

УДК 303.732.4

UDC 303.732.4

**ПРОГНОЗИРОВАНИЕ
ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ
ПАЦИЕНТОВ, ПЕРЕНЕСШИХ СЕРДЕЧНЫЙ
ПРИСТУП, ПО ДАННЫМ
ЭХОКАРДИОГРАММЫ НА ОСНОВЕ БАЗЫ
ДАННЫХ РЕПОЗИТОРИЯ UCI¹**

**PREDICTION OF SURVIVAL IN PATIENTS
WHO HAVE HAD A HEART ATTACK,
ACCORDING TO ECHO-CARDIOGRAPHIC
DATABASE OF UCI REPOSITORY**

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Кубанский государственный аграрный университет, Россия, 350044, Краснодар, Калинина, 13,
prof.lutsenko@gmail.com

Lutsenko Eugeny Veniaminovich
Dr.Sci.Econ., Cand.Tech.Sci., professor
Kuban State Agrarian University, Krasnodar, Russia

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».

Since there are many artificial intelligence systems, there is a need of comparable quality assessment of their mathematical models. For this purpose, these systems can be tested on the same database source data, for which it is very convenient to use a public database of the UCI repository. This article provides a detailed example of using the UCI database repository to assess the quality of the mathematical models used in the ASC-analysis and its software toolkit system of artificial intelligence called "Aidos"

Ключевые слова: АСК-АНАЛИЗ, МОДЕЛИ, ИДЕНТИФИКАЦИЯ, ЖИВЫЕ СУЩЕСТВА, ФЕНОТИПИЧЕСКИЕ ПРИЗНАКИ

Keywords: ASC-ANALYSIS, MODELS, IDENTIFICATION, LIVING CREATURES, PHENOTYPIC TRAITS

СОДЕРЖАНИЕ

1. КРАТКАЯ ТЕОРИЯ	2
1.1. ПРОБЛЕМАТИКА.....	2
1.2. ИСХОДНЫЕ ДАННЫЕ.....	2
1.3. ИНСТРУМЕНТАРИЙ.....	3
1.4. МЕТРИЗАЦИЯ ШКАЛ.....	3
2. СИНТЕЗ И ВЕРИФИКАЦИЯ МОДЕЛЕЙ	4
2.1. ОПИСАНИЕ РЕШЕНИЯ.....	4
2.2. ПРЕОБРАЗОВАНИЕ ИСХОДНЫХ ДАННЫХ ИЗ HTML-ФОРМАТА В ФАЙЛ ИСХОДНЫХ ДАННЫХ MS EXCEL.....	4
2.3. ЭТАПЫ АСК-АНАЛИЗА И ПРЕОБРАЗОВАНИЕ ИСХОДНЫХ ДАННЫХ В ИНФОРМАЦИЮ, А ЕЕ В ЗНАНИЯ В СИСТЕМЕ "ЭЙДОС".....	12
2.4. СКАЧИВАНИЕ И ИНСТАЛЛЯЦИЯ СИСТЕМЫ «ЭЙДОС».....	15
2.5. АВТОМАТИЗИРОВАННАЯ ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ ПУТЕМ ИМПОРТА ИСХОДНЫХ ДАННЫХ ИЗ ВНЕШНИХ БАЗ ДАННЫХ В СИСТЕМУ "ЭЙДОС".....	17
2.6. СИНТЕЗ И ВЕРИФИКАЦИЯ СТАТИСТИЧЕСКИХ И ИНТЕЛЛЕКТУАЛЬНЫХ МОДЕЛЕЙ.....	20
2.7. ЧАСТНЫЕ КРИТЕРИИ И ВИДЫ МОДЕЛЕЙ СИСТЕМЫ «ЭЙДОС».....	22
2.8. ЦЕННОСТЬ ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ИДЕНТИФИКАЦИИ И ПРОГНОЗИРОВАНИЯ.....	26
2.9. ИНТЕГРАЛЬНЫЕ КРИТЕРИИ СИСТЕМЫ «ЭЙДОС».....	26
2.10. РЕЗУЛЬТАТЫ ВЕРИФИКАЦИИ МОДЕЛЕЙ.....	28
3. РЕШЕНИЕ ЗАДАЧ ИДЕНТИФИКАЦИИ И ПРОГНОЗИРОВАНИЯ В НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	29
4. ИССЛЕДОВАНИЕ МОДЕЛИРУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ ПУТЕМ ИССЛЕДОВАНИЯ ЕЕ МОДЕЛИ	32
4.1. ИНФОРМАЦИОННЫЕ ПОРТРЕТЫ КЛАССОВ И ПРИЗНАКОВ.....	32
4.2. КОГНИТИВНЫЕ ФУНКЦИИ.....	35
4.3. SWOT И PEST МАТРИЦЫ И ДИАГРАММЫ.....	40
4.4. НЕЛОКАЛЬНЫЕ НЕЙРОНЫ.....	42
ЗАКЛЮЧЕНИЕ	43
ЛИТЕРАТУРА	44

¹ *Материал подготовлен по результатам исследований, проведенных при финансовой поддержке РГНФ, проект №13-02-00440а*

1. Краткая теория

1.1. Проблематика

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 работах [1, 2] рассматривалось решение этой задачи для оценки качества модели, реализованной в универсальной автоматизированной системе «Эйдос», представляющей собой программный инструментарий автоматизированного системно-когнитивного анализа (АСК-анализ) [3, 4]. Данная статья является продолжением этого подхода.

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.

Для решения этой задачи необходимы тестовые задачи. Удачным выбором является база данных тестовых задач для систем искусственного интеллекта репозитория UCI: <http://archive.ics.uci.edu/ml/datasets.html>. «Репозиторий UCI (UCI Machine Learning Repository) — крупнейший репозиторий реальных и модельных задач машинного обучения. Содержит реальные данные по прикладным задачам в области биологии, медицины, физики, техники, социологии, и др. Задачи (наборы данных, data set) именно этого репозитория чаще всего используются научным сообществом для эмпирического анализа алгоритмов машинного обучения. Репозиторий UCI создан в университете г.Ирвин (Калифорния, США).»²

1.2. Исходные данные

В данной работе использована база данных "Echocardiogram Data Set": <http://archive.ics.uci.edu/ml/datasets/Echocardiogram> [5] из банка исходных данных по задачам искусственного интеллекта - репозитория UCI, который выполнен сотрудниками Школы информатики и компьютерных исследований Калифорнийского университета США. База данных дополнена значениями и приводится полностью в приложении.

На момент написания работы этот банк включает в себя 290 баз исходных данных по различным задачам.

² См., например: <http://www.machinelearning.ru/wiki/index.php?title=UCI>

1.3. Инструментарий

Для решения задачи используем стандартные возможности Microsoft Office Word и Excel. А так же систему искусственного интеллекта "Эйдос-Х++".

Математическая модель системы "Эйдос" была разработана в 1979 и впервые прошла экспериментальную апробацию в 1981 году. С 1981 по 1992 система "Эйдос" неоднократно реализовалась на платформе Wang (на компьютерах Wang-2200C). Для IBM-совместимых персональных компьютеров система "Эйдос" впервые реализована в 1992 году и с тех пор совершенствуется постоянно, вплоть до настоящего времени [6, 7, 8].

1.4. Метризация шкал

Измерительные шкалы - это инструмент создания формальных моделей реальных объектов, а так же и инструмент, который позволяет повысить степень формализации этих моделей до такого уровня, при котором возможна их реализация на компьютерах [9].

Чем выше степень формализации модели, тем точнее решаются различные задачи с использованием этих моделей, тем проще использовать модели для проектирования искусственной систем и их создания, а также тем более точные и развитые математические методы могут применяться в этих моделях.

Однако, если разные факторы измеряются в различных единицах измерения, то результат сравнения объектов будет зависеть от единиц измерения факторов, что недопустимо. Поэтому возникает проблема сопоставимости результатов измерений разнородных величин (измеряемых в разных единицах измерения) в шкалах различных типов (номинальных, порядковых и числовых).

В АСК-анализе и его инструментарии интеллектуальной системе "Эйдос" эта проблема решается путем метризации шкал [9]. Система «Эйдос» в настоящее время поддерживает 7 способов метризации всех типов шкал, которые обеспечивают корректную совместную сопоставимую количественную обработку разнородных по своей природе факторов, различных единиц измерения. Факторы описываются шкалами, а значения факторов – градациями шкал. В АСК-анализе все они рассматриваются с единой точки зрения: сколько информации содержится в их значениях о переходе объекта, на который они действуют, в определенное состояние, при этом направление влияния и сила значений факторов измеряется в одних общих для факторов единицах измерения (единицах количества информации). По этой причине можно складывать силу и направление влияния всех действующих на объект значений факторов, и определять результат их совместного влияния на объект. При этом объект является нелиней-

ным, а факторы взаимодействуют друг с другом внутри него (не выполняется принцип суперпозиции).

Измерение – Отображение реальных объектов в формальных шкалах. Система "Эйдос" представляет – средство для построения и применения измерительных инструментов в различных предметных областях. В ней реализованы разнообразные технологии метризации, которые позволяют любые свойства объектов (количественные и качественные) исследовать в наиболее сильных абсолютных шкалах знаний.

2. Синтез и верификация моделей

2.1. 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.

2.2. 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 "Echocardiogram", которую оставим без изменений [5].

Общее описание задачи:

(файл: [echocardiogram.names](http://archive.ics.uci.edu/ml/machine-learning-databases/echocardiogram/echocardiogram.names) по адресу: <http://archive.ics.uci.edu/ml/machine-learning-databases/echocardiogram/echocardiogram.names>)

1. Title: Echocardiogram Data
2. Source Information:
 - Donor: Steven Salzberg (salzberg@cs.jhu.edu)
 - Collector:
 - Dr. Evlin Kinney
 - The Reed Institute
 - P.O. Box 402603
 - Maimi, FL 33140-0603
 - Date Received: 28 February 1989
3. Past Usage:
 - 1. Salzberg, S. (1988). Exemplar-based learning: Theory and implementation (Technical Report TR-10-88). Harvard University,

Center for Research in Computing Technology, Aiken Computation Laboratory (33 Oxford Street; Cambridge, MA 02138).

- Steve applied his EACH program to predict survival (i.e., life or death), did not use the wall-motion attribute, and recorded 87 correct and 29 incorrect in an incremental application to this database. He also showed that, by tuning EACH to this domain, EACH was able to derive (non-incrementally) a set of 28 hyper-rectangles that could perfectly classify 119 instances.
- 2. Kan, G., Visser, C., Kooler, J., & Dunning, A. (1986). Short and long term predictive value of wall motion score in acute myocardial infarction. British Heart Journal, 56, 422-427.
- They predicted the same variable (whether patients will live one year after a heart attack) using a different set of 345 instances. Their statistical test recorded a 61% accuracy in predicting that a patient will die (post-hoc fit).
- 3. Elvin Kinney (in communication with Steven Salzberg) reported that a Cox regression application recorded a 60% accuracy in predicting that a patient will die.

4. Relevant Information:

- All the patients suffered heart attacks at some point in the past. Some are still alive and some are not. The survival and still-alive variables, when taken together, indicate whether a patient survived for at least one year following the heart attack.

The problem addressed by past researchers was to predict from the other variables whether or not the patient will survive at least one year. The most difficult part of this problem is correctly predicting that the patient will NOT survive. (Part of the difficulty seems to be the size of the data set.)

5. Number of Instances: 132

6. Number of Attributes: 13 (all numeric-valued)

7. Attribute Information:

1. survival -- the number of months patient survived (has survived, if patient is still alive). Because all the patients had their heart attacks at different times, it is possible that some patients have survived less than one year but they are still alive. Check the second variable to confirm this. Such patients cannot be used for the prediction task mentioned above.
2. still-alive -- a binary variable. 0=dead at end of survival period, 1 means still alive
3. age-at-heart-attack -- age in years when heart attack occurred
4. pericardial-effusion -- binary. Pericardial effusion is fluid around the heart. 0=no fluid, 1=fluid
5. fractional-shortening -- a measure of contractility around the heart
lower numbers are increasingly abnormal
6. epss -- E-point septal separation, another measure of contractility.
Larger numbers are increasingly abnormal.
7. lvdd -- left ventricular end-diastolic dimension. This is a measure of the size of the heart at end-diastole.
Large hearts tend to be sick hearts.
8. wall-motion-score -- a measure of how the segments of the left ventricle are moving
9. wall-motion-index -- equals wall-motion-score divided by number of segments seen. Usually 12-13 segments are seen in an echocardiogram. Use this variable INSTEAD of the wall motion score.

- 10. mult -- a derivate var which can be ignored
- 11. name -- the name of the patient (I have replaced them with "name")
- 12. group -- meaningless, ignore it
- 13. alive-at-1 -- Boolean-valued. Derived from the first two attributes.
0 means patient was either dead after 1 year or had been followed for less than 1 year. 1 means patient was alive at 1 year.

8. Missing Attribute Values: (denoted by "?")

Attribute #:	Number of Missing Values: (total: 132)
1	2
2	1
3	5
4	1
5	8
6	15
7	11
8	4
9	1
10	4
11	0
12	22
13	58

9. Distribution of attribute number 2: still-alive

Value	Number of instances with this value
0	88 (dead)
1	43 (alive)
?	1
Total	132

10. Distribution of attribute number 13: alive-at-1

Value	Number of instances with this value
0	50
1	24
?	58
Total	132

Обучающая выборка

(файл: «[echocardiogram.data](http://archive.ics.uci.edu/ml/machine-learning-databases/echocardiogram/)» по адресу: <http://archive.ics.uci.edu/ml/machine-learning-databases/echocardiogram/>):

```
11,0,71,0,0.260,9,4.600,14,1,1,name,1,0
19,0,72,0,0.380,6,4.100,14,1.700,0.588,name,1,0
16,0,55,0,0.260,4,3.420,14,1,1,name,1,0
57,0,60,0,0.253,12.062,4.603,16,1.450,0.788,name,1,0
19,1,57,0,0.160,22,5.750,18,2.250,0.571,name,1,0
26,0,68,0,0.260,5,4.310,12,1,0.857,name,1,0
13,0,62,0,0.230,31,5.430,22.5,1.875,0.857,name,1,0
50,0,60,0,0.330,8,5.250,14,1,1,name,1,0
19,0,46,0,0.340,0,5.090,16,1.140,1.003,name,1,0
25,0,54,0,0.140,13,4.490,15.5,1.190,0.930,name,1,0
10,1,77,0,0.130,16,4.230,18,1.800,0.714,name,1,1
52,0,62,1,0.450,9,3.600,16,1.140,1.003,name,1,0
52,0,73,0,0.330,6,4,14,1,1,name,1,0
44,0,60,0,0.150,10,3.730,14,1,1,name,1,0
0.5,1,62,0,0.120,23,5.800,11.67,2.330,0.358,name,1,1
```

24,0,55,1,0.250,12.063,4.290,14,1,1,name,1,0
 0.5,1,69,1,0.260,11,4.650,18,1.640,0.784,name,1,1
 0.5,1,62.529,1,0.070,20,5.200,24,2,0.857,name,1,1
 22,1,66,0,0.090,17,5.819,8,1.333,0.429,name,1,0
 1,1,66,1,0.220,15,5.400,27,2.250,0.857,name,1,1
 0.75,1,69,0,0.150,12,5.390,19.5,1.625,0.857,name,1,1
 0.75,1,85,1,0.180,19,5.460,13.83,1.380,0.710,name,1,1
 0.5,1,73,0,0.230,12.733,6.060,7.5,1.500,0.360,name,1,1
 5,1,71,0,0.170,0,4.650,8,1,0.570,name,1,1
 48,0,64,0,0.190,5.9,3.480,10,1.110,0.640,name,2,?
 29,0,54,0,0.300,7,3.850,10,1.667,0.430,name,2,?
 29,0,35,0,0.300,5,4.170,14,1,1,name,2,?
 29,0,55,0,?,7,?,2,1,2,name,2,?
 0.25,1,75,0,?,?,?,?,1,?,name,2,?
 36,0,55,1,0.210,4.2,4.160,14,1.560,0.640,name,2,0
 1,1,65,0,0.150,?,5.050,10,1,0.710,name,2,?
 1,1,52,1,0.170,17.200,5.320,14,1.170,0.857,name,2,?
 03,1,?,0,?,12,?,6,3,0.140,name,2,?
 27,0,47,0,0.400,5.120,3.100,12,1,0.857,name,2,?
 35,0,63,0,?,10,?,14,1.170,0.857,name,2,0
 26,0,61,0,0.610,13.100,4.070,13,1.625,0.571,name,2,0
 16,0,63,1,?,?,5.310,5,1,0.357,name,2,0
 1,1,65,0,0.060,23.600,?,21.500,2.150,0.714,name,2,1
 19,0,68,0,0.510,?,3.880,15,1.670,0.640,name,2,0
 31,0,80,0,0.410,5.400,4.360,?,1,?,name,2,?
 32,0,54,0,0.350,9.300,3.630,11,1.222,0.640,name,2,0
 16,0,70,1,0.270,4.700,4.490,22,2,0.786,name,2,0
 40,0,79,0,0.150,17.500,4.270,13,1.300,0.714,name,2,0
 46,0,56,0,0.330,?,3.590,14,1,1,name,2,0
 2,1,67,1,0.440,9,3.960,17.500,1.450,0.857,name,2,?
 37,0,64,0,0.090,?,?,12,2,0.428,name,2,?
 19.5,1,81,0,0.120,?,?,9,1.250,0.570,name,2,0
 20,1,59,0,0.030,21.300,6.290,17,1.310,0.928,name,2,0
 0.25,1,63,1,?,?,?,23,2.300,0.714,name,2,1
 ,?,?,77,?,?,?,?,2,?,name,2,?
 2,1,56,1,0.040,14,5,?,?,?,name,2,1
 7,1,61,1,0.270,?,?,9,1.500,0.428,name,2,1
 10,0,57,0,0.240,14.800,5.260,18,1.380,0.812,name,2,?
 12,0,58,0,0.300,9.400,3.490,14,1,1,name,2,0
 1,1,60,0,0.010,24.600,5.650,39,3,0.928,name,2,1
 10,0,66,0,0.290,15.600,6.150,14,1,1,name,2,0
 45,0,63,0,0.15,13,4.57,13,1.08,0.857,name,2,0
 22,0,57,0,0.13,18.6,4.37,12.33,1.37,0.642,name,2,0
 53,0,70,0,0.10,9.8,5.30,23,2.30,0.714,name,2,0
 38,0,68,0,0.29,?,4.41,14,1.167,0.857,name,2,?
 26,0,79,0,0.17,11.9,5.15,10.5,1.05,0.714,name,2,0
 9,0,73,0,0.12,?,6.78,16.67,1.39,0.857,name,2,?
 26,0,72,0,0.187,12,5.02,13,1.18,0.785,name,2,0
 0.5,1,59,0,0.13,16.4,4.96,17.83,1.37,0.928,name,2,?
 12,0,67,1,0.11,10.3,4.68,11,1,0.785,name,2,?
 49,0,51,0,0.16,13.2,5.26,11,1,0.786,name,2,0
 0.75,1,50,0,0.14,11.4,4.75,10,2.5,0.28,name,2,?
 49,0,70,1,0.25,9.7,5.57,5.5,1.10,0.357,name,2,0
 47,0,65,0,0.36,8.8,5.78,12,1,0.857,name,2,0
 41,0,78,0,0.06,16.1,5.62,13.67,1.367,0.714,name,2,0
 .25,1,86,0,0.225,12.2,5.20,24,2.18,0.786,name,2,1
 33,0,56,0,0.25,11,4.72,11,1,0.785,name,2,0
 29,0,60,0,0.12,10.2,4.31,15,1.67,0.64,name,2,0
 41,0,59,0,0.29,7.5,4.75,13,1.08,0.857,name,2,0
 26,0,50,0,0.06,30.1,5.95,21.5,2.39,0.643,name,2,?
 15,0,54,0,0.217,17.9,4.54,16.5,1.18,1,name,2,0

.25,1,68,0,0.22,21.7,4.85,15,1.15,0.928,name,2,?
 .03,1,?,0,0.26,19.4,4.77,21,2.1,0.714,name,2,1
 12,0,64,0,0.20,7.1,4.58,14,1,1,name,2,0
 32,0,63,0,0.20,5,5.20,8,1,0.57,name,2,?
 32,0,65,0,0.06,23.6,6.74,12,1.09,0.785,name,2,?
 27,0,54,1,0.07,16.8,4.16,18,1.5,0.857,name,2,0
 23,0,62,0,0.25,6,4.48,11,1,0.786,name,2,?
 0.75,1,78,0,0.05,10,4.44,15,1.36,0.786,name,2,1
 0.75,1,61,0,?,?,?,28,2.33,0.857,name,2,1
 34,0,52,0,0.14,25,6.21,11.5,1.15,0.714,name,2,?
 1,1,73,0,0.05,14.8,4.14,15.5,1.41,0.786,name,2,?
 21,1,70,1,0.16,19.2,5.25,11,1,0.786,name,2,?
 55,0,55,0,0.28,5.5,4.48,22,1.83,0.857,name,2,0
 15,1,60,0,0.18,8.7,4.56,13.5,1.04,0.928,name,2,?
 0.5,1,67,0,0.155,11.3,5.16,13,1,0.928,name,2,?
 35,0,64,0,0.30,6.6,4.36,14,1.27,0.786,name,2,?
 53,0,59,0,0.344,9.1,4.04,9,1,0.643,name,2,0
 33,0,46,0,0.272,16.5,5.36,12.67,1.06,0.857,name,2,?
 ?,1,61,0,0.20,9.4,4.02,15.67,1.42,0.786,name,2,1
 33,0,63,0,0.25,5.6,3.87,18,1.50,0.857,name,2,?
 40,1,74,0,0.20,4.8,4.56,12.5,1.04,0.857,name,2,0
 33,0,59,0,0.50,9.1,3.42,18,1.5,0.857,name,2,?
 5,1,65,1,0.16,8.5,5.47,16,1.45,0.786,name,2,1
 4,1,58,0,0.17,28.9,6.73,26.08,2.01,0.928,name,2,1
 31,0,53,0,0.17,?,4.69,10,1,0.71,name,2,?
 33,0,66,0,0.20,?,4.23,12,1,0.857,name,2,0
 22,0,70,0,0.38,0,4.55,10,1,0.714,name,2,0
 25,0,62,0,0.258,11.8,4.87,11,1,0.786,name,2,?
 1.25,1,63,0,0.30,6.9,3.52,18.16,1.51,0.857,name,2,1
 24,0,59,0,0.17,14.3,5.49,13.5,1.50,0.643,name,2,0
 25,0,57,0,0.228,9.7,4.29,11,1,0.786,name,2,0
 24,0,57,0,0.036,7,4.12,13.5,1.23,0.786,name,2,?
 .75,1,78,0,0.23,40,6.23,14,1.4,0.714,name,2,1
 3,1,62,0,0.26,7.6,4.42,14,1,1,name,2,1
 27,0,62,0,0.22,12.1,3.92,11,1,0.785,name,?,?,
 13,0,66,0,0.24,13.6,4.38,22,2.20,0.714,name,?,?,
 36,0,61,0,0.27,9.00,4.06,12,1,0.857,name,?,?,
 25,0,59,1,0.40,9.20,5.36,12,1,0.857,name,?,?,
 27,0,57,0,0.29,9.40,4.77,9,1,0.64,name,?,?,
 34,0,62,1,0.19,28.9,6.63,19.5,1.95,0.714,name,?,?,
 37,0,?,0,0.26,0,4.38,9,1,0.64,name,?,?,
 34,0,54,0,0.43,9.30,4.79,10,1,0.714,name,?,?,
 28,1,62,1,0.24,28.6,5.86,21.5,1.95,0.786,name,?,?,
 28,0,?,0,0.23,19.1,5.49,12,1.20,0.71,name,?,?,
 17,0,64,0,0.15,6.60,4.17,14,1.27,0.786,name,?,?,
 38,0,57,1,0.12,0,2.32,16.5,1.375,0.857,name,?,?,
 31,0,61,0,0.18,0,4.48,11,1.375,0.57,name,?,?,
 12,0,61,1,0.19,13.2,5.04,19,1.73,0.786,name,?,?,
 36,0,48,0,0.15,12,3.66,10,1,0.714,name,?,?,
 17,0,?,0,0.09,6.80,4.96,13,1.08,0.857,name,?,?,
 21,0,61,0,0.14,25.5,5.16,14,1.27,0.786,name,?,?,
 7.5,1,64,0,0.24,12.9,4.72,12,1,0.857,name,?,?,
 41,0,64,0,0.28,5.40,5.47,11,1.10,0.714,name,?,?,
 36,0,69,0,0.20,7.00,5.05,14.5,1.21,0.857,name,?,?,
 22,0,57,0,0.14,16.1,4.36,15,1.36,0.786,name,?,?,
 20,0,62,0,0.15,0,4.51,15.5,1.409,0.786,name,?,?,

Для преобразования данных из HTML-формата в Excel необходимо:

1. Скопировать текст обучающей выборки в MS Word.

2. Нажать Ctrl+N. В открывшемся окне выбрать вкладку "Заменить". В поле "Найти" поставить символ ",". Выбрать поле "Заменить на", нажать "Больше", затем "Специальный", выбрать "Символ табуляции" и нажать "Заменить все".

3. Преобразовать полученный после этой замены текст в таблицу, используя в качестве разделителя символ табуляции.

4. В полученной таблице заменить:

– точку на запятую, чтобы в MS Excel числовые по смыслу данные и фактически воспринимались как числовые;

– во всех колонках, кроме колонки «Group», заменить 1 на «Yes», а 0 на «No»;

– в колонке «Group» заменить «1» на «one», а «2» на «two»;

– во всех колонках заменить «?» на 0, если колонка числовая, и на пробел, если она текстовая, что будет означать отсутствие данных;

– во всех числовых колонках сделать число десятичных разрядов 2 или 3 в зависимости от количества значащих цифр в исходных данных.

5. Скопировать получившуюся таблицу из MS Word в MS Excel.

6. Используя информацию о наименованиях полей, содержащуюся в общем описании задачи, привести таблицу к виду, представленному в таблице 1. Колонка с именами пациентов переставлена на первое место и вместо реальных имен поставлены условные.

Ниже приведен машинный перевод наименований полей (классификационных и описательных шкал):

1. Имя пациента (вместо реальных имен поставлены условные).

2. Количество месяцев, которое пациент прожил после инфаркта.

3. Еще жив – бинарная переменная. 0 = мертвым в конце периода выживания, 1 означает еще жив.

4. Возраст в годах, когда произошел сердечный приступ.

5. Перикарда-плеврит - двоичный. Перикардит является жидкость вокруг сердца. 0 = нет жидкости, 1 = жидкость.

6. Дробно-укорочение - мера *contracility* вокруг сердца более низкие цифры более ненормальным.

7. EPSS - E-точка разделения перегородки, другая мера сократительной. Большие числа все более ненормальным.

8. LVDD - левого желудочка конечного диастолического размера. это мерой размера сердца на конец диастолы. Большие сердца, как правило, больные сердца.

9. Стена-движение-оценка - мера того, насколько сегментах слева желудочек движутся.

10. Стена-движение-индекс - равна настенного движение-балл, разделенный на число сегменты видел. Обычно 12-13 сегменты видел в эхо-

кардиограмме. Используйте эту переменную ВМЕСТО партитуры движения стенки.

- 11. Мульти - производное вар которые могут быть проигнорированы.
- 12. Группа - бессмысленно, игнорировать его.
- 13. живы-на-1 - булевозначный. Полученный из первых двух атрибутов:

тов:

- 0 означает, пациент либо умер после 1 года или были;
- последовали менее 1 года. 1 означает пациента;
- был жив в течение 1 года.

В результате получаем таблицу исходных данных (таблица 1):

Таблица 1 – Исходные данные (полностью)

The object of training sample	Survival	Still alive	Age at heart attack	Pericardial effusion	Fractional shortening	EPSS	LVDD	Wall motion score	Wall motion index	Mult	Group	Alive at 1
name-1	11,00	No	71,000	No	0,260	9,00	4,600	14,00	1,000	1,000	one	No
name-2	19,00	No	72,000	No	0,380	6,00	4,100	14,00	1,700	0,588	one	No
name-3	16,00	No	55,000	No	0,260	4,00	3,420	14,00	1,000	1,000	one	No
name-4	57,00	No	60,000	No	0,253	12,06	4,603	16,00	1,450	0,788	one	No
name-5	19,00	Yes	57,000	No	0,160	22,00	5,750	18,00	2,250	0,571	one	No
name-6	26,00	No	68,000	No	0,260	5,00	4,310	12,00	1,000	0,857	one	No
name-7	13,00	No	62,000	No	0,230	31,00	5,430	22,50	1,875	0,857	one	No
name-8	50,00	No	60,000	No	0,330	8,00	5,250	14,00	1,000	1,000	one	No
name-9	19,00	No	46,000	No	0,340	0,00	5,090	16,00	1,140	1,003	one	No
name-10	25,00	No	54,000	No	0,140	13,00	4,490	15,50	1,190	0,930	one	No
name-11	10,00	Yes	77,000	No	0,130	16,00	4,230	18,00	1,800	0,714	one	Yes
name-12	52,00	No	62,000	Yes	0,450	9,00	3,600	16,00	1,140	1,003	one	No
name-13	52,00	No	73,000	No	0,330	6,00	4,000	14,00	1,000	1,000	one	No
name-14	44,00	No	60,000	No	0,150	10,00	3,730	14,00	1,000	1,000	one	No
name-15	0,50	Yes	62,000	No	0,120	23,00	5,800	11,67	2,330	0,358	one	Yes
name-16	24,00	No	55,000	Yes	0,250	12,06	4,290	14,00	1,000	1,000	one	No
name-17	0,50	Yes	69,000	Yes	0,260	11,00	4,650	18,00	1,640	0,784	one	Yes
name-18	0,50	Yes	62,529	Yes	0,070	20,00	5,200	24,00	2,000	0,857	one	Yes
name-19	22,00	Yes	66,000	No	0,090	17,00	5,819	8,00	1,333	0,429	one	No
name-20	1,00	Yes	66,000	Yes	0,220	15,00	5,400	27,00	2,250	0,857	one	Yes
name-21	0,75	Yes	69,000	No	0,150	12,00	5,390	19,50	1,625	0,857	one	Yes
name-22	0,75	Yes	85,000	Yes	0,180	19,00	5,460	13,83	1,380	0,710	one	Yes
name-23	0,50	Yes	73,000	No	0,230	12,73	6,060	7,50	1,500	0,360	one	Yes
name-24	5,00	Yes	71,000	No	0,170	0,00	4,650	8,00	1,000	0,570	one	Yes
name-25	48,00	No	64,000	No	0,190	5,90	3,480	10,00	1,110	0,640	two	
name-26	29,00	No	54,000	No	0,300	7,00	3,850	10,00	1,667	0,430	two	
name-27	29,00	No	35,000	No	0,300	5,00	4,170	14,00	1,000	1,000	two	
name-28	29,00	No	55,000	No	0,000	7,00	0,000	2,00	1,000	2,000	two	
name-29	0,25	Yes	75,000	No	0,000	0,00	0,000	0,00	1,000	0,000	two	
name-30	36,00	No	55,000	Yes	0,210	4,20	4,160	14,00	1,560	0,640	two	No
name-31	1,00	Yes	65,000	No	0,150	0,00	5,050	10,00	1,000	0,710	two	
name-32	1,00	Yes	52,000	Yes	0,170	17,20	5,320	14,00	1,170	0,857	two	
name-33	3,00	Yes	0,000	No	0,000	12,00	0,000	6,00	3,000	0,140	two	
name-34	27,00	No	47,000	No	0,400	5,12	3,100	12,00	1,000	0,857	two	
name-35	35,00	No	63,000	No	0,000	10,00	0,000	14,00	1,170	0,857	two	No
name-36	26,00	No	61,000	No	0,610	13,10	4,070	13,00	1,625	0,571	two	No
name-37	16,00	No	63,000	Yes	0,000	0,00	5,310	5,00	1,000	0,357	two	No
name-38	1,00	Yes	65,000	No	0,060	23,60	0,000	21,50	2,150	0,714	two	Yes
name-39	19,00	No	68,000	No	0,510	0,00	3,880	15,00	1,670	0,640	two	No
name-40	31,00	No	80,000	No	0,410	5,40	4,360	0,00	1,000	0,000	two	
name-41	32,00	No	54,000	No	0,350	9,30	3,630	11,00	1,222	0,640	two	No
name-42	16,00	No	70,000	Yes	0,270	4,70	4,490	22,00	2,000	0,786	two	No
name-43	40,00	No	79,000	No	0,150	17,50	4,270	13,00	1,300	0,714	two	No

name-44	46,00	No	56,000	No	0,330	0,00	3,590	14,00	1,000	1,000	two	No
name-45	2,00	Yes	67,000	Yes	0,440	9,00	3,960	17,50	1,450	0,857	two	
name-46	37,00	No	64,000	No	0,090	0,00	0,000	12,00	2,000	0,428	two	
name-47	19,50	Yes	81,000	No	0,120	0,00	0,000	9,00	1,250	0,570	two	No
name-48	20,00	Yes	59,000	No	0,030	21,30	6,290	17,00	1,310	0,928	two	No
name-49	0,25	Yes	63,000	Yes	0,000	0,00	0,000	23,00	2,300	0,714	two	Yes
name-50	0,00		0,000		0,000	0,00	0,000	0,00	0,000	2,000		
name-51	2,00	Yes	56,000	Yes	0,040	14,00	5,000	0,00	0,000	0,000	two	Yes
name-52	7,00	Yes	61,000	Yes	0,270	0,00	0,000	9,00	1,500	0,428	two	Yes
name-53	10,00	No	57,000	No	0,240	14,80	5,260	18,00	1,380	0,812	two	
name-54	12,00	No	58,000	No	0,300	9,40	3,490	14,00	1,000	1,000	two	No
name-55	1,00	Yes	60,000	No	0,010	24,60	5,650	39,00	3,000	0,928	two	Yes
name-56	10,00	No	66,000	No	0,290	15,60	6,150	14,00	1,000	1,000	two	No
name-57	45,00	No	63,000	No	0,150	13,00	4,570	13,00	1,080	0,857	two	No
name-58	22,00	No	57,000	No	0,130	18,60	4,370	12,33	1,370	0,642	two	No
name-59	53,00	No	70,000	No	0,100	9,80	5,300	23,00	2,300	0,714	two	No
name-60	38,00	No	68,000	No	0,290	0,00	4,410	14,00	1,167	0,857	two	
name-61	26,00	No	79,000	No	0,170	11,90	5,150	10,50	1,050	0,714	two	No
name-62	9,00	No	73,000	No	0,120	0,00	6,780	16,67	1,390	0,857	two	
name-63	26,00	No	72,000	No	0,187	12,00	5,020	13,00	1,180	0,785	two	No
name-64	0,50	Yes	59,000	No	0,130	16,40	4,960	17,83	1,370	0,928	two	
name-65	12,00	No	67,000	Yes	0,110	10,30	4,680	11,00	1,000	0,785	two	
name-66	49,00	No	51,000	No	0,160	13,20	5,260	11,00	1,000	0,786	two	No
name-67	0,75	Yes	50,000	No	0,140	11,40	4,750	10,00	2,500	0,280	two	
name-68	49,00	No	70,000	Yes	0,250	9,70	5,570	5,50	1,100	0,357	two	No
name-69	47,00	No	65,000	No	0,360	8,80	5,780	12,00	1,000	0,857	two	No
name-70	41,00	No	78,000	No	0,060	16,10	5,620	13,67	1,367	0,714	two	No
name-71	0,25	Yes	86,000	No	0,225	12,20	5,200	24,00	2,180	0,786	two	Yes
name-72	33,00	No	56,000	No	0,250	11,00	4,720	11,00	1,000	0,785	two	No
name-73	29,00	No	60,000	No	0,120	10,20	4,310	15,00	1,670	0,640	two	No
name-74	41,00	No	59,000	No	0,290	7,50	4,750	13,00	1,080	0,857	two	No
name-75	26,00	No	50,000	No	0,060	30,10	5,950	21,50	2,390	0,643	two	
name-76	15,00	No	54,000	No	0,217	17,90	4,540	16,50	1,180	1,000	two	No
name-77	0,25	Yes	68,000	No	0,220	21,70	4,850	15,00	1,150	0,928	two	
name-78	0,03	Yes	0,000	No	0,260	19,40	4,770	21,00	2,100	0,714	two	Yes
name-79	12,00	No	64,000	No	0,200	7,10	4,580	14,00	1,000	1,000	two	No
name-80	32,00	No	63,000	No	0,200	5,00	5,200	8,00	1,000	0,570	two	
name-81	32,00	No	65,000	No	0,060	23,60	6,740	12,00	1,090	0,785	two	
name-82	27,00	No	54,000	Yes	0,070	16,80	4,160	18,00	1,500	0,857	two	No
name-83	23,00	No	62,000	No	0,250	6,00	4,480	11,00	1,000	0,786	two	
name-84	0,75	Yes	78,000	No	0,050	10,00	4,440	15,00	1,360	0,786	two	Yes
name-85	0,75	Yes	61,000	No	0,000	0,00	0,000	28,00	2,330	0,857	two	Yes
name-86	34,00	No	52,000	No	0,140	25,00	6,210	11,50	1,150	0,714	two	
name-87	1,00	Yes	73,000	No	0,050	14,80	4,140	15,50	1,410	0,786	two	
name-88	21,00	Yes	70,000	Yes	0,160	19,20	5,250	11,00	1,000	0,786	two	
name-89	55,00	No	55,000	No	0,280	5,50	4,480	22,00	1,830	0,857	two	No
name-90	15,00	Yes	60,000	No	0,180	8,70	4,560	13,50	1,040	0,928	two	
name-91	0,50	Yes	67,000	No	0,155	11,30	5,160	13,00	1,000	0,928	two	
name-92	35,00	No	64,000	No	0,300	6,60	4,360	14,00	1,270	0,786	two	
name-93	53,00	No	59,000	No	0,344	9,10	4,040	9,00	1,000	0,643	two	No
name-94	33,00	No	46,000	No	0,272	16,50	5,360	12,67	1,060	0,857	two	
name-95	0,00	Yes	61,000	No	0,200	9,40	4,020	15,67	1,420	0,786	two	Yes
name-96	33,00	No	63,000	No	0,250	5,60	3,870	18,00	1,500	0,857	two	
name-97	40,00	Yes	74,000	No	0,200	4,80	4,560	12,50	1,040	0,857	two	No
name-98	33,00	No	59,000	No	0,500	9,10	3,420	18,00	1,500	0,857	two	
name-99	5,00	Yes	65,000	Yes	0,160	8,50	5,470	16,00	1,450	0,786	two	Yes
name-100	4,00	Yes	58,000	No	0,170	28,90	6,730	26,08	2,010	0,928	two	Yes
name-101	31,00	No	53,000	No	0,170	0,00	4,690	10,00	1,000	0,710	two	
name-102	33,00	No	66,000	No	0,200	0,00	4,230	12,00	1,000	0,857	two	No
name-103	22,00	No	70,000	No	0,380	0,00	4,550	10,00	1,000	0,714	two	No
name-104	25,00	No	62,000	No	0,258	11,80	4,870	11,00	1,000	0,786	two	
name-105	1,25	Yes	63,000	No	0,300	6,90	3,520	18,16	1,510	0,857	two	Yes
name-106	24,00	No	59,000	No	0,170	14,30	5,490	13,50	1,500	0,643	two	No
name-107	25,00	No	57,000	No	0,228	9,70	4,290	11,00	1,000	0,786	two	No
name-108	24,00	No	57,000	No	0,036	7,00	4,120	13,50	1,230	0,786	two	
name-109	0,75	Yes	78,000	No	0,230	40,00	6,230	14,00	1,400	0,714	two	Yes

name-110	3,00	Yes	62,000	No	0,260	7,60	4,420	14,00	1,000	1,000	two	Yes
name-111	27,00	No	62,000	No	0,220	12,10	3,920	11,00	1,000	0,785		
name-112	13,00	No	66,000	No	0,240	13,60	4,380	22,00	2,200	0,714		
name-113	36,00	No	61,000	No	0,270	9,00	4,060	12,00	1,000	0,857		
name-114	25,00	No	59,000	Yes	0,400	9,20	5,360	12,00	1,000	0,857		
name-115	27,00	No	57,000	No	0,290	9,40	4,770	9,00	1,000	0,640		
name-116	34,00	No	62,000	Yes	0,190	28,90	6,630	19,50	1,950	0,714		
name-117	37,00	No	0,000	No	0,260	0,00	4,380	9,00	1,000	0,640		
name-118	34,00	No	54,000	No	0,430	9,30	4,790	10,00	1,000	0,714		
name-119	28,00	Yes	62,000	Yes	0,240	28,60	5,860	21,50	1,950	0,786		
name-120	28,00	No	0,000	No	0,230	19,10	5,490	12,00	1,200	0,710		
name-121	17,00	No	64,000	No	0,150	6,60	4,170	14,00	1,270	0,786		
name-122	38,00	No	57,000	Yes	0,120	0,00	2,320	16,50	1,375	0,857		
name-123	31,00	No	61,000	No	0,180	0,00	4,480	11,00	1,375	0,570		
name-124	12,00	No	61,000	Yes	0,190	13,20	5,040	19,00	1,730	0,786		
name-125	36,00	No	48,000	No	0,150	12,00	3,660	10,00	1,000	0,714		
name-126	17,00	No	0,000	No	0,090	6,80	4,960	13,00	1,080	0,857		
name-127	21,00	No	61,000	No	0,140	25,50	5,160	14,00	1,270	0,786		
name-128	7,50	Yes	64,000	No	0,240	12,90	4,720	12,00	1,000	0,857		
name-129	41,00	No	64,000	No	0,280	5,40	5,470	11,00	1,100	0,714		
name-130	36,00	No	69,000	No	0,200	7,00	5,050	14,50	1,210	0,857		
name-131	22,00	No	57,000	No	0,140	16,10	4,360	15,00	1,360	0,786		
name-132	20,00	No	62,000	No	0,150	0,00	4,510	15,50	1,409	0,786		

2.3. Этапы АСК-анализа и преобразование исходных данных в информацию, а ее в знания в системе "Эйдос"

АСК-анализ включает следующие этапы [3, 4]:

1. Когнитивная структуризация предметной области (неформализованный этап). На этом этапе решается, что мы хотим прогнозировать и на основе чего. В нашей задаче мы хотим прогнозировать продолжительность жизни пациента после перенесенного им инфаркта на основе анализа эхокардиограммы.

2. Формализация предметной области. На этом этапе разрабатываются классификационные и описательные шкалы и градации, а затем с их использованием исходные данные кодируются и представляются в форме баз событий, между которыми могут быть выявлены причинно-следственные связи.

3. Синтез и верификация моделей (оценка достоверности, адекватности). Повышение качества модели. Выбор наиболее достоверной модели для решения в ней задач.

4. Решение задач идентификации и прогнозирования.

5. Решение задач принятия решений и управления.

6. Решение задач исследования моделируемой предметной области путем исследования ее модели.

На рисунке 1 приведены автоматизированные в системе «Эйдос» этапы АСК-анализа, которые обеспечивают последовательное повышение степени формализации модели путем преобразования исходных данных в информацию, а далее в знания (рисунок 1):

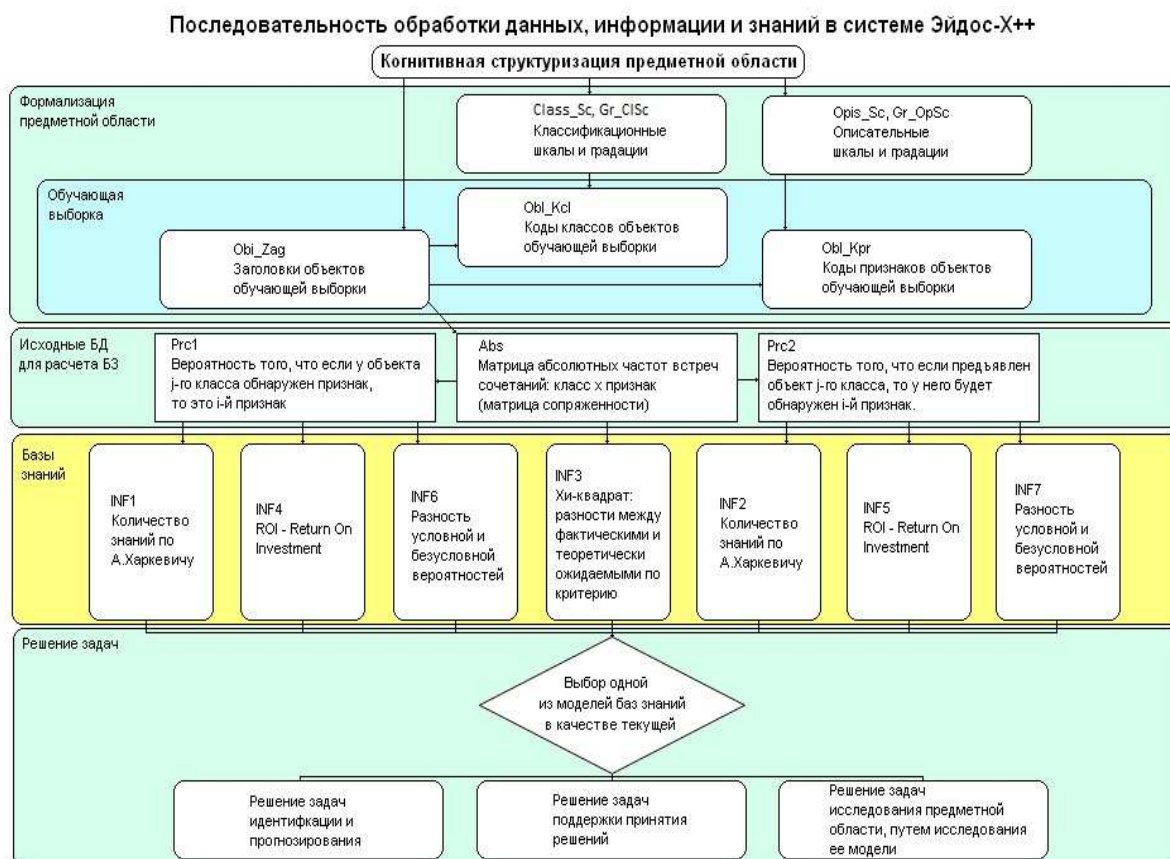


Рисунок 1 – Этапы последовательного преобразования данных в информацию, а ее в знания в системе "Эйдос"

Подробно этот процесс описан в работе [10]. Суть этого процесса в следующем:

1. *Информация* рассматривается как осмысленные исходные данные.
2. *Смысл*, согласно концепции Шенка-Абельсона [12] считается известным, когда выявлены причинно-следственные связи.
3. *Анализ* – это операция выявления смысла из исходных данных.
4. Причинно-следственные связи существуют не между элементами исходных данных, а между реальными событиями, которые они отражают (моделируют), т.е. причинно-следственные связи – это характеристика реальной области, а не абстрактных моделей. Иначе говоря, анализ самих исходных данных невозможен, а возможен только анализ событий, описанных этими исходными данными.
5. Поэтому перед анализом исходных данных необходимо предварительно преобразовать их в базы событий, т.е. в эвентологические базы.
6. Это преобразование осуществляется с помощью справочников событий, факторов и их значений, т.е. с помощью классификационных и описательных шкал и градаций, которые также необходимо разработать.
7. Формализация предметной области представляет собой разработку справочников классификационных и описательных шкал и градаций и преобразование с их помощью баз исходных данных в базы событий (т.е.

обучающую выборку), и является первым автоматизированным в системе «Эйдос» этапом АСК-анализа.

8. Затем следуют остальные перечисленные выше этапы АСК-анализа:

– синтез и верификация моделей и выбор наиболее достоверной из них;

– решение в ней задач идентификации, прогнозирования, принятия решений и исследования предметной области, т.е. преобразование информации в знания.

Этап синтеза и верификации моделей завершает процесс анализа исходных данных и преобразования их в информацию, а ее в знания.

В АСК-анализе есть несколько режимов, обеспечивающих решение задачи принятия решений для *управления* или достижения *целей*, которая представляет собой обратную задачу прогнозирования: это и режим 4.2.1, позволяющий формировать информационные портреты классов, а также режим 4.4.8, поддерживающий количественный автоматизированный SWOT и –PEST анализ, включая построение SWOT и –PEST матриц и диаграмм [13], а также режим 4.4.10, визуализирующий нейросетевую интерпретацию модели знаний системы «Эйдос» [14]. Эти режимы обеспечивают преобразование информации в знания, т.к. **знания представляют собой информацию, полезную для достижения целей, т.е. по сути технологию, в частности ноу-хау** [10]. Наличие цели является ключевым моментом для преобразования информации в знания. А постановка целей (целеполагание) не мыслима без мотивации, которая в настоящее время является слабо формализованным этапом.

Итак, в процессе *анализа* исходные данные представляются в форме базы событий, между которыми выявляются причинно-следственные связи, и, таким образом, исходные данные преобразуются в информацию, представляющую собой осмысленные данные (смысл есть знание причинно-следственных связей), а затем информация используется для достижения целей (управления), т.е. преобразуется в знания.

Формализация предметной области включает разработку классификационных и описательных шкал и градаций и преобразование с их использованием исходных данных (таблица 1) в обучающую выборку. Этот этап *полностью автоматизируется* программным интерфейсом системы «Эйдос» с внешними табличными базами исходных данных (режим 2.3.2.2).

Но перед выполнением этого этапа АСК-анализа, естественно, необходимо сначала скачать и установить систему «Эйдос».

2.4. Скачивание и инсталляция системы «Эйдос»

Для скачивания и инсталляции системы «Эйдос» необходимо по адресу: http://lc.kubagro.ru/aidos/_Aidos-X.htm открыть и выполнить следующую инструкцию³:

ИНСТРУКЦИЯ по скачиванию и установке системы «Эйдос» (объем около 50 Мб)

Система не требует инсталляции, не меняет никаких системных файлов и содержимого папок операционной системы,

т.е. является портативной (portable) программой. Но чтобы она работала необходимо аккуратно выполнить следующие пункты.

1. Скачать самую новую на текущий момент версию системы «Эйдос-X++» по ссылке:

<http://lc.kubagro.ru/a.rar> (ссылки для обновления системы даны в режиме 6.2)

2. Разархивировать этот архив в любую папку с правами на запись с коротким латинским именем и путем доступа,

включающим только папки с такими же именами (лучше всего в корневой каталог какого-нибудь диска).

3. Запустить систему. Файл запуска: `_AIDOS-X.exe`

4. Задать имя: 1 и пароль: 1 (потом их можно поменять в режиме 1.2).

5. Выполнить режим 1.11 (только 1-й раз при установке системы в данную папку)

6. Перед тем как запустить новый режим НЕОБХОДИМО ЗАВЕРШИТЬ предыдущий (Help можно не закрывать)

Разработана программа: «_START_AIDOS.exe», полностью снимающая с пользователя системы «Эйдос-X++» заботу о проверке наличия и скачивании обновлений.

Эту программу надо просто скачать по ссылке: http://lc.kubagro.ru/Install_Aidos-X/_START_AIDOS.exe, поместить в папку с исполнимым модулем системы и всегда запускать систему с помощью этого файла.

При запуске программы `_START_AIDOS.EXE` система Эйдос не должна быть запущена, т.к. она содержится в файле обновлений и при его разархивировании возникнет конфликт, если система будет запущена.

1. Программа `_START_AIDOS.exe` определяет дату системы Эйдос в текущей папке, и дату обновлений на FTP-сервере не скачивая их, и, если система Эйдос устарела, скачивает обновления. (Если в текущей папке нет исполнимого модуля системы Эйдос, то программа пытается скачать полную инсталляцию системы, но не может этого сделать из-за ограниченной функциональности демо-версии библиотеки `Xb2NET.DLL`).

2. После этого появляется диалоговое окно с сообщением, что надо сначала разархивировать систему, заменяя все файлы (опция: «Yes to All» или «OverWrite All»), и только после этого закрыть данное окно. Появление окна архиватора происходит на мгновение позже появления диалогового окна, практически одновременно с ним и поверх него. Поэтому чтобы прочитать окно с этим сообщением нужно сдвинуть окно архиватора.

3. Потом программа `_START_AIDOS.exe` запускает обновления на разархивирование. После окончания разархивирования окно архиватора с отображением стадии процесса исчезает.

4. После закрытия диалогового окна с инструкцией (см. п.2), происходит запуск обновленной версии системы Эйдос на исполнение.

Для работы программы `_START_AIDOS.exe` необходима библиотека: `Xb2NET.DLL`, которую можно скачать по ссылке: http://lc.kubagro.ru/Install_Aidos-X/Xb2NET.DLL. Перед первым запуском этой программы данную библиотеку необходимо скачать и поместить либо в папку с этой программой, а значит и исполнимым модулем системы «Эйдос-X++», либо в любую другую папку, на которую в операционной системе прописаны пути поиска файлов, например в папку: `c:\Windows\System32`. Эта библиотека стоит около 500\$ и у меня ее нет, поэтому я даю только бесплатную демо-версию, которая выдает сообщение об ограниченной функциональности, но для наших целей ее достаточно.

По этим ссылкам всегда размещена наиболее полная на момент скачивания незащищенная от несанкционированного копирования портативная (portable) версия системы (не требующая инсталляции) с исходными

³ Имеется и форум автора системы проф.Е.В.Луценко для пользователей системы:
<http://proflutsenko.vdforum.ru/>

текстами, находящаяся в полном открытом бесплатном доступе (объем около 50 Мб). Обновление имеет объем около 3 Мб.

Для преобразования исходных данных обучающей выборки в базы данных системы "Эйдос" необходимо файл MS Excel, который содержит базу исходных данных, скопировать в папку: Aidos-X\AID_DATA\Inp_data и назвать Inp_data.xls. Этот файл организован следующим образом (рисунок 2):

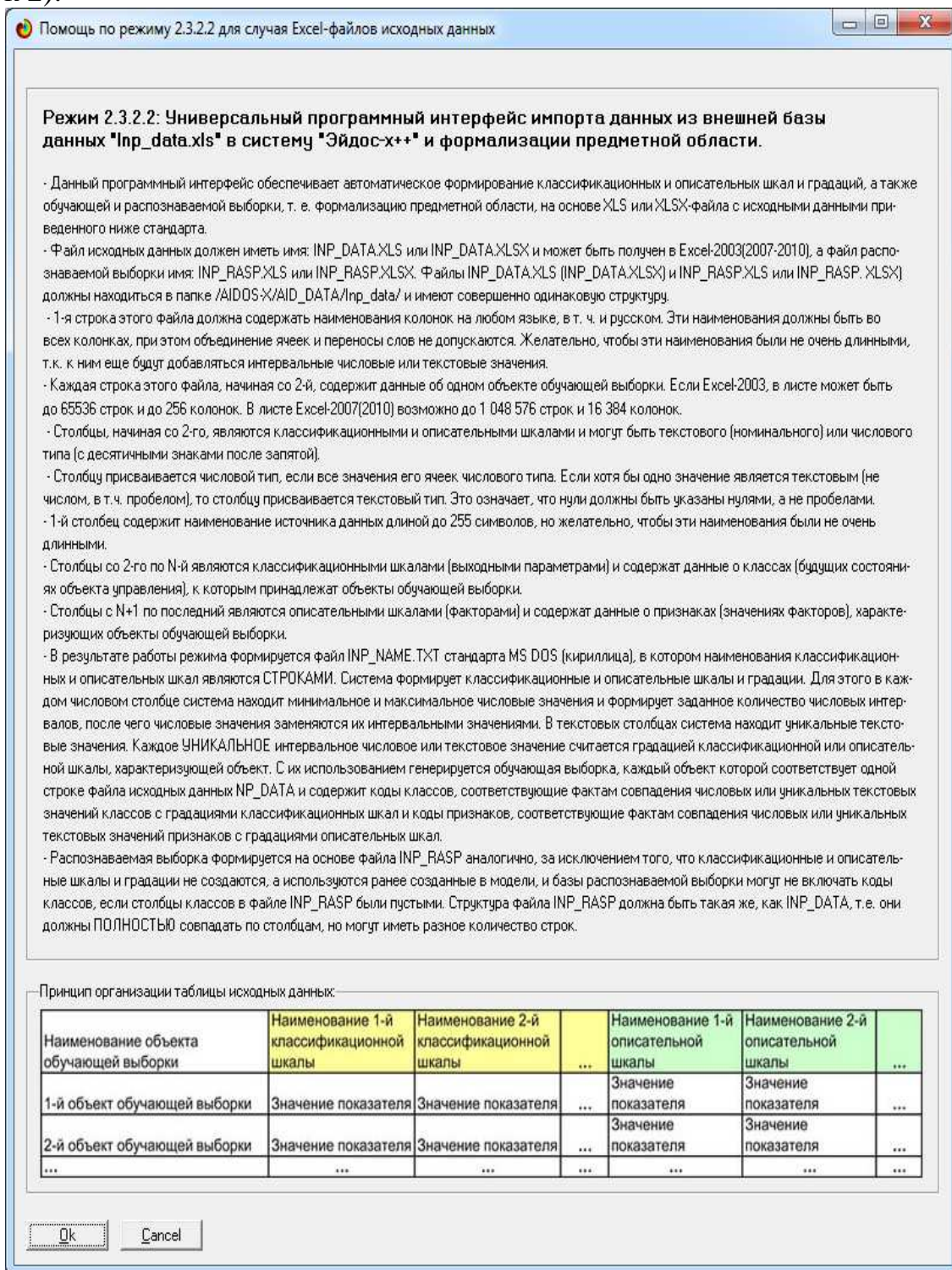


Рисунок 2 – Help по режиму 2.3.2.2 системы «Эйдос»

2.5. 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 (таблица 1), т.е. режимом 2.3.2.2 (рисунок 3):

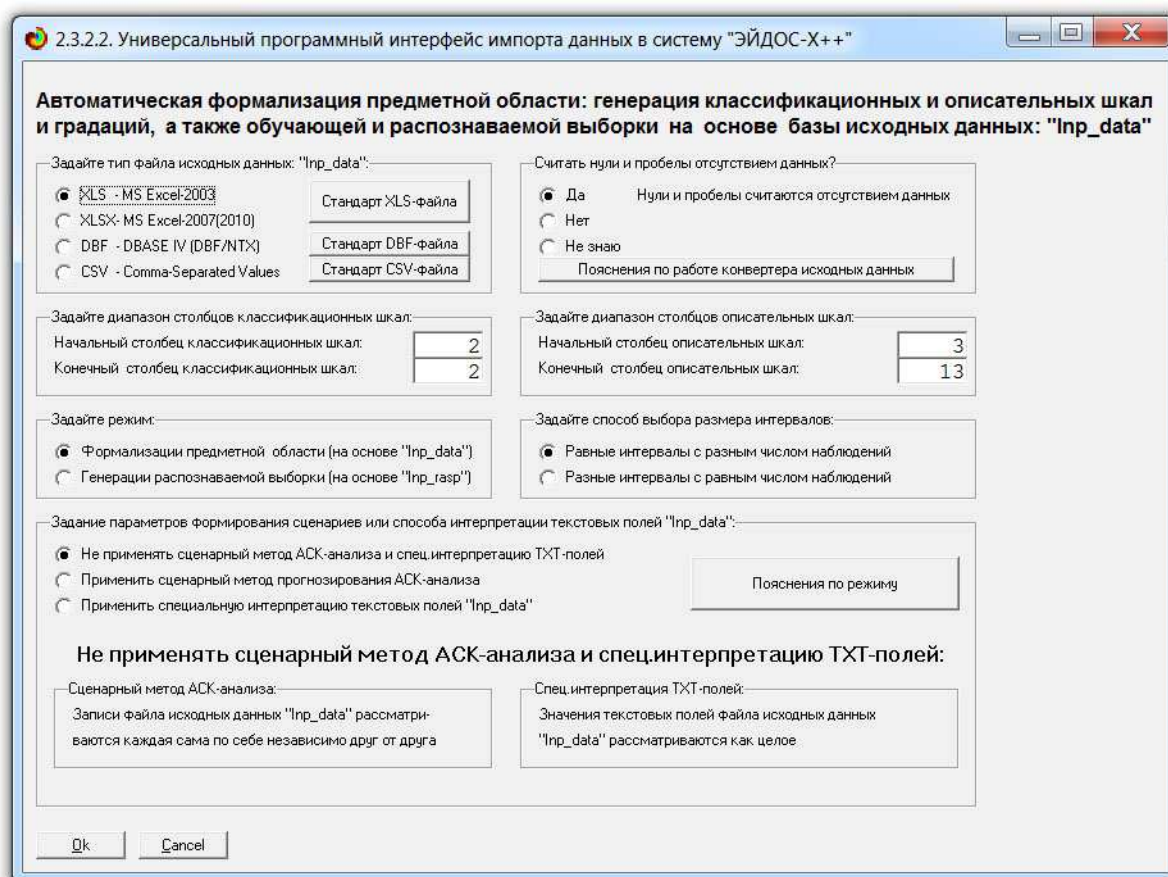


Рисунок 3 – Экранная форма Универсального программного интерфейса импорта данных в систему "Эйдос" (режим 2.3.2.2.)

В экранной форме, приведенной на рисунке 2, 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" – 2, "Конечный столбец классификационных шкал" – 2;
- "Задайте диапазон столбцов описательных шкал": "Начальный столбец описательных шкал" – 3, "Конечный столбец описательных шкал" – 13;

– "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 4). В этом окне необходимо нажать кнопку "Выйти на создание модели".

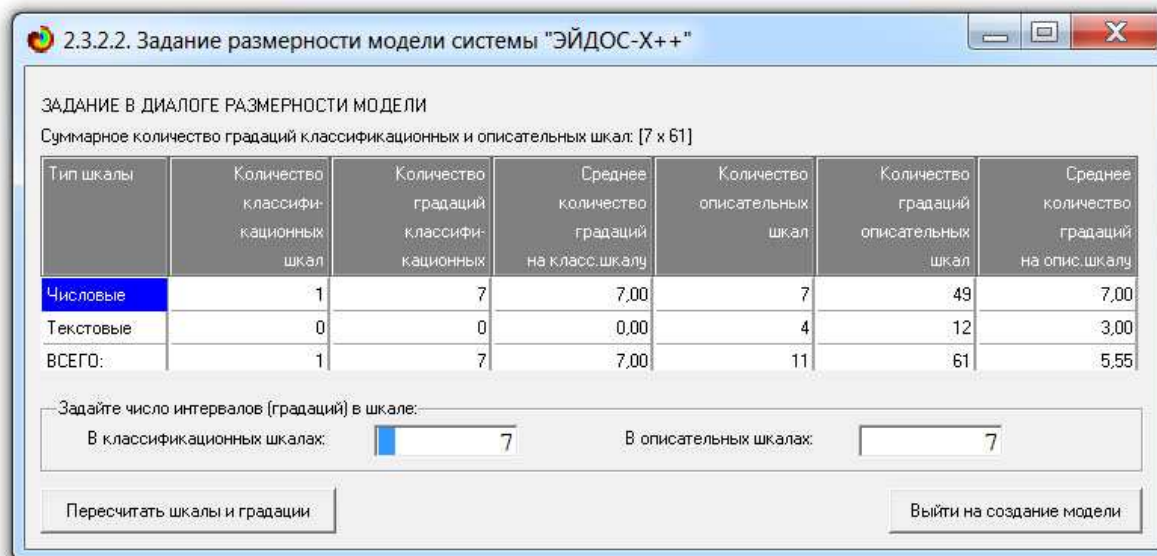


Рисунок 4 – Задание размерности модели системы "Эйдос"

Далее открывается окно, отображающее стадию процесса импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос" (рисунок 5), а также прогноз времени завершения этого процесса. В том окне необходимо дождаться завершения формализации предметной области и нажать кнопку "ОК".

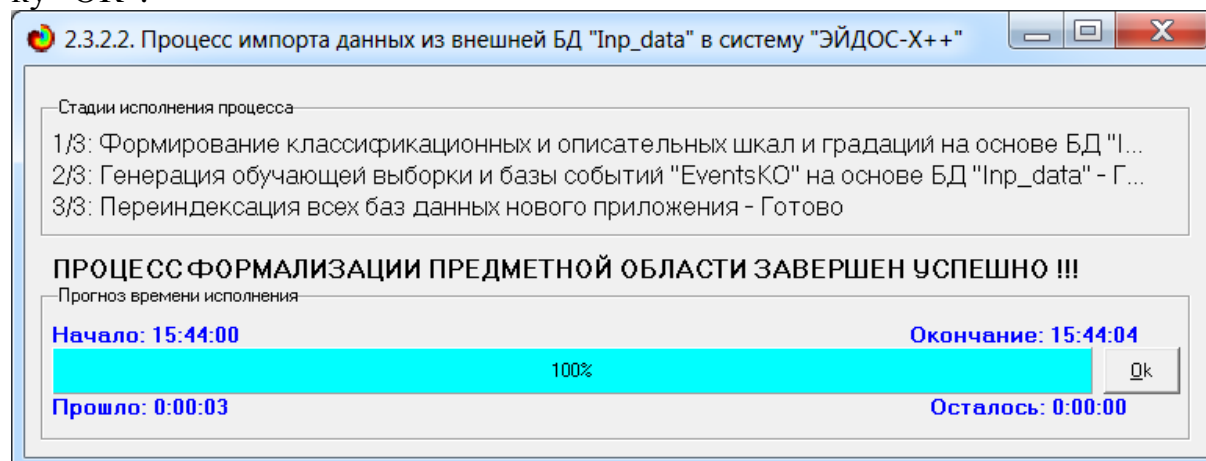


Рисунок 5 – Процесс импорта данных из внешней БД "Inp_data.xls" в систему "Эйдос"

Для просмотра классификационных шкал и градаций необходимо запустить режим 2.1 (рисунок 6):

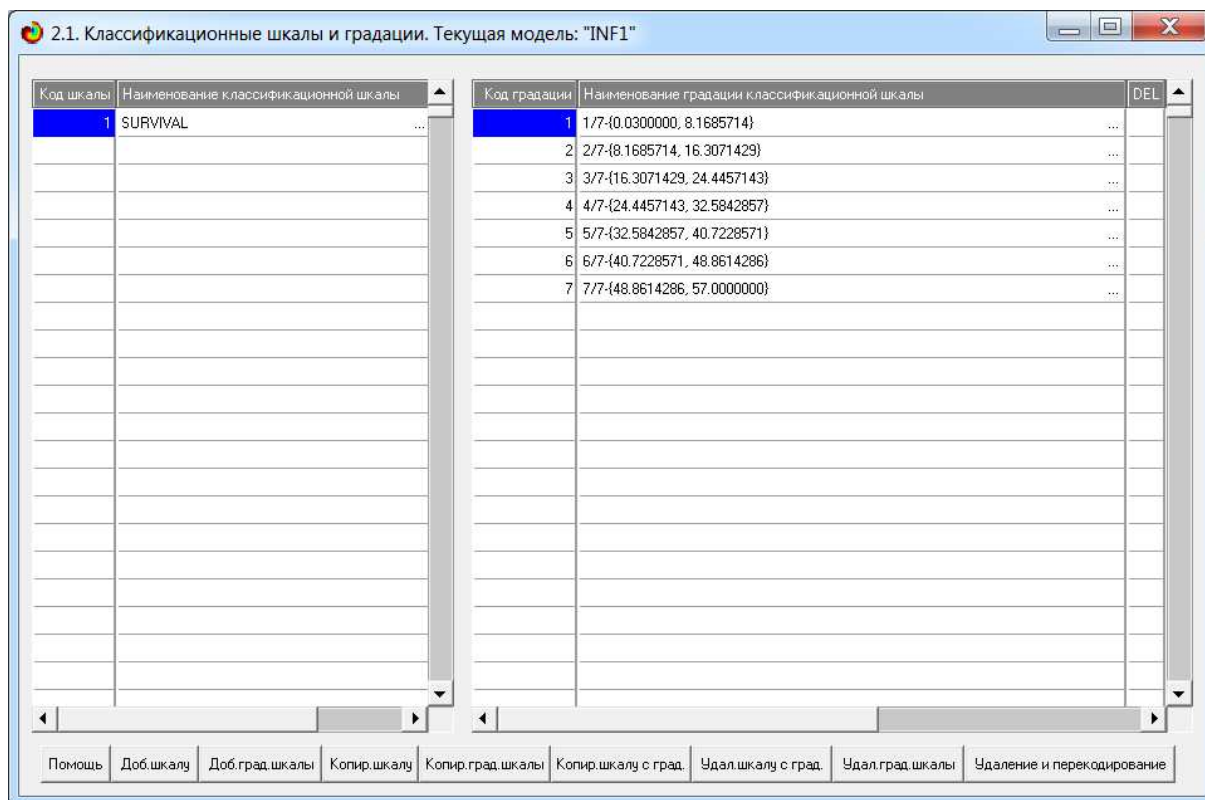


Рисунок 6 – Классификационные шкалы и градации (фрагмент)

Для просмотра описательных шкал и градаций необходимо запустить режим 2.2 (рисунок 7):

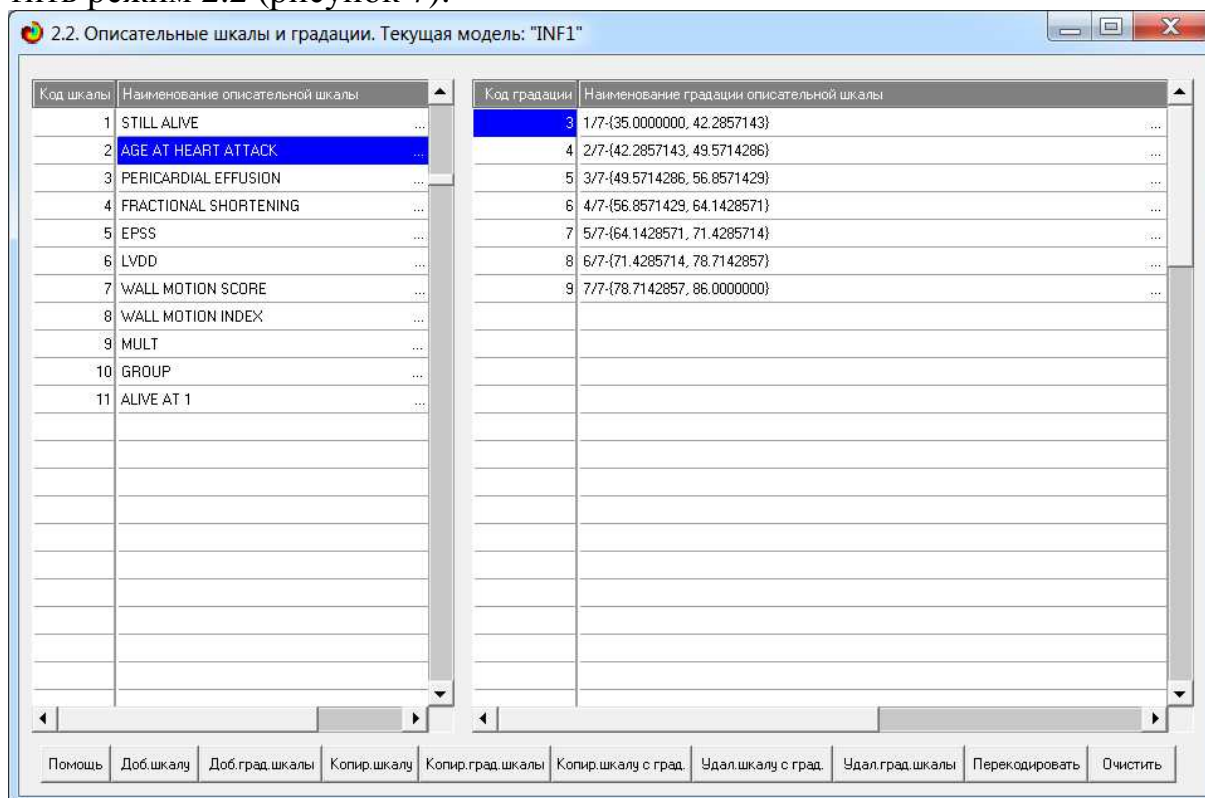


Рисунок 7 – Описательные шкалы и градации (фрагмент)

Для просмотра обучающей выборки необходимо запустить режим 2.3.1. (рисунок 8):

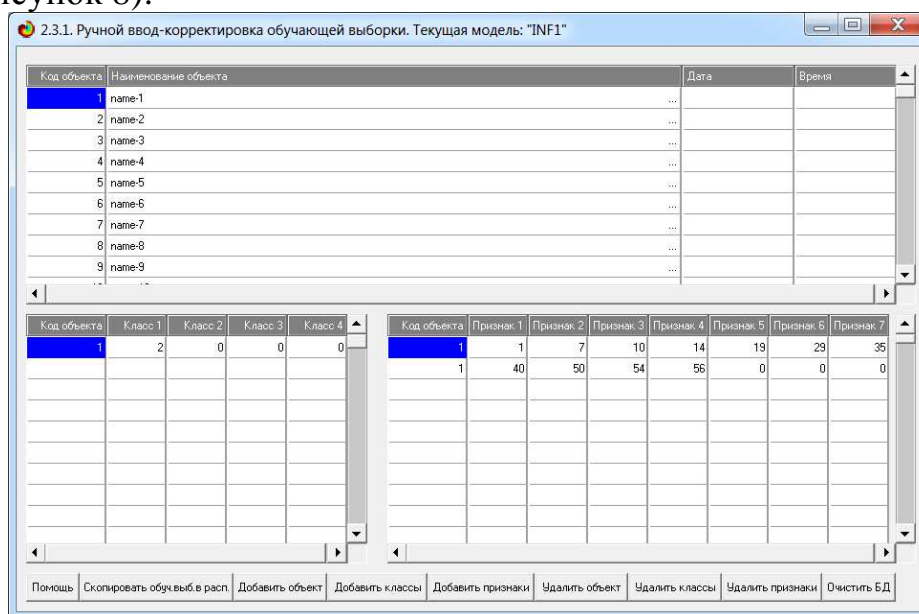


Рисунок 8 – Обучающая выборка (фрагмент)

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]).

2.6. Синтез и верификация статистических и интеллектуальных моделей

Далее запускаем систему "Эйдос" из папки "Aidos-X" файлом _aidos-x.exe. Система попросит ввести логин и пароль (рисунок 9). Необходимо ввести: логин – 1, пароль – 1.

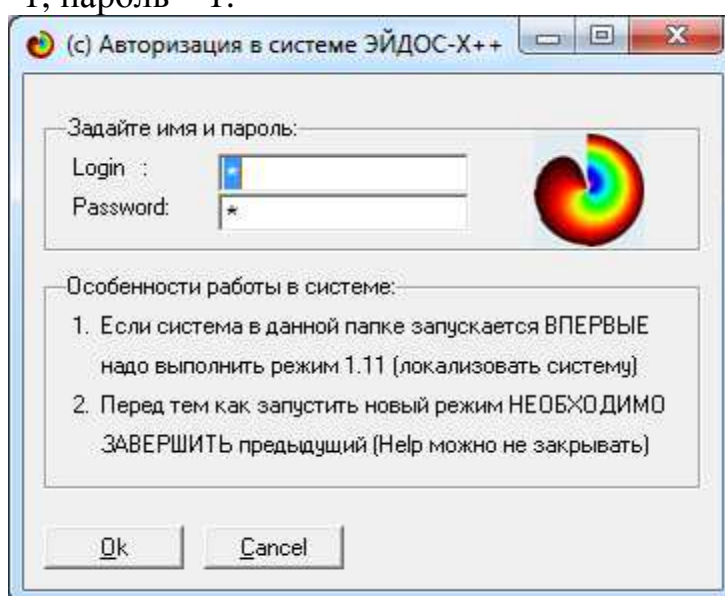


Рисунок 9 – Авторизация в системе "Эйдос"

После откроется главное окно программы. Далее необходимо выполнить режим 1.11 для сброса приложений и локализации системы (рисунок 10).

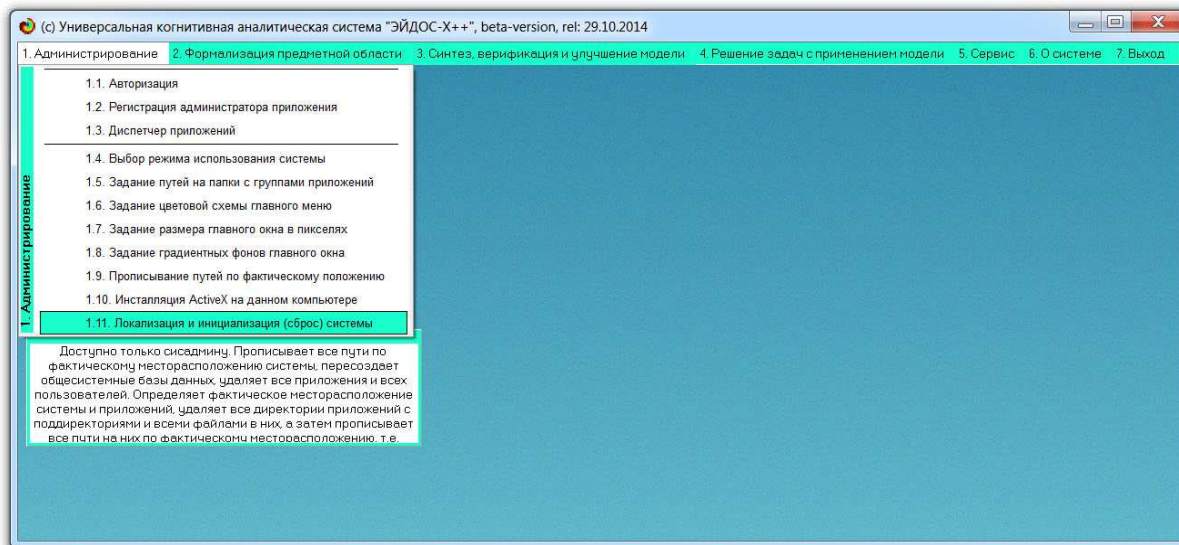


Рисунок 10 – Главное окно программы "Эйдос"

Далее запускаем режим 3.5, в котором происходит выбор моделей для синтеза и верификации (рисунок 11) и нажмем кнопку "ОК". После успешного завершения, также необходимо нажать кнопку "ОК" (рисунок 12).

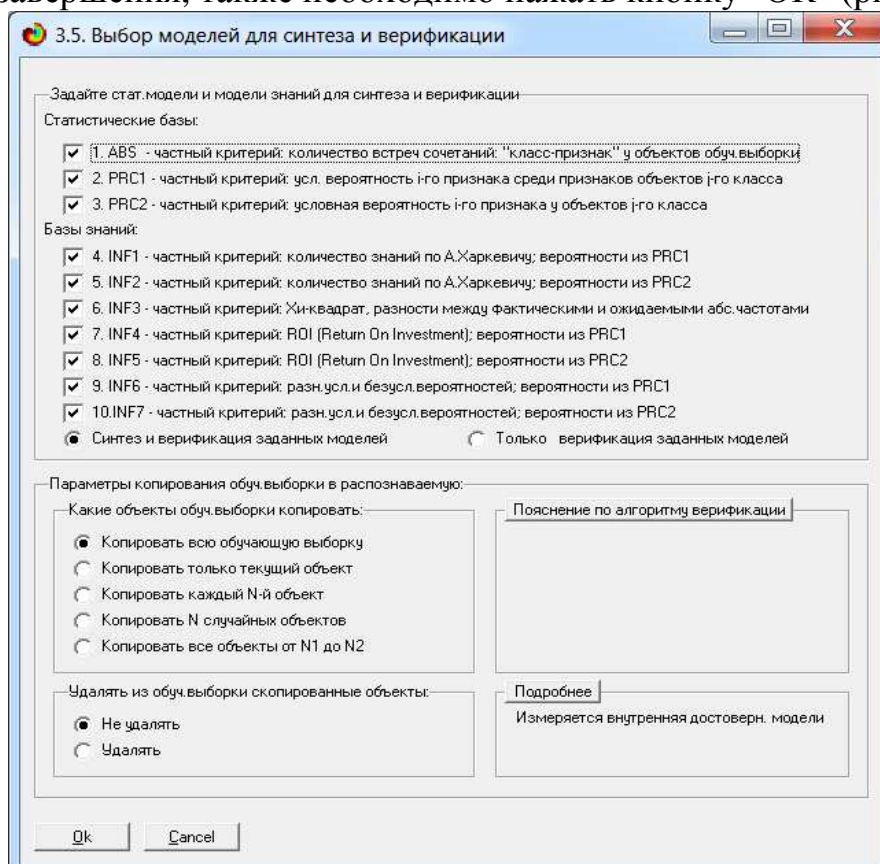


Рисунок 11 – Выбор моделей для синтеза и верификации

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 11.

В результате выполнения режима 3.5 (рисунок 12) созданы все модели, со всеми частными критериями, перечисленные на рисунке 11, но ниже мы приведем лишь некоторые из них (таблицы 3-5).

Предварительно рассмотрим частные и интегральные критерии, применяемые в настоящее время в системе «Эйдос».

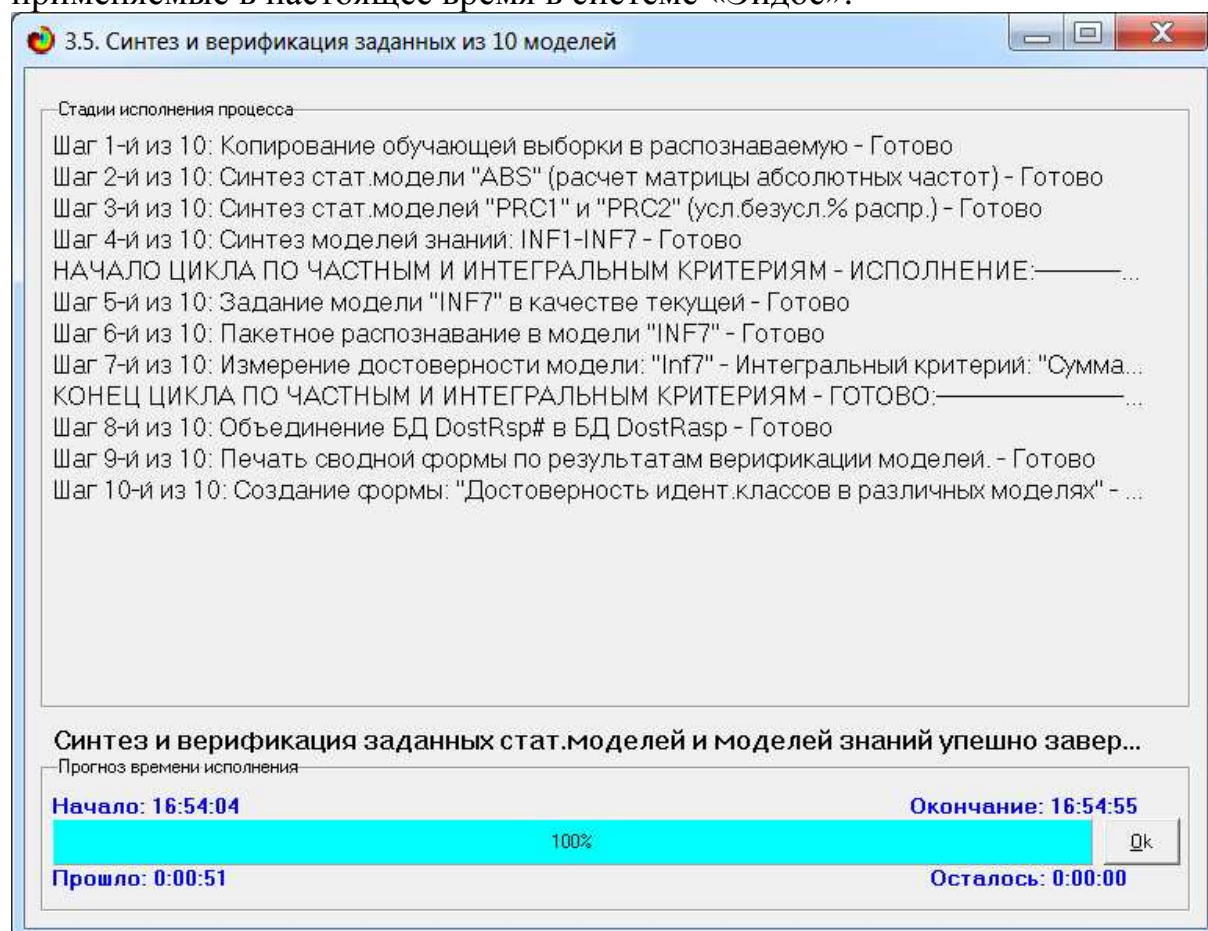


Рисунок 12 – Синтез и верификация статистических моделей и моделей знаний

Интересно заметить (см. рисунок 12), что синтез и верификация всех 10 моделей заняла 51 секунду.

2.7. Частные критерии и 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. Это так называемые *частные* критерии сходства, приведенные в таблице 2.

Таблица 2 – Частные критерии знаний, используемые в настоящее время в АСК-анализе и системе «Эйдос-Х++»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	через относительные частоты	через абсолютные частоты
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу. Относительная частота того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_iN_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу. Относительная частота того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.	$I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_iN_j}$
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	---	$I_{ij} = N_{ij} - \frac{N_iN_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$
INF6 , частный критерий: разность условной и безусловной относительных частот, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$
INF7 , частный критерий: разность условной и безусловной относительных частот, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$

Обозначения:

i – значение прошлого параметра;

j – значение будущего параметра;

N_{ij} – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра;

M – суммарное число значений всех прошлых параметров;

W – суммарное число значений всех будущих параметров.

N_i – количество встреч i -м значения прошлого параметра по всей выборке;

N_j – количество встреч j -го значения будущего параметра по всей выборке;

N – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра по всей выборке.

I_{ij} – частный критерий знаний: количество знаний в факте наблюдения i -го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее j -му значению будущего параметра;

Ψ – нормировочный коэффициент (Е.В.Луценко, 1979, впервые опубликовано в 1993 году [15]), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;

P_i – безусловная относительная частота встречи i -го значения прошлого параметра в обучающей выборке;

P_{ij} – условная относительная частота встречи i -го значения прошлого параметра при j -м значении будущего параметра.

По сути, частные критерии представляют собой просто формулы для преобразования матрицы абсолютных частот (таблица 3)⁴ в матрицы ус-

⁴ Которая является также матрицей сопряженности или корреляционной матрицей.

ловных и безусловных процентных распределений (таблицы 4 и 4) и матрицы знаний (проф. В.И.Лойко).

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

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

Код признака	Наименование описательной шкалы и градации	1. SURVIVAL... 8.1685714}	2. SURVIVAL... 16.3071429}	3. SURVIVAL... 24.4457143}	4. SURVIVAL... 32.5842857}	5. SURVIVAL... 40.7228571}	6. SURVIVAL... 48.8614286}	7. SURVIVAL... 57.0000000}
1	STILL ALIVE-No		14	14	24	19	8	9
2	STILL ALIVE-Yes	33	2	5	1	1		
3	AGE AT HEART ATTACK-1/7-{35.0000000, 42.2857143}				1			
4	AGE AT HEART ATTACK-2/7-{42.2857143, 49.5714286}			1	1	2		
5	AGE AT HEART ATTACK-3/7-{49.5714286, 56.8571429}	3	2	1	7	4	1	2
6	AGE AT HEART ATTACK-4/7-{56.8571429, 64.1428571}	11	7	10	10	8	5	4
7	AGE AT HEART ATTACK-5/7-{64.1428571, 71.4285714}	10	5	4	2	3	1	2
8	AGE AT HEART ATTACK-6/7-{71.4285714, 78.7142857}	5	2	1	1	1	1	1
9	AGE AT HEART ATTACK-7/7-{78.7142857, 86.0000000}	2		1	2	1		
10	PERICARDIAL EFFUSION-No	23	12	17	22	17	8	7
11	PERICARDIAL EFFUSION-Yes	10	4	2	3	3		2
12	FRACTIONAL SHORTENING-1/7-{0.0100000, 0.0957143}	6		4	3	1	1	
13	FRACTIONAL SHORTENING-2/7-{0.0957143, 0.1814286}	11	4	9	5	4	2	2
14	FRACTIONAL SHORTENING-3/7-{0.1814286, 0.2671429}	9	8	2	8	8	1	2
15	FRACTIONAL SHORTENING-4/7-{0.2671429, 0.3528571}	2	3	1	4	4	3	4
16	FRACTIONAL SHORTENING-5/7-{0.3528571, 0.4385714}			2	3	1	1	
17	FRACTIONAL SHORTENING-6/7-{0.4385714, 0.5242857}	1		1		1		1
18	FRACTIONAL SHORTENING-7/7-{0.5242857, 0.6100000}							
19	EPSS-1/7-{4.0000000, 9.1428571}	4	5	5	7	7	4	5

5.5. Модель: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"

Код признака	Наименование описательной шкалы и градации	1. SURVIVAL... 8.1685714}	2. SURVIVAL... 16.3071429}	3. SURVIVAL... 24.4457143}	4. SURVIVAL... 32.5842857}	5. SURVIVAL... 40.7228571}	6. SURVIVAL... 48.8614286}	7. SURVIVAL... 57.0000000}
1	STILL ALIVE-No		87.500	73.684	96.000	95.000	100.000	100.000
2	STILL ALIVE-Yes	100.000	12.500	26.316	4.000	5.000		
3	AGE AT HEART ATTACK-1/7-{35.0000000, 42.2857143}				4.000			
4	AGE AT HEART ATTACK-2/7-{42.2857143, 49.5714286}			5.263	4.000	10.000		
5	AGE AT HEART ATTACK-3/7-{49.5714286, 56.8571429}	9.091	12.500	5.263	28.000	20.000	12.500	22.222
6	AGE AT HEART ATTACK-4/7-{56.8571429, 64.1428571}	33.333	43.750	52.632	40.000	40.000	62.500	44.444
7	AGE AT HEART ATTACK-5/7-{64.1428571, 71.4285714}	30.303	31.250	21.053	8.000	15.000	12.500	22.222
8	AGE AT HEART ATTACK-6/7-{71.4285714, 78.7142857}	15.152	12.500	5.263	4.000	5.000	12.500	11.111
9	AGE AT HEART ATTACK-7/7-{78.7142857, 86.0000000}	6.061		5.263	8.000	5.000		
10	PERICARDIAL EFFUSION-No	69.697	75.000	89.474	88.000	85.000	100.000	77.778
11	PERICARDIAL EFFUSION-Yes	30.303	25.000	10.526	12.000	15.000		22.222
12	FRACTIONAL SHORTENING-1/7-{0.0100000, 0.0957143}	18.182		21.053	12.000	5.000	12.500	
13	FRACTIONAL SHORTENING-2/7-{0.0957143, 0.1814286}	33.333	25.000	47.368	20.000	20.000	25.000	22.222
14	FRACTIONAL SHORTENING-3/7-{0.1814286, 0.2671429}	27.273	50.000	10.526	32.000	40.000	12.500	22.222
15	FRACTIONAL SHORTENING-4/7-{0.2671429, 0.3528571}	6.061	18.750	5.263	16.000	20.000	37.500	44.444
16	FRACTIONAL SHORTENING-5/7-{0.3528571, 0.4385714}			10.526	12.000	5.000	12.500	
17	FRACTIONAL SHORTENING-6/7-{0.4385714, 0.5242857}	3.030		5.263		5.000		11.111
18	FRACTIONAL SHORTENING-7/7-{0.5242857, 0.6100000}				4.000			
19	EPSS-1/7-{4.0000000, 9.1428571}	12.121	31.250	26.316	28.000	35.000	50.000	55.556

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

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

Код признака	Наименование описательной шкалы и градации	1. SURVIVAL: 8.1685714}	2. SURVIVAL: 16.3071429}	3. SURVIVAL: 24.4457143}	4. SURVIVAL: 32.5842857}	5. SURVIVAL: 40.7228571}	6. SURVIVAL: 48.8614286}	7. SURVIVAL: 57.0000000}
1	STILL ALIVE-No		0.089	0.036	0.143	0.153	0.134	0.116
2	STILL ALIVE-Yes	0.445	-0.383	-0.078	-0.811	-0.710		
3	AGE AT HEART ATTACK-1/7-{35.0000000, 42.28...}				0.652			
4	AGE AT HEART ATTACK-2/7-{42.2857143, 49.57...}			0.213	0.109	0.482		
5	AGE AT HEART ATTACK-3/7-{49.5714286, 56.85...}	-0.203	-0.093	-0.418	0.241	0.123	-0.100	0.107
6	AGE AT HEART ATTACK-4/7-{56.8571429, 64.14...}	-0.091	0.001	0.088	-0.015	-0.002	0.134	-0.018
7	AGE AT HEART ATTACK-5/7-{64.1428571, 71.42...}	0.151	0.148	0.008	-0.367	-0.107	-0.218	-0.011
8	AGE AT HEART ATTACK-6/7-{71.4285714, 78.71...}	0.197	0.107	-0.218	-0.321	-0.220	0.100	0.036
9	AGE AT HEART ATTACK-7/7-{78.7142857, 86.00...}	0.109		0.054	0.222	0.052		
10	PERICARDIAL EFFUSION-No	-0.059	-0.044	0.039	0.036	0.037	0.061	-0.055
11	PERICARDIAL EFFUSION-Yes	0.197	0.107	-0.218	-0.162	-0.061		0.036
12	FRACTIONAL SHORTENING-1/7-{0.0100000, 0.0...}	0.181		0.238	0.022	-0.307	0.013	
13	FRACTIONAL SHORTENING-2/7-{0.0957143, 0.1...}	0.065	-0.063	0.202	-0.132	-0.118	-0.070	-0.134
14	FRACTIONAL SHORTENING-3/7-{0.1814286, 0.2...}	-0.024	0.198	-0.397	0.042	0.143	-0.351	-0.144
15	FRACTIONAL SHORTENING-4/7-{0.2671429, 0.3...}	-0.381	0.047	-0.437	0.003	0.104	0.311	0.359
16	FRACTIONAL SHORTENING-5/7-{0.3528571, 0.4...}			0.265	0.320	-0.009	0.311	
17	FRACTIONAL SHORTENING-6/7-{0.4385714, 0.5...}	-0.003		0.213		0.210		0.466
18	FRACTIONAL SHORTENING-7/7-{0.5242857, 0.6...}				0.652			
19	EPSS-1/7-{4.0000000, 9.1428571}	-0.331	0.025	-0.028	0.000	0.101	0.202	0.225

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

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

Код признака	Наименование описательной шкалы и градации	1. SURVIVAL: 8.1685714}	2. SURVIVAL: 16.3071429}	3. SURVIVAL: 24.4457143}	4. SURVIVAL: 32.5842857}	5. SURVIVAL: 40.7228571}	6. SURVIVAL: 48.8614286}	7. SURVIVAL: 57.0000000}
1	STILL ALIVE-No	-22.186	2.839	1.216	7.360	6.148	2.318	2.304
2	STILL ALIVE-Yes	22.411	-3.327	-1.101	-6.942	-5.134	-2.712	-3.196
3	AGE AT HEART ATTACK-1/7-{35.0000000, 42.28...}	-0.252	-0.127	-0.145	0.811	-0.146	-0.065	-0.076
4	AGE AT HEART ATTACK-2/7-{42.2857143, 49.57...}	-1.008	-0.507	0.419	0.244	1.416	-0.258	-0.304
5	AGE AT HEART ATTACK-3/7-{49.5714286, 56.85...}	-2.042	-0.537	-1.905	3.218	1.079	-0.291	0.478
6	AGE AT HEART ATTACK-4/7-{56.8571429, 64.14...}	-2.866	0.025	2.010	-0.400	-0.032	1.449	-0.185
7	AGE AT HEART ATTACK-5/7-{64.1428571, 71.42...}	3.193	1.576	0.078	-3.105	-0.943	-0.743	-0.055
8	AGE AT HEART ATTACK-6/7-{71.4285714, 78.71...}	1.975	0.478	-0.743	-1.269	-0.752	0.225	0.087
9	AGE AT HEART ATTACK-7/7-{78.7142857, 86.00...}	0.487	-0.761	0.128	0.865	0.124	-0.387	-0.457
10	PERICARDIAL EFFUSION-No	-3.724	-1.444	1.601	1.957	1.520	1.156	-1.066
11	PERICARDIAL EFFUSION-Yes	3.949	0.956	-1.487	-1.538	-0.505	-1.550	0.174
12	FRACTIONAL SHORTENING-1/7-{0.0100000, 0.0...}	2.218	-1.902	1.821	0.164	-1.191	0.032	-1.141
13	FRACTIONAL SHORTENING-2/7-{0.0957143, 0.1...}	1.672	-0.693	3.625	-1.996	-1.404	-0.389	-0.816
14	FRACTIONAL SHORTENING-3/7-{0.1814286, 0.2...}	-0.580	3.181	-3.520	0.815	2.450	-1.453	-0.892
15	FRACTIONAL SHORTENING-4/7-{0.2671429, 0.3...}	-3.294	0.337	-2.051	0.029	0.933	1.644	2.402
16	FRACTIONAL SHORTENING-5/7-{0.3528571, 0.4...}	-1.765	-0.888	0.983	1.676	-0.022	0.548	-0.533
17	FRACTIONAL SHORTENING-6/7-{0.4385714, 0.5...}	-0.008	-0.507	0.419	-0.756	0.416	-0.258	0.696
18	FRACTIONAL SHORTENING-7/7-{0.5242857, 0.6...}	-0.252	-0.127	-0.145	0.811	-0.146	-0.065	-0.076
19	EPSS-1/7-{4.0000000, 9.1428571}	-5.328	0.307	-0.375	0.004	1.596	1.611	2.184

2.8. Ценность описательных шкал и градаций для решения задач идентификации и прогнозирования

Для любой из моделей системой «Эйдос» рассчитывается *ценность*⁵ градации описательной шкалы, т.е. признака, для идентификации или прогнозирования. Количественной мерой ценности признака в той или иной модели является вариабельность по классам частного критерия (таблица 2) для этого признака. Мер вариабельности может быть много, но наиболее известными является среднее модулей отклонения от среднего, дисперсия и среднеквадратичное отклонение. Последняя мера и используется в АСК-анализе и системе «Эйдос».

В системе «Эйдос» ценность признаков нарастающим итогом выводится в графической форме.

При большом объеме обучающей выборки можно без ущерба для достоверности модели удалить из нее малозначимые признаки. Для этого в системе «Эйдос» также есть соответствующие инструменты.

2.9. Интегральные критерии системы «Эйдос»

Но если нам известно, что объект обладает не одним, а несколькими признаками, то как посчитать их *общий* вклад в сходство с теми или иными классами? Для этого в системе «Эйдос» используется 2 аддитивных интегральных критерия: «Сумма знаний» и «Семантический резонанс знаний».

Интегральный критерий «Семантический резонанс знаний» представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 3.3:

$$I_j = (\vec{I}_{ij}, \vec{L}_i).$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид:

$$I_j = \sum_{i=1}^M I_{ij} L_i,$$

где: М – количество градаций описательных шкал (признаков);

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j-го класса;

⁵ Эта ценность в АСК-анализе называется также интегральной информативностью, дифференцирующей способностью и селективной силой, т.е. эти термины являются синонимами.

$\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i - \text{й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i - \text{й фактор действует с истинностью } n; \\ 0, & \text{если } i - \text{й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n, если он присутствует у объекта с интенсивностью n, т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

Интегральный критерий «Семантический резонанс знаний» представляет собой *нормированное* суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 3.3 и имеет вид:

$$I_j = \frac{1}{\sigma_I \sigma_L M} \sum_{i=1}^M (I_{ij} - \bar{I}_j) (L_i - \bar{L}),$$

где:

M – количество градаций описательных шкал (признаков);

\bar{I}_j – средняя информативность по вектору класса;

\bar{L} – среднее по вектору объекта;

σ_I – среднее квадратичное отклонение частных критериев знаний вектора класса;

σ_L – среднее квадратичное отклонение по вектору распознаваемого объекта.

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j–го класса;

$\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i - \text{й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i - \text{й фактор действует с истинностью } n; \\ 0, & \text{если } i - \text{й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n, если он присутствует у объекта с интенсивностью n, т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

Приведенное выражение для интегрального критерия «Семантический резонанс знаний» получается непосредственно из выражения для критерия «Сумма знаний» после замены координат перемножаемых векторов их стандартизированными значениями:

$$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, \quad L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}.$$

Свое наименование интегральный критерий сходства «Семантический резонанс знаний» получил потому, что по своей математической форме является корреляцией двух векторов: состояния j-го класса и состояния распознаваемого объекта.

2.10. Результаты верификации моделей

Результаты верификации (оценки достоверности) моделей, отличающихся частными критериями (таблица 2) с двумя приведенными выше интегральными критериями с приведены на рисунке 13:

Наименование модели и частного критерия	Интегральный критерий	Вероятность правильной идентифика...	Вероятность правильной не идентифика...	Средняя вероятность правильного результата	Дата получения результата	Время получения резуль...
AB5 - частный критерий: количество встреч сочетаний: "клас...	Корреляция абс. частот с обр...	100.000	2.708	51.354	31.10.2014	13:3
1. AB5 - частный критерий: количество встреч сочетаний: "клас...	Сумма абс. частот по признак...	100.000	0.735	50.368	31.10.2014	13:3
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн.частот с о...	100.000	2.708	51.354	31.10.2014	13:3
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн.частот по приз...	100.000	0.735	50.368	31.10.2014	13:3
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн.частот с о...	100.000	2.708	51.354	31.10.2014	13:3
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн.частот по приз...	100.000	0.735	50.368	31.10.2014	13:3
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	78.462	77.442	77.952	31.10.2014	13:3
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	88.462	65.261	76.861	31.10.2014	13:3
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	76.923	77.125	77.024	31.10.2014	13:3
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	88.462	65.058	76.760	31.10.2014	13:3
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Семантический резонанс зна...	93.846	56.084	74.965	31.10.2014	13:3
6. INF3 - частный критерий: Хи-квадрат, разности между фактич...	Сумма знаний	93.846	56.084	74.965	31.10.2014	13:3
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	66.923	88.183	77.553	31.10.2014	13:3
7. INF4 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	96.923	47.822	72.372	31.10.2014	13:3
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Семантический резонанс зна...	64.615	87.927	76.271	31.10.2014	13:3
8. INF5 - частный критерий: ROI (Return On Investment); вероятно...	Сумма знаний	96.154	46.626	71.390	31.10.2014	13:3
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Семантический резонанс зна...	93.077	58.060	75.568	31.10.2014	13:3
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер...	Сумма знаний	93.846	46.688	70.267	31.10.2014	13:3
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Семантический резонанс зна...	93.077	57.847	75.462	31.10.2014	13:3
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве...	Сумма знаний	95.385	47.619	71.502	31.10.2014	13:3

Рисунок 13 – Оценки достоверности моделей

Наиболее достоверной в данном приложении оказались модели INF1 при интегральном критерии «Резонанс знаний». Уровень достоверности прогнозирования с применением модели выше, чем экспертных оценок, достоверность которых считается равной около 70%.

Для оценки достоверности моделей в АСК-анализе и системе «Эйдос» используется метрика, сходная с F-критерием (рисунок 14):

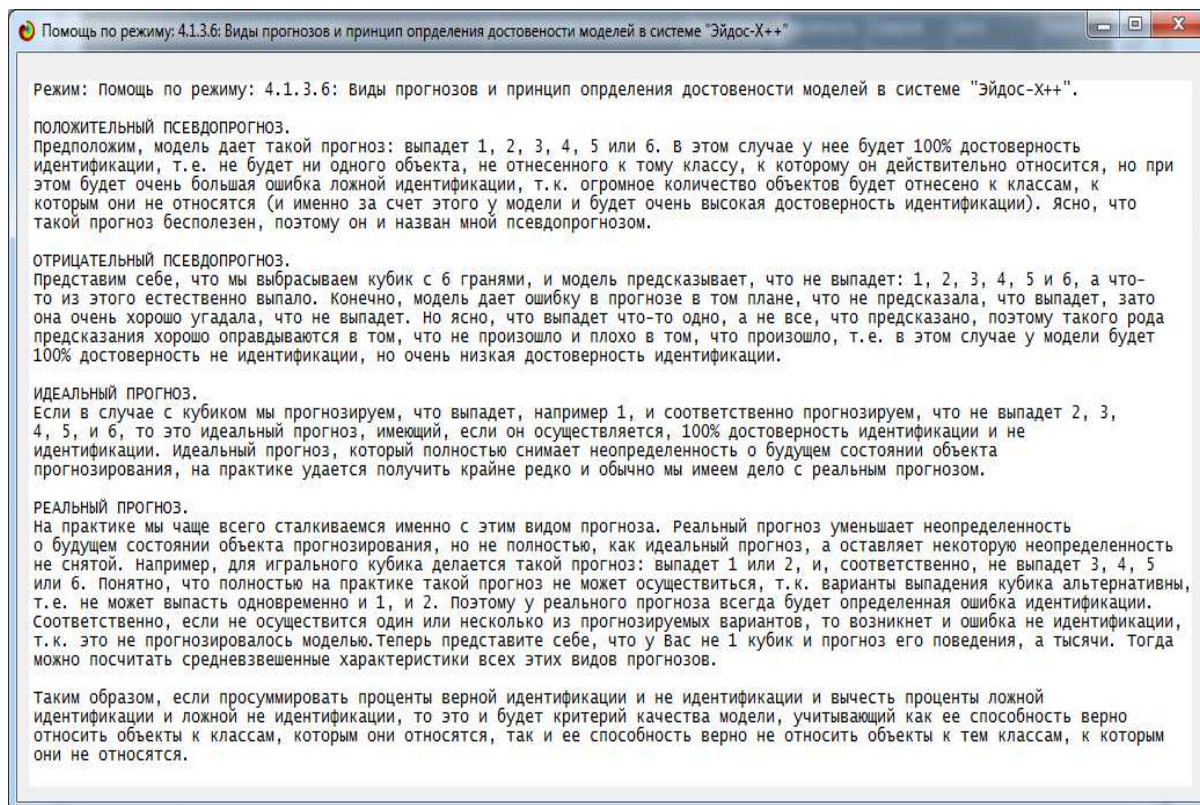


Рисунок 14 – Виды прогнозов и принцип определения достоверности моделей по авторскому варианту метрики, сходной с F-критерием

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.

3. Решение задач идентификации и прогнозирования в наиболее достоверной модели

Режим 4.1.3 обеспечивает отображение результатов идентификации 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.

Рассмотрим некоторые из них.

На рисунке 15 приведен пример прогноза для пациента 1 в наиболее достоверной модели:

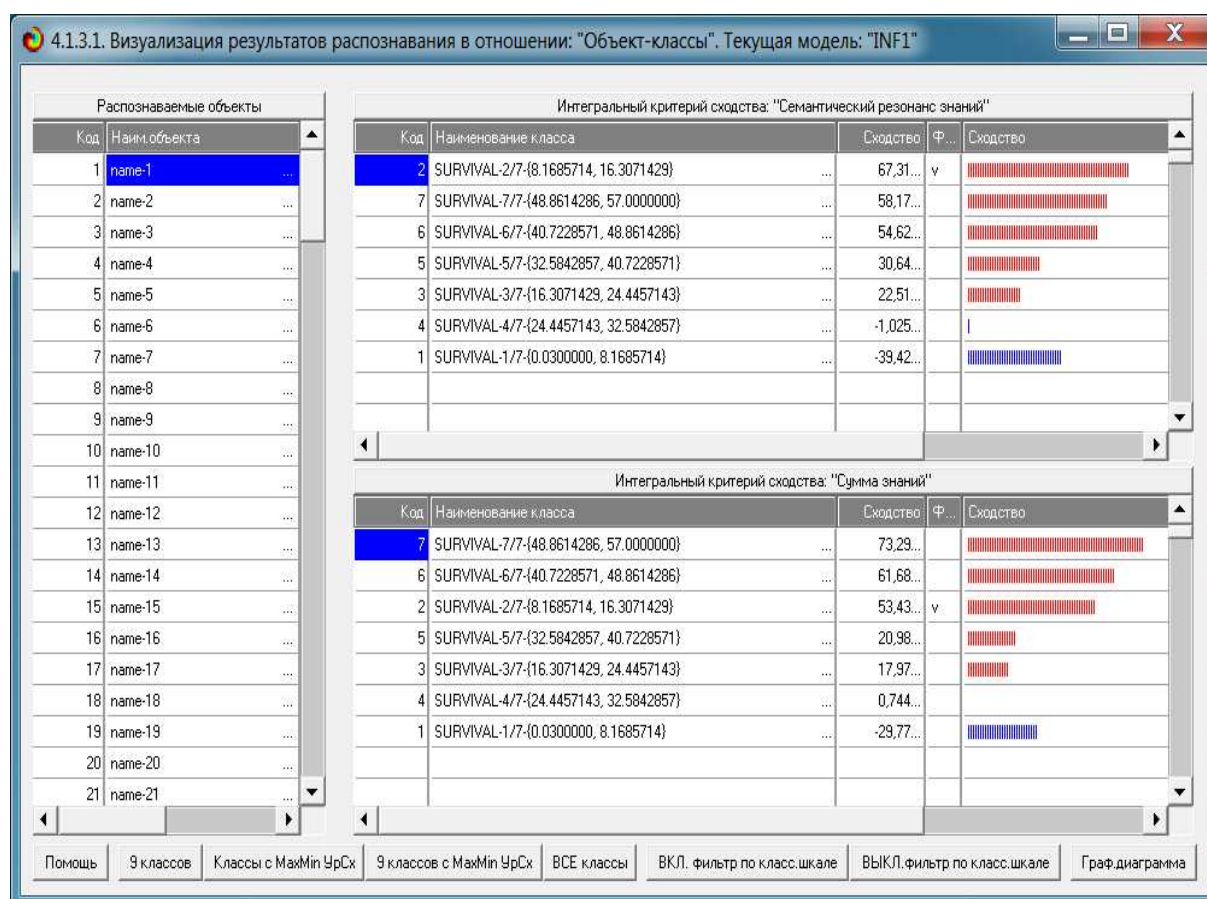


Рисунок 15 – Пример прогноза для пациента 1

На рисунке 16 приведены результаты прогнозирования малой продолжительности жизни после перенесения инфаркта, который, очевидно, прогнозируется с очень высокой достоверностью.

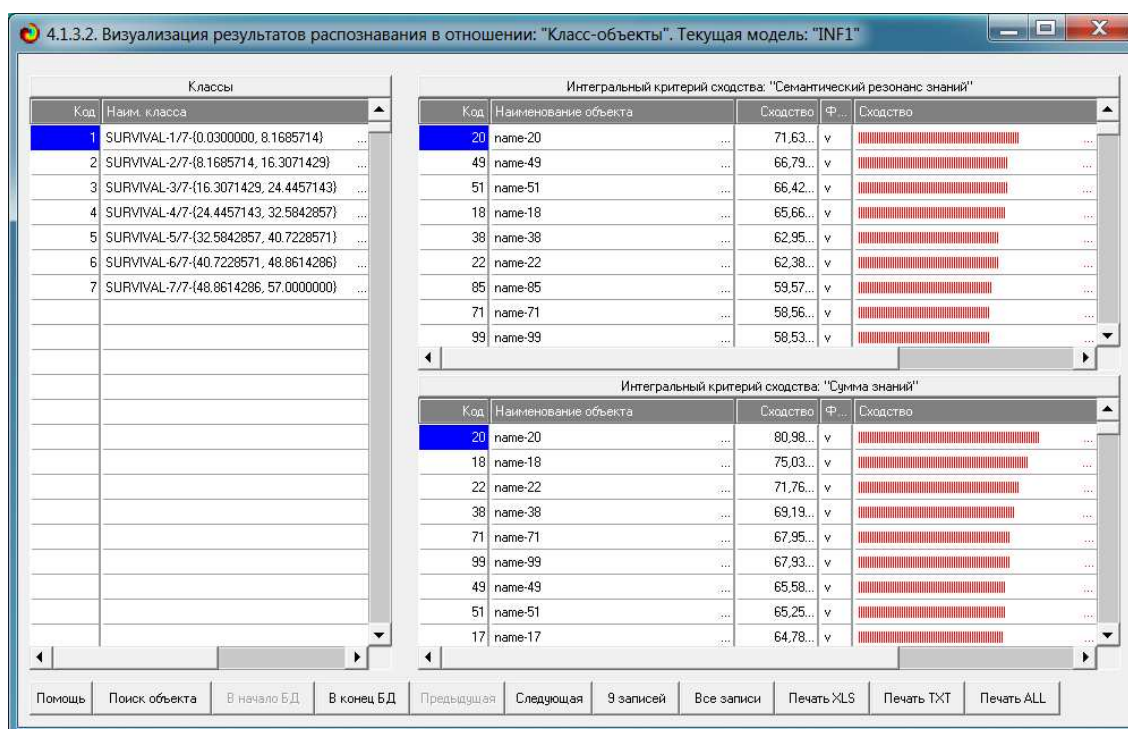


Рисунок 16 – Пример прогнозирования малой продолжительности жизни пациентов после перенесенного инфаркта

На рисунке 17 приведены частные распределения уровней сходства и различия для верно и ошибочно идентифицированных и неидентифицированных объектов в наиболее достоверной модели INF1:

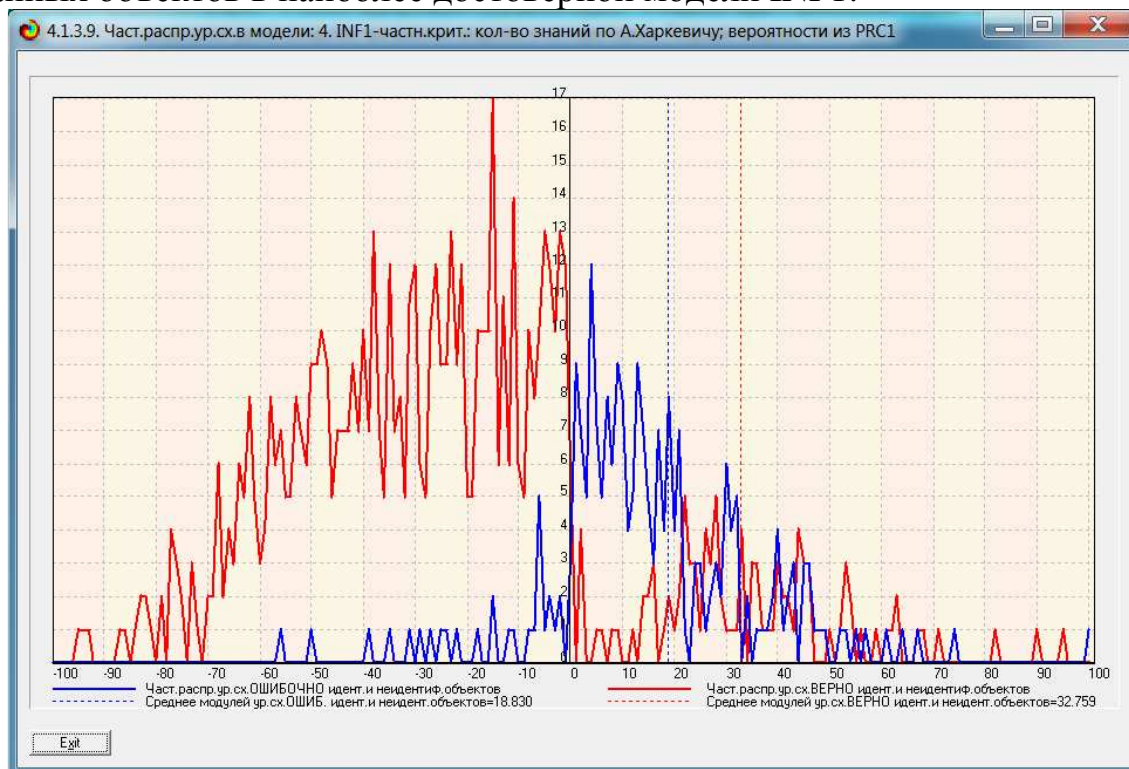


Рисунок 17 – Частное распределение уровня сходства в модели INF1

Из рисунка 17 видно, что:

- наиболее достоверная модель INF1 лучше определяет непринадлежность объекта к классу, чем принадлежность;
- модуль уровня сходства-различия в наиболее достоверной модели для верно идентифицированных и верно неидентифицированных объектов значительно выше, чем для ошибочно идентифицированных и ошибочно неидентифицированных.

4. Исследование моделируемой предметной области путем исследования ее модели

Так созданные модели показали достаточно высокую адекватность, то их исследование обоснованно можно считать исследованием самой моделируемой предметной области.

Это исследование в частности включает:

- информационные портреты классов и признаков;
- когнитивные функции;
- SWOT и PEST матрицы и диаграммы;
- нелокальные нейроны.

4.1. Информационные портреты классов и признаков

Информационный портрет класса – это список факторов, ранжированных в порядке убывания силы их влияния на переход объекта управления в состояние, соответствующее данному классу. Информационный портрет класса отражает систему его детерминации. Генерация информационного портрета класса представляет собой решение обратной задачи прогнозирования, т.к. при прогнозировании по системе факторов определяется спектр наиболее вероятных будущих состояний объекта управления, в которые он может перейти под влиянием данной системы факторов, а в информационном портрете мы наоборот, по заданному будущему состоянию объекта управления определяем систему факторов, детерминирующих это состояние, т.е. вызывающих переход объекта управления в это состояние. В начале информационного портрета класса идут факторы, оказывающие положительное влияние на переход объекта управления в заданное состояние, затем факторы, не оказывающие на это существенного влияния, и далее – факторы, препятствующие переходу объекта управления в это состояние (в порядке возрастания силы препятствования). Информационные портреты классов могут быть *отфильтрованы* по диапазону факторов, т.е. мы можем отобразить влияние на переход объекта управления в данное состояние не всех отраженных в модели факторов, а

только тех, коды которых попадают в определенный диапазон, например, относящиеся к определенным описательным шкалам.

Информационный (семантический) портрет фактора – это список классов, ранжированный в порядке убывания силы влияния данного фактора на переход объекта управления в состояния, соответствующие данным классам. Информационный портрет фактора называется также его *семантическим портретом*, т.к. в соответствии с концепцией смысла системно-когнитивного анализа, являющейся обобщением концепции смысла Шенка-Абельсона, *смысл фактора состоит в том, какие будущие состояния объекта управления он детерминирует*. Сначала в этом списке идут состояния объекта управления, на переход в которые данный фактор оказывает наибольшее влияние, затем состояния, на которые данный фактор не оказывает существенного влияния, и далее состояния – переходу в которые данный фактор препятствует. Информационные портреты факторов могут быть *отфильтрованы* по диапазону классов, т.е. мы можем отобразить влияние данного фактора на переход объекта управления не во все возможные будущие состояния, а только в состояния, коды которых попадают в определенный диапазон, например, относящиеся к определенным классификационным шкалам.

Рассмотрим информационные портреты классов в режиме 4.2.1 (рисунок 18):

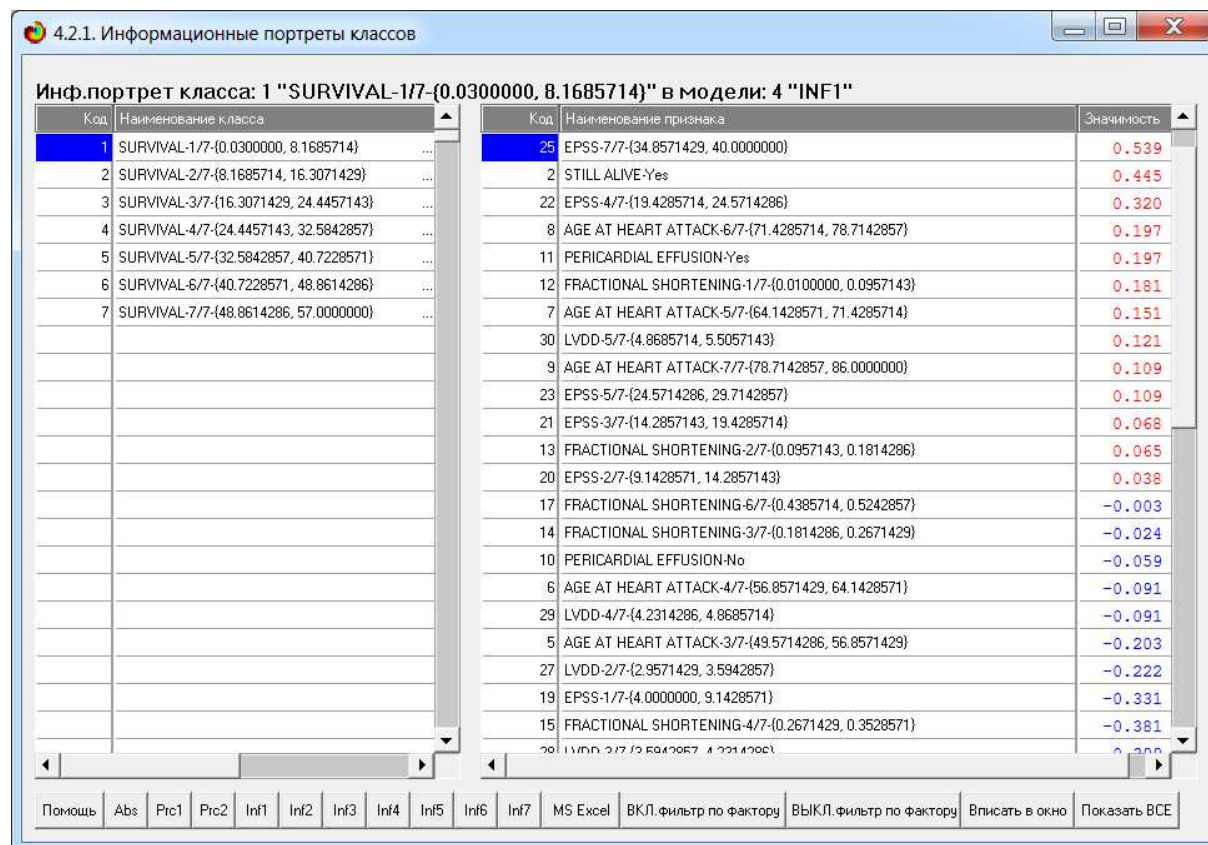


Рисунок 18 – Пример информационного портрета класса

Из этого портрета мы можем сделать вывод о том, что на основе *определений конкретных* объектов обучающей выборки (т.е. конкретных онтологий), в каждом из которых были с одной стороны приведены признаки объекта, а с другой указана его принадлежность к обобщающим классам, система «Эйдос» смогла сформировать *обобщенные определения* классов (обобщенные онтологии), в которых для каждого класса указана степень характерности и не характерности для него всех имеющихся в модели признаков. В данном случае в качестве объектов обучающей выборки выступали описания пациентов, перенесших инфаркт, но в принципе ими могут быть любые объекты или их состояния из любой предметной области (например, [9]⁶), т.к. *система «Эйдос» разрабатывалась в обобщенной постановке, независимой от предметной области.*

Частные критерии:

- ABS – частный критерий: количество встреч сочетаний: "класс-признак" у объектов обучающей выборки;
- PRC1 – частный критерий: условная вероятность *i*-го признака среди признаков объектов *j*-го класса;
- PRC2 – частный критерий: условная вероятность *i*-го признака у объектов *j*-го класса;
- INF1 – частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1;
- INF2 – частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2;
- INF3 – частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсолютными частотами;
- INF 4 – частный критерий: ROI (Return on investment); вероятности из PRC1;
- INF 5 – частный критерий: ROI (Return on investment); вероятности из PRC2;
- INF6 – частный критерий: разности условной и безусловной вероятностей; вероятности из PRC1;
- INF7 - частный критерий: разности условной и безусловной вероятностей; вероятности из PRC2.

Информационные портреты классов могут быть отфильтрованы по одной из описательных шкал, т.е. по одному из факторов, в этом случае информационный портрет будет отражать влияние на переход объекта управления в состояние, соответствующее классу, не всех отраженных в модели факторов, а значений (т.е. градаций описательных шкал или признаков) одного из них:

⁶ Система «Эйдос» была применена во многих предметных областях:
<http://lc.kubagro.ru/> <http://ej.kubagro.ru/a/viewaut.asp?id=11>

- выбор класса осуществляется установкой курсора на записи с нужным классом в левом окне;
- выбор статической модели или модели знаний осуществляется нажатием соответствующей кнопки внизу окна;
- выбор фактора (описательной шкалы) для фильтрации осуществляется путем указания его курсором в правом окне и последующего нажатия кнопки "вкл. фильтр по фактору".

Рассмотрим информационные или семантические (смысловые) портреты признаков в режиме 4.3.1. Данный режим показывает *какой смысл* имеет тот или иной признак, т.е., например, для модели INF1, какое количество информации содержится в факте наблюдения данного признака у некоторого объекта о том, что этот объект принадлежит различным классам (рисунок 19).

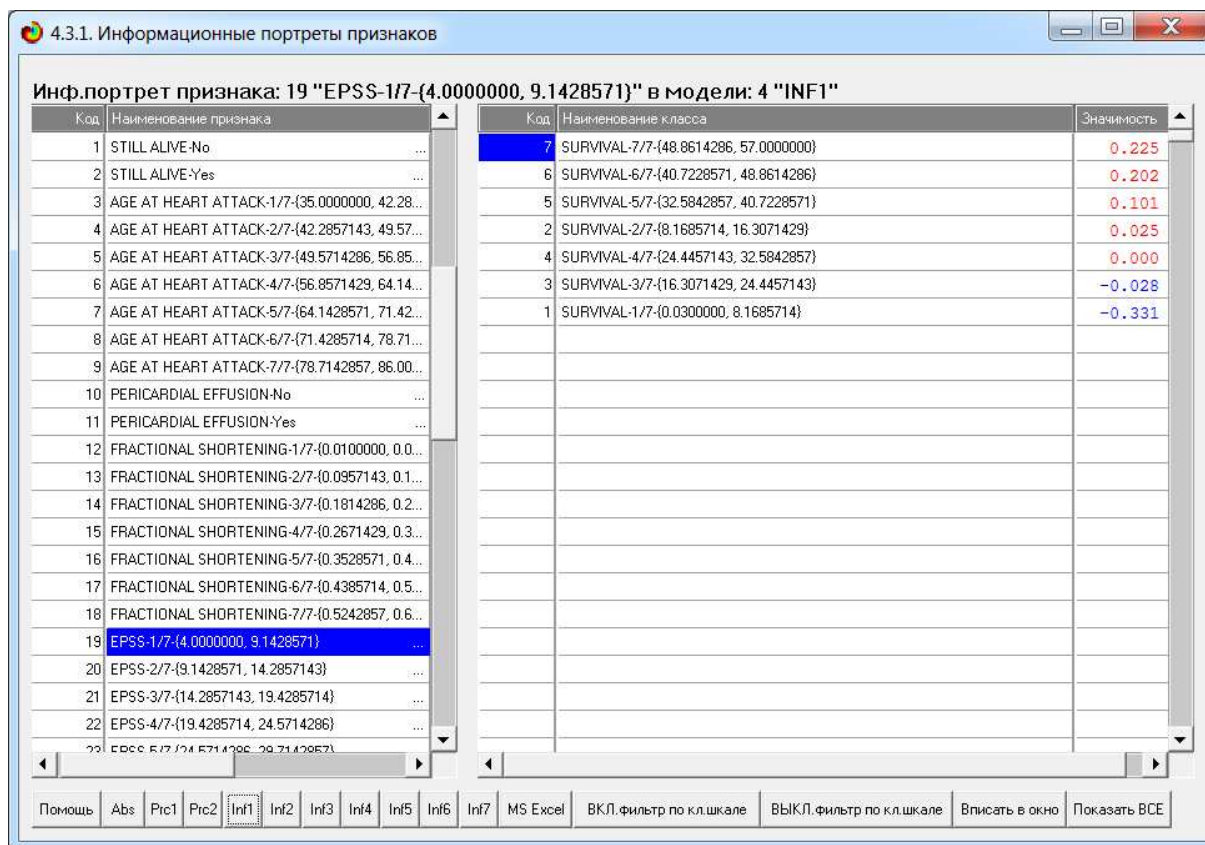


Рисунок 19 – Пример информационного (семантического) портрета признака

4.2. Когнитивные функции

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

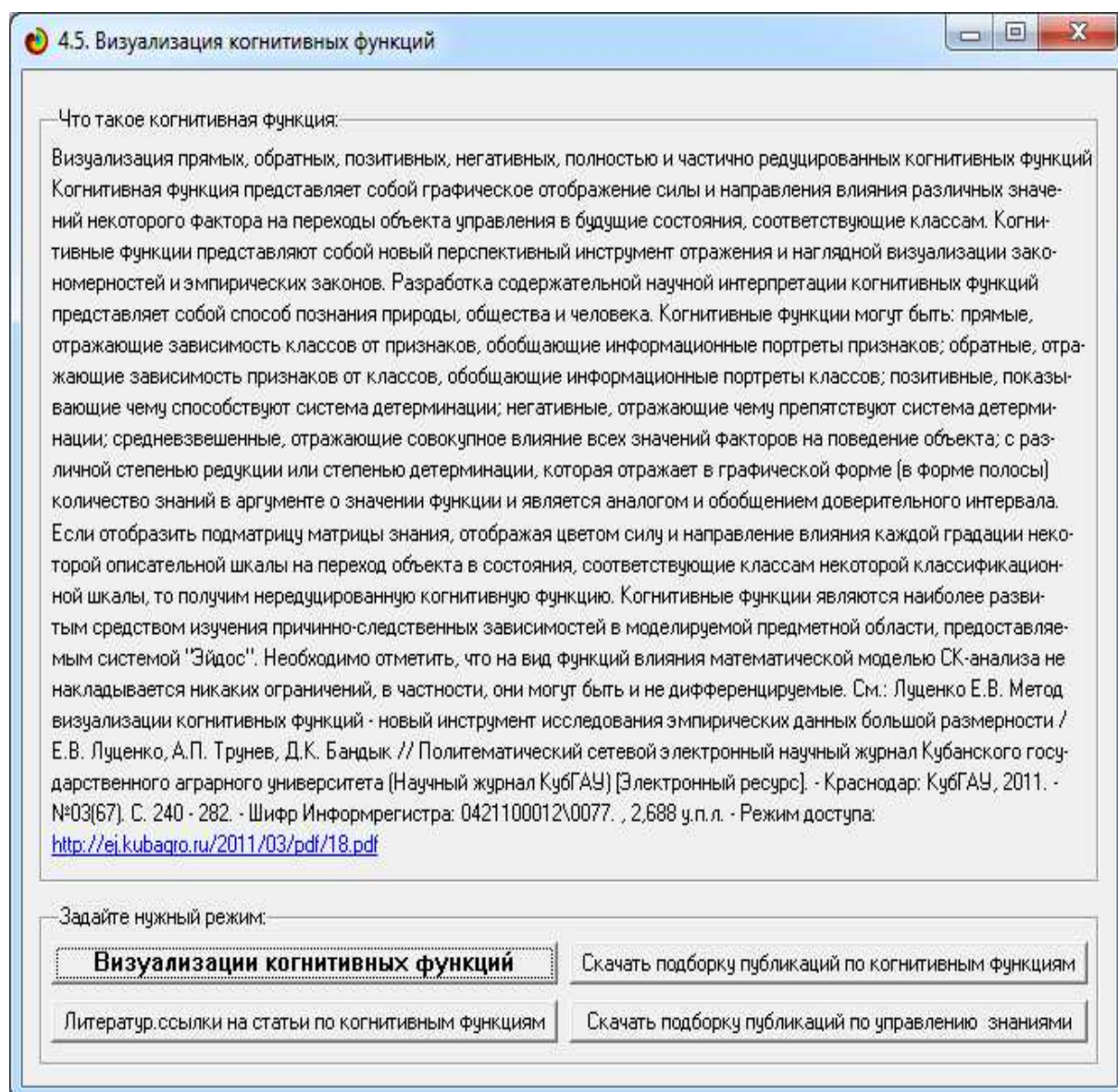
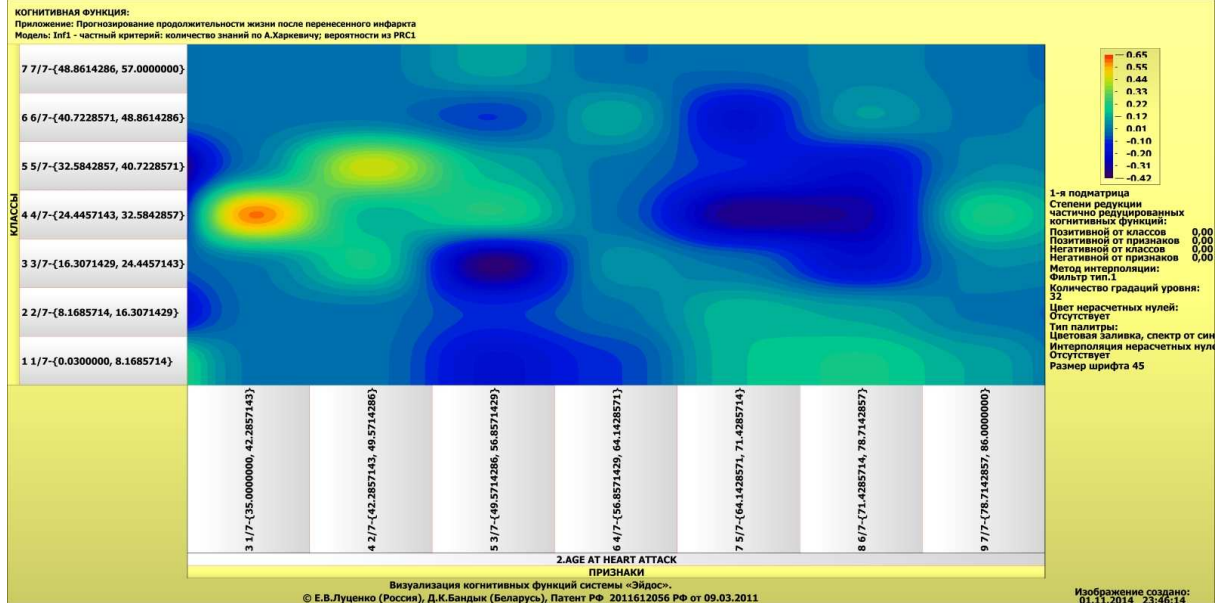
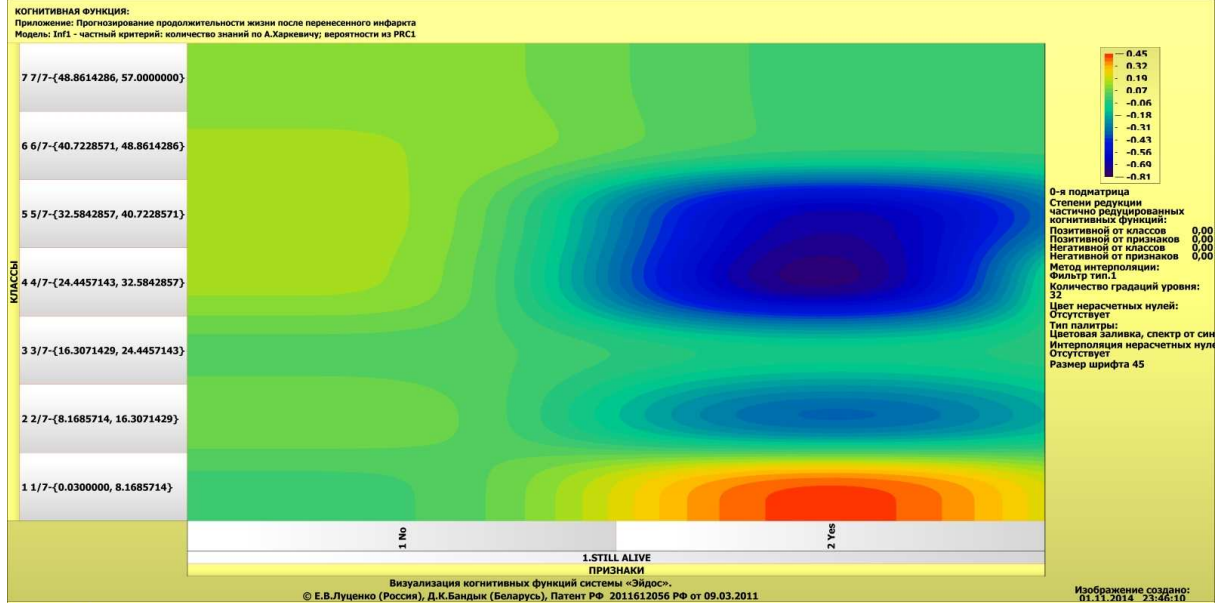
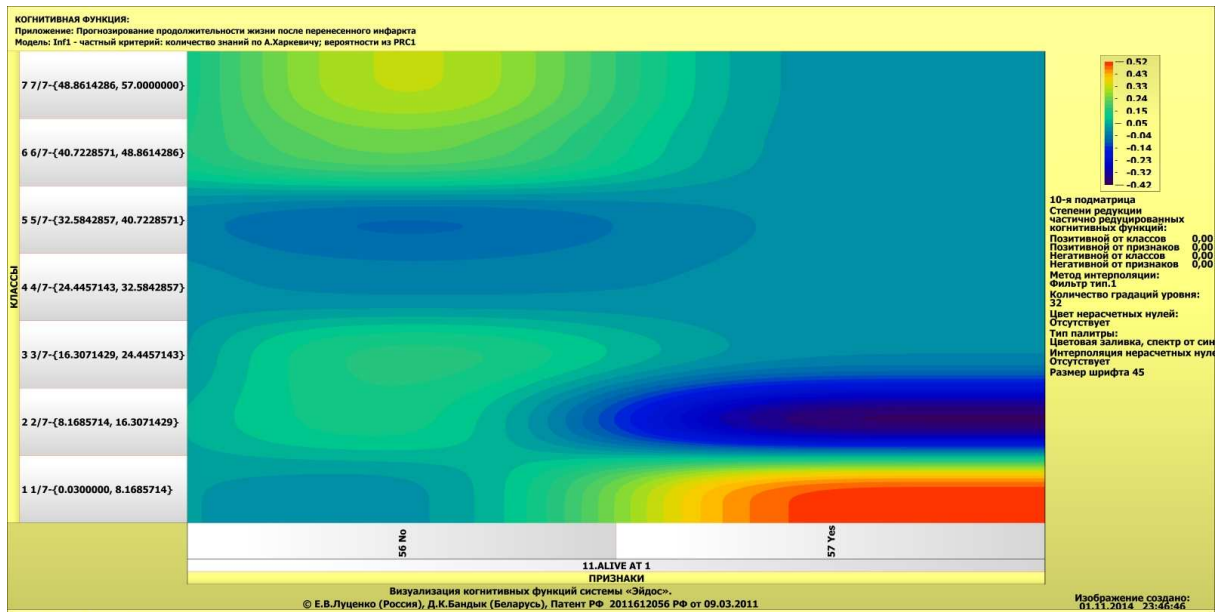


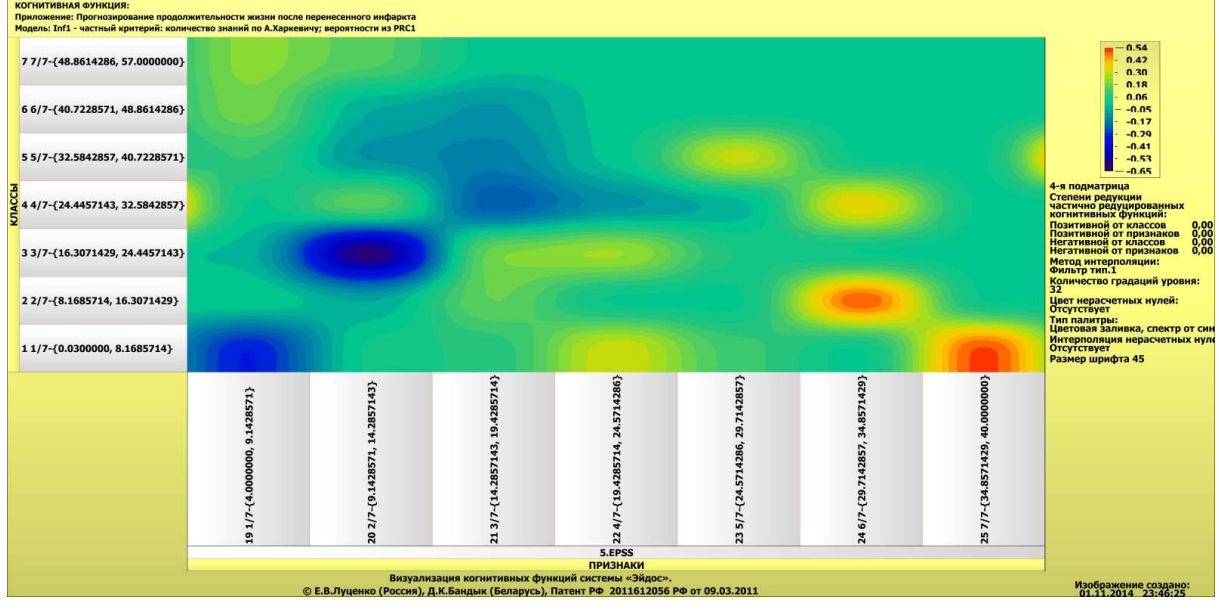
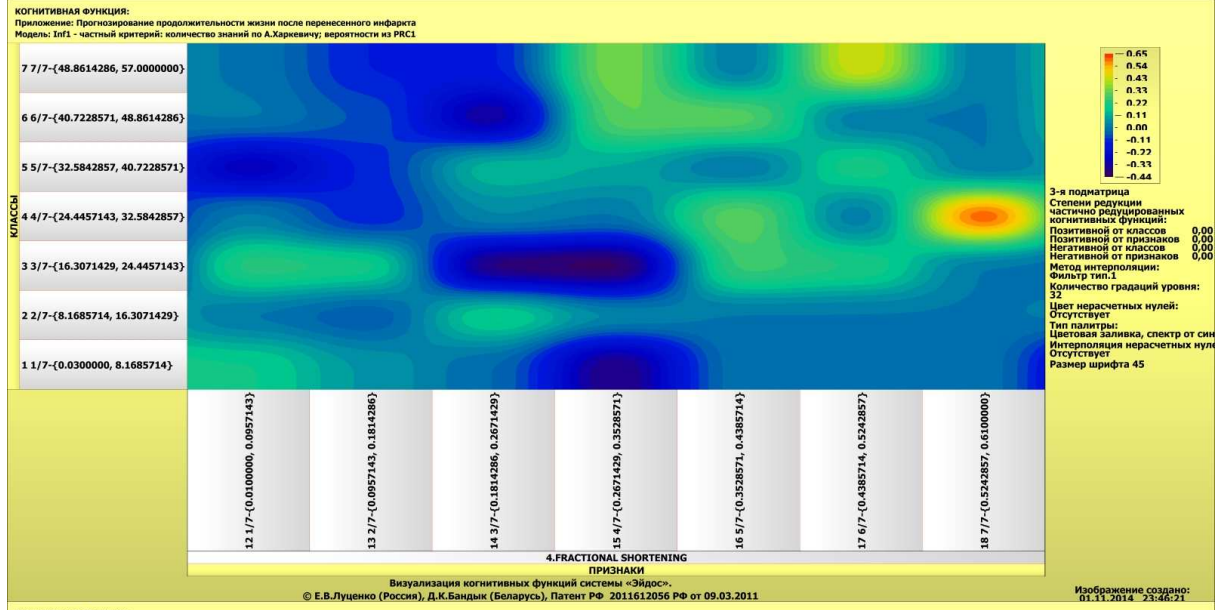
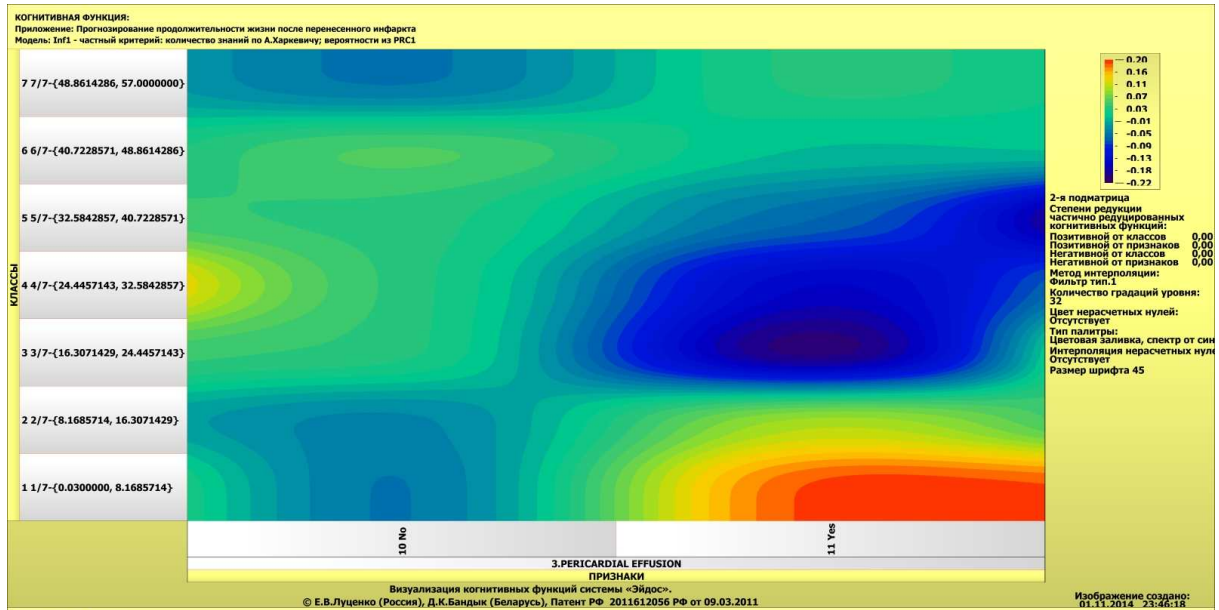
Рисунок 20 – Экранная форма режима 4.5 системы «Эйдос-X++» «Визуализация когнитивных функций»

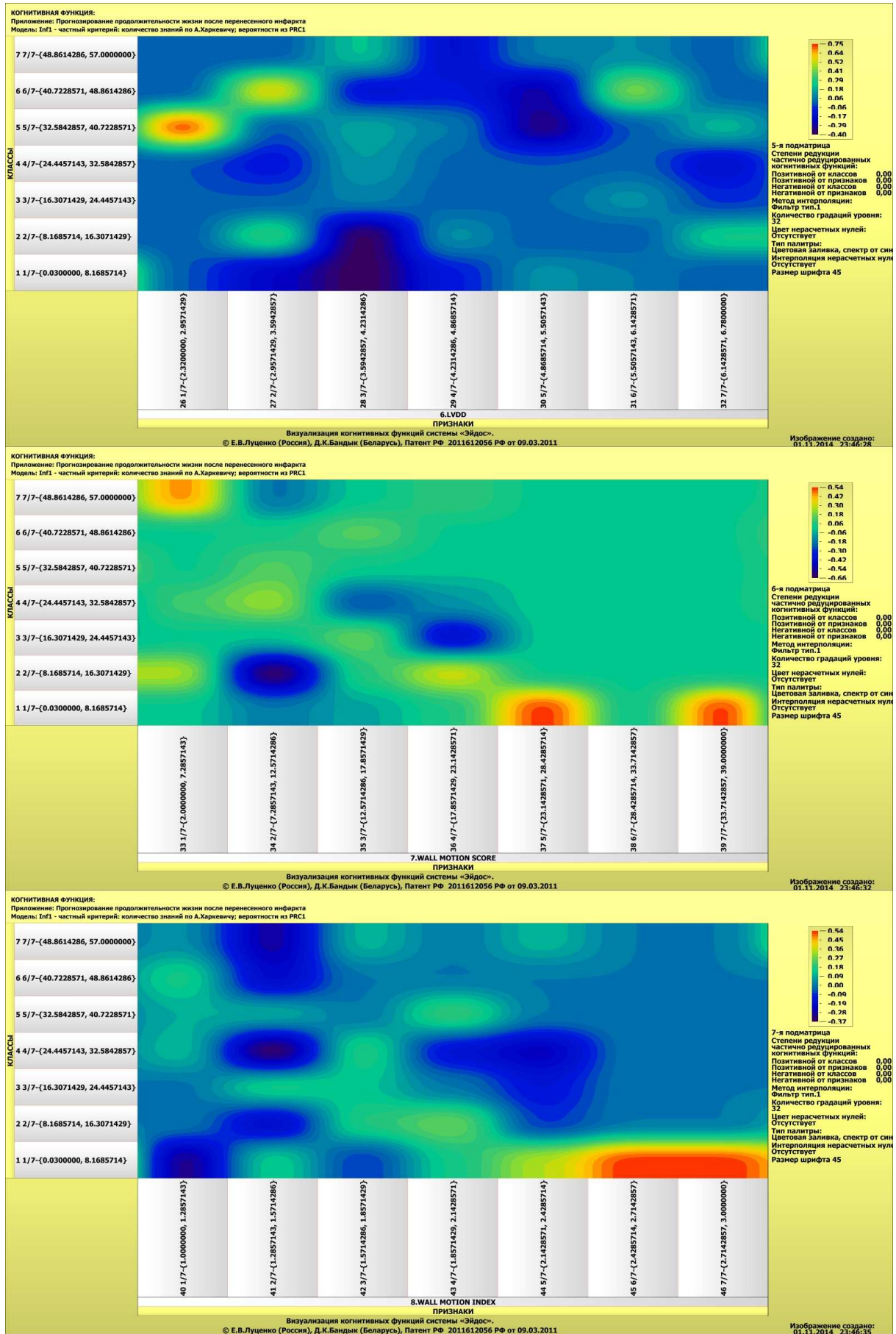
Когнитивным функциям посвящено много работ автора⁷, но наиболее новой и обобщающей из них является работа [16]. Поэтому здесь не будем останавливаться на описании того, что представляют собой когнитивные функции в АСК-анализе.

Для классификационной шкалы «Продолжительность жизни после инфаркта» (это обобщенные классы) когнитивные функции и всех описательных шкал для модели INF1 приведены на рисунке 21:

⁷ См., например: <http://www.twirpx.com/file/775236/>







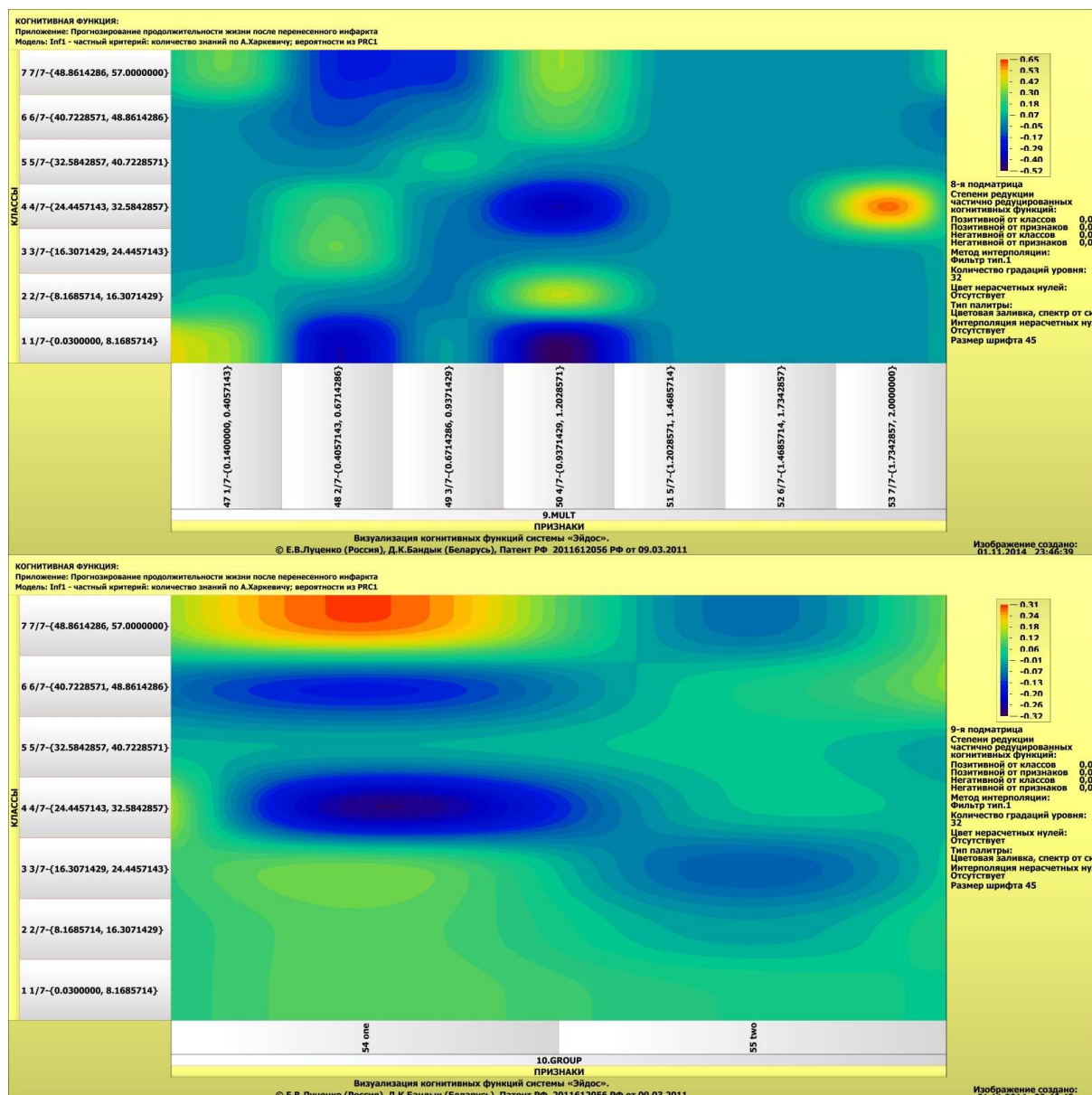


Рисунок 21 – Визуализация когнитивных функций для обобщенных классов и всех описательных шкал для модели INF1

4.3. 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 [13] (рисунок 22):

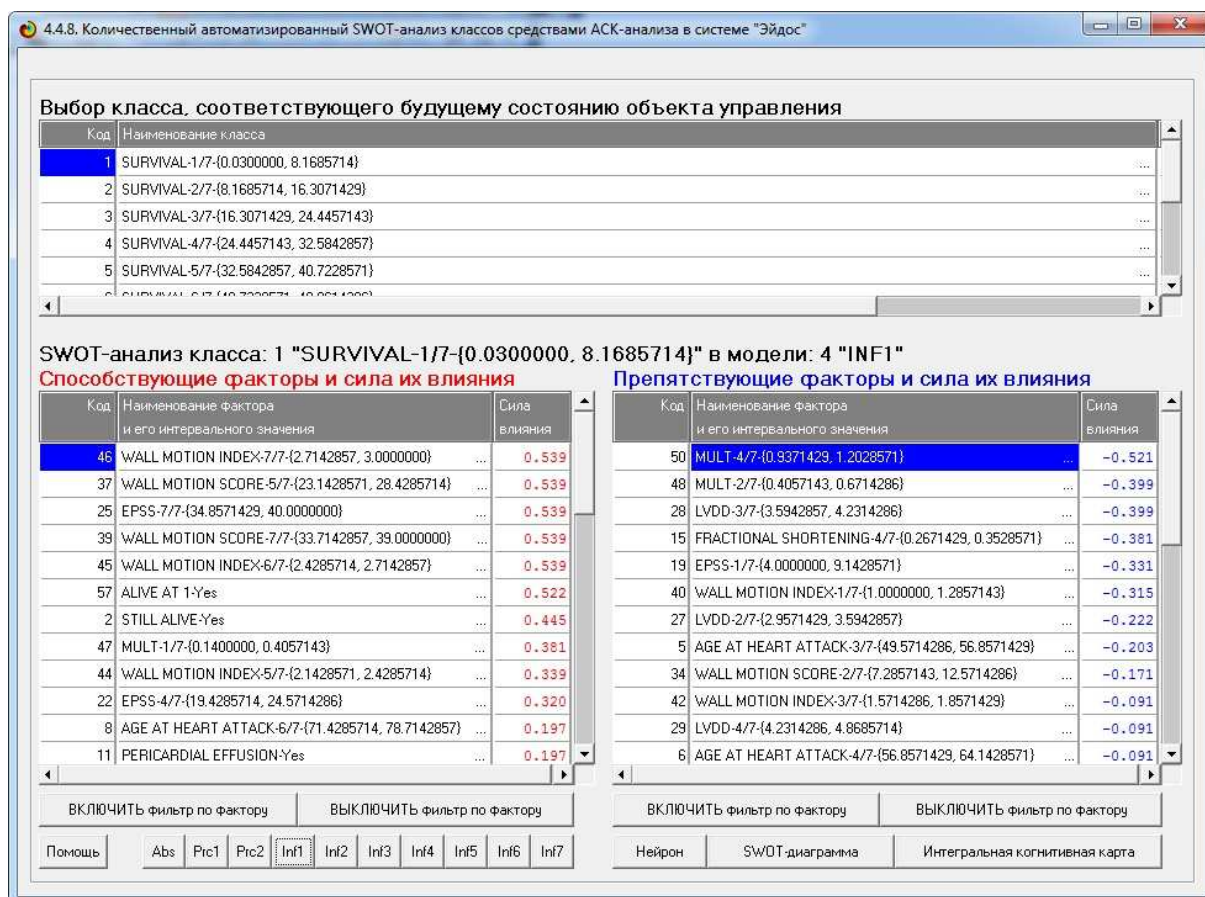


Рисунок 22 – Пример SWOT-матрицы в модели INF1

На рисунке 23 приведена графическая SWOT-диаграмма, соответствующая SWOT-матрице, представленной на рисунке 22.

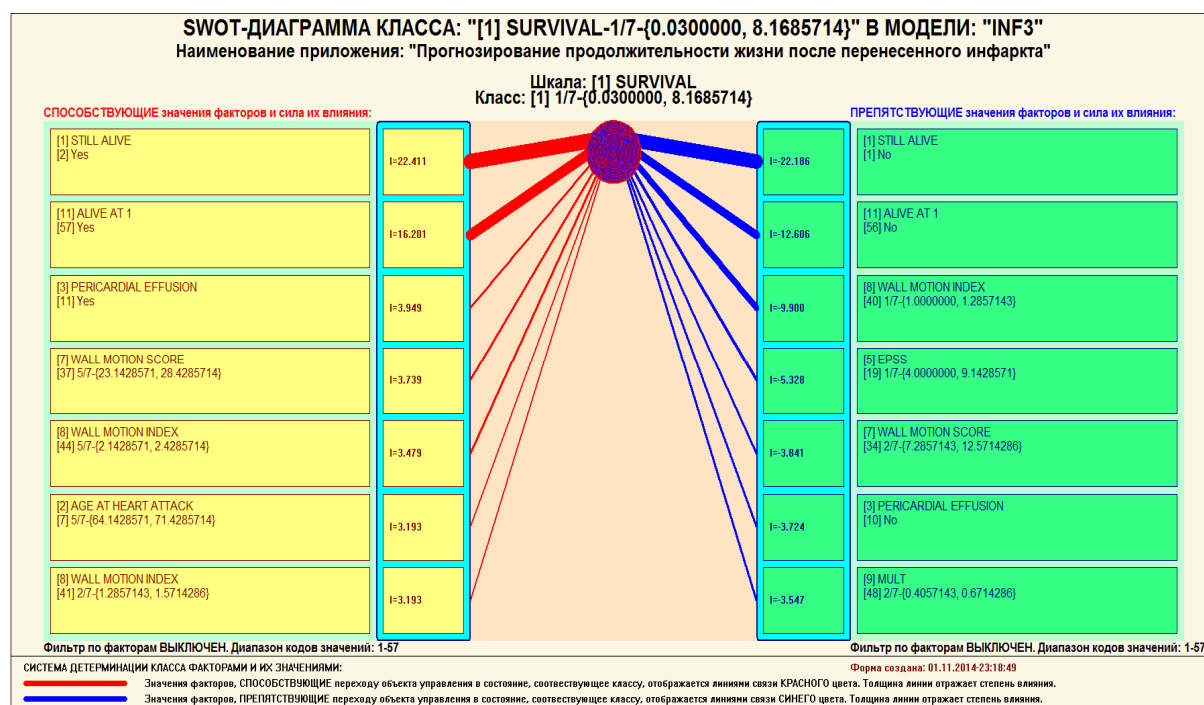


Рисунок 23 – Пример SWOT-диаграммы в модели INF1

PEST-анализ является SWOT-анализом с более детализированной классификацией внешних факторов. Поэтому выводы, полученные на примере SWOT-анализа, можно распространить и на PEST-анализ.

4.4. Нелокальные нейроны

Нелокальные нейроны и интерпретируемые нейронные сети позволяют в наглядной форме отобразить систему детерминации будущих состояний. *Нелокальный нейрон* представляет собой будущее состояние объекта управления с изображением наиболее сильно влияющих на него факторов с указанием силы и направления (способствует-препятствует) их влияния. Нейронная сеть представляет собой совокупность взаимосвязанных нейронов. В классических нейронных сетях связь между нейронами осуществляется по входным и выходным сигналам, а в нелокальных нейронных сетях – на основе общего информационного поля, реализуемого семантической информационной моделью. Система «Эйдос» обеспечивает построение любого подмножества многослойной нейронной сети с заданными или выбираемыми по заданным критериям рецепторами и нейронами, связанными друг с другом связями любого уровня опосредованности [14].

На рисунке 24 представлен пример нелокального нейрона в модели INF1:

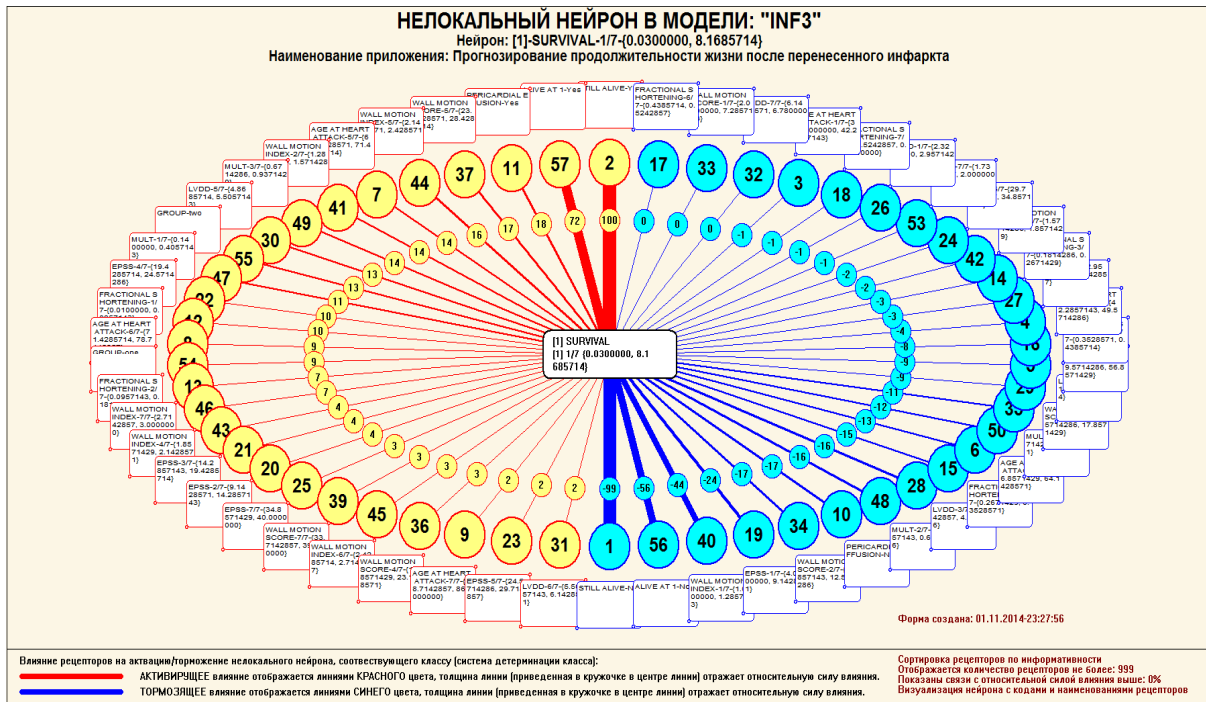


Рисунок 24 – Пример нелокального нейрона в модели INF1

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 INF1, основанная на семантической мере целесообразности информации А.Харкевича при интегральном критерии «Резонанс знаний». Уровень достоверности прогнозирования принадлежности объекта к классу с применением данной составляет 78,5%, а достоверность прогнозирования не-принадлежности объекта к классу – 77,4%, 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 27% 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.

Ясно, что если на основе базы данных UCI, рассмотренной в данной статье, 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: 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: КубГАУ, 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: КубГАУ, 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 " Echocardiogram Data Set". Режим доступа: <http://archive.ics.uci.edu/ml/datasets/Echocardiogram>, 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. 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.

10. Луценко Е.В. Методологические аспекты выявления, представления и использования знаний в АСК-анализе и интеллектуальной системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №06(070). С. 233 – 280. – Шифр Информрегистра: 0421100012\0197, IDA [article ID]: 0701106018. – Режим доступа: <http://ej.kubagro.ru/2011/06/pdf/18.pdf>, 3 у.п.л.

11. 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.

12. Луценко Е.В. Системно-когнитивный анализ как развитие концепции смысла Шенка – Абельсона / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №03(005). С. 44 – 65. – IDA [article ID]: 0050403004. – Режим доступа: <http://ej.kubagro.ru/2004/03/pdf/04.pdf>, 1,375 у.п.л.

13. Луценко Е.В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы «Эйдос-Х++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 1368 – 1410. – IDA [article ID]: 1011407090. – Режим доступа: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 у.п.л.

14. 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.

15. Луценко Е.В. Автоматизированная система распознавания образов, математическая модель и опыт применения. В сб.: "В.И.Вернадский и современность (к 130-летию со дня рождения)". Тезисы научно-практической конференции. – Краснодар: КНА, 1993. – С. 37-42.

16. Луценко Е.В. Когнитивные функции как обобщение классического понятия функциональной зависимости на основе теории информации в системной нечеткой интервальной математике / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. –

№01(095). С. 122 – 183. – IDA [article ID]: 0951401007. – Режим доступа: <http://ej.kubagro.ru/2014/01/pdf/07.pdf>, 3,875 у.п.л.

References

1. Lucenko E.V. Metodika ispol'zovaniya repozitarija UCI dlja ocenki kachestva matematicheskikh modelej sistem iskusstvennogo intellekta / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: 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 setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 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 sopostavimoj ocenki jeffektivnosti vuzov / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 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 primenenie v issledovanii jekonomiceskikh, social'no-psihologicheskikh, tehnologicheskikh i organizacionno-tehnicheskikh sistem): Monografija (nauchnoe izdanie). – Krasnodar: KubGAU. 2002. – 605 s.

5. Repozitarij UCI [Jelektronnyj resurs]. Stat'ja " Echocardiogram Data Set". Rezhim dostupa: <http://archive.ics.uci.edu/ml/datasets/Echocardiogram>, 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 otechestvennykh universal'nyh sistem iskusstvennogo intellekta, shiroko primenjaemyh i razvivajushhihsja i v nastojashhee vremja / E.V. Lucenko // Politematicheskij setевой 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 setевой 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. Metrizacija izmeritel'nyh shkal razlichnyh tipov i sovmestnaja sopostavimaja kolichestvennaja obrabotka raznorodnyh faktorov v sistemno-kognitivnom analize i sisteme «Jejdos» / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU)

[Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №08(092). S. 859 – 883. – IDA [article ID]: 0921308058. – Rezhim dostupa: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 u.p.l.

10. Lucenko E.V. Metodologicheskie aspekty vyjavlenija, predstavlenija i ispol'zovanija znaniy v ASK-analize i intellektual'noj sisteme «Jejdos» / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2011. – №06(070). S. 233 – 280. – Shifr Informregistra: 0421100012(0197, IDA [article ID]: 0701106018. – Rezhim dostupa: <http://ej.kubagro.ru/2011/06/pdf/18.pdf>, 3 u.p.l.

11. Lucenko E.V. Modelirovanie slozhnyh mnogofaktornyh nelinejnyh ob#ektov upravlenija na osnove fragmentirovannyh zashumlennyh jempiricheskikh dannyh bol'shoj razmernosti v sistemno-kognitivnom analize i intellektual'noj sisteme «Jejdos-H++» / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №07(091). S. 164 – 188. – IDA [article ID]: 0911307012. – Rezhim dostupa: <http://ej.kubagro.ru/2013/07/pdf/12.pdf>, 1,562 u.p.l.

12. Lucenko E.V. Sistemno-kognitivnyj analiz kak razvitie koncepcii smysla Shenka – Abel'sona / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2004. – №03(005). S. 44 – 65. – IDA [article ID]: 0050403004. – Rezhim dostupa: <http://ej.kubagro.ru/2004/03/pdf/04.pdf>, 1,375 u.p.l.

13. Lucenko E.V. Kolichestvennyj avtomatizirovannyj SWOT- i PEST-analiz sredstvami ASK-analiza i intellektual'noj sistemy «Jejdos-H++» / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №07(101). S. 1368 – 1410. – IDA [article ID]: 1011407090. – Rezhim dostupa: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 u.p.l.

14. Lucenko E.V. Sistemnaja teorija informacii i nelokal'nye interpretiruemye nejronnye seti prjamogo scheta / E.V. Lucenko // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2003. – №01(001). S. 79 – 91. – IDA [article ID]: 0010301011. – Rezhim dostupa: <http://ej.kubagro.ru/2003/01/pdf/11.pdf>, 0,812 u.p.l.

15. Lucenko E.V. Avtomatizirovannaja sistema raspoznavanija obrazov, matematicheskaja model' i opyt primenenija. V sb.: "V.I.Vernadskij i sovremennost' (k 130-letiju so dnja rozhdenija)". Tezisy nauchno-prakticheskoy konferencii. – Krasnodar: KNA, 1993. – S. 37-42.

16. Lucenko E.V. Kognitivnye funkcii kak obobshhenie klassicheskogo ponjatija funkcional'noj zavisimosti na osnove teorii informacii v sistemnoj nechetkoj interval'noj matematike / E.V. Lucenko, A.I. Orlov // Politematicheskij setевой jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №01(095). S. 122 – 183. – IDA [article ID]: 0951401007. – Rezhim dostupa: <http://ej.kubagro.ru/2014/01/pdf/07.pdf>, 3,875 u.p.l.