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 Каминского Владислава Николаевича курса 3 очной формы обучения направления подготовки Информационные системы и технологии.
направленность (профиль) академический бакалавриат.
Наименование темы «Исследование и разработка системно-когнитивной модели статистики штрафных бросков nba в определенный период времени».
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:
Исследование и разработка системно-когнитивной модели статистики штрафных бросков nba в определенный период времени
выполнил студент группы ИТ1401
Каминский Владислав Николаевич

Допущена к защите _____

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

_____
(подпись, расшифровка подписи)

Нормоконтролер Николаева Ирина Валентиновна, к.т.н., доцент
(подпись, расшифровка подписи)

Защищена _____ Оценка _____
(дата)
Члены комиссии _____ В.И. Лойко
_____ Е.В. Луценко
_____ И. В. Николаева
(подпись, дата, расшифровка подписи)

Краснодар
2017 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:
Зав. кафедрой КТС_____ В. И. Лойко

ЗАДАНИЕ
на курсовую работу

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

Каминского Владислава Николаевич
(Ф.И.О.)

Тема проекта: **Разработка и исследование системно когнитивной модели зависимости количества преступлений в Швеции от количества населения**

Содержание задания: Проанализировать методы формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами, принятия решений и исследования моделируемой предметной области путем исследования ее модели

Объем работы:

а) пояснительная записка к работе _____ листа формата А4
б) графическая часть _____ лист формата А4
Рекомендуемая литература: _ Луценко Е.В. Лабораторный практикум по интеллектуальным информационным системам: Учебное пособие для бакалавриата. 7-е изд., перераб. и доп.- Краснодар: КубГАУ – 2016, – 615 с., в электронном виде на сайте автора: http://lc.kubagro.ru/aidos/p14.htm
Срок выполнения проекта: с “ ” _____ по “ ” _____2017 г.
Срок защиты: “___” _____2017 г.
Дата выдачи задания: “ ” _____2017 г.
Дата сдачи проекта на кафедру: “___” _____2017 г.
Руководитель проекта:_____ Луценко Евгений Вениаминович,
д.э.н., к.т.н., профессор
(подпись, Ф.И.О., звание, степень)
Задание принял студент _____ “___” _____2017 г.
(подпись, дата)

Краснодар
2017 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 5
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ 6
1.1. Описание решения 6
1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel 6
1.3. Синтез и верификация статистических и интеллектуальных моделей 12
1.4. Виды моделей системы «Эйдос» 14
1.5. Результаты верификации моделей 15
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ 22
2.1. Решение задачи идентификации 22
2.2. Когнитивные функции 25
2.3. SWOT и PEST матрицы и диаграммы 26
2.4. НЕЛОКАЛЬНЫЕ НЕЙРОННЫЕ СЕТИ И НЕЙРОНЫ. 29
2.5 КЛАСТЕРНЫЙ И КОНСТРУКТИВНЫЙ АНАЛИЗ 30
ЗАКЛЮЧЕНИЕ 31
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 32

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 по поиску зависимости состава продуктов быстрого питания от выбранной категории.

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

В данной курсовой работе использована база данных «menu» из банка исходных данных по задачам искусственного интеллекта – репозитория Kaggle.

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

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

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

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

Из электронного ресурса баз данных Kaggle возьмем базу штрафных бросков – «free_throws.csv», которую оставим без изменений.

Ссылка для скачивания исходных данных: http:// /www.kaggle.com/sebastianmantey/nba-free-throws

Обучающая выборка:

Таблица 1 – crime.xls

A	B	C	D	E	F	G	H	I	J	K	L
	MIL - DET	261101008	2.0	Charlie Vi	regular	48 - 39	2006 - 200	1	2:54	Charlie Vi	105 - 97
	MIL - DET	261101008	2.0	Michael R	regular	51 - 41	2006 - 200	1	1:29	Michael R	105 - 97
	MIL - DET	261101008	2.0	Andrew B	regular	51 - 41	2006 - 200	0	1:29	Andrew B	105 - 97
	MIL - DET	261101008	2.0	Andrew B	regular	51 - 41	2006 - 200	0	1:29	Andrew B	105 - 97
	MIL - DET	261101008	3.0	Chauncey	regular	55 - 49	2006 - 200	1	10:33	Chauncey	105 - 97
	MIL - DET	261101008	3.0	Chauncey	regular	55 - 50	2006 - 200	1	9:51	Chauncey	105 - 97
	MIL - DET	261101008	3.0	Chauncey	regular	55 - 51	2006 - 200	1	9:51	Chauncey	105 - 97
	MIL - DET	261101008	3.0	Chauncey	regular	57 - 52	2006 - 200	1	9:07	Chauncey	105 - 97
	MIL - DET	261101008	3.0	Chauncey	regular	57 - 53	2006 - 200	1	9:07	Chauncey	105 - 97
	MIL - DET	261101008	3.0	Richard Ha	regular	61 - 58	2006 - 200	1	7:11	Richard Ha	105 - 97
	MIL - DET	261101008	3.0	Richard Ha	regular	61 - 58	2006 - 200	0	7:11	Richard Ha	105 - 97
	MIL - DET	261101008	3.0	Michael R	regular	62 - 58	2006 - 200	1	7:03	Michael R	105 - 97
	MIL - DET	261101008	3.0	Charlie Vi	regular	63 - 58	2006 - 200	1	7:03	Charlie Vi	105 - 97
	MIL - DET	261101008	3.0	Charlie Vi	regular	64 - 58	2006 - 200	1	7:03	Charlie Vi	105 - 97
	MIL - DET	261101008	3.0	Tayshaun	regular	73 - 64	2006 - 200	0	1:53	Tayshaun	105 - 97
	MIL - DET	261101008	3.0	Tayshaun	regular	73 - 65	2006 - 200	1	1:53	Tayshaun	105 - 97
	MIL - DET	261101008	3.0	Michael R	regular	76 - 65	2006 - 200	1	1:40	Michael R	105 - 97
	MIL - DET	261101008	3.0	Ronald Mu	regular	78 - 68	2006 - 200	1	0:31	Ronald Mu	105 - 97
	MIL - DET	261101008	3.0	Ronald Mu	regular	78 - 69	2006 - 200	1	0:31	Ronald Mu	105 - 97
	MIL - DET	261101008	4.0	Jason Max	regular	80 - 72	2006 - 200	0	11:23	Jason Max	105 - 97
	MIL - DET	261101008	4.0	Jason Max	regular	80 - 73	2006 - 200	1	11:23	Jason Max	105 - 97
	MIL - DET	261101008	4.0	Charlie Be	regular	85 - 75	2006 - 200	1	7:50	Charlie Be	105 - 97
	MIL - DET	261101008	4.0	Charlie Be	regular	86 - 75	2006 - 200	1	7:50	Charlie Be	105 - 97
	MIL - DET	261101008	4.0	Chauncey	regular	88 - 78	2006 - 200	1	7:01	Chauncey	105 - 97
	MIL - DET	261101008	4.0	Steve Blak	regular	88 - 78	2006 - 200	0	6:41	Steve Blak	105 - 97
	MIL - DET	261101008	4.0	Richard Ha	regular	96 - 83	2006 - 200	1	4:09	Richard Ha	105 - 97
	MIL - DET	261101008	4.0	Richard Ha	regular	96 - 84	2006 - 200	1	4:09	Richard Ha	105 - 97
	MIL - DET	261101008	4.0	Richard Ha	regular	96 - 85	2006 - 200	1	3:46	Richard Ha	105 - 97
	MIL - DET	261101008	4.0	Richard Ha	regular	96 - 86	2006 - 200	1	3:46	Richard Ha	105 - 97
	MIL - DET	261101008	4.0	Andrew B	regular	96 - 86	2006 - 200	0	2:55	Andrew B	105 - 97
	MIL - DET	261101008	4.0	Andrew B	regular	97 - 86	2006 - 200	1	2:55	Andrew B	105 - 97
	MIL - DET	261101008	4.0	Michael R	regular	98 - 86	2006 - 200	1	2:31	Michael R	105 - 97
	MIL - DET	261101008	4.0	Michael R	regular	99 - 86	2006 - 200	1	2:31	Michael R	105 - 97
	MIL - DET	261101008	4.0	Tayshaun	regular	99 - 86	2006 - 200	0	1:52	Tayshaun	105 - 97
	MIL - DET	261101008	4.0	Tayshaun	regular	99 - 87	2006 - 200	1	1:52	Tayshaun	105 - 97
	MIL - DET	261101008	4.0	Tayshaun	regular	99 - 87	2006 - 200	0	1:32	Tayshaun	105 - 97
	MIL - DET	261101008	4.0	Tayshaun	regular	99 - 88	2006 - 200	1	1:32	Tayshaun	105 - 97
	MIL - DET	261101008	4.0	Michael R	regular	102 - 88	2006 - 200	1	1:15	Michael R	105 - 97
	MIL - DET	261101008	4.0	Andrew B	regular	102 - 91	2006 - 200	0	0:44	Andrew B	105 - 97

Всего в базе 11 столбцов, 2 классификационные и 9 описательных.

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 csv-файла в xls-файл, для удобства дальнейшего редактирования и выполним следующие операции.

Конвертируем исходный файл из csv формата в xls и запишем его с именем: Inp_data.xls в папку: c:\Aidos-X\AID_DATA\Inp_data\. В результате получим таблицу исходных данных, полностью подготовленную для обработки в системе «Эйдос» и записанную в нужную папку в виде файла нужного типа с нужным именем.

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

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода

данных из внешних баз данных табличного вида, т.е. режимом 2.3.2.2.

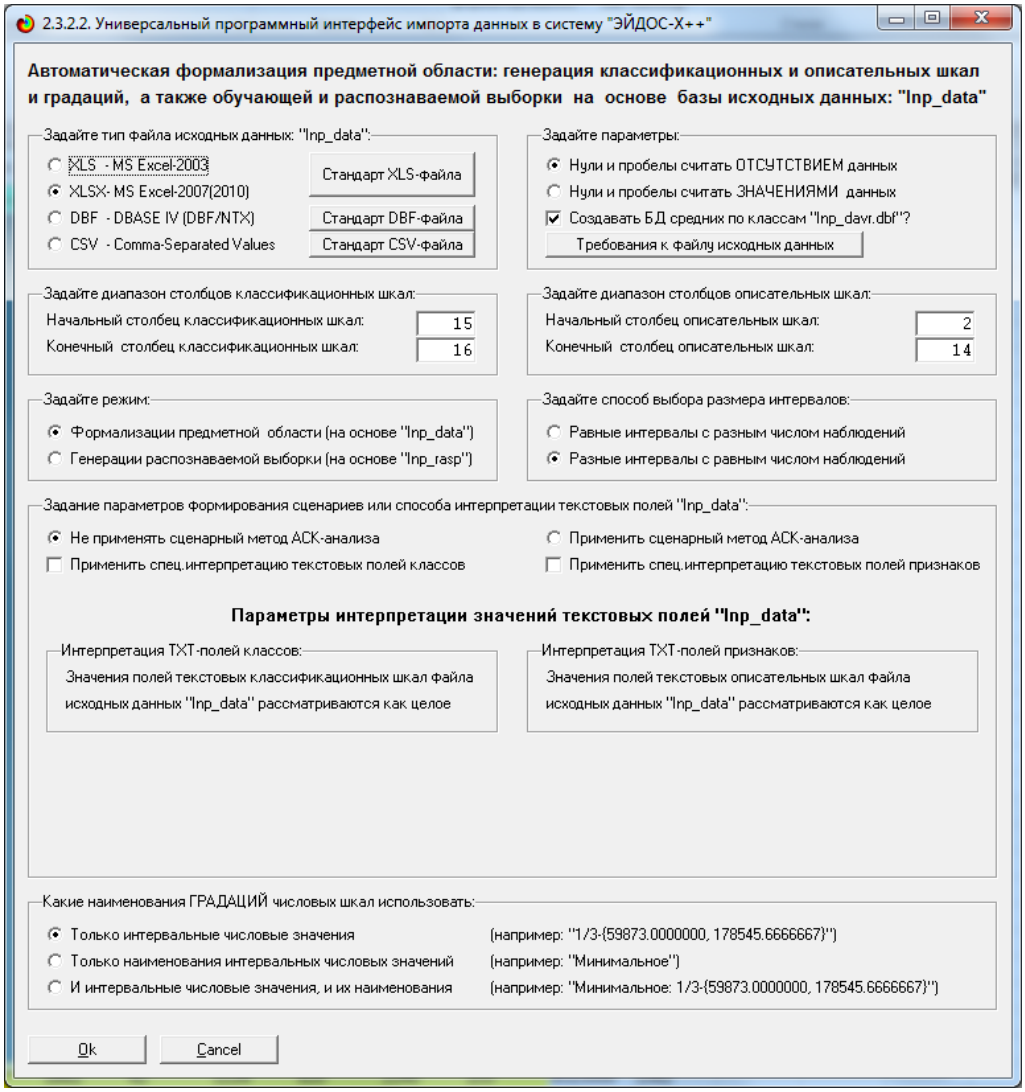


Рисунок 4. Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

В экранной форме, приведенной на рисунке 5, задать настройки, показанные на рисунке:

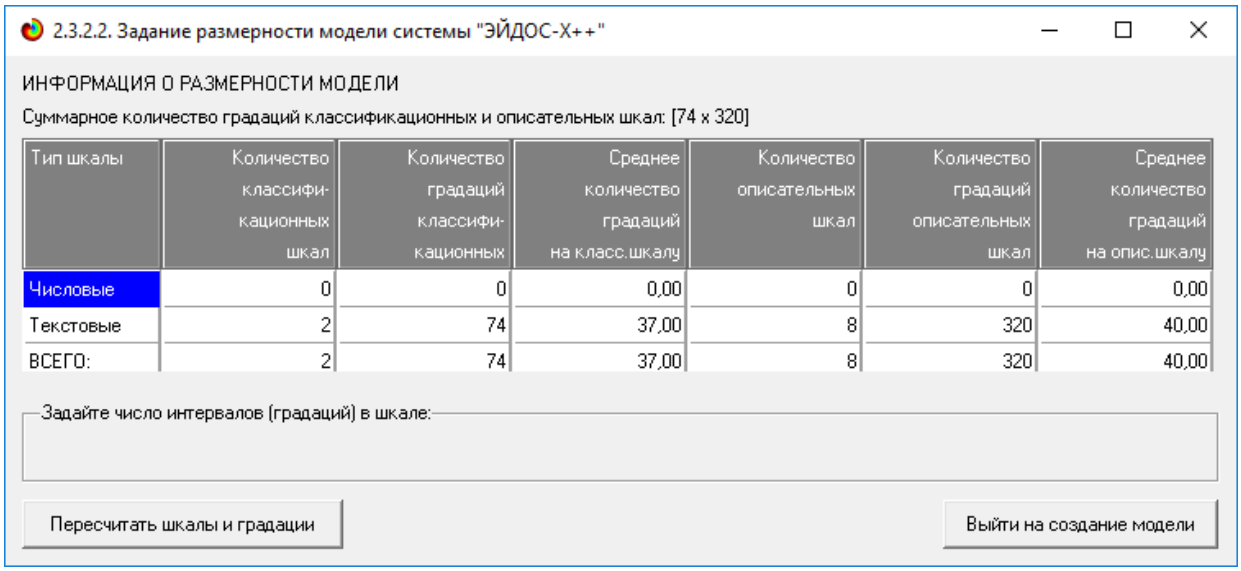
– "Задайте тип файла исходных данных Inp_data": "XLS – MS Excel-2003";

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

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 3, "Конечный столбец описательных шкал" – 24;

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

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 классифи-кационных шкал	Количество градаций классифи-кационных	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	0	0	0,00	0	0	0,00
Текстовые	2	74	37,00	8	320	40,00
ВСЕГО:	2	74	37,00	8	320	40,00

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".

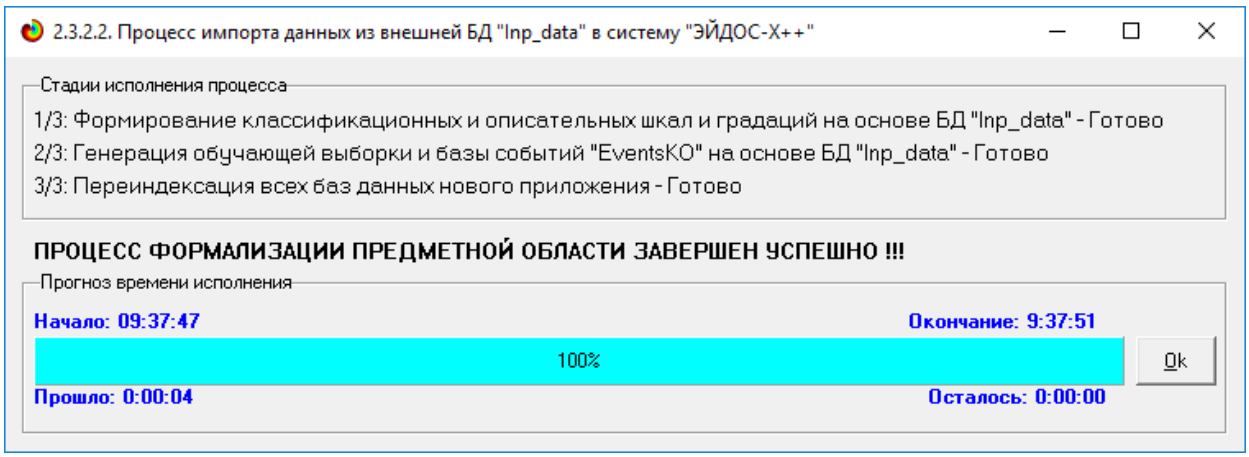


Рисунок 6. Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

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).

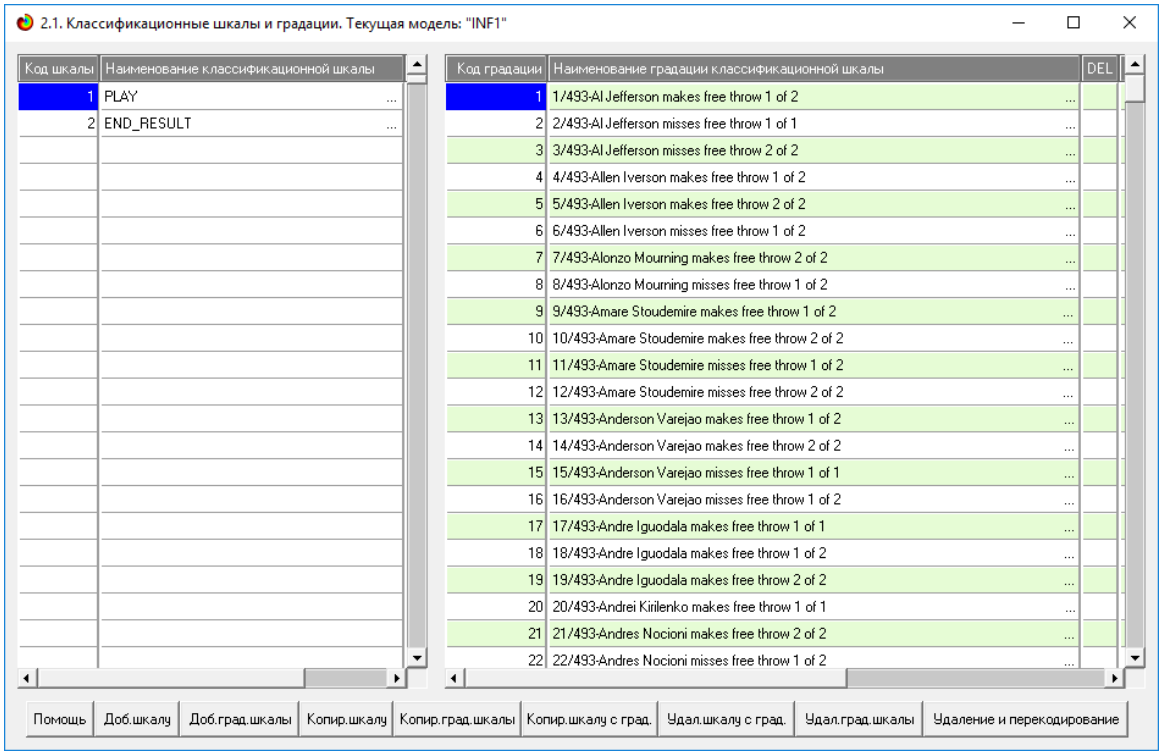


Рисунок 7. Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 8):

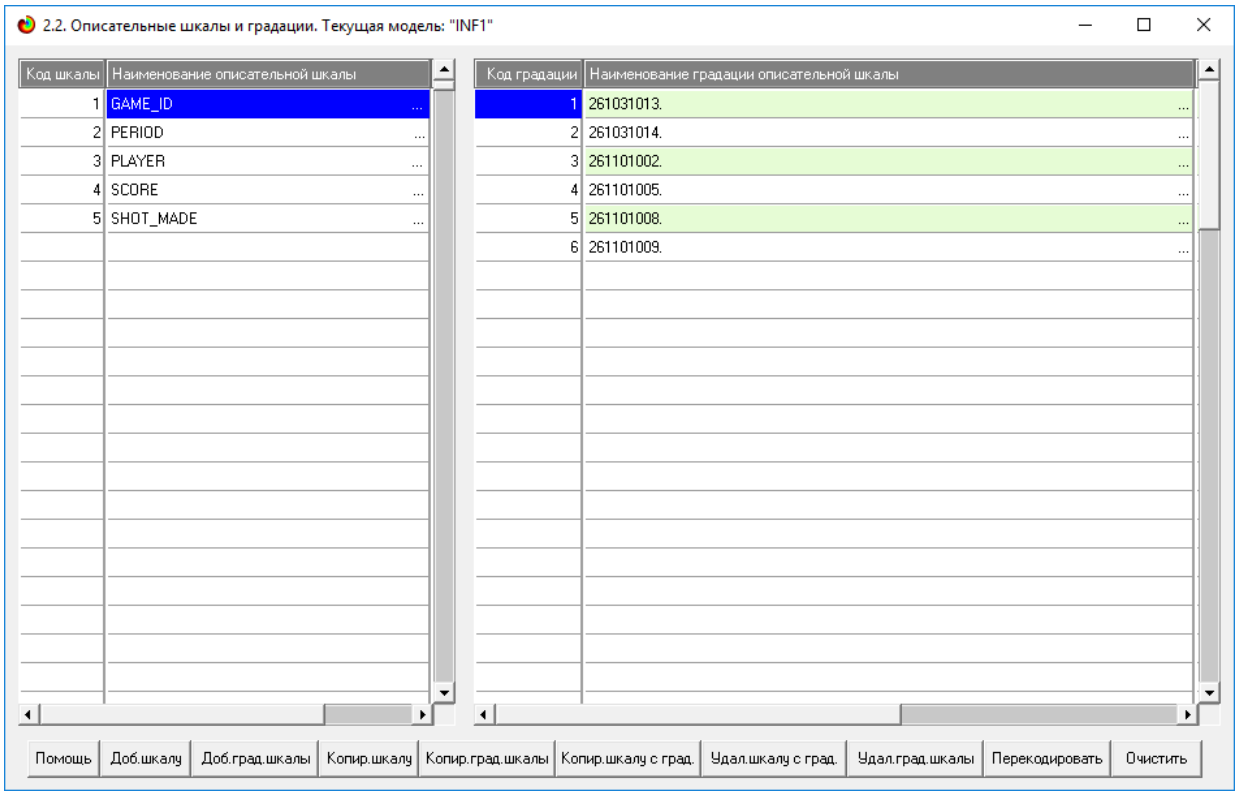


Рисунок 8. Описательные шкалы и градации (фрагмент)

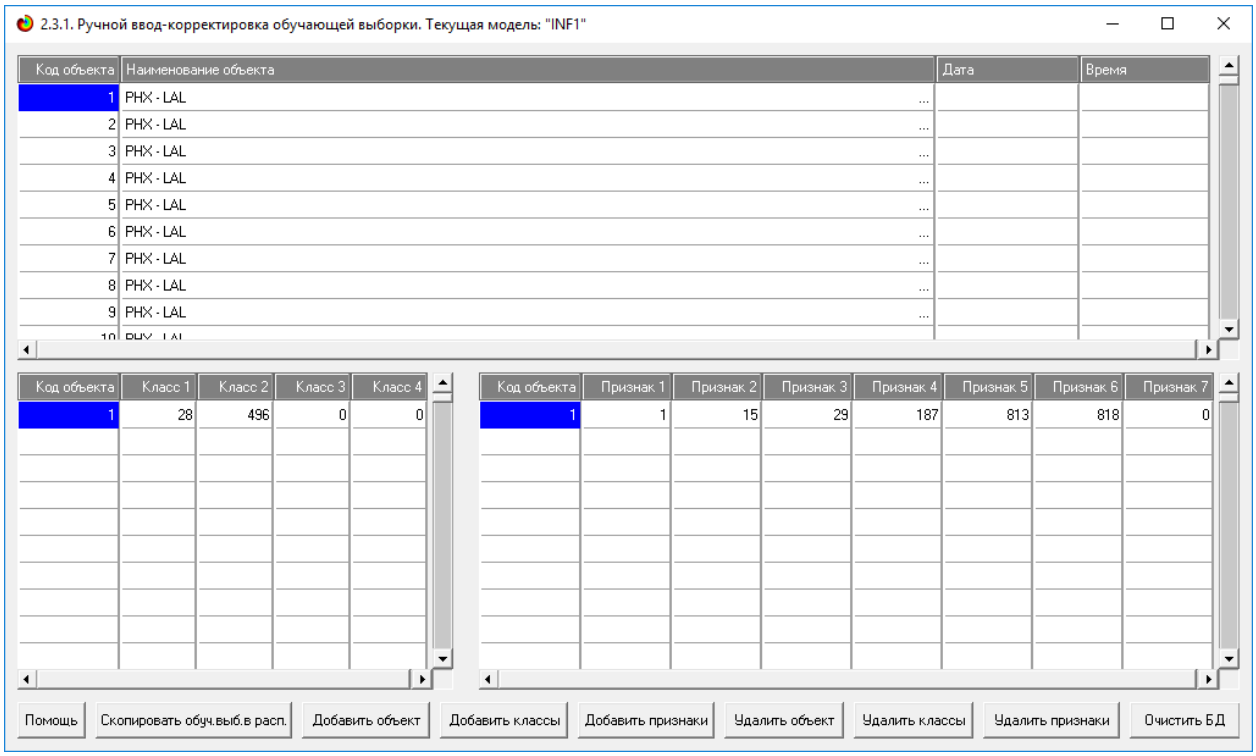


Рисунок 9. Обучающая выборка (фрагмент)

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).

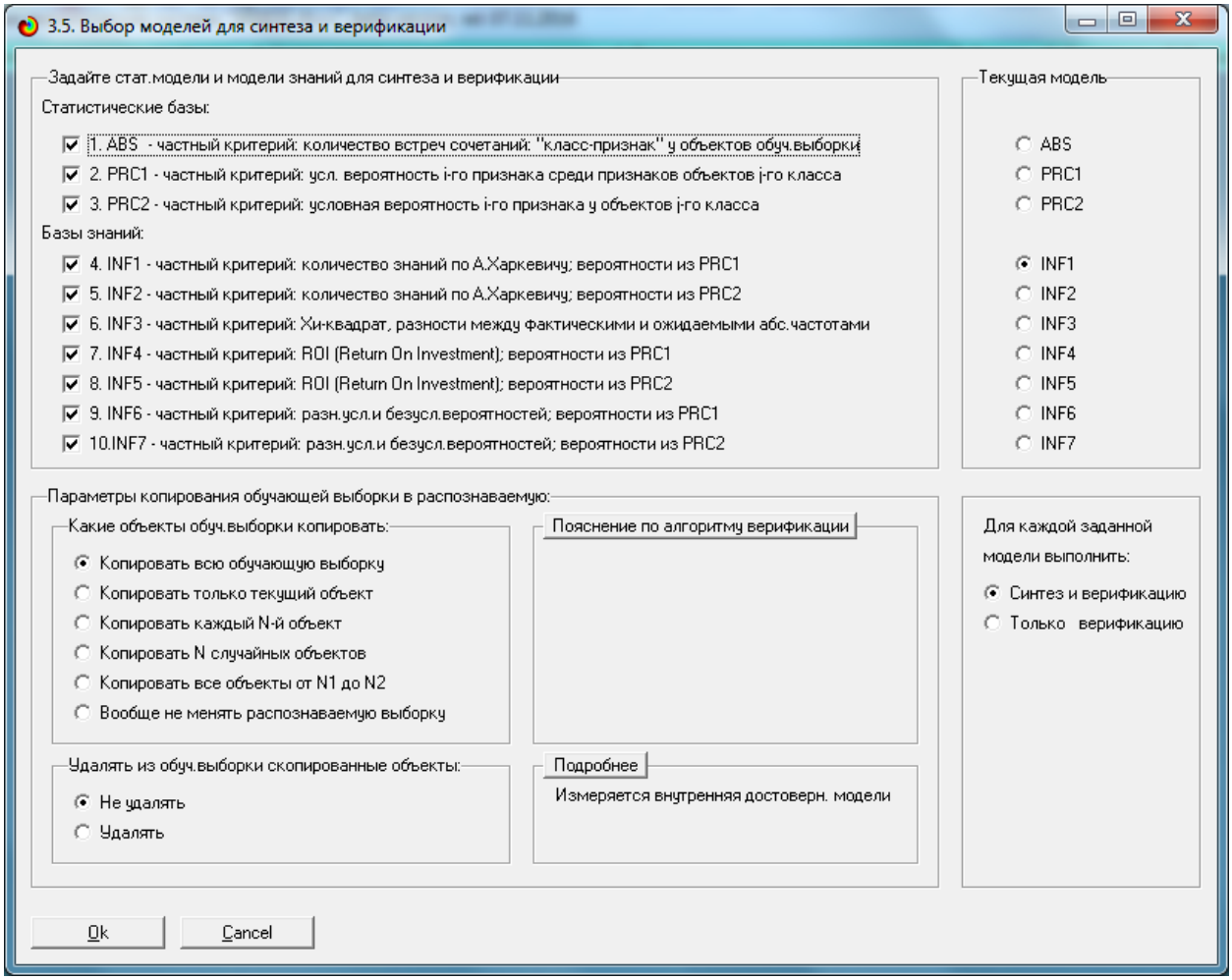


Рисунок 10. Выбор моделей для синтеза и верификации, а также текущей модели

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.

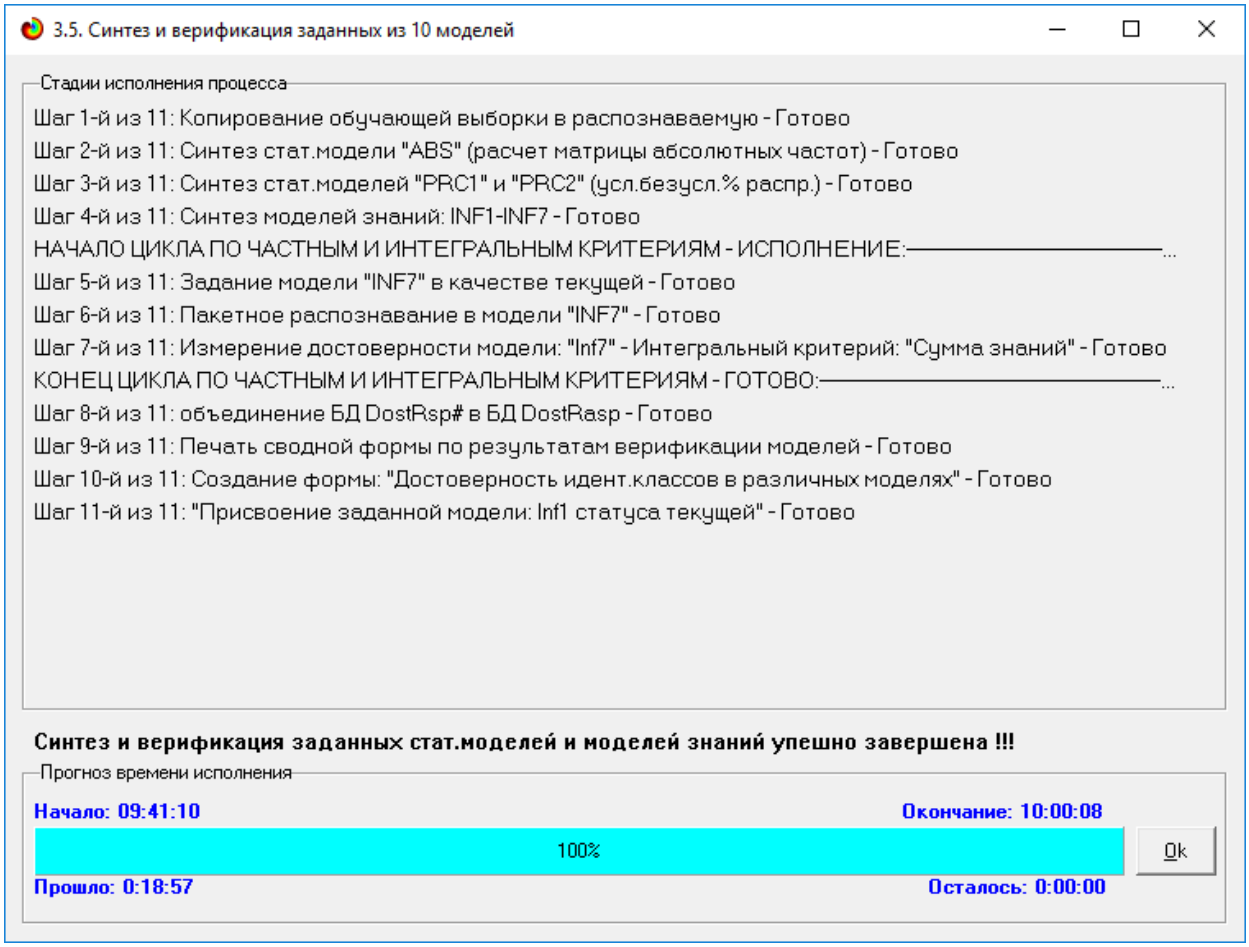


Рисунок 11. Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 32 секунды. При этом верификация (оценка достоверности моделей) проводилась на всех 260 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 INF3, 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: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"

Код признака	Наименование описательной шкалы и градации	1. PLAY AL JEFFERS	2. PLAY ALONZO MOU	3. PLAY AMARE STOU	4. PLAY ANDERSON V	5. PLAY ANDREW BOG	6. PLAY ANDREW BYN	7. PLAY ANTAWN JAM	8. PLAY ANTOINE WA	9. PLAY ANTONIO DA	10. PLAY BARON DAVI	11. PLAY BEN GORDON	12. PLAY BEN WALLAC	13. PLAY BOBBY JACK	14 BR HA
1	GAME_ID-261031013.			100.000			71.429								
2	GAME_ID-261031014.		100.000						100.000			100.000	100.000		
3	GAME_ID-261101002.	100.000												100.000	
4	GAME_ID-261101005.				100.000			100.000		100.000					
5	GAME_ID-261101008.					100.000									
6	GAME_ID-261101009.						28.571				100.000				
7	PERIOD-1.0	33.333		50.000		14.286	85.714	50.000	33.333		100.000	80.000	100.000	50.000	
8	PERIOD-2.0		100.000	50.000	55.556	28.571	14.286		66.667			20.000		50.000	
9	PERIOD-3.0									33.333					
10	PERIOD-4.0	66.667			44.444	57.143		50.000		66.667					
11	PLAYER-Al Jeffers	100.000													
12	PLAYER-Alonzo Mou		100.000												
13	PLAYER-Amare Stou			100.000											
14	PLAYER-Anderson V				100.000										
15	PLAYER-Andrew Bog					100.000									
16	PLAYER-Andrew Byn						100.000								
17	PLAYER-Antawn Jam							100.000							
18	PLAYER-Antoine Wa								100.000						
19	PLAYER-Antonio Da									100.000					
20	PLAYER-Baron Davi										100.000				
21	PLAYER-Ben Gordon											100.000			
22	PLAYER-Ben Wallac												100.000		
23	PLAYER-Bobby Jack													100.000	
24	PLAYER-Brendan Ha														

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

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

Код признака	Наименование описательной шкалы и градации	1. PLAY AL JEFFERS	2. PLAY ALONZO MOU	3. PLAY AMARE STOU	4. PLAY ANDERSON V	5. PLAY ANDREW BOG	6. PLAY ANDREW BYN	7. PLAY ANTAWN JAM	8. PLAY ANTOINE WA	9. PLAY ANTONIO DA	10. PLAY BARON DAVI	11. PLAY BEN GORDON	12. PLAY BEN WALLAC	13. PLAY BOBBY JACK
1	GAME_ID-261031013.			4			5							
2	GAME_ID-261031014.		2						6			5	1	
3	GAME_ID-261101002.	3												2
4	GAME_ID-261101005.				9			4		6				
5	GAME_ID-261101008.					7								
6	GAME_ID-261101009.						2				2			
7	PERIOD-1.0	1		2		1	6	2	2		2	4	1	1
8	PERIOD-2.0		2	2	5	2	1		4			1		1
9	PERIOD-3.0									2				
10	PERIOD-4.0	2			4	4		2		4				
11	PLAYER-Al Jeffers	3												
12	PLAYER-Alonzo Mou		2											
13	PLAYER-Amare Stou			4										
14	PLAYER-Anderson V				9									
15	PLAYER-Andrew Bog					7								
16	PLAYER-Andrew Byn						7							
17	PLAYER-Antawn Jam							4						
18	PLAYER-Antoine Wa								6					
19	PLAYER-Antonio Da									6				
20	PLAYER-Baron Davi										2			
21	PLAYER-Ben Gordon											5		
22	PLAYER-Ben Wallac												1	
23	PLAYER-Bobby Jack													2
24	PLAYER-Brendan Ha													

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

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

Код признака	Наименование описательной шкалы и градации	1. PLAY AL JEFFERS	2. PLAY ALONZO MOU	3. PLAY AMARE STOU	4. PLAY ANDERSON V	5. PLAY ANDREW BOG	6. PLAY ANDREW BYN	7. PLAY ANTAWN JAM	8. PLAY ANTOINE WA	9. PLAY ANTONIO DA	10. PLAY BARON DAVI	11. PLAY BEN GORDON	12. PLAY BEN WALLAC	13. PLAY BOBBY JACK	14 BR HA
1	GAME_ID-261031013.			1.584			1.315								
2	GAME_ID-261031014.		1.364						1.364			1.364	1.364		
3	GAME_ID-261101002.	1.134												1.134	
4	GAME_ID-261101005.				1.293			1.293		1.293					
5	GAME_ID-261101008.					1.293									
6	GAME_ID-261101009.						1.501				2.501				
7	PERIOD-1.0	0.246		0.570		−0.431	1.000	0.570	0.246		1.123	0.945	1.123	0.570	
8	PERIOD-2.0		1.021	0.467	0.551	0.020	−0.533		0.697			−0.264		0.467	
9	PERIOD-3.0									0.364					
10	PERIOD-4.0	0.736			0.413	0.613		0.507		0.736					
11	PLAYER-Al Jeffers	3.672													
12	PLAYER-Alonzo Mou		3.996												
13	PLAYER-Amare Stou			3.442											
14	PLAYER-Anderson V				2.795										
15	PLAYER-Andrew Bog					2.995									
16	PLAYER-Andrew Byn						2.995								
17	PLAYER-Antawn Jam							3.442							
18	PLAYER-Antoine Wa								3.118						
19	PLAYER-Antonio Da									3.118					
20	PLAYER-Baron Davi										3.996				
21	PLAYER-Ben Gordon											3.264			
22	PLAYER-Ben Wallac												4.549		
23	PLAYER-Bobby Jack													3.996	
24	PLAYER-Brendan Ha														

1.5. Результаты верификации моделей

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

4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит.. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма моду... уровней схо... истинно-поло... решений (ST...	Сумма моду... уровней схо... истинно-отри... решений (ST...	Сумма моду... уровней схо... ложно-поло... решений (SFP)	Сумма моду... уровней схо... ложно-отриц... решений (SF...	S-Точность модели	S-Полнота модели	L1-мера проф Е.В.Луценко	Средний модуль уровней сходс... истинно-полож... решений	Средний модуль уровней сходс... истинно-отрицат. решений
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс.частот с обр....	0.048	1.000	0.092	429.975	138.006	2391.982		0.152	1.000	0.264	0.721	
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс.частот по признак...	0.048	1.000	0.092	250.407		457.085		0.354	1.000	0.523	0.420	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	0.048	1.000	0.092	429.975	138.006	2391.982		0.152	1.000	0.264	0.721	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	0.048	1.000	0.092	429.969	138.006	2391.968		0.152	1.000	0.264	0.721	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	0.074	1.000	0.138	403.194	1799.085	1322.555		0.234	1.000	0.379	0.677	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	0.074	1.000	0.138	254.215	455.539	201.323		0.558	1.000	0.716	0.427	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	0.058	1.000	0.109	409.732	230.222	1771.362		0.188	1.000	0.316	0.687	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	0.058	1.000	0.109	409.732	230.222	1771.362		0.188	1.000	0.316	0.687	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	

Помощь

Рисунок 12. Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели ABS, INF3 при интегральном критерии «Семантический резонанс знаний». При этом точность модели составляет 0,950. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 69%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием (рисунок 13).

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

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

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

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

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

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

Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не
идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно
относить объекты к классам, которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Этот критерий предложен и
реализован в системе "Эйдос" проф. Е.В.Луценко в 1994 году. Эта мера достоверности модели предполагает два варианта нормировки: {-1, +1} и {0, 1}:
L1 = (      TP + TN - FP - FN ) / ( TP + TN + FP + FN )        (нормировка: {-1,+1})
L2 = ( 1 + ( TP + TN - FP - FN ) / ( TP + TN + FP + FN ) ) / 2   (нормировка: { 0, 1})
где: TP - истино-положительное решение; TN - истино-отрицательное решение; FP - ложно-положительное  решение; FN - ложно-отрицательное  решение;

Классическая F-мера достоверности моделей Ван Ризбергена:

Precision = TP/(TP+FP) - точность модели;
Recall    = TP/(TP+FN) - полнота  модели;
F-mera    = 2*(Precision*Recall)/(Precision+Recall).

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

SPrecision = STP/(STP+SFP) - нечеткая мультиклассовая точность модели;
SRecall    = STP/(STP+SFN) - нечеткая мультиклассовая полнота  модели;
L-mera     = 2*(SPrecision*SRecall)/(SPrecision+SRecall) - нечеткая мультиклассовая достоверность модели.

где:
STP - Сумма модулей сходства истино-положительных решений; STN - Сумма модулей сходства истино-отрицательных решений;
SFP - Сумма модулей сходства ложно-положительных  решений; SFN - Сумма модулей сходства ложно-отрицательных  решений.
```

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 не идентифицированных ситуаций в наиболее достоверной модели INF5.

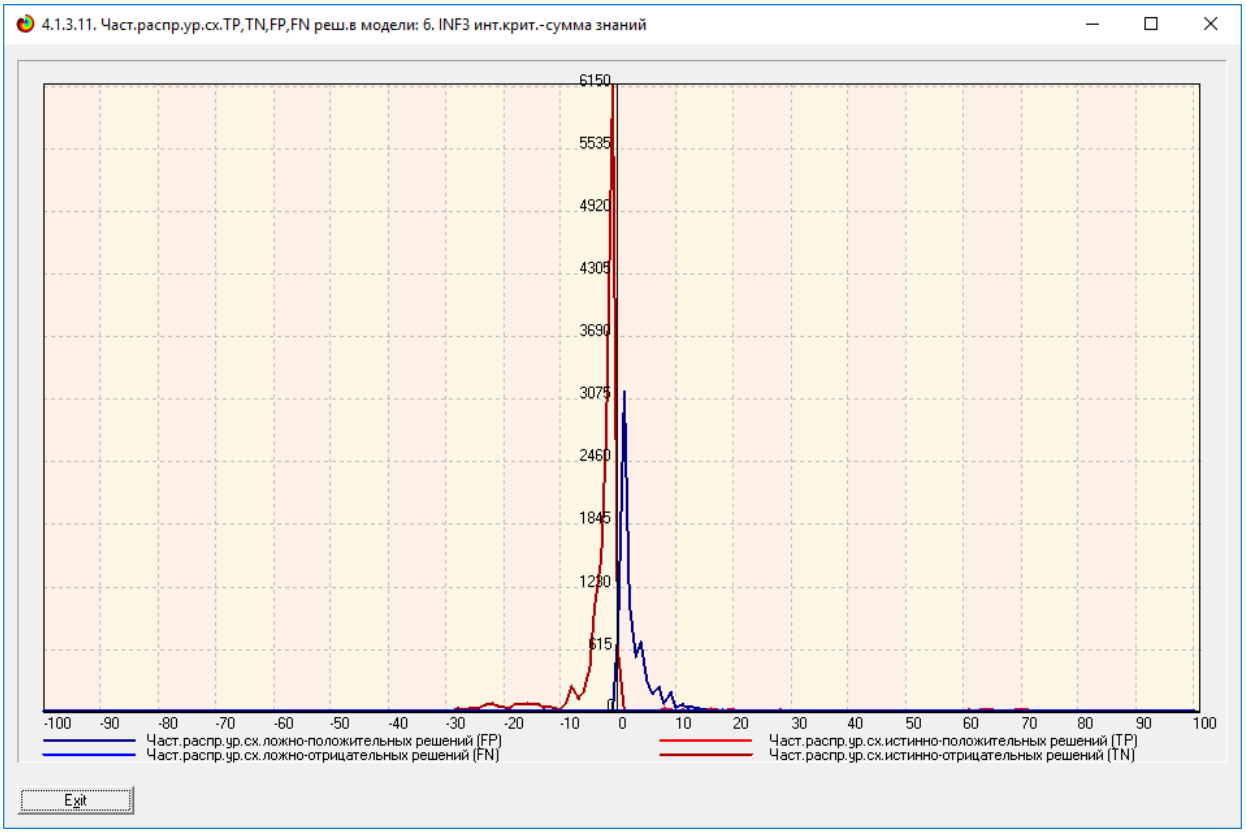


Рисунок 14. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF3

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 уровня сходства выше 70% все положительные решения являются истинными. Для очень больших значений уровней сходства-различия (более 70%) 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.

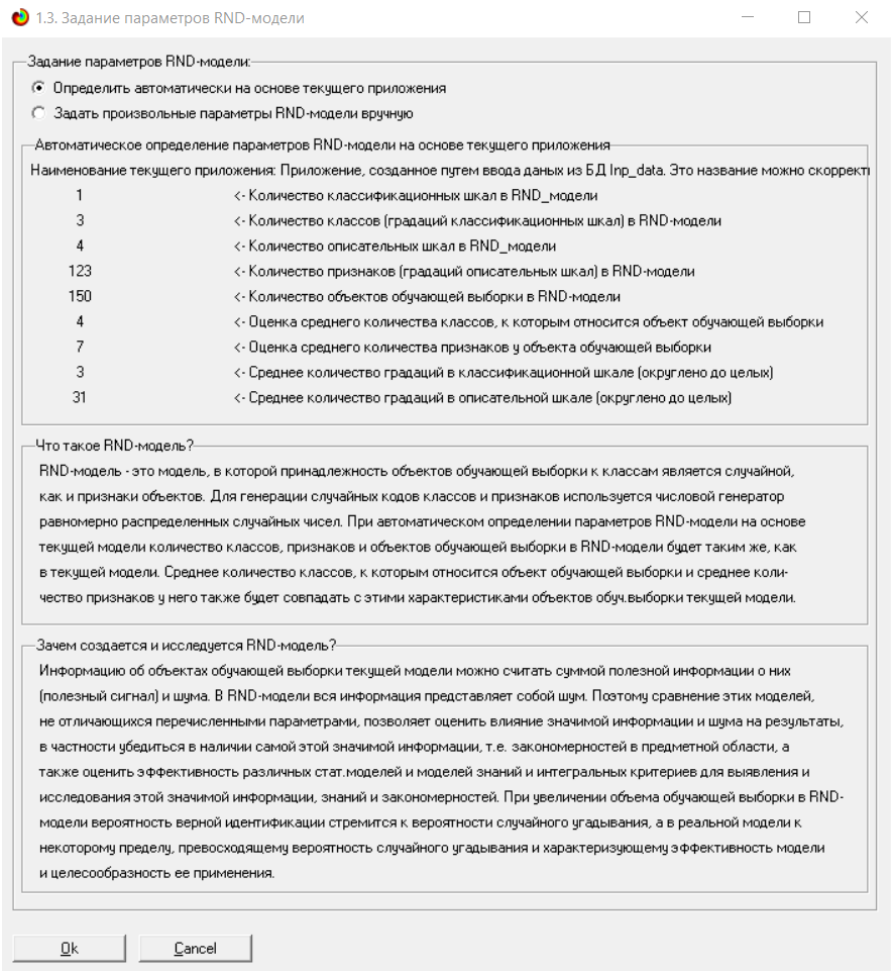


Рисунок 15. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

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

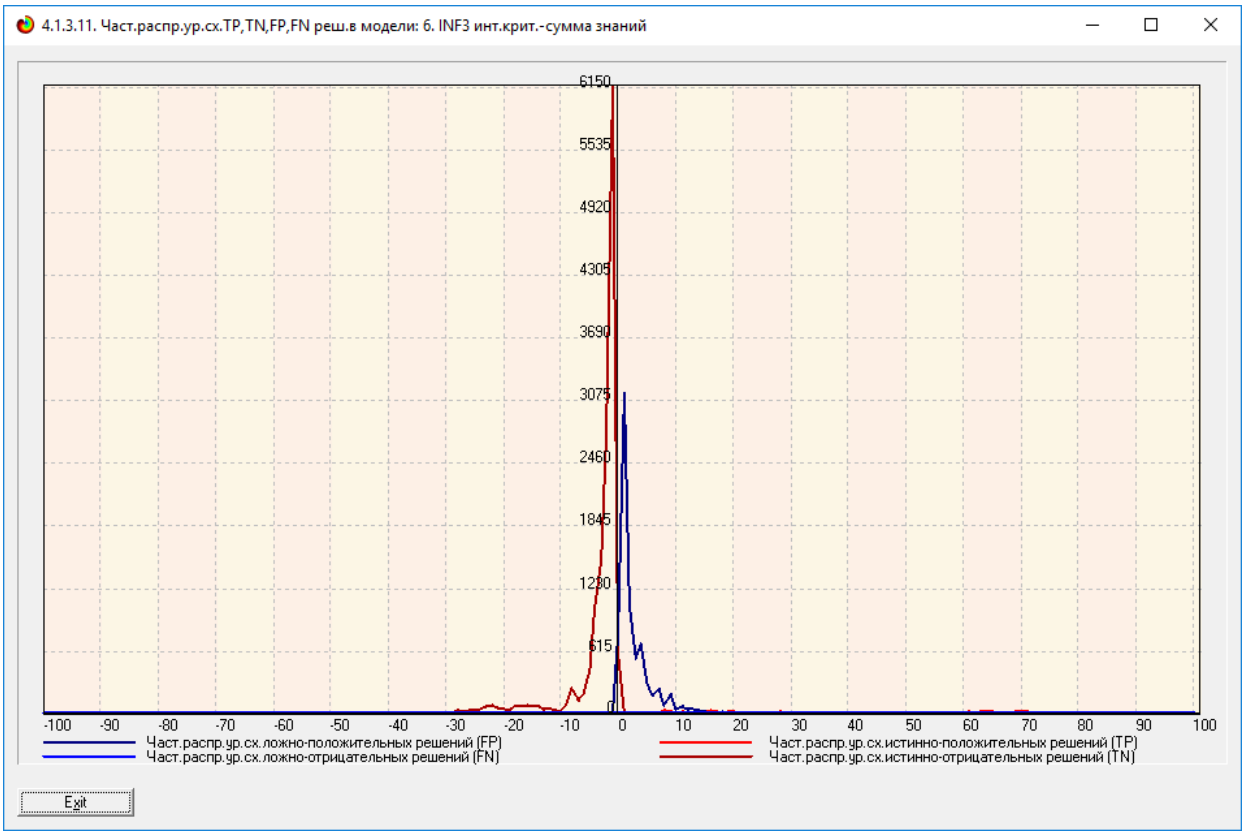


Рисунок 16. Частное распределение сходства-различия верно и ошибочно идентифицированных и не идентифицированных состояний в случайной модели INF3

Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и не идентифицированных 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. Если же такой информации в модели нет, то и распределение получается типа, приведенного на рисунке 13.

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

4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит.. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма моду... уровней схо... истино-поло... решений (ST...	Сумма моду... уровней схо... истино-отри... решений (ST...	Сумма моду... уровней схо... ложно-поло... решений (SFP)	Сумма моду... уровней схо... ложно-отриц... решений (SF...	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко	Средний модуль уровней сходс... истино-полож... решений	Средний мо... уровней схо... истино-отри... решений
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс.частот с обр....	0.048	1.000	0.092	429.975	138.006	2391.982		0.152	1.000	0.264	0.721	
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс.частот по признак...	0.048	1.000	0.092	250.407		457.085		0.354	1.000	0.523	0.420	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	0.048	1.000	0.092	429.975	138.006	2391.982		0.152	1.000	0.264	0.721	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	0.048	1.000	0.092	429.969	138.006	2391.968		0.152	1.000	0.264	0.721	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	0.074	1.000	0.138	403.194	1799.085	1322.555		0.234	1.000	0.379	0.677	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	0.074	1.000	0.138	254.215	455.539	201.323		0.558	1.000	0.716	0.427	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	0.058	1.000	0.109	409.732	230.222	1771.362		0.188	1.000	0.316	0.687	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	0.058	1.000	0.109	409.732	230.222	1771.362		0.188	1.000	0.316	0.687	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	

Помощь

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 В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

В соответствии с технологией АСК-анализа зададим текущей модель INF3 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.1.2.

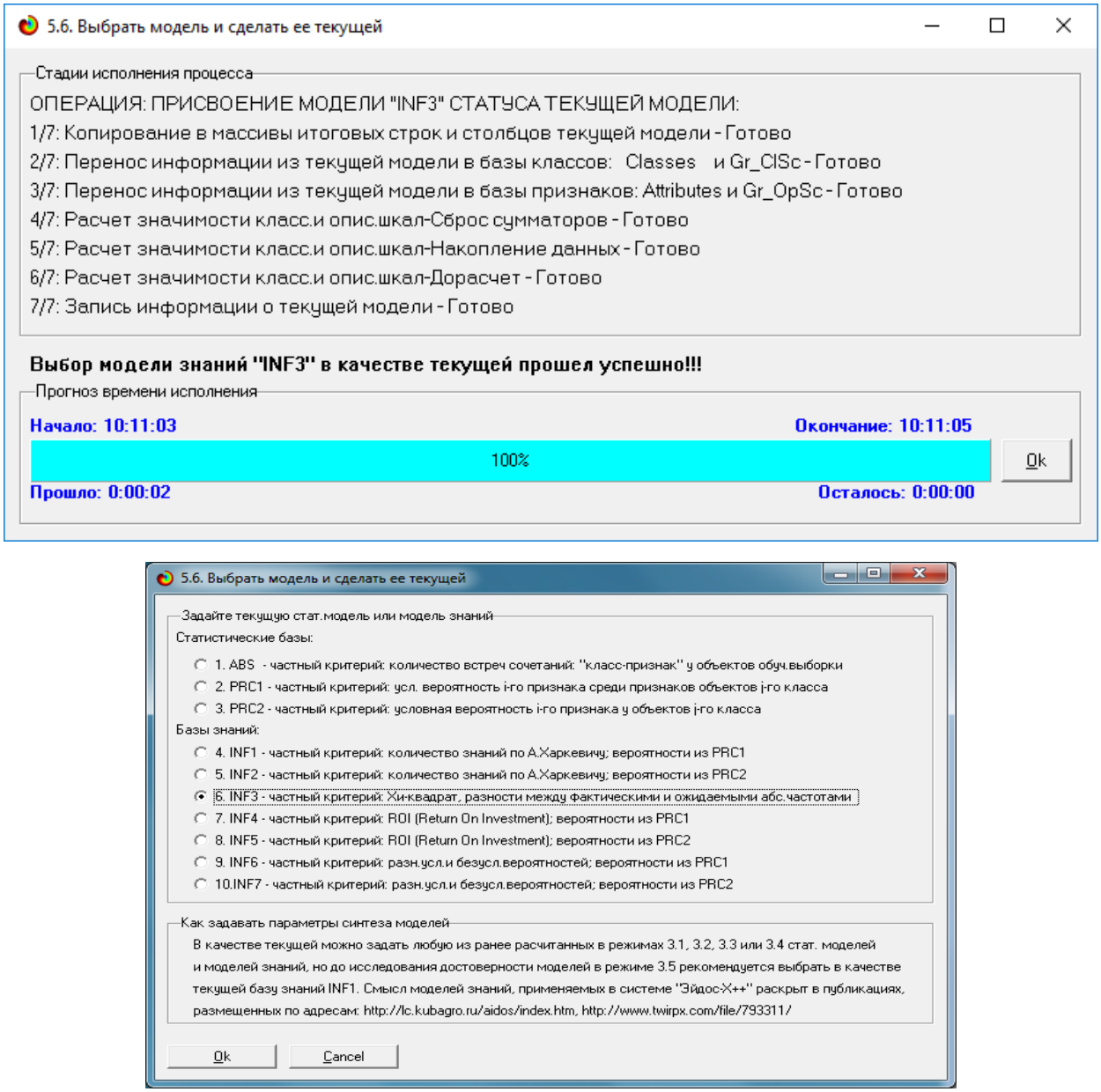


Рисунок 18. Экранные формы режима задания модели в качестве текущей

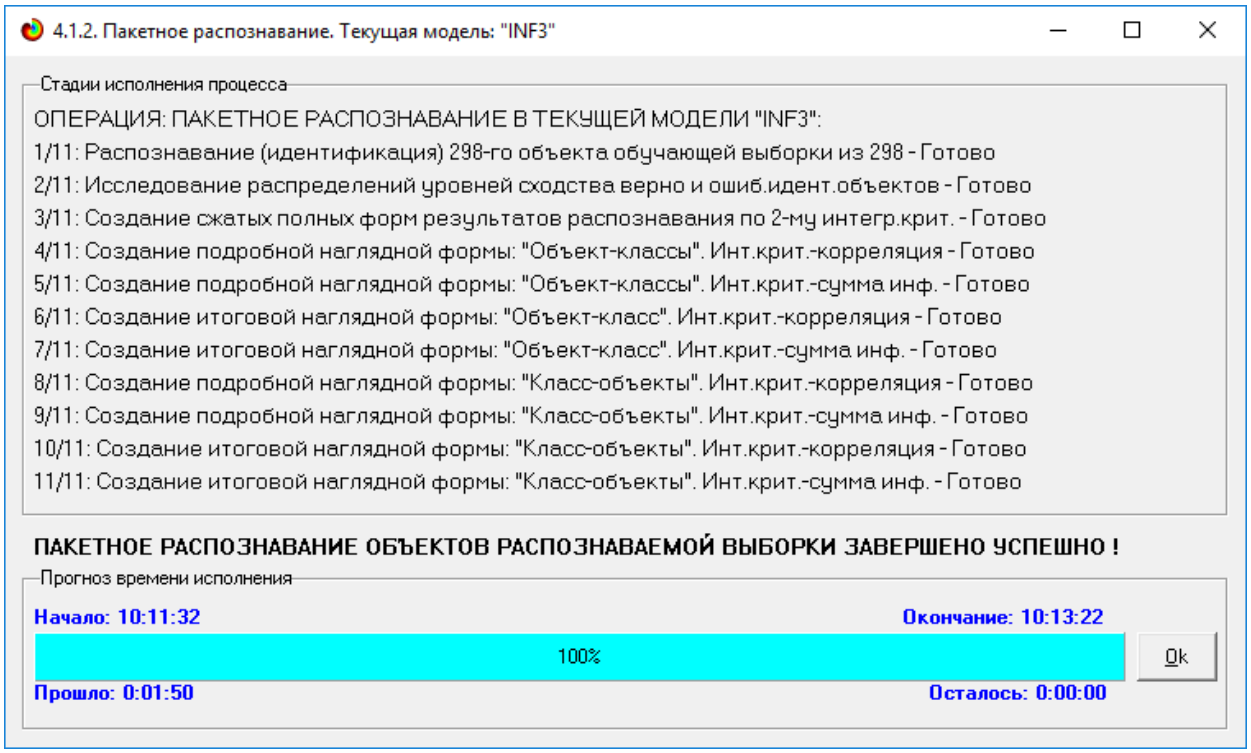


Рисунок 19. Экранная форма режима пакетного распознавания в текущей модели

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:

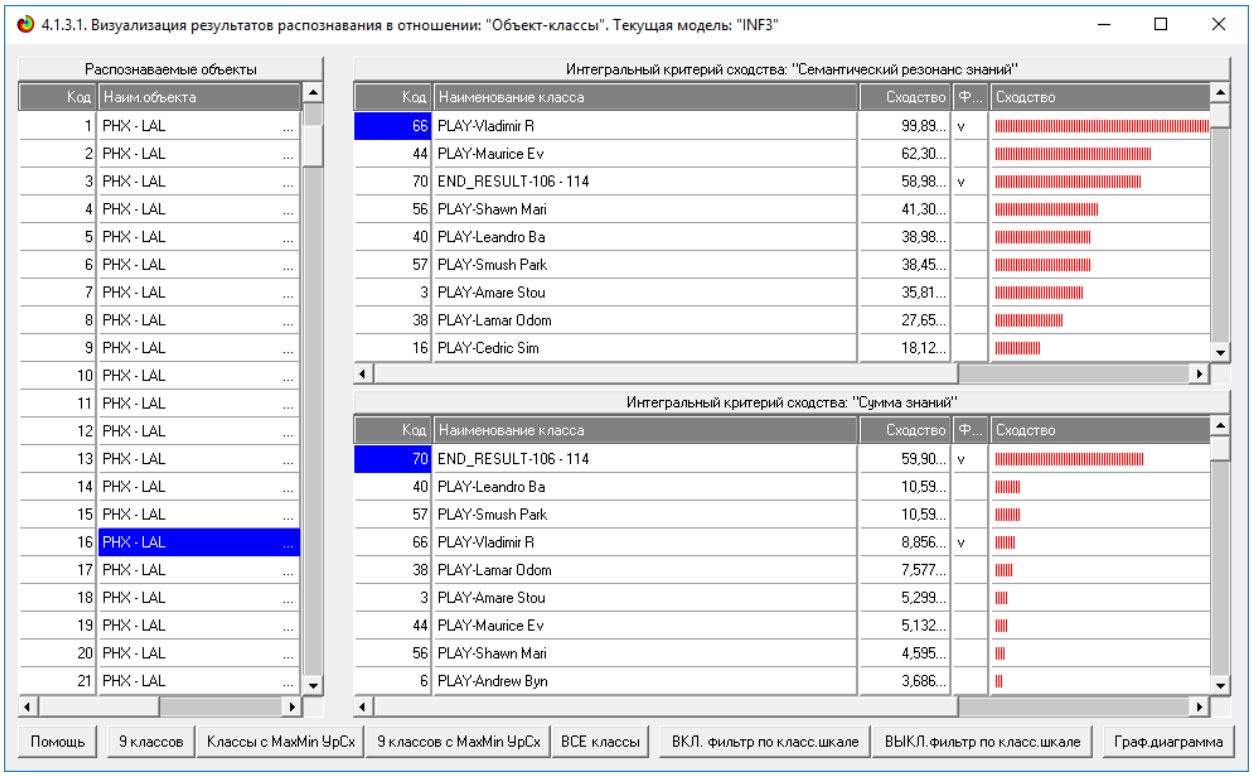


Рисунок 20. Пример идентификации классов в модели INF3

4.1.3.1. Визуализация результатов распознавания в отношении: "Объект-классы". Текущая модель: "INF3"

Распознаваемые объекты

Код	Наим.объекта
1	PHX - LAL
2	PHX - LAL
3	PHX - LAL
4	PHX - LAL
5	PHX - LAL
6	PHX - LAL
7	PHX - LAL
8	PHX - LAL
9	PHX - LAL
10	PHX - LAL
11	PHX - LAL
12	PHX - LAL
13	PHX - LAL
14	PHX - LAL
15	PHX - LAL
16	PHX - LAL
17	PHX - LAL
18	PHX - LAL
19	PHX - LAL
20	PHX - LAL
21	PHX - LAL

Интегральный критерий сходства: "Семантический резонанс знаний"

Код	Наименование класса	Сходство	Ф...	Сходство
6	PLAY-Andrew Byn	82,18...	v	
70	END_RESULT-106 - 114	45,49...	v	
3	PLAY-Amare Stou	36,60...		
56	PLAY-Shawn Mari	30,93...		
72	END_RESULT-110 - 98	26,78...		
38	PLAY-Lamar Odom	24,07...		
40	PLAY-Leandro Ba	21,57...		
57	PLAY-Smush Park	21,28...		
43	PLAY-Luol Deng	18,79...		

Интегральный критерий сходства: "Сумма знаний"

Код	Наименование класса	Сходство	Ф...	Сходство
70	END_RESULT-106 - 114	46,20...	v	
6	PLAY-Andrew Byn	20,02...	v	
72	END_RESULT-110 - 98	12,01...		
38	PLAY-Lamar Odom	6,598...		
40	PLAY-Leandro Ba	5,865...		
57	PLAY-Smush Park	5,865...		
3	PLAY-Amare Stou	5,415...		
43	PLAY-Luol Deng	4,399...		
56	PLAY-Shawn Mari	3,441...		

Помощь | 9 классов | Классы с MaxMin УрСх | 9 классов с MaxMin УрСх | ВСЕ классы | ВКЛ. фильтр по класс.шкале | ВЫКЛ.фильтр по класс.шкале | Граф.диаграмма

Рисунок 21. Пример идентификации классов в модели INF3

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 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 функции) 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-Х++» «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 INF5.

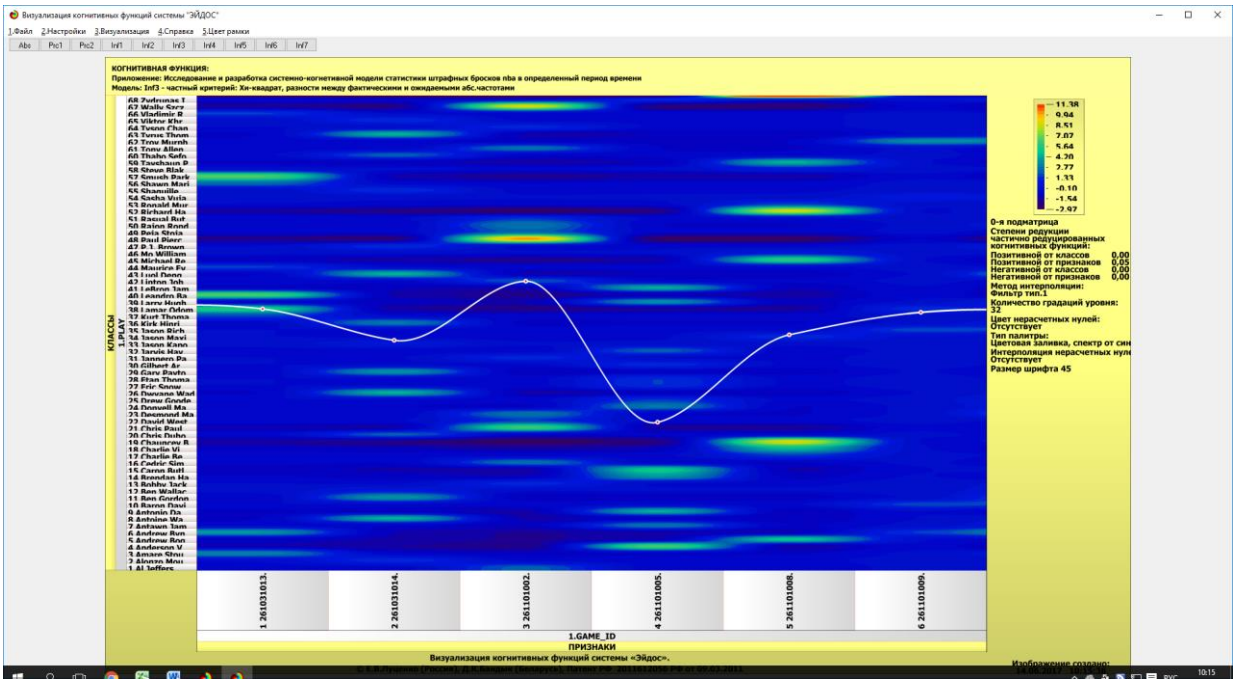


Рисунок 23. Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF3

2.3. SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-

анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают не формализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей уже около 30 лет, но она малоизвестна – это интеллектуальная система «Эйдос». Данная система всегда обеспечивала возможность проведения количественного автоматизированного SWOT-анализа без использования экспертных оценок непосредственно на основе эмпирических данных. Результаты SWOT-анализа выводились в форме информационных портретов. В версии системы под MS Windows: «Эйдос-Х++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 24).

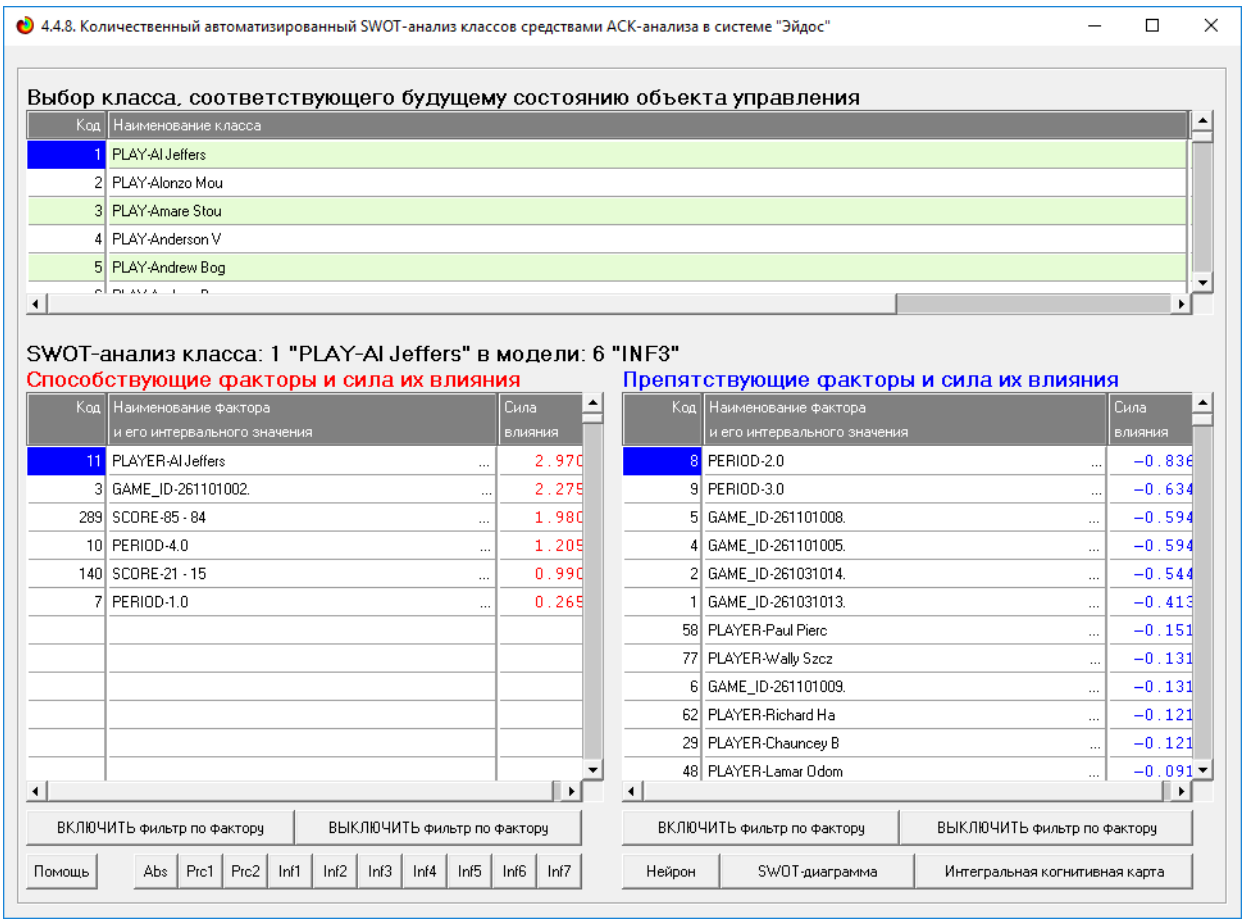


Рисунок 24. Пример SWOT-матрицы в модели INF3

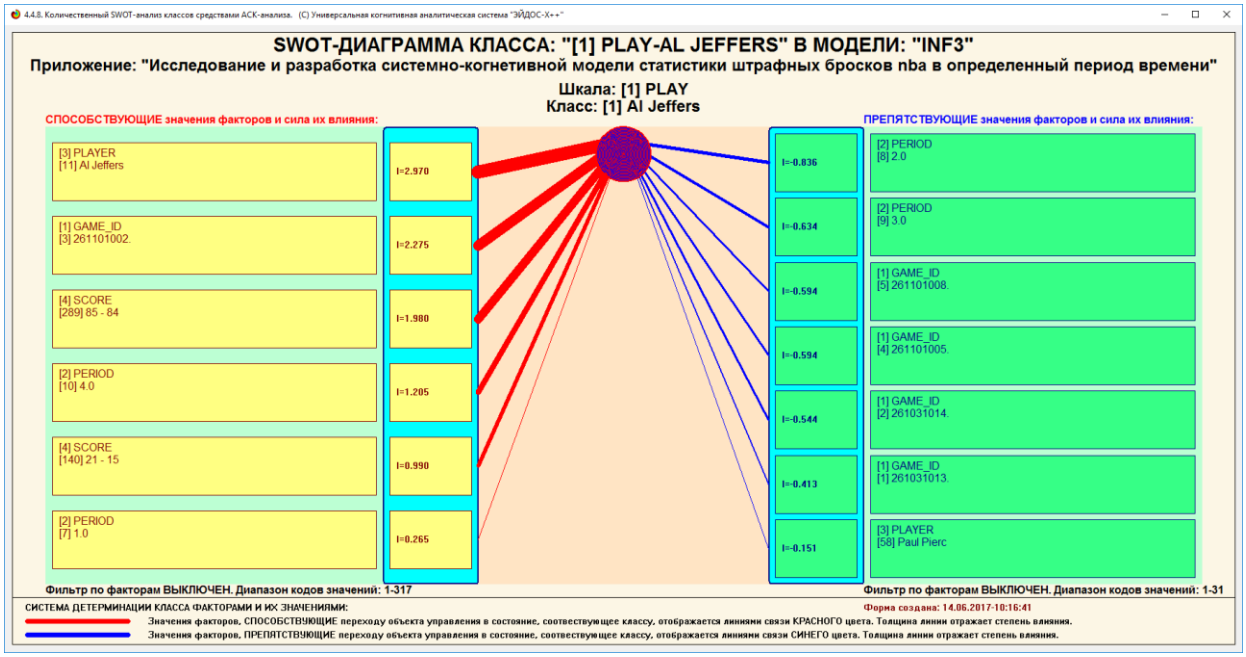


Рисунок 25 - Пример SWOT-матрицы в модели INF3

2.4. НЕЛОКАЛЬНЫЕ НЕЙРОННЫЕ СЕТИ И НЕЙРОНЫ.

Каждому классу системно-когнитивной модели соответствует не локальный нейрон совокупность, которых образует не локальную нейронную сеть [9]. В качестве примеров приведены рисунок 26-27.

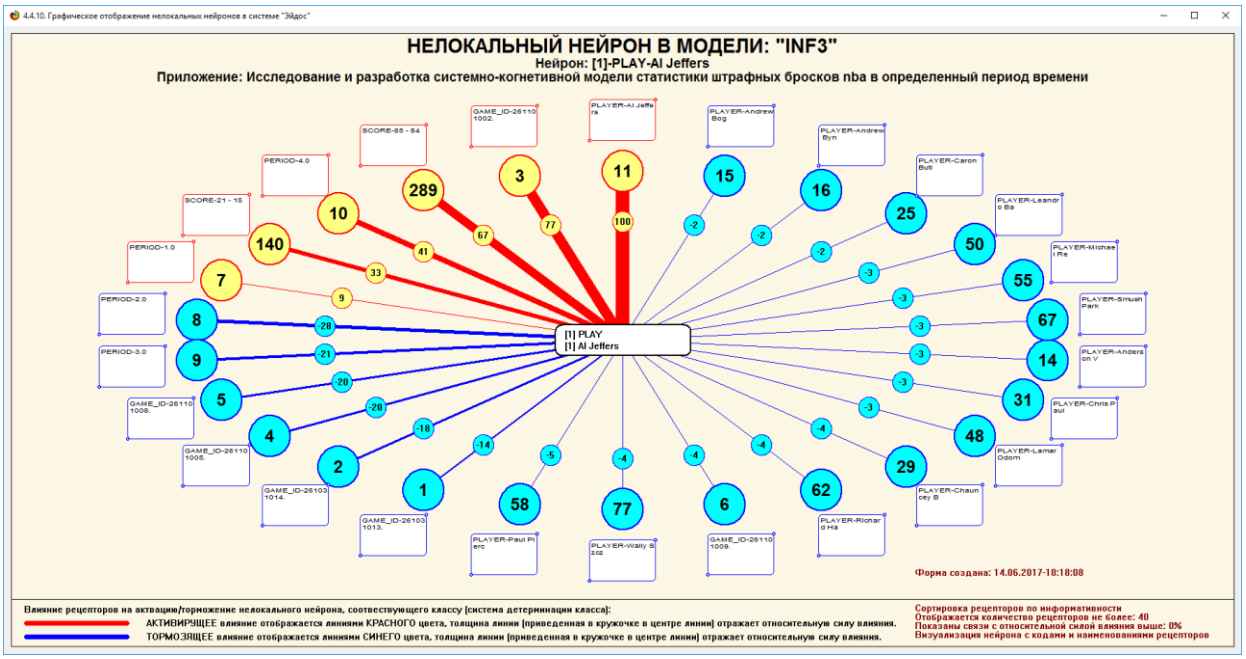


Рисунок 26. Пример нелокальные нейронные сети и нейроны модели INF3

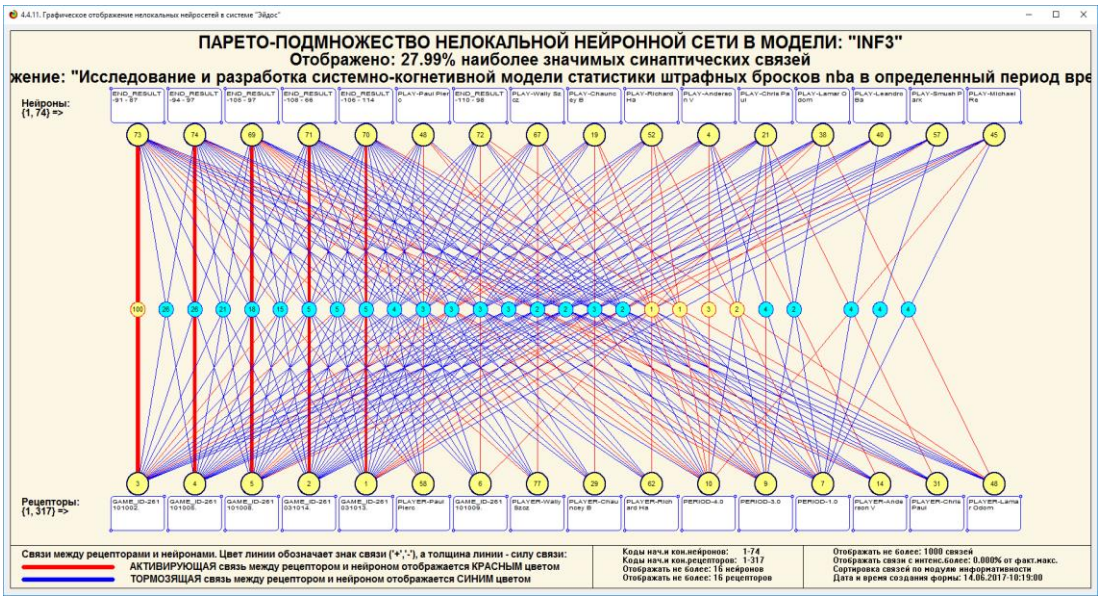


Рисунок 26. Пример нелокальные нейронные сети и нейроны модели INF3

2.5 КЛАСТЕРНЫЙ И КОНСТРУКТИВНЫЙ АНАЛИЗ

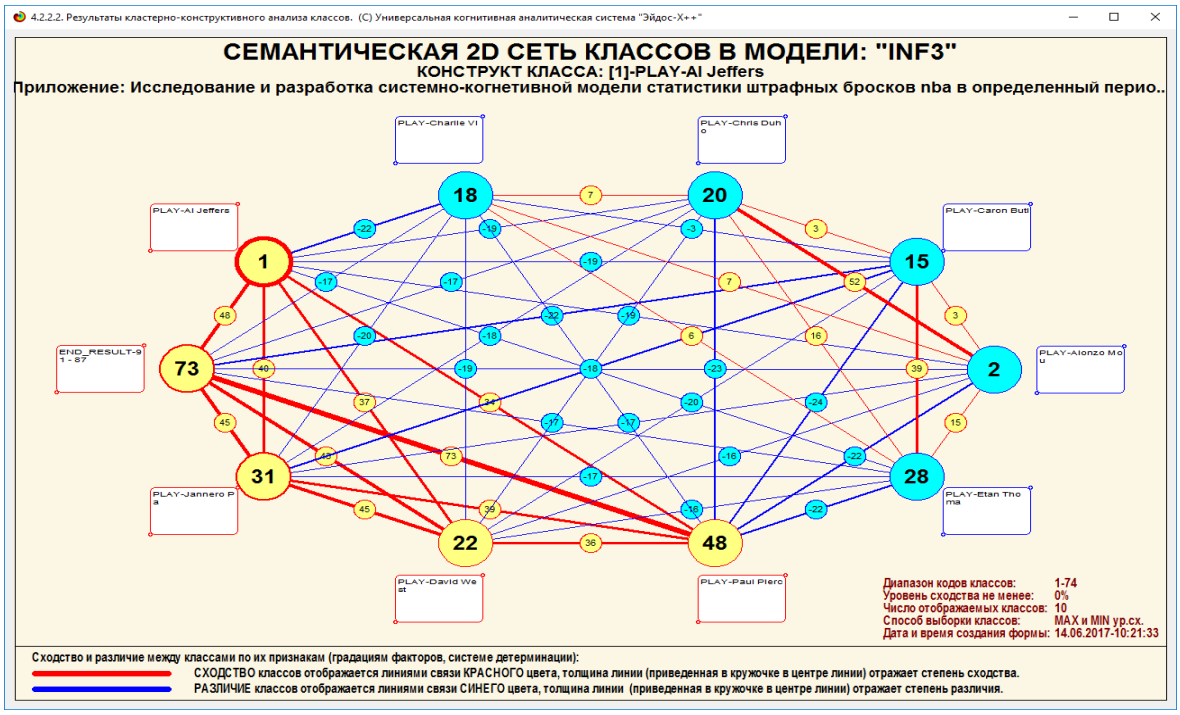


Рисунок 27. Пример кластерный и конструктивный анализ модели INF5

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 Kaggle. В данной курсовой работе приводится развернутый пример использования базы данных Kaggle 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 INF3, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,909, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 91%. 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 Kaggle, рассмотренной в данной курсовой работе, построить модели прогнозирования не с помощью АСК-анлаиза 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 автомати- зированного 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 аграрно- го 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.

5. [Электронный ресурс]. Статья "baza dannix sotrudnikov": http://allexcel.ru/gotovye-tablitsy-excel-besplatno, свободный. - Загл. с экрана. Яз. анг.

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

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

8. 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.

9. Луценко Е.В., Боровко А.Ю. ПРОГНОЗИРОВАНИЕ КОЛИЧЕСТВА И КЛАССОВ СОЛНЕЧНЫХ ВСПЫШЕК НА ОСНОВЕ ИХ ПРЕДЫСТОРИИ ПО ДАННЫМ РЕПОЗИТОРИЯ UCI 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