

Рындина С.В. Представление данных в application-скоринг. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей VIII Всерос. научно-техн. конф. – Пенза: ПДЗ, 2008. – С. 214-215.

ПРЕДСТАВЛЕНИЕ ДАННЫХ В APPLICATION-СКОРИНГЕ

С.В. Рындина

Пензенский государственный педагогический университет
им. В.Г. Белинского,
г. Пенза

Application-скоринг – метод классификации заемщиков по вероятности возвращения кредита. При разбиении заемщиков на группы анализируются социально-финансовые показатели, значения которых получают из анкет, заполняемых заемщиками при получении кредита. Эти данные дополняются также информацией из других источников.

К социальным показателям относятся: возраст, пол, место проживания, сколько лет проживает по данному адресу, место работы, профессия, сколько лет работает на последнем месте работы, количество детей, возраст детей, возраст родителей, семейное положение, возраст супруга, наличие судимости.

К финансовым показателям относятся: уровень подтвержденных доходов клиента, доход супруга или как вариант суммарный доход семьи, стоимость жилья, стоимость машины, стоимость другого имущества (дача, гараж и т.п.), находится ли имущество под арестом или в залоге.

Достоверность полученных от потенциального заемщика данных должна проверяться кредитным менеджером. Даже при экспресс-кредитовании есть возможность сделать контрольные звонки по месту работы и месту проживания и сверить по электронным справочникам номера стационарных телефонов с указанными адресами и названиями организаций.

В профиль заемщика следует ввести также дополнительные поля: идентификатор кредитного менеджера, выдавшего кредит, и личное впечатление (экспертная оценка). Скоринг-программа не общается с клиентом напрямую и не учитывает многих психологических моментов, поведенческих реакций. Эта информация считывается кредитным менеджером в той мере, которая зависит от его профессионализма. Насколько можно доверять экспертной оценке конкретного кредитного менеджера, покажет дальнейшая история отношений с его клиентами. Так как эти данные регистрируются, но на начальных этапах не учитываются программой скоринга, то в будущем можно будет оценить процент верных попаданий. Правда, только на тех заемщиках, вероятность возвращения кредита которыми достаточно высока для того, чтобы он был ими получен. В дальнейшем, если экспертная оценка окажется значимой характеристикой, ее можно использовать для уточнения оценочной карты. Причем вес этого критерия должен зависеть от того, кто из кредитных менеджеров дал экспертную оценку потенциальному заемщику на основании предыдущей истории его работы с клиентами. Барьером на пути злоупотреблений станет анализ эффективности предыдущей работы и совпадение ранее проставленных оценок с реальностью. Если процент верных оценок снизится, то и вес последующих оценок кредитного менеджера будет уменьшен.

Также в модели как дополнительный фактор может учитываться положительная история кредитования по предыдущим займам. Банки крайне неохотно обмениваются информацией о заемщиках, бюро кредитных историй пополняется тоже медленно. Поэтому данный фактор будет иметь значение скорее в том случае, если предыдущий кредит заемщик брал в этом же банке.

Часть факторов, учитываемых в модели, может описываться количественными переменными: возраст, доход, стоимость имущества, сколько лет заемщик работает на последнем месте и т.д. Часть факторов: пол, профессия, место работы – качественные факторы, которые задаются с помощью фиктивных переменных. Причем для учета влияния качественных факторов на малых выборках лучше провести предварительную группировку их значений. Профессия, место работы – это значимые факторы, однако принимающие довольно большое количество значений (имеющие большое число уровней градации). Для описания влияния качественных факторов обычно вводится некоторое количество бинарных фиктивных переменных (на единицу меньше, чем уровней градации). Так, для пяти возможных значений профессии (продавец, учитель, врач, программист, водитель) потребуются четыре бинарные переменные (каждая принимает два возможных значения, например, 1 – если заемщик продавец, и 0 – если заемщик не продавец); пятый уровень соответствует случаю, когда все четыре введенные переменные равны нулю (не продавец, не учитель, не врач, не программист – значит водитель). При постоянном объеме выборки наблюдений значимость факторов снижается, если увеличивается число переменных, их описывающих. Поэтому предварительно стоит провести социально-демографическое исследование региона, в котором работает банк, а также учесть отраслевую специфику региона. Это исследование позволит объединить некоторые значения категориальных факторов в один уровень. Отраслевая специфика региона – важный показатель для такого объединения. Одна и та же профессия в разных регионах может относиться к разному уровню градации. Аналогичным образом разбиения на уровни проводятся и для других качественных показателей.