МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»**

Факультет прикладной информатики

Кафедра компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии
на тему: АСК-анализ доклада о мировом счастье до 2020 года на основе данных портала Kaggle
Выполнил студент группы: ИТ1941 Аветисян Эдуард Альбертович
Допущен к защите: _____

Руководитель проекта: д.э.н., к. т. н., профессор Луценко Е.В. (_____)
(подпись, расшифровка подписи)

Защищен_____15/02/2021_____
(дата)
Оценка_____отлично_____

Краснодар 2021

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 курсовую работу

Студента Аветисяна Эдуарда Альбертовича
курса 2 очной формы обучения группы ИТ1941
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ доклада о мировом счастье до 2020 года на основе данных портала Kaggle»
Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	*Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)*
1	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	*5*
2	Уровень и корректность использования в работе различных методов исследований	*5*
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	*5*
4	Ясность, четкость, последовательность и обоснованность изложения	*5*
5	Применение современных технологий обработки информации	*5*
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	*5*
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	*5*
8	Ответы на вопросы при защите	*5*

Достоинства работы _____
_____
_____

Недостатки работы _____
_____
_____

Итоговая оценка при защите _____
Рецензент _____ (Е.В. Луценко)
«_____» _____2021 г.

РЕФЕРАТ

Курсовая работа содержит: 35 страниц, 29 рисунков, 17 литературных источников.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, ШКАЛЫ, КЛАССЫ.

Целью работы провести АСК-анализ доклада о мировом счастье до 2020 года на основе данных портала Kaggle. В данной курсовой работе необходимо 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.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 5

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ 6

1.1. Описание решения 6

1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel 8

1.3. Ввод выборки в систему Aidos-X 9

1.4. Синтез и верификация статистических и интеллектуальных моделей 14

1.5. Виды моделей системы Aidos-X 15

1.6. Результаты верификации моделей 17

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ 20

2.1. Решение задачи идентификации 20

2.2. Кластерно-конструктивный анализ 24

2.3. Нелокальные нейронные сети и нейтроны 27

2.4. SWOT и PEST матрицы и диаграммы 29

ЗАКЛЮЧЕНИЕ 32

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ 34

ВВЕДЕНИЕ

Одним из важных и перспективных направлений развития современных информационных технологий является создание систем искусственного интеллекта. Существует множество альтернатив систем искусственного интеллекта, но возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи АСК-анализ доклада о мировом счастье до 2020 года на основе данных портала Kaggle. 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 «World Happiness Report up to 2020» из банка исходных данных по задачам искусственного интеллекта – Kaggle.

Для решения задачи используем стандартные возможности MS Office Word и Excel, а также систему искусственного интеллекта "Aidos-X++".

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

В качестве метода исследования, решения проблемы и достижения цели было решено использовать новый метод искусственного интеллекта: Автоматизированный системно-когнитивный анализ или АСК-анализ.

Основной причиной выбора АСК-анализа является то, что он включает теорию и метод количественного выявления в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал и единицах измерения.

Очень важным является также то, что АСК-анализ имеет свой развитый и доступный программный инструментарий, в качестве которого в настоящее время выступает Универсальная когнитивная аналитическая система Aidos-X.

Преимущества данной системы:

- универсальность;
- доступность, то есть данная система находится в полном открытом бесплатном доступе, причем с актуальными исходными текстам;
- одна из первых отечественных систем искусственного интеллекта персонального уровня, т.е. она не требует от пользователя специальной подготовки в области технологий искусственного интеллекта (акт внедрения системы Aidos-X 1987 года);
- обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения;

- содержит большое количество локальных и облачных учебных и научных приложений (в настоящее время их 31 и 251, соответственно);
- мультиязычная поддержка интерфейса (больше 50 языков);
- поддерживает online среду накопления знаний и широко используется во всем мире;
- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз;
- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение задач классификации, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе;
- хорошо имитирует человеческий стиль мышления: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции.

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-X.

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 набор данных «World Happiness Report up to 2020»

https://www.kaggle.com/mathurinache/world-happiness-report

Csv файл содержит следующие данные:

Country name – Страна.

Regional indication – Регион страны.

Ladder score – Коэфициент сравниваемой позиции в таблице.

Standard error of ladder score Стандартные отклонения от сравниваемой позиции.

Upper whisker Верхняя граница.

Lower whisker Нижняя граница.

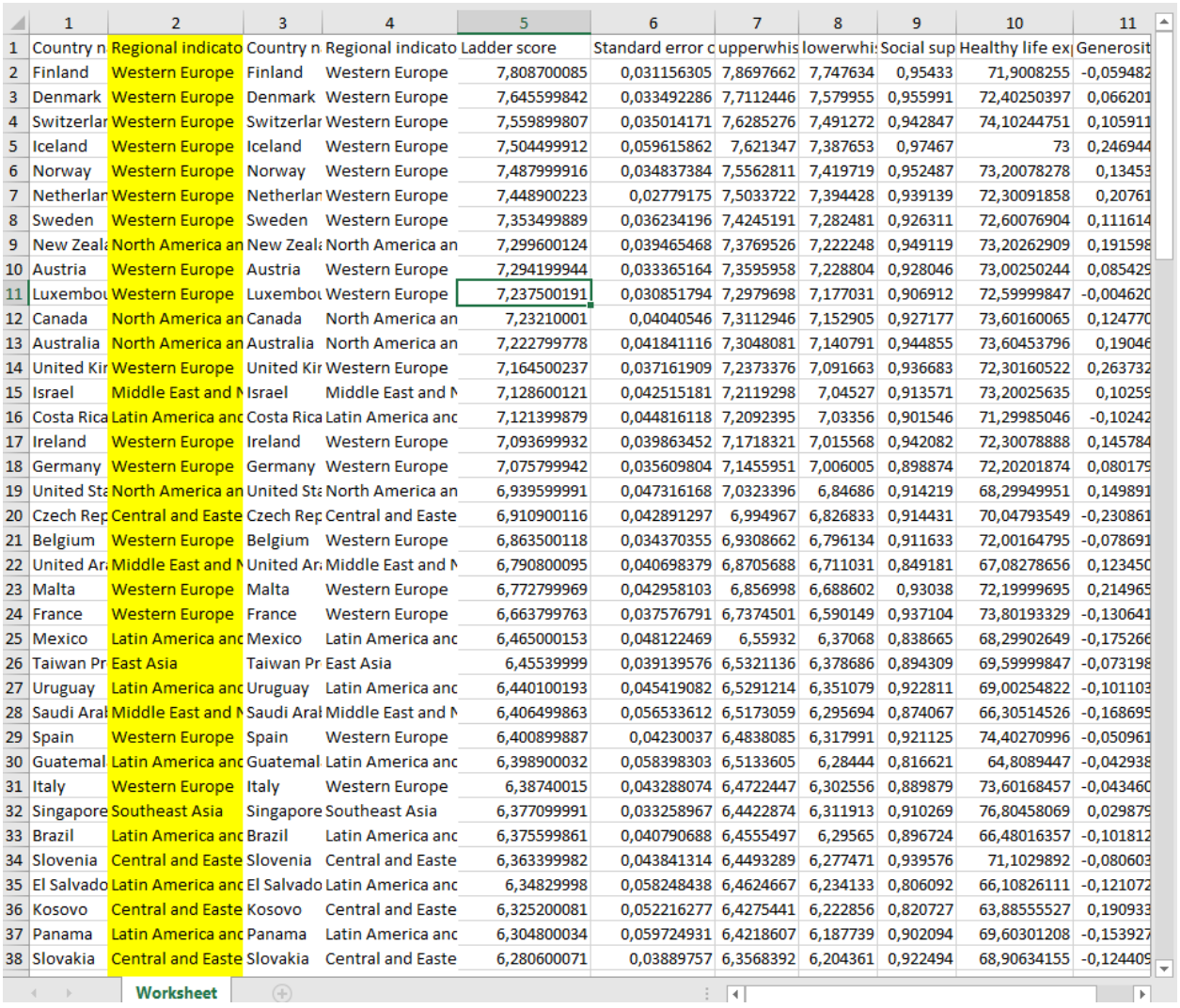
Social support Социальная поддержка.

Healthy life expectancy Ожидаемая продолжительность здоровой жизни.

Generosity Щедрость.

Для конвертации csv-файла в xls был использован онлайн конвертер: https://document.online-convert.com/ru/convert/csv-to-excel

После конвертации необходимо добавить еще один столбец, в данном случае было решено дублировать столбец, содержащий названия регионов. Также данный столбец был переименован в «Regional indication» и размещен в таблице между столбцами «Country name» и «Ladder score» (Рисунок 1).



	1	2	3	4	5	6	7	8	9	10	11
1	Country n	Regional indicato	Country n	Regional indicato	Ladder score	Standard error c	upperwhis	lowerwhi	Social sup	Healthy life ex	Generosit
2	Finland	Western Europe	Finland	Western Europe	7,808700085	0,031156305	7,8697662	7,747634	0,95433	71,9008255	-0,059482
3	Denmark	Western Europe	Denmark	Western Europe	7,645599842	0,033492286	7,7112446	7,579955	0,955991	72,40250397	0,066201
4	Switzerlar	Western Europe	Switzerlar	Western Europe	7,559899807	0,035014171	7,6285276	7,491272	0,942847	74,10244751	0,105911
5	Iceland	Western Europe	Iceland	Western Europe	7,504499912	0,059615862	7,621347	7,387653	0,97467	73	0,246944
6	Norway	Western Europe	Norway	Western Europe	7,487999916	0,034837384	7,5562811	7,419719	0,952487	73,20078278	0,13453
7	Netherlar	Western Europe	Netherlar	Western Europe	7,448900223	0,02779175	7,5033722	7,394428	0,939139	72,30091858	0,20761
8	Sweden	Western Europe	Sweden	Western Europe	7,353499889	0,036234196	7,4245191	7,282481	0,926311	72,60076904	0,111614
9	New Zeal	North America an	New Zeal	North America an	7,299600124	0,039465468	7,3769526	7,222248	0,949119	73,20262909	0,191598
10	Austria	Western Europe	Austria	Western Europe	7,294199944	0,033365164	7,3595958	7,228804	0,928046	73,00250244	0,085429
11	Luxembou	Western Europe	Luxembou	Western Europe	7,237500191	0,030851794	7,2979698	7,177031	0,906912	72,59999847	-0,004620
12	Canada	North America an	Canada	North America an	7,23210001	0,04040546	7,3112946	7,152905	0,927177	73,60160065	0,124770
13	Australia	North America an	Australia	North America an	7,222799778	0,041841116	7,3048081	7,140791	0,944855	73,60453796	0,19046
14	United Kir	Western Europe	United Kir	Western Europe	7,164500237	0,037161909	7,2373376	7,091663	0,936683	72,30160522	0,263732
15	Israel	Middle East and N	Israel	Middle East and N	7,128600121	0,042515181	7,2119298	7,04527	0,913571	73,20025635	0,10259
16	Costa Rica	Latin America anc	Costa Rica	Latin America anc	7,121399879	0,044816118	7,2092395	7,03356	0,901546	71,29985046	-0,10242
17	Ireland	Western Europe	Ireland	Western Europe	7,093699932	0,039863452	7,1718321	7,015568	0,942082	72,30078888	0,145784
18	Germany	Western Europe	Germany	Western Europe	7,075799942	0,035609804	7,1455951	7,006005	0,898874	72,20201874	0,080179
19	United Sta	North America an	United Sta	North America an	6,939599991	0,047316168	7,0323396	6,84686	0,914219	68,29949951	0,149891
20	Czech Rep	Central and Easte	Czech Rep	Central and Easte	6,910900116	0,042891297	6,994967	6,826833	0,914431	70,04793549	-0,230861
21	Belgium	Western Europe	Belgium	Western Europe	6,863500118	0,034370355	6,9308662	6,796134	0,911633	72,00164795	-0,078691
22	United Ar	Middle East and N	United Ar	Middle East and N	6,790800095	0,040698379	6,8705688	6,711031	0,849181	67,08278656	0,123450
23	Malta	Western Europe	Malta	Western Europe	6,772799969	0,042958103	6,856998	6,688602	0,93038	72,19999695	0,214965
24	France	Western Europe	France	Western Europe	6,663799763	0,037576791	6,7374501	6,590149	0,937104	73,80193329	-0,130641
25	Mexico	Latin America anc	Mexico	Latin America anc	6,465000153	0,048122469	6,55932	6,37068	0,838665	68,29902649	-0,175266
26	Taiwan Pr	East Asia	Taiwan Pr	East Asia	6,45539999	0,039139576	6,5321136	6,378686	0,894309	69,59999847	-0,073198
27	Uruguay	Latin America anc	Uruguay	Latin America anc	6,440100193	0,045419082	6,5291214	6,351079	0,922811	69,00254822	-0,101103
28	Saudi Ara	Middle East and N	Saudi Ara	Middle East and N	6,406499863	0,056533612	6,5173059	6,295694	0,874067	66,30514526	-0,168695
29	Spain	Western Europe	Spain	Western Europe	6,400899887	0,04230037	6,4838085	6,317991	0,921125	74,40270996	-0,050961
30	Guatemal	Latin America anc	Guatemal	Latin America anc	6,398900032	0,058398303	6,5133605	6,28444	0,816621	64,8089447	-0,042938
31	Italy	Western Europe	Italy	Western Europe	6,38740015	0,043288074	6,4722447	6,302556	0,889879	73,60168457	-0,043460
32	Singapore	Southeast Asia	Singapore	Southeast Asia	6,377099991	0,033258967	6,4422874	6,311913	0,910269	76,80458069	0,029879
33	Brazil	Latin America anc	Brazil	Latin America anc	6,375599861	0,040790688	6,4555497	6,29565	0,896724	66,48016357	-0,101812
34	Slovenia	Central and Easte	Slovenia	Central and Easte	6,363399982	0,043841314	6,4493289	6,277471	0,939576	71,1029892	-0,080603
35	El Salvado	Latin America anc	El Salvado	Latin America anc	6,34829998	0,058248438	6,4624667	6,234133	0,806092	66,10826111	-0,121072
36	Kosovo	Central and Easte	Kosovo	Central and Easte	6,325200081	0,052216277	6,4275441	6,222856	0,820727	63,88555527	0,190933
37	Panama	Latin America anc	Panama	Latin America anc	6,304800034	0,059724931	6,4218607	6,187739	0,902094	69,60301208	-0,153927
38	Slovakia	Central and Easte	Slovakia	Central and Easte	6,280600071	0,03889757	6,3568392	6,204361	0,922494	68,90634155	-0,124409

Worksheet

Рисунок 1 - Фрагмент обучающей выборки

Таким образом, в качестве классификационной шкалы решено использовать столбец 2. Описательные шкалы столбцы 3-11.

1.3. Ввод выборки в систему Aidos-X

Теперь, когда имеется обучающая выборка в формате *.xls, можно импортировать ее в систему Aidos-X.

Скопируем данную выборку в папку Inp_data и переименуем ее в Inp_data. Далее был использован универсальный программный интерфейс импорта данных в систему Aidos-X (режим 2.3.2.2), результат заполнения данной формы представлен на рисунке 2.

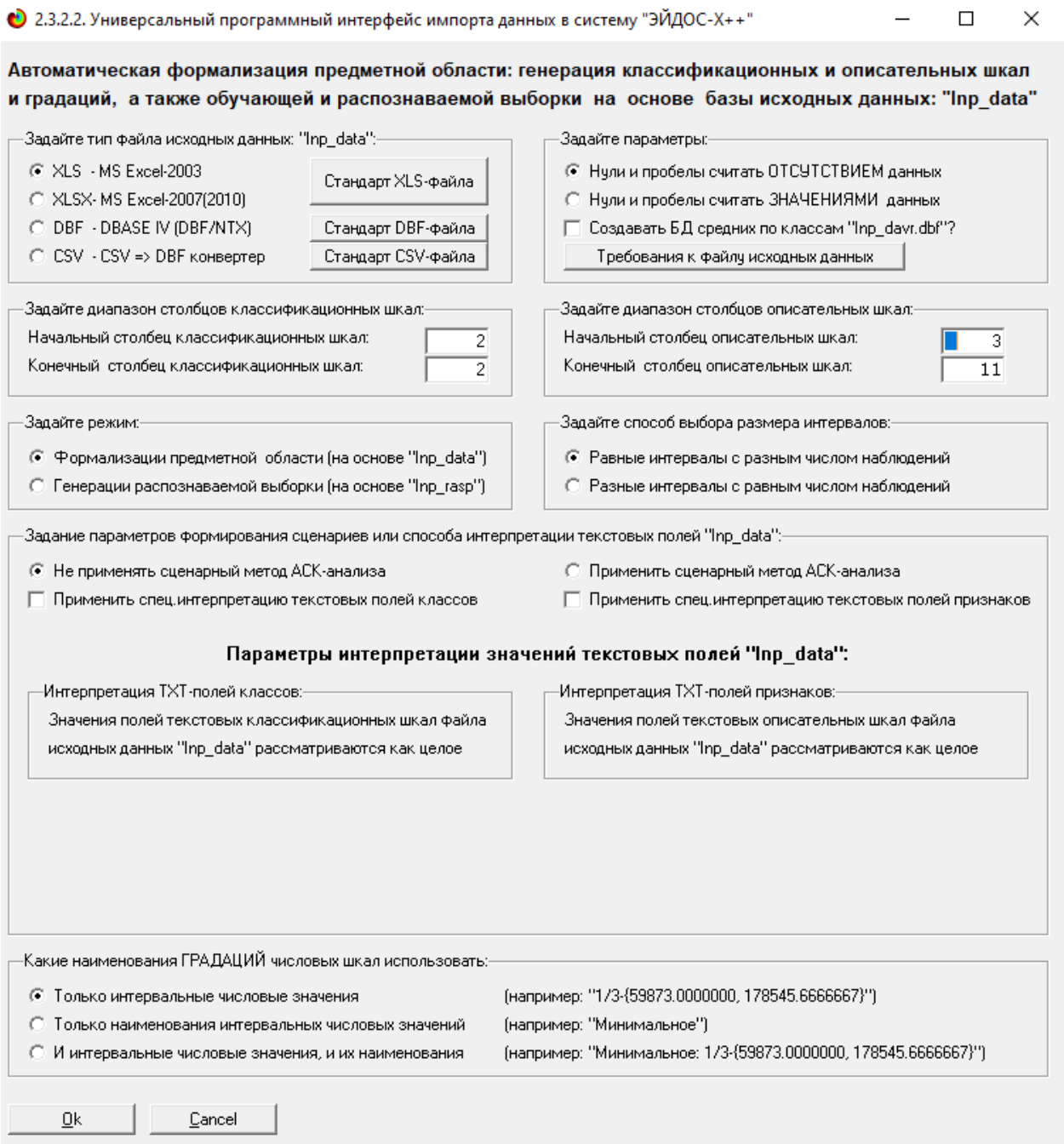


Рисунок 2 - Ввод обучающей выборки

Стоит отметить следующие настройки:

- Тип файла – xls;
- Классификационные шкалы – 2;
- Описательные шкалы – 3-11;
- Применяется спец. интерпцретация текстовых полей признаков;

- В качестве признаков рассматриваются слова, без проведения лемитизации.

В форме задания размерности модели системы оставляем всё без изменений (Рисунок 3).

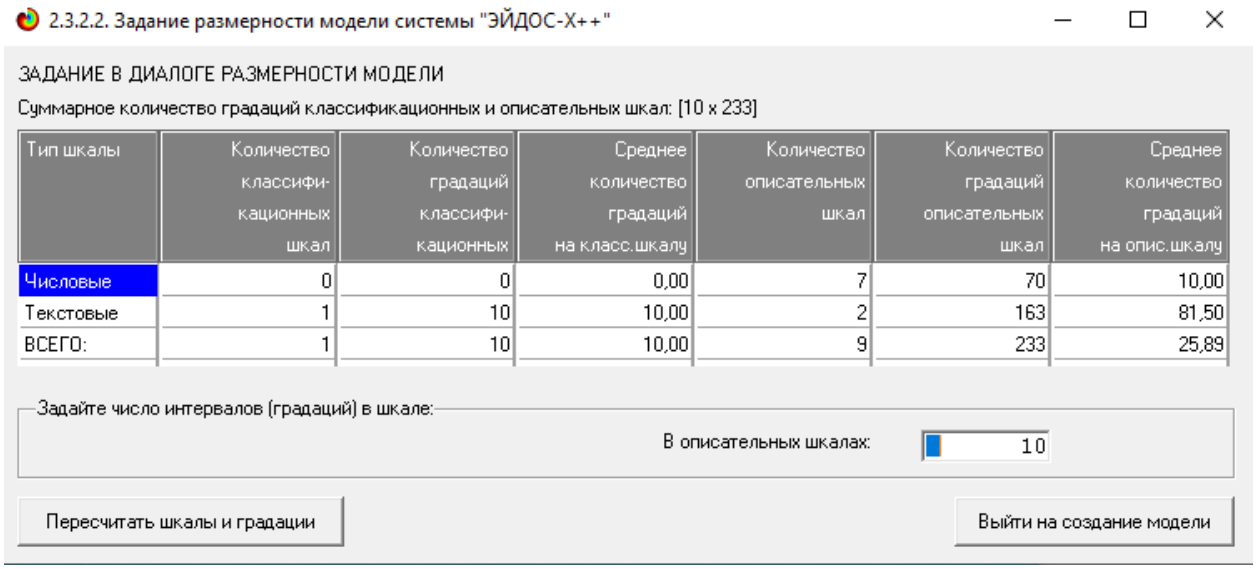


Рисунок 3 - Задание размерности модели системы Aidos-X

Процесс импорта данных из внешней БД "Inp_data.xls" в систему Aidos-X представлен на рисунке 4.

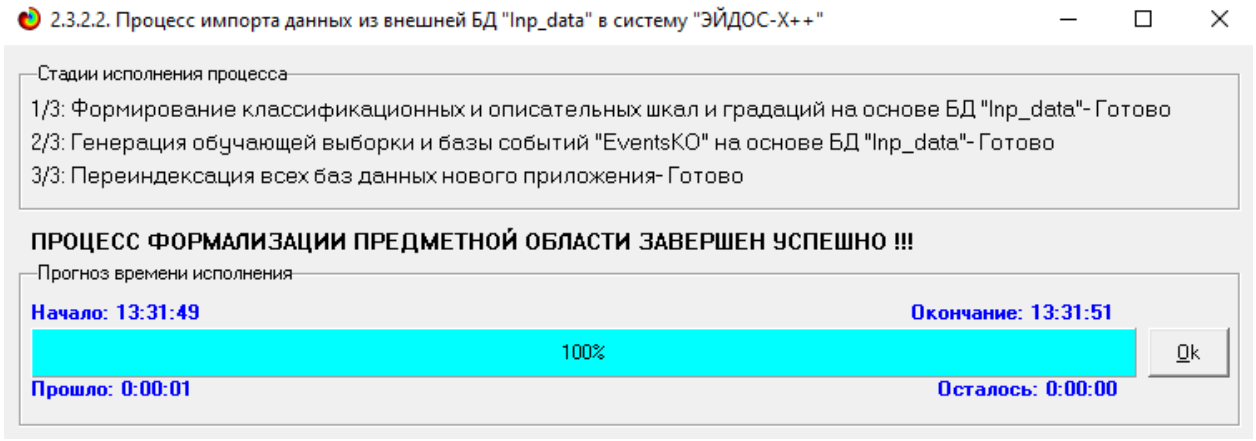


Рисунок 4 – Процесс импорта данных

Так как предварительно числовые шкалы были разбиты на интервалы, то 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. Тем самым этап формализации предметной области выполняется полностью автоматизировано.

Классификационная шкала представлена на рисунке 5, её можно просмотреть в режиме 2.1. Описательные шкалы – в режиме 2.2 (рисунок 6).

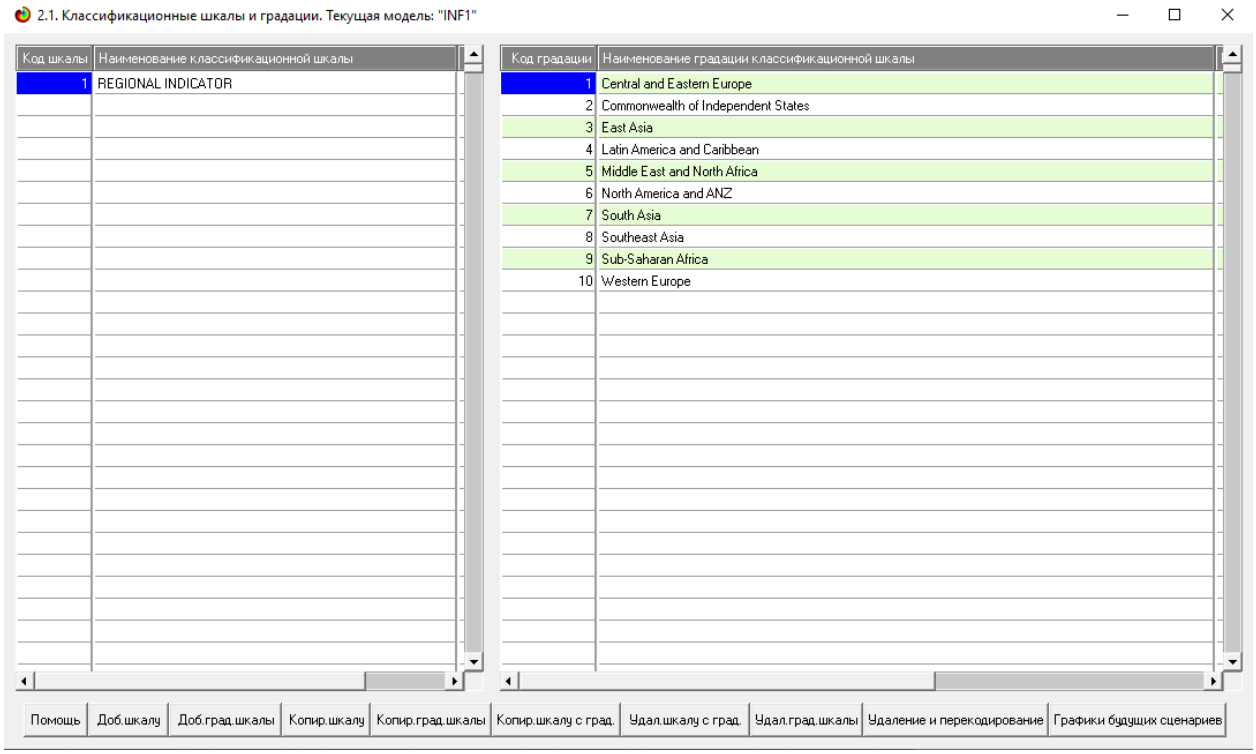


Рисунок 5 - Классификационные шкалы и градации

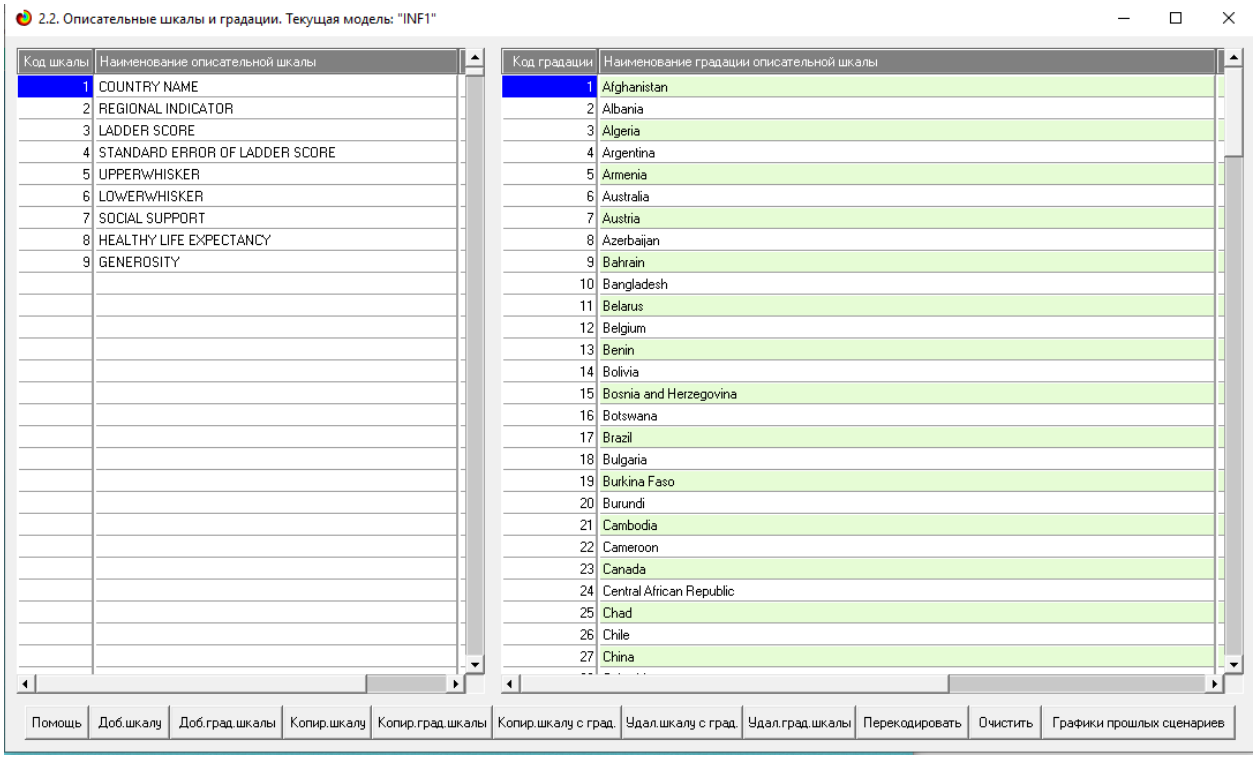


Рисунок 6 - Описательные шкалы и градации (фрагмент)

Для ручного ввода-корректировки обучающей выборки существует режим 2.3.1, он представлен на рисунке 7. Установка значений описательных и классификационных шкал объектов осуществляется по их номерам.

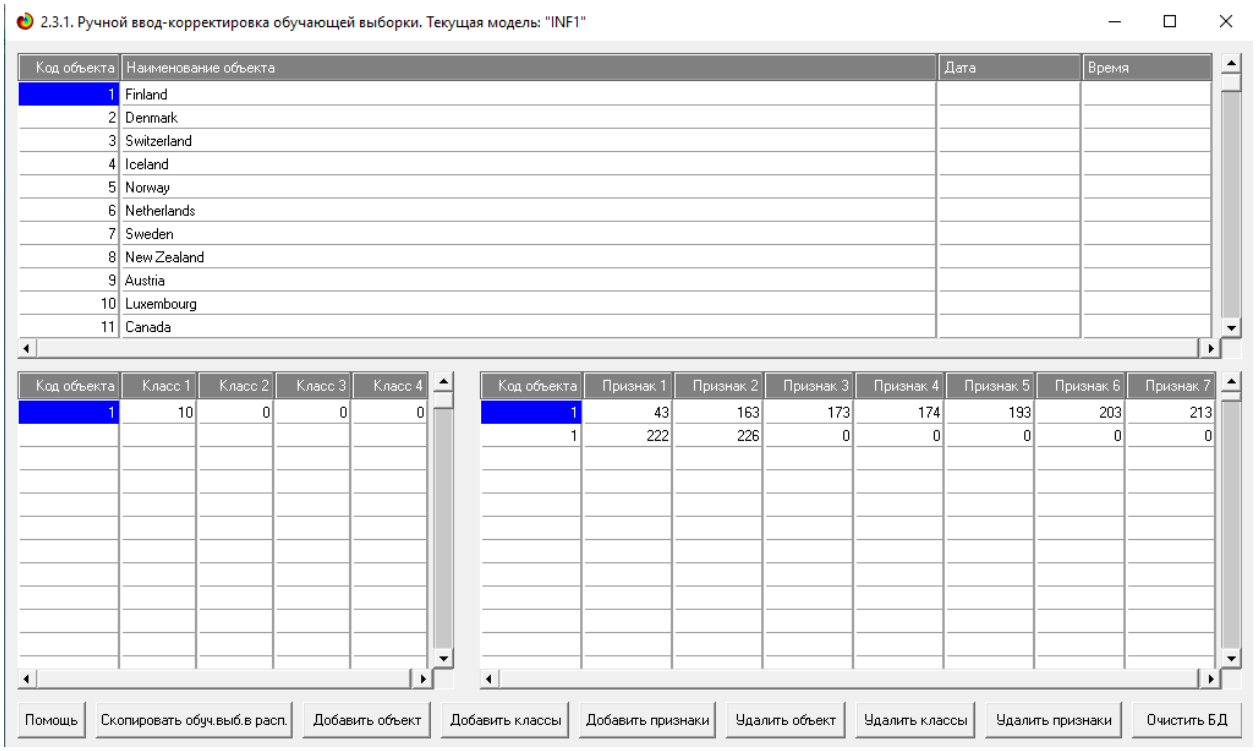


Рисунок 7 - Обучающая выборка (фрагмент)

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 (с учетом нелинейности системы).

1.4. 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 8).

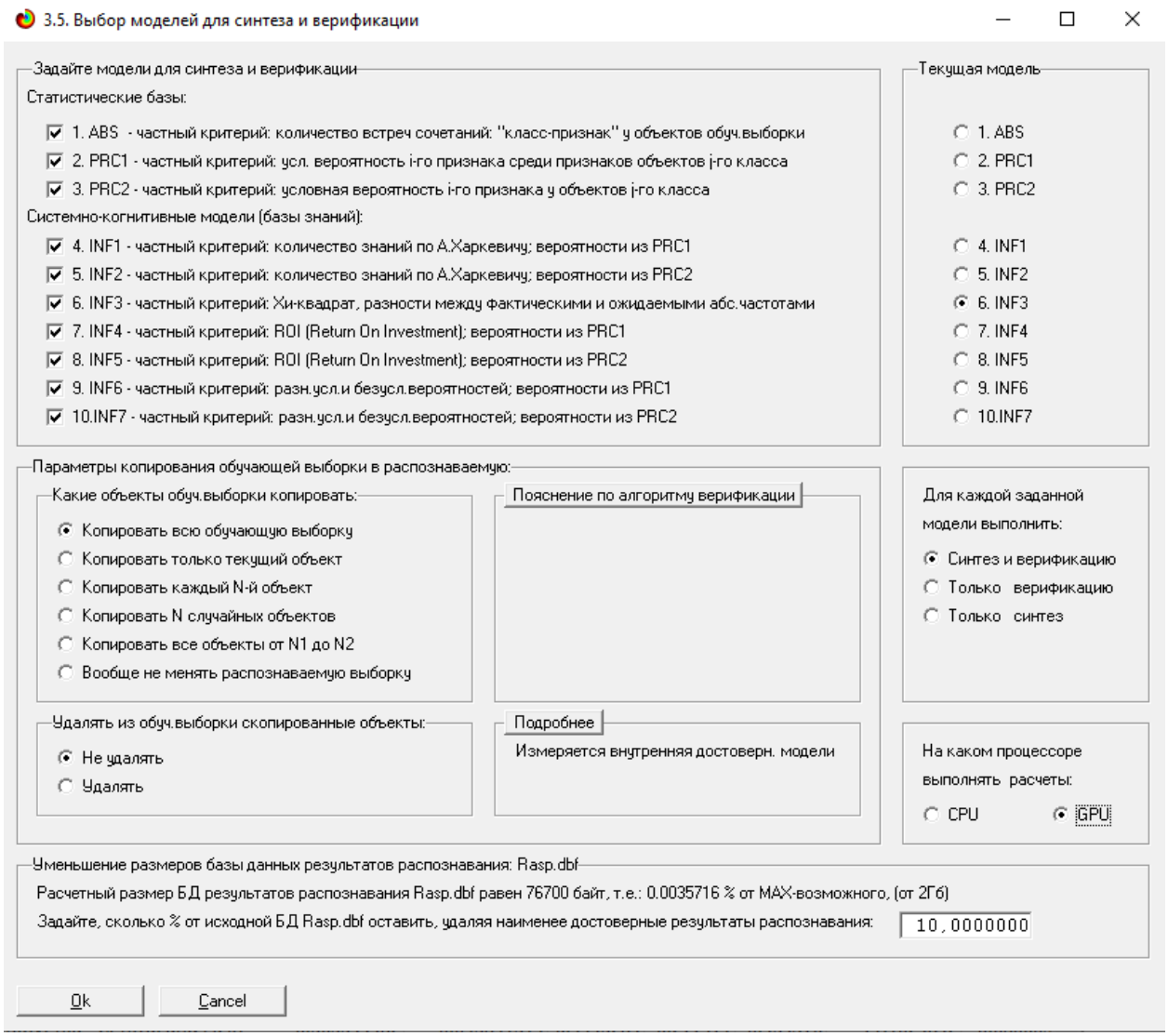


Рисунок 8 - Выбор моделей для синтеза и верификации

В данном режиме имеется много различных методов верификации моделей. Но мы используем параметры, приведенные на рисунке 8. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

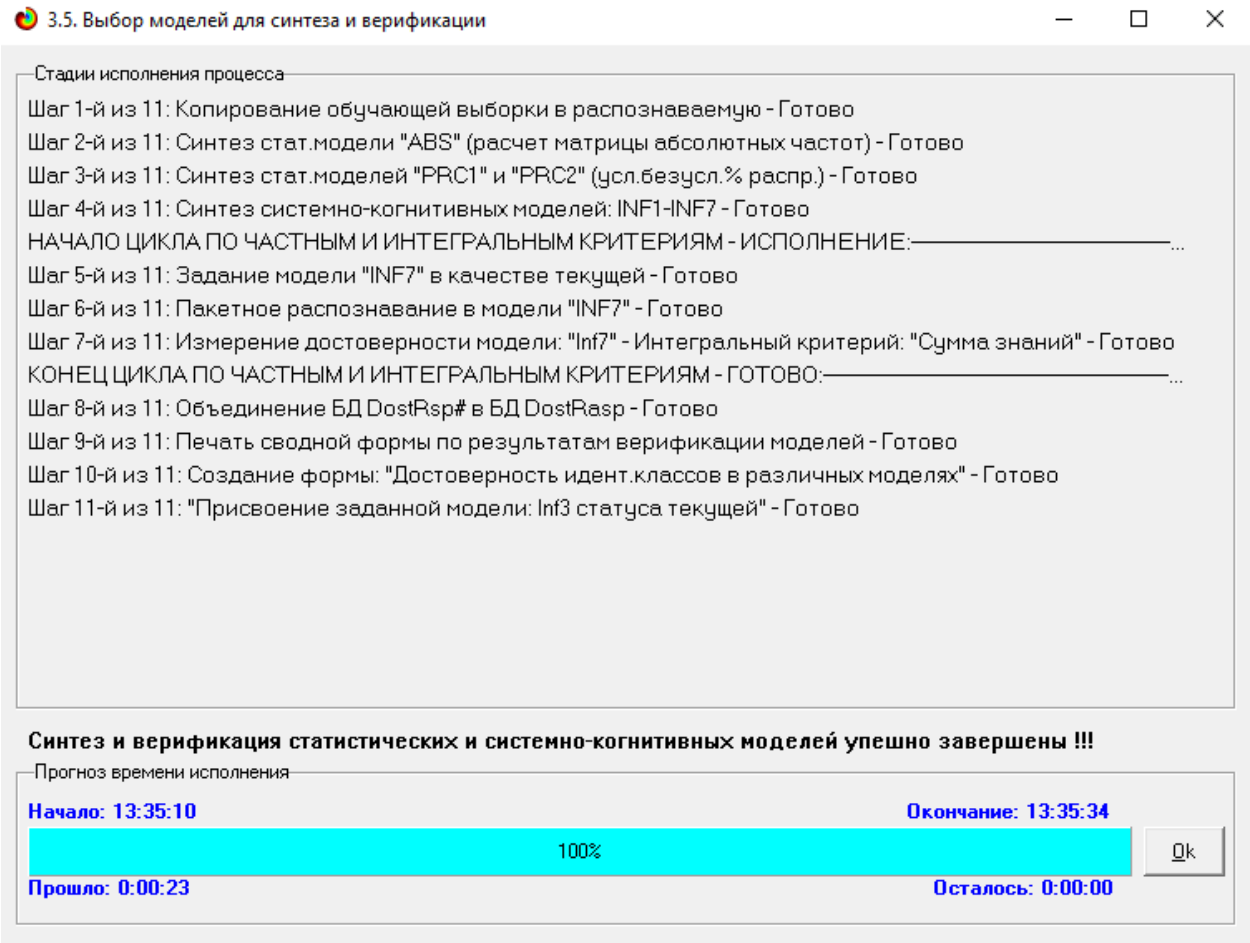


Рисунок 8 - Процесс исполнения режима 3.5

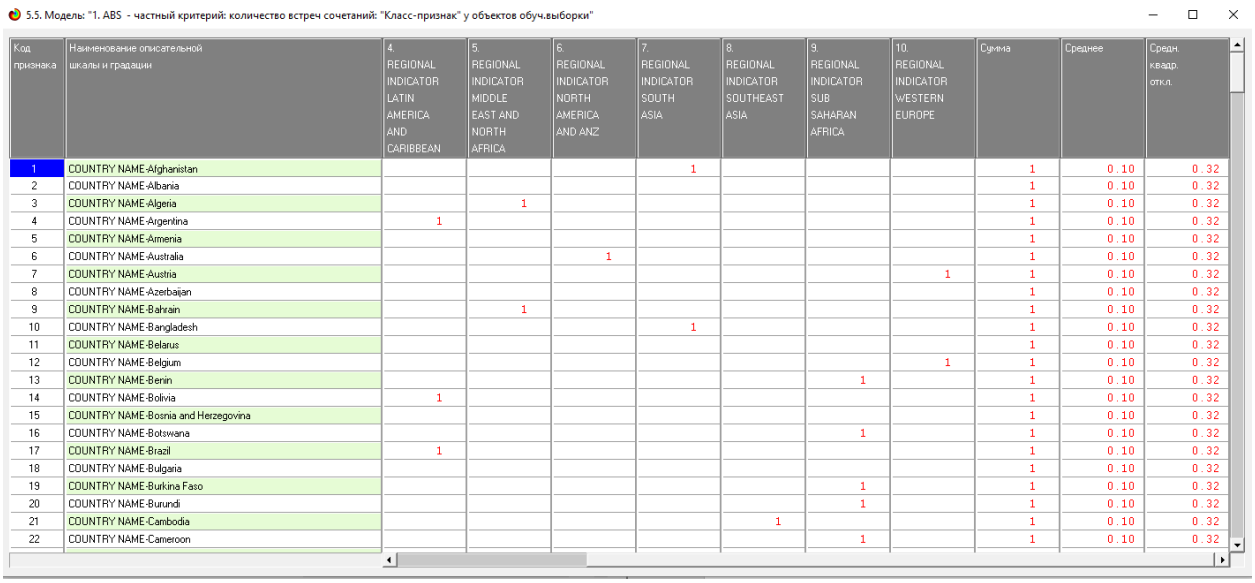
Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 23 секунды. При этом оценка достоверности моделей проводилась на 154 примерах наблюдения из обучающей выборки. После этого можно перейти непосредственно к выбору наиболее достоверной модели.

1.5. Виды моделей системы Aidos-X

Рассмотрим решение задачи идентификации на примере нескольких моделей, в которой рассчитано количество информации по А. Харкевичу,

которое было получено по принадлежности идентифицируемого объекта к каждому из классов, если известно, что у этого объекта есть некоторый признак.

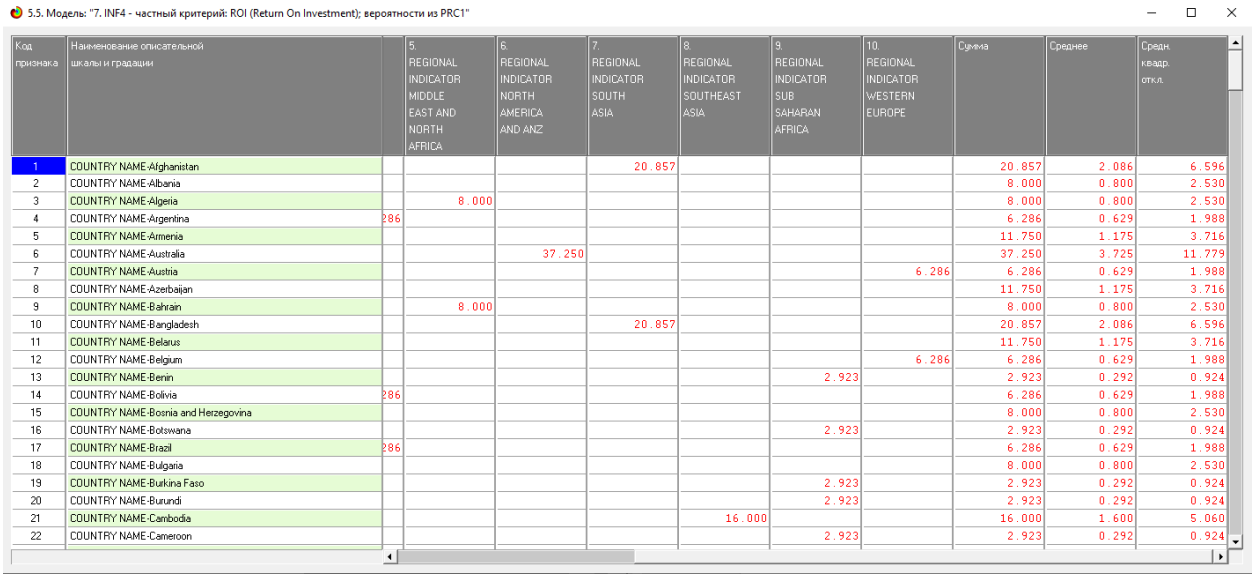
То есть частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот (рисунок 9) в матрицы условных и безусловных процентных распределений, и матрицы знаний (рисунок 10 и 11).



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

Код признака	Наименование описательной шкалы и градации	4. REGIONAL INDICATOR LATIN AMERICA AND CARIBBEAN	5. REGIONAL INDICATOR MIDDLE EAST AND NORTH AFRICA	6. REGIONAL INDICATOR NORTH AMERICA AND ANZ	7. REGIONAL INDICATOR SOUTH ASIA	8. REGIONAL INDICATOR SOUTHEAST ASIA	9. REGIONAL INDICATOR SUB SAHARAN AFRICA	10. REGIONAL INDICATOR WESTERN EUROPE	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY NAME-Afghanistan				1				1	0.10	0.32
2	COUNTRY NAME-Albania								1	0.10	0.32
3	COUNTRY NAME-Algeria		1						1	0.10	0.32
4	COUNTRY NAME-Argentina	1							1	0.10	0.32
5	COUNTRY NAME-Armenia								1	0.10	0.32
6	COUNTRY NAME-Australia			1					1	0.10	0.32
7	COUNTRY NAME-Austria							1	1	0.10	0.32
8	COUNTRY NAME-Azerbaijan								1	0.10	0.32
9	COUNTRY NAME-Bahrain		1						1	0.10	0.32
10	COUNTRY NAME-Bangladesh				1				1	0.10	0.32
11	COUNTRY NAME-Belarus								1	0.10	0.32
12	COUNTRY NAME-Belgium							1	1	0.10	0.32
13	COUNTRY NAME-Benin						1		1	0.10	0.32
14	COUNTRY NAME-Bolivia	1							1	0.10	0.32
15	COUNTRY NAME-Bosnia and Herzegovina								1	0.10	0.32
16	COUNTRY NAME-Botswana						1		1	0.10	0.32
17	COUNTRY NAME-Brazil	1							1	0.10	0.32
18	COUNTRY NAME-Bulgaria								1	0.10	0.32
19	COUNTRY NAME-Burkina Faso						1		1	0.10	0.32
20	COUNTRY NAME-Burundi						1		1	0.10	0.32
21	COUNTRY NAME-Cambodia					1			1	0.10	0.32
22	COUNTRY NAME-Cameroon						1		1	0.10	0.32

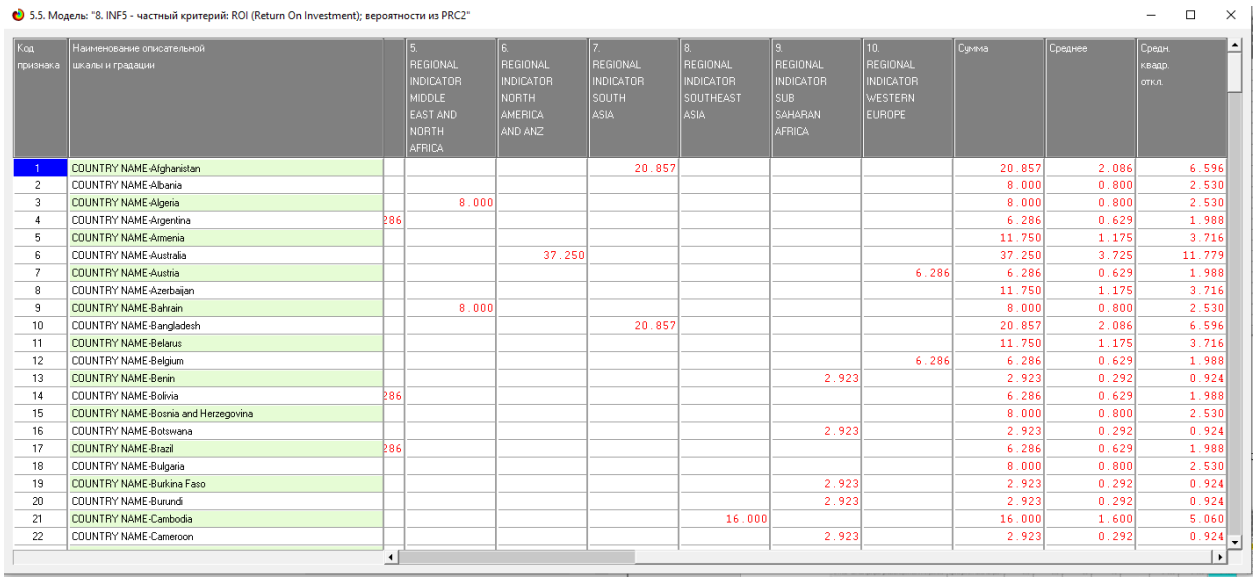
Рисунок 9 - Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений



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

Код признака	Наименование описательной шкалы и градации		5. REGIONAL INDICATOR MIDDLE EAST AND NORTH AFRICA	6. REGIONAL INDICATOR NORTH AMERICA AND ANZ	7. REGIONAL INDICATOR SOUTH ASIA	8. REGIONAL INDICATOR SOUTHEAST ASIA	9. REGIONAL INDICATOR SUB SAHARAN AFRICA	10. REGIONAL INDICATOR WESTERN EUROPE	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY NAME-Afghanistan				20.857				20.857	2.086	6.596
2	COUNTRY NAME-Albania								8.000	0.800	2.530
3	COUNTRY NAME-Algeria		8.000						8.000	0.800	2.530
4	COUNTRY NAME-Argentina	286							6.286	0.629	1.988
5	COUNTRY NAME-Armenia								11.750	1.175	3.716
6	COUNTRY NAME-Australia			37.250					37.250	3.725	11.779
7	COUNTRY NAME-Austria							6.286	6.286	0.629	1.988
8	COUNTRY NAME-Azerbaijan								11.750	1.175	3.716
9	COUNTRY NAME-Bahrain		8.000						8.000	0.800	2.530
10	COUNTRY NAME-Bangladesh				20.857				20.857	2.086	6.596
11	COUNTRY NAME-Belarus								11.750	1.175	3.716
12	COUNTRY NAME-Belgium							6.286	6.286	0.629	1.988
13	COUNTRY NAME-Benin						2.923		2.923	0.292	0.924
14	COUNTRY NAME-Bolivia	286							6.286	0.629	1.988
15	COUNTRY NAME-Bosnia and Herzegovina								8.000	0.800	2.530
16	COUNTRY NAME-Botswana						2.923		2.923	0.292	0.924
17	COUNTRY NAME-Brazil	286							6.286	0.629	1.988
18	COUNTRY NAME-Bulgaria								8.000	0.800	2.530
19	COUNTRY NAME-Burkina Faso						2.923		2.923	0.292	0.924
20	COUNTRY NAME-Burundi						2.923		2.923	0.292	0.924
21	COUNTRY NAME-Cambodia					16.000			16.000	1.600	5.060
22	COUNTRY NAME-Cameroon						2.923		2.923	0.292	0.924

Рисунок 10 - Модель INF4 (фрагмент)



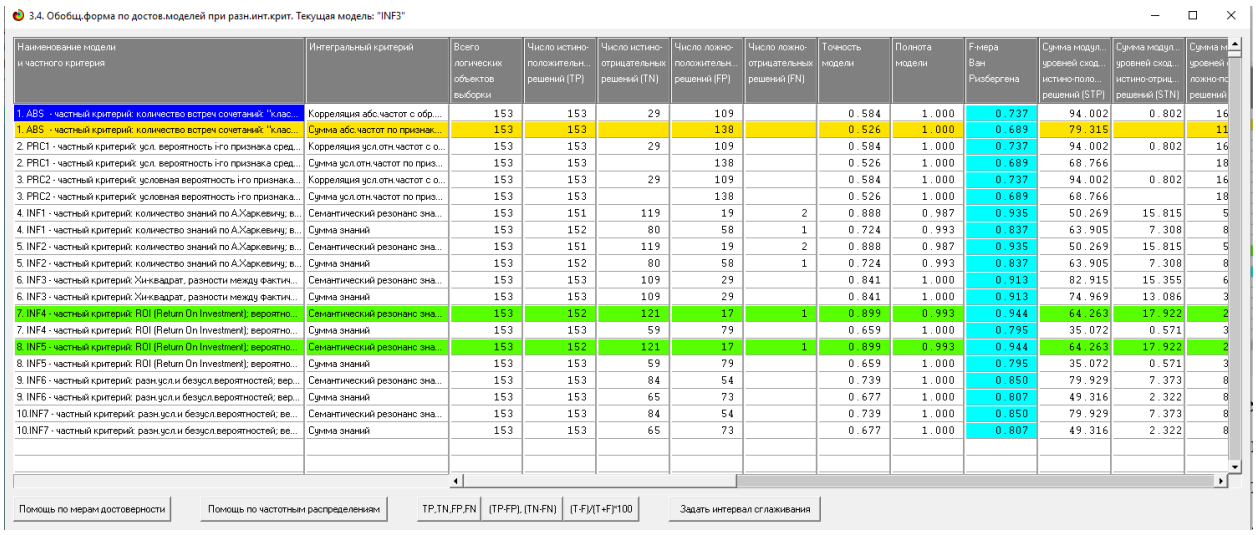
5.5. Модель: "8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2"

Код признака	Наименование описательной шкалы и градации		5. REGIONAL INDICATOR MIDDLE EAST AND NORTH AFRICA	6. REGIONAL INDICATOR NORTH AMERICA AND ANZ	7. REGIONAL INDICATOR SOUTH ASIA	8. REGIONAL INDICATOR SOUTHEAST ASIA	9. REGIONAL INDICATOR SUB SAHARAN AFRICA	10. REGIONAL INDICATOR WESTERN EUROPE	Сумма	Среднее	Средн. квадр. откл.
1	COUNTRY NAME-Afghanistan				20.857				20.857	2.086	6.596
2	COUNTRY NAME-Albania								8.000	0.800	2.530
3	COUNTRY NAME-Algeria		8.000						8.000	0.800	2.530
4	COUNTRY NAME-Argentina	286							6.286	0.629	1.988
5	COUNTRY NAME-Armenia								11.750	1.175	3.716
6	COUNTRY NAME-Australia			37.250					37.250	3.725	11.779
7	COUNTRY NAME-Austria							6.286	6.286	0.629	1.988
8	COUNTRY NAME-Azerbaijan								11.750	1.175	3.716
9	COUNTRY NAME-Bahrain		8.000						8.000	0.800	2.530
10	COUNTRY NAME-Bangladesh				20.857				20.857	2.086	6.596
11	COUNTRY NAME-Belarus								11.750	1.175	3.716
12	COUNTRY NAME-Belgium							6.286	6.286	0.629	1.988
13	COUNTRY NAME-Benin						2.923		2.923	0.292	0.924
14	COUNTRY NAME-Bolivia	286							6.286	0.629	1.988
15	COUNTRY NAME-Bosnia and Herzegovina								8.000	0.800	2.530
16	COUNTRY NAME-Botswana						2.923		2.923	0.292	0.924
17	COUNTRY NAME-Brazil	286							6.286	0.629	1.988
18	COUNTRY NAME-Bulgaria								8.000	0.800	2.530
19	COUNTRY NAME-Burkina Faso						2.923		2.923	0.292	0.924
20	COUNTRY NAME-Burundi						2.923		2.923	0.292	0.924
21	COUNTRY NAME-Cambodia					16.000			16.000	1.600	5.060
22	COUNTRY NAME-Cameroon						2.923		2.923	0.292	0.924

Рисунок 11 - Модель INF5 (фрагмент)

1.6. Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями представлены на рисунке 12.



3.4. Обобщ.форма по достов.моделей при разн.инт.крит. Текущая модель: "INF3"

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительн... решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительн... решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма модул... уровней сход... истинно-поло... решений (STP)	Сумма модул... уровней сход... истинно-отриц... решений (STN)	Сумма м... уровней ... ложно-по... решений
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс.частот с обр....	153	153	29	109		0.584	1.000	0.737	94.002	0.802	16
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс.частот по признак...	153	153		138		0.526	1.000	0.689	79.315		11
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	153	153	29	109		0.584	1.000	0.737	94.002	0.802	16
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	153	153		138		0.526	1.000	0.689	68.766		18
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	153	153	29	109		0.584	1.000	0.737	94.002	0.802	16
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	153	153		138		0.526	1.000	0.689	68.766		18
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	153	151	119	19	2	0.888	0.987	0.935	50.269	15.815	5
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	153	152	80	58	1	0.724	0.993	0.837	63.905	7.308	8
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	153	151	119	19	2	0.888	0.987	0.935	50.269	15.815	5
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	153	152	80	58	1	0.724	0.993	0.837	63.905	7.308	8
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	153	153	109	29		0.841	1.000	0.913	82.915	15.355	6
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	153	153	109	29		0.841	1.000	0.913	74.969	13.086	3
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	153	152	121	17	1	0.899	0.993	0.944	64.263	17.922	2
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	153	153	59	79		0.659	1.000	0.795	35.072	0.571	3
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	153	152	121	17	1	0.899	0.993	0.944	64.263	17.922	2
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	153	153	59	79		0.659	1.000	0.795	35.072	0.571	3
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	153	153	84	54		0.739	1.000	0.850	79.929	7.373	8
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	153	153	65	73		0.677	1.000	0.807	49.316	2.322	8
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	153	153	84	54		0.739	1.000	0.850	79.929	7.373	8
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	153	153	65	73		0.677	1.000	0.807	49.316	2.322	8

Помощь по мерам достоверности | Помощь по частотным распределениям | TP,TN,FP,FN | (TP-FP), (TN-FN) | (T-F)/(T+F)*100 | Задать интервал сглаживания

Рисунок 12 - Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF4 и INF 5 при интегральном критерии «Семантический резонанс знаний», но в дальнейшем будем описывать только INF 4. При этом точность модели (F-мера Ван Ризбергена) составляет 0,944 а точность модели (L1-мера профессора Луценко) - 0,982. L1-мера профессора Луценко является более

достоверной, по сравнению с F-мерой Ван Ризбергена. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 100%. Для оценки достоверности моделей в АСК-анализе и системе Aidos-X используется F-мера Ван Ризбергена и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Е.В.Луценко (рисунок 13).

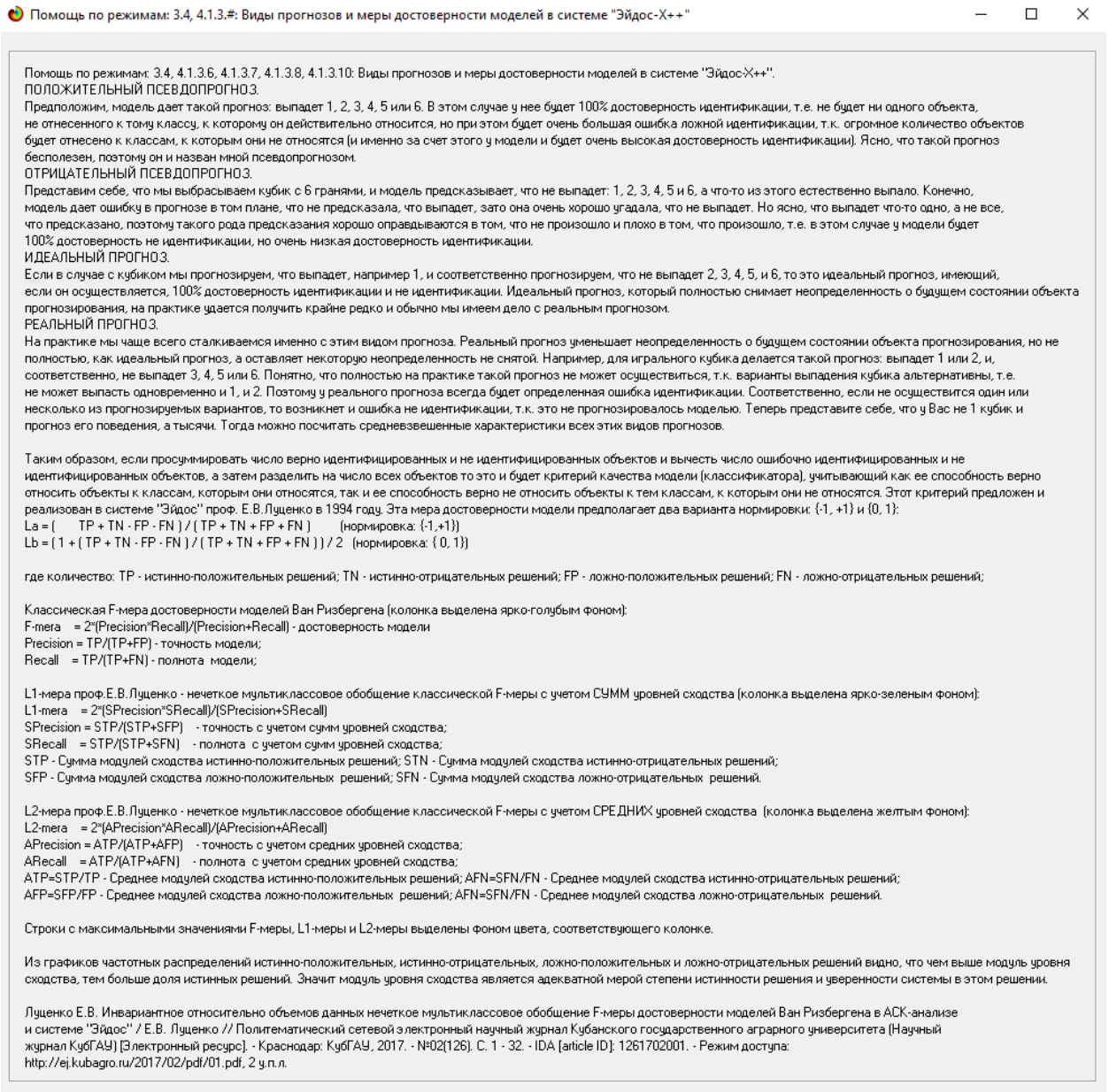


Рисунок 13- Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

Также необходимо обратить внимание на то, 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 21 интеллектуальных технологий. На рисунке 14 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и не идентифицированных ситуаций в наиболее достоверной модели INF 4.

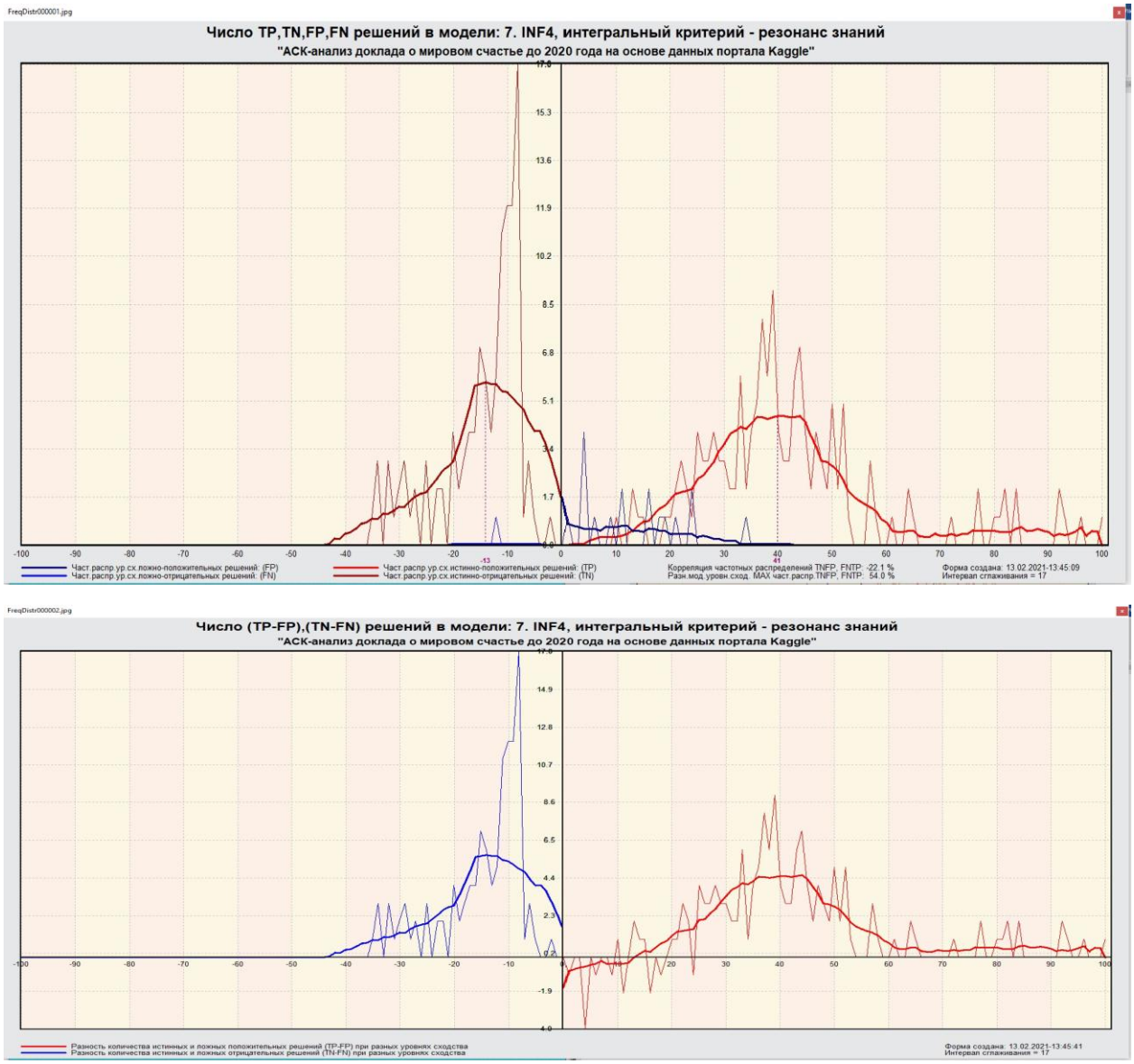


Рисунок 14 - Частотные распределения числа верно и ошибочно идентифицированных и не идентифицированных состояний объекта моделирования в зависимости от уровня сходства в модели INF4

Из рисунка 14 видно, что:

– наиболее модель INF4 лучше определяет непринадлежность объекта к классу, чем принадлежность;

– модуль уровня сходства-различия в модели INF4 для верно идентифицированных и верно не идентифицированных объектов значительно выше, чем для ошибочно идентифицированных и ошибочно не идентифицированных. Это верно практически для всего диапазона уровней сходства-различия, кроме небольших по модулю значений в диапазоне от 0 до 20% уровня сходства. Для уровней сходства-различия более 25% ошибочно идентифицированные и не идентифицированными ситуации практически отсутствует.

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.

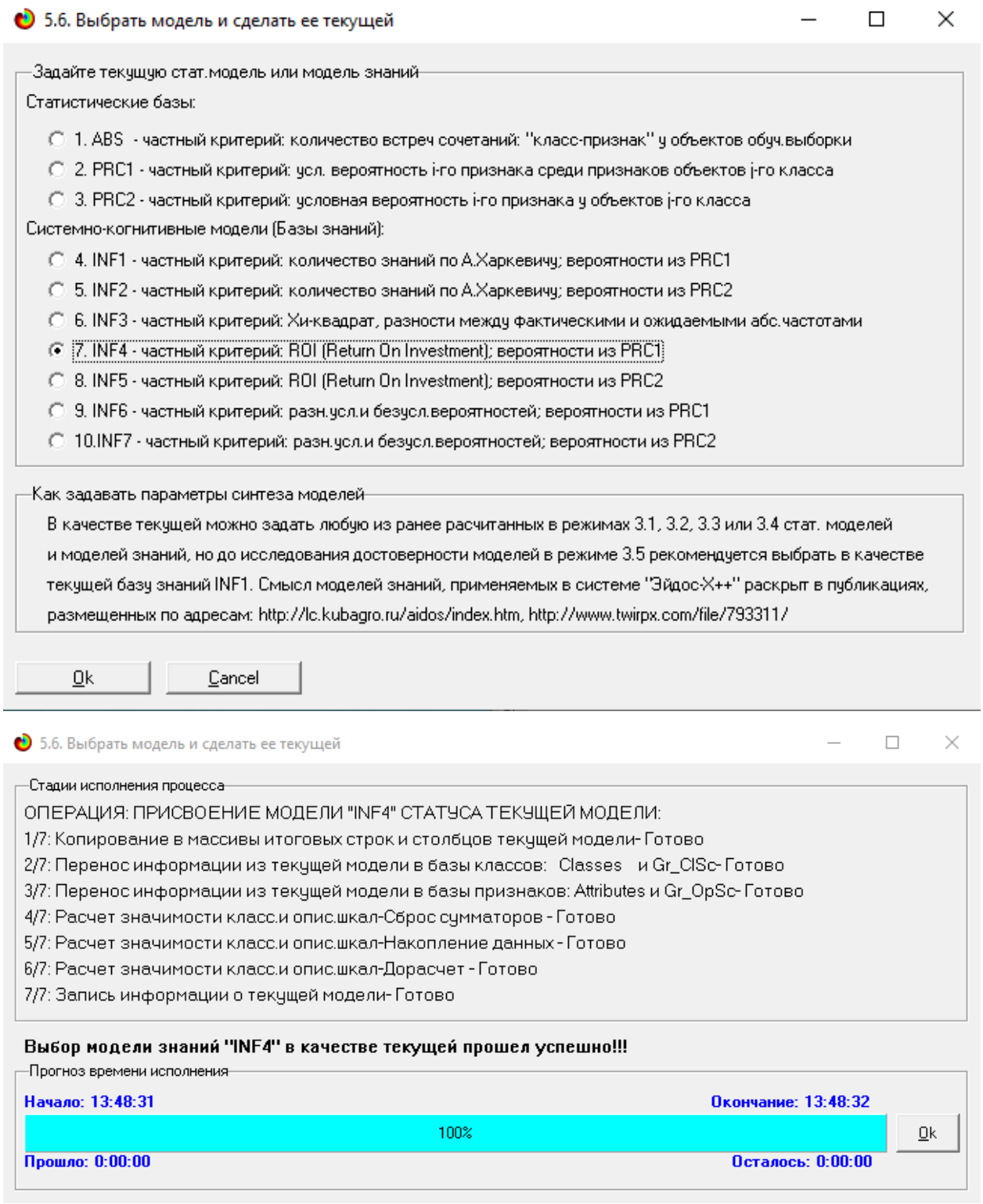


Рисунок 15 - Экранные формы режима задания модели в качестве текущей

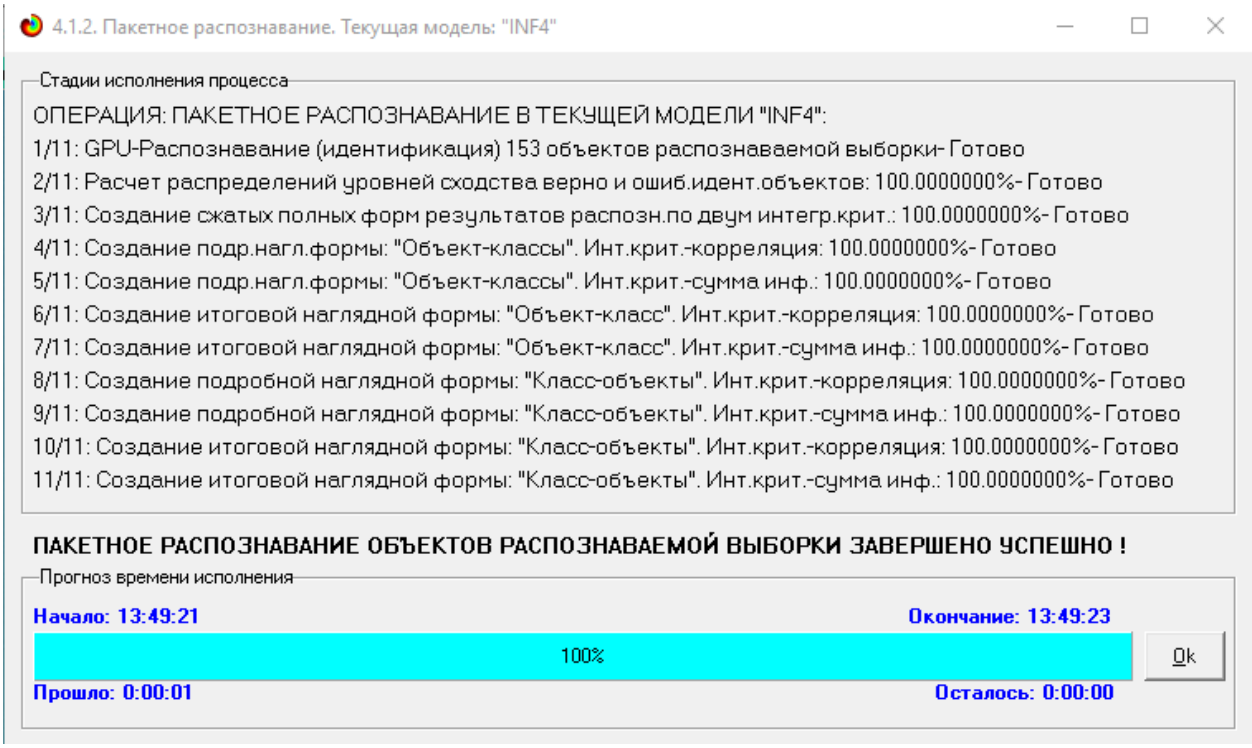


Рисунок 16 - Экранная форма режима пакетного распознавания в текущей модели

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 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 17 и 18 приведены примеры прогнозов высокой и низкой достоверности частоты объектов и классов в модели ABS на основе наблюдения предыстории их развития:

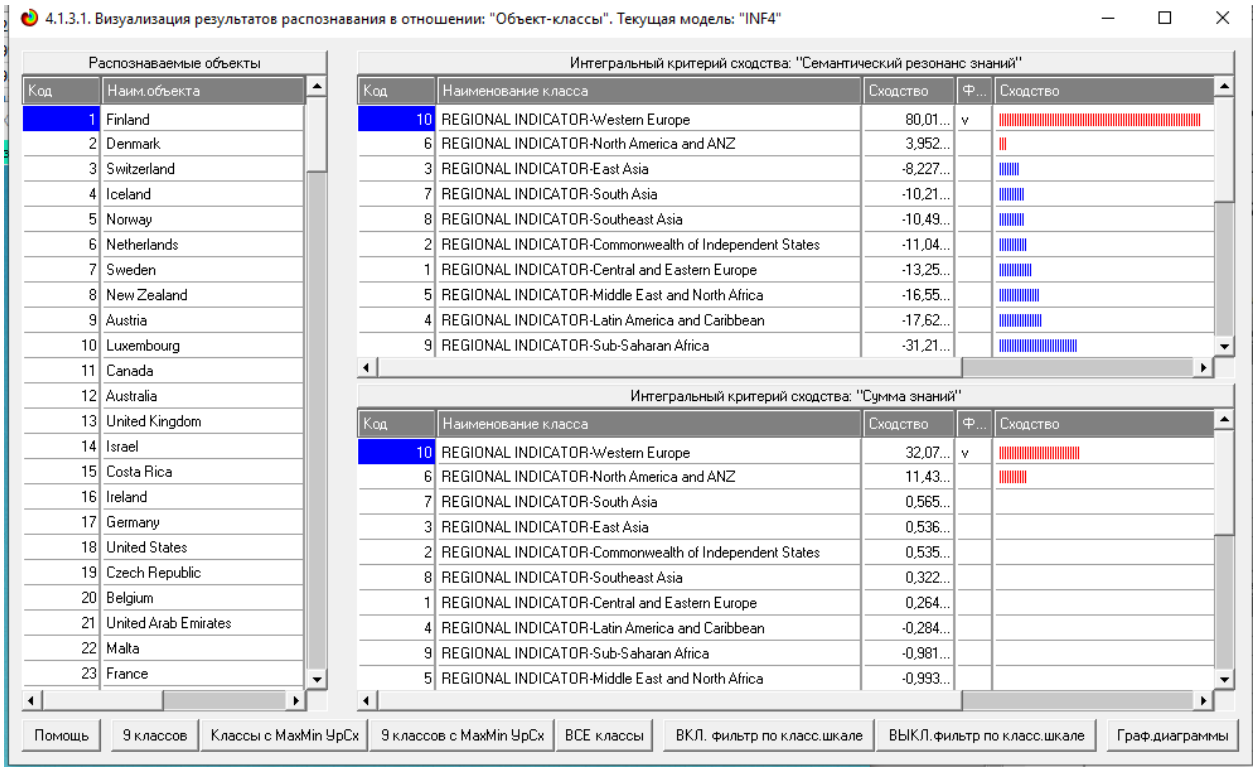


Рисунок 17. Пример идентификации объектов в модели ABS

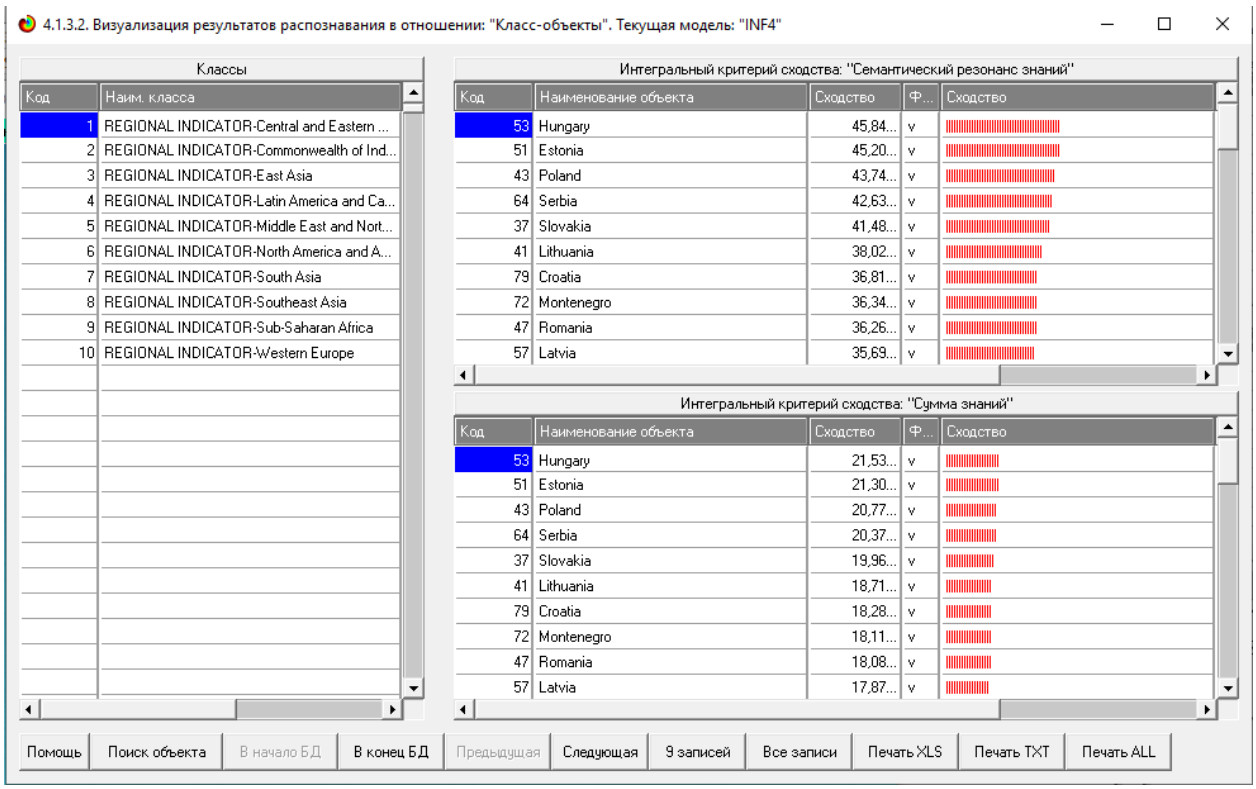


Рисунок 18. Пример идентификации классов в модели ABS

2.2. Кластерно-конструктивный анализ

Сходство-различие обобщенных образов различных результатов научной деятельности по характерным для них системам значений показателей. Результаты сравнения классов по системе характерных приведены на рисунке 19:

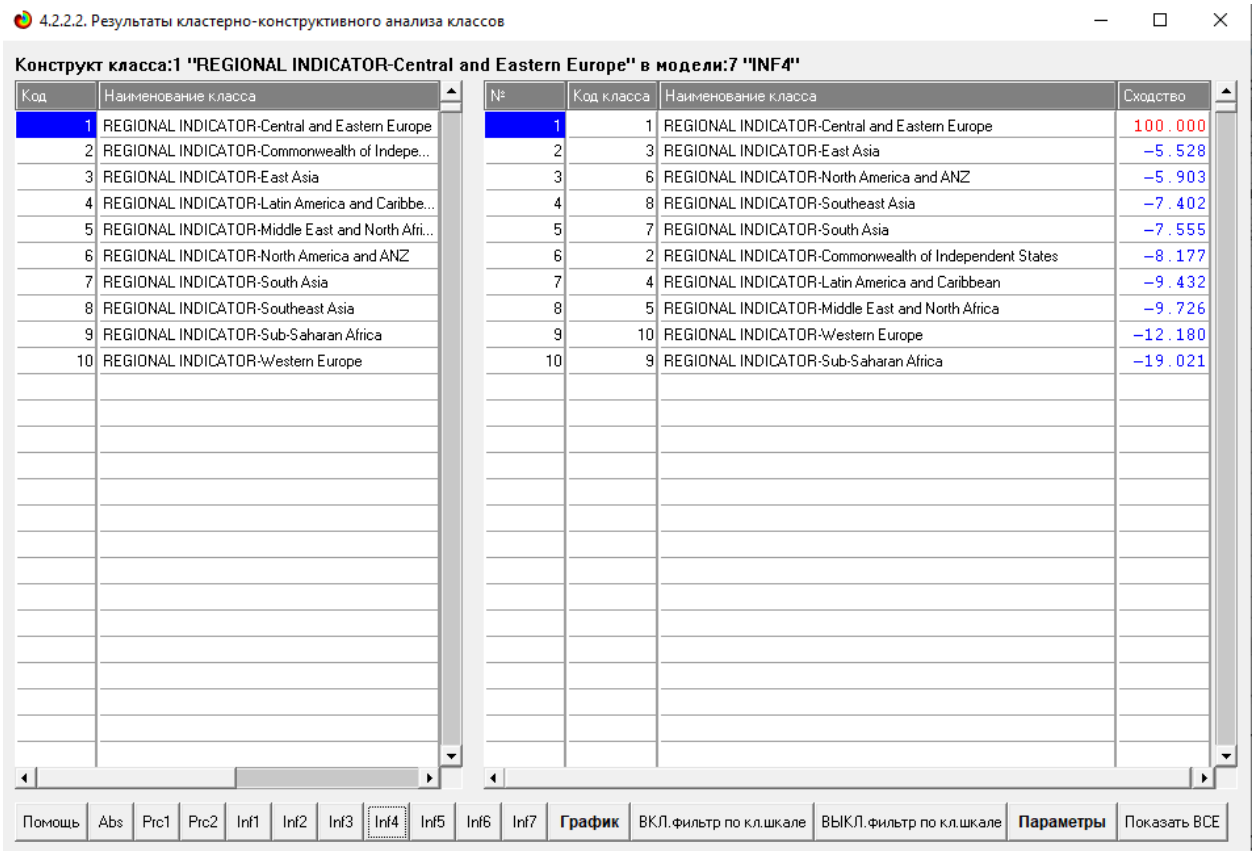


Рисунок 19 - Результаты кластерно-конструктивного анализа классов

На рисунке 20 представлена семантическая сеть классов в модели «INF4».

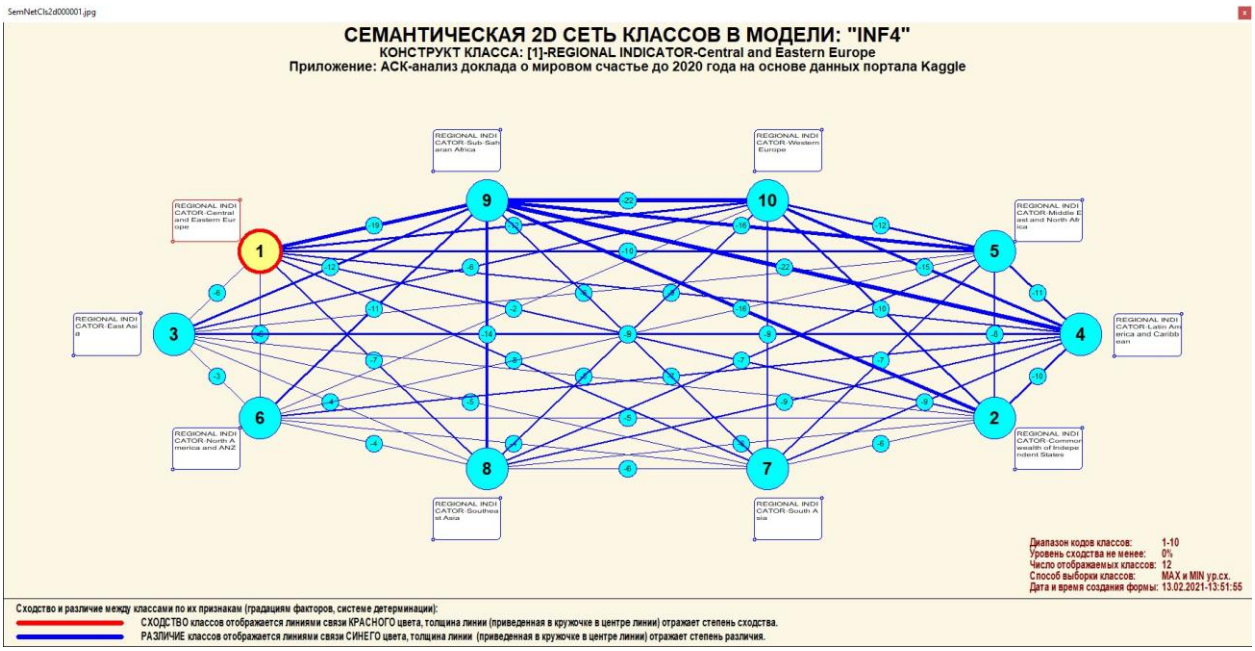


Рисунок 20 - Семантическая 2D сеть классов

Рисунок 21 – Агломеративная дендрограмма классов

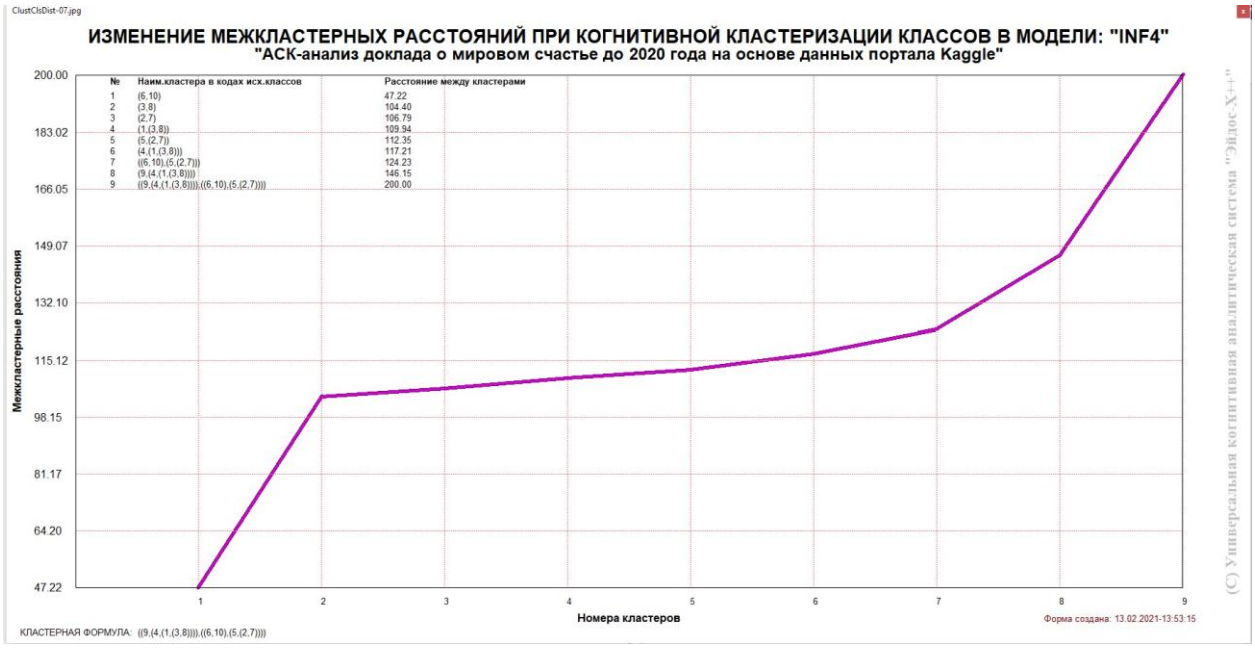


Рисунок 22 – График межкластерных расстояний классов

2.3. Нелокальные нейронные сети и нейтроны

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

Рассмотрим пару примеров, возвращаясь к нашим задачам.

На рисунке 23 изображено графическое отображение нелокальных нейронов в системе Aidos-X

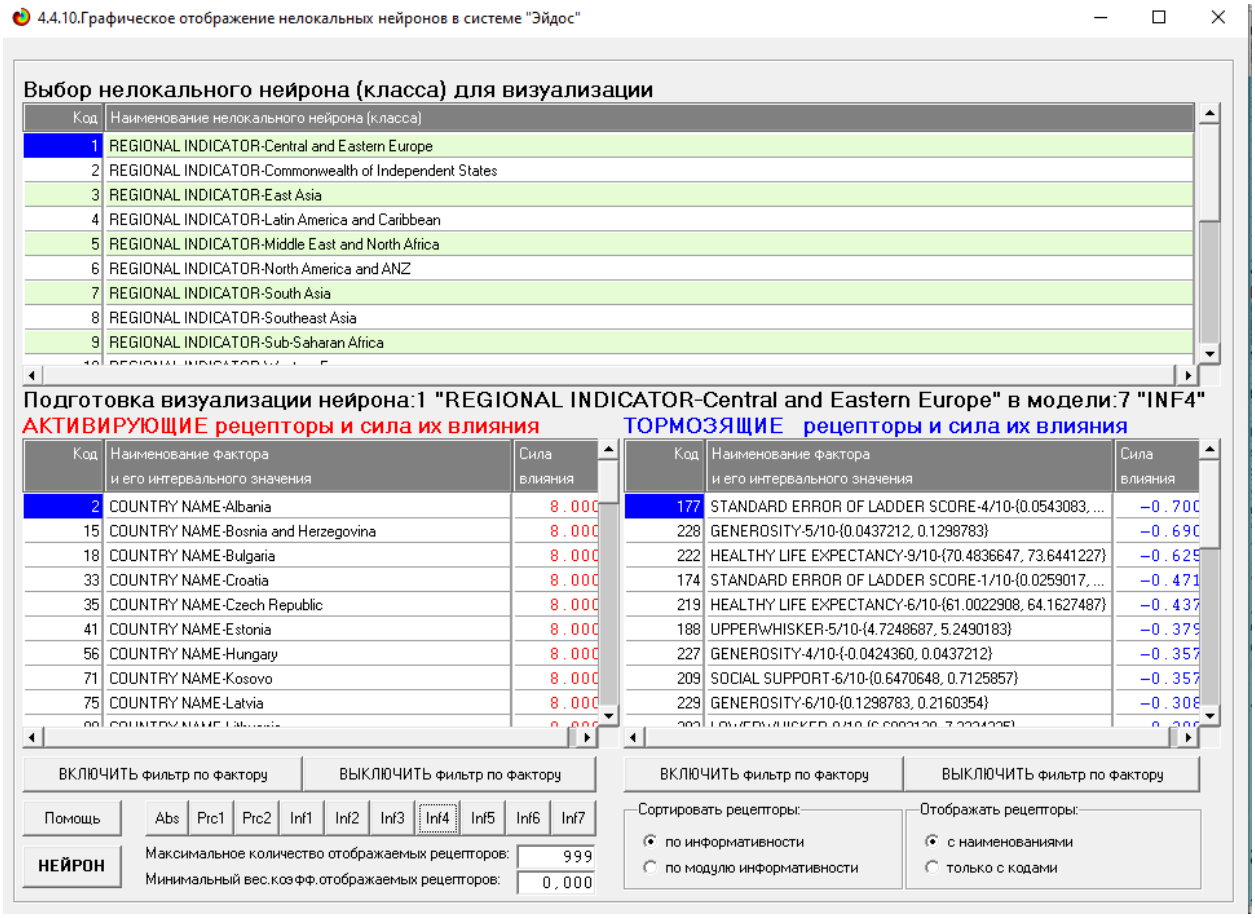


Рисунок 24 - Графическое отображение нелокальных нейронов в системе Aidos-X

Для каждого технологического фактора в соответствии с предложенной моделью определяется величина и направление его влиянии на осуществление всех желаемых и не желаемых хозяйственных ситуаций. Для каждой ситуации эта информация отображается в различных текстовых и графических формах, в частности в форме нелокального нейрона (рисунок

25). На данной диаграмме цвет линии означает знак связи (красный – положительная, синий – отрицательная), а толщина –ее модуль.

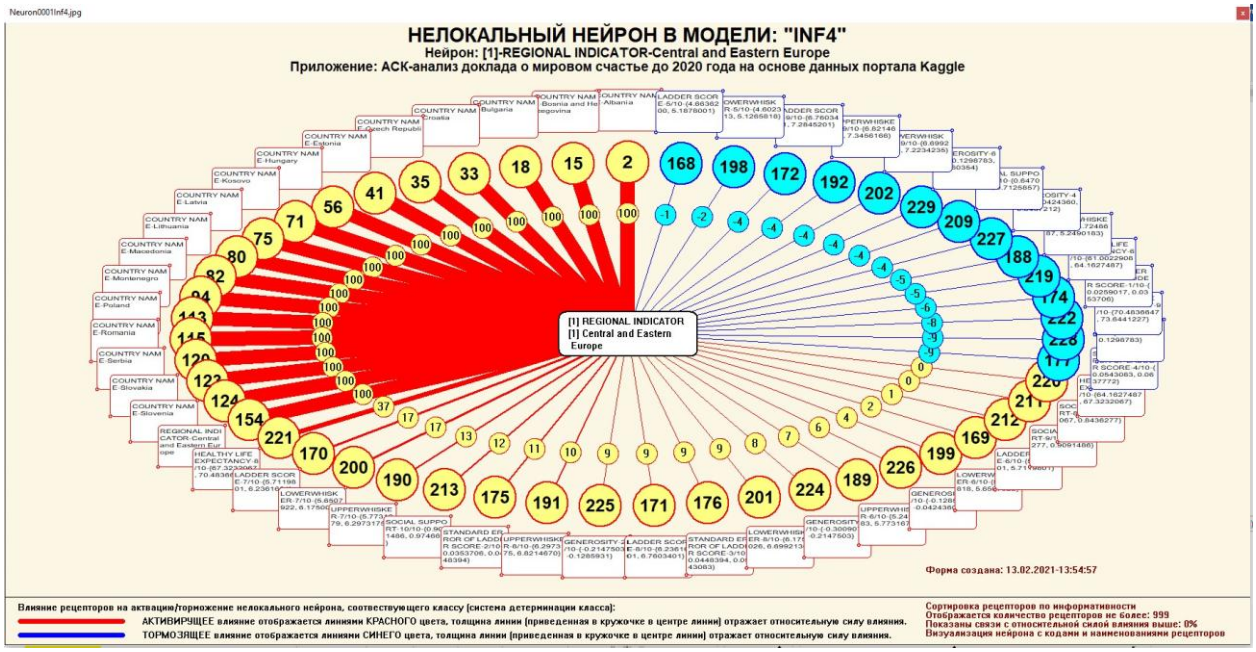


Рисунок 25 - Нелокальный нейрон в модели “INF3”

Дополнение модели нейрона связями факторов позволяет построить классическую когнитивную карту ситуации (будущего состояния АОУ). Детальная внутренняя структура любой связи отображается в форме инвертированной когнитивной диаграммы (рисунок 26). Необходимо отметить, что все указанные графические формы генерируются системой Aidos-X автоматически в соответствии с созданной моделью[9].

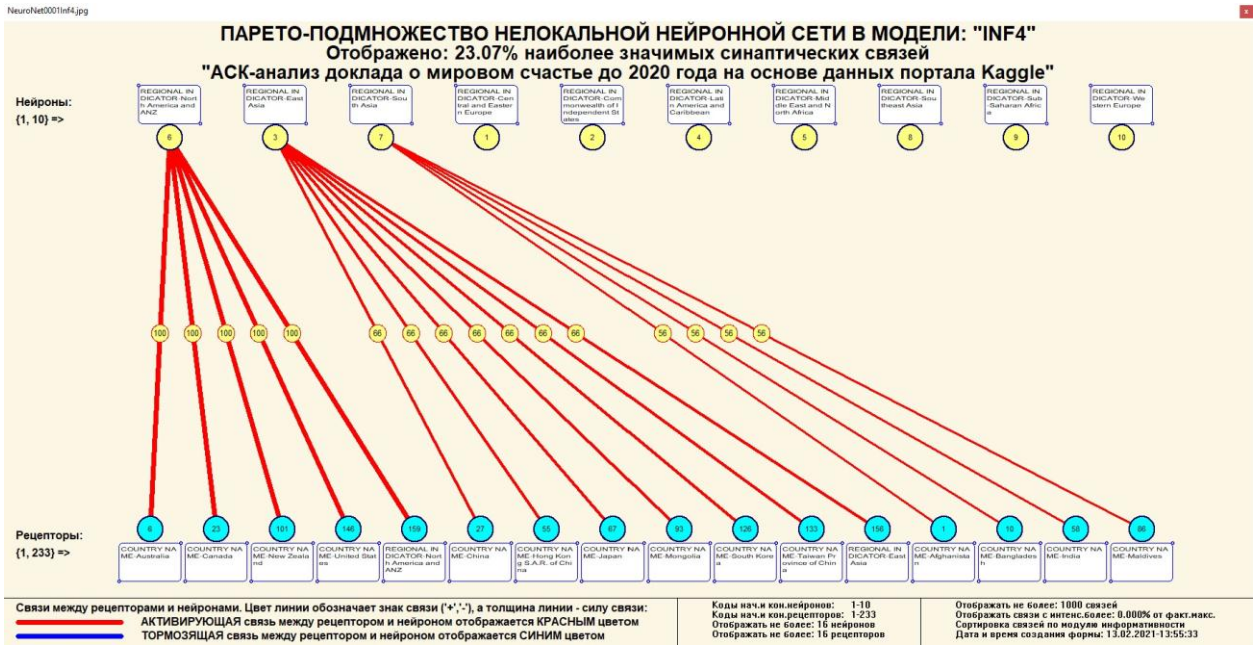


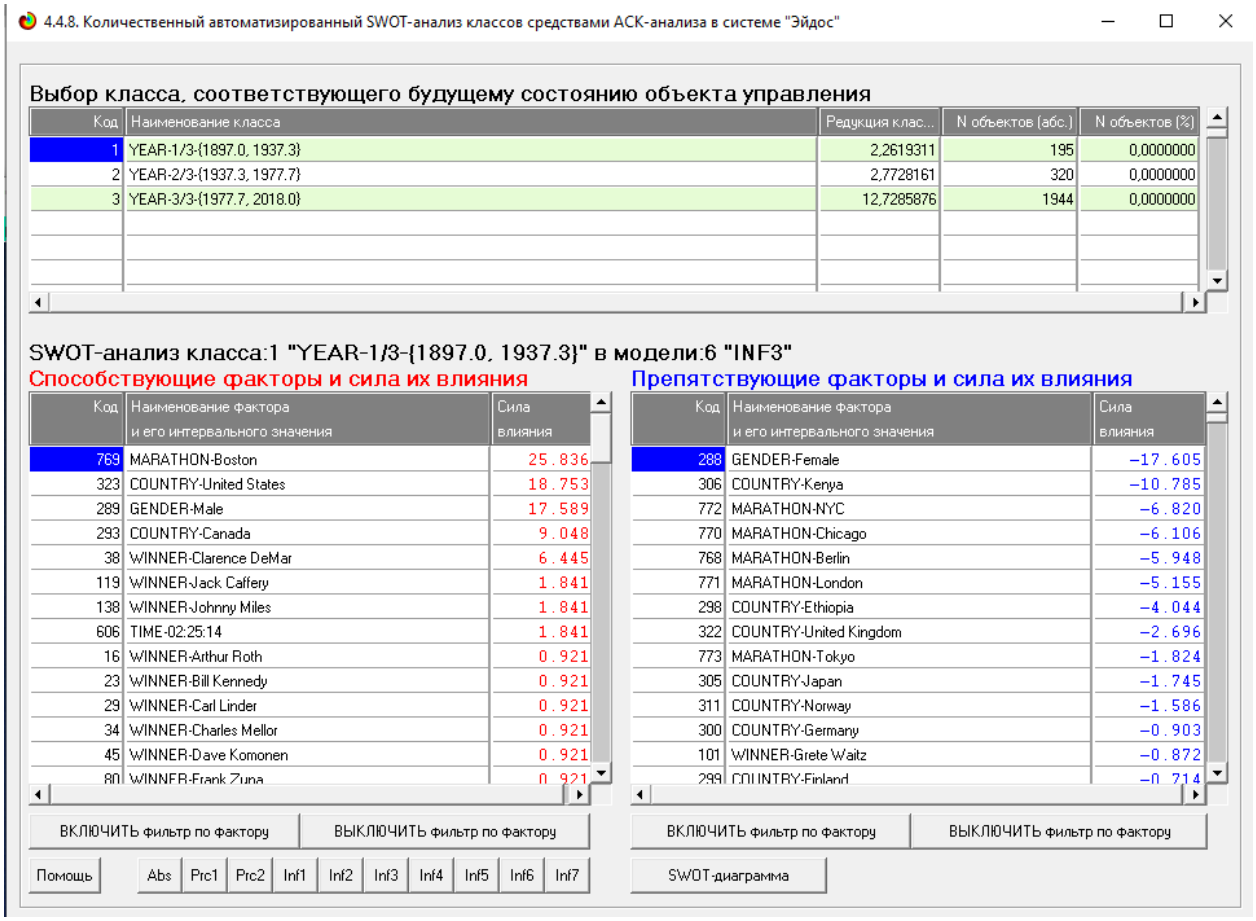
Рисунок 26 - Паретто – подмножество нелокальной нейронной сети

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 не формализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных. Подобная технология разработана давно, ей уже около 30 лет, но она малоизвестна – это интеллектуальная система Aidos-X. Данная система

всегда обеспечивала возможность проведения количественного автоматизированного SWOT-анализа без использования экспертных оценок непосредственно на основе эмпирических данных. Результаты SWOTанализа выводились в форме информационных портретов. 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 27).

На рисунке 28 приведен пример табличной выходной формы количественного автоматизированного SWOT- и PEST- анализа средствами системы Aidos.



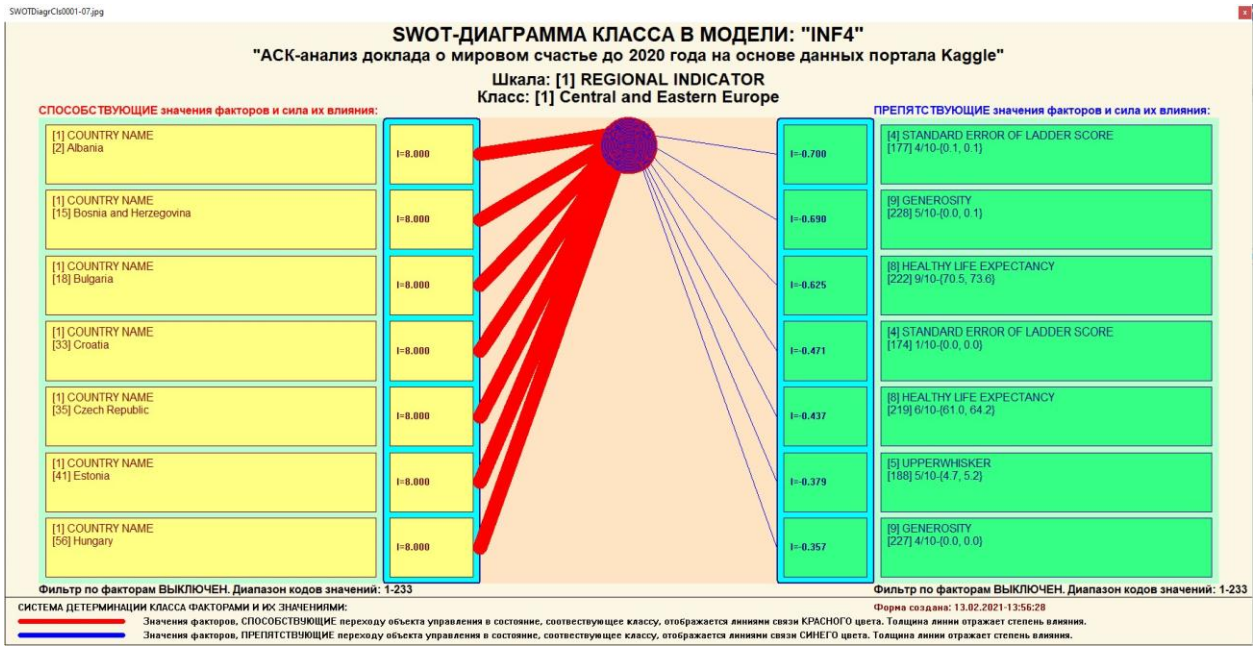


Рисунок 27 - Пример SWOT-матрицы в модели INF4

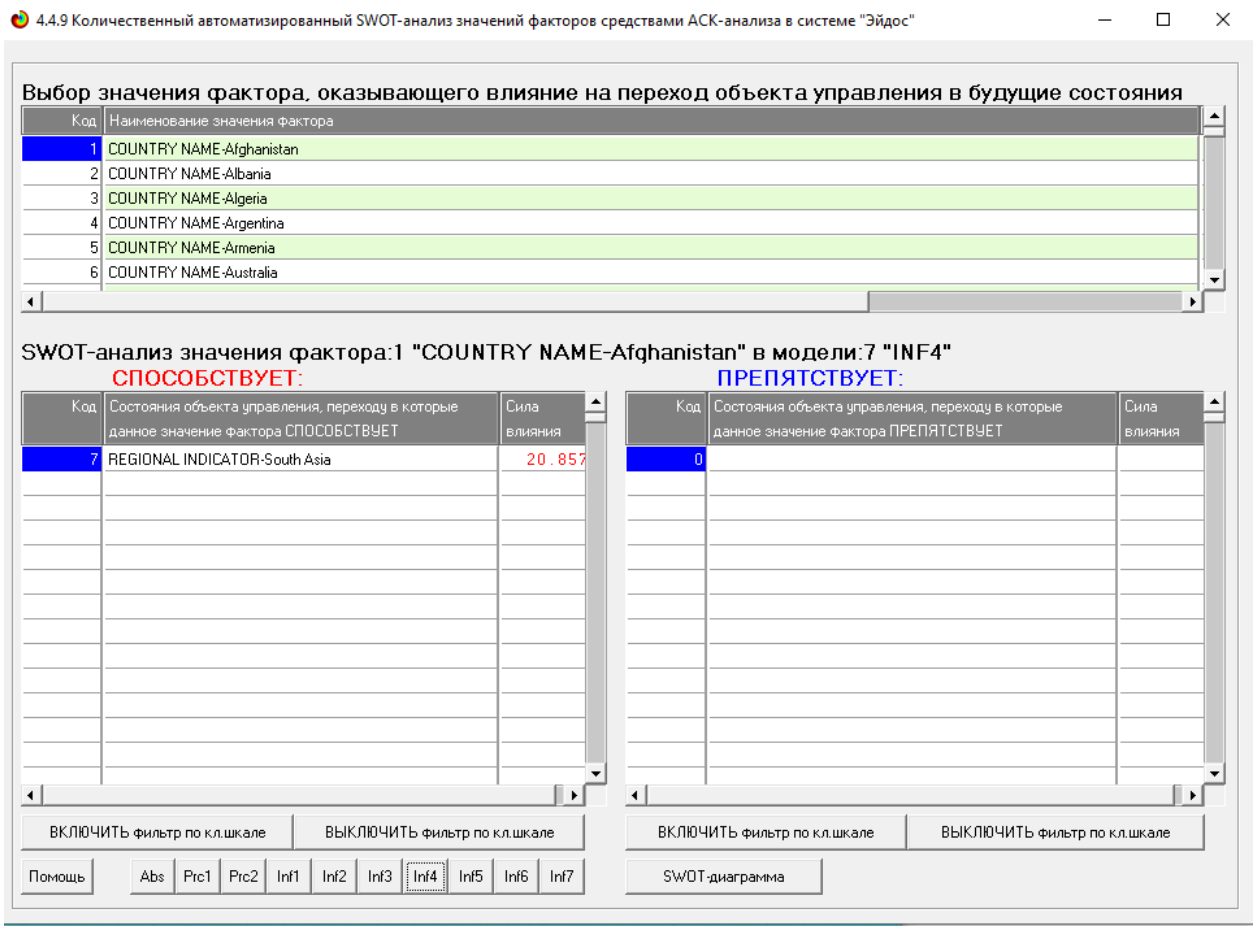


Рисунок 28 - Табличная выходная форма количественного автоматизированного SWOT- и PEST- анализа средствами системы Aidos

На рисунке 29 приведен пример графической выходной формы количественного автоматизированного SWOT- и PEST- анализа средствами системы Aidos.

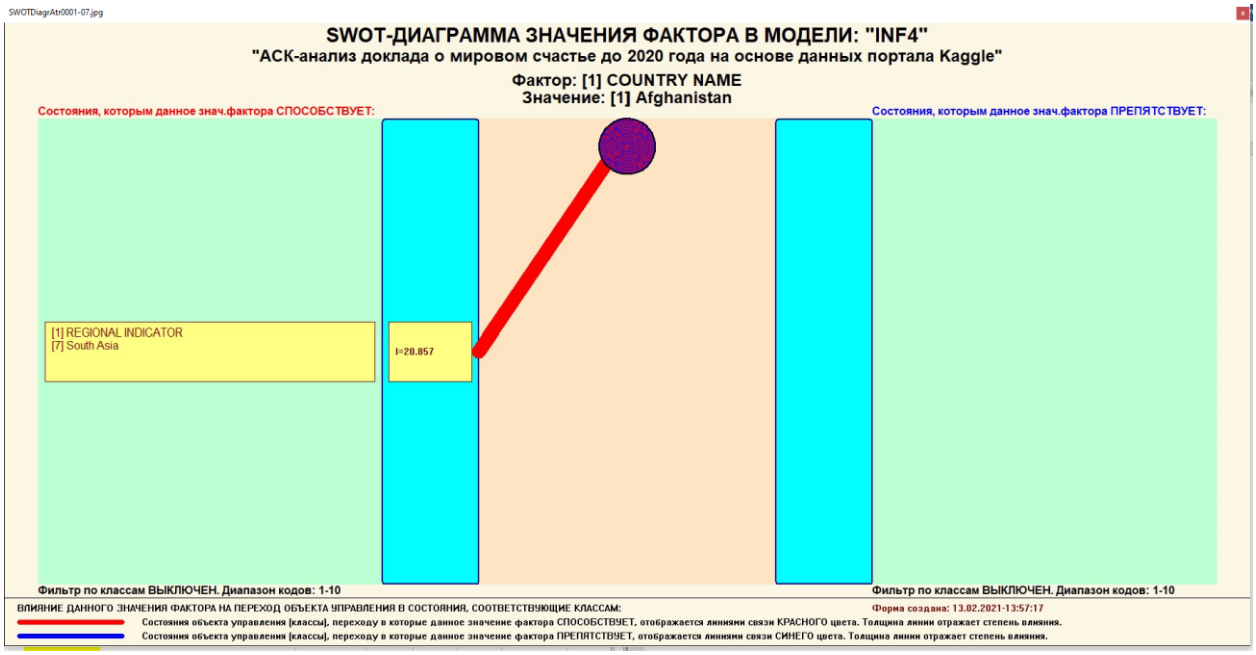


Рисунок 29 - Графическая выходная форма количественного автоматизированного SWOT- и PEST-анализа средствами системы Aidos

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 электронную базу Kaggle.

Сверхзадачей искусственного интеллекта является построение компьютерной интеллектуальной системы, которая обладала бы уровнем эффективности решений неформализованных задач, сравнимым с человеческим или превосходящим его.

Самую существенную часть систем искусственного интеллекта составляют экспертные системы. Экспертная система обычно определяется

как программа ЭВМ, моделирующая действия эксперта человека при решении задач в узкой предметной области: составление базы знаний и накопления их.

В данной курсовой работе был показано построение модели зависимости регионов стран от показателей мирового счастья системой искусственного интеллекта "Aidos-X++". При этом наиболее достоверной в данном приложении оказалась модель INF4, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Семантический резонанс знаний». Точность модели составляет 0,944, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 49,4%.

АСК-анализ текстов позволяет:

- формировать обобщенные лингвистические образы классов (семантические ядра) на основе фрагментов или примеров относящихся к ним текстов на любом языке;

- количественно сравнивать лингвистический образ конкретного человека, или описание объекта, процесса с обобщенными лингвистическими образами групп (классов);

- сравнивать обобщенные лингвистические образы классов друг с другом и создавать их кластеры и конструкты;

- исследовать моделируемую предметную область путем исследования ее лингвистической системно-когнитивной модели;

- проводить интеллектуальную атрибуцию текстов, т.е. определять вероятное авторство анонимных и псевдонимных текстов, датировку, жанр и смысловую направленность содержания текстов;

- все это можно делать для любого естественного или искусственного языка или системы кодирования.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: http://ej.kubagro.ru/2016/02/pdf/01.pdf, 3,75 у.п.л.3.
2. Луценко Е.В. АСК-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: Куб- ГАУ, 2014. – №06(100). С. 1346 – 1395. – IDA [article ID]: 1001406090. – Режим доступа: http://ej.kubagro.ru/2014/06/pdf/90.pdf, 3,125 у.п.л.
3. Луценко Е.В. 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. [Электронный ресурс]. Статья"Emergency – 911 Calls": https://www.kaggle.com/mchirico/montcoalert, свободный. - Загл. сэкрана. Яз.анг.
5. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: http://lc.kubagro.ru/, свободный. - Загл. с экрана. Яз.рус.
6. Луценко Е.В. Количественная оценка степени манипулирования индексом Хирша и его модификация, устойчивая к манипулированию / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №07(121). С. 202 – 234. – IDA [article ID]: 1211607005. – Режим доступа: http://ej.kubagro.ru/2016/07/pdf/05.pdf, 2,062 у.п.л.. http://dx.doi.org/10.21515/1990-4665-121-0057.
7. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- X++» / 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.
8. 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.
9. Луценко Е.В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы «Эйдос» (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е.В. Луценко, Н.В. Андрафанова, Н.В. Потапова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2019. – №01(145). С. 31 – 102. – IDA [article ID]: 1451901033. – Режим доступа: http://ej.kubagro.ru/2019/01/pdf/33.pdf, 4,5 у.п.л.

10. Луценко Е.В. Формирование семантического ядра ветеринарии путем Автоматизированного системно-когнитивного анализа паспортов научных специальностей ВАК РФ и автоматическая классификация текстов по направлениям науки / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2018. – №10(144). С. 44 – 102. – IDA [article ID]: 1441810033. – Режим доступа: http://ej.kubagro.ru/2018/10/pdf/33.pdf, 3,688 у.п.л.

11. Луценко Е.В. Интеллектуальная привязка некорректных ссылок к литературным источникам в библиографических базах данных с применением АСК-анализа и системы «Эйдос» (на примере Российского индекса научного цитирования – РИНЦ) / Е.В. Луценко, В.А. Глухов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №01(125). С. 1 – 65. – IDA [article ID]: 1251701001. – Режим доступа: http://ej.kubagro.ru/2017/01/pdf/01.pdf, 4,062 у.п.л.

12. Луценко Е.В. Применение АСК-анализа и интеллектуальной системы "Эйдос" для решения в общем виде задачи идентификации литературных источников и авторов по стандартным, нестандартным и некорректным библиографическим описаниям / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №09(103). С. 498 – 544. – IDA [article ID]: 1031409032. – Режим доступа: http://ej.kubagro.ru/2014/09/pdf/32.pdf, 2,938 у.п.л.

13. Луценко Е.В. АСК-анализ проблематики статей Научного журнала КубГАУ в динамике / Е.В. Луценко, В.И. Лойко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №06(100). С. 109 – 145. – IDA [article ID]: 1001406007. – Режим доступа: http://ej.kubagro.ru/2014/06/pdf/07.pdf, 2,312 у.п.л.

14. Луценко Е.В. Атрибуция анонимных и псевдонимных текстов в системно-когнитивном анализе / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №03(005). С. 44 – 64. – IDA [article ID]: 0050403003. – Режим доступа: http://ej.kubagro.ru/2004/03/pdf/03.pdf, 1,312 у.п.л.

15. Луценко Е.В. Атрибуция текстов, как обобщенная задача идентификации и прогнозирования / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №02(002). С. 146 – 164. – IDA [article ID]: 0020302013. – Режим доступа: http://ej.kubagro.ru/2003/02/pdf/13.pdf, 1,188 у.п.л.

16. Луценко Д.С., Луценко Е.В. Интеллектуальная датировка текста, определение авторства и жанра на примере русской литературы XIX и XX веков, 2020 // Статья в открытом архиве. 38 с. – DOI: 10.13140/RG.2.2.28824.01281, https://www.elibrary.ru/item.asp?id=43796415

17. Lutsenko D.S., Lutsenko E.V. Intellectual attribution of literary texts (finding the dates of the text, determining authorship and genre on the example of russian literature of the XIX and XX centuries), 2020 // Статья в открытом архиве. 9 p. – DOI: 10.13140/RG.2.2.15349.81122, https://www.elibrary.ru/item.asp?id=43794562