

ОСОБЛИВОСТІ МОДЕЛЮВАННЯ КРЕДИТНИХ РИЗИКІВ ФІЗИЧНИХ ОСІБ

Дослідження з побудови моделей поведінкового скорингу позичальників-фізичних осіб ґрунтувалось на інформації щодо соціо-демографічних та фінансових характеристик позичальників на момент отримання кредиту (без можливості ідентифікації самої особи). Всього в дослідженні було проаналізовано більше 80 характеристик позичальника як узятих із його кредитної заявки та офіційних документів, так і розрахованих на основі даних із баз бюро кредитних історій, що містять інформацію щодо виконання позичальником попередніх та діючих кредитних угод, звернень до кредитно-фінансових установ тощо. За кожним позичальником враховувалась уся відома щодо нього кредитна історія за максимально тривалий проміжок часу (у наявності була інформація більш ніж за двадцятирічний період щодо десятків мільйонів кредитів).

Навчання та тестування скорингових моделей проводилось на статистиці за нових економічних реалій – на кредитах, виданих протягом 2015 року (кількох десятків тисяч, за якими вже була інформація щодо факту виконання або невиконання кредитних зобов'язань). При цьому, щоб модель не підлаштувалась під конкретний клас позичальника, до навчальної вибірки угоди, за якими зобов'язання були виконані, та які виявились дефолтними, були взяті в однаковій кількості. Якщо ж формувати навчальну вибірку пропорційно класам кредитів в реальності, то модель може налаштуватись навіть так, щоб усі кредити визнавати надійними (оскільки дефолтних кредитів майже у десять разів менше, то при оптимізації моделі ними може бути знехтувано, оскільки при незначному зменшенні альфа-помилки суттєво більше починає зростати бета-помилка, знижуючи загальну точність класифікації).

Оскільки і навчальна, і тестова вибірки склались з дійсно великої кількості спостережень, то базовою гіпотезою при побудові скорингових моделей було те, що для врахування значного різноманіття характеристик кредиту й позичальника та виявлення складних нелінійних залежностей результуючого показника (класу кредиту, що приймає значення 0 для кредитів, за якими зобов'язання були виконані, та 1 – для дефолтних) від вхідних факто-

рів необхідно конструювати досить масштабні економіко-математичні моделі з великою кількістю внутрішніх параметрів. Найбільш відповідним цим вимогам інструментарієм є нейронні мережі персептронного типу, де регулювати кількість параметрів можна за рахунок зміни числа прихованих шарів та їх нейронів.

Однак, в результаті проведення численних модельних експериментів чітко проявився ефект перенавчання, коли зі збільшенням кількості параметрів у нейронній мережі (причому неважливо, за рахунок збільшення кількості нейронів у одному прихованому шарі або ж одночасного збільшення кількості шарів) похибка відтворення вихідних даних на навчальній вибірці суттєво зменшувалась, але при цьому помилка класифікації на тестовій вибірці зростала.

Результати тестування логістичних регресій та нейронних мереж різної конфігурації (персептронів з одним прихованим шаром із різною кількістю нейронів) за коефіцієнтом Джині подано у табл. 1.

Таблиця 1

РЕЗУЛЬТАТИ ОЦІНЮВАННЯ АДЕКВАТНОСТІ ЛОГІСТИЧНИХ РЕГРЕСІЙ ТА НЕЙРОННИХ МЕРЕЖ РІЗНОЇ КОНФІГУРАЦІЇ ЗА КОЕФІЦІЄНТОМ ДЖИНІ

Вихідні дані	Початкові		Переведені у WOE	
Тип вибірки	Навчальна	Тестова	Навчальна	Тестова
Обсяг навчальної вибірки: 1000 значень				
Логістична регресія	0,793	0,820	0,659	0,562
2 нейрони	0,799	0,840	0,764	0,506
3 нейронів	0,798	0,828	0,798	0,500
5 нейронів	0,843	0,767	0,876	0,459
10 нейронів	0,908	0,648	0,986	0,345
Обсяг навчальної вибірки: 2000 значень				
Логістична регресія	0,740	0,830	0,620	0,594
2 нейрони	0,775	0,873	0,667	0,573
3 нейронів	0,773	0,855	0,725	0,521
5 нейронів	0,761	0,680	0,760	0,549
10 нейронів	0,843	0,593	0,908	0,397

Закінчення табл. 1

Вихідні дані	Початкові		Переведені у WOE	
Тип вибірки	Навчальна	Тестова	Навчальна	Тестова
Обсяг навчальної вибірки: 4000 значень				
Логістична регресія	0,798	0,813	0,612	0,604
2 нейрони	0,825	0,867	0,654	0,580
3 нейронів	0,844	0,828	0,698	0,553
5 нейронів	0,832	0,755	0,778	0,534
10 нейронів	0,858	0,562	0,846	0,419
Обсяг навчальної вибірки: 6000 значень				
Логістична регресія	0,789	0,830	0,613	0,606
2 нейрони	0,798	0,808	0,625	0,601
3 нейронів	0,794	0,606	0,678	0,579
5 нейронів	0,855	0,503	0,739	0,549
10 нейронів	0,873	0,476	0,818	0,472
Обсяг навчальної вибірки: 8000 значень				
Логістична регресія	0,819	0,835	0,617	0,579
2 нейрони	0,829	0,787	0,650	0,602
3 нейронів	0,865	0,602	0,674	0,581
5 нейронів	0,889	0,488	0,691	0,554
10 нейронів	0,893	0,424	0,742	0,533

Зауважимо, що логістична регресія (яка є найбільш поширеним у банківській практиці способом реалізації скорингової моделі) за своєю сутністю є повним аналогом одного єдиного нейрона із сигмоїдною функцією активації. Але навіть така проста конструкція моделі виявляється дуже ефективною та у багатьох випадках демонструє вищу адекватність за нейронні мережі, що видно з табл. 1. Тобто, зі збільшенням параметрів моделі зростала здатність до відтворення залежностей у навчальному наборі даних, але таким чином відбувалось звичайне підлаштування моделі під цю статистику зі зменшенням здатності до узагальнення

(що видно по падінню коефіцієнту Джині при перевірці на тестовій вибірці).

Обсяг навчальної вибірки не здійснював якогось помітного впливу на точність моделі при її тестуванні. Тобто, побудовані моделі ефективно виявляли закономірності в поведінці позичальників вже за досить невеликих обсягів навчальної вибірки. Цікавим виявився факт суттєвого зниження адекватності всіх побудованих моделей при застосуванні прийнятого в практиці конструювання скорингових карт біннінгу показників із переведенням їх значень у WOE. Причому для цього спеціально був розроблений алгоритм категоризації, але дослідження у цьому напрямку ще триватимуть.

Найкращий результат моделювання кредитних ризиків продемонстрували персеPTRони найпростішої конфігурації (з двома трьома нейронами на єдиному прихованому шарі). Результати перевірки таких моделей на тестовій вибірці сягали коефіцієнту Джині 0,87, що є надзвичайно високим показником адекватності скорингової моделі. Причому ці результати відрізнялись стійкістю за різних умов налаштування та тестування моделей.

Маханець Л. Л.

к.е.н., доцент

*Чернівецький національний університет
імені Юрія Федьковича*

ВИЗНАЧЕННЯ МІРИ МАНЕВРНОСТІ ОБСЯГУ ЗОВНІШНЬОГО БОРГУ УКРАЇНИ

Сучасна економічна ситуація в Україні характеризується загрозливим зростанням боргового навантаження, що виступає одним з базових факторів гальмування розвитку економіки країни. Хоча валовий зовнішній борг України продовжує скорочуватися другий рік поспіль, однак це відбувається через скорочення зобов'язань у приватному секторі, тоді як державний зовнішній борг на кінець 2015 року зріс до історичного максимуму у \$43 млрд. [1].

Для того щоб уникнути нового рефінансування, необхідно розробити й змодельовати стратегію й тактику поведінки держави на зовнішньому і внутрішньому кредитному ринках. Це можна