

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И Т. ТРУБИЛИНА»**

Факультет заочного обучения

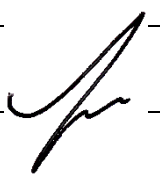
Кафедра компьютерных систем и технологий

КУРСОВАЯ РАБОТА

по дисциплине: Интеллектуальные системы и технологии
на тему: АСК- анализ рейтинга ресторанного бизнеса на основе данных с
портала Kaggle

Выполнил студент группы: ИТз1941 Костров Даниил Антонович

Допущен к защите: _____

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

Защищен 07.04.2021
(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 заочного обучения

РЕЦЕНЗИЯ

на курсовую работу

Студент Костров Даниил Антонович

курса 2 заочной формы обучения группы ИТ31941

Направление подготовки 09.04.02 «Информационные системы и технологии»

Наименование темы АСК- анализ рейтинга ресторанного бизнеса на основе данных с 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 _____отлично_____

Рецензент _____(Е.В. Луценко)

«07» 04 2021 г.



РЕФЕРАТ

Курсовая работа содержит: 37 страниц, 25 рисунков, 6 литературных источников.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, НЕЛОКАЛЬНЫЕ НЕЙТРОННЫЕ СЕТИ, НЕЙРОНЫ.

Целью работы является проведение АСК-анализ рейтинга ресторанного бизнеса на основе данных с 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 MS Excel .	6
1.3. Ввод выборки в систему Aidos-X.....	7
1.4. Синтез и верификация статистических и интеллектуальных моделей	11
1.5. Виды моделей системы Aidos-X.....	13
1.6. Результаты верификации моделей	15
2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	19
2.1. Решение задачи идентификации.....	19
2.2. Когнитивные функции.....	23
2.3. SWOT и PEST матрицы и диаграммы.....	28
2.4. Кластерно-конструктивный анализ классов и признаков.....	31
2.5. Нелокальные нейроны и нейронные сети.....	32
ЗАКЛЮЧЕНИЕ	35
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	37

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 исследования сведений о рейтинге ресторанного бизнеса на основе АСК-анализа данных с 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 «Restaurant Business Rankings 2020» из банка исходных данных по задачам искусственного интеллекта — Kaggle.

Для решения задачи используются стандартные возможности MS Office Word и Excel, а также система искусственного интеллекта «Aidos-X++».

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

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

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

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

С электронного ресурса [kaggle.com](https://www.kaggle.com/michau96/restaurant-business-rankings-2020) возьмем набор данных «Restaurant Business Rankings 2020» <https://www.kaggle.com/michau96/restaurant-business-rankings-2020>

Этот набор данных взят из электронного ресурса [kaggle.com](https://www.kaggle.com). Цель набора данных — предсказать, есть ли дальнейший рост ресторана и предпочтения людей, на основе определенных измерений, включенных в набор данных. На выбор этих экземпляров из более крупной базы данных было наложено несколько ограничений. В частности, все рестораны здесь из Штатов Америки.

CSV файл содержит следующие данные:

Rank
Restaurant
Sales
YOY_Sales %

Units
YOY_Units %
Headquarters
Segment_Category

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

Результат конвертации представлен на рисунке 1. Таким образом, в качестве классификационной шкалы решено использовать столбец В — «Rank». Остальные шкалы являются описательными.

Rank	Rank	Restaurant	Sales	YOY_Sales %	Units	YOY_Units %	Headquarters	Segment_Category
1	1	McDonald's	40412	4,90%	13846	-0,50%	NA	Quick Service & Burger
2	2	Starbucks	21380	8,60%	15049	3,00%	NA	Quick Service & Coffee Cafe
3	3	Chick-fil-A	11320	13,00%	2470	5,00%	NA	Quick Service & Chicken
4	4	Taco Bell	11293	9,00%	6766	2,70%	NA	Quick Service & Mexican
5	5	Burger King	10204	2,70%	7346	0,20%	NA	Quick Service & Burger
6	6	Subway	10200	-2,00%	23801	-4,00%	NA	Quick Service & Sandwich
7	7	Wendy's	9762	4,20%	5852	0,70%	NA	Quick Service & Burger
8	8	Dunkin'	9228	5,00%	9630	2,20%	NA	Quick Service & Coffee Cafe
9	9	Domino's	7044	6,90%	6126	4,30%	NA	Quick Service & Pizza
10	10	Panera Bread	5890	4,00%	2160	3,20%	NA	Fast Casual & Bakery Cafe
11	11	Pizza Hut	5558	0,60%	7306	-2,40%	NA	Quick Service & Pizza
12	12	Chipotle Mexican Grill	5509	14,80%	2584	5,30%	NA	Fast Casual & Mexican
13	13	Sonic Drive-In	4687	4,60%	3526	-2,10%	NA	Quick Service & Burger
14	14	KFC	4546	2,50%	4065	-0,20%	NA	Quick Service & Chicken
15	15	Olive Garden	4287	5,00%	866	1,30%	NA	Casual Dining & Italian/Pizza
16	16	Applebee's	4085	-3,00%	1665	-1,70%	NA	Casual Dining & Varied Menu
17	17	Panda Express	3946	12,00%	2209	4,90%	NA	Fast Casual & Asian/Noodle
18	18	Arby's	3884	1,90%	3359	0,90%	NA	Quick Service & Sandwich
19	19	Popeyes Louisiana Kitchen	3812	18,30%	2476	5,50%	NA	Quick Service & Chicken
20	20	Little Caesars	3811	2,70%	4237	-0,60%	NA	Quick Service & Pizza
21	21	Dairy Queen	3760	2,90%	4381	-0,60%	NA	Quick Service & Frozen Desserts

Рисунок 1 — Фрагмент обучающей выборки

1.3. Ввод выборки в систему Aidos-X

Теперь, когда имеется обучающая выборка в формате *.xlsx, 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.

2.3.2.2. Универсальный программный интерфейс импорта данных в систему "ЭЙДОС-Х++"

Автоматическая формализация предметной области: генерация классификационных и описательных шкал и градаций, а также обучающей и распознаваемой выборки на основе базы исходных данных: "Inp_data"

Задайте тип файла исходных данных: "Inp_data":

☐ XLS - MS Excel-2003 Стандарт XLS-файла
☒ XLSX- MS Excel-2007(2010) Стандарт XLSX-файла
☐ DBF - DBASE IV (DBF/NTX) Стандарт DBF-файла
☐ CSV - CSV => DBF конвертер Стандарт CSV-файла

Задайте параметры:

☒ Нули и пробелы считать ОТСУТСТВИЕМ данных
☐ Нули и пробелы считать ЗНАЧЕНИЯМИ данных
☐ Создавать БД средних по классам "Inp_davg.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

☒ Формализации предметной области (на основе "Inp_data")
☐ Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

☐ Равные интервалы с разным числом наблюдений
☒ Разные интервалы с равным числом наблюдений

Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp_data":

☒ Не применять сценарный метод АСК-анализа ☐ Применить сценарный метод АСК-анализа

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:

Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

☒ Только интервальные числовые значения (например: "1/3-(59873.0000000, 178545.6666667)")
☐ Только наименования интервальных числовых значений (например: "Минимальное")
☐ И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-(59873.0000000, 178545.6666667)")

Ok Cancel

Рисунок 2 — Ввод обучающей выборки

Стоит отметить следующие настройки:

- Тип файла — xlsx;
- Классификационные шкалы — 2;
- Описательные шкалы — 3-9;
- Применяется спец. интерпретация текстовых полей признаков;
- В качестве признаков рассматриваются слова, без проведения лемитизации.

В форме задания размерности вносим корректировки (рисунок 3) и пересчитываем шкалы и градации.

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

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы
1	RANK	1	1/10-(1.0000000, 25.0000000)
		2	2/10-(25.0000000, 50.0000000)
		3	3/10-(50.0000000, 75.0000000)
		4	4/10-(75.0000000, 100.0000000)
		5	5/10-(100.0000000, 125.0000000)
		6	6/10-(125.0000000, 150.0000000)
		7	7/10-(150.0000000, 175.0000000)
		8	8/10-(175.0000000, 200.0000000)
		9	9/10-(200.0000000, 225.0000000)
		10	10/10-(225.0000000, 250.0000000)

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 Графики будущих сценариев

Рисунок 5 — Классификационные шкалы и градации

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 6), а обучающей выборки режим 2.3.1. (рисунок 7):

2.2. Описательные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	RESTAURANT	1	1/250-54th Street Restaurant & Drafthouse
2	SALES	2	2/250-A&W All-American Food
3	YDY_SALES %	3	3/250-Anthony's Coal Fired Pizza
4	UNITS	4	4/250-Applebee's
5	YDY_UNITS %	5	5/250-Arby's
6	HEADQUARTERS	6	6/250-Au Bon Pain
7	SEGMENT_CATEGORY	7	7/250-Auntie Anne's
		8	8/250-Bahama Breeze Island Grille
		9	9/250-Bar Louie
		10	10/250-Baskin-Robbins
		11	11/250-Beef 'O' Brady's
		12	12/250-Benihana
		13	13/250-Big Boy/Frisch's Big Boy
		14	14/250-Biggby Coffee
		15	15/250-Bill Miller Bar-B-Q
		16	16/250-BJ's Restaurant & Brewhouse
		17	17/250-Black Angus Steakhouse
		18	18/250-Black Bear Diner
		19	19/250-Blaze Pizza
		20	20/250-Bob Evans
		21	21/250-Bojangles'
		22	22/250-Bonefish Grill
		23	23/250-Boston Market
		24	24/250-Braum's Ice Cream & Dairy Stores
		25	25/250-Brio Tuscan Grille
		26	26/250-Bruegger's Bagels
		27	27/250-Bubba Gump Shrimp Co.

Помощь Доб. шкалу Доб. град. шкалы Копир. шкалу Копир. град. шкалы Копир. шкалу с град. Удал. шкалу с град. Удал. град. шкалы Перекодировать Очистить Графики прошлых сценариев

Рисунок 6 — Описательные шкалы и градации (фрагмент)

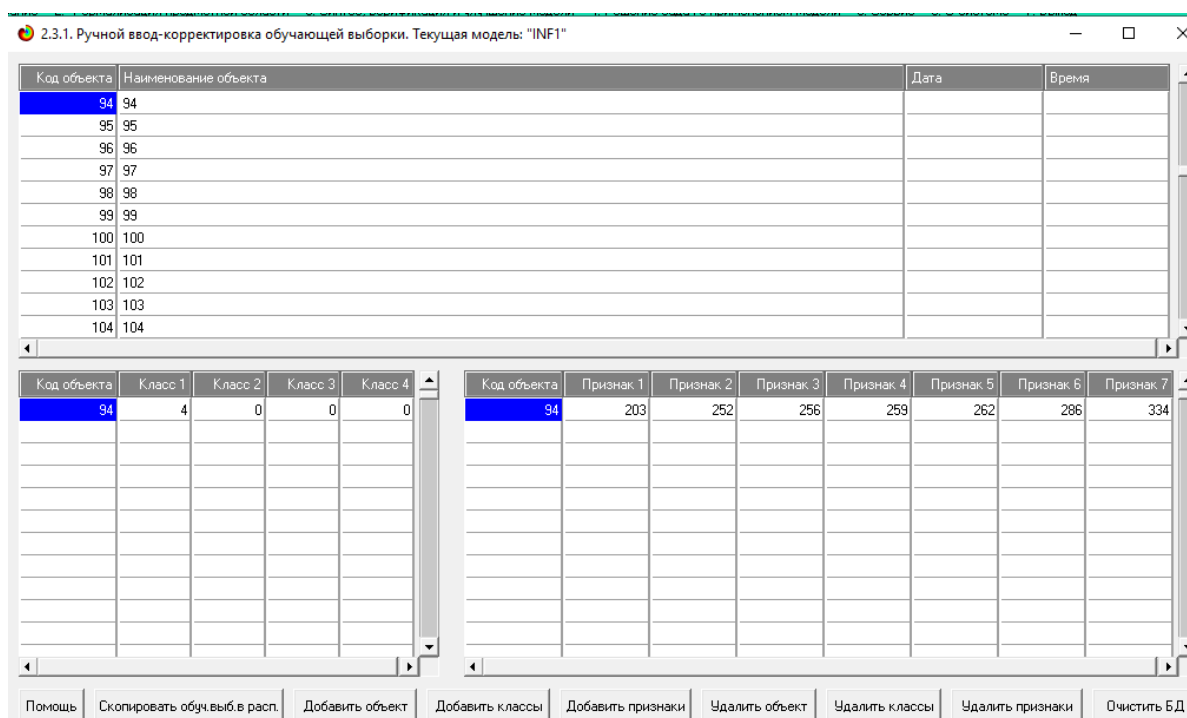


Рисунок 7 — Корректировка обучающей выборки (фрагмент)

Для просмотра эвентологических баз данных (баз событий) необходимо запустить режим 2.4 (рисунок 8),

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

	Наименование объекта	2. RANK	3. RESTAURANT	4. SALES	5. YOY_SALES %	6. UNITS	7. YOY_UNITS %	8. HEADQUARTERS	9. SEGMENT
1	1	1	137	253	256	259	261	286	
2	2	1	207	253	256	259	262	286	
3	3	1	41	253	256	259	262	286	
4	4	1	211	253	256	259	261	286	
5	5	1	30	253	255	259	261	286	
6	6	1	209	253	254	259	260	286	
7	7	1	240	253	255	259	261	286	
8	8	1	68	253	256	259	261	286	
9	9	1	65	253	256	259	262	286	
10	10	1	161	253	255	259	262	286	
11	11	1	172	253	255	259	260	286	
12	12	1	44	253	256	259	262	286	
13	13	1	204	253	255	259	260	286	
14	14	1	116	253	255	259	261	286	
15	15	1	156	253	256	259	261	286	
16	16	1	4	253	254	259	260	286	
0	1	2	3	4	5	6	7	8	9

Рисунок 8 — Просмотр базы событий (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 [11]).

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 9).

Рисунок 9 — Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию, приведенные на рисунке 9. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 10.

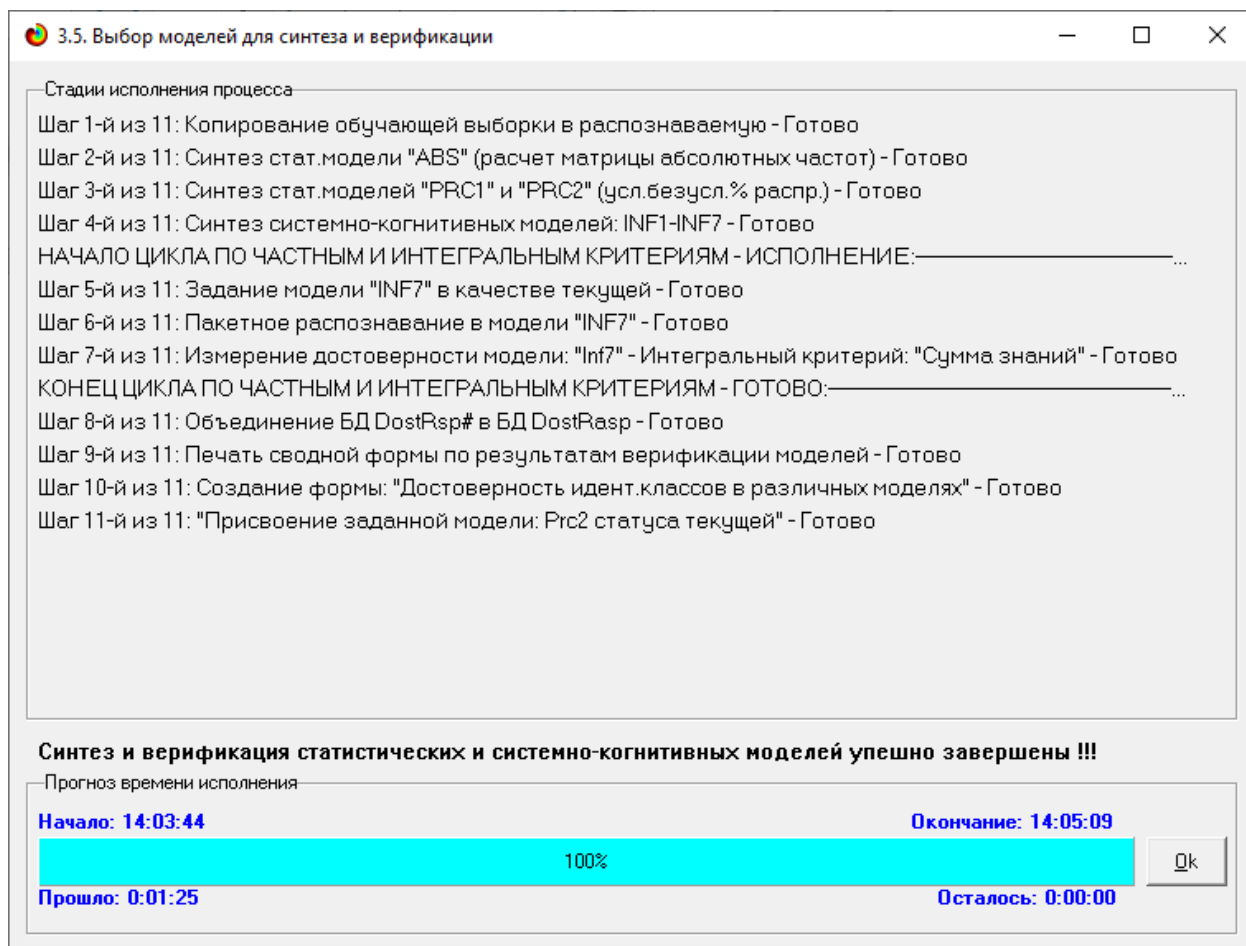


Рисунок 10 — Синтез и верификация статистических моделей
и моделей знаний

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 11) в матрицы условных и безусловных процентных распределений, и матрицы знаний (рисунок 12 и 13).

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

Код признака	Наименование описательной шкалы и градаций	1. RANK 1/10 (1.0, 25.0)	2. RANK 2/10 (25.0, 50.0)	3. RANK 3/10 (50.0, 75.0)	4. RANK 4/10 (75.0, 100.0)	5. RANK 5/10 (100.0, 125.0)	6. RANK 6/10 (125.0, 150.0)	7. RANK 7/10 (150.0, 175.0)	8. RANK 8/10 (175.0, 200.0)	9. RANK 9/10 (200.0, 225.0)	10. RANK 10/10 (225.0, 250.0)	Сумма
1	1/250-54th Street Restaurant & Drafthouse										1	
2	RESTAURANT-2/250-A&W All-American Food							1				
3	RESTAURANT-3/250-Anthony's Coal Fired Pizza									1		
4	RESTAURANT-4/250-Applebee's	1										
5	RESTAURANT-5/250-Arby's	1										
6	RESTAURANT-6/250-Au Bon Pain						1					
7	RESTAURANT-7/250-Auntie Anne's				1							
8	RESTAURANT-8/250-Bahama Breeze Island Grille							1				
9	RESTAURANT-9/250-Bar Louie						1					
10	RESTAURANT-10/250-Baskin-Robbins			1								
11	RESTAURANT-11/250-Beef 'O' Brady's								1			
12	RESTAURANT-12/250-Benihana					1						
13	RESTAURANT-13/250-Big Boy/Frisch's Big Boy					1						
14	RESTAURANT-14/250-Bigby Coffee										1	
15	RESTAURANT-15/250-Bill Miller Bar-B-Q											
16	16/250-BJ's Restaurant & Brewhouse		1								1	
17	RESTAURANT-17/250-Black Angus Steakhouse									1		
18	RESTAURANT-18/250-Black Bear Diner					1						
19	RESTAURANT-19/250-Blaze Pizza					1						
20	RESTAURANT-20/250-Bob Evans			1								
21	RESTAURANT-21/250-Bojangles'		1									
22	RESTAURANT-22/250-Bonafish Grill			1								
23	RESTAURANT-23/250-Boston Market				1							
24	RESTAURANT-24/250-Braum's Ice Cream & Dairy Stores						1					
25	RESTAURANT-25/250-Brio Tuscan Grille							1				

Рисунок 11 — Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений

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

Код признака	Наименование описательной шкалы и градаций	1. RANK 1/10 (1.0, 25.0)	2. RANK 2/10 (25.0, 50.0)	3. RANK 3/10 (50.0, 75.0)	4. RANK 4/10 (75.0, 100.0)	5. RANK 5/10 (100.0, 125.0)	6. RANK 6/10 (125.0, 150.0)	7. RANK 7/10 (150.0, 175.0)	8. RANK 8/10 (175.0, 200.0)	9. RANK 9/10 (200.0, 225.0)	10. RANK 10/10 (225.0, 250.0)	Сумма
1	1/250-54th Street Restaurant & Drafthouse											
2	RESTAURANT-2/250-A&W All-American Food							1.024				
3	RESTAURANT-3/250-Anthony's Coal Fired Pizza									1.024		
4	RESTAURANT-4/250-Applebee's	1.024										
5	RESTAURANT-5/250-Arby's	1.024										
6	RESTAURANT-6/250-Au Bon Pain						1.024					
7	RESTAURANT-7/250-Auntie Anne's				1.024							
8	RESTAURANT-8/250-Bahama Breeze Island Grille							1.024				
9	RESTAURANT-9/250-Bar Louie						1.024					
10	RESTAURANT-10/250-Baskin-Robbins			1.024								
11	RESTAURANT-11/250-Beef 'O' Brady's								1.024			
12	RESTAURANT-12/250-Benihana					1.024						
13	RESTAURANT-13/250-Big Boy/Frisch's Big Boy					1.024						
14	RESTAURANT-14/250-Bigby Coffee											
15	RESTAURANT-15/250-Bill Miller Bar-B-Q										1.024	
16	16/250-BJ's Restaurant & Brewhouse		1.024									
17	RESTAURANT-17/250-Black Angus Steakhouse										1.024	
18	RESTAURANT-18/250-Black Bear Diner					1.024						
19	RESTAURANT-19/250-Blaze Pizza					1.024						
20	RESTAURANT-20/250-Bob Evans			1.024								
21	RESTAURANT-21/250-Bojangles'		1.024									
22	RESTAURANT-22/250-Bonafish Grill			1.024								
23	RESTAURANT-23/250-Boston Market				1.024							
24	RESTAURANT-24/250-Braum's Ice Cream & Dairy Stores						1.024					
25	RESTAURANT-25/250-Brio Tuscan Grille							1.024				

Рисунок 12 — Модель INF1 (фрагмент)

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

Код признака	Наименование описательной шкалы и градаций	1. RANK 1/10 (1.0, 25.0)	2. RANK 2/10 (25.0, 50.0)	3. RANK 3/10 (50.0, 75.0)	4. RANK 4/10 (75.0, 100.0)	5. RANK 5/10 (100.0, 125.0)	6. RANK 6/10 (125.0, 150.0)	7. RANK 7/10 (150.0, 175.0)	8. RANK 8/10 (175.0, 200.0)	9. RANK 9/10 (200.0, 225.0)	10. RANK 10/10 (225.0, 250.0)
1	1/250-54th Street Restaurant & Draffhouse										
2	RESTAURANT-2/250-AAW/ All-American Food							9.000			
3	RESTAURANT-3/250-Anthony's Coal Fired Pizza									9.000	
4	RESTAURANT-4/250-Applebee's	9.000									
5	RESTAURANT-5/250-Airby's	9.000									
6	RESTAURANT-6/250-Au Bon Pain						9.000				
7	RESTAURANT-7/250-Auntie Anne's				9.000						
8	RESTAURANT-8/250-Bahama Breeze Island Grille							9.000			
9	RESTAURANT-9/250-Bar Louie						9.000				
10	RESTAURANT-10/250-Baskin-Robbins			9.000							
11	RESTAURANT-11/250-Beef 'O' Brady's								9.000		
12	RESTAURANT-12/250-Benihana					9.000					
13	RESTAURANT-13/250-Big Boy/Frisch's Big Boy					9.000					
14	RESTAURANT-14/250-Bigby Coffee										
15	RESTAURANT-15/250-Bill Miller Bar-B-Q									9.000	
16	16/250-BJ's Restaurant & Brewhouse		9.000								
17	RESTAURANT-17/250-Black Angus Steakhouse									9.000	
18	RESTAURANT-18/250-Black Bear Diner					9.000					
19	RESTAURANT-19/250-Blaze Pizza					9.000					
20	RESTAURANT-20/250-Bob Evans			9.000							
21	RESTAURANT-21/250-Bojangles'		9.000								
22	RESTAURANT-22/250-Bonfish Grill			9.000							
23	RESTAURANT-23/250-Boston Market				9.000						
24	RESTAURANT-24/250-Braum's Ice Cream & Dairy Stores						9.000				
25	RESTAURANT-25/250-Brio Tuscan Grille							9.000			

Рисунок 13 — Модель INF5 (фрагмент)

1.6. Результаты верификации моделей

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

3.4. Обобщ.форма по достов.моделям при разн.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергена	Степень модальн. уровня истинно-поло. решений (STP)	Степень модальн. уровня истинно-отриц. решений (STN)	Степень модальн. уровня ложно-поло. решений
1. ABS - частный критерий: количество встреч сочетаний: "класс"	Корреляция абс.частот с абс. частот по признакам	250	250			2250	0.100	1.000	0.182	201.941		1066
1. ABS - частный критерий: количество встреч сочетаний: "класс"	Степень абс. частот по признакам	250	250			2250	0.100	1.000	0.182	180.611		980
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Корреляция усл.отн.частот с о. частот по признакам	250	250			2250	0.100	1.000	0.182	201.941		1066
2. PRC1 - частный критерий: усл. вероятность иго признака сред.	Степень усл.отн.частот по признакам	250	250			2250	0.100	1.000	0.182	180.611		980
3. PRC2 - частный критерий: условная вероятность иго признака сред.	Корреляция усл.отн.частот с о. частот по признакам	250	250			2250	0.100	1.000	0.182	201.941		1066
3. PRC2 - частный критерий: условная вероятность иго признака сред.	Степень усл.отн.частот по признакам	250	250			2250	0.100	1.000	0.182	180.611		980
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс знаний	250	249	1839	411	1	0.377	0.996	0.547	159.002	613.529	61
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Степень знаний	250	250	955	1295		0.162	1.000	0.279	159.885	160.650	185
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс знаний	250	249	1839	411	1	0.377	0.996	0.547	159.002	613.529	61
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Степень знаний	250	250	955	1295		0.162	1.000	0.279	159.885	160.650	185
6. INF3 - частный критерий: Хинквардат, разности между факт.	Семантический резонанс знаний	250	249	1441	809	1	0.235	0.996	0.381	164.533	468.253	345
6. INF3 - частный критерий: Хинквардат, разности между факт.	Степень знаний	250	249	1441	809	1	0.235	0.996	0.381	134.279	403.701	269
7. INF4 - частный критерий: ROI (Return On Investment); вероятно.	Семантический резонанс знаний	250	250	2170	80		0.758	1.000	0.862	137.776	604.063	4
7. INF4 - частный критерий: ROI (Return On Investment); вероятно.	Степень знаний	250	250	765	1485		0.144	1.000	0.252	151.028	18.901	108
8. INF5 - частный критерий: ROI (Return On Investment); вероятно.	Семантический резонанс знаний	250	250	2170	80		0.758	1.000	0.862	137.776	604.063	4
8. INF5 - частный критерий: ROI (Return On Investment); вероятно.	Степень знаний	250	250	765	1485		0.144	1.000	0.252	151.028	18.901	108
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер.	Семантический резонанс знаний	250	249	820	1430	1	0.148	0.996	0.258	162.917	140.887	447
9. INF6 - частный критерий: разн.усл.и безусл.вероятностей; вер.	Степень знаний	250	249	701	1549	1	0.138	0.996	0.243	134.279	104.281	380
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; вер.	Семантический резонанс знаний	250	249	820	1430	1	0.148	0.996	0.258	162.917	140.887	447
10. INF7 - частный критерий: разн.усл.и безусл.вероятностей; вер.	Степень знаний	250	249	706	1544	1	0.139	0.996	0.244	134.279	104.281	380

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

Рисунок 14 — Оценки достоверности моделей

Наиболее достоверной в данном приложении оказалась модель INF3 при интегральном критерии «Семантический резонанс знаний». При этом точность модели (F-мера Ван Ризбергена) составляет 0.862, 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 15).

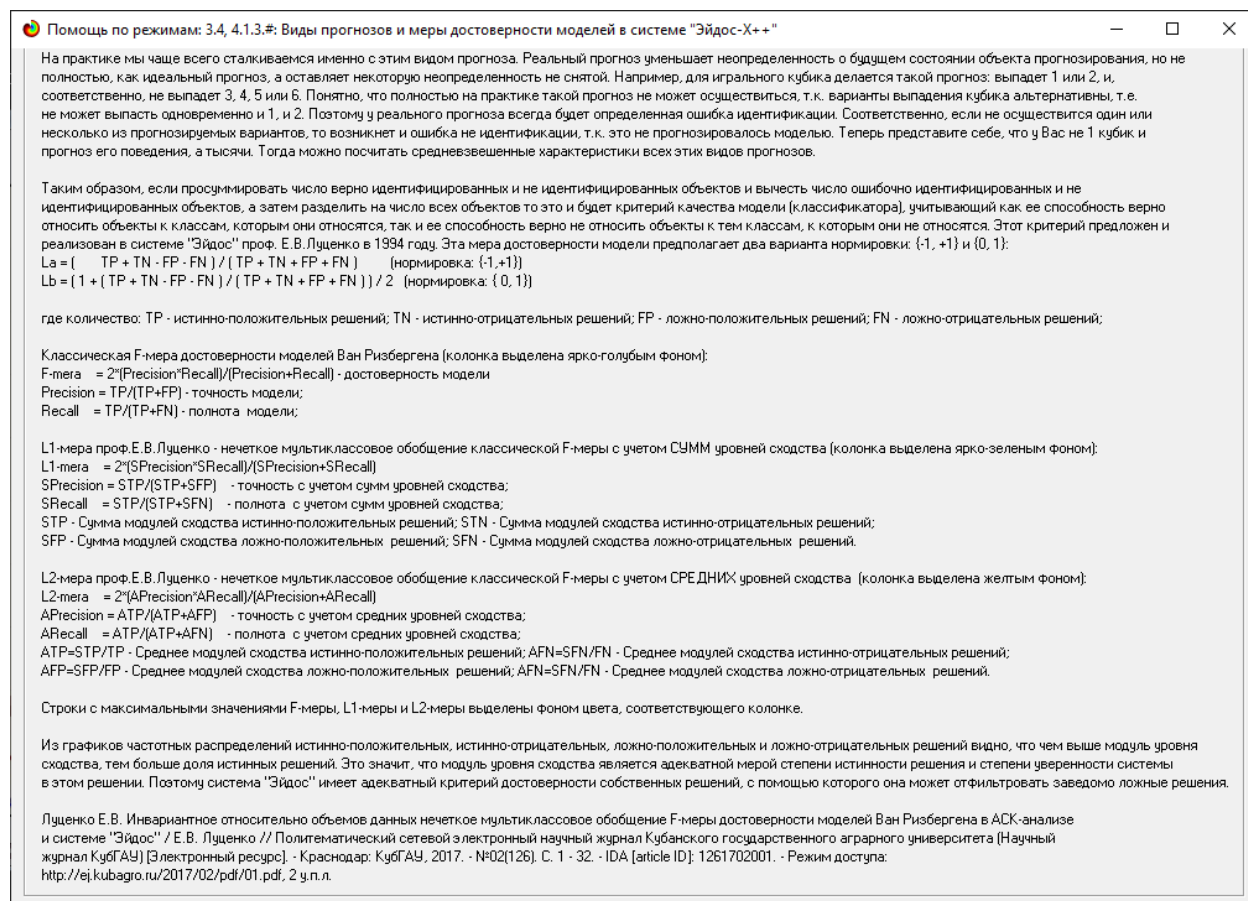


Рисунок 15 — Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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 ИИ интеллектуальных технологий.

На рисунке 16 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам идентификации характеристик оружия в СК-модели INF3 по данным обучающей выборки:

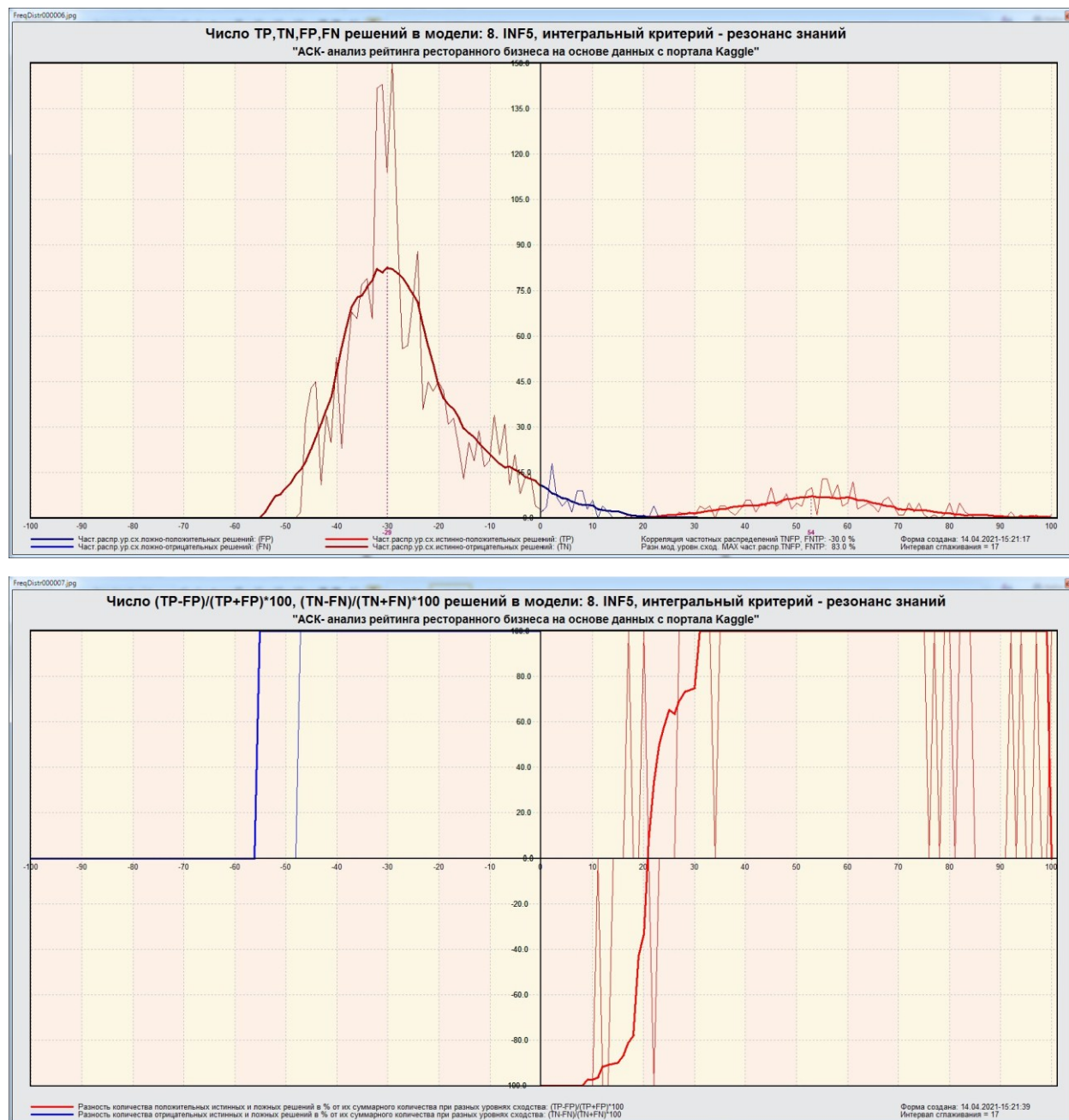


Рисунок 16 — Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf5

Рисунок 16 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе.

Левое распределение, большее по амплитуде включает истинно-отрицательные и ложноположительные решения, а правое, меньшее по амплитуде, включает ложные отрицательные и истинно-положительные решения.

Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу идентификации вида оружия по его характеристикам и другие задачи.

Видно, что для отрицательных решений количество истинных решений всегда значительно превосходит количество ложных решений, причем при уровнях различия больше примерно 5% ложные отрицательные решения вообще практически отсутствуют.

Видно также, что для положительных решений картина более сложная и включает 3 диапазона уровней сходства

- 1) при уровнях сходства от 0% до 30% количество ложных решений больше числа истинных;
- 2) при уровнях сходства выше 30% ложные положительные решения не встречаются.

2. РЕШЕНИЕ ЗАДАЧ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи идентификации

В соответствии со схемой обработки данных, информации и знаний в системе «Эйдос», присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (рисунок 17):

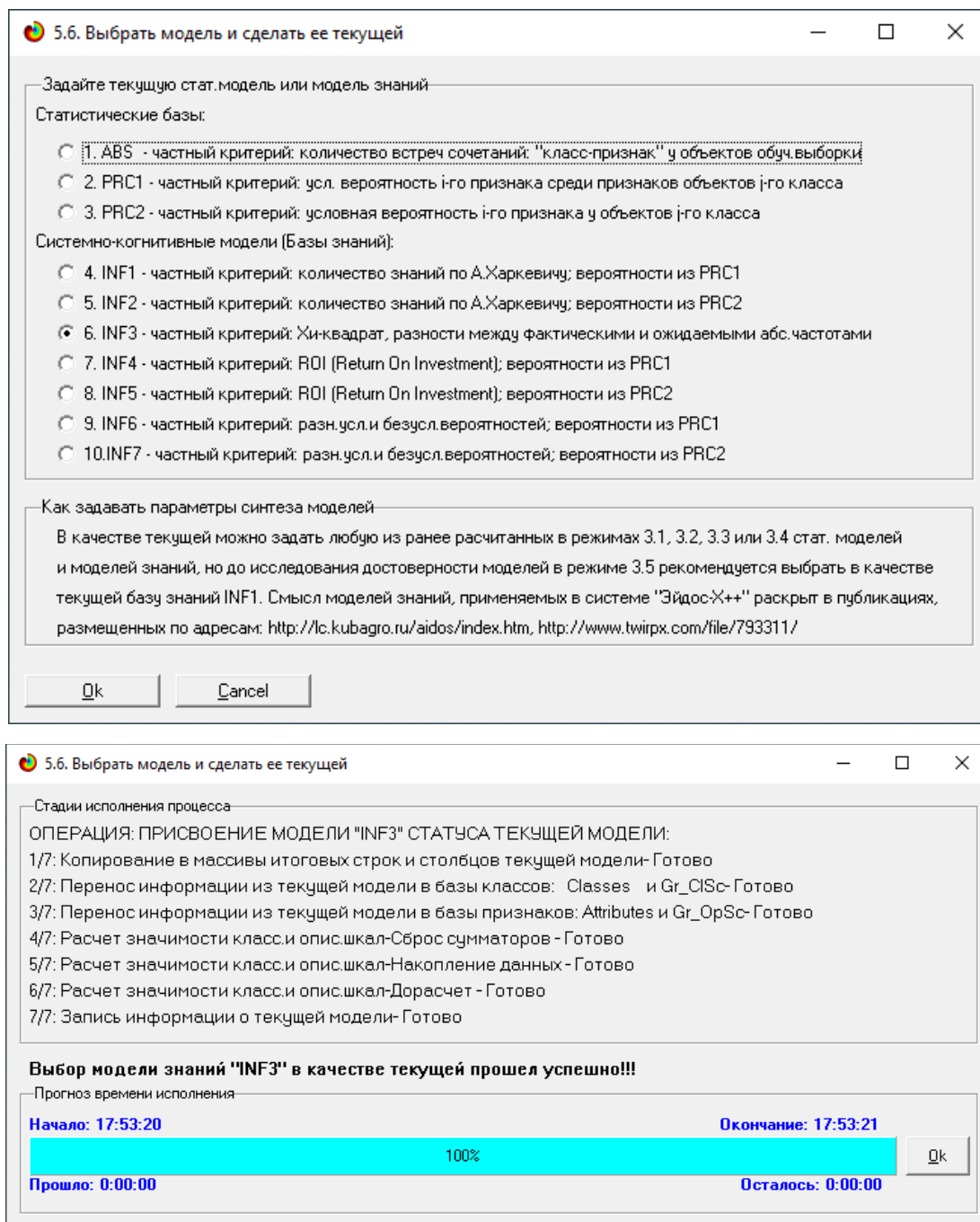


Рисунок 17. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

Решим задачу системной идентификации, т.е. определение класса оружия на основе обучающей выборки в наиболее достоверной СК-модели INF3 на GPU (рисунок 18).

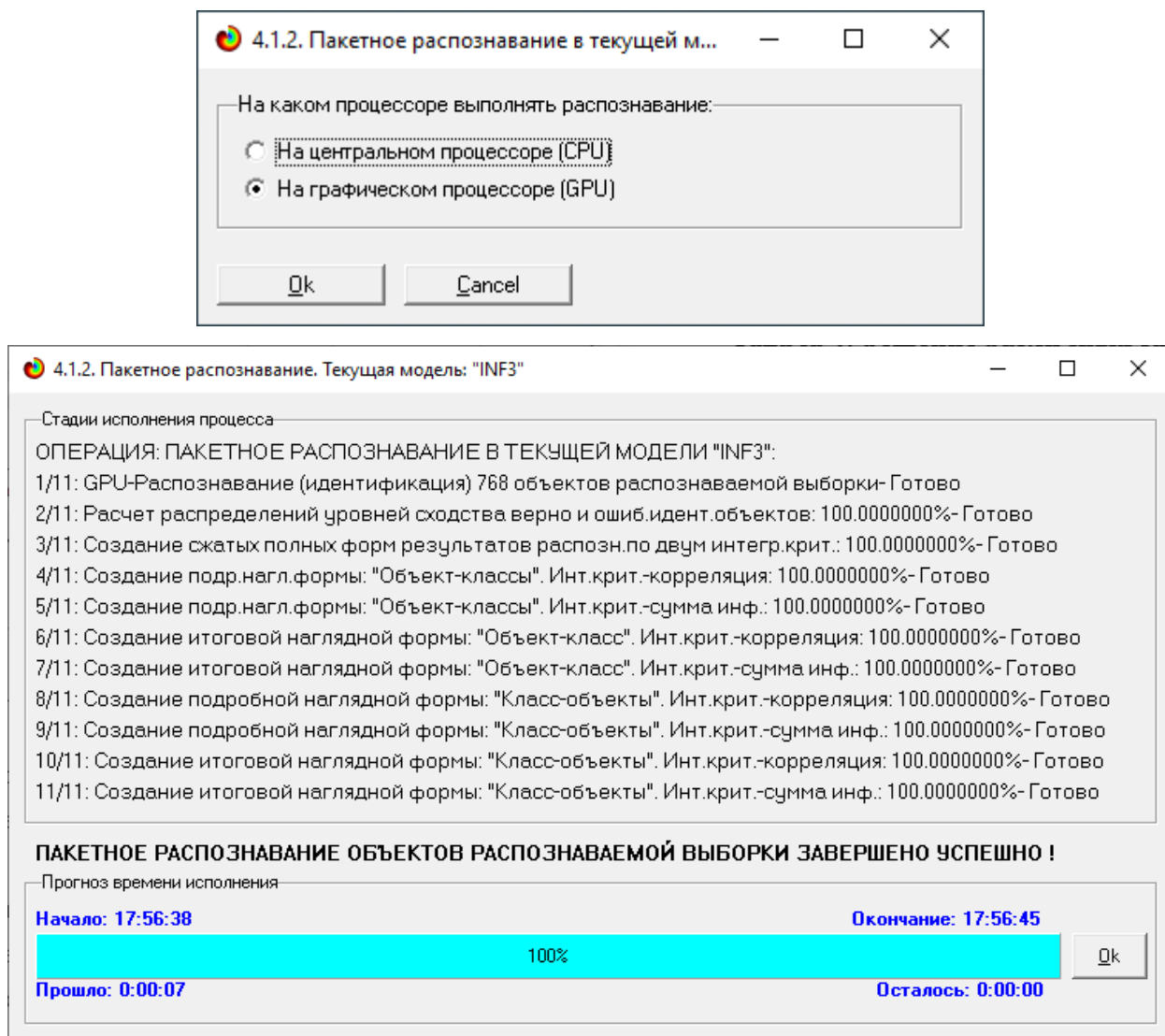


Рисунок 18. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Из рисунка 18 видно, что процесс идентификации занял 7 секунд.

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.

Отметим, что 99% этого времени заняла не сама идентификация на GPU, а создание 10 выходных форм на основе результатов этого прогнозирования. Эти формы отражают результаты прогнозирования в различных разрезах и обобщениях:

На рисунке 19 приведены примеры прогнозов высокой и низкой достоверности частоты объектов и классов в модели INF3 на основе наблюдения предыстории их развития:

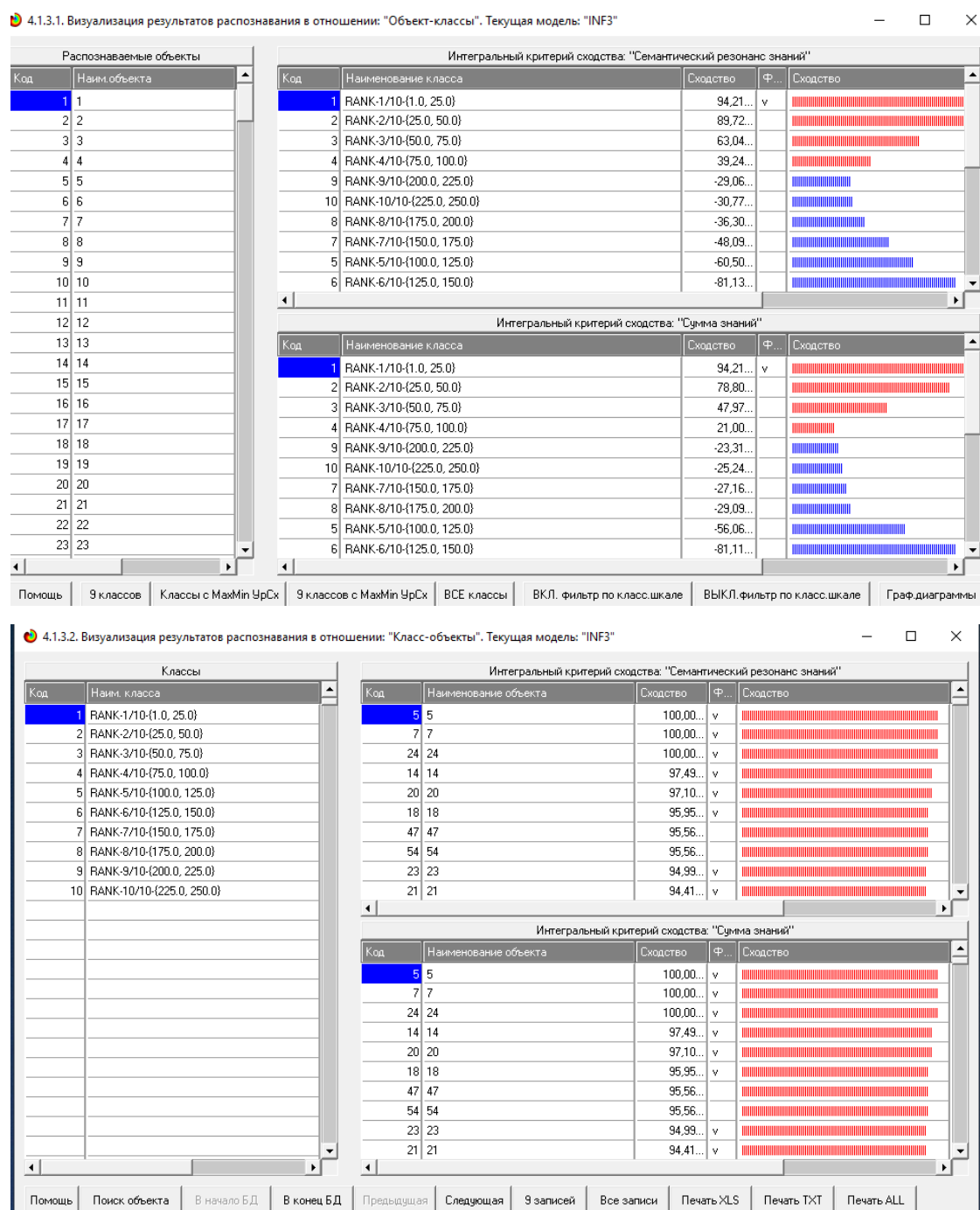


Рисунок 19. Выходные формы по результатам идентификации вида оружия по его характеристикам

Символ «√» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка 13 видно, что результаты идентификации являются отличными, естественно при учете информации из рисунка 9 о том, что достоверные прогнозы в данной модели

имеют уровень сходства выше 42%, т.е., по сути, результаты с более низким уровнем сходства надо просто игнорировать.

2.2. Когнитивные функции

Рассмотрим режим 4.5, в котором реализована возможность визуализации когнитивных функций для любых моделей и любых сочетаний классификационных и описательных шкал (рисунок 20).

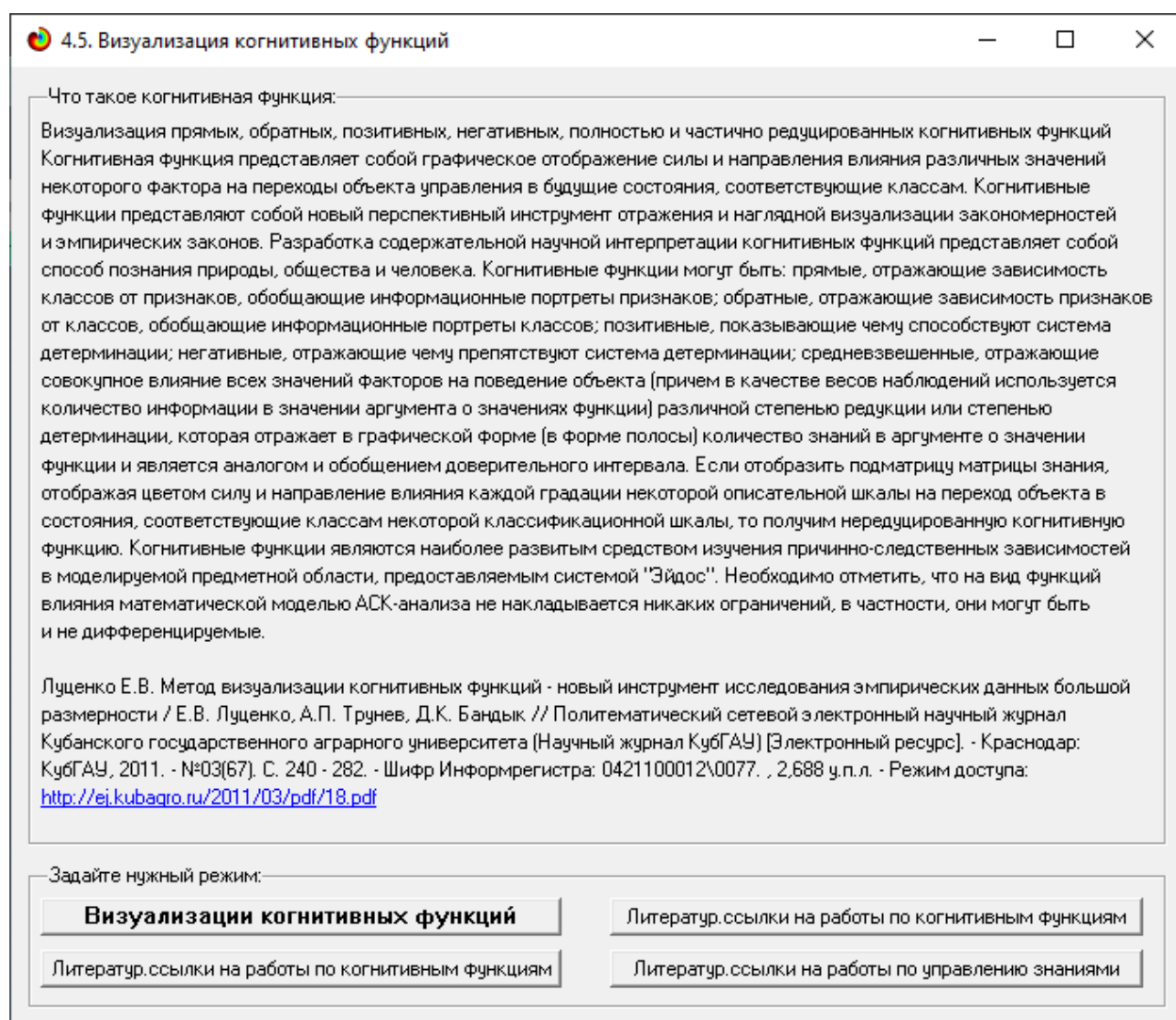
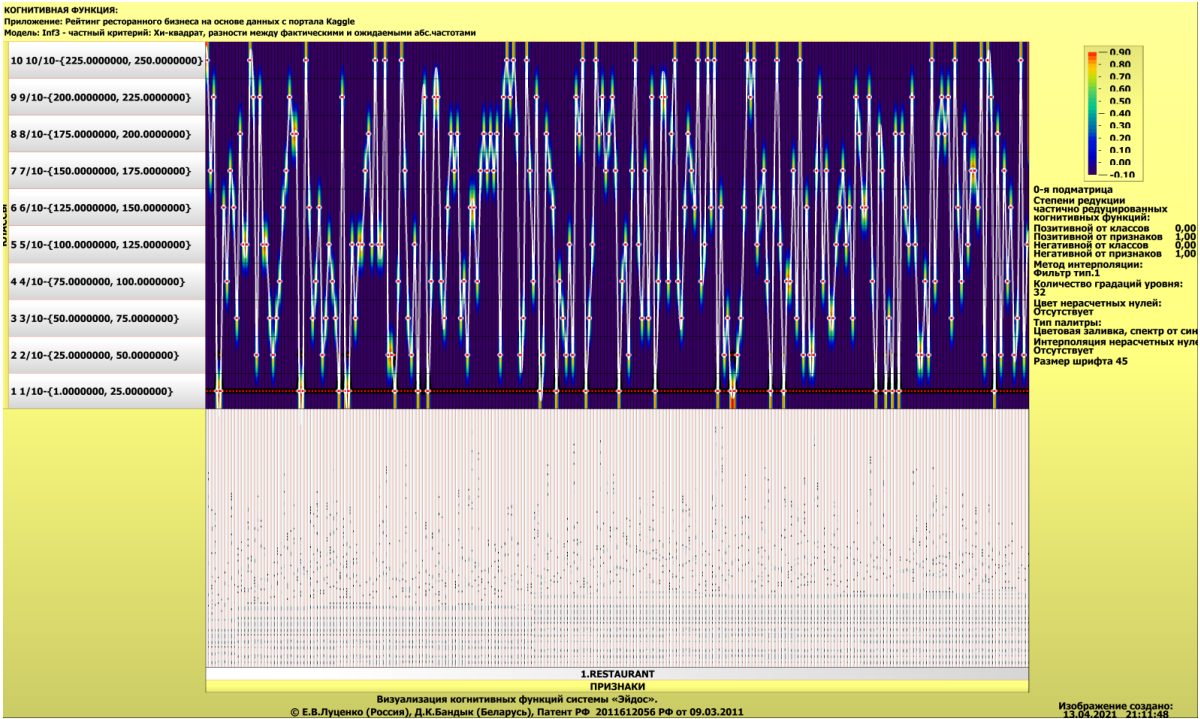
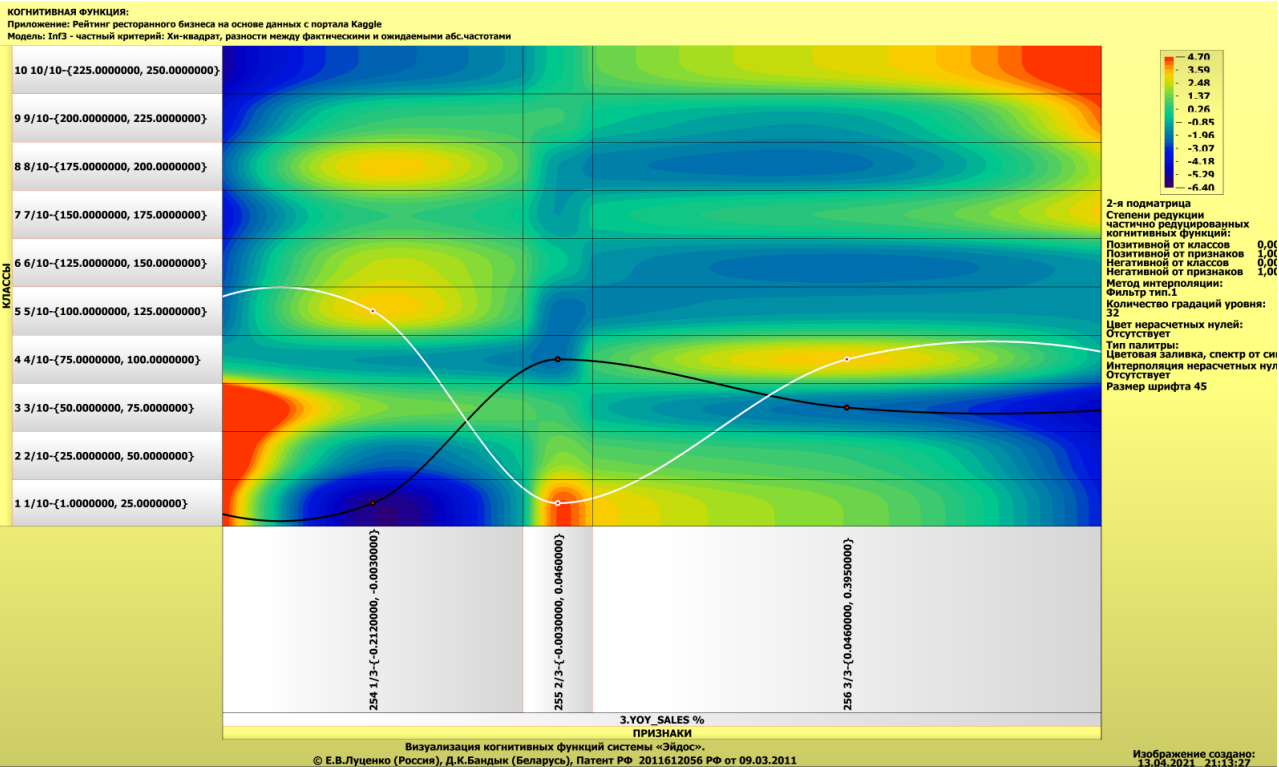
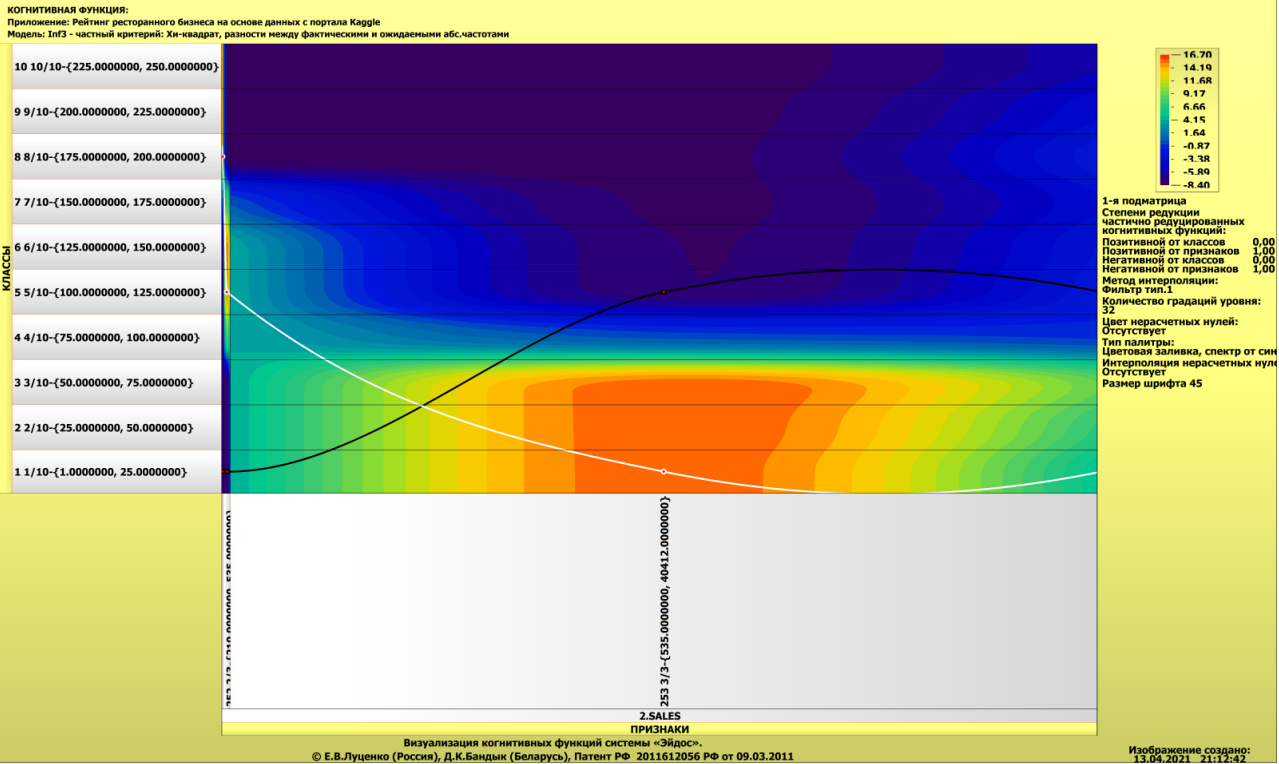
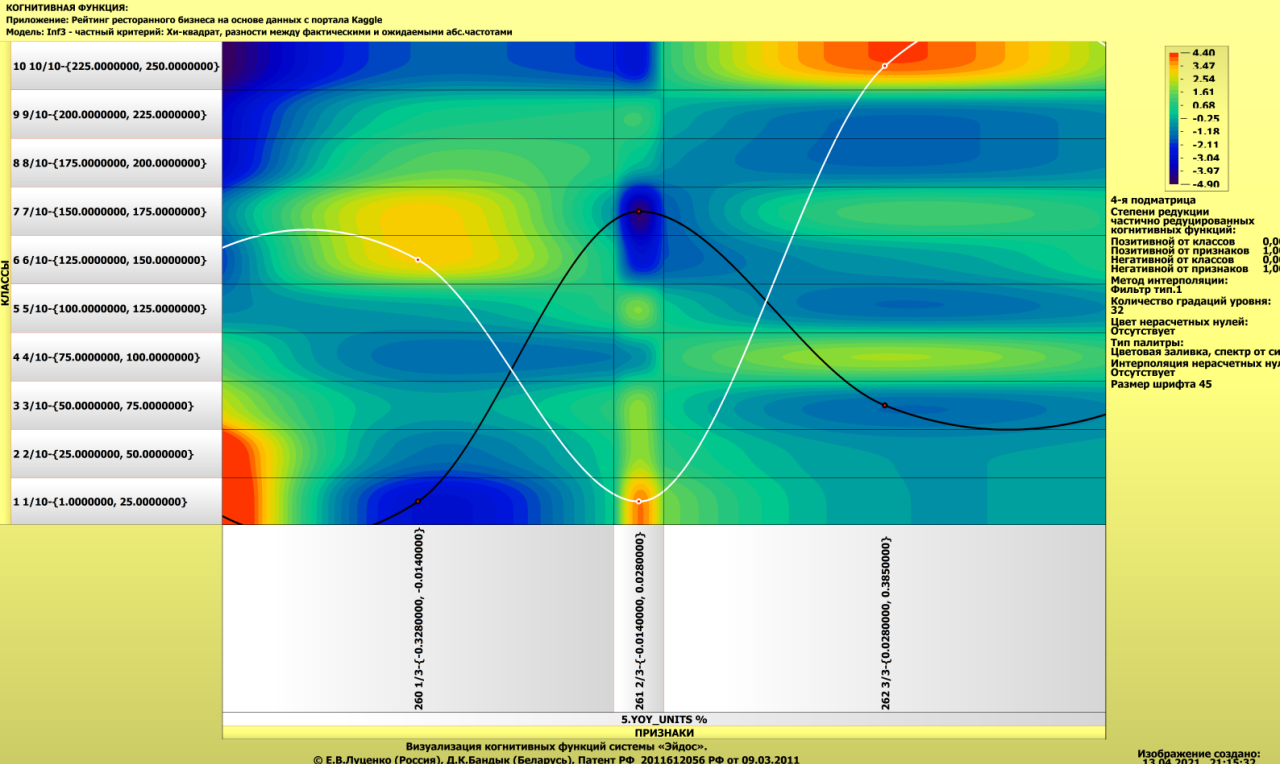
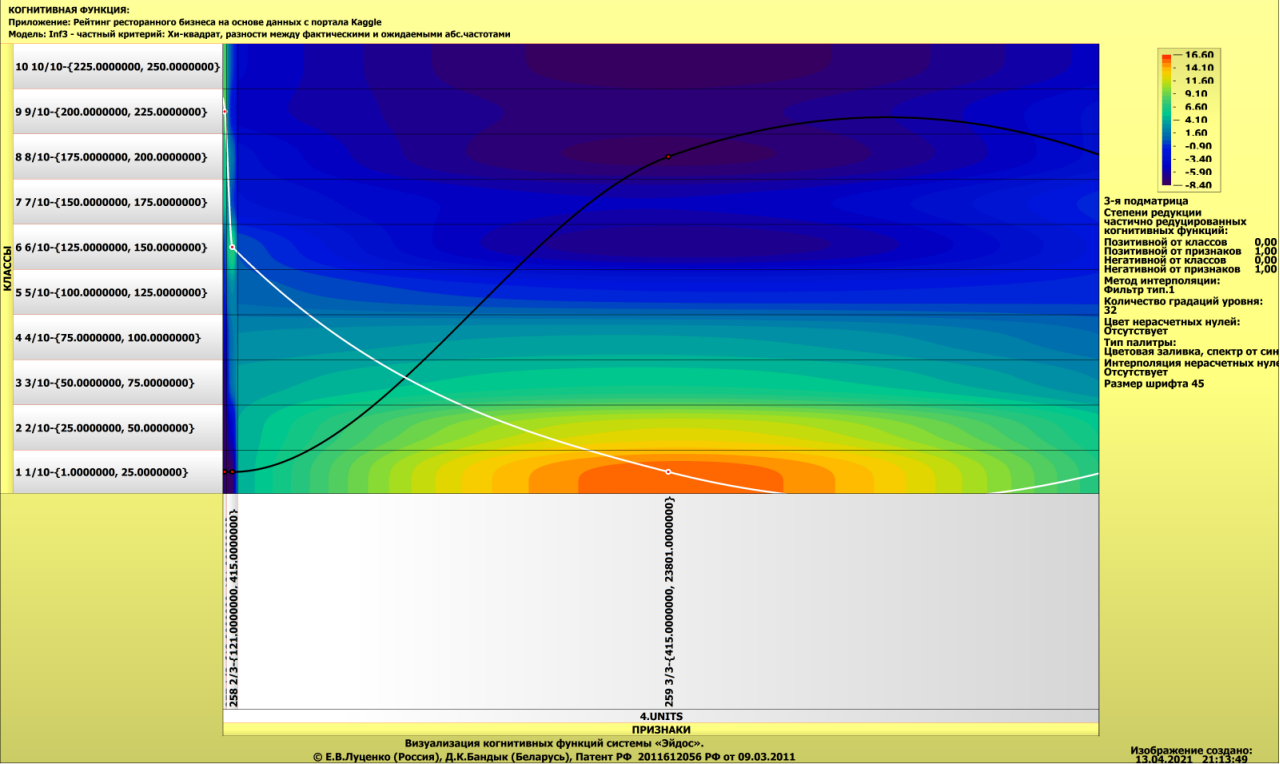


Рисунок 20 — Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»

Применительно к задаче, рассматриваемой в данной работе, когнитивная функция показывает, какое количество информации содержится в различных значениях факторов о том, что объект моделирования перейдет в те или иные будущие состояния. Поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 21 приведены визуализации всех когнитивных функций данного приложения для модели INF3.







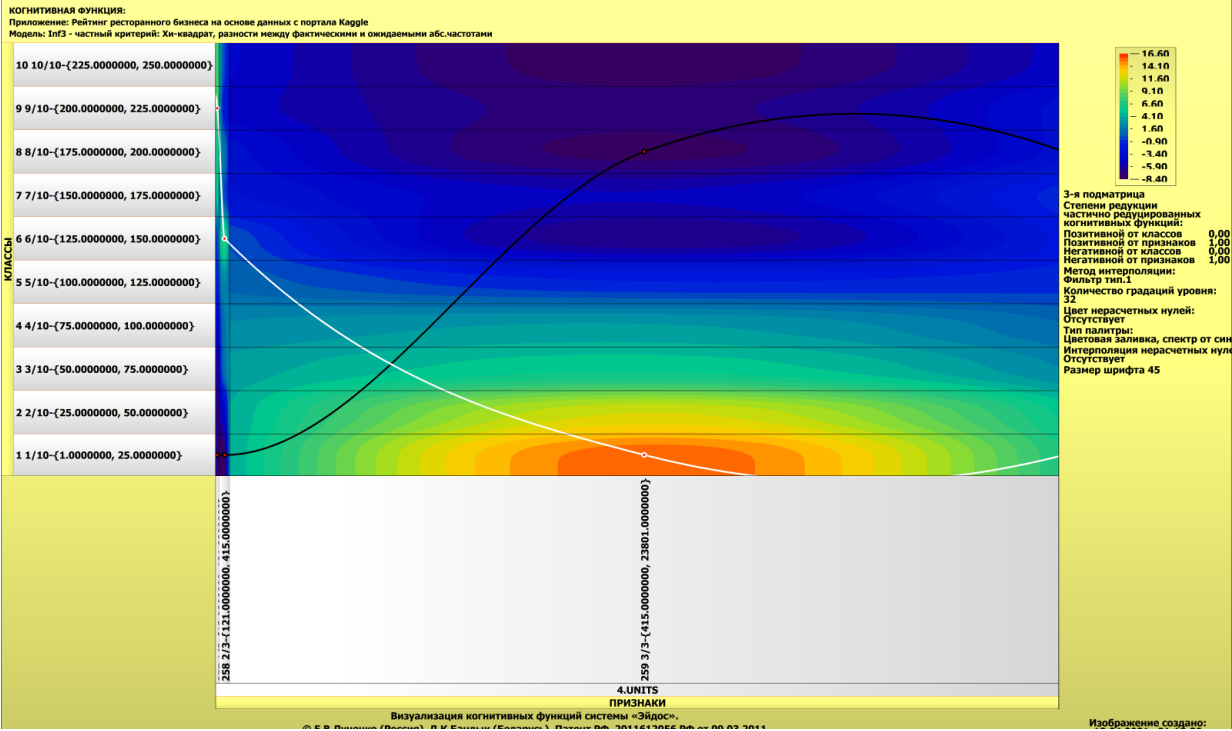
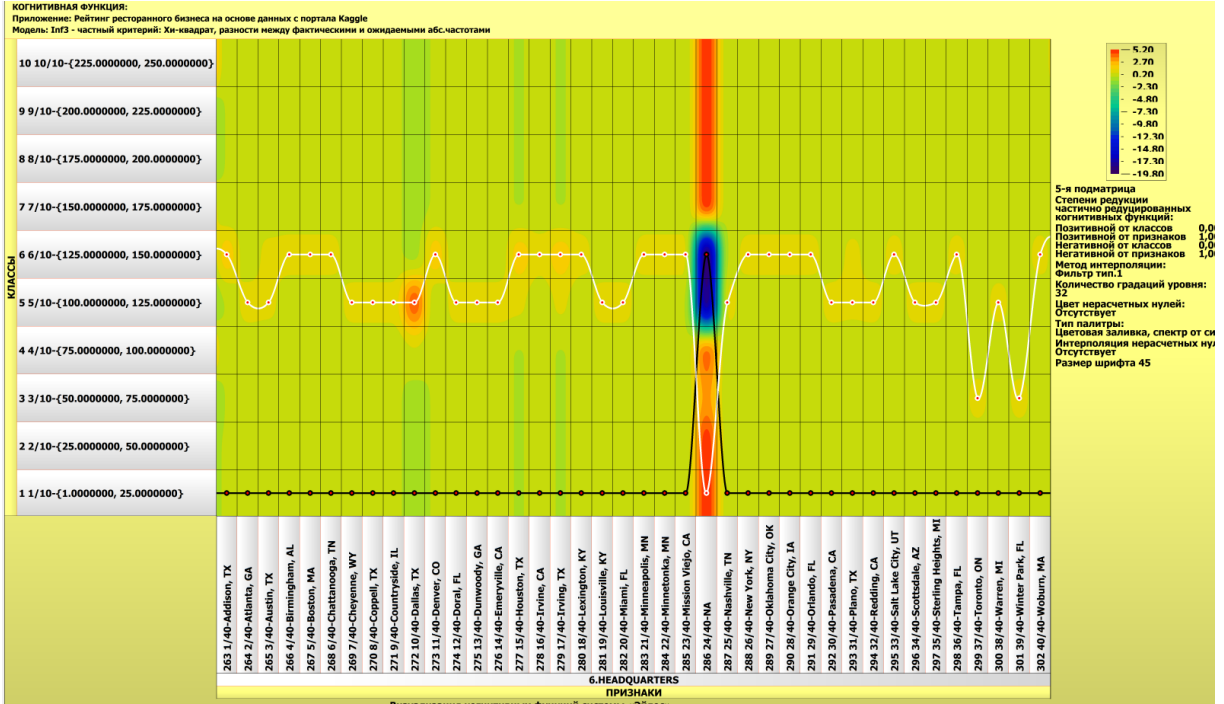


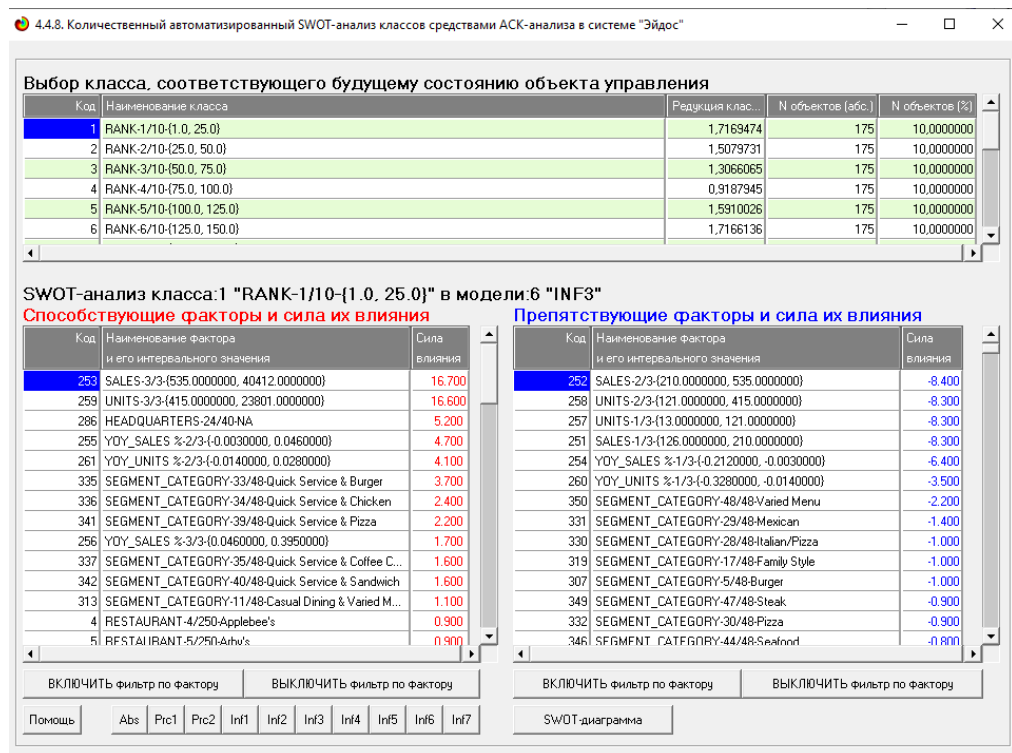
Рисунок 21 — Визуализация когнитивных функций для обобщенных классов и описательных шкал в модели INF4

2.3. SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов.

Ясно, что эксперты это делают 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 «Эйдос».

Данная система всегда обеспечивала возможность проведения количественного автоматизированного SWOT-анализа без использования экспертных оценок непосредственно на основе эмпирических данных. Результаты SWOT-анализа выводились в форме информационных портретов. В версии системы под MS Windows: «Эйдос-X++» предложено автоматизированное количественное решение прямой и обратной задач SWOT-анализа с построением традиционных SWOT-матриц и диаграмм (рисунок 22).



SWOTDiagrCls0001-06.jpg

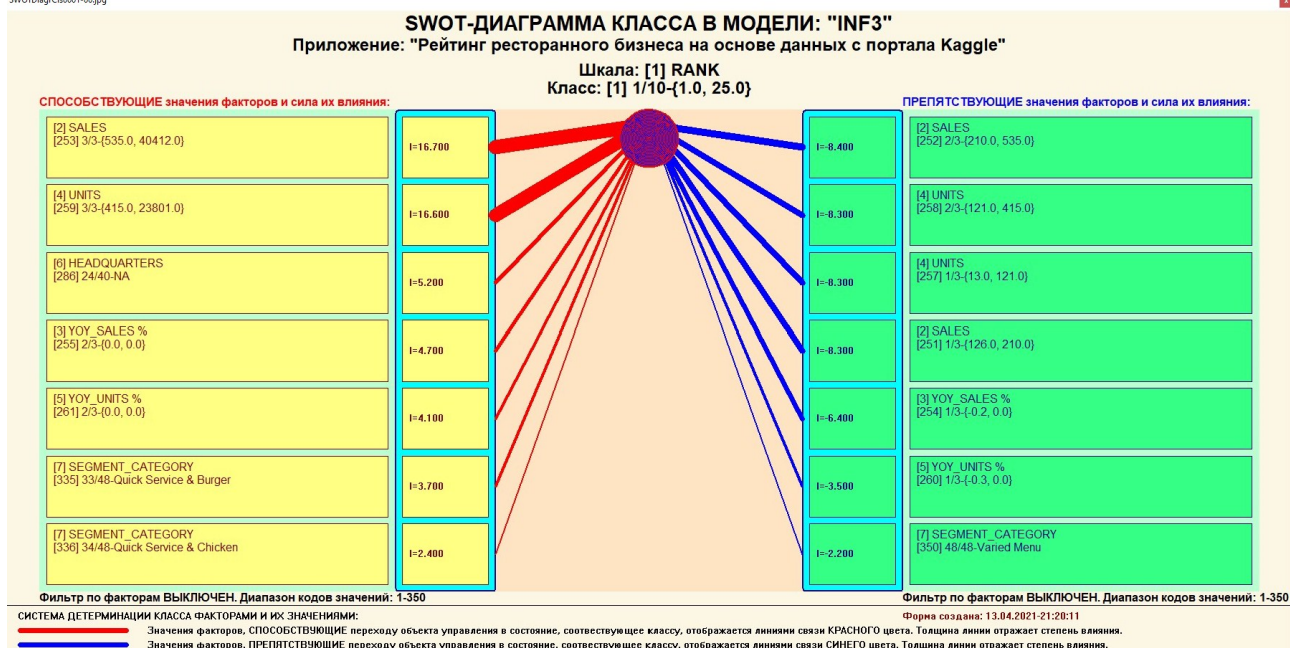
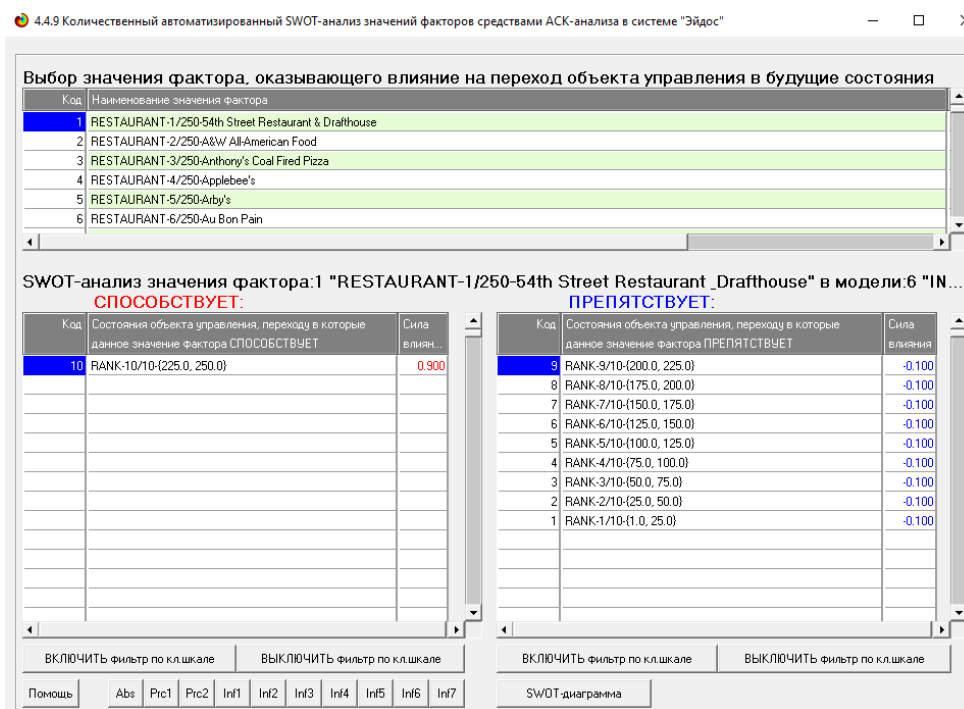


Рисунок 22 — Пример SWOT-матрицы в модели INF3

На рисунке 23 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF3.



SWOTdiagAtr0001-06.jpg

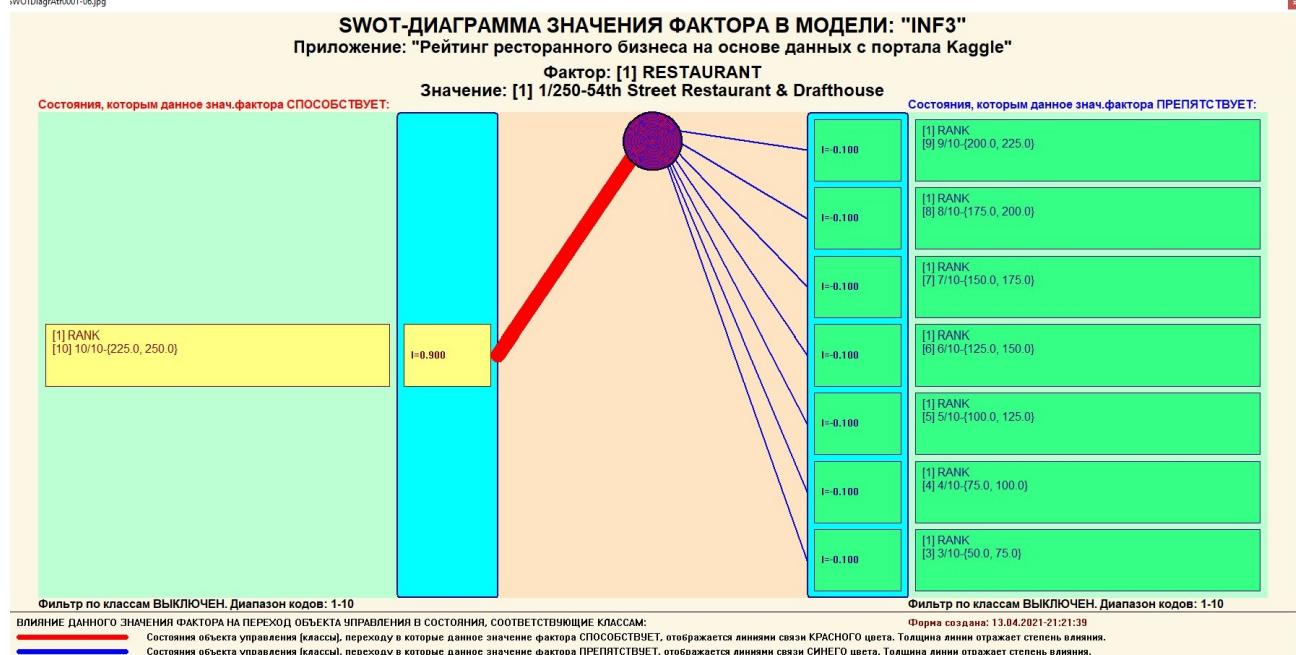


Рисунок 23 — Пример инвертированной SWOT-диаграммы в модели INF3

2.4. Кластерно-конструктивный анализ классов и признаков

В режиме 4.3.2.1, после расчета матриц сходства, кластеров и конструкторов, строим агломеративные дендрограммы классов и графики изменения межклатерных расстояний (рисунок 24):.

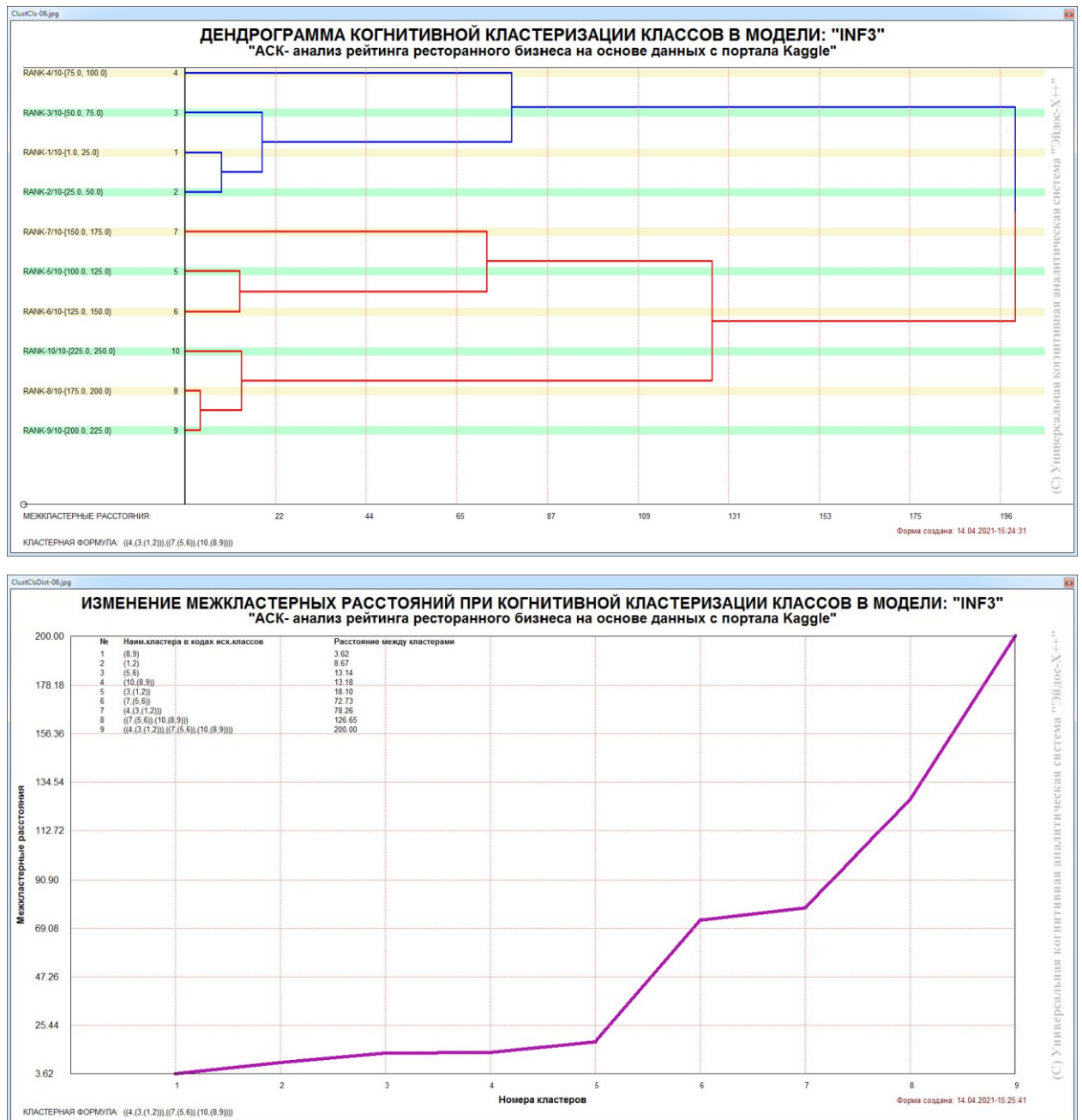


Рисунок 24. Аагломеративные дендрограммы классов и графики изменения межклатерных расстояний

В режиме 4.3.2.2, после расчета матриц сходства, кластеров и конструкторов, строим 2D сеть классов в выбранной модели знаний, для наглядного представления сети классов.

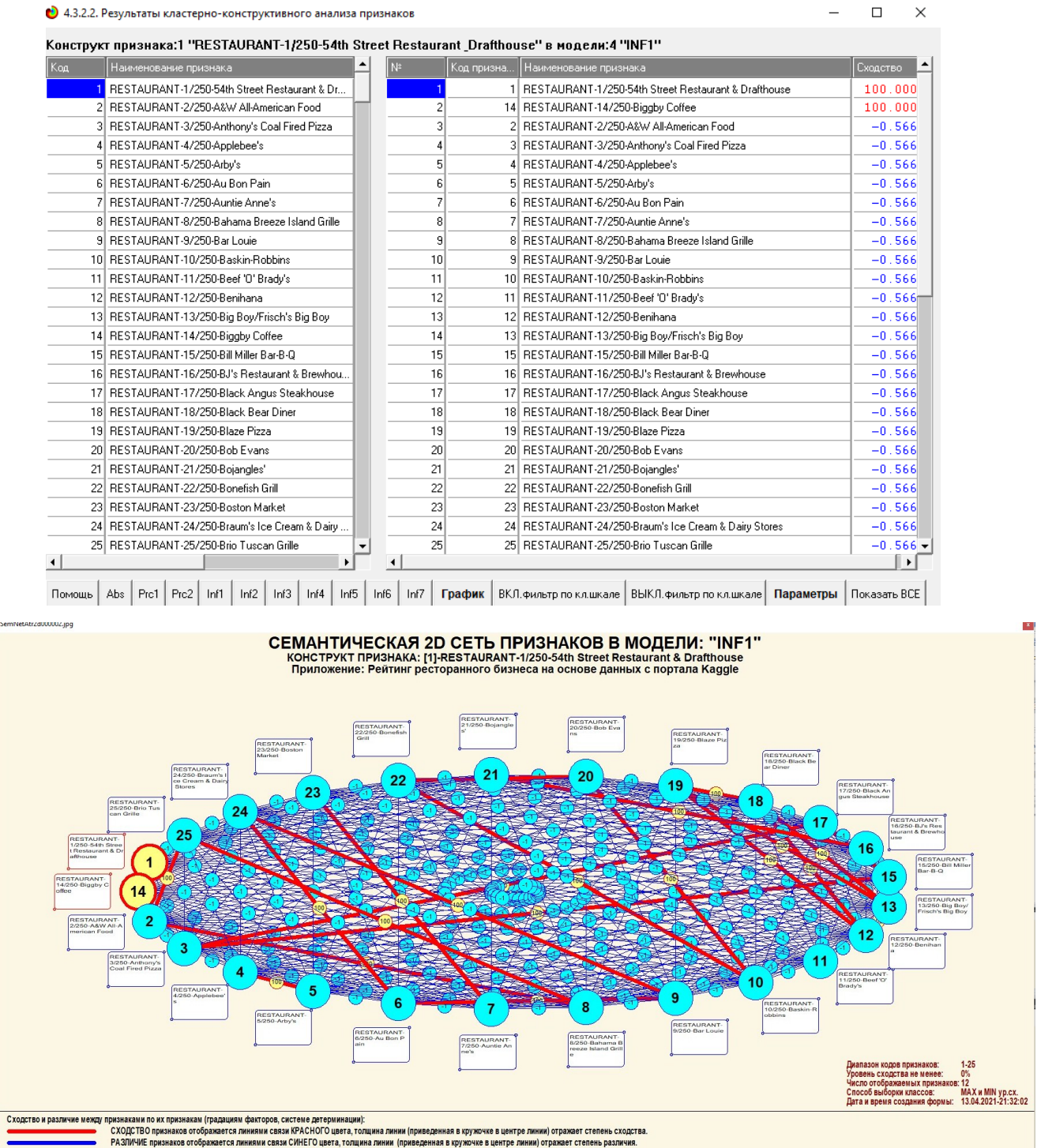


Рисунок 24 — Результаты кластерно-конструктивного анализа признаков

2.5. Нелокальные нейроны и нейронные сети

Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует не локальную нейронную сеть. Нелокальные нейроны и нелокальные нейронные сети представлены на рисунке 25:

4.4.10.Графическое отображение нелокальных нейронов в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	RANK-1/10-{1.0, 25.0}
2	RANK-2/10-{25.0, 50.0}
3	RANK-3/10-{50.0, 75.0}
4	RANK-4/10-{75.0, 100.0}
5	RANK-5/10-{100.0, 125.0}
6	RANK-6/10-{125.0, 150.0}
7	RANK-7/10-{150.0, 175.0}
8	RANK-8/10-{175.0, 200.0}
9	RANK-9/10-{200.0, 225.0}
10	RANK-10/10-{225.0, 250.0}

Подготовка визуализации нейрона: 1 "RANK-1/10-{1.0, 25.0}" в модели: 6 "INF3"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
253	SALES-3/3-{535.0000000, 40412.0000000}	16.700
258	UNITS-2/3-{415.0000000, 23801.0000000}	16.600
286	HEADQUARTERS-24/40-NA	5.200
255	YOY_SALES %2/3-{0.0030000, 0.0460000}	4.700
261	YOY_UNITS %2/3-{0.0140000, 0.0280000}	4.100
335	SEGMENT_CATEGORY-33/48-Quick Service & Burger	3.700
336	SEGMENT_CATEGORY-34/48-Quick Service & Chicken	2.400
341	SEGMENT_CATEGORY-39/48-Quick Service & Pizza	2.200
256	YOY_SALES %3/3-{0.0460000, 0.3950000}	1.700
337	SEGMENT_CATEGORY-40/48-Quick Service & Cafe	1.600

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
252	SALES-2/3-{210.0000000, 535.0000000}	-8.400
258	UNITS-2/3-{121.0000000, 415.0000000}	-8.300
257	UNITS-1/3-{13.0000000, 121.0000000}	-8.300
251	SALES-1/3-{126.0000000, 210.0000000}	-8.300
254	YOY_SALES %1/3-{0.2120000, -0.0030000}	-6.400
260	YOY_UNITS %1/3-{0.3280000, -0.0140000}	-3.500
350	SEGMENT_CATEGORY-48/48-Variety Menu	-2.200
331	SEGMENT_CATEGORY-29/48-Mexican	-1.400
330	SEGMENT_CATEGORY-28/48-Italian/Pizza	-1.000
310	SEGMENT_CATEGORY-17/48-Fast Food	-1.000

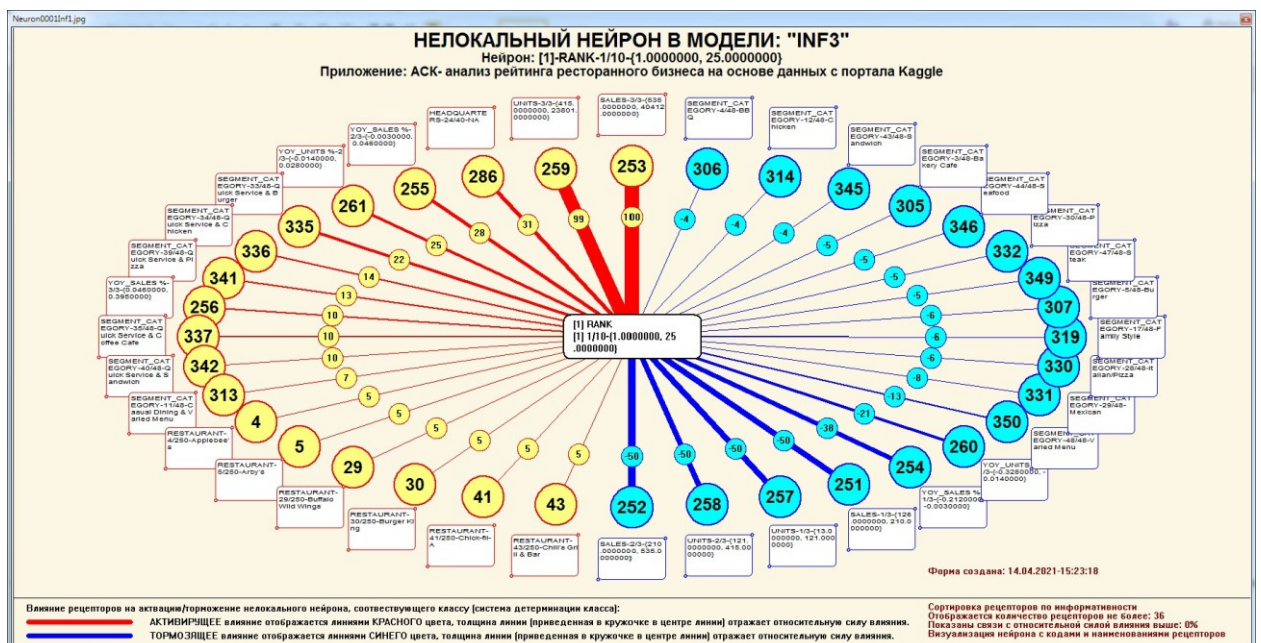
ВКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

НЕЙРОН Максимальное количество отображаемых рецепторов: 999 Минимальный вес коэф. отображаемых рецепторов: 0,000

Сортировать рецепторы: ☒ по информативности ☐ по модулю информативности

Отображать рецепторы: ☒ с наименованиями ☐ только с кодами



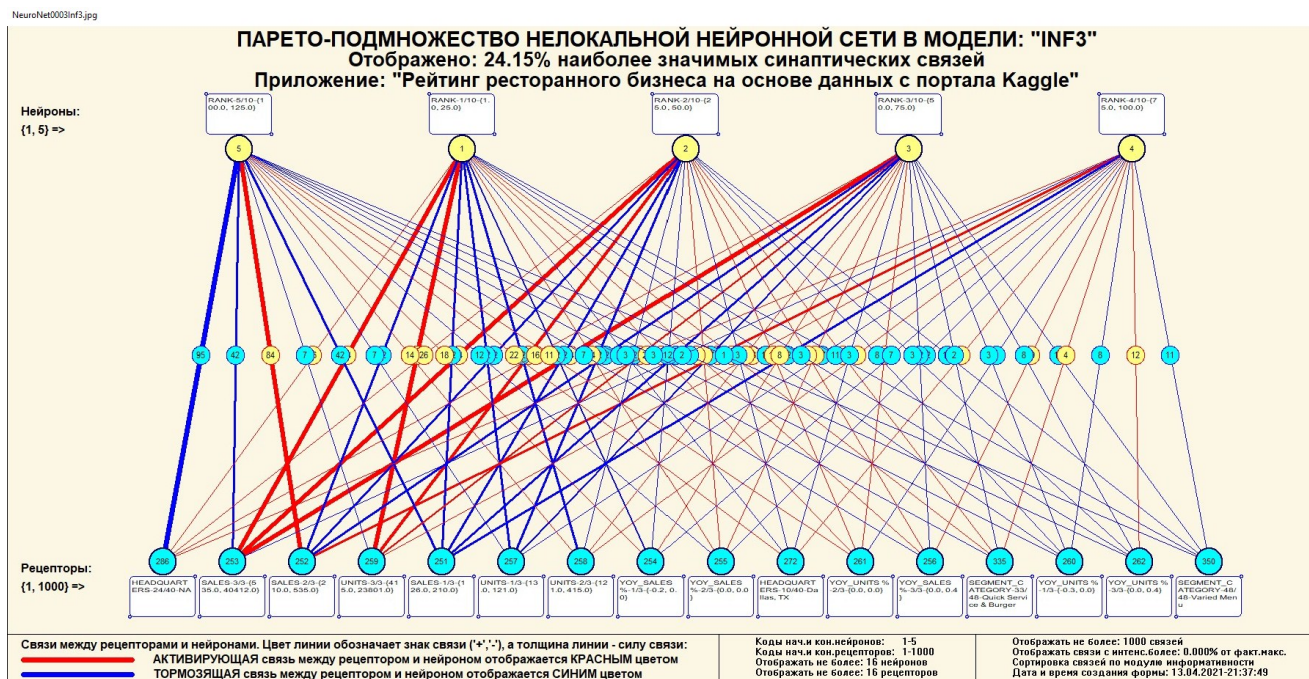
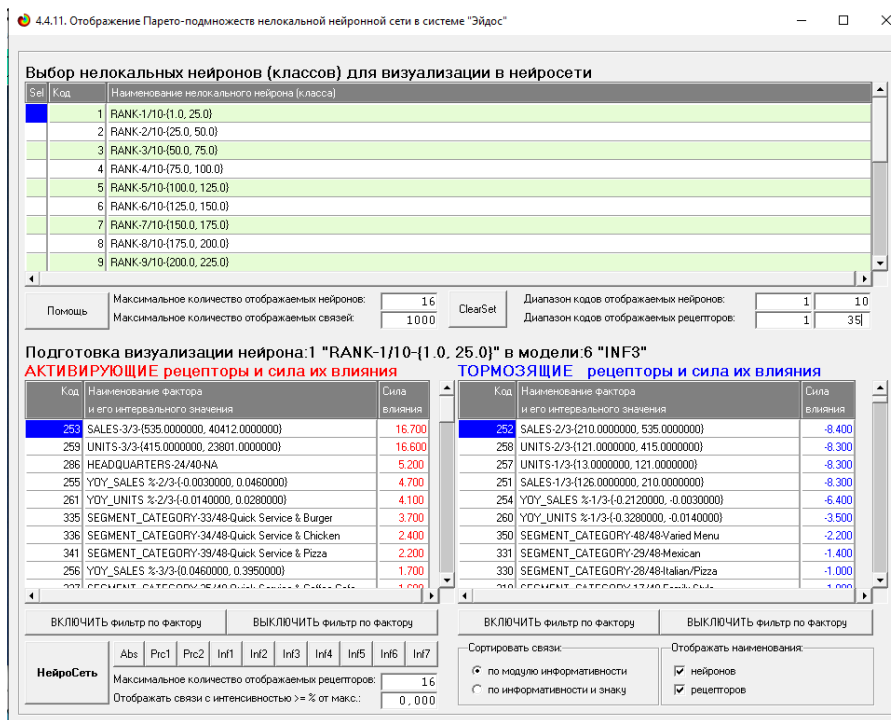


Рисунок 25 — Нелокальные нейроны и нелокальные нейронные сети

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

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 продемонстрировано построение модели зависимости уровня мирового счастья от различных характеристик стран системой искусственного интеллекта «Aidos-X++» с использованием общедоступной базы данных «Pima Indians Diabetes Database». При этом наиболее достоверной в данном приложении оказалась модель INF3. Точность модели по F-мере Ван Ризбергена составляет 0.744.

АСК-анализ 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. 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.

2. 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 [article ID]: 0020302012. — Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 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. 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.

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

6. 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.