

ОЦЕНКА КРЕДИТНОГО РИСКА (CREDIT SCORING)

Игорь Енюков

генеральный директор ООО «СтатПоинт»

В настоящее время кредитование населения и малого бизнеса становится одним из основных направлений активности в банковском секторе. Предоставление кредита является достаточно рискованной операцией, поэтому желание максимально уменьшить связанные с ней риски вполне естественно. Один из подходов, получивший название Credit Scoring и позволяющий существенно продвинуться в этом направлении, базируется на применении методов классификации многомерных наблюдений. Этот подход широко используется зарубежными банками и, вероятно, будет полезен для российских банков.

Credit Score (CS) представляет собой числовой индикатор, оценивающий кредитный риск, связанный с предоставлением кредита физическому или юридическому лицу. Величина индикатора (*score*) обычно вычисляется на основе некоторого набора объективных показателей, характеризующих объект кредитования (заемщика), значения которых известны. В качестве таких показателей для физических лиц используются возраст, семейное положение, уровень дохода и т. д., для юридических лиц — балансовые показатели и, возможно, какие-то личностные оценки руководства фирмы.

Индикатор обычно нормируется так, что его значения расположены в интервале от 0 до 1 и интерпретируются как вероятность невозврата кредита заемщиком. Иногда в качестве индикатора используется величина, получаемая вычитанием значения вышеопределенного индикатора из 1. Она интерпретируется как вероятность возврата кредита — 0 (кредит не будет возвращен) и 1 (кредит будет возвращен). В данной статье индикатор интерпретируется как вероятность

невозврата кредита (т. е. как индикатор риска).

Возможность оценки кредитного риска позволяет банку минимизировать финансовые потери, увеличить доходность, уменьшить процентные ставки по кредиту и тем самым улучшить маркетинговую привлекательность. Далее рассматриваются два подхода к решению данной проблемы.

ПОДХОД, БАЗИРУЮЩИЙСЯ НА ИСПОЛЬЗОВАНИИ РЕШАЮЩИХ ПРАВИЛ

Задача построения подобного индикатора тесно связана с разработкой решающих правил (РП) (или в другой терминологии — правил классификации), позволяющих отнести потенциального заемщика в один из двух классов (категорий) — «хороший» и «плохой».

«Хороший» заемщик означает возврат кредита в срок (*duly*), «плохой» — невозврат (*default*) или какие-либо иные осложнения. Результат применения РП можно рассматривать как вырожденный индикатор, который принимает только

2 значения — 0 (минимальный риск) или 1 (максимальный риск).

В большинстве случаев РП достаточно для принятия решения о предоставлении кредита или отказе в нем. Однако, если необходимо, РП можно преобразовать и в непрерывную индикаторную функцию (*scoring function*), принимающую непрерывный ряд значений в интервале от 0 до 1.

Это можно сделать, исходя из следующих соображений. Поскольку РП (для двух классов) [1] имеет вид:

где f — функция РП, зависящая от типа РП, c — пороговое значение, то в качестве индикаторной можно использовать, например, функцию вида

$$I = \begin{cases} 0 & \text{если } f \leq c \\ 1 & \text{иначе} \end{cases} \quad (2)$$

где множитель b выбирают таким образом, чтобы индикатор принимал значения в интервале от 0 до 1 и значения,

близкие к 0 для «хороших» объектов и близкие к 1 для «плохих».

В настоящее время известно большое число типов РП. Оценка параметров РП проводится на основе так называемых обучающих выборок (ОВ), которые в рамках данной задачи представляют собой наборы кредитных историй с измеренными показателями и характеристикой возврата *duly* или *default*.

Окончательная оценка качества РП должна обязательно проводиться на тестовой выборке (экзамене, независимом контроле), т. е. на объектах (кредитных историях), не используемых при обучении (при оценке параметров РП).

При выборе одного из РП среди нескольких выбирается то, для которого ошибка классификации (*error rate*), оцененная на тестовой выборке минимальна. Если такую выборку организовать трудно, например, из-за недостаточного числа кредитных историй, то в качестве оценки ошибки на независимом контроле можно использовать оценки ошибок классификации, полученные с помощью методов кросс-валидации или *bootstrap*.

При кросс-валидации обучение повторяется несколько раз, и при каждом обучении некоторое фиксированное количество объектов, выделяется для теста. При этом объекты, уже использованные для тестирования на предыдущих обучениях, в новую тестовую выборку не включаются. Объем такой тестовой выборки мал (это может быть даже один объект). Оценку ошибки на контроле получают усреднением ошибок на тестовых выборках по всем повторам обучения.

Bootstrap осуществляется аналогично с той только разницей, что каждый раз тестовые объекты выбирают из совокупности всех имеющихся объектов случайным образом, следовательно, тестовые выборки на повторях могут содержать одинаковые объекты.

Подход, базирующийся на формализации экспертных мнений

Применение предыдущего подхода требует наличия ОВ, т. е. некоторого множества кредитных историй с известным завершением. Как быть, если таких выборок нет или их размер мал? Здесь возможно построение индикатора на основе того или иного способа формализации мнений экспертов. Далее рассмотрим два подхода: формирование псевдообучающих выборок (псевдо-ОВ) и подход, основанный на формировании экспертами образа идеального заемщика («Идеал») и его противоположности — очень плохого заемщика («Антипод»).

Псевдообучающие выборки. В этом случае необходимо иметь некоторую выборку кредитных историй, но не обязательно завершенных. Эксперт, исходя из своих экспертных представлений, присваивает им завершения — *duly* или *default*. Это может делаться и на стадии решения о предоставлении кредита, так что для потенциально плохих заемщиков кредит может в действительности не предоставляться. Так, получаются две выборки, которые можно назвать псевдообучающими. Далее задача решается так же, как и при нали-

используются как дискретные количественные.

Поскольку для примера 1 имеются ОВ, воспользуемся этим, чтобы оценить эффективность данного подхода. Для этого определим $Y(X)$ как отношение расстояния Махаланобиса от объекта X до «Идеала» к расстоянию до «Антипода», и проведем линейный ДА с одной переменной $Y(X)$. Таблица частот классификации представлена в табл. 2 (использовалась система *TeleStat* [2]). Клетки вне диагонали содержат частоты ошибочной классификации.

Таблица 1. ХАРАКТЕРИСТИКИ ОБРАЗОВ «ИДЕАЛА» И «АНТИПОДА»

Показатель	Значение для «Идеала»	Значение для «Антипода»
Возраст	2 (23–40 лет)	1 (менее 23 лет)
Семейное положение	2 (состоит в браке)	3 (разведен)
Стаж клиента в банке	5 (более 12 лет)	1 (нет или менее 1 года)
Получение зарплаты через банк	1 (получает)	2 (не получает)
Сберегательный счет в банке	4 (более 100 КФ)	1 (менее 10 КФ)
Служебное положение	1 (руководитель)	2 (служащий)
Текущий счет (в среднем)	3 (более 5 КФ)	1 (менее 2 КФ)
Изменения счета (в среднем)	4 (более 50 КФ)	1
Накопленный дебит	1	3
Перерасход	2 (не допускал)	1 (допускал)
Чеки	1 (имеет чековую книжку)	2 (не имеет чековой книжки)

Таблица 2. ЧАСТОТА КЛАССИФИКАЦИИ ДЛЯ РП, ИСПОЛЬЗУЮЩЕГО ОБРАЗЫ «ИДЕАЛА» И «АНТИПОДА»

Показатель	«Хороший»	«Плохой»	% ошибки
«Хороший»	171	66	0,2785
«Плохой»	61	170	0,2641
% ошибки	0,2629	0,2797	0,2714

чи реальных ОВ. Следует заметить, что РП, полученные таким способом, скорее моделируют работу эксперта, чем непосредственно отражают реальные риски. Но это оправдано, если, например, псевдо-ОВ формируются опытными экспертами, а полученные РП затем используются менее опытными работниками банка.

«Идеал» и «Антипод». В этом случае эксперты формируют две искусственные кредитные истории, одна из которых представляет собой историю для «Идеала», вторая — для «Антипода». Необходимо иметь также и набор историй без завершения, по сути, просто заполненных реальными заемщиками форм со значениями показателей x_1, \dots, x_p . Эти истории необходимы для оценки матрицы связей (ковариаций) между переменными x_1, \dots, x_p и введения метрики Махаланобиса [1].

В табл. 1 приведены возможные образы для «Идеала» и «Антипода» для показателей примера 1, см. ниже.

Заметим, что при вычислении расстояний значения номинальных переменных предварительно кодируются (оцифровываются) [1] и далее используются как количественные. Ординальные переменные

Доля ошибочной классификации равна 0,2741, что в данном случае существенно меньше, чем при случайном угадывании (вероятность ошибки 0,5).

Мониторинг РП. По мере накопления массива кредитных историй ранее разработанные РП необходимо корректировать. Корректировка может проводиться как с целью уточнения РП с прежним набором показателей (пересчет параметров по новой ОВ), так и с добавлением в РП новых переменных.

Пример 1. Попытаемся разработать РП для определения надежности клиентов, обращающихся в банк с просьбой о предоставлении потребительского кредита, используя данные о 468 клиентах одного из французских банков, которым этот банк выдавал кредит в начале 1990-х гг. Клиенты разделены на две группы: «Хорошие» (с которыми банк не имел проблем) и «Плохие» (с которыми возникали проблемы).

Количество «Хороших» составило 237, «Плохих» — 231.

Заметим, что эта выборка представляет собой часть большого архива кре-

дитных историй и сформирована следующим образом: количество хороших клиентов почти равно количеству плохих, что не отражает реального распределения типов клиентов – обычно число хороших клиентов намного превышает количество плохих (пример 2).

Для каждого клиента имеется 13 показателей, из них 11 ординальные или номинальные и 2 количественные. Всего имеется 14 показателей (1-я дополнительная переменная – тип клиента), описание которых приведено ниже:

Таблица 3. РЕЗУЛЬТАТЫ ИСПОЛЬЗОВАНИЯ НЕСКОЛЬКИХ ТИПОВ РП ДЛЯ ДАННЫХ ПРИМЕРА 1

РП	Обучение	Тест	Ранг
Линейные дискриминантные функции (ДФ)	0,209	0,250	5
Квадратичные ДФ	0,20	0,218	2
Нейронная сеть	0,17	0,18	1
Радиальные базисные функции	0,215	0,222	3
Дерево классификаций (CART)	0,153	0,304	7
Метод ближайших соседей (3 соседа, нормализованные переменные)	0,258	0,272	6
Логистическая регрессия	0,167	0,226	4

- 1. Тип.** 2 градации:
 - 1) «Хороший», («Надежный клиент»),
 - 2) «Плохой», («Ненадежный клиент»).
- 2. Возраст.** 4 градации (ординальная):
 - 1) меньше 23 лет,
 - 2) от 23 до 40 лет,
 - 3) от 40 до 50 лет,
 - 4) более 50 лет.
- 3. Семейное положение.** 4 градации (номинальная):
 - 1) «Холост»,
 - 2) «Брак»,
 - 3) «Разведен»,
 - 4) «Вдов».
- 4. Стаж клиента в банке.** 5 градаций (ординальная):
 - 1) 1 год или меньше,
 - 2) от 2 до 4 лет,
 - 3) от 4 до 6 лет,
 - 4) от 6 до 12 лет,
 - 5) более 12 лет.
- 5. Получение зарплаты через банк.** 2 градации:
 - 1) получает,
 - 2) не получает.
- 6. Сберегательный счет в банке.** 4 градации (ординальная):
 - 1) нет сбережений,
 - 2) менее 10 КФ,
 - 3) от 10 до 100 КФ,
 - 4) более 100 КФ.
- 7. Служебное положение.** 3 градации (номинальная):
 - 1) руководитель,
 - 2) служащий,
 - 3) другое.

- 8. Текущий счет (в среднем).** 3 градации (ординальная):
 - 1) меньше 2 КФ,
 - 2) от 2 до 5 КФ,
 - 3) более 5 КФ.
- 9. Изменения счета (в среднем).** 4 градации (ординальная):
 - 1) менее 10 КФ,
 - 2) от 10 до 30 КФ,
 - 3) от 30 до 50 КФ,
 - 4) более 50 КФ.
- 10. Накопленный дебит.** 3 градации (ординальная):
 - 1) менее 40 КФ,
 - 2) от 40 до 100 КФ,
 - 3) более 100 КФ.
- 11. Перерасход.** 2 градации:
 - 1) не было,
 - 2) имел место.
- 12. Чеки.** 2 градации:
 - 1) имеется чековая книжка,
 - 2) нет чековой книжки.

Далее идут 2 количественные переменные, значения которых определяются самим банком и не расшифровываются.

- 13. «XXXX».**
- 14. «YYYY».**

Возможно, что это индексы надежности клиентов, разработанные в самом банке.

РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ МЕТОДОВ ПОСТРОЕНИЯ РП, РЕАЛИЗОВАННЫХ В СИСТЕМЕ TELESTAT

Для обработки вышеописанного набора данных были использованы методы построения РП, реализованные в систе-

ме TeleStat [2] (версия 2.2). Для получения оценки ошибки на независимом контроле использовался *bootstrap*. Для этого обучение проводилось 10 раз. Каждый раз 50 объектов случайным образом выделялись для тестовой выборки. Окончательная оценка ошибки на тесте получена усреднением ошибки по 10 тестовым выборкам размера 50. Ошибка обучения приведена для РП, оцененного по всем объектам.

Процедуры оценки РП использовали без дополнительной настройки, т. е. с теми параметрами (число соседей и радиальных функций, тип расстояния, число слоев нейронов и т. д.), которые по умолчанию предлагались системой TeleStat. В частности, не производилась оптимизация по составу переменных.

Показатели XXXX и YYYY при получении РП не использовались.

В табл. 3 приведены результаты для 7 типов РП. В столбцах «Обучение» и «Тест» приведены соответственно доли ошибочной классификации для ОБ и на контроле.

В столбце «Ранг» указан ранг метода в соответствии с возрастанием ошибки на тесте.

Описание использованных РП можно найти в [1–4].

Наименьшая ошибка на контроле отмечена у нейронной сети (ранг 1). По умолчанию использовалась нейронная сеть без скрытых слоев, так что ее можно рассматривать как вариант линейной ДА [5]. Отличие состоит в разном подходе к обработке неколичественных показателей.

На последнем месте «Дерево классификаций»: этот метод дает наименьшую ошибку на обучении и наибольшую на тесте.

Результаты применения РП, в особенности нейронной сети, следует признать вполне удовлетворительными.

УЧЕТ НЕРАВНЫХ ЦЕН ОШИБОК

В реальной ситуации ошибки в принятии решения могут иметь различную цену. В ситуации «предоставить кре-

Возможность оценки кредитного риска позволяет банку минимизировать финансовые потери, увеличить доходность, уменьшить процентные ставки по кредиту и тем самым улучшить маркетинговую привлекательность.

дит/отказать в кредите» потери, связанные с предоставлением кредита «плохому» заемщику, скорее всего (в среднем), превышают потери при отказе в кредите «хорошему» заемщику.

Относительная стоимость потерь задается с помощью матрицы стоимостей ошибок (*cost matrix*). Для примера приведем матрицу стоимостей ошибок, предложенную экспертами одного из банков:

	«Хороший»	«Плохой»
«Хороший»	0	1
«Плохой»	5	0

Таким образом, ошибка отнесения «плохого» заемщика к «хорошим» признана в 5 раз более существенной, чем ошибка отнесения «хорошего» заемщика к «плохим».

При наличии матрицы стоимости качество принятия решения оценивается уже не по величине *error rate*, а по средней стоимости потерь, которая получается в результате суммирования произведений частот ошибок на соответствующие им стоимости и последующего деления на объем выборки. Рассмотрим применение матрицы стоимостей ошибок для примера 2.

Пример 2. Далее будут использованы данные о клиентах одного из немецких банков. Подробное описание и ссылки, где можно найти эти данные приведены на сайте [6]. Как и в примере 1, имеется две группы клиентов — «Хорошие» и «Плохие» — 700 и 300 соответственно.

Таким образом, размер группы «Плохих» существенно меньше, чем группы «Хороших». Предположим, что это соотношение отражает реальное распределение клиентов по типу. Тогда качество РП имеет смысл соотносить с так называемым правилом классификации по умолчанию.

Когда стоимости ошибок одинаковы, правило по умолчанию предлагает относить объект к наиболее часто встречаемой категории. Применительно к данным примера 2, использование этого правила состоит в том, что любого клиента считают хорошим (кредит выдают всем). Тогда доля ошибок составит 0,3. Однако с учетом стоимостей ошибок получаем условные потери: $5 \cdot 0,3 = 1,5$.

При неравных стоимостях ошибок правило по умолчанию выбирается из условия минимизации условных потерь. Это

Таблица 4. ЧАСТОТА КЛАССИФИКАЦИИ ДЛЯ ЛДА

	«Хороший»	«Плохой»
«Хороший»	754	282
«Плохой»	106	358

достигается, если всех клиентов считать плохими (т. е. не выдавать кредита никому), условные потери составят 0,7. Таким образом, в этом случае данное правило и будет правилом по умолчанию. Желательно, чтобы РП, выбранное для задачи оценки кредитного риска, давало меньший уровень потерь, чем правило по умолчанию.

Для разработки РП можно использовать 20 переменных ((7 количественных, 13 качественных (ординальных и номинальных)). Некоторые из них близки по смыслу с показателями примера 1. Данные из этого примера подробно анализировались в работе [6]. При анализе было использовано 22 статистических системы и было найдено, что наилучший результат достигается при применении линейного дискриминантного анализа (ЛДА). Величина условных потерь на контроле составила 0,535 (Table 9.18 из [6]).

Система *TeleStat* дает следующий результат. В табл. 4 представлена матрица частот классификации для контроля (*bootstrap*, 10 повторений с случайным выбором тестовой выборки размера 150).

Подсчет условных потерь дает $(282 + 5 \cdot 106)/1500 = 0,541$, что с учетом статистической погрешности, практически совпадает с результатом из [6] и ниже уровня потерь по умолчанию (0,7). Таким образом, можно заключить, что применение РП может уменьшить потери, связанные с выдачей кредита.

Приведенные в статье примеры применения статистического подхода в сочетании с методами формализации экспертных мнений для оценки кредитного риска показывают, что этот подход может помочь в принятии правильного решения о предоставлении/отказе в кредитовании и тем самым уменьшить потери кредитора. ■

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная

статистика. Классификация и сокращение размерностей. М.: Финансы и статистика, 1989.

2. Енюков И. С., Ретинская И. В., Скуратов А. К. Статистический анализ и мониторинг научно-образовательных Интернет-сетей. М.: Финансы и статистика, 2004.

3. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002.

4. Day N. E., Kerridge D. F. A general maximum likelihood discriminant. 1967. *Biometrics*. Vol. 23. P. 313–324.

5. Masters T. Neural, Novel & Hybrid Algorithms for Time Series Prediction. NY: John Wiley & Sons, 1995.

6. Machine Learning, Neural and Statistical Classification //Ed. D. Michie, D. J. Spiegelhalter, C. C. Taylor. 1994, February 17. <http://www.amsta.leed.ac.uk/~charles/statlog/>

Ошибка отнесения «плохого» заемщика к «хорошим» признана в 5 раз более существенной, чем ошибка отнесения «хорошего» заемщика к «плохим».