

БАНКОВСКОЕ ДЕЛО

М. А. Маракуева

канд. экон. наук, доцент факультета экономики Европейского университета в Санкт-Петербурге

СКОРИНГ В РОЗНИЧНОМ КРЕДИТОВАНИИ: РАСПРОСТРАНЕННЫЕ ОШИБКИ И ИХ СТОИМОСТЬ

Введение¹

Рыночные механизмы связаны с асимметрией информации как одним из важнейших признаков, формирующих прибыль или убыток контрагентов. Например, потенциальный заемщик, приходя за кредитом в банк, не раскрывает всей информации о себе, о своих платежах и доходах. Банк для правильного принятия решения о том, одобрить заемщику выдачу кредита или нет, вынужден достраивать недостающую информацию или покупать ее извне. От того, насколько точно банк достроит информацию, релевантную для принятия решения, зависит финансовый результат: размер будущей выручки, размер убытков от невозврата кредита и сумма резервирования средств под плохие долги.

Скоринговая оценка представляет собой вероятность наступления моделируемого события для конкретного наблюдения из генеральной совокупности. Скоринговая модель — это математическая модель, с помощью которой рассчитывается вероятность события на основании набора факторов. Для ситуации выдачи кредита основным моделируемым событием является невозврат кредита, наблюдением — отдельный заемщик, генеральной совокупностью — все потенциальные заемщики в экономике, а факторами — те характеристики заемщика, сложившиеся под влиянием экономической конъюнктуры рынка и предыдущих действий заемщика, которые влияют на вероятность невозврата кредита. Уловить сигналы недобросовестности заемщика, сконцентрироваться на важнейших факторах, отбросив «шумы», — основная задача аналитика, разрабатывающего скоринговые модели.

Скоринговые модели физических лиц широко используются в банковских и небанковских финансовых институтах, а также в бюро кредитных историй (БКИ), государственных организациях, фирмах, формирующих цифровые профили и портреты физических лиц. Цифровизация и задачи моментального принятия кредитного решения приводят к тому, что все больше кредитов одобряются автоматически, без участия человека, так что скоринг лежит в основе системы автоматического принятия кредитных решений.

В данной статье разбираются опыт банковского кредитования физических лиц и некоторые скрытые проблемы, с которыми сталкиваются банковские математики в процессе построения и внедрения скоринговых моделей. Многолетняя практика руководства подразделениями, разрабатывающими скоринговые модели, убедила нас в необходимости еще раз напомнить о важных этапах построения моделей и разобрать типичные промахи в процессе разработки скорингов в банках. Разбор примеров на основе реальной банковской выборки заемщиков позволяет продемонстрировать в статье существенность обычно опускаемых при моделировании

¹ Автор благодарит за помощь в подготовке расчетов к статье аналитика данных Кирилла Руслановича Артемова.

процедур, оценить уровень материальности для прибыли банка неточностей в скорингах, стоимость неверно принятых в результате управленческих решений. Статья может быть полезна как для сотрудников кредитных организаций, занимающихся построением моделей, так и руководителей среднего и высшего звена.

Методика данной работы следующая. На основании реальных данных банков сформирована некая учебная упрощенная выборка для демонстрации обсуждаемых эффектов. Упрощение касалось лишь снижения размерности: уменьшено количество наблюдений до 6 тысяч, выборка оперирует только 10 готовыми для скоринга переменными, 3 из которых представляют собой демографические данные. На основе выборки разработана «эталонная» модель логистической регрессии, которая сравнивается в статье с ситуацией реализации той или иной обсуждаемой ошибки или неэффективности в построении модели. В результате сравнения демонстрируется размер эффекта от ошибки в модели.

Далее оценивается «стоимость» этой неточности для банка, по упрощенной схеме, приведенной, например, в книге (Thomas, Edelman, Crook, 2002, p. 109). В статье мы исходим из предположения, что на «хорошем» клиенте банк может заработать 5% от выданного кредита, потери от «плохого» клиента составляют 80% выданного кредита, в случае дефолта после одобрения кредита. Если принять, что средняя сумма кредита составляет 100 тыс. руб., можно рассчитать сумму потерь при установлении определенного уровня отсека по скорингу. Ошибки скоринговых моделей в сравнении с «безошибочной» моделью порождают дополнительные потери банка, которые составят сумму потерь от ошибок 1-го и 2-го рода: одобренных плохих клиентов и недополученной прибыли от отказанных заявок хороших клиентов. Доход же банк получает в этой упрощенной схеме от одобренных «хороших» клиентов. В стандартных расчетах в данной статье ошибки 2-го рода учитываться не будут, за исключением нескольких примеров, поскольку в реальности банки не наблюдают дефолты в отказных заявках. Некую аппроксимацию уровня ошибок 2-го рода можно сделать, проследив альтернативное кредитование отказного клиента в других банках по БКИ. Но этот эксперимент ограничен сроком действия согласия заемщика на проверку в БКИ. Второй вариант — организовать временное одобрение части отказных заявок по скорингу (*score override*), но этот довольно дорогой эксперимент может себе позволить не каждый банк. Мы ограничимся подсчетом прибыли с учетом потерь от одобренных заявок в тех случаях, где модельные ошибки 2-го рода не являются основной причиной снижения прибыли. Чтобы показать «стоимость» конкретной обсуждаемой в статье ошибки скоринга, достаточно сравнить доход и потери от «эталонной» модели и этой же модели, но с внедренной ошибкой, при одном уровне отсека (*cut-off*) заявок по скорингу.

Для формирования учебной выборки искусственно симулировались лишь незначительные ее параметры, никакой подгонки под конкретные описываемые примеры ошибок не осуществлялось. Таким образом, в статье показано, что даже если на редуцированной по количеству наблюдений и составу переменных выборке столь существенны различия в эффективности моделей при реализации описанных ошибок построения, то в реальных процессах кредитования эффект от ошибок будет значительно более ощутимый.

Особенности применения скоринга

С математической точки зрения скоринг — это задача по формированию новой величины, так называемой «скоринговой оценки», на основе разнородных массивов данных о действиях потенциального потребителя. Цель решения задачи:

оценить потенциальное будущее поведение потребителя на основе исторической информации. Самый простой и известный пример применения скоринга — решение банка о выдаче кредита. Скоринг применяется не только в выдачах, но и в управлении портфелем, в работе по сбору просроченной задолженности, формировании новых кредитных и некредитных предложений клиентам банка и прочих направлениях деятельности банка.

Важно отметить, что скоринговая модель должна быть доступна для понимания как со стороны менеджмента банка (который утверждает ее использование для принятия решений), так и для регулятора. ЦБ РФ с 2004 г. объявил о приверженности базельским принципам, которые регулируют размер рисков, принимаемых на себя кредитными организациями. В соответствии с базельскими принципами, ЦБ РФ выдвигает конкретные требования к скоринговым моделям банков и имеет полномочия по проверке качества моделей.

Методами построения скоринговых моделей традиционно являются логистическая регрессия, деревья решений, нейронные сети. Последние 7–5 лет начали применяться более продвинутое и сложные модели. Однако, на наш взгляд, умение работать именно с классическими моделями позволяет не тиражировать ошибки и в более сложные модельные вычисления. Поэтому в статье мы приводим типичные проблемы в построении скоринга на примере логистической бинарной регрессии.

Неэффективность работы со скоринговой моделью может привести как в прямым потерям банка от кредитования неблагонадежных заемщиков, так и к упущенной прибыли, если модель ошибочно будет отказывать потенциально доходным для банка клиентам, клиентская лояльность к банку снижается, клиенты просто уйдут в конкурирующий банк. Неверная оценка риска заемщика приводит также к некорректному расчету суммы и ставки процента по выдаваемому кредиту и неверному начислению резервов на кредит в случае применения продвинутого подхода Базель в оценке кредитных рисков. В результате модельная неточность порождает мультипликативный эффект на прибыль банка.

Разберем отдельные виды ошибок, которые присутствуют в моделях неявно и не могут быть обнаружены как самим аналитиком, так и руководством банка без дополнительного исследования.

Мы не будем останавливаться на необходимости проверки основных гипотез перед применением логистической регрессии, таких как нормальность распределения количественных переменных, независимость наблюдений, независимость остатков, отсутствие мультиколлинеарности. Последствия пренебрежения соблюдением данных предпосылок достаточно показаны в других работах (например, Сорокин, 2014, с. 3). Хотелось бы только упомянуть о предпосылке независимости наблюдений для построения логистической регрессии, поскольку часто в базе данных для скоринга оказываются несколько записей одного клиента о выдаче кредитов на интервале наблюдений. Достаточно напомнить о необходимости агрегировать данные по клиенту в этом случае.

Сравнение моделей в случае исключения или включения переменных (вложенных моделей) некорректно проводить только по уровню значимости модели и коэффициентов при факторах. В учебнике Магнуса, Катешева, Пересецкого (Магнус, Катышев, Пересецкий, 2004, с. 130) показано, что гипотеза нулевого коэффициента в этом случае не работает. Аналогично некорректно сравнивать показатели эффективности модели (R^2 для линейной регрессии, $GINI$ для логистической), нужно оценивать значимость различий. В учебнике предложена метрика оценки среднеквадратического отклонения для сравнения двух вложенных моделей,

для того чтобы оценить необходимость включения в модель взаимодействия факторов. Мы оперируем сравнением эффективности в терминах коэффициента *GINI*, однако наряду с этим показываем разницу в распределении по скорингу и результат в виде изменения прибыли банка от принятия кредитных решений на основе скоринговых моделей.

Следует сказать о параметрах «эталонной» модели, на основе которой в статье демонстрируется стоимость той или иной ошибки. Логистическая регрессия построена на 6 тыс. наблюдений учебной выборки. Выборку составляют кредитные заявки, обогащенные данными о кредитном поведении потенциального заемщика до момента заявки и его демографическими характеристиками (рис. 1). Дефолт — событие выхода в просрочку свыше 90 дней в течение 12 месяцев после выдачи кредита. Эталонная модель имеет следующие параметры: мера разделяющей способности модели — коэффициент *GINI*, равен 65% на построении, 66% на валидации *out-of-sample*, что является хорошим показателем для поведенческих скоринговых моделей и свидетельством устойчивости модели. Уровень отсекающей установлен на 6%-ной вероятности дефолта. Для оценки «стоимости» ошибок, порождаемых различными сдвигами в скоринге, мы оставляем тот же скорбалл в качестве уровня отсекающей, что и в «эталонной» модели, как это происходит на боевых процессах.

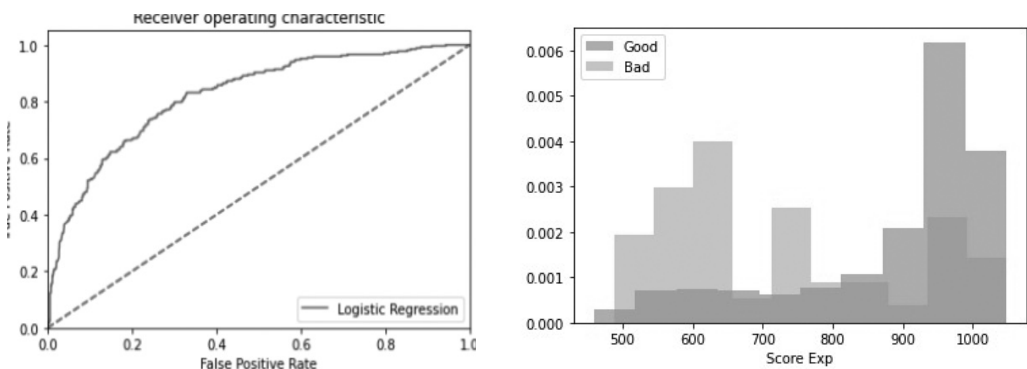


Рис. 1. Качество эталонной модели: ROC-кривая и разделяющая способность по скорингу: частота «хороших» и «плохих» заемщиков в распределении по скорбаллу

Перейдем к описанию тех ошибок, которые считаем актуальными и критическими в текущей практике построения скоринговых моделей в банках.

Ошибка 1. Ограничение построения скоринговой модели только теми переменными, которые имеются в выгрузке для создания модели

Практически все российские банки за последнее десятилетие осознали, что хорошие скоринговые карты являются крупнейшим конкурентным преимуществом при кредитовании физических лиц, наряду с накопленной клиентской базой. А в борьбе за качество скоринговых моделей доступность максимально полной информации о клиенте и возможность ее использования в моделях так же важны, как и собственно качество модели.

Мы стояли у истоков применения внешних для банковского сектора данных в скоринге и кредитных процедурах и на опыте убедились в существенном их положительном эффекте на качество моделей. Но и внутренние ресурсы в банках также далеко не исчерпаны. В настоящее время практически не используется в скоринговых моделях информация о поведении клиента при выходе в просрочку

(данные подразделений *collection*) или параметров лояльности, разрабатываемых в *CRM* подразделениях, тех же откликов на звонки или классификации заемщиков по поведенческим признакам в портфеле. Транзакционный скоринг, учитывающий информацию по карточным транзакциям клиента, также редок вследствие сложности обработки большого массива данных транзакций по картам. Мы на опыте знаем, как ценны эти переменные и насколько они улучшают качество сегментирования клиента по уровню риска. Данные о депозитах, например, увеличивают эффективность моделей, измеряемую после внедрения, более, чем на 15%, учет карточных транзакций — почти вдвое.

Ограничение набора данных ведет к тому, что поведение клиента описывается не полностью, а модель теряет свою разделяющую способность, увеличивается вероятность ошибок 1-го и 2-го рода. Вероятность этой ситуации выше в случае, когда скорингист не имеет доступа к исходным банковским хранилищам данных, а пользуется витриной для скоринга с фиксированным набором переменных или заказывает выгрузки для построения моделей у ИТ-подразделения.

Однако даже при ограниченном наборе характеристик заемщика во власти аналитика проявить инициативу и изобретательность и сгенерировать дополнительные факторы, которые дадут возможность описать поведение клиента максимально полно. Речь идет о детальном описании поведения клиента и его платежей. За счет формирования новых переменных из небольшого массива данных можно значительно увеличить объем информации о клиенте, имеющейся в распоряжении аналитика. Все скорингисты знают, например, как эффективна переменная «длина кредитной истории клиента в банке или в целом на рынке». Это искусственная переменная, рассчитывается на основании даты выдачи первого кредита. Чем дольше клиент демонстрирует свою кредитную историю, тем надежнее предсказание его поведения и устойчивее имеющиеся тренды в поведении. Но таких переменных может быть множество. И всегда выигрывает тот аналитик (или команда скорингистов, что мы видели при организации конкурсов на построение моделей), который потратил время на создание максимального числа таких синтетических переменных.

Еще один пример. В хранилище присутствует поле даты входа клиента в просрочку. Переменные, которые, скорее всего, будут значимы в скоринге, сформированные из этого поля, могут выглядеть следующим образом: частота просрочек, длительность просрочек, интервалы между просрочками (мин, макс и пр.), недавний минимальный интервал между просрочками, недавний максимальный интервал между просрочками, последовательности просрочек, частота коротких и частота длинных просрочек, давность и количество входа в просрочку 30+, 60+ и пр., *FPD*, *SPD*, *TPD* и многие другие.

Чтобы продемонстрировать выигрыш модели от подобных синтетических переменных, а вернее, потери от их отсутствия, обратимся к учебной выборке. В нашем учебном примере фактор «длина последней просрочки» — эффективная переменная в модели. Заменим ее на довольно близкую, но более простую переменную — «максимальная длина просрочки», рассчитанную за всю кредитную историю клиента. Снижение эффективности модели, измеренным коэффициентом *GINI*, из-за упрощения переменной составляет 9 п.п., до 56%. Графически снижение эффективности видно на результирующей *ROC*-кривой модели или кривой ошибок, как ее называют (рис. 2). В деньгах снижение эффективности модели при включении более примитивной переменной составляет 34% от прибыли или около 1 млн руб. на выборке из 6 тыс. входящих анкет. В месячном выражении для среднего банка потеря составит не менее 60 млн руб.

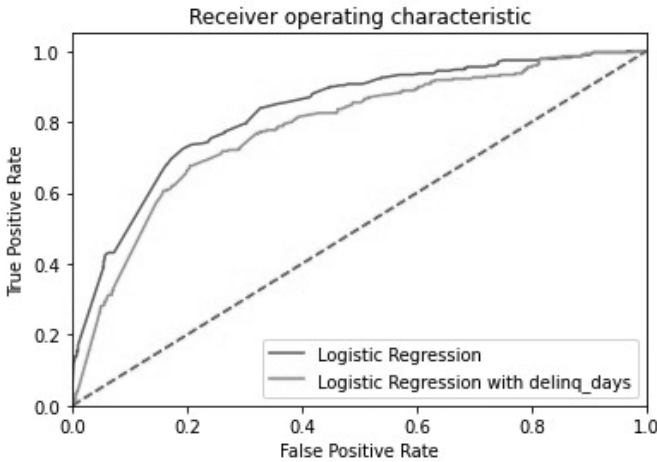


Рис 2. Смещение ROC-кривой при применении упрощенной переменной в скоринге (*delinq_days*)

Ошибка 2. При отборе факторов не учтены межфакторные взаимодействия

Дилемма выбора переменных, стоящая перед скорингистом, не имеет однозначного решения. Состав переменных определяется, конечно, в первую очередь имеющимися данными. Но вот как выбрать те факторы, которые максимально опишут процесс формирования зависимой переменной? При отсеивании значимых факторов мы получим смещенную оценку, при включении незначимых — неэффективную.

При выборе набора факторов на первом месте стоит логика, так как исходя из нее на предыдущих этапах подготовки данных рассчитываются синтетические переменные, являющиеся характерными признаками зависимой переменной. И этот этап крайне важен для интерпретируемости модели, для ее представления руководству банка и защиты перед бизнесом, а также для дальнейшего мониторинга работы модели. Если логика включения переменной не ясна и коэффициент при ней неинтерпретируем, то не будет предсказуемой и работа модели при малейшем изменении кредитных процедур или сдвиге входящей популяции.

Выбор факторов логистической регрессии стандартно осуществляется путем отсеивания переменных с незначимыми коэффициентами. Его можно разделить на три этапа: одномерный анализ — когда оценивается взаимосвязь каждого из предполагаемых факторов модели с зависимой переменной, двухмерный — по парам переменных и многомерный — когда выбирается финальный набор факторов. На этапе многомерного анализа факторов применяется как метод последовательных включений, так и метод последовательных исключений. Метод последовательных исключений переменных или метод последовательных включений содержит логическое ограничение — оптимальный набор факторов может не быть достигнут из-за ошибки на промежуточных этапах. Часто мы видим, что исключенная ранее переменная может быть значима в финальной модели. Причин может быть несколько. Методы последовательных включений и исключений зависят от первоначального набора переменных. Также возможна корреляция переменных, в т.ч. и скрытая, и тогда оценки коэффициентов и их значимости смещаются. Причина может быть также в том, что факторы могут взаимодействовать при влиянии на зависимую переменную.

Анализ межфакторного взаимодействия как раз относится к этапу двумерного анализа факторов. Взаимодействие факторов — это изменение зависимости результирующей переменной от одного фактора под влиянием другого фактора. Эту нелинейность необходимо учитывать, чтобы не заложить ошибочные тренды в модель.

Если мультиколлинеарность (высокую корреляцию) факторов скорингисты привыкли проверять, понимая, что от наличия корреляции могут появляться смещения в оценках регрессии, то взаимодействие факторов — более сложный признак. Это ситуация, когда влияние на зависимую переменную одного фактора зависит от значений другого фактора. Грубо говоря, это слом тренда зависимости результирующей переменной от фактора. При этом значимые переменные становятся менее эффективными или незначимыми из-за слома тренда влияния, который не учтен в модели. В английском языке для этого случая применяется термин — *interactions*. Наиболее наглядно она демонстрируется на интервальных и номинальных переменных. Сложность обнаружения такой ситуации «на глаз» или из общей логики приводит к необходимости дополнительных тестов на этапе изучения исходных данных перед построением модели.

К сожалению, даже в ТОП-10 розничных банков (занимающихся кредитованием физических лиц) аналитики не всегда включают этап поиска межфакторных взаимодействий в процесс построения скорингов. Чаще используются ранее найденные зависимости, ставшие общепринятыми. Например, переменная «пол — возраст», внедренная в российскую практику моделирования консультантами McKinsey. Разная склонность к риску (зависимая переменная) для женщин и мужчин разных возрастов. Мужчины снижают склонность к риску, а женщины демонстрируют нелинейную зависимость от возраста или более пологий тренд. Есть масса других примеров. В то же время в зарубежных банках (например, в GE-Money bank) уже с начала 2000-х гг. стоят отдельные ИТ-модули для автоматического поиска взаимодействий факторов при построении скоринга.

Поскольку преобразование количественных переменных для логистической регрессии в интервальные (*binning*) стало общей практикой, мы не будем останавливаться на этапах сначала первичного разбиения (*fine classing*), затем его финализации (*coarse classing*), а применим результаты последней стадии для поиска взаимодействий факторов между собой (иногда их называют взаимодействием второго порядка). Для поиска взаимодействий проще всего применять таблицы сопряженности факторов и рассчитывать значимость оценки хи-квадрат, демонстрирующую значимость асимметрии распределения наблюдений в таблице.

Учет взаимодействия факторов возможен несколькими способами. В случае, если межфакторное взаимодействие очень сильное или, если одна переменная меняет поведение сразу нескольких факторов в модели, то логично разбивать данные на подвыборки по этому фактору и строить отдельные модели. Часто так делают с региональной разбивкой, когда в разных регионах на риск заемщика влияют разные факторы и зависимости отличаются. В случае, если данных недостаточно или межфакторное взаимодействие локальное, оно может быть учтено с помощью переменных взаимодействия или созданием новых переменных совместного распределения факторов. Традиционный пример переменной «пол — возраст» является переменной взаимодействия. При этом, конечно, отдельные переменные пола и возраста исключаются из модели во избежание дублирования информации и смещения оценок модели.

Возможны более сложные варианты переменных, когда создается синтетическая переменная. Например, частота и глубина просрочки. Интуитивно понятно, что

частые просрочки не могут быть длинными по сроку. Такого рода неравномерность в зависимостях имеет дополнительную выгоду. Совместное распределение взаимодействующих факторов несет в себе больше информации, чем каждый фактор в отдельности. Одновременно исключается дублирование информации в двух факторах: интервал переменной частоты просрочек (бакет) с частыми просрочками содержит практически те же наблюдения, что бакет переменной «длина просрочки» с короткими интервалами неплатежей. Включение обоих бакетов увеличивает в два раза значимость ситуации коротких просрочек в модели, что некорректно. Аналогично для группы заемщиков без просрочки, которые составляют отдельный бакет в одной и в другой переменной одновременно.

Совместная сложная переменная, отражающая взаимодействие факторов, будет включать отдельно: ситуации коротких частых просрочек («технические» просрочки), ситуации редких коротких неплатежей, ситуации редких длинных просрочек, пограничные ситуации, когда просрочка не является «технической», но и не несет в себе признаки дефолта. Причем ситуации редких длинных просрочек также можно разбить на случаи: просрочки были недавно или были давно. Такие комбинированные логические переменные очень помогают адекватно и полно описать поведение клиентов. Они несколько напоминают деревья решений, с помощью которых, собственно, и можно выявлять специфические «ветви» для включения в комбинированные переменные.

Реализация поиска и включения взаимодействия переменных в российских банках лежит полностью на усмотрении аналитика. Однако давно существует практика стандартизации этого этапа построения модели. Необходимо создать процедуру, которая проверяет все межфакторные взаимодействия. Во многие статистические пакеты этот этап уже включен, его просто надо использовать. При работе с логистической регрессией анализ взаимодействий факторов осуществляют после этапа *coarse classing*, поскольку получившиеся в результате интервальные переменные значительно сокращают время поиска взаимодействий по сравнению с непрерывными. При этом точность не теряется — исходные факторы все равно были бы включены в регрессию с теми же интервалами (бакетами) значений. Результирующий вид сложных переменных взаимодействия формируется, конечно, вручную, исходя из экономической логики. Практика показывает, что затраты на создание сложных переменных с лихвой окупаются ростом эффективности модели.

В качестве примера, в учебной выборке было найдено взаимодействие переменных длины кредитной истории и глубины последней просрочки. Построена новая переменная, учитывающая взаимодействие факторов (рис. 3). Она выделяет совершенно новые тренды изменения уровня дефолтов (*bad rate*) в зависимости от значения переменной «длина кредитной истории заемщика» (длина КИ). Новые тренды образуют группы наблюдений, разнящиеся по уровню последней просрочки. Для заемщиков без просрочек в истории тренд близок к общему — риск снижается с ростом длины кредитной истории. Но для заемщиков с просрочкой свыше 100 дней (дефолт) это снижение значительно более резкое. Находится группа с низким риском даже среди заемщиков, уходивших в прошлом в дефолт, — это заемщики с долгой кредитной историей (у нас нет переменной — давность просрочки, пересечение с ней дало бы возможность более точно описать зависимость). Риск же тех, кто имеет просрочку до 100 дней в истории, практически не зависит от длины кредитной истории.

Мы нашли взаимодействие переменных, характерное для конкретной выборки. Оно не универсально для всех выборок, кредитных продуктов и регионов. Для каждой модели аналитик конструирует специфические взаимодействия, интерпретируя их.

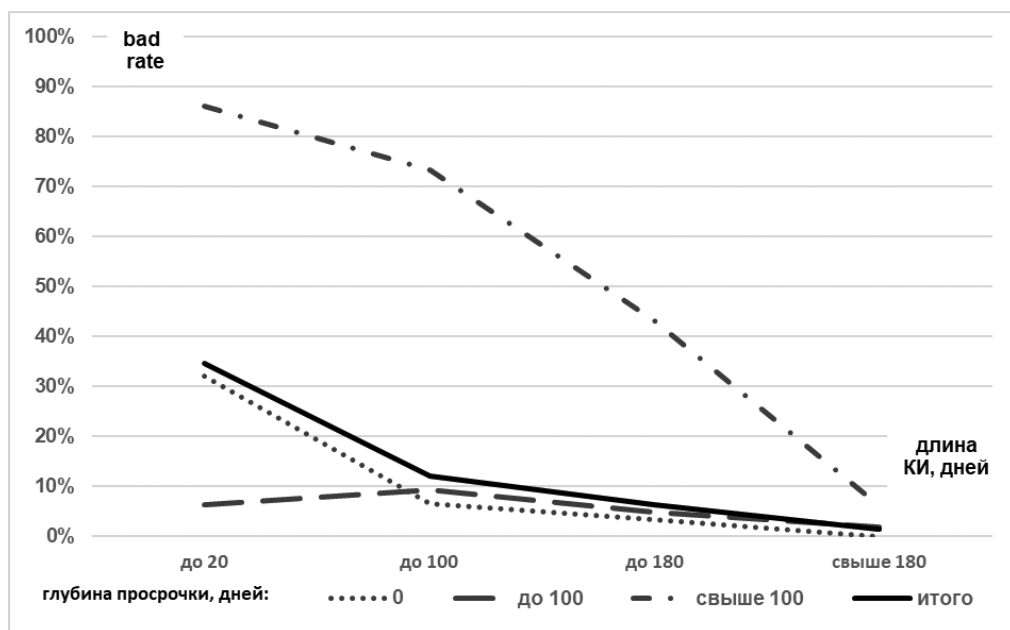


Рис. 3. Зависимость уровня риска (*bad rate*) от длины кредитной истории при разном уровне последней просрочки заемщика

Если бы аналитик не включил взаимодействие факторов в модель, а включил бы оба фактора по отдельности, то даже на учебном примере эффективность модели уменьшилась бы до 64% *GINI*, снижение эффективности составило бы 19 п.п. Стоимость данной ошибки в денежном выражении — 28% от прибыли при одобрениях по скоркарте или 800 тыс. руб. (на выборке из 6 тыс. наблюдений). Для среднего банка это означает около 48 млн руб. недополученной прибыли в месяц.

Ошибка 3. Соотнесение скоринга и уровня текущего риска — калибровка

Скоринг применяется в банках не только для принятия решения — давать кредит или нет на основании уровня отсечения (так называемый «заявочный скоринг»). Широко применяется также портфельный скоринг — скорбалл, присвоенный клиенту, имеющему кредит в текущем кредитном портфеле банка. Он используется для расчета резервов по кредитному портфелю и прогнозирования динамики портфеля, бюджетирования и других задач управления портфелем. Оба скоринга могут применяться для установления лимитов кредитования и процентных ставок, но каждый в своей зоне — выдачи кредитов или управления портфелем (Конева, 2021, с. 54; Маракуева, 2020, с. 85). Все эти варианты применения скоринга связаны не только с уровнем отсечения (*cut-off*), они основаны на полной шкале скорбалла. Тогда для экономического эффекта использования скоринга становится крайне важной верная калибровка всей шкалы скоринга.

Калибровка — соотнесение скорингового балла, посчитанного по модели, с актуальной вероятностью дефолта. Калибровка является ключевым моментом для корректного внедрения скоринговой модели, от нее зависит прибыль после внедрения скоринга не меньше, чем от качества модели. Мы не останавливаемся на начальном этапе калибровки, когда соотносятся скоринг и модельный риск.

С каким уровнем дефолта соотносить скорбалл для внедрения скоркарты в процессы банка? Самая распространенная ошибка — применять распределение

по дефолтам, которое характерно для обучающей выборки (модельное). Практически всегда входящая популяция заемщиков для банка характеризуется сдвигом уровня риска и сдвигом распределения его по профилям заемщика со временем. Макроэкономическая ситуация не стоит на месте, разрабатываются новые банковские продукты, формируются новые таргет-группы заемщиков — довольно много факторов, способствующих сдвигу. Одновременно та часть входящей популяции заемщиков, которая доходит до скоринга, подвергается дополнительному смещению в части уровня дефолта из-за постоянной донастройки кредитных процедур, ввода новых проверок и отсечений (отказов) до скоринга, изменения продуктовых политик и пр.

Распространенные практики, в т.ч. описываемые обычно в статьях по калибровке скоринга, — это выбор некоего периода, похожего по параметрам макроэкономики и кредитным процедурам, или аналитическое преобразование выборки, как предложено в учебнике «Credit scoring and its implementation» (Thomas, Edelman, Crook, 2002, p. 148). Еще один вариант — цензурирование выборки, т.е. использование более свежих наблюдений, чем позволяет интервал наблюдения зависимой переменной (обычно 12 мес.). Возможно также симулирование тестовой выборки под текущие параметры входящей на скоринг популяции¹. Скорингисты сами определяют, каким методом пользоваться.

В идеале скоринговые команды банка совместно с финансовыми подразделениями должны постоянно тестировать качество калибровки риска по скорингу и осуществлять ее актуализацию. При этом в качестве актуального уровня риска имеет смысл брать не сложившийся на сегодня, а прогнозный уровень риска на период работы скоринговой модели. Следует проводить прогнозирование риска популяции на будущее в соответствии с планами продаж, риск-аппетитом и планируемыми изменениями в кредитных процедурах. Для этого в банке внедряется тестовая среда системы принятия кредитных решений, на которой постоянно просчитывается риск текущей и вызревшей популяции на актуальных кредитных процедурах и с учетом их планируемых изменений.

Существует ряд методов, позволяющих оценить качество калибровки: показатель Бриера, показатель надежности, проверка значимости отклонений частоты дефолтов от ожидаемых: биномиальный калибровочный тест, тестирование на основе корреляций дефолтов, проверочная калибровка в соответствии со стандартным нормальным распределением.

В нашем примере смоделирован сдвиг популяции, как показано на рис. 4: предполагаем, что экономическая ситуация улучшилась, и половина тех, кто ранее демонстрировал дефолт через год после выдачи, удержалась в рамках небольшой просрочки. Полученное снижение риска входящей популяции (на рис. 4 — смещенная выборка) ведет к несоответствию модельного уровня риска и фактического для новой популяции, особенно в группах с прогнозируемым высоким риском.

Если бы при данном сдвиге модель продолжала бы работать при старой калибровке и уровне отсечения, прибыль банка с учетом потерь по риску снизилась бы на 43% относительно «эталонной» модели, потери в прибыли составили бы 75 млн руб. в месяц. Если же аналитик заметил сдвиг популяции и осуществил перекалибровку модели в соответствии с новым уровнем риска, банк бы снизил потери до 23%, т.е. до 40,5 млн руб. Конечно, при снижении прибыли банка в два раза в результате сдвига популяции модель требует перестроения, а не только

¹ Методы калибровки собраны, например, в статье (Широбокова, 2019, с. 406). Методы изменения исходной выборки под известный уровень риска (перезвешивание и прореживание выборки) описаны в статье (Сорокин, 2014, с. 8).

перекалибровки. Рекомендации по своевременной актуализации модели — постоянный мониторинг показателя *population stability index (PSI)*. Если он показывает превышение порога сдвига в 25% в течение нескольких периодов, то модель требует перестроения (Thomas, Edelman, Crook, 2002, p. 155).



Рис. 4. Сдвиг популяции и необходимость калибровки модели по риску

Ошибка 4. Пересечение кредитных процедур с факторами модели — инверсия в факторах из-за кредитных процедур до скоринга

Оторванность моделирующих подразделений от ИТ-решений, поддерживающих кредитные процедуры, играет злую шутку с крупными банками. Неудачная практика — отдавать модель на внедрение другим подразделением без детального контроля внедрения разработчиком модели и без тестирования ее работы в боевом режиме. На опыте встречались довольно трагичные случаи, когда некорректное внедрение модели приводило к плохой работе отлично построенных скоркарт. Обычно причина кроется в некорректной подаче переменных на вход, неверной трактовке результата или встраивании скоринга в узкий процесс вместо планируемого широкого, или наоборот. Возможно, что в процессе работы модели и критического изменения кредитных процедур или продуктовой политики (речь о кредитных продуктах) произойдет следующий за этим существенный сдвиг подающейся на скоринг популяции. Поэтому тестовая среда крайне необходима, в том числе, для контроля внедрения и работы скоринга. Тестирование изменений кредитных процедур нужно осуществлять с контролем качества работы скоринга на новых стратегиях.

Но бывает ситуация, когда аналитик, не видящий процедур и не участвующий в их изменении, осуществляет построение модели без учета уже свершившихся сдвигов в процедурах. Ошибка скорингиста в этом случае обусловлена функциональным разделением обязанностей, но это не снимает с него ответственности.

Другой случай, с которым мы сталкивались, — это модели, построенные внешними консультантами по заказу. Часто для улучшения качества модели консультанты неправомерно включают в модели дополнительное число «плохих» наблюдений. Например, включают в выборку для построения модели клиентов на текущей просрочке в банке или с большими историческими просрочками, т. е.

заранее отказные случаи. Это искусственно «растаскивает» выборку на плохих и добросовестных заемщиков. Конечно, при внедрении такие модели сразу сильно теряют в эффективности относительно построения, так как заявки со столь высоким риском не доходят до этапа скоринга.

Рассмотрим пример. До скоринга организовалась дополнительная проверка в кредитной процедуре, например, переменная из скоринга встала в качестве параметра отказа. Пусть это будет, например, историческая просрочка по данным БКИ, и при просрочке свыше 100 дней клиент получает отказ. В этом случае около 70–80% «плохих» по скорингу клиентов не доходит до скоринга. Эффективность модели при этом может упасть вдвое, как показывает опыт. Требуются перестроение модели на новых данных и срочная замена действующей модели.

На учебном примере мы видим ухудшение качества скоринга в случае отказов до скоринга по критерию максимальной исторической просрочки клиента с 65–66% до 57–60%, что для работающей модели очень существенно. До скоринга отсеялись 85% заемщиков с будущим дефолтом, что дало сильное смещение оценки вероятности дефолта на выходе модели. Потери в прибыли здесь в основном связаны с недополученной прибылью от отказанных заявок добросовестных заемщиков, так как при значительном улучшении входящей на скоринг популяции уровень отсека остался прежним. В примере мы отсекали 19% излишне хороших клиентов, недополучив 4,5 млн прибыли или в пересчете на размер бизнеса среднего банка — более 150 млн руб. Одновременно, в модели оказался смещенным портрет «плохого» заемщика. Благодаря смещению мы видим потери на учебной выборке как дополнительные 2% от эффективности. Ухудшение эффективности модели показано на рис. 5.

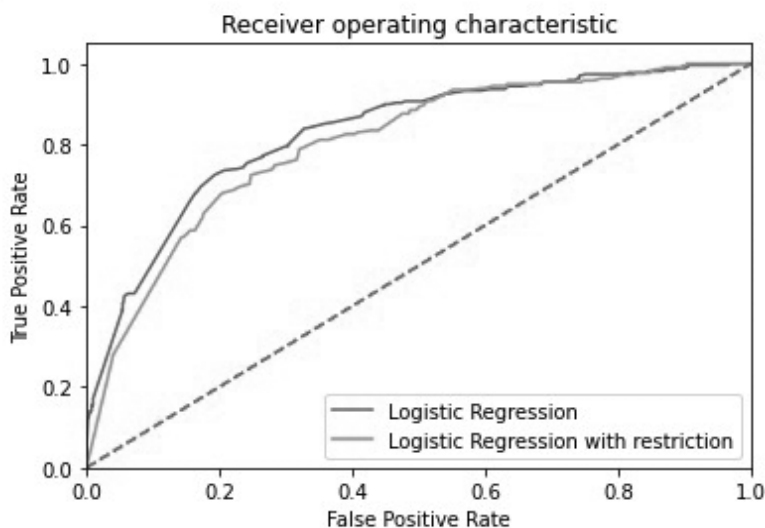


Рис. 5. ROC-кривая: потеря в эффективности модели при пересечении ее с проверками до скоринга (with restriction)

Мы рекомендуем всегда соотносить внедряемые проверки со скорингом и включенными в него факторами. Все изменения кредитных процедур, даже самые незначительные, обязательно нужно тестировать на тестовых средах систем принятия кредитных решений и при тестировании оценивать эффективность текущих моделей скоринга. Также необходимо перестраивать скоринг в случае, когда тест показал снижение его эффективности. И делать это до внедрения изменений в кредитные процедуры.

Ошибка 5. Устойчивость факторов и модели в целом

Еще одна довольно распространенная ошибка при работе скоринга: ответственные за внедрение моделей сотрудники не проконтролировали изменение в работе одного или нескольких факторов модели. Частый и наиболее болезненный для банка пример: сдвиг значений внешней информации, включенной в модель. Нам встречались на практике как ситуации осознанной перекалибровки используемых внешних скорингов и данных без предупреждения об этом пользователей, так и непровольного сдвига данных из-за технических сбоев на стороне поставщика данных.

Другие ситуации и их причины рассмотрены подробно, например, в (Thomas, Edelman, Crook, 2002, p. 162–163), где указано, что переменная может сдвигаться довольно постепенно, например, количество счетов клиента при вводе новых банковских продуктов. Это создает проблему обнаружения сдвига переменной. В период экономической нестабильности риск бедности преимущественно возрастает среди домохозяйств, обремененных кредитами (Ниворожкина, Макаренко, Полякова, 2020, с. 54). Значит, в кризисные периоды перераспределение риска по шкале значений показателя дохода клиента, если он включен в модель в виде переменных *debt to income (DTI)* и др., будет причиной нестабильной работы переменной.

Также один из наиболее ярких примеров последних нескольких лет — переменная количества заявок заемщика на кредит на коротком горизонте времени. Ранее эта переменная указывала на наличие недавних отказов в других банках и МФО, нежелательном хождении заемщика по банкам. С переходом на массовые онлайн заявки, расширением работы брокеров веерная рассылка заявок заемщиком или посредником стала рутинным случаем, не несущим в себе негативной информации о заемщике. При этом далеко не все банки обновили свои скоркарты. Устаревшая переменная, ранее значимая, продолжает генерировать отказы хорошим заемщикам.

Постоянный мониторинг и трекинг работы отдельных переменных, а не только всей скоркарты — решение данной проблемы. Все описанные случаи сдвига нам удалось вовремя выявить именно потому, что стояли сигналы на существенный сдвиг переменных относительно параметров при построении модели. Оценка чувствительности скоринговой модели к изменению распределения одного из факторов — также хорошая практика, необходимая для определения порогов изменения для срабатывания сигналов.

Покажем на цифровом примере из нашей практики, какова стоимость данной ошибки в рублевом эквиваленте. Допустим, из анкеты на кредит был исключен вопрос «число иждивенцев». Это, собственно, исторический факт, так как за последние несколько лет кредитные анкеты всех банков претерпели существенные сокращения. Если команда математиков банка не отследила данное изменение и не была информирована, то скоркарта продолжает работать с пустыми значениями поля. В нашем учебном примере этот случай повлек за собой незначительное снижение эффективности карты до значений коэффициента *GINI* 62–64%, так как эта переменная, хотя и значимая, но не самая эффективная в модели. На прибыли это отразилось в виде потери 12% прибыли, т. е. около 21 млн руб. в месяц для среднего банка.

Калибровка скоринга по вероятности дефолта, как мы выяснили, принципиально важна для корректной работы скоркарты. Но что делать, если все модели корректно откалиброваны, но не соотносятся между собой? Скажем, один и тот

же клиент по модели скоринга для выдачи кредитных карт имеет вероятность дефолта 45% и получит отказ, по модели для выдачи кредита наличными 7% и, скорее всего, получит одобрение, а по портфельному скорингу, используемому для мониторинга качества портфеля, — всего лишь 3%? Эта ситуация, которой озадачен каждый скорингист — соотнесение результатов различных моделей и совместное использование скорингов. Кредитные карты — высокорисковый продукт, и даже хороший клиент автоматически получает завышенную оценку риска при прогнозе выдачи ему кредитной карты. Таких примеров немало.

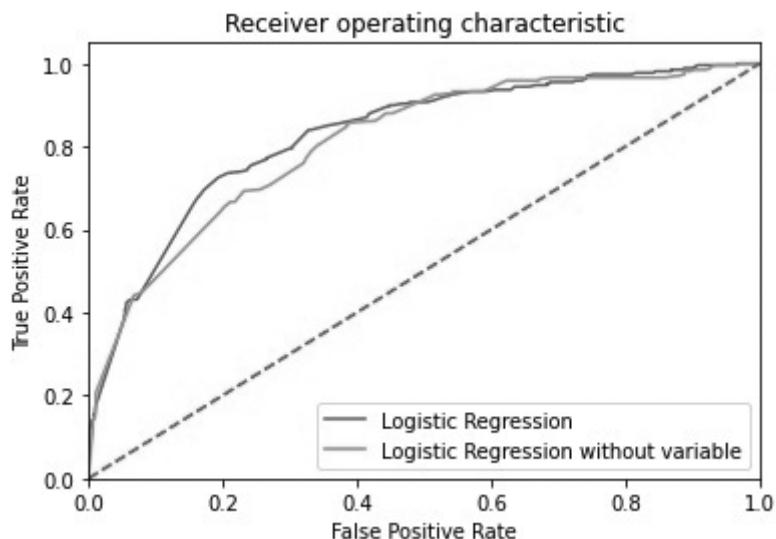


Рис. 6. ROC-кривая: потеря в эффективности модели при смещении фактора (*without variable*)

Ошибка 6. Соотнесение скоринга по разным кредитным продуктам при подборе продукта: риск продукта или риск клиента

Решением этой задачи занимаются как подразделения риск-менеджмента, так и внешние консультанты. Рекомендации типа создания матриц пересечения моделей, в ячейках которых рассчитана «реальная» вероятность дефолта, или применение портфельных моделей в процессе выдач, это в большей части временные «заплатки».

Как показала практика, частью решения этого вопроса может быть учет продуктовой истории клиента во всех скоркартах. Такие переменные, как продуктовый набор закрытых кредитов, продуктовый набор действующих кредитов, заявочная продуктовая история, помогают приблизить распределение по риску заемщиков в разных скоркартах. Эти переменные дают возможность разнести продуктовый риск и риск, присущий самому клиенту (*inherent risk*). Переменная берет на себя именно объяснение продуктовых рисков. Можно усложнить решение — учитывать продуктовую переменную совместно с переменными просрочек. Таким образом, можно найти и включить в модели закономерности продуктового поведения клиентов. Кто-то из клиентов банка постоянно просчитывается в длине *grace period* и сроках платежей по карте, а кто-то из заемщиков, аккуратно используя карту, не справляется с крупным кредитом наличными. Отчасти поэтому обычно эффективность транзакционных данных по картам для оценки уровня кредитного риска очень высокая.

Второе решение — пытаться все-таки соотносить калибровки риска по разным продуктам. И необходимость такого решения диктуют CRM активности, в рамках которых надо сделать в качестве предложения клиенту выбор того продукта, который максимально прибылен для банка с точки зрения поведения клиента. Если не говорить о построении специализированной модели вероятной прибыли клиента с учетом рисков, сложность информационного насыщения которой сильно ограничивает ее применение, то решение в том, чтобы приближать калибровки продуктовых скорингов. Необходимо искать общие специфические группы клиентов, которые имеют высокие риски по всем продуктам, выделяя их в худшие бакеты всех моделей. Также можно выделять и хороших заемщиков, имеющих низкий риск независимо от продуктового набора. Портреты этих клиентов можно рекомендовать включать во все скоркарты банка. Это поможет избежать как разницы в продуктовых скорингах одного клиента, так и резких скачков прогноза риска клиента после одобрения ему очередного кредита.

Источники

- Белусов В.* Нелинейность и взаимодействие переменных в моделях кредитного скоринга // Риск-менеджмент в кредитной организации. 2017. № 1 (25).
- Конева Е.* «Заболели» ли скоринговые модели во время пандемии? // Банковское кредитование. 2017. № 3. URL: http://www.reglament.net/bank/credit/2021_3/get_article.htm?id=7124.
- Магнус Я. Р., Катышев П. К., Пересецкий А. А.* Эконометрика. Начальный курс. М., 2004.
- Маракуева М. А.* Практические вопросы применения скоринговых моделей при управлении кредитными рисками банка. М., 2020.
- Ниворожкина Л. И., Макаренко Е. Н., Полякова И. А.* На чем экономим? Динамика кредитного поведения российских домашних хозяйств // Финансы и бизнес. 2020. Т. 16. № 1. С. 46–58.
- Сорокин А. С.* Построение скоринговых карт с использованием модели логистической регрессии // Интернет-журнал «Науковедение». 2014. № 2 (21).
- Широбоква М. А.* Калибровка скоринговой модели с учетом цензурированных данных // Вестник Пермского университета. Экономика. 2019. Т. 14. № 3. 2019. С. 406–420.
- Hosmer D. W., Lemeshow S.* Applied Logistic Regression. New Jersey, 2000.
- Thomas L., Edelman D., Crook J.* Credit Scoring and its Applications. Philadelphia, 2002.

References

- Belousov V. Nelineinost i vzaimodejstvie peremennih v modeljah kreditnogo skoringa [Nonlinearity and Interaction in Credit Scoring Models]. *Risk-menedzhment v kreditnoy organizatsii* [Risk management in a credit institution], 2017, N 1 (25). (In Russian)
- Hosmer D. W., Lemeshow S. *Applied logistic regression*. Philadelphia, 2000.
- Koneva E. «Zaboleli» li skoringovie modeli vo vremja pandemii? [Did the scoring models get sick during the pandemic?]. *Bankovskoe kreditovanie* [Bank lending], 2017, N 3. Available at: http://www.reglament.net/bank/credit/2021_3/get_article.htm?id=7124. (In Russian)
- Magnus J. R., Katishev P. K., Peresetskii A. A. *Ekonometrika. Nachalnii kurs* [Econometrics. Initial course]. Moscow, 2004. (In Russian)
- Marakueva M. A. *Prakticheskie voprosi primeneniya skoringovih modelei pri upravlenii kreditnimi riskami banka*. Moscow, 2020. (In Russian)
- Nivorogkina L. I., Makarenko E. N., Poljakova I. A. Na chem ekonomim? Dinamika kreditnogo povedeniya rossiiskih domashnih hozjastv [What are We Saving on? Dynamics of the Credit Behavior of Russian households]. *Finansi i biznes* [Finance and Business], 2020, vol. 16, iss. 1, pp. 46–58. (In Russian)
- Shirobokova M. A. Kalibrovka skoringovoj modeli s uchetom cenzurirovannyh dannyh [Calibration of the scoring model taking into account censored data]. *Vestnik Permskogo universiteta. Ekonomika* [Bulletin of Perm University. Economy], 2019, vol. 14, N 3, pp. 406–420. (In Russian)
- Sorokin A. S. Postroenie skoringovih kart s ispolzovaniem modeli logisticheskoi regressii [Building Scoring Maps Using the Logistic Regression Model]. *Internet-gurnal «Naukovedenie»* [Naukovedenie Internet Journal], 2014, N 2(21).
- Thomas L., Edelman D., Crook J. *Credit scoring and its applications*. Philadelphia, 2002.