

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 Кукса Владислава Юрьевича курса 3 очной формы обучения
направления подготовки Информационные системы и технологии.
направленность (профиль) академический бакалавриат.

Наименование темы «Исследование и разработка системно-когнитивной модели
классификации землетрясений на основе года и магнитуды».

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

Оценка качества выполнения курсового проекта (работы)

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

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 и исследование системно когнитивной модели классификации
землетрясений на основе года и магнитуды**

выполнил студент группы ИТ1401

Кукса Владислав Юрьевич

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)

Кукса Владислав Юрьевич

(Ф.И.О.)

Тема проекта: **Разработка и исследование системно когнитивной модели
классификации землетрясений на основе года и магнитуды**

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

Объем работы:

а) пояснительная записка к работе _____ листа формата А4

б) графическая часть _____ лист формата А4

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

Срок выполнения проекта: с “ ” _____ по “ ” _____ 2017 г.

Срок защиты: “ ” _____ 2017 г.

Дата выдачи задания: “ ” _____ 2017 г.

Дата сдачи проекта на кафедру: “ ” _____ 2017 г.

Руководитель проекта: _____ Луценко Евгений Вениаминович.

д.э.н., к.т.н., профессор

(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ “ ” _____ 2017 г.

(подпись, дата)

Краснодар
2017 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ.....	7
1.1. Описание решения	7
1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel	6
1.3. Синтез и верификация статистических и интеллектуальных моделей	14
1.4. Виды моделей системы «Эйдос»	16
1.5. Результаты верификации моделей	17
2.РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	24
2.1. Решение задачи идентификации	24
2.2. Когнитивные функции.....	27
2.3. SWOT и PEST матрицы и диаграммы	29
2.4. Нелокальные Нейронные Сети И Нейроны.....	31
2.5 Кластерный и Конструктивный Анализ.....	33
ЗАКЛЮЧЕНИЕ	34
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	35

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 землетрясений– «database.csv», которую оставим без изменений.

<https://www.kaggle.com/usgs/earthquake-database>

Обучающая выборка:

Таблица 1 – database.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	Date	Time	Latitude	Longitude	Type	Depth	Depth Err	Depth Sel	Magnitude	Magnitude	Magnitude	Magnitude	Azimuthal	Horizontal	Horizontal	Root Mean	ID	Source	Location S	Magnitude	Status	
2	01/02/196	13:44:18	19,246	145,616	Earthquake	131,6			6	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
3	01/04/196	11:29:49	1,863	127,352	Earthquake	80			5,8	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
4	01/05/196	18:05:58	-20,579	-173,972	Earthquake	20			6,2	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
5	01/08/196	18:49:43	-59,076	-23,557	Earthquake	15			5,8	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
6	01/09/196	13:32:50	11,938	126,427	Earthquake	15			5,8	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
7	01/10/196	13:36:32	-13,405	166,629	Earthquake	35			6,7	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
8	01/12/196	13:32:25	27,357	87,867	Earthquake	20			5,9	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
9	01/15/196	23:17:42	-13,309	166,212	Earthquake	35			6	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
10	01/16/196	11:32:37	-56,452	-27,043	Earthquake	95			6	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
11	01/17/196	10:43:17	-24,563	178,487	Earthquake	565			5,8	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
12	01/17/196	20:57:41	-6,807	108,988	Earthquake	227,9			5,9	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
13	01/24/196	00:11:17	-2,608	125,952	Earthquake	20			8,2	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
14	01/29/196	09:35:30	54,636	161,703	Earthquake	55			5,5	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
15	02/01/196	05:27:06	-18,697	-177,864	Earthquake	482,9			5,6	MW								ISCGEM86	ISCGEM	ISCGEM	Automatic	
16	02/02/196	15:56:51	37,523	73,251	Earthquake	15			6	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
17	02/04/196	03:25:00	-51,84	139,741	Earthquake	10			6,1	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
18	02/04/196	05:01:22	51,251	178,715	Earthquake	30,3			8,7	MW								OFFICIAL	OFFICIAL	ISCGEM	Automatic	
19	02/04/196	06:04:59	51,639	175,055	Earthquake	30			6	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
20	02/04/196	06:37:06	52,528	172,007	Earthquake	25			5,7	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
21	02/04/196	06:39:32	51,626	175,746	Earthquake	25			5,8	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
22	02/04/196	07:11:23	51,037	177,848	Earthquake	25			5,9	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
23	02/04/196	07:14:59	51,73	173,975	Earthquake	20			5,9	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
24	02/04/196	07:23:12	51,775	173,058	Earthquake	10			5,7	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
25	02/04/196	07:43:43	52,611	172,588	Earthquake	24			5,7	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
26	02/04/196	08:06:17	51,831	174,368	Earthquake	31,8			5,7	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
27	02/04/196	08:33:41	51,948	173,969	Earthquake	20			5,6	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
28	02/04/196	08:40:44	51,443	179,605	Earthquake	30			7,3	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
29	02/04/196	12:06:08	52,773	171,974	Earthquake	30			6,5	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
30	02/04/196	12:50:59	51,772	174,696	Earthquake	20			5,6	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
31	02/04/196	14:18:29	52,975	171,091	Earthquake	25			6,4	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
32	02/04/196	15:51:25	52,99	170,874	Earthquake	25			5,8	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
33	02/04/196	18:34:12	51,536	175,045	Earthquake	25			5,8	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
34	02/04/196	19:44:04	13,245	-44,922	Earthquake	10			5,8	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	
35	02/04/196	22:30:03	51,812	174,206	Earthquake	10			5,7	MW								ISCGEM85	ISCGEM	ISCGEM	Automatic	

Всего в базе 21 столбцов, 7 классификационные и 4 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.

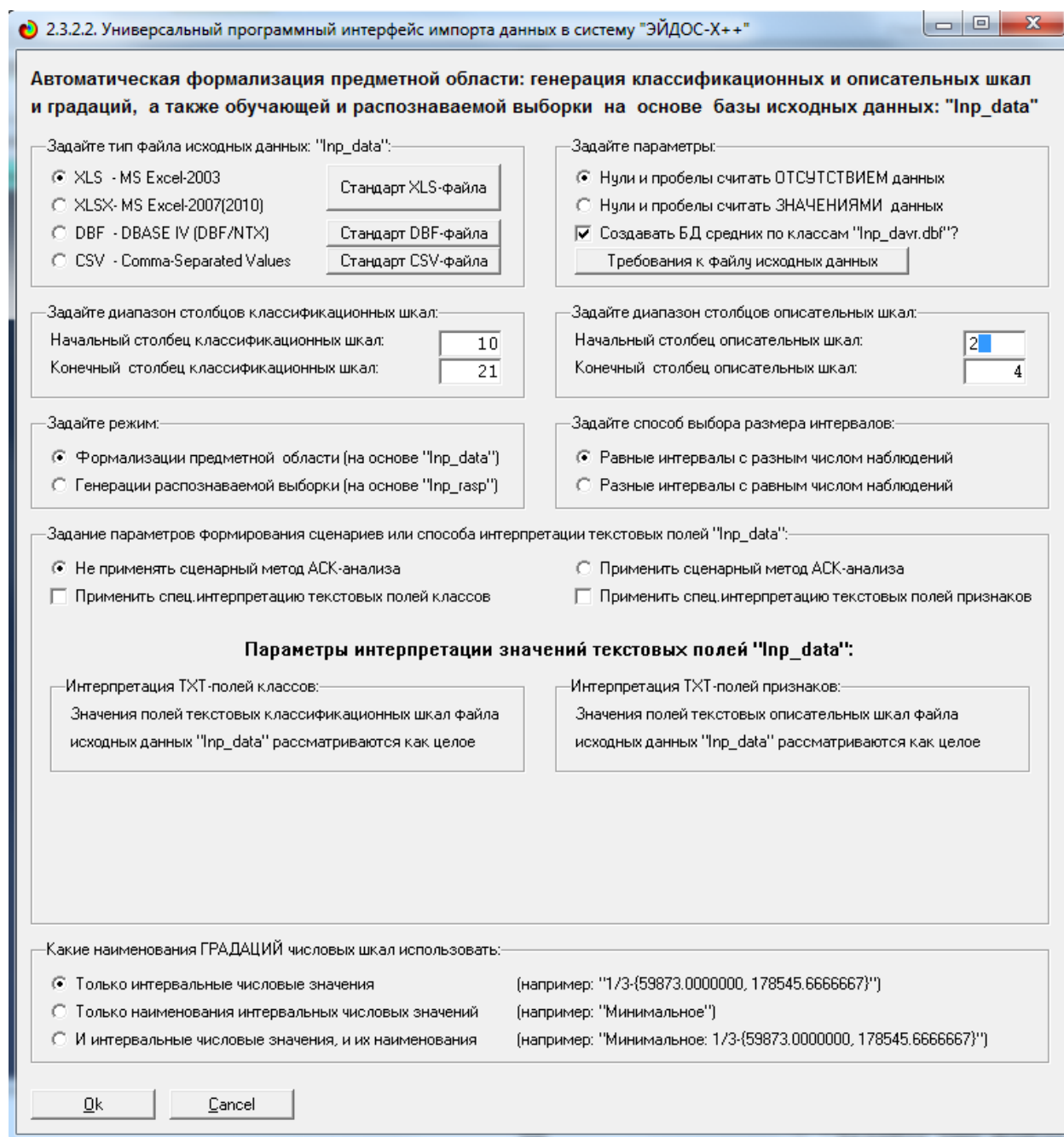


Рисунок 4. Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

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".

2.3.2.2. Задание размерности модели системы "Эйдос-Х++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [353 x 344]

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	0	0	0,00	2	6	3,00
Текстовые	12	353	29,42	1	338	338,00
ВСЕГО:	12	353	29,42	3	344	114,67

Задайте число интервалов (градаций) в шкале:

В описательных шкалах:

Пересчитать шкалы и градации

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

"OK".

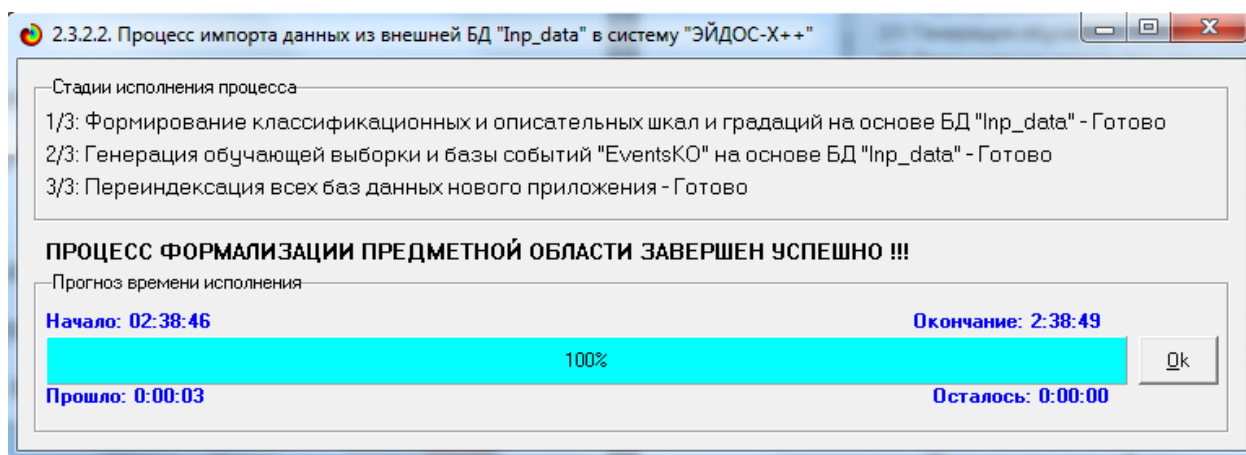


Рисунок 6. Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

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).

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

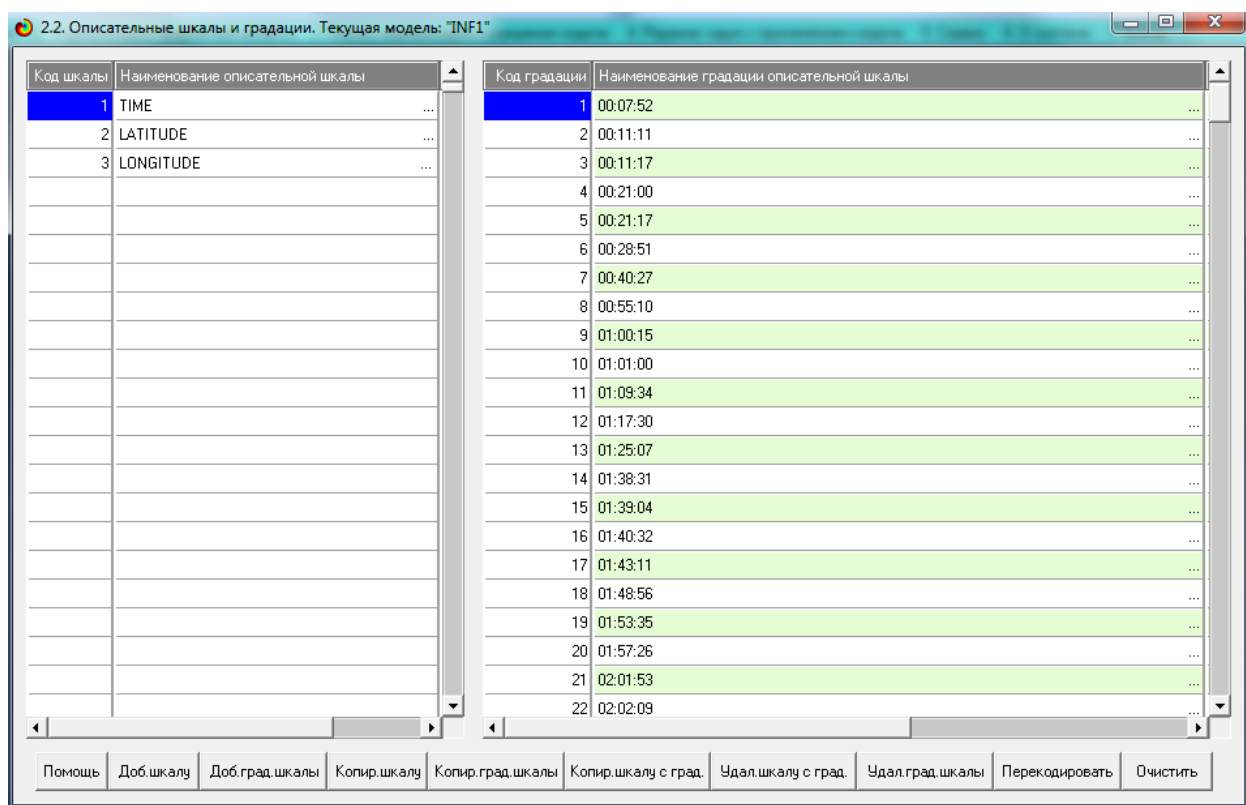


Рисунок 8. Описательные шкалы и градации (фрагмент)

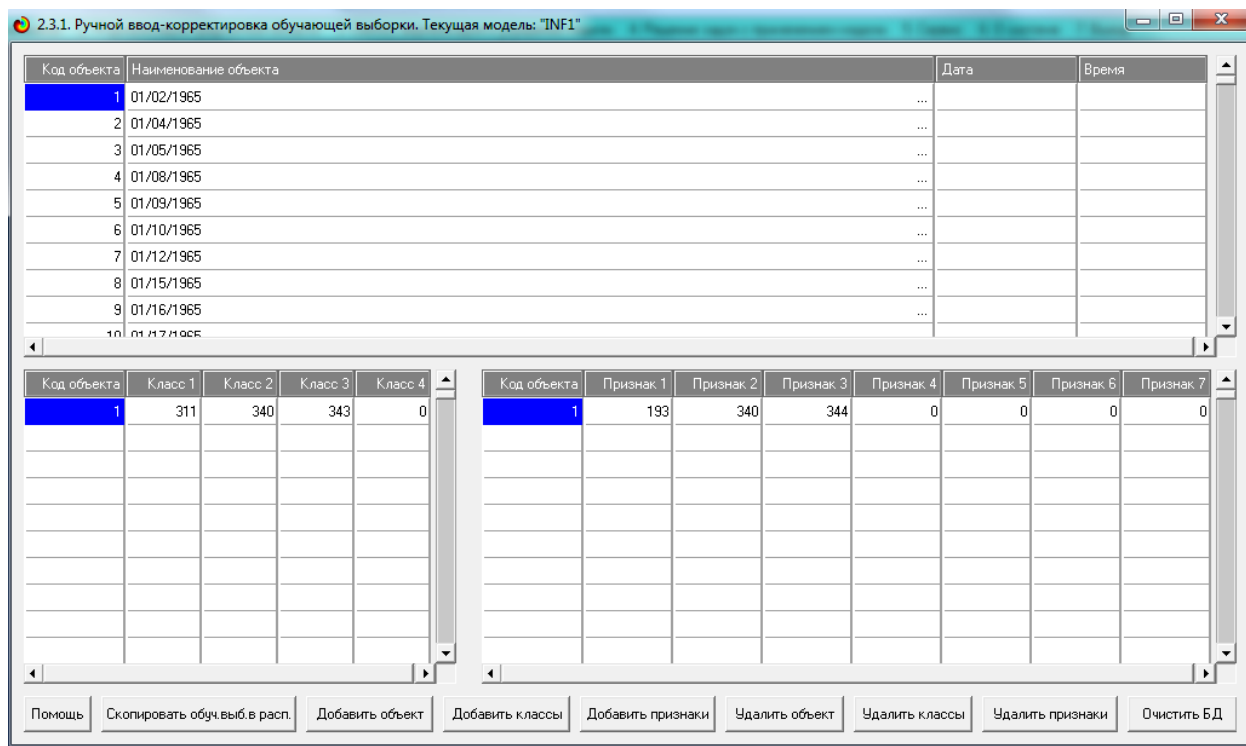


Рисунок 9. Обучающая выборка (фрагмент)

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).

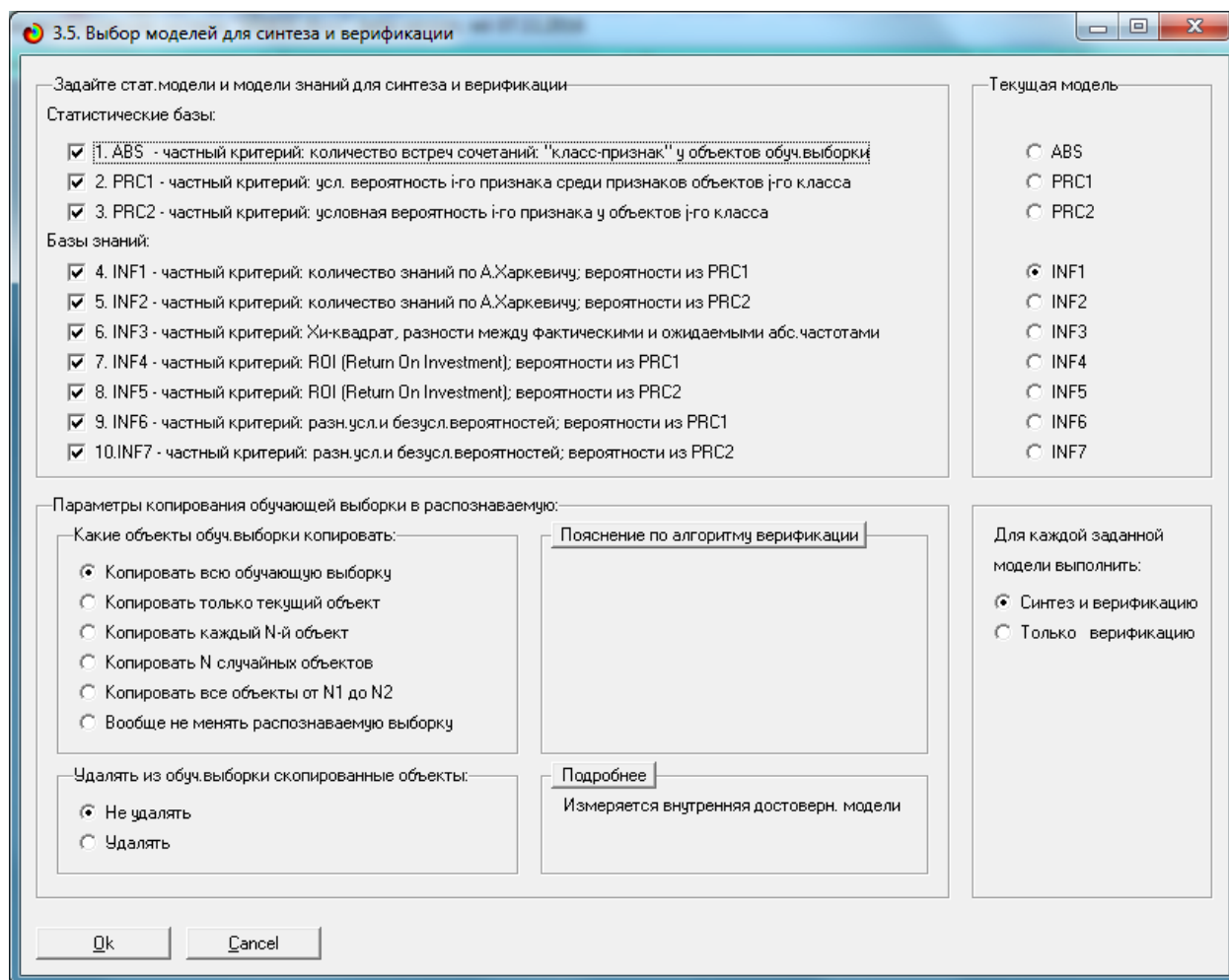


Рисунок 10. Выбор моделей для синтеза и верификации, а также текущей модели

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.

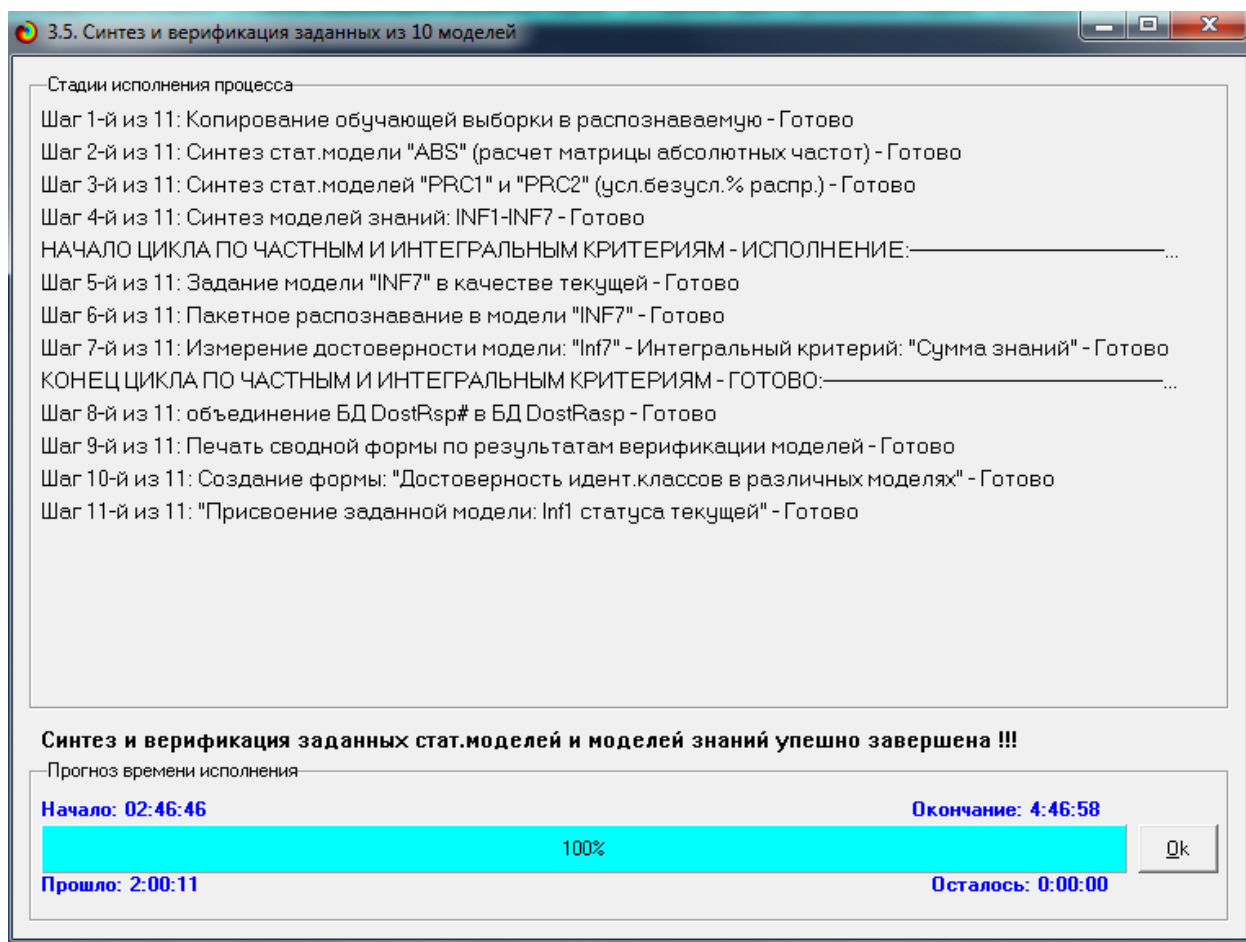


Рисунок 11. Синтез и верификация статистических моделей и моделей знаний

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)

Код признака	Наименование описательной шкалы и градации	1. ID ISCGEM851...	2. ID ISCGEM851...	3. ID ISCGEM851...	4. ID ISCGEM851...	5. ID ISCGEM851...	6. ID ISCGEM851...	7. ID ISCGEM851...	8. ID ISCGEM851...	9. ID ISCGEM851...	10. ID ISCGEM851...	11. ID ISCGEM851...	12. ID ISCGEM851...	13. ID ISCGEM851...	14. ID ISCGEM851...
1	TIME-00:07:52														
2	TIME-00:11:11														
3	TIME-00:11:17														
4	TIME-00:21:00														
5	TIME-00:21:17														
6	TIME-00:28:51														
7	TIME-00:40:27												100.000		
8	TIME-00:55:10														
9	TIME-01:00:15														
10	TIME-01:01:00														
11	TIME-01:09:34														
12	TIME-01:17:30														
13	TIME-01:25:07														
14	TIME-01:38:31														
15	TIME-01:39:04														
16	TIME-01:40:32														
17	TIME-01:43:11														
18	TIME-01:48:56														
19	TIME-01:53:35														
20	TIME-01:57:26														
21	TIME-02:01:53														
22	TIME-02:02:09														
23	TIME-02:06:35														

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

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

Код признака	Наименование описательной шкалы и градации	1. ID ISCGEM851...	2. ID ISCGEM851...	3. ID ISCGEM851...	4. ID ISCGEM851...	5. ID ISCGEM851...	6. ID ISCGEM851...	7. ID ISCGEM851...	8. ID ISCGEM851...	9. ID ISCGEM851...	10. ID ISCGEM851...	11. ID ISCGEM851...	12. ID ISCGEM851...	13. ID ISCGEM851...
1	TIME-00:07:52													
2	TIME-00:11:11													
3	TIME-00:11:17													
4	TIME-00:21:00													
5	TIME-00:21:17													
6	TIME-00:28:51													
7	TIME-00:40:27												1	
8	TIME-00:55:10													
9	TIME-01:00:15													
10	TIME-01:01:00													
11	TIME-01:09:34													
12	TIME-01:17:30													
13	TIME-01:25:07													
14	TIME-01:38:31													
15	TIME-01:39:04													
16	TIME-01:40:32													
17	TIME-01:43:11													
18	TIME-01:48:56													
19	TIME-01:53:35													
20	TIME-01:57:26													
21	TIME-02:01:53													
22	TIME-02:02:09													
23	TIME-02:06:35													

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. ID ISCGEM851...	2. ID ISCGEM851...	3. ID ISCGEM851...	4. ID ISCGEM851...	5. ID ISCGEM851...	6. ID ISCGEM851...	7. ID ISCGEM851...	8. ID ISCGEM851...	9. ID ISCGEM851...	10. ID ISCGEM851...	11. ID ISCGEM851...	12. ID ISCGEM851...	13. ID ISCGEM851...	14. ID ISCGEM851...
1	TIME-00:07:52														
2	TIME-00:11:11														
3	TIME-00:11:17														
4	TIME-00:21:00														
5	TIME-00:21:17														
6	TIME-00:28:51														
7	TIME-00:40:27												6.119		
8	TIME-00:55:10														
9	TIME-01:00:15														
10	TIME-01:01:00														
11	TIME-01:09:34														
12	TIME-01:17:30														
13	TIME-01:25:07														
14	TIME-01:38:31														
15	TIME-01:39:04														
16	TIME-01:40:32														
17	TIME-01:43:11														
18	TIME-01:48:56														
19	TIME-01:53:35														
20	TIME-01:57:26														
21	TIME-02:01:53														
22	TIME-02:02:09														
23	TIME-02:06:35														

1.5. Результаты верификации моделей

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

Наименование модели и частного критерия	Интегральный критерий	Всего 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	Фигера Ван Ризбергера	Сумма модулей уровней сходства истинно-отрицательных решений (ST)	Сумма модулей уровней сходства ложно-положительных решений (SFP)	Сумма модулей уровней сходства ложно-отрицательных решений (SF)
1. ABS - частный критерий: количество встреч сочетаний "класс"	Корреляция абс. частот с абс. частотами по признаку...	1017	1017	36384	79215	0	0.013	1.000	0.025	712.597	320.094	34941.191
1. ABS - частный критерий: количество встреч сочетаний "класс"	Средняя абс. частот по признаку...	1017	1017	36384	79215	0	0.013	1.000	0.025	527.131	301.128	
2. PRIC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отн. частот с о...	1017	1017	36384	79215	0	0.013	1.000	0.025	712.586	320.094	34941.190
2. PRIC1 - частный критерий: усл. вероятность иго признака сред.	Средняя усл.отн. частот по признаку...	1017	1017	36384	79215	0	0.013	1.000	0.025	549.416	35262.414	
3. PRIC2 - частный критерий: условная вероятность иго признака.	Корреляция усл.отн. частот с о...	1017	1017	36384	79215	0	0.013	1.000	0.025	712.570	320.094	34941.185
3. PRIC2 - частный критерий: условная вероятность иго признака.	Средняя усл.отн. частот по признаку...	1017	1017	36384	79215	0	0.013	1.000	0.025	549.416	35262.414	
4. INF1 - частный критерий: количество знаний по АХаркевичу, в.	Семантический резонанс зна...	1017	427	36705	78894	590	0.005	0.420	0.011	263.915	278.266	6814.349
4. INF1 - частный критерий: количество знаний по АХаркевичу, в.	Средняя знаний	1017	865	36603	78996	152	0.011	0.851	0.021	257.589	18.765	6950.291
5. INF2 - частный критерий: количество знаний по АХаркевичу, в.	Семантический резонанс зна...	1017	427	36705	78894	590	0.005	0.420	0.011	263.945	278.297	6815.110
5. INF2 - частный критерий: количество знаний по АХаркевичу, в.	Средняя знаний	1017	865	36603	78996	152	0.011	0.851	0.021	257.589	18.765	6950.291
6. INF3 - частный критерий: Хинкватер, разности между факти...	Семантический резонанс зна...	1017	710	71155	44444	307	0.016	0.698	0.031	400.613	14004.212	15494.772
6. INF3 - частный критерий: Хинкватер, разности между факти...	Средняя знаний	1017	710	71155	44444	307	0.016	0.698	0.031	177.171	4512.628	4398.201
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	1017	427	106547	9052	590	0.045	0.420	0.081	211.264	443.108	33.646
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Средняя знаний	1017	888	36580	79019	129	0.011	0.873	0.022	327.541	0.287	348.504
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	1017	427	106547	9052	590	0.045	0.420	0.081	211.264	443.108	33.646
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Средняя знаний	1017	888	36580	79019	129	0.011	0.873	0.022	327.541	0.287	348.504
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер.	Семантический резонанс зна...	1017	710	36580	79019	307	0.009	0.698	0.018	392.675	352.801	22264.262
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер.	Средняя знаний	1017	710	36580	79019	307	0.009	0.698	0.018	256.434	12.452	17620.195
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	1017	710	36580	79019	307	0.009	0.698	0.018	392.675	352.801	22264.262
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Средняя знаний	1017	710	36580	79019	307	0.009	0.698	0.018	256.434	12.452	17620.195

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 осуществится, 100% достоверность идентификации и не идентификации. Идеальный прогноз, который полностью снимает неопределенность о будущем состоянии объекта прогнозирования, на практике удается получить крайне редко и обычно мы имеем дело с реальным прогнозом. Реальный прогноз. На практике мы чаще всего сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой. Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов. Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно относить объекты к классам, которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Этот критерий предложен и реализован в системе "Эйдос" проф. Е.В. Луценко в 1994 году. Эта мера достоверности модели предполагает два варианта нормировки: {-1, +1} и {0, 1}: $L1 = \frac{TP + TN - FP - FN}{TP + TN + FP + FN} \quad (\text{нормировка: } \{-1, +1\})$ $L2 = \frac{1 + (TP + TN - FP - FN)}{2} \quad (\text{нормировка: } \{0, 1\})$ где количество: TP - истинно-положительных решений; TN - истинно-отрицательных решений; FP - ложно-положительных решений; FN - ложно-отрицательных решений; Классическая F-мера достоверности моделей Ван Ризбергера (колонка выделена ярко-голубым фоном): $F\text{-мера} = 2 * (Precision * Recall) / (Precision + Recall)$ $Precision = TP / (TP + FP) \quad \text{— точность модели;}$ $Recall = TP / (TP + FN) \quad \text{— полнота модели;}$ L1-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом сумм уровней сходства (колонка выделена ярко-зеленым фоном): $L1\text{-мера} = 2 * (SPrecision * SRecall) / (SPrecision + SRecall)$ $SPrecision = STP / (STP + SFP) \quad \text{— точность с учетом сумм уровней сходства;}$ $SRecall = STP / (STP + SFN) \quad \text{— полнота с учетом сумм уровней сходства;}$ STP - сумма модулей сходства истинно-положительных решений; STN - сумма модулей сходства истинно-отрицательных решений; SFP - сумма модулей сходства ложно-положительных решений; SFN - сумма модулей сходства ложно-отрицательных решений. L2-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом средних уровней сходства (колонка выделена желтым фоном): $L2\text{-мера} = 2 * (APrecision * ARecall) / (APrecision + ARecall)$ $APrecision = ATP / (ATP + AFP) \quad \text{— точность с учетом средних уровней сходства;}$ $ARecall = ATP / (ATP + AFN) \quad \text{— полнота с учетом средних уровней сходства;}$ ATP-STP/TP - Среднее модулей сходства истинно-положительных решений; AFN-SFN/FN - Среднее модулей сходства истинно-отрицательных решений; AFP-SFP/FP - Среднее модулей сходства ложно-положительных решений; AFN-SFN/FN - Среднее модулей сходства ложно-отрицательных решений. Строки с максимальными значениями F-меры, L1-меры и L2-меры выделены фоном цвета, соответствующего колонке. Луценко Е.В. Инвариантно относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе "Эйдос" // Е.В. Луценко / Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. - Краснодар: КубГАУ, 2017. - №02(128). С. 1 - 32. - doi [Article ID]: 1261702021. - Режим доступа: http://ej.kubagro.ru/2017/02/pdf/01.pdf, 2 у.н.н.

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

модели

INF5.

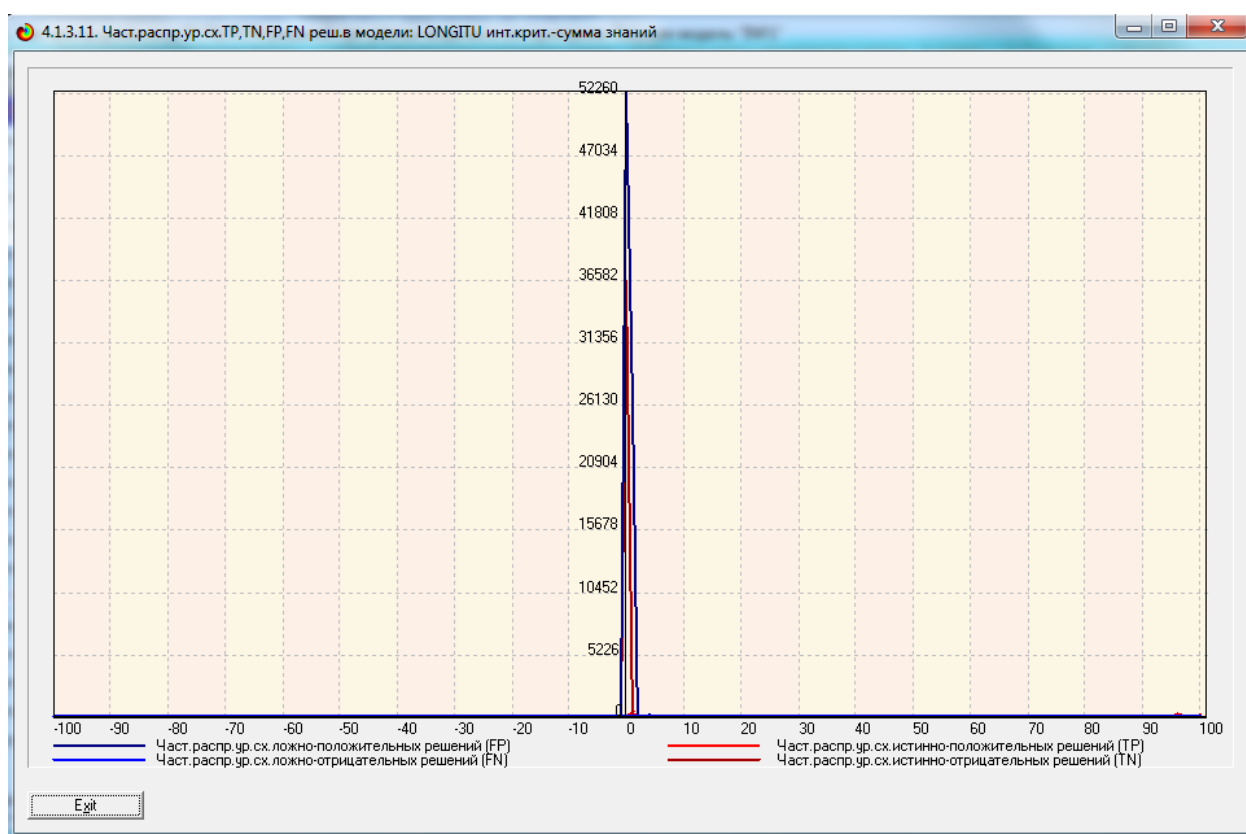


Рисунок 14. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF3

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.

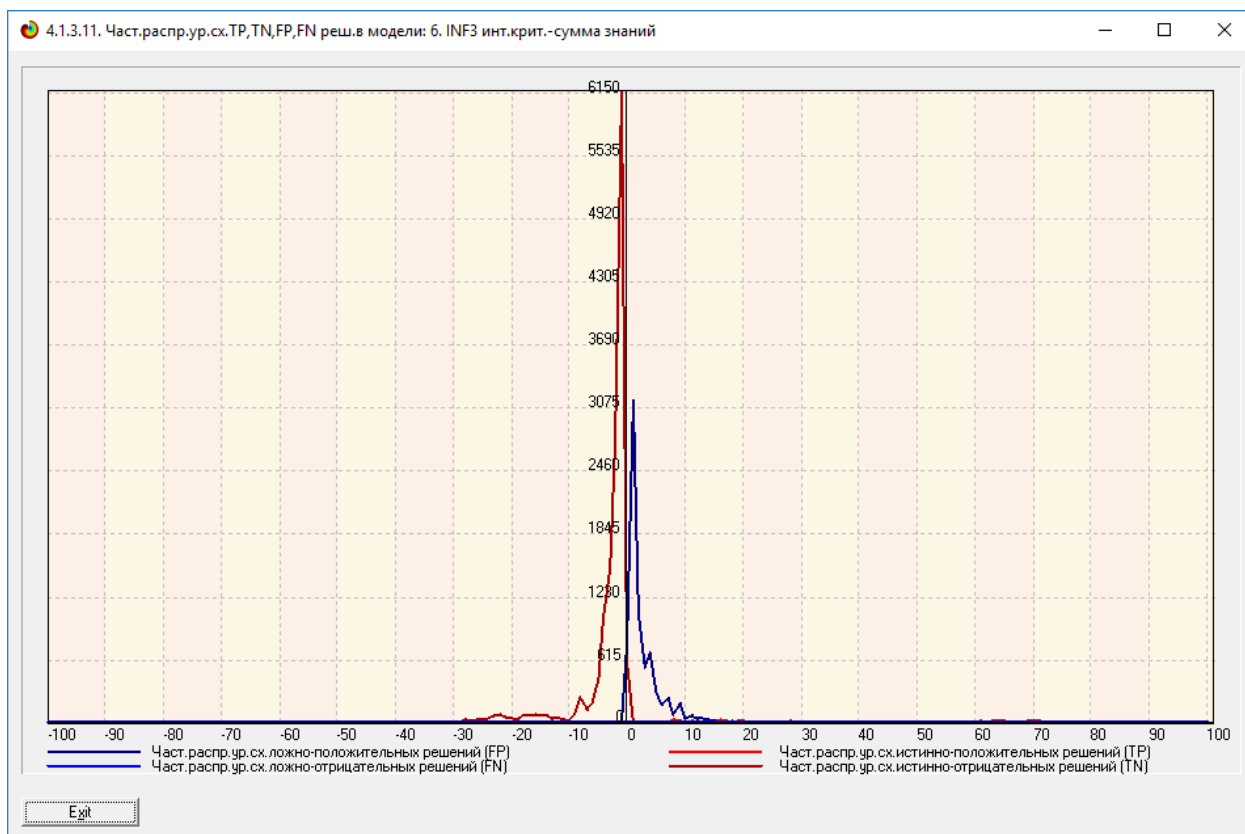


Рисунок 16. Частное распределение сходства-различия верно и ошибочно идентифицированных и не идентифицированных состояний в случайной модели INF3

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. крит. Тенущая модель: "INF1"													
Наименование модели и частного критерия	Интегральный критерий	Точность модели	Полнота модели	Фигура Рэндалга	Среднее между-уровней сп. истинно-поло-решений (ST)	Среднее между-уровней сп. истинно-отри-решений (ST)	Среднее между-уровней сп. ложно-поло-решений (SFP)	Среднее между-уровней сп. ложно-отри-решений (SF)	S-Точность модели	S-Полнота модели	LT-мера проф. Е. В. Пучкова	Средняя нагрузка уровней сп. истинно-поло-решений	Средняя нагрузка уровней сп. истинно-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	
2. PRIC1 - частный критерий: усл. вероятность i-го признака сред...	Средняя абс. частот по призна...	0.048	1.000	0.092	250.407		457.085		0.354	1.000	0.523	0.420	
3. PRIC2 - 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	
4. INF1 - частный критерий: количество знаний по АХаркевичу, в...	Средняя усл. отн. частот по при...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
5. INF2 - частный критерий: количество знаний по АХаркевичу, в...	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	
6. INF3 - частный критерий: Хинкватер, разности между фактин...	Средняя усл. отн. частот по при...	0.048	1.000	0.092	324.371		2201.486		0.128	1.000	0.228	0.544	
7. INF4 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
8. INF5 - частный критерий: ROI (Return On Investment), веро...	Семантический резонанс зна...	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.066	1.000	0.124	235.693	291.390	601.005		0.282	1.000	0.440	0.395	
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.057	1.000	0.107	290.310	45.691	764.644		0.275	1.000	0.432	0.487	
11. INF8 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.074	1.000	0.138	403.194	1799.085	1322.555		0.234	1.000	0.379	0.677	
12. INF9 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.074	1.000	0.138	254.215	455.539	201.323		0.558	1.000	0.716	0.427	
13. INF10 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
14. INF11 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
15. INF12 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.205	1.000	0.340	250.112	367.178	26.580		0.904	1.000	0.950	0.420	
16. INF13 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.057	1.000	0.107	71.867	0.892	46.101		0.609	1.000	0.757	0.121	
17. INF14 - частный критерий: разн. усл. и безуслов. 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	
18. INF15 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	
19. INF16 - частный критерий: разн. усл. и безуслов. 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	
20. INF17 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	0.057	1.000	0.107	274.607	37.609	1555.969		0.150	1.000	0.261	0.461	

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 INF3 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.1.2.

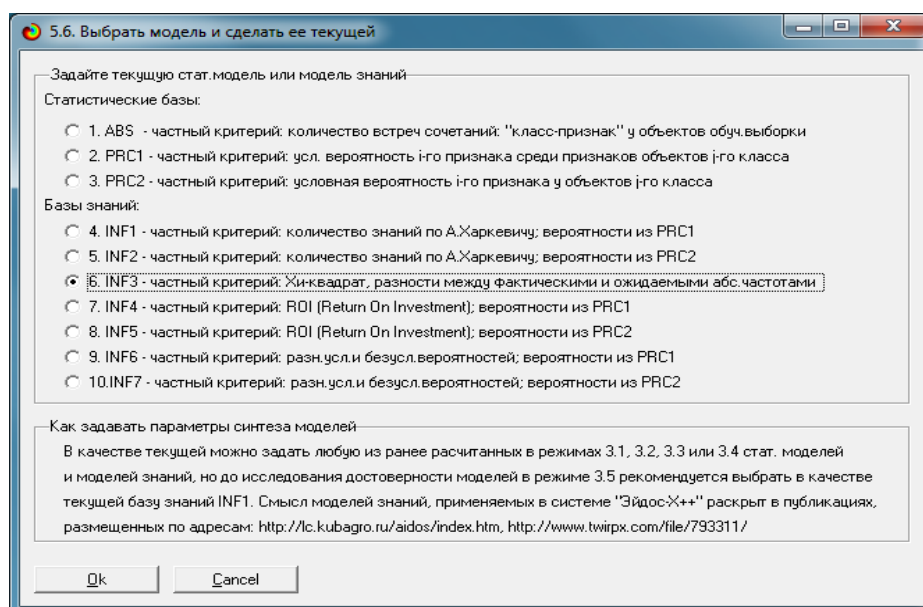
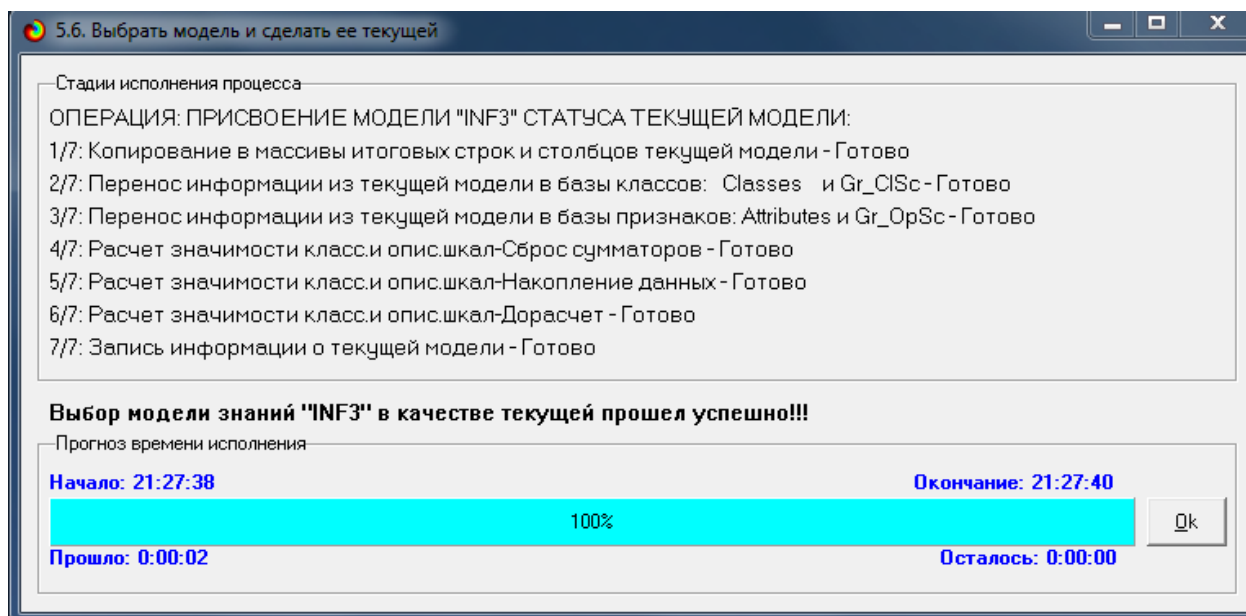


Рисунок 18. Экранные формы режима задания модели в качестве текущей

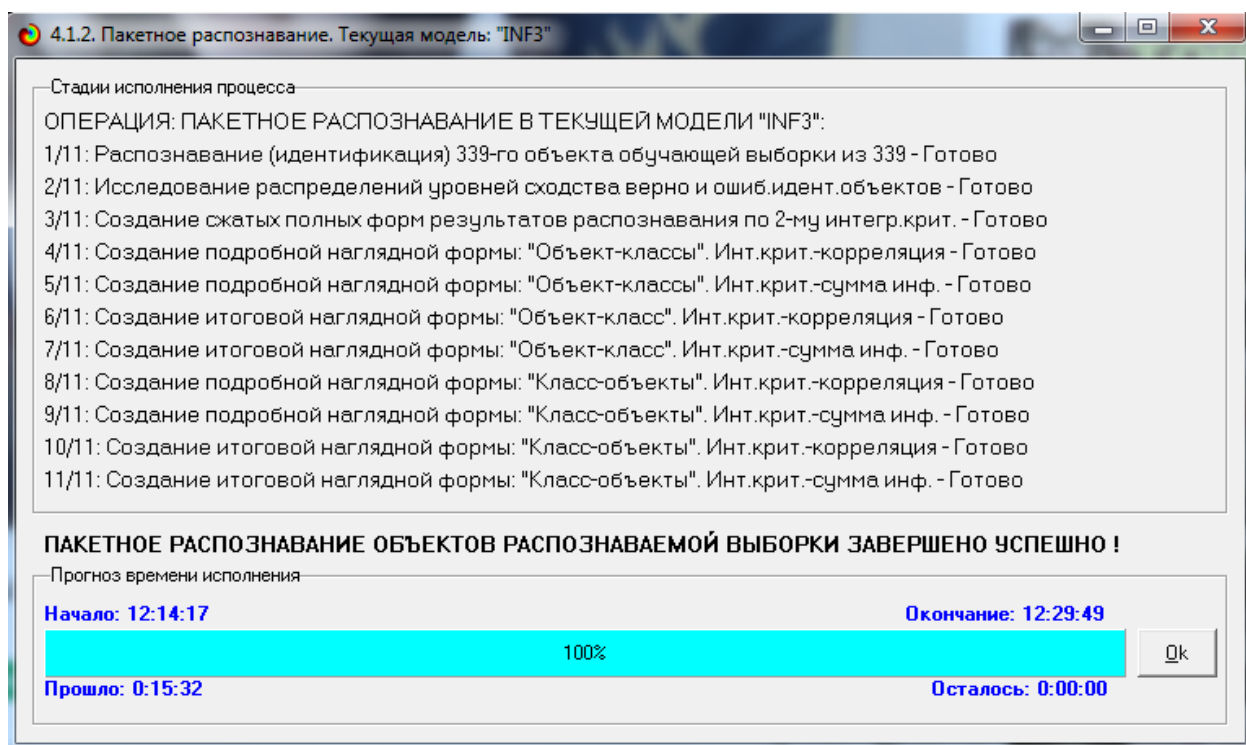


Рисунок 19. Экранная форма режима пакетного распознавания в текущей модели

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:

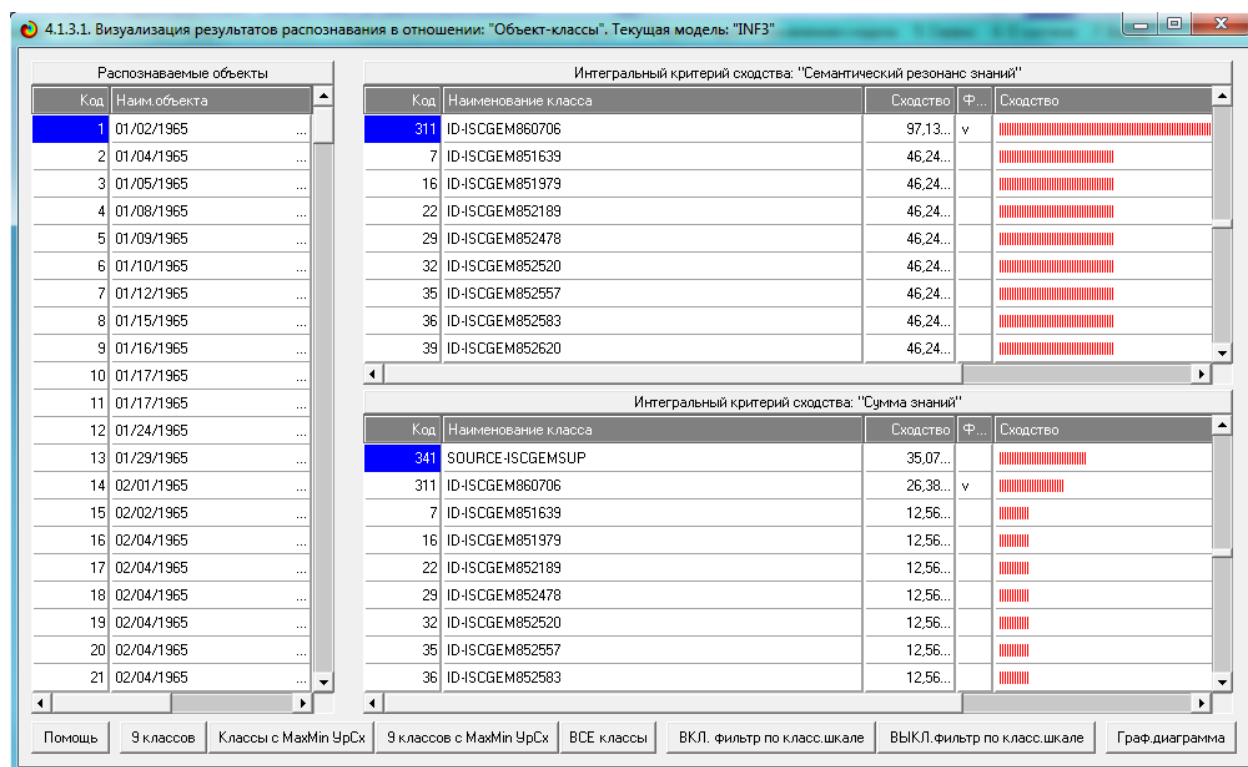


Рисунок 20. Пример идентификации классов в модели INF3

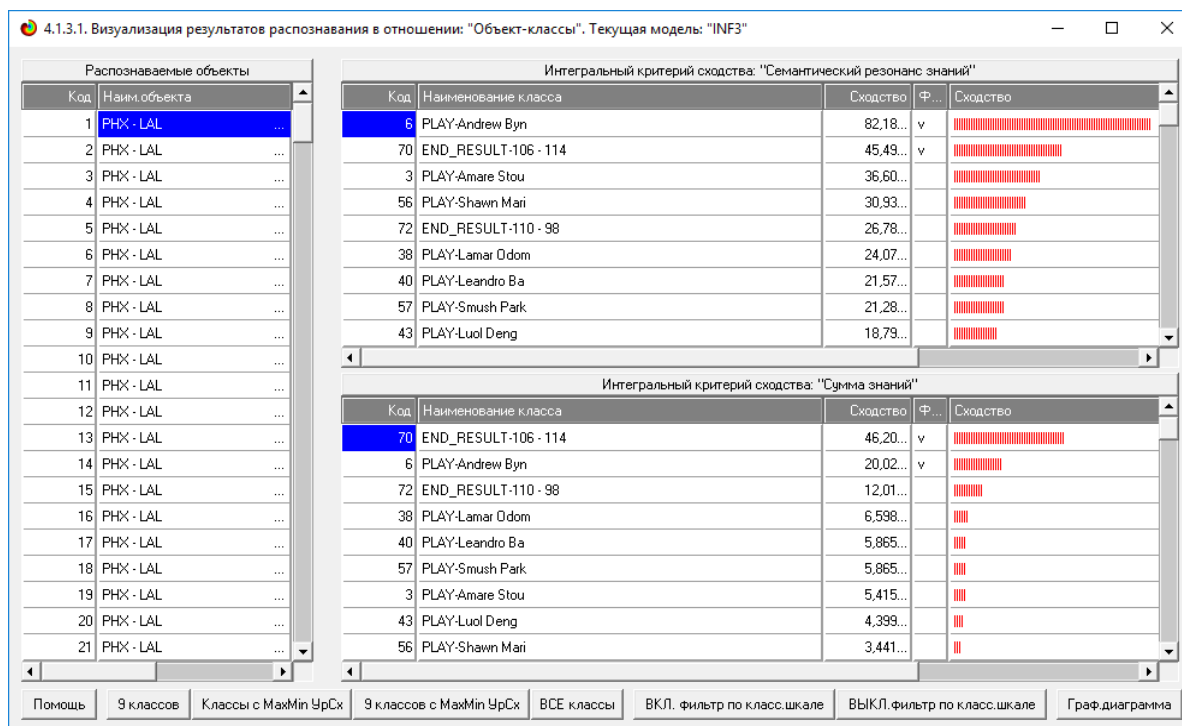


Рисунок 21. Пример идентификации классов в модели INF3

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)

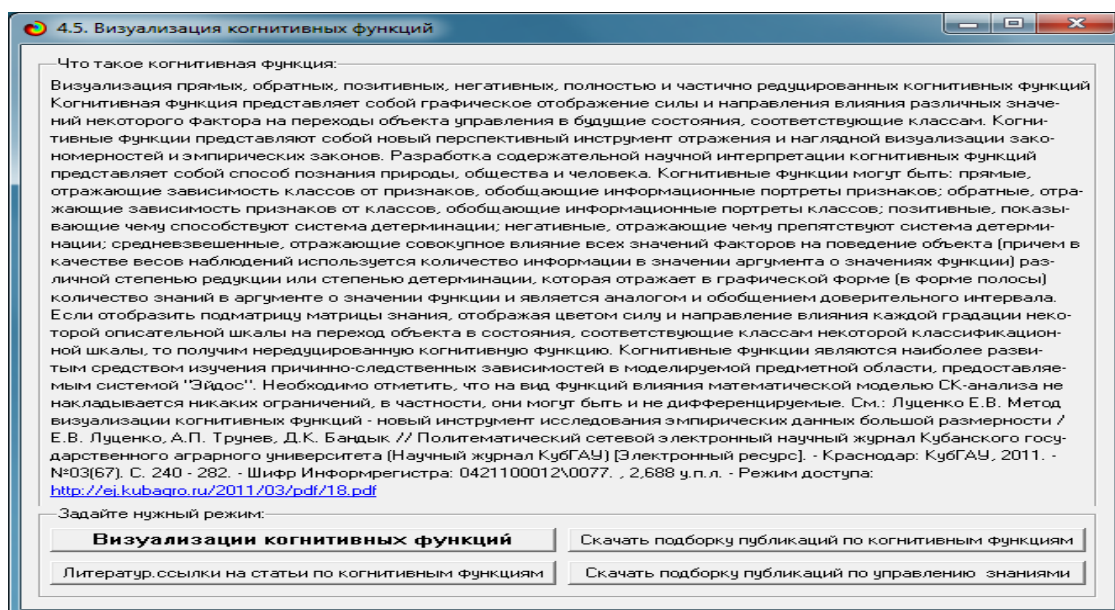


Рисунок 22. Экранная форма режима 4.5 системы «Эйдос-X++» «Визуализация когнитивных функций»

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 INF5.

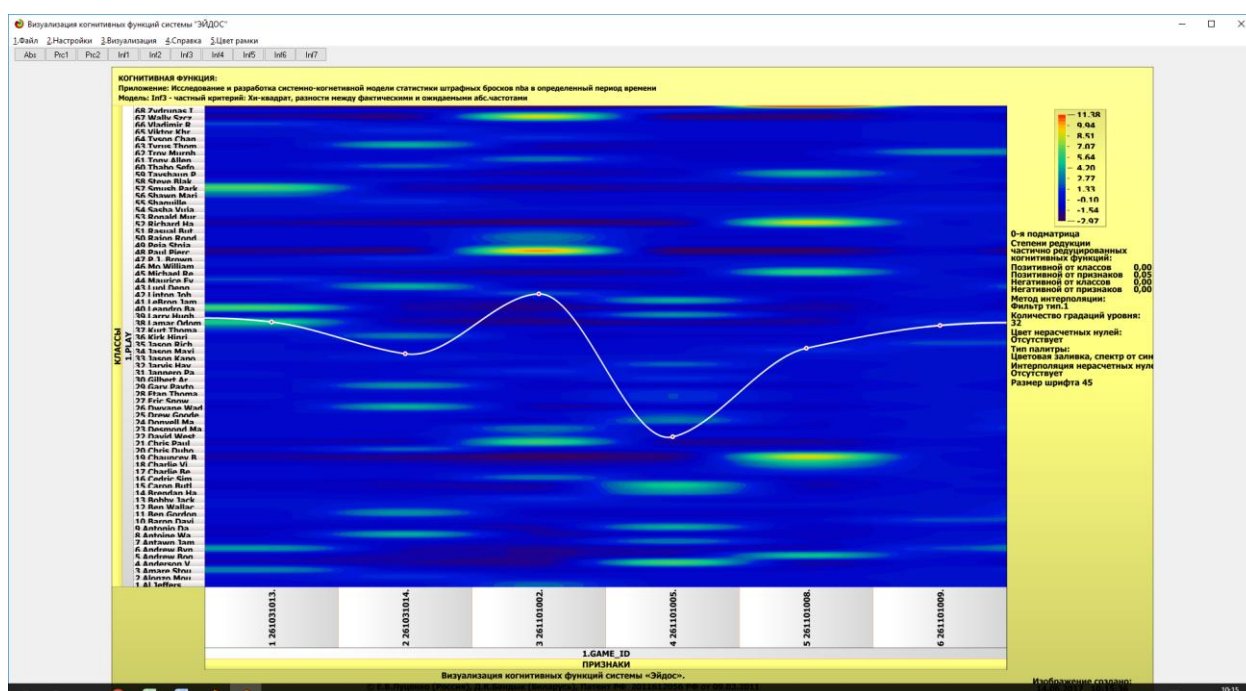


Рисунок 23. Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF3

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-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 24).

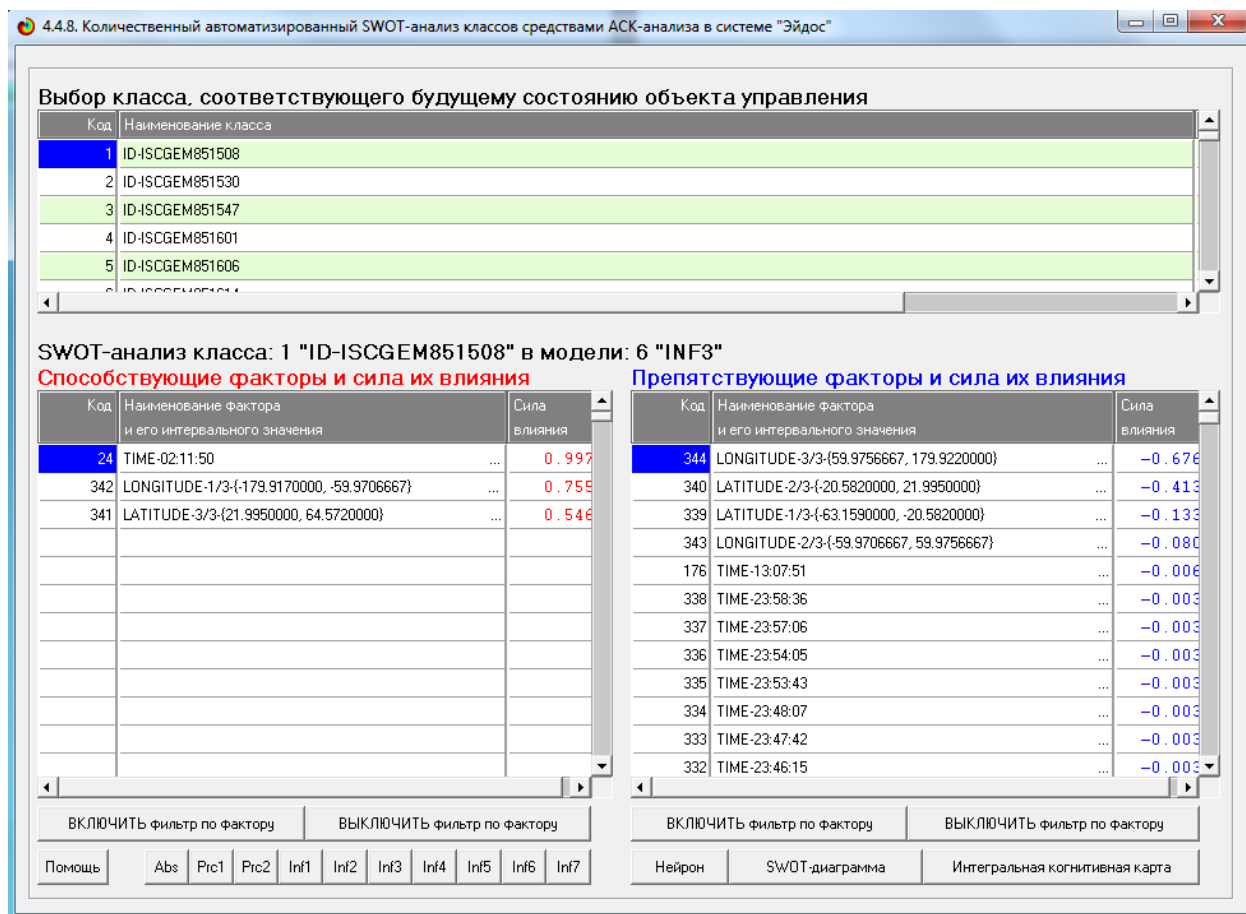


Рисунок 24. Пример SWOT-матрицы в модели INF3

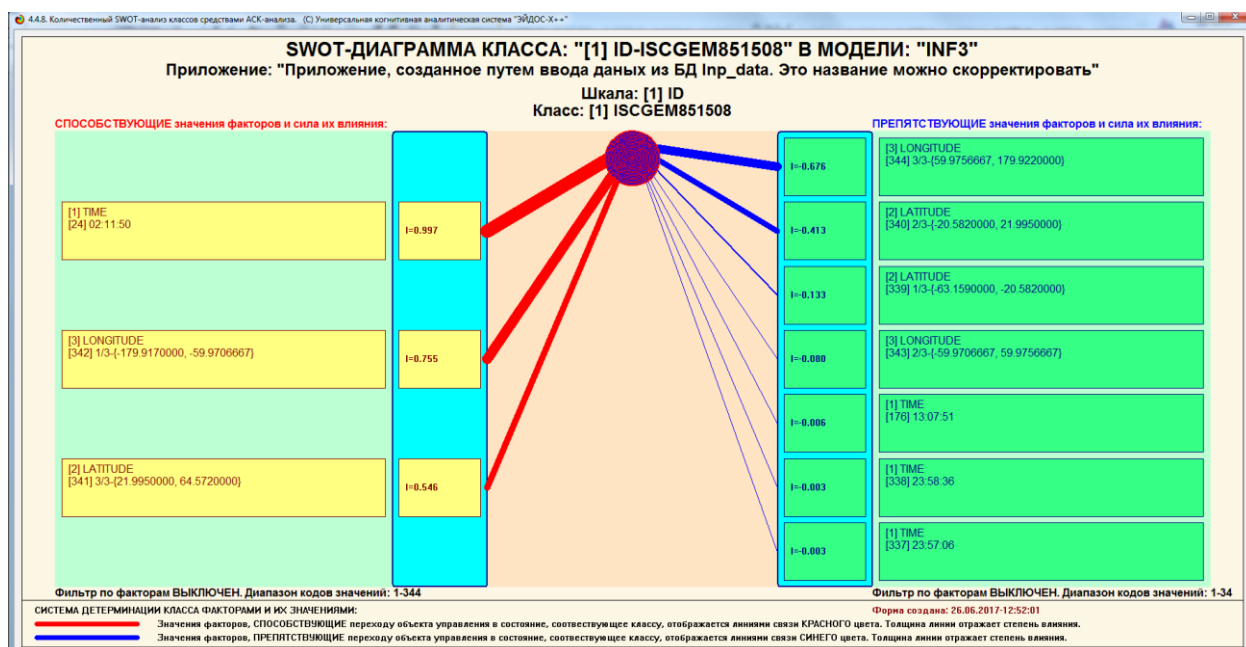


Рисунок 25 - Пример SWOT-матрицы в модели INF3

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.

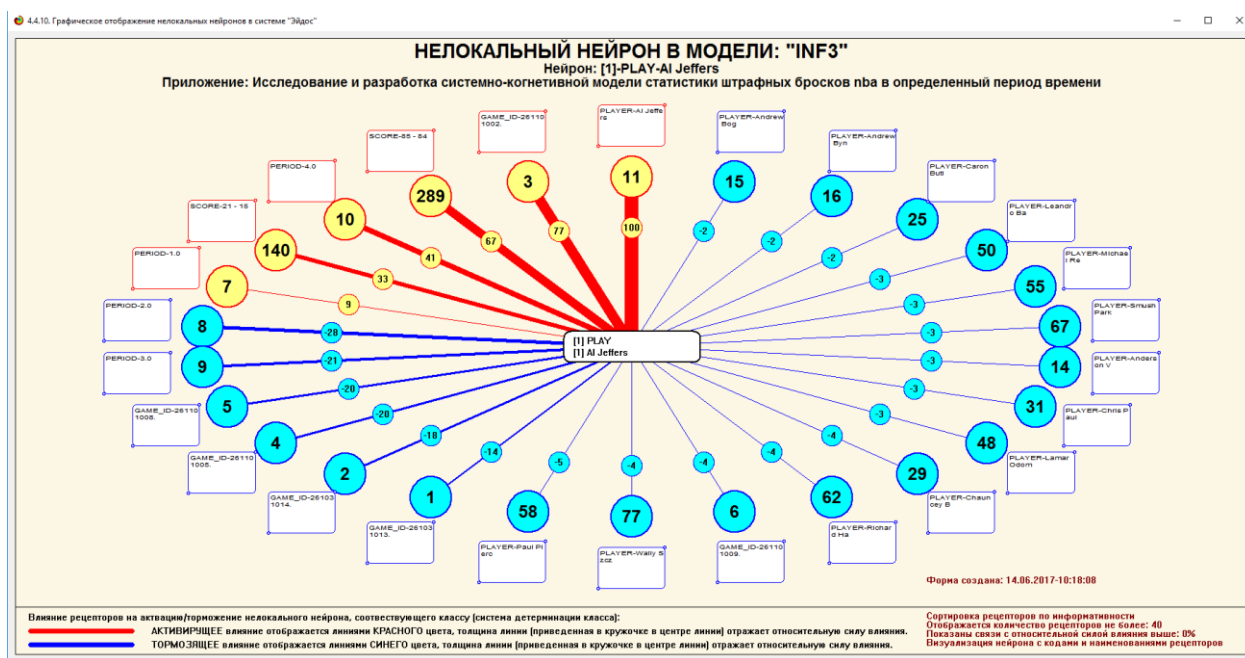


Рисунок 26. Пример нелокальные нейронные сети и нейроны модели INF3

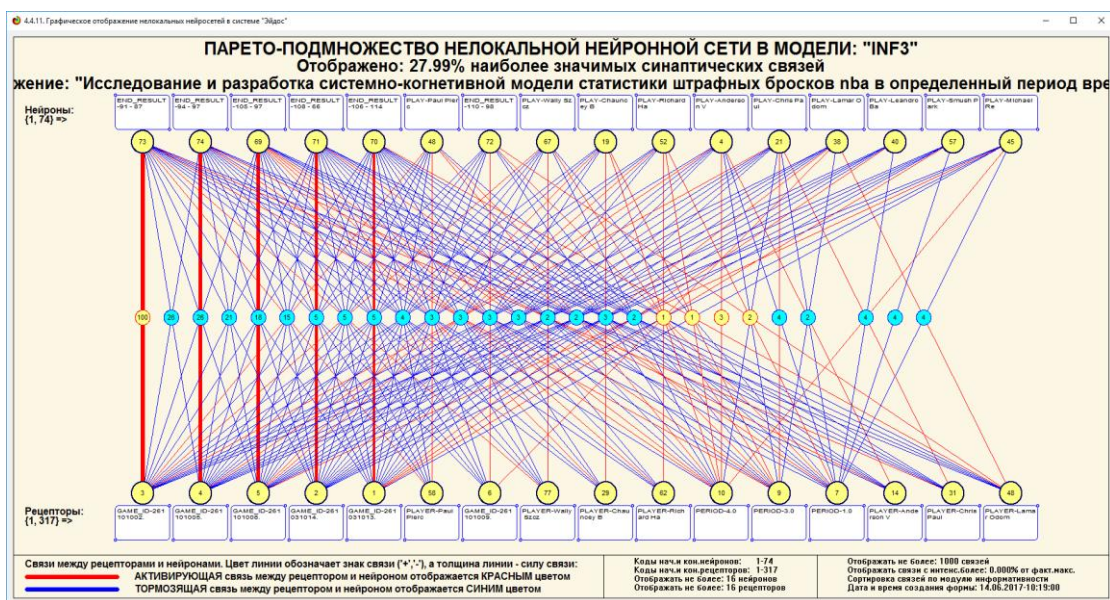


Рисунок 26. Пример нелокальные нейронные сети и нейроны модели INF3

2.5 КЛАСТЕРНЫЙ И КОНСТРУКТИВНЫЙ АНАЛИЗ

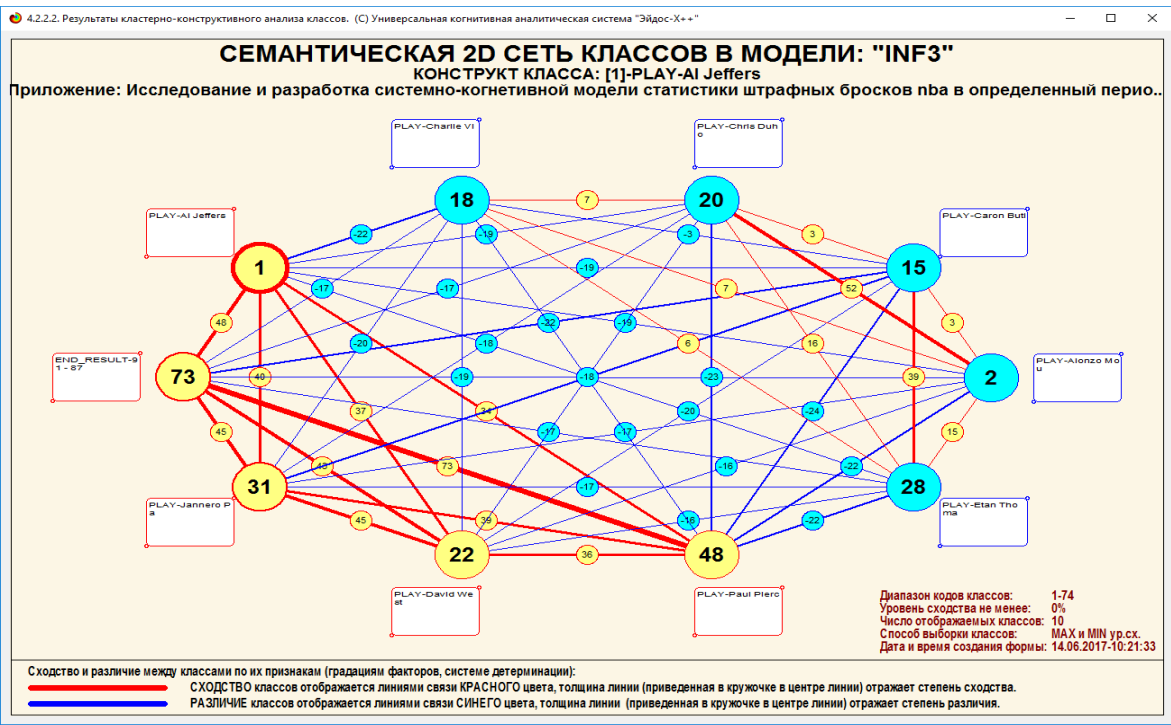


Рисунок 27. Пример кластерный и конструктивный анализ модели INF5

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-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.

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

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>