


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 дефолта заемщика на основе сведений о клиенте»
Выполнил студент группы: ИТ2141 Рылов Михаил Леонидович

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: 35 страниц, 29 рисунка, 10 литературных источников.

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

Цель данной курсовой работы состоит в проведении АСК-анализа данных о дефолте заемщика, а также построение и отображение причинно-следственных связей между параметрами исходных данных, путем создания интеллектуальной модели.

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

РЕФЕРАТ	2
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	12
1.6 Результаты верификации моделей	14
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	18
2.1 Решение задач идентификации.....	18
2.2 SWOT и PEST матрицы и диаграммы	22
2.3 Исследование моделируемой предметной области путем исследования ее модели.....	23
2.3.1 Когнитивные диаграммы классов	23
2.3.2 Когнитивные диаграммы значений факторов.....	24
2.3.3 Агломеративная когнитивная кластеризация значений факторов.....	25
2.3.4 Нелокальные нейроны и нелокальные нейронные сети	27
2.3.5 3D-интегральные когнитивные карты	28
2.3.6 Когнитивные функции	29
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 дефолта заемщика по сведениям о клиенте. Выявление зависимостей между наблюдаемыми данными позволит облегчить процесс решения предоставления займа финансовым организациям.

Из всего вышесказанного можно выделить цель работы. Целью данной работы является разработка интеллектуальных моделей, отражающих реально существующие причинно-следственные взаимосвязи между сведениями о клиенте и его дефолте.

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 375, соответственно);
- мультязычная 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 текстами;
- 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:

- client_id – идентификатор заявителя;
- education – образование;
- sex – пол;
- age – возраст;
- car – наличие машины;
- decline_app_cnt – количество отказанных прошлых заявок;
- good_work – наличие «хорошей» работы;
- score_bki – скор бал по данным бюро кредитных историй;
- income – доход;
- default – флаг дефолта по кредиту;

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

	A	B	C	D	E	F	G	H	I	J
1	client_id	education	sex	age	car	decline_ap	good_wor	score_bki	income	default
2	25905	SCH	M	62	Y	0	Y	-2,00875	18000	Y
3	63161	SCH	F	59	N	0	Y	-1,53228	19000	Y
4	25887	SCH	M	25	Y	2	Y	-1,40814	30000	Y
5	16222	SCH	F	53	N	0	Y	-2,05747	10000	Y
6	101655	GRD	M	48	N	0	N	-1,24472	30000	Y
7	41415	SCH	M	27	Y	0	N	-2,03226	15000	Y
8	28436	SCH	M	39	N	0	Y	-2,225	28000	Y
9	68769	SCH	F	39	N	0	Y	-1,52274	45000	Y
10	38424	SCH	F	50	Y	1	Y	-1,67606	30000	Y
11	4496	UGR	F	54	N	0	Y	-2,69518	24000	Y
12	26995	UGR	M	50	Y	2	Y	-1,7001	60000	Y
13	91883	GRD	M	29	N	0	Y	-1,87002	40000	Y
14	2828	GRD	F	32	N	0	Y	-1,69675	23000	Y
15	70724	GRD	M	30	N	0	N	-1,40851	30000	Y
16	68278	SCH	F	48	N	0	Y	-2,02321	32000	Y
17	14992	SCH	M	23	N	0	Y	-0,79481	33000	Y
18	23070	UGR	F	37	N	0	N	-2,06784	33000	Y
19	61842	SCH	F	62	N	0	Y	-1,95936	8000	Y
20	26038	GRD	F	36	N	0	Y	-1,97191	60000	N
21	96972	SCH	M	41	N	0	Y	-2,05064	25000	Y
22	1806	SCH	M	28	Y	0	Y	-1,05467	50000	Y
23	75967	GRD	M	58	N	0	Y	-1,6889	60000	Y
24	98251	GRD	F	59	N	0	Y	-1,79412	14000	Y
25	51615	SCH	M	27	Y	0	Y	-2,17854	55000	N
26	35287	SCH	F	26	N	0	Y	-1,43658	26000	N
27	82132	GRD	M	45	N	0	Y	-2,67748	23000	Y
28	75528	SCH	F	44	N	0	Y	-2,29297	20000	Y
29	59924	SCH	F	43	N	0	Y	-2,40153	25000	Y
30	15681	UGR	M	37	Y	0	Y	-1,74068	100000	Y
31	92176	GRD	F	26	N	0	Y	-2,56406	15000	Y
32	60374	SCH	F	56	N	0	Y	-2,40472	10000	Y
33	98030	GRD	F	28	N	0	N	-1,53331	15000	Y
34	53675	GRD	M	62	N	0	Y	-2,82274	14800	N
35	24303	UGR	M	26	N	2	Y	-1,10978	70000	Y
36	65933	GRD	F	40	Y	0	Y	-1,58615	45000	Y

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.xls.

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).

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

Задайте тип файла исходных данных: "Inp_data":

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

Задайте параметры:

Нули и пробелы считать ОТСУТСТВИЕМ данных
 Нули и пробелы считать ЗНАЧЕНИЯМИ данных
 Создавать БД средних по классам "Inp_davr.dbf"?
Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:
 Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:
 Конечный столбец описательных шкал:

Задайте режим:

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

Задайте способ выбора размера интервалов:

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

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

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

Параметры интерпретации значений текстовых полей "Inp_data":

В качестве классов рассматриваются:

Значения полей целиком
 Элементы значений полей - слова > символов:
 Элементы значений полей - символы

Выделять уникальные значения и сортировать
 Не выделять уникальных значений и не сортировать

В качестве признаков рассматриваются:

Значения полей целиком
 Элементы значений полей - слова > символов:
 Элементы значений полей - символы

Проводить лемматизацию
 Не проводить лемматизацию

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}")

Рисунок 2 – Интерфейс импорта данных

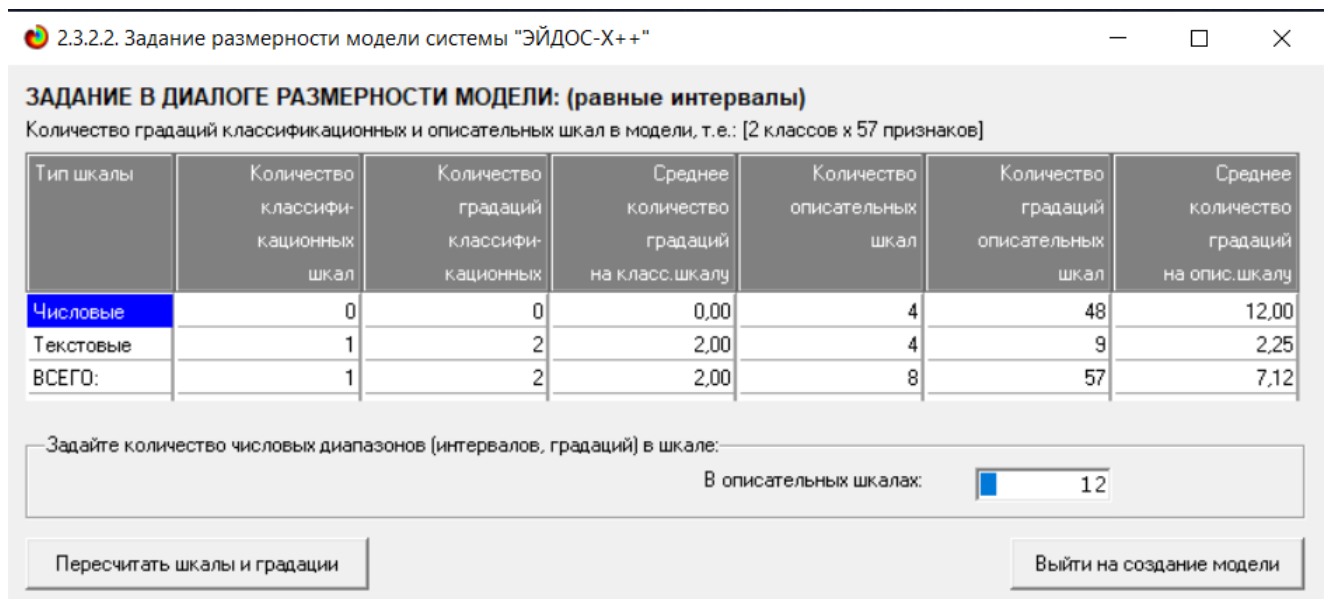


Рисунок 3 – Задание размерности модели

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

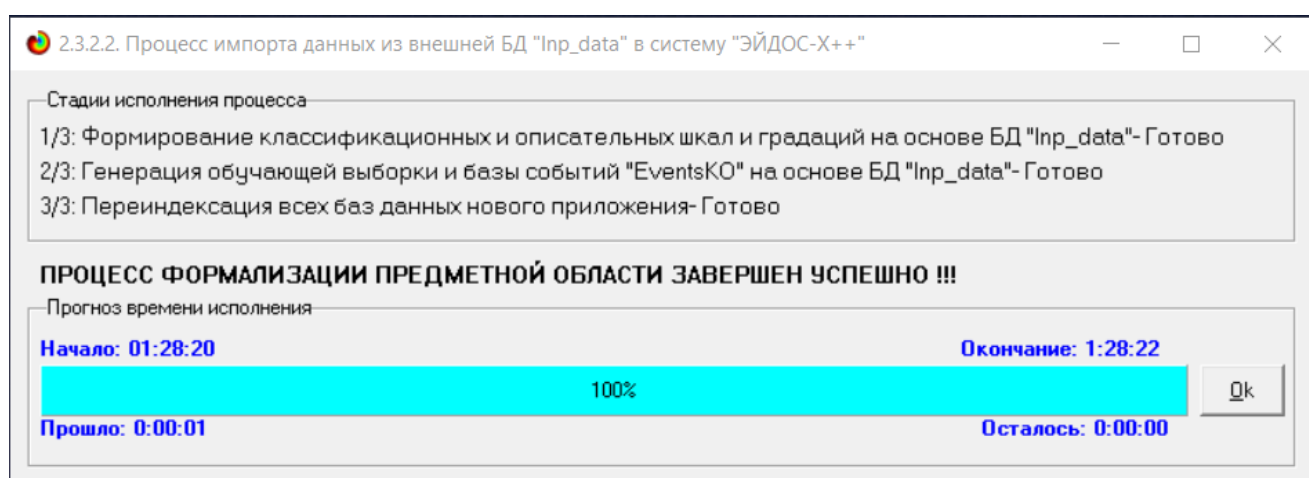


Рисунок 4 – Процесс импорта данных

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).

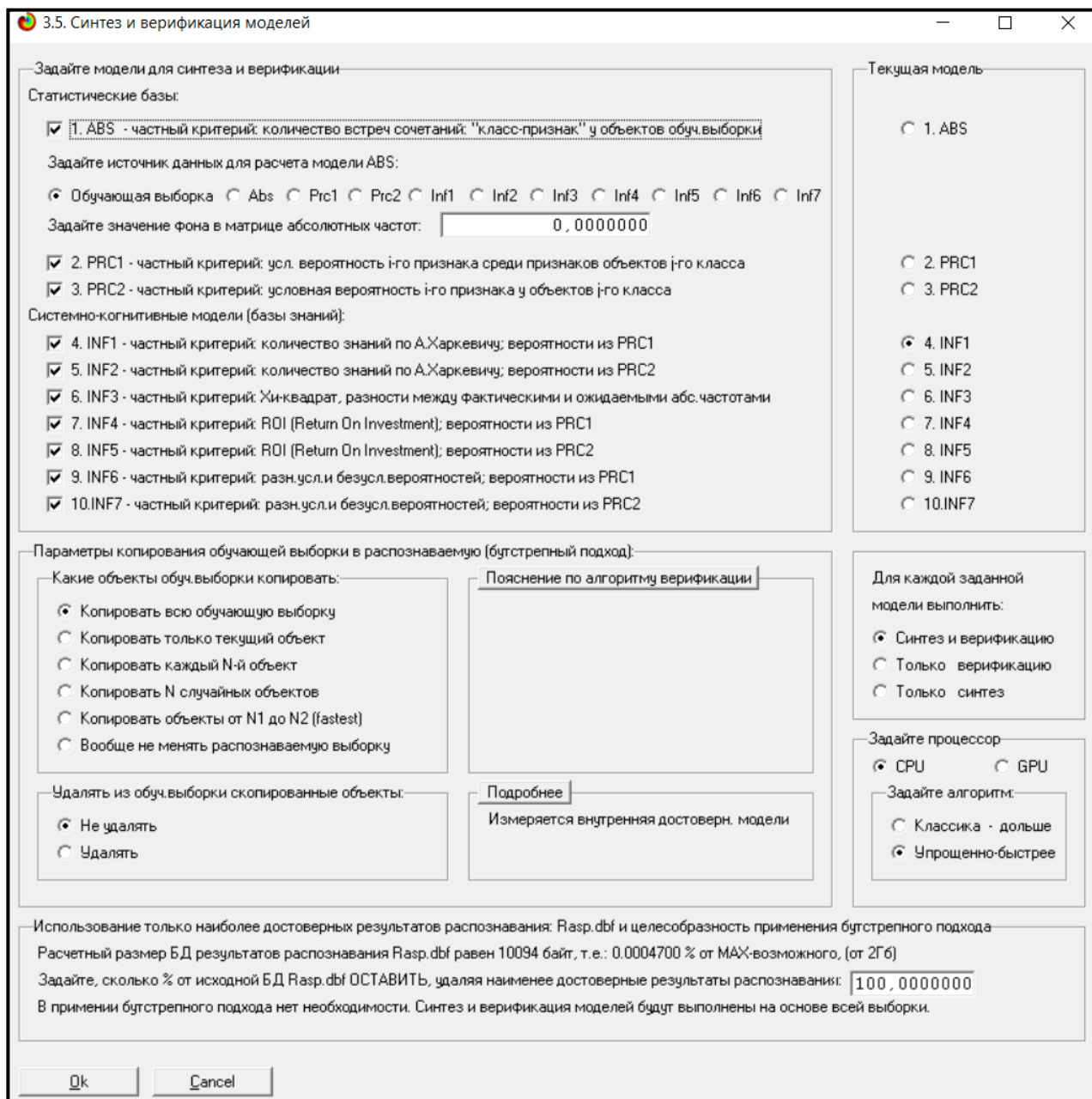


Рисунок 5 – Выбор моделей для синтеза и верификации

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.

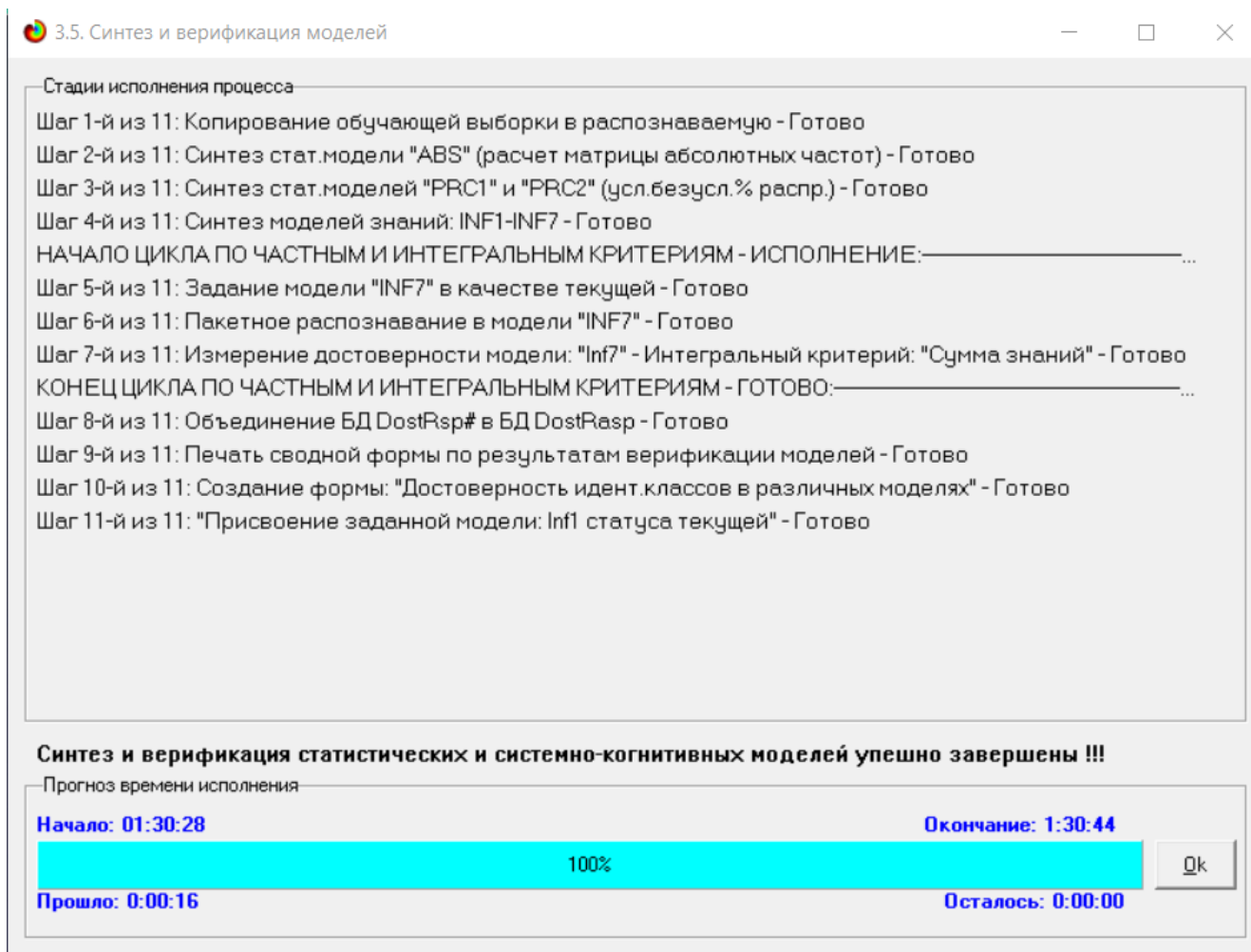


Рисунок 6 – Процесс исполнения режима 3.5

Стоит отметить, что синтез и верификация всех 10 моделей на данной задаче заняли 16 секунд. 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.

Код признака	Наименование описательной шкалы и градации	1. DEFAULT N	2. DEFAULT Y	Сумма	Среднее	Средн. квадр. откл.
1.0	EDUCATION-GRD	3.0	45.0	48.0	24.00	29.70
2.0	EDUCATION-SCH	12.0	63.0	75.0	37.50	36.06
3.0	EDUCATION-UGR	1.0	15.0	16.0	8.00	9.90
4.0	SEX-F	9.0	65.0	74.0	37.00	39.60
5.0	SEX-M	7.0	58.0	65.0	32.50	36.06
6.0	AGE-1/12-{21.0000000, 24.8333333}		10.0	10.0	5.00	7.07
7.0	AGE-2/12-{24.8333333, 28.6666667}	3.0	19.0	22.0	11.00	11.31
8.0	AGE-3/12-{28.6666667, 32.5000000}	3.0	19.0	22.0	11.00	11.31
9.0	AGE-4/12-{32.5000000, 36.3333333}	2.0	10.0	12.0	6.00	5.66
10.0	AGE-5/12-{36.3333333, 40.1666667}	6.0	17.0	23.0	11.50	7.78
11.0	AGE-6/12-{40.1666667, 44.0000000}		10.0	10.0	5.00	7.07
12.0	AGE-7/12-{44.0000000, 47.8333333}		5.0	5.0	2.50	3.54
13.0	AGE-8/12-{47.8333333, 51.6666667}		10.0	10.0	5.00	7.07
14.0	AGE-9/12-{51.6666667, 55.5000000}		5.0	5.0	2.50	3.54
15.0	AGE-10/12-{55.5000000, 59.3333333}	1.0	9.0	10.0	5.00	5.66
16.0	AGE-11/12-{59.3333333, 63.1666667}	1.0	7.0	8.0	4.00	4.24
17.0	AGE-12/12-{63.1666667, 67.0000000}		2.0	2.0	1.00	1.41
18.0	CAR-N	12.0	84.0	96.0	48.00	50.91
19.0	CAR-Y	4.0	39.0	43.0	21.50	24.75
20.0	DECLINE_APP_CNT-1/12-{1.0000000, 1.2500000}	2.0	9.0	11.0	5.50	4.95
21.0	DECLINE_APP_CNT-2/12-{1.2500000, 1.5000000}					
22.0	DECLINE_APP_CNT-3/12-{1.5000000, 1.7500000}					
23.0	DECLINE_APP_CNT-4/12-{1.7500000, 2.0000000}		5.0	5.0	2.50	3.54
24.0	DECLINE_APP_CNT-5/12-{2.0000000, 2.2500000}					
25.0	DECLINE_APP_CNT-6/12-{2.2500000, 2.5000000}					
26.0	DECLINE_APP_CNT-7/12-{2.5000000, 2.7500000}					

Рисунок 7 – Фрагмент матрицы абсолютных частот

5.5. Модель: "5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2"

Код признака	Наименование описательной шкалы и градации	1. DEFAULT N	2. DEFAULT Y	Сумма	Среднее	Средн. квадр. откл.
1.0	EDUCATION-GRD	-0.124	0.012	-0.112	-0.056	0.096
2.0	EDUCATION-SCH	0.067	-0.011	0.056	0.028	0.055
3.0	EDUCATION-UGR	-0.124	0.012	-0.112	-0.056	0.096
4.0	SEX-F	0.011	-0.001	0.010	0.005	0.009
5.0	SEX-M	-0.013	0.002	-0.012	-0.006	0.011
6.0	AGE-1/12-{21.0000000, 24.8333333}		0.025	0.025	0.012	0.018
7.0	AGE-2/12-{24.8333333, 28.6666667}	0.034	-0.005	0.029	0.015	0.028
8.0	AGE-3/12-{28.6666667, 32.5000000}	0.034	-0.005	0.029	0.015	0.028
9.0	AGE-4/12-{32.5000000, 36.3333333}	0.075	-0.012	0.063	0.031	0.062
10.0	AGE-5/12-{36.3333333, 40.1666667}	0.166	-0.036	0.129	0.065	0.143
11.0	AGE-6/12-{40.1666667, 44.0000000}		0.025	0.025	0.012	0.018
12.0	AGE-7/12-{44.0000000, 47.8333333}		0.025	0.025	0.012	0.018
13.0	AGE-8/12-{47.8333333, 51.6666667}		0.025	0.025	0.012	0.018
14.0	AGE-9/12-{51.6666667, 55.5000000}		0.025	0.025	0.012	0.018
15.0	AGE-10/12-{55.5000000, 59.3333333}	-0.029	0.003	-0.025	-0.013	0.023
16.0	AGE-11/12-{59.3333333, 63.1666667}	0.017	-0.002	0.014	0.007	0.013
17.0	AGE-12/12-{63.1666667, 67.0000000}		0.025	0.025	0.012	0.018
18.0	CAR-N	0.017	-0.002	0.014	0.007	0.013
19.0	CAR-Y	-0.043	0.005	-0.038	-0.019	0.034
20.0	DECLINE_APP_CNT-1/12-{1.0000000, 1.2500000}	0.093	-0.016	0.077	0.038	0.077
21.0	DECLINE_APP_CNT-2/12-{1.2500000, 1.5000000}					
22.0	DECLINE_APP_CNT-3/12-{1.5000000, 1.7500000}					
23.0	DECLINE_APP_CNT-4/12-{1.7500000, 2.0000000}		0.025	0.025	0.012	0.018
24.0	DECLINE_APP_CNT-5/12-{2.0000000, 2.2500000}					
25.0	DECLINE_APP_CNT-6/12-{2.2500000, 2.5000000}					
26.0	DECLINE_APP_CNT-7/12-{2.5000000, 2.7500000}					

Рисунок 8 – Фрагмент модели INF2

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 достов. моделям при разн. инт. крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Ю-лык (J)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модуль... истинно-полож. решений (STP)	Сумма модуль... истинно-отриц. решений (STN)	Сумма модуль... ложно-полож. решений (SFP)	Сумма модуль... ложно-отриц. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко	Сред. ур. истин. реш.
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	0.500	1.000	0.667	96.501	89.715	0.518	1.000	0.683				
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс частот по призна...	0.500	1.000	0.667	101.296	27.761	0.785	1.000	0.879				
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Корреляция усл отн частот с о...	0.500	1.000	0.667	96.501	89.715	0.518	1.000	0.683				
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Сумма усл отн частот по приз...	0.500	1.000	0.667	95.625	99.713	0.490	1.000	0.657				
3. PRC2 - частный критерий: условная вероятность его признака.	Корреляция усл отн частот с о...	0.500	1.000	0.667	96.500	89.715	0.518	1.000	0.683				
3. PRC2 - частный критерий: условная вероятность его признака.	Сумма усл отн частот по при...	0.500	1.000	0.667	95.777	99.734	0.490	1.000	0.659				
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.436	0.295	0.352	6.757	21.461	9.755	16.522	0.409	0.290	0.340		
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.526	0.576	0.550	8.439	9.747	16.504	2.247	0.338	0.790	0.474		
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.441	0.295	0.353	6.743	21.573	9.681	16.500	0.411	0.290	0.340		
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.529	0.583	0.555	8.449	9.923	16.339	2.237	0.341	0.791	0.476		
6. INF3 - частный критерий: Хи-квадрат: разности между факти...	Семантический резонанс зна...	0.338	0.338	0.338	14.611	14.611	30.208	30.208	0.326	0.326	0.326		
6. INF3 - частный критерий: Хи-квадрат: разности между факти...	Сумма знаний	0.338	0.338	0.338	14.696	14.696	30.552	30.552	0.325	0.325	0.325		
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.443	0.252	0.321	5.314	19.225	5.714	21.046	0.482	0.202	0.284		
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.492	0.640	0.556	6.575	2.769	14.834	1.062	0.307	0.861	0.453		
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.443	0.252	0.321	5.311	19.313	5.677	21.028	0.483	0.202	0.285		
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.503	0.655	0.569	6.580	2.825	14.703	1.055	0.309	0.862	0.455		
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	0.343	0.338	0.341	13.730	13.837	29.686	29.119	0.316	0.320	0.318		
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	0.311	0.338	0.324	9.794	3.473	36.847	3.980	0.210	0.711	0.324		
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	0.343	0.338	0.341	13.814	13.955	29.487	28.950	0.319	0.323	0.321		
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	0.311	0.338	0.324	9.788	3.568	36.555	3.955	0.211	0.712	0.326		

Рисунок 9 – Оценка достоверности моделей (F-критерий)

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

Наименование модели и частного критерия	Интегральный критерий	Полнота модели	L1-мера проф. Е.В. Луценко	Средний модуль... истинно-полож. решений	Средний модуль... истинно-отриц. решений	Средний модуль... ложно-полож. решений	Средний модуль... ложно-отриц. решений	A-Точность модели A-Precision = ATR/A TP+	A-Полнота модели A-Recall = ATR/A TP+	L2-мера проф. Е.В. Луценко	Процент правильной идентификац...	Процент прав. на иде...
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	1.000	0.683	0.694	0.645	0.645	0.518	1.000	0.683	100.000	100.000	
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс частот по призна...	1.000	0.879	0.729	0.200	0.200	0.785	1.000	0.879	100.000	100.000	
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Корреляция усл отн частот с о...	1.000	0.683	0.694	0.645	0.645	0.518	1.000	0.683	100.000	100.000	
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Сумма усл отн частот по приз...	1.000	0.657	0.688	0.717	0.717	0.490	1.000	0.657	100.000	100.000	
3. PRC2 - частный критерий: условная вероятность его признака.	Корреляция усл отн частот с о...	1.000	0.683	0.694	0.645	0.645	0.518	1.000	0.683	100.000	100.000	
3. PRC2 - частный критерий: условная вероятность его признака.	Сумма усл отн частот по приз...	1.000	0.658	0.689	0.718	0.718	0.490	1.000	0.658	100.000	100.000	
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.290	0.340	0.165	0.250	0.184	0.169	0.472	0.494	0.483	29.496	9
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.790	0.474	0.105	0.145	0.229	0.038	0.315	0.735	0.441	57.554	9
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.290	0.340	0.164	0.248	0.186	0.168	0.469	0.494	0.481	29.496	9
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.791	0.476	0.104	0.148	0.227	0.039	0.315	0.730	0.440	58.273	9
6. INF3 - частный критерий: Хи-квадрат: разности между факти...	Семантический резонанс зна...	0.326	0.326	0.311	0.311	0.328	0.328	0.486	0.486	0.486	33.813	8
6. INF3 - частный критерий: Хи-квадрат: разности между факти...	Сумма знаний	0.325	0.325	0.313	0.332	0.332	0.332	0.485	0.485	0.485	33.813	8
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.202	0.284	0.152	0.202	0.130	0.202	0.539	0.429	0.478	25.180	9
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.861	0.453	0.074	0.059	0.161	0.021	0.314	0.777	0.447	64.029	8
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.202	0.285	0.152	0.203	0.129	0.202	0.540	0.429	0.478	25.180	9
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.862	0.455	0.072	0.058	0.163	0.022	0.307	0.767	0.438	65.468	8
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	0.320	0.318	0.292	0.282	0.330	0.317	0.470	0.480	0.475	33.813	8
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	0.711	0.324	0.208	0.099	0.354	0.043	0.370	0.828	0.512	33.813	8
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	0.323	0.321	0.294	0.285	0.328	0.315	0.473	0.483	0.478	33.813	8
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	0.712	0.326	0.208	0.102	0.351	0.043	0.372	0.829	0.514	33.813	8

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=0,667 максимуме 1,000), по критерию L1 профессора Е.В. Луценко наиболее достоверной является СК-модель ABS с интегральным критерием «Сумма абсолютных частот» (L1=0,879 при максимуме 1,000), а по критерию L2 профессора Е.В. Луценко

наиболее достоверной является СК- модель ABS с интегральным критерием «Сумма абсолютных частот» ($L2=0,879$ 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.

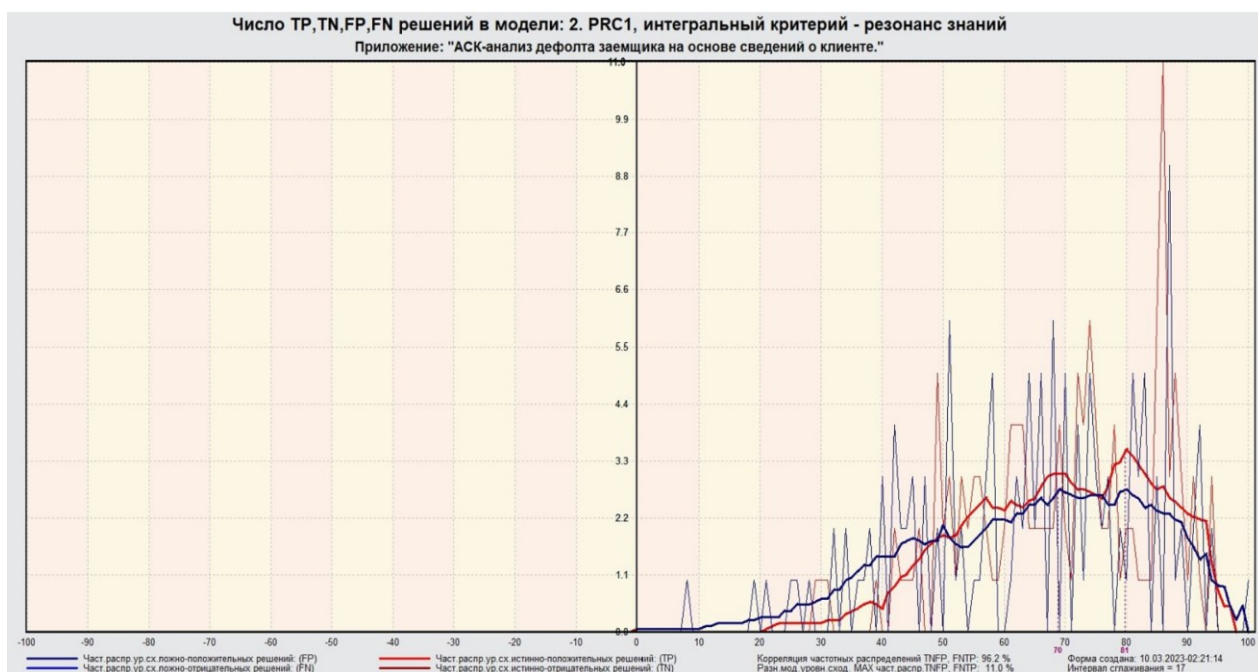


Рисунок 12 – Частотные распределения модели PRC1 (резонанс знаний)

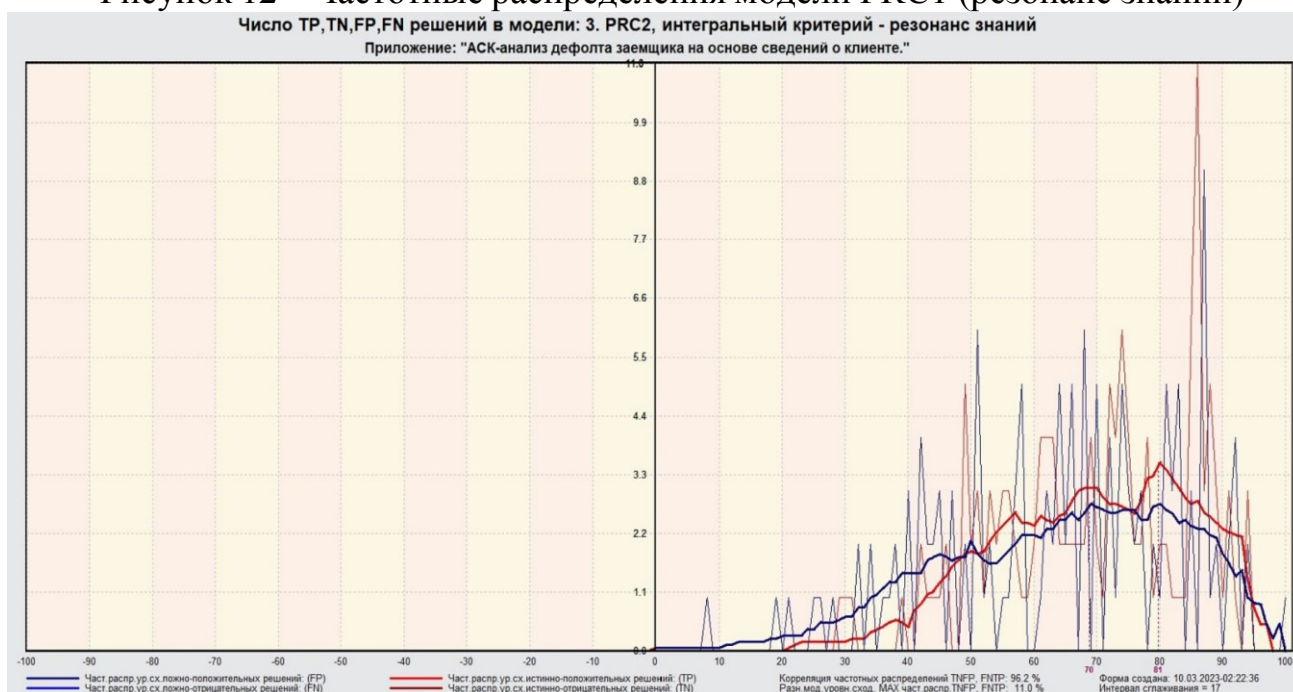


Рисунок 13 – Частотные распределения модели PRC2 (резонанс знаний)

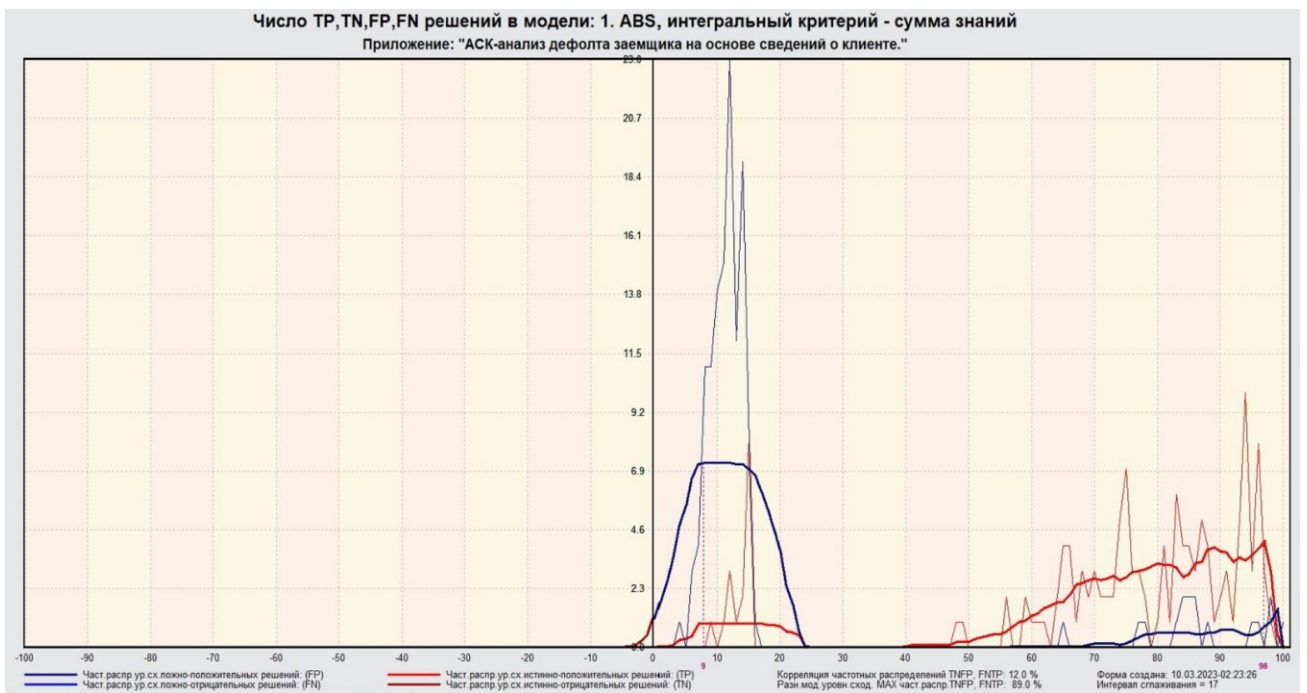


Рисунок 14 – Частотные распределения модели ABS (сумма знаний)

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1 Решение задач идентификации

Решим задачу системной идентификации, т.е. определим к каким респондентам к каким категориям относятся на основе обучающей выборки в наиболее достоверной СК-модели ABS (режим 5.6) (рисунок 15, 16).

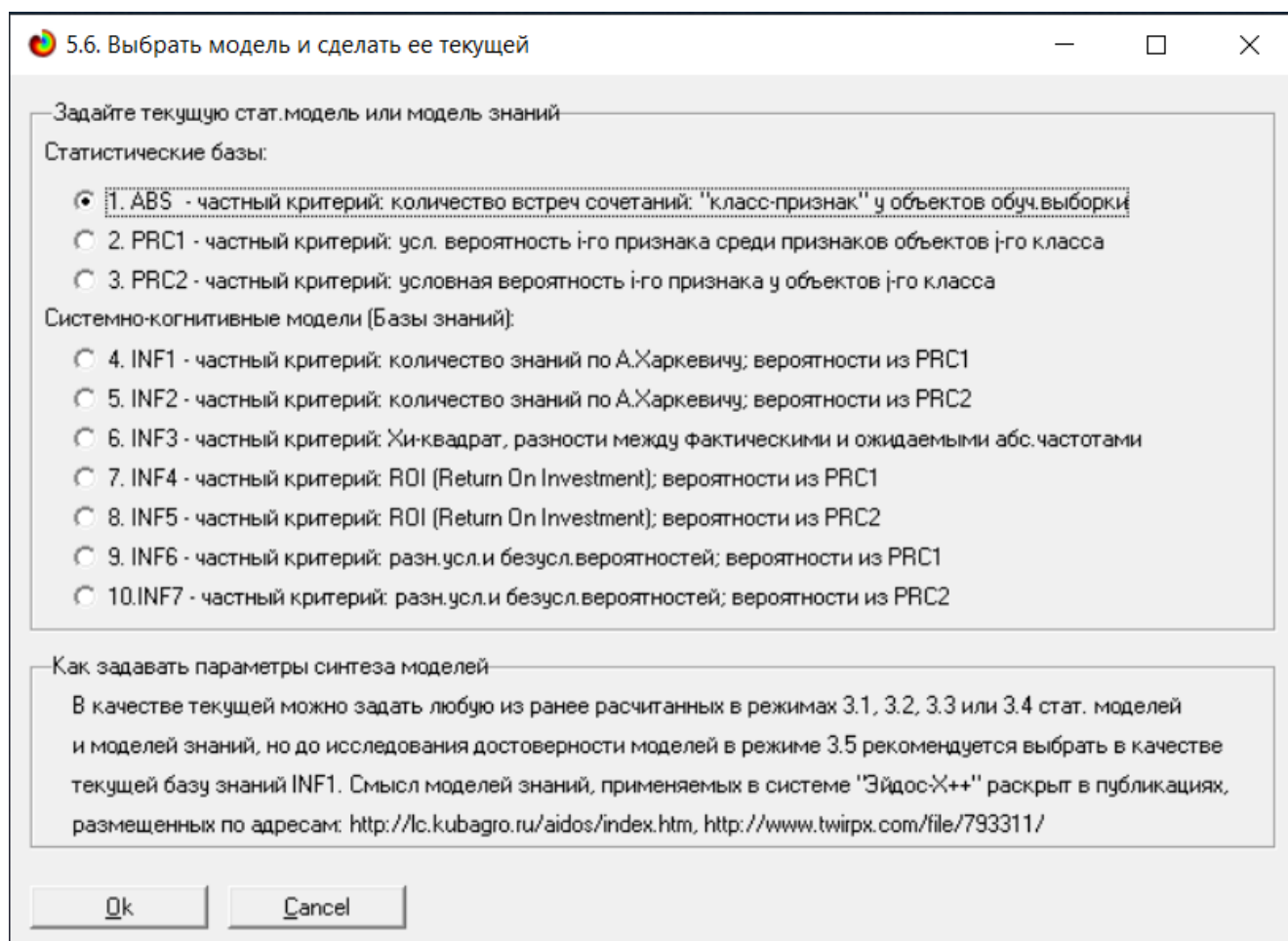


Рисунок 15 – Интерфейс задания модели

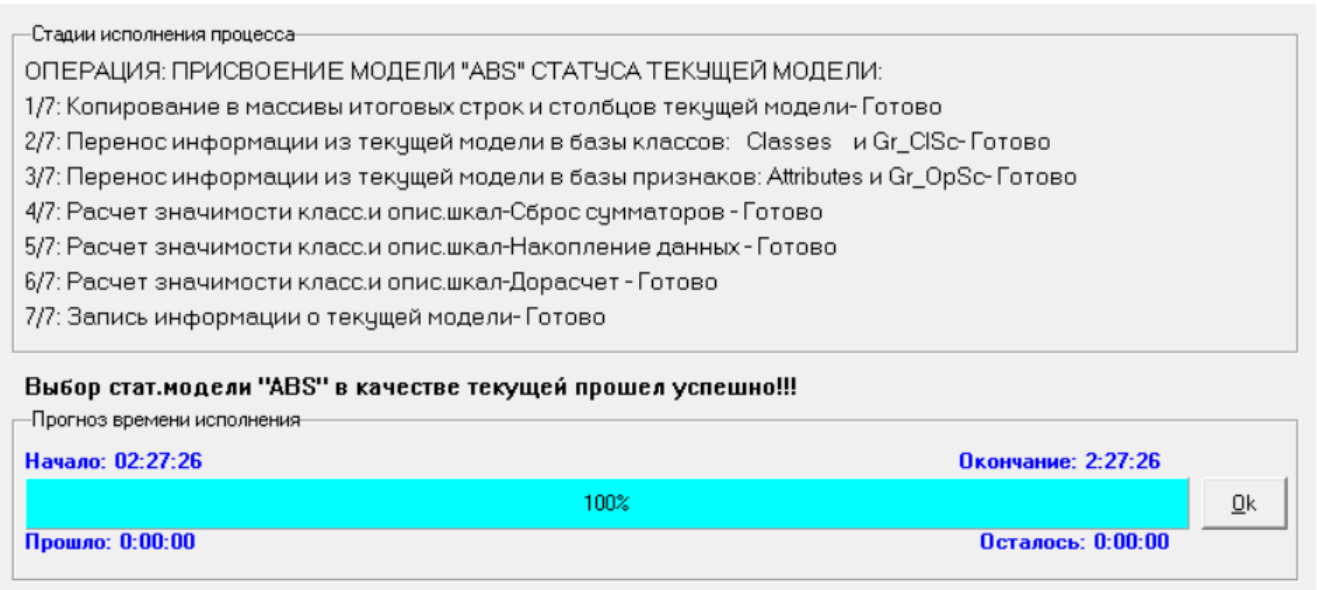


Рисунок 16 – Процесс загрузки модели

Затем следует провести пакетное распознавание в режиме 4.1.2.

(рисунок 17).

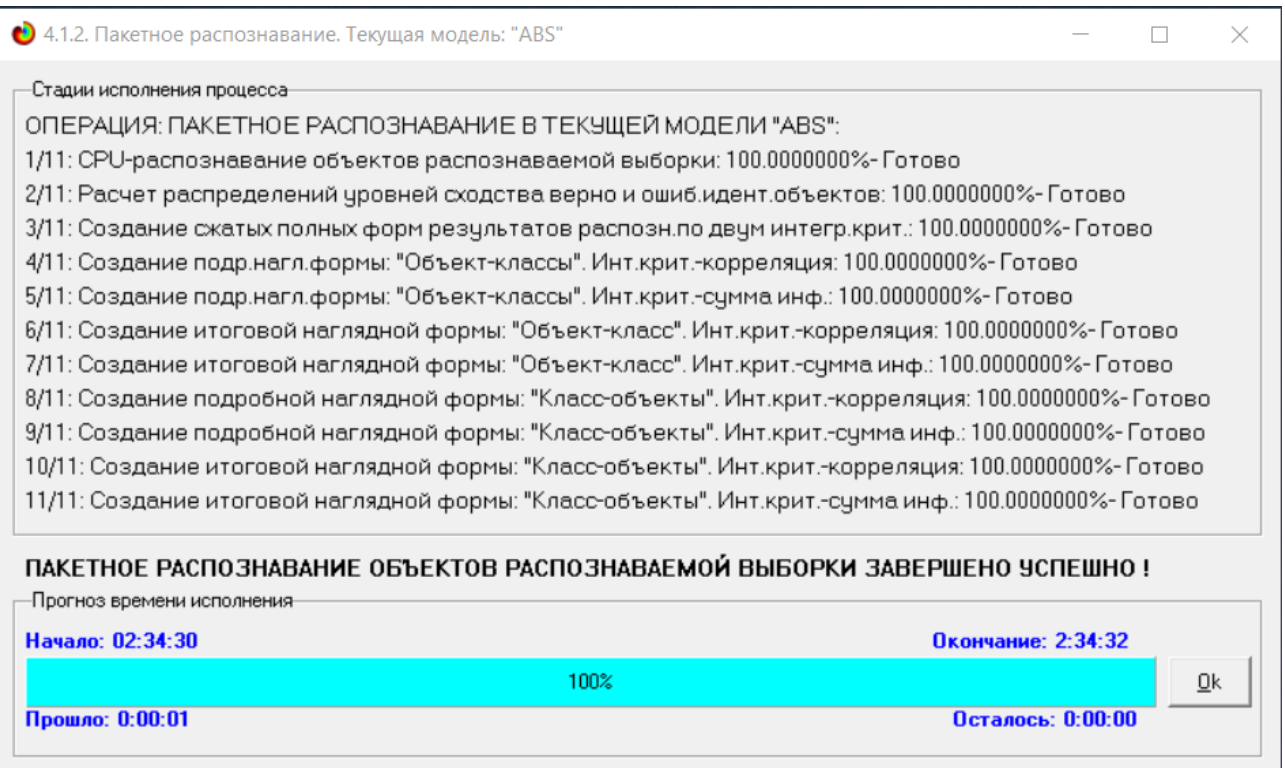


Рисунок 17 – Процесс пакетного распознавания модели

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».

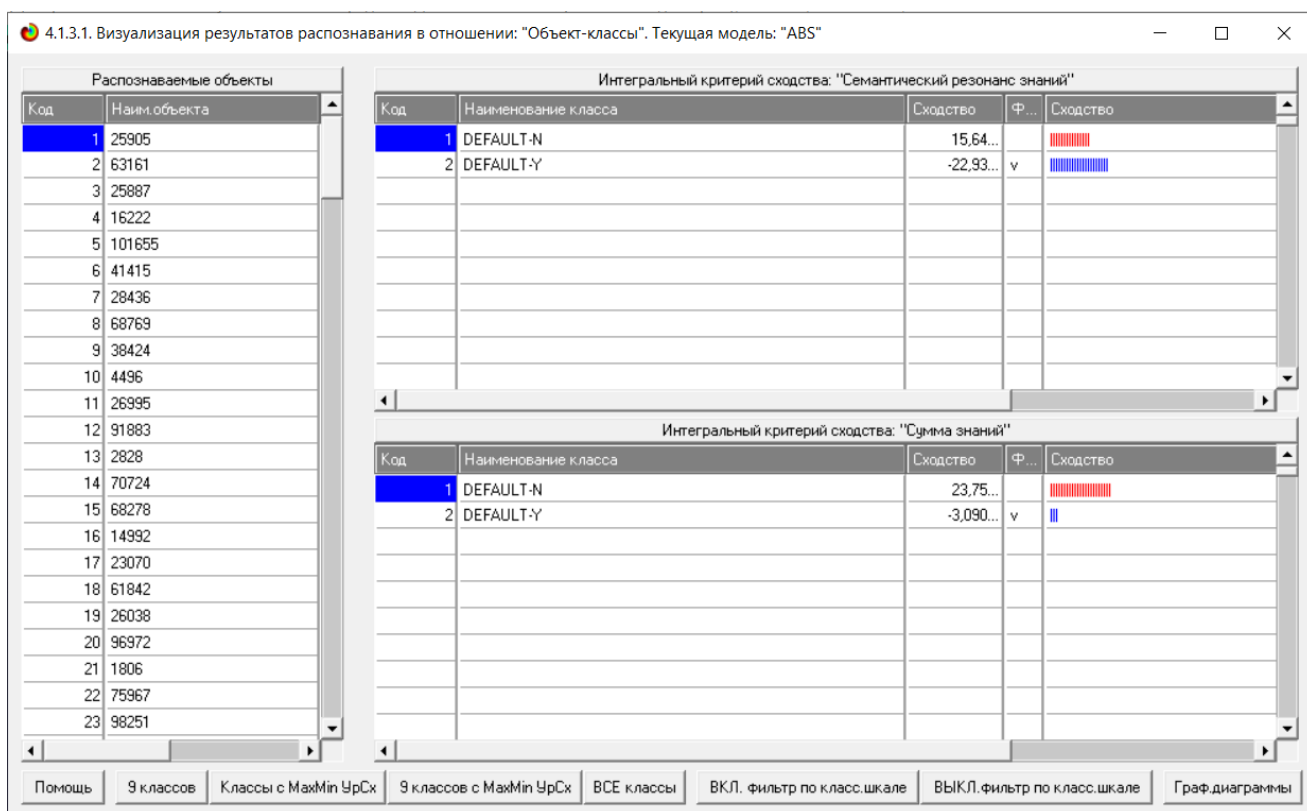


Рисунок 18 – Визуализация «Объект-классы»

Классы	
Код	Наим. класса
1	DEFAULT-N
2	DEFAULT-Y

Интегральный критерий сходства: "Семантический резонанс знаний"				
Код	Наименование объекта	Сходство	Ф...	Сходство
8	68769	100,00...		
7	28436	94,33...		
59	27149	94,33...	v	
113	23436	94,33...	v	
121	20426	94,33...		
131	29936	94,33...	v	
25	35287	92,45...	v	
54	77041	92,45...		
60	50586	92,45...		
94	47779	92,45...		

Интегральный критерий сходства: "Сумма знаний"				
Код	Наименование объекта	Сходство	Ф...	Сходство
8	68769	15,72...		
125	48580	15,25...	v	
7	28436	15,02...		
59	27149	15,02...	v	
113	23436	15,02...	v	
121	20426	15,02...		
131	29936	15,02...	v	
25	35287	14,78...	v	
48	84681	14,78...		
54	77041	14,78...		

Помощь Поиск объекта В начало БД В конец БД Предыдущая Следующая 9 записей Все записи Печать XLS Печать TXT Печать ALL

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 дефолт заемщика (рисунок 20).

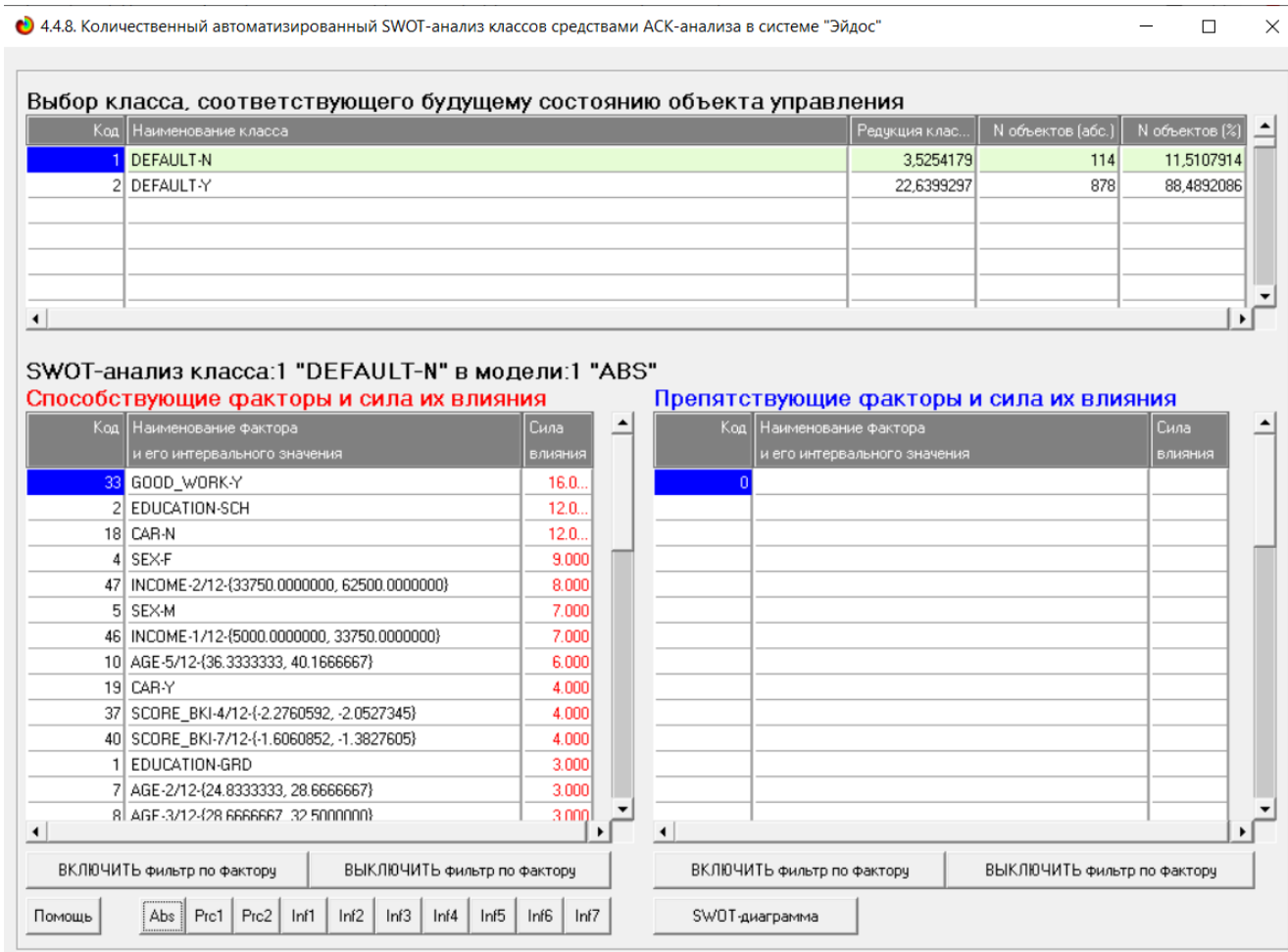


Рисунок 20 – SWOT-матрица

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

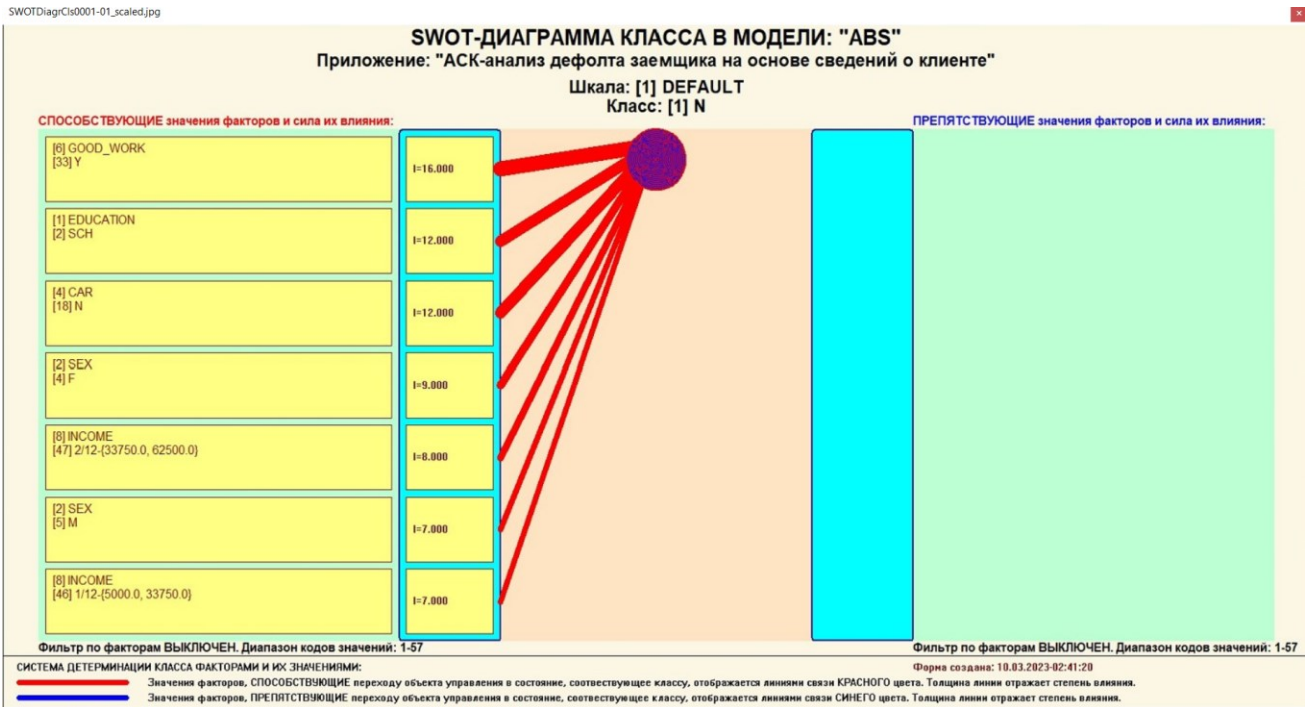


Рисунок 21 – SWOT-диаграмма

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.

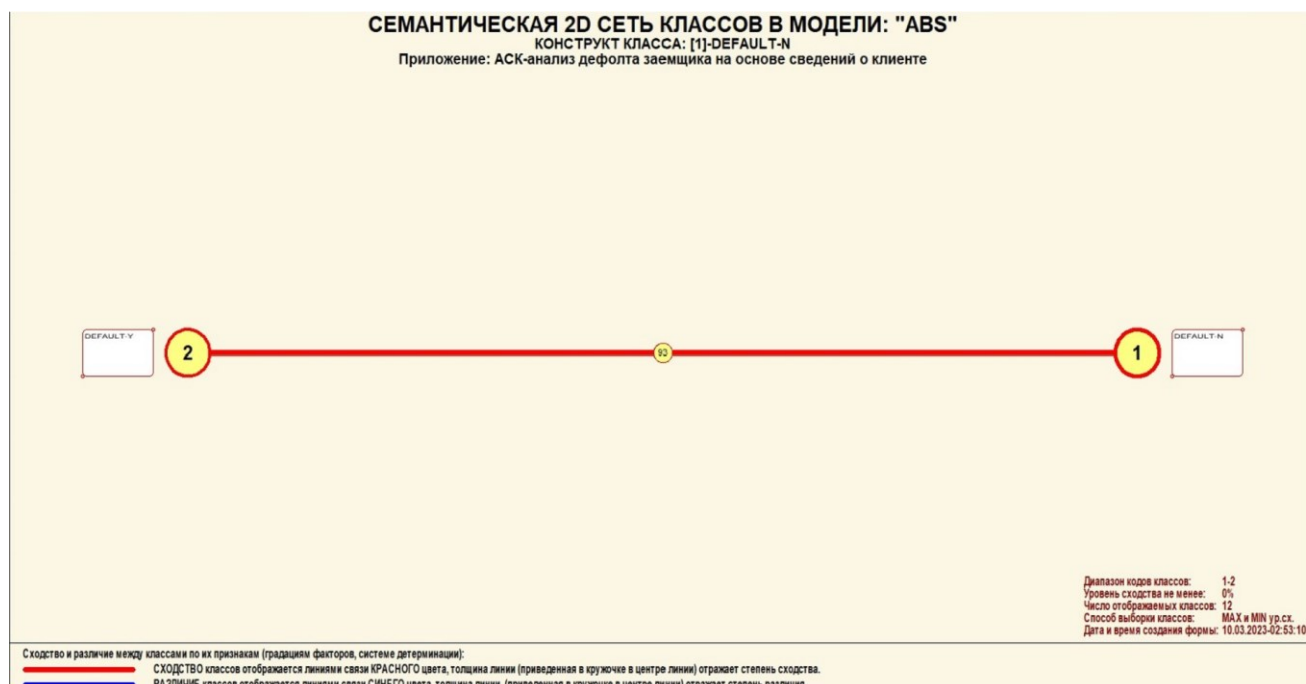


Рисунок 22 – Когнитивная диаграмма классов

2.3.2 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 23).

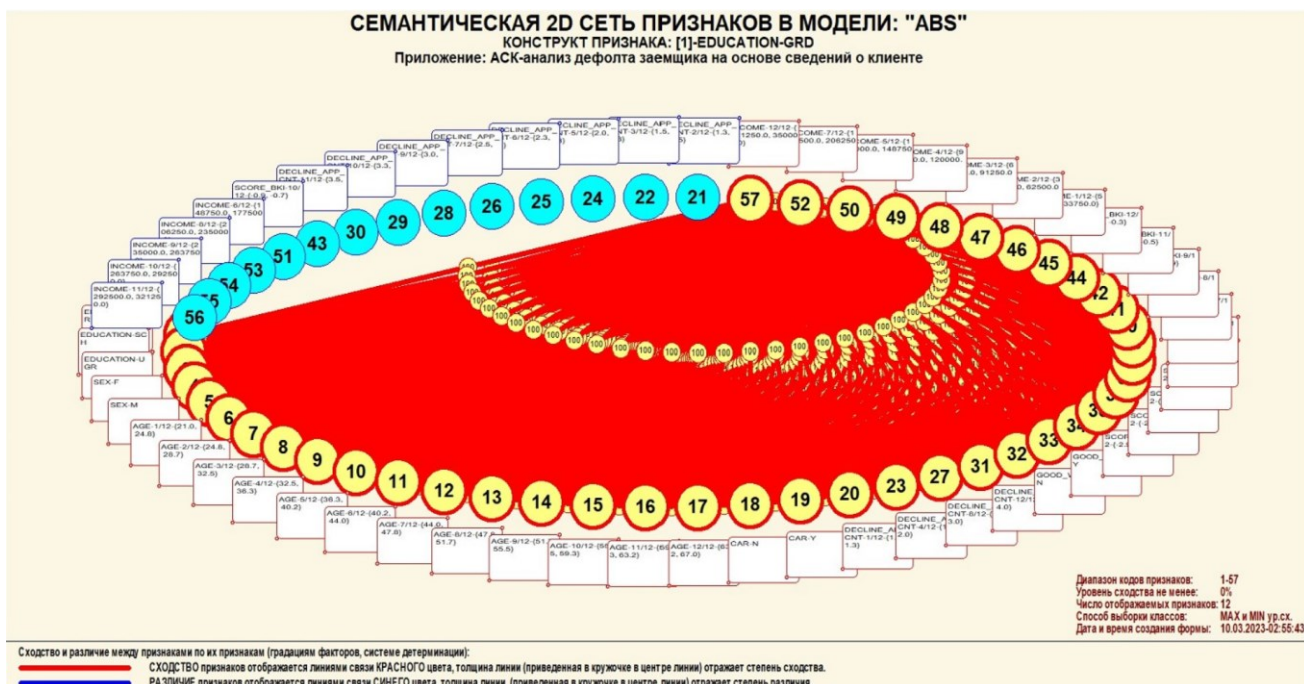


Рисунок 23 – График изменения межкластерных расстояний

2.3.3 Агломеративная когнитивная кластеризация значений факторов

На рисунках 24-25 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 дефолта заемщика на основе сведений о клиенте"

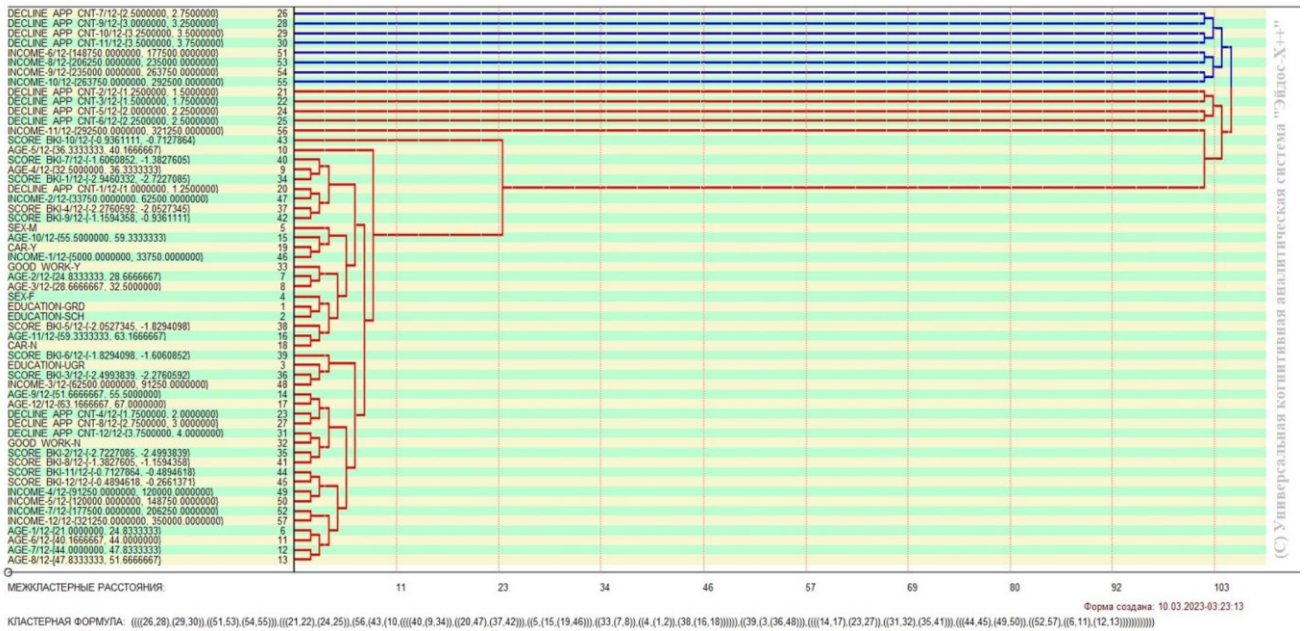


Рисунок 24 – Дендрограмма когнитивной кластеризации признаков

ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "ABS"
 Приложение: "АСК-анализ дефолта заемщика на основе сведений о клиенте"

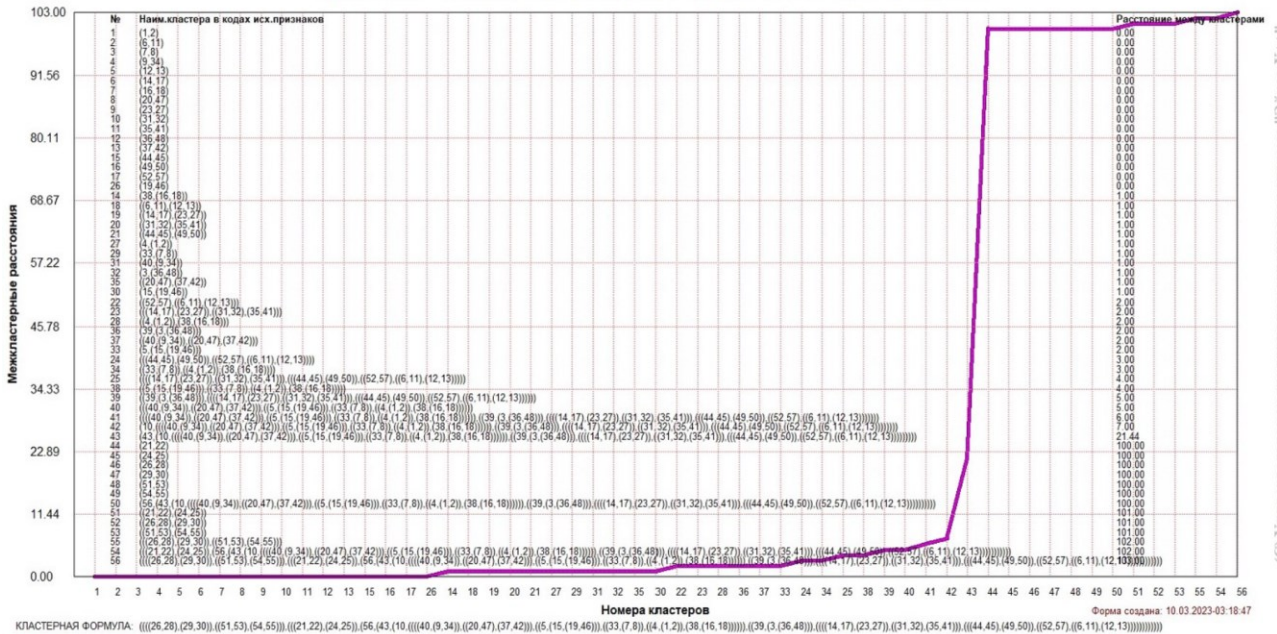


Рисунок 25 – График изменения межкластерных расстояний при когнитивной кластеризации признаков

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

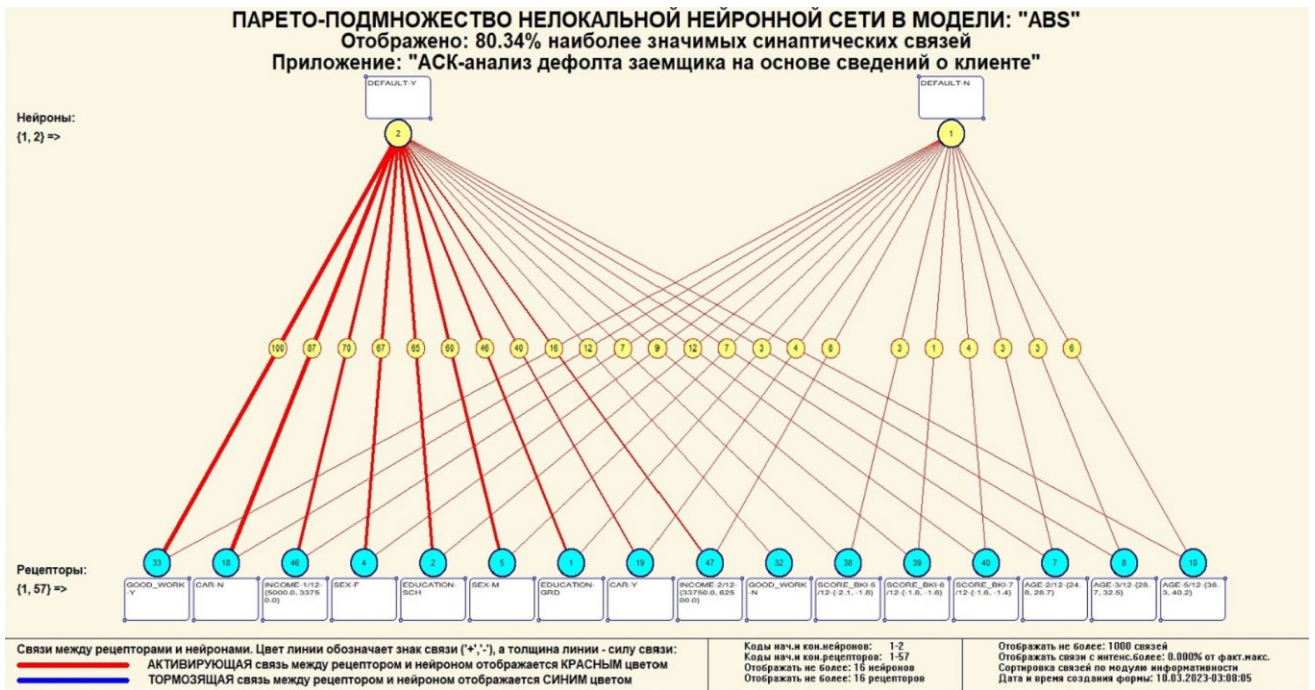


Рисунок 27 – Один слой нелокальной нейронной сети

2.3.5 3D-интегральные когнитивные карты

На рисунке 28 приведен фрагмент 3D-интегральной когнитивной карты, отражающий фрагмент СК-модели ABS.

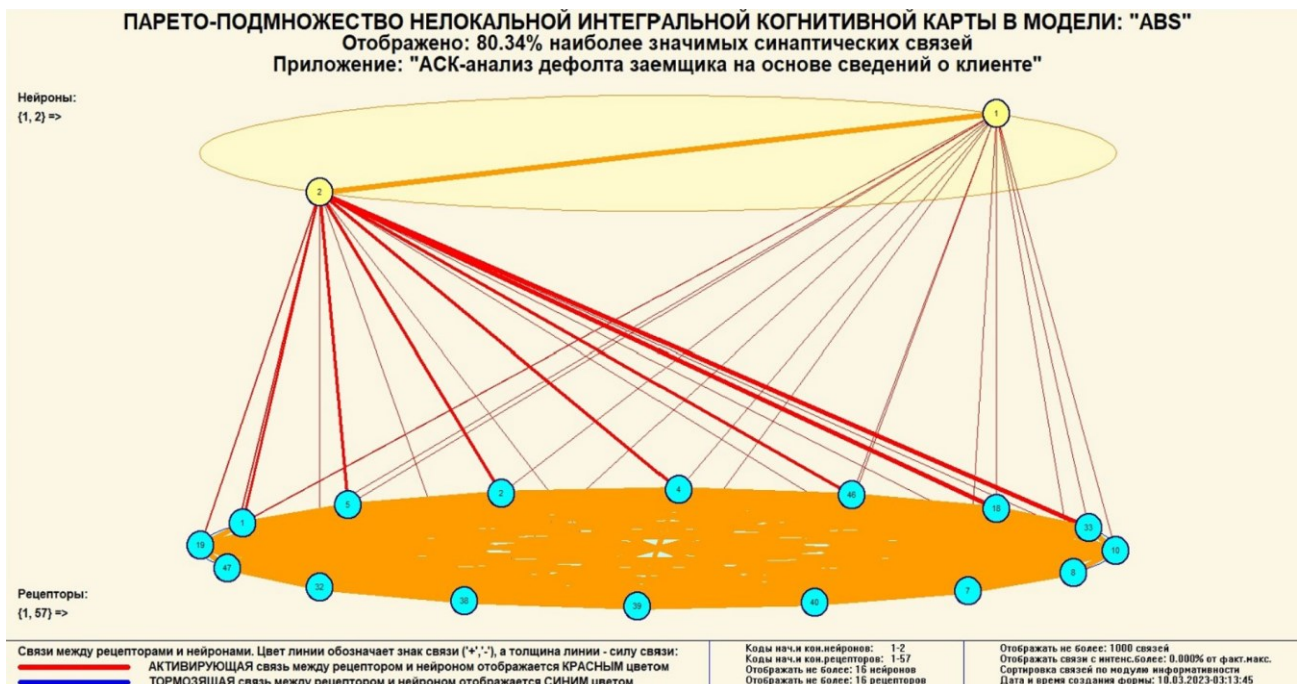


Рисунок 28 – 3D-интегральная когнитивная карта

2.3.6 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 29.

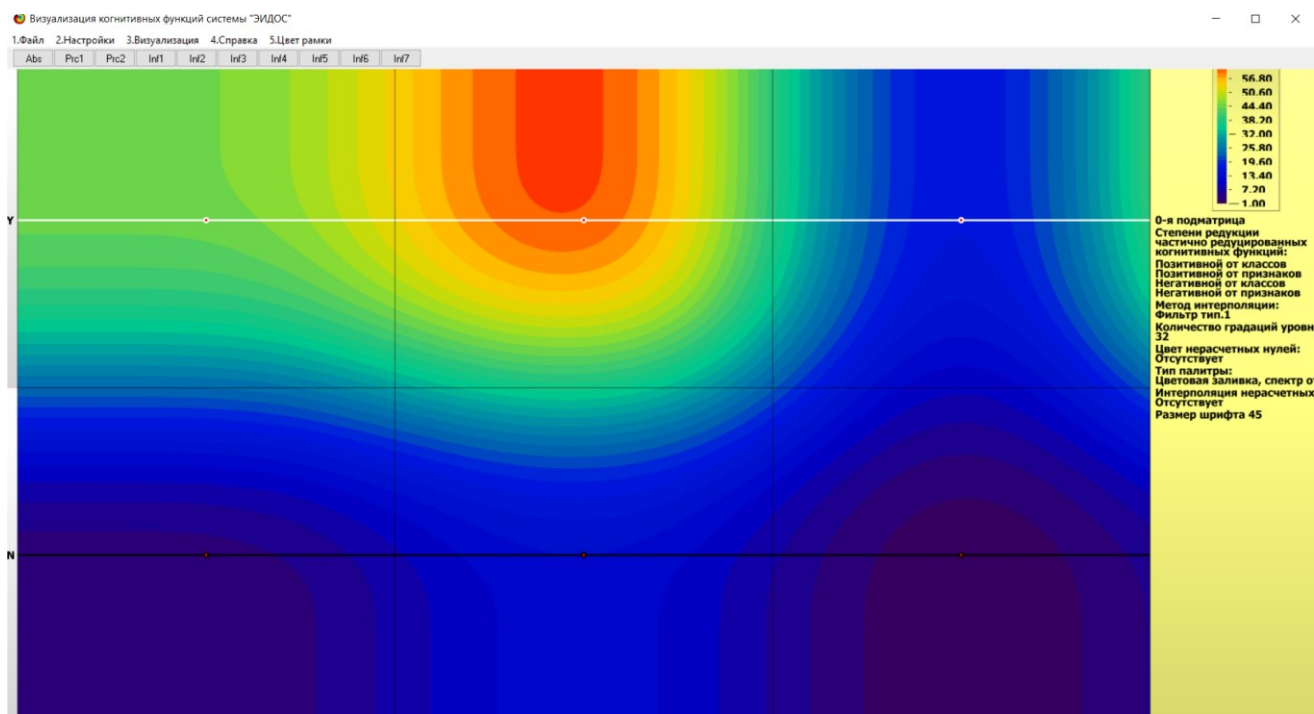


Рисунок 29 – Пример визуализации когнитивной функции по признаку 1
Female и 2 Male

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, 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. Оpubл. 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.