

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

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

Факультет заочного обучения
Кафедра компьютерных технологий и систем

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к курсовой работе

по дисциплине: СИСТЕМЫ ПРЕДСТАВЛЕНИЯ И ПРИОБРЕТЕНИЯ
ЗНАНИЙ

на тему:

**«Автоматизированный системно-когнитивный анализ количества
лесных пожаров, зарегистрированных в Бразилии, на основе данных
портала Kaggle»**

выполнил студент группы ИТ31841 Диденко Денис Владимирович

Допущен к защите _____

Руководитель проекта _____ Луценко Евгений
Вениаминович, д.э.н., к.т.н.,
профессор
(подпись, расшифровка подписи)

Нормоконтролер Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(подпись, расшифровка подписи)

Защищена _____ Оценка _____

(дата)

Члены комиссии _____ В.И. Лойко
_____ Е.В. Луценко
_____ И. В. Николаева
(подпись, дата, расшифровка подписи)

Краснодар

2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Факультет заочного обучения
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:

Зав. кафедрой КТС _____ В. И. Лойко

ЗАДАНИЕ
на курсовую работу

Студенту: ИТ31841 группы 2 курса
Факультета заочного
Специальности: 09.04.02 Информационные системы и технологии
(шифр)

Диденко Денису Владимировичу
(Ф.И.О.)

Тема проекта: **Автоматизированный системно-когнитивный анализ количества лесных пожаров, зарегистрированных в Бразилии, на основе данных портала Kaggle**

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

Объем работы:

а) пояснительная записка к работе _____ листа формата А4

б) графическая часть _____ лист формата А4

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

Срок выполнения проекта: с “ ” _____ по “ ” _____ 2019 г.

Срок защиты: _____ “ ” _____ 2019 г.

Дата выдачи задания: _____ “ ” _____ 2019 г.

Дата сдачи проекта на кафедру: _____ “ ” _____ 2019 г.

Руководитель проекта: _____ Луценко Евгений Вениаминович,

д.э.н., к.т.н., профессор

(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ “ ” _____ 2019 г.

(подпись, дата)

Краснодар
2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т.ТРУБИЛИНА»

Факультет заочного обучения

РЕЦЕНЗИЯ
на курсовой проект (работу)

Студента (ки) Диденко Денис Владимирович

курса 2 заочной формы обучения

направления подготовки Информационные системы и технологии

направленность(профиль) Информационные системы и технологии в науке и управлении

Наименование темы «Автоматизированный системно-когнитивный анализ количества лесных пожаров, зарегистрированных в Бразилии, 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 курсового проекта (работы)

№ п/п	Показатель	Оценка соответствия заданию (по 5-и балльной шкале)
1.	Актуальность тематики работы	5
2.	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	5
3.	Уровень и корректность использования в работе различных методов исследований	4
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	4
5.	Ясность, четкость, последовательность и обоснованность изложения	5
6.	Применение современных технологий обработки информации	4
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	5
8.	Ответы на вопросы при защите	5

Достоинства работы

Актуальность темы применения современных математических моделей и программного инструментария для проведения АСК-анализа на основе базы данных

Недостатки работы

Отдельные погрешности в оформлении

Итоговая оценка при защите 5(отл.)

Рецензент

Е.В. Луценко

« »

2019 г.

РЕФЕРАТ

Курсовая работа содержит: 27 страниц, 19 рисунков, 4 таблицы, 9 литературных источников.

Ключевые слова: КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, ШКАЛЫ, КЛАССЫ.

В данной курсовой работе были проанализированы 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.....	6
1. Синтез и верификация моделей.....	7
1.1. Описание решения	7
1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel	7
1.3. Ввод выборки в систему Aidos-X.....	8
1.4. Синтез и верификация статистических и интеллектуальных моделей	12
1.5. Виды моделей системы «Aidos».....	13
1.5. Результаты верификации моделей	15
2. РЕШЕНИЕ ЗАДАЧ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	17
2.1. Решение задачи идентификации.....	17
2.2. Выбор модели и пакетное распознавание	17
2.3 Визуализация когнитивных функций	20
2.4 SWOT и PEST матрицы и диаграммы	20
2.5 Нелокальные нейронные сети и нейтроны.....	23
2.6 Кластерный и конструктивный анализ	25
Заключение	26
Список используемых источников.....	27

ВВЕДЕНИЕ

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи по определению зависимостей лесных пожаров, зарегистрированных в Бразилии.

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 «Forest Fires in Brazil» из банка исходных данных по задачам искусственного интеллекта – репозитория Kaggle.

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

1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1.Описание решения

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

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

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

С электронного ресурса **kaggle.com** возьмем набор данных «Forest Fires in Brazil» <https://www.kaggle.com/gustavomodelli/forest-fires-in-brazil>

Данный файл содержит:

- year - Год, когда происходят лесные пожары
- state - Штат бразильского государства
- month - Месяц, когда происходят лесные пожары
- number - Количество зарегистрированных лесных пожаров
- date - Дата возникновения лесных пожаров где сообщалось

В качестве разделителей используются запятые.

С помощью онлайн-конвертера <https://onlineconvertfree.com/ru/convert-format/csv-to-xls> трансформируем формат csv в формат xls.

Добавим в таблицу столбец A и назовем его Number. В итоге получается следующая таблица для работы:

Таблица 1 – Фрагмент обучающей выборки amazon.xls

	A	B	C	D	E	F
1	Number	year	state	month	number	date
2	1	1998	Acre	Janeiro	0	1998-01-01
3	2	1999	Acre	Janeiro	0	1999-01-01
4	3	2000	Acre	Janeiro	0	2000-01-01
5	4	2001	Acre	Janeiro	0	2001-01-01
6	5	2002	Acre	Janeiro	0	2002-01-01
7	6	2003	Acre	Janeiro	10	2003-01-01
8	7	2004	Acre	Janeiro	0	2004-01-01
9	8	2005	Acre	Janeiro	12	2005-01-01
10	9	2006	Acre	Janeiro	4	2006-01-01
11	10	2007	Acre	Janeiro	0	2007-01-01
12	11	2008	Acre	Janeiro	0	2008-01-01
13	12	2009	Acre	Janeiro	0	2009-01-01
14	13	2010	Acre	Janeiro	1	2010-01-01
15	14	2011	Acre	Janeiro	0	2011-01-01
16	15	2012	Acre	Janeiro	0	2012-01-01
17	16	2013	Acre	Janeiro	0	2013-01-01
18	17	2014	Acre	Janeiro	0	2014-01-01
19	18	2015	Acre	Janeiro	1	2015-01-01
20	19	2016	Acre	Janeiro	12	2016-01-01
21	20	2017	Acre	Janeiro	0	2017-01-01
22	21	1998	Acre	Fevereiro	0	1998-01-01
23	22	1999	Acre	Fevereiro	0	1999-01-01
24	23	2000	Acre	Fevereiro	0	2000-01-01
25	24	2001	Acre	Fevereiro	0	2001-01-01
26	25	2002	Acre	Fevereiro	1	2002-01-01
27	26	2003	Acre	Fevereiro	0	2003-01-01
28	27	2004	Acre	Fevereiro	3	2004-01-01
29	28	2005	Acre	Fevereiro	0	2005-01-01
30	29	2006	Acre	Fevereiro	0	2006-01-01
31	30	2007	Acre	Fevereiro	5	2007-01-01

Классификационные шкалы следующие: Год, Бразильский штат. Месяц, Количество пожаров, Дата – описательные.

1.3. Ввод выборки в систему Aidos-X

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

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

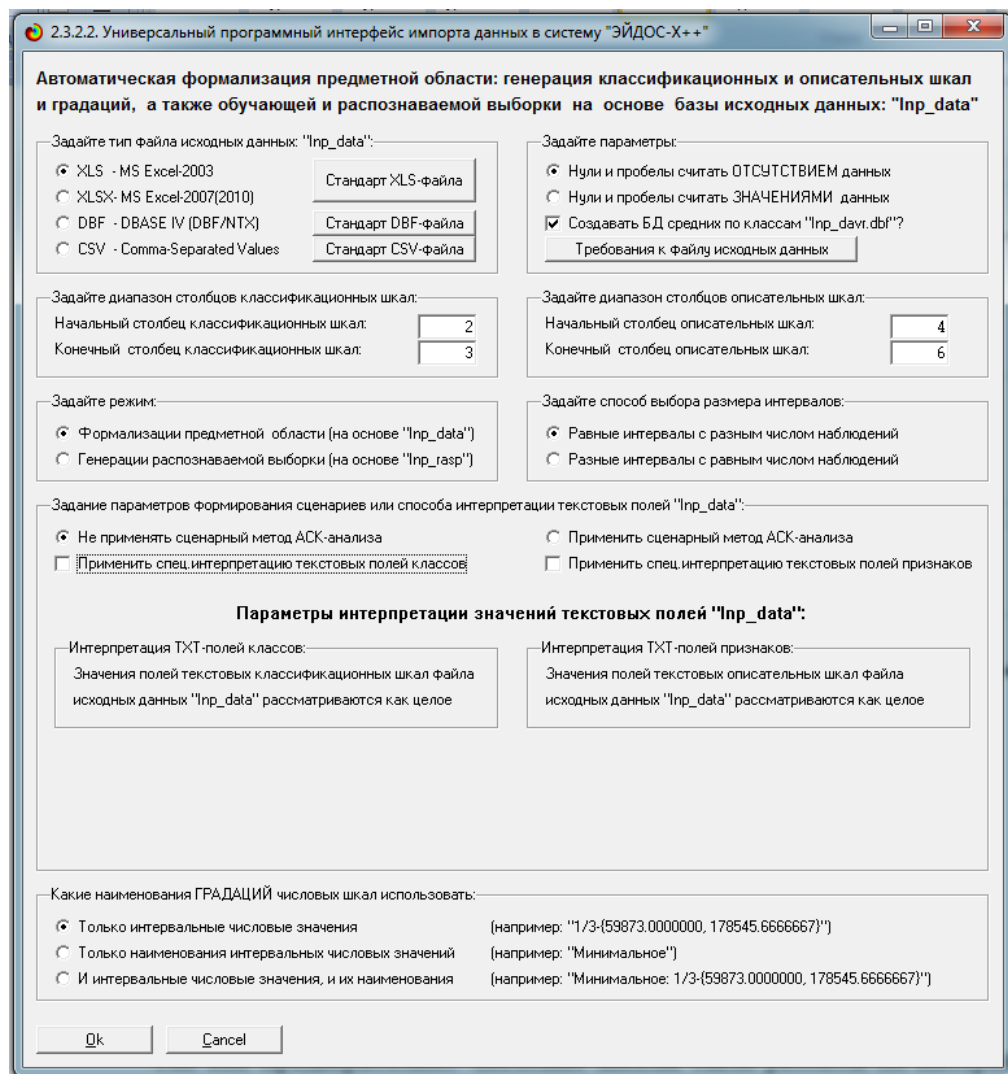


Рисунок 1 – Ввод обучающей выборки

Стоит отметить следующие настройки:

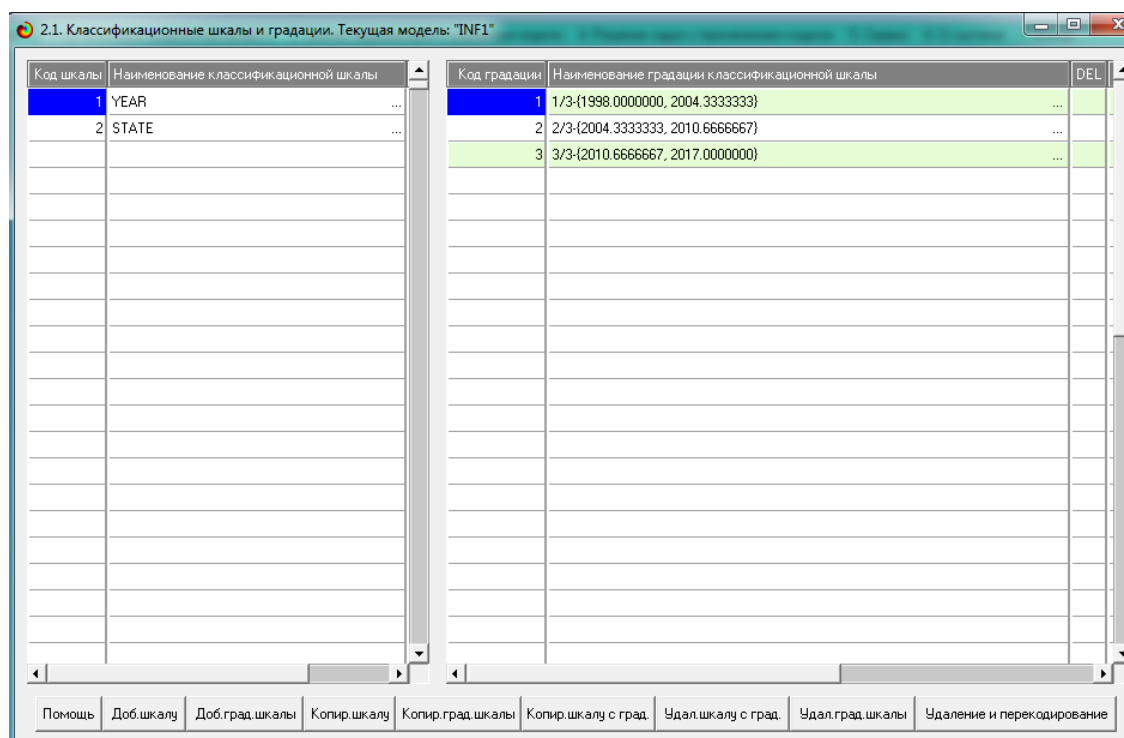
- Тип файла – xls.
- Классификационные шкалы – 2-3.
- Описательные шкалы – 4-6.
- Не применяется спец. интерпретация текстовых полей классов.

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

После импорта автоматически формируются классификационные и описательные шкалы, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Этим самым

полностью автоматизировано выполняется этап «Формализация предметной области».

Классификационные шкалы можно просмотреть в режиме 2.1 (рисунок 2).



2.1. Классификационные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование классификационной шкалы
1	YEAR
2	STATE

Код градации	Наименование градации классификационной шкалы	DEL
1	1/3-{1998.0000000, 2004.3333333}	...
2	2/3-{2004.3333333, 2010.6666667}	...
3	3/3-{2010.6666667, 2017.0000000}	...

Помощь Доб. шкалу Доб. град. шкалы Копир. шкалу Копир. град. шкалы Копир. шкалу с град. Удал. шкалу с град. Удал. град. шкалы Удаление и перекодирование

Рисунок 2 – Классификационные шкалы

Описательные шкалы – в режиме 2.2 (рисунок 3).

Для ручного ввода-корректировки выборки существует режим 2.3.1, его можно увидеть на рисунке 4. Установка значений описательных и классификационных шкал объектов осуществляется по их номерам.

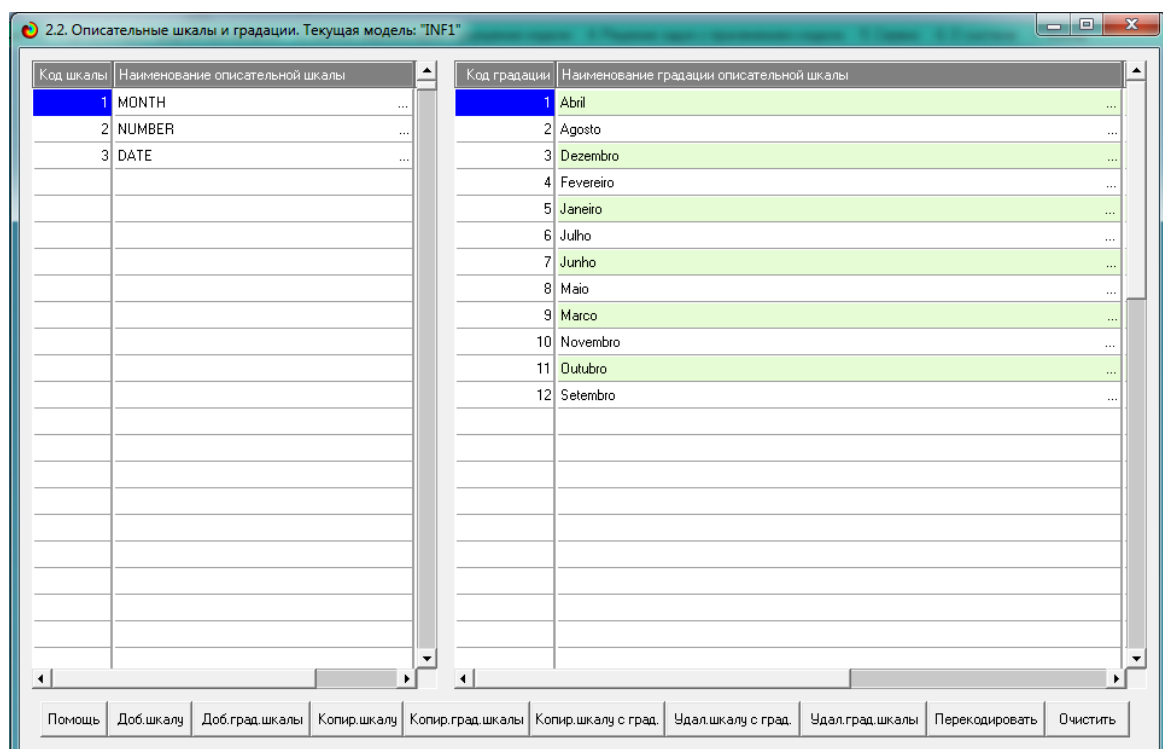


Рисунок 3 – Описательные шкалы

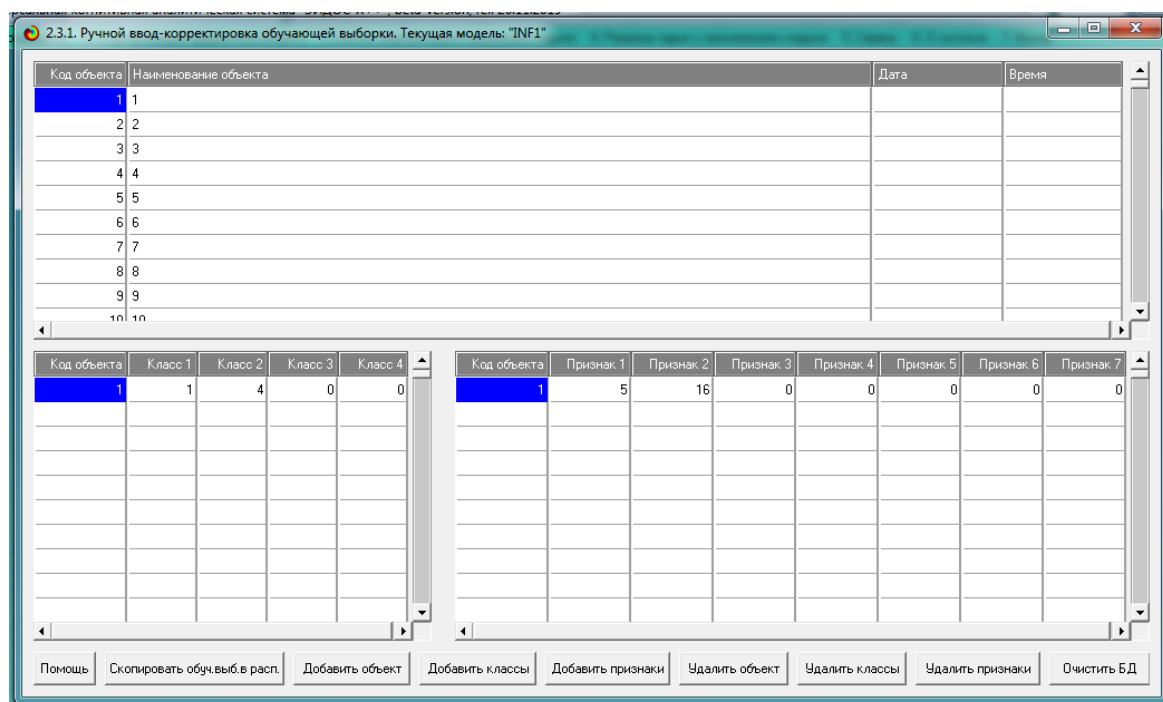


Рисунок 4 – Ручной ввод-корректировка выборки

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

1.4. Синтез и верификация статистических и интеллектуальных моделей

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

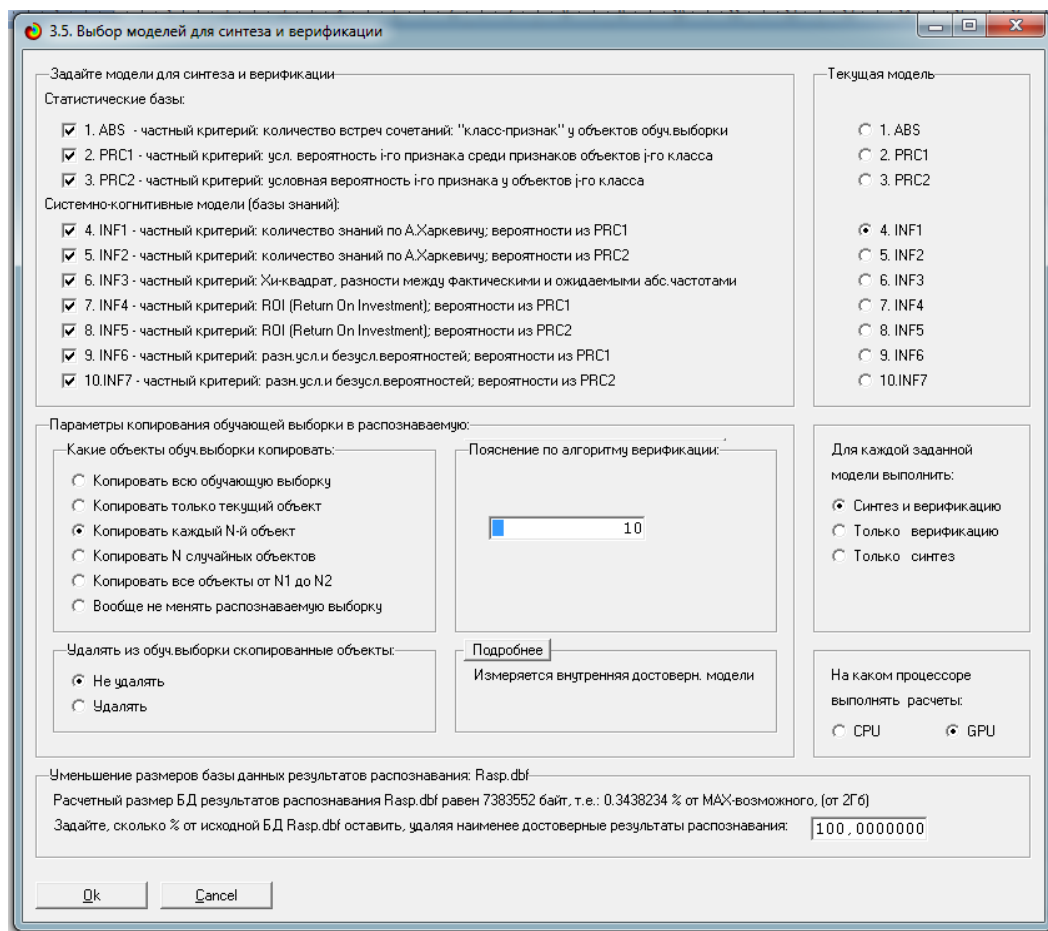


Рисунок 5 – Настройка синтеза и верификации моделей

Так как в выборке около 6450 значений, для распознавания будем использовать каждый 10й объект. Из-за сортировки таблицы по названию Бразильского штата это будет надежнее, чем использование рандомных записей.

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

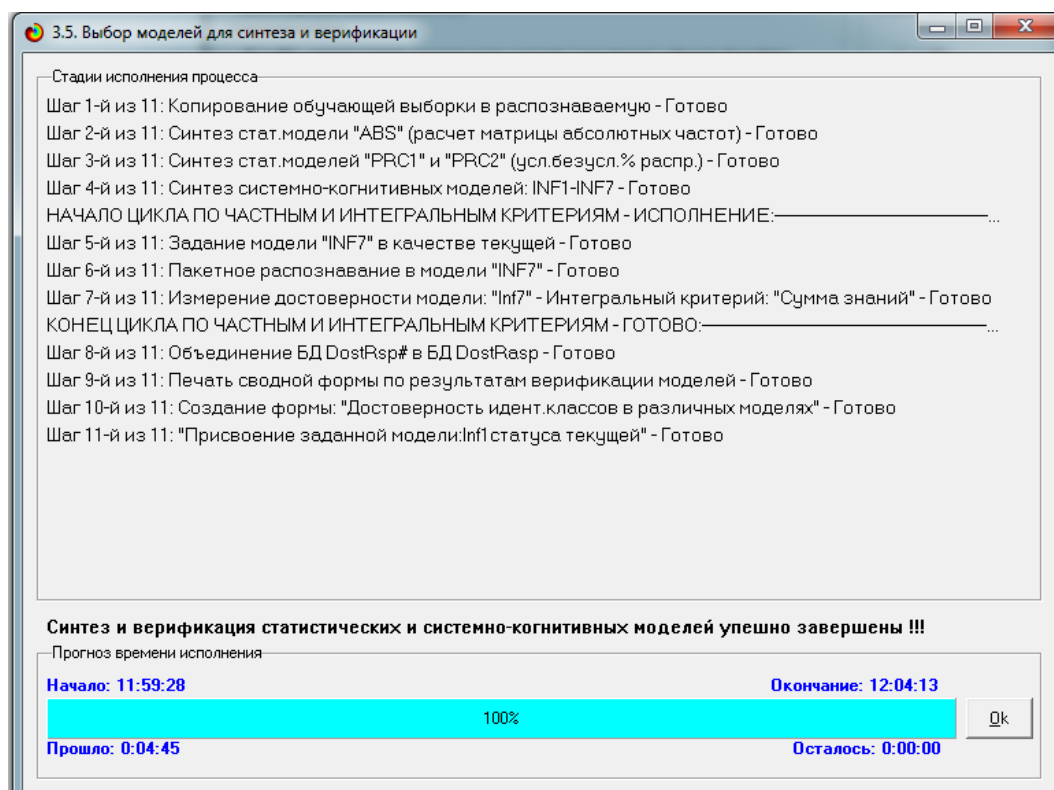


Рисунок 6 – Завершение процесса синтеза и верификации моделей

После этого можно перейти непосредственно к выбору наиболее достоверной модели.

1.5. Виды моделей системы «Aidos»

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

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

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

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

Код признака	Наименование описательной шкалы и градации	1. YEAR 1/3 (1998.0, 2004.3)	2. YEAR 2/3 (2004.3, 2010.7)	3. YEAR 3/3 (2010.7, 2017.0)	4. STATE ACRE	5. STATE ALAGOAS	6. STATE AMAPA	7. STATE AMAZONAS	8. STATE BAHIA	9. STATE CEARA	10. STATE DISTRITO FEDERAL	11. STATE ESPIRITO SANTO
1	MONTH:Abil	189	162	189	20	20	20	20	20	20	20	20
2	MONTH:Agosto	189	162	189	20	20	20	20	20	20	20	20
3	MONTH:Dezembro	189	162	162	19	19	19	19	19	19	19	19
4	MONTH:Fevereiro	189	162	189	20	20	20	20	20	20	20	20
5	MONTH:Janeiro	189	162	190	20	21	20	20	20	20	20	20
6	MONTH:Julho	189	162	189	20	20	20	20	20	20	20	20
7	MONTH:Junho	189	162	189	20	20	20	20	20	20	20	20
8	MONTH:Maior	189	162	189	20	20	20	20	20	20	20	20
9	MONTH:Marco	189	162	189	20	20	20	20	20	20	20	20
10	MONTH:Novembro	189	162	189	20	20	20	20	20	20	20	20
11	MONTH:Outubro	189	162	189	20	20	20	20	20	20	20	20
12	MONTH:Setembro	189	162	189	20	20	20	20	20	20	20	20
13	NUMBER-1/3-(1.0000000, 333.3333333)	1692	1564	1093	148	195	137	200	195	188	152	226
14	NUMBER-2/3-(333.3333333, 665.6666667)	146	128	178	15		20	11	27	23		
15	NUMBER-3/3-(665.6666667, 998.0000000)	84	66	86	8		9	17	12	11		
16	DATE-1998-01-01	324			12	12	12	12	12	12	12	12
17	DATE-1999-01-01	324			12	12	12	12	12	12	12	12
18	DATE-2000-01-01	324			12	12	12	12	12	12	12	12
19	DATE-2001-01-01	324			12	12	12	12	12	12	12	12
20	DATE-2002-01-01	324			12	12	12	12	12	12	12	12
21	DATE-2003-01-01	324			12	12	12	12	12	12	12	12
22	DATE-2004-01-01	324			12	12	12	12	12	12	12	12
23	DATE-2005-01-01		324		12	12	12	12	12	12	12	12

Таблица 3 – Матрица информативностей в битах

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

Код признака	Наименование описательной шкалы и градации	1. YEAR 1/3 (1998.0, 2004.3)	2. YEAR 2/3 (2004.3, 2010.7)	3. YEAR 3/3 (2010.7, 2017.0)	4. STATE ACRE	5. STATE ALAGOAS	6. STATE AMAPA	7. STATE AMAZONAS	8. STATE BAHIA	9. STATE CEARA	10. STATE DISTRITO FEDERAL	11. STATE ESPIRITO SANTO
1	MONTH:Abil	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
2	MONTH:Agosto	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
3	MONTH:Dezembro	0.030	0.021	-0.051	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
4	MONTH:Fevereiro	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
5	MONTH:Janeiro	0.006	-0.003	-0.004	0.029	0.034	0.033	-0.008	-0.012	-0.005	0.043	
6	MONTH:Julho	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
7	MONTH:Junho	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
8	MONTH:Maior	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
9	MONTH:Marco	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
10	MONTH:Novembro	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
11	MONTH:Outubro	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
12	MONTH:Setembro	0.007	-0.002	-0.005	0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
13	NUMBER-1/3-(1.0000000, 333.3333333)	-0.021	0.004	0.017	-0.083	0.022	-0.114	0.014	-0.001	-0.010	-0.058	
14	NUMBER-2/3-(333.3333333, 665.6666667)	-0.029	-0.028	0.047	-0.019		0.113	-0.195	0.202	0.138		
15	NUMBER-3/3-(665.6666667, 998.0000000)	0.015	-0.033	0.013	-0.009		0.047	0.289	0.130	0.099		
16	DATE-1998-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
17	DATE-1999-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
18	DATE-2000-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
19	DATE-2001-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
20	DATE-2002-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
21	DATE-2003-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
22	DATE-2004-01-01	0.476			0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	
23	DATE-2005-01-01		0.536		0.030	0.013	0.034	-0.007	-0.011	-0.004	0.043	

Таблица 4 – Матрица знаний

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

Код признака	Наименование описательной шкалы и градации	1. YEAR 1/3 (1998.0, 2004.3)	2. YEAR 2/3 (2004.3, 2010.7)	3. YEAR 3/3 (2010.7, 2017.0)	4. STATE ACRE	5. STATE ALAGOAS	6. STATE AMAPA	7. STATE AMAZONAS	8. STATE BAHIA	9. STATE CEARA	10. STATE DISTRITO FEDERAL	11. STATE ESPIRITO SANTO
1	MONTH:Abil	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
2	MONTH:Agosto	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
3	MONTH:Dezembro	12.262	7.484	-19.746	1.239	0.527	1.375	-0.321	-0.486	-0.157	1.759	
4	MONTH:Fevereiro	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
5	MONTH:Janeiro	2.615	-0.949	-1.666	1.269	1.519	1.413	-0.376	-0.549	-0.203	1.818	
6	MONTH:Julho	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
7	MONTH:Junho	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
8	MONTH:Maior	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
9	MONTH:Marco	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
10	MONTH:Novembro	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
11	MONTH:Outubro	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
12	MONTH:Setembro	2.960	-0.648	-2.312	1.304	0.555	1.448	-0.338	-0.511	-0.165	1.851	
13	NUMBER-1/3-(1.0000000, 333.3333333)	-81.926	13.120	68.806	-30.272	9.587	-39.898	6.071	-0.577	4.281	-21.053	
14	NUMBER-2/3-(333.3333333, 665.6666667)	-9.722	-8.143	17.865	-0.649	-16.276	4.471	-6.024	9.831	6.121	-15.191	
15	NUMBER-3/3-(665.6666667, 998.0000000)	2.694	-5.083	2.390	-0.171	-8.498	0.892	8.111	3.036	2.187	-7.932	
16	DATE-1998-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
17	DATE-1999-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
18	DATE-2000-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
19	DATE-2001-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
20	DATE-2002-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
21	DATE-2003-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
22	DATE-2004-01-01	212.376	-97.589	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	
23	DATE-2005-01-01	-111.624	226.411	-114.787	0.782	0.333	0.869	-0.203	-0.307	-0.099	1.111	

1.5. Результаты верификации моделей

Результаты верификации моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями, приведены на рисунке 6.

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Ф-мера Ван Рубергена	Сумма знан... уровней сход... истинно-поло... решений (STP)	Сумма знан... уровней сход... истинно-отриц... решений (STN)	Сумма знан... уровней сход... ложно-полож... решений
1. ABS - частный критерий: количество встреч сочетаний: "клас..."	Корреляция абс частот с абс...	1290	1176	2957	12523	114	0.086	0.912	0.157	1008.411	224.310	11508
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн частот с о...	1290	1176	2957	12523	114	0.086	0.912	0.157	1008.411	224.310	11508
3. PRC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн частот с о...	1290	1176	2957	12523	114	0.086	0.912	0.157	1008.411	224.310	11508
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в...	Семантический резонанс зна...	1290	1029	8832	6648	261	0.134	0.798	0.230	171.724	1860.547	1356
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в...	Семантический резонанс зна...	1290	1017	8068	7412	273	0.121	0.788	0.209	562.159	471.588	475
6. INF3 - частный критерий: Хинкварт: разности между фактик...	Семантический резонанс зна...	1290	1109	5333	10147	181	0.099	0.860	0.177	572.152	613.155	544
7. INF4 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	1290	1008	9329	6151	282	0.141	0.781	0.239	173.704	1809.255	1374
8. INF5 - частный критерий: ROI (Return On Investment), веротно...	Семантический резонанс зна...	1290	1017	8066	7414	273	0.121	0.788	0.209	568.306	226.358	320
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей, вер...	Семантический резонанс зна...	1290	1040	8456	7024	250	0.129	0.806	0.204	485.535	2143.682	1755
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей, ве...	Семантический резонанс зна...	1290	912	7457	8023	378	0.102	0.707	0.178	321.078	2497.720	4105
10. INF8 - частный критерий: разн.усл.и безусл.вероятностей, ве...	Семантический резонанс зна...	1290	1109	5301	10179	181	0.098	0.860	0.176	384.159	2007.018	1697

Рисунок 6 – Результаты верификации моделей

Наиболее достоверной в данном приложении оказалась модель INF4 при интегральном критерии «Сумма знаний», если учитывать L-1 меру проф. Е.В. Луценко. Точность модели составляет 0,776, то есть уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 49,4%.

Также статистические модели обычно дают более низкую средневзвешенную достоверность идентификации и неидентификации, чем модели знаний, и практически никогда – более высокую. Соответственно, это оправдывает применение интеллектуальных технологий и моделей знаний. Посмотрим, каковы частные распределения уровней сходства и различия для верно и ошибочно неидентифицированных и идентифицированных ситуаций в наиболее достоверной модели INF4 (рисунок 7).

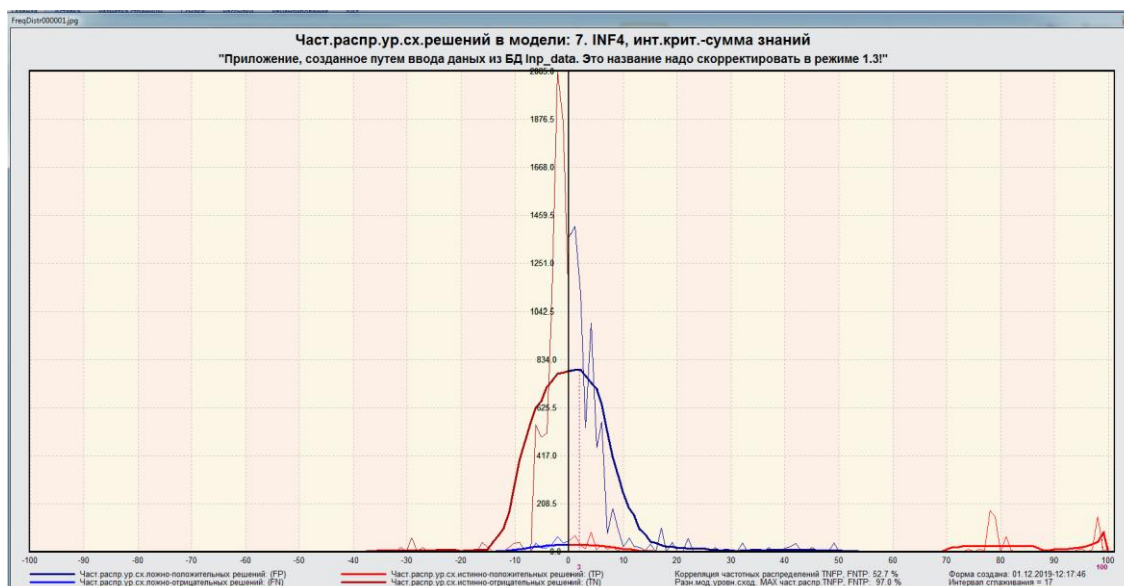


Рисунок 7 – Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF4

Как можно видеть, при модели INF4 непринадлежность объекта к классу лучше определяется, чем принадлежность. При этом существует вероятность ложного определения непринадлежности или принадлежности.

Выберем модель INF4 в качестве основной и перейдем к решению задач идентификации.

2.РЕШЕНИЕ ЗАДАЧ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

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

Во-первых, зависимость того – в каком штате случается наибольшее количество пожаров. Ведь зная наиболее пожароопасные штаты, можно оптимально распределять технические и человеческие ресурсы для борьбы со стихийным бедствием.

Во-вторых, при этом следует анализировать также и количество пожаров в штатах в зависимости от времени года.

2.2. Выбор модели и пакетное распознавание

Зададим модель INF4 в качестве текущей в режиме 5.6 (рисунок 8), после чего проведем пакетное распознавание в режиме 4.1.2 (рисунок 9).

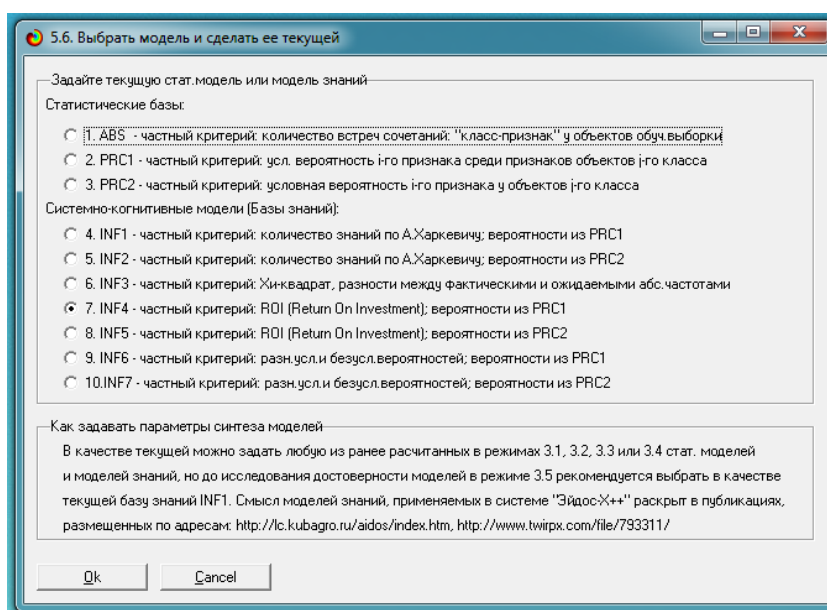


Рисунок 8 – Выбор текущей модели

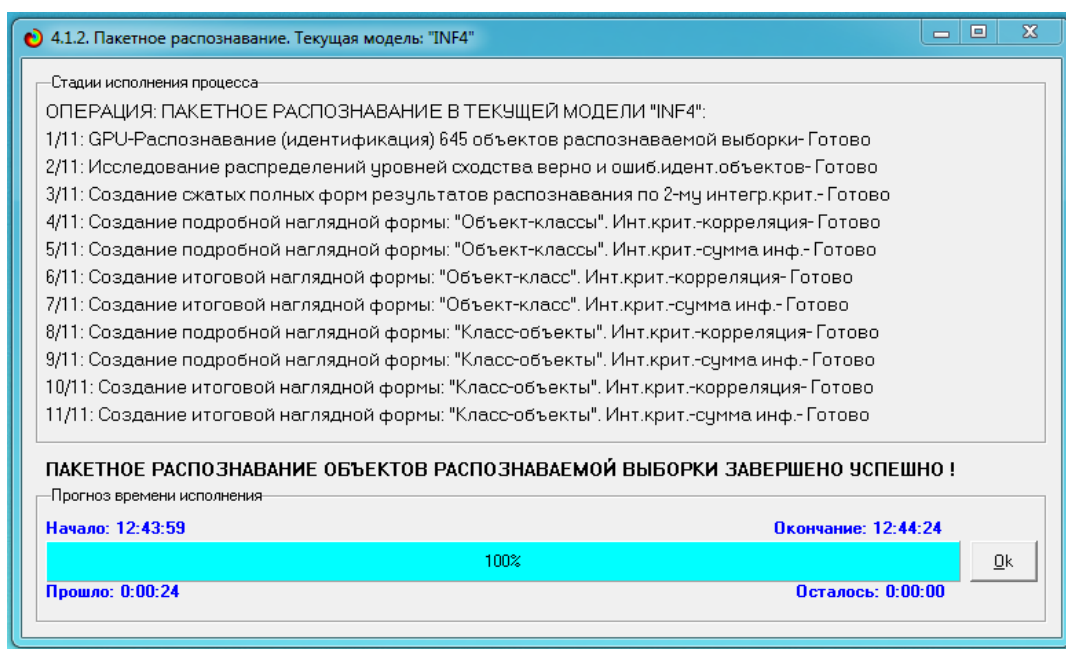


Рисунок 9 – Пакетное распознавание в новой модели

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

Режим 4.1.3 системы «Aidos-X» обеспечивает отображение результатов идентификации и прогнозирования в различных формах:

1. Подробно наглядно: "Объект – классы".
2. Подробно наглядно: "Класс – объекты".
3. Итоги наглядно: "Объект – классы".
4. Итоги наглядно: "Класс – объекты".
5. Подробно сжато: "Объект – классы".
6. Обобщенная форма по достоверности моделей при разных интегральных критериях.
7. Обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям.
8. Статистический анализ результатов идентификации по классам, моделям и интегральным критериям.
9. Распознавание уровня сходства при разных моделях и интегральных критериях.

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

Ниже кратко рассмотрим некоторые из них.

На рисунках 10 и 11 приведены примеры прогнозов высокой и низкой достоверности классов, а также итоги распознавания классов объектов в модели INF4:

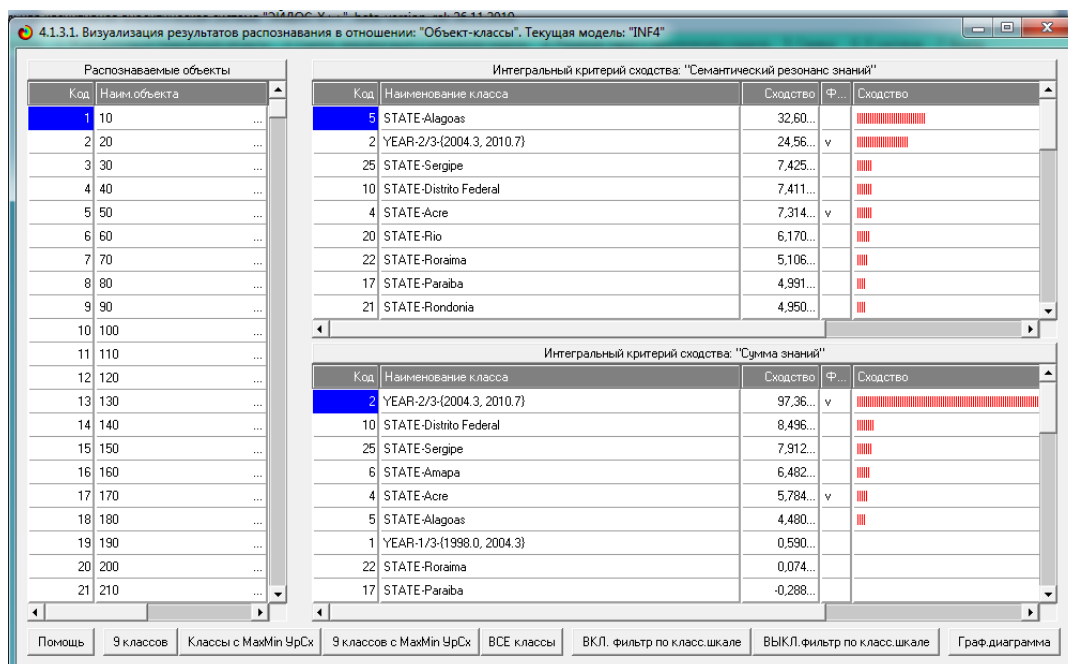


Рисунок 10 – Прогнозы достоверности классов объектов

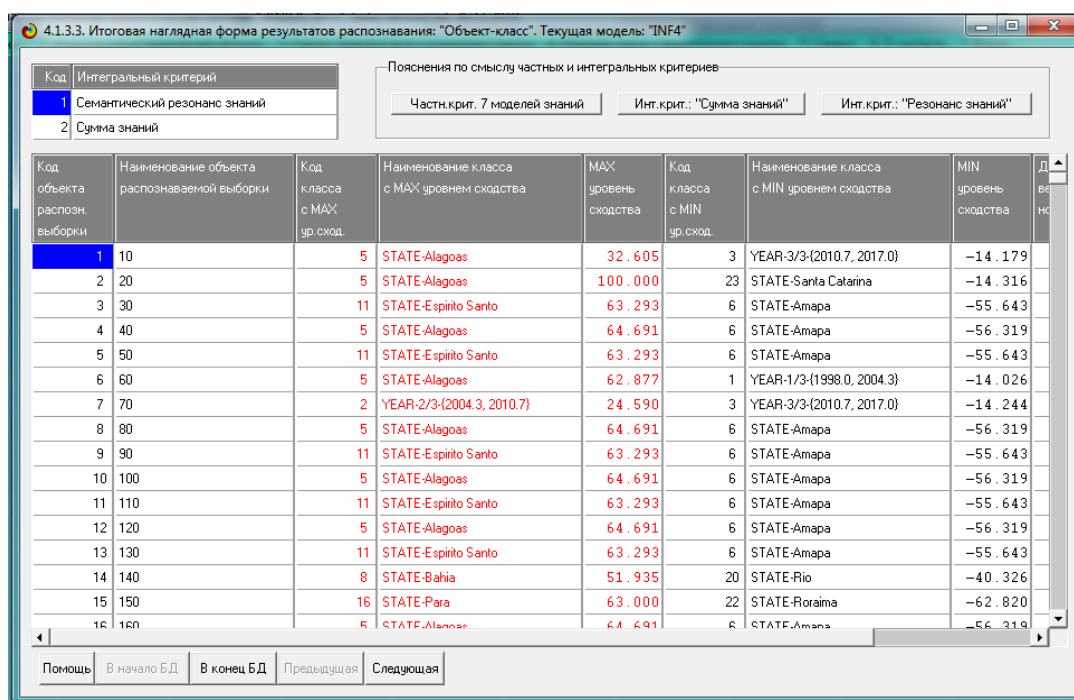


Рисунок 11 – Итоги распознавания классов объектов

2.3 Визуализация когнитивных функций

Применительно к задачам, рассматриваемым в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния.

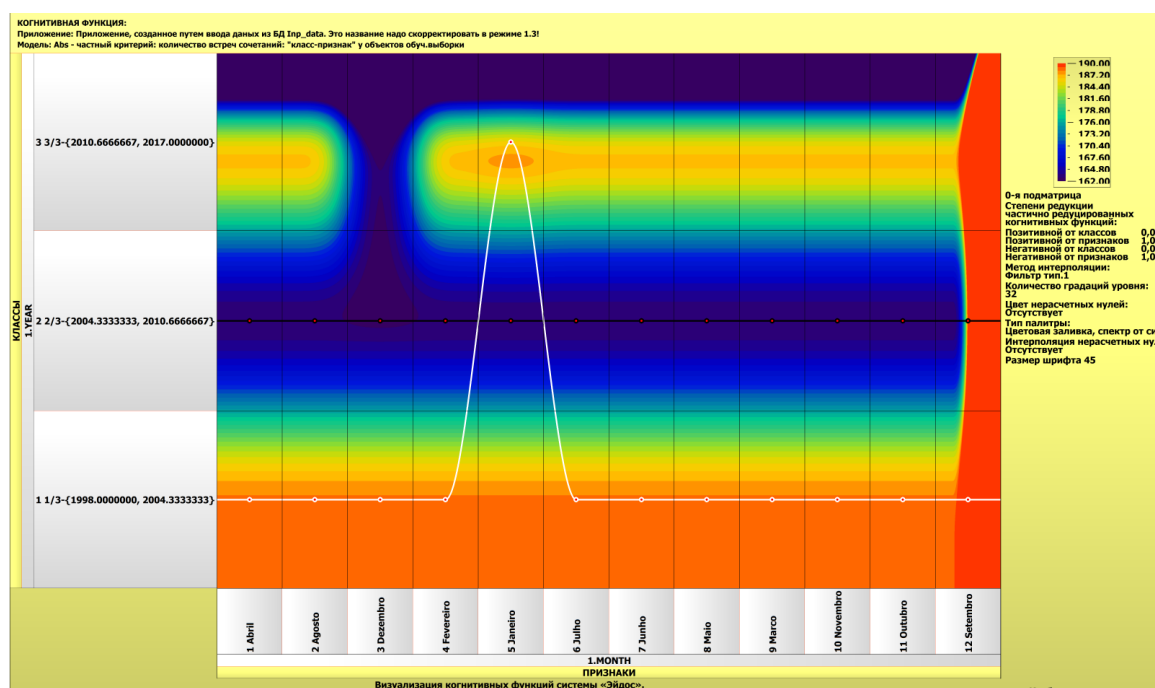


Рисунок 12 – Визуализация зависимости количества пожаров от времени года

Рассмотрим визуализацию когнитивных функций в рамках определенных задач.

Зависимость количества пожаров в Бразильских штатах от времени года показывает нам насколько количество пожаров в январе выше, чем во все остальные месяцы.

2.4 SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он

подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных.

В версии системы под MS Windows: «Aidos-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 13).

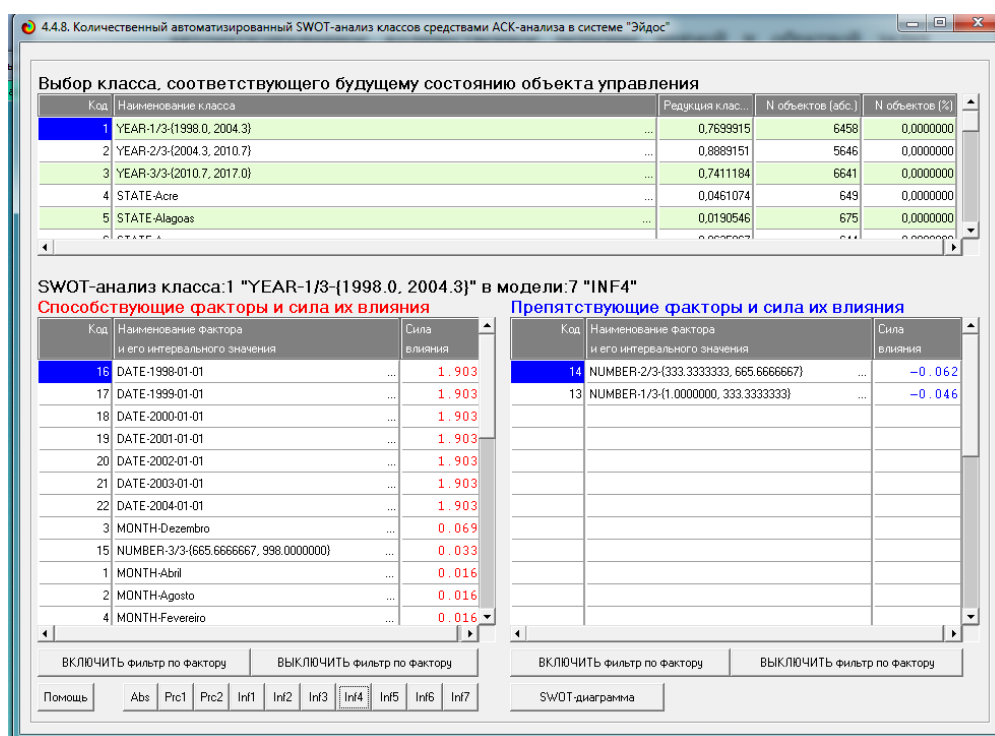


Рисунок 13 – Пример SWOT-Матрицы в модели INF4

На рисунке 14 приведен пример инвертированной SWOT-матрицы.

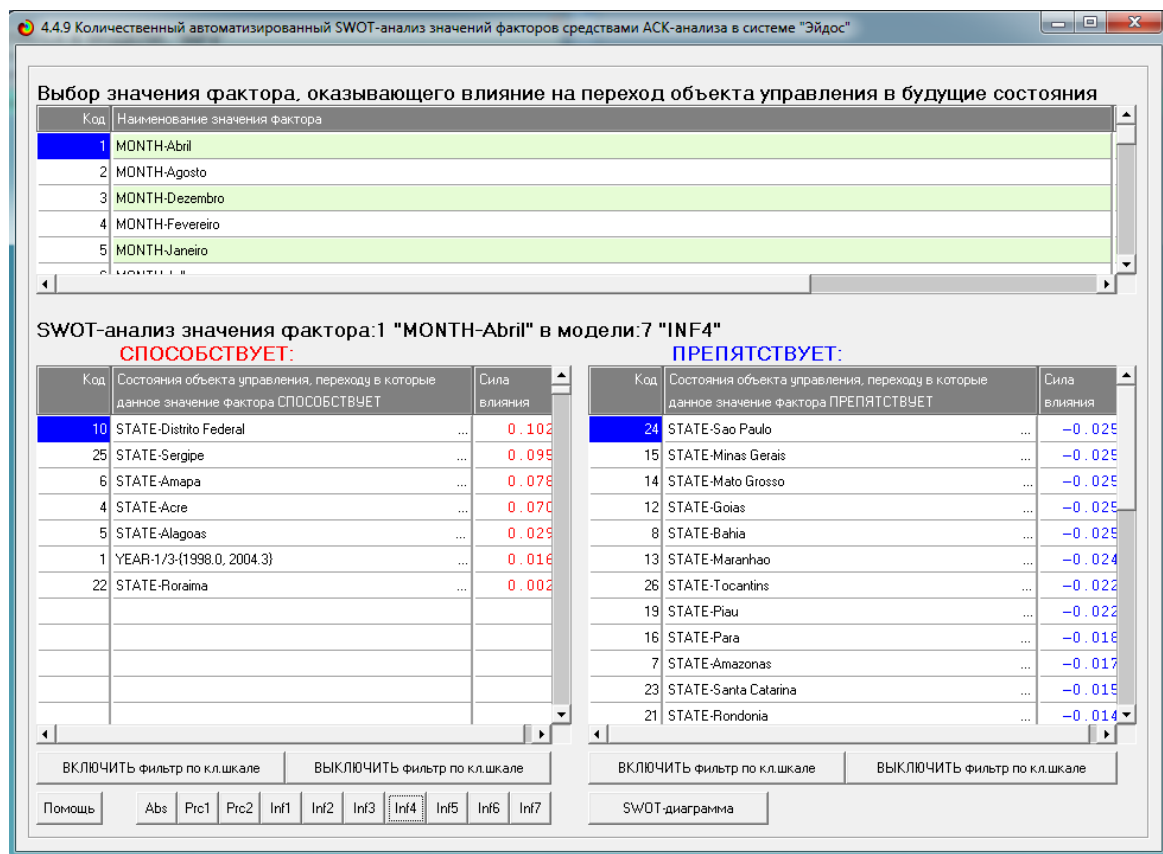


Рисунок 14 – Пример инвертированной SWOT-Матрицы в модели INF4

А теперь рассмотрим месяцы, в которые чаще всего случались пожары (рисунок 15):

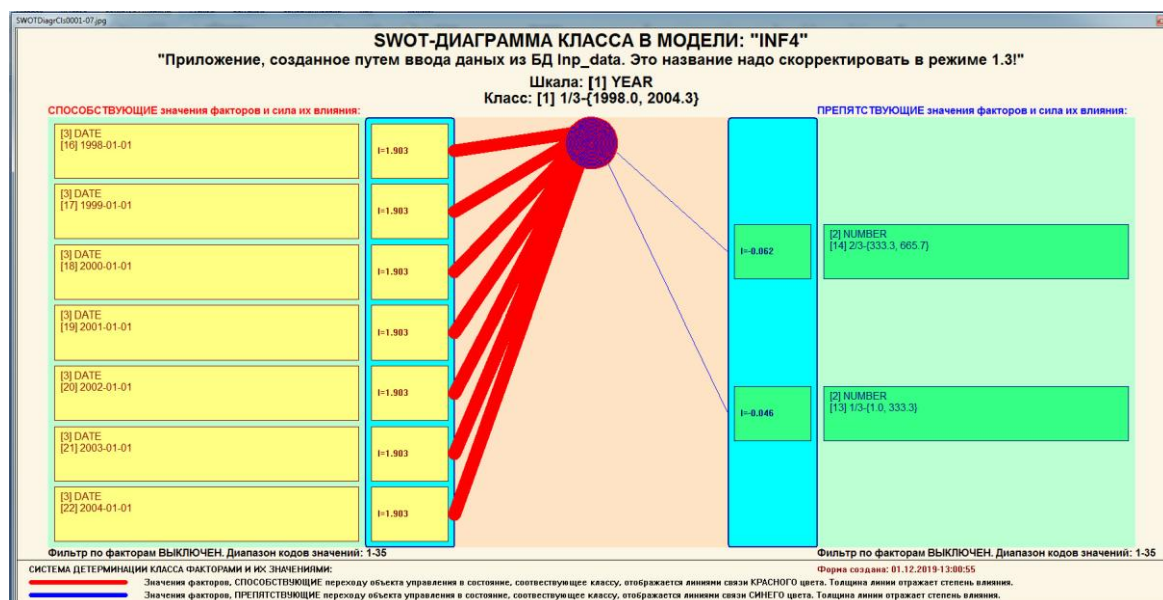


Рисунок 15 –SWOT-матрица

Как можно видеть, большая часть пожаров приходилась на первый месяц года.

А теперь узнаем, в каких штатах произошло наибольшее количество пожаров (рисунок 20).

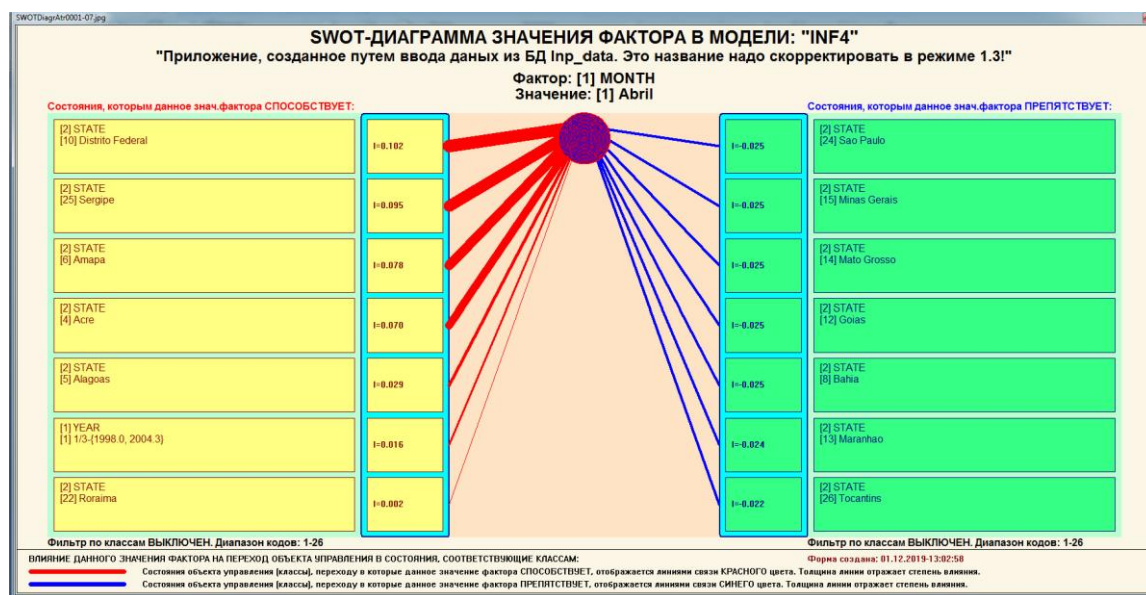


Рисунок 16 – SWOT-диаграмма разницы счета

2.5 Нелокальные нейронные сети и нейтроны

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

Рассмотрим пару примеров, возвращаясь к нашим задачам.

На рисунке 17 изображен нелокальный нейрон для модели INF4.

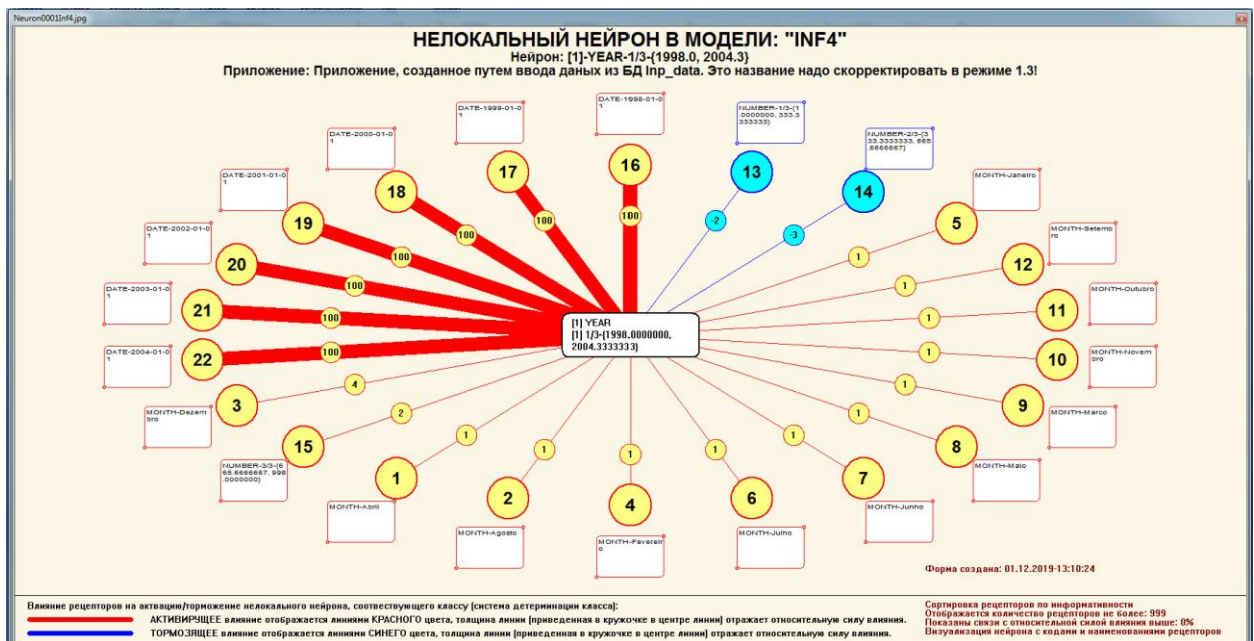


Рисунок 17 – Нейрон для модели INF4

На его основании можно определить, что количество пожаров в Бразильских штатах возрастает с сентября месяца и пик пожаров приходится на январь.

На основании нелокальных нейронов строится нейронная сеть, пример которой представлен на рисунке 18.

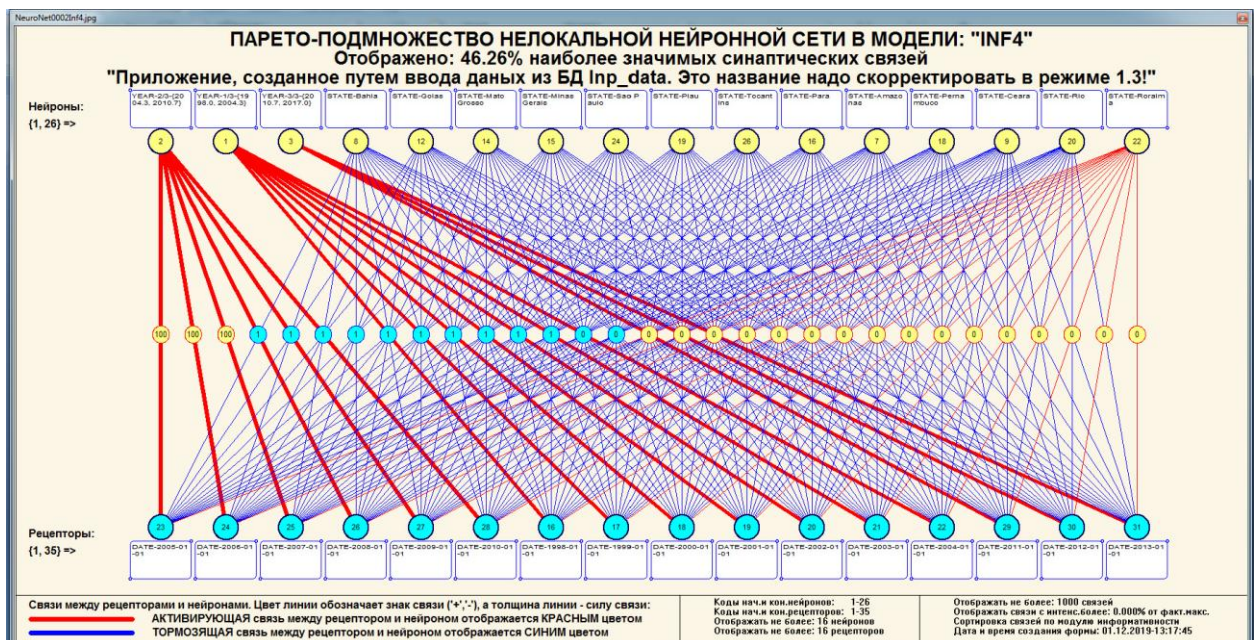


Рисунок 18 – Пример нелокальной нейронной сети в модели INF4

По ней можно, в основном, определить зависимость выбора штата от даты и количества пожаров.

2.6 Кластерный и конструктивный анализ

В режиме 4.2.2.2 после расчета матриц сходства, кластеров и конструкторов (4.2.2.1) строим 2D сеть классов в выбранной модели знаний (рисунок 24).

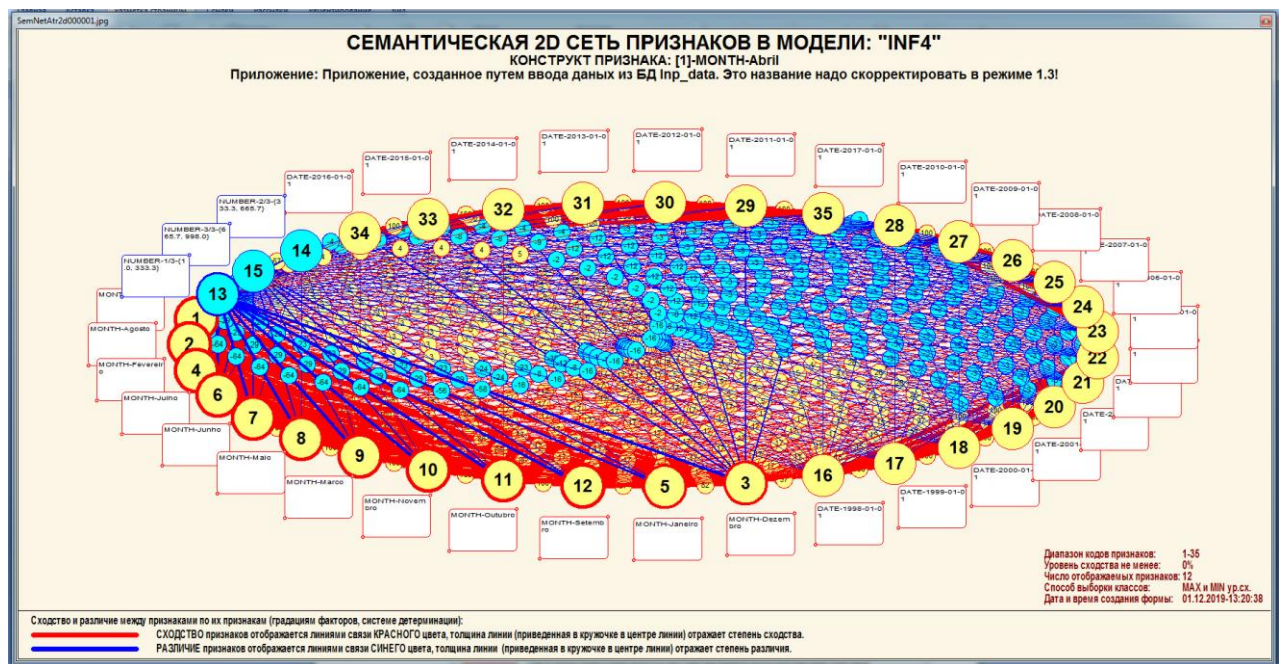


Рисунок 19 – Семантическая 2D сеть классов

ЗАКЛЮЧЕНИЕ

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

В данной курсовой работе был показан пример использования общедоступной базы данных «Количество лесных пожаров, зарегистрированных в Бразилии» для построения модели зависимости количества пожаров от времени года в Бразильских штатах. При этом наиболее достоверной в данном приложении оказалась модель INF4, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,776, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 49,4%.

При рассмотрении модели в перспективе ретроанализа можно убедиться в достоверности моделей: анализ количества произошедших пожаров подтверждает выводы, получаемые в процессе выполнения задач курсовой работы.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Луценко Е.В. Методика использования репозитория UCI для оценки качества математических моделей систем искусственного интеллекта / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: http://lc.kubagro.ru/My_training_schedule.doc КубГАУ, 2003. – №02(002). С. 120 – 145. – IDA [articleID]: 0020302012. – Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 у.п.л. Луценко Е.В. АСК-анализ, моделирование и идентификация живых существ на основе их фенотипических признаков / Е.В. Луценко, Ю.Н. Пенкина // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: Куб- ГАУ, 2014. – №06(100). С. 1346 – 1395. – IDA [article ID]: 1001406090. – Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/90.pdf>, 3,125 у.п.л.

2. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: Куб- ГАУ, 2013. – №04(088). С. 340 – 359. – IDA [article ID]: 0881304022. – Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.

3. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических,

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

5. [Электронный ресурс]. Статья "bazadannixsotrudnikov": <http://allexcel.ru/gotovye-tablitsy-excel-besplatno>, свободный. - Загл. с экрана. Яз.анг.

6. Официальный сайт WorldofWarcraft [Электронный ресурс]. Режим доступа: <http://eu.battle.net/wow/ru>, свободный. - Загл. с экрана. Яз.рус.

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

8. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос- Х++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №09(083). С. 328 – 356. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 у.п.л.

9. Луценко Е.В., Боровко А.Ю. ПРОГНОЗИРОВАНИЕ КОЛИЧЕСТВА И КЛАССОВ СОЛНЕЧНЫХ ВСПЫШЕК НА ОСНОВЕ ИХ ПРЕДЫСТОРИИ ПО ДАННЫМ РЕПОЗИТОРИЯ UCI С ПРИМЕНЕНИЕМ АСК- АНАЛИЗА И ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «ЭЙДОС» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2014/10/pdf/99.pdf>