

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к курсовой работе

по дисциплине: ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ
СИСТЕМЫ

на тему:

**Решение задач идентификации и прогнозирования с применением АСК-
анализа и системы «Эйдос» на основе данных репозитория UCI: UFO Sightings**

выполнил студент группы ИТ1421 **Рассулов Антон Андреевич**

Допущена к защите _____

Руководитель проекта Луценко Евгений Вениаминович,

д.э.н., к.т.н., профессор

(подпись, расшифровка подписи)

Нормоконтролер _____ Николаева Ирина Валентиновна, к.т.н.,
доцент

(подпись, расшифровка подписи)

Защищена _____

Оценка

(дата)

Члены комиссии _____

В.И. Лойко

Е.В. Луценко

И. В. Николаева

(подпись, дата, расшифровка подписи)

Краснодар

2017 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:

Зав. кафедрой КТС _____ В. И. Лойко

**ЗАДАНИЕ
на курсовую работу**

Студентке: ИТ1421 группы 3 курса
Факультета прикладной
информатики
Специальности: 09.03.02 Информационные системы и
технологии

(шифр)

Рассулов Антон Андреевич
(Ф.И.О.)

Тема проекта: **Решение задач идентификации и прогнозирования с**
применением АСК-анализа и системы «Эйдос» на основе данных репозитория
UCI: UFO Sightings

Содержание задания: 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

Объем работы:

а) пояснительная записка к работе _____ листа формата
А4

б) графическая часть _____ лист формата А4

Рекомендуемая литература: Луценко Е.В. Лабораторный практикум по
интеллектуальным информационным системам: Учебное пособие для бакалавриата. 7-е
изд., перераб. и доп.- Краснодар: КубГАУ – 2016, – 615 с., в электронном виде на сайте
автора: <http://lc.kubagro.ru/aidos/p14.htm>

Срок выполнения проекта: с « » _____ по « » _____ 2017 г.

Срок защиты: « » _____ 2017 г.

Дата выдачи задания: « » _____ 2017 г.

Дата сдачи проекта на кафедру: « » _____ 2017 г.

Руководитель проекта: Луценко Евгений Вениаминович

д.э.н., к.т.н., профессор
(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ « » _____ 2017 г.

(подпись, дата)

Краснодар
2017 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ.....	7
1.1. Описание решения	7
1.2. Подготовка данных в промежуточные файлы MS Excel	7
1.3. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос"	10
1.3. Синтез и верификация статистических и интеллектуальных моделей .	14
1.4. Виды моделей системы «Эйдос»	16
1.5. Результаты верификации моделей	17
2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	24
2.1. Решение задачи	24
2.2. Когнитивные функции	27
2.3. SWOT и PEST матрицы и диаграммы	31
2.4. Нелокальные нейроны и нейронная сеть	33
2.5. Кластерный и конструктивный анализ классов	33
ЗАКЛЮЧЕНИЕ	36
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	37

ВВЕДЕНИЕ

На сегодняшний день создание систем искусственного интеллекта – это перспективный шаг к расширению возможностей компьютерных наук и автоматизации разумного поведения, опирающийся на теоретические и прикладные принципы. Однако возникает необходимость оценки качества математических моделей этих систем.

В данной курсовой работе рассмотрено решение задачи прогнозирования появления неопознанных летающих объектов в разных странах мира.

Целью работы является разработка системно-когнитивной модели, основывающейся на следующих признаках: тип объекта, длительность появления, место, дата.

Задачами курсовой работы являются:

- 1) систематизация, закрепление и расширение теоретических и практических знаний по дисциплине "Интеллектуальные информационные системы и технологии";
- 2) изучение интеллектуальной информационной системы "Эйдос";
- 3) решение поставленной цели с помощью интеллектуальной информационной системы "Эйдос".

Объектом исследования является база данных "MUP blagoustroistvo".

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 данной работе использована база данных " UFO Sightings ". База данных дополнена значениями и приводится полностью в приложении.

На момент написания курсовой работы этот банк включает в себя более 80 тысяч исходных данных по различным задачам, однако рассмотрена будет лишь часть усредненных данных. Решение поставленной задачи проведем в четыре этапа:

1. Подготовка данных 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 Подготовка данных в промежуточные файлы MS Excel

Из электронного ресурса баз данных рассматриваемого предприятия возьмем базу данных сотрудников – «MUP blagoustroistvo.xls», которую оставим без изменений.

Общее описание задачи:

1. Дата
2. Тип объекта
3. Длительность (секунды)
4. Город
5. Штат
6. Страна

Столбцы 4-6 – это описательные шкалы.

Столбцы 2-3 являются классификационными шкалами.

Обучающая выборка:

Date/time	Shape	duration (seconds)	City	State	Coutry
07.03.2010 21:45	other	300	austintown	oh	us
10/26/2007 05:00	oval	2700	otis	me	us
11.07.2008 20:15	?	4	palm springs	cam	us
10/22/2003 21:00	fireball	1800	abingdon	va	us
3/28/2005 22:30	unknown	120	melbourne	fl	us
10.03.2008 20:00	light	600	kodiak	ak	us
11/20/1996 21:30	circle	600	parma (??)	oh	us
8/17/2012 22:00	light	120	hermiston	o	us
4/19/2008 11:00	light	60	lake park	fl	us
8/25/1998 13:00	disk	10	sandpoint	id	us
06.11.2000 16:38	light	90	glenrothes (uk/scotland)	-	gb
07.10.1993 0:00	other	60	rabwah (pakistan)	-	us
7/19/1996 10:45	disk	900	northampton	pa3	us
8/14/2013 21:00	light	300	buckeye	az	us
09.12.2011 21:30	triangle	1	providence	ri	us
4/23/2013 00:00	light	7200	exeter	ne	us
2/15/1955 19:00	light	300	new york city (brooklyn)	ny	us
6/13/2009 21:30	fireball	20	marengo	il	us
11.05.1975 0:00	fireball	180	winslow	az	us
9/15/2001 21:30	oval	2400	beja (portugal)	-	us
4/30/2010 20:00	circle	120	broadview heights	oh	us
2/29/2012 23:15	triangle	0	ferrum	va	us
2/16/2004 20:35	unknown	10	commack	ny	us
10/18/2005 01:15	formation	15	hickory hills	il	us
03.06.2000 16:30	sphere	60	irvine	cam	us
06.01.1978 18:00	fireball	14400	fulton	mo	us
5/13/1991 22:00	disk	180	sechelt (canada)	bc	cam
1/17/2013 20:00	light	1500	winchester	va	us
05.05.2012 19:45	oval	300	berkeley	cam	us
1/25/2004 15:00	egg	180	hayward	cam	us
01.01.2014 0:05	circle	120	elon	nc	us
12.01.2001 21:10	light	120	mazatlán (mexico)	-	us
05.11.2003 9:30	disk	1800	laguna niguel	cam	us
7/15/1961 23:00	unknown	1800	mount hope	wv	us
3/16/2013 20:15	triangle	10	oswego	ny	us
05.01.2005 13:00	other	2	stevens pass	wa	us
12/28/2012 23:00	fireball	300	fort myers	fl	us
12/27/2013 17:30	light	120	marietta	mn	us
8/15/1995 17:00	teardrop	120	chatsworth (fort mountain community)	ga	us

06.01.2009 23:00	light	2700	hamilton	mi	us
05.12.2013 1:29	formation	420	joliet	il	us
8/25/2001 21:00	circle	90	lorane	o	us
6/15/1994 22:00	triangle	600	montgomery	al	us
6/18/2013 00:05	circle	0	chicago	il	us
9/13/2011 04:20	light	3600	bismarck	mo	us
10/22/2000 19:30	other	15	cooper city	fl	us
01.03.2003 3:00	other	300	norwood	ma	us
7/21/2007 11:00	triangle	6	whitby (canada)	on	cam
07.03.2008 16:00	diamond	30	el reno	ok	us
1/20/2011 17:30	unknown	60	tegucigalpa (honduras)	-	us
9/13/2006 21:00	sphere	5400	lumberton	tx	us
11/22/2012 19:05	light	180	fairbanks	ak	us
6/29/2000 00:10	circle	15	chico	cam	us
6/22/2010 20:45	fireball	420	east bridgewater	ma	us
09.07.1999 2:00	fireball	120	witchita falls/ft worth (between)	tx	us
4/24/2014 08:45	oval	10	rio rancho	nm	us
10.10.1980 19:00	sphere	180	houston	tx	us
7/19/2012 21:50	sphere	20	emerald isle	nc	us
06.01.1973 0:00	circle	300	sultan (near)	wa	us
10/23/2013 19:45	chevron	300	wilmington	ma	us
6/19/2006 16:10	sphere	3	spring hill	fl	us
7/13/2002 22:00	?	0		tx	us
12/30/2007 01:00	unknown	1800	perth amboy	nj	us
11/22/2012 18:36	fireball	600	hopkinton	ma	us
6/13/1947 20:18	unknown	600	oakmont	pa3	us
7/23/2006 22:30	light	10	kiel	wi	us
02.10.1999 20:40	sphere	1200	north pole	ak	us
12.12.2008 19:12	circle	0	in orbit	-	us
11/24/2004 00:00	changing	120	winter garden	fl	us
9/26/2008 17:30	changing	300	denver	co	us
09.01.2004 22:00	light	180	destin	fl	us
8/15/2002 22:15	unknown	1200	macomb	mi	us
11.05.2012 0:00	disk	900	spring creek	nv	us
7/23/2001 18:17	fireball	120	latrobe	pa3	us
1/25/2006 22:30	fireball	2	muncie	in	us
3/26/2011 21:30	unknown	600	san diego	cam	us
8/27/2012 21:27	unknown	80	saskatoon (canada)	sk	cam
07.04.2013 23:47	light	60	redding	cam	us
4/18/2005 13:20	triangle	240	los angeles	cam	us
10/26/2005 09:45	oval	900	tracy	cam	us
6/14/2010 18:10	circle	600	burbank	cam	us
09.07.2006 18:20	cylinder	1200	harbor city	cam	us
8/17/2012 23:00	circle	180	warren	mi	us
12/31/2002 13:00	other	30	burke	va	us
12.05.1996 22:00	?	30	seattle	wa	us
08.06.2011 0:14	triangle	5	canyon lake	cam	us
5/28/2010 22:12	unknown	45	green bay	wi	us

10/21/2006 19:30	unknown	120	brookings	o	us
11/14/2010 19:00	light	3600	safety harbor	fl	us
07.04.1977 14:00	sphere	60	marley park	md	us
10/18/2013 19:15	unknown	2700	pinecrest/sonora	cam	us
8/31/2004 21:00	other	1200	wakefield	nh	us
7/30/2009 21:00	sphere	60	bucks	pa	us
6/28/2001 23:00	light	2700	spearfish	sd	us
1/25/2002 20:00	triangle	120	bigelow	ar	us
11/20/2013 16:55	fireball	600	berlin	vt	us
12/31/2009 23:59	triangle	600	el cajon	cam	us
1/27/2009 03:30	cone	600	n. hollywood	cam	us
10/18/2005 23:00	other	10	orlando	fl	us

Таблица 1 – UFO Sightings.xls

1.3 Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос".

Для преобразования исходных данных обучающей выборки в базы данных системы "Эйдос" необходимо файл MicrosoftOfficeExcel, который содержит базу данных компьютерных процессоров скопировать в Aidos-X\AID_DATA\Inp_data и назвать Inp_data.xlsx.

Далее запускаем систему "Эйдос" из папки "Aidos-X" файлом _aidos-h.exe. Система попросит ввести логин и пароль. Необходимо ввести: логин – 1, пароль – 1. После откроется главное окно программы.

Для загрузки базы данных необходимо зайти в режим 2.3.2.2 и в настройках изменить (рисунок 1):

- "Задайте тип файла исходных данных Inp_data": "XLSX-MSExcels-2003";
- "Задайте диапазон столбцов классификационных шкал": "Начальный столбец классификационных шкал" – 2, "Конечный столбец классификационных шкал" – 3;
- "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 4, "Конечный столбец описательных шкал" – 6;

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

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) Стандарт DBF-файла

☐ DBF - DBASE IV (DBF/NTX) Стандарт CSV-файла

☐ CSV - Comma-Separated Values

Задайте параметры:

☒ Нули и пробелы считать ОТСУТСТВИЕМ данных

☐ Нули и пробелы считать ЗНАЧЕНИЯМИ данных

☒ Создавать БД средних по классам "Inp_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

☒ Формализации предметной области (на основе "Inp_data")

☐ Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

☒ Равные интервалы с разным числом наблюдений

☐ Разные интервалы с равным числом наблюдений

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

☒ Не применять сценарный метод АСК-анализа

☐ Применить сценарный метод АСК-анализа

☒ Применить спец. интерпретацию текстовых полей классов

☐ Применить спец. интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

В качестве классов рассматриваются:

☐ Значения полей целиком

☒ Элементы значений полей - слова

☐ Элементы значений полей - символы

☐ Проводить лемматизацию

☒ Не проводить лемматизацию

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 1 – Универсальный программный интерфейс импорта данных в систему "Эйдос"

После нажать кнопку "Ок". Далее открывается окно, где размещена информация о размерности модели (рисунок 2). В этом окне необходимо задать число интервалов равным 3 и нажать кнопку "Выйти на создание модели".

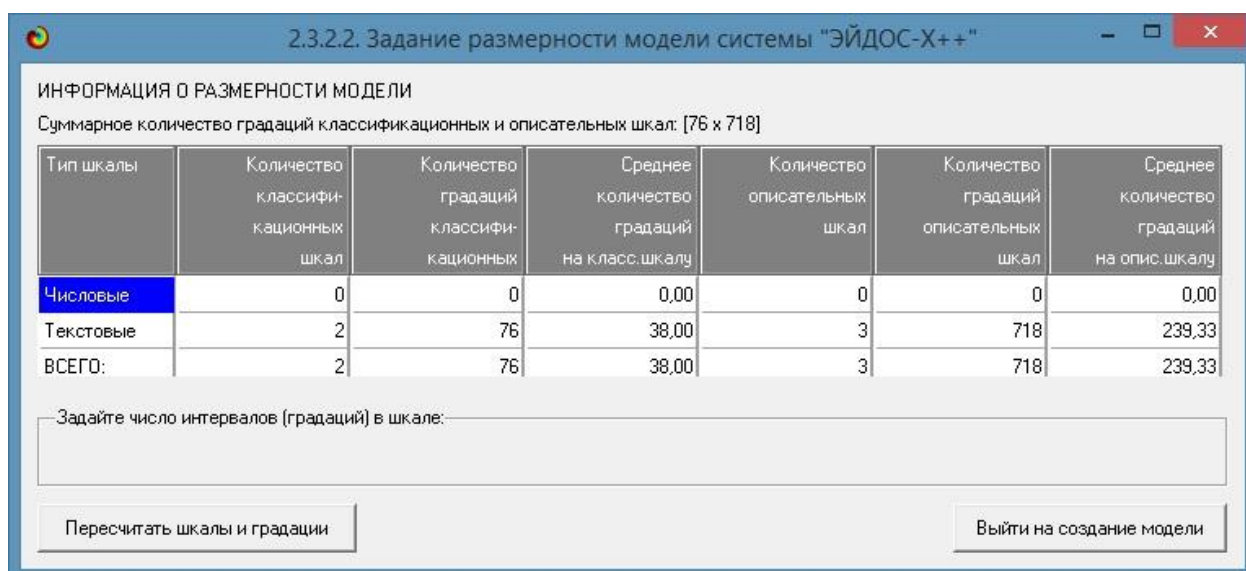


Рисунок 2 – Задание размерности модели системы "Эйдос"

Далее открывается окно, где происходит процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос" (рисунок 3). В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "Ок".

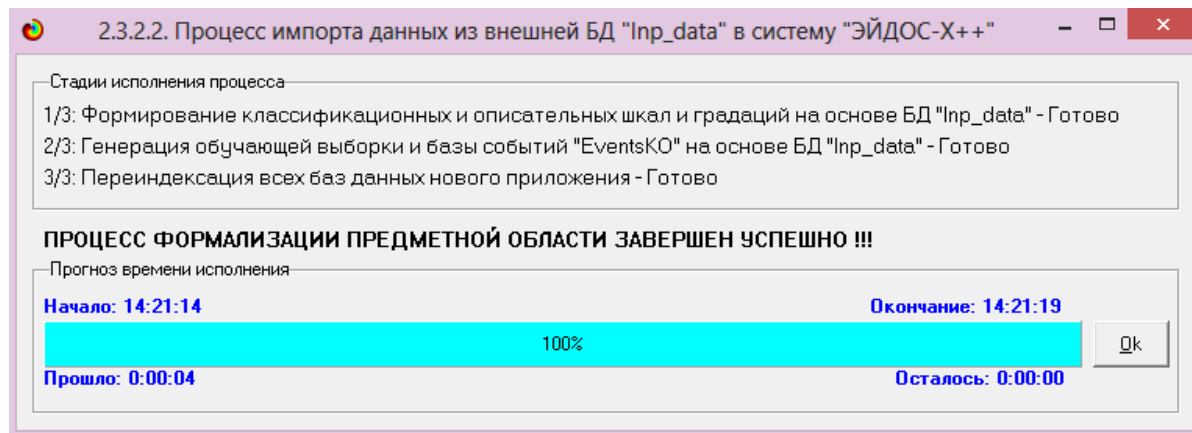


Рисунок 3 – Процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос"

В результате этого процесса будут сформированы классификационные и описательные шкалы и градации. Их применение позволит закодировать исходные данные и представить в форме эвентологических баз данных. Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1 (рисунок 4). Для просмотра описательных шкал и градаций

необходимо запустить режим 2.2 (рисунок 5). На рисунке 6 представлена обучающая выборка.

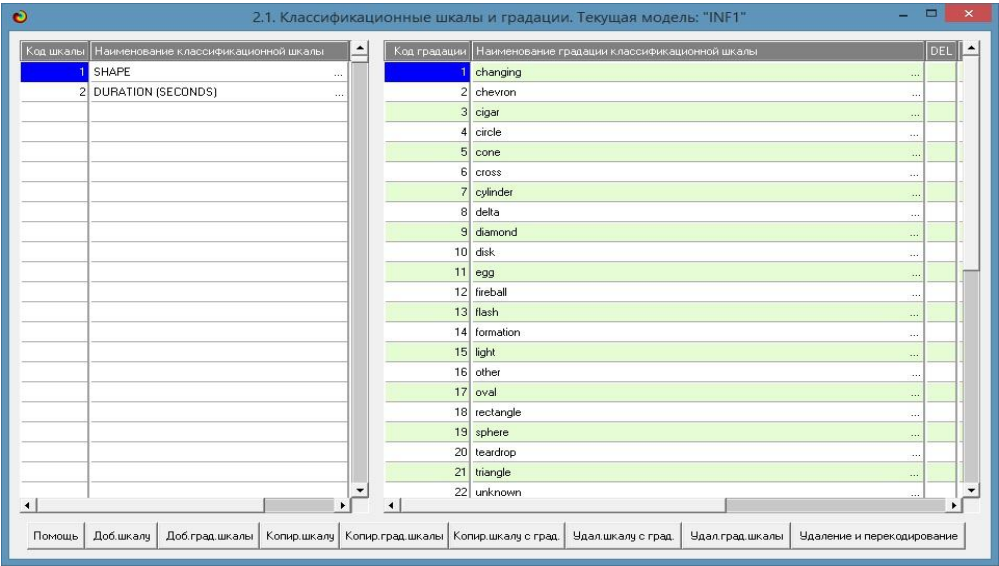


Рисунок 4 – Классификационные шкалы и градации

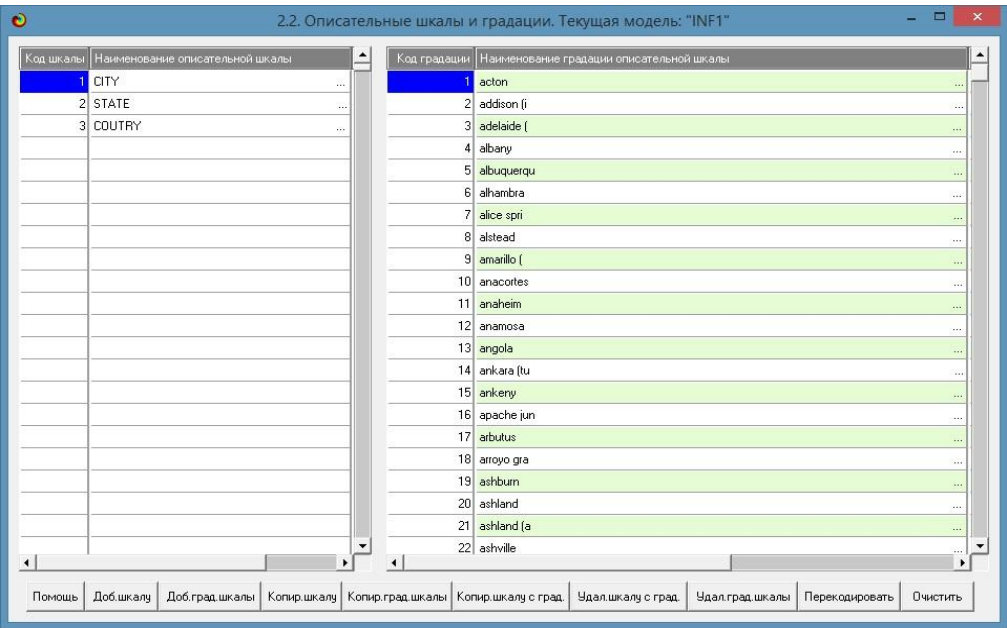


Рисунок 5 – Описательные шкалы и градации

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

Код объекта	Наименование объекта	Дата	Время
1	07.03.2010	...	...
2	10/26/2007	...	...
3	11.07.2008	...	...
4	10/22/2003	...	...
5	3/28/2005	...	...
6	10.03.2008	...	...
7	11/20/1996	...	...
8	8/17/2012	...	...
9	4/19/2008	...	...
10	0/25/1999	...	...

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	16	54	0	0

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	37	1000	1025	0	0	0	0

Помощь Скопировать обуч. выбо. расп. Добавить объект Добавить классы Добавить признаки Удалить объект Удалить классы Удалить признаки Очистить БД

Рисунок 6 – Обучающая выборка (фрагмент)

Таким образом, созданы 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).

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).

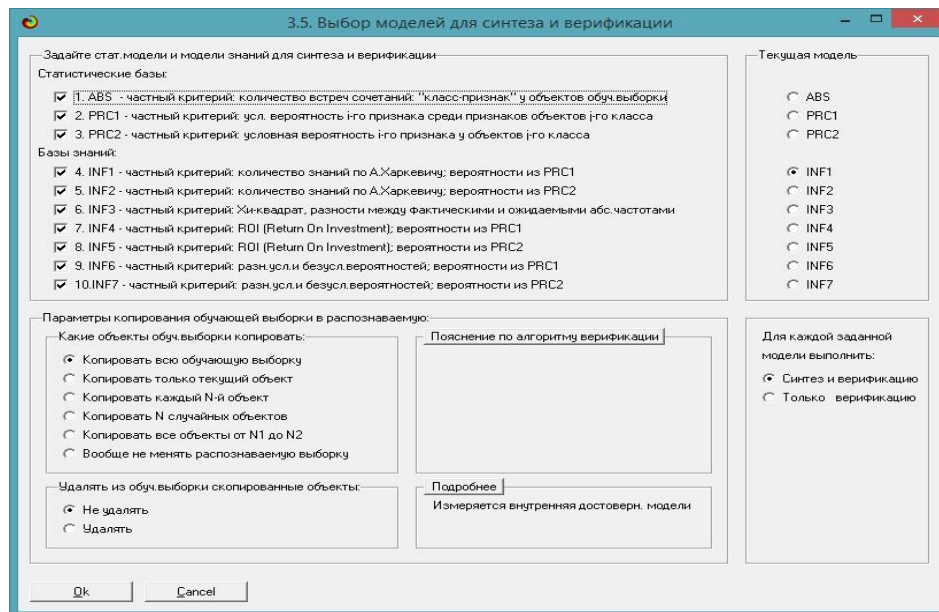


Рисунок 7 – Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

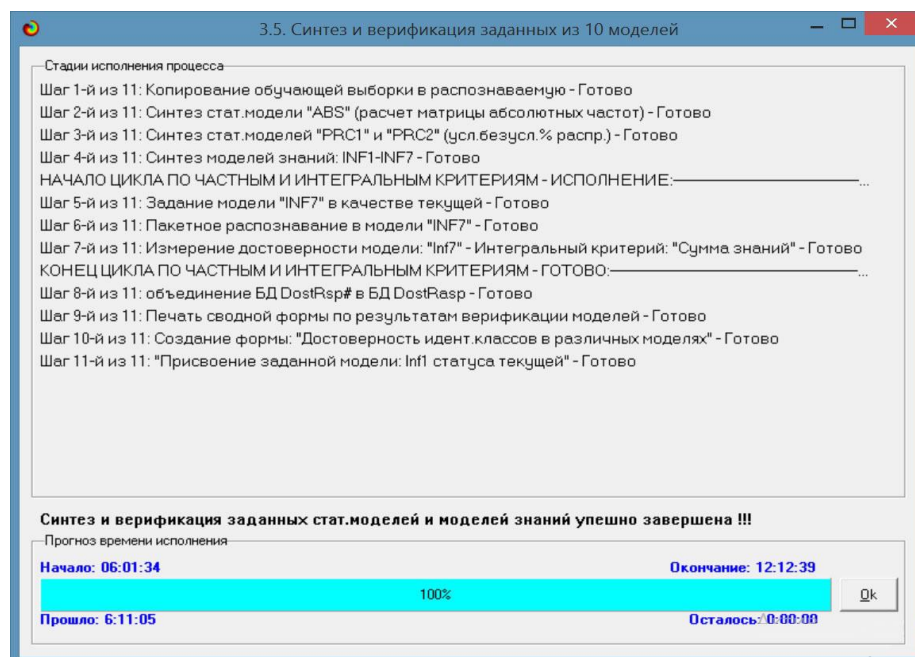


Рисунок 8 – Синтез и верификация статистических моделей и моделей знаний

Синтез и верификация всех моделей на данной задаче заняли 2,31 минут. При этом верификация (оценка достоверности моделей) проводилась на 1200 примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 8, но ниже 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. SHAPE ?	2. SHAPE CHANGING	3. SHAPE CHEVRON	4. SHAPE DIGAR	5. SHAPE CIRCLE	6. SHAPE CONE	7. SHAPE CROSS	8. SHAPE CYLINDER	9. SHAPE DIAMOND	10. SHAPE DISK	11. SHAPE EGG	12. SHAPE FIREBALL	13. SHAPE FLASH
133	QTY-chaco cary													
134	QTY-chadon													
135	QTY-champaign												1	
136	QTY-chandler	1				1				1				
137	QTY-charleston													
138	QTY-charlotte					1								1
139	QTY-chatsworth													
140	QTY-cherry val													
141	QTY-chicago					1								
142	QTY-chico					1								
143	QTY-chilton													
144	QTY-chinook	1												
145	QTY-chippewa f													
146	QTY-chugiak													
147	QTY-cincinnati													
148	QTY-circlevill													
149	QTY-clearlake													
150	QTY-cleburne	1												
151	QTY-cleveland								1	1			1	
152	QTY-clinton							1						
153	QTY-clive													
154	QTY-colchester	1												
155	QTY-colbranv													

Таблица 3 – Матрица информативностей (модель INF1) в битах (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	1. SHAPE ?	2. SHAPE CHANGING	3. SHAPE CHEVRON	4. SHAPE CIGAR	5. SHAPE CIRCLE	6. SHAPE CONE	7. SHAPE CROSS	8. SHAPE CYLINDER	9. SHAPE DIAMOND	10. SHAPE DISK	11. SHAPE EGG	12. SHAPE FIREBALL	13. SHAPE FLASH	14. SHAPE FO
164	QTY-conway														
165	QTY-coolidge	1.868													
166	QTY-cooper cit														
167	QTY-coral gabl	2.384													
168	QTY-cornith									2.259					
169	QTY-corner bro														
170	QTY-corning				2.709										
171	QTY-corpus chr												1.394		
172	QTY-cortland														
173	QTY-covallis												1.910		
174	QTY-covington					1.782									
175	QTY-creston														
176	QTY-crofton														
177	QTY-crow agenc														
178	QTY-culoden	2.384													
179	QTY-culpeper														
180	QTY-cumberland														
181	QTY-cunning				2.193									2.701	
182	QTY-dahlonga												1.910		
183	QTY-dallas										1.270				
184	QTY-dallas/ir										2.088				
185	QTY-dandridge										1.572				
186	QTY-darvers														

Таблица 4 – Матрица знаний (модель INF3) (фрагмент)

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсчастотами"															
Код признака	Наименование описательной шкалы и градации	1. SHAPE ?	2. SHAPE CHANGING	3. SHAPE CHEVRON	4. SHAPE CIGAR	5. SHAPE CIRCLE	6. SHAPE CONE	7. SHAPE CROSS	8. SHAPE CYLINDER	9. SHAPE DIAMOND	10. SHAPE DISK	11. SHAPE EGG	12. SHAPE FIREBALL	13. SHAPE FLASH	14. SHAPE FO
185	QTY-dandridge	-0.081	-0.055	-0.020	-0.053	-0.182	-0.003	-0.005	-0.035	-0.048	0.879	-0.030	-0.154	-0.027	
186	QTY-darvers	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
187	QTY-darville	-0.081	-0.055	-0.020	-0.053	-0.182	-0.003	-0.005	-0.035	-0.048	-0.121	-0.030	-0.154	-0.027	
188	QTY-dartmouth	-0.122	-0.082	-0.030	0.921	-0.274	-0.005	-0.007	-0.052	-0.072	-0.182	-0.045	-0.231	-0.040	
189	QTY-dayton	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	0.939	-0.015	-0.077	-0.013	
190	QTY-deerfield	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
191	QTY-delay bea	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	0.923	-0.013	
192	QTY-deltona	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
193	QTY-dennis	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
194	QTY-denton	-0.041	-0.027	-0.010	-0.026	0.909	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
195	QTY-denver	-0.163	0.891	-0.040	-0.105	-0.365	-0.007	-0.010	-0.070	-0.096	-0.242	-0.060	-0.307	-0.053	
196	QTY-destin	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
197	QTY-dewey beac	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
198	QTY-dexter	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	0.976	-0.061	-0.015	-0.077	-0.013	
199	QTY-diamond ba	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
200	QTY-dickson ci	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	0.983	-0.024	-0.061	-0.015	-0.077	-0.013	
201	QTY-dieks	-0.041	-0.027	-0.010	0.974	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
202	QTY-dom felici	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
203	QTY-donnacona	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	0.939	-0.015	-0.077	-0.013	
204	QTY-douglas	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
205	QTY-douglasvvi	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	
206	QTY-dover	-0.041	-0.027	-0.010	-0.026	-0.091	-0.002	-0.002	-0.017	0.976	-0.061	-0.015	-0.077	-0.013	
207	QTY-downey	-0.041	-0.027	-0.010	-0.026	0.909	-0.002	-0.002	-0.017	-0.024	-0.061	-0.015	-0.077	-0.013	

Таблица 5 – Матрица знаний (модель INF3) (фрагмент)

1.5. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей (режим 4.1.3.6), отличающихся частными критериями с двумя приведенными выше интегральными критериями приведены на рисунке 9.

4.1.3.6. Обобщенная форма по достов.моделей при разн.инт.крит. Текущая модель: "INF1"													
Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Гибрида Ван Ризбергера	Сумма модулей сходства истинно-положительных решений (ST)	Сумма модулей сходства истинно-отрицательных решений (ST)	Сумма модулей сходства ложно-положительных решений (SFP)	Сумма модулей сходства ложно-отрицательных решений (SFN)
1. ABS - частный критерий: количество встреч сочетаний "у.лас..."	Корреляция абс частот с обр...	2418	2418	9755	106309		0.022	1.000	0.044	1389.588	31.214	48964.453	
1. ABS - частный критерий: количество встреч сочетаний "у.лас..."	Сигма абс частот по признаку...	2418	2418	9755	106309		0.022	1.000	0.044	731.091		9432.879	
2. PRIC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн частот с о...	2418	2418	9755	106309		0.022	1.000	0.044	1389.589	31.214	48964.488	
2. PRIC1 - частный критерий: усл. вероятность его признака сред...	Сигма усл.отн частот по признаку...	2418	2418	9755	106309		0.022	1.000	0.044	811.206		34544.269	
3. PRIC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн частот с о...	2418	2418	9755	106309		0.022	1.000	0.044	1389.557	31.213	48963.719	
3. PRIC2 - частный критерий: условная вероятность его признака...	Сигма усл.отн частот по признаку...	2418	2418	9755	106309		0.022	1.000	0.044	808.912		34285.530	
4. INF1 - частный критерий: количество знаний по А/Ориентации, в...	Семантический резонанс зна...	2418	2329	62112	53952	89	0.041	0.963	0.079	236.107	762.062	1187.501	1.14
4. INF1 - частный критерий: количество знаний по А/Ориентации, в...	Сигма знаний	2418	2403	34635	81429	15	0.029	0.994	0.056	496.217	419.191	1848.439	0.38
5. INF2 - частный критерий: количество знаний по А/Ориентации, в...	Семантический резонанс зна...	2418	2329	62890	53174	89	0.042	0.963	0.080	236.045	762.146	1137.599	1.14
5. INF2 - частный критерий: количество знаний по А/Ориентации, в...	Сигма знаний	2418	2401	34134	81930	17	0.028	0.993	0.055	496.015	417.774	1820.837	0.39
6. INF3 - частный критерий: Хинкватер, разности между факти...	Семантический резонанс зна...	2418	1977	49385	66679	441	0.029	0.818	0.056	385.377	6136.837	5312.389	37.88
6. INF3 - частный критерий: Хинкватер, разности между факти...	Сигма знаний	2418	1977	49387	66677	441	0.029	0.818	0.056	375.564	3510.985	3190.195	54.77
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	2418	2266	108570	7494	152	0.232	0.937	0.372	262.426	895.265	166.817	2.22
7. INF4 - частный критерий: ROI (Return On Investment): веро...	Сигма знаний	2418	2407	34567	81497	11	0.029	0.995	0.056	62.410	3.386	74.232	0.00
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Семантический резонанс зна...	2418	2266	109687	6377	152	0.262	0.937	0.410	262.433	897.095	165.144	2.22
8. INF5 - частный критерий: ROI (Return On Investment): веро...	Сигма знаний	2418	2406	33984	82080	12	0.028	0.995	0.055	62.271	3.368	73.388	0.00
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей, вер...	Семантический резонанс зна...	2418	1921	37053	79011	497	0.024	0.794	0.046	360.708	3327.366	5978.172	45.11
9. INF6 - частный критерий: разн.усл.и безуслов.вероятностей, вер...	Сигма знаний	2418	1977	35599	80465	441	0.024	0.818	0.047	73.693	786.946	2506.495	3.14
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей, ве...	Семантический резонанс зна...	2418	1953	36668	79196	465	0.024	0.808	0.047	362.586	3275.341	5774.446	41.25
10. INF7 - частный критерий: разн.усл.и безуслов.вероятностей, ве...	Сигма знаний	2418	1999	34594	81470	419	0.024	0.827	0.047	73.031	779.970	2319.702	2.94
Помощь													

Рисунок 9 Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF4 и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели INF5 составляет 0,465. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется нечеткое мультиклассовое обобщение F-меры, предложенное проф.Е.В.Луценко (L-мера), при этом точность модели составляет 0,824 и 0,911 по L1 и L2-мерам. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%.

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 бесполезен, 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 всего 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 отнестись к классам, к 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}: $L1 = \frac{(TP + TN - FP - FN)}{(TP + TN + FP + FN)} \quad \text{(нормировка: \{-1, +1\})}$ $L2 = \frac{(1 + (TP + TN - FP - FN))}{(TP + TN + FP + FN)} \quad \text{(нормировка: \{0, 1\})}$ где: TP - истинно-положительное решение; TN - истинно-отрицательное решение; FP - ложно-положительное решение; FN - ложно-отрицательное решение; классическая F-мера достоверности моделей Ван Ризбергера: $Precision = TP / (TP + FP) \quad \text{— точность модели;}$ $Recall = TP / (TP + FN) \quad \text{— полнота модели;}$ $F\text{-мера} = 2 * (Precision * Recall) / (Precision + Recall).$ в АСК-анализе и системе "Эйдос" предлагается L-мера, представляющая собой нечеткое мультиклассовое обобщение классической F-меры достоверности моделей Ван Ризбергера: $SPrecision = STP / (STP + SFP) \quad \text{— нечеткая мультиклассовая точность модели;}$ $SRecall = STN / (STN + SFN) \quad \text{— нечеткая мультиклассовая полнота модели;}$ $L\text{-мера} = 2 * (SPrecision * SRecall) / (SPrecision + SRecall) \quad \text{— нечеткая мультиклассовая достоверность модели.}$ 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 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 INF5.

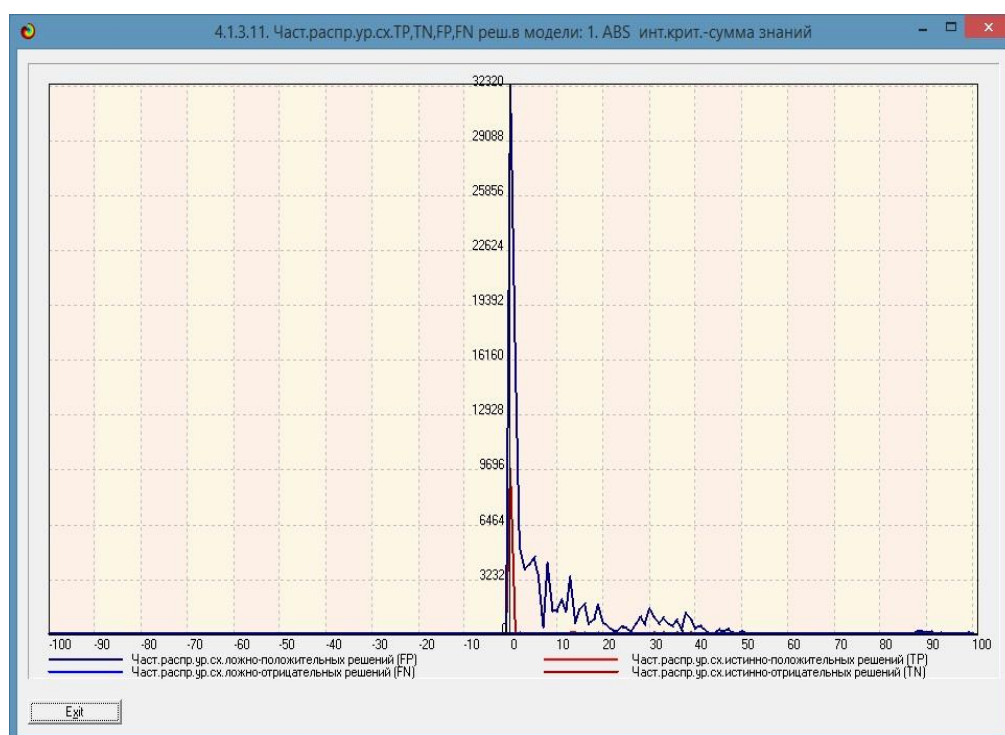


Рисунок 11. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF5

Из рисунка 11 видно, что:

- наиболее достоверная модель INF5 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 9);
- модуль уровня сходства-различия в наиболее достоверной модели INF5 для верно идентифицированных и верно неидентифицированных объектов 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 больших значений уровней сходства-различия (более 50%) различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями настолько велико, что учитывая уровень сходства-различия не представляет никакой проблемы разделить истинные положительные и отрицательные решения от ложных положительных и отрицательных решений. Это и стало основой для нечеткого обобщения F-меры Ван Ризбергена, т.е. такого ее обобщения, которое учитывает не только сам факт ложно или истинно положительного или отрицательного решения, но и его надежность.

Возьмем любые данные о наблюдениях как сумму истинного значения и шума, однако ни первое, ни второе неизвестны. Поэтому необходимо 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.

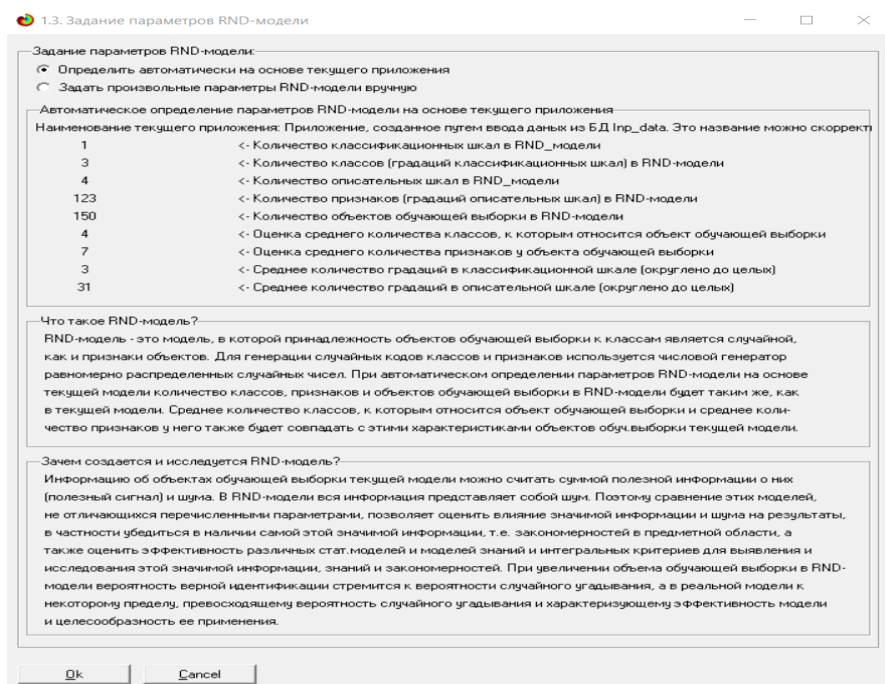


Рисунок 12. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 13 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF5.

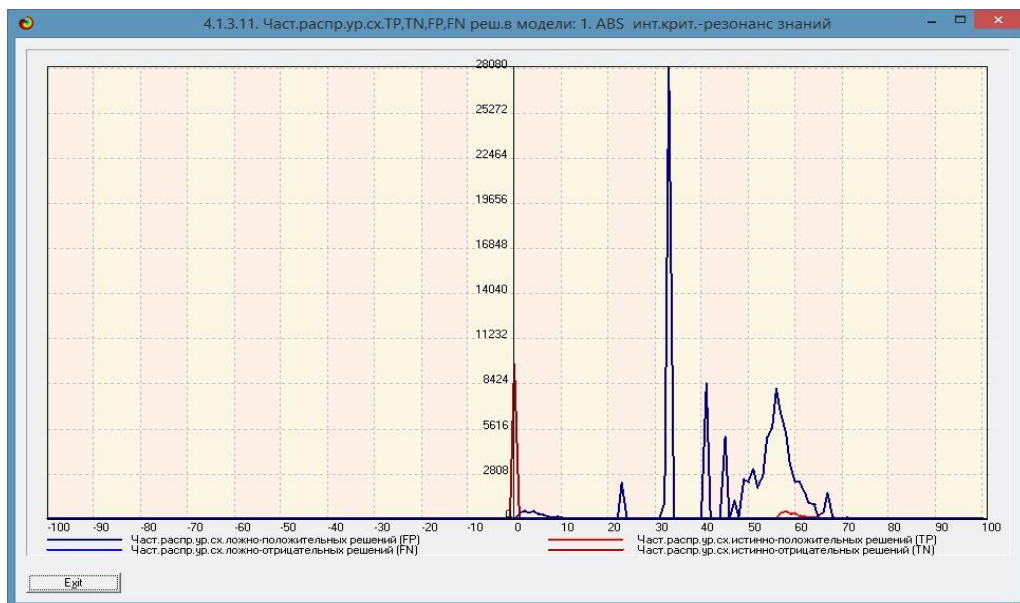


Рисунок 13. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF4

Различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели 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.

На рисунке 14 приведены данные по достоверности статистических и когнитивных моделей, созданных на основе случайной выборки.

4.1.3.6. Обобщенная форма по достов. моделей при разн. инт. крит. Текущая модель: "INF1"													
Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Риббергена	Сумма мод. уровней схо. истинно-полож. решений (ST)	Сумма мод. уровней схо. истинно-отриц. решений (ST)	Сумма мод. уровней схо. ложно-полож. решений (SFP)	Сумма мод. уровней схо. ложно-отриц. решений (SFP)
1. ABS - частный критерий: количество встреч сочетаний "улас"	Корреляция абс частот с абс частот по признакам...	2418	2418	9755	106309		0.022	1.000	0.044	1389.588	31.214	48964.453	
1. ABS - частный критерий: количество встреч сочетаний "улас"	Смена абс частот по признакам...	2418	2418	9755	106309		0.022	1.000	0.044	731.091		9432.879	
2. PRCT - частный критерий: усл. вероятность того признака сред.	Корреляция усл.отн частот с о...	2418	2418	9755	106309		0.022	1.000	0.044	1389.589	31.214	48964.488	
2. PRCT - частный критерий: усл. вероятность того признака сред.	Смена усл.отн частот по признакам...	2418	2418	9755	106309		0.022	1.000	0.044	811.206		34544.269	
3. PRCT - частный критерий: усл. вероятность того признака сред.	Корреляция усл.отн частот с о...	2418	2418	9755	106309		0.022	1.000	0.044	1389.557	31.213	48963.719	
3. PRCT - частный критерий: усл. вероятность того признака сред.	Смена усл.отн частот по признакам...	2418	2418	9755	106309		0.022	1.000	0.044	808.912		34285.530	
4. INF1 - частный критерий: количество знаний по А/Харкевичу; в.	Семантический резонанс зна...	2418	2329	62112	53952	89	0.041	0.963	0.079	236.107	762.062	1187.501	1.14
4. INF1 - частный критерий: количество знаний по А/Харкевичу; в.	Смена знаний	2418	2403	34635	81429	15	0.029	0.994	0.056	496.217	419.191	1848.439	0.38
5. INF2 - частный критерий: количество знаний по А/Харкевичу; в.	Семантический резонанс зна...	2418	2329	62890	53174	89	0.042	0.963	0.080	236.045	762.146	1137.599	1.14
5. INF2 - частный критерий: количество знаний по А/Харкевичу; в.	Смена знаний	2418	2401	34134	81930	17	0.028	0.993	0.055	496.015	417.774	1820.837	0.39
6. INF3 - частный критерий: Уинкватер, разности между фактич...	Семантический резонанс зна...	2418	1977	49385	66679	441	0.029	0.818	0.056	385.377	6136.837	5312.389	37.88
6. INF3 - частный критерий: Уинкватер, разности между фактич...	Смена знаний	2418	1977	49387	66677	441	0.029	0.818	0.056	375.564	3510.985	3190.195	54.77
7. INF4 - частный критерий: ROI (Return On Investment); веро...	Семантический резонанс зна...	2418	2266	108570	7494	152	0.232	0.937	0.372	262.426	895.265	166.817	2.22
7. INF4 - частный критерий: ROI (Return On Investment); веро...	Смена знаний	2418	2407	34567	81497	11	0.029	0.995	0.056	62.410	3.386	74.232	0.00
8. INF5 - частный критерий: ROI (Return On Investment); веро...	Семантический резонанс зна...	2418	2266	109687	6377	152	0.262	0.937	0.410	262.433	897.095	165.144	2.22
8. INF5 - частный критерий: ROI (Return On Investment); веро...	Смена знаний	2418	2406	33984	82080	12	0.028	0.995	0.055	62.271	3.368	73.388	0.00
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер.	Семантический резонанс зна...	2418	1921	37053	79011	497	0.024	0.794	0.046	360.708	3327.366	5978.172	45.11
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер.	Смена знаний	2418	1977	35599	80465	441	0.024	0.818	0.047	73.693	786.946	2506.495	3.14
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	2418	1953	36868	79196	465	0.024	0.808	0.047	362.586	3275.341	5774.446	41.25
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Смена знаний	2418	1999	34594	81470	419	0.024	0.827	0.047	73.031	779.970	2319.702	2.94
Поиск													

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 INF5, отражающей реальный объект моделирования, примерно на 17% 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.

2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи

В соответствии с технологией АСК-анализа зададим текущей модель INF4 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1 (рисунок 16).

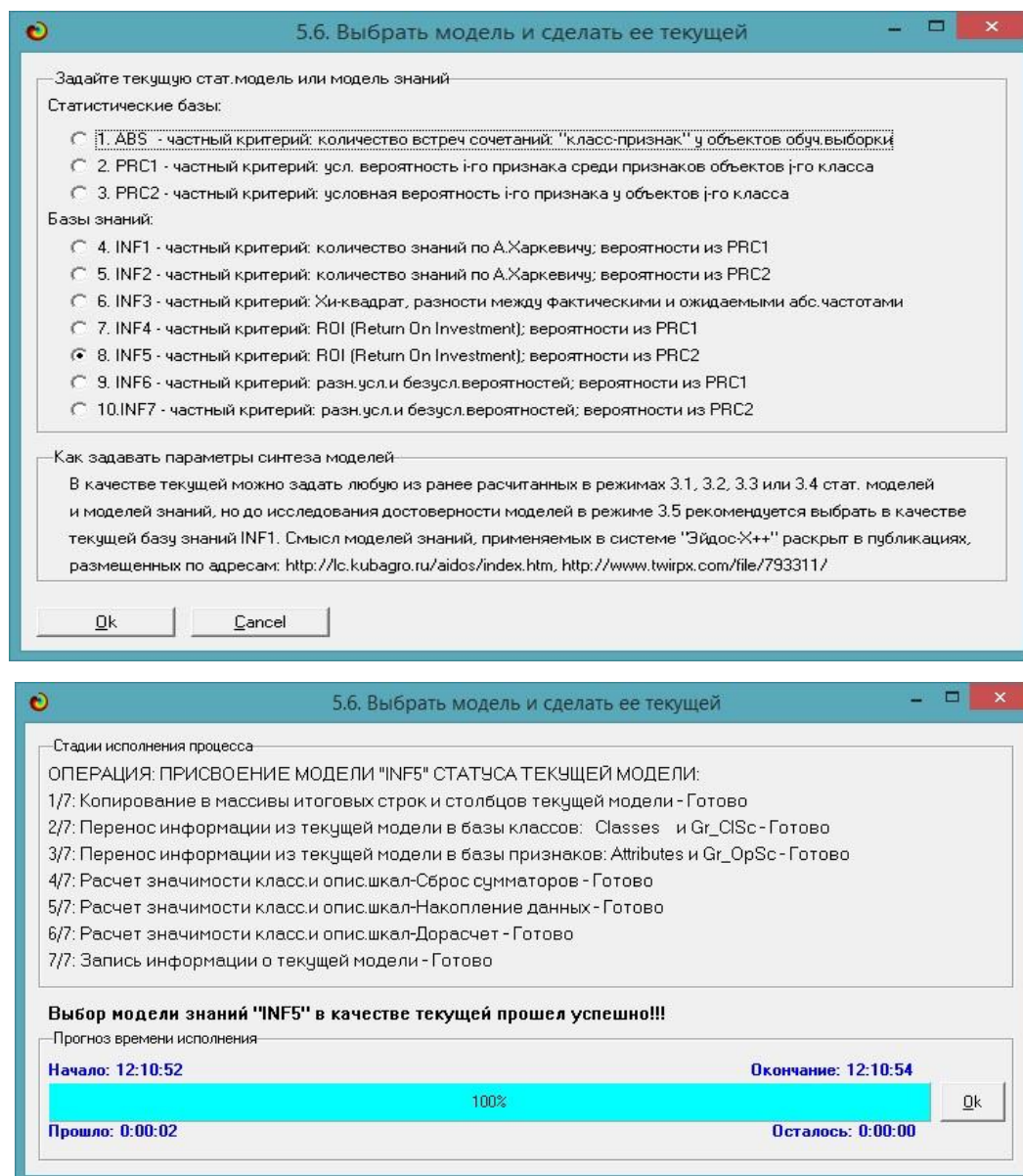


Рисунок 15. Экранные формы режима задания модели в качестве текущей

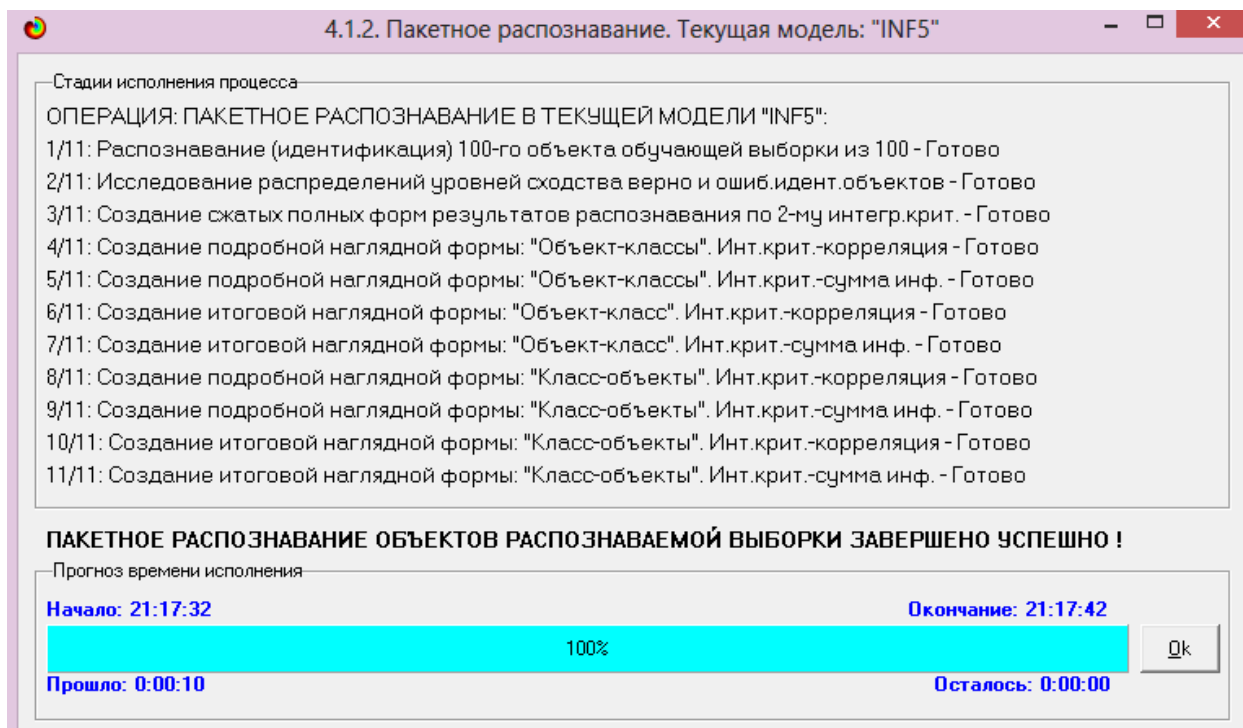


Рисунок 16. Экранная форма режима пакетного распознавания в текущей модели

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 по оплате услуг в зависимости от жилищных условий, района жилья и состава семьи в наиболее достоверной модели INF5 на основе наблюдения предыстории оплаты платежей:

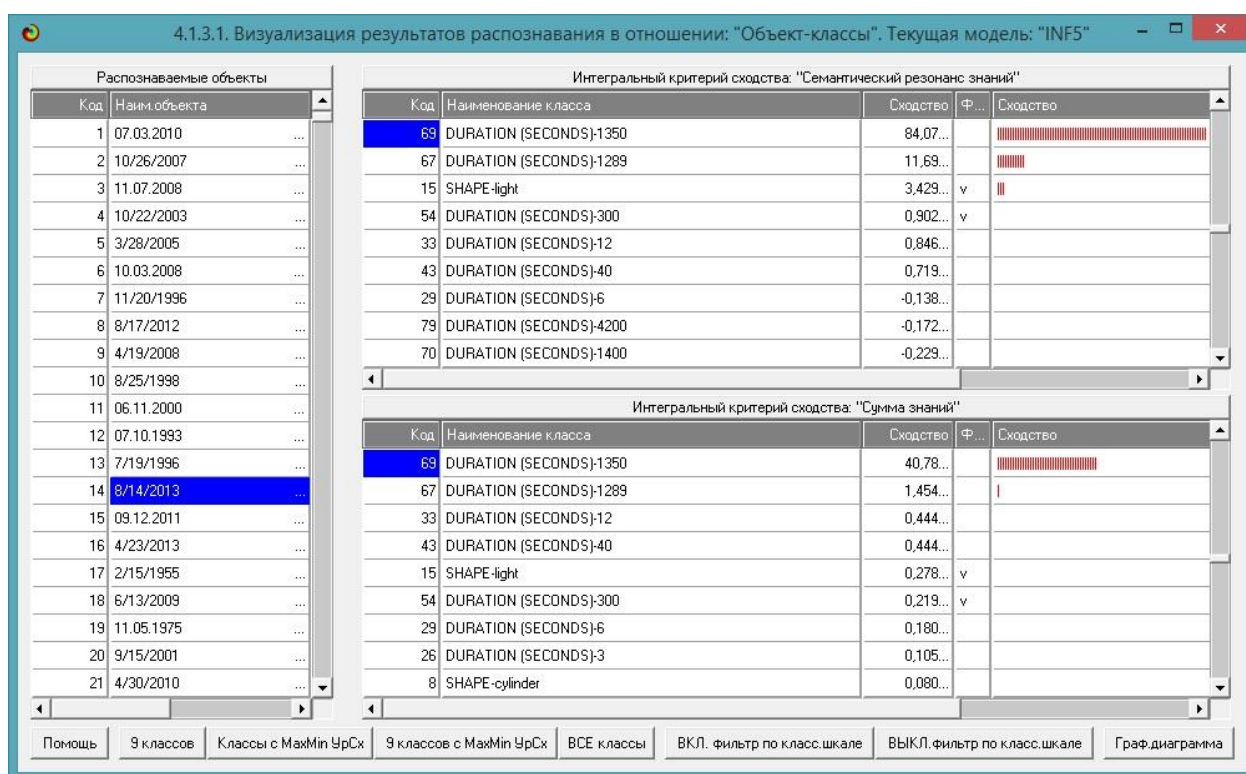


Рисунок 17. Пример идентификации дисциплины оплаты услуг в модели INF5

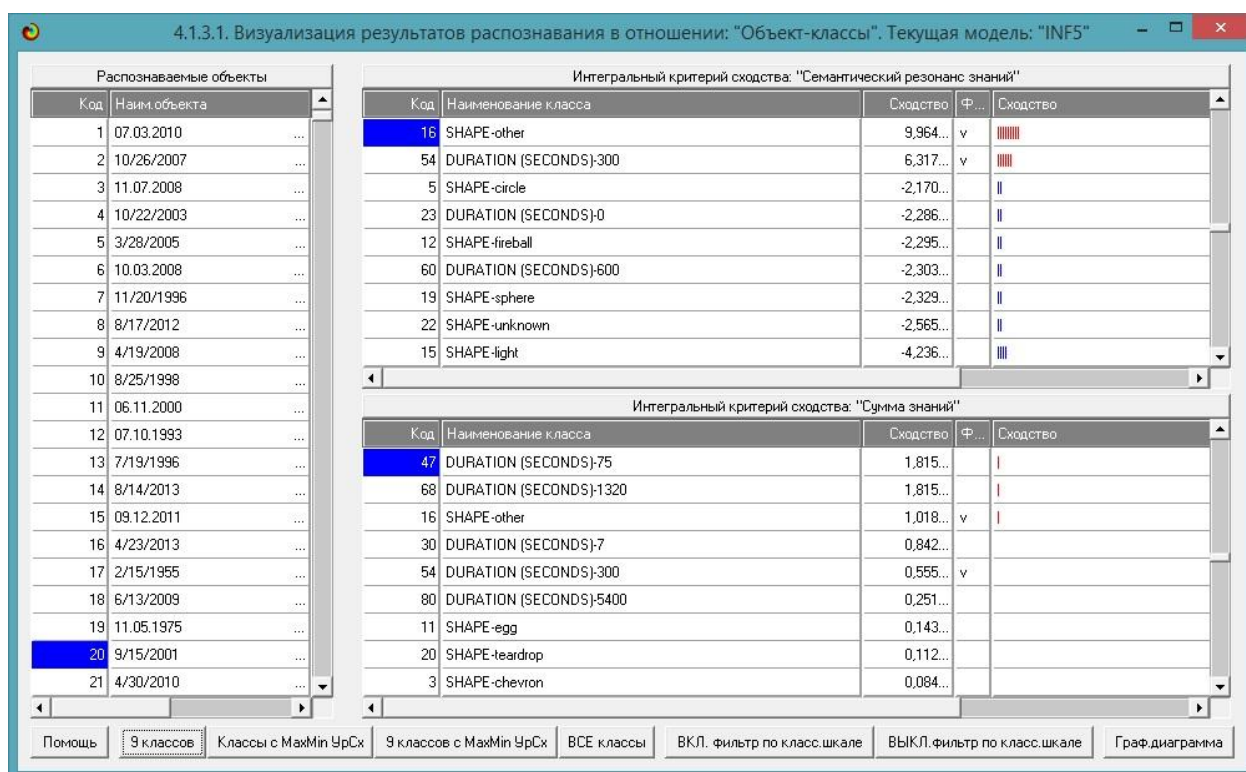


Рисунок 18. Пример идентификации дисциплины оплаты услуг в модели INF5

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.

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. Когнитивным функциям посвящено много работ, поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 20 приведены визуализации всех когнитивных функций данного приложения для модели INF5.

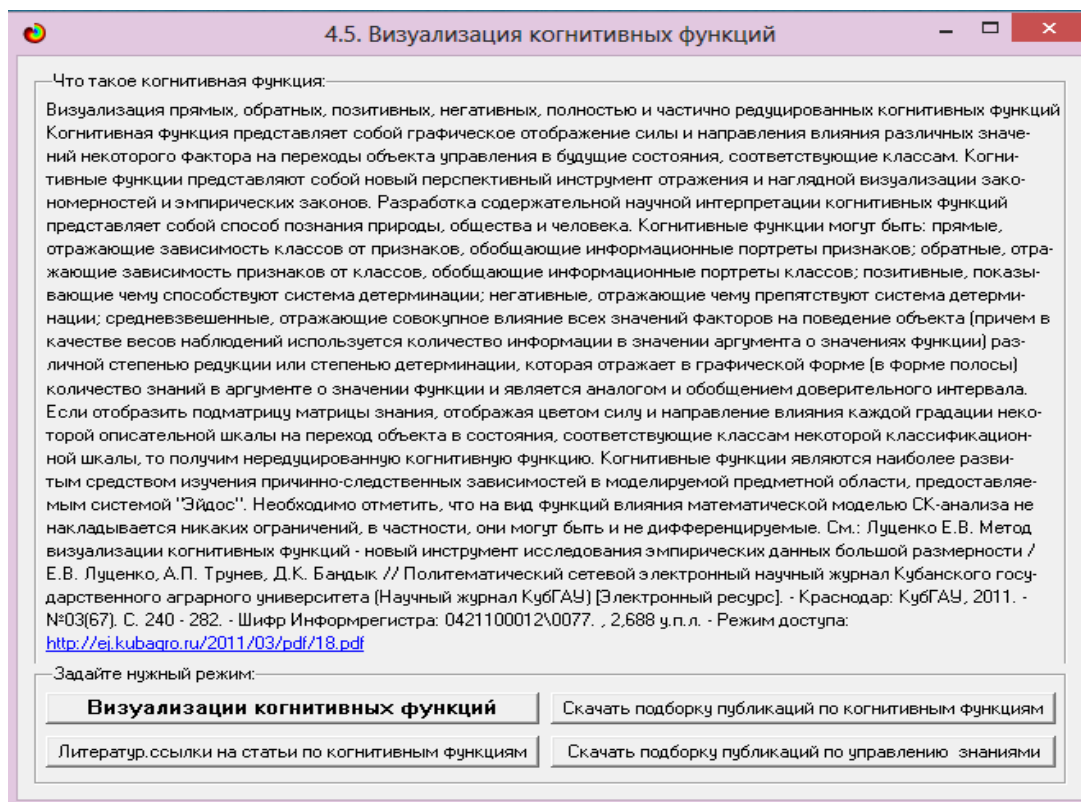
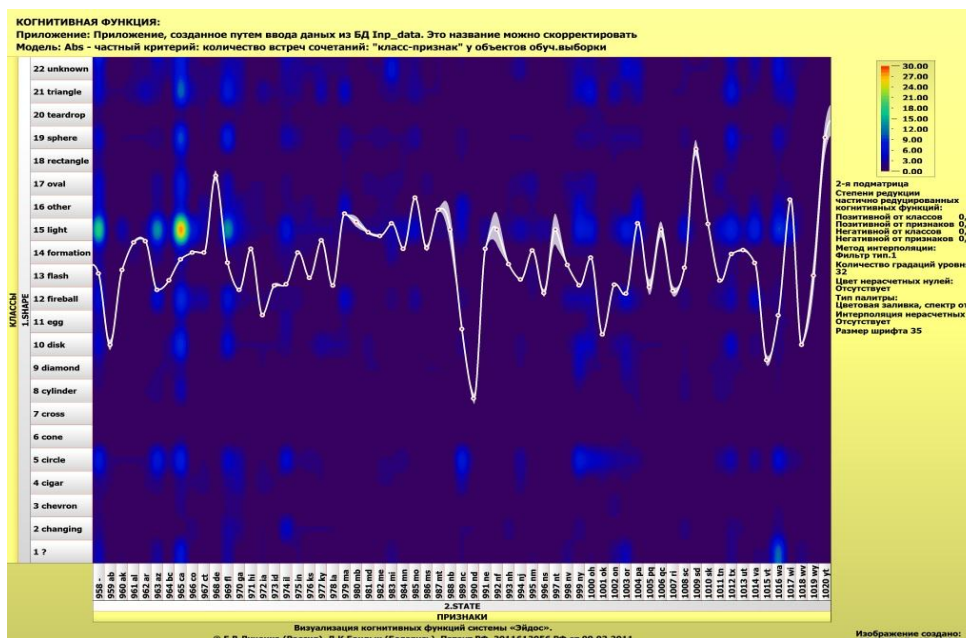
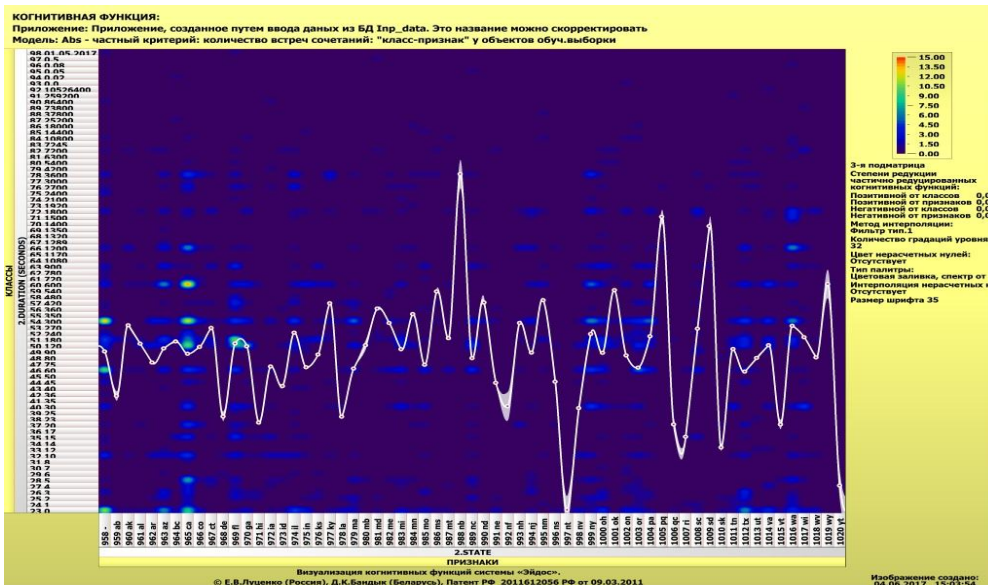
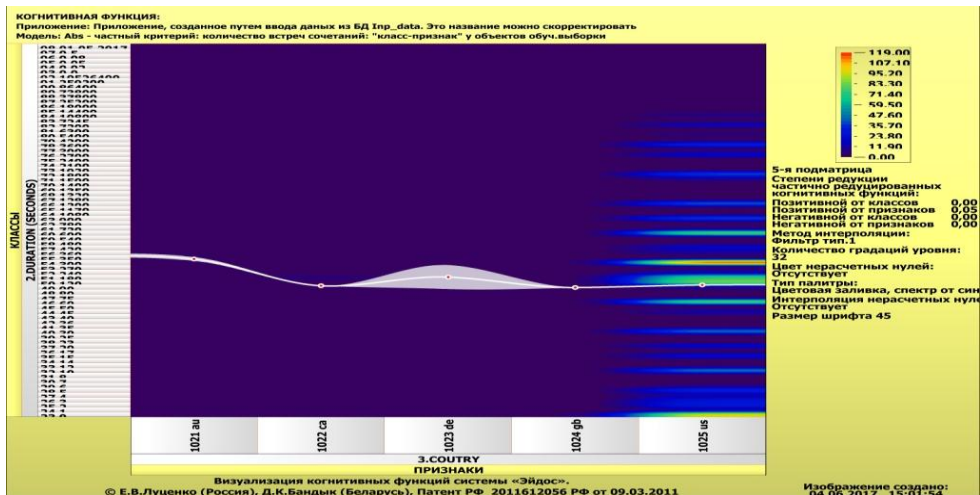
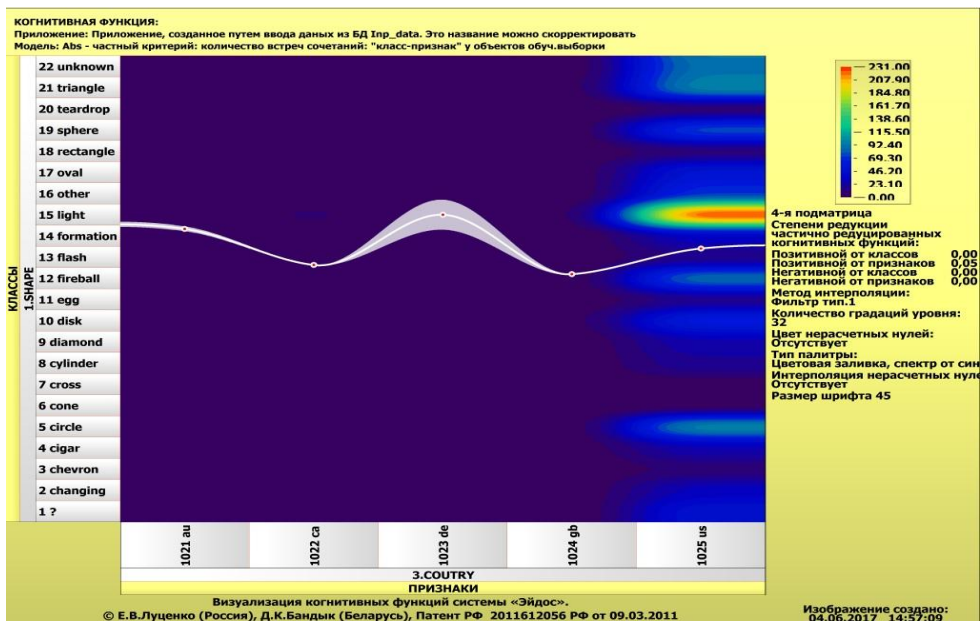


Рисунок 19. Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»





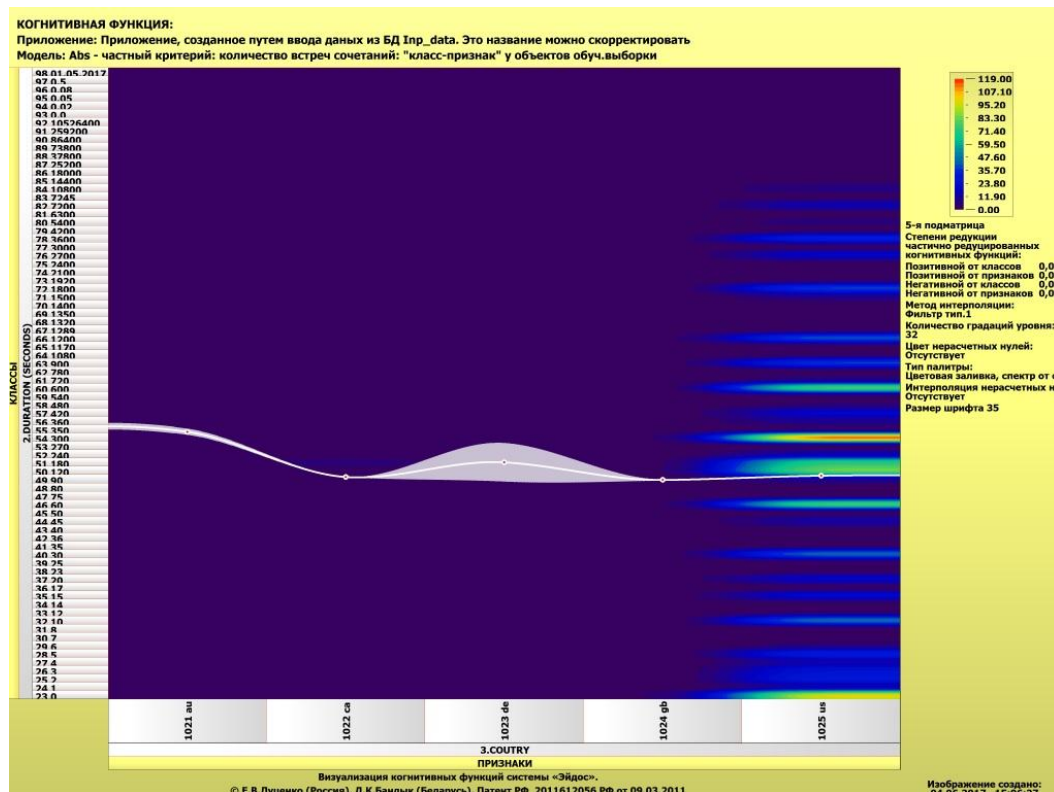
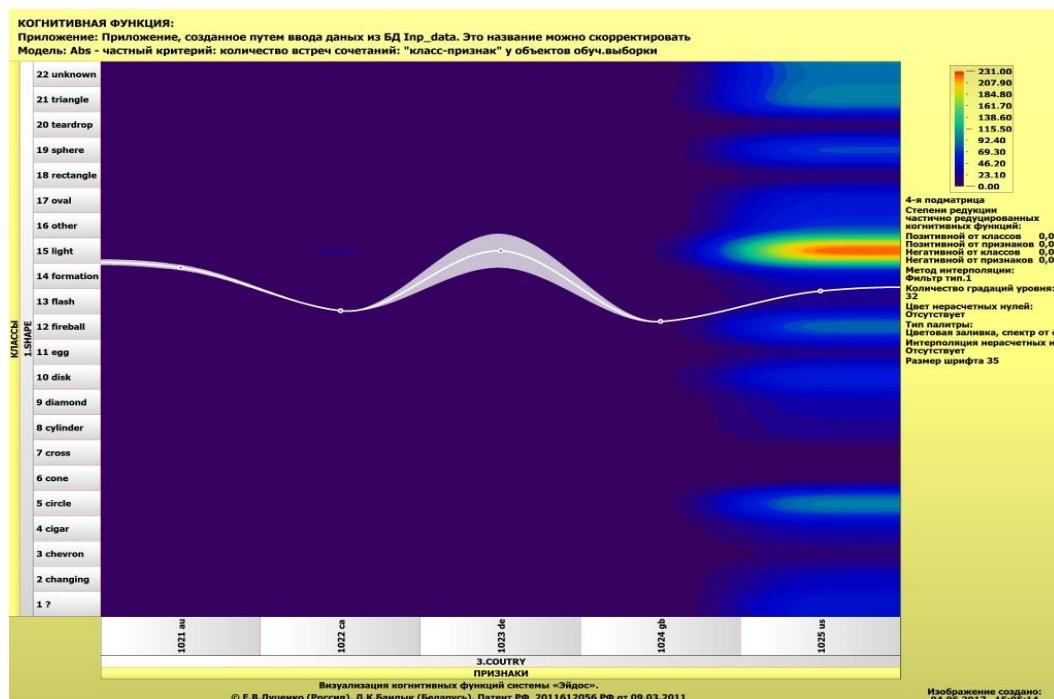


Рисунок 20. Визуализация когнитивных функций для обобщенных классов и описательных шкал для модели INF5

На основе полученных данных сделаем следующие прогнозы по дисциплине появление неопознанных летающих объектов в разных частях мира:

- 1). Наибольшее число появлений в Германии.
- 2). Самый частый появляющийся объект – свечение.
- 3). Среднее время появления 300 сек.

В выборке были представлены множество различных объектов. Визуализация показала, что наиболее частым объектом появления является свечение. Более подробные данные по улицам представлены на рисунке 20.

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 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 21).

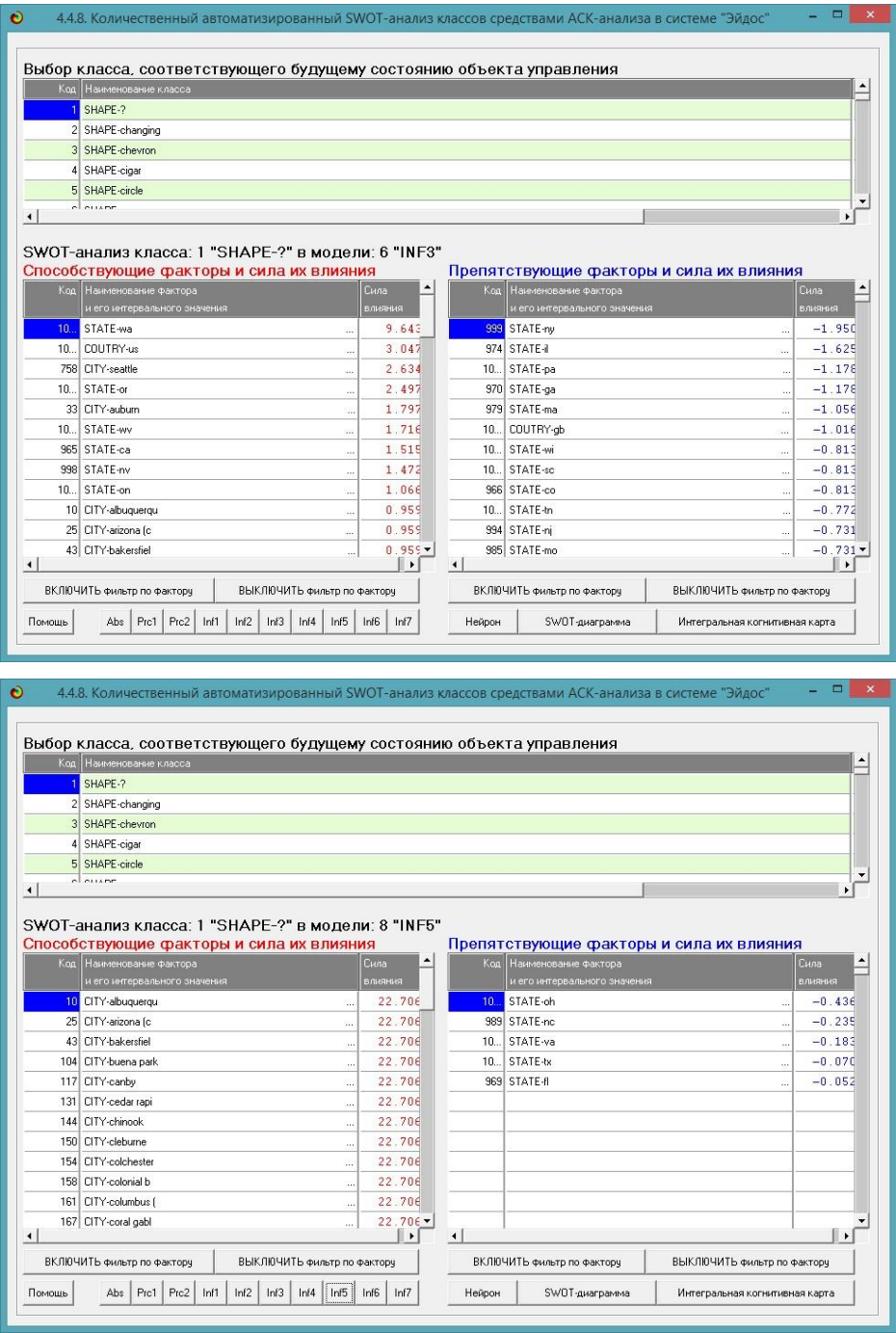


Рисунок 21. Пример SWOT-матрицы в модели INF5

На рисунке 22 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF3.

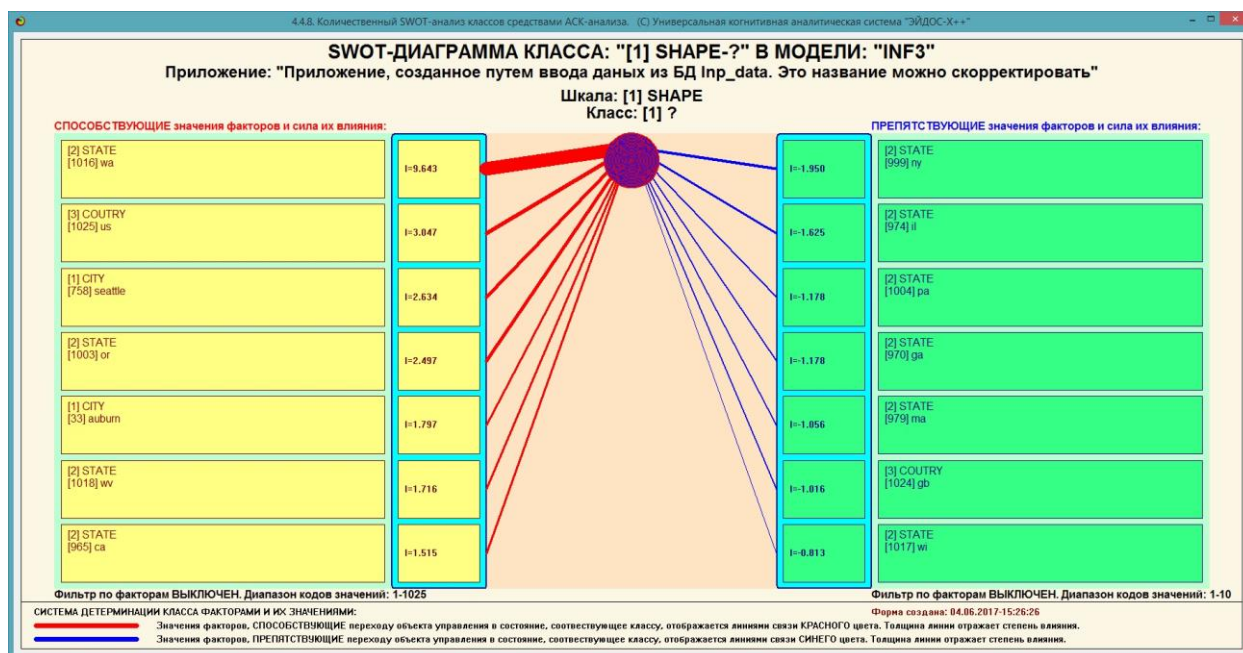


Рисунок 22. Пример SWOT-матрицы в модели INF3

2.4. Нелокальные нейроны и нейронная сеть

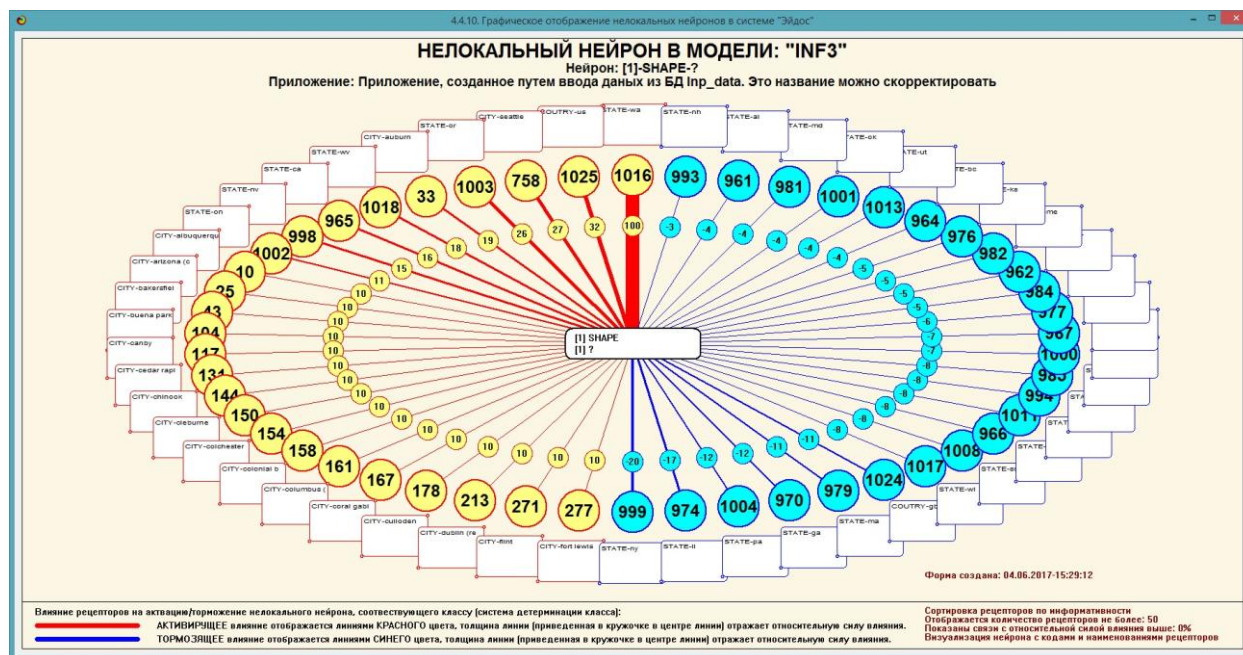


Рисунок 23. Пример графического отображения нелокальных нейронов

В качестве примера приведено графическое отображение нелокальных нейронов в модели INF3 на основе признака «Оплата-Большое». Каждому классу соответствует нейрон, совокупность которых образует нелокальную нейронную сеть [9].

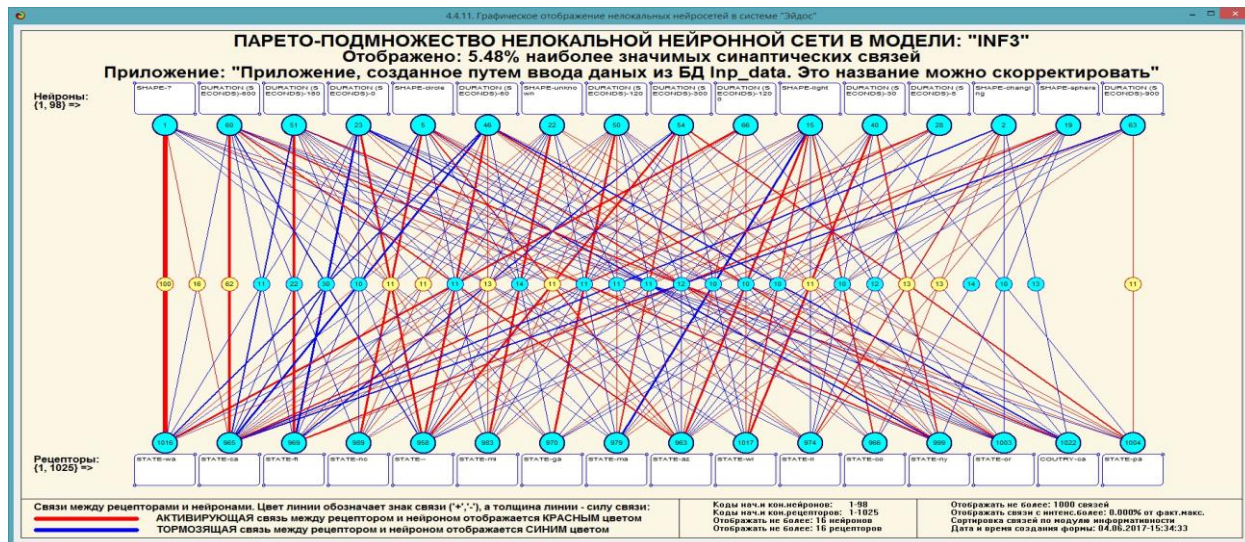


Рисунок 24. Пример графического отображения парето-подмножества нелокальной нейронной сети

На рисунке 24 в графическом отображения парето-подмножества нелокальной нейронной сети в модели INF3 у классов слева сильнее связи между значениями факторов и результатом [9].

Из рисунка 26 видно, что, например, 6 и 9 (переплата и оплата) признаки имеют сходство.

2.5. Кластерный и конструктивный анализ классов

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 1 "SHAPE-?" в модели: 4 "INF1"

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	SHAPE-?	1	1	SHAPE-?	100.000
2	SHAPE-changing	2	81	DURATION (SECONDS)-6300	18.345
3	SHAPE-chevron	3	93	DURATION (SECONDS)-0.0	14.804
4	SHAPE-cigar	4	83	DURATION (SECONDS)-7245	14.040
5	SHAPE-circle	5	23	DURATION (SECONDS)-0	13.954
6	SHAPE-cone	6	70	DURATION (SECONDS)-1400	13.262
7	SHAPE-cross	7	77	DURATION (SECONDS)-3000	10.090
8	SHAPE-cylinder	8	44	DURATION (SECONDS)-45	7.909
9	SHAPE-diamond	9	71	DURATION (SECONDS)-1500	7.593
10	SHAPE-disk	10	53	DURATION (SECONDS)-270	6.491
11	SHAPE-egg	11	43	DURATION (SECONDS)-40	6.267
12	SHAPE-fireball	12	30	DURATION (SECONDS)-7	5.473
13	SHAPE-flash	13	27	DURATION (SECONDS)-4	5.201
14	SHAPE-formation	14	63	DURATION (SECONDS)-900	4.913
15	SHAPE-light	15	98	DURATION (SECONDS)-01.05.2017	4.573
16	SHAPE-other	16	24	DURATION (SECONDS)-1	4.210
17	SHAPE-oval	17	96	DURATION (SECONDS)-0.08	2.831
18	SHAPE-rectangle	18	86	DURATION (SECONDS)-18000	2.826
19	SHAPE-sphere	19	26	DURATION (SECONDS)-3	2.493
20	SHAPE-teardrop	20	52	DURATION (SECONDS)-240	2.367
21	SHAPE-triangle	21	79	DURATION (SECONDS)-4200	2.303
22	SHAPE-unknown	22	56	DURATION (SECONDS)-360	2.027
23	DURATION (SECONDS)-0	23	76	DURATION (SECONDS)-2700	1.880

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

Рисунок 25. Результаты кластерного и конструктивного анализа классов

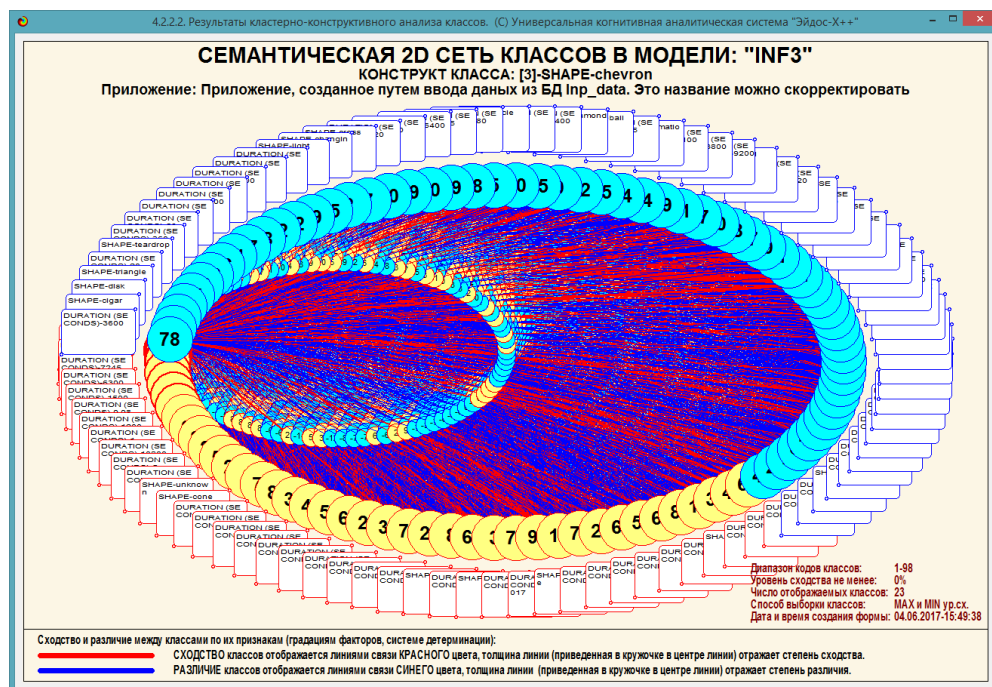


Рисунок 26. Графические результаты кластерного и конструктивного анализа классов

ЗАКЛЮЧЕНИЕ

На сегодняшний день все больше проявилась потребность в системах, не только представляющих информацию, но и выполняющих некоторый ее предварительный анализ, способных давать некоторые советы и рекомендации, осуществлять прогнозирование развития ситуаций, отбирать наиболее перспективные альтернативы решений, взять на себя значительную часть рутинных операций, а также функции предварительного анализа и оценок.

В данной курсовой работе была разработана системно-когнитивной модель прогноза дисциплины появления неопознанных летающих объектов в разных странах мира с помощью программного инструментария системы искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении оказались модели INF5, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,911, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием. Также обращает на себя внимание, что статистические модели в данном приложении дают примерно на 17% более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, что, как правило, наблюдается и в других приложениях. Этим и оправдано применение моделей знаний.

Обработка данных, их систематизация и визуализация позволили сделать прогнозы и обоснованные выводы по дисциплине неопознанные летающие объекты. Таким образом, повышены качество и оперативность принятия решений, а также эффективность управления знаниями.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. 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.
2. 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.
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. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://lc.kubagro.ru/>, свободный. - Загл. с экрана. Яз. рус.
5. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- Х++» / 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.

6. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 у.п.л.

7. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 164 – 188. – IDA [article ID]: 0911307012. – Режим доступа: <http://ej.kubagro.ru/2013/07/pdf/12.pdf>, 1,562 у.п.л.

8. Луценко Е.В. Нечеткое мультиклассовое обобщение классической F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе «Эйдос» / Луценко Е.В. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №09(123). – Режим доступа: <http://ej.kubagro.ru/2016/09/pdf/01.pdf>, 1,813 у.п.л. – IDA [article ID]: 1231609001. <http://dx.doi.org/10.21515/1990-4665-123-001>