

# **Анализ и прогнозирование расторжения договоров накопительного страхования жизни**

**Т.К. Богданова, доцент НИУ ВШЭ, [tanbog@hse.ru](mailto:tanbog@hse.ru)**

**А.С. Есаулова, выпускница НИУ ВШЭ**

## **Введение.**

В настоящее время на Российском страховом рынке сложилась весьма негативная ситуация. По данным ЦБ РФ темпы прироста страховых взносов (по всем видам страхования), начиная с 2013 года неуклонно падают: совокупное количество страховых договоров в 1-ом полугодии 2015 года уменьшилось на 6,6% по сравнению с 1-ым полугодием 2014 года. Как отмечает РАЕХ (Эксперт РА) рост страхового рынка происходит лишь в номинальном выражении. Так за 1 полугодие 2015 года темпы прироста страховых взносов составили всего 2,1% при инфляции 8,5% (Рейтинговое агентство РАЕХ, 2015). Совокупная величина собранных страховых взносов выросла на 10,5 млрд. рублей, при этом в совокупности рост сегментов страхования жизни, страхования имущества физических лиц и ДМС составил 11,8 млрд. рублей. Рынок ОСАГО вырос на 26,5 млрд. рублей, но исключительно за счет значительного увеличения тарифов. При этом его рост с трудом покрывает сокращение взносов в страховании автокаско (-15,5 млрд. рублей), страховании от НС и болезней (-7,8 млрд. рублей), страховании сельскохозяйственных рисков (-3,4 млрд. рублей) и страховании СМР (-2,9 млрд. рублей). По прогнозу РАЕХ (Эксперт РА), в 2016 году негативные тенденции на страховом рынке сохранятся.

Тем не менее, несмотря на экономический кризис, по итогам 2014 года сегмент страхования жизни обеспечил наибольший прирост взносов по сравнению с другими сегментами рынка. Согласно данным РАЕХ («Эксперт РА») за 2014 год сумма взносов по страхованию жизни составила 10,85 млрд. руб., т.е. по сравнению с 2013 годом прирост взносов составил 27,8%. В 1

полугодии 2015 года темпы прироста взносов по страхованию жизни замедлились до 12,5%, но, несмотря на замедление темпов прироста взносов, сегмент страхования жизни растет заметно быстрее страхового рынка в целом (2,1% за 1 полугодие 2015 года) (Рейтинговое агентство RAEX, 2015).

По базовому прогнозу RAEX (Эксперт РА), темпы прироста страхования жизни в 2016 году составят 9-12%, по негативному прогнозу значения показателей будут равны 2-5%. При этом в 2016 году ожидается ускорение роста выплат в страховании жизни в связи с наступлением сроков выплат по среднесрочным договорам страхования жизни. Таким образом, страховщики, специализирующиеся на страховании жизни, столкнутся с ростом убыточности и сокращением рентабельности собственных средств, что будет оказывать негативное влияние на рейтинги страховых компаний (Рейтинговое агентство RAEX, 2015).

В силу того, что средняя длительность договора накопительного страхования жизни составляет 10-15 лет, одной из особенностей такого рода страхования является долговременное сотрудничество клиентов и страховых компаний. Как следствие, одним из важнейших факторов эффективного функционирования бизнеса является повышение лояльности со стороны клиентов.

*Актуальность исследования* определяется тем, что своевременный и регулярный комплексный анализ договоров страхования, входящих в портфель страховщика, изучение факторов, оказывающих влияние на факт расторжения договоров, а также детальная характеристика портрета клиента, позволят более качественно организовать работу по заключению и сопровождению договоров страхования и взаимодействию с клиентами. Это, в конечном счете, должно способствовать уменьшению процента расторжения договоров и повышению прибыльности страховщиков.

## **1. Причины расторжения договоров страхования жизни**

Страховым случаем в классическом страховании жизни считается смерть или продолжающаяся жизнь (дожитие) застрахованного до окончания

срока страхования (Лельчук А.Л, 2010). Таким образом, риском является не сама смерть, а время ее наступления. На практике договоры страхования жизни заключаются чаще всего как договоры смешанного страхования жизни, где кроме двух основных рисков – дожития или смерти застрахованного, - также осуществляется страхование по риску потери трудоспособности в результате несчастного случая (Архипов А. П., 2012). Такие договоры страхования более привлекательны для клиентов, т.к. в понимании людей, риск наступления несчастного случая и последующей травмы более высок, чем классические риски страхования жизни.

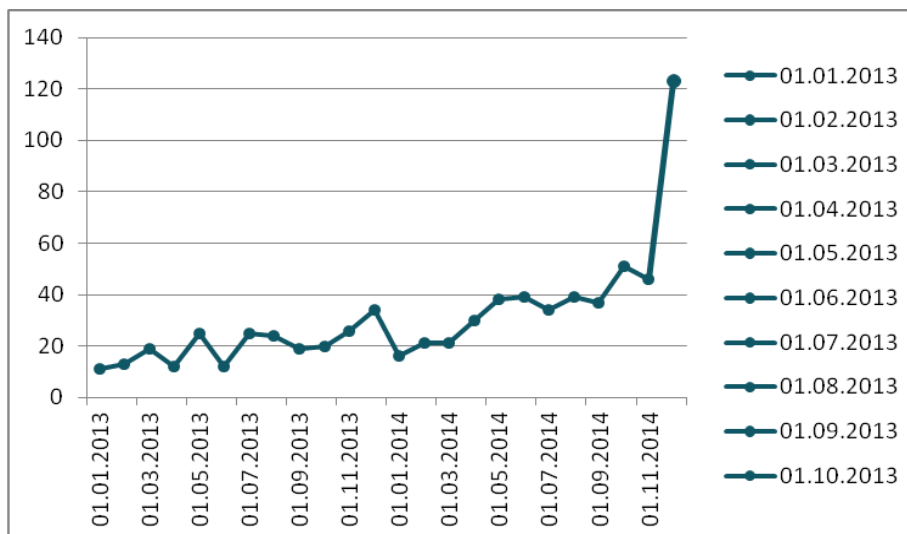
Удержать клиента, который приобрел договор страхования жизни, бывает довольно трудно. Особенно, это касается договоров с регулярной оплатой, когда клиент берет на себя обязательство с определенной периодичностью оплачивать свои страховые взносы страховой компании в течение многих лет. В случае расторжения договора клиент получает тот размер выкупной суммы, который был определен в договоре страхования плюс некоторый инвестиционный доход (Закон РФ «Об организации страхового дела в Российской Федерации», 2015 г.). Размер выкупной суммы даже с дополнительным доходом будет меньше той суммы, которую он заплатил страховой компании. Страховая компания же при досрочном расторжении договора вынуждена изымать средства из страхового резерва, тем самым, теряя средства для инвестирования.

Анализ причин, по которым договоры страхования могут быть расторгнуты, начиная со второго и последующих взносов, показывает, что чаще всего расторгаются договоры из-за неоплаты премии (68%) и по желанию клиента (26%) (рис. 1). Решение клиента о досрочном расторжении договора может стать следствием экономического кризиса в стране. В период нестабильной экономической ситуации, наиболее подвержены риску клиенты, которые приобрели договоры страхования в иностранной валюте. Для них рублевые эквиваленты страховых взносов могут увеличиваться порой в несколько раз.



**Рис. 1. Причины расторжения договоров страхования и их доля среди всех причин расторжения.**

Как видно из рис. 2, к декабрю 2014 года количество заявлений на расторжение договоров, поступивших от клиентов, выросло почти в 5 раз по сравнению со средним значением за предыдущие два года деятельности компании.



**Рис. 2. Частота заявлений на расторжение договоров страхования.**

Анализ соотношения действующих и расторгнутых договоров с регулярной оплатой показал, что действующими являются только 60,67% всего портфеля договоров с регулярной оплатой, а 39,33% договоров являются расторгнутыми. При этом 14,79% составляют договора, расторгнутые по другим причинам. Это расторжения договоров с начала их

действия, отказы от договоров на стадии заявления, а также дожитие договоров до даты окончания. Расторжения договоров с начала действия или отказ от страхования на стадии заявления является не самым убыточным и частым видом расторжений и не входит в рамки данного исследования.

Наибольшую проблему с точки зрения нормального функционирования компании составляет высокая доля расторжений договоров по причинам «расторжение по желанию клиента» и «расторжение из-за неоплаты премии», т.е. ситуация, при которой компания не получает очередных взносов клиентов. Доля расторжений по этим причинам составляет 17,71% и 6,83% соответственно. Именно анализу и прогнозированию договоров расторжения страхования жизни по этим двум причинам посвящено данное исследование.

Конечно, в период кризиса показатель расторжения договоров, безусловно, будет очень высоким. Но, имея инструмент, позволяющий прогнозировать расторжения договоров страхования, можно будет удержать некоторых клиентов, предложив им соответствующие льготы по оплате премий. В качестве вариантов таких льгот страховые компании могут предлагать смену валюты договора или, так называемые, «финансовые каникулы», когда клиенту предоставляется возможность оплатить свой взнос позже, не расторгая при этом договор. Направленная работа с клиентами по договорам, находящимся в группе риска, может повысить лояльность клиентов и понизить долю расторгнутых договоров в общем портфеле компании.

Задачами построения подобных предсказательных моделей занимается прогнозная аналитика. Она представляет собой набор статистических методов моделирования, машинного обучения и Data mining, которые изучают исторические данные для прогнозирования будущих событий (Nuse C., 2007).

## **2. Информационная база исследования**

Для исследования была сформирована информационная база, основой которой послужил портфель договоров страхования одной из крупнейших

страховых компаний России за 2008-2015 гг., состоящий из 15 688 договоров накопительного страхования жизни и 11 переменных, 6 из которых являются характеристикам договоров, а 5 - характеристиками клиентов.

Как видно из табл.1, действуют 9 518 договоров (60,7%), 751 договор (4,8%) аннулирован и 5 419 договоров (34,5%) расторгнуто. Таким образом, на расторгнутые договоры приходится треть всей выборки.

**Таблица 1. Состав договоров страхования.**

Статус	Частота	Проценты	Накопленный процент
Действует	9518	60,7	60,7
Расторгнут	5419	34,5	95,2
Аннулирован	751	4,8	100,0
Всего	15688	100,0	

Договоры, имеющие статус «аннулирован», это договоры, которые были расторгнуты на стадии заявления на страхование и никогда не получали статус договора. Доля таких договоров очень низкая, поэтому они не влияют на эффективность работы страховой компании, и анализ их не имеет смысла.

Среди расторгнутых договоров имеется определенная доля договоров, которые были аннулированы в течение 30 дней с момента действия с полным возвратом премии и договоров, которые уже закончили свое действие в силу наступления определенного страхового события (дожитие клиента до конца срока действия договора, смерть клиента в течение срока страхования). Для последующего анализа эти договора были отнесены в группу прочие причины расторжения (табл. 2).

**Таблица 2. Типы расторжения договоров страхования жизни.**

Тип расторжения договора	Частота	Проценты	Накопленный процент
Расторгнутые по желанию клиентов	1072	17,4	17,4
Расторгнутые из-за неоплаты	2778	45,0	62,4
Прочие причины расторжения	2320	37,6	100,0
Всего	6170	100,0	

Наибольший интерес для исследования представляют две причины расторжения договоров: расторжение по желанию клиента и расторжение из-за неоплаты премии. Из таблицы 2, в которой представлен частотный анализ

типов расторжения договоров страхования жизни, видно, что совокупная доля договоров, расторгнутых по желанию клиента или из-за неоплаты, составляет 62,4%.

Для прогнозирования расторжения договоров по этой причине был проведен анализ на отсутствие взаимосвязей между фактом расторжения договора и различными характеристиками самого договора или клиента, который приобрел договор страхования.

Для анализа были использованы

- следующие характеристики договоров: программа страхования, страховой продукт, город заключения договора, валюта ответственности по договору, периодичность платежа, длительность договора.
- социодемографические характеристики клиентов: пол, возраст на дату заключения договора, семейное положение, должность, сфера занятости.

Анализ взаимосвязи факторов, проведенный с использованием таблиц сопряженности критерия Хи-квадрат (Тюрин Ю.Н., Макаров А.А., 2003), показал, что:

- в целом во всех городах договоры расторгают с одинаковой частотой. В среднем в Москве, Екатеринбурге, Волгограде и Уфе процент расторжения составляет 37,4%. Немного выше - средний процент расторжений в городах Санкт-Петербург, Н.Новгород, Самара и Ростов-на-Дону – 41,8%;
- реже всего среди договоров с регулярной оплатой расторгаются договоры с ежегодной оплатой (34,0%), а чаще всего – договоры с ежеквартальной оплатой (67,8%);
- самый высокий процент расторжений у возрастных категорий от 18 до 24 лет и от 25 до 35 лет – 66,9% и 49,2% расторжений соответственно. С увеличением возраста клиентов число расторжений

договоров уменьшается, так для возрастной группы 75 лет и старше число расторгнутых договоров менее 10%;

- самый высокий – 68,0% процент расторжений договоров у клиентов, которые указали свое семейное положение как «разведен/-на».

Между всеми анализируемыми параметрами выявлена высокозначимая взаимосвязь, т.е. все параметры влияют на факт расторжения договоров.

Также была выявлена высокозначимая связь между фактом расторжения договора и видом страхового продукта.

### 3. Моделирование расторжения договоров страхования жизни.

*Кластерный анализ.* Проведенный по методу Варда кластерный анализ (Бююль А., Цефель П., 2005) исходных данных позволил выделить 4 кластера, примерно равных по числу наблюдений, при 5-ти кластерном решении (табл.3).

Таблица 3. Результаты кластеризации по методу Варда.

Кластеры	Частота	Проценты	Валидный процент	Накопленный процент
Допустимо 1	364	23,5	23,5	23,5
2	540	34,9	34,9	58,4
3	387	25,0	25,0	83,5
4	253	16,3	16,4	99,8
5	3	,2	,2	100,0
Всего	1547	99,9	100,0	
Пропущенные Системные	1	,1		
Всего	1548	100,0		

Полученные кластеры имели значимые различия средних значений по всем исследуемым показателям: по программе страхования, возрасту клиентов, сроку действия договоров, премии первого года и годовой премии за текущий период, доли изменения премии. Значимость различий средних значений показателей была подтверждена результатами непараметрического теста с использованием критерия U Манна-Уитни.



Однако анализ результатов кластеризации показал, что во все кластеры попали как действующие, так и расторгнутые договора. Полученные кластеры не имеют значимых различий по факту расторжения договоров.

**Бинарная логистическая регрессия.** Положительное решение было найдено с использованием бинарной логистической регрессионной модели (Громько Г.Л., 2009).

Вероятность наступления события рассчитывается по формуле:

$$p = \frac{1}{1 + e^{-z}}$$

где  $z$  - стандартное уравнение регрессии, которое рассчитывается по формуле:

$$z = \alpha_0 + k_1 * x_1 + k_2 * x_2 + \dots + k_n * x_n$$

$x_n$  - значения независимых переменных,  $k_n$  - коэффициенты, рассчитанные с помощью бинарной логистической регрессии,  $\alpha_0$  - константа.

Уравнение регрессии имеет следующий вид:

$$z = -1,728 - 0,04 * \text{Программа страхования} +$$

$$+ \begin{bmatrix} -0,485 \\ -0,078 \\ 0,547 \end{bmatrix} * \begin{bmatrix} \text{Страховой продукт (1)} \\ \text{Страховой продукт (2)} \\ \text{Страховой продукт (3)} \end{bmatrix} +$$

$$\begin{bmatrix} 1,217 \\ 0,783 \\ 0,528 \\ 0,616 \\ 0,708 \end{bmatrix} * \begin{bmatrix} \text{Возрастная категория (1)} \\ \text{Возрастная категория (2)} \\ \text{Возрастная категория (3)} \\ \text{Возрастная категория (4)} \\ \text{Возрастная категория (5)} \end{bmatrix} + \begin{bmatrix} -0,449 \\ -0,246 \\ -0,306 \\ -0,461 \\ -0,156 \\ -0,483 \\ -0,049 \end{bmatrix} * \begin{bmatrix} \text{Город (1)} \\ \text{Город (2)} \\ \text{Город (3)} \\ \text{Город (4)} \\ \text{Город (5)} \\ \text{Город (6)} \\ \text{Город (7)} \end{bmatrix} -$$

$$- 0,5 * \text{Длительность в годах} +$$

$$+ \begin{bmatrix} -0,821 \\ -1,014 \\ -0,687 \\ -0,154 \end{bmatrix} * \begin{bmatrix} \text{Категория длительности (1)} \\ \text{Категория длительности (2)} \\ \text{Категория длительности (3)} \\ \text{Категория длительности (4)} \end{bmatrix} +$$

$$+ 0,497 * \text{Длительность текущая в годах} +$$

$$+0,0001 * \text{Премия текущая (руб.)} - 0,360 * \text{Изменение премии} + \\ +1,113 * \text{Регулярность оплаты}$$

Для анализа качества предсказания модели было проведено моделирование с использованием 3-х порогов отсечения: 0,5, 0,45 и 0,4 (табл.4, табл.5, табл.6).

Как видно из табл.4, при пороге отсечения равном 0,5, общая прогнозная точность модели составила 70,2%, действующих договоров 76,3%, расторгнутых договоров – 60,8%.

**Таблица 4. Итоговый результат обработки данных методом бинарной логистической регрессии с пороговым отсечением 0,5.**

Наблюдённые		Предсказанные		
		Статус договоров		Доля правильных предсказаний
		Действующие	Расторгнутые	
Статус договора	Действующие	7262	2251	76,3
	Расторгнутые	2418	3752	60,8
Общая процентная доля				70,2

а. Значение отсечения - ,500

**Таблица 5. Итоговый результат обработки данных методом бинарной логистической регрессии с пороговым отсечением 0,45.**

Наблюдённые		Предсказанные		
		Статус договоров		Доля правильных предсказаний
		Действующие	Расторгнутые	
Статус договора	Действующие	6695	2818	70,4
	Расторгнутые	1837	4333	70,2
Общая процентная доля				70,3

а. Значение отсечения - ,450

Как видно из табл.5, при пороге отсечения равном 0,45, общая прогнозная точность модели составила 70,3%, т.е. практически такая же, как и в предыдущем случае. Но точность предсказания действующих договоров стала хуже и составила 70,4%, однако значительно повысилась точность предсказания расторгнутых договоров – 70,2%.

Как видно из табл.6, при пороге отсечения равном 0,4, общая прогнозная точность модели стала несколько хуже 69,7%, чем в предыдущих случаях. Существенно ухудшилась точность предсказания действующих договоров – 64,9%, однако значительно повысилась точность предсказания расторгнутых договоров – 77,1%.

**Таблица 6. Итоговый результат обработки данных методом бинарной логистической регрессии с пороговым отсечением 0,4.**

Наблюдаемые		Предсказанные		
		Статус договоров		Доля правильных предсказаний
		Действующие	Расторгнутые	
Статус договора	Действующие	6174	3339	64,9
	Расторгнутые	1411	4759	77,1
Общая процентная доля				69,7

а. Значение отсечения - ,400

Таким образом, с уменьшением порога отсечения увеличивается доля правильно предсказанных расторгнутых договоров страхования, но уменьшается доля правильно предсказанных действующих договоров.

Выбор уровня порога отсечения зависит от того, предсказание какого из двух состояний договора представляет интерес для исследователя. Но оптимальным можно считать значение порога отсечения равное 0,45, т.к. в этом случае общая прогнозная точность предсказания как действующих, так и расторгнутых договоров приблизительно одинаковая, порядка 70%.

Прогнозную точность модели бинарной логистической регрессии можно повысить, если учесть влияние таких факторов, как семейное положение, должность и сфера занятости. Поскольку между этими факторами и статусом договора была выявлена высокосвязная взаимосвязь. В данном расчете они не учитывались т.к. для большинства договоров значения этих переменных отсутствовали.

***Предсказание расторжения договора страхования по методу дерева решений.*** В качестве независимых переменных для модели дерева решений (Quinlan J. R., 1986) были выбраны следующие переменные: программа

страхования, группа страхового продукта, пол клиента, возраст клиента, семейное положение, должность, сфера занятости клиента, город заключения договора, длительность договора, валюта ответственности по договору, периодичность оплаты, факт расторжения договора, размер премии первого года в рублях, текущий размер годовой премии в рублях, процент изменения текущей премии относительно премии первого года.

Анализируемая выборка для построения дерева решений была разделена на обучающую (70%) и контрольную (30%). С использованием двух методов: метода CHAID (Kass G. V., 1980) и метода CRT (Breiman L., Friedman J. H., Olshen R. A., & Stone C. J., 1984) были построены деревья решений глубиной 5 уровней.

Общая прогнозная точность моделей, как для обучающей, так и для контрольной выборок, при использовании обоих методов около 70%. Оба метода дали очень хорошую прогнозную точность для предсказания действующих договоров: метод CHAID - 82,9% для обучающей выборки и 82,2% - для контрольной; метод CRT – 84,3% для обучающей выборки и 83,8% - для контрольной. Но практически эти методы не работают для предсказания расторгнутых договоров: метод CHAID - 51,9% для обучающей выборки и 49,2% - для контрольной; метод CRT – 48,5% для обучающей выборки и 48,2% - для контрольной.

## **Заключение**

В данном исследовании проведена детальная классификация договоров страхования жизни, дана оценка текущего состояния страхового рынка в России и, в частности, сегмента страхования жизни; проанализирована структура страхового портфеля компании. Выявлено, что наибольшую доходность компании приносят договоры с регулярной оплатой премии, что по очередным взносам доходность компании выше, чем по взносам первого года, что почти треть договоров с регулярной оплатой премии являются расторгнутыми.

Было выявлено высокозначимое взаимовлияние параметров договоров и характеристик клиентов на факт расторжения договоров страхования.

На основе иерархического метода кластеризации договоров страхования с использованием метода Варда были выявлены кластеры, значимо различающиеся по программе страхования, возрасту клиентов, сроку действия договоров, премии первого года и премии текущей, доли изменения премии, но не имеющие значимых различий по факту расторжения договоров.

Для прогнозирования расторжения договоров страхования были разработаны модели на основе бинарной логистической регрессии и метода деревьев решений.

Модель бинарной логистической регрессии при пороговом отсечении равном 0,4 показала достаточно высокие результаты для прогнозирования расторжений договоров страхования, точность предсказания расторгнутых договоров составила 77,1%.

Модели деревьев решений, построенные по методу CHAID и CRT с глубиной дерева равной 5 уровням, показали высокие результаты для предсказания действующих договоров страхования. Для модели, построенной по методу CHAID, процент правильных предсказаний для действующих договоров составил около 82, как для обучающей, так и для контрольной выборки. Но модель не позволяет прогнозировать расторжение договоров.

***Практическая значимость исследования.*** Безусловно, само по себе предсказание расторжения и/или продолжающего действия договора не может повысить эффективность работы компании. Но на основе разработанных моделей и полученных результатов можно выработать рекомендации менеджменту страховой компании и другим заинтересованным лицам, касающиеся продуктовой линейки и страховых тарифов. С маркетинговой точки зрения, результаты проведенных исследований можно использовать для создания портретов потребителей с

различной степенью «надежности» клиентов. Использование предсказанных значений будущих статусов договоров в работе по сбору очередных страховых взносов позволит страховой компании снизить показатели убыточности договоров, повысить лояльность своих клиентов, и это, в конечном счете, будет способствовать повышению эффективности бизнеса и конкурентоспособности компании.

### *Литература.*

1. Рейтинговое агентство RAEX – Исследование российского рынка страхования [Электронный ресурс]. - RAEX, 2015. – Режим доступа: <http://www.raexpert.ru/researches/insurance>, свободный.
2. Закон РФ «Об организации страхового дела в Российской Федерации» № 4015-1 от 27.11.1992 г. в ред. от 08.03.2015 г.
3. Архипов А. П. Страхование: Учебник. – Кнорус, 2012. –288 с.
4. Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. - СПб.: ООО "ДиаСофтЮП", 2005, 608 с.
5. Громько Г.Л. Теория статистики. Практикум. – М.: ИНФРА-М, 2009. 240 с.
6. Лельчук А.Л. Страхование жизни - М.: Анкил, 2010, 456 с.
7. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. / Под ред. В.Э. Фигурнова. - М.: ИНФРА-М, 2003, 544 с.
8. Breiman L., Friedman J. H., Olshen R. A., & Stone C. J. Classification and regression trees. Florence, KY: Wadsworth, 1984.
9. Kass G. V. An Exploratory Technique for Investigating Large Quantities of Categorical Data // Applied Statistics, 1980, Vol. 29, № 2, pp. 119-127.
10. Nyce C. Predictive Analytics. – American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, 2007.
11. Quinlan J. R. Induction of Decision Trees // Machine Learning, 1986, Vol. 1, № 1, pp. 81-106.