

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) Ворожейкина Дмитрия Сергеевича

курса 2 заочной формы обучения

направления подготовки 09.04.02 «Информационные системы и технологии»

направленность(профиль) «Информационные системы и технологии в науке и управлении» (академическая магистратура)

Наименование темы «Автоматизированный системно-когнитивный анализ мирового рейтинга университетов 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	
2.	Степень полноты обзора состояния проблемы и корректность постановки цели и задач исследования	
3.	Уровень и корректность использования в работе различных методов исследований	
4.	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	
5.	Ясность, четкость, последовательность и обоснованность изложения	
6.	Применение современных технологий обработки информации	
7.	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	
8.	Ответы на вопросы при защите	

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 _____

Рецензент _____  Е.В. Луценко

« _____ » _____ 201 _____

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 мирового рейтинга
университетов на основе данных портала Kaggle**

выполнил студент группы ИТ31841
Ворожейкин Дмитрий Сергеевич

Допущена 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

_____ В.Н. Лаптев

(подпись, дата, расшифровка подписи)

Краснодар
2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГБОУ ВО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:

Зав. кафедрой КТС _____ В. И. Лойко

ЗАДАНИЕ

на курсовую работу

Студенту: ИТ31841 группы 2 курса
Факультета прикладной информатики
Специальности: 09.04.02 «Информационные системы и технологии»
(шифр)

Ворожейкину Дмитрию Сергеевичу
(Ф.И.О.)

Тема проекта: **Автоматизированный системно-когнитивный анализ
мирового рейтинга университетов на основе данных портала Kaggle**

Содержание задания: Проанализировать современные методы и средства
формирования обобщенных образов классов и решения задач идентификации
конкретных объектов с классами

Объем работы:

а) пояснительная записка к работе _____ 42 листов формата А4

б) графическая часть _____ 33 лист формата А4

Рекомендуемая литература: Луценко Е.В. Универсальная когнитивная
аналитическая система «Aidos- X++» 2017 г.

Срок выполнения проекта: с “ _____ ” _____ по “ _____ ” _____ 20__ г.

Срок защиты: “ _____ ” _____ 20__ г.

Дата выдачи задания: “ _____ ” _____ 20__ г.

Дата сдачи проекта на кафедру: “ _____ ” _____ 20__ г.

Руководитель проекта: Луценко Евгений Вениаминович, д.э.н., к.т.н.,
профессор

(подпись, Ф.И.О., звание, степень)

Задание принял студент

(подпись, дата)

Реферат

Пояснительная записка курсового проекта (работы) 42 с., 33 рис., 2 табл., 9 источников.

Ключевые слова: РЕЙТИНГ, МЕЖДУНАРОДНЫЕ ПЕРСПЕКТИВЫ, ПРОГНОЗ, AIDOS X, СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ГРАДАЦИИ, КЛАССИФИКАЦИОННЫЕ ШКАЛЫ, ОПИСАТЕЛЬНЫЕ ШКАЛЫ

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.

Оглавление

1. Синтез и верификация моделей.....	7
1.1. Описание решения	7
1.2. Подготовка и нормализация исходных данных.....	7
1.3. Синтез и верификация статистических и интеллектуальных моделей.....	16
1.4. Верификация статистических и системно-когнитивных моделей....	19
2. Решение задач идентификации в наиболее достоверной модели.....	25
2.1. Выбор наиболее достоверной модели и присвоение ей статуса текущей	25
2.2. Прогнозирование (диагностика, классификация, распознавание, идентификация)	26
2.3. Визуализация когнитивных функций	29
2.4. Нелокальные нейроны и нейронная сеть.....	35
2.5. SWOT и PEST матрицы и диаграммы.....	37
Список литературы	42

Введение

Каждый человек рано или поздно в своей жизни встает перед вопросами: куда пойдешь учиться? как сделать выбор будущего учебного заведения и факультета? В настоящее время существует множество рейтингов высших учебных заведений, изучая которые будущие студенты определяют, где они продолжат свое обучение. Зачастую на анализ тратится немало времени, ведь необходимо обработать большой объем информации, определиться с направлением обучения, расположением вузов, рейтингом, престижностью и перспективами учебных заведений.

Одним из способов решения данной проблемы становится использование систем искусственного интеллекта, которые рассматривают обучающие выборки данных.

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 анализу мирового рейтинга университетов.

Для аналитической работы системы необходимы свободный доступ к тестовым исходным данным и методика, которая поможет преобразовать эти данные в форму, которая нужна для работы в системе искусственного интеллекта. Удачным выбором является база данных тестовых задач для систем искусственного интеллекта Kaggle.

В данной курсовой работе использована база данных «World University Rankings 2020» из банка исходных данных по задачам искусственного интеллекта – Kaggle. Для решения задачи используем стандартные возможности Microsoft Office Excel, а также систему искусственного интеллекта «Aidos-X++».

1. Синтез и верификация моделей

1.1. Описание решения

Для достижения поставленной цели в данной работе предлагается применить Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Aidos».

АСК-анализ представляет собой метод искусственного интеллекта, разработанный проф. Е.В. Луценко в 2002 году для решения широкого класса задач идентификации, прогнозирования, классификации, диагностики, поддержки принятия решений и исследования моделируемой предметной области путем исследования ее модели. АСК-анализ доведен до инновационного уровня благодаря тому, что имеет свой программный инструментарий – универсальную когнитивную аналитическую систему «Aidos-X++» [2].

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. Подготовка и нормализация исходных данных

Исходные данные для данной работы взяты из базы данных «World University Rankings 2020» из банка исходных данных по задачам искусственного интеллекта – Kaggle [1]. Фрагмент данных из файла приведен на рисунке (Рисунок 1).

Данные представляют собой мировой рейтинг университетов по версии Times Higher Education.

Описание полей:

- University: Университет
- Country: Страна
- Number students: Количество студентов
- Number students per Staff: Количество студентов на одного сотрудника
- International Students: Зарубежные студенты
- Percentage Female: Процент женщин
- Percentage Male: Процент мужчин
- Teaching: Обучение
- Research: Исследование
- Industry Income: Доходы от промышленности
- International Outlook: Международные перспективы
- Overall Ranking: Общий рейтинг

Rank_ Char	Score_ Rank	University	Country	Number_ students	Numb_students_ per_Staff	International_ Students	Percentage_ Female	Percentage_ Male
1	1	University of Oxford	United Kingdom	20664	11,2	41%	46%	54%
2	2	California Institute of Technology	United States	2240	6,4	30%	34%	66%
3	3	University of Cambridge	United Kingdom	18978	10,9	37%	47%	53%
4	4	Stanford University	United States	16135	7,3	23%	43%	57%
5	5	Massachusetts Institute of Technology	United States	11247	8,6	34%	39%	61%
6	6	Princeton University	United States	7983	8,1	25%	45%	55%
7	7	Harvard University	United States	20823	9,2	24%	49%	51%
8	8	Yale University	United States	12402	5,4	20%	50%	50%
9	9	University of Chicago	United States	13833	5,7	28%	46%	54%
10	10	Imperial College London	United Kingdom	16760	11,7	56%	38%	62%
11	11	University of Pennsylvania	United States	20578	6,5	21%	52%	48%
12	12	Johns Hopkins University	United States	16171	4,5	26%	52%	48%
13	13	University of California, Berkeley	United States	41081	13,7	17%	50%	50%
13	13	ETH Zurich	Switzerland	18956	12,5	39%	32%	68%
15	14	UCL	United Kingdom	32665	10,6	52%	57%	43%
16	15	Columbia University	United States	26586	5,8	37%	0%	0%
17	16	University of California, Los Angeles	United States	41066	9,4	17%	54%	46%
18	17	University of Toronto	Canada	73370	20,1	21%	59%	41%
19	18	Cornell University	United States	22319	9,8	25%	49%	51%
20	19	Duke University	United States	15309	4,3	19%	49%	51%
21	20	University of Michigan-Ann Arbor	United States	42982	8,3	17%	49%	51%

22	21	Northwestern University	United States	17951	12,8	20%	49%	51%
23	22	Tsinghua University	China	38783	12	11%	34%	66%
24	23	Peking University	China	39575	11,5	15%	46%	54%
25	24	National University of Singapore	Singapore	30869	17,9	28%	51%	49%
26	25	University of Washington	United States	45692	11,1	16%	53%	47%
27	26	Carnegie Mellon University	United States	13430	13,6	48%	41%	59%
27	26	London School of Economics and Political Science	United Kingdom	10570	12,1	71%	53%	47%
29	27	New York University	United States	44466	8,9	33%	57%	43%
30	28	University of Edinburgh	United Kingdom	29433	12,8	41%	60%	40%
31	29	University of California, San Diego	United States	33579	13	23%	46%	54%
32	30	LMU Munich	Germany	34519	34,7	17%	61%	39%
32	30	University of Melbourne	Australia	47385	26,3	46%	56%	44%

Рисунок 1 – Фрагмент данных из базы данных «World University Rankings 2020»

С исходными данными были выполнены следующие преобразования:

удалены столбцы: Rank_Char, Score_Rank, Citations, Score_Result, т.к.

они несущественны для результата анализа

- добавлено поле Number
- файл сохранен с именем inp_data.xls.

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

Фрагмент подготовленной и нормализованной выборке приведен на рисунке (Рисунок 2).

Number	University	Country	Number_students	Numb_students_per_Staff	International_Students	Percentage_Female	Percentage_Male
1	University of Oxford	United Kingdom	20664	11,2	41%	46%	54%
2	California Institute of Technology	United States	2240	6,4	30%	34%	66%
3	University of Cambridge	United Kingdom	18978	10,9	37%	47%	53%
4	Stanford University	United States	16135	7,3	23%	43%	57%
5	Massachusetts Institute of Technology	United States	11247	8,6	34%	39%	61%
6	Princeton University	United States	7983	8,1	25%	45%	55%
7	Harvard University	United States	20823	9,2	24%	49%	51%
8	Yale University	United States	12402	5,4	20%	50%	50%
9	University of Chicago	United States	13833	5,7	28%	46%	54%
10	Imperial College London	United Kingdom	16760	11,7	56%	38%	62%
11	University of Pennsylvania	United States	20578	6,5	21%	52%	48%
12	Johns Hopkins University	United States	16171	4,5	26%	52%	48%
13	University of California, Berkeley	United States	41081	13,7	17%	50%	50%
14	ETH Zurich	Switzerland	18956	12,5	39%	32%	68%
15	UCL	United Kingdom	32665	10,6	52%	57%	43%
16	Columbia University	United States	26586	5,8	37%	0%	0%
17	University of California, Los Angeles	United States	41066	9,4	17%	54%	46%
18	University of Toronto	Canada	73370	20,1	21%	59%	41%
19	Cornell University	United States	22319	9,8	25%	49%	51%

20	Duke University	United States	15309	4,3	19%	49%	51%
21	University of Michigan-Ann Arbor	United States	42982	8,3	17%	49%	51%
22	Northwestern University	United States	17951	12,8	20%	49%	51%
23	Tsinghua University	China	38783	12	11%	34%	66%
24	Peking University	China	39575	11,5	15%	46%	54%
25	National University of Singapore	Singapore	30869	17,9	28%	51%	49%
26	University of Washington	United States	45692	11,1	16%	53%	47%
27	Carnegie Mellon University	United States	13430	13,6	48%	41%	59%
28	London School of Economics and Political Science	United Kingdom	10570	12,1	71%	53%	47%
29	New York University	United States	44466	8,9	33%	57%	43%
30	University of Edinburgh	United Kingdom	29433	12,8	41%	60%	40%
31	University of California, San Diego	United States	33579	13	23%	46%	54%
32	LMU Munich	Germany	34519	34,7	17%	61%	39%
33	University of Melbourne	Australia	47385	26,3	46%	56%	44%
34	University of British Columbia	Canada	52108	18,1	32%	55%	45%
35	University of Hong Kong	Hong Kong	18260	18,2	44%	54%	46%
36	King's College London	United Kingdom	26057	12,5	44%	62%	38%
37	The University of Tokyo	Japan	25913	10,6	12%	0%	0%
38	École Polytechnique Fédérale de Lausanne	Switzerland	10492	11,4	59%	28%	72%
39	Georgia Institute of Technology	United States	21372	18,3	27%	33%	67%
40	University of Texas at Austin	United States	49165	17,2	10%	52%	48%
41	Karolinska Institute	Sweden	7777	9	20%	68%	32%
42	McGill University	Canada	31294	13	29%	58%	42%
43	Technical University of Munich	Germany	40472	47,8	27%	34%	66%
44	Heidelberg University	Germany	20800	16,1	18%	54%	46%
45	KU Leuven	Belgium	45049	36,3	15%	50%	50%

Рисунок 2 – Фрагмент данных обучающей выборки

Классификационные шкалы следующие: Университет, Страна. Количество студентов, количество студентов на одного сотрудника, зарубежные студенты, процент женщин, процент мужчин, обучение, исследование, доходы от промышленности, международные перспективы, общий рейтинг – описательные.

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.

Для загрузки базы исходных данных в систему «Aidos» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, т.е. режимом 2.3.2.2.

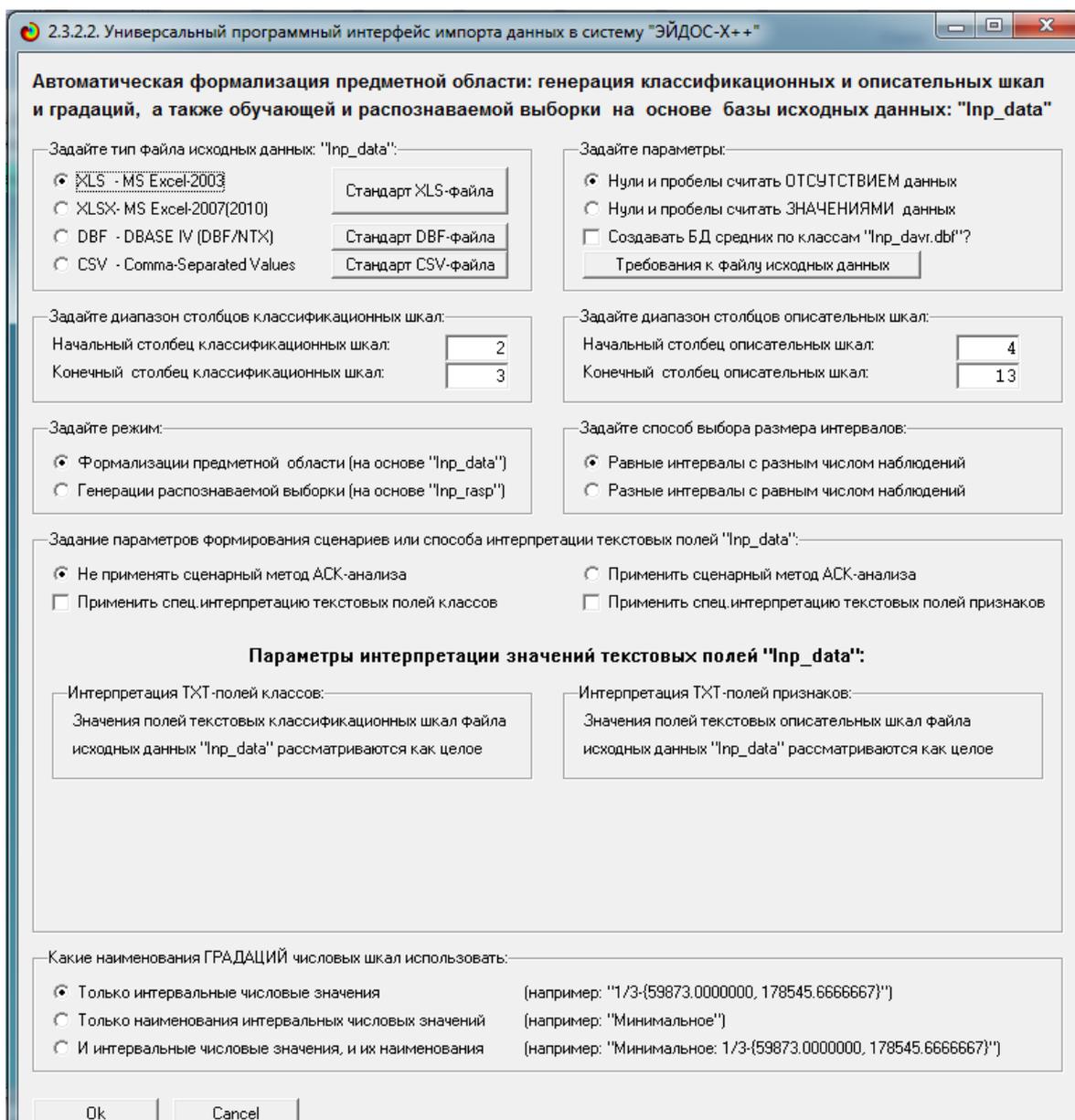


Рисунок 3 – Экранная форма универсального программного интерфейса импорта данных в систему "Aidos" (режим 2.3.2.2)

В экранной форме, приведенной на рисунке (Рисунок 3), задаем следующие настройки:

- «Задайте тип файла исходных данных «Inp_data»: «XLS – MS Excel- 2003»;
- «Задайте диапазон столбцов классификационных шкал»: «Начальный столбец классификационных шкал» – 2; «Конечный столбец классификационных шкал» – 3;

- «Задать диапазон столбцов описательных шкал»: «Начальный столбец описательных шкал» – 4, «Конечный столбец описательных шкал» – 13;
- «Задание параметров формирования сценариев или способа интерпретации текстовых полей»: «Не применять сценарный метод АСК- анализа».

После заполнения необходимых данных, нажать кнопку «Ок». Открывшаяся Открывшаяся экранная форма (

Рисунок 4) показывает сколько текстовых и числовых классификационных и описательных шкал система обнаружила при заданных параметрах и сколько в них обнаружено градаций.

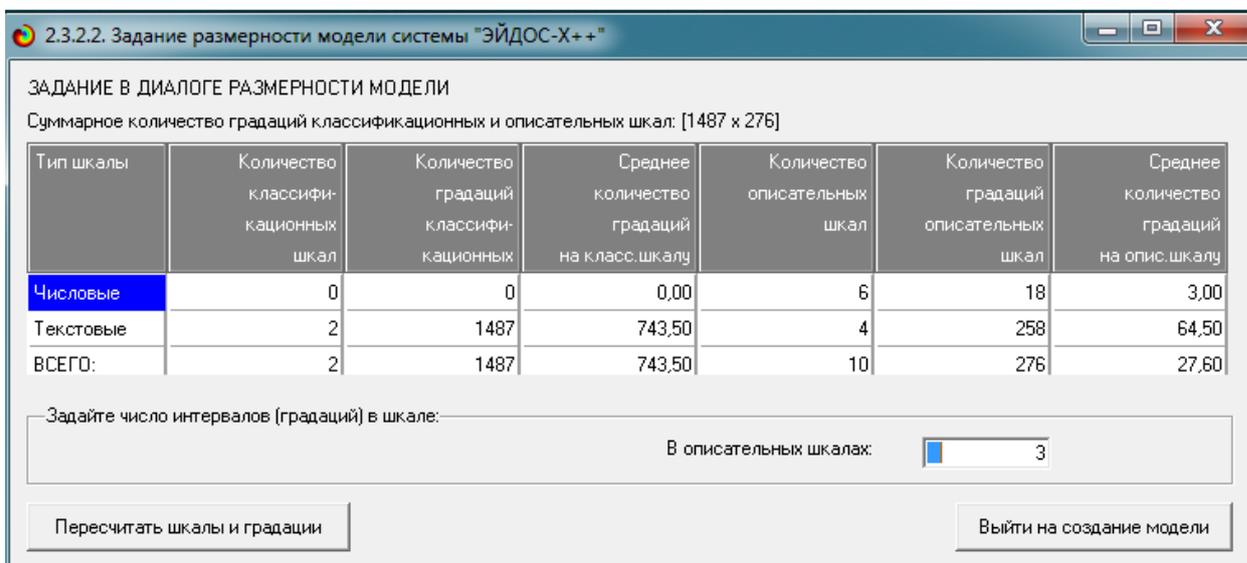


Рисунок 4 – Экранная форма внутреннего калькулятора режима 2.3.2.2

Обратим внимание на то, что заданы адаптивные интервалы, учитывающее неравномерность распределения данных по диапазону значений, что важно при относительно небольшом числе наблюдений. Если бы интервалы были заданы равными по величине, то в них бы учитывалось сильно отличающееся число наблюдений, а в некоторых интервалах их бы могло не оказаться вовсе.

Если обнаружены шкалы числового типа, то появляется возможность задать количество интервальных числовых значений в этих шкалах. Это делается отдельно для классификационных и описательных шкал, таким образом, число интервальных числовых значений в классификационных и описательных шкалах может отличаться. В результате выполнения данного режима формируются классификационные, описательные шкалы, градации и обучающая выборка.

После нажатия на кнопку «Выйти на создание модели» запускается процесс процесс импорта данных из внешней БД «Inp_data.xlsx» в систему «Aidos» (

Рисунок 5), а также прогноз времени завершения этого процесса.

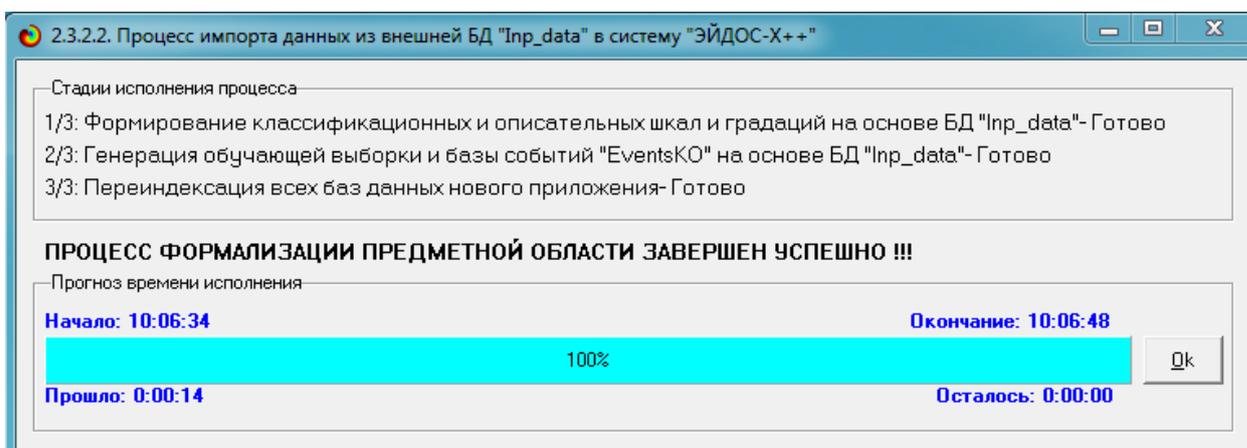


Рисунок 5 – Экранная форма импорта данных режима 2.3.2.2

В результате работы режима сформировано 1 классификационная шкала с шкала с суммарным количеством градаций (классов): 2 (

Таблица 1) и 7 описательных шкал с суммарным числом градаций:67 (

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы	DEL
1	UNIVERSITY	1	?Shahid Chamran University of Ahvaz	
2	COUNTRY	2	Aalborg University	
		3	Aalto University	
		4	Aarhus University	
		5	Aberystwyth University	
		6	Ac?badem University	
		7	Acharya Nagarjuna University	
		8	Adam Mickiewicz University, Pozna?	
		9	AGH University of Science and Technology	
		10	Aichi Medical University	
		11	Ain Shams University	
		12	Aix-Marseille University	
		13	Ajou University	
		14	Akdeniz University	
		15	Akita University	
		16	Al-Azhar University	
		17	Al-Balqa Applied University	
		18	Al-Farabi Kazakh National University	
		19	Alexandria University	
		20	Alexandru Ioan Cuza University	
		21	Alfaisal University	
		22	Aligarh Muslim University	

Таблица 2).

С использованием классификационных и описательных шкал и градаций исходные данные (Рисунок 2) были закодированы и в результате получена обучающая выборка. Для просмотра классификационных шкал и градаций используется режим 2.1, для просмотра описательных шкал и градаций используется режим 2.2

Таблица 1– Классификационные шкалы и градации

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

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы	DEL
1	UNIVERSITY	1	?Shahid Chamran University of Ahvaz	
2	COUNTRY	2	Aalborg University	
		3	Aalto University	
		4	Aarhus University	
		5	Aberystwyth University	
		6	Ac?badem University	
		7	Acharya Nagarjuna University	
		8	Adam Mickiewicz University, Pozna?	
		9	AGH University of Science and Technology	
		10	Aichi Medical University	
		11	Ain Shams University	
		12	Aix-Marseille University	
		13	Ajou University	
		14	Akdeniz University	
		15	Akita University	
		16	Al-Azhar University	
		17	Al-Balqa Applied University	
		18	Al-Farabi Kazakh National University	
		19	Alexandria University	
		20	Alexandru Ioan Cuza University	
		21	Alfaisal University	
		22	Aligarh Muslim University	

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

Таблица 2 – Описательные шкалы и градации

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

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	NUMBER_STUDENTS	1	1/3-{558.0000000, 277073.3333333}
2	NUMB_STUDENTS_PER_STAFF	2	2/3-{277073.3333333, 553588.6666667}
3	INTERNATIONAL_STUDENTS	3	3/3-{553588.6666667, 830104.0000000}
4	PERCENTAGE_FEMALE		
5	PERCENTAGE_MALE		
6	TEACHING		
7	RESEARCH		
8	INDUSTRY_INCOME		
9	INTERNATIONAL_OUTLOOK		
10	OVERALL_RANKING		

Помощь Доб. шкалу Доб. град. шкалы Копир. шкалу Копир. град. шкалы Копир. шкалу с град. Удал. шкалу с град. Удал. град. шкалы Перекодировать Очистить

Обучающая выборка, по сути, представляет собой нормализованную с помощью классификационных и описательных шкал и градаций базу

исходных данных. 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 [3]).

1.3. Синтез и верификация статистических и интеллектуальных моделей

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) моделей осуществляется в режиме 3.5 системы «Aidos» (Рисунок 6).

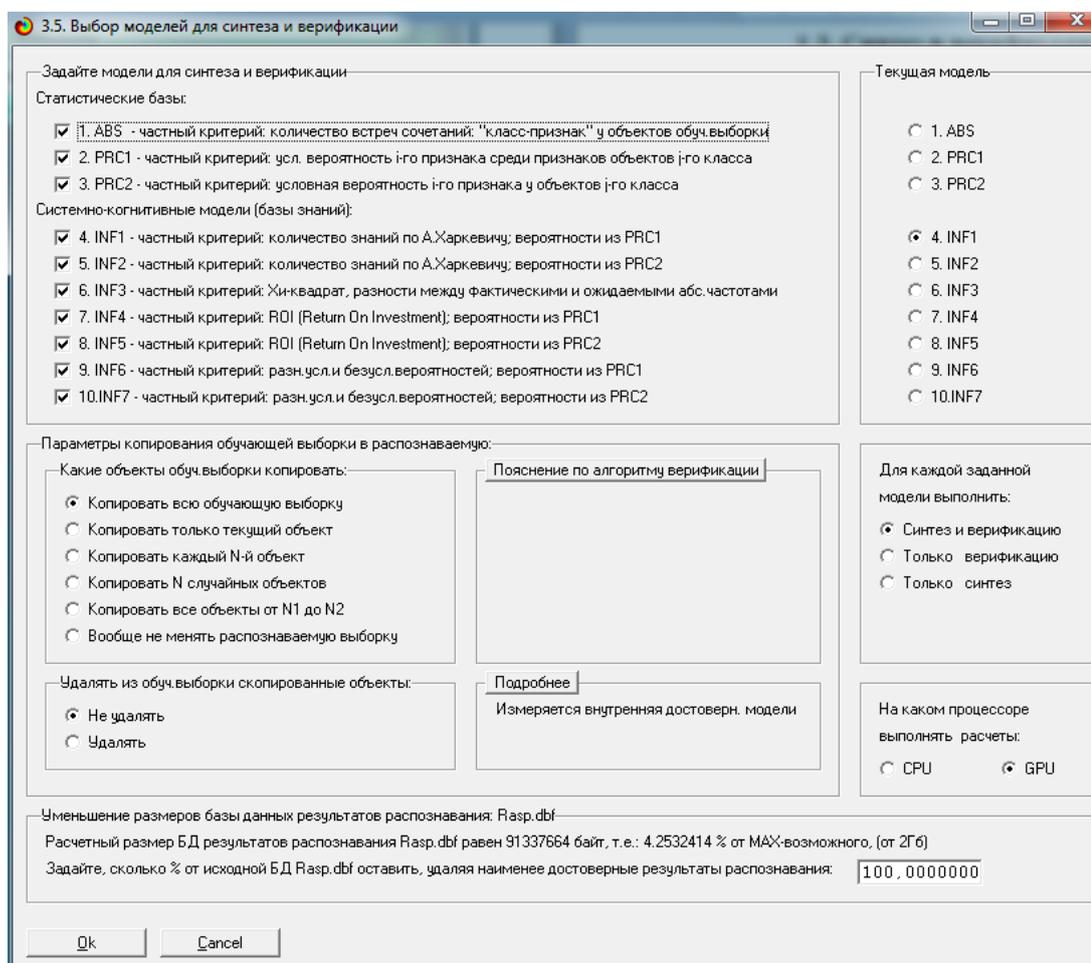


Рисунок 6 – Экранная форма выбора моделей для синтеза и верификации режима 3.5

Необходимо отметить, что синтез и верификация всех моделей на ресурсах заняли 8 часов 25 минут 05 секунд (

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

Код признака	Наименование описательной шкалы и градации	72. UNIVERSITY BANDUNG INSTITUTE OF TECHNOLOGY (ITB)	73. UNIVERSITY BANGOR UNIVERSITY	74. UNIVERSITY BAR ILAN UNIVERSITY	75. UNIVERSITY BASHKIR STATE UNIVERSITY	76. UNIVERSITY BAUMAN MOSCOW STATE TECHNICAL UNIVERSITY	77. UNIVERSITY BAYLOR UNIVERSITY	78. UNIVERSITY BEIHANG UNIVERSITY	79. UNIVERSITY BEIJING INSTITUTE OF TECHNOLOGY	80. UNIVERSITY BEIJING JIAOTONG UNIVERSITY	81. UNIVERSITY BEIJING NORMAL UNIVERSITY	82. UNIVERSITY BELARUSIAN STATE UNIVERSITY
1	NUMBER_STUDENTS-1/3-(598.0000000, 277073.333...	1						1				
2	NUMBER_STUDENTS-2/3-(277073.3333333, 553588...											
3	NUMBER_STUDENTS-3/3-(553588.6666667, 830104...											
4	NUMB_STUDENTS_PER_STAFF-1/3-(0.9000000, 165...	1						1				
5	NUMB_STUDENTS_PER_STAFF-2/3-(165.1000000, 3...											
6	NUMB_STUDENTS_PER_STAFF-3/3-(329.3000000, 4...											
7	INTERNATIONAL_STUDENTS-%											
8	INTERNATIONAL_STUDENTS-0%											
9	INTERNATIONAL_STUDENTS-1%											
10	INTERNATIONAL_STUDENTS-10%											
11	INTERNATIONAL_STUDENTS-11%											
12	INTERNATIONAL_STUDENTS-12%											
13	INTERNATIONAL_STUDENTS-13%											
14	INTERNATIONAL_STUDENTS-14%											
15	INTERNATIONAL_STUDENTS-15%											
16	INTERNATIONAL_STUDENTS-16%											
17	INTERNATIONAL_STUDENTS-17%											

Рисунок 8 – Матрица абсолютных частот (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	258. UNIVERSITY GDA'YSK UNIVERSITY OF TECHNOLOGY	259. UNIVERSITY GEBZE TECHNICAL UNIVERSITY	260. UNIVERSITY GEORGE MASON UNIVERSITY	261. UNIVERSITY GEORGE WASHINGTON UNIVERSITY	262. UNIVERSITY GEORGETOWN UNIVERSITY	263. UNIVERSITY GEORGIA INSTITUTE OF TECHNOLOGY	264. UNIVERSITY GEORGIA STATE UNIVERSITY	265. UNIVERSITY GHENT UNIVERSITY	266. UNIVERSITY GIFU UNIVERSITY	267. UNIVERSITY GITAM UNIVERSITY	268. UNIVERSITY GLAZGOW UNIVERSITY
25	INTERNATIONAL_STUDENTS-24%											
26	INTERNATIONAL_STUDENTS-25%											
27	INTERNATIONAL_STUDENTS-26%											
28	INTERNATIONAL_STUDENTS-27%											
29	INTERNATIONAL_STUDENTS-28%						5.164					
30	INTERNATIONAL_STUDENTS-29%											
31	INTERNATIONAL_STUDENTS-3%											
32	INTERNATIONAL_STUDENTS-30%											
33	INTERNATIONAL_STUDENTS-31%											
34	INTERNATIONAL_STUDENTS-32%											
35	INTERNATIONAL_STUDENTS-33%											
36	INTERNATIONAL_STUDENTS-34%											
37	INTERNATIONAL_STUDENTS-35%											
38	INTERNATIONAL_STUDENTS-36%											
39	INTERNATIONAL_STUDENTS-37%											
40	INTERNATIONAL_STUDENTS-38%											
41	INTERNATIONAL_STUDENTS-39%											

Рисунок 9 – Матрица информативностей INF1 (фрагмент)

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

Код признака	Наименование описательной шкалы и градации	306. UNIVERSITY HUMBOLDT UNIVERSITY OF BERLIN	307. UNIVERSITY HUNAN UNIVERSITY	308. UNIVERSITY HYOGO COLLEGE OF MEDICINE	309. UNIVERSITY I AKADEMI UNIVERSITY	310. UNIVERSITY I CENTRALE DE LYON	311. UNIVERSITY I DES MINES DE SAINT I	312. UNIVERSITY I DES PONTS PARISTECH	313. UNIVERSITY I NORMALE SUPI DE LYON	314. UNIVERSITY I NORMALE SUPI PARIS SACLAY	315. UNIVERSITY I POLYTECHNIQUE
1	NUMBER_STUDENTS-1/3-(598.0000000, 277073...			-0.013		-0.013	-0.013	-0.013	-0.013	-0.013	0.088
2	NUMBER_STUDENTS-2/3-(277073.3333333, 553...			-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
3	NUMBER_STUDENTS-3/3-(553588.6666667, 830...			-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
4	NUMB_STUDENTS_PER_STAFF-1/3-(0.9000000, 1...			-0.014		-0.014	-0.014	-0.014	-0.014	-0.014	0.087
5	NUMB_STUDENTS_PER_STAFF-2/3-(165.100000...										
6	NUMB_STUDENTS_PER_STAFF-3/3-(329.300000...			-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
7	INTERNATIONAL_STUDENTS-%			-0.001		-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
8	INTERNATIONAL_STUDENTS-0%			-0.074		-0.074	-0.074	-0.074	-0.074	-0.074	-0.067
9	INTERNATIONAL_STUDENTS-1%			0.892		-0.108	-0.108	-0.108	-0.108	-0.108	-0.098
10	INTERNATIONAL_STUDENTS-10%			-0.027		-0.027	-0.027	-0.027	-0.027	-0.027	-0.025
11	INTERNATIONAL_STUDENTS-11%			-0.039		-0.039	-0.039	-0.039	-0.039	-0.039	-0.035
12	INTERNATIONAL_STUDENTS-12%			-0.031		-0.031	-0.031	-0.031	-0.031	-0.031	-0.028
13	INTERNATIONAL_STUDENTS-13%			-0.025		-0.025	-0.025	-0.025	0.975	-0.025	-0.023
14	INTERNATIONAL_STUDENTS-14%			-0.020		-0.020	-0.020	-0.020	-0.020	0.980	-0.018
15	INTERNATIONAL_STUDENTS-15%			-0.025		-0.025	-0.025	-0.025	-0.025	-0.025	-0.023
16	INTERNATIONAL_STUDENTS-16%			-0.020		-0.020	-0.020	-0.020	-0.020	-0.020	-0.018
17	INTERNATIONAL_STUDENTS-17%			-0.025		-0.025	-0.025	-0.025	-0.025	-0.025	-0.023

Рисунок 10 – Модель INF3 (фрагмент)

Отметим, что в АСК-анализе и СК-моделях степень выраженности различных свойств объектов наблюдения рассматривается с единственной точки зрения: с точки зрения того, какое количество информации содержится в них о том, к каким обобщающим категориям (классам) будут принадлежать или не принадлежать эти объекты. Поэтому не играет никакой роли в каких единицах измерения измеряются те или иные свойства объектов наблюдения, а также в каких единицах измеряются результаты влияния этих свойств, натуральных, в процентах или стоимостных [4]. Это и есть решение проблемы сопоставимости в АСК-анализе и системе «Aidos», отличающее их от других интеллектуальных технологий.

1.4. Верификация статистических и системно-когнитивных моделей

Оценка достоверности моделей в системе «Aidos» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1-L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры [5]. В режиме 3.4 системы «Aidos» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности.

Из рисунка (Рисунок 11) мы видим, что в данном интеллектуальном приложении по F-критерию Ван Ризбергена наиболее достоверной является СК-модель INF3 с интегральным критерием «Сумма знаний» ($F=0,019$ при максимуме 1), по критерию L1 проф.Е.В.Луценко наиболее достоверной также является СК-модель INF3 с интегральным критерием «Сумма знаний» ($L1=0,159$ при максимуме 1).

Наименование модели и частного критерия	Интегральный критерий	ложноположительный (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модулей уровней сход. истинно-полож. решений (STP)	Сумма модулей уровней сход. истинно-отриц. решений (STN)	Сумма модулей ложнополож. решений (SFP)	Сумма модулей ложноотриц. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В.Луценко
1. ABS - частный критерий: количество встреч сонетаний: "клас...	Корреляция абс частот с обр...	52	0.001	0.966	0.002	0.000		*****.***			1.000	
1. ABS - частный критерий: количество встреч сонетаний: "клас...	Сумма абс частот по признак...		0.005	1.000	0.010	498.016		10943.688		0.044	1.000	0.083
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн частот с о...	52	0.001	0.966	0.002	0.000		*****.***			1.000	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн частот по приз...		0.005	1.000	0.010	926.373		141530.554		0.007	1.000	0.013
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн частот с о...	52	0.001	0.966	0.002	0.000		*****.***			1.000	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн частот по приз...		0.005	1.000	0.010	913.795		139946.041		0.006	1.000	0.013
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	1351	0.000	0.115	0.000			*****.***				
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	85	0.005	0.944	0.010	342.398	1134.002	15605.703	4.115	0.021	0.988	0.042
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	1350	0.000	0.116	0.000			*****.***				
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	88	0.005	0.942	0.010	353.460	1189.961	16335.938	4.415	0.021	0.988	0.041
6. INF3 - частный критерий: Хинкварт, разности между фактик...	Семантический резонанс зна...	883	0.000	0.422	0.001	0.000	0.000	*****.***			1.000	
6. INF3 - частный критерий: Хинкварт, разности между фактик...	Сумма знаний	132	0.010	0.914	0.019	335.279	3851.484	3537.121	20.916	0.087	0.941	0.159
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	1348	0.000	0.117	0.000			*****.***				
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	43	0.005	0.972	0.010	23.840	4.238	410.228	0.011	0.055	1.000	0.104
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	1348	0.000	0.117	0.000			*****.***				
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	42	0.005	0.972	0.010	23.383	4.219	404.247	0.012	0.055	1.000	0.104
9. INF6 - частный критерий: разн усл и безул.вероятностей; вер...	Семантический резонанс зна...	952	0.000	0.377	0.001	0.000		*****.***			1.000	
9. INF6 - частный критерий: разн усл и безул.вероятностей; вер...	Сумма знаний	132	0.005	0.914	0.010	278.834	1172.643	28931.455	3.895	0.010	0.986	0.019
10. INF7 - частный критерий: разн усл и безул.вероятностей; ве...	Семантический резонанс зна...	900	0.000	0.411	0.001	0.000	0.000	*****.***			1.000	
10. INF7 - частный критерий: разн усл и безул.вероятностей; ве...	Сумма знаний	163	0.005	0.893	0.009	291.018	1323.595	30704.352	5.524	0.009	0.981	0.019

Рисунок 11 – Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф.Е.В.Луценко (фрагмент)

Хотелось бы обратить внимание, что статистические модели, как правило, дают более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, и практически никогда – более высокую. Этим и оправдано применение моделей знаний и интеллектуальных технологий.

Это позволяет сделать обоснованные выводы о том, что:

- причинно-следственная зависимость адекватно отражена в системно-когнитивных моделях INF3 и INF5;
- так как L2-мера проф. Е.В. Луценко для модели INF5 составляет 0,104, а для INF3 – 0,159, то предпочтение отдаем модели INF3.

На рисунке (Рисунок 12) приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам оценки данных датчиков в СК-модели INF3 по данным обучающей выборки.



Рисунок 12 – Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели INF3

На рисунке (Рисунок 13) изображены частотные распределения разностей долей разностей уровней сходимости решений в СК-модели INF3.

Как можно видеть, при модели INF3 непринадлежность объекта к классу лучше определяется, чем принадлежность. При этом существует вероятность ложного определения непринадлежности или принадлежности.

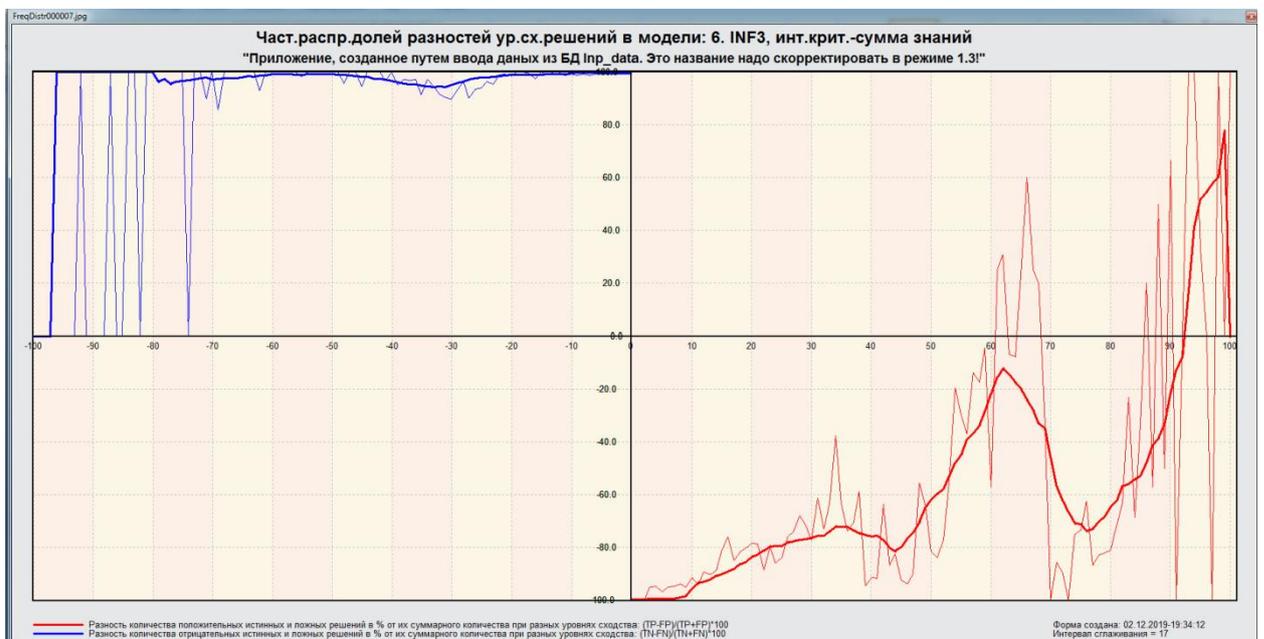


Рисунок 13 – Частотные распределения разностей долей разностей уровней сходимости решений в СК-модели INF3

- видно, что для отрицательных решений количество истинных решений всегда значительно превосходит количество ложных решений;
- также видно, что для положительных решений при уровнях сходства от 0% до примерно 62% есть и истинные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства. При уровнях сходства выше 62% ложно-положительных решений практически не встречается;
- модель INF3 лучше определяет непринадлежность объекта к классу, чем принадлежность (процент правильной неидентификации выше).

Любые данные о наблюдениях можно считать суммой истинного значения и шума, причем ни первое, ни второе неизвестны. Поэтому имеет смысл сравнить созданные модели с чисто случайными моделями, совпадающими по основным характеристикам. В системе «Aidos» есть лабораторная работа № 2.01: «Исследование RND-модели при различных объемах выборки». Если данная работа устанавливается при отсутствии текущего приложения, то все параметры создаваемых моделей задаются вручную, если же текущая модель существует, как в нашем случае, то все основные ее параметры определяются автоматически (Рисунок 14).

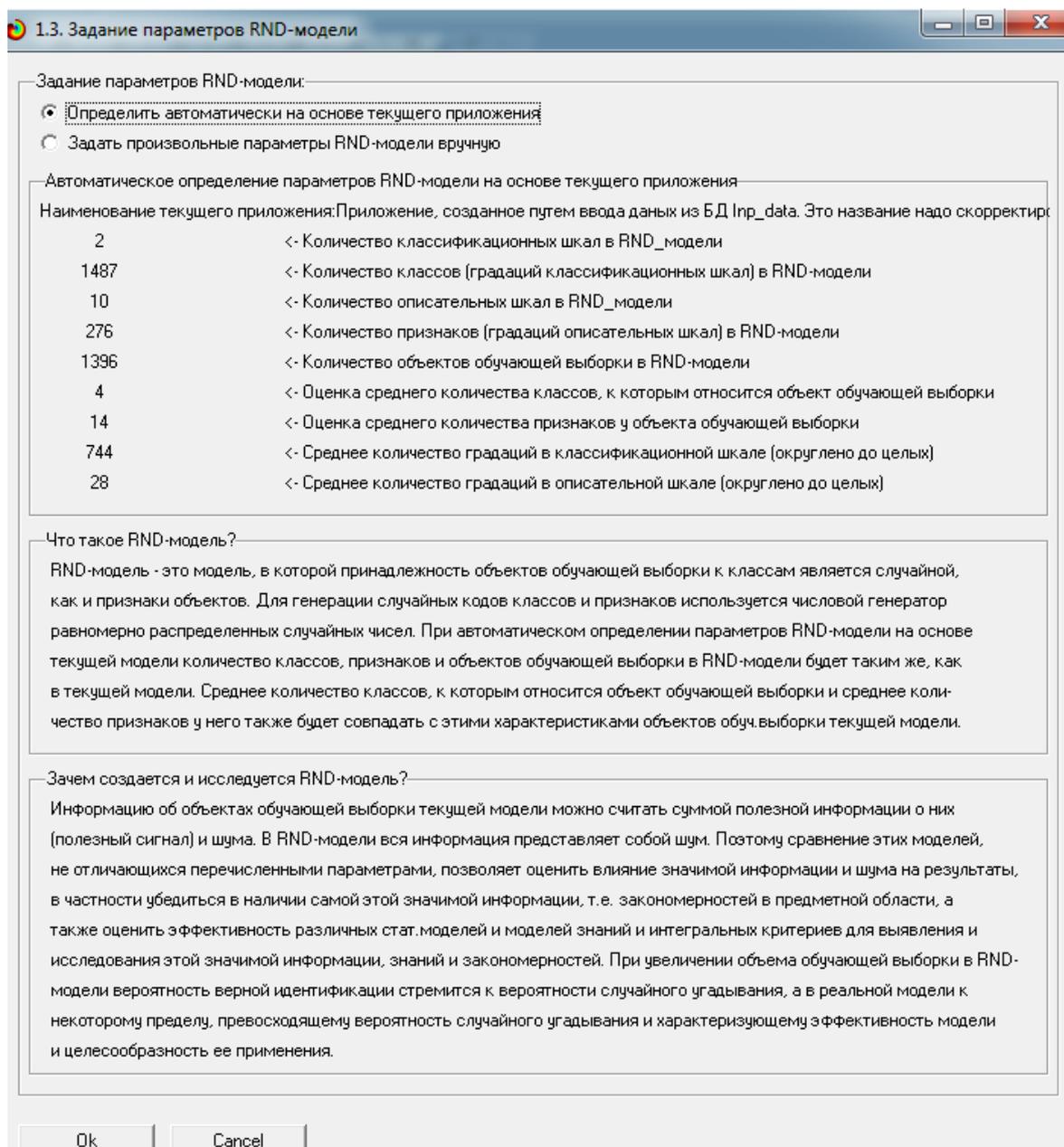


Рисунок 14 – Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке (Рисунок 16) показано частотное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF3. Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели (Рисунок 15) объясняется тем, что в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной

стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны. Если же такой информации в модели нет, то и распределение получается типа, приведенного на рисунке (Рисунок 16).

413.6. Обобщенная форма по достов. моделям при разн. крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	ть	Полного модели	F-мера Ван Ризбергера	Сумма модул. уровней сход. истинно полож. решений (STP)	Сумма модул. уровней сход. истинно отриц. решений (STM)	Сумма модул. уровней сход. ложно полож. решений (SFP)	Сумма модул. уровней сход. ложно отрицат. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Луценко	Средняя модуль. уровнев. сход. истинно полож. решения
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	0.007	0.808	0.014	2860.000							
1. ABS - частный критерий: количество встреч сочетаний: "клас...	Сумма абс. частот по призна...	0.003	1.000	0.006	22.000	980.385		0.022	1.000	0.043	0.846	
2. FRC1 - частный критерий: усл. вероятность иго признака сред...	Корреляция усл. отн. частот с о...	0.007	0.808	0.014	2860.000							
2. FRC1 - частный критерий: усл. вероятность иго признака сред...	Сумма усл. отн. частот по при...	0.003	1.000	0.006	9.962	493.290		0.020	1.000	0.039	0.383	
3. FRC2 - частный критерий: условная вероятность иго признака...	Корреляция усл. отн. частот с о...	0.007	0.808	0.014	2860.000							
3. FRC2 - частный критерий: условная вероятность иго признака...	Сумма усл. отн. частот по при...	0.003	1.000	0.006	9.763	483.260		0.020	1.000	0.039	0.375	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.007	0.769	0.014	2860.000							
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.003	1.000	0.006	17.581	844.386		0.020	1.000	0.040	0.676	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	0.007	0.769	0.014	2860.000							
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	0.003	1.000	0.006	17.486	839.239		0.020	1.000	0.040	0.673	
6. INF3 - частный критерий: Хи-квадрат: разности между фактик...	Семантический резонанс зна...	0.008	0.846	0.015	2860.000							
6. INF3 - частный критерий: Хи-квадрат: разности между фактик...	Сумма знаний	0.004	1.000	0.008	21.821	487.934	466.114	0.045	1.000	0.086	0.839	
7. INF4 - частный критерий: ROI (Return On Investment); вероатно...	Семантический резонанс зна...	0.008	0.846	0.015	2860.000							
7. INF4 - частный критерий: ROI (Return On Investment); вероатно...	Сумма знаний	0.003	1.000	0.006	8.965	436.031		0.020	1.000	0.039	0.345	
8. INF5 - частный критерий: ROI (Return On Investment); вероатно...	Семантический резонанс зна...	0.008	0.846	0.015	2860.000							
8. INF5 - частный критерий: ROI (Return On Investment); вероатно...	Сумма знаний	0.003	1.000	0.006	8.784	426.994		0.020	1.000	0.040	0.338	
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	0.008	0.846	0.015	2860.000							
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	0.003	1.000	0.006	9.450	468.819		0.020	1.000	0.039	0.363	
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	0.008	0.846	0.015	2860.000							
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	0.003	1.000	0.006	9.256	459.047		0.020	1.000	0.039	0.356	

Рисунок 15 – Экранная форма с информацией о достоверности RND-моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф. Е.В. Луценко (фрагмент)

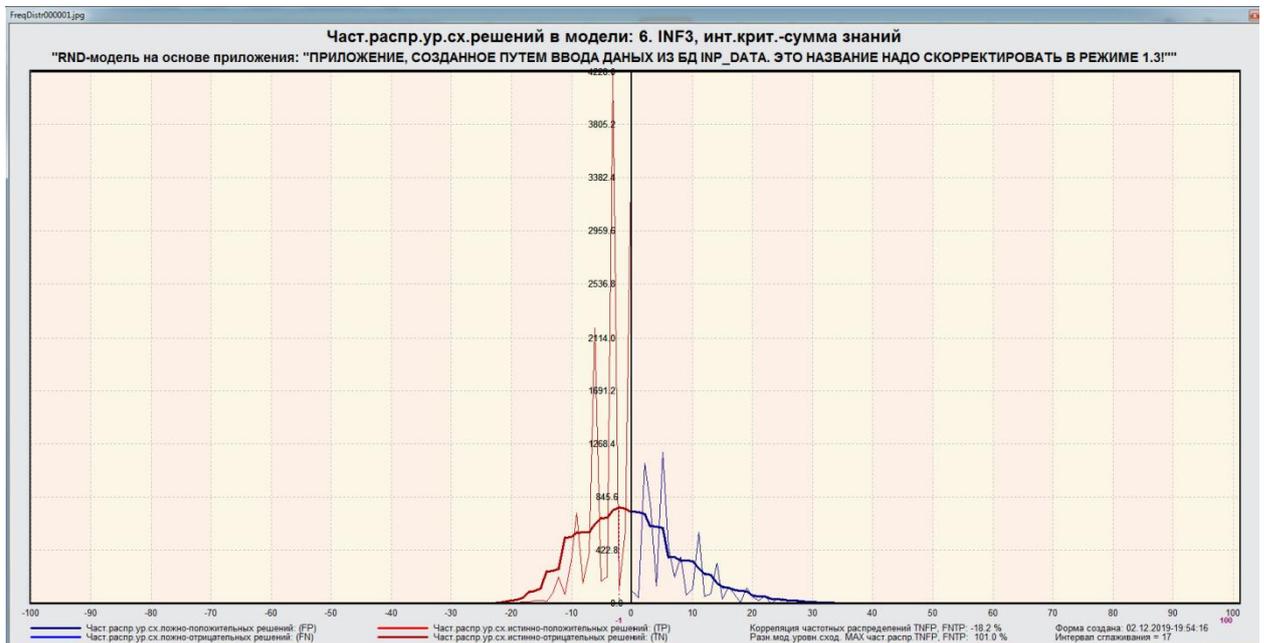


Рисунок 16 – Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в RND-модели INF3

На основе его сравнения с рисунком (Рисунок 12) можно сделать следующие выводы:

- RND-модель дает ложно-положительные и ложно-отрицательные прогнозы практически на всем диапазоне уровней сходства-различия;

- достоверность лучшей модели INF3, отражающей реальный объект моделирования выше, чем аналогичной случайной модели;
- различие между достоверностью статистических моделей и моделей знаний, созданных на основе случайной выборки, значительно меньше, чем у моделей, отражающих реальный объект моделирования;
- в реальных моделях кроме шума есть также и информация об истинных причинно-следственных взаимосвязях факторов и их значений с одной стороны, и состояниями объекта моделирования, которые ими обуславливаются, с другой стороны, причем примерно 1/3 достоверности обусловлена отражением в реальных моделях закономерностей предметной области, а 2/3 достоверности обусловлено наличием шума в исходных данных.

На основании этого можно предположить, что в исходных данных уровень сигнала о реальных причинно-следственных связях в моделируемой предметной области примерно в два раза ниже уровня шума[6].

2. Решение задач идентификации в наиболее достоверной модели

2.1. Выбор наиболее достоверной модели и присвоение ей статуса текущей

В соответствии со схемой обработки данных, информации и знаний в системе «Aidos», присвоим СК-модели INF3 статус текущей модели. Для этого запустим режим 5.6 с параметрами, приведенными на экранной форме (Рисунок 17), результаты операции изображены на рисунке (Рисунок 18).

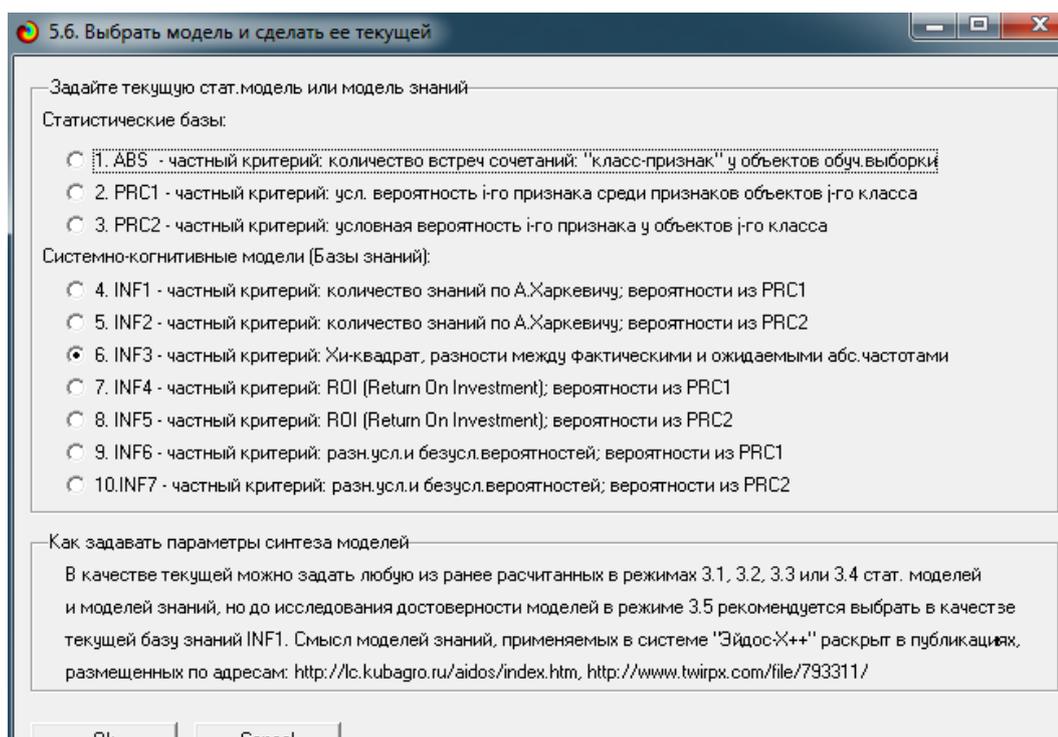


Рисунок 17 – Экранная форма придания наиболее достоверной по F-критерию СК-модели INF3 статуса текущей модели

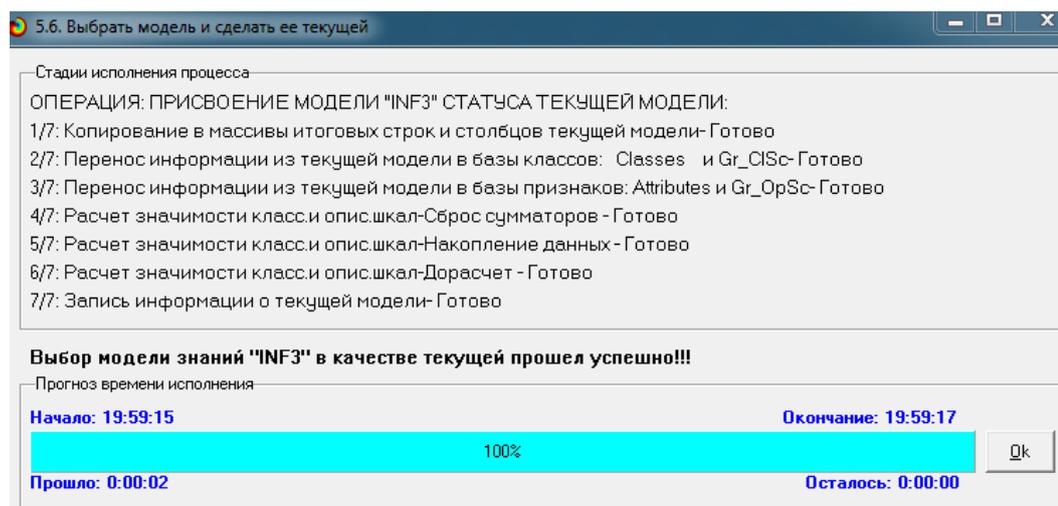


Рисунок 18 – Экранная формы результата присвоения статуса текущей модели

2.2. Прогнозирование (диагностика, классификация, распознавание, идентификация)

Решим задачу системной идентификации, т.е. определение приоритетного приоритетного университета на основе обучающей выборки в наиболее достоверной СК-модели INF3 (

Рисунок 7).

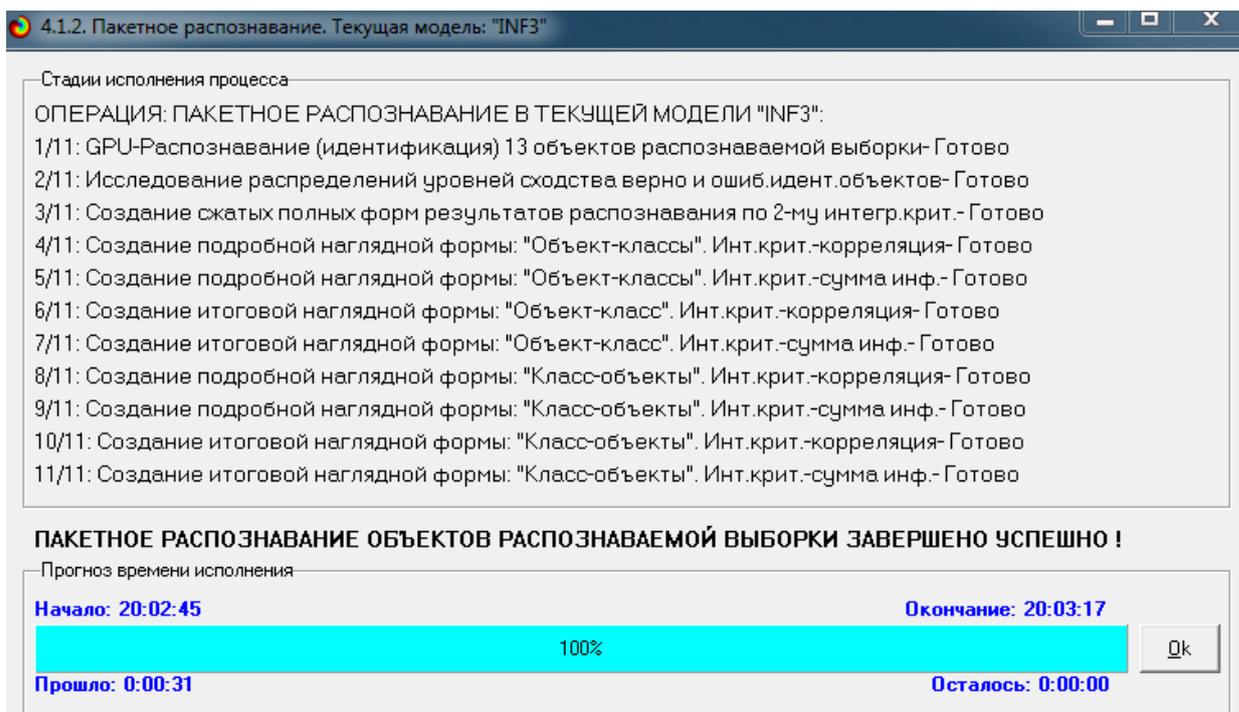


Рисунок 19 – Экранная форма с результатами работы пакетного распознавания модели INF3(режим 4.1.2)

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:

- подробно наглядно: «Объект – классы»;
- подробно наглядно: «Класс – объекты»;
- итоги наглядно: «Объект – классы»;
- итоги наглядно: «Класс – объекты»;
- подробно сжато: «Объект – классы»;
- обобщенная форма по достоверности моделей при разных интегральных критериях;
- обобщенный статистический анализ результатов идентификации по моделям и интегральным критериям;
- статистический анализ результатов идентификации по классам, моделям и интегральным критериям;

- распознавание уровня сходства при разных моделях и интегральных критериях;
- достоверность идентификации классов при разных моделях и интегральных критериях.

На рисунках (Рисунок 20, Рисунок 21) приведены примеры прогнозов высокой и низкой достоверности классов, а также итоги распознавания классов объектов в модели INF3.

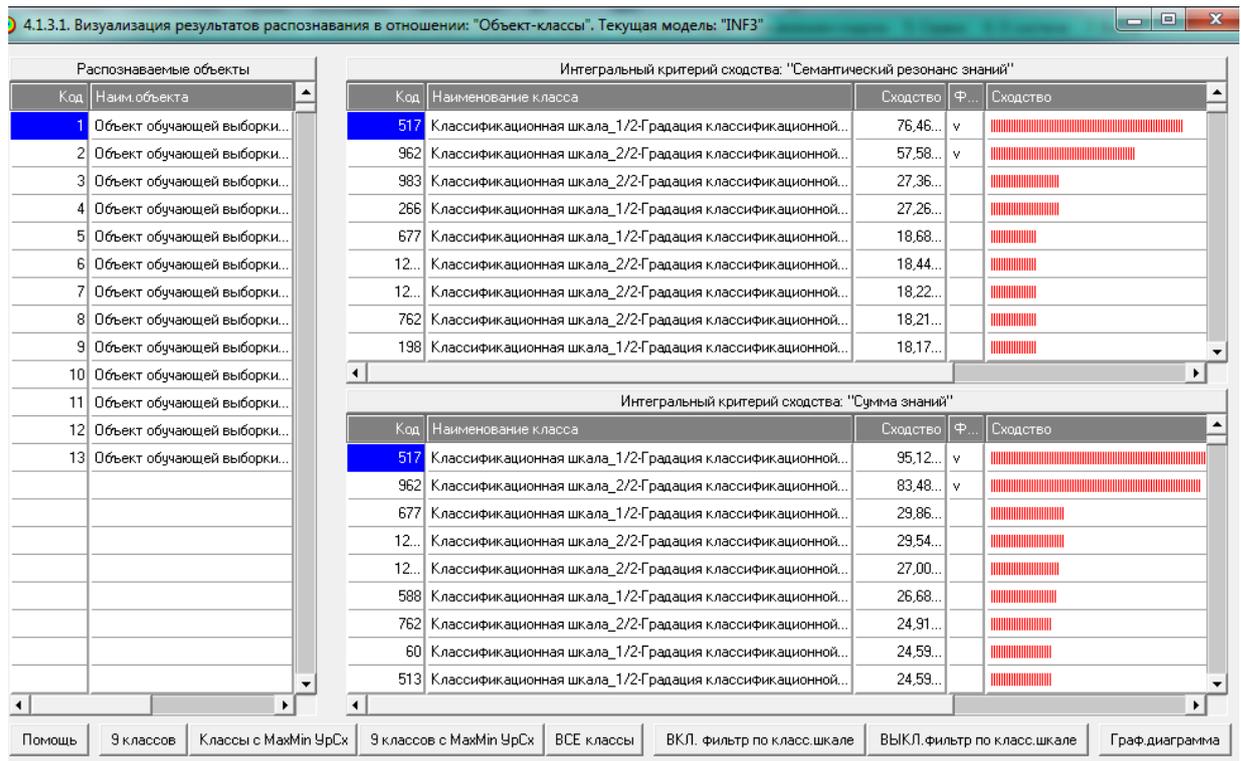


Рисунок 20 – Экранная форма визуализации результатов распознавания в отношении «Объект-классы» модели INF3(режим 4.1.3.1)

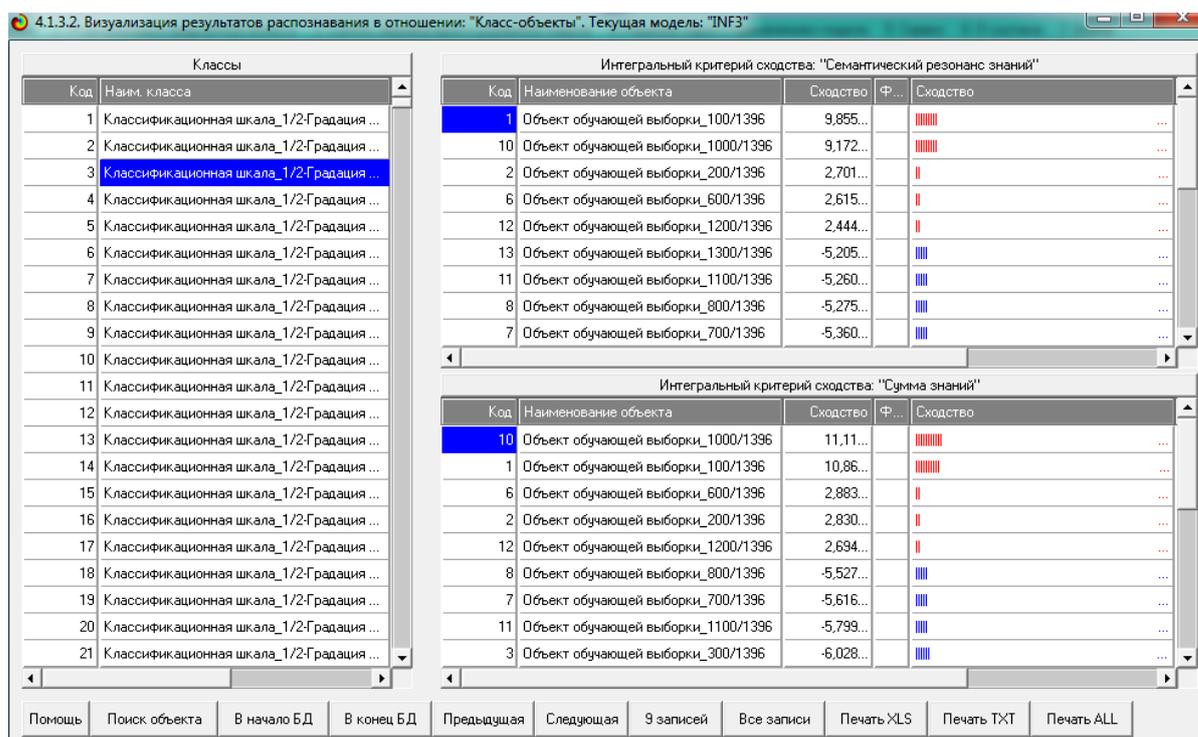


Рисунок 21 – Экранная форма визуализации результатов распознавания в отношении «Классы-объекты» модели INF3(режим 4.1.3.2)

Символ «v» стоит против тех результатов идентификации, которые подтвердились на опыте, т.е. соответствуют факту. Из рисунка (Рисунок 20) видно, что результаты идентификации являются хорошими, естественно при учете информации из рисунка (Рисунок 12) о том, что достоверные прогнозы в данной модели имеют уровень сходства выше 62%, т.е. по сути результаты с более низки уровнем сходства надо просто игнорировать.

2.3. Визуализация когнитивных функций

Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам.

Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека.

Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации; негативные, отражающие чему препятствуют система детерминации; средневзвешенные, отражающие совокупное влияние всех значений факторов на поведение объекта (причем в качестве весов наблюдений используется количество информации в значении аргумента о значениях функции) различной степенью редукции или степенью детерминации, которая отражает в графической форме (в форме полосы) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала.

Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию.

Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой «Aidos». Необходимо отметить, что на вид функций влияния математической моделью АСК-анализа не накладывается никаких ограничений, в частности, они могут быть не дифференцируемые.

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. В системе «Aidos» за формирование когнитивных функций отвечает режим 4.5.

На рисунках (Рисунок 22, Рисунок 23, Рисунок 24, Рисунок 28, Рисунок 25, Рисунок 26, Рисунок 27) приведены примеры некоторых когнитивных функций, наглядно отражающих силу и направление влияния значений (т.е. степени выраженности) различных критериев мирового рейтинга университетов.

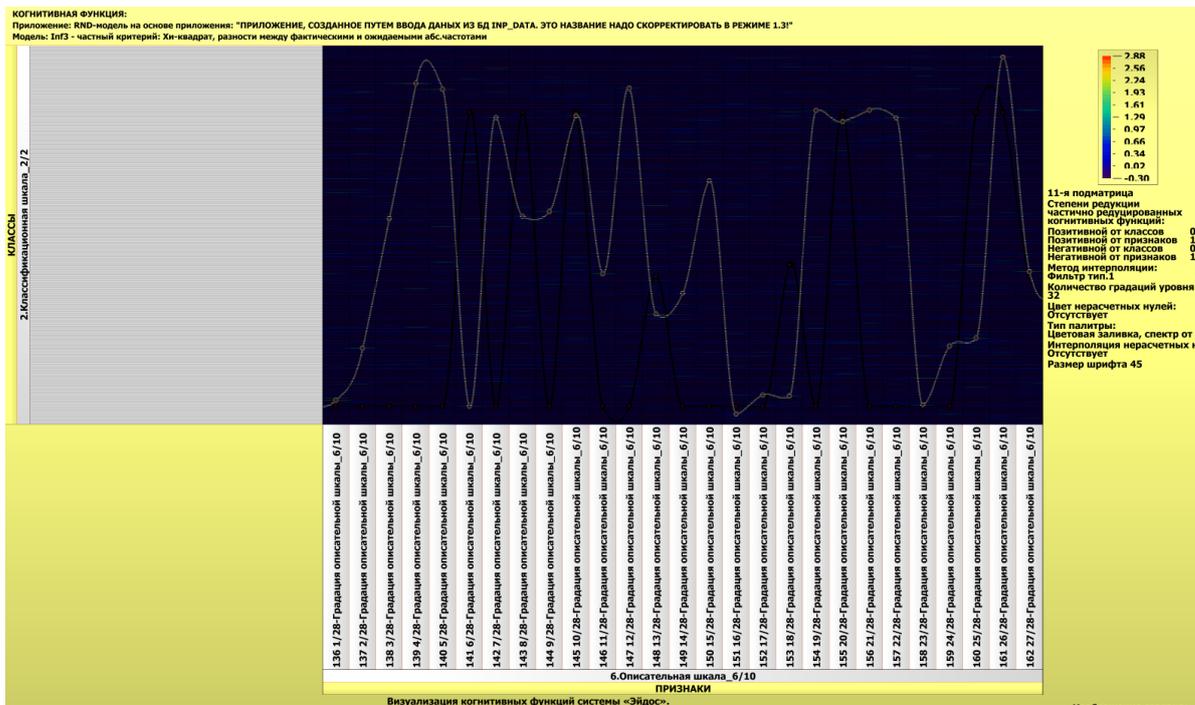


Рисунок 22 – Когнитивная функция, отражающая зависимость рейтинга университета от качества обучения

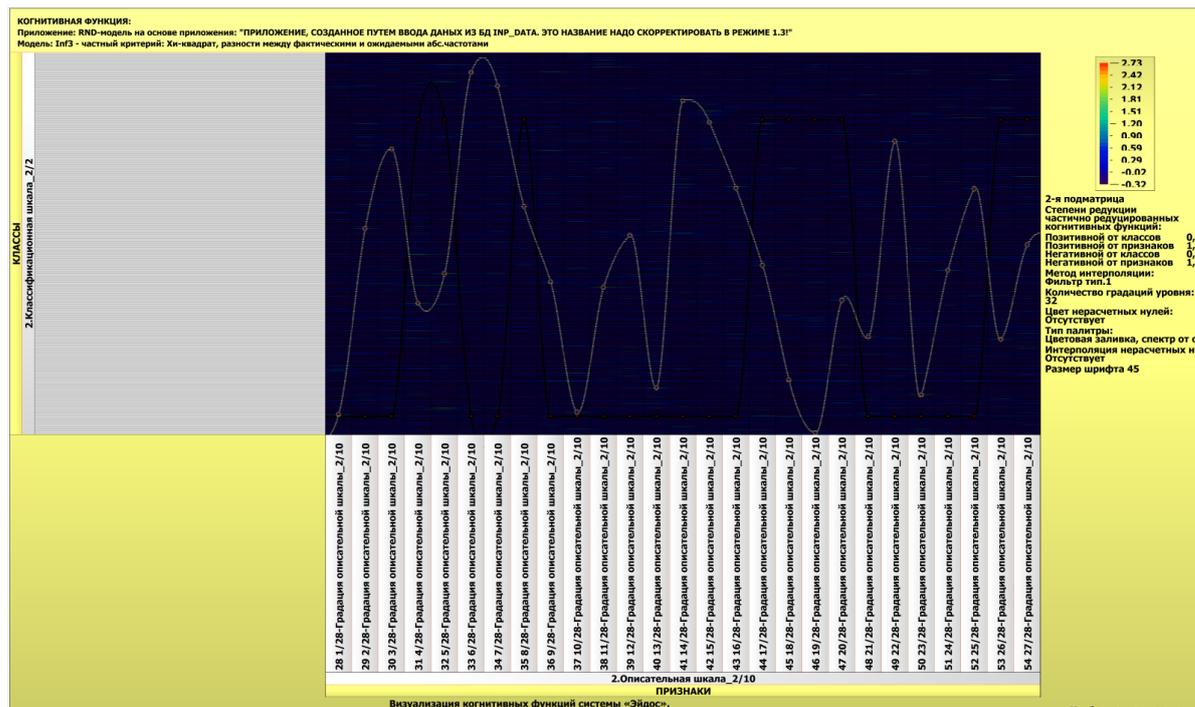


Рисунок 23 – Когнитивная функция, отражающая зависимость рейтинга университета от количества студентов на одного сотрудника

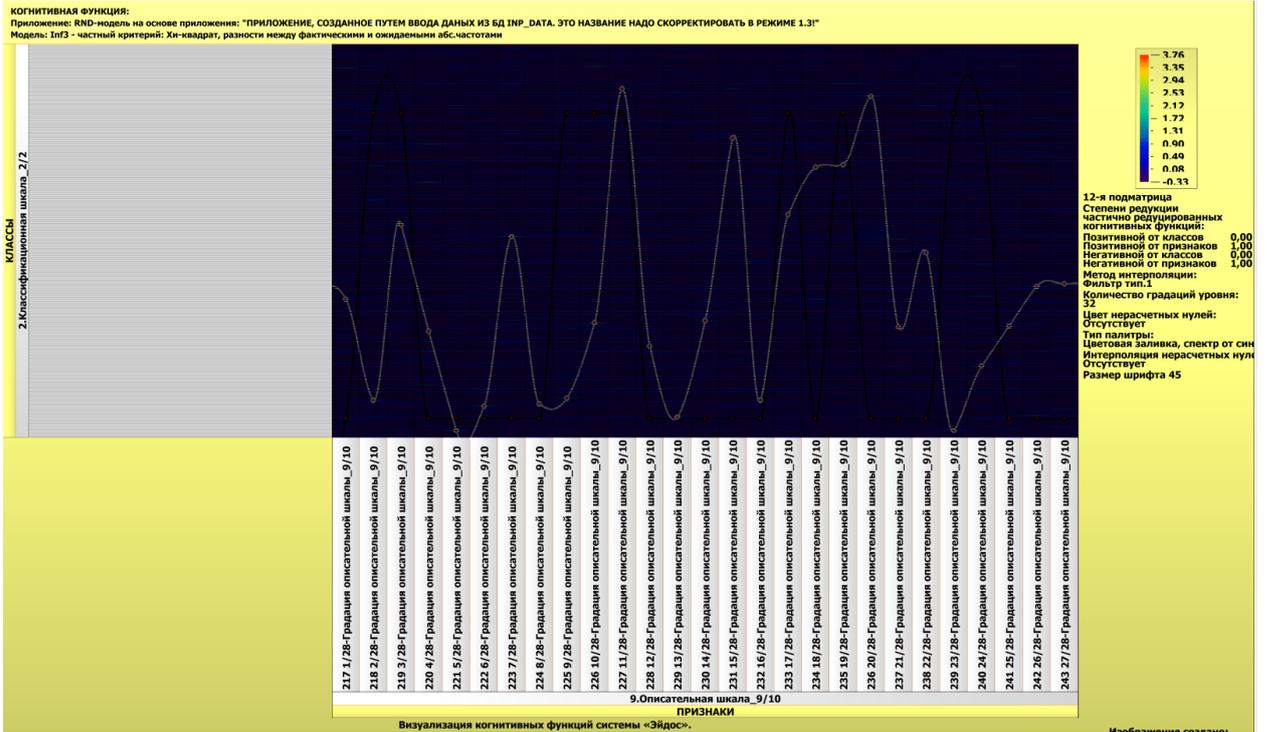


Рисунок 24 – Когнитивная функция, отражающая зависимость рейтинга университета от его международных перспектив

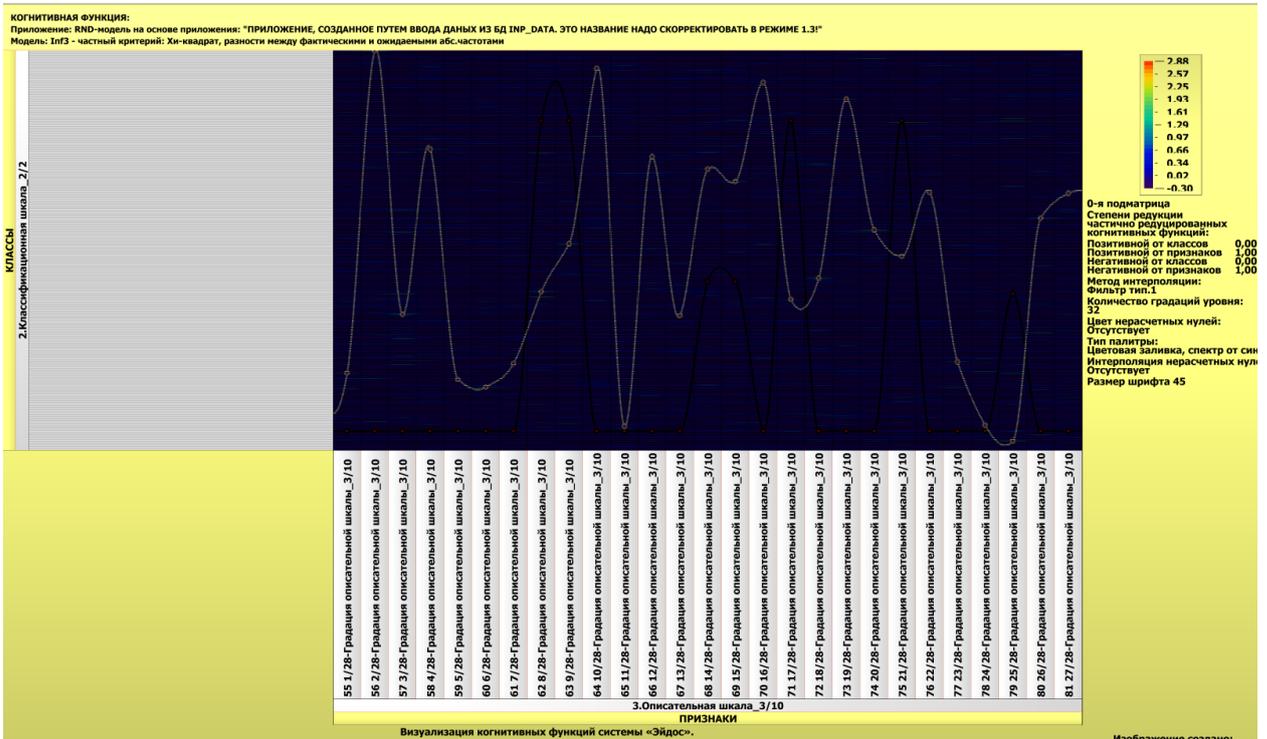


Рисунок 25 – Когнитивная функция, отражающая зависимость рейтинга университета от количества зарубежных студентов

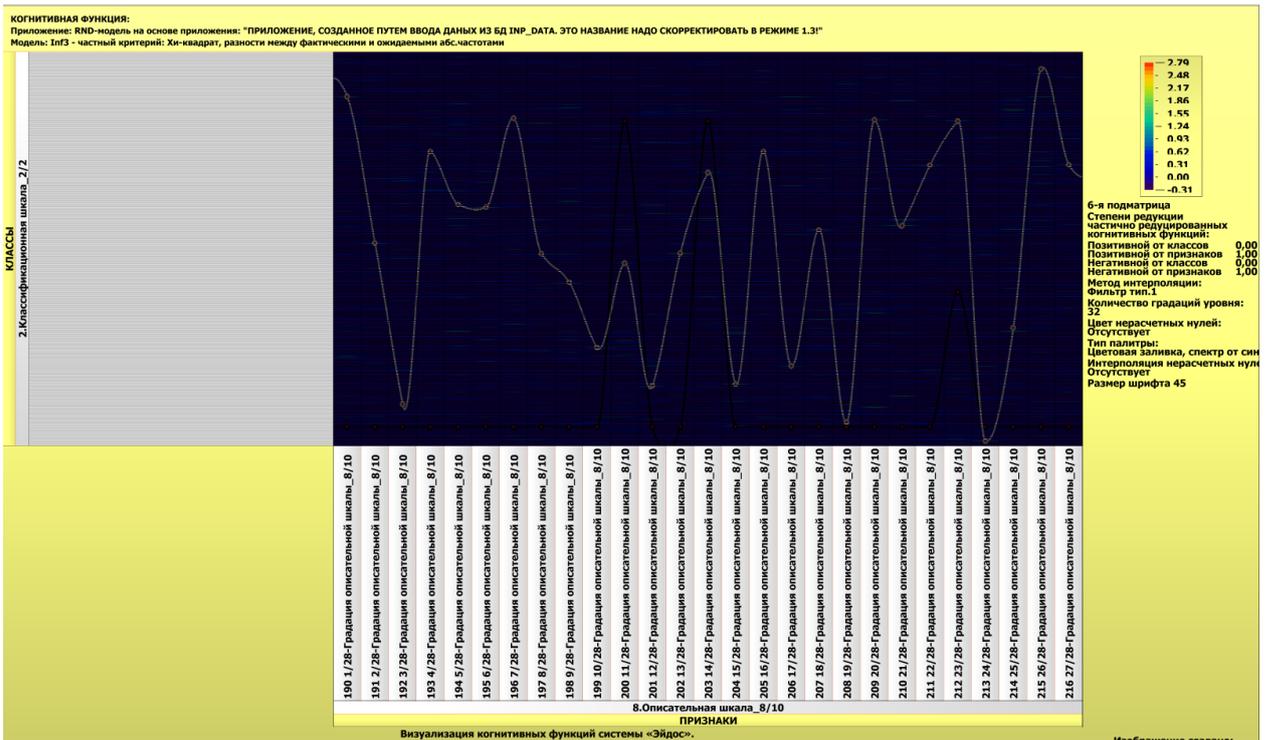


Рисунок 26 – Когнитивная функция, отражающая зависимость рейтинга университета от доходов промышленности

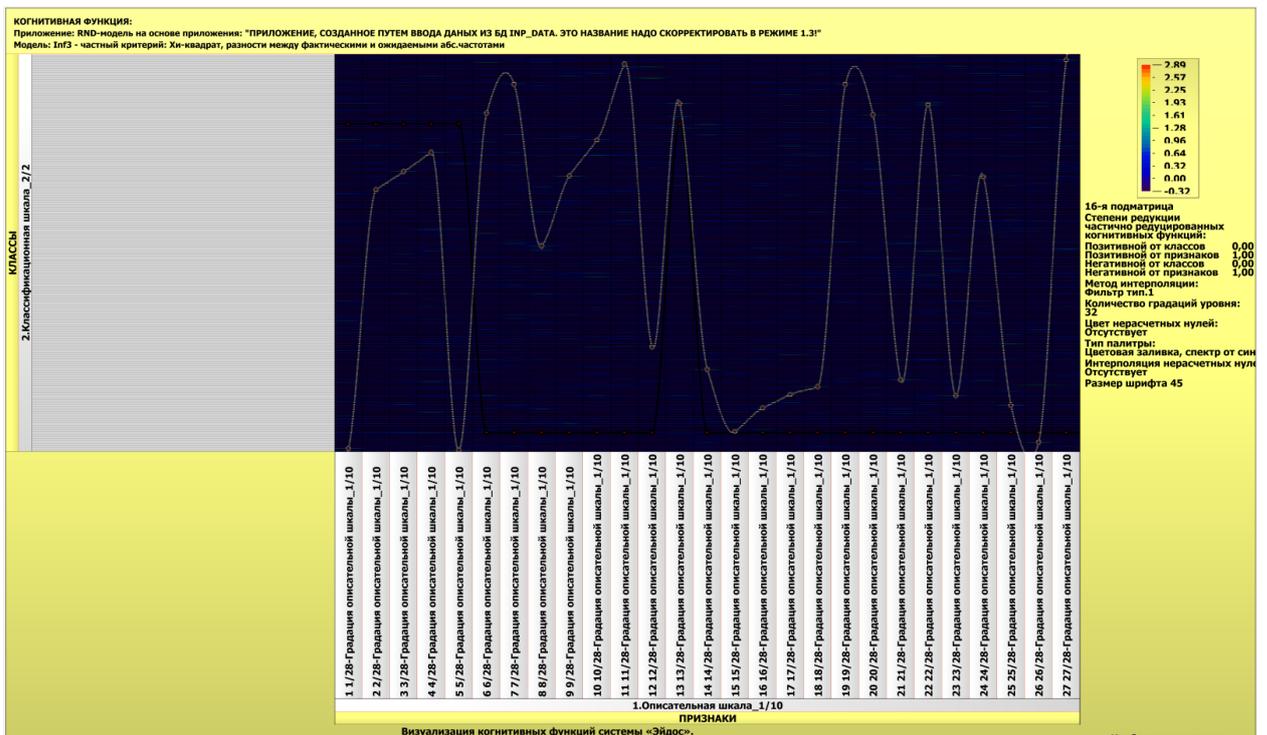


Рисунок 27 – Когнитивная функция, отражающая зависимость рейтинга университета от количества студентов

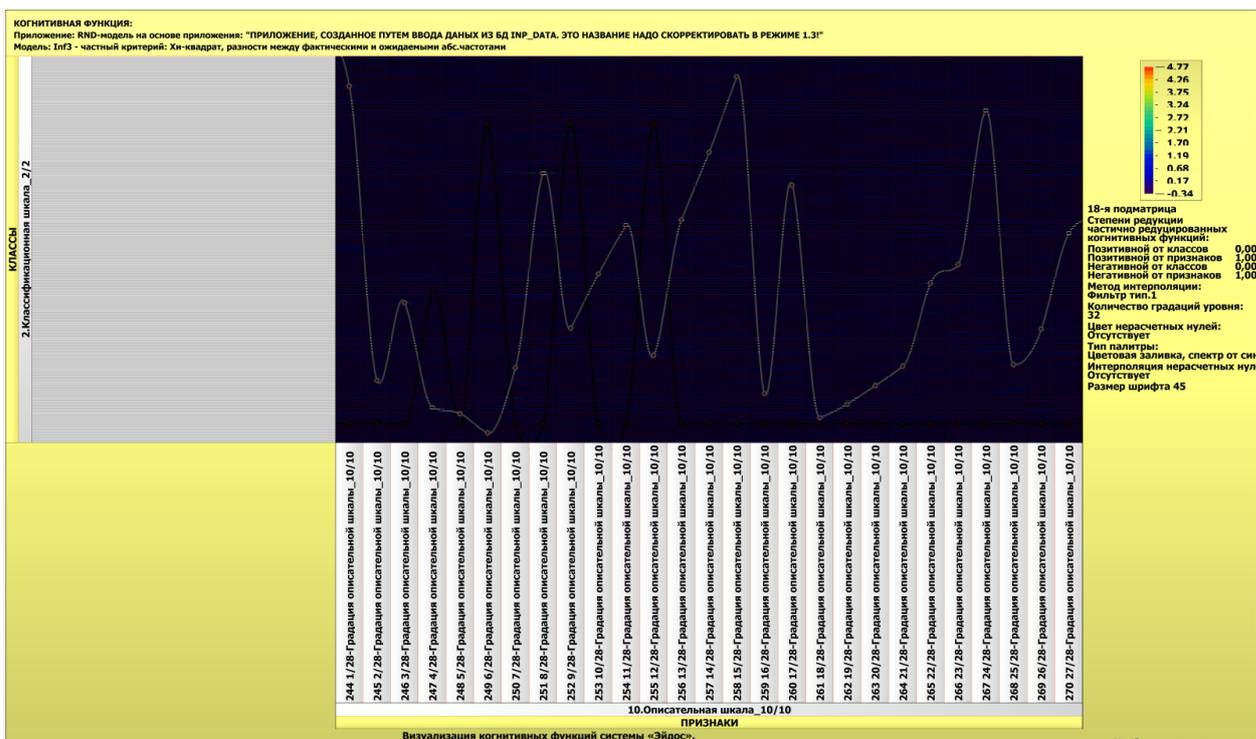


Рисунок 28 – Когнитивная функция, отражающая общий рейтинг университета

В результате анализа полученных данных можно сделать следующие ВЫВОДЫ:

- когнитивная функция отражает зависимость рейтинга университета от качества обучения и показывает, что на первом месте стоит Стенфордский университет (Рисунок 22);
- когнитивные функции, отражающая зависимость рейтинга университета от количества студентов на одного сотрудника (Рисунок 23) и от количества студентов (Рисунок 26) указывают на о то, что лидером является Университет Анатолу;
- согласно когнитивной функции, отражающей зависимость рейтинга университета от его международных перспектив, лидирует Университет Люксембурга (Рисунок 24);
- когнитивная функция, (Рисунок 28) показывает, что самое большое количество зарубежных студентов учится в Американском университете Шарджи и Восточно-Средиземноморском Университете;

- рейтинг Азиатского Университета (Тайвань), Стамбульского Технического Университета, Мюнхенского технического университета наиболее зависит от доходов промышленности (Рисунок 25);
- тем не менее когнитивная функция, отражающая общий рейтинг университета показывает, что самый высокий рейтинг имеет Университет Оксфорда (Рисунок 27).

Инф. портрет класса:1 "Классификационная шкала_1/2-Градация классификационной шкалы_1/744" в модели:6 "INF3"

Код	Наименование признака	Значимость
0		

Код	Наименование признака
1	Классификационная шкала_1/2-Градация клас...
2	Классификационная шкала_1/2-Градация клас...
3	Классификационная шкала_1/2-Градация клас...
4	Классификационная шкала_1/2-Градация клас...
5	Классификационная шкала_1/2-Градация клас...
6	Классификационная шкала_1/2-Градация клас...
7	Классификационная шкала_1/2-Градация клас...
8	Классификационная шкала_1/2-Градация клас...
9	Классификационная шкала_1/2-Градация клас...
10	Классификационная шкала_1/2-Градация клас...
11	Классификационная шкала_1/2-Градация клас...
12	Классификационная шкала_1/2-Градация клас...
13	Классификационная шкала_1/2-Градация клас...
14	Классификационная шкала_1/2-Градация клас...
15	Классификационная шкала_1/2-Градация клас...
16	Классификационная шкала_1/2-Градация клас...
17	Классификационная шкала_1/2-Градация клас...
18	Классификационная шкала_1/2-Градация клас...
19	Классификационная шкала_1/2-Градация клас...
20	Классификационная шкала_1/2-Градация клас...
21	Классификационная шкала_1/2-Градация клас...
22	Классификационная шкала_1/2-Градация клас...
23	Классификационная шкала_1/2-Градация клас...

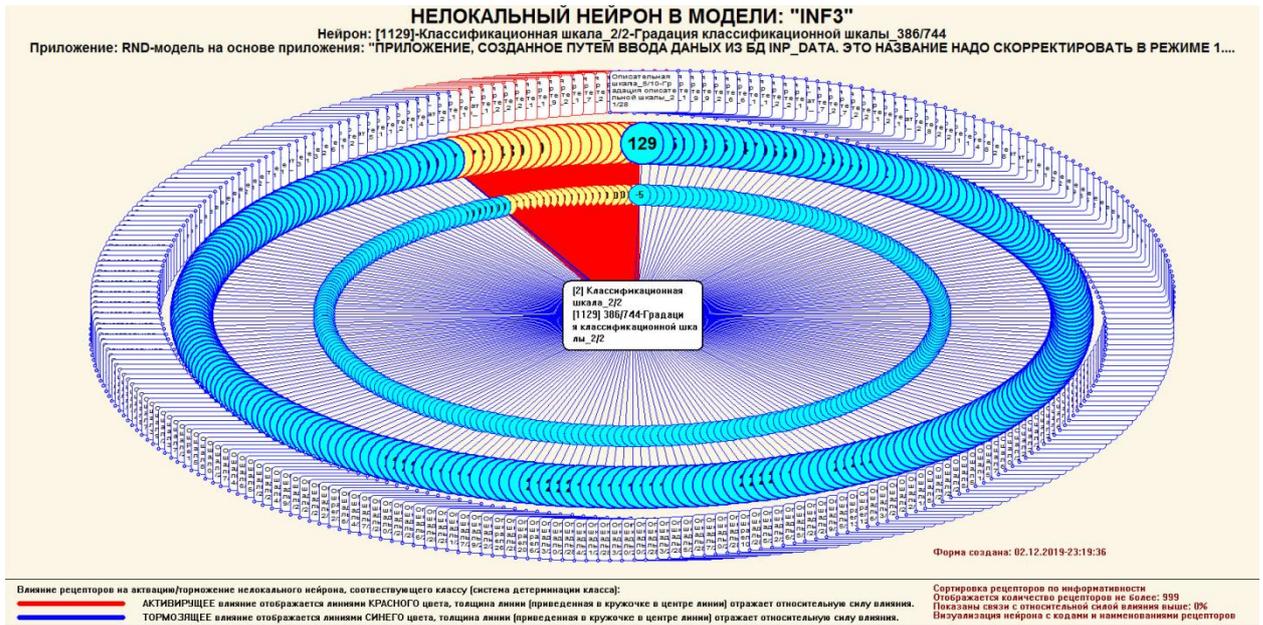
Помощь Abs Prс1 Prс2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 MS Excel ВКЛ. фильтр по фактору ВЫКЛ. фильтр по фактору Вписать в окно Показать ВСЕ

Рисунок 29 – Информационный портрет рейтинга университетов

2.4. Нелокальные нейроны и нейронная сеть

Для каждого технологического фактора в соответствии с предложенной моделью определяется величина и направление его влияния на осуществление всех предпочтений абитуриента. Для каждой ситуации эта информация отображается в различных текстовых и графических формах, в частности, в форме нелокального нейрона (Рисунок 0). На данной диаграмме цвет линии означает знак связи (красный – положительная, синий – отрицательная), а толщина – ее модуль.

Рисунок 30 – Графическое отображение нелокальных нейронов в модели INF3 для нейрона OCCUPANCY-No



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 29). Необходимо отметить, что все указанные графические формы генерируются системой «Aidos» автоматически в соответствии с созданной моделью[8].

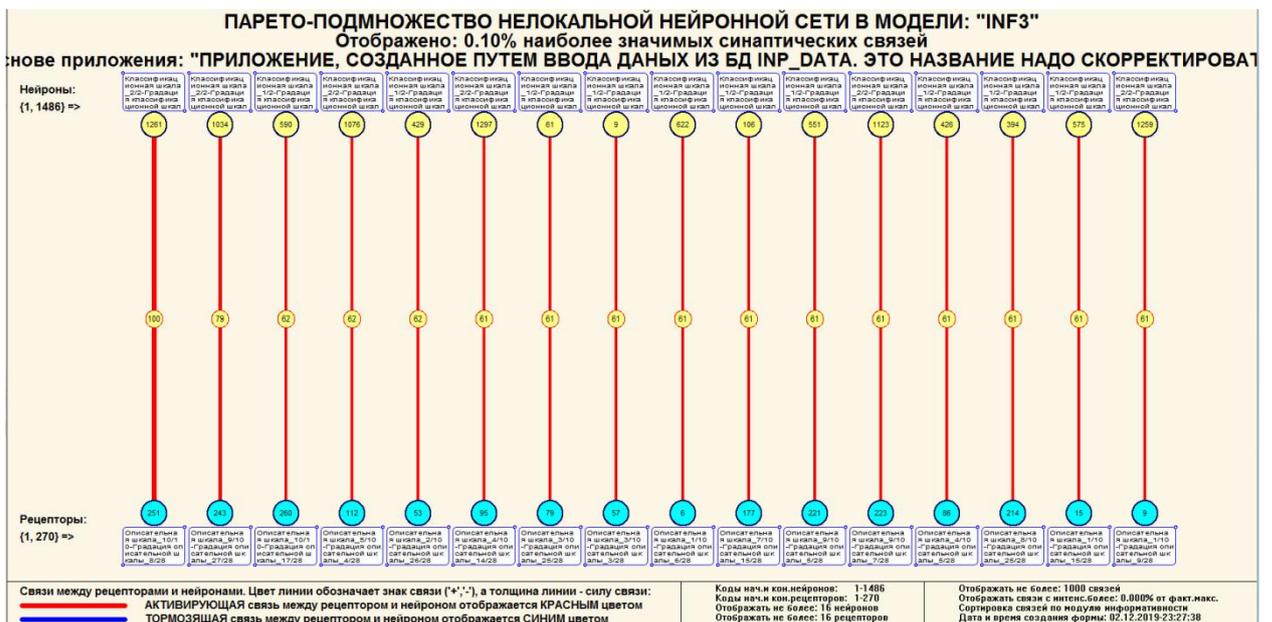


Рисунок 29 –Парето-подмножество нелокальной нейронной сети в модели INF3

Модель знаний системы «Aidos» относится к нечетким декларативным гибридным моделям и объединяет в себе некоторые особенности нейросетевой и фреймовой моделей представления знаний. Классы в этой модели соответствуют нейронам и фреймам, а признаки рецепторам и шпациям (описательные шкалы – слотам).

От фреймовой модели представления знаний модель системы «Aidos» отличается своей эффективной и простой программной реализацией, полученной за счет того, что разные фреймы отличаются друг от друга не набором слотов и шпаций, а лишь информацией в них. Поэтому в системе «Aidos» при увеличении числа фреймов само количество баз данных не увеличивается, а увеличивается лишь их размерность.

2.5. SWOT и PEST матрицы и диаграммы

При принятии решений определяется сила и направление влияния значений факторов на принадлежность состояний объекта моделирования к тем или иным классам, соответствующим различным будущим состояниям. По сути, это решение задачи SWOT-анализа [9].

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.

Применительно к задаче, решаемой в данной работе, SWOT-анализ показывает степень влияния различных значений параметров рейтинг университетов.

В системе «Aidos» в режиме 4.4.8 поддерживается решение этой задачи. При этом выявляется система детерминации заданного класса, т.е. система значений факторов, обуславливающих переход объекта моделирования и управления в состояние, соответствующее данному классу, а также препятствующих этому переходу.

На рисунках (Рисунок 302, Рисунок 313) приведены SWOT-диаграммы, наглядно отражающие силу и направление влияния различных значений параметров рейтинг университетов.

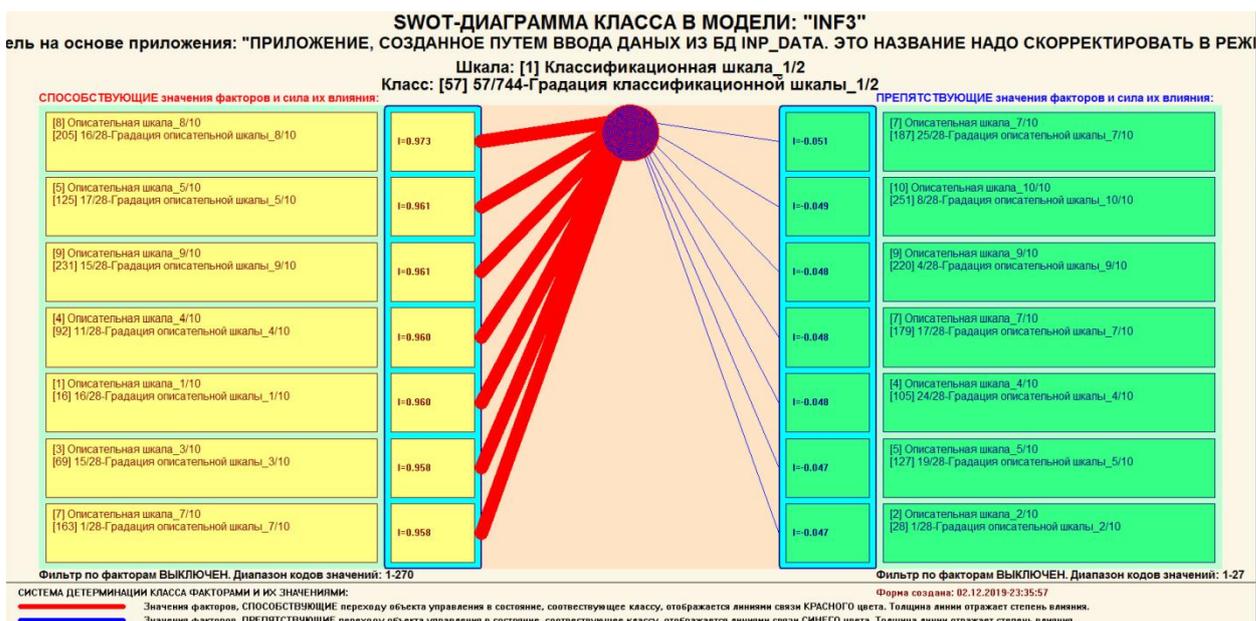


Рисунок 30 –SWOT-диаграмма в модели INF3

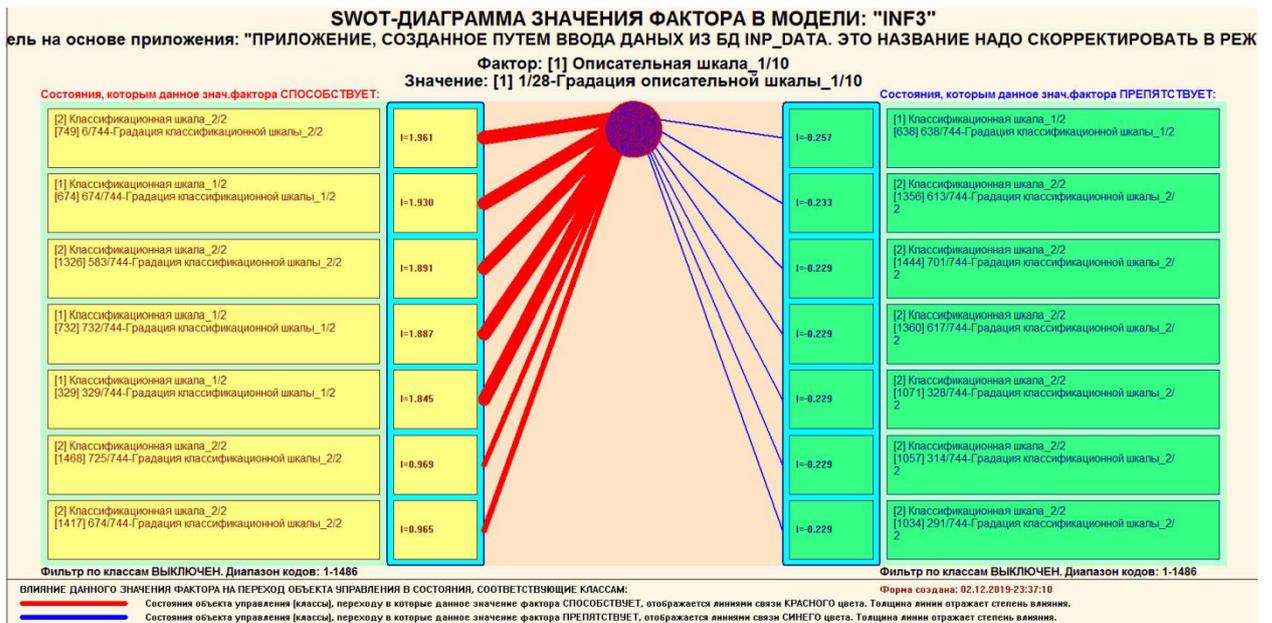


Рисунок 31 –SWOT-диаграмма в модели INF3

Заключение

Как показывает анализ результатов численного эксперимента предложенное и реализованное в системе «Aidos» решение поставленных задач является вполне эффективным, что позволяет обоснованно утверждать, что цель работы достигнута, поставленная проблема решена.

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 приводится развернутый пример использования базы данных Kaggle для оценки мирового рейтинга учебных заведений, применяемых в АСК-анализе и его программном инструментарии системе искусственного интеллекта «Aidos». При этом наиболее достоверной в данном приложении оказались модель INF3, основанная на частном критерии Хи-квадрат, разности между фактическими и ожидаемыми абсолютными частотами при интегральном критерии «Сумма знаний». Точность модели составляет 0,159, что является хорошим результатом. Для оценки достоверности моделей в АСК-анализе и системе «Aidos» используется метрика, сходная с F-критерием. Также необходимо обратить внимание, что статистические модели в данном приложении дают примерно на 21% более низкую средневзвешенную достоверность идентификации и не идентификации, чем модели знаний, что, как правило, наблюдается и в других приложениях. Этим и оправдано применение моделей знаний.

В результате проделанной работы, с помощью системы «Aidos» были созданы статистические и системно-когнитивные модели, в которых непосредственно на основе эмпирических данных сформированы обобщенные образы классов по различным значениям показателей рейтинга университетов, изучено влияние значений различных параметров на эти прогнозы, и, на основе этого, решены задачи квалиметрии, классификации и

исследования моделируемой предметной области путем исследования ее модели.

Список литературы

1. World University Rankings 2020// [Электронный ресурс] URL: <https://www.kaggle.com/joeshamen/world-university-rankings-2020>
2. Сайт профессора Е.В. Луценко // [Электронный ресурс] intuit.ru. URL: <http://lc.kubagro.ru/>
3. 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 «AIDOS» // 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>
4. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Aidos» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 у.п.л.
5. Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе «Aidos» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №02(126). С. 1 – 32. – IDA [article ID]: 1261702001. – Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 у.п.л.

6. 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.

7. Луценко Е.В. Метод визуализации когнитивных функций - новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. - Краснодар: КубГАУ, 2011. - №03(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077, 2,688 у.п.л. - Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>

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. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы «Aidos-X++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 1367 – 1409. – IDA [article ID]: 1011407090. – Режим доступа: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 у.п.л.