

УДК 004.9:519.226

**Н. В. Кузнецова, П. І. Бідюк**

Навчально-науковий комплекс «Інститут прикладного системного аналізу»

Національного технічного університету України «КПІ»

пр. Перемоги, 37, 03056 Київ, Україна

e-mail: natalia-17@mail.ru

## Нейронні та мережі Байєса у задачі аналізу кредитних ризиків

*Робота присвячена аналізу дефолтів позичальників кредиту фінансової установи з використанням трьох типів математичних моделей і фактичних даних з банківської установи. Представлено результати побудови та практичного застосування моделей у формі нейронної мережі зворотного розповсюдження, статичної байєсівської мережі та інтегрованої моделі, яка складається з двох указаних структур. Виконано ряд обчислювальних експериментів стосовно прогнозування дефолтів позичальників кредитів з використанням кожної побудованої моделі окремо, а також комбінованої (інтегрованої) моделі. Показано, що кращий результат на використаних вибірках даних забезпечує комбінована модель, і встановлено, що для розв'язання задачі прогнозування дефолтів клієнтів банку доцільно застосовувати множину різних моделей, інтегроване використання яких дає можливість підвищити якість оцінок прогнозів.*

**Ключові слова:** інтегрований підхід, нейронні мережі, мережі Байєса, кредитні ризики.

На сьогоднішньому етапі розвитку світової економіки спостерігається зростання динаміки фінансових процесів, що потребує виконання поглибленого своєчасного аналізу поточних ситуацій та прискореного прийняття рішень аналітиками в реальному часі. Серед задач фінансового аналізу особливе місце займають задачі оцінювання прогнозів [1]. Для підвищення якості остаточних результатів існуючі математичні методи інтелектуального аналізу даних (ІАД) потребують додаткової модифікації та обробки даних, виконання їхніх певних трансформацій тощо [2]. На виконання дій, що спрямовані на підготовку даних, необхідні високий рівень математичної підготовки та поглиблене володіння спеціальними сучасними інструментальними засобами, які можуть бути використані для реалізації цих методів.

Особлива увага фінансових аналітиків приділяється розвитку і практичному застосуванню методів аналізу фінансових ризиків, зокрема ризиків кредитування, як одних із найбільш поширених у банківській системі та інших фінансових організаціях, діяльність яких пов'язана з кредитуванням фізичних та юридичних осіб [3–5]. Позитивні результати моделювання і прогнозування кредитних ризиків досягнуті при використанні сучасних інформаційних технологій і моделей та методів інтелектуального аналізу даних. До класу цих методів відносять такі: нелінійна байєсівська регресія і байєсівські мережі, метод опорних векторів, нечітка логіка, дерева рішень, нейронні мережі та комбінації вказаних методів. У результаті виконання попередніх досліджень встановлено, що прийнятні результати класифікації клієнтів на дві групи (надійні та ненадійні) можна отримати за допомогою нелінійних моделей, нечіткої логіки та байєсівських мереж у комбінації з використанням сучасних інформаційних технологій для збору, попередньої обробки, модифікації і передачі даних. Дослідження присвячене аналізу та порівнянню вибраних методів оцінювання кредитних ризиків, що ґрунтуються на комбінуванні різних методів.

### **Постановка завдання**

Мета роботи полягає у виконанні таких завдань: вибір, аналіз і застосування до фактичних даних обраних методів оцінювання кредитних ризиків; комбінування двох поширених методів з метою підвищення якості оцінок прогнозів; виконання порівняльного аналізу отриманих результатів і формулювання рекомендацій стосовно практичного застосування обраних методів.

### **Вибір методів оцінювання кредитного ризику**

Для аналізу даних і порівняння результатів обрано нейронні мережі та мережі Байєса. Ідея порівняння мереж Байєса та нейронних мереж пов'язана з інтуїтивною подібністю обох методів інтелектуального аналізу даних, а саме: встановлення певних взаємозв'язків і кількості шарів для прогнозування вихідної події/змінної або нейрону та налаштування вихідних результатів залежно від значень інших змінних. Найкраще ці мережі можуть спрацювати саме у випадку застосування їх до прогнозування фінансових процесів, оскільки фінансові процеси та фінансові дані, що з ними пов'язані, мають таку характеристичну природу та визначаються сукупністю факторів, взаємопов'язаних між собою.

Автори пропонують розширити використання раніше запропонованого інтегрованого підходу [6, 7] на нейронні мережі, оскільки очікується, що комбінація кількох методів дасть можливість отримати кращі результати оцінювання та менші похибки прогнозів порівняно з використанням окремо цих моделей, як це вдалося досягнути у попередніх дослідженнях з використанням інших методів інтелектуального аналізу даних. Передбачається застосування інтегрованого підходу, де для попереднього аналізу в якості додаткового шару нейромережі пропонується мережа Байєса. Мережі Байєса використовуються з одного боку для коректного відбору вхідних змінних і зменшення часу, необхідного на навчання нейронної мережі, а з іншого боку — це своєрідний інструмент модифікації та обробки вхідних даних, подібний до нечіткої нейронної мережі.

Як предметну область обрано дані стосовно оцінювання кредитоспроможності позичальників-фізичних осіб, отримані з реальних українських банків, які відображають часову та ситуативну необхідність інтелектуальної підтримки прийняття рішень при аналізі протікання фінансових процесів у реальному часі.

## Особливості предметної області

Сьогодні у фінансових системах накопичуються все більш розширені дані стосовно позичальників кредитів, при цьому зберігається інформація, яка, на перший погляд, ніяким чином не впливає на остаточний результат. Це можна пояснити тим, що на початковому етапі ще не відомо, що саме може впливати на поведінку позичальника, а у подальшому деякі з накопичених даних-характеристик стосовно позичальника можуть надати пояснення його діям.

Усі засоби оцінювання кредитного ризику розрізняються методами та формами оцінювання, але в якості найбільш вживаних моделей оцінювання кредитного ризику можна виділити експертні та скорингові моделі [8].

Експертні моделі будуються на основі знань і досвіду одного чи кількох експертів-фахівців в області кредитування та встановлених ними вагових коефіцієнтів і обраних характеристик позичальника. Ці моделі слід застосовувати не окремо на рівні банку, а лише як початкову модель у випадку відсутності статистики — бази даних за кредитами банку (наприклад, у новому банку). В процесі роботи банку і накопичення достатньої статистики по кредитах фінансовій установі слід перейти на автоматичну модель прийняття рішень — скорингову модель.

Відбір ефективної скорингової моделі є доволі непростим завданням. Як інструментарій останнім часом використовують скорингові карти, логістичну регресію, дерева рішень, нейронні мережі, мережі Байеса, які спрямовані на визначення кількісної оцінки можливого ризику невиконання взятих на себе зобов'язань контрагентом банку — позичальником. Скорингові моделі показали свою високу ефективність при роботі з великими масивами вхідних даних (змінних моделі), але також потребують періодичного оновлення для запобігання діяльності зловмисників і шахраїв, що можуть спробувати прорахувати «правильні» відповіді на запитання в анкеті для отримання кредиту.

Скоринговий підхід на основі статистики дефолтів за попередні періоди часу передбачає визначення ймовірності реалізації певної події на основі таких припущень:

— об'єкти, для аналізу яких пропонується використати статистичні дані, і об'єкти, на яких збираються статистичні дані, є еквівалентними (вимога еквівалентності об'єктів);

— умови, в яких пропонується використовувати статистичні дані й умови їхнього збору є еквівалентними (вимога еквівалентності умов);

— обсяги вибірок статистичних даних є достатніми, методи обробки — коректними, а джерело інформації — надійним (вимога переконливості).

Скорингові моделі також висувають певні вимоги до навчальної вибірки, на основі якої будується модель. Так, при побудові моделі має використовуватись однорідна вибірка за однаковими продуктами (споживчий, іпотечний кредит тощо) за певний період; у навчальній вибірці мають бути наявні як позитивні, так і

негативні випадки (повернення кредиту і дефолти); періодично (кожні 2 роки) навчальна вибірка та модель мають оновлюватися.

При роботі з експертними та скоринговими моделями необхідно мати на увазі, що модель, яка будується для оцінки позичальників, буде суттєво залежати від кредитної політики самої фінансової установи. Для фінансових установ, що проводять агресивну кредитну політику і видають кредити на досить лояльних умовах, межа, що встановлюється для видачі кредиту, буде значно нижчою (ймовірність повернення кредиту на рівні 0,7–0,8) порівняно з фінансовими установами, що проводять жорстку кредитну політику і видають кредити лише найнадійнішим позичальникам (наприклад, імовірність повернення кредиту яких більша за 0,9).

## Інструментальні засоби моделювання

### Нейронні мережі

*Нейронні мережі* (НМ) — це системи, які здатні до самонавчання та імітують діяльність людського мозку. Існує багато різновидів нейронних мереж: нейронні мережі зворотного розповсюдження, мережі Хопфілда, стохастичні мережі тощо [9]. Всі вони складаються з великої кількості зв'язаних між собою однотипних нейронів, які імітують нейрони головного мозку, а відрізняються за їхньою будовою, топологією зв'язків між ними та алгоритмами навчання.

Нейронні мережі зворотного розповсюдження (НМЗР) — це інструмент пошуку закономірностей, прогнозування, якісного аналізу. Свою назву вони отримали завдяки алгоритму навчання, в якому похибка розповсюджується від вихідного шару до вхідного, тобто у напрямі, протилежному розповсюдженню сигналу при нормальному функціонуванні мережі. Першим методом навчання, розробленим для НМЗР, був градієнтний метод навчання; він простий у реалізації, але має такі недоліки: характеризується повільною збіжністю та знаходить лише локальний екстремум [9].

У НМ всі вхідні і вихідні змінні представлені у вигляді чисел з плаваючою точкою у діапазоні  $[0...1]$ . У той же час дані предметної області часто мають інше представлення — числа з будь-якого діапазону, дати, символічні рядки, тобто дані можуть бути кількісними і якісними. Для представлення якісних змінних у вигляді числових можна використати такий підхід: у випадку наявності деякої якості (наприклад, наявності залогу у позичальника) кодують 1 (є залог), а у випадку відсутності — 0 (немає залогу). Іншим способом кодування є кодування двійковими вагами, але набір можливих значень може виявитися досить нерівномірним. Тут раціональнішим способом буде поставити у відповідність кожному значенню свою вагу, що відрізняється на 1 від ваги сусіднього значення. Наприклад, віку людини (до 25 років, 25–39 років, 40–49 років, 50–59 років, більше 59 років) можна ставити у відповідність різні ваги. Так, віку 50–59 років можна поставити у відповідність число 3. Таким чином, вік буде закодований числами у проміжку  $[0...4]$ .

Так само можна ставити у відповідність числа і неупорядкованим даним, але це може призвести до небажаної упорядкованості, спотворити дані і сильно ускладнити процес навчання НМ. Для розв'язання цієї проблеми можна поставити у відповідність кожному значенню один із входів НМ (за наявності значення від-

повідний йому вхід встановлюється в 1, або в 0 — у протилежному випадку). За наявності великої кількості варіантів вхідного значення кількість входів НМ значно зростає, що різко збільшує обчислювальні витрати на навчання НМ. У такому випадку кожному значенню вхідного параметра можна поставити у відповідність бінарний вектор, кожний розряд якого відповідає окремому значенню НМ. Перевагою НМ є висока точність прогнозування, а недоліком — складність налаштування та висока тривалість навчання, а також неможливість вербальної інтерпретації процесу моделювання, оскільки система функціонує як «чорний ящик».

### Інструмент аналізу даних — мережі Байєса

Мережа Байєса (МБ) подається як трійка  $N = \langle V, G, J \rangle$ , першою компонентою якої є множина змінних  $V$ ; другою — спрямований ациклічний граф  $G$ , вузли якого відповідають випадковим змінним модельованого процесу;  $J$  — спільний розподіл імовірностей змінних  $V = \{X_1, X_2, \dots, X_n\}$ . При цьому виконується така марковська умова: кожна змінна мережі не залежить від усіх інших змінних, за винятком батьківських попередників цієї змінної [8].

При побудові мереж Байєса можливі випадки, коли попередня структура мережі задана емпірично (отримана від експертів, отримана за допомогою інших методів чи алгоритмів). Якщо структура невідома, то її потрібно побудувати (розв'язати задачу навчання структури). За наявності статистичної інформації можна скористатись алгоритмом побудови мережі Байєса на основі статистичних даних, який у першому випадку скоригує структуру мережі, а в другому — побудує структуру мережі, яку потім можна буде модифікувати, використовуючи знання експертів.

Запропонована у роботі [7] інформаційна технологія передбачає отримання та обробку даних про позичальника, прогнозування його кредитоспроможності з використанням мережі Байєса і надання висновку-рекомендації щодо видачі кредиту. Відділення банку видають кредити, збираючи усю множину даних — фінансовий стан, соціально-демографічні характеристики позичальника тощо — на основі розроблених і встановлених у банку форм-анкет (кредитних заявок). Надані позичальником дані мають ретельно перевірятися менеджером кредитного відділу в процесі обробки кредитних заявок перед прийняттям рішення щодо видачі кредиту, уточнюються та вводяться в інформаційну систему банку. Також до системи вводиться інформація щодо суми кредиту, процентної ставки, дати видачі, строку кредиту. В процесі життєвого циклу кредиту (тобто протягом терміну обслуговування кредиту в банку до моменту його погашення) вноситься інформація щодо вчасності та повноти внесень щомісячної оплати кредиту, а наприкінці строку обслуговування кредиту відмічається, чи був повернутий цей кредит чи ні, тобто чи відбувся дефолт. Таким чином, банком збирається база позичальників банку, з якої у будь-який момент може бути отримана інформація для аналізу та побудови моделей. Інформація для побудови моделі має стосуватися лише тих кредитів, за якими вже відомо, чи були вони повернуті, чи ні. Вибірки слід обирати таким чином, щоб це була найновіша інформація стосовно кредитів за один і той самий проміжок часу за одних і тих самих умов.

Статистичні дані, що отримані з бази даних банку, попередньо обробляються і завантажуються до блоку аналізу даних, де на основі відомого методу будується модель. У цій інформаційній технології пропонується використати мережі Байєса для побудови оціночної моделі. В блоці побудови моделі — будується і навчається МБ, встановлюються причинно-наслідкові зв'язки між окремими змінними — характеристиками клієнта. Отримана структура мережі використовується для аналізу характеристик моделі та оцінювання кредитоспроможності позичальника. Як приклад розглянемо моделювання та оцінювання фінансової стабільності позичальників з фактичної тестової вибірки за реальними даними банку.

### Приклад моделювання

Український банк видає споживчі кредити на купівлю мобільних телефонів, побутової техніки, електроніки, меблів і збирає статистику, чи були вони повернуті. Задача полягає у тому, щоб виявити модель, яка найефективніше розрізняє добросовісних і недобросовісних клієнтів.

Для побудови моделі використана навчальна вибірка з 2000 записів, які містять інформацію про клієнтів (вік, кількість дітей, заробітна плата, місце роботи, сума кредиту тощо) та інформацію — були кредити повернуті чи ні (0 — клієнт повернув кредит, 1 — дефолт). У навчальній вибірці 75 % (1500 кредитів) були повернуті, а 25 % (тобто 500 кредитів) — не були повернуті. Побудована модель перевіряється на вибірці з 200 записів, з яких 75 % — кредити повернуті, 25 % — дефолти. Для перевіркової вибірки по кожному клієнту обчислюємо ймовірність дефолту та порівнюємо спрогнозовані значення з фактичною інформацією, чи був повернутий кредит (0 — якщо кредит повернули, 1 — в іншому випадку). Незалежно від того, яка саме модель або набір моделей використовується в інформаційній технології, результатом моделювання повинна бути ймовірність несплати (дефолту), визначена для кожного клієнта у портфелі [7].

Для аналізу кредитоспроможності позичальників кредиту на основі наведених даних побудуємо моделі оцінювання кредитоспроможності позичальника, а для аналізу їхньої якості обчислимо загальну точність кожної моделі, помилки I-го та II-го роду, а також GINI-індекс [10]. На основі отриманих результатів зробимо висновки стосовно доцільності й перспективності використання цих моделей для оцінювання кредитоспроможності позичальників.

Побудова нейронної мережі здійснювалася за допомогою аналітичної платформи Deductor. Кількість прихованих шарів — 1; за алгоритм навчання обрано Back Propagation. Час навчання нейронної мережі становив 7 хвилин 43 секунди.

Обчислення загальної точності моделі, побудованої за допомогою нейронної мережі, та помилок першого та другого роду здійснено з використанням відповідних математичних виразів. Результати виконання обчислювальних експериментів для різних порогів відсікання занесені в табл. 1.

Очевидно, що максимальна точність моделі на рівні 0,76 досягається при значенні порогу відсікання 0,2; 0,25 та 0,3. Для порогу 0,3 модель пропускає 12 дефолтів і відсіює 24 % добросовісних позичальників. Побудована ROC-крива представлена на рис. 1. Значення площі під кривою становить:  $AUC = 0,823$ , а індекс GINI відповідно:  $GINI = 2 * AUC - 1 = 0,646$ . Значення точності для цієї моделі є

досить високим, але кількість дефолтів, які пропускає модель, є більшою порівняно з логістичною регресією чи деревом рішень [8].

Таблиця 1. Загальна точність моделі та помилки I-го та II-го роду для нейронної мережі

Факт повернення чи неповернення кредиту	Прогноз: Повернення кредиту (0)	Прогноз: Дефолт (1)	Процент точності
<b>Cut-off=0,1</b>			
Факт: Повернення кредиту (0)	100	50	0,666667
Факт: Дефолт (1)	7	43	0,86
Загальна точність моделі			<b>0,715</b>
<b>Cut-off=0,15</b>			
Факт: Повернення кредиту (0)	101	49	0,673333
Факт: Дефолт (1)	8	42	0,84
Загальна точність моделі			<b>0,715</b>
<b>Cut-off=0,2</b>			
Факт: Повернення кредиту (0)	113	37	0,753333
Факт: Дефолт (1)	11	39	0,78
Загальна точність моделі			<b>0,76</b>
<b>Cut-off=0,25</b>			
Факт: Повернення кредиту (0)	113	37	0,753333
Факт: Дефолт (1)	11	39	0,78
Загальна точність моделі			<b>0,76</b>
<b>Cut-off=0,3</b>			
Факт: Повернення кредиту (0)	114	36	0,76
Факт: Дефолт (1)	12	38	0,76
Загальна точність моделі			<b>0,76</b>

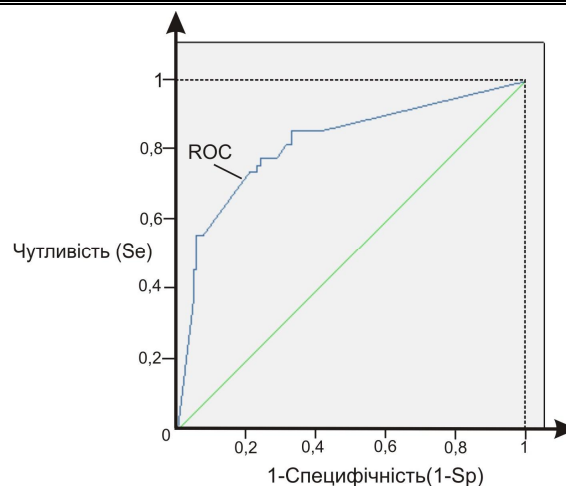


Рис. 1. ROC-крива для нейронної мережі

### Мережі Байеса

Помилки I-го та II-го роду при різних значеннях порогу відсікання для мережі Байеса наведені у табл. 2.

Таблиця 2. Загальна точність моделі та помилки I-го та II-го роду для різних рівнів порогу відсікання, отримані для мереж Байєса

Факт повернення чи неповернення кредиту	Прогноз: Повернення кредиту (0)	Прогноз: Дефолт (1)	Процент точності
<b>Cut-off=0,1</b>			
Факт: Повернення кредиту (0)	55	95	0,366666667
Факт: Дефолт (1)	3	47	0,94
Загальна точність моделі			<b>0,51</b>
<b>Cut-off=0,15</b>			
Факт: Повернення кредиту (0)	82	68	0,546667
Факт: Дефолт (1)	7	43	0,86
Загальна точність моделі			<b>0,625</b>
<b>Cut-off=0,2</b>			
Факт: Повернення кредиту (0)	98	52	0,653333333
Факт: Дефолт (1)	8	42	0,84
Загальна точність моделі			<b>0,7</b>
<b>Cut-off=0,25</b>			
Факт: Повернення кредиту (0)	100	50	0,666666667
Факт: Дефолт (1)	8	42	0,84
Загальна точність моделі			<b>0,71</b>
<b>Cut-off=0,3</b>			
Факт: Повернення кредиту (0)	110	40	0,733333333
Факт: Дефолт (1)	10	40	0,8
Загальна точність моделі			<b>0,75</b>

МБ побудована на основі навчальної вибірки у 2000 випадків. Для перевіркової вибірки (200 випадків) обчислено ймовірності дефолтів, які потім були порівняні з фактичними результатами. Найбільша точність моделі досягається на рівні 0,75 при встановленні порогу 0,3; при цьому було пропущено 10 дефолтів.

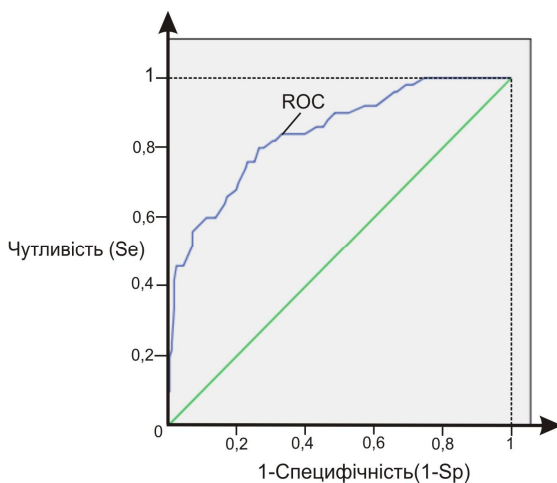


Рис. 2. ROC-крива для мережі Байєса

Очевидно, що мережа Байєса має схильність до «перестраховки», тобто вона частіше помиляється, відкидаючи тих клієнтів, які би повернули кредит. Точність моделі та кількість помилок I-го та II-го роду залежать від порогу відсікання. Встановлюючи поріг відсікання, ми визначаємо не лише відсоток відсіяних клієнтів, а й нижню границю ймовірності повернення кредиту. Значення ймовірності дефолту 0,1 або 0,2 для клієнта є досить незначними та статистично малими, а тому поріг відсікання доцільно було б встановити на рівні 0,25–0,3. Побудована для МБ ROC-крива представлена на рис. 2.

Для порівняння моделей, використаємо індекс GINI. Значення площі під кривою становить:  $AUC = 0,839$ , а  $GINI = 2 * AUC - 1 = 0,678$ .



## Комбіноване використання мереж Байєса та нейронних мереж

Для перевірки доцільності використання запропонованого інтегрованого підходу виконано ряд обчислювальних експериментів, у яких на початковому етапі застосовувалася мережа Байєса, що встановлювала причинно-наслідкові зв'язки між змінними-характеристиками та результуючою змінною і визначала ті змінні, які доцільно використовувати у подальшому для моделювання. При встановленні причинно-наслідкових зв'язків між змінними розглянуто такі типи змінних: ті, що взагалі не впливають на результат (так звана «зайва» інформація у вибірці); змінні, які не впливають на результат, але зазнають впливу сукупності характеристик, що, зокрема, впливають на ключову змінну; змінні, які опосередковано впливають на результат («батьківські» змінні характеристик, що визначають результуючу змінну) та змінні, характеристики, які безпосередньо впливають на результуючу змінну.

Розглянемо конкретний приклад. У результаті виконання обчислень було відібрано для включення у нейронну мережу змінні, що за побудованою на попередньому етапі мережею Байєса взагалі не впливають на результат (таких було дві змінні). Всі інші змінні подавалися на вхід нейронної мережі. В результаті час навчання нейромережі скоротився до 7 хвилин 28 секунд, що становить 3–4 % від часу навчання. В результаті використання інтегрованого підходу збільшилася точність моделі, зменшилася кількість помилок і зріс індекс GINI. Отримані результати обчислювальних експериментів представлені у табл. 3.

Таблиця 3. Загальна точність моделі та помилки I-го та II-го роду для нейронної мережі

Факт повернення чи неповернення кредиту	Прогноз: Повернення кредиту (0)	Прогноз: Дефолт (1)	Процент точності
<b>Cut-off=0,1</b>			
Факт: Повернення кредиту (0)	111	39	0,74
Факт: Дефолт (1)	8	42	0,84
Загальна точність моделі			<b>0,765</b>
<b>Cut-off=0,15</b>			
Факт: Повернення кредиту (0)	111	39	0,74
Факт: Дефолт (1)	9	41	0,82
Загальна точність моделі			<b>0,76</b>
<b>Cut-off=0,2</b>			
Факт: Повернення кредиту (0)	113	37	0,753333
Факт: Дефолт (1)	9	41	0,82
Загальна точність моделі			<b>0,77</b>
<b>Cut-off=0,25</b>			
Факт: Повернення кредиту (0)	114	36	0,76
Факт: Дефолт (1)	9	41	0,82
Загальна точність моделі			<b>0,775</b>
<b>Cut-off=0,3</b>			
Факт: Повернення кредиту (0)	114	36	0,76
Факт: Дефолт (1)	9	41	0,82
Загальна точність моделі			<b>0,775</b>

Побудована за результатами експериментів ROC-крива представлена на рис. 3. Значення площі під кривою становить:  $AUC = 0,842$ , а  $GINI = 2 * AUC - 1 = 0,684$ , що є вищим показником порівняно з обома методами — мережами Байєса та нейронними мережами.

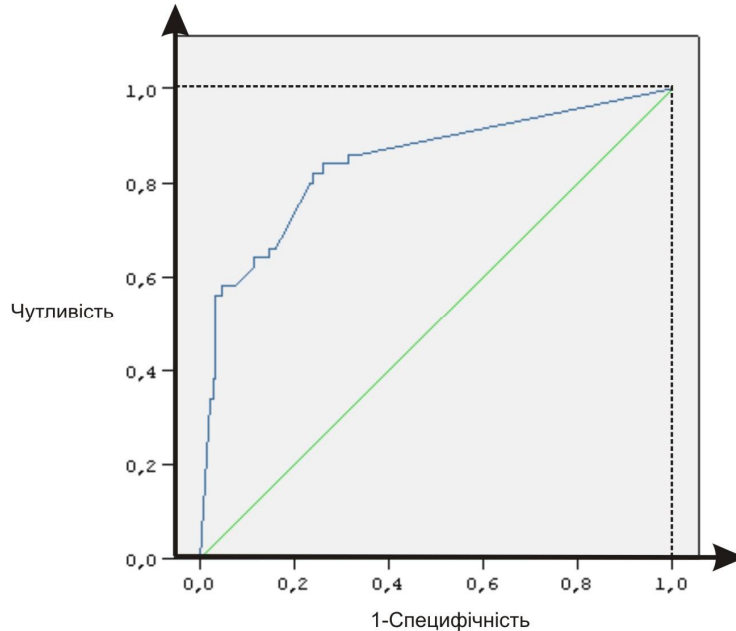


Рис. 3. ROC-крива для інтегрованої моделі на основі мережі Байєса та нейронної мережі

## Висновки

На основі фактичних статистичних даних стосовно позичальників кредитів банку побудовано математичні моделі у формі нейронних мереж зворотного розповсюдження та статичної байєсівської мережі. Виконано ряд обчислювальних експериментів стосовно прогнозування дефолту позичальників кредитів із використанням кожної побудованої моделі окремо, а також комбінованої (інтегрованої) моделі, яка складається з обох згаданих структур. Показано, що кращі результати забезпечує комбінована модель (у прикладі — із загальною точністю моделі 0,775 та індексом Джіні 0,684).

Результати виконаних експериментів з використанням інтегрованого підходу свідчать про доцільність застосування додаткових засобів інтелектуального аналізу даних для попередньої обробки вхідних даних. Як показали експериментальні дослідження, за рахунок використання мереж Байєса на попередньому етапі вдається зменшити час на обробку вхідних даних і навчання нейромережі, але найголовніше, вдається суттєво покращити результати розпізнавання і зменшити значення помилок другого роду та підняти загальну точність моделі. Крім того, суттєво збільшився індекс GINI, і в результаті модель забезпечує високу якість прогнозування ризику втрат.

При виконанні подальших досліджень доцільно створити систему підтримки прийняття рішень, розширену критеріями аналізу якості обчислень на кожному

етапі та додатковими методами аналізу даних позичальників кредитів. Згідно з попередньо виконаними дослідженнями, кількість методів прогнозування повинна бути не меншою чотирьох.

1. *Tsay R.S.* Analysis of financial time series / R.S. Tsay. — New York: John Wiley & Sons, Inc., 2010. — 715 p.
2. *Tuffery S.* Data mining and statistics for decision making / S. Tuffery. — New York: John Wiley & Sons, Inc., 2011. — 704 p.
3. *Siddiqi N.* Credit risks scorecards / N. Siddiqi. — New York: John Wiley & Sons, Inc., 2006. — 210 p.
4. *Jones S.* Advances in credit risk modeling and corporate bankruptcy prediction / S. Jones, D.A. Hensher. — Eds. — Cambridge: Cambridge University Press, 2008. — 310 p.
5. *Anderson R.* The Credit scoring toolkit / R. Anderson. — Oxford: Oxford University Press, 2007. — 790 p.
6. *Кузнєцова Н.В.* Інтегрований підхід до оцінювання кредитних ризиків / Н.В. Кузнєцова // Тр. Одес. политехн. ун-та. — Одеса, 2010. — Вып. 1(33). — 2(34). — С. 187–192.
7. *Кузнєцова Н.В.* Інформаційна технологія аналізу фінансових даних на основі інтегрованого методу / Н.В. Кузнєцова, П.І. Бідюк // Системні дослідження та інформаційні технології. — 2011. — № 1. — С. 22–33.
8. *Кузнєцова Н.В.* Системний підхід до аналізу кредитних ризиків з використанням мереж Байеса / Н. В. Кузнєцова, П. І. Бідюк // Наукові вісті НТУУ «КПІ». — 2008. — № 3. — С. 11–24.
9. *Зайченко Ю.П.* Основи проектування інтелектуальних систем / Ю.П. Зайченко. — К: Слово, 2006. — 352 с.
10. *Кузнєцова Н.В.* Порівняльний аналіз характеристик моделей оцінювання ризиків кредитування / Н.В. Кузнєцова, П.І. Бідюк // Наукові вісті НТУУ «КПІ». — 2010. — № 1. — С. 42–53.

Надійшла до редакції 27.03.2015