

УДК 303.732.4

**ПРОГНОЗИРОВАНИЕ РИСКОВ ОСАГО (АНДЕРАЙТИНГ)
С ПРИМЕНЕНИЕМ СИСТЕМНО-КОГНИТИВНОГО АНАЛИЗА
(по данным ООО Росгосстрах-ЮГ Краснодарского края)**

Луценко Е.В., – д.э.н., к.т.н., профессор

Подставкин Н.А., – студент-дипломник

Кубанский государственный аграрный университет

В статье описываются результаты применения системно-когнитивного анализа для прогнозирования рисков совершения дорожно-транспортных происшествий и сумм страховых выплат в системе обязательного автострахования

Ключевые слова: АНДЕРАЙТИНГ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ
ПРОГНОЗИРОВАНИЕ РИСКОВ ДТП СУММ СТРАХОВЫХ ВЫПЛАТ
ОБЯЗАТЕЛЬНОЕ АВТОСТРАХОВАНИЕ

В соответствии с действующим законодательством страховая компания, оказывающая услуги обязательного автострахования, обязана оказывать эти услуги всем обращающимся за этим клиентам и не может отказать кому-либо из них, при условии, что клиент предоставил все необходимые документы. Вместе с тем компания имеет ряд законных способов повлиять на контингент обращающихся к ней клиентов. Среди этих способов необходимо отметить прежде всего целенаправленную рекламу, ориентированную на определенные группы населения, а также маркетинговые мероприятия. В результате использования подобных технологий страховая компания уже сейчас может повлиять на приоритеты и механизм принятия решений различных категорий потенциальных клиентов в таком направлении, которое приведет к относительному увеличению доли желательных клиентов и уменьшения доли нежелательных. Кроме того возможно в будущем страховые компании получат большие права по отбору клиентов. Уже сегодня компании имеют право уменьшать оплату за автострахования безаварийным водителям и повышать для водителей, совершивших до-

рожно-транспортные происшествия (ДТП), повлекшие страховые выплаты со стороны компании.

В связи со всем этим, когда в страховую компанию обращается очередной клиент то возникает резонные вопросы о том:

1. Какова вероятность совершения ДТП этим водителем или владельцем автотранспортного средства (конкретно именно этим водителем и конкретно именно на данном автотранспортном средстве).

2. Какова возможная тяжесть этого ДТП и наиболее вероятная сумма соответствующей страховой выплаты.

В теории и практике страхования эта задача получила название "*андеррайтинг*". Обладание технологией решения этой задачи уже сегодня позволило бы компании существенно уменьшить риски автострахования и соответственно, уменьшить страховые выплаты и увеличить свою прибыль. По-видимому, в перспективе возможность решения этой задачи в реальном времени непосредственно во время приема клиента на начальной стадии оформления документов будет играть еще большую роль.

Однако, *проблема* состоит в том, что не смотря на все эти достаточно очевидные выгоды и преимущества в реальной практике большинства страховых компаний системы андеррайтинга не применяются.

На наш взгляд это обусловлено следующими обстоятельствами:

1. Эти системы недостаточно технологичны для их применения в реальном времени, непосредственно во время обслуживания клиента.

2. Существующие системы разработаны за рубежом или в мегаполисах (в основном в Москве и Санкт-Петербурге) и очень слабо отражают региональную специфику (т.е. *нелокализованы*), вернее вообще ее практически не отражают, из-за чего и имеют очень низкую достоверность прогнозирования, близкую и статистически незначимо отличающуюся от вероятности случайного угадывания без использования этих систем или другой априорной информации.

3. Эти системы не обладают *адаптивностью* и не учитывают динамику предметной области, которая чрезвычайно высока, особенно в Южном Федеральном Округе (ЮФО). В результате даже первоначально хоро-

шо работающие системы очень быстро теряют адекватность модели и качество прогнозов.

4. Стоимость этих систем настолько высока, что их приобретение и использование чаще всего мало или вообще нерентабельно.

Целью данной работы является решение поставленной проблемы путем разработки адаптивной методики прогнозирования рисков автострахования и сумм страховых выплат, и, на этой основе, поддержки принятия решений в страховой компании.

Для достижения поставленной цели был выбран метод системно-когнитивного анализа (СК-анализ). Этот выбор был обусловлен тем, что данный метод является непараметрическим, позволяет сопоставимо обрабатывать тысячи градаций факторов и будущих состояний объекта управления при неполных (фрагментированных), зашумленных данных различной природы, т.е. измеряемых в различных единицах измерения. Для метода СК-анализа разработаны и методика численных расчетов, и соответствующий программный инструментарий, а также технология и методика их применения. Они прошли успешную апробацию при решении ряда задач в различных предметных областях [1-33]. Наличие инструментария СК-анализа (базовая система "Эйдос") позволяет не только осуществить синтез семантической информационной модели (СИМ), но и периодически проводить адаптацию и синтез ее новых версий, обеспечивая тем самым отслеживание динамики предметной области и сохраняя высокую адекватность модели в изменяющихся условиях. Важной особенностью СК-анализа является возможность единообразной числовой обработки разнотипных по смыслу и единицам измерения числовых и нечисловых данных. Это обеспечивается тем, что нечисловым величинам тем же методом, что и числовым, приписываются сопоставимые в пространстве и времени, а также между собой, количественные значения, позволяющие обрабатывать их как числовые: на первых двух этапах СК-анализа числовые величины сводятся к интервальным оценкам, как и информация об объектах нечисловой природы (фактах, событиях) (этот этап реализуется и в методах интервальной статистики); на третьем этапе СК-анализа всем этим величинам по единой методике, основанной на системном обобщении семантической теории информации А.Харкевича, сопоставляются количественные вели-

<http://ej.kubagro.ru/2007/05/pdf/08.pdf>

чины (имеющие смысл количества информации в признаке о принадлежности объекта к классу), с которыми в дальнейшем и производятся все операции моделирования (этот этап является уникальным для СК-анализа).

В работах [5, 6] приведен перечень этапов системно-когнитивного анализа, которые необходимо выполнить, чтобы осуществить синтез модели и исследование модели объекта управления. Учитывая эти этапы СК-анализа выполним *декомпозицию* цели работы в последовательность **задач**, решение которых обеспечит ее поэтапное достижение:

1. Когнитивная структуризация предметной области и формальная постановка задачи, проектирование структуры и состава исходных данных.

2. Получение исходных данных запланированного состава в той форме, в которой они накапливаются в поставляющей их организации (обычно в форме базы данных).

3. Разработка стандартной Excel-формы для представления исходных данных.

4. Преобразование исходных данных из исходных баз данных в стандартную электронную Excel-форму.

5. Контроль достоверности исходных данных и исправление ошибок.

6. Разработка и использование *программного интерфейса* для преобразования исходных данных из стандартной Excel-формы в базы данных, используемые в инструментарии системно-когнитивного анализа (СК-анализ) – универсальной когнитивной аналитической системе "Эйдос" (система "Эйдос").

7. Синтез семантической информационной модели (СИМ).

8. Оптимизация СИМ.

9. Измерение адекватности СИМ.

10. **Задача 1:** *"Многокритериальная типизация"* автомобилей клиентов по типам: "Не совершившие ДТП", "Совершившие ДТП" и категориям, отражающим суммы страховых выплат.

11. **Задача 2:** "Разработка методики *прогнозирования* риска совершения ДТП и суммы страховой выплаты на основе информации о клиенте и его автомобиле".

12. **Задача 3:** "Разработка методики *поддержки принятия решений* по выбору контингента клиентов, наиболее предпочтительных и нежелательных для автострахования".

13. Разработка принципов оценки экономической эффективности разработанных технологий при их применении в реальной страховой компании.

14. Исследование ограничений разработанной технологии и перспектив ее развития.

Кратко рассмотрим решение этих задач.

1. Когнитивная структуризация предметной области это 1-й этап формальной постановки задачи, на котором решается, какие параметры будут рассматриваться в качестве причин, а какие – следствий. На этом этапе было решено в качестве следствий, т.е. классов, рассматривать: безаварийность, аварийность, сумму страховой выплаты, а в качестве причин: марку и модель автотранспортного средства, его цвет, Российское или иностранное производство автомобиля, водительский стаж владельца.

На этапе формальной постановки задачи, исходя из результатов когнитивной структуризации, было осуществлено проектирование структуры и состава исходных данных.

2. Затем исходные данные запланированного состава *были получены* в той форме, в которой они накапливаются в поставляющей их организации (обычно в форме базы данных). В нашем случае этой организацией выступила компания ООО Росгосстрах-ЮГ Краснодарского края. Здесь необходимо отметить, что в полученной базе данных представлено **65535** примеров (это максимальное количество строк в листе Excel) застрахованных автотранспортных средств, из которых **540** участвовали в различного рода ДТП и по этим случаям *были произведены* страховые выплаты. Этого было более чем достаточно для целей данной работы, за что авторы благодарны руководству данной страховой компании.

3. Была разработана стандартная Excel-форма для представления исходных данных (см. рис. 1).

4. Исходные данные из Excel-формы, представленной на рис.1, были преобразованы средствами Excel в стандартную для программного интерфейса системы "Эйдос" электронную Excel-форму (см. рис. 2).

№	Генератор	мощность	Мощность в кВт	мощность в МВт	мощность в МВт	Имя	Статус	Сумма выработки
1	С-174425	УТН 1376078723	27	28,35	ЕАЭ	1-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	18 856,50
2	С-174425	УТН 1376078723	17	21,33	ЕАЭ	1-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	34 563,00
3	С-174425	УТН 1376078723	06	7,03	ЕАЭ	3-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	13 301,72
4	С-174425	УТН 1376078723	0	0	ЕАЭ	3-06	ТМ-Берёв (ТТЭ-0179У-0-0910К)	2 064,35
5	С-174425	УТН 1376078723	70	67,20	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 443,00
6	С-174425	УТН 1376078723	75	65,12	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	54 341,00
7	С-174425	УТН 1376078723	25	22,65	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	17 328,65
8	С-174425	УТН 1376078723	35	33,15	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 654,00
9	С-174425	УТН 1376078723	35	33,15	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 615,40
10	С-174425	УТН 1376078723	75	65,12	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	11 307,00
11	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	8 654,00
12	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	31 854,00
13	С-174425	УТН 1376078723	106	102,21	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	4 313,00
14	С-174425	УТН 1376078723	102	73	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 545,00
15	С-174425	УТН 1376078723	72	62,36	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	17 058,14
16	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 547,00
17	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	13 360,90
18	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	8 764,00
19	С-174425	УТН 1376078723	07	6,99	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	9 523,95
20	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 633,00
21	С-174425	УТН 1376078723	17	14,88	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	24 364,40
22	С-174425	УТН 1376078723	21	18,90	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 663,00
23	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	49 131,99
24	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 493,00
25	С-174425	УТН 1376078723	21	18,90	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 921,00
26	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	877,00
27	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	120 000,00
28	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	6 219,03
29	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	87 543,00
30	С-174425	УТН 1376078723	06	4,37	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	10 478,18
31	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	66 553,00
32	С-174425	УТН 1376078723	66	60	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	21 917,50

Рисунок 1. Excel-форма с исходными данными

№	Генератор	мощность	Мощность в кВт	мощность в МВт	мощность в МВт	Имя	Статус	Сумма выработки	Color
1	С-174425	УТН 1376078723	27	28,35	ЕАЭ	1-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	18 856,50	ТМ-Берёв (ТТЭ-0179У-0-0910К)
2	С-174425	УТН 1376078723	17	21,33	ЕАЭ	1-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	34 563,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
3	С-174425	УТН 1376078723	06	7,03	ЕАЭ	3-	ТМ-Берёв (ТТЭ-0179У-0-0910К)	13 301,72	ТМ-Берёв (ТТЭ-0179У-0-0910К)
4	С-174425	УТН 1376078723	0	0	ЕАЭ	3-06	ТМ-Берёв (ТТЭ-0179У-0-0910К)	2 064,35	ТМ-Берёв (ТТЭ-0179У-0-0910К)
5	С-174425	УТН 1376078723	70	67,20	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 443,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
6	С-174425	УТН 1376078723	75	65,12	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	54 341,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
7	С-174425	УТН 1376078723	25	22,65	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	17 328,65	ТМ-Берёв (ТТЭ-0179У-0-0910К)
8	С-174425	УТН 1376078723	35	33,15	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 654,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
9	С-174425	УТН 1376078723	35	33,15	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 615,40	ТМ-Берёв (ТТЭ-0179У-0-0910К)
10	С-174425	УТН 1376078723	75	65,12	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	11 307,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
11	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	8 654,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
12	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	31 854,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
13	С-174425	УТН 1376078723	106	102,21	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	4 313,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
14	С-174425	УТН 1376078723	102	73	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 545,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
15	С-174425	УТН 1376078723	72	62,36	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	17 058,14	ТМ-Берёв (ТТЭ-0179У-0-0910К)
16	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 547,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
17	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	13 360,90	ТМ-Берёв (ТТЭ-0179У-0-0910К)
18	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	8 764,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
19	С-174425	УТН 1376078723	07	6,99	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	9 523,95	ТМ-Берёв (ТТЭ-0179У-0-0910К)
20	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 633,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
21	С-174425	УТН 1376078723	17	14,88	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	24 364,40	ТМ-Берёв (ТТЭ-0179У-0-0910К)
22	С-174425	УТН 1376078723	21	18,90	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	7 663,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
23	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	49 131,99	ТМ-Берёв (ТТЭ-0179У-0-0910К)
24	С-174425	УТН 1376078723	20	17,40	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 493,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
25	С-174425	УТН 1376078723	21	18,90	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	5 921,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
26	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	877,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
27	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	120 000,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
28	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	6 219,03	ТМ-Берёв (ТТЭ-0179У-0-0910К)
29	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	87 543,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
30	С-174425	УТН 1376078723	06	4,37	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	10 478,18	ТМ-Берёв (ТТЭ-0179У-0-0910К)
31	С-174425	УТН 1376078723	0	0	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	66 553,00	ТМ-Берёв (ТТЭ-0179У-0-0910К)
32	С-174425	УТН 1376078723	66	60	ЕАЭ	21079	ТМ-Берёв (ТТЭ-0179У-0-0910К)	21 917,50	ТМ-Берёв (ТТЭ-0179У-0-0910К)

Рисунок 2. Excel-форма, стандартная для программного интерфейса системы "Эйдос"

5. На этапе контроля достоверности исходных данных было обнаружено, что в исходной базе данных мощность двигателя у некоторых особо престижных моделей автомобилей доходила до 728690 л.с. В данной работе было принято решение не использовать в модели этот параметр, а не исправлять его. Хотя в принципе зная марку и модель автотранспортного средства можно было бы точно узнать и мощность его двигателя, и колесную формулу. Но для целей данной работы этого не требовалось и сделано не было.

6. Затем Excel-форма, приведенная на рисунке 2 средствами Excel была записана в стандарте DBF-4, dBASE IV (*.dbf), разработан и использован *программный интерфейс* для преобразования исходных данных из стандартной Excel-формы в базы данных, используемые в инструментарии системно-когнитивного анализа (СК-анализ) – универсальной когнитивной аналитической системе "Эйдос" (система "Эйдос").

Ниже приводится исходный текст программы данного программного интерфейса, работающий с получившимся dbf-файлом:

```
*****
***** ФОРМИРОВАНИЕ ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ ***
***** И ОБУЧАЮЩЕЙ ВЫБОРКИ ИЗ DBF-Excel-файла ***
***** Подставкин Н.А. 05/22/07 11:49am *****

PARAMETERS File_name

**** БЛОК-1. ОТОБРАЖЕНИЕ ИНФОРМАЦИИ О ФУНКЦИЯХ ПРОГРАММНОГО ИНТЕРФЕЙСА ****

scr23 = SAVESCREEN(0,0,24,79)

SET CURSOR OFF
SET DATE ITALIAN
SET DECIMALS TO 15
SET ESCAPE On

FOR J=0 TO 24
  @J,0 SAY REPLICATE("█",80) COLOR "gb+/N"
NEXT

SHOWTIME(0,60,.F., "rg+/n", .F.,.F.)

FOR J=0 TO 24
  @J,0 SAY REPLICATE("█",80) COLOR "gb+/N"
NEXT

**** ГЕНЕРАЦИЯ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ ***

COPY FILE ("Card.dbf") TO ("Card_sh.dbf")

CLOSE ALL
USE Object EXCLUSIVE NEW;ZAP
USE Priz_ob EXCLUSIVE NEW;ZAP
USE Priz_per EXCLUSIVE NEW;ZAP
USE Card EXCLUSIVE NEW
USE Card_sh EXCLUSIVE NEW;ZAP
APPEND BLANK
APPEND BLANK
APPEND BLANK

SELECT Card
N_Rec = RECCOUNT()

M_KodK1 = 0
```

M_KodSh = 0
M_KodGr = 0
N_Grad = 10

```
* Структура базы данных: card.dbf 01.05.07 08:56:20
* =====
* | N | Имя поля | Тип | Ширина | Дес. | Примечание |
* =====
* | 1 | SUMMA | N | 10 | 2 |
* | 2 | STAG | N | 3 | 0 |
* | 3 | POWER | N | 8 | 0 |
* | 4 | PROIZVODST | C | 9 | 0 |
* | 5 | MARKA | C | 11 | 0 |
* | 6 | MARKA_MODE | C | 31 | 0 |
* | 7 | COLOR | C | 51 | 0 |
* =====
* В С Е Г О длина записи: 124 байтов.
* =====
```

```
A_FNRus := {}
AADD(A_FNRus,"Сумма страховой выплаты: ")
AADD(A_FNRus,"Стаж владельца")
AADD(A_FNRus,"Мощность двигателя (л.с.)")
AADD(A_FNRus,"Производство")
AADD(A_FNRus,"Марка")
AADD(A_FNRus,"Марка и модель")
AADD(A_FNRus,"Цвет")
```

```
@24,0 SAY REPLICATE("■",80) COLOR "rb/n"
```

```
FOR ff=1 TO 7 && Начало цикла по полям Card.dbf
```

```
*** Числовые столбцы
```

```
IF 1 <= ff .AND. ff <= 3
```

```
SELECT Card
INDEX ON STR(999999999-FIELDGET(ff),10) TO Mrk_funi UNIQUE
DBGOTOP()
F_MaxSH = FIELDGET(ff)
DEGOTOBOTOM();DBSKIP(-1)
F_MinSH = FIELDGET(ff) && Не брать 0, т.к. для него отдельный класс
Delta = (F_MaxSH-F_MinSH)/N_Grad
```

```
SELECT Card_sh
DBGOTO(1);FIELDPUT(ff,F_MaxSH)
DBGOTO(2);FIELDPUT(ff,F_MinSH)
DBGOTO(3);FIELDPUT(ff,Delta)
```

```
** Классы
```

```
IF ff = 1
```

```
SELECT Object
APPEND BLANK
F_MinGR = 0
F_MaxGR = 0
M_Name = A_FNRus[ff]+: {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
REPLACE Kod WITH ++M_KodKl
REPLACE Name WITH M_Name
SELECT Object
APPEND BLANK
F_MinGR = F_MinSH
F_MaxGR = F_MaxSH
M_Name = A_FNRus[ff]+: {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
REPLACE Kod WITH ++M_KodKl
REPLACE Name WITH M_Name
SELECT Object
FOR gr = 1 TO N_Grad
SELECT Object
APPEND BLANK
F_MinGR = F_MinSH+(gr-1)*Delta
F_MaxGR = F_MinSH+(gr )*Delta
M_Name = A_FNRus[ff]+: {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
REPLACE Kod WITH ++M_KodKl
REPLACE Name WITH M_Name
NEXT
ENDIF
```

```
** Признаки числовые столбцы
```

```
IF ff = 2 && Только стаж, а мощность не берем, т.к. у нее бредовые значения
```

```
SELECT Priz_ob
APPEND BLANK
REPLACE Kod WITH ++M_KodSh
REPLACE Name WITH A_FNRus[ff]
FOR gr=1 TO N_Grad
SELECT Priz_per
APPEND BLANK
F_MinGR = F_MinSH+(gr-1)*Delta
F_MaxGR = F_MinSH+(gr )*Delta
M_Name = A_FNRus[ff]+: {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
REPLACE Kod WITH ++M_KodGr
REPLACE Kod_ob_pr WITH M_KodSh
REPLACE Name WITH M_Name
SELECT Priz_ob
FIELDPUT(gr+2,M_KodGr)
NEXT
ENDIF
ENDIF
```

```
** Признаки текстовые столбцы
```



```

IF 4 <= ff .AND. ff <= 7

SELECT Card
INDEX ON FIELDGET(ff) TO Mrk_funi UNIQUE

** Признаки

SELECT Priz_ob
APPEND BLANK
REPLACE Kod WITH ++M_KodSh
REPLACE Name WITH A_FNRus[ff]

SELECT Card
SET ORDER TO 1
DBGOTOP()
gr = 0
DO WHILE .NOT. EOF()

M_Name = A_FNRus[ff]+"-"+ALLTRIM(FIELDGET(ff))

SELECT Priz_per
APPEND BLANK
REPLACE Kod WITH ++M_KodGr
REPLACE Kod_ob_pr WITH M_KodSh
REPLACE Name WITH M_Name

SELECT Priz_ob
FIELDPUT(++gr+2,M_KodGr)

SELECT Card
DBSKIP(1)
ENDDO
ENDIF
NEXT

*WAIT

CLOSE ALL

*** ГЕНЕРАЦИЯ ОБУЧАЮЩЕЙ ВЫБОРКИ *****

CLOSE ALL
USE Object EXCLUSIVE NEW
INDEX ON Name TO Obj_name
CLOSE ALL
USE Priz_per EXCLUSIVE NEW
INDEX ON Name TO Prpe_nam

CLOSE ALL
USE Object INDEX Obj_name EXCLUSIVE NEW
USE Priz_per INDEX Prpe_nam EXCLUSIVE NEW
USE Card EXCLUSIVE NEW
USE Card_sh EXCLUSIVE NEW
USE ObInfZag EXCLUSIVE NEW;ZAP
USE ObInfKpr EXCLUSIVE NEW;ZAP

N_Rec = RECCOUNT()
DBGOTOP()

@24,0 SAY REPLICATE("■",80) COLOR "rb/n"

M_KodIst = 0

SELECT Card

N_Rec = RECCOUNT()
Num = 0

DBGOTOP()
DO WHILE .NOT. EOF()

SELECT Card_sh
DBGOTO(1);F_MaxSH = FIELDGET(1)
DBGOTO(2);F_MinSH = FIELDGET(1)

***** База заголовков
SELECT Card
M_NameIst = STR(FIELDGET(1),15,2)

**** Формирование массива кодов классов из БД Card

A_Obj := {}
M_Summa = FIELDGET(1)

F_MinGR = 0
F_MaxGR = 0
IF F_MinGR <= M_Summa .AND. M_Summa <= F_MaxGR
M_Name = A_FNRus[1]+"": {"+ALLTRIM(STR(F_MinGR,11,3))+"", "+ALLTRIM(STR(F_MaxGR,11,3))+"}
SELECT Object;SET ORDER TO 1;T=DBSEEK(M_Name)
IF T
IF ASCAN(A_Obj, Kod) = 0
AADD(A_Obj, Kod)
ENDIF
ENDIF
ENDIF

F_MinGR = F_MinSH
F_MaxGR = F_MaxSH
IF F_MinGR <= M_Summa .AND. M_Summa <= F_MaxGR
M_Name = A_FNRus[1]+"": {"+ALLTRIM(STR(F_MinGR,11,3))+"", "+ALLTRIM(STR(F_MaxGR,11,3))+"}
SELECT Object;SET ORDER TO 1;T=DBSEEK(M_Name)
IF T

```

```

        IF ASCAN(A_Obj, Kod) = 0
            AADD(A_Obj, Kod)
        ENDIF
    ENDIF
ENDIF

SELECT Card_sh
DBGOTO(1);F_MaxSH = FIELDGET(1)
DBGOTO(2);F_MinSH = FIELDGET(1)
Delta = (F_MaxSH-F_MinSH)/N_Grad
FOR gr=1 TO N_Grad
    F_MinGR = F_MinSH+(gr-1)*Delta
    F_MaxGR = F_MinSH+(gr )*Delta
    IF F_MinGR <= M_Summa .AND. M_Summa <= F_MaxGR
        M_Name = A_FNRus[1]+": {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
        SELECT Object;SET ORDER TO 1;T=DBSEEK(M_Name)
        IF T
            IF ASCAN(A_Obj, Kod) = 0
                AADD(A_Obj, Kod)
            ENDIF
        ENDIF
    ENDIF
NEXT

SELECT ObInfZag
APPEND BLANK
REPLACE Kod_ist WITH ++M_KodIst
REPLACE Name_ist WITH M_NameIst
FOR jj=1 TO LEN(A_Obj)
    FIELDPUT(jj+2,A_Obj[jj])
NEXT

***** формирование массива кодов признаков из БД Card
M_KodPr = {}
FOR jj=2 TO 7          && Начало цикла по полям БД Card
    SELECT Card
    Fv =FIELDGET(jj)
    *IF 2 <= jj .AND. jj <= 3
        IF jj = 2      && берем только стаж, т.к. у мощности бредовые значения
            SELECT Card_sh
            DBGOTO(1);F_MaxSH = FIELDGET(jj)
            DBGOTO(2);F_MinSH = FIELDGET(jj)
            Delta = (F_MaxSH-F_MinSH)/N_Grad
            FOR gr=1 TO N_Grad
                F_MinGR = F_MinSH+(gr-1)*Delta
                F_MaxGR = F_MinSH+(gr )*Delta
                IF F_MinGR <= Fv .AND. Fv <= F_MaxGR
                    M_Name = A_FNRus[jj]+": {"+ALLTRIM(STR(F_MinGR,11,3))+", "+ALLTRIM(STR(F_MaxGR,11,3))+"}"
                    SELECT Priz_per;SET ORDER TO 1;T=DBSEEK(M_Name)
                    IF T
                        IF ASCAN(M_KodPr, Kod) = 0
                            AADD(M_KodPr, Kod)
                        ENDIF
                    ENDIF
                ENDIF
            NEXT
        ENDIF
    ENDIF
NEXT

IF jj > 3
    M_Name = A_FNRus[jj]+"-"+FIELDGET(jj)
    SELECT Priz_per;SET ORDER TO 1;T=DBSEEK(M_Name)
    IF T
        IF ASCAN(M_KodPr, Kod) = 0
            AADD(M_KodPr, Kod)
        ENDIF
    ENDIF
ENDIF
NEXT

***** Запись массива кодов признаков из БД &Fns в БД ObInfKpr
SELECT ObInfKpr
APPEND BLANK
FIELDPUT(1,M_KodIst)
k=2
FOR jj=1 TO LEN(M_KodPr)
    IF k <= 12
        FIELDPUT(k++,M_KodPr[jj])
    ELSE
        APPEND BLANK
        FIELDPUT(1,M_KodIst)
        k=2
        FIELDPUT(k ,M_KodPr[jj])
    ENDIF
NEXT

p=++Num/N_Rec*100;p=IF(p<=100,p,100)
@24,0 SAY STR(p,3)+"%" COLOR "w+r+"
@24,4 SAY REPLICATE("█",0.76*p) COLOR "rg+/n"

SELECT Card
DBSKIP(1)
ENDDO

@24,0 SAY REPLICATE("█",80) COLOR "gb+/n"
Mess = " ПРОЦЕСС ГЕНЕРАЦИИ ЗАВЕРШЕН УСПЕШНО !!! "
@24,40-LEN(Mess)/2 SAY Mess COLOR "rg+/rb"

INKEY(0)

RESTSCREEN(0,0,24,79,scr23)
CLOSE ALL
QUIT

```

В результате работы данного программного интерфейса автоматически получают исходный справочник классов распознавания, справочник признаков, а также обучающая выборка, представляющая собой закодированные в соответствии с этими справочниками страховые случаи (см. таб.1, таб.2, таб.3):

Таблица 1 – СПРАВОЧНИК КЛАССОВ РАСПОЗНАВАНИЯ

22-05-07 12:05:20 г. Краснодар

N п/п	Код класса	Наименование класса распознавания	Степень редукции образа	Абсол. кол-во	% к кол физичес анкет
1	1	Сумма страховой выплаты: : {0.000, 0.000}.....	0.01916	64995	99.200
2	2	Сумма страховой выплаты: : {34.000, 573575.000}.....	0.15524	540	0.800
3	3	Сумма страховой выплаты: : {34.000, 57388.100}.....	0.15300	469	0.700
4	4	Сумма страховой выплаты: : {57388.100, 114742.200}....	0.13388	58	0.100
5	5	Сумма страховой выплаты: : {114742.200, 172096.300}...	0.05552	11	
6	6	Сумма страховой выплаты: : {172096.300, 229450.400}...	0.00000		
7	7	Сумма страховой выплаты: : {229450.400, 286804.500}...	0.05630	1	
8	8	Сумма страховой выплаты: : {286804.500, 344158.600}...	0.00000		
9	9	Сумма страховой выплаты: : {344158.600, 401512.700}...	0.00000		
10	10	Сумма страховой выплаты: : {401512.700, 458866.800}...	0.00000		
11	11	Сумма страховой выплаты: : {458866.800, 516220.900}...	0.00000		
12	12	Сумма страховой выплаты: : {516220.900, 573575.000}...	0.03674	1	

Универсальная когнитивная аналитическая система НПП *ЭЙДОС*

Таблица 2 – СПРАВОЧНИК НАИМЕНОВАНИЙ ШКАЛ И ГРАДАЦИЙ ПРИЗНАКОВ (ФРАГМЕНТ)

22-05-07 12:05:25 г. Краснодар

N п/п	Код призна	Наименования шкал и градаций признаков	Интегр. инф-сть	Абсол. кол-во	% к кол физ. анк
[1]	СТАЖ ВЛАДЕЛЬЦА			
1	1	Стаж владельца: {1.000, 5.300}.....	0.078	7474	11.405
2	2	Стаж владельца: {5.300, 9.600}.....	0.004	5820	8.881
3	3	Стаж владельца: {9.600, 13.900}.....	0.038	6005	9.163
4	4	Стаж владельца: {13.900, 18.200}.....	0.193	7348	11.212
5	5	Стаж владельца: {18.200, 22.500}.....	0.067	6022	9.189
6	6	Стаж владельца: {22.500, 26.800}.....	0.012	6004	9.162
7	7	Стаж владельца: {26.800, 31.100}.....	0.038	7436	11.347
8	8	Стаж владельца: {31.100, 35.400}.....	0.225	5952	9.082
9	9	Стаж владельца: {35.400, 39.700}.....	0.134	5948	9.076
10	10	Стаж владельца: {39.700, 44.000}.....	0.039	7442	11.356
[2]	ПРОИЗВОДСТВО			
11	11	Производство-Иномарка.....	0.068	13569	20.705
12	12	Производство-Неизвестно.....	0.001	1493	2.278
13	13	Производство-Россия.....	0.030	50473	77.017
[3]	МАРКА			
14	14	Марка-.....	0.001	1493	2.278
15	15	Марка-Alfa-Romeo.....	0.001	16	0.024
16	16	Марка-Aro.....	0.001	1	0.002
17	17	Марка-Asia.....	0.001	12	0.018
18	18	Марка-Astra.....	0.001	2	0.003
19	19	Марка-Audi.....	0.205	792	1.209
20	20	Марка-Austin.....	0.001	1	0.002
21	21	Марка-BMW.....	0.043	697	1.064
22	22	Марка-Bentley.....	0.001	1	0.002
23	23	Марка-Buick.....	0.001	1	0.002
24	24	Марка-Cadillac.....	0.001	4	0.006

[4]		МАРКА И МОДЕЛЬ			
114	114	Марка и модель--.....	0.001	1493	2.278
115	115	Марка и модель-Alfa-Romeo-156.....	0.001	2	0.003
116	116	Марка и модель-Alfa-Romeo-166.....	0.001	1	0.002
117	117	Марка и модель-Alfa-Romeo-2106.....	0.001	11	0.017
118	118	Марка и модель-Alfa-Romeo-33.....	0.001	1	0.002
119	119	Марка и модель-Alfa-Romeo-75.....	0.001	1	0.002
120	120	Марка и модель-Aro-10.....	0.001	1	0.002
121	121	Марка и модель-Asia-2106.....	0.001	12	0.018
122	122	Марка и модель-Astra-2106.....	0.001	2	0.003
123	123	Марка и модель-Audi-100.....	0.001	177	0.270
124	124	Марка и модель-Audi-200.....	0.001	6	0.009
125	125	Марка и модель-Audi-2106.....	0.001	58	0.089
126	126	Марка и модель-Audi-80.....	0.271	250	0.381
127	127	Марка и модель-Audi-90.....	0.001	6	0.009
128	128	Марка и модель-Audi-A2.....	0.001	1	0.002
129	129	Марка и модель-Audi-A3.....	0.001	14	0.021
130	130	Марка и модель-Audi-A4.....	0.001	82	0.125
131	131	Марка и модель-Audi-A4 Avant.....	0.001	2	0.003
132	132	Марка и модель-Audi-A6.....	0.021	158	0.241

[5]		ЦВЕТ			
888	888	Цвет-Бежевый (оттенки бежевого и светло-золо	0.109	3664	5.591
889	889	Цвет-Белый (оттенки белого).....	0.057	3632	5.542
890	890	Цвет-Бордовый(оттенки бордового и вишневого)	0.030	3594	5.484
891	891	Цвет-Голубой (оттенки голубого).....	0.231	3696	5.640
892	892	Цвет-Жёлтый (оттенки жёлтого и светлозолотис	0.240	3676	5.609
893	893	Цвет-Зеленый (оттенки светло-зеленого).....	0.121	3701	5.647
894	894	Цвет-Коричневый (оттенки коричневого).....	0.018	3523	5.376
895	895	Цвет-Красный (оттенки красного).....	0.045	3671	5.602
896	896	Цвет-Оранжевый (оттенки оранжевого).....	0.132	3640	5.554
897	897	Цвет-Розовый (оттенки розового).....	0.064	3673	5.605
898	898	Цвет-Светло-серый (оттенки светло-серого и с	0.009	3763	5.742
899	899	Цвет-Синий (оттенки синего).....	0.053	3708	5.658
900	900	Цвет-Сиреневый (оттенки сиреневого).....	0.012	3512	5.359
901	901	Цвет-Тёмно-Серый (оттенки тёмно-серого).....	0.020	3708	5.658
902	902	Цвет-Тёмно-зеленый (оттенки тёмно-зеленого).	0.130	3726	5.686
903	903	Цвет-Фиолетовый (оттенки фioletового).....	0.052	3550	5.417
904	904	Цвет-Хамелеон или несколько цветов без преоб	0.026	3508	5.353
905	905	Цвет-Чёрный.....	0.040	3590	5.478

Универсальная когнитивная аналитическая система

НПП *ЭЙДОС*

Таблица 3 – АНКЕТА обучающей выборки № 1

22-05-07 18:16:12

г.Краснодар

Код	Наименования классов распознавания
2	Сумма страховой выплаты: : {34.000, 573575.000}
3	Сумма страховой выплаты: : {34.000, 57388.100}

Код	Градации описательных шкал (признаки)
10	Стаж владельца: {39.700, 44.000}
13	Производство-Россия
93	Марка-ВАЗ
741	Марка и модель-ВАЗ-1111
901	Цвет-Тёмно-Серый (оттенки тёмно-серого)

Универсальная когнитивная аналитическая система

НПП *ЭЙДОС*

Таким образом данным программным интерфейсом полностью автоматизируется этап СК-анализа, называемый "Формализация предметной области".

7. Затем стандартными средствами системы "Эйдос" (режим: _235) был выполнен синтез семантической информационной модели (СИМ).

8. В системе "Эйдос" реализовано много различных методов оптимизации модели: это и исключение из модели статистически малопредставленных классов и факторов (артефактов), и исключение незначимых факторов, и ремонт (взвешивание) данных, что обеспечивает не только классическую, но и структурную репрезентативность исследуемой выборки по отношению к генеральной совокупности, и итерационное разделение классов на типичную и нетипичную части. Последний метод и был использован для оптимизация СИМ. В результате категории были разделены на типичные и нетипичные части и был получен следующий справочник классов (таблица 4):

Таблица 4 – СПРАВОЧНИК КЛАССОВ ОПТИМИЗИРОВАННОЙ МОДЕЛИ

22-05-07 18:57:05

г. Краснодар

N п/п	Код класса	Наименование класса распознавания	Степень редукции образа	Абсол. кол-во	% к кол физичес анкет
1	1	Сумма страховой выплаты : {0.000, 0.000}.....	0.00000		
2	16	Сумма страховой выплаты : {0.000, 0.000} it={1}.....	0.11539	48368	73.800
3	23	Сумма страховой выплаты : {0.000, 0.000} it={1,2}.....	0.31439	1218	1.900
4	22	Сумма страховой выплаты : {0.000, 0.000} it={2}.....	0.25730	15409	23.500
5	5	Сумма страховой выплаты : {114742.200, 172096.300}.....	0.07006	11	
6	6	Сумма страховой выплаты : {172096.300, 229450.400}.....	0.00000		
7	7	Сумма страховой выплаты : {229450.400, 286804.500}.....	0.07104	1	
8	8	Сумма страховой выплаты : {286804.500, 344158.600}.....	0.00000		
9	2	Сумма страховой выплаты : {34.000, 573575.000}.....	0.26875	235	0.400
10	13	Сумма страховой выплаты : {34.000, 573575.000} it={1}....	0.08553	236	0.400
11	18	Сумма страховой выплаты : {34.000, 573575.000} it={1,2}..	0.13139	37	0.100
12	19	Сумма страховой выплаты : {34.000, 573575.000} it={2}....	0.07634	32	
13	3	Сумма страховой выплаты : {34.000, 57388.100}.....	0.27906	164	0.300
14	14	Сумма страховой выплаты : {34.000, 57388.100} it={1}.....	0.09376	231	0.400
15	17	Сумма страховой выплаты : {34.000, 57388.100} it={1,2}....	0.13006	40	0.100
16	21	Сумма страховой выплаты : {34.000, 57388.100} it={2}.....	0.07720	34	0.100
17	9	Сумма страховой выплаты : {344158.600, 401512.700}.....	0.00000		
18	10	Сумма страховой выплаты : {401512.700, 458866.800}.....	0.00000		
19	11	Сумма страховой выплаты : {458866.800, 516220.900}.....	0.00000		
20	12	Сумма страховой выплаты : {516220.900, 573575.000}.....	0.04636	1	
21	4	Сумма страховой выплаты : {57388.100, 114742.200}.....	0.18025	46	0.100
22	15	Сумма страховой выплаты : {57388.100, 114742.200} it={1}..	0.05121	9	
23	20	Сумма страховой выплаты : {57388.100, 114742.200} it={2}..	0.06247	3	

Универсальная когнитивная аналитическая система

НПП *Эйдос*

При этом средняя по всей выборке объемом 65535 страховых случаев вероятность правильного отнесения страхового случая к тем классам, к которым он действительно относится, на 2-й итерации составила **88,953%** (на 1-й итерации она составляла всего 23,692%). При дальнейших итерациях эта величина стабилизировалась, поэтому в этот процесс был остановлен. Достигнутая степень адекватности (достоверности) модели оценива-

ется нами как довольно высокая и достаточная для того, чтобы исследование этой модели считать исследованием самой моделируемой предметной области, и выводы, полученные путем исследования модели считать относящимися к самой предметной области.

9. Контрольное измерение адекватности СИМ было проведено на тестовой выборке, в которую вошли **2160** страховых случаев, представляющие все классы, в т.ч. **все** с совершенными ДТП и по остальным классам не более 540 случаев. При этом были получены результаты, представленные ниже:

- ИЗМЕРЕНИЕ АДЕКВАТНОСТИ СИМ МЕТОДАМИ И ТЕХНИКАМИ ЭКСПЕРИМЕНТАЛЬНО-СТАТИСТИЧЕСКОГО И КОМПЬЮТЕРНО-СТАТИСТИЧЕСКОГО ПОДХОДОВ
- Всего объектов в классе: 2160 (100% от п.15)
 Всего объектов в классе: 2160
1. Средняя достоверность измерения значений признаков объектов: 4,88%
 2. Средняя достоверность измерения значений признаков объектов: 0,0%
 3. Средняя достоверность измерения значений признаков объектов: 0,31%
 4. Средняя достоверность измерения значений признаков объектов: 0,88%
 5. Средняя достоверность измерения значений признаков объектов: 5,88%
 6. Средняя достоверность измерения значений признаков объектов: 4,01%
 7. Средняя достоверность измерения значений признаков объектов: 1,72% (100% для п.11 и п.14)
 8. Средняя достоверность измерения значений признаков объектов: 1,72% (100% для п.11 и п.14)
 9. Средняя достоверность измерения значений признаков объектов: 4,01%
 10. Средняя достоверность измерения значений признаков объектов: 1,72% (100% для п.11 и п.14)
 11. Средняя достоверность измерения значений признаков объектов: 2,00% (100% для п.11 и п.14)
 12. Средняя достоверность измерения значений признаков объектов: 1,72% (100% для п.11 и п.14)
 13. Средняя достоверность измерения значений признаков объектов: 4,01%
 14. Средняя достоверность измерения значений признаков объектов: 1,72% (100% для п.11 и п.14)
 15. Средняя достоверность измерения значений признаков объектов: 10,0%
 16. Средняя достоверность измерения значений признаков объектов: 10,0%

22.05.07 17:42:09 Таблица

№ п/п	Класс	Имя признака	Достоверность измерения значений признаков объектов	Средняя достоверность измерения значений признаков объектов по классам				Достоверность измерения значений признаков объектов	Число объектов в классе	Количество объектов в классе					Средняя достоверность измерения значений признаков объектов	Число объектов в классе
				1	2	3	4			5	6	7	8	9		
1	13	Служебный автомобиль	7,4	0,6	0,0	1,4	0,2	30,2	226	226	14	67	23	10,35	0,61	
2	14	Служебный автомобиль	7,3	0,6	0,0	1,3	0,1	39,5	221	216	15	63	23	10,34	0,74	
3	23	Служебный автомобиль	4,5	1,4	0,0	0,4	1,5	69,6	5+0	515	35	30	11,00	1,07		
4	16	Служебный автомобиль	1,0	0,9	0,1	0,9	1,1	37,0	5+0	471	29	60	11,00	1,08		
5	3	Служебный автомобиль	2,5	0,2	0,0	0,5	2,0	56,5	134	129	35	45	13,73	1,18		
6	2	Служебный автомобиль	2,4	0,2	0,0	0,4	2,5	54,0	225	125	70	40	17,00	0,61		
7	21	Служебный автомобиль	1,7	0,2	0,0	1,0	4,5	16,9	14	14	0	67	28	1,74	0,70	
8	17	Служебный автомобиль	1,1	0,2	0,0	1,1	4,0	4,4	4+0	12	0	104	18	1,92	0,38	
9	10	Служебный автомобиль	0,9	0,2	0,0	1,2	1,9	4,4	17	10	7	105	19	1,20	0,31	
10	19	Служебный автомобиль	0,7	0,2	0,0	1,2	1,7	6,4	12	12	0	101	11	1,01	0,32	
11	4	Служебный автомобиль	0,7	0,1	0,0	1,6	0,0	22,0	46	45	1	107	37	1,17	0,38	
12	5	Служебный автомобиль	1,9	0,1	0,0	6,0	1,9	35,0	11	11	0	179	30	1,33	1,01	
13	7	Служебный автомобиль	5,3	0,0	0,0	5,6	0,2	34,4	1	1	0	134	0	0,71	0,10	
14	20	Служебный автомобиль	0,7	0,1	0,0	0,9	0,2	11,7	3	3	0	172	67	1,13	0,31	
15	15	Служебный автомобиль	10,0	0,1	0,0	10,3	0,1	2,2	9	9	0	103	1	1,07	0,00	
16	12	Служебный автомобиль	10,0	0,0	0,0	11,0	0,2	30,1	1	1	0	100	37	1,13	0,10	
Среднеарифметическое значение			4,0	0,6	0,0	0,0	5,4	46,0	197,6	160,4	107,2	45,7	11,0	10,45	1,03	

Информация: количество объектов в классе

ФОРМУЛЫ РАСЧЕТА ПОКАЗАТЕЛЯ ДИСТАНЦИОННОЙ АДЕКВАТНОСТИ (ГО АИИСО 1500):
 C0100 = (C0101 + C0102 + C0103 + C0104) / (C0101 + C0102 + C0103 + C0104 + C0105)
 C0200 = (C0201 + C0202 + C0203 + C0204) / (C0201 + C0202 + C0203 + C0204 + C0205)
 C0300 = (C0301 + C0302) / (C0301 + C0302)
 C0400 = (C0401 + C0402) / (C0401 + C0402)
 C0500 = (C0501 + C0502) / (C0501 + C0502)
 C0600 = (C0601 + C0602) / (C0601 + C0602)
 где E – класс соответствия строки;
 где N – количество объектов в классе; где N – количество объектов в рассматриваемой выборке

ФОРМУЛЫ РАСЧЕТА ПОКАЗАТЕЛЯ ИТЕРАЦИОННОЙ АДЕКВАТНОСТИ (ГО АИИСО 1500) ПО ВСЕМ ОБЪЕКТАМ:
 C1 = C100 / (C100 + C101 + C102 + C103 + C104 + C105)
 где C1 = 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16
 где N – количество объектов в рассматриваемой выборке
 ПРИМЕЧАНИЕ: данные только результаты измерения объектов в рассматриваемой выборке

Из анализа этой формы можно сделать следующие выводы:

– хорошо представленные классы можно использовать при прогнозировании, т.к. достоверность идентификации по этим классам достаточно высокая;

- результаты прогнозирования по слабо представленным классам учитывать в принятии решений нецелесообразно;
- применение модели обеспечивает во много раз более высокую достоверность, чем случайное угадывание или не использование модели;
- общая вероятность достоверной идентификации оказалась несколько ниже, чем по всей выборке, по всей видимости из-за того, что в тестовой выборке не было возможности указать все страховые случаи по которым не было ДТП, из-за огромного количества таких случаев.

10. По сути задача 1: "**Многокритериальная типизация** автомобилей клиентов по типам: "Не совершившие ДТП", "Совершившие ДТП" и категориям, отражающим суммы страховых выплат была решена при синтезе модели на 7-м этапе. Результатом этого этапа и решением 1-й задачи является матрица информативностей, *фрагмент* которой приводится на рисунке 3:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	KOD_PR	OBJ_2	OBJ_3	OBJ_4	OBJ_5	OBJ_7	OBJ_12	OBJ_13	OBJ_14	OBJ_15	OBJ_16	OBJ_17	OBJ_18	OBJ_19	OBJ_20	OBJ_21	OBJ_22	OBJ_23
2	1	0,41	0,46	0,30	-0,08	0,00	0,00	-0,92	-1,16	0,00	0,02	0,31	0,31	0,36	0,00	0,44	-0,29	0,60
3	2	0,02	0,01	-0,11	0,01	0,00	0,00	0,01	0,05	0,08	0,01	-0,45	-0,42	0,13	0,47	0,01	-0,02	-0,14
4	3	-0,23	-0,47	0,13	0,00	0,00	0,00	0,08	0,08	0,07	-0,00	-0,46	-0,43	-0,38	0,00	-0,40	0,03	-0,35
5	4	-0,42	-0,64	-0,33	0,18	0,00	0,78	0,10	0,09	0,24	-0,01	-0,14	-0,11	-0,45	0,00	-0,47	0,04	-0,30
6	5	-0,31	-0,27	-0,26	0,00	0,00	0,00	-0,15	-0,26	0,32	-0,03	0,18	0,06	0,00	0,00	0,00	0,11	-0,46
7	6	-0,01	-0,02	0,13	0,00	0,00	0,00	0,04	0,05	0,00	0,01	-0,07	-0,04	0,12	0,00	0,18	-0,04	-0,20
8	7	-0,31	-0,65	0,05	-0,07	0,00	0,00	0,06	0,07	-0,01	-0,00	-0,14	-0,12	-0,06	0,38	-0,22	0,03	-0,41
9	8	-0,17	-0,14	0,00	0,25	0,85	0,00	-0,00	-0,01	0,07	-0,00	0,04	0,07	-0,13	0,46	-0,39	0,03	-0,40
10	9	-0,23	-0,10	0,00	0,00	0,00	0,00	-0,00	-0,01	0,07	-0,01	-0,21	-0,18	-0,37	0,00	-0,39	0,06	-0,40
11	10	0,07	0,05	0,11	0,17	0,00	0,00	-0,03	0,00	0,00	0,01	-0,04	-0,01	-0,06	0,00	-0,08	-0,02	-0,20
12	11	0,12	0,15	0,14	-0,29	0,00	0,00	-0,47	-0,43	0,00	-0,29	0,26	0,29	-0,42	0,00	-0,44	0,33	0,48
13	12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
14	13	-0,02	-0,04	-0,03	0,07	0,09	0,09	0,08	0,08	0,09	0,06	-0,10	-0,12	0,08	0,09	0,08	-0,25	-0,49
15	14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
16	15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
17	16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
18	17	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
19	18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
20	19	-0,12	-0,24	0,00	0,73	0,00	0,00	0,00	0,00	0,00	0,00	0,26	0,29	0,00	0,00	0,00	0,49	0,51
21	20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
22	21	0,31	0,38	0,26	0,00	0,00	0,00	0,00	0,00	0,00	-0,08	0,31	0,34	0,00	0,00	0,00	0,12	0,45
23	22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
24	23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
25	24	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	0,00
26	25	0,00	0,00	0,00	0,00	0,00	0,00	0,25	0,26	0,00	-0,02	0,00	0,00	0,00	0,00	0,00	0,07	-0,09

Рисунок 3. Фрагмент матрицы информативностей

В этой матрице столбцы соответствуют классам распознавания, строки – градациям факторов, а в клетках на их пересечении приведено *количество информации* в битах, которое содержится в факте обнаружения в страховом случае определенной градации фактора (например, водительского стажа, марки, модели или цвета автомобиля) о том, что этот случай относится к определенному классу. Приведен лишь фрагмент этой матрицы, т.к. в ней **905** строк.

11. **Задача 2:** "Разработка методики *прогнозирования* риска совершения ДТП и суммы страховой выплаты на основе информации о клиенте и его автомобиле", решается по сути автоматически при синтезе модели на 7-м этапе СК-анализа. В системе "Эйдос" есть стандартный режим _42, обеспечивающий подсчет для каждого страхового случая (представленного в распознаваемой выборке) суммарного количества информации, которое содержится в его признаках о принадлежности данного случая к каждому из классов. Все классы сортируются (ранжируются) в порядке убывания суммарного количества информации, содержащегося в описании страхового случая, о принадлежности к ним. Эта информация представляется в виде экранной формы и файла (см. рис. 4):

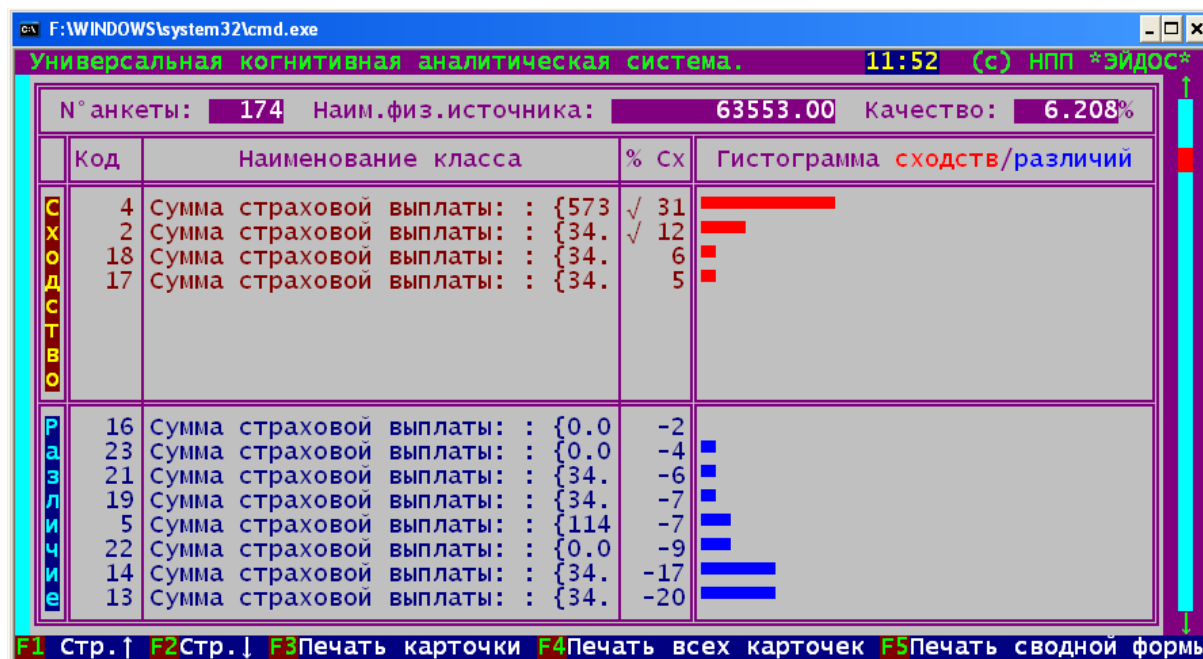


Рисунок 4. Пример экранной формы с результатами прогнозирования для тестового страхового случая с номером 174.

На рисунке 4 птичками "√" отмечены классы, к которым данный страховой случай действительно относится. Эта же информация приводится и в файле, в котором наименования классов приводятся не в сокращенном варианте:

Результаты идентификации информационного источника с классами распознавания 11:54:17
22-05-07

Номер анкеты: 174 Паин.физ.источника: 63553.00 Качество результата распознавания: 6.200%

Код	Наименование класса распознавания	% Сх	Гистограмма сходств/различий
4	Сумма страховой выплаты: : {57388.100, 114742.200}.....	√ 31	
2	Сумма страховой выплаты: : {34.000, 573575.000}.....	√ 12	
18	Сумма страховой выплаты: : {34.000, 573575.000} it={1,2}.....	6	
17	Сумма страховой выплаты: : {34.000, 57388.100} it={1,2}.....	5	
12	Сумма страховой выплаты: : {516220.900, 573575.000}.....	-0	
7	Сумма страховой выплаты: : {229450.400, 286804.500}.....	0	
20	Сумма страховой выплаты: : {57388.100, 114742.200} it={1,2}.....	1	
15	Сумма страховой выплаты: : {57388.100, 114742.200} it={1}.....	-1	
3	Сумма страховой выплаты: : {34.000, 57308.100}.....	-1	
16	Сумма страховой выплаты: : {0.000, 0.000} it={1,2}.....	2	
23	Сумма страховой выплаты: : {0.000, 0.000} it={1,2}.....	-4	
21	Сумма страховой выплаты: : {34.000, 57308.100} it={2}.....	-6	
19	Сумма страховой выплаты: : {34.000, 573575.000} it={1,2}.....	7	
5	Сумма страховой выплаты: : {114742.200, 172896.300}.....	7	
22	Сумма страховой выплаты: : {0.000, 0.000} it={2}.....	-9	
14	Сумма страховой выплаты: : {34.000, 57308.100} it={1}.....	-17	
13	Сумма страховой выплаты: : {34.000, 573575.000} it={1,2}.....	20	

Универсальная когнитивная аналитическая система НПП «Эйдос»

Если в распознаваемой выборке представлено сразу несколько страховых случаев, то может представлять интерес другая форма вывода информации о результатах прогнозирования по ним:

F:\WINDOWS\system32\cmd.exe Универсальная когнитивная аналитическая система. 11:59 (с) НПП «Эйдос»

Класс: 2 Сумма страховой выплаты: : {34.000, 573575.0} Качество: 2.35%

	Код	Информационный источник	% Сход	Гистограмма сходств/различий
Сходство	2	34563.00	√ 15	
	105	71058.14	√ 14	
	116	1942.00	√ 12	
	174	63553.00	√ 12	
	1158	0.00	10	
	77	45624.00	√ 10	
	135	10284.72	√ 10	
	182	767.00	√ 10	
Различие	557	0.00	-12	
	1379	0.00	-12	
	1480	0.00	-12	
	1504	0.00	-12	
	1305	0.00	-13	
	1750	0.00	-13	
	2864	0.00	-13	
16	7547.00	-14		

F1 Стр.↑ F2 Стр.↓ F3 Печать карточки F4 Печать всех карточек F5 Печать сводной формы

Рисунок 5. Пример карточки идентификации страховых случаев с классом: код 2, "Сумма страховой выплаты: 34-573575 рублей"

По сути этот класс эквивалентен по смыслу классу: "Совершит ДТП".

12. Для решения задачи 3: "Разработка методики *поддержки принятия решений* по выбору контингента клиентов, наиболее предпочтительных и нежелательных для автострахования", необходимо исследовать модель. Это можно сделать используя как возможности системы "Эйдос", так и просто загрузив матрицу информативности в Excel. В результате получаем следующие формы (см. рис. 5-7):

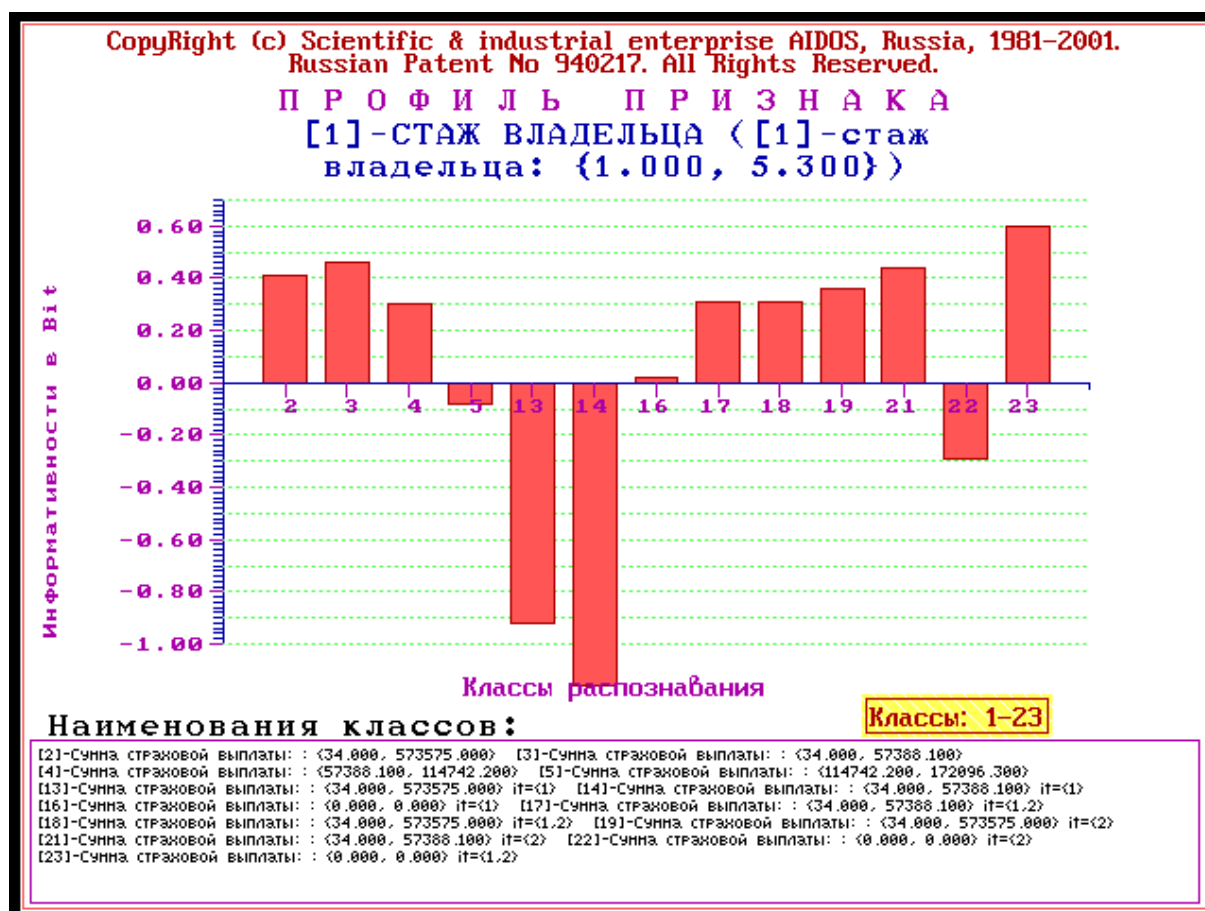


Рисунок 6. Влияние малого стажа на аварийность

Из этой формы видно, каким образом водительский стаж 1-5 лет обуславливает принадлежность страхового случая к классам оптимизированной модели (таблица 4).

Рассмотрим влияние водительского стажа на безаварийность (рис.7) и цвета автомобиля на безаварийность (рис.8).

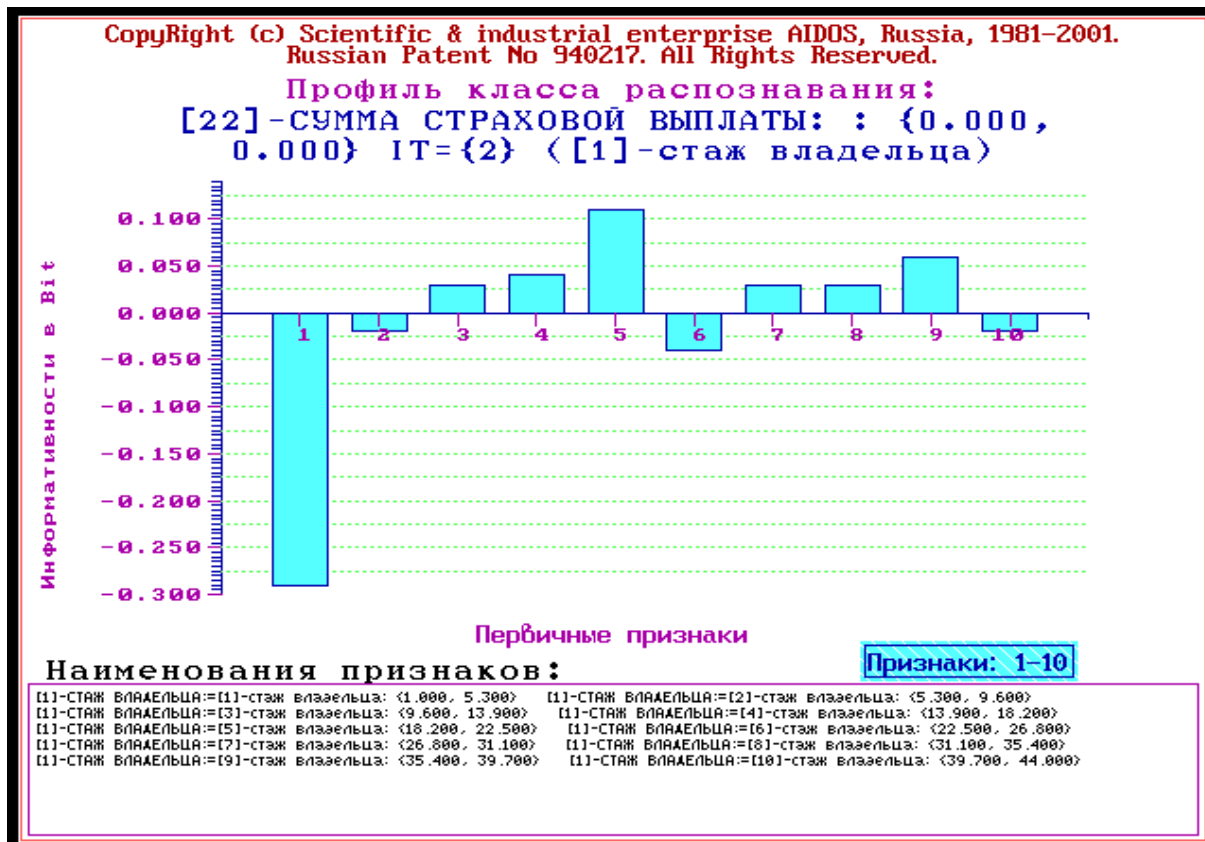


Рисунок 7. Влияние водительского стажа на безаварийность

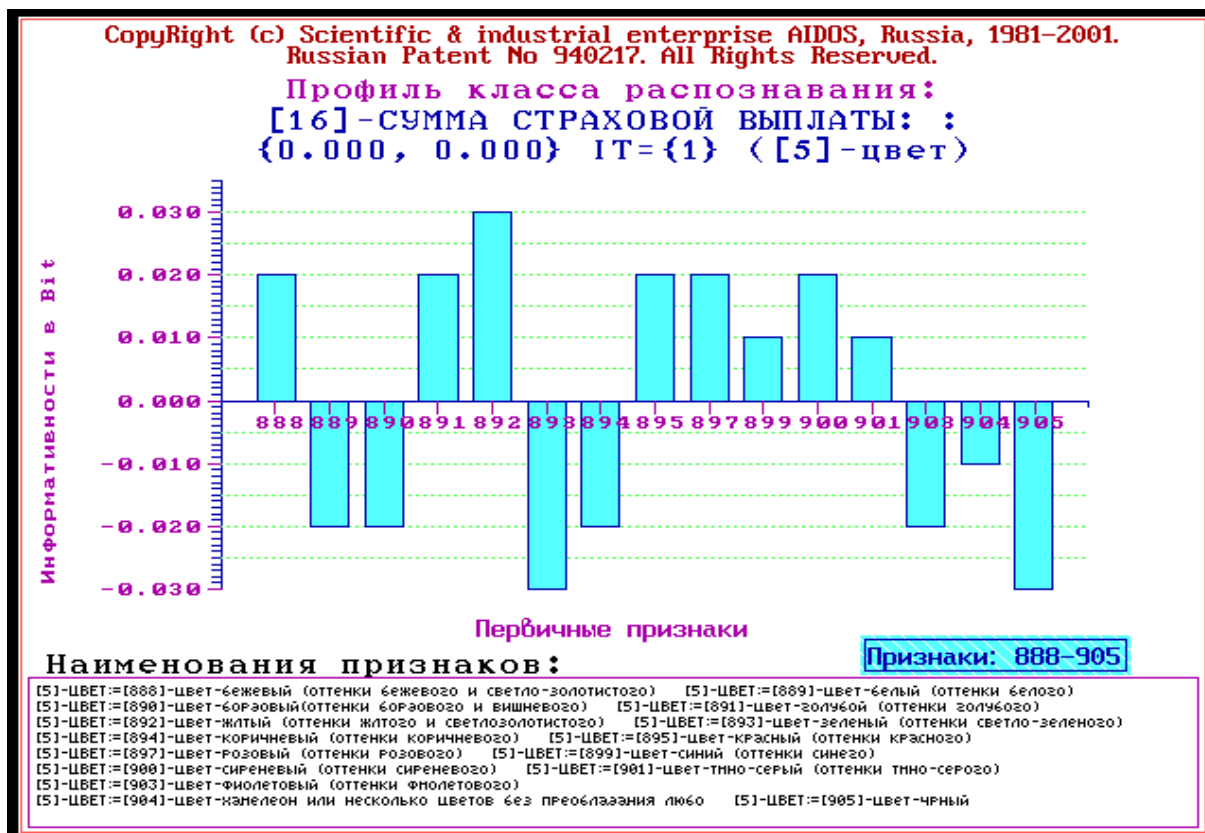


Рисунок 8. Влияние цвета автомобиля на безаварийность

Из рис.7 видно, что обывательское представление о том, что чем больше водительский стаж, тем меньше аварийность, не совсем соответствует действительности. Вернее оно соответствует действительности на интервалах: {1, 22.5} лет и {22.5, 39.7} лет. А вот стаж 22.5-26 лет как это ни парадоксально, несет информацию о том, что водитель не относится к безаварийному классу (как и очень малый стаж от 1 до 5 лет, но в меньшей степени). Правда дальнейшее увеличение стажа до 39.7 лет также постепенно приводит к меньшей аварийности, но стаж 39.7 - 44 года опять говорит о склонности к аварийности. Если причины высокой аварийности при очень малом (до 5 лет) и малом (до 9 лет) стаже понятны: это неопытность и лихачество, то причины аварийности опытных водителей с большим (22-26 лет) и очень большим стажем (39-44 года) видимо кроются в состоянии здоровья, связанном с кризисом среднего возраста и с наступлением старости.

Необходимо отметить, что задача выявления фактически имеющихся зависимостей, и задача содержательного объяснения причин существования именно обнаруженных зависимостей, а не каких-либо других, т.е. задача *содержательной интерпретации обнаруженных зависимостей*, – это совершенно разные задачи. Авторы считают, что задача интерпретации должна решаться специалистами в моделируемой предметной области, в данном случае – специалистами в области автострахования.

Получены также функции влияния на аварийность и безаварийность марки и модели автомобиля, но эти формы не приводятся из-за большой размерности (т.е. очень большого количества моделей и марок). Отметим лишь, что из этих форм следует гипотеза о том, чем дороже автомобиль, тем больше (при всех прочих равных условиях) вероятность обращения в страховую компанию при участии в ДТП.

На рисунке 9 приведем форму, показывающую влияние того, произведен ли автомобиль в России или за рубежом на принадлежность к новым категориям (таб.4).

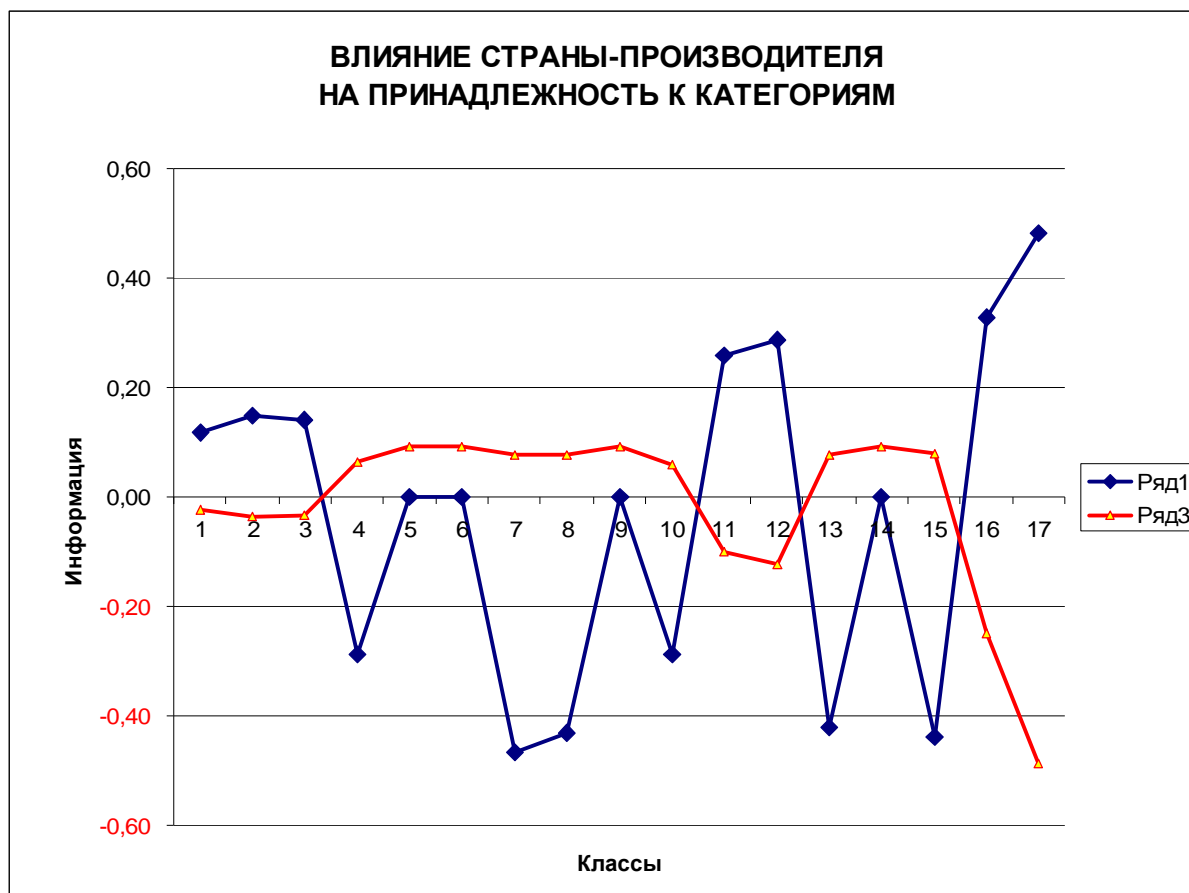


Рисунок 9. Влияние российского (ряд 3) и зарубежного (ряд 1) производства автомобиля на принадлежность к классам оптимизированной модели (таб.4)

13. Основной принцип оценки экономической эффективности разработанной методики (при условии ее применения в деятельности реальной страховой компании) состоит в том, что данная методика позволяет создать научно обоснованный *образ желаемого клиента* (как и образ *нежелаемого*) и это, в сочетании с методами формирования контингента клиентов, позволяют улучшить состав портфеля договоров автострахования, уменьшив в нем долю убыточных и увеличив долю прибыльных договоров, повысив таким образом рентабельность и прибыль компании.

14. При планировании данного исследования авторы ставили цель лишь оценить возможность применения технологии СК-анализа для решения задачи андеррайтинга. Данное исследование показало, что это возможно и перспективно. Но для того, чтобы иметь основания сделать этот вы-

вод не было необходимости проводить полномасштабное научное исследование. Поэтому, естественно, что представленный в работе вариант имеет ряд ограничений и недостатков, в преодолении которых и состоит перспектива ее развития.

В частности можно было бы увеличить объем обучающей выборки, взять *значительно большее* количество параметров, характеризующих как автотранспортное средство, так и его владельца, а также **локализовать** задачу для других регионов. Например, можно было бы учитывать является ли автомобиль переднеприводным, заднеприводным или полноприводным, мощность его двигателя, год изготовления, более детально можно было бы учитывать страну-производитель и т.д. Владельца вообще можно было бы исследовать как личность, как это делается в транспортной психологии, например применив подход, описанный в статьях: [8-15]. Но для достижения целей данной работы этого не требовалось и не делалось.

Таким образом, на основе исследования разработанного упрощенного варианта подсистемы айдерейтинга можно сделать вывод о возможности полномасштабного решения этой задачи методом системно-когнитивного анализа.

Литература

1. Луценко Е.В. Универсальная автоматизированная система распознавания образов "Эйдос" (версия 4.1).-Краснодар: КЮИ МВД РФ, 1995.- 76с.
2. Луценко Е.В. Теоретические основы и технология адаптивного семантического анализа в поддержке принятия решений (на примере универсальной автоматизированной системы распознавания образов "ЭЙДОС-5.1"). - Краснодар: КЮИ МВД РФ, 1996. - 280с.
3. Симанков В.С., Луценко Е.В. Адаптивное управление сложными системами на основе теории распознавания образов. Монография (научное издание). – Краснодар: ТУ КубГТУ, 1999. - 318с.
4. Симанков В.С., Луценко Е.В., Лаптев В.Н. Системный анализ в адаптивном управлении: Монография (научное издание). /Под науч. ред. В.С.Симанкова. – Краснодар: ИСТЭК КубГТУ, 2001. – 258с.
5. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организацион-

- но-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.
6. Луценко Е.В. Интеллектуальные информационные системы: Учебное пособие для студентов специальности 351400 "Прикладная информатика (по отраслям)". – Краснодар: КубГАУ. 2004. – 633 с.
 7. Луценко Е.В., Лойко В.И. Семантические информационные модели управления агропромышленным комплексом. Монография. – Краснодар: КубГАУ. 2005. – 480 с.
 8. Луценко Е.В., Лаптев В.Н., Третьяк В.Г. Прогнозирование качества специальной деятельности методом подсознательного (подпорогового) тестирования на основе семантического резонанса В сб.: "Проблемы совершенствования систем защиты информации, энергоснабжения военных объектов и образовательных технологий подготовки специалистов". Материалы II межвузовской научно-технической конференции. – Краснодар: КВИ, 2001. – С. 127-128.
 9. Луценко Е.В., Третьяк В.Г., Александров С.Г. Модель профессионально-значимых качеств личности сотрудников ОВД Вестник учебного отдела КЮИ МВД РФ. – 2001. – №1. – С. 37-40.
 10. Луценко Е.В., Третьяк В.Г. Анализ профессиональных траекторий специалистов с применением системы "Эйдос" Личность и ее бытие (социально-психологические аспекты бытия личности в местном сообществе): сборник научных работ / Под.ред. З.И.Рябикиной. – Краснодар: КубГУ, 2002. –С. 43-49.
 11. Луценко Е.В., Щукин Т.Н., Дорохов В. Б., Лебедев А.Н. ЭЭГ прогноз успешности выполнения психомоторного теста при снижении уровня бодрствования: постановка задачи Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №04(6). – Режим доступа: <http://ej.kubagro.ru/2004/04/20/p20.asp>.
 12. Луценко Е.В., Щукин Т.Н., Дорохов В. Б., Лебедев А.Н. ЭЭГ прогноз успешности выполнения психомоторного теста при снижении уровня бодрствования: описание эксперимента Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №04(6). – Режим доступа: <http://ej.kubagro.ru/2004/04/21/p21.asp>.
 13. Луценко Е.В., Щукин Т.Н., Дорохов В. Б., Лебедев А.Н. ЭЭГ прогноз успешности выполнения психомоторного теста при снижении уровня бодрствования: анализ результатов исследования Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №04(6). – Режим доступа: <http://ej.kubagro.ru/2004/04/22/p22.asp>.
 14. Луценко Е.В., Лебедев А.Н. Диагностика и прогнозирование профессиональных и творческих способностей методом АСК-анализа электроэнцефалограмм в системе "Эйдос" Межвузовский сборник научных трудов, том 1. – Краснодар: КВИ, 2003. – С. 227-229.
 15. Луценко Е.В., Наприев И.Л., Синтез многоуровневых семантических информационных моделей активных объектов управления в системно-когнитивном анализе Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2007. – №04(28). – Режим доступа: <http://ej.kubagro.ru/2007/04/pdf/11.pdf>

Примечание: для удобства читателей эти и другие работы автора размещены на его сайте по адресу: <http://lc.kubagro.ru/aidos/eidos.htm>