

Перспективи розвитку методу лежать у площині використання алгоритмів машинного навчання для динамічного калібрування параметрів моделей та впровадження хмарних обчислень для оптимізації швидкості симуляцій. Метод Монте-Карло симуляції імплементований в різних інструментах від універсальних (R, Python, MatLab, MS Excel тощо) до вузькоспеціалізованих та веб-інструментах, що використовують хмарні калькулятори та платформи.

Список використаних джерел

1. Парфенцева Н. О., Голубова Г. В. Моделювання фінансових ризиків на основі статистичних методів оцінювання. Науковий вісник Національної академії статистики, обліку та аудиту: зб. наук. пр.. 2022. №1-2. С.14-20. doi:10.31767/nasoa.1-2-2022.02
2. Герасименко С.С., Голубова Г.В. Статистична оцінка ризиків банківської діяльності. Науковий вісник Національної академії статистики, обліку та аудиту, (3-4), 5-14. <https://doi.org/10.31767/nasoa.3-4-2022.01>
3. Голубова Г.В. Індекс фінансового стресу як індикатор фінансової стабільності України в умовах війни. Статистика України. 2022. № 2. С. 4–11. Doi: 10.31767/su.2(97)2022.02.01

Грабарєв А. В.,
к.е.н., доцент,

академік Академії економічних наук України,
Київський національний економічний університет імені Вадима Гетьмана,

Баранюк М. Р.,
здобувач третього (наукового) рівня вищої освіти,
Київський національний економічний університет імені Вадима Гетьмана

СКОРИНГОВІ ПІДХОДИ ТА МОДЕЛІ В УПРАВЛІННІ КРЕДИТНИМ РИЗИКОМ БАНКУ

Сучасний розвиток банківської системи супроводжується посиленням ролі кількісних методів оцінювання кредитного ризику, оскільки саме кредитний ризик залишається одним із ключових джерел потенційних втрат банку. Зростання обсягів даних, цифровізація кредитного процесу, ускладнення поведінки позичальників та підвищення вимог до якості внутрішнього контролю зумовлюють необхідність використання більш формалізованих, відтворюваних і технологічно придатних інструментів ухвалення кредитних рішень. У цьому контексті скорингові моделі набувають особливого значення як аналітичний механізм, що дає змогу трансформувати інформацію про позичальника у кількісну оцінку рівня ризику, ймовірність дефолту та відповідний управлінський сигнал [1; 2].

Актуальність теми для української банківської практики посилюється тим, що кредитні рішення ухвалюються в умовах макроекономічної волатильності, воєнних ризиків, структурних зламів, зміни платіжної поведінки клієнтів і нерівномірної якості історичних даних. За таких умов скоринг не може розглядатися лише як технічний алгоритм класифікації заявок. Він має виконувати ширшу функцію: забезпечувати ризик-орієнтоване ранжування клієнтів, підтримувати андеррайтинг, моніторинг портфеля, раннє виявлення погіршення кредитної якості та обґрунтування рішень у межах внутрішньої кредитної політики банку.

Метою дослідження є узагальнення прикладних можливостей традиційних і сучасних скорингових моделей в управлінні кредитним ризиком банку, а також обґрунтування доцільності використання гібридного підходу в умовах українського банківського середовища. Методологічну основу становлять порівняльний аналіз, систематизація наукових підходів, матричне оцінювання моделей та імітаційна апробація на синтезованому демонстраційному наборі даних, сформованому за логікою реального роздрібного кредитного портфеля. Застосування демонстраційних даних є методично виправданим, оскільки первинна

банківська інформація про позичальників, кредитні операції та кредитні історії має обмежений доступ і охороняється режимом банківської таємниці.

У практиці кредитного ризик-менеджменту традиційні скорингові моделі, зокрема логістична регресія, дискримінантний аналіз, probit-моделі та класичні scorecards, зберігають важливе значення завдяки прозорості, відносній простоті впровадження, зручності валідації та високій регуляторній прийнятності. Їх перевага полягає в можливості пояснення впливу окремих факторів на результат і документування логіки кредитного рішення. Водночас такі моделі часто обмежені у здатності відображати нелінійні зв'язки, складні взаємодії ознак та швидкі зміни поведінки позичальників у кризовому середовищі [3; 4].

Сучасні моделі машинного навчання, серед яких decision trees, Random Forest, gradient boosting, XGBoost, LightGBM, нейронні мережі та ансамблеві алгоритми, мають вищий потенціал щодо виявлення складних закономірностей у великих і неоднорідних масивах даних [5]. Їх застосування може підвищувати точність прогнозування дефолту, особливо за наявності поведінкових, транзакційних і нетрадиційних ознак. Проте використання таких моделей у банках пов'язане з низкою обмежень: складністю пояснення індивідуального рішення, підвищеними вимогами до якості даних, необхідністю регулярної ревалідації, контролю data drift та впровадження інструментів explainable AI.

Для порівняльного оцінювання скорингових моделей доцільно застосовувати не лише показники прогнозної точності, а ширшу систему критеріїв. До таких критеріїв належать дискримінаційна сила моделі, чутливість до дефолтного класу, якість калібрування прогнозних імовірностей, інтерпретованість, придатність до аудиту, вимоги до даних, складність технологічної реалізації, відповідність внутрішнім процедурам банку та стійкість до нестабільного середовища. Саме такий підхід дає змогу оцінювати модель не як ізольований математичний алгоритм, а як елемент системи кредитного ризик-менеджменту.

Таблиця 1

Порівняльна характеристика скорингових підходів в управлінні кредитним ризиком банку

Підхід	Основні переваги	Ключові обмеження	Практична доцільність
Логістична регресія, scorecards	Прозорість, стабільність, простота валідації, можливість пояснення факторів ризику	Обмежена гнучкість за нелінійних залежностей і структурних зламів	Базова інтерпретована модель для кредитної політики, PD-оцінювання та внутрішнього контролю
Дискримінантний аналіз	Зрозуміла статистична логіка, придатність як контрольний класичний підхід	Чутливість до припущень щодо розподілу даних, нижча адаптивність	Допоміжна модель для порівняння результатів і методичної перевірки
Random Forest	Вища точність, робота з нелінійностями, стійкість до частини шумових ознак	Менша прозорість, потреба в додаткових інструментах пояснення	Підсилення скорингу у складніших сегментах кредитного портфеля
XGBoost / LightGBM	Висока прогностична сила, чутливість до складних взаємозв'язків, гнучкість налаштування	Складніша валідація, ризик “чорної скриньки”, високі вимоги до governance	Поглиблене ризикове уточнення за умови калібрування та explainability-контролю
Гібридний підхід logit + ML re-ranking	Поєднання пояснюваного статистичного ядра з ML-уточненням у зоні невизначеності	Потреба у складнішій архітектурі, моніторингу та регулярній ревалідації	Найбільш збалансований варіант для волатильного середовища та прикордонних кредитних рішень

Складено авторами на основі узагальнення підходів до скорингового моделювання [1-5].

У межах демонстраційної апробації, узагальненої у вихідній статті, було сформовано умовний масив із 1000 спостережень, що імітує структуру роздрібного кредитного портфеля банку. Частка дефолтних випадків становила 22,0 %, недефолтних — 78,0 %, що відображає типовий для скорингових задач помірний дисбаланс класів. До набору змінних включено вік позичальника, стаж роботи, дохід, боргове навантаження, кількість активних кредитів,

історичну прострочку, суму і строк кредиту, наявність забезпечення, тип продукту, регіональний ризик та ознаку шокового середовища. Для перевірки моделей використано часовий розподіл вибірки 60/20/20, 5-fold cross-validation, out-of-time test, вагове коригування класів, SMOTE для ML-моделей та калібрування ймовірнісних виходів.

Результати апробації показали, що XGBoost забезпечив найвищу прогнозну точність: AUC-ROC становив 0,861, Gini — 0,722, recall для дефолтного класу — 0,781, Brier score — 0,141. Логістична регресія продемонструвала нижчу формальну точність (AUC-ROC 0,781), проте зберегла суттєву перевагу з позицій прозорості, валідації та регуляторного використання. Гібридна модель logit + ML re-ranking досягла AUC-ROC 0,848, Gini 0,696, recall 0,773 і Brier score 0,144, що свідчить про її здатність забезпечити кращий баланс між точністю, пояснюваністю, каліброваністю та практичною придатністю.

Операційна логіка гібридного підходу полягає у двоступеневій побудові кредитного рішення. На першому етапі логістична регресія формує базову ймовірність дефолту і забезпечує прозоре первинне ранжування заявок. На другому етапі складніший ML-алгоритм застосовується не до всієї вибірки, а до так званої “сірої зони”, тобто сегмента заявок із прикордонним рівнем ризику. У цьому сегменті ML re-ranking виконує повторне впорядкування позичальників за складнішими нелінійними ознаками. Така архітектура зменшує ризик неконтрольованого використання “чорної скриньки”, оскільки базове рішення залишається пояснюваним, а складніший алгоритм використовується як інструмент поглибленого уточнення.

Для українських банків такий підхід є особливо релевантним. В умовах війни, інфляційних шоків, зміни регіональних ризиків, міграції клієнтів, трансформації джерел доходів і фрагментарності історичних даних модель, побудована виключно на минулих закономірностях, може швидко втрачати прогностичну силу. Тому скорингова система повинна містити механізми регулярної рекалібровки, перегляду cut-off thresholds, контролю стабільності змінних, моніторингу drift-ефектів і перевірки результатів на out-of-time вибірках. Це дає змогу підтримувати не лише математичну точність моделі, а й її управлінську корисність.

Отже, порівняльне оцінювання скорингових моделей засвідчує, що вибір оптимального підходу не може ґрунтуватися лише на максимізації одного показника точності. Для банку важливо поєднати прогнозну силу, пояснюваність, якість калібрування, технологічну реалізованість і стійкість до структурних змін. У цьому контексті класичні моделі доцільно розглядати як прозоре статистичне ядро скорингової системи, ML-моделі — як інструмент підвищення чутливості до складних ризикових патернів, а гібридні рішення — як найбільш збалансований формат для нестабільного середовища.

References

1. Thomas L. C., Crook J. N., Edelman D. B. Credit Scoring and Its Applications. Philadelphia : SIAM, 2002. 248 p.
2. Crook J. N., Edelman D. B., Thomas L. C. Recent developments in consumer credit risk assessment. *European Journal of Operational Research*. 2007. Vol. 183, Issue 3. P. 1447-1465. DOI: <https://doi.org/10.1016/j.ejor.2006.09.100>
3. Banasik J., Crook J. Reject inference, augmentation, and sample selection. *European Journal of Operational Research*. 2007. Vol. 183, Issue 3. P. 1582-1594. DOI: <https://doi.org/10.1016/j.ejor.2006.06.072>
4. Pang S., Hou X., Xia L. Borrowers' credit quality scoring model and applications, with default discriminant analysis based on the extreme learning machine. *Technological Forecasting and Social Change*. 2021. Vol. 165. Art. 120462. DOI: <https://doi.org/10.1016/j.techfore.2020.120462>
5. Gambacorta L., Huang Y., Qiu H., Wang J. How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability*. 2024. Vol. 73. Art. 101284. DOI: <https://doi.org/10.1016/j.jfs.2024.101284>