

Министерство сельского хозяйства российской федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
им. И.Т. Трубилина

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

Лабораторная работа

по дисциплине: Интеллектуальные информационные системы

на тему:

АСК-анализ президентских выборов в России 2018 года по
Краснодарскому краю

выполнил студент группы: ПИ1521

Карзенкова Ольга Сергеевна

Руководитель работы:
профессор Луценко Е.В.

Краснодар 2018

Содержание

ВВЕДЕНИЕ.....	3
1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	4
1.1. Описание решения	4
1.2. Преобразование исходных данных в файл MSExcel.....	4
1.3. Синтез и верификация статистических и интеллектуальных моделей.....	9
1.4. Виды моделей системы «Эйдос».....	10
1.5. Результаты верификации моделей	12
2.РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	16
2.1.Решение задачи	16
2.2 Когнитивные функции.....	19
2.3. SWOT и PEST матрицы и диаграммы	21
2.4. Нелокальные нейронные сети	23
2.5 Кластерный и конструктивный анализ.....	24
ЗАКЛЮЧЕНИЕ	25
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	26

ВВЕДЕНИЕ

Одним из важнейших и перспективных направлений в развитии современных информационных технологий, на сегодняшний день, является создание систем искусственного интеллекта. В результате развития технологий в данном направлении, возникла необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи анализа результатов президентских выборов в Российской Федерации, проведенных в 2018 году на базе данных Краснодарского края.

При выполнении работы нужно учесть, что необходимы свободный доступ к тестовым исходным данным и методика, которая необходима для преобразования этих данных в форму, которая необходима для работы в системе искусственного интеллекта. Среди баз данных текстовых задач для систем искусственного интеллекта удачным выбором является репозиторий «Kaggle».

В данной лабораторной работе использована база данных «*Russian Presidential Elections 2018*» из банка исходных данных по задачам искусственного интеллект-репозитория «Kaggle».

Для решения задачи используем стандартные возможности Microsoft Office Excel, блокнот, а также систему искусственного интеллекта «Aidos-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 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 в файл MS Excel

Исходные данные взяты с сайта Kaggle, режим доступа по ссылке: <https://www.kaggle.com/rexhaif/ru-elections-2018/data>.

Из электронного ресурса баз данных «Kaggle» возьмем базу данных «*Russian Presidential Elections 2018*», которую приведем к виду, показанному на рисунке 1.

	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	onsite_ballots	offsite_ballots	unused_ballots	found_offsite_ballots	found_onsite_ballots	invalid_ballots	valid_ballots	lost_ballots	unaccounted_ballots	baburnin	grudinin	zhirnovsky	putin
2	1447	11	542	11	1447	9	1449	0	0	4	176	79	1136
3	1470	14	516	14	1470	14	1470	0	0	2	128	87	1214
4	1490	7	503	7	1490	27	1470	0	0	5	171	94	1162
5	1065	48	387	48	1065	20	1093	0	0	5	98	72	882
6	1171	13	316	13	1171	10	1174	0	0	7	124	105	902
7	1302	13	585	13	1302	14	1301	0	0	1	138	96	1022
8	1546	62	292	62	1546	10	1598	0	0	8	191	95	1250
9	4557	46	5703	740	4557	175	5122	0	0	46	159	80	4779
10	3235	52	313	52	3235	136	3151	0	0	18	32	22	3062
11	2713	17	6005	282	2713	149	2846	0	0	19	122	68	2607
12	1807	12	881	12	1807	111	1708	0	0	12	50	27	1597
13	2691	8	5114	195	2691	6	2880	0	0	13	52	30	2762
14	1746	94	659	295	1746	26	2015	0	0	13	21	20	1937
15	411	14	628	761	411	43	1129	0	0	5	130	68	890
16	1175	160	0	1525	1175	88	2612	0	0	8	144	180	2220
17	188	0	67	15	187	1	201	0	0	4	26	7	143
18	747	239	1667	239	746	8	977	0	0	9	71	13	684
19	263	127	60	127	263	3	387	0	0	1	39	5	260
20	198	0	152	0	198	0	198	0	0	3	20	6	134
21	236	0	114	0	236	5	231	0	0	3	28	4	151
22	329	126	245	126	328	4	450	0	0	2	47	8	308
23	2244	6	950	6	2242	37	2211	0	0	8	228	39	1409
24	237	0	663	0	237	0	237	0	0	1	28	4	152
25	1431	9	7125	86	1431	14	1503	0	0	5	119	35	1252
26	715	0	1530	0	715	6	709	0	0	5	60	20	558
27	392	1	662	172	392	5	559	0	0	3	27	14	483
28	469	1	930	1	469	7	463	0	0	1	20	12	428
29	311	0	350	192	311	3	500	0	0	3	39	13	425
30	454	0	876	26	451	11	466	0	0	2	40	12	355
31	159	0	41	0	159	1	158	0	0	1	7	2	143
32	29	12	9	12	29	0	41	0	0	0	8	0	31
33	4724	28	1121	155	4720	15	4860	0	0	10	158	51	4417
34	651	0	1808	481	650	2	1129	0	0	9	85	77	936
35	140	0	201	0	140	1	139	0	0	2	12	1	111
36	535	33	6432	33	535	7	561	0	0	2	20	8	394
37	84	0	33	0	84	1	83	0	0	0	11	0	69
38	1303	8	426	8	1303	9	1302	0	0	2	28	5	1242

Рисунок 1 – Исходные данные в файле MS Excel

Поскольку ввод исходных данных в систему «Эйдос» планируется осуществить с помощью ее универсального программного интерфейса импорта данных из внешних баз данных, который работает с файлами MS Excel, то преобразуем данные из html-файла в xls-файл, для чего выполним следующие операции.

Скопируем таблицу в MS Excel и запишем ее с именем: *Inp_data.xlsx* в папку: C:\Aidos-X\AID_DATA\Inp_data\. В файле *Inp_data.xlsx* добавим пустую колонку на позиции «А» и автоматически пронумеруем все строки. В результате получим таблицу исходных данных, полностью подготовленную для обработки в системе «Эйдос» и записанную в нужную папку в виде файла нужного типа с нужным именем.

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, расположенный в режиме 2.3.2.2.

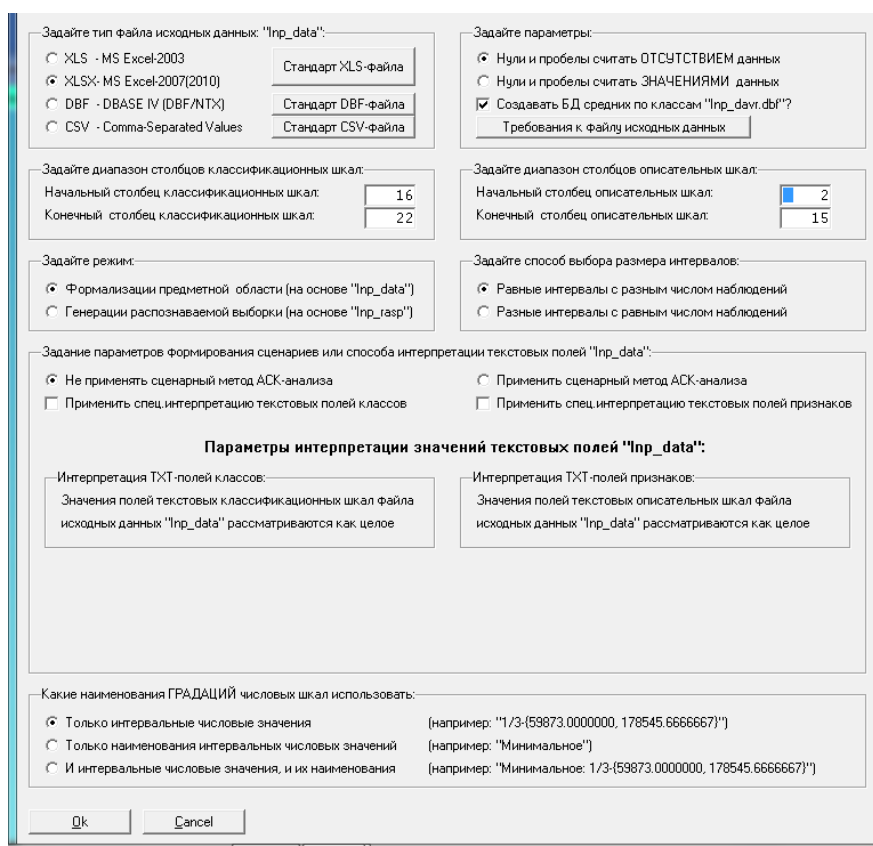


Рисунок 2 – Экранная форма универсального программного интерфейса импорта данных в систему «Эйдос»

В экранной форме, приведенной на рисунке 2, необходимо задать настройки:

— «Тип файла исходных данных Inp_data»: «XLS – MS Excel-2007»;

— «Диапазон столбцов классификационных шкал»: «Начальный столбец классификационных шкал» – 16, «Конечный столбец классификационных шкал» – 22;

— «Диапазон столбцов описательных шкал»: «Начальный столбец описательных шкал» – 2, «Конечный столбец описательных шкал» – 15;

— «Параметры формирования сценариев или способа интерпретации текстовых полей»: «Не применять сценарный метод АСК-анализа и спец. интерпретацию ТХТ-полей».

После нажать кнопку «ОК». После нажатия кнопки открывается окно, где размещена информация о размерности модели. В этом окне необходимо нажать кнопку «Выйти на создание модели», что показано на рисунке 3.

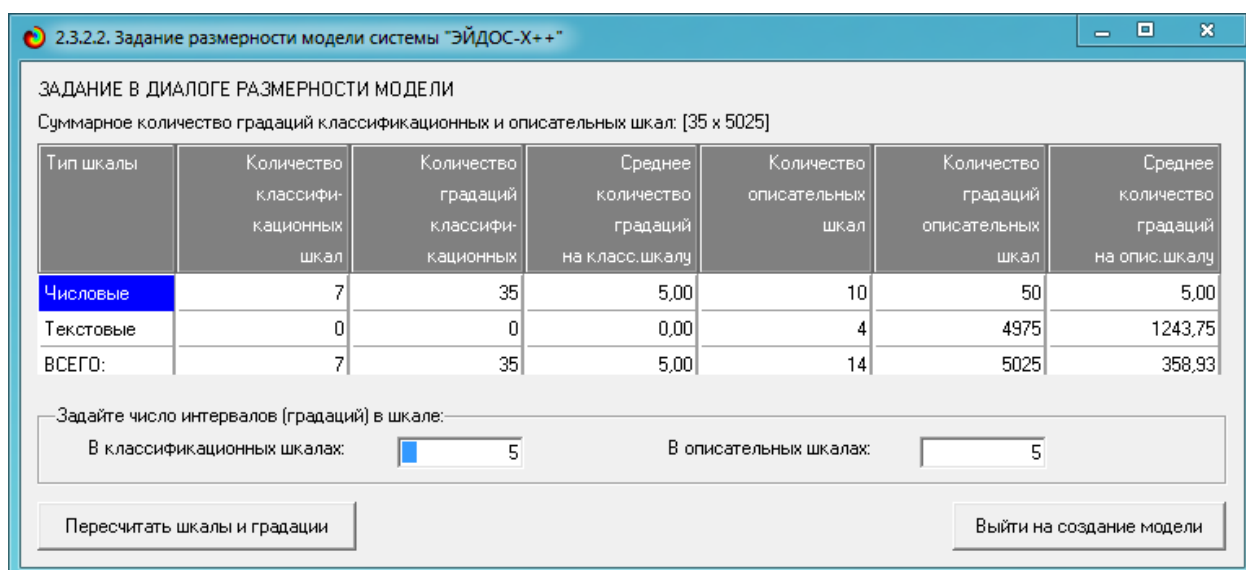


Рисунок 3 – Задание размерности модели системы «Эйдос»

Далее следует открытие окна, которое отображает стадию процесса импорта данных из внешней БД «Inp_data.xlsx» в систему «Эйдос», а также прогноз времени завершения этого процесса. В данном окне необходимо

дождаться завершения формализации предметной области и нажать кнопку «ОК», как показано на рисунке 4.

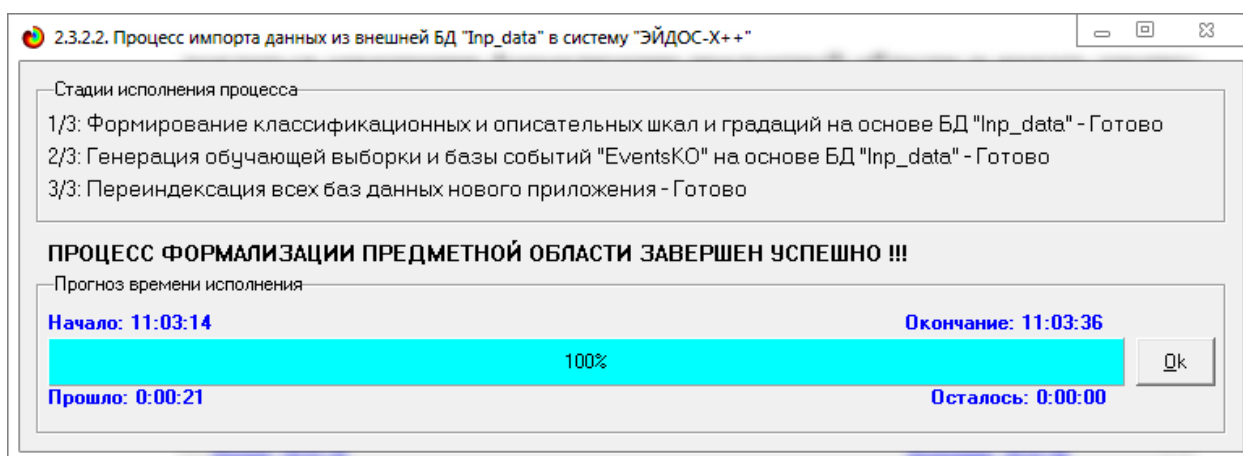


Рисунок 4 – Процесс импорта данных из внешней БД «Inp_data.xlsx» в систему «Эйдос»

В результате выполненных действий формируются классификационные и описательные шкалы, градации, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Эти действия полностью автоматизированы и выполняются на втором этапе АСК-анализа «Формализация предметной области». Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1, фрагмент показан на рисунке 5.

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы	DEL
1	BABURIN	1	1/5-(1.0000000, 23.0000000)	
2	GRUDININ	2	2/5-(23.0000000, 45.0000000)	
3	ZHIRINOVSKY	3	3/5-(45.0000000, 67.0000000)	
4	PUTIN	4	4/5-(67.0000000, 89.0000000)	
5	SOVCHAK	5	5/5-(89.0000000, 111.0000000)	
6	SURAYKIN			
7	TITOV			

Рисунок 5 – Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2, фрагмент которых отображен на рисунке 6.

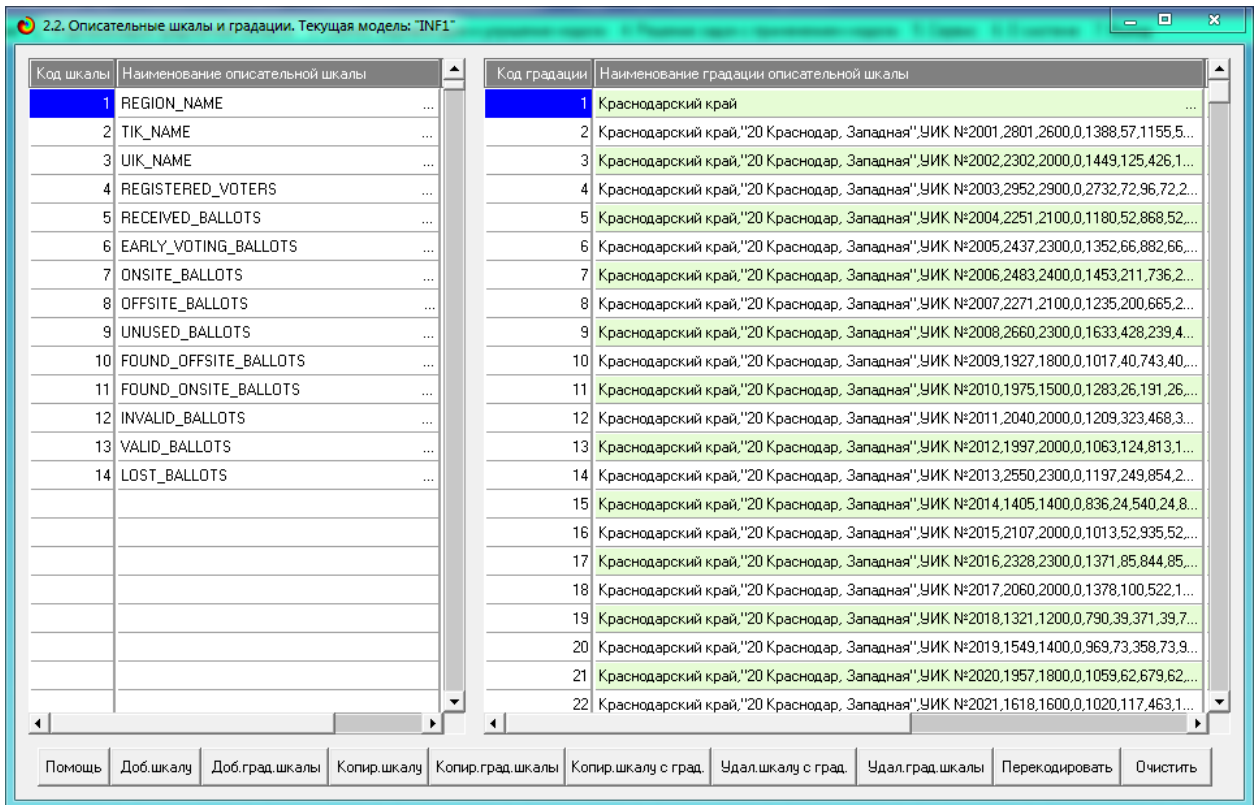


Рисунок 6 – Описательные шкалы и градации (фрагмент)

Для просмотра обучающей выборки необходимо запустить режим 2.3.1, показанный на рисунке 7.

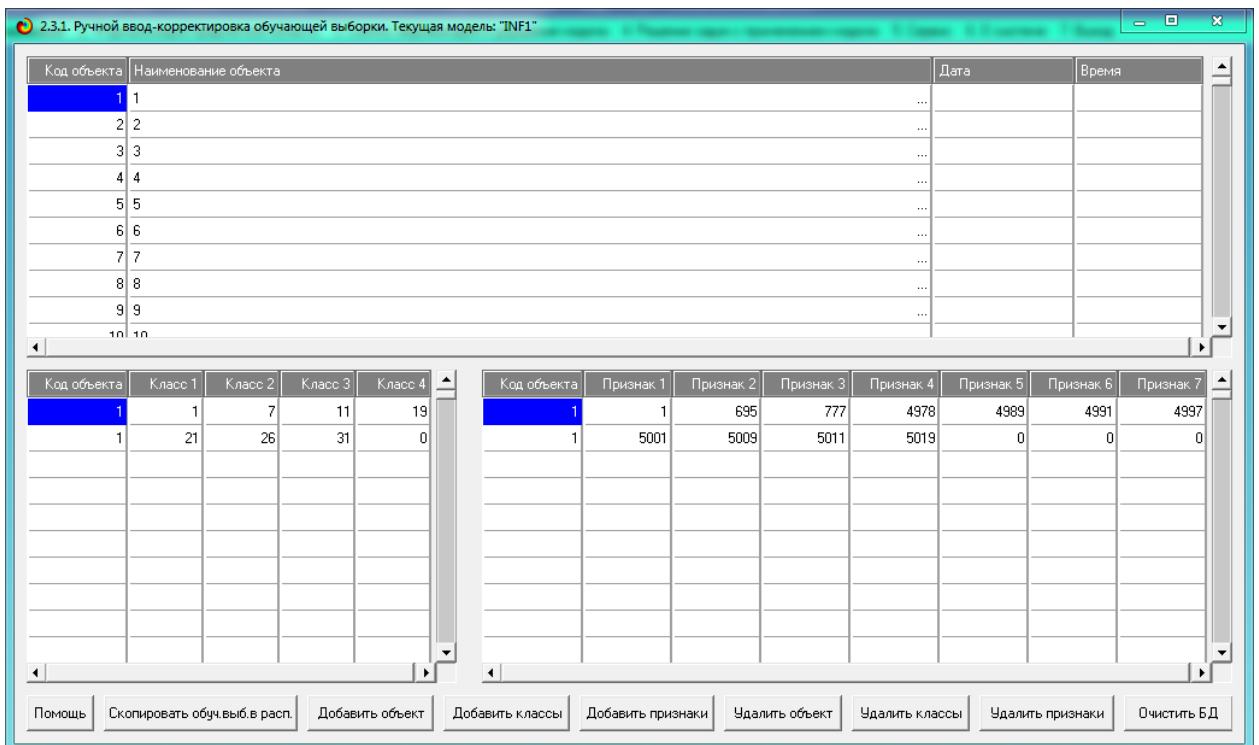


Рисунок 7 – Обучающая выборка (фрагмент)

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.

1.3. Синтез и верификация статистических и интеллектуальных моделей

Далее необходимо запустить режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей, что показано на рисунке 8.

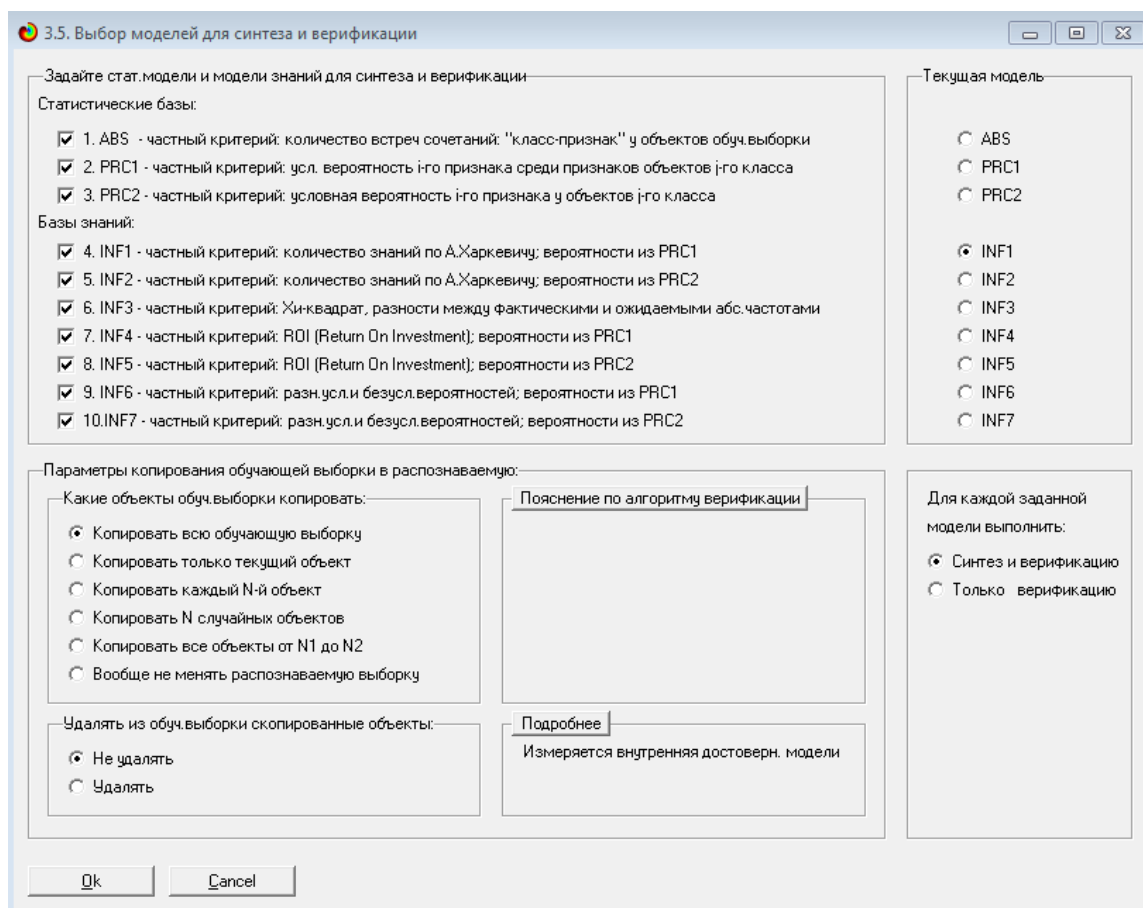


Рисунок 8 – Выбор моделей для синтеза и верификации, а также текущей модели

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 8. Стадия

процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 9.

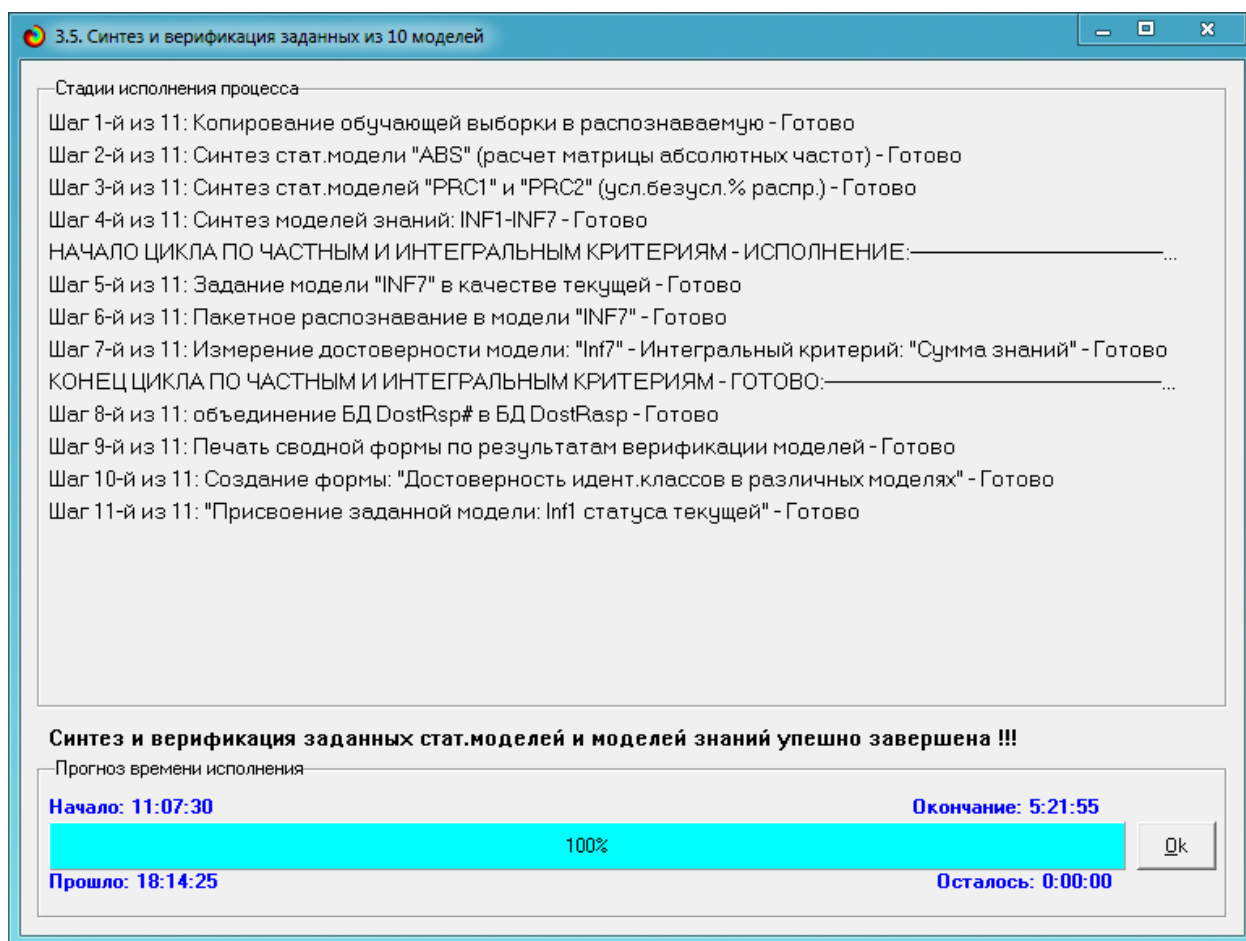


Рисунок 9 – Синтез и верификация статистических моделей и моделей знаний

Верификация (оценка достоверности моделей) проводилась на всех примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 8.

1.4. Виды моделей системы «Эйдос»

Рассмотрим решение задачи идентификации на примере модели INF1, в которой рассчитано количество информации по А. Харкевичу, которое мы получаем о принадлежности идентифицируемого объекта к каждому из классов, если знаем, что у этого объекта есть некоторый признак.

По сути, частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот в матрицы условных и безусловных процентных распределений, и матрицы знаний.

Код признака	Наименование описательной шкалы и градации	1. BABURIN 1/5 (1 0, 23 0)	2. BABURIN 2/5 (23 0, 45 0)	3. BABURIN 3/5 (45 0, 67 0)	4. BABURIN 4/5 (67 0, 89 0)	5. BABURIN 5/5 (89 0, 111 0)	6. GRUDININ 1/5 (1 0, 112 4)	7. GRUDININ 2/5 (112 4, 223 8)	8. GRUDININ 3/5 (223 8, 335 2)	9. GRUDININ 4/5 (335 2, 446 6)	10. GRUDININ 5/5 (446 6, 558 0)	11. ZHIRIN 1/5 (1 0, 82 4)
1	REGION_NAME-Краснодарский край	1856	8	3		1	1238	730	113	9	4	1
2	REGION_NAME-Краснодарский край,"20 Краснод...						1					
3	REGION_NAME-Краснодарский край,"20 Краснод...						1					
4	REGION_NAME-Краснодарский край,"20 Краснод...						1					
5	REGION_NAME-Краснодарский край,"20 Краснод...						1					
6	REGION_NAME-Краснодарский край,"20 Краснод...						1					
7	REGION_NAME-Краснодарский край,"20 Краснод...						1					
8	REGION_NAME-Краснодарский край,"20 Краснод...						1					
9	REGION_NAME-Краснодарский край,"20 Краснод...						1					
10	REGION_NAME-Краснодарский край,"20 Краснод...						1					
11	REGION_NAME-Краснодарский край,"20 Краснод...						1					
12	REGION_NAME-Краснодарский край,"20 Краснод...						1					
13	REGION_NAME-Краснодарский край,"20 Краснод...						1					
14	REGION_NAME-Краснодарский край,"20 Краснод...						1					
15	REGION_NAME-Краснодарский край,"20 Краснод...						1					
16	REGION_NAME-Краснодарский край,"20 Краснод...						1					
17	REGION_NAME-Краснодарский край,"20 Краснод...						1					
18	REGION_NAME-Краснодарский край,"20 Краснод...						1					
19	REGION_NAME-Краснодарский край,"20 Краснод...						1					
20	REGION_NAME-Краснодарский край,"20 Краснод...						1					
21	REGION_NAME-Краснодарский край,"20 Краснод...						1					
22	REGION_NAME-Краснодарский край,"20 Краснод...						1					

Рисунок 10 – Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений (фрагмент)

Код признака	Наименование описательной шкалы и градации	12. ZHIRINOVSKY 2/5 (82 4, 163 8)	13. ZHIRINOVSKY 3/5 (163 8, 245 2)	14. ZHIRINOVSKY 4/5 (245 2, 326 6)	15. ZHIRINOVSKY 5/5 (326 6, 408 0)	16. PUTIN 1/5 (1 0, 543 4)	17. PUTIN 2/5 (543 4, 1085 8)	18. PUTIN 3/5 (1085 8, 1628 2)	19. PUTIN 4/5 (1628 2, 2170 6)	20. PUTIN 5/5 (2170 6, 2713 0)	21. SOBCHAK 1/5 (1 0, 513 8)
1	REGION_NAME-Краснодарский край	-0.174	-1.522		-0.789	-0.213	0.115	0.116	0.119	0.124	0.10
2	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
3	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
4	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
5	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
6	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
7	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
8	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
9	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
10	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
11	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
12	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
13	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
14	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
15	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
16	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
17	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
18	REGION_NAME-Краснодарский край,"20 Краснод...		1.057			0.339					
19	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
20	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
21	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					
22	REGION_NAME-Краснодарский край,"20 Краснод...	0.651				0.339					

Рисунок 11 – Матрица информативности (модель INF1) в битах (фрагмент)

5. Модель: 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсолютными

д изнака	Наименование описательной шкалы и градации	1. BABURIN 1/5 (1.0, 23.0)	2. BABURIN 2/5 (23.0, 45.0)	3. BABURIN 3/5 (45.0, 67.0)	4. BABURIN 4/5 (67.0, 89.0)	5. BABURIN 5/5 (89.0, 111.0)	6. GRUDININ 1/5 (1.0, 112.4)	7. GRUDININ 2/5 (112.4, 223.8)	8. GRUDININ 3/5 (223.8, 335.2)	9. GRUDININ 4/5 (335.2, 446.6)	10. GRUDININ 5/5 (446.6, 558.0)
1	REGION_NAME-Краснодарский край	452.845	2.114	0.688		0.229	-236.063	172.479	26.748	2.063	1.051
2	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
3	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
4	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
5	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
6	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
7	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
8	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
9	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
10	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
11	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
12	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
13	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
14	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
15	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
16	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
17	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
18	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
19	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
20	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
21	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001
22	REGION_NAME-Краснодарский край,"20 Красн...	-0.611	-0.003	-0.001		0.000	0.358	-0.243	-0.038	-0.003	-0.001

Рисунок 12 – Матрица знаний (модель INF3) (фрагмент)

5. Модель: 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1

д изнака	Наименование описательной шкалы и градации	6. GRUDININ 1/5 (1.0, 112.4)	7. GRUDININ 2/5 (112.4, 223.8)	8. GRUDININ 3/5 (223.8, 335.2)	9. GRUDININ 4/5 (335.2, 446.6)	10. GRUDININ 5/5 (446.6, 558.0)	11. ZHIRINOVSKY 1/5 (1.0, 82.4)	12. ZHIRINOVSKY 2/5 (82.4, 163.8)	13. ZHIRINOVSKY 3/5 (163.8, 245.2)	14. ZHIRINOVSKY 4/5 (245.2, 326.6)	15. ZHIRINOVSKY 5/5 (326.6, 408.0)
1	REGION_NAME-Краснодарский край	-0.160	0.309	0.310	0.297	0.359	0.245	-0.339	-0.973		-0.84
2	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
3	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
4	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
5	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
6	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
7	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
8	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
9	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
10	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
11	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
12	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
13	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
14	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
15	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
16	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
17	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
18	REGION_NAME-Краснодарский край,"20 Красн...	0.558							11.329		
19	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
20	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
21	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			
22	REGION_NAME-Краснодарский край,"20 Красн...	0.558						3.702			

Рисунок 13 – Матрица знаний (модель INF4) (фрагмент)

1.5. Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями, приведены на рисунке 14.

Наименование модели и частного критерия	Интегральный критерий	F-мера Ван Ризбергена	Сумма модул... уровней сход... истинно-поло... решений (STP)	Сумма модул... уровней сход... истинно-отриц... решений (STN)	Сумма модул... уровней сход... ложно-полож... решений (SFP)	Сумма модул... уровней сход... ложно-отрицат... решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко	Средний модуль... уровней сход... истинно-полож... решений
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс. частот с обр...	0.318	11587.181	0.848	27813.980		0.294	1.000	0.455	0.648
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс. частот по признак...	0.318	10437.512		5326.235		0.662	1.000	0.797	0.584
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл. частот с о...	0.318	11587.192	0.848	27814.008		0.294	1.000	0.455	0.648
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл. частот по приз...	0.318	7582.075		19993.195		0.275	1.000	0.431	0.424
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл. частот с о...	0.318	11587.178	0.848	27813.982		0.294	1.000	0.455	0.648
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл. частот по приз...	0.318	6942.169		18534.512		0.272	1.000	0.428	0.389
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.343	433.400	3051.942	1396.510	302.170	0.237	0.589	0.338	0.044
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.454	1580.110	5265.247	2223.801	84.115	0.415	0.949	0.578	0.101
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс зна...	0.347	420.833	2961.259	1389.356	301.289	0.232	0.583	0.332	0.043
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.454	1546.163	5019.423	2233.899	81.353	0.409	0.950	0.572	0.098
6. INF3 - частный критерий: Минимум разности между фактич...	Семантический резонанс зна...	0.589	7132.546	19821.355	4524.093	763.048	0.612	0.903	0.730	0.474
6. INF3 - частный критерий: Минимум разности между фактич...	Сумма знаний	0.589	5810.813	6593.017	1200.125	417.921	0.829	0.933	0.878	0.387
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.543	581.830	2081.508	127.015	436.374	0.821	0.571	0.674	0.059
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.419	36.539	12.290	47.471	0.244	0.435	0.993	0.605	0.002
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.538	567.975	2074.491	128.444	441.486	0.816	0.563	0.666	0.059
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.421	34.561	10.911	45.319	0.223	0.433	0.994	0.603	0.002
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	0.400	6540.581	7982.957	7119.345	666.111	0.479	0.908	0.627	0.435
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	0.398	1856.234	3757.120	4320.674	110.884	0.301	0.944	0.456	0.123
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	0.407	6320.363	7446.010	6983.602	618.861	0.475	0.911	0.624	0.420
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	0.405	1636.348	3206.200	3966.600	90.884	0.292	0.947	0.446	0.109

Рисунок 14 – Оценки достоверности моделей

Наиболее достоверной в данном приложении оказалась модель INF3 при интегральном критерии «Сумма знаний». При этом точность модели составляет 0,878. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием, описанным на рисунке 15.

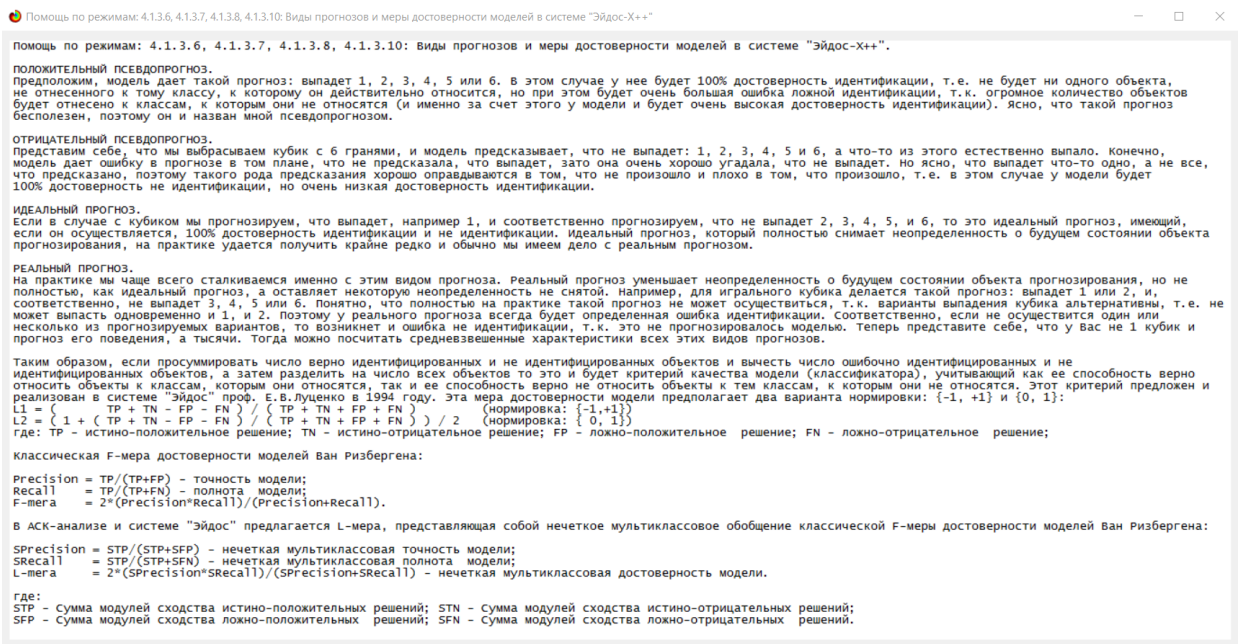


Рисунок 15 – Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

Стоит отметить, 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 16 приведены частные распределения уровней сходства и различия для верно- и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF3.

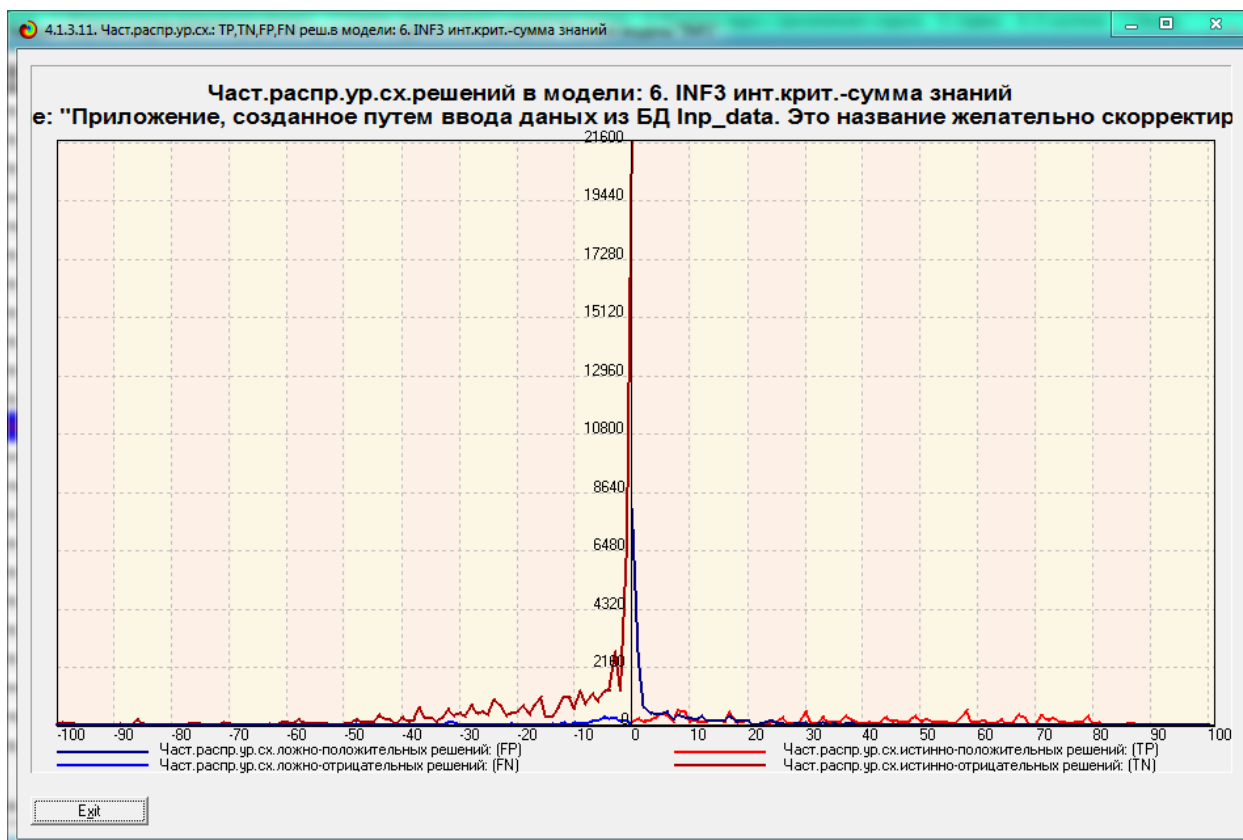


Рисунок 16 – Частное распределение ложно-положительных, ложно-отрицательных, истинно-положительных и истинно-отрицательных решений объекта моделирования в модели INF3

Из рисунка 16 видно, что:

- наиболее достоверная модель INF3 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 14);

- 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. Для очень больших значений уровней сходства-различия (более 95%) также различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями практически отсутствует.

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) и проведем пакетное распознавание в режиме 4.1.2, как показано на рисунке 18.

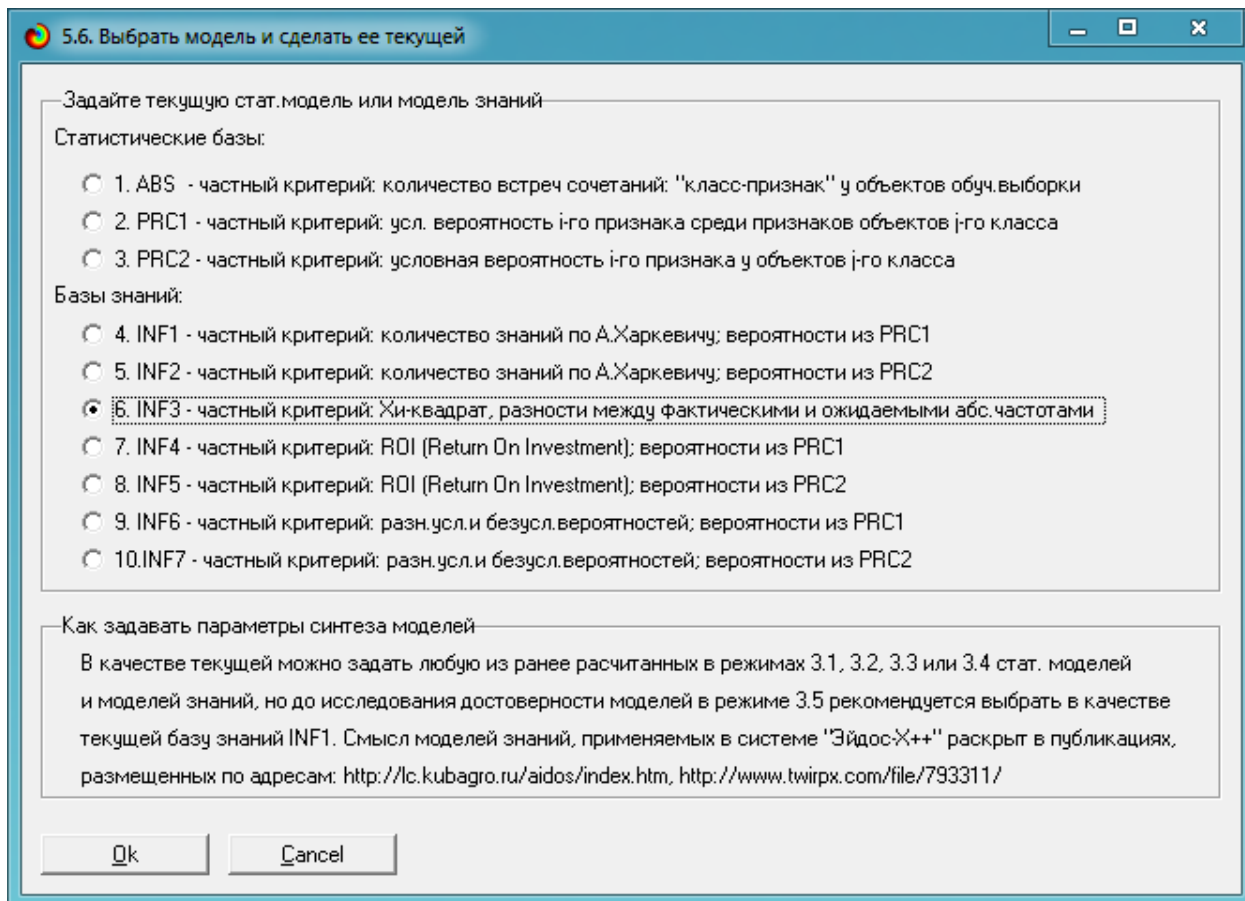


Рисунок 17 – Экранная форма режима задания модели в качестве текущей

Следующим шагом необходимо произвести пакетное распознавание в текущей модели. Если распознавание производится впервые, то вы увидите сообщение, уведомляющее об этом.

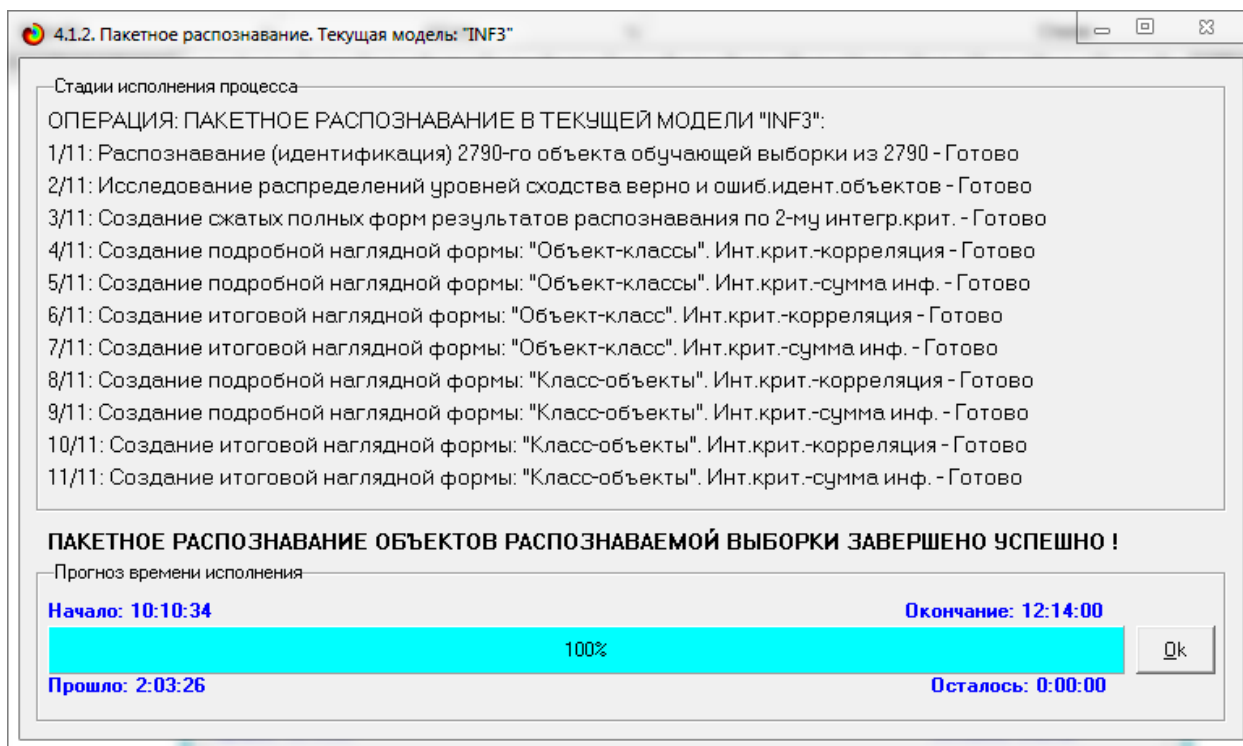


Рисунок 18 – Экранная форма режима пакетного распознавания в текущей модели

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.

Для наглядного примера, кратко рассмотрим некоторые из них.

На рисунках 20 и 21 приведены примеры прогнозов высокой и низкой достоверности частоты и классов бейсбольных команд в наиболее достоверной модели INF3 на основе описательных шкал.

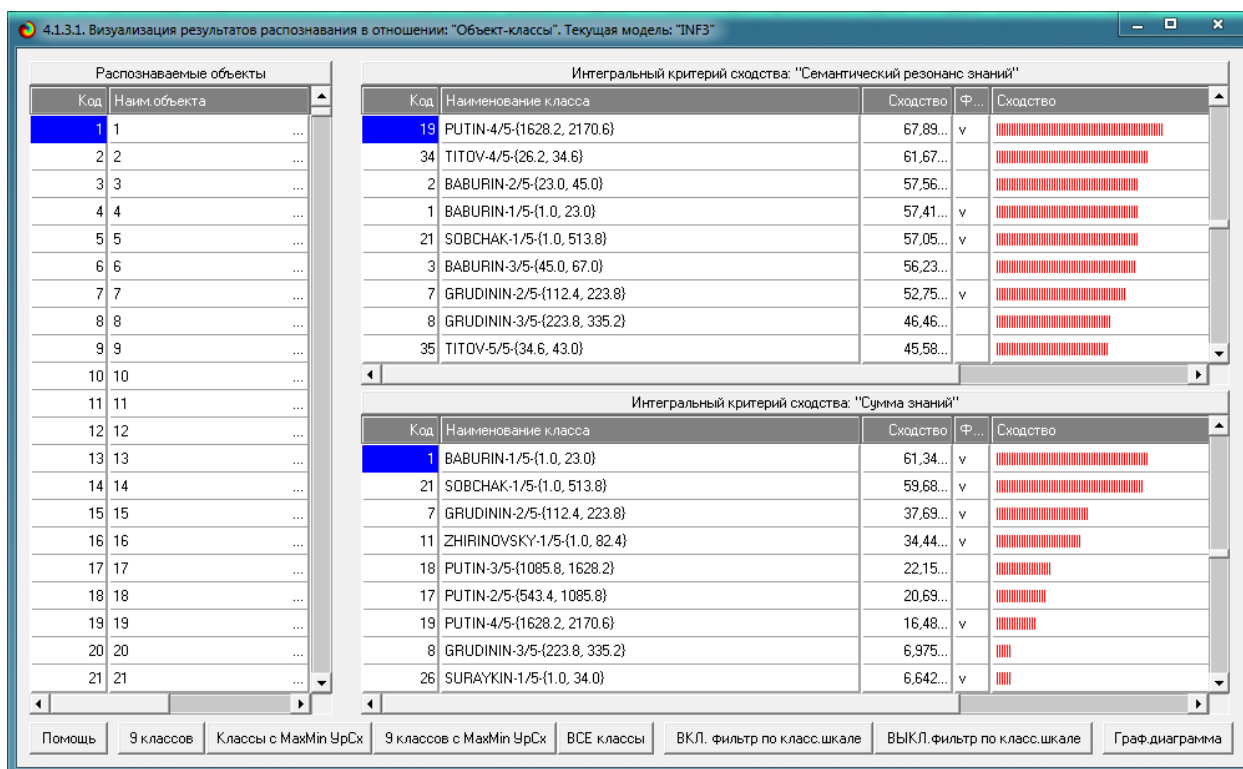


Рисунок 19 – Пример идентификации классов в модели INF3

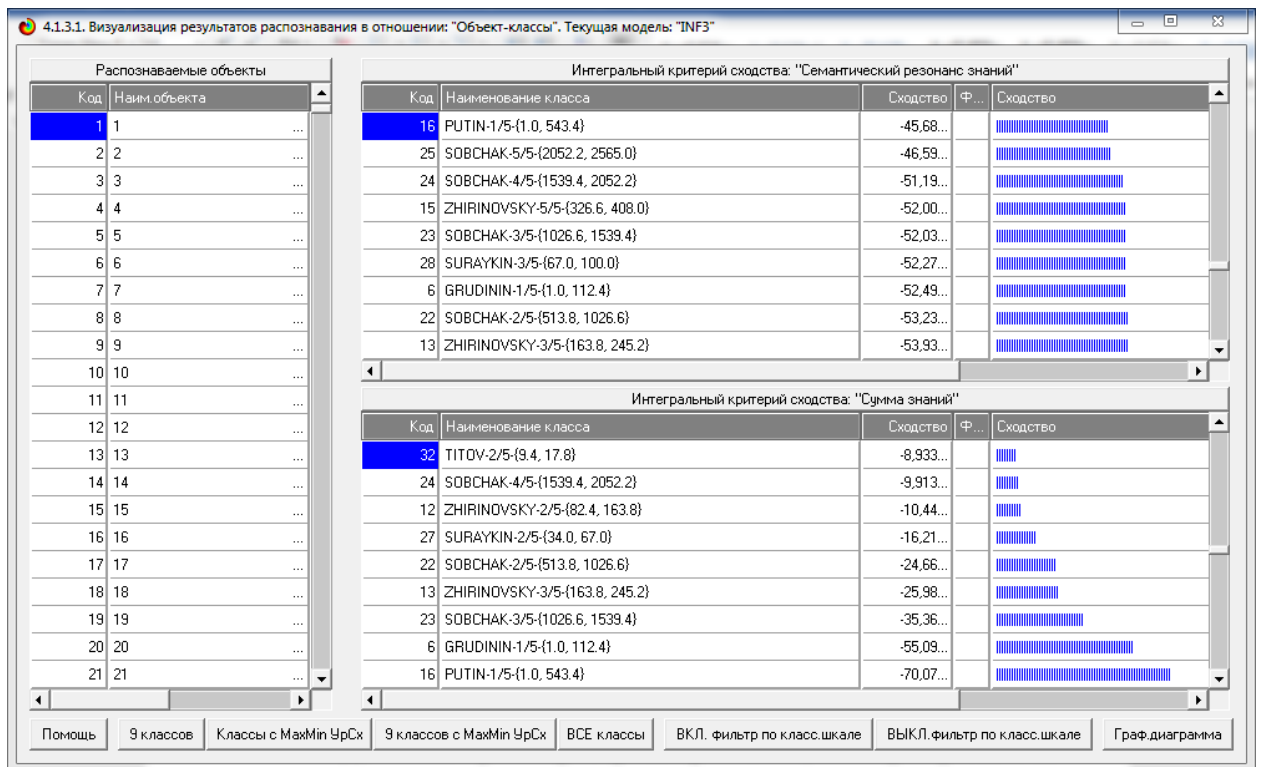


Рисунок 20 – Пример идентификации классов в модели INF3

2.2 Когнитивные функции

Далее стоит рассмотреть режим работы 4.5, в котором реализована возможность визуализации когнитивных функций для любых моделей и любых сочетаний классификационных и описательных шкал. Начало работы в этом режиме начинается с понятия когнитивных функций, а также предоставления возможности перехода на статьи по данной проблематике и скачивания публикаций, как это продемонстрировано на рисунке 21.

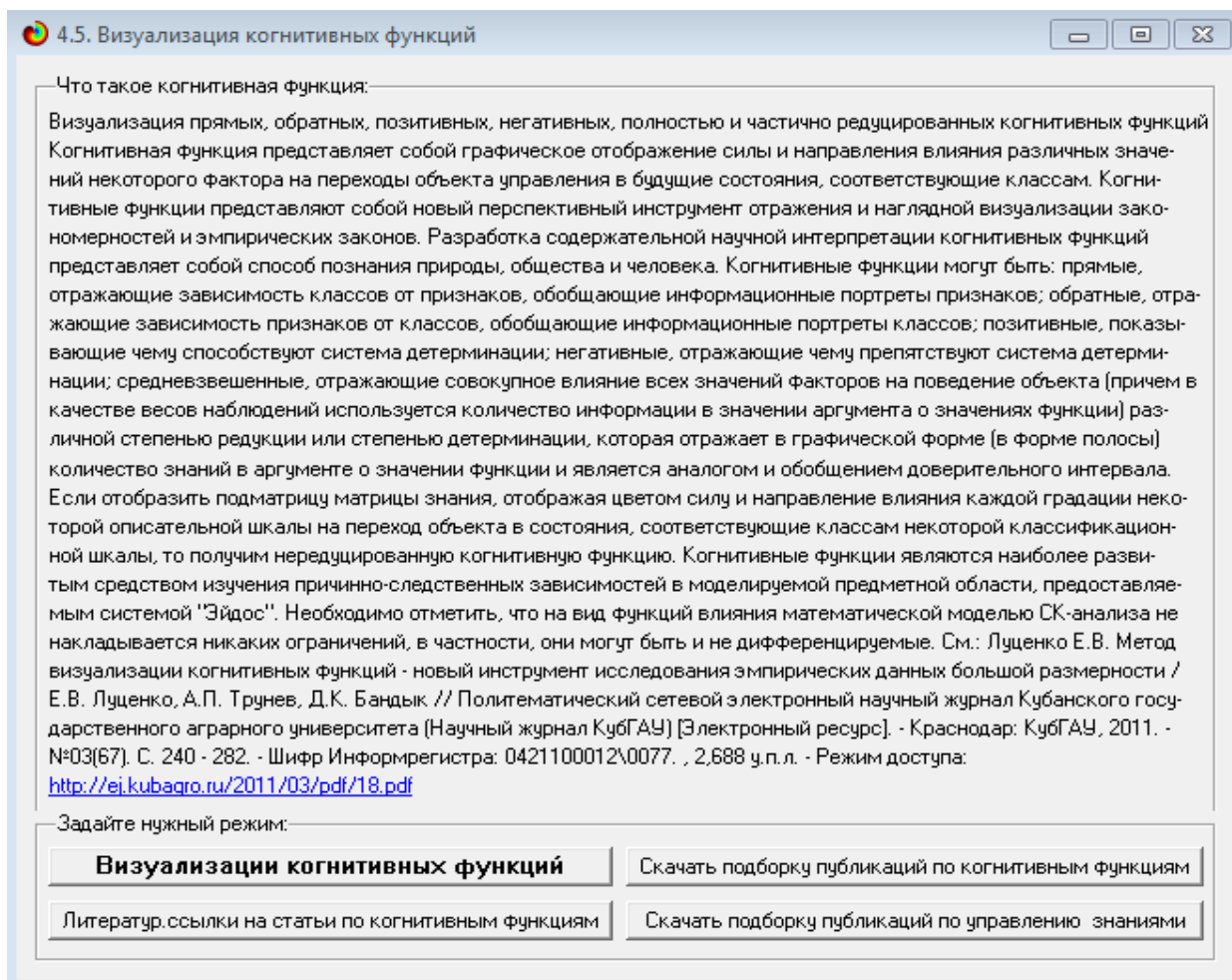


Рисунок 21– Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»

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 множество работ, поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 22 приведены визуализации всех когнитивных функций данного приложения для модели INF3.

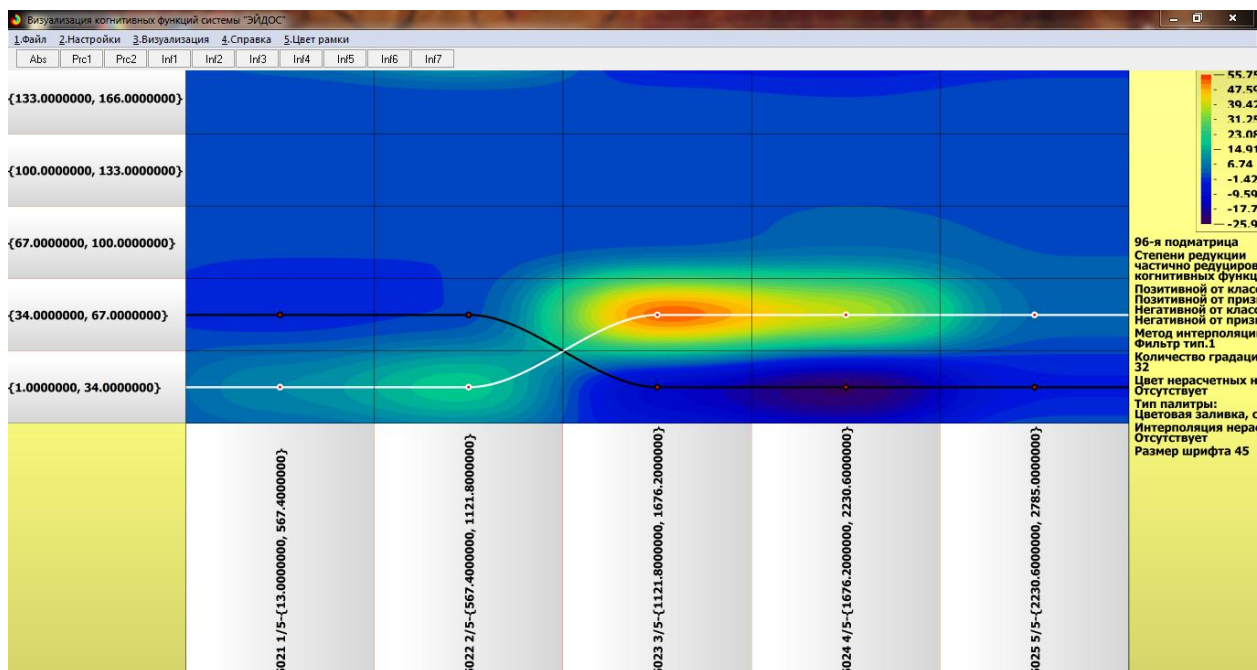


Рисунок 22 – Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF4

2.3. SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT- анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных.

Подобная технология разработана давно, 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 23).

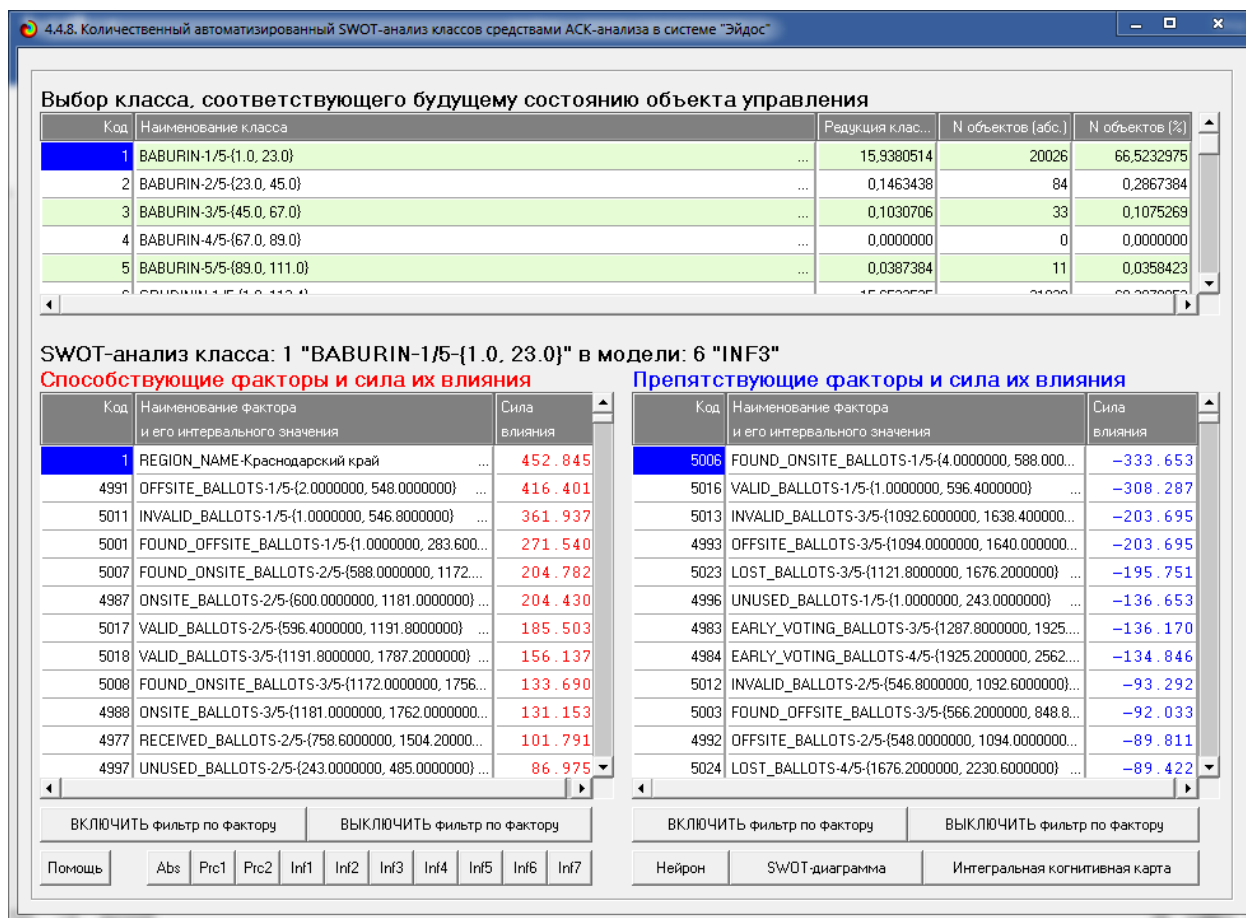


Рисунок 23 – Пример SWOT-Матрицы в модели INF3

Следующим шагом является построение SWOT-диаграммы, для наглядной демонстрации работы, продемонстрированной на рисунке 24.

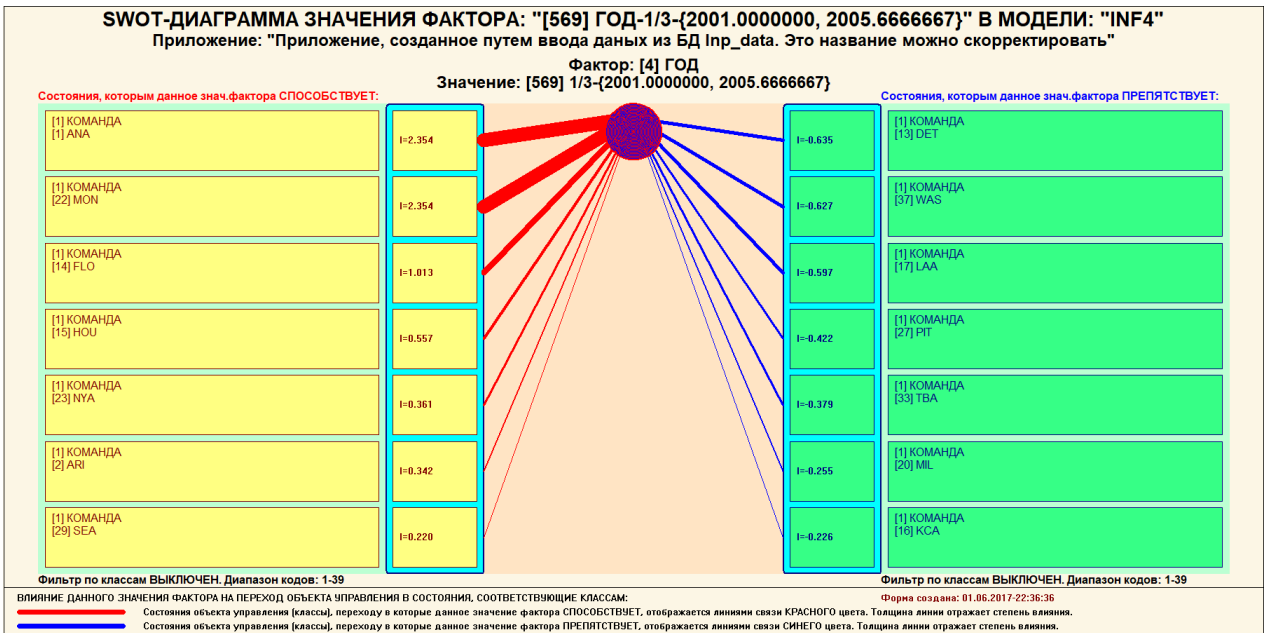


Рисунок 24 – SWOT-матрица, построенная для 2001 года

2.4. Нелокальные нейронные сети

Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует не локальную нейронную сеть. Рассмотрим пример на рисунке 25, отображающий режим 4.4.10.



Рисунок 25 – Нейрон для президентских выборов

Благодаря данному нейрону видно, как различные кандидаты влияют на состояние выборов: какие оказывают активирующее влияние, а какие тормозящее.

2.5 Кластерный и конструктивный анализ

В режиме 4.2.2.2, после расчета матриц сходства, кластеров и конструкторов, строим 2D сеть классов в выбранной модели знаний, для наглядного представления сети классов. На рисунке 26 наглядно видно, что наибольшее сходство наблюдается в классе «Baburin 1/5».

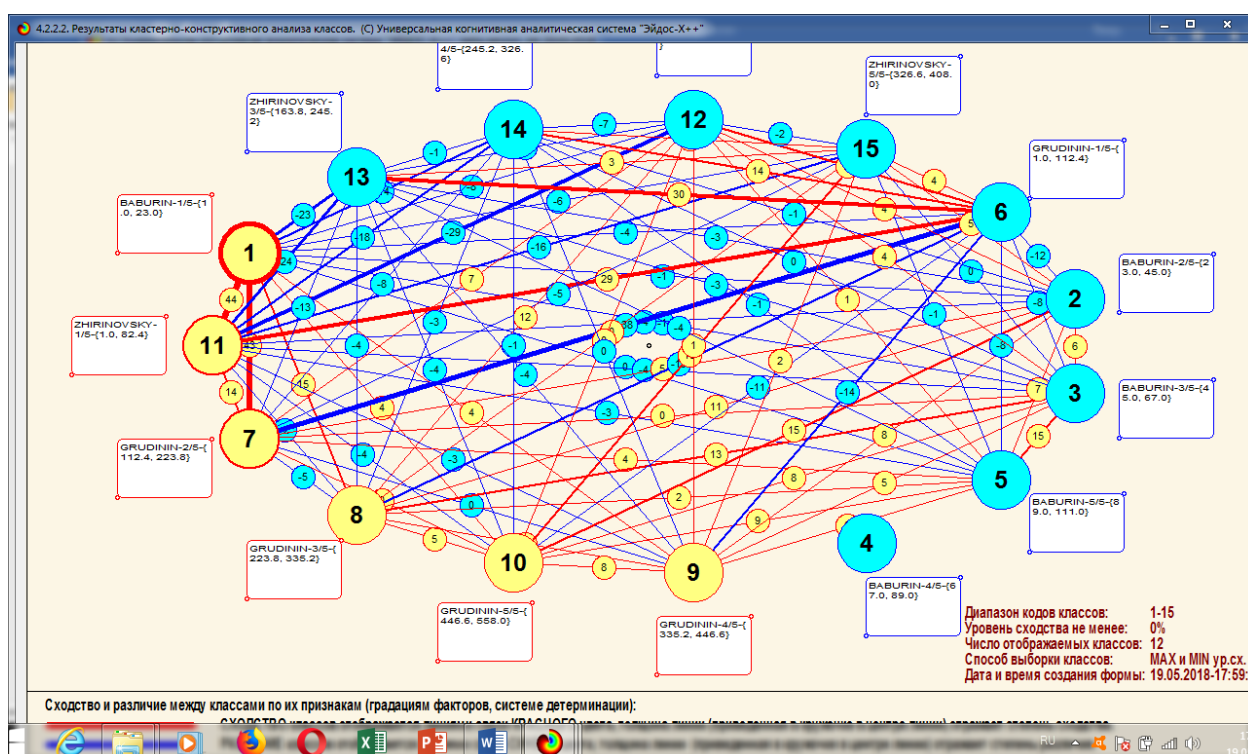


Рисунок 26 – Семантическая 2D сеть классов

ЗАКЛЮЧЕНИЕ

В связи с существованием множества систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную базу репозитория «Kaggle».

В данной лабораторной приводится развернутый пример использования базы данных репозитория «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 0,878? что заметно выше, чем достоверность экспертных оценок, которая считается равной около 60%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием.

На основе базы данных «Kaggle», рассмотренной в данной лабораторной работе, построены модели прогнозирования с помощью АСК-анализа и реализующей его системы «Эйдос».

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.

6. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://ic.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.