

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ

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

Факультет прикладной информатики
Кафедра компьютерных технологий и систем

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к курсовой работе

по дисциплине: ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ
СИСТЕМЫ

на тему:

**Решение задач идентификации и прогнозирования с применением АСК-
анализа и системы «Эйдос» на основе данных репозитория UCI: Airlines
Delay**

выполнил студент группы ИТ1421 **Стригин Вячеслав Сергеевич**

Допущена к защите _____

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

(подпись, расшифровка подписи)

Нормоконтролер _____ **Николаева Ирина Валентиновна, к.т.н.,**
доцент

(подпись, расшифровка подписи)

Защищена _____

Оценка

(дата)

Члены комиссии _____ **В.И. Лойко**

_____ **Е.В. Луценко**

_____ **И. В. Николаева**

(подпись, дата, расшифровка подписи)

Краснодар

2017 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

Студентке: ИТ1421 группы 3 курса
Факультета прикладной
информатики
Специальности: 09.03.02 Информационные системы и
технологии

(шифр)

Стригин Вячеслав Сергеевич

(Ф.И.О.)

Тема проекта: **Решение задач идентификации и прогнозирования с**
применением АСК-анализа и системы «Эйдос» на основе данных
репозитория UCI: Airlines Delay

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

Объем работы:

а) пояснительная записка к работе _____ листа формата
А4

б) графическая часть _____ лист формата А4

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

Срок выполнения проекта: с « » _____ по « » _____ 2017 г.

Срок защиты: « » _____ 2017 г.

Дата выдачи задания: « » _____ 2017 г.

Дата сдачи проекта на кафедру: « » _____ 2017 г.

Руководитель проекта: Луценко Евгений Вениаминович

д.э.н., к.т.н., профессор
(подпись, Ф.И.О., звание, степень)

Задание принял студент _____ « » _____ 2017 г.
(подпись, дата)

Краснодар
2017 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	7
1.1. Описание решения	7
1.2. Подготовка данных в промежуточные файлы MS Excel	7
1.3. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос"	10
1.4. Синтез и верификация статистических и интеллектуальных моделей	14
1.5. Виды моделей системы «Эйдос»	16
1.6. Результаты верификации моделей	18
2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	24
2.1. Решение задачи	24
2.2. Когнитивные функции	27
2.3. SWOT и PEST матрицы и диаграммы	31
2.4. Нелокальные нейроны и нейронная сеть	34
2.5. Кластерный и конструктивный анализ классов	36
ЗАКЛЮЧЕНИЕ	37
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	39

ВВЕДЕНИЕ

На сегодняшний день создание систем искусственного интеллекта – это перспективный шаг к расширению возможностей компьютерных наук и автоматизации разумного поведения, опирающийся на теоретические и прикладные принципы. Однако возникает необходимость оценки качества математических моделей этих систем.

В данной курсовой работе рассмотрено решение задач идентификации и прогнозирования с применением АСК-анализа и системы «Эйдос» на основе данных репозитория UCI: Airlines Delay.

Целью работы является разработка системно-когнитивной модели, основывающейся на следующих признаках: дни месяца, недели, названии борта, время полета и дистанция.

Задачами курсовой работы являются:

- 1) систематизация, закрепление и расширение теоретических и практических знаний по дисциплине "Интеллектуальные информационные системы и технологии";
- 2) изучение интеллектуальной информационной системы "Эйдос";
- 3) решение поставленной цели с помощью интеллектуальной информационной системы "Эйдос".

Объектом исследования является база данных "Airlines Delay".

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

1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ

1.1. Описание решения

В данной работе использована база данных "Airlines Delay", предоставленная репозиторием UCI. База данных дополнена значениями и приводится полностью в приложении.

На момент написания курсовой работы этот банк включает в себя более 30 тысяч исходных данных по различным задачам, однако рассмотрена будет лишь часть усредненных данных. Решение поставленной задачи проведем в четыре этапа:

1. Подготовка 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 Подготовка данных в промежуточные файлы MS Excel

Из электронного репозитория UCI баз данных возьмем базу данных полетов на самолетах – «Airlines Delay.xls», которую оставим без изменений.

Общее описание задачи:

1. ID
2. Origin
3. Dest
4. DayOfMouth
5. DayOfWeek
6. TailNum
7. AirTime
8. Distance

Столбцы 2-3 являются классификационными шкалами.

Обучающая выборка:

Таблица 1 – Airlines Delay.xls

ID	Origin	Dest	DayofMonth	DayOfWeek	TailNum	AirTime	Distance
0	IAD	TPA	3	4	N712SW	116.0	810
1	IAD	TPA	3	4	N772SW	113.0	810
2	IND	BWI	3	4	N428WN	76.0	515
4	IND	BWI	3	4	N464WN	77.0	515
5	IND	JAX	3	4	N726SW	87.0	688
6	IND	LAS	3	4	N763SW	230.0	1591
10	IND	MCO	3	4	N690SW	106.0	828
11	IND	MCO	3	4	N334SW	107.0	828
15	IND	MDW	3	4	N263WN	37.0	162
16	IND	PHX	3	4	N286WN	213.0	1489
17	IND	PHX	3	4	N778SW	205.0	1489
18	IND	TPA	3	4	N674AA	110.0	838
19	ISP	BWI	3	4	N643SW	49.0	220
21	ISP	BWI	3	4	N724SW	47.0	220
22	ISP	BWI	3	4	N786SW	49.0	220
23	ISP	BWI	3	4	N714CB	47.0	220
25	ISP	BWI	3	4	N394SW	50.0	220
26	ISP	FLL	3	4	N215WN	143.0	1093
27	ISP	FLL	3	4	N243WN	155.0	1093
30	ISP	MCO	3	4	N798SW	134.0	972
33	ISP	MCO	3	4	N247WN	140.0	972
34	ISP	MCO	3	4	N707SA	134.0	972
35	ISP	MCO	3	4	N443WN	134.0	972
37	ISP	MDW	3	4	N779SW	118.0	765
38	ISP	MDW	3	4	N704SW	114.0	765
39	ISP	MDW	3	4	N709SW	113.0	765
40	ISP	PBI	3	4	N459WN	143.0	1052
41	ISP	PBI	3	4	N621SW	138.0	1052
42	ISP	PBI	3	4	N206WN	140.0	1052
43	ISP	RSW	3	4	N280WN	151.0	1101
44	ISP	TPA	3	4	N241WN	143.0	1034
45	ISP	TPA	3	4	N200WN	144.0	1034
48	JAN	BWI	3	4	N475WN	121.0	888
49	JAN	HOU	3	4	N792SW	59.0	359
52	JAN	HOU	3	4	N252WN	60.0	359
53	JAN	MCO	3	4	N346SW	81.0	587
54	JAN	MDW	3	4	N299WN	97.0	666
56	JAX	BHM	3	4	N487WN	59.0	365
57	JAX	BNA	3	4	N243WN	77.0	484
58	JAX	BNA	3	4	N479WN	72.0	484
64	JAX	FLL	3	4	N726SW	54.0	318
69	JAX	FLL	3	4	N467WN	49.0	318

70	JAX	HOU	3	4	N643SW	125.0	816
71	JAX	IND	3	4	N263WN	104.0	688
74	JAX	PHL	3	4	N521SW	113.0	742
75	JAX	PHL	3	4	N663SW	120.0	742
76	JAX	TPA	3	4	N376SW	39.0	180
77	JAX	TPA	3	4	N791SW	36.0	180
78	JAX	TPA	3	4	N392SW	38.0	180
79	LAS	ABQ	3	4	N799SW	60.0	487
81	LAS	ABQ	3	4	N618WN	60.0	487
82	LAS	ABQ	3	4	N682SW	65.0	487
84	LAS	ABQ	3	4	N509SW	60.0	487
86	LAS	ALB	3	4	N473WN	243.0	2237
87	LAS	AMA	3	4	N351SW	91.0	758
88	LAS	AUS	3	4	N238WN	131.0	1090
89	LAS	AUS	3	4	N499WN	127.0	1090
90	LAS	AUS	3	4	N309SW	135.0	1090
91	LAS	BDL	3	4	N245WN	253.0	2298
94	LAS	BNA	3	4	N275WN	178.0	1588
95	LAS	BNA	3	4	N271WN	183.0	1588
96	LAS	BNA	3	4	N434WN	177.0	1588
98	LAS	BOI	3	4	N619SW	73.0	520
99	LAS	BUF	3	4	N290WN	221.0	1987
101	LAS	BUR	3	4	N383SW	46.0	223
102	LAS	BUR	3	4	N608SW	46.0	223
104	LAS	BUR	3	4	N626SW	48.0	223
105	LAS	BUR	3	4	N369SW	43.0	223
106	LAS	BUR	3	4	N396SW	50.0	223
107	LAS	BUR	3	4	N293	46.0	223
108	LAS	BUR	3	4	N509SW	47.0	223
109	LAS	BUR	3	4	N409WN	46.0	223
111	LAS	BUR	3	4	N695SW	47.0	223
113	LAS	BWI	3	4	N220WN	244.0	2106
114	LAS	BWI	3	4	N784SW	230.0	2106
116	LAS	CLE	3	4	N244WN	213.0	1825
117	LAS	CMH	3	4	N455WN	195.0	1772
118	LAS	CMH	3	4	N297WN	201.0	1772
120	LAS	DEN	3	4	N481WN	79.0	629
121	LAS	DEN	3	4	N203WN	77.0	629
122	LAS	DEN	3	4	N491WN	88.0	629
123	LAS	DEN	3	4	N480WN	78.0	629
124	LAS	DEN	3	4	N318SW	85.0	629
125	LAS	DEN	3	4	N647SW	80.0	629
126	LAS	ELP	3	4	N313SW	78.0	584
127	LAS	ELP	3	4	N352SW	73.0	584
128	LAS	GEG	3	4	N321SW	116.0	806
129	LAS	GEG	3	4	N660SW	116.0	806

130	LAS	HOU	3	4	N414WN	143.0	1235
131	LAS	HOU	3	4	N408WN	147.0	1235
132	LAS	HOU	3	4	N475WN	146.0	1235
133	LAS	HOU	3	4	N321SW	150.0	1235
134	LAS	IAD	3	4	N230WN	232.0	2066
135	LAS	IND	3	4	N761RR	176.0	1591
136	LAS	IND	3	4	N464WN	179.0	1591
137	LAS	ISP	3	4	N272WN	269.0	2283
138	LAS	LAX	3	4	N499WN	42.0	236
139	LAS	LAX	3	4	N244WN	40.0	236
140	LAS	LAX	3	4	N335SW	41.0	236
141	LAS	LAX	3	4	N488WN	42.0	236

1.3. Преобразование исходных данных из промежуточных файлов MS Excel в базы данных системы "Эйдос".

Для преобразования исходных данных обучающей выборки в базы данных системы "Эйдос" необходимо файл Microsoft Office Excel, который содержит базу данных перелетов скопировать в Aidos-X\AID_DATA\Inp_data и назвать Inp_data.xlsx.

Далее запускаем систему "Эйдос" из папки "Aidos-X" файлом _aidos-x.exe. Система попросит ввести логин и пароль. Необходимо ввести: логин – 1, пароль – 1. После откроется главное окно программы.

Для загрузки базы данных необходимо зайти в режим 2.3.2.2 и в настройках изменить (рисунок 1):

- "Задайте тип файла исходных данных Inp_data": "XLSX-MSExcel-2003";
- "Задайте диапазон столбцов классификационных шкал": "Начальный столбец классификационных шкал" – 2, "Конечный столбец классификационных шкал" – 3;
- "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 4, "Конечный столбец описательных шкал" – 8;

– "Задание параметров формирования сценариев или способа интерпретации текстовых полей": "Не применять сценарный метод АСК-анализа" и Интерпретацию TXT-полей".

Рисунок 1 – Универсальный программный интерфейс импорта данных в систему "Эйдос"

После нажать кнопку "Ок". Далее открывается окно, где размещена информация о размерности модели (рисунок 2). В этом окне необходимо задать число интервалов равным 5 и нажать кнопку "Выйти на создание модели".

2.3.2.2. Задание размерности модели системы "Эйдос-Х++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [129 x 856]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	0	0	0,00	3	15	5,00
Текстовые	2	129	64,50	2	841	420,50
ВСЕГО:	2	129	64,50	5	856	171,20

Задайте число интервалов (градаций) в шкале:

В описательных шкалах:

Пересчитать шкалы и градации

Выйти на создание модели

Рисунок 2 – Задание размерности модели системы "Эйдос"

Далее открывается окно, где происходит процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос" (рисунок 3). В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "Ок".

2.3.2.2. Процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос-Х++"

Стадии исполнения процесса

1/3: Формирование классификационных и описательных шкал и градаций на основе БД "Inp_data" - Готово

2/3: Генерация обучающей выборки и базы событий "EventsKO" на основе БД "Inp_data" - Готово

3/3: Переиндексация всех баз данных нового приложения - Готово

ПРОЦЕСС ФОРМАЛИЗАЦИИ ПРЕДМЕТНОЙ ОБЛАСТИ ЗАВЕРШЕН УСПЕШНО !!!

Прогноз времени исполнения

Начало: 11:50:48

Окончание: 11:51:29

100%

Прошло: 0:00:40

Осталось: 0:00:00

Ok

Рисунок 3 – Процесс импорта данных из внешней БД "Inp_data" в систему "Эйдос"

В результате этого процесса будут сформированы классификационные и описательные шкалы и градации. Их применение позволит закодировать исходные данные и представить в форме эвентологических баз данных. Для просмотра классификационных шкал и градаций необходимо запустить

режим 2.1 (рисунок 4). Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 5). На рисунке 6 представлена обучающая выборка.

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

Код шкалы	Наименование классификационной шкалы	Код градации	Наименование градации классификационной шкалы	DEL
1	ORIGIN	1	ABE	
2	DEST	2	ABI	
		3	ABQ	
		4	ABY	
		5	ACK	
		6	ACT	
		7	ACV	
		8	ACY	
		9	ADK	
		10	ADQ	
		11	AEX	
		12	AGS	
		13	AKN	
		14	ALB	
		15	ALO	
		16	AMA	
		17	ANC	
		18	ASE	
		19	ATL	
		20	ATW	
		21	AUS	
		22	AVL	

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

Рисунок 4 – Классификационные шкалы и градации

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

Код шкалы	Наименование описательной шкалы	Код градации	Наименование градации описательной шкалы
1	MONTH	1	1/5-(1.0000000, 2.0000000)
2	DAYOFMONTH	2	2/5-(2.0000000, 3.0000000)
3	DAYOFWEEK	3	3/5-(3.0000000, 4.0000000)
4	TAILNUM	4	4/5-(4.0000000, 5.0000000)
5	AIRTIME	5	5/5-(5.0000000, 6.0000000)
6	DISTANCE		

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

Рисунок 5 – Описательные шкалы и градации

2.3.1. Ручной ввод-корректировка обучающей выборки. Текущая модель: "INF4"

Код объекта	Наименование объекта	Дата	Время
1	0	...	
2	1	...	
3	2	...	
4	4	...	
5	5	...	
6	6	...	
7	10	...	
8	11	...	
9	15	...	
10	16	...	

Код объекта	Класс 1	Класс 2	Класс 3	Класс 4
1	23	127	0	0

Код объекта	Признак 1	Признак 2	Признак 3	Признак 4	Признак 5	Признак 6	Признак 7
1	1	8	436	545	853	0	0

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

Рисунок 6 – Обучающая выборка (фрагмент)

Таким образом, созданы все необходимые и достаточные предпосылки для выявления силы и направления причинно-следственных связей между значениями факторов и результатами их совместного системного воздействия (с учетом нелинейности системы).

1.4. Синтез и верификация статистических и интеллектуальных моделей

Далее запускаем режим 3.5, в котором задаются модели для синтеза и верификации, а также задается модель, которой по окончании режима присваивается статус текущей (рисунок 7).

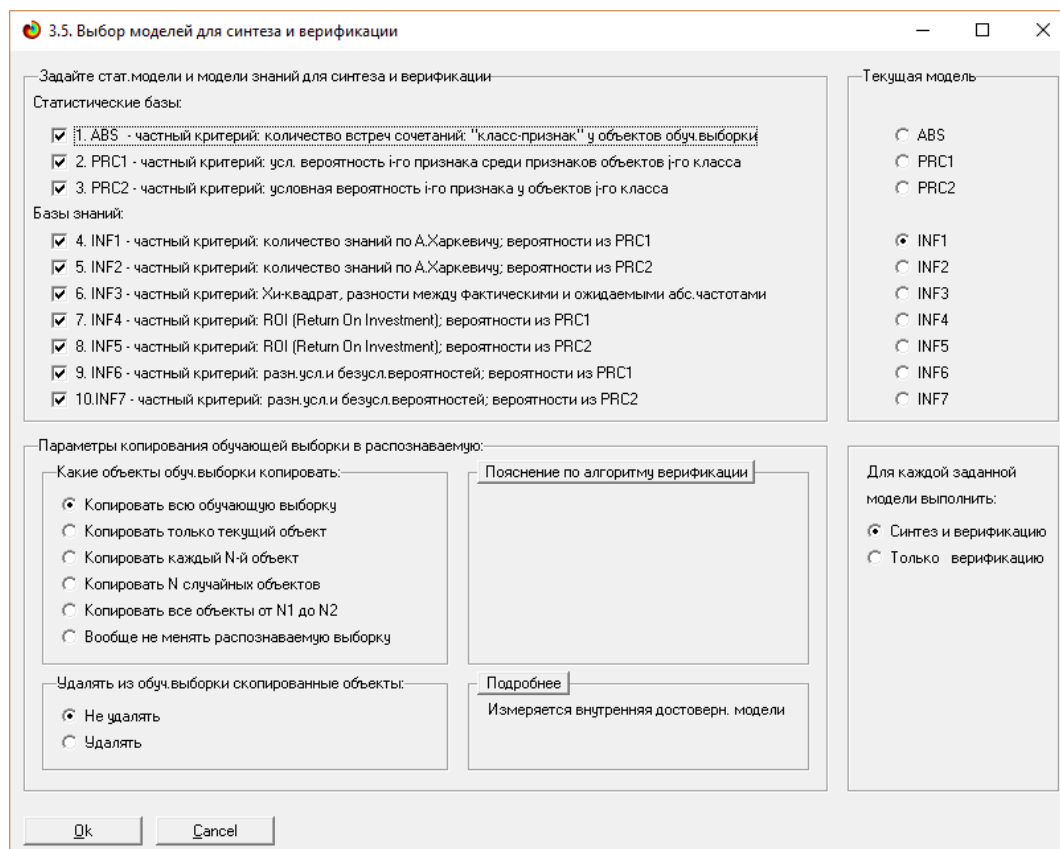


Рисунок 7 – Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 8.

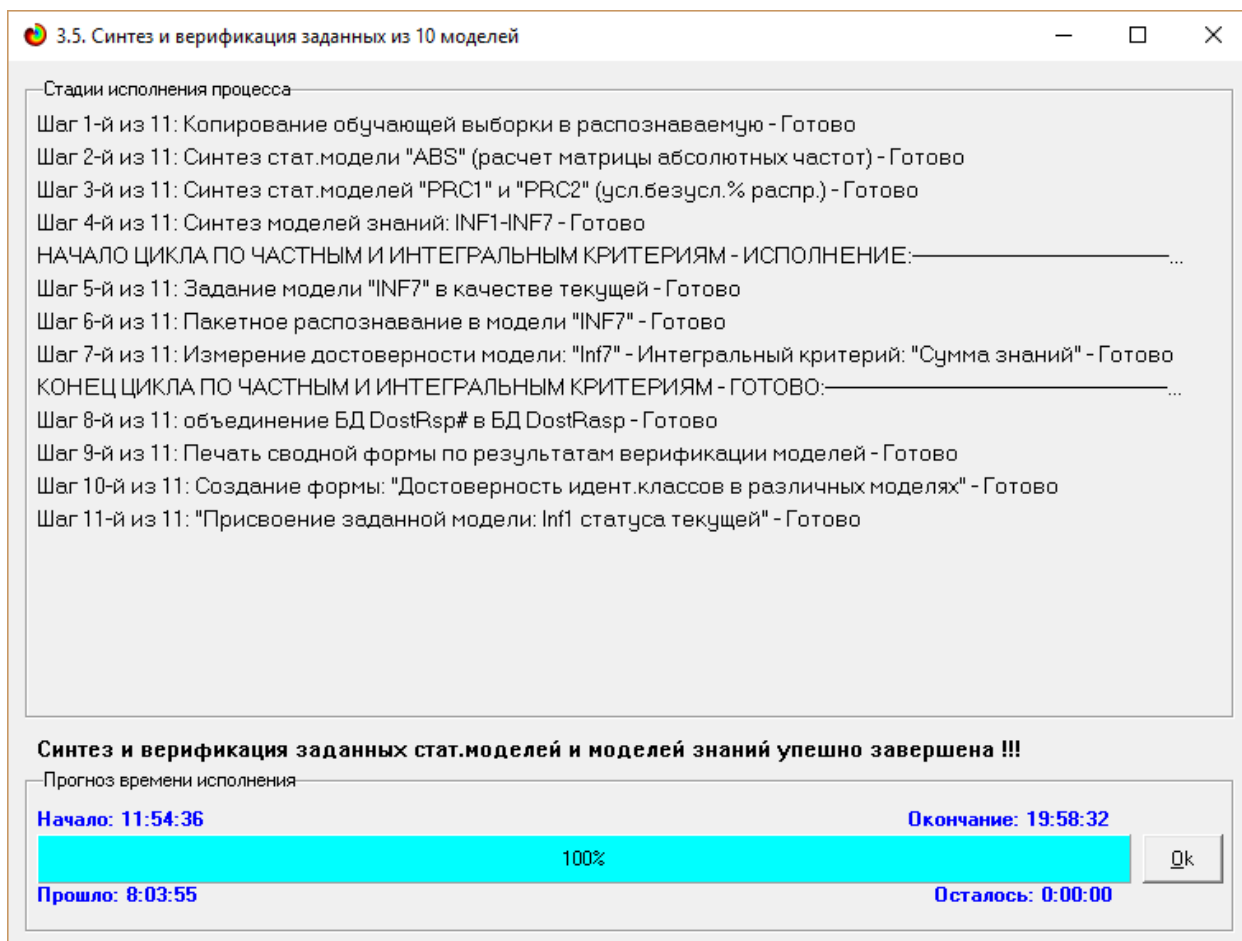


Рисунок 8 – Синтез и верификация статистических моделей и моделей знаний

Синтез и верификация всех моделей на данной задаче заняли 8 часов и 4 минуты. При этом верификация (оценка достоверности моделей) проводилась на всех 1000 примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 8, но ниже приведем лишь некоторые из них (таблицы 2, 3, 4).

1.5. 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 (таблица 2) в матрицы условных и безусловных процентных распределений, и матрицы знаний (таблицы 3 и 4) (проф. В.И.Лойко).

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

Код признака	Наименование описательной шкалы и градации	1. ORIGIN ABQ	2. ORIGIN ALB	3. ORIGIN AMA	4. ORIGIN AUS	5. ORIGIN BDL	6. ORIGIN BHM	7. ORIGIN BNA	8. ORIGIN BOI	9. ORIGIN BUF	10. ORIGIN BUR	11. ORIGIN BWI	12. ORIGIN CLE
1	DAYOFMONTH-1/5 (3.000000, 8.600000)	144	25	14	56	33	24	139	49	33	160	307	18
2	DAYOFMONTH-2/5 (8.600000, 14.200000)	64	11	3	39	23	23	110	21	23	54	267	16
3	DAYOFMONTH-3/5 (14.200000, 19.800000)	61	11	7	65	20	31	134	20	31	67	291	20
4	DAYOFMONTH-4/5 (19.800000, 25.400000)	123	10	12	84	16	50	120	30	30	140	275	23
5	DAYOFMONTH-5/5 (25.400000, 31.000000)	121	16	23	93	20	36	130	38	38	137	246	28
6	DAYOFWEEK-1/5 (1.000000, 2.200000)	128	13	12	82	29	33	153	48	40	143	324	32
7	DAYOFWEEK-2/5 (2.200000, 3.400000)	44	8	6	42	10	17	60	16	25	53	150	9
8	DAYOFWEEK-3/5 (3.400000, 4.600000)	91	12	12	72	17	45	128	21	28	82	277	20
9	DAYOFWEEK-4/5 (4.600000, 5.800000)	102	10	10	65	25	22	121	28	23	131	251	15
10	DAYOFWEEK-5/5 (5.800000, 7.000000)	148	30	19	76	31	47	171	45	39	149	384	29
11	TAILNUM-N12157	1											
12	TAILNUM-N200wN		1									2	
13	TAILNUM-N201LV							2	1		2	2	1
14	TAILNUM-N202wN							3			1	3	2
15	TAILNUM-N203wN							1			1	6	
16	TAILNUM-N204wN							1		1	3	5	
17	TAILNUM-N205wN						1	2				2	
18	TAILNUM-N206wN	1			1						4	5	
19	TAILNUM-N207wN		1								1	3	
20	TAILNUM-N208wN	1					2	3				3	1
21	TAILNUM-N209wN	1					2	2		2	2	3	
22	TAILNUM-N210wN	1		1			1	1			1	1	
23	TAILNUM-N211wN			1		1						3	
24	TAILNUM-N212wN						1	1		2		4	

Таблица 3 – Матрица информативностей (модель INF1) в битах (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. ORIGIN ABQ	2. ORIGIN ALB	3. ORIGIN AMA	4. ORIGIN AUS	5. ORIGIN BDL	6. ORIGIN BHM	7. ORIGIN BNA	8. ORIGIN BOI	9. ORIGIN BUF	10. ORIGIN BUR	11. ORIGIN BWI	12. ORIGIN CLE	13. ORIGIN CMH	14. ORIGIN CR
1	DAYOFMONTH-1/5 (3.000000, 8.600000)	0.000	0.110	-0.093	-0.291	0.027	-0.361	-0.136	0.055	-0.153	0.012	-0.131	-0.274	-0.299	
2	DAYOFMONTH-2/5 (8.600000, 14.200000)	-0.036	0.069	-0.534	-0.078	0.240	0.029	0.148	-0.001	0.060	-0.177	0.205	0.075	0.250	
3	DAYOFMONTH-3/5 (14.200000, 19.800000)	-0.159	-0.028	-0.160	0.109	0.066	0.098	0.161	-0.125	0.129	-0.154	0.156	0.102	0.073	
4	DAYOFMONTH-4/5 (19.800000, 25.400000)	0.063	-0.248	-0.029	0.084	-0.225	0.196	-0.068	-0.067	-0.056	0.088	-0.042	0.013	0.119	
5	DAYOFMONTH-5/5 (25.400000, 31.000000)	0.055	0.014	0.334	0.142	-0.099	0.015	-0.022	0.066	0.076	0.077	-0.103	0.123	-0.093	
6	DAYOFWEEK-1/5 (1.000000, 2.200000)	0.023	-0.164	-0.090	0.009	0.044	-0.096	0.006	0.132	0.042	0.038	-0.013	0.134	0.202	
7	DAYOFWEEK-2/5 (2.200000, 3.400000)	-0.025	0.111	0.070	0.183	-0.002	0.081	0.031	0.068	0.326	0.032	0.104	-0.025	-0.031	
8	DAYOFWEEK-3/5 (3.400000, 4.600000)	-0.130	-0.172	-0.054	-0.027	-0.217	0.112	-0.057	-0.290	-0.120	-0.235	-0.064	-0.091	-0.039	
9	DAYOFWEEK-4/5 (4.600000, 5.800000)	0.050	-0.157	-0.039	0.033	0.114	-0.168	0.028	-0.014	-0.112	0.142	-0.002	-0.134	-0.039	
10	DAYOFWEEK-5/5 (5.800000, 7.000000)	0.047	0.243	0.108	-0.090	0.024	0.043	0.010	0.040	-0.029	0.004	0.025	0.023	-0.179	
11	TAILNUM-N12157	2.270													
12	TAILNUM-N200wN		1.278									0.029		0.786	
13	TAILNUM-N201LV							0.248	0.633		0.317	-0.187	0.860	0.955	
14	TAILNUM-N202wN							0.560			0.020	0.125	1.332		
15	TAILNUM-N203wN							-0.110			-0.040	0.450		0.598	
16	TAILNUM-N204wN							-0.128		0.653	0.552	0.330			
17	TAILNUM-N205wN						0.699	0.335				-0.100		0.657	
18	TAILNUM-N206wN	-0.064			0.170						0.659	0.278			
19	TAILNUM-N207wN		1.128							-0.001	0.104				
20	TAILNUM-N208wN	-0.002					1.015	0.491				0.056	0.878		
21	TAILNUM-N209wN	-0.029					0.988	0.239		1.020	0.309	0.029		0.561	
22	TAILNUM-N210wN	0.329		1.530			0.962	0.213			0.283	-0.222		0.920	
23	TAILNUM-N211wN			1.370		1.014						0.228			
24	TAILNUM-N212wN						0.586	-0.163		1.002		0.171			

Таблица 4 – Матрица знаний (модель INF3) (фрагмент)

Код признака	Наименование описательной шкалы и градации	1. ORIGIN ABQ	2. ORIGIN ALB	3. ORIGIN AMA	4. ORIGIN AUS	5. ORIGIN BDL	6. ORIGIN BHM	7. ORIGIN BNA	8. ORIGIN BOI	9. ORIGIN BUF	10. ORIGIN BUR	11. ORIGIN BWI	12. ORIGIN CLE	13. ORIGIN CMH	14. ORIGIN CR
1	DAYOFMONTH-1/5-(3.000000, 8.600000)	0.013	4.511	-2.560	-38.588	1.564	-22.031	-38.668	4.653	-10.505	3.383	-82.017	-11.471	-20.680	
2	DAYOFMONTH-2/5-(6.600000, 14.200000)	-4.303	1.280	-4.856	-5.870	8.088	1.164	25.720	-0.037	2.363	-20.294	82.462	2.020	13.433	
3	DAYOFMONTH-3/5-(14.200000, 19.800000)	-20.274	-0.565	-2.347	11.610	2.256	5.018	33.715	-5.032	6.444	-21.403	71.418	3.365	3.958	
4	DAYOFMONTH-4/5-(19.800000, 25.400000)	13.140	-5.633	-0.635	11.830	-7.985	14.879	-15.559	-3.836	-3.194	20.503	-21.816	0.514	9.095	
5	DAYOFMONTH-5/5-(25.400000, 31.000000)	11.424	0.407	10.398	21.017	-3.923	0.970	-5.208	4.251	4.892	17.812	-50.047	5.572	-5.807	
6	DAYOFWEEK-1/5-(1.000000, 2.200000)	5.235	-4.469	-2.119	1.354	2.198	-6.246	1.519	10.189	2.907	9.467	-7.680	6.873	18.643	
7	DAYOFWEEK-2/5-(2.200000, 3.400000)	-1.993	1.455	0.710	11.786	-0.041	2.297	3.249	1.835	11.104	2.973	25.739	-0.414	-0.869	
8	DAYOFWEEK-3/5-(3.400000, 4.600000)	-24.032	-4.369	-1.230	-3.567	-8.114	8.226	-13.940	-14.429	-6.756	-43.123	-33.788	-3.545	-2.689	
9	DAYOFWEEK-4/5-(4.600000, 5.800000)	8.776	-3.266	-0.722	3.759	4.647	-7.803	5.969	-0.712	-5.167	29.598	-0.868	-4.081	-2.165	
10	DAYOFWEEK-5/5-(5.800000, 7.000000)	12.013	10.649	3.360	-13.332	1.311	3.527	3.204	3.117	-2.088	1.085	16.598	1.166	-12.919	
11	TAILNUM-N12157	0.983	-0.002	-0.002	-0.011	-0.004	-0.005	-0.021	-0.005	-0.018	-0.045	-0.003	-0.006	-0.006	
12	TAILNUM-N200VN	-0.703	0.900	-0.081	-0.462	-0.153	-0.225	-0.867	-0.217	-0.212	-0.765	0.101	-0.144	0.757	
13	TAILNUM-N201LV	-1.038	-0.148	-0.119	-0.682	-0.227	-0.332	0.720	0.680	-0.314	0.871	-0.804	0.788	1.642	
14	TAILNUM-N202VN	-0.887	-0.126	-0.102	-0.583	-0.194	-0.284	1.905	-0.273	-0.268	0.035	0.603	1.818	-0.306	
15	TAILNUM-N203VN	-0.987	-0.141	-0.114	-0.649	-0.216	-0.316	-0.218	-0.304	-0.298	-0.074	3.332	-0.202	0.659	
16	TAILNUM-N204VN	-1.021	-0.145	-0.117	-0.671	-0.223	-0.326	-0.260	-0.314	0.692	1.889	2.242	-0.209	-0.352	
17	TAILNUM-N205VN	-0.887	-0.126	-0.102	-0.583	-0.194	0.716	0.905	-0.273	-0.268	-0.965	-0.397	-0.182	0.694	
18	TAILNUM-N206VN	-0.121	-0.160	-0.129	0.263	-0.245	-0.358	-1.384	-0.345	-0.339	2.780	1.970	-0.230	-0.387	
19	TAILNUM-N207VN	-0.921	0.869	-0.106	-0.605	-0.201	-0.294	-1.136	-0.284	-0.278	-0.001	0.513	-0.188	-0.318	
20	TAILNUM-N208VN	-0.004	-0.143	-0.115	-0.660	-0.219	1.679	1.761	-0.309	-0.303	-1.092	0.287	0.794	-0.346	
21	TAILNUM-N209VN	-0.054	-0.150	-0.121	-0.693	-0.230	1.663	0.699	-0.325	1.681	0.853	0.151	-0.216	0.636	
22	TAILNUM-N210VN	0.448	-0.079	0.936	-0.363	-0.121	0.823	0.318	-0.170	-0.167	0.399	-0.492	-0.113	0.809	
23	TAILNUM-N211VN	-0.736	-0.105	0.915	-0.484	0.839	-0.235	-0.909	-0.227	-0.223	-0.801	1.010	-0.151	-0.254	
24	TAILNUM-N212VN	-1.088	-0.155	-0.125	-0.715	-0.238	0.652	-0.342	-0.335	1.671	-1.183	1.061	-0.223	-0.375	

1.6. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей (режим 4.1.3.6), отличающихся частными критериями с двумя приведенными выше интегральными критериями приведены на рисунке 9.

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложно-положительных решений (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Полнота модели	Фигера Ван Розенберга	Синга моду-уровней спо- истинно-поло- решений (ST)	Синга моду-уровней спо- ложно-отри- решений (ST)	Синга моду-уровней спо- ложно-поло- решений (SFP)	Синга моду-уровней спо- ложно-отри- решений (SFP)
1. ABS - частный критерий: количество встреч сонетаний "1/лас...	Коррекция абс. частот с обр...	2000	2000	475	126525		0.016	1.000	0.031	1084.336	3.494	60514.382	
1. ABS - частный критерий: количество встреч сонетаний "1/лас...	Синга абс. частот по призна...	2000	2000	475	126525		0.016	1.000	0.031	570.397		15243.227	
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	2000	2000	475	126525		0.016	1.000	0.031	1084.336	3.494	60514.418	
2. PR1 - частный критерий: усл. вероятность i-го признака сред...	Синга усл.отн.частот по при...	2000	2000	475	126525		0.016	1.000	0.031	487.925		27840.742	
3. PR2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	2000	2000	475	126525		0.016	1.000	0.031	1084.315	3.494	60513.246	
3. PR2 - частный критерий: условная вероятность i-го признака...	Синга усл.отн.частот по при...	2000	2000	475	126525		0.016	1.000	0.031	487.925		27840.742	
4. INF1 - частный критерий: количество знаний по АХаркевичу, в...	Семантический резонанс зна...	2000	1561	94931	32069	439	0.046	0.781	0.088	162.756	8858.091	1988.754	25.011
4. INF1 - частный критерий: количество знаний по АХаркевичу, в...	Синга знаний	2000	1707	54019	72981	293	0.023	0.854	0.045	318.301	4335.939	7349.404	20.771
5. INF2 - частный критерий: количество знаний по АХаркевичу, в...	Семантический резонанс зна...	2000	1561	94931	32069	439	0.046	0.781	0.088	162.756	8858.092	1988.754	25.011
5. INF2 - частный критерий: количество знаний по АХаркевичу, в...	Синга знаний	2000	1707	54019	72981	293	0.023	0.854	0.045	318.301	4335.938	7349.403	20.771
6. INF3 - частный критерий: Хинкватер, разности между факти...	Семантический резонанс зна...	2000	1446	65348	61652	554	0.023	0.723	0.044	546.669	17146.514	20112.249	149.021
6. INF3 - частный критерий: Хинкватер, разности между факти...	Синга знаний	2000	1446	65348	61652	554	0.023	0.723	0.044	317.930	6868.332	6657.804	107.401
7. INF4 - частный критерий: ROI (Return On Investment), верою...	Семантический резонанс зна...	2000	1265	105059	21941	735	0.055	0.633	0.100	165.371	7752.511	1595.000	36.901
7. INF4 - частный критерий: ROI (Return On Investment), верою...	Синга знаний	2000	1805	45433	81567	195	0.022	0.903	0.042	31.487	109.377	717.882	0.461
8. INF5 - частный критерий: ROI (Return On Investment), верою...	Семантический резонанс зна...	2000	1265	105059	21941	735	0.055	0.633	0.100	165.372	7752.518	1595.001	36.901
8. INF5 - частный критерий: ROI (Return On Investment), верою...	Синга знаний	2000	1805	45433	81567	195	0.022	0.903	0.042	31.486	109.377	717.882	0.461
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Семантический резонанс зна...	2000	1438	60504	66496	562	0.021	0.719	0.041	525.953	14629.621	20277.910	143.351
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Синга знаний	2000	1446	58595	68415	554	0.021	0.723	0.040	85.955	2442.967	3709.700	19.691
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	2000	1438	60504	66496	562	0.021	0.719	0.041	525.953	14629.621	20277.910	143.351
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Синга знаний	2000	1446	58595	68415	554	0.021	0.723	0.040	85.955	2442.967	3709.700	19.691

Рисунок 9 Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF4 и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели INF4 составляет 0,055. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется нечеткое мультиклассовое обобщение F-меры, предложенное проф.Е.В.Луценко (L-

мера), при этом точность модели составляет 0,169 и 0,680 по L1 и L2-мерам. Таким образом, уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной примерно 70%.

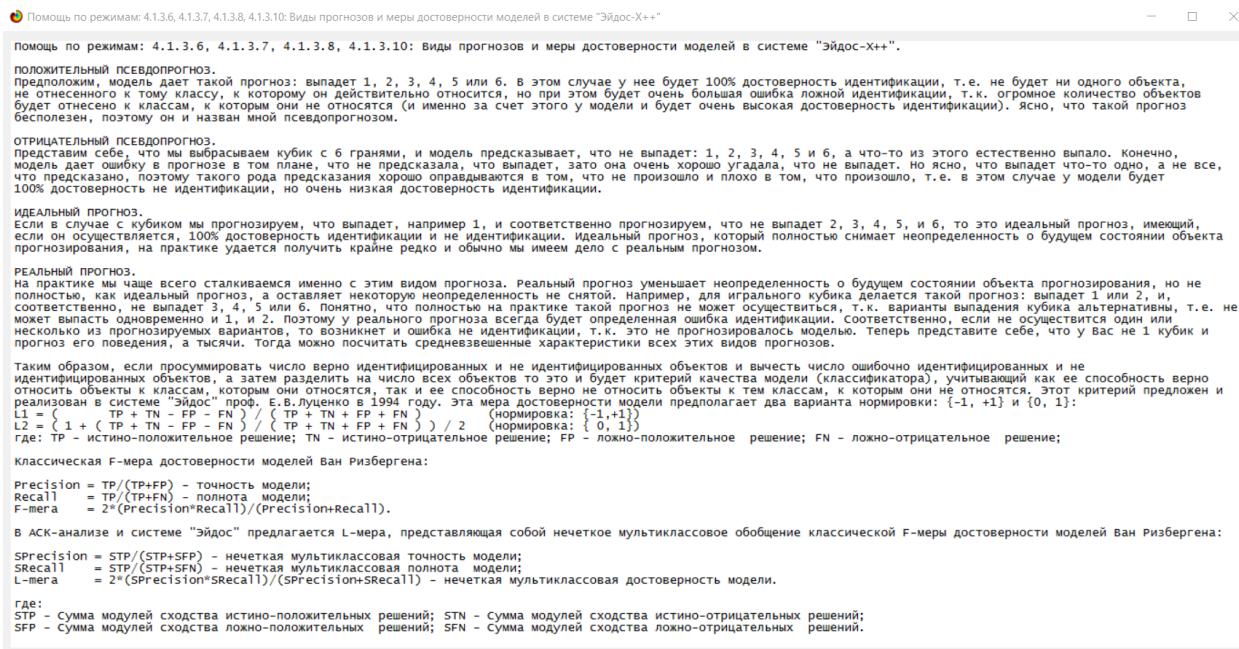


Рисунок 10. Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

Также 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. На рисунке 11 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF4.

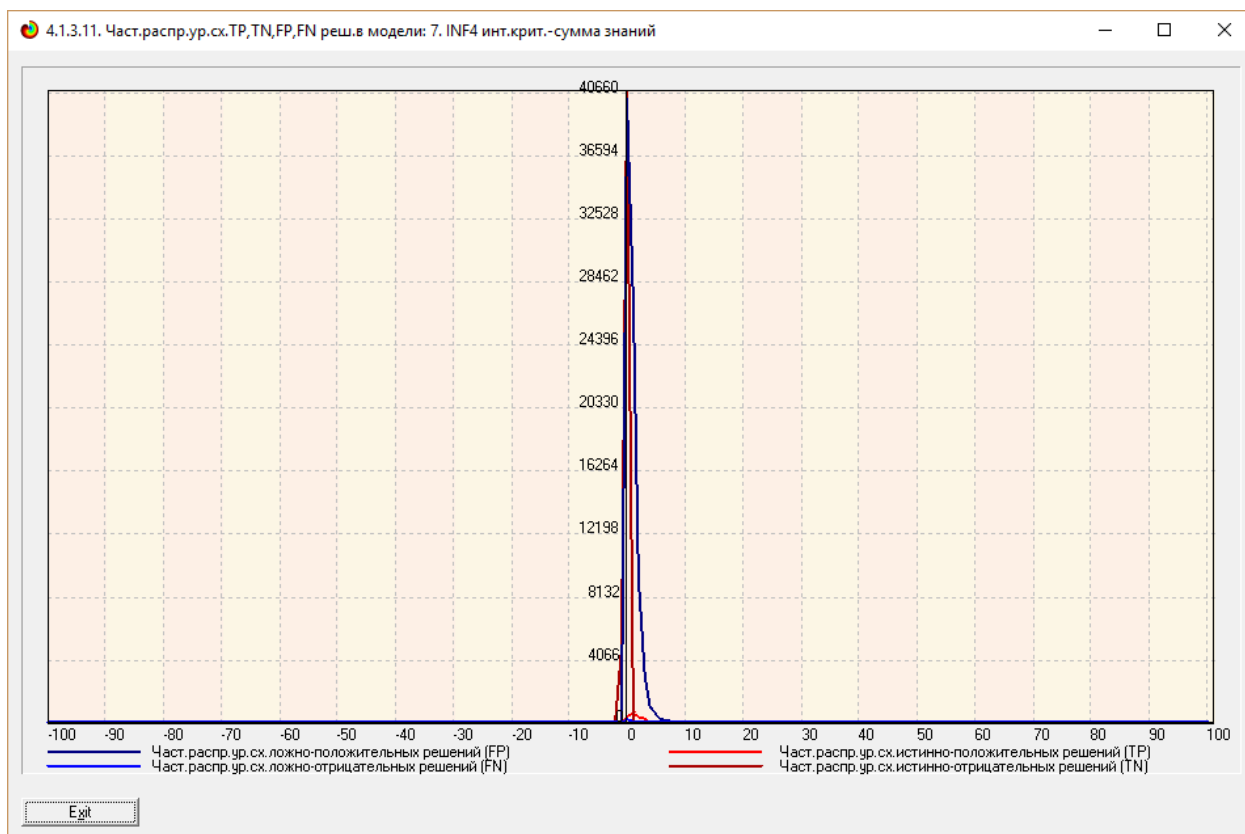


Рисунок 11. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF4

Из рисунка 11 видно, что:

- наиболее достоверная модель INF4 лучше определяет непринадлежность объекта к классу, чем принадлежность (что видно также из рисунка 9);

- модуль уровня сходства-различия в наиболее достоверной модели INF4 для верно идентифицированных и верно неидентифицированных объектов выше, чем для ошибочно идентифицированных и ошибочно неидентифицированных. Это верно практически для всего диапазона уровней сходства-различия, кроме небольших по модулю значений в диапазоне от 0 до 15% уровня сходства. Для больших значений уровней сходства-различия (более 50%) различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями настолько велико, что учитывая уровень сходства-различия не представляет никакой проблемы разделить

истинные положительные и отрицательные решения от ложных положительных и отрицательных решений. Это и стало основой для нечеткого обобщения F-меры Ван Ризбергена, т.е. такого ее обобщения, которое учитывает не только сам факт ложно или истинно положительного или отрицательного решения, но и его надежность.

Возьмем любые данные о наблюдениях как сумму истинного значения и шума, однако ни первое, ни второе неизвестны. Поэтому необходимо сравнить созданные модели с чисто случайными моделями, совпадающими по основным характеристикам. В системе «Эйдос» 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 12. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

Различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели 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.

На рисунке 14 приведены данные по достоверности статистических и когнитивных моделей, созданных на основе случайной выборки.

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

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложн. положительных решений (FP)	Число ложн. отрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Риббергена	Сумма модулей уровней схожести истинно-положительных решений (ST)	Сумма модулей уровней схожести истинно-отрицательных решений (ST)	Сумма модулей уровней схожести ложно-положительных решений (SFP)	Сумма модулей уровней схожести ложно-отрицательных решений (SF)
1. ABS - частный критерий: количество встреч сопоставлений "клас...	Корреляция абс. частот с абс. частотами по признакам...	2000	2000	475	126525		0.016	1.000	0.031	1084.336	3.494	60514.382	
1. ABS - частный критерий: количество встреч сопоставлений "клас...	Сумма абс. частот по признакам...	2000	2000	475	126525		0.016	1.000	0.031	570.397		15243.227	
2. PRIC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн. частот с о...	2000	2000	475	126525		0.016	1.000	0.031	1084.336	3.494	60514.418	
2. PRIC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн. частот по признакам...	2000	2000	475	126525		0.016	1.000	0.031	487.925		27840.742	
3. PRIC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн. частот с о...	2000	2000	475	126525		0.016	1.000	0.031	1084.315	3.494	60513.246	
3. PRIC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн. частот по признакам...	2000	2000	475	126525		0.016	1.000	0.031	487.925		27840.742	
4. INF1 - частный критерий: количество знаний по АХаркевичу: в...	Семантический резонанс зна...	2000	1561	94931	32069	439	0.046	0.781	0.088	162.756	8858.091	1988.754	25.014
4. INF1 - частный критерий: количество знаний по АХаркевичу: в...	Сумма знаний	2000	1707	54019	72981	293	0.023	0.854	0.045	318.301	4335.939	7349.404	20.771
5. INF2 - частный критерий: количество знаний по АХаркевичу: в...	Семантический резонанс зна...	2000	1561	94931	32069	439	0.046	0.781	0.088	162.756	8858.092	1988.754	25.014
5. INF2 - частный критерий: количество знаний по АХаркевичу: в...	Сумма знаний	2000	1707	54019	72981	293	0.023	0.854	0.045	318.301	4335.938	7349.403	20.771
6. INF3 - частный критерий: "Хинкватрат", разности между факти...	Семантический резонанс зна...	2000	1446	65348	61652	554	0.023	0.723	0.044	546.669	17146.514	20112.249	149.021
6. INF3 - частный критерий: "Хинкватрат", разности между факти...	Сумма знаний	2000	1446	65348	61652	554	0.023	0.723	0.044	317.930	6868.332	6657.804	107.401
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	2000	1265	105059	21941	735	0.055	0.633	0.100	165.371	7752.511	1595.000	36.901
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	2000	1805	45433	81567	195	0.022	0.903	0.042	31.487	109.377	717.882	0.461
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	2000	1265	105059	21941	735	0.055	0.633	0.100	165.372	7752.518	1595.001	36.901
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	2000	1805	45433	81567	195	0.022	0.903	0.042	31.486	109.377	717.882	0.461
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Семантический резонанс зна...	2000	1438	60504	66496	562	0.021	0.719	0.041	525.953	14629.621	20277.910	143.351
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер...	Сумма знаний	2000	1446	58585	68415	554	0.021	0.723	0.040	85.955	2442.967	3709.700	19.691
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	2000	1438	60504	66496	562	0.021	0.719	0.041	525.953	14629.621	20277.910	143.351
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Сумма знаний	2000	1446	58585	68415	554	0.021	0.723	0.040	85.955	2442.967	3709.700	19.691

Рисунок 14. Достоверность статистических и когнитивных моделей, созданных на основе случайной выборки

На основе его сравнения с рисунком 9 можно сделать следующие выводы:

– достоверность лучшей модели INF4, отражающей реальный объект моделирования, примерно на 17% 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.

2. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ

2.1. Решение задачи

В соответствии с технологией АСК-анализа зададим текущей модель INF4 (режим 5.6) (рисунок 15) и проведем пакетное распознавание в режиме 4.2.1 (рисунок 16).

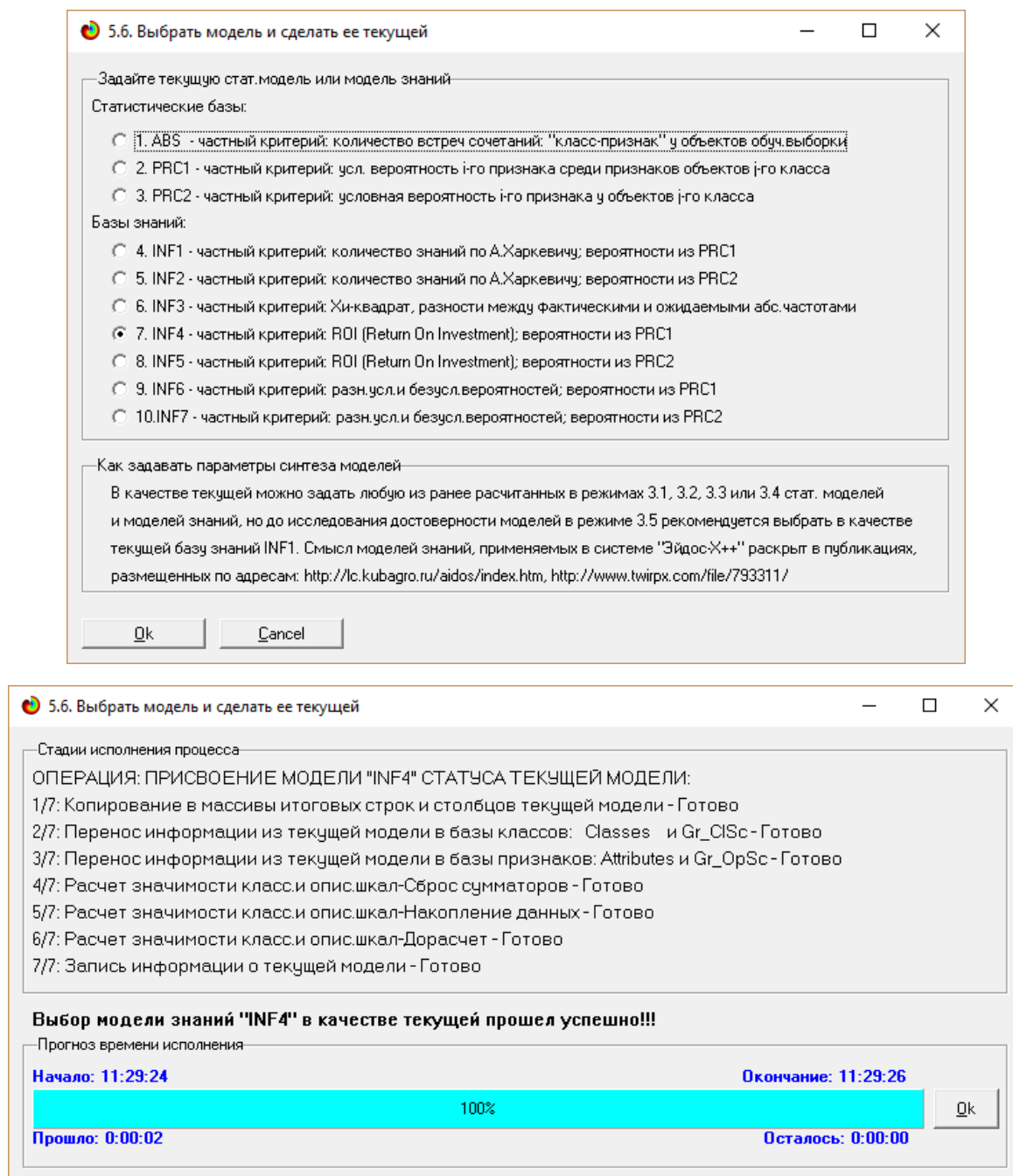


Рисунок 15. Экранные формы режима задания модели в качестве текущей

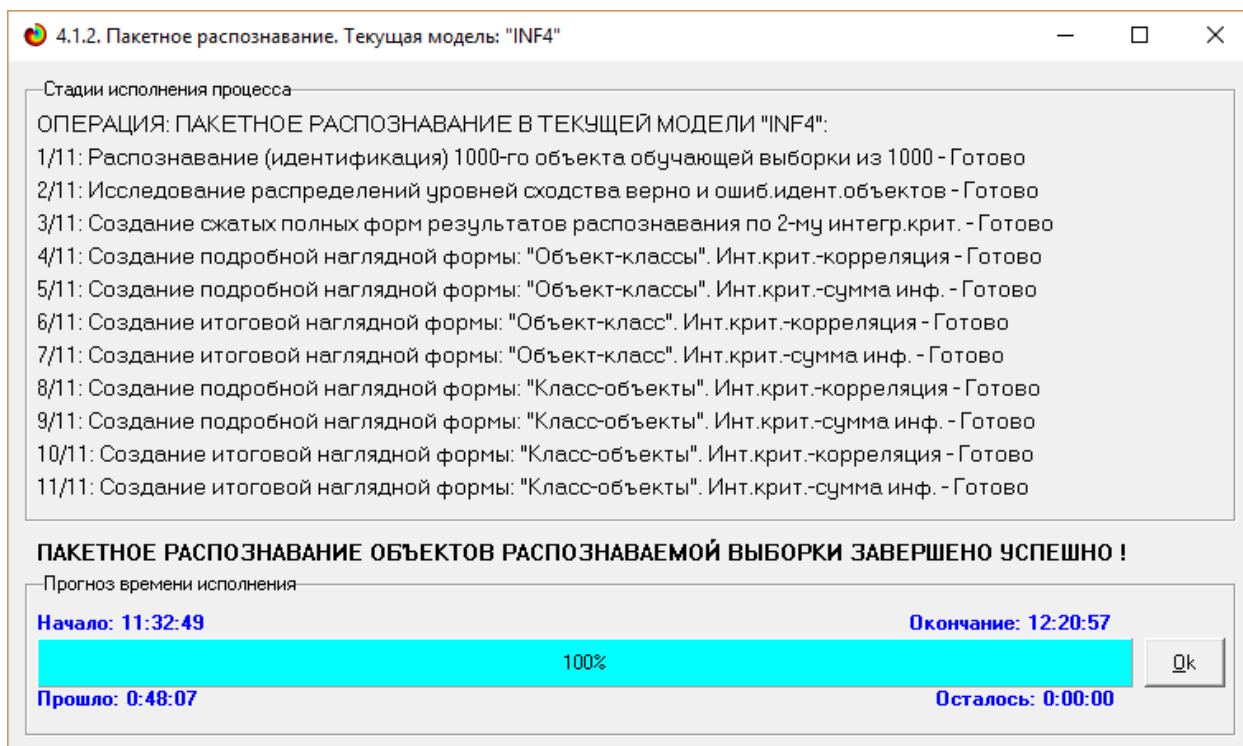


Рисунок 16. Экранная форма режима пакетного распознавания в текущей модели

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

Режим 4.1.3 системы «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.

Ниже кратко рассмотрим некоторые из них.

На рисунках 17 и 18 приведены примеры прогнозов по оплате услуг в зависимости от жилищных условий, района жилья и состава семьи в наиболее достоверной модели INF5 на основе наблюдения предыстории оплаты платежей:

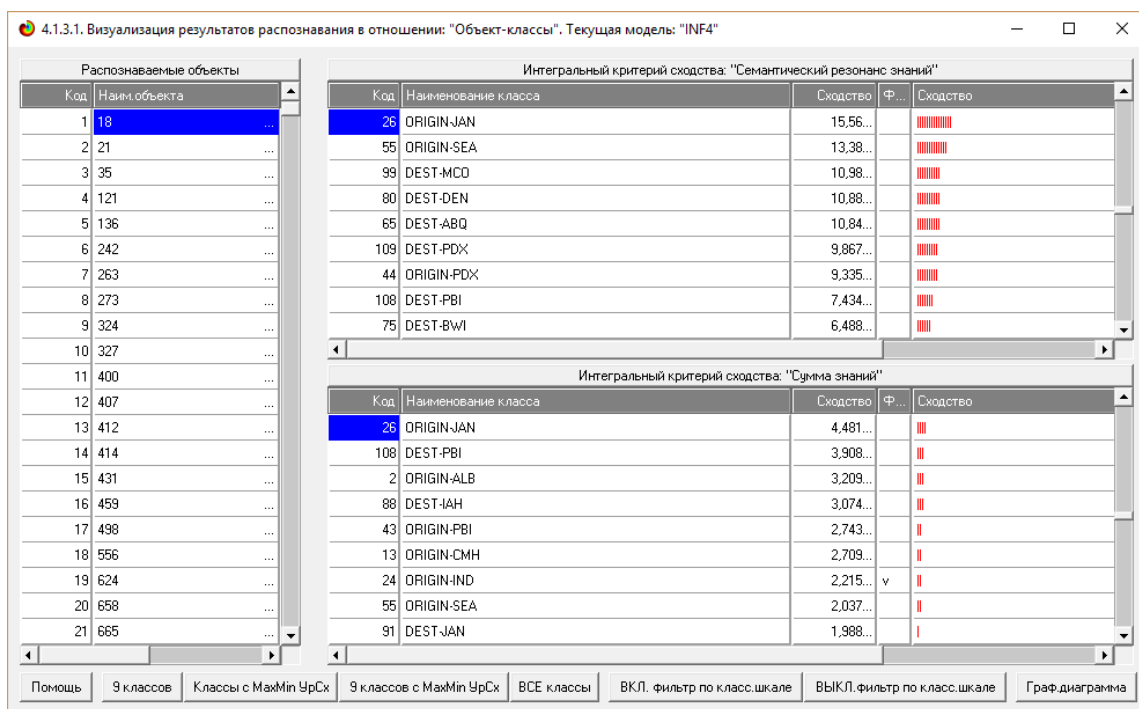


Рисунок 17. Пример идентификации вылетов и прилёты в аэропорт

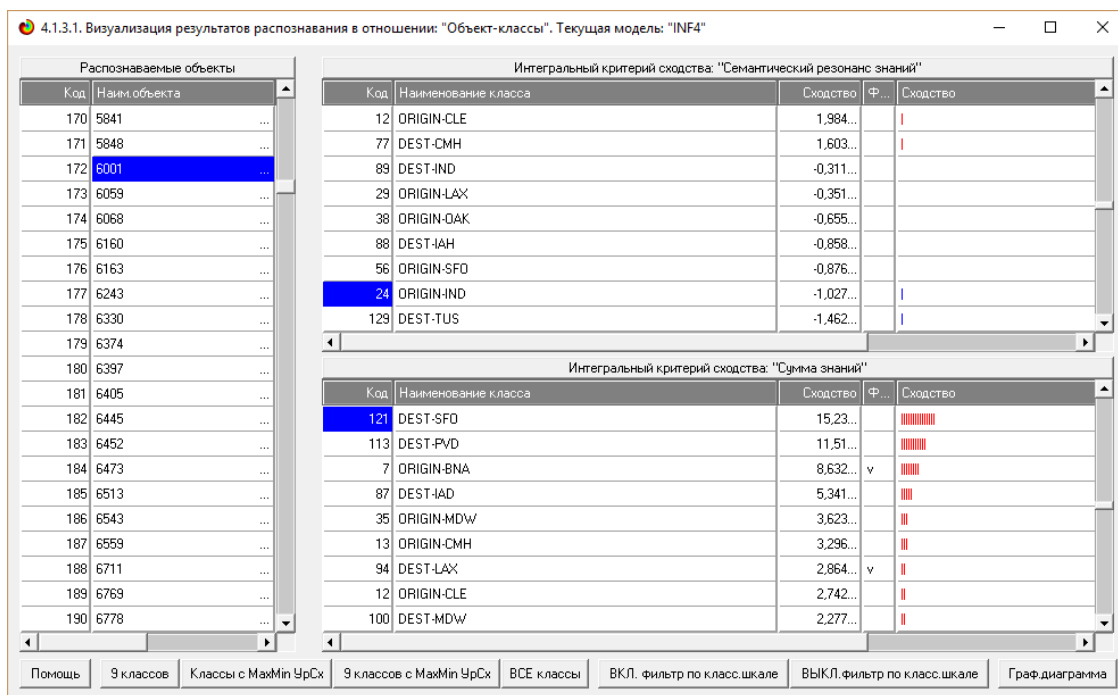


Рисунок 18. Пример вылетов и прилёты в аэропорт в модели INF4

2.2. Когнитивные функции

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

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 много работ, поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 20 приведены визуализации всех когнитивных функций данного приложения для модели INF4.

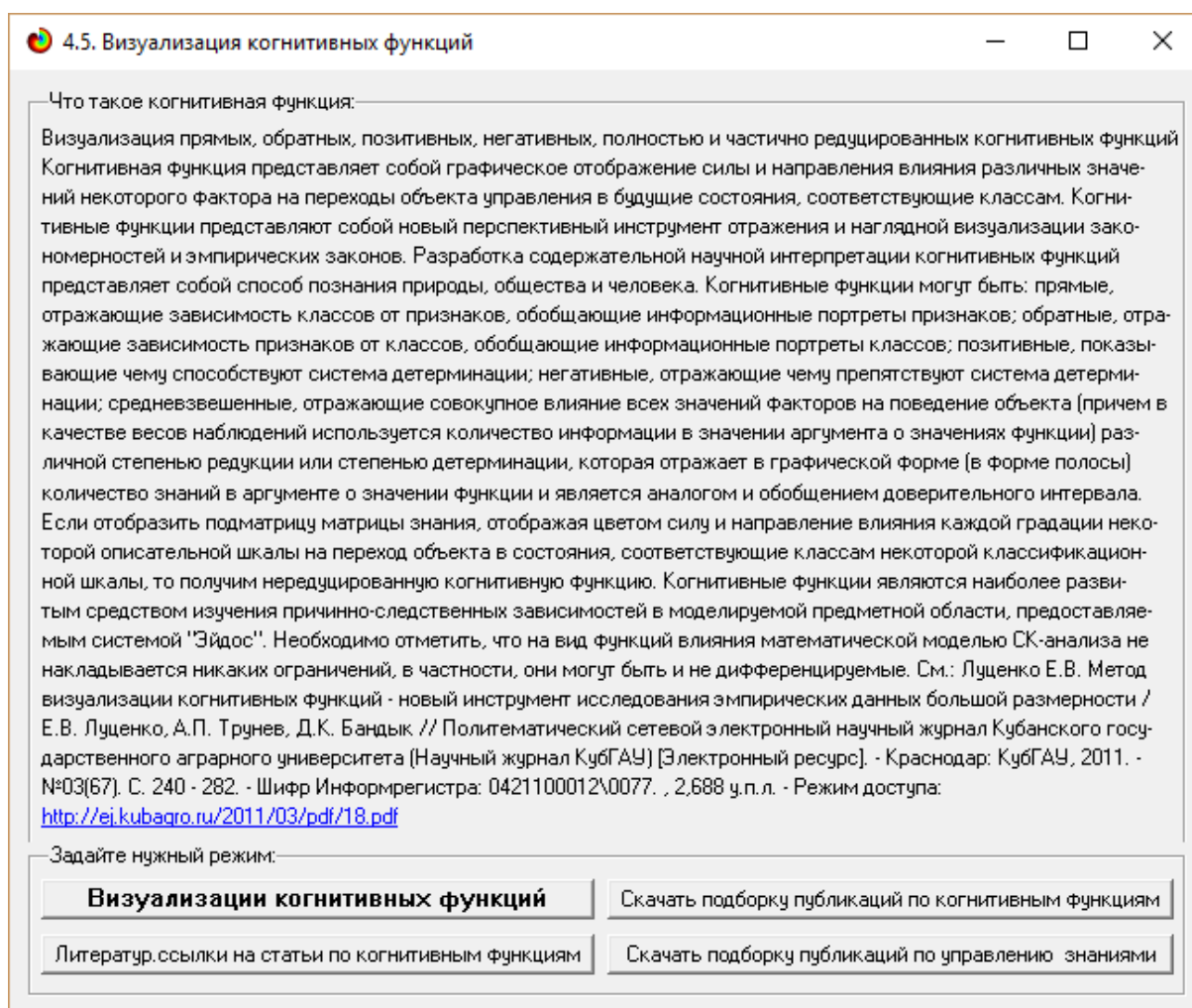
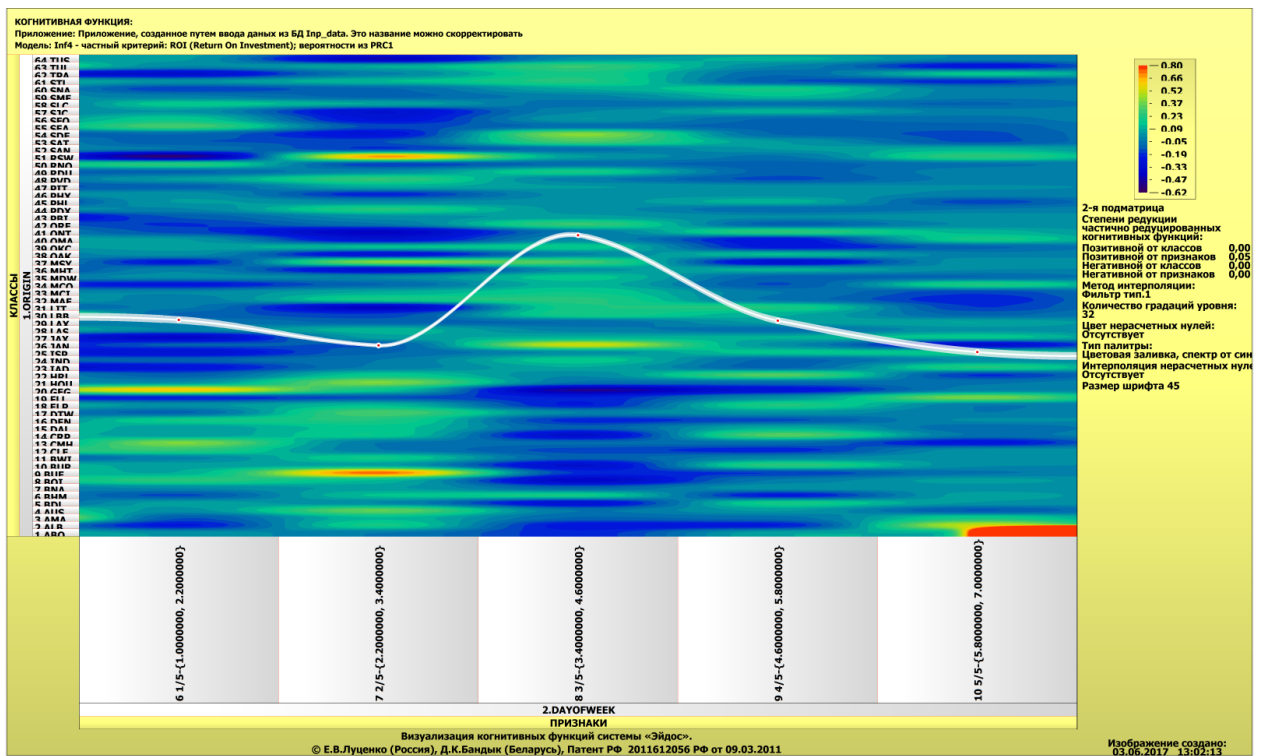
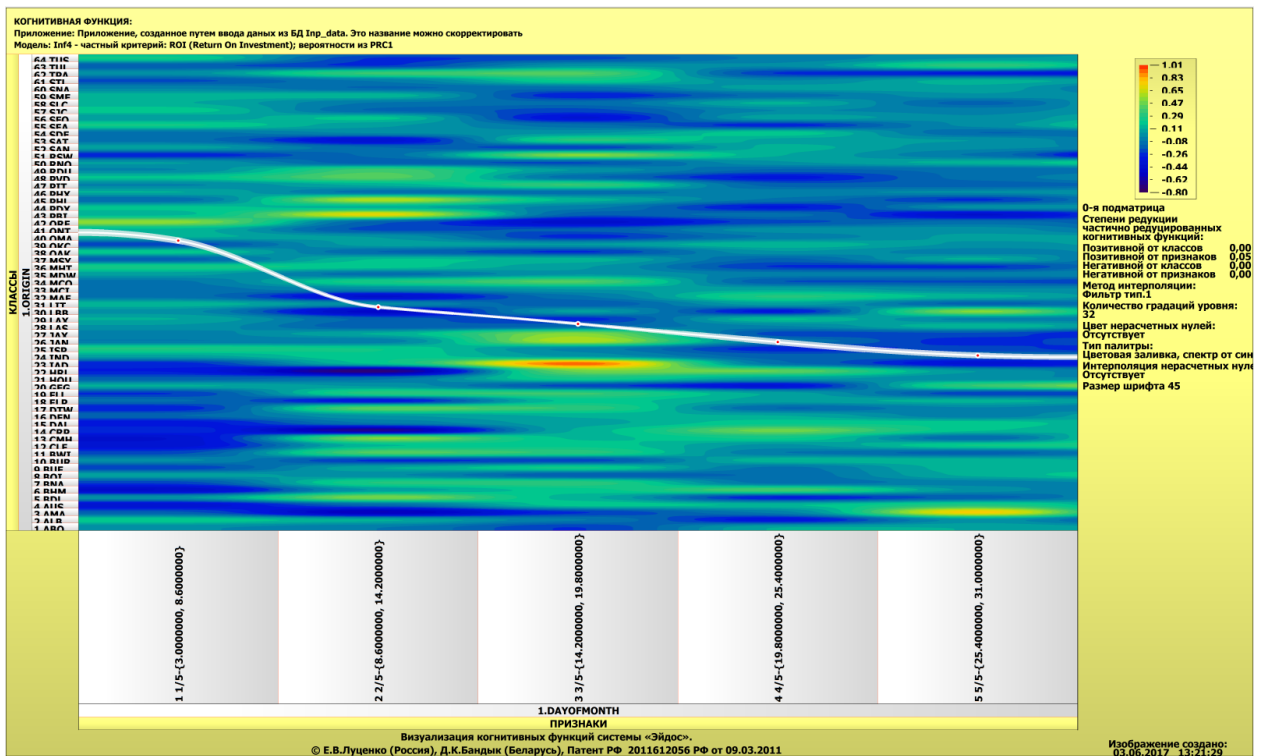
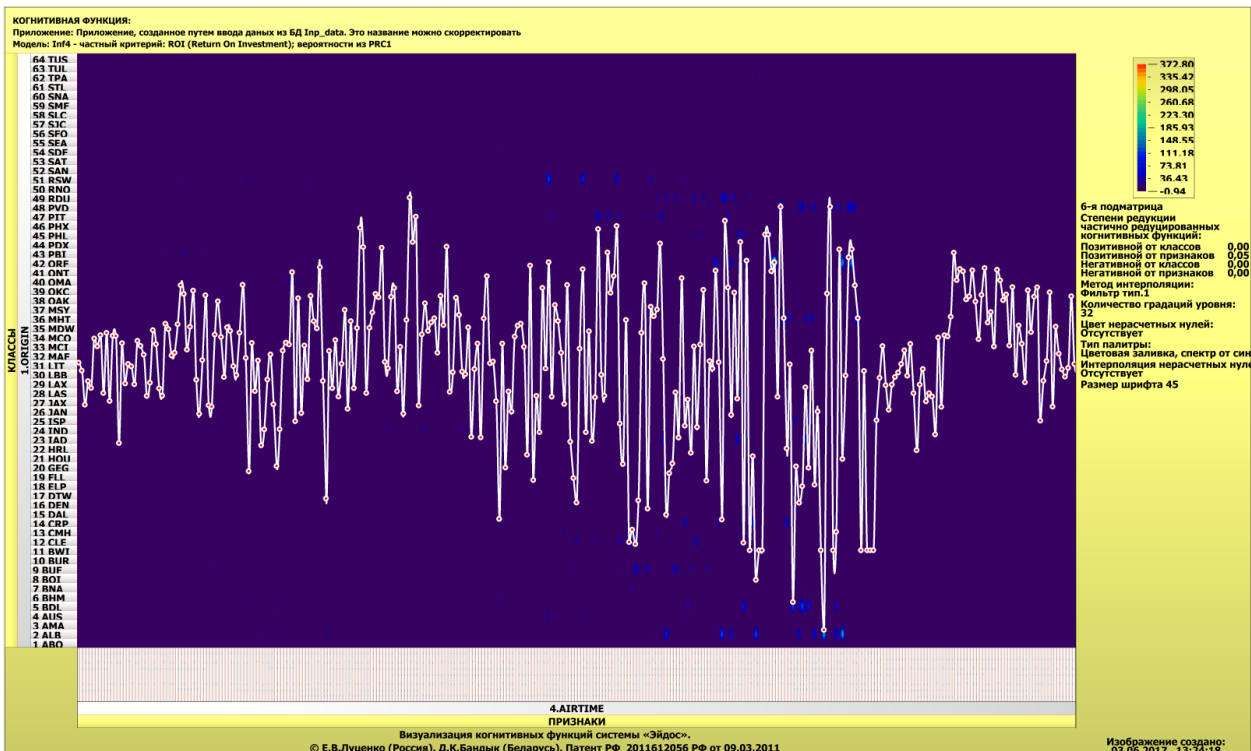
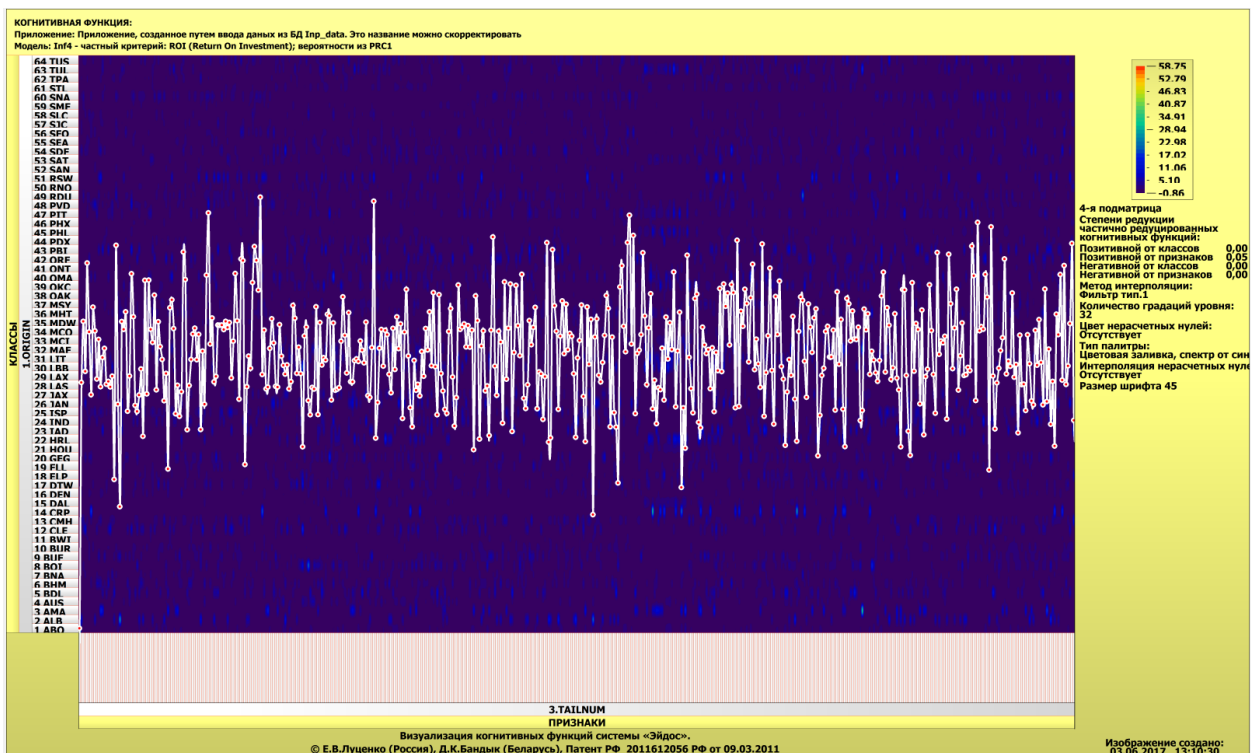


Рисунок 19. Экранная форма режима 4.5 системы «Эйдос-X++» «Визуализация когнитивных функций»





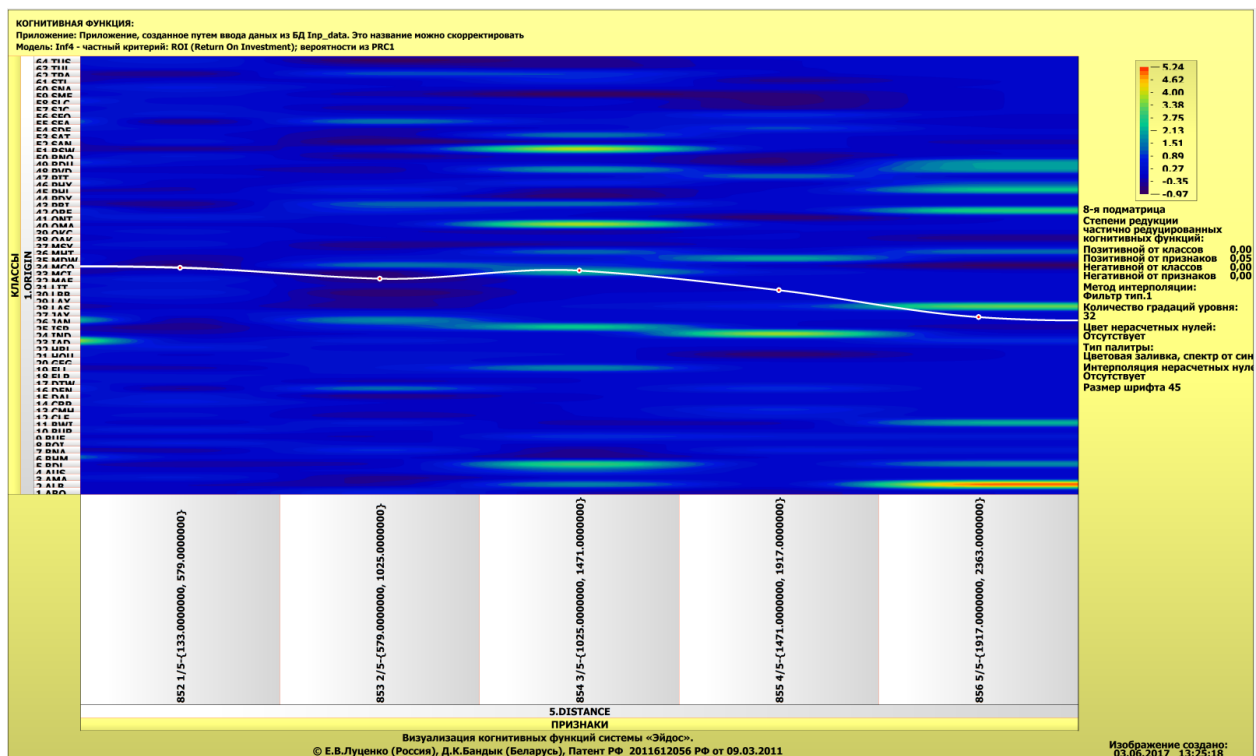


Рисунок 20. Визуализация когнитивных функций для обобщенных классов и описательных шкал для модели INF4

На основе полученных данных сделаем следующие прогнозы по полетам на самолетах:

1). Максимальная загруженность в аэропортах приходится на середину месяцев и выходные дни недели;

2). Наиболее часто летают на средние и большие дистанции.

В выборке были представлены аэропорта отправления и прилетов.

2.3. SWOT и PEST матрицы и диаграммы

SWOT-анализ является широко известным и общепризнанным методом стратегического планирования. Однако это не мешает тому, что он подвергается критике, часто вполне справедливой, обоснованной и хорошо аргументированной. В результате критического рассмотрения SWOT-анализа выявлено довольно много его слабых сторон (недостатков),

источником которых является необходимость привлечения экспертов, в частности для оценки силы и направления влияния факторов. Ясно, что эксперты это делают 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 21).

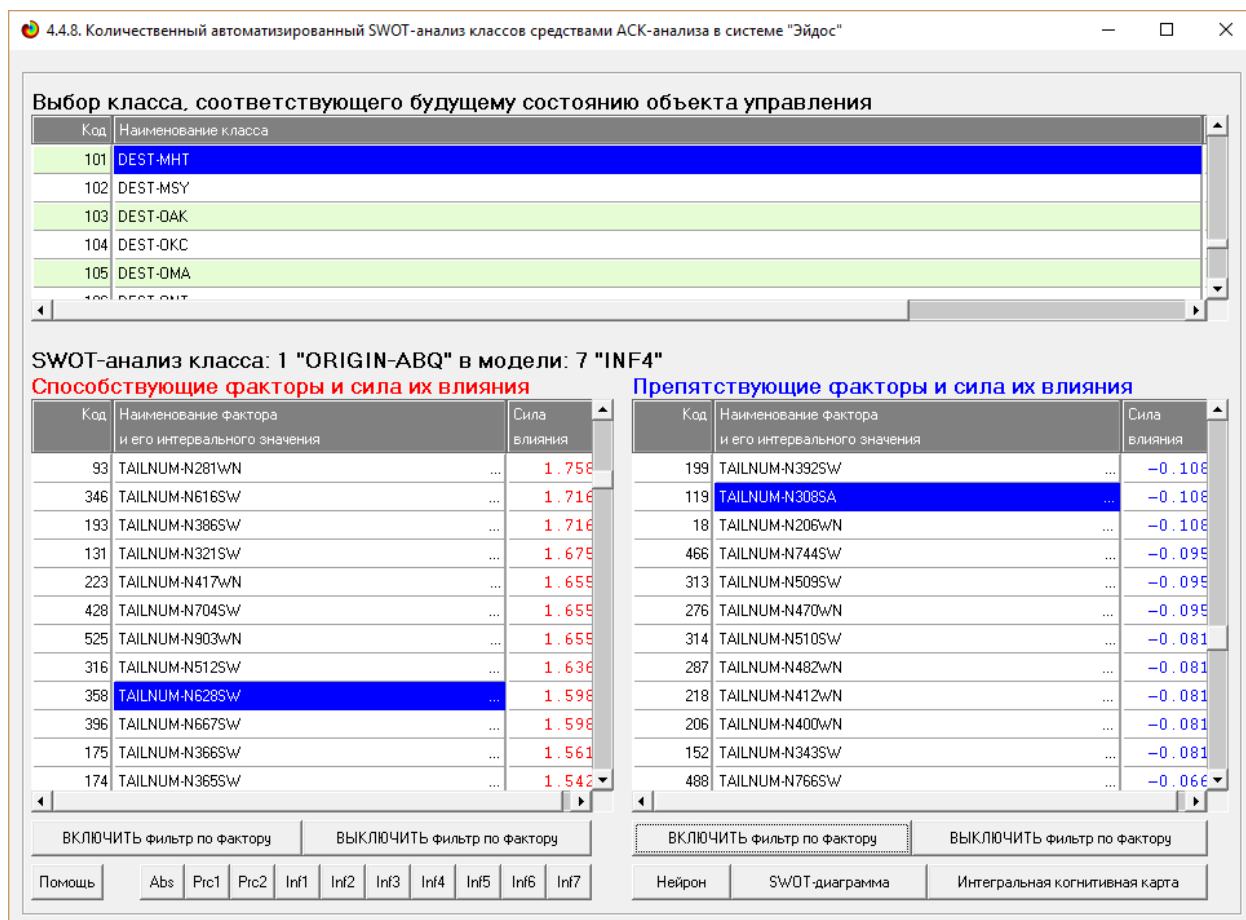
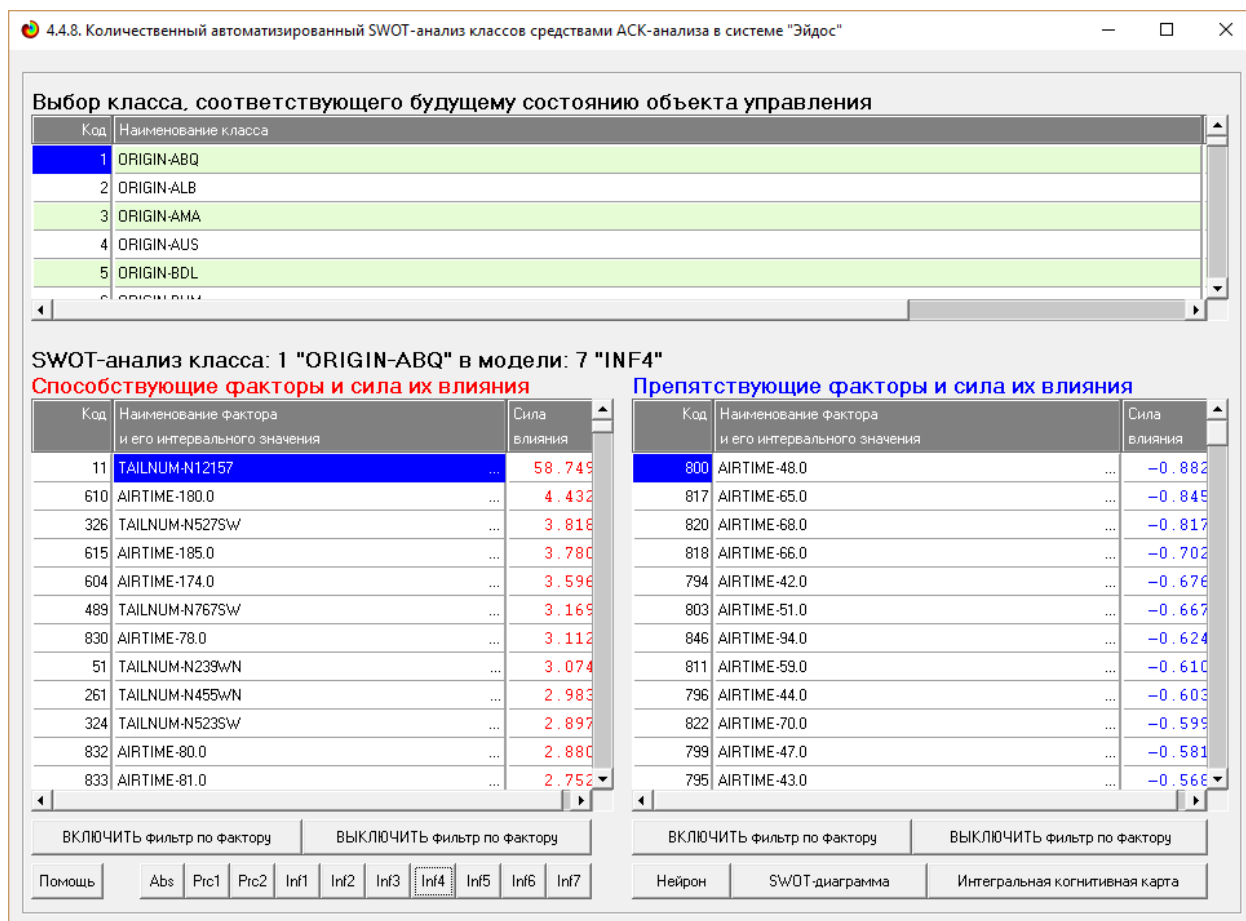


Рисунок 21. Пример SWOT-матрицы в модели INF4

На рисунке 22 приведены примеры инвертированной SWOT- матрицы и инвертированной SWOT-диаграммы в модели INF5.

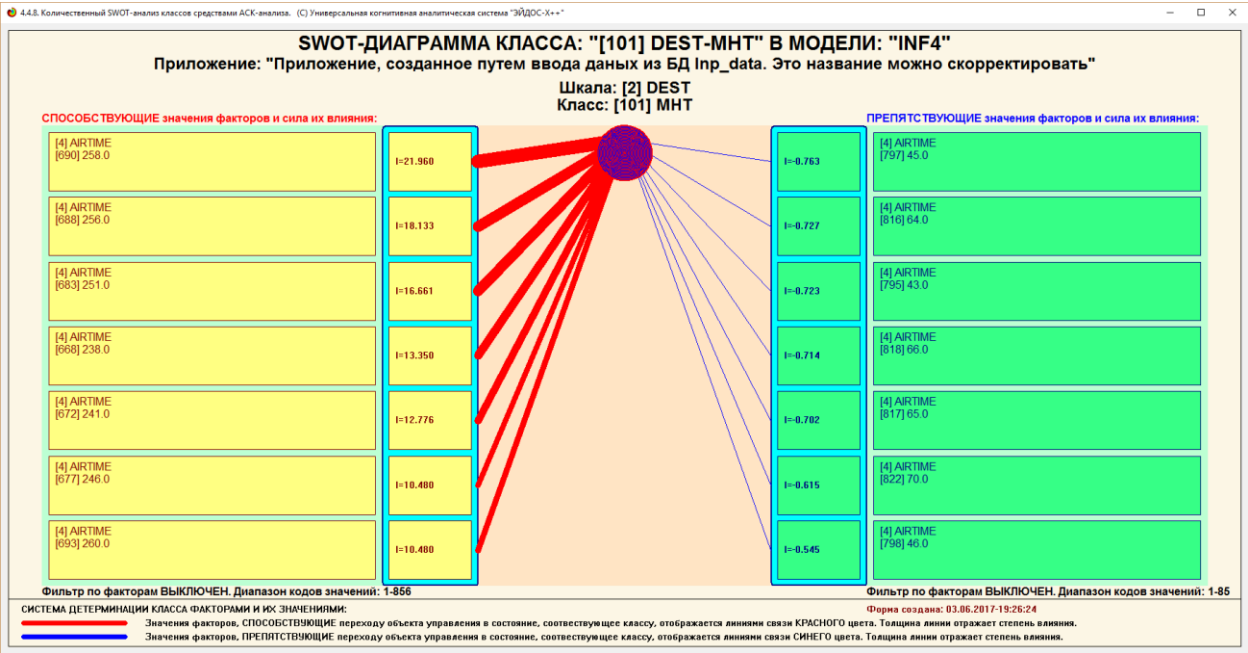


Рисунок 22. Пример SWOT-матрицы в модели INF5

2.4. Нелокальные нейроны и нейронная сеть

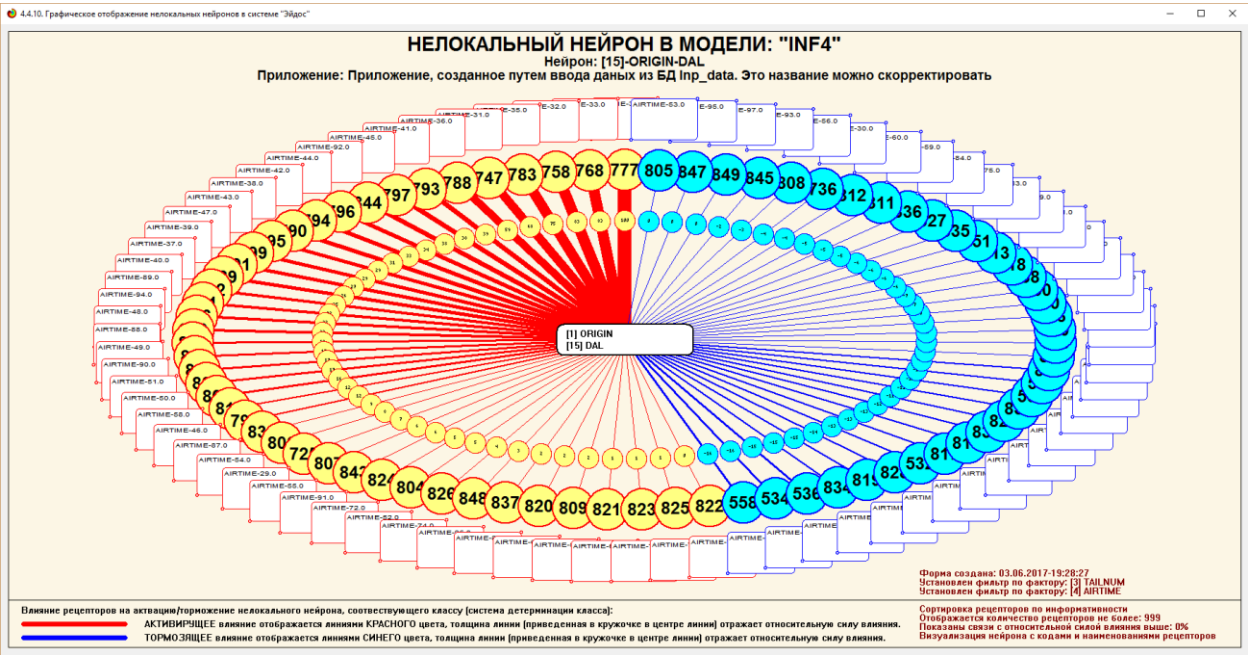


Рисунок 23. Пример графического отображения нелокальных нейронов

В качестве примера приведено графическое отображение нелокальных нейронов в модели INF4 на основе признака «Оплата-Большое». Каждому классу соответствует нейрон, совокупность которых образует нелокальную нейронную сеть [9].

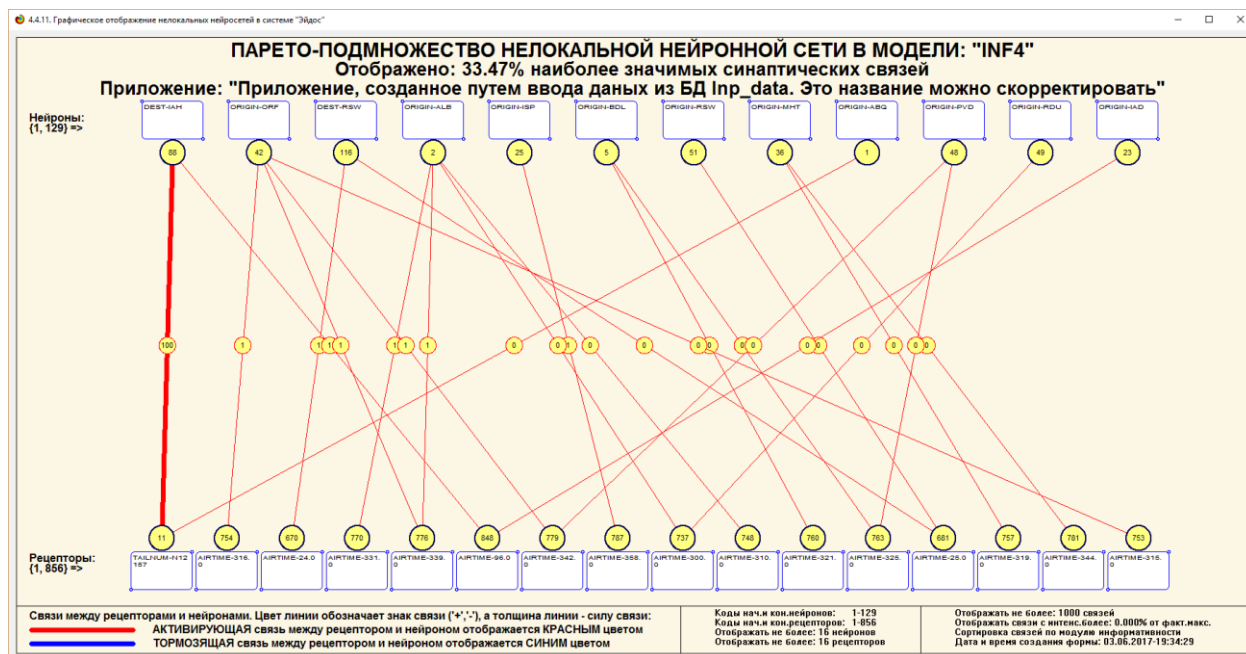


Рисунок 24. Пример графического отображения парето-подмножества нелокальной нейронной сети

На рисунке 24 в графическом отображения парето-подмножества нелокальной нейронной сети в модели INF4 у классов слева сильнее связи между значениями факторов и результатом [9].

Из рисунка 26 видно, что, например, 6 и 9 (переплата и оплата) признаки имеют сходство.

2.5. Кластерный и конструктивный анализ классов

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 1 "ORIGIN-ABQ" в модели: 7 "INF4"

Код	Наименование класса	№	Код класса	Наименование класса	Сходство
1	ORIGIN-ABQ	1	1	ORIGIN-ABQ	100.000
2	ORIGIN-ALB	2	88	DEST-IAH	94.714
3	ORIGIN-AMA	3	65	DEST-ABQ	8.301
4	ORIGIN-AUS	4	79	DEST-DAL	4.258
5	ORIGIN-BDL	5	16	ORIGIN-DEN	3.920
6	ORIGIN-BHM	6	67	DEST-AMA	3.134
7	ORIGIN-BNA	7	53	ORIGIN-SAT	2.969
8	ORIGIN-BOI	8	63	ORIGIN-TUL	2.722
9	ORIGIN-BUF	9	44	ORIGIN-PDX	2.463
10	ORIGIN-BUR	10	82	DEST-ELP	2.171
11	ORIGIN-BWI	11	64	ORIGIN-TUS	1.936
12	ORIGIN-CLE	12	70	DEST-BHM	1.739
13	ORIGIN-CMH	13	123	DEST-SLC	1.713
14	ORIGIN-CRP	14	89	DEST-IND	1.704
15	ORIGIN-DAL	15	128	DEST-TUL	1.622
16	ORIGIN-DEN	16	109	DEST-PDX	1.539
17	ORIGIN-DTW	17	52	ORIGIN-SAN	1.511
18	ORIGIN-ELP	18	21	ORIGIN-HOU	1.510
19	ORIGIN-FLL	19	55	ORIGIN-SEA	1.308
20	ORIGIN-GEG	20	119	DEST-SDF	1.230
21	ORIGIN-HOU	21	46	ORIGIN-PHX	1.192
22	ORIGIN-HRL	22	30	ORIGIN-LBB	1.068
23	ORIGIN-IAD	23	61	ORIGIN-STL	0.999

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ.фильтр по кл.шкале ВЫКЛ.фильтр по кл.шкале Вписать в окно Показать ВСЕ

Рисунок 25. Результаты кластерного и конструктивного анализа классов

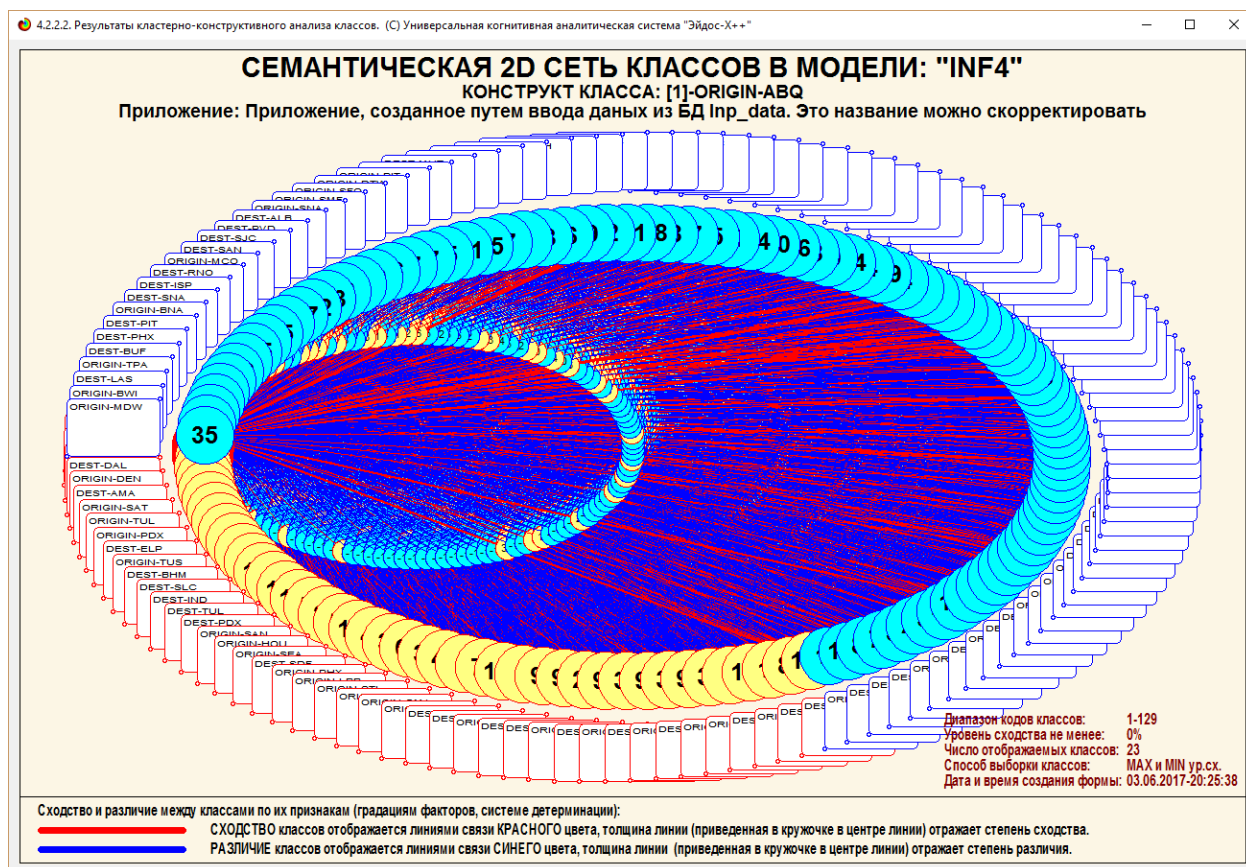


Рисунок 26. Графические результаты кластерного и конструктивного анализа классов

ЗАКЛЮЧЕНИЕ

На сегодняшний день все больше проявилась потребность в системах, не только представляющих информацию, но и выполняющих некоторый ее предварительный анализ, способных давать некоторые советы и рекомендации, осуществлять прогнозирование развития ситуаций, отбирать наиболее перспективные альтернативы решений, взять на себя значительную часть рутинных операций, а также функции предварительного анализа и оценок.

В данной курсовой работе была разработана системно-когнитивной модель прогноза загруженности аэропортов, востребованности рейсов с помощью программного инструментария системы искусственного интеллекта «Эйдос». При этом наиболее достоверной в данном приложении

оказались модели INF4, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Сумма знаний». Точность модели составляет 0,055, что заметно 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. Также обращает на себя внимание, что статистические модели в данном приложении дают примерно на 17% 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.

Обработка данных, их систематизация и визуализация позволили сделать прогнозы и обоснованные выводы по перелетам на самолетах. Таким образом, повышены качество и оперативность принятия решений, а также эффективность управления знаниями.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

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

6. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / 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.

7. Луценко Е.В. Моделирование сложных многофакторных нелинейных объектов управления на основе фрагментированных зашумленных эмпирических данных большой размерности в системно-когнитивном анализе и интеллектуальной системе «Эйдос-Х++» / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 164 – 188. – IDA [article ID]: 0911307012. – Режим доступа: <http://ej.kubagro.ru/2013/07/pdf/12.pdf>, 1,562 у.п.л.

8. Луценко Е.В. Нечеткое мультиклассовое обобщение классической F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе «Эйдос» / Луценко Е.В. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №09(123). – Режим доступа: <http://ej.kubagro.ru/2016/09/pdf/01.pdf>, 1,813 у.п.л. – IDA [article ID]: 1231609001. <http://dx.doi.org/10.21515/1990-4665-123-001>

9. 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.