

## ПРИМЕНЕНИЕ СИСТЕМЫ КРЕДИТНОГО СКОРИНГА ДЛЯ ОЦЕНКИ КРЕДИТОСПОСОБНОСТИ ЗАЕМЩИКА БАНКА

### APPLICATION SYSTEM CREDIT SCORING TO ASSESS THE CREDITWORTHINESS BORROWER BANK

*Макалада кредитти алууда кардардын кайтарып берүү мүмкүнчүлүгүн баалоонун скоринг-моделдердин киргизүүнүн заманбап түрлөрү каралган.*

*Ачкыч сөздөр:* скоринг моделдер, кредит скоринги, кайтарып берүү мүмкүнчүлүгүн баалоо, кредиттөө.

*В статье рассматривается современный подход оценки платежеспособности клиента при кредитовании с использованием скоринговых моделей для оптимизации процесса принятия решения при кредитном запросе.*

*Ключевые слова:* скоринговые модели, кредитный скоринг, оценка платежеспособности, кредитование.

*The modern approach for the assessment the client's solvency for loaning is considered in the article. This approach explains the use of scoring models to optimize the decision-making process for the loan request.*

*Keywords:* scoring models, credit scoring, evaluation of solvency loans.

Исследование вопросов, связанных с оценкой и прогнозированием кредитных рисков и потерь коммерческого банка, стало особенно актуальным в период кризиса. Снижение платежеспособности заемщиков, уменьшение объемов привлеченных банками средств приводят к тому, что необходимо оценивать кредитные риски более точно с целью формирования адекватных резервов под возможные потери. Одним из важнейших направлений деятельности коммерческого банка является розничное кредитование и, как следствие, оценка рисков заемщиков – физических лиц. В связи с бурным ростом сектора потребительского кредитования в Кыргызстане в последнем десятилетии оценка кредитного риска физических лиц является особенно актуальной.

Традиционные методы оценки физических лиц экспертным путем теряют свою эффективность по мере увеличения объемов розничного кредитования. Рост предложения новых банковских услуг и кредитных продуктов требует частичной или полной автоматизации процессов оценки платежеспособности клиента и выдачи кредита с учетом правильности оценки возможных потерь и построения эффективной системы контроля и управления рисками во всех сферах деятельности банка. Все вышперечисленное заставляет банки более серьезно задуматься над вопросом применения современных методик автоматизированной оценки кредитного риска физических лиц, а именно скоринга новых клиентов.

Скоринговая система оценки платежеспособности розничного клиента – это статистическая модель, оценивающая вероятность того, что заемщик не заплатит по своим обязательствам в срок, т.е. представляет собой методику оценки заемщика. Большинство систем проверки клиента построено таким образом, что при обращении человека в банк за различными услугами, он проходит несколько этапов проверки, одним из которых является проверка платежеспособности заемщика, основываясь на социально-демографических данных и данных кредитной истории. В результате анализа переменных получают показатель, который оценивает степень кредитоспособности заемщика по

ранговой шкале: «хороший» или «плохой». Каждый банк подстраивает систему проверки и принятия решений, исходя из своих возможностей, специфики предоставляемых продуктов и определенной политике кредитования, которая может быть охарактеризована по разным признакам: качеству кредитного портфеля, сфере деятельности, территориальному принципу и другим, учитывая риски и свои финансовые обеспечения. Качество заемщика оценивается определенными баллами, отражающими степень его кредитоспособности. Таким образом банк формирует на основе этих данных определенный уровень отсека, который определяет качество обратившегося за услугой клиента и возможность дальнейшего рассмотрения его заявки.

Так, в случае положительного решения банки продолжают работать с заявкой. Если же рейтинг заемщика, рассчитанный моделями кредитного скоринга, оказывается ниже уровня отсека, банки отказываются от дальнейшего рассмотрения заявки или отправляют заявку клиента на рассмотрение кредитным специалистом и/или андеррайтером для осуществления, так называемого скоринга экспертной оценки, когда решение принимается непосредственно специалистом в данной области. Философия скоринга заключается не в поиске объяснений, почему этот человек не платит. Система вообще ничего не знает о том, вернет ли кредит данный заемщик, но в ее арсенале множество кредитных историй «прошлых» клиентов, проанализировав которые, она делает, к примеру, вывод о том, что заемщики такого возраста со сходным семейным положением и образованием имеют тенденцию возвращать или не возвращать взятую сумму. Итак, скоринг представляет собой математическую или статистическую модель и выступает классификатором всех заемщиков на различные группы риска.

В западной банковской системе, когда человек обращается за кредитом, банк располагает следующей информацией для анализа:

- анкетой, которую заполняет заемщик;
- информацией на данного заемщика из кредитного бюро, в котором хранится кредитная история взрослого населения страны;
- данными движения по счетам, если речь идет о клиенте банка.

Кредитные аналитики оперируют следующими понятиями: характеристики-признаки клиентов и градации-значения, которые принимает признак. В анкете клиента характеристиками-признаками являются вопросы - возраст, семейное положение, профессия, а градациями-значениями - ответы на эти вопросы. В упрощенном виде скоринговая модель дает взвешенную сумму определенных характеристик. В результате получают интегральный показатель (score); чем он выше, тем выше надежность клиента. Интегральный показатель каждого клиента сравнивается с неким заданным уровнем показателя. Если показатель выше этого уровня, то выдается кредит, если ниже этой линии нет.

Конкуренция на рынке потребительских кредитов заметно обострилась - банки в борьбе за свою долю рынка предлагают разнообразные продукты, снижают процентные ставки, предлагают привлекательные условия кредитования. В результате конкуренции обеспечить успех бизнеса розничного кредитования не так просто. Главный вопрос: кому давать кредит, а кому нет. Если раньше риск невозврата еще покрывался величиной процентной ставки, то сейчас ситуация иная. Вводить же жесткие ограничения для получателей кредита - значит упустить прибыль, которая могла быть при более гибких ограничениях.

Чтобы работа на рынке розничного кредитования приносила прибыль, необходима эффективная система оценки рисков, которая позволила бы заранее отсекал ненадежных заемщиков и не отказывать надежным, обоснованно определяла бы размер взноса в потребительском кредите или лимит по кредитной карте. Именно такая система создает запас прочности банку, который позволяет ему выводить на рынок привлекательные для заемщиков продукты.

Для снижения рисков при обслуживании большого количества клиентов в банках применяются автоматизированные системы принятия решения о предоставлении кредита на основе скоринговой системы, базирующейся на статистической обработке большого массива данных о клиентах.

Эффективность использования скоринговой системы зависит от качества исходной скоринговой программы, которая дает оценку заемщику с учетом всех имеющихся о нем данных, от качества настройки скоринговых карт (набора параметров и оценки их значимости), которые могут модифицироваться по мере наращивания объемов статистики по выданным кредитам. Скоринговая система становится эффективной тогда, когда имеется возможность проанализировать большой объем статистических данных. От этого зависят настройки системы и ее точность.

Профессионализм использования скоринговой модели дает возможность банкам максимально точно оценивать кредитные риски, что позволяет повысить доходность кредитных операций. Среди преимуществ предполагаемое обеспечение скоринговых систем западные банкиры указывают в первую очередь на снижение уровня невозврата кредита. Далее отмечается быстрота и беспристрастность в принятии решений, возможность эффективного управления кредитным портфелем, определение оптимального соотношения между доходностью кредитных операций и уровнем риска.

Одним из основных этапов построения модели является выбор и анализ независимых переменных. Основным источником данных являются анкетные данные клиента на момент подачи кредитной заявки, например:

- Демографические показатели: возраст, пол, национальность, место проживание, длительность проживания в актуальном месте жительства, образование, профессия, длительность трудоустройства, наличие собственности, семейное положение, наличие детей и др.

- Данные по запрашиваемому кредиту: цель кредита, общий размер займа, срок финансирования, первоначальный взнос, отношение размера займа к размеру обеспечения по кредиту и др.

- Финансовые показатели: при учете количественных характеристик в модели рекомендуется избегать абсолютных величин и использовать коэффициенты, как например: сумма задолженности к доходу, ежемесячные выплаты по кредиту к ежемесячному доходу, месячный свободно располагаемый бюджет к ежемесячному доходу, ежемесячные выплаты по кредиту к месячному свободно располагаемому бюджету и др.

- Маркетинговые показатели: источник поступления кредитной анкеты, проводимая программа, побуждающий мотив и др.

Следующим основным источником информации является *внутренняя кредитная история банка и информация, полученная в бюро кредитных историй на момент подачи анкеты*. Используемыми скоринговыми переменными могут являться: количество текущих счетов клиента, количество и наличие кредитных карт, общая сумма всех кредитов, время получения последнего кредита, наличие у клиента других продуктов этой финансовой организации, состояние текущего счета, утилизация существующих лимитов, рейтинги бюро кредитных историй и др.

Анализ скоринговых показателей необходимо начинать с проверки их состоятельности и поиска возможных ошибок. В случае неполного наличия данных в выборке необходимо проанализировать причину их отсутствия. Если данные были потеряны или не собраны по вине финансового института, то клиентов с отсутствующими данными следует исключить из дальнейшего анализа или присвоить среднее значение по категории. Если информация была не указана физическими лицами в анкете сознательно, то все пустые значения необходимо отнести в отдельную самостоятельную категорию и анализировать наряду с остальными показателями модели на статистическую значимость. Например, если при анализе переменной «семейное положение» некоторые клиенты не

указали своего семейного положения, то наряду с имеющимися категориями: «женат», «холост», «разведен», «вдовец», появится дополнительно группа «не указано». Кроме того значения переменных необходимо проверить на предмет экстремальных величин, и при наличии таковых удалить их из дальнейшего анализа или присвоить средние результаты по группе.

Корреляционный анализ является важным этапом оценки скоринговых характеристик. Все используемые в модели переменные необходимо проверить на наличие между ними корреляции. В случае отсутствия проблемы мультиколлинеарности необходимо найти оптимальное сочетание между удалением статистически незначимых характеристик и группировкой или выбором одной общей переменной из каждого коррелирующего кластера.

Следующим этапом анализа независимых переменных является проверка их статистической значимости (bi-variate analysis). Данный анализ заключается в проверке наличия и силы связи между одной зависимой и независимой переменными, что позволяет определить, какие переменные являются наиболее точными предсказателями модели. Основные статистические методы, используемые в анализе отдельных характеристик, основываются на сравнении распределения «плохих» и «хороших» клиентов; ими являются:

- 1) статистика  $\chi^2$
- 2) коэффициент Крамера V
- 3) информационное Значение (IV – Information Value) Статистика  $\chi^2$  используется для сравнения двух выборок и определения их схожести. В данном случае сравнивается распределение «платежеспособных» и «неплатежеспособных» клиентов для каждой анализируемой отдельно переменной. Статистика  $t$  определяется по формуле:

$$X^2(v) = \sum (x_i - E)^2 / E,$$

где  $v$  - число степеней свободы,  $x_i$  - наблюдаемое значение,  $E$  - ожидаемое значение.

Если  $\chi^2$  меньше критического значения (по грубому правилу 0,5), то гипотеза  $H_0$  об одинаковом распределении «платежеспособных» и «неплатежеспособных» клиентов подтверждается, и, следовательно, анализируемая переменная не является статистически значимой.

Коэффициент Крамера V - это коэффициент, полученный на основе статистики  $\chi^2$ , который принимает значения от 0 до 1, где 0 говорит о том, что две выборки идентичны, а 1, что они абсолютно разные. Коэффициент рассчитывается по формуле:

$$\text{Коэффициент Крамера V} = \sqrt{\chi^2 / N(k - 1)},$$

где  $\chi^2$  - статистика  $\chi^2$ ,  $N$  - объем выборки,  $k$  - наименьшее число строк и столбцов в выборке.

Данный коэффициент позволяет таким образом сравнивать относительную значимость отдельных скоринговых показателей. Обычно значения его лежат в промежутке 0,06 - 0,35, чем больше значение, тем статистически значимее переменная.

Информационное Значение (IV) - считается самой распространенной мерой определения значимости переменных и измерения разницы в распределении «плохих» и «хороших» клиентов. Информационное Значение определяется по формуле:

$$IV = I(G_i - B_i) \ln(G_i/B_i),$$

где  $G_i$  - процент всех «хороших» случаев,  $B_i$  - процент всех «плохих» случаев.

По грубому правилу значения данного коэффициента трактуются следующим образом:

- ✓ менее 0,02 - статистически незначимая переменная;
- ✓ 0,02 - 0,1 - статистически мало значимая переменная;
- ✓ 0,1 - 0,3 - статистически значимая переменная;

✓ 0,3 и более - статистически сильная переменная.

Заключительным моментом анализа скоринговых показателей является группировка внутри переменных на классы. Особенно это касается интервальных и качественных переменных с большим количеством характеристик, например профессия. Группировка внутри показателей позволяет легче понять существующие зависимости в модели, скоринговые баллы становятся более транспарентны и устойчивы к небольшим изменениям в клиентской базе, она решает проблему экстремальных величин и редких значений и придает больше статистической значимости скоринговым переменным. При группировке на классы внутри независимых показателей следует руководствоваться следующими основными критериями:

1. размер группы – в каждом классе должно содержаться не менее 4-5% выборки;
2. коэффициент дефолта (bad rate) – процентное соотношение «неплатежеспособных» клиентов ко всем клиентам в группе;
3. WOE (Weight of Evidence) – измеряет статистическую значимость каждого класса переменной и рассчитывается как  $\ln(G_i/B_i)$ , где  $G_i$  – процент всех «хороших» случаев,  $B_i$  – процент всех «плохих» случаев;

Таким образом, если внутри скоринговой переменной существуют две подгруппы небольшого размера, которые ведут себя одинаково, имеют сравнимый коэффициент дефолта и одинаковый WOE, то имеет смысл сгруппировать их в один класс, если это не противоречит здравому смыслу. При анализе интервальных переменных следует построить график коэффициента дефолта или WOE и сгруппировать схожие значения в классы. Важно иметь в виду, что любая группировка должна быть в первую очередь логична и полученные связи между различными подгруппами и платежеспособностью клиентов должны быть объяснимы с точки зрения ведения бизнеса.

Анализ скоринговых переменных позволил выделить ряд самых сильных и качественных характеристик (обычно около 12–18 штук), на которых основывается построение статистической модели. Следует иметь в виду, что качественные показатели необходимо преобразовать в фиктивные переменные (dummies), принимающие значение 0 или 1. Для сгруппированных скоринговых показателей мы рекомендуем использовать в качестве значений WOE каждого отдельного класса. Фиктивные переменные предполагают, что разница между одной группой качественной переменной и другой одинакова, использование WOE позволяет решить эту проблему тем, что отражает точное направление и масштаб зависимости между различными классами сгруппированных характеристик. Зависимой характеристикой является в модели платежеспособность клиента, которая также отображается в виде фиктивной переменной, например «неплатежеспособный» клиент получает значение 100, «платежеспособный» клиент – 200 в случае линейной регрессии, или «неплатежеспособный» клиент приобретает значение 0, «платежеспособный» клиент – 1 в случае логарифмической регрессии.

За основу банковской скоринговой системы анализа платежеспособности физических лиц могут быть взяты различные статистические модели. Основными из них являются:

- множественная линейная регрессия,
- множественная логарифмическая регрессия,
- нейросети,
- деревья решений.

*Множественная линейная регрессия* связывает поведение зависимой переменной (платежеспособность клиента) с линейной функцией ряда независимых переменных (скоринговыми характеристиками). Она находит наилучшую линейную зависимость путем минимизации суммы стандартных отклонений и имеет следующую формулу:

$$Y_1 = \beta_1 X_{11} + \beta_2 X_{12} + \beta_j X_{1j} + \dots + \beta_m X_{1mj}$$

где  $y_i$  - зависимая переменная (платежеспособный/неплатежеспособный клиент),  $x_i$  - независимые переменные (скоринговые характеристики),  $\beta$  - параметры модели (скоринговые веса).

Полученные в результате построения модели  $\beta$  значения - это параметры, которые определяют характер связи между наблюдаемым значением переменной «платежеспособность клиента» и соответствующими скоринговыми характеристиками.  $\beta$  значения являются скоринговыми весами.

Множественная логарифмическая регрессия также как и множественная линейная регрессия связывает поведение зависимой переменной с функцией ряда независимых переменных. Она использует принцип максимального правдоподобия, а не наименьших квадратов для того, чтобы достигнуть равновесия. Множественная логарифмическая регрессия находит наилучшее статистическое соответствие и имеет формулу:

$$\text{Logit}(\pi_i) = \alpha + \beta_1 x_i + \beta_2 X_j + \dots + \beta_m x_{mj}$$

где  $\pi_i = \text{Prob}(y_i=y_i/x_i)$  - вероятность позитивного случая,  $\alpha$  - независимый параметр логарифмической регрессии,  $X_i$  - независимые переменные (скоринговые характеристики),  $\beta$  - параметры модели (скоринговые веса).

В логарифмической регрессии скоринговые баллы трансформированы в вероятности со значениями от 0 до 1.

Нейросети - это статистическая модель, которая состоит из множества нейронов, сгруппированных в слои, создающие сеть. Каждый нейрон - это обрабатывающий элемент с заданной единичной функцией. Связи между нейронами создают сеть, позволяющую определить взаимосвязи между отдельными данными. Типичный пример структуры нейросети показан на рисунке 1.

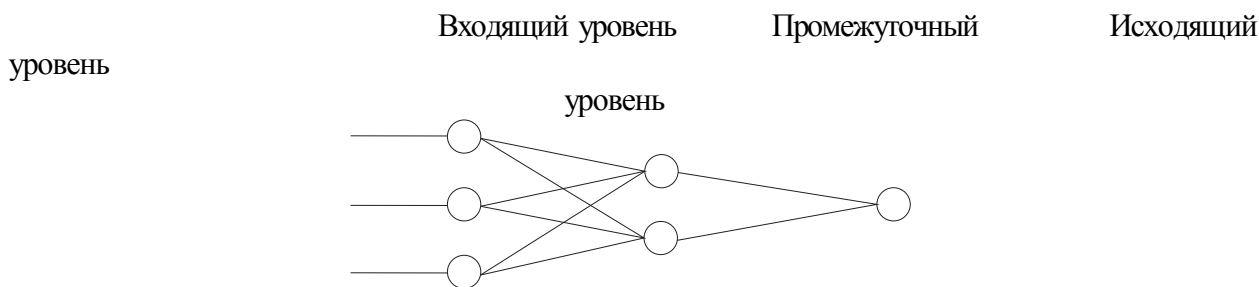


Рис. 1. Упрощенная схема механизма нейросети

Применительно к скоринговой модели: - входящий уровень состоит из всех характеристик, используемых для предсказания платежеспособности. Исходящий уровень представляет собой ожидаемый результат (платежеспособность клиента). Входящими и исходящими данными модели могут являться как количественные, так и качественные показатели, трансформированные в фиктивные переменные.

Деревья решений – это модель, строящаяся на логической цепочке правил, которые пытаются описать отдельные взаимосвязи между данными относительно ожидаемого результата. Структура деревьев решений открыто показывает аргументацию правил и поэтому позволяет легко понять процесс принятия решения. Графически упрощенная модель показана на рисунке 2.

При построении большинства из выше перечисленных моделей необходимо руководствоваться принципом подбора входящих переменных. Изначальный отбор входящих переменных был сделан выше в процессе анализа скоринговых переменных, однако не все из этих отобранных характеристик должны попасть в конечную модель, а лишь статистически значимые. Это связано с тем, что, признанные статистически значимыми в процессе индивидуального анализа, переменные могут потерять силу в модели при взаимодействии с остальными входящими показателями. Целью подбора переменных является поиск независимых характеристик, которые обеспечивают оптимальную комбинацию. Современные статистические программы, как правило,

осуществляют процесс подбора автоматически. Например, один из способов – это построение пробной модели со всеми возможными показателями. Далее осуществляется поиск самой статистически сильной переменной и включение ее в новую окончательную модель. Затем система ищет следующую самую сильную характеристику и строит модель с двумя переменными и т. д., пока в модель не будут включены все значимые характеристики. Значимость показателей определяется на основе F-статистик, t-статистик и др.

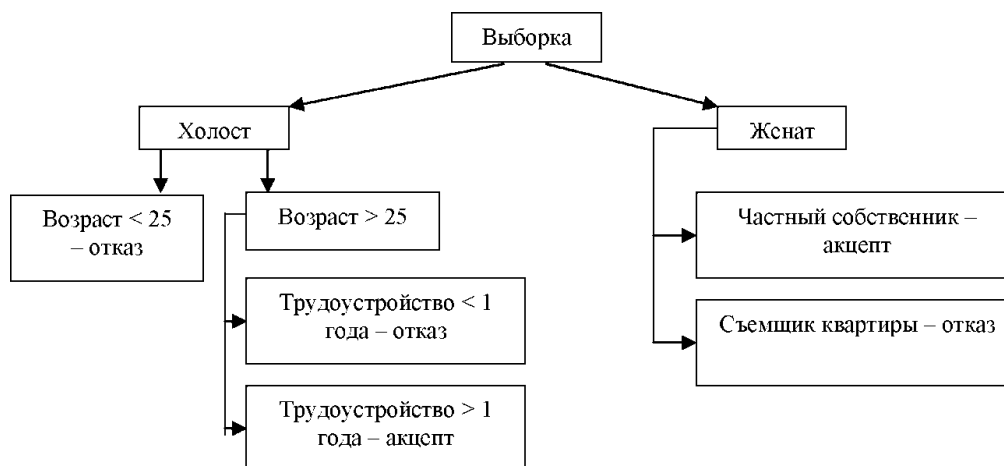


Рис. 2. Упрощенная структура деревьев решений

Финансовые институты свободны в выборе любой статистической модели для построения своих скоринговых систем. Выбор модели может быть продиктован путем сравнения статистических показателей качества модели, основными из которых являются статистика Колмогорова-Смирнова, коэффициент Джини (Gini coefficient) и область под ROC кривой (Receiver Operating Curve). Чем выше каждая из этих статистик, тем качественнее считается построенная модель.

С целью оптимизации процесса оценки платежеспособности клиента, и принятия верных и эффективных решений относительно выдачи кредита важно понимать и успешно применять все компоненты скоринговой системы. Правильно применяя данную технологию, любой банк сможет сократить время принятия решений по выдаче кредита, управлять кредитными рисками и минимизировать субъективность при рассмотрении заявок. Общими ошибками при внедрении скоринговой системы являются недостаточная работа объектов интеллектуальной собственности, невнимание к обучению персонала, неправильная оценка влияния скоринговой модели и слабая заинтересованность участие руководства в процессе, а также отсутствие координации общих усилий по внедрению.

Знания основных категорий и трудностей, на которые необходимо обратить внимание при установке новой или обновленной скоринговой модели и системы в целом, позволяет кредитной организации более качественно подготовиться к успешному внедрению со всеми необходимыми элементами управления. Скоординированное участие отделов управления рисками, маркетинга, обслуживания клиентов, контроля качества, обучения, системной поддержки, конечных пользователей, а также операционного и юридического отделов важно для успеха любого внедрения скоринговой системы.

### Список литературы

1. Ишина И.В. Скоринг – модель оценки кредитного риска [Текст] / И.В.Ишина // Аудит и финансовый анализ. - 2007. - № 4.
2. Ежедневная аналитическая газета «РБК daily».[Электронный ресурс] Режим доступа: <http://rbcdaily.ru/2011/07/21/finance/562949980687405>
3. Мэйз Э. Руководство по кредитному скорингу [Текст] / Э.Мэйз. –

Минск: Издательство «Гревцов Паблшер», 2008. – 464 с.

4. Ковалев М. Методика построения банковской скоринговой модели для оценки кредитоспособности физических лиц [Текст] / М. Ковалев, В. Корженевская // Банки Казахстана. – 2008. - №1. – с. 43–48.

5. Сорокин А.С. К вопросу оценки согласованности мнений экспертов при использовании методов экспертного оценивания в кредитном скоринге [Текст] / А.С. Сорокин // Сб. ст. по мат. IX междунар. научн. конгресса. – М.: «Эдитус», 2014. – с. 281-283.