

Оценка качества и сравнение скоринговых карт

© 2009 С.В. Уланов

кандидат экономических наук, доцент
Сарапульский политехнический институт

В статье предлагается комплексная методика оценки кредитных скоринговых карт, основанная на таких критериях, как обычное и кумулятивное распределение кредитов, статистика К-С, ROC-анализ и Lift-кривая. На примере показывается, как сравнивать две скоринговые карты, построенные разными методами - статистическим и экспертным.

Ключевые слова: банковский менеджмент, скоринг, оценка кредитоспособности, математические методы.

Введение

Кредитный скоринг как технология массовой оценки кредитоспособности заемщика получила распространение в нашей стране значительно позже, чем на Западе, в силу того, что активно банки стали развивать розничный бизнес только последнее десятилетие¹. Скоринг же позволяет решить важную задачу - обеспечить хорошее качество кредитного портфеля при высокой скорости принятия решений о предоставлении кредитов, поэтому число отечественных банков, внедряющих скоринг в программы потребительского и беззалогового кредитования, с каждым годом увеличивается, хотя официальная статистика на этот счет отсутствует.

Из-за того, что технология скоринга имеет зарубежное происхождение, до сих пор в ней нет единой терминологии, поэтому часто специалисты по розничным банковским рискам используют англоязычные понятия без перевода на русский язык (примеры: *approval rate*, *проблема reject inference*, *cut-off*, *LTI* и др.). Кроме того, по-разному трактуют и само понятие скоринга, включая в него процедуры верификации и андеррайтинга кредитных заявок. Говоря о скоринге, мы будем отталкиваться от следующих определений². **Скоринг** - это метод оценки благонадежности клиента на основании обработки информации о поведении аналогичных клиентов в прошлом либо экспертных знаний. **Скоринговая модель** - это математическая модель, предсказывающая, вернет или нет клиент кредит в срок. Такие модели строятся либо на основе статистических данных по кредитным историям, либо (реже) с применением методов обработки экспертных оценок. На выходе модель формирует некоторый интегральный показатель (рейтинг), указывающий на степень риска, связанного с

данным клиентом, либо сразу выносятся решение о принадлежности клиента к одному из двух классов - "плохой", "хороший". Для перехода от интегрального показателя каждого клиента к его классу делают сравнение с точкой отсечения, которая рассчитывается из отношения, сколько в среднем нужно заемщиков, которые не допускают просроченной задолженности, для того чтобы компенсировать убытки от одного дефолтного кредита³.

Классическим представителем скоринговой модели является *скоринговая карта*. Анализируя литературу по скорингу, в которой сосредоточен мировой опыт его использования⁴, можно сказать, что балльные скоринговые методики лидируют по количеству и частоте использования, уступая нейросетевым и другим моделям машинного обучения. Отчасти причины этого находятся в консервативном восприятии новых алгоритмов в банковском риск-менеджменте.

Как правило, при наличии кредитных историй современные скоринговые карты строятся с использованием аппарата множественной логической регрессии. Будем называть такие карты *статистическими*. И если алгоритмическая база скоринга сегодня хорошо проработана, то вопросы об оценке их качества и сравнения часто не включаются в стандартный инструментальный кредитного аналитика. Из-за того, что современные скоринговые технологии представляют собой сплав из различных областей, таких как математическая статистика, машинное обучение, методы оптимизации, регрессионный анализ, финансы, возникает потребность в наличии комплексной методики, позволяющей проводить оценку и сравнение нескольких скоринговых карт для выбора наилучшей.

¹ Коновалихин М.Ю., Сергиенко Д.О., Кулик В.В., Кремлева И.В. Подходы к построению скоринговых моделей // Управление финансовыми рисками. 2007. № 01(09).

² Там же.

³ Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. СПб., 2009.

⁴ Руководство по кредитному скорингу / Под ред. Э. Мэйнз. Минск, 2008.

Методы валидации и сравнения скоринговых карт

Большинство методов валидации и сравнения скоринговых карт основываются на анализе качества модели с непрерывным выходом, который трактуется как рейтинговый балл. Разделение клиентов на два класса - благонадежные и неблагонадежные - производится при помощи порогового балла, что представляет собой отдельную задачу и не связано напрямую с понятием качества скоринговой карты.

Основой для построения скоринговой карты является использование двух множеств - рабочего и тестового. На тестовом множестве оценивается обобщающая сила карты, т.е. определяется, насколько модель уловила закономерности исследуемого процесса при установленном пороге отсека. Это, по сути, оценка работы бинарного классификатора на основе трех главных показателей.

1. Общая ошибка классификации

$$E = (1 - k / m) \cdot 100\%,$$

где m - общее число примеров в тестовом множестве;
 k - число верно классифицированных примеров.

2. Уровень одобрений скоринговой карты

$$AR = (1 - b_m / m) \cdot 100\%,$$

где b_f - число отказов клиентам по скорингу.

3. Уровень дефолтов скоринговой карты

$$BR = (b / (k - b_m)) \cdot 100\%,$$

где b - фактическое число дефолтных кредитов.

На рабочем множестве анализируется то, насколько эффективно скоринговая карта разделяет заемщиков на классы. К обязательным процедурам оценки скоринговой карты можно отнести статистическую оценку значимости коэффициентов и самой логрегрессионной модели, а также расчет доверительных интервалов. Однако эти процедуры подходят только для статистических скоринговых карт (т.е. полученных на прошлых кредитных историях). Возникает проблема, как сравнить две карты, одна из которых статистическая, а другая экспертная. Нужны критерии, пригодные для любых скоринговых карт. Выделим четыре таких критерия:

1. **Кривые распределения для хороших и плохих кредитов.** Это зависимости числа наблюдений от скорингового балла. Чем более явно разделены два данных распределения, тем более эффективна скоринговая карта.

2. **Распределение кумулятивных процентов для хороших и плохих кредитов.** Такой график показывает зависимость скорингового балла от CE и CNE (CE - доля хороших кредитов накопленным итогом, CNE - доля плохих кредитов накопленным итогом). Вертикальное расстояние

между точками кумулятивных кривых есть уровень статистики $K - C$. Эта статистика измеряет способность скоринговой карты разделять классы, и на практике ее диапазон варьирования - от 25 до 75. Чем выше значение статистики $K - C$, тем лучше качество скоринговой карты. В месте максимального значения $K - C$ лежит оптимальный скоринговый балл, минимизирующий суммарную ошибку классификации I и II рода.

3. **ROC-кривая и индекс AUC.** Предварительно введем понятия ошибок I и II рода. Для каждого скорингового балла отсека $c_i = c_{i-1} + \Delta$, $i = 1, \dots, n$ получим соответствующий бинарный классификатор, характеристиками которого являются два типа ошибок:

ошибка I рода P_1 - интересное событие ошибочно не обнаружено (выдача кредита заемщику, у которого будут значительные просрочки или дефолт по кредиту);

ошибка II рода P_2 - интересное событие присутствует, но не обнаружено (отказ хорошему заемщику).

Формулы для расчета данных ошибок имеют вид

$$P_1 = bg / (bb + bg); P_2 = gg / (gg + gb),$$

где bb и gg - количество плохих и хороших кредитов, соответственно, предсказанных верно;

bg - число плохих кредитов, предсказанных как хорошие;

gb - число хороших, предсказанных как плохие.

Совокупность точек (P_1^i, P_2^i) образует ROC-кривую. Численный показатель площади под ROC-кривой называется AUC-индексом и рассчитывается любым численным методом, например методом трапеций:

$$AUC = \sum_i (0,5 \cdot (P_1^{i+1} + P_1^i) (P_2^{i+1} - P_2^i)).$$

Значение AUC-индекса, большее 0,8, свидетельствует о хорошей прогностической силе скоринговой карты. Часто ROC-кривую строят отдельно на рабочем и на тестовом множестве.

4. **Lift-кривая.** Формируется на основе лифт-фактора, который в машинном обучении при решении задачи оптимизации массовой рассылки был впервые определен как показатель, отражающий увеличение числа откликов относительно числа почтовых отправлений. Lift-кривая строится следующим образом: по горизонтальной оси откладывается размер выборки в процентах от общего числа наблюдений в списке примеров, упорядоченном по убыванию вероятности наступления положительного исхода. По вертикальной оси фиксируется кумулятивное число положительных исходов в каждой подвыборке. Вообще, Lift-кривая призвана выбрать подмножество при-

меров, которые имеют наибольшую вероятность положительного исхода. Эта задача не так актуальна в скоринге, как, к примеру, в директ-маркетинге, однако по площади под кривой можно тоже проводить сравнение моделей.

Практический пример

Пусть имеются две скоринговые карты, параметры которых приведены в таблице. Баллы первой получены на основе модели множественной логистической регрессии и представляют собой регрессионные коэффициенты, баллы второй - путем опроса эксперта. Эксперту предлагалось только проставить веса в уже упорядоченные признаки заемщика и решать, что больше влияет на положительную кредитоспособность, например "женат/замужем" или "другое" в семейном положении, ему не приходилось. То есть эксперт выносил решение о том, насколько больше влияние одного признака на другой. Заметим, что на практике правильно (с точки зрения статистических закономерностей в данных) упорядочить значения признаков тоже представляет собой трудную задачу, результаты решения которой прямым образом влияют на качество скоринговой карты.

Требуется оценить эффективность каждой скоринговой карты для выбора наилучшей.

Принцип интерпретации коэффициентов по статистической скоринговой карте следующий. Например, рассмотрим образование заемщика. Рассчитаем отношения шансов: $OR(\text{среднее}) = e^0 = 1$; $OR(\text{среднее специальное}) = e^{0,94} = 2,6$; $OR(\text{высшее}) = e^{1,86} = 6,2$. Их интерпретация следующая: при фиксированных значениях других переменных у клиента с высшим образованием шансы стать благонадежным заемщиком в 6,2 раза выше по сравнению с тем, у кого образование среднее. В экспертной же модели клиент с высшим образованием получит 10 баллов к общему рейтингу.

Для оценки эффективности моделей по описанным выше для каждой скоринговой карты критериям понадобится два поля:

1) фактический класс заемщика (плохой, хороший);

2) кредитный рейтинг, рассчитанный на основе скоринговой карты.

На рис. 1 приведены графики кривых распределения кредитов. Видно, что в экспертной модели хуже разделяются распределения плохих и хороших кредитов, но в статистической ско-

Параметры двух скоринговых карт

Группа признаков заемщика	Признак	Коэффициент	Экспертный балл
Начальный балл (константа)	-	-6,39	0
Проживание	Аренда	0	0
	Муниципальное жилье	1,42	5
	Собственник	1,94	10
Срок проживания в регионе	Менее года	0	0
	От 1 до 5 лет	1,06	5
	Свыше 5 лет	1,47	10
Стаж работы на последнем месте	Менее года	0	0
	От 1 до 5 лет	0,63	5
	Свыше 5 лет	1,12	10
Возраст	От 50	0	0
	До 35	0,35	5
	От 35 до 49	0,90	10
Семейное положение	Разведен/вдовство	0	0
	Другое	0,6	0
	Холост/не замужем	1,25	5
	Женат/замужем	1,79	10
Образование	Среднее	0	0
	Среднее специальное	0,94	5
	Высшее	1,82	10
Уровень должности	Сотрудник	0	0
	Руководитель среднего звена	0,4	5
	Руководитель высшего звена	0,47	10
Кредитная история	Отрицательная	0	0
	Нет данных	1,99	5
	Положительная	4,22	20
Отношение "платеж/доход", %	От 40	0	0
	От 30 до 40	0,45	5
	От 20 до 30	1,15	7
	До 20	1,60	10



а)



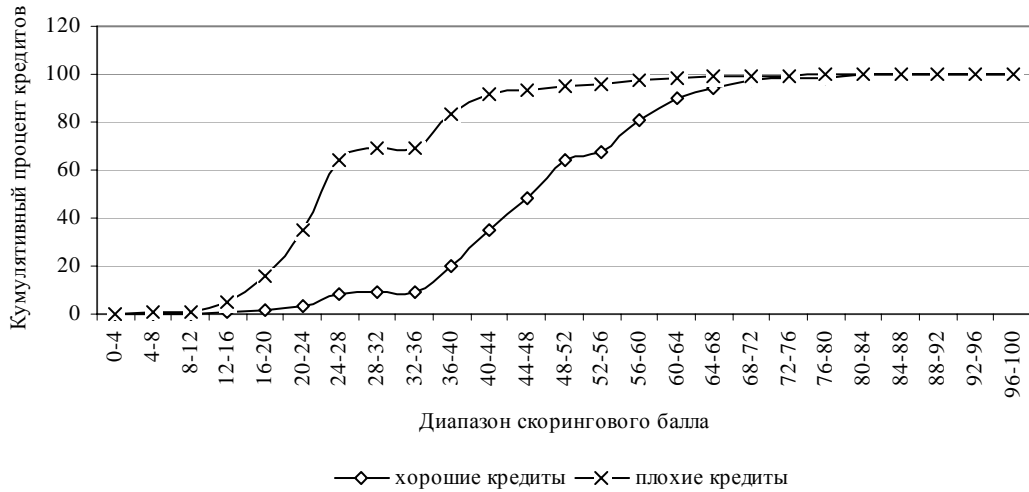
б)

Рис. 1. Кривые распределения для кредитов:
а) - логистическая регрессия; б) - экспертная модель



а)

Рис. 2. Кумулятивные кривые распределения для кредитов (начало):
а) - логистическая регрессия



б) **Рис. 2. Кумулятивные кривые распределения для кредитов (окончание):**
б) - экспертная модель

ринговой карте большое число плохих кредитов сконцентрировано в верхнем диапазоне скорингового балла, что можно объяснить характером логит-преобразования в уравнении регрессии. Это подтверждает и форма кривых распределения кумулятивных процентов (рис. 2), из которых следует, что экспертная модель демонстрирует более крутой подъем, поскольку в ней большее число плохих кредитов принадлежит области самых низких значений скорингового балла. Значение статистики для логрессионной скоринговой карты К-С максимально в диапазоне 76-80 баллов и равно 72,6. Для экспертной модели оно составляет 63 в диапазоне 36-40 баллов.

При визуальной оценке ROC-кривых расположение относительно друг друга указывает на их сравнительную эффективность. Кривая, расположенная выше и левее, свидетельствует о

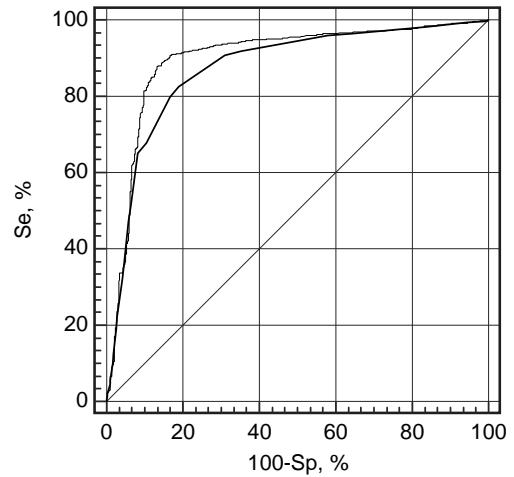


Рис. 3. Сравнение ROC-кривых для двух скоринговых карт

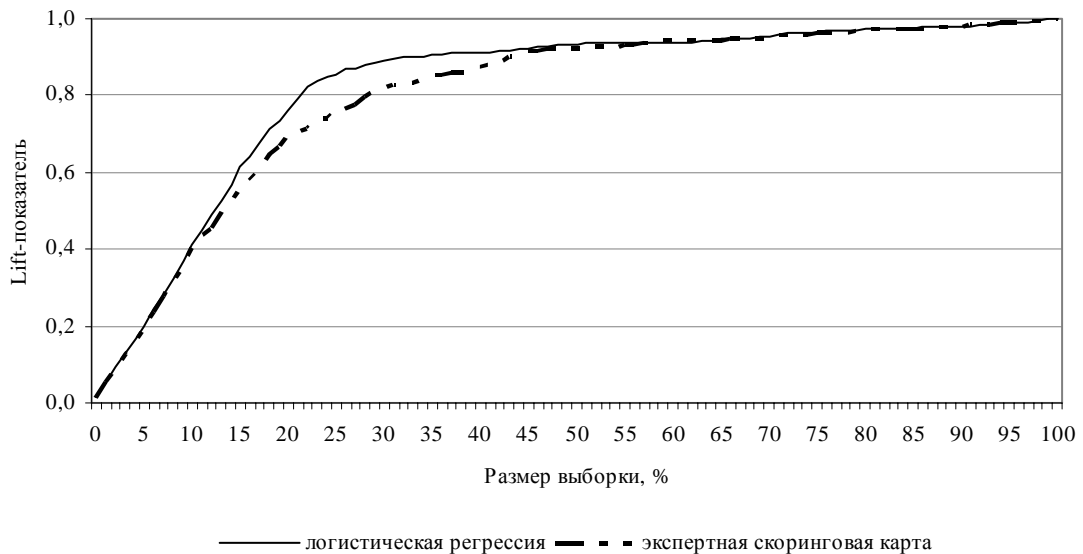


Рис. 4. Сравнение Lift-кривых для двух скоринговых карт

большей предсказательной способности модели. Так, на рис. 3 видно, что статистическая скоринговая карта с $AUC = 0,9$ превосходит экспертную модель, а следовательно, обладает лучшей прогностической силой.

Lift-кривые имеют схожую с ROC-кривой форму: с 10-го по 40-й процентилю статистическая скоринговая карта имеет больший лифт-показатель, чем экспертная модель (рис. 4).

Предложенную методику сравнения и оценки скоринговых карт целесообразно применять в следующих случаях:

- сравнение скоринговых моделей, разработанных различными методами, в том числе экспертными, для принятия решения о выборе наилучшей;
- оценка пригодности использования сторонней скоринговой карты в новых условиях (другой регион, другая страна и т.п.).

Поступила в редакцию 04.08.2009 г.