

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

Лабораторная работа

по дисциплине: Интеллектуальные информационные системы

на тему:

АСК- анализ влияния различных факторов на индекс счастья

выполнил студент группы: ПИ1521

Захаров Илья Олегович

Руководитель работы:
профессор Луценко Е.В.

Краснодар 2018

Содержание

Захаров Илья Олегович	1
ВВЕДЕНИЕ.....	3
1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	4
1.1. Описание решения	4
1.2. Преобразование исходных данных в файл MS Excel.....	4
1.3. Синтез и верификация статистических и интеллектуальных моделей.....	10
1.4. Виды моделей системы «Эйдос».....	11
1.5. Результаты верификации моделей	13
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ .	16
2.1. Решение задачи идентификации	16
2.2 Когнитивные функции.....	19
2.3. SWOT и PEST матрицы и диаграммы	21
2.4. Нелокальные нейронные сети	24
2.5 Кластерный и конструктивный анализ.....	24
ЗАКЛЮЧЕНИЕ	27
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	28

ВВЕДЕНИЕ

Одним из важнейших и перспективных направлений в развитии современных информационных технологий, на сегодняшний день, является создание систем искусственного интеллекта. В результате развития технологий в данном направлении, возникла необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи идентификации классов индекса счастья.

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

В данной лабораторной работе использована база данных «Happiness rank» из банка исходных данных по задачам искусственного интеллект-репозитория «Kaggle».

Для решения задачи используем стандартные возможности Microsoft Office Excel, блокнот, а также систему искусственного интеллекта «Aidos-X».

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

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

1. Преобразование исходных данных в промежуточные файлы MS Excel.
2. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы «Эйдос».
3. Синтез и верификация моделей предметной области.
4. Применение моделей для решения задач идентификации, прогнозирования и исследования предметной области.

1.2. Преобразование исходных данных в файл MS Excel

Из электронного ресурса баз данных «Kaggle» возьмем базу данных «Happiness rank», которую приведем к виду на рисунке 1.

The screenshot shows a Microsoft Excel spreadsheet titled 'Inp_data - Microsoft Excel'. The data is contained in a single sheet named 'Лист1'. The columns represent various socio-economic and happiness-related metrics for different countries. The first few columns are labeled A through S, and the last few columns are labeled M through Z. The data includes fields such as Country, Happiness Rank, Region, Happiness Score, Lower Confidence Interval, Upper Confidence Interval, Economy (GDP per Capita), Family, Health (Life Expectancy), Freedom, Trust (Government Corruption), Generosity, and Dystopia Residual. The data spans from row 1 to approximately row 140, with many rows containing country names like Denmark, Switzerland, Norway, and so on. The Excel ribbon at the top shows standard tabs like Home, Insert, Page Layout, Formulas, Data, Review, View, and Help. The status bar at the bottom right indicates the date as 03.05.2018 and the time as 10:34.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Country	Happiness Rank	Country	Region	Happiness Score	Lower Confidence Interval	Upper Confidence Interval	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity	Dystopia Residual				
2	Denmark	1	Denmark	Western Europe	7,526	7,48	7,59	1,4478	0,79041	0,84453	0,36171	2,70559						
3	Switzerland	2	Switzerland	Western Europe	7,509	7,428	7,59	1,52733	1,1452	0,85557	0,41200	0,28076	2,60485					
4	Austria	3	Austria	Western Europe	7,501	7,337	7,469	1,42833	1,13533	0,85733	0,40564	0,28739	2,61237					
5	Norway	4	Norway	Western Europe	7,498	7,421	7,575	1,57744	1,1299	0,79797	0,59009	0,35778	0,37989	2,66465				
6	Finland	5	Finland	Western Europe	7,413	7,351	7,475	1,40598	1,1346	0,81091	0,57104	0,41004	0,25492	2,82596				
7	Canada	6	Canada	North America	7,401	7,359	7,475	1,46473	1,0961	0,81091	0,57107	0,31329	0,44814	2,70485				
8	Iceland	7	Iceland	Iceland	7,339	7,258	7,384	1,44699	1,0920	0,81231	0,57221	0,31321	0,44815	2,70485				
9	New Zealand	8	New Zealand	Australia and New Zealand	7,334	7,264	7,404	1,36006	1,1728	0,85096	0,58147	0,41904	0,49001	2,47553				
10	Australia	9	Australia	Australia and New Zealand	7,315	7,241	7,427	1,3555	1,08781	0,81121	0,58183	0,31351	0,47607	2,4489				
11	Sweden	10	Sweden	Western Europe	7,291	7,257	7,355	1,41813	1,0920	0,85096	0,58147	0,31321	0,44814	2,70485				
12	Portugal	11	Portugal	Europe	7,267	7,198	7,355	1,35794	1,08549	0,84817	0,56452	0,29718	0,32288	3,51202				
13	Austria	12	Austria	Western Europe	7,119	7,045	7,195	1,45018	1,08183	0,80565	0,54355	0,21548	0,32886	2,69543				
14	United States	13	United States	North America	7,104	7,02	7,188	1,50796	1,0478	0,84813	0,41688	0,41077	2,72782					
15	United States	14	United States	Latin America and Caribbean	7,087	6,999	7,095	1,60479	1,04221	0,84813	0,41688	0,41077	2,72782					
16	Puerto Rico	15	Puerto Rico	Latin America and Caribbean	7,059	6,794	7,284	1,35943	1,0811	0,77798	0,48523	0,12278	0,22202	3,30798				
17	Germany	16	Germany	Western Europe	6,994	6,93	7,058	1,44787	1,0977	0,81487	0,33466	0,28551	0,30452	2,90811				
18	Ireland	17	Ireland	Latin America and Caribbean	6,952	6,875	7,029	1,08181	1,0934	0,81433	0,40423	0,16168	0,40776	5,97751				
19	Belgium	18	Belgium	Western Europe	6,929	6,905	7,027	1,42539	1,0920	0,85096	0,58147	0,31321	0,44814	2,70485				
20	Ireland	19	Ireland	Western Europe	6,907	6,858	6,978	1,48341	1,0811	0,81435	0,40408	0,29794	0,44818	2,13988				
21	Luxembourg	20	Luxembourg	Western Europe	6,871	6,804	6,938	1,69752	1,04	0,84542	0,54847	0,35339	0,37771	2,11059				
22	Mexico	21	Mexico	Latin America and Caribbean	6,778	6,88	6,876	1,11308	0,7348	0,71143	0,57049	0,18359	0,11733	3,53606				
23	United Kingdom	22	United Kingdom	Western Europe	6,759	6,573	6,824	1,44581	0,73079	0,74771	0,57049	0,18359	0,11733	3,53606				
24	United Kingdom	23	United Kingdom	Western Europe	6,725	6,647	6,803	1,40283	1,0887	0,80991	0,30036	0,27399	0,30156	2,14899				
25	China	24	China	Latin America and Caribbean	6,705	6,615	6,795	1,16167	0,9059	0,81883	0,37789	0,11481	0,31595	2,98050				
26	Panama	25	Panama	Latin America and Caribbean	6,701	6,601	6,801	1,08059	0,8891	0,79337	0,48927	0,08423	0,34118	5,05559				
27	Argentina	26	Argentina	Latin America and Caribbean	6,65	6,54	6,74	1,12337	0,95561	0,87711	0,32328	0,08423	0,34118	5,13989				
28	Czech Republic	27	Czech Republic	Central and Eastern Europe	6,598	6,515	6,877	1,30915	1,0979	0,79376	0,41418	0,05986	0,09929	2,90111				
29	United Arab Emirates	28	United Arab Emirates	Middle East and Northern Africa	6,573	6,494	6,652	1,57352	0,8711	0,79393	0,56215	0,35561	0,26951	2,23507				
30	Uruguay	29	Uruguay	Latin America and Caribbean	6,543	6,456	6,654	1,18337	1,0314	0,72185	0,54388	0,21394	0,18056	2,67159				
31	Malta	30	Malta	Western Europe	6,488	6,409	6,547	1,30302	0,82038	0,82044	0,44737	0,18359	0,18359	1,87023				
32	Colombia	31	Colombia	Latin America and Caribbean	6,481	6,384	6,578	1,50302	1,0217	0,59569	0,44738	0,05399	0,15626	3,17471				
33	France	32	France	Western Europe	6,478	6,397	6,559	1,36488	1,0034	0,83795	0,46562	0,17748	0,12118	2,47464				
34	United Arab Emirates	33	United Arab Emirates	Middle East and Northern Africa	6,472	6,399	6,552	1,48953	0,84433	0,84527	0,44935	0,18359	0,18359	2,77901				
35	Saudi Arabia	34	Saudi Arabia	Middle East and Northern Africa	6,379	6,287	6,572	1,48953	0,84433	0,84527	0,79044	0,30008	0,15457	2,61482				
36	Taiwan	35	Taiwan	Eastern Asia	6,379	6,305	6,455	1,39729	0,9262	0,79565	0,33377	0,0665	0,25453	2,61523				
37	Qatar	36	Qatar	Middle East and Northern Africa	6,375	6,178	6,438	1,82427	0,8796	0,79567	0,33377	0,0665	0,23289	1,58224				
38	UAE	37	UAE	Western Europe	6,361	6,288	6,424	1,24233	1,0230	0,81763	0,79567	0,0665	0,23289	2,58060				
39	Algeria	38	Algeria	Middle East and Northern Africa	6,325	6,227	6,483	1,05266	0,83511	0,81804	0,21006	0,16157	0,07044	3,48004				
40	Guatemala	39	Guatemala	Latin America and Caribbean	6,324	6,213	6,435	0,83454	0,8712	0,54039	0,50379	0,08701	0,28808	3,19863				
41	Suriname	40	Suriname	Latin America and Caribbean	6,269	6,075	6,445	1,06886	0,7877	0,50935	0,52324	0,12992	0,16665	5,05852				
42	Yemen	41	Yemen	Middle East and Northern Africa	6,239	6,154	6,324	1,61714	0,8776	0,55999	0,21616	0,23669	0,15945	2,23600				
43	Barbados	42	Barbados	Middle East and Northern Africa	6,218	6,178	6,308	1,40204	0,8341	0,55998	0,41275	0,28772	0,17447	2,77405				

Рисунок 1 - Исходные данные в файле MS Excel

Поскольку ввод исходных данных в систему «Эйдос» планируется осуществлять с помощью ее универсального программного интерфейса импорта данных из внешних баз данных, который работает с файлами MS

Excel, то преобразуем данные из html-файла в xls-файл, для чего выполним следующие операции.

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

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

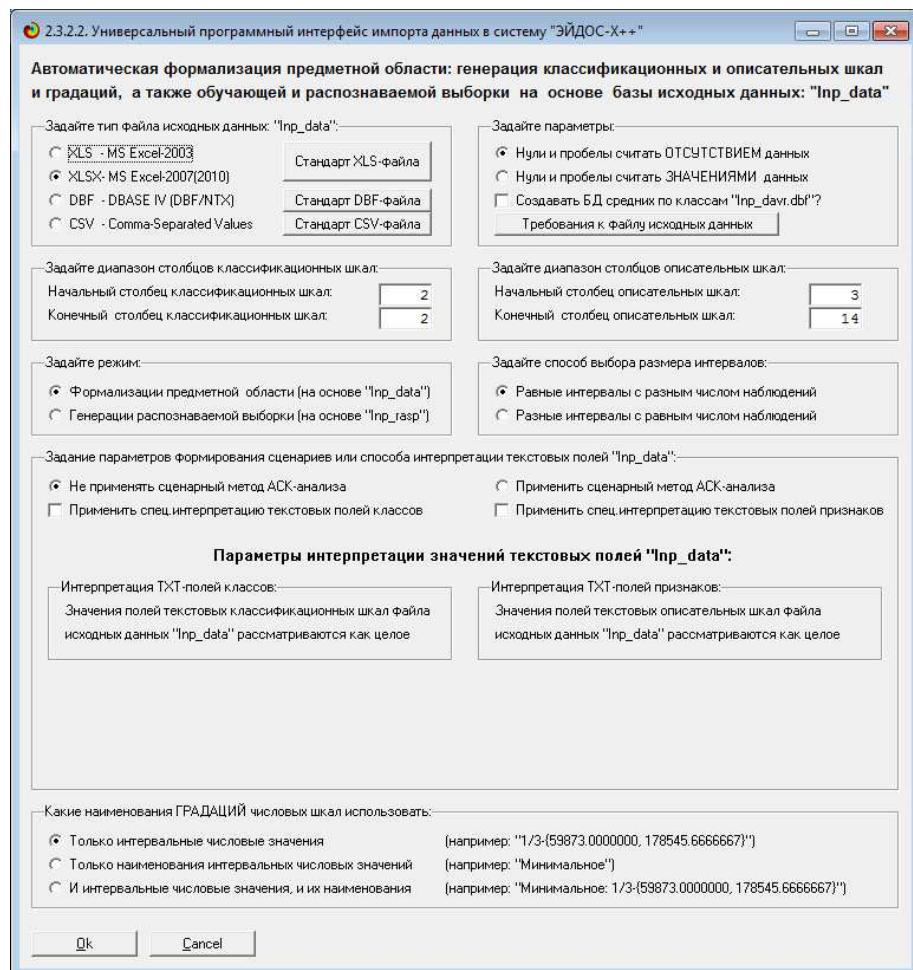


Рисунок 2- Экранная форма универсального программного интерфейса импорта данных в систему «Эйдос»

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

- «Тип файла исходных данных Inp_data»: «XLS – MS Excel-2003»;
- «Диапазон столбцов классификационных шкал»: «Начальный столбец классификационных шкал» – 2, «Конечный столбец классификационных шкал» – 2;
- «Диапазон столбцов описательных шкал»: «Начальный столбец описательных шкал»-3, «Конечный столбец описательных шкал»-14;
- «Параметры формирования сценариев или способа интерпретации текстовых полей»: «Не применять сценарный метод АСК-анализа и спец.интерпретацию TXT-полей».

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

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

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [5 x 217]

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	10	50	5,00
Текстовые	0	0	0,00	2	167	83,50
ВСЕГО:	1	5	5,00	12	217	18,08

Задайте число интервалов (градаций) в шкале:

В классификационных шкалах: В описательных шкалах:

Рисунок 3- Задание размерности модели системы «Эйдос»

Далее следует открытие окна, которое отображает стадию процесса импорта данных из внешней БД «Inp_data.xls» в систему «Эйдос», а также прогноз времени завершения этого процесса. В данном окне необходимо дождаться завершения формализации предметной области и нажать кнопку «OK», как показано на рисунке 4.

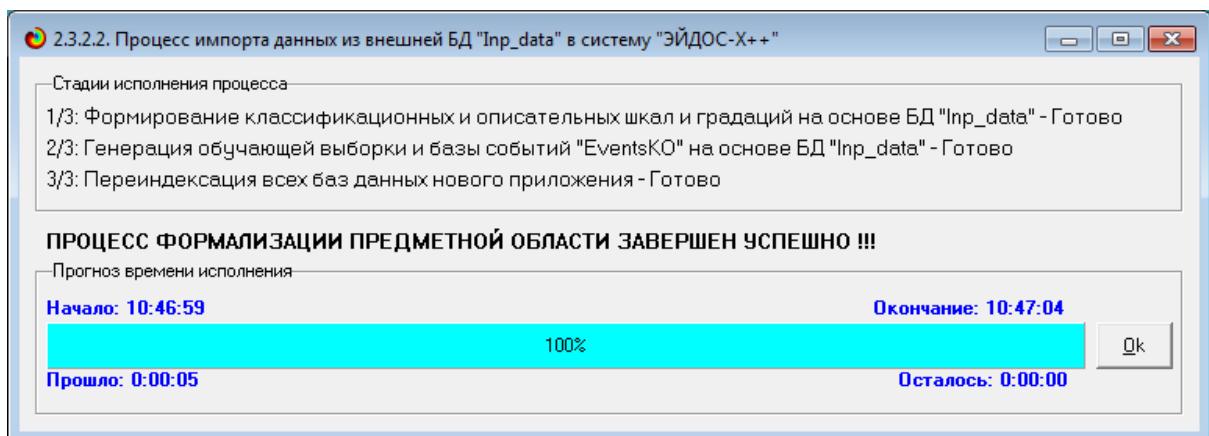


Рисунок 4- Процесс импорта данных из внешней БД «Inp_data.xls» в систему «Эйдос»

В результате выполненных действий формируются классификационные и описательные шкалы, градации, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Эти действия полностью автоматизированы и выполняются на 2-й этапе АСК- анализа «Формализация предметной области». Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1, рисунок 5.

2.1. Классификационные шкалы и градации. Текущая модель: "INFI"	
Код шкалы	Наименование классификационной шкалы
Код градации	Наименование градации классификационной шкалы
1	HAPPINESS RANK
1	1/5-{1.000000, 32.200000}
2	2/5-{32.200000, 63.400000}
3	3/5-{63.400000, 94.600000}
4	4/5-{94.600000, 125.800000}
5	5/5-{125.800000, 157.000000}

Рисунок 5- Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2, рисунок 6.

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	COUNTRY	1	Afghanistan
2	REGION	2	Albania
3	HAPPINESS SCORE	3	Algeria
4	LOWER CONFIDENCE INTERVAL	4	Angola
5	UPPER CONFIDENCE INTERVAL	5	Argentina
6	ECONOMY (GDP PER CAPITA)	6	Armenia
7	FAMILY	7	Australia
8	HEALTH (LIFE EXPECTANCY)	8	Austria
9	FREEDOM	9	Azerbaijan
10	TRUST (GOVERNMENT CORRUPTION)	10	Bahrain
11	GENEROSITY	11	Bangladesh
12	DYSTOPIA RESIDUAL	12	Belarus
		13	Belgium
		14	Belize
		15	Benin
		16	Bhutan
		17	Bolivia
		18	Bosnia and Herzegovina
		19	Botswana
		20	Brazil
		21	Bulgaria
		22	Burkina Faso

Рисунок 6- Описательные шкалы и градации (фрагмент)

Для просмотра обучающей выборки необходимо запустить режим 2.3.1., рисунок 7.

2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: "INF1"

Код объекта	Наименование объекта				Дата	Время						
Код объекта	Класс 1	Класс 2	Класс 3	Класс 4	Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	Denmark	1	38	167	172	177	182	186	192
2	Switzerland	1	197	202	207	210	216	0	0
3	Iceland
4	Norway
5	Finland
6	Canada
7	Netherlands
8	New Zealand
9	Australia
10	Sweden
11	Israel
12	Austria
13	United States
0	1	2	3	4	5	6	7	8	9

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

2.4. Просмотр эвентологических баз данных (баз событий). Текущая модель: "INF1"

№	Наименование объекта	2. HAPPINESS RANK	3. COUNTRY	4. REGION	5. HAPPINESS SCORE	6. LOWER CONFIDENCE INTERVAL	7. UPPER CONFIDENCE INTERVAL	8. ECONOMY (GDP PER CAPITA)	9. FAMILY
1	Denmark	1	38	167	172	177	182	186	192
2	Switzerland	1	135	167	172	177	182	187	192
3	Iceland	1	58	167	172	177	182	186	192
4	Norway	1	104	167	172	177	182	187	192
5	Finland	1	45	167	172	177	182	186	192
6	Canada	1	26	163	172	177	182	186	192
7	Netherlands	1	98	167	172	177	182	186	192
8	New Zealand	1	99	158	172	177	182	186	192
9	Australia	1	7	158	172	177	182	186	192
10	Sweden	1	134	167	172	177	182	186	192
11	Israel	1	64	162	172	177	182	186	192
12	Austria	1	8	167	172	177	182	186	192
13	United States	1	150	163	172	177	182	187	192
0	1	2	3	4	5	6	7	8	9

Рисунок 7- Обучающая выборка (фрагмент)

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

1.3. Синтез и верификация статистических и интеллектуальных моделей

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

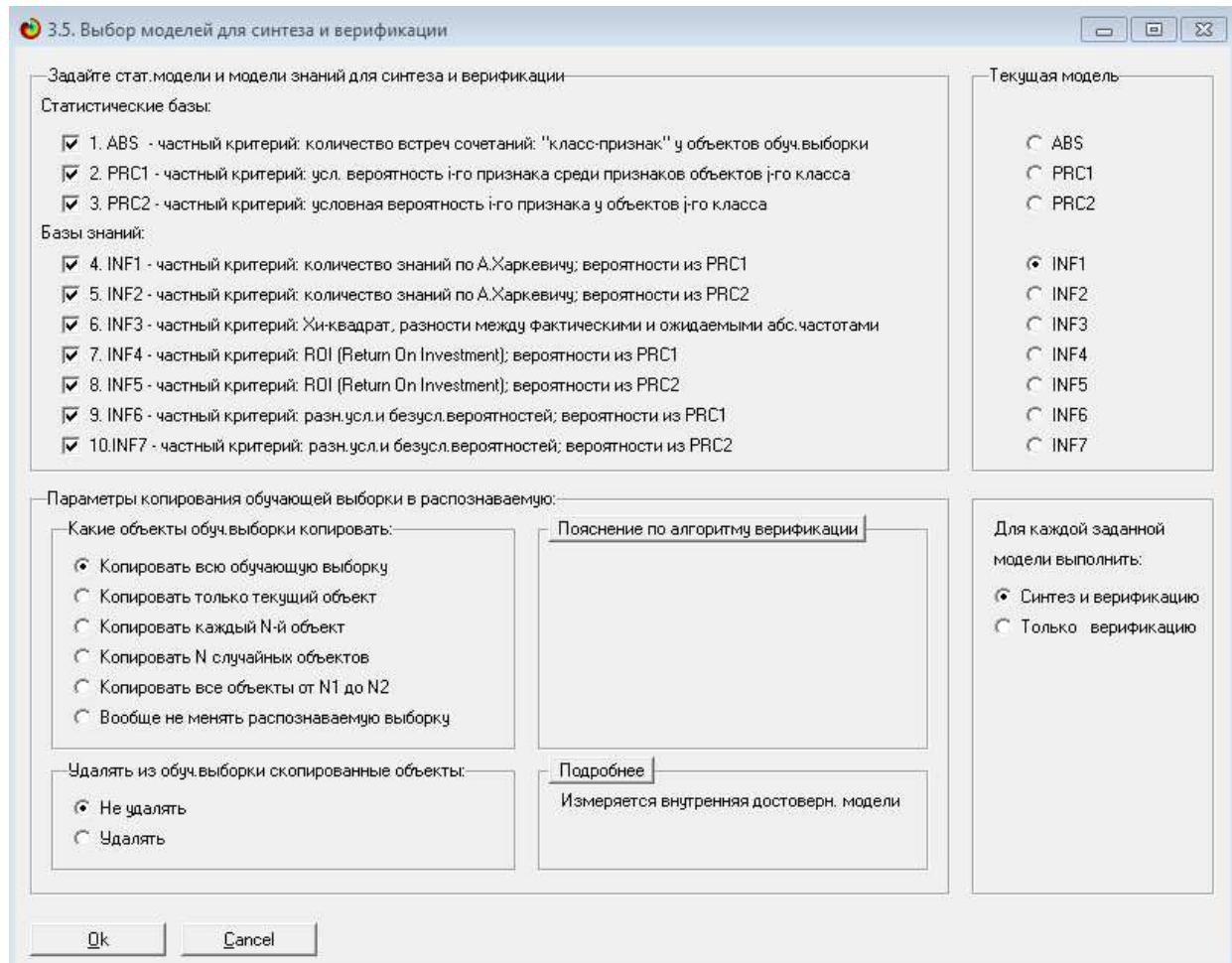


Рисунок 8- Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию, приведенные на рисунке 8. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 9.

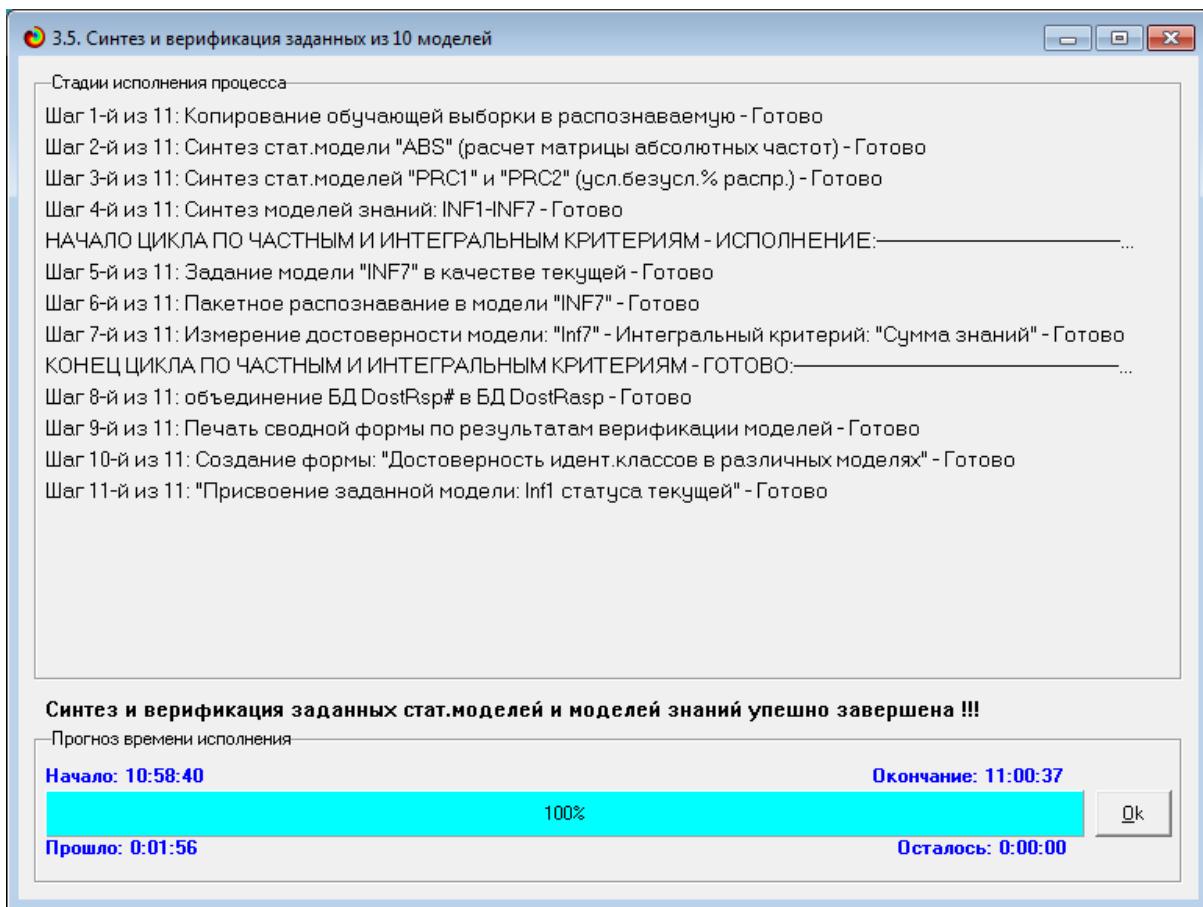


Рисунок 9- Синтез и верификация статистических моделей и моделей знаний

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

1.4. Виды моделей системы «Эйдос»

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

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

Код признака	Наименование описательной шкалы и градации	1. HAPPINESS RANK 1/5 {1.0, 32.2}	2. HAPPINESS RANK 2/5 {32.2, 63.4}	3. HAPPINESS RANK 3/5 {63.4, 94.6}	4. HAPPINESS RANK 4/5 {94.6, 125.8}	5. HAPPINESS RANK 5/5 {125.8, 157.0}	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY-Afghanistan					1	1	0.20	0.45
2	COUNTRY-Albania				1		1	0.20	0.45
3	COUNTRY-Algeria		1				1	0.20	0.45
4	COUNTRY-Angola					1	1	0.20	0.45
5	COUNTRY-Argentina	1					1	0.20	0.45
6	COUNTRY-Armenia				1		1	0.20	0.45
7	COUNTRY-Australia	1					1	0.20	0.45
8	COUNTRY-Austria	1					1	0.20	0.45
9	COUNTRY-Azerbaijan			1			1	0.20	0.45
10	COUNTRY-Bahrain		1				1	0.20	0.45
11	COUNTRY-Bangladesh				1		1	0.20	0.45
12	COUNTRY-Belarus		1				1	0.20	0.45
13	COUNTRY-Belgium	1					1	0.20	0.45
14	COUNTRY-Belize		1				1	0.20	0.45
15	COUNTRY-Benin					1	1	0.20	0.45
16	COUNTRY-Bhutan			1			1	0.20	0.45
17	COUNTRY-Bolivia		1				1	0.20	0.45
18	COUNTRY-Bosnia and Herzegovina			1			1	0.20	0.45
19	COUNTRY-Botswana					1	1	0.20	0.45
20	COUNTRY-Brazil	1					1	0.20	0.45
21	COUNTRY-Bulgaria					1	1	0.20	0.45
22	COUNTRY-Burkina Faso					1	1	0.20	0.45

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

Код признака	Наименование описательной шкалы и градации	1. HAPPINESS RANK 1/5 {1.0, 32.2}	2. HAPPINESS RANK 2/5 {32.2, 63.4}	3. HAPPINESS RANK 3/5 {63.4, 94.6}	4. HAPPINESS RANK 4/5 {94.6, 125.8}	5. HAPPINESS RANK 5/5 {125.8, 157.0}	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY-Afghanistan					0.491	0.491	0.098	0.219
2	COUNTRY-Albania				0.500		0.500	0.100	0.224
3	COUNTRY-Algeria		0.499				0.499	0.100	0.223
4	COUNTRY-Angola					0.491	0.491	0.098	0.219
5	COUNTRY-Argentina	0.489					0.489	0.098	0.219
6	COUNTRY-Armenia				0.500		0.500	0.100	0.224
7	COUNTRY-Australia	0.489					0.489	0.098	0.219
8	COUNTRY-Austria	0.489					0.489	0.098	0.219
9	COUNTRY-Azerbaijan			0.500			0.500	0.100	0.224
10	COUNTRY-Bahrain		0.499				0.499	0.100	0.223
11	COUNTRY-Bangladesh				0.500		0.500	0.100	0.224
12	COUNTRY-Belarus		0.499				0.499	0.100	0.223
13	COUNTRY-Belgium	0.489					0.489	0.098	0.219
14	COUNTRY-Belize		0.499				0.499	0.100	0.223
15	COUNTRY-Benin					0.491	0.491	0.098	0.219
16	COUNTRY-Bhutan			0.500			0.500	0.100	0.224
17	COUNTRY-Bolivia		0.499				0.499	0.100	0.223
18	COUNTRY-Bosnia and Herzegovina			0.500			0.500	0.100	0.224
19	COUNTRY-Botswana					0.491	0.491	0.098	0.219
20	COUNTRY-Brazil	0.489					0.489	0.098	0.219
21	COUNTRY-Bulgaria					0.491	0.491	0.098	0.219
22	COUNTRY-Burkina Faso					0.491	0.491	0.098	0.219

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

Код признака	Наименование описательной шкалы и градации	1. HAPPINESS RANK 1/5 {1,0, 32,2}	2. HAPPINESS RANK 2/5 {32,2, 63,4}	3. HAPPINESS RANK 3/5 {63,4, 94,6}	4. HAPPINESS RANK 4/5 {94,6, 125,8}	5. HAPPINESS RANK 5/5 {125,8, 157,0}	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY-Afghanistan	-0.204	-0.198	-0.197	-0.197	0.797			0.445
2	COUNTRY-Albania	-0.204	-0.198	-0.197	0.803	-0.203			0.449
3	COUNTRY-Algeria	-0.204	0.802	-0.197	-0.197	-0.203			0.448
4	COUNTRY-Angola	-0.204	-0.198	-0.197	-0.197	0.797			0.445
5	COUNTRY-Argentina	0.796	-0.198	-0.197	-0.197	-0.203			0.445
6	COUNTRY-Armenia	-0.204	-0.198	-0.197	0.803	-0.203			0.449
7	COUNTRY-Australia	0.796	-0.198	-0.197	-0.197	-0.203			0.445
8	COUNTRY-Austria	0.796	-0.198	-0.197	-0.197	-0.203			0.445
9	COUNTRY-Azerbaijan	-0.204	-0.198	0.803	-0.197	-0.203			0.449
10	COUNTRY-Bahrain	-0.204	0.802	-0.197	-0.197	-0.203			0.448
11	COUNTRY-Bangladesh	-0.204	-0.198	-0.197	0.803	-0.203			0.449
12	COUNTRY-Belarus	-0.204	0.802	-0.197	-0.197	-0.203			0.448
13	COUNTRY-Belgium	0.796	-0.198	-0.197	-0.197	-0.203			0.445
14	COUNTRY-Belize	-0.204	0.802	-0.197	-0.197	-0.203			0.448
15	COUNTRY-Benin	-0.204	-0.198	-0.197	-0.197	0.797			0.445
16	COUNTRY-Bhutan	-0.204	-0.198	0.803	-0.197	-0.203			0.449
17	COUNTRY-Bolivia	-0.204	0.802	-0.197	-0.197	-0.203			0.448
18	COUNTRY-Bosnia and Herzegovina	-0.204	-0.198	0.803	-0.197	-0.203			0.449
19	COUNTRY-Botswana	-0.204	-0.198	-0.197	-0.197	0.797			0.445
20	COUNTRY-Brazil	0.796	-0.198	-0.197	-0.197	-0.203			0.445
21	COUNTRY-Bulgaria	-0.204	-0.198	-0.197	-0.197	0.797			0.445
22	COUNTRY-Burkina Faso	-0.204	-0.198	-0.197	-0.197	0.797			0.445

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

Код признака	Наименование описательной шкалы и градации	1. HAPPINESS RANK 1/5 {1,0, 32,2}	2. HAPPINESS RANK 2/5 {32,2, 63,4}	3. HAPPINESS RANK 3/5 {63,4, 94,6}	4. HAPPINESS RANK 4/5 {94,6, 125,8}	5. HAPPINESS RANK 5/5 {125,8, 157,0}	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY-Afghanistan					3.916	3.916	0.783	1.751
2	COUNTRY-Albania				4.076		4.076	0.815	1.823
3	COUNTRY-Algeria		4.048				4.048	0.810	1.810
4	COUNTRY-Angola					3.916	3.916	0.783	1.751
5	COUNTRY-Argentina	3.891					3.891	0.778	1.740
6	COUNTRY-Armenia	3.891			4.076		4.076	0.815	1.823
7	COUNTRY-Australia	3.891					3.891	0.778	1.740
8	COUNTRY-Austria	3.891					3.891	0.778	1.740
9	COUNTRY-Azerbaijan			4.076			4.076	0.815	1.823
10	COUNTRY-Bahrain		4.048				4.048	0.810	1.810
11	COUNTRY-Bangladesh				4.076		4.076	0.815	1.823
12	COUNTRY-Belarus		4.048				4.048	0.810	1.810
13	COUNTRY-Belgium	3.891					3.891	0.778	1.740
14	COUNTRY-Belize		4.048				4.048	0.810	1.810
15	COUNTRY-Benin				3.916		3.916	0.783	1.751
16	COUNTRY-Bhutan			4.076			4.076	0.815	1.823
17	COUNTRY-Bolivia		4.048				4.048	0.810	1.810
18	COUNTRY-Bosnia and Herzegovina			4.076			4.076	0.815	1.823
19	COUNTRY-Botswana				3.916		3.916	0.783	1.751
20	COUNTRY-Brazil	3.891					3.891	0.778	1.740
21	COUNTRY-Bulgaria					3.916	3.916	0.783	1.751
22	COUNTRY-Burkina Faso					3.916	3.916	0.783	1.751

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

1.5. Результаты верификации моделей

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

4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит. Текущая модель: "INF1"													
Наименование модели и частного критерия	Интегральный критерий	Число единиц	Полнота модели	F-мера Ван Ризбергена	Сумма модуля уровней сходства истинно-положительных решений (STP)	Сумма модуля уровней сходства истинно-отрицательных решений (SFP)	Сумма модуля уровней сходства ложно-положительных решений (SLP)	Сумма модуля уровней сходства ложно-отрицательных решений (SFL)	S-точность модели	S-полнота модели	L1-мера проф. Е.В.Лущенко	Среднее значение	
1. ABS - частный критерий: количество встреч сочтаний: "классификации";	Корреляция абсолютной частоты с общим количеством встреч сочтаний	0.203	1.000	0.337	119.130	0.358	166.605		0.417	1.000	0.588		
1. ABS - частный критерий: количество встреч сочтаний: "классификации";	Сумма абсолютной частоты по признакам	0.200	1.000	0.333	115.268		195.616		0.371	1.000	0.541		
2. PRC1 - частный критерий: усл. вероятность i-го признака среди остальных	Корреляция условной частоты с остаточной частотой	0.203	1.000	0.337	119.130	0.358	166.605		0.417	1.000	0.588		
2. PRC1 - частный критерий: усл. вероятность i-го признака среди остальных	Сумма условной частоты с остаточной частотой	0.200	1.000	0.333	117.789		200.461		0.370	1.000	0.540		
3. PRC2 - частный критерий: условная вероятность i-го признака среди остальных	Корреляция условной частоты с остаточной частотой	0.203	1.000	0.337	119.130	0.358	166.605		0.417	1.000	0.588		
3. PRC2 - частный критерий: условная вероятность i-го признака среди остальных	Сумма условной частоты с остаточной частотой	0.200	1.000	0.333	117.442		199.785		0.370	1.000	0.540		
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс знаний	0.746	0.936	0.831	62.696	255.175	7.081	1.504	0.899	0.977	0.936		
4. INF1 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.463	0.987	0.630	80.884	100.694	25.129	0.301	0.763	0.996	0.864		
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Семантический резонанс знаний	0.750	0.936	0.833	62.603	254.807	7.040	1.505	0.899	0.977	0.936		
5. INF2 - частный критерий: количество знаний по АХаркевичу; в...	Сумма знаний	0.463	0.987	0.630	80.875	100.511	25.057	0.304	0.763	0.996	0.864		
6. INF3 - частный критерий: Хиквадрат, различия между фактическими и ожидаемыми значениями	Семантический резонанс знаний	0.542	0.987	0.700	95.926	128.291	35.317	0.015	0.731	1.000	0.844		
6. INF3 - частный критерий: Хиквадрат, различия между фактическими и ожидаемыми значениями	Сумма знаний	0.542	0.987	0.700	79.343	104.986	25.656	0.012	0.756	1.000	0.861		
7. INF4 - частный критерий ROI (Return On Investment); вероятно...	Семантический резонанс знаний	0.835	0.936	0.883	60.519	232.653	2.634	0.766	0.958	0.988	0.973		
7. INF4 - частный критерий ROI (Return On Investment); вероятно...	Сумма знаний	0.375	1.000	0.545	82.987	20.995	25.506		0.765	1.000	0.867		
8. INF5 - частный критерий ROI (Return On Investment); вероятно...	Семантический резонанс знаний	0.835	0.936	0.883	60.486	232.563	2.625	0.770	0.958	0.987	0.973		
8. INF5 - частный критерий ROI (Return On Investment); вероятно...	Сумма знаний	0.374	1.000	0.544	82.653	20.941	25.397		0.765	1.000	0.867		
9. INF6 - частный критерий: разн.усл. и безразн.вероятностей; вер...	Семантический резонанс знаний	0.431	0.981	0.599	94.485	65.411	40.615	0.140	0.699	0.999	0.823		
9. INF6 - частный критерий: разн.усл. и безразн.вероятностей; вер...	Сумма знаний	0.364	0.987	0.532	80.933	34.190	37.937	0.013	0.681	1.000	0.810		
10.INF7 - частный критерий: разн.усл. и безразн.вероятностей; вер...	Семантический резонанс знаний	0.431	0.981	0.599	94.353	65.449	40.524	0.138	0.700	0.999	0.823		
10.INF7 - частный критерий: разн.усл. и безразн.вероятностей; вер...	Сумма знаний	0.363	0.987	0.531	80.600	34.078	37.753	0.012	0.681	1.000	0.810		

Рисунок 14- Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF5 при интегральном критерии «Сумма знаний». При этом точность модели составляет 0,916. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием, описанным на рисунке 15.

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, что 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, который не будет точно отнесен к классам, к которым он относится. Но при этом будет быть 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 будет иметь высокую достоверность идентификации).													
ПРЕДСЛОВЛЕНЫЙ 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 то, 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 мы выбрасываем 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 не осуществляется, но практика получает его.													
На практике мы имеем дело с реальным прогнозом. Реальный прогноз уменьшает неопределенность, но не полностью, как идеальный прогноз. Реальный прогноз не снимает опасения, что ошибочно идентифицированный кубик, делается из-за прогноза, что не выпадет, или же, наоборот, что выпадет, но не снимает опасения, что ошибочно не идентифицированный кубик, делается из-за прогноза, что выпадет.													
ПОДСТАВЛЕНЫЙ ПСЕВДОПРОГНОЗ													
Представим себе, что мы выбрасываем кубиком 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 то, 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.													
ПОЛНОСТЬЮ ПОДСТАВЛЕНЫЙ ПСЕВДОПРОГНОЗ													
Таким образом, если просуммировать число верно идентифицированных объектов, и вычесть число ошибочно идентифицированных объектов, то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно идентифицировать все объекты, так и ее способность не идентифицировать все объекты. Критерий, предложенный в 1969 году, это метод F-меры, предполагающий две варианты нормировки: $\{-1, +1\}$ и $\{0, 1\}$.													
ПОЛНОСТЬЮ ИДЕНТИФИЦИРОВАННЫЙ ПСЕВДОПРОГНОЗ													
Предположим, что модель Ван Ризбергена имеет одинаковую достоверность идентификации и не идентификации. Тогда можно посчитать средневзвешенную характеристику всех этих видов прогнозов.													
ПОЛНОСТЬЮ ИДЕНТИФИЦИРОВАННЫЙ ПСЕВДОПРОГНОЗ													
Классическая F-мера достоверности моделей Ван Ризбергена (колонка выделена ярко-зеленым фоном):													
F-мера = $2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$													
Precision = $\frac{\text{TP}}{\text{TP} + \text{FN}}$													
Recall = $\frac{\text{TP}}{\text{TP} + \text{TN}}$													
L1-мера проф. Е.В.Лущенко – нечеткое мультиклассовое обобщение классической F-меры с учетом сумм уровней сходства (колонка выделена ярко-зеленым фоном):													
L1-мера = $2 \times \frac{\text{S}(precision \times recall)}{\text{S}(precision) + \text{S}(recall)}$													
S(precision) = $\frac{\text{TP}}{\text{TP} + \text{FP}}$													
S(recall) = $\frac{\text{TP}}{\text{TP} + \text{FN}}$													
STP = Сумма модулей сходства истинно-положительных решений; STN = Сумма модулей сходства истинно-отрицательных решений;													
SFP = Сумма модулей сходства ложно-положительных решений; SFL = Сумма модулей сходства ложно-отрицательных решений;													
L2-мера проф. Е.В.Лущенко – нечеткое мультиклассовое обобщение классической F-меры с учетом средних уровней сходства (колонка выделена желтым фоном):													
L2-мера = $2 \times \frac{\text{A}(\text{precision} \times \text{recall})}{\text{A}(\text{precision}) + \text{A}(\text{recall})}$													
A(precision) = $\frac{\text{TP}}{\text{TP} + \text{FP}}$													
A(recall) = $\frac{\text{TP}}{\text{TP} + \text{FN}}$													
ATP-STP/TP = Среднее модулей сходства истинно-положительных решений; AEP-SFN/FN = Среднее модулей сходства истинно-отрицательных решений;													
ATP-STP/TP = Среднее модулей сходства ложно-положительных решений; AEP-SFP/FN = Среднее модулей сходства ложно-отрицательных решений;													
Строки с максимальными значениями F-меры, L1-меры и L2-меры выделены фоном цвета, соответствующим колонке.													
ЛУЧЕНКО Е.В. Ивариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе, и системе "Эйдос-Х++" // Политехнический сетевой электронный научный журнал // КубГАУ, 2017. – №02(126). С. 1 – 32. – IDA [article ID]: 12617/02001. – Режим доступа: http://ej3.kubagro.ru/2017/02/pdf/01.pdf . – 2, у. л.п.													

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

Стоит отметить, что статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не

идентификации, чем модели знаний, и практически никогда – более высокую. Этим и оправдано применение моделей знаний и интеллектуальных технологий. На рисунке 16 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF5.

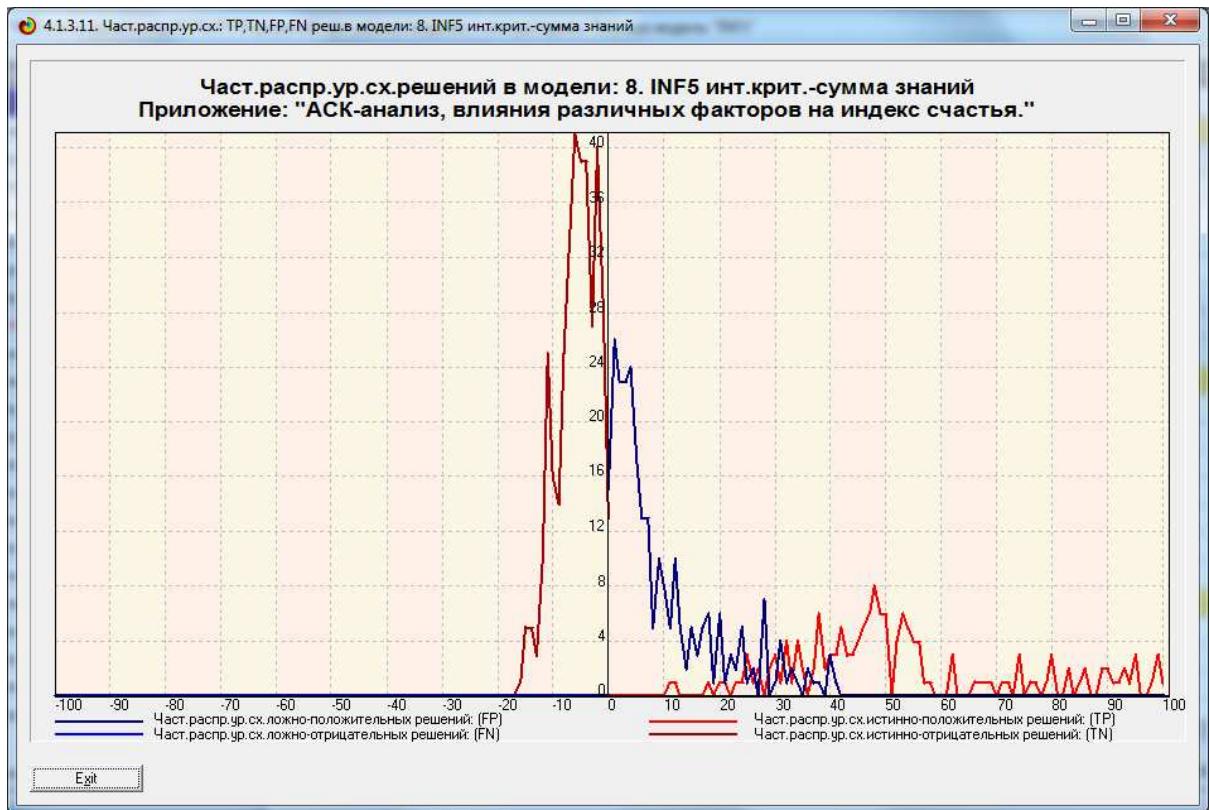


Рисунок 16- Частное распределение ложно-положительных решений, ложно-отрицательных решений, истинно-положительных решений и истинно-отрицательных решений объекта моделирования в модели INF5

Из рисунка 16 видно, что:

- наиболее достоверная модель INF5 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 14);
- Отрицательные решения всегда истинные. В положительных решениях до 10 идут только ложные решения, с 10 до 40 идут и ложные и истинные решения при чем, чем больше уровень сходства, тем больше истинных решений. После 40 идут только истинные решения.

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

В соответствии с технологией АСК-анализа зададим текущей модель INF5 (режим 5.6) и проведем пакетное распознавание в режиме 4.2.1, как показано на рисунке 18.

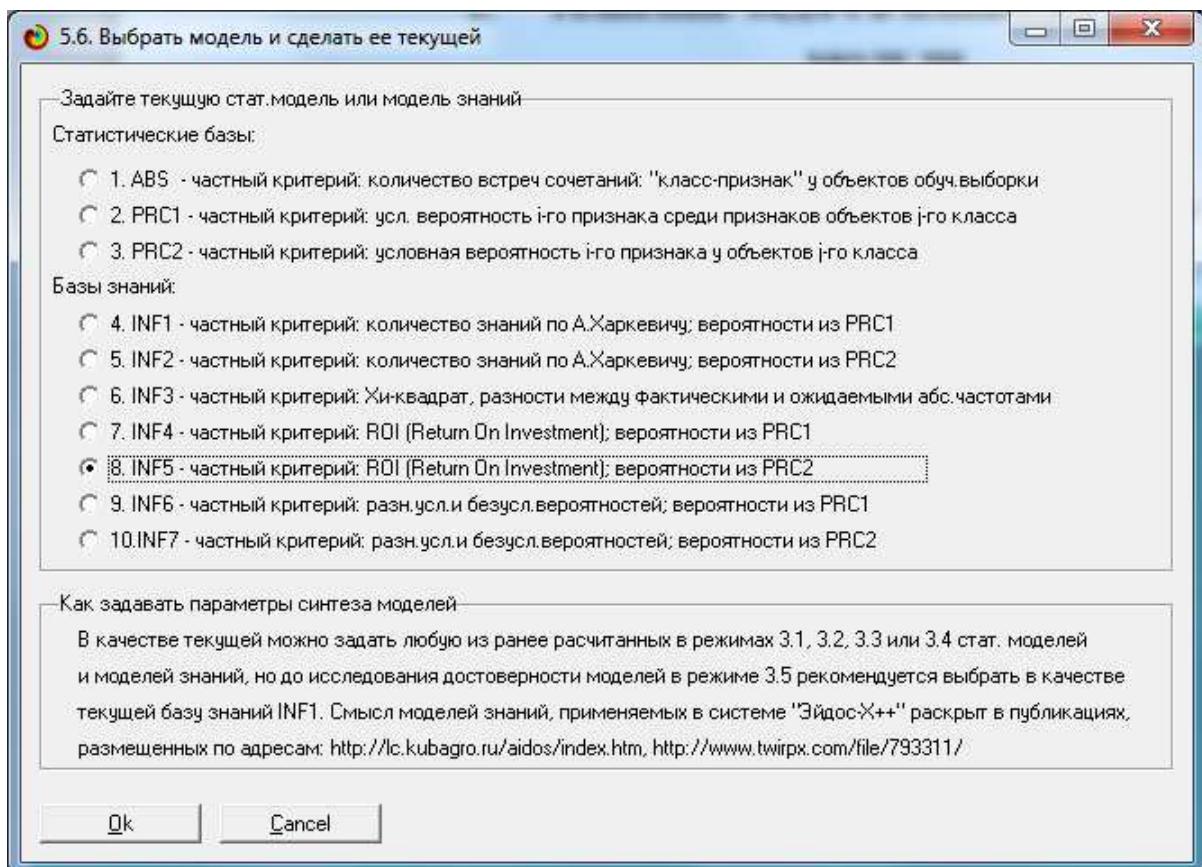


Рисунок 17- Экранная форма режима задания модели в качестве текущей

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

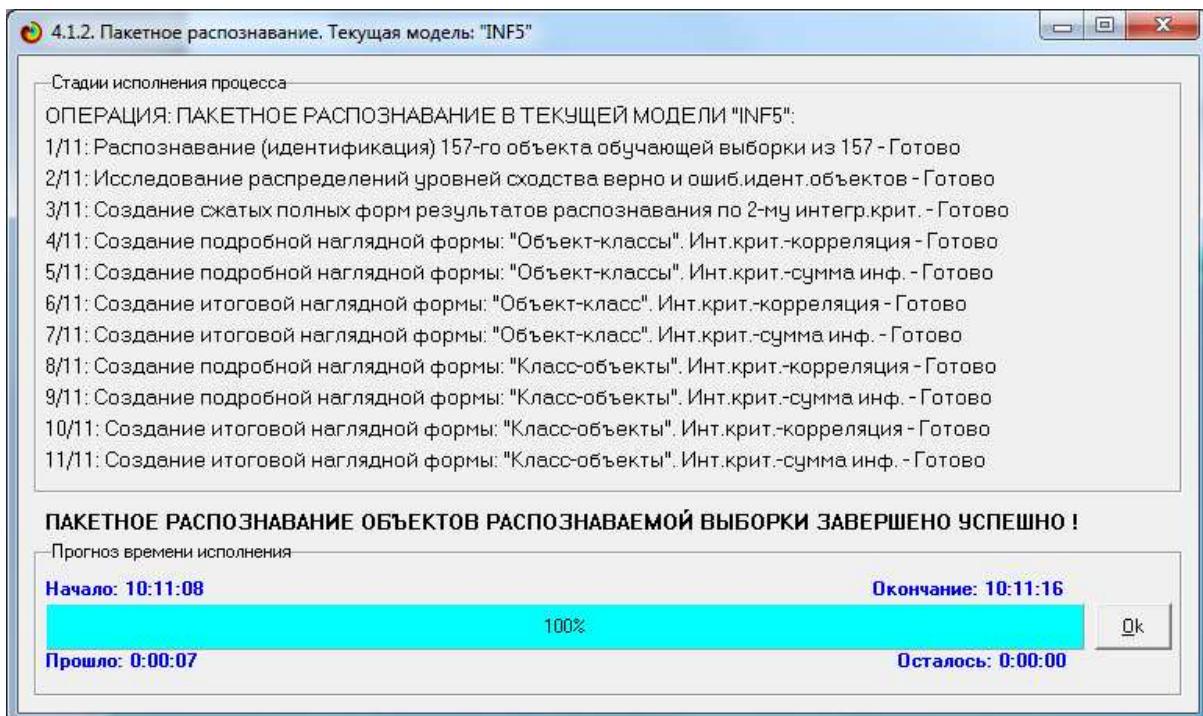


Рисунок 18- Экранная форма режима пакетного распознавания в текущей модели

В результате пакетного распознавания в текущей модели создается ряд баз данных, которые визуализируются в выходных экранах 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.

11. Распределение уровня сходства при разных моделях в интегральных критериях.

12. Объединение в одной БД строк по самым достоверным моделям.

Для наглядного примера, кратко рассмотрим некоторые из них.

На рисунках 19 и 20 приведены примеры прогнозов высокой и низкой достоверности частоты и классов уровня счастья в наиболее достоверной модели INF5 на основе описательных шкал.

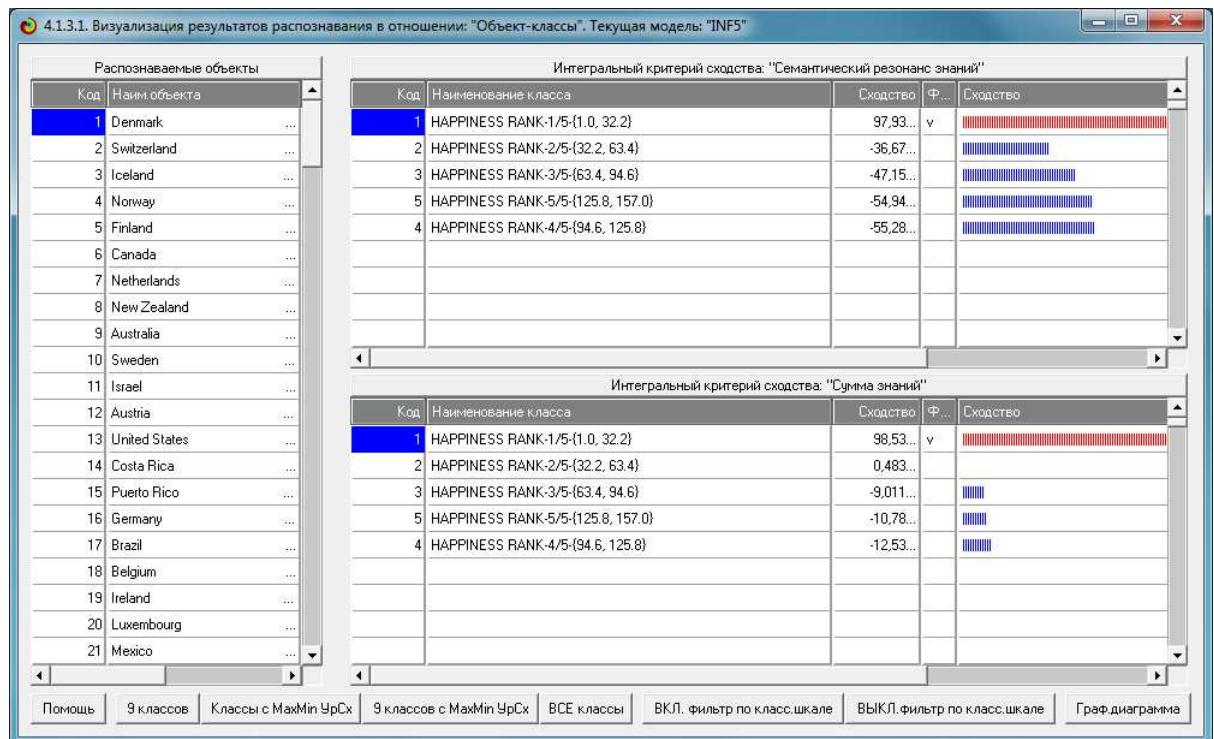


Рисунок 19- Пример идентификации классов в модели INF5

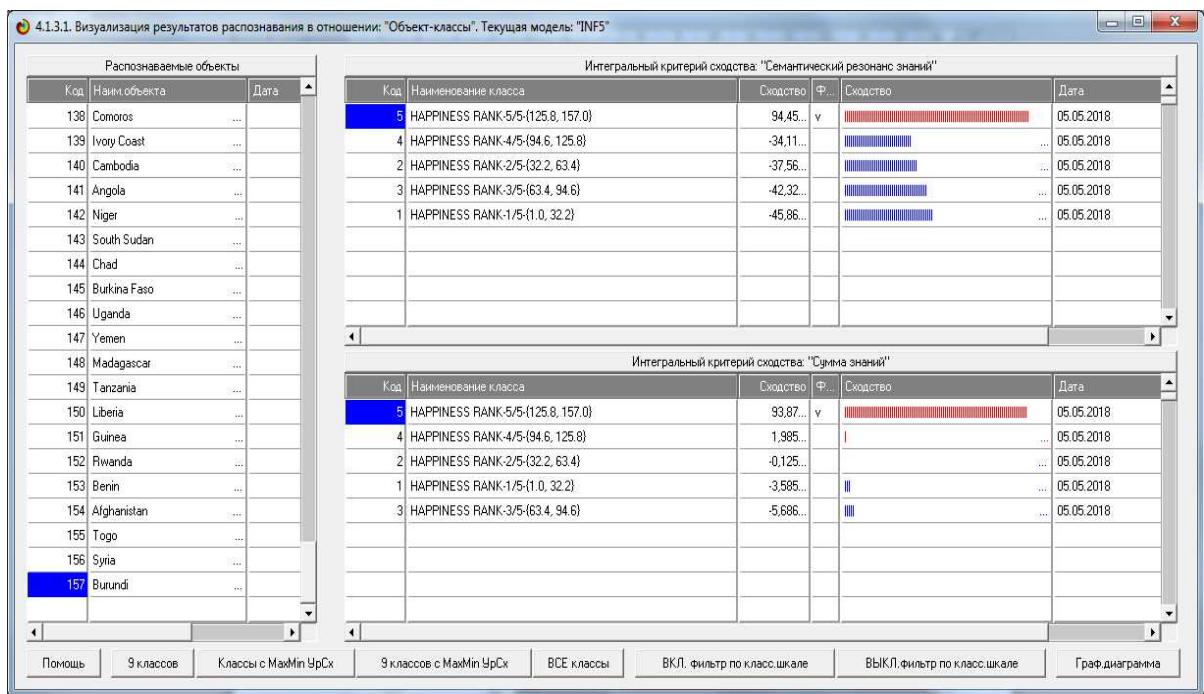


Рисунок 20- Пример идентификации классов в модели INF5

2.2 Когнитивные функции

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

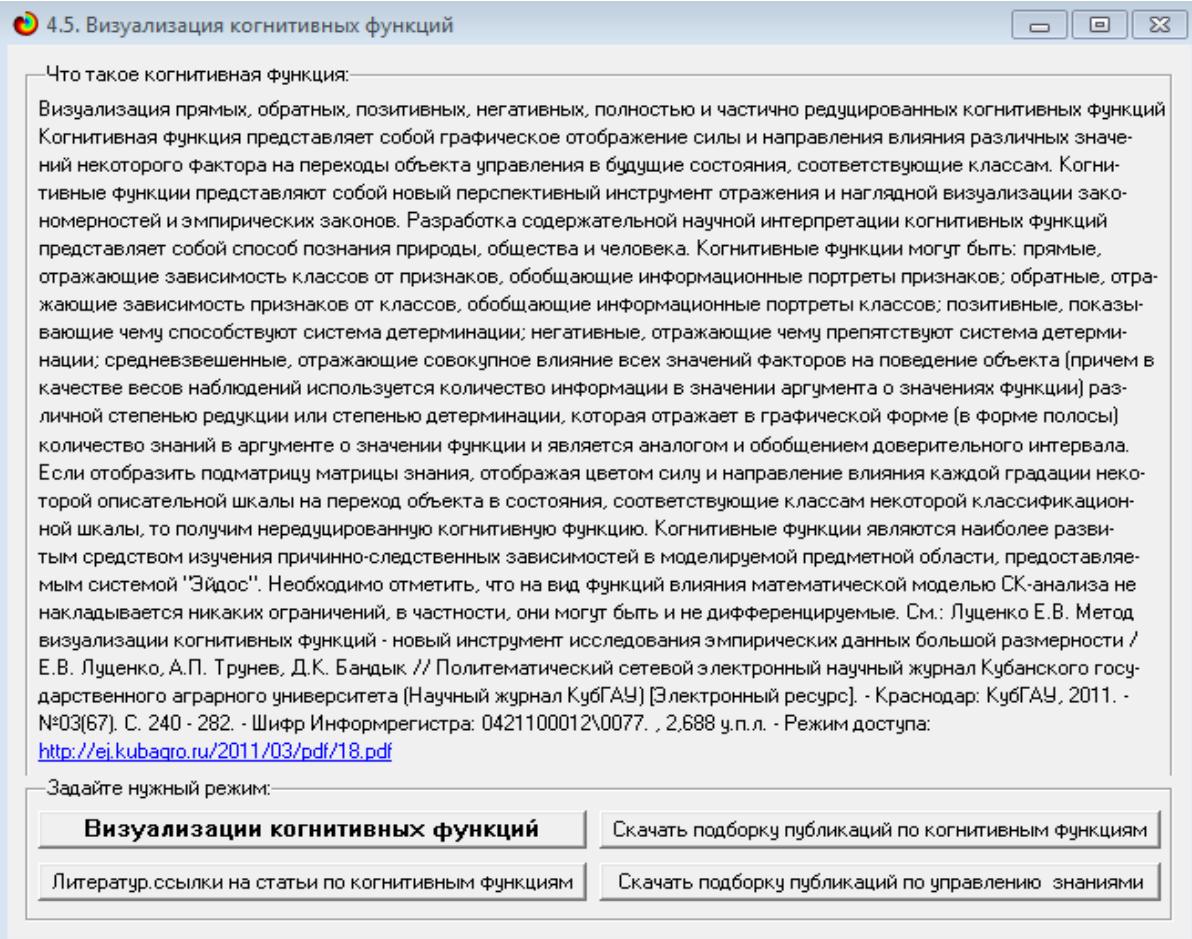


Рисунок 21- Экранная форма режима 4.5 системы «Эйдос-Х++»
«Визуализация когнитивных функций»

Применительно к задаче, рассматриваемой в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния. Когнитивным функциям посвящено множество работ, поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 22 приведены визуализации всех когнитивных функций данного приложения для модели INF5.

На рисунке 22 показано, что, чем выше уровень конфиденциальности, тем ниже уровень счастья. Люди больше реагируют на минимальном и максимальном уровне счастья, а на промежуточных значениях более нейтрально.

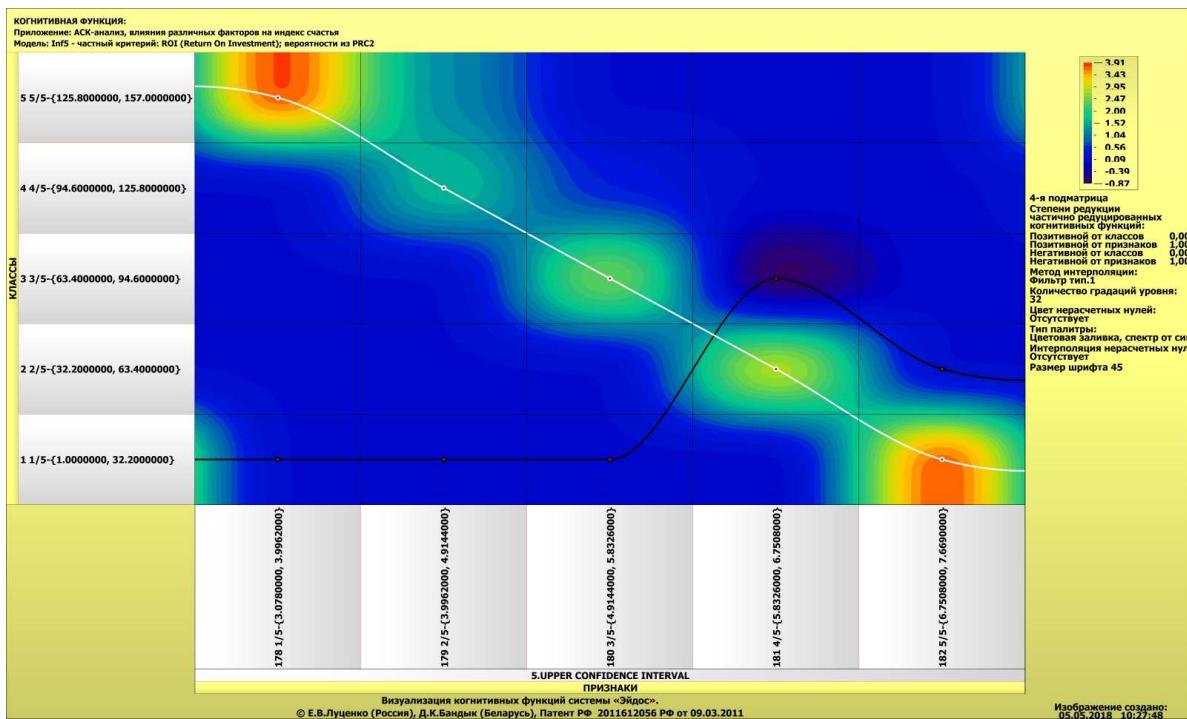


Рисунок 22-Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF5

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-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 23).

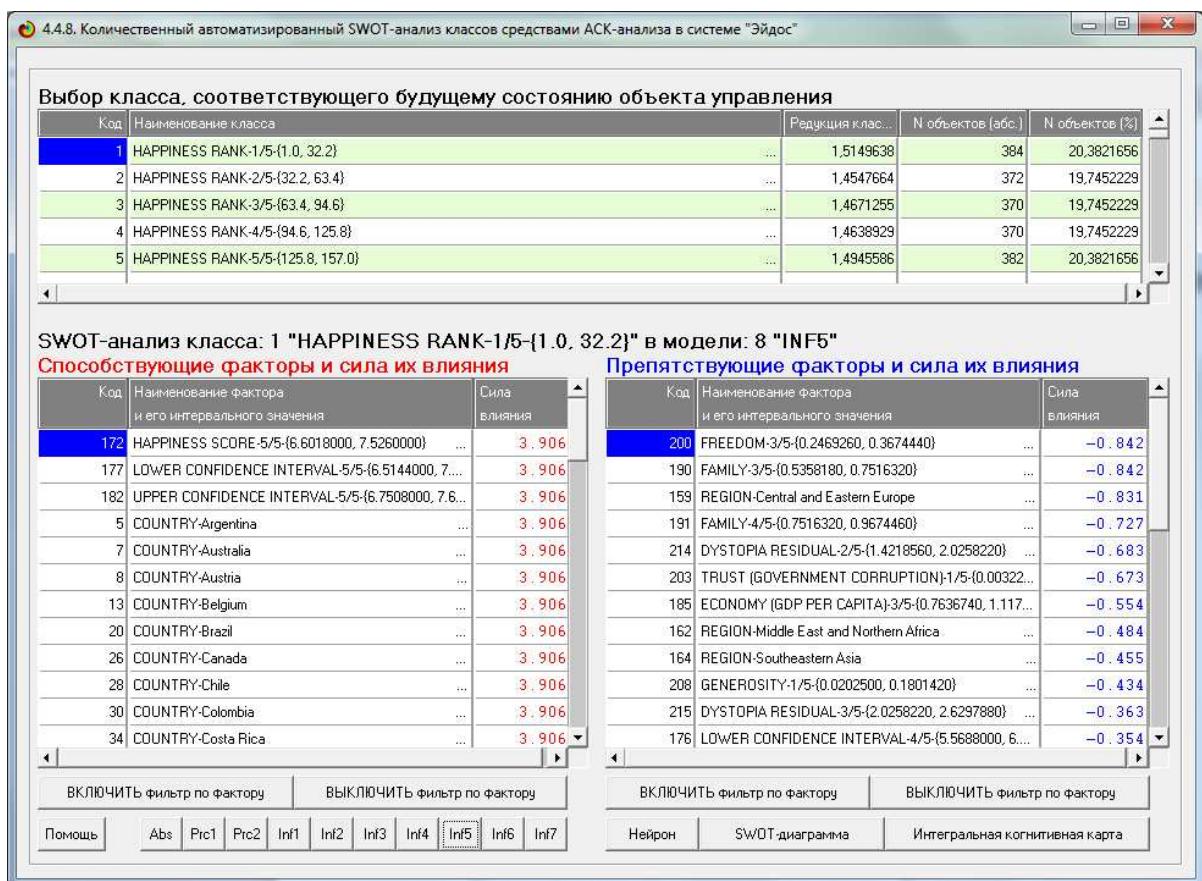


Рисунок 23-Пример SWOT-Матрицы в модели INF5

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

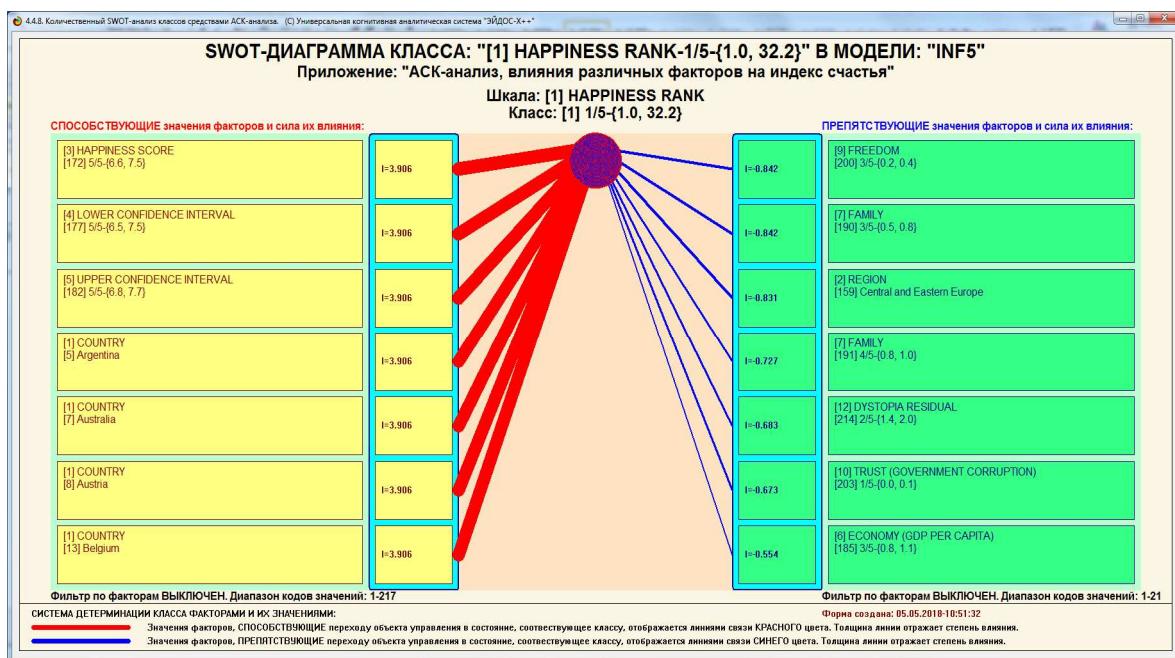


Рисунок 24— SWOT-матрица

На диаграмме показываются способствующие и препятствующие факторы. К наиболее способствующим факторам относятся:

- Индекс счастья[172]5/5-{6.6,7.5}, где $i=3.906$;
- Нижний интервал конфиденциальности[177]5/5-{6.5,7.5}, где $i=3.906$;
- Верхний интервал конфиденциальности[182]5/5-{6.8,7.7}, где $i=3.906$;
- Уровень счастья в Аргентине, где $i=3.906$;
- Уровень счастья в Австралии, где $i=3.906$;
- Уровень счастья в Австрии, где $i=3.906$;
- Уровень счастья в Бельгии, где $i=3.906$;

К наиболее препятствующим относятся:

- Уровень свободы[200]3/5-{0.2,0.4}, где $i=-0.842$;
- Семейное благополучие[190]3/5-{0.5,0.8}, где $i=-0.842$;
- Регион[159] Центральная и восточная Европа, где $i=-0.831$;
- Семейное благополучие[191]4/5-{0.8,1.0}, где $i=-0.727$;
- Уровень остаточной дистопии[214]2/5-{1.4,2.0}, где $i=-0.683$;
- Уровень доверия[203]1/5-{0.0,0.1}, где $i=-0.673$;

- Уровень экономики [185] 3/5 - {0.8, 1.1}, где $i = -0.554$.

2.4. Нелокальные нейронные сети

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

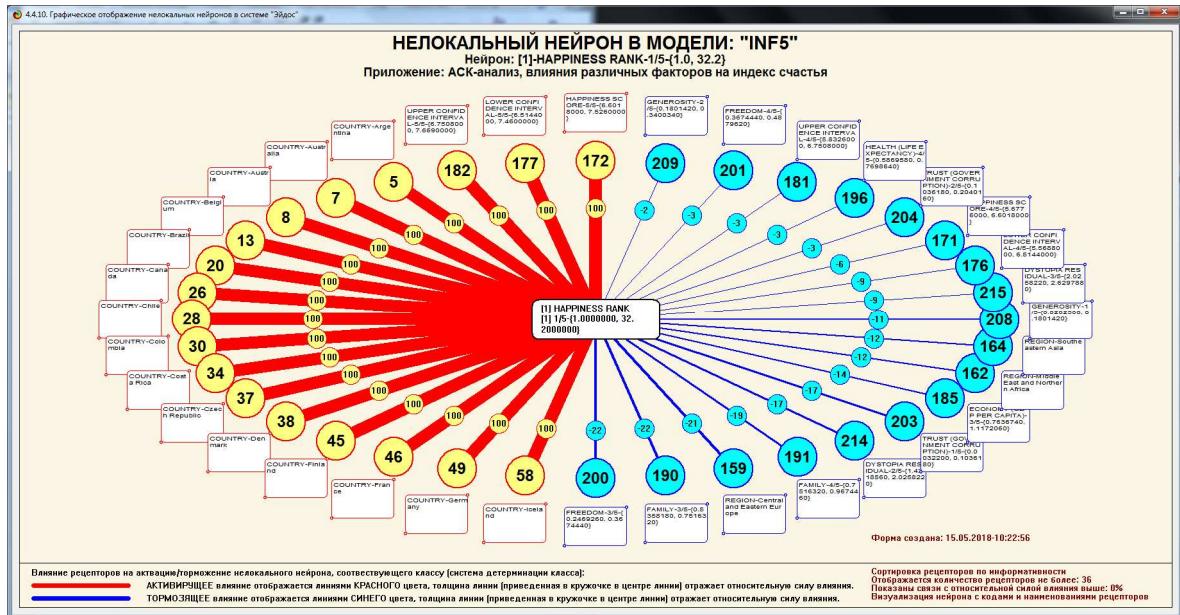


Рисунок 25- Нейрон для Happiness Rank

Благодаря данному нейрону видно, как различные факторы влияют на индекс счастья людей, какие оказывают положительное влияние, а какие отрицательное.

2.5 Кластерный и конструктивный анализ

В режиме 4.3.2.2, после расчета матриц сходства, кластеров и конструкторов, строим 2D сеть классов в выбранной модели знаний, для наглядного представления сети классов. На рисунке 26 наглядно видно, что наибольшее сходство наблюдается в классе «Burkina Faso», когда как наибольшее различие наблюдается в классе «Dystopia residual».

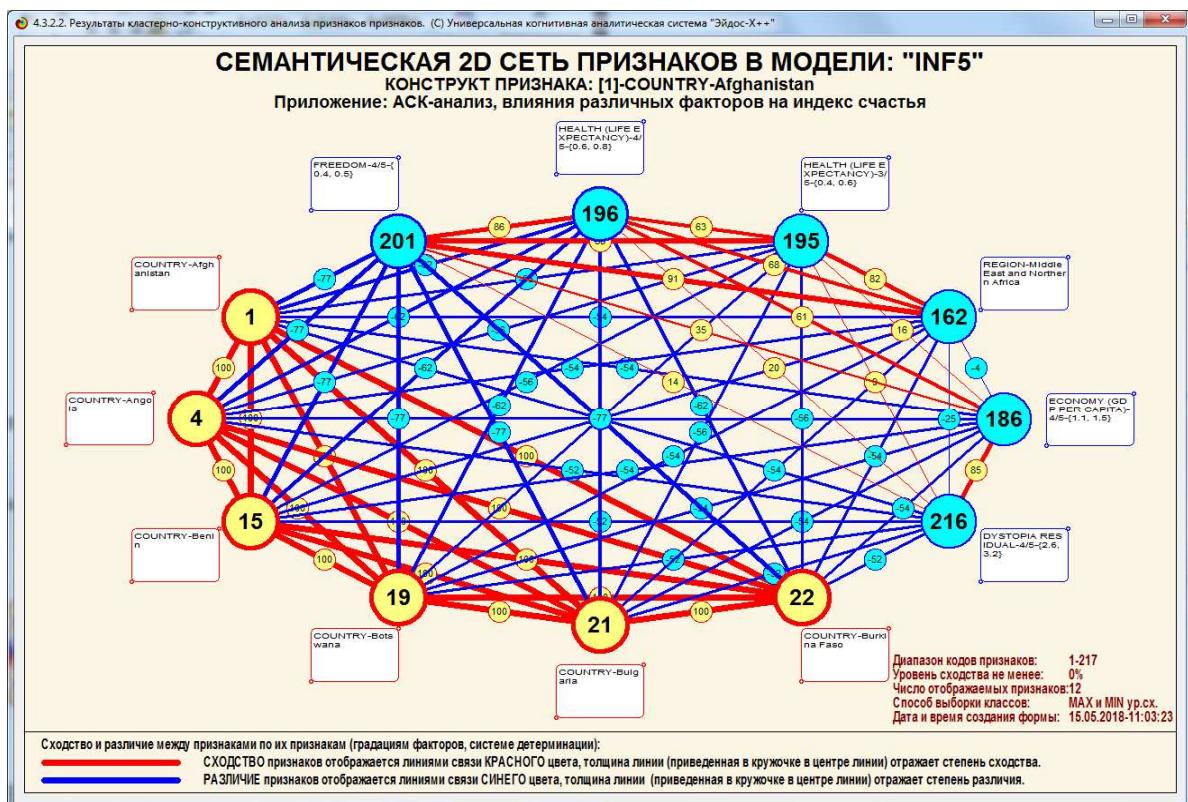


Рисунок 26– Семантическая 2D сеть классов

На рисунках 27 и 28 приведены дендрограмма агломеративной когнитивной кластеризации классов и график изменения межкластерных расстояний:

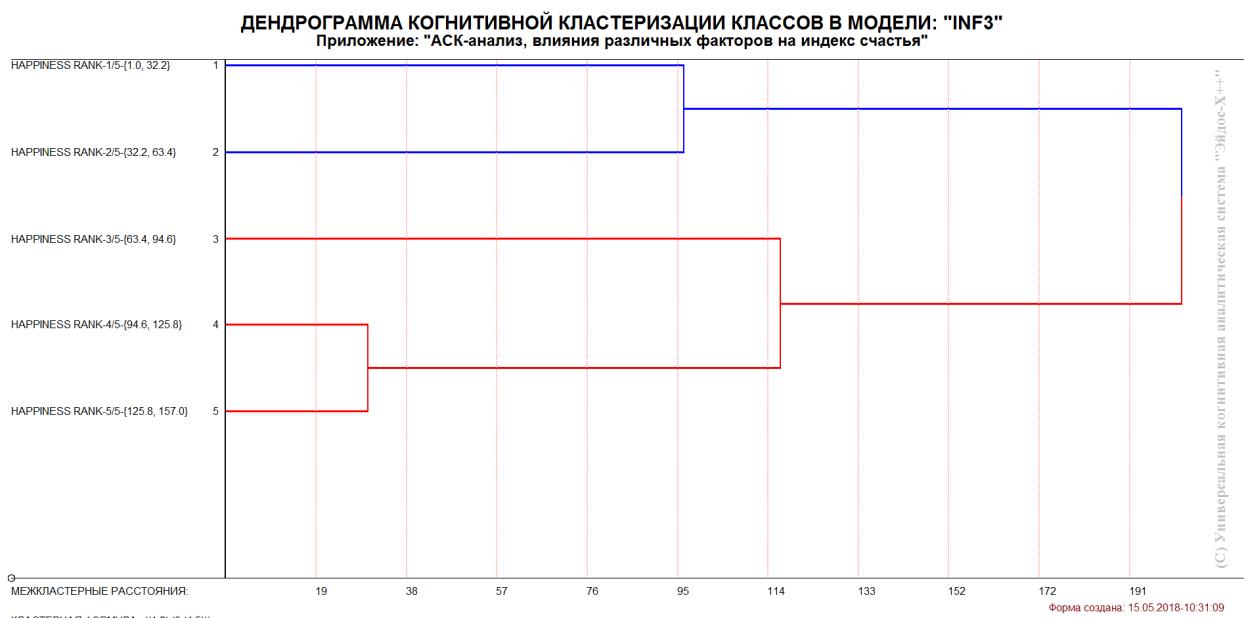


Рисунок 27– Дендрограмма агломеративной когнитивной кластеризации классов

ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ КЛАССОВ В МОДЕЛИ: "INF3"
Приложение: "ACK-анализ, влияния различных факторов на индекс счастья"

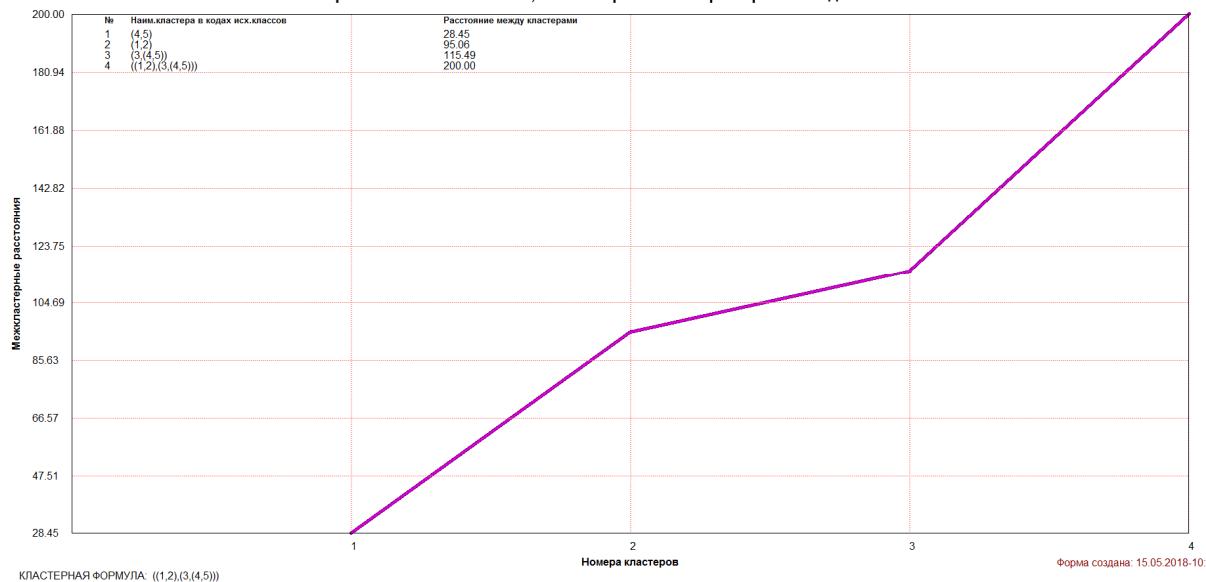


Рисунок 28– График изменения межклusterных расстояний

ЗАКЛЮЧЕНИЕ

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

В данной лабораторной приводится развернутый пример использования базы данных репозитория «Kaggle» для оценки качества математических моделей, применяемых в АСК-анализе и его программном инструментарии 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 0,916, 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.

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

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Методика использования репозитория UCI для оценки качества математических моделей систем искусственного интеллекта / Е.В. Луценко // Политетматический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: http://lc.kubagro.ru/My_training_schedule.doc КубГАУ, 2003. – №02(002). С. 120 – 145. – IDA [article ID]: 0020302012. – Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 у.п.л.
2. Луценко Е.В. ACK-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политетматический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №06(100). С. 1346 – 1395. – IDA [article ID]: 1001406090. – Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/90.pdf>, 3,125 у.п.л.
3. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политетматический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №04(088). С. 340 – 359. – IDA [article ID]: 0881304022. – Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.
4. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических,

технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.

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

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