

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И Т. ТРУБИЛИНА»**

Факультет заочного обучения

Кафедра компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии
на тему: АСК-предсказание начала диабета на основе диагностических мер
(база данных диабета индейцев пима) на основе данных с портала Kaggle
Выполнил студент группы: ИТз1941 Оленич Сергей Алексеевич

Допущен к защите: _____

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

Защищен 07.04.2021
(дата)

Оценка отлично

Краснодар 2021.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ ИМЕНИ И.Т.
ТРУБИЛИНА»

Факультет заочного обучения

РЕЦЕНЗИЯ

на курсовую работу

Студент Оленич Сергей Алексеевич

курса 2 заочной формы обучения группы ИТз1941

Направление подготовки 09.04.02 «Информационные системы и технологии»

Наименование темы АСК-предсказание начала диабета на основе
диагностических мер (база данных диабета индейцев пима) на основе данных
с портала Kaggle

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор

(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)
1	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
2	Уровень и корректность использования в работе различных методов исследований	5
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
4	Ясность, четкость, последовательность и обоснованность изложения	5
5	Применение современных технологий обработки информации	5
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	5
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	5
8	Ответы на вопросы при защите	5

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____отлично_____

Рецензент _____ (Е.В. Луценко)

«07»_04_2021 г.

РЕФЕРАТ

Курсовая работа содержит: 37 страниц, 25 рисунков, 6 литературных источников.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, НЕЛОКАЛЬНЫЕ НЕЙТРОННЫЕ СЕТИ, НЕЙРОНЫ.

Целью работы является проведение АСК-предсказание начала диабета на основе диагностических мер (база данных диабета индейцев пима) на основе данных с портала Kaggle.

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	5
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	6
1.1. Описание решения	6
1.2. Преобразование исходных данных из CSV-формата в файл MS Excel	6
1.3. Ввод выборки в систему Aidos-X.....	8
1.4. Синтез и верификация статистических и интеллектуальных моделей ...	12
1.5. Виды моделей системы Aidos-X.....	13
1.6. Результаты верификации моделей	15
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	19
2.1. Решение задачи идентификации.....	19
2.2. Когнитивные функции.....	23
2.3. SWOT и PEST матрицы и диаграммы.....	28
2.4. Кластерно-конструктивный анализ признаков	31
2.5. Нелокальные нейроны и нейронные сети.....	32
ЗАКЛЮЧЕНИЕ	34
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	36

ВВЕДЕНИЕ

Одним из важных и перспективных направлений развития современных информационных технологий является создание систем искусственного интеллекта. Существует множество альтернатив систем искусственного интеллекта, но возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи исследования сведений о мировом счастье на основе АСК-анализа данных с портала Kaggle.

Для достижения поставленной цели необходимы свободный доступ к тестовым исходным данным и методика, которая поможет преобразовать эти данные в форму, которая необходима для работы в системе искусственного интеллекта. Удачным выбором является сборник баз данных Kaggle.

В данной курсовой работе использована база данных «World Happiness Report 2021» из банка исходных данных по задачам искусственного интеллекта — Kaggle.

Для решения задачи используются стандартные возможности MS Office Word и Excel, а также система искусственного интеллекта «Aidos-X++».

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

В соответствии с методологией АСК-анализа решение поставленной задачи проведем в три этапа:

1. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы «Эйдос».
2. Синтез и верификация моделей предметной области.
3. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

1.2. Преобразование исходных данных из CSV-формата в файл MS Excel

С электронного ресурса [kaggle.com](https://www.kaggle.com/uciml/pima-indians-diabetes-database) возьмем набор данных «Pima Indians Diabetes Database» <https://www.kaggle.com/uciml/pima-indians-diabetes-database>

Этот набор данных взят из Национального института диабета, болезней органов пищеварения и почек. Цель набора данных — диагностически предсказать, есть ли у пациента диабет, на основе определенных диагностических измерений, включенных в набор данных. На выбор этих экземпляров из более крупной базы данных было наложено несколько ограничений. В частности, все пациенты здесь — женщины не моложе 21 года, принадлежащие к индейцам пима.

Наборы данных состоят из нескольких медицинских предикторов и одной целевой переменной «Результат» (Outcome). Предикторы включают в себя количество беременностей были у пациентки, их ИМТ, уровня инсулина, возраст и так далее.

CSV файл содержит следующие данные:

— Pregnancies — Количество беременностей;

- Glucose — Концентрация глюкозы в плазме крови после 2-часового орального теста на толерантность к глюкозе;
- Blood Pressure — Диастолическое артериальное давление (мм/НГ);
- Skin Thickness — Толщина трицепса (мм);
- Insulin — 2 часа сывороточного инсулина (м.е./мл);
- BMI — Индекс массы тела (кг/м в квадрате);
- Diabetes Pedigree Function — Функция, которая определяет риск развития диабета 2 типа на основе семейного анамнеза. Чем больше функция, тем выше риск развития диабета 2 типа;
- Age — Возраст пациента;
- Outcome — Результат, диагностирован ли у человека диабет 2 типа (1 — да, 0 — нет).

Для конвертации CSV-файла в XLS был использован онлайн конвертер: <https://document.online-convert.com/ru/convert/csv-to-excel>

Результат конвертации представлен на рисунке 1. Таким образом, в качестве классификационной шкалы решено использовать столбец I — «Outcome». Остальные шкалы являются описательными.

	A	B	C	D	E	F	G	H	I
1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	6	148	72	35	0	33,6	0,627	50	Yes
3	1	85	66	29	0	26,6	0,351	31	No
4	8	183	64	0	0	23,3	0,672	32	Yes
5	1	89	66	23	94	28,1	0,167	21	No
6	0	137	40	35	168	43,1	2,288	33	Yes
7	5	116	74	0	0	25,6	0,201	30	No
8	3	78	50	32	88	31	0,248	26	Yes
9	10	115	0	0	0	35,3	0,134	29	No
10	2	197	70	45	543	30,5	0,158	53	Yes
11	8	125	96	0	0	0	0,232	54	Yes
12	4	110	92	0	0	37,6	0,191	30	No
13	10	168	74	0	0	38	0,537	34	Yes
14	10	139	80	0	0	27,1	1,441	57	No
15	1	189	60	23	846	30,1	0,398	59	Yes

Рисунок 1 — Фрагмент обучающей выборки

1.3. Ввод выборки в систему Aidos-X

Теперь, когда имеется обучающая выборка в формате *.xls, можно импортировать ее в систему Aidos-X.

Скопируем данную выборку в папку Inp_data и переименуем ее в Inp_data. Далее был использован универсальный программный интерфейс импорта данных в систему Aidos-X (режим 2.3.2.2), результат заполнения данной формы представлен на рисунке 2.

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

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

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:

Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

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

Ok Cancel

Рисунок 2 — Ввод обучающей выборки

Стоит отметить следующие настройки:

- Тип файла — xls;
- Классификационные шкалы — 9;
- Описательные шкалы — 2-8;
- Применяется спец. интерпретация текстовых полей признаков;
- В качестве признаков рассматриваются слова, без проведения лемитизации.

В форме задания размерности вносим корректировки (рисунок 3) и пересчитываем шкалы и градации.

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	0	0	0,00	7	35	5,00
Текстовые	1	2	2,00	0	0	0,00
ВСЕГО:	1	2	2,00	7	35	5,00

Задать число интервалов (градаций) в шкале: В описательных шкалах:

Рисунок 3 — Задание размерности модели системы Aidos-X

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

Стадии исполнения процесса

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

ПРОЦЕСС ФОРМАЛИЗАЦИИ ПРЕДМЕТНОЙ ОБЛАСТИ ЗАВЕРШЕН УСПЕШНО !!!

Прогноз времени исполнения

Начало: 10:28:56 Окончание: 10:29:01

100%

Прошло: 0:00:05 Осталось: 0:00:00

Рисунок 4 — Процесс импорта данных из внешней БД «Inp_data.xls» в систему «Эйдос»

В результате формируются классификационные и описательные шкалы и градации, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Этим самым полностью автоматизировано выполняется 2-й этап АСК-анализа «Формализация предметной области». Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1 (рисунок 5).

2.1. Классификационные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы
1	OUTCOME	1	1/2-No
		2	2/2-Yes

Помощь | Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Удаление и перекодирование | Графики будущих сценариев

Рисунок 5 — Классификационные шкалы и градации

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 6), а обучающей выборки режим 2.3.1. (рисунок 7):

2.2. Описательные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	GLUCOSE	1	1/5-{44.0000000, 95.0000000}
2	BLOODPRESSURE	2	2/5-{95.0000000, 109.0000000}
3	SKINTHICKNESS	3	3/5-{109.0000000, 125.0000000}
4	INSULIN	4	4/5-{125.0000000, 147.0000000}
5	BMI	5	5/5-{147.0000000, 199.0000000}
6	DIABETESPEDIGREEFUNCTION		
7	AGE		

Помощь | Доб. шкалу | Доб. град. шкалы | Копир. шкалу | Копир. град. шкалы | Копир. шкалу с град. | Удал. шкалу с град. | Удал. град. шкалы | Перекодировать | Очистить | Графики прошлых сценариев

Рисунок 6 — Описательные шкалы и градации (фрагмент)

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

Код объекта	Наименование объекта	Дата	Время
1	6		
2	1		
3	8		
4	1		
5	0		
6	5		
7	3		
8	10		
9	2		
10	8		
11	4		

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4	Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	2	0	0	0	1	5	8	14	23	29	35	0

Помощь | Скопировать обуч.выб.в расп. | Добавить объект | Добавить классы | Добавить признаки | Удалить объект | Удалить классы | Удалить признаки | Очистить БД

Рисунок 7 — Корректировка обучающей выборки (фрагмент)

Для просмотра эвентологических баз данных (баз событий) необходимо запустить режим 2.4 (рисунок 8),

2.4. Просмотр эвентологических баз данных (баз событий). Текущая модель: "INF1"

№	Наименование объекта	9. OUTCOME	2. GLUCOSE	3. BLOODPRESSURE	4. SKINTHICKNESS	5. INSULIN	6. BMI	7. DIABETESPEC
1	6	2	5	8	14		23	
2	1	1	1	7	13		22	
3	8	2	5	7			21	
4	1	1	1	7	12	17	22	
5	0	2	4	6	14	19	25	
6	5	1	3	8			21	
7	3	2	1	6	13	17	23	
8	10	1	3				24	
9	2	2	5	7	15	20	23	
10	8	2	3	10				
11	4	1	3	10			24	
12	10	2	5	8			25	
13	10	1	4	9			22	
14	1	2	5	6	12	20	22	
15	5	2	5	8	11	19	21	
16	7	2	2				22	
17	0	2	3	10	15	20	25	
0	1	9	2	3	4	5	6	7

Рисунок 8 — Просмотр базы событий (фрагмент)

Тем самым создаются все необходимые и достаточные предпосылки для выявления силы и направления причинно-следственных связей между значениями факторов и результатами их совместного системного воздействия (с учетом нелинейности системы [11]).

1.4. Синтез и верификация статистических и интеллектуальных моделей

Далее запускаем режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей (рисунок 9).

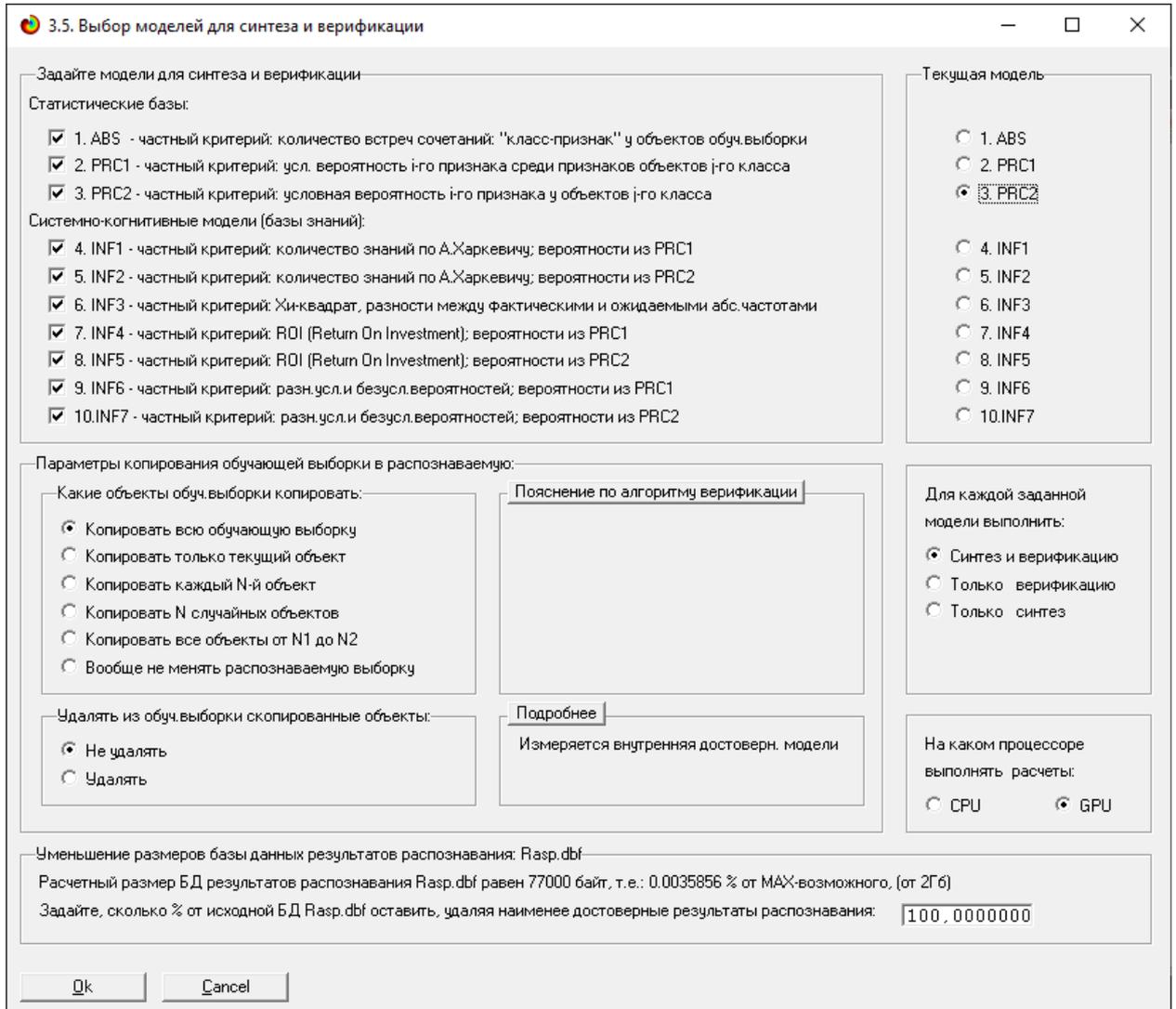


Рисунок 9 — Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию, приведенные на рисунке 9. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 10.

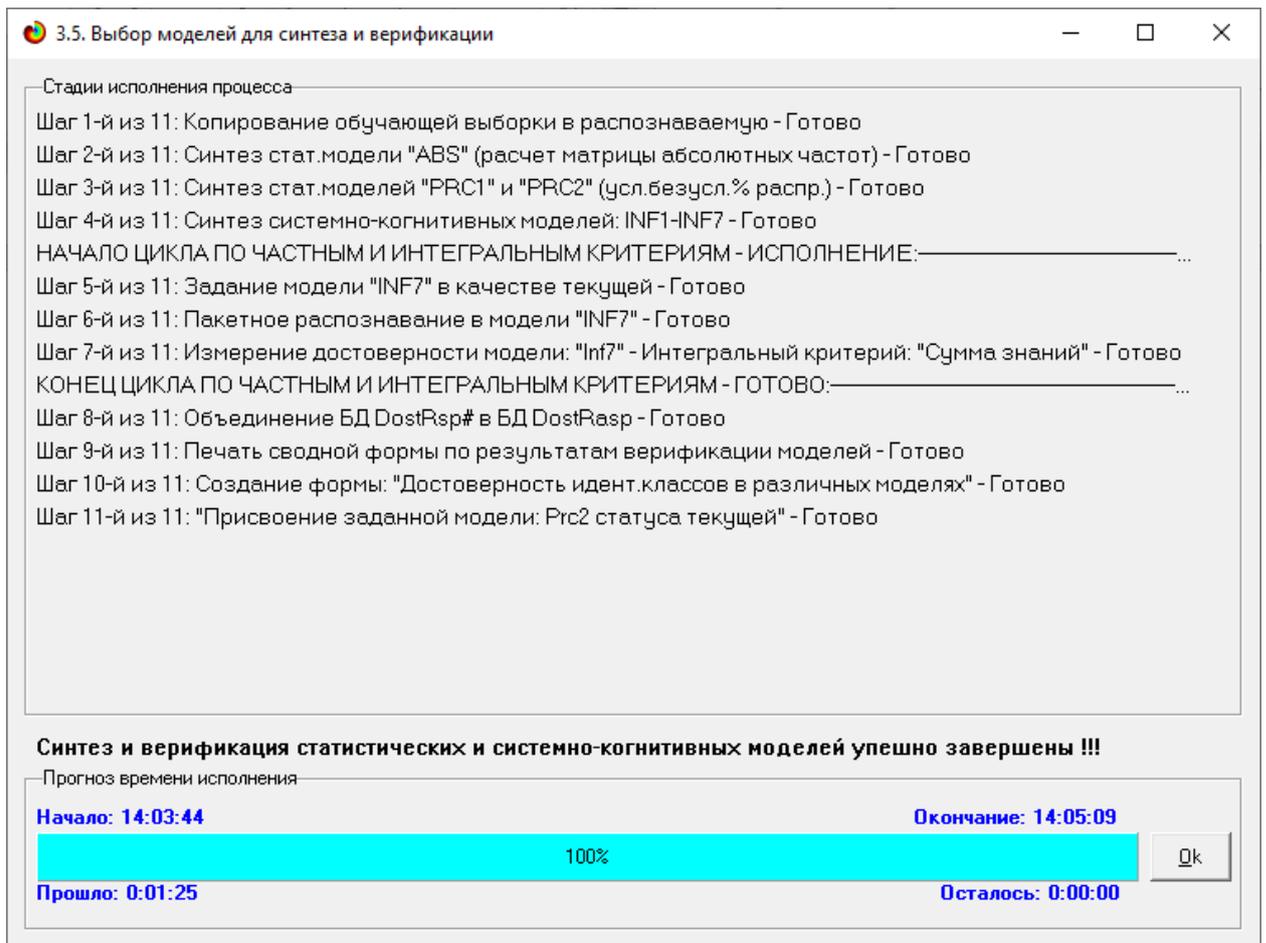


Рисунок 10 — Синтез и верификация статистических моделей
и моделей знаний

1.5. Виды моделей системы Aidos-X

Рассмотрим решение задачи идентификации на примере нескольких моделей, в которой 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.

То есть частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот (рисунок 11) в матрицы условных и безусловных процентных распределений, и матрицы знаний (рисунок 12 и 13).

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

Код признака	Наименование описательной шкалы и градации	1. OUTCOME 1/2 NO	2. OUTCOME 2/2 YES	Сумма	Среднее	Средн. квадрат. откл.
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	143	12	155	77.50	92.63
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	124	30	154	77.00	66.47
3	GLUCOSE-3/5-(109.0000000, 125.0000000)	109	48	157	78.50	43.13
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	83	66	149	74.50	12.02
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	38	110	148	74.00	50.91
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	125	33	158	79.00	65.05
7	BLOODPRESSURE-2/5-(62.0000000, 70.0000000)	122	60	182	91.00	43.84
8	BLOODPRESSURE-3/5-(70.0000000, 75.0000000)	70	34	104	52.00	25.46
9	BLOODPRESSURE-4/5-(75.0000000, 82.0000000)	93	61	154	77.00	22.63
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	71	64	135	67.50	4.95
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	96	15	111	55.50	57.28
12	SKINTHICKNESS-2/5-(19.0000000, 26.0000000)	76	29	105	52.50	33.23
13	SKINTHICKNESS-3/5-(26.0000000, 32.0000000)	89	48	137	68.50	28.99
14	SKINTHICKNESS-4/5-(32.0000000, 38.0000000)	42	38	80	40.00	2.83
15	SKINTHICKNESS-5/5-(38.0000000, 99.0000000)	58	50	108	54.00	5.66
16	INSULIN-1/5-(14.0000000, 67.0000000)	72	6	78	39.00	46.67
17	INSULIN-2/5-(67.0000000, 105.0000000)	70	14	84	42.00	39.60
18	INSULIN-3/5-(105.0000000, 145.0000000)	46	28	74	37.00	12.73
19	INSULIN-4/5-(145.0000000, 210.0000000)	42	41	83	41.50	0.71
20	INSULIN-5/5-(210.0000000, 846.0000000)	34	41	75	37.50	4.95
21	BMI-1/5-(18.2000000, 26.1000000)	139	12	151	75.50	89.80
22	BMI-2/5-(26.1000000, 30.2000000)	109	42	151	75.50	47.38
23	BMI-3/5-(30.2000000, 33.8000000)	85	70	155	77.50	10.61
24	BMI-4/5-(33.8000000, 37.8000000)	84	64	148	74.00	14.14
25	BMI-5/5-(37.8000000, 67.1000000)	74	78	152	76.00	2.83
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.2190000)	120	34	154	77.00	60.81

Рисунок 11 — Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"

Код признака	Наименование описательной шкалы и градации	1. OUTCOME 1/2 NO	2. OUTCOME 2/2 YES	Сумма	Среднее	Средн. квадрат. откл.
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	0.040	-0.177	-0.136	-0.068	0.154
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	0.024	-0.068	-0.043	-0.022	0.065
3	GLUCOSE-3/5-(109.0000000, 125.0000000)	0.007	-0.014	-0.007	-0.004	0.015
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	-0.019	0.030	0.010	0.005	0.034
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	-0.111	0.091	-0.020	-0.010	0.142
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	0.022	-0.059	-0.037	-0.019	0.058
7	BLOODPRESSURE-2/5-(62.0000000, 70.0000000)	0.003	-0.005	-0.003	-0.001	0.006
8	BLOODPRESSURE-3/5-(70.0000000, 75.0000000)	0.003	-0.006	-0.003	-0.002	0.007
9	BLOODPRESSURE-4/5-(75.0000000, 82.0000000)	-0.010	0.016	0.007	0.003	0.018
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	-0.026	0.038	0.012	0.006	0.045
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	0.033	-0.111	-0.078	-0.039	0.102
12	SKINTHICKNESS-2/5-(19.0000000, 26.0000000)	0.012	-0.026	-0.014	-0.007	0.027
13	SKINTHICKNESS-3/5-(26.0000000, 32.0000000)	-0.001	0.002	0.001	0.000	0.002
14	SKINTHICKNESS-4/5-(32.0000000, 38.0000000)	-0.026	0.038	0.012	0.006	0.045
15	SKINTHICKNESS-5/5-(38.0000000, 99.0000000)	-0.023	0.035	0.011	0.006	0.041
16	INSULIN-1/5-(14.0000000, 67.0000000)	0.041	-0.177	-0.137	-0.068	0.154
17	INSULIN-2/5-(67.0000000, 105.0000000)	0.028	-0.086	-0.058	-0.029	0.081
18	INSULIN-3/5-(105.0000000, 145.0000000)	-0.006	0.011	0.005	0.002	0.012
19	INSULIN-4/5-(145.0000000, 210.0000000)	-0.030	0.042	0.012	0.006	0.052
20	INSULIN-5/5-(210.0000000, 846.0000000)	-0.043	0.054	0.011	0.005	0.069
21	BMI-1/5-(18.2000000, 26.1000000)	0.040	-0.174	-0.133	-0.067	0.151
22	BMI-2/5-(26.1000000, 30.2000000)	0.011	-0.025	-0.014	-0.007	0.026
23	BMI-3/5-(30.2000000, 33.8000000)	-0.021	0.032	0.011	0.005	0.037
24	BMI-4/5-(33.8000000, 37.8000000)	-0.017	0.027	0.010	0.005	0.031
25	BMI-5/5-(37.8000000, 67.1000000)	-0.035	0.047	0.012	0.006	0.058
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.2190000)	0.021	-0.053	-0.032	-0.016	0.052

Рисунок 12 — Модель INF1 (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. OUTCOME 1/2 NO	2. OUTCOME 2/2 YES	Сумма	Среднее	Средн. квадрат. откл.
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	0.417	-0.778	-0.361	-0.181	0.845
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	0.237	-0.442	-0.205	-0.102	0.480
3	GLUCOSE-3/5-(109.0000000, 125.0000000)	0.066	-0.124	-0.057	-0.029	0.135
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	-0.144	0.269	0.125	0.062	0.293
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	-0.606	1.130	0.524	0.262	1.227
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	0.215	-0.401	-0.186	-0.093	0.436
7	BLOODPRESSURE-2/5-(62.0000000, 70.0000000)	0.030	-0.055	-0.026	-0.013	0.060
8	BLOODPRESSURE-3/5-(70.0000000, 75.0000000)	0.034	-0.063	-0.029	-0.015	0.069
9	BLOODPRESSURE-4/5-(75.0000000, 82.0000000)	-0.072	0.135	0.063	0.031	0.147
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	-0.192	0.359	0.166	0.083	0.389
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	0.328	-0.613	-0.284	-0.142	0.666
12	SKINTHICKNESS-2/5-(19.0000000, 26.0000000)	0.112	-0.209	-0.097	-0.048	0.226
13	SKINTHICKNESS-3/5-(26.0000000, 32.0000000)	-0.002	0.004	0.002	0.001	0.004
14	SKINTHICKNESS-4/5-(32.0000000, 38.0000000)	-0.194	0.361	0.168	0.084	0.392
15	SKINTHICKNESS-5/5-(38.0000000, 99.0000000)	-0.175	0.327	0.152	0.076	0.355
16	INSULIN-1/5-(14.0000000, 67.0000000)	0.418	-0.780	-0.362	-0.181	0.847
17	INSULIN-2/5-(67.0000000, 105.0000000)	0.280	-0.522	-0.242	-0.121	0.567
18	INSULIN-3/5-(105.0000000, 145.0000000)	-0.045	0.084	0.039	0.020	0.092
19	INSULIN-4/5-(145.0000000, 210.0000000)	-0.223	0.416	0.193	0.096	0.451
20	INSULIN-5/5-(210.0000000, 846.0000000)	-0.304	0.567	0.263	0.131	0.615
21	BMI-1/5-(18.2000000, 26.1000000)	0.414	-0.772	-0.358	-0.179	0.839
22	BMI-2/5-(26.1000000, 30.2000000)	0.109	-0.203	-0.094	-0.047	0.220
23	BMI-3/5-(30.2000000, 33.8000000)	-0.158	0.294	0.136	0.068	0.320
24	BMI-4/5-(33.8000000, 37.8000000)	-0.128	0.239	0.111	0.055	0.260
25	BMI-5/5-(37.8000000, 67.1000000)	-0.252	0.471	0.218	0.109	0.511
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.2190000)	0.197	-0.367	-0.170	-0.085	0.399

Рисунок 13 — Модель INF5 (фрагмент)

1.6. Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями представлены на рисунке 14.

Наименование модели и частного критерия	Интегральный критерий	Юлота модели	Ф-мера Ван Ризбергера	Сумма модул. уровней сход. истинно-полож. решений (STP)	Сумма модул. уровней сход. истинно-отриц. решений (STN)	Сумма модул. уровней ложно-полож. решений (SFP)	Сумма модул. уровней ложно-отриц. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Ладенко	Средний уровень истинности решений
1. ABS - частный критерий: количество встреч сонеганий "Клас...	Корреляция абс частот с обр...	0.819	0.745	245.382	143.599	69.880	24.270	0.778	0.910	0.839	0.
1. ABS - частный критерий: количество встреч сонеганий "Клас...	Сумма абс частот по признак...	1.000	0.667	481.633		327.476		0.595	1.000	0.746	0.
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн частот с о...	0.819	0.744	245.382	143.599	69.880	24.270	0.778	0.910	0.839	0.
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн частот по приз...	1.000	0.667	513.186		417.338		0.552	1.000	0.711	0.
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн частот с о...	0.819	0.744	245.382	143.599	69.880	24.270	0.778	0.910	0.839	0.
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн частот по приз...	1.000	0.667	518.825		419.845		0.553	1.000	0.712	0.
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.768	0.740	175.614	181.680	45.682	35.234	0.794	0.833	0.813	0.
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.637	0.695	63.180	162.796	15.800	25.755	0.800	0.710	0.753	0.
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.768	0.740	175.614	181.680	45.682	35.234	0.794	0.833	0.813	0.
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.646	0.702	62.151	163.551	14.622	25.125	0.810	0.712	0.758	0.
6. INF3 - частный критерий: Хинквадат, разности между факти...	Семантический резонанс зна...	0.737	0.737	222.819	222.819	47.704	47.704	0.824	0.824	0.824	0.
6. INF3 - частный критерий: Хинквадат, разности между факти...	Сумма знаний	0.737	0.737	211.858	211.858	45.173	45.173	0.824	0.824	0.824	0.
7. INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.723	0.723	212.958	212.958	46.547	46.547	0.821	0.821	0.821	0.
7. INF4 - частный критерий: ROI (Return On Investment), веротно...	Сумма знаний	0.719	0.719	148.120	164.083	43.146	27.184	0.774	0.845	0.808	0.
8. INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	0.723	0.723	212.958	212.958	46.547	46.547	0.821	0.821	0.821	0.
8. INF5 - частный критерий: ROI (Return On Investment), веротно...	Сумма знаний	0.732	0.732	145.571	165.351	39.877	26.199	0.785	0.847	0.815	0.
9. INF6 - частный критерий: разн.усли безул.вероятностей; вер...	Семантический резонанс зна...	0.737	0.737	222.819	222.819	47.704	47.704	0.824	0.824	0.824	0.
9. INF6 - частный критерий: разн.усли безул.вероятностей; вер...	Сумма знаний	0.737	0.737	152.191	171.280	41.935	27.036	0.784	0.849	0.815	0.
10. INF7 - частный критерий: разн.усли безул.вероятностей; ве...	Семантический резонанс зна...	0.737	0.737	222.874	222.874	47.581	47.581	0.824	0.824	0.824	0.
10. INF7 - частный критерий: разн.усли безул.вероятностей; ве...	Сумма знаний	0.742	0.742	149.660	172.744	38.774	26.096	0.794	0.852	0.822	0.

Рисунок 14 — Оценки достоверности моделей

Наиболее достоверной в данном приложении оказалась модель INF3 при интегральном критерии «Семантический резонанс знаний». При этом достоверность модели (F-мера Ван Ризбергена) составляет 0.737, а достоверность модели (L1-мера профессора Луценко) — 0.824. L1-мера профессора Луценко является более достоверной, по сравнению с F-мерой Ван Ризбергена. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 100%. 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 15).

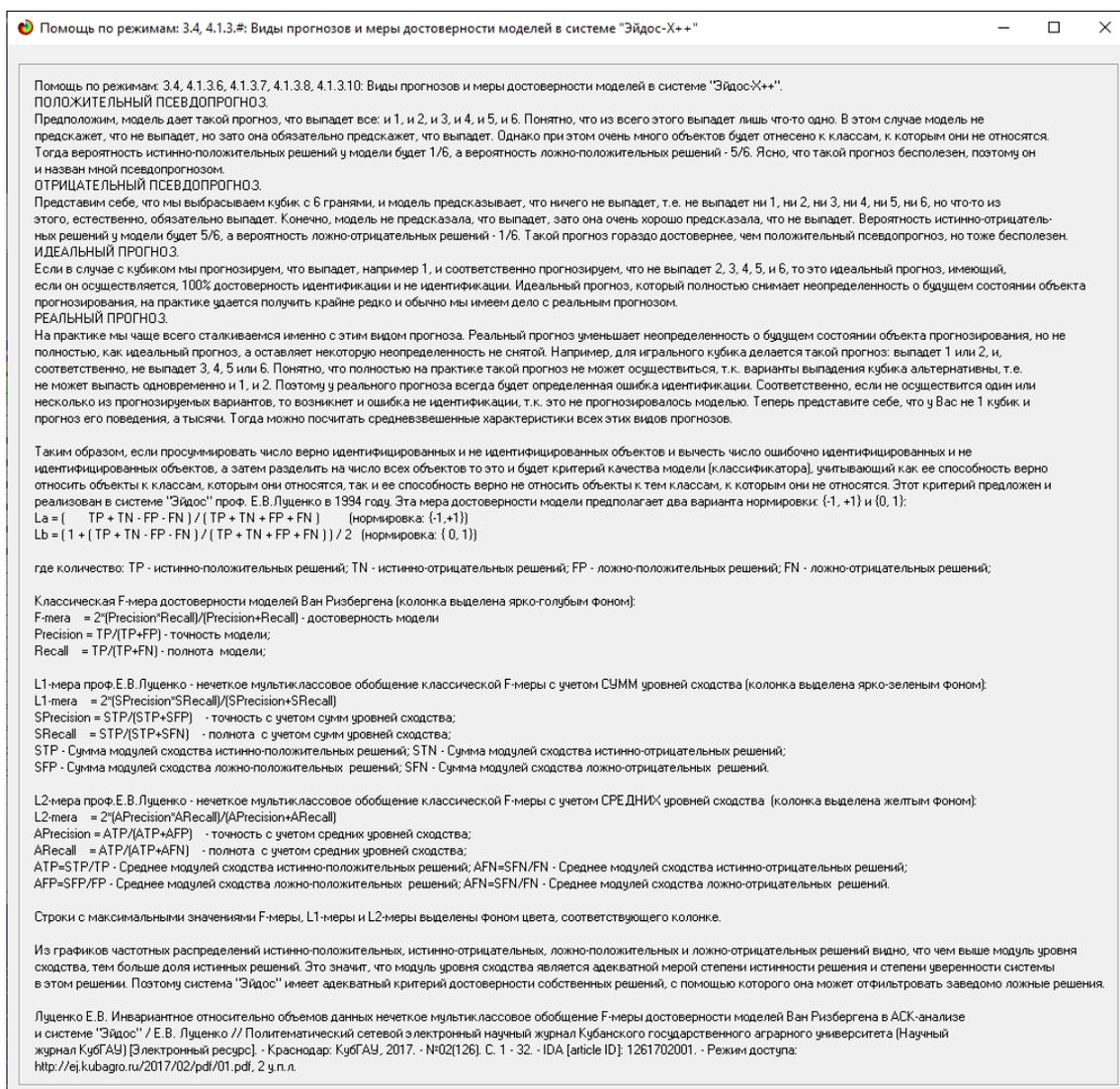


Рисунок 15 — Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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. Этим и оправдано применение моделей знаний и 21 интеллектуальных технологий.

На рисунке 16 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам идентификации характеристик оружия в СК-модели INF3 по данным обучающей выборки:

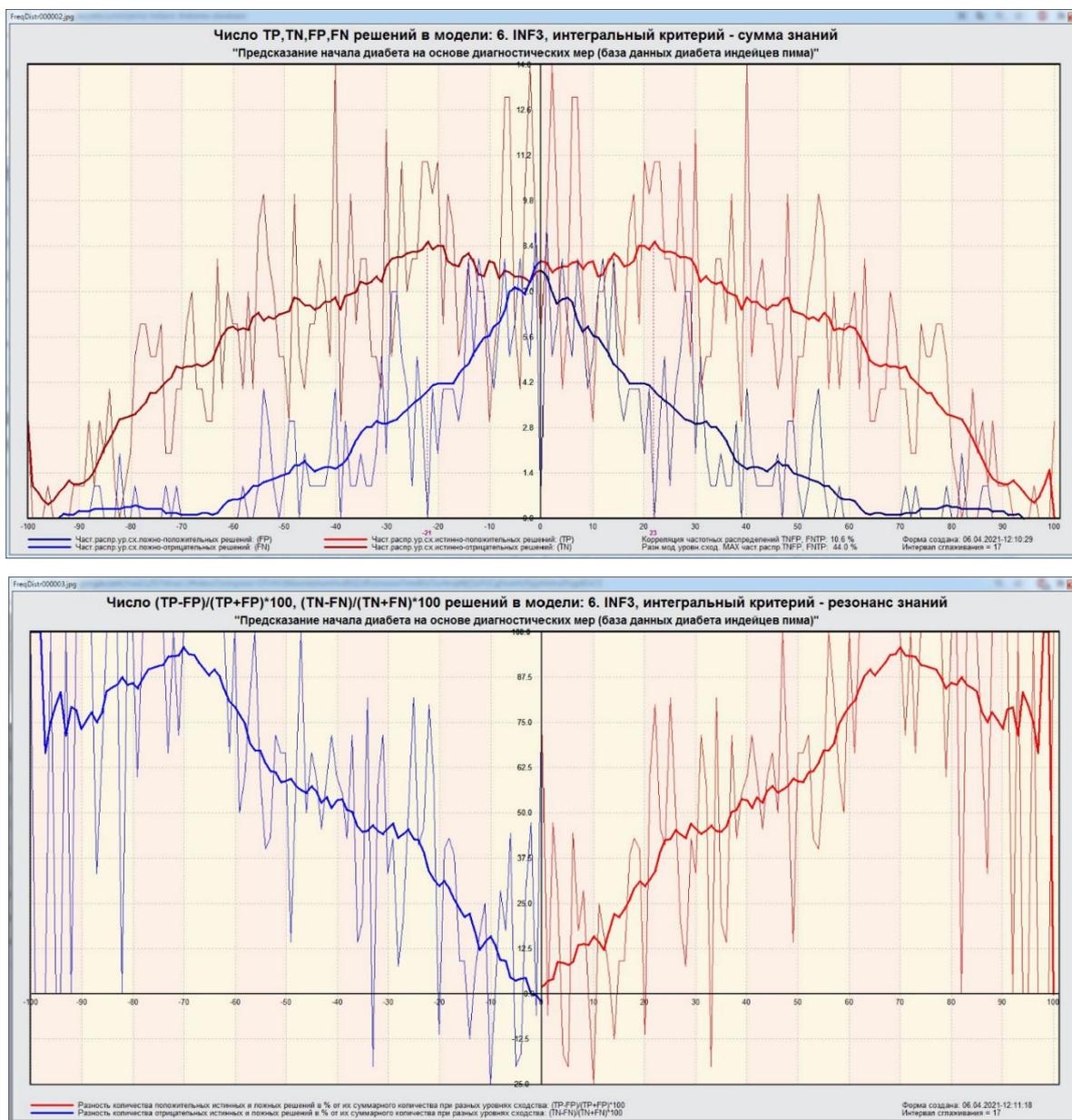


Рисунок 16 — Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf3

Рисунок 16 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе.

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

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу идентификации вида оружия по его характеристикам и другие задачи.

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

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

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

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

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

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос», присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (рисунок 17):

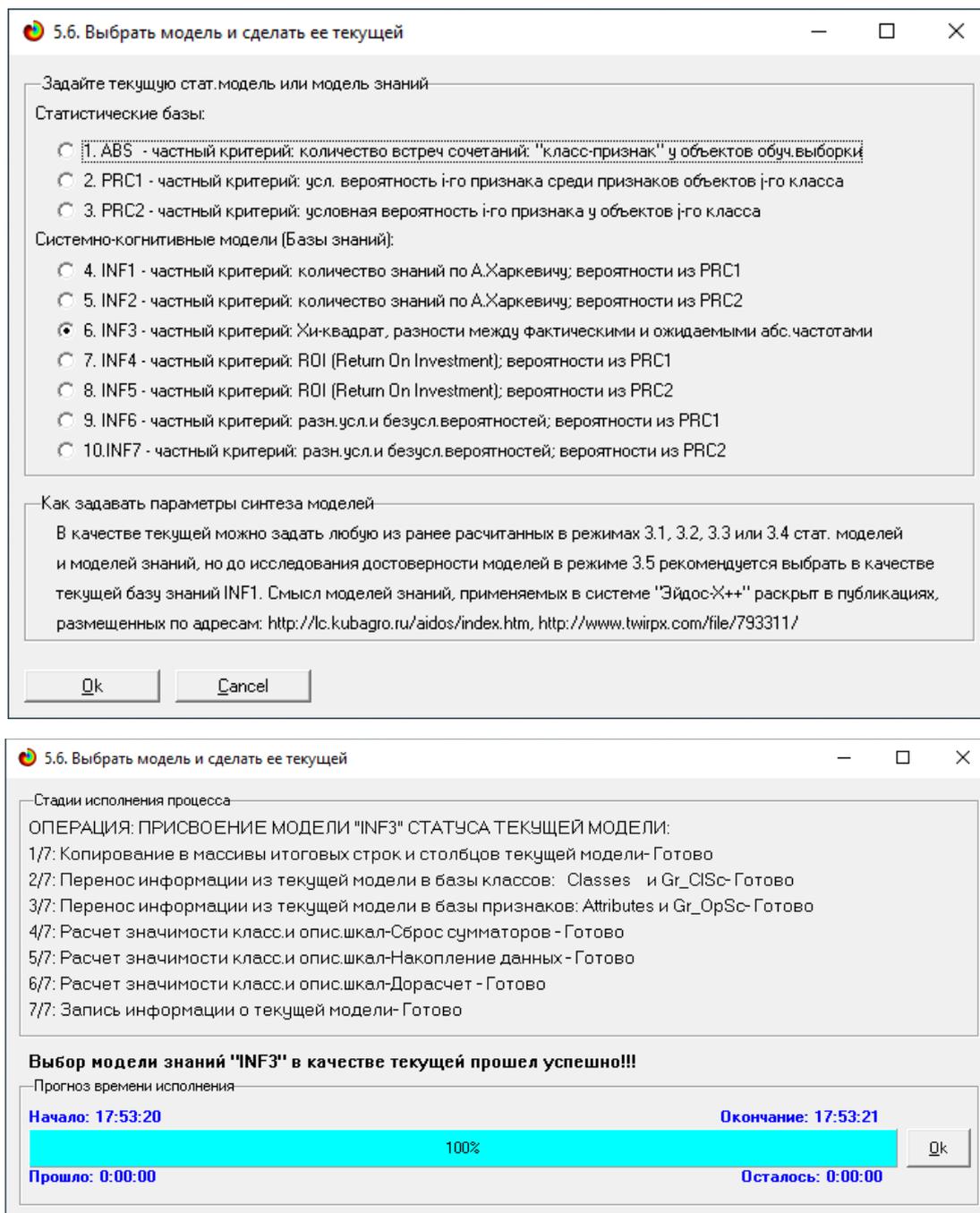


Рисунок 17. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

Решим задачу системной идентификации, т.е. определение класса оружия на основе обучающей выборки в наиболее достоверной СК-модели INF3 на GPU (рисунок 18).

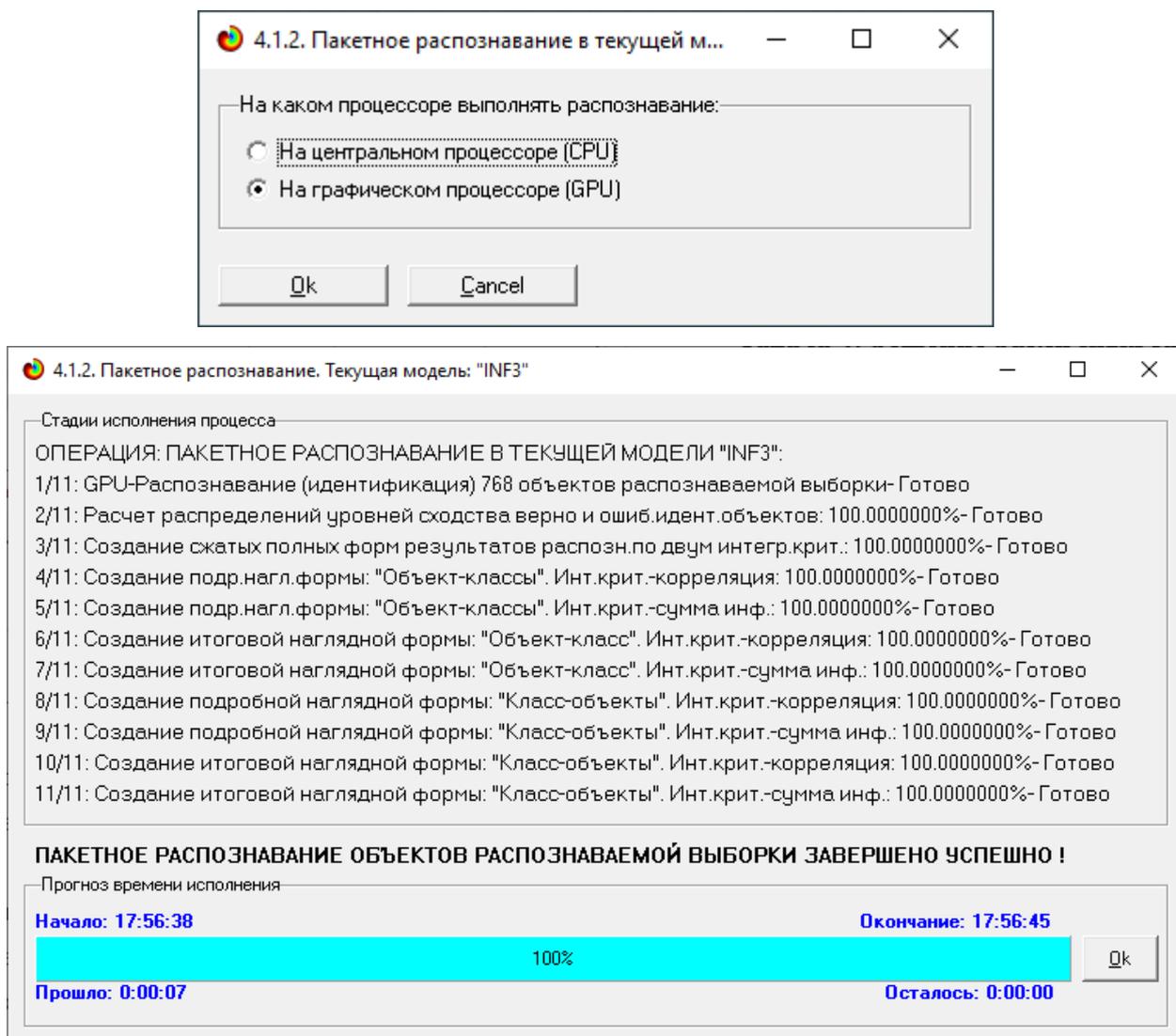


Рисунок 18. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Из рисунка 18 видно, что процесс идентификации занял 7 секунд.

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:

1. Подробно наглядно: «Объект — классы».
2. Подробно наглядно: «Класс — объекты».
3. Итоги наглядно: «Объект — классы».
4. Итоги наглядно: «Класс — объекты».
5. Подробно сжато: «Объект — классы».
6. Обобщенная форма по достоверности моделей при разных интегральных критериях.
7. Обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям.
8. Статистический анализ результатов идентификации по классам, моделям и интегральным критериям.
9. Распознавание уровня сходства при разных моделях и интегральных критериях.
10. Достоверность идентификации классов при разных моделях и интегральных критериях.

Отметим, что 99% этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этого прогнозирования. Эти формы отражают результаты прогнозирования в различных разрезах и обобщениях:

На рисунке 19 приведены примеры прогнозов высокой и низкой достоверности частоты объектов и классов в модели INF3 на основе наблюдения предыстории их развития:

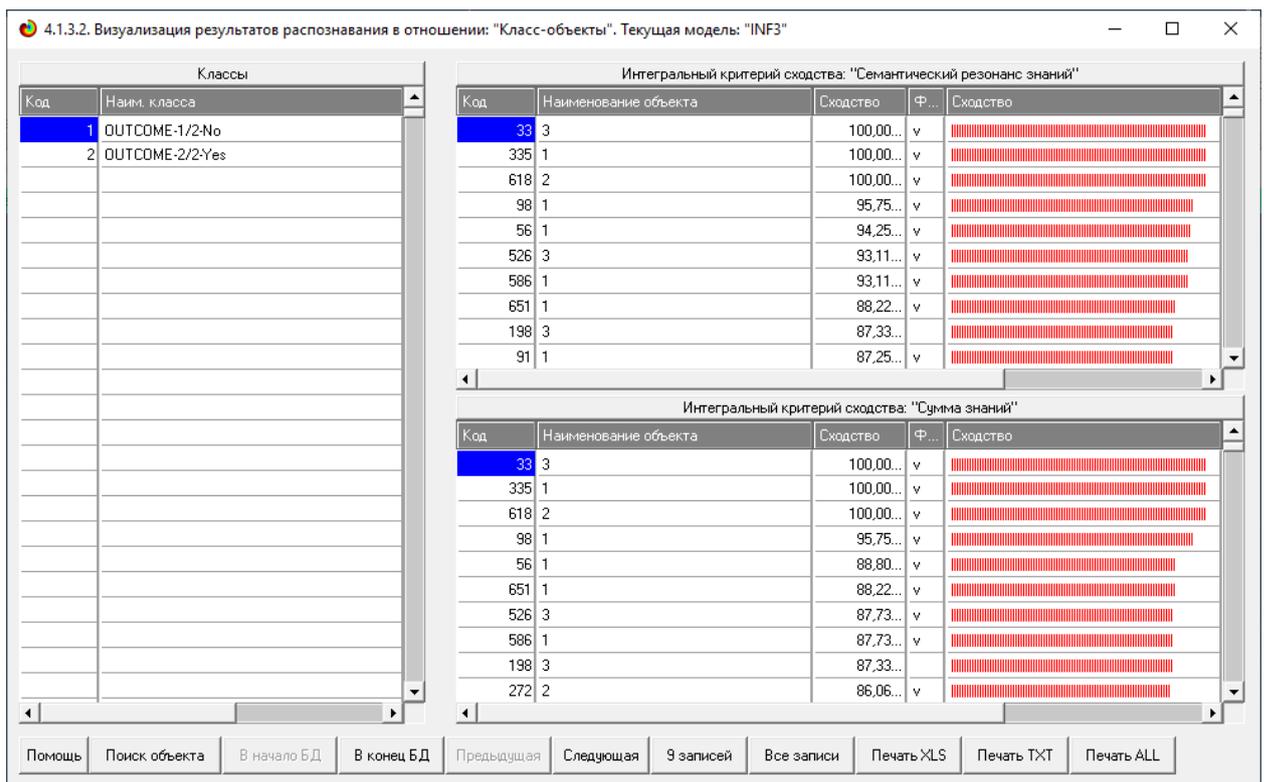
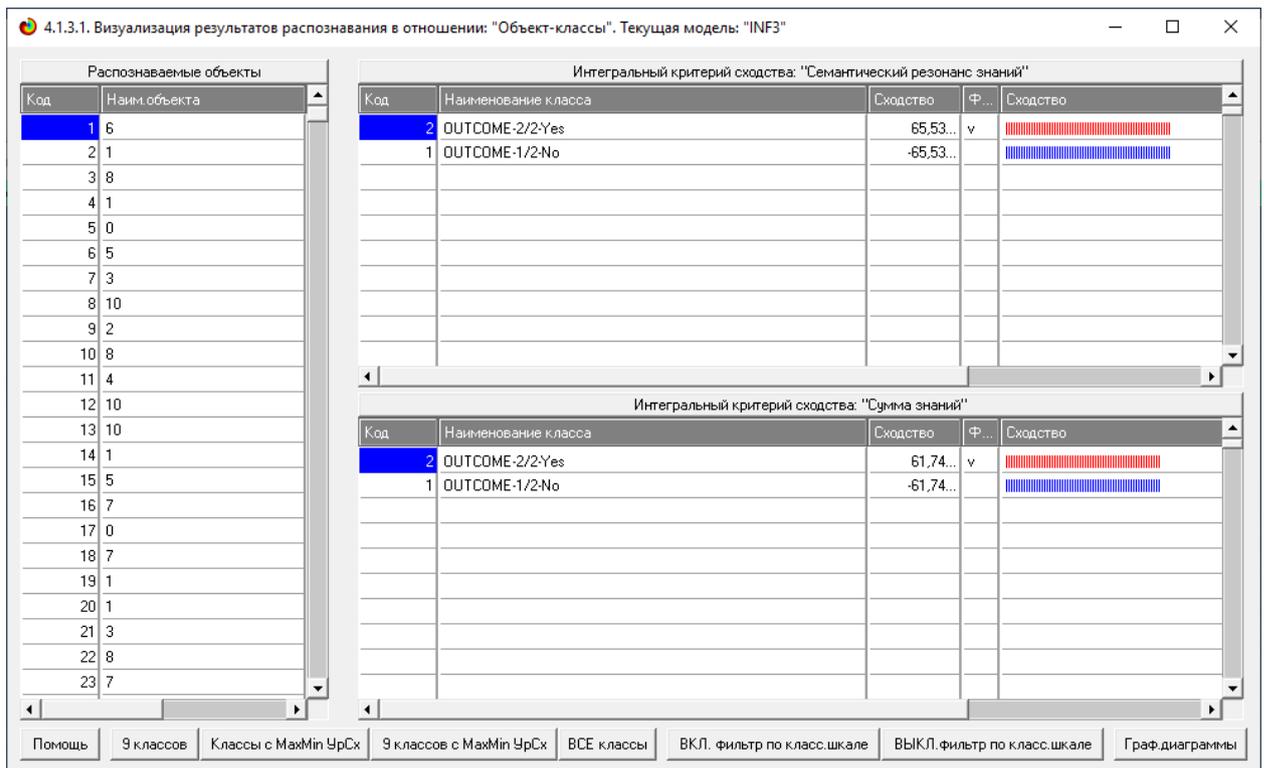


Рисунок 19. Выходные формы по результатам идентификации вида оружия по его характеристикам

Символ «√» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка 13 видно, что

результаты идентификации являются отличными, естественно при учете информации из рисунка 9 о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 42%, т.е., по сути, результаты с более низким уровнем сходства надо просто игнорировать.

2.2. Когнитивные функции

Рассмотрим режим 4.5, в котором реализована возможность визуализации когнитивных функций для любых моделей и любых сочетаний классификационных и описательных шкал (рисунок 20).

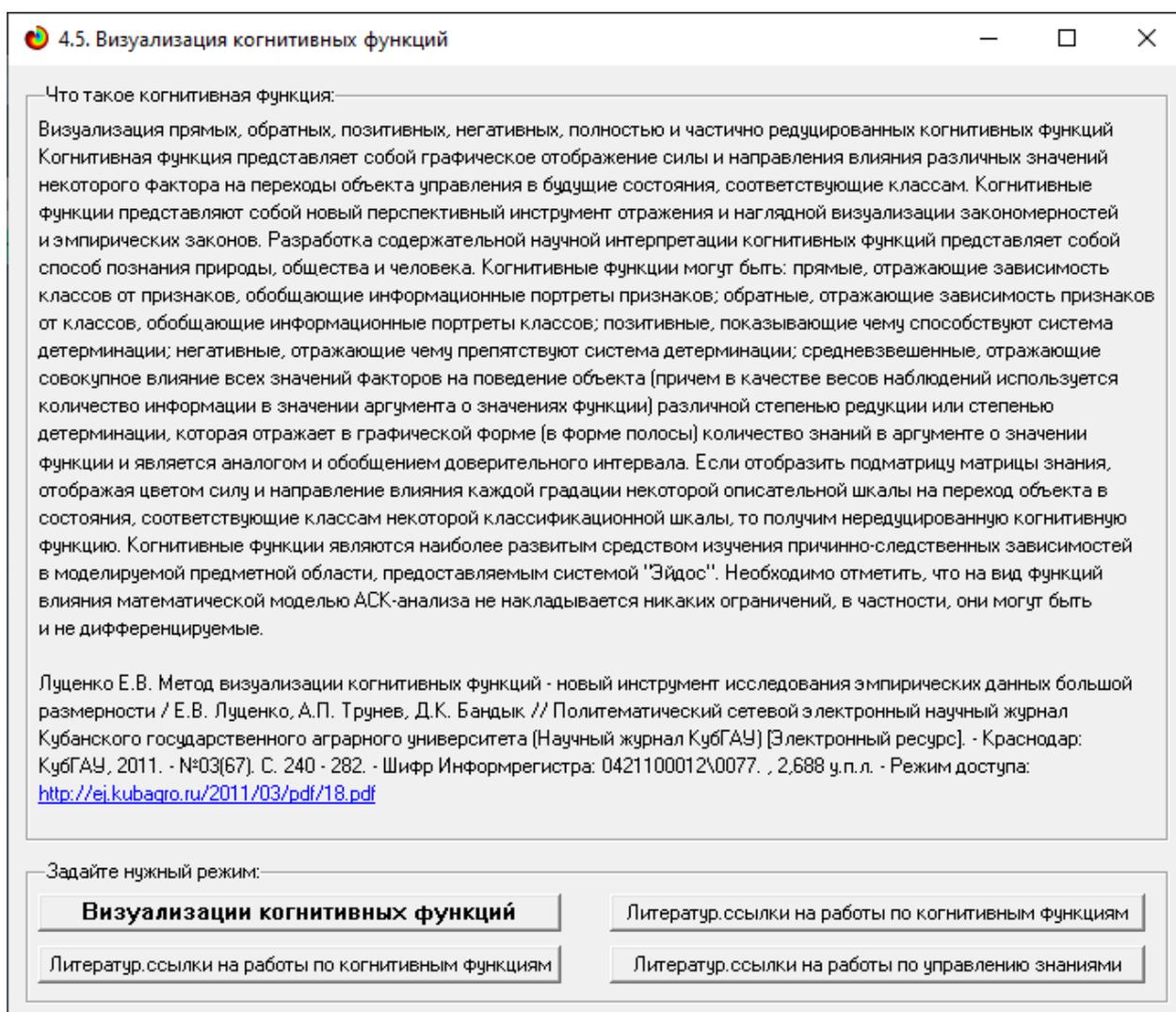
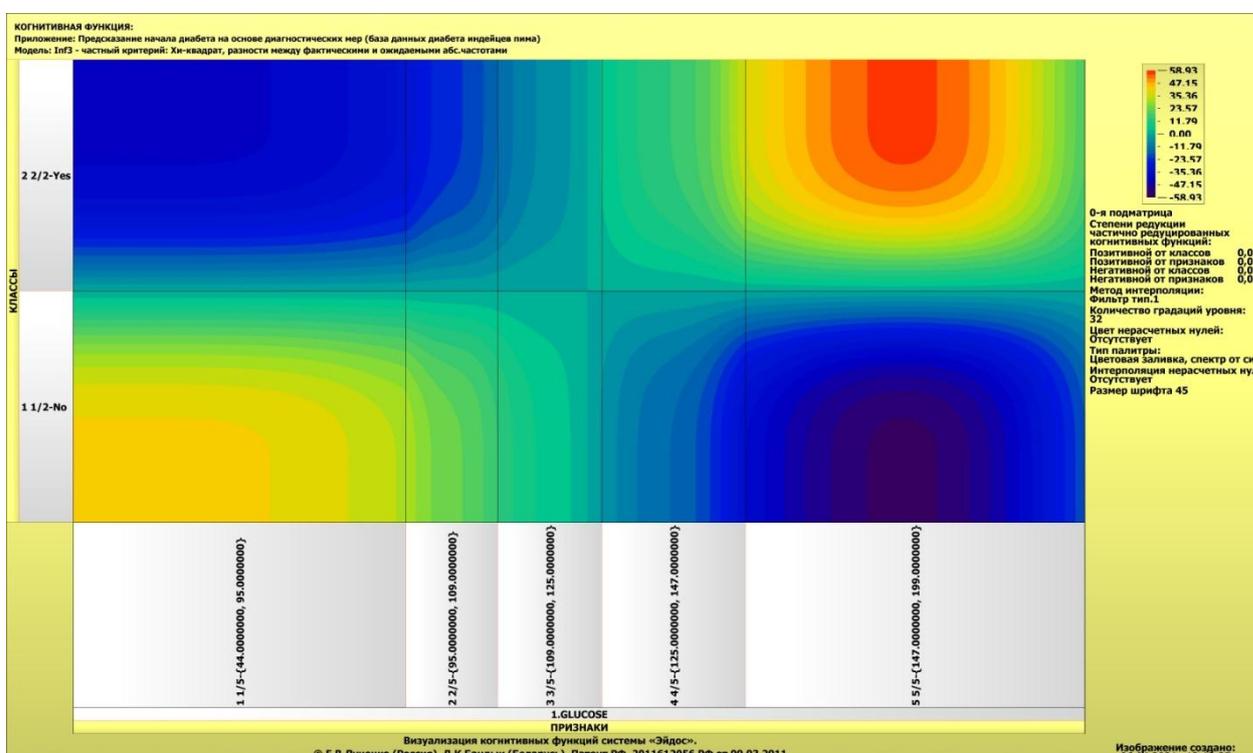
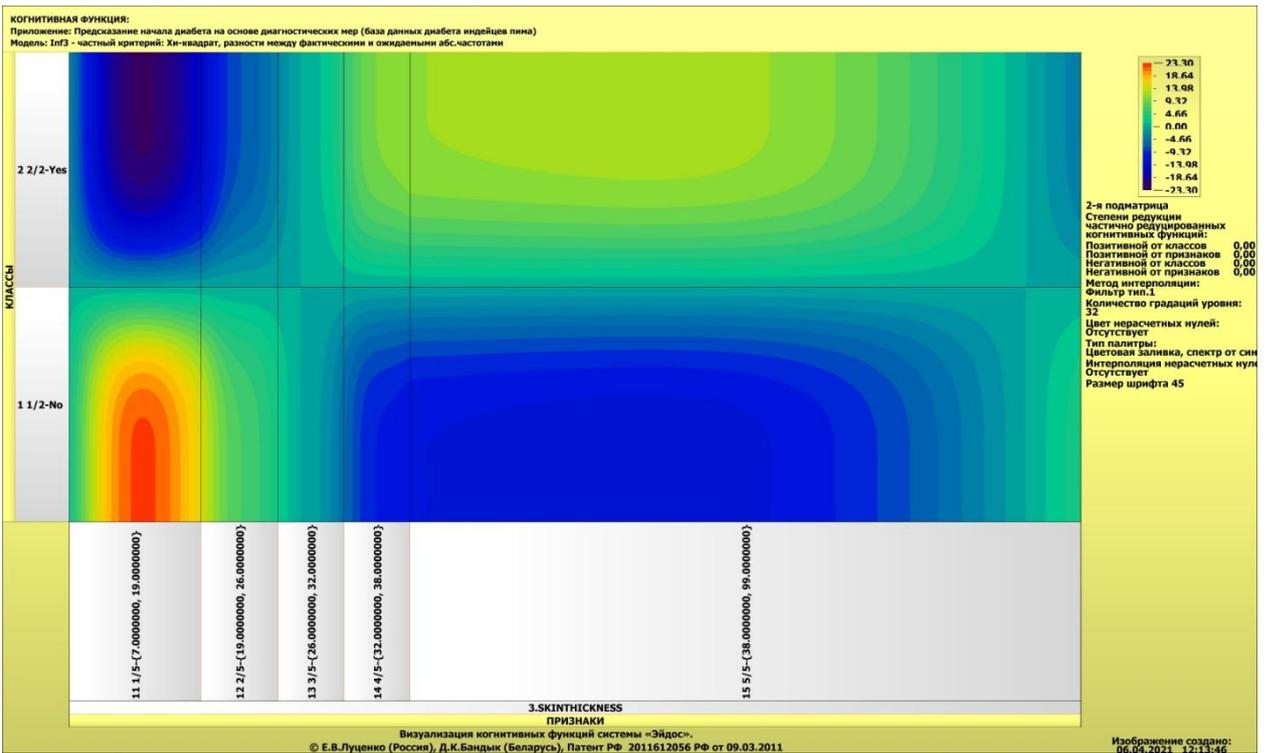
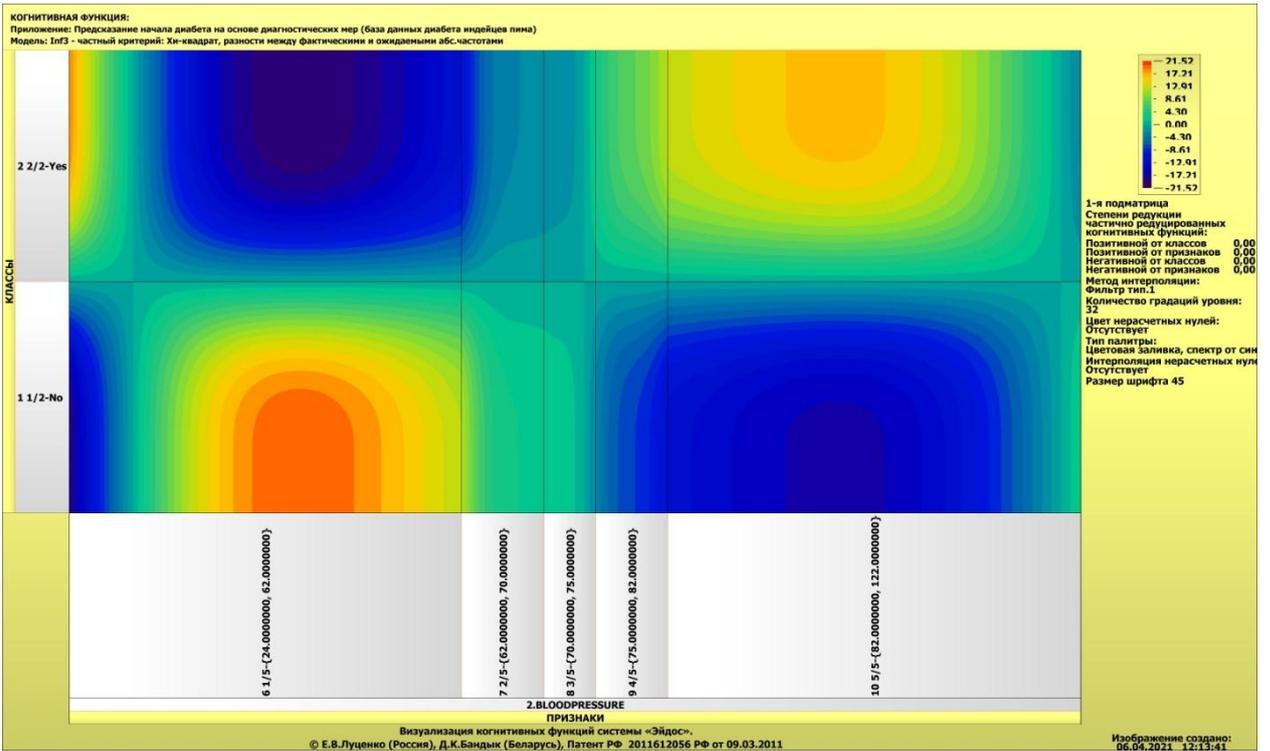
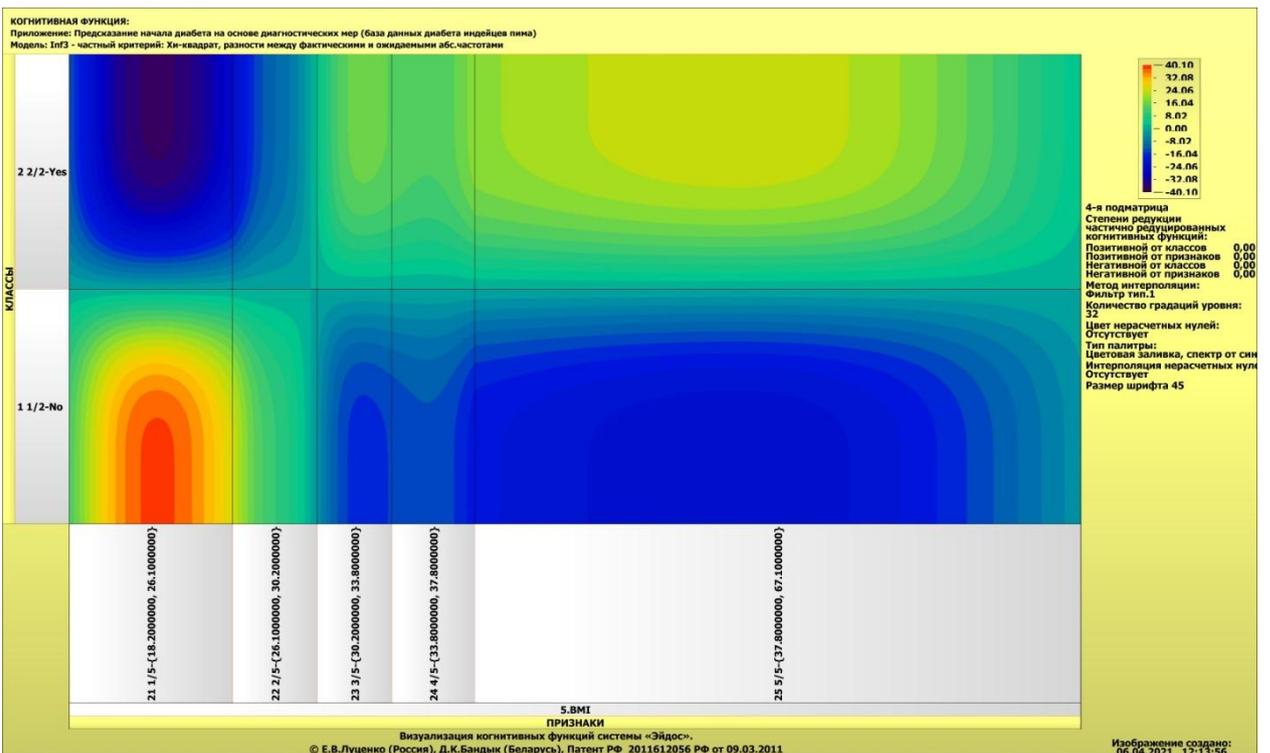
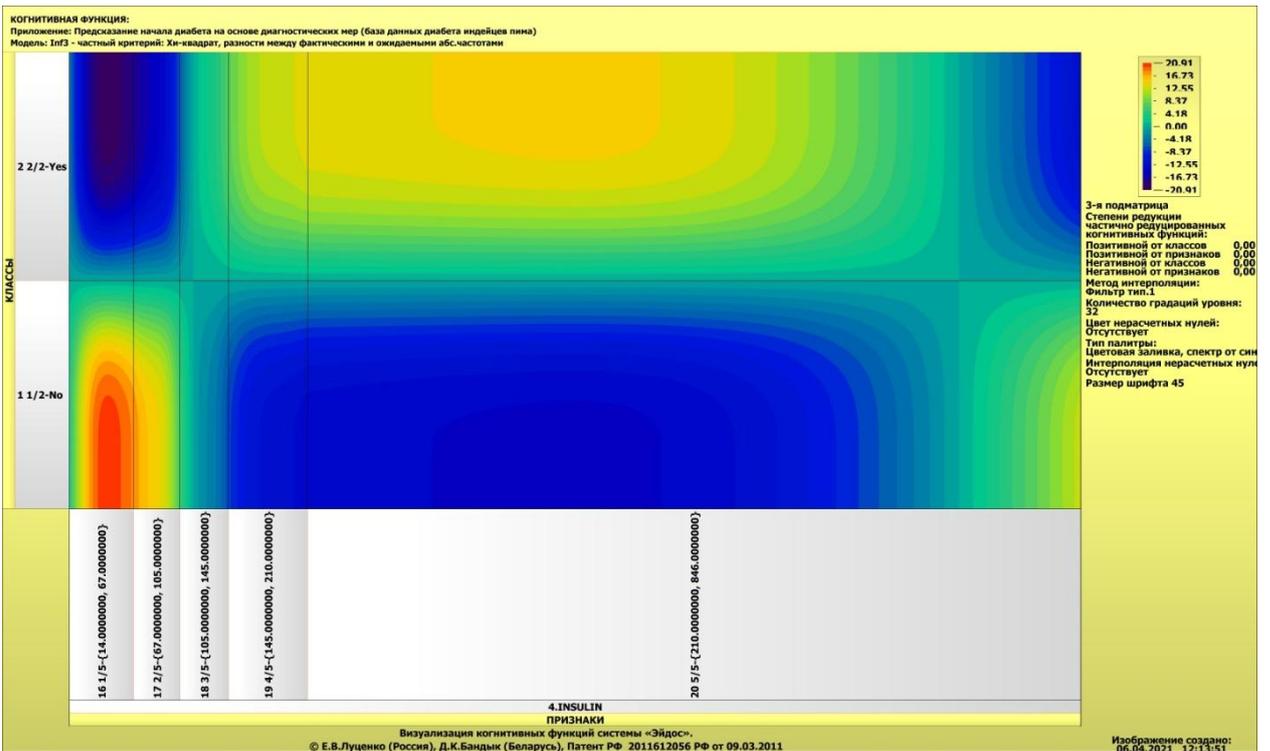


Рисунок 20 — Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»

Применительно к задаче, рассматриваемой в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния. Поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 21 приведены визуализации всех когнитивных функций данного приложения для модели INF3.







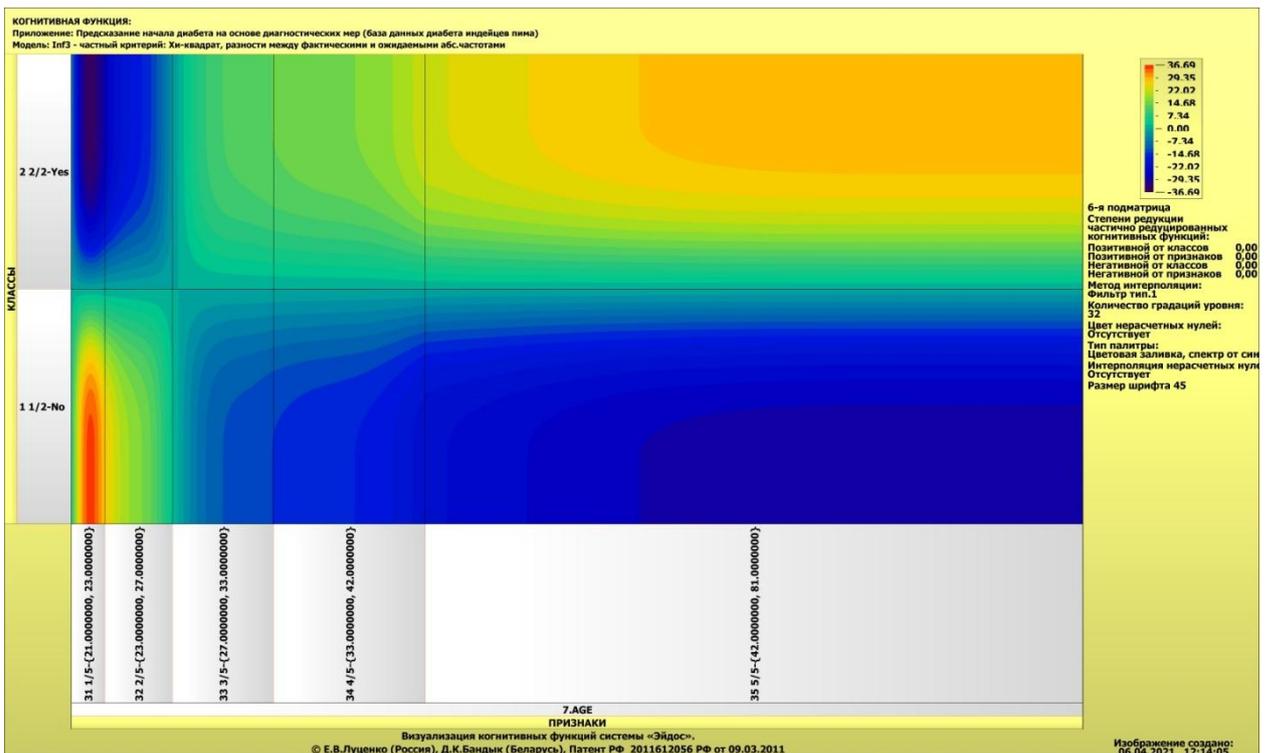
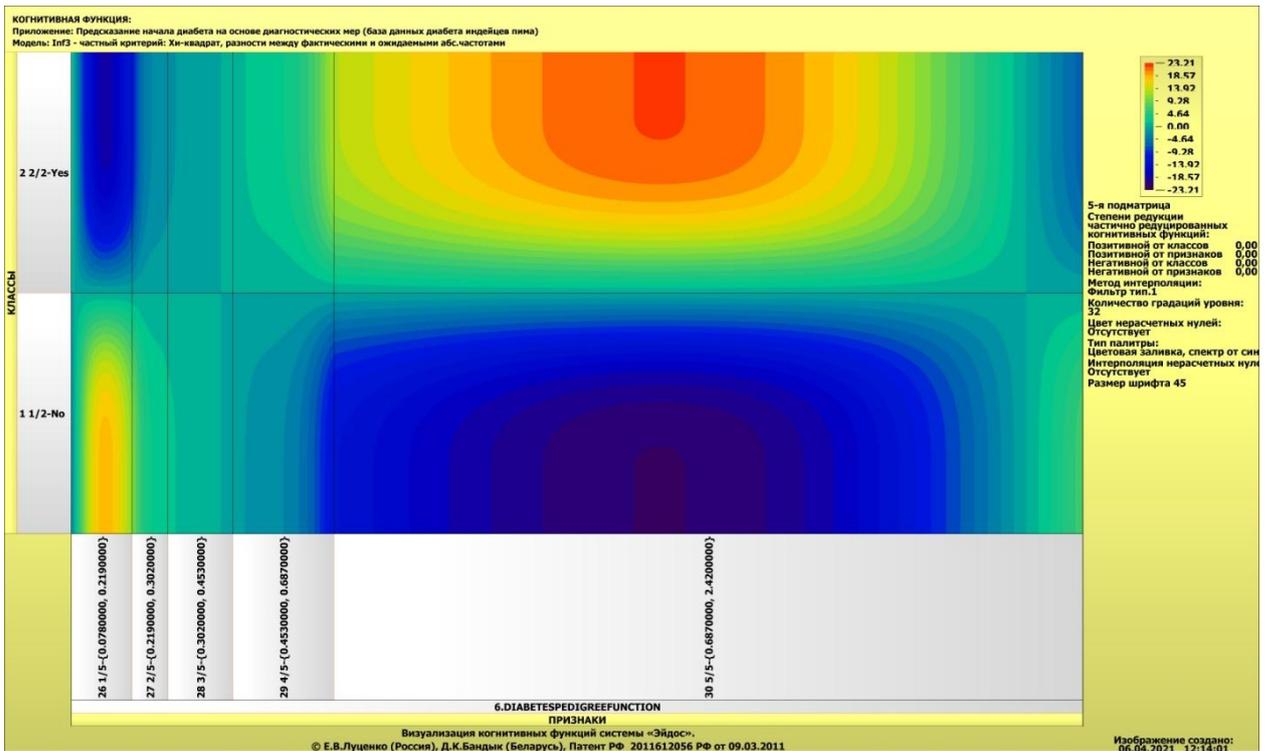


Рисунок 21 — Визуализация когнитивных функций для обобщенных классов и описательных шкал в модели INF4

2.3. SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов.

Ясно, что эксперты это делают не формализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей уже около 30 лет, но она малоизвестна — это интеллектуальная система «Эйдос».

Данная система всегда обеспечивала возможность проведения количественного автоматизированного SWOT-анализа без использования экспертных оценок непосредственно на основе эмпирических данных. Результаты SWOT-анализа выводились в форме информационных портретов. В версии системы под MS Windows: «Эйдос-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 22).

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	OUTCOME-1/2-No	21,2216382	3094	0,0000000
2	OUTCOME-2/2-Yes	21,2216382	1630	0,0000000

SWOT-анализ класса:1 "OUTCOME-1/2-No" в модели:6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	41.482
21	BMI-1/5-(18.2000000, 26.1000000)	40.102
31	AGE-1/5-(21.0000000, 23.0000000)	36.693
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	23.300
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	23.137
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	21.517
16	INSULIN-1/5-(14.0000000, 67.0000000)	20.914
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.219...	19.137
32	AGE-2/5-(23.0000000, 27.0000000)	16.862
17	INSULIN-2/5-(67.0000000, 105.0000000)	14.984
22	BMI-2/5-(26.1000000, 30.2000000)	10.102
12	SKINTHICKNESS-2/5-(19.0000000, 26.0000000)	7.230
3	GLUCOSE-3/5-(109.0000000, 125.0000000)	6.172
7	BLOODPRESSURE-2/5-(62.0000000, 70.0000000)	2.798

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	-58.933
35	AGE-5/5-(42.0000000, 81.0000000)	-28.863
25	BMI-5/5-(37.8000000, 67.1000000)	-25.553
30	DIABETESPEDIGREEFUNCTION-5/5-(0.6870000, 2.420...	-23.208
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	-17.419
23	BMI-3/5-(30.2000000, 33.8000000)	-16.518
34	AGE-4/5-(33.0000000, 42.0000000)	-15.693
20	INSULIN-5/5-(210.0000000, 846.0000000)	-15.122
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	-14.588
24	BMI-4/5-(33.8000000, 37.8000000)	-12.933
15	SKINTHICKNESS-5/5-(38.0000000, 99.0000000)	-12.735
19	INSULIN-4/5-(145.0000000, 210.0000000)	-12.361
33	AGE-3/5-(27.0000000, 33.0000000)	-12.003
14	SKINTHICKNESS-4/5-(32.0000000, 38.0000000)	-10.396

ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма

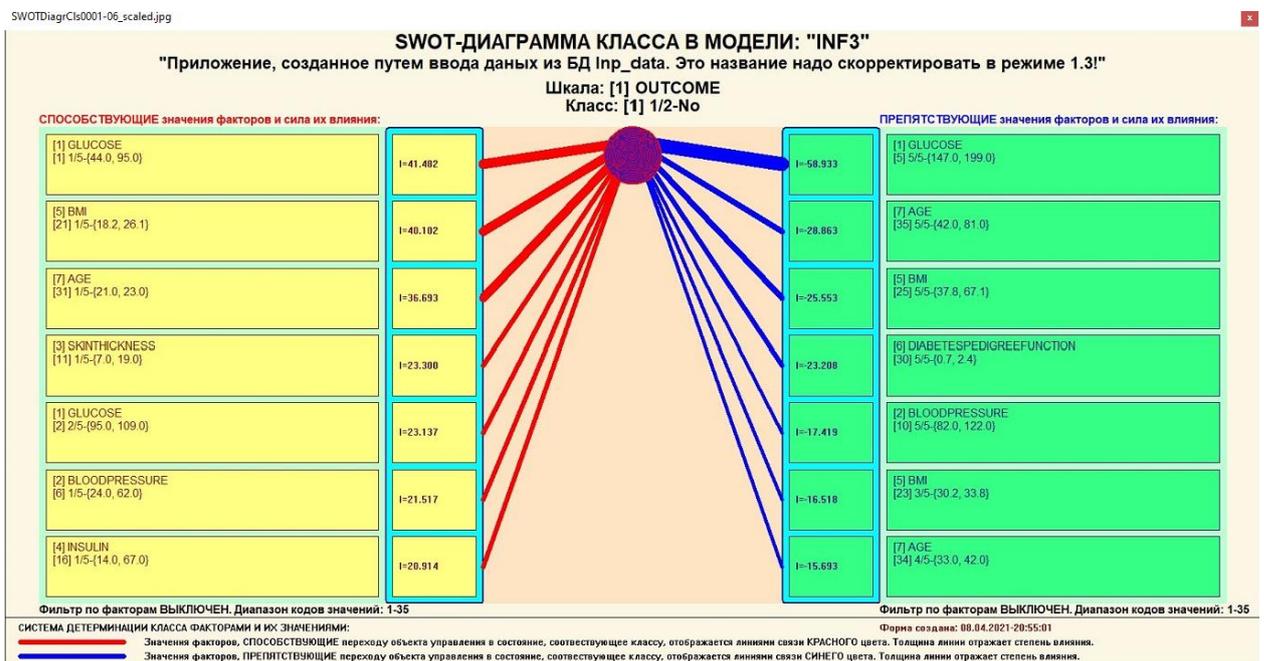


Рисунок 22 — Пример SWOT-матрицы в модели INF3

На рисунке 23 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF3.

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами АСК-анализа в системе "Эйдос"

Выбор значения фактора, оказывающего влияние на переход объекта управления в будущие состояния

Код	Наименование значения фактора
1	GLUCOSE-1/5-{44.0000000, 95.0000000}
2	GLUCOSE-2/5-{95.0000000, 109.0000000}
3	GLUCOSE-3/5-{109.0000000, 125.0000000}
4	GLUCOSE-4/5-{125.0000000, 147.0000000}
5	GLUCOSE-5/5-{147.0000000, 199.0000000}
6	BLOODPRESSURE-1/5-{24.0000000, 62.0000000}

SWOT-анализ значения фактора:1 "GLUCOSE-1/5-{44.0000000, 95.0000000}" в модели:6 "INF3"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влияния
1	OUTCOME-1/2-№	41.482

ПРЕПЯТСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
2	OUTCOME-2/2-Yes	-41.482

ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале ВКЛЮЧИТЬ фильтр по кл.шкале ВЫКЛЮЧИТЬ фильтр по кл.шкале

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма

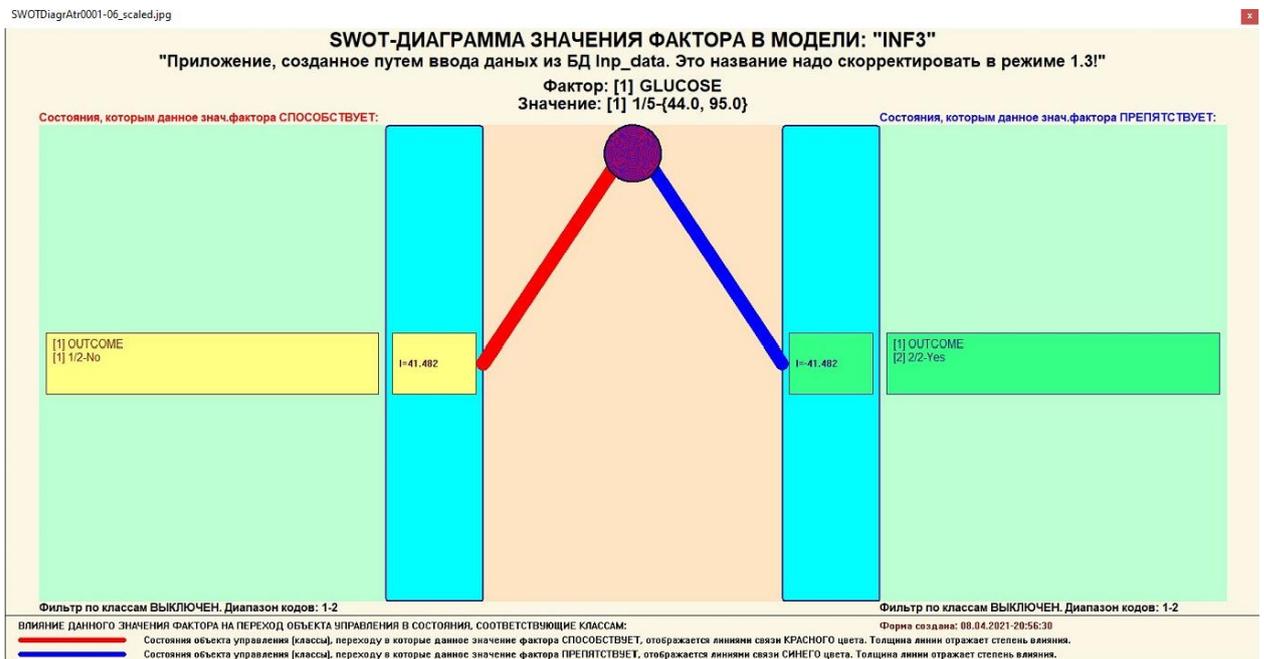


Рисунок 23 — Пример инвертированной SWOT-диаграммы в модели INF3

2.4. Кластерно-конструктивный анализ признаков

В режиме 4.3.2.2, после расчета матриц сходства, кластеров и конструкторов, строим 2D сеть классов в выбранной модели знаний, для наглядного представления сети классов.

4.3.2.2. Результаты кластерно-конструктивного анализа признаков

Конструктор признака:1 "GLUCOSE-1/5-{44.0, 95.0}" в модели:4 "INF1"

Код	Наименование признака	№	Код призна...	Наименование признака	Сходство
1	GLUCOSE-1/5-{44.0000000, 95.0000000}	1	1	GLUCOSE-1/5-{44.0000000, 95.0000000}	100.000
2	GLUCOSE-2/5-{95.0000000, 109.0000000}	2	16	INSULIN-1/5-{14.0000000, 67.0000000}	100.000
3	GLUCOSE-3/5-{109.0000000, 125.0000000}	3	21	BMI-1/5-{18.2000000, 26.1000000}	100.000
4	GLUCOSE-4/5-{125.0000000, 147.0000000}	4	31	AGE-1/5-{21.0000000, 23.0000000}	99.826
5	GLUCOSE-5/5-{147.0000000, 199.0000000}	5	11	SKINTHICKNESS-1/5-{7.0000000, 19.0000000}	99.814
6	BLOODPRESSURE-1/5-{24.0000000, 62.0000000}	6	17	INSULIN-2/5-{67.0000000, 105.0000000}	99.583
7	BLOODPRESSURE-2/5-{62.0000000, 70.0000000}	7	2	GLUCOSE-2/5-{95.0000000, 109.0000000}	99.312
8	BLOODPRESSURE-3/5-{70.0000000, 75.0000000}	8	6	BLOODPRESSURE-1/5-{24.0000000, 62.0000000}	99.155
9	BLOODPRESSURE-4/5-{75.0000000, 82.0000000}	9	26	DIABETESPEDIGREEFUNCTION-1/5-{0.0780000, 0.2190000}	99.010
10	BLOODPRESSURE-5/5-{82.0000000, 122.0000000}	10	32	AGE-2/5-{23.0000000, 27.0000000}	98.770
11	SKINTHICKNESS-1/5-{7.0000000, 19.0000000}	11	12	SKINTHICKNESS-2/5-{19.0000000, 26.0000000}	98.206
12	SKINTHICKNESS-2/5-{19.0000000, 26.0000000}	12	22	BMI-2/5-{26.1000000, 30.2000000}	98.173
13	SKINTHICKNESS-3/5-{26.0000000, 32.0000000}	13	3	GLUCOSE-3/5-{109.0000000, 125.0000000}	97.688
14	SKINTHICKNESS-4/5-{32.0000000, 38.0000000}	14	27	DIABETESPEDIGREEFUNCTION-2/5-{0.2190000, 0.3020000}	97.280
15	SKINTHICKNESS-5/5-{38.0000000, 99.0000000}	15	8	BLOODPRESSURE-3/5-{70.0000000, 75.0000000}	97.277
16	INSULIN-1/5-{14.0000000, 67.0000000}	16	7	BLOODPRESSURE-2/5-{62.0000000, 70.0000000}	97.222
17	INSULIN-2/5-{67.0000000, 105.0000000}	17	28	DIABETESPEDIGREEFUNCTION-3/5-{0.3020000, 0.4530000}	97.010
18	INSULIN-3/5-{105.0000000, 145.0000000}	18	5	GLUCOSE-5/5-{147.0000000, 199.0000000}	-80.967
19	INSULIN-4/5-{145.0000000, 210.0000000}	19	20	INSULIN-5/5-{210.0000000, 846.0000000}	-90.927
20	INSULIN-5/5-{210.0000000, 846.0000000}	20	35	AGE-5/5-{42.0000000, 81.0000000}	-91.466
21	BMI-1/5-{18.2000000, 26.1000000}	21	25	BMI-5/5-{37.8000000, 67.1000000}	-92.168
22	BMI-2/5-{26.1000000, 30.2000000}	22	30	DIABETESPEDIGREEFUNCTION-5/5-{0.6870000, 2.4200000}	-92.737
23	BMI-3/5-{30.2000000, 33.8000000}	23	19	INSULIN-4/5-{145.0000000, 210.0000000}	-92.830
24	BMI-4/5-{33.8000000, 37.8000000}	24	14	SKINTHICKNESS-4/5-{32.0000000, 38.0000000}	-93.452
25	BMI-5/5-{37.8000000, 67.1000000}	25	10	BLOODPRESSURE-5/5-{82.0000000, 122.0000000}	-93.482

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ. фильтр по кл.шкале ВЫКЛ. фильтр по кл.шкале Параметры Показать ВСЕ

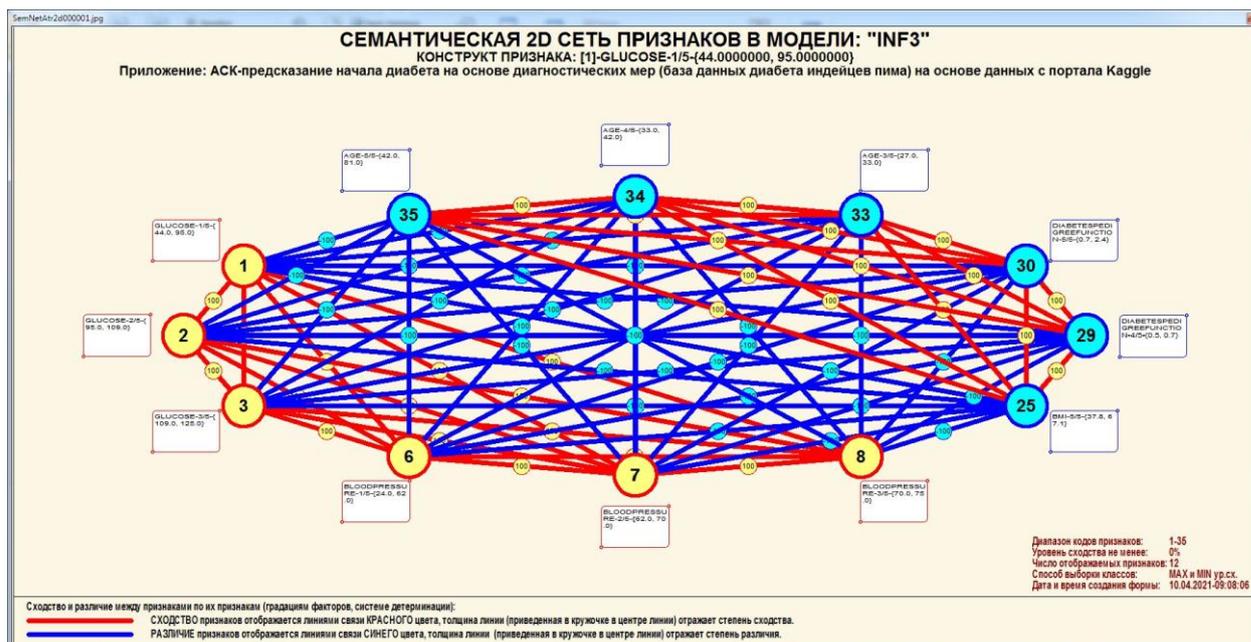


Рисунок 24 — Результаты кластерно-конструктивного анализа признаков

2.5. Нелокальные нейроны и нейронные сети

Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует не локальную нейронную сеть. Нелокальные нейроны и нелокальные нейронные сети представлены на рисунке 25:

4.4.10.Графическое отображение нелокальных нейронов в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	OUTCOME-1/2-No
2	OUTCOME-2/2-Yes

Подготовка визуализации нейрона:1 "OUTCOME-1/2-No" в модели:6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	41.482
21	BMI-1/5-(18.2000000, 26.1000000)	40.102
31	AGE-1/5-(21.0000000, 23.0000000)	36.693
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	23.300
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	23.137
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	21.517
16	INSULIN-1/5-(14.0000000, 67.0000000)	20.914
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.2190000)	19.137
32	AGE-2/5-(23.0000000, 27.0000000)	16.862
17	INSULIN-2/5-(67.0000000, 106.0000000)	14.000

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	-58.933
35	AGE-5/5-(42.0000000, 81.0000000)	-28.863
25	BMI-5/5-(37.8000000, 67.1000000)	-25.553
30	DIABETESPEDIGREEFUNCTION-5/5-(0.6870000, 2.4200000)	-23.208
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	-17.419
23	BMI-3/5-(30.2000000, 33.8000000)	-16.518
34	AGE-4/5-(33.0000000, 42.0000000)	-15.693
20	INSULIN-5/5-(210.0000000, 846.0000000)	-15.122
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	-14.588
34	BMI-4/5-(33.0000000, 37.0000000)	-13.000

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

НЕЙРОН Максимальное количество отображаемых рецепторов: 999 Минимальный вес коэф.ф.отображаемых рецепторов: 0,000

Сортировать рецепторы: по информативности по модулю информативности с наименованиями только с кодами

Neuron0001Inf3_scaled.jpg



4.4.11. Отображение Парето-подмножеств нелокальной нейронной сети в системе "Эйдос"

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	OUTCOME-1/2-No
<input checked="" type="checkbox"/>	2	OUTCOME-2/2-Yes

Помощь Максимальное количество отображаемых нейронов: ClearSet Диапазон кодов отображаемых нейронов: -
 Максимальное количество отображаемых связей: Диапазон кодов отображаемых рецепторов: -

Подготовка визуализации нейрона: 1 "OUTCOME-1/2-No" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния **ТОРМОЗЯЩИЕ рецепторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
1	GLUCOSE-1/5-(44.0000000, 95.0000000)	41.482
21	BMI-1/5-(18.2000000, 26.1000000)	40.102
31	AGE-1/5-(21.0000000, 23.0000000)	36.693
11	SKINTHICKNESS-1/5-(7.0000000, 19.0000000)	23.300
2	GLUCOSE-2/5-(95.0000000, 109.0000000)	23.137
6	BLOODPRESSURE-1/5-(24.0000000, 62.0000000)	21.517
16	INSULIN-1/5-(14.0000000, 67.0000000)	20.914
26	DIABETESPEDIGREEFUNCTION-1/5-(0.0780000, 0.2190000)	19.137
32	AGE-2/5-(23.0000000, 27.0000000)	16.862
17	INSULIN-2/5-(67.0000000, 105.0000000)	14.000

Код	Наименование фактора и его интервального значения	Сила влияния
5	GLUCOSE-5/5-(147.0000000, 199.0000000)	-58.933
35	AGE-5/5-(42.0000000, 81.0000000)	-28.863
25	BMI-5/5-(37.8000000, 67.1000000)	-25.553
30	DIABETESPEDIGREEFUNCTION-5/5-(0.6870000, 2.4200000)	-23.208
10	BLOODPRESSURE-5/5-(82.0000000, 122.0000000)	-17.419
23	BMI-3/5-(30.2000000, 33.8000000)	-16.518
34	AGE-4/5-(33.0000000, 42.0000000)	-15.693
20	INSULIN-5/5-(210.0000000, 846.0000000)	-15.122
4	GLUCOSE-4/5-(125.0000000, 147.0000000)	-14.588
24	BMI-4/5-(33.0000000, 37.0000000)	-13.000

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

НейроСеть Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Максимальное количество отображаемых рецепторов: Отображать связи с интенсивностью >= % от макс.:

Сортировать связи: по модулю информативности по информативности и знаку Отображать наименования: нейронов рецепторов

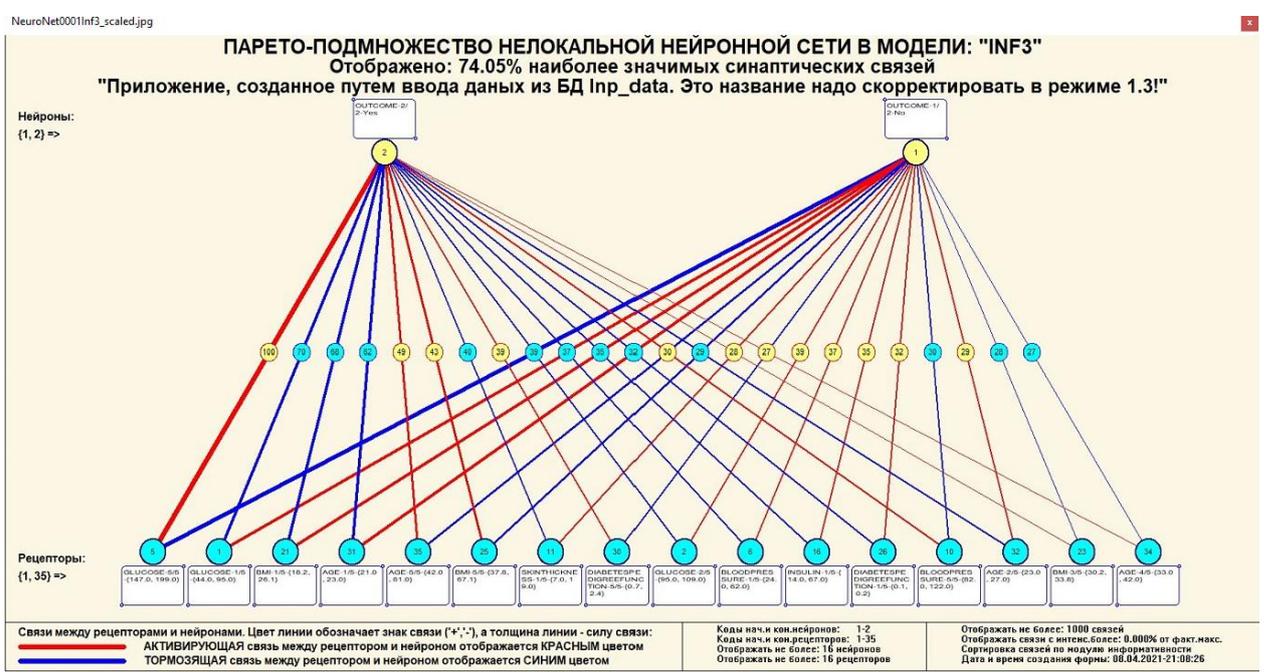


Рисунок 25 — Нелокальные нейроны и нелокальные нейронные сети
 Благодаря данному нейрону видно, как различные факторы влияют на модель, какие оказывают положительное влияние, а какие отрицательное.

ЗАКЛЮЧЕНИЕ

Так как существует множество систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную электронную базу Kaggle.

Сверхзадачей искусственного интеллекта является построение компьютерной интеллектуальной системы, которая обладала бы уровнем эффективности решений неформализованных задач, сравнимым с человеческим или превосходящим его.

Самую существенную часть систем искусственного интеллекта составляют экспертные системы. Экспертная система обычно определяется как программа ЭВМ, моделирующая действия эксперта человека при решении задач в узкой предметной области: составление базы знаний и накопления их.

В данной курсовой работе был продемонстрировано построение модели зависимости уровня мирового счастья от различных характеристик стран системой искусственного интеллекта «Aidos-X++» с использованием общедоступной базы данных «Pima Indians Diabetes Database». При этом наиболее достоверной в данном приложении оказалась модель INF3. Точность модели по F-мере Ван Ризбергена составляет 0.744.

АСК-анализ позволяет:

- формировать обобщенные лингвистические образы классов (семантические ядра) на основе фрагментов или примеров, относящихся к ним текстов на любом языке;
- количественно сравнивать лингвистический образ конкретного человека, или описание объекта, процесса с обобщенными лингвистическими образами групп (классов);

- сравнивать обобщенные лингвистические образы классов друг с другом и создавать их кластеры и конструкторы;
- исследовать моделируемую предметную область путем исследования ее лингвистической системно-когнитивной модели;
- проводить интеллектуальную атрибуцию текстов, т.е. определять вероятное авторство анонимных и псевдонимных текстов, датировку, жанр и смысловую направленность содержания текстов;
- все это можно делать для любого естественного или искусственного языка, или системы кодирования.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Луценко Е.В. 30 лет системе «Эйдос» — одной из старейших отечественных универсальных систем искусственного интеллекта, широко применяемых и развивающихся и в настоящее время / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2009. — №10(054). С. 48 — 77. — Шифр Информрегистра: 0420900012\0110, IDA [article ID]: 0540910004. — Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/04.pdf>, 1,875 у.п.л.

2. Луценко Е.В. Методика использования репозитория UCI для оценки качества математических моделей систем искусственного интеллекта / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: http://lc.kubagro.ru/My_training_schedule.doc КубГАУ, 2003. — №02(002). С. 120 — 145. — IDA [article ID]: 0020302012. — Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 у.п.л.

3. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). — Краснодар: КубГАУ. 2002. — 605 с.

4. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: Куб- ГАУ, 2013. —

№04(088). С. 340 — 359. — IDA [article ID]: 0881304022. — Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.

5. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://ic.kubagro.ru/>, свободный. - Загл. с экрана. Яз. рус.

6. Луценко Е.В. АСК-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: Куб- ГАУ, 2014. — №06(100). С. 1346 — 1395. — IDA [article ID]: 1001406090. — Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/90.pdf>, 3,125 у.п.л.