

# СЕКЦИЯ «ПРОБЛЕМЫ ЭЛЕКТРОННОГО БИЗНЕСА»

## АЛГОРИТМЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ СОЗДАНИЯ СИСТЕМ КРЕДИТНОГО СКОРИНГА

*Ковалёв Н.М., Поддубный Д.П.*

*Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Ермакова Е.В. – канд. экон. наук, доцент*

Кредиторы используют систему кредитного скоринга или числовую систему, чтобы измерить вероятность того, что заемщик будет платить за деньги, которые он или она занимает. Он создается путем присвоения баллов различным атрибутам, связанным с кредитоспособностью заявителя.

Кредитный скоринг — система оценки кредитоспособности лица, основанная на численных статистических методах. Кредитный скоринг широко используется как крупными банками в кредитовании. Скоринг заключается в присвоении баллов по заполнению некой анкеты, а также на основе предсказаний конкретной модели. По результатам набранных баллов системой автоматически принимается решение об одобрении или отказе в выдаче кредита.

Система кредитного скоринга позволяет кредиторам и другим финансовым учреждениям определять кредитоспособность человека. Некоторые финансовые организации устанавливают собственный метод кредитного скоринга.

Многие факторы влияют на кредитные рейтинги, присвоенные через системы. Факторы включают в себя проценты по платежам, продолжительность использования кредита, сумму долга человека и виды долга, который есть у этого человека.

Кредиторы используют эти методы, чтобы определить, какому риску их подвергает конкретный заемщик, если они решат предоставить ему кредит. Эти цифры основаны на риске.

Если у человека низкий кредитный рейтинг, он или она, скорее всего, заплатит больше, чтобы занять деньги на покупку дома или на покупку автомобиля, чем человек с более высоким кредитным рейтингом. В то время как системы кредитного скоринга устанавливают руководящие принципы, отдельные кредиторы определяют, какой уровень является приемлемым и сколько взимать процентов.

Кредитный скоринг, применяемый для потенциального заемщика, делится на несколько видов:

- заявочный скоринг – используется для оценки платежеспособности потенциального заемщика, когда тот подает заявку на кредит, применяется в связи с высоким процентом невозврата кредитов, таким образом помогая принимать решение о выдаче или отказе в кредите, в том числе предложению дополнительных условий для заемщика.

- поведенческий скоринг – выдает прогноз изменения платежеспособности потенциального заемщика в результате влияния внешних факторов и определяет наиболее оптимальные условия для погашения кредита. В данном случае источником данных является информация о финансовых операциях потенциального заемщика в течении определенного времени.

- скоринг мошенничества – данный тип скоринга используется банками наряду с другими видами скоринга для определения риска мошеннических действий со стороны потенциального заемщика.

При выдаче кредита учитываются многие факторы. К ним относятся характеристики заемщика, их экономическое положение, сумма запрашиваемого кредита, его цель и тип залога. Разнообразие этих факторов означает, что риск оценивается с использованием элементов количественного и качественного анализа [1].

Количественный анализ включает в себя, прежде всего, оценку материального положения клиента на основе его доходов и ежемесячных расходов. Он также может включать анализ движения денежных средств по счетам клиентов и кредитной истории. В то время как качественная оценка учитывает, среди прочего, семейное положение, образование или форму занятости - для физических лиц, а для предприятий - организационно-правовую форму, отрасль, в которой они работают, или способ ведения бухгалтерского учета [2].

Мы часто слышим такие словесные конструкции, как «машинное обучение», «нейронные сети». Алгоритмы машинного обучения могут решать множество различных типов задач, в том числе помогать банкам. Использование скоринговых систем, дает возможность быстро принять решение о возможности предоставления кредита. Они позволяют оценить кредитоспособность заемщика, основываясь на статистических методах. Методы машинного обучения позволяют в значительной степени автоматизировать эти процессы.

Банки обычно выдают кредиты на основе модели кредитного скоринга, которая сочетает в себе качественный и количественный анализ. Кредитный скоринг основан на статистических методах, благодаря которым можно прогнозировать вероятность наступления определенного события в будущем – в данном случае дефолта по кредиту.

В процессе скоринга используется информация о покупателе, собранная на этапе подачи заявки, в основном данные, характеризующие клиента, но также и информация об его прошлом поведении. Каждое кредитное учреждение рассматривает разный набор характеристик и присваивает им разное количество баллов. Например, высокообразованный человек обычно набирает больше баллов, чем человек, не окончивший колледж, но точное количество баллов и его влияние на окончательный балл могут варьироваться в зависимости от банка. Сумма баллов за определенные характеристики обычно является итоговой оценкой.

Рассматриваемая задача прогнозирования надежности заемщика представляет собой задачу классификации. Классификация — задача разделения множества наблюдений или объектов на группы. Таким образом клиентов можно разделить на некоторое число групп, например, надежные и сомнительные. Таким образом для достижения более лучшего результата необходимо учитывать множество факторов о клиенте и иметь большую выборку, которая используется для обучения нейронной сети. Чем больше выборка - тем выше точность кредитного скоринга.

В основе моделей для решения данной задачи всегда лежит классификатор. Классификатор - это алгоритм, который сопоставляет входные данные с определенной категорией. Вид классификатора определяется математическими методами, которые разделяют данные на категории на основе определенных характеристик, которые выбираются на этапе сбора данных [3].

Для построения простейших моделей кредитного скоринга достаточно использовать классический бинарный классификатор, который будет делить всех заемщиков на тех, которым мы можем дать займ и тех, которым будет отказано в займе. Для принятия такого решения необходимы как данные, полученные напрямую от заемщика или государственных органов: налоговая отчетность; паспортные данные владельцев, ген. директора, гл. бухгалтера; выписки ЕГРЮЛ, ЕГРИП; правоустанавливающие документы; данные по задолженностям; данные по судебным делам и прочие. Также нам будут необходимы более специфические данные, которые мы можем получить при помощи математических методов: взаимодействие с гос. органами – подряд, субподряд или поставка; взаимодействие с компаниями из топ-100; наличие в окружении заемщика компаний банкротов, должников, компаний с низким скоринговым баллом; участие в благотворительных организациях и прочие.

Выделив основные характеристики, мы можем использовать алгоритмы логистической регрессии для того, чтобы сделать классификацию. Логистическая регрессия - метод построения линейного классификатора, позволяющий оценивать апостериорные вероятности принадлежности объектов классам. Данный алгоритм классификации использует сигмовидную функцию в качестве функции активации и позволяет дать вероятностную оценку принадлежности объекта каждому классу. Тем самым мы можем выделить вышеописанные классы. Более того, мы можем увеличить количество классов потенциальных заемщиков.

Для повышения точности классификатора мы можем также использовать алгоритмы дерева решений, случайного леса и метода опорных векторов. Как правило в реальных системах используется не один алгоритм, а несколько. Это называется ассамблея алгоритмов.

Вывод:

Так благодаря алгоритмам машинного обучения мы можем создать более точные модели кредитного скоринга, которые могут учитывать множество факторов и принимать решения о выдаче кредита быстрее и качественнее.

*58-я научная конференция аспирантов, магистрантов и студентов БГУИР, 2022 г*

**Список использованных источников:**

1. Кредитный скоринг. Сокрушающий эффект ИИ при оценке кредитоспособности заёмщика [Электронный ресурс] - Электронные данные. - Режим доступа: <https://habr.com/ru/post/652987/> - Дата доступа: 01.03.2022
2. Краткий курс машинного обучения или как создать нейронную сеть для решения скоринг задачи [Электронный ресурс] - Электронные данные. - Режим доступа: <https://habr.com/ru/post/340792/> - Дата доступа: 01.03.2022
3. Artificial Intelligence Programs [Электронный ресурс] - Электронные данные. - Режим доступа: <https://learn.stanford.edu/Artificial-Intelligence-Programs.htm> - Дата доступа: 01.03.2022