

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:

«АСК анализ рейтинга различных кухонь мира основе данных
репозитория UCI»

Выполнил студент группы ИТ31841 Цепилов Даниил Олегович

Допущен к защите _____

Руководитель проекта _____ Луценко 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)

Краснодар
2019 г.

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Факультет заочного обучения
Кафедра компьютерных технологий и систем

УТВЕРЖДАЮ:

Зав. кафедрой КТС _____ В. И. Лойко

**ЗАДАНИЕ
на курсовую работу**

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

Цепилову Даниилу Олеговичу
(Ф.И.О.)

Тема проекта: **АСК анализ рейтинга различных кухонь мира основе данных репозитория UCI**

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

Объем работы:

а) пояснительная записка к работе _____ 32 _____ листа формата А4

б) графическая часть _____ 26 _____ рисунков.

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

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

Срок защиты: “ ” _____ 2019 г.

Дата выдачи задания: “ ” _____ 2019 г.

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

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

(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ “ ” _____ 2019 г.

(подпись, дата)

Краснодар
2019 г.

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

Факультет заочного обучения

РЕЦЕНЗИЯ
на курсовой проект (работу)

Студента Цепилова Даниила Олеговича

курса 2 заочной формы обучения

направления подготовки Информационные системы и технологии

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

Наименование темы «АСК анализ рейтинга различных кухонь мира основе данных репозитория UCI».

Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсового проекта (работы)

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

Достоинства работы

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

Недостатки работы

Отдельные погрешности в оформлении

Итоговая оценка при защите 5(отл.)

Рецензент _____

Е.В. Луценко

« ____ » _____ 2019 г.

РЕФЕРАТ

Курсовая работа содержит: 32 страницы, 26 рисунков, 9 литературных источников.

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

В данной курсовой работе были проанализированы методы формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами, принятия решений и исследования моделируемой предметной области путем исследования ее модели.

СОДЕРЖАНИЕ

Введение.....	6
1. Синтез и верификация моделей.....	7
1.1. Описание решения	7
1.2. Преобразование исходных данных из CSV-формата в файл исходных данных MS Excel	7
1.3. Синтез и верификация статистических и интеллектуальных моделей	13
1.4. Виды моделей системы «Aidos».....	15
1.5. Результаты верификации моделей	18
2. РЕШЕНИЕ ЗАДАЧ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ.....	20
2.1. Решение задачи	20
2.2 Визуализация когнитивных функций	23
2.3 SWOT и PEST матрицы и диаграммы	25
2.4 Нелокальные нейронные сети и нейтроны.....	27
2.5 Кластерный и конструктивный анализ	29
Заключение	30
Список используемых источников.....	31

ВВЕДЕНИЕ

Одним из важнейших и перспективных направлений в развитии современных информационных технологий, на сегодняшний день, является создание систем искусственного интеллекта. В результате развития технологий в данном направлении, возникла необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи идентификации классов рейтинга мировых кухонь.

При выполнении работы нужно учесть, что необходимы свободный доступ к тестовым исходным данным и методика, которая необходима для преобразования этих данных в форму, которая необходима для работы в системе искусственного интеллекта. Среди баз данных текстовых задач для систем искусственного интеллекта удачным выбором является репозиторий «UCI».

В данной курсовой работе использована база данных «Restaurant & consumer data Data Set» из банка исходных данных по задачам искусственного интеллекта - репозитория «UCI».

Для решения задачи используем стандартные возможности Microsoft Office Excel, блокнот, а также систему искусственного интеллекта «Aidos-X++».

1.СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1.Описание решения

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

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

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

Из электронного ресурса баз данных «UCI» возьмем набор данных «Restaurant & consumer data Data Set» <http://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data>, которую приведем к виду, представленному на рисунке 1.

	A	B	C	D	E	
1	ID Пользователя	ID Кухня	Рейтинг	Рейтинг еды	Рейтинг сервиса	
2	U1077	135085	2	2	2	
3	U1077	135038	2	2	1	
4	U1077	132825	2	2	2	
5	U1077	135060	1	2	2	
6	U1068	135104	1	1	2	
7	U1068	132740	0	0	0	
8	U1068	132663	1	1	1	
9	U1068	132732	0	0	0	
10	U1068	132630	1	1	1	
11	U1067	132584	2	2	2	
12	U1067	132733	1	1	1	
13	U1067	132732	1	2	2	
14	U1067	132630	1	0	1	
15	U1067	135104	0	0	0	
16	U1067	132560	1	0	0	
17	U1103	132584	1	2	1	
18	U1103	132732	0	0	2	
19	U1103	132630	1	2	0	
20	U1103	132613	2	2	2	
21	U1103	132667	1	2	2	
22	U1103	135104	1	2	0	
23	U1103	132663	1	0	2	
24	U1103	132733	2	2	2	
25	U1107	132660	2	2	1	

Рисунок 1 - Исходные данные в файле MS Excel

Поскольку ввод исходных данных в систему «Эйдос» планируется осуществить с помощью ее универсального программного интерфейса импорта данных из внешних баз данных, который работает с файлами MS Excel, то преобразуем данные из csv-файла в xls-файл, для чего выполним следующие операции.

Скопируем таблицу в MS Excel и запишем ее с именем: Inp_data.xls в папку: C:\Aidos-X\AID_DATA\Inp_data\. В файле Inp_data.xls добавим нумерацию на позиции «А» и автоматически укажем номера для всех строк. В результате получим таблицу исходных данных, полностью подготовленную для обработки системой «Эйдос» и записанную в нужной папке в виде файла нужного типа с нужным именем.

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, расположенный в режиме 2.3.2.2.

Рисунок 2- Экранная форма универсального программного интерфейса импорта данных в систему «Эйдос»

В экранной форме, приведенной на рисунке 2, необходимо задать настройки:

— «Тип файла исходных данных Inp_data»: «XLS – MS Excel-2003»;

— «Диапазон столбцов классификационных шкал»: «Начальный столбец классификационных шкал» – 2, «Конечный столбец классификационных шкал» – 3;

— «Диапазон столбцов описательных шкал»: «Начальный столбец описательных шкал» - 4, «Конечный столбец описательных шкал» - 6;

— «Параметры формирования сценариев или способа интерпретации текстовых полей»: «Не применять сценарный метод АСК-анализа и спец.интерпретацию ТХТ-полей».

После нажатия кнопки «ОК» следует открытие окна, где размещается информация о размерности загруженной модели. На данной форме необходимо нажать кнопку «Выйти на создание модели», что показано на рисунке 3.

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных шкал	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	1	3	3,00	3	30	10,00
Текстовые	1	138	138,00	0	0	0,00
ВСЕГО:	2	141	70,50	3	30	10,00

Задайте число интервалов (градаций) в шкале:

В классификационных шкалах: В описательных шкалах:

Рисунок 3- Задание размерности модели системы «Эйдос»

Далее происходит открытие формы, отображающей стадию процесса импорта данных из внешней БД «Inp_data.xls» в систему «Эйдос», а также прогноз времени завершения этого процесса. В данном окне необходимо дождаться завершения формализации предметной области и нажать кнопку «ОК», как показано на рисунке 4.

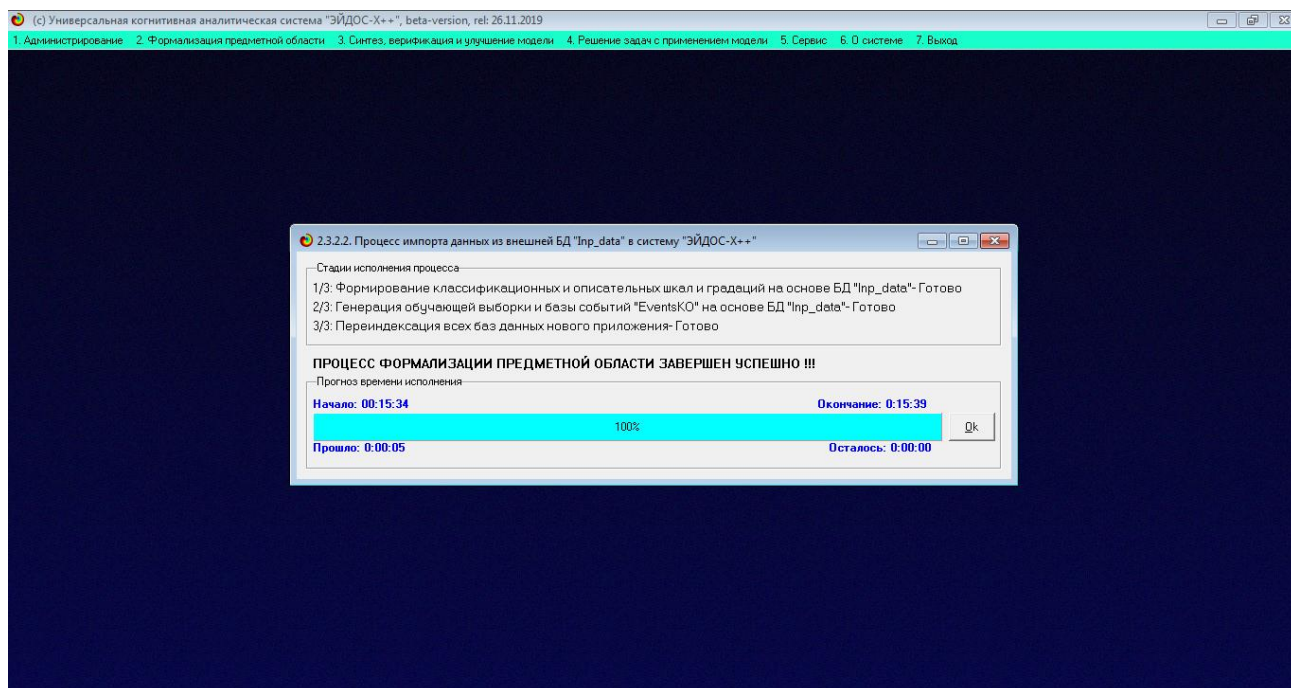


Рисунок 4- Процесс импорта данных из внешней БД «Inp_data.xls» в систему «Эйдос»

В результате выполненных действий формируются классификационные и описательные шкалы, градации, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Эти действия полностью автоматизированы и выполняются на 2-й этапе АСК- анализа «Формализация предметной области». Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1, рисунок 5.

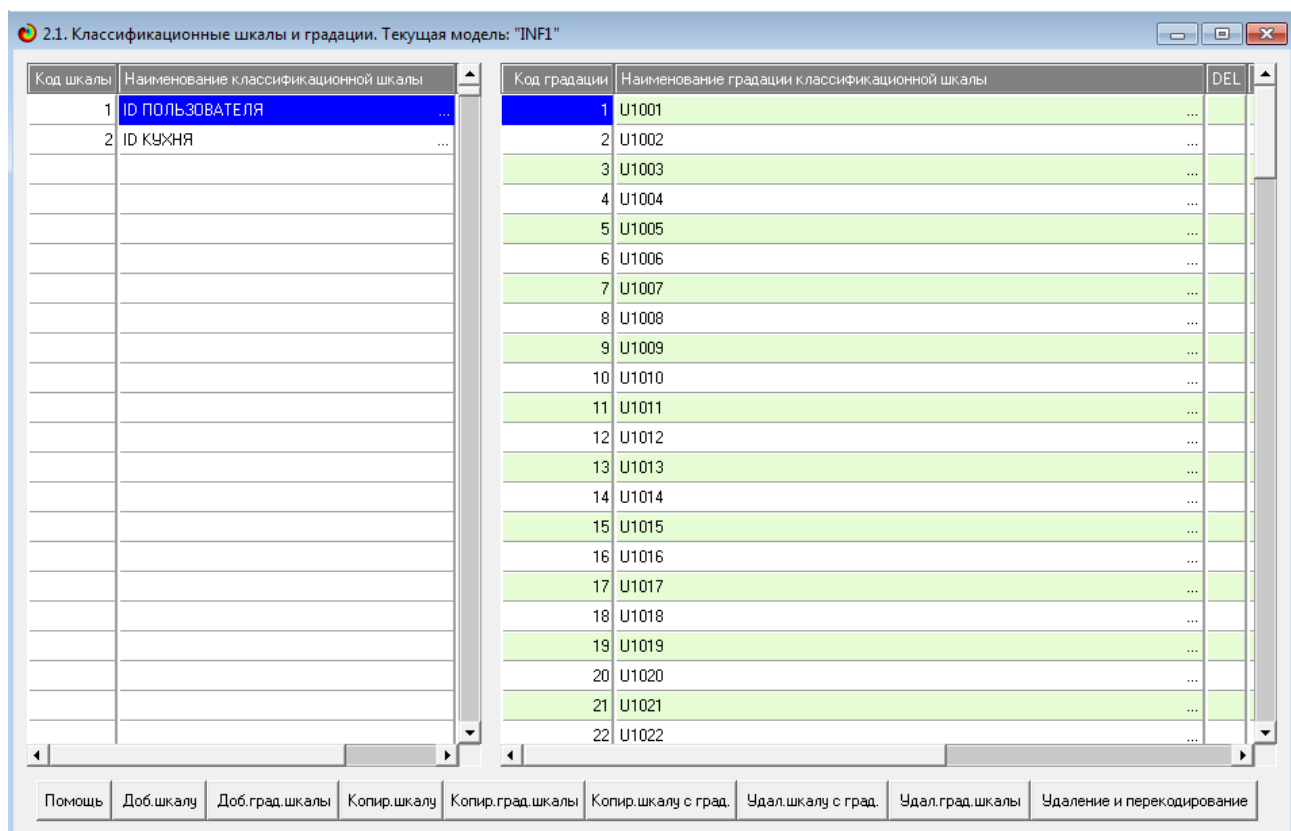


Рисунок 5- Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2, рисунок 6.

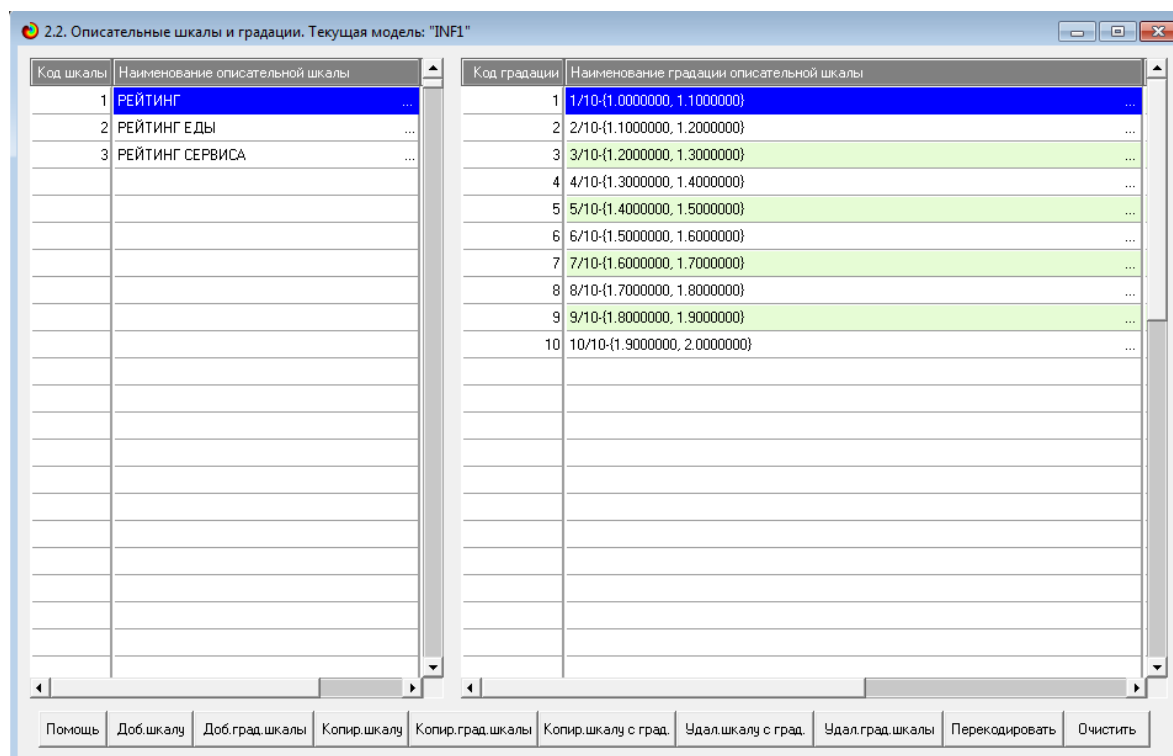


Рисунок 6 - Описательные шкалы и градации (фрагмент)

Для просмотра обучающей выборки необходимо запустить режим 2.3.1., рисунок 7.

The screenshot shows a software window titled "2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: 'TNF1'". The interface includes two data tables and a control bar at the bottom.

Top Table:

Код объекта	Наименование объекта	Дата	Время
1	1		
2	2		
3	3		
4	4		
5	5		
6	6		
7	7		
8	8		
9	9		
10	10		
11	11		
12	12		

Bottom Left Table:

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	77	141	0	0

Bottom Right Table:

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	10	20	30	0	0	0	0

Control Bar:

Помощь Скопировать обуч. выб. в расп. Добавить объект Добавить классы Добавить признаки Удалить объект Удалить классы Удалить признаки Очистить БД

Рисунок 7- Обучающая выборка (фрагмент)

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

1.3. Синтез и верификация статистических и интеллектуальных моделей

Далее необходимо запустить режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей, что показано на рисунке 8.

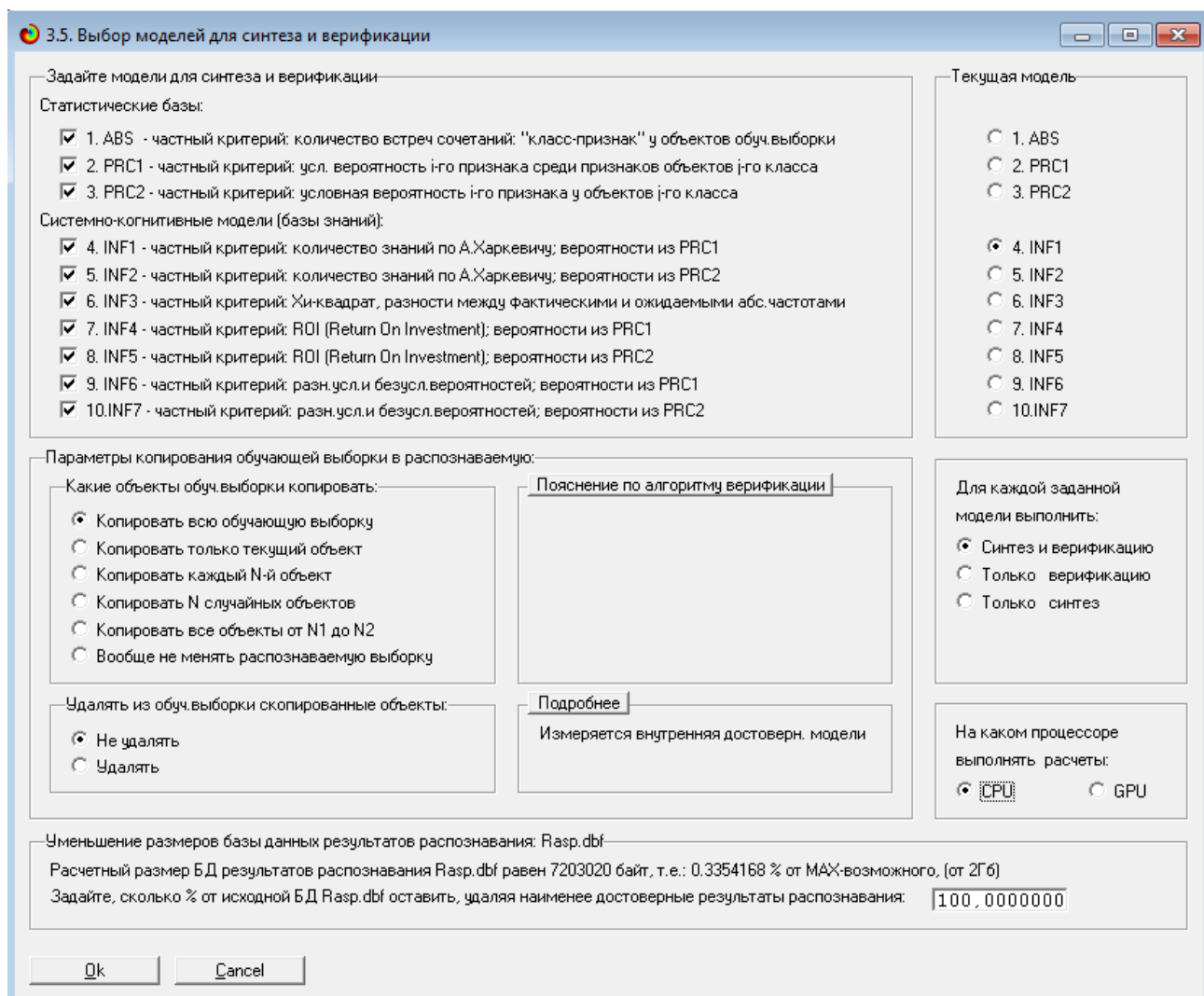


Рисунок 8- Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Для синтеза и верификации выбираются все имеющиеся модели, после чего необходимо переключить выполнение расчетов с графического процессора на системный процессор, т.к. тактовая частота на данном устройстве системного превышает частоту графического процессора.

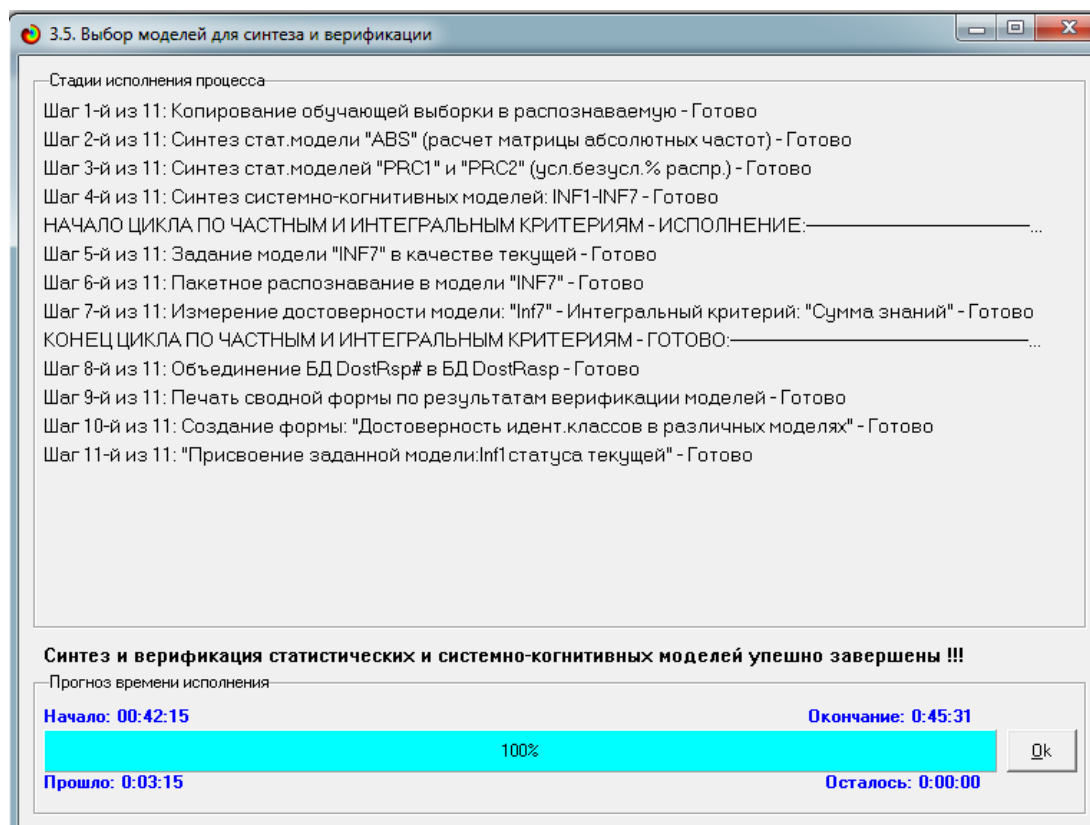


Рисунок 9 – Завершение процесса синтеза и верификации моделей

После этого можно перейти непосредственно к выбору наиболее достоверной модели.

1.4. Виды моделей системы «Эйдос»

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

По сути, частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот в матрицы условных и безусловных процентных распределений, и матрицы знаний.

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

Код признака	Наименование описательной шкалы и градации	134. ID Пользователя	135. ID Пользователя	136. ID Пользователя	137. ID Пользователя	138. ID Пользователя	139. ID КЧ/НЧ 1/3 (132560.0, 133409.7)	140. ID КЧ/НЧ 2/3 (133409.7, 134259.3)	141. ID КЧ/НЧ 3/3 (134259.3, 135109.0)	Сумма	Среднее	Средн. квадрат. откл.
14	РЕЙТИНГ ЕДЫ-4/10-(1.3000000, 1.4000000)											
15	РЕЙТИНГ ЕДЫ-5/10-(1.4000000, 1.5000000)											
16	РЕЙТИНГ ЕДЫ-6/10-(1.5000000, 1.6000000)											
17	РЕЙТИНГ ЕДЫ-7/10-(1.6000000, 1.7000000)											
18	РЕЙТИНГ ЕДЫ-8/10-(1.7000000, 1.8000000)											
19	РЕЙТИНГ ЕДЫ-9/10-(1.8000000, 1.9000000)											
20	РЕЙТИНГ ЕДЫ-10/10-(1.9000000, 2.0000000)	5		7	12	3	199		317	1032	7.32	31.1
21	РЕЙТИНГ СЕРВИСА-1/10-(1.0000000, 1.1000000)	7		2	1	2	175		251	852	6.04	25.4
22	РЕЙТИНГ СЕРВИСА-2/10-(1.1000000, 1.2000000)											
23	РЕЙТИНГ СЕРВИСА-3/10-(1.2000000, 1.3000000)											
24	РЕЙТИНГ СЕРВИСА-4/10-(1.3000000, 1.4000000)											
25	РЕЙТИНГ СЕРВИСА-5/10-(1.4000000, 1.5000000)											
26	РЕЙТИНГ СЕРВИСА-6/10-(1.5000000, 1.6000000)											
27	РЕЙТИНГ СЕРВИСА-7/10-(1.6000000, 1.7000000)											
28	РЕЙТИНГ СЕРВИСА-8/10-(1.7000000, 1.8000000)											
29	РЕЙТИНГ СЕРВИСА-9/10-(1.8000000, 1.9000000)											
30	РЕЙТИНГ СЕРВИСА-10/10-(1.9000000, 2.0000000)	6		8	13	1	140		280	840	5.96	26.1
	Сумма числа признаков	41		30	40	9	984		1664	5296		
	Среднее	1		1	1	0	33		56		1.25	
	Среднеквадратичное отклонение	3		2	4	1	68		114			12.3
	Сумма числа объектов обуч.выборки	14		10	14	3	361		607	1936		

Рисунок 10- Матрица абсолютных частот (модель ABS) и условных и безусловных процентных распределений (фрагмент)

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"											
Код признака	Наименование описательной шкалы и градации	135. ID Пользователя U1135	136. ID Пользователя U1136	137. ID Пользователя U1137	138. ID Пользователя U1138	139. ID КЧ/НЧ 1/3 (132560.0, 133409.7)	140. ID КЧ/НЧ 2/3 (133409.7, 134259.3)	141. ID КЧ/НЧ 3/3 (134259.3, 135109.0)	Сумма	Среднее	Средн. квадрат. откл.
13	РЕЙТИНГ ЕДЫ-3/10-(1.2000000, 1.3000000)										
14	РЕЙТИНГ ЕДЫ-4/10-(1.3000000, 1.4000000)										
15	РЕЙТИНГ ЕДЫ-5/10-(1.4000000, 1.5000000)										
16	РЕЙТИНГ ЕДЫ-6/10-(1.5000000, 1.6000000)										
17	РЕЙТИНГ ЕДЫ-7/10-(1.6000000, 1.7000000)										
18	РЕЙТИНГ ЕДЫ-8/10-(1.7000000, 1.8000000)										
19	РЕЙТИНГ ЕДЫ-9/10-(1.8000000, 1.9000000)										
20	РЕЙТИНГ ЕДЫ-10/10-(1.9000000, 2.0000000)	90	0.150	0.359	0.447	0.031		-0.019	0.772	0.005	0.3
21	РЕЙТИНГ СЕРВИСА-1/10-(1.0000000, 1.1000000)	50	-0.733	-1.550	0.269	0.083		-0.054	-1.947	-0.014	0.4
22	РЕЙТИНГ СЕРВИСА-2/10-(1.1000000, 1.2000000)										
23	РЕЙТИНГ СЕРВИСА-3/10-(1.2000000, 1.3000000)										
24	РЕЙТИНГ СЕРВИСА-4/10-(1.3000000, 1.4000000)										
25	РЕЙТИНГ СЕРВИСА-5/10-(1.4000000, 1.5000000)										
26	РЕЙТИНГ СЕРВИСА-6/10-(1.5000000, 1.6000000)										
27	РЕЙТИНГ СЕРВИСА-7/10-(1.6000000, 1.7000000)										
28	РЕЙТИНГ СЕРВИСА-8/10-(1.7000000, 1.8000000)										
29	РЕЙТИНГ СЕРВИСА-9/10-(1.8000000, 1.9000000)										
30	РЕЙТИНГ СЕРВИСА-10/10-(1.9000000, 2.0000000)	67	0.433	0.597	-0.296	-0.090		0.049	-1.022	-0.007	0.4
	Сумма	23	-0.524	-1.571	0.281	-0.021		-0.008	-8.016		
	Среднее	04	-0.017	-0.052	0.009	-0.001		0.000		-0.002	
	Среднеквадратичное отклонение	11	0.172	0.421	0.128	0.036		0.021			0.2

Рисунок 11- Матрица информативности (модель INF1) в битах (фрагмент)

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами"										
Код признака	Наименование описательной шкалы и градации	135. ID ПОЛЬЗОВАТЕЛЯ U1135	136. ID ПОЛЬЗОВАТЕЛЯ U1136	137. ID ПОЛЬЗОВАТЕЛЯ U1137	138. ID ПОЛЬЗОВАТЕЛЯ U1138	139. ID КУХНЯ 1/3 (132560.0, 133409.7)	140. ID КУХНЯ 2/3 (133409.7, 134259.3)	141. ID КУХНЯ 3/3 (134259.3, 135109.0)	Сумма	Среднее
13	РЕЙТИНГ ЕДЫ-3/10-(1.2000000, 1.3000000)									
14	РЕЙТИНГ ЕДЫ-4/10-(1.3000000, 1.4000000)									
15	РЕЙТИНГ ЕДЫ-5/10-(1.4000000, 1.5000000)									
16	РЕЙТИНГ ЕДЫ-6/10-(1.5000000, 1.6000000)									
17	РЕЙТИНГ ЕДЫ-7/10-(1.6000000, 1.7000000)									
18	РЕЙТИНГ ЕДЫ-8/10-(1.7000000, 1.8000000)									
19	РЕЙТИНГ ЕДЫ-9/10-(1.8000000, 1.9000000)									
20	РЕЙТИНГ ЕДЫ-10/10-(1.9000000, 2.0000000)	89	1.154	4.205	1.246	7.254		-7.254		
21	РЕЙТИНГ СЕРВИСА-1/10-(1.0000000, 1.1000000)	04	-2.826	-5.435	0.552	16.698		-16.698		
22	РЕЙТИНГ СЕРВИСА-2/10-(1.1000000, 1.2000000)									
23	РЕЙТИНГ СЕРВИСА-3/10-(1.2000000, 1.3000000)									
24	РЕЙТИНГ СЕРВИСА-4/10-(1.3000000, 1.4000000)									
25	РЕЙТИНГ СЕРВИСА-5/10-(1.4000000, 1.5000000)									
26	РЕЙТИНГ СЕРВИСА-6/10-(1.5000000, 1.6000000)									
27	РЕЙТИНГ СЕРВИСА-7/10-(1.6000000, 1.7000000)									
28	РЕЙТИНГ СЕРВИСА-8/10-(1.7000000, 1.8000000)									
29	РЕЙТИНГ СЕРВИСА-9/10-(1.8000000, 1.9000000)									
30	РЕЙТИНГ СЕРВИСА-10/10-(1.9000000, 2.0000000)	03	3.242	6.656	-0.427	-16.073		16.073		
	Сумма									
	Среднее									
	Среднеквадратичное отклонение	03	0.878	2.535	0.372	6.992		6.992		

Рисунок 12- Матрица знаний (модель INF3) (фрагмент)

5.5. Модель: "7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1"											
Код признака	Наименование описательной шкалы и градации	135. ID ПОЛЬЗОВАТЕЛЯ U1135	136. ID ПОЛЬЗОВАТЕЛЯ U1136	137. ID ПОЛЬЗОВАТЕЛЯ U1137	138. ID ПОЛЬЗОВАТЕЛЯ U1138	139. ID КУХНЯ 1/3 (132560.0, 133409.7)	140. ID КУХНЯ 2/3 (133409.7, 134259.3)	141. ID КУХНЯ 3/3 (134259.3, 135109.0)	Сумма	Среднее	Средн. квадрат. откл.
13	РЕЙТИНГ ЕДЫ-3/10-(1.2000000, 1.3000000)										
14	РЕЙТИНГ ЕДЫ-4/10-(1.3000000, 1.4000000)										
15	РЕЙТИНГ ЕДЫ-5/10-(1.4000000, 1.5000000)										
16	РЕЙТИНГ ЕДЫ-6/10-(1.5000000, 1.6000000)										
17	РЕЙТИНГ ЕДЫ-7/10-(1.6000000, 1.7000000)										
18	РЕЙТИНГ ЕДЫ-8/10-(1.7000000, 1.8000000)										
19	РЕЙТИНГ ЕДЫ-9/10-(1.8000000, 1.9000000)										
20	РЕЙТИНГ ЕДЫ-10/10-(1.9000000, 2.0000000)	74	0.197	0.540	0.711	0.038		-0.022	15.184	0.108	0.4
21	РЕЙТИНГ СЕРВИСА-1/10-(1.0000000, 1.1000000)	61	-0.586	-0.845	0.381	0.105		-0.062	17.486	0.124	0.9
22	РЕЙТИНГ СЕРВИСА-2/10-(1.1000000, 1.2000000)										
23	РЕЙТИНГ СЕРВИСА-3/10-(1.2000000, 1.3000000)										
24	РЕЙТИНГ СЕРВИСА-4/10-(1.3000000, 1.4000000)										
25	РЕЙТИНГ СЕРВИСА-5/10-(1.4000000, 1.5000000)										
26	РЕЙТИНГ СЕРВИСА-6/10-(1.5000000, 1.6000000)										
27	РЕЙТИНГ СЕРВИСА-7/10-(1.6000000, 1.7000000)										
28	РЕЙТИНГ СЕРВИСА-8/10-(1.7000000, 1.8000000)										
29	РЕЙТИНГ СЕРВИСА-9/10-(1.8000000, 1.9000000)										
30	РЕЙТИНГ СЕРВИСА-10/10-(1.9000000, 2.0000000)	77	0.681	1.049	-0.299	-0.103		0.061	15.276	0.108	0.4
	Сумма	06	-0.080	0.689	0.702	0.001		-0.001	106.344		
	Среднее	04	-0.003	0.023	0.023	0.000		0.000		0.025	
	Среднеквадратичное отклонение	35	0.183	0.341	0.172	0.043		0.025			0.2

Рисунок 13 -Матрица знаний (модель INF4) (фрагмент)

1.5. Результаты верификации моделей

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

4.13.6. Обобщ.форма по достов.моделей при разн.инт.крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Итоги	L1-мера проф. Е.В. Луценко	Средний модуль уровней сходств истинно-полож. решений	Средний модуль уровней сходств истинно-отрицат. решений	Средний модуль уровней сходств ложно-положит. решений	Средний модуль уровней сходств ложно-отрицат. решений	A-Точность модели APrecision = ATP/(ATP+...	A-Полнота модели ARecall = ATP/(ATP+...	L2-мера проф. Е.В. Луценко	Процент правильно идентифицированных
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс. частот с обр...	1.000	0.040	0.701	0.078	0.566		0.553	1.000	0.712	100.
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс. частот по призна...	1.000	0.430	0.363		0.016		0.959	1.000	0.979	100.
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн. частот с о...	1.000	0.040	0.701	0.078	0.566		0.553	1.000	0.712	100.
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн. частот по приз...	1.000	0.036	0.549		0.481		0.533	1.000	0.695	100.
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн. частот с о...	1.000	0.040	0.701	0.078	0.566		0.553	1.000	0.712	100.
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн. частот по приз...	1.000	0.036	0.503		0.434		0.537	1.000	0.699	100.
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	0.694	0.043	0.437	0.411	0.406	0.412	0.518	0.515	0.517	68.
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	0.804	0.031	0.154	0.221	0.200	0.078	0.435	0.662	0.525	67.
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	0.698	0.043	0.445	0.408	0.411	0.405	0.520	0.524	0.522	67.
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Сумма знаний	0.809	0.033	0.145	0.198	0.185	0.071	0.440	0.672	0.532	67.
6. INF3 - частный критерий: Хинквадрат, разности между факти...	Семантический резонанс зна...	0.745	0.047	0.529	0.463	0.459	0.422	0.536	0.556	0.546	69.
6. INF3 - частный критерий: Хинквадрат, разности между факти...	Сумма знаний	0.625	0.133	0.294	0.082	0.080	0.412	0.787	0.417	0.545	69.
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.755	0.041	0.521	0.344	0.469	0.388	0.526	0.573	0.549	69.
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.898	0.029	0.175	0.118	0.221	0.045	0.442	0.794	0.567	69.
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.759	0.042	0.530	0.336	0.480	0.379	0.525	0.583	0.553	69.
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.900	0.031	0.214	0.140	0.259	0.054	0.452	0.798	0.577	69.
9. INF6 - частный критерий: разн. усл.и безуслов. вероятностей; вер...	Семантический резонанс зна...	0.756	0.041	0.522	0.349	0.470	0.392	0.526	0.571	0.548	69.
9. INF6 - частный критерий: разн. усл.и безуслов. вероятностей; вер...	Сумма знаний	0.897	0.029	0.175	0.122	0.222	0.047	0.441	0.789	0.565	69.
10. INF7 - частный критерий: разн. усл.и безуслов. вероятностей; ве...	Семантический резонанс зна...	0.759	0.042	0.530	0.347	0.475	0.387	0.527	0.578	0.552	69.
10. INF7 - частный критерий: разн. усл.и безуслов. вероятностей; ве...	Сумма знаний	0.898	0.031	0.232	0.154	0.281	0.059	0.452	0.796	0.576	69.

Помощь по меркам достоверности | Помощь по частным 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 INF5 при интегральном критерии «Сумма знаний». При этом точность модели составляет 0,798. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием.

На рисунке 15 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF5.

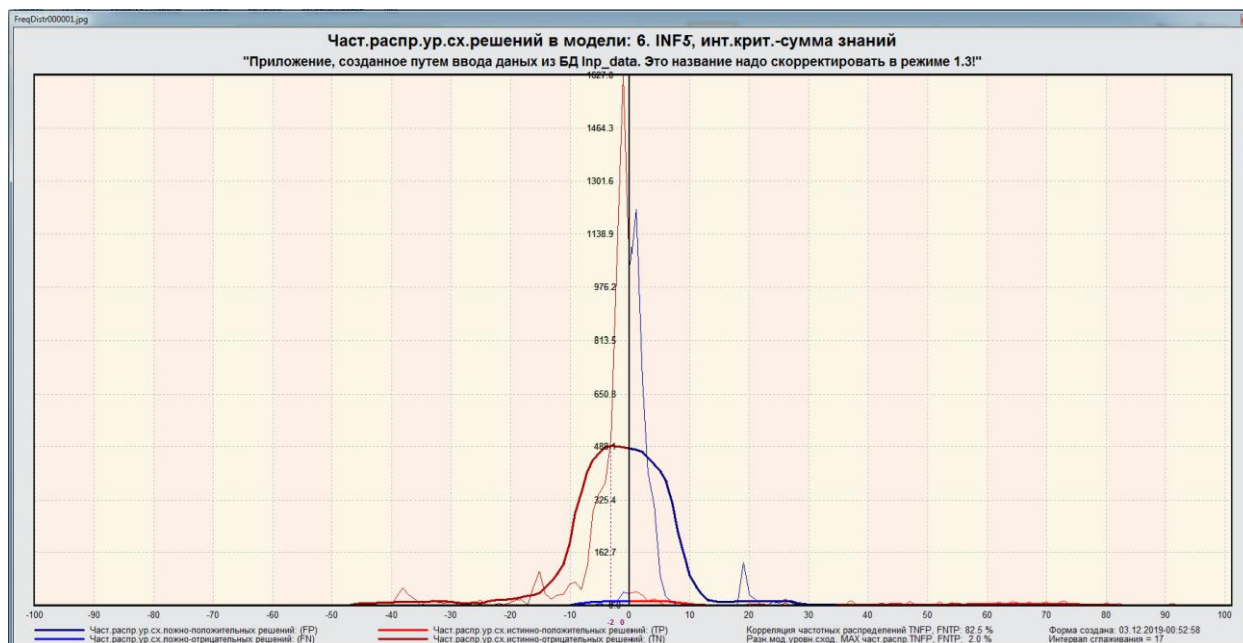


Рисунок 15 – Частотное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF5

Как можно видеть, при модели INF5 непринадлежность объекта к классу лучше определяется, чем принадлежность. При этом существует вероятность ложного определения непринадлежности или принадлежности.

Выберем модель INF5 в качестве основной и перейдем к решению задач идентификации.

2.РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи

В соответствии с технологией АСК-анализа зададим текущей модель INF5 (режим 5.6) и проведем пакетное распознавание в режиме 4.2.1, как показано на рисунке 16.

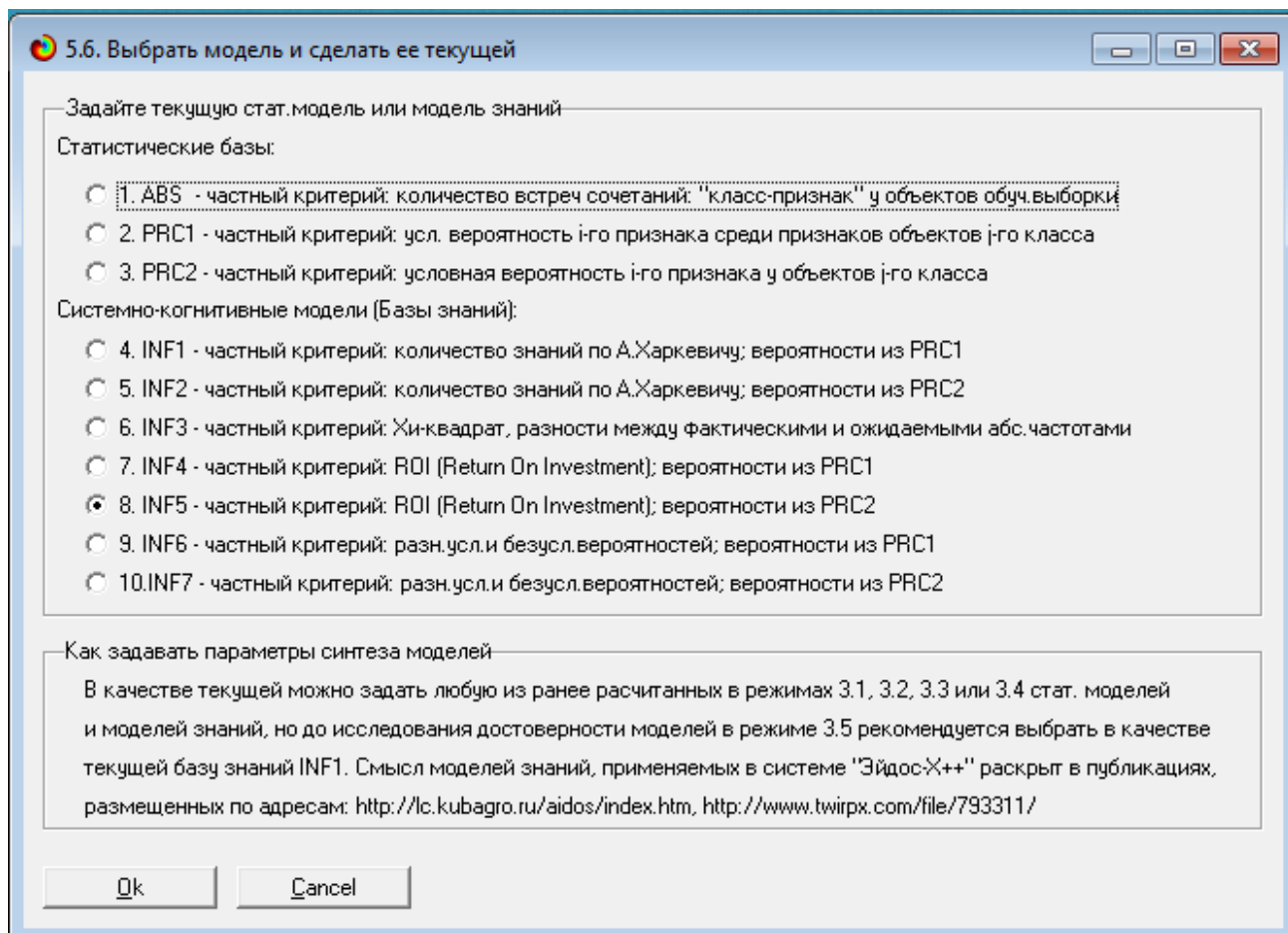


Рисунок 16 - Экранная форма режима задания модели в качестве текущей

В результате установки модели текущей пользователь увидит форму, представленную на рисунке 17.

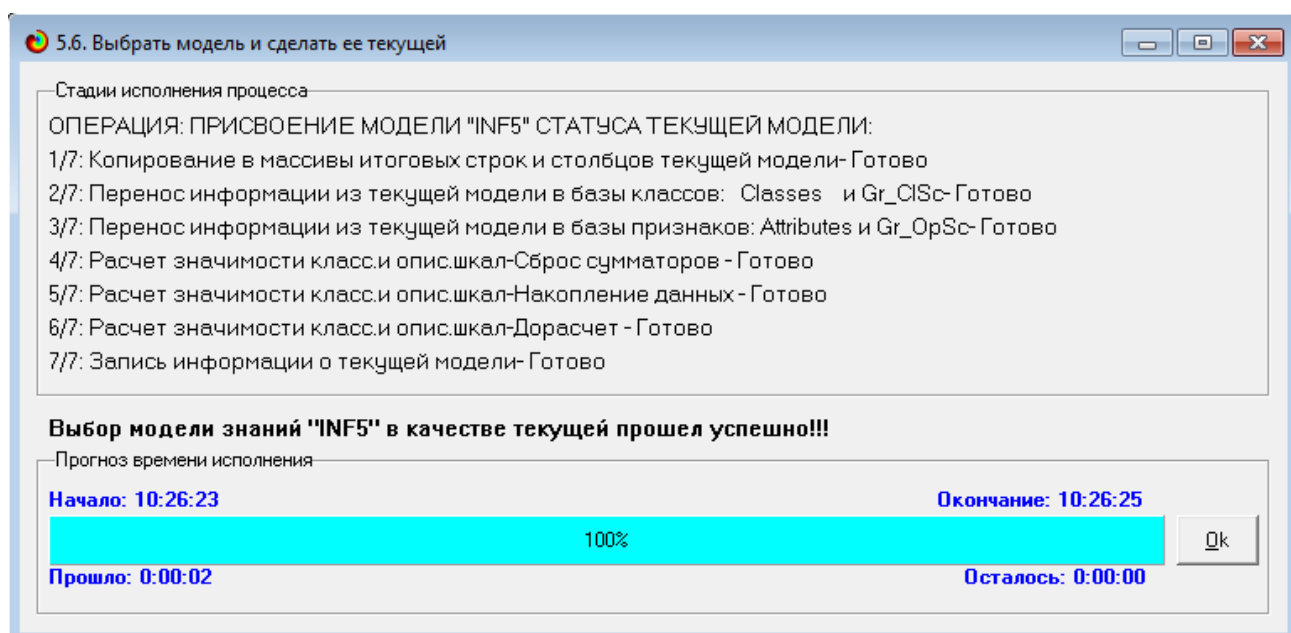


Рисунок 17 – Присвоение модели статуса текущей модели

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

Режим 4.1.3 системы «Эйдос» 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.

Для наглядного примера, кратко рассмотрим некоторые из них.

На рисунках 18 и 19 приведены примеры прогнозов высокой и низкой достоверности частоты и классов в наиболее достоверной модели INF5 на основе описательных шкал.

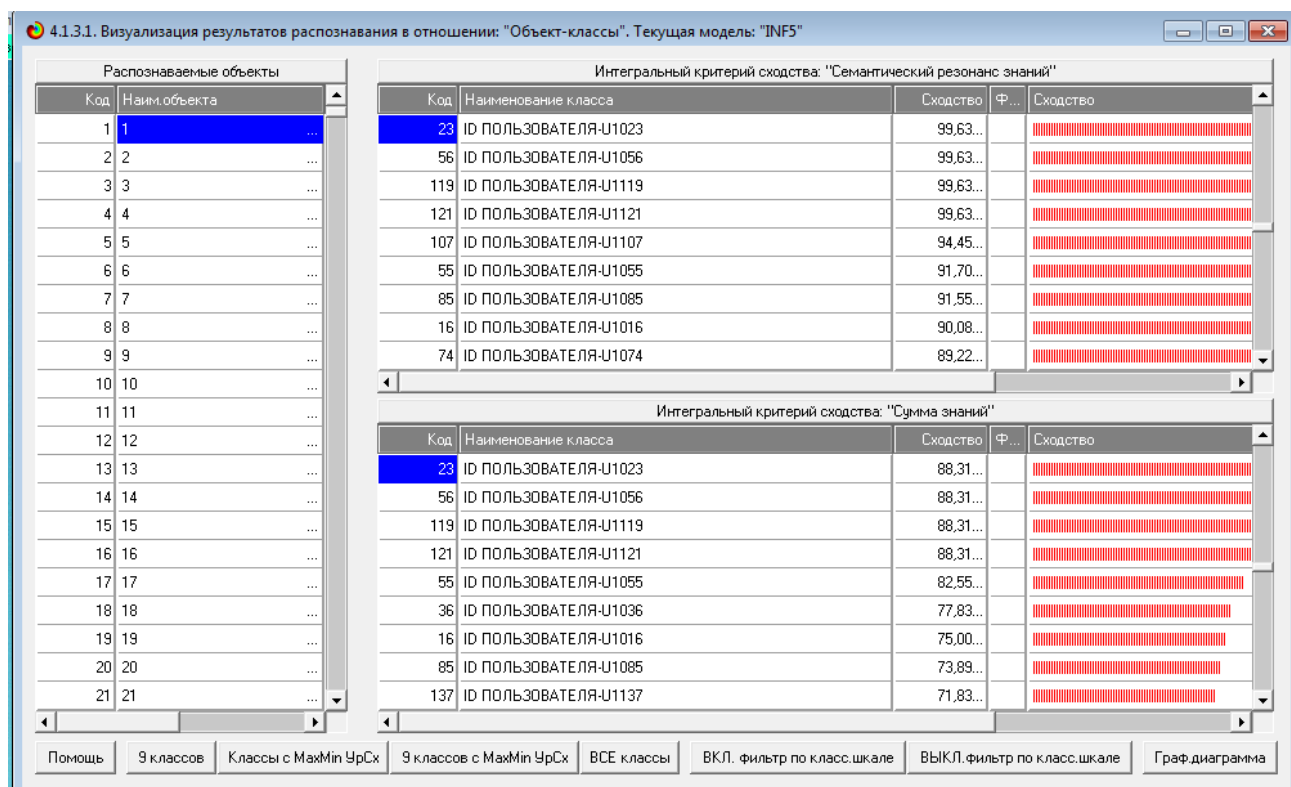


Рисунок 18 – Прогнозы достоверности классов объектов

4.1.3.3. Итоговая наглядная форма результатов распознавания: "Объект-класс". Текущая модель: "INF5"

Код Интегральный критерий

1 Семантический резонанс знаний

2 Сумма знаний

Пояснения по смыслу частных и интегральных критериев

Частн. крит.: 7 моделей знаний

Инт. крит.: "Сумма знаний"

Инт. крит.: "Резонанс знаний"

Код объекта распознаем. выборки	Наименование объекта распознаваемой выборки	Код класса с МАХ ур. сход.	Наименование класса с МАХ уровнем сходства	МАХ уровень сходства	Код класса с MIN ур. сход.	Наименование класса с MIN уровнем сходства	MIN уровень сходства
187	187	22	ID ПОЛЬЗОВАТЕЛЯ-U1022 ...	81.777	139	ID КУХНЯ-1/3-{132560.0000000, ...	-72.466
188	188	130	ID ПОЛЬЗОВАТЕЛЯ-U1130 ...	99.601	25	ID ПОЛЬЗОВАТЕЛЯ-U1025 ...	-73.955
189	189	22	ID ПОЛЬЗОВАТЕЛЯ-U1022 ...	81.777	139	ID КУХНЯ-1/3-{132560.0000000, ...	-72.466
190	190	22	ID ПОЛЬЗОВАТЕЛЯ-U1022 ...	81.777	139	ID КУХНЯ-1/3-{132560.0000000, ...	-72.466
191	191	102	ID ПОЛЬЗОВАТЕЛЯ-U1102 ...	99.819	64	ID ПОЛЬЗОВАТЕЛЯ-U1064 ...	-82.433
192	192	130	ID ПОЛЬЗОВАТЕЛЯ-U1130 ...	99.601	25	ID ПОЛЬЗОВАТЕЛЯ-U1025 ...	-73.955
193	193	120	ID ПОЛЬЗОВАТЕЛЯ-U1120 ...	82.803	97	ID ПОЛЬЗОВАТЕЛЯ-U1097 ...	-66.660
194	194	21	ID ПОЛЬЗОВАТЕЛЯ-U1021 ...	94.318	32	ID ПОЛЬЗОВАТЕЛЯ-U1032 ...	-64.614
195	195	22	ID ПОЛЬЗОВАТЕЛЯ-U1022 ...	81.777	139	ID КУХНЯ-1/3-{132560.0000000, ...	-72.466
196	196	22	ID ПОЛЬЗОВАТЕЛЯ-U1022 ...	81.777	139	ID КУХНЯ-1/3-{132560.0000000, ...	-72.466
197	197	21	ID ПОЛЬЗОВАТЕЛЯ-U1021 ...	94.318	32	ID ПОЛЬЗОВАТЕЛЯ-U1032 ...	-64.614
198	198	134	ID ПОЛЬЗОВАТЕЛЯ-U1134 ...	75.287	46	ID ПОЛЬЗОВАТЕЛЯ-U1046 ...	-69.231
199	199	23	ID ПОЛЬЗОВАТЕЛЯ-U1023 ...	99.636	18	ID ПОЛЬЗОВАТЕЛЯ-U1018 ...	-81.926
200	200	130	ID ПОЛЬЗОВАТЕЛЯ-U1130 ...	99.601	25	ID ПОЛЬЗОВАТЕЛЯ-U1025 ...	-73.955
201	201	114	ID ПОЛЬЗОВАТЕЛЯ-U1114 ...	98.934	131	ID ПОЛЬЗОВАТЕЛЯ-U1131 ...	-92.231
202	202	29	ID ПОЛЬЗОВАТЕЛЯ-U1129 ...	84.958	126	ID ПОЛЬЗОВАТЕЛЯ-U1126 ...	-60.147

Помощь В начало БД В конец БД Предыдущая Следующая

Рисунок 19 – Итоги распознавания классов объектов

2.2 Визуализация когнитивных функций

Далее стоит рассмотреть режим работы 4.5, в котором реализована возможность визуализации когнитивных функций для любых моделей и любых сочетаний классификационных и описательных шкал. Начало работы в этом режиме начинается с понятия когнитивных функций, а также предоставления возможности перехода на статьи по данной проблематике и скачивания публикаций, как это продемонстрировано на рисунке 20.

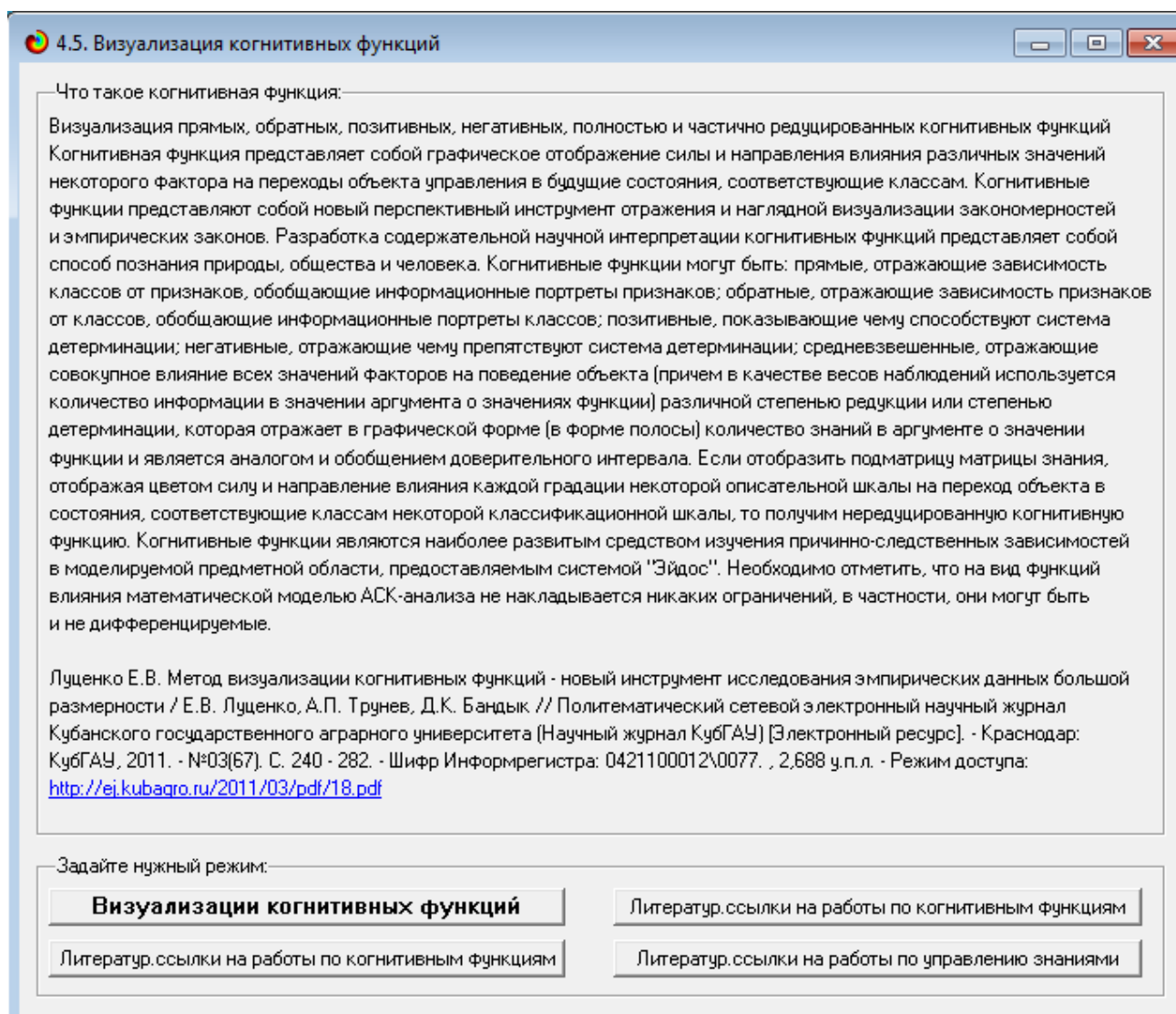


Рисунок 20 - Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»

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. Когнитивным функциям посвящено множество работ, 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 INF5.

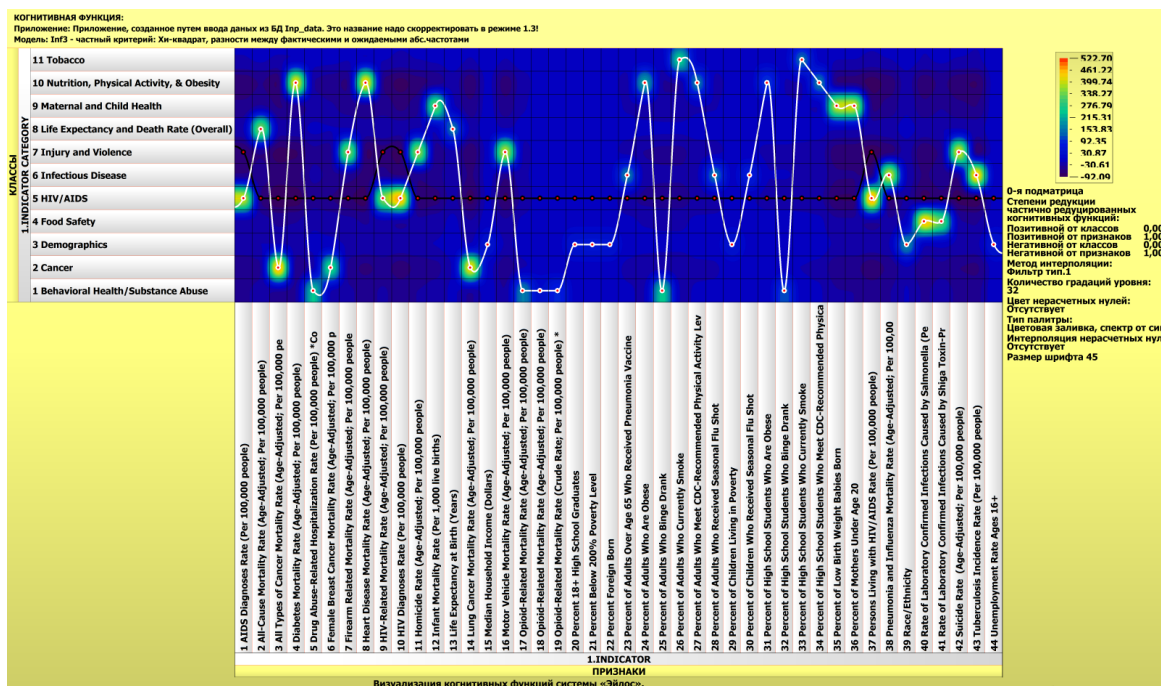


Рисунок 21 – Визуализация зависимости рейтинга от качества

Рассмотрим визуализацию когнитивных функций в рамках определенных задач.

2.3 SWOTи PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков), источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают неформализуемым путем (интуитивно), на основе своего профессионального опыта и компетенции. Но возможности экспертов имеют свои ограничения и часто по различным причинам они не могут и не

хотят это сделать. Таким образом, возникает проблема проведения SWOT-анализа без привлечения экспертов. Эта проблема может решаться путем автоматизации функций экспертов, т.е. путем измерения силы и направления влияния факторов непосредственно на основе эмпирических данных.

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).

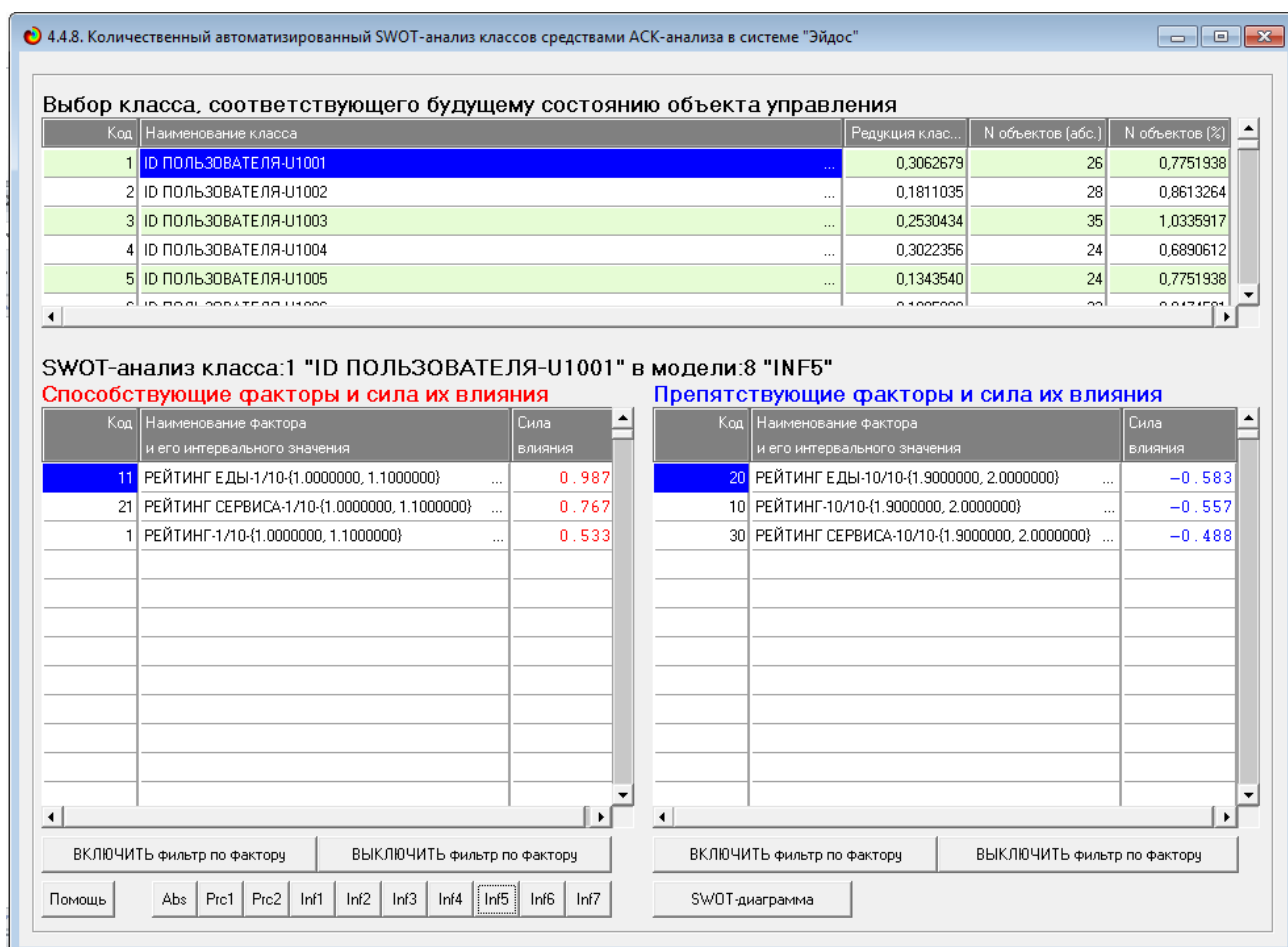


Рисунок 22 – Пример SWOT-Матрицы в модели INF5

Следующим шагом является построение SWOT-диаграммы, для наглядной демонстрации работы, продемонстрированной на рисунке 23.

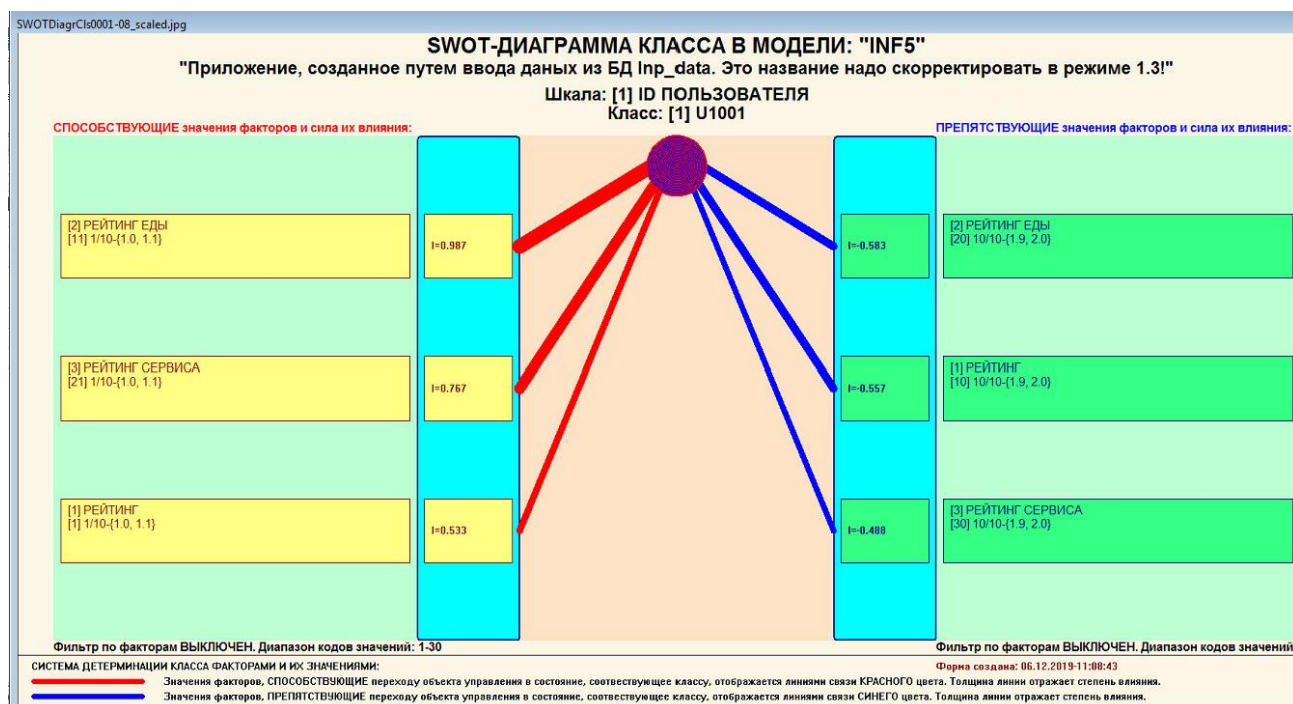


Рисунок 23 –SWOT-диаграмма

2.4 Нелокальные нейронные сети и нейтроны

Каждому классу системно-когнитивной модели соответствует нелокальный нейрон, совокупность которых образует не локальную нейронную сеть. Рассмотрим пример на рисунке 24.

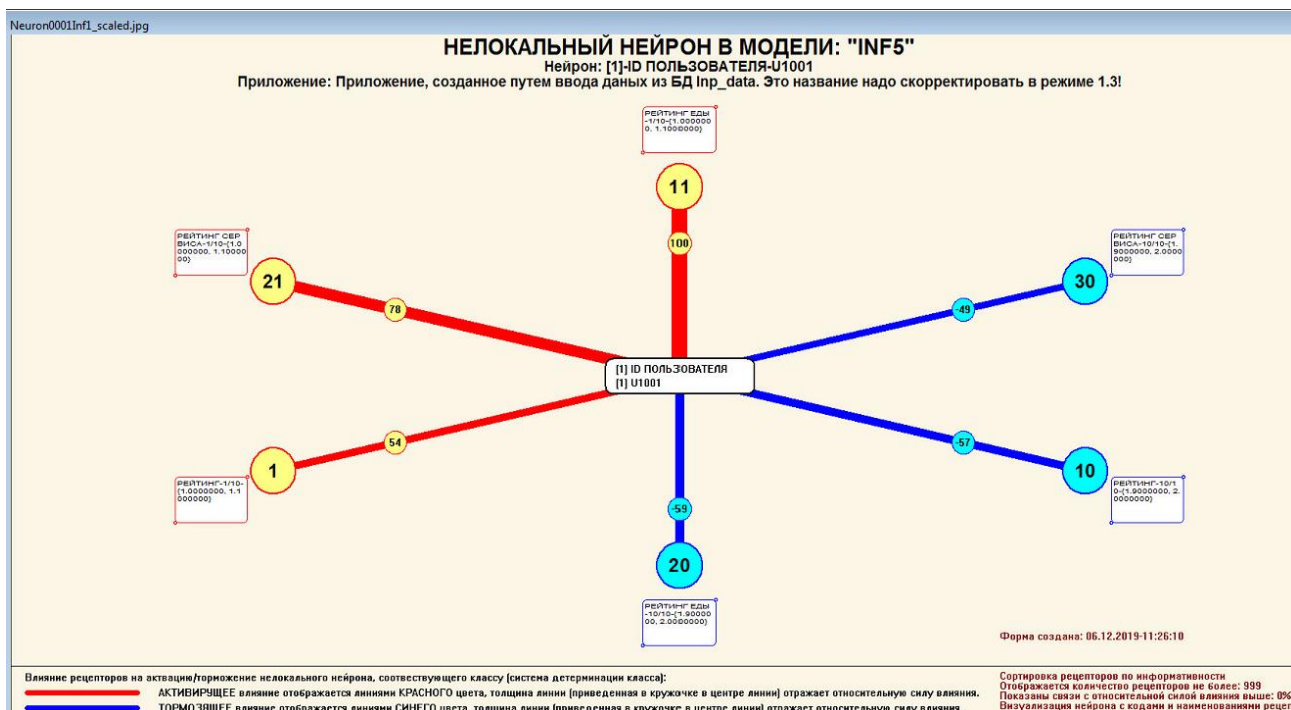


Рисунок 24 – Нейрон для модели INF5

Благодаря данному нейрону видно, на что обращает внимание посетитель при составлении рейтинга кухонь влияют, какие оказывают активизирующее влияние, а какие тормозящее.

На основании нелокальных нейронов строится нейронная сеть, пример которой представлен на рисунке 25.

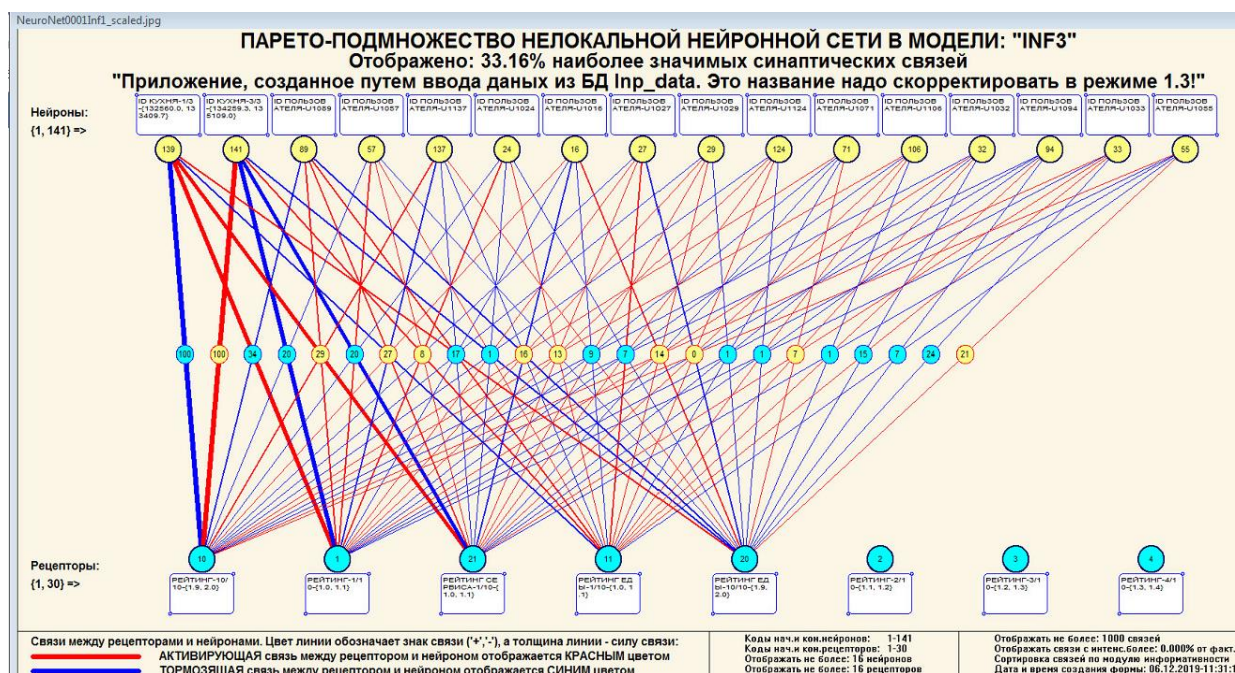


Рисунок 25 – Пример нелокальной нейронной сети в модели INF5

2.5 Кластерный и конструктивный анализ

В режиме 4.2.2.2, после расчета матриц сходства, кластеров и конструкторов, строим 2D сеть классов в выбранной модели знаний, для наглядного представления сети классов.

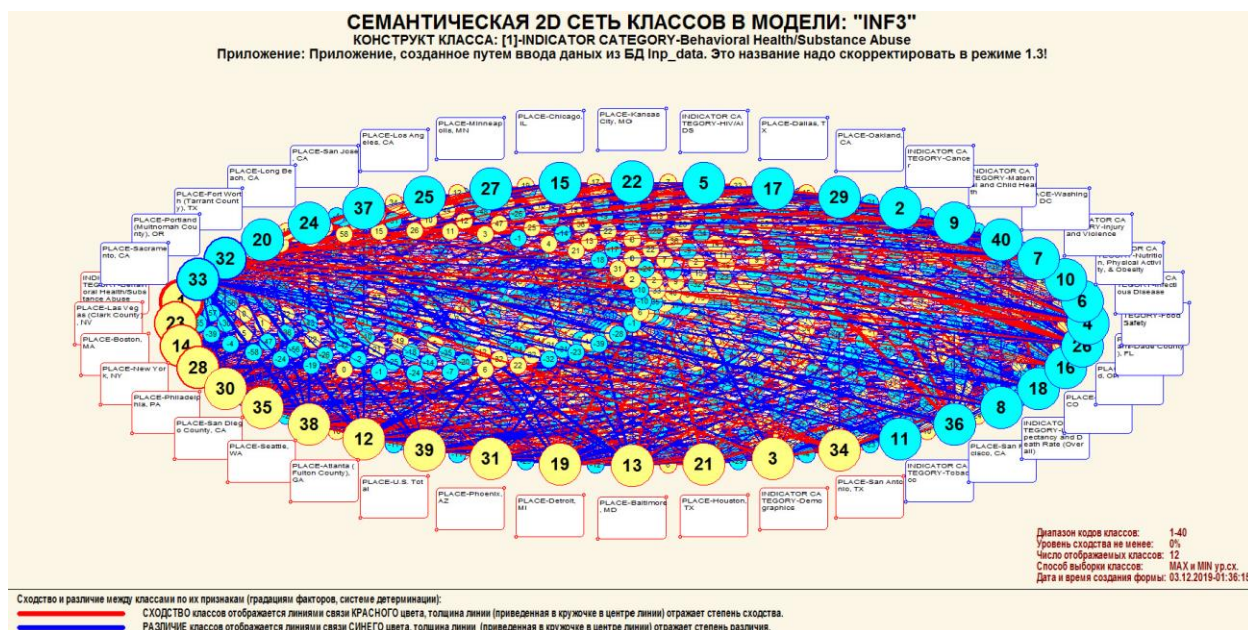


Рисунок 26 – Семантическая 2D сеть классов

ЗАКЛЮЧЕНИЕ

В связи с существованием множества 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 репозитория «UCI».

В данной курсовой работе приводится развернутый пример использования базы данных репозитория «UCI» для оценки качества математических моделей, применяемых в АСК-анализе и его программном инструментарии системе искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении оказались модели INF5, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,798, что выше, чем достоверность экспертных оценок, которая считается равной около 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием.

На основе базы данных «ICU», рассмотренной в данной курсовой работе, построен модели прогнозирования с помощью АСК-анализа и реализующей его системы «Эйдос».

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

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

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

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

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

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

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

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

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

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