

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

Факультет прикладной информатики

**РЕЦЕНЗИЯ
на курсовой проект (работу)**

Студента (ки) Гапона Дмитрия Сергеевича
курса 3 очной (заочной) формы обучения
направления подготовки Информационные системы и технологии
направленность(профиль)Интеллектуальные информационные системы и технологии.
Наименование темы «Исследование причин пропусков пациентами
запланированных встреч с врачами с использованием АСК-анализа и системы
Эйдос на основе данных репозитория UCI».

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	3
3.	Уровень и корректность использования в работе различных методов исследований	4
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	3
5.	Ясность, четкость, последовательность и обоснованность изложения	4
6.	Применение современных технологий обработки информации	3
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	4
8.	Ответы на вопросы при защите	4

Достоинства работы

Актуальность темы применения современных математических моделей и
программного инструментария для проведения АСК-анализа на основе базы
данных

Недостатки работы

Отдельные погрешности в оформлении

Итоговая оценка при защите 4 (хорошо)

Рецензент 

Е.В. Луценко

« » 201 г.

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 И
ТЕХНОЛОГИИ

на тему:

**ИССЛЕДОВАНИЕ ПРИЧИН ПРОПУСКОВ ПАЦИЕНТАМИ
ЗАПЛАНИРОВАННЫХ ВСТРЕЧ С ВРАЧАМИ С ИСПОЛЬЗОВАНИЕМ
АСК-АНАЛИЗА И СИСТЕМЫ ЭЙДОС НА ОСНОВЕ ДАННЫХ
РЕПОЗИТОРИЯ UCI**

выполнил студент группы ИТ1421 Гапон Дмитрий Сергеевич

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

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

Студенту: ИТ1421 группы 3 курса

Факультета прикладной информатики

Специальности: 09.03.02 Информационные системы и технологии
(шифр)

Гапону Дмитрию Сергеевичу

(Ф.И.О.)

Тема проекта: **Исследование причин пропусков пациентами запланированных встреч с врачами с использованием АСК-анализа и системы Эйдос на основе данных репозитория UCI**

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 _____ 36 _____ листа формата А4

б) графическая часть _____ 27 _____ 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

2017 г.

РЕФЕРАТ

36 страниц, 31 рисунок, 4 таблицы, 9 литературных источников

ЭЙДОС X, СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ГРАДАЦИИ,
КЛАССИФИКАЦИОННЫЕ ШКАЛЫ, ОПИСАТЕЛЬНЫЕ ШКАЛЫ

Целью работы является исследование причин пропусков пациентами запланированных встреч с врачами с использованием АСК-анализа и системы Эйдос на основе данных репозитория UCI.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	7
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	8
1.1. Описание решения.....	8
1.2. Преобразование исходных данных из HTML-формата в файл исходных данных MS Excel	8
1.3. Синтез и верификация статистических и интеллектуальных моделей .	14
1.4. Виды моделей системы «Эйдос»	16
1.5. Результаты верификации моделей	18
2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	25
2.1. Решение задачи	25
2.2. Когнитивные функции	28
ЗАКЛЮЧЕНИЕ	37
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	38

ВВЕДЕНИЕ

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи идентификации посещения/непосещения пациентами больниц.

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 репозитория UCI.

В данной курсовой работе использована база данных «no show medicine» из банка исходных данных по задачам искусственного интеллекта – репозитория UCI.

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 allexcel возьмем базу данных пациентов – «no show medicine», которую оставим без изменений.

Общее описание задачи:

1. Status-статус посещения
2. Age-возраст пациента
3. DayOfTheWeek-день недели
4. AppointmentRegistration-дата регистрации
5. ApointmentData-дата приема
6. Alcoolism-индекс алкоголизма
7. HiperTension-гипертония
8. Handcap-дефекты
9. Smokes – курение
- 10.Scholarship – степень заболевания
- 11.Sms_Reminder – смс – оповещение

12. AwaitingTime – время ожидания

Столбцы 1-2 классификационные шкалы. Так как классификация производится по статусу посещения (посетил, не посетил) и по возрасту посетителя.

Столбец 3 - 11 является описательными шкалой..

Обучающая выборка:

Таблица 1 – no show medicine.xls

nomer	Status	Age	DayOfTheWeek	AppointmentRegistration	ApointmentData	Alcoholism	HiperTensii	Handcap	Smokes	Scholarshij	Sms_Reminder	AwaitingTime
1	Show-Up	19	Wednesday	2014-12-16T14:46:25Z	2015-01-14T00:00:00Z	0	0	0	0	0	0	-29
2	Show-Up	24	Wednesday	2015-08-18T07:01:26Z	2015-08-19T00:00:00Z	0	0	0	0	0	0	-1
3	Show-Up	4	Tuesday	2014-02-17T12:53:46Z	2014-02-18T00:00:00Z	0	0	0	0	0	0	-1
4	Show-Up	5	Thursday	2014-07-23T17:02:11Z	2014-08-07T00:00:00Z	0	0	0	0	0	1	-15
5	Show-Up	38	Tuesday	2015-10-21T15:20:09Z	2015-10-27T00:00:00Z	0	0	0	0	0	1	-6
6	No-Show	5	Tuesday	2014-06-17T06:47:27Z	2014-07-22T00:00:00Z	0	0	0	0	0	1	-35
7	Show-Up	46	Tuesday	2015-06-19T19:53:58Z	2015-07-07T00:00:00Z	0	0	0	0	0	1	-18
8	Show-Up	4	Friday	2015-08-07T08:19:19Z	2015-08-21T00:00:00Z	0	0	0	0	1	1	-14
9	Show-Up	20	Tuesday	2014-10-14T11:46:36Z	2014-10-28T00:00:00Z	0	0	0	0	0	0	-14
10	Show-Up	51	Tuesday	2014-06-13T07:50:43Z	2014-06-17T00:00:00Z	0	1	0	0	0	1	-4
11	Show-Up	33	Monday	2015-10-01T08:43:18Z	2015-11-09T00:00:00Z	0	0	0	0	0	0	-39
12	Show-Up	58	Thursday	2014-04-30T08:04:27Z	2014-05-22T00:00:00Z	0	1	0	0	1	1	-22
13	No-Show	62	Monday	2014-07-25T15:02:33Z	2014-08-11T00:00:00Z	0	1	0	0	0	0	-17
14	Show-Up	62	Monday	2014-07-08T12:56:25Z	2014-07-14T00:00:00Z	0	0	0	0	0	1	-6
15	No-Show	38	Thursday	2014-03-21T08:23:16Z	2014-03-27T00:00:00Z	0	0	0	0	0	1	-6
16	Show-Up	73	Thursday	2015-04-01T16:27:05Z	2015-04-02T00:00:00Z	0	1	0	0	0	0	-1
17	Show-Up	48	Monday	2015-03-07T10:45:25Z	2015-03-30T00:00:00Z	0	0	0	0	0	1	-23
18	No-Show	52	Thursday	2014-06-03T10:20:37Z	2014-06-05T00:00:00Z	0	0	0	0	0	0	-2
19	Show-Up	64	Monday	2014-02-17T14:32:34Z	2014-02-24T00:00:00Z	0	1	0	0	0	1	-7
20	Show-Up	1	Wednesday	2014-04-01T13:39:00Z	2014-04-02T00:00:00Z	0	0	0	0	0	0	-1
21	Show-Up	33	Saturday	2014-03-12T10:26:26Z	2014-03-15T00:00:00Z	0	0	0	0	0	0	-3
22	Show-Up	44	Friday	2015-04-08T12:31:28Z	2015-04-10T00:00:00Z	0	0	0	0	0	0	-2
23	Show-Up	5	Thursday	2015-11-11T09:44:38Z	2015-11-12T00:00:00Z	0	0	0	0	0	0	-1
24	Show-Up	69	Friday	2014-11-24T08:25:35Z	2014-11-28T00:00:00Z	0	1	0	0	0	0	-4
25	No-Show	16	Tuesday	2014-10-23T16:49:42Z	2014-10-28T00:00:00Z	0	0	1	0	0	0	-5
26	No-Show	8	Monday	2015-05-11T14:56:01Z	2015-05-25T00:00:00Z	0	0	0	0	0	1	-14
27	Show-Up	7	Thursday	2015-04-01T08:45:19Z	2015-04-16T00:00:00Z	0	0	0	0	0	0	-15
28	Show-Up	51	Thursday	2014-07-29T10:33:20Z	2014-07-31T00:00:00Z	0	0	0	0	0	0	-2

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 файле Inp_data.xls добавим пустую колонку на позиции «А» и автоматически пронумеруем все строки. 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.

The screenshot displays a complex configuration window for data import. It is organized into several sections:

- File Type Selection:** Radio buttons for XLS (MS Excel-2003), XLSX (MS Excel-2007(2010)), DBF (DBASE IV), and CSV (Comma-Separated Values). Each has a corresponding 'Стандарт' button.
- Parameters:** Radio buttons for handling zeros and blanks (as missing or as values), and a checked checkbox for creating a database for averages. A 'Требования к файлу' button is also present.
- Column Ranges:** Input fields for starting and ending columns for classification and descriptive scales.
- Mode:** Radio buttons for formalization of the subject area (selected) or generation of a recognizable selection.
- Interval Selection:** Radio buttons for equal intervals with varying observation counts (selected) or different intervals with equal observation counts.
- Scenario Parameters:** Checkboxes for applying the ASC analysis method and special text interpretation for classes and signs.
- Text Field Interpretation:** Two boxes for defining how text fields in classification and descriptive scales are interpreted (e.g., as integers).
- Scale Naming:** Radio buttons for using only numerical values, only names, or both, with example strings provided.

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- 2003";

– "Задайте диапазон столбцов классификационных шкал": "Начальный столбец классификационных шкал" – 2, "Конечный столбец классификационных шкал" – 3(второй столбец в таблице);

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 4 "Конечный столбец описательных шкал" – 13;

– "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".

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	1	6	6,00	5	75	15,00
Текстовые	0	0	0,00	3	19	6,33
ВСЕГО:	1	6	6,00	8	94	11,75

Задайте число интервалов (градаций) в шкале:

В классификационных шкалах: В описательных шкалах:

Пересчитать шкалы и градации Параметры числ. шкал и градаций 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".

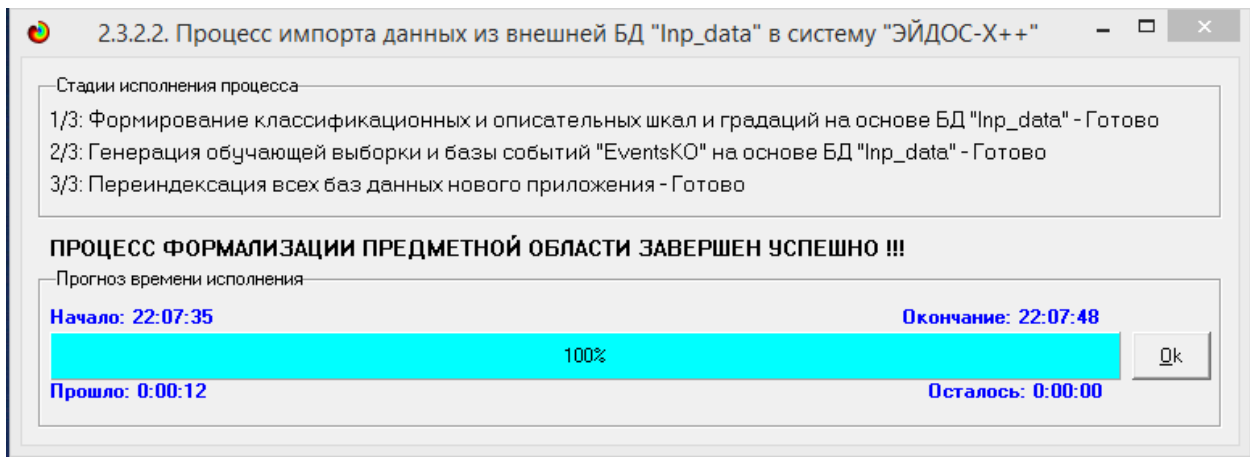


Рисунок 3. Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

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).

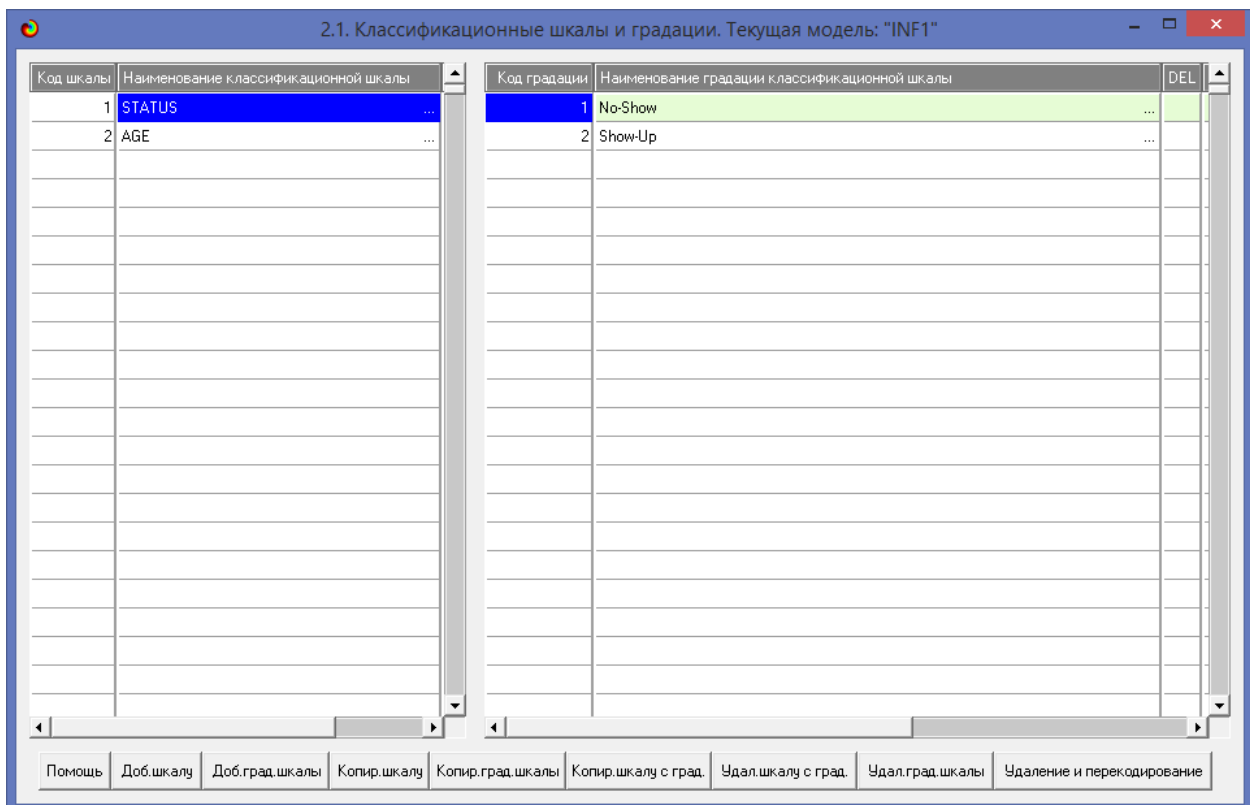


Рисунок 4. Классификационные шкалы и градации (фрагмент)

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

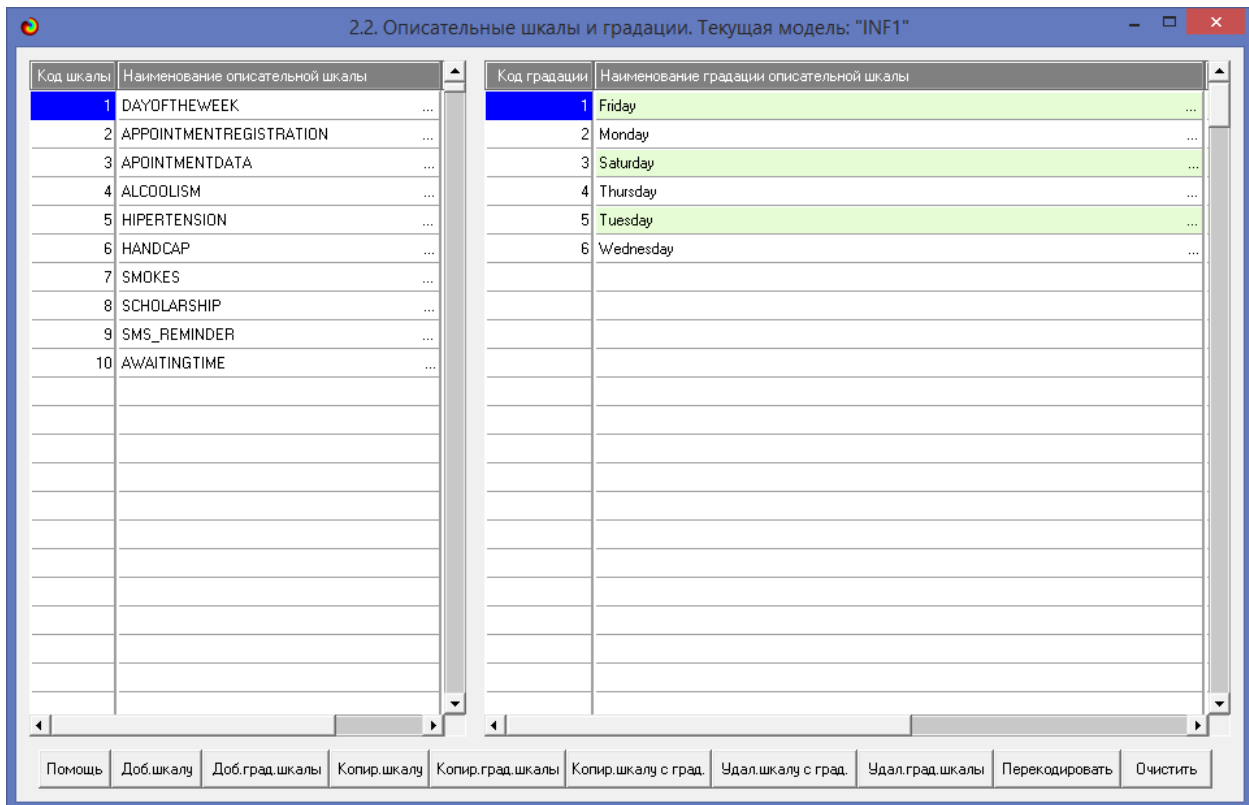


Рисунок 5. Описательные шкалы и градации (фрагмент)

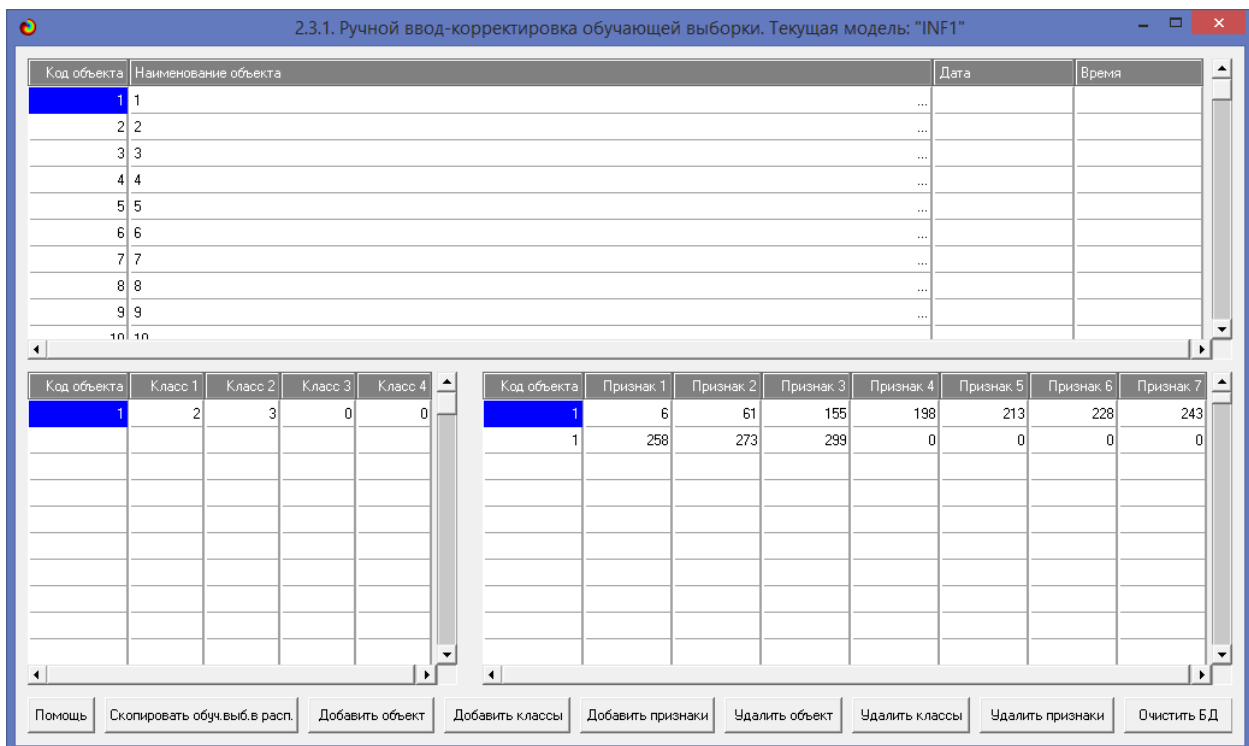


Рисунок 6. Обучающая выборка (фрагмент)

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).

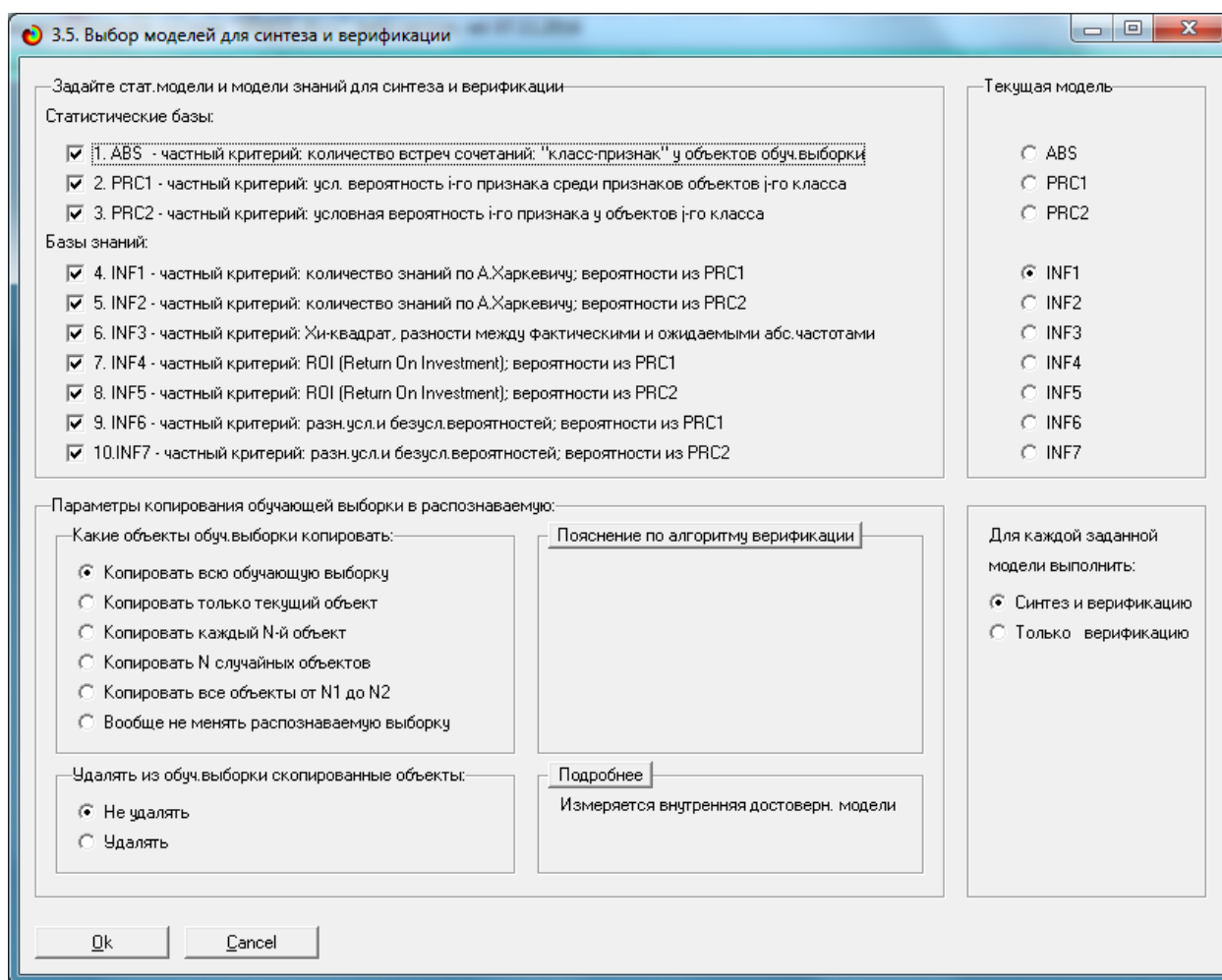


Рисунок 7. Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры

по умолчанию, приведенные на рисунке 10. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

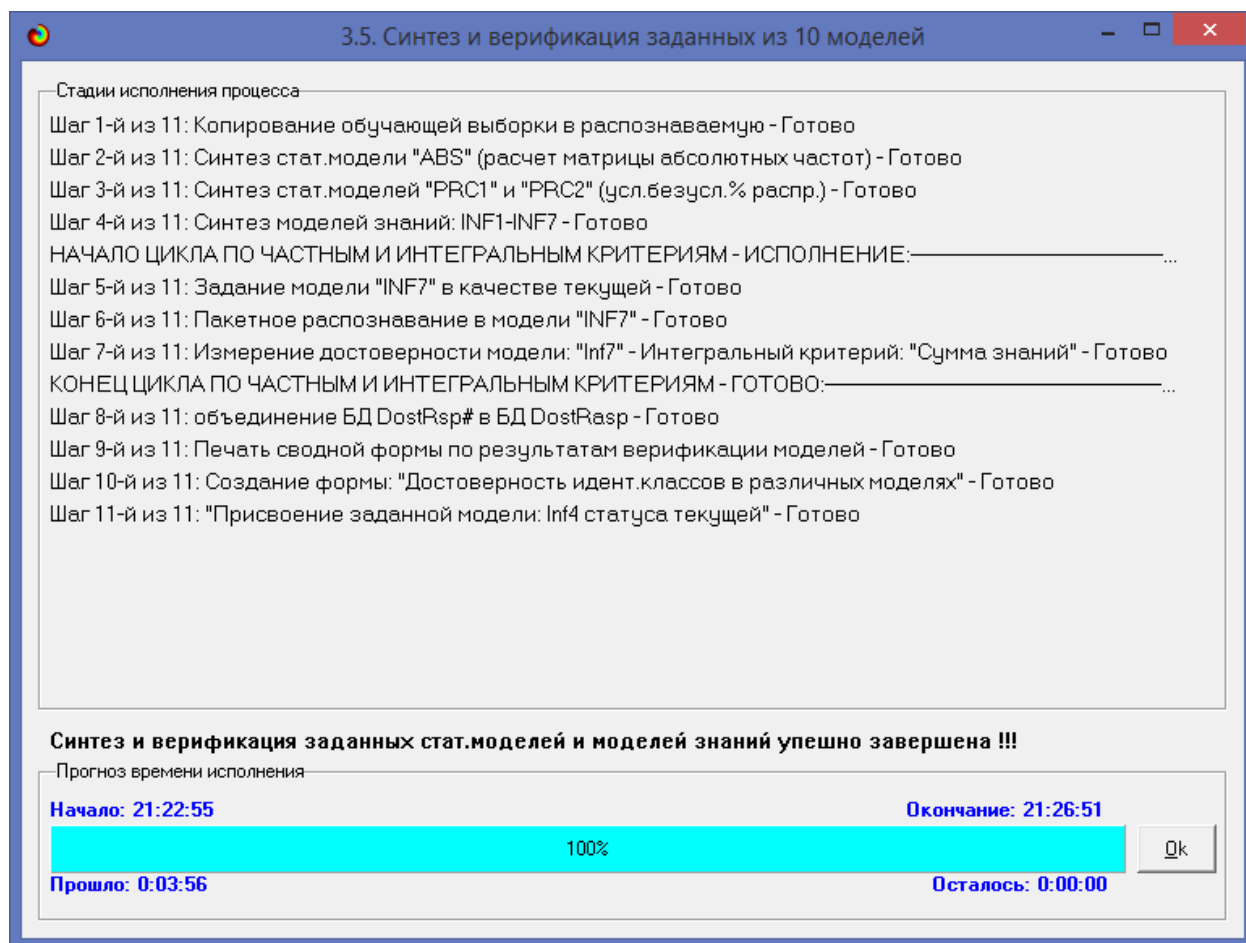


Рисунок 8. Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 42 минут, так как обрабатывается большой объем данных (500 строк). При этом верификация (оценка достоверности моделей) проводилась на всех 131 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. LAT 1/6 (39.9866944, 40.0612338)	2. LAT 2/6 (40.0612338, 40.1357732)	3. LAT 3/6 (40.1357732, 40.2103127)	4. LAT 4/6 (40.2103127, 40.2848521)	5. LAT 5/6 (40.2848521, 40.3593915)	6. LAT 6/6 (40.3593915, 40.4339309)	7. LNG 1/6 { 75.6744381, 75.5689601)	8. LNG 2/6 { 75.5689601, 75.4634822)	9. LNG 3/6 { 75.4634822, 75.3580042)	10. LNG 4/6 { 75.3580042, 75.2525262)	11. LNG 5/6 { 75.2525262, 75.1470483)	12. LNG 6/6 { 75.1470483, 75.0415703)	Сумма
1	TIMESTAMP-2015-12-10 14:39:21		1								1			2
2	TIMESTAMP-2015-12-10 15:39:04				1						1			2
3	TIMESTAMP-2015-12-10 16:17:05				1					1				2
4	TIMESTAMP-2015-12-10 16:32:10		1								1			2
5	TIMESTAMP-2015-12-10 16:46:48			1									1	2
6	TIMESTAMP-2015-12-10 16:47:36		1								1			2
7	TIMESTAMP-2015-12-10 16:51:42					1				1				2
8	TIMESTAMP-2015-12-10 16:56:52				1			1						2
9	TIMESTAMP-2015-12-10 17:09:49		1							1				2
10	TIMESTAMP-2015-12-10 17:10:52					1		1						2
11	TIMESTAMP-2015-12-10 17:12:47		1										1	2
12	TIMESTAMP-2015-12-10 17:15:49			1									1	2
13	TIMESTAMP-2015-12-10 17:29:21				1						1			2
14	TIMESTAMP-2015-12-10 17:31:00				1							1		2
15	TIMESTAMP-2015-12-10 17:33:50				1							1		2
16	TIMESTAMP-2015-12-10 17:35:41		1							1				2
17	TIMESTAMP-2015-12-10 17:41:54					1		1						2
18	TIMESTAMP-2015-12-10 17:42:44				1						1			2
19	TIMESTAMP-2015-12-10 17:43:29		1								1			2
20	TIMESTAMP-2015-12-10 17:43:45		1								1			2
21	TIMESTAMP-2015-12-10 17:45:23			1									1	2
22	TIMESTAMP-2015-12-10 17:53:22		1									1		2
23	TIMESTAMP-2015-12-10 17:58:22			1								1		2

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. STATUS NO SHOW	2. STATUS SHOW UP	3. AGE ОЧЕНЬ МАЛОЕ: 1/5 (0.0000000, 19.0000000)	4. AGE МАЛОЕ: 2/5 (19.0000000, 38.0000000)	5. AGE СРЕДНЕЕ: 3/5 (38.0000000, 57.0000000)	6. AGE БОЛЬШОЕ: 4/5 (57.0000000, 76.0000000)	7. AGE ОЧЕНЬ БОЛЬШОЕ: 5/5 (76.0000000, 95.0000000)	Сумма	Среднее	Средн. квадр. откл.
1	DAYOFTHEWEEK-Friday	0.056	-0.027	0.056	-0.161	-0.117	0.171	-0.011	-0.033	-0.005	0.113
2	DAYOFTHEWEEK-Monday	0.035	-0.016	-0.071	0.181	-0.545		0.224	-0.192	-0.027	0.252
3	DAYOFTHEWEEK-Saturday		0.134		0.481				0.614	0.088	0.180
4	DAYOFTHEWEEK-Thursday	-0.086	0.032	-0.086	-0.197	0.185	0.028	-0.154	-0.278	-0.040	0.131
5	DAYOFTHEWEEK-Tuesday	-0.039	0.016	0.044	-0.150	0.194	-0.074		-0.010	-0.001	0.107
6	DAYOFTHEWEEK-Wednesday	0.079	-0.040	0.079	0.175	-0.501	-0.213	0.268	-0.154	-0.022	0.261
7	APPOINTMENTREGISTRATION-2014-01-03T10:1...		0.134		0.481				0.614	0.088	0.180
8	APPOINTMENTREGISTRATION-2014-01-10T09:2...	0.442					0.556		0.998	0.143	0.246
9	APPOINTMENTREGISTRATION-2014-01-15T09:2...		0.134			0.524			0.658	0.094	0.196
10	APPOINTMENTREGISTRATION-2014-02-12T08:2...	0.442					0.556		0.998	0.143	0.246
11	APPOINTMENTREGISTRATION-2014-02-17T12:5...		0.134	0.592					0.725	0.104	0.221
12	APPOINTMENTREGISTRATION-2014-02-17T14:3...		0.134				0.556		0.690	0.099	0.208
13	APPOINTMENTREGISTRATION-2014-02-25T09:2...		0.134				0.556		0.690	0.099	0.208
14	APPOINTMENTREGISTRATION-2014-03-12T10:2...		0.134		0.481				0.614	0.088	0.180
15	APPOINTMENTREGISTRATION-2014-03-13T08:2...		0.134			0.524			0.658	0.094	0.196
16	APPOINTMENTREGISTRATION-2014-03-21T08:2...	0.442			0.481				0.922	0.132	0.225
17	APPOINTMENTREGISTRATION-2014-03-25T08:4...		0.134		0.481				0.614	0.088	0.180
18	APPOINTMENTREGISTRATION-2014-03-31T10:1...		0.134		0.481				0.614	0.088	0.180
19	APPOINTMENTREGISTRATION-2014-04-01T13:3...		0.134	0.592					0.725	0.104	0.221
20	APPOINTMENTREGISTRATION-2014-04-24T15:2...		0.134					1.037	1.170	0.167	0.387
21	APPOINTMENTREGISTRATION-2014-04-30T08:0...		0.134				0.556		0.690	0.099	0.208

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

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами"											
Код признака	Наименование описательной шкалы и градации	1. STATUS NO SHOW	2. STATUS SHOW UP	3. AGE ОЧЕНЬ МАЛОЕ: 1/5 (0.0000000, 19.0000000)	4. AGE МАЛОЕ: 2/5 (19.0000000, 38.0000000)	5. AGE СРЕДНЕЕ: 3/5 (38.0000000, 57.0000000)	6. AGE БОЛЬШОЕ: 4/5 (57.0000000, 76.0000000)	7. AGE ОЧЕНЬ БОЛЬШОЕ: 5/5 (76.0000000, 95.0000000)	Сумма	Среднее	Средн. квадр. откл.
1	DAYOFTHEWEEK-Friday	0.848	-0.848	0.566	-1.636	-1.121	2.222	-0.030			1.331
2	DAYOFTHEWEEK-Monday	0.545	-0.545	-0.636	3.091	-3.364		0.909			1.945
3	DAYOFTHEWEEK-Saturday	-0.303	0.303	-0.202	0.727	-0.242	-0.222	-0.061			0.380
4	DAYOFTHEWEEK-Thursday	-1.576	1.576	-1.051	-2.818	3.939	0.444	-0.515			2.236
5	DAYOFTHEWEEK-Tuesday	-0.667	0.667	0.556	-2.000	3.667	-0.889	-1.333			1.880
6	DAYOFTHEWEEK-Wednesday	1.152	-1.152	0.768	2.636	-2.879	-1.556	1.030			1.913
7	APPOINTMENTREGISTRATION-2014-01-03T10:1...	-0.303	0.303	-0.202	0.727	-0.242	-0.222	-0.061			0.380
8	APPOINTMENTREGISTRATION-2014-01-10T09:2...	0.697	-0.697	-0.202	-0.273	-0.242	0.778	-0.061			0.541
9	APPOINTMENTREGISTRATION-2014-01-15T09:2...	-0.303	0.303	-0.202	-0.273	0.758	-0.222	-0.061			0.393
10	APPOINTMENTREGISTRATION-2014-02-12T08:2...	0.697	-0.697	-0.202	-0.273	-0.242	0.778	-0.061			0.541
11	APPOINTMENTREGISTRATION-2014-02-17T12:5...	-0.303	0.303	0.798	-0.273	-0.242	-0.222	-0.061			0.410
12	APPOINTMENTREGISTRATION-2014-02-17T14:3...	-0.303	0.303	-0.202	-0.273	-0.242	0.778	-0.061			0.401
13	APPOINTMENTREGISTRATION-2014-02-25T09:2...	-0.303	0.303	-0.202	-0.273	-0.242	0.778	-0.061			0.401
14	APPOINTMENTREGISTRATION-2014-03-12T10:2...	-0.303	0.303	-0.202	0.727	-0.242	-0.222	-0.061			0.380
15	APPOINTMENTREGISTRATION-2014-03-13T08:2...	-0.303	0.303	-0.202	-0.273	0.758	-0.222	-0.061			0.393
16	APPOINTMENTREGISTRATION-2014-03-21T08:2...	0.697	-0.697	-0.202	0.727	-0.242	-0.222	-0.061			0.525
17	APPOINTMENTREGISTRATION-2014-03-25T08:4...	-0.303	0.303	-0.202	0.727	-0.242	-0.222	-0.061			0.380
18	APPOINTMENTREGISTRATION-2014-03-31T10:1...	-0.303	0.303	-0.202	0.727	-0.242	-0.222	-0.061			0.380
19	APPOINTMENTREGISTRATION-2014-04-01T13:3...	-0.303	0.303	0.798	-0.273	-0.242	-0.222	-0.061			0.410
20	APPOINTMENTREGISTRATION-2014-04-24T15:2...	-0.303	0.303	-0.202	-0.273	-0.242	-0.222	0.939			0.464
21	APPOINTMENTREGISTRATION-2014-04-30T08:0...	-0.303	0.303	-0.202	-0.273	-0.242	0.778	-0.061			0.401

1.5. Результаты верификации моделей

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

Наименование модели и частного критерия	Интегральный критерий	Всего логически объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма уровней истинности
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	198	198		495		0.286	1.000	0.444	176
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс. частот по признак...	198	198		495		0.286	1.000	0.444	104
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн. частот с о...	198	198		495		0.286	1.000	0.444	176
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн. частот по при...	198	198		495		0.286	1.000	0.444	168
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн. частот с о...	198	198		495		0.286	1.000	0.444	176
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн. частот по при...	198	198		495		0.286	1.000	0.444	168
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	198	105	483	12	93	0.897	0.530	0.667	19
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	198	195	322	173	3	0.530	0.985	0.689	55
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	198	105	483	12	93	0.897	0.530	0.667	19
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	198	195	322	173	3	0.530	0.985	0.689	55
6. INF3 - частный критерий: Хинквадрат, разности между факти...	Семантический резонанс зна...	198	153	300	195	45	0.440	0.773	0.560	77
6. INF3 - частный критерий: Хинквадрат, разности между факти...	Сумма знаний	198	153	302	193	45	0.442	0.773	0.563	72
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	198	113	494	1	85	0.991	0.571	0.724	20
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	198	198	269	226		0.467	1.000	0.637	30
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	198	113	494	1	85	0.991	0.571	0.724	20
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	198	198	269	226		0.467	1.000	0.637	30
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	198	148	296	199	50	0.427	0.747	0.543	74
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	198	153	276	219	45	0.411	0.773	0.537	28
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	198	148	296	199	50	0.427	0.747	0.543	74
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	198	153	276	219	45	0.411	0.773	0.537	28

Рисунок 9. Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF4и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели (F-мера Ван Ризбергера) составляет 0,724,а точность модели (L1-мера профессора Луценко 0,92). 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 мультимодальное обобщение, предложенное профессором Е.В.Луценко [7] (рисунок 10).

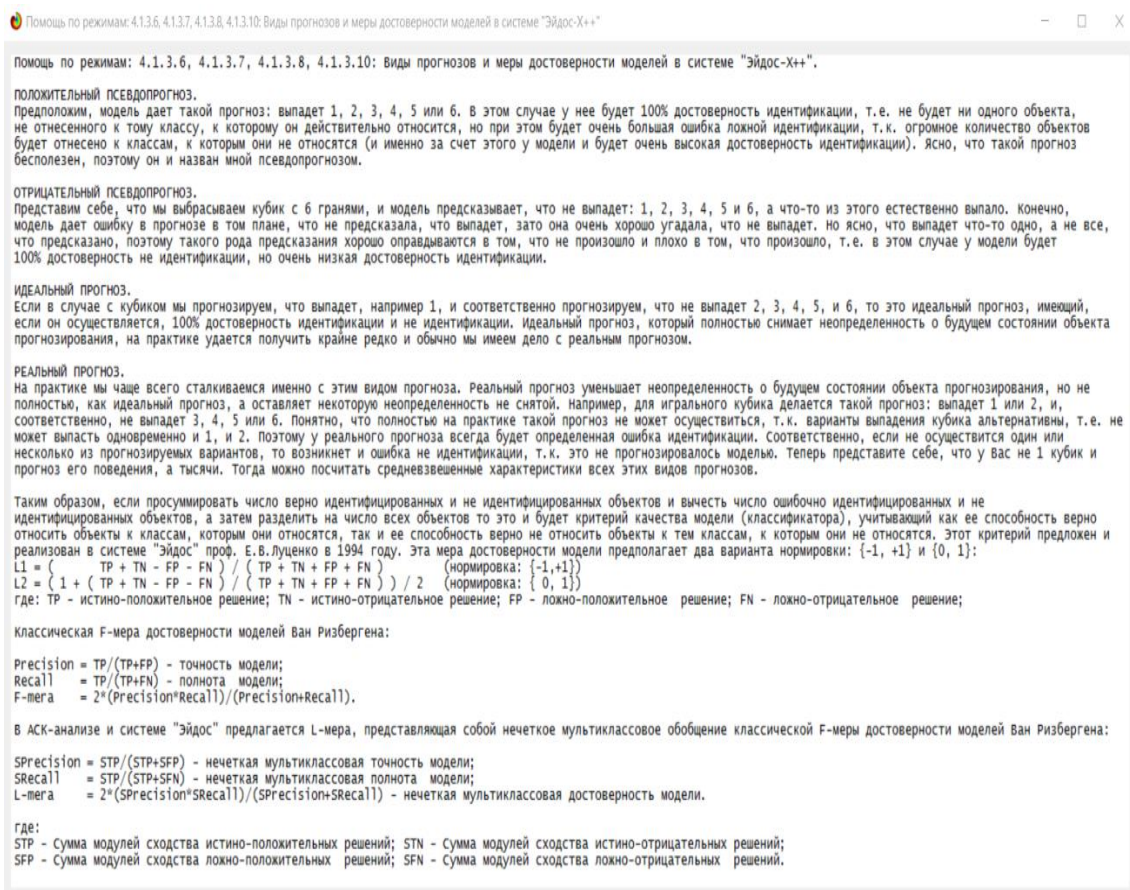


Рисунок 10. Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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 INF4.

4.1.3.11. Част.распр.уровн.сход.ТР,TN,FP,FN решений при разных моделях и инт.критериях. Текущая модель: "INF4"

Наименование частного и интегрального критерия	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
1. ABS - ЧАСТНЫЙ КРИТЕРИЙ: КОЭФФИЦИЕНТ ВЕРНОСТИ РЕШЕНИЙ													
Интегральный критерий: КОРРЕЛЯЦИЯ АБС. ЧАСТОТ С ОБР.ОБЪЕКТА													
Уровни схождения (Ур.Сх.) (%):	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Схождения истинных решений (TP+TN)													
Част.распр.Уровней Схождения ложных решений (FP+FN)													
Част.распр.Уровней Схождения истинно-положительных решений (TP)													
Част.распр.Уровней Схождения истинно-отрицательных решений (TN)													
Част.распр.Уровней Схождения ложно-положительных решений (FP)													
Част.распр.Уровней Схождения ложно-отрицательных решений (FN)													
Интегральный критерий: СУММА АБС.ЧАСТОТ ПО ПРИЗНАКАМ ОБЪЕКТА													
Уровни схождения (Ур.Сх.) (%):	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Схождения истинных решений (TP+TN)													
Част.распр.Уровней Схождения ложных решений (FP+FN)													
Част.распр.Уровней Схождения истинно-положительных решений (TP)													
Част.распр.Уровней Схождения истинно-отрицательных решений (TN)													
Част.распр.Уровней Схождения ложно-положительных решений (FP)													
Част.распр.Уровней Схождения ложно-отрицательных решений (FN)													
2. PRC1 - ЧАСТНЫЙ КРИТЕРИЙ: УСЛ. ВЕРОЯТНОСТЬ I-ГО ПРИЗНАКА СРЕДИ ПРИЗНАКОВ ОБ...													
Интегральный критерий: КОРРЕЛЯЦИЯ УСЛ.ОТН.ЧАСТОТ С ОБР.ОБЪЕКТА													
Уровни схождения (Ур.Сх.) (%):	-100	-99	-98	-97	-96	-95	-94	-93	-92	-91	-90	-89	-88
Част.распр.Уровней Схождения истинных решений (TP+TN)													
Част.распр.Уровней Схождения ложных решений (FP+FN)													
Част.распр.Уровней Схождения истинно-положительных решений (TP)													

Помощь Графики TP,TN,FP,FN, инт.крит.-резонанс знаний Графики TP,TN,FP,FN, инт.крит.-сумма знаний Excel

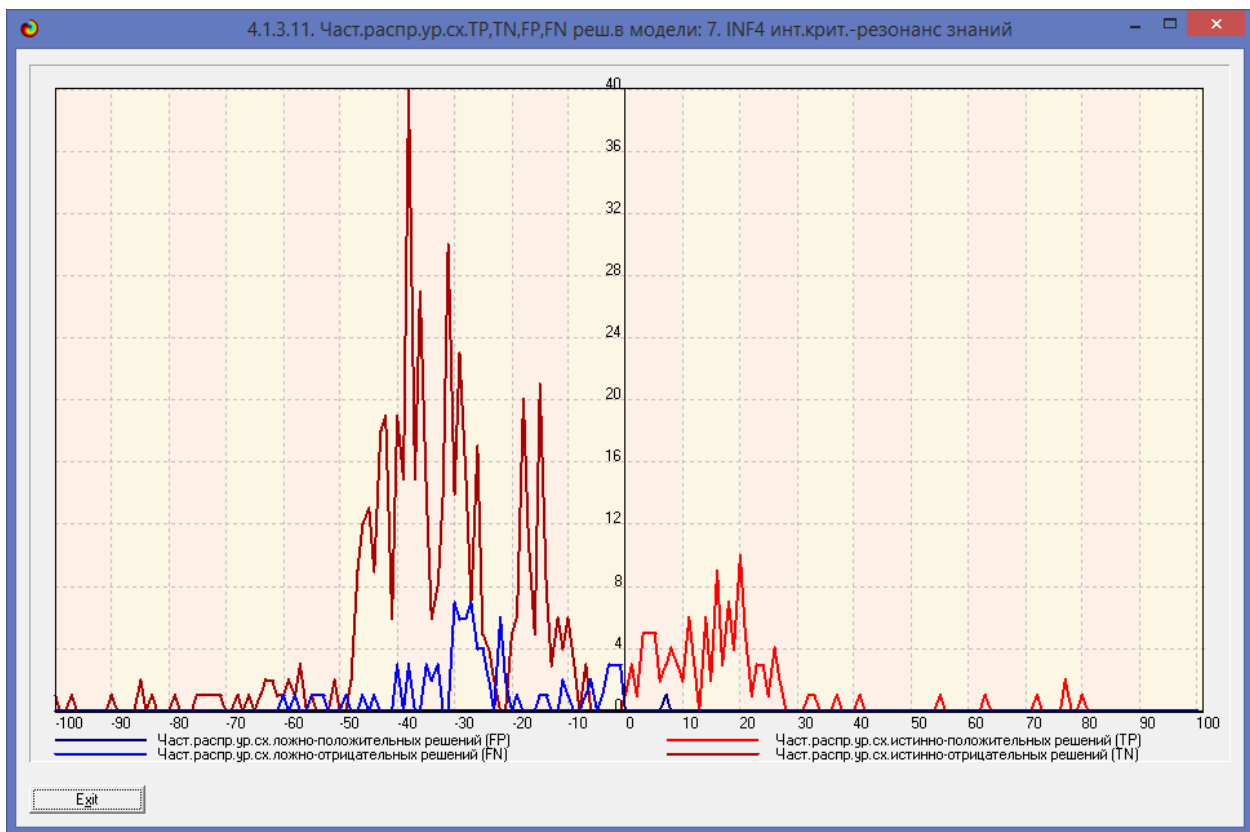


Рисунок 11. Частное распределение схождения-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF4

Из рисунка 11 видно, что:

– наиболее достоверная модель INF4 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 9);

– модуль уровня сходства-различия в наиболее достоверной модели INF3 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):

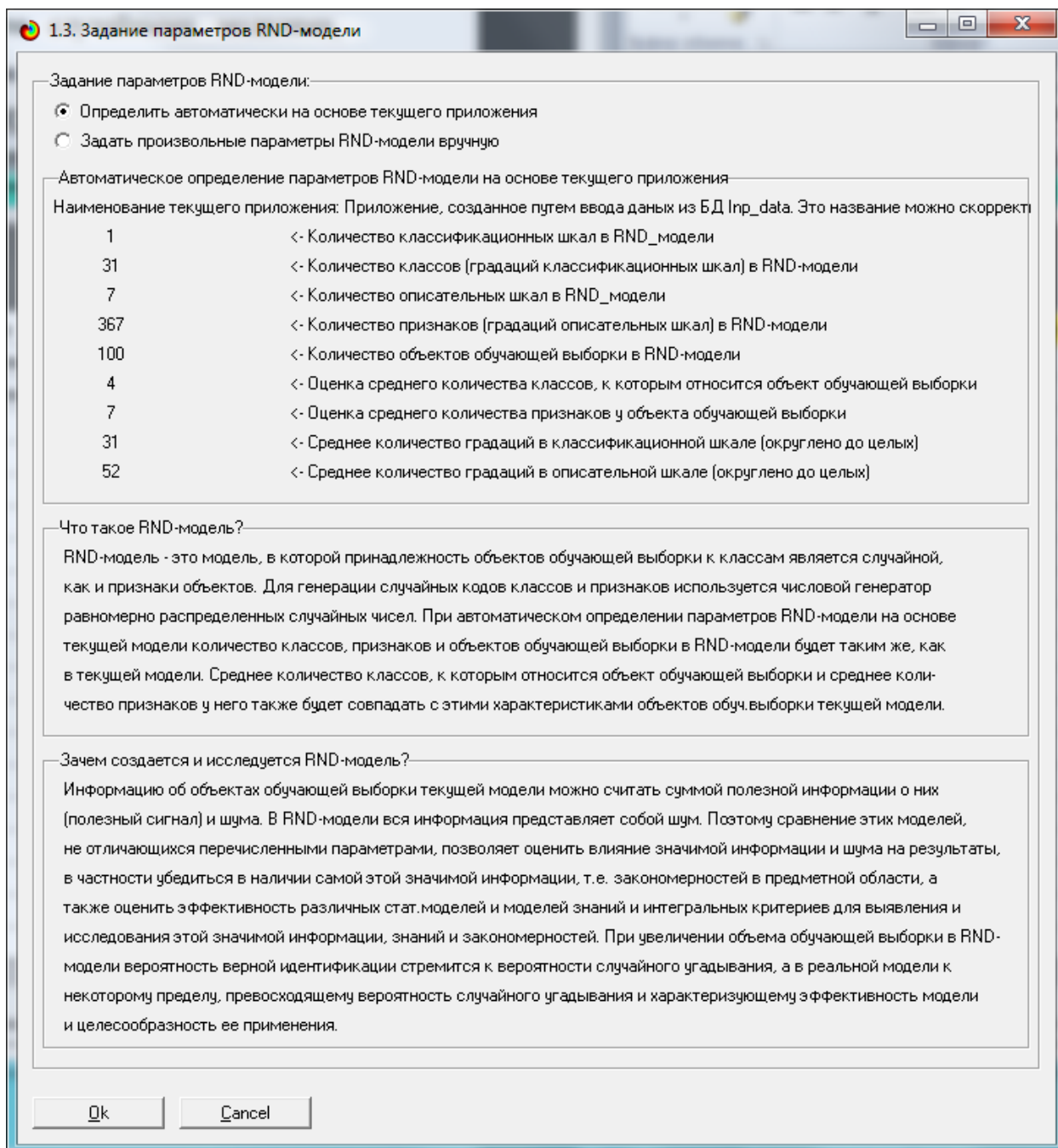
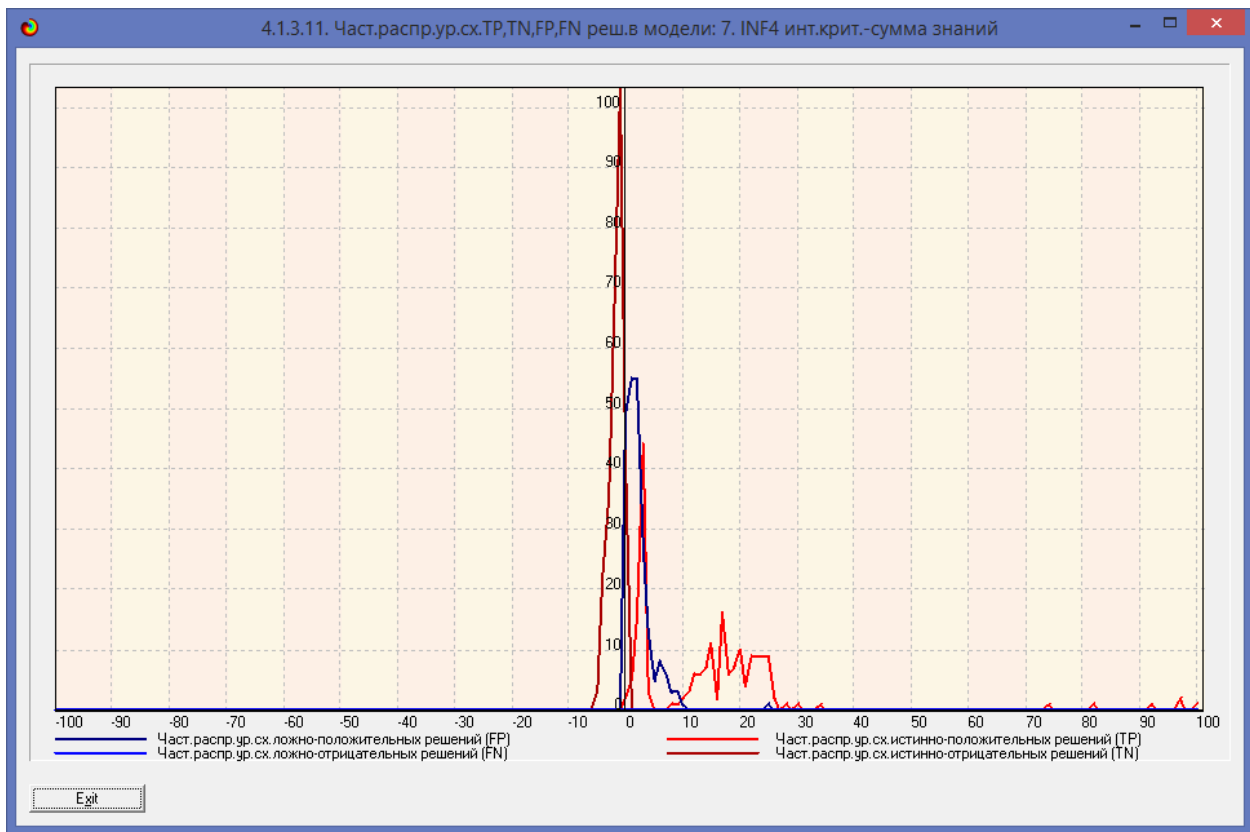


Рисунок 12. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 13 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF4



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].

4.1.3.6. Обобщенная форма по довер. моделям при разн. инт. крит. Текущая модель: "INF4"									
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	Ф-мера Ван Ризбергера
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс. частот с обр...	198	198		495		0.286	1.000	0.444
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс. частот по признак...	198	198		495		0.286	1.000	0.444
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл. отн. частот с о...	198	198		495		0.286	1.000	0.444
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл. отн. частот по приз...	198	198		495		0.286	1.000	0.444
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл. отн. частот с о...	198	198		495		0.286	1.000	0.444
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл. отн. частот по приз...	198	198		495		0.286	1.000	0.444
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	198	105	483	12	93	0.897	0.530	0.667
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	198	195	322	173	3	0.530	0.985	0.689
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	198	105	483	12	93	0.897	0.530	0.667
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	198	195	322	173	3	0.530	0.985	0.689
6. INF3 - частный критерий: Хинкватрат, разности между фактич...	Семантический резонанс зна...	198	153	300	195	45	0.440	0.773	0.560
6. INF3 - частный критерий: Хинкватрат, разности между фактич...	Сумма знаний	198	153	302	193	45	0.442	0.773	0.563
7. INF4 - частный критерий: ROI (Return On Investment), вероятн...	Семантический резонанс зна...	198	113	494	1	85	0.991	0.571	0.724
7. INF4 - частный критерий: ROI (Return On Investment), вероятн...	Сумма знаний	198	198	269	226		0.467	1.000	0.637
8. INF5 - частный критерий: ROI (Return On Investment), вероятн...	Семантический резонанс зна...	198	113	494	1	85	0.991	0.571	0.724
8. INF5 - частный критерий: ROI (Return On Investment), вероятн...	Сумма знаний	198	198	269	226		0.467	1.000	0.637
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	198	148	296	199	50	0.427	0.747	0.543
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	198	153	276	219	45	0.411	0.773	0.537
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	198	148	296	199	50	0.427	0.747	0.543
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	198	153	276	219	45	0.411	0.773	0.537

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

На основе его сравнения с рисунком 9 можно сделать следующие выводы:

- достоверности лучшей модели INF4, отражающей реальный объект моделирования, примерно на 20% 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 INF4 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1.

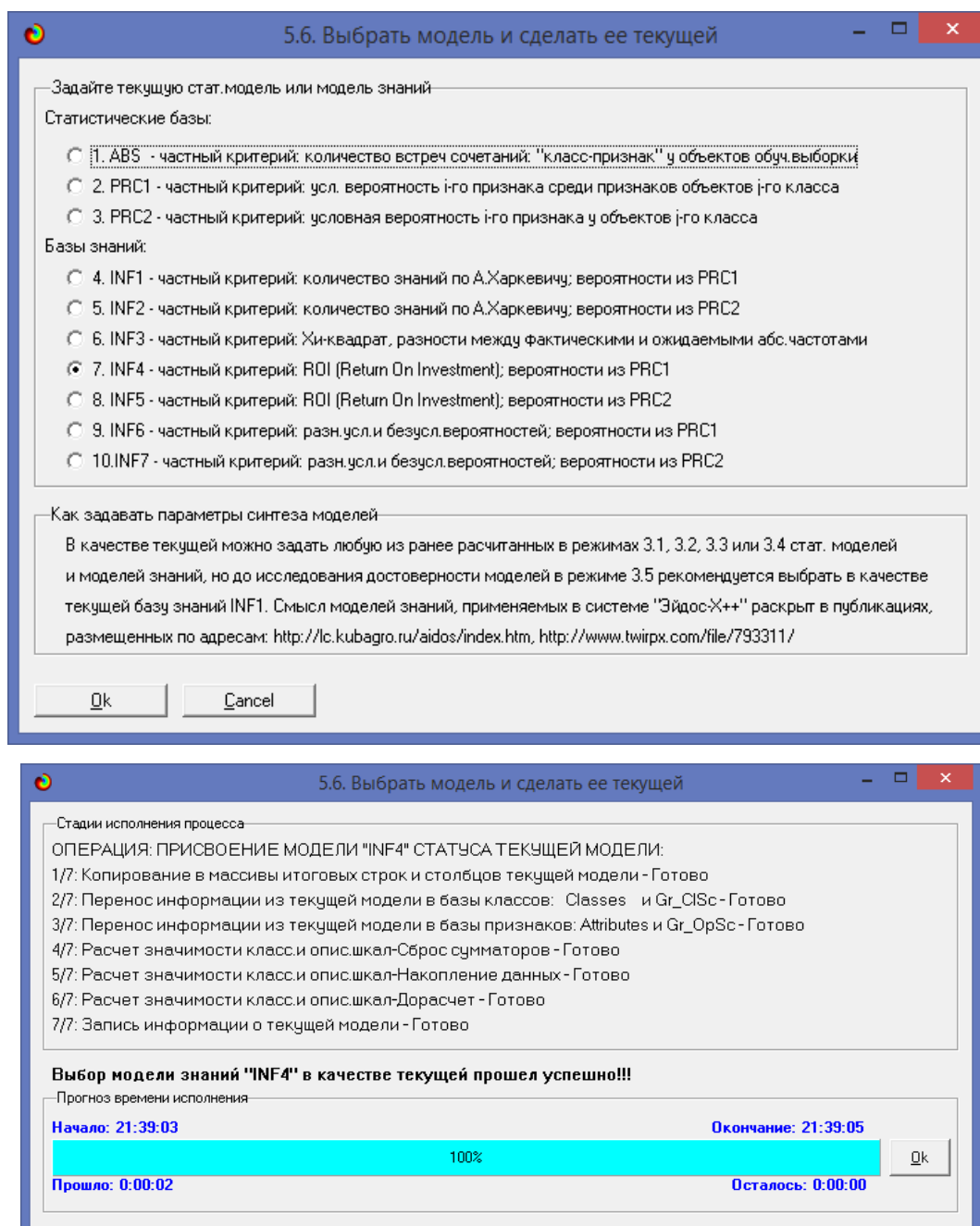


Рисунок 15. Экранные формы режима задания модели в качестве текущей

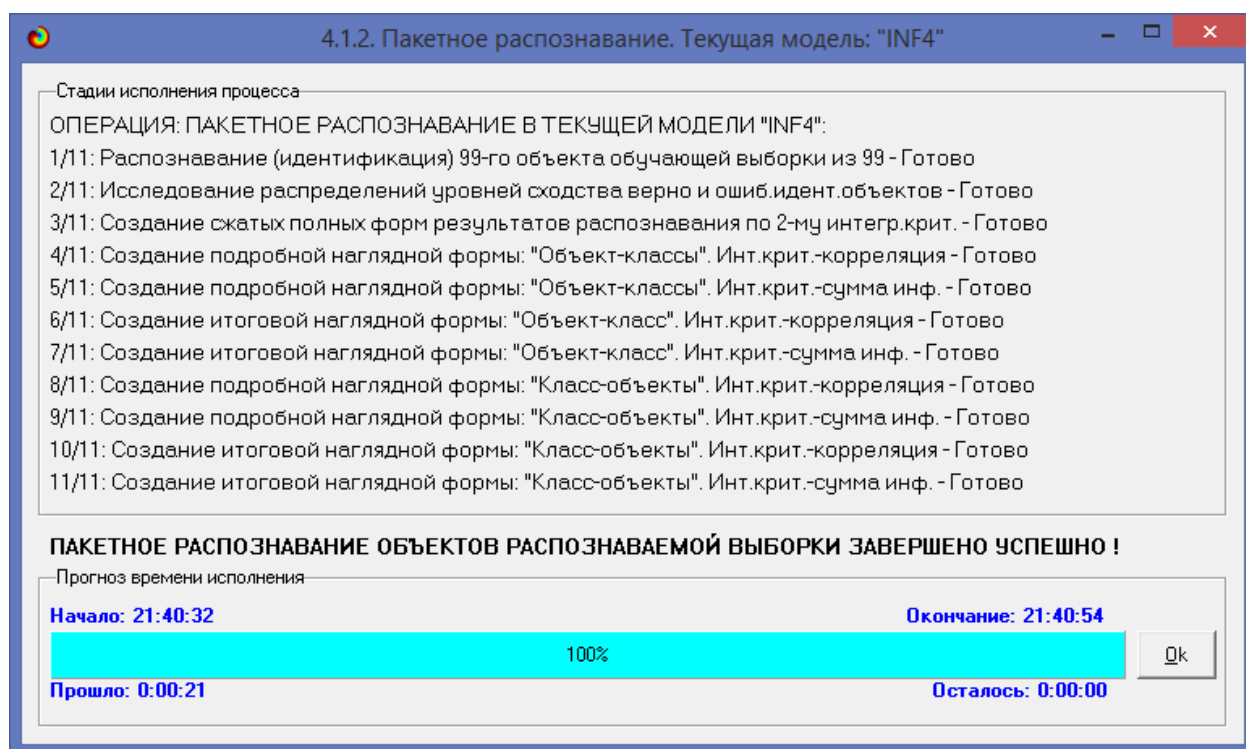


Рисунок 16. Экранная форма режима пакетного распознавания в текущей модели

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 15 и 16 приведены примеры прогнозов высокой и низкой достоверности частоты и классов ирисов в наиболее достоверной модели INF4 на основе наблюдения предыстории их развития:

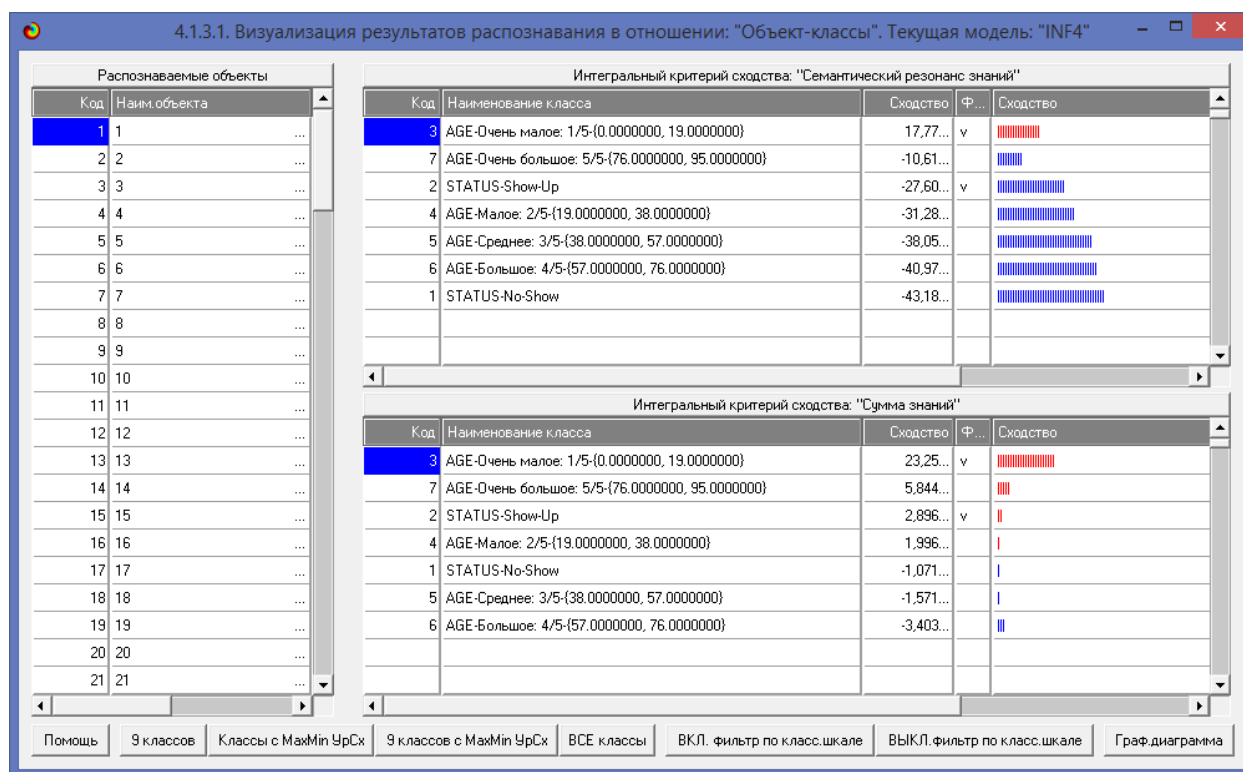


Рисунок 17. Пример идентификации классов в модели INF4

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 18)

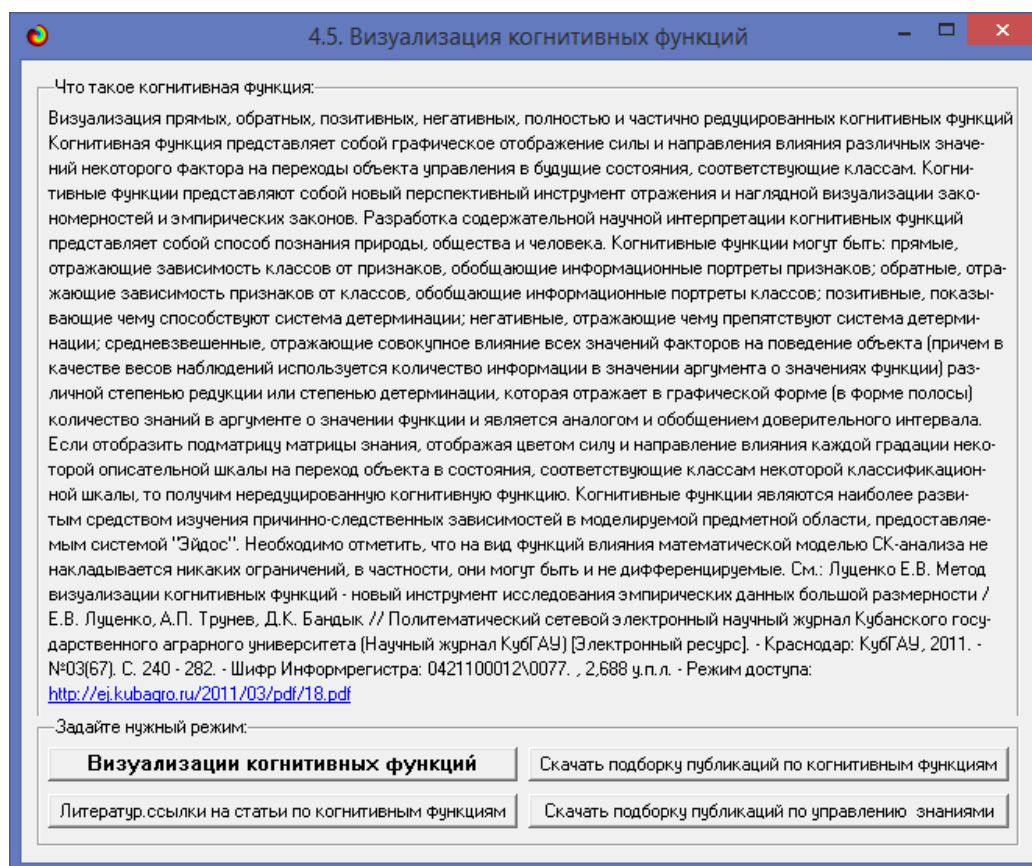


Рисунок 18. Экранная форма режима визуализации когнитивных функций

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. 19–24):

На рисунке 19 представлена когнитивная функция, показывающая, что в меньшей широте и долготе преобладают вызовы по чрезвычайные ситуации,

связанных с травмами людей. В средней широте и долготе вызовы связанные с кровоизлиянием, а в большей долготе вызовы, связанные с тошнотой [5].

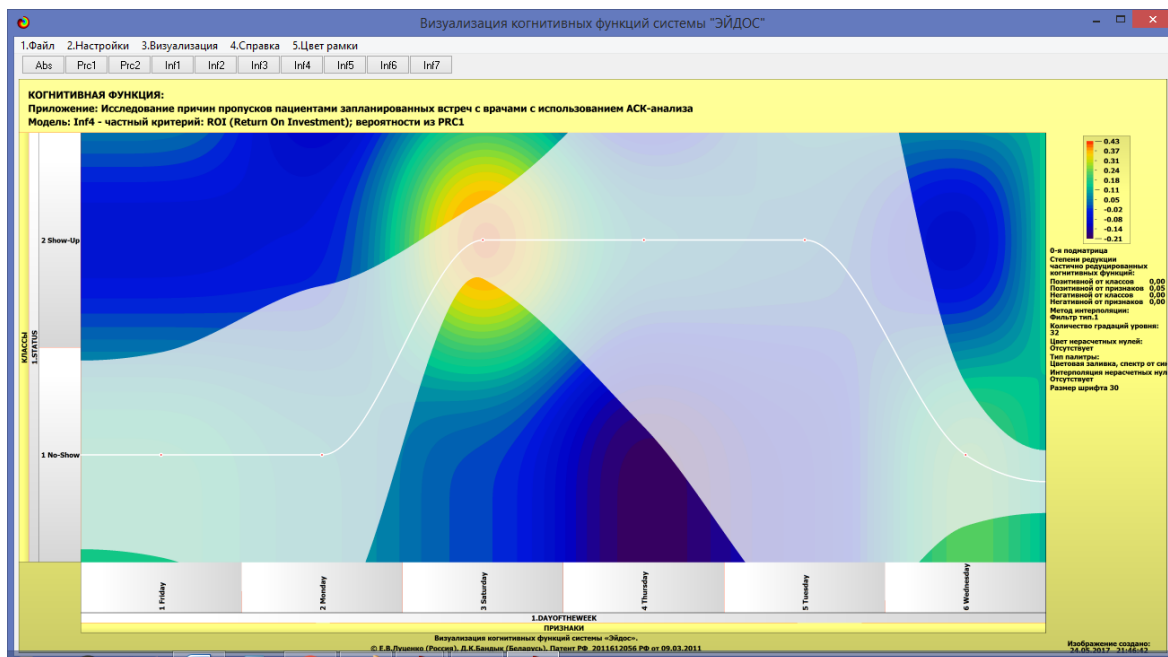


Рисунок 19. Когнитивная функция, отражающая взаимосвязь посещения/не посещения, в зависимости от дня недели

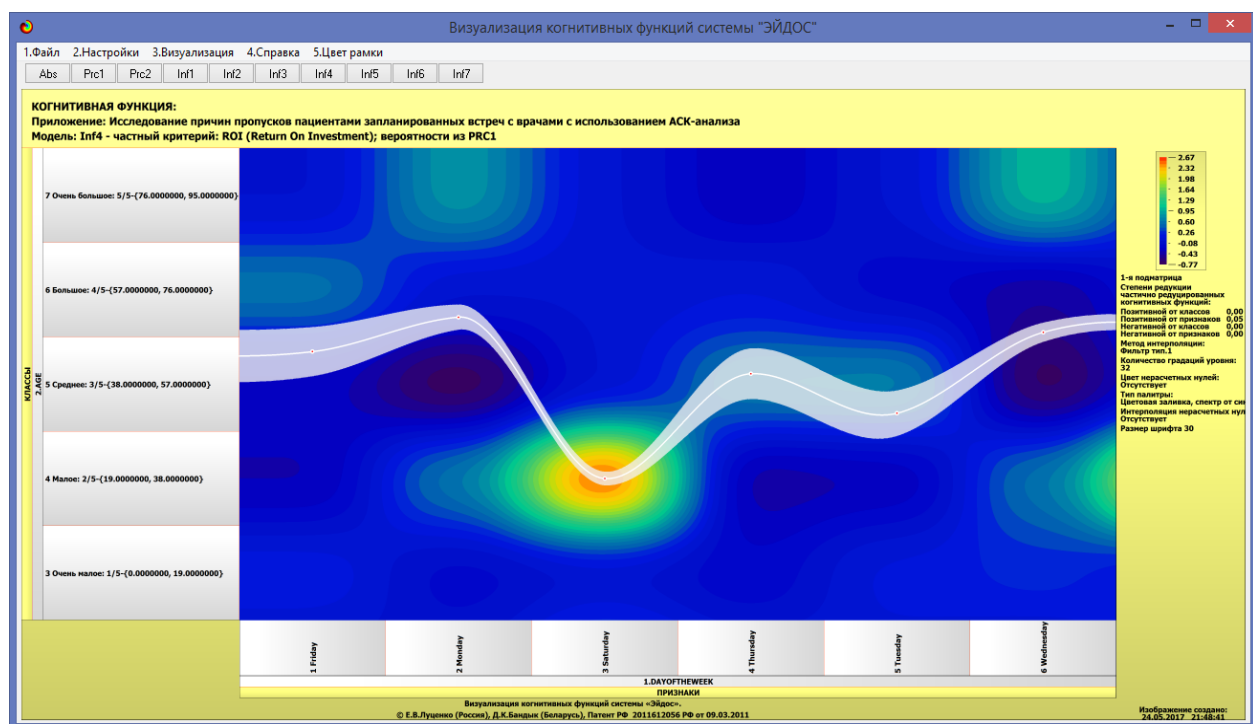


Рисунок 20. Когнитивная функция, отражающая взаимосвязь возраста посетителя и дня недели

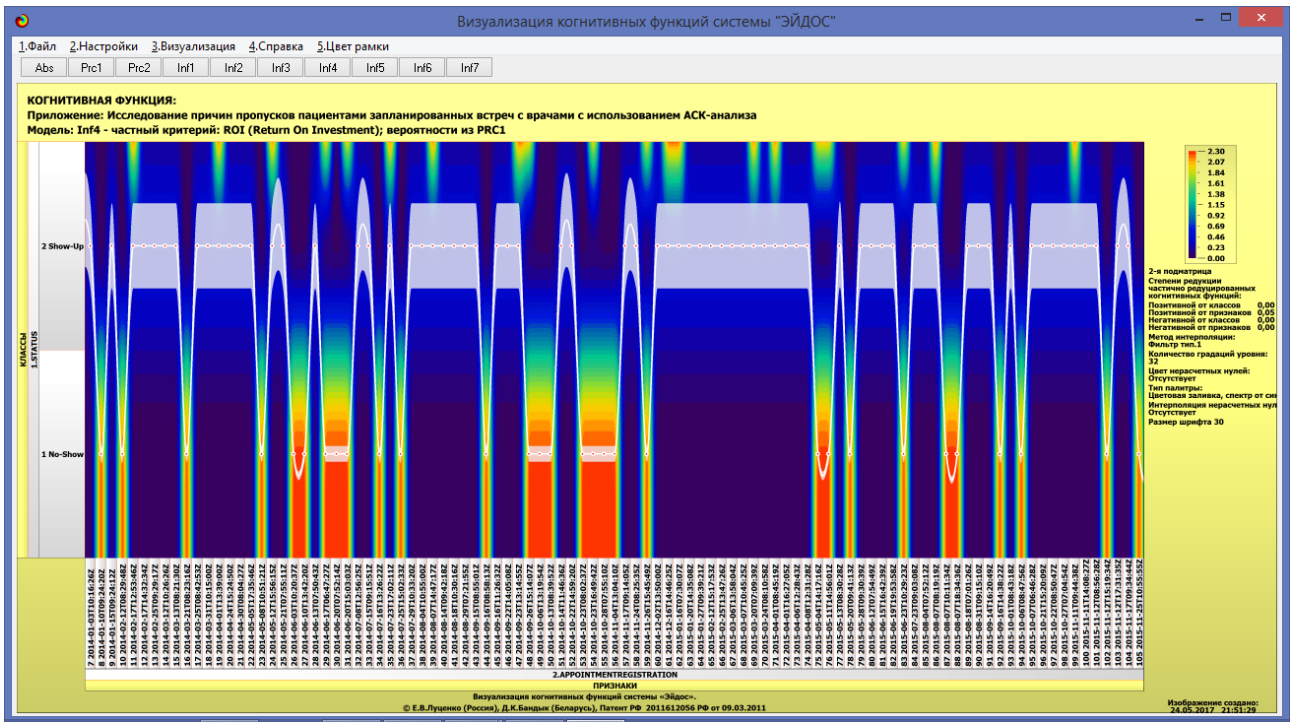


Рисунок 21. Когнитивная функция, отражающая взаимосвязь посещения/непосещения с датой посещения

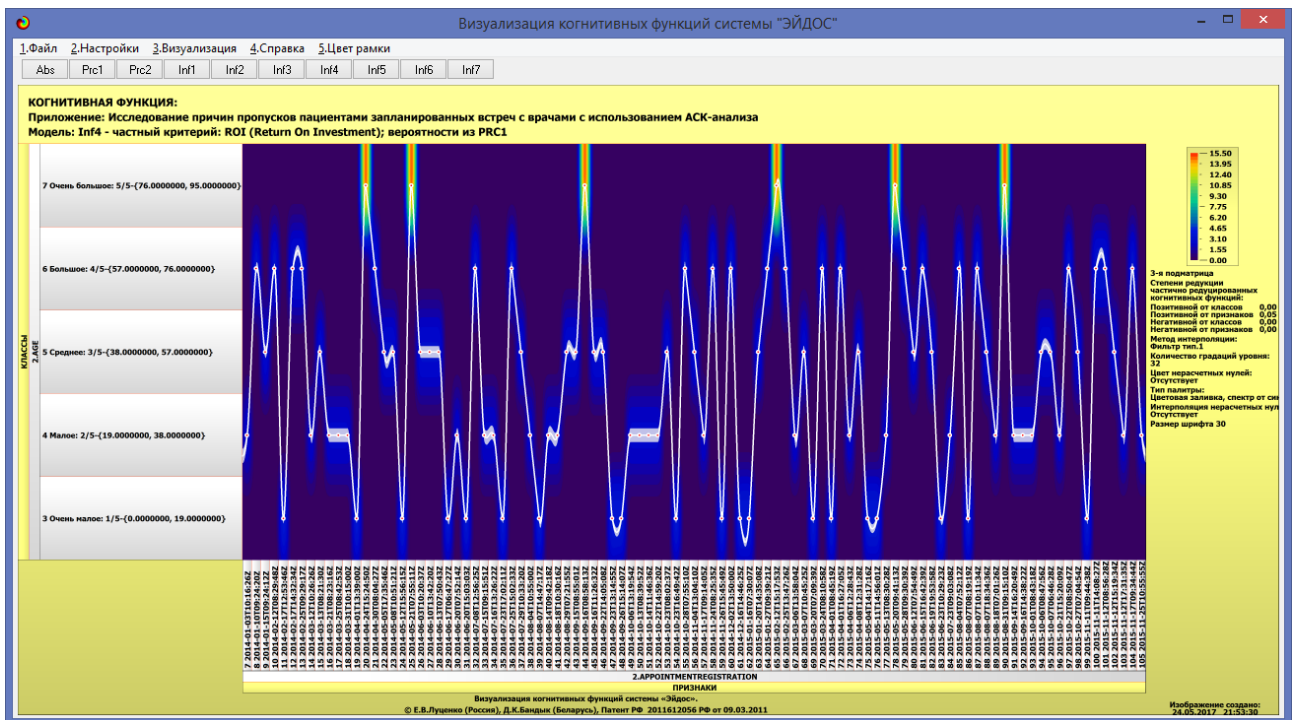


Рисунок 22. Когнитивная функция, отражающая взаимосвязь возраста пациента от даты регистрации

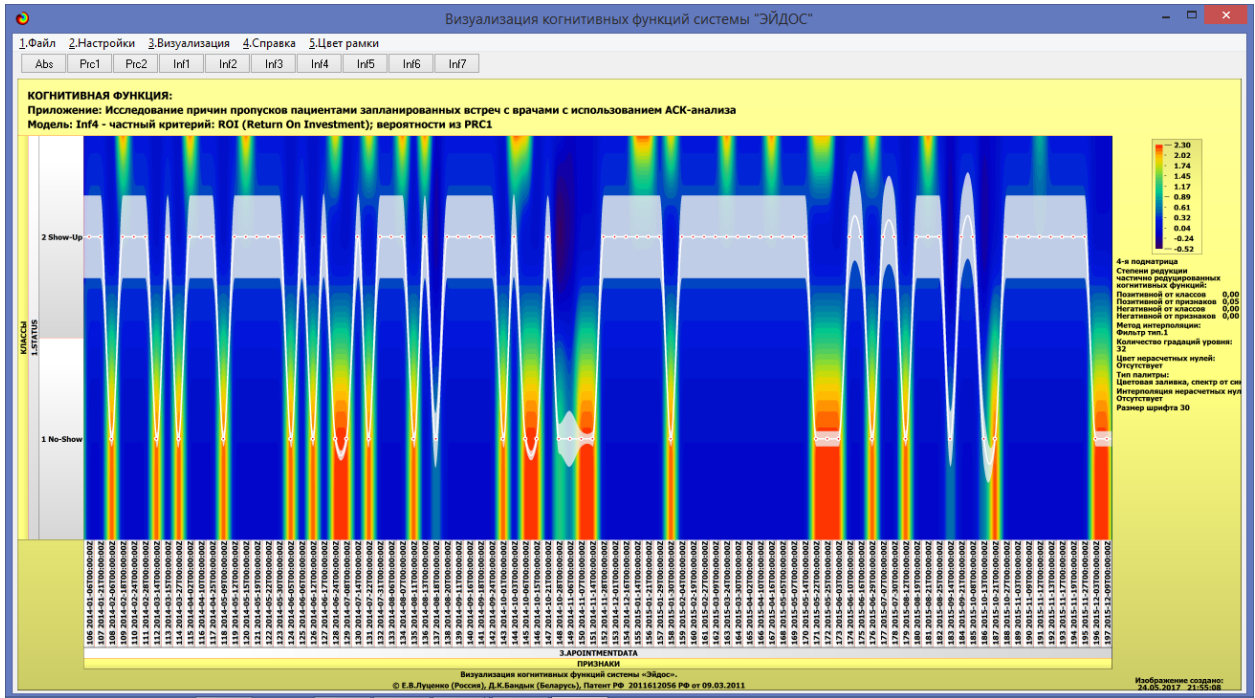


Рисунок 23. Когнитивная функция, отражающая взаимосвязь посещения/непосещения с датой

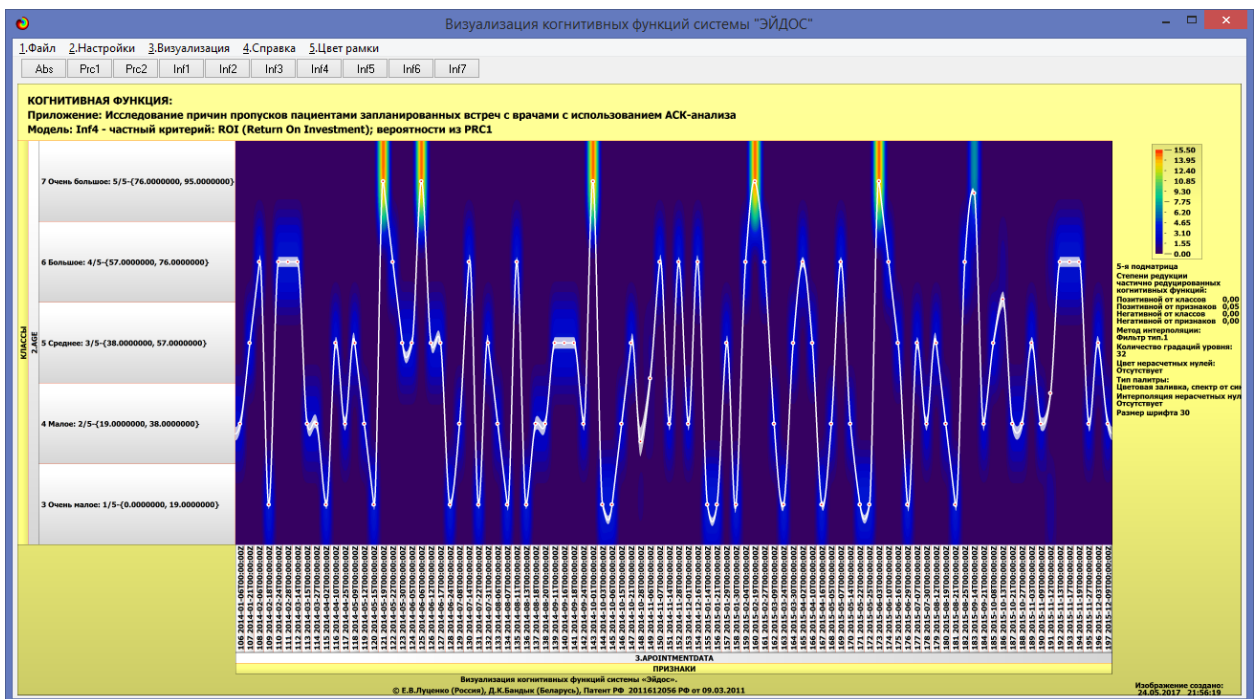


Рисунок 24. Когнитивная функция, отражающая взаимосвязь возраста пациента с датой регистрации

Сходство-различие обобщенных образов различных результатов научной деятельности по характерным для них системам значений показателей. Результаты

сравнения классов по системе характерных для типов чрезвычайных ситуаций приведены на рисунок 25:

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 2 "STATUS-Show-Up" в модели: 8 "INF5"

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	STATUS-No-Show	1	2	STATUS-Show-Up	100.000
2	STATUS-Show-Up	2	6	AGE-Большое: 4/5-(57.0000000, 76.0000000)	24.964
3	AGE-Очень малое: 1/5-(0.0000000, 19.0000000) ...	3	4	AGE-Малое: 2/5-(19.0000000, 38.0000000)	22.784
4	AGE-Малое: 2/5-(19.0000000, 38.0000000) ...	4	5	AGE-Среднее: 3/5-(38.0000000, 57.0000000)	19.245
5	AGE-Среднее: 3/5-(38.0000000, 57.0000000) ...	5	3	AGE-Очень малое: 1/5-(0.0000000, 19.0000000)	10.779
6	AGE-Большое: 4/5-(57.0000000, 76.0000000) ...	6	7	AGE-Очень большое: 5/5-(76.0000000, 95.0000000)	5.305
7	AGE-Очень большое: 5/5-(76.0000000, 95.0000000) ...	7	1	STATUS-No-Show	-45.116

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

Рисунок 25 - Результаты сравнения классов по системе характерных для типов чрезвычайных ситуаций

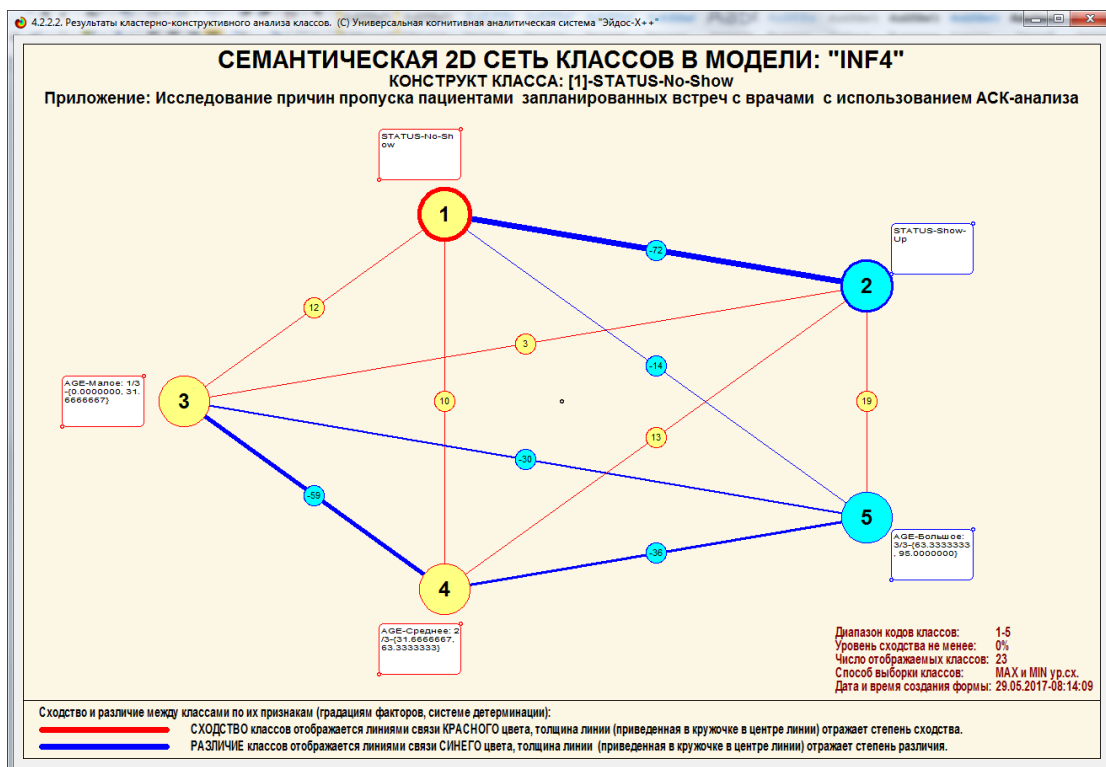


Рисунок 26. Результаты сравнения классов по системе характерных для значений чрезвычайных ситуаций

Из когнитивной диаграммы, приведенной на рисунке 26, мы видим, что, посещение пациентов зависит от их возраста, в следствии чего возраст (малое, среднее, большое) образу общую группу, что в дальнейшем отражается на посещение или непосещение пациентами мед. служб. Эти группы образуют противоположный кластер с более высокой вариабельностью объектов, внутри него. Отметим также, что приведенная когнитивная диаграмма формируется системой «Эйдос» автоматически на основе созданных моделей.

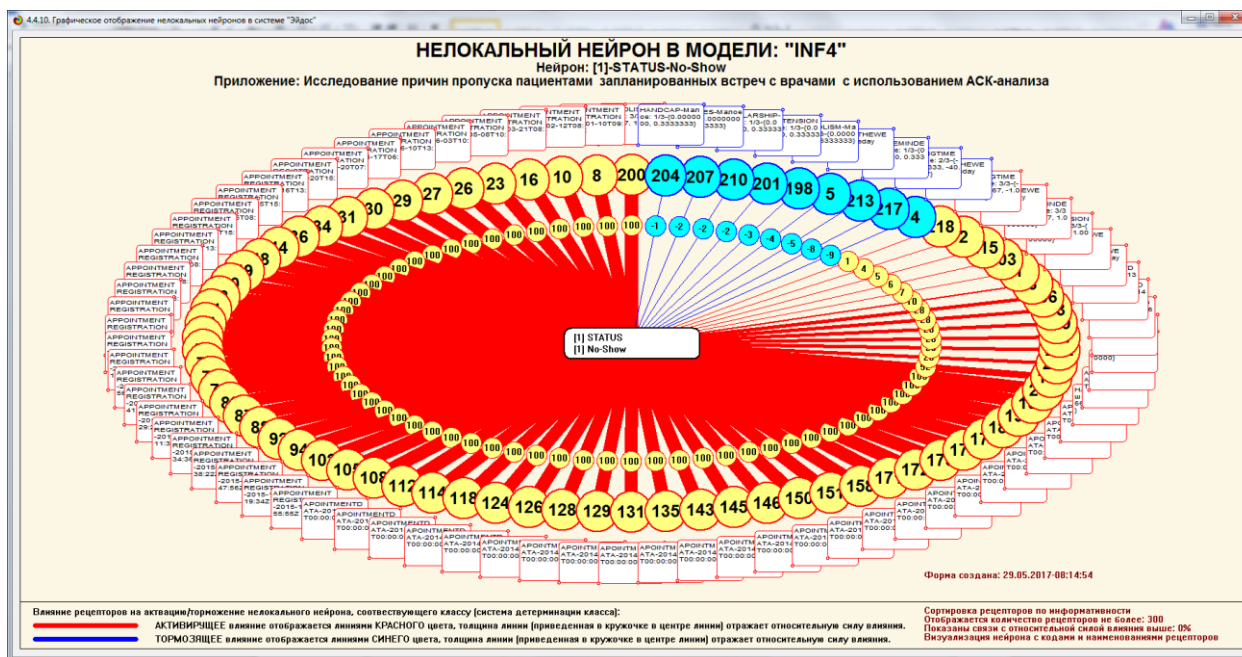


Рисунок 27. Графическое отображение нелокальных нейронов

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 27). 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 28). Необходимо отметить, что все указанные графические формы генерируются системой "Эйдос" автоматически в соответствии с созданной моделью[9].

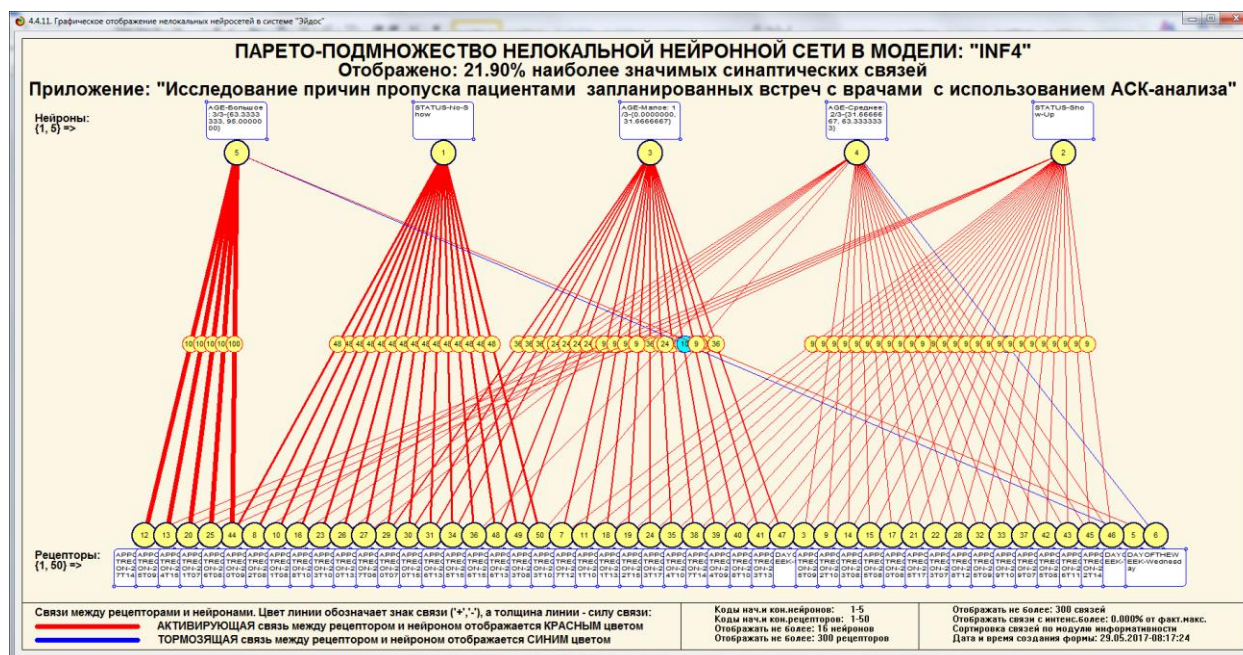


Рисунок 28 – Паретто – подмножеств нелокальной нейронной сети в системе «Эйдос - X»

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).

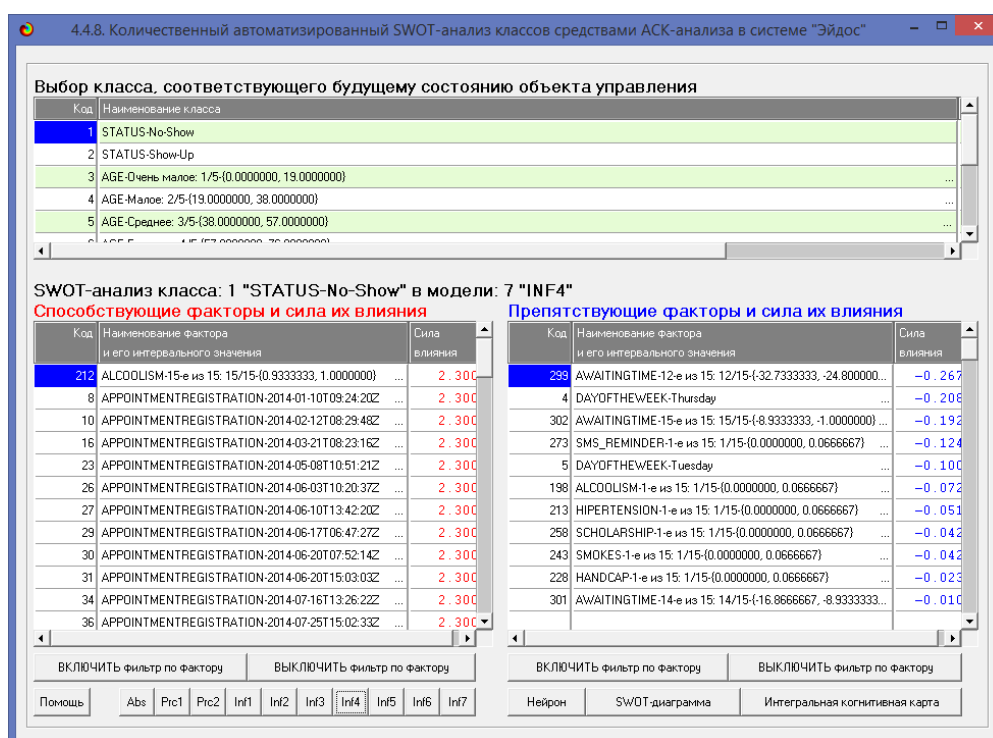


Рисунок 29. Пример SWOT-матрицы в модели INF4

На рисунке 30 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF4.

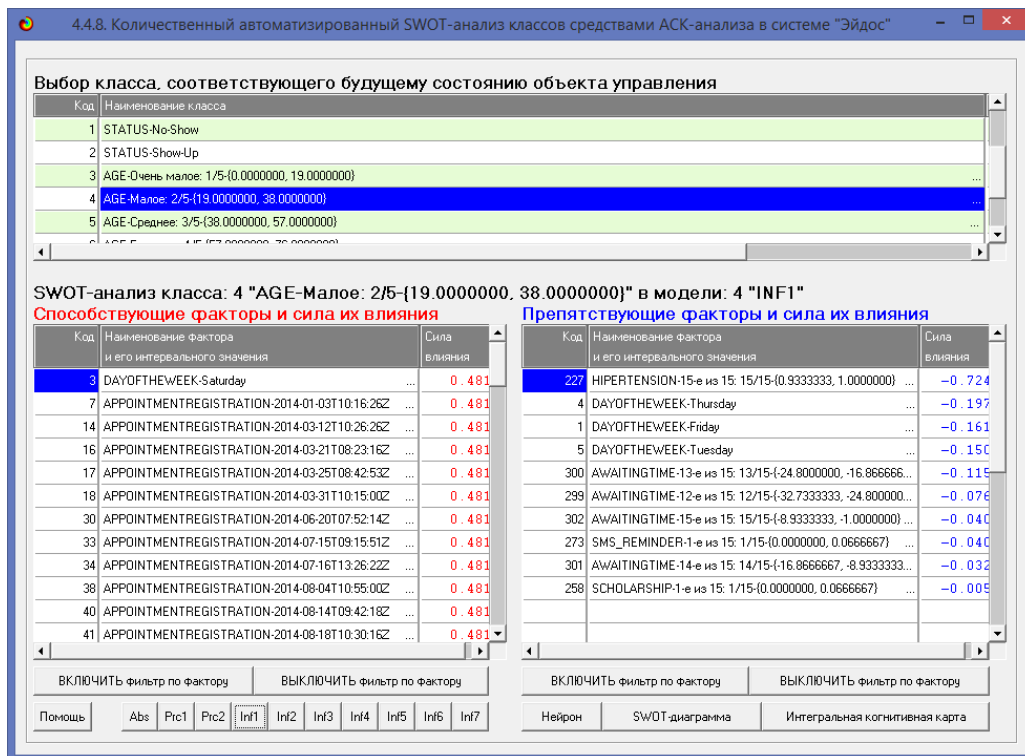


Рисунок 30. Табличная выходная форма количественного автоматизированного SWOT- и PEST-анализа средствами системы «Эйдос»

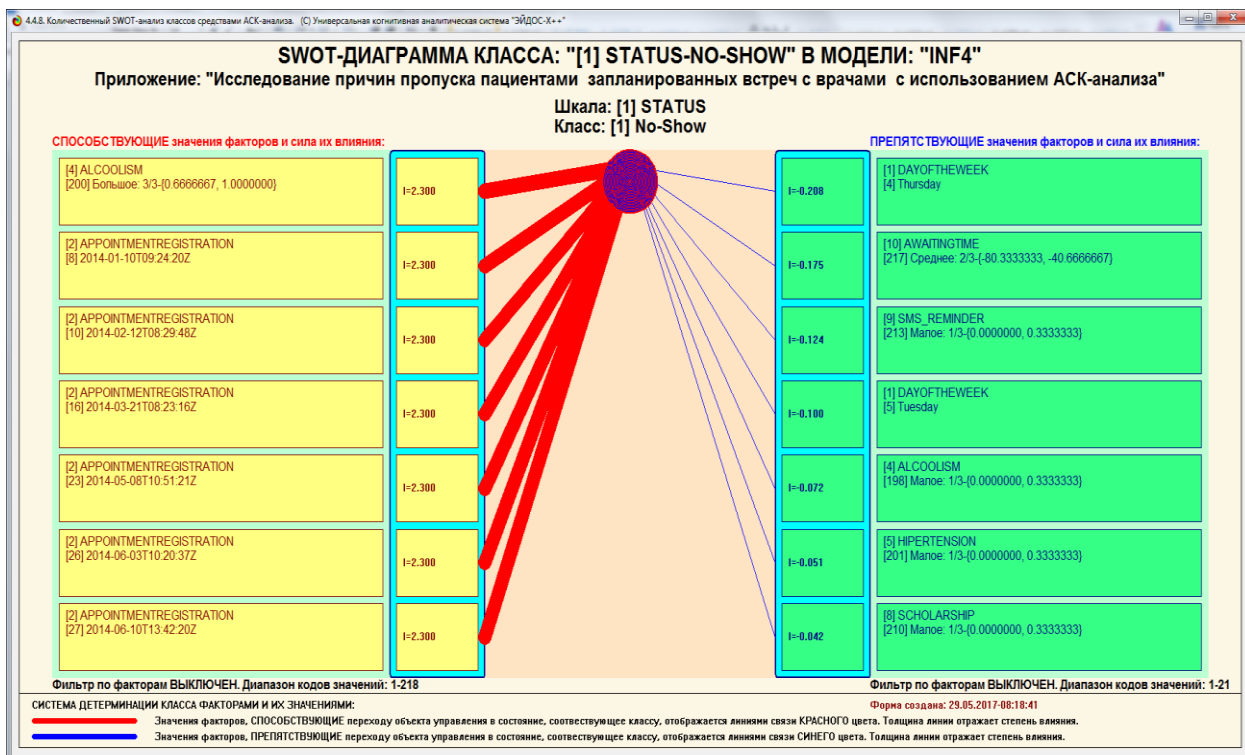


Рисунок 31. Графическая выходная форма количественного автоматизированного SWOT- и PEST-анализа средствами системы «Эйдос»

ЗАКЛЮЧЕНИЕ

Так как существует множество систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную базу репозитория UCI. В данной курсовой работе приводится развернутый пример использования базы данных репозитория UCI для оценки качества математических моделей, применяемых в АСК-анализе и его программном инструментарии системе искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении оказались модели INF4, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,724, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 73%. 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 25% 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.

На основе базы данных UCI, рассмотренной в данной курсовой работе, построить модели прогнозирования не с помощью АСК-анализа и реализующей его системы «Эйдос», а с применением других математических методов и реализующих их программных систем, то можно сопоставимо сравнить их качество.

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. 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.

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. с экрана. Яз. англ.

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-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.

8. Луценко Е.В., Боровко А.Ю. ПРОГНОЗИРОВАНИЕ КОЛИЧЕСТВА И КЛАССОВ СОЛНЕЧНЫХ ВСПЫШЕК НА ОСНОВЕ ИХ ПРЕДЫСТОРИИ ПО ДАННЫМ РЕПОЗИТОРИЯ UC1 С ПРИМЕНЕНИЕМ АСК- АНАЛИЗА И ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «ЭЙДОС» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2014/10/pdf/99.pdf>

9. 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.