

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

РЕЦЕНЗИЯ

на курсовой проект (работу)

Студента (ки) Лысенко Валерий Владимирович
курса 2 заочной формы обучения
направления подготовки Информационные системы и технологии
направленность(профиль)Информационные системы и технологии в науке и
управлении

Наименование темы «Исследование преступности в Лондоне с использованием
АСК системы Эйдос на основе данных web-сервера Kaggle».

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсового проекта (работы)

№ п/п	Показатель	Оценка соответствия заданию (по 5-и балльной шкале)
1.	Актуальность тематики работы	4
2.	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
3.	Уровень и корректность использования в работе различных методов исследований	5
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
5.	Ясность, четкость, последовательность и обоснованность изложения	4
6.	Применение современных технологий обработки информации	5
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	4
8.	Ответы на вопросы при защите	5

Достоинства работы

Применения современных систем искусственного интеллекта,
математических моделей и программного инструментария для проведения
АСК-анализа на основе базы данных

Недостатки работы

Отдельные погрешности в оформлении

Итоговая оценка при защите 5 (отлично)

Рецензент _____
« _____ » _____ 2019 г.

Е.В. Луценко

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»

Заочный факультет
Кафедра компьютерных технологий и систем

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
к курсовой работе

по дисциплине: СИСТЕМЫ ПРЕДСТАВЛЕНИЯ И ПРИОБРЕТЕНИЯ
ЗНАНИЙ

на тему:

ИССЛЕДОВАНИЕ ПРЕСТУПНОСТИ В ЛОНДОНЕ С
ИСПОЛЬЗОВАНИЕМ АСК СИСТЕМЫ ЭЙДОС НА ОСНОВЕ
ДАНЫХ WEB-СЕВЕРА KAGGLE

выполнил студент группы ИТ31841 Лысенко Валерий Владимирович

Допущена к защите _____

Руководитель проекта Луценко Евгений Вениаминович, д.э.н., к.т.н.,
профессор

(подпись, расшифровка подписи)

Нормоконтролер Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор

(подпись, расшифровка подписи)

Защищена _____

Оценка _____

(дата)

Члены комиссии _____

В.И. Лойко

Е.В. Луценко

В.Н. Лаптев

(подпись, дата, расшифровка подписи)

Краснодар

2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»

Заочный факультет
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:

Зав. кафедрой КТС _____ В. И. Лойко

ЗАДАНИЕ

на курсовую работу

Студенту: ИТ31841 группы 2 курса
Факультета заочного
Специальности: 09.04.02 Информационные системы и технологии
(шифр)

Лысенко Валерию Владимировичу

(Ф.И.О.)

Тема проекта: **Исследование преступности в Лондоне с использованием АСК системы Эйдос на основе данных web-сервера Kaggle**

Содержание задания: Проанализировать современные методы и средства формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами

Объем работы:

а) пояснительная записка к работе _____ 42 _____ листа формата А4
б) графическая часть _____ 35 _____ лист формата А4

Рекомендуемая литература: Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- X++» 2017 г.

Срок выполнения проекта: с “___” _____ по “___” _____ 20__ г.

Срок защиты: “___” _____ 20__ г.

Дата выдачи задания: “___” _____ 20__ г.

Дата сдачи проекта на кафедру: “___” _____ 20__ г.

Руководитель проекта: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(подпись, Ф.И.О., звание, степень)

Задание принял студент _____
(подпись, дата)

Краснодар

2019 г.

РЕФЕРАТ

42 страницы, 31 рисунок, 4 таблицы, 8 литературных источников
СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ЭЙДОС X,
ГРАДАЦИИ, КЛАССИФИКАЦИОННЫЕ ШКАЛЫ, ОПИСАТЕЛЬНЫЕ
ШКАЛЫ

Целью работы является исследование преступности в Лондоне с использованием АСК системы Эйдос на основе данных web-сервера Kaggle.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	8
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	9
1.1. Описание решения.....	9
1.2. Преобразование исходных данных из HTML-формата в файл исходных данных MS Excel	9
1.3. Синтез и верификация статистических и интеллектуальных моделей .	15
1.4. Виды моделей системы «Эйдос».....	17
1.5. Результаты верификации моделей	19
2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	27
2.1. Решение задачи	27
2.2. Когнитивные функции	30
2.3 Кластерно-конструктивного анализа	34
2.5 SWOT и PEST матрицы и диаграммы	37
ЗАКЛЮЧЕНИЕ	40
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	41

ВВЕДЕНИЕ

Системы искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. В данной работе рассмотрен анализ массовых расстрелов в Америке.

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 база данных тестовых задач для систем искусственного интеллекта электронного ресурса баз данных Kaggle.

Для решения задачи используем стандартные возможности Microsoft Office Word и Excel, блокнот, а также систему искусственного интеллекта "Эйдос- X++".

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

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

1. Преобразование исходных данных из HTML-формата в промежуточные файлы MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос".
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

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

Из электронного ресурса баз данных Kaggle возьмем базу данных «London Crime Data».

Общее описание задачи:

1. lsoa_code - код зоны вывода супер-уровня нижнего уровня в соответствии с Управлением национальной статистики
2. borough - район
3. major_category - основная категория
4. minor_category - второстепенная категория
5. value - значение (Сводная информация о количестве преступлений за месяц)
6. year - год
7. month - месяц

Добавим столбец А и пронумеруем события согласно порядковых номеров, назовем колонку Number.

Столбцы 2-5 являются классификационной шкалой.

Столбцы 6-8 описательные шкалы.

Обучающая выборка:

Таблица 1 – London_Crime_Data.xls

A	B	C	D	E	F	G	H
Number	Isoa_code Lower Layer Super Output Area code according to the Office for National Statistics	borough	major_category	minor_category	value Summary of the number of crimes for the month	year	month
1	E01032739	City of London	Violence Against the Person	Harassment	0	2010	12
2	E01032739	City of London	Violence Against the Person	Harassment	0	2010	7
3	E01032740	City of London	Violence Against the Person	Harassment	0	2014	12
4	E01000005	City of London	Violence Against the Person	Harassment	0	2013	4
5	E01032740	City of London	Violence Against the Person	Harassment	0	2016	1
6	E01032739	City of London	Violence Against the Person	Harassment	1	2014	9
7	E01000001	City of London	Violence Against the Person	Harassment	0	2010	4
8	E01000002	City of London	Violence Against the Person	Harassment	0	2008	5
9	E01032740	City of London	Violence Against the Person	Harassment	2	2015	7
10	E01000005	City of London	Violence Against the Person	Harassment	0	2016	10
11	E01032740	City of London	Violence Against the Person	Harassment	0	2015	1
12	E01000005	City of London	Violence Against the Person	Harassment	1	2011	9

Поскольку ввод исходных данных в систему «Эйдос» планируется осуществить с помощью ее универсального программного интерфейса импорта данных из внешних баз данных, который работает с файлами MS Excel, то преобразуем данные из html-файла в xls-файл, для чего выполним следующие операции.

Скопируем получившуюся таблицу из MS Word в MS Excel и запишем ее с именем: Inp_data.xls в папку: c:\Aidos-X\AID_DATA\Inp_data\. В результате получим таблицу исходных данных, полностью подготовленную для обработки в системе «Эйдос» и записанную в нужную папку в виде файла нужного типа с нужным именем.

Автоматизированная формализация предметной области путем импорта исходных данных из внешних баз данных в систему "Эйдос".

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, т.е. режимом 2.3.2.2.

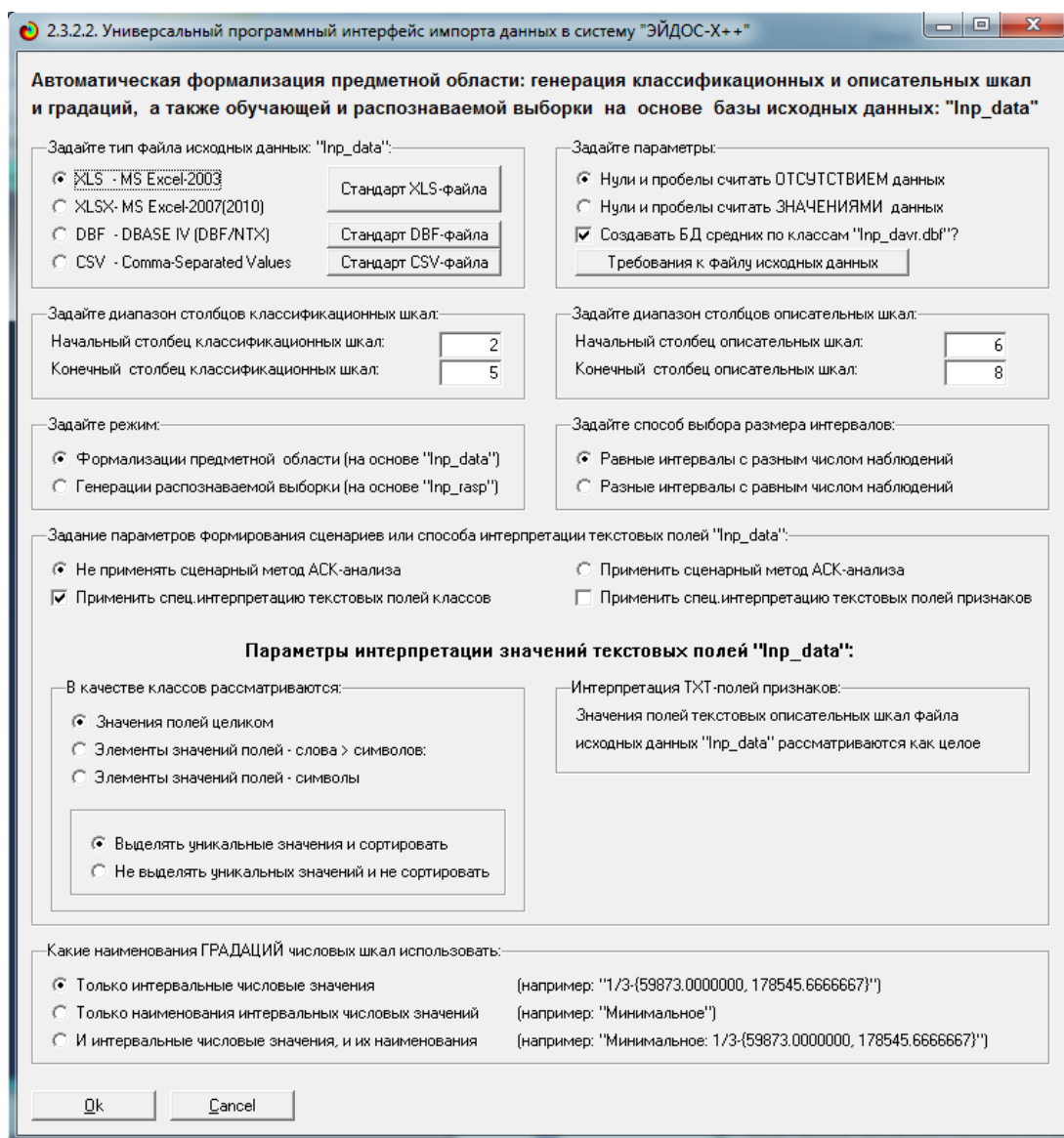


Рисунок 1 Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

В экранной форме, приведенной на рисунке 1, задать настройки, показанные на рисунке:

– "Задайте тип файла исходных данных Inp_data": "XLS – MS Excel-2007(2010)";

– "Задайте диапазон столбцов классификационных шкал": "Начальный столбец классификационных шкал" – 2, "Конечный столбец классификационных шкал" – 4(второй столбец в таблице);

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 5, "Конечный столбец описательных шкал" – 9;

– "Задание параметров формирования сценариев или способа интерпретации текстовых полей": "Не применять сценарный метод АСК-анализа и спец.интерпретацию ТХТ-полей".

После нажать кнопку "ОК". Далее открывается окно, где размещена информация о размерности модели (рисунок 2). В этом окне необходимо нажать кнопку "Выйти на создание модели".

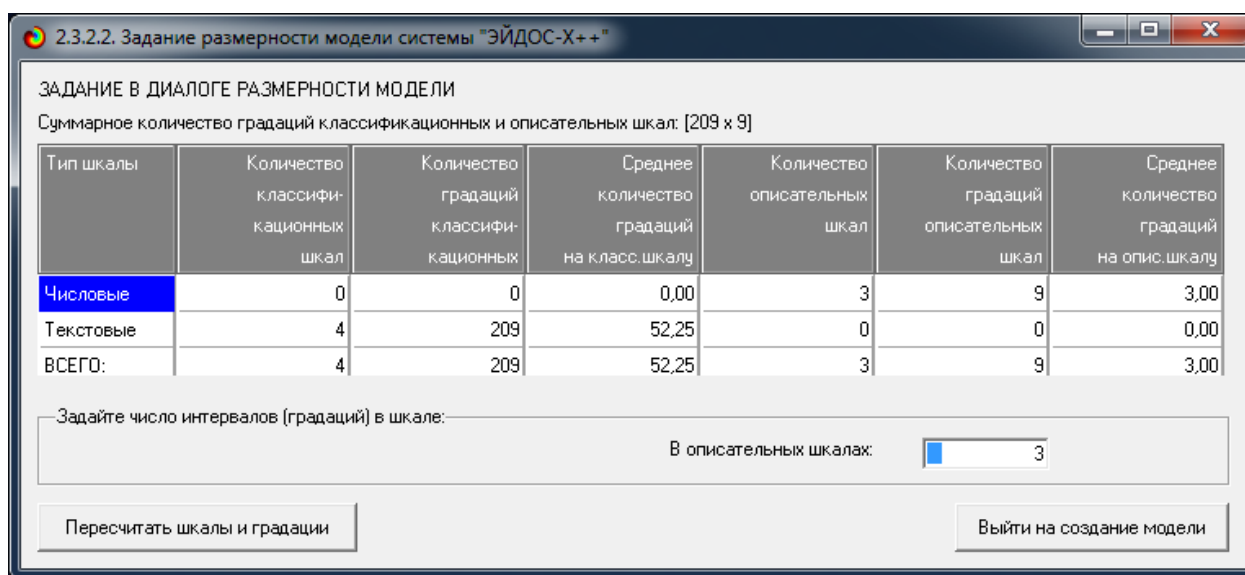


Рисунок 2. Задание размерности модели системы "Эйдос"

Далее открывается окно, отображающее стадию процесса импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос" (рисунок 3), а также прогноз времени завершения этого процесса. В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "ОК".

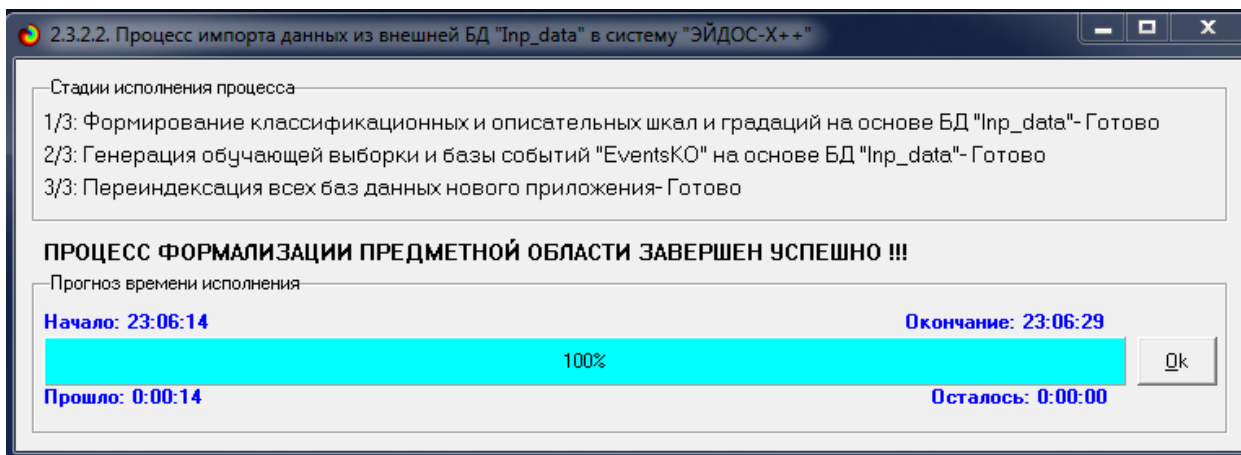


Рисунок 3. Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

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 4).

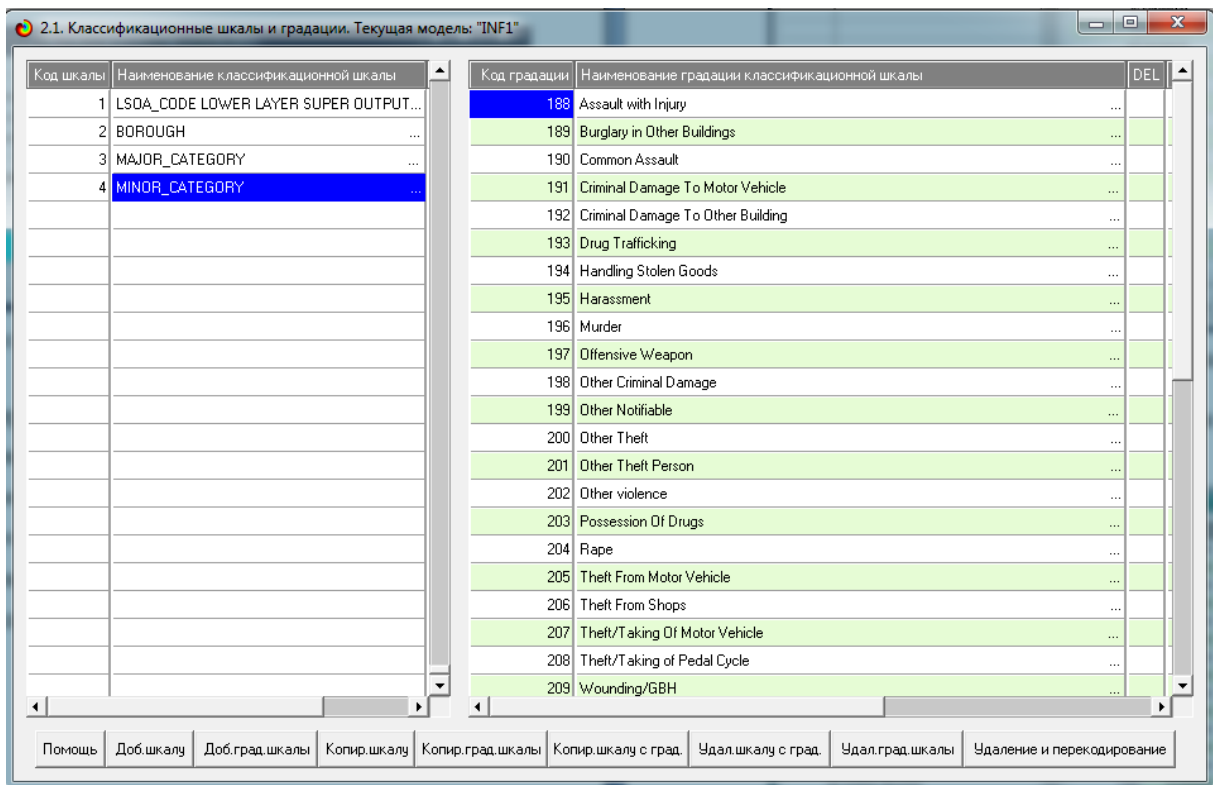


Рисунок 4. Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 5):

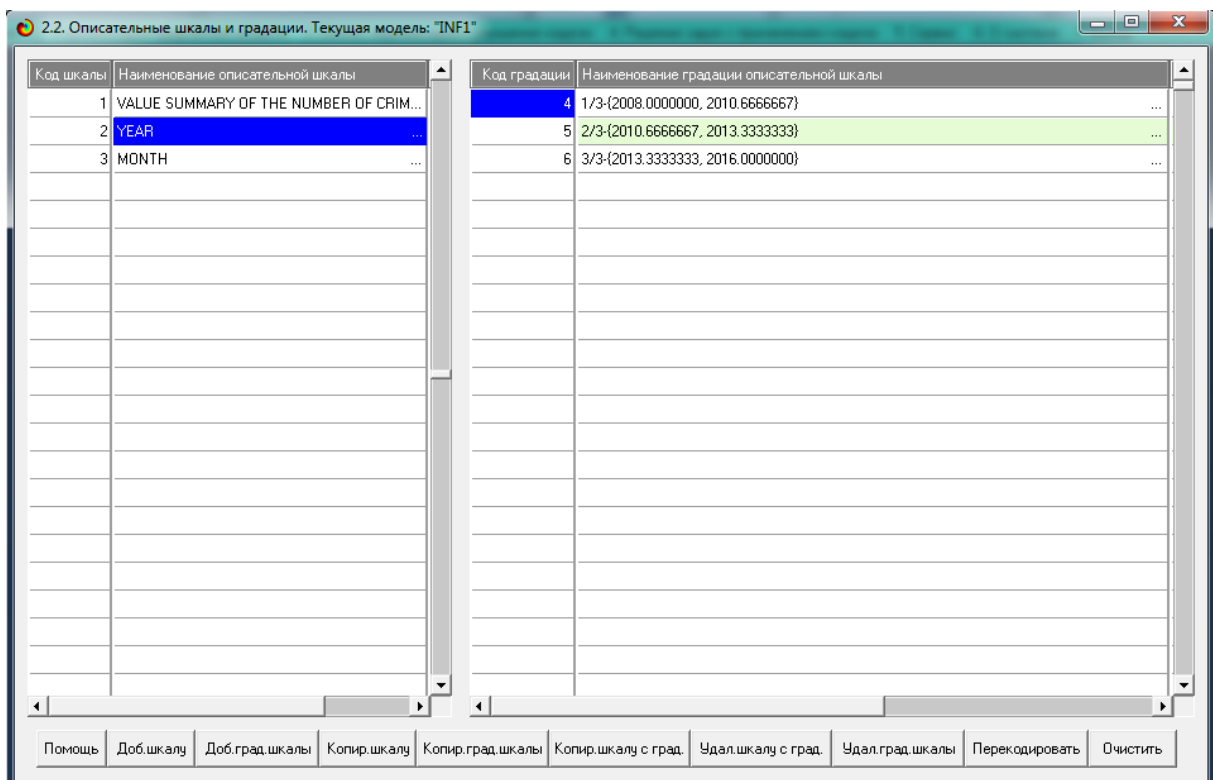


Рисунок 5. Описательные шкалы и градации (фрагмент)

2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: "TNF1"

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

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	170	180	187	195

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	4	9	0	0	0	0	0

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

Рисунок 6. Обучающая выборка (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 [8]).

1.3. 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 7).

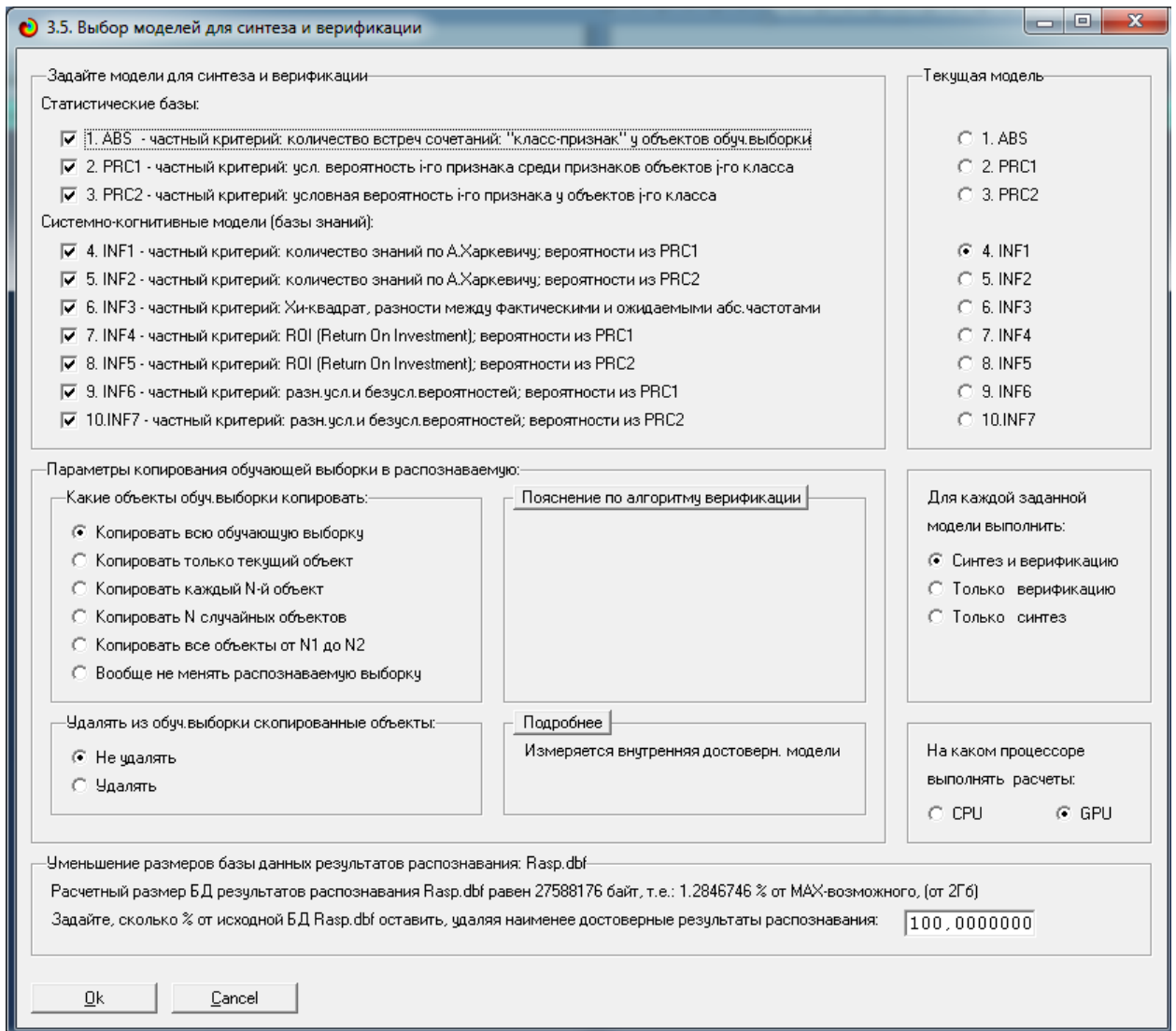


Рисунок 7. Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей. Но мы используем параметры, приведенные на рисунке 7. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

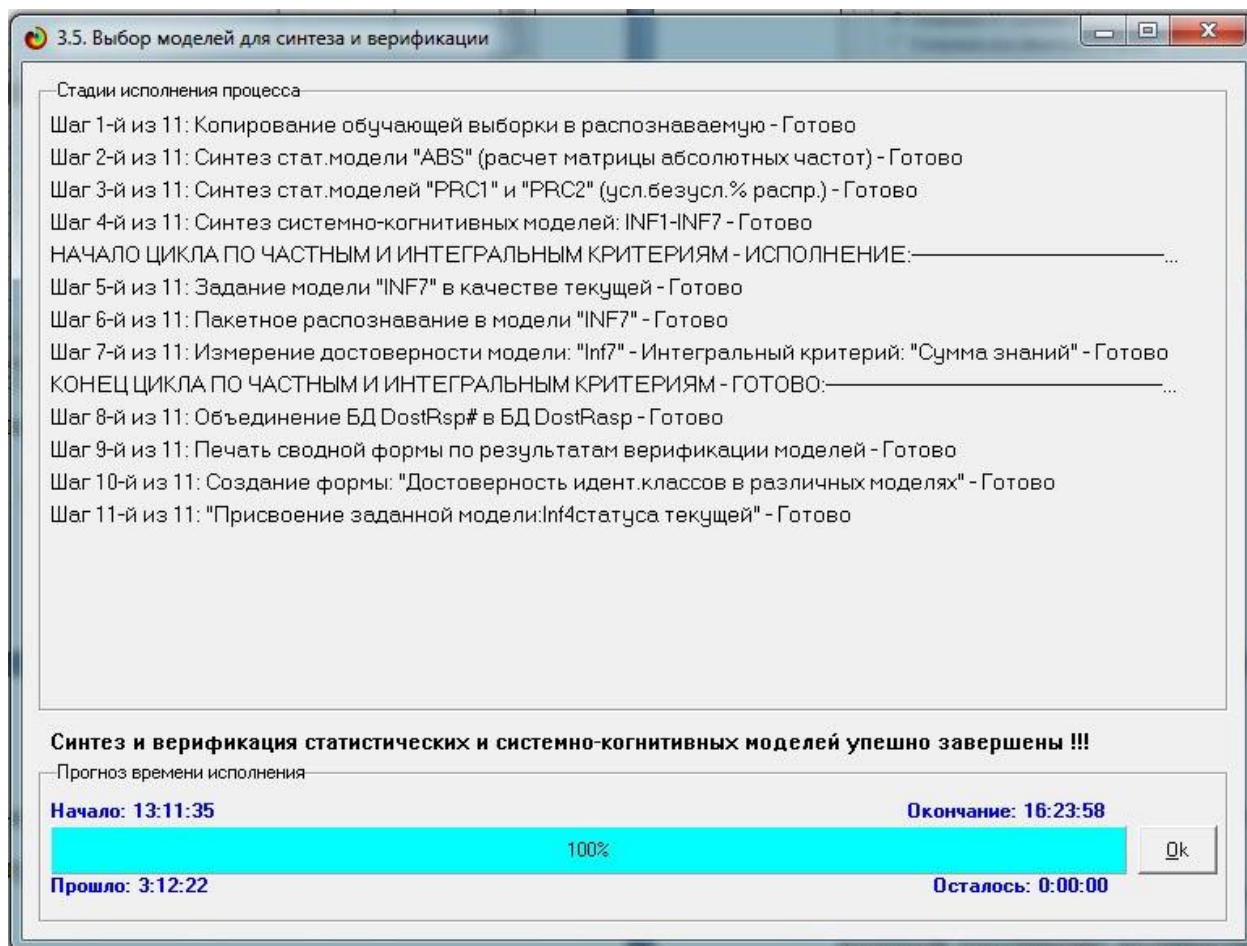


Рисунок 8. Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 3 часа 12 минут, так как обрабатывается большой объем данных. При этом верификация (оценка достоверности моделей) проводилась на 50 примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 10, но ниже мы приведем лишь некоторые из них (таблицы 2, 3, 4).

1.4. Виды моделей системы «Эйдос»

Рассмотрим решение задачи идентификации на примере модели INF1, в которой рассчитано количество информации по А. Харкевичу, которое мы

получаем о принадлежности идентифицируемого объекта к каждому из классов, если знаем, что у этого объекта есть некоторый признак.

По сути, частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот (таблица 2) в матрицы условных и безусловных процентных распределений, и матрицы знаний (таблицы 3 и 4) (проф. В.И.Лойко).

Таблица 2 – Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений (фрагменты)

Код признака	Наименование описательной шкалы и градации	1. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000001	2. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000002	3. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000003	4. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000005	5. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000471	6. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000472	7. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000473	8. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000474	9. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000475	10. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000476	11. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000477
1	VALUE SUMMARY OF THE NUMBER OF CRIMES FO...	2	5	3	14	2	5	5	3	4	4	7
2	VALUE SUMMARY OF THE NUMBER OF CRIMES FO...											
3	VALUE SUMMARY OF THE NUMBER OF CRIMES FO...											
4	YEAR-1/3-(2008.0000000, 2010.6666667) ...	59	40	29	78	8	11	8	2		3	3
5	YEAR-2/3-(2010.6666667, 2013.3333333) ...	69	51	40	84	10	9	5	7	4	2	5
6	YEAR-3/3-(2013.3333333, 2016.0000000) ...	66	54	23	77	5	7	10	2	4	2	3
7	MONTH-1/3-(1.0000000, 4.6666667) ...	64	47	25	73	8	6	7	5	1	3	3
8	MONTH-2/3-(4.6666667, 8.3333333) ...	59	43	40	83	7	13	10	3		1	4
9	MONTH-3/3-(8.3333333, 12.0000000) ...	71	55	27	83	8	8	6	3	7	3	4
	Сумма числа признаков	390	295	187	492	48	59	51	25	20	18	29
	Среднее	43	33	21	55	5	7	6	3	2	2	3
	Среднеквадратичное отклонение	32	24	16	38	4	4	4	2	3	1	2
	Сумма числа объектов обуч.выборки	194	145	92	239	23	27	23	11	8	7	11

Таблица 3 – Модель INF4 (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000001	2. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000002	3. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000003	4. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000005	5. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000471	6. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000472	7. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000473	8. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000474	9. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000475	10. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000476	11. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000477
1	VALUE SUMMARY OF THE NUMBER OF CRIME...	-0.940	-0.802	-0.813	-0.668	-0.513	-0.010	0.146	0.402	1.337	1.597	
2	VALUE SUMMARY OF THE NUMBER OF CRIME...											
3	VALUE SUMMARY OF THE NUMBER OF CRIME...											
4	YEAR-1/3-(2008.000000, 2010.666667) ...	0.008	-0.097	0.033	0.056	0.111	0.242	0.045	-0.467			0.111
5	YEAR-2/3-(2010.666667, 2013.333333) ...	0.166	0.139	0.410	0.125	0.373	0.005	-0.354	0.845	0.318	-0.268	
6	YEAR-3/3-(2013.333333, 2016.000000) ...	0.103	0.193	-0.198	0.020	-0.321	-0.227	0.278	-0.479	0.304	-0.276	
7	MONTH-1/3-(1.000000, 4.666667) ...	0.085	0.053	-0.116	-0.019	0.102	-0.328	-0.093	0.322	-0.670	0.102	
8	MONTH-2/3-(4.666667, 8.333333) ...	-0.001	-0.037	0.412	0.114	-0.037	0.455	0.295	-0.208			-0.633
9	MONTH-3/3-(8.333333, 12.000000) ...	0.194	0.223	-0.053	0.106	0.093	-0.111	-0.229	-0.213	1.295	0.093	
	Сумма	-0.386	-0.328	-0.325	-0.265	-0.193	0.027	0.088	0.203	2.584	0.725	
	Среднее	-0.043	-0.036	-0.036	-0.029	-0.021	0.003	0.010	0.023	0.287	0.081	
	Среднеквадратичное отклонение	0.345	0.306	0.361	0.245	0.258	0.234	0.216	0.435	0.648	0.620	

Таблица 4 – Модель INF5 (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000001	2. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000002	3. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000003	4. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000005	5. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000471	6. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000472	7. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000473	8. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000474	9. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000475	10. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000476	11. LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS E01000477
1	VALUE SUMMARY OF THE NUMBER OF CRIME...	-0.945	-0.817	-0.827	-0.688	-0.537	-0.015	0.156	0.451	1.660	2.040	
2	VALUE SUMMARY OF THE NUMBER OF CRIME...											
3	VALUE SUMMARY OF THE NUMBER OF CRIME...											
4	YEAR-1/3-(2008.000000, 2010.666667) ...	-0.077	-0.163	-0.044	-0.010	0.055	0.236	0.055	-0.448			0.300
5	YEAR-2/3-(2010.666667, 2013.333333) ...	0.067	0.055	0.304	0.054	0.304		-0.348	0.909	0.500	-0.143	
6	YEAR-3/3-(2013.333333, 2016.000000) ...	0.010	0.105	-0.258	-0.044	-0.355	-0.231	0.290	-0.460	0.484	-0.152	
7	MONTH-1/3-(1.000000, 4.666667) ...	-0.007	-0.025	-0.182	-0.081	0.047	-0.331	-0.084	0.368	-0.624	0.290	
8	MONTH-2/3-(4.666667, 8.333333) ...	-0.086	-0.109	0.307	0.044	-0.085	0.447	0.307	-0.180			-0.571
9	MONTH-3/3-(8.333333, 12.000000) ...	0.092	0.132	-0.124	0.037	0.038	-0.116	-0.221	-0.186	1.612	0.279	
	Сумма	-0.947	-0.820	-0.824	-0.688	-0.533	-0.009	0.155	0.452	3.631	2.043	
	Среднее	-0.105	-0.091	-0.092	-0.076	-0.059	-0.001	0.017	0.050	0.403	0.227	
	Среднеквадратичное отклонение	0.320	0.288	0.338	0.233	0.247	0.232	0.218	0.450	0.771	0.735	

1.5. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями приведены на рисунке 9.

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

Наименование модели и частного критерия	Интегральный критерий	точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модул. уровней сход. истинно полож. решений (STR)	Сумма модул. уровней сход. истинно отриц. решений (STN)	Сумма модул. уровней сход. ложно полож. решений (SFP)	Сумма модул. уровней сход. ложно отрицат. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко	Средний м. уровень сход. истинно полож. решений
1. ABS - частный критерий: количество встреч сомнений; "клас...	Корреляция абс. частот с обр...	0.025	0.986	0.049	4284.545	27150.137	158171.668	19.889	0.026	0.995	0.051	0.3
1.ABS - частный критерий: количество встреч сомнений; "клас...	Сумма абс. частот по признак...	0.019	1.000	0.038	4259.563	27150.137	9735.733	0.304	1.000	0.467	0.3	
2. PRC1 - частный критерий: усл. вероятность по признаку сред...	Корреляция усл. отн. частот с о...	0.025	0.986	0.049	4284.544	27150.132	158171.642	19.889	0.026	0.995	0.051	0.3
2.PRC1 - частный критерий: усл. вероятность по признаку сред...	Сумма усл.отн. частот по призна...	0.019	1.000	0.038	4087.739		191872.468		0.021	1.000	0.041	0.3
3. PRC2 - частный критерий: условная вероятность i-го признака.	Корреляция усл.отн. частот с о...	0.025	0.986	0.049	4284.545	27150.137	158171.668	19.889	0.026	0.995	0.051	0.3
3.PRC2 - частный критерий: условная вероятность i-го признака.	Сумма усл.отн. частот по призна...	0.019	1.000	0.038	3007.664		149502.516		0.020	1.000	0.039	0.2
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс: зна...	0.022	0.540	0.042	1791.309	104128.390	86404.790	1464.590	0.020	0.550	0.039	0.2
4.INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.024	0.631	0.046	471.837	38229.058	29454.627	177.581	0.016	0.727	0.031	0.0
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс: зна...	0.021	0.536	0.041	1717.450	100703.023	87398.237	1467.463	0.019	0.539	0.037	0.2
5.INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.022	0.620	0.043	398.316	27134.772	32235.750	133.722	0.012	0.749	0.024	0.0
6. INF3 - частный критерий: Уникалдат; разности между факти...	Семантический резонанс: зна...	0.026	0.633	0.049	2812.263	115308.797	104175.080	1140.887	0.026	0.711	0.051	0.3
6.INF3 - частный критерий: Уникалдат; разности между факти...	Сумма знаний	0.026	0.633	0.049	1586.820	5530.894	4966.552	1022.479	0.242	0.608	0.346	0.2
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс: зна...	0.023	0.519	0.044	1937.304	96997.932	88571.696	1506.798	0.021	0.562	0.041	0.3
7.INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.022	0.639	0.043	40.682	1230.587	2458.252	6.880	0.016	0.855	0.032	0.0
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс: зна...	0.023	0.515	0.043	1795.463	95757.242	87693.769	1524.801	0.020	0.541	0.039	0.2
8.INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.021	0.629	0.041	32.496	764.605	2418.061	4.236	0.013	0.885	0.026	0.0
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс: зна...	0.025	0.626	0.049	2631.726	101852.629	100741.434	1106.228	0.025	0.704	0.049	0.3
9.INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	0.022	0.633	0.043	551.461	26242.042	44856.039	139.087	0.012	0.799	0.024	0.0
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс: зна...	0.024	0.610	0.046	2490.509	99722.540	101909.905	988.228	0.024	0.716	0.046	0.3
10.INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	0.021	0.627	0.041	348.364	13732.744	37969.746	69.961	0.009	0.833	0.018	0.0

Помощь по меркам достоверности | Помощь по частотным распределениям | TP, TN, FP, FN | (TP,FP), (TN, FN) | (TP, FN)/(TP+FN) | Задать интервал сглаживания

Рисунок 9. Оценки достоверности моделей

Наиболее достоверной в данном приложении оказалась модель ABS при интегральном критерии «Сумма знаний». При этом точность модели (F-мера Ван Ризбергера) составляет 0,038, а точность модели (L1-мера профессора Луценко) - 0,467. L1-мера профессора Луценко является более достоверной, по сравнению с F-мерой Ван Ризбергера. 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 «Эйдос» используется F-мера Ван Ризбергера и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Е.В.Луценко [7] (рисунок 10).

Помощь по режимам: 4.1.3.6, 4.1.3.7, 4.1.3.8, 4.1.3.10: Виды прогнозов и меры достоверности моделей в системе "Эйдос-Х++".

ПОЛОЖИТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.

Предположим, модель дает такой прогноз: выпадет 1, 2, 3, 4, 5 или 6. В этом случае у нее будет 100% достоверность идентификации, т.е. не будет ни одного объекта, не отнесенного к тому классу, к которому он действительно относится, но при этом будет очень большая ошибка ложной идентификации, т.к. огромное количество объектов будет отнесено к классам, к которым они не относятся (и именно за счет этого у модели и будет очень высокая достоверность идентификации). Ясно, что такой прогноз бесплезен, поэтому он и назван мной псевдопрогнозом.

ОТРИЦАТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.

Представим себе, что мы выбрасываем кубик с 6 гранями, и модель предсказывает, что не выпадет: 1, 2, 3, 4, 5 и 6, а что-то из этого естественно выпало. Конечно, модель дает ошибку в прогнозе в том плане, что не предсказала, что выпадет, зато она очень хорошо угадала, что не выпадет. Но ясно, что выпадет что-то одно, а не все, что предсказано, поэтому такого рода предсказания хорошо оправдываются в том, что не произошло и плохо в том, что произошло, т.е. в этом случае у модели будет 100% достоверность не идентификации, но очень низкая достоверность идентификации.

ИДЕАЛЬНЫЙ ПРОГНОЗ.

Если в случае с кубиком мы прогнозируем, что выпадет, например 1, и соответственно прогнозируем, что не выпадет 2, 3, 4, 5, и 6, то это идеальный прогноз, имеющий, если он осуществляется, 100% достоверность идентификации и не идентификации. Идеальный прогноз, который полностью снимает неопределенность о будущем состоянии объекта прогнозирования, на практике удается получить крайне редко и обычно мы имеем дело с реальным прогнозом.

РЕАЛЬНЫЙ ПРОГНОЗ.

На практике мы чаще сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой. Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у Вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов.

Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно относить объекты к классам, которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Этот критерий предложен и реализован в системе "Эйдос" проф. Е.В.Луценко в 1994 году. Эта мера достоверности модели предполагает два варианта нормировки: $\{-1, +1\}$ и $\{0, 1\}$:

$$L1 = \frac{TP + TN - FP - FN}{TP + TN + FP + FN} \quad (\text{нормировка: } \{-1, +1\})$$

$$L2 = \frac{1 + (TP + TN - FP - FN)}{(TP + TN + FP + FN)} / 2 \quad (\text{нормировка: } \{0, 1\})$$

где: TP - истинно-положительное решение; TN - истинно-отрицательное решение; FP - ложно-положительное решение; FN - ложно-отрицательное решение;

Классическая F-мера достоверности моделей Ван Ризбергена:

Precision = TP/(TP+FP) - точность модели;

Recall = TP/(TP+FN) - полнота модели;

F-мера = $2 * (Precision * Recall) / (Precision + Recall)$.

В АСК-анализе и системе "Эйдос" предлагается L-мера, представляющая собой нечеткое мультиклассовое обобщение классической F-меры достоверности моделей Ван Ризбергена:

SPrecision = STP/(STP+SFP) - нечеткая мультиклассовая точность модели;

SRecall = STP/(STP+SFN) - нечеткая мультиклассовая полнота модели;

L-мера = $2 * (SPrecision * SRecall) / (SPrecision + SRecall)$ - нечеткая мультиклассовая достоверность модели.

где:

STP - Сумма модулей сходства истинно-положительных решений; STN - Сумма модулей сходства истинно-отрицательных решений;

SFP - Сумма модулей сходства ложно-положительных решений; SFN - Сумма модулей сходства ложно-отрицательных решений.

Рисунок 10. Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

Также обращает на себя внимание, 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 интеллектуальных технологий. На рисунке 11 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели ABS.

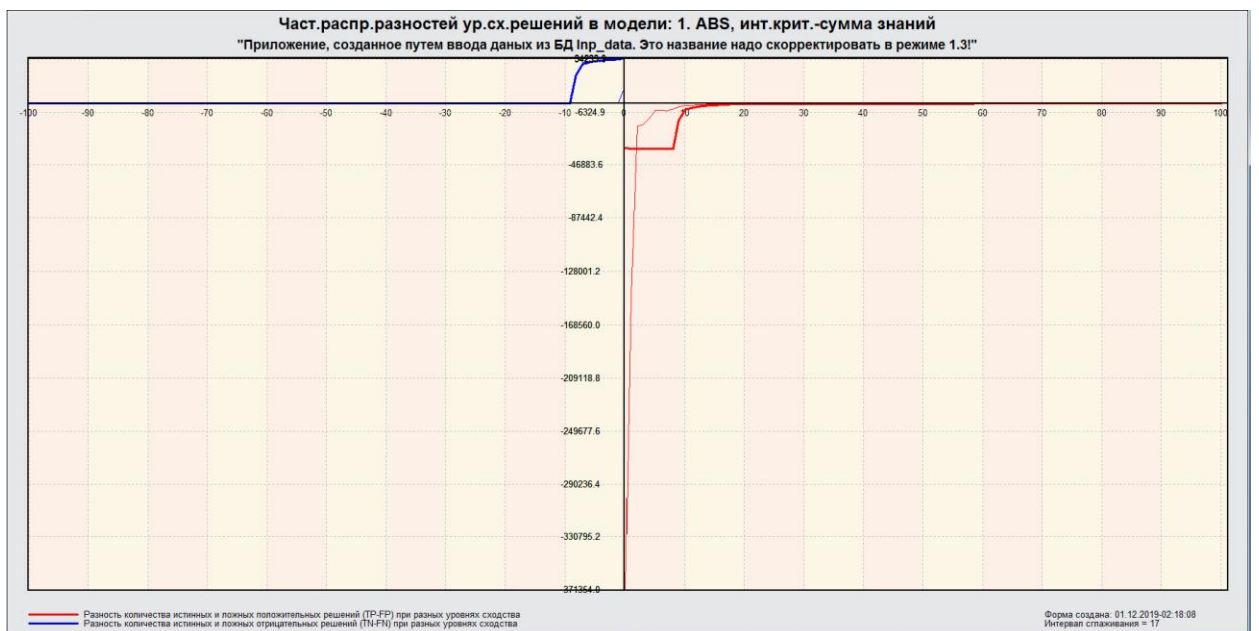
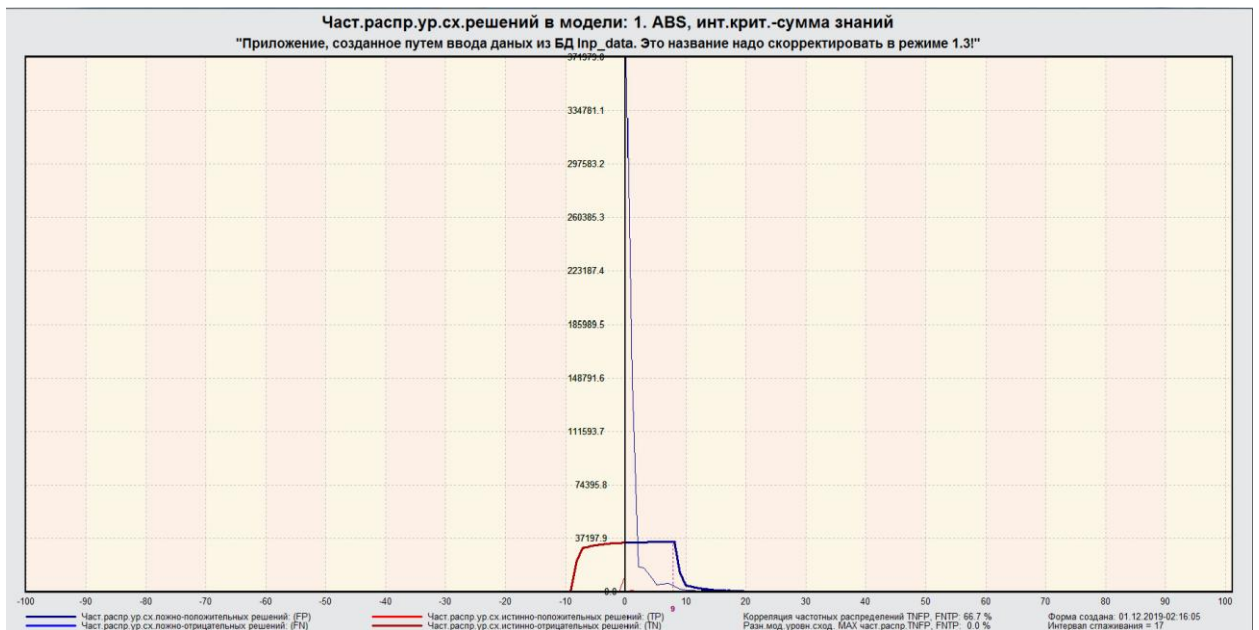


Рисунок 11. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели ABS.

Из рисунка 11 видно, что:

- наиболее достоверная модель ABS лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 9);

– модуль уровня сходства-различия в наиболее достоверной модели ABS для верно идентифицированных и верно неидентифицированных объектов значительно выше, чем для ошибочно идентифицированных и ошибочно неидентифицированных. Это верно практически для всего диапазона уровней сходства-различия, кроме небольших по модулю значений в диапазоне от 0 до 15% уровня сходства. Для очень больших значений уровней сходства-различия (более 90%) также различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями практически отсутствует[1].

Любые данные о наблюдениях можно считать суммой истинного значения и шума, причем ни первое, ни второе неизвестны. Поэтому имеет смысл сравнить созданные модели с чисто случайными моделями, совпадающими по основным характеристикам. В системе «Эйдос» есть лабораторная работа № 2.01: «Исследование RND-модели при различных объемах выборки». Если данная работа устанавливается при отсутствии текущего приложения, то все параметры создаваемых моделей задаются вручную, если же текущая модель существует, как в нашем случае, то все основные ее параметры определяются автоматически (рисунок 12):

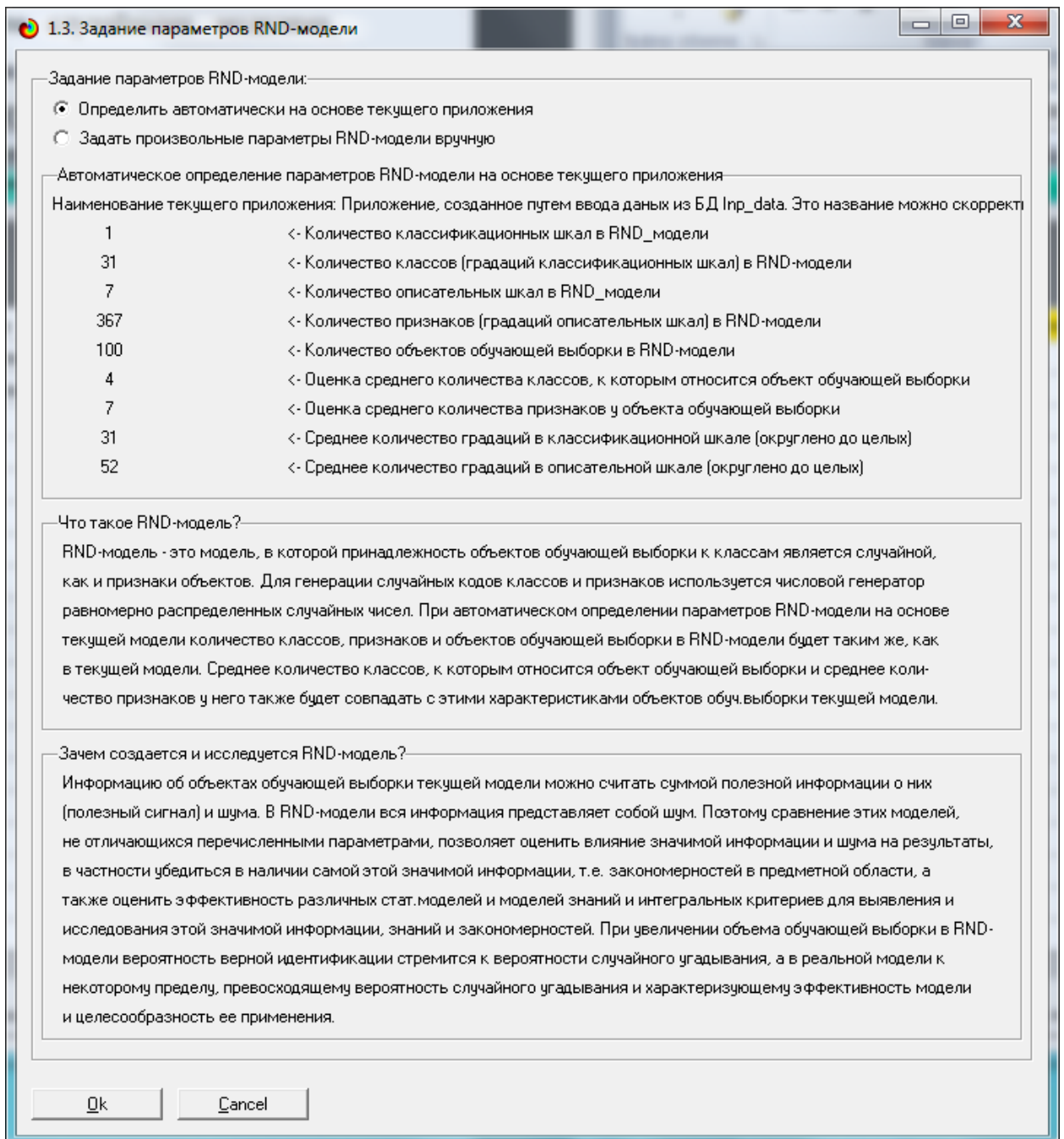
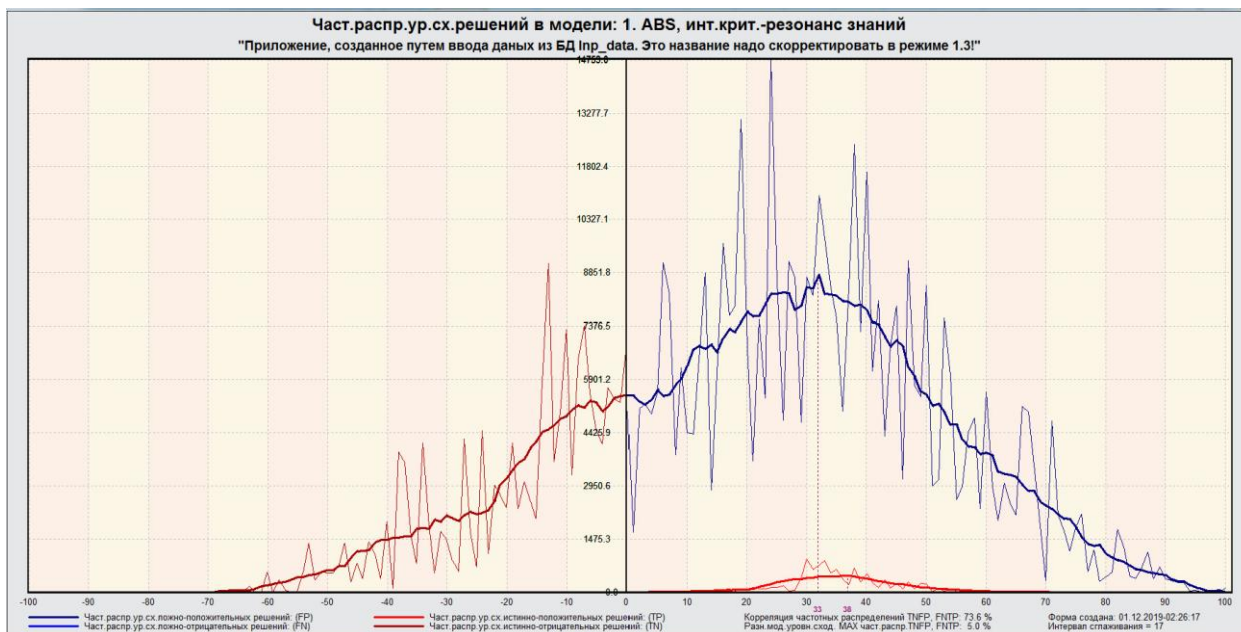


Рисунок 12. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 13 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели ABS.



На рисунке 13 приведены данные по достоверности статистических и когнитивных моделей, созданных на основе случайной выборки.

Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели (рисунки 12 и 13) объясняется тем, что в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны. Если же такой информации в модели нет, то и распределение получается типа, приведенного на рисунке 14[4].

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Фильтрация Фишбергена	Сигнальный уровень истинно-положительных решений (STP)	Сигнальный уровень истинно-отрицательных решений (STN)	Сигнальный уровень ложно-положительных решений
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	12000	11832	154394	460606	168	0.025	0.986	0.049	4284.545	27150.137	158171
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сигна абс. частот по признак...	12000	12000	10075	604925		0.019	1.000	0.038	4259.563		9735
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн. частот с о...	12000	11828	156417	458583	172	0.025	0.986	0.049	4284.544	27150.132	158171
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Сигна усл.отн. частот по при...	12000	12000	10075	604925		0.019	1.000	0.038	4087.739		191872
3. PRC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн. частот с о...	12000	11828	155444	459556	172	0.025	0.986	0.049	4284.545	27150.137	158171
3. PRC2 - частный критерий: условная вероятность его признака...	Сигна усл.отн. частот по при...	12000	12000	10075	604925		0.019	1.000	0.038	3007.664		149502
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12000	6476	327500	287500	5524	0.022	0.540	0.042	1791.309	104128.390	86404
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сигна знаний	12000	7568	304446	310554	4432	0.024	0.631	0.046	471.837	38229.058	29454
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	12000	6432	320724	294276	5568	0.021	0.536	0.041	1717.450	100703.023	87398
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сигна знаний	12000	7440	286230	328770	4560	0.022	0.620	0.043	398.316	27134.772	32235
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	12000	7596	325995	289005	4404	0.026	0.633	0.049	2812.268	115308.797	104175
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Сигна знаний	12000	7596	325995	289005	4404	0.026	0.633	0.049	1586.820	5530.894	4966
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12000	6231	348155	266845	5769	0.023	0.519	0.044	1937.304	96997.932	88571
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сигна знаний	12000	7672	278529	336471	4328	0.022	0.639	0.043	40.683	1230.587	2458
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	12000	6177	347730	267270	5823	0.023	0.515	0.043	1795.463	95757.242	87693
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сигна знаний	12000	7553	263976	351024	4447	0.021	0.629	0.041	32.496	764.605	2418
9. INF6 - частный критерий: разн. усл.и безуслов. вероятностей; вер...	Семантический резонанс зна...	12000	7517	326634	288366	4483	0.025	0.626	0.049	2631.726	101852.629	100741
9. INF6 - частный критерий: разн. усл.и безуслов. вероятностей; вер...	Сигна знаний	12000	7596	281725	333275	4404	0.022	0.633	0.043	551.461	26242.042	44856
10. INF7 - частный критерий: разн. усл.и безуслов. вероятностей; ве...	Семантический резонанс зна...	12000	7314	319621	295379	4686	0.024	0.610	0.046	2490.509	99722.540	101909
10. INF7 - частный критерий: разн. усл.и безуслов. вероятностей; ве...	Сигна знаний	12000	7527	264992	350008	4473	0.021	0.627	0.041	348.364	13732.744	37969

Рисунок 14. Достоверность статистических и когнитивных моделей, созданных на основе случайной выборки

На основе его сравнения с рисунком 9 можно сделать следующие **ВЫВОДЫ**:

- достоверность лучшей модели ABS, отражающей реальный объект моделирования, примерно на 21% выше, чем аналогичной случайной модели;
- различие между достоверностью статистических моделей и моделей знаний, созданных на основе случайной выборки, значительно меньше, чем у моделей, отражающих реальный объект моделирования;
- в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны, причем примерно 1/3 достоверности обусловлена отражением в реальных моделях закономерностей предметной области, а 2/3 достоверности обусловлено наличием шума в исходных данных. На основании этого можно предположить, что в исходных данных уровень сигнала о реальных причинно-следственных связях в моделируемой предметной области примерно в два раза ниже уровня шума[3].

2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи

В соответствии с технологией АСК-анализа зададим текущей модель ABS (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1.

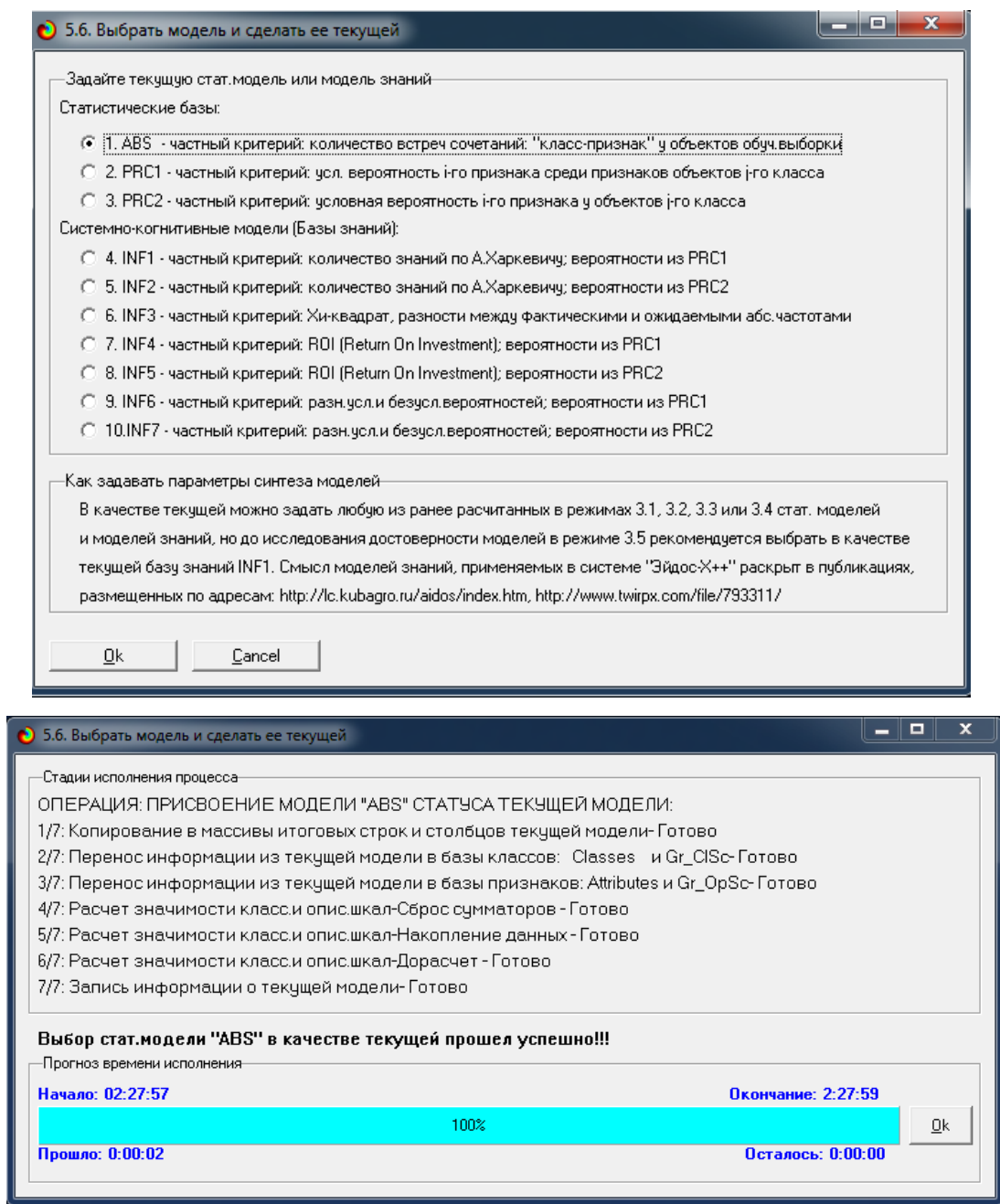


Рисунок 15. Экранные формы режима задания модели в качестве текущей

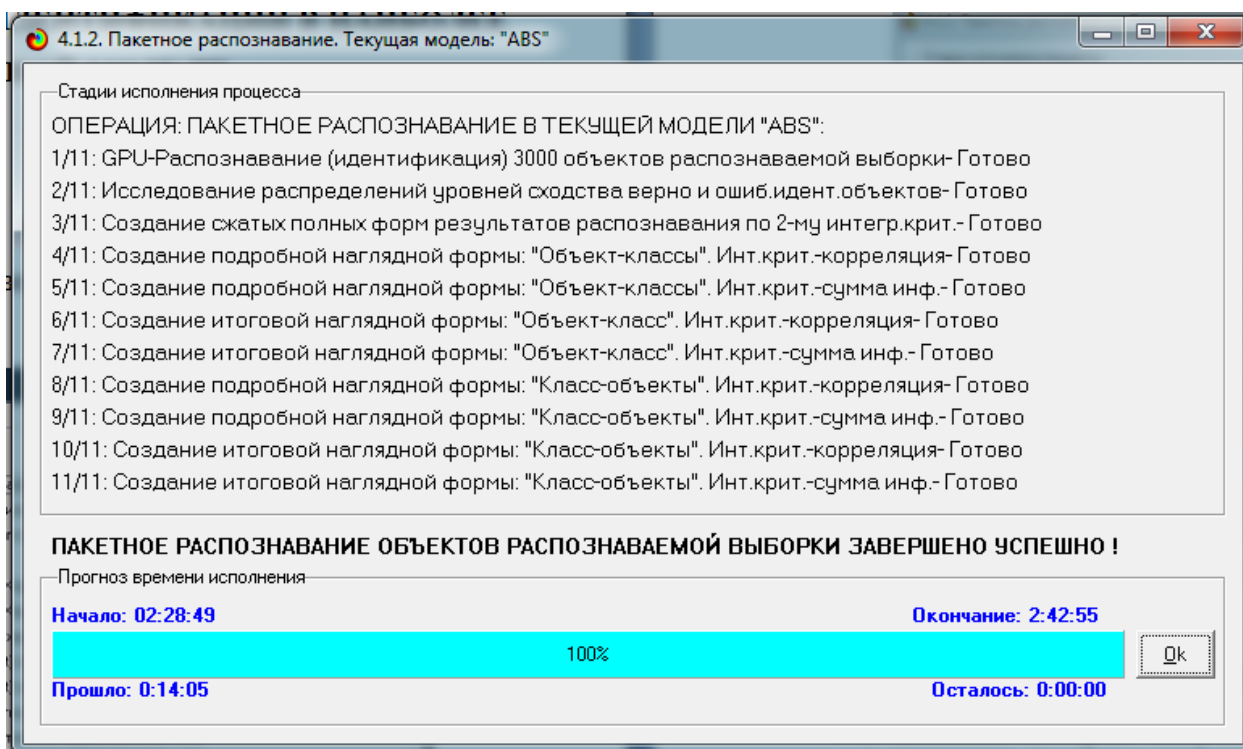


Рисунок 16. Экранная форма режима пакетного распознавания в текущей модели

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 «Эйдос» 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.

Ниже кратко рассмотрим некоторые из них.

На рисунках 17 и 18 приведены примеры прогнозов высокой и низкой достоверности частоты объектов и классов в наиболее достоверной модели ABS на основе наблюдения предыстории их развития:

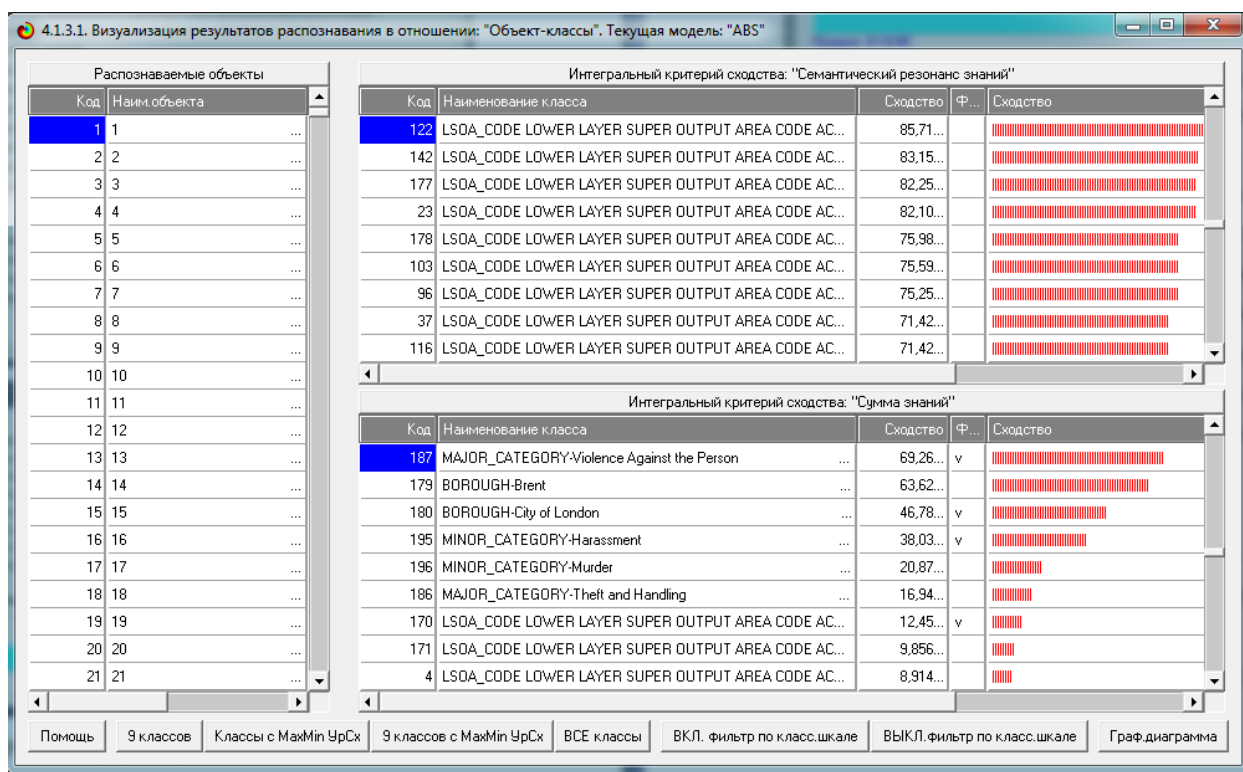


Рисунок 17. Пример идентификации объектов в модели ABS

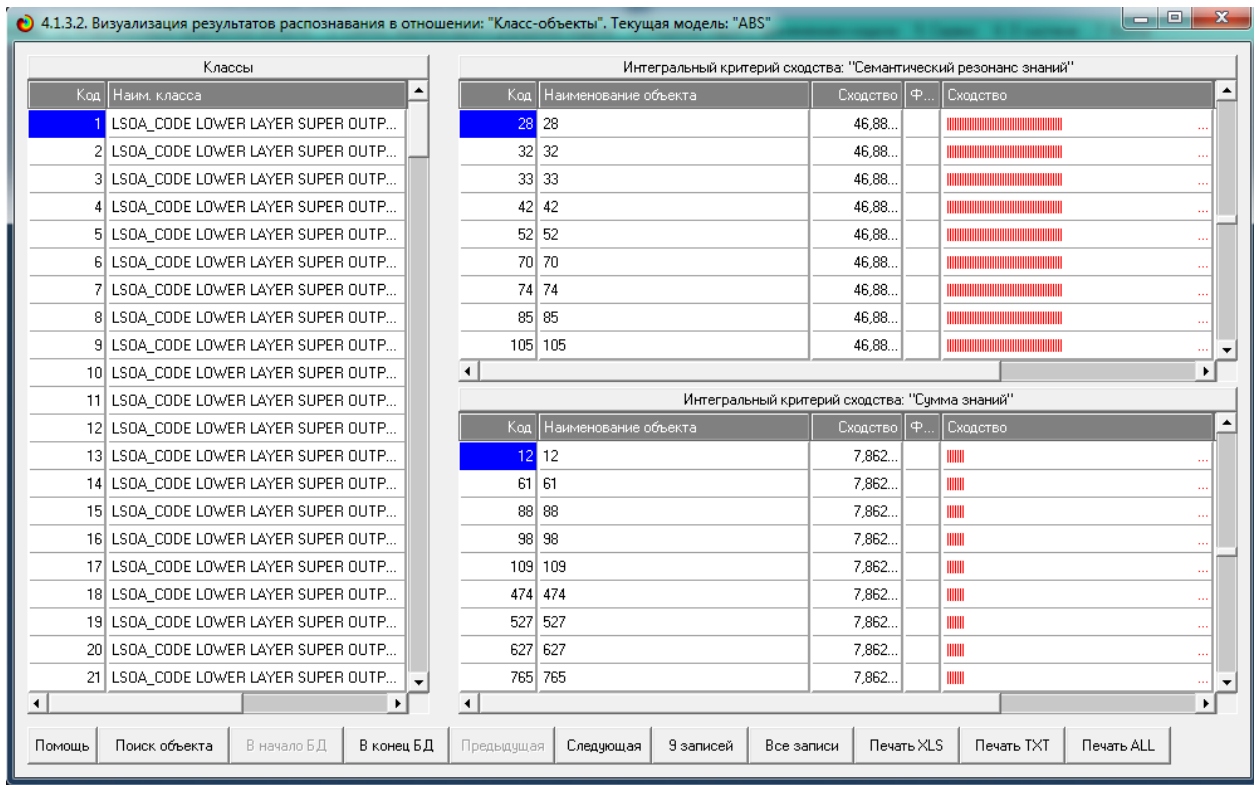


Рисунок 18. Пример идентификации классов в модели ABS

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 19)

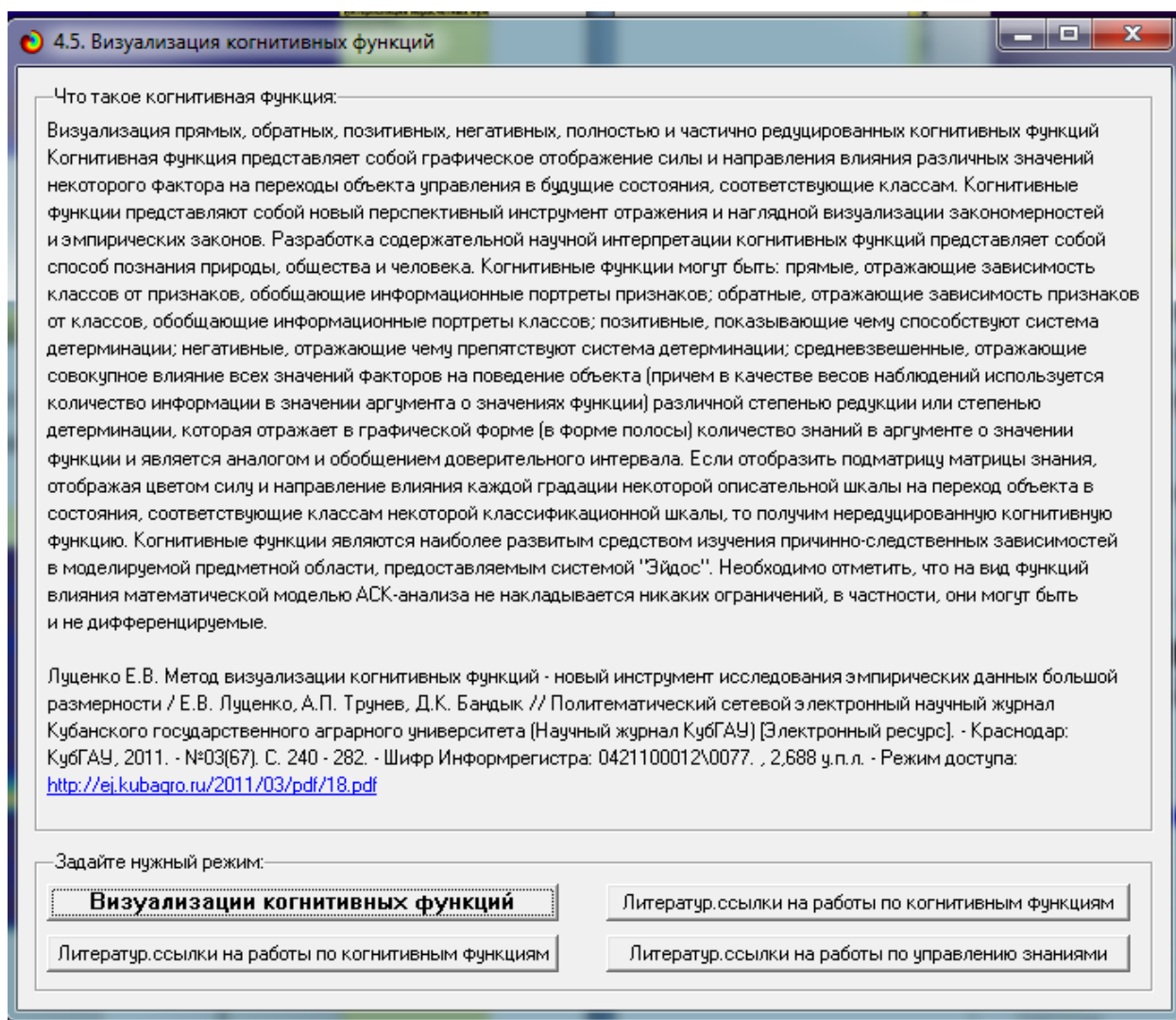


Рисунок 19. Экранная форма режима визуализации когнитивных функций

В когнитивных функциях количество информации в значениях аргумента о значениях функции отображается цветом (красным максимальное, синим минимальное), линией соединены значения функции о которых в значении аргумента содержится максимальное количество информации, ширина линии (аналог доверительного интервала) отражает степень неопределенности значения функции, которое тем ниже, чем больше информации о нем в значении функции (рис. 20–24):

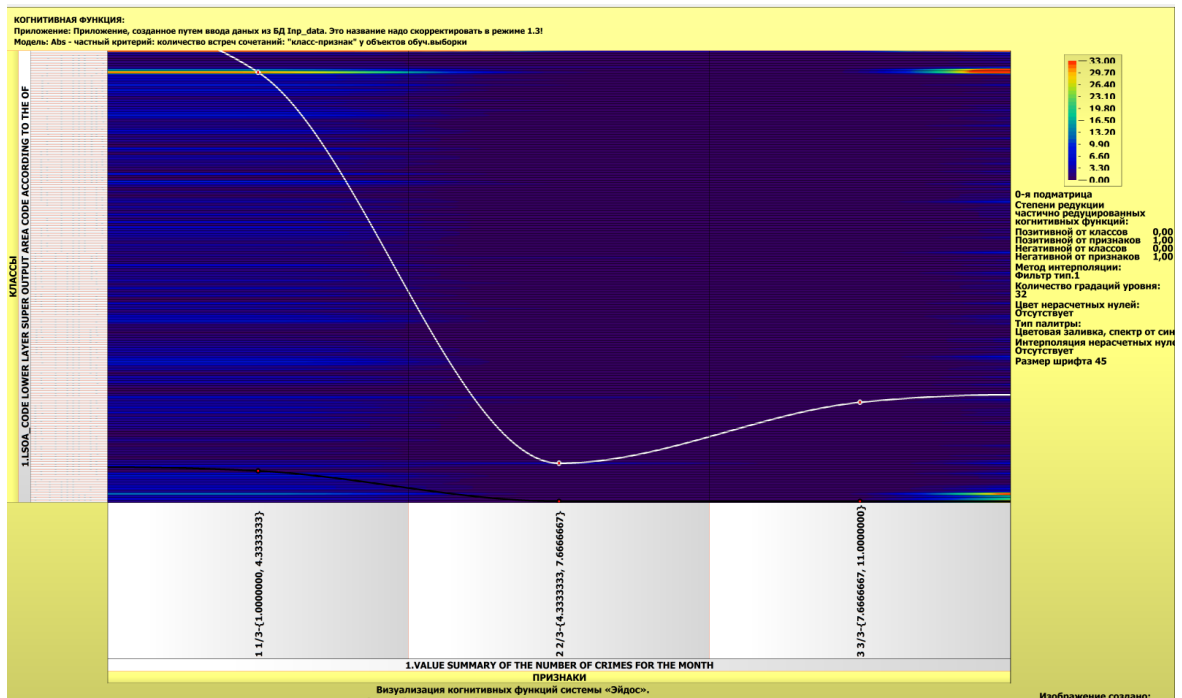


Рисунок 20. Когнитивная функция, отражающая взаимосвязь кода зоны вывода супер-уровня и количества преступлений за месяц.

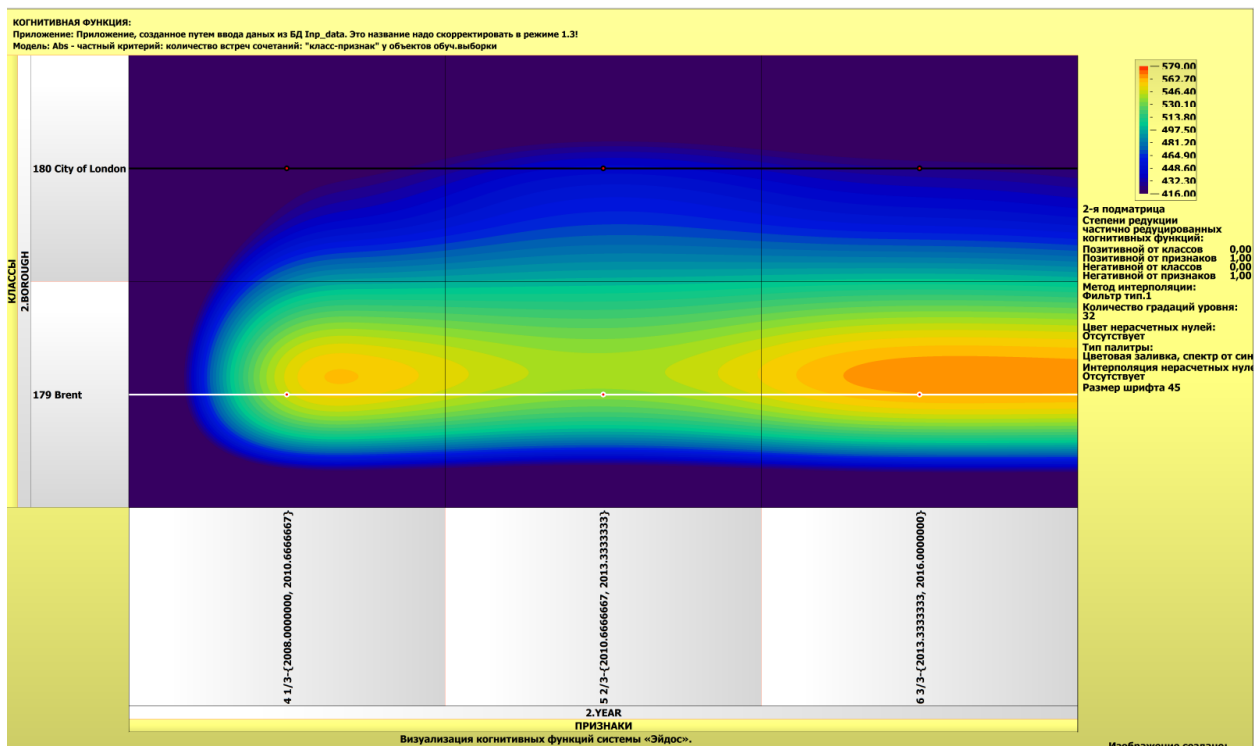


Рисунок 21. Когнитивная функция, отражающая взаимосвязь района и года

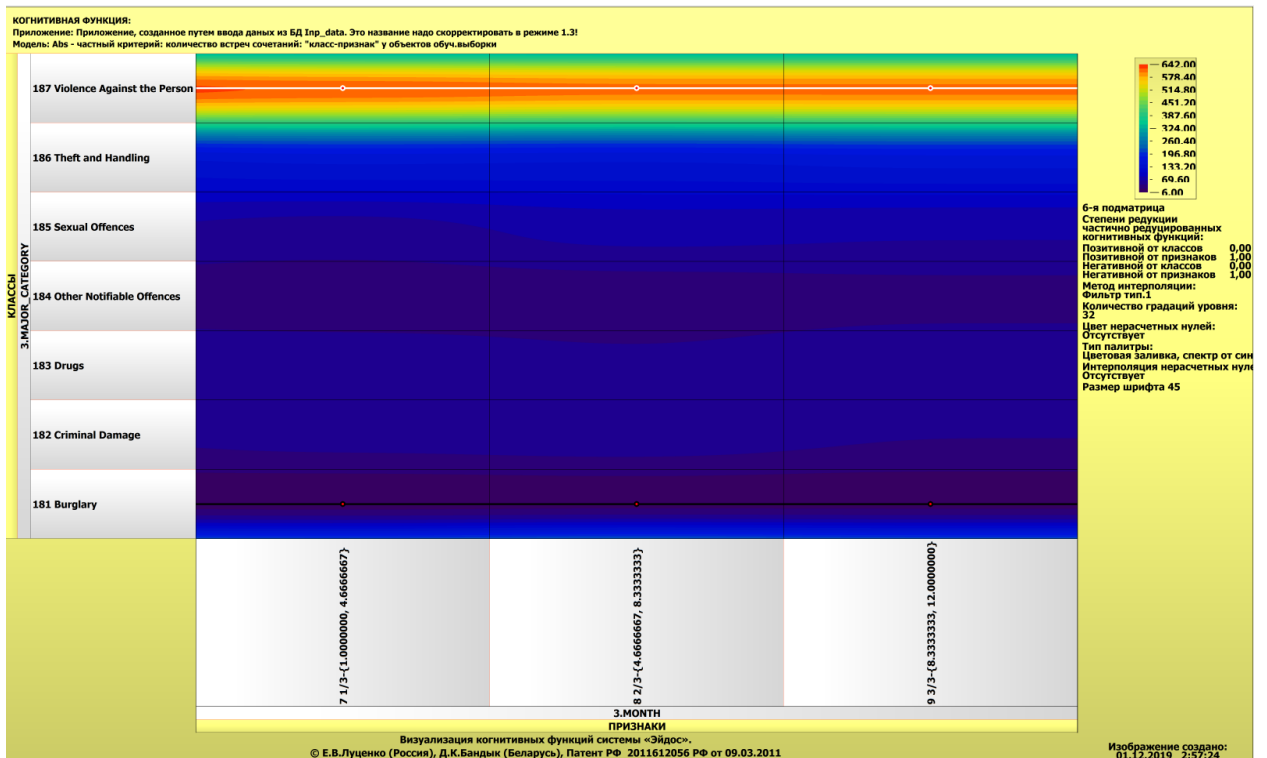


Рисунок 22. Когнитивная функция, отражающая взаимосвязь основной категории и месяца.

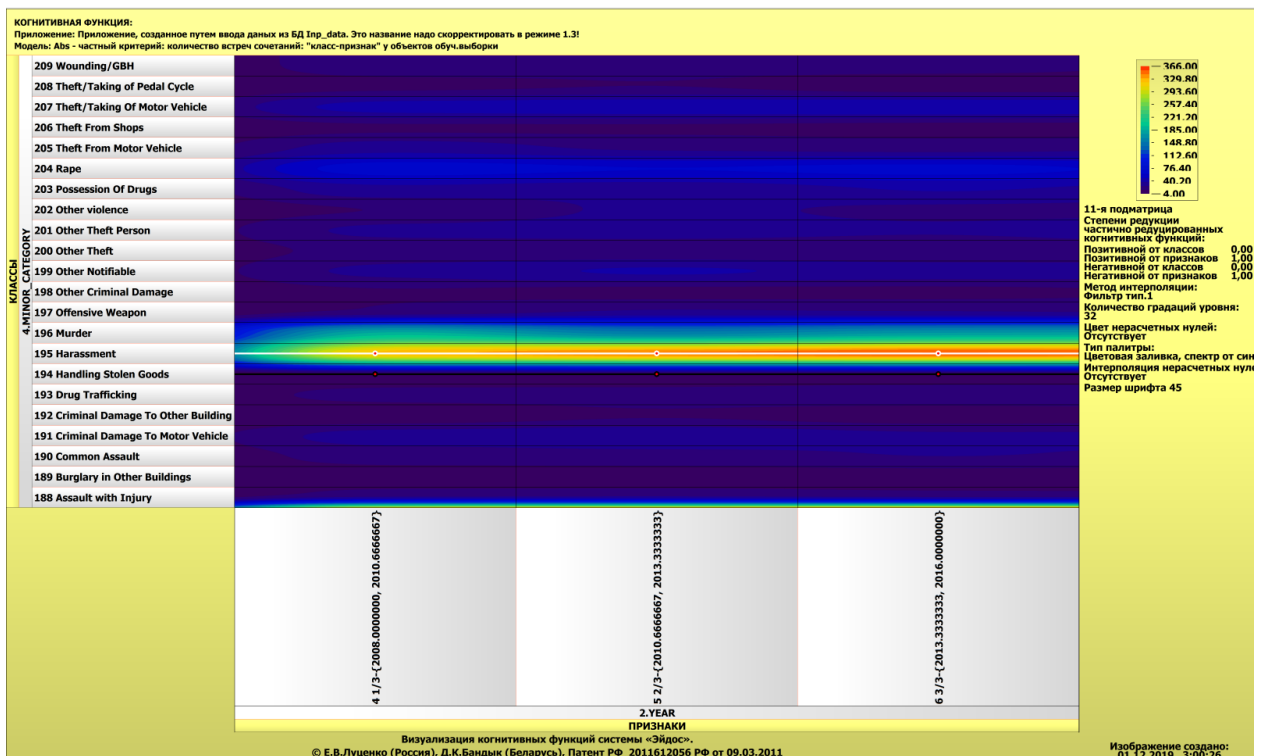


Рисунок 23. Когнитивная функция, отражающая взаимосвязь дополнительной категории и года.

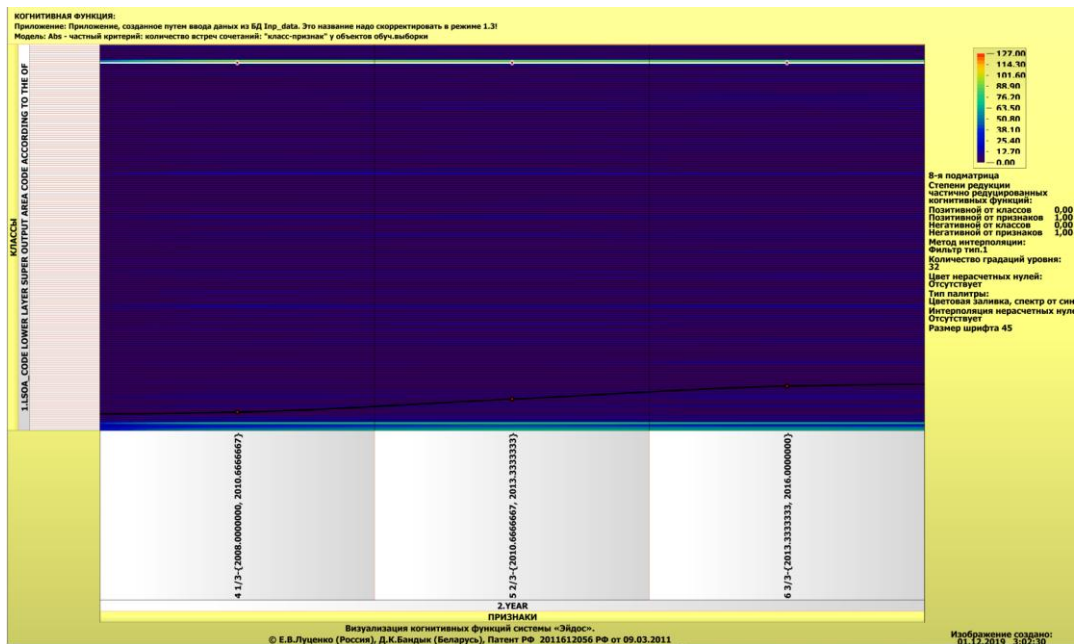


Рисунок 24. Когнитивная функция, отражающая взаимосвязь кода зоны вывода супер-уровня и года.

2.3 Кластерно-конструктивный анализ

Сходство-различие обобщенных образов различных результатов научной деятельности по характерным для них системам значений показателей. Результаты сравнения классов по системе характерных для смерти от огнестрельного оружия приведены на рисунок 25:

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса:1 "LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCORDING TO THE OFFICE FOR NATIONAL STATISTICS-E01..."

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	1	1	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA ACCO...	100.000
2	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	2	207	MINOR_CATEGORY-Theft/Taking Of Motor Vehicle	99.905
3	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	3	196	MINOR_CATEGORY-Murder	99.552
4	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	4	185	MAJOR_CATEGORY-Sexual Offences	99.437
5	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	5	204	MINOR_CATEGORY-Rape	99.437
6	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	6	2	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCO...	99.287
7	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	7	4	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCO...	99.255
8	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	8	193	MINOR_CATEGORY-Drug Trafficking	99.103
9	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	9	184	MAJOR_CATEGORY-Other Notifiable Offences	98.784
10	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	10	199	MINOR_CATEGORY-Other Notifiable	98.784
11	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	11	182	MAJOR_CATEGORY-Criminal Damage	98.778
12	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	12	205	MINOR_CATEGORY-Theft From Motor Vehicle	98.562
13	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	13	191	MINOR_CATEGORY-Criminal Damage To Motor Vehicle	98.547
14	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	14	202	MINOR_CATEGORY-Other violence	98.252
15	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	15	190	MINOR_CATEGORY-Common Assault	97.164
16	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	16	183	MAJOR_CATEGORY-Drugs	97.101
17	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	17	198	MINOR_CATEGORY-Other Criminal Damage	96.448
18	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	18	197	MINOR_CATEGORY-Offensive Weapon	96.373
19	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	19	171	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCO...	96.069
20	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	20	188	MINOR_CATEGORY-Assault with Injury	94.967
21	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	21	3	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCO...	94.385
22	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	22	99	LSOA_CODE LOWER LAYER SUPER OUTPUT AREA CODE ACCO...	94.232
23	LSOA_CODE LOWER LAYER SUPER OUTPUT A...	23	206	MINOR_CATEGORY-Theft From Shoos	93.895

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 25 - Результаты кластерно-конструктивного анализа классов

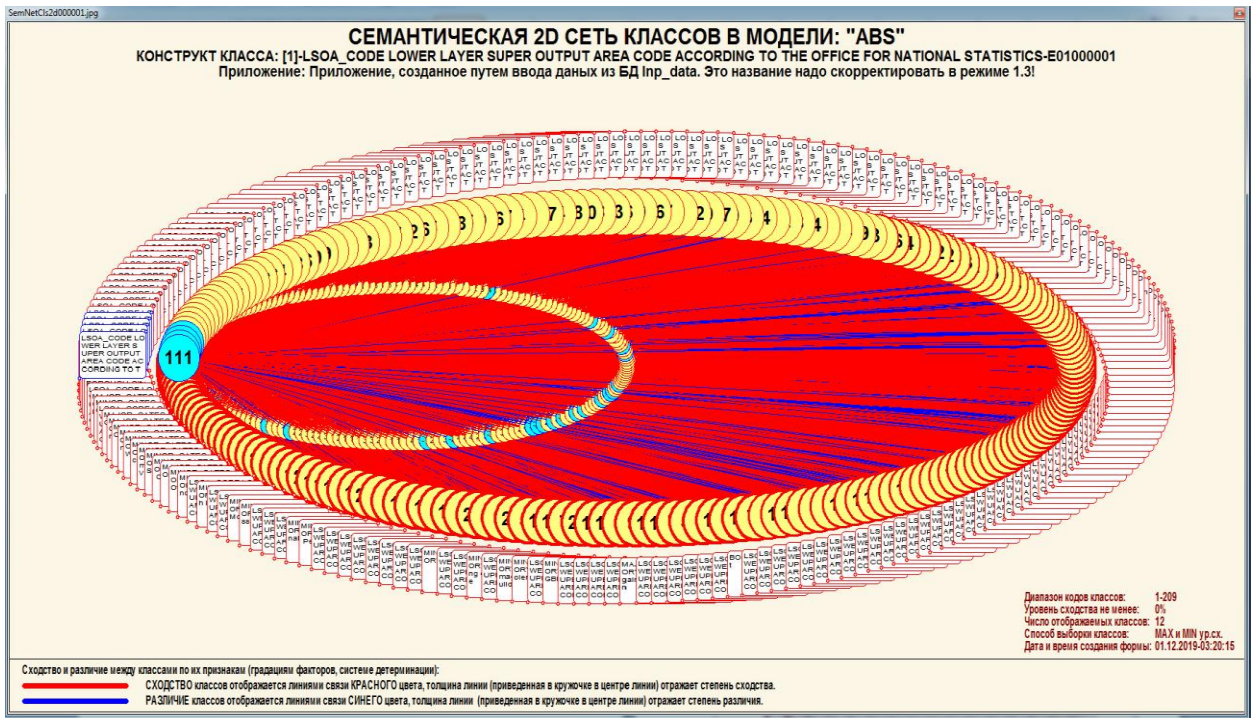


Рисунок 26. Результаты семантическая 2D сеть классов

2.4 Нейронная сеть

Отметим также, что приведенная когнитивная диаграмма формируется системой «Эйдос» автоматически на основе созданных моделей.

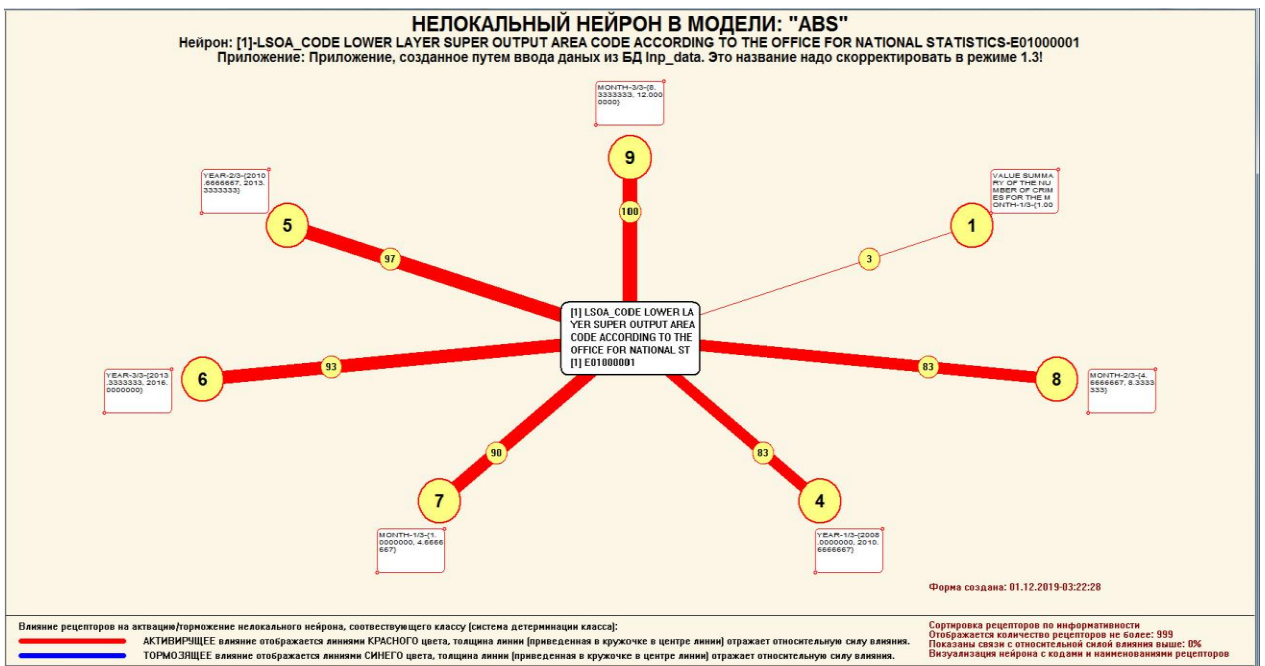


Рисунок 27. Графическое отображение нелокальных нейронов

Для каждого технологического фактора в соответствии с предложенной моделью определяется величина и направление его влияния на осуществление всех желаемых и не желаемых хозяйственных ситуаций. Для каждой ситуации эта информация отображается в различных текстовых и графических формах, в частности в форме нелокального нейрона (рисунок 26). На данной диаграмме цвет линии означает знак связи (красный – положительная, синий – отрицательная), а толщина – ее модуль.

Дополнение модели нейрона связями факторов позволяет построить классическую когнитивную карту ситуации (будущего состояния АОУ). Детальная внутренняя структура любой связи отображается в форме инвертированной когнитивной диаграммы (рисунок 27). Необходимо отметить, что все указанные графические формы генерируются системой "Эйдос" автоматически в соответствии с созданной моделью[9].

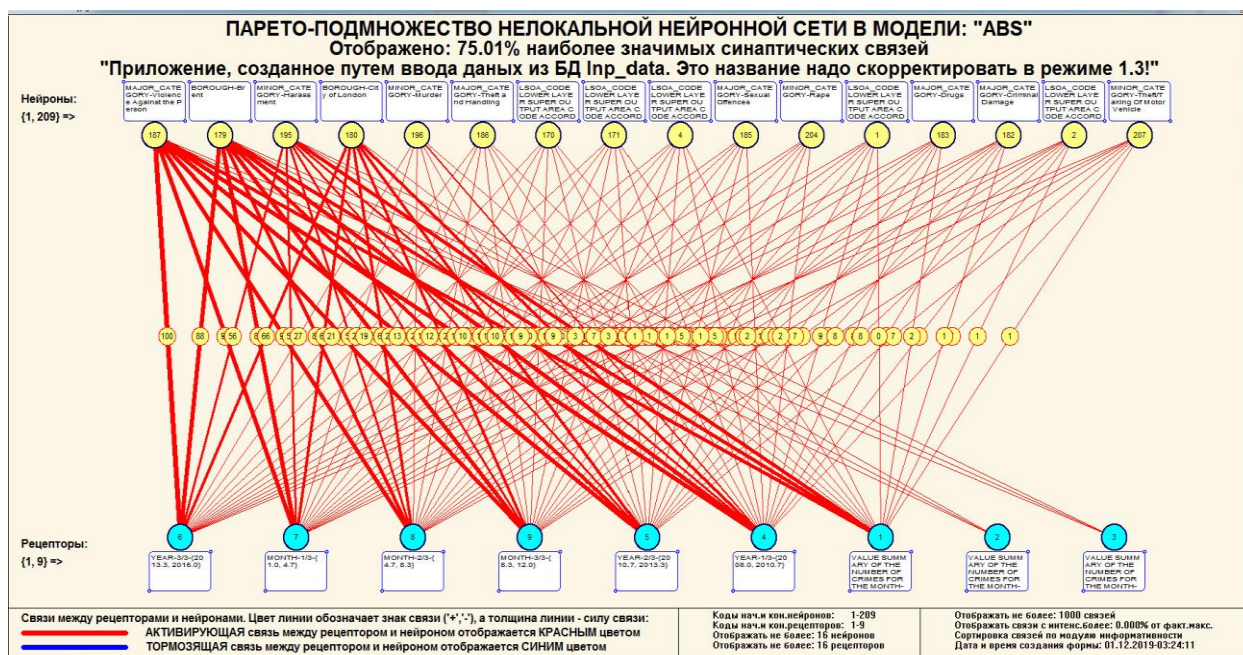


Рисунок 28 – Паретто – подмножеств нелокальной нейронной сети в системе «Эйдос - X»

2.5 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 неформализуемым 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 29).

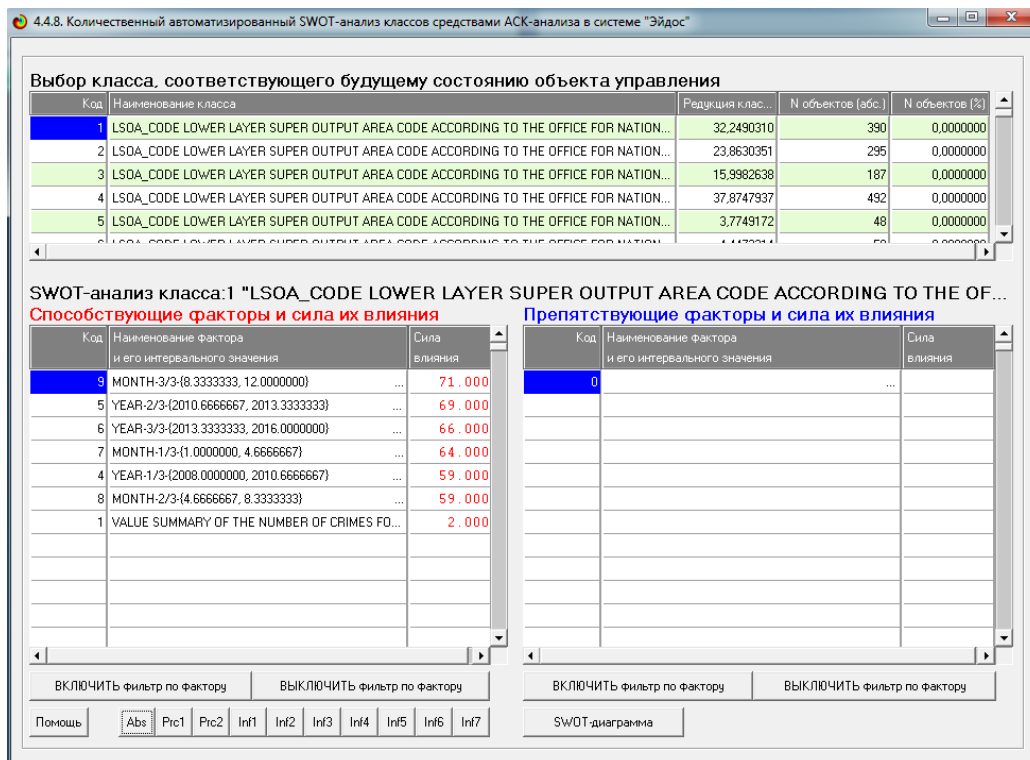


Рисунок 29. Пример SWOT-матрицы в модели ABS

На рисунке 30 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели ABS.

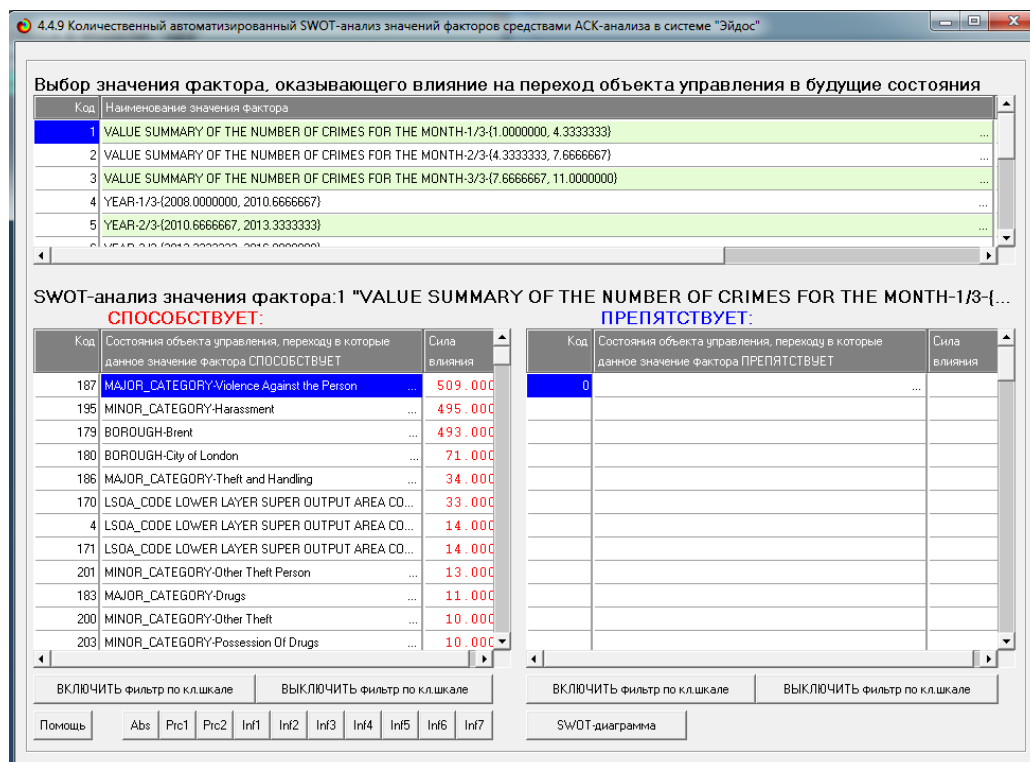


Рисунок 30. Табличная выходная форма количественного автоматизированного SWOT- и PEST- анализа средствами системы «Эйдос»

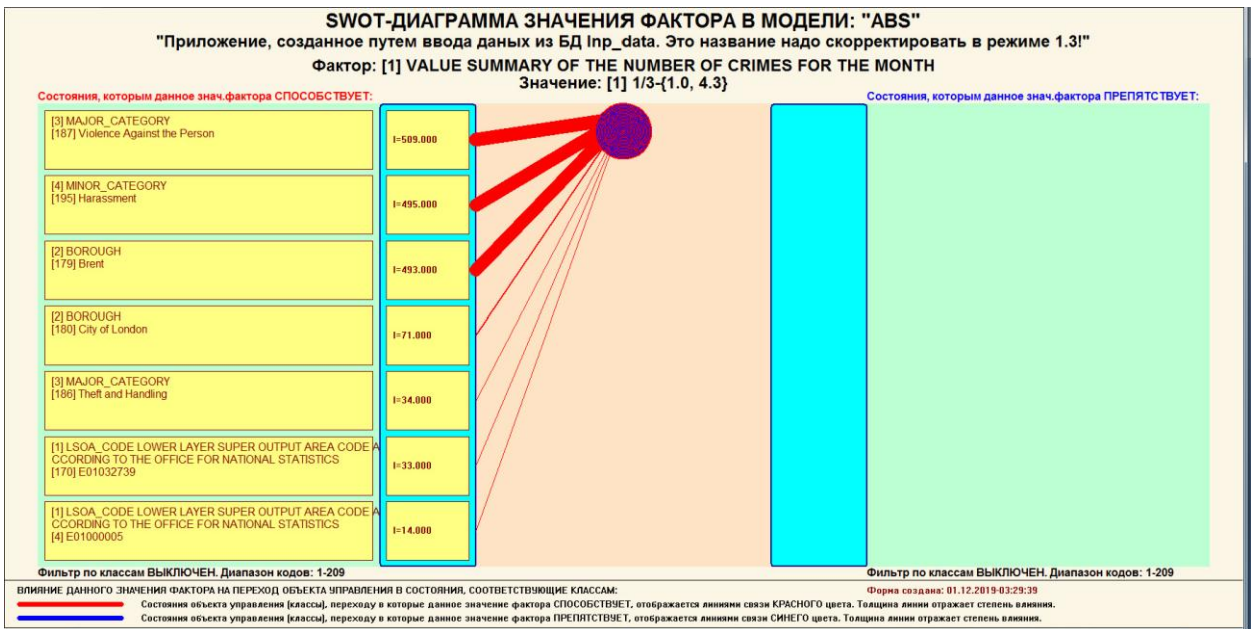
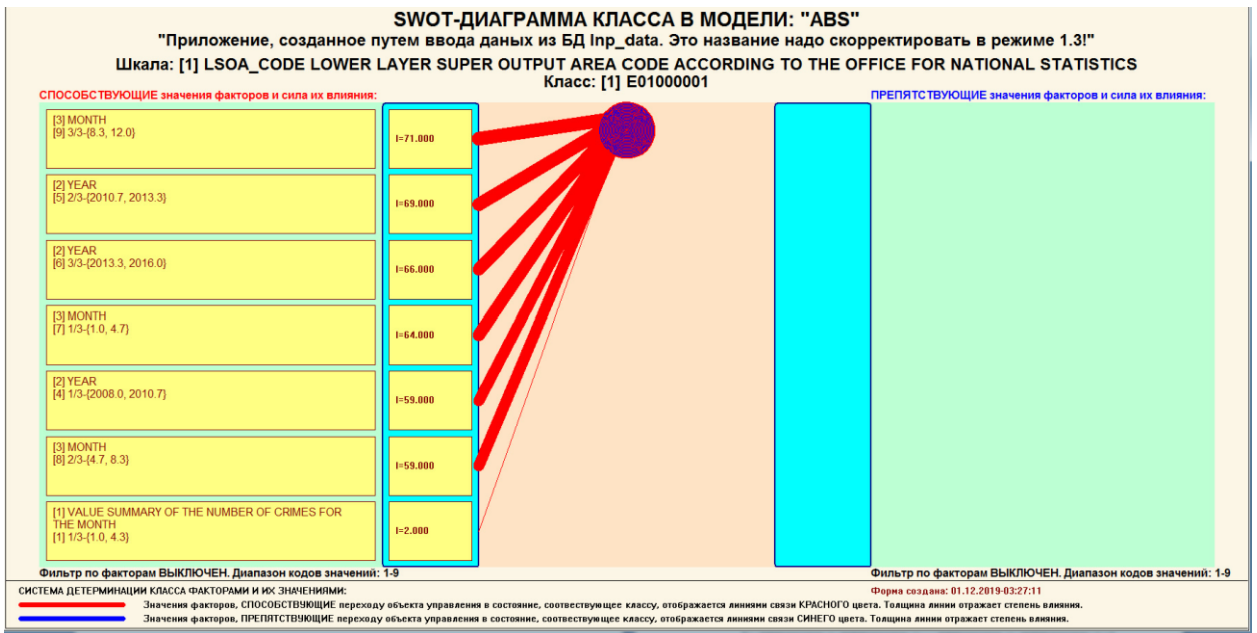


Рисунок 31. Графическая выходная форма количественного автоматизированного SWOT- и PEST- анализа средствами системы «Эйдос»

ЗАКЛЮЧЕНИЕ

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

Благодаря системам искусственного интеллекта, есть возможность рассчитать большой объем данных для получения статистики и вероятностей этих событий. В данной курсовой работе приводится пример использования базы данных и инструментария системы искусственного интеллекта «Эйдос».

Основное направление искусственного интеллекта - это представление знаний и разработка систем, основанных на знаниях. Оно связано с разработкой моделей представления знаний, созданием баз знаний, образующих ядро экспертных систем.

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.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/01.pdf>, 3,75 у.п.л.3.

2. 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.

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. [Электронный ресурс]. Статья "Emergency – 911 Calls": <https://www.kaggle.com/mchirico/montcoalert>, свободный. - Загл. сэкрана. Яз.анг.

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

6.Луценко Е.В. Количественная оценка степени манипулирования индексом Хирша и его модификация, устойчивая к манипулированию / Е.В.

Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №07(121). С. 202 – 234. – IDA [article ID]: 1211607005. – Режим доступа: <http://ej.kubagro.ru/2016/07/pdf/05.pdf>, 2,062 у.п.л.
<http://dx.doi.org/10.21515/1990-4665-121-0057>.

7. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- X++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №09(083). С. 328 – 356. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 у.п.л.

8. 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.