

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 Османова Рустама Мехтиевича курса 3 очной формы обучения
направления подготовки Информационные системы и технологии,
направленность (профиль) академический бакалавриат.

Наименование темы «Разработка системно системно-когнитивной модели по
употреблению алкоголя среди молодежи США 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	5
2.	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
3.	Уровень и корректность использования в работе различных методов исследований	5
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	5
5.	Ясность, четкость, последовательность и обоснованность изложения	5
6.	Применение современных технологий обработки информации	4
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

Применение современных математических моделей и программного инструмента

Недостатки работы

Недостаточно качественное оформление

Итоговая оценка при защите _____

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

на тему:

**Разработка системно системно-когнитивной модели по употреблению
алкоголя среди молодежи США на основе данных репозитория UC1.**

выполнил студент группы ИТ1401

Османов Рустам Мехтиевич

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 _____

(дата)

Оценка _____

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

Студента: ИТ1401 группы 3 курса
Факультета прикладной информатики
Специальности: 09.03.02 Информационные системы и технологии
(шифр)

Османов Рустам Мехтиевич (Ф.И.О.)

Тема проекта: **Разработка системно системно-когнитивной модели по употреблению алкоголя среди молодежи США на основе данных репозитория UCI**

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://ic.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.

РЕФЕРАТ

Курсовая работа содержит 29 страниц, 24 рисунка, 3 используемых источников.

Ключевые слова: РАЗРАБОТКА СИСТЕМНО-КОГНИТИВНОЙ МОДЕЛИ. ВЕРБАЛЬНОЕ ОПИСАНИЕ. ИДЕНТИФИКАЦИЯ ИЗОБРАЖЕНИЙ ПО ИХ ВЕРБАЛЬНОМУ ОПИСАНИЮ.

Цель работы – исследование системно-когнитивной модели по употреблению алкоголя среди молодежи.

Основные результаты:

– **задача решена**, т.е. проведено исследование в ходе которого созданы модели системно-когнитивной модели по употреблению алкоголя среди молодежи .

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	6
1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	7
1.1. Описание решения.....	7
1.2. Преобразование исходных данных из HTML-формата в файл исходных данных MS Excel	7
1.3. Синтез и верификация статистических и интеллектуальных моделей ...	13
1.4. Виды моделей системы «Эйдос»	14
1.5. Результаты верификации моделей	16
2.РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	21
2.1. Решение задачи	21
2.2. Когнитивные функции	Ошибка! Закладка не определена.
2.3. SWOT и PEST матрицы и диаграммы	25
ЗАКЛЮЧЕНИЕ	28
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	29

ВВЕДЕНИЕ

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий.

Автоматическая обработка визуальной информации является одним из важнейших направлений в области искусственного интеллекта. Интерес к проблемам компьютерной обработки определяется расширением возможностей как самих компьютерных систем, так и разработкой новых технологий обработки, анализа и идентификации различных видов изображений.

Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено употребление алкоголя среди молодежи основанное на данных репозитория UCI/.

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 «**baza dannix sotrudnikov**» 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 сотрудников – «[baza dannix scholnicov.xls](#)», которую оставим без изменений.

Столбцы 1-15 описательные шкалы.
Столбец 9 является классификационной шкалой. Этот столбец показывает месячный доход сотрудников.

Обучающая выборка:

Рисунок 1 – baza dannah scholnicov.xls

n	возраст	внеклассные занятия	занятия спортом	доступ в интернет	свободное время	встречи с друзьями	Потребление алкоголя в будни	Потребление алкоголя в выходные	здоровье	кол-во занятий в школе
1	18	yes	no	no	3	4	1	1	3	6
2	17	no	no	yes	3	3	1	1	3	4
3	15	yes	no	yes	3	2	2	3	3	10
4	15	no	yes	yes	2	2	1	1	5	2
5	16	no	no	no	3	2	1	2	5	4
6	16	no	yes	yes	4	2	1	2	5	10
7	16	no	no	yes	4	4	1	1	3	0
8	17	yes	no	no	1	4	1	1	1	6
9	15	no	no	yes	2	2	1	1	1	0
10	15	no	yes	yes	5	1	1	1	5	0
11	15	no	no	yes	3	3	1	2	2	0
12	15	no	yes	yes	2	2	1	1	4	4
13	15	no	yes	yes	3	3	1	3	5	2
14	15	no	no	yes	4	3	1	2	3	2
15	15	no	no	yes	5	2	1	1	3	0
16	16	no	no	yes	4	4	1	2	2	4
17	16	no	yes	yes	2	3	1	2	2	6
18	16	yes	yes	no	3	2	1	1	4	4
19	17	no	yes	yes	5	5	2	4	5	16
20	16	no	yes	yes	1	3	1	3	5	4
21	15	no	no	yes	4	1	1	1	1	0
22	15	no	no	yes	4	2	1	1	5	0
23	16	no	yes	yes	5	1	1	3	5	2
24	16	no	yes	yes	4	4	2	4	5	0
25	15	yes	yes	yes	3	2	1	1	5	2
26	16	no	no	yes	2	2	1	3	5	14
27	15	no	no	yes	2	2	1	2	5	2
28	15	no	no	yes	2	4	2	4	1	4
29	16	yes	yes	yes	3	3	1	1	5	4
30	16	no	yes	yes	4	5	5	5	5	16
31	15	no	no	yes	4	2	3	4	5	0
32	15	no	yes	yes	3	1	1	1	5	0
33	15	no	yes	yes	5	2	1	1	5	0
34	15	no	yes	yes	3	2	1	1	2	0
35	16	no	no	yes	4	3	1	1	5	0
36	15	no	yes	no	5	1	1	1	5	0

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.

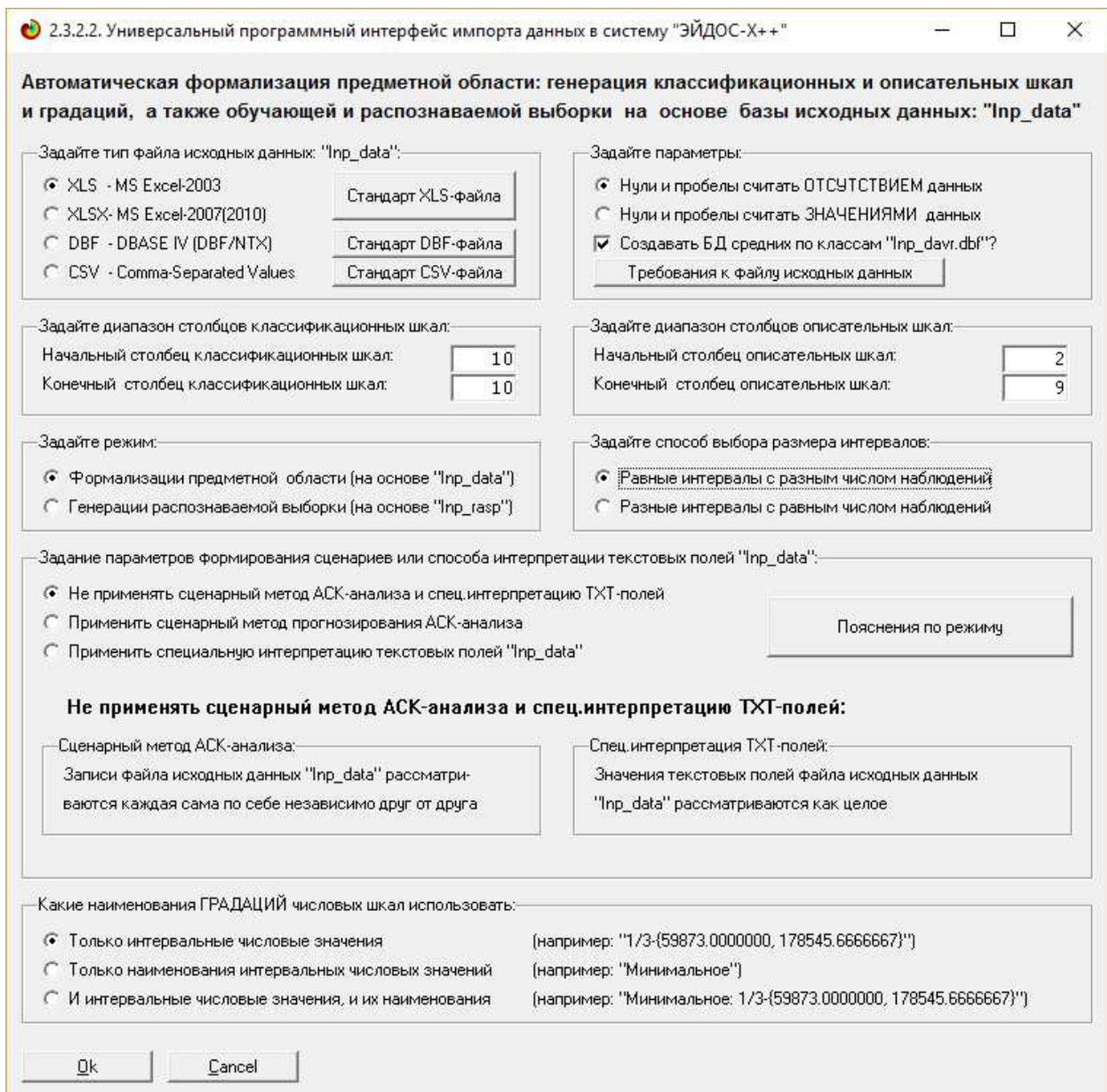


Рисунок 4. Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

В экранной форме, приведенной на рисунке 5, 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" – 10, "Конечный столбец классификационных шкал" – 10(последний столбец в таблице);

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 2, "Конечный столбец описательных шкал" – 9;

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

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

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 5. 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".

Стадии исполнения процесса

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

ПРОЦЕСС ФОРМАЛИЗАЦИИ ПРЕДМЕТНОЙ ОБЛАСТИ ЗАВЕРШЕН УСПЕШНО !!!

Прогноз времени исполнения

Начало: 23:15:10 Окончание: 23:15:14

Прошло: 0:00:03 Осталось: 0:00:00

100% Ok

Рисунок 6. 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 7).

Код шкалы	Наименование классификационной шкалы
1	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ ...
2	ЗДОРОВЬЕ ...
3	КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ ...

Код градации	Наименование градации классификационной шкалы	DEL
1	Малое: 1/3-{1.0000000, 2.3333333}	...
2	Среднее: 2/3-{2.3333333, 3.6666667}	...
3	Большое: 3/3-{3.6666667, 5.0000000}	...

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

Рисунок 7. Классификационные шкалы и градации (фрагмент)

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

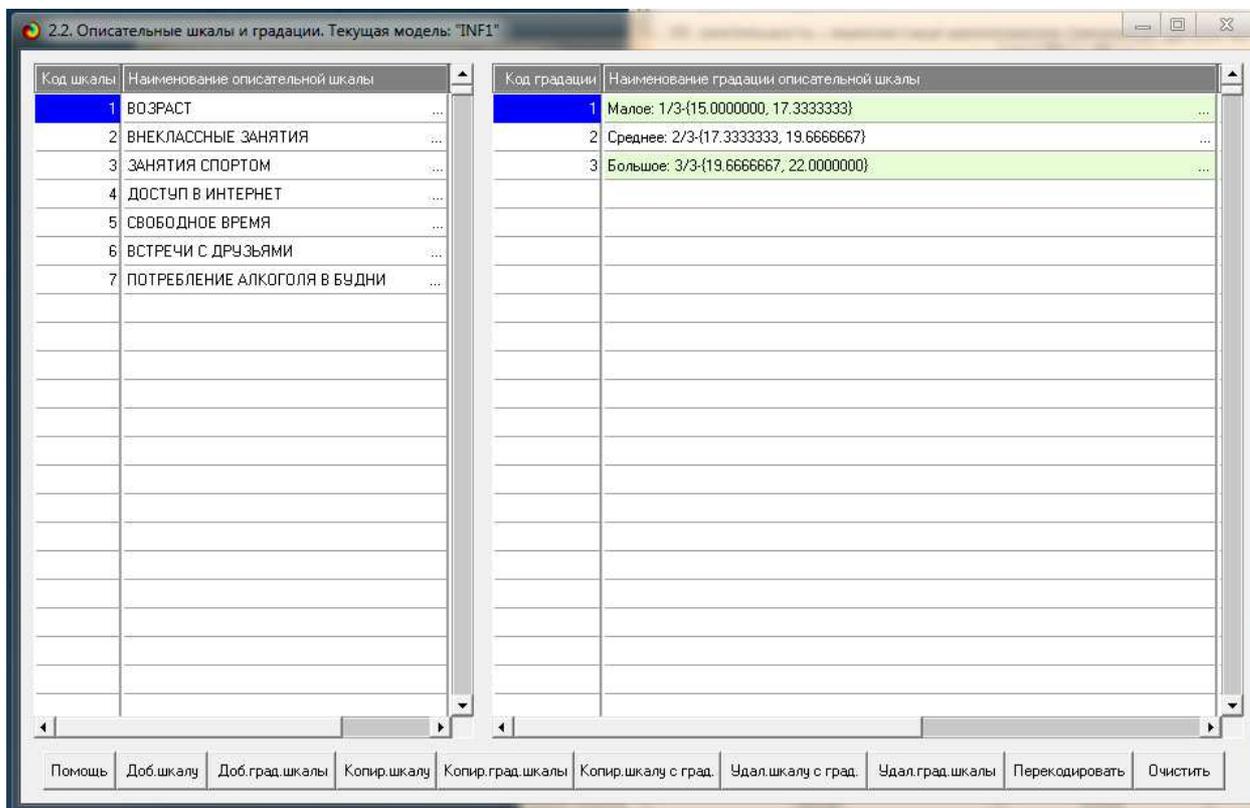


Рисунок 8. Описательные шкалы и градации (фрагмент)

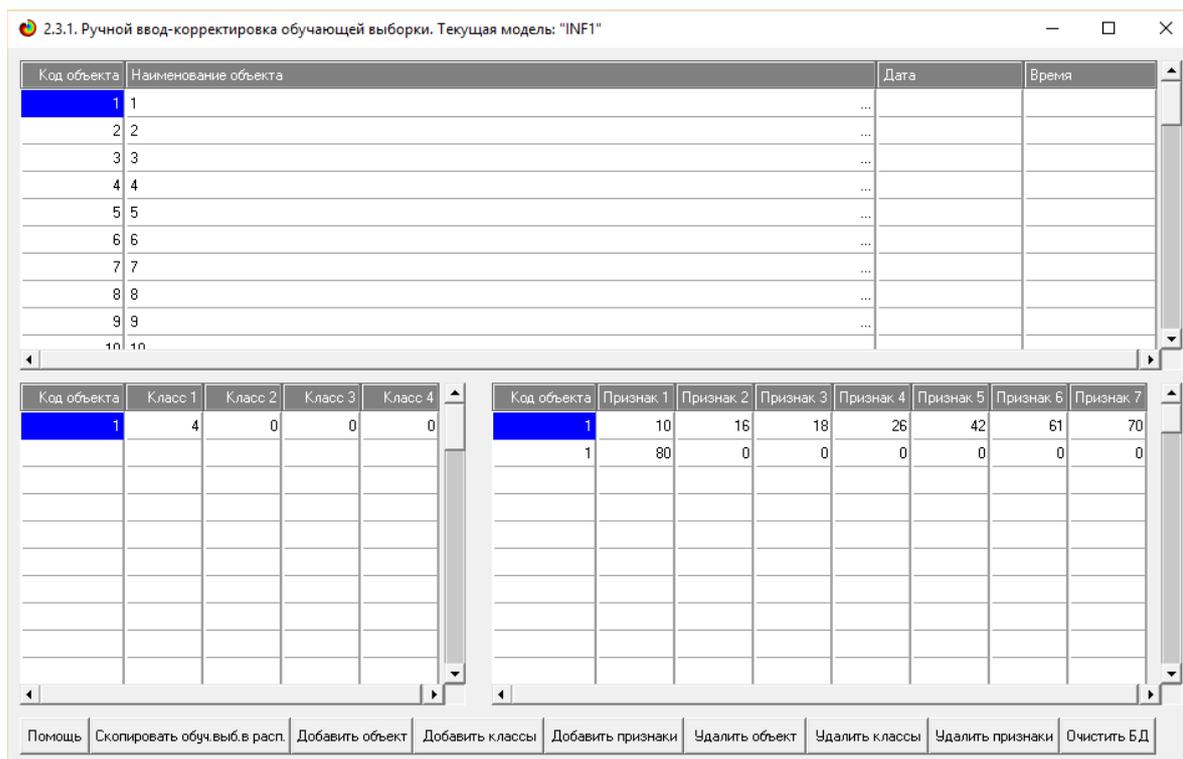


Рисунок 9. Обучающая выборка (фрагмент)

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 [11]).

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 10).

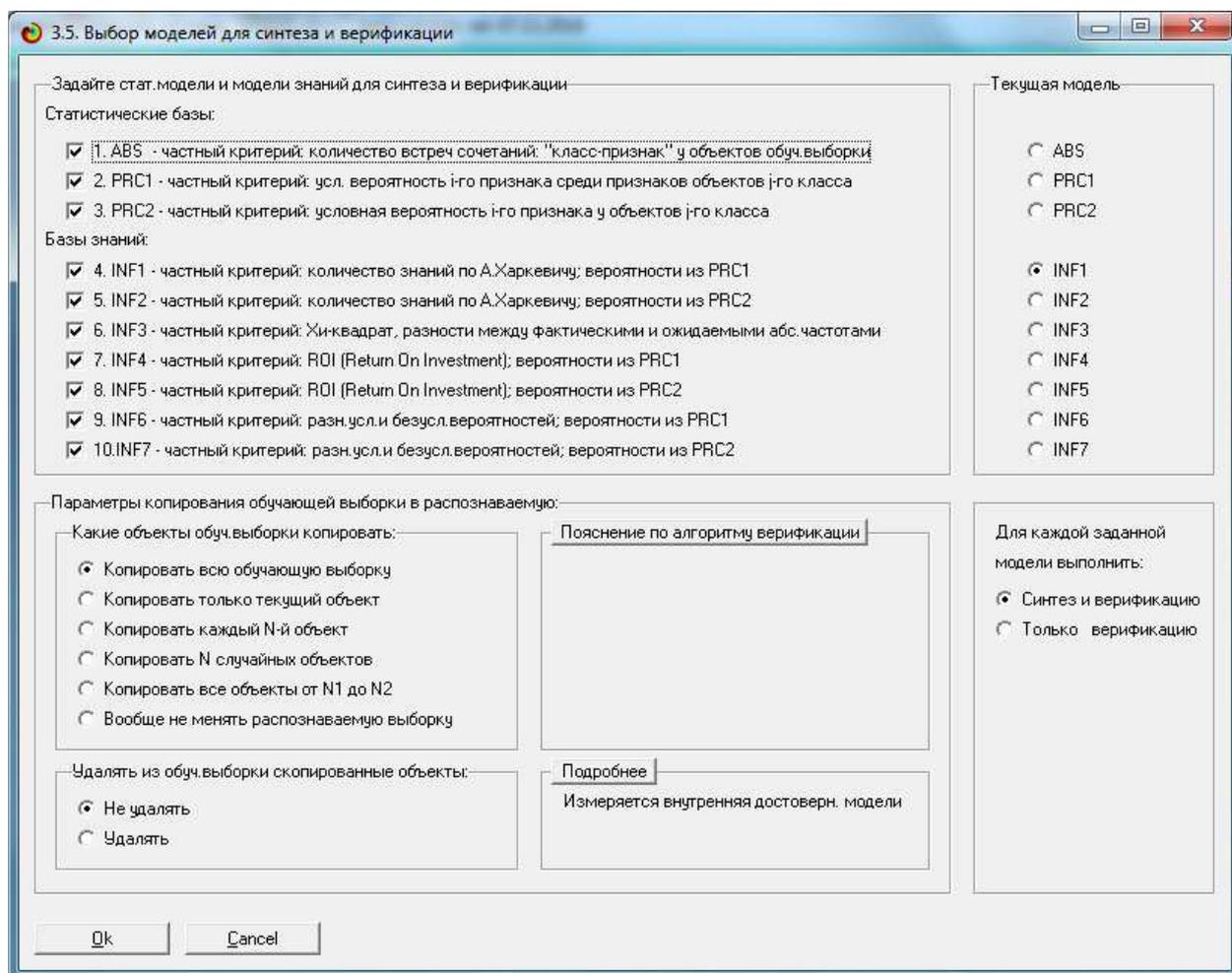


Рисунок 10. Выбор моделей для синтеза и верификации, а также текущей модели

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 11.

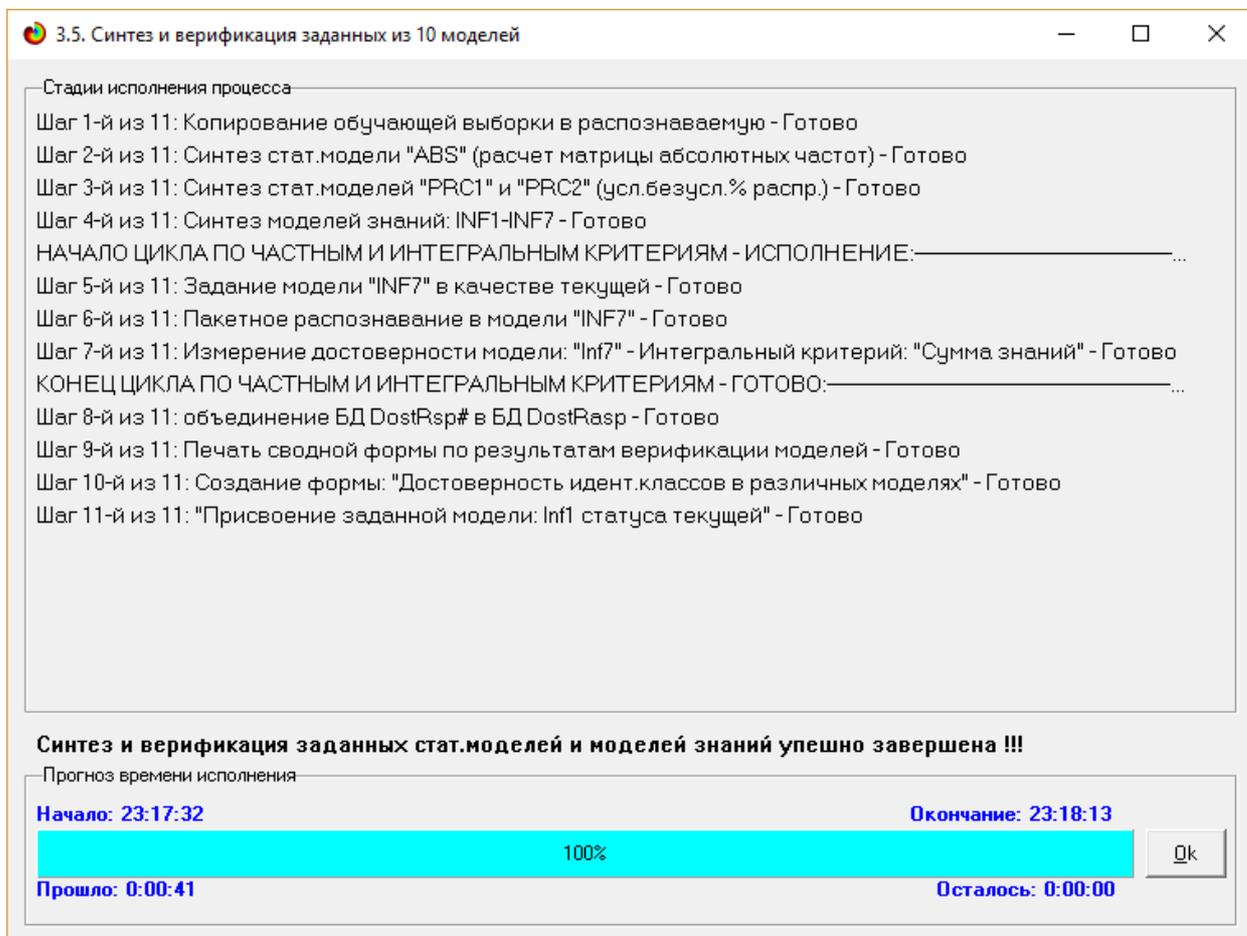


Рисунок 11. Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 41 секунд. 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)

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

Код признака	Наименование описательной шкалы и градации	1. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	2. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ СРЕДНЕЕ: 2/3 (2.3333333, 3.6666667)	3. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	4. ЗДОРОВЬЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	5. ЗДОРОВЬЕ СРЕДНЕЕ: 2/3 (2.3333333, 3.6666667)	6. ЗДОРОВЬЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	7. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ МАЛОЕ: 1/3 (0.0000000, 25.0000000)	8. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ СРЕДНЕЕ: 2/3 (25.0000000, 50.0000000)	9. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ БОЛЬШОЕ: 3/3 (50.0000000, 75.0000000)	Сумма	Среднее	Средн. квадрат. откл.
1	ВОЗРАСТ-Малое: 1/3(15.0000000, 17.3333333) ...	171	55	58	66	61	157	279	3	2	852	94.667	90.522
2	ВОЗРАСТ-Среднее: 2/3(17.3333333, 19.6666667) ...	63	24	19	25	28	53	103	2	1	318	35.333	32.600
3	ВОЗРАСТ-Большое: 3/3(19.6666667, 22.0000000) ...	2	1	2	1	2	2	5			15	1.667	1.500
4	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-но ...	203	70	71	79	79	186	338	4	2	1032	114.667	108.487
5	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-yes ...	33	10	8	13	12	26	49	1	1	153	17.000	15.984
6	ЗАНЯТИЯ СПОРТОМ-но ...	113	43	38	43	53	98	189	4	1	582	64.667	59.609
7	ЗАНЯТИЯ СПОРТОМ-yes ...	123	37	41	49	38	114	198	1	2	603	67.000	65.027
8	ДОСТУП В ИНТЕРНЕТ-но ...	40	13	13	9	19	38	66			198	22.000	21.863
9	ДОСТУП В ИНТЕРНЕТ-yes ...	196	67	66	83	72	174	321	5	3	987	109.667	102.740
10	СВОБОДНОЕ ВРЕМЯ-Малое: 1/3(1.0000000, 2.3333333) ...	60	11	12	21	23	39	81	1	1	249	27.667	27.455
11	СВОБОДНОЕ ВРЕМЯ-Среднее: 2/3(2.3333333, 3.6666667) ...	99	32	26	35	42	80	153	2	2	471	52.333	49.571
12	СВОБОДНОЕ ВРЕМЯ-Большое: 3/3(3.6666667, 5.0000000) ...	77	37	41	36	26	93	153	2	2	465	51.667	48.719
13	ВСТРЕЧИ С ДРУЗЬЯМИ-Малое: 1/3(1.0000000, 2.3333333) ...	102	14	10	24	34	68	124	1	1	378	42.000	45.495
14	ВСТРЕЧИ С ДРУЗЬЯМИ-Среднее: 2/3(2.3333333, 3.6666667) ...	84	35	11	30	37	63	127	1	2	290	43.333	41.644
15	ВСТРЕЧИ С ДРУЗЬЯМИ-Большое: 3/3(3.6666667, 5.0000000) ...	50	31	58	38	20	81	136	3		417	46.333	42.447
16	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Малое: 1/3(1.0000000, 2.3333333) ...	234	71	46	83	83	185	344	4	3	1053	117.000	114.606
17	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Среднее: 2/3(2.3333333, 3.6666667) ...	1	8	17	7	4	15	25	1		78	8.667	8.617
18	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Большое: 3/3(3.6666667, 5.0000000) ...	1	1	16	2	4	12	18			54	6.000	7.263

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

5.5. Модель: "4. INF1 - частный критерий: количество знаний по АХаркевичу; вероятности из PRCL1"

Код признака	Наименование описательной шкалы и градации	1. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	2. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ СРЕДНЕЕ: 2/3 (2.3333333, 3.6666667)	3. ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	4. ЗДОРОВЬЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	5. ЗДОРОВЬЕ СРЕДНЕЕ: 2/3 (2.3333333, 3.6666667)	6. ЗДОРОВЬЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	7. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ МАЛОЕ: 1/3 (0.0000000, 25.0000000)	8. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ СРЕДНЕЕ: 2/3 (25.0000000, 50.0000000)	9. КОЛ-ВО ЗАНЯТИЙ В ШКОЛЕ БОЛЬШОЕ: 3/3 (50.0000000, 75.0000000)	Сумма	Среднее	Средн. квадрат. откл.
1	ВОЗРАСТ-Малое: 1/3(15.0000000, 17.3333333) ...	0.003	-0.016	0.007	-0.001	-0.025	0.010	0.001	-0.064	-0.027	-0.110	-0.012	0.024
2	ВОЗРАСТ-Среднее: 2/3(17.3333333, 19.6666667) ...	-0.002	0.039	-0.038	0.004	0.048	-0.025	-0.003	0.140	0.076	0.240	0.027	0.056
3	ВОЗРАСТ-Большое: 3/3(19.6666667, 22.0000000) ...	-0.141	-0.004	0.244	-0.054	0.194	-0.103	0.007			0.142	0.016	0.127
4	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-но ...	-0.004	0.002	0.011	-0.005	-0.001	0.003	0.001	-0.030	-0.094	-0.118	-0.013	0.032
5	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-yes ...	0.028	-0.011	-0.085	0.032	0.007	-0.018	-0.007	0.154	0.333	0.432	0.048	0.124
6	ЗАНЯТИЯ СПОРТОМ-но ...	-0.009	0.032	-0.007	-0.017	0.060	-0.021	-0.002	0.171	-0.136	0.070	0.008	0.081
7	ЗАНЯТИЯ СПОРТОМ-yes ...	0.008	-0.034	0.007	0.016	-0.069	0.019	0.002	-0.328	0.095	-0.284	-0.032	0.120
8	ДОСТУП В ИНТЕРНЕТ-но ...	0.005	-0.010	-0.005	-0.188	0.078	0.025	0.007			-0.088	-0.010	0.072
9	ДОСТУП В ИНТЕРНЕТ-yes ...	-0.001	0.002	0.001	0.028	-0.018	-0.005	-0.001	0.064	0.064	0.134	0.015	0.030
10	СВОБОДНОЕ ВРЕМЯ-Малое: 1/3(1.0000000, 2.3333333) ...	0.067	-0.149	-0.114	0.029	0.065	-0.047	-0.001	-0.017	0.162	-0.095	-0.001	0.096
11	СВОБОДНОЕ ВРЕМЯ-Среднее: 2/3(2.3333333, 3.6666667) ...	0.019	0.002	-0.066	-0.015	0.053	-0.018	-0.002	0.002	0.182	0.156	0.017	0.069
12	СВОБОДНОЕ ВРЕМЯ-Большое: 3/3(3.6666667, 5.0000000) ...	-0.065	0.058	0.098	-0.001	-0.111	0.039	0.003	0.007		0.027	0.003	0.062
13	ВСТРЕЧИ С ДРУЗЬЯМИ-Малое: 1/3(1.0000000, 2.3333333) ...	0.107	-0.211	-0.325	-0.071	0.056	0.002	0.002	-0.164	0.015	-0.589	-0.065	0.141
14	ВСТРЕЧИ С ДРУЗЬЯМИ-Среднее: 2/3(2.3333333, 3.6666667) ...	0.028	0.100	-0.302	-0.003	0.074	-0.036	-0.001	-0.175	0.248	-0.068	-0.008	0.158
15	ВСТРЕЧИ С ДРУЗЬЯМИ-Большое: 3/3(3.6666667, 5.0000000) ...	-0.178	0.034	0.258	0.056	-0.165	0.029	0.000	0.187		0.221	0.025	0.142
16	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Малое: 1/3(1.0000000, 2.3333333) ...	0.038	0.000	-0.148	0.005	0.009	-0.006	0.000	-0.037	0.041	-0.098	-0.011	0.057
17	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Среднее: 2/3(2.3333333, 3.6666667) ...	-0.964	0.147	0.416	0.051	-0.142	0.025	-0.007	0.390		-0.082	-0.009	0.403
18	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Большое: 3/3(3.6666667, 5.0000000) ...	-0.834	-0.454	0.524	-0.260	-0.013	0.076	0.007			-0.954	-0.106	0.379

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. ПОТРЕБЛЕ. АЛКОГОЛЯ В ВЫХОДНЫЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	2. ПОТРЕБЛЕ. АЛКОГОЛЯ В ВЫХОДНЫЕ СРЕДНЕЕ: 2/3 (3.6666667, 5.0000000)	3. ПОТРЕБЛЕ. АЛКОГОЛЯ В ВЫХОДНЫЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	4. ЗДОРОВЬЕ МАЛОЕ: 1/3 (1.0000000, 2.3333333)	5. ЗДОРОВЬЕ СРЕДНЕЕ: 2/3 (2.3333333, 3.6666667)	6. ЗДОРОВЬЕ БОЛЬШОЕ: 3/3 (3.6666667, 5.0000000)	7. КОЛ ВО ЗАНЯТИЙ В ШКОЛЕ МАЛОЕ: 1/3 (0.0000000, 25.0000000)	8. КОЛ ВО ЗАНЯТИЙ В ШКОЛЕ СРЕДНЕЕ: 2/3 (25.0000000, 50.0000000)	9. КОЛ ВО ЗАНЯТИЙ В ШКОЛЕ БОЛЬШОЕ: 3/3 (50.0000000, 75.0000000)	Случая	Среднее	Средн. к-вадр. откл.
1	ВОЗРАСТ-Малое: 1/3 (15.0000000, 17.3333333)...	1.319	-2.519	1.200	-0.147	-4.428	4.575	0.752	-0.595	-0.157		2.525	
2	ВОЗРАСТ-Среднее: 2/3 (17.3333333, 19.6666666)...	-0.332	2.532	-2.200	0.311	3.580	-3.891	-0.853	0.658	0.195		2.253	
3	ВОЗРАСТ-Большое: 3/3 (19.6666667, 22.0000000)...	-0.987	-0.013	1.000	-0.165	0.848	-0.684	0.101	-0.063	-0.038		0.633	
4	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-но	-2.529	0.329	2.200	-1.122	-0.251	1.372	0.967	-0.354	-0.613		1.414	
5	ВНЕКЛАССНЫЕ ЗАНЯТИЯ-yes	2.529	-0.329	-2.200	1.122	0.251	-1.372	-0.967	0.354	0.613		1.414	
6	ЗАНЯТИЯ СПОРТОМ-но	-2.909	3.709	-0.800	-2.185	8.306	-6.122	-1.071	1.544	-0.473		4.151	
7	ЗАНЯТИЯ СПОРТОМ-yes	2.909	-3.709	0.800	2.185	-8.306	6.122	1.071	-1.544	0.473		4.151	
8	ДОСТУП В ИНТЕРНЕТ-но	0.567	-0.367	-0.200	-6.372	3.795	2.577	1.337	-0.835	-0.501		2.848	
9	ДОСТУП В ИНТЕРНЕТ-yes	-0.567	0.367	0.200	6.372	-3.795	-2.577	-1.337	0.835	0.501		2.848	
10	СВОБОДНОЕ ВРЕМЯ-Малое: 1/3 (1.0000000, 2.3333333)...	10.410	-5.810	-4.600	1.668	3.878	-5.547	-0.319	-0.051	0.370		5.149	
11	СВОБОДНОЕ ВРЕМЯ-Среднее: 2/3 (2.3333333, 3.6666667)...	5.197	0.203	-5.400	-1.567	5.830	-4.263	-0.820	0.013	0.808		3.714	
12	СВОБОДНОЕ ВРЕМЯ-Большое: 3/3 (3.6666667, 5.0000000)...	-15.608	5.608	10.000	-0.101	-9.709	9.810	1.139	0.038	-1.177		8.428	
13	ВСТРЕЧИ С ДРУЗЬЯМИ-Малое: 1/3 (1.0000000, 2.3333333)...	26.719	-11.519	-15.200	-5.347	4.972	0.375	0.552	-0.595	0.043		11.894	
14	ВСТРЕЧИ С ДРУЗЬЯМИ-Среднее: 2/3 (2.3333333, 3.6666667)...	6.329	8.671	-15.000	-0.278	7.051	-6.772	-0.367	-0.646	1.013		7.395	
15	ВСТРЕЧИ С ДРУЗЬЯМИ-Большое: 3/3 (3.6666667, 5.0000000)...	-33.048	2.848	30.200	5.625	-12.023	6.397	-0.185	1.241	-1.056		16.704	
16	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Малое: 1/3 (1.0000000, 2.3333333)...	24.289	-0.089	-24.200	1.248	2.137	-3.385	0.109	-0.443	0.334		12.214	
17	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Среднее: 2/3 (2.3333333, 3.6666667)...	-14.534	2.734	11.800	0.944	-1.990	1.046	-0.473	0.671	-0.197		6.751	
18	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Большое: 3/3 (3.6666667, 5.0000000)...	-9.754	-2.646	12.400	-2.192	-0.147	2.339	0.365	-0.228	-0.137		5.771	

1.5. Результаты верификации моделей

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

413.6. Обобщенная по достов. моделям при разн. 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 Рисбергера	Средня мод. уровнев. со. истинно-поло. решений (ST)	Средня мод. уровнев. со. истинно-отриц. решений (ST)	Средня мод. уровнев. со. ложно-поло. решений (SFP)	Средня мод. уровнев. со. ложно-отриц. решений (SF)
1. ABS - частный критерий: количество встреч со знаком: '1' класс	Корреляция абс. частот с обр...	1185	1171	74	2296	14	0.338	0.988	0.503	768.576	8.047	1365.933	1.231
1. ABS - частный критерий: количество встреч со знаком: '1' класс	Средня абс. частот по призна...	1185	1185		2370		0.333	1.000	0.500	651.079		426.392	
2. PRC1 - частный критерий: усл. вероятность 1-го признака сред...	Корреляция усл.отн. частот с о...	1185	1171	74	2296	14	0.338	0.988	0.503	768.576	8.047	1365.933	1.231
2. PRC1 - частный критерий: усл. вероятность 1-го признака сред...	Средня усл.отн. частот по призна...	1185	1185		2370		0.333	1.000	0.500	916.426		1783.913	
3. PRC2 - частный критерий: условная вероятность 1-го признака...	Корреляция усл.отн. частот с о...	1185	1171	74	2296	14	0.338	0.988	0.503	768.579	8.047	1365.947	1.231
3. PRC2 - частный критерий: условная вероятность 1-го признака...	Средня усл.отн. частот по призна...	1185	1185		2370		0.333	1.000	0.500	916.426		1783.913	
4. INF1 - частный критерий: количество знаний по А/Аркевичу, в...	Семантический резонанс зна...	1185	750	1368	1002	435	0.428	0.633	0.511	225.754	436.546	267.841	84.361
4. INF1 - частный критерий: количество знаний по А/Аркевичу, в...	Средня знаний	1185	720	1270	1100	465	0.396	0.608	0.479	60.762	280.865	160.262	26.47
5. INF2 - частный критерий: количество знаний по А/Аркевичу, в...	Семантический резонанс зна...	1185	750	1368	1002	435	0.428	0.633	0.511	225.754	436.546	267.842	84.361
5. INF2 - частный критерий: количество знаний по А/Аркевичу, в...	Средня знаний	1185	720	1270	1100	465	0.396	0.608	0.479	60.762	280.865	160.262	26.47
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	1185	777	1257	1113	408	0.411	0.656	0.505	266.830	446.479	313.900	104.49
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Средня знаний	1185	777	1257	1113	408	0.411	0.656	0.505	183.835	264.600	125.669	44.90
7. INF4 - частный критерий: ROI (Return On Investment), вероатно...	Семантический резонанс зна...	1185	715	1469	901	470	0.442	0.603	0.511	223.803	507.791	233.640	104.79
7. INF4 - частный критерий: ROI (Return On Investment), вероатно...	Средня знаний	1185	738	1115	1255	447	0.370	0.623	0.464	53.794	106.978	149.355	11.301
8. INF5 - частный критерий: ROI (Return On Investment), вероатно...	Семантический резонанс зна...	1185	715	1469	901	470	0.442	0.603	0.511	223.803	507.791	233.640	104.79
8. INF5 - частный критерий: ROI (Return On Investment), вероатно...	Средня знаний	1185	738	1115	1255	447	0.370	0.623	0.464	53.794	106.978	149.355	11.301
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Семантический резонанс зна...	1185	776	1377	993	409	0.439	0.655	0.525	265.858	463.860	279.701	104.88
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Средня знаний	1185	777	1377	993	408	0.385	0.656	0.495	70.622	222.376	221.207	18.59
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	1185	776	1377	993	409	0.439	0.655	0.525	265.858	463.860	279.701	104.88
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Средня знаний	1185	777	1377	993	408	0.385	0.656	0.495	70.622	222.376	221.207	18.59

Рисунок 12. Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF1, INF2, INF4 и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели составляет 0,892. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 90%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием (рисунок 13).

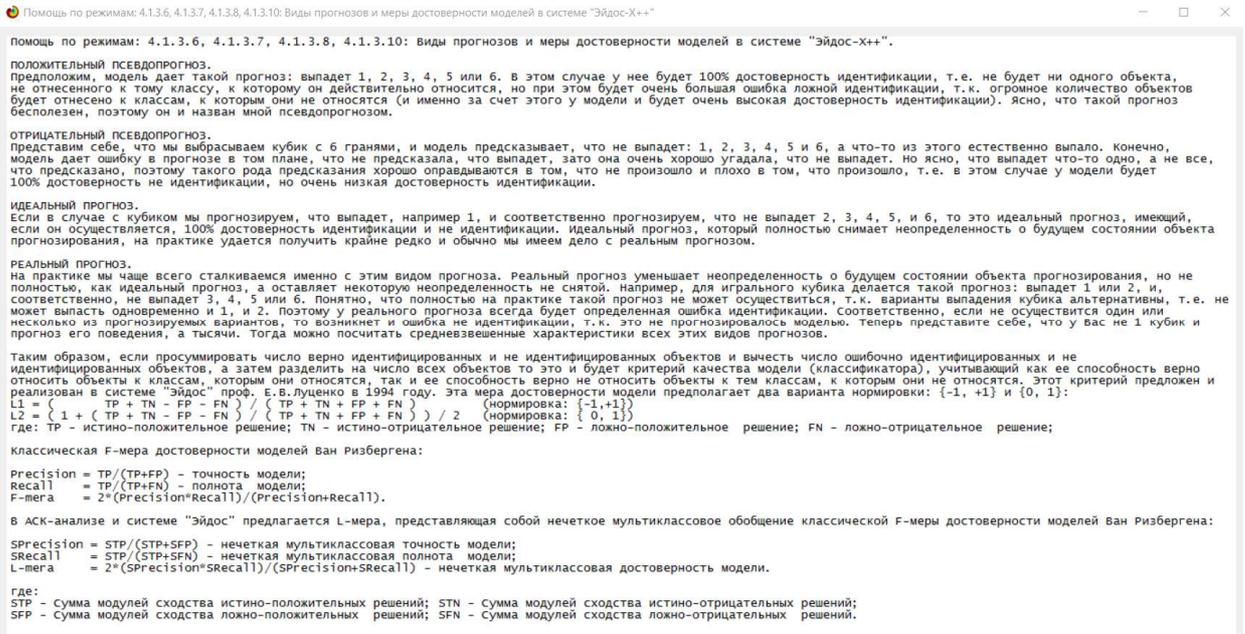


Рисунок 13. Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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.

Код класса	Наименование объекта распознанной выборки	MAX досто-верность	Модель с MAX достоверностью	Интегр. критерий с MAX досто-верностью	Модель: ABS, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: FRC1, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: FRC2, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF1, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF2, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF3, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF4, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF5, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF6, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: INF7, интегр. критер. РЕЗОНАНС ЗНАНИЙ	Модель: ABS, интегр. критер. Сумма ЗНАНИЙ	Модель: FRC1, интегр. критер. Сумма ЗНАНИЙ
1	1	0.500	ABS	Сумма	0.333	0.333	0.333	0.444	0.444	0.333	0.444	0.444	0.286	0.286	0.500	0
2	2	0.571	INF1	Резонанс	0.500	0.500	0.500	0.571	0.571	0.444	0.571	0.571	0.500	0.500	0.500	0
3	3	0.500	ABS	Резонанс	0.500	0.500	0.500	0.250	0.250	0.286	0.286	0.286	0.286	0.286	0.500	0
4	4	0.857	INF6	Резонанс	0.500	0.500	0.500	0.667	0.667	0.750	0.667	0.667	0.857	0.857	0.500	0
5	5	0.857	INF1	Резонанс	0.500	0.500	0.500	0.857	0.857	0.667	0.857	0.857	0.667	0.667	0.500	0
6	6	0.857	INF1	Резонанс	0.500	0.500	0.500	0.857	0.857	0.857	0.857	0.857	0.857	0.857	0.500	0
7	7	0.500	ABS	Резонанс	0.500	0.500	0.500	0.222	0.222	0.222	0.250	0.250	0.222	0.222	0.500	0
8	8	0.571	INF3	Резонанс	0.545	0.545	0.545	0.500	0.500	0.571	0.571	0.571	0.571	0.571	0.571	0
9	9	0.667	INF1	Резонанс	0.500	0.500	0.500	0.667	0.667	0.500	0.667	0.667	0.667	0.667	0.500	0
10	10	0.857	INF1	Резонанс	0.500	0.500	0.500	0.857	0.857	0.857	0.857	0.857	0.857	0.857	0.500	0
11	11	0.571	INF1	Резонанс	0.500	0.500	0.500	0.571	0.571	0.444	0.571	0.571	0.500	0.500	0.500	0
12	12	0.857	INF6	Резонанс	0.500	0.500	0.500	0.667	0.667	0.750	0.667	0.667	0.857	0.857	0.500	0
13	13	0.500	ABS	Резонанс	0.500	0.500	0.500	0.286	0.286	0.500	0.286	0.286	0.500	0.500	0.500	0
14	14	0.571	INF6	Резонанс	0.500	0.500	0.500	0.286	0.286	0.500	0.333	0.333	0.571	0.571	0.500	0
15	15	0.667	INF6	Резонанс	0.500	0.500	0.500	0.500	0.500	0.571	0.571	0.571	0.667	0.667	0.500	0
16	16	0.500	ABS	Резонанс	0.500	0.500	0.500	0.444	0.444	0.444	0.500	0.500	0.444	0.444	0.500	0
17	17	0.857	INF3	Резонанс	0.500	0.500	0.500	0.571	0.571	0.857	0.571	0.571	0.857	0.857	0.500	0
18	18	0.750	INF1	Резонанс	0.600	0.600	0.600	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.500	0
19	19	0.750	INF3	Резонанс	0.500	0.500	0.500	0.667	0.667	0.750	0.500	0.500	0.750	0.750	0.500	0
20	20	0.500	ABS	Резонанс	0.500	0.500	0.500	0.286	0.286	0.286	0.286	0.286	0.286	0.286	0.500	0
21	21	0.857	INF4	Резонанс	0.500	0.500	0.500	0.750	0.750	0.571	0.857	0.857	0.667	0.667	0.500	0
22	22	1.000	INF6	Резонанс	0.500	0.500	0.500	0.750	0.750	0.857	0.857	0.857	1.000	1.000	0.500	0

Рисунок 14. F-мера- верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF4

Из рисунка 11 видно, что:

– наиболее достоверная модель INF3 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.

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.

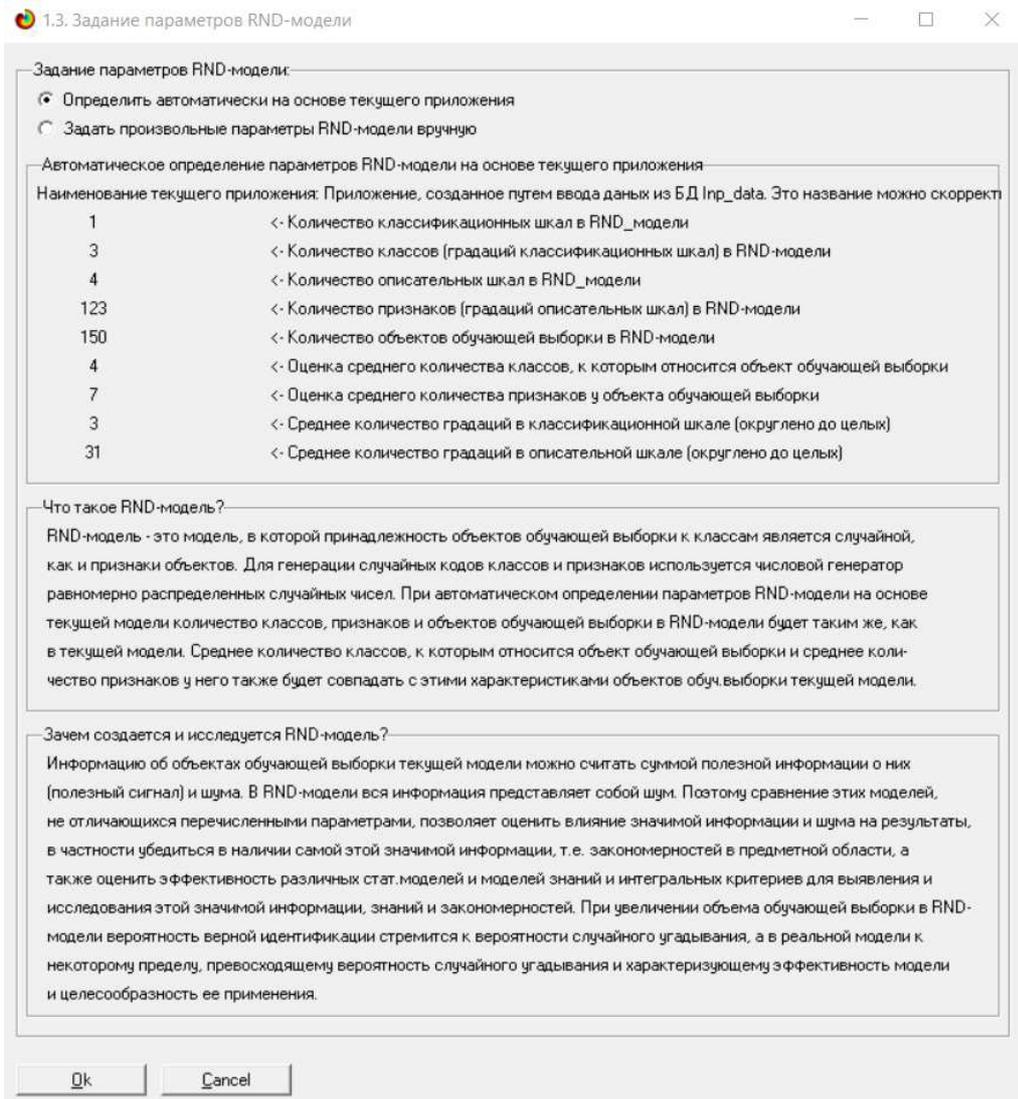


Рисунок 15. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 13 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF3.

Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели (рисунки 11 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.

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

Наименование модели и частного критерия	Интегральный критерий	Всего логически 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	Фигера Ван Риббергена	Сумма модальных истинно-положительных решений (ST)	Сумма модальных истинно-отрицательных решений (ST)	Сумма модальных ложно-положительных решений (SFP)	Сумма модальных ложно-отрицательных решений (SF)
1. ABS - частный критерий: количество встреч соитаний: "клас...	Корреляция абс частот с обр...	1185	1171	74	2296	14	0.338	0.988	0.503	768.576	8.047	1365.933	1.231
1. ABS - частный критерий: количество встреч соитаний: "клас...	Сумма абс частот по призна...	1185	1185		2370		0.333	1.000	0.500	651.079		426.392	
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн частот с о...	1185	1171	74	2296	14	0.338	0.988	0.503	768.576	8.047	1365.933	1.231
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Сумма усл.отн частот по призна...	1185	1185		2370		0.333	1.000	0.500	916.426		1783.913	
3. PRC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн частот с о...	1185	1171	74	2296	14	0.338	0.988	0.503	768.579	8.047	1365.947	1.231
3. PRC2 - частный критерий: условная вероятность его признака...	Сумма усл.отн частот по призна...	1185	1185		2370		0.333	1.000	0.500	916.426		1783.913	
4. INF1 - частный критерий: количество знаний по АХаркевичу: в...	Семантический резонанс зна...	1185	750	1368	1002	435	0.428	0.633	0.511	225.754	436.546	267.841	84.361
4. INF1 - частный критерий: количество знаний по АХаркевичу: в...	Сумма знаний	1185	720	1270	1100	465	0.396	0.608	0.479	60.762	280.865	160.262	26.471
5. INF2 - частный критерий: количество знаний по АХаркевичу: в...	Семантический резонанс зна...	1185	750	1368	1002	435	0.428	0.633	0.511	225.754	436.546	267.842	84.361
5. INF2 - частный критерий: количество знаний по АХаркевичу: в...	Сумма знаний	1185	720	1270	1100	465	0.396	0.608	0.479	60.762	280.865	160.262	26.471
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	1185	777	1257	1113	408	0.411	0.656	0.505	266.380	446.479	313.900	104.491
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Сумма знаний	1185	777	1257	1113	408	0.411	0.656	0.505	183.835	264.600	125.669	44.901
7. INF4 - частный критерий: ROI (Return On Investment): вероатно...	Семантический резонанс зна...	1185	715	1469	901	470	0.442	0.603	0.511	223.803	507.791	233.640	104.791
7. INF4 - частный критерий: ROI (Return On Investment): вероатно...	Сумма знаний	1185	738	1115	1255	447	0.370	0.623	0.464	53.794	106.978	149.355	11.301
8. INF5 - частный критерий: ROI (Return On Investment): вероатно...	Семантический резонанс зна...	1185	715	1469	901	470	0.442	0.603	0.511	223.803	507.791	233.640	104.791
8. INF5 - частный критерий: ROI (Return On Investment): вероатно...	Сумма знаний	1185	738	1115	1255	447	0.370	0.623	0.464	53.794	106.978	149.355	11.301
9. INF6 - частный критерий: разн усл и безуслов вероятностей, вер...	Семантический резонанс зна...	1185	776	1377	993	409	0.439	0.655	0.525	265.858	469.860	279.701	104.881
9. INF6 - частный критерий: разн усл и безуслов вероятностей, вер...	Сумма знаний	1185	777	1131	1239	408	0.385	0.656	0.485	70.622	222.376	221.207	18.591
10. INF7 - частный критерий: разн усл и безуслов вероятностей, ве...	Семантический резонанс зна...	1185	776	1377	993	409	0.439	0.655	0.525	265.858	469.860	279.701	104.881
10. INF7 - частный критерий: разн усл и безуслов вероятностей, ве...	Сумма знаний	1185	777	1131	1239	408	0.385	0.656	0.485	70.622	222.376	221.207	18.591

Рисунок 17. Достоверность статистических и когнитивных моделей, созданных на основе случайной выборки

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

- достоверность лучшей модели INF3, 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.

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.

4.2.1. Информационные портреты классов

Инф.портрет класса: 1 "ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В ВЫХОДНЫЕ-Малое: 1/3-{1.0000000, 2.3333333}" в модели: 6 "INF3"

Код	Наименование признака	Значимость
13	ВСТРЕЧИ С ДРУЗЬЯМИ-Малое: 1/3-{1.0000000, 2.3333333}	26.719
16	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Малое: 1/3-{1.0000000, 2.3333333}	24.289
10	СВОБОДНОЕ ВРЕМЯ-Малое: 1/3-{1.0000000, 2.3333333}	10.410
14	ВСТРЕЧИ С ДРУЗЬЯМИ-Среднее: 2/3-{2.3333333, 3.6666667}	6.329
11	СВОБОДНОЕ ВРЕМЯ-Среднее: 2/3-{2.3333333, 3.6666667}	5.197
7	ЗАНЯТИЯ СПОРТОМ-yes	2.909
5	ВНЕКЛАСНЫЕ ЗАНЯТИЯ-yes	2.529
1	ВОЗРАСТ-Малое: 1/3-{15.0000000, 17.3333333}	1.319
8	ДОСТУП В ИНТЕРНЕТ-no	0.567
2	ВОЗРАСТ-Среднее: 2/3-{17.3333333, 19.6666667}	-0.332
9	ДОСТУП В ИНТЕРНЕТ-yes	-0.567
3	ВОЗРАСТ-Большое: 3/3-{19.6666667, 22.0000000}	-0.987
4	ВНЕКЛАСНЫЕ ЗАНЯТИЯ-no	-2.529
6	ЗАНЯТИЯ СПОРТОМ-no	-2.909
18	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Большое: 3/3-{3.6666667, 5.0000000}	-9.754
17	ПОТРЕБЛЕНИЕ АЛКОГОЛЯ В БУДНИ-Среднее: 2/3-{2.3333333, 3.6666667}	-14.534
12	СВОБОДНОЕ ВРЕМЯ-Большое: 3/3-{3.6666667, 5.0000000}	-15.608
15	ВСТРЕЧИ С ДРУЗЬЯМИ-Большое: 3/3-{3.6666667, 5.0000000}	-33.048

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 MS Excel ВКЛ.фильтр по фактору ВЫКЛ.фильтр по фактору Вписать в окно Показать ВСЕ

Рисунок 11- пакетное распознавание

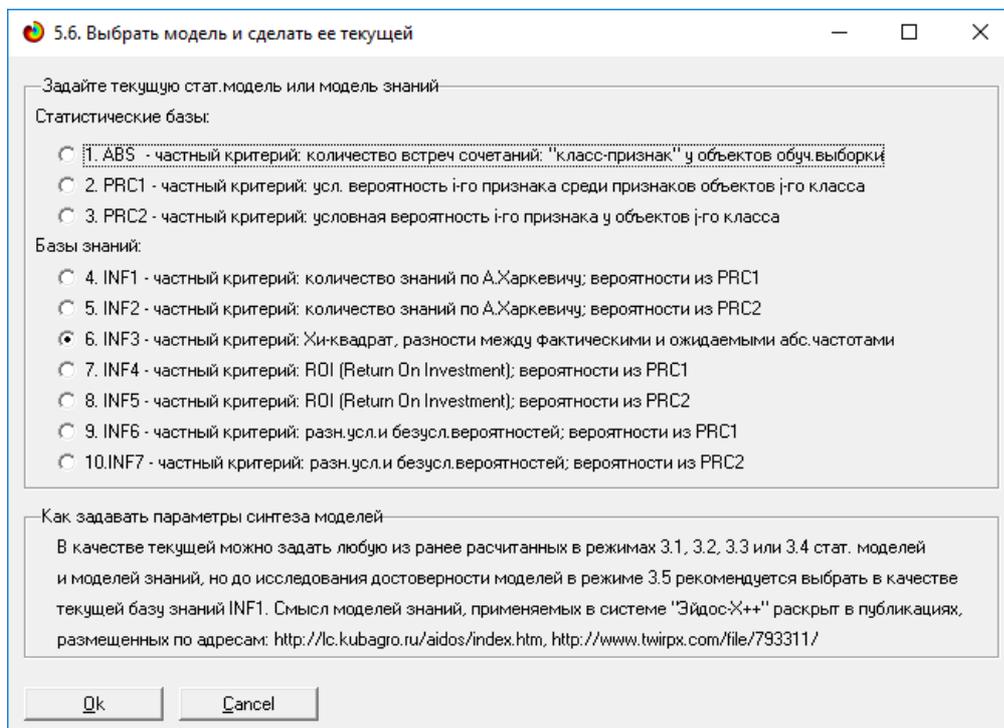


Рисунок 18. Экранные формы режима задания модели в качестве текущей

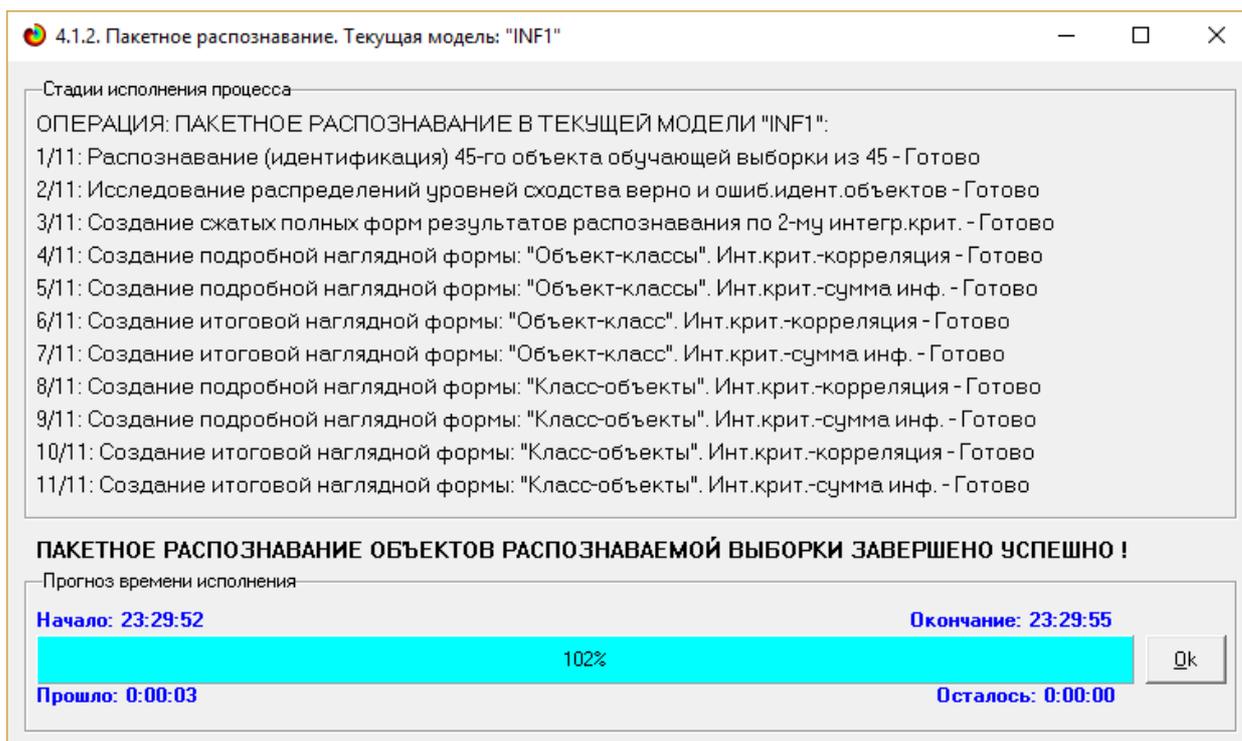


Рисунок 19. Экранная форма режима пакетного распознавания в текущей модели

В результате пакетного распознавания в текущей модели создается ряд баз данных, которые визуализируются в выходных экранных формах, отражающих результаты решения задачи идентификации и прогнозирования.

Режим 4.1.3 системы «Эйдос» обеспечивает отображение результатов идентификации и прогнозирования в различных формах:

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

Ниже кратко рассмотрим некоторые из них.

На рисунках 17 и 18 приведены примеры прогнозов высокой и низкой достоверности частоты и классов ирисов в наиболее достоверной модели INF3 на основе наблюдения предыстории их развития:

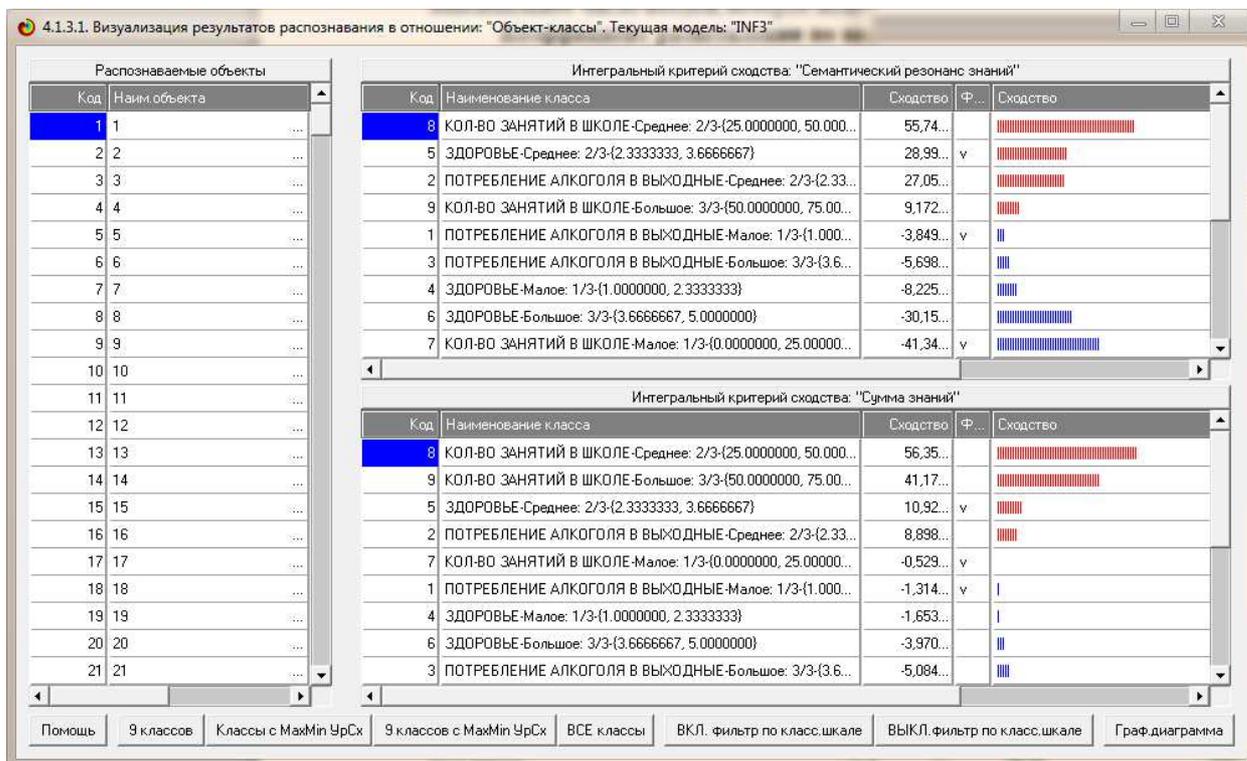


Рисунок 20. Пример идентификации классов в модели INF3

Рассмотрим режим 4.5, в котором реализована возможность визуализации когнитивных функций для любых моделей и любых сочетаний классификационных и описательных шкал (рисунок 22)

4.5. Визуализация когнитивных функций

Что такое когнитивная функция:

Визуализация прямых, обратных, позитивных, негативных, полностью и частично редуцированных когнитивных функций Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам. Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека. Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации; негативные, отражающие чему препятствуют система детерминации; средневзвешенные, отражающие совокупное влияние всех значений факторов на поведение объекта (причем в качестве весов наблюдений используется количество информации в значении аргумента о значениях функций) различной степенью редукции или степенью детерминации, которая отражает в графической форме (в форме полосы) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала. Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию. Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой "Эйдос". Необходимо отметить, что на вид функций влияния математической моделью СК-анализа не накладывается никаких ограничений, в частности, они могут быть и не дифференцируемые. См.: Луценко Е.В. Метод визуализации когнитивных функций - новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. - Краснодар: КубГАУ, 2011. - №03(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077. , 2,688 у.п.л. - Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>

Задайте нужный режим:

Визуализации когнитивных функций | Скачать подборку публикаций по когнитивным функциям

Литератур.ссылки на статьи по когнитивным функциям | Скачать подборку публикаций по управлению знаниями

Рисунок 22. Экранная форма режима 4.5 системы «Эйдос-X++» «Визуализация когнитивных функций»

Применительно к задаче, рассматриваемой в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния. Когнитивным функциям посвящено много работ автора [9], но наиболее новой и обобщающей из них является работа. Поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 23. приведены визуализации всех когнитивных функций данного приложения для модели INF1.

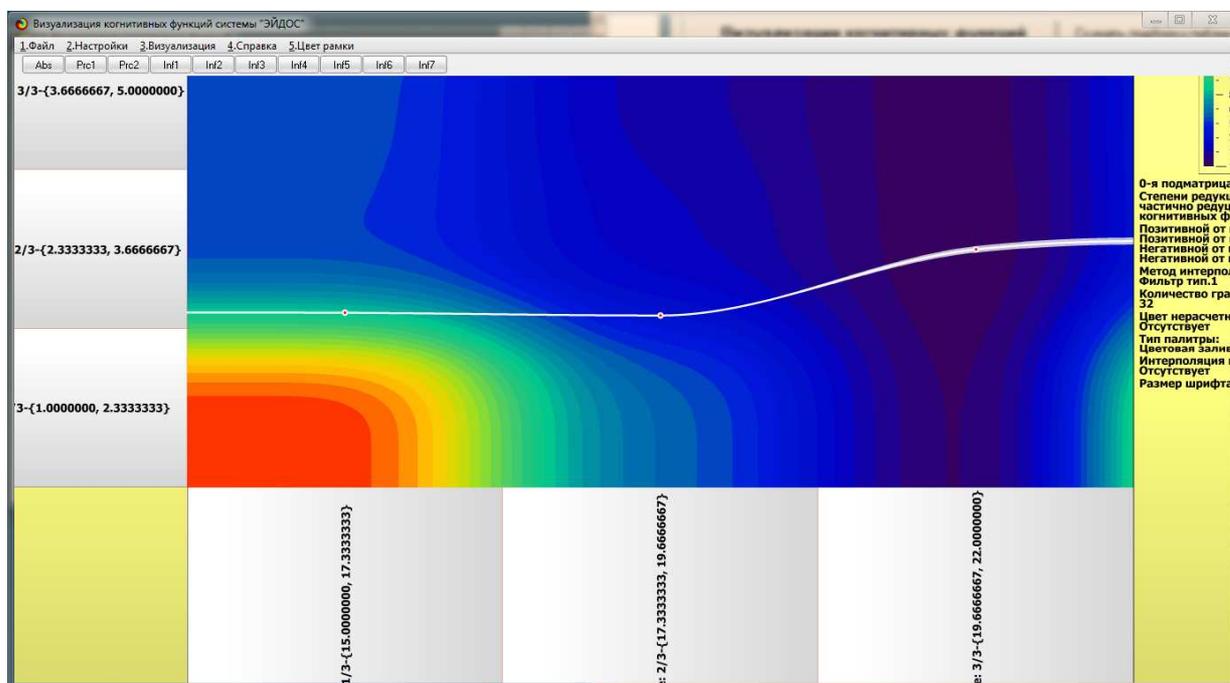


Рисунок 23. Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF1

2.2. 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 24).

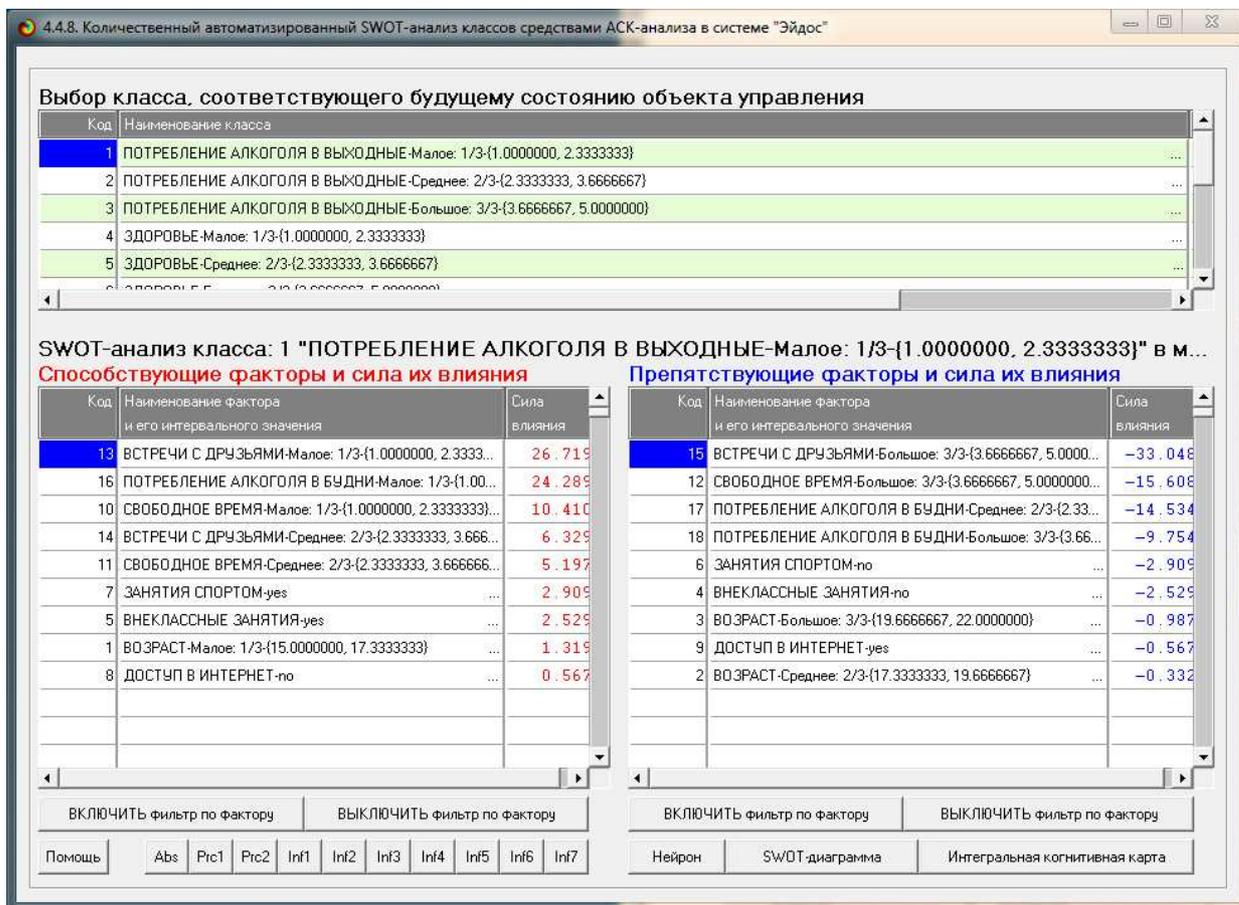


Рисунок 24. Пример SWOT-матрицы в модели INF2

На рисунке 25 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF4.

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 0,960, 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 21% 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. Методика использования репозитория UCI для оценки качества математических моделей систем искусственного интеллекта / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: http://lc.kubagro.ru/My_training_schedule.doc

КубГАУ, 2003. – №02(002). С. 120 – 145. – IDA [article ID]: 0020302012. – Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 у.п.л.

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

3. Луценко Е.В. 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.

4. 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