УДК 004.942

Informatics and Mathematical Methods in Simulation
Vol. 8 (2018), No. 1, pp. 16-25

АНАЛІЗ ТА ПРОГНОЗУВАННЯ РИЗИКІВ ШАХРАЙСТВА З КРЕДИТНИМИ КАРТКАМИ

Н.В. Кузнєцова

Київський політехнічний інститут імені Ігоря Сікорського,
пр. Перемоги, 37, Київ, 03056, Україна; e-mail: natalia-kpi@ukr.net

У роботі досліджено основні підходи виявлення шахрайських операцій з кредитними картками та запропоновано комбіноване виконання поведінкового оцінювання клієнтів-власників кредитних карт та моніторингу операцій з кредитними картами з метою виявлення можливої спроби шахрайства. Було сформовано вхідну вибірку з 4 тисяч активних кредитних карт, за якими проводилось 55 тисяч транзакцій, класифіковано основні типи кредитних карт за їх поведінкою. За результатами проведеного дослідження поведінки клієнтів кредитних карток будується поведінкова модель, що описує, як вони будуть діяти в наступні моменти часу, які операції проводити, на яку суму та з якими лімітами, і таким чином дозволяє виявляти нетипові запити, що є свідченням спроби шахрайських операцій. Поряд зі стандартними характеристиками, що описують клієнта та кредитну карту, формувалися агреговані характеристики, які описували поведінку клієнта (баланс, доходи, витрати, ліміти) в динаміці, і за рахунок їх застосування вдалось отримати більш якісні моделі. При аналізі ризику шахрайських транзакцій використовувались операції за транзакціями за попередні періоди (типові та шахрайські), будувалися моделі на основі логістичної регресії, регресійного рівняння, регресії з інтегрованим ковзним середнім, а також застосовувались багатошаровий перцептрон та радіально-базисна функція. Такі моделі здійснювали класифікацію транзакції на предмет шахрайства, а далі прогнозували суму транзакції. Найкращим методом класифікації транзакцій виявився багатошаровий перцептрон, а найвищі значення точності прогнозу показала модель авторегресії з інтегрованим ковзним середнім (ARIMA) із значенням похибки 4.46% для прогнозування суми транзакції. Запропонований комбінований підхід виявився ефективним для аналізу та моніторингу кредитних карт та виявлення шахрайських операцій. Перспективним може бути його використання для інших типів фінансових ризиків (для білінгових та платіжних систем, систем переведення грошей та інших торгівельних операцій).

Ключові слова: ризики шахрайства, поведінковий скоринг, комбінований підхід, кредитні картки, моделі оцінювання ризиків, регресійні моделі

Вступ

Ризики шахрайських операцій з кредитними картками є одними з найбільш складно модельованих, оскільки характеризуються великою кількістю особливостей як самих операцій, транзакцій, які здійснюються, так і великою кількістю додаткових параметрів, які потребують урахування. Найчастіше моделюють ризики шахрайства, за якими доступні вже певні статистичні дані; цим займаються відділи моніторингу та протидії шахрайським операціям комерційних банків, трансфертних та білінгових систем, систем сплати послуг, перерахунку валют, платіжних систем тощо.

В Україні за даними Національного банку України у 2016 році було в обігу 31,1 млн. активних банківських карток (з 70 млн. виданих). Близько 5 млн. клієнтів мають негативну кредитну історію, яка пов’язана не тільки з шахрайством, а й з банкрутством неплатоспроможних банків. Ще донедавна вся відповідальність за можливі шахрайські операції з кредитними картками повністю лежала на плечах самих власників цих кредитних карток. Проте на вимогу найбільших світових платіжних систем (Visa,

Mastercard), які мають однакові правила обслуговування всіх клієнтів платіжних систем незалежно від країни, ця відповідальність була повністю перекладена на самі українські банки. І, якщо раніше моніторинг фінансових операцій здійснювався лише з точки зору використання в рамках ліміту, валюти та часу доби, то тепер набагато ретельніше перевіряються не лише великі оперції, а й самі факти підняття ліміту.

Ризик шахрайських операцій з кредитними картками може розглядатися у декількох контекстах: підняття лімітів операцій, використання кредитних ліній з непогашеними заборгованостями, перевірка транзакцій за операціями – чи не є вони шахрайськими.

Метою даної роботи є аналіз існуючих методів та підходів до моделювання ризиків шахрайства та дослідження можливості застосування їх на реальних статистичних даних.

Для моделювання у роботі вирішувались завдання аналізу поведінки клієнтів з метою виявлення шахрайства з їх кредитними картками на прикладі українського банку, здійснювалось прогнозування шахрайської операції на основі вхідних характеристик для міжнародних кредитних карток.

Основна частина

Оцінювання ризиків банківської діяльності називають ще скорингом, оскільки при цьому розробляються скорингові моделі, за якими оцінюють власне клієнтів та кредити. Виділяють попередній скоринг, поведінковий скоринг, колекшн-скоринг, а також скоринг шахрайства [1,2].

Поведінковий скоринг (Behavioral scoring) – це динамічна оцінка стану кредитоспроможності існуючого позичальника, заснована на даних про історію операцій по його рахунках (графік погашення заборгованості, запити нових кредитів, оборот за поточними рахунками, і т. д.). Результатом поведінкового скорингу зазвичай є пропозиція банку скористатися іншими банківськими послугами: кредитна карта, кредит готівкою за зниженою процентною ставкою, автокредитування та ін. Схвалення наступних кредитів у банку для позичальника – це також результат успішного проведення поведінкового скорингу [1].

Скоринг шахрайства (Fraud scoring) – це вид скорингу, який надає статистичну оцінку ймовірності шахрайських дій з боку потенційного позичальника. Як правило, застосовується спільно з іншими видами дослідження клієнтів, такими як аплікаційний скоринг і поведінковий скоринг. Fraud-scoring використовується перш за все як певний бар'єр на шляху шахраїв при отриманні кредиту. Так, система в автоматичному режимі може порівнювати дані клієнтів з так званими «чорними» і «сірими» списками, робити запити в бюро кредитних історій і інші зовнішні бази даних. Ще одна функція fraud-scoring – перевірка наданих даних на протиріччя, причому як в рамках самої анкети, коли можуть порівнюватися відповіді на різні питання, так і через порівняння отриманих результатів з встановленими даними і статистикою (наприклад, про розміри доходів в різних галузях). Також цифри можуть порівнюватися з тими, які надали інші клієнти, з усередненими показниками по всьому портфелю кредитів банку [1].

Всі види скорингу обов’язково присутні у всіх банках на момент перевірки кредитних заявок, проте подальше дослідження і перевірка клієнтів, які вже користуються кредитними продуктами банку, і є доволі корисним, але було не дуже розповсюдженим в українських банках. Зі зміною вимог Національного банку України до нормативів капіталу банку і забезпеченості кредитів ситуація дещо змінилась. Тепер все частіше банки перевіряють позичальників не лише на етапі обробки кредитних заявок, а й здійснюють відслідковування транзакцій з метою недопущення проведення шахрайських операцій, а також здійснюють поведінковий скоринг з метою забезпечити

лояльність клієнтів, вчасно збільшивши кредитні ліміти, або навпаки, заблокувавши кредитні карти, по яким проводяться підозрілі маніпуляції.

Для цього здійснюється відслідковування і класифікація позичальників за їх діями протягом певного періоду з кредитними лімітами, платежами, надходженнями тощо. Транзакція вважається шахрайською, якщо вона відрізняється від звичайної поведінки користувача. Це пов'язано з тим, що передбачається, що зловмисники будуть вести себе зовсім по-іншому, ніж власник облікового запису. Отже, спочатку необхідно напрацювати і виявити модель поведінки користувача кредитної карти, а потім вже виявляти шахрайство [3]. Іншим варіантом є класифікація самих транзакцій з метою виявлення, чи є вони нормальними (типовими), чи є шахрайськими. Для вирішення цієї задачі можуть використовуватись методи дерев рішень, нейронні мережі, правила виведення тощо [4,5].

Важливо підкреслити ключові відмінності між аналізом поведінки користувачів та підходами щодо аналізу шахрайства. Метод аналізу шахрайства дозволяє виявити відомі шахрайські хитрощі з низьким хибним позитивним рівнем. Проте, оскільки доля випадків шахрайських заявок є зазвичай доволі низькою, то сам класифікатор базується на обмежених записах про шахрайство і не може виявити нові фальсифікації. В результаті, пропуски шахрайства при класифікації можуть бути надзвичайно високими залежно від того, наскільки винахідливими є шахраї. Аналіз поведінки користувачів в цьому випадку справляється набагато краще з виявленням нових фальсифікацій. Будь-яка діяльність, яка суттєво відрізняється від моделі, буде розглядатися як можливе шахрайство.

Системи виявлення шахрайства також стикаються з певними труднощами та обмеженнями: незбалансовані дані, різне значення неправильної класифікації, відсутність адаптивності, тощо [3]. Незбалансованість даних пов’язана з невеликою кількістю шахрайських операцій в загальній вибірці. Різна вартість помилкового виявлення пов’язана з втратами. Так, шахрайська операція сприйнята як нормальна призведе до втрат, рівних сумі транзакції. Проте, якщо система сприймає нормальну транзакцію як шахрайську операцію і відхиляє її, то це призводить до незначних втрат, пов’язаних, наприклад, з уточненням і додатковою аутентифікацією клієнта. Відсутність адаптивності передбачає, що класифікаційні алгоритми, як правило, стикаються з проблемою появи та, відповідно, складнощами виявлення нових типів звичайних чи шахрайських зразків.

Далі ми виконаємо аналіз кредитних карток та порівняємо обидва підходи з точки зору поведінкового скорингу та виявлення шахрайства.

Аналіз поведінки клієнтів-власників кредитних карток з точки зору використання лімітів. Для аналізу поведінки клієнтів були використані реальні дані українського банку за кредитними картами, виданими у 2013-2015 роках. Були відібрані спостереження за 4 тисячами активних кредитних карт, які проводили різні транзакції. Загальна кількість транзакцій складає 55 тисяч записів. Наявні такі поведінкові параметри: залишок за тілом кредиту, залишок за комісією/відсотками, прострочена заборгованість за тілом, прострочена заборгованість за комісією/відсотками, кількість днів прострочки за тілом, тощо. Аплікаційні дані: тип клієнта, вік, рік видачі, тривалість угоди, ліміт на початку, сума при подачі заявки, ознака рівності ліміту та суми, адреса прописки, адреса місця постійного проживання, кількість дітей, і т.п. Агреговані дані: максимальна кількість прострочених місяців, середнє значення щомісячного зняття, максимальне значення відношення простроченої заборгованості до встановленого ліміту, і т.п.

У результаті попереднього аналізу можна виділити декілька типових кредитних карт: кредитні карти, що безперебійно погашають заборгованості та картки, що перейшли в дефолт (за міжнародною практикою, коли кредитна заборгованість перевищує 90 днів). Для більшості недефолтних кредитних карт все ж спостерігалися

прострочені платежі протягом невеликого періоду в межах 1-7 днів. Тут можна виділити додаткову класифікацію: з постійним лімітом, зменшенням та збільшенням ліміту. Спільною рисою для цих трьох випадків є відносно постійний темп платежів, що залежить від величини ліміту та/або заборгованості. Разом з тим зняття з кредитної карти є певною випадковою величиною із очевидним математичним сподіванням, а структура в послідовності прострочених платежів взагалі не спостерігається. Також серед користувачів можна виділити тих, що залишали баланс кредитних карт на високому рівні та тих, що намагалися звести баланс до нуля якнайшвидше. Приклад кредитної карти без заборгованостей, проте зі збільшенням кредитного ліміту наведено на рис. 1.

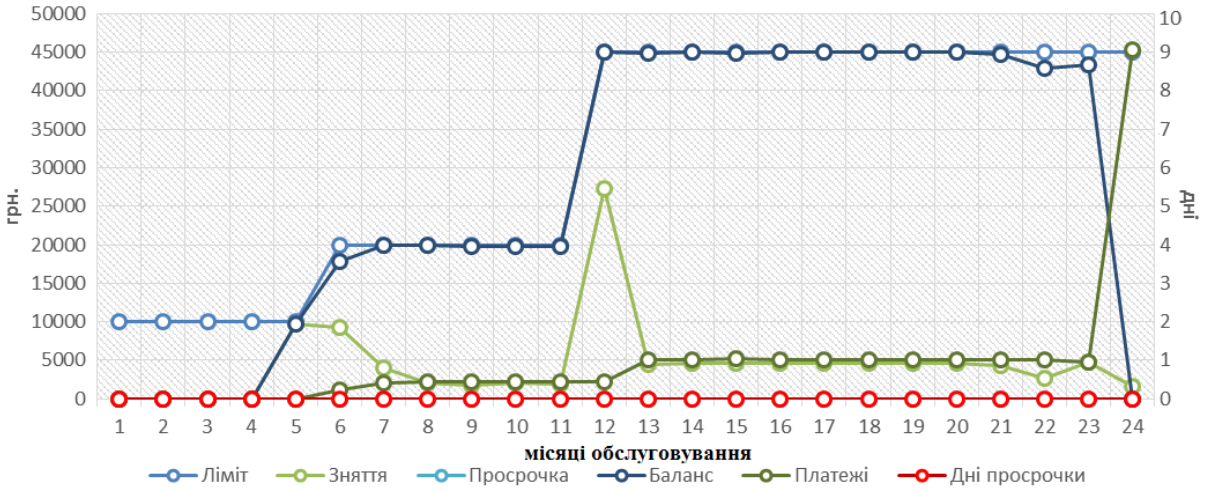

Рис. 1. Приклад кредитної карти без прострочених платежів (зі збільшення ліміту)

Приклад карти другого типу зі зведенням до мінімуму балансу наведено на рис. 2. Окремо можна виділити випадки простроченої комісії при нульовому балансі на карті, які не вважаються дефолтними, проте також є доволі типовими для клієнтів.

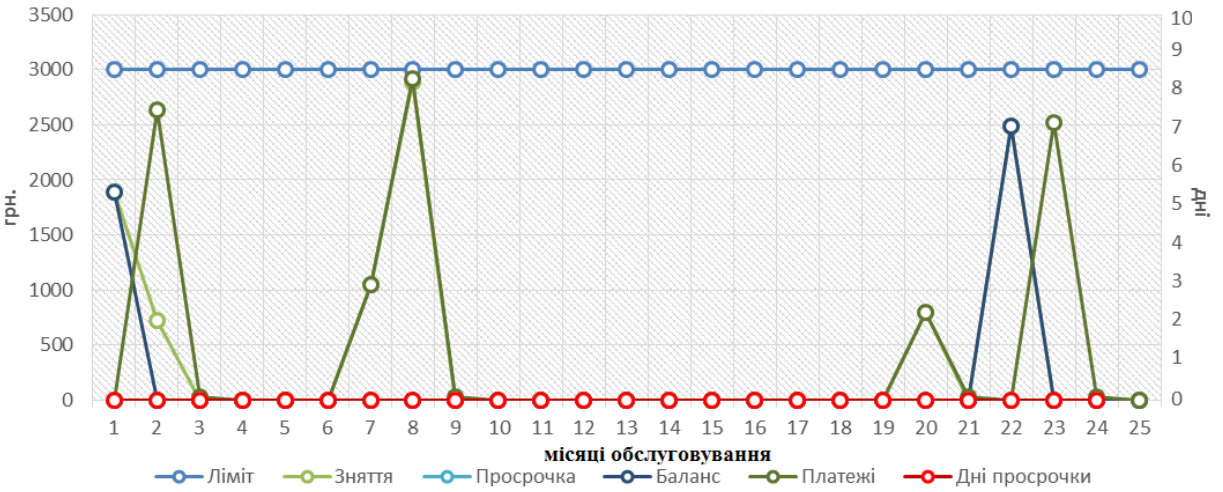

Рис. 2. Кредитна карта другого типу

Наступним варіантом є так звані дефолтні картки, які мають заборгованість більш ніж 3 місяці, і можуть характеризуватися постійним, збільшеним або навіть зменшеним кредитним лімітом. Для таких клієнтів необхідно виявити факт дефолту і заблокувати можливість використання кредитного ліміту в наступному місяці, щоб зменшити фінансові втрати банку.

Згідно з описаним критерієм кредитна карта на рис. 3 переходить у стан дефолту на 21-му місяці свого життя, при цьому довгий період власник карти був взірцем, дотримувався графіку виплат досить тривалий час.

Однак бувають випадки, коли кредитна карта переходить у заборгованість і прострочку, проте потім клієнт успішно покриває заборгованість та закриває карту.

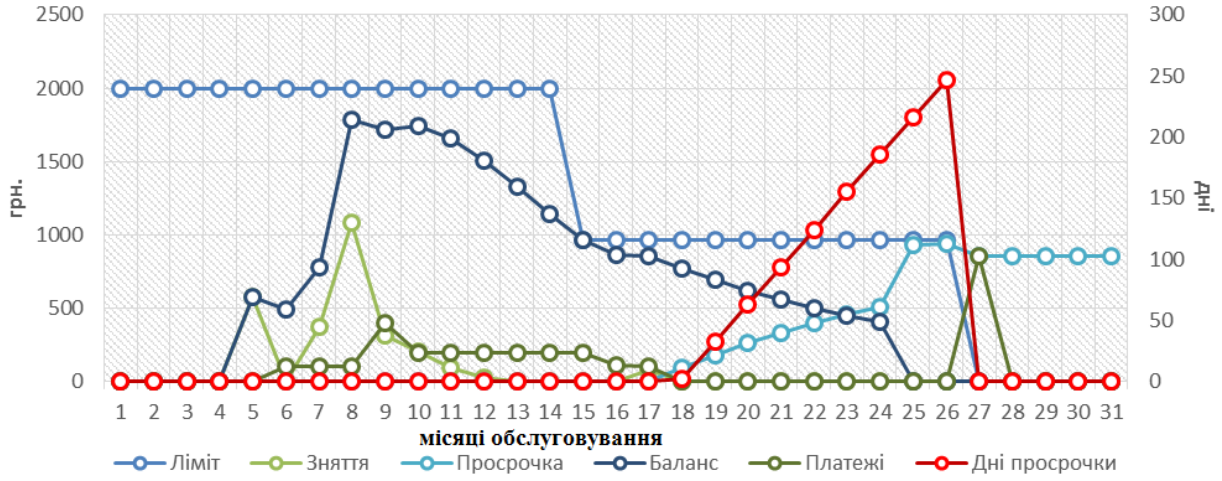

Рис. 3. Приклад дефолтної кредитної карти зі зменшення ліміту

За такими критеріями були відібрані і сформовані види поведінки клієнтів, і тепер можна безпосередньо виявляти атипову або шахрайську поведінку, яка суттєво відрізняється від описаної вище. Перейдемо до наступного етапу виявлення шахрайських операцій за кредитними картами.

Прогнозування шахрайської операції за кредитними картками

Вхідні дані: Набір даних містить єврооблігаційні транзакції, зроблені кредитними картками у вересні 2013 року. Набір даних дуже незбалансований, шахрайство становить 0.172% всіх транзакцій. Набір даних містить тільки числові вхідні змінні. Змінні V1, V2, ... , V29 є основними компонентами, отриманими за допомогою методу головних компонент, а єдиною неперетвореною функцією є «Час». Функція «Час» містить секунди, що минули між кожною транзакцією і першої транзакцією в наборі даних. Функція «Сума» – це сума транзакції, яка виступає в якості цільової змінної, наприклад, для навчання з урахуванням витрат. Функція «Клас» – є цільовою змінною відповіді, і вона приймає значення «1» в разі шахрайства, «0» – в іншому випадку.

1 Етап. Задача виявлення хибної транзакції є задачею класифікації, а тому доцільно перевіряти точність за основі індексу GINI та області під кривою точності (AUC). Для того, щоб визначити важливі змінні, які мають бути включені до аналізу, здійснюємо попередній аналіз змінних-характеристик [1]. В ідеалі ми маємо сформувати вибірку таким чином, щоб отримати збалансовані розподіли за кожною змінною класу. Графіки на рис. 4 показують, наскільки різнорідні обидва класи щодо кожного розподілу змінних. Тільки V13, V15, V22, V23, V24, V25 і V26 можна розглядати як беззмістовні змінні, а всі інші змінні – потенційно корисні змінні. Проте, на початковому етапі використаємо всі змінні для побудови класифікатора.

Для цього побудуємо логістичну регресію [6] в середовищі R як класифікатор з двома виходами (0 та 1) за допомогою вбудованої функції evalmod із пакету precrec. Кількісну інтерпретацію ROC дає показник AUC − площа, обмежена ROC-кривою і віссю частки помилкових позитивних класифікацій. Чим вище показник AUC, тим якісніше класифікатор. В нашому випадку значення дорівнює 0.7589, тому дані та такий класифікатор можна використовувати для прогнозування суми по транзакціях.

2 Етап. Прогнозування суми транзакції з метою виявлення шахрайської операції.

Моделювання здійснювалось за допомогою пакету SPSS Statistic, налаштовуючи вихідну змінну (цільову або Target) як змінну V30 – загальна сума по транзакціям.

Прогнозування суми транзакції було виконано різними методами за допомогою регресійної моделі, нейронних мереж, тощо.

Було побудовано кілька різних моделей авторегресії, авторегресії з ковзним середнім, інтегрованої авторегресії з ковзним середнім. Для кожної з обраних моделей були побудовані статистичні критерії якості та таблиці залишків. Найкращі значення показала модель ARIMA ($R^2 = 0.987$, $BIC = 6.343$).

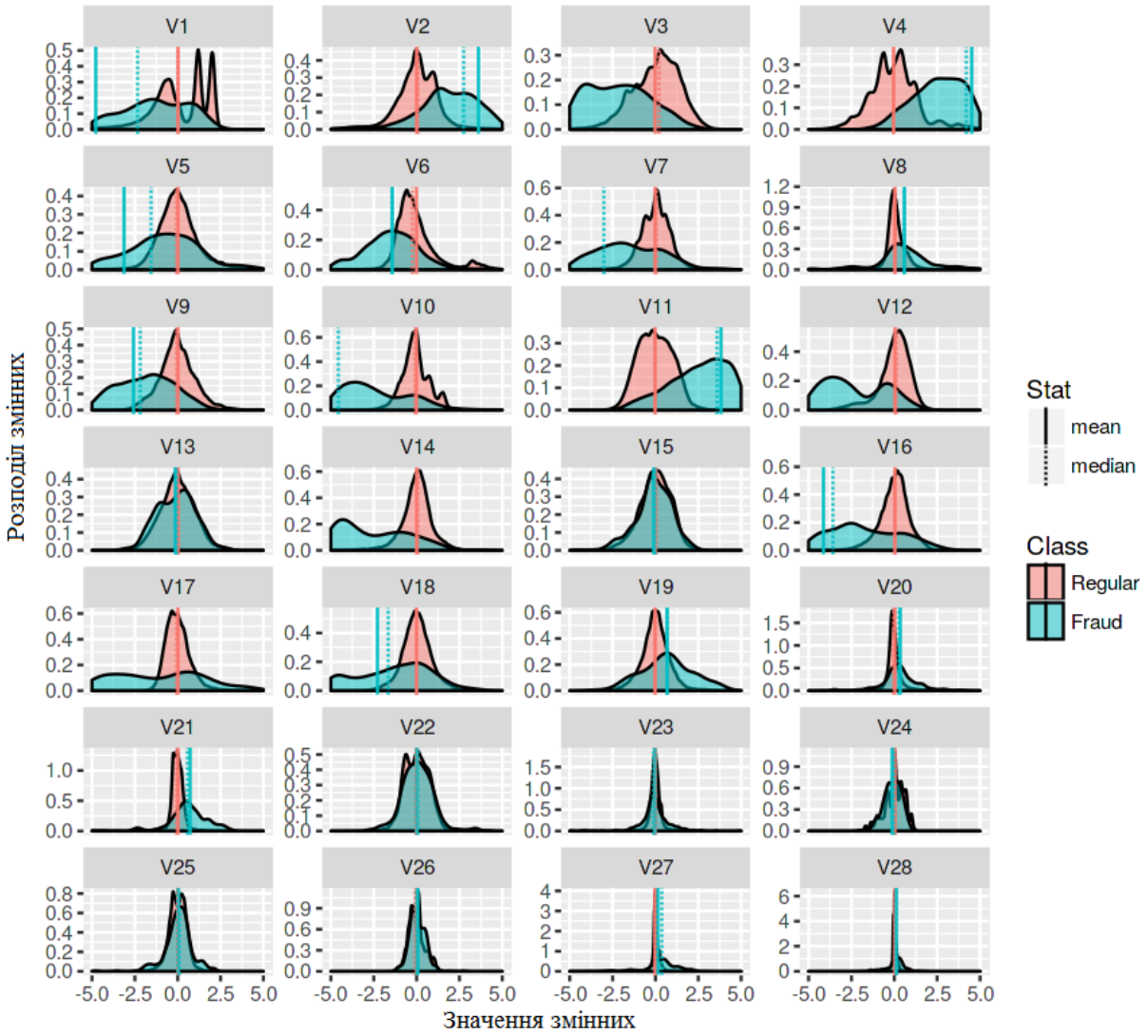

Рис. 4. Попередній аналіз вхідних характеристик

Для оцінювання якості побудови моделі та прогнозу застосовують критерії[7]:

1. Середня відсоткова похибка: $MPE = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - \hat{y}_i)}{y_i} \times 100\%$.
2. Середня відсоткова абсолютна похибка: $MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{|y_i|} \times 100\%$.
3. Середня похибка: $ME = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$.

4. Середня абсолютна похибка: $MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i |$.

5. Середньоквадратична похибка (Standard Error або RMSE – root mean square error): $SE = \sqrt{\frac{1}{N}\sum_{i=1}^{T}(y_i - \hat{y}_i)^2}$.

6. Середньоквадратична похибка (MSE – mean square error): $MSE = E\left((y - \hat{y})^2\right) = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}$.

7. Сума квадратів похибок: $\sum_{k=1}^{N} e^2(k) = \sum_{k=1}^{N} [\hat{y}(k) - y(k)]^2 \to \min_{\hat{\theta}}$.

Для моделі ARIMA були отримані наступні значення: $MAPE = 4.4469\%,\ MAE = 2.81,\ MSE = 7.89,\ SE = 62.24$.

Далі здійснювалась побудова регресійної моделі і були отримані такі показники якості моделі отримали такі: $R^2 = 0.996,\ R_{adj}{}^2 = 0.991$.

Надалі було здійснено побудову багатошарового перцептрону, де користувач може самостійно налаштувати мережу, вказавши кількість нейронів, прихованих шарів, тощо, для досягнення вищого результату (рис. 5).

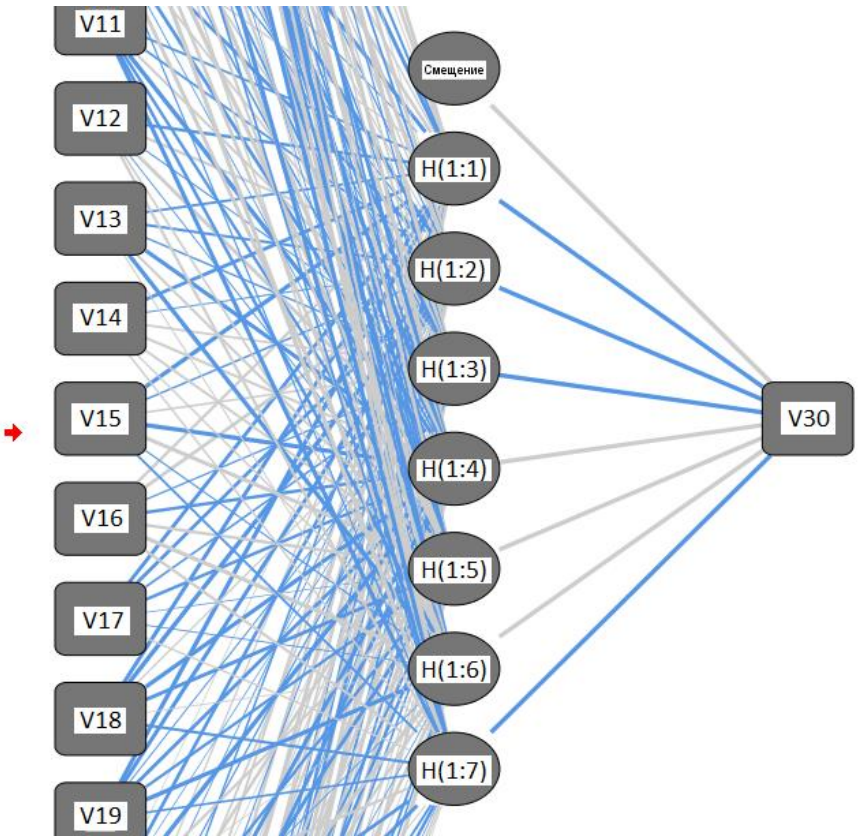

Рис. 5. Нейронна мережа із цільовою змінною

Вибірку було розділено на навчальну та перевірочну, здійснено оцінку суми квадратів похибок та відносної похибки, $ME = 0.06$ та $RMSE = 5.151$.

Також було здійснено моделювання на основі радіально-базисної функції (РБФ). Для налаштування функції було визначено 2 вхідних нейрона, функція активація для

прихованого шару Softmax, кількість нейронів, параметри моделі та прогнозу, додаткові графіки, тощо.

Зведені характеристики якості прогнозу на навчальній та перевірочній вибірці наведено у таблиці 1.

Таблиця 1.

Зведена таблиця похибок прогнозу за всіма моделями

Відносна похибка	Навчальна вибірка	Перевірочна вібірка
ARIMA	0.987	0.044
Регресійне рівняння	0.1757	0.991
Багатошаровий перцептрон	0.148	0.060
Радіально-базисна функція	0.439	0.931

Таким чином, найкращою моделлю виявився багатошаровий перцептрон, за ним був побудований прогноз з доволі якісними показниками. Найкращою за значенням прогнозу виявилась модель ARIMA із значення похибки 4.46%. Найгірше значення було отримано при застосуванні регресійної моделі.

Можна сказати, що якщо використовувати багатошаровий перцептрон або модель ARIMA, то в нас буде найвище, найточніше значення прогнозу, тобто оптимістичний сценарій для 20% перевірочної вибірки. Якщо ми будемо використовувати РБФ, то отримаємо базовий сценарій для результативних значень. Якщо ми будемо використовувати регресійну модель, то отримаємо значення прогнозу, які дуже сильно, з великою похибкою, будуть відрізнятися від реальних. В даному випадку ми будемо мати песимістичний прогноз і великий ризик того, що частина грошей буде загублена або отримана шахраями.

Повертаючись до початкових розподілених класів (де 0 – все добре, 1 – шахрайство), при застосуванні ARIMA та багатошарового перцептрону ймовірність шахрайства буде мінімальною, при РБФ – середньою, а при регресії – найбільшою.

Висновки

Задача виявлення шахрайства є однією з найактуальніших у банківському та фінансовому секторі. Саме ризики шахрайства важко структурувати та формалізувати, а тому потрібні нові комбіновані підходи та методи. У роботі використано системний підхід до опрацювання ризиків шахрайства з кредитними картками, який передбачає комбінацію поведінкового оцінювання та виявлення шахрайства за транзакціями. Такий підхід із застосуванням аналізу поведінки власника кредитної карти дозволяє виявити спробу шахрайської операції ще на етапі її виникнення, а подальше застосування методів регресійного аналізу, нейронних мереж дозволяє спрогнозувати ймовірність такої операції. Це дозволяє також збалансувати вхідні дані, отримавши агреговані дані по заборгованостям, часу обслуговування, кількості днів прострочки кредитного ліміту, тощо, та отримати більш точні оцінки прогнозу. Наступним етапом для прогнозування ризиків шахрайства є прогнозування суми операції, оскільки всі засоби запобігання шахрайству також розподілені в залежності від суми операції, а прогнозування шахрайської операції на суму 1 грн. та запобігання їй є недоцільним. Проте сам факт виявлення такої шахрайської операції є важливим сигналом про можливу спробу подальших транзакцій та є сигналом для відслідковування таких кредитних карт.

Перспективним є подальше дослідження ризиків шахрайських операцій та розширення класу задач прогнозування, з використанням комбінації поведінкового скорингу та моніторингу шахрайських операцій, на торгівельні операції, платіжні

системи, системи перерахування грошей з точки зору виявлення типових та нетипових операцій та блокування можливих шахрайських дій.

Список літератури

1. Siddiqi, N. Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring / N. Siddiqi // John Wiley & Sons, Hoboken. — 2005. — 208 p.
2. Kuznietsova, N.V. Scoring Technology for Risk Assessment of Fraud in Banking / Selected Papers of the XVI International Scientific and Practical Conference "Information Technologies and Security" (ITS 2016). — 2016. — Pp. 54-61 .
3. Sorournejad, S. A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective / S. Sorournejad, Z. Zojaji, R.E. Atani, Amir Hassan Monadjemi / Cornel University Library, 2016. Mode of access: https://arxiv.org/ftp/arxiv/papers/1611/1611.06439.pdf .
4. Воронцов, К. Математические методы обучения по прецедентам. Курс лекций / К. Воронцов. — М: МФТИ, 2006. — 58 с.
5. Чубукова, И. А. Data Mining / Чубукова И. А. — М.: Бином ЛБЗ, 2008. — 384 с.
6. Beninel, F. Transfer Learning Using Logistic Regression in Credit Scoring / F. Beninel, W. Bouaguel, G. Belmufti / Cornel University Library, 2012. Mode of access: https://arxiv.org/pdf/1212.6167.pdf.
7. Бідюк, П.І. Аналіз часових рядів: навч. Посіб / П.І. Бідюк, В.Д. Романенко, О.Л. Тимощук. — К.: НТУУ «КПІ», 2013. — 600 с.

АНАЛИЗ И ПРОГНОЗИРОВАНИЕ РИСКОВ МОШЕННИЧЕСТВА С КРЕДИТНЫМИ КАРТОЧКАМИ

Н.В. Кузнецова

Киевский политехнический институт имени Игоря Сикорского,
пр. Победы, 37, Киев, 03056, Украина; e-mail: natalia-kpi@ukr.net

В работе исследованы основные подходы к выявлению мошеннических операций с кредитными карточками и предложено комбинированное исполнение поведенческого оценивания клиентов-владельцев кредитных карт и мониторинга операций по кредитным картам с целью выявления возможной попытки мошенничества. Была сформирована входная выборка из 4 тысяч активных кредитных карт, по которым проводилось 55 тысяч транзакций, класифицированы основные типы кредитных карт по их поведению. По результатам проведенного исследования поведения клиентов кредитных карточек строится поведенческая модель, описывающая, как они будут действовать в последующие моменты времени, какие операции проводить, на какую сумму и с какими лимитами, и таким образом позволяющая выявлять нетипичные запросы, которые свидетельствуют о попытках мошеннических операций. Наряду со стандартными характеристиками, описывающими клиента и кредитную карту, формировались агрегированные характеристики, описывающие поведение клиента (баланс, доходы, расходы, лимиты) в динамике, и за счет их использования удавалось получить более качественные модели. При анализе риска мошеннических транзакций использовались операции по транзакциям за предыдущие периоды (типичные и мошеннические), строились модели на основе логистической регрессии, регрессионного уравнения, регрессии с интегрированным скользящим средним, а также применялись многослойный персептрон и радиально-базисная функция. Такие модели осуществляли классификацию транзакции на предмет мошенничества, а дальше прогнозировали сумму транзакции. Лучшим методом классификации транзакций оказался многослойный персептрон, а высокие значения точности прогноза показала модель авторегрессии с интегрированным скользящим средним (ARIMA) со значением погрешности 4.46% для прогнозирования суммы транзакции. Предложенный комбинированный подход оказался эффективным для анализа и мониторинга кредитных карт и выявления мошеннических операций, и перспективным может быть его использование для других типов финансовых рисков (для биллинговых и платежных систем, систем перевода денег и других торговых операций).

Ключевые слова: риски мошенничества, поведенческий скоринг, комбинированный подход, кредитные карточки, модели оценки рисков, регрессионные модели

ANALYSIS AND FORECASTING OF CREDIT CARDS' FRAUD RISKS

N.V. Kuznietsova

Igor Sikorsky Kyiv Polytechnic Institute,
37, Peremohy Ave., Kyiv, 03056, Ukraine; e-mail: natalia-kpi@ukr.net

The paper investigates the main approaches for detecting fraudulent operations with credit cards and proposes a combined implementation of behavioral assessment of credit card holders and credit card transactions in order to detect possible fraud attempts. An incoming sample of 4,000 active credit cards was generated, for which 55 thousand transactions were carried out, the main types of credit cards were distinguished by their behavior. According to the results of a study on the behavior of credit card clients, a behavioral model was built. It describes how they will act at the next moments, what transactions carry out, of what amount and with what limits, and thereby non-typical queries could be identified that are evidence of fraudulent attempts. Along with the standard characteristics describing the client and credit card, aggregated characteristics were formed for client's behavior describing (balance, income, expenses, limits) in dynamics and due to this managed to obtain more qualitative models. In analyzing the risk of fraudulent transactions, transaction transactions for previous periods (typical and fraudulent) were used, models based on logistic regression, regression equation, regression with integrated slip medium, and also a multi-layer perceptron and radial-basic function were used. Such models carried out the classification of the transaction for fraud, and then predicted the amount of the transaction. The best method of classification of transactions was a multilayer perceptron, and the highest accuracy values of the forecast showed autoregression integrated moving average (ARIMA) model with an error value of 4.46% for forecasting transaction amount. The proposed combined approach proved to be quite effective in analyzing and monitoring of credit cards and detecting fraudulent transactions, and it could be implemented on other types of financial risks (for billing and payment systems, money transfer systems and other trading operations).

Keywords: fraud risks, behavioral scoring, credit cards, combined approach, risk assessment models, regression models