

УДК 004.8

01.00.00 Физико-математические науки

**ИДЕНТИФИКАЦИЯ СОРТОВ ИРИСОВ ПО
ВНЕШНИМ ПРИЗНАКАМ С
ПРИМЕНЕНИЕМ АСК-АНАЛИЗА И
ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ
«ЭЙДОС» (ДАННЫЕ РЕПОЗИТОРИЯ UCI)**

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Scopus Author ID: 57191193316
РИНЦ SPIN-код: 9523-7101
prof.lutsenko@gmail.com

*Кубанский государственный аграрный
университет имени И.Т.Трубилина, Краснодар,
Россия*

Муренько Иван Викторович
студент 4-го курса
murenkoivan@yandex.ru

*Кубанский государственный аграрный
университет имени И.Т.Трубилина, Краснодар,
Россия*

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив математических моделей систем искусственного интеллекта, то возникает необходимость оценки качества этих моделей, для чего необходимо их сравнение. Для достижения поставленной цели необходимы свободный доступ к исходным данным и методика, которая позволяет преобразовать эти данные в форму, необходимую для их обработки в системе искусственного интеллекта. Удачным выбором для этих целей является база данных тестовых задач для систем искусственного интеллекта репозитория UCI. В данной работе использована база данных «Iris Data Set» из банка исходных данных по задачам искусственного интеллекта – репозитория UCI, на основе которой решаются задачи формализации предметной области (разработки классификационных и описательных шкал и градаций и кодирование исходных данных с их использованием, в результате чего формируется обучающая выборка, по сути представляющая собой нормализованные исходные данные), синтеза и верификации статистических и системно-когнитивных моделей предметной области, идентификации конкретных цветов с классами, в качестве которых выступают сорта Ириса, а также исследования предметной области путем исследования ее модели. Для решения этих задач применяется автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальная

UDC 004.8

Physics and mathematical sciences

**IDENTIFICATION OF VARIETIES OF IRISES
BY THEIR APPEARANCE WITH THE USE
OF ASC-ANALYSIS AND "EIDOS"
INTELLECTUAL SYSTEM (REPOSITORY
UCI DATA)**

Lutsenko Eugeny Veniaminovich
Dr.Sci.Econ., Cand.Tech.Sci., professor
Scopus Author ID: 57191193316
RSCI SPIN-code: 9523-7101
prof.lutsenko@gmail.com

*Kuban State Agrarian University, Krasnodar,
Russia*

Mureniko Ivan Viktorovich
student of the 4th year
murenkoivan@yandex.ru

*Kuban State Agrarian University, Krasnodar,
Russia*

The creation of artificial intelligence systems is one of important and perspective directions of development of modern information technology. Since there are many alternatives of mathematical models of systems of artificial intelligence, there is a need to assess the quality of these models, which requires their comparison. To achieve this goal we require free access to the source data and methodology, which allows to convert these data into a form needed for processing in artificial intelligence. A good choice for these purposes is a database of test problems for systems of artificial intelligence of repository of UCI. In this work we used the database "Iris Data Set" from the bank's original task of artificial intelligence – UCI repository, which solved the problem of formalization of the subject area (development of classification and descriptive dials and graduations and the encoding of the source data, resulting training sample, essentially representing a normalized source data), synthesis and verification statistical and system-cognitive models of the subject area, identify colors with classes, which serve varieties of Iris, as well as studies of the subject area by studying its model. To solve these problems we used the automated system-cognitive analysis (ASC-analysis) and its programmatic Toolkit – intellectual system called "Eidos"

система «Эйдос»

Ключевые слова: ИДЕНТИФИКАЦИЯ СОРТОВ ИРИСОВ, ВНЕШНИЕ ПРИЗНАКИ, АСК-АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «ЭЙДОС», РЕПОЗИТОРИЙ UCI

Keywords: IDENTIFICATION OF VARIETIES OF IRISES, OUTWARD SIGNS, ASK-ANALYSIS, INTELLECTUAL SYSTEM, "EIDOS", UCI REPOSITORY

Doi: 10.21515/1990-4665-123-121

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	2
1. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	3
1.1. ОПИСАНИЕ РЕШЕНИЯ	3
1.2. ПРЕОБРАЗОВАНИЕ ИСХОДНЫХ ДАННЫХ ИЗ HTML-ФОРМАТА В ФАЙЛ ИСХОДНЫХ ДАННЫХ MS EXCEL	3
1.3. СИНТЕЗ И ВЕРИФИКАЦИЯ СТАТИСТИЧЕСКИХ И ИНТЕЛЛЕКТУАЛЬНЫХ МОДЕЛЕЙ.....	14
1.4. ВИДЫ МОДЕЛЕЙ СИСТЕМЫ «ЭЙДОС»	16
1.5. 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	30
ЗАКЛЮЧЕНИЕ	32
ЛИТЕРАТУРА	32

ВВЕДЕНИЕ

Создание систем искусственного интеллекта является одним из важных и перспективных направлений развития современных информационных технологий. Так как существует множество альтернатив систем искусственного интеллекта, то возникает необходимость оценки качества математических моделей этих систем. В данной работе рассмотрено решение задачи идентификации классов цветов ириса.

Для достижения поставленной цели необходимы свободный доступ к тестовым исходным данным и методика, которая поможет преобразовать эти данные в форму, которая необходима для работы в системе искусственного интеллекта. Удачным выбором является база данных тестовых задач для систем искусственного интеллекта репозитория UCI.

В данной работе использована база данных «Iris Data Set» из банка исходных данных по задачам искусственного интеллекта – репозитория UCI.

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

1. Преобразование исходных данных из HTML-формата 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 HTML-формата в файл исходных данных MS Excel

Из банка исходных данных по задачам искусственного интеллекта – репозитория UCI получаем исходную информацию по базе данных «Iris Data Set», которую оставим без изменений.

Общее описание задачи:

Файл «iris.names»:

1. Title: Iris Plants Database
Updated Sept 21 by C.Blake - Added discrepancy information

2. Sources:

- (a) Creator: R.A. Fisher
- (b) Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)
- (c) Date: July, 1988

3. Past Usage:

- Publications: too many to mention!!! Here are a few.

1. Fisher, R.A. "The use of multiple measurements in taxonomic problems" Annual Eugenics, 7, Part II, 179-188 (1936); also in "Contributions to Mathematical Statistics" (John Wiley, NY, 1950).

2. Duda, R.O., & Hart, P.E. (1973) Pattern Classification and Scene Analysis.

(Q327.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page 218.

3. Dasarathy, B.V. (1980) "Nosing Around the Neighborhood: A New System Structure and Classification Rule for Recognition in Partially

Exposed

Environments". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 1, 67-71.

-- Results:

-- very low misclassification rates (0% for the setosa class)

4. Gates, G.W. (1972) "The Reduced Nearest Neighbor Rule". IEEE Transactions on Information Theory, May 1972, 431-433.

-- Results:

-- very low misclassification rates again

5. See also: 1988 MLC Proceedings, 54-64. Cheeseman et al's AUTOCLASS

II

conceptual clustering system finds 3 classes in the data.

4. Relevant Information:

--- This is perhaps the best known database to be found in the pattern recognition literature. Fisher's paper is a classic in the field and is referenced frequently to this day. (See Duda & Hart, for example.) The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are NOT linearly separable from each other.

--- Predicted attribute: class of iris plant.

--- This is an exceedingly simple domain.

--- This data differs from the data presented in Fishers article

(identified by Steve Chadwick, spchadwick@espeedaz.net)

The 35th sample should be: 4.9,3.1,1.5,0.2,"Iris-setosa"

where the error is in the fourth feature.

The 38th sample: 4.9,3.6,1.4,0.1,"Iris-setosa"

where the errors are in the second and third features.

5. Number of Instances: 150 (50 in each of three classes)

6. Number of Attributes: 4 numeric, predictive attributes and the class

7. Attribute Information:

1. sepal length in cm

2. sepal width in cm

3. petal length in cm

4. petal width in cm

5. class:

-- Iris Setosa

-- Iris Versicolour

-- Iris Virginica

8. Missing Attribute Values: None

Summary Statistics:

	Min	Max	Mean	SD	Class Correlation
sepal length:	4.3	7.9	5.84	0.83	0.7826
sepal width:	2.0	4.4	3.05	0.43	-0.4194
petal length:	1.0	6.9	3.76	1.76	0.9490 (high!)
petal width:	0.1	2.5	1.20	0.76	0.9565 (high!)

9. Class Distribution: 33.3% for each of 3 classes.

Обучающая выборка представлена в таблице 1:

Таблица 1 – iris.data

n	sepal length	sepal width	petal length	petal width	class
1	5,1	3,5	1,4	0,2	Iris-setosa
2	4,9	3,0	1,4	0,2	Iris-setosa
3	4,7	3,2	1,3	0,2	Iris-setosa
4	4,6	3,1	1,5	0,2	Iris-setosa
5	5,0	3,6	1,4	0,2	Iris-setosa
6	5,4	3,9	1,7	0,4	Iris-setosa
7	4,6	3,4	1,4	0,3	Iris-setosa
8	5,0	3,4	1,5	0,2	Iris-setosa
9	4,4	2,9	1,4	0,2	Iris-setosa
10	4,9	3,1	1,5	0,1	Iris-setosa
11	5,4	3,7	1,5	0,2	Iris-setosa
12	4,8	3,4	1,6	0,2	Iris-setosa
13	4,8	3,0	1,4	0,1	Iris-setosa
14	4,3	3,0	1,1	0,1	Iris-setosa
15	5,8	4,0	1,2	0,2	Iris-setosa
16	5,7	4,4	1,5	0,4	Iris-setosa
17	5,4	3,9	1,3	0,4	Iris-setosa
18	5,1	3,5	1,4	0,3	Iris-setosa
19	5,7	3,8	1,7	0,3	Iris-setosa
20	5,1	3,8	1,5	0,3	Iris-setosa
21	5,4	3,4	1,7	0,2	Iris-setosa
22	5,1	3,7	1,5	0,4	Iris-setosa
23	4,6	3,6	1,0	0,2	Iris-setosa
24	5,1	3,3	1,7	0,5	Iris-setosa
25	4,8	3,4	1,9	0,2	Iris-setosa
26	5,0	3,0	1,6	0,2	Iris-setosa
27	5,0	3,4	1,6	0,4	Iris-setosa
28	5,2	3,5	1,5	0,2	Iris-setosa
29	5,2	3,4	1,4	0,2	Iris-setosa
30	4,7	3,2	1,6	0,2	Iris-setosa
31	4,8	3,1	1,6	0,2	Iris-setosa
32	5,4	3,4	1,5	0,4	Iris-setosa
33	5,2	4,1	1,5	0,1	Iris-setosa
34	5,5	4,2	1,4	0,2	Iris-setosa
35	4,9	3,1	1,5	0,1	Iris-setosa
36	5,0	3,2	1,2	0,2	Iris-setosa
37	5,5	3,5	1,3	0,2	Iris-setosa
38	4,9	3,1	1,5	0,1	Iris-setosa
39	4,4	3,0	1,3	0,2	Iris-setosa
40	5,1	3,4	1,5	0,2	Iris-setosa
41	5,0	3,5	1,3	0,3	Iris-setosa
42	4,5	2,3	1,3	0,3	Iris-setosa
43	4,4	3,2	1,3	0,2	Iris-setosa

44	5,0	3,5	1,6	0,6	Iris-setosa
45	5,1	3,8	1,9	0,4	Iris-setosa
46	4,8	3,0	1,4	0,3	Iris-setosa
47	5,1	3,8	1,6	0,2	Iris-setosa
48	4,6	3,2	1,4	0,2	Iris-setosa
49	5,3	3,7	1,5	0,2	Iris-setosa
50	5,0	3,3	1,4	0,2	Iris-setosa
51	7,0	3,2	4,7	1,4	Iris-versicolor
52	6,4	3,2	4,5	1,5	Iris-versicolor
53	6,9	3,1	4,9	1,5	Iris-versicolor
54	5,5	2,3	4,0	1,3	Iris-versicolor
55	6,5	2,8	4,6	1,5	Iris-versicolor
56	5,7	2,8	4,5	1,3	Iris-versicolor
57	6,3	3,3	4,7	1,6	Iris-versicolor
58	4,9	2,4	3,3	1,0	Iris-versicolor
59	6,6	2,9	4,6	1,3	Iris-versicolor
60	5,2	2,7	3,9	1,4	Iris-versicolor
61	5,0	2,0	3,5	1,0	Iris-versicolor
62	5,9	3,0	4,2	1,5	Iris-versicolor
63	6,0	2,2	4,0	1,0	Iris-versicolor
64	6,1	2,9	4,7	1,4	Iris-versicolor
65	5,6	2,9	3,6	1,3	Iris-versicolor
66	6,7	3,1	4,4	1,4	Iris-versicolor
67	5,6	3,0	4,5	1,5	Iris-versicolor
68	5,8	2,7	4,1	1,0	Iris-versicolor
69	6,2	2,2	4,5	1,5	Iris-versicolor
70	5,6	2,5	3,9	1,1	Iris-versicolor
71	5,9	3,2	4,8	1,8	Iris-versicolor
72	6,1	2,8	4,0	1,3	Iris-versicolor
73	6,3	2,5	4,9	1,5	Iris-versicolor
74	6,1	2,8	4,7	1,2	Iris-versicolor
75	6,4	2,9	4,3	1,3	Iris-versicolor
76	6,6	3,0	4,4	1,4	Iris-versicolor
77	6,8	2,8	4,8	1,4	Iris-versicolor
78	6,7	3,0	5,0	1,7	Iris-versicolor
79	6,0	2,9	4,5	1,5	Iris-versicolor
80	5,7	2,6	3,5	1,0	Iris-versicolor
81	5,5	2,4	3,8	1,1	Iris-versicolor
82	5,5	2,4	3,7	1,0	Iris-versicolor
83	5,8	2,7	3,9	1,2	Iris-versicolor
84	6,0	2,7	5,1	1,6	Iris-versicolor
85	5,4	3,0	4,5	1,5	Iris-versicolor
86	6,0	3,4	4,5	1,6	Iris-versicolor
87	6,7	3,1	4,7	1,5	Iris-versicolor
88	6,3	2,3	4,4	1,3	Iris-versicolor
89	5,6	3,0	4,1	1,3	Iris-versicolor
90	5,5	2,5	4,0	1,3	Iris-versicolor
91	5,5	2,6	4,4	1,2	Iris-versicolor
92	6,1	3,0	4,6	1,4	Iris-versicolor
93	5,8	2,6	4,0	1,2	Iris-versicolor
94	5,0	2,3	3,3	1,0	Iris-versicolor
95	5,6	2,7	4,2	1,3	Iris-versicolor
96	5,7	3,0	4,2	1,2	Iris-versicolor
97	5,7	2,9	4,2	1,3	Iris-versicolor
98	6,2	2,9	4,3	1,3	Iris-versicolor

99	5,1	2,5	3,0	1,1	Iris-versicolor
100	5,7	2,8	4,1	1,3	Iris-versicolor
101	6,3	3,3	6,0	2,5	Iris-virginica
102	5,8	2,7	5,1	1,9	Iris-virginica
103	7,1	3,0	5,9	2,1	Iris-virginica
104	6,3	2,9	5,6	1,8	Iris-virginica
105	6,5	3,0	5,8	2,2	Iris-virginica
106	7,6	3,0	6,6	2,1	Iris-virginica
107	4,9	2,5	4,5	1,7	Iris-virginica
108	7,3	2,9	6,3	1,8	Iris-virginica
109	6,7	2,5	5,8	1,8	Iris-virginica
110	7,2	3,6	6,1	2,5	Iris-virginica
111	6,5	3,2	5,1	2,0	Iris-virginica
112	6,4	2,7	5,3	1,9	Iris-virginica
113	6,8	3,0	5,5	2,1	Iris-virginica
114	5,7	2,5	5,0	2,0	Iris-virginica
115	5,8	2,8	5,1	2,4	Iris-virginica
116	6,4	3,2	5,3	2,3	Iris-virginica
117	6,5	3,0	5,5	1,8	Iris-virginica
118	7,7	3,8	6,7	2,2	Iris-virginica
119	7,7	2,6	6,9	2,3	Iris-virginica
120	6,0	2,2	5,0	1,5	Iris-virginica
121	6,9	3,2	5,7	2,3	Iris-virginica
122	5,6	2,8	4,9	2,0	Iris-virginica
123	7,7	2,8	6,7	2,0	Iris-virginica
124	6,3	2,7	4,9	1,8	Iris-virginica
125	6,7	3,3	5,7	2,1	Iris-virginica
126	7,2	3,2	6,0	1,8	Iris-virginica
127	6,2	2,8	4,8	1,8	Iris-virginica
128	6,1	3,0	4,9	1,8	Iris-virginica
129	6,4	2,8	5,6	2,1	Iris-virginica
130	7,2	3,0	5,8	1,6	Iris-virginica
131	7,4	2,8	6,1	1,9	Iris-virginica
132	7,9	3,8	6,4	2,0	Iris-virginica
133	6,4	2,8	5,6	2,2	Iris-virginica
134	6,3	2,8	5,1	1,5	Iris-virginica
135	6,1	2,6	5,6	1,4	Iris-virginica
136	7,7	3,0	6,1	2,3	Iris-virginica
137	6,3	3,4	5,6	2,4	Iris-virginica
138	6,4	3,1	5,5	1,8	Iris-virginica
139	6,0	3,0	4,8	1,8	Iris-virginica
140	6,9	3,1	5,4	2,1	Iris-virginica
141	6,7	3,1	5,6	2,4	Iris-virginica
142	6,9	3,1	5,1	2,3	Iris-virginica
143	5,8	2,7	5,1	1,9	Iris-virginica
144	6,8	3,2	5,9	2,3	Iris-virginica
145	6,7	3,3	5,7	2,5	Iris-virginica
146	6,7	3,0	5,2	2,3	Iris-virginica
147	6,3	2,5	5,0	1,9	Iris-virginica
148	6,5	3,0	5,2	2,0	Iris-virginica
149	6,2	3,4	5,4	2,3	Iris-virginica
150	5,9	3,0	5,1	1,8	Iris-virginica

На рисунках 1-3 приведены примеры цветов Ирисов разных классов.



Рисунок 1. Ирис класса *setosa*



Рисунок 2. Ирис класса *versicolor*



Рисунок 3. Ирис класса *virginica*

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 html-файла в xls-файл, для чего выполним следующие операции.

Скопируем получившуюся таблицу из MS Word 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 пустую колонку на позиции «А» и автоматически пронумеруем все строки. В результате получим таблицу исходных данных, полностью подготовленную для обработки в системе «Эйдос» и записанную в нужную папку в виде файла нужного типа с нужным именем.

Автоматизированная формализация предметной области путем импорта исходных данных из внешних баз данных в систему "Эйдос".

Для загрузки базы исходных данных в систему «Эйдос» необходимо воспользоваться универсальным программным интерфейсом для ввода данных из внешних баз данных табличного вида, т.е. режимом 2.3.2.2 (рисунок 4):

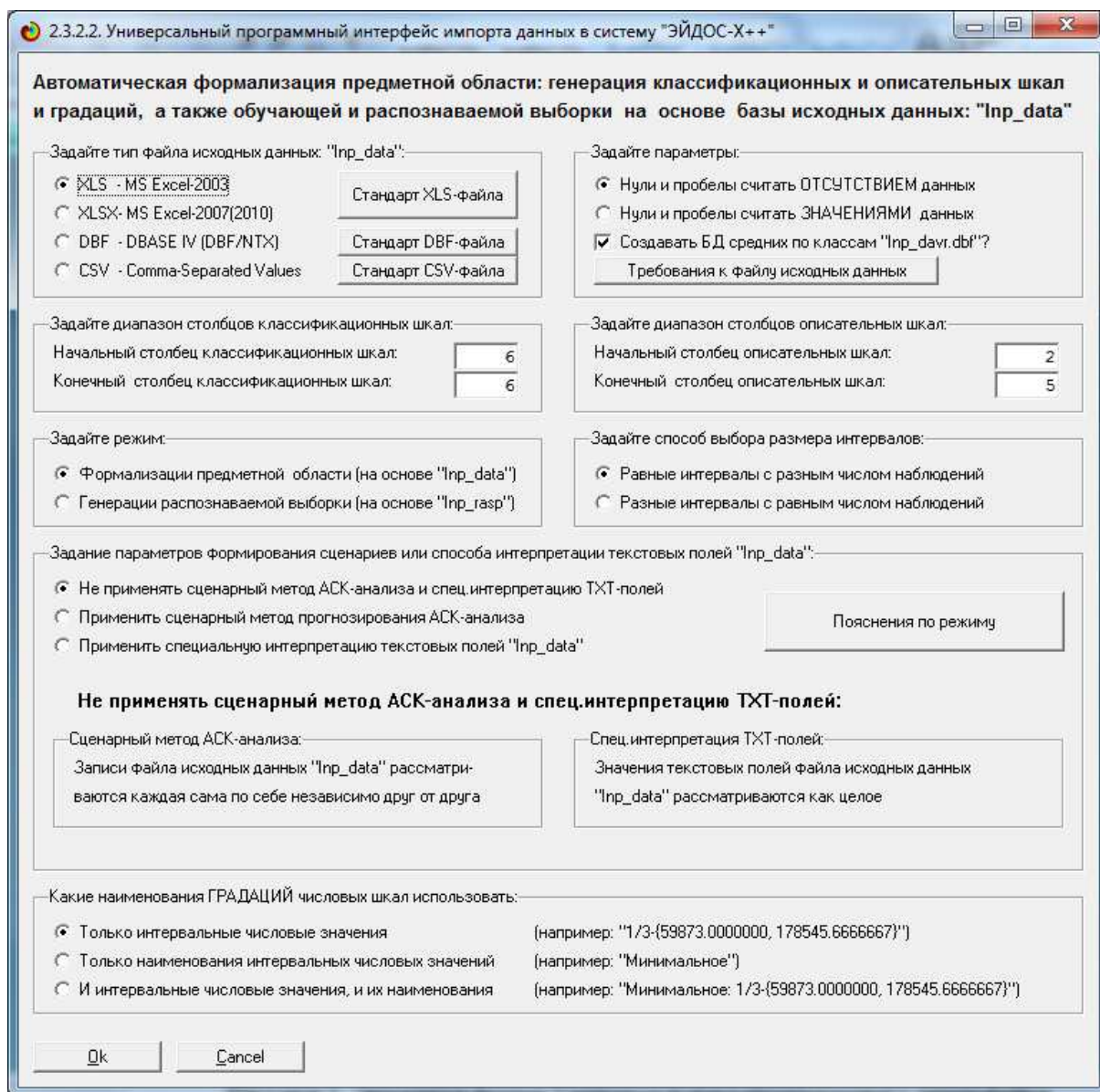


Рисунок 4. Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

В экранной форме, приведенной на рисунке 4 необходимо задать настройки, показанные на рисунке:

– "Задайте тип файла исходных данных Inp_data": "XLS – MS Excel-2003";

– "Задайте диапазон столбцов классификационных шкал": "Начальный столбец классификационных шкал" – 6, "Конечный столбец классификационных шкал" – 7 (последний столбец в таблице);

– "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 2, "Конечный столбец описательных шкал" – 5;

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

После нажать кнопку "ОК". Далее открывается окно, где размещена информация о размерности модели (рисунок 5). В этом окне необходимо нажать кнопку "Выйти на создание модели".

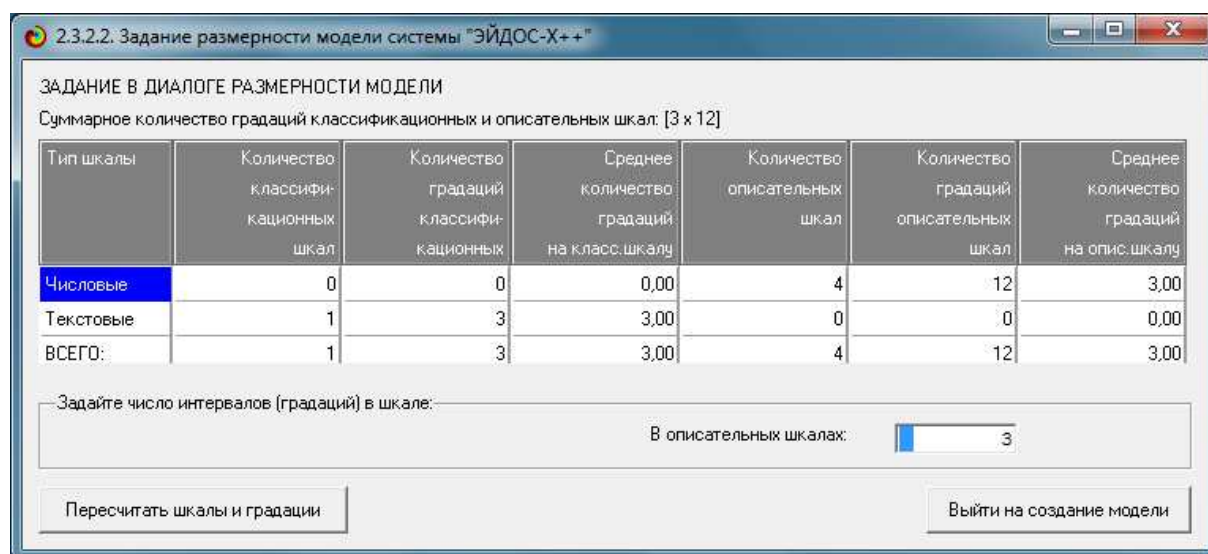


Рисунок 5. Задание размерности модели системы "Эйдос"

Далее открывается окно, отображающее стадию процесса импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос" (рисунок 6), а также прогноз времени завершения этого процесса. В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "ОК".

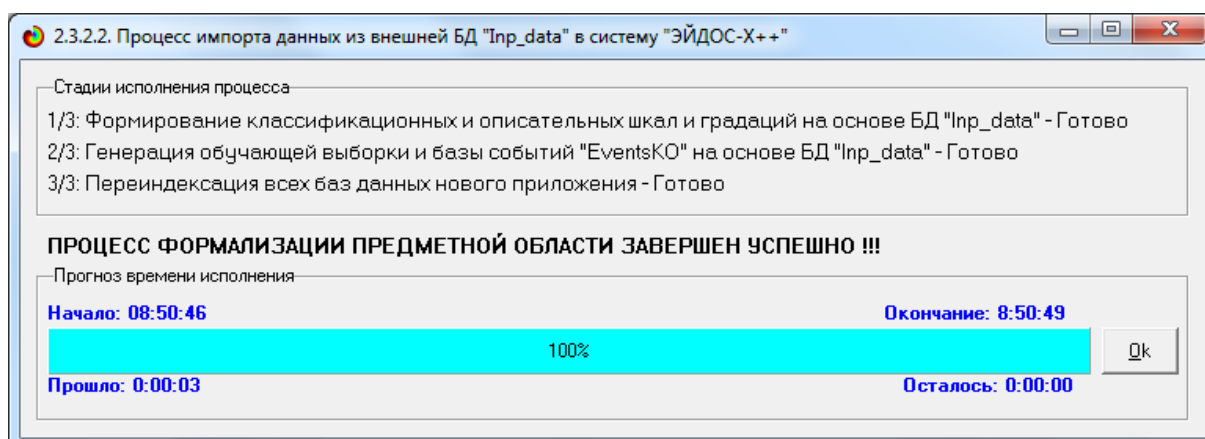


Рисунок 6. Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

В результате формируются классификационные и описательные шкалы и градации, с применением которых исходные данные кодируются и представляются в форме эвентологических баз данных. Этим самым полностью автоматизировано выполняется 2-й этап АСК-анализа «Формализация предметной области». Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1 (рисунок 7).

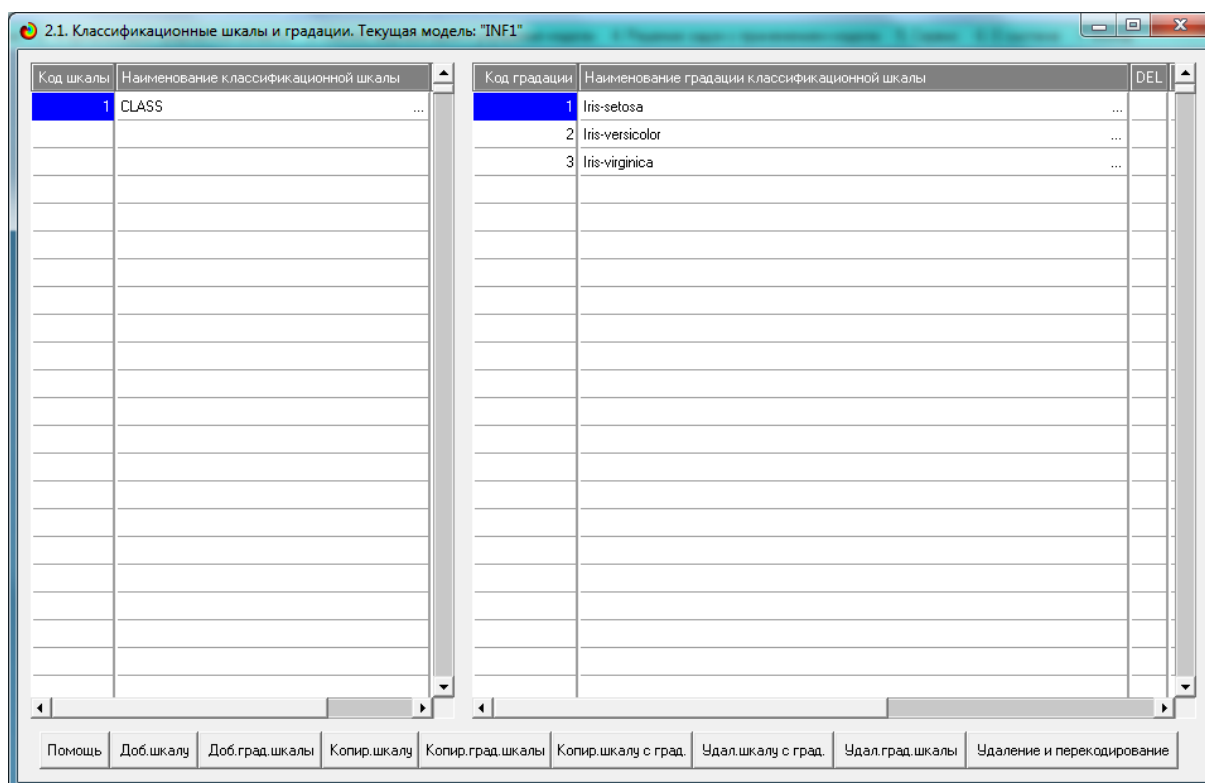


Рисунок 7. Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 8), а для просмотра обучающей выборки – режим 2.3.1 (рисунок 9):

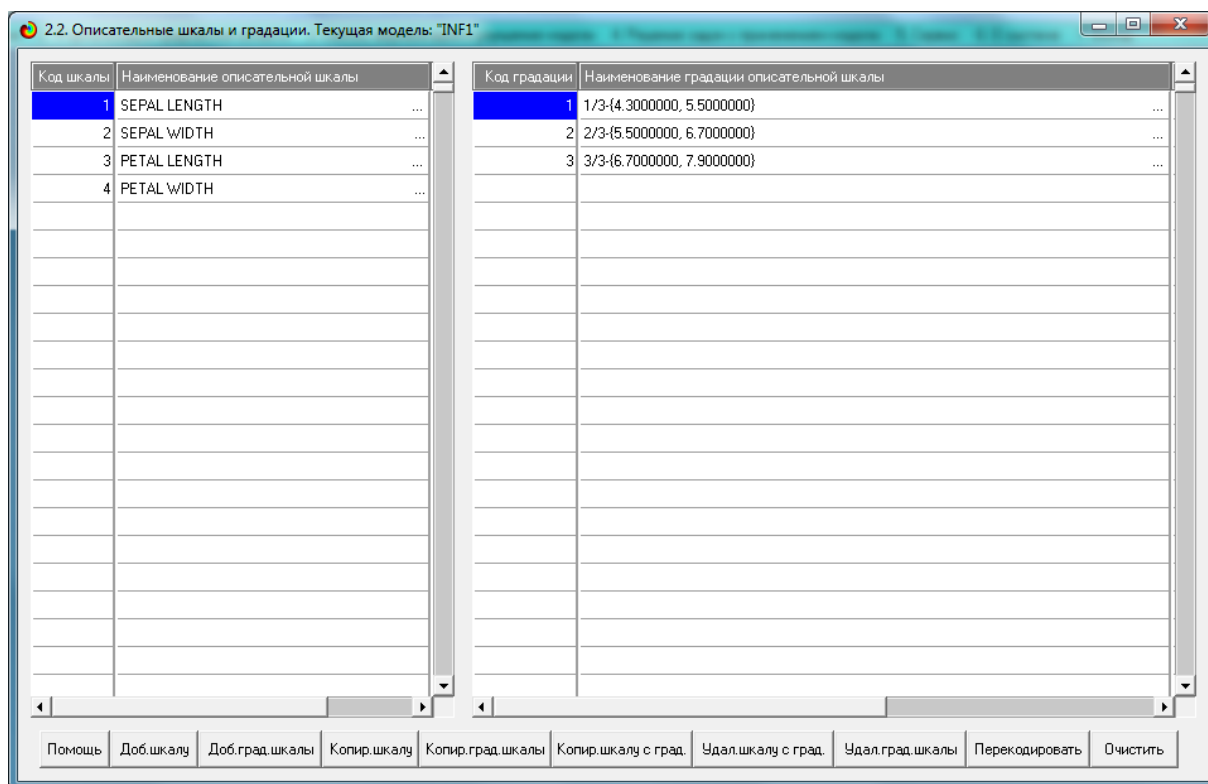


Рисунок 8. Описательные шкалы и градации (фрагмент)

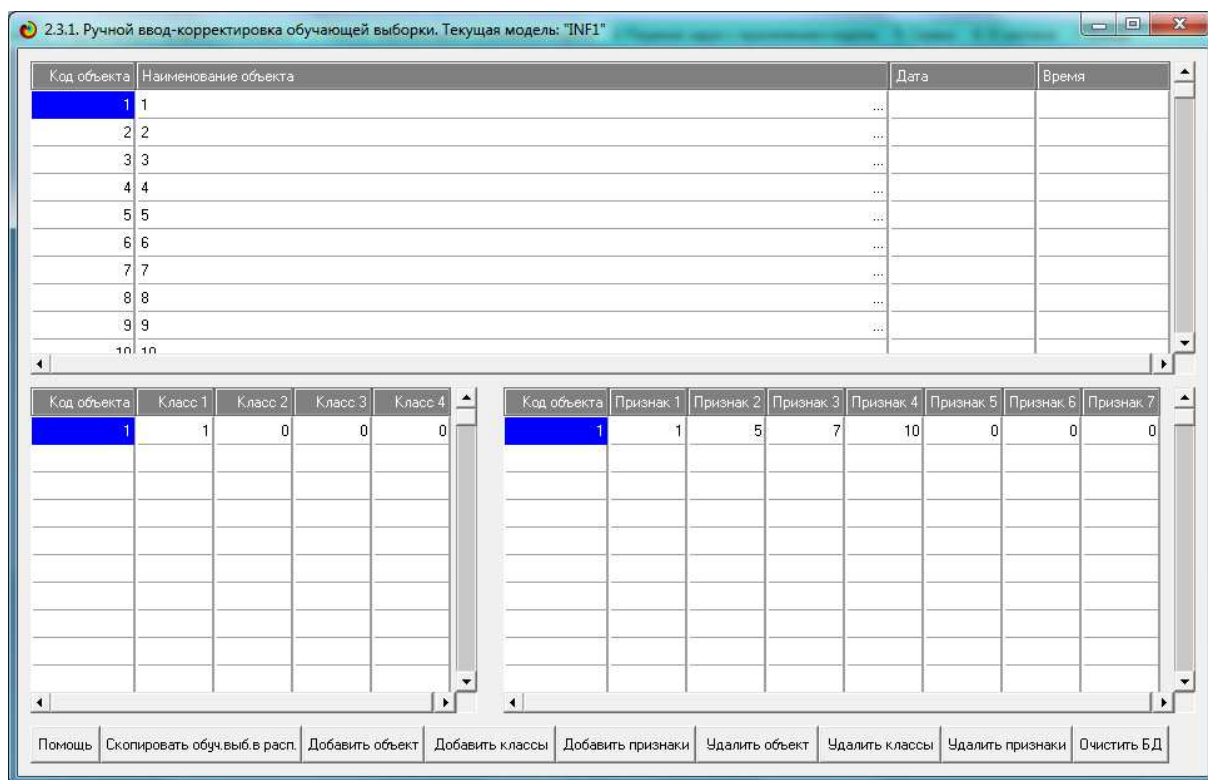


Рисунок 9. Обучающая выборка (фрагмент)

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 (с учетом нелинейности системы [11]).

1.3. 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 10).

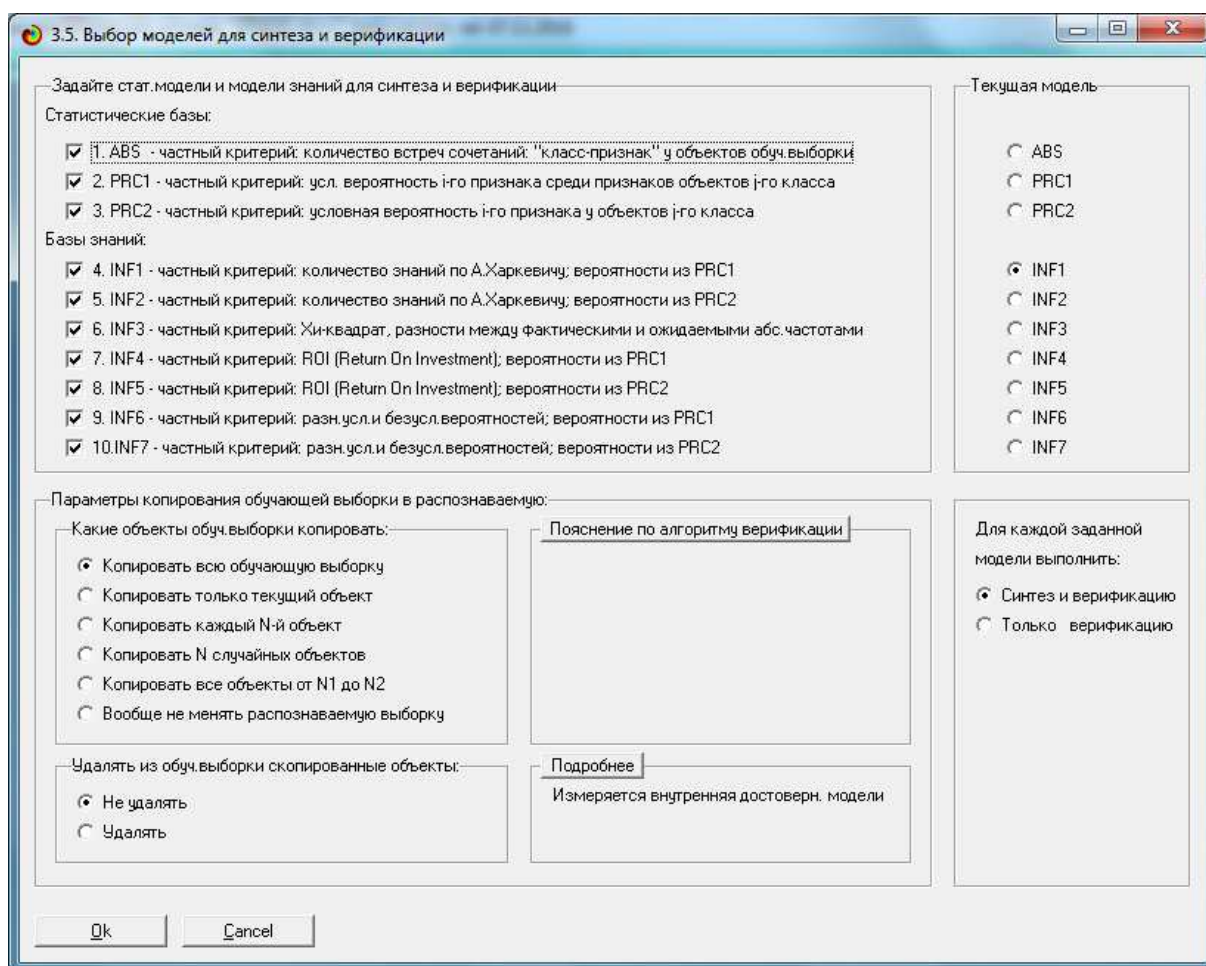


Рисунок 10. Выбор моделей для синтеза и верификации, а также текущей модели

В данном режиме имеется много различных методов верификации моделей, в том числе и поддерживающие бутстрепный метод. Но мы используем параметры по умолчанию, приведенные на рисунке 10. Стадия процесса исполнения режима 3.5 и прогноз времени его окончания отображаются на экранной форме, приведенной на рисунке 11.

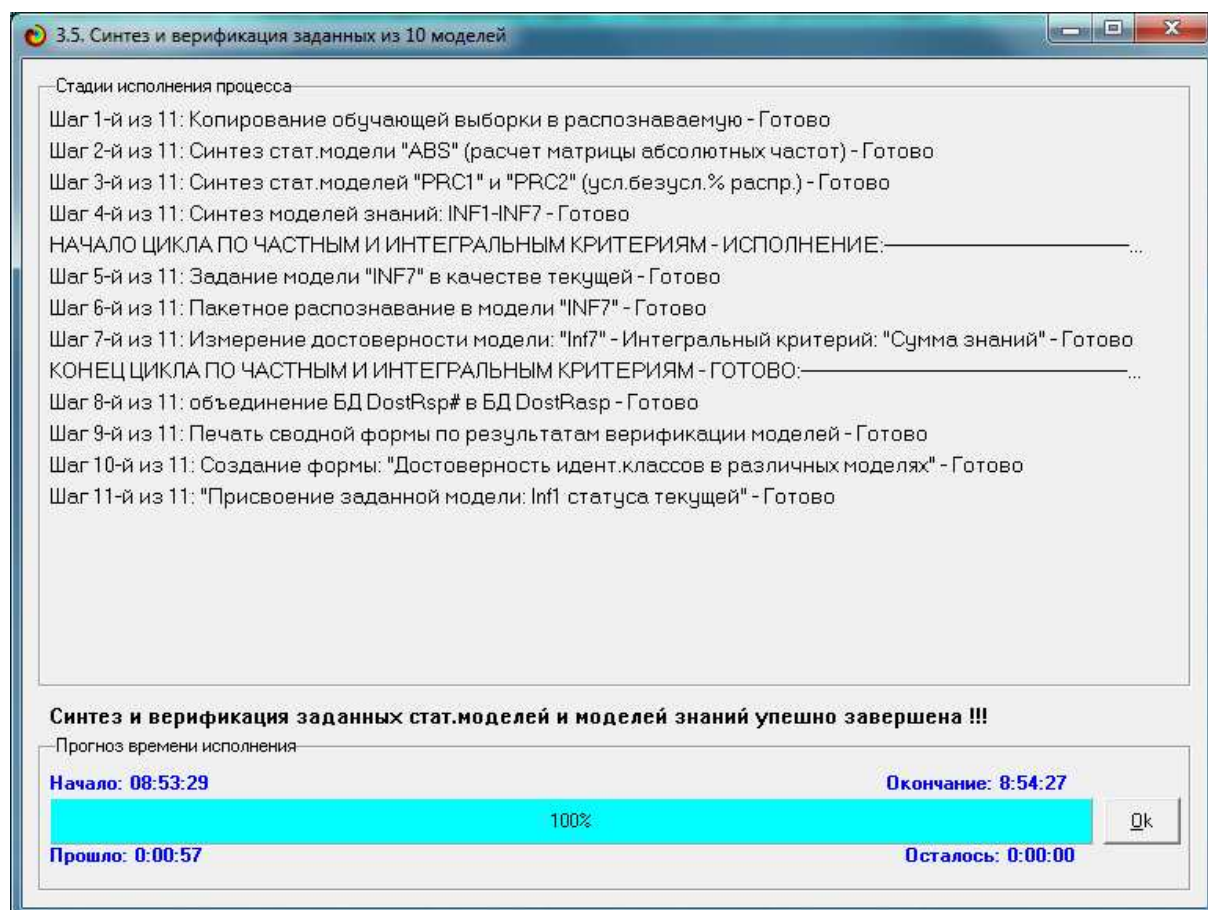


Рисунок 11. Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 11), что синтез и верификация всех 10 моделей на данной задаче заняли 57 секунд. При этом верификация (оценка достоверности моделей) проводилась на всех 150 примерах наблюдения из обучающей выборки. В результате выполнения режима 3.5 созданы все модели, со всеми частными критериями, перечисленные на рисунке 10, но ниже мы приведем лишь некоторые из них (таблицы 2, 3, 4).

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 (таблица 2) в матрицы условных и безусловных процентных распределений, и матрицы знаний (таблицы 3 и 4).

Таблица 2 – Матрица абсолютных частот (модель ABS)

Код признака	Наименование описательной шкалы и градации	1. CLASS IRIS SETOSA	2. CLASS IRIS VERSICO...	3. CLASS IRIS VIRGINICA	Сумма	Среднее	Средн. квадрат. откл.
1	SEPAL LENGTH-1/3-{4.3000000, 5.5000000}	47	11	1	59	19.667	24.194
2	SEPAL LENGTH-2/3-{5.5000000, 6.7000000}	3	36	32	71	23.667	18.009
3	SEPAL LENGTH-3/3-{6.7000000, 7.9000000}		3	17	20	6.667	9.074
4	SEPAL WIDTH-1/3-{2.0000000, 2.8000000}	1	27	19	47	15.667	13.317
5	SEPAL WIDTH-2/3-{2.8000000, 3.6000000}	36	23	29	88	29.333	6.506
6	SEPAL WIDTH-3/3-{3.6000000, 4.4000000}	13		2	15	5.000	7.000
7	PETAL LENGTH-1/3-{1.0000000, 2.9666667}	50			50	16.667	28.868
8	PETAL LENGTH-2/3-{2.9666667, 4.9333333}		48	6	54	18.000	26.153
9	PETAL LENGTH-3/3-{4.9333333, 6.9000000}		2	44	46	15.333	24.846
10	PETAL WIDTH-1/3-{0.1000000, 0.9000000}	50			50	16.667	28.868
11	PETAL WIDTH-2/3-{0.9000000, 1.7000000}		49	5	54	18.000	26.963
12	PETAL WIDTH-3/3-{1.7000000, 2.5000000}		1	45	46	15.333	25.697
	Сумма числа признаков	200	200	200	600		
	Среднее	17	17	17		16.667	
	Среднеквадратичное отклонение	22	19	17			18.989
	Сумма числа объектов обуч.выборки	50	50	50	150		

Таблица 3 – Матрица информативностей (модель INF1) в битах

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

Код признака	Наименование описательной шкалы и градации	1. CLASS IRIS SETOSA	2. CLASS IRIS VERSICO...	3. CLASS IRIS VIRGINICA	Сумма	Среднее	Средн. квадр. откл.
1	SEPAL LENGTH-1/3-(4.3000000, 5.5000000)	0.216	-0.144	-0.738	-0.666	-0.222	0.482
2	SEPAL LENGTH-2/3-(5.5000000, 6.7000000)	-0.512	0.104	0.075	-0.333	-0.111	0.347
3	SEPAL LENGTH-3/3-(6.7000000, 7.9000000)		-0.198	0.232	0.034	0.011	0.215
4	SEPAL WIDTH-1/3-(2.0000000, 2.8000000)	-0.682	0.135	0.048	-0.499	-0.166	0.448
5	SEPAL WIDTH-2/3-(2.8000000, 3.6000000)	0.051	-0.060	-0.003	-0.012	-0.004	0.056
6	SEPAL WIDTH-3/3-(3.6000000, 4.4000000)	0.237		-0.227	0.010	0.003	0.232
7	PETAL LENGTH-1/3-(1.0000000, 2.9666667)	0.272			0.272	0.091	0.157
8	PETAL LENGTH-2/3-(2.9666667, 4.9333333)		0.243	-0.272	-0.029	-0.010	0.258
9	PETAL LENGTH-3/3-(4.9333333, 6.9000000)		-0.505	0.261	-0.243	-0.081	0.389
10	PETAL WIDTH-1/3-(0.1000000, 0.9000000)	0.272			0.272	0.091	0.157
11	PETAL WIDTH-2/3-(0.9000000, 1.7000000)		0.248	-0.317	-0.069	-0.023	0.283
12	PETAL WIDTH-3/3-(1.7000000, 2.5000000)		-0.676	0.267	-0.410	-0.137	0.486
	Сумма	-0.146	-0.853	-0.675	-1.674		
	Среднее	-0.012	-0.071	-0.056		-0.047	
	Среднеквадратичное отклонение	0.298	0.280	0.291			0.283

Помощь MS Excel MS Word

Таблица 4 – Матрица знаний (модель INF3)

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

Код признака	Наименование описательной шкалы и градации	1. CLASS IRIS SETOSA	2. CLASS IRIS VERSICO...	3. CLASS IRIS VIRGINICA	Сумма	Среднее	Средн. квадр. откл.
1	SEPAL LENGTH-1/3-(4.3000000, 5.5000000)	27.333	-8.667	-18.667			24.194
2	SEPAL LENGTH-2/3-(5.5000000, 6.7000000)	-20.667	12.333	8.333			18.009
3	SEPAL LENGTH-3/3-(6.7000000, 7.9000000)	-6.667	-3.667	10.333			9.074
4	SEPAL WIDTH-1/3-(2.0000000, 2.8000000)	-14.667	11.333	3.333			13.317
5	SEPAL WIDTH-2/3-(2.8000000, 3.6000000)	6.667	-6.333	-0.333			6.506
6	SEPAL WIDTH-3/3-(3.6000000, 4.4000000)	8.000	-5.000	-3.000			7.000
7	PETAL LENGTH-1/3-(1.0000000, 2.9666667)	33.333	-16.667	-16.667			28.868
8	PETAL LENGTH-2/3-(2.9666667, 4.9333333)	-18.000	30.000	-12.000			26.153
9	PETAL LENGTH-3/3-(4.9333333, 6.9000000)	-15.333	-13.333	28.667			24.846
10	PETAL WIDTH-1/3-(0.1000000, 0.9000000)	33.333	-16.667	-16.667			28.868
11	PETAL WIDTH-2/3-(0.9000000, 1.7000000)	-18.000	31.000	-13.000			26.963
12	PETAL WIDTH-3/3-(1.7000000, 2.5000000)	-15.333	-14.333	29.667			25.697
	Сумма						
	Среднее						
	Среднеквадратичное отклонение	21.068	17.146	16.829			17.914

Помощь MS Excel MS Word

1.5. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей, отличающихся частными критериями с двумя приведенными выше интегральными критериями приведены на рисунке 12.

Наименование модели и частного критерия	Интегральный критерий	Всего логически объектов выборки	Число истинно-положительн. решений (TP)	Число истинно-отрицательн. решений (TN)	Число ложн. положительных решений (FP)	Число ложн. отрицательн. решений (FN)	Точность модели	Полнота модели	Фильтра Ван Ризбергера	Сумма модулей уровней схожести истинно-полож. решений (ST...)	Сумма модулей уровней схожести истинно-отриц. решений (ST...)	Сумма модулей схожести уровней схожести истинно-полож. решений (SFP)	Сумма модулей схожести уровней схожести истинно-отриц. решений (SF...)	ST... мод.
ABS - частный критерий: количество встреч сочетаний "класс..."	Корреляция abs частот с обр...	150	149	268	32	1	0.823	0.993	0.900	122.076	88.234	6.895	0.396	
1. ABS - частный критерий: количество встреч сочетаний "класс..."	Сумма abs частот по признак...	150	150		300		0.333	1.000	0.500	123.607		63.082		
2. PRC1 - частный критерий: усл. вероятность n-го признака сред...	Корреляция усл.отн. частот с о...	150	149	268	32	1	0.823	0.993	0.900	122.076	88.234	6.895	0.396	
2. PRC1 - частный критерий: усл. вероятность n-го признака сред...	Сумма усл.отн. частот по при...	150	150		300		0.333	1.000	0.500	123.607		63.082		
3. PRC2 - частный критерий: усл. вероятность n-го признака...	Корреляция усл.отн. частот с о...	150	149	268	32	1	0.823	0.993	0.900	122.071	88.231	6.895	0.396	
3. PRC2 - частный критерий: усл. вероятность n-го признака...	Сумма усл.отн. частот по при...	150	150		300		0.333	1.000	0.500	123.607		63.082		
4. INF1 - частный критерий: количество знаний по А/Архивному в...	Семантический резонанс зна...	150	148	220	80	2	0.649	0.987	0.783	101.586	118.539	9.274	0.970	
4. INF1 - частный критерий: количество знаний по А/Архивному в...	Сумма знаний	150	146	276	24	4	0.859	0.973	0.913	65.422	131.641	1.663	1.320	
5. INF2 - частный критерий: количество знаний по А/Архивному в...	Семантический резонанс зна...	150	148	220	80	2	0.649	0.987	0.783	101.586	118.539	9.274	0.970	
5. INF2 - частный критерий: количество знаний по А/Архивному в...	Сумма знаний	150	146	276	24	4	0.859	0.973	0.913	65.422	131.641	1.663	1.320	
6. INF3 - частный критерий: %инквадрат, разности между факти...	Семантический резонанс зна...	150	149	287	13	1	0.920	0.993	0.955	123.059	127.849	5.135	0.485	
6. INF3 - частный критерий: %инквадрат, разности между факти...	Сумма знаний	150	149	287	13	1	0.920	0.993	0.955	110.513	114.366	4.248	0.395	
7. INF4 - частный критерий: ROI (Return On Investment); веротно...	Семантический резонанс зна...	150	149	288	12	1	0.925	0.993	0.958	105.655	119.937	3.331	0.612	
7. INF4 - частный критерий: ROI (Return On Investment); веротно...	Сумма знаний	150	149	282	48	1	0.756	0.993	0.859	97.615	40.567	5.490	0.304	
8. INF5 - частный критерий: ROI (Return On Investment); веротно...	Семантический резонанс зна...	150	149	288	12	1	0.925	0.993	0.958	105.655	119.937	3.331	0.612	
8. INF5 - частный критерий: ROI (Return On Investment); веротно...	Сумма знаний	150	149	282	48	1	0.756	0.993	0.859	97.615	40.567	5.490	0.304	
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	150	149	287	13	1	0.920	0.993	0.955	117.635	125.625	3.982	0.725	
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	150	149	282	48	1	0.756	0.993	0.859	110.513	50.072	7.565	0.395	
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	150	149	287	13	1	0.920	0.993	0.955	117.635	125.625	3.982	0.725	
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	150	149	282	48	1	0.756	0.993	0.859	110.513	50.072	7.565	0.395	

Рисунок 12. Оценки достоверности моделей

Наиболее достоверными в данном приложении оказались модели INF4 и INF5 при интегральном критерии «Семантический резонанс знаний». При этом точность модели (F-мера Ван Ризбергера) составляет **0,958**. 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 F-мера Ван Ризбергера и L-мера, представляющая собой ее нечеткое мультиклассовое обобщение, предложенное профессором Е.В.Луценко [10] (рисунок 13).

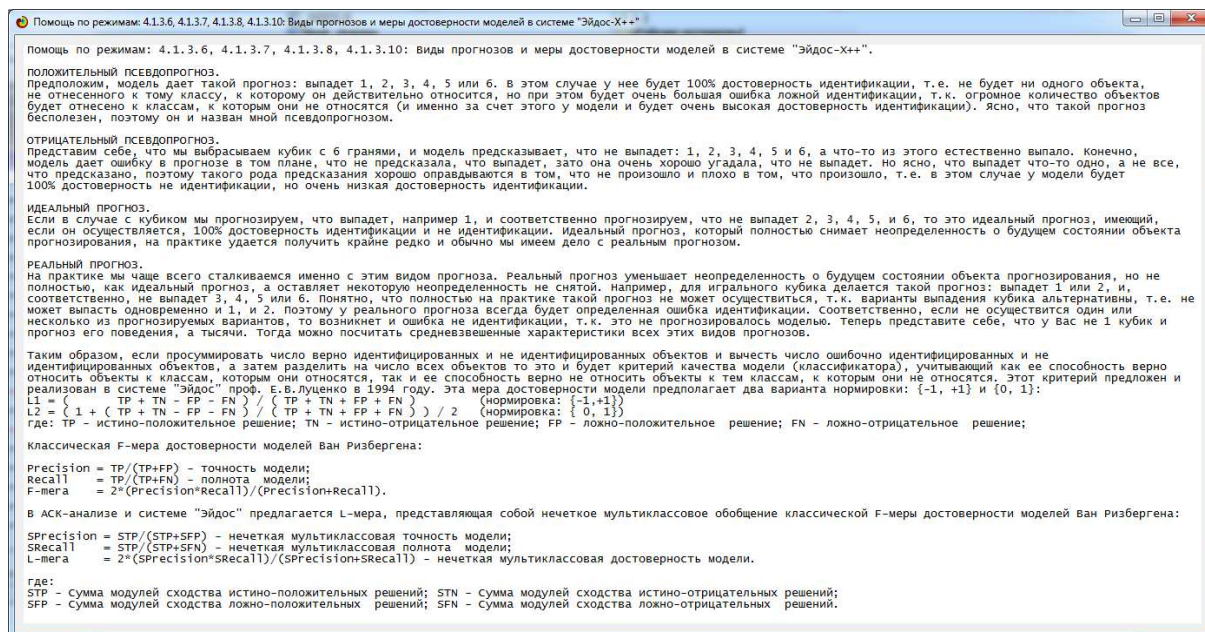


Рисунок 13. Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

Также обращает на себя 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. На рисунке 14 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных ситуаций в наиболее достоверной модели INF4.

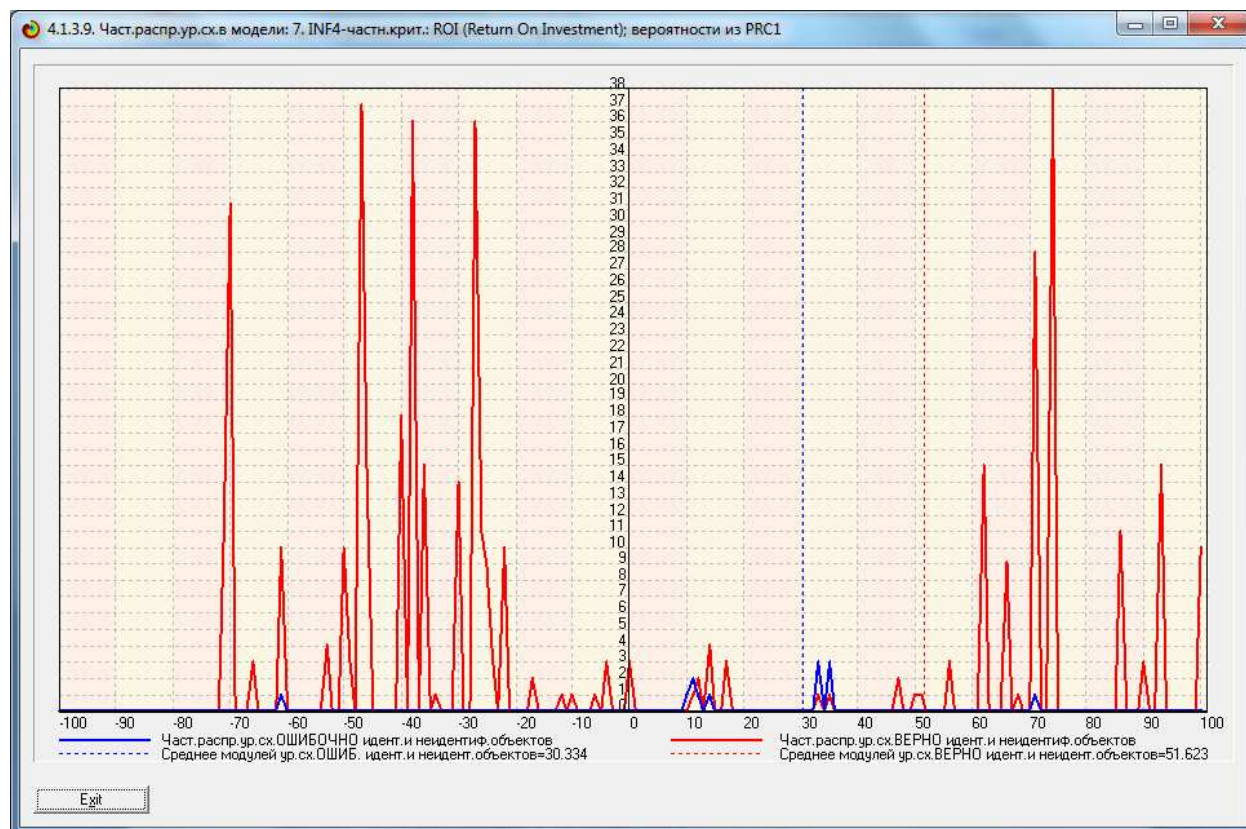


Рисунок 14. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования в модели INF4

Из рисунка 14 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 значительно 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 очень больших значений уровней сходства-различия (более 70%) также различие между верно и ошибочно идентифицированными и неидентифицированными ситуациями практически отсутствует.

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 «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 15):



Рисунок 15. Экранная форма управления созданием случайных моделей, совпадающих с текущей по размерностям основных баз данных

На рисунке 16 показано частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF4.

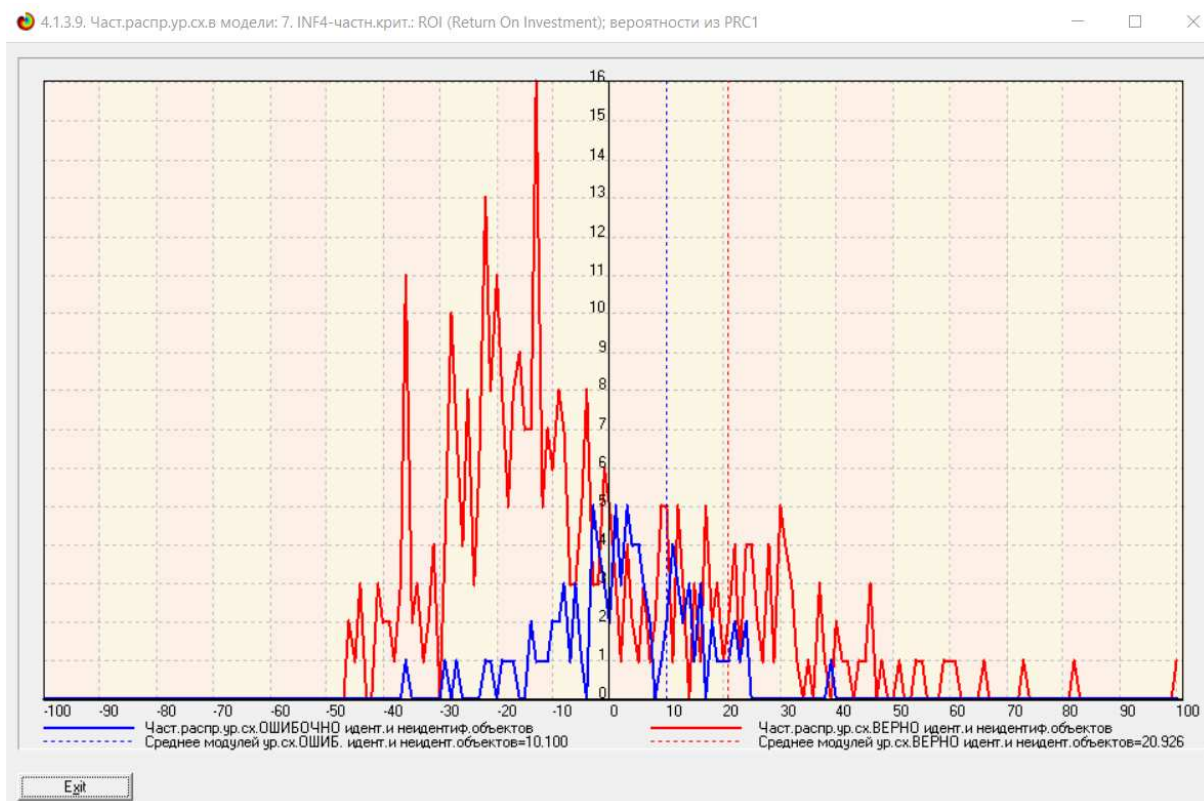


Рисунок 16. Частное распределение сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний в случайной модели INF4

Совершенно очевидное различие частотных распределений уровней сходства-различия верно и ошибочно идентифицированных и неидентифицированных состояний объекта моделирования и случайной модели (рисунки 15 и 16) 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 17.

На рисунке 17 приведены данные по достоверности статистических и когнитивных моделей, созданных на основе случайной выборки.

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

Наименование модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинно-положительных решений (TP)	Число истинно-отрицательных решений (TN)	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	F-мера Ван Ризбергера	Сумма модул. уровней сход. истинно-полож. решений (STP)	Сумма модул. уровней сход. истинно-отриц. решений (STN)	Сумма модул. уровней сход. ложно-полож. решений (SFP)
1. ABS - частный критерий: количество встреч сочетаний "класс"	Корреляция абс. частот с обр...	150	143	154	146	7	0.495	0.953	0.651	57.968	33.516	28.531
1. ABS - частный критерий: количество встреч сочетаний "класс"	Сумма абс. частот по признак...	150	150		300		0.333	1.000	0.500	88.059		111.412
2. PRC1 - частный критерий: усл. вероятность i-го признака сред.	Корреляция усл. отн. частот с о...	150	143	154	146	7	0.495	0.953	0.651	57.968	33.516	28.531
2. PRC1 - частный критерий: усл. вероятность i-го признака сред.	Сумма усл. отн. частот по при...	150	150		300		0.333	1.000	0.500	89.060		114.684
3. PRC2 - частный критерий: условная вероятность i-го признака.	Корреляция усл. отн. частот с о...	150	143	154	146	7	0.495	0.953	0.651	57.971	33.518	28.532
3. PRC2 - частный критерий: условная вероятность i-го признака.	Сумма усл. отн. частот по при...	150	150		300		0.333	1.000	0.500	89.062		114.562
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	150	116	225	75	34	0.607	0.773	0.680	38.277	67.310	10.989
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в.	Сумма знаний	150	128	190	110	22	0.538	0.853	0.660	35.126	37.220	13.812
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Семантический резонанс зна...	150	116	225	75	34	0.607	0.773	0.680	38.266	67.311	10.985
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в.	Сумма знаний	150	128	193	107	22	0.545	0.853	0.665	35.095	37.201	13.722
6. INF3 - частный критерий: "Квадрат разности между факти...	Семантический резонанс зна...	150	137	219	81	13	0.628	0.913	0.745	47.227	57.715	12.016
6. INF3 - частный критерий: "Квадрат разности между факти...	Сумма знаний	150	137	219	81	13	0.628	0.913	0.745	37.432	45.774	9.457
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	150	112	243	57	38	0.663	0.747	0.702	27.917	46.370	5.885
7. INF4 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	150	141	156	144	9	0.495	0.940	0.648	30.078	12.656	12.997
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Семантический резонанс зна...	150	112	243	57	38	0.663	0.747	0.702	27.916	46.374	5.886
8. INF5 - частный критерий: ROI (Return On Investment), вероятно...	Сумма знаний	150	141	156	144	9	0.495	0.940	0.648	30.039	12.646	12.923
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер.	Семантический резонанс зна...	150	118	224	76	32	0.608	0.787	0.686	36.716	50.195	11.033
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей, вер.	Сумма знаний	150	137	156	144	13	0.488	0.913	0.636	34.058	18.903	15.806
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Семантический резонанс зна...	150	118	224	76	32	0.608	0.787	0.686	36.712	50.208	11.033
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей, ве...	Сумма знаний	150	137	155	145	13	0.486	0.913	0.634	34.001	18.883	15.698

Рисунок 17. Достоверность статистических и когнитивных моделей, созданных на основе случайной выборки

На основе сравнения рисунка 17 с рисунком 12 можно сделать следующие выводы:

- достоверность лучшей модели INF5, отражающей реальный объект моделирования, примерно на 28% 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 18) и проведем пакетное распознавание в режиме 4.2.1 (рисунок 19).

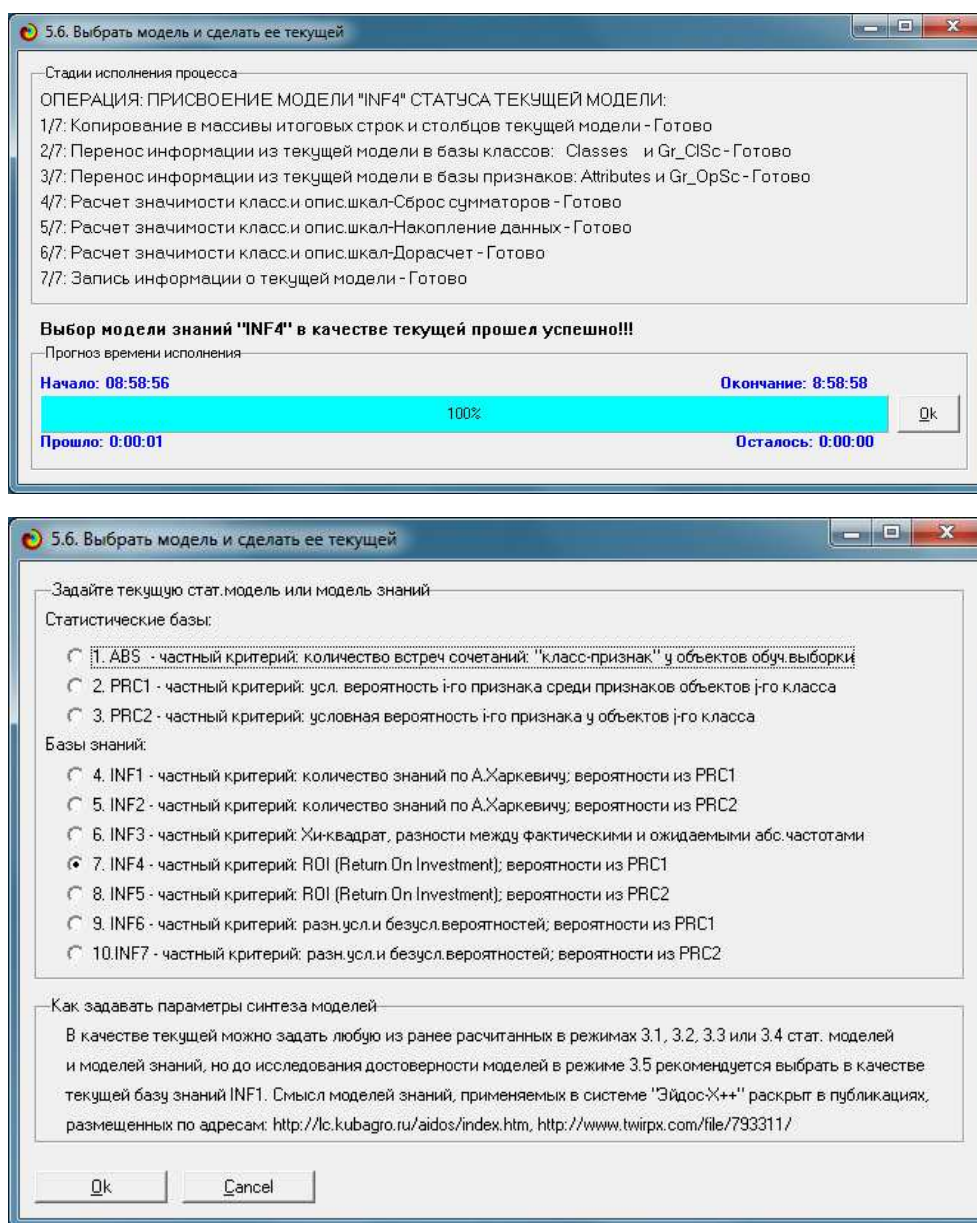


Рисунок 18. Экранные формы режима задания модели в качестве текущей

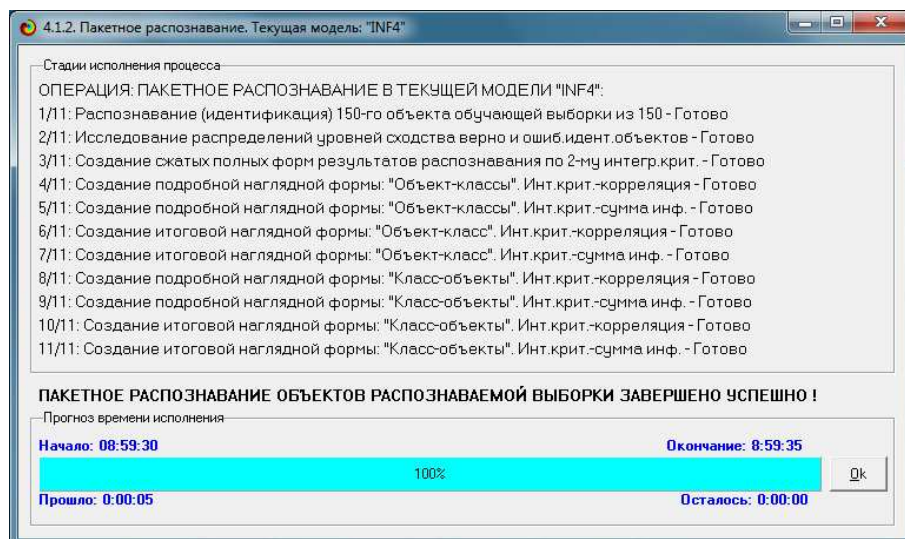


Рисунок 19. Экранная форма режима пакетного распознавания в текущей модели

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 20 и 21 приведены примеры прогнозов высокой и низкой достоверности частоты и классов Ирисов в наиболее достоверной модели INF4 на основе наблюдения предыстории их развития:

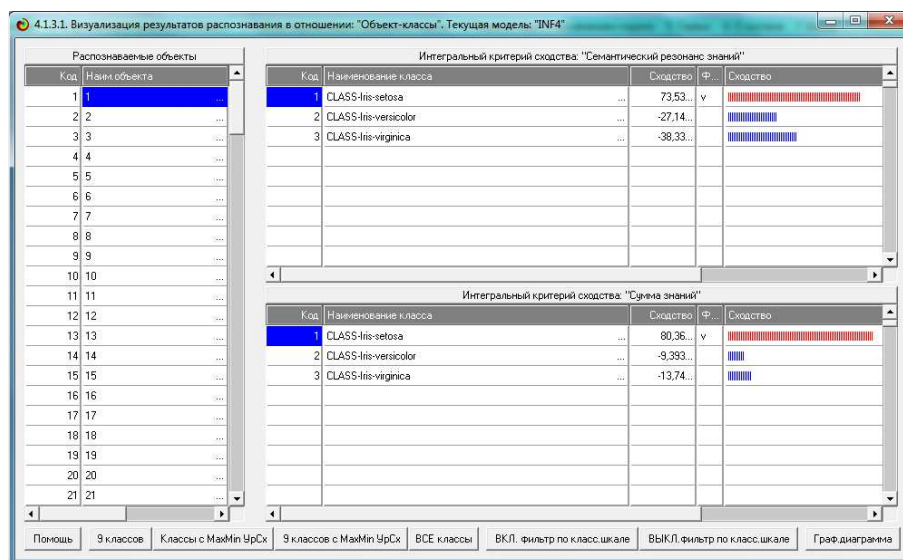


Рисунок 20. Пример идентификации классов ириса в модели INF4

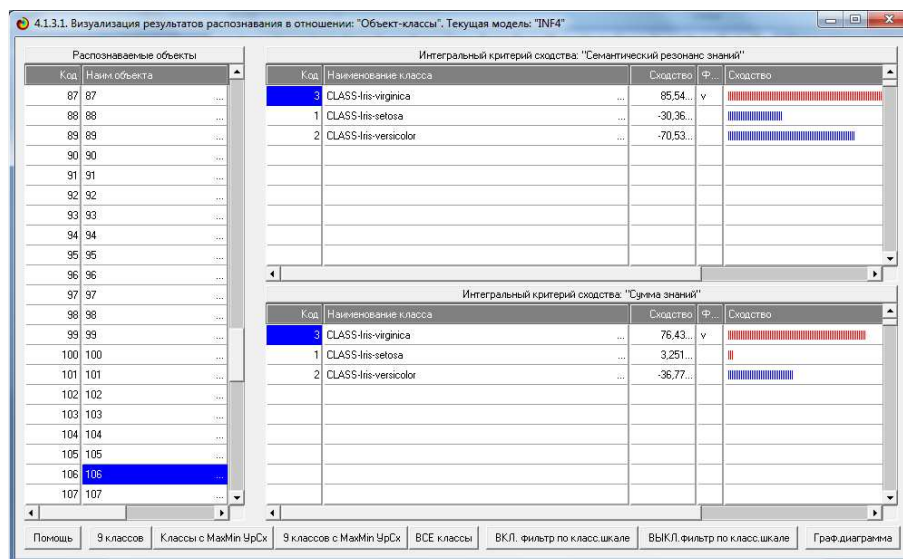


Рисунок 21. Пример идентификации классов ириса в модели INF4

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 (рисунок 22)

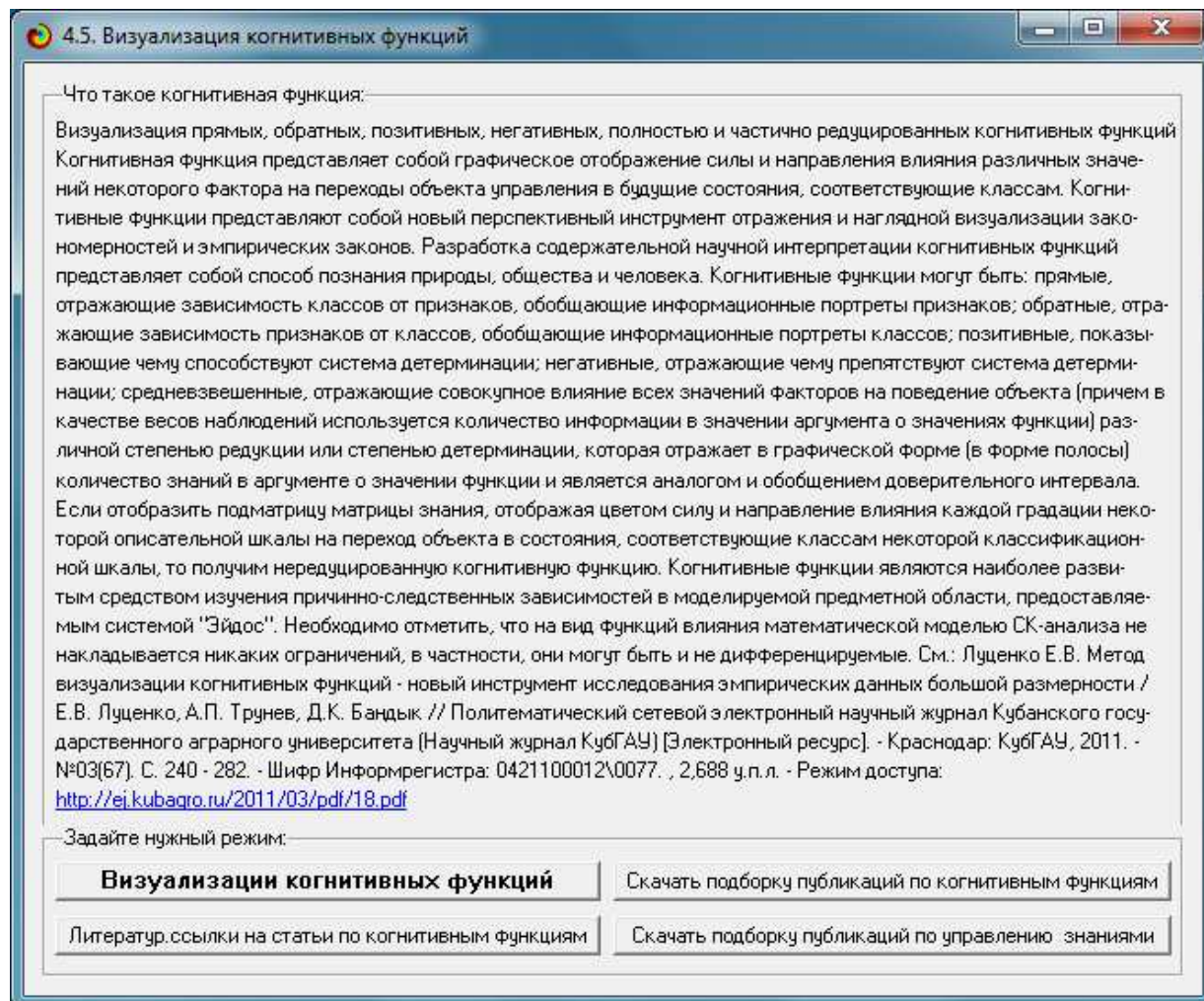
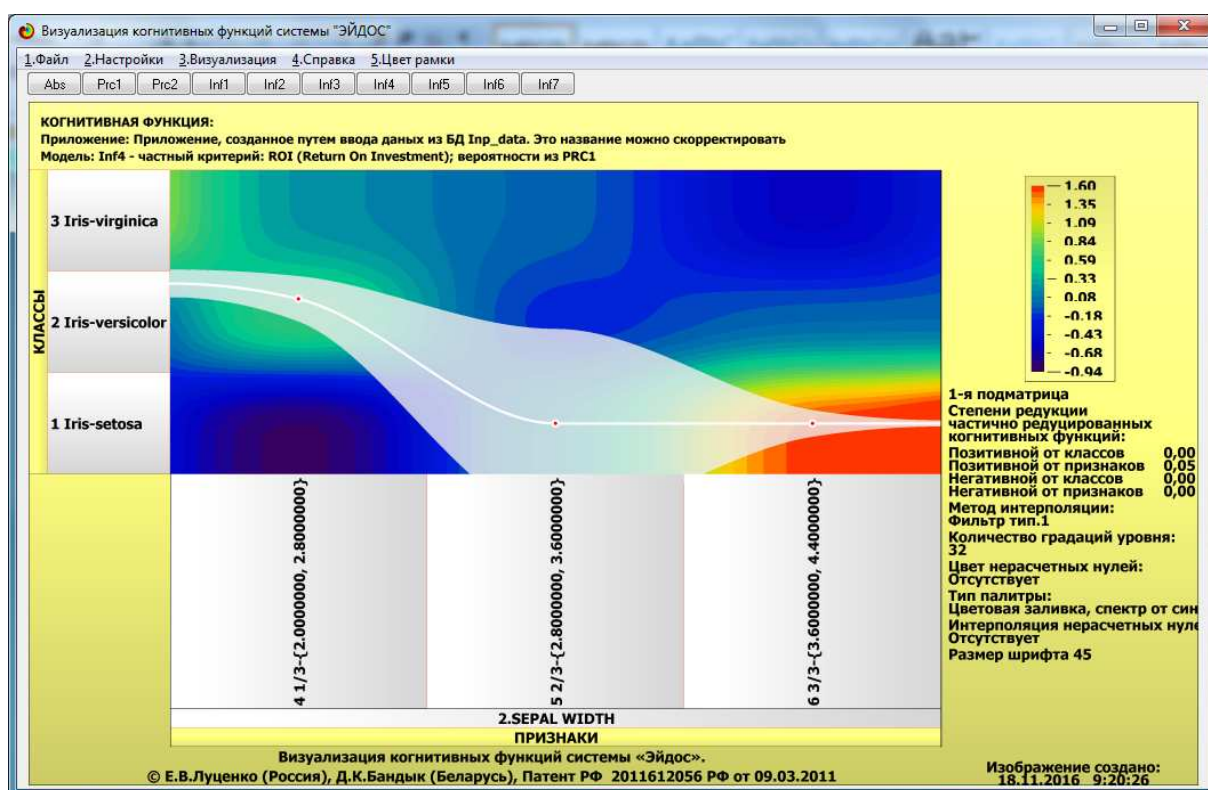
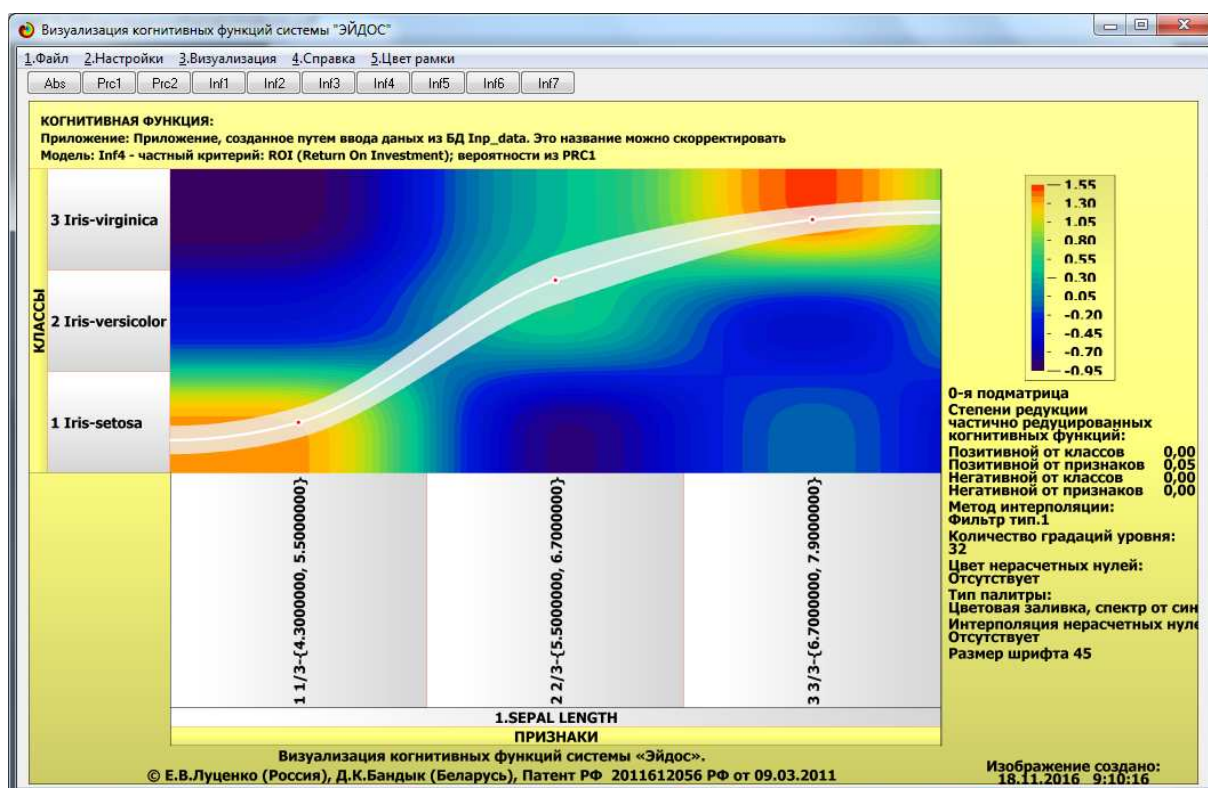


Рисунок 22. Экранная форма режима 4.5 системы «Эйдос-X++»
«Визуализация когнитивных функций»

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 много работ автора, но наиболее новой и обобщающей из них является работа [11]. Поэтому здесь не будем

останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе. На рисунке 23 приведены визуализации всех когнитивных функций данного приложения для модели INF3.



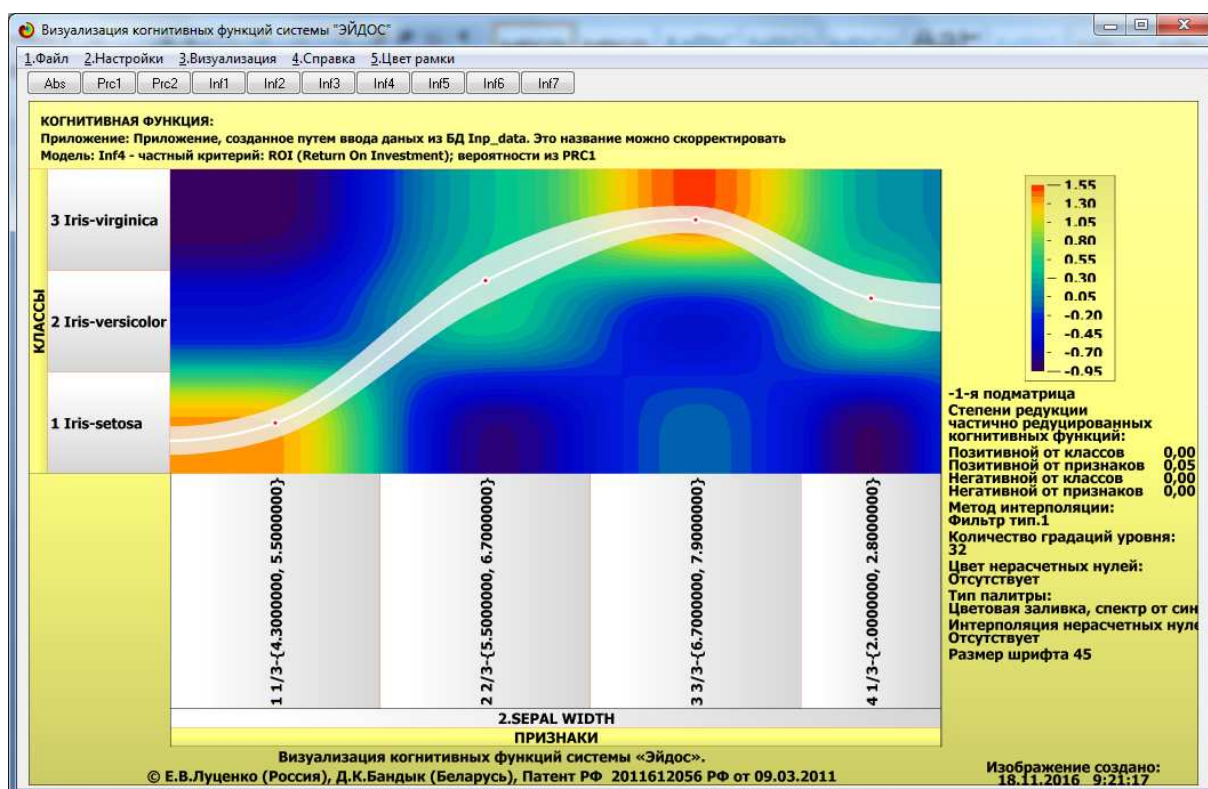
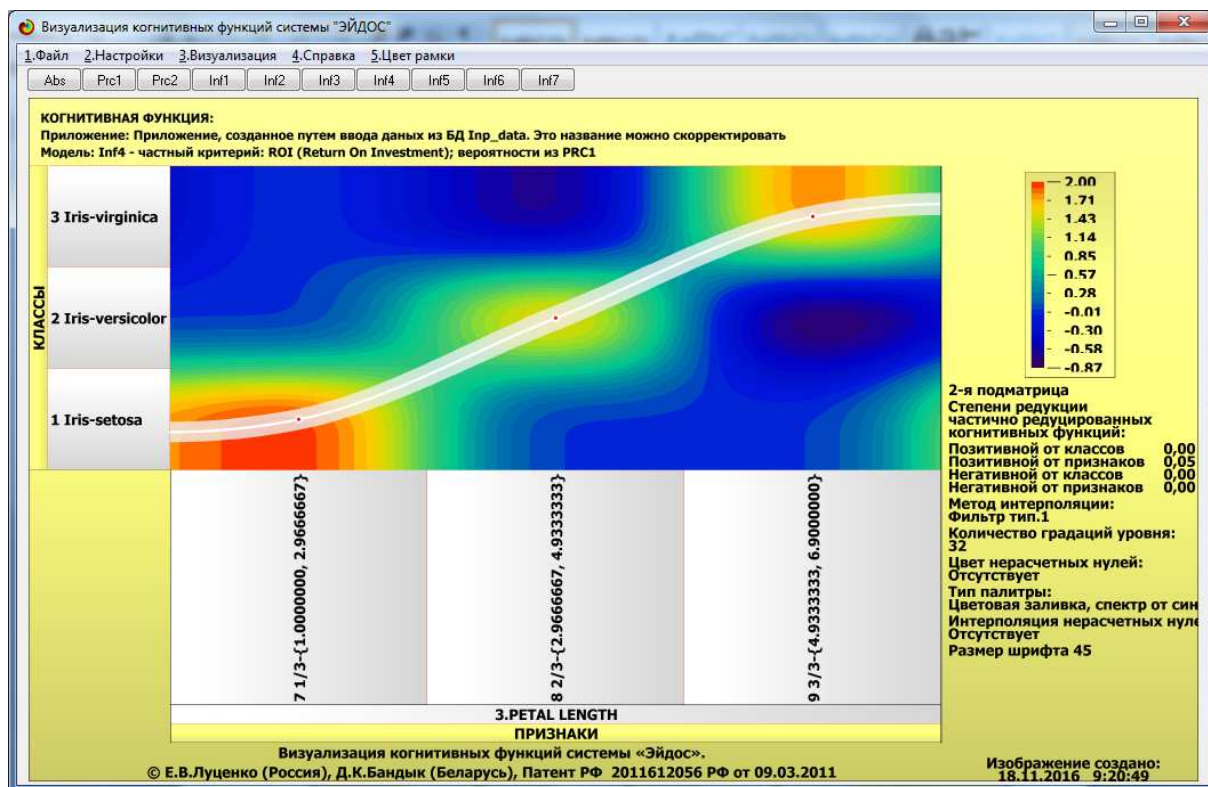


Рисунок 23. Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF3

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 24).

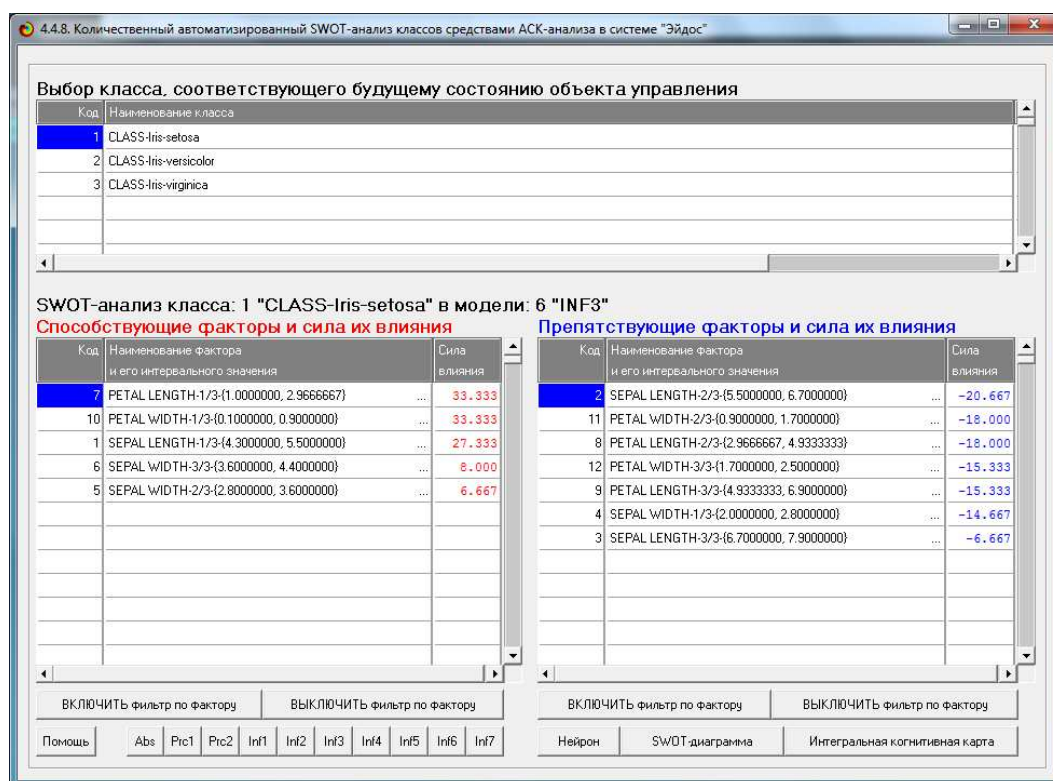


Рисунок 24. Пример SWOT-матрицы в модели INF3

На рисунке 25 приведены примеры инвертированной SWOT-матрицы и инвертированной SWOT-диаграммы в модели INF3.

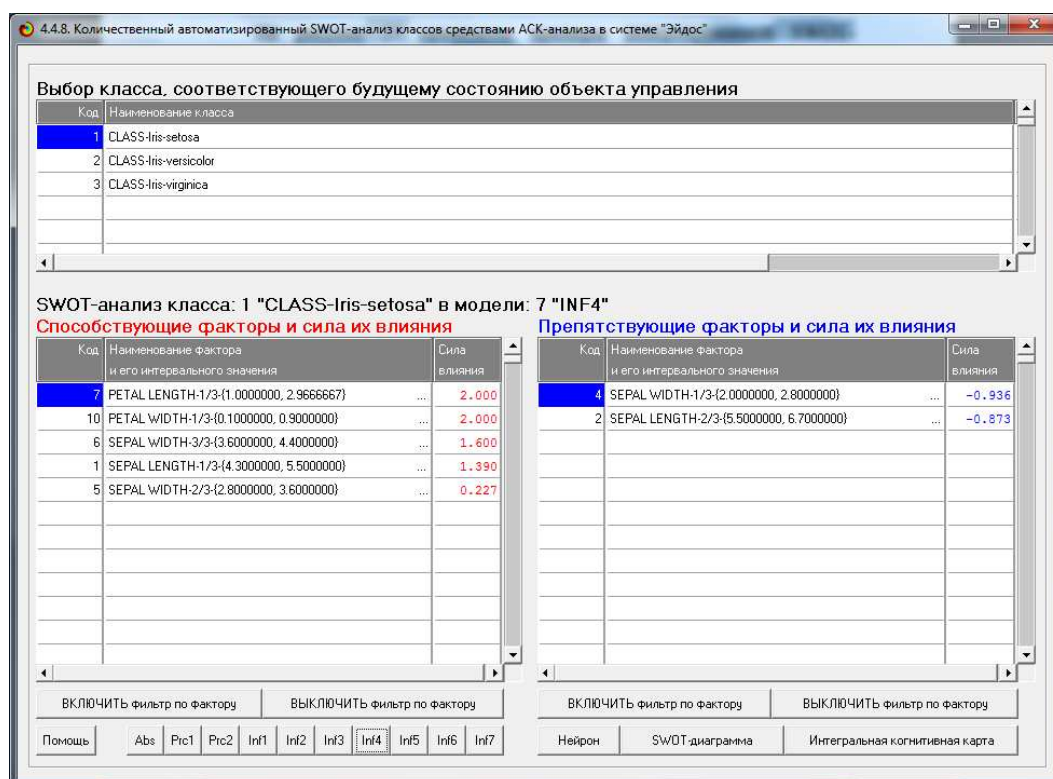


Рисунок 25. Пример SWOT-матрицы в модели INF4

ЗАКЛЮЧЕНИЕ

Так как существует множество систем искусственного интеллекта, то возникает необходимость сопоставимой оценки качества их математических моделей. Одним из вариантов решения этой задачи является тестирование различных системы на общей базе исходных данных, для чего очень удобно использовать общедоступную базу репозитория UCI. В данной 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 INF4 и INF5 при интегральном критерии «Резонанс знаний». Точность модели составляет 0,968, что заметно выше, чем достоверность экспертных оценок, которая считается равной около 70%. Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется F-мера Ван Ризбергена и ее нечеткое мультиклассовое обобщение L-мера, предложенная проф.Е.В.Луценко [10].

На основе базы данных UCI, рассмотренной в данной работе, построить модели прогнозирования не с помощью АСК-анализа и реализующей его системы «Эйдос», а с применением других математических методов и реализующих их программных систем, то можно сопоставимо сравнить их качество.

ЛИТЕРАТУРА

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 [article ID]: 0020302012. – Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 у.п.л.

2. 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.

3. 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.

4. 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. Репозиторий UCI [Электронный ресурс]. Статья "Iris Data Set": <http://archive.ics.uci.edu/ml/datasets/Iris>, свободный. - Загл. с экрана. Яз. англ.

6. Сайт профессора Е.В.Луценко [Электронный ресурс]. Режим доступа: <http://lc.kubagro.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>

10. 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]: 1221608001.
<http://dx.doi.org/10.21515/1990-4665-123-001>

11. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с. ISBN 978-5-94672-757-0. <http://elibrary.ru/item.asp?id=21358220>

LITERATURA

1. Lucenko E.V. Metodika ispol'zovaniya repozitorija UCI dlja ocenki kachestva matematicheskikh modelej sistem iskusstvennogo intellekta / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: http://lc.kubagro.ru/My_training_schedule.doc

KubGAU, 2003. – №02(002). S. 120 – 145. – IDA [article ID]: 0020302012. – Rezhim dostupa: <http://ej.kubagro.ru/2003/02/pdf/12.pdf>, 1,625 u.p.l.

2. Lucenko E.V. ASK-analiz, modelirovanie i identifikacija zhivykh sushhestv na osnove ih fenotipicheskikh priznakov / E.V. Lucenko, Ju.N. Penkina // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: Kub- GAU, 2014. – №06(100). S. 1346 – 1395. – IDA [article ID]: 1001406090. – Rezhim dostupa: <http://ej.kubagro.ru/2014/06/pdf/90.pdf>, 3,125 u.p.l.

3. Lucenko E.V. Teoreticheskie osnovy, tehnologija i instrumentarij avtomatizirovannogo sistemno-kognitivnogo analiza i vozmozhnosti ego primeneniya dlja sopostavimoy ocenki jeffektivnosti vuzov / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: Kub- GAU, 2013. – №04(088). S. 340 – 359. – IDA [article ID]: 0881304022. – Rezhim dostupa: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 u.p.l.

4. Lucenko E.V. Avtomatizirovannyj sistemno-kognitivnyj analiz v upravlenii aktivnymi ob#ektami (sistemnaja teorija informacii i ee primeneniye v issledovanii jekonomicheskikh, social'no-psihologicheskikh, tehnologicheskikh i organizacionno-tehnicheskikh sistem): Monografija (nauchnoe izdanie). – Krasnodar: KubGAU. 2002. – 605 s.

5. Repozitorij UCI [Jelektronnyj resurs]. Stat'ja "Iris Data Set": <http://archive.ics.uci.edu/ml/datasets/Iris>, svobodnyj. - Zagl. s jekrana. Jaz. ang.

6. Sajt professora E.V.Lucenko [Jelektronnyj resurs]. Rezhim dostupa: <http://lc.kubagro.ru/>, svobodnyj. - Zagl. s jekrana. Jaz. rus.

7. Lucenko E.V. 30 let sisteme «Jejdos» – odnoj iz starejsih otechestvennyh universal'nyh sistem iskusstvennogo intellekta, shiroko primenjaemyh i razvivajushhihsja i v nastojashhee vremja / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2009. – №10(054). S. 48 – 77. – Shifr Informregistra: 0420900012\0110, IDA [article ID]: 0540910004. – Rezhim dostupa: <http://ej.kubagro.ru/2009/10/pdf/04.pdf>, 1,875 u.p.l.

8. Lucenko E.V. Universal'naja kognitivnaja analiticheskaja sistema «Jejdos- H++» / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2012. – №09(083). S. 328 – 356. – IDA [article ID]: 0831209025. – Rezhim dostupa: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 u.p.l.

9. Lucenko E.V., Borovko A.Ju. Prognozirovaniye kolichestva i klassov solnechnyh vspyshek na osnove ih predystorii po dannym repozitorija UCI s primeneniem ASK-analiza i intellektual'noj sistemy «Jejdos» // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal

Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – IDA [article ID]: 0831209025. – Rezhim dostupa: <http://ej.kubagro.ru/2014/10/pdf/99.pdf>

10. Lucenko E.V. Nechetkoe mul'tiklassovoe obobshhenie klassicheskoy F-mery dostovernosti modelej Van Rizbergena v ASK-analize i sisteme «Jejdos» / Lucenko E.V. // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2016. – №09(123). – Rezhim dostupa: <http://ej.kubagro.ru/2016/09/pdf/01.pdf>, 1,813 u.p.l. – IDA [article ID]: 1221608001. <http://dx.doi.org/10.21515/1990-4665-123-001>

11. Orlov A.I., Lucenko E.V. Sistemnaja nechetkaja interval'naja matematika. Monografija (nauchnoe izdanie). – Krasnodar, KubGAU. 2014. – 600 s. ISBN 978-5-94672-757-0. <http://elibrary.ru/item.asp?id=21358220>