МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования

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

Факультет прикладной информатики

РЕЦЕНЗИЯ
на курсовой проект (работу)

Студента (ки) Колкнев Сергей Александрович
курса 3 очной (заочной) формы обучения
направления подготовки информационные системы и технологии
направленность (профиль) Интеллектуальные информационные системы.
Наименование темы «Разработка и исследование системно-когнитивной модели классификации преступлений на основе данных репозитория UCI за 1975-2015 годы по видам насильственных преступлений на душу населения».
Рецензент: Луценко Е.В., 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	5
3.	Уровень и корректность использования в работе различных методов исследований	4
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	5
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	4
8.	Ответы на вопросы при защите	5

Достоинства работы :
Применение современных математических моделей и программного инструментария

Недостатки работы :
Мелкие недочеты в оформлении

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 преступлений на основе данных репозитория UCI за 1975-2015 годы по видам насильственных преступлений на душу населения

выполнил студент группы ИТ1401 **Колкнев Сергей Александрович**

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

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

Колкнев Сергей Александрович
(Ф.И.О.)

Тема проекта: **Разработка и исследование системно-когнитивной модели классификации преступлений на основе данных репозитория UCI за 1975-2015 годы по видам насильственных преступлений на душу 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

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ 8

1.1. Описание решения 8

1.2 Подготовка данных в промежуточные файлы MS Excel 8

1.3 Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос". 10

1.3. Синтез и верификация статистических и интеллектуальных моделей 15

1.4. Виды моделей системы «Эйдос» 16

1.5. Результаты верификации моделей 18

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 31

2.4. Нелокальные нейроны и нейронная сеть 34

2.5. Кластерный и конструктивный анализ классов 36

ЗАКЛЮЧЕНИЕ 37

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 39

ВВЕДЕНИЕ

На сегодняшний день создание систем искусственного интеллекта – это перспективный шаг к расширению возможностей компьютерных наук и автоматизации разумного поведения, опирающийся на теоретические и прикладные принципы. Однако возникает необходимость оценки качества математических моделей этих систем.

В данной курсовой работе рассмотрено решение задачи прогнозирования преступлений на основе данных 1975-2015 годов по видам насильственных преступлений на душу населения.

Целью работы является разработка системно-когнитивной модели, основывающейся на следующих признаках: отчетный год, население, насильственные преступления.

Задачами курсовой работы являются:

1) систематизация, закрепление и расширение теоретических и практических знаний по дисциплине "Интеллектуальные информационные системы и технологии";

2) изучение интеллектуальной информационной системы "Эйдос";

3) решение поставленной цели с помощью интеллектуальной информационной системы "Эйдос".

Объектом исследования является база данных "MUP Crime".

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 данной работе использована база данных " MUP Crime", предоставленная МУП «Криминал». База данных дополнена значениями и приводится полностью в приложении.

На момент написания курсовой работы этот банк включает в себя более 2 тысяч исходных данных по различным задачам, однако рассмотрена будет лишь часть усредненных данных. Решение поставленной задачи проведем в четыре этапа:

1. Подготовка 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 Подготовка данных в промежуточные файлы MS Excel

Из электронного ресурса баз данных рассматриваемого предприятия возьмем базу данных сотрудников – «MUP Crime.xls», которую оставим без изменений. Данная база была взята с сайта https://www.kaggle.com/marshallproject/crime-rates.

Общее описание задачи:

1. Number
2. Report_year
3. Agency_code
4. Agency_jurisdiction
5. Population
6. Violent_crimes
7. Homicides

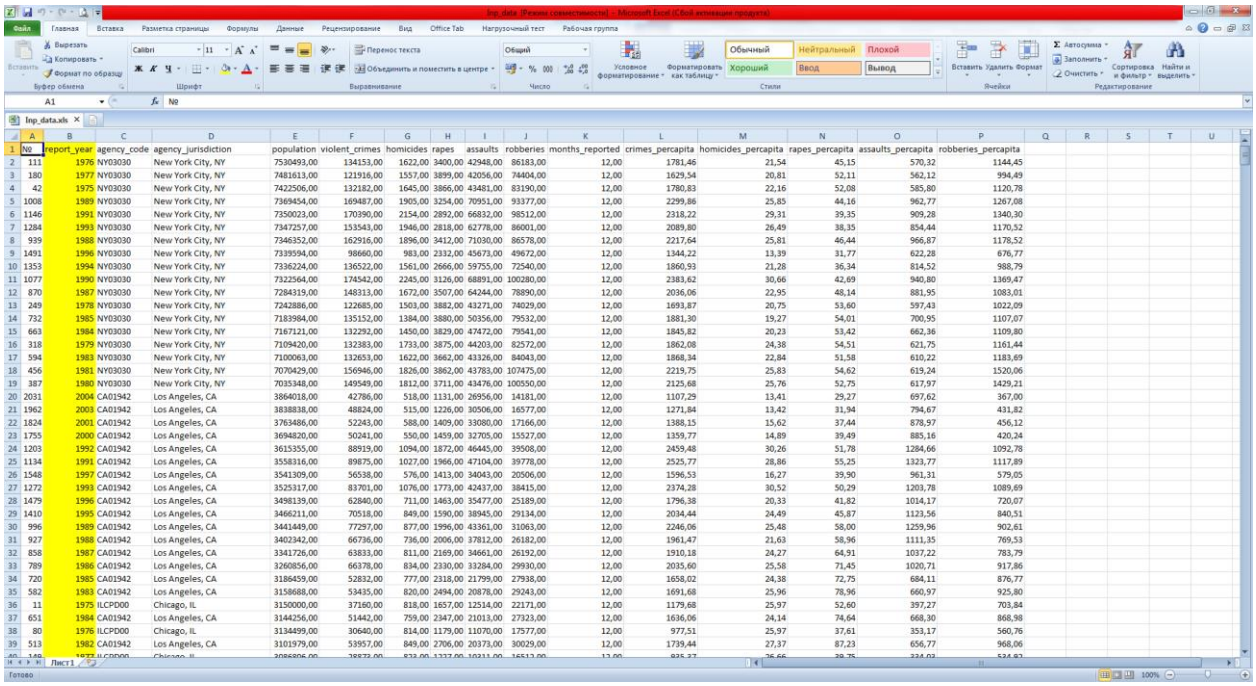
8. Rapes
9. Assaults
10. Robberies
11. Months_reported
12. Crimes_percapita
13. Homicides_percapita
14. Rapes_percapita
15. Assaults_percapita
16. Robberies_percapita

Столбцы 3-16 – это описательные шкалы.

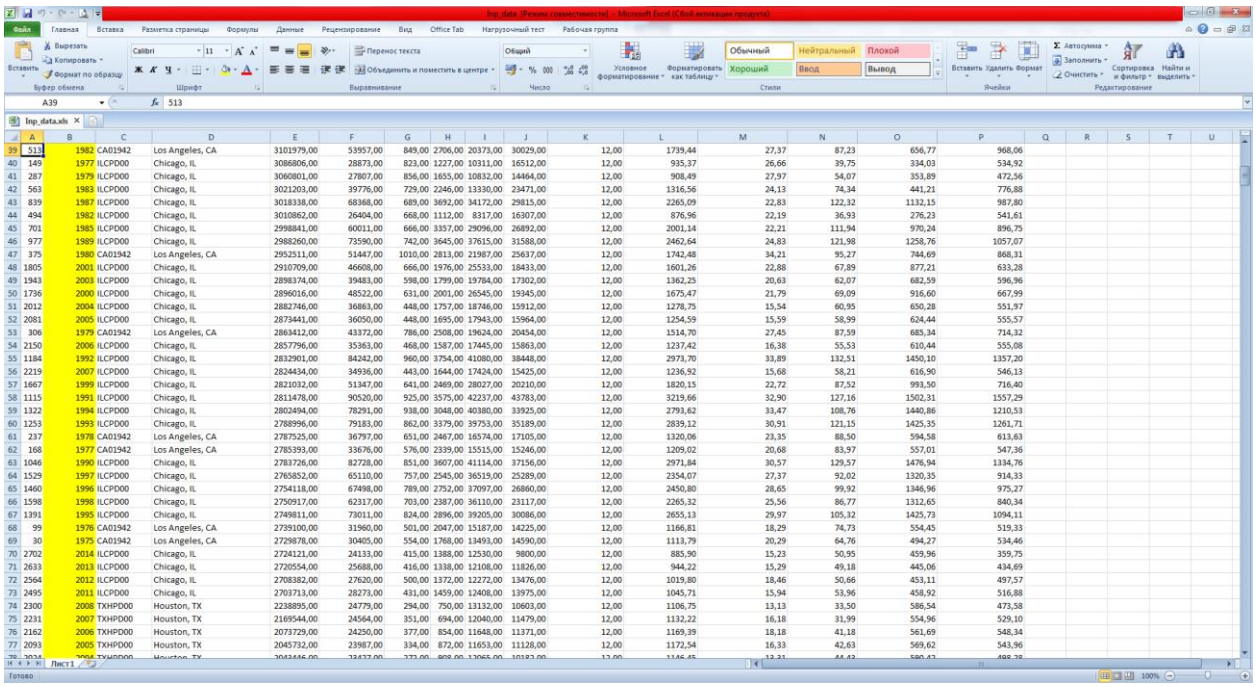
Столбцы 2-2 являются классификационными шкалами.

Обучающая выборка:

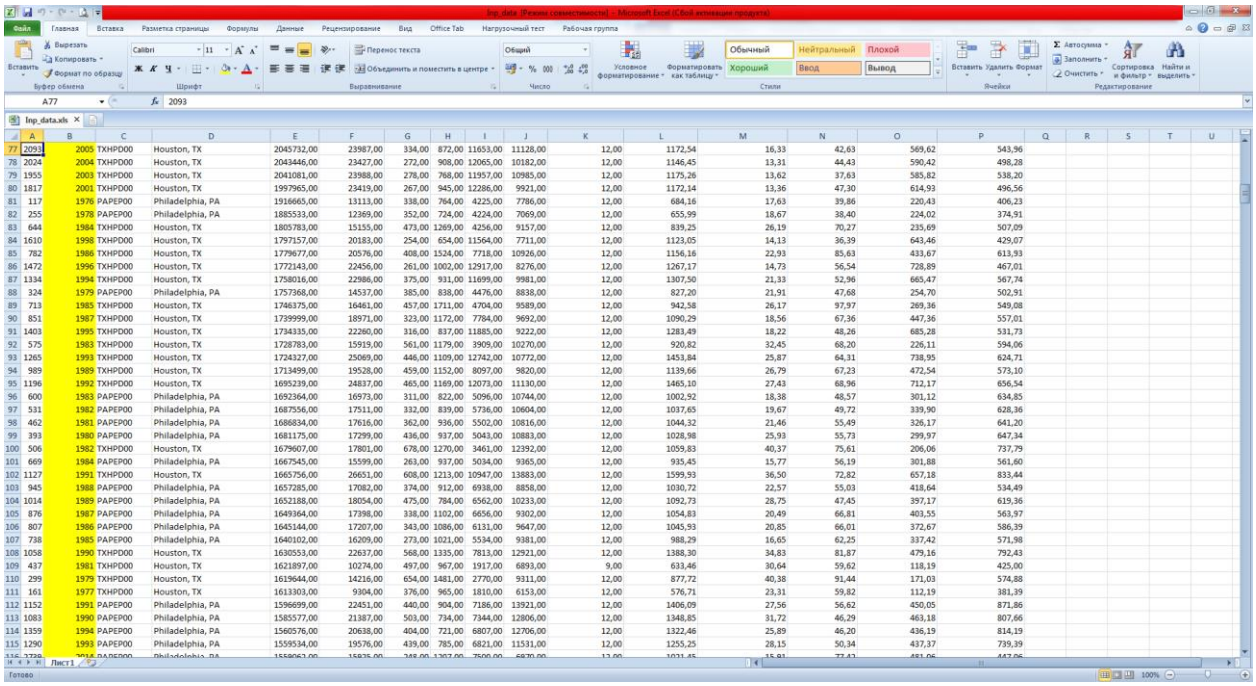
Таблица 1 – MUP Crime.xls



№	report_year	agency_code	agency_jurisdiction	population	violent_crimes	homicides	rapes	assaults	robberies	months_reported	crimes_percapita	homicides_percapita	rapes_percapita	assaults_percapita	robberies_percapita
111	1976	NY03030	New York City, NY	7530493,00	134153,00	1622,00	3400,00	42948,00	86183,00	12,00	1781,46	21,54	45,15	570,32	1144,45
180	1977	NY03030	New York City, NY	7481613,00	121916,00	1557,00	3899,00	42056,00	74404,00	12,00	1629,54	20,81	52,11	562,12	994,49
42	1975	NY03030	New York City, NY	7422506,00	132182,00	1645,00	3866,00	43481,00	83190,00	12,00	1780,83	22,16	52,08	585,80	1120,78
1008	1989	NY03030	New York City, NY	7369454,00	169487,00	1905,00	3254,00	70951,00	93377,00	12,00	2299,86	25,85	44,16	962,77	1267,08
1146	1991	NY03030	New York City, NY	7350023,00	170390,00	2154,00	2892,00	66832,00	98512,00	12,00	2318,22	29,31	39,35	909,28	1340,30
1284	1993	NY03030	New York City, NY	7347257,00	153543,00	1946,00	2818,00	62778,00	86001,00	12,00	2089,80	26,49	38,35	854,44	1170,52
939	1988	NY03030	New York City, NY	7346352,00	162916,00	1896,00	3412,00	71030,00	86578,00	12,00	2217,64	25,81	46,44	966,87	1178,52
1491	1996	NY03030	New York City, NY	7339594,00	98660,00	983,00	2332,00	45673,00	49672,00	12,00	1344,22	13,39	31,77	622,28	676,77
1353	1994	NY03030	New York City, NY	7336224,00	136522,00	1561,00	2666,00	59755,00	72540,00	12,00	1860,93	21,28	36,34	814,52	988,79
1077	1990	NY03030	New York City, NY	7322564,00	174542,00	2245,00	3126,00	68891,00	100280,00	12,00	2383,62	30,66	42,69	940,80	1369,47
870	1987	NY03030	New York City, NY	7284319,00	148313,00	1672,00	3507,00	64244,00	78890,00	12,00	2036,06	22,95	48,14	881,95	1083,01
249	1978	NY03030	New York City, NY	7242886,00	122685,00	1503,00	3882,00	43271,00	74029,00	12,00	1693,87	20,75	53,60	597,43	1022,09
732	1985	NY03030	New York City, NY	7183984,00	135152,00	1384,00	3880,00	50356,00	79532,00	12,00	1881,30	19,27	54,01	700,95	1107,07
663	1984	NY03030	New York City, NY	7167121,00	132292,00	1450,00	3829,00	47472,00	79541,00	12,00	1845,82	20,23	53,42	662,36	1109,80
318	1979	NY03030	New York City, NY	7109420,00	132383,00	1733,00	3875,00	44203,00	82572,00	12,00	1862,08	24,38	54,51	621,75	1161,44
594	1983	NY03030	New York City, NY	7100063,00	132653,00	1622,00	3662,00	43326,00	84043,00	12,00	1868,34	22,84	51,58	610,22	1183,69
456	1981	NY03030	New York City, NY	7070429,00	156946,00	1826,00	3862,00	43783,00	107475,00	12,00	2219,75	25,83	54,62	619,24	1520,06
387	1980	NY03030	New York City, NY	7035348,00	149549,00	1812,00	3711,00	43476,00	100550,00	12,00	2125,68	25,76	52,75	617,97	1429,21
2031	2004	CA01942	Los Angeles, CA	3864018,00	42786,00	518,00	1131,00	26956,00	14181,00	12,00	1107,29	13,41	29,27	697,62	367,00
1962	2003	CA01942	Los Angeles, CA	3838838,00	48824,00	515,00	1226,00	30506,00	16577,00	12,00	1271,84	13,42	31,94	794,67	431,82
1824	2001	CA01942	Los Angeles, CA	3763486,00	52243,00	588,00	1409,00	33080,00	17166,00	12,00	1388,15	15,62	37,44	878,97	456,12
1755	2000	CA01942	Los Angeles, CA	3694820,00	50241,00	550,00	1459,00	32705,00	15527,00	12,00	1359,77	14,89	39,49	885,16	420,24
1203	1992	CA01942	Los Angeles, CA	3615355,00	88919,00	1094,00	1872,00	46445,00	39508,00	12,00	2459,48	30,26	51,78	1284,66	1092,78
1134	1991	CA01942	Los Angeles, CA	3558316,00	89875,00	1027,00	1966,00	47104,00	39778,00	12,00	2525,77	28,86	55,25	1323,77	1117,89
1548	1997	CA01942	Los Angeles, CA	3541309,00	56538,00	576,00	1413,00	34043,00	20506,00	12,00	1596,53	16,27	39,90	961,31	579,05
1272	1993	CA01942	Los Angeles, CA	3525317,00	83701,00	1076,00	1773,00	42437,00	38415,00	12,00	2374,28	30,52	50,29	1203,78	1089,69
1479	1996	CA01942	Los Angeles, CA	3498139,00	62840,00	711,00	1463,00	35477,00	25189,00	12,00	1796,38	20,33	41,82	1014,17	720,07
1410	1995	CA01942	Los Angeles, CA	3466211,00	70518,00	849,00	1590,00	38945,00	29134,00	12,00	2034,44	24,49	45,87	1123,56	840,51
996	1989	CA01942	Los Angeles, CA	3441449,00	77297,00	877,00	1996,00	43361,00	31063,00	12,00	2246,06	25,48	58,00	1259,96	902,61
927	1988	CA01942	Los Angeles, CA	3402342,00	66736,00	736,00	2006,00	37812,00	26182,00	12,00	1961,47	21,63	58,96	1111,35	769,53
858	1987	CA01942	Los Angeles, CA	3341726,00	63833,00	811,00	2169,00	34661,00	26192,00	12,00	1910,18	24,27	64,91	1037,22	783,79
789	1986	CA01942	Los Angeles, CA	3260856,00	66378,00	834,00	2330,00	33284,00	29930,00	12,00	2035,60	25,58	71,45	1020,71	917,86
720	1985	CA01942	Los Angeles, CA	3186459,00	52832,00	777,00	2318,00	21799,00	27938,00	12,00	1658,02	24,38	72,75	684,11	876,77
582	1983	CA01942	Los Angeles, CA	3158688,00	53435,00	820,00	2494,00	20878,00	29243,00	12,00	1691,68	25,96	78,96	660,97	925,80
11	1975	ILCPD00	Chicago, IL	3150000,00	37160,00	818,00	1657,00	12514,00	22171,00	12,00	1179,68	25,97	52,60	397,27	703,84
651	1984	CA01942	Los Angeles, CA	3144256,00	51442,00	759,00	2347,00	21013,00	27323,00	12,00	1636,06	24,14	74,64	668,30	868,98
80	1976	ILCPD00	Chicago, IL	3134499,00	30640,00	814,00	1179,00	11070,00	17577,00	12,00	977,51	25,97	37,61	353,17	560,76
513	1982	CA01942	Los Angeles, CA	3101979,00	53957,00	849,00	2706,00	20373,00	30029,00	12,00	1739,44	27,37	87,23	656,77	968,06



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
39	513	1982	CA01942	Los Angeles, CA	3101979,00	53957,00	849,00	2706,00	20373,00	30029,00	12,00	1739,44	27,37	87,23	656,77	968,06
40	149	1977	ILCPD00	Chicago, IL	3086806,00	28873,00	823,00	1227,00	10311,00	16512,00	12,00	935,37	26,66	39,75	334,03	534,92
41	287	1979	ILCPD00	Chicago, IL	3060801,00	27807,00	856,00	1655,00	10832,00	14464,00	12,00	908,49	27,97	54,07	353,89	472,56
42	563	1983	ILCPD00	Chicago, IL	3021203,00	39776,00	729,00	2246,00	13330,00	23471,00	12,00	1316,56	24,13	74,34	441,21	776,88
43	839	1987	ILCPD00	Chicago, IL	3018338,00	68368,00	689,00	3692,00	34172,00	29815,00	12,00	2265,09	22,83	122,32	1132,15	987,80
44	494	1982	ILCPD00	Chicago, IL	3010862,00	26404,00	668,00	1112,00	8317,00	16307,00	12,00	876,96	22,19	36,93	276,23	541,61
45	701	1985	ILCPD00	Chicago, IL	2998841,00	60011,00	666,00	3357,00	29096,00	26892,00	12,00	2001,14	22,21	111,94	970,24	896,75
46	977	1989	ILCPD00	Chicago, IL	2988260,00	73590,00	742,00	3645,00	37615,00	31588,00	12,00	2462,64	24,83	121,98	1258,76	1057,07
47	375	1980	CA01942	Los Angeles, CA	2952511,00	51447,00	1010,00	2813,00	21987,00	25637,00	12,00	1742,48	34,21	95,27	744,69	868,31
48	1805	2001	ILCPD00	Chicago, IL	2910709,00	46608,00	666,00	1976,00	25533,00	18433,00	12,00	1601,26	22,88	67,89	877,21	633,28
49	1943	2003	ILCPD00	Chicago, IL	2898374,00	39483,00	598,00	1799,00	19784,00	17302,00	12,00	1362,25	20,63	62,07	682,59	596,96
50	1736	2000	ILCPD00	Chicago, IL	2896016,00	48522,00	631,00	2001,00	26545,00	19345,00	12,00	1675,47	21,79	69,09	916,60	667,99
51	2012	2004	ILCPD00	Chicago, IL	2882746,00	36863,00	448,00	1757,00	18746,00	15912,00	12,00	1278,75	15,54	60,95	650,28	551,97
52	2081	2005	ILCPD00	Chicago, IL	2873441,00	36050,00	448,00	1695,00	17943,00	15964,00	12,00	1254,59	15,59	58,99	624,44	555,57
53	306	1979	CA01942	Los Angeles, CA	2863412,00	43372,00	786,00	2508,00	19624,00	20454,00	12,00	1514,70	27,45	87,59	685,34	714,32
54	2150	2006	ILCPD00	Chicago, IL	2857796,00	35363,00	468,00	1587,00	17445,00	15863,00	12,00	1237,42	16,38	55,53	610,44	555,08
55	1184	1992	ILCPD00	Chicago, IL	2832901,00	84242,00	960,00	3754,00	41080,00	38448,00	12,00	2973,70	33,89	132,51	1450,10	1357,20
56	2219	2007	ILCPD00	Chicago, IL	2824434,00	34936,00	443,00	1644,00	17424,00	15425,00	12,00	1236,92	15,68	58,21	616,90	546,13
57	1667	1999	ILCPD00	Chicago, IL	2821032,00	51347,00	641,00	2469,00	28027,00	20210,00	12,00	1820,15	22,72	87,52	993,50	716,40
58	1115	1991	ILCPD00	Chicago, IL	2811478,00	90520,00	925,00	3575,00	42237,00	43783,00	12,00	3219,66	32,90	127,16	1502,31	1557,29
59	1322	1994	ILCPD00	Chicago, IL	2802494,00	78291,00	938,00	3048,00	40380,00	33925,00	12,00	2793,62	33,47	108,76	1440,86	1210,53
60	1253	1993	ILCPD00	Chicago, IL	2788996,00	79183,00	862,00	3379,00	39753,00	35189,00	12,00	2839,12	30,91	121,15	1425,35	1261,71
61	237	1978	CA01942	Los Angeles, CA	2787525,00	36797,00	651,00	2467,00	16574,00	17105,00	12,00	1320,06	23,35	88,50	594,58	613,63
62	168	1977	CA01942	Los Angeles, CA	2785393,00	33676,00	576,00	2339,00	15515,00	15246,00	12,00	1209,02	20,68	83,97	557,01	547,36
63	1046	1990	ILCPD00	Chicago, IL	2783726,00	82728,00	851,00	3607,00	41114,00	37156,00	12,00	2971,84	30,57	129,57	1476,94	1334,76
64	1529	1997	ILCPD00	Chicago, IL	2765852,00	65110,00	757,00	2545,00	36519,00	25289,00	12,00	2354,07	27,37	92,02	1320,35	914,33
65	1460	1996	ILCPD00	Chicago, IL	2754118,00	67498,00	789,00	2752,00	37097,00	26860,00	12,00	2450,80	28,65	99,92	1346,96	975,27
66	1598	1998	ILCPD00	Chicago, IL	2750917,00	62317,00	703,00	2387,00	36110,00	23117,00	12,00	2265,32	25,56	86,77	1312,65	840,34
67	1391	1995	ILCPD00	Chicago, IL	2749811,00	73011,00	824,00	2896,00	39205,00	30086,00	12,00	2655,13	29,97	105,32	1425,73	1094,11
68	99	1976	CA01942	Los Angeles, CA	2739100,00	31960,00	501,00	2047,00	15187,00	14225,00	12,00	1166,81	18,29	74,73	554,45	519,33
69	30	1975	CA01942	Los Angeles, CA	2729878,00	30405,00	554,00	1768,00	13493,00	14590,00	12,00	1113,79	20,29	64,76	494,27	534,46
70	2702	2014	ILCPD00	Chicago, IL	2724121,00	24133,00	415,00	1388,00	12530,00	9800,00	12,00	885,90	15,23	50,95	459,96	359,75
71	2633	2013	ILCPD00	Chicago, IL	2720554,00	25688,00	416,00	1338,00	12108,00	11826,00	12,00	944,22	15,29	49,18	445,06	434,69
72	2564	2012	ILCPD00	Chicago, IL	2708382,00	27620,00	500,00	1372,00	12272,00	13476,00	12,00	1019,80	18,46	50,66	453,11	497,57
73	2495	2011	ILCPD00	Chicago, IL	2703713,00	28273,00	431,00	1459,00	12408,00	13975,00	12,00	1045,71	15,94	53,96	458,92	516,88
74	2300	2008	TXHPD00	Houston, TX	2238895,00	24779,00	294,00	750,00	13132,00	10603,00	12,00	1106,75	13,13	33,50	586,54	473,58
75	2231	2007	TXHPD00	Houston, TX	2169544,00	24564,00	351,00	694,00	12040,00	11479,00	12,00	1132,22	16,18	31,99	554,96	529,10
76	2162	2006	TXHPD00	Houston, TX	2073729,00	24250,00	377,00	854,00	11648,00	11371,00	12,00	1169,39	18,18	41,18	561,69	548,34
77	2093	2005	TXHPD00	Houston, TX	2045732,00	23987,00	334,00	872,00	11653,00	11128,00	12,00	1172,54	16,33	42,63	569,62	543,96



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
77	2093	2005	TXHPD00	Houston, TX	2045732,00	23987,00	334,00	872,00	11653,00	11128,00	12,00	1172,54	16,33	42,63	569,62	543,96
78	2024	2004	TXHPD00	Houston, TX	2043446,00	23427,00	272,00	908,00	12065,00	10182,00	12,00	1146,45	13,31	44,43	590,42	498,28
79	1955	2003	TXHPD00	Houston, TX	2041081,00	23988,00	278,00	788,00	11957,00	10985,00	12,00	1175,26	13,62	37,63	585,82	538,20
80	1817	2001	TXHPD00	Houston, TX	1997965,00	23419,00	267,00	945,00	12286,00	9921,00	12,00	1172,14	13,36	47,30	614,93	496,56
81	117	1976	PAPEP00	Philadelphia, PA	1916665,00	13113,00	338,00	764,00	4225,00	7786,00	12,00	684,16	17,63	39,86	220,43	406,23
82	255	1978	PAPEP00	Philadelphia, PA	1885533,00	12369,00	352,00	724,00	4224,00	7069,00	12,00	655,99	18,67	38,40	224,02	374,91
83	644	1984	TXHPD00	Houston, TX	1805783,00	15155,00	473,00	1269,00	4256,00	9157,00	12,00	839,25	26,19	70,27	235,69	507,09
84	1610	1998	TXHPD00	Houston, TX	1797157,00	20183,00	254,00	654,00	11564,00	7711,00	12,00	1123,05	14,13	36,39	643,46	429,07
85	782	1986	TXHPD00	Houston, TX	1779677,00	20576,00	408,00	1524,00	7718,00	10926,00	12,00	1156,16	22,93	85,63	433,67	613,93
86	1472	1996	TXHPD00	Houston, TX	1772143,00	22456,00	261,00	1002,00	12917,00	8276,00	12,00	1267,17	14,73	56,54	728,89	467,01
87	1334	1994	TXHPD00	Houston, TX	1758016,00	22986,00	375,00	931,00	11699,00	9981,00	12,00	1307,50	21,33	52,96	665,47	567,74
88	324	1979	PAPEP00	Philadelphia, PA	1757368,00	14537,00	385,00	838,00	4476,00	8838,00	12,00	827,20	21,91	47,68	254,70	502,91
89	713	1985	TXHPD00	Houston, TX	1746375,00	16461,00	457,00	1711,00	4704,00	9589,00	12,00	942,58	26,17	97,97	269,36	549,08
90	851	1987	TXHPD00	Houston, TX	1739999,00	18971,00	323,00	1172,00	7784,00	9692,00	12,00	1090,29	18,56	67,36	447,36	557,01
91	1403	1995	TXHPD00	Houston, TX	1734335,00	22260,00	316,00	837,00	11885,00	9222,00	12,00	1283,49	18,22	48,26	685,28	531,73
92	575	1983	TXHPD00	Houston, TX	1728783,00	15919,00	561,00	1179,00	3909,00	10270,00	12,00	920,82	32,45	68,20	226,11	594,06
93	1265	1993	TXHPD00	Houston, TX	1724327,00	25069,00	446,00	1109,00	12742,00	10772,00	12,00	1453,84	25,87	64,31	738,95	624,71
94	989	1989	TXHPD00	Houston, TX	1713499,00	19528,00	459,00	1152,00	8097,00	9820,00	12,00	1139,66	26,79	67,23	472,54	573,10
95	1196	1992	TXHPD00	Houston, TX	1695233,00	24837,00	465,00	1169,00	12073,00	11130,00	12,00	1465,10	27,43	68,96	712,17	656,54
96	600	1983	PAPEP00	Philadelphia, PA	1692364,00	16973,00	311,00	822,00	5096,00	10744,00	12,00	1002,92	18,38	48,57	301,12	634,85
97	531	1982	PAPEP00	Philadelphia, PA	1687556,00	17511,00	332,00	839,00	5736,00	10604,00	12,00	1037,65	19,67	49,72	339,90	628,36
98	462	1981	PAPEP00	Philadelphia, PA	1686834,00	17616,00	362,00	936,00	5502,00	10816,00	12,00	1044,32	21,46	55,49	326,17	641,20
99	393	1980	PAPEP00	Philadelphia, PA	1681175,00	17299,00	436,00	937,00	5043,00	10883,00	12,00	1028,98	25,93	55,73	299,97	647,34
100	506	1982	TXHPD00	Houston, TX	1679607,00	17801,00	678,00	1270,00	3461,00	12392,00	12,00	1059,83	40,37	75,61	206,06	737,79
101	669	1984	PAPEP00	Philadelphia, PA	1667545,00	15599,00	263,00	937,00	5034,00	9365,00	12,00	935,45	15,77	56,19	301,88	561,60
102	1127	1991	TXHPD00	Houston, TX	1665756,00	26651,00	608,00	1213,00	10947,00	13883,00	12,00	1599,93	36,50	72,82	657,18	833,44
103	945	1988	PAPEP00	Philadelphia, PA	1657285,00	17082,00	374,00	912,00	6938,00	8858,00	12,00	1030,72	22,57	55,03	418,64	534,49
104	1014	1989	PAPEP00	Philadelphia, PA	1652188,00	18054,00	475,00	784,00	6562,00	10233,00	12,00	1092,73	28,75	47,45	397,17	619,36
105	876	1987	PAPEP00	Philadelphia, PA	1649364,00	17398,00	338,00	1102,00	6656,00	9302,00	12,00	1054,83	20,49	66,81	403,55	563,97
106	807	1986	PAPEP00	Philadelphia, PA	1645144,00	17207,00	343,00	1086,00	6131,00	9647,00	12,00	1045,93	20,85	66,01	372,67	586,39
107	738	1985	PAPEP00	Philadelphia, PA	1640102,00	16209,00	273,00	1021,00	5534,00	9381,00	12,00	988,29	16,65	62,25	337,42	571,98
108	1058	1990	TXHPD00	Houston, TX	1630553,00	22637,00	568,00	1335,00	7813,00	12921,00	12,00	1388,30	34,83	81,87	479,16	792,43
109	437	1981	TXHPD00	Houston, TX	1621897,00	10274,00	497,00	967,00	1917,00	6893,00	9,00	633,46	30,64	59,62	118,19	425,00
110	299	1979	TXHPD00	Houston, TX	1619644,00	14216,00	654,00	1481,00	2770,00	9311,00	12,00	877,72	40,38	91,44	171,03	574,88
111	161	1977	TXHPD00	Houston, TX	1613303,00	9304,00	376,00	965,00	1810,00	6153,00	12,00	576,71	23,31	59,82	112,19	381,39
112	1152	1991	PAPEP00	Philadelphia, PA	1596699,00	22451,00	440,00	904,00	7186,00	13921,00	12,00	1406,09	27,56	56,62	450,05	871,86
113	1083	1990	PAPEP00	Philadelphia, PA	1585577,00	21387,00	503,00	734,00	7344,00	12806,00	12,00	1348,85	31,72	46,29	463,18	807,66
114	1359	1994	PAPEP00	Philadelphia, PA	1560576,00	20638,00	404,00	721,00	6807,00	12706,00	12,00	1322,46	25,89	46,20	436,19	814,19
115	1290	1993	PAPEP00	Philadelphia, PA	1559534,00	19576,00	439,00	785,00	6821,00	11531,00	12,00	1255,25	28,15	50,34	437,37	739,39

Таблица 1 – MUP Crime.xls

1.3 Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос".

Для преобразования исходных данных обучающей выборки в базы данных системы "Эйдос" необходимо файл MicrosoftOfficeExcel, который

содержи базу данных компьютерных процессоров скопировать в Aidos-X\AID_DATA\Inp_data и назвать Inp_data.xlsx.

Далее запускаем систему "Эйдос" из папки "Aidos-X" файлом _aidos-x.exe. Система попросит ввести логин и пароль. Необходимо ввести: логин – 1, пароль – 1. После откроется главное окно программы.

Для загрузки базы данных необходимо зайти в режим 2.3.2.2 и в настройках изменить (рисунок 1):

– "Задайте тип файла исходных данных Inp_data": "XLSX-MSExcel-2003";

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

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 3, "Конечный столбец описательных шкал" – 16;

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

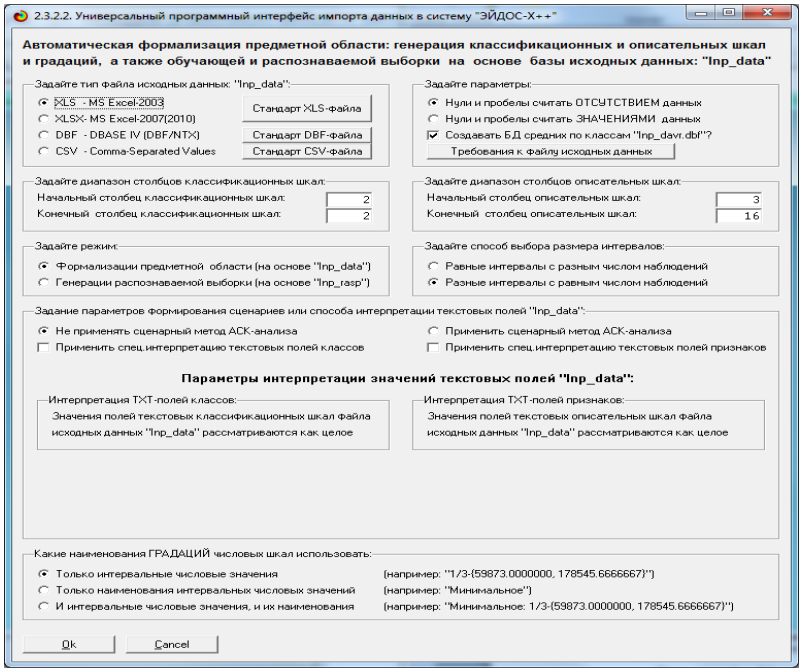
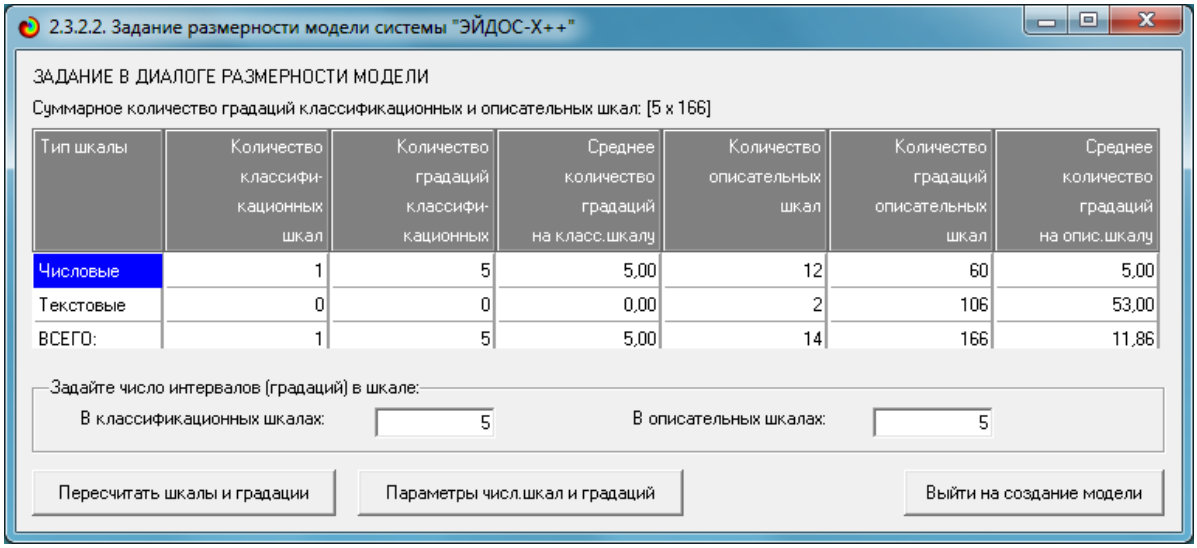


Рисунок 1 – Универсальный программный интерфейс импорта данных в систему "Эйдос"

После нажать кнопку "Ок". Далее открывается окно, где размещена информация о размерности модели (рисунок 2). В этом окне необходимо задать число интервалов равным 5 и нажать кнопку "Выйти на создание модели".



Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	1	5	5,00	12	60	5,00
Текстовые	0	0	0,00	2	106	53,00
ВСЕГО:	1	5	5,00	14	166	11,86

Рисунок 2 – Задание размерности модели системы "Эйдос"

Далее открывается окно, где происходит процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос" (рисунок 3). В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "Ок".

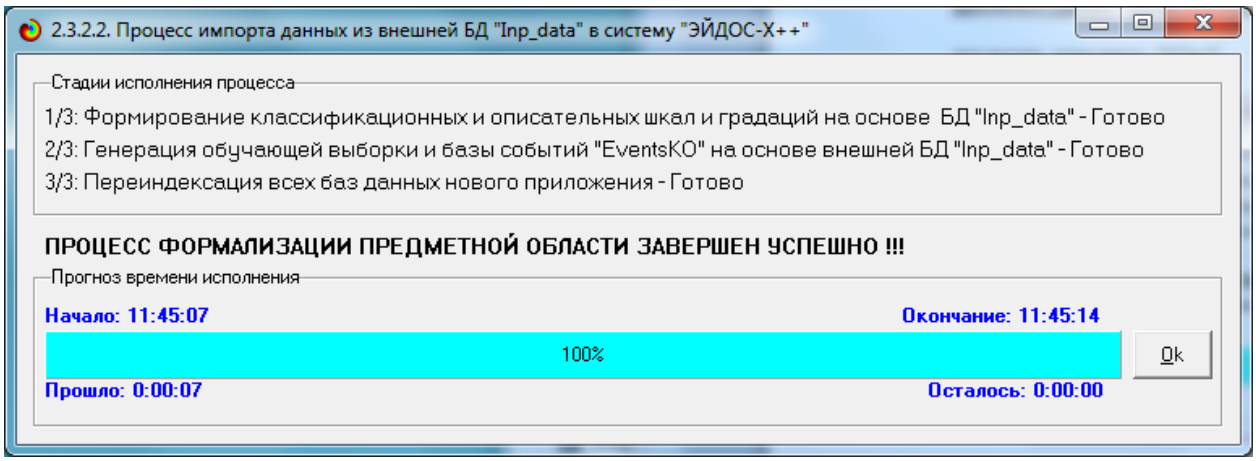


Рисунок 3 – Процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос"

В результате этого процесса будут сформированы классификационные и описательные шкалы и градации. Их применение позволит закодировать исходные данные и представить в форме эвентологических баз данных. Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1 (рисунок 4). Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 5). На рисунке 6 представлена обучающая выборка.

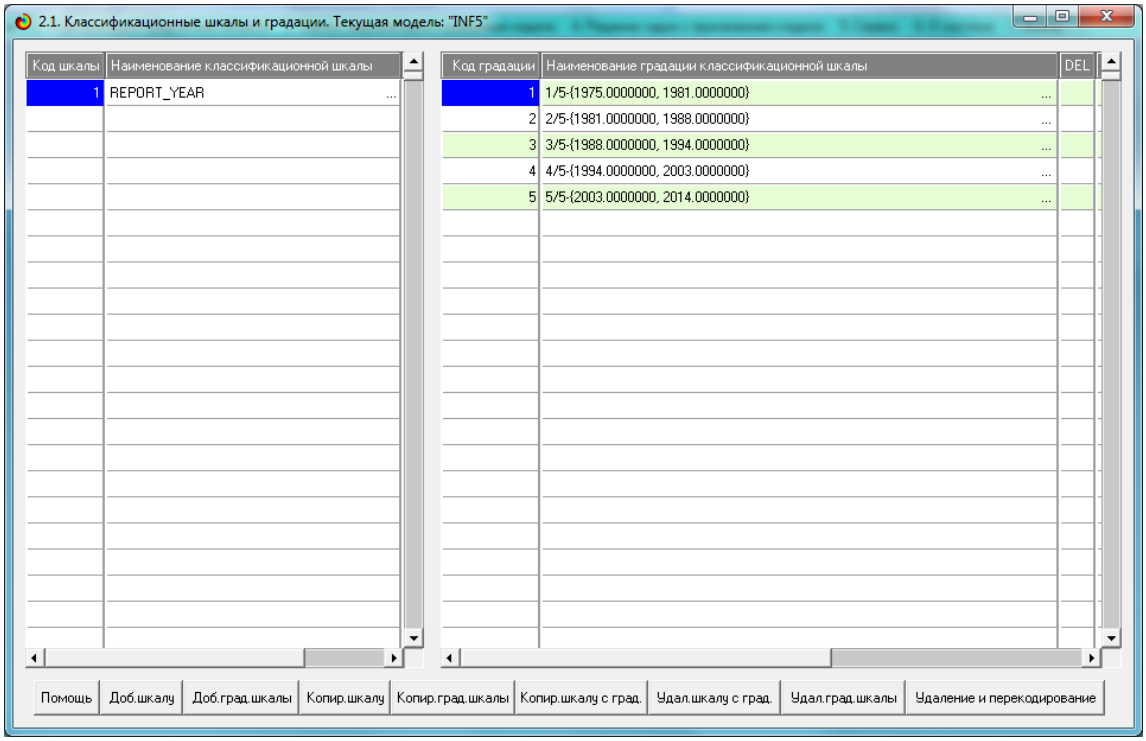


Рисунок 4 – Классификационные шкалы и градации

2.2. Описательные шкалы и градации. Текущая модель: "INF5"

Код шкалы	Наименование описательной шкалы
1	AGENCY_CODE
2	AGENCY_JURISDICTION
3	POPULATION
4	VIOLENT_CRIMES
5	HOMICIDES
6	RAPES
7	ASSAULTS
8	ROBBERIES
9	MONTHS_REPORTED
10	CRIMES_PERCAPITA
11	HOMICIDES_PERCAPITA
12	RAPES_PERCAPITA
13	ASSAULTS_PERCAPITA
14	ROBBERIES_PERCAPITA

Код градации	Наименование градации описательной шкалы
1	1/53-CA01900
2	2/53-CA01941
3	3/53-CA01942
4	4/53-CA03404
5	5/53-CA03711
6	6/53-CA03801
7	7/53-CODPD00
8	8/53-DCMPD00
9	9/53-FL01300
10	10/53-FL01306
11	11/53-FL01602
12	12/53-FL02902
13	13/53-FL04804
14	14/53-GAAPD00
15	15/53-ILCPD00
16	16/53-INIPD00
17	17/53-KS08703
18	18/53-LANPD00
19	19/53-MA01301
20	20/53-MD01604
21	21/53-MD01721
22	22/53-MDBPD00

Помощь | Доб.шкалу | Доб.град.шкалы | Копир.шкалу | Копир.град.шкалы | Копир.шкалу с град. | Удал.шкалу с град. | Удал.град.шкалы | Перекодировать | Очистить

Рисунок 5 – Описательные шкалы и градации

2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: "INF5"

Код объекта	Наименование объекта	Дата	Время
1	111		
2	180		
3	42		
4	1008		
5	1146		
6	1284		
7	939		
8	1491		
9	1353		
10	1077		

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	1	0	0	0

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	33	87	111	116	121	126	131
1	136	137	145	149	152	158	166

Помощь | Скопировать обуч.выб.в расп. | Добавить объект | Добавить классы | Добавить признаки | Удалить объект | Удалить классы | Удалить признаки | Очистить БД

Рисунок 6 – Обучающая выборка (фрагмент)

Таким образом, созданы 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).

1.3. Синтез и верификация статистических и интеллектуальных моделей

Далее запускаем режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей (рисунок 7).

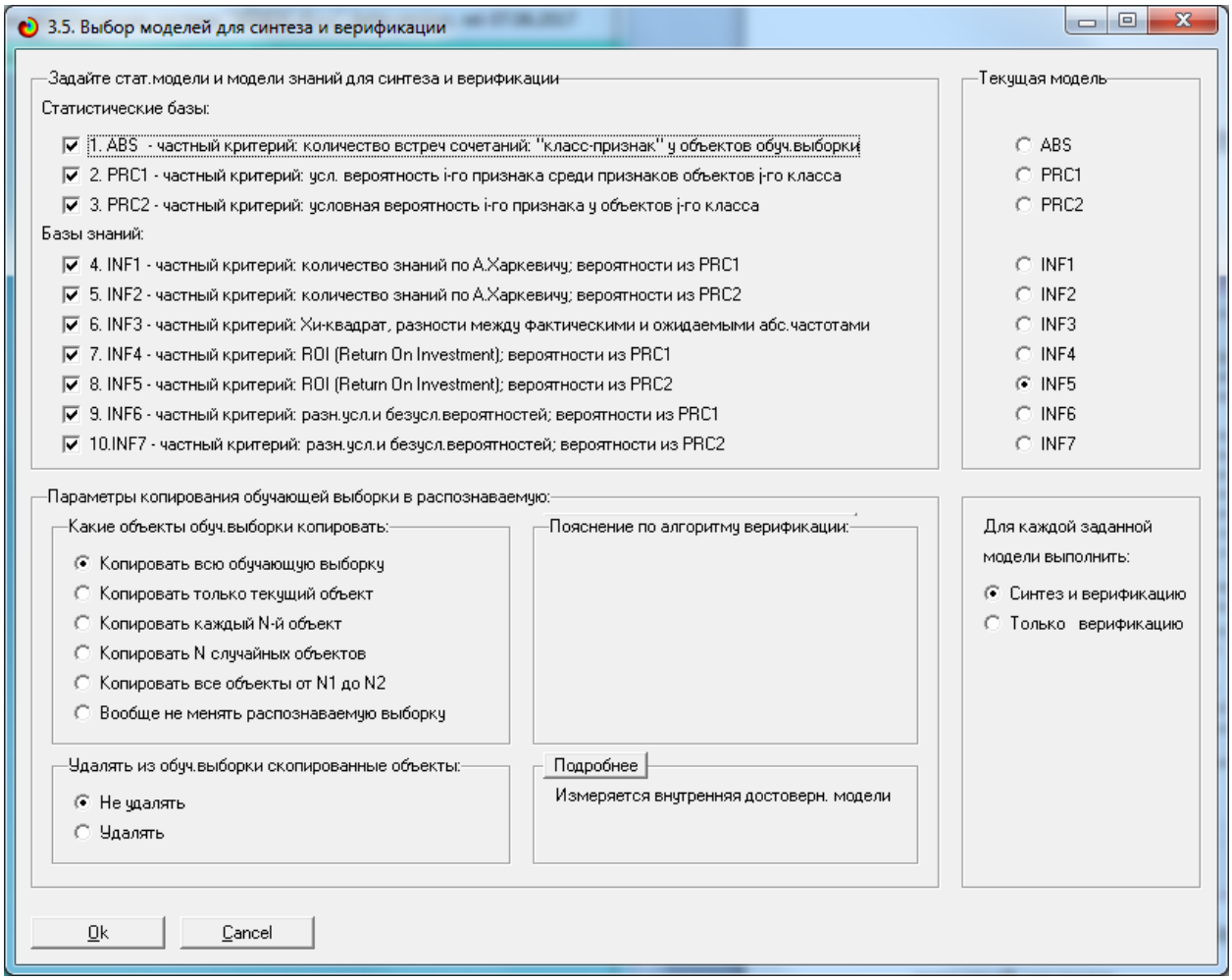


Рисунок 7 – Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

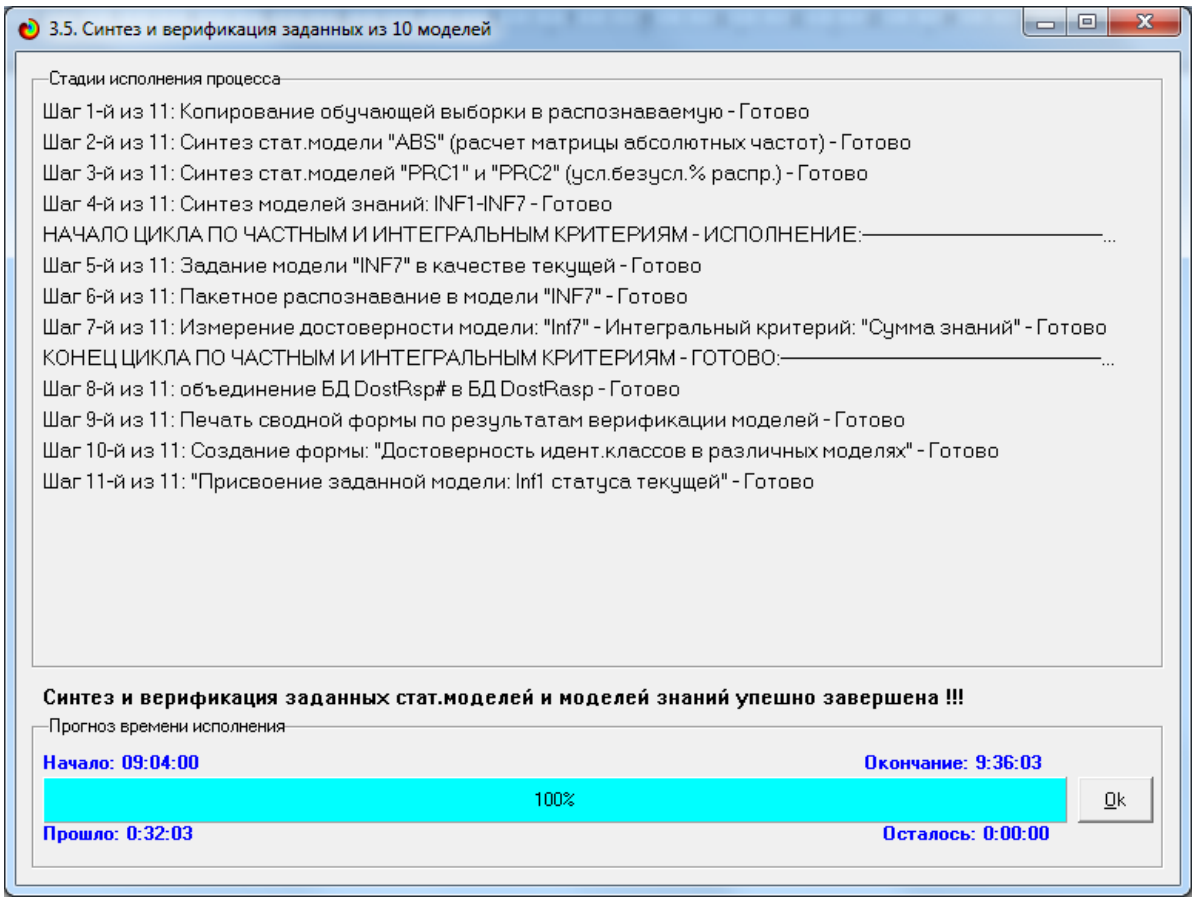


Рисунок 8 – Синтез и верификация статистических моделей и моделей знаний

Синтез и верификация всех моделей на данной задаче заняли 32 минуты. При этом верификация (оценка достоверности моделей) проводилась на всех 100 примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 8, но ниже приведем лишь некоторые из них (таблицы 2, 3, 4).

1.4. Виды моделей системы «Эйдос»

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

По сути, частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот (таблица 2) в матрицы условных и безусловных процентных распределений, и матрицы знаний (таблицы 3 и 4) (проф. В.И.Лойко).

Таблица 2 – Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений (фрагменты)

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

Код признака	Наименование описательной шкалы и градации	1. REPORT_Y... 1/5 {1975.00000... 1981.00000...	2. REPORT_Y... 2/5 {1981.00000... 1988.00000...	3. REPORT_Y... 3/5 {1988.00000... 1994.00000...	4. REPORT_Y... 4/5 {1994.00000... 2003.00000...	5. REPORT_Y... 5/5 {2003.00000... 2014.00000...	Сумма	Среднее	Средн. квадр. откл.
1	AGENCY_CODE-1/53-CA01900	4	5	6	4	1	20	4.000	1.871
2	AGENCY_CODE-2/53-CA01941	6	5	6	2		19	3.800	2.683
3	AGENCY_CODE-3/53-CA01942	6	7	4	6	1	24	4.800	2.387
4	AGENCY_CODE-4/53-CA03404	5	6	2	2		15	3.000	2.449
5	AGENCY_CODE-5/53-CA03711		1	1			2	0.400	0.548
6	AGENCY_CODE-6/53-CA03801	6	3	3	1		13	2.600	2.302
7	AGENCY_CODE-7/53-CODPD00	6	5	5		1	17	3.400	2.702
8	AGENCY_CODE-8/53-DCMPD00	6	6	5	9	9	35	7.000	1.871
9	AGENCY_CODE-9/53-FL01300	4	6	6	2		18	3.600	2.608
10	AGENCY_CODE-10/53-FL01306	5	7	6	5	5	28	5.600	0.894
11	AGENCY_CODE-11/53-FL01602	5	4	5		4	18	3.600	2.074
12	AGENCY_CODE-12/53-FL02902	4	6	5	3		18	3.600	2.302
13	AGENCY_CODE-13/53-FL04804	4	3	1	1	2	11	2.200	1.304
14	AGENCY_CODE-14/53-GAAPD00	7	7	6	9	8	37	7.400	1.140
15	AGENCY_CODE-15/53-ILCPD00	4	4	6	8	8	30	6.000	2.000
16	AGENCY_CODE-16/53-INIPD00	5	3	2	6	5	21	4.200	1.643
17	AGENCY_CODE-17/53-KS08703	2		2	1		5	1.000	1.000
18	AGENCY_CODE-18/53-LANPD00	6	6	6	9	9	36	7.200	1.643
19	AGENCY_CODE-19/53-MA01301	4	5	3	1	1	14	2.800	1.789
20	AGENCY_CODE-20/53-MD01604					1	1	0.200	0.447
21	AGENCY_CODE-21/53-MD01721			3	7	5	15	3.000	3.082
22	AGENCY_CODE-22/53-MDBPD00	6	7	6	9	11	39	7.800	2.168

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

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

Код признака	Наименование описательной шкалы и градации	1. REPORT_Y... 1/5 {1975.00000... 1981.00000...	2. REPORT_Y... 2/5 {1981.00000... 1988.00000...	3. REPORT_Y... 3/5 {1988.00000... 1994.00000...	4. REPORT_Y... 4/5 {1994.00000... 2003.00000...	5. REPORT_Y... 5/5 {2003.00000... 2014.00000...	Сумма	Среднее	Средн. квадр. откл.
1	AGENCY_CODE-1/53-CA01900	-0.006	0.040	0.105	0.009	-0.331	-0.184	-0.037	0.170
2	AGENCY_CODE-2/53-CA01941	0.105	0.052	0.117	-0.148		0.127	0.025	0.107
3	AGENCY_CODE-3/53-CA01942	0.048	0.077	-0.038	0.063	-0.375	-0.225	-0.045	0.190
4	AGENCY_CODE-4/53-CA03404	0.118	0.154	-0.093	-0.090		0.089	0.018	0.115
5	AGENCY_CODE-5/53-CA03711		0.209	0.229			0.438	0.088	0.120
6	AGENCY_CODE-6/53-CA03801	0.197	0.020	0.041	-0.224		0.034	0.007	0.151
7	AGENCY_CODE-7/53-CODPD00	0.132	0.079	0.100		-0.291	0.020	0.004	0.172
8	AGENCY_CODE-8/53-DCMPD00	-0.044	-0.052	-0.076	0.070	0.068	-0.034	-0.007	0.070
9	AGENCY_CODE-9/53-FL01300	0.019	0.110	0.130	-0.135		0.125	0.025	0.105
10	AGENCY_CODE-10/53-FL01306	-0.034	0.040	0.023	-0.019	-0.020	-0.011	-0.002	0.032
11	AGENCY_CODE-11/53-FL01602	0.074	0.011	0.086		0.033	0.204	0.041	0.038
12	AGENCY_CODE-12/53-FL02902	0.019	0.110	0.086	-0.036		0.179	0.036	0.061
13	AGENCY_CODE-13/53-FL04804	0.139	0.061	-0.186	-0.184	-0.016	-0.185	-0.037	0.146
14	AGENCY_CODE-14/53-GAAPD00	-0.020	-0.028	-0.045	0.056	0.026	-0.011	-0.002	0.042
15	AGENCY_CODE-15/53-ILCPD00	-0.105	-0.113	0.006	0.079	0.077	-0.057	-0.011	0.094
16	AGENCY_CODE-16/53-INIPD00	0.036	-0.096	-0.175	0.095	0.050	-0.090	-0.018	0.113
17	AGENCY_CODE-17/53-KS08703	0.162		0.175	0.009		0.346	0.069	0.091
18	AGENCY_CODE-18/53-LANPD00	-0.051	-0.059	-0.038	0.063	0.062	-0.024	-0.005	0.062
19	AGENCY_CODE-19/53-MA01301	0.080	0.127	0.023	-0.242	-0.244	-0.256	-0.051	0.179
20	AGENCY_CODE-20/53-MD01604					0.399	0.399	0.080	0.179
21	AGENCY_CODE-21/53-MD01721			0.006	0.215	0.132	0.353	0.071	0.098
22	AGENCY_CODE-22/53-MDBPD00	-0.070	-0.041	-0.058	0.043	0.091	-0.035	-0.007	0.071

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

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами"

Код признака	Наименование описательной шкалы и градации	1. REPORT_Y... 1/5 (1975.00000... 1981.00000...	2. REPORT_Y... 2/5 (1981.00000... 1988.00000...	3. REPORT_Y... 3/5 (1988.00000... 1994.00000...	4. REPORT_Y... 4/5 (1994.00000... 2003.00000...	5. REPORT_Y... 5/5 (2003.00000... 2014.00000...	Сумма	Среднее	Средн. квадр. отк.л.
1	AGENCY_CODE-1/53-CA01900	-0.107	0.755	2.096	0.139	-2.883			1.824
2	AGENCY_CODE-2/53-CA01941	2.098	0.967	2.292	-1.668	-3.689			2.597
3	AGENCY_CODE-3/53-CA01942	1.072	1.906	-0.684	1.366	-3.660			2.265
4	AGENCY_CODE-4/53-CA03404	1.920	2.816	-0.928	-0.896	-2.912			2.333
5	AGENCY_CODE-5/53-CA03711	-0.411	0.576	0.610	-0.386	-0.388			0.541
6	AGENCY_CODE-6/53-CA03801	3.331	0.241	0.463	-1.510	-2.524			2.237
7	AGENCY_CODE-7/53-CODPD00	2.509	1.392	1.682	-3.282	-2.301			2.604
8	AGENCY_CODE-8/53-DCMPD00	-1.187	-1.428	-1.831	2.242	2.204			2.043
9	AGENCY_CODE-9/53-FL01300	0.304	2.180	2.487	-1.475	-3.495			2.521
10	AGENCY_CODE-10/53-FL01306	-0.750	1.057	0.535	-0.406	-0.437			0.762
11	AGENCY_CODE-11/53-FL01602	1.304	0.180	1.487	-3.475	0.505			2.017
12	AGENCY_CODE-12/53-FL02902	0.304	2.180	1.487	-0.475	-3.495			2.208
13	AGENCY_CODE-13/53-FL04804	1.741	0.665	-1.147	-1.124	-0.136			1.232
14	AGENCY_CODE-14/53-GAAPD00	-0.598	-0.853	-1.222	1.856	0.816			1.293
15	AGENCY_CODE-15/53-ILCPD00	-2.160	-2.367	0.145	2.208	2.175			2.230
16	AGENCY_CODE-16/53-INIPD00	0.688	-1.457	-2.099	1.945	0.923			1.706
17	AGENCY_CODE-17/53-KS08703	0.973	-1.061	1.024	0.035	-0.971			1.008
18	AGENCY_CODE-18/53-LANPD00	-1.392	-1.641	-1.026	2.049	2.010			1.866
19	AGENCY_CODE-19/53-MA01301	1.125	2.029	0.267	-1.703	-1.718			1.681
20	AGENCY_CODE-20/53-MD01604	-0.205	-0.212	-0.195	-0.193	0.806			0.451
21	AGENCY_CODE-21/53-MD01721	-3.080	-3.184	0.072	4.104	2.088			3.195
22	AGENCY_CODE-22/53-MDBPD00	-2.008	-1.277	-1.612	1.470	3.428			2.354

1.5. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей (режим 4.1.3.6), отличающихся частными критериями с двумя приведенными выше интегральными критериями приведены на рисунке 9.

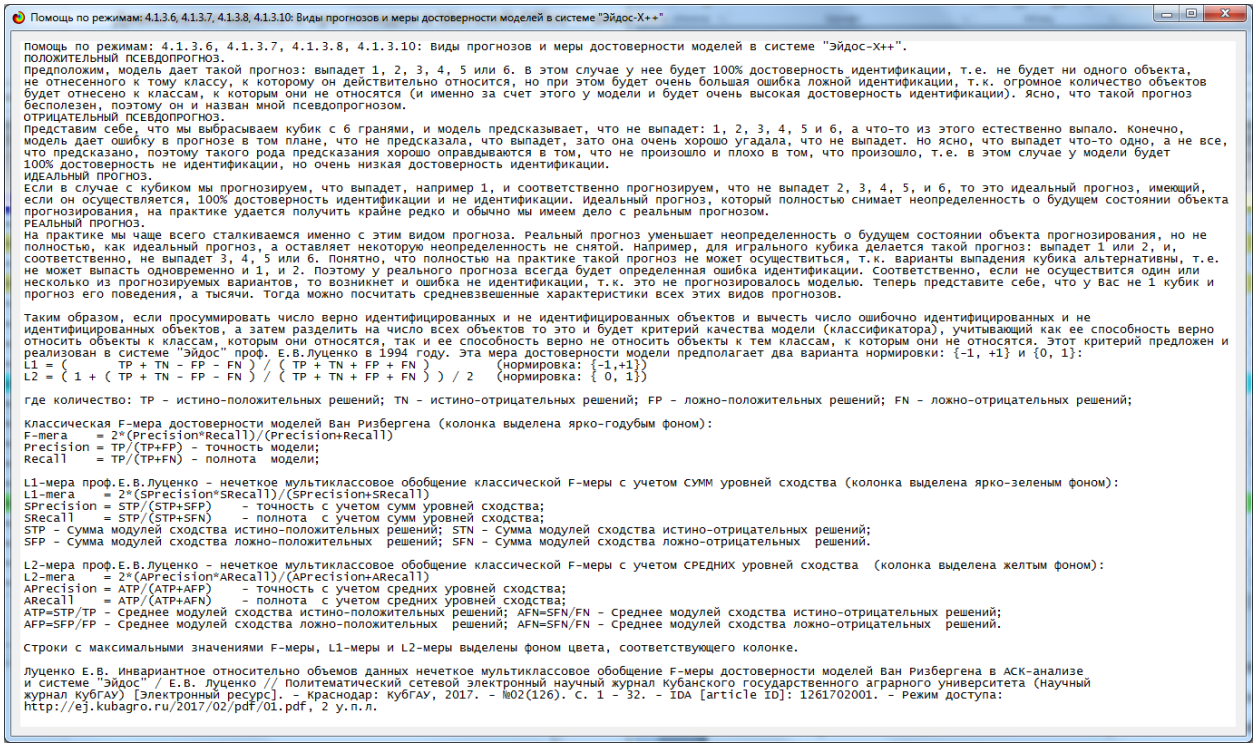
4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит.. Текущая модель: "INF5"

Наименование модели и частного критерия	Интегральный критерий	ота ли	F-мера Ван Рызбергена	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 модуль уровней сходс... истино-отрицат. решений	Средний модуль уровней сходс... ложно-положит. решений
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс.частот с обр....	1.000	0.333	696.312		2402.845		0.225	1.000	0.367	0.708		0.610
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс.частот по признак...	1.000	0.333	771.977		2802.548		0.216	1.000	0.355	0.785		0.712
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	1.000	0.333	696.312		2402.845		0.225	1.000	0.367	0.708		0.610
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	1.000	0.333	749.007		2721.380		0.216	1.000	0.355	0.761		0.691
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	1.000	0.333	696.311		2402.843		0.225	1.000	0.367	0.708		0.610
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	1.000	0.333	748.798		2720.618		0.216	1.000	0.355	0.761		0.691
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.605	0.451	114.827	800.494	144.512	54.702	0.443	0.677	0.535	0.193	0.141	0.136
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.731	0.479	153.939	703.100	197.616	37.804	0.438	0.803	0.567	0.214	0.143	0.152
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.605	0.451	114.844	800.534	144.531	54.713	0.443	0.677	0.535	0.193	0.141	0.136
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.731	0.479	153.993	703.313	197.653	37.802	0.438	0.803	0.567	0.214	0.143	0.152
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	0.794	0.449	269.580	657.121	418.957	35.524	0.392	0.884	0.543	0.345	0.175	0.245
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	0.794	0.449	219.495	526.795	335.383	28.083	0.396	0.887	0.547	0.281	0.138	0.196
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.411	0.423	78.041	1025.040	74.623	105.540	0.511	0.425	0.464	0.193	0.182	0.143
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.821	0.456	155.308	273.865	212.573	11.938	0.422	0.929	0.580	0.192	0.068	0.121
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0.411	0.423	78.051	1024.933	74.637	105.540	0.511	0.425	0.464	0.193	0.182	0.143
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	0.820	0.456	155.210	273.640	212.425	11.911	0.422	0.929	0.580	0.192	0.067	0.121
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	0.773	0.447	261.299	668.541	402.606	39.198	0.394	0.870	0.542	0.343	0.176	0.243
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	0.794	0.447	215.049	481.864	332.826	27.495	0.393	0.887	0.544	0.275	0.135	0.193
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	0.773	0.447	261.011	668.469	402.053	39.198	0.394	0.869	0.542	0.343	0.176	0.243
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	0.794	0.447	214.854	481.279	332.484	27.455	0.393	0.887	0.544	0.275	0.135	0.192

Помощь

Рисунок 9 Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF2 и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели INF5 составляет 0,465. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется нечеткое мультиклассовое обобщение F-меры, предложенное проф.Е.В.Луценко (L-мера), при этом точность модели составляет 0,580 по L1-меры. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%.



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 количество: TP - истино-положительных решений; TN - истино-отрицательных решений; FP - ложно-положительных решений; FN - ложно-отрицательных решений;

Классическая F-мера достоверности моделей Ван Ризбергена (колонка выделена ярко-годубым фоном):
F-mera = 2*(Precision*Recall)/(Precision+Recall)
Precision = TP/(TP+FP) - точность модели;
Recall = TP/(TP+FN) - 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-mera = 2*(SPrecision*SRecall)/(SPrecision+SRecall)
SPrecision = STP/(STP+SFP) - точность с учетом сумм уровней сходства;
SRecall = STP/(STP+SFN) - полнота с учетом сумм уровней сходства;
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.

L2-мера проф.Е.В.Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СРЕДНИХ уровней сходства (колонка выделена желтым фоном):
L2-mera = 2*(APrecision*ARecall)/(APrecision+ARecall)
APrecision = ATP/(ATP+AFP) - точность с учетом средних уровней сходства;
ARecall = ATP/(ATP+AFN) - полнота с учетом средних уровней сходства;
ATP=STP/TP - Среднее модулей сходства истино-положительных решений; AFN=SFN/FN - Среднее модулей сходства истино-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. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. - Краснодар: КубГАУ, 2017. - №02(126). С. 1 - 32. - IDA [article ID]: 1261702001. - Режим доступа: http://ej.kubagro.ru/2017/02/pdf/01.pdf, 2 у.п.л.

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

Также статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую. Этим и оправдано применение моделей знаний и интеллектуальных технологий. На рисунке 11 приведены частные распределения уровней сходства и различия для верно и

ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF5.

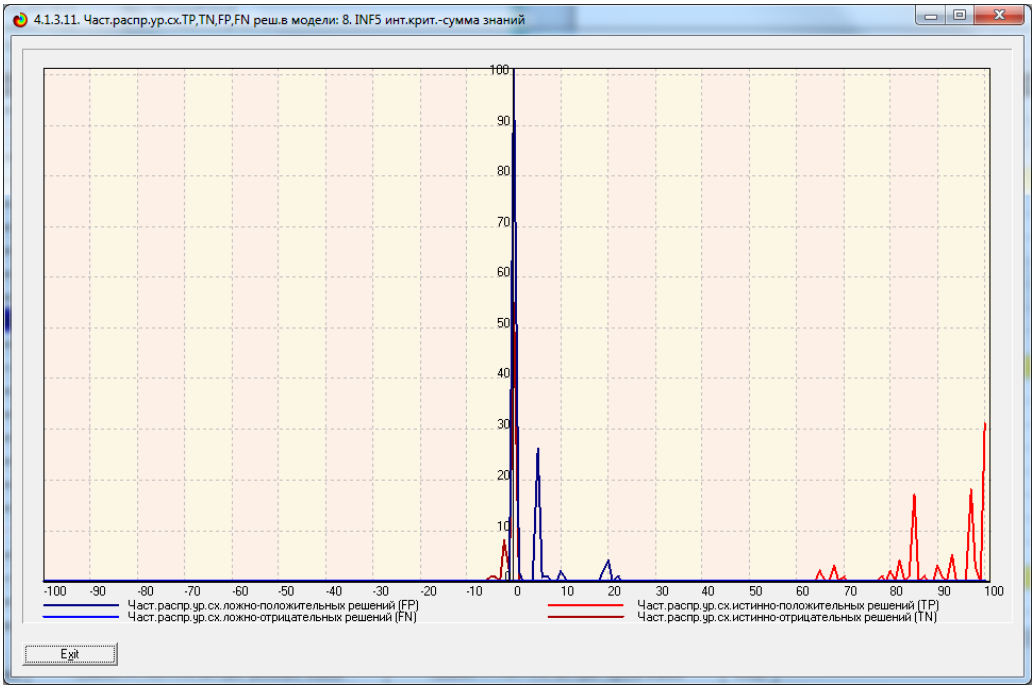


Рисунок 11. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF5

Из рисунка 11 видно, что:

– наиболее достоверная модель INF5 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 9);

– модуль уровня сходства-различия в наиболее достоверной модели INF5 для верно идентифицированных и верно неидентифицированных объектов 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 больших значений уровней сходства-различия (более 50%) различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями настолько велико, что учитывая

уровень сходства-различия не представляет никакой проблемы разделить истинные положительные и отрицательные решения от ложных положительных и отрицательных решений. Это и стало основой для нечеткого обобщения F-меры Ван Ризбергена, т.е. такого ее обобщения, которое учитывает е только сам факт ложно или истинно положительного или отрицательного решения, но и его надежность.

Возьмем любые данные о наблюдениях как сумму истинного значения и шума, однако ни первое, ни второе неизвестны. Поэтому необходимо сравнить созданные модели с чисто случайными моделями, совпадающими по основным характеристикам. В системе «Эйдос» есть лабораторная работа № 2.01: «Исследование RND-модели при различных объемах выборки». Если данная работа устанавливается при отсутствии текущего приложения, то все параметры создаваемых моделей задаются вручную, если же текущая модель существует, как в нашем случае, то все основные ее параметры определяются автоматически.

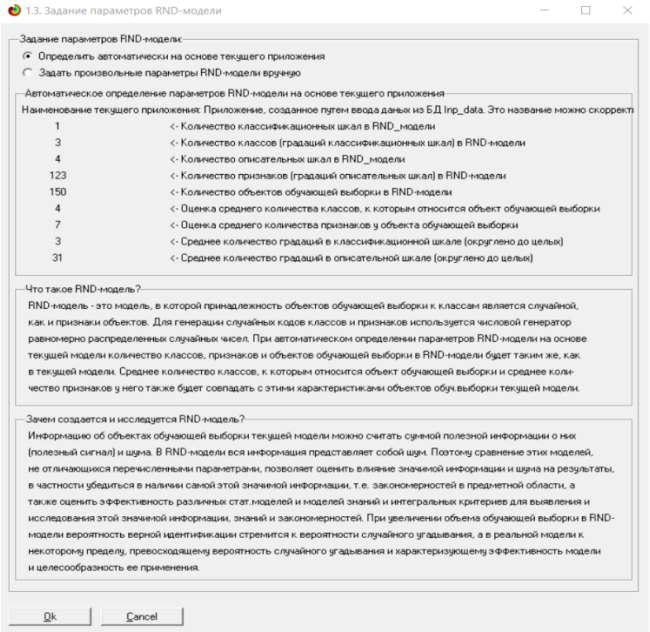


Рисунок 12. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 13 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF5.

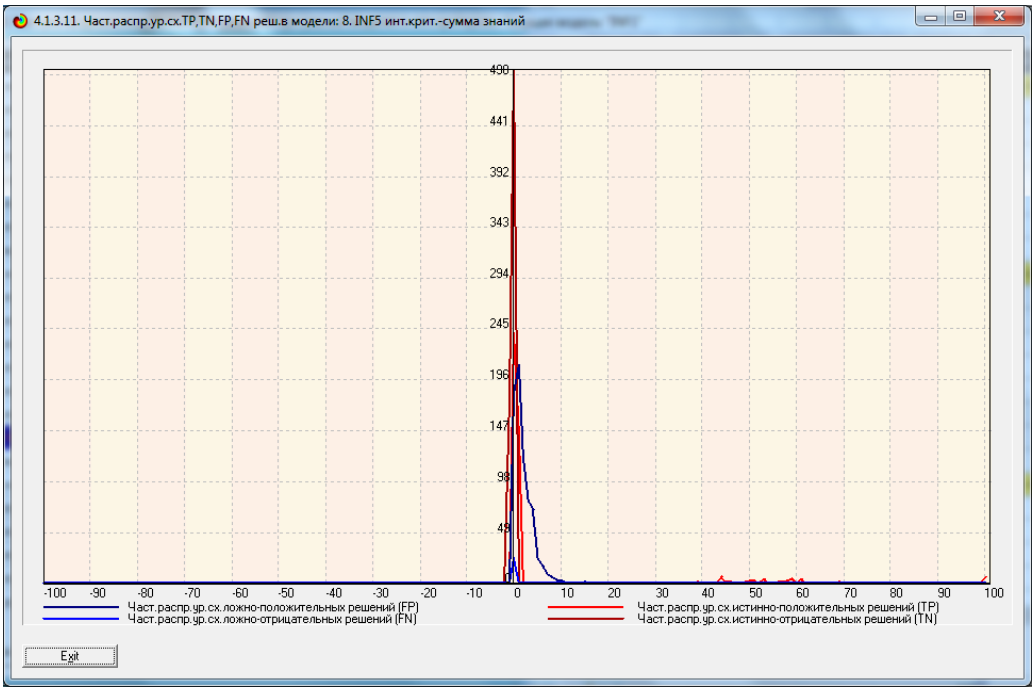


Рисунок 13. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF4

Различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели объясняется тем, что в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны.

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

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

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительн... решений (TP)	Число истинно-отрицательных решений (TN)	Число ложн... положитель... решений (FP)	Число ложн... отрицатель... решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма моду... уровней схо... истинно-поло... решений (ST...	Су ур ис ре
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс.частот с обр....	162	162	509	729		0.182	1.000	0.308	81.260	
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс.частот по признак...	162	162	445	793		0.170	1.000	0.290	75.596	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	162	162	509	729		0.182	1.000	0.308	81.260	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	162	162	445	793		0.170	1.000	0.290	46.536	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	162	162	509	729		0.182	1.000	0.308	81.258	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	162	162	445	793		0.170	1.000	0.290	43.435	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	162	107	913	325	55	0.248	0.660	0.360	27.699	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	162	157	712	526	5	0.230	0.969	0.372	36.794	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	162	107	926	312	55	0.255	0.660	0.368	27.637	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	162	157	718	520	5	0.232	0.969	0.374	36.122	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	162	141	866	372	21	0.275	0.870	0.418	60.925	
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	162	141	866	372	21	0.275	0.870	0.418	56.137	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	162	107	1130	108	55	0.498	0.660	0.568	26.877	
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	162	159	698	540	3	0.227	0.981	0.369	13.069	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	162	106	1130	108	56	0.495	0.654	0.564	26.844	
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	162	159	700	538	3	0.228	0.981	0.370	12.254	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	162	141	803	435	21	0.245	0.870	0.382	61.316	
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	162	141	705	533	21	0.209	0.870	0.337	24.449	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	162	141	809	429	21	0.247	0.870	0.385	62.105	
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	162	141	704	534	21	0.209	0.870	0.337	22.683	

Помощь

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

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

– достоверность лучшей модели INF5, отражающей реальный объект моделирования, примерно на 17% 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 INF4 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1 (рисунок 16).

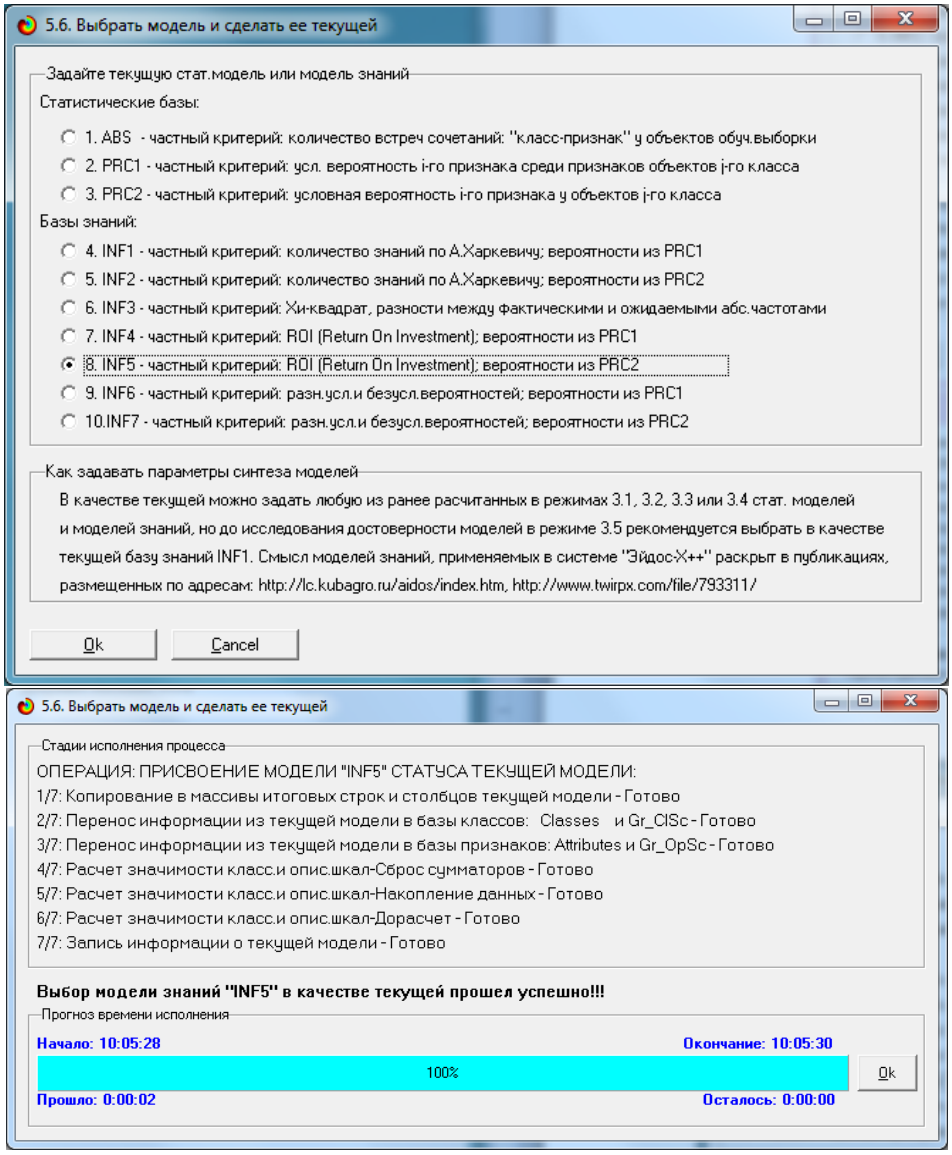


Рисунок 15. Экранные формы режима задания модели в качестве текущей

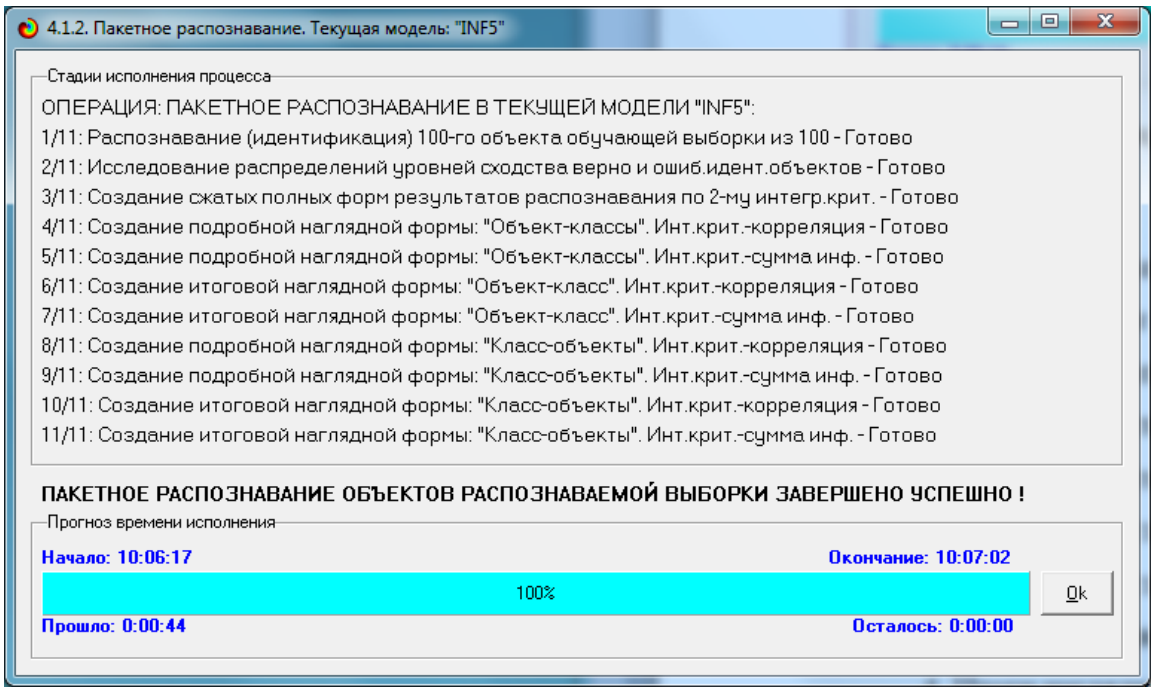


Рисунок 16. Экранная форма режима пакетного распознавания в текущей модели

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

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

1. Подробно наглядно: "Объект – классы".

2. Подробно наглядно: "Класс – объекты".

3. Итоги наглядно: "Объект – классы".

4. Итоги наглядно: "Класс – объекты".

5. Подробно сжато: "Объект – классы".

6. Обобщенная форма по достоверности моделей при разных интегральных критериях.

7. Обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям.

8. Статистический анализ результатов идентификации по классам, моделям и интегральным критериям.

9. Распознавание уровня сходства при разных моделях и интегральных критериях.

10.Достоверность идентификации классов при разных моделях и интегральных критериях.

Ниже кратко рассмотрим некоторые из них.

На рисунках 17 и 18 приведены примеры прогнозов преступности в зависимости от отчетного года, населения, насильственных преступлений в наиболее достоверной модели INF5 на основе наблюдения 1975-2015 г.:

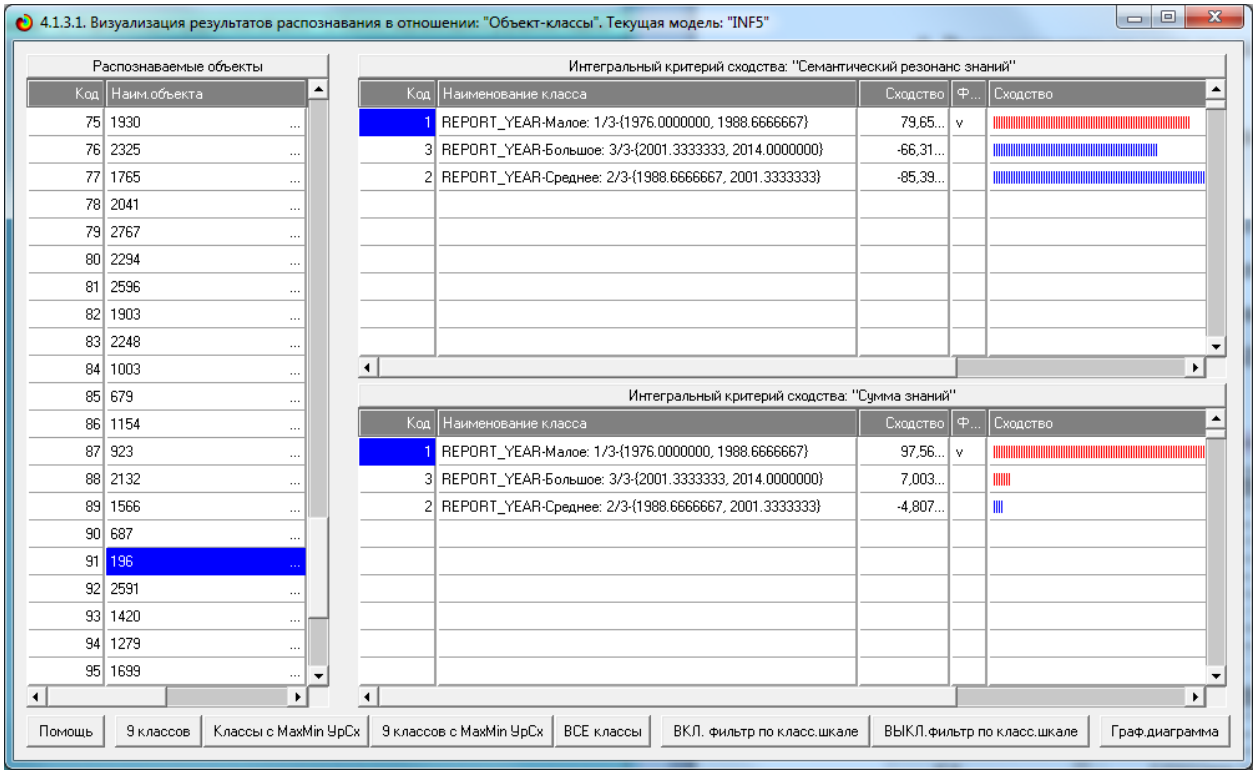


Рисунок 17. Пример идентификации дисциплины отчетный год в модели INF5

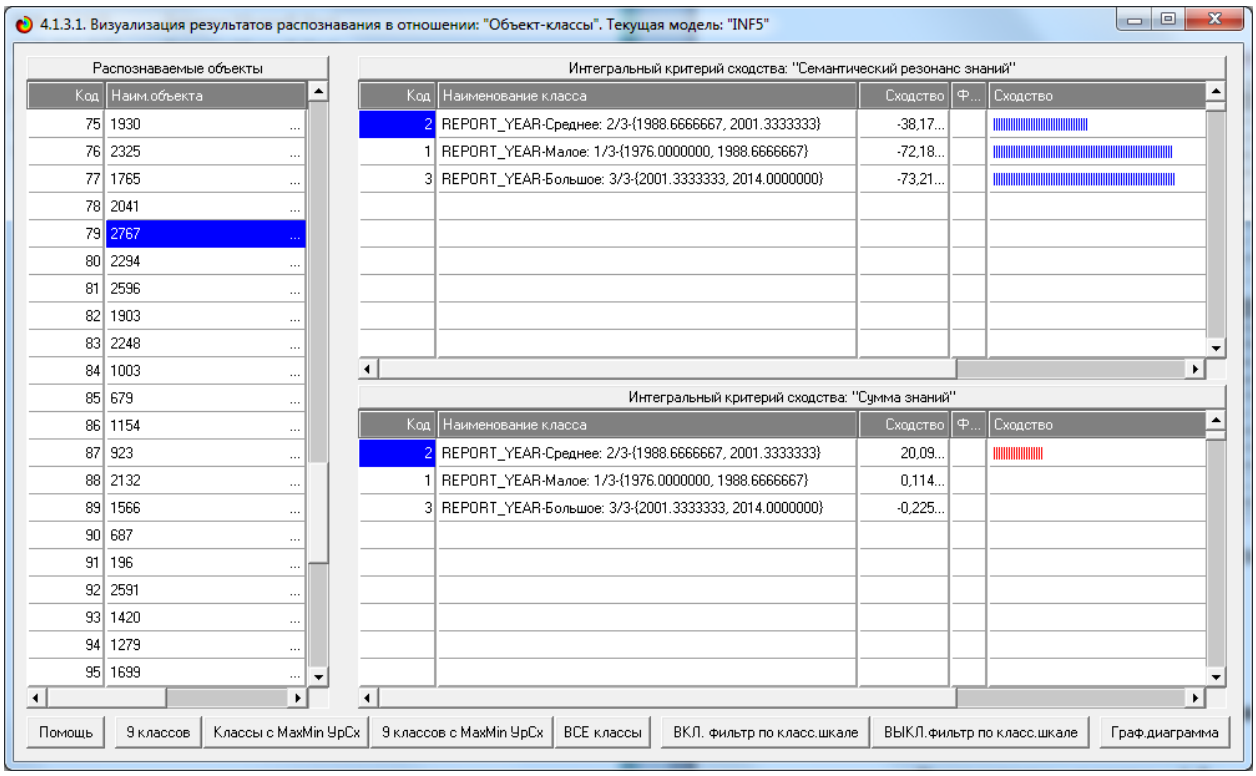


Рисунок 18. Пример идентификации дисциплины отчетный год в модели INF5

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.

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, поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 20 приведены визуализации всех когнитивных функций данного приложения для модели INF5.

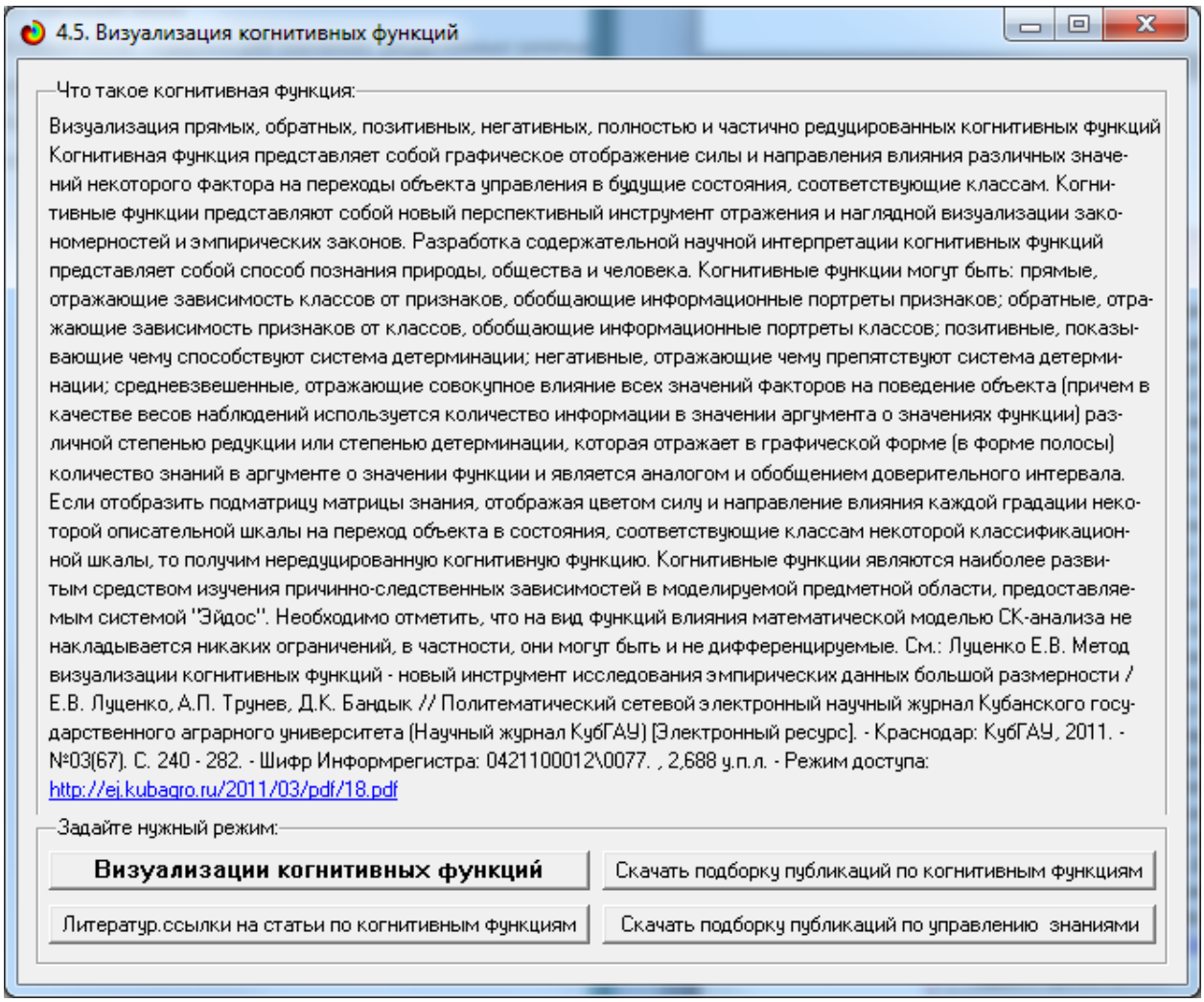
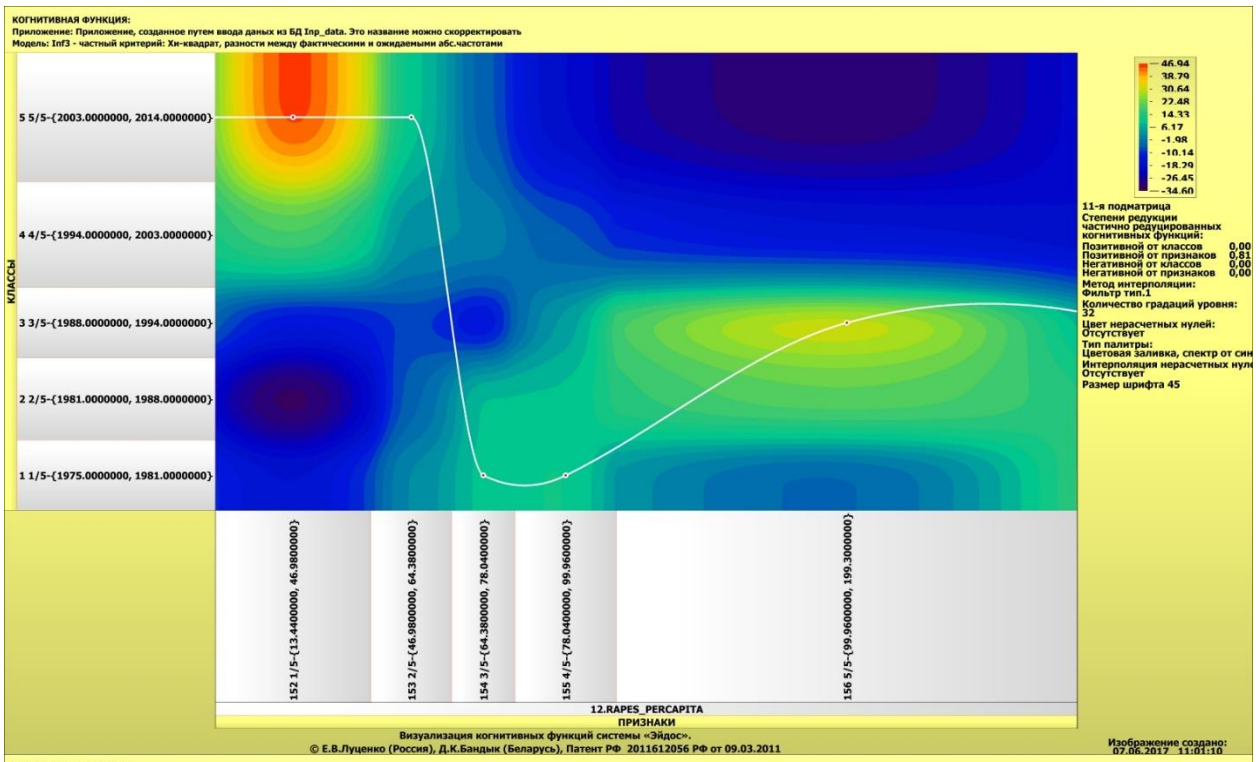
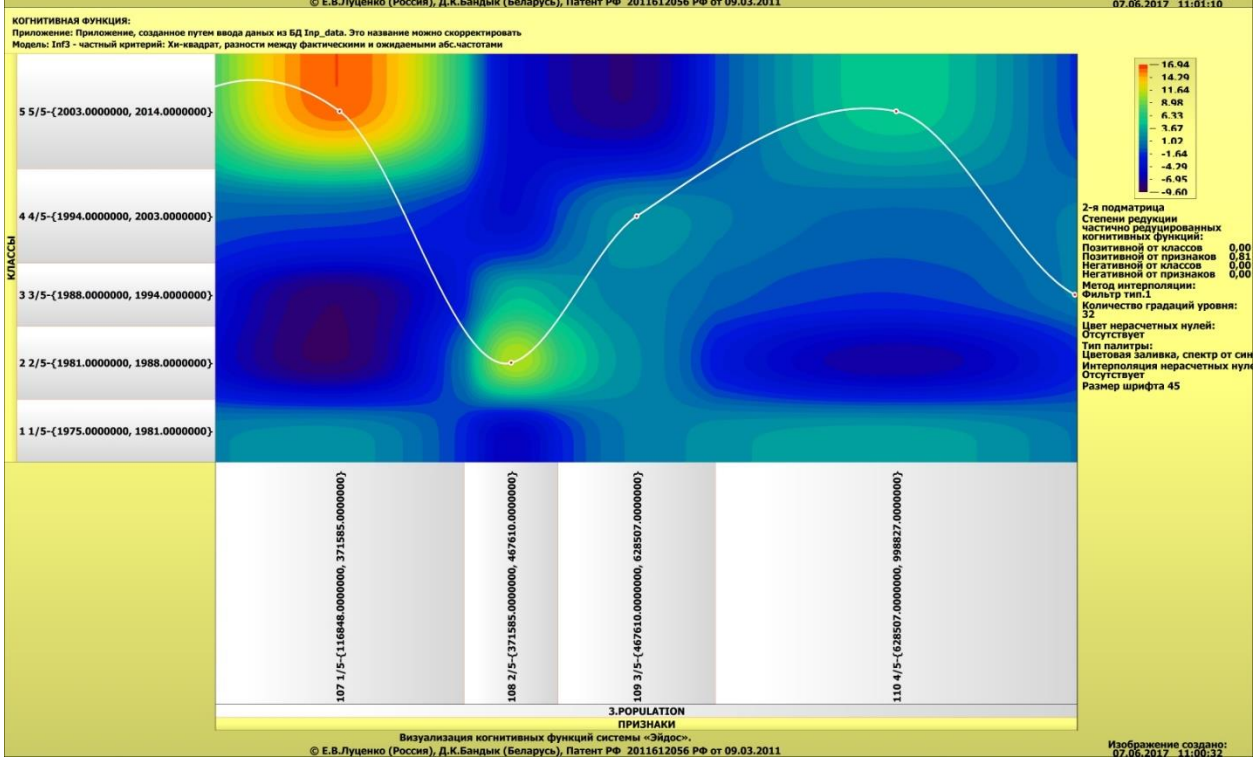
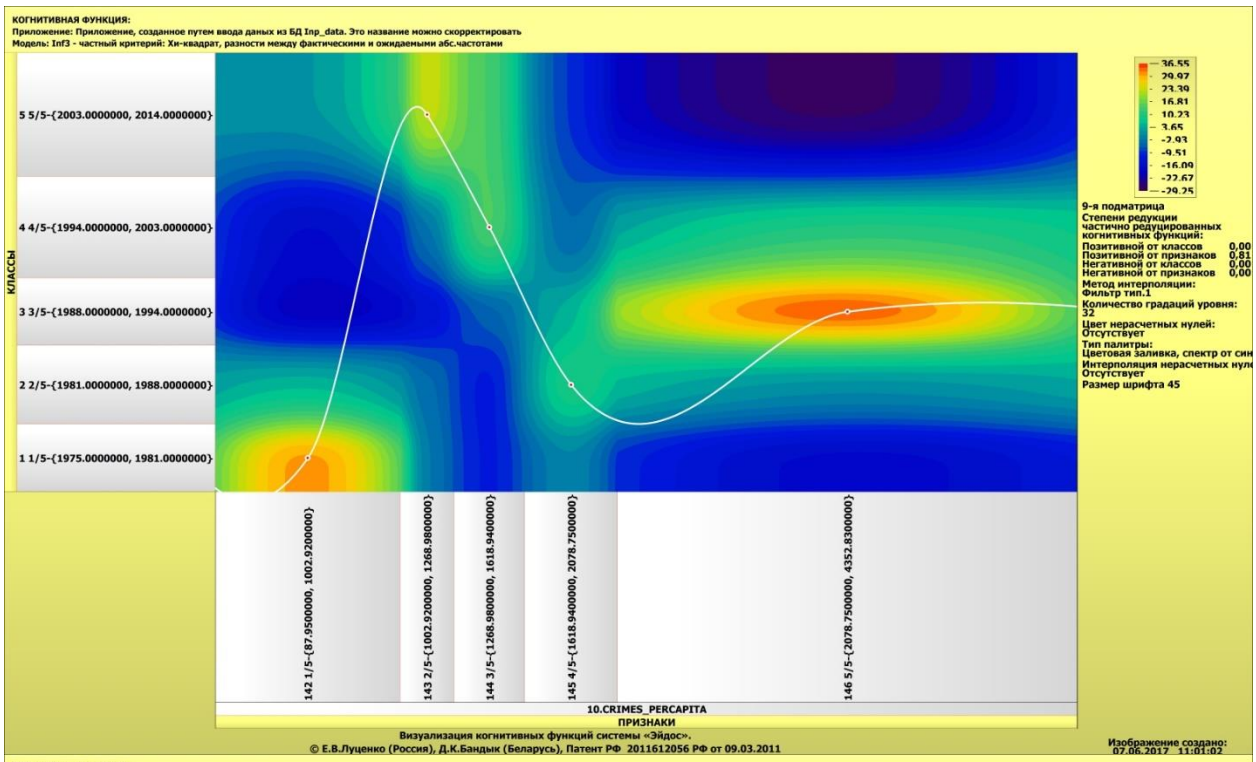
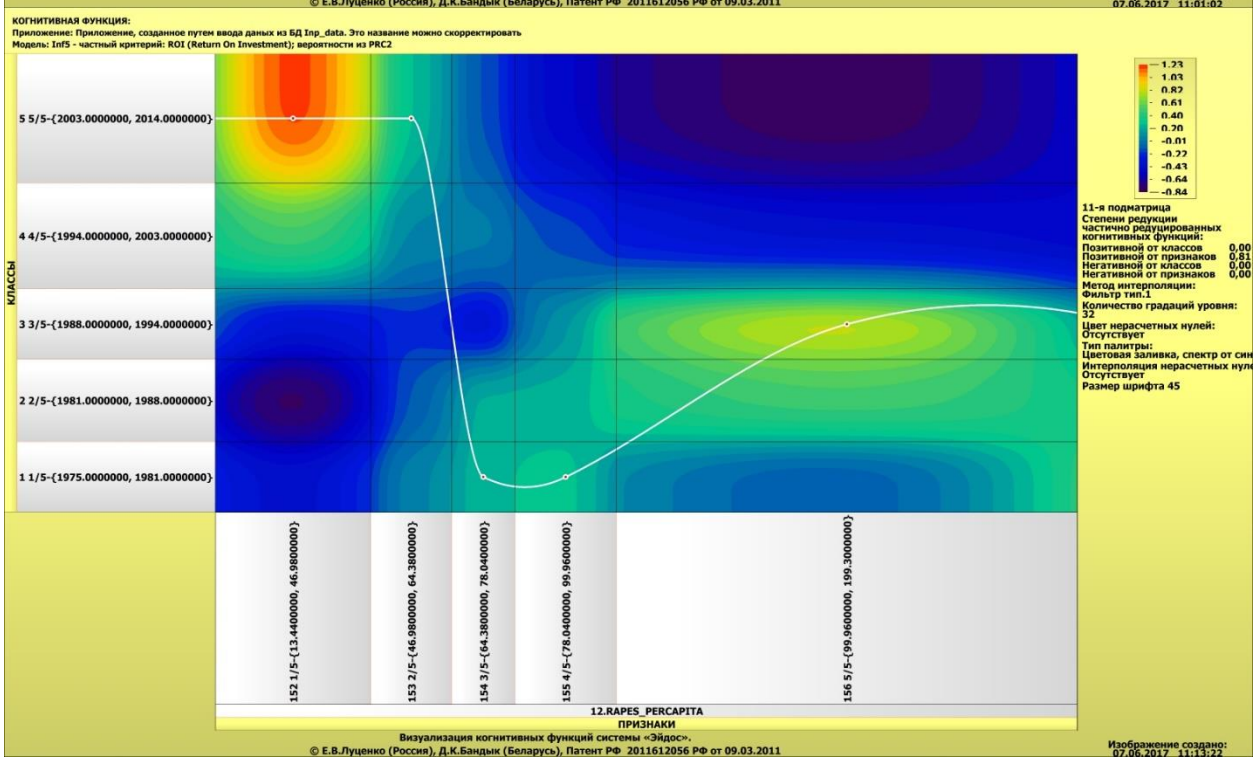


Рисунок 19. Экранная форма режима 4.5 системы «Эйдос-Х++» «Визуализация когнитивных функций»


КОГНИТИВНАЯ ФУНКЦИЯ:
Приложение: Приложение, созданное путем ввода даных из БД Inp_data. Это название можно скорректировать
Модель: Inf3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами
КЛАССЫ
5 5/5-{2003.0000000, 2014.0000000}
4 4/5-{1994.0000000, 2003.0000000}
3 3/5-{1988.0000000, 1994.0000000}
2 2/5-{1981.0000000, 1988.0000000}
1 1/5-{1975.0000000, 1981.0000000}
152 1/5-{13.4400000, 46.9800000}
153 2/5-{46.9800000, 64.3800000}
154 3/5-{64.3800000, 78.0400000}
155 4/5-{78.0400000, 99.9600000}
156 5/5-{99.9600000, 199.3000000}
12.RAPES_PERCAPITA
ПРИЗНАКИ
46.94
38.79
30.64
22.48
14.33
6.17
-1.98
-10.14
-18.29
-26.45
-34.60
11-я подматрица
Степени редукции
частично редуцированных
когнитивных функций:
Позитивной от классов 0,00
Позитивной от признаков 0,81
Негативной от классов 0,00
Негативной от признаков 0,00
Метод интерполяции:
Фильтр тип.1
Количество градаций уровня:
32
Цвет нерасчетных нулей:
Отсутствует
Тип палитры:
Цветовая заливка, спектр от син
Интерполяция нерасчетных нуле
Отсутствует
Размер шрифта 45
Визуализация когнитивных функций системы «Эйдос».
© Е.В.Луценко (Россия), Д.К.Бандык (Беларусь), Патент РФ 2011612056 РФ от 09.03.2011
Изображение создано:
07.06.2017 11:01:10


КОГНИТИВНАЯ ФУНКЦИЯ:
Приложение: Приложение, созданное путем ввода даных из БД Inp_data. Это название можно скорректировать
Модель: Inf3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами
КЛАССЫ
5 5/5-{2003.0000000, 2014.0000000}
4 4/5-{1994.0000000, 2003.0000000}
3 3/5-{1988.0000000, 1994.0000000}
2 2/5-{1981.0000000, 1988.0000000}
1 1/5-{1975.0000000, 1981.0000000}
107 1/5-{116848.0000000, 371585.0000000}
108 2/5-{371585.0000000, 467610.0000000}
109 3/5-{467610.0000000, 628507.0000000}
110 4/5-{628507.0000000, 998827.0000000}
3.POPULATION
ПРИЗНАКИ
16.94
14.29
11.64
8.98
6.33
3.67
1.02
-1.64
-4.29
-6.95
-9.60
2-я подматрица
Степени редукции
частично редуцированных
когнитивных функций:
Позитивной от классов 0,00
Позитивной от признаков 0,81
Негативной от классов 0,00
Негативной от признаков 0,00
Метод интерполяции:
Фильтр тип.1
Количество градаций уровня:
32
Цвет нерасчетных нулей:
Отсутствует
Тип палитры:
Цветовая заливка, спектр от син
Интерполяция нерасчетных нуле
Отсутствует
Размер шрифта 45
Визуализация когнитивных функций системы «Эйдос».
© Е.В.Луценко (Россия), Д.К.Бандык (Беларусь), Патент РФ 2011612056 РФ от 09.03.2011
Изображение создано:
07.06.2017 11:00:32


КОГНИТИВНАЯ ФУНКЦИЯ:
Приложение: Приложение, созданное путем ввода даных из БД Inp_data. Это название можно скорректировать
Модель: Inf3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами
КЛАССЫ
5 5/5-{2003.0000000, 2014.0000000}
4 4/5-{1994.0000000, 2003.0000000}
3 3/5-{1988.0000000, 1994.0000000}
2 2/5-{1981.0000000, 1988.0000000}
1 1/5-{1975.0000000, 1981.0000000}
142 1/5-{87.9500000, 1002.9200000}
143 2/5-{1002.9200000, 1268.9800000}
144 3/5-{1268.9800000, 1618.9400000}
145 4/5-{1618.9400000, 2078.7500000}
146 5/5-{2078.7500000, 4352.8300000}
10.CRIMES_PERCAPITA
ПРИЗНАКИ
36.55
29.97
23.39
16.81
10.23
3.65
-2.93
-9.51
-16.09
-22.67
-29.25
9-я подматрица
Степени редукции
частично редуцированных
когнитивных функций:
Позитивной от классов 0,00
Позитивной от признаков 0,81
Негативной от классов 0,00
Негативной от признаков 0,00
Метод интерполяции:
Фильтр тип.1
Количество градаций уровня:
32
Цвет нерасчетных нулей:
Отсутствует
Тип палитры:
Цветовая заливка, спектр от син
Интерполяция нерасчетных нуле
Отсутствует
Размер шрифта 45
Визуализация когнитивных функций системы «Эйдос».
© Е.В.Луценко (Россия), Д.К.Бандык (Беларусь), Патент РФ 2011612056 РФ от 09.03.2011
Изображение создано:
07.06.2017 11:01:02


КОГНИТИВНАЯ ФУНКЦИЯ:
Приложение: Приложение, созданное путем ввода даных из БД Inp_data. Это название можно скорректировать
Модель: Inf5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
КЛАССЫ
5 5/5-{2003.0000000, 2014.0000000}
4 4/5-{1994.0000000, 2003.0000000}
3 3/5-{1988.0000000, 1994.0000000}
2 2/5-{1981.0000000, 1988.0000000}
1 1/5-{1975.0000000, 1981.0000000}
152 1/5-{13.4400000, 46.9800000}
153 2/5-{46.9800000, 64.3800000}
154 3/5-{64.3800000, 78.0400000}
155 4/5-{78.0400000, 99.9600000}
156 5/5-{99.9600000, 199.3000000}
12.RAPES_PERCAPITA
ПРИЗНАКИ
1.23
1.03
0.82
0.61
0.40
0.20
-0.01
-0.22
-0.43
-0.64
-0.84
11-я подматрица
Степени редукции
частично редуцированных
когнитивных функций:
Позитивной от классов 0,00
Позитивной от признаков 0,81
Негативной от классов 0,00
Негативной от признаков 0,00
Метод интерполяции:
Фильтр тип.1
Количество градаций уровня:
32
Цвет нерасчетных нулей:
Отсутствует
Тип палитры:
Цветовая заливка, спектр от син
Интерполяция нерасчетных нуле
Отсутствует
Размер шрифта 45
Визуализация когнитивных функций системы «Эйдос».
© Е.В.Луценко (Россия), Д.К.Бандык (Беларусь), Патент РФ 2011612056 РФ от 09.03.2011
Изображение создано:
07.06.2017 11:13:22

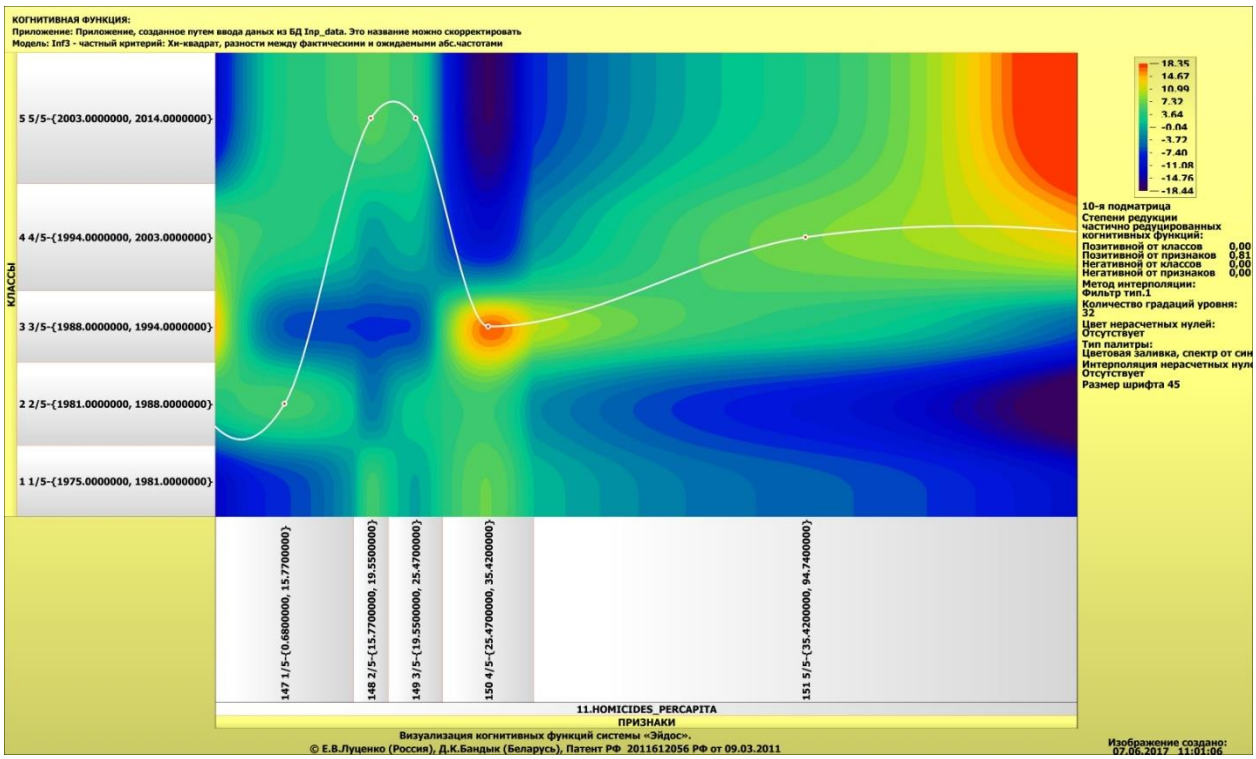


Рисунок 20. Визуализация когнитивных функций для обобщенных классов и описательных шкал для модели INF5

На основе полученных данных сделаем следующие прогнозы по классификации преступлений за 1975-2015г.:

1). С каждым годом уровень преступности повышается ;

2 Большее количество преступлений происходит в 1988-1994г.;

Визуализация показала, что за период 1975-2015г. самым совершаемым видом преступление оказался изнасилование. Более подробные данные по преступлениям представлены на рисунке 20.

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, 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 21).

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса
1	REPORT_YEAR-1/5-{1975.0000000, 1981.0000000}
2	REPORT_YEAR-2/5-{1981.0000000, 1988.0000000}
3	REPORT_YEAR-3/5-{1988.0000000, 1994.0000000}
4	REPORT_YEAR-4/5-{1994.0000000, 2003.0000000}
5	REPORT_YEAR-5/5-{2003.0000000, 2014.0000000}

SWOT-анализ класса: 1 "REPORT_YEAR-1/5-{1975.0000000, 1981.0000000}" в модели: 6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
157	ASSAULTS_PERCAPITA-1/5-{36.3100000, 422.7800000}	38.753
142	CRIMES_PERCAPITA-1/5-{87.9500000, 1002.9200000}	32.753
127	ASSAULTS-1/5-{376.0000000, 2038.0000000}	28.753
112	VIOLENT_CRIMES-1/5-{960.0000000, 4650.0000000}	20.753
162	ROBBERIES_PERCAPITA-1/5-{36.3100000, 429.3300000}	17.753
128	ASSAULTS-2/5-{2038.0000000, 3369.0000000}	13.547
158	ASSAULTS_PERCAPITA-2/5-{422.7800000, 584.7700000}	12.547
132	ROBBERIES-1/5-{304.0000000, 1976.0000000}	11.753
155	RAPES_PERCAPITA-4/5-{78.0400000, 99.9600000}	11.547
166	ROBBERIES_PERCAPITA-5/5-{987.8000000, 2337.520000	10.547
154	RAPES_PERCAPITA-3/5-{64.3800000, 78.0400000}	9.547
120	HOMICIDES-4/5-{165.0000000, 298.0000000}	8.753

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
161	ASSAULTS_PERCAPITA-5/5-{1099.0700000, 2368.2200000}	-34.453
131	ASSAULTS-5/5-{7735.0000000, 71030.0000000}	-21.453
160	ASSAULTS_PERCAPITA-4/5-{789.2800000, 1099.0700000}	-19.453
134	ROBBERIES-3/5-{3114.0000000, 4447.0000000}	-19.453
146	CRIMES_PERCAPITA-5/5-{2078.7500000, 4352.8300000}	-18.453
152	RAPES_PERCAPITA-1/5-{13.4400000, 46.9800000}	-17.247
130	ASSAULTS-4/5-{4704.0000000, 7735.0000000}	-14.453
165	ROBBERIES_PERCAPITA-4/5-{710.9900000, 987.8000000}	-12.453
116	VIOLENT_CRIMES-5/5-{16097.0000000, 174542.0000000}	-12.453
119	HOMICIDES-3/5-{111.0000000, 165.0000000}	-11.658
144	CRIMES_PERCAPITA-3/5-{1268.9800000, 1618.9400000}	-11.453
114	VIOLENT_CRIMES-3/5-{7181.0000000, 10160.0000000}	-11.453

ВКЛЮЧИТЬ фильтр по фактору | ВЫКЛЮЧИТЬ фильтр по фактору | ВКЛЮЧИТЬ фильтр по фактору | ВЫКЛЮЧИТЬ фильтр по фактору

Помощь | Abs | Prc1 | Prc2 | Inf1 | Inf2 | Inf3 | Inf4 | Inf5 | Inf6 | Inf7 | Нейрон | SWOT-диаграмма | Интегральная когнитивная карта

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса
1	REPORT_YEAR-1/5-{1975.0000000, 1981.0000000}
2	REPORT_YEAR-2/5-{1981.0000000, 1988.0000000}
3	REPORT_YEAR-3/5-{1988.0000000, 1994.0000000}
4	REPORT_YEAR-4/5-{1994.0000000, 2003.0000000}
5	REPORT_YEAR-5/5-{2003.0000000, 2014.0000000}

SWOT-анализ класса: 1 "REPORT_YEAR-1/5-{1975.0000000, 1981.0000000}" в модели: 6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
35	AGENCY_CODE-35/53-OHCLP00	0.634
63	AGENCY_JURISDICTION-10/53-Cleveland, OH	0.634
133	ROBBERIES-2/5-{1976.0000000, 3114.0000000}	0.547
47	AGENCY_CODE-47/53-TXDPD00	0.429
65	AGENCY_JURISDICTION-12/53-Dallas, TX	0.429
9	AGENCY_CODE-9/53-FL01300	0.304
12	AGENCY_CODE-12/53-FL02902	0.304
80	AGENCY_JURISDICTION-27/53-Miami-Dade County, FL	0.304
102	AGENCY_JURISDICTION-49/53-Tampa, FL	0.304
25	AGENCY_CODE-25/53-MOKPD00	0.224
74	AGENCY_JURISDICTION-21/53-Kansas City, MO	0.224

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
84	AGENCY_JURISDICTION-31/53-Montgomery County, MD	-0.205
69	AGENCY_JURISDICTION-16/53-Fairfax County, VA	-0.205
68	AGENCY_JURISDICTION-15/53-El Paso, TX	-0.205
52	AGENCY_CODE-52/53-VA12800	-0.205
51	AGENCY_CODE-51/53-VA02901	-0.205
44	AGENCY_CODE-44/53-TX07102	-0.205
39	AGENCY_CODE-39/53-OR02602	-0.205
28	AGENCY_CODE-28/53-NC09201	-0.205
20	AGENCY_CODE-20/53-MD01604	-0.205
77	AGENCY_JURISDICTION-24/53-Los Angeles County, CA	-0.107
1	AGENCY_CODE-1/53-CA01900	-0.107

ВКЛЮЧИТЬ фильтр по фактору | ВЫКЛЮЧИТЬ фильтр по фактору | ВКЛЮЧИТЬ фильтр по фактору | ВЫКЛЮЧИТЬ фильтр по фактору

Помощь | Abs | Prc1 | Prc2 | Inf1 | Inf2 | Inf3 | Inf4 | Inf5 | Inf6 | Inf7 | Нейрон | SWOT-диаграмма | Интегральная когнитивная карта

Рисунок 21. Пример SWOT-матрицы в модели INF3

На рисунке 22 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF3.

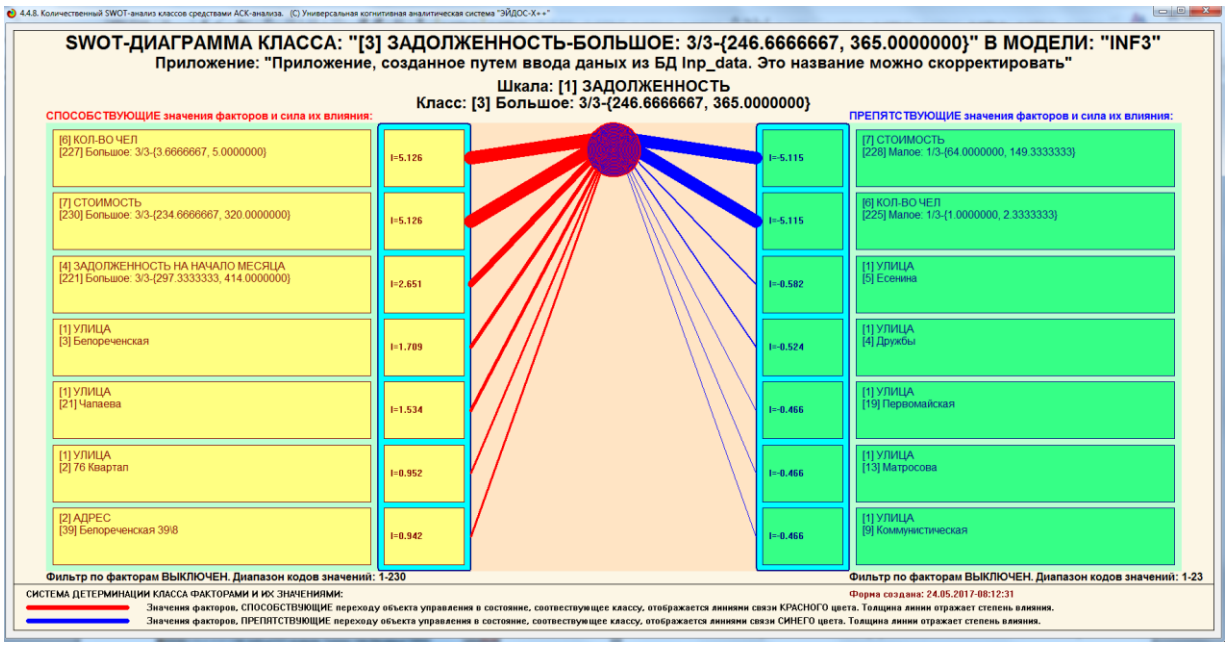


Рисунок 22. Пример SWOT-матрицы в модели INF3

2.4. Нелокальные нейроны и нейронная сеть

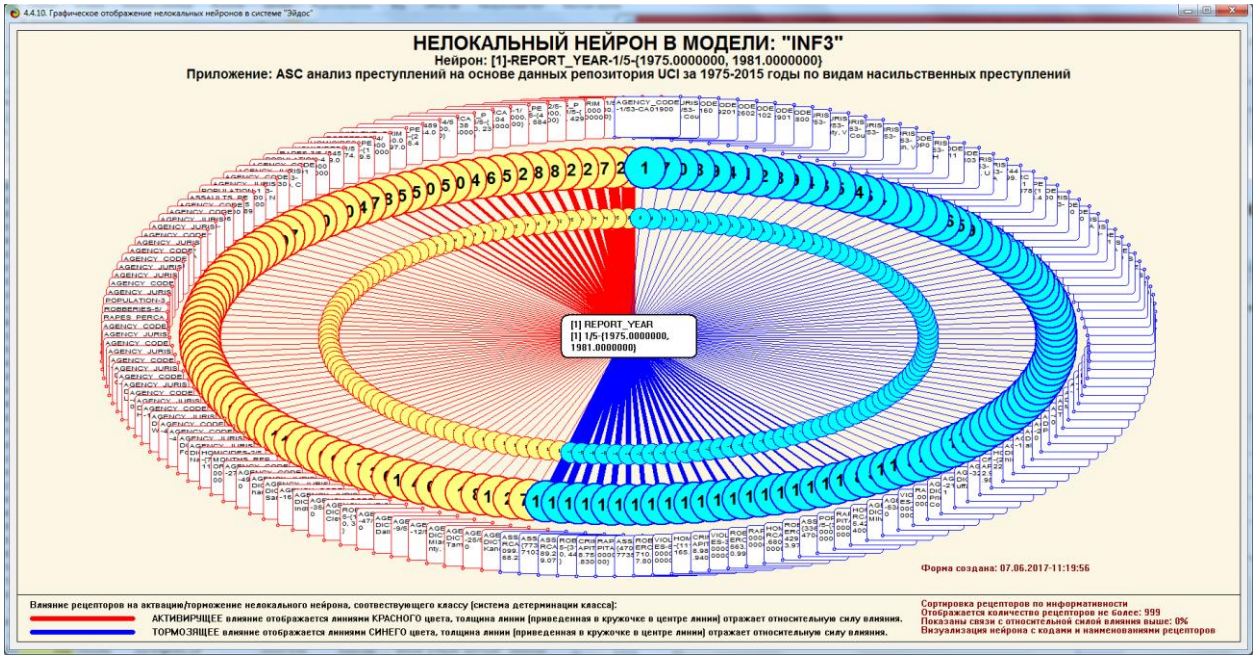


Рисунок 23. Пример графического отображения нелокальных нейронов

На рис23 надписи на рецепторах не читабельны из-за их большого количества, поэтому на рисунке 23б применены изображение того же нейрона с меньшим числом рецепторов.

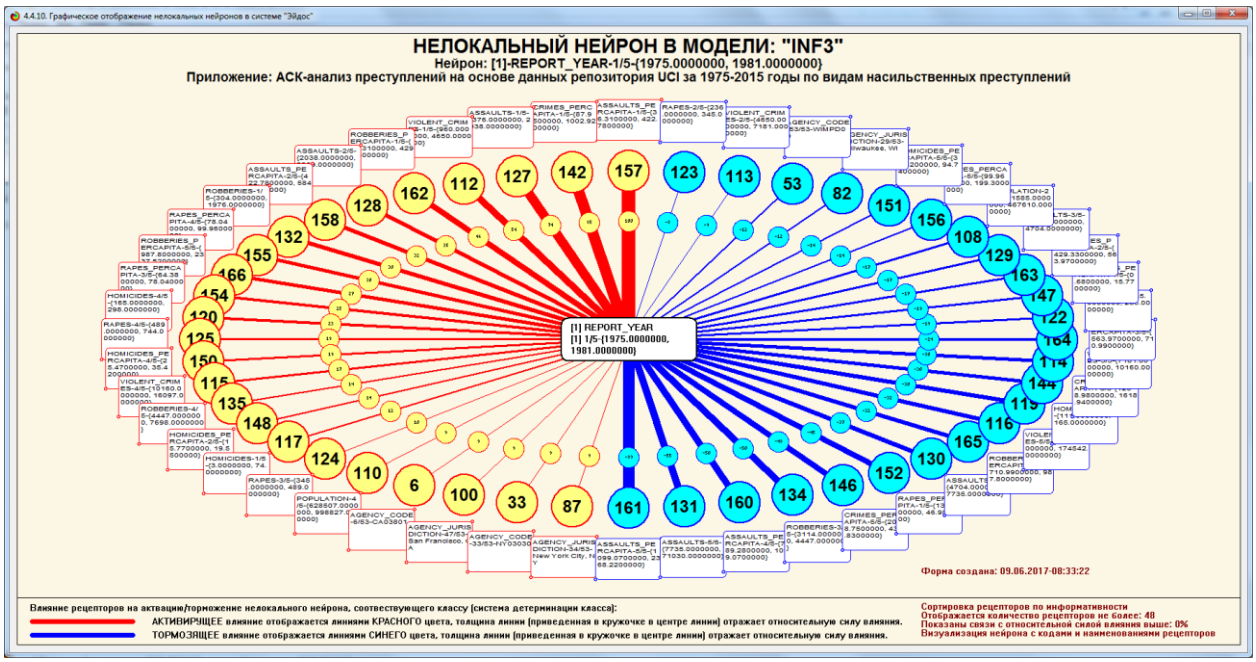


Рисунок 23б. Пример графического отображения нелокальных нейронов

В качестве примера приведено графическое отображение нелокальных нейронов в модели INF3 на основе признака «Отчетный-год». Каждому классу соответствует нейрон, совокупность которых образует нелокальную нейронную сеть [9].

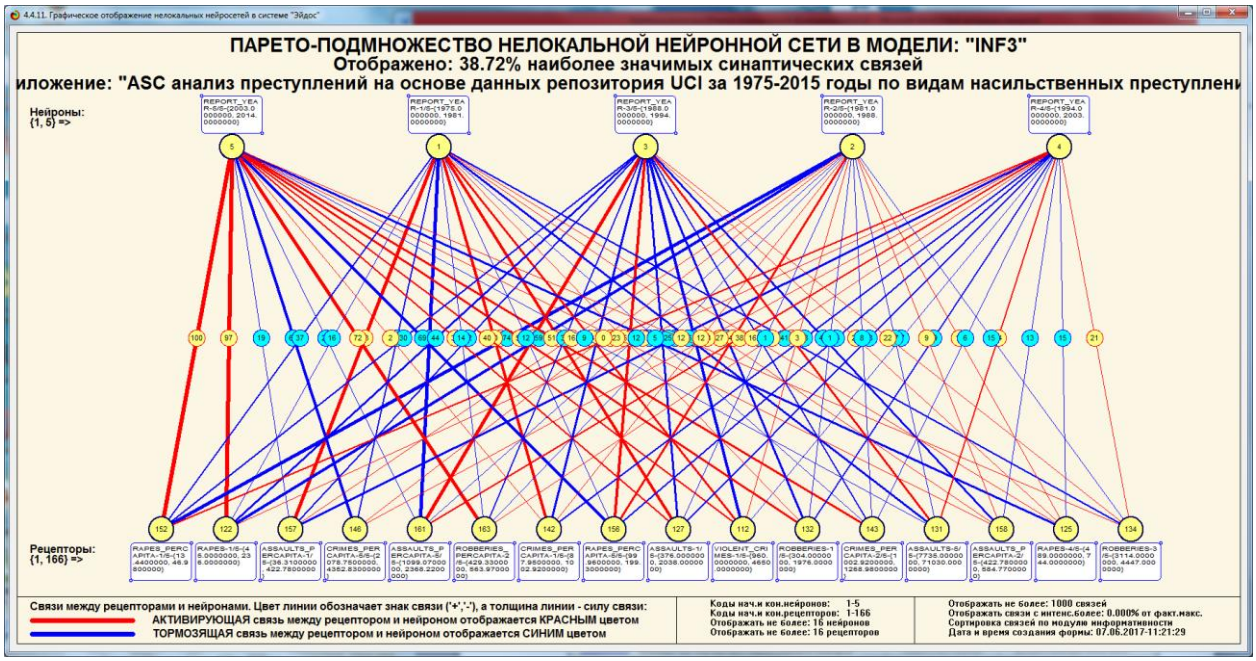


Рисунок 24. Пример графического отображения парето-подмножества нелокальной нейронной сети

На рисунке 24 в графическом отображения парето-подмножества нелокальной нейронной сети в модели INF3 у классов слева сильнее связи между значениями факторов и результатом [9].

2.5. Кластерный и конструктивный анализ классов

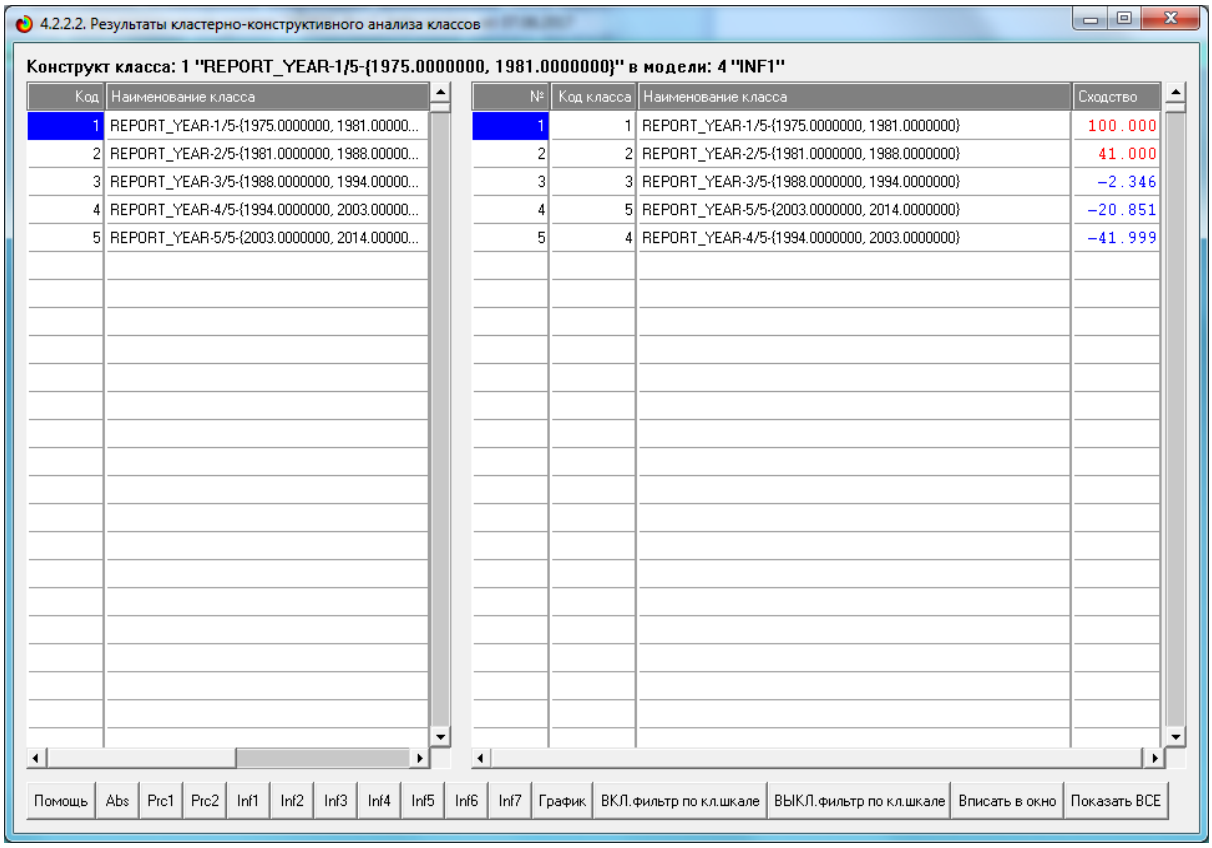


Рисунок 25. Результаты кластерного и конструктивного анализа классов

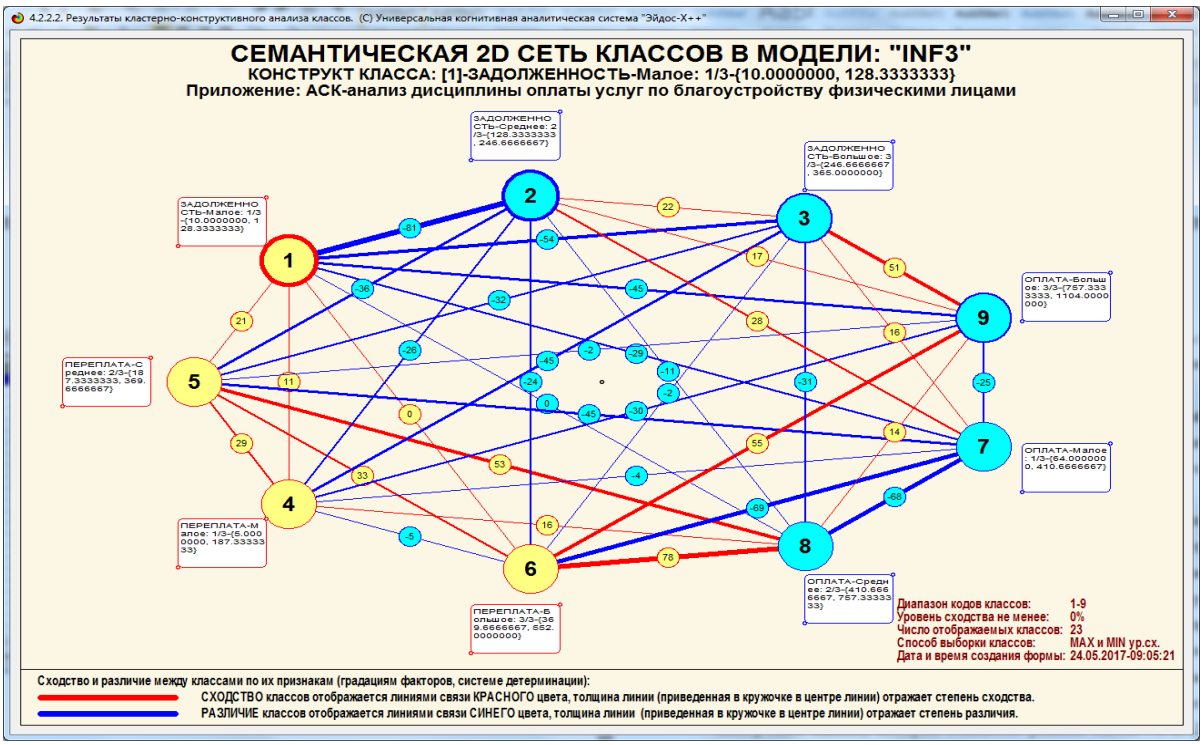


Рисунок 26. Графические результаты кластерного и конструктивного анализа классов

ЗАКЛЮЧЕНИЕ

На сегодняшний день все больше проявилась потребность в системах, не только представляющих информацию, но и выполняющих некоторый ее предварительный анализ, способных давать некоторые советы и рекомендации, осуществлять прогнозирование развития ситуаций, отбирать наиболее перспективные альтернативы решений, взять на себя значительную часть рутинных операций, а также функции предварительного анализа и оценок.

В данной курсовой работе была разработана системно-когнитивной модель классификации преступлений за 1975-2015г. с помощью программного инструментария системы искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении оказались модели INF5, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,911, что заметно выше, чем достоверность экспертных оценок,

которая считается равной около 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием. Также обращает на себя внимание, что статистические модели в данном приложении дают примерно на 17% 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.

Обработка данных, их систематизация и визуализация позволили сделать прогнозы и обоснованные выводы по дисциплине оплаты услуг. Таким образом, повышены качество и оперативность принятия решений, а также эффективность управления знаниями.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

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

2. Луценко Е.В. 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.

3. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.

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

5. 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.

6. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: http://ej.kubagro.ru/2013/08/pdf/58.pdf, 1,562 у.п.л.

7. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 164 – 188. – IDA [article ID]: 0911307012. – Режим доступа: http://ej.kubagro.ru/2013/07/pdf/12.pdf, 1,562 у.п.л.

8. Луценко Е.В. Нечеткое мультиклассовое обобщение классической F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе «Эйдос» / Луценко Е.В. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №09(123). – Режим доступа: http://ej.kubagro.ru/2016/09/pdf/01.pdf, 1,813 у.п.л. – IDA [article ID]: 1231609001. http://dx.doi.org/10.21515/1990-4665-123-001

9. Луценко Е.В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №01(001). С. 79 – 91. – IDA [article ID]: 0010301011. – Режим доступа: http://ej.kubagro.ru/2003/01/pdf/11.pdf, 0,812 у.п.л.