

УДК 004.8

08.00.13 - Математические и инструментальные методы экономики (экономические науки)

ПРОГНОЗИРОВАНИЕ РИСКОВ НЕВОЗВРАТА ССУДЫ С ПРИМЕНЕНИЕМ ИНТЕЛЛЕКТУАЛЬНОГО ИТЕРАЦИОННОГО АЛГОРИТМА УЧЕТА НЕТИПИЧНЫХ СЛУЧАЕВ

Луценко Евгений Вениаминович
 д.э.н., к.т.н., профессор
 Web of Science ResearcherID S-8667-2018
 Scopus Author ID: 57188763047
 РИНЦ id=123162, SPIN-code: 9523-7101
prof.lutsenko@gmail.com <http://lc.kubagro.ru>
<https://www.researchgate.net/profile/Eugene-Lutsenko>

В данной работе ставится, рассматривается и решается актуальная задача достоверного прогнозирования рисков невозврата ссуды. От успешности решения этой задачи самым непосредственным образом зависят прибыльность, успешность и устойчивость субсидирующих организаций. Предлагается теоретическое и практическое решение этой задачи путем применения интеллектуального итерационного алгоритма учета нетипичных случаев субсидирования, которые обычно и приводят к понижению достоверности прогнозирования. Предлагаемое решение основано на применении автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария – интеллектуальной системы «Эйдос». Приводится подробный численный пример, основанный на данных портала Каггл

Ключевые слова: АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, АСК-АНАЛИЗ, СИСТЕМА «ЭЙДОС»

DOI: <http://dx.doi.org/10.21515/1990-4665-170-009>

Введение

В данной работе ставится, подробно рассматривается и решается актуальная задача достоверного прогнозирования рисков невозврата ссуды. От успешности решения этой задачи самым непосредственным образом зависят прибыльность, успешность и устойчивость субсидирующих организаций. Предлагается теоретическое и практическое решение этой задачи путем применения интеллектуального итерационного алгоритма учета нетипичных случаев субсидирования, которые обычно и приводят к понижению достоверности прогнозирования. Предлагаемое решение основано на применении методов, ссылка на которые дается во введении работы [35].

UDC 004.8

08.00.13 - Mathematical and instrumental methods of Economics (Economics)

FORECASTING THE RISKS OF LOAN NON-REPAYMENT USING AN INTELLIGENT ITERATIVE ALGORITHM FOR ACCOUNTING FOR ATYPICAL CASES

Lutsenko Evgeniy Veniaminovich
 Dr.Sci.Econ., Cand.Tech.Sci., professor
 Web of Science ResearcherID S-8667-2018
 Scopus Author ID: 57188763047
 RSCI id=123162, SPIN-code: 9523-7101
prof.lutsenko@gmail.com <http://lc.kubagro.ru>
<https://www.researchgate.net/profile/Eugene-Lutsenko>

In this work we set, examine and solve the actual problem of reliable forecasting of the risks of non-repayment of a loan. The profitability, success, and sustainability of subsidizing organizations directly depend on the success of this task. We have also proposed a theoretical and practical solution to this problem is by applying an intelligent iterative algorithm for accounting for atypical cases of subsidies, which usually lead to a decrease in the reliability of forecasting. The proposed solution is based on the use of automated system-cognitive analysis (ASC-analysis) and its software tools which is "Eidos", the intelligent system. We also provide a detailed numerical example based on data from the Kaggl portal

Keywords: AUTOMATED SYSTEM-COGNITIVE ANALYSIS, ASC-ANALYSIS, "EIDOS" SYSTEM

1. Проблематика работы

1.1. Описание предметной области, объект, предмет, проблема, цель и задачи работы

Задача прогнозирования рисков невозврата финансовых средств (ссуд, кредитов, инвестиций)¹ физическими и юридическими лицами является весьма актуальной для всех, систематически занимающихся этим видом деятельности. Если достоверно прогнозировать риск невозврата ссуды и предоставлять ее только тогда, когда риск ее невозврата мал и не предоставлять ее, когда он велик, то прибыль от субсидирования будет высока². Понятно, что чем ниже достоверность прогнозирования риска невозврата ссуды, тем ниже прибыльность от субсидирования. Более того, при высоком риске невозврата субсидирование может привести к убыточности от этого вида деятельности или даже к банкротству («неоправданное кредитование»)³.

Поэтому прибыль от субсидирования самым непосредственным образом зависит от достоверности прогнозирования риска невозврата ссуды.

Из этого следует объект и предмет исследования, а также проблема, решаемая в работе, ее цель и задачи.

Объект исследования – прогнозирование риска невозврата ссуды.

Предмет исследования – прогнозирование рисков невозврата ссуды с применением интеллектуального итерационного алгоритма учета нетипичных случаев.

Проблема – возникающая при прогнозировании риска невозврата ссуды – ложно-отрицательные решения, возникающие, когда модель не относит ситуацию к классу (обобщенной категории), к которой она в действительности относится. Этот вид ошибок возникает в нетипичных ситуациях, резко снижает достоверность прогнозирования и приводит к ошибочным рекомендациям о выдаче или невыдаче ссуды, противоположным адекватным. *Фактически* стандартные алгоритмы прогнозирования основаны на обобщении типовых ситуаций и, поэтому, нетипичные ситуации ими обрабатываются ошибочно. *Желательно* было бы, чтобы алгоритм прогнозирования корректно обрабатывал не только типовые, но и нетипичные случаи.

Таким образом, налицо противоречие между фактическим и желаемым и это противоречие *неприемлемо*.

Постановка целей и задач работы подробно описывается в разделе 1.1 работы [35].

¹ Далее в работе будем говорить о ссудах

² Это очень напоминает вечный двигатель, основанный на «демоне Максвелла», непосредственно преобразующий информацию в энергию.

³ Случаи целенаправленного банкротства в данной работе не рассматриваются.

1.2. Обоснование требований к методу решения проблемы

Конкретно требования к методу решения проблемы приведены в разделе 1.2 работы [35]. Очень важно также, чтобы данный метод был реализован в реально доступной пользователям программной системе.

1.3. Выбор метода решения проблемы, соответствующего обоснованным требованиям

В работах [4-9], посвященных вопросам прогнозирования рисков невозврата заемных средств (скоринг) само понятие «нетипичные случаи» вообще не встречается. В работе [10] оно встречается, но не приводится ни способа выявления таких нетипичных случаев с применением технологий искусственного интеллекта, ни *способа их обработки и учета в интеллектуальных моделях* (о которых в работе [10] вообще не упоминается).

В работах автора [11-18] приводятся исследования и разработки учета нетипичных случаев в интеллектуальных скоринговых моделях (см. также раздел 1.3 работы [35]).

Эти разработки соответствуют всем обоснованным в разделе 1.2 требованиям, за исключением одного, но критически важного требования: описанная в них DOS-версия системы «Эйдос» морально устарела и давно отсутствует в открытом доступе.

Поэтому с 2012 года автором разрабатывается и непрерывно совершенствуется органичная для MS Windows новая версия этой системы: «Эйдос-Х++» (подробнее о системе «Эйдос» см. также раздел 1.3 работы [35]).

В рамках решения поставленной в данной работе проблемы в системе «Эйдос» должен быть разработан новый режим, реализующий предложенные в данной работе идею, концепцию и интеллектуальный итерационный алгоритм учета в модели прогнозирования нетипичных случаев возврата и невозврата ссуды. Описание данных идеи, концепции, алгоритма и режима в открытой печати осуществляется впервые. В этом состоит актуальность данной работы.

1.4. Выводы

По результатам 1-й главы сформулированы объект, предмет, проблема, цель и задачи работы. В качестве метода решения поставленной проблемы выбран АСК-анализ, который полностью соответствует обоснованным требованиям.

Вместе с тем этот метод должен быть доработан в следующих аспектах:

1. Должна быть предложена идея и разработана концепция учета в модели прогнозирования нетипичных случаев возврата и невозврата ссуды.

2. Данная концепция должна быть конкретизирована и детализирована до уровня, описанного в разделе 1.4 работы [35].

В этой доработке наиболее подходящего метода решения проблемы и его апробации на основе реальных данных и состоит содержание последующих разделов данной работы.

2. Научное (теоретическое) решение проблемы

2.1. Идея и концепция научного решения проблемы

Идея научного решения проблемы состоит в следующем. Обобщенные образы классов создаются путем обобщения всех относящихся к ним наблюдений обучающей выборки. При верификации моделей выясняется, что некоторые наблюдения модель правильно относит к тем классам, к которым они действительно принадлежат, а некоторые нет. Естественно считать, что верно идентифицируются те наблюдения, которые являются типичными для данной предметной области. Нетипичные же наблюдения, являющиеся исключениями из правила (т.е. выявленных и отраженных в модели закономерностей), модель не относит к тем классам, к которым они в действительности относятся, т.е. дает ложно-отрицательные решения. Предлагается создать собственные классы с теми же наименованиями для нетипичных наблюдений (исключений).

Концепция научного решения проблемы. Формулировка концепции решения проблемы дана в разделе 2.1 работы [35]. Необходимо также провести верификацию моделей путем распознавания в них наблюдений из обучающей выборки. Выбрать модель, наиболее достоверную по F-критерию Ван Ризбергена. Если количество ложно-отрицательных решений в результатах распознавания в наиболее достоверной модели не уменьшилось, то переход на выход. Иначе для наблюдений обучающей выборки, идентификация которых привела к ложно-отрицательным решениям (т.е. они не были отнесены к классам, к которым в действительности относятся) создать для них в исходных данных свои классы с теми же наименованиями, но с добавлением к названию символов, означающих, что этот класс создан для исключений. Перейти на начало. Выход.

2.2. Описание общей теории метода решения проблемы: суть метода АСК-анализа

Суть метода АСК-анализа описана в разделе 2.2 работы [35].

Понятие данных, по-видимому, является одним из предельно общих понятий, выработанных человечеством, может быть даже это наиболее общее понятие. По мнению автора, понятие данных является даже более общим, чем понятия материи и сознания, бытия и небытия, пространства и времени. О чем бы мы не говорили, о чем бы не думали, что бы не

представляли, все это мы делаем лишь оперируя определенными данными. Таким образом, человек находится как бы в мыльном пузыре или коконе из данных, за пределы которого он в принципе не может выйти. И цветные движущиеся узоры на поверхности этого мыльного пузыря человек и принимает за реальный мир. А чтобы определить, что такое данные именно это и нужно сделать, т.е. необходимо выйти за пределы данных к чему-то еще более общему. Но как это сделать не совсем понятно, даже в принципе. Это создает большие сложности для традиционного определения понятия данных (см. раздел 2.2 работы [35])⁴.

Поэтому предлагается два варианта.

1-й вариант: указать только специфический признак данных, а более общее понятие не указывать. Этот вариант приводит к такому определению: **данные** – это все, связанное с теми или иными различиями в наличии или степени выраженности тех или иных свойств.

2-й вариант: определить понятие данных через частное понятие, входящее в него, и сделать это не смотря на то, что это не принято⁵.

Подробнее смысловое содержание понятий: «Данные», «Информация», «Знания» в сопоставлении друг с другом раскрывается в разделе 2.2 работы [35].

2.3. Описание личного вклада автора в теоретическое решение проблемы (научная новизна)

- *предлагается* идея и концепция учета нетипичных случаев;
- которая *в отличие от традиционного подхода* (Луценко, 1995, 1996);
- *путем* разделения классов на типичную и нетипичную части (нетипичными считаются наблюдения, порождающие ложно-отрицательные решения);
- *обеспечивает 61,5%:* повышение достоверности прогнозирования рисков невозврата ссуды по критерию L1 проф.Е.В.Луценко (мера L1 является нечетким мультиклассовым обобщением известной F-меры Ван Ризбергера [24]).

В этом и состоит научная новизна исследования.

2.4. Выводы

В данном разделе автором *предложен* интеллектуальный итерационный алгоритм учета нетипичных случаев, который *в отличие от*

⁴ Например, как определяют млекопитающее: это животное (более общее понятие), выкармливающее своих детенышей молоком (специфический признак, выделяющий из более общего понятия определяемое подмножество).

⁵ Например, животные (более общее понятие), это и млекопитающие, и не млекопитающие, т.е. как выкармливающее своих детенышей молоком (специфический признак, выделяющий из более общего понятия известное частное понятие), так и не молоком и вообще не выкармливающие.

традиционного подхода (Луценко, 1995, 1996) путем разделения классов на типичную и нетипичную части (нетипичными считаются наблюдения, порождающие ложно-отрицательные решения) *обеспечивает 61,5%: повышение достоверности прогнозирования рисков невозврата ссуды.*

Этим самым поставленная проблема решена теоретически и задача главы выполнена.

Однако для применения предложенного теоретического решения поставленной проблемы на практике необходимо разработать и создать инфраструктуру практического применения этого решения. Решению этой задачи и посвящен следующий раздел.

3. Инженерное (практическое) решение проблемы: технология (программный инструментарий) и методика решения проблемы

3.1. Техничко-экономическое обоснование целесообразности работы (ТЭО)

Ответим на вопрос: «*Зачем это нужно?*». Ответ на этот вопрос, по сути, является основным содержанием «Техничко-экономического обоснования» (ТЭО).

Задача прогнозирования рисков невозврата финансовых средств (ссуд, кредитов, инвестиций) физическими и юридическими лицами является весьма актуальной для всех, систематически занимающихся этим видом деятельности. Если достоверно прогнозировать риск невозврата ссуды и предоставлять ее только тогда, когда риск ее невозврата мал и не предоставлять ее, когда он велик, то прибыль от субсидирования будет высока. Понятно, что чем ниже достоверность прогнозирования риска невозврата ссуды, тем ниже прибыльность от субсидирования. Более того, при высоком риске невозврата субсидирование может привести к убыточности от этого вида деятельности или даже к банкротству («неоправданное кредитование»). Поэтому прибыль от субсидирования самым непосредственным образом зависит от достоверности прогнозирования риска невозврата ссуды.

Из этого следует целесообразность проведения исследования.

3.2. Техническое задание (ТЗ): функциональное описание того, что необходимо сделать (ТЗ)

Ответим на вопрос: «*Что нужно сделать*», чтобы решить поставленную проблему, т.е. какие *функции* необходимо реализовать в программной системе, чтобы она обеспечивала решение поставленной проблемы. Ответ на этот вопрос является основным содержанием «Технического задания» (ТЗ).

В системе «Эйдос» необходимо реализовать режим, обеспечивающий реализацию сформулированных выше идеи и концепции решения проблемы.

3.3. Технический проект (ТП) и рабочий проект (РП): структуры и отношения данных, алгоритмы их обработки, реализация

Подробнее содержание технического проекта описано в разделе 3.3 работы [35].

3.3.1. Суть математической модели АСК-анализа и частные критерии

Суть математической модели АСК-анализа описана в разделе 3.3.1 работы [35].

3.3.2. Интегральные критерии и принятие управляющих решений

Интегральные критерии, применяемые в настоящее время в АСК-анализе, описаны в разделе 3.3.2 работы [35].

3.3.3. Интеллектуальный итерационный алгоритм учета нетипичных случаев

Предлагается следующий алгоритм решения проблемы (рисунок 1):

Шаг 1. Содержание данного шага приведено в разделе 3.3.3 работы [35].

Шаг 2. Выполнить синтез моделей [35].

Шаг 3. Провести верификацию моделей путем распознавания в них наблюдений из обучающей выборки.

Шаг 4. Выбрать модель, наиболее достоверную по F-критерию Ван Ризбергена.

Шаг 5. Если количество ложно-отрицательных решений в результатах распознавания в наиболее достоверной модели не уменьшилось, то переход на шаг 7 (выход).

Шаг 6. Для наблюдений обучающей выборки, идентификация которых привела к ложно-отрицательным решениям (т.е. они не были отнесены к классам, к которым в действительности относятся) создать для них в исходных данных свои классы с теми же наименованиями, но с добавлением к названию символов, означающих, что этот класс создан для исключений. Перейти на шаг 1.

Шаг 7. Выход.

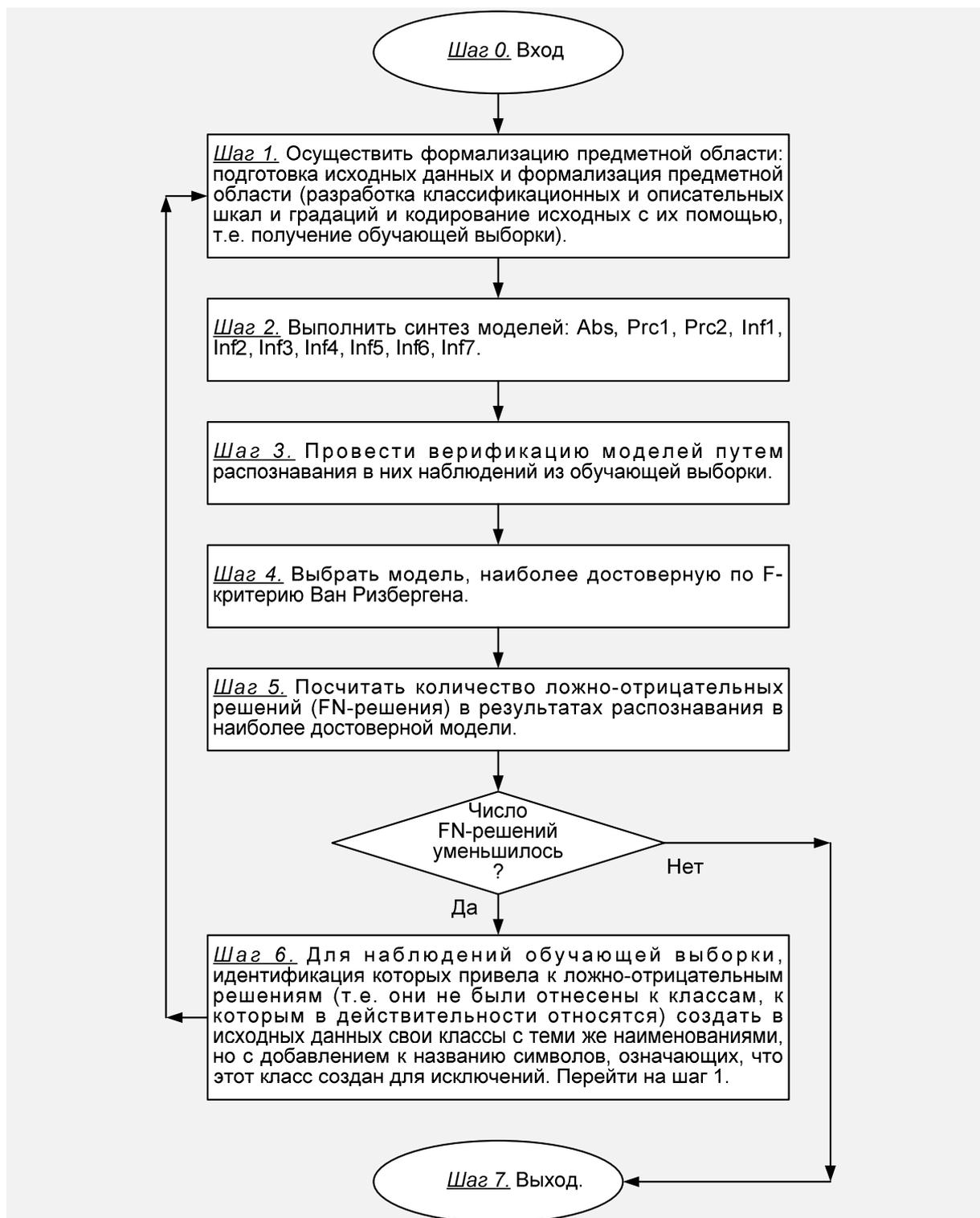


Рисунок 1. Интеллектуальный итерационный алгоритм учета нетипичных случаев

Рабочий проект (РП) – это сама программная реализация данного алгоритма и соответствующих структур данных.

В этапы разработки технического и рабочего проекта АСК-анализа и интеллектуальной системы «Эйдос» приведены в разделе 3.3.3 работы [35].

Обычно считается, что программирование является не научной, а инженерной работой (приложение 1). Автор всю сознательную жизнь профессионально занимается программированием (с 1976 года) и совершенно не согласен с этой распространённой точкой зрения. По мнению автора, программирование, являясь инженерным трудом, безусловно, включает серьёзную научную компоненту, а также является видом искусства.

3.4. Выводы

По результатам данного раздела можно сделать обоснованный вывод о том, что инфраструктура применения на практике ранее разработанного научного решения проблемы успешно спроектирована и создана, чем научное решение доведено до статуса инновации.

Следующим шагом является *внедрение, применение на практике* новой версии системы «Эйдос», в состав которой входит режим, реализующий предложенные в данной работе идею, концепцию и интеллектуальный алгоритм учета нетипичных случаев при прогнозировании рисков невозврата ссуды.

В данной работе в качестве применения на практике предложенных решений рассмотрим подробный *численный пример*, основанный на реальных данных, находящихся в общем доступе. Для этого необходимо разработать все виды обеспечения внедрения и осуществить само внедрение, что позволит проверить эффективность всех ранее принятых решений и оценить их значение для науки и практики. Все это и является содержанием следующего раздела.

4. Применение разработанной технологии на практике, ее внедрение и оценка эффективности решения проблемы

4.1. Методика, план и обеспечение внедрения

Методика внедрения – это перечень этапов работ или задач, которые необходимо выполнить для осуществления внедрения.

В разделе данной работы путем декомпозиции ее цели получена определенная последовательность *задач*, подробнее о которых можно прочитать в разделе 4.1 работы [35].

План внедрения – это утвержденный руководителем внедрения перечень мероприятий с датами выполнения и ответственными по каждому пункту. В принципе план внедрения это методика внедрения, дополненная различной информацией, необходимой для ее реализации.

В данном случае автор в одном лице совмещал роли и руководителя, и ответственного исполнителя внедрения. Как исполнитель автор решал все поставленные задачи в сроки, которые он заранее жестко не определял, но так, чтобы успеть подготовить материал к очередному номеру журнала.

Виды обеспечения внедрения:

- финансовое (финансирование работ по внедрению);
- организационное (описание схем и регламентов информационного, финансового и материального взаимодействия участников внедрения, логистика);
- юридическое (документы, придающие внедрению и формируемым в его результате отношениям и документам необходимый юридический статус);
- кадровое (специалисты, участвующие во внедрении, их количество и компетенции);
- техническое (компьютерная и орг.техника, компьютерные сети, доступ в Internet, а транспортное обеспечение, здания и сооружения, служебные помещения).

Все эти виды обеспечения внедрения необходимы для его успеха.

Данная работа выполнялась по инициативе и за счет автора на его собственном компьютере.

4.2. Методика оценки эффективности внедрения

Эффективность внедрения может рассматриваться в натуральном или/и стоимостном выражении.

В натуральном выражении эффективность это в основном количество и качество продукции.

В стоимостном выражении эффективность это в основном прибыль и рентабельность.

В случае данной работы эффективность предложенных решений в *натуральном* выражении может измеряться повышением достоверности моделей прогнозирования, в частности уменьшением количества ложно-отрицательных решений (ошибок не отнесения ситуаций к классам, к которым они на самом деле относятся). В данной работе мы оценим этот вид эффективности в численном примере, основанном на реальных данных.

В стоимостном выражении эффективность предлагаемых решений можно определить путем сравнения прибыли от субсидирования за определённый период до их применения и после их применения. Ожидается увеличение этой прибыли за счет снижения доли неадекватных решений.

4.3. Описание внедрения и его результатов (численный пример)

На этом численном примере наглядно проиллюстрируем принятые решения в той форме, в которой они в настоящее время реализованы в системе «Эйдос», и исследуем их эффективность [35].

4.3.1. Задача 1: когнитивная структуризация предметной области

Система «Эйдос» выявляет *эмпирические закономерности* в моделируемой предметной области и отображает их в различных формах: табличной, графической и аналитической. Это соответствует эмпирическому этапу развития. Этим самым она вплотную подводит исследователя к теоретическому уровню познания [23, 35]⁶

О том, что это значит, можно подробнее ознакомиться в разделе 4.3.1 работы [35].

Данная работа основана на исходных данных, размещенных на портале Kaggle: <https://www.kaggle.com/kmlDas/loan-default-prediction>.

По описанию задачи на портале Kaggle можно сделать вывод о том, что ее смысл состоит в том, чтобы по характеристикам получателя ссуды спрогнозировать риск ее невозврата. На основе этого прогноза можно принять обоснованное решение о выдаче или невыдаче ссуды данному конкретному ссудополучателю.

Исходные данные содержат следующие параметры:

- серийный номер или уникальный идентификатор ссудополучателя;
- занятость ссудополучателя (работает ли он или является безработным);
- банковский баланс ссудополучателя;
- годовая зарплата ссудополучателя;
- результат возврата/невозврата ссуды.

О том, что в данной работе выбрано в качестве классификационных и описательных шкал и градаций можно прочесть в разделе 4.3.1 работы [35] (таблицы 1 и 2):

Таблица 1 – Классификационная шкала

Код	Наименование
1	DEFAULTED?

Таблица 2 – Описательные шкалы

Код	Наименование
1	EMPLOYED
2	BANK BALANCE
3	ANNUAL SALARY

4.3.2. Задача 2: подготовка исходных данных и формализация предметной области

Исходные данные для данной работы получены непосредственно с портала Kaggle по прямой ссылке: <https://www.kaggle.com/kmlDas/loan-default-prediction/download>.

Эти данные представлены в виде CSV-файла. После скачивания этого файла для ввода в систему «Эйдос» с ним было выполнено несколько простых преобразований:

⁶ См., также: http://lc.kubagro.ru/aidos/Works_on_identification_presentation_and_use_of_knowledge.htm

1. CSV-файл был переименован с «Default_Fin.csv» на «Inp_data.csv» и размещен в папке: ..\Aidos-X\AID_DATA\Inp_data\ системы «Эйдос» для исходных данных табличного типа.

2. CSV-файл был преобразован в XLS-файл для удобства дальнейшей корректировки и ввода в систему «Эйдос».

Само CSV-XLS преобразование (конвертирование) может быть осуществлено онлайн с помощью одного из онлайн-конвертеров. Рекомендуется использовать следующие CSV-XLS-онлайн конвертеры, которые очень хорошо работают со стандартными CSV-файлами:

<https://convertio.co/ru/csv-xls/>,

<https://onlineconvertfree.com/ru/convert-format/csv-to-xls/>;

<https://document.online-convert.com/ru/convert/csv-to-excel>.

В простейшем случае CSV-файл это текст, состоящий из строк, в каждой из которых содержится *одинаковое* количество элементов, разделенных каким-либо разделителем, чаще всего запятой. Таким образом, строки CSV-файла можно поставить в соответствие строкам таблицы, а элементы строк – колонкам таблицы.

Но следует иметь в виду, что сам CSV-стандарт (форматированный текст) еще не совсем устоялся. Но в CSV-файлах в качестве разделителя могут быть использованы и другие символы, например, точка с запятой или табуляция. Иногда, когда необходимо, чтобы внутри элементов использовалась запятая, эти элементы выделяют кавычками. Поэтому иногда (достаточно редко) встречаются CSV-файлы с необычными форматами, которые не всякий конвертер сможет корректно преобразовать. В этом случае рекомендуется попробовать подобрать другой конвертер, которых очень много в открытом доступе. Потратив на это некоторое время, обычно удается получить желаемый результат.

3. После преобразования CSV-файла в XLS-файл в нем средствами MS-Excel были произведены следующие корректировки:

– в колонках Employed и Defaulted логические переменные 0/1 замены на Yes, No соответственно;

– колонка Defaulted, советующая классификационной шкале, выделена желтым фоном;

– в числовых колонках сделано два знака после запятой (по числу максимальному числу разрядов после запятой, которое встречается в исходных данных в данной колонке);

– при разработке реальных научных интеллектуальных приложений убедительно рекомендуется в числовых колонах в обязательном порядке указывать единицы измерения, в нашем случае это доллары США;

– XLS-файл (стандарт MS Excel-2003) записан в стандарте более новых версий MS Excel как XLSX. Это сделано потому, что в новом стандарте файл имеет размер примерно в два раза меньше, чем в старом.

В результате всех этих операций получилась таблица исходных данных (см. таблицу 3 в работе [35]):

В таблице 3 работы [35] приведен лишь небольшой фрагмент исходных данных, т.к. в этой таблице 10000 строк. Полностью ее можно скачать из Эйдос-облака по прямой ссылке:

http://aidos.byethost5.com/Source_data_applications/Applications-000293/Inp_data.xlsx.

Запустим режим 2.3.2.2 системы «Эйдос» [35] (рисунок 2):

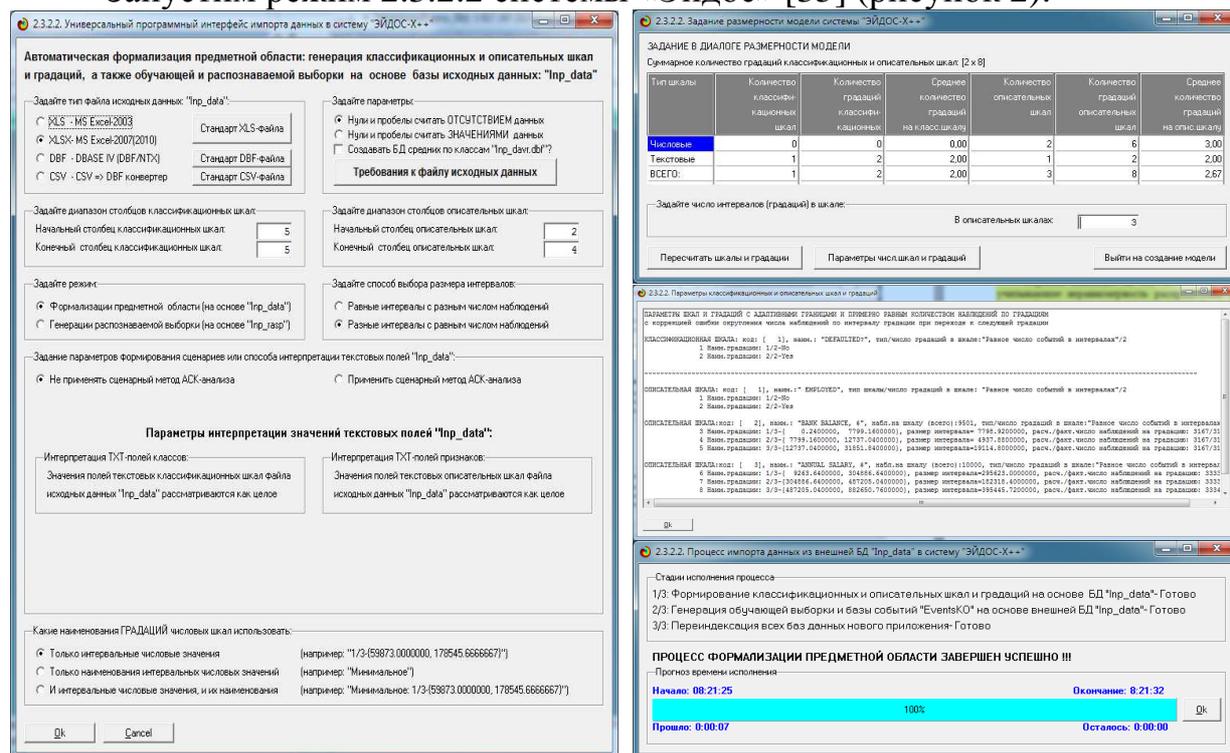


Рисунок 2. Экранные форма программного интерфейса (API) 2.3.2.2 системы «Эйдос» с внешними данными табличного типа

О выводах, которые можно сделать на основе рисунка 4 и о смысле и требованиях API-2.3.2.2 (рисунок 3) можно узнать из раздела 4.3.2 работы [35].

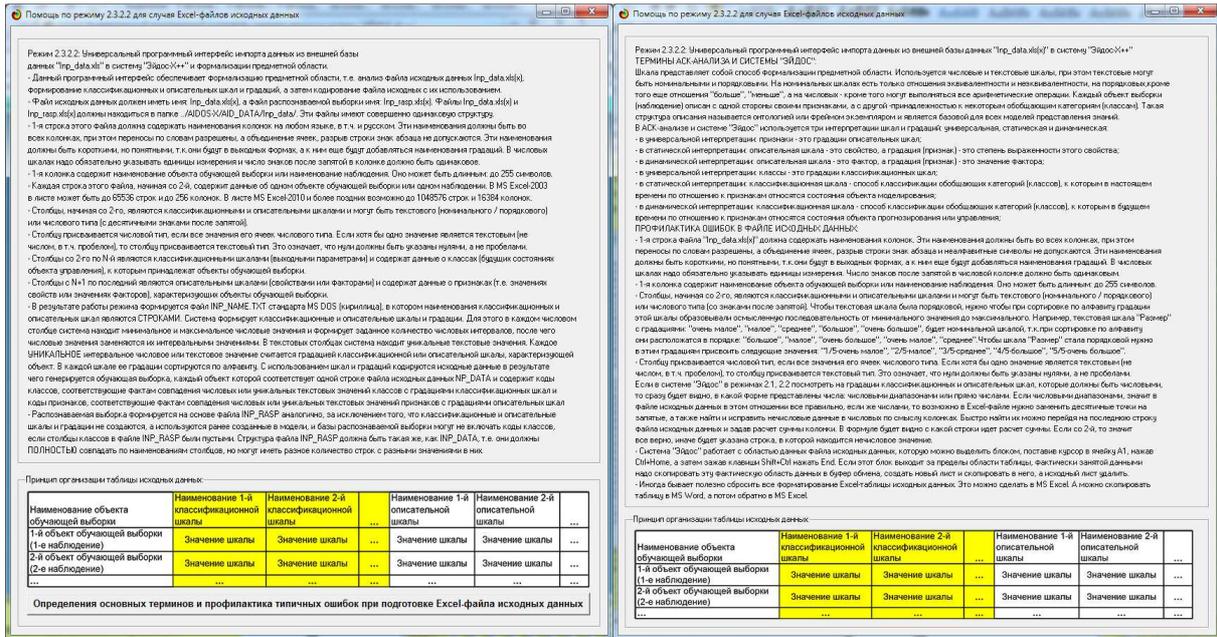


Рисунок 3. Экранные формы HELP программного интерфейса (API) 2.3.2.2

О результатах работы API-2.3.2.2, а также о том, что представляет собой обучающая выборка и какие возможности дальнейшей обработки открывает API-2.3.2.2 можно прочитать в разделе 4.3.2 работы [35] (таблицы 3, 4, 5).

Таблица 3 – Классификационные шкалы и градации (возврат/невозврат ссуды)

Код	Наименование
1	DEFAULTED?-1/2-No
2	DEFAULTED?-2/2-Yes

Таблица 4 – Описательные шкалы и градации (характеристики ссудополучателей)

Код	Наименование
1	EMPLOYED-1/2-No
2	EMPLOYED-2/2-Yes
3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}
4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}
5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}
6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}
7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}
8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}

Таблица 5 – Обучающая выборка (фрагмент)

ID ссудополучателя	EMPLOYED	BANK BALANCE, \$-	ANNUAL SALARY, \$-	DEFAULTED?-?
1	2	4	8	1
2	1	4	6	1
3	2	5	7	1
4	2	3	7	1
5	2	4	7	1
6	1	4	6	1
7	2	4	6	1
8	1	4	6	1
9	2	5	7	1
***	***	***	***	***
25	2	4	7	1

4.3.3. Задача 3: применение итерационного алгоритма разделения обобщенных образов классов на типичную и нетипичную части

В соответствии с предложенным интеллектуальным итерационным алгоритмом учета нетипичных случаев при прогнозировании невозврата ссуды сначала создадим базовую (исходную) модель, а затем будем повышать ее достоверность путем создания в исходных данных классов для нетипичных случаев и синтеза и верификации моделей в последовательных итерациях. Этот процесс останавливается тогда, когда модели перестают улучшаться в его результате.

4.3.3.1. Исходная модель

4.3.3.1.1. Синтез и верификация статистических и системно-когнитивных моделей

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) осуществляется в режиме 3.5 системы «Эйдос» (рисунок 4).

Обратим внимание на то, что на рисунке 6 в правом нижнем углу окна задана опция: «Расчеты проводить на графическом процессоре (GPU)».

Из рисунка 6 видно, что весь процесс синтеза и верификации моделей занял 5 минут 34 секунды. Отметим, что при синтезе и верификации моделей использовался графический процессор (GPU) видеокарты. На центральном процессоре (CPU) выполнение этих операций занимает значительно большее время (на некоторых задачах это происходит в десятки, сотни и даже тысячи раз дольше). Таким образом, неграфические вычисления на графических процессорах видеокарты делает возможной обработку больших объемов исходных данных за разумное время. В процесс синтеза и верификации моделей осуществляется также расчет 10 выходных форм, на что уходит более 99% времени исполнения.

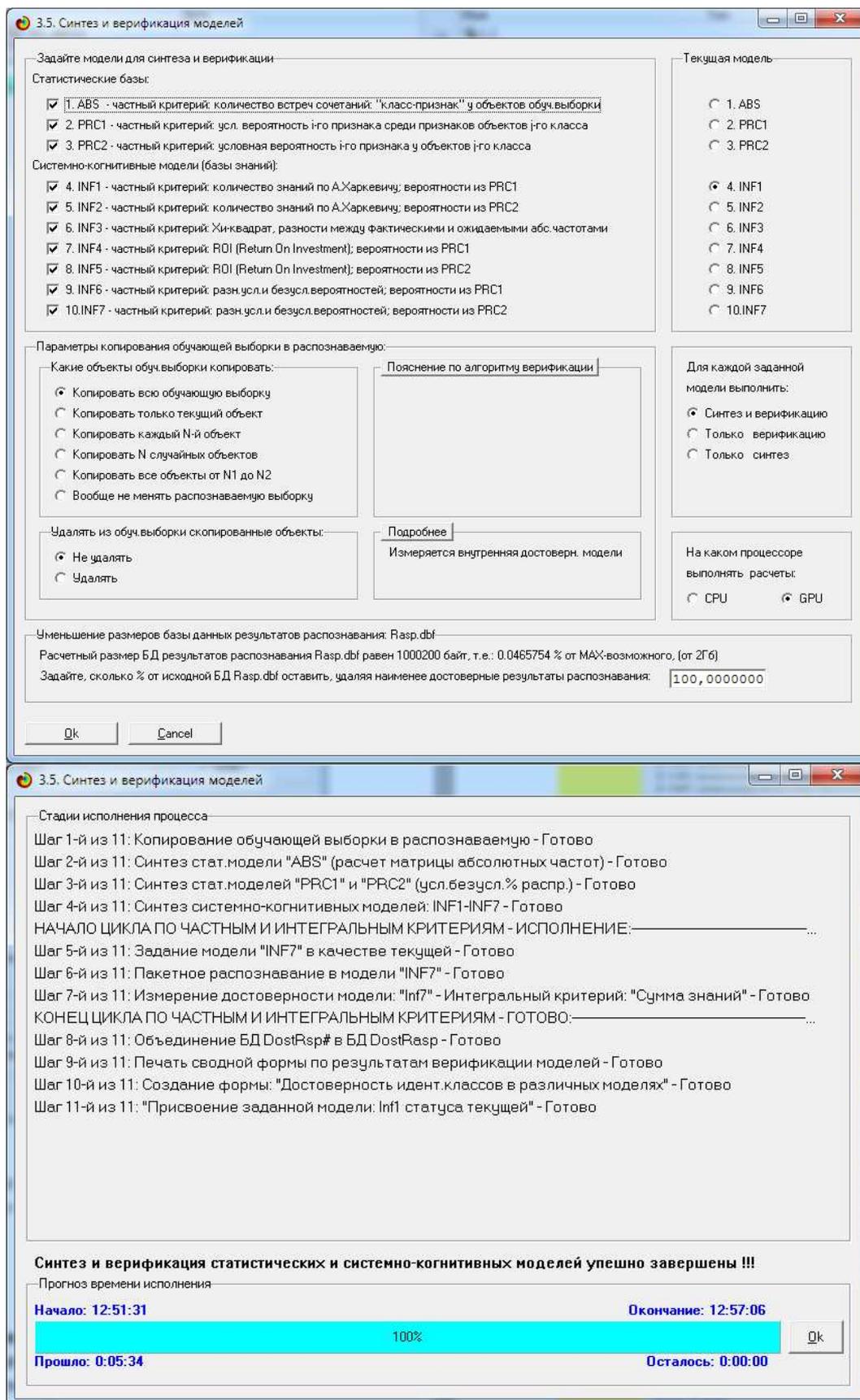


Рисунок 4. Экранные формы режима синтеза и верификации моделей системы «Эйдос» (режим 3.5)

Некоторые созданные статистические и системно-когнитивные модели (СК-модели) приведены на рисунках 5 - 8:

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

Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/2 NO	2. DEFAULTED? 2/2 YES	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	2817	127	2944	1472.00	1902.12
2	EMPLOYED-2/2-Yes	6850	206	7056	3528.00	4698.02
3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}	3167		3167	1583.50	2239.41
4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}	3158	9	3167	1583.50	2226.68
5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}	2843	324	3167	1583.50	1781.20
6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}	3199	134	3333	1666.50	2167.28
7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}	3235	98	3333	1666.50	2218.19
8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}	3233	101	3334	1667.00	2214.66
	Сумма числа признаков	28502	999	29501		
	Среднее	3563	125		1843.81	
	Среднеквадратичное отклонение	1339	105			917.45
	Сумма числа объектов обуч.выборки	9667	333	10000		

Рисунок 5. Матрица абсолютных частот: статистическая Модель ABS

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

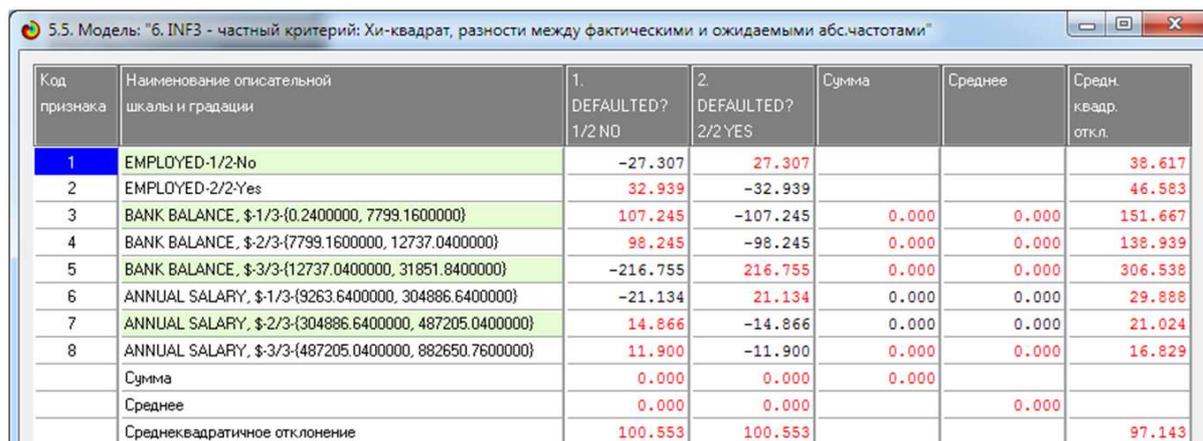
Код признака	Наименование описательной шкалы и градации	1. DEFAULT... 1/2 NO	2. DEFAULT... 2/2 YES	Безусл. вероятн.	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	29.140	38.138	29.440	33.639	6.362
2	EMPLOYED-2/2-Yes	70.860	61.862	70.560	66.361	6.362
3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}	32.761		31.670	16.380	23.165
4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}	32.668	2.703	31.670	17.685	21.189
5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}	29.409	97.297	31.670	63.353	48.004
6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}	33.092	40.240	33.330	36.666	5.055
7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}	33.464	29.429	33.330	31.447	2.853
8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}	33.444	30.330	33.340	31.887	2.201
	Сумма	294.838	300.000	594.838		
	Среднее	36.855	37.500		37.177	
	Среднеквадратичное отклонение	13.851	31.421			23.457

Рисунок 6. Матрица условных и безусловных процентных распределений: статистическая модель PRC2

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

Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/2 NO	2. DEFAULTED? 2/2 YES	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	-0.001	0.024	0.023	0.011	0.017
2	EMPLOYED-2/2-Yes	0.000	-0.014	-0.014	-0.007	0.011
3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}	0.003		0.003	0.002	0.002
4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}	0.003	-0.241	-0.238	-0.119	0.172
5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}	-0.007	0.107	0.100	0.050	0.081
6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}	-0.001	0.017	0.016	0.008	0.012
7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}	0.000	-0.014	-0.013	-0.007	0.010
8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}	0.000	-0.011	-0.010	-0.005	0.008
	Сумма	-0.001	-0.132	-0.133		
	Среднее	0.000	-0.017		-0.008	
	Среднеквадратичное отклонение	0.003	0.099			0.068

Рисунок 7. Матрица информативностей: СК-модель INF1



Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/2 NO	2. DEFAULTED? 2/2 YES	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	-27.307	27.307			38.617
2	EMPLOYED-2/2-Yes	32.939	-32.939			46.583
3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}	107.245	-107.245	0.000	0.000	151.667
4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}	98.245	-98.245	0.000	0.000	138.939
5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}	-216.755	216.755	0.000	0.000	306.538
6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}	-21.134	21.134	0.000	0.000	29.888
7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}	14.866	-14.866	0.000	0.000	21.024
8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}	11.900	-11.900	0.000	0.000	16.829
	Сумма	0.000	0.000	0.000		
	Среднее	0.000	0.000		0.000	
	Среднеквадратичное отклонение	100.553	100.553			97.143

Рисунок 8. Матрица хи-квадрат: СК-модель INF3

Отметим, что в АСК-анализе и СК-моделях степень выраженности различных свойств объектов наблюдения рассматривается с единственной точки зрения: с точки зрения того, какое **количество информации** содержится в них о том, к каким обобщающим категориям (классам) будут принадлежать или не принадлежать эти объекты. Поэтому не играет никакой роли, в каких единицах измерения измеряются те или иные свойства объектов наблюдения, а также в каких единицах измеряются результаты влияния этих свойств, натуральных, в процентах или стоимостных [21]. Это и есть решение проблемы сопоставимости в АСК-анализе и системе «Эйдос», отличающее их от других интеллектуальных технологий.

4.3.3.1.2. Оценка достоверности моделей

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных и ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1- L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры [24]. В режиме 3.4 системы «Эйдос» изучается достоверность каждой частной модели в соответствии с этими мерами достоверности (рисунок 9).

Из рисунка 11 мы видим, что в данном интеллектуальном приложении по F-критерию Ван Ризбергена наиболее достоверной является СК-модель INF3 с интегральным критерием «Сумма знаний» ($F=0,708$ при максимуме 1,000), что является неплохим результатом для моделируемой предметной области.

Это подтверждает наличие и адекватное отражение в СК-модели INF3 сильной причинно-следственной зависимости между

характеристиками ссудополучателей и возвратом или невозвратом ими полученной ссуды.

3.4. Обобщенная форма по доствам моделей при разн. крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	Число ложноположительных решений (FP)	Число ложноотрицательных решений (FN)	Точность модели	Полнота модели	Финера Ван Ризбергана	Сумма мод. уровней соо. истинно-поло. решений (ST)	Сумма мод. уровней соо. истинно-отриц. решений (ST)	Сумма мод. уровней соо. ложно-поло. решений (SF)	Сумма мод. уровней соо. ложно-отриц. решений (SF)	S-Точность модели	S-Полнота модели	L1-мера пров. Е.В. Писцеева
1. ABS - частный критерий: количество встреч сомнений; "Клас.	Корреляция абс. частот с обр...	3546	2826	0.669	0.717	0.692	4934.698	1688.196	2560.961	1242.544	0.658	0.799	0.722
1. ABS - частный критерий: количество встреч сомнений; "Клас.	Сумма абс. частот по призна...	10000		0.500	1.000	0.667	8624.811		572.086		0.938	1.000	0.368
2. PRC1 - частный критерий: усл. вероятность i-го признака сред.	Корреляция усл.отриц. частот с о...	3546	2826	0.669	0.717	0.692	4934.698	1688.196	2560.961	1242.544	0.658	0.799	0.722
2. PRC1 - частный критерий: усл. вероятность i-го признака сред.	Сумма усл.отриц. частот по при...	10000		0.500	1.000	0.667	6325.932		5908.893		0.517	1.000	0.682
3. PRC2 - частный критерий: условная вероятность i-го признака.	Корреляция усл.отриц. частот с о...	3546	2826	0.669	0.717	0.692	4934.699	1688.196	2560.961	1242.545	0.658	0.799	0.722
3. PRC2 - частный критерий: условная вероятность i-го признака.	Сумма усл.отриц. частот по при...	10000		0.500	1.000	0.667	6222.304		5905.449		0.513	1.000	0.678
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в.	Семантический резонанс: зна...	6518	2918	0.521	0.708	0.600	3137.609	2699.264	2272.899	2366.918	0.580	0.570	0.575
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в.	Сумма знаний	3803	2939	0.650	0.706	0.677	210.809	3166.842	1248.929	85.914	0.144	0.710	0.240
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в.	Семантический резонанс: зна...	6518	2918	0.521	0.708	0.600	3138.461	2692.606	2268.804	2366.900	0.580	0.570	0.575
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в.	Сумма знаний	3803	2939	0.650	0.706	0.677	216.466	3133.694	1336.445	89.206	0.139	0.708	0.233
6. INF3 - частный критерий: "Хинквадрат" - разности между фактич.	Семантический резонанс: зна...	2318	2918	0.708	0.708	0.708	3273.554	3273.585	2257.545	2257.545	0.592	0.592	0.592
6. INF3 - частный критерий: "Хинквадрат" - разности между фактич.	Сумма знаний	2318	2918	0.708	0.708	0.708	3264.956	3264.984	2256.155	2256.154	0.591	0.591	0.591
7. INF4 - частный критерий: ROI (Return On Investment); вероятн.	Семантический резонанс: зна...	3546	2939	0.666	0.706	0.685	3152.927	3201.641	2128.251	2360.404	0.597	0.572	0.584
7. INF4 - частный критерий: ROI (Return On Investment); вероятн.	Сумма знаний	3803	2939	0.650	0.706	0.677	359.901	1488.645	2491.570	86.980	0.126	0.805	0.218
8. INF5 - частный критерий: ROI (Return On Investment); вероятн.	Семантический резонанс: зна...	3546	2939	0.666	0.706	0.685	3153.733	3202.236	2132.427	2360.375	0.597	0.572	0.584
8. INF5 - частный критерий: ROI (Return On Investment); вероятн.	Сумма знаний	3803	2939	0.650	0.706	0.677	353.157	1361.469	2510.981	85.657	0.123	0.805	0.214
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер.	Семантический резонанс: зна...	3546	2918	0.666	0.708	0.687	3246.230	3314.485	2033.411	2259.539	0.615	0.590	0.602
9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер.	Сумма знаний	3580	2918	0.664	0.708	0.686	359.706	1855.558	2368.342	83.468	0.132	0.812	0.227
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве.	Семантический резонанс: зна...	3546	2918	0.666	0.708	0.687	3236.554	3302.302	2049.359	2272.274	0.612	0.588	0.600
10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве.	Сумма знаний	3580	2918	0.664	0.708	0.686	352.923	1674.687	2401.822	82.710	0.128	0.810	0.221

Рисунок 9. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергана

На рисунке 10 приведено частотное распределения числа истинных и ложных положительных и отрицательных решений по результатам прогнозирования возврата/невозврата ссуды в СК-модели INF3 по данным обучающей выборки:

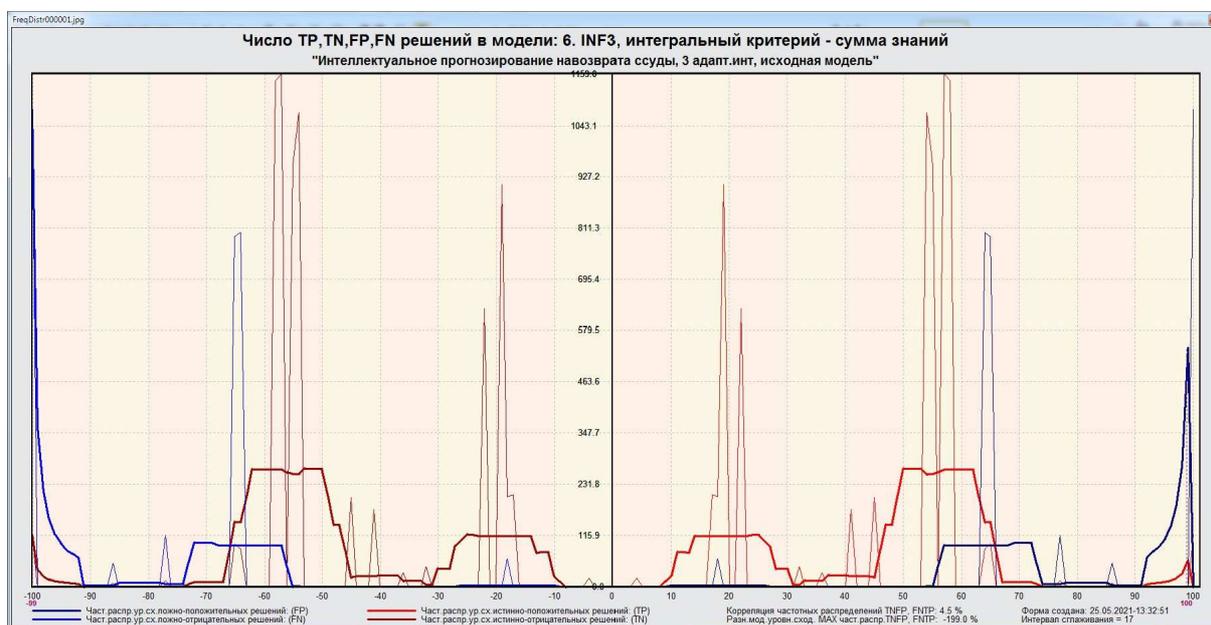


Рисунок 10. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели Inf3

Рисунок 10 содержит изображения частотных распределений количества истинных и ложных положительных и отрицательных решений в зависимости от уровня сходства. Из этого рисунка мы видим, что в районе высоких уровней сходства есть аномальный выброс ложноположительных и ложноотрицательных решений, которого не должно

быть в разумной модели. Возможно, этот выброс связан с изменениями, внесенными их автором, разместившим их на портале Kaggle. Он упоминает об этих изменениях и объясняет их необходимостью обезличивания данных.

4.3.3.1.3. Задание текущей модели и распознавание в ней

Важно иметь в виду, что корректируется (путем разделения классов на типичную и нетипичную части) всегда та модель, в которой проводилось распознавание. Если данный режим 3.7.6 запускается сразу после режима синтеза и верификации моделей 3.5 – то это будет СК-модель INF7, как последняя, в которой проводилась верификация в режиме 3.5. Если мы в режиме 3.4 выбрали другую модель для улучшения как наиболее достоверную, то надо ее задать в качестве текущей в режиме 5.6, а затем выполнить **распознавание** в этой модели в режиме 4.1.2.

Поэтому далее выполняем режим 5.6 (рисунки 1 и 11), а затем режим **распознавания** 4.1.2 в текущей модели INF3 (рисунок 12).

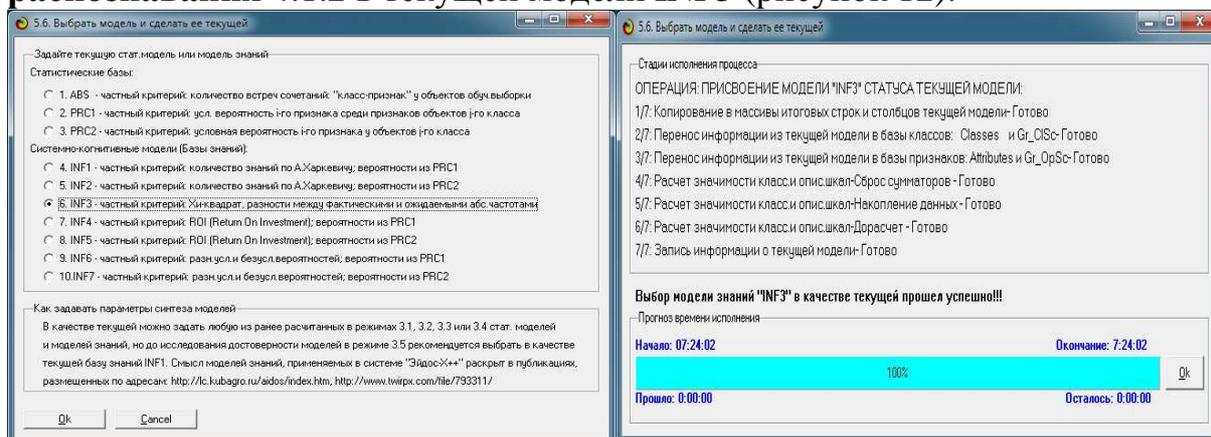


Рисунок 11. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

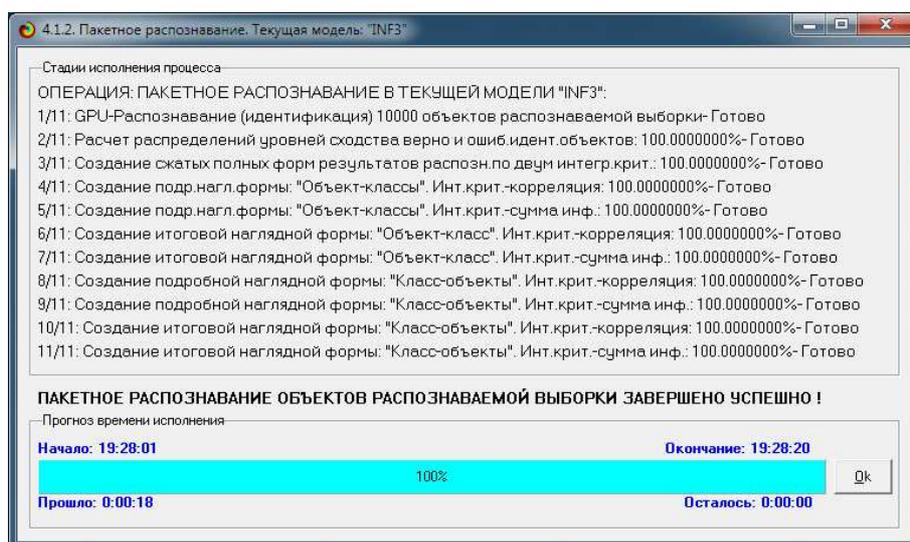


Рисунок 12. Экранная форма пакетного распознавания в текущей СК-модели Inf3

Из экранной формы на рисунке 14 видно, что процесс идентификации 10000 объектов распознаваемой выборки шел 18 секунд. Для самого распознавания использовался графический процессор (GPU), а точнее 1500 шейдерных процессоров видеокарты NVIDIA GeForce GTX 770. Для расчета 10 выходных форм по результатам распознавания использовался центральный процессор (CPU) i7. В основном время было затрачено именно на расчет выходных форм.

4.3.3.2. Первая итерация. Синтез и верификация статистических и системно-когнитивных моделей. Оценка достоверности моделей. Задание текущей модели и распознавание в ней

Для выполнения 1-й итерации запускаем режим 3.7.6 (рисунок 13):

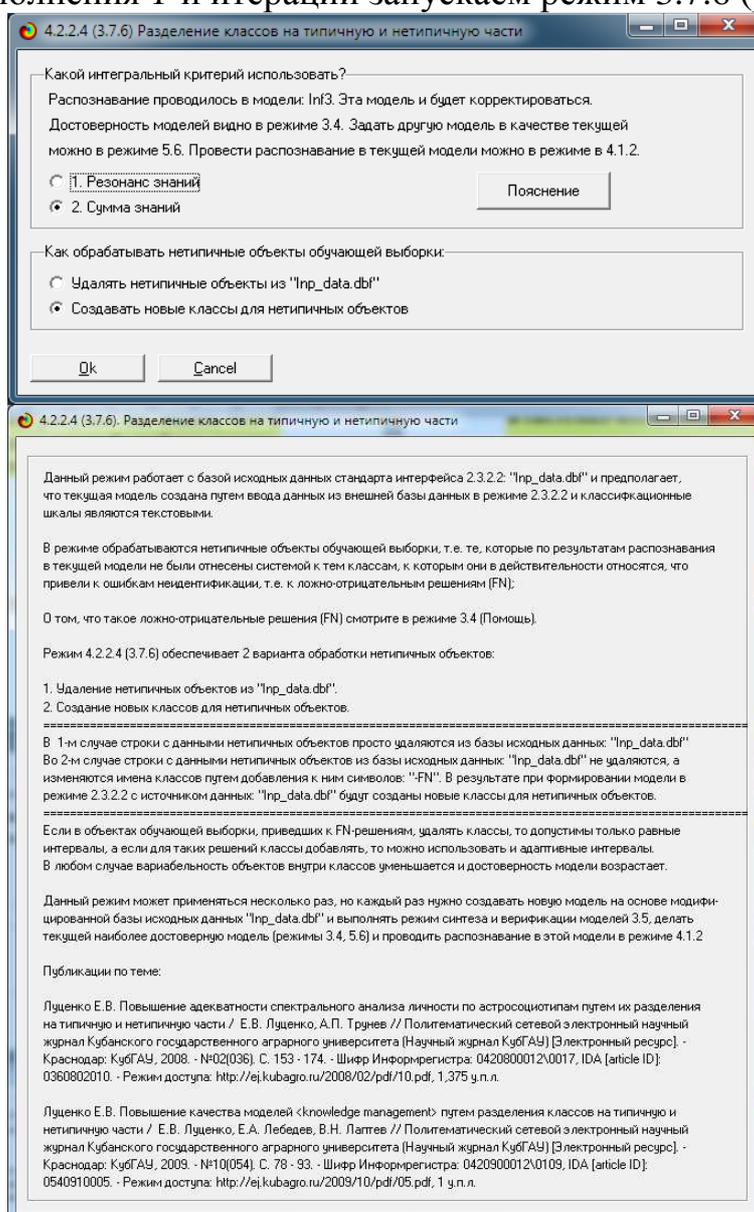


Рисунок 13. Экранные формы режима разделения классов на типичную и нетипичную части

После запуска режима появляется прогресс-бар и после него окно с сообщением и запускается API-2.3.2.2 с параметрами, сформированными данным режимом (рисунок 14). Затем выполняем режимы, следуя инструкциям системы (рисунки 15)⁷.

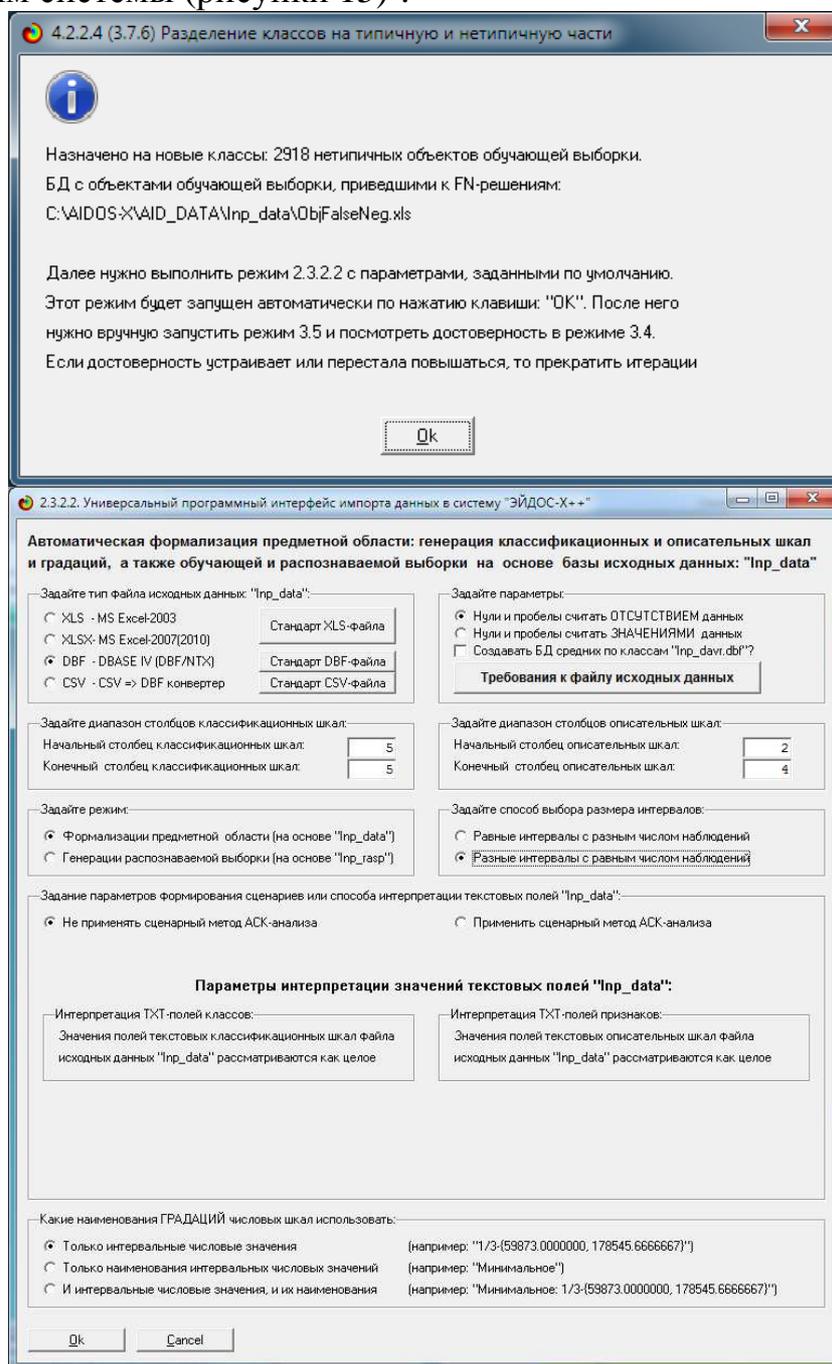


Рисунок 14. Экранные формы режима разделения классов на типичную и нетипичную части

⁷ Мы эти рисунки, полученные на 1-й итерации, не комментируем, т.к. их наименования есть в наименовании экранных форм и они полностью аналогичны полученным в исходной модели.

2.3.2.2. Задание размерности модели системы "ЭЙДОС-X++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ

Суммарное количество градаций классификационных и описательных шкал: [4 x 8]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	0	0	0,00	2	6	3,00
Текстовые	1	4	4,00	1	2	2,00
ВСЕГО:	1	4	4,00	3	8	2,67

Задайте число интервалов (градаций) в шкале:

В описательных шкалах:

Пересчитать шкалы и градации Параметры числ. шкал и градаций Выйти на создание модели

2.3.2.2. Параметры классификационных и описательных шкал и градаций

ПАРАМЕТРЫ ШКАЛ И ГРАДАЦИЙ С АДАПТИВНЫМИ ГРАНИЦАМИ И ПРИМЕРНО РАВНЫМ КОЛИЧЕСТВОМ НАБЛЮДЕНИЙ ПО ГРАДАЦИЯМ с коррекцией ошибки округления числа наблюдений по интервалу градации при переходе к следующей градации

КЛАССИФИКАЦИОННАЯ ШКАЛА: код: [1], наименование: "DEFAULTED?", тип/число градаций в шкале: "Равное число событий в интервалах"/4

- Наим. градации: 1/4-No
- Наим. градации: 2/4-No_FN
- Наим. градации: 3/4-Yes
- Наим. градации: 4/4-Yes_FN

ОПИСАТЕЛЬНАЯ ШКАЛА: код: [1], наименование: "EMPLOYED", тип шкалы/число градаций в шкале: "Равное число событий в интервалах"/2

- Наим. градации: 1/2-No
- Наим. градации: 2/2-Yes

ОПИСАТЕЛЬНАЯ ШКАЛА: код: [2], наименование: "BANK BALANCE, €", наблюд. на шкалу (всего): 9501, тип/число градаций в шкале: "Равное число событий в интервалах"

- Наим. градации: 1/3-[0.2400000, 7799.1600000], размер интервала= 7799.9200000, расч./факт. число наблюдений на градации: 3167/31
- Наим. градации: 2/3-[7799.1600000, 12797.0400000], размер интервала= 4997.8800000, расч./факт. число наблюдений на градации: 3167/31
- Наим. градации: 3/3-[12797.0400000, 31851.8400000], размер интервала=19114.8000000, расч./факт. число наблюдений на градации: 3167/31

ОПИСАТЕЛЬНАЯ ШКАЛА: код: [3], наименование: "ANNUAL SALARY, €", наблюд. на шкалу (всего): 10000, тип/число градаций в шкале: "Равное число событий в интервалах"

- Наим. градации: 1/3-[9263.6400000, 304886.6400000], размер интервала=295623.0000000, расч./факт. число наблюдений на градации: 3333

Ok

2.3.2.2. Процесс импорта данных из внешней БД "Inp_data" в систему "ЭЙДОС-X++"

Стадии исполнения процесса

- 1/3: Формирование классификационных и описательных шкал и градаций на основе БД "Inp_data"- Готово
- 2/3: Генерация обучающей выборки и базы событий "EventsKO" на основе внешней БД "Inp_data"- Готово
- 3/3: Переиндексация всех баз данных нового приложения- Готово

ПРОЦЕСС ФОРМАЛИЗАЦИИ ПРЕДМЕТНОЙ ОБЛАСТИ ЗАВЕРШЕН УСПЕШНО !!!

Прогноз времени исполнения

Начало: 14:16:10 Окончание: 14:16:17

100%

Прошло: 0:00:07 Осталось: 0:00:00

Ok

4.2.2.4 (3.7.6) Разделение классов на типичную и нетипичную части

Далее необходимо в режиме 3.5 создать и верифицировать модели: Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7.

Затем в режиме 3.4 необходимо определить модель и интегральный критерий, при котором достигается максимальная достоверность.

Режим 4.2.2.4 (3.7.6) можно повторять до достижения необходимого достаточно высокого уровня достоверности моделей.

Если достоверность модели достаточно высока, то в ней корректно можно решать задачи идентификации и прогнозирования (4.1.2), принятия решений (4.4.8) и исследования объекта моделирования путем исследования его модели (режимы: 4.4.9, 4.4.10, 4.4.11, 4.4.12, 4.2.1, 4.2.2.1, 4.2.2.2, 4.2.3, 4.2.3.1, 4.2.3.2, 4.2.3.3, 4.5, 3.7.5, 3.7.4., 3.7.3, 3.7.9 и т.д.)

Ok

3.5. Синтез и верификация моделей

Стадии исполнения процесса

- Шаг 1-й из 11: Копирование обучающей выборки в распознаваемую - Готово
- Шаг 2-й из 11: Синтез стат. модели "ABS" (расчет матрицы абсолютных частот) - Готово
- Шаг 3-й из 11: Синтез стат. моделей "PRC1" и "PRC2" (учл. безучл. % распр.) - Готово
- Шаг 4-й из 11: Синтез системно-когнитивных моделей: INF1-INF7 - Готово
- НАЧАЛО ЦИКЛА ПО ЧАСТНЫМ И ИНТЕГРАЛЬНЫМ КРИТЕРИЯМ - ИСПОЛНЕНИЕ: _____
- Шаг 5-й из 11: Задание модели "INF7" в качестве текущей - Готово
- Шаг 6-й из 11: Пакетное распознавание в модели "INF7" - Готово
- Шаг 7-й из 11: Измерение достоверности модели: "Inf7" - Интегральный критерий: "Сумма знаний" - Готово
- КОНЕЦ ЦИКЛА ПО ЧАСТНЫМ И ИНТЕГРАЛЬНЫМ КРИТЕРИЯМ - ГОТОВО: _____
- Шаг 8-й из 11: Объединение БД DostRsp# в БД DostRasp - Готово
- Шаг 9-й из 11: Печать сводной формы по результатам верификации моделей - Готово
- Шаг 10-й из 11: Создание формы: "Достоверность идент. классов в различных моделях" - Готово
- Шаг 11-й из 11: "Присвоение заданной модели: Inf1 статуса текущей" - Готово

Синтез и верификация статистических и системно-когнитивных моделей успешно завершены !!!

Прогноз времени исполнения

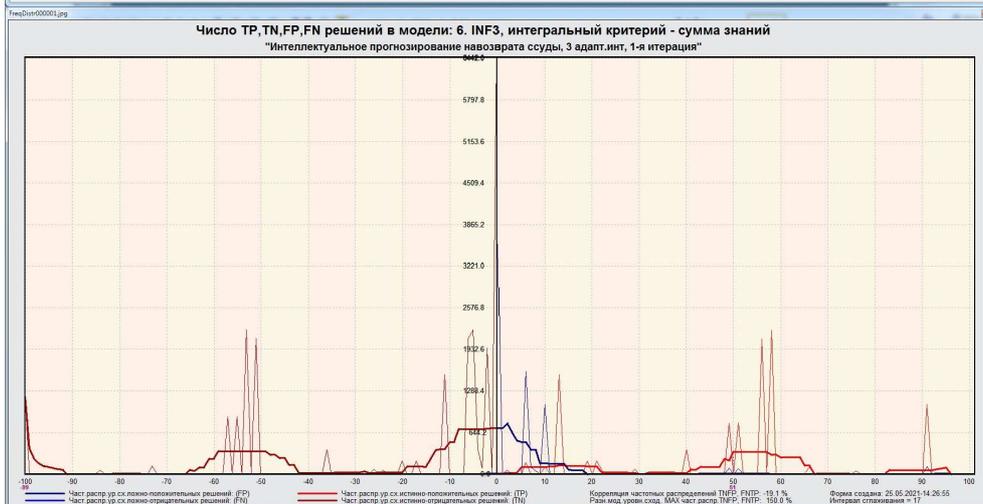
Начало: 14:17:18 Окончание: 14:25:34

100%

Прошло: 0:08:16 Осталось: 0:00:00

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

Наименование модели и частного критерия	Интегральный критерий	коэффициент (FP)	Число ложно-отрицательных решений (FN)	Точность модели	Помеха модели	Фигура Вилл Ризберга	Сумма мод. уровней сч. истинно-полож. решений (ST)	Сумма мод. уровней сч. истинно-отриц. решений (ST)	Сумма мод. уровней сч. ложнополож. решений (ST)	Сумма мод. уровней сч. ложноотриц. решений (ST)	S-Точность модели	S-Помеха модели	S-Цифра проф. Е.В. Пиджака
1. ABS - частный критерий: количество встреч символов "числ."	Корреляция абс. частот с абс...	3553	1620	0.382	0.838	0.325	4004.708	5695.962	5718.843	250.630	0.512	0.960	0.668
1. ABS - частный критерий: количество встреч символов "числ."	Сумма абс. частот по призна...	30000		0.250	1.000	0.400	7444.221		3896.364		0.656	1.000	0.793
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Корреляция усл.отч.частот с о...	3553	1620	0.382	0.838	0.325	4004.708	5695.962	5718.842	250.630	0.512	0.960	0.668
2. PRC1 - частный критерий: усл. вероятность его признака сред.	Сумма усл.отч.частот по при...	30000		0.250	1.000	0.400	6726.961		14519.487		0.517	1.000	0.481
3. PRC2 - частный критерий: условная вероятность его признака.	Корреляция усл.отч.частот с о...	3553	1620	0.382	0.838	0.325	4004.708	5695.962	5718.842	250.630	0.512	0.960	0.668
3. PRC2 - частный критерий: условная вероятность его признака.	Сумма усл.отч.частот по при...	30000		0.250	1.000	0.400	6820.532		14471.726		0.314	1.000	0.479
4. INF1 - частный критерий: количество значений по А/Харвингу в...	Семантический резонанс зна...	10061	1976	0.381	0.802	0.336	3539.090	9132.238	4502.111	469.816	0.440	0.888	0.587
4. INF1 - частный критерий: количество значений по А/Харвингу в...	Сумма значений	5795	22	0.357	0.998	0.358	3995.874	4212.047	5299.590	1.112	0.363	1.000	0.532
5. INF2 - частный критерий: количество значений по А/Харвингу в...	Семантический резонанс зна...	10061	2022	0.379	0.798	0.314	3556.484	9157.179	4541.348	476.123	0.439	0.882	0.586
5. INF2 - частный критерий: количество значений по А/Харвингу в...	Сумма значений	5795	22	0.357	0.998	0.358	3044.787	4108.077	5888.564	1.203	0.353	1.000	0.521
6. INF3 - частный критерий: Угол раздв. разности между фазам...	Семантический резонанс зна...	6812		0.998	1.000	0.746	6294.447	9895.100	4900.672		0.516	1.000	0.681
6. INF3 - частный критерий: Угол раздв. разности между фазам...	Сумма значений	4822		0.998	1.000	0.747	6366.264	9985.716	4482.266		0.496	1.000	0.646
7. INF4 - частный критерий: ROI (Return On Investment): вероят...	Семантический резонанс зна...	6745	1975	0.543	0.803	0.448	4275.115	10046.218	4397.877	446.335	0.508	0.908	0.641
7. INF4 - частный критерий: ROI (Return On Investment): вероят...	Сумма значений	6745	21	0.374	0.998	0.344	3708.562	2182.608	6587.166	0.399	0.360	1.004	0.530
8. INF5 - частный критерий: ROI (Return On Investment): вероят...	Семантический резонанс зна...	6745	1975	0.543	0.803	0.448	4288.469	10103.656	4243.337	453.649	0.503	0.904	0.646
8. INF5 - частный критерий: ROI (Return On Investment): вероят...	Сумма значений	6745	21	0.374	0.998	0.344	3657.689	2057.250	6763.835	0.462	0.351	1.000	0.520
9. INF6 - частный критерий: разность безуслов. вероятностей: вер...	Семантический резонанс зна...	10819	1767	0.431	0.823	0.366	4341.858	9469.654	4190.474	461.721	0.509	0.904	0.653
9. INF6 - частный критерий: разность безуслов. вероятностей: вер...	Сумма значений	6038		0.324	1.000	0.359	3455.624	2653.600	6672.084		0.341	1.000	0.509
10. INF7 - частный критерий: разность безуслов. вероятностей: ве...	Семантический резонанс зна...	10819	1767	0.431	0.823	0.366	4346.831	9706.285	4280.319	466.090	0.504	0.903	0.647
10. INF7 - частный критерий: разность безуслов. вероятностей: ве...	Сумма значений	6038		0.354	1.000	0.355	3334.429	2415.072	6736.398		0.331	1.000	0.497



5.5. Модель: "1. ABS - частный критерий: количество встреч сочетаний "Класс-признак" у объектов обуч.выборки"								
Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/4 NO	2. DEFAULTED? 2/4 NO_FN	3. DEFAULTED? 3/4 YES	4. DEFAULTED? 4/4 YES_FN	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	1618	1199	126	1	2944	736.00	796.79
2	EMPLOYED-2/2-Yes	5140	1710	198	8	7056	1764.00	2376.00
3	BANK BALANCE, \$-1/3-(0.2400000, 7799.1600000)	3167				3167	791.75	1583.50
4	BANK BALANCE, \$-2/3-(7799.1600000, 12737.0400000)	3158				3167	791.75	1577.51
5	BANK BALANCE, \$-3/3-(12737.0400000, 31851.8400000)		2843	324		3167	791.75	1376.00
6	ANNUAL SALARY, \$-1/3-(9263.6400000, 304886.6400000)	1939	1260	133	1	3333	833.25	928.76
7	ANNUAL SALARY, \$-2/3-(304886.6400000, 487205.0400000)	2379	856	94	4	3333	833.25	1099.09
8	ANNUAL SALARY, \$-3/3-(487205.0400000, 882650.7600000)	2440	793	97	4	3334	833.50	1127.38
	Сумма числа признаков	19841	8661	972	27	29501		
	Среднее	2480	1083	122	3		921.91	
	Среднеквадратичное отклонение	1474	927	106	4			828.91
	Сумма числа объектов обуч.выборки	6758	2909	324	9	10000		

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. DEFAULT... 1/4 NO	2. DEFAULT... 2/4 NO_FN	3. DEFAULT... 3/4 YES	4. DEFAULT... 4/4 YES_FN	Безусл. вероятн.	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	23.942	41.217	38.889	11.111	29.440	28.790	14.053
2	EMPLOYED-2/2-Yes	76.058	58.783	61.111	88.889	70.560	71.210	14.053
3	BANK BALANCE, \$-1/3-(0.2400000, 7799.1600000)	46.863				31.670	11.716	23.431
4	BANK BALANCE, \$-2/3-(7799.1600000, 12737.0400000)	46.730			100.000	31.670	36.682	47.614
5	BANK BALANCE, \$-3/3-(12737.0400000, 31851.8400000)		97.731	100.000		31.670	49.433	57.088
6	ANNUAL SALARY, \$-1/3-(9263.6400000, 304886.6400000)	28.692	43.314	41.049	11.111	33.330	31.042	14.759
7	ANNUAL SALARY, \$-2/3-(304886.6400000, 487205.0400000)	35.203	29.426	29.012	44.444	33.330	34.521	7.194
8	ANNUAL SALARY, \$-3/3-(487205.0400000, 882650.7600000)	36.105	27.260	29.938	44.444	33.340	34.437	7.631
	Сумма	293.593	297.731	300.000	300.000	1191.324		
	Среднее	36.699	37.216	37.500	37.500		37.229	
	Среднеквадратичное отклонение	21.809	31.871	32.558	39.368			30.431

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"								
Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/4 NO	2. DEFAULTED? 2/4 NO_FN	3. DEFAULTED? 3/4 YES	4. DEFAULTED? 4/4 YES_FN	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	-0.039	0.064	0.051	-0.193	-0.117	-0.029	0.118
2	EMPLOYED-2/2-Yes	0.016	-0.037	-0.031	0.042	-0.011	-0.003	0.038
3	BANK BALANCE, \$-1/3-(0.2400000, 7799.1600000)	0.077				0.077	0.019	0.039
4	BANK BALANCE, \$-2/3-(7799.1600000, 12737.0400000)	0.077			0.220	0.297	0.074	0.104
5	BANK BALANCE, \$-3/3-(12737.0400000, 31851.8400000)		0.217	0.220		0.437	0.109	0.126
6	ANNUAL SALARY, \$-1/3-(9263.6400000, 304886.6400000)	-0.028	0.049	0.037	-0.217	-0.159	-0.040	0.123
7	ANNUAL SALARY, \$-2/3-(304886.6400000, 487205.0400000)	0.012	-0.026	-0.030	0.053	0.008	0.002	0.039
8	ANNUAL SALARY, \$-3/3-(487205.0400000, 882650.7600000)	0.016	-0.041	-0.024	0.053	0.004	0.001	0.042
	Сумма	0.130	0.226	0.223	-0.042	0.536		
	Среднее	0.016	0.028	0.028	-0.005		0.017	
	Среднеквадратичное отклонение	0.043	0.085	0.083	0.141			0.090

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами"								
Код признака	Наименование описательной шкалы и градации	1. DEFAULTED? 1/4 NO	2. DEFAULTED? 2/4 NO_FN	3. DEFAULTED? 3/4 YES	4. DEFAULTED? 4/4 YES_FN	Сумма	Среднее	Средн. квадр. откл.
1	EMPLOYED-1/2-No	-361.997	334.691	29.001	-1.694	0.000	0.000	285.134
2	EMPLOYED-2/2-Yes	394.463	-361.524	-34.481	1.542	0.000	0.000	309.565
3	BANK BALANCE, \$-1/3-(0.2400000, 7799.1600000)	1037.023	-929.778	-104.346	-2.899	0.000	0.000	806.391
4	BANK BALANCE, \$-2/3-(7799.1600000, 12737.0400000)	1028.023	-929.778	-104.346	6.101	0.000	0.000	802.547
5	BANK BALANCE, \$-3/3-(12737.0400000, 31851.8400000)	-2129.977	1913.222	219.654	-2.899	0.000	0.000	1657.858
6	ANNUAL SALARY, \$-1/3-(9263.6400000, 304886.6400000)	-302.621	281.487	23.184	-2.050	0.000	0.000	238.995
7	ANNUAL SALARY, \$-2/3-(304886.6400000, 487205.0400000)	137.379	-122.513	-15.816	0.950	0.000	0.000	106.667
8	ANNUAL SALARY, \$-3/3-(487205.0400000, 882650.7600000)	197.707	-185.807	-12.849	0.949	0.000	0.000	156.820
	Сумма	0.000	0.000	0.000	0.000	0.000	0.000	
	Среднее	0.000	0.000	0.000	0.000		0.000	
	Среднеквадратичное отклонение	1007.489	907.184	102.125	3.054			646.060

Рисунок 15. Экранные формы режима разделения классов на типичную и нетипичную части 1-й итерации

В данном интеллектуальном приложении по F-критерию Ван Ризбергена наиболее достоверной является СК-модель INF3 с интегральным критерием «Сумма знаний».

Из сравнения рисунков 9 и 15 мы видим, что в результате 1-й итерации достоверность этой модели :

– по критерию Ван Ризбергена выросла на 5,3%: с 0,708 до 0,746;

– по критерию L1 проф.Е.В.Луценко [24] выросла на **61,5%**: с 0,592 до 0,956, что вообще является отличным показателем.

Частотное распределение истинных и ложных положительных и отрицательных решений в зависимости от уровня сходства также стало гораздо более разумным, т.к. мы видим, что отмеченная аномалия исчезла.

Видно, что для отрицательных решений ложные отрицательные решения вообще отсутствуют. Это и есть результат работы режима разделения классов на типичную и нетипичную части, реализующего предлагаемый в данной работе интеллектуальный алгоритм (рисунок 3).

Из рисунка 15 видно также, что для положительных решений картина более сложная и включает 3 диапазона уровней сходства

1) при уровнях сходства от 0% до 10% количество ложных решений больше числа истинных;

2) при уровнях сходства от 10% до примерно 20% есть и истинные и ложные положительные решения, но число истинных решений больше числа ложных и доля истинных решений возрастает при увеличении уровня сходства;

3) при уровнях сходства выше 20% ложные положительные решения практически не встречаются.

На рисунке 16 приведен Help режима 3.4, в котором описаны меры достоверности моделей, применяемые в системе «Эйдос»/

Также мы видим, что процесс синтеза и верификации моделей продолжался заметно дольше: если исходные модели формировались и верифицировались 5 минут 34 секунды, о модели 1-й итерации уже 8 минут 16 секунд. Это связано с увеличением количества классов в два раза в связи с добавлением классов для нетипичных случаев (таблица 6).

Из рисунка 15 мы видим, что из 10000 наблюдений обучающей выборки довольно значительная часть: 2918 оказались нетипичными, т.е. не были отнесены к классам, к которым они в действительности относятся. Для этих наблюдений и были созданы собственные классы с теми же наименованиями, но с добавлением символов: _FN (это обозначение ложно-отрицательных решений по Ван Ризбергену) (таблица 6).

Таблица 6 – Классификационные шкалы и градации 1-й итерации

Код	Наименование
1	DEFAULTED?-1/4-No
2	DEFAULTED?-2/4-No_FN
3	DEFAULTED?-3/4-Yes
4	DEFAULTED?-4/4-Yes_FN

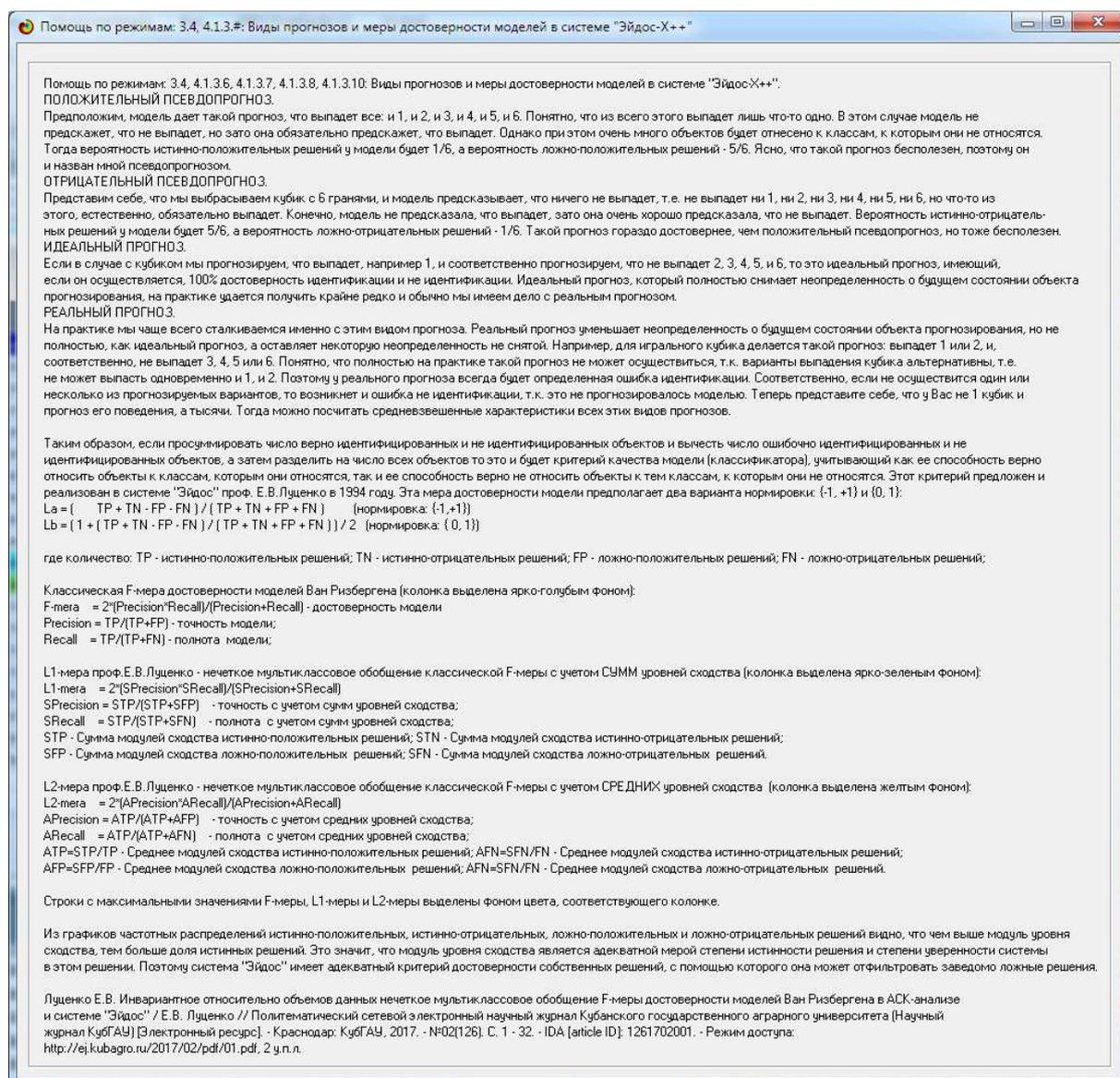


Рисунок 16. Help режима 3.4, в котором описаны меры достоверности моделей, применяемые в системе «Эйдос»

Важно иметь в виду, что корректируется (путем разделения классов на типичную и нетипичную части) всегда та модель, в которой проводилось распознавание. Если данный режим 3.7.6 запускается сразу после режима синтеза и верификации моделей 3.5 – то это будет СК-модель INF7, как последняя, в которой проводилась верификация в режиме 3.5. Если мы в режиме 3.4 выбрали другую модель для улучшения как наиболее достоверную, то надо ее задать в качестве текущей в режиме 5.6, а затем выполнить **распознавание** в этой модели в режиме 4.1.2.

Поэтому далее выполняем режим 5.6 (рисунки 1 и 17), а затем режим **распознавания** 4.1.2 в текущей модели INF3 (рисунок 18):

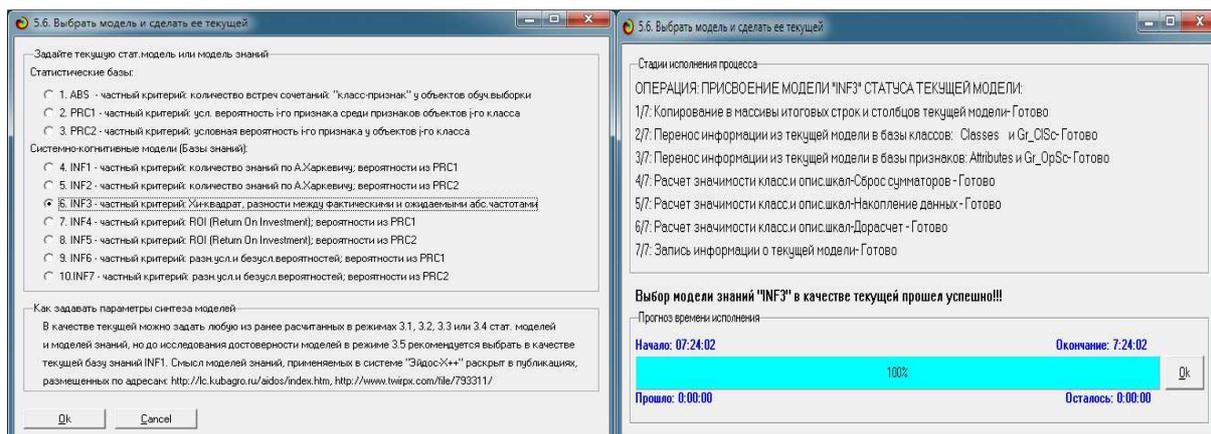


Рисунок 17. Экранные формы придания наиболее достоверной СК-модели Inf3 статуса текущей модели

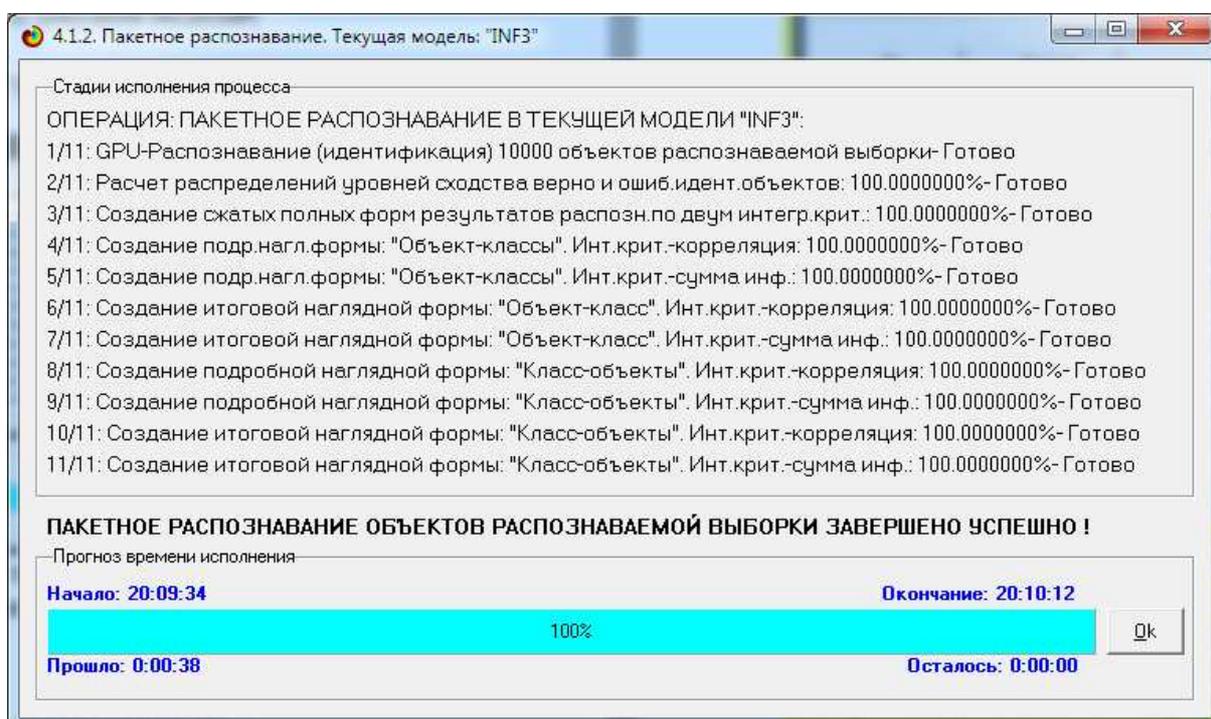


Рисунок 18. Экранная форма пакетного распознавания в текущей СК-модели Inf3

Из сравнения рисунков 12 и 18 видно, что на 1-й итерации распознавание 10000 объектов распознаваемой выборки длилось уже не 8, а 38 секунд. Это связано с увеличением числа классов с двух до 4 за счет создания классов для нетипичных случаев.

4.3.3.3. Вторая итерация. Синтез и верификация статистических и системно-когнитивных моделей. Оценка достоверности моделей. Задание текущей модели и решение в ней всех поставленных задач

Для 2-й итерации запускаем еще раз режим 3.7.6 (рисунок 19).

В результате работы данного режима мы обнаруживаем, что новых нетипичных объектов в обучающей выборки не найдено и, соответственно, новых классов не создано. Поэтому модели 2-й итерация ничем не отличается от моделей 1-й итерации, которая и является последней. То, что так и получится, было видно и до 2-й итерации уже из формы по достоверности моделей 1-й итерации, приведенной на рисунке 15. Из этой формы видно, что в 1-й итерации уже нет ложно-отрицательных решений.

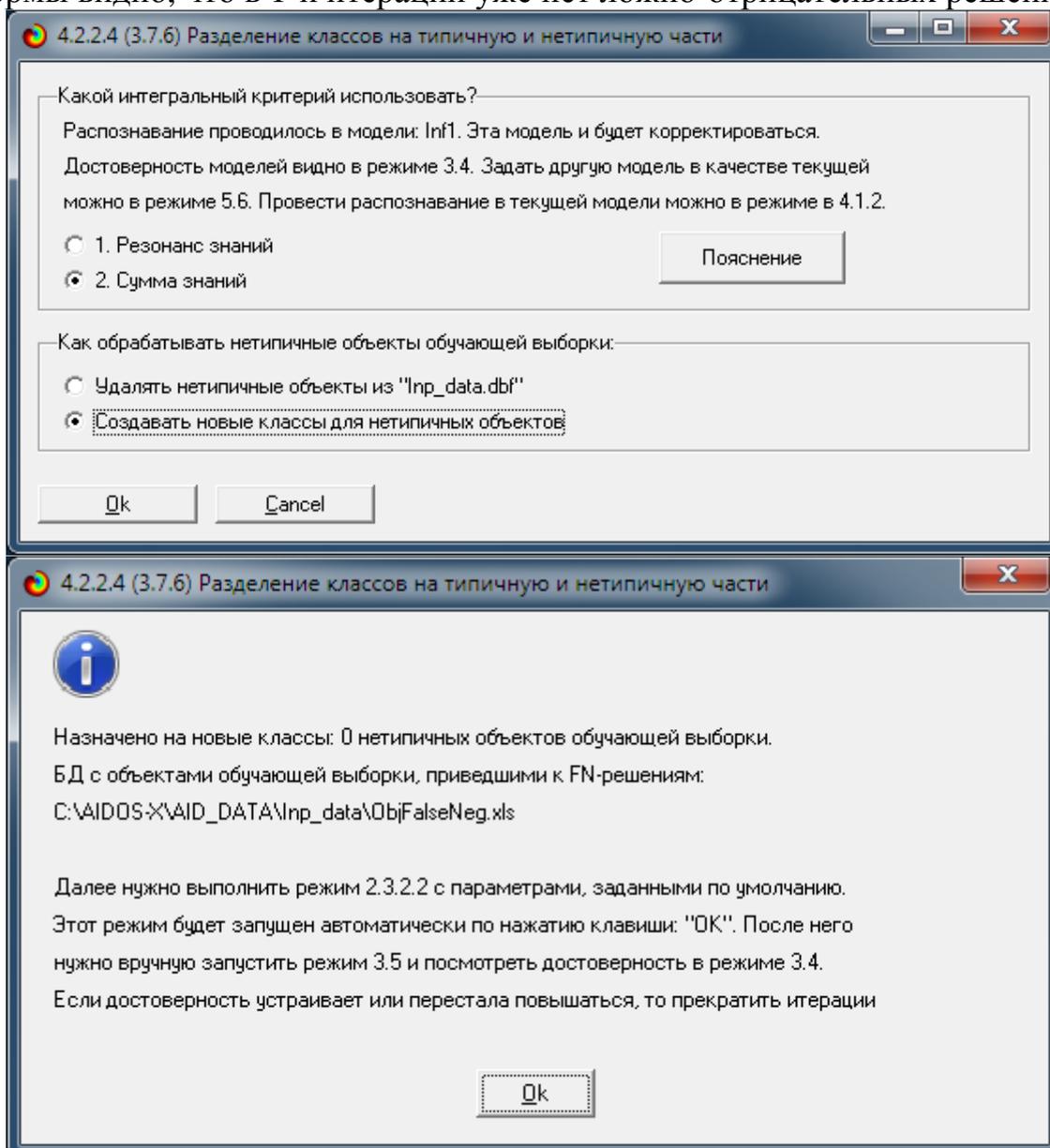


Рисунок 19. Экранные формы режима разделения классов на типичную и нетипичную части 2-й итерации

Поэтому на этом этапе выходим из интеллектуального алгоритма разделения классов на типичную и нетипичную части (рисунок 3), создаем все модели в режиме 3.5, выбираем наиболее достоверную модель INF3 в качестве текущей в режиме 5.6 и решаем в ней все последующие задачи. Этому и посвящены последующие разделы данной работы.

4.3.4. Задача 4: решение различных задач в наиболее достоверной модели

4.3.4.1. Подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация)

Далее выполняется режим 4.1.2 (см. раздел 4.3.4.1 работы [35] (рисунок 20)).

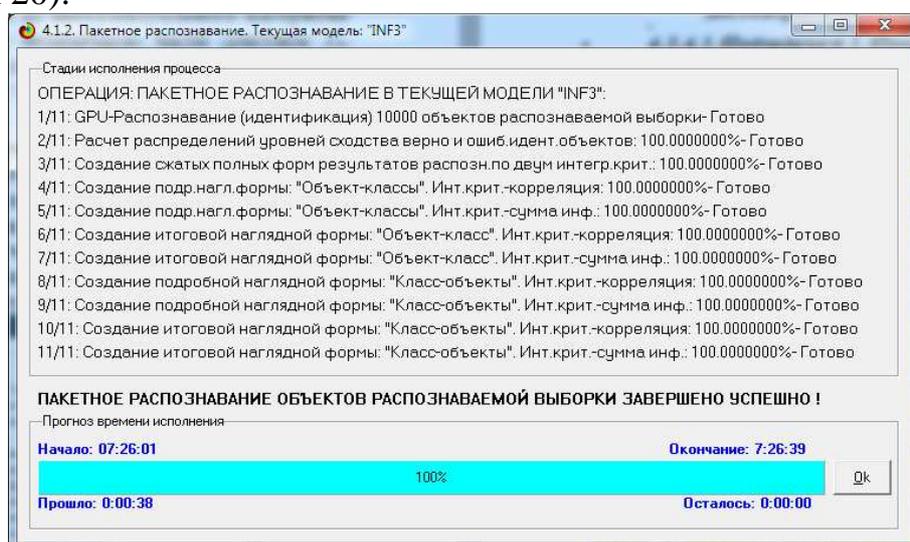


Рисунок 20. Экранные формы, которые отображают процесс решения задачи системной идентификации в текущей модели

Выводы, которые можно сделать из экранных форм на рисунке 21 см. в разделе 4.3.4.1 работы [35] (рисунок 21).

4.1.3.1. Визуализация результатов распознавания в отношении: "Объект-классы". Текущая модель: "INF3"

Код	Наименование объекта	Сходство	Ф	Сходство
1	1	99.96	v	████████████████████
2	2	57.15	v	████████████████████
3	3	92.29	v	████████████████████
4	4	57.96	v	████████████████████
5	5			
6	6			
7	7			
8	8			
9	9			
10	10			
11	11			
12	12			
13	13			
14	14	57.97	v	████████████████████
15	15	9.897	v	████████████████████
16	16	5.427	v	████████████████████
17	17	82.85	v	████████████████████
18	18			
19	19			
20	20			
21	21			
22	22			
23	23			

4.1.3.2. Визуализация результатов распознавания в отношении: "Класс-объекты". Текущая модель: "INF3"

Код	Наименование объекта	Сходство	Ф	Сходство
14	14	57.46	v	████████████████████
15	15	57.46	v	████████████████████
19	19	57.46	v	████████████████████
21	21	57.46	v	████████████████████
27	27	57.46	v	████████████████████
31	31	57.46	v	████████████████████
46	46	57.46	v	████████████████████
51	51	57.46	v	████████████████████
72	72	57.46	v	████████████████████
80	80	57.46	v	████████████████████

Рисунок 21. Выходные формы по результатам прогнозирования риска невозврата ссуды по характеристикам получателей

Достоверные прогнозы в данной модели имеют уровень сходства выше 10% по интегральному критерию «Сумма знаний» (нижнее правое окно в экранных формах на рисунке 19), т.е., по сути, результаты с более низки уровнем сходства надо просто игнорировать (см. раздел 4.3.4.1 работы[35]).

4.3.4.2. Подзадача 4.2. Поддержка принятия решений в простейшем варианте (SWOT-анализ)

Для решения данной подзадачи используем режим 4.4.8 системы «Эйдос» (см. раздел 4.3.4.2 работы [35]). Результаты приведены на рисунках 22.

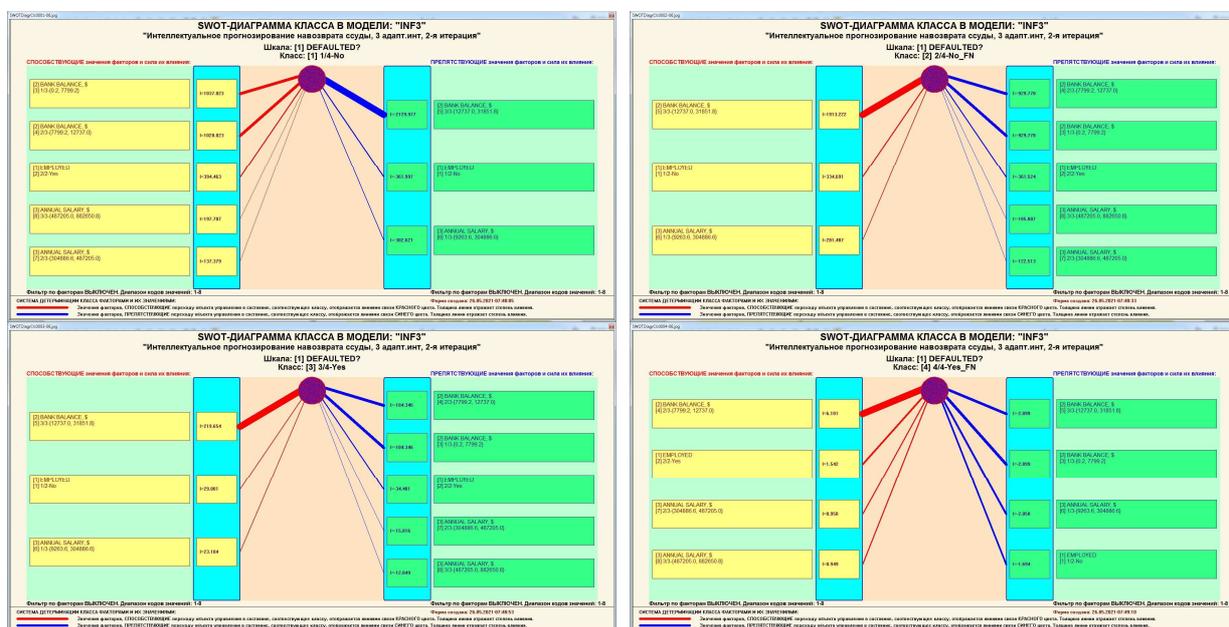


Рисунок 22. SWOT-диаграммы детерминации риска невозврата ссуды характеристиками получателей

Смысл SWOT-диаграмм поясняется в разделе 4.3.4.2 работы [35]. Впервые акт внедрения системы, являющейся далеким предшественником системы «Эйдос», был получен в 1987 году (см. 2-й акт внедрения по ссылке: <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>). Уже в той версии системы был режим, выполняющий функции, реализованные в текущей версии при генерации SWOT-диаграмм и инвертированных SWOT-диаграмм (термин автора) (рисунок 23).

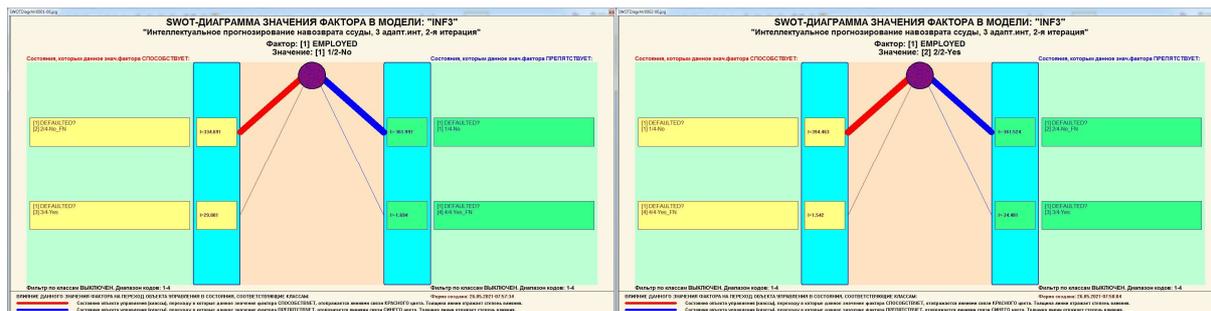


Рисунок 23. Примеры SWOT-диаграмм, отражающих силу и направление влияния различных характеристик ссудополучателя на риск невозврата полученной им ссуды

Из инвертированных SWOT-диаграмм, приведенных на рисунке 21, видно, как влияет безработный ссудополучатель или нет на риск невозврата полученной им ссуды.

Отметим, что аналогичные инвертированные SWOT-диаграммы могут быть получены для всех характеристик ссудополучателей и здесь они не приводятся только из-за ограничений на объем работы. Но они могут быть получены любым желающим, если он скачает систему «Эйдос» с сайта ее автора и разработчика проф.Е.В.Луценко по ссылке: http://lc.kubagro.ru/aidos/_Aidos-X.htm, установит ее на своем компьютере, а затем в режиме 1.3 установит интеллектуальное облачное Эйдос-приложение №293, просчитает модели в режиме 3.5 и перейдет в режим 4.4.9.

Преимущества автоматизированного SWOT-анализа, реализованного в АСК-анализе и системе «Эйдос» перед традиционным SWOT-анализом подробнее описан в разделе 4.3.4.2 работы [35].

4.3.4.3. Подзадача 4.2. Развитый алгоритм принятия решений

В предыдущем разделе кратко описан лишь очень упрощенный вариант принятия решений. В развитой форме, полностью поддерживаемой системой «Эйдос», он приведен в данном разделе (рисунок 24).

Подробное пояснение данного алгоритма (который в принципе и так вполне понятен) не входит в задачи данной работы и дано в других работах автора, например [20], а также в видеозанятиях:

– в Пермском национальном университете:

<https://bigbluebutton.pstu.ru/b/w3y-2ir-ukd-bqn>

– в Кубанском государственном университете и Кубанском государственном аграрном университете:

<https://disk.yandex.ru/d/knISAD5qzV83Ng?w=1>;

https://studio.youtube.com/channel/UC_QF84d8SCaWxsnXnexNFzg.

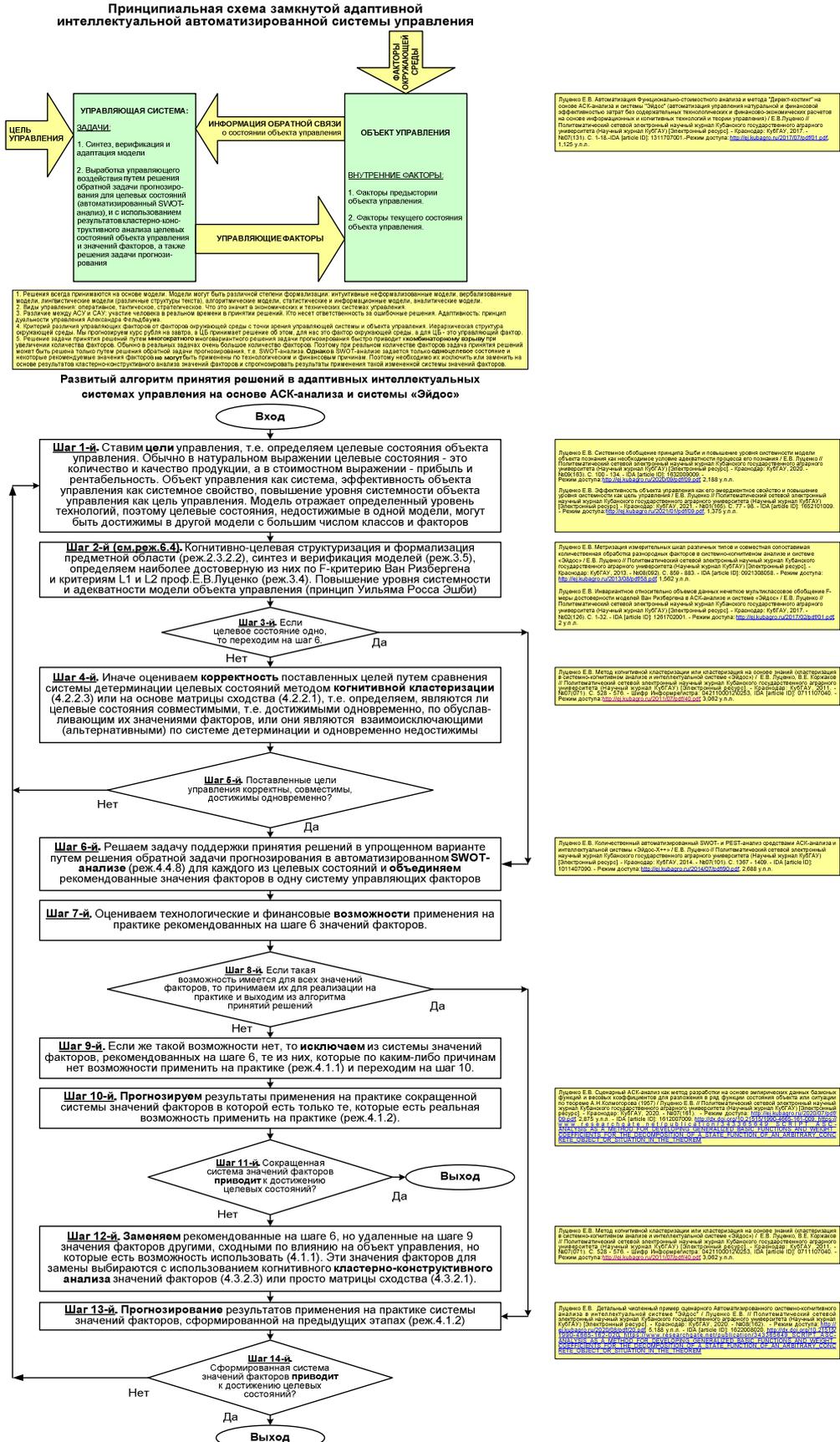


Рисунок 24. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

4.3.4.4. Подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели

О решении подзадачи 4.3 подробнее можно ознакомиться в разделе 4.3.4.4. работы [35].

4.3.4.4.1. Когнитивные диаграммы классов

Эти диаграммы отражают сходство/различие классов. Мы получаем их в режимах 4.2.2.1 и 4.2.2.2 (рисунки 25, 26) [35].

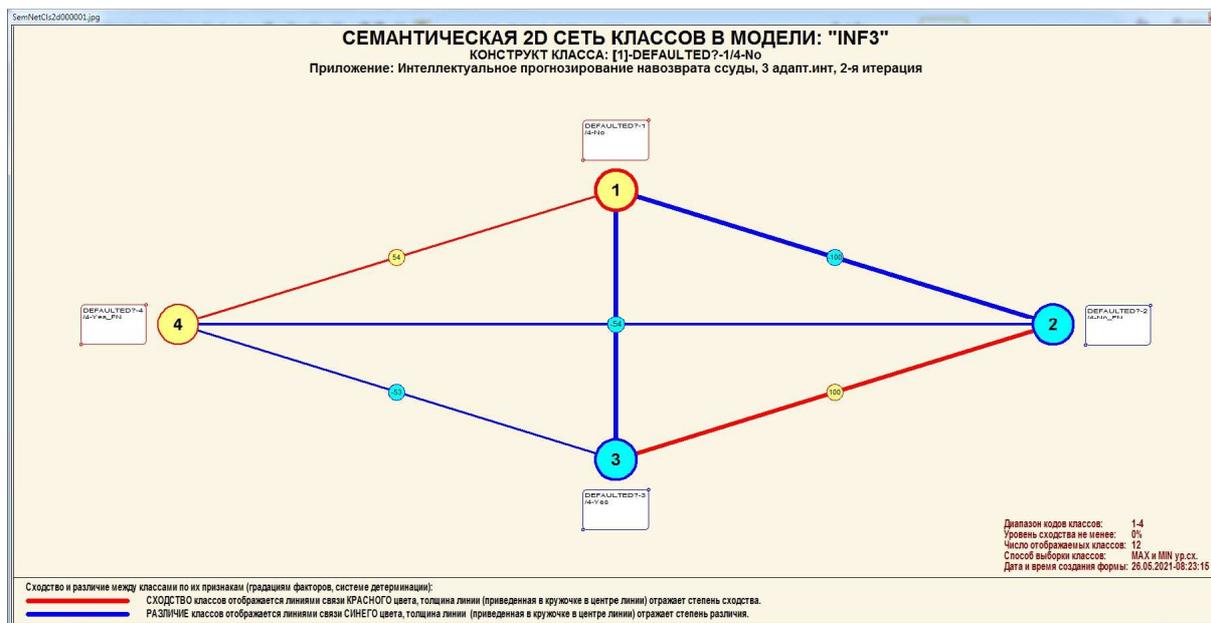


Рисунок 25. Когнитивная диаграмма классов, отражающая сходство/различие рисков невозврата ссуды по связанным с ними значениям характеристик ссудополучателей

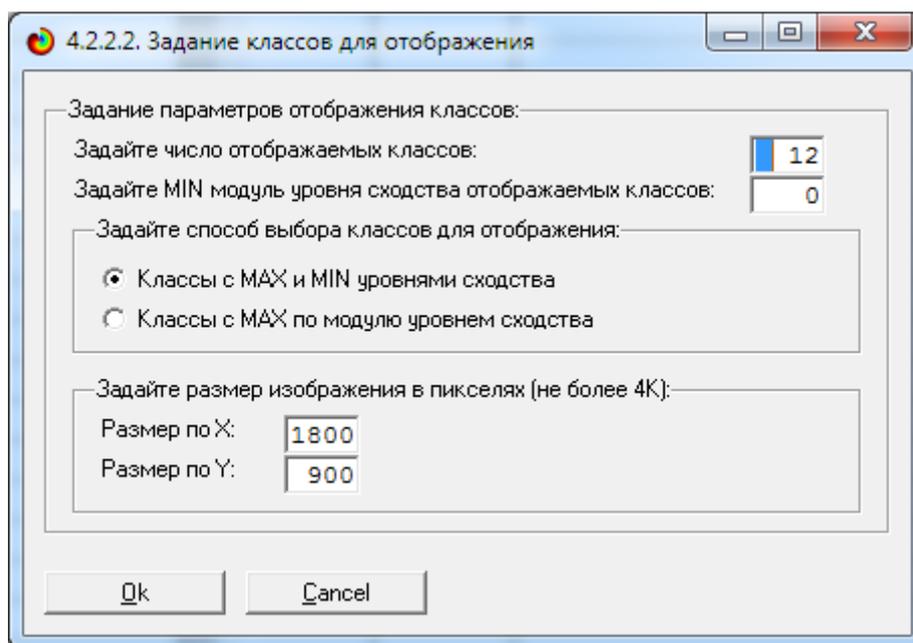


Рисунок 26. Диалоговое окно управления параметрами формирования и вывода изображения когнитивной диаграммы классов

4.3.4.4.2. Агломеративная когнитивная кластеризация классов

Для получения агломеративных дендрограмм классов выполняем режимы 4.2.2.1 и 4.2.2.3 системы «Эйдос» (рисунки 27, 28).

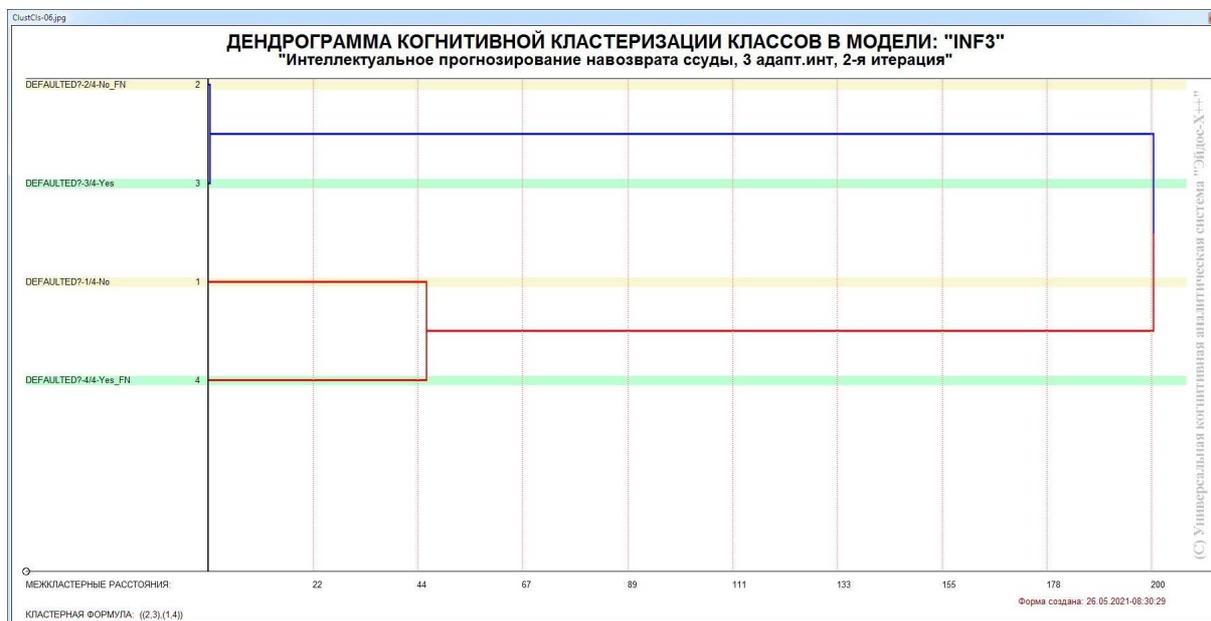


Рисунок 27. Дендрограмма когнитивной агломеративной кластеризации, отражающая сходство/различие рисков невозврата ссуды по системе их детерминации характеристиками ссудополучателей

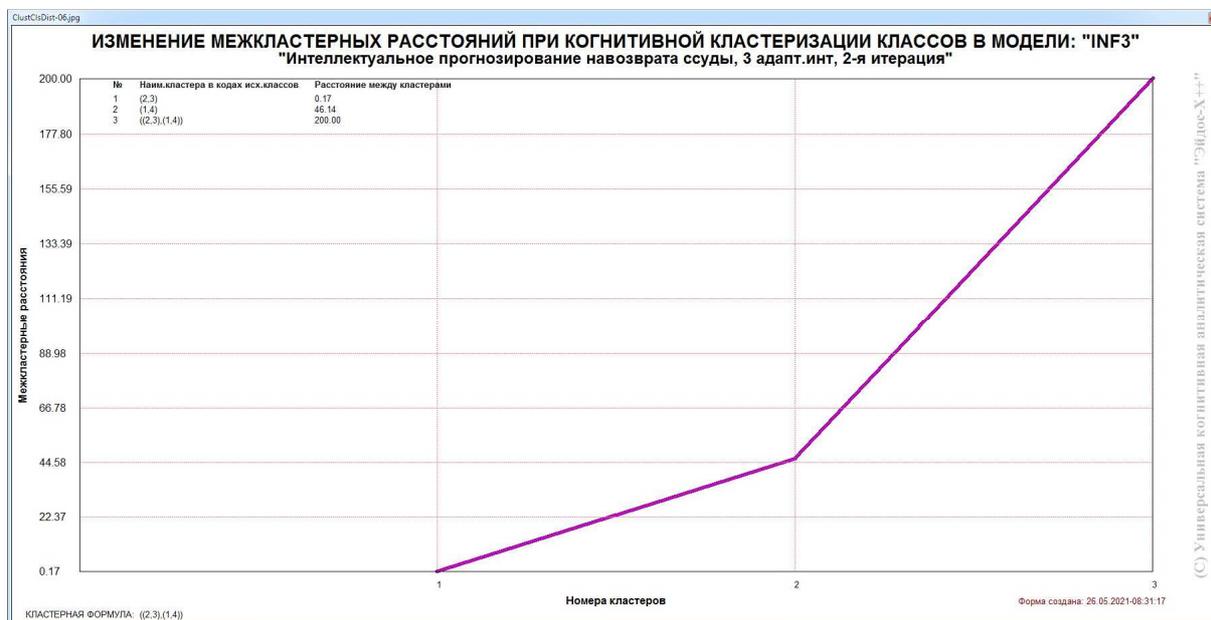


Рисунок 28. График изменения межкластерных расстояний

Необходимые пояснения даны в разделе 4.3.4.4.2 работы [35].

4.3.4.4.3. Когнитивные диаграммы значений факторов

Для получения этих диаграмм запускаем режимы 4.3.2.1 и 4.3.2.2 системы «Эйдос» (рисунки 29, 30).

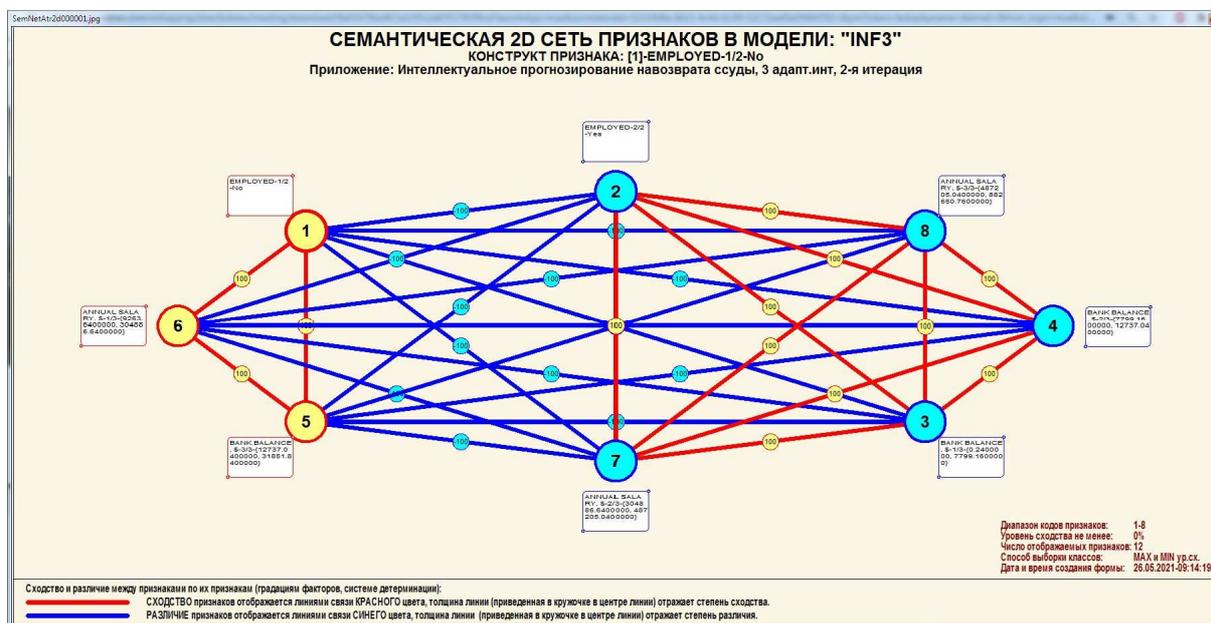


Рисунок 29. Сходство/различие характеристик ссудополучателей по их влиянию на риск невозврата ссуды

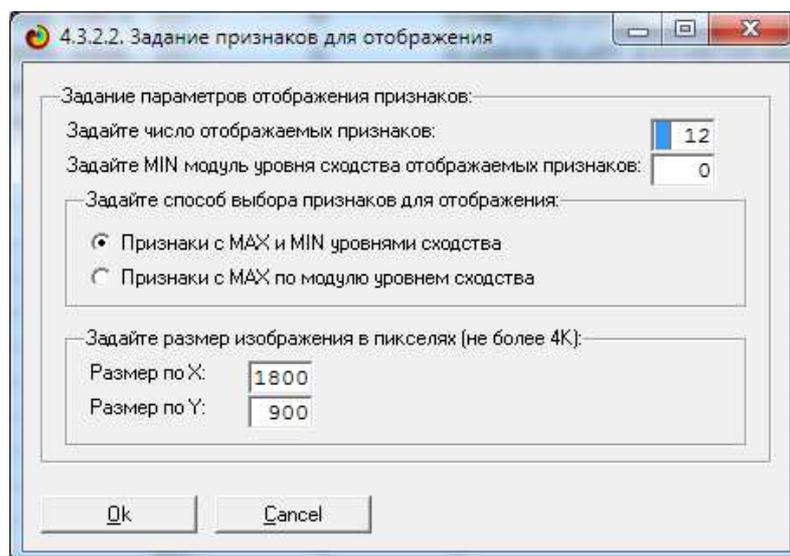


Рисунок 30. Параметры отображения когнитивной диаграммы, приведенной на рисунке 27

Необходимые пояснения даны в разделе 4.3.4.4.3 работы [35].

4.3.4.4.4. Агломеративная когнитивная кластеризация значений факторов

Для получения агломеративных дендрограмм значений факторов выполняем режимы 4.3.2.1 и 4.3.2.3 системы «Эйдос» (рисунки 31, 32).

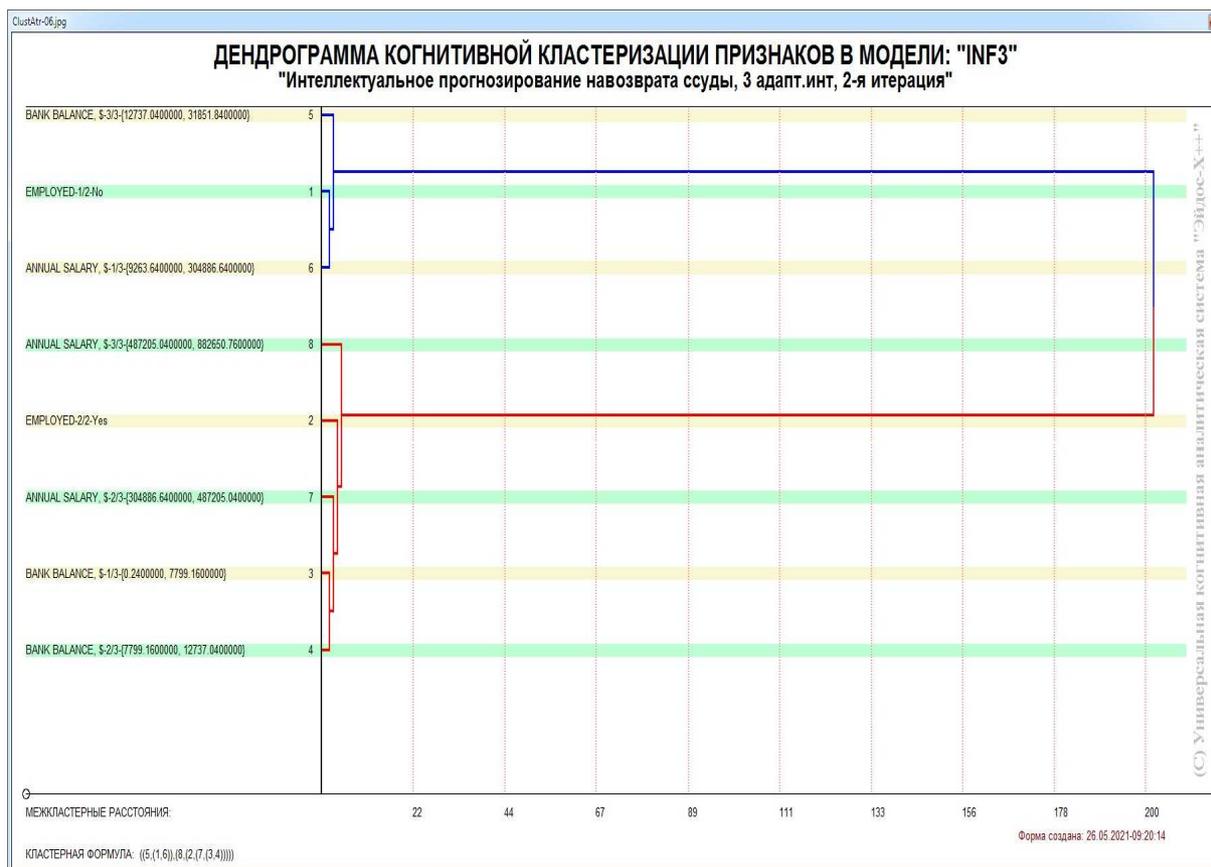


Рисунок 31. Дендрограмма агломеративной когнитивной кластеризации характеристик ссудополучателей по их влиянию на риск невозврата ссуды

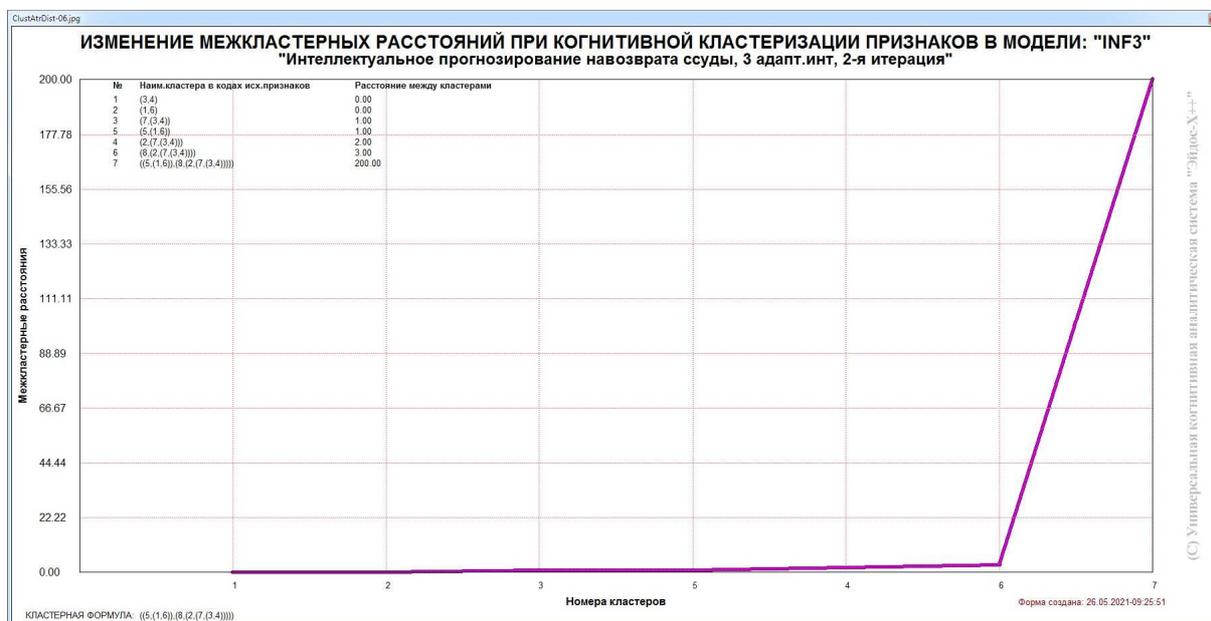


Рисунок 32. График изменения межкластерных расстояний при когнитивной кластеризации значений факторов

Необходимые пояснения даны в разделе 4.3.4.4.4 работы [35].

4.3.4.4.5. Нелокальные нейроны и нелокальные нейронные сети

На рисунке 33 приведён пример нелокального нейрона, а на рисунке 34 – фрагмент одного слоя нелокальной нейронной сети [27]:

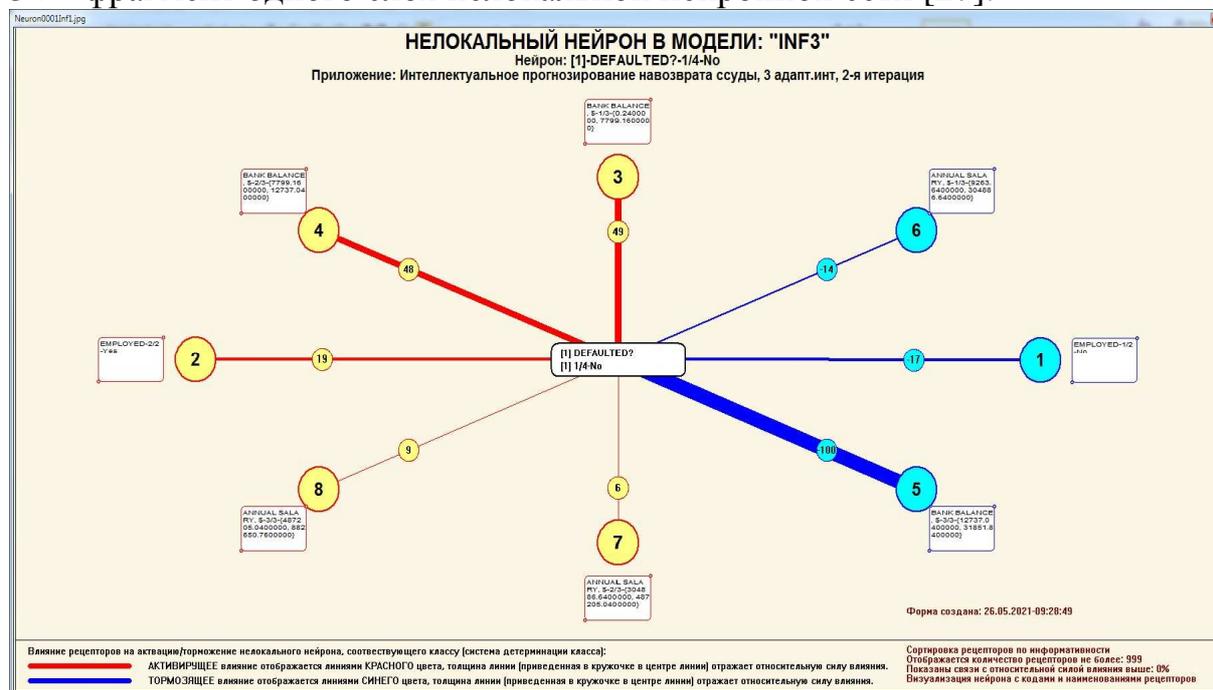


Рисунок 33. Пример нелокального нейрона, отражающего силу и направление влияния значений характеристик ссудополучателей на риск невозврата полученной ими ссуды

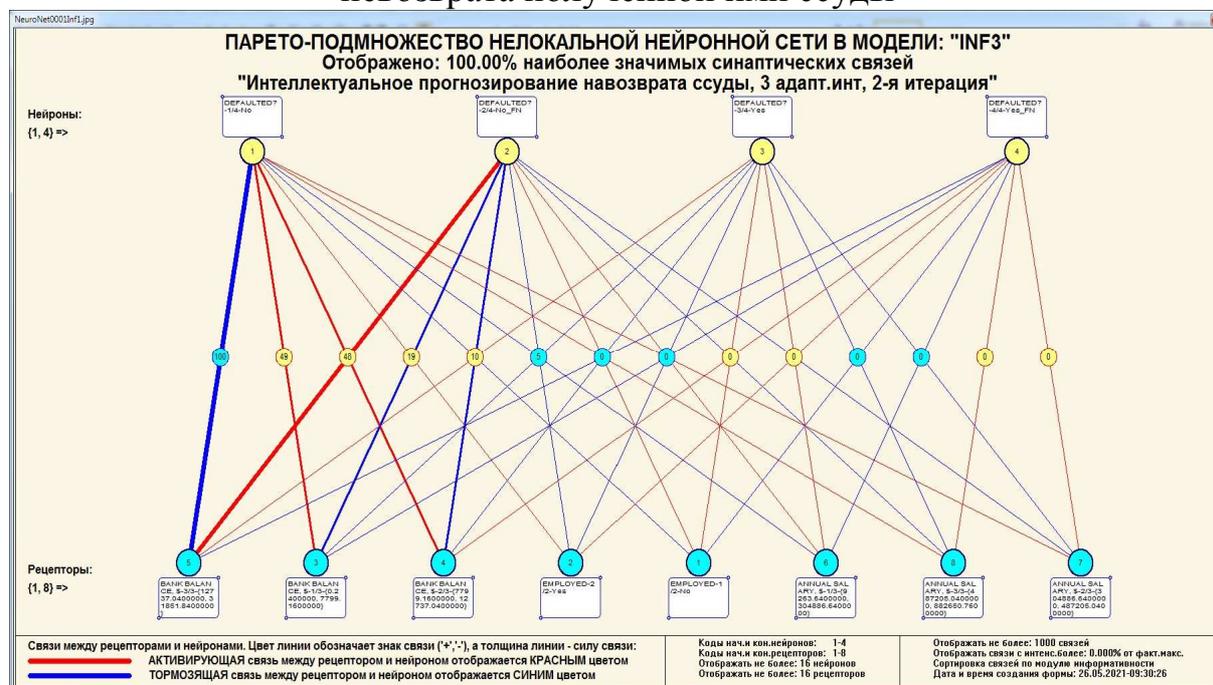


Рисунок 34. Один слой нелокальной нейронной сети, отражающий силу и направление влияния характеристик ссудополучателей на риск невозврата полученной ими ссуды

Необходимые пояснения даны в разделе 4.3.4.4.5 работы [35].

4.3.4.4.6. 3d-интегральные когнитивные карты

На рисунке 32 приведена 3d-интегральная когнитивная карта в СК-модели Inf3.

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивных диаграмм классов и значений факторов, отображенных соответственно на рисунках 25 и 29, и одного слоя нейронной сети, приведенного на рисунке 34.

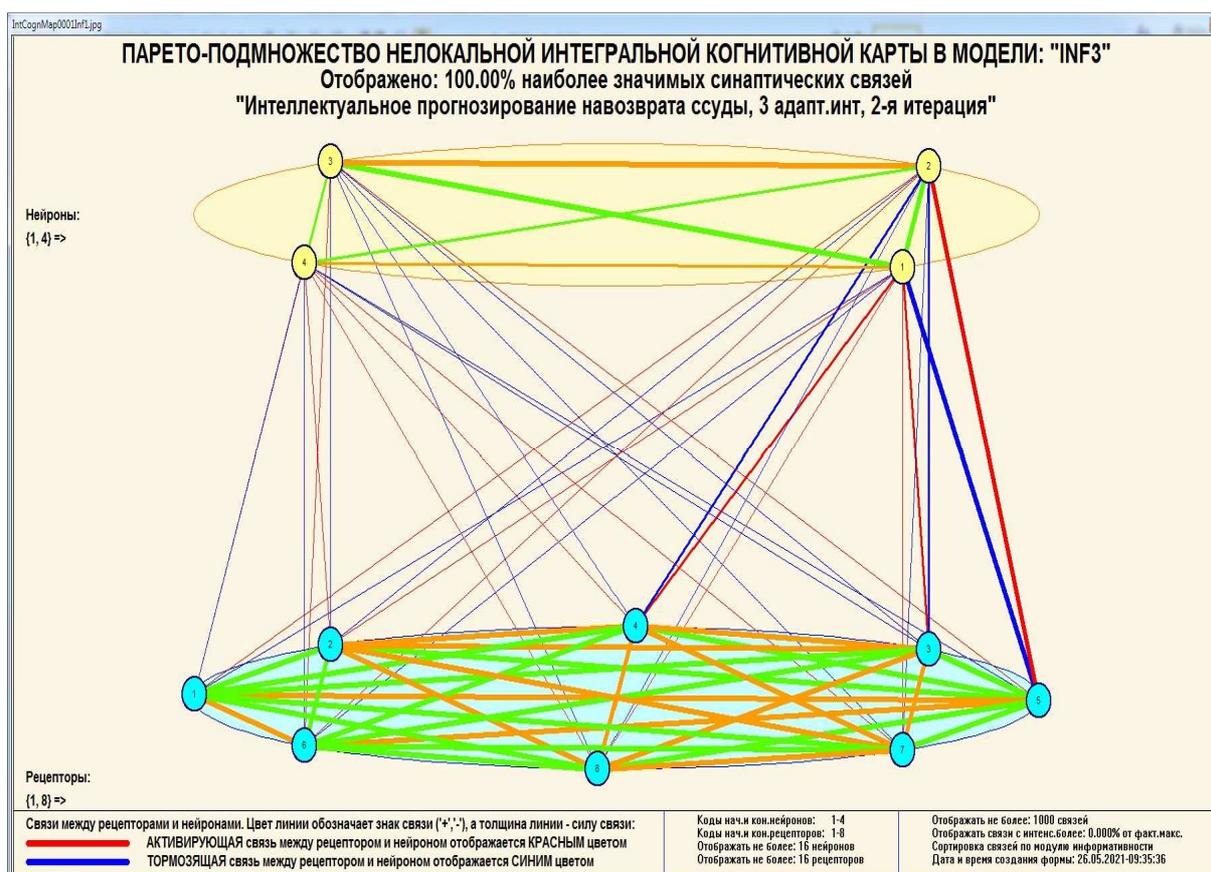


Рисунок 35. 3d-интегральная когнитивная карта в СК-модели Inf3

4.3.4.4.7. Когнитивные функции

Вместо описания того, что представляют собой когнитивные функции, приведем help соответствующего режима системы «Эйдос» (рисунок 36) и сошлемся на работы, в которых описан этот подход [22]⁸.

⁸ Подборка публикаций проф.Е.В.Луценко & С^o по когнитивным функциям:
http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm.

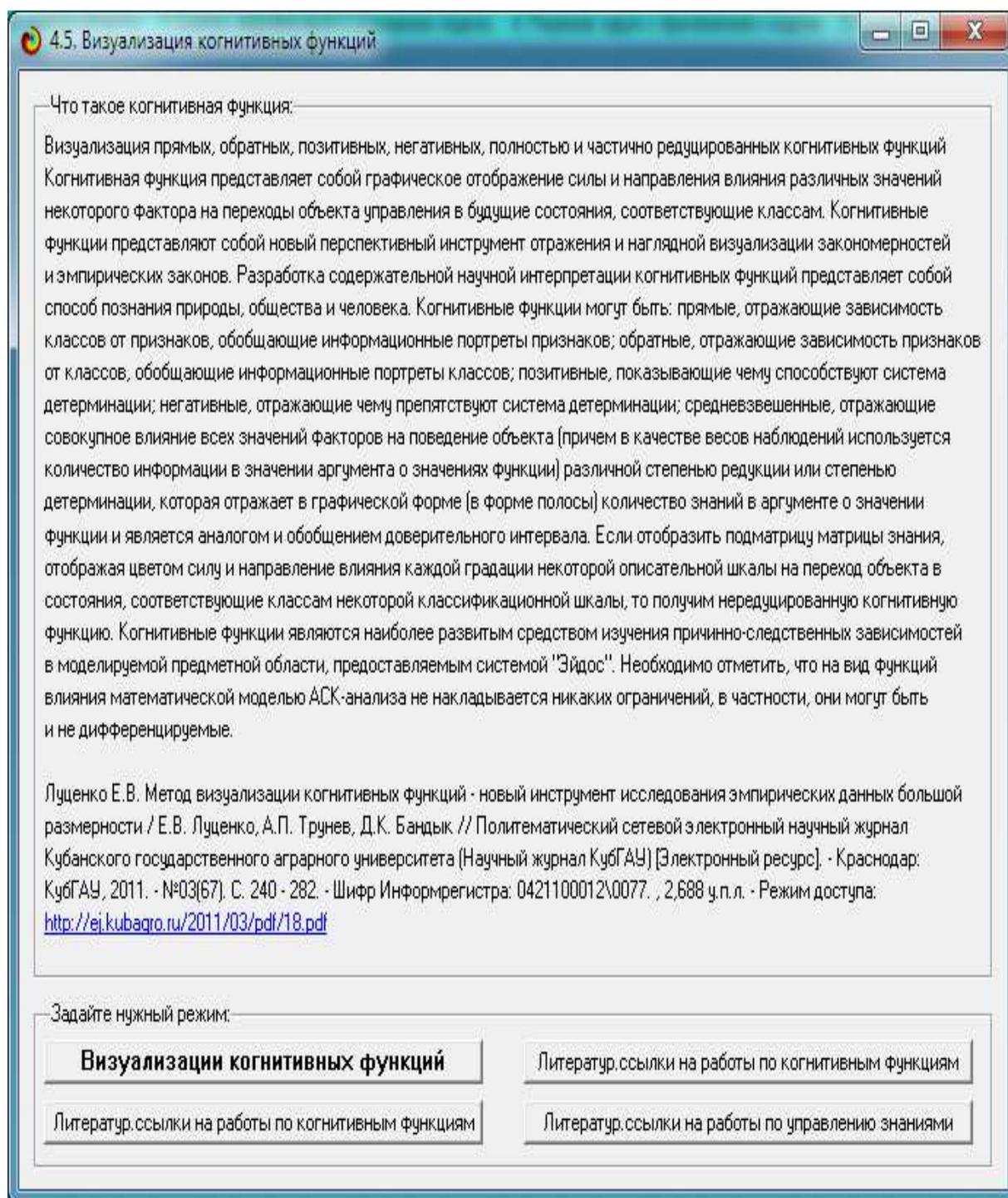
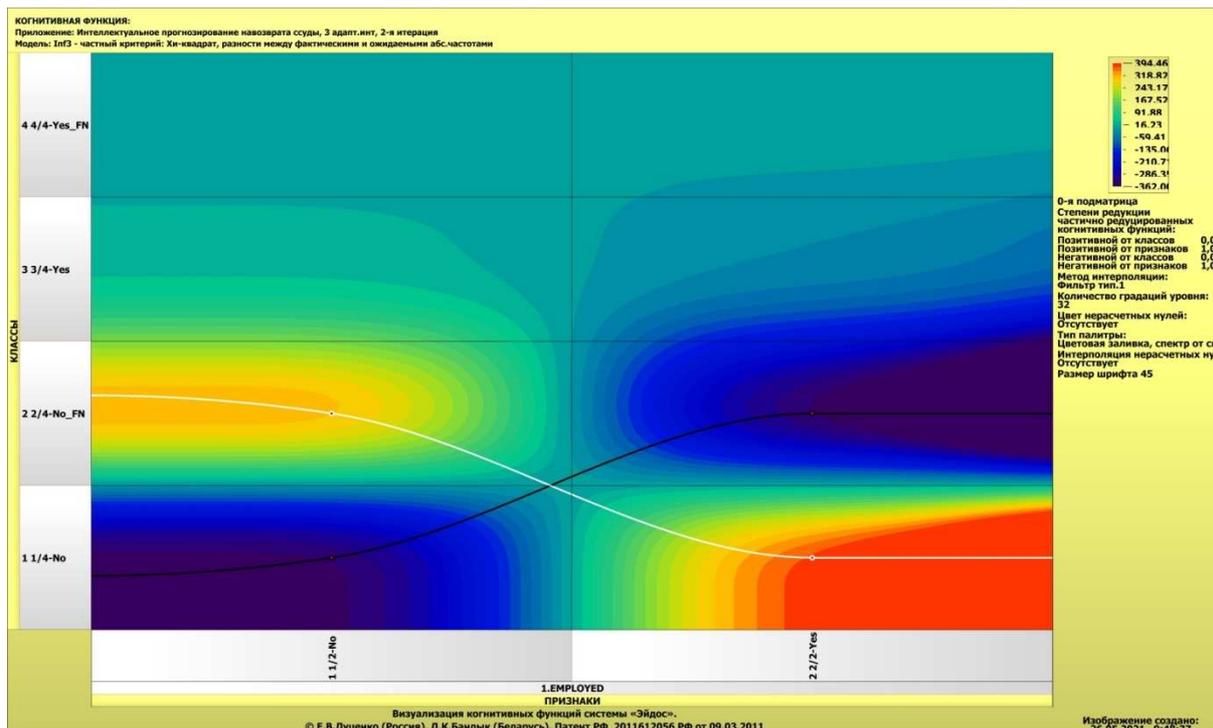


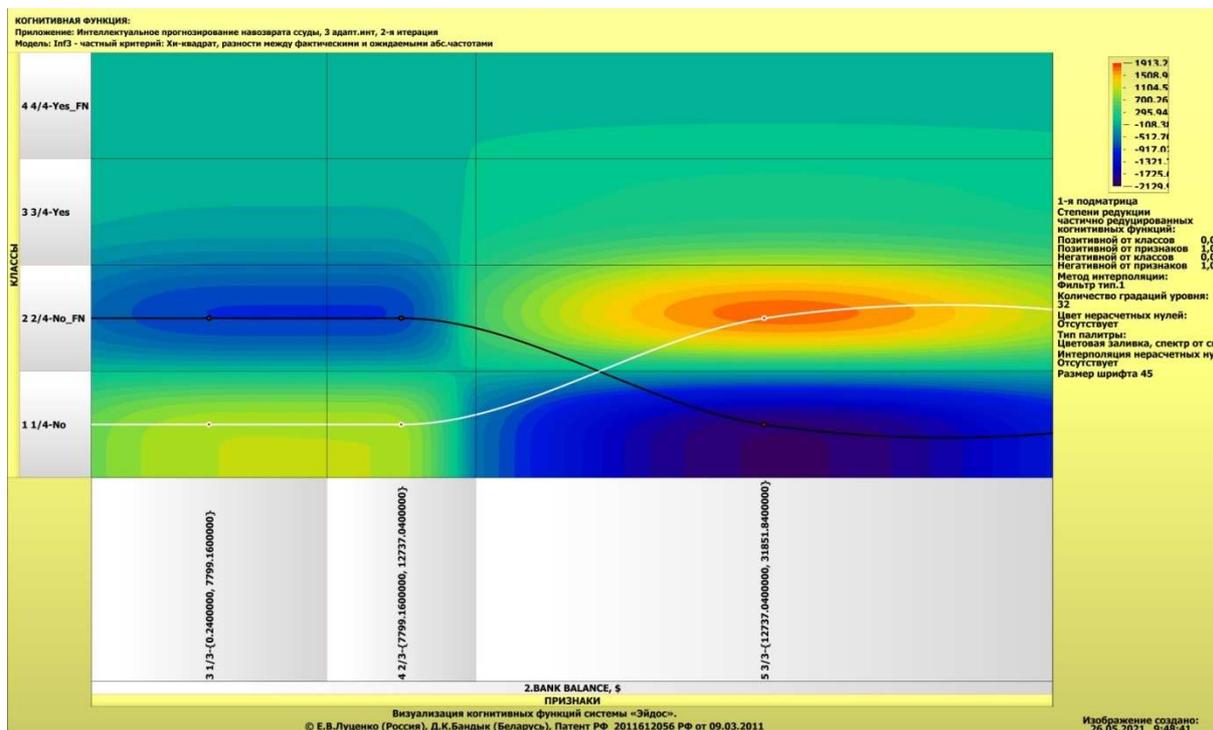
Рисунок 36. Help режима визуализации когнитивных функций

Необходимые пояснения по когнитивным функциям (рисунок 37) даны в разделе 4.3.4.4.7 работы [35].

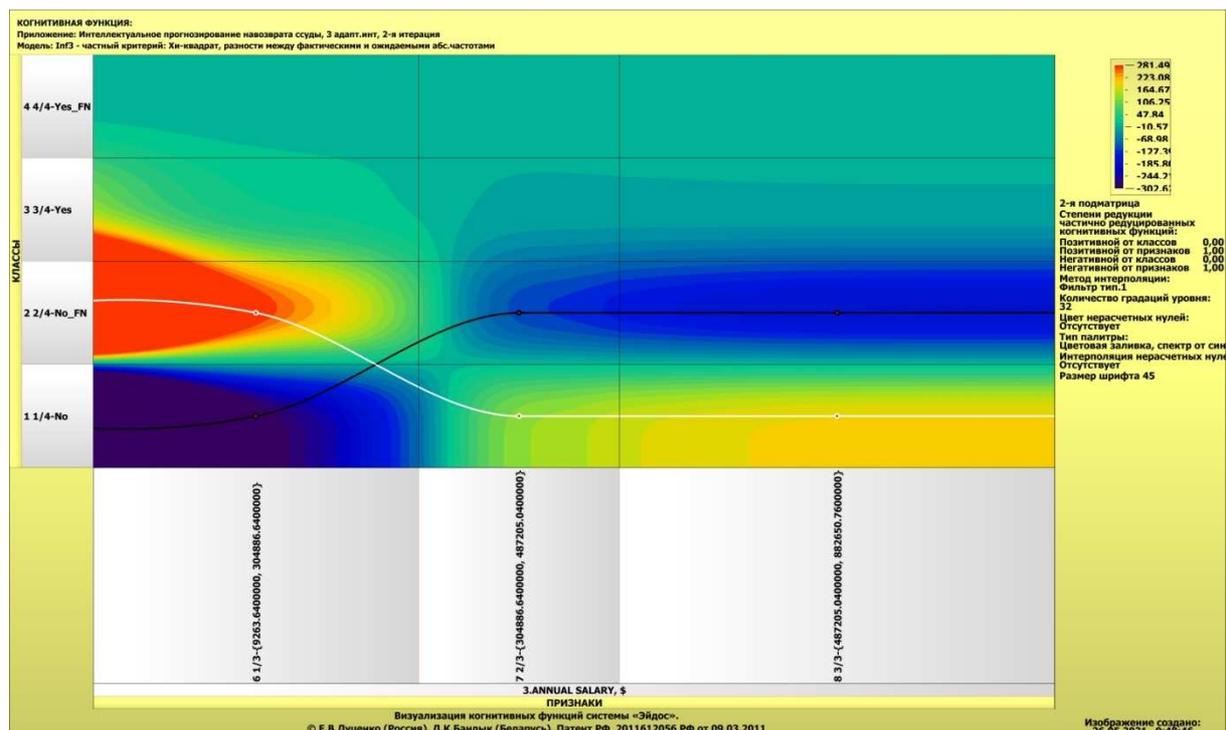
В них есть несколько интересных неожиданных моментов, требующих специальной содержательной интерпретации. Эта содержательная интерпретация является делом специалистов по субсидированию и не входит в задачи данной работы.



А)



Б)



В)

Рисунок 37. Примеры когнитивных функций в СК-модели INF3, отражающих силу и направление влияния значений характеристик ссудополучателей на риск невозврата полученной ими ссуды

4.3.4.4.8. Сила и направление влияния значений факторов на принадлежность к классам

Для изучения ценности значений факторов (признаков) для решения задачи прогнозирования рисков субсидирования запускаем режим 3.7.5 системы «Эйдос» (рисунок 38, таблица 7).

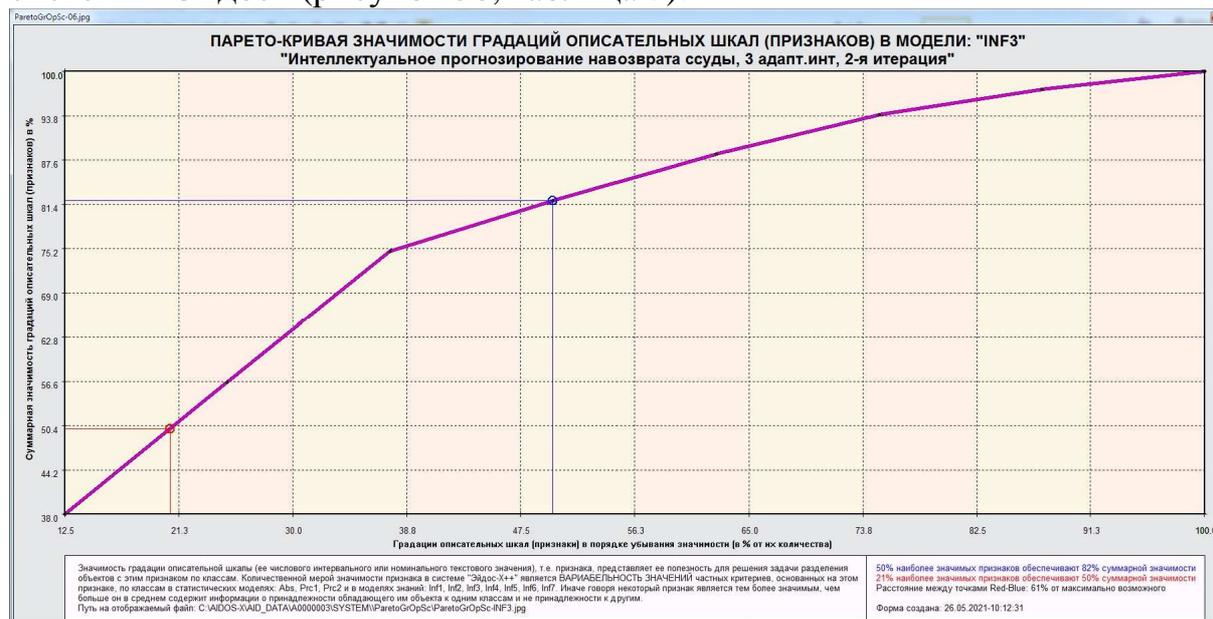


Рисунок 38. Парето-кривая значимости градаций описательных шкал

Таблица 7 – Парето-таблица значимости градаций описательных шкал, т.е. сила влияния значений характеристик ссудополучателя на риск невозврата ссуды в СК-модели INF3

№	Код	Наименование	Код шкалы	Значимость, %	Значимость нараст. итогом, %
1	5	BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000}	2	37,990	37,990
2	3	BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000}	2	18,478	56,468
3	4	BANK BALANCE, \$-2/3-{7799.1600000, 12737.0400000}	2	18,390	74,858
4	2	EMPLOYED-2/2-Yes	1	7,094	81,952
5	1	EMPLOYED-1/2-No	1	6,534	88,486
6	6	ANNUAL SALARY, \$-1/3-{9263.6400000, 304886.6400000}	3	5,477	93,962
7	8	ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000}	3	3,594	97,556
8	7	ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}	3	2,444	100,000

Данные, приведенные на рисунке 38 и в таблице 7, находится в XLS-файлах, созданных в режиме 3.7.5. Информация об этом содержится в экранной форме на рисунке 39:

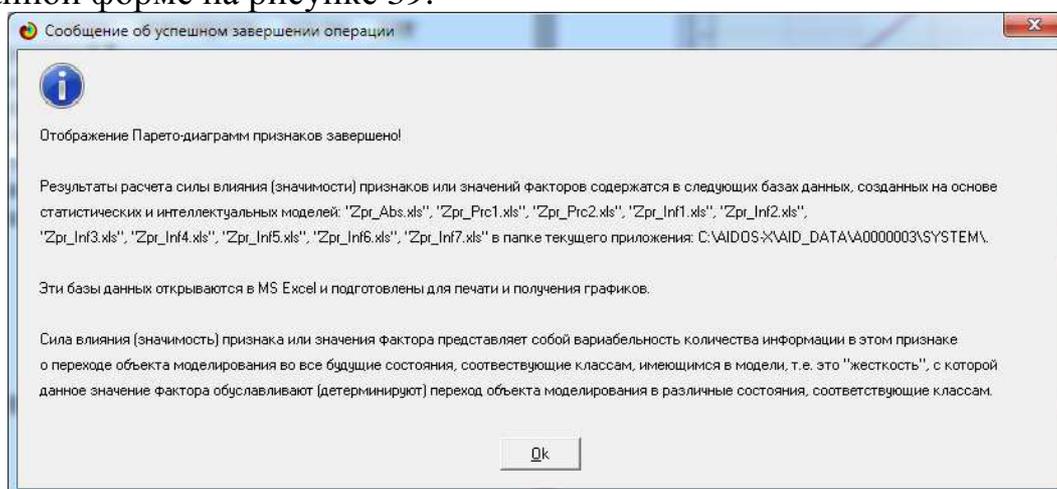


Рисунок 39. Информация о XLS-файлах

Необходимые пояснения даны в разделе 4.3.4.4.8 работы [35].

Из таблицы 7 видно, что наиболее сильное влияние на риск невозврата ссуды оказывают следующие характеристики ссудополучателя:

- BANK BALANCE, \$-3/3-{12737.0400000, 31851.8400000};
- BANK BALANCE, \$-1/3-{0.2400000, 7799.1600000};

а наиболее низкое:

- ANNUAL SALARY, \$-3/3-{487205.0400000, 882650.7600000};
- ANNUAL SALARY, \$-2/3-{304886.6400000, 487205.0400000}.

При этом сила влияния наиболее и наименее значимых значений факторов классов отличается в 15 раз, что очень существенно.

Ценность же характеристик ссудополучателей (всей описательной шкалы или фактора), для решения этих задач можно количественно оценивать как среднее от ценности значений этого параметра (таблица 8). Это можно сделать в режиме 3.7.4 (рисунок 40):

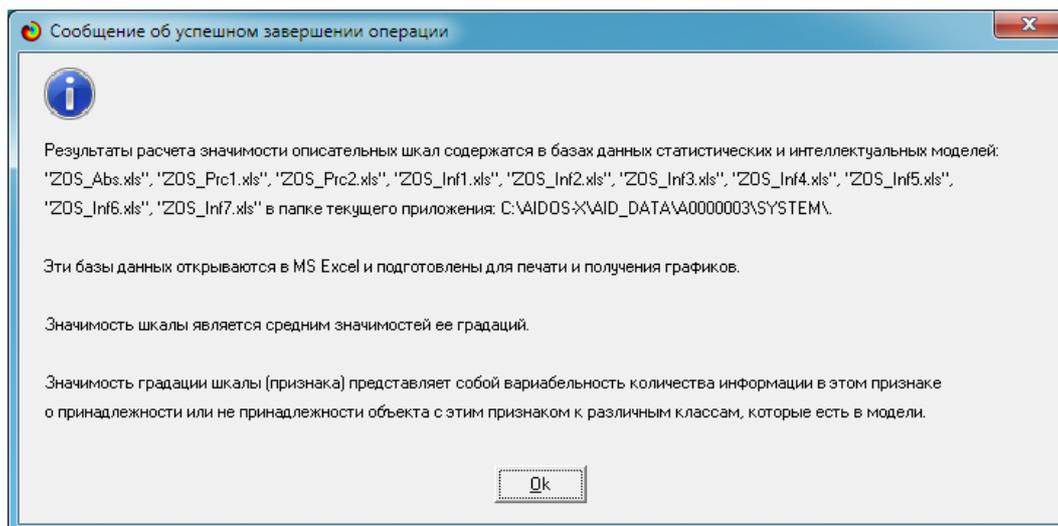


Рисунок 40. Информация о XLS-файлах

Таблица 8 – Парето-таблица значимости описательных шкал, т.е. сила влияния характеристик ссудополучателя на риск невозврата ссуды в СК-модели INF3

№	Код	Название описательной шкалы	Значи_ мость, %	Значимость нарастающим итогом
1	2	BANK BALANCE, \$	70,0829613	70,0829613
2	1	EMPLOYED	19,1372199	89,2201812
3	3	ANNUAL SALARY, \$	10,7798188	100,0000000

Из таблицы 8 видно, что наиболее сильное влияние на прогнозирование риска невозврата ссуды оказывает баланс счета в банке, который дает более 2/3 суммарной значимости, в три с лишним раза меньшее влияние на это оказывает то, является ли ссудополучатель работающим или безработным, и еще в два раза меньшее влияние оказывает заработная плата за год.

При этом сила влияния наиболее и наименее значимых факторов классов отличается в 7 раз, что очень существенно.

4.3.4.4.9. Степень детерминированности классов значениями обуславливающих их факторов

Для изучения степени детерминированности классов значениями характеристик ссудополучателя запускаем режим 3.7.3 системы «Эйдос» (таблица 9) (рисунок 41).



Рисунок 41. Парето-кривая степени детерминированности классов

Эта информация есть и в табличной форме (рисунок 42, таблица 9):

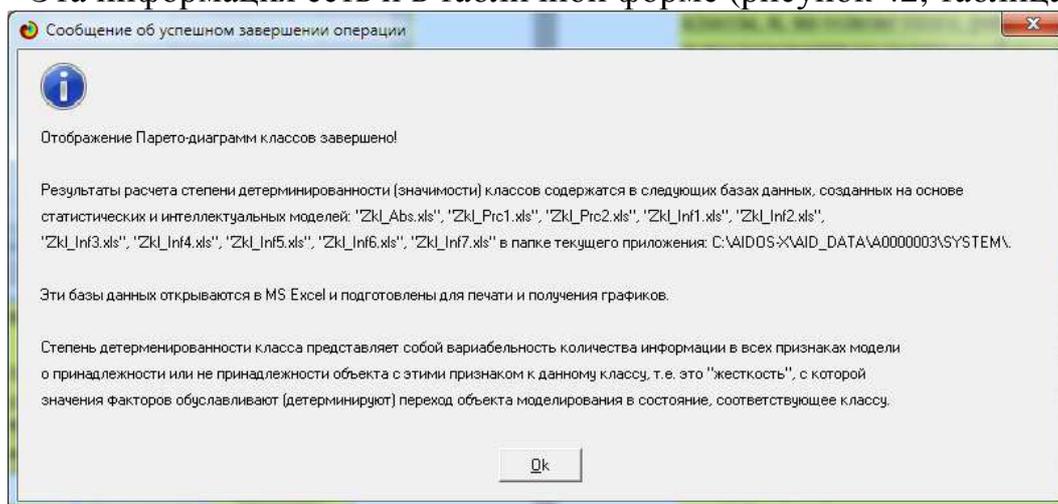


Рисунок 42. Информация о XLS-файлах

Таблица 9 – Парето-таблица степеней детерминированности (обусловленности) классов в СК-модели INF3

№	Код класса	Название класса	Значимость, %	Значимость нарастающим итогом
1	1	DEFAULTED?-1/4-No	49,879	49,879
2	2	DEFAULTED?-2/4-No_FN	44,913	94,793
3	3	DEFAULTED?-3/4-Yes	5,056	99,849
4	4	DEFAULTED?-4/4-Yes_FN	0,151	100,000

Из таблицы 9 мы видим, что значения характеристик ссудополучателей наиболее сильно детерминируют (обуславливают) такие

- DEFAULTED?-1/4-No;
- DEFAULTED?-2/4-No_FN;

а наименее сильно:

- DEFAULTED?-3/4-Yes;
- DEFAULTED?-4/4-Yes_FN.

При этом степень детерминированности наиболее и наименее детерминированных классов отличается в 6622 раз, что очень существенно.

Чем выше степень детерминированности класса степени риска невозврата ссуды значениями характеристик ссудополучателя, тем легче определить этот класс по этим характеристикам.

4.3.5. Выводы

Со всеми моделями, созданными в данной статье, можно ознакомиться установив облачное Эйдос-приложение №293 в режиме 1.3 системы «Эйдос». Саму систему можно бесплатно скачать с сайта ее автора и разработчика проф.Е.В.Луценко по ссылке: http://lc.kubagro.ru/aidos/_Aidos-X.htm (см. раздел 4.3.5 работы [35]).

4.4. Выводы

По результатам данного раздела делается обоснованный вывод о том, что поставленные в работе задачи решены, цель достигнута, поставленная проблема успешно решена, эффективность внедрения подтверждена модельным численным примером, основанным на реальных данных портала Kaggle.

На основании этого делается обоснованный вывод том, что научное решение, ранее доведенное до статуса новаций, в данном разделе до инновационного уровня, т.е. полностью готово для широкого применения на практике.

Заключение

Объект исследования – прогнозирование риска невозврата ссуды.

Предмет исследования – прогнозирование рисков невозврата ссуды с применением интеллектуального итерационного алгоритма учета нетипичных случаев.

Проблема – возникающая при прогнозировании риска невозврата ссуды – ложно-отрицательные решения, возникающие, когда модель не относит ситуацию к классу (обобщенной категории), к которой она в действительности относится.

Целью данной работы является решение поставленной проблемы путем разработки алгоритма, обеспечивающего формирование обобщенных образов классов не только на основе типичных относящихся к ним наблюдений, но и на основе исключений, т.е. нетипичных случаев.

Путем декомпозиции цели получим последовательность **задач**, решение которых является этапами достижения цели: задача 1: когнитивная структуризация предметной области; задача 2: подготовка

исходных данных и формализация предметной области; задача 3: синтез и верификация статистических и системно-когнитивных моделей. Улучшение наиболее достоверной модели путем применения итерационного алгоритма разделения обобщенных образов классов на типичную и нетипичную части; задача 4: решение различных задач в наиболее достоверной модели: подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация); подзадача 4.2. Поддержка принятия решений; подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели.

Все поставленные задачи были решены.

Задача 1. В качестве классификационной шкалы выбран риск невозврата ссуды, в качестве описательных шкал: характеристики ссудополучателя работает ли он или нет, какой у него счет в банке и годовой доход.

Задача 2. Разработаны градации классификационных и описательных шкал и с их помощью закодированы исходные данные, в результате чего получена обучающая выборка.

Задача 3. Осуществлен синтез и верификация 3 статистических и 7 системно-когнитивных моделей, выбрана наиболее достоверная из них. Ей оказалась СК-модель INF3. Затем было повышено качество наиболее достоверной модели путем применения итерационного алгоритма разделения обобщенных образов классов на типичную и нетипичную части. В результате ее достоверность по L1 критерию повысилась на 61%.

Задача 4. В наиболее достоверной модели: были решены следующие подзадачи: подзадача 4.1. Прогнозирование (диагностика, классификация, распознавание, идентификация); подзадача 4.2. Поддержка принятия решений; подзадача 4.3. Исследование моделируемой предметной области путем исследования ее модели: когнитивные диаграммы классов и значений факторов, агломеративная когнитивная кластеризация классов и значений факторов, нелокальные нейроны и нейронные сети, 3d-интегральные когнитивные карты, когнитивные функции), исследование силы и направления влияния факторов и степени детерминированности классов, обуславливающими их значениями факторов.

На основании всего этого можно сделать обоснованный вывод о том, что поставленные в работе задачи решены, цель достигнута, поставленная проблема успешно решена. Достигнутая в работе цель и решенная проблема соответствуют поставленным.

Литература

1. Lutsenko E.V. Methods of writing scientific papers, logic and the manner in which scientific statements // February 2021, DOI: [10.13140/RG.2.2.23546.41920](https://doi.org/10.13140/RG.2.2.23546.41920), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), <https://www.researchgate.net/publication/349039044>
2. Луценко Е. В. Научная публицистика : учебник / Е. В. Луценко. – Краснодар : ВЦСКИ «Эйдос», 2020. – 188 с. DOI 10.13140/RG.2.2.36089.01126, <https://www.researchgate.net/publication/340264083>.
3. Луценко Е. В. Современные информационно-коммуникационные технологии в научно-исследовательской деятельности и образовании: учеб. пособие для асп. напр. подг.: 40.06.01 "Юриспруденция"/ Е. В. Луценко, Г. М. Меретуков, В. И. Лойко. – Краснодар: КубГАУ, 2019. – 256с. ISBN 978-5-907294-46-2, <https://www.researchgate.net/publication/339285519>.
4. Прокопьева А.А. Применение информационных технологий и математического моделирования в управлении банковскими рисками // Магист. дисс., Режим доступа: https://dspace.spbu.ru/bitstream/11701/12281/1/Master_s_dissertation-7.pdf (обращение: 23.05.2021).
5. Беляков А.В.. Резервы на возможные потери - экономическая сущность и учёт для целей налогообложения. Дата обращения: 27 февраля 2010. [Архивировано](#) 12 апреля 2012 года.
6. Курносенко А.А.. Особенности правового регулирования банковскими рисками в условиях рыночной экономики // Банковское право. — 2008. — Вып. 5.
7. Погорелова О.С.. Проблемы прогнозирования кредитных рисков // Банковское кредитование. — 2008. — Вып. 3.
8. Слуцкий А.А.. Концепция определения значения минимального резерва по ссудам // Банковское кредитование. — 2008. — Вып. 4.
9. Сухов А.В.. Управление кредитными рисками в России и Европе: сравнительный анализ // Управление в кредитной организации. — 2008. — Вып. 6.
10. Информационный бюллетень // Выпуск № 1(44), Москва 2013 год, Режим доступа: http://www.ksp.mos.ru/upload/documents/inf_bulletin/KSP-144.pdf, (обращение: 23.05.2021).
11. Луценко Е.В., Лебедев Е.А., Подсистема автоматического формирования двоичного дерева классов семантической информационной модели (Подсистема "Эйдос-Tree"). Пат. № 2008610096 РФ. Заяв. № 2007613721 РФ. Оpubл. от 09.01.2008. – Режим доступа: <http://lc.kubagro.ru/aidos/2008610096.jpg>, 3,125 / 2,500 у.п.л.
12. Луценко Е.В. Системно-когнитивный анализ временных рядов на примере фондового рынка (когнитивная структуризация и формализация предметной области) / Е.В. Луценко, Е.А. Лебедев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №07(051). С. 1 – 37. – Шифр Информрегистра: 0420900012\0073, IDA [article ID]: 0510907001. – Режим доступа: <http://ej.kubagro.ru/2009/07/pdf/01.pdf>, 2,312 у.п.л.
13. Луценко Е.В. Системно-когнитивный анализ временных рядов на примере фондового рынка (синтез и верификация семантической информационной модели) / Е.В. Луценко, Е.А. Лебедев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №07(051). С. 38 – 46. – Шифр Информрегистра: 0420900012\0072, IDA [article ID]: 0510907002. – Режим доступа: <http://ej.kubagro.ru/2009/07/pdf/02.pdf>, 0,562 у.п.л.
14. Луценко Е.В. Системно-когнитивный анализ временных рядов на примере фондового рынка (прогнозирование, принятие решений и исследование предметной

области) / Е.В. Луценко, Е.А. Лебедев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №07(051). С. 47 – 82. – Шифр Информрегистра: 0420900012\0071, IDA [article ID]: 0510907003. – Режим доступа: <http://ej.kubagro.ru/2009/07/pdf/03.pdf>, 2,25 у.п.л.

15. Луценко Е.В. Повышение качества моделей «knowledge management» путем разделения классов на типичную и нетипичную части / Е.В. Луценко, Е.А. Лебедев, В.Н. Лаптев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №10(054). С. 78 – 93. – Шифр Информрегистра: 0420900012\0109, IDA [article ID]: 0540910005. – Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/05.pdf>, 1 у.п.л.

16. Луценко, Е. В. Определение кредитоспособности физических лиц и риска их кредитования / Е. В. Луценко, Е. А. Лебедев // Финансы и кредит. – 2006. – № 32(236). – С. 75-83. <https://elibrary.ru/item.asp?id=9290406>

17. Луценко Е.В. Астросоциотипология и спектральный анализ личности по астросоциотипам с применением семантических информационных мультимodelей / Е.В. Луценко, А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №01(035). С. 101 – 151. – Шифр Информрегистра: 0420800012\0002, IDA [article ID]: 0350801010. – Режим доступа: <http://ej.kubagro.ru/2008/01/pdf/10.pdf>, 3,188 у.п.л.

18. Луценко Е.В. Повышение адекватности спектрального анализа личности по астросоциотипам путем их разделения на типичную и нетипичную части / Е.В. Луценко, А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №02(036). С. 153 – 174. – Шифр Информрегистра: 0420800012\0017, IDA [article ID]: 0360802010. – Режим доступа: <http://ej.kubagro.ru/2008/02/pdf/10.pdf>, 1,375 у.п.л.

19. 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. <http://elibrary.ru/item.asp?id=18632909>

20. Lutsenko E.V. SYSTEM ANALYSIS AND DECISION MAKING (Automated system-cognitive analysis and solving problems of identification, decision-making and research of the simulated subject area) // August 2020, DOI: [10.13140/RG.2.2.27247.05289](https://doi.org/10.13140/RG.2.2.27247.05289), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), <https://www.researchgate.net/publication/343998862>

21. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системнокогнитивном 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.

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

23. Луценко Е.В. Проблемы и перспективы теории и методологии научного познания и автоматизированный системно-когнитивный анализ как

автоматизированный метод научного познания, обеспечивающий содержательное феноменологическое моделирование / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №03(127). С. 1 – 60. – IDA [article ID]: 1271703001. – Режим доступа: <http://ej.kubagro.ru/2017/03/pdf/01.pdf>, 3,75 у.п.л.

24. Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №02(126). С. 1 – 32. – IDA [article ID]: 1261702001. – Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf> 2 у.п.л.

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

26. Луценко Е.В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе «Эйдос») / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №07(071). С. 528 – 576. – Шифр Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf> 3,062 у.п.л.

27. 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.

28. 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-X++» / 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.

29. Луценко Е.В. Открытая масштабируемая интерактивная интеллектуальная online среда для обучения и научных исследований на базе АСК-анализа и системы «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – №06(130). С. 1 – 55. – IDA [article ID]: 1301706001. – Режим доступа: <http://ej.kubagro.ru/2017/06/pdf/01.pdf>, 3,438 у.п.л. http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

30. Луценко Е.В., Открытая масштабируемая интерактивная интеллектуальная online среда «Эйдос» («Эйдос-online»). Свид. РосПатента РФ на программу для ЭВМ,

Заявка № 2017618053 от 07.08.2017, Гос.рег.№ 2017661153, зарегистр. 04.10.2017. – Режим доступа: <http://lc.kubagro.ru/aidos/2017661153.jpg> 2 у.п.л.

31. Lutsenko E.V. Theoretical foundations, mathematical model and software tools for Automated system-cognitive analysis // July 2020, DOI: [10.13140/RG.2.2.21918.15685](https://doi.org/10.13140/RG.2.2.21918.15685), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), <https://www.researchgate.net/publication/343057312>

32. Сайт проф.Е.В.Луценко: <http://lc.kubagro.ru/>

33. Блог Е.В.Луценко в RG <https://www.researchgate.net/profile/Eugene-Lutsenko>

34. Канал проф.Е.В.Луценко на YouTube: https://youtube.com/channel/UC_QF84d8SCaWxsnXnexNFzg

35. Lutsenko E.V. Forecasting the risks of loan non-repayment using an intelligent iterative algorithm for accounting for atypical cases // May 2021, DOI: [10.13140/RG.2.2.32991.38560](https://doi.org/10.13140/RG.2.2.32991.38560), <https://www.researchgate.net/publication/351924470>.

Literatura

1. Lutsenko E.V. Methods of writing scientific papers, logic and the manner in which scientific statements // February 2021, DOI: [10.13140/RG.2.2.23546.41920](https://doi.org/10.13140/RG.2.2.23546.41920), License [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/), <https://www.researchgate.net/publication/349039044>

2. Lucenko E. V. Nauchnaya publicistika : uchebnik / E. V. Lucenko. – Krasnodar : VCzSKI «E`jdos», 2020. – 188 s. DOI [10.13140/RG.2.2.36089.01126](https://doi.org/10.13140/RG.2.2.36089.01126), <https://www.researchgate.net/publication/340264083>.

3. Lucenko E. V. Sovremennyy`e informacionno-kommunikacionny`e texnologii v nauchno-issledovatel`skoj deyatel`nosti i obrazovanii: ucheb. posobie dlya asp. napr. podg.: 40.06.01 "Yurisprudenciya"/ E. V. Lucenko, G. M. Meretukov, V. I. Lojko. – Krasnodar: KubGAU, 2019. – 256s. ISBN 978-5-907294-46-2, <https://www.researchgate.net/publication/339285519>.

4. Prokop`eva A.A. Primenenie informacionny`x texnologij i matematicheskogo modelirovaniya v upravlenii bankovskimi riskami // Magist. diss., Rezhim dostupa: https://dspace.spbu.ru/bitstream/11701/12281/1/Master_s_dissertation-7.pdf (obrashhenie: 23.05.2021).

5. Belyakov A.V.. Rezervy` na vozmozhny`e poteri - e`konomicheskaya sushhnost` i uchyot dlya celej nalogooblozheniya. Data obrashheniya: 27 fevralya 2010. Arxivirovano 12 aprelya 2012 goda.

6. Kurnosenko A.A.. Osobennosti pravovogo regulirovaniya bankovskimi riskami v usloviyax ry`nochnoj e`konomiki // Bankovskoe pravo. — 2008. — Vy`p. 5.

7. Pogorelova O.S.. Problemy` prognozirovaniya kreditny`x riskov // Bankovskoe kreditovanie. — 2008. — Vy`p. 3.

8. Sluczkiy A.A.. Konceptiya opredeleniya znacheniya minimal`nogo rezerva po ssudam // Bankovskoe kreditovanie. — 2008. — Vy`p. 4.

9. Suxov A.V.. Upravlenie kreditny`mi riskami v Rossii i Evrope: sravnitel`ny`j analiz // Upravlenie v kreditnoj organizacii. — 2008. — Vy`p. 6.

10. Informacionny`j byulleten` // Vy`pusk № 1(44), Moskva 2013 god, Rezhim dostupa: http://www.ksp.mos.ru/upload/documents/inf_bulletin/KSP-144.pdf, (obrashhenie: 23.05.2021).

11. Lucenko E.V., Lebedev E.A., Podсистема avtomaticheskogo formirovaniya dvoichnogo dereva klassov semanticheskoy informacionnoj modeli (Podсистема "E`jdos-Tree"). Pat. № 2008610096 RF. Zayav. № 2007613721 RF. Opubl. ot 09.01.2008. – Rezhim dostupa: <http://lc.kubagro.ru/aidos/2008610096.jpg>, 3,125 / 2,500 у.п.л.

12. Lucenko E.V. Sistemno-kognitivny`j analiz vremenny`x ryadov na primere fondovogo ry`nka (kognitivnaya strukturizaciya i formalizaciya predmetnoj oblasti) / E.V. Lucenko, E.A. Lebedev // Politematicheskij setевой e`lektronny`j nauchny`j zhurnal

Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2009. – №07(051). S. 1 – 37. – Shifr Informregistra: 0420900012\0073, IDA [article ID]: 0510907001. – Rezhim dostupa: <http://ej.kubagro.ru/2009/07/pdf/01.pdf>, 2,312 u.p.l.

13. Lucenko E.V. Sistemno-kognitivny`j analiz vremenny`x ryadov na primere fondovogo ry`nka (sintez i verifikaciya semanticheskoy informacionnoj modeli) / E.V. Lucenko, E.A. Lebedev // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2009. – №07(051). S. 38 – 46. – Shifr Informregistra: 0420900012\0072, IDA [article ID]: 0510907002. – Rezhim dostupa: <http://ej.kubagro.ru/2009/07/pdf/02.pdf>, 0,562 u.p.l.

14. Lucenko E.V. Sistemno-kognitivny`j analiz vremenny`x ryadov na primere fondovogo ry`nka (prognozirovanie, prinyatie reshenij i issledovanie predmetnoj oblasti) / E.V. Lucenko, E.A. Lebedev // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2009. – №07(051). S. 47 – 82. – Shifr Informregistra: 0420900012\0071, IDA [article ID]: 0510907003. – Rezhim dostupa: <http://ej.kubagro.ru/2009/07/pdf/03.pdf>, 2,25 u.p.l.

15. Lucenko E.V. Povy`shenie kachestva modelej «knowledge management» putem razdeleniya klassov na tipichnuyu i netipichnuyu chasti / E.V. Lucenko, E.A. Lebedev, V.N. Laptev // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2009. – №10(054). S. 78 – 93. – Shifr Informregistra: 0420900012\0109, IDA [article ID]: 0540910005. – Rezhim dostupa: <http://ej.kubagro.ru/2009/10/pdf/05.pdf>, 1 u.p.l.

16. Lucenko, E. V. Opreделение kreditosposobnosti fizicheskix licz i riska ix kreditovaniya / E. V. Lucenko, E. A. Lebedev // Finansy` i kredit. – 2006. – № 32(236). – S. 75-83. <https://elibrary.ru/item.asp?id=9290406>

17. Lucenko E.V. Astrosociotipologiya i spektral`ny`j analiz lichnosti po astrosociotipam s primeneniem semanticheskix informacionny`x mul`timodelej / E.V. Lucenko, A.P. Trunev // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2008. – №01(035). S. 101 – 151. – Shifr Informregistra: 0420800012\0002, IDA [article ID]: 0350801010. – Rezhim dostupa: <http://ej.kubagro.ru/2008/01/pdf/10.pdf>, 3,188 u.p.l.

18. Lucenko E.V. Povy`shenie adekvatnosti spektral`nogo analiza lichnosti po astrosociotipam putem ix razdeleniya na tipichnuyu i netipichnuyu chasti / E.V. Lucenko, A.P. Trunev // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2008. – №02(036). S. 153 – 174. – Shifr Informregistra: 0420800012\0017, IDA [article ID]: 0360802010. – Rezhim dostupa: <http://ej.kubagro.ru/2008/02/pdf/10.pdf>, 1,375 u.p.l.

19. Lucenko E.V. Avtomatizirovanny`j sistemno-kognitivny`j analiz v upravlenii aktivny`mi ob`ektami (sistemnaya teoriya informacii i ee primenenie v issledovanii e`konomicheskix, social`no-psixologicheskix, texnologicheskix i organizacionno-texnicheskix sistem): Monografiya (nauchnoe izdanie). – Krasnodar: KubGAU. 2002. – 605 s. <http://elibrary.ru/item.asp?id=18632909>

20. Lutsenko E.V. SYSTEM ANALYSIS AND DECISION MAKING (Automated system-cognitive analysis and solving problems of identification, decision-making and research of the simulated subject area) // August 2020, DOI:

10.13140/RG.2.2.27247.05289, License CC BY-SA 4.0, <https://www.researchgate.net/publication/343998862>

21. Lucenko E.V. Metrizaciya izmeritel'ny`x shkal razlichny`x tipov i sovmešt'naya sopostavimaya kolichestvennaya obrabotka raznorodny`x faktorov v sistemnokognitivnom analize i sisteme «E`jdos» / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j 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.

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

23. Lucenko E.V. Problemy` i perspektivy` teorii i metodologii nauchnogo poznaniya i avtomatizirovanny`j sistemno-kognitivny`j analiz kak avtomatizirovanny`j metod nauchnogo poznaniya, obespechivayushhij soderzhatel`noe fenomenologicheskoe modelirovanie / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2017. – №03(127). S. 1 – 60. – IDA [article ID]: 1271703001. – Rezhim dostupa: <http://ej.kubagro.ru/2017/03/pdf/01.pdf>, 3,75 u.p.l.

24. Lucenko E.V. Invariantnoe otnositel'no ob`emov danny`x nechetkoe mul'tiklassovoe obobshhenie F-mery` dostovernosti modelej Van Rizbergena v ASKANalize i sisteme «E`jdos» / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2017. – №02(126). S. 1 – 32. – IDA [article ID]: 1261702001. – Rezhim dostupa: <http://ej.kubagro.ru/2017/02/pdf/01.pdf> 2 u.p.l.

25. Lucenko E.V. Kolichestvenny`j avtomatizirovanny`j SWOT- i PEST-analiz sredstvami ASK-analiza i intellektual`noj sistemy` «E`jdos-X++» / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2014. – №07(101). S. 1367 – 1409. – IDA [article ID]: 1011407090. – Rezhim dostupa: <http://ej.kubagro.ru/2014/07/pdf/90.pdf> 2,688 u.p.l.

26. Lucenko E.V. Metod kognitivnoj klasterizacii ili klasterizaciya na osnove znaniy (klasterizaciya v sistemno-kognitivnom analize i intellektual`noj sisteme «E`jdos») / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2011. – №07(071). S. 528 – 576. – Shifr Informregistra: 0421100012\0253, IDA [article ID]: 0711107040. – Rezhim dostupa: <http://ej.kubagro.ru/2011/07/pdf/40.pdf> 3,062 u.p.l.

27. Lucenko E.V. Sistemnaya teoriya informacii i nelokal`ny`e interpretiruemy`e nejronny`e seti pryamogo scheta / E.V. Lucenko // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j 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.

28. Lucenko E.V. Modelirovanie slozhny`x mnogofaktorny`x nelinejny`x ob`ektov upravleniya na osnove fragmentirovanny`x zashumlenny`x e`mpiricheskix danny`x bol'shoj razmernosti v sistemno-kognitivnom analize i intellektual`noj sisteme «E`jdos-X++» / E.V. Lucenko, V.E. Korzhakov // Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j 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.

29. Lucenko E.V. Otkry`taya masshtabiruemaya interaktivnaya intellektual`naya online sreda dlya obucheniya i nauchny`x issledovaniy na baze ASK-analiza i sistemy` «E`jdos» / E.V. Lucenko // Politematicheskij setevoy e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchny`j zhurnal KubGAU) [E`lektronny`j resurs]. – Krasnodar: KubGAU, 2017. – №06(130). S. 1 – 55. – IDA [article ID]: 1301706001. – Rezhim dostupa: <http://ej.kubagro.ru/2017/06/pdf/01.pdf>, 3,438 u.p.l. http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf

30. Lucenko E.V., Otkry`taya masshtabiruemaya interaktivnaya intellektual`naya online sreda «E`jdos» («E`jdos-online»). Svid. RosPatenta RF na programmu dlya E`VM, Zayavka № 2017618053 ot 07.08.2017, Gos.reg.№ 2017661153, zaregistr. 04.10.2017. – Rezhim dostupa: <http://lc.kubagro.ru/aidos/2017661153.jpg> 2 u.p.l.

31. Lutsenko E.V. Theoretical foundations, mathematical model and software tools for Automated system-cognitive analysis // July 2020, DOI: 10.13140/RG.2.2.21918.15685, License CC BY-SA 4.0, <https://www.researchgate.net/publication/343057312>

32. Sajt prof.E.V.Lucenko: <http://lc.kubagro.ru/>

33. Blog E.V.Lucenko v RG <https://www.researchgate.net/profile/Eugene-Lutsenko>

34. Kanal prof.E.V.Lucenko na YouTube: https://youtube.com/channel/UC_QF84d8SCaWxsnXnexNFzg

35. Lutsenko E.V. Forecasting the risks of loan non-repayment using an intelligent iterative algorithm for accounting for atypical cases // May 2021, DOI: 10.13140/RG.2.2.32991.38560, <https://www.researchgate.net/publication/351924470>.

Приложение 1. Программная реализация в системе «Эйдос» интеллектуального итерационного алгоритма учета нетипичных случаев

```

*****
***** Из файла исходных данных "Inp_data.dbf" стандарта программного интерфейса 2.3.2.2 удаляются объекты обучающей выборки
***** с отрицательным уровнем сходства с классом, к которому они относятся. В данном режиме используются результаты распознавания.
*****
FUNCTION F3_7_6()

LOCAL GetList := {}

*Razrab()
*Running(.F.)
*RETURN NIL

Running(.T.)

IF M_KodAdmAppls = 0 // Выйти из системы если нет авторизации
  LB_Warning(L("Вы не авторизовались в системе (режим 1.1) и не можете ей пользоваться!"))
  Running(.F.)
  RETURN NIL
ENDIF

*IF ApplChange("3.7.6()") // Если не запущен режим, работающий с БД, то перейти в папку выбранного приложения
* Running(.F.)
* RETURN NIL
*ENDIF

IF ApplChange("") // Если не запущен режим, работающий с БД, то перейти в папку выбранного приложения
  Running(.F.)
  RETURN NIL
ENDIF

***** Проверки на наличие необходимых баз данных и сообщения, если их нет

IF .NOT. FILE(Disk_dir + "\AID_DATA\Inp_data\Inp_data.dbf")
  LB_Warning(L('В папке: нет базы данных: "Inp_data.dbf"!'), L('3.7.6. Разделение классов на типичную и нетипичную части'))
  Running(.F.)
  RETURN NIL
ENDIF

IF .NOT. FILE(Disk_dir + "\_2_3_2_2.arx")
  LB_Warning(L("Необходимо создать модель в режиме 2.3.2.2."), L('3.7.6. Разделение классов на типичную и нетипичную части' ))
ELSE
  aSoftInt = DC_ARestore(M_PathAppl + "\_2_3_2_2.arx")
  aSoftInt[ 2] = 1 // Нули и пробелы считать отсутствием данных
  aSoftInt[15] = 1 // 1-равные интервалы, 2-адаптивные интервалы (с примерно равным количеством наблюдений, по Котельникову)

```

```

// Если удалять классы, то можно использовать равные интервалы, а если добавлять, то можно и
адаптивные
aSoftInt[27] = 3 // Использовать Inp_data.dbf
DC_ASave(aSoftInt , M_PathAppl+"\_2_3_2_2.arx")
DC_ASave(aSoftInt , Disk_dir +"\_2_3_2_2.arx")
ENDIF

CloseAll() // Закрытие всех баз данных с ожиданием завершения операций
USE Inp_data EXCLUSIVE NEW

mFlagTXT = .T.
FOR mClSc = aSoftInt[3] TO aSoftInt[4] // Цикл по классификационным шкалам
    IF FIELDTYPE(mClSc)="C" // Символьные столбцы
        mFlagTXT = .F.
        EXIT
    ENDIF
NEXT
IF mFlagTXT // Нет текстовых классификационных шкал
    aMess := {}
    AADD(aMess, L('В файле исходных данных "Inp_data" нет текстовых классификационных шкал, '))
    AADD(aMess, L('а данный режим работает только с текстовыми классификационными шкалами. '))
    LB Warning(aMess, L('3.7.6. Разделение классов на типичную и нетипичную части' ))
    Running(.F.)
    RETURN NIL
ENDIF

*MsgBox(M_PathAppl+"Rsp2i.dbf")
IF .NOT. FILE(M_PathAppl+"Rsp2i.dbf") .OR.;
    .NOT. FILE(M_PathAppl+"Rsp2k.dbf")
    aMess := {}
    AADD(aMess, L('В папке: ')+ALLTRIM(M_PathAppl)+L(' нет базы данных: "Rsp2i.dbf"!'))
    AADD(aMess, L('Необходимо выполнить режим 3.5, чтобы сформировать ее.'))
    LB Warning(aMess, L('3.7.6. Разделение классов на типичную и нетипичную части' ))
    Running(.F.)
    RETURN NIL
ENDIF

*****

mIntKrit = 1
IF .NOT. FILE('_DelObj.txt')
    StrFile(ALLTRIM(STR(1)), '_mIntKrit.txt') // Запись текстового файла _mIntKrit.txt
ENDIF
mIntKrit = VAL(FileStr('_mIntKrit.txt')) // Загрузка текстового файла _mIntKrit.txt

mDelObj = 1
IF .NOT. FILE('_DelObj.txt')
    StrFile(ALLTRIM(STR(1)), '_DelObj.txt') // Запись текстового файла _DelObj.txt
ENDIF

```

```

mDelObj = VAL(FileStr('_DelObj.txt')) // Загрузка текстового файла _DelObj.txt

mIntKrit = 2
@ 1, 0 DCGROUP oGroup1 CAPTION L('Какой интегральный критерий использовать? ') SIZE 80.0, 3.5
@ 1, 2 DCRADIO mIntKrit VALUE 1 PROMPT L('1. Резонанс знаний ') PARENT oGroup1
@ 2, 2 DCRADIO mIntKrit VALUE 2 PROMPT L('2. Сумма знаний ') PARENT oGroup1

@1.2,50 DCPUSHBUTTON CAPTION L("Пояснение") SIZE 15, 1.5 ACTION {||Help376(), DC_GetRefresh(GetList)} PARENT oGroup1

@ 5, 0 DCGROUP oGroup2 CAPTION L('Как обрабатывать нетипичные объекты обучающей выборки:') SIZE 80.0, 3.5
@ 1, 2 DCRADIO mDelObj VALUE 1 PROMPT L('Удалять нетипичные объекты из "Inp_data.dbf" ') PARENT oGroup2
@ 2, 2 DCRADIO mDelObj VALUE 2 PROMPT L('Создавать новые классы для нетипичных объектов ') PARENT oGroup2

DCREAD GUI;
TO lExit ;
FIT;
ADDBUTTONS;
MODAL;
TITLE L("4.2.2.4 (3.7.6) Разделение классов на типичную и нетипичную части")

*****
IF lExit
** Button Ok
ELSE
*****
**** БД, открытые перед запуском главного меню
**** Восстанавливать их после выхода из функций главного меню
*****
CloseAll() // Закрытие всех баз данных с ожиданием завершения операций
DIRCHANGE(Disk_dir) // Перейти в папку с исполнимым модулем системы
USE PathGrAp EXCLUSIVE NEW
USE ApplS EXCLUSIVE NEW
USE Users EXCLUSIVE NEW
*****
Running(.F.)
RETURN NIL
ENDIF
*****

ERASE('_IntKrit.txt');StrFile(ALLTRIM(STR(mIntKrit)), '_IntKrit.txt') // Запись текстового файла _ObjErr.txt
ERASE('_DelObj.txt');StrFile(ALLTRIM(STR(mDelObj)), '_DelObj.txt') // Запись текстового файла _DelObj.txt
*****

aSoftInt = DC_ARestore(M_PathAppl+"\_2_3_2_2.arx")

// Если в объектах обучающей выборки, приведших к FN-решениям, удалять классы, то допустимы только равные интервалы,
// а если для таких решений классы добавлять, то можно использовать и адаптивные интервалы
IF mDelObj = 1 // Удалять нетипичные объекты из "Inp_data.dbf"
aSoftInt[ 2] = 1 // Нули и пробелы считать отсутствием данных

```

```

        aSoftInt[15] = 1                                // 1-равные интервалы, 2-адаптивные интервалы (с примерно равным количеством наблюдений, по
Котельникову) <<<==#####
        ENDIF
        aSoftInt[27] = 3                                // Использовать Inp_data.dbf
        DC_ASave(aSoftInt , M_PathAppl+"\_2_3_2_2.arx")
        DC_ASave(aSoftInt , Disk_dir +"\_2_3_2_2.arx")

CloseAll()      // Закрытие всех баз данных с ожиданием завершения операций
COPY FILE ("Inp_data.dbf") TO ("ObjFalseNeg.dbf")
COPY FILE ("Inp_data.dbf") TO ("InpDataSource.dbf")

CloseAll()      // Закрытие всех баз данных с ожиданием завершения операций
USE Inp_data    EXCLUSIVE NEW
USE ObjFalseNeg EXCLUSIVE NEW;ZAP

DO CASE
    CASE mIntKrit=1
        USE Rsp2k EXCLUSIVE NEW
    CASE mIntKrit=2
        USE Rsp2i EXCLUSIVE NEW
ENDCASE

nMax = RECCOUNT()
Mess = L('4.2.2.4 (3.7.6). Разделение классов на типичную и нетипичную части')
@ 4,5 DCPROGRESS oProgress SIZE 70,1.1 MAXCOUNT nMax COLOR GRA_CLR_CYAN PERCENT EVERY 100
DCREAD GUI TITLE Mess PARENT @oDialog FIT EXIT
oDialog:show()
nTime = 0
DC_GetProgress(oProgress,0,nMax)

mNObjFN = 0

DBGOTOP()
DO WHILE .NOT. EOF()

    mKodObj = KOD_OBJ
    mKodClSc = KOD_CLSC

    mNameClsOld = ALLTRIM(NAME_CLS)
    mPos = RAT('-', mNameClsOld)
    mNameCls = SUBSTR(mNameClsOld, mPos+1, LEN(mNameClsOld)-mPos)

    IF LEN(ALLTRIM(Fakt)) > 0 .AND. IF(mIntKrit=1, KORR, SUM_INF) < 0

*       MsgBox('Kod_obj='+ALLTRIM(STR(KOD_OBJ))+ ' '+ALLTRIM(NAME_OBJ)+' '+ALLTRIM(STR(KOD_CLS))+ ' '+ALLTRIM(NAME_CLS)+' '+ALLTRIM(STR(KOD_CLSC))+
'+ALLTRIM(STR(KORR))+ ' '+ALLTRIM(STR(SUM_INF))+ ' '+ALLTRIM(ФАКТ))

        SELECT Inp_data
        DBGOTO(mKodObj)

```

```

*      MsgBox('Kod_obj='+ALLTRIM(STR(FIELDGET(1)))+ ' '+ALLTRIM(FIELDGET(2))+ ' '+ALLTRIM(STR(FIELDGET(3)))+ ' '+ALLTRIM(STR(FIELDGET(4)))+
'+ALLTRIM(FIELDGET(5)))

      mNumClSc = aSoftInt[3]+mKodClSc-1          // Номер колонки классификационной шкалы в БД Inp_data
      IF FIELDTYPE(mNumClSc) = "C"             // Текстовые классификационные шкалы
          mNObjFN++
          IF mDelObj = 1                       // Удалять объекты обучающей выборки
              FIELDPUT(mNumClSc, '' )         // Класс - отсутствие данных
          ELSE
              FIELDPUT(mNumClSc, mNameCls+'_FN' ) // Класс - отсутствие данных
          ENDIF
          aObj := {}
          FOR j=1 TO FCOUNT()
              AADD(aObj, FIELDGET(j))
          NEXT
          SELECT ObjFalseNeg                   // БД с объектами обучающей выборки, приведшими к ложно-отрицательным решениям
          APPEND BLANK
          FOR j=1 TO LEN(aObj)
              FIELDPUT(j, aObj[j])
          NEXT
      ENDIF

      ENDIF

      DO CASE
          CASE mIntKrit=1
              SELECT Rsp2k
          CASE mIntKrit=2
              SELECT Rsp2i
          ENDCASE

      DC_GetProgress(oProgress, ++nTime, nMax)
      DBSKIP(1)

      ENDDO

      *MsgBox('STOP')
      DC_GetProgress(oProgress, nMax, nMax)
      oDialog:Destroy()

      **** После корректировки файла Inp_data.dbf в папке приложения записать его в ..\AID_DATA\Inp_data\

      *MsgBox(M_ApplsPath+"\Inp_data\Inp_data.dbf")

      CloseAll() // Закрытие всех баз данных с ожиданием завершения операций
      COPY FILE ("Inp_data.dbf") TO (M_ApplsPath+"\Inp_data\Inp_data.dbf")
      COPY FILE ("ObjFalseNeg.dbf") TO (M_ApplsPath+"\Inp_data\ObjFalseNeg.dbf")
      COPY FILE ("ObjFalseNeg.dbf") TO (M_ApplsPath+"\Inp_data\ObjFalseNeg.xls")

      aMess := {}
    
```

```

IF mDelObj = 1
  AADD(aMess, L('Удалено:')+' '+ALLTRIM(STR(mNObjFN))+' '+L('нетипичных объектов обучающей выборки.'))
ENDIF
IF mDelObj = 2
  AADD(aMess, L('Назначено на новые классы:')+' '+ALLTRIM(STR(mNObjFN))+' '+L('нетипичных объектов обучающей выборки.'))
ENDIF
AADD(aMess, L('БД с объектами обучающей выборки, приведшими к FN-решениям:')+' '+M_ApplsPath+"\Inp_data\ObjFalseNeg.xls")
AADD(aMess, L(' '))
AADD(aMess, L('Далее нужно выполнить режим 2.3.2.2 с параметрами, заданными по умолчанию.'))
LB_Warning(aMess, L('4.2.2.4 (3.7.6) Разделение классов на типичную и нетипичную части'))

*****
***** ЗАПИСАТЬ ПАРАМЕТРЫ ДЛЯ 2.3.2.2, ЧТОБЫ ЗАГРУЗКА ШЛА ИЗ INP_DATA.DBF
*****

IF FILE("_2_3_2_2.arx")

  aSoftInt = DC_ARestore(Disk_dir +"\_2_3_2_2.arx")
  aSoftInt[ 2] = 1
  aSoftInt[27] = 3
  DC_ASave(aSoftInt , Disk_dir +"\_2_3_2_2.arx")

ENDIF

*****
***** Запустить 2.3.2.2, 3.5
*****

F2_3_2_2("", "") // Запуск универсального программного интерфейса с внешними базами данных

* Возникает ошибка в отображении кода исполнения. Так и не смог разобраться
* F3_5('GPU', '', '', '') // Какая модель?
* F3_5('GPU', 'SintRec', '3.7.6', 'ALL') // Какая модель?
* F3_5('GPU', 'SintRec', '3.5', 'ALL')

*****
***** БД, открытые перед запуском главного меню
***** Восстанавливать их после выхода из функций главного меню
*****
CLOseAll() // Закрытие всех баз данных с ожиданием завершения операций
DIRCHANGE(Disk_dir) // Перейти в папку с исполнимым модулем системы
USE PathGrAp EXCLUSIVE NEW
USE Appls EXCLUSIVE NEW
USE Users EXCLUSIVE NEW
*****

AADD(aMess, L(' '))
AADD(aMess, L('Далее необходимо в режиме 3.5 создать и верифицировать модели: Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7. '))
AADD(aMess, L('Затем в режиме 3.4 необходимо определить модель и интегральный критерий, при которых достигается максимальная достоверность.'))
AADD(aMess, L(' '))

```

```
AADD(aMess, L('Режим 4.2.2.4 (3.7.6) можно повторять до достижения необходимого достаточно высокого уровня достоверности моделей. '))
AADD(aMess, L(' '))
AADD(aMess, L('Если достоверность модели достаточно высока, то в ней корректно можно решать задачи идентификации и прогнозирования (4.1.2),'))
AADD(aMess, L(' принятия решений (4.4.8) и исследования объекта моделирования путем исследования его модели (режимы: 4.4.9, 4.4.10, 4.4.11,'))
AADD(aMess, L(' 4.4.12, 4.2.1, 4.2.2.1, 4.2.2.2, 4.2.2.3, 4.2.3, 4.3.2.1, 4.3.2.2, 4.3.2.3, 4.5, 3.7.5, 3.7.4., 3.7.3, 3.7.9 и т.д.) '))
LB_Warning(aMess, L('4.2.2.4 (3.7.6) Разделение классов на типичную и нетипичную части'))
```

```
Running(.F.)
RETURN NIL
```
