

СКОРИНГОВЫЕ МОДЕЛИ ДЛЯ ОЦЕНКИ КРЕДИТОСПОСОБНОСТИ ЗАЕМЩИКОВ

Лукьянец С.П., Кирий В.В., Маркус А.Т.

Харьковский национальный университет радиоэлектроники

In the banking sector in the management of credit risk is one of the key objectives - the credit rating of the borrowers. Evaluation results individual risks are the basis for the risk analysis of the total loan portfolio. Assessment of risk of loan default by the borrower to a particular practice is carried out in the two main approaches - based on the subjective opinions of experts or based automated scoring systems.

Кредитным скорингом называется быстрая, точная и устойчивая процедура оценки кредитного риска, имеющая научное обоснование. Скоринговая модель является математической моделью, позволяющей сопоставить характеристикам заемщика численное значение – скоринговый рейтинг, характеризующий кредитоспособность (вероятность успешного исхода кредитной сделки). Моделей скоринга множество, каждая из них использует свой набор факторов, характеризующих риск, связанный с кредитованием заемщика, и получает в результате пороговую оценку, которая и позволяет разделять заемщиков на «плохих» и «хороших».

Скоринг можно отнести к классификационным задачам, в которых на основе имеющихся данных требуется получить функцию, которая разделяет клиентов на две эти группы наиболее точно. Однако имеющиеся данные сперва должны быть представлены в форме, подходящей для анализа. Как с качественными, так и с количественными характеристиками можно работать, используя два подхода:

1. Каждый из признаков преобразуется в двоичную переменную. При таком подходе количество переменных может стать очень высоким, что становится неудобным. С другой стороны, этот подход не создает никаких дополнительных отношений между независимыми и зависимой переменными.

2. Каждая характеристика может быть преобразована в переменную, принимающую такие значения, которые соответствуют отношению числа неблагонадежных клиентов к числу благонадежных с таким же признаком или логарифм от этого отношения. Каждый признак в этом случае получает числовую характеристику уровня его «рискованности».

Существует множество различных методов классификации, среди которых можно отметить следующие: статистические методы (линейная регрессия, логистическая регрессия); самоорганизующиеся карты; дерево классификации; нейронные сети.

Наиболее распространены среди них регрессионные методы, а особенно многофакторная регрессия:

$$P = w_0 + w_1x_1 + \dots + w_nx_n, \quad (1)$$

где P – вероятность невозврата кредита, w - весовые коэффициенты, x - характеристики заемщика.

Недостатком модели является то, что в левой части уравнения должна быть строго положительная величина (вероятность может принадлежать отрезку от 0 до 1), а переменные в правой части могут быть как больше, так и меньше нуля.

Этой проблемы позволяет избежать логистическая регрессия:

$$\log(P / (1 - P)) = w_0 + w_1x_1 + \dots + w_nx_n. \quad (2)$$

Логистическая регрессия – это разновидность множественной регрессии, общее назначение которой состоит в анализе связи между несколькими независимыми переменными (называемыми регрессорами или предикторами) и зависимой переменной. Бинарная логистическая регрессия, как следует из названия, применяется в случае, когда зависимая переменная является бинарной (т.е. может принимать только два значения). Иными словами, с помощью логистической регрессии можно оценивать вероятность того, что событие наступит для конкретного испытуемого (больной/здоровый, возврат кредита/дефолт и т.д.).

Как дерево классификации, так и нейронные сети являются системами, разделяющими клиентов на группы с одинаковым уровнем риском и максимальным отличием от уровня риска прочих групп. Нейронные сети часто используют при оценке кредитоспособности юридических лиц, в которой используются выборки меньшего размера чем для аналогичной задачей для физических лиц, а также во Fraud-скоринге, благодаря возможности таких сетей выявлять нестандартные ситуации.

Каждый из этих методов имеет собственные преимущества и недостатки а их выбор зависит от стратегии, которой следует банк и от того, какие требования при разработке моделей для него наиболее приоритетны.

Подобные модели могут использоваться как по отдельности, так и совместно, повышая качество модели. Они позволяют: формализовать процесс оценки рисков; снизить уровень риска выдачи кредита недобросовестному заемщику; обрабатывать значительно большее количество заявок за краткие сроки; легко тиражировать модели; сократить расходы на содержание штата сотрудников, оценивающих заявки; постоянно адаптировать модель под меняющиеся условия.