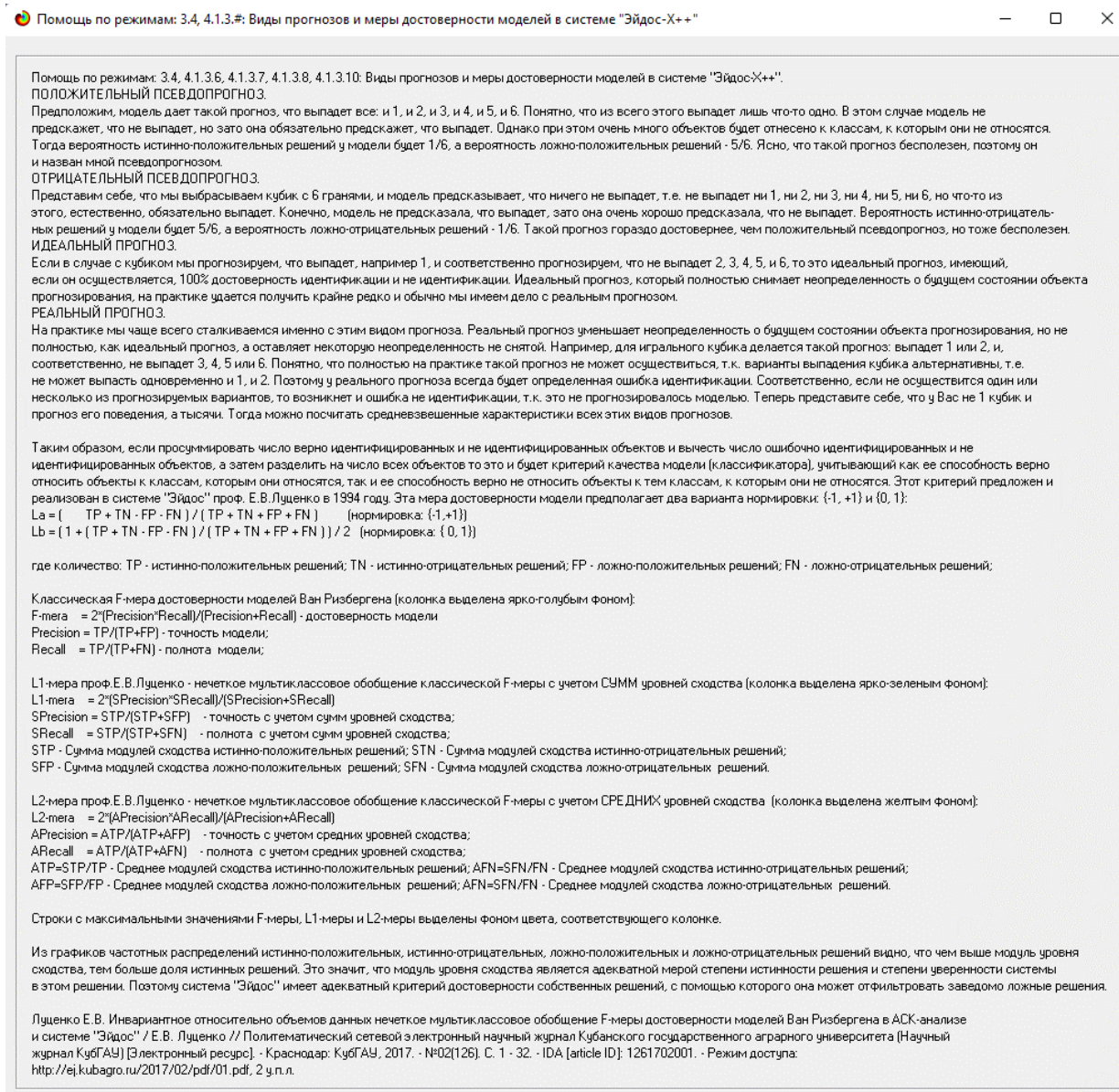


Рисунок 11 – Виды прогнозов и принципы определения

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 12-14 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.

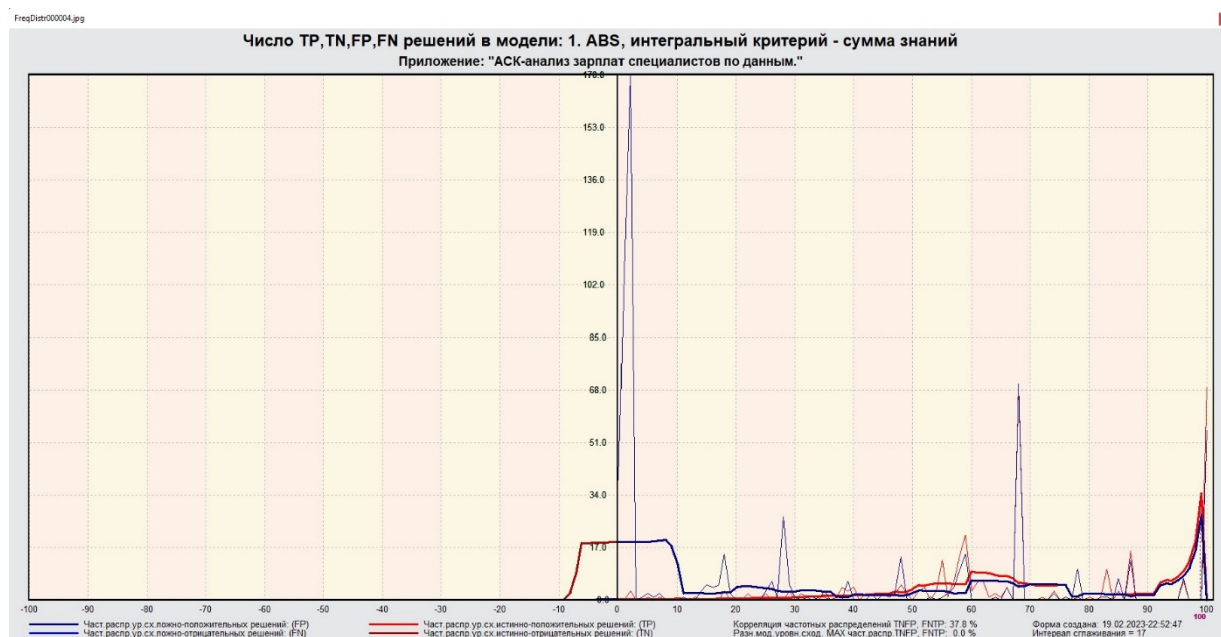


Рисунок 12 – Частотные распределения модели ABS

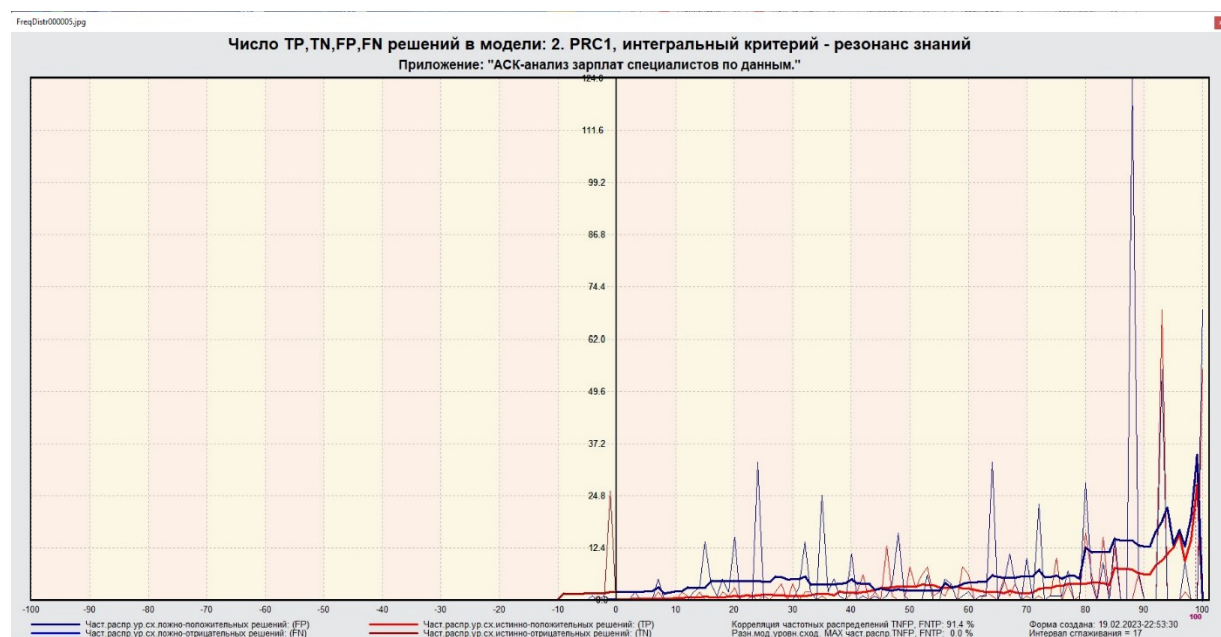


Рисунок 13 – Частотные распределения модели PRC1



Рисунок 14 – Частотные распределения модели PRC2

Видно, что модели PRC1 и PRC2 на основе интегрального критерия «резонанс знаний» при проценте схожести более 90% практически не дают ложноположительные решения. При этом модели ABS на основе интегрального критерия «сумма знаний» при проценте схожести меньше примерно 85% дают меньше ложноотрицательных, 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 ABS (режим 5.6) (рисунок 15, 16).

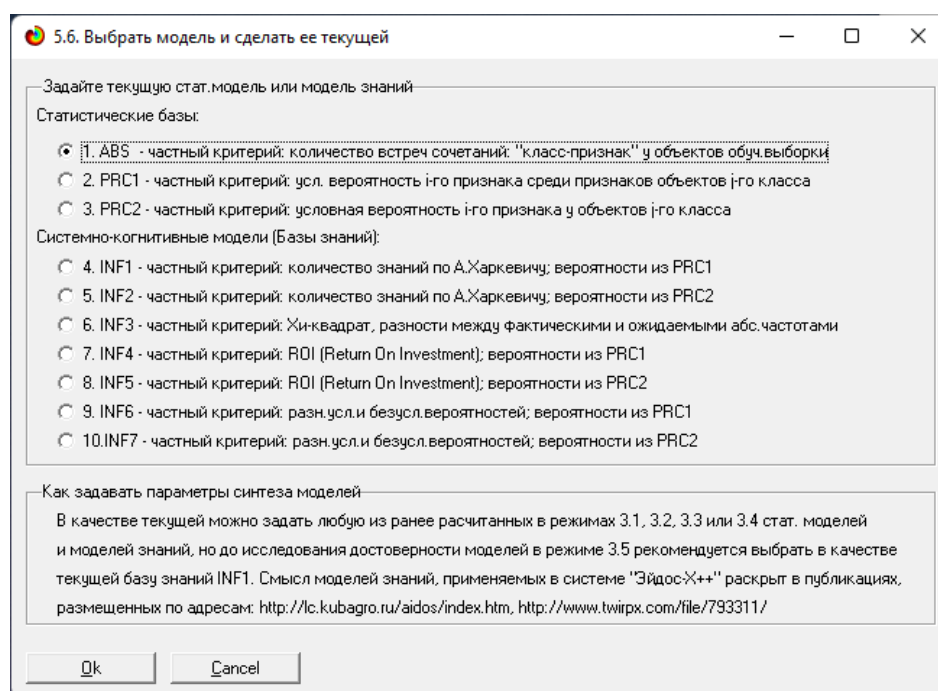


Рисунок 15 – Интерфейс задания модели

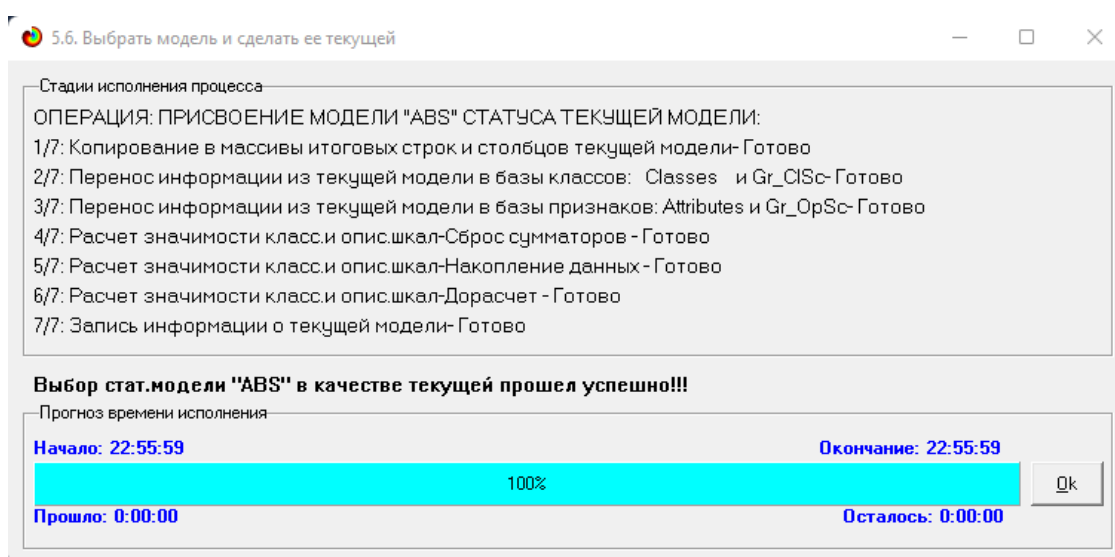


Рисунок 16 – Процесс загрузки модели

Затем следует провести пакетное распознавание в режиме 4.1.2. (рисунок 17).

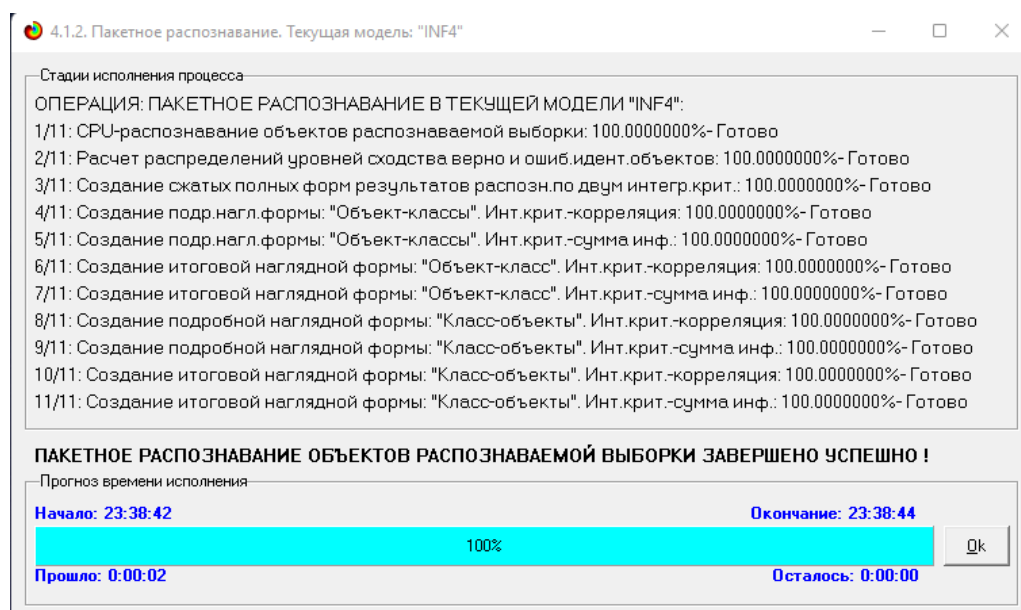


Рисунок 17 – Процесс пакетного распознавания модели

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

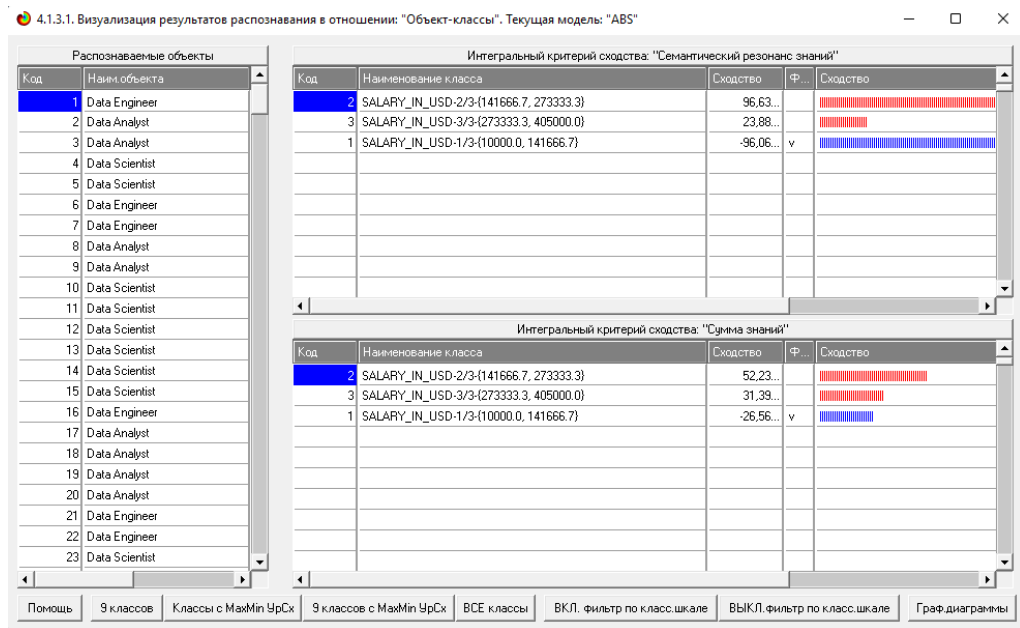


Рисунок 18 – Визуализация «Объект-классы»

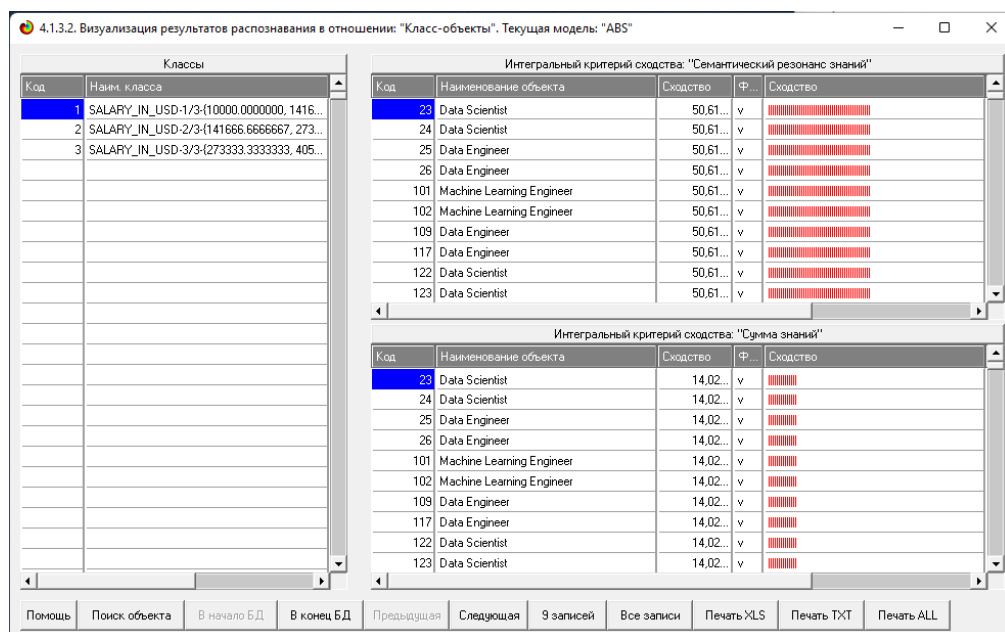


Рисунок 19 – Визуализация «Класс-объекты»

## 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 20).

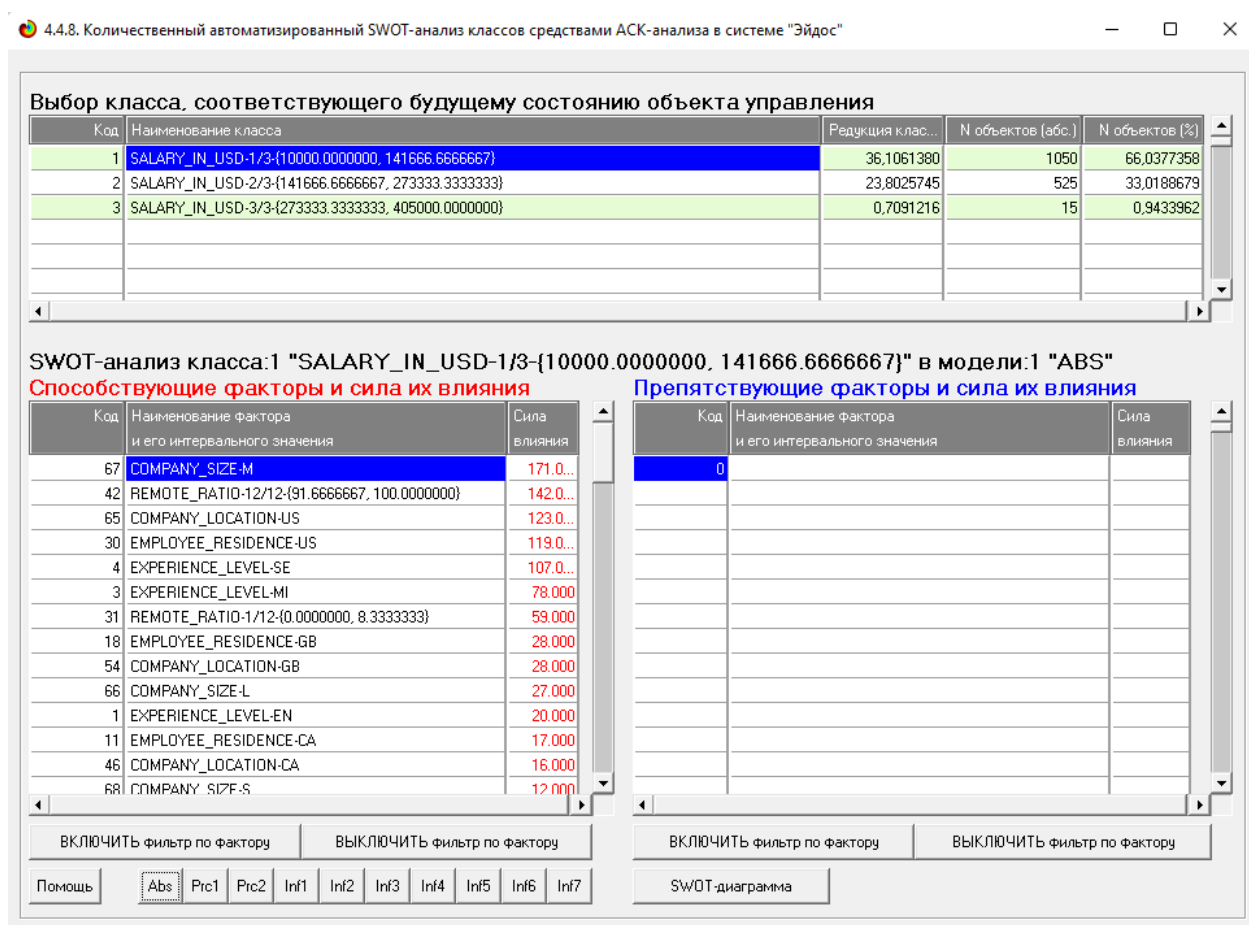


Рисунок 20 – SWOT-матрица

На рисунке 21 приведен пример табличной выходной формы количественного автоматизированного SWOT-анализа и PEST-анализа средствами системы Aidos-X.

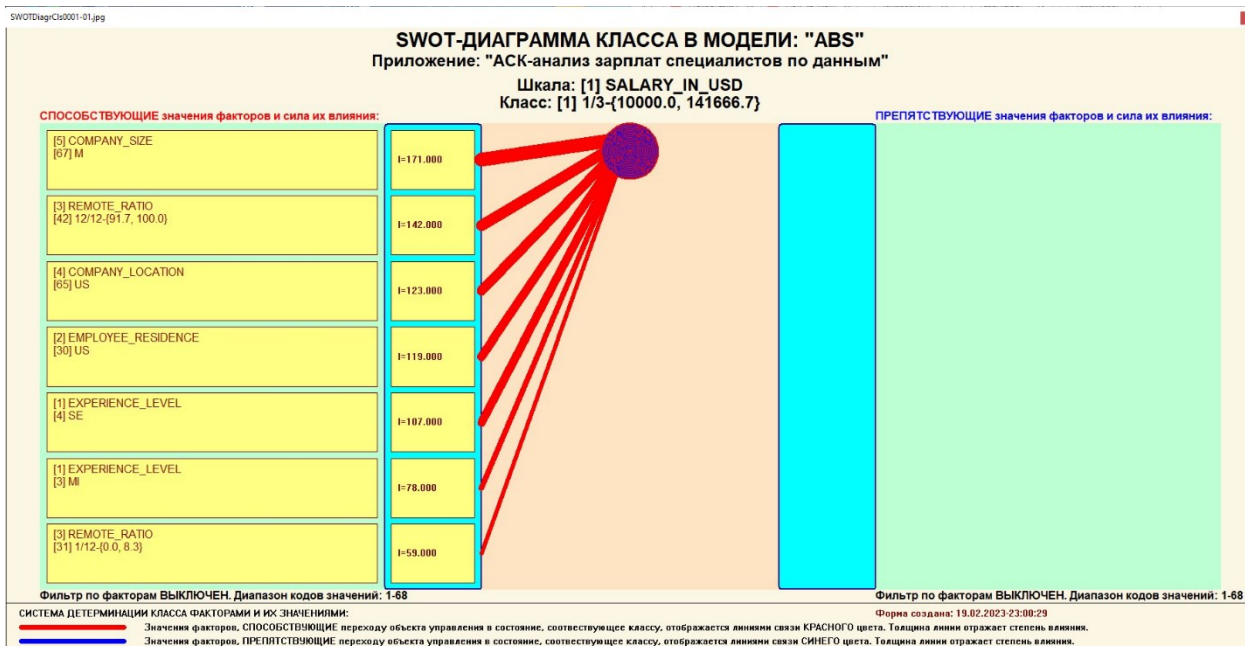


Рисунок 21 – SWOT-диаграмма

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 22, 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.

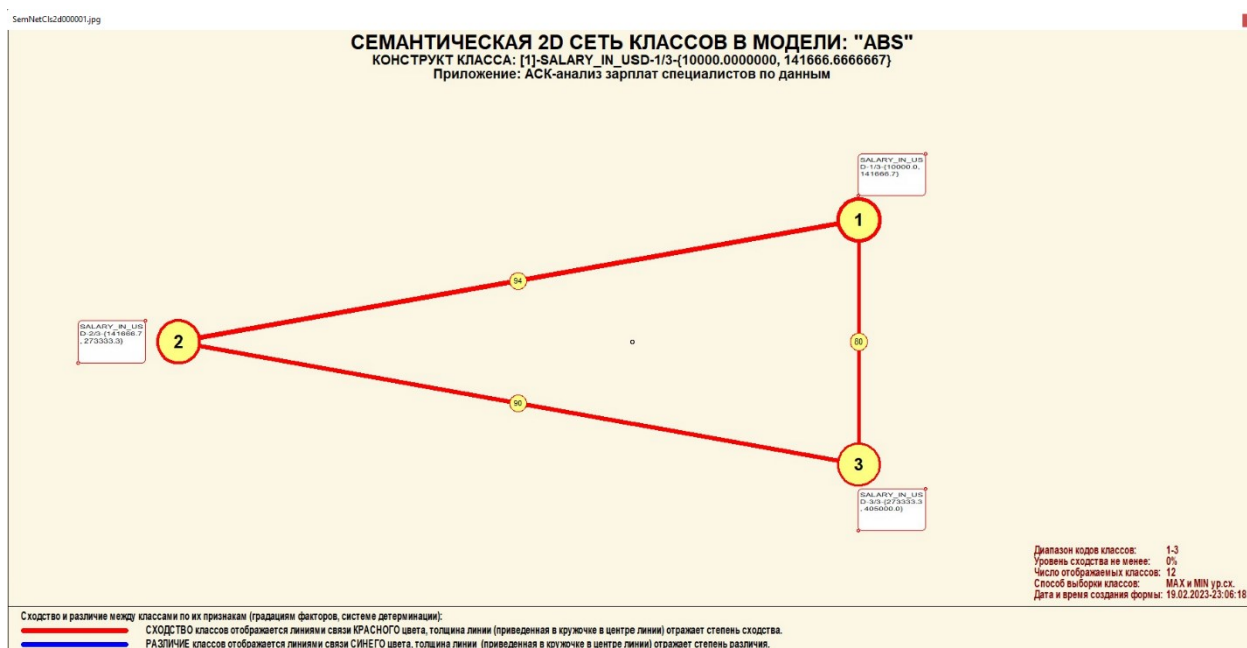


Рисунок 22 – Когнитивная диаграмма классов

### 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 23 и 24).

**ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ В МОДЕЛИ: "ABS"**  
**Приложение: "АСК-анализ зарплат специалистов по данным"**

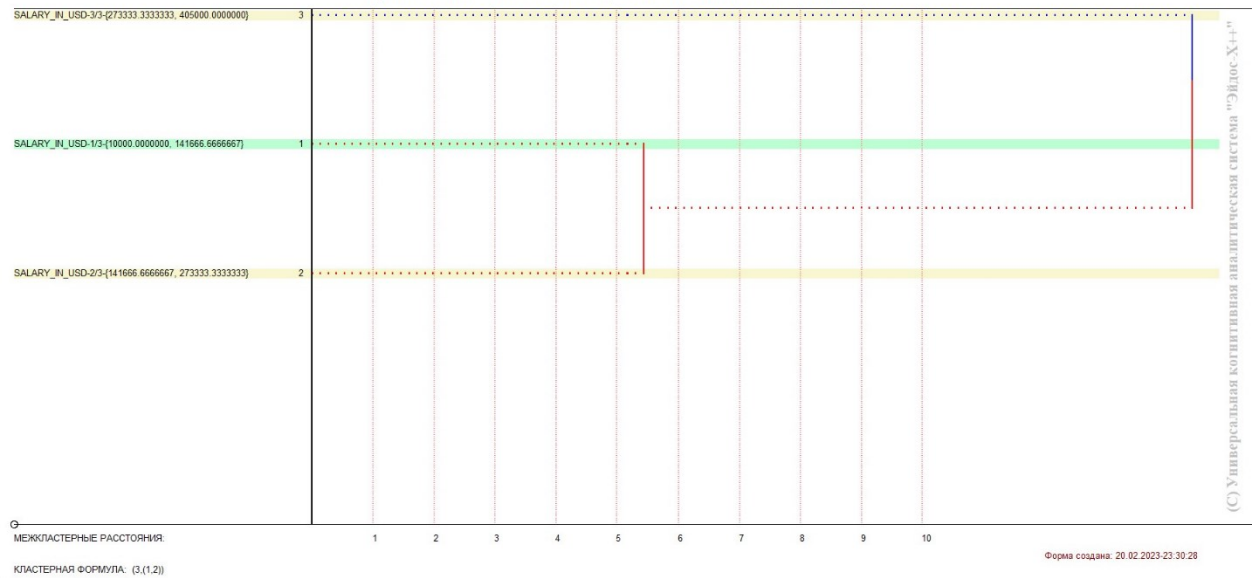


Рисунок 23 – Дендрограмма когнитивной кластеризации классов

**ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ В МОДЕЛИ: "ABS"**  
**Приложение: "АСК-анализ зарплат специалистов по данным"**

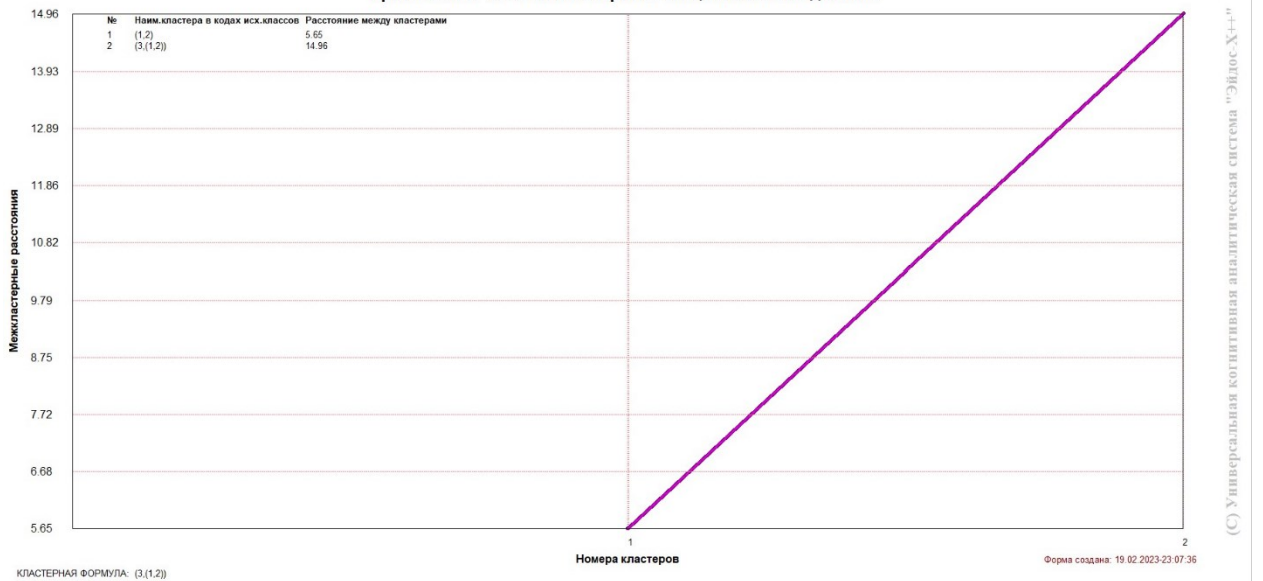


Рисунок 24 – График изменения межкластерных расстояний при когнитивной кластеризации классов

Из рисунка 23 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 25).

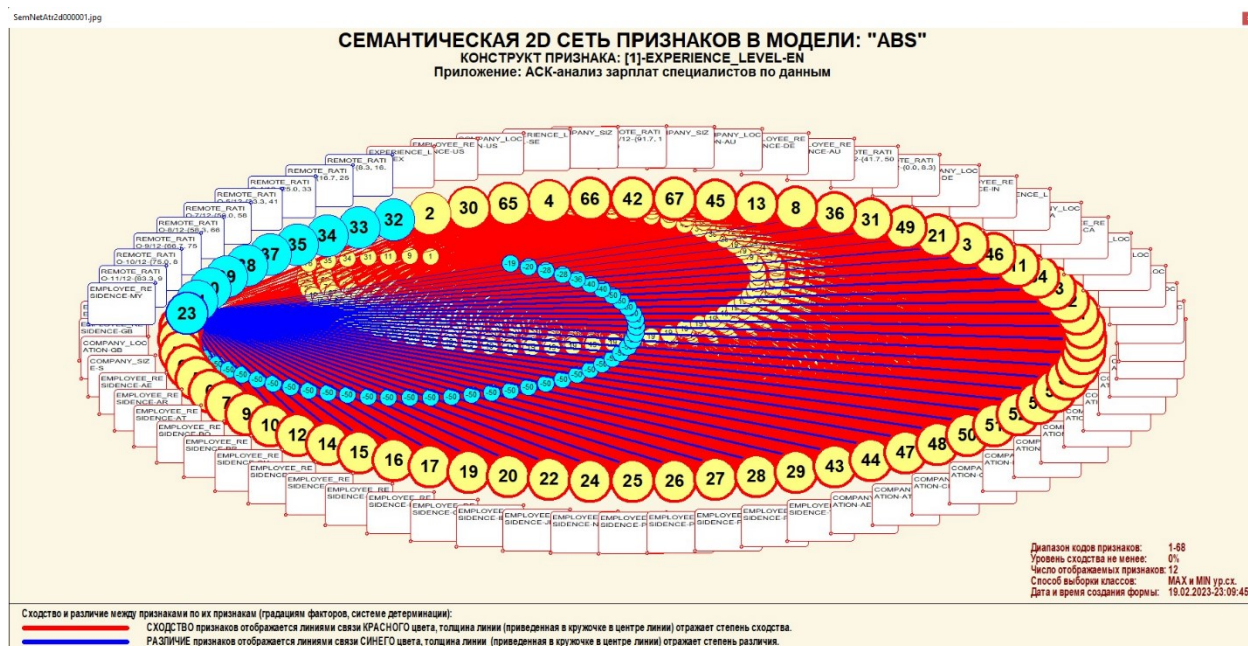


Рисунок 25 – График изменения межкластерных расстояний

### 2.3.4 Агломеративная когнитивная кластеризация значений факторов

На рисунках 26-27 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: "ABS" Приложение: "АСК-анализ зарплат специалистов по данным"

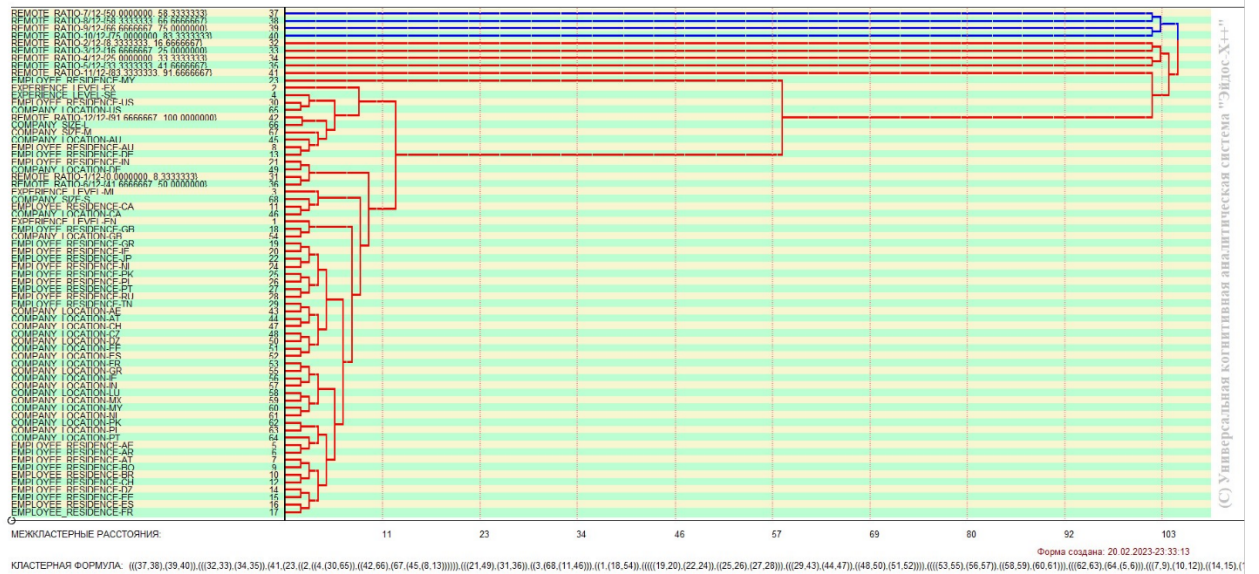


Рисунок 26 – Дендрограмма когнитивной кластеризации признаков

### ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "ABS" Приложение: "АСК-анализ зарплат специалистов по данным"

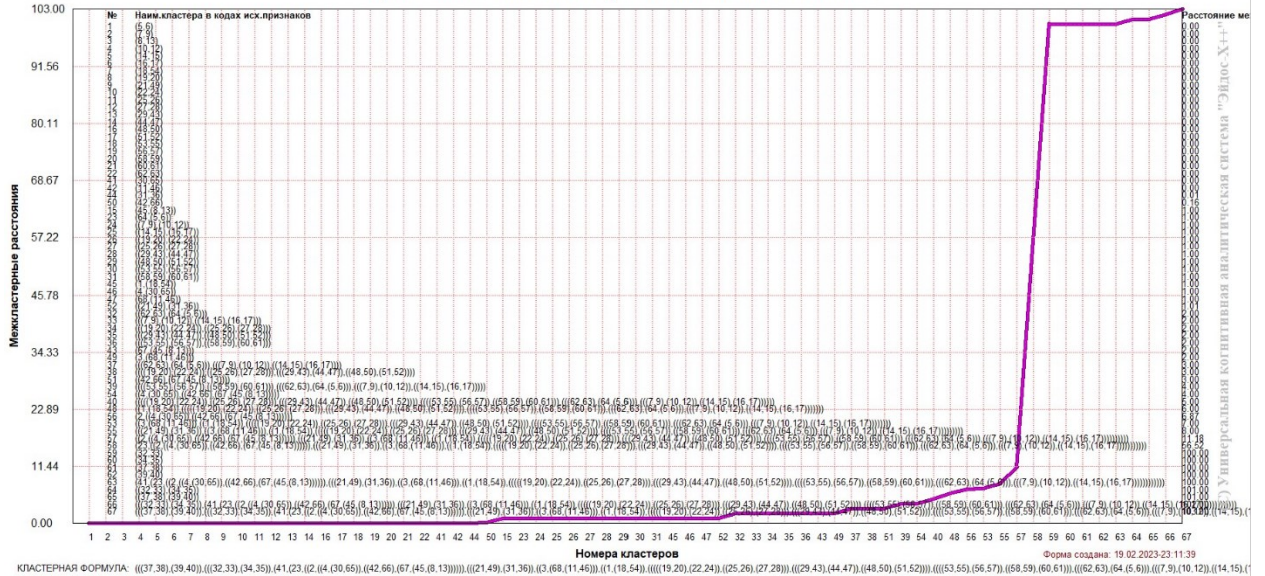


Рисунок 27 – График изменения межкластерных расстояний при когнитивной кластеризации признаков

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 28 приведён пример нелокального нейрона, а на рисунке 29 – 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.

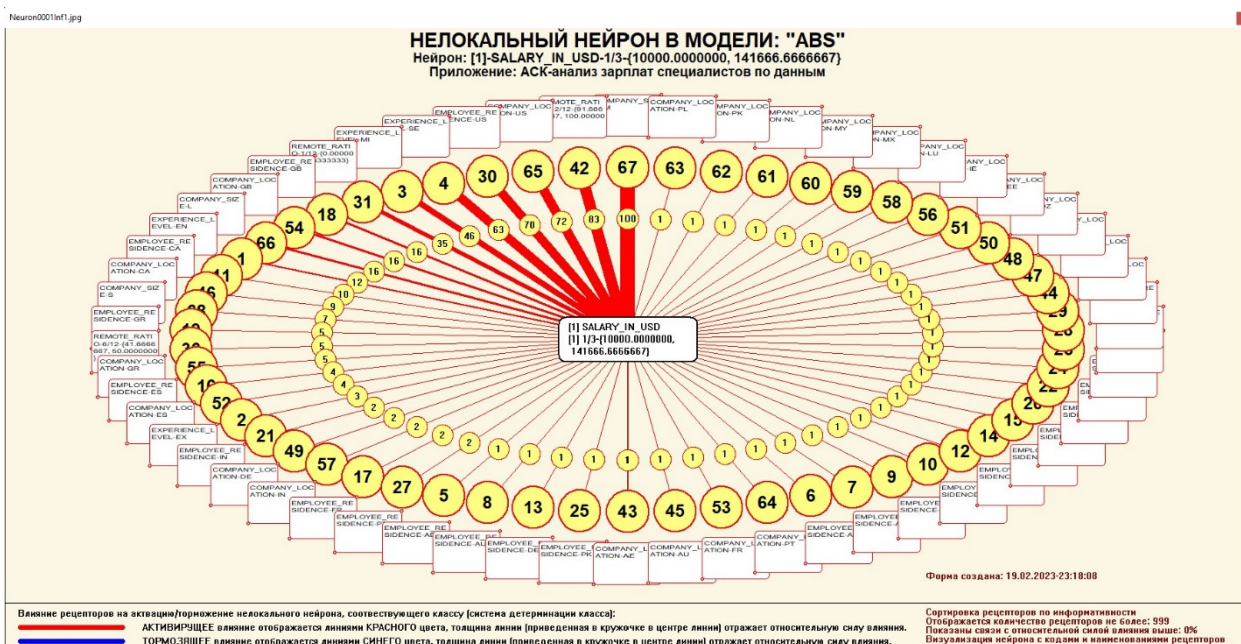


Рисунок 28 – Пример нелокального нейрона

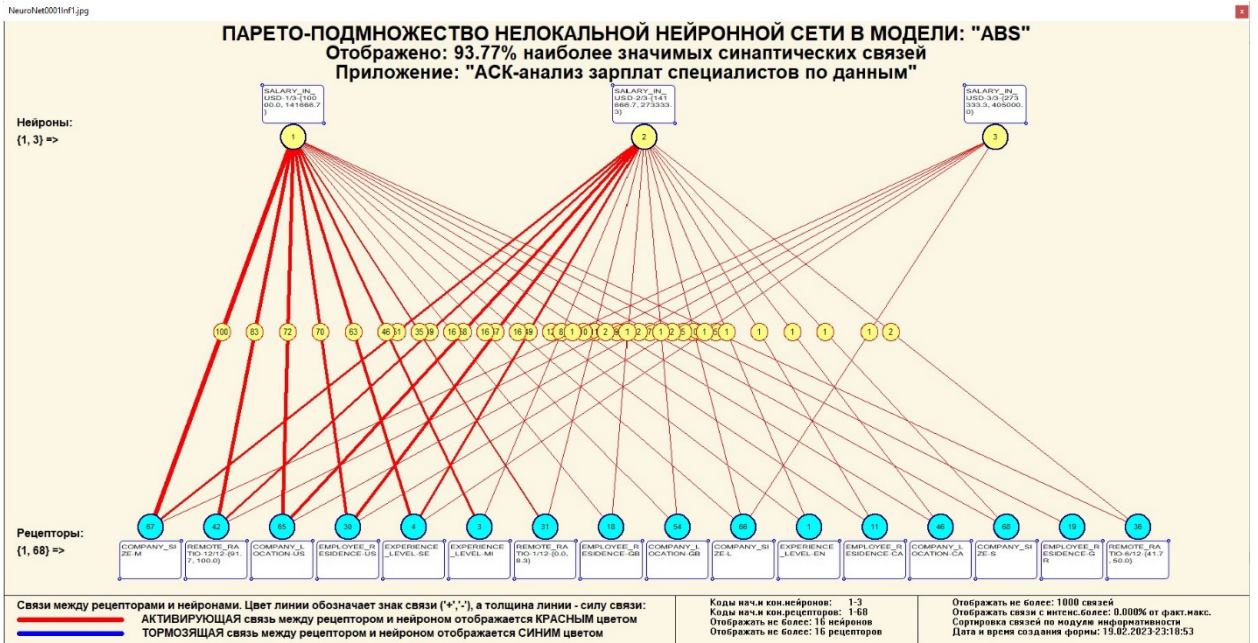


Рисунок 29 – Один слой нелокальной нейронной сети

### 2.3.6 3D-интегральные когнитивные карты

На рисунке 30 приведен фрагмент 3D-интегральной когнитивной карты, отражающий фрагмент СК-модели ABS.

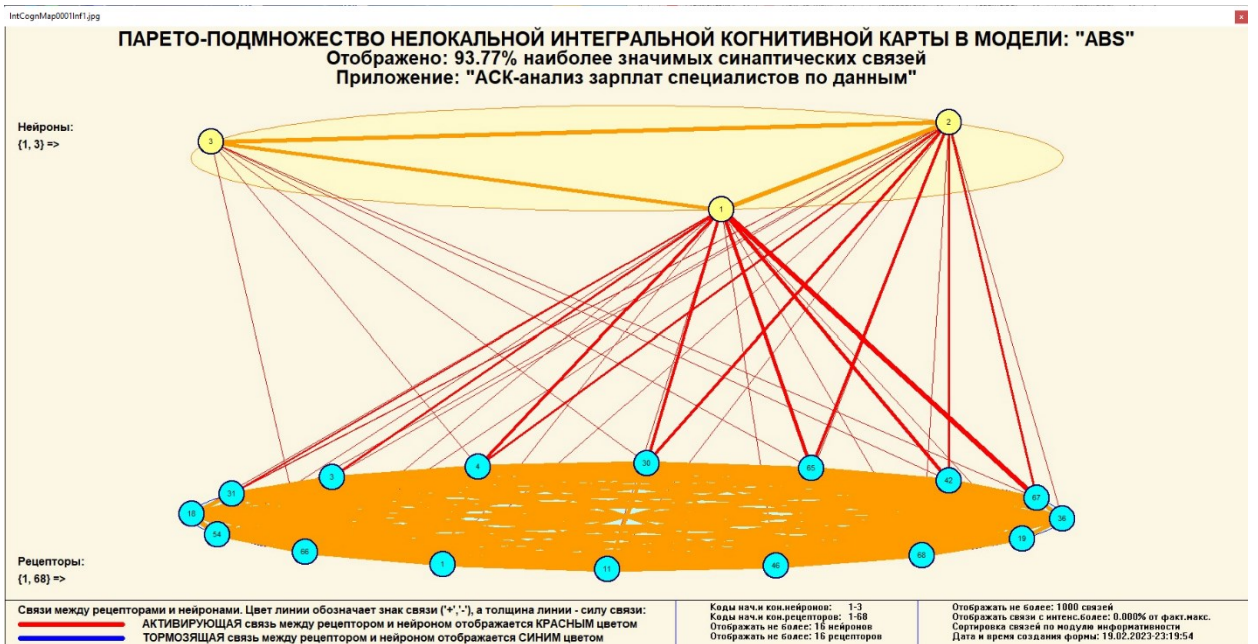


Рисунок 30 – 3D-интегральная когнитивная карта

## 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 31, 32, 33, 34, 35.

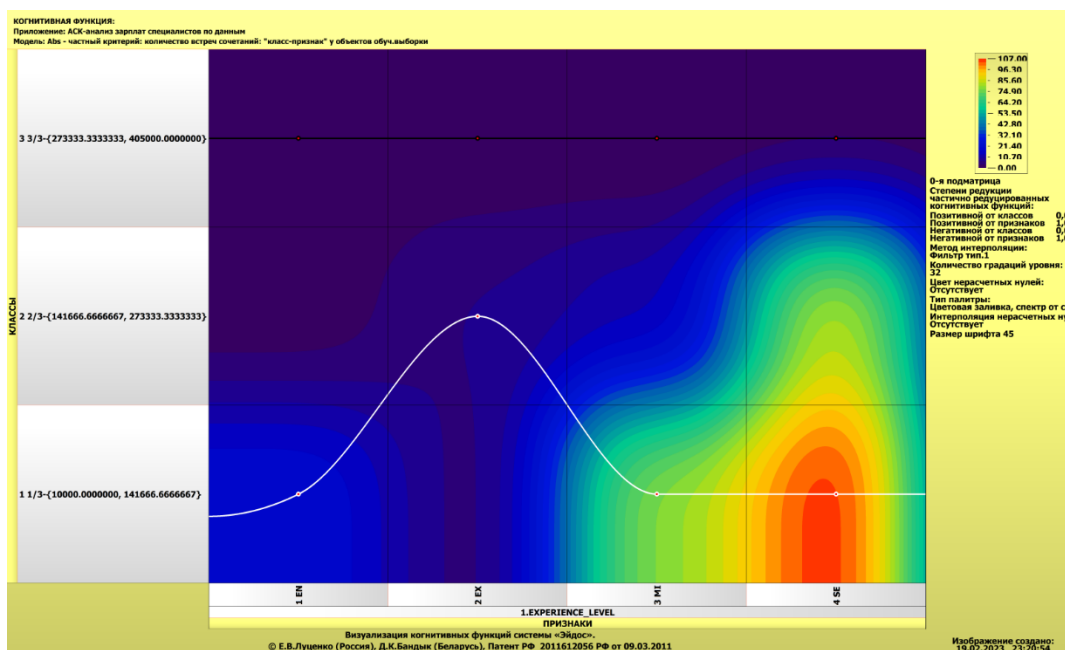


Рисунок 31 – Пример визуализации когнитивной функции по признаку 1

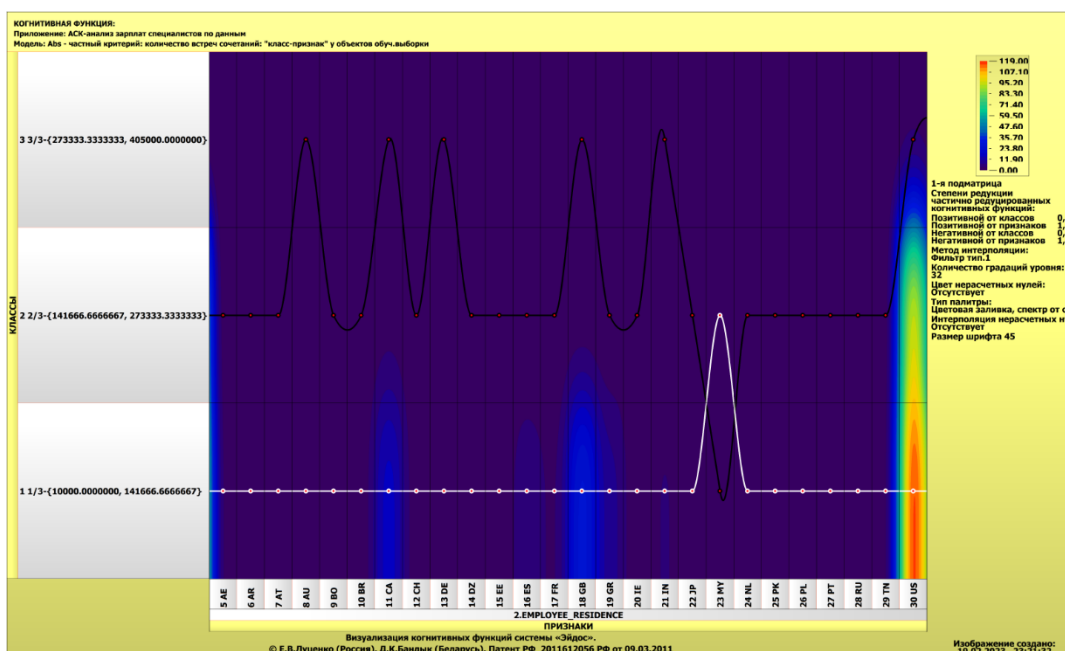


Рисунок 32 – Пример визуализации когнитивной функции по признаку 2

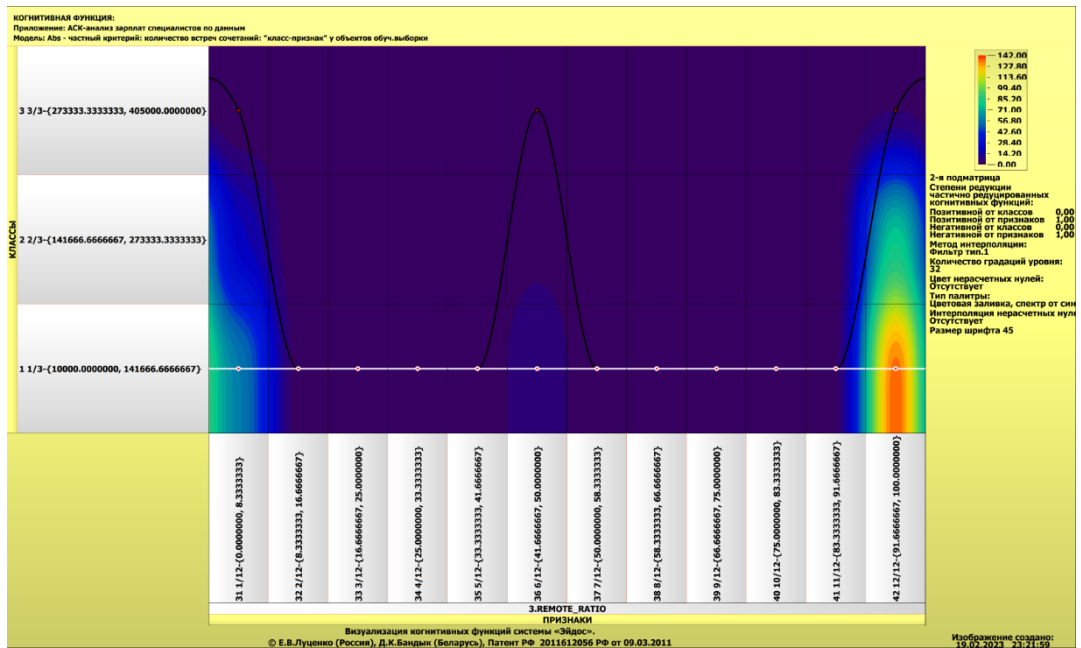


Рисунок 33 – Пример визуализации когнитивной функции по признаку 3

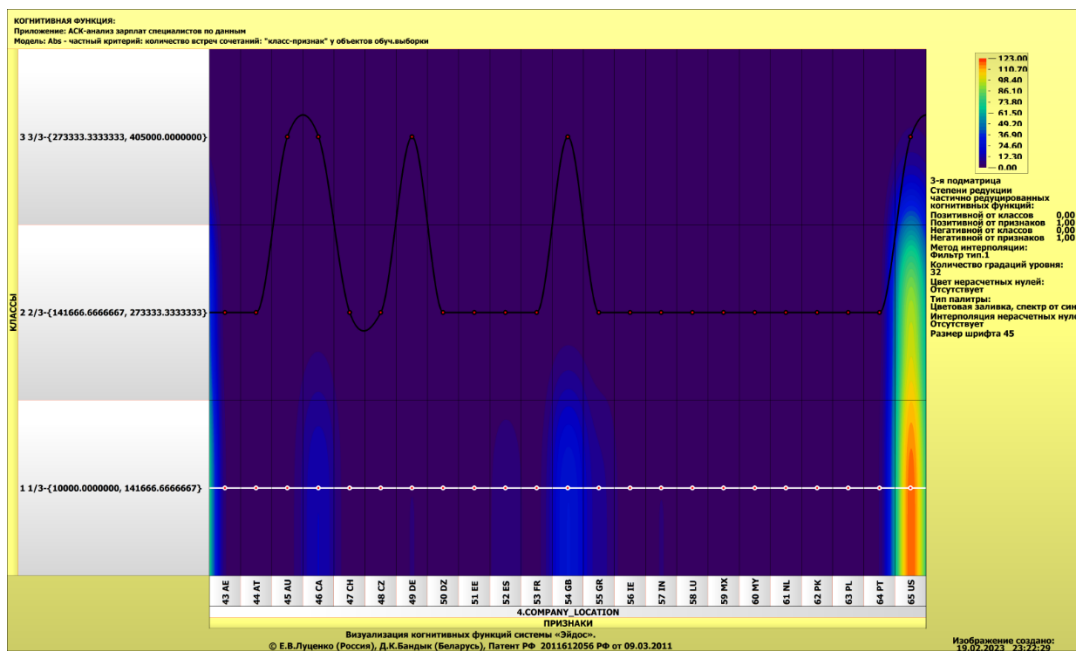


Рисунок 34 – Пример визуализации когнитивной функции по признаку 4

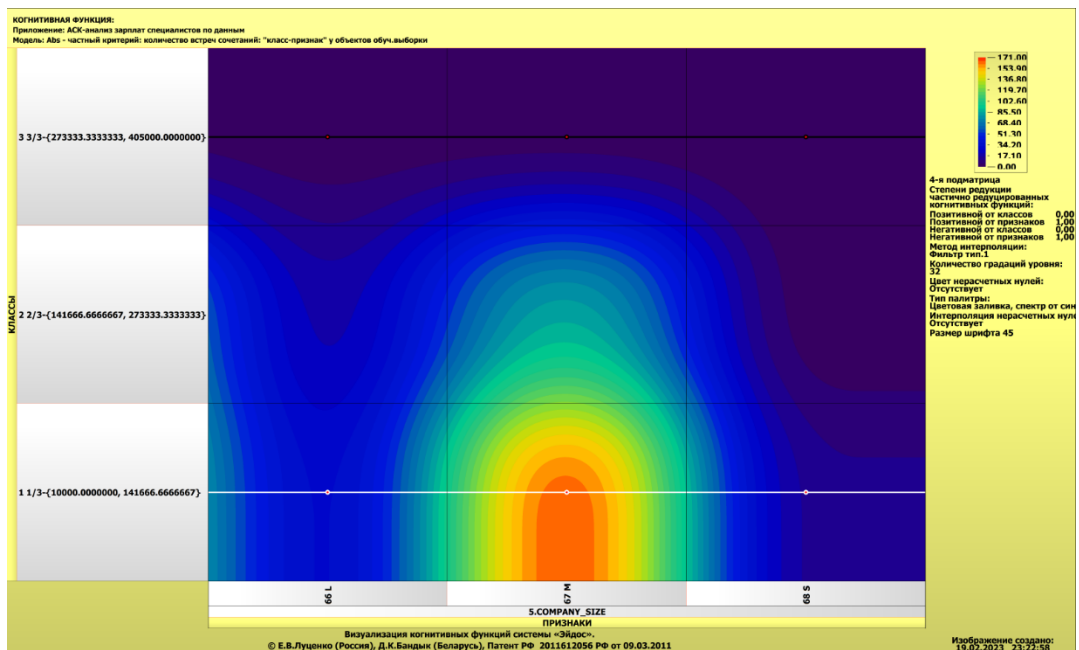


Рисунок 35 – Пример визуализации когнитивной функции по признаку 4

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.