

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к курсовой работе

по дисциплине: ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ
на тему:

Исследование активности мобильных телефонов в городе с
использованием АСК –анализа и системы Эйдос на основе данных
репозитория UC1

выполнил студент группы ИТ1421 Нагорный Валерий Александрович

Допущен 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**

Студенту: ИТ1421 группы 3 курса
Факультета прикладной информатики
Специальности: 09.03.02 Информационные системы и технологии
(шифр)

Нагорному Валерию Александровичу
(Ф.И.О.)

Тема проекта: **Исследование активности мобильных телефонов в городе с использованием АСК –анализа и системы Эйдос на основе данных репозитория UC1**

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

Объем работы:

а) пояснительная записка к работе _____ листа формата А4 б)
графическая часть _____ лист формата А4

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

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

Срок защиты: “ ” _____ 2017 г.

Дата выдачи задания: “ ” _____ 2017 г.

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

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

(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ “ ” _____ 2017 г.

(подпись, дата)

Краснодар
2017 г.

СОДЕРЖАНИЕ

Введение	4
1. Синтез и верификация моделей.....	5
1.1. Описание решения	5
1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel	5
1.3. Ввод выборки в систему Aidos-X.....	7
1.4. Синтез и верификация статистических и интеллектуальных моделей.....	11
1.5. Виды моделей системы «Aidos».....	12
1.5. Результаты верификации моделей	14
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.	16
2.1. Решение задачи идентификации	16
2.2. Выбор модели и пакетное распознавание	16
2.3 Визуализация когнитивных функций	19
2.4 SWOT и PEST матрицы и диаграммы	21
2.5 Нелокальные нейронные сети и нейтроны.....	23
2.6 Кластерный и конструктивный анализ.....	25
Заключение	26
Список используемых источников.....	27

ВВЕДЕНИЕ

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи по определению использования за неделю мобильных телефонов в Испанских провинциях.

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

В данной курсовой работе использована база данных «Mobile phone activity in a city» из банка исходных данных по задачам искусственного интеллекта – репозитория Kaggle.

Для решения задачи используем стандартные возможности Microsoft Office Word и Excel, 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 из csv-формата в промежуточные файлы MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Aidos".
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

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

С электронного ресурса kaggle.com возьмем базу данных: mobile-phone-activity.csv по ссылке: <https://www.kaggle.com/marcodena/mobile-phone-activity>

В этой базе данных хранится подробную информацию о звонках за неделю во всех провинциях Испании. Она содержит:

- Название провинции.
- Общее количество часов звонков.
- Количество часов телефонного общения мужчин.
- Количество часов телефонного общения женщин.
- Количество часов телефонного общения людей не состоящих в браке.
- Количество часов телефонного общения людей состоящих в браке.

Это нам позволит понять, какого пола и в каком семейном положении люди чаще общаются по телефону.

Преобразуем файлы в формат *.xlsx, выполнив объединение двух таблиц по совпадению полей «Название зоны». В итоге получается следующая таблица для работы:

Таблица 1 – Фрагмент обучающей выборки

1	Провинция	Провинция	Общее	Мужчины	Женщины	Не состоящие в браке	Женаты/замужем
2	Alessandria	Alessandria	427229	204983	222246	158222	198253
3	Asti	Asti	217573	105370	112203	82119	102538
4	Biella	Biella	182192	86821	95371	69188	82431
5	Cuneo	Cuneo	586378	287786	298592	231452	276983
6	Novara	Novara	365559	176361	189198	143302	170741
7	Torino	Torino	2247780	1081298	1166482	881971	1041312
8	Verbano-Cusio-Ossola	Verbano-Cusio-Ossola	160264	77219	83045	62913	71909
9	Vercelli	Vercelli	176941	85150	91791	67575	80931
10	Valle d'Aosta/Vallée d'Aoste	Valle d'Aosta/Vallée d'Aoste	126806	61861	64945	54474	52935
11	Bergamo	Bergamo	1086277	536201	550076	456200	508965
12	Brescia	Brescia	1238044	607910	630134	510780	580195
13	Como	Como	586735	286101	300634	240306	275282
14	Cremona	Cremona	357623	174506	183117	142230	167883
15	Lecco	Lecco	336310	164901	171409	138804	159074
16	Lodi	Lodi	223755	109771	113984	90994	105209
17	Mantova	Mantova	408336	198711	209625	161084	193167
18	Milano	Milano	3038420	1455050	1583370	1257677	1375762
19	Monza e della Brianza	Monza e della Brianza	840129	409618	430511	343808	400253
20	Pavia	Pavia	535822	258625	277197	208864	246303
21	Sondrio	Sondrio	180814	88357	92457	75636	82299
22	Varese	Varese	871886	422752	449134	348107	412094
23	Bolzano/Bozen	Bolzano/Bozen	504643	248322	256321	249598	200884
24	Trento	Trento	524832	255731	269101	227249	233360
25	Belluno	Belluno	210001	101135	108866	87454	92046
26	Padova	Padova	921361	447352	474009	378556	434387
27	Rovigo	Rovigo	242349	117281	125068	93602	115688
28	Treviso	Treviso	876790	429287	447503	362088	414444
29	Venezia	Venezia	846962	408590	438372	335842	401070
30	Verona	Verona	900542	439225	461317	375843	415627
31	Vicenza	Vicenza	859205	421812	437393	360744	400998
32	Gorizia	Gorizia	140143	67830	72313	53707	63584
33	Pordenone	Pordenone	310811	151951	158860	125627	145531
34	Trieste	Trieste	232601	109555	123046	87281	101043
35	Udine	Udine	535430	258689	276741	213860	242510
36	Genova	Genova	855834	402267	453567	327146	382945
37	Imperia	Imperia	214502	102095	112407	82015	94732
38	La Spezia	La Spezia	219330	104390	114940	79519	103153
39	Savona	Savona	281028	133220	147808	103108	128615
40	Bologna	Bologna	976243	466757	509486	403096	429283
41	Ferrara	Ferrara	353481	168345	185136	133658	163299
42	Forli-Cesena	Forli-Cesena	390738	189911	200827	160507	179071
43	Modena	Modena	685777	333069	352708	273877	319289
44	Parma	Parma	427434	206072	221362	171671	195260
45	Piacenza	Piacenza	284616	137838	146778	112005	131644
46	Ravenna	Ravenna	384761	185521	199240	155806	171879
47	Reggio nell'Emilia	Reggio nell'Emilia	517316	252659	264657	212174	239445
48	Rimini	Rimini	321769	154991	166778	133847	146221
49	Arezzo	Arezzo	343676	166507	177169	130916	167941
50	Firenze	Firenze	973145	463574	509571	393196	447491
51	Grosseto	Grosseto	220564	105585	114979	82500	103512
52	Livorno	Livorno	335247	160393	174854	124339	158833
53	Lucca	Lucca	388327	186183	202144	150013	181190
54	Massa-Carrara	Massa-Carrara	199650	95754	103896	76470	92149

56	Pistoia	Pistoia	287866	138054	149812	109805	138391
57	Prato	Prato	245916	119088	126828	97999	118587
58	Siena	Siena	266621	127633	138988	103331	126476
59	Perugia	Perugia	655844	315217	340627	254748	316040
60	Terni	Terni	228424	108840	119584	83979	110751
61	Ancona	Ancona	473865	228105	245760	188813	223470
62	Ascoli Piceno	Ascoli Piceno	210407	101728	108679	83600	100671
63	Fermo	Fermo	174857	84840	90017	69314	84395
64	Macerata	Macerata	319607	154770	164837	126599	152917
65	Pesaro e Urbino	Pesaro e Urbino	362583	176501	186082	148981	168590
66	Frosinone	Frosinone	492661	240388	252273	194210	242020
67	Latina	Latina	544732	266395	278337	224015	259199
68	Rieti	Rieti	155164	75762	79402	60852	72955
69	Roma	Roma	3997465	1902725	2094740	1701651	1743838
70	Viterbo	Viterbo	312864	151880	160984	121406	148403
71	Chieti	Chieti	387956	188479	199477	149827	192278
72	L'Aquila	L'Aquila	298343	145340	153003	119913	141319
73	Pescara	Pescara	314661	151154	163507	125681	149997
74	Teramo	Teramo	306349	149424	156925	123469	147479
75	Campobasso	Campobasso	226419	110207	116212	89814	110740
76	Isernia	Isernia	87241	42587	44654	34623	42296
77	Avellino	Avellino	429157	209491	219666	177979	207230
78	Benevento	Benevento	284900	138308	146592	117358	137440
79	Caserta	Caserta	904921	441070	463851	397738	429519
80	Napoli	Napoli	3054956	1475144	1579812	1371131	1399146
81	Salerno	Salerno	1092876	531769	561107	467450	514615
82	Bari	Bari	1247303	607155	640148	516003	615881
83	Barletta-Andria-Trani	Barletta-Andria-Trani	391723	193354	198369	166387	194905
84	Brindisi	Brindisi	400801	192599	208202	168178	192203
85	Foggia	Foggia	626072	305162	320910	267853	300317
86	Lecce	Lecce	802018	382317	419701	334226	386757

Классификационная шкала следующая: Провинции. Остальные – описательные.

1.3. Ввод выборки в систему Aidos-X

Теперь, когда мы имеем обучающую выборку в формате *.xlsx, можно импортировать ее в систему Aidos-X.

Скопируем выборку в папку d:\Aidos-X\AID_DATA\Inp_data\ и переименуем ее в Inp_data. Затем, запустив систему, воспользуемся универсальным программным интерфейсом для ввода данных из внешних баз табличного вида (режим 2.3.2.2). Настройка ввода будет такая, как на рисунке 1 – ввод обучающей выборки.

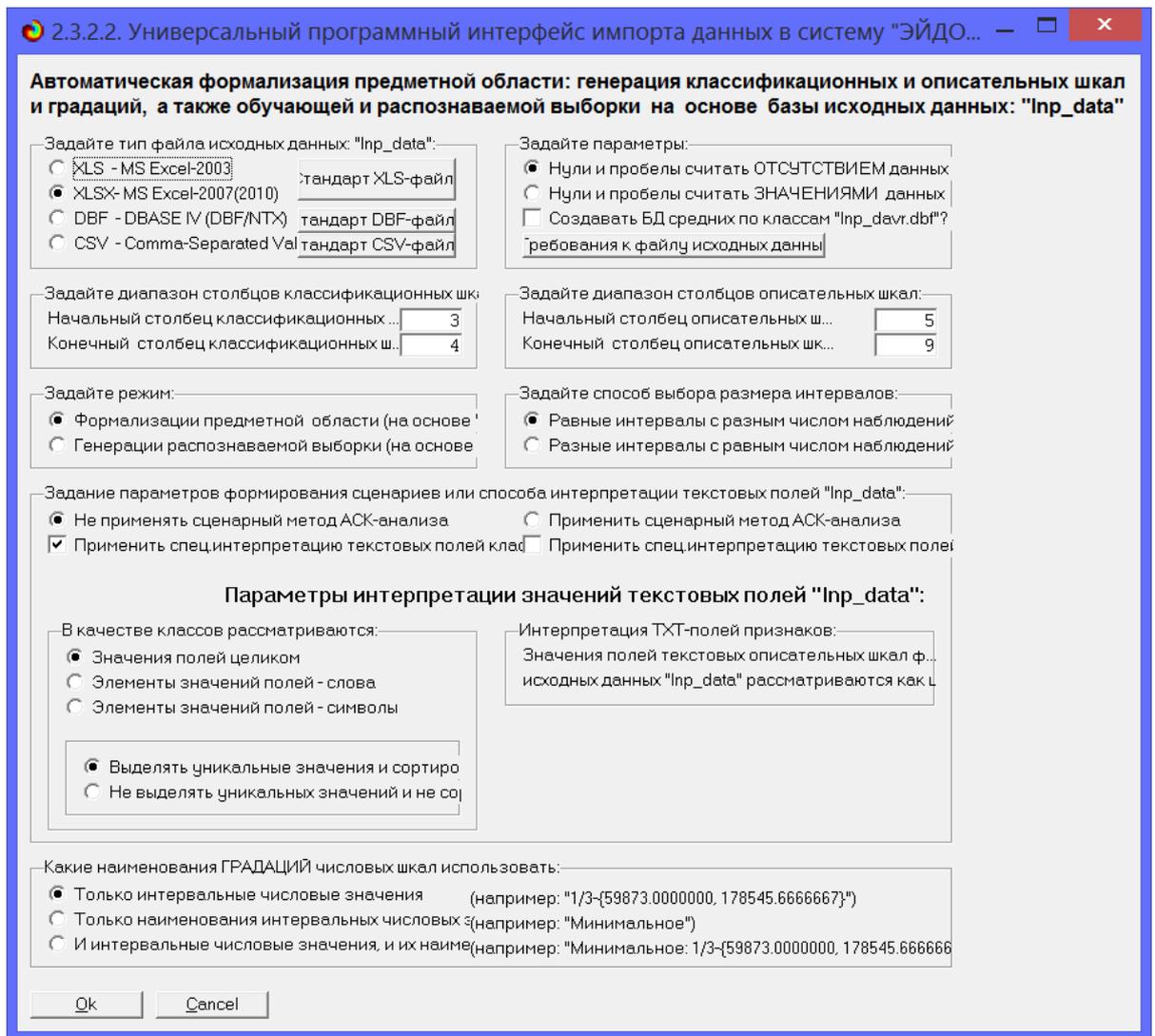


Рисунок 1 – Ввод обучающей выборки

Так как предварительно числовые шкалы были разбиты на интервалы так, как это необходимо для исследования, перерасчет шкал после ввода выборки производить не надо.

После импорта автоматически формируются классификационные и описательные шкалы, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Этим самым полностью автоматизировано выполняется этап «Формализация предметной области».

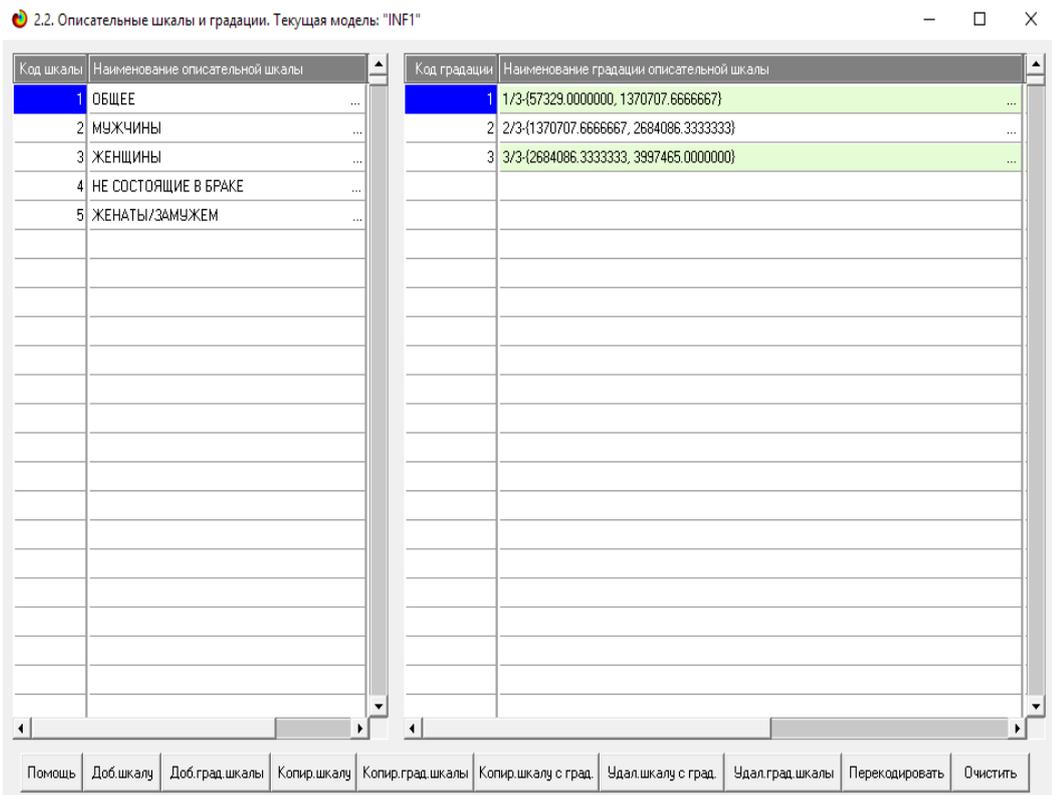


Рисунок 3 – Описательные шкалы

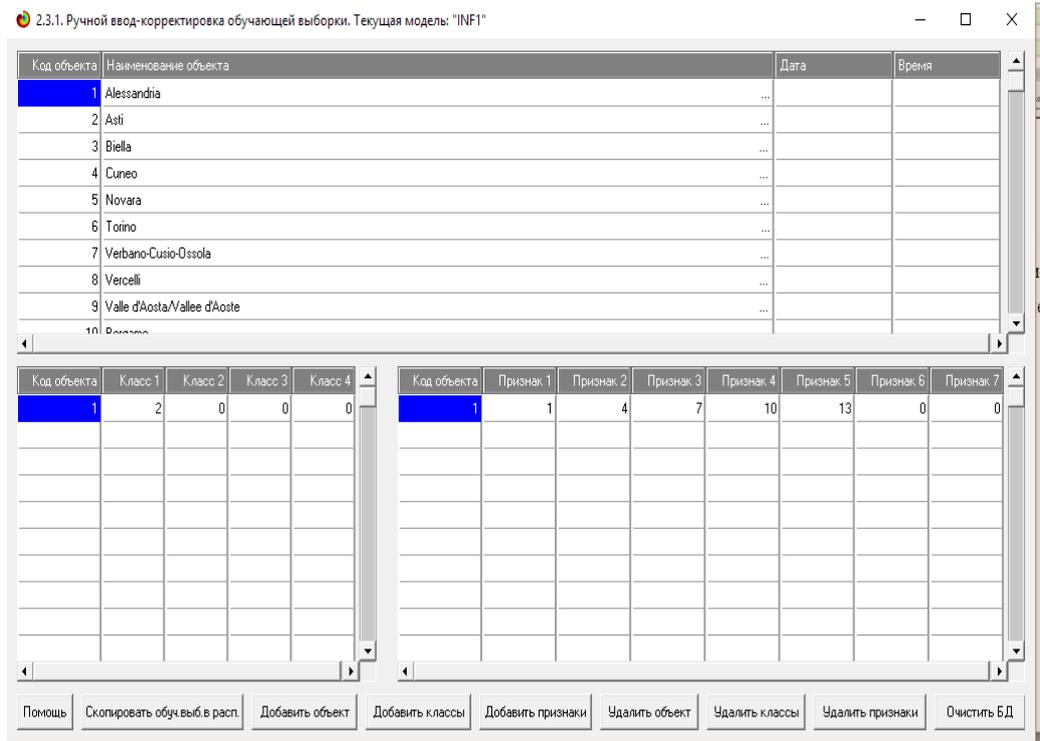


Рисунок 4 – Ручной ввод-корректировка выборки

Тем самым создаются все необходимые и достаточные предпосылки для выявления связей между значениями факторов и результатами их совместного системного воздействия.

1.4. Синтез и верификация статистических и интеллектуальных моделей

Далее запускаем режим 3.5 – Синтез и верификация моделей. Настраиваем так, как показано на рисунке 5.

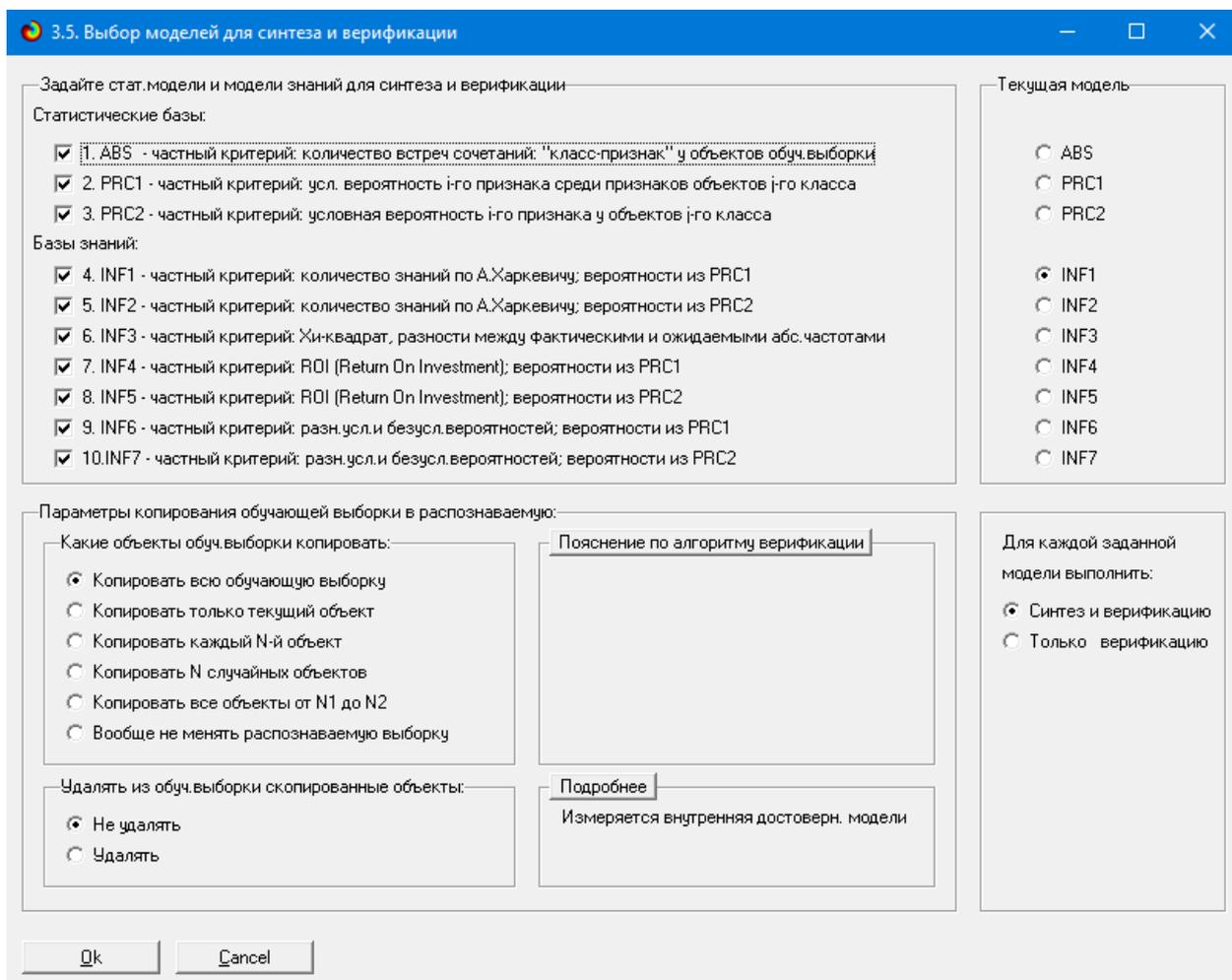


Рисунок 5 – Настройка синтеза и верификации моделей

После построения моделей нажимаем «ОК» на форме, которая показана на рисунке 6.

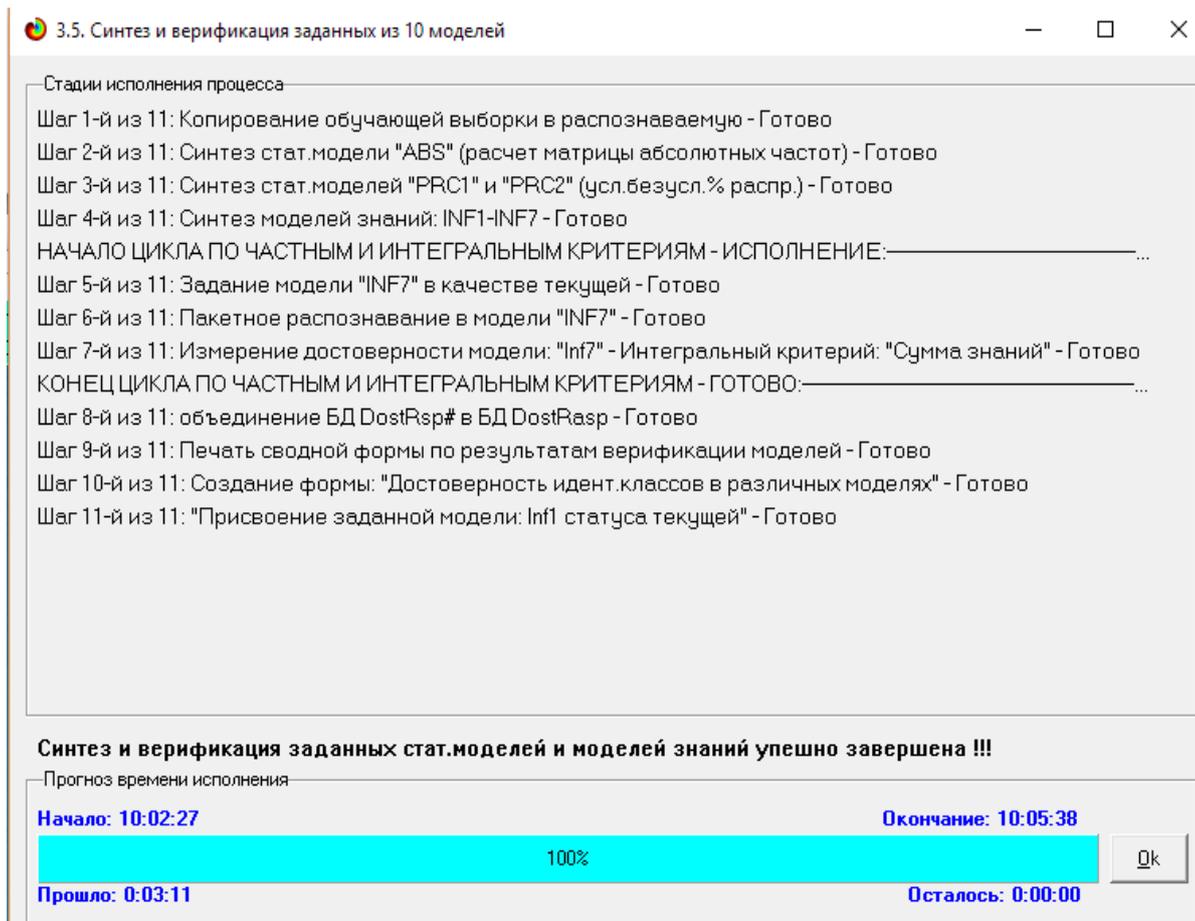


Рисунок 6 – Завершение процесса синтеза и верификации моделей

После этого можно перейти непосредственно к выбору наиболее достоверной модели.

1.5. Виды моделей системы «Aidos»

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).

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

Код признака	Наименование описательной шкалы и градации	1. ПРОВИНЦИЯ ABRIGENTO	2. ПРОВИНЦИЯ ALESSAND...	3. ПРОВИНЦИЯ ANCONA	4. ПРОВИНЦИЯ AREZZO	5. ПРОВИНЦИЯ ASCOLI PICENO	6. ПРОВИНЦИЯ ASTI	7. ПРОВИНЦИЯ AVELLINO	8. ПРОВИНЦИЯ BARI	9. ПРОВИНЦИЯ BARLETTA ANDRIA TRANI	10. ПРОВИНЦИЯ BELLUNO	11. ПРОВИНЦИЯ BENEVENTO	12. ПРОВИНЦИЯ BERGAMO
1	ОБЩЕЕ-1/3-(57329.0000000, 1370707.6666667) ...	1	1	1	1	1	1	1	1	1	1	1	1
2	ОБЩЕЕ-2/3-(1370707.6666667, 2684086.3333333) ...												
3	ОБЩЕЕ-3/3-(2684086.3333333, 3997465.0000000) ...												
4	МУЖЧИНЫ-1/3-(28300.0000000, 653108.3333333) ...	1	1	1	1	1	1	1	1	1	1	1	1
5	МУЖЧИНЫ-2/3-(653108.3333333, 1277916.6666667) ...												
6	МУЖЧИНЫ-3/3-(1277916.6666667, 1902725.0000000) ...												
7	ЖЕНЩИНЫ-1/3-(29029.0000000, 717599.3333333) ...	1	1	1	1	1	1	1	1	1	1	1	1
8	ЖЕНЩИНЫ-2/3-(717599.3333333, 1406169.6666667) ...												
9	ЖЕНЩИНЫ-3/3-(1406169.6666667, 2094740.0000000) ...												
10	НЕ СОСТОЯЩИЕ В БРАКЕ-1/3-(27160.0000000, 585...	1	1	1	1	1	1	1	1	1	1	1	1
11	НЕ СОСТОЯЩИЕ В БРАКЕ-2/3-(585323.6666667, 11...												
12	НЕ СОСТОЯЩИЕ В БРАКЕ-3/3-(1143487.3333333, 1...												
13	ЖЕНАТЫ/ЗАМУЖЕМ-1/3-(24615.0000000, 597689.3...	1	1	1	1	1	1	1		1	1	1	1
14	ЖЕНАТЫ/ЗАМУЖЕМ-2/3-(597689.3333333, 117076...								1				
15	ЖЕНАТЫ/ЗАМУЖЕМ-3/3-(1170763.6666667, 17438...												
	Сумма числа признаков	5	5	5	5	5	5	5	5	5	5	5	5
	Среднее	0	0	0	0	0	0	0	0	0	0	0	0
	Среднеквадратичное отклонение	1	1	1	1	1	1	1	1	1	1	1	1
	Сумма числа объектов обучающей выборки	1	1	1	1	1	1	1	1	1	1	1	1

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

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

Код признака	Наименование описательной шкалы и градации	1. ПРОВИНЦИЯ ABRIGENTO	2. ПРОВИНЦИЯ ALESSAND...	3. ПРОВИНЦИЯ ANCONA	4. ПРОВИНЦИЯ AREZZO	5. ПРОВИНЦИЯ ASCOLI PICENO	6. ПРОВИНЦИЯ ASTI	7. ПРОВИНЦИЯ AVELLINO	8. ПРОВИНЦИЯ BARI	9. ПРОВИНЦИЯ BARLETTA ANDRIA TRANI	10. ПРОВИНЦИЯ BELLUNO	11. ПРОВИНЦИЯ BENEVENTO	12. ПРОВИНЦИЯ BERGAMO
1	ОБЩЕЕ-1/3-(57329.0000000, 1370707.6666667) ...	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040
2	ОБЩЕЕ-2/3-(1370707.6666667, 2684086.3333333) ...												
3	ОБЩЕЕ-3/3-(2684086.3333333, 3997465.0000000) ...												
4	МУЖЧИНЫ-1/3-(28300.0000000, 653108.3333333) ...	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040
5	МУЖЧИНЫ-2/3-(653108.3333333, 1277916.6666667) ...												
6	МУЖЧИНЫ-3/3-(1277916.6666667, 1902725.0000000) ...												
7	ЖЕНЩИНЫ-1/3-(29029.0000000, 717599.3333333) ...	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040
8	ЖЕНЩИНЫ-2/3-(717599.3333333, 1406169.6666667) ...												
9	ЖЕНЩИНЫ-3/3-(1406169.6666667, 2094740.0000000) ...												
10	НЕ СОСТОЯЩИЕ В БРАКЕ-1/3-(27160.0000000, 585...	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040	0.040
11	НЕ СОСТОЯЩИЕ В БРАКЕ-2/3-(585323.6666667, 11...												
12	НЕ СОСТОЯЩИЕ В БРАКЕ-3/3-(1143487.3333333, 1...												
13	ЖЕНАТЫ/ЗАМУЖЕМ-1/3-(24615.0000000, 597...	0.050	0.050	0.050	0.050	0.050	0.050	0.050		0.050	0.050	0.050	0.050
14	ЖЕНАТЫ/ЗАМУЖЕМ-2/3-(597689.3333333, 11...								4.307				
15	ЖЕНАТЫ/ЗАМУЖЕМ-3/3-(1170763.6666667, 17438...												
	Сумма	0.209	0.209	0.209	0.209	0.209	0.209	0.209	4.466	0.209	0.209	0.209	0.209
	Среднее	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.298	0.014	0.014	0.014	0.014
	Среднеквадратичное отклонение	0.021	0.021	0.021	0.021	0.021	0.021	0.021	1.109	0.021	0.021	0.021	0.021

Таблица 3 – Матрица информативностей (модель INF1) в битах

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

Код признака	Наименование описательной шкалы и градации	1. ПРОВИНЦИЯ ABRIGENTO	2. ПРОВИНЦИЯ ALESSAND...	3. ПРОВИНЦИЯ ANCONA	4. ПРОВИНЦИЯ AREZZO	5. ПРОВИНЦИЯ ASCOLI PICENO	6. ПРОВИНЦИЯ ASTI	7. ПРОВИНЦИЯ AVELLINO	8. ПРОВИНЦИЯ BARI	9. ПРОВИНЦИЯ BARLETTA ANDRIA TRANI	10. ПРОВИНЦИЯ BELLUNO	11. ПРОВИНЦИЯ BENEVENTO	12. ПРОВИНЦИЯ BERGAMO
1	ОБЩЕЕ-1/3-(57329.0000000, 1370707.6666667) ...	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
2	ОБЩЕЕ-2/3-(1370707.6666667, 2684086.3333333) ...	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
3	ОБЩЕЕ-3/3-(2684086.3333333, 3997465.0000000) ...	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027
4	МУЖЧИНЫ-1/3-(28300.0000000, 653108.3333333) ...	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
5	МУЖЧИНЫ-2/3-(653108.3333333, 1277916.6666667) ...	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
6	МУЖЧИНЫ-3/3-(1277916.6666667, 1902725.0000000) ...	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027
7	ЖЕНЩИНЫ-1/3-(29029.0000000, 717599.3333333) ...	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
8	ЖЕНЩИНЫ-2/3-(717599.3333333, 1406169.6666667) ...	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
9	ЖЕНЩИНЫ-3/3-(1406169.6666667, 2094740.0000000) ...	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027
10	НЕ СОСТОЯЩИЕ В БРАКЕ-1/3-(27160.0000000, 585...	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
11	НЕ СОСТОЯЩИЕ В БРАКЕ-2/3-(585323.6666667, 11...	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
12	НЕ СОСТОЯЩИЕ В БРАКЕ-3/3-(1143487.3333333, 1...	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027
13	ЖЕНАТЫ/ЗАМУЖЕМ-1/3-(24615.0000000, 597...	0.045	0.045	0.045	0.045	0.045	0.045	0.045	-0.955	0.045	0.045	0.045	0.045
14	ЖЕНАТЫ/ЗАМУЖЕМ-2/3-(597689.3333333, 11...	-0.018	-0.018	-0.018	-0.018	-0.018	-0.018	-0.018	0.982	-0.018	-0.018	-0.018	-0.018
15	ЖЕНАТЫ/ЗАМУЖЕМ-3/3-(1170763.6666667, 17438...	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027	-0.027
	Сумма												
	Среднее												
	Среднеквадратичное отклонение	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.367	0.029	0.029	0.029	0.029

Таблица 4 – Матрица знаний (модель INF3)

1.5. Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями, приведены на рисунке 7.

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-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризенбергена	Сумма мод.уровней сходства истинно-пол.
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс.частот с обр...	30	30	600	270		0.100	1.000	0.182	30.00
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс.частот по признак...	30	30	530	340		0.081	1.000	0.150	19.80
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	30	30	600	270		0.100	1.000	0.182	30.00
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	30	30	530	340		0.081	1.000	0.150	30.00
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	30	30	600	270		0.100	1.000	0.182	29.90
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	30	30	530	340		0.081	1.000	0.150	19.80
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	30	30	600	270		0.100	1.000	0.182	29.10
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	20.70
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	30	30	598	272		0.099	1.000	0.181	29.40
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	17.20
6. INF3 - частный критерий: Хинквдрат, разности между факти...	Семантический резонанс зна...	30	30	600	270		0.100	1.000	0.182	29.50
6. INF3 - частный критерий: Хинквдрат, разности между факти...	Сумма знаний	30	30	600	270		0.100	1.000	0.182	20.50
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	30	30	621	249		0.108	1.000	0.194	27.40
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	20.10
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	30	30	611	259		0.104	1.000	0.188	27.80
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	13.30
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	30	30	600	270		0.100	1.000	0.182	29.70
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	24.20
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	30	30	600	270		0.100	1.000	0.182	29.90
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	30	30	530	340		0.081	1.000	0.150	18.90

Помощь

Рисунок 7 – Результаты верификации моделей

Наиболее достоверной в данном приложении оказалась модель INF5 при интегральном критерии «Семантический резонанс знаний», если учитывать F-меру Ван Ризенбергена. Точность модели составляет 0,019, то есть уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 19%.

Также статистические модели обычно дают более низкую средневзвешенную достоверность идентификации и неидентификации, чем модели знаний, и практически никогда – более высокую. Соответственно, это оправдывает применение интеллектуальных технологий и моделей знаний. Посмотрим, каковы частные распределения уровней сходства и различия для верно и ошибочно неидентифицированных и идентифицированных ситуаций

в наиболее достоверной модели INF5 (рисунок 8). Для этого зайдём в режим 4.1.3.11, выберем модель INF5 и запросим график по сумме знаний.

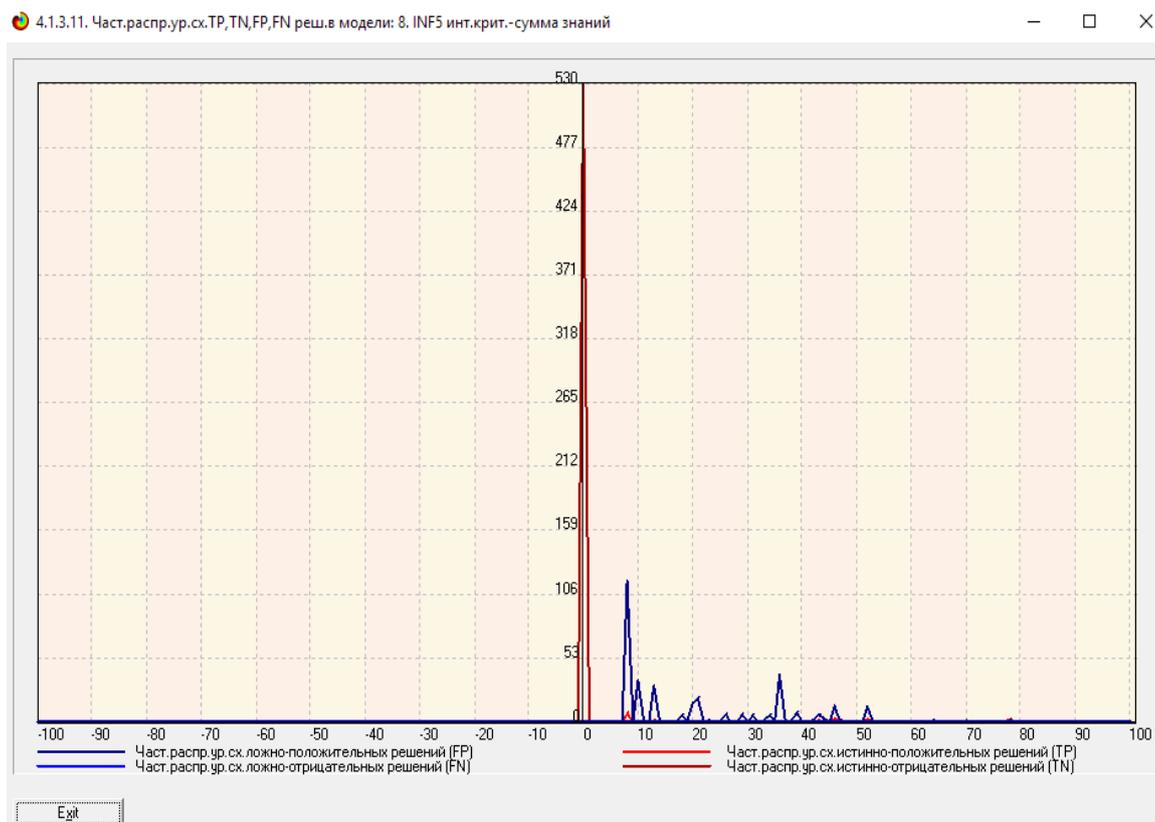


Рисунок 8 – Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF5

Как можно видеть, при модели INF5 непринадлежность объекта к классу лучше определяется, чем принадлежность. При этом существует вероятность ложного определения непринадлежности или принадлежности.

Выберем модель INF5 в качестве основной и перейдем к решению задач идентификации.

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

Если смотреть с позиции аналитика, то какие задачи при анализе баз данных пользователей можно выделить?

Во-первых, зависимость количества часов общения по мобильному телефону от провинции. Зачастую количество часов общения больше в тех провинциях Испании, где больше материальный достаток.

Во-вторых, зависимость количества часов общения по мобильному телефону от пола людей проживающих в Испании. Зачастую женщины тратят времени больше нежели мужчины.

В- третьих, зависимость количества часов общения по мобильному телефону от семейного положения людей проживающих в Испании. Зачастую люди состоящие в браке тратят времени на общение по телефону больше нежели не состоящие в браке.

Таким образом, в первую очередь, нас интересуют следующие связи:

- Провинции – общее.
- Мужчины – Женщины.
- Состоящие в браке – не состоящие в браке.
-

2.2. Выбор модели и пакетное распознавание

Зададим модель INF5 в качестве текущей в режиме 5.6 (рисунок 9), после чего проведем пакетное распознавание в режиме 4.1.2 (рисунок 10).

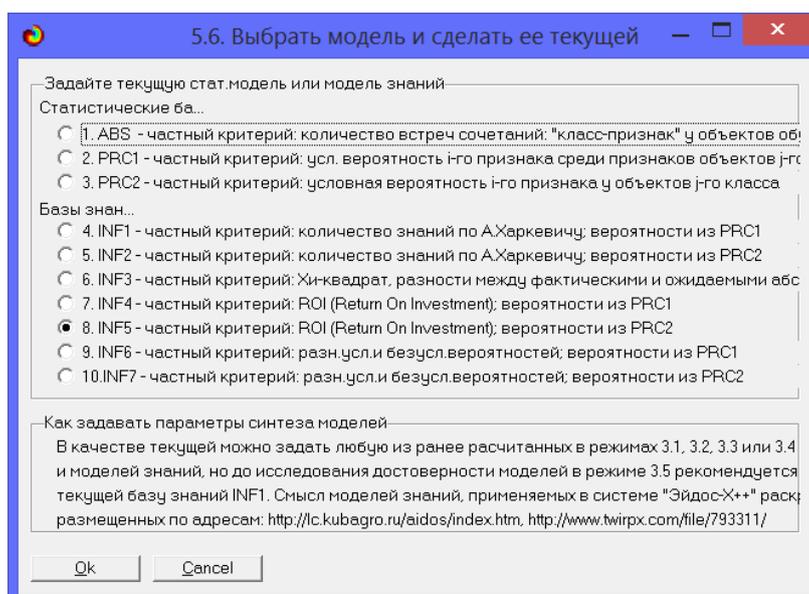


Рисунок 9 – Выбор текущей модели

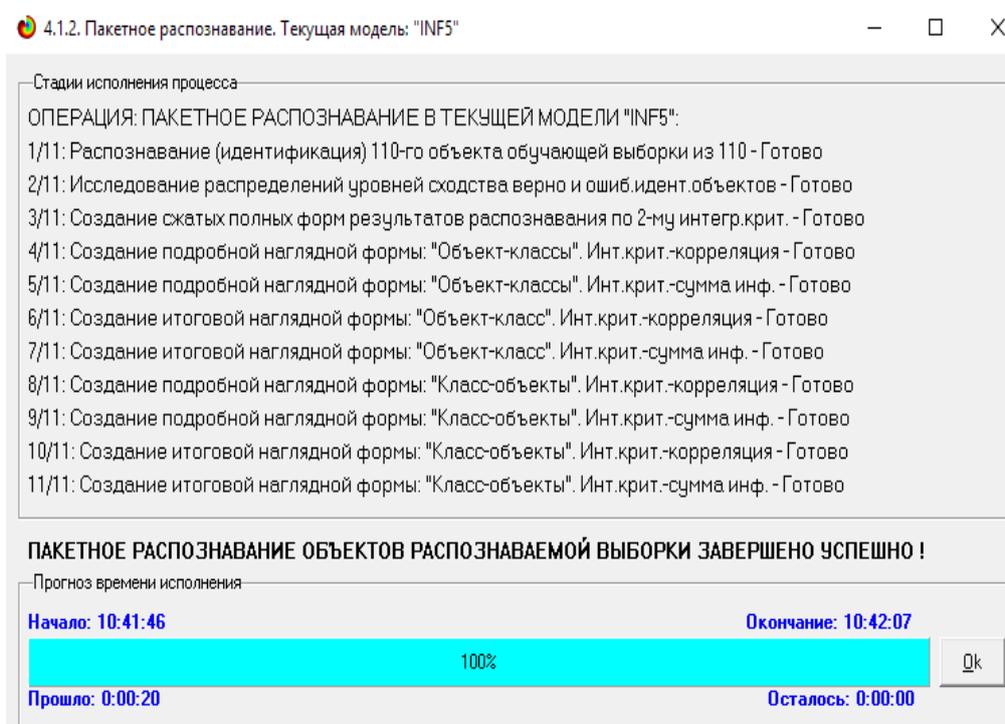


Рисунок 10 – Пакетное распознавание в новой модели

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 «Aidos-X» 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 11 и 12 приведены примеры прогнозов высокой и низкой достоверности классов, а также итоги распознавания классов объектов в модели INF5:

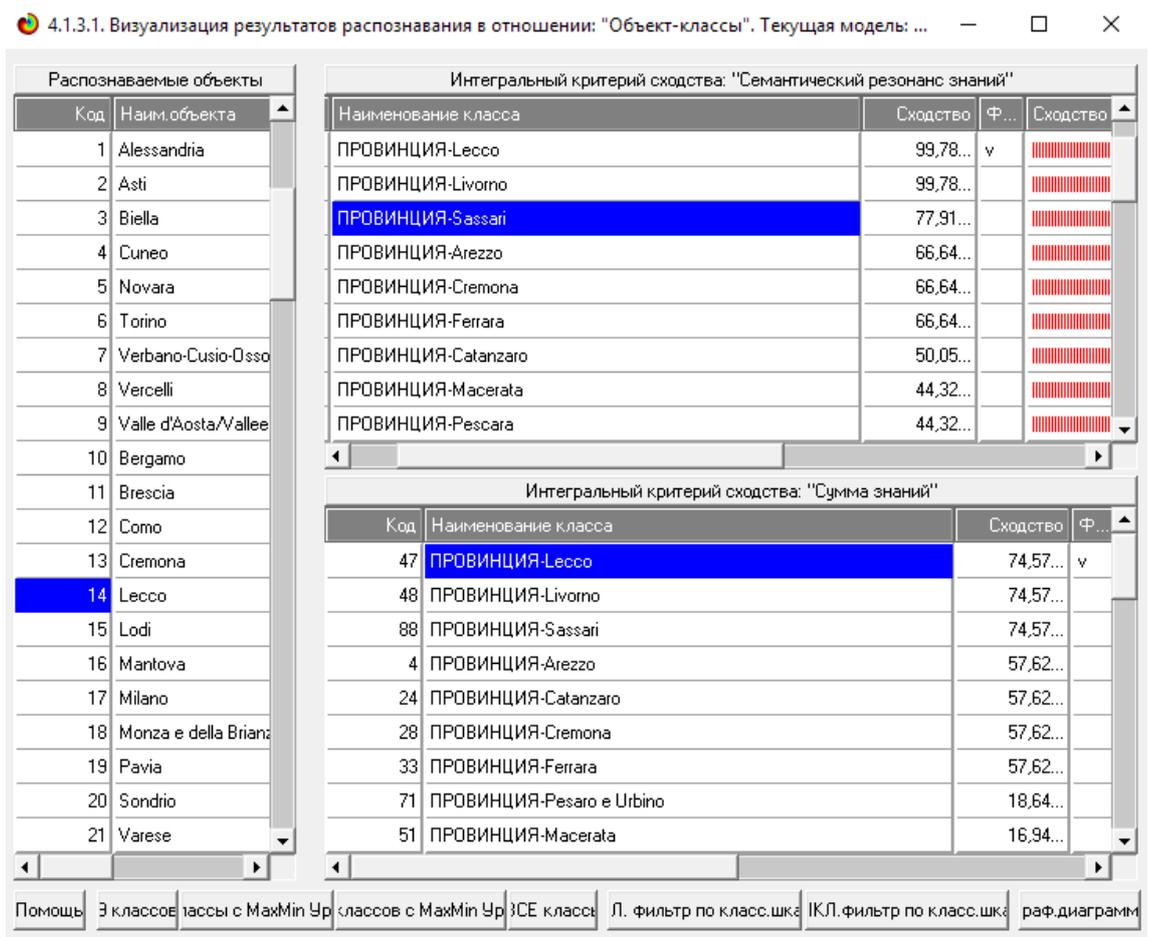


Рисунок 11 – Пример идентификации дисциплины оплаты услуг в модели INF5

2.3 Визуализация когнитивных функций

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 12 приведены визуализации всех когнитивных функций данного приложения для модели INF5.

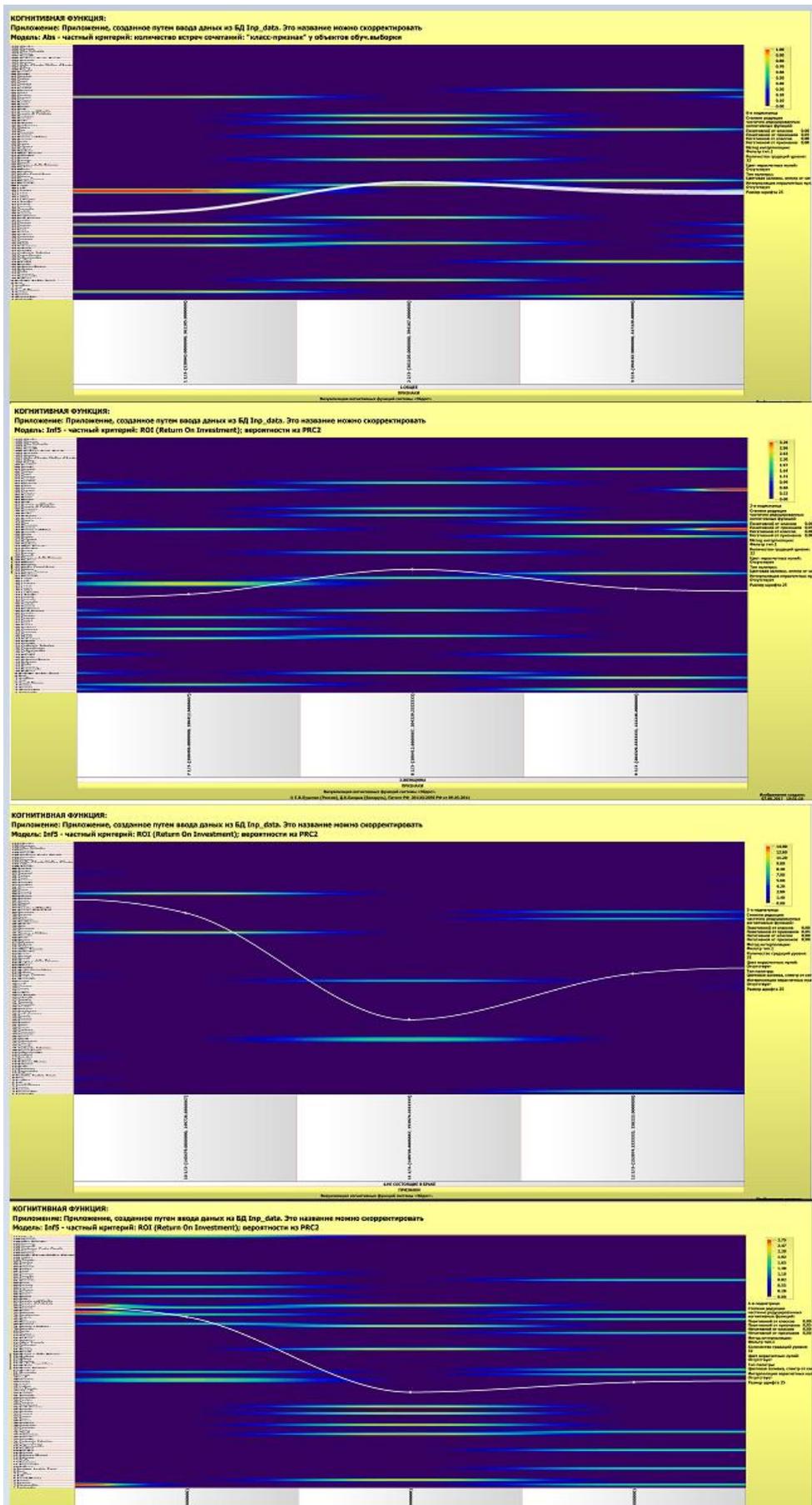


Рисунок 13 – Визуализация когнитивных функций для обобщённых классов описательных шкал для модели INF5

На основе полученных данных сделаем следующие прогнозы по времени использования мобильных телефонов в провинциях Испании:

- 1) Мужчины общаются по телефону меньше нежели женщины;
- 2) Не состоящие в браке общаются по телефону меньше нежели состоящие;
- 3) Больше всего по телефону общаются в провинции Roma

2.4 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.

В версии системы под MS Windows: «Aidos-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 17).

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса
1	ПРОВИНЦИЯ-Agrigento
2	ПРОВИНЦИЯ-Alessandria
3	ПРОВИНЦИЯ-Ancona
4	ПРОВИНЦИЯ-Arezzo
5	ПРОВИНЦИЯ-Ascoli Piceno
6	ПРОВИНЦИЯ-A...

SWOT-анализ класса: 2 "ПРОВИНЦИЯ-Alessandria" в модели: 6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила вли...
12	НЕ СОСТОЯЩИЕ В БРАКЕ-3/3-{153974.3333333, 158222...}	...
3	ОБЩЕЕ-3/3-{394167.0000000, 427229.0000000}	...
6	МУЖЧИНЫ-3/3-{189837.6666667, 204983.0000000}	...
9	ЖЕНЩИНЫ-3/3-{204329.3333333, 222246.0000000}	...
15	ЖЕНАТЫ/ЗАМУЖЕМ-3/3-{180625.0000000, 198253.000...}	...

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила вли...
13	ЖЕНАТЫ/ЗАМУЖЕМ-1/3-{145369.0000000, 162997.00000...}	...
14	ЖЕНАТЫ/ЗАМУЖЕМ-2/3-{162997.0000000, 180625.00000...}	...
8	ЖЕНЩИНЫ-2/3-{186412.6666667, 204329.3333333}	...
7	ЖЕНЩИНЫ-1/3-{168496.0000000, 186412.6666667}	...
2	ОБЩЕЕ-2/3-{361105.0000000, 394167.0000000}	...
4	МУЖЧИНЫ-1/3-{159547.0000000, 174692.3333333}	...
1	ОБЩЕЕ-1/3-{328043.0000000, 361105.0000000}	...
5	МУЖЧИНЫ-2/3-{174692.3333333, 189837.6666667}	...
11	НЕ СОСТОЯЩИЕ В БРАКЕ-2/3-{149726.6666667, 153974.3...}	...
10	НЕ СОСТОЯЩИЕ В БРАКЕ-1/3-{145479.0000000, 149726.6...}	...

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 17 – Пример SWOT-Матрицы в модели INF3

На рисунке 18 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF3.

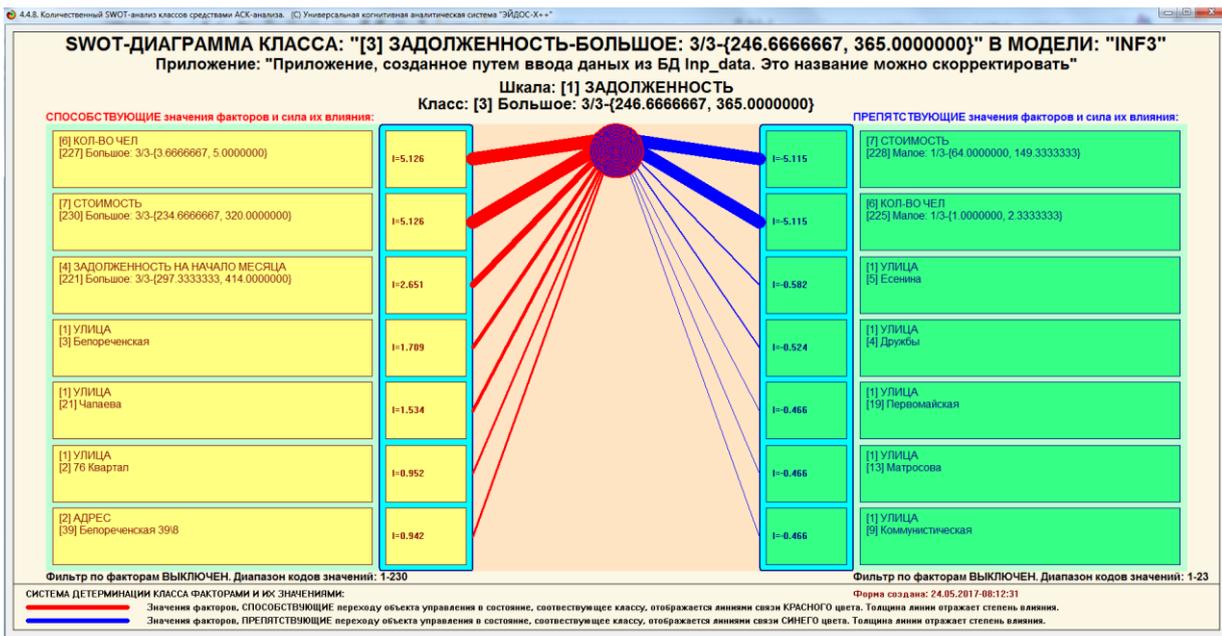


Рисунок 18. Пример SWOT-матрицы в модели INF5

На рисунке 19 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF5.

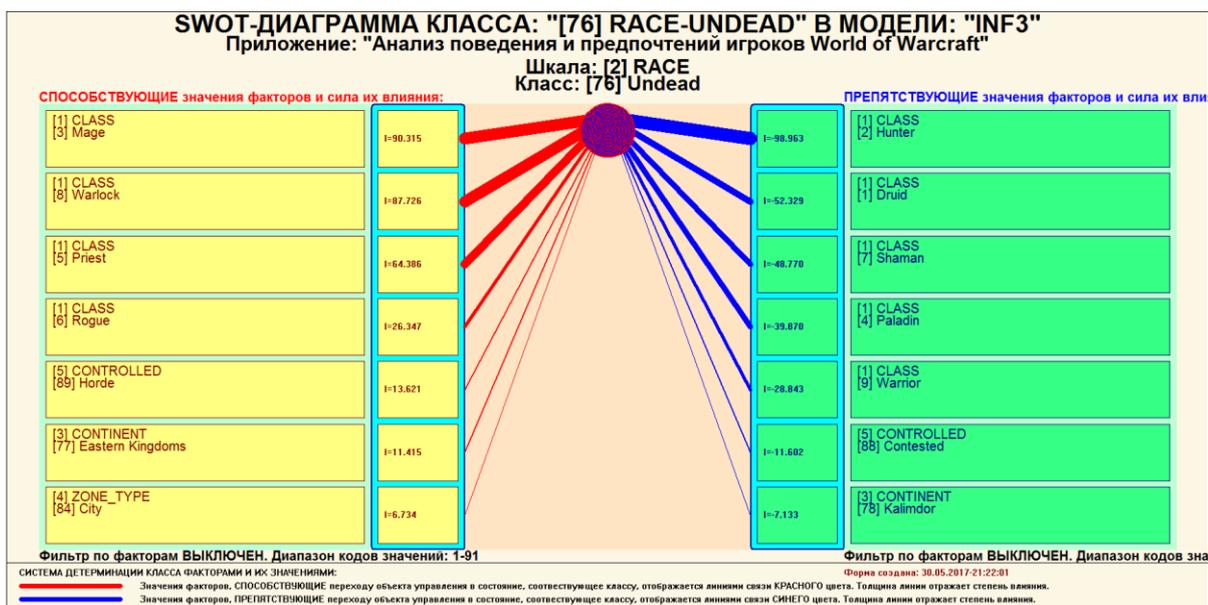


Рисунок 19 – SWOT-диаграмма признаков отрекшихся

2.5 Нелокальные нейронные сети и нейтроны

Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует не локальную нейронную сеть.

Рассмотрим пару примеров, возвращаясь к нашим задачам.

На рисунке 21 изображен нелокальный нейрон для 30-39 уровней.

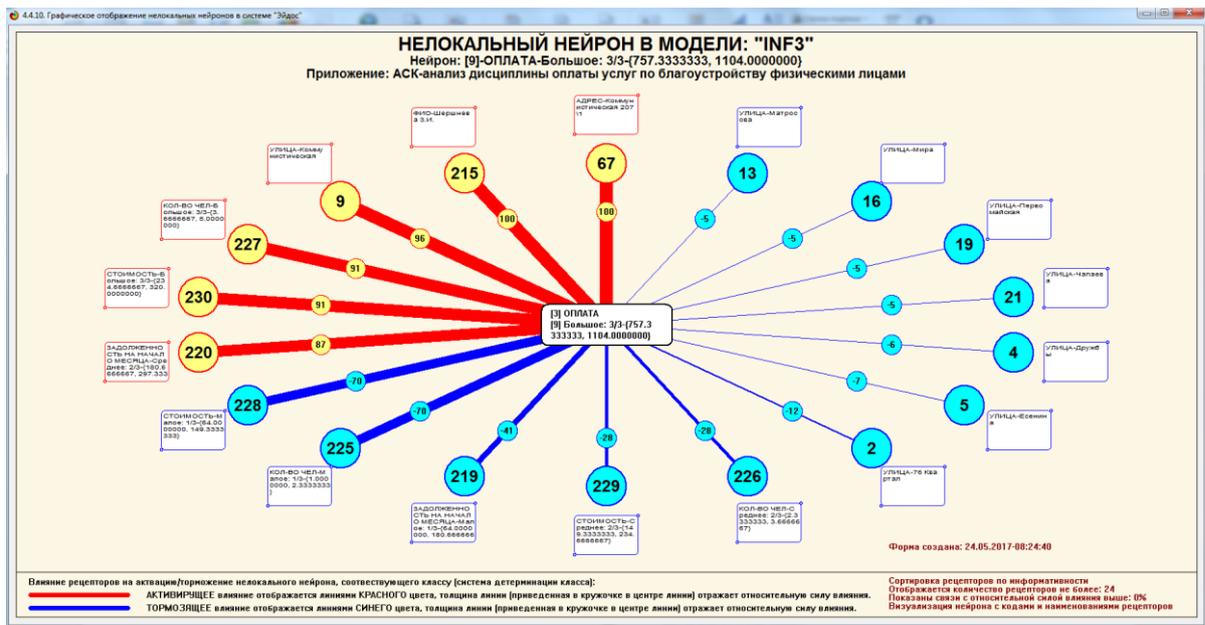


Рисунок 20 – Пример графического отображения нелокальных нейронов

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

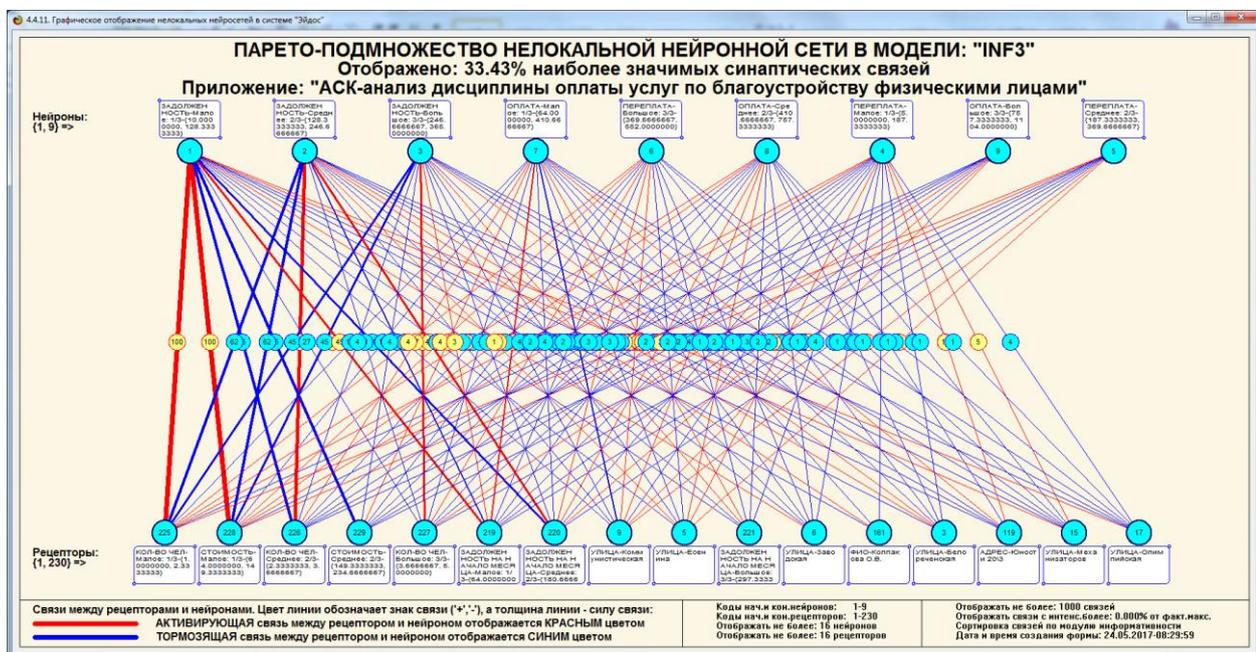


Рисунок 21 – Пример графического отображения парето-подмножества нелокальной нейронной сети

На рисунке 22 в графическом отображения парето-подмножества нелокальной нейронной сети в модели INF3 у классов слева сильнее связи между значениями факторов и результатом [9].

2.6 Кластерный и конструктивный анализ

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 1 "ПРОВИНЦИЯ-Agrigento" в модели: 4 "INF1"

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	ПРОВИНЦИЯ-Agrigento	1	1	ПРОВИНЦИЯ-Agrigento	
2	ПРОВИНЦИЯ-Alessandria	2	2	ПРОВИНЦИЯ-Alessandria	
3	ПРОВИНЦИЯ-Ancona	3	3	ПРОВИНЦИЯ-Ancona	
4	ПРОВИНЦИЯ-Arezzo	4	4	ПРОВИНЦИЯ-Arezzo	
5	ПРОВИНЦИЯ-Ascoli Piceno	5	5	ПРОВИНЦИЯ-Ascoli Piceno	
6	ПРОВИНЦИЯ-Asti	6	6	ПРОВИНЦИЯ-Asti	
7	ПРОВИНЦИЯ-Avellino	7	7	ПРОВИНЦИЯ-Avellino	
8	ПРОВИНЦИЯ-Bari	8	8	ПРОВИНЦИЯ-Bari	
9	ПРОВИНЦИЯ-Barletta-Andria-Trani	9	9	ПРОВИНЦИЯ-Barletta-Andria-Trani	
10	ПРОВИНЦИЯ-Belluno	10	10	ПРОВИНЦИЯ-Belluno	
11	ПРОВИНЦИЯ-Benevento	11	11	ПРОВИНЦИЯ-Benevento	
12	ПРОВИНЦИЯ-Bergamo	12	12	ПРОВИНЦИЯ-Bergamo	
13	ПРОВИНЦИЯ-Biella	13	13	ПРОВИНЦИЯ-Biella	
14	ПРОВИНЦИЯ-Bologna	14	14	ПРОВИНЦИЯ-Bologna	
15	ПРОВИНЦИЯ-Bolzano/Bozen	15	15	ПРОВИНЦИЯ-Bolzano/Bozen	
16	ПРОВИНЦИЯ-Brescia	16	16	ПРОВИНЦИЯ-Brescia	
17	ПРОВИНЦИЯ-Brindisi	17	17	ПРОВИНЦИЯ-Brindisi	
18	ПРОВИНЦИЯ-Cagliari	18	18	ПРОВИНЦИЯ-Cagliari	
19	ПРОВИНЦИЯ-Caltanissetta	19	19	ПРОВИНЦИЯ-Caltanissetta	
20	ПРОВИНЦИЯ-Campobasso	20	20	ПРОВИНЦИЯ-Campobasso	
21	ПРОВИНЦИЯ-Carbonia-Iglesias	21	21	ПРОВИНЦИЯ-Carbonia-Iglesias	
22	ПРОВИНЦИЯ-Caserta	22	22	ПРОВИНЦИЯ-Caserta	
23	ПРОВИНЦИЯ-Catania	23	23	ПРОВИНЦИЯ-Catania	

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ. фильтр по кл.шкале ВЫКЛ. фильтр по кл.шкале Вписать в окно Показать ВСЕ

Рисунок 22. Результаты кластерного и конструктивного анализа классов

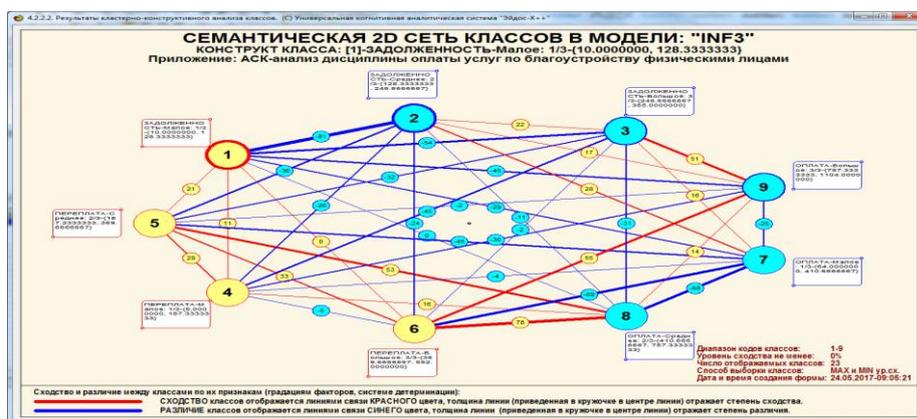


Рисунок 23. Графические результаты кластерного и конструктивного анализа классов

ЗАКЛЮЧЕНИЕ

Так как существует множество систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную базу Kaggle.

В данной курсовой работе была разработана системно-когнитивной модель диагностики использования мобильных телефонов с помощью программного инструментария системы 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 0,911, 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 ИСПОЛЬЗУЕМЫХ 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>