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 случаев заболевания лихорадкой денге на Филиппинах на основе данных портала Kaggle
Выполнил студент группы: ИТ1941 Рубцов Александр Сергеевич

Допущен к защите: _____

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

Защищен_____12.02.2021_____
(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 Рубцова Александра Сергеевича
курса 2 очной формы обучения группы ИТ1941
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ случаев заболевания лихорадкой денге на Филиппинах 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: 36 страницы, 29 рисунков, 17 литературных источников.

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

Целью работы АСК-анализ случаев заболевания лихорадкой денге на Филиппинах 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 10

1.4. Синтез и верификация статистических и интеллектуальных моделей 14

1.5. Виды моделей системы Aidos-X 16

1.6. Результаты верификации моделей 18

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ 21

2.1. Решение задачи идентификации 21

2.2. Кластерно-конструктивный анализ 25

2.3. Нелокальные нейронные сети и нейтроны 27

2.4. SWOT и PEST матрицы и диаграммы 30

ЗАКЛЮЧЕНИЕ 33

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 случаев заболевания лихорадкой денге на Филиппинах 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 «Dengue Cases in the Philippines» 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 «Dengue Cases in the Philippines»

https://www.kaggle.com/grosvenpaul/dengue-cases-in-the-philippines

Csv файл содержит следующие данные:

Month – Месяц.

Year Год.

Region Регион.

Dengue_Cases Коэффициент эффективности болезни.

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 наименования игр. Также данный столбец был переименован в «Month» и размещен в таблице между столбцами «Year» и «Region» (Рисунок 1).

	1	2	3	4	5
1	Month	Year	Month	Region	Dengue_C
2	Jan	2008	Jan	Region.I	2,953926
3	Feb	2008	Feb	Region.I	2,183336
4	Mar	2008	Mar	Region.I	0,97241
5	Apr	2008	Apr	Region.I	9,357156
6	May	2008	May	Region.I	7,320599
7	Jun	2008	Jun	Region.I	4,513452
8	Jul	2008	Jul	Region.I	6,82522
9	Aug	2008	Aug	Region.I	9,210377
10	Sep	2008	Sep	Region.I	13,30184
11	Oct	2008	Oct	Region.I	1,724652
12	Nov	2008	Nov	Region.I	1,32101
13	Dec	2008	Dec	Region.I	1,0458
14	Jan	2009	Jan	Region.I	5,23598
15	Feb	2009	Feb	Region.I	3,719351
16	Mar	2009	Mar	Region.I	2,473549
17	Apr	2009	Apr	Region.I	3,051313
18	May	2009	May	Region.I	4,658217
19	Jun	2009	Jun	Region.I	6,644278
20	Jul	2009	Jul	Region.I	1,263857
21	Aug	2009	Aug	Region.I	1,516629
22	Sep	2009	Sep	Region.I	0,902755
23	Oct	2009	Oct	Region.I	4,333225
24	Nov	2009	Nov	Region.I	3,159643
25	Dec	2009	Dec	Region.I	2,166613
26	Jan	2010	Jan	Region.I	1,97192
27	Feb	2010	Feb	Region.I	3,197708
28	Mar	2010	Mar	Region.I	5,276218
29	Apr	2010	Apr	Region.I	0,657307
30	May	2010	May	Region.I	0,728367
31	Jun	2010	Jun	Region.I	0,604011
32	Jul	2010	Jul	Region.I	4,796562
33	Aug	2010	Aug	Region.I	3,286533
34	Sep	2010	Sep	Region.I	2,700287
35	Oct	2010	Oct	Region.I	1,954155
36	Nov	2010	Nov	Region.I	3,712894
37	Dec	2010	Dec	Region.I	4,672207
38	Jan	2011	Jan	Region.I	1,817482

Рисунок 1 - Фрагмент обучающей выборки

Таким образом, в качестве классификационной шкалы решено использовать столбец 2. Описательные шкалы столбцы 3-5.

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.

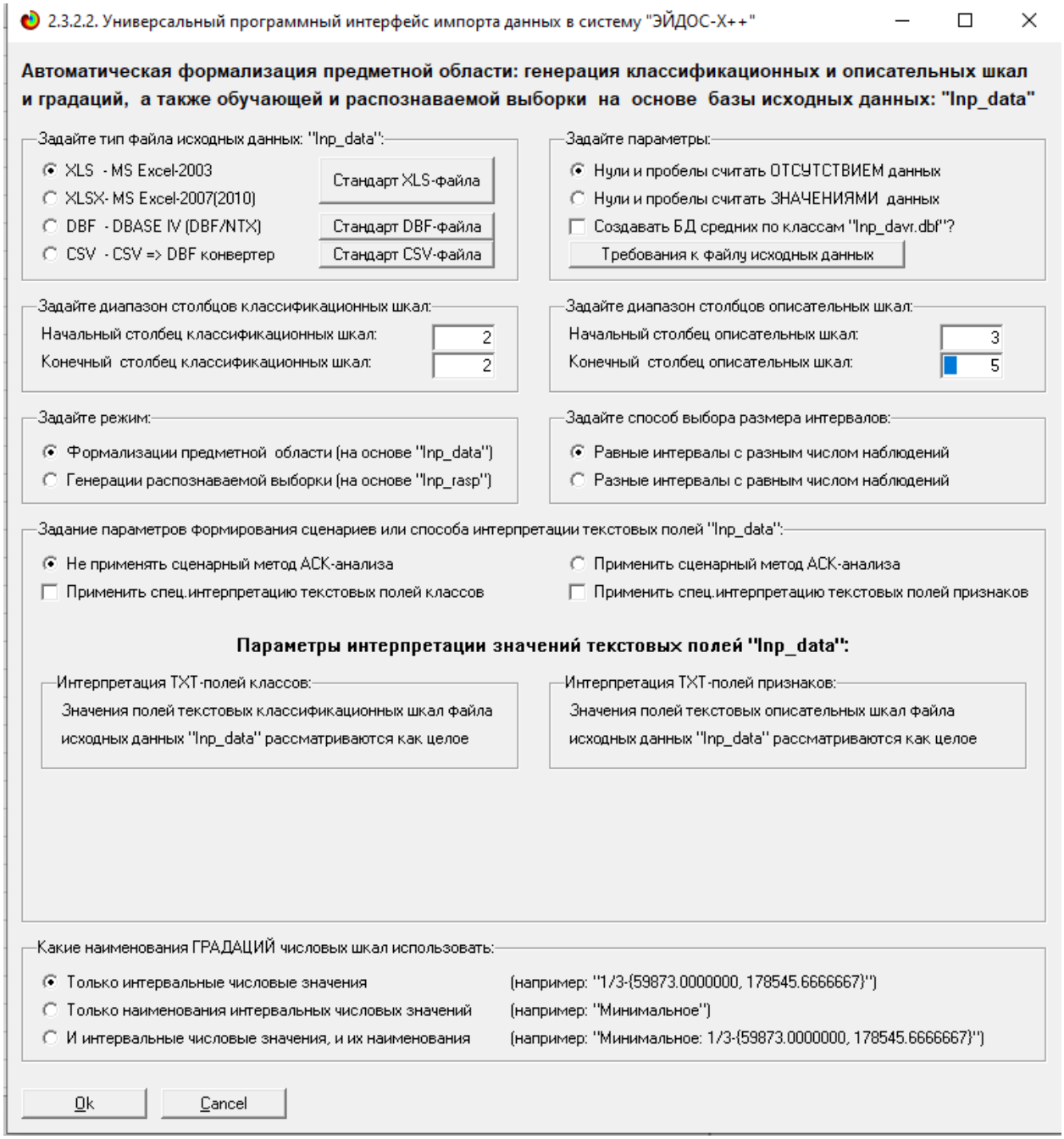


Рисунок 2 - Ввод обучающей выборки

Стоит отметить следующие настройки:

- Тип файла – xls;
- Классификационные шкалы – 2;
- Описательные шкалы – 3-5;
- 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).

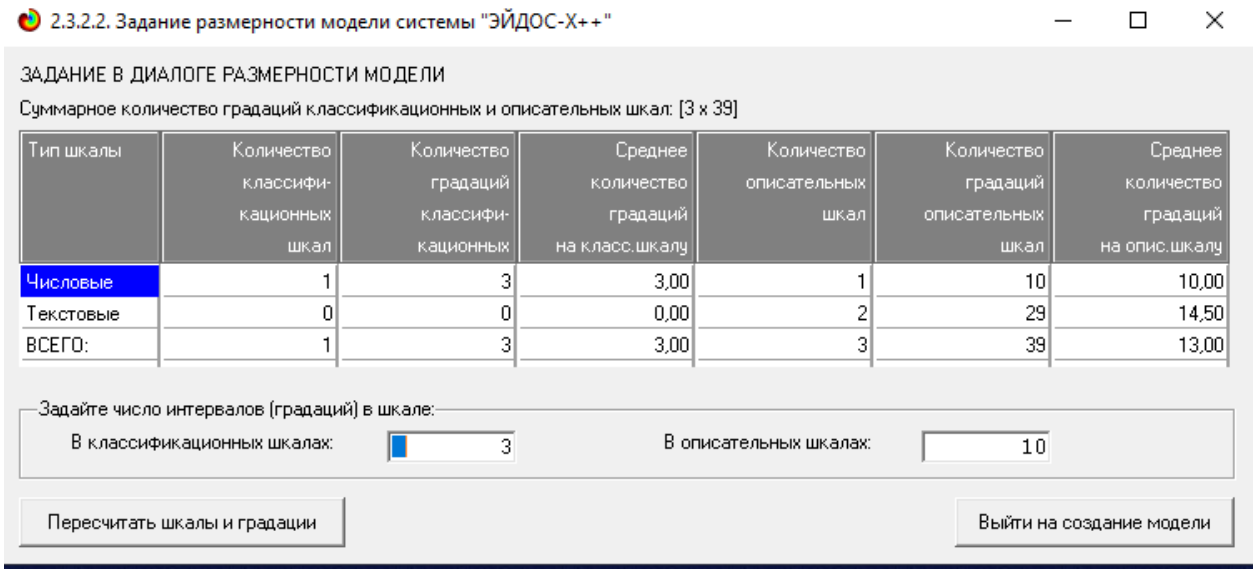


Рисунок 3 - Задание размерности модели системы Aidos-X

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

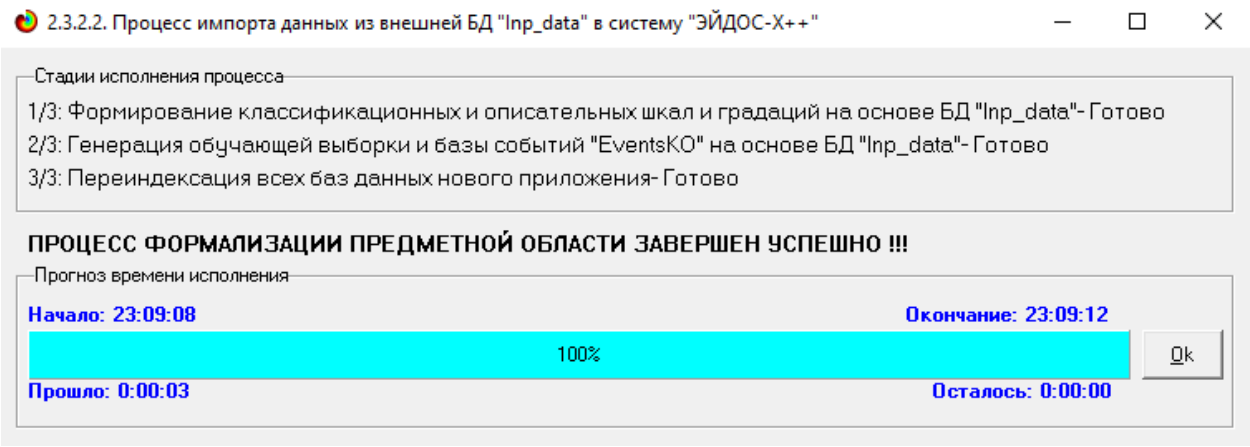


Рисунок 4 – Процесс импорта данных

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).

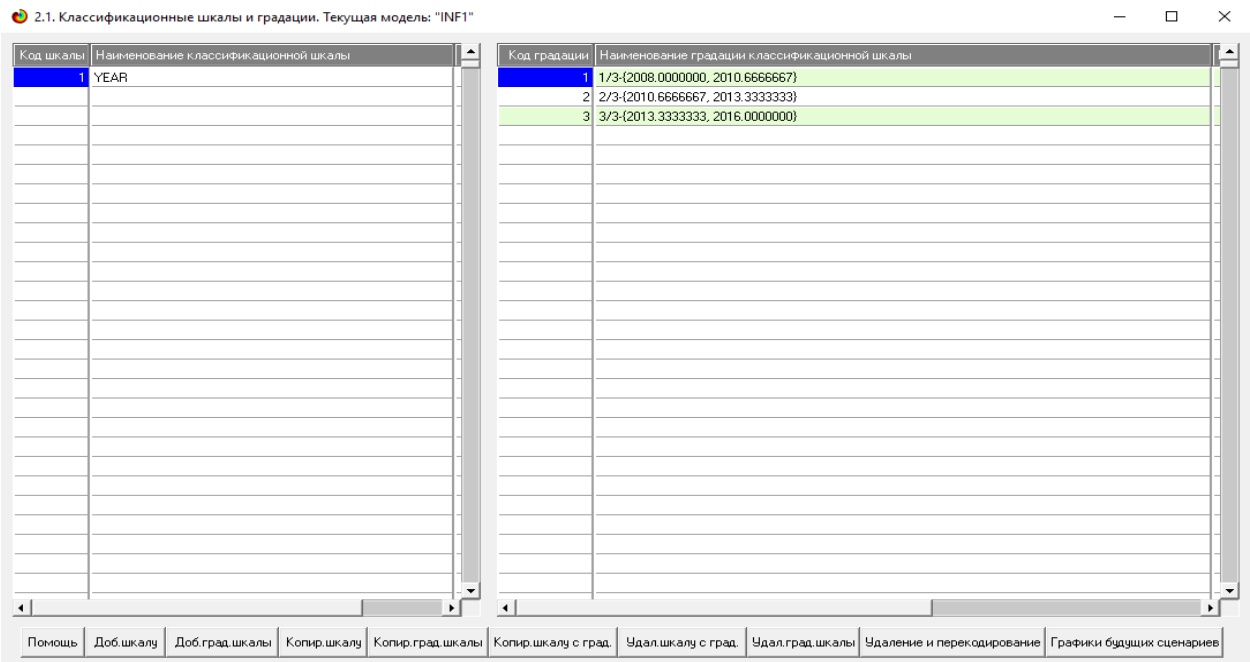


Рисунок 5 - Классификационные шкалы и градации

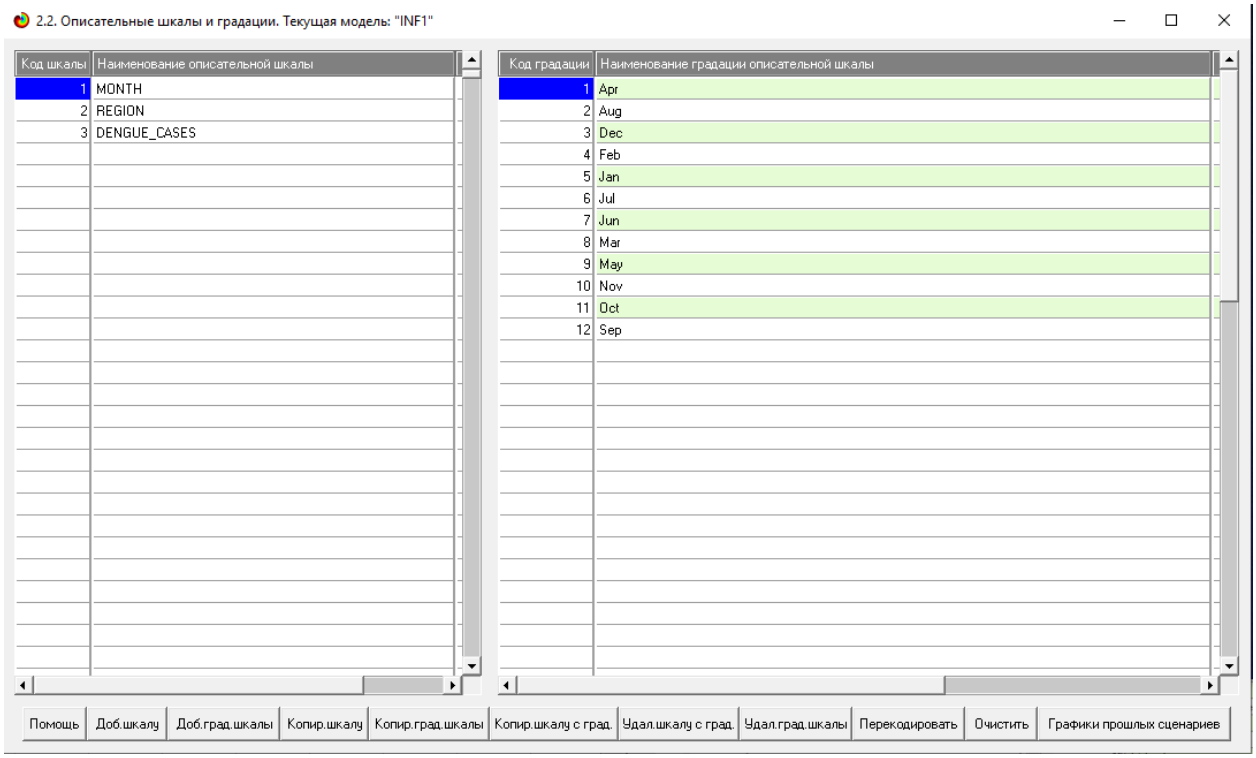


Рисунок 6 - Описательные шкалы и градации (фрагмент)

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.

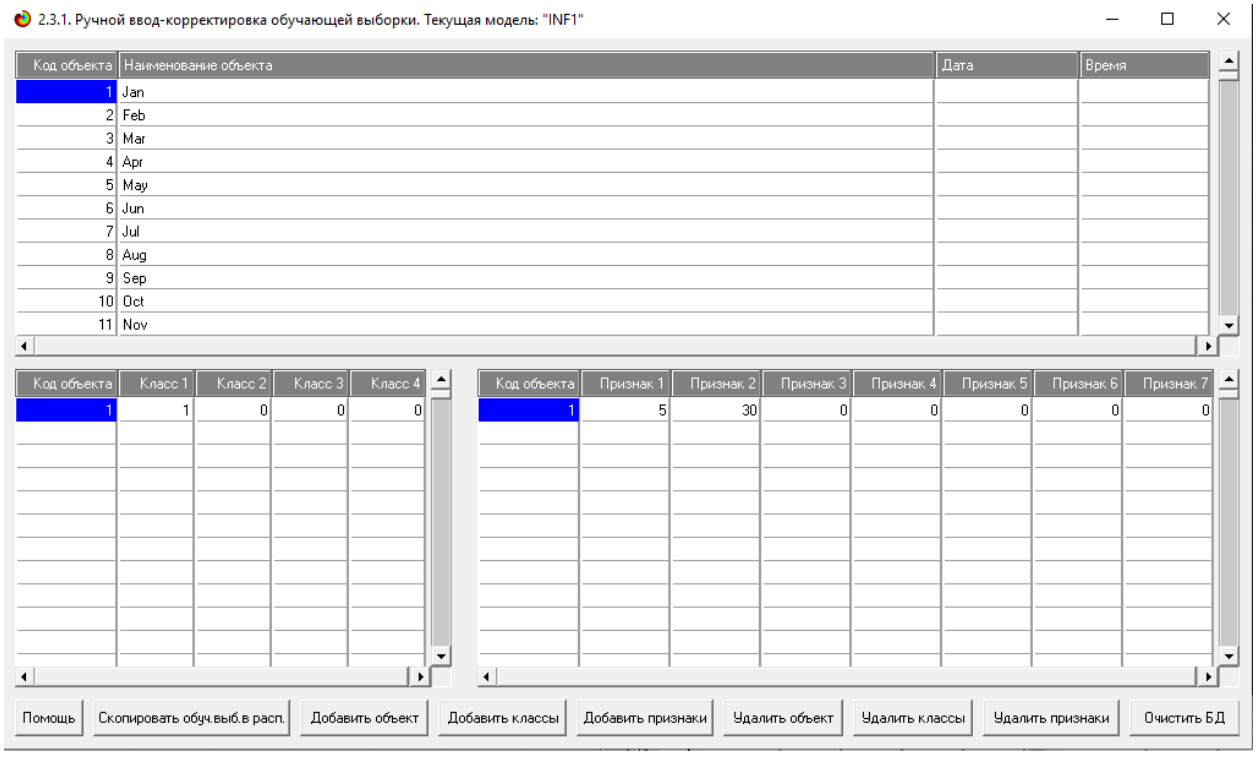


Рисунок 7 - Обучающая выборка (фрагмент)

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).

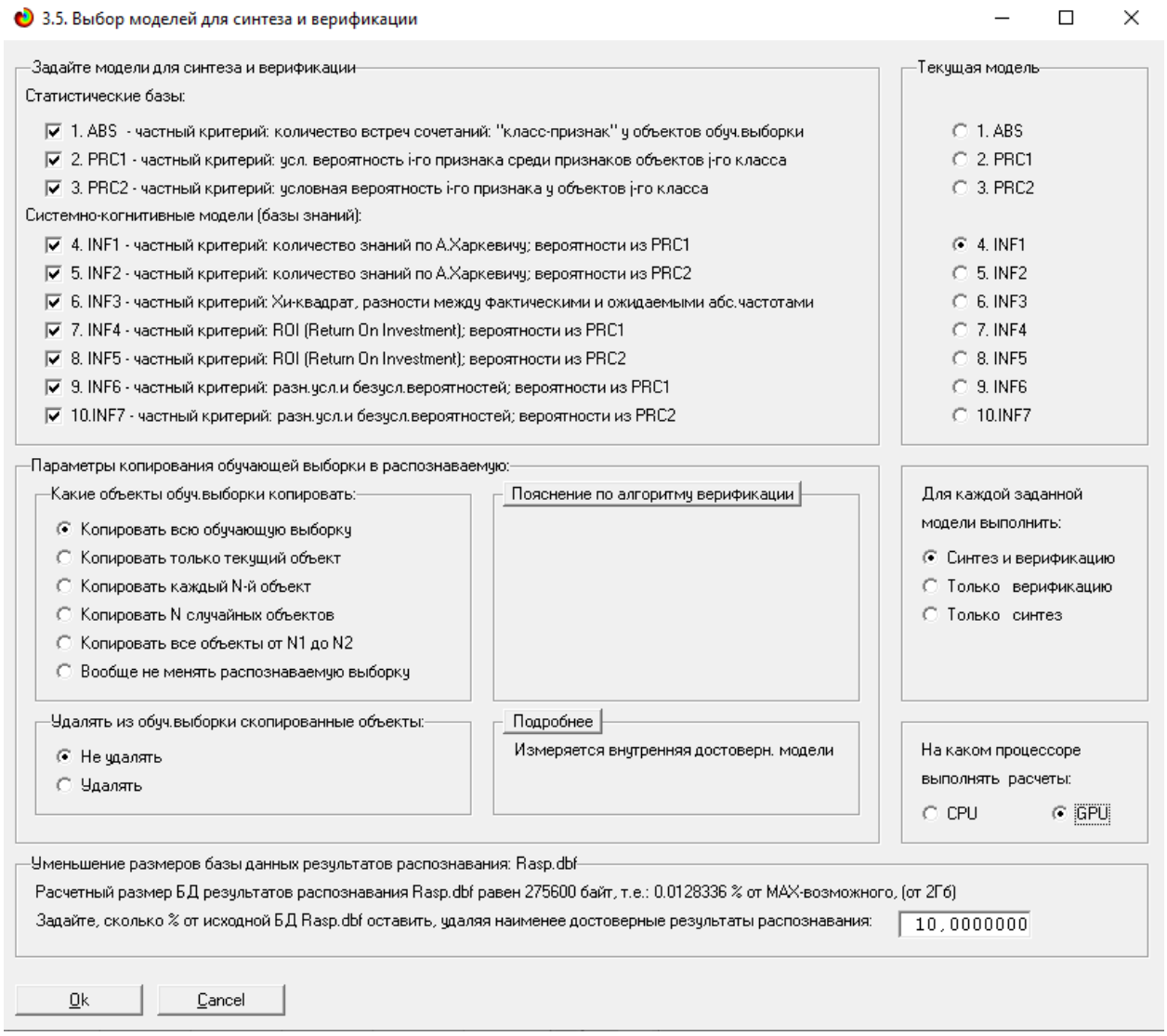


Рисунок 8 - Выбор моделей для синтеза и верификации

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.

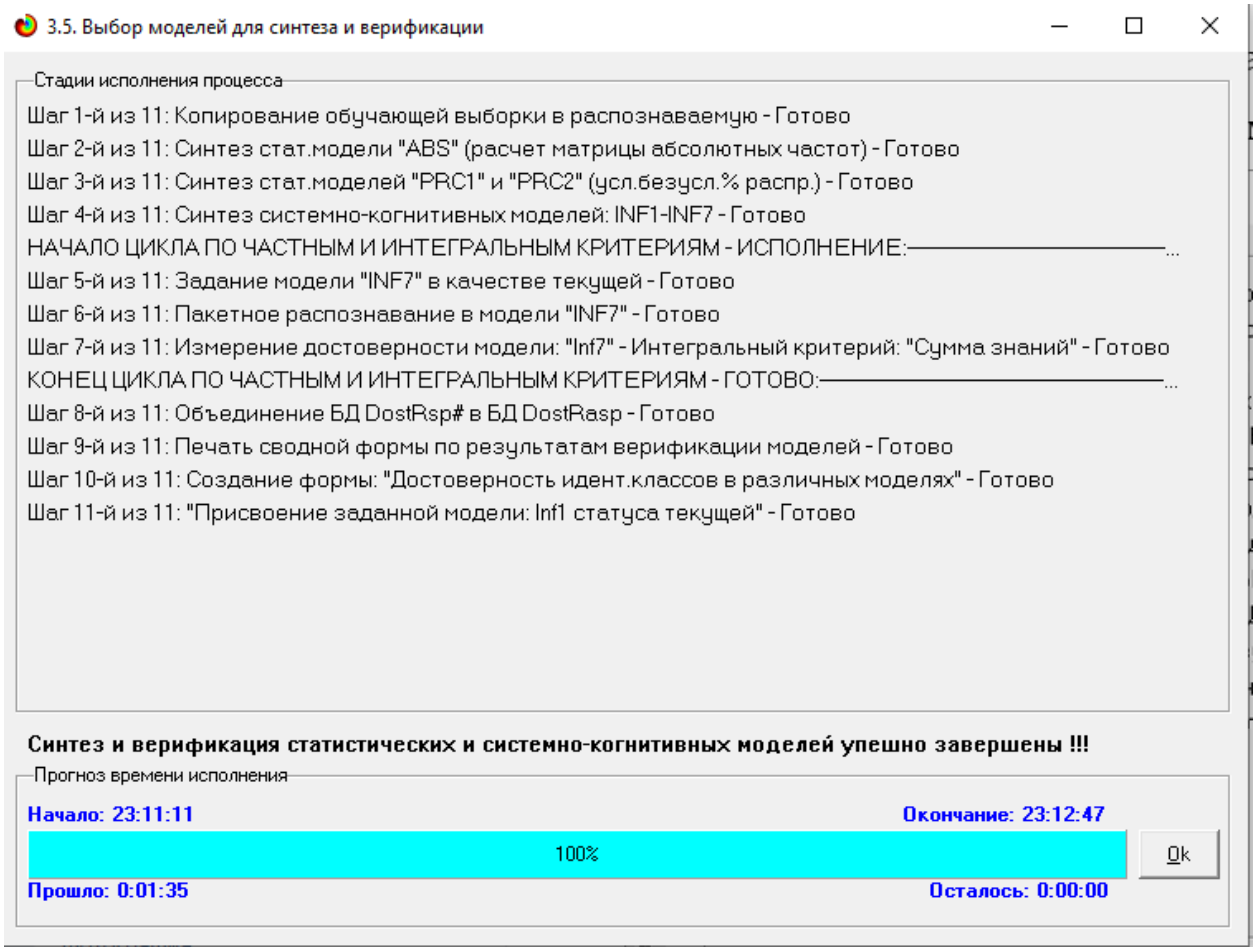


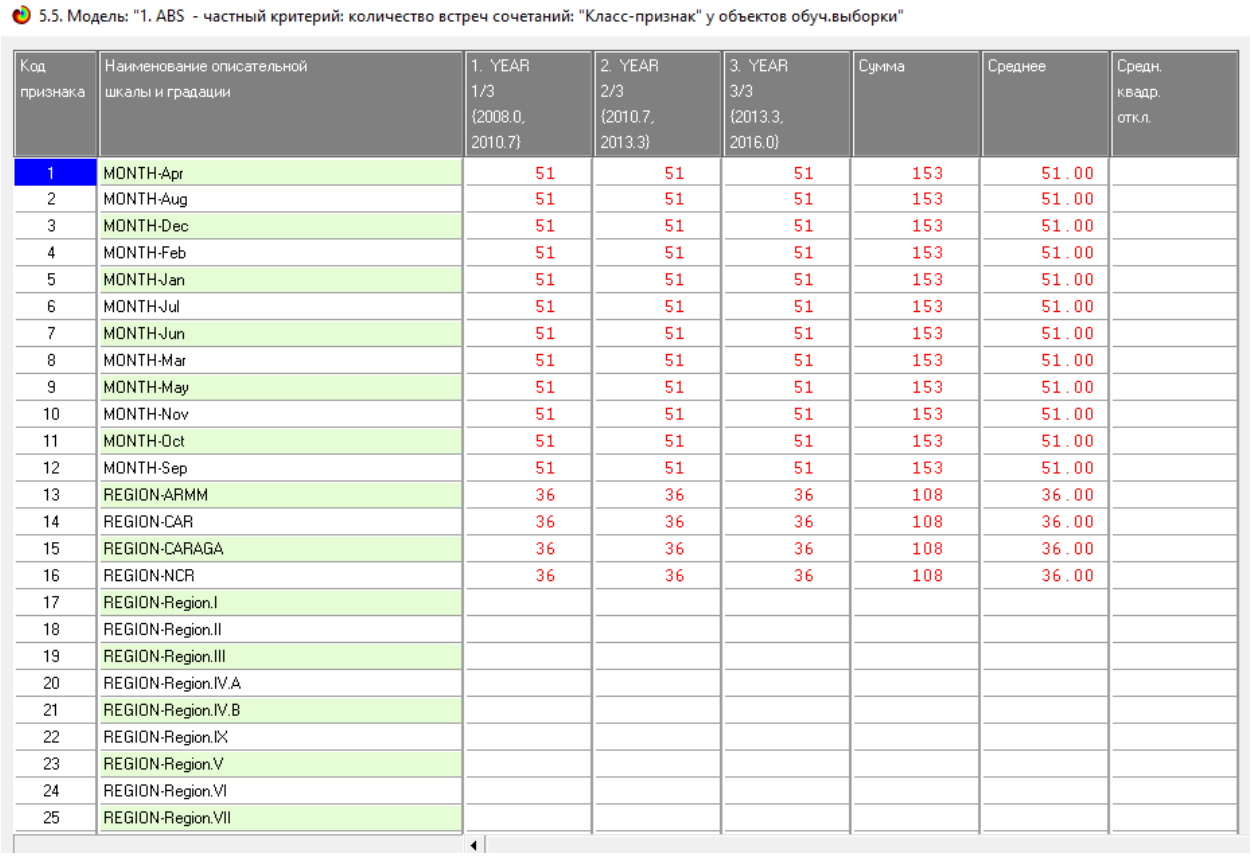
Рисунок 8 - Процесс исполнения режима 3.5

Интересно заметить (см. рисунок 8), что синтез и верификация всех 10 моделей на данной задаче заняли 1 минуту 35 секунд. При этом оценка достоверности моделей проводилась на 1837 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	1. YEAR 1/3 {2008.0, 2010.7}	2. YEAR 2/3 {2010.7, 2013.3}	3. YEAR 3/3 {2013.3, 2016.0}	Сумма	Среднее	Средн. квадр. откл.
1	MONTH-Apr	51	51	51	153	51.00	
2	MONTH-Aug	51	51	51	153	51.00	
3	MONTH-Dec	51	51	51	153	51.00	
4	MONTH-Feb	51	51	51	153	51.00	
5	MONTH-Jan	51	51	51	153	51.00	
6	MONTH-Jul	51	51	51	153	51.00	
7	MONTH-Jun	51	51	51	153	51.00	
8	MONTH-Mar	51	51	51	153	51.00	
9	MONTH-May	51	51	51	153	51.00	
10	MONTH-Nov	51	51	51	153	51.00	
11	MONTH-Oct	51	51	51	153	51.00	
12	MONTH-Sep	51	51	51	153	51.00	
13	REGION-ARMM	36	36	36	108	36.00	
14	REGION-CAR	36	36	36	108	36.00	
15	REGION-CARAGA	36	36	36	108	36.00	
16	REGION-NCR	36	36	36	108	36.00	
17	REGION-Region.I						
18	REGION-Region.II						
19	REGION-Region.III						
20	REGION-Region.IV.A						
21	REGION-Region.IV.B						
22	REGION-Region.IX						
23	REGION-Region.V						
24	REGION-Region.VI						
25	REGION-Region.VII						

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	1. YEAR 1/3 {2008.0, 2010.7}	2. YEAR 2/3 {2010.7, 2013.3}	3. YEAR 3/3 {2013.3, 2016.0}	Сумма	Среднее	Средн. квадр. откл.
1	MONTH-Apr	−0.001	−0.001	0.002	0.000	0.000	0.002
2	MONTH-Aug	−0.001	−0.001	0.002	0.000	0.000	0.002
3	MONTH-Dec	−0.001	−0.001	0.002	0.000	0.000	0.002
4	MONTH-Feb	−0.001	−0.001	0.002	0.000	0.000	0.002
5	MONTH-Jan	−0.001	−0.001	0.002	0.000	0.000	0.002
6	MONTH-Jul	−0.001	−0.001	0.002	0.000	0.000	0.002
7	MONTH-Jun	−0.001	−0.001	0.002	0.000	0.000	0.002
8	MONTH-Mar	−0.001	−0.001	0.002	0.000	0.000	0.002
9	MONTH-May	−0.001	−0.001	0.002	0.000	0.000	0.002
10	MONTH-Nov	−0.001	−0.001	0.002	0.000	0.000	0.002
11	MONTH-Oct	−0.001	−0.001	0.002	0.000	0.000	0.002
12	MONTH-Sep	−0.001	−0.001	0.002	0.000	0.000	0.002
13	REGION-ARMM	−0.001	−0.001	0.002	0.000	0.000	0.002
14	REGION-CAR	−0.001	−0.001	0.002	0.000	0.000	0.002
15	REGION-CARAGA	−0.001	−0.001	0.002	0.000	0.000	0.002
16	REGION-NCR	−0.001	−0.001	0.002	0.000	0.000	0.002
17	REGION-Region.I						
18	REGION-Region.II						
19	REGION-Region.III						
20	REGION-Region.IV.A						
21	REGION-Region.IV.B						
22	REGION-Region.IX						
23	REGION-Region.V						
24	REGION-Region.VI						
25	REGION-Region.VII						

Рисунок 10 - Модель INF4 (фрагмент)

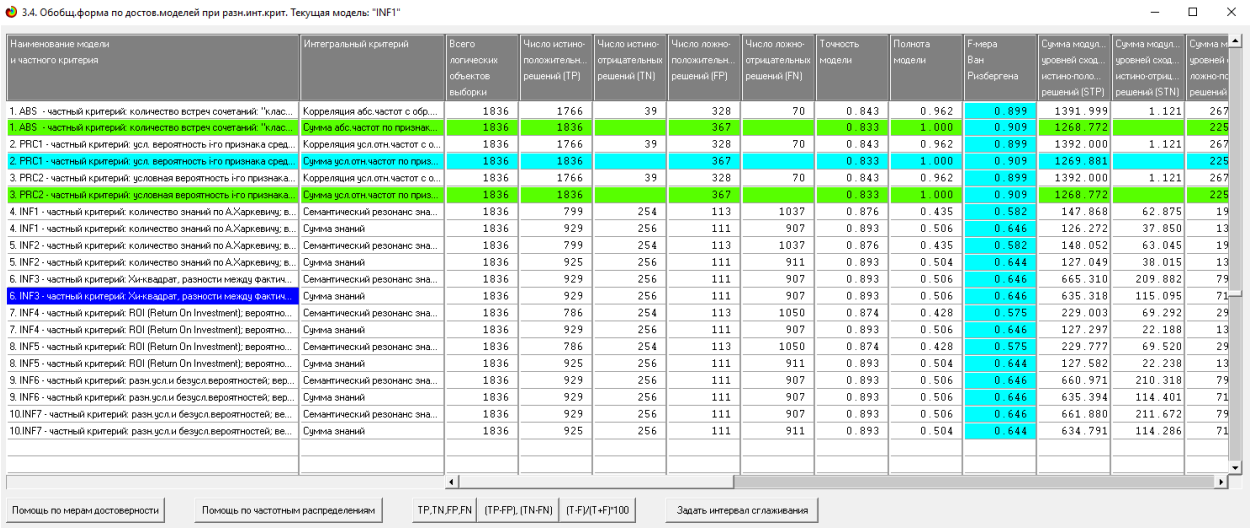
5.5. Модель: "8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2"

Код признака	Наименование описательной шкалы и градации	1. YEAR 1/3 {2008.0, 2010.7}	2. YEAR 2/3 {2010.7, 2013.3}	3. YEAR 3/3 {2013.3, 2016.0}	Сумма	Среднее	Средн. квадр. откл.
1	MONTH-Apr						
2	MONTH-Aug						
3	MONTH-Dec						
4	MONTH-Feb						
5	MONTH-Jan						
6	MONTH-Jul						
7	MONTH-Jun						
8	MONTH-Mar						
9	MONTH-May						
10	MONTH-Nov						
11	MONTH-Oct						
12	MONTH-Sep						
13	REGION-ARMM						
14	REGION-CAR						
15	REGION-CARAGA						
16	REGION-NCR						
17	REGION-Region.I						
18	REGION-Region.II						
19	REGION-Region.III						
20	REGION-Region.IV.A						
21	REGION-Region.IV.B						
22	REGION-Region.IX						
23	REGION-Region.V						
24	REGION-Region.VI						
25	REGION-Region.VII						

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.



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....	1836	1766	39	328	70	0.843	0.962	0.899	1391.999	1.121	267
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс.частот по признак...	1836	1836		367		0.833	1.000	0.909	1268.772		225
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	1836	1766	39	328	70	0.843	0.962	0.899	1392.000	1.121	267
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	1836	1836		367		0.833	1.000	0.909	1269.881		225
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	1836	1766	39	328	70	0.843	0.962	0.899	1392.000	1.121	267
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	1836	1836		367		0.833	1.000	0.909	1268.772		225
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	1836	799	254	113	1037	0.876	0.435	0.582	147.868	62.875	19
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	1836	929	256	111	907	0.893	0.506	0.646	126.272	37.850	13
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	1836	799	254	113	1037	0.876	0.435	0.582	148.052	63.045	19
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	1836	925	256	111	911	0.893	0.504	0.644	127.049	38.015	13
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	1836	929	256	111	907	0.893	0.506	0.646	665.310	209.882	79
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	1836	929	256	111	907	0.893	0.506	0.646	635.318	115.095	71
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	1836	786	254	113	1050	0.874	0.428	0.575	229.003	69.292	29
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	1836	929	256	111	907	0.893	0.506	0.646	127.297	22.188	13
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	1836	786	254	113	1050	0.874	0.428	0.575	229.777	69.520	29
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	1836	925	256	111	911	0.893	0.504	0.644	127.582	22.238	13
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Семантический резонанс зна...	1836	929	256	111	907	0.893	0.506	0.646	660.971	210.318	79
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер...	Сумма знаний	1836	929	256	111	907	0.893	0.506	0.646	635.394	114.401	71
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Семантический резонанс зна...	1836	929	256	111	907	0.893	0.506	0.646	661.880	211.672	79
10.INF7 - частный критерий: разн.усл.и безусл.вероятностей; ве...	Сумма знаний	1836	925	256	111	911	0.893	0.504	0.644	634.791	114.286	71

Рисунок 12 - Оценки достоверности моделей

Наиболее достоверной в данном приложении оказалась модель INF1 при интегральном критерии «Сумма знаний». При этом точность модели (F-мера Ван Ризбергена) составляет 0,690 а точность модели (L1-мера профессора Луценко) - 0,976. 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).

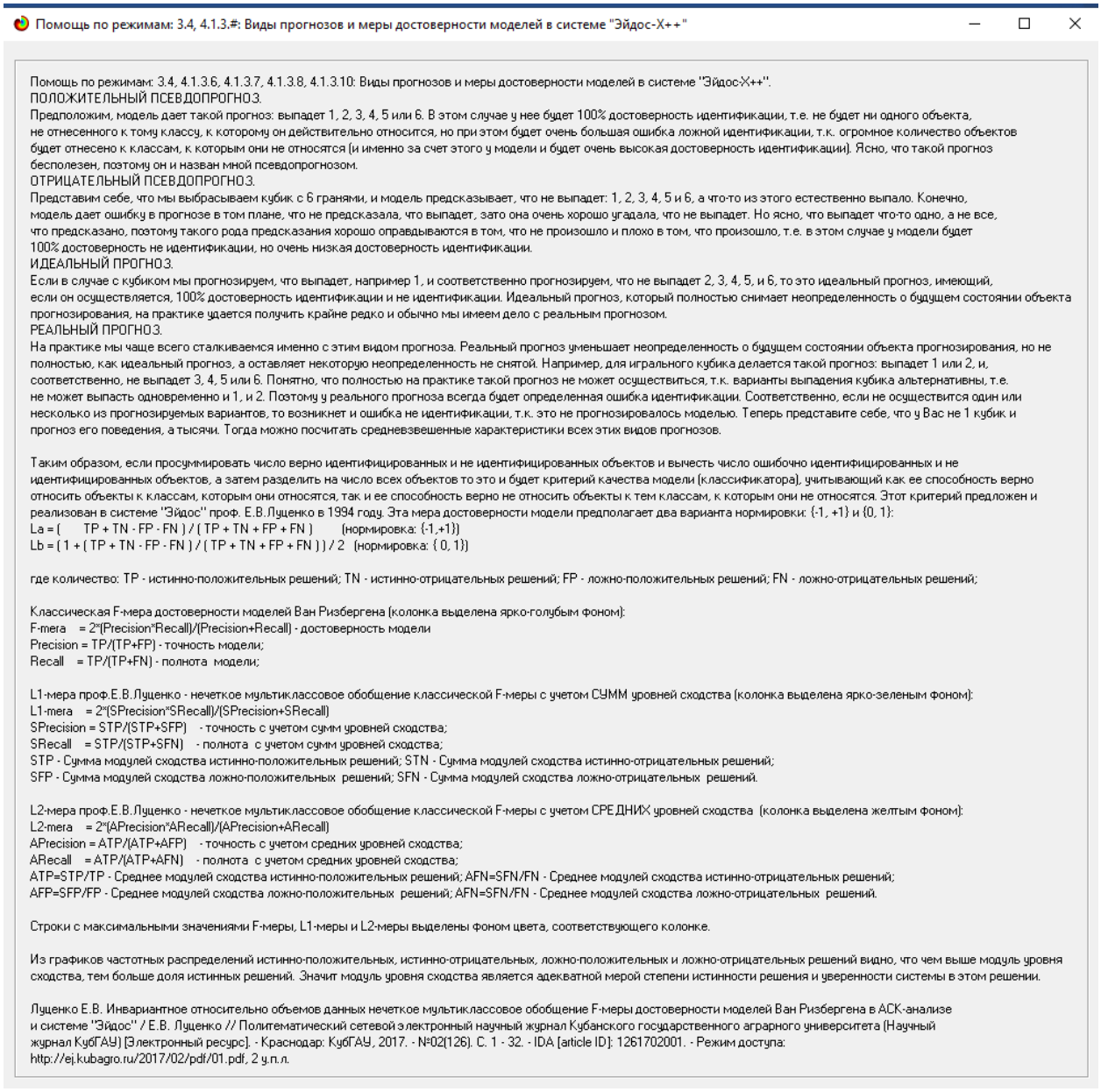
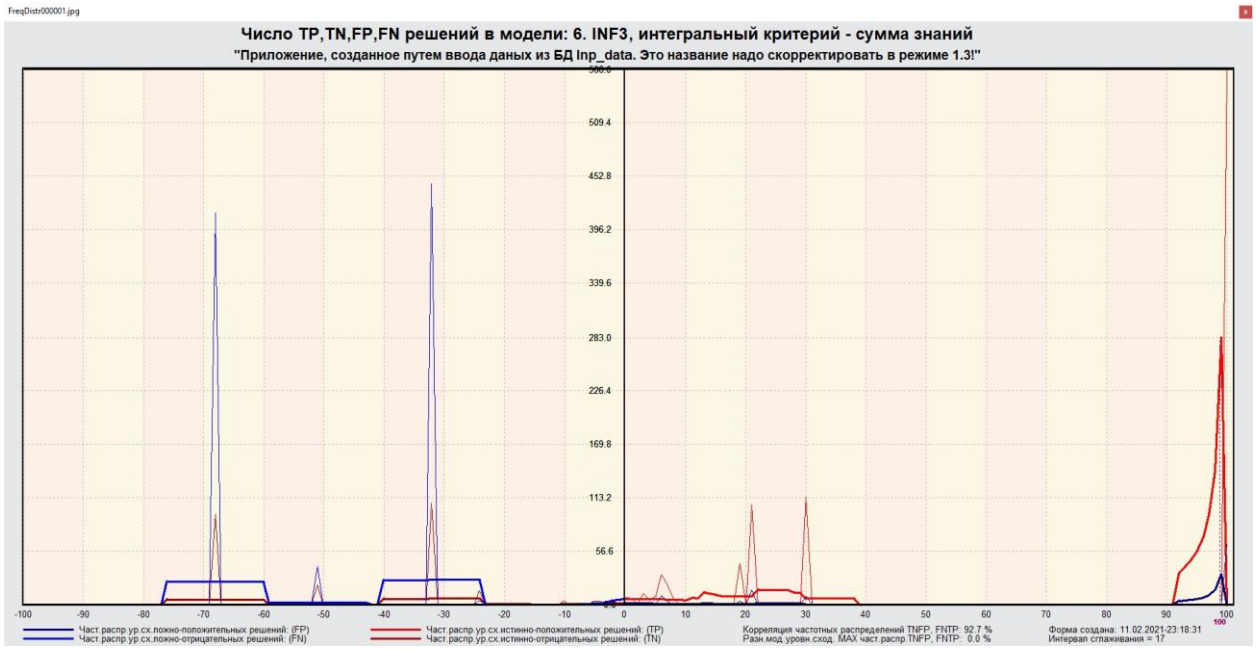


Рисунок 13- Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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 INF3.



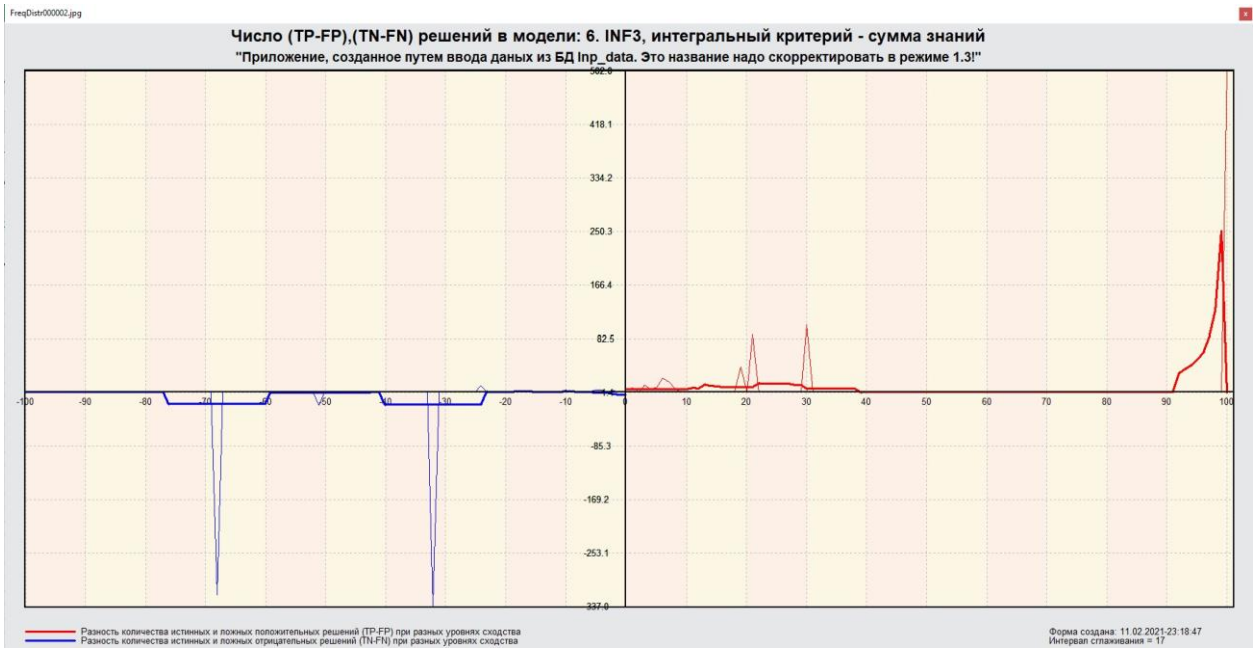


Рисунок 14 - Частотные распределения числа верно и ошибочно идентифицированных и не идентифицированных состояний объекта моделирования в зависимости от уровня сходства в модели INF3

Из рисунка 14 видно, что:

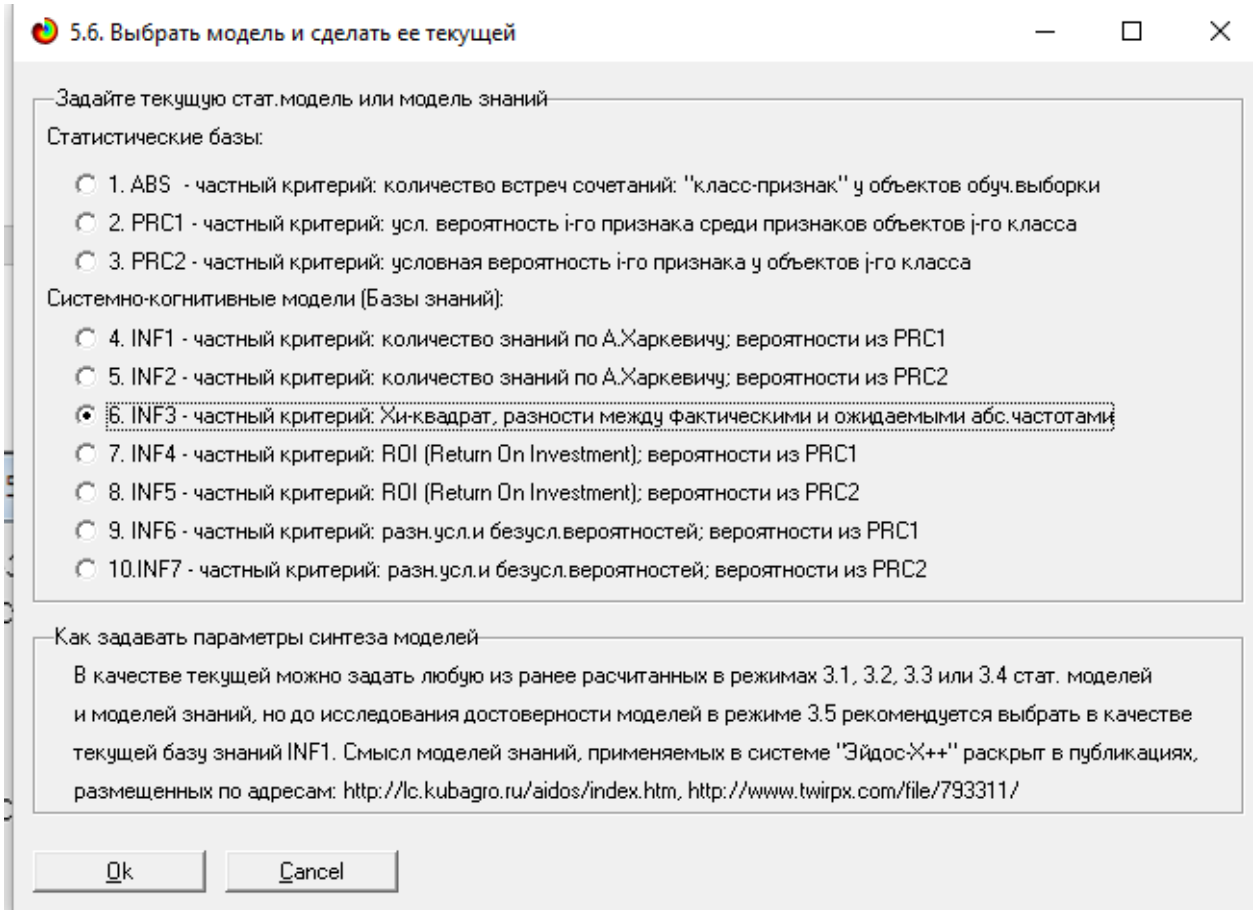
– наиболее модель INF3 лучше определяет непринадлежность объекта к классу, чем принадлежность;

– модуль уровня сходства-различия в модели INF3 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 INF3 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1.



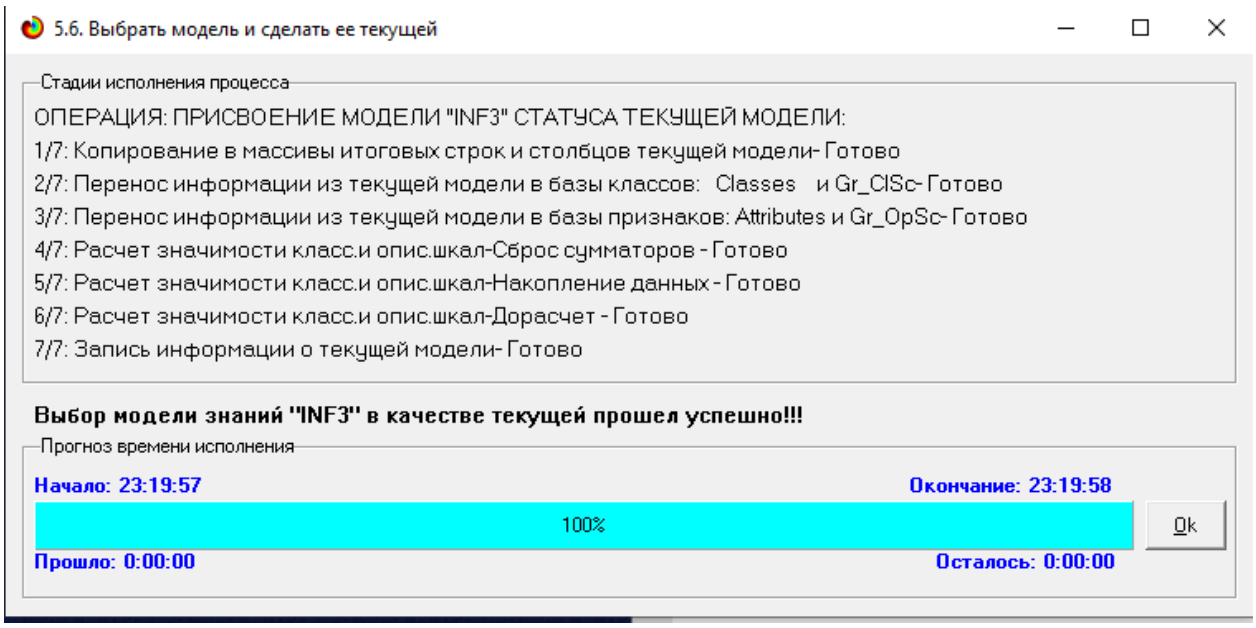


Рисунок 15 - Экранные формы режима задания модели в качестве текущей

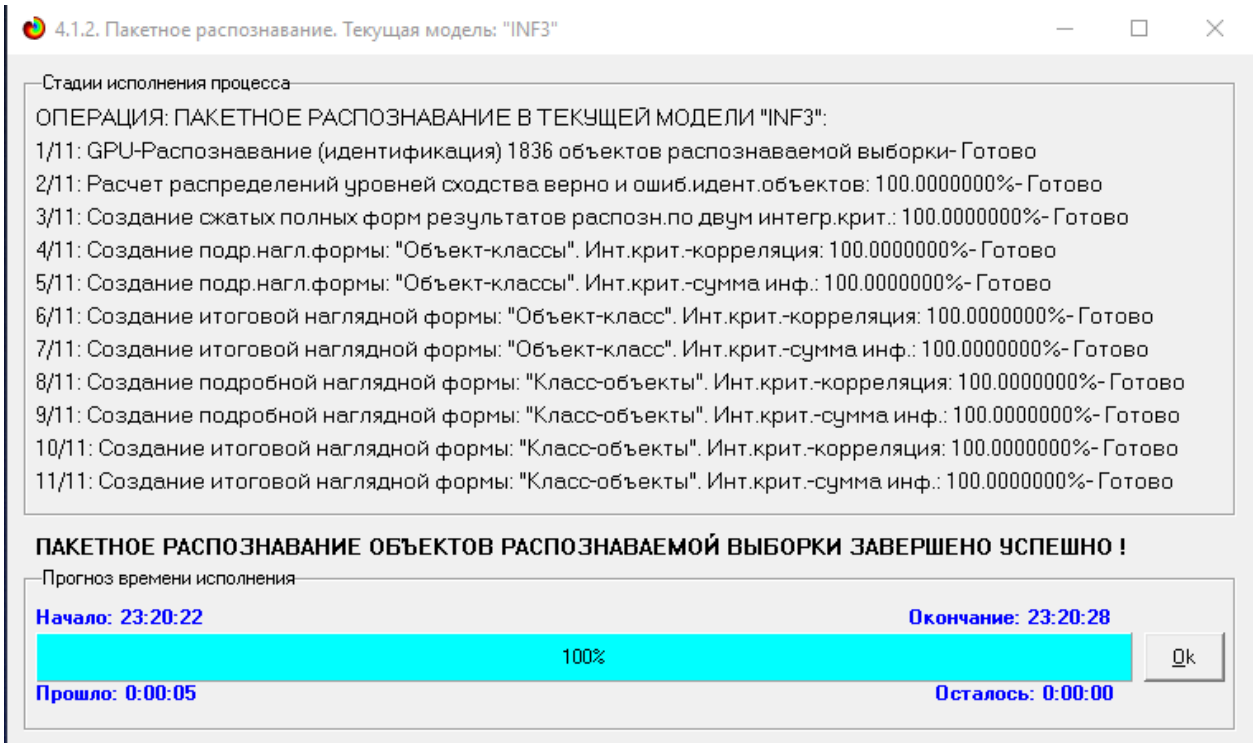


Рисунок 16 - Экранная форма режима пакетного распознавания в текущей модели

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 INF3 на основе наблюдения предыстории их развития:

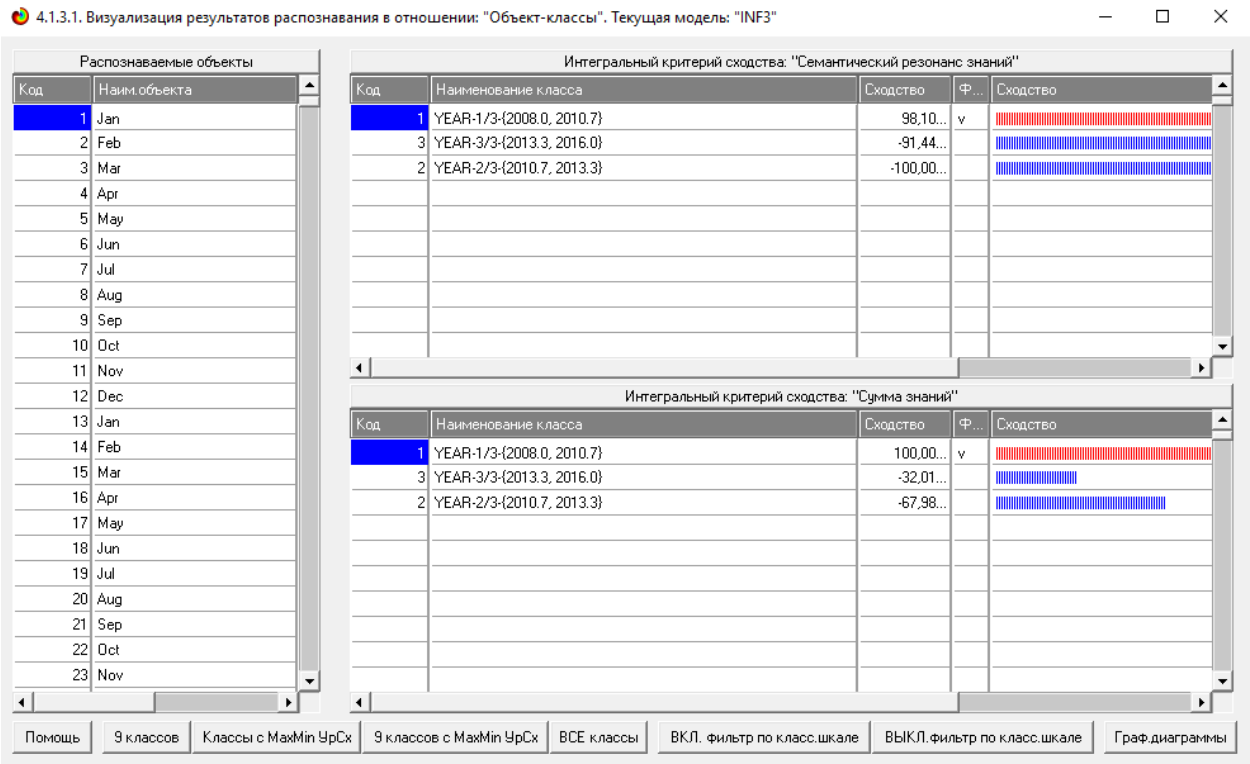


Рисунок 17. Пример идентификации объектов в модели INF3

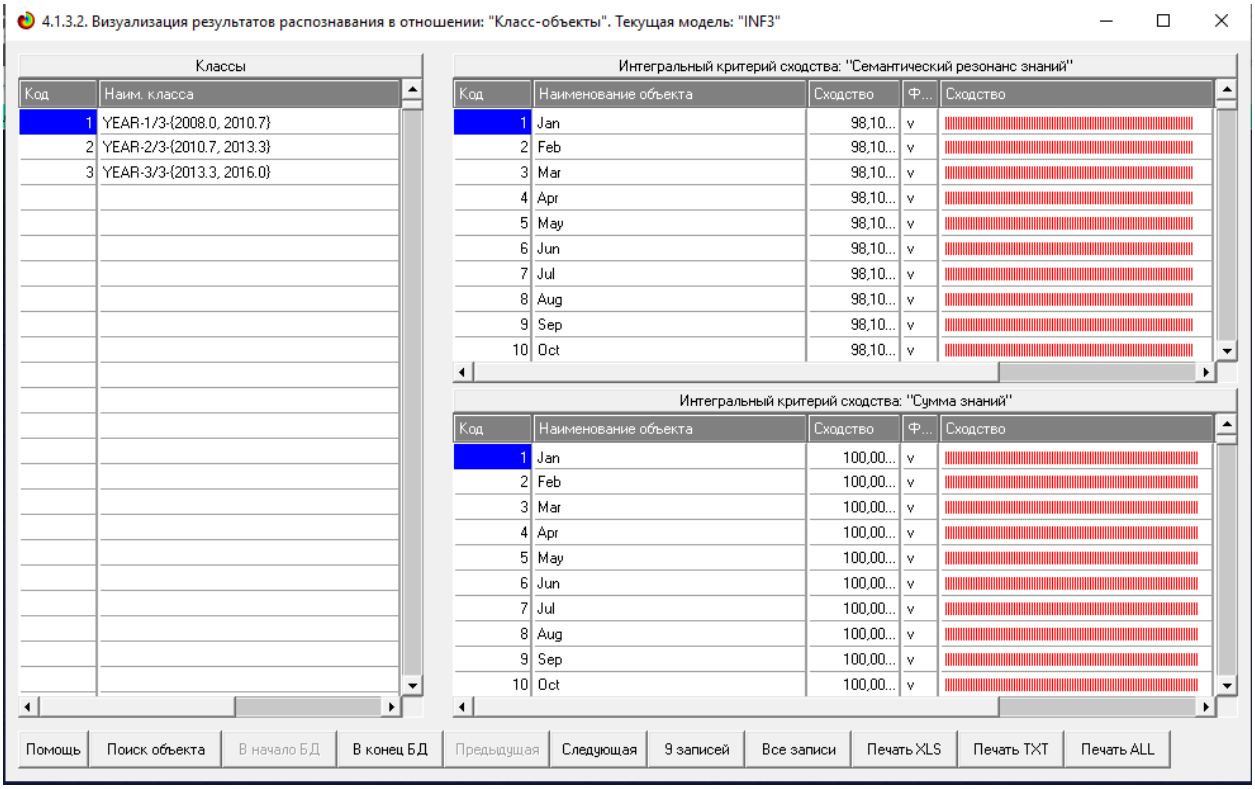


Рисунок 18. Пример идентификации классов в модели INF3

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:

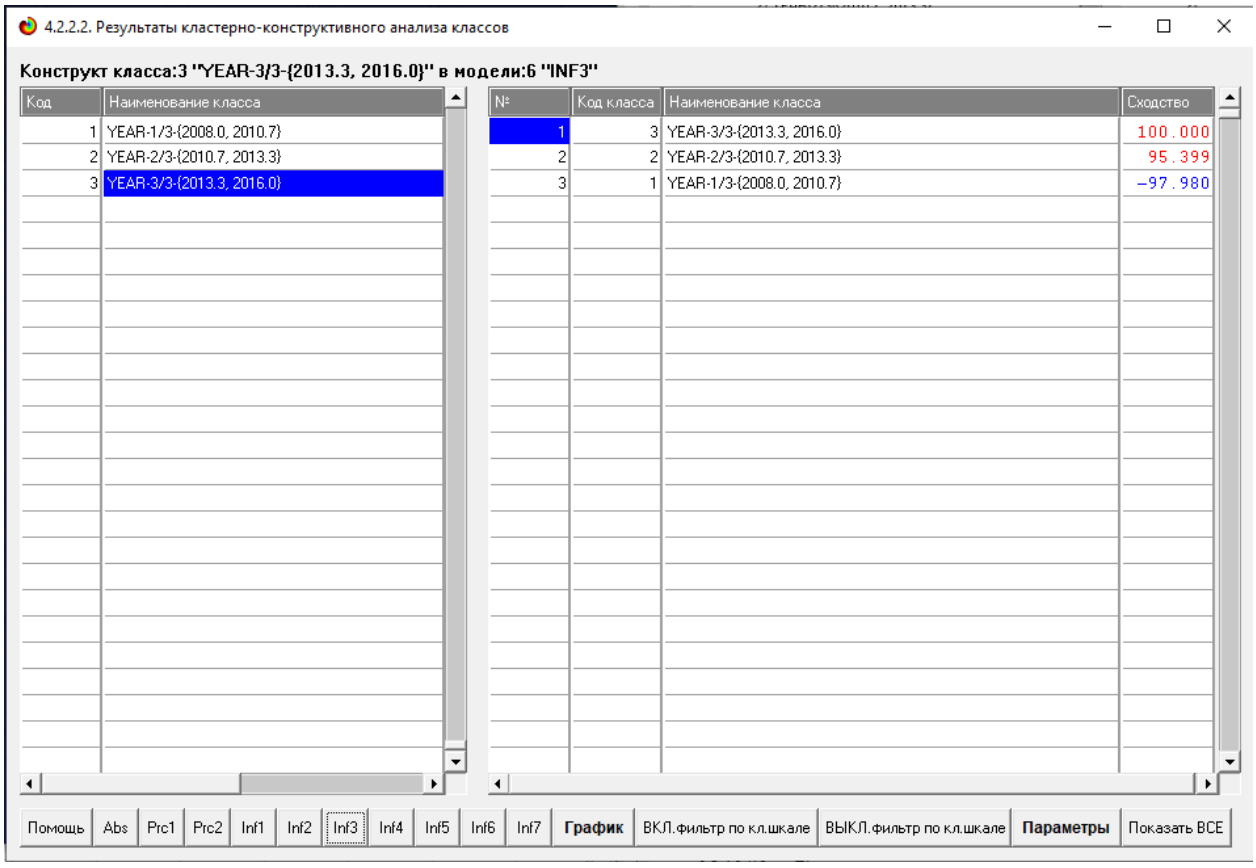


Рисунок 19 - Результаты кластерно-конструктивного анализа классов

На рисунке 20 представлена семантическая сеть классов в модели «INF3».

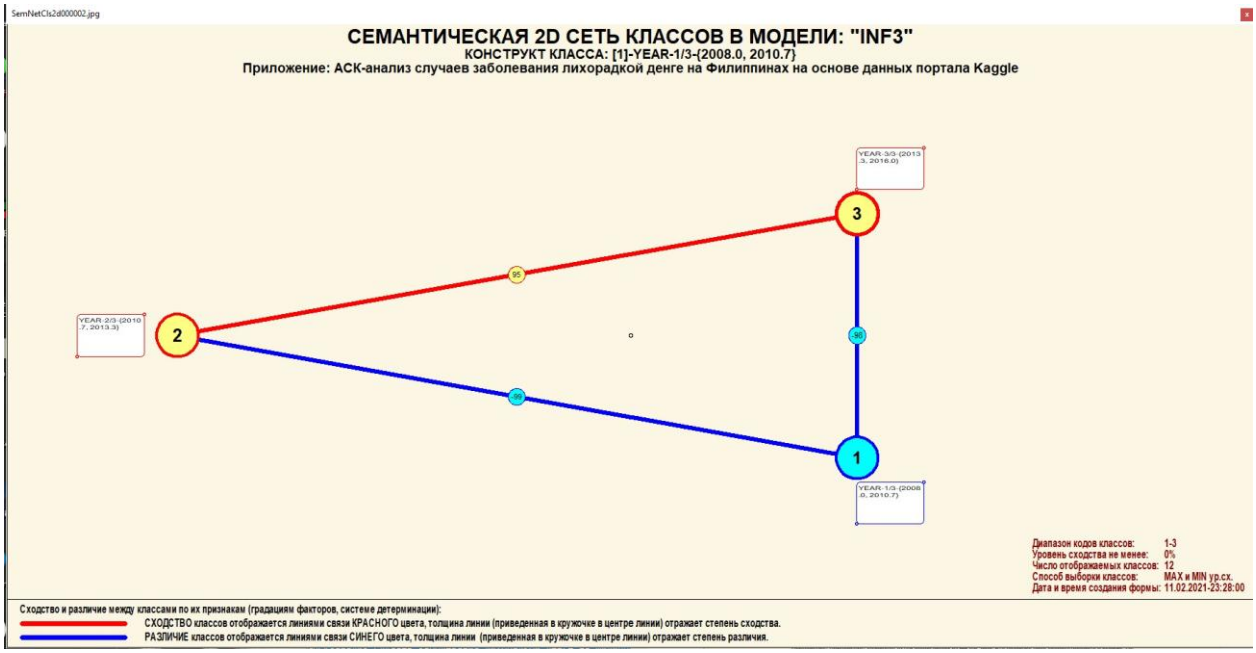


Рисунок 20 - Семантическая 2D сеть классов

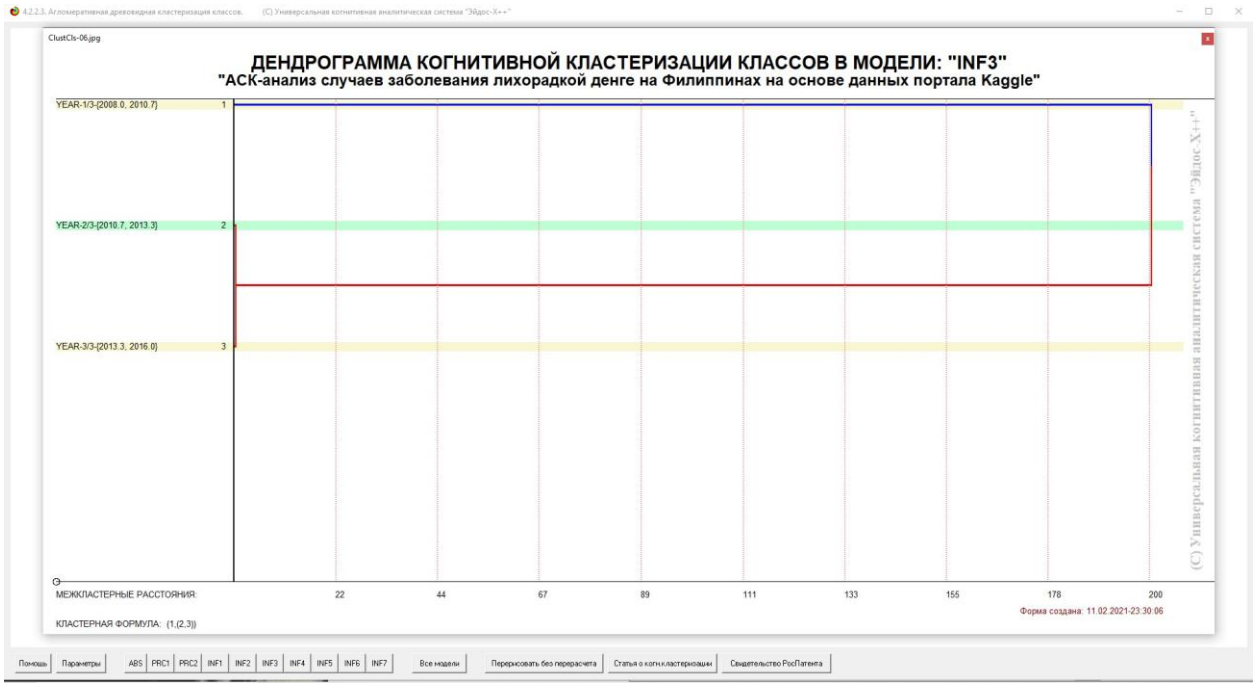


Рисунок 21 – Агломеративная дендрограмма классов

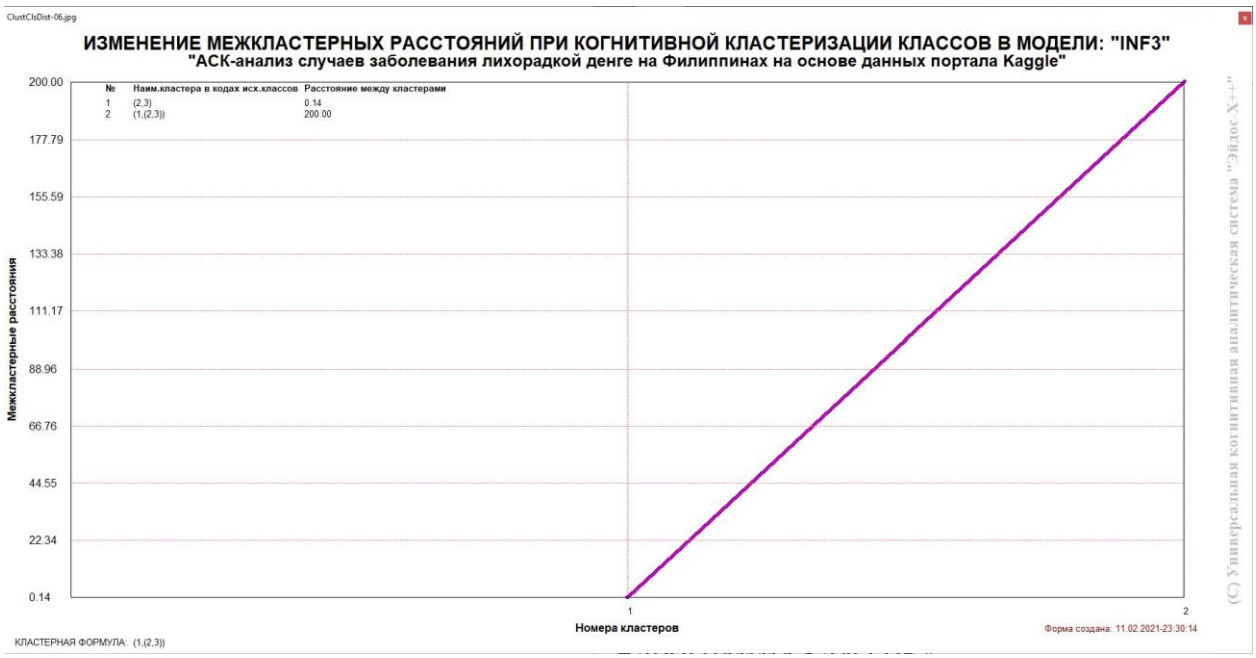


Рисунок 22 – График межкластерных расстояний классов

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

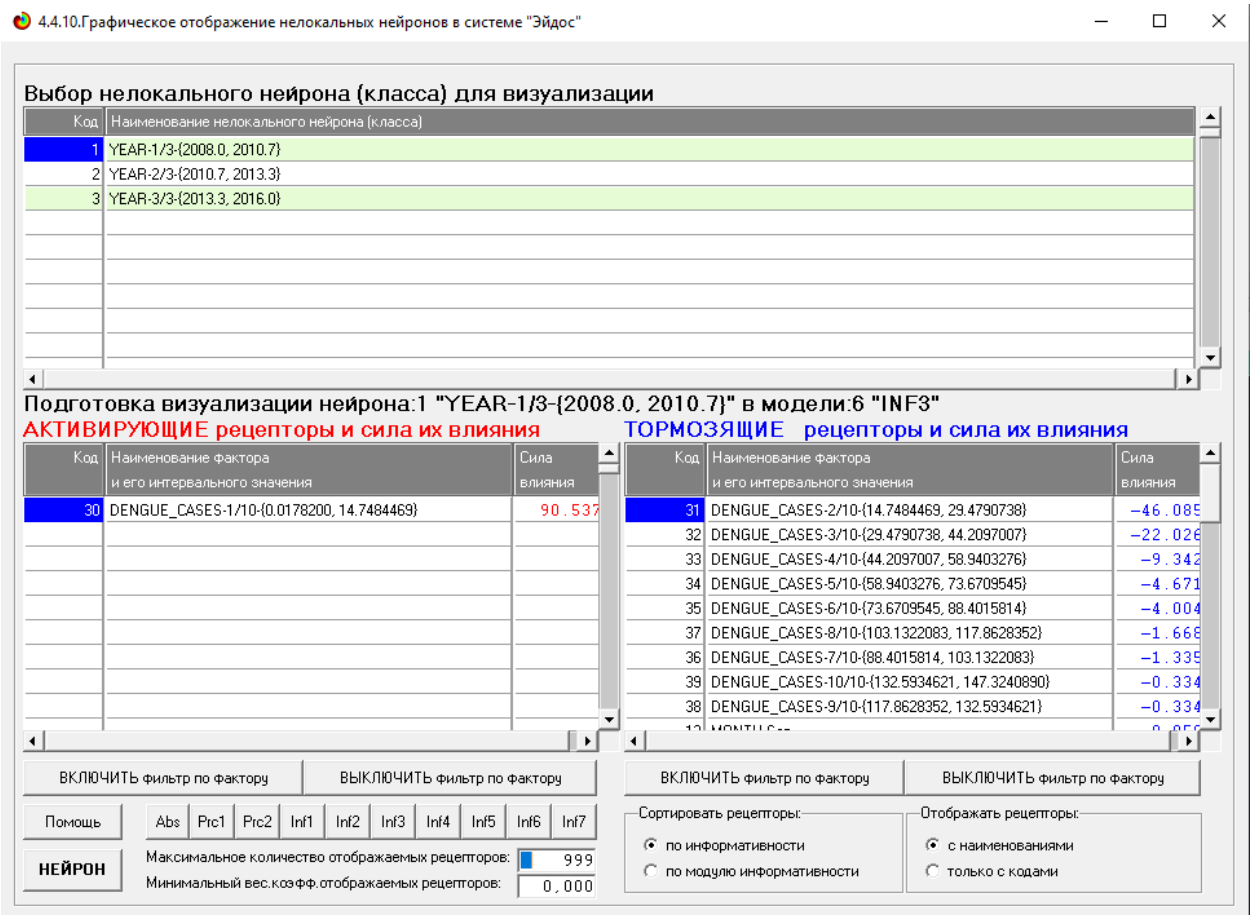


Рисунок 24 - Графическое отображение нелокальных нейронов в системе Aidos-X

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.

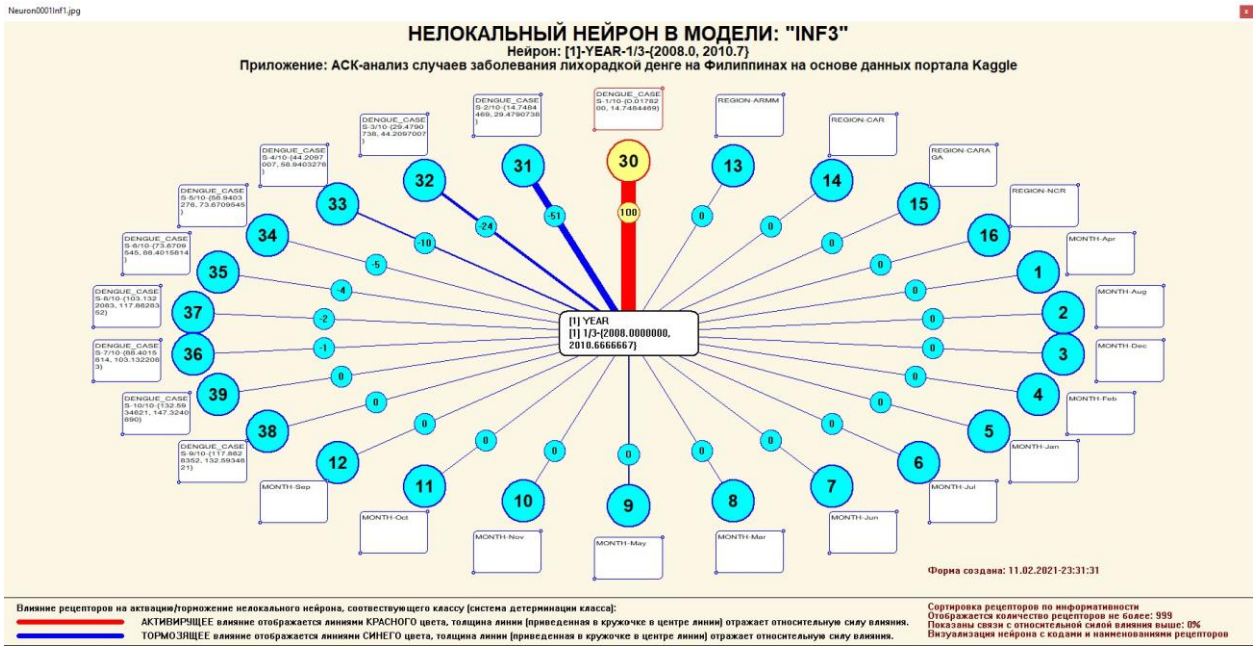


Рисунок 25 - Нелокальный нейрон в модели “INF3”

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].

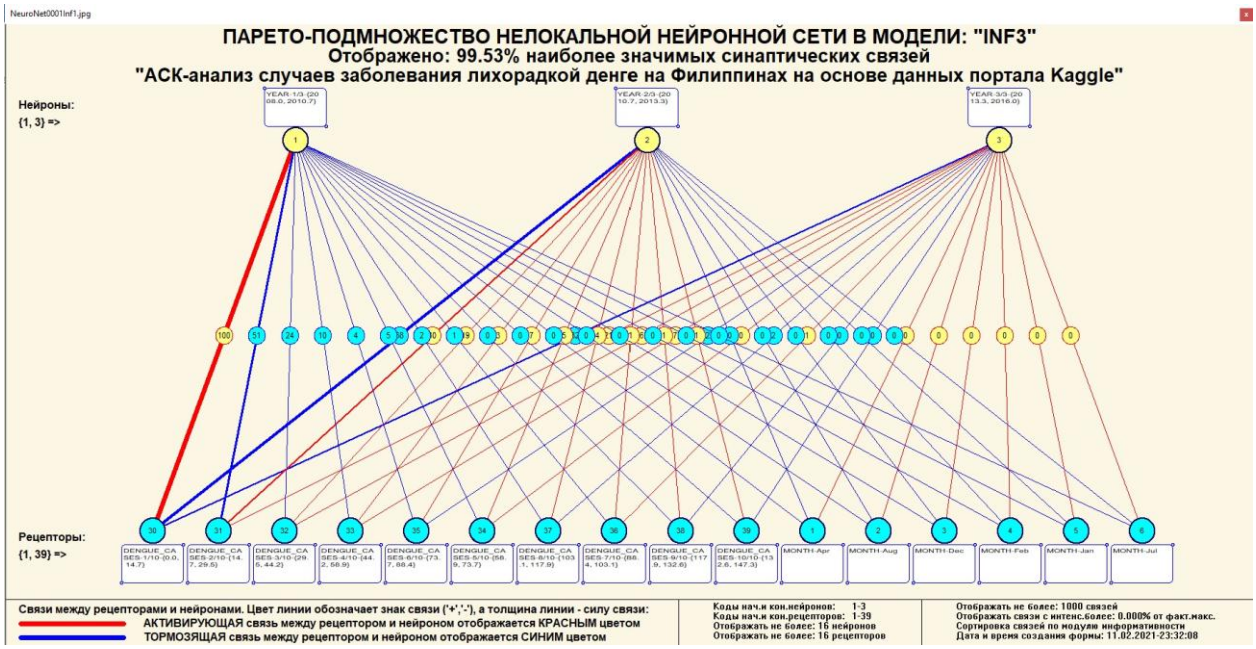
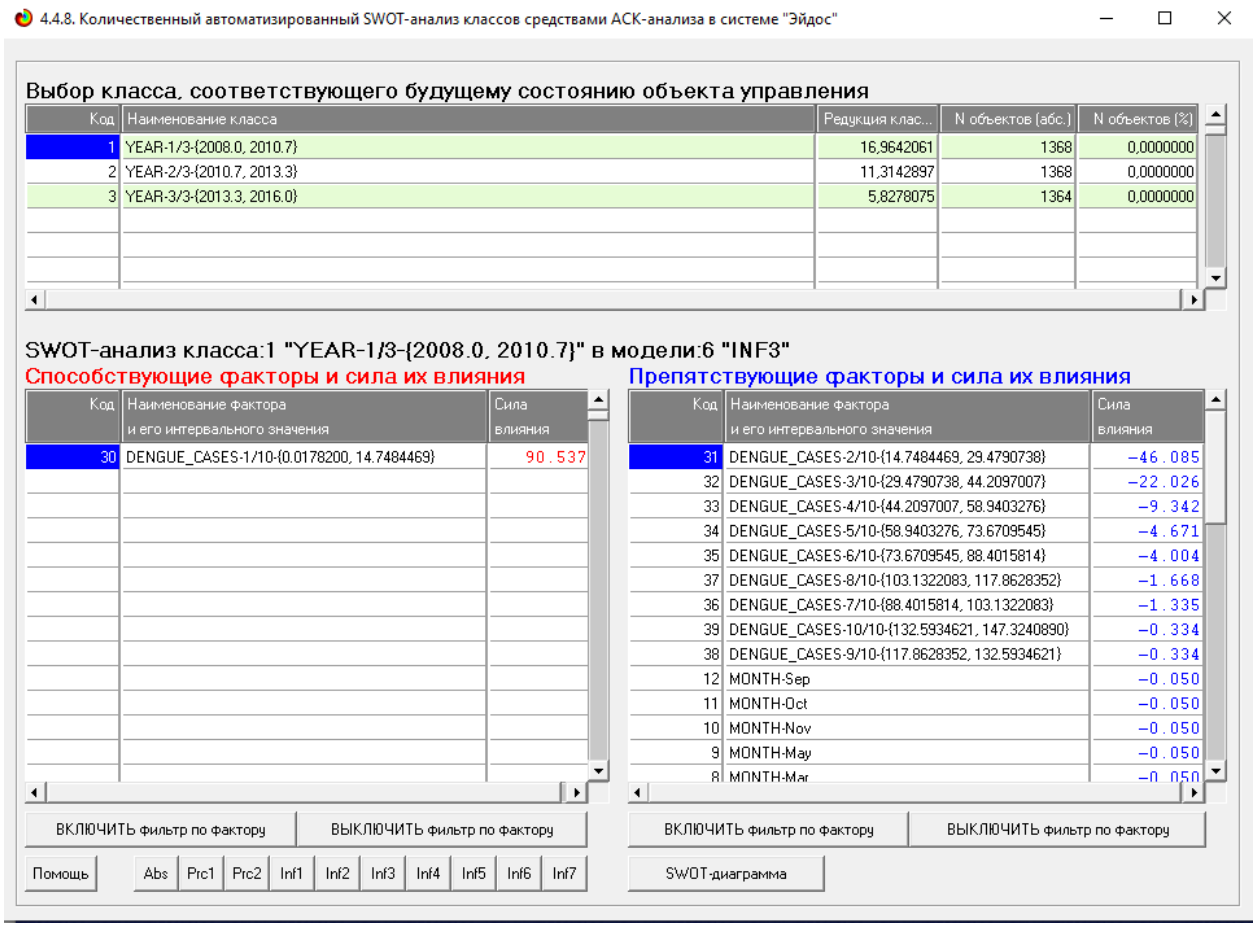


Рисунок 26 - Паретто – подмножество нелокальной нейронной сети

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.



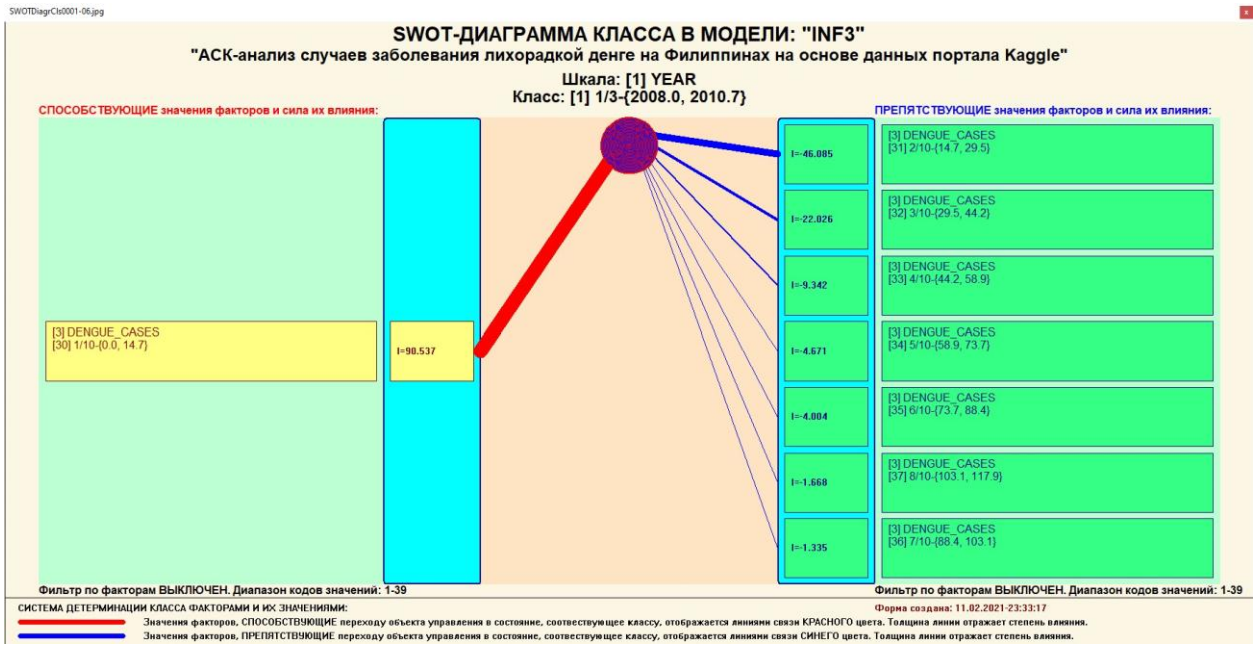


Рисунок 27 - Пример SWOT-матрицы в модели INF3

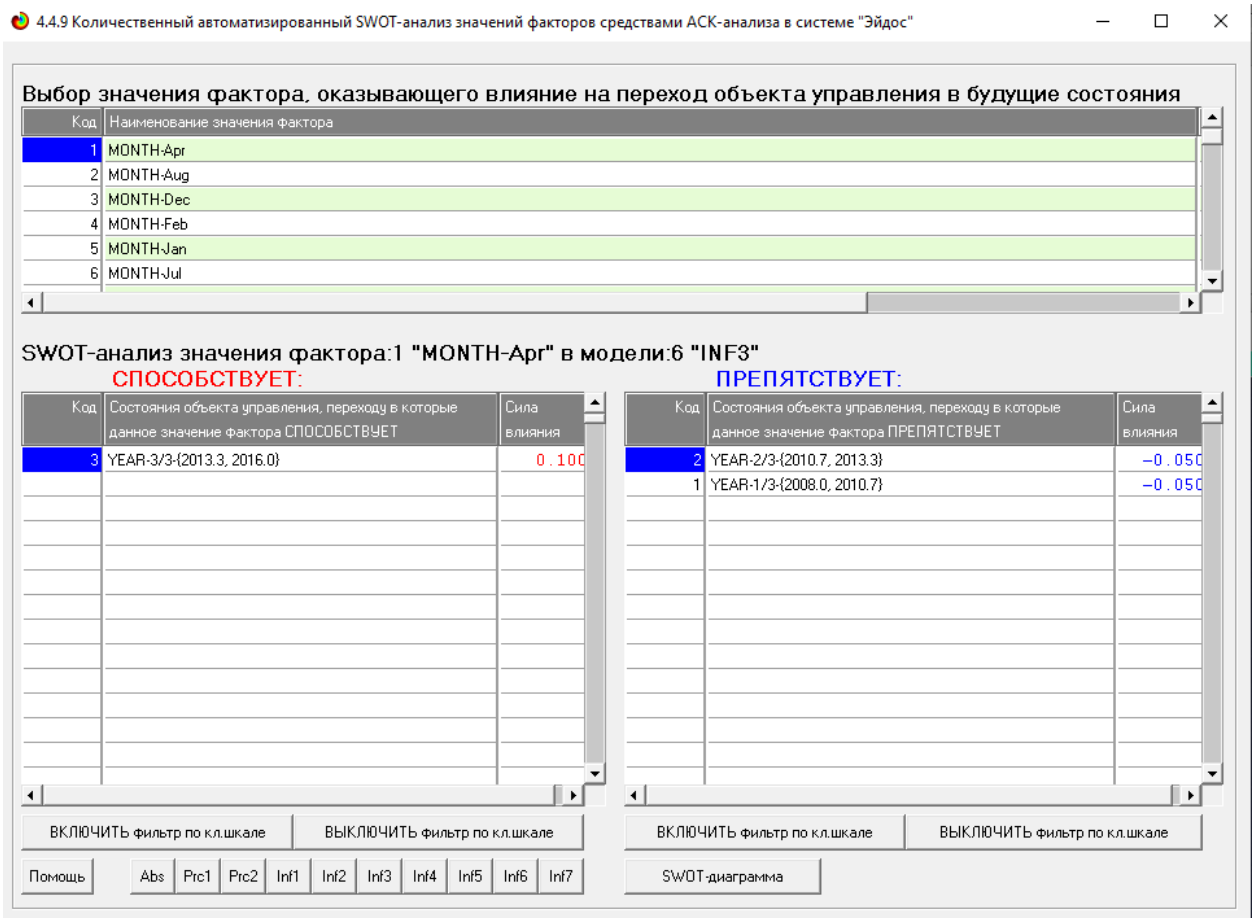


Рисунок 28 - Табличная выходная форма количественного автоматизированного SWOT- и PEST- анализа средствами системы Aidos

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

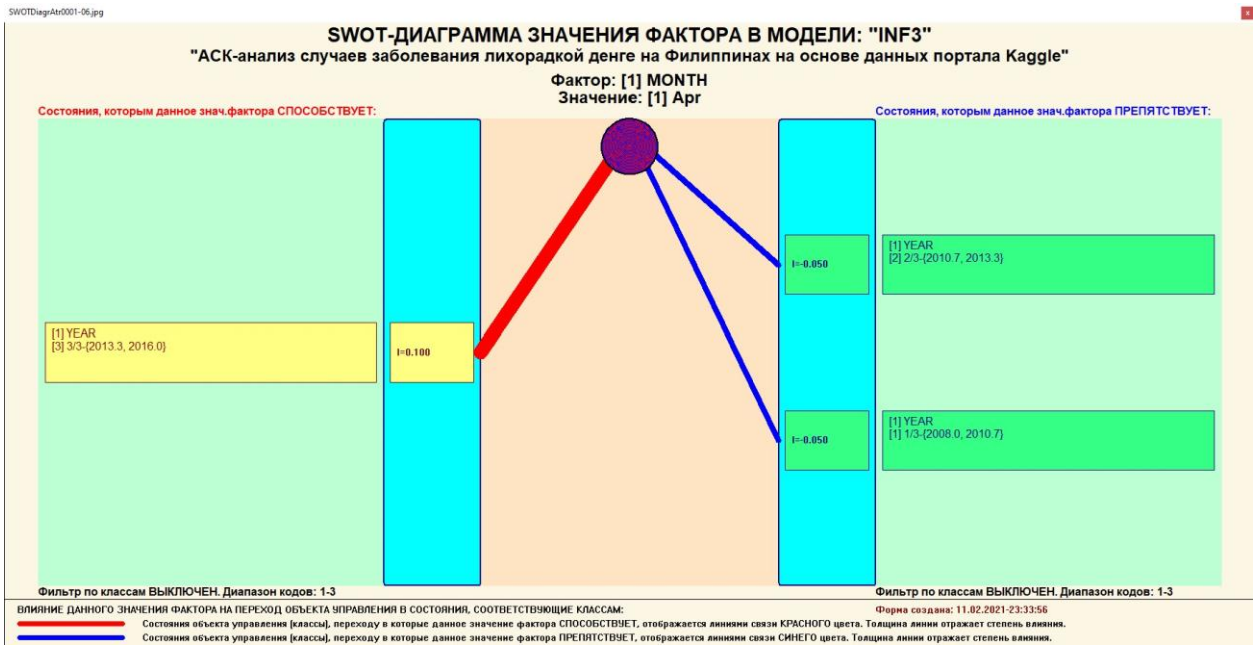


Рисунок 29 - Графическая выходная форма количественного автоматизированного SWOT- и PEST-анализа средствами системы Aidos

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 оценки заболевания денге системой искусственного интеллекта "Aidos-X++". При этом наиболее достоверной в данном приложении оказалась модель INF3, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,976, 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