

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Острозька академія»
Економічний факультет
Кафедра економіко-математичного моделювання та інформаційних
технологій

Кваліфікаційна робота

на здобуття освітнього ступеня магістра

на тему: **«Моделювання кредитоспроможності
позичальників комерційного банку»**

Виконала: студентка 2 курсу, групи МЕК-61
спеціальності 051 «Економіка»
освітньо-професійної програми
«Економічна кібернетика»
Другого (магістерського) рівня вищої освіти
Корева Маргарита Василівна

Керівник – доктор економічних наук, професор
Матвійчук Андрій Вікторович
Рецензент – доктор економічних наук, професор
Аверкина Марина Федорівна

"РОБОТА ДОПУЩЕНА ДО ЗАХИСТУ"

**Завідувач кафедри економіко-математичного моделювання
та інформаційних технологій _____ (проф. Власюк А. П.)**
(підпис)

Протокол № ____ від « ____ » _____ 2020 р.

Острог, 2020

Міністерство освіти і науки
Національний університет «Острозька академія»

Економічний факультет
Кафедра економіко-математичного моделювання та інформаційних технологій
Спеціальність 051 «Економіка»
Освітньо-професійна програма «Економічна кібернетика»

ЗАТВЕРДЖУЮ
Завідувач кафедри економіко-
математичного моделювання
та інформаційних технологій

проф., д.т.н. Власюк А.П.
“20” січня 2020 року

**З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТА**

Кореви Маргарити Василівни

(прізвище, ім'я, по батькові)

1. Тема роботи Моделювання кредитоспроможності позичальників
комерційного банку

керівник роботи Матвійчук Андрій Вікторович, доктор економічних наук,
професор,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджено наказом Національного Університету «Острозька академія» від
“20” грудня 2019 року № 501-ст

2. Строк подання студентом роботи 07.12.2020 року

3. Вихідні дані до роботи _____

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно
розробити) _____

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Розділ 1	Матвійчук Андрій Вікторович, д.е.н., проф.	02.03.2020	02.03.2020
Розділ 2	Матвійчук Андрій Вікторович, д.е.н., проф.	20.04.2020	20.04.2020
Розділ 3	Матвійчук Андрій Вікторович, д.е.н., проф.	25.05.2020	25.05.2020

7. Дата видачі завдання 17.01.2020

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломної роботи	Строк виконання етапів роботи	Примітка
1.	Вивчення літератури	До 30.01.2020	
2.	Робробка змісту (плану)	До 19.02.2020	
3.	Ознайомлення керівника із текстом дипломної роботи (чорновий варіант):		
3.1	Розділ 1	До 16.03.2020	
3.2	Розділ 2	До 11.05.2020	
3.3	Розділ 3	До 07.09.2020	
4.	Ознайомлення керівника із текстом дипломної роботи із врахуванням зауважень	До 26.10.2020	
5.	Попередній захист дипломної роботи	До 23.11.2020	
6.	Рецензування дипломної роботи	До 11.12.2020	
7.	Здача роботи на кафедрі Реєстрація на Moodle	До 07.12.2020	

Студент _____ **Корева М.В.** _____
(підпис) (прізвище та ініціали)

Керівник роботи _____ **Матвійчук А.В.** _____
(підпис) (прізвище та ініціали)

АНОТАЦІЯ
кваліфікаційної роботи
на здобуття освітньо-кваліфікаційного рівня магістра

Тема: *Моделювання кредитоспроможності позичальників комерційного банку*

Автор: *Корева М.В.*

Науковий керівник: *Матвійчук А.В.*

Захищена “.....”.....202...року.

Короткий зміст праці: Робота присвячена дослідженню моделювання кредитоспроможності позичальників комерційного банку. Структура роботи включає у себе вступ, три розділи, висновок, список використаних джерел та додатки.

У першому розділі оцінено сучасний показників кредитної діяльності банків України за останні 7 років; систематизовано поняття та сутність «кредитоспроможності»; виокремлено метод кредитного скорингу, як основного для побудови майбутніх моделей, подано характеристику та порівняльний аналіз основних методів та моделей оцінки кредитоспроможності позичальників комерційних банків.

У другому розділі розкрито теоретичні аспекти основ кредитування позичальників; розкрито алгоритм побудови моделей за допомогою логістичної регресії, дерев рішень, кластеризації на основі k- та G-means, карт Кохонена на основі SOM-алгоритмів; побудовано моделі оцінки кредитоспроможності фізичних осіб за допомогою логістичної регресії, побудовано моделі оцінювання кредитного ризику за технологією машинного навчання. та обрано модель із кращою точністю прогнозів.

У третьому розділі подано аналіз показників фінансового стану підприємства; відібрано показники та побудови моделі; розкрито сутність дискримінантного аналізу; побудовано модель оцінки кредитоспроможності юридичних осіб.

Корева М.В.

Зміст

ВСТУП	6
РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ МОДЕЛЮВАННЯ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКІВ БАНКУ	10
1.1. Економічна характеристика та аналіз фактичного стану показників кредитної діяльності банків України за 2013-2019рр.....	10
1.2. Поняття, механізм та необхідність оцінки кредитоспроможності позичальника.	25
1.3. Характеристика та порівняльний аналіз основних методів оцінювання кредитоспроможності позичальника.....	34
РОЗДІЛ 2. ПОБУДОВА МОДЕЛІ ОЦІНКИ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКА-ФІЗИЧНОЇ ОСОБИ	48
2.1. Алгоритм побудови скорингових моделей.....	48
2.2. Побудова логістичної регресії та скорингової карти.	55
2.3. Побудова моделей оцінювання кредитного ризику за технологією машинного навчання.....	70
РОЗДІЛ 3. ПОБУДОВА МОДЕЛІ ОЦІНКИ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКА-ЮРИДИЧНОЇ ОСОБИ.....	82
3.1. Аналіз показників фінансового стану підприємства та підбір коефіцієнтів для побудови моделі.	82
3.2. Розробка методики проведення розрахунків.....	88
3.3. Побудова дискримінантної функції та оцінка якості моделі.....	94
ВИСНОВКИ.....	102
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	105
ДОДАТКИ.....	111

ВСТУП

Кредитування як зовнішнє джерело фінансування поточних витрат, проектів та бізнес-ідей використовується давно, досить широко та є звичною справою, що вже вкоренилася у повсякденне життя, як звичайних людей, так і великих підприємств.

Незважаючи на широку розповсюдження кредитування та його висвітлення у економічній літературі, досі лишається безліч невирішених проблем, що пов'язані із розробкою ефективних методик оцінки кредитоспроможності позичальника та відбору показників для їх оцінки, простроченої заборгованості та значних процентних ставок. Варто відзначити високі ризики по кредитах через зростання проблемних кредитів.

Кредитування фізичних та юридичних осіб – одна із найбільш затребуваних послуг у сфері банківського кредитування. Залучення позичкових коштів дозволяє комерційним організаціям реалізовувати нові проекти, без використання засобів із обороту, збільшувати капітал, розширювати масштаби діяльності.

На сьогодні актуальність кредитування юридичних та фізичних осіб значно зросла. Таким чином, актуальність даної дипломної роботи можна визначити наступним чином. Кредитування слугує інструментом для встановлення довготривалих відносин між банком та клієнтом. Розвиток цих відносин та комплексність при наданні послуг дозволяють скоротити ризик коливань залишків на рахунках клієнтів банку, зроблять їх більш прогнозованими та спланованими і, таким чином, знизять кредитні ризики банку. Тому покращення кредитної політики комерційного банку є необхідною умовою його успішної діяльності.

Майже всі компанії рано чи пізно беруть кредит у банку для підтримки та розвитку свого бізнесу. Безліч фізичних осіб також звертаються до банків по кредит, коли їм не вистачає власним коштів (наприклад, на покупку квартири, машини або ж відкриття власного бізнесу). У свою чергу, банки постійно

посилюють правила оцінки кредитоспроможності позичальників, щоб підвищити ймовірність повернення наданих у кредит коштів.

Усе це обумовлює необхідність серйозного наукового дослідження вказаних проблем. Пошук оптимальних форм кредитування має здійснювати як спеціаліст банку, так і експерт та аналітик. Дані форми мають стимулювати залучення у процес кредитування все більше нових позичальників, зробити кредит доступним засобом підвищення добробуту кожного, при цьому зменшивши ризику комерційних банків. Кредитування для комерційних банків є одним із значних сегментів послуг, що приносить стабільно високий прибуток.

Ішина І.В., Сазонова М.Н. у статті «Скоринг – модель оцінки кредитного ризику» [1] розглядають можливість застосування скорингових систем із метою мінімізації ризику у кредитних організаціях, де обслуговується велика кількість клієнтів, а також його застосування на початковій стадії оцінки юридичних осіб. Автори розглядають можливість збору інформації для скорингової оцінки через анкети. Різний набір питань у анкеті дозволить для заздалегідь заданого рівня ризику встановити відповідний ліміт кредитоспроможності позичальника. Було виявлено, що кожен банк має свої методи для оцінки кредитоспроможності клієнта.

Мета даного дослідження полягає у побудові моделей для оцінки кредитоспроможності позичальників комерційного банку на базі економіко-математичного моделювання та визначення теоретичних аспектів проблематики.

Для досягнення поставленої мети потрібно виконати ряд завдань:

- визначити сутність кредитоспроможності, кредитного скорингу та дискримінантної функції;
- дослідити та порівняти існуючі методи та моделі оцінки кредитоспроможності позичальників;
- провести аналіз макроекономічних показників сучасного стану кредитної діяльності на прикладі комерційних банків України;
- дослідити механізм впливу кредитного рейтингу на процес кредитування;

- провести аналіз факторів, що впливають на показник кредитного рейтингу;
- провести факторний аналіз, визначити взаємозалежності та розрахувати показники, що будуть використовуватися для побудови моделі із використанням дискримінантного аналізу;
- побудувати модель процесу кредитування фізичних осіб із застосування кредитного рейтингу;
- визначити принципи роботи програмного забезпечення Deductor Studio та доцільність його використання для побудови моделей;
- побудувати модель логістичної регресії;
- визначити кращу модель для оцінки кредитоспроможності позичальників-фізичних осіб;
- побудувати модель оцінки кредитоспроможності позичальників-юридичних осіб на основі дискримінантної функції;
- визначити кращу модель для оцінки кредитоспроможності позичальників-фізичних осіб;
- розробити методичні рекомендації та пропозиції щодо вдосконалення процесу оцінки кредитоспроможності позичальників комерційного банку.

Об'єктом даного дослідження є кредитоспроможність позичальників комерційних банків.

Предметом є економіко-математичні та економетричні методи та моделі оцінювання кредитоспроможності позичальників банківських установ.

Процес дослідження спирався на широке використання загальнонаукових методів пізнання. У рамках системного підходу було застосовано методи аналізу та синтезу, групування та порівняння, використано аналітичні та статистичні методи, графічні та табличні методи досліджень. Для дослідження було використано моделі кредитного скорингу та дискримінантного аналізу.

Структура роботи включає у себе вступ, три розділи, висновок, список використаних джерел та додатки.

У першому розділі оцінено сучасний показників кредитної діяльності банків України за останні 7 років; систематизовано поняття та сутність «кредитоспроможності»; подано характеристику та порівняльний аналіз основних методів та моделей оцінки кредитоспроможності позичальників комерційних банків.

У другому розділі розкрито теоретичні аспекти основ кредитування позичальників; побудовано моделі оцінки кредитоспроможності фізичних осіб та обрано модель із кращою точністю прогнозів.

У третьому розділі подано аналіз показників фінансового стану підприємства; відібрано показники та побудови моделі; розкрито сутність дискримінантного аналізу; побудовано модель оцінки кредитоспроможності юридичних осіб.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ МОДЕЛЮВАННЯ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКІВ БАНКУ

1.1. Економічна характеристика та аналіз фактичного стану показників кредитної діяльності банків України за 2013-2019рр.

Необхідною умовою нормального функціонування економіки будь-якої країни є розвиток банківського сектору. Здійснити реструктуризацію економіки, яка необхідна Україні, можна через ефективну банківську систему. Також, надійний та розвинений банківський сектор здатен відіграти важливу роль у стабілізації економіки країни. Тут головна роль в утвердженні розвитку належить комерційним банкам, як одним із інструментів ринкової економіки.

Система комерційних банків охоплює важливі сфери ринкової економіки, такі як виробництво, розподіл, обмін та споживання, і є невід'ємною складовою діяльності як фізичних, так і юридичних осіб.

Комерційні банки повинні залучати тимчасово вільні кошти та ефективно використовувати акумульовані ресурси для задоволення потреб реального сектору економіки та подальшого росту і розвитку. Тому наявна необхідність досліджувати їхню діяльність.

У переважній більшості країн діє дворівнева банківська система. До цих країн відноситься і Україна. Суть дворівневої системи: на першому рівні міститься центральний банк (в Україні НБУ), а на другому знаходяться комерційні банки. Метою діяльності останніх є отримання прибутку шляхом обслуговування фізичних та юридичних осіб та через іншу інвестиційну діяльність.

Станом на теперішній час основу банківської системи України складають державні банки («Ощадбанк», «ПриватБанк», «Укресімбанк» та «Укргазбанк»). На рис.1.1 подано динаміку зміни кількості банків України із 1 січня 2013 по 1 січня 2020 рр. Після переходу «Приватбанку» у статус державного банку частка державних банків у структурі банківської системи сильно зросла.

Спостерігається зменшення частки банків іноземних банківських груп. Насамперед це пов'язане із виходом із українського ринку банків із російським капіталом.

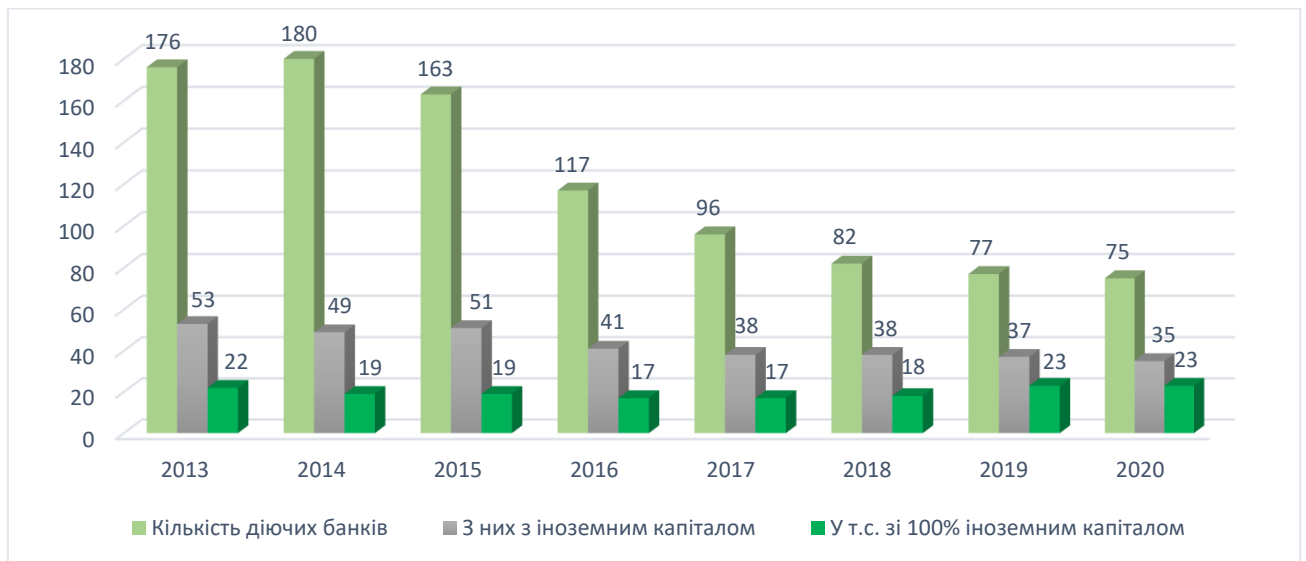


Рис.1.1 Динаміка зміни кількості банків України за 2013-2019рр.

Джерело: сформовано автором на основі [2]

Протягом 2014– 2017 рр. банківська система України функціонувала в умовах високої економіко-політичної нестабільності, ведення бойових дій на сході України та зростання кредитної заборгованості в банках України, що, своєю чергою, безпосередньо погіршило якість активних ресурсів банків.

Депозитні, кредитні та розрахункові операції займають основну частку у діяльності комерційних банків. Однак важливим є не лише вміння формування ресурсів, але й ефективне їх розміщення. Для розуміння кредитної ситуації необхідно розглянути активи банків України, що представлені у таблиці 1.1. Для того, щоб візуально побачити тенденцію змін активів, розглянемо рис. 1.2.

Як бачимо із таблиці 1.1, протягом досліджуваного періоду активи банків збільшувалися, проте незначно (найбільший приріст був на 9,83% на кінець 2019 року). Помітне зменшення відбулося на початку 2016 року (на 4,74% у порівнянні із початком 2015 року). Вагомими чинниками такого коливання у обсязі активів є зменшення кількості комерційних банків, зміна курсів національної валюти, скорочення кредитних портфелів та рух клієнтських

коштів. Висока частка активів у іноземній валюті пояснює ключовий вплив зміни курсів національної валюти на обсяг активів.

Таблиця 1.1

Зміна обсягів активів комерційних банків України в 2013-2019 рр., млн грн

Показники/роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.01. 2020
Активи:							
млн грн	1278095	1316852	1254385	1256299	1336358	1359703	1493298
у % до попереднього року		103,03	95,26	100,15	106,37	101,75	109,83
Резерви для відшкодування можливих втрат за активними банківськими операціями							
млн грн	131252	204931	321303	484383	516985	556445	492229
у % до активів	10,27	15,56	25,61	38,56	38,69	40,92	32,96
у % до попереднього року		156,14	156,79	150,76	106,73	107,63	88,46

**Джерело: розраховано автором на основі [2]*

На рівень кредитування на початку 2014 року мала вплив внутрішньополітична криза. Спостерігається зменшення частки кредитного портфеля у активах банків. Наслідком цього стала нестабільна економічна ситуація в Україні, зниження платоспроможності позичальників та нерегульованість законодавства, що стало наслідком високого кредитного ризику, який, у свою чергу, потребує формування більших резервів банків.

Як видно з рис 1.2., активи загалом залишалися на стабільному рівні, але частка резервів збільшувалася щорічно і на початок 2019 року сягнула 40,92%, що більше на 30,65% у порівнянні з 2014 роком. Це свідчить про те, що банки реагують на ризики під час кредитування клієнтів, і, тим самим, вони страхують себе від можливих загроз.

Тепер можна аналізувати статті активу, а саме – структуру кредитного портфелю. Вона показує стан кредитування на сучасному етапі та проблеми, що

виникли в процесі. Аналіз кредитного портфеля важливий з точки зору оцінки стійкості та ліквідності банків. Переважну частину прибутку банки формують саме за рахунок позик. Але важливим є не лише те, який обсяг кредитів був наданий, але й ризики, які виникли у зв'язку з видачою позик. Потрібно оцінити чи платоспроможний клієнт, чи буде погашено заборгованість клієнта і чи не виникне нової проблеми заборгованості.

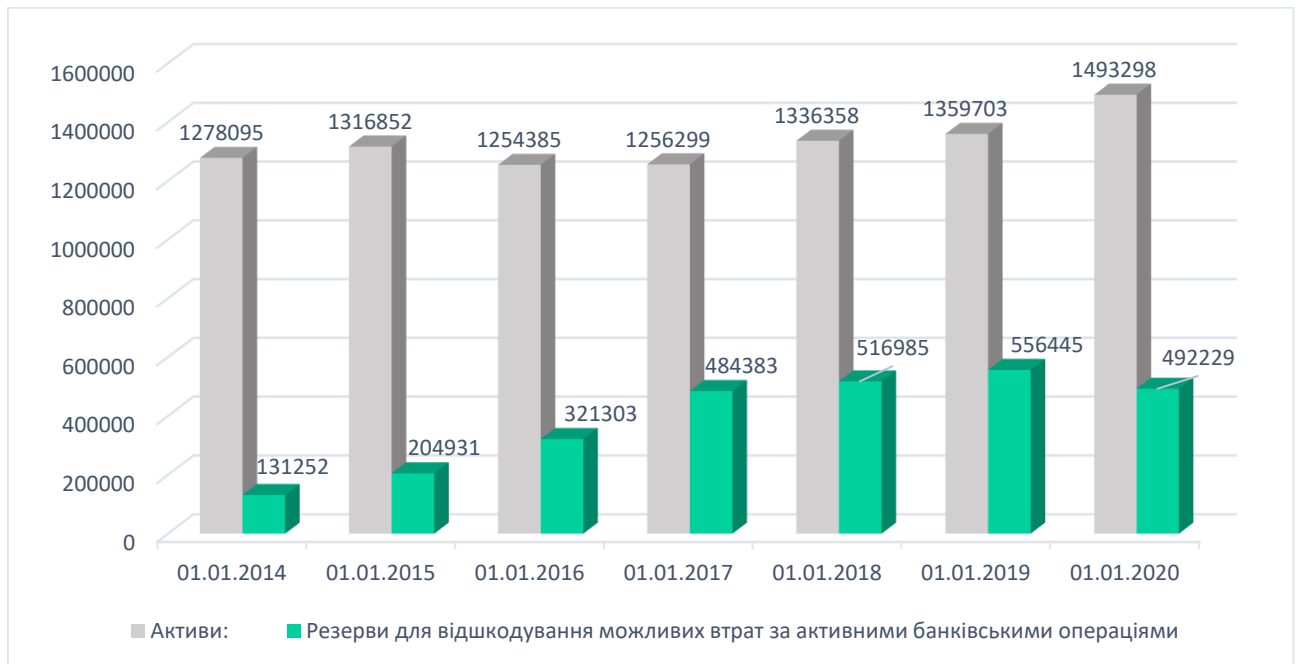


Рис.1.2. Динаміка обсягу активів комерційних банків України за 2013-2019рр.

Джерело: сформовано автором на основі [2]

Ситуація кредитування в Україні за досліджуваний період представлена в таблиці 1.2.

Як можна побачити з таблиці 1.2, резерви для відшкодування можливих втрат збільшувалися щорічно і станом на 01.01.2019 становили 49,73% від кредитного портфеля. Зменшення на 2,1% відбулося на початку 2020 року. Це свідчить про обмежену видачу кредитів та перевірку банками клієнтів на кредитоспроможність не лише на дату видачі позики, а й на весь період виплат.

Таблиця 1.2

Аналіз якості кредитного портфеля банків України за 2013-2019рр., млн грн

Показники/роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.01. 2020
Кредитний портфель, млн грн	911402	1006358	1009768	1005923	1036745	1118860	1033430
резерви для відшкодування можливих втрат за активними банківськими операціями, млн грн	131252	204931	321303	484383	511062	556445	492229
відношення резервів під кредитні ризики до кредитного портфелю, %	14,4	20,36	31,82	48,15	49,29	49,73	47,63
проблемні кредити, млн грн	70205	135921	213311	307812	594999	630767	530780
питома вага проблемних кредитів у кредитному портфелі, %	7,7	13,51	21,12	30,6	57,39	56,38	51,36

**Джерело: сформовано автором на основі [2]*

Зниження обороту та обсягів виробництва основних галузей, зниження реальних доходів населення при високій девальвації національної валюти, а також інші зміни в економіці можуть привести до появи труднощів з обслуговуванням позичальників кредитних зобов'язань, що негативно вплине на ліквідність та капіталізацію банківських установ. Дані фактори привели до нездатності багатьох позичальників вчасно виконувати свої зобов'язання, що вплинуло на зростання проблемної заборгованості в Україні впродовж 2014-2019рр. [2]. Динаміку простроченої заборгованості за кредитами представлено на рис 1.3.

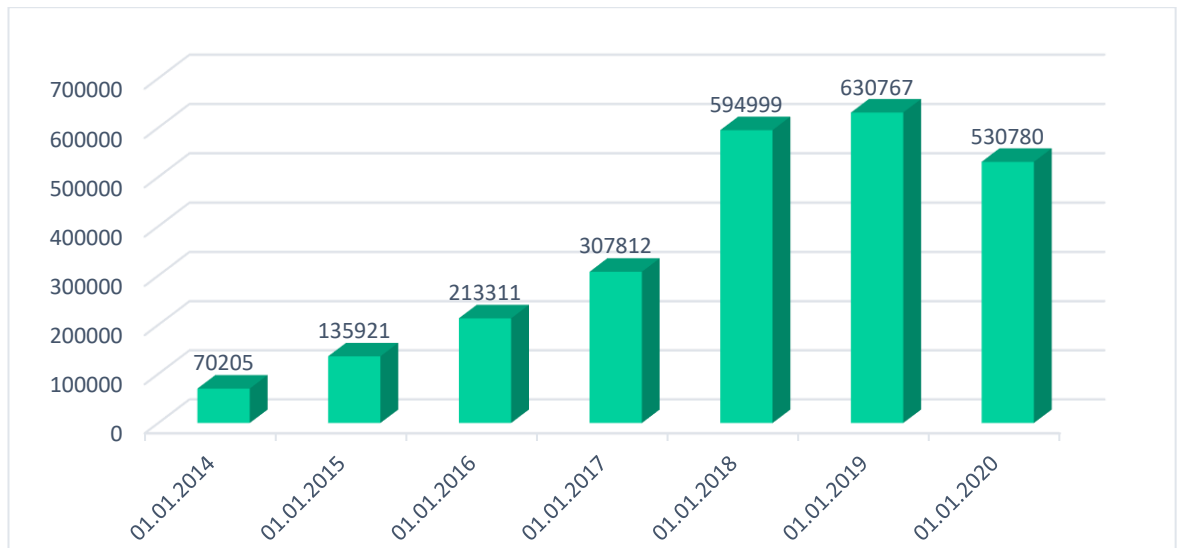


Рис.1.3. Динаміка простроченої заборгованості за кредитами у 2013-2019рр.

Джерело: сформовано автором на основі [2]

Якщо говорити про проблемні кредити протягом досліджуваного періоду, то їх питома вага у кредитному портфелі зростала шаленими темпами із кожним роком, і вже на кінець 2019 року становила 51,36% від усіх наданих кредитів. Таке зростання зумовлене високою економіко-політичною нестабільністю, наприклад девальвацією гривні, що сильно вплинуло на кредитоспроможність клієнтів. Таке збільшення частки проблемної заборгованості спричинює значні відрахування в резерви для відшкодування можливих втрат, а це, у свою чергу, приводить до того, що банки менш ефективно використовують свій капітал.

Важливою функцією банків є перерозподіл коштів, які тимчасово вивільняються в процесі кругообігу фондів підприємств та грошових доходів приватних осіб [3].

Банки надають кредити як для фізичних, так і для юридичних осіб. Власне кажучи, без позичальників неможливий сам процес кредитування, адже саме клієнти є підґрунтям цього процесу. Отож, за допомогою таблиці 1.3, розглянемо обсяги кредитування фізичних та юридичних осіб протягом 2013-2019рр.

Таблиця 1.3

Динаміка обсягів наданих кредитів юридичним та фізичним особам
банками України за період 2013-2019рр, млн грн

Показник/Роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.01. 2020
Кредити надані клієнтам	911402	1006358	1009768	1005923	1042798	1118860	1033430
Кредити, що надані суб'єктам господарювання:	698777	802582	830631	847092	870302	919071	821936
у % до всіх наданих кредитів	76,67	79,75	82,26	84,21	83,46	82,14	79,53
у % до попереднього року		114,86	103,49	101,98	102,74	105,6	89,43
кредити, що надані фізичним особам:	167773	179711	175711	157385	170938	196859	206737
у % до всіх наданих кредитів	18,41	17,86	17,4	15,65	16,39	17,59	20
у % до попереднього року		107,12	97,77	89,57	108,61	115,16	105,02

**Джерело: сформовано автором на основі [2]*

З таблиці 1.3. видно, що кредити, які надаються суб'єктам господарювання, переважають над кредитами, що надаються фізичним особам, на 59,53% та на 01.01.2020 становлять 79,53% від усіх наданих кредитів. Кредити, надані суб'єктам господарювання, показують позитивну тенденцію протягом усього досліджуваного періоду, однак на кінець 2019 року даних показник скоротився на 10,57% у порівнянні із кінцем 2018 року. Обсяги кредитування фізичних осіб у % до всіх наданих кредитів знижуються до початку 2018 року, тоді відбувається зростання на 8,61%. На початку 2019 року цей показник є найбільшим та складає 15,16%. Варто відзначити, що у 2014 році сума наданих кредитів фізичним особам банками становила 167773 млн грн, а на початок 2020 року вона зросла на 38964 млн грн, тобто обсяг наданих кредитів фізичним особам за 6 років зросла на 23,22%. Сума наданих кредитів суб'єктам господарювання за 6 років зросла на 17,62% (або на 123159 млн грн). Більша

частка у кредитуванні юридичних осіб зумовлена кризовим становищем економіки України. Банки віддають перевагу кредитуванню фізичних осіб. Таким чином, вони зменшують ризик самого процесу кредитування та намагаються повернути надані кошти разом із процентами [2].

Значно жорсткішими стали умови надання кредитів фізичним особам. Тому далі доцільно буде розглянути кредитування саме фізичних осіб.

Розглянемо для початку кредити, що були надані фізичним особам у період з 01.02.2017 по 01.08.2020. Динаміку представлено на рис 1.4.

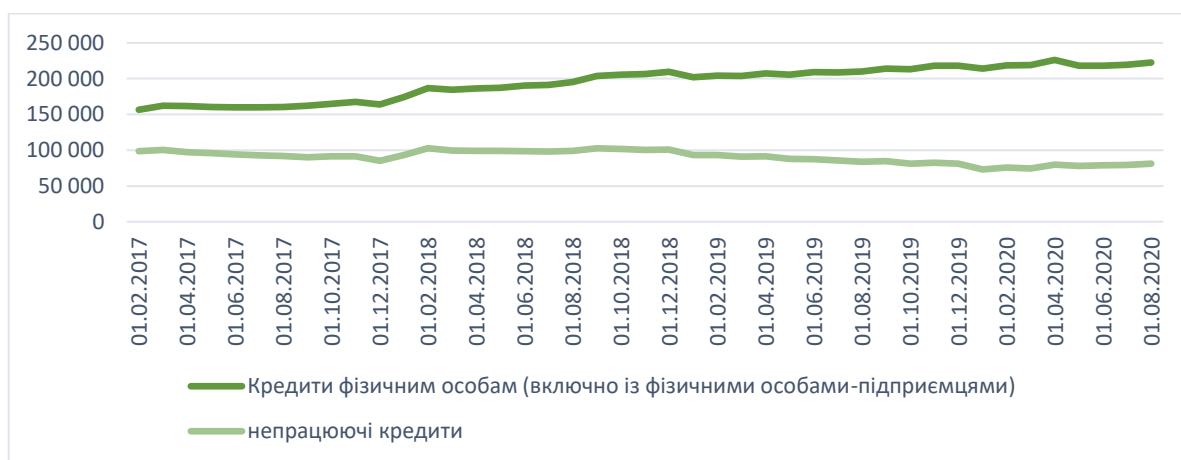


Рис.1.4. Динаміка кредитів, наданих фізичним особам у 2017-2020рр.

Джерело: сформовано автором на основі [2]

Як видно з рис.1.4, обсяги кредитування фізичних осіб показують позитивну тенденцію із початку 2017 року і до серпня цього року. Видно, що частка непрацюючих кредитів до кінця 2018 року зростала разом із обсягом наданих кредитів, однак є тенденція до зниження проблемних кредитів у наступних місяцях.

Важливо також проаналізувати споживчі кредити, що були надані домогосподарствам банками України. Динаміку можна побачити в таблиці 1.4.

З таблиці 1.4. можна побачити, що частка споживчих кредитів до початку 2017 року зменшується. Надалі спостерігається позитивна тенденція до зростання споживчих кредитів. Вже у жовтні 2019 року видно, що частка споживчих кредитів сягає 80% від обсягу усіх наданих кредитів домогосподарствам. Якщо у 2015 році темп приросту споживчих кредитів

становив -2252 млн грн, або -1,64%, у порівнянні із 2014 роком, то вже станом на 01.01.2019 темп приросту дорівнював 22154,12 млн грн, або 24,22% у порівнянні із 2019 роком. Разом із тим зменшувався і обсяг наданих кредитів домогосподарствам. Станом на 01.01.2017 він становив 163333 млн грн, що менше на 30196 млн грн, або на 18,5%, у порівнянні із початком досліджуваного періоду. Після 2017 року помітне зростання наданих кредитів домогосподарствам. Якщо порівнювати обсяг наданих кредитів на початку 2017 року із жовтнем місяцем 2019 року, то їхнє значення зросло на 53825 млн грн (або на 32,95%).

Таблиця 1.4.

Динаміка споживчих кредитів домогосподарствам України, надані вітчизняними банками у 2014-2019рр.

Показники/роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.10. 2019
Всього кредитів домогосподарствам, млн грн	193529	211215	174869	163 333	174 182	201 102	217 158
Споживчі кредити, млн грн	137346	135094	104879	101 528	122 066	151 625	173 779
Частка споживчих кредитів,%	70,97	63,96	59,98	62,16	70,08	75,4	80,02
Темп приросту споживчих кредитів, млн грн:		-2252	-30215	-14536	20537,58	29558,69	22154,12
%		-1,64	-22,37	-3,19	20,23	24,22	14,61

**Джерело: сформовано автором на основі [2]*

Комерційні банки надають кредити і для юридичних осіб. І так як їхня частка є більшою за частку кредитів, наданих фізичним особам, доцільно буде проаналізувати динаміку кредитів, наданих юридичним особам. Дані подано на рис. 1.5.

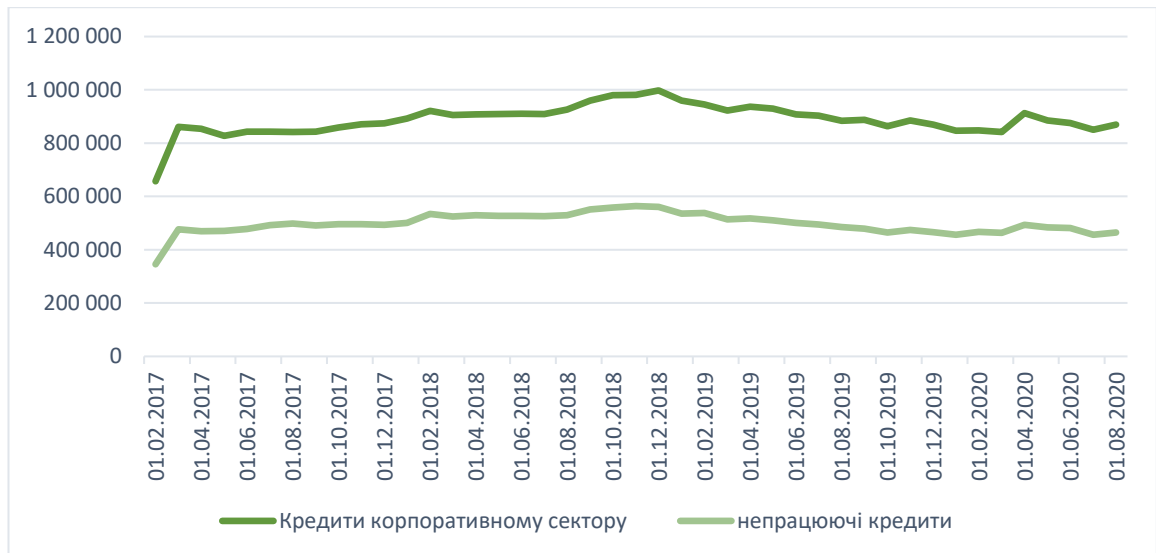


Рис.1.5. Динаміка наданих кредитів юридичним особам за період 2017-2020рр.

Джерело: сформовано автором на основі [2]

Як видно з рис.1.5, обсяги кредитування юридичних осіб показують позитивну тенденцію із початку 2017 року і до кінця 2018 року. У наступні періоди помітний незначний спад кредитування юридичних осіб. Видно, що обсяг непрацюючих кредитів цілком залежить від суми наданих кредитів.

У таблиці 1.5 розглянуто кредити, що були надані нефінансовим корпораціям, у розрізі строків їх погашення. Очікувано, що найбільшу частку будуть займати кредити із строком погашення до 1 року та від 1 до 5 років. У кредитів із строком погашення понад 5 років спостерігається зростання у частці усіх кредитів, але лише до початку 2018 року, тоді цей показник знову йде на зниження. Також видно, що зменшується частка кредитів строком погашення від 1 до 5 років. У порівнянні із початком досліджуваного періоду їхня частка скоротилася більше, аніж на 10%. Помітна позитивна тенденція у частці кредитів із строком погашення до 1 року, оскільки такі кредити є більш надійними, мають великий попит та швидко повертаються.

Таблиця 1.5

Динаміка кредитів нефінансовим корпораціям України, надані комерційними банками у 2014-2019рр.

Показник/Роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.10. 2019
Кредити нефінансовим корпораціям	778 841	787 795	822 114	829 932	859 740	845 506	781 882
За строками погашення							
до 1 року	339 631	387 900	306 002	354 989	413 695	406 323	391 878
у % до усіх кредитів	43,61	49,24	37,22	42,77	48,12	48,06	50,12
у % до попереднього року		114,21	78,89	116,01	116,54	98,22	96,44
від 1 року до 5 років	333 646	289 883	353 748	288 950	278 126	272 221	252 054
у % до усіх кредитів	42,84	36,8	43,03	34,82	32,35	32,2	32,24
у % до попереднього року		86,88	122,03	81,68	96,25	97,88	92,59
більше 5 років	105 564	110 012	162 364	185 993	167 920	166 963	137 950
у % до усіх кредитів	13,55	13,96	19,75	22,41	19,53	19,75	17,64
у % до попереднього року		104,21	147,59	114,55	90,28	99,43	82,62

**Джерело: сформовано автором на основі [2]*

Кредитний портфель представлено переважно коротко- та середньостроковими кредитами нефінансовим корпораціям, що було подано у табл.1.3. Кредитний портфель суб'єктів господарювання за видами економічної діяльності зображено на рис.1.6. Як свідчать дані, у жовтні 2019 року він складався у переважній більшості із кредитів, наданих представникам сфери «Переробна промисловість» та «Оптова та роздрібна торгівля». Вагомою є частка кредитів, наданих організаціям, що займаються постачанням електроенергії та газу і операціями із нерухомим майном.



Рис.1.6. Кредитний портфель суб'єктів господарювання за видами економічної діяльності у жовтні 2019 року, %

Джерело: сформовано автором на основі [2]

Було розглянуто структуру та динаміку кредитів, що надають комерційні банки для своїх позичальників. Однак для чіткого розуміння ситуації, що складається на фінансовому ринку, варто розглянути структуру ресурсів банків. Це пов'язано із тим, що джерелами формування ресурсів банків є створення власного капіталу, а не лише акумуляція грошових коштів позичальників. Як свідчать дані таблиці 1.6, головне джерело ресурсів банків – кошти вкладників та кредиторів. Видно, що зобов'язання зростають із кожним роком, а це може свідчити про зростання депозитної бази банків. Отже, попри усі кризові явища та ситуації, банки здатні підтримувати довіру населення.

Зобов'язання у структурі пасивів банків займають понад 80% за увесь досліджуваний період. На початку 2017 року їхня позначка сягнула 90,15%. Загалом, після 2016 року помітне зростання обсягів капіталу банків у порівнянні із попередніми роками. Однак, на кінець досліджуваного періоду маємо наступну ситуацію у порівнянні із початком 2014 року: пасиви збільшилися на 215203 млн грн (або на 16,84%), із них капітал зріс всього на 7322 млн грн (або на 3,8%), а зобов'язання аж на 207881 млн грн (або на 19,15%).

Таблиця 1.6

Структура зобов'язань та власного капіталу комерційних банків України у
2014-2019рр., млн грн

Показник/Роки	01.01. 2014	01.01. 2015	01.01. 2016	01.01. 2017	01.01. 2018	01.01. 2019	01.01. 2020
Пасиви	127809 5	1316852	1254385	1256299	1333831	1359703	1493298
Власний капітал	192599	148023	103713	123784	161108	154960	199921
у % до пасивів	15,07	11,24	8,27	9,85	12,08	11,4	13,39
у % до попереднього року		76,86	70,07	119,35	130,15	96,18	129,01
Зобов'язання	108549 6	1168829	1150672	1132515	1172723	1204743	1293377
у % до пасивів	84,93	88,76	91,73	90,15	87,92	88,6	86,61
у % до попереднього року		107,68	98,45	98,42	103,55	102,73	107,76

**Джерело: сформовано автором на основі [2]*

Проблемою банківської діяльності є виконання своїх зобов'язань. Динаміку та структуру зобов'язань комерційних банків України можна побачити на рис.1.7. Сукупний обсяг зобов'язань комерційних банків на кінець досліджуваного періоду збільшився на 7,76%, тобто до 1293377 млн грн., при цьому вага депозитів у зобов'язаннях банківського сектору продовжує зростати. Як видно із рис.1.7, спостерігається збільшення частки коштів суб'єктів господарювання. Їхня частка поступово зрівнюється із депозитами населення. Частка кредитів НБУ залишається символічною.

Проблемами розвитку кредитної діяльності банків є невизначена інституційна організація кредитного ринку, дефіцит кредитних ресурсів, підвищений кредитний ризик, недовіра контрагентів [4].

Основними напрямками підвищення кредитної активності банків для підтримки реального сектору економіки повинні стати заходи макроекономічного та мікроекономічного спрямування.

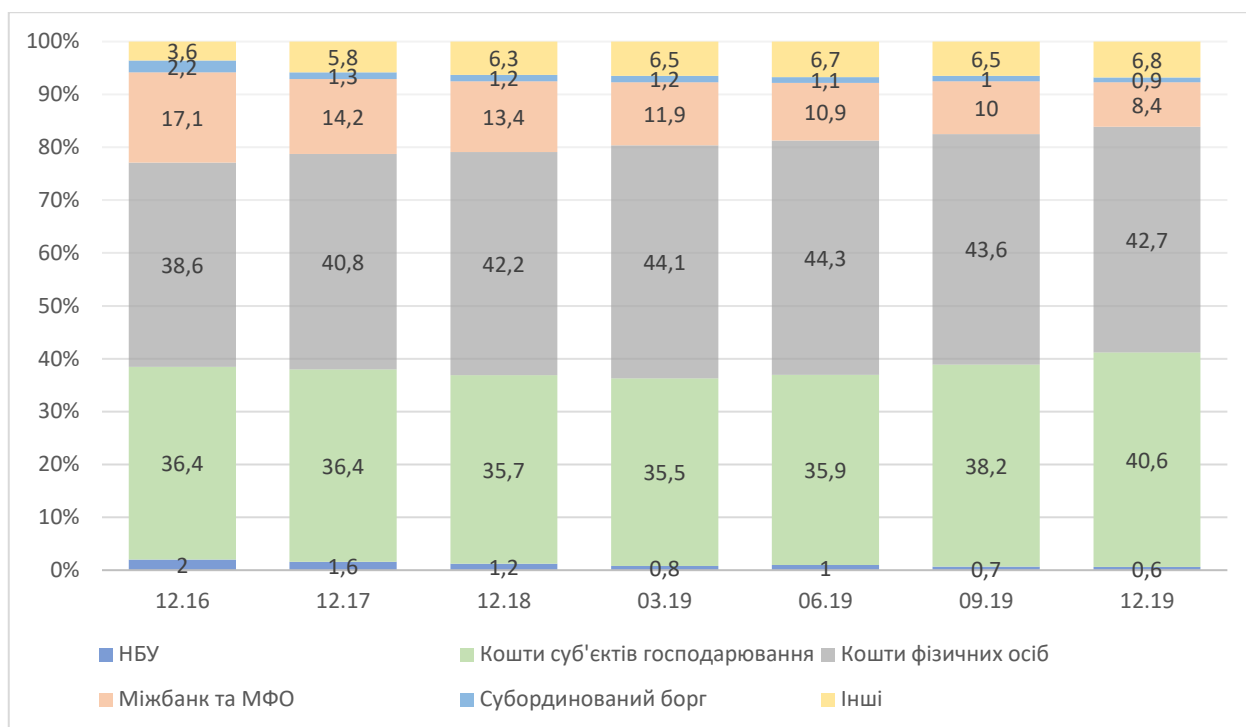


Рис.1.7. Структура зобов'язань комерційних банків України у 2014-2019рр.

Джерело: сформовано автором на основі [2]

До заходів макроекономічного спрямування слід віднести, у першу чергу, подальшу стабілізацію кредитного ринку України. Ефективне функціонування кредитного ринку та його структурних елементів значною мірою залежить від організації регулювання та нагляду за кредитними установами. Важливою передумовою цього є розробка стратегії розвитку кредитного ринку щодо його інституціональної структури. Удосконалення механізмів регулювання та нагляду впливає на підвищення ліквідності кредитних установ; посилення інвестиційної та інноваційної їх діяльності; забезпечення надійної системи гарантування вкладів і захисту клієнтів; підвищення капіталізації кредитних установ шляхом збільшення статутного капіталу.

До заходів мікроекономічного спрямування належать: підвищення ефективності кредитних операцій і поліпшення схеми надання позик на рівні взаємовідносин «банк-клієнт»; на рівні НБУ – удосконалення діючого механізму рефінансування; на рівні комерційного банку – вдосконалення відповідних аспектів фінансового менеджменту, що охоплюють продукування кредитних

послуг і підтримання стабільності банківської системи шляхом мінімізації ризиків та оптимального управління портфелем кредитних вкладень загалом.

Необхідно відзначити, що у період виходу з кризи у розрізі банківського напрямку виокремлюються наступні можливі шляхи удосконалення кредитних взаємовідносин, а саме: застосування відповідних заходів щодо стимулювання довгострокового кредитування банками суб'єктів підприємницької діяльності; оптимізація процентної політики банку щодо встановлення диференційованих підходів до кредитування підприємств малого та середнього бізнесу; удосконалення механізму зменшення або унеможливлення кредитних ризиків шляхом створення єдиної нормативної бази для визначення фінансового стану підприємств і системи рейтингів надійності позичальників, поліпшення умов забезпечення банківських позичок; удосконалення юридичної основи реструктуризації заборгованості та відповідних економічних механізмів видозміни заборгованості [5].

Отже, тепер можна підсумувати поданий аналіз. Під час кредитування в рамках банківської системи України виникає багато проблем. Зокрема, це:

- Високий рівень валютного та кредитного ризику;
- Значний рівень операційного та регулярного ризиків унаслідок політичної нестабільності;
- Збереження великої питомої ваги неробочих активів унаслідок відкладених кредитних ризиків;
- Внаслідок економічної кризи зріс рівень взаємної недовіри між банками і позичальниками;
- Обсяг проблемних кредитів, як і раніше, досить високий. Тому банки змушені збільшувати резерви під кредитні ризики, наслідком цього стає замороження частини ресурсів, які могли б стати джерелом кредитів;
- Надзвичайно довга процедура розгляду можливості отримання кредиту.

I, на додачу до останнього пункту, часто ця процедура не дає надійних результатів, що призводить до зростання проблемних кредитів. Тому на сьогодні є актуальним питання вибору та побудови ефективних методів та моделей для оцінки кредитоспроможності позичальників банку.

1.2. Поняття, механізм та необхідність оцінки кредитоспроможності позичальника.

Найбільша частина доходів банку формується за рахунок надання кредитів.

Варто зазначити, що єдиної системи оцінки кредитоспроможності позичальника не існує, кожен банк самостійно розробляє методику оцінки кредитоспроможності своїх позичальників. В основі методики оцінки кредитоспроможності підприємств-позичальників лежать методичні рекомендації НБУ щодо оцінки банками кредитоспроможності та фінансової стабільності позичальника.

Існує нагальна необхідність дослідження підходів банків та різних методів для оцінки кредитоспроможності позичальників та проведення їх порівняльного аналізу. Метою цього дослідження є підвищення ефективності відбору позичальників для кредитування.

Рівень ризику, що виникає у процесі кредитування позичальника, банки оцінюють виходячи із багатьох факторів. Найважливіші – це плато- та кредитоспроможність потенційного клієнта. Може здатися, що ці два поняття є тотожними, але насправді між ними простежується суттєва відмінність.

Платоспроможність – це здатність позичальника платити за своїми зобов'язаннями за визначений проміжок часу [6]. Мова йде про платежі за конкретну дату або ж за попередній період. До таких зобов'язань належать рахунки оплати для постачальників обладнання, повернення позичених ресурсів (кредитів), виплата заробітної плати працівникам, податки та збори.

Оцінюють цей фінансовий параметр у вигляді коефіцієнта платоспроможності, що рівний відношенню наявних грошових коштів до суми платежів за минулий період або визначену дату. Якщо цей коефіцієнт більший або рівний одиниці, то позичальника прийнято вважати платоспроможним. У іншому ж випадку мова йде про низький рівень платоспроможності клієнта, що може привести до прострочених платежів за кредитами [7].

Для аналізу платоспроможності юридичної особи банку необхідно ознайомитися із фінансовою звітністю компанії/фірми (баланс та звіти про прибуток та збиток).

Таким чином, платоспроможність – важливий сигнальний фактор, що відображає фінансовий стан клієнта за визначений момент часу.

Кредитоспроможність – здатність підприємства повністю виконати свої зобов'язання по кредитному договору, вчасно розрахуватися по сумі кредиту та нарахованим відсоткам. На відміну від платоспроможності, що оцінює минулий та поточний фінансовий стан підприємства, кредитоспроможність оцінює майбутні зобов'язання позичальника, у тому числі і ризики неповернення кредиту.

Для оцінки такої важливої характеристики позичальника використовується набагато більше факторів, і платоспроможність є лише одним із них. Якщо мова йде про підприємство, то виконується кількісний та якісний аналіз усього бізнесу, усієї схеми діяльності компанії.

При цьому оцінка кредитоспроможності будується не просто на фінансовому результаті підприємства за конкретний звітний період, але і на прогнозованих показниках діяльності на увесь період кредитування.

У даний час переважна більшість економістів визнає, що кредитоспроможність – це спроможність позичальника за конкретних умов кредитування в повному обсязі та у визначений кредитною угодою термін розрахуватися за своїми борговими зобов'язаннями [8, 9].

Для розуміння важливості кредитування у банківській сфері необхідно розглянути функції, які воно виконує. Для початку це функція перерозподілу

грошових коштів. Вона полягає у тому, що розміщені у банку тимчасово вільні кошти одних підприємств, організацій та населення перерозподіляються банками (за умови повернення) між іншими підприємствами, організаціями та населенням, які потребують додатковим ресурсів для здійснення процесу виробництва.

Наступна функція – це заміщення готівки у платіжному обігу. За допомогою кредиту підприємства та установи мають можливість здійснювати розрахунки не готівкою, а шляхом перерахування суми платежу на рахунок одержувача у безготівковому вигляді.

Останньою згадаємо контрольну функцію. Вона виявляється у процесі перерозподілу кредитних ресурсів, контролю за їх цільовим та ефективним використанням.

Спочатку пропонується розглянути кредитування позичальників-юридичних осіб.

Кредит стимулює розвиток підприємств, прискорює формування ресурсів капіталу для розширення відтворення на основі досягнень науково-технічного прогресу [10].

Без кредитної підтримки складно забезпечити швидкий розвиток домогосподарств, підприємств, провадження інших видів підприємницької діяльності. Необхідність кредитування обумовлена також комерційною організацією управління в умовах ринку, коли на кожному підприємстві постійно виникає додаткова потреба у ресурсах.

За допомогою кредитного механізму підприємства отримують засоби, необхідні їм для нормальної роботи. Кредит має велике значення також у розвитку економічних зв'язків між галузями та регіонами, у підвищенні ефективності виробництва, у створенні та використанні доходів та прибутків.

Кредит здатен активно діяти на об'єм та структуру грошової маси, платіжного обороту та швидкості обертання грошей. Завдяки кредиту відбувається більш швидкий процес капіталізації прибутку, а, відповідно, і концентрації виробництва.

Кредитоспроможність клієнта комерційного банку розуміє під собою можливість позичальника здійснювати виплати по взятих у борг коштам вчасно та у повному об'ємі, із врахуванням відсотків банку.

Так як банк зацікавлений у своєму прибутку, то перед наданням кредиту він, звісно, проводить ретельну перевірку кредитоспроможності позичальника та прогнозує ймовірність повернення коштів назад у банк через деякий час. Рівень неплатоспроможності потенційного клієнта у минулому буде враховуватися банком, однак існує ряд інших критеріїв, котрі використовують для оцінки ймовірності повернення коштів назад до банку.

На законодавчому рівні поняття кредитоспроможності та базову методичку її оцінювання визначено у Постанові правління Національного банку України "Про затвердження Положення про порядок формування та використання банками України резервів для відшкодування можливих втрат за активними банківськими операціями" від 25 січня 2012 року № 23. Положення розроблено з урахуванням основних вимог міжнародних стандартів фінансової звітності, рекомендацій Базельського комітету з банківського нагляду та має на меті забезпечення надійності й стабільності банківської системи, захисту інтересів вкладників і кредиторів банків. Згідно з даним Положенням кредитоспроможність визначається як наявність у позичальника (контрагента банку) передумов для проведення кредитної операції і його спроможність повернути борг у повному обсязі та в обумовлені договором строки [11].

Банк може для реалізації власної кредитної політики самостійно розробляти порядок оцінки фінансового стану позичальника-юридичної особи, але із урахуванням основних показників, встановлених НБУ. Додатково банк може визначати методичку врахування додаткових показників в оцінці кредитоспроможності своїх клієнтів (попередньо обґрунтувавши їх доцільність). Це дозволить банку мінімізувати кредитні ризики від неповернення боргів та визначити потенційно необхідну суму для формування резерву відшкодування можливих втрат.

Аналіз кредитоспроможності проводиться із тією метою, щоб визначити можливість надання кредиту клієнту або ж припинення кредитних зв'язків із даним позичальником за результатами оцінки його фінансової діяльності.

Для точнішого визначення характеристик окремого кредиту їх можна згрупувати за напрямками надання (цим і визначається ступінь ризику даного кредиту):

- кредити для початку бізнесу. Використовуються вони для вкладень у основні та оборотні кошти, для капітальних вкладень. Такі кредити мають високий ступінь ризику;
- кредити для стимулювання операційної діяльності. Наявний досить серйозний ризик у втраті, що пояснюється зміною ринкової кон'юнктури, попиту, сезонності товару;
- кредити під короткострокові проекти. Такі проекти, зазвичай, мають високий рівень прибутковості, однак існує ризик у характері цього проекту. Оскільки чим вищий очікуваний прибуток від проекту, тим більший ризик втрат по ньому;
- кредити під довгострокові проекти. Тут присутні два види ризиків: перший пов'язаний із можливою інфляцією, другий – із моральним старінням об'єкту кредитування.

Умовами надання кредиту позичальнику є його здатність у залученні коштів, достатніх для погашення боргу та сплати нарахованих відсотків.

Оцінюючи кредитоспроможність, банк аналізує такі характеристики [12]:

- Доходи підприємства. Оцінка прибутку, який одержує підприємство, пов'язується із ступенем ризику при кредитуванні. Банк оцінює прибуток підприємства з точки зору майбутнього погашення боргу підприємства за рахунок цього прибутку;
- Вагомість підприємства. Характеризується якістю звітів, що надаються підприємством, своєчасність погашення попередніх кредитів, компетентність його керівництва;
- Мета використання кредиту;

- Погашення кредиту. Здійснюється аналіз повернення кредиту за рахунок реалізації продукції та інших матеріальних цінностей, надання гарантій та використання заставного права;
- Сума кредиту. Співвідносяться власні та позичені кошти підприємства для оцінки його ліквідності;
- Характер клієнта: проводиться аналіз репутації клієнта та його топ-менеджерів. Вивчається, як довго ця фірма перебуває на ринку, як вона розвивається у порівнянні із середніми показниками у своєму сегменті, досвід роботи її керівників, історія погашення попередніх боргів компанії, бізнес-план витрат та шляхи досягнення прибутків;
- Контроль: вивчається чи легально працює позичальник, як може вплинути на його фінансовий стан очікувані зміни у законодавстві, оцінюється наскільки інформація про позичальника та позику, що наявні у кредитній заявці, відповідають стандартам банку та стандартам органів банківського нагляду;
- Здатність клієнта позичати гроші. Доволі формальний критерій, що показує, чи має посадова особа право підписувати договір із банком від імені своєї фірми.

Окрім цього, можна навести ще групу критеріїв аналізу кредитоспроможності позичальника:

- можливість підприємства випускати конкурентоспроможну продукцію;
- капітал клієнта;
- забезпечення кредиту;
- умови, у яких здійснюється кредитна операція.

Таким чином, кредитоспроможність можна назвати прогнозом майбутньої платоспроможності підприємства. Рішення банку буде базуватися на основі результатів цього аналізу.

Процес кредитування має враховувати зміни у платіжних можливостях позичальника за увесь період існування заборгованості. Із урахування цього

фактору, рівень кредитоспроможності є інтегральною характеристикою фінансового стану позичальника. Оцінювати її прийнято у два етапи:

1. Визначається клас позичальника. На його основі говорять про можливість надання кредиту даному клієнту.

2. Аналіз можливості повернення кредиту. Тут здійснюють оцінку стану обслуговування боргу позичальником та визначаються ризики кредитної заборгованості.

Для оцінки рівня кредитоспроможності використовують різну інформацію про підприємство. Можна виділити наступні джерела інформації: фінансова звітність; статистична звітність; дані бухгалтерського обліку; експертні оцінки та інша додаткова інформація, що надається за погодженням обох сторін.

Складність оцінки кредитного ризику для банку полягає у тому, що цей ризик є комплексним, і управління такими ризиками є складним завданням. Важливим інструментом управління кредитним ризиком для комерційного банку є його кредитна політика. Посилення конкуренції на ринку банківських послуг із надання кредиту підкреслює значимість управління кредитними ризиками. Тому система управління кредитним ризиком є важливою для комерційних банків. Оскільки банк повинен забезпечувати надійність та бути основою стабільності економічної системи, своєчасна ідентифікація кредитного ризику, якісне управління ним і визначення чинників ризику набувають особливого значення [13].

На рисунку 1.8 зображено порядок оцінки кредитоспроможності позичальника-юридичної особи.

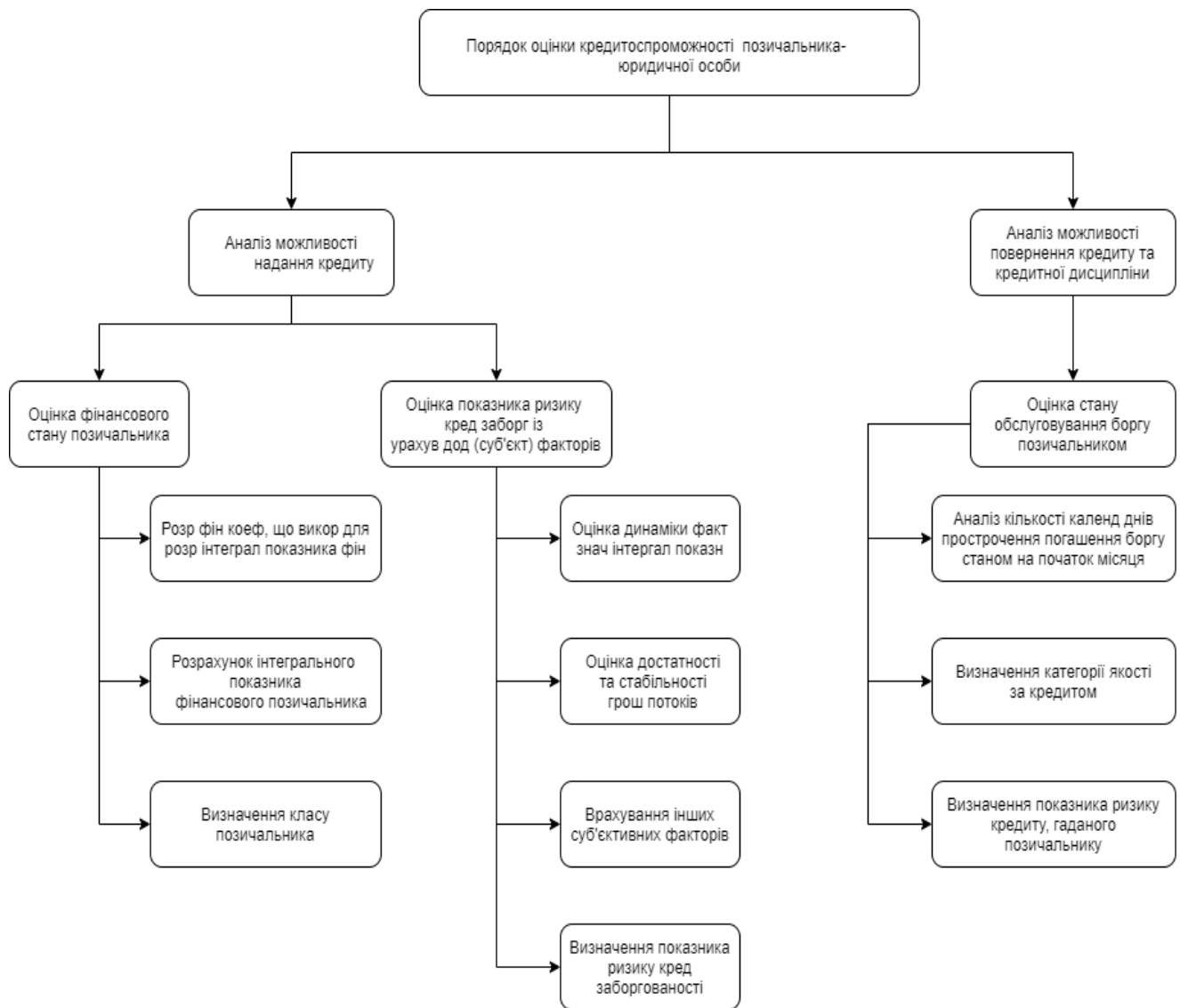


Рис.1.8. Порядок оцінки кредитоспроможності позичальника-юридичної особи.

Джерело: [13]

Як видно процес оцінки є складним та комплексним. Прийняття рішення про надання кредиту відбувається на етапі аналізу можливості надання кредиту після того, як було дано оцінку фінансовому стану позичальника та оцінку показника ризику кредитної заборгованості. У разі позитивного рішення про надання кредиту конкретному позичальнику переходять до оцінки стану обслуговування боргу позичальником.

Далі перейдемо до розгляду кредитоспроможності позичальників-фізичних осіб.

Найкращим прикладом оцінки кредитоспроможності фізичних осіб є скорингові карти. Поняття «скорингова карта» означає набір характеристик (вік,

доход, професія, стаж роботи, наявність майна й т.д.) позичальника й відповідних вагових коефіцієнтів, виражених у балах. Потенційний позичальник повідомляє про себе необхідні відомості і йому нараховується певна кількість балів. Чим більше балів – тим меншим є ризик (і більшим є шанс отримати кредит).

Фрагмент скорингової карти представлено в таблиці 1.7.

Таблиця 1.7

Фрагмент скорингової карти

Назва характеристики	Атрибут	Скоринговий бал
Вік	18-23	63
Вік	24-26	76
Вік	27-30	79
Вік	31-34	83
Вік	35-40	86
Вік	40-51	89
Вік	51-...	92
Дохід	<3200	75
Дохід	3200-5000	83
Дохід	5000-7000	89
Дохід	7000-9000	95
Дохід	>9000	107

Скоринговий бал формується на основі різноманітних характеристик позичальника. Такими характеристиками, як правило, виступають:

- соціально-демографічні (вік, освіта, шлюбний статус, приналежність до певної соціальної групи, регіон проживання, тощо);
- професійно-кваліфікаційні (стаж роботи на останньому місці роботи, посада, сукупний досвід роботи тощо);
- добробуту (величина та структура доходів, наявність рухомого та нерухомого майна тощо);
- поведінкові (повернення кредитів в минулому, наявність та тип прострочення тощо).

Кожна з характеристик включає в себе низку атрибутів, які відображають певні параметри характеристики (наприклад, атрибутами характеристики «вік» є вікові категорії 18 – 22 роки, 23 – 30 років, 31 – 40 років тощо).

У спрощеному вигляді скорингова модель - це зважена сума визначених характеристик позичальника: вік, сімейний стан, місце роботи, дохід та багато інших факторів. Дана методика може застосовуватися і для фізичних, і для юридичних осіб.

Для того, щоб побудувати скорингову модель, яка буде давати найточніші прогнози неповернення за кредитами, необхідно, щоб популяція клієнтів, на якій розробляється дана модель, була якнайбільш однорідною. Мається на увазі, що варто враховувати відмінності у менталітеті кожної окремої нації.

Також потрібно знати, що не можна автоматично перенести модель із однієї країни в іншу або з одного банку в інший, зважаючи на особливості кредитування та профіль клієнтів.

Наприклад, в США є плюсом те, що людина змінила багато місць роботи, оскільки це свідчить про те, що вона є затребувана на ринку праці. В Україні, навпаки, дана обставина є мінусом, оскільки говорить про те, що людина не є цінним працівником, що підвищує ймовірність втрати роботи і прострочки платежів.

Таким чином, скорингові карти необхідно адаптовувати для різних періодів часу як і для різних країн, навіть для різних регіонів країни. Навіть всередині одного банку існують різноманітні моделі для різних груп клієнтів і видів кредиту.

1.3. Характеристика та порівняльний аналіз основних методів оцінювання кредитоспроможності позичальника.

Сучасні практичні підходи до аналізу кредитоспроможності позичальника комерційного банку засновані на комплексному застосуванні фінансових та нефінансових критеріїв. Умовно такі підходи можна поділити на дві категорії – класифікаційні та комплексні – як представлено на рис. 1.9.

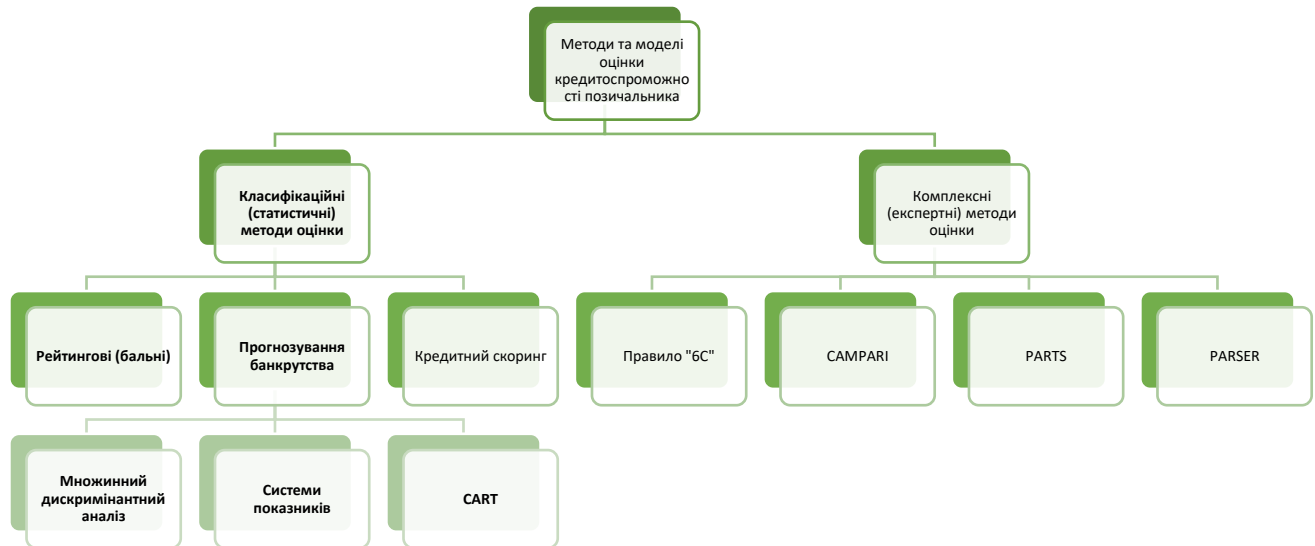


Рис 1.9. Методи оцінки кредитоспроможності позичальника.

Джерело: [14]

Класифікаційні моделі дозволяють поділити позичальників на групи (класи) і являються допоміжним інструментом при визначенні можливості схвалення кредитної заявки. Рейтингові моделі ділять позичальників на «поганих» та «хороших», а моделі прогнозування намагаються диференціювати фірми-банкроти та стійкі компанії.

Західними банками активно застосовуються рейтингові (бальні) системи оцінки, що створюють самі банки на основі емпіричного підходу із використанням регресійного математичного або факторного аналізу. Ці системи використовують історичні дані про банківські «надійні» та «погані» позики та дають змогу визначити критеріальний рівень оцінки позичальника.

Етапи проведення рейтингової моделі схематично зображено на рис.1.10.

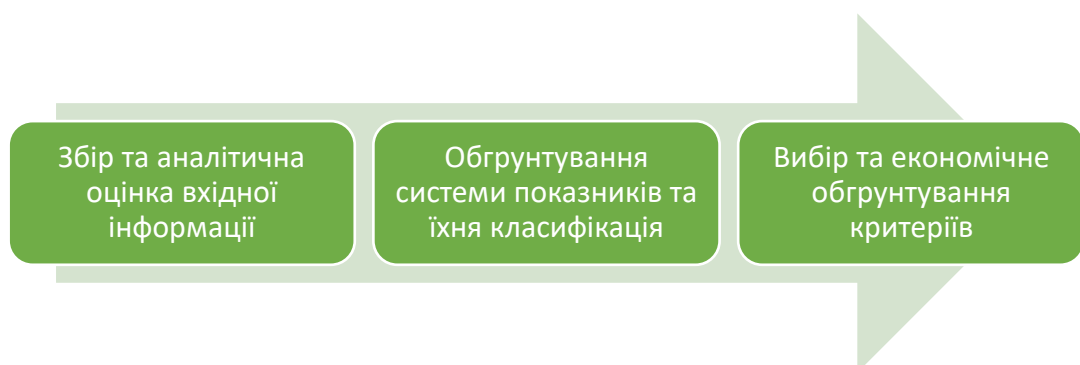


Рис. 1.10. Етапи проведення рейтингової моделі .

Джерело: [15]

Показник рейтингової оцінки може бути надійним критерієм порівняльної оцінки діяльності різних підприємств та їхніх підрозділів, конкурентоспроможності їхньої продукції, показником ефективності прийнятих раніше управлінських рішень, основою вибору можливих варіантів розвитку виробництва, критерієм інвестиційної привабливості різних господарських об'єктів, помічником у виборі банківською установою надійного ділового партнера. Перевагою рейтингової (бальної моделі) є її простота: достатньо розрахувати фінансові коефіцієнти та зважити їх, щоб визначити клас позичальника.

Рейтингова оцінка позичальника-підприємства розраховується на основі отриманих фінансових коефіцієнтів та виражається у балах. Бали вираховуються шляхом множення значення будь-якого показника на його вагу у інтегральному показнику (рейтингу).

Загальний вигляд рейтингової оцінки [16]:

$$K_0 = \sum_{i=1}^n A_i * K_i, \quad (1.1)$$

де K_0 – інтегральний показник (рейтинг);

A_i – вага окремого i -го показника; $\sum A_i = 1$

K_i – значення i -го показника;

n – кількість показників.

Модифікацією рейтингової оцінки є кредитний скоринг - це математична або статистична модель, за допомогою якої кредитна установа може визначити, наскільки велика ймовірність того, що конкретний позичальник поверне кредит у визначений термін. Це робиться на основі кредитної історії попередніх клієнтів. Тобто, в основі скорингових систем лежить припущення, що люди зі схожими соціальними показниками поведуться однаково.

Відмінність кредитного скорингу від рейтингової оцінки полягає у тому, що у формулу рейтингової оцінки (1.1) замість K_i підставляють B_i – окрема бальна оцінка i -го показника. При цьому для кожного показника визначають декілька інтервалів значень, і кожному інтервалу надається певна кількість балів або визначається клас (1, 2, 3...).

Якщо отриманий позичальником рейтинг (кредитний скоринг) нижчий раніше встановленого спеціалістами та експертами банку значення, то такому позичальнику буде відмовлено у наданні кредиту. Якщо ж його оцінка задовольняє встановлені нормативи, то його кредитна заявка буде схвалена. При введених допустимих інтервалів значень оцінки можна одночасно визначати відповідну кожному інтервалу відсоткову ставку та вид забезпечення.

Переваги методу [17]:

- 1) Об'єктивність (мінімальний вплив людського фактору на прийняття рішення).
- 2) Адаптованість.
- 3) Мінімальні затрати на навчання персоналу.
- 4) Виявлення і запобігання шахрайства.

Недоліки методу [17]:

- 1) Необхідність частої адаптації у зв'язку зі змінами внутрішніх умов.
- 2) Важкість визначення характеристик і значимості показників у балах.
- 3) Оцінка кредитоспроможності потенційних позичальників проводиться на основі наявної інформації про попередні, видані кредити, а відомостей, що характеризують можливу поведінку позичальників, яким було відмовлено у кредиті, немає.
- 4) Потреба у постійній доробці та оновленні.

Прогнозні моделі використовують для оцінки якості потенційних позичальників. Базою прогнозних моделей є статистичні методи, найбільш розповсюдженими серед яких є множинний дискримінантний аналіз (МДА, відомий так само як «кластерний аналіз»).

При оцінці кредитоспроможності позичальника множинним дискримінантним аналізом використовується дискримінантна функція (1.2). Дискримінантна функція враховує коефіцієнти регресії та фактори, що характеризують фінансовий стан позичальника. Коефіцієнти регресії розраховуються шляхом статистичного аналізу даних. Даними для аналізу

служить выборка фирм, які вижили протягом досліджуваного періоду або стали за цей час банкрутами. Якщо значення показника Z близьке до показника середнього значення фірми-банкрута, то, за умови подальшого погіршення ситуації у фірмі, вона теж збанкрутує. Можливий і варіант, коли фірмі вдасться уникнути банкрутства, шляхом старань менеджерів у питанні усунення фінансових труднощів компанії. Отже, Z -значення являється своєрідним сигналом для виявлення збанкрутілих (або ж близьких до банкрутства) фірм.

Загальний вигляд дискримінантної функції:

$$Z = a_0 + \sum_{i=1}^n a_i * f_i, \quad (1.2)$$

де a_0 та a_i – коефіцієнти регресії,

f_i – фінансові коефіцієнти, що характеризують фінансовий стан позичальника.

Графічно лінія відсікання має вигляд прямої, що ділить компанії на дві групи: на потенційних банкрутів та на тих, кому банкрутство не загрожує.

Відомими моделями множинного дискримінантного аналізу є моделі Альтмана та Чессера.

У двадцятому столітті Альтманом, Хольдерманом та Нарайаной було введено поняття « Z -аналізу» [18], в основу якого покладено рівняння (1.3):

$$Z = 1,2 * X_1 + 1,4 * X_2 + 0,6 * X_3 + 1,0 * X_4 + 3,3 * X_5, \quad (1.3)$$

де X_1 – відношення власних оборотних засобів до суми активів;

X_2 – відношення реінвестованого прибутку до суми активів;

X_3 – відношення ринкової вартості акцій до позичкового капіталу;

X_4 – відношення виручки від реалізації до суми активів;

X_5 – відношення прибутку до оподаткування та вирахування відсотків до суми активів.

Залежно від значення показника Z відбувається віднесення підприємства до певної групи [18, 19]:

$Z \leq 1,8$ – ймовірність банкрутства дуже висока;

$1,8 < Z \leq 2,7$ – ймовірність банкрутства висока;

$2,7 < Z \leq 3,0$ – ймовірність банкрутства низька;

$3,0 < Z$ – ймовірність банкрутства дуже низька.

Модель Альтмана, що складається із п'яти факторів, була побудована на основі аналізу фінансового становища 66 компаній. Модель давала прогноз банкрутства на 2-3 роки вперед із точністю до 95%. Подальші роботи Альтмана та його колег покращили напрацьовані дослідження. У їхній останній спільній роботі було досліджено фактор капіталізації зобов'язань оренди та було застосовано прийом згладжування даних (це дозволило вирівняти випадкові коливання). Нова модель дозволяє визначити ймовірність настання банкрутства із точністю до 70% на п'ять років уперед [18, 19].

Прогнозом випадків невиконання умов договору про кредит займається модель нагляду за позиками Чессера. «Невиконання умов» – це не лише випадки непогашення позики, але і усі інші відхилення від умов надання кредиту, що можуть зробити позику менш вигідною для банку.

Модель Чессера розглядає шість параметрів [20]:

X_1 – відношення вартості цінних паперів до суми активів;

X_2 – відношення чистої суми продажів до суми вартості цінних паперів;

X_3 – відношення прибутку до оподаткування та відрахування відсотків до суми активів;

X_4 – відношення зобов'язань до суми активів;

X_5 – відношення основного капіталу до чистих активів;

X_6 – відношення оборотного капіталу до чистої суми продажів.

Рівняння моделі Чессера (1.4) виглядає наступним чином [20]:

$$Y = -2,0434 - 5,24 * X_1 + 0,0053 * X_2 - 6,6507 * X_3 + 4,4009 * X_4 - 0,0791 * X_5 - 0,1020 * X_6 \quad (1.4)$$

Змінна Y – представляє лінійну комбінацію незалежних змінних, що використовується у рівнянні (1.5) для оцінки ймовірності невиконання умов договору [16, 20]:

$$Z = \frac{1}{1+e^{-Y}}, \quad (1.5)$$

де $e = 2,71828$ (число Ейлера – основа натуральних логарифмів).

Високе значення оцінки Z свідчить про високу ймовірність невиконання умов договору позичальником. У моделі Чессера застосовуються наступні критерії:

- якщо $Z \geq 0,50$, то позичальник відноситься до групи тих, хто ймовірно порушить умови договору;
- якщо $Z < 0,50$, то позичальника можна вважати надійним у відношенні виконання умов договору.

По результатах досліджень Чессера по даних 37 «позитивних» та 37 «поганих» позиках, йому вдалося правильно визначити три із кожних чотирьох досліджуваних випадків [16, 20].

Складність застосування множинного дискримінантного аналізу полягає у тому, що для вірних та адекватних результатів необхідна наявність великої вибірки підприємства, що будуть диференційовані за галузями та розмірами. Недоліком є те, що часто інформацію про стан компаній усередині певної галузі важко, а то і неможливо, знайти; і часто відсутні дані для достатньої для аналізу кількості підприємств, що збанкрутували. Це значно погіршує розрахунки коефіцієнтів регресії.

Модель CART також відноситься до класифікаційних методів оцінки кредитоспроможності позичальника. CART (Classification and regression tree) – непараметрична модель, основними перевагами якої є можливість широкого застосування, легкість у розумінні та простота розрахунків. Суть методу: компанії-позичальники розділяються на «гілки» залежно від значень певного фінансового коефіцієнта, потім йде поділ на «гілки» знову у відповідності до інших значень коефіцієнтів. Точність таких моделей складає близько 90%. Приклад класифікаційного дерева для виявлення фірм-банкрутів представлено на рис.1.11 [21].

Використовуючи усі вище описані методи оцінки надійності клієнта, варто пам'ятати, що надання кредитів це не лише механічний акт. Це складний процес, де важливим аспектом виступають людські відносини між сторонами.

Математичні методи та моделі не враховують людського фактору при прийнятті рішень, однак на практиці цей аспект необхідно враховувати.

Недоліком класифікаційних моделей є переоцінка важливості кількісних факторів, емпіризм (довільність вибору системою базових кількісних показників), недостовірність вихідних даних, громіздкі дані.

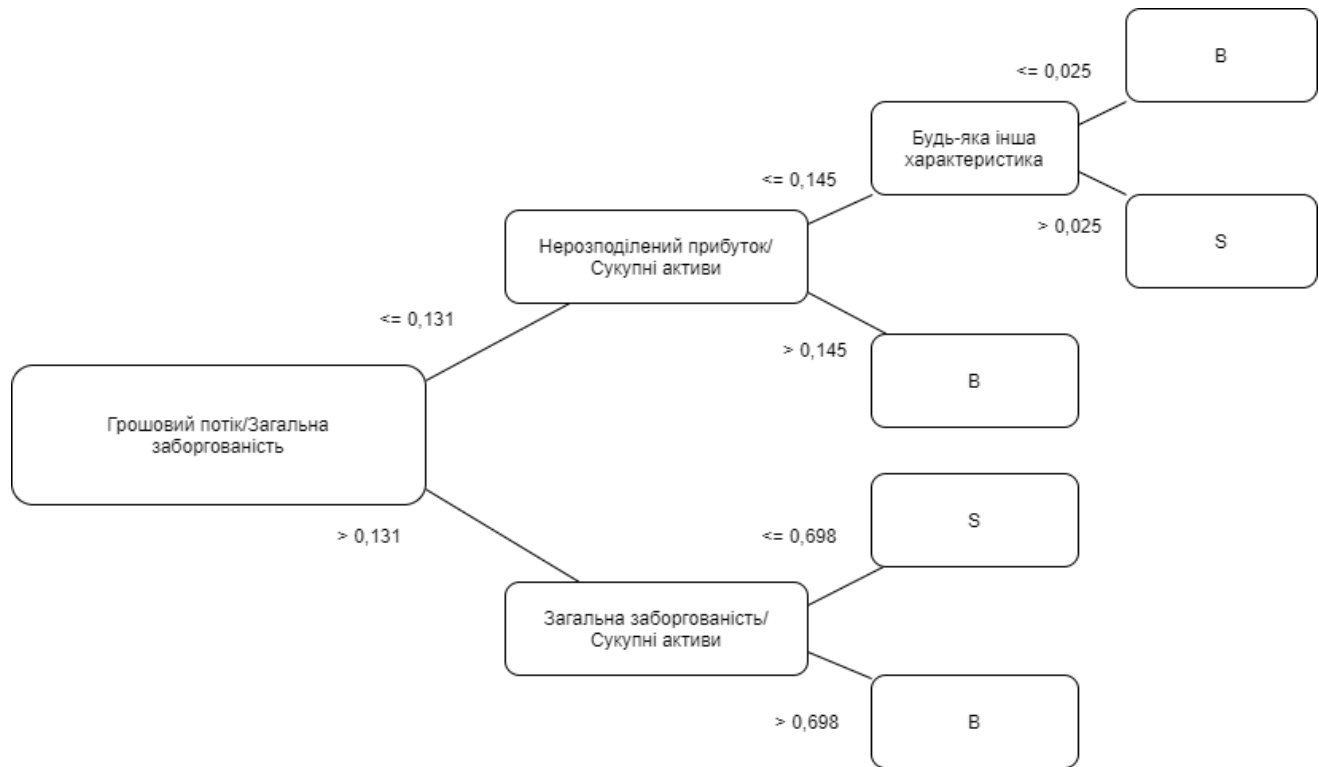


Рис.1.11. Приклад використання моделі CART (B – ймовірний банкрут, S – ймовірно стійка компанія)

Джерело: [21]

Поєднати кількісні та якісні характеристики позичальника дозволяють моделі комплексного аналізу. До них належать: правило «шести Сі», CAMPARI, PARTS, оціночна система аналізу.

У правилі «шести Сі», який досить популярний у банках США, використовуються характеристики, що розпочинаються на англійську літеру «С», для відбору позичальників. Критерії, що використовуються у моделі, – scarcity (здатність брати позику), character (репутація позичальника), capital/cash (здатність отримувати дохід), collateral (забезпечення активами), conditions (стан економічної кон'юнктури), control (чутливість позичальника до зміни макроекономічних параметрів).

Далі пропонується коротко розглянути два популярних у світовій практиці підходи у оцінці кредитоспроможності позичальників. Ці підходи мають назви PARSER (рис.1.12) та CAMPARI (рис.1.13) [22]. Аналіз методом CAMPARI полягає у почерговому виділенні із кредитної заявки та фінансових документів найбільш суттєві фактори, що визначають діяльність клієнта

PARSER



- P - Person - Інформація про особу позичальника та його репутацію;
- A - Amount - необхідна сума кредиту;
- R - Repayment - можливості погашення;
- S - Security - оцінка забезпечення;
- E - Expediency - доцільність кредиту;
- R - Remuneration - процентна ставка.

Рис.1.12. Розшифрування методики PARSER

Джерело: [22]

CAMPARI



- C - Character - характеристика клієнта;
- A - Ability - здатність погасити позику;
- M - Margin - дохідність;
- P - Purpose - мета взяття кредиту;
- A - Ammount - сума позики;
- R - Repayment - умови погашення позики;
- I - Insurance - страхування ризику непогашення позики.

Рис.1.13. Розшифрування методики CAMPARI

Джерело: [22]

Оціночна система застосовується при аналізі індивідуальних позичальників. Тут головна роль відводиться на досвід та проникливість експертів та спеціалістів банку. Вони оцінюють характер позичальника, наміри використання кредиту та джерела погашення кредиту. Характер позичальника може визначатися із його кредитної історії, місцем та протяжністю роботи там, типом та місцем проживання. Вік клієнта може розглядатися як фактор, від якого залежать майбутні доходи, він також визначає час, що лишився позичальнику до пенсії, та очікувану тривалість життя позичальника, що враховується при визначенні терміну кредиту. Дохід клієнта майже завжди є основним джерелом погашення кредиту. Експерти по кредитах повинні бути максимально об'єктивними та не звертати увагу на суб'єктивні цінності та власні вподобання.

Західними спеціалістами складено перелік із 20 характеристик, за якими прийнято оцінювати заявку на кредит від фізичної особи. Для кожної із характеристик використовують градацію від 2 до 11. Перелік цих характеристик наведено у таблиці 1.8 [23].

Загалом можна назвати наступні недоліки методики оціночної системи відбору суб'єктів кредитування:

- Суб'єктивізм експертів. Рішення експертів базується на їх особистому досвіді, інтуїції та знаннях, а це приводить до достатньо суб'єктивних результатів.
- Нестабільність результатів. Оцінка може залежати від особистих вподобань та емоційного стану експертів.
- Некерованість експертизи. Якість тут буде випадковою величиною, яку не можна буде покращити чи погіршити.
- Відсутність механізму навчання експертів. Стати хорошим експертом можна лише завдяки накопиченню досвіду, передати який не можна, оскільки відсутні будь-які ефективні методики навчання.

- Проблема підвищення кваліфікації експерта. Підвищити кваліфікацію експерта можливо лише через накопичення досвіду, як позитивного, так і негативного. А негативний досвід – це нові проблемні кредити.
- Висока вартість експертизи. Оскільки тут бере участь вище управління банку.
- Обмеження мінімального розміру кредитної заявки. Це впливає із попереднього недоліка, оскільки проводиться дорога експертиза.
- Обмежується число можливих заявок для розгляду. Оскільки експерти не можуть фізично розглянути та проаналізувати усі заявки, що надходять до банку.
- Втрачена вигода. Цей недолік впливає із попереднього судження, оскільки розглядається обмежена кількість заявок, що неодмінно приведе до обмеження потоку заявок та упущення вигідної угоди.

Таблиця 1.8

Перелік характеристик для оцінки кредитоспроможності фізичних осіб

1	Наявність рахунку у банку	11	Тривалість проживання на одному місці
2	Термін, на який береться кредит	12	Фінансовий стан
3	Кредитна історія клієнта	13	Вік клієнта
4	Мета використання кредиту	14	Борги клієнта
5	Сума кредиту	15	Характеристики квартири/будинку
6	Сума на рахунках у банках	16	Кількість попередній кредитів
7	Тривалість роботи на одному місці	17	Спеціальність
8	Часткова сплата представленої суми	18	Кількість осіб на утриманні
9	Сімейний стан	19	Наявність телефону
10	Можливості поручителя	20	Належність до жителів чи гостей

**Джерело: [23]*

Перевагою комплексних моделей є те, що вони беруть до уваги як кількісні, так і якісні характеристики позичальника. Комплексні методики оцінки мають ряд недоліків, попри те, що їх використовує велике число комерційних банків. До недоліків даних методів можна віднести їх «емпіричний» характер, акцент на суб'єктивну думку експертів, орієнтація здебільшого на якісні чинники.

Було подано характеристику класифікаційних та комплексних моделей оцінки кредитоспроможності позичальників комерційного банку. Це дозволяє виділити їх основні плюси та мінуси у порівнянні один із одним. Результати представлено у таблиці 1.9.

Таблиця 1.9

Переваги та недоліки класифікаційних та комплексних методів оцінки кредитоспроможності позичальників

Моделі	Суть	Переваги	Недоліки
Класифікаційні	Для оцінки беруться усі показники, які характеризують кредитоспроможності клієнта. Показники можуть бути проаналізовані у даний момент часу або ж за певний визначений період часу (в динаміці).	Об'єктивність результатів	Аналіз вимагає наявності значного масиву даних. Якісні чинники важко врахувати у даній моделі.
Комплексні	Для кожного показника визначається його значимість. Значимість виражено у балах.	Використання математичного інструментарію, що дозволяє швидко та із мінімальними витратами праці проаналізувати великий обсяг кредитних заявок. Використання як кількісним, так і якісних характеристик позичальника.	Визначення шкали оцінки. Вони вимагають постійного оновлення інформації. Не враховуються специфічні особливості окремих позичальників. Суб'єктивізм.

*Джерело: [23]

Якщо говорити про оцінку кредитоспроможності позичальника-фізичної особи, то доцільно обрати для оцінки кредитний скоринг. Однак, тут існує безліч методів та моделей. Далі пропонується розглянути суть, переваги та недоліки основних методів та моделей кредитного скорингу.

Виділяють три основні типи методів кредитного скорингу: статичні моделі, методи штучного інтелекту та гібридні методи. У таблиці 1.10 наведено основну інформацію та відмінності даних методик та виокремлено їх переваги та недоліки.

Оцінивши існуючі методики оцінки кредитоспроможності та перелічивши їхні основні недоліки, можна судити про доцільність створення такого підходу до оцінювання кредитоспроможності позичальників-юридичних осіб, що буде побудований на базі сучасних математичних методів та адаптований до українського банківського законодавства. Було вирішено за основу взяти метод CAMPARI та провести його адаптацію для можливості використання банківськими установами України.

Таблиця 1.10

Порівняльний аналіз методик кредитного скорингу

Метод	Суть	Переваги	Недоліки
Статистичні моделі			
LDA (лінійний дискримінантний аналіз)	Відноситься до класифікаційних методів. Використовує лінійну дискримінантну функцію для класифікації позичальників на два чи більше класів.	Широко вживана методика; простота у використанні.	Потребує нормального розподілу змінних; лінійний зв'язок між залежними та незалежними змінними; невисока точність результатів.
Логістична регресія	Деформація лінійної регресії.	Може використовуватися для якісних характеристик; не потребує нормального розподілу змінних.	Висока кореляція; невисока точність результатів.
MARS (сплайни багатовимірної адаптивної регресії)	Нелінійна та непараметрична регресія.	Більш гнучка у порівнянні із лінійною регресією, легка інтерпретація результатів, працює із дискретними та неперервними даними	Дає гірші результати у порівнянні із деревом рішень, не працює із пропущеними значеннями, не можна побудувати довірчий інтервал.
Модель Байеса	Статистична класифікація із принципом роботи «білий ящик».	Висока точність передбачення.	Кореляція змінних.

Продовження табл. 1.10

Дерева рішень	Ієрархічна деревовидна структура, що складається із правил типу «якщо..., то...»	Формують чіткі правила класифікації, візуалізація, швидко навчаються, підтримують числові та категоріальні ознаки.	Поступається якістю класифікації, можливість «перенавчання» дерева, важкий пошук оптимального дерева рішень.
Модель Маркова	Для прогнозу використовується попередня історія.	Простота моделювання	Вузька сфера застосування моделей
Методи штучного інтелекту			
Штучні нейронні мережі	Працює за принципом «чорної скриньки».	Висока швидкість навчання; відсутня кореляція між змінними.	Складність методу та важкість у тлумаченні результатів.
SVM (метод опорних векторів)	Основа – мінімізація структурних ризиків.	Немає обмежень по розподілу даних; хороша робастність;	Краще підходить для невеликих вибірок даних;
Генетичні алгоритми та програмування	Комп'ютерна модель, що симулює природній відбір, та покликана знайти оптимальне рішення.	Самоадаптивний метод; здатний працювати із великим масивом даних; висока робастність.	Висока витрата ресурсів; низька ефективність; розрахунки, які не гарантують отримання оптимального рішення.
Метод k-найближчого сусіда	Метод кластерного аналізу.	Стійкість до аномалій, простота реалізації, легка інтерпретація результатів.	Вимога репрезентативності даних, «невідокремленість» моделі від даних.

*Джерело: складено автором на основі [24,25,26,27]

РОЗДІЛ 2. ПОБУДОВА МОДЕЛІ ОЦІНКИ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКА-ФІЗИЧНОЇ ОСОБИ

2.1. Алгоритм побудови скорингових моделей

У попередньому розділі було розглянуто рейтингові моделі оцінки кредитоспроможності позичальників. Далі пропонується розглянути конкретні методи, що будуть застосовані для побудови моделей у наступних питаннях. Кожен із має схожі етапи побудови моделі:

1. Впорядковуємо початкову вибірку.
2. Розраховуємо статистичні характеристики для кожної змінної.
3. Розбиваємо її на навчальну та тестову вибірки.
4. Проводимо аналіз та коригування змінних моделі.
5. Будуємо модель.
6. Розраховуємо показники достовірності та точності моделі.

Гістограма статистичних характеристик із неперервним типом даних розраховується по формулі 2.1 [27]:

$$[1 + 3,2 * \lg(N)], \quad (2.1)$$

де N – загальне число спостережень.

Для оцінки достовірності моделей кредитного скорингу пропонується використовувати ROC-криву та коефіцієнт Джині. Достовірність характеризує здатність моделі розрізняти «хороших» позичальників від «поганих». ROC-крива (від англ. – Receiver Operator Characteristic) – показує залежність кількості правильно класифікованих позитивних випадків від кількості неправильно класифікованих негативних випадків. Числовий показник площі під кривою називається AUC (Area Under Curve) [50]. Чим більша є цей показник, тим кращу прогнозу силу має модель. Вважається, що значення площі від 0,9 до 1 відповідає відмінній якості моделі, від 0,8 до 0,9 – дуже хорошему, 0,7-0,8 – хорошему, 0,6-0,7 – середньому, 0,5-0,6 – незадовільному.

За значенням площі під ROC-кривою можна розрахувати індекс Джині. Цей показник переводить значення площі під кривою у діапазон від 0 до 1. Чим вища його величина, тим вища дискримінаційна здатність моделі. Коефіцієнт Джині розраховується за такою формулою [8]:

$$\text{Gini} = 2\text{AUC} - 1 \quad (2.2)$$

За $\text{Gini} = 0$ класифікація здійснюється випадковим чином, а при $\text{Gini} = 1$ класифікатор працює ідеально (існує такий поріг відсікання, за якого всі позичальники з розрахунком скорингової моделі, вищим за цей поріг, є насправді дефолтними, а всі позичальники, для яких розрахований моделлю рейтинг менше порогу, є надійними).

Для початку розглянемо логістичну регресія – різновид множинної регресії, що призначена для класифікації на основі значень вхідних полів. При цьому вихідна змінна є категоріальною або бінарною.

При бінарній класифікації кожний об'єкт або спостереження повинен бути віднесений до одного із двох класів (наприклад, А і Б). Результатом буде ймовірність того, що один об'єкт належить до класу А, а інший – до класу Б.

Логістична регресія – найбільш поширена статистична модель для побудови скорингових карт і представлена у вигляді наступної формули [50]:

$$\ln\left(\frac{p_i}{1-p_i}\right) = b_0 + b_1 x_i^{(1)} + b_2 x_i^{(2)} + \dots + b_k x_i^{(k)} + e_i, \quad (2.3)$$

де p_i – імовірність настання дефолту по кредиту для і-го позичальника; $x_i^{(j)}$ – значення j-ої незалежної змінної; b_0 – незалежна константа моделі; b_j – параметри моделі; e_i – компонент випадкової помилки.

Рівняння 2.3 відображає лінійну залежність імовірності настання прострочки по кредиту в залежності від значення незалежних змінних. Константа в моделі відображає природній рівень ризику настання модельованої події при рівності всіх незалежних змінних нулю. Значення коефіцієнтів при незалежних змінних використовують для побудови скорингової моделі.

Для побудови моделі необхідно зібрати достатню і якісну інформацію про позичальників банку. Від якості даних залежить точність прогнозування та успіх розробки моделі.

Також потрібно визначити залежну змінну, яка відображає мету побудови моделі. Мета може бути різною, як загальною (скорочення втрат за новими кредитними рахунками), так і конкретною (скорочення кількості неповернених кредитів за схваленими заявками).

Далі потрібно визначити незалежні змінні. Це можуть бути дані із кредитної заяви або анкети позичальника: соціально-демографічні (вік, сімейний стан, стать, посада і т.д.), інформація про запитуваний кредит (термін погашення, сума, розмір щомісячних платежів, мета кредиту і т.д.).

Наступний крок – формування навчальної та тестової вибірки. Це робиться для перевірки адекватності і точності передбачення скорингової моделі. Навчальна вибірка – спостереження, на основі яких буде будуватися модель. Тестова вибірка – спостереження, за якими відоме значення залежної змінної, але вони не беруть участі у побудові моделі, а використовуються для перевірки точності передбачення. Формують їх випадково, найчастіше у відношенні 70-80% і 30-20% відповідно від початкового об'єму вибірки.

Коли вже є сформовані дві вибірки, можна переходити до категоризації кількісних змінних.

Категоризацію здійснюється за алгоритмом:

1. Змінну розбивають на декілька груп;
2. У кожній групі рахують долю «поганих» і «хороших» кредитів;
3. Розраховують ваговий показник WOE (від англ. Weight of Evidence)

для кожної категорії за наступною формулою:

$$WOE_i = \ln \left(\frac{d_i^{(1)}}{d_i^{(2)}} \right), \quad (2.4)$$

де $d_i^{(1)}$ і $d_i^{(2)}$ – частка «поганих» і «хороших» кредитів відповідно в i -ї групі; $i = \overline{1, k}$, k – кількість категорій змінної.

4. Аналіз отриманих вагових показників, об'єднання сусідніх категорій з близькими значеннями WOE в одну категорію.

5. Обрахунок показників WOE повторюється.

При об'єднанні категорій керуються такими правилами:

- У кожній групі повинно бути не менше 5% від усіх спостережень;
- Не повинно бути груп із кількістю «поганих» і «хороших» кредитів рівних 0;
- Значення WOE мають мати зростаючий чи спадаючий тренд при переході від однієї категорій до іншої [28].

Наступною моделлю пропонується розглянути кластеризацію на основі k-means [29] та G-means [30]. Вхідними параметрами буде слугувати початкова вибірка та рівень значимості. Відмінним етапом цієї моделі буде ініціалізація множини центрів кластерів та їх поділ. Розподіл кластерів відбувається по гаусівському закону із заданим рівнем значимості.

Для побудови моделі потрібно задати параметри:

- рівень значимості, % - ймовірність істинності нульової гіпотези про те, що значення у вибірці розподілені за нормальним законом (використовується алгоритм G-means);
- кількість кластерів – потрібну кількість кластерів (використовується алгоритм k-means).

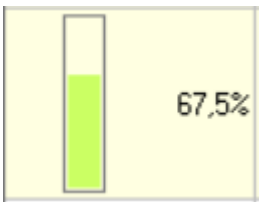


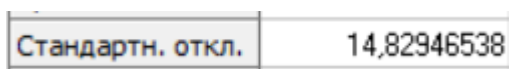
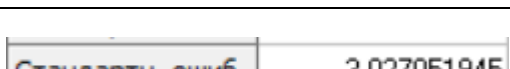
Основним фактором, що визначає склад кластера, є значимість властивостей, виражена у відсотках. Загальна значимість розглянутого поля визначається варіабельністю її розглянутих параметрів [31].

Значимість атрибутів показує їхній вплив на формування того чи іншого кластеру. Для неперервних і дискретних полів значимість визначається порізному. Для неперервних полів вона установлюється залежно від відхилення середнього значення розглянутої групи кластерів від загального середнього всієї вибірки (чим більше виражене дане відхилення, тим більше його значимість). Значимість для дискретних полів визначається наявністю індивідуальних відмінностей між розглянутими групами (чим більше виражені відмінності, тим

більше значимість). Для кожної розглянутої властивості в кластері обчислюється: довірчий інтервал, середнє, стандартне відхилення й стандартна помилка (таблиця 2.1).

Таблиця 2.1

Визначення показників кластерів

Показник	Приклад	Опис
Значимість		Виражається у відсотках. Для неперервних полів використовується t-критерій Стьюдента, а для дискретних – критерій хі-квадрат. Загальна значимість поля визначається за F-критерієм Фішера.
Довірчий інтервал		Графічно зображено 95% довірчого інтервалу для середнього значення кластера (темно-сіра область). Окрім цього показуються: <ul style="list-style-type: none"> • середнє значення кластера – червона лінія; • середнє значення усієї вибірки – синя штрих-пунктирна лінія.
Середнє		Середнє значення поля, розраховане для об'єктів, що належать кластеру.
Стандартне відхилення		Стандартне відхилення поля, розраховане для об'єктів, що належать кластеру.
Стандартна похибка		Стандартна похибка поля, розраховане для об'єктів, що належать кластеру.

*Джерело: [31]

Додатково пропонується розглянути алгоритм побудови та навчання нейронної мережі Кохонена, їх ще називають самоорганізованими картами. Вони дозволяють представити результати кластеризації у вигляді двомірних карт, де відстань між об'єктами відповідає відстані між їхніми векторами у багатомірному просторі, а власне значення ознак відображаються різними кольорами та відтінками [29, 32]. Відмінним етапом є навчання мережі Кохонена та побудова відповідної карти, у результаті чого спостереження розділяються на комірки.

Для налаштування та побудови карти необхідно вказати параметри:

- розмір осі X – кількість комірок по горизонталі;

- розмір осі Y – кількість комірок по вертикалі;
- спосіб початкової ініціалізації карти – випадковим чином, із навчальної вибірки чи із власних векторів;
- рівень значимості – параметр визначення кластерів (чим вище значення, тим більше кластерів);
- помилку, менше за яку приклад вважається нерозпізнаним, - відстань від вектору ознаки прикладу x до вектору ознаки найближчої комірки u ;
- епоху – один із варіантів зупинки навчання.

Результати подаються у вигляді карти Кохонена (приклад можна побачити на рис.2.15), де кожному нейрону відповідає свій вектор ознак, залежно від якого відповідна комірка має свій колір на карті [27].

Алгоритм функціонування самоорганізуючих карт (Self Organizing Maps – SOM) являє собою один із варіантів кластеризації багатомірних векторів – алгоритм проектування із збереженням топологічної схожості.

Важливою відмінністю алгоритму SOM є те, що усі нейрони (вузли, центри класів) впорядковані у певну структуру (двовірну мережу). При цьому, в ході навчання, модифікуються не лише нейрон-еталон (нейрон карти, який найбільше відповідає вектору вхідних параметрів та визначає, до якого класу віднести спостереження), але і його сусіди. При використанні цього алгоритму вектори, що є близькими до вихідного простору, будуть знаходитися поруч на отриманій карті [53].

Зазвичай, для представлення результатів використовують одно- або двовірні мережі. Кожен нейрон представляє собою n -мірний вектор-стовпець, де n визначається розмірністю вихідного простору. Нейрони розташовані у вузлах із прямокутними або шестикутними комірками. При цьому, нейрони взаємодіють між собою, тому величина цієї взаємодії визначається відстанню між нейронами на карті [33].

Загальний принцип цього способу аналізу полягає у тому, що нове спостереження проганяється через побудовану карту та потрапляє у певний

кластер, де визначається відношення «хороших» та «поганих» позичальників та на основі цього робиться висновок про ймовірність повернення кредиту [33].

Наступна модель, що пропонується для оцінки, - це дерева рішень. У першому розділі було подано та описано структуру даної моделі. Дерева рішень – є ефективним інструментом інтелектуального аналізу даних та прогнозної аналітики.

Для побудови моделі ми будемо використовувати модифікований алгоритм побудови дерев рішень на основі алгоритму C4.5 [34]. Відмінним етапом у побудові даної моделі буде розбиття вхідної множини T на дві - T_1 та T_2 . Тут важливо, що на першій ітерації у T_1 потрапить лише перше спостереження, а всі решта – у T_3 . Наступне розбиття отримується шляхом переміщення першого елемента із T_3 у T_1 .

Основна задача при побудові дерев рішень – розбити навчальну вибірку на підмножини із застосуванням правил у вузлах. Розбивати вибірку варто до тих пір, доки усі вузли не стануть листами [27].

Вузол стає листом у двох випадках [27]:

1. Природнім шляхом – коли він містить один об'єкт або об'єкт одного класу;
2. Після досягнення заданої алгоритмом умови.

Для налаштування вузла потрібно задати параметри:

- мінімальна кількість прикладів у вузлі, при якому створюється новий – використовується для побудови дерева, якщо у вузол потрапляє більше прикладів, то він видаляється;
- рівень довіри, що використовується при відсіченні вузлів – параметр використовується для оптимізації дерева після його побудови [34].

Для кожного вузла знаходиться додаткова кількість помилок, що і впливає на його формування.

Після побудови дерева класифікаційних правил можна розрахувати значимість атрибутів – показник, що характеризує наскільки сильно екзогенна

змінна залежить від кожної ендогенної. Формула для розрахунку значимості m -ого атрибуту має наступний вигляд [27]:

$$\text{Значимість}_m = \frac{\sum_{j=1}^{k_m} (E_{m,j} - \sum_{i=1}^{n_{m,j}} E_{m,j,i} * \frac{N_{m,j,i}}{N_{m,j}})}{\sum_{l=1}^g \sum_{j=1}^{k_l} (E_{l,j} - \sum_{i=1}^{n_{l,j}} E_{l,j,i} * \frac{N_{l,j,i}}{N_{l,j}})} * 100\%, \quad (2.5)$$

де g – кількість вхідних атрибутів, k – кількість вузлів, які були розбиті по атрибуту 1, $E_{l,j}$ – ентропія батьківського вузла, розбитого по атрибуту 1, $E_{l,j,i}$ – дочірній вузол для j -го, який був розбитий по атрибуту 1, $N_{l,j}$, $N_{l,j,i}$ – кількість прикладів у відповідних вузлах, $n_{l,j}$ – кількість дочірніх вузлів для j -го батьківського.

Для побудови моделей ми використаємо програмне забезпечення Deductor Studio. Початкова наша вибірка має 4446 спостережень та 13 змінних, що містять інформацію (сума кредиту, вік та сімейний стан позичальника, стаж роботи та ін.) по клієнтах банку, що були позичальниками у одному із американських банків [35].

Дані містять 13 факторних ознак: стаж роботи, термін погашення кредиту, вік, тип володіння житлом, сімейний стан, наявність записів, тип роботи, витрати, доходи, активи, борг, сума кредиту, сума повернення за кредитом.

2.2. Побудова логістичної регресії та скорингової карти.

Перейдемо до побудови логістичної регресії. Фрагмент даних, що використовуються для побудови моделі представлено на рис. 2.1. Дані містять 13 факторних ознак, із них неперервні – «Стаж роботи», «Термін погашення кредиту», «Вік», «Витрати», «Доходи», «Активи», «Борг», «Сума кредиту», «Сума повернення за кредитом», та дискретні – «Тип володіння житлом», «Сімейний стан», «Наявність записів», «Тип роботи».

Age	Marital	Records	Job	Expenses	Income	Assets	Debt	Amount	Price	Give credit?	Credit status [bool]
46	married	absent	freelance	90	200	3000	0	2000	2985	no	0
41	married	present	part-time	90	80	0	0	1200	1468	no	0
21	single	absent	part-time	35	50	0	0	400	500	no	0
36	married	present	part-time	45	130	750	0	1100	1511	no	0
25	single	present	fixed	46	107	0	0	1500	2189	no	0
45	married	present	part-time	105	112	2000	500	600	1332	no	0
23	married	present	fixed	75	85	5000	0	600	1600	no	0
23	single	present	part-time	45	122	2500	0	400	400	no	0
36	married	present	fixed	70	99	0	0	950	950	no	0
31	single	present	fixed	44	90	0	0	1300	1700	no	0
32	married	absent	fixed	45	142	7000	3000	1250	1566	no	0
43	married	present	part-time	75	71	3000	0	1500	1552	no	0
27	separated	present	fixed	48	128	0	0	450	545	no	0
29	married	absent	freelance	85	100	0	0	1000	1345	no	0
27	single	present	fixed	35	70	0	0	600	926	no	0
29	married	present	fixed	60	70	4500	0	850	1100	no	0
25	single	present	part-time	40	50	0	0	1040	1048	no	0
62	married	present	fixed	60	146	0	0	1000	1292	no	0
32	single	absent	fixed	35	105	0	0	1490	1490	no	0
46	married	present	fixed	45	138	0	0	250	270	no	0
29	married	present	part-time	78	180	0	0	650	650	no	0

Рис.2.1. Фрагмент генеральної сукупності для побудови моделі

Джерело: [35]

Для вибірки було побудовано гістограму статистичних характеристик, де для полів із дискретним типом даних число стовпців рівне числу унікальних значень, а для неперервних – розраховується по формулі 2.1. Тут ми можемо побачити об'єднання по категоріях кожного показника, який відсоток у кожній категорії конкретного значення показника, частоту, діапазон. Для прикладу наведено гістограму (рис.2.2) для змінної «Тип роботи». Решту показників було винесено у Додаток А.

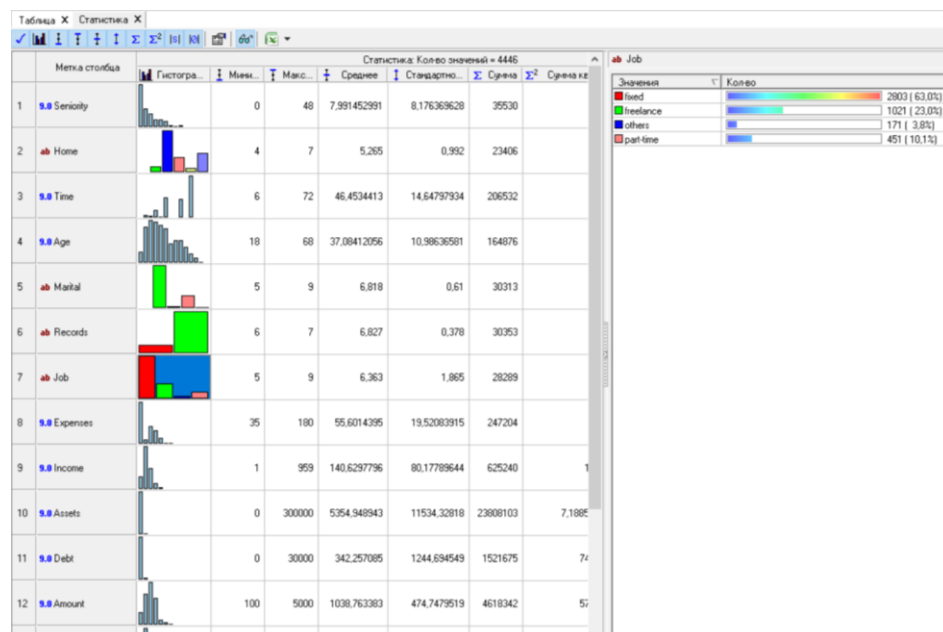


Рис.2.2. Статистика характеристик моделі для дискретної змінної «Тип роботи»

Оскільки дані для моделі бралися вже попередньо оброблені, то крок із підготовкою даних для моделі можна пропустити. Аномалій, пропущених або незаповнених значень, невірно заповнених комірок вибірка немає.

Наступним кроком пропонується здійснити кореляційний аналіз, щоб побачити чи багато у нас незначних факторів. Для вибору факторів використовується коефіцієнт кореляції Пірсона. Із порогом значимості 0,1 отримали результати представлені у таблиці 2.2. Як видно, аналіз показав, що змінні «Стаж роботи», «Час, на який береться кредит», «Наявність записів», «Тип роботи», «Доходи» та «Сума кредиту» мають найбільшу вагу для нашої вибірки.

Таблиця 2.2

Коефіцієнти кореляції для генеральної сукупності

Вхідні значення	Кореляція із вихідним значенням	Статус
Стаж роботи	0,260	Значимий
Тип володіння житлом	-0,053	Незначимий
Термін, на який береться кредит	-0,101	Значимий
Вік	0,095	Незначимий
Сімейний стан	-0,010	Незначимий
Наявність записів	-0,278	Значимий
Тип роботи	-0,153	Значимий
Витрати	-0,030	Незначимий
Доходи	0,144	Значимий
Активи	0,097	Незначимий
Борги	-0,010	Незначимий
Сума кредиту	-0,154	Значимий
Вартість	-0,010	Незначимий

**Джерело: розраховано автором на основі [35]*

Ми маємо дані, що не містять ПІБ позичальника. Однак, коли аналітик отримує початкові дані про потенційних позичальників, там обов'язково буде така інформація. Можливий випадок, коли інформація про ПІБ позичальника, його рахунок тощо буде містити дублікати, які потрібно усувати перед побудовою моделі. Для цього можна скористатися очисткою даних від дублікатів та протиріччя. Застосуємо даний аналіз для нашої вибірки. На рис.2.3 зображено окремо дублікати та протиріччя. Ми маємо два дублікати та одне протиріччя. Для нашої вибірки дублікат це нормально, оскільки у нас загальні дані, котрі можуть легко повторитися у більше ніж 4 тис спостережень. А ось

протиріччя – це вже цікаво. Вийшло, що ми маємо два спостереження із абсолютно ідентичними вхідними параметрами, однак із різним результатом: одним позичальник кошти повернув, а інший – ні.

Що робити із дублікатами та протиріччями, має вирішувати аналітик, що займається побудовою моделі. Ми вирішуємо не відкидати ні дублікати, ні протиріччя.

Дублікати				Протиріччя				Вхідні поля											Вихідні поля
Признак	Група	Признак	Група	Seniority	Home	Time	Age	Marital	Records	Job	Expenses	Income	Assets	Debt	Amount	Price	Give credit?		
<input type="checkbox"/>		<input checked="" type="checkbox"/>	1	1	rent	60	37	single	present	fixed	45	100	0	0	1150	1555	no		
<input checked="" type="checkbox"/>	2	<input type="checkbox"/>		5	owner	60	36	married	present	freelance	90	62	3000	0	650	1295	no		
<input checked="" type="checkbox"/>	2	<input type="checkbox"/>		5	owner	60	36	married	present	freelance	90	62	3000	0	650	1295	no		
<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	1	1	rent	60	37	single	present	fixed	45	100	0	0	1150	1555	yes		
<input checked="" type="checkbox"/>	1	<input type="checkbox"/>		4	owner	60	37	married	present	freelance	35	128	18000	0	800	1560	yes		
<input checked="" type="checkbox"/>	1	<input type="checkbox"/>		4	owner	60	37	married	present	freelance	35	128	18000	0	800	1560	yes		

Рис. 2.3. Аналіз дублікатів та протиріччів генеральної сукупності

Підготувавши дані, переходимо до етапу побудови власне самої моделі. Побудуємо дві логістичні моделі: одну для початкових даних, іншу – на даних, що ми отримали після кореляційного аналізу.

Спочатку будуємо модель для вибірки із вхідними параметрами без кореляційного аналізу.

Формуємо навчаючу та тестову вибірку у відношенні 80:20% відповідно від початкового об'єму вибірки. Виходить, що навчаюча вибірка міститиме 3557, а тестова – 889 спостережень. Пропорції «хороших» і «поганих» позичальників потрібно залишити такими ж. Виходить, що у навчаючій вибірці «хороших» заявок буде 2558, «поганих» - 999, а у тестовій вибірці – 639 і 250 відповідно.

Параметри для побудови моделі налаштуємо наступним чином:

- метод відбору змінних ми залишаємо як повне включення;
- максимальне число ітерацій 600;
- подія: схвалення кредиту;
- поріг відсіювання - 0,75;
- референсна точка (число балів, при яких буде досягтися потрібне відношення шансів) визначено як 600, відношення шансів – 50.

В результаті було отримано коефіцієнти, що представлені у таблиці 2.3.

Таблиця 2.3.

Статистики логістичної регресії для навчаючої вибірки, що застосовуються для важливості предикторів

Показники /Статистика	<i>coeff b</i>	<i>s.e.</i>	<i>Wald</i>	<i>p-value</i>	<i>exp(b)</i>
Константа	0,843779	0,053941	244,6879	3,74E-55	2,325136
WOE(стаж роботи)	-0,80796	0,078988	104,6322	1,47E-24	0,445765
WOE(тип володіння житлом)	-0,51572	0,118388	18,97654	1,32E-05	0,597068
WOE(термін погашення)	-0,36552	0,15824	5,33574	0,020892	0,693835
WOE(вік)	0,820453	0,255291	10,32845	0,00131	2,271528
WOE(сімейний стан)	-0,58476	0,221367	6,978074	0,008251	0,557237
WOE(наявність записів)	-1,14479	0,079508	207,3145	5,29E-47	0,318289
WOE(тип роботи)	-0,73651	0,090164	66,7242	3,12E-16	0,478783
WOE(витрати)	-1,18717	0,255543	21,58225	3,39E-06	0,305083
WOE(доходи)	-0,91219	0,094354	93,46388	4,14E-22	0,401645
WOE(активи)	-0,61009	0,112805	29,2501	6,36E-08	0,543303
WOE(борг)	-8,79109	1,840951	22,80346	1,79E-06	0,000152
WOE(величина кредиту)	-1,13346	0,121082	87,62994	7,89E-21	0,321918
WOE(сума погашення)	0,028133	0,20091	0,019608	0,888637	1,028533

*Джерело: розраховано автором на основі [35]

Перший стовпчик – це коефіцієнти логістичної регресії, а другий – стандартна помилка. Статистика Вальда (третій стовпчик) – це критерій значимості коефіцієнта для відповідного предиктора. Чим більше його значення, тим більша значимість.

Використовуючи формулу 2.3, виведемо формулу моделі логістичної регресії:

$$\ln\left(\frac{p_i}{1-p_i}\right) = 0,8438 - 0,8080x_1 - 0,5157x_2 - 0,3655x_3 + 0,8205x_4 - 0,5848x_5 - 1,1448x_6 - 0,7365x_7 - 1,1872x_8 - 0,9122x_9 - 0,6101x_{10} - 8,7911x_{11} - 1,1335x_{12} + 0,0281x_{13} \quad (2.6)$$

Значення моделі потрібно коригувати аналітику в залежності від мети: чи пріоритетом є надання кредиту як найменшій кількості «поганих»

позичальників, чи уникнення помилки відмови у кредитуванні для «хорошого» позичальника. Для прикладу ми пропонуємо побудову двох логістичних регресій із підбором параметрів. У першому випадку пріоритетом було уникнення помилки відмови у кредитуванні «хорошим» позичникам» (результати представлені на рис.2.4). У другому випадку переслідувалася мета уникнення надання кредиту «поганим» позичальникам. (результати на рис.2.5).

Give credit?			
Класифіцировано			
Фактически	no	yes	Итого
no	919	330	1249
yes	698	2499	3197
Итого	1617	2829	4446

Рис.2.4. Таблиця спряженості для випадку зменшення помилок I роду

Give credit?			
Класифіцировано			
Фактически	no	yes	Итого
no	1015	234	1249
yes	970	2227	3197
Итого	1985	2461	4446

Рис.2.5. Таблиця спряженості для випадку зменшення помилок II роду

Тепер розглянемо отримані результати після побудови моделі, із врахування зменшення помилок I роду.

На рис.2.6 представлено коефіцієнти регресії, що дозволяє нам судити про вагу кожного показника у формуванні загального балу. Як видно, позитивними вважаються високий стаж роботи, тип володіння житлом: власник, середній вік, сімейний стан: одружений, короткостроковість кредиту, наявність записів, вищі доходи, менші витрати, менша сума кредиту, із типів роботи кращим є повний робочий день. Варто звернути увагу на значимість кожного показника. За формулою 2.3 можна робити наступні висновки:

- Сильну прогнозну здатність мають такі незалежні змінні: тип володіння житлом, вік;
- Середню: сімейний стан;

- Слабку: вік, період кредитування;
- Не мають прогнозної здатності: активи, борги, стаж роботи, наявність записів, тип роботи.

Модель є статистично значимою ($P > 0,0001$), як і більшість коефіцієнтів. Також тут розраховано показник «Відношення шансів» (OR) - це відношення ймовірності того, що подія відбудеться до ймовірності того, що подія не відбудеться.

Зведений звіт по регресії представлено у додатку Б.

Атрибут	Коефіцієнт	Стандартна помилка	Коефіцієнт Вальда	Значимість	Отношение шансов	Нижняя граница ДИ	Верхняя граница ДИ	Балл
9.0 <Константа>	-0,4152051106							413,442797
9.0 Age	-0,01291690433	0,005576051307	5,366156856	0,02053116565	0,9871661608	0,9764361036	0,9980141307	-0,3727030763
9.0 Amount	-0,001912804...	0,0001930085433	98,21737133	3,748924606E-23	0,9980890233	0,9977115209	0,9984666686	-0,055191884...
9.0 Assets	3,043768571E-5	7,721219443E-6	15,54001614	8,077710714E-5	1,000030438	1,000015304	1,000045572	0,000878245...
9.0 Debt	-0,000164160...	4,047672874E-5	16,44849485	4,998985628E-5	0,999835853	0,9997565348	0,9999151776	-0,004736668...
9.0 Expenses	-0,01455819595	0,00293675832	24,57416135	7,150411986E-7	0,9855472622	0,9798907019	0,9912364759	-0,4200607419
ab Home = ignore	0							0
ab Home = others	0,01679481839	0,6449325367	0,0006781441952	0,9792244752	1,016936644	0,2872871751	3,599743491	0,484596024
ab Home = owner	0,9326282833	0,6293405863	2,196067491	0,1383633242	2,541179347	0,7401671308	8,724505868	26,90996399
ab Home = parents	0,7728708262	0,6388788282	1,46344645	0,2263825496	2,165975476	0,6191970771	7,576665518	22,30033816
ab Home = privat	0,1185466772	0,6476978651	0,03349915402	0,854776252	1,125859425	0,3163389105	4,006966586	3,420534066
ab Home = rent	0,2603681441	0,6336357303	0,1688481775	0,681137504	1,297407632	0,3747268617	4,491982655	7,512636606
9.0 Income	0,00704932829...	0,0007917156003	79,27778001	5,39647343E-19	1,007074188	1,005512659	1,008638141	0,2033993109
ab Job = fixed	0							42,24235905
ab Job = freelance	-0,7240242223	0,1140964472	40,26817515	2,213861316E-10	0,4847973927	0,3876496703	0,6062910148	21,35143595
ab Job = others	-0,6428636724	0,2219479026	8,389492288	0,003773963098	0,5257845913	0,3403157005	0,8123323023	23,6932344
ab Job = part-time	-1,464008604	0,1404830116	108,6024544	1,983311475E-25	0,231307195	0,175633594	0,3046286147	0
ab Marital = divorced	0							19,45743305
ab Marital = married	0,4215475326	0,4686024792	0,8058094221	0,369362104	1,524318664	0,6072144068	3,826568282	31,62072375
ab Marital = separated	-0,6743432431	0,5186378942	1,690571448	0,1935256569	0,5094909224	0,1843584136	1,408023615	0
ab Marital = single	0,04336578255	0,4746807685	0,008346244478	0,9272083059	1,044319819	0,411886059	2,647829079	20,70870504
ab Marital = widow	-0,3019015883	0,5874099246	0,264148553	0,6072842616	0,739410828	0,233815008	2,338294609	10,74639457
9.0 Price	0,00090933377...	0,0001428355078	40,53013478	1,936069083E-10	1,000909751	1,000629578	1,001190003	0,02623794043
ab Records = absent	0							0
ab Records = present	1,741150588	0,1119623176	241,8404309	1,561047599E-54	5,703902491	4,580024949	7,103564715	50,23898638
9.0 Seniority	0,09021503185	0,008545282653	111,4563839	4,700380687E-26	1,094409591	1,076232243	1,112893951	2,603055581
9.0 Time	-0,001709152...	0,003911574531	0,1909231178	0,6621494227	0,9982923071	0,9906679773	1,005975315	-0,049315722...

Рис.2.6. Коефіцієнти регресії логістичної регресії 1

Фрагмент результатів по розрахунку балів та віднесенню позичальників генеральної сукупності до одного із класів подано на рис.2.7. Тут можна аналізувати і приклад помилкового віднесення позичальника не у потрібну групу. Найважливішими для нас є чотири останні стовпці:

- give credit? – вихідна змінна, де no – означає, що у кредитуванні варто відмовити, а yes – кредитування варто схвалити;
- give credit?_OUT – модельоване значення кредитоспроможності позичальника, отримане через логістичну регресію;
- give credit?_Ймовірність події – ймовірність настання події «yes», тобто надання кредиту позичальнику;

- give_credit?_Бал – бал, що отримав позичальник шляхом моделювання його кредитоспроможності.

Seniority	Home	Time	Age	Marital	Records	Job	Expenses	Income	Assets	Debt	Amount	Price	Give credit?	Credit status (bool)	Give credit?_OUT	Give credit?_Вероятность события	Give credit?_Бал
48	owner	36	62	married	present	freelance	60	76	25000	0	180	350	yes	1	yes	0.997044714145399	655.087210441862
47	owner	48	61	married	present	fixed	35	110	5500	30	1200	1200	yes	1	yes	0.994905384020783	639.31171856652
45	others	48	62	married	present	fixed	45	100	0	0	450	935	yes	1	yes	0.993175273450917	630.825542044105
45	owner	30	64	married	present	freelance	45	157	20000	0	1000	1598	yes	1	yes	0.996761536977816	652.438782554381
45	owner	12	60	married	present	freelance	35	52	19000	0	550	1240	yes	1	yes	0.996767067049018	652.488256313315
43	owner	60	57	married	absent	fixed	45	283	8000	0	900	1168	no	0	yes	0.992632748964534	628.602673005262
43	others	24	60	married	present	fixed	60	185	0	0	319	1669	yes	1	yes	0.997903998801962	665.025223787675
40	owner	60	60	married	present	fixed	45	232	5000	0	1130	1275	yes	1	yes	0.996737613144702	652.225718075371
41	others	24	59	married	present	fixed	60	177	0	0	500	1633	yes	1	yes	0.996176523792212	647.630330314142
40	owner	36	62	married	absent	freelance	45	157	13500	0	1000	1300	no	0	yes	0.956604390293471	576.368904474001
40	rent	18	62	married	present	freelance	77	166	10000	0	900	2040	yes	1	yes	0.98952181802443	618.348300000921
40	owner	48	60	married	present	freelance	75	183	5000	0	1200	1308	yes	1	yes	0.981146719328369	601.154504133535
40	owner	24	60	married	present	freelance	45	80	7500	0	325	1365	yes	1	yes	0.995951512675039	645.973850929923
40	owner	24	63	married	present	fixed	45	250	55000	0	1000	1751	yes	1	yes	0.999627409510708	714.914554129979
40	owner	36	59	married	present	freelance	35	184	17000	3000	500	1571	yes	1	yes	0.997590032502332	660.988667218874
40	others	60	61	married	absent	fixed	60	242	0	0	1000	1100	yes	1	yes	0.934695474890377	563.90770702975
40	owner	60	54	married	absent	fixed	45	220	13500	0	1543	1668	yes	1	yes	0.973902189522597	591.558400365083
40	owner	36	58	married	present	freelance	45	51	26000	0	1500	2554	yes	1	yes	0.991013493337471	622.822889400065
40	owner	36	58	married	present	freelance	75	214	50000	0	1000	1160	yes	1	yes	0.997083881199737	655.473307671672
40	owner	42	61	married	present	fixed	90	275	12000	0	1000	1335	yes	1	yes	0.99674315863551	652.274966958326
40	rent	36	59	married	present	fixed	59	142	0	0	1000	1800	yes	1	yes	0.990612116135754	621.550405439247
40	owner	24	60	married	present	freelance	75	186	8000	0	900	1307	yes	1	yes	0.990792144467405	622.114344636052
40	owner	24	62	married	present	freelance	60	130	15000	0	250	758	yes	1	yes	0.995650726409355	643.897305407215
39	owner	60	58	married	present	fixed	35	113	3500	0	1380	1623	yes	1	yes	0.989814236052378	619.173510170502
38	owner	48	56	widow	present	fixed	35	112	3000	0	850	960	yes	1	yes	0.985194397257633	608.246742080266
38	rent	12	66	married	present	fixed	79	340	0	0	700	1568	yes	1	yes	0.997258606222757	657.261161901542
38	rent	60	57	married	present	fixed	56	210	0	0	1350	1520	yes	1	yes	0.983080376120293	604.333675650257
38	owner	24	56	married	absent	fixed	75	98	7000	0	300	1308	yes	1	yes	0.982542542059717	603.414965653237
38	owner	48	57	married	present	freelance	45	58	12000	0	1500	3262	yes	1	yes	0.991690482109599	625.102506272516
37	parents	36	52	married	present	fixed	60	240	0	0	1510	1625	yes	1	yes	0.98363664630695	617.91143298337
37	owner	60	56	widow	present	fixed	35	300	10000	0	1300	1750	yes	1	yes	0.995860581743646	645.3303150637
37	owner	48	63	married	present	fixed	60	150	3500	0	700	808	yes	1	yes	0.991905854453685	625.866488082698
37	owner	60	52	married	present	fixed	45	324	7000	0	1550	1590	yes	1	yes	0.996191716828645	647.74565172309

Рис.2.7. Фрагмент таблиці розрахованих балів для кожного спостереження

Розглянемо таблицю спряженості, що будується на основі результатів класифікації моделі та фактичної приналежності прикладів до класів. Розпочнемо із результатами навчальної вибірки (таблиця 2.4).

Таблиця 2.4

Таблиця спряженості для навчальної вибірки, кількість позичальників

Позичальники	Класифіковані за допомогою логістичної регресії 1		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
Фактичні значення			
У видачі кредиту відмовлено	800	199	999
Схвалено надання кредиту	713	1845	2558
Всього позичальників	1513	2044	3557
Ассурасу	0,8008	0,7213	-
Поріг відсічення	0,75		
Подія	Схвалення кредиту		

З таблиці 2.4 можна зробити наступні висновки:

- 1845 – правильно класифіковані «хороші» клієнти;

- 800 – правильно класифіковані «погані» клієнти;
- 713 – «хороші» клієнти класифіковані як «погані» (похибка I роду);
- 199 – «погані» клієнти, класифіковані як «хороші» (похибка II роду).

Показник Ассурасу визначає точність класифікації, тобто це співвідношення правильних прогнозів до загальних прогнозів.

Розглянемо результати для тестової вибірки (таблиця 2.5).

Таблиця 2.5

Таблиця спряженості для тестової вибірки

Позичальники	Класифіковані за допомогою логістичної регресії 1		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	184	66	250
Схвалено надання кредиту	147	492	639
Всього позичальників	331	558	889
Ассурасу	0,7360	0,7700	-
Поріг відсічення	0,75		
Подія	Схвалення кредиту		

З таблиці 2.5 можна зробити наступні висновки:

- 492 – правильно класифіковані «хороші» клієнти;
- 184 – правильно класифіковані «погані» клієнти;
- 147 – «хороші» клієнти класифіковані як «погані» (похибка I роду);
- 66 – «погані» клієнти, класифіковані як «хороші» (похибка II роду).

Як бачимо, початкова та тестова вибірки показали дуже схожі результати по точності класифікації. Навчальна вибірка краще класифікує «поганих» позичальників, а тестова – «хороших».

Розглянемо результати для всієї сукупності загалом (таблиця 2.6).

Таблиця 2.6

Таблиця спряженості для генеральної сукупності

Позичальники	Класифіковані за допомогою логістичної регресії 1		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	984	265	1249
Схвалено надання кредиту	860	2337	3197
Всього позичальників	1844	2602	4446
Ассигасу	0,7878	0,7310	-
Поріг відсічення	0,75		
Подія	Схвалення кредиту		

З таблиці 2.6 можна зробити наступні висновки:

- 2337 – правильно класифіковані «хороші» клієнти;
- 984 – правильно класифіковані «погані» клієнти;
- 860 – «хороші» клієнти класифіковані як «погані» (похибка I роду);
- 265 – «погані» клієнти, класифіковані як «хороші» (похибка II роду).

Тобто, у загальному модель добре визначає помилки I та II родів та допомагає зменшити кількість кредитів, наданих »поганим» позичальникам, та збільшити частку коректної класифікації «хороших» позичальників. У середньому виходить 75% коректно визначених випадків.

Наступним кроком необхідно оцінити достовірність моделі логістичної регресії за допомогою ROC-кривої. Як видно із рис.2.8, числовий показник AUC для навчальної вибірки дорівнює 0,84, а для тестової – 0,83. Це свідчить про високу прогностну силу моделі для обох вибірок.

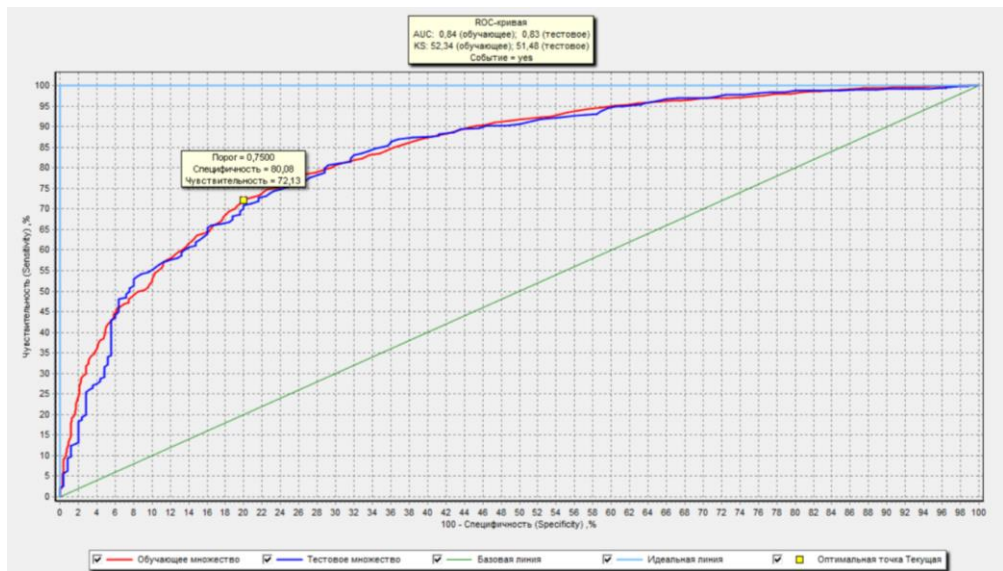


Рис.2.8. ROC-крива для логістичної регресії 1.

За значенням площі під ROC-кривою можна розрахувати індекс Джині (рис.2.9).

За формулою (2.2) коефіцієнт Gini дорівнює 0,67 та 0,66 для початкової та тестової вибірок відповідно. Це свідчить про майже рівноцінний розподіл між «хорошими» та «поганими» позичальниками у двох вибірках. Показник вказує на хорошу дискримінаційну здатність моделі.

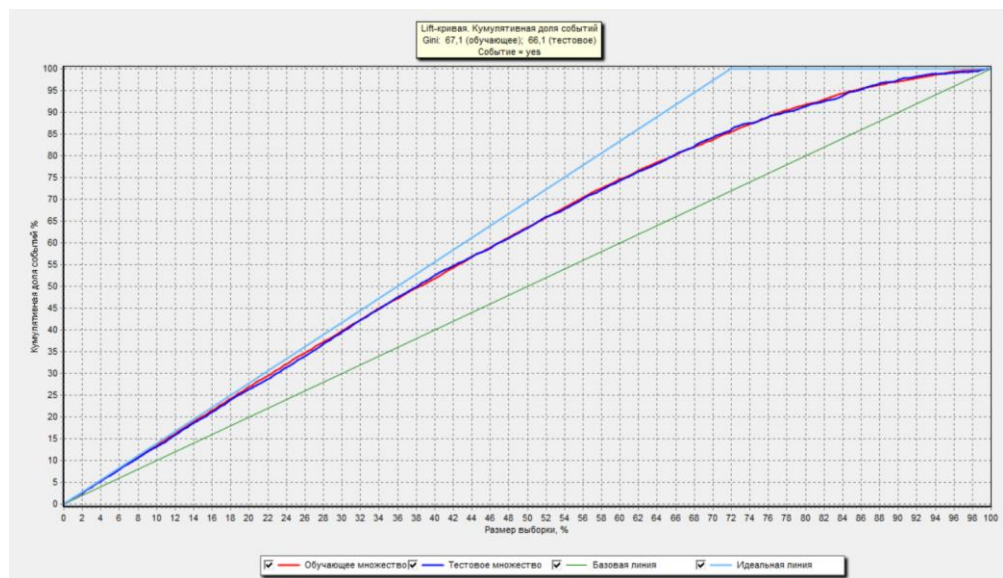


Рис.2.9. Lift-крива для логістичної регресії 1

Далі побудуємо логістичну регресію для вибірки із корегованими входними параметрами після кореляційного аналізу.

Параметри моделі лишаються як і для випадку логістичної регресії 1.

У таблиці 2.7 представлено коефіцієнти регресії, що дозволяють нам судити про вагу кожного показника у формуванні загального балу. Як видно, позитивними вважаються високий стаж роботи, короткостроковість кредиту, наявність записів, вищі доходи та менша сума кредиту, із типів роботи кращим є повний робочий день. Варто звернути увагу на значимість кожного показника. Модель не є статистично значимою ($P < 0,0001$), як і більшість коефіцієнтів. Також тут розраховано показник «Відношення шансів».

Таблиця 2.7

Коефіцієнти регресії для логістичної регресії 2

Показник	Коефіцієнт	Стандартна помилка	Коефіцієнт Вальда	Значимість	Відношення шансів	Бал
Константа	-0,08					443,62
Стаж роботи	0,09	0,01	142,81	6,4E-33	1,11	2,62
Час	-0,01	0,003	10,75	0,001	0,99	-0,33
Наявність записів						
- відсутні	0					0
- наявні	1,71		263,82	2,5E-59	5,55	49,44
Тип роботи:						
- повний робочий день	0					41,17
- фріланс	-0,60	0,11	32,26	1,3E-8	0,55	23,76
- інше	-0,54	0,19	7,68	0,01	0,40	25,48
- неповний робочий день	-1,43	0,13	112,33	3,02E-26	0,18	0
Доходи	0,01	0,001	75,84	3,07-E18	1,00	0,18
Сума кредиту	-0,001	0,0001	55,40	9,84-E14	0,99	-0,02

Для логістичної регресії 2 побудуємо таблицю спряженості для кожної із вибірок (табл.2.8). Вона допоможе нам розглянути як точність класифікації позичальників до потрібного класу, так і, власне, кількість позичальників, що були віднесені до того чи іншого класу фактично та за допомогою моделі.

Таблиця 2.8

Зведена таблиця спряженості вибірок для логістичної регресії 2

Позичальники	Класифіковані за допомогою регресії		
Фактичні значення	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
Навчальна вибірка			
У видачі кредиту відмовлено	756	243	999
Схвалено надання кредиту	814	1744	2558
Всього позичальників	1570	1987	3557
Ассигасу	0,7568	0,6818	-
Поріг відсічення	0,75		
Подія	Схвалення		
Тестова вибірка			
У видачі кредиту відмовлено	168	82	250
Схвалено надання кредиту	140	499	639
Всього позичальників	308	581	889
Ассигасу	0,6720	0,7809	-
Поріг відсічення	0,75		
Подія	Схвалення		
Генеральна сукупність			
У видачі кредиту відмовлено	924	325	1249
Схвалено надання кредиту	954	2243	3197
Всього позичальників	1878	2568	4446
Ассигасу	0,7398	0,7016	-
Поріг відсічення	0,75		
Подія	Схвалення		

З таблиці 2.8 можна зробити наступні висновки для навчальної вибірки:

- 1744 – правильно класифіковані «хороші» клієнти;
- 756 – правильно класифіковані «погані» клієнти;
- 814 – «хороші» клієнти класифіковані як «погані» (похибка I роду);
- 243 – «погані» клієнти, класифіковані як «хороші» (похибка II роду).

У висновку до таблиці можна вказати, що представлена модель коректно класифікує понад 70% спостережень. Як бачимо, навчальна та тестова вибірки

показали дуже схожі результати по точності класифікації. Тестова вибірка показує кращу точність при визначенні помилок II роду, ніж першого.

Наступним кроком необхідно оцінити достовірність моделі логістичної регресії 2 за допомогою ROC-кривої. Як видно із рис.2.10 числовий показник AUC для навчальної та тестової вибірки однаковий і дорівнює 0,8. Це свідчить про дуже хорошу прогностну силу моделі для обох вибірок.

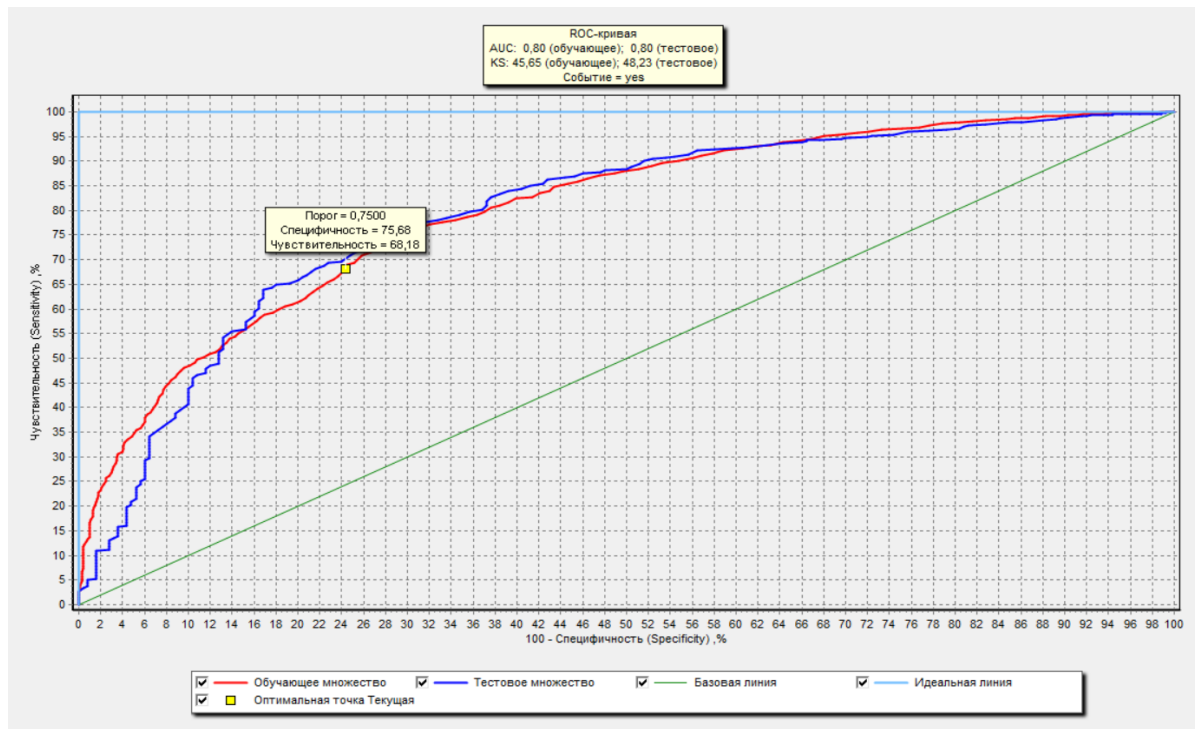


Рис.2.10. ROC-крива для логістичної регресії 2

За значенням площі під ROC-кривою можна розрахувати індекс Джині (рис.2.11).

За формулою 2.2 коефіцієнт Gini дорівнює 0,60 та 0,59 для навчальної та тестової вибірок відповідно. Це свідчить про майже рівноцінний розподіл між «хорошими» та «поганими» позичальниками у двох вибірках. Даний показник свідчить про хорошу дискримінаційну здатність моделі.

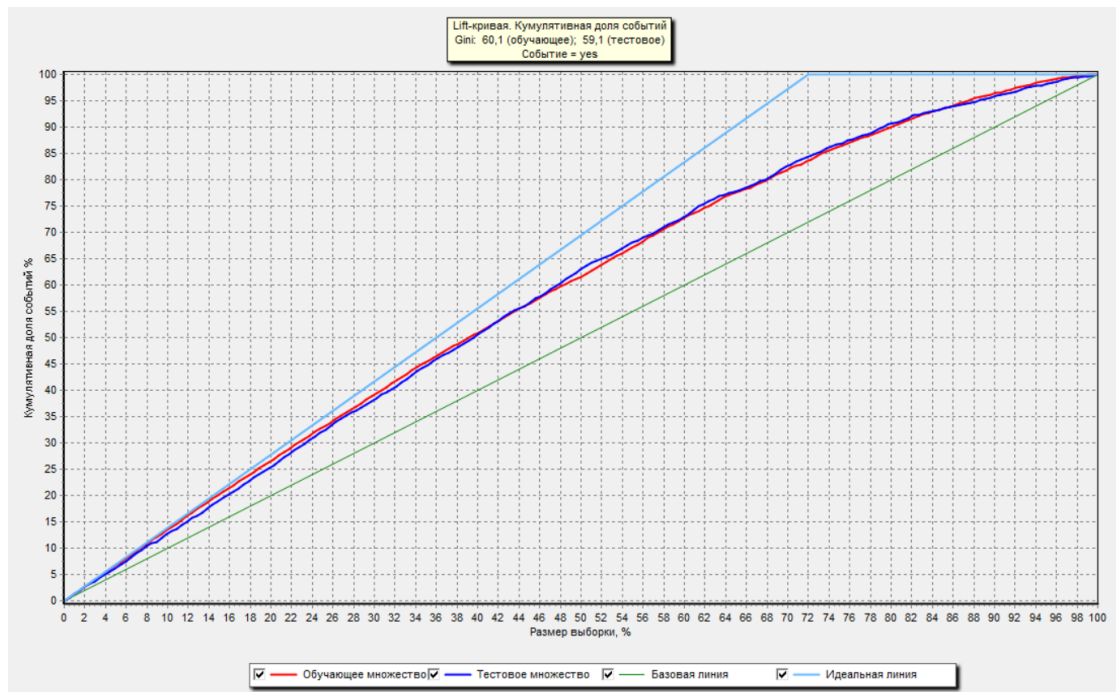


Рис.2.11. Lift-крива для логістичної регресії 2

Побудувавши дві моделі логістичної регресії, першу із всіма 13-ма показниками, а другу із значимими 6-ма, можна порівняти дані моделі та визначити, котра ж дає кращі результати. Зведені результати класифікації обох моделей для кожної із вибірок подано у таблиці 2.9.

Таблиця 2.9

Зведені результати класифікації позичальників із використанням логістичної регресії

Вибірка	Коректно класифіковано, %	Некоректно класифіковано, %	AUC	Gini
Логістична регресія 1				
Навчальна	74,36	25,64	0,84	0,67
Тестова	74,58	25,42	0,83	0,66
Генеральна	75,98	24,02		
Логістична регресія 2				
Навчальна	72,90	27,1	0,80	0,60
Тестова	71,88	28,12	0,80	0,59
Генеральна	72,69	27,31		

За результатами, представленими у таблиці 2.9, робимо висновок, що кращу прогнозу здатність має перша модель. Тобто кореляційний аналіз не врахував усі значимі коефіцієнти. Звісно, кожен із моделей можна покращувати далі, шукати кращий поріг відсічення та кращу точність моделі. Однак, 75% вірно класифікованих позичальників можна вважати дуже хорошим показником.

2.3. Побудова моделей оцінювання кредитного ризику за технологією машинного навчання.

У даному розділі пропонується порівняти три моделі (дерева рішень, кластеризацію та карти Кохена), що було описано у п.2.1 та побудовано у програмному середовищі Deductor із результатами, отриманими у п.2.2, та визначити який алгоритм дає кращі результати при оцінці кредитоспроможності позичальників комерційного банку.

Дані та їхню підготовку до побудови кожної моделі використовуємо аналогічно як і у п. 2.2. Поділ на навчальну та тестову вибірки теж буде дорівнювати 80% та 20% відповідно.

Дерева рішень. Шляхом перебору та оцінки параметрів навчання дерева (приклади інших варіантів подано у Додатку В) рішень було прийнято встановити наступні параметри:

- Мінімальна кількість прикладів у вузлі, при якому буде створено новий = 10;
- Будувати дерево із більш достовірними правилами попри компактність дерева: false;
- Відсікати вузли дерева: false;
- Рівень довіри, що використовується при відсіченні вузлів дерева: disable

Побудована модель виявила правила, за якими у майбутньому можна приймати рішення: давати чи ні кредит позичальнику. Правила представлені у вигляді дерева (рис.2.12).

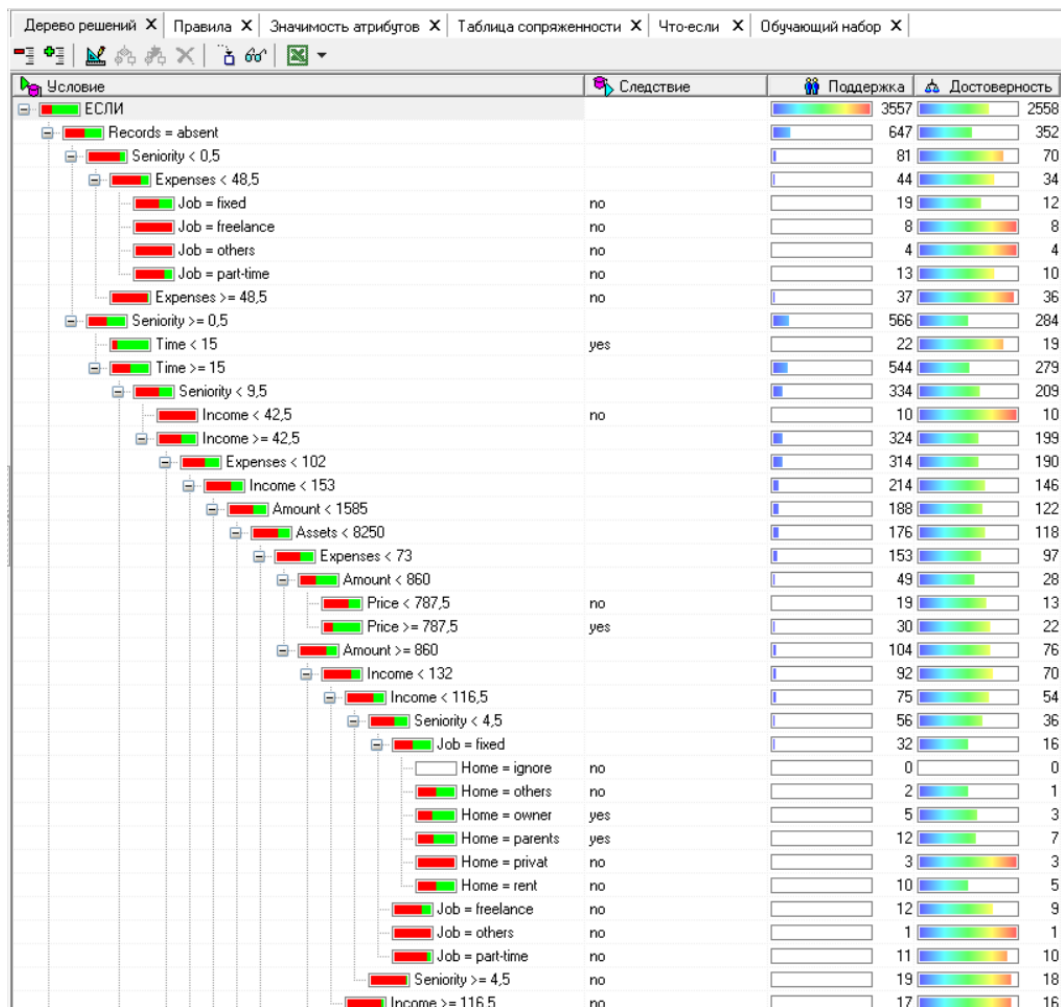


Рис.2.12. Фрагмент дерева рішень

Із рис.2.12 можна переглянути механізм отримання детальної інформації по окремому вузлу за подіями схвалення або відмови по кредиту. Тут міститься інформація про структуру позичальників, що підпадають кожне із правил, а також - сформовані правила.

Для прикладу пропонується розглянути правила із подією «no» (відхилення заявки по кредиту), що представлено на рис. 2.13. Було сформовано наступне правило для гілки: якщо [Записи відсутні І стаж роботи $\geq 0,5$, але $< 9,5$ І Період кредитування ≥ 15 І Доходи $\geq 42,5$ І Витрати < 102 І Доходи < 153 І Сума кредиту ≥ 1585], то [відмовити у кредитуванні]. Під це правило підходить 24 (або 92,3%) спостереження із такими характеристиками.

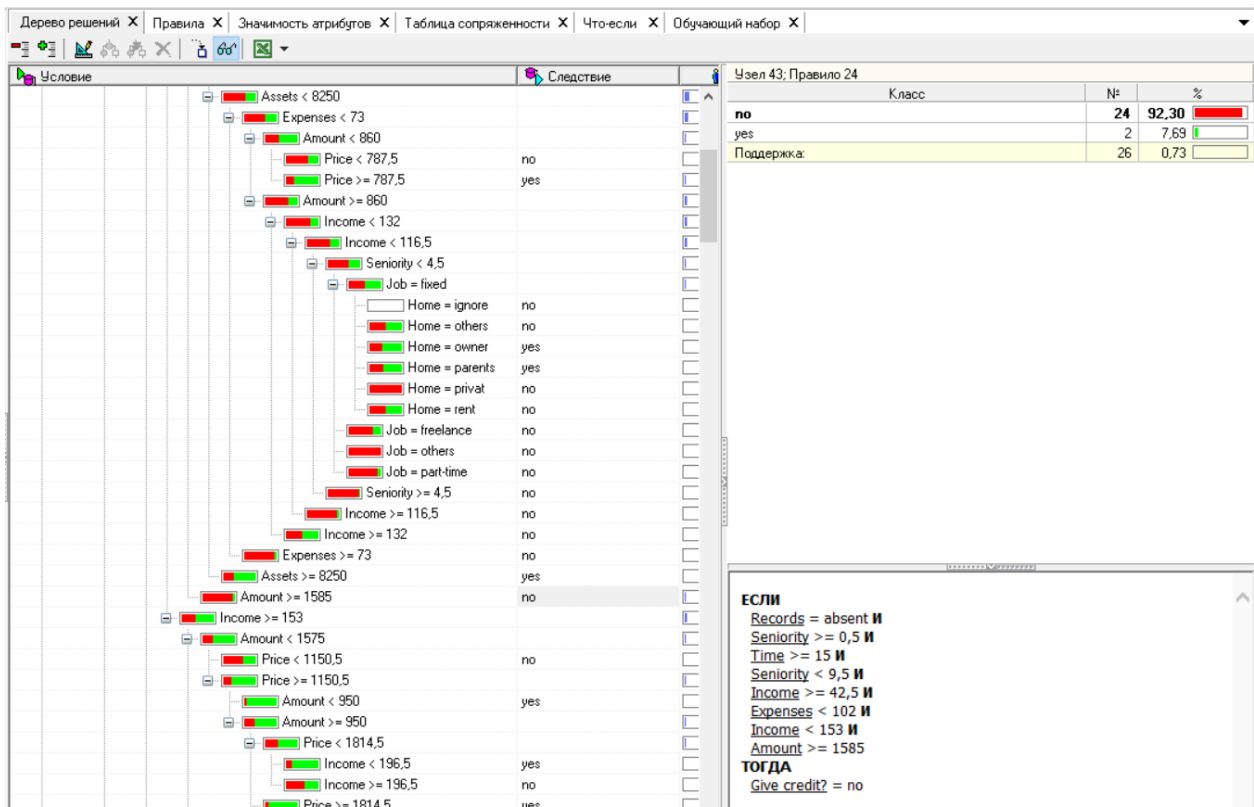


Рис.2.13. Приклад правила для випадку відмови у кредитуванні позичальнику

Правила об'єднані у окрему таблицю (Додаток Г). Загалом ми отримали 172 правила, за якими відбувається оцінка кредитоспроможності позичальників-фізичних осіб.

За формулою 2.5 ми отримали табличку 2.10 із значимістю кожного атрибуту.

Таблиця 2.10

Значимість атрибутів для моделі дерева рішень

Номер параметра	Параметр	Значимість, %
1	Стаж роботи	20,33
9	Доходи	15,34
6	Наявність записів	14,52
12	Сума кредиту	10,71
13	Вартість кредиту	8,06
8	Витрати	6,84
10	Активи	6,64
7	Тип роботи	6,46
4	Вік	3,74
3	Період кредитування	2,81
2	Тип володіння житлом	2,40
5	Сімейний стан	1,35
11	Борги	0,80

Отже, найбільший вплив мають змінні «Стаж роботи», «Доходи» та «Сума кредиту», а найменше впливають на результат – «Сімейний стан», «Тип володіння житлом» та «Борги» позичальника.

Розглянемо таблицю спряженості, що будується на основі результатів класифікації моделі та фактичної приналежності прикладів з навчальної вибірки до класів (таблиця 2.11).

Таблиця 2.11

Таблиця спряженості для навчаючої вибірки

Позичальники	Класифіковані за допомогою дерева рішень		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	687	312	999
Схвалено надання кредиту	259	2299	2558
Всього позичальників	946	2611	3557
Ассурасу	0,6877	0,8987	-
Параметри	(10; false; false; 0)		

Розглянемо результати для тестової вибірки (таблиця 2.12).

Таблиця 2.12

Таблиця спряженості для тестової вибірки

Позичальники	Класифіковані за допомогою дерева рішень		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	123	127	250
Схвалено надання кредиту	85	554	639
Всього позичальників	208	681	889
Ассурасу	0,4920	0,8670	-
Параметри	(10; false; false; 0)		

Як бачимо, початкова та тестова вибірки показали різні результати по точності класифікації «поганих» позичальників. Тестова вибірка набагато гірше визначає позичальників, що не повернуть кредит. Навчальна вибірка краще класифікує «хороших» та «поганих» позичальників.

Розглянемо результати для всієї сукупності загалом (таблиця 2.13).

Таблиця 2.13

Таблиця спряженості для генеральної сукупності

Позичальники	Класифіковані за допомогою дерева рішень		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	810	439	1249
Схвалено надання кредиту	344	2853	3197
Всього позичальників	1154	3292	4446
Ассигасу	0,6485	0,8924	-
Параметри	(10; false; false; 0)		

Як видно із таблиці 2.13, класифікація здійснюється добре лише для навчальної вибірки. Якщо розглядати тестову та сукупну вибірки, то результати по точності класифікації моделі будуть не надто хорошими. Модель доцільно використовувати для зменшення помилок I роду, тут вона дає досить позитивні результати по класифікації.

Кластеризація. Вибір екзогенних та ендогенних змінних та формування навчальної/тестової вибірок проводимо за такою ж схемою, що і для логістичної регресії. Далі налаштовуємо параметри кластеризації, тобто визначаємо, на яку кількість кластерів буде розподілятися вхідна множина прикладів. Модель кластеризації будується за алгоритмом k-means. Шляхом перебору різних випадків та параметрів, було вирішено будувати модель із наступними налаштуваннями:

- Рівень значимості = 5%, при цьому розбиття відбулося на 218 кластерів.

При ручному заданні кількості кластерів результати розподілу були набагато гіршими. У додатку Д подано результати моделі із різними параметрами кластеризації.

Фрагмент загальної структури сформованих алгоритмом кластерів можна переглянути на рис.2.14. Тут представлено всі розглянуті властивості разом із характером впливу їх на склад кластера.

		Кластеры									
		108	105	74	21	4					
		24 (0,5%)		24 (0,5%)		24 (0,5%)		24 (0,5%)		24 (0,5%)	
- + Поля		Показатели		☒	☒	☒	☒	☒	☒	☒	
+ ab Marital	Значимость		93,8%		93,8%		100,0%		93,8%		100,0%
	Доверительный интервал										
+ 9.0 Age	Значимость		100,0%		31,7%		99,9%		100,0%		74,1%
	Доверительный интервал										
+ 9.0 Time	Значимость		100,0%		98,8%		69,1%		99,9%		67,5%
	Доверительный интервал										
	Среднее		59,5		54		49,5		56,5		43,5
	Стандартн. откл.		4,303689014		9,362273971		13,03173384		5,571667458		14,82946538
	Стандартн. ошиб.		0,8784868414		1,911066172		2,660091532		1,137311857		3,027051945
+ ab Home	Значимость		100,0%		100,0%		100,0%		100,0%		100,0%
	Доверительный интервал										
+ 9.0 Seniority	Значимость		89,4%		49,6%		99,9%		99,9%		61,4%
	Доверительный интервал										

Рис.2.14. Фрагмент характеристики властивостей кластерів

Також було побудовано матрицю порівнянь, що подано у додатку Е, де можна побачити схожість побудованих кластерів між собою. Тут у відсотковому вираженні демонструється наскільки кластери схожі один із одним та мають однакові елементи. За допомогою матриці можна визначити найсильніші та

найслабші зв'язки, визначити потрібний діапазон критерія та вірно представити кластери.

Було побудовано таблиці спряженості для кожної окремої вибірки (таблиці 2.14-2.16).

Таблиця 2.14

Таблиця спряженості для навчаючої вибірки

Позичальники	Класифіковані за допомогою алгоритму k-means		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	465	534	999
Схвалено надання кредиту	240	2318	2558
Всього позичальників	705	2852	3557
Ассигасу	0,4655	0,9062	-
Параметри	(p=5; k=218)		

Таблиця 2.15

Таблиця спряженості для тестової вибірки

Позичальники	Класифіковані за допомогою алгоритму k-means		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	102	148	250
Схвалено надання кредиту	59	580	639
Всього позичальників	161	728	889
Ассигасу	0,4080	0,9077	-
Параметри	(p=5; k=218)		

Таблиця 2.16

Таблиця спряженості для генеральної сукупності

Позичальники	Класифіковані за допомогою алгоритму k-means		
	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
У видачі кредиту відмовлено	567	682	1249

Продовження табл. 2.16

Схвалено надання кредиту	299	2898	3197
Всього позичальників	866	3580	4446
Ассигасу	0,4540	0,9065	-
Параметри	(p=5; k=218)		

Як бачимо із табл.2.14 та 2.15, початкова та тестова вибірки показали подібні результати по точності класифікації позичальників. При цьому точність діагностування позичальників, що не повернуть кредит, є значно гіршою.

Як видно із загальної таблиці 2.16, класифікація здійснюється добре лише для позичальників, що повернуть кредит. Якщо розглядати тестову та сукупну вибірки, то результати по точності класифікації моделі будуть не надто хорошими. Модель доцільно використовувати для зменшення помилок I роду, тут вона дає досить позитивні результати по класифікації.

Останньою побудуємо модель із використанням карт Кохонена. Виконуємо перші дії аналогічно як і для решти моделей. Однак тут дещо потрібно змінити вхідні дані, оскільки модель не працює із дискретними параметрами. Замінюємо їх на неперервні наступним чином:

- Для змінної «вид володіння житлом» було виділено категорію «1», якщо мова йде про орендування житла (rent), «2», якщо клієнт є власником помешкання (owner), «3» - інформацію приховано (priv), «4» - дані пропущено (ignore), «5» - клієнт живе у батьків (parnts), «6» -інше (other).
- Для змінної «сімейний стан» – «1», якщо клієнт є самотнім/самотньою (single), «2» - одружений/заміжня (married), «3» - вдівець/вдова(widow), «4» - живуть окремо(separated), «5» - розлучений/розлучена(divorced)
- Для змінної «наявність записів» введемо лише два позначення «1» для відсутності записів про клієнта та «2», якщо записи наявні.

- Для змінної «тип роботи» позначення будуть наступні: «1» - фіксований графік(fixed), «2» - неповний робочий день (part-time), «3» - фріланс(freelance), «4» - інше(others).

Налаштування параметрів карт Кохена:

- розмір осі X: 16;
- розмір осі Y: 12;
- помилка < 0,001
- епоха = 2000
- рівень значимості: 5.

Приклади результатів побудови моделі із використанням карт Кохонена наведено у Додатку Ж, де було вказано різні варіанти параметрів моделі, проте жодна із них не давала результатів кращих, аніж із обраними вище показниками..

Після побудови моделі утворилося 10 кластерів. Фрагмент для карт Кохонена розміщено у Додатках Й-Л.

На рис.2.15 представлено карти Кохонена для кожної змінної та кластерів. Спроековано багатовимірний простір на площу. Тут можна провести паралель із географічними картами: коли більш високі поверхні, що знаходяться над рівнем моря зображуються червоним кольором, а ті, що нижче рівня моря – голубим. Принцип тут схожий.

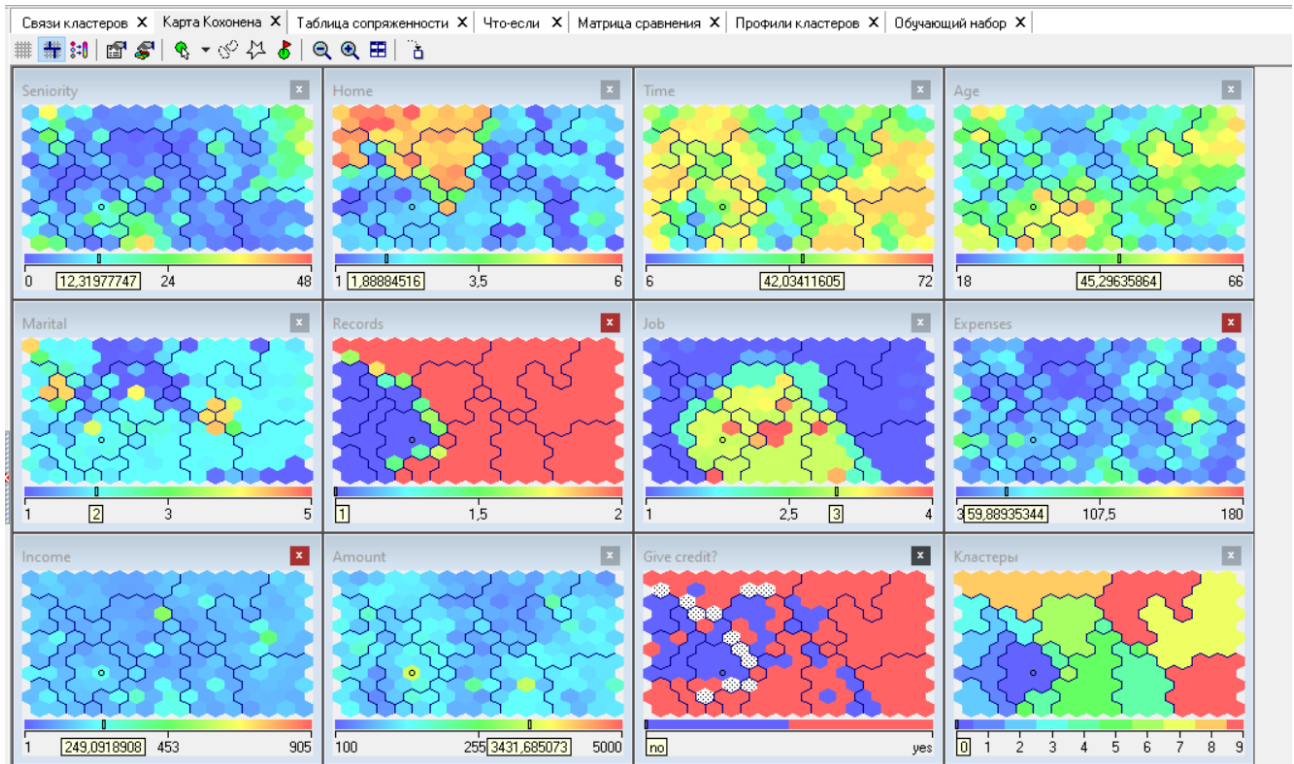


Рис.2.15. Карти Кохонена

Перше вікно – це «Стаж роботи», де розподіл значень знаходиться від мінімального (нуль) до максимального (48) значення. Для кожного інтервалу є свій колір, тобто можна судити, що для більшості спостережень у вибірці стаж роботи коливається від 0 до 24. Було проведено групування та побудовано кластери (остання карта). Кластери позначені блакитним та близьким до нього кольором містять більше «поганих» позичальників. Можна проаналізувати, хто відноситься до кластеру «1». Візьмемо довільну точку у цьому кластері і одразу на інших картках з'являються точки відношення до усіх інших карт. Для випадку, коли у кредитуванні буде відмовлено, параметри наступні: стаж роботи 2,5 роки; тип володіння житлом - оренда; сімейний стан - одружений; записи - присутні; доходи – 213 доларів; сума кредиту – 1568 доларів; період кредитування – 56 днів; вік – 39 років; тип роботи - повний робочий день; витрати – 102 долари. Аналогічно можна аналізувати решту результатів, що показують карти Кохонена.

Таблиця 2.17

Зведена таблиця спряженості вибірок для карт Кохонена

Позичальники	Класифіковані за допомогою алгоритму SOM		
Фактичні значення	У видачі кредиту відмовлено	Схвалено надання кредиту	Всього позичальників
Навчальна вибірка			
У видачі кредиту відмовлено	500	499	999
Схвалено надання кредиту	290	2268	2558
Всього позичальників	790	2767	3557
Ассигасу	0,5005	0,8866	
Параметри	(0,001; 2000; 0,001;5)		
Тестова вибірка			
У видачі кредиту відмовлено	113	137	250
Схвалено надання кредиту	64	575	639
Всього позичальників	177	712	889
Ассигасу	0,4520	0,8998	
Параметри	(0,001; 2000; 0,001;5)		
Генеральна сукупність			
У видачі кредиту відмовлено	613	636	1249
Схвалено надання кредиту	354	2843	3197
Всього позичальників	967	3479	4446
Ассигасу	0,4908	0,8893	
Параметри	(0,001; 2000; 0,001;5)		

На основі отриманих результатів було побудовано таблицю спряженості для кожної із вибірок. Загальний вигляд цієї таблиці із використанням карт Кохонена подано у табл.2.17.

Як бачимо, навчальна та тестова вибірки показали подібні результати по точності класифікації позичальників. Вибірки набагато гірше визначають точність розподілу позичальників, що не повернуть кредит.

Як видно із таблиці 2.17, класифікація здійснюється добре лише для позичальників, що повернуть кредит. Якщо розглядати тестову та сукупну вибірки, то результати по точності класифікації моделі будуть не надто

хорошими. Модель доцільно використовувати для зменшення помилок I роду, тут вона дає досить позитивні результати по класифікації.

На основі побудованих моделей можна сформуванати таблицю (табл.2.18), для порівняння якості результатів та визначити, яка ж із моделей підходить краще.

Таблиця 2.18

Порівняльна таблиця отриманих результатів

Назва моделі	Коректно класифіковано % позичальників		Середня точність моделі	Орієнтація на зменшення помилок
	У видачі кредиту відмовлено	Схвалено надання кредиту		
Логістична регресія 1	0,24	0,76	0,75	I та II роду
Дерева рішень	0,18	0,82	0,77	I роду
Кластеризація	0,22	0,78	0,68	I роду
Карти Кохонена	0,22	0,78	0,69	I роду

Як видно із табл.2.18, логістична регресія дає кращі результати для вирішення помилок II роду. Однак логістична регресія показує дещо гірші, у порівнянні із іншими методами, результати по вирішенню помилок I роду. І тут вже аналітику необхідно вирішувати, мінімізацію яких ризиків варто обрати. Вважаємо, що доцільно виділити логістичну регресію як кращу із наявних моделей.

РОЗДІЛ 3. ПОБУДОВА МОДЕЛІ ОЦІНКИ КРЕДИТОСПРОМОЖНОСТІ ПОЗИЧАЛЬНИКА-ЮРИДИЧНОЇ ОСОБИ

3.1. Аналіз показників фінансового стану підприємства та підбір коефіцієнтів для побудови моделі.

Основним інструментом при оцінці кредитоспроможності підприємства слугує система показників, що здатна відобразити конкретну сторону його економічної діяльності. Їхня сукупність показує фінансово-економічний стан майбутнього клієнта. Існує ряд вимог, котрих варто дотримуватися при оцінці кредитоспроможності позичальника, щодо процесу та результатів досліджень:

- розраховується ряд показників, а не один показник;
- показники розглядаються у взаємозв'язку, так як кожен елемент впливає на всі інші;
- варто врахувати можливі похибки, що можуть виникнути через непередбачувані фактори, при отриманні результатів.

Мікуліна М.О. запропонувала власний порядок для оцінки кредитоспроможності підприємства. Цю схему зображено на рис. 3.1. Для ідентифікації позичальника оцінюють характер його діяльності – форму власності, спеціалізацію, тривалість існування та становище на ринку. Увесь аналіз проводиться на основі формальних та неформальних показників [36].

Формальні показники обчислюються на базі фінансової звітності. Для їх обрахунку існують формули і вони приймають конкретні числові значення. Ці показники і служать для комплексної оцінки фінансового стану.

Для врахування неформальних показників використовують експертні оцінки. Оцінка ділової репутації позичальника є комплексним експертним висновком, який служить для вирішення про надання кредиту клієнту, встановлення робочих зв'язків або припинення співпраці із позичальником.

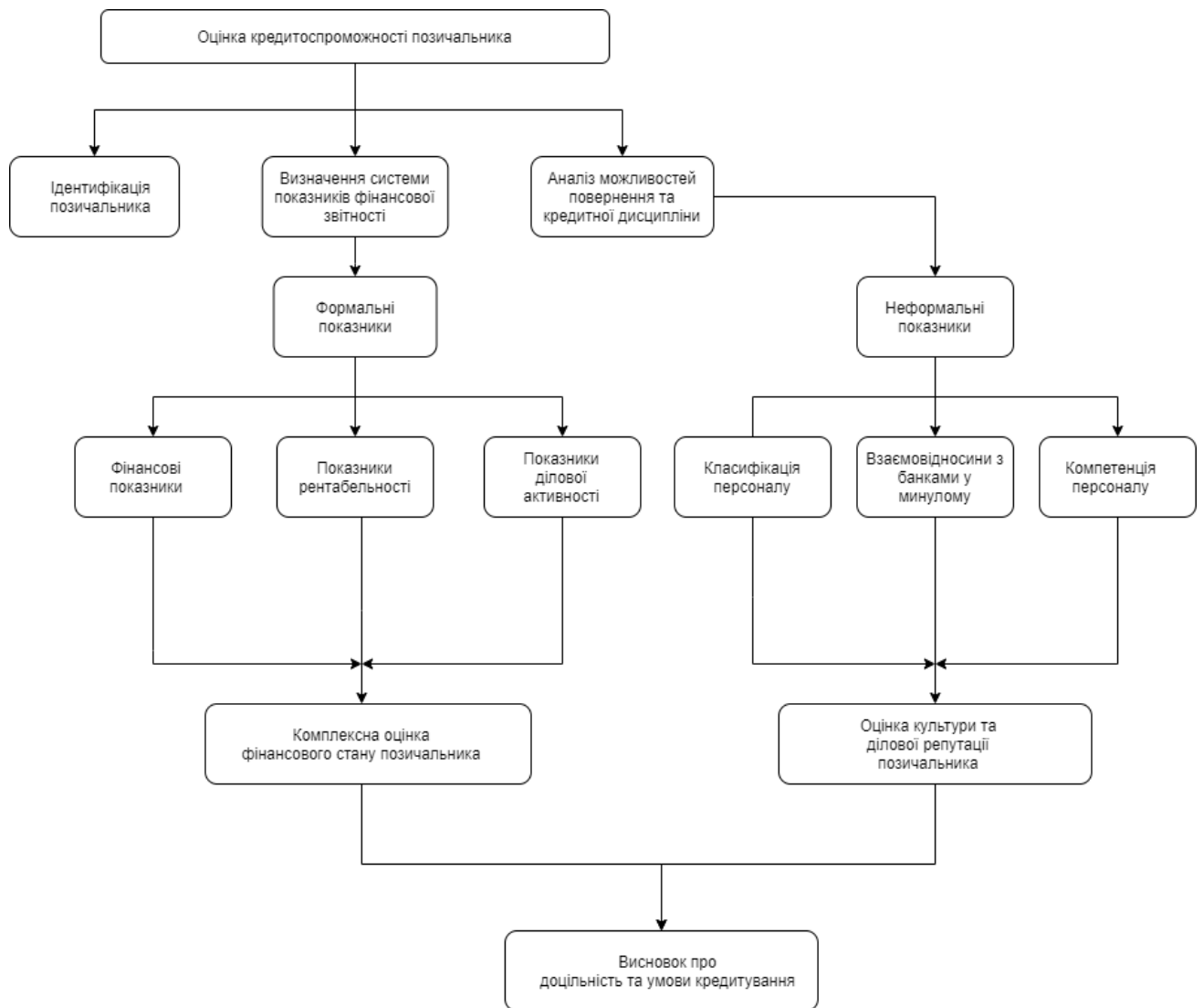


Рис. 3.1. Порядок оцінки кредитоспроможності позичальника.

Джерело: [36]

Основні вимоги до кредитоспроможності позичальника визначені «Положенням про порядок формування та використання резерву для відшкодування можливих втрат за кредитними операціями банків», затвердженим постановою Правління НБУ від 13.04.2011 р. № 114. У ньому для оцінки кредитоспроможності позичальника-юридичної особи комерційному банку пропонують використовувати для критеріїв економічної оцінки фінансової діяльності клієнта п'ять показників (коефіцієнт загальної ліквідності, коефіцієнт абсолютної ліквідності, коефіцієнт співвідношення залучених та власних коштів, коефіцієнт фінансової незалежності та коефіцієнт маневреності власних коштів) [37].

Однак для оцінки кредитоспроможності позичальника доцільно збільшити кількість фінансових коефіцієнтів.

Для побудови моделі вирішено використати 21 показник, що відображають фінансовий стан позичальника-юридичної особи.

Таблиця 3.1

Початковий набір ознак, що будуть застосовуватися при аналізі
кредитоспроможності

№	Показник	Позначення	Нормативне значення	Формула	
Показники ліквідності					
1.	Коефіцієнт покриття	$K_{п}$	x_1	1-3	Оборотні активи/Поточні зобов'язання
2.	Коефіцієнт абсолютної ліквідності	$K_{ал}$	x_2	0,1-0,2	Грошові кошти та їх еквіваленти/Поточні зобов'язання
3.	Коефіцієнт швидкої ліквідності	$K_{шл}$	x_3	$\geq 0,5$	(Чиста реалізаційна вартість + Кошти та їх еквіваленти у нац. вал.)/Поточні зобов'язання
4.	Коефіцієнт мобільності активів	$K_{ма}$	x_4	Тенденція до зростання	Оборотні активи/Необоротні активи
5.	Коефіцієнт поточної заборгованості	$K_{пз}$	x_5	Тенденція до зниження	Поточні зобов'язання/Активи
6.	Коефіцієнт покриття боргів власним капіталом	$K_{пб}$	x_6	Тенденція до зростання	Власний капітал/(Забезпечення майбутніх витрат + Довгострокові зобов'язання + Короткострокові зобов'язання)
Показники фінансової стійкості					
1.	Коефіцієнт фінансової автономії	K_a	x_7	0,4-0,6	Власний капітал/Пасиви
2.	Коефіцієнт фінансової залежності	$K_{фз}$	x_8	1,67-2,5	Пасиви/Власний капітал
3.	Коефіцієнт маневреності власного капіталу	K_m	x_9	$\geq 0,1$	(Власний капітал-Необоротні активи)/Власний капітал
4.	Коефіцієнт концентрації позикового капіталу	$K_{кпк}$	x_{10}	0,4-0,6	(Довгострокові зобов'язання + Поточні зобов'язання)/Активи

Продовження таблиці 3.1

5.	Коефіцієнт покриття інвестицій	$K_{\text{пі}}$	x_{11}	0,75-0,9	(Капітал + Чиста реалізаційна вартість)/Активи
6.	Коефіцієнт фінансового ризику	$K_{\text{р}}$	x_{12}	≤ 1	(Забезпечення майбутніх витрат + Довгострокові зобов'язання + Короткострокові зобов'язання)/Капітал
7.	Коефіцієнт забезпечення власними оборотними коштами	$K_{\text{звк}}$	x_{13}	$\geq 0,1$	(Оборотні активи-Поточні зобов'язання)/Оборотні активи
Показники оцінки майнового стану					
1.	Коефіцієнт зносу основних засобів	$K_{\text{з}}$	x_{14}	$< 0,5$	Знос/Первинна вартість
Показники рентабельності					
1.	Коефіцієнт покриття активів	$K_{\text{па}}$	x_{15}	Тенденція до зростання	Чистий дохід від реалізації продукції/Активи
2.	Період окупності активів	$K_{\text{оа}}$	x_{16}	Залежить від галузі (тенденція до зменшення)	Активи/ Чистий дохід від реалізації продукції
Показники ділової активності					
1.	Оборотність оборотних активів	$K_{\text{ооа}}$	x_{17}	Тенденція до збільшення	Чистий дохід від реалізації продукції/Оборотні активи
2.	Оборотність кредитної заборгованості	$K_{\text{окз}}$	x_{18}	Залежить від галузі (вважають позитивним тенденцію до зниження)	Чистий дохід від реалізації продукції/Поточні зобов'язання
3.	Оборотність власного капіталу	$K_{\text{овк}}$	x_{19}	Тенденція до зростання	Чистий дохід від реалізації продукції/Власний капітал
4.	Оборотність основних засобів	$K_{\text{ооз}}$	x_{20}	Залежить від галузі	Чистий дохід від реалізації продукції/Залишкова вартість
5.	Оборотність дебіторської заборгованості	$K_{\text{одз}}$	x_{21}	Залежить від галузі (позитивним вважається тенденція до зростання)	Чистий дохід від реалізації продукції/Дебіторська заборгованість

*Джерело: сформовано автором на основі [38, 39, 40, 41, 42, 43, 44]

Перераховані коефіцієнти відносяться до однієї із чотирьох груп показників, що застосовуються для визначення кредитоспроможності: показники

ліквідності, ділової активності, фінансової стійкості та рентабельності [45]. У кожній групі аналізується окремий аспект діяльності підприємства.

Оцінку кредитоспроможності позичальника-юридичної особи варто проводити у комплексі усіх цих показників. Ліквідність підприємства характеризується співвідношенням активів та платіжних зобов'язань. Розрахунок коефіцієнтів ліквідності здійснюється за допомогою активів із другого розділу балансу. Для аналізу беруться саме оборотні активи, оскільки вони і використовуються для розрахунків по боргам підприємства. Необоротні активи слугують вторинним джерелом для погашення позики. Тому банку доцільно розглядати їх, коли вони виступають у ролі забезпечення кредиту (як приклад, це можуть бути будинки, обладнання тощо).

Для розрахунку показників ліквідності беруться до уваги три види оборотних активів, що різняться ступенем ліквідності, та три види поточних пасивів, що поділені за терміном їх погашення. Підприємства, зазвичай, таку інформацію надають банку на самостійній основі, але, якщо такі дані не надійшли банку, він зобов'язаний звертатися до балансу, а саме до поточних зобов'язань позичальника.

Із перерахованих у таблиці 3.4 показників ліквідності особливе місце займає коефіцієнт поточної ліквідності (коефіцієнт покриття). Він є основою при визнанні структури балансу підприємства незадовільною. Його використовують для оцінки межі кредитування клієнта. Якщо коефіцієнт покриття ≤ 1 , то банк відмовиться від надання позики, або ж припинить (якщо такий наявний) надання поточного кредиту. Оскільки таке значення коефіцієнту поточної ліквідності свідчить про те, що підприємству немає чим оплачувати свої поточні зобов'язання [39].

Далі детальніше зупинимося на показниках фінансової стійкості та ділової активності підприємства.

Ступінь використання власних та залучених коштів при формуванні ресурсів підприємства відображають показники фінансової стійкості. Вони показують наскільки підприємство залежне від типу джерел коштів, що дає

зможу аналізувати спроможність до погашення боргу. Дамо коротку характеристику основних показників групи.

Коефіцієнт фінансової автономії показує яку частину власних активів підприємство спроможне фінансувати власними фінансовими ресурсами. Для прикладу, якщо значення коефіцієнту дорівнює 0,25%, то підприємство здатне профінансувати 25% активів за рахунок власного капіталу. Низьке значення показника є сигналом високого ризику та низької стійкості позичальника у перспективі. Невисока частка власного капіталу супроводжується значними фінансовими витратами, що може привести до ситуації перевищення суми операційних та фінансових витрат над фінансовим результатом підприємства. Вплив даного фактору у перспективі може привести компанію до банкрутства [38].

Коефіцієнт фінансової залежності є оберненим до попереднього показника. Значення показника вказує яка кількість фінансових ресурсів використовується підприємством на кожну гривню власного капіталу. При оцінюванні кількісного вираження даного коефіцієнта варто порівнювати його значення із іншими суб'єктами ринку. Однак, занадто висока залежність говорить про значний рівень фінансових ризиків, а занадто низька – неповне використання можливостей підприємством [40].

Варто також розглянути коефіцієнт концентрації позикового капіталу, що демонструє рівень використання фінансових інструментів із метою підвищення потенційної рентабельності інвестицій. Високе значення коефіцієнта говорить нам про наявність значних фінансових ризиків підприємства. Однак, залучення позикового капіталу, при вірній політиці підприємства, здатен забезпечити ріст компанії. При аналізі показника варто порівнювати підприємство із іншими учасниками ринку. Проте, занадто низьке значення свідчить про неповне використання фінансового та виробничого потенціалу компанії. Якщо ж коефіцієнт виявився більший одиниці, то це свідчить про те, що у підприємства більше боргів, аніж активів, що є сигналом до потенційного банкрутства [41].

Описані групи показників відображають стан позичальника на час складання звітності. Для оцінки ефективності роботи підприємства у динаміці та відображення реального ходу виробничої діяльності варто розглянути показники оборотності. Коротко опишемо деякі показники даної групи.

Коефіцієнт оборотності оборотних активів вимірює ефективність використання оборотних активів підприємства. Показник демонструє швидкість переходу оборотного капіталу із стадії виробничих запасів у готову продукцію, а останнього – у грошову форму. Нормативного значення для показника немає, тому варто порівнювати його із конкурентами: високе значення свідчить про інтенсивне використання оборотних активів [42].

Порівнюючи оборотність кредиторської та дебіторської заборгованості можна визначити якість політики комерційного (товарного) кредитування в компанії. Перевищення кредиторської заборгованості над дебіторською означає, що компанія використовує кошти кредиторів в якості джерела фінансування своїх дебіторів, а інша частина грошей використовується фірмою для фінансування своїх інших операцій [43].

Описані вище показники дозволяють оцінити вчасність повернення боргу, ліквідність та реальність оборотних активів, можливе зменшення прибутку позичальника при якому він досі зможе здійснювати платежі по кредиту, оцінити загальний фінансових стан та фінансову стійкість підприємства.

Кредитоспроможність позичальника оцінюють при зіставленні фактичного значення коефіцієнта та показників, що характеризують репутацію позичальника, його фінансовий стан, грошові потоки та діловий ризик, з їхніми критеріальними рівнями. За результатами аналізу підприємство відноситься до певного класу позичальника із якими банк готовий працювати або ж якому кредит видавати не рекомендується [52].

3.2. Розробка методики проведення розрахунків

Методи дискримінантного аналізу широко використовуються у різних галузях наук, таких як економіка, соціологія, психологія. Вивченням

дискримінантного аналізу займалися Аренс та Лейтер [46], Кендалл та Стюарт [47], Г. Хотеллінг [48] та ін. Основні сфери застосування дискримінантного аналізу:

- Статистичний аналіз об'єктів;
- Моделювання залежностей між об'єктами у багатовимірному просторі ознак;
- Підбір змінних, що найкраще дискримінують групи об'єктів;
- Передбачення поведінки нових об'єктів (висновок робиться на основі початкової вибірки);
- Уточнення результатів класифікації (якщо банку необхідно віднести нових позичальників до однієї із наперед визначених груп «хороших» або «поганих»).

Дискримінантний аналіз – сукупність методів багатовимірного статистичного аналізу, задача яких розуміє під собою кластеризацію багатовимірних об'єктів за заданим набором характеристик. Відмінність дискримінантного від кластерного аналізу полягає у тому, що він не створює нові кластери, а лише виводить правило, за яким нові об'єкти відносять до одного із уже існуючих кластерів.

Основне завдання дискримінантного аналізу – вироблення рекомендацій щодо віднесення нових об'єктів до одного із раніше сформованих класів.

Постановка задачі дискримінантного аналізу. Нехай існує множина M , що складається із n об'єктів спостереження, кожен із яких характеризується певним набором ознак x_{ij} – значення j -ої змінної для i -го об'єкта, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$. Множина M включає у себе q (за умови, що $q \geq 2$) навчаючих підмножин M_k розміром n_k кожна і підмножину Z об'єктів, що підлягають дискримінації, де k – номер підмножини ($k = 1, 2, \dots, q$).

Дискримінантні ознаки x_{ij} використовують для поділу об'єктів підмножин. Кількість змінних може бути необмежена, однак кращим варіантом буде їх вибір на підставі логічного аналізу вихідної інформації. Потрібно дотримуватися

нерівності, що число об'єктів спостереження буде більшим, аніж кількість дискримінантних змінних ($n > m$).

Необхідно встановити правило розподілу p об'єктів підмножини Z із відповідними ознаками по підмножинам M_k . Залежно від геометричного розташування класів у просторі дискримінантних змінних обирається вид функції (лінійна чи нелінійна). Коефіцієнти функції a_i визначаються так, зоб дискримінантні функції для кожного із класів як найбільше відрізнялися один від одного. На основі цих коефіцієнтів для кожного об'єкта із кожної підмножини розраховується значення дискримінантних функцій, середні значення підмножин та дискримінантні константи. Виходить, що кожному спостереженню, що на початку описувалося змінними x_1, x_2, \dots, x_m , буде відповідати лише одне значення дискримінантної функції. Порівнюючи значення дискримінантних констант для кожного із класів із конкретними значення дискримінантних функцій нових об'єктів, роблять висновок про їх належність до однієї із існуючих підмножин M_k [49].

Геометрична інтерпретація постановки задачі дискримінантного аналізу представлена на рис.3.2, де M_1 та M_2 – це дві навчальні вибірки, а Z – підмножина нових об'єктів, що необхідно поділити на класи.

Найчастіше використовується лінійна дискримінантна функція, що представлена у вигляді скалярного добутку вектору дискримінантних множників $A = (a_1, a_2, \dots, a_m)$ і вектору дискримінантних змінних $X^T = (x_{i1}, x_{i2}, \dots, x_{im})$:

$$F_i = AX_i^T, \quad (3.1)$$

або

$$F_i = a_1x_{i1} + a_2x_{i2} + \dots + a_mx_{im}, \quad (3.2)$$

де X^T – транспонований вектор дискримінантних функцій x_{ij} .

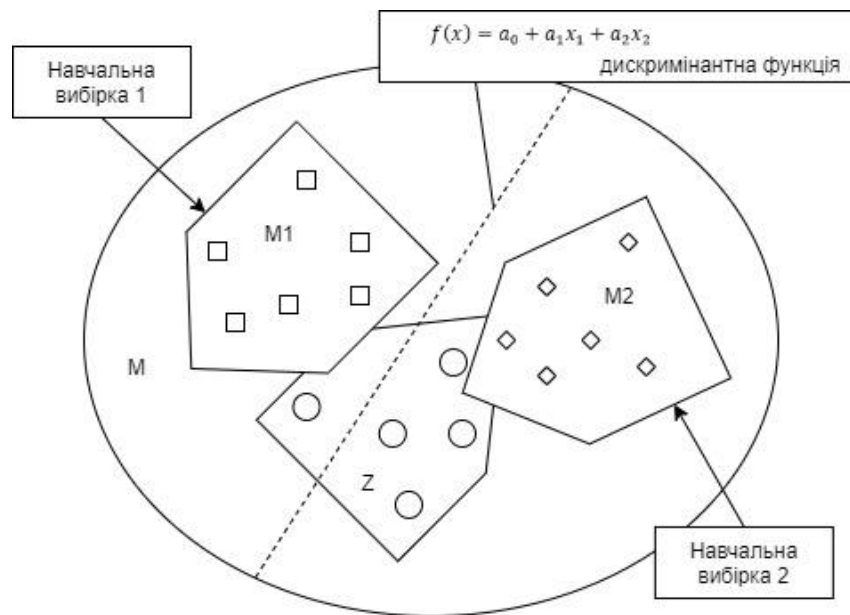


Рис.3.2. Геометрична інтерпретація задачі дискримінантного аналізу

Джерело: [49]

Основні припущення дискримінантного аналізу:

- загальна множина об'єктів M розбита на декілька навчаючих підмножин;
- у кожній підмножині знаходяться два об'єкти спостереження;
- кількість об'єктів спостереження n має перевищувати кількість дискримінантних змінних m не менше, ніж на 2 одиниці ($0 < m < n - 2$);
- змінна класифікації вимірюється за шкалою найменувань (наприклад, «хороші» та «погані» позичальники);
- припускається відсутність мультиколінеарності;
- усі підмножини гомоскедастичні [51].

Основні проблеми застосування дискримінантного аналізу пов'язані із відбором дискримінантних змінних та вибором виду дискримінантної функції.

Алгоритм дискримінантного аналізу для випадків двох класів. Існує два попередньо сформовані класи об'єктів із відомими характеристиками – дві навчальні вибірки. Необхідно:

- вивести класифікаційне правило у вигляді лінійної дискримінантної функції;

– здійснити класифікацію нових об'єктів.

1. На основі статистичних спостережень будують дві вибірки X та Y . Вибірка X – містить k об'єктів першого класу, а Y – p об'єктів другого класу. Разом вони складають сукупну вибірку із n об'єктів ($k+p=n$):

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{k1} & \cdots & x_{km} \end{pmatrix} - \text{об'єкти 1 класу};$$

$$Y = \begin{pmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{p1} & \cdots & y_{pm} \end{pmatrix} - \text{об'єкти 2 класу}.$$

2. Визначають вектори середніх значень по кожній із ознак для кожного із об'єктів класів:

$$\bar{X} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_m \end{pmatrix}, \bar{Y} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_m \end{pmatrix}.$$

3. Знаходять оцінки коваріаційних матриць S_x, S_y (розмірності $m \times m$) для кожного із двох класів:

$$S_x = \begin{pmatrix} cov^x(1,1) & cov^x(2,1) & \dots & cov^x(1,m) \\ cov^x(2,1) & cov^x(2,2) & \dots & cov^x(2,m) \\ \dots & \dots & \dots & \dots \\ cov^x(m,1) & cov^x(m,2) & \dots & cov^x(m,m) \end{pmatrix};$$

$$S_y = \begin{pmatrix} cov^y(1,1) & cov^y(2,1) & \dots & cov^y(1,m) \\ cov^y(2,1) & cov^y(2,2) & \dots & cov^y(2,m) \\ \dots & \dots & \dots & \dots \\ cov^y(m,1) & cov^y(m,2) & \dots & cov^y(m,m) \end{pmatrix}.$$

4. Визначають незміщену оцінку сумарної коваріаційної матриці S :

$$S = \frac{1}{k+p-2} (kS_x + pS_y) \quad (3.3)$$

5. Знаходять матрицю S^{-1} , обернену до сумарної коваріаційної матриці S .

6. Визначають вектор-стовпець оцінок коефіцієнтів дискримінантної функції із урахуванням усіх елементів навчаючих підмножин:

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = S^{-1}(\bar{X} - \bar{Y}) \quad (3.4)$$

Формула 3.4 отримана за МНК із умови забезпечення найбільшої різниці між дискримінантними функціями. При цьому найкращий розподіл навчаючих

підмножин забезпечується у випадку, коли внутрішньогрупова варіація є мінімальною, а міжгрупова варіація – максимальна.

7. За значенням вектору A визначають оцінки дискримінантної функції для усіх об'єктів першої та другої підмножин:

$$F_1 = XA, \quad (3.5)$$

$$F_2 = YA \quad (3.6)$$

Наприклад, для першого об'єкта першої підмножини дискримінантна функція матиме вигляд:

$$f(x_1) = a_1x_{11} + a_2x_{12} + \dots + a_mx_{1m}.$$

8. Визначають середні значення оцінок дискримінантних функцій для кожної з двох підмножин:

$$\bar{F}_1, \bar{F}_2$$

9. Визначають дискримінантну константу (межу дискримінації, яка розділяє дві навчаючі підмножини):

$$\bar{F} = \frac{1}{2}(\bar{F}_1 + \bar{F}_2) \quad (3.7)$$

10. Здійснюють розподіл нових об'єктів по навчальних вибірках. Нехай існує об'єкт Z із заданим набором ознак $Z = (z_1 + z_2, \dots, z_m)$, який необхідно віднести до 1 або 2 класу. Визначають значення дискримінантної функції для об'єкта Z :

$$F_z = f(Z) = ZA \quad (3.8)$$

або

$$f_x = f(Z) = a_1z_1 + a_2z_2 + \dots + a_mz_m$$

і порівнюють його із дискримінантною константою \bar{F} .

За результатами порівняння роблять наступні висновки:

- якщо $\bar{F}_1 > \bar{F}_2$ і значення дискримінантної функції $F_z \geq$ дискримінантної константи \bar{F} ($F_z - \bar{F} \geq 0$), то його відносять до першого класу, у іншому випадку ($F_z < \bar{F}$) – до 2 класу;
- якщо $\bar{F}_1 < \bar{F}_2$ і $F_z < \bar{F}$, то об'єкт відносять до 1 класу, інакше – до другого.

11. Оцінюють якість розподілу нових об'єктів по класах. Для цього використовують підходи:

- ступінь вкладу кожної змінної у дискримінантну функцію може бути оцінений на основі аналізу коефіцієнтів дискримінантної функції; однак їх значення залежить від масштабів одиниць вимірювання ознак, тому такий підхід не завжди є зручним;
- вивчають дискримінантні навантаження ознак як парні лінійні коефіцієнти кореляції між розрахованими значеннями дискримінантної функції та значеннями ознак, які були обрані для її побудови;
- оцінка якості класифікації може бути перевірена шляхом порівняння фактичної приналежності об'єктів навчаючих підмножин із результатами, отриманими на основі обрахунку дискримінантних функцій [49].

3.3. Побудова дискримінантної функції та оцінка якості моделі

Для побудови моделі було відібрано 125 спостережень по підприємствах із відомими результатом їхньої діяльності – чи стали підприємства банкрутами, чи залишилися на ринку.

Використовуючи етапи, що були описані у п.3.2, побудуємо дискримінантну функцію для оцінки кредитоспроможності позичальників-юридичних осіб.

Навчальну підмножину формуємо як матрицю X , куди відносимо ті підприємства, що стали банкрутами. Дана матриця містить 41-е спостереження та подана у Додатку М.

Матрицю Y формуємо як іншу навчаючу підмножину, куди відносимо ті підприємства, що є «хорошими» позичальниками. Дана матриця містить 48 спостережень та подана у Додатку Н.

У матрицю Z віднесемо нові підприємства, що підлягають класифікації, - це буде наша тестова вибірка. Дана матриця містить 36 спостережень та подана у

Додатку П. Загальна кількість підприємств навчальної вибірки становить $n=89$ (41+48).

Середні значення для кожної ознаки кожного із класу навчальної підмножини подано у табл. 3.2.

Для підмножин X та Y було розраховано коваріаційні матриці розмірністю 21×21 . Отримані коваріаційні матриці S_x та S_y подано у Додатку Р. Кожну із коваріаційних матриць, також, було перемножено на кількість спостережень у кожній із вибірок.

Таблиця 3.2

Середні значення кожної ознаки підмножин

№	Підмножини	
	X	Y
1	0,315142	0,343364
2	4,293324	10,76923
3	0,436848	0,877236
4	1,58004	1,112746
5	1,521242	1,96364
6	2,924116	8,44557
7	5,198932	1,025436
8	0,485249	0,832024
9	8,616183	1,205599
10	0,372681	0,109835
11	2,151866	2,094099
12	2,094765	5,156256
13	1,749171	1,463997
14	0,708675	0,846488
15	0,594782	0,90441
16	7,613935	0,20344
17	-2,02032	0,385552
18	0,27081	0,758825
19	87,52385	59,58076
20	0,504557	0,16459
21	1,759716	5,873547

За формулою 3.3 розраховали об'єднану коваріаційну матрицю $\hat{S} = \frac{1}{41+48-2} * (41S_x + 48S_y)$. У таблиці 3.3 подано фрагмент результатів обрахунків. Показники позначені через $x_1 - x_{21}$ описано у п.3.1. Усі результати винесено у Додаток С.

Таблиця 3.3

Фрагмент сумарної коваріаційної матриці

Показники	x14	x2	x3	x4	x17	x18	x19	x7	x8
x14	0,03	-0,10	0,01	0,03	0,03	-0,03	0,07	0,00	-0,34
x2	-0,10	15,15	0,72	-0,91	-0,90	7,99	-6,30	0,36	-9,87
x3	0,01	0,72	0,55	0,31	-0,24	0,21	-0,26	0,03	0,25
x4	0,03	-0,91	0,31	1,53	-0,55	-0,78	0,82	-0,03	2,92
x17	0,03	-0,90	-0,24	-0,55	2,19	2,35	0,94	-0,02	-1,66
x18	-0,03	7,99	0,21	-0,78	2,35	11,36	-2,57	0,20	-7,18
x19	0,07	-6,30	-0,26	0,82	0,94	-2,57	63,36	-0,86	113,65
x7	0,00	0,36	0,03	-0,03	-0,02	0,20	-0,86	0,03	-1,50
x8	-0,34	-9,87	0,25	2,92	-1,66	-7,18	113,65	-1,50	291,21
x5	0,00	-0,38	-0,02	0,05	0,01	-0,22	0,76	-0,03	1,22
x20	0,11	-1,57	0,12	0,84	0,65	0,66	1,75	-0,04	-0,28
x1	-0,01	6,38	0,54	0,33	-1,05	3,19	-2,76	0,17	-3,97
x16	-0,03	0,23	0,05	0,08	-0,62	-1,27	-0,39	0,00	1,89
x15	0,02	-0,43	-0,05	-0,06	0,55	0,72	0,47	-0,01	-0,42
x11	0,00	0,29	0,10	0,06	-0,05	0,14	-0,88	0,03	-1,50
x12	-0,34	-9,87	0,25	2,92	-1,66	-7,17	113,66	-1,50	291,22
x9	0,06	3,02	0,06	-0,33	0,03	2,01	-33,42	0,47	-76,50
x13	-0,01	0,69	0,07	0,05	-0,18	0,33	-1,27	0,05	-1,44
x21	-2,25	78,51	-35,40	-32,88	101,29	209,54	-100,18	2,85	-242,35
x10	0,00	-0,36	-0,03	0,03	0,02	-0,20	0,87	-0,03	1,53
x6	0,00	6,94	0,54	-0,22	-0,56	3,02	-3,72	0,27	-5,95

Знаходимо матрицю \hat{S}^{-1} , обернену до матриці \hat{S} . У табл. 3.4 подано фрагмент результатів обчислень. Усі результати винесено у Додаток Т.

Таблиця 3.4

Фрагмент оберненої матриці \hat{S}^{-1}

Показники	x14	x2	x3	x4	x17	x18	x19	x7
x14	72,00	-1,98	4,97	2,71	4,22	2,30	-0,22	47,34
x2	-1,98	0,84	-0,28	0,47	-0,86	-0,47	-0,05	-1,82
x3	4,97	-0,28	9,97	2,83	0,42	0,29	-0,12	89,92
x4	2,71	0,47	2,83	5,68	-2,35	0,02	-0,06	46,86
x17	4,22	-0,86	0,42	-2,35	6,74	-0,01	-0,07	-3,75
x18	2,30	-0,47	0,29	0,02	-0,01	0,68	0,05	-0,13
x19	-0,22	-0,05	-0,12	-0,06	-0,07	0,05	0,18	-2,61
x7	47,34	-1,82	89,92	46,86	-3,75	-0,13	-2,61	4504,95
x8	-196,00	21,39	-20,57	2,20	-30,83	-5,78	-2,34	2743,33
x5	-16,36	1,57	7,12	-19,58	27,23	-2,29	-2,52	39,00
x20	-3,53	-0,23	-1,32	-3,98	2,50	0,03	-0,11	-26,35
x1	-0,27	-0,70	-1,25	-1,46	1,80	-0,09	0,04	-11,81
x16	-0,67	-0,32	-1,31	-1,45	-0,38	0,35	0,05	-40,59

Продовження табл. 3.4

x15	-26,60	5,45	-2,76	11,28	-23,08	-3,68	-0,27	2,72
x11	-34,96	0,91	-59,97	-28,69	-1,66	-1,96	1,47	-792,36
x12	195,91	-21,35	20,29	-2,31	30,65	5,76	2,44	-2746,16
x9	-0,61	0,07	-1,08	-0,49	-0,72	0,01	0,64	-12,48
x13	5,48	-1,48	1,77	-13,40	18,40	-0,59	-1,52	-13,23
x21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,01
x10	38,78	-3,73	25,01	12,86	-1,18	-1,11	-1,14	3615,79
x6	0,34	-0,17	-0,23	0,01	-0,12	0,07	0,01	-3,08

За формулою 3.4 було розраховано коефіцієнти дискримінантної функції А для усіх елементів навчаючих підмножин. У результаті ми отримуємо вектор-стовпець із 21-го елемента, що поданий у таблиці 3.5 як вектор-рядок.

Таблиця 3.5

Коефіцієнти дискримінантної функції

№	1	2	3	4	5	6	7	8	9	10	11
A	-1,55	0,09	-1,54	-0,03	-1,04	-0,34	0,07	-42,37	-4,93	-14,14	-0,31
№	12	13	14	15	16	17	18	19	20	21	
A	-0,20	-0,45	2,22	13,42	5,13	0,94	-6,11	0,00	-15,22	0,10	

Дискримінантна функція матиме наступний вигляд $F = -1.55x_1 + 0.09x_2 - 1.54x_3 - 0.03x_4 - 1.04x_5 - 0.34x_6 + 0.07x_7 - 42.37x_8 - 4.93x_9 - 14.14x_{10} - 0.31x_{11} - 0.2x_{12} - 0.45x_{13} + 2.22x_{14} + 13.42x_{15} + 5.13x_{16} + 0.94x_{17} - 6.11x_{18} - 15.22x_{20} + 0.1x_{21}$.

За формулами 3.5 та 3.6 визначили значення дискримінантної функції F_x та F_y для кожного із об'єктів класу. Обраховані дані наведено у таблиці 3.6.

Таблиця 3.6

Значення дискримінантної функції для навчаючих підмножин

F _x											
№	1	2	3	4	5	6	7	8	9	10	11
F _x	-40,8	-41,3	-39,3	-39,8	-36,6	-37,3	-40,6	-39,7	-36,7	-38,2	-39,1
№	12,0	13,0	14,0	15,0	16,0	17,0	18,0	19,0	20,0	21,0	
F _x	-37,2	-36,5	-38,0	-37,2	-36,0	-35,1	-34,4	-35,1	-35,4	-35,9	
№	22,0	23,0	24,0	25,0	26,0	27,0	28,0	29,0	30,0	31,0	
F _x	-35,6	-32,9	-33,6	-33,7	-33,3	-30,4	-35,7	-32,2	-33,5	-33,5	

Продовження табл.3.6

№	32,0	33,0	34,0	35,0	36,0	37,0	38,0	39,0	40,0	41,0	
F _x	-33,4	-32,5	-32,9	-29,6	-29,6	-30,8	-31,5	-34,4	-33,7	-33,0	
F _y											
№	1	2	3	4	5	6	7	8	9	10	11
F _y	-43,0	-43,2	-46,4	-43,3	-45,5	-44,3	-43,1	-42,0	-46,1	-40,5	-40,7
№	12,0	13,0	14,0	15,0	16,0	17,0	18,0	19,0	20,0	21,0	22,0
F _y	-40,7	-40,3	-40,1	-40,8	-44,2	-43,6	-44,6	-42,1	-42,5	-44,4	-41,6
№	23,0	24,0	25,0	26,0	27,0	28,0	29,0	30,0	31,0	32,0	33,0
F _y	-41,8	-40,6	-41,7	-43,1	-39,8	-39,1	-41,5	-40,9	-41,1	-41,5	-41,6
№	34,0	35,0	36,0	37,0	38,0	39,0	40,0	41,0	42,0	43,0	44,0
F _y	-43,2	-39,6	-41,3	-37,6	-43,6	-43,2	-40,9	-40,6	-39,7	-40,0	-42,0
№	45,0	46,0	47,0	48,0							
F _y	-38,9	-38,3	-39,3	-40,5							

Знаходимо середні значення оцінок дискримінантної функції для кожної із навчальних підмножин. Результати обрахунків:

$$\bar{F}_1 = -35,26582228$$

$$\bar{F}_2 = -41,76343213$$

Скористаємося формулою 3.7 для визначення дискримінантної константи \bar{F} . У нашому випадку вона дорівнює -38,51462721.

На основі отриманих раніше результатів можна переходити до розподілу нових об'єктів. За допомогою формули 3.8, розраховуємо значення дискримінантних функцій для нашої тестової вибірки Z. Результати представлені у таблиці 3.7.

Ми маємо, що $\bar{F}_1 > \bar{F}_2$. Тепер варто визначити для яких підприємств виконується нерівність $F_Z - \bar{F} \geq 0$ та віднести їх до першого класу, у іншому випадку – до другого. Значення різниці подано у Додатку У.

Таблиця 3.7

Значення дискримінантної функції для підмножини нових об'єктів Z

№	1	2	3	4	5	6	7	8	9	10	11	12
Fz	-43,8	-39,7	-41,7	-39,7	-25,3	-38,4	-35,1	-35,4	-35,9	-32,5	-32,9	-29,6
№	13,0	14,0	15,0	16,0	17,0	18,0	19,0	20,0	21,0	22,0	23,0	24,0
Fz	-43,5	-45,2	-40,9	-41,2	-47,5	-41,8	-43,1	-39,4	-44,2	-41,5	-46,2	-45,6
№	25,0	26,0	27,0	28,0	29,0	30,0	31,0	32,0	33,0	34,0	35,0	36,0
Fz	-40,5	-40,4	-44,5	-40,7	-49,1	-41,7	-43,0	-38,9	-45,0	-39,5	-44,4	-41,4

Наступним кроком побудуємо таблицю відношень підприємств до класів. Підмножина X у нас відповідала за підприємства-банкрути, тому умовно позначимо цей клас як 1. Підмножина Y – це «хороші» позичальники, їх позначаємо як 0. Результати класифікації подано у таблиці 3.8.

Таблиця 3.8

Результати класифікації

№ підприємства	Клас (модель)	Клас (фактично)	Вірно класифіковано?
1	0	1	Ні
2	0	1	Ні
3	0	1	Ні
4	0	1	Ні
5	1	1	Так
6	1	1	Так
7	1	1	Так
8	1	1	Так
9	1	1	Так
10	1	1	Так
11	1	1	Так
12	1	1	Так
13	0	0	Так
14	0	0	Так
15	0	0	Так
16	0	0	Так
17	0	0	Так
18	0	0	Так
19	0	0	Так
20	0	0	Так
21	0	0	Так
22	0	0	Так
23	0	0	Так
24	0	0	Так

Продовження табл.3.8

25	0	0	Так
26	0	0	Так
27	0	0	Так
28	0	0	Так
29	0	0	Так
30	0	0	Так
31	0	0	Так
32	0	0	Так
33	0	0	Так
34	0	0	Так
35	0	0	Так
36	0	0	Так

Тепер перейдемо до оцінки якості класифікації об'єктів. Для цього порівнюємо значення дискримінантних функцій F_x та F_y для обох навчальних вибірок із константою дискримінації \bar{F} . Для моделі із високою якістю класифікації буде виконуватися нерівність $F_x > \bar{F}$ (3.9) для усіх значень F_x та $F_y < \bar{F}$ (3.10) для усіх значень F_y . Для нашої моделі із усіх об'єктів дискримінантної функції F_x під нерівність 3.9 підпадають 34 спостереження із 41-го. А із усіх об'єктів дискримінантної функції F_y під нерівність 3.10 підпадають 46 спостережень із 48-ми. Тому ми можемо судити про достатнього хорошу якість класифікації нових об'єктів підмножини на основі отриманої дискримінантної функції.

Оскільки ми знаємо фактичний результат для нових об'єктів, що було класифіковано у таблиці 3.8, то ми можемо судити про якість класифікації спираючись на те, яку кількість нових об'єктів було віднесено до правильного класу. Зведені дані подано у таблиці 3.9.

Таблиця 3.9

Таблиця спряженості тестової вибірки

	Вірно класифіковано	Невірно класифіковано	у %
"Хороших" позичальників	24	0	100
"Поганих" позичальників	8	4	67
Усього	32	4	0,89

Як видно із таблиці 3.9, модель вірно класифікувала усіх «хороших» позичальників. Однак, «погані» позичальники були класифіковані із помилкою у чотири спостереження. Загалом, це не поганий показник, якщо врахувати, що у загальному правильно було віднесено до потрібного класу 89% із усіх спостережень.

ВИСНОВКИ

У даній кваліфікаційній роботі була досліджена тема моделювання кредитоспроможності позичальників комерційного банку. При вивченні теми було поставлено конкретну мету – розглянути сутність кредитоспроможності та кредитування, дослідити методи та моделі оцінки кредитоспроможності фізичних на юридичних осіб, дослідити доцільність використання кредитного скорингу банками при вирішенні питань кредитування позичальників, побудувати моделі та провести аналіз отриманих результатів.

У першому розділі було подано економічну характеристику та проведено аналіз фактичного стану показників кредитної діяльності банків України за останні сім років. Було виявлено значну частку простроченої заборгованості за кредитами у структурі кредитного портфеля і, відповідно, зростання частки резервів на відшкодування даних ризиків у активах банків. Аналіз динаміки наданих кредитів показав, що кредити, які надаються суб'єктам господарювання, переважають над кредитами, що надаються фізичним особам. Дане явище зумовлене кризовим становищем економіки України. Банки віддають перевагу кредитуванню фізичних осіб, оскільки так вони зменшують ризик самого процесу кредитування та намагаються повернути надані кошти разом із процентами. Також було розглянуто поняття кредитоспроможності, важливість кредитування для банків та позичальників, процес кредитування.

Перед побудовою моделей необхідно ознайомитися із існуючими методиками, які використовуються для оцінки кредитоспроможності. Було подано перелік та аналіз видів методик оцінки кредитоспроможності позичальників-юридичних осіб. Для кожного із видів було подано порівняльний аналіз, що дає можливість обрати потрібну модель. Окремо було наведено принципи та моделі оцінки кредитоспроможності позичальників-фізичних осіб. Для них також подано порівняльний аналіз у вигляді таблиці, із виокремленими перевагами та недоліками, що дозволяє судити про доцільність використання кожного із методів.

У другому розділі було розглянуто теоретичні основи та математичні моделі оцінки кредитоспроможності фізичних осіб. Детально описано принципи побудови моделі на основі логістичної регресії, подано основні характеристики визначення якості моделі. Подано характеристику та алгоритми методів, які будуть застосовуватися у третьому питанні другого розділу при побудові моделі, а саме дерева рішень, карти Кохонена та кластеризація на основі алгоритму k-means. Виокремлено спільні та відмінні етапи при побудові кожної із моделей. Побудовано модель оцінки кредитного ризику із врахуванням бінарної змінної, що характеризує прийняття рішення про видачу кредиту, а також наведено приклад реалізації моделі на реальних даних. Отримані результати можуть використовуватися для розробки ефективної системи ризик-менеджменту у кредитних організаціях. Побудовано моделі оцінювання кредитного ризику за технологією машинного навчання із використанням алгоритмів дерев рішень, k-та G-means, SOM. Було порівняно точність результатів класифікації моделей та виявлено кращу модель.

У третьому розділі було проведено аналіз показників фінансового стану підприємства та відібрано коефіцієнтів для побудови моделі. Розглянули дискримінантну функцію, алгоритм її побудови та геометричну інтерпретацію для побудови моделі оцінки кредитоспроможності позичальників-юридичних осіб. На основі відомих значень по 125 підприємствах було побудовано дискримінантну модель, що дала дуже хороші результати – 89% усіх спостережень тестової вибірки було коректно віднесено до потрібного класу.

Інтелектуальні технології моделювання кредитоспроможності позичальників комерційних банків дозволять:

- підвищити ефективність робіт в області модельних досліджень та методів зниження кредитного ризику конкретного позичальника та диверсифікувати кредитний ризик банківської установи;
- забезпечити підтримку рішень, що мають відношення до перерозподілу фінансових активів, врахування ступеня взаємовпливу кредитного ризику та ризику ліквідності, організації комплексної

системи оперативного управління та стратегічного менеджменту у комерційному банку;

- наблизитися до вирішення проблем забезпечення необхідного ступеня якості кредитного портфеля банку, вибору оптимального відношення між рентабельністю та ліквідністю банку.

У зв'язку з вище сказаним, слід зосередитися на подальшому вдосконаленні скорингу як методу оцінки ризику. З метою підвищення ефективності використання цього методу можна запропонувати наступні рекомендації:

- ✓ Для різних видів споживчого кредиту слід застосовувати різні моделі кредитного скорингу. Розробляти моделі потрібно із врахуванням властивих різним кредитним продуктам видів кредитного ризику і факторів кредитоспроможності;
- ✓ Для одних і тих же кредитних продуктів можна побудувати різні скорингові моделі. І вже з них, на основі показників якості моделі, обрати найкращу;
- ✓ Для побудови скорингової моделі необхідно використовувати навчаючу вибірку з позитивними і негативними випадками, а не на генеральну сукупність. Це дасть змогу уникнути систематичних помилок;
- ✓ Система скорингу має постійно оновлюватися, оскільки змінюються соціально-економічні та макроекономічні умови;
- ✓ Скоринговий метод оцінки повинен не тільки відмовляти клієнту у випадку недостатньої кредитоспроможності останнього, а й пропонувати альтернативний варіант позички (з меншою сумою або на більш жорстких умовах), що буде відповідати кредитоспроможності клієнта. Це дозволить зберегти потенційного клієнта і ефективно використовувати банківські ресурси;
- ✓ Скоринговий метод має не лише оцінити клієнта в даний момент, але і спрогнозувати його поведінку у випадку виникнення яких-небудь труднощів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ишина І.В., Сазанова М.Н. Скоринг – модель оцінки кредитного ризику. *Аудит и финансовый анализ*. – 2007. - №4.
2. Офіційний сайт НБУ. URL: <https://bank.gov.ua> (дата звернення 12.10.2020)
3. Усоскін В. Н. Сучасний комерційний банк: управління і операції. М: ІПЦ "Вазар- СЕРО", 1994. 320 с.
4. Коваленко В.В. Кредитна діяльність банків України: проблеми та перспективи розвитку: наукове вид. Одеса : Атлант, 2015. 217 с.
5. Малахова О.Л. Банківська система у механізмі кредитного забезпечення підприємницької діяльності. Тернопіль, 2004. 274 с.
6. Виленский С.А. Оценка кредитных рисков: теория и практика. М.: Финансы, 2011. 315 с.
7. Вишневская О.Р. Методы и модели оценки кредитоспособности заемщика. СПб.: Невский проспект, 2015. 458 с.
8. Берегова, Г., Лабецька Л. Методи аналізу кредитного ризику та побудова моделі оцінки кредитоспроможності позичальника. *Регіональна економіка*, 2005. №4. С. 113-122.
9. Галасюк В.В. Оцінка кредитоспроможності позичальників: що оцінюємо? *Вісник НБУ*, 2001. №5. С.54-56.
- 10.Абрамченко Р.Б. Кредитные риски в Российской Федерации. М.: Инфра-М, 2014. 405с.
- 11.Про затвердження Положення про визначення банками України розміру кредитного ризику за активними банківськими операціями від 30.06.2016 № 351. URL: <https://zakon.rada.gov.ua/laws/show/v0351500-16#Text> (дата звернення 15.11.2020).
- 12.Агеева Т.А. Оценка кредитных рисков. *Вопросы экономики*, 2016. №6. С. 13-19.

- 13.Ильясов С. М. Об оценке кредитоспособности банковского заемщика. *Деньги и кредит*, 2006. № 9. С. 28–34.
- 14.Бордюг В.В. Теоретичні основи оцінки кредитоспроможності позичальника банку. *Вісник Університету банківської справи Національного банку України*, 2008. № 3. С. 112-115.
- 15.Кредитний ризик комерційного банку: навч.посіб. /Вітлінський В.В, Пернарівський О.В., Наконечний Я.С, Великоіваненко Г.І. К.: Т-во «Знання», КОО, 2000. 251 с.
- 16.Едророва В. Н., Хасянова С. Ю. Модели анализа кредитоспособности заемщиков. *Финансы и кредит*, 2002. С. 9.
- 17.Кузнєцова Н.В., Бідюк П.І. ПОРІВНЯЛЬНИЙ АНАЛІЗ ХАРАКТЕРИСТИК МОДЕЛЕЙ ОЦІНЮВАННЯ РИЗИКІВ КРЕДИТУВАННЯ. *Наукові вісті НТУУ “КПІ”*, 2010. №1. С. 42 – 53.
- 18.Лашенко В. А. Діагностика ймовірності банкрутства як експрес-аналіз економічної безпеки підприємства. *Управління розвитком*, 2013. Вип. №. 161. С. 54-58.
- 19.Мороз С. І. Порівняльний аналіз існуючих моделей прогнозування ймовірності банкрутства підприємств. *Економічні науки*. Сер.: Облік і фінанси, 2012. . 9 (2). С. 437-442.
- 20.Погребняк А. Ю. Порівняльний аналіз методик оцінки ймовірності банкрутства в системі антикризового управління на підприємстві. *Бізнес Інформ*, 2014. №. 7. С. 344-352.
- 21.Джозеф Ф. Синки, мл. Управление финансами в коммерческих банках [Текст] : пер. с англ. 4-го изд. / Д. Ф. Синки ; пер. С. М. Лощатова [и др.]. - М. : Catallaxy, 1994. 820 с.
- 22.Кривоконь О.С., Буряк І.М. Організаційні аспекти аналізу кредитоспроможності підприємств-позичальників. *Ефективна економіка*, 2012. №6. С. 15-18 URL: <http://www.economy.nayka.com.ua/index.php?operation=1&iid=1223> (дата звернення 15.09.2020).

23. Оцінка кредитоспроможності та інвестиційної привабливості суб'єктів господарювання: монографія / Єпіфанов А. О., Дехтяр Н. А., Мельник Т. М., Школьник І. О. та ін. За ред. доктора економічних наук А. О. Єпіфанова. Суми: УАБС НБУ, 2007. 286 с.
24. Xiao-Lin Li, Yu Zhong. An Overview of Personal Credit Scoring: Techniques and Future Work. *International Journal of Intelligence Science*, 2012. №2. С. 181–189.
25. Веб-інтерфейс, для перегляду статей. URL: <https://www.wikiwand.com> (дата звернення 20.11.2020)
26. Сайт кафедри обчислювальної математики факультету кібернетики Київського національного університету ім. Тараса Шевченка. URL: <http://om.univ.kiev.ua> (дата звернення 15.10.2020).
27. Інтернет видавництво про бізнес, стартапи, інновації, маркетинг та технології. скорингу. URL: <https://vc.ru> (дата звернення 21.11.2020).
28. Матвійчук А. В., Клебан Ю. В. Біннінг кількісних змінних з формуванням тренду для задач скорингу, 2017. URL: <http://ir.kneu.edu.ua/bitstream/2010/23780/3/213-229.pdf> (дата звернення 15.11.2020).
29. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям (+CD): Учеб. Пособие. 2-е изд., перераб. и доп. СПб.: Питер, 2010. 704 с.: ил.
30. Namerly G., Elkan C. Learning the k in k-means In Proc. 17th NIPS, 2003.
31. Біла Н.І. Інформаційні системи та технології в управлінні. Методичні вказівки, теоретичні відомості і завдання до лабораторних робіт для студентів та магістрів денної форми навчання спеціальності 7.803060101 Менеджмент організацій і адміністрування. Частина 2. Кластерний аналіз у бізнес-аналітиці. Запоріжжя: ЗНТУ, 2014. 38 с.
32. Kohonen, T. Self-Organizing Maps, 3rd ed. New York: Springer-Verlag, 2001.

- 33.Офіційний сайт компанії BaseGroup Labs. URL: <http://basegroup.ru/deductor/description> (дата звернення 01.10.2020).
- 34.Quinlan, J.R. C4.5 Programs for Machine Learning, San Mateo, CA: Morgan Kaufmann, 1993.
- 35.Веб-сервіс для розробки програмного забезпечення. URL: <https://github.com/gastonstat/CreditScoring> (дата звернення 20.10.2020).
- 36.Мікуліна М. О. Кредитоспроможність та методичні засади її визначення *Вісник УАБС*, 2003. № 2(15). С. 103-107.
- 37.Положення «Про порядок формування і використання резерву для відшкодування можливих збитків по кредитних операціях банків» затверджене ухвалою Правління НБУ від 13.04.2011 №114.
- 38.Кобилецький В. Р. Коефіцієнт фінансової автономії (Коефіцієнт фінансової незалежності). *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovnik-ekonomichnikh-pokaznikov/346-pokaznik-finansovoji-avtonomiji-pokaznik-finansovoji-nezalezhnosti> (дата звернення 10.10.2020).
- 39.Кобилецький В. Р., Коефіцієнт поточної ліквідності (Коефіцієнт покриття). *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovnik-ekonomichnikh-pokaznikov/256-pokaznik-potochnoji-likvidnosti-pokaznik-pokrittya> (дата звернення 11.10.2020).
- 40.Кобилецький В. Р., Коефіцієнт фінансової. *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovnik-ekonomichnikh-pokaznikov/347-pokaznik-finansovoji-zalezhnosti> (дата звернення 12.10.2020).
- 41.Кобилецький В. Р., Коефіцієнт концентрації позикового капіталу. *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovnik-ekonomichnikh-pokaznikov/264-koefitsient-kontsentratsiji-pozikovogo-kapitalu> (дата звернення 13.10.2020).
- 42.Кобилецький В. Р., Коефіцієнт оборотності оборотних активів. *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovnik-ekonomichnikh-pokaznikov/265-pokaznik-oborotnosti-oborotnih-aktiviv>

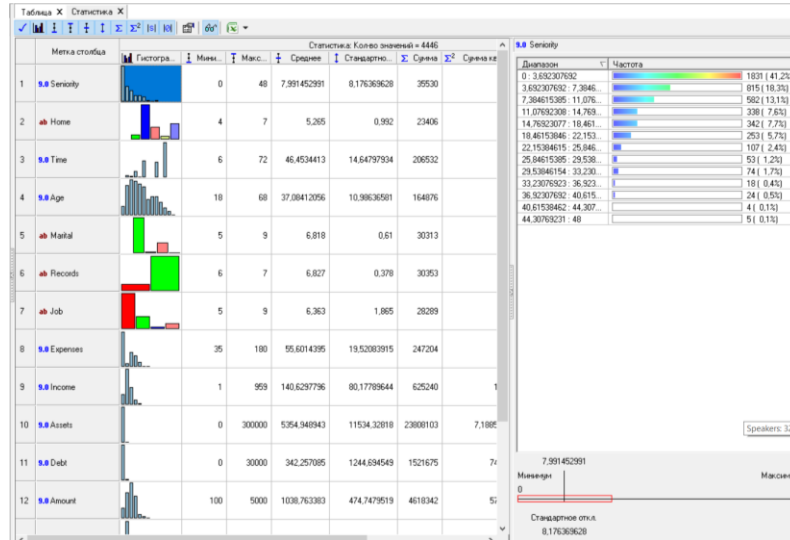
- <https://www.finalon.com/slovník-ekonomichnikh-pokaznikov/292-oborotnist-oborotnikh-aktiviv> (дата звернення 14.10.2020).
- 43.Кобилецький В. Р., Коефіцієнт оборотності кредиторської заборгованості. *Онлайн-журнал «Financial Analysis online»* – URL: <https://www.finalon.com/slovník-ekonomichnikh-pokaznikov/294-oborotnist-kreditorskoji-zaborgovanosti> (дата звернення 15.10.2020).
- 44.Сайт аудиторської фірми «Авдеев и Ко»: аудиторські та бухгалтерські послуги. URL: <https://www.audit-it.ru> (дата звернення 10.11.2020).
- 45.Фурик В.Г., Кулик І.М. Фінанси підприємств: Практикум. Вінниця: ВНТУ, 2010. 93 с.
- 46.Аренс Х. Лейтер Ю. Многомерный дисперсионный анализ. М: *Финансы и статистика*, 1985. 230 с.
- 47.Кендалл М. Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
- 48.Hotelling, Harold (September 1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology (American Psychological Association)* 24 (6): 417–441.
- 49.Олійник А.О., Субботін С. О., Олійник О. О. Інтелектуальний аналіз даних : навчальний посібник. Запоріжжя :ЗНТУ, 2012. 278 с.
- 50.Сорокин А. С. Построение скоринговых карт с использованием модели логистической регрессии. *Интернет-журнал «Наукоедение»*, 2014. Вып. 2. С. 1–29. URL: <http://naukovedenie.ru/PDF/180EVN214.pdf> (дата звернення 17.10.2020).
- 51.Сайт персональних навчальних систем Харківського національного економічного університету ім. Семена Кузнеця. URL: <https://pns.hneu.edu.ua> (дата звернення 15.10.2020).
- 52.Рекомендації з приводу оцінки комерційними банками кредитоспроможності і фінансової стабільності позичальника / Національний банк України. № 23011/79 від 02.06.94 р. URL: <http://www.bank.gov.ua> (дата звернення 15.11.2020).

53.SOM Toolbox for Matlab URL: <http://www.cis.hut.fi/projects/somtoolbox>
(дата звернення 14.10.2020).

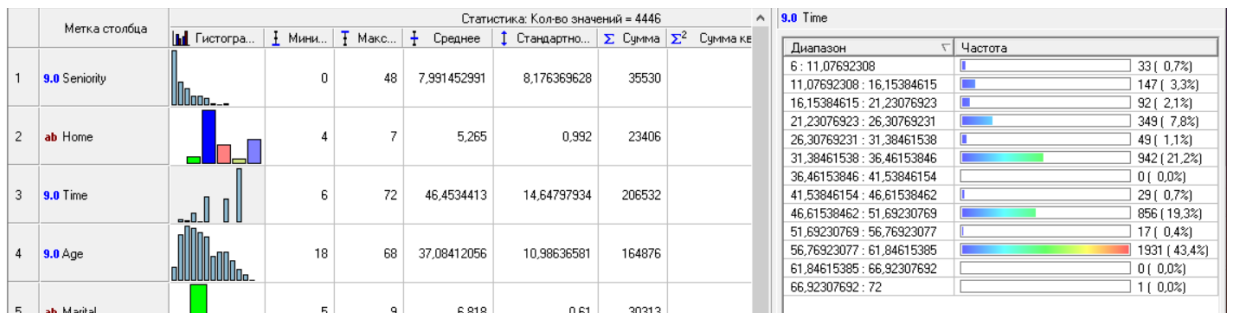
ДОДАТКИ

Додаток А

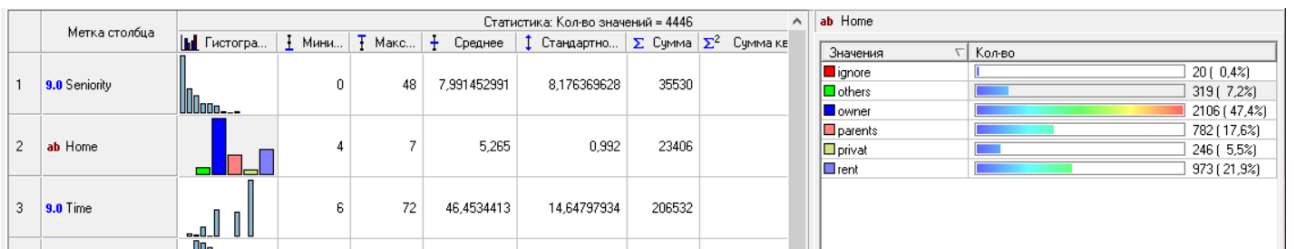
Гістограма для показників моделі



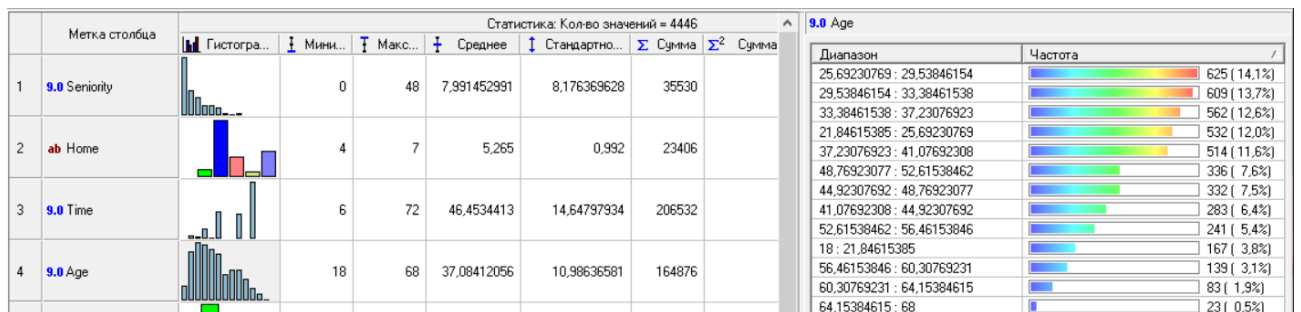
Seniority (22; 0.75)



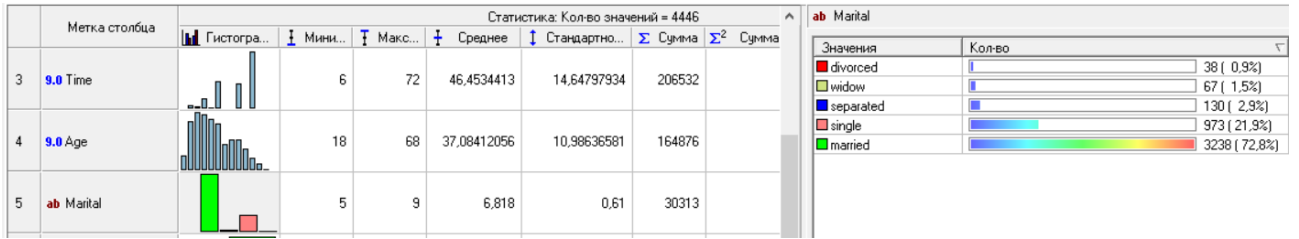
Time (22; 0.75)



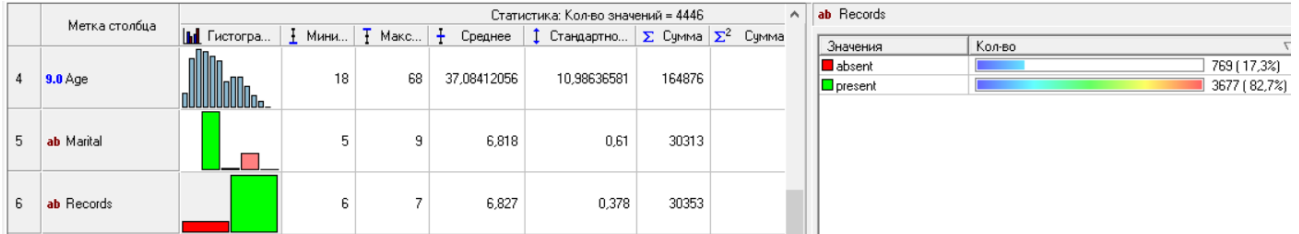
Home (22; 0.75)



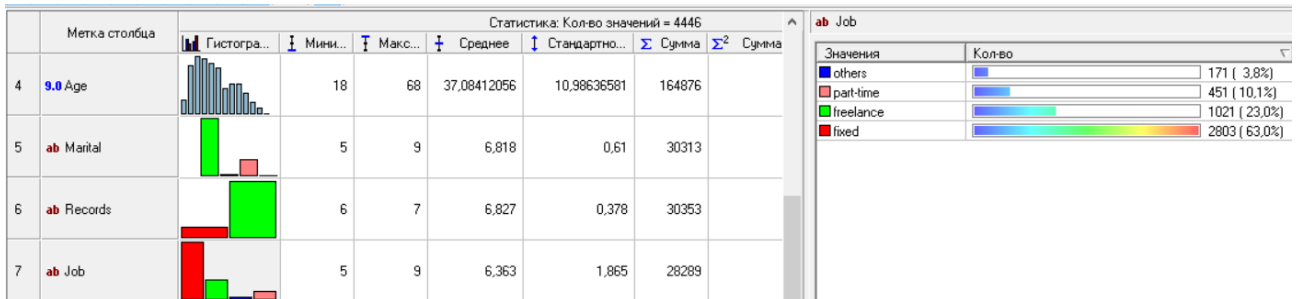
Age (22; 0.75)



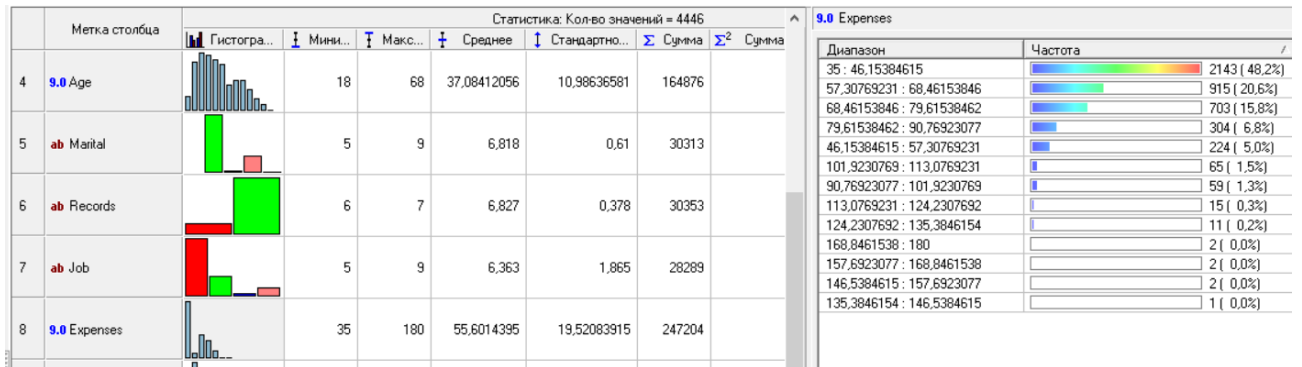
Marital (22; 0.75)



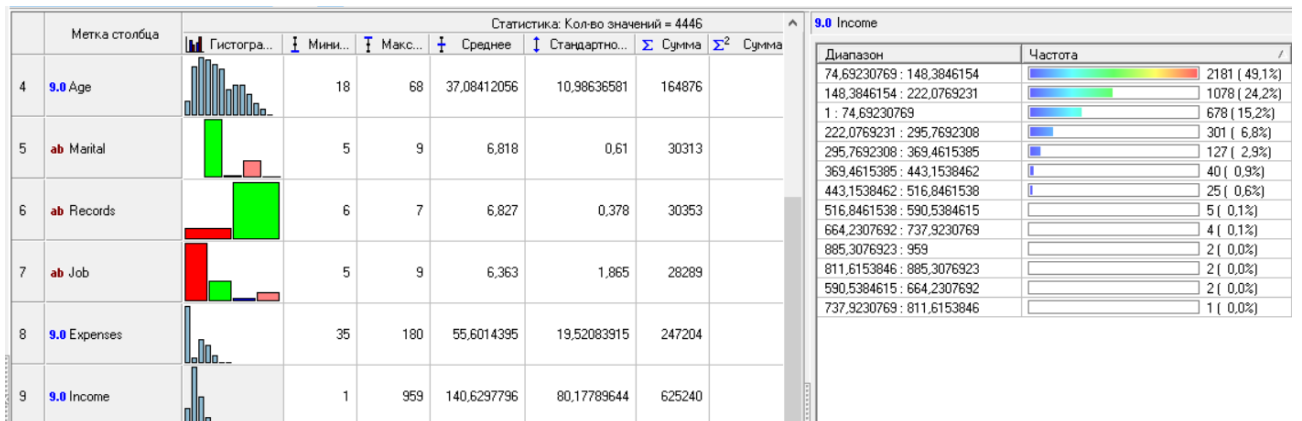
Records (22; 0.75)



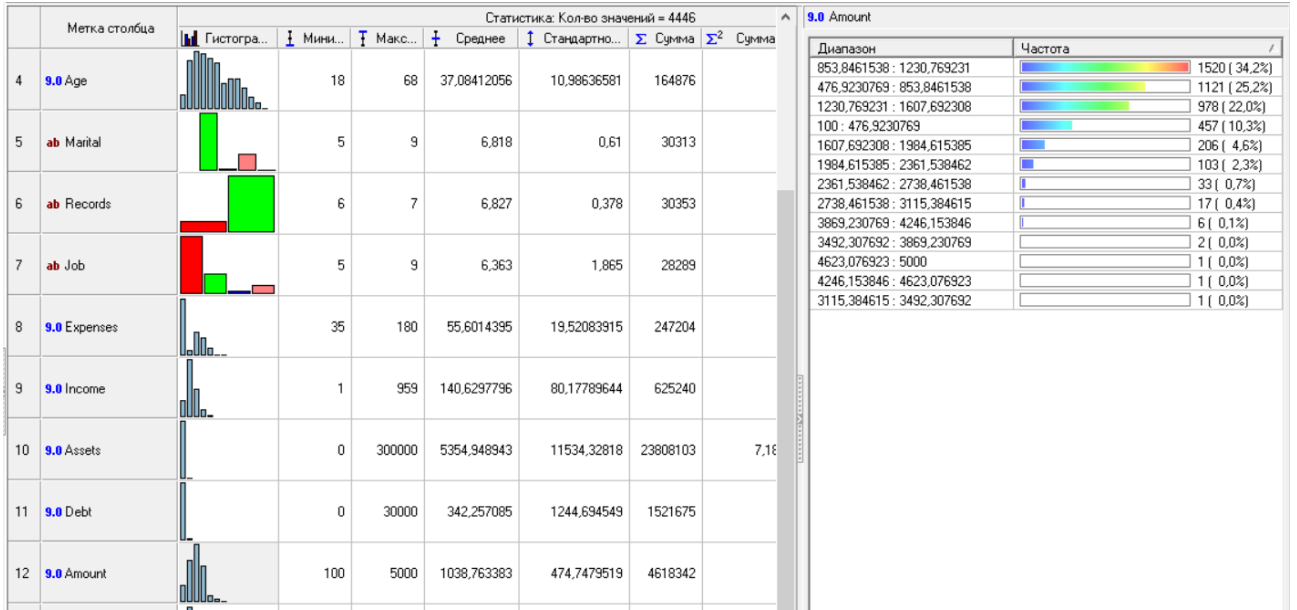
Job (22; 0.75)



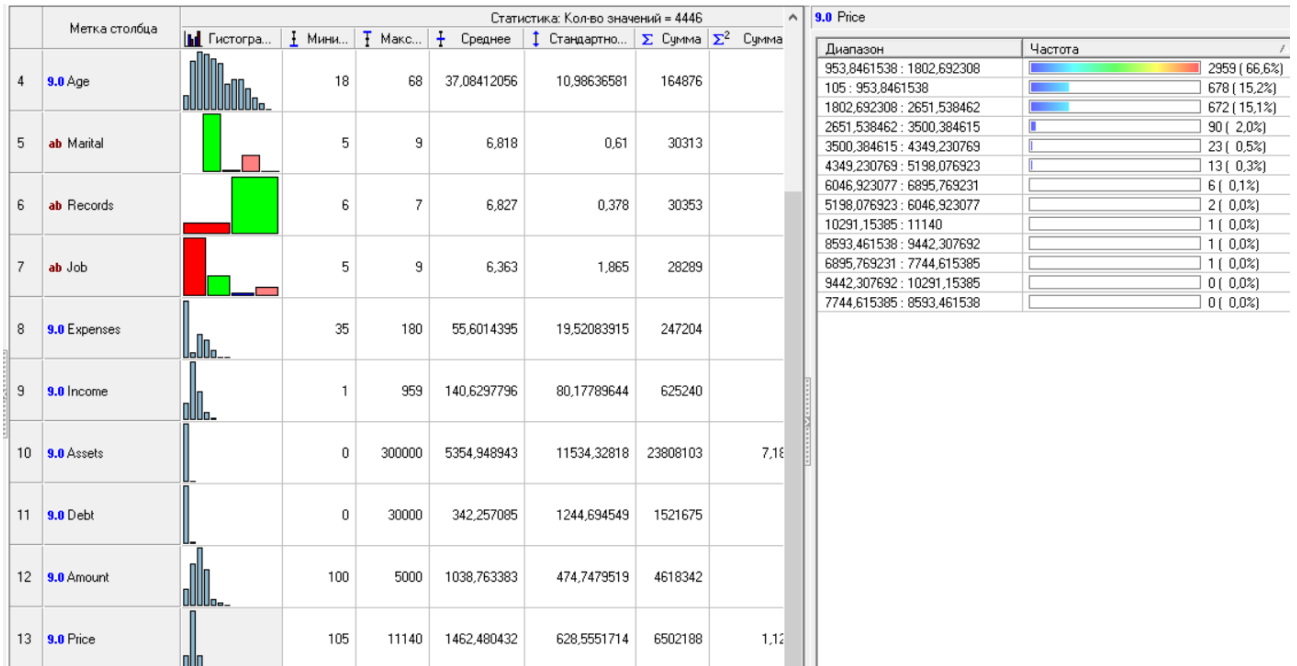
Expenses (22; 0.75)



Income (22; 0.75)



Amount (22; 0.75)



Price (22; 0.75)

Додаток Б

Зведений звіт по регресії

Фактор	Коэффициент	Стандартная ошибка	Коэффициент Вальда	Значимость	Отношение шансов	95% доверительный интервал отношения шансов	
						Минимум	Максимум
Константа	-0.41521						
Seniority	0.09022	0.0095	111.4564	0.0000	1.09441	1.07623	1.11289
Home = ignore	0.0	0.0000			1.0	1.0	1.0
Home = others	0.01679	0.6449	0.0007	0.9792	1.01694	0.28729	3.59974
Home = owner	0.93263	0.6293	2.1961	0.1384	2.54118	0.74017	8.72491
Home = parents	0.77287	0.6389	1.4634	0.2264	2.16598	0.6192	7.57667
Home = privat	0.11855	0.6477	0.0335	0.8548	1.12586	0.31634	4.00697
Home = rent	0.26037	0.6336	0.1688	0.6811	1.29741	0.37473	4.49198
Time	-0.00171	0.0039	0.1909	0.6621	0.99829	0.99067	1.00598
Age	-0.01292	0.0056	5.3662	0.0205	0.98717	0.97644	0.99801
Marital = divorced	0.0	0.0000			1.0	1.0	1.0
Marital = married	0.42195	0.4696	0.8058	0.3694	1.52432	0.60721	3.82657
Marital = separated	-0.67434	0.5186	1.6306	0.1935	0.50949	0.19436	1.40802
Marital = single	0.04337	0.4747	0.0083	0.9272	1.04432	0.41189	2.64783
Marital = widow	-0.3019	0.5974	0.2641	0.6073	0.73941	0.23382	2.23829
Records = absent	0.0	0.0000			1.0	1.0	1.0
Records = present	1.74115	0.1120	241.8404	0.0000	5.7039	4.58002	7.10356
Job = fixed	0.0	0.0000			1.0	1.0	1.0
Job = freelance	-0.72402	0.1141	40.2682	0.0000	0.4848	0.38765	0.60629
Job = others	-0.64286	0.2219	8.3895	0.0038	0.52578	0.34032	0.81233
Job = part-time	-1.46401	0.1405	108.6025	0.0000	0.23131	0.17563	0.30463
Expenses	-0.01456	0.0029	24.5742	0.0000	0.98595	0.97989	0.99124
Income	0.00705	0.0008	79.2778	0.0000	1.00707	1.00551	1.00864
Assets	0.00003	0.0000	15.5400	0.0001	1.00003	1.00002	1.00005
Debt	-0.00016	0.0000	16.4485	0.0000	0.99984	0.99976	0.99992
Amount	-0.00191	0.0002	98.2174	0.0000	0.99809	0.99771	0.99847
Price	0.00091	0.0001	40.5301	0.0000	1.00091	1.00063	1.00119

Додаток В

Деревя рішень при різних наборах параметрів навчання для тестової вибірки

Фактически	Классифицировано		
	no	yes	Итого
no	93	157	250
yes	58	581	639
Итого	151	738	889

(2; true; true; 20)

Фактически	Классифицировано		
	no	yes	Итого
no	109	141	250
yes	78	561	639
Итого	187	702	889

(5; false; true; 40)

Фактически	Классифицировано		
	no	yes	Итого
no	110	140	250
yes	62	577	639
Итого	172	717	889

(10; false; true; 40)

Give credit?			
Фактически	Классифицировано		
	no	yes	Итого
no	117	133	250
yes	65	574	639
Итого	182	707	889

(12; false; false; 0)

Додаток Г

Фрагмент правил дерева рiшень

Дерево решений X Правила X Значимость атрибутов X Таблица сопряженности X Что-если X Обучающий набор X									
Правил: 172 из 172 Фильтр: Без фильтрации									
№	Номер правила	Условие			Следствие	Поддержка		Достоверность	
		Показатель	Знак	Значение		ab Give credit?	Кол-во	%	Кол-во
1	1	ab Records	=	absent	no	19	0,53	12	63,16
		9.0 Seniority	<	0,5					
		9.0 Expenses	<	48,5					
		ab Job	=	fixed					
2	2	ab Records	=	absent	no	8	0,22	8	100,00
		9.0 Seniority	<	0,5					
		9.0 Expenses	<	48,5					
		ab Job	=	freelance					
3	3	ab Records	=	absent	no	4	0,11	4	100,00
		9.0 Seniority	<	0,5					
		9.0 Expenses	<	48,5					
		ab Job	=	others					
4	4	ab Records	=	absent	no	13	0,37	10	76,92
		9.0 Seniority	<	0,5					
		9.0 Expenses	<	48,5					
		ab Job	=	part-time					
5	5	ab Records	=	absent	no	37	1,04	36	97,30
		9.0 Seniority	<	0,5					
		9.0 Expenses	>=	48,5					
6	6	ab Records	=	absent	yes	22	0,62	19	86,36
		9.0 Seniority	>=	0,5					
		9.0 Time	<	15					
7	7	ab Records	=	absent	no	10	0,28	10	100,00
		9.0 Seniority	>=	0,5					
		9.0 Time	>=	15					
		9.0 Seniority	<	9,5					
		9.0 Income	<	42,5					

Додаток Д

Варіанти кластеризації вибірки

Фактически	Классифицировано		
	no	yes	Итого
no	80	170	250
yes	43	596	639
Итого	123	766	889

($p=0,01; k= 109$)

Фактически	Классифицировано		
	no	yes	Итого
no	79	171	250
yes	62	577	639
Итого	141	748	889

($k= 15$)

Фактически	Классифицировано		
	no	yes	Итого
no		250	250
yes		639	639
Итого	0	889	889

($k= 3$)

Додаток Е

Фрагмент матриці порівнянь

∇	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	100,00%	72,61%	72,34%	71,74%	61,70%	51,31%	53,66%	61,70%	71,54%	58,28%	71,43%	51,76%	51,65%	59,42%	60,85%	52,61%	44,67%
1	72,61%	100,00%	88,10%	85,78%	72,39%	59,54%	61,58%	72,70%	62,04%	49,16%	60,05%	61,02%	60,35%	51,95%	53,35%	61,83%	52,69%
2	72,34%	88,10%	100,00%	87,11%	71,87%	59,59%	61,91%	71,98%	59,85%	52,46%	61,96%	58,60%	59,74%	51,75%	52,11%	59,99%	50,73%
3	71,74%	85,78%	87,11%	100,00%	69,92%	55,76%	60,67%	69,47%	58,32%	48,57%	59,21%	56,82%	56,11%	48,61%	49,96%	57,78%	48,79%
4	61,70%	72,39%	71,87%	69,92%	100,00%	71,68%	72,94%	62,02%	53,08%	58,97%	53,15%	71,27%	71,83%	61,16%	46,12%	53,16%	45,77%
5	51,31%	59,54%	59,59%	55,76%	71,68%	100,00%	59,96%	52,32%	44,90%	51,24%	45,51%	59,32%	61,79%	53,21%	39,22%	45,21%	39,09%
6	53,66%	61,58%	61,91%	60,67%	72,94%	59,96%	100,00%	53,50%	45,85%	51,43%	46,27%	64,81%	60,34%	52,20%	39,60%	45,80%	51,91%
7	61,70%	72,70%	71,98%	69,47%	62,02%	52,32%	53,50%	100,00%	53,29%	44,09%	53,04%	52,60%	53,09%	45,96%	61,95%	72,75%	61,20%
8	71,54%	62,04%	59,85%	58,32%	53,08%	44,90%	45,85%	53,29%	100,00%	65,48%	86,02%	61,45%	60,73%	70,86%	72,69%	62,12%	53,19%
9	58,28%	49,16%	52,46%	48,57%	58,97%	51,24%	51,43%	44,09%	65,48%	100,00%	71,66%	64,68%	68,42%	83,76%	57,67%	49,45%	42,05%
10	71,43%	60,05%	61,96%	59,21%	53,15%	45,51%	46,27%	53,04%	86,02%	71,66%	100,00%	58,97%	60,55%	71,26%	70,56%	60,13%	51,36%
11	51,76%	61,02%	58,60%	56,82%	71,27%	59,32%	64,81%	52,60%	61,45%	64,68%	58,97%	100,00%	88,72%	70,78%	52,98%	61,45%	57,05%
12	51,65%	60,35%	59,74%	56,11%	71,83%	61,79%	60,34%	53,09%	60,73%	68,42%	60,55%	88,72%	100,00%	73,21%	53,36%	61,49%	53,28%
13	59,42%	51,95%	51,75%	48,61%	61,16%	53,21%	52,20%	45,96%	70,86%	83,76%	71,26%	70,78%	73,21%	100,00%	61,43%	52,99%	46,14%
14	60,85%	53,35%	52,11%	49,96%	46,12%	39,22%	39,60%	61,95%	72,69%	57,67%	70,56%	52,98%	53,36%	61,43%	100,00%	73,15%	61,71%
15	52,61%	61,83%	59,99%	57,78%	53,16%	45,21%	45,80%	72,75%	62,12%	49,45%	60,13%	61,45%	61,49%	52,99%	73,15%	100,00%	72,76%
16	44,67%	52,69%	50,73%	48,79%	45,77%	39,09%	51,91%	61,20%	53,19%	42,05%	51,36%	57,05%	53,28%	46,14%	61,71%	72,76%	100,00%
17	52,12%	59,96%	61,26%	57,40%	52,99%	45,93%	45,82%	72,16%	60,05%	52,58%	61,87%	59,50%	61,64%	53,41%	71,56%	89,03%	70,55%
18	61,18%	72,96%	70,44%	68,24%	61,08%	51,47%	52,85%	61,76%	73,00%	56,51%	69,92%	71,37%	70,84%	60,21%	61,86%	72,93%	61,04%
19	61,69%	72,63%	70,51%	71,22%	61,26%	50,02%	53,18%	61,09%	72,11%	56,06%	69,86%	70,47%	68,61%	58,69%	60,39%	71,25%	59,76%
20	45,56%	52,53%	52,04%	51,54%	45,47%	37,07%	52,85%	45,23%	52,63%	42,72%	51,90%	55,40%	50,49%	43,88%	44,67%	52,00%	58,94%
21	60,39%	72,14%	69,63%	66,64%	61,58%	53,03%	52,58%	61,72%	72,68%	57,10%	70,33%	72,25%	72,68%	61,68%	62,04%	73,10%	62,18%
22	60,79%	70,47%	72,69%	68,06%	61,40%	52,71%	53,23%	61,47%	70,40%	61,11%	73,16%	69,29%	71,55%	61,19%	60,78%	70,95%	59,72%
23	52,70%	60,53%	61,56%	59,09%	52,38%	43,97%	46,06%	52,81%	60,44%	51,28%	61,22%	58,87%	59,30%	51,49%	52,13%	60,34%	50,95%
24	61,83%	53,77%	52,60%	51,72%	46,44%	38,62%	40,09%	62,03%	53,62%	43,19%	52,43%	39,68%	39,42%	45,43%	61,81%	53,51%	45,87%

Карти Кохена

Фактически	Классифицировано		
	no	yes	Итого
no	101	149	250
yes	58	581	639
Итого	159	730	889

(0,05; 500; 0; 0,1)

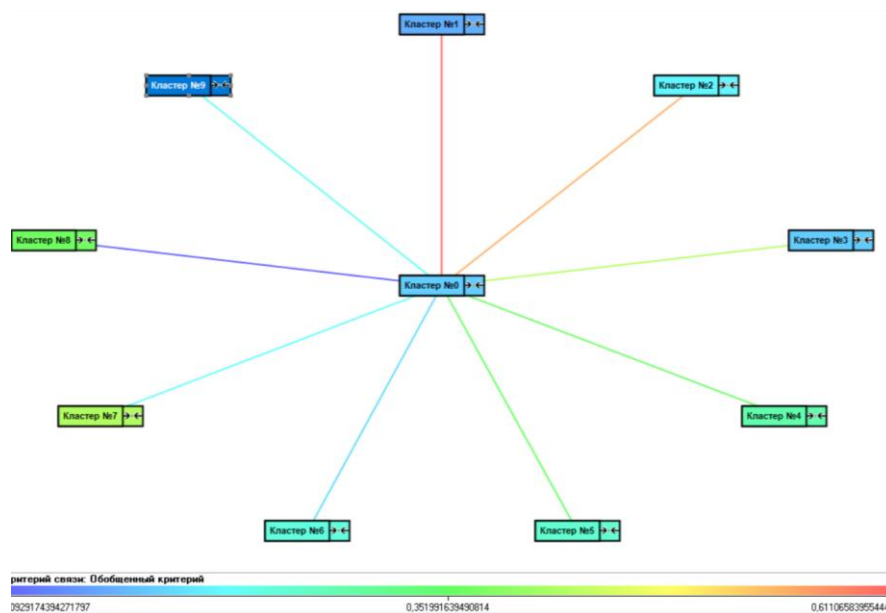
Фактически	Классифицировано		
	no	yes	Итого
no	108	142	250
yes	66	573	639
Итого	174	715	889

(0,05; 5000; 0,04; 5)

Фактически	Классифицировано		
	no	yes	Итого
no	105	145	250
yes	64	575	639
Итого	169	720	889

(0,01; 2000; 0,02; 5)

Зв'язок кластерів для карти Кохонена



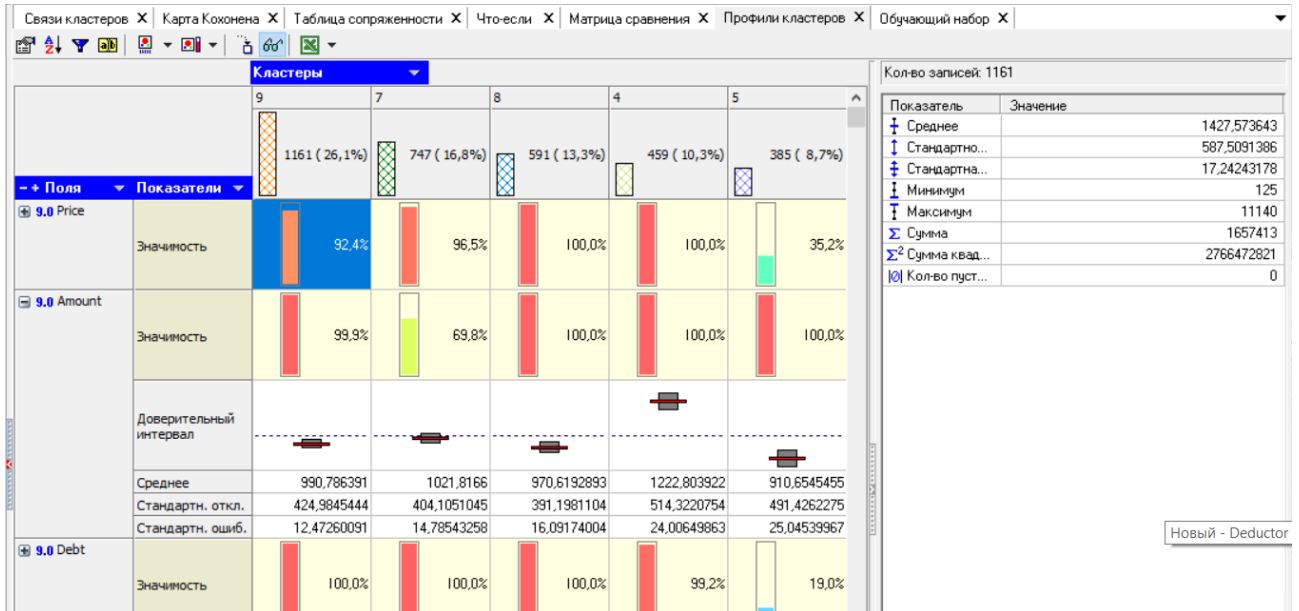
Додаток К

Матриця порівняння для карти Кохонена

κ	0	1	2	3	4	5	6	7	8	9
0	100,00%	61,11%	54,83%	42,18%	35,85%	35,62%	19,37%	22,41%	9,29%	23,21%
1	61,11%	100,00%	80,98%	50,95%	27,51%	22,63%	16,04%	33,08%	20,27%	36,18%
2	54,83%	80,98%	100,00%	52,12%	25,11%	18,34%	13,61%	35,90%	21,47%	35,30%
3	42,18%	50,95%	52,12%	100,00%	18,02%	12,11%	32,50%	19,19%	35,88%	21,69%
4	35,85%	27,51%	25,11%	18,02%	100,00%	73,39%	54,25%	55,26%	39,92%	60,95%
5	35,62%	22,63%	18,34%	12,11%	73,39%	100,00%	45,70%	50,38%	30,33%	47,64%
6	19,37%	16,04%	13,61%	32,50%	54,25%	45,70%	100,00%	36,97%	65,50%	45,05%
7	22,41%	33,08%	35,90%	19,19%	55,26%	50,38%	36,97%	100,00%	49,17%	74,91%
8	9,29%	20,27%	21,47%	35,88%	39,92%	30,33%	65,50%	49,17%	100,00%	54,77%
9	23,21%	36,18%	35,30%	21,69%	60,95%	47,64%	45,05%	74,91%	54,77%	100,00%

Додаток Л

Профілі кластерів для карти Кохонена



Додаток М

Матриця Х для підприємств-банкрутів

ID	x14	x2	x3	x4	x17	x18	x19	x7	x8	x5	x20	x1	x16	x15	x11	x12	x9	x13	x21	x10	x6	D
1	0,214628	9,165623	0,665908	0,431721	2,227554	6,156517	0,755061	0,889593	1,12411	0,109103	1,200688	2,763802	1,488768	0,671696	0,955291	0,12411	0,214854	0,63818	10,22401	0,109103	8,057368	1
2	0,480266	13,11654	0,161057	1,577644	0,866733	6,915671	0,598116	0,881515	1,13441	0,07624	1,800264	7,979008	1,89664	0,527248	0,886261	0,13441	0,562588	0,874671	111,093	0,087829	7,439916	1
3	0,363036	5,890563	0,414199	0,456499	1,409116	2,601558	0,534757	0,825886	1,21082	0,169763	2,733127	1,846234	2,264244	0,441649	0,895203	0,21082	0,168678	0,458357	6,371501	0,174114	4,743373	1
4	0,303024	6,685531	0,296014	1,055064	0,848376	2,911909	0,534339	0,815127	1,226803	0,149577	1,199595	3,432333	2,295927	0,435554	0,843688	0,226803	0,403034	0,708653	15,25	0,184873	4,409118	1
5	0,416655	5,630495	3,812347	4,274287	0,633736	2,891712	0,665467	0,771176	1,29574	0,177604	2,951168	4,562958	1,947114	1,44866	0,230129	0,754329	0,780844	0,758725	0,177604	4,345391	1	
6	0,318811	16,062	0,021258	1,016681	1,582039	12,81045	0,999724	0,797783	1,253473	0,062259	2,055571	8,097431	1,25382	0,797563	0,798445	0,253473	0,378448	0,876504	1205,25	0,19389	3,945187	1
7	0,403007	12,00244	0,584756	0,698068	1,416214	6,987805	0,757836	0,768238	1,30168	0,083316	1,127398	4,934146	1,71627	0,582199	0,807966	0,30168	0,233435	0,797331	14,65473	0,231762	3,314774	1
8	0,361302	6,116738	0,345682	1,279868	1,098535	4,257996	0,908295	0,766405	1,304793	1,632486	1,95234	3,876066	1,43653	0,696122	0,788279	0,304793	0,322031	0,742006	31,8247	0,224139	3,280918	1
9	0,423799	4,12035	1,876268	2,331671	0,749103	1,260135	0,692273	0,757302	1,320477	0,242698	2,785788	2,883628	1,90745	0,52426	1,212572	0,320477	0,603659	0,653215	1,151538	0,242698	3,12035	1
10	0,281304	3,880333	0,17476	0,750433	1,644254	2,735298	0,949646	0,74229	1,347182	0,25771	1,894826	1,663549	1,418614	0,704913	0,782672	0,347182	0,230372	0,398876	17,45626	0,25771	2,880333	1
11	0,624697	3,716486	0,1817	1,084029	2,689534	5,199321	2,035152	0,687412	1,454731	0,269071	3,671152	1,933168	0,714802	1,398988	0,710347	0,439403	0,301962	0,482714	60,99847	0,293306	2,275814	1
12	0,238961	6,733095	0,078999	0,353483	2,276206	4,002601	0,858997	0,692048	1,444987	0,14852	0,933788	1,758453	1,68218	0,594467	0,697842	0,444987	-0,06761	0,431318	102,6	0,304959	2,247256	1
13	0,176966	3,587398	0,056668	0,592204	1,482436	1,97801	0,799635	0,689536	1,450251	0,278754	1,92787	1,334297	1,81364	0,551377	0,705273	0,450251	0,089155	0,250541	35,0362	0,278754	2,220985	1
14	0,262635	8,155575	0,40884	0,541657	1,839971	5,272329	0,9957	0,649261	1,540213	0,122616	1,00984	2,865441	1,546864	0,646469	0,692459	0,540213	0,000937	0,651014	14,9652	0,350739	1,851122	1
15	0,319105	4,092366	0,160826	1,274498	1,350788	3,097528	1,193427	0,634227	1,576722	0,244357	2,166033	2,293127	1,321172	0,756904	0,662507	0,576722	0,306782	0,563914	26,76526	0,343933	1,733938	1
16	0,184273	3,845467	0,379759	0,730632	1,358747	2,205879	0,914635	0,627169	1,594467	0,260046	1,969923	1,623465	1,743281	0,573631	0,725731	0,594467	0,078679	0,384034	5,820026	0,372831	1,68218	1
17	0,088183	2,595201	0,00516	1,168355	1,167897	1,633127	1,023775	0,614673	1,62688	0,385327	1,378785	1,398349	1,5891	0,629287	0,616165	0,62688	0,249717	0,284871	422	0,385327	1,595201	1
18	0,390623	2,799306	1,62681	9,112827	0,577357	1,456382	0,856022	0,607771	1,645356	0,357231	5,411408	2,522499	1,922096	0,520265	1,187252	0,645356	0,8373	0,603568	0,897813	0,3919	1,549532	1
19	0,259331	2,711686	0,455413	1,239765	0,842631	1,264776	0,84083	0,55471	1,802744	0,368774	1,419552	1,500985	2,144005	0,466417	0,722502	0,802744	0,195119	0,333771	2,77974	0,44529	1,245728	1
20	0,273854	5,450651	0,803975	0,737492	1,752481	4,054489	1,398794	0,531782	1,880468	0,183464	1,310769	2,313571	1,34335	0,743854	0,678592	0,880468	-0,08229	0,567768	5,066809	0,468218	1,135759	1
21	0,156339	2,983735	0,254752	0,392066	1,524811	3,954137	0,628308	0,508953	1,964817	0,33515	1,222882	2,104113	1,327156	0,319779	0,594179	0,964817	0,420761	0,52474	3,752324	0,491047	1,036466	1
22	0,240778	3,262743	0,217734	0,738725	0,478424	4,821889	2,9106	0,607752	1,969465	0,306491	2,569598	1,386228	0,676652	1,477864	0,57425	0,969465	-0,13271	0,278618	22,22429	0,492248	1,031497	1
23	0,299049	3,223612	0,2957	0,753722	1,579797	2,188741	1,436621	0,472617	2,115878	0,310211	1,202491	1,385457	1,472816	0,678972	0,515256	1,115878	-0,20651	0,278217	15,92378	0,410914	0,896155	1
24	0,772805	2,157012	0,52117	0,954959	2,837826	2,989984	3,119407	0,44437	2,250379	0,463604	2,717646	1,053618	0,721413	1,386169	0,684549	1,250379	-0,15107	0,050889	5,771398	0,55563	0,799757	1
25	0,13088	3,114642	0,045646	0,95218	0,853498	1,296579	0,972335	0,428129	2,335745	0,321064	1,018274	1,519134	2,402201	0,416285	0,430394	1,326625	-0,19645	0,34173	183,8161	0,567966	0,753793	1
26	0,276426	5,773499	0,153596	0,305088	1,602266	2,162514	0,910749	0,411264	2,431527	0,173205	0,913855	1,34966	2,669809	0,374559	0,437451	1,431527	-0,86311	0,259073	14,30337	0,588736	0,698555	1
27	0,257718	1,938446	0,081206	1,043955	1,171497	1,159589	1,701525	0,35157	2,844382	0,515877	1,668075	0,989835	1,671667	0,598205	0,352269	1,844382	-0,39128	-0,01027	855,5833	0,64843	0,542187	1
28	0,483014	3,967043	0,405367	0,791587	4,520367	7,88936	5,732809	0,346902	2,882655	0,252077	3,586687	1,745292	0,502835	1,988725	0,388678	1,882655	-0,60212	0,42703	47,60511	0,653098	0,531165	1
29	0,147665	4,841747	0,799546	0,612531	1,945737	3,57561	2,404195	0,30717	3,25553	0,206537	1,244433	1,837663	1,354104	0,738496	0,470525	2,25553	-1,01724	0,455831	4,520797	0,69283	0,443355	1
30	0,247801	1,456858	0,076347	1,564802	0,385376	0,272109	0,612619	0,304885	3,279926	0,686408	0,616381	0,706088	5,353945	0,186778	0,356436	2,279926	-0,01589	-0,41625	3,623145	0,695115	0,438611	1
31	0,685695	1,434685	0,286383	1,797592	1,609193	1,483445	3,412687	0,302983	3,300514	0,697017	3,657218	0,921857	0,967131	1,033986	0,498219	2,300514	-0,17977	-0,08477	5,296085	0,697017	0,434685	1
32	0,290926	1,308326	0,018756	1,642006	1,231506	1,002795	3,640587	0,210535	4,7498	0,764336	2,033408	0,814284	1,304679	0,766472	0,223971	3,7498	-0,79359	-0,22807	57,04603	0,789465	0,266681	1
33	0,179823	1,780236	0,124541	1,530757	0,636851	0,685759	1,837269	0,209663	4,769566	0,561723	1,444069	1,076796	2,596008	0,385207	0,278962	3,767425	-0,88464	0,071319	5,558601	0,789889	0,265433	1
34	0,23969	1,628627	0,192617	3,240472	0,545327	0,678692	0,201503	0,49627	0,614014	1,24456	2,399656	1,4456	2,596008	0,416726	0,319693	3,9627	-0,17032	0,196503	3,525908	0,798497	0,252353	1
35	0,199066	2,274921	0,011095	1,32174	1,439315	1,864037	4,342154	0,188705	5,299282	0,439576	1,929275	1,295086	1,220427	0,819385	0,193427	4,299282	-1,28246	0,227851	173,5082	0,690718	0,232597	1
36	0,020966	2,371062	0,344425	0,639317	1,473634	1,362655	3,103184	0,185198	5,399637	0,421752	0,942119	0,92469	1,740031	0,574702	0,330385	4,399637	-2,29383	-0,08144	3,958355	0,814802	0,227291	1
37	0,349181	1,204386	0,387884	6,660714	0,760598	0,796475	7,184154	0,092052	10,86348	0,830298	6,931301	1,04717	1,512145	0,661312	0,413614	9,863483	-0,41808	0,045046	2,056556	0,907948	0,101384	1
38	0,81062	1,070512	0,025122	0,428166	4,944785	1,586985	2,250647	0,065888	15,18191	0,934132	6,826115	0,320941	0,674557	1,482453	0,088417	14,18191	-9,63035	-2,11584	65,74283	0,934132	0,070512	1
39	0,218545	1,165075	0,237565	1,461401	0,910288	0,62966	2,273901	0,023767	42,07457	0,858314	1,351777	0,691715	1,850325	0,540446	0,225394	41,07457	-16,0932	-0,44568	2,680427	0,976233	0,024346	1
40	0,461922	1,295594	0,081658	1,741453	1,756251	1,445395	54,87697	0,02033	49,1895	0,771847	3,062809	0,823	0,89636	1,115624	0,082492	48,1895	-16,9429	-0,21507	17,94699	0,97967	0,020751	1
41	0,067464	1,415669	0,839471	3,045311	0,421208	0,448868	48,70398	0,00651	153,6059	0,70638	1,283543	1,065668	3,153868	0,317071	0,109783	152,6059	-36,9696	0,061621	3,070215	0,99349	0,006553	1

Додаток Н

Матриця У підприємств, що є «хорошими» позичальниками

ID	x14	x2	x3	x4	x7	x18	x19	x7	x8	x5	x20	x1	x16	x15	x11	x12	x9	x13	x21	x10	x6	D
1	0.366742	18.93316	0.869751	0.718385	1.197358	9.477292	0.529491	0.945372	1.057784	0.052817	1.032102	7.915167	1.99774	0.500566	0.965515	0.055869	0.384431	0.87366	25.30892	0.052817	17.89889	0
2	0.353374	16.02694	0.533299	0.74027	1.688591	11.51195	0.768609	0.934528	1.070059	0.062395	1.508192	6.817489	1.392201	0.718287	0.960793	0.070059	0.385119	0.853318	27.34783	0.062395	14.27374	0
3	0.549756	15.72483	2.929788	2.43084	0.76382	8.510067	0.5911	0.915559	1.092239	0.063594	3.028477	11.41446	1.847792	0.541187	0.982285	0.071682	0.681644	0.910245	7.053487	0.066529	13.95048	0
4	0.171939	14.11706	2.247185	0.71947	1.45482	8.541003	0.652434	0.927321	0.972321	0.070836	2.714991	1.47086	1.65285	0.605010	1.086465	0.078335	0.370089	0.829667	3.801698	0.070836	12.75913	0
5	0.250751	17.41025	1.328949	0.833948	1.992724	15.77627	0.984796	0.920331	1.086794	0.057437	2.675864	7.916936	1.103572	0.906149	0.992017	0.086794	0.407400	0.873689	12.60647	0.079862	11.52157	0
6	0.340363	14.89045	0.985558	1.194	1.602495	12.98591	0.94842	0.919525	1.087518	0.067157	1.913338	8.103558	1.146662	0.872096	0.96447	0.087518	0.50432	0.876597	19.40368	0.079694	11.42622	0
7	0.221919	17.87402	0.078176	0.4576	1.784195	6.176293	0.341001	0.911945	1.096536	0.077207	1.340208	4.066432	1.315699	0.310974	0.981324	0.09656	0.293475	0.858444	4.482243	0.088057	10.35624	0
8	0.105263	17.87402	0.078176	0.4576	1.784195	6.176293	0.341001	0.911945	1.096536	0.077207	1.340208	4.066432	1.315699	0.310974	0.981324	0.09656	0.293475	0.858444	4.482243	0.088057	10.35624	0
9	0.463454	6.931114	0.373995	1.032147	1.319201	6.444087	0.787655	0.850671	1.175543	0.144277	1.632694	3.203379	1.49246	0.670035	0.904216	0.175543	0.421527	0.71594	12.51353	0.149329	5.696617	0
10	0.340211	6.548011	0.548011	1.767113	1.206	5.951688	0.914731	0.84196	1.187705	0.129403	2.176068	4.935065	1.29842	0.70167	0.865622	0.187705	0.570778	0.797368	32.5483	0.15804	5.32752	0
11	0.435485	9.169512	0.172764	0.695888	1.70052	4.402439	0.57024	0.841956	1.187711	0.10957	0.860421	3.762602	2.082825	0.480117	0.858093	0.187711	0.299652	0.734226	29.75275	0.15804	5.32752	0
12	0.190361	9.862036	0.067122	0.614078	1.953122	7.328154	0.88256	0.841946	1.187725	0.101399	1.260126	3.75202	1.345772	0.743068	0.84642	0.187725	0.264147	0.733477	166.0704	0.158054	5.326954	0
13	0.499062	6.609352	0.699313	0.91457	1.984886	6.261598	1.128775	0.832087	1.201798	0.099713	3.988677	7.394376	1.267375	0.948159	0.952223	0.190491	0.378194	0.683007	10.27997	0.151424	5.249593	0
14	0.341166	12.80061	0.487338	1.928424	1.397254	1.110917	1.034086	0.839256	1.191531	0.078121	2.375649	7.95072	1.152256	0.867863	0.98312	0.191531	0.548553	0.874225	10.27997	0.151424	5.249593	0
15	0.33004	15.10747	0.312551	1.077383	1.867448	14.63167	1.153531	0.832923	1.192901	0.066192	2.138353	7.835111	1.032519	0.968056	0.857895	0.192901	0.482576	0.772369	49.4072	0.161707	5.184016	0
16	0.588159	17.83116	1.603439	0.483546	2.079082	12.08338	0.808508	0.838155	1.193096	0.056082	1.821524	5.811881	1.475677	0.677655	0.917149	0.193096	0.195781	0.827939	8.578616	0.161845	5.178765	0
17	0.220502	13.16615	0.487338	1.928424	1.397254	1.110917	1.034086	0.839256	1.191531	0.078121	2.375649	7.95072	1.152256	0.867863	0.98312	0.191531	0.548553	0.874225	10.27997	0.151424	5.249593	0
18	0.220502	13.16615	0.487338	1.928424	1.397254	1.110917	1.034086	0.839256	1.191531	0.078121	2.375649	7.95072	1.152256	0.867863	0.98312	0.191531	0.548553	0.874225	10.27997	0.151424	5.249593	0
19	0.427975	10.02874	0.209266	0.979266	4.288469	1.72974	1.128775	0.832087	1.201798	0.099713	3.988677	7.394376	1.267375	0.948159	0.952223	0.190491	0.378194	0.683007	10.27997	0.151424	5.249593	0
20	0.601135	6.07459	0.734262	0.613503	5.983842	13.81181	0.98257	0.832087	1.201798	0.099713	3.988677	7.394376	1.267375	0.948159	0.952223	0.190491	0.378194	0.683007	10.27997	0.151424	5.249593	0
21	0.601135	6.07459	0.734262	0.613503	5.983842	13.81181	0.98257	0.832087	1.201798	0.099713	3.988677	7.394376	1.267375	0.948159	0.952223	0.190491	0.378194	0.683007	10.27997	0.151424	5.249593	0
22	0.649218	6.41401	0.745277	2.043937	1.73663	7.479437	1.406013	0.829373	1.20573	0.155909	4.292183	4.306867	0.857553	1.166109	0.8333	0.20573	0.603891	0.767813	296.957	0.155909	4.860745	0
23	0.221298	7.893333	3.379394	2.21978	1.971935	10.73091	1.640356	0.828302	1.206596	0.172689	4.500763	5.441818	2.53071	1.35949	1.248771	0.206596	0.625255	0.764238	3.236629	0.171222	4.840359	0
24	0.252595	12.62746	1.729838	0.508058	1.1729	4.989674	0.47054	0.828302	1.206596	0.172689	4.500763	5.441818	2.53071	1.35949	1.248771	0.206596	0.625255	0.764238	3.236629	0.171222	4.840359	0
25	0.265099	11.18354	0.575158	0.918827	1.79997	9.639241	1.046016	0.823995	1.213599	0.089417	1.77687	5.525322	1.16021	0.861913	0.873797	0.20729	0.199441	0.764935	3.055567	0.171698	4.681672	0
26	0.192082	16.35777	0.604268	1.831986	1.140013	11.78814	0.87557	0.823058	1.214981	0.061133	2.07751	10.58169	1.387646	0.720645	0.847302	0.214981	0.570979	0.905497	29.72431	0.176942	4.651564	0
27	0.260998	8.77642	0.266956	0.785452	0.907928	3.505422	0.487691	0.81899	1.221017	0.113942	0.846381	3.860903	2.50367	0.399414	0.839522	0.221017	0.31613	0.740993	19.45241	0.18101	4.524542	0
28	0.550569	5.747125	0.879001	0.927998	1.448131	8.287582	0.854224	0.816795	1.224297	0.084189	1.346469	5.722948	0.971726	0.697726	0.874712	0.224297	0.365583	0.825265	23.00806	0.183205	4.458368	0
29	0.261459	11.878	0.590051	0.927998	1.448131	8.287582	0.854224	0.816795	1.224297	0.084189	1.346469	5.722948	0.971726	0.697726	0.874712	0.224297	0.365583	0.825265	23.00806	0.183205	4.458368	0
30	0.279182	7.41217	0.874341	1.124418	2.036846	7.990773	1.32298	0.814878	1.227177	0.134914	2.335559	3.923111	0.927586	1.078068	0.918198	0.227177	0.422347	0.7451	10.43431	0.185122	4.401857	0
31	0.496342	8.624092	1.48241	0.821829	2.021268	7.863415	1.121496	0.813018	1.229985	0.115954	2.121372	3.890337	1.096736	0.911796	0.824944	0.233401	0.324863	0.742953	7.645619	0.186982	4.348112	0
32	0.298045	11.5354	0.525195	0.699756	1.690854	8.029689	0.85856	0.810766	1.233401	0.08669	2.106876	4.748895	1.437441	0.656593	0.696091	0.840712	0.274366	0.789425	23.24483	0.189234	4.28447	0
33	0.308133	15.49844	1.415443	0.821127	0.775217	5.416224	0.431184	0.810487	1.233826	0.064523	0.894321	6.986723	2.861484	0.349469	0.899469	0.233826	0.322625	0.856871	4.063438	0.189513	4.276684	0
34	0.371471	14.89646	2.261377	0.849445	1.514664	10.36318	0.862247	0.806823	1.239429	0.06713	1.442479	8.841903	1.437441	0.656593	0.696091	0.840712	0.274366	0.789425	23.24483	0.189234	4.28447	0
35	0.326745	6.343804	0.343804	0.47071	1.318883	1.97373	4.320228	0.854779	0.96715	0.157634	2.000757	3.608088	1.468396	0.681015	0.909671	0.251513	0.5482819	0.722845	6.029051	0.203285	3.91921	0
36	0.271469	9.347116	0.735309	1.751454	1.224778	7.283951	0.980718	0.940968	1.257912	0.107035	3.013279	5.94716	1.282644	0.77964	0.851631	0.257912	0.542819	0.722845	6.029051	0.203285	3.91921	0
37	0.631122	4.827872	2.668409	2.6344	0.827588	2.771171	0.724677	0.792069	1.262517	0.207131	2.0458	3.94146	1.742178	0.557994	0.825256	0.262517	0.613313	0.701358	1.228616	0.207931	3.809279	0
38	0.239197	4.76496	0.029833	0.335115	11.09571	13.28043	3.524066	0.79029	1.265358	0.20971	3.718343	1.196897	0.359062	2.785035	0.796547	0.265358	0.052248	0.164506	445.16	0.20971	3.768496	0
39	0.200066	11.10733	1.618845	0.729973	2.938418	13.72682	1.565566	0.789385	1.266809	0.090121	2.301499	4.671499	0.80917	0.628764	0.830616	0.266809	0.26952	0.785936	8.78786	0.210615	3.747995	0
40	0.505221	12.25802	0.518234	0.623565	1.667236	7.707386	0.796964	0.788969	1.267477	0.081579	1.047929	4.682283	1.590424	0.628764	0.830616	0.266809	0.26952	0.785936	8.78786	0.210615	3.747995	0
41	0.400149	10.64138	0.773506	0.787833	1.257976	5.898986	0.703326	0.788175	1.268753	0.093973	1.101057	4.689269	1.803934	0.554344	0.843981	0.268753	0.233401	0.783663	15.07743	0.202639	3.738642	0
42	0.483283	5.682823	0.158095	1.165633	1.41676	4.333484	0.970693	0.785581	1.272943	0.175969	1.670265	3.058729	1.311375	0.762558	0.802415	0.272943						

Додаток П

Матриця Z нових спостережень, що підлягають класифікації

ID	x14	x2	x3	x4	x17	x18	x19	x7	x8	x5	x20	x1	x16	x15	x11	x12	x9	x13	x21	x10	x6
1	0,371573	16,17989	0,604384	1,246859	1,463321	13,13883	0,935072	0,868433	1,151499	0,061805	1,824555	8,978775	1,231456	0,812047	0,903809	0,151499	0,487507	0,888626	22,95502	0,129653	6,600686
2	0,369836	6,559251	3,819826	4,284599	0,718579	3,82144	0,687402	0,847544	1,17988	0,152456	3,078824	5,31805	1,716434	0,582603	1,397903	0,17988	0,776732	0,811961	1,058587	0,152456	5,559251
3	0,300723	10,84523	0,664708	0,498279	2,438659	8,79569	0,957615	0,846916	1,180754	0,092206	1,304295	3,606773	1,233016	0,811019	0,907871	0,180754	0,211926	0,722744	13,30525	0,153084	5,532367
4	0,238198	7,82266	2,189001	1,058905	2,132495	8,566941	1,525399	0,761794	1,392875	0,417834	2,327741	4,017333	0,913122	1,095144	0,991072	0,392875	0,324479	0,751079	4,009574	0,28206	2,545342
5	0,02644	3,875594	0,125834	0,096254	2,685107	0,913717	0,334407	0,70501	1,418419	0,258025	0,28767	1,306028	4,241596	0,23576	0,7372	0,418419	-0,29388	-1,93868	7,32399	0,273034	2,89947
6	0,343544	4,163211	0,910903	0,766791	2,7241	4,922019	1,686456	0,701035	1,426461	0,240199	2,83724	1,806842	0,845834	1,182265	0,843526	0,426461	0,192626	0,446548	8,29716	0,298965	2,344878
7	0,259331	2,711686	0,455413	1,239765	0,842631	1,264776	0,84083	0,55471	1,802744	0,368774	1,419552	1,500985	2,144005	0,466417	0,722502	0,802744	0,195119	0,333771	2,77974	0,44529	1,245728
8	0,273854	5,450651	0,803975	0,737492	1,752481	4,054489	1,398794	0,531782	1,880468	0,183464	1,310769	2,313571	1,34435	0,743854	0,678592	0,880468	-0,08229	0,567768	5,066809	0,468218	1,135759
9	0,156339	2,983735	0,254752	2,392066	0,453463	0,954137	0,628308	0,508953	1,964817	0,33515	1,222882	2,104113	3,127156	0,319779	0,594175	0,964817	0,420761	0,52474	3,752324	0,491047	1,036466
10	0,179823	1,780236	0,124541	1,530757	0,636851	0,685759	1,837269	0,209663	4,769566	0,561723	1,044069	1,076796	2,596008	0,385207	0,278962	3,767425	-0,88464	0,071319	5,586601	0,789889	0,265433
11	0,23969	1,628627	0,192617	3,240472	0,545327	0,678692	2,068088	0,201503	4,9627	0,614014	2,177681	1,24456	2,399656	0,416726	0,319693	3,9627	-0,17032	0,196503	3,525908	0,798497	0,252353
12	0,199066	2,274921	0,011095	1,32174	1,439315	1,864037	4,342154	0,188705	5,299282	0,439576	1,929275	1,295086	1,220427	0,819385	0,193427	4,299282	-1,28246	0,227851	173,5082	0,690718	0,232597
13	0,354222	11,13496	0,627212	1,13037	2,013293	11,89491	1,173652	0,910193	1,098668	0,089807	2,608685	5,908186	0,936111	1,06825	0,947049	0,098668	0,484283	0,830743	28,98383	0,089807	10,13496
14	0,191338	12,1009	2,024664	1,432177	2,197294	15,65695	1,424375	0,908375	1,100867	0,082639	3,511692	7,125561	0,772877	1,293867	1,070132	0,100867	0,547374	0,85966	7,998855	0,091625	9,914055
15	0,464173	18,81261	0,167437	0,387642	1,249717	6,567578	0,391472	0,891776	1,121358	0,053156	0,509372	2,525252	2,864467	0,349105	0,895615	0,105442	0,191914	0,809714	90,93293	0,09401	9,483843
16	0,308755	10,44524	0,448001	1,238362	0,733009	4,235885	0,484668	0,904263	1,105873	0,095737	1,158905	5,778764	2,658984	0,405532	0,940165	0,105873	0,505946	0,826953	1,129551	0,095737	9,445242
17	0,670811	10,35278	2,463341	2,895361	2,240927	17,24407	1,843737	0,903408	1,10692	0,096592	6,984836	7,69506	6,60368	1,665646	1,077544	0,10692	0,715836	0,870047	9,565156	0,095692	9,523781
18	0,45042	1,01892	0,529189	0,735246	1,581961	7,318919	0,742813	0,902376	1,108185	0,091584	1,390429	4,626486	1,491876	0,670297	0,950446	0,108185	0,361367	0,783853	13,94439	0,097624	9,243408
19	0,356631	17,7359	0,893685	0,485487	1,563125	9,057732	0,567841	0,899372	1,111887	0,056383	1,35181	5,79463	1,958095	0,5107	0,930541	0,111887	0,251735	0,827426	16,38491	0,100628	8,937561
20	0,284254	9,660831	0,14935	0,183451	0,877893	1,31466	0,151797	0,896472	1,115484	0,103511	0,541334	1,497518	7,348541	0,136081	0,914477	0,115484	0,05746	0,332228	9,068833	0,103511	8,659193
21	0,340164	16,77912	1,667641	0,906538	1,495081	11,92818	0,794681	0,894565	1,117861	0,059598	1,459338	7,978288	1,406679	0,710894	0,988155	0,117861	0,413669	0,87466	7,595852	0,090952	8,484541
22	0,197343	12,87907	0,335548	0,384698	2,16546	7,748173	0,675334	0,890832	1,122546	0,077645	1,221559	3,578073	1,662207	0,60161	0,914203	0,122546	0,189321	0,72052	25,74172	0,109168	8,160208
23	0,257646	15,2251	4,306513	1,774965	1,232124	11,99904	0,891538	0,883989	1,131236	0,065681	2,186976	9,738506	1,268859	0,788109	1,155709	0,131236	0,592342	0,897315	2,90044	0,116011	7,619848
24	0,203215	1,765558	3,471369	1,157695	1,031645	9,772744	0,626703	0,883227	1,132211	0,056639	1,28353	9,47297	1,806615	0,553521	1,032523	0,132211	0,475268	0,894436	3,707553	0,115316	7,563656
25	0,503786	9,170342	0,189495	0,996777	1,486846	6,806442	0,883366	0,839942	1,190559	0,109047	1,924118	4,577771	1,347303	0,742223	0,857853	0,132728	0,40376	0,781553	41,43993	0,114484	7,534204
26	0,19408	9,182692	0,086976	0,477185	2,168852	6,433566	0,795332	0,880914	1,135185	0,108901	1,400571	2,966346	1,42731	0,700619	0,887673	0,135185	0,231522	0,662885	103,662	0,119086	7,397282
27	0,225746	16,43732	1,896818	1,276528	1,535258	14,15043	0,978821	0,879499	1,137011	0,060837	2,205953	9,216972	1,161612	0,860872	0,968819	0,137011	0,50055	0,891504	9,638095	0,060837	7,298685
28	0,274354	8,102967	0,437648	0,802184	1,628365	5,873143	0,826858	0,876588	1,140786	0,123442	1,413958	3,606773	1,379665	0,724814	0,927988	0,140786	0,366998	0,722744	14,10161	0,123442	7,102967
29	0,571577	10,12382	1,695148	1,450228	4,387069	26,28747	2,97468	0,872899	1,145608	0,098777	6,757818	5,992035	3,8512	2,596595	1,011301	0,145608	0,532448	0,833112	18,76124	0,127101	6,867755
30	0,27818	7,756517	3,503805	5,582703	0,960806	6,320376	0,935449	0,871078	1,148005	0,128924	5,405056	6,578199	1,272224	0,814847	1,319121	0,148005	0,825603	0,847983	1,818673	0,138924	6,756617
31	0,337863	16,2697	0,775152	0,918871	1,354181	10,5503	0,746228	0,868979	1,150763	0,061464	1,458323	7,790909	1,542107	0,648463	0,916223	0,150763	0,400291	0,871645	13,72871	0,121011	6,632926
32	0,320564	7,467397	0,121628	0,237431	0,907619	1,30041	0,201122	0,865868	1,154911	0,133915	0,688063	1,43277	5,74234	0,174145	0,881144	0,154911	0,066706	0,302051	1,140016	0,133915	6,455337
33	0,594861	13,94149	1,242818	1,72143	1,404124	12,38246	1,02769	0,864242	1,157083	0,071728	3,864067	8,818636	1,125907	0,888173	0,937103	0,157083	0,574825	0,886604	12,19004	0,135758	6,366062
34	0,390525	8,659476	0,069133	0,257127	0,878081	1,555218	0,208252	0,862404	1,159549	-0,11548	0,863492	1,771156	5,568015	0,179597	0,870129	0,159549	0,077627	0,435397	23,24816	0,137586	6,267652
35	0,258946	13,10899	0,804799	0,510685	3,391209	15,02806	1,32952	0,862262	1,159741	0,076284	1,779115	4,431476	0,872301	1,46394	0,922072	0,159741	0,232308	0,774342	19,16701	0,137738	6,260135
36	0,324509	11,09541	1,36183	0,562658	1,45775	5,822766	0,608887	0,861885	1,160247	0,090127	1,330517	3,994352	1,905522	0,524791	0,973534	0,160247	0,25765	0,749647	4,700372	0,138115	6,240347

Додаток Р

Коваріаційні матриці підмножин X та Y

Table with 21 columns and 21 rows of numerical data representing covariance matrices for subsets X and Y.

Коваріаційна матриця Sx

Table with 21 columns and 21 rows of numerical data representing the covariance matrix Sx.

Коваріаційна матриця Sy

Додаток С

Сумарна коваріаційна матриця S

Table with 21 columns and 21 rows of numerical data representing the sum of covariance matrices S.

Додаток Т

Обернена матриця \hat{S}^{-1}

72,00427227	-1,97726042	4,973372	2,70929	4,221649	2,296732	-0,21837	47,33679	-196,002	-16,3623	-3,53103	-0,27171	-0,67352	-26,5999	-34,9588	195,9092	-0,61362	5,477845	-0,00035	38,78439	0,344728
-1,977260423	0,836908898	-0,28093	0,470432	-0,85531	-0,46602	-0,04519	-1,82036	21,38966	1,567257	-0,22667	-0,70152	-0,32205	5,452769	0,912224	-21,3492	0,072073	-1,47732	0,000347	-3,73057	-0,17069
4,973372325	-0,28093207	9,969302	2,830903	0,416585	0,288376	-0,11506	89,91915	-20,5748	7,116748	-1,31974	-1,25433	-1,30754	-2,76247	-59,9678	20,29029	-1,07728	1,773365	0,000844	25,01012	-0,22764
2,709290025	0,470431763	2,830903	5,681944	-2,35077	0,015764	-0,05559	46,85734	2,199131	-19,5802	-3,97831	-1,46185	-1,4453	11,28138	-28,6894	-2,31402	-0,48678	-13,4043	0,00031	12,85638	0,00994
4,221648857	-0,85530721	0,416585	-2,35077	6,739891	-0,01431	-0,06581	-3,74622	-30,8298	27,23044	2,495516	1,803459	-0,37831	-23,0848	-1,66431	30,64876	-0,71636	18,39698	-0,00495	-1,18097	-0,11511
2,296732016	-0,4660195	0,288376	0,015764	-0,01431	0,681105	0,049129	-0,12976	-5,77514	-2,29214	0,027126	-0,08839	0,34901	-3,68121	-1,96066	5,755083	0,00861	-0,59328	-0,00059	-1,11488	0,067057
-0,218366676	-0,04518751	-0,11506	-0,05559	-0,06581	0,049129	0,179361	-2,60535	-2,34459	-2,51967	-0,10984	0,040017	0,045863	-0,27382	1,472833	2,444754	0,637405	-1,5241	7,29E-05	-1,13728	0,009128
47,33678867	-1,82036063	89,91915	46,85734	-3,74622	-0,12976	-2,60535	4504,946	2743,332	38,99977	-26,3463	-11,809	-40,5866	2,721692	-792,357	-2746,16	-12,4778	-13,2346	-0,01363	3615,791	-3,08127
-196,0024939	21,38965568	-20,5748	2,199131	-30,8298	-5,77514	-2,34459	2743,332	12824,7	95,89355	2,657576	-24,9433	-87,9629	60,66549	-202,806	-12820	14,8492	-74,4323	-0,04557	2495,022	-4,53203
-16,36226895	1,567257461	7,116748	-19,5802	27,23044	-2,29214	-2,51967	38,99977	95,89355	538,0174	12,75893	5,239496	-1,4524	-68,9274	-36,0647	-98,8883	-14,0877	185,0311	-0,02559	-157,992	-4,35558
-3,53102522	-0,22666743	-1,31974	-3,97831	2,495516	0,027126	-0,10984	-26,3463	2,657576	12,75893	5,157745	0,45149	1,71604	-13,2541	9,954133	-2,82949	-0,85165	14,76768	-0,00061	-9,58308	-0,04449
-0,271713644	-0,70152217	-1,25433	-1,46185	1,803459	-0,08839	0,040017	-11,809	-24,9433	5,239496	0,45149	1,971685	0,2512	-4,88441	10,51892	25,00341	0,294052	2,829704	-0,00104	0,668317	-0,00456
-0,67351927	-0,32205233	-1,30754	-1,4453	-0,37831	0,34901	0,045863	-40,5866	-87,9629	-1,4524	1,71604	0,2512	6,474862	4,402359	11,60314	87,6844	-1,00055	5,414817	0,001649	-26,3388	0,08283
-26,59991348	5,452768812	-2,76247	11,28138	-23,0848	-3,68121	-0,27382	2,721692	60,66549	-68,9274	-13,2541	-4,88441	4,402359	121,9448	13,88787	-60,3123	0,734509	-58,903	0,015432	-12,6062	0,031606
-34,95875562	0,912223697	-59,9678	-28,6894	-1,66431	-1,96066	1,472833	-792,357	-202,806	-36,0647	9,954133	10,51892	11,60314	13,88787	501,3929	205,4616	11,15935	-11,4357	0,01181	-249,221	1,501309
195,9091825	-21,3492264	20,29029	-2,31402	30,64876	5,755083	2,444754	-2746,16	-12820	-98,8883	-2,82949	25,00341	87,6844	-60,3123	205,4616	12815,43	-13,9192	71,63793	0,045562	-2494,77	4,539946
-0,613617635	0,072073352	-1,07728	-0,48678	-0,71636	0,00861	0,637405	-12,4778	14,8492	-14,0877	-0,85165	0,294052	-1,00055	0,734509	11,15935	-13,9192	4,489422	-12,5269	-4,5E-06	1,216641	0,057407
5,477845132	-1,47731713	1,773365	-13,4043	18,39698	-0,59328	-1,5241	-13,2346	-74,4323	185,0311	14,76768	2,829704	5,414817	-58,903	-11,4357	71,63793	-12,5269	127,2266	-0,00836	-15,1165	-0,52602
-0,00035452	0,00034697	0,000844	0,00031	-0,00495	-0,00059	7,29E-05	-0,01363	-0,04557	-0,02559	-0,00061	-0,00104	0,001649	0,015432	0,01181	0,045562	-4,5E-06	-0,00836	3,84E-05	0,00668	0,000519
38,78439194	-3,73057258	25,01012	12,85638	-1,18097	-1,11488	-1,13728	3615,791	2495,022	-157,992	-9,58308	0,668317	-26,3388	-12,6062	-249,221	-2494,77	1,216641	-15,1165	0,00668	3514,955	2,509584
0,344727661	-0,17069328	-0,22764	0,00994	-0,11511	0,067057	0,009128	-3,08127	-4,53203	-4,35558	-0,04449	-0,00456	0,08283	0,031606	1,501309	4,539946	0,057407	-0,52602	0,000519	2,509584	0,328977

Додаток У

$F_Z - \bar{F}$

№	1	2	3	4	5	6	7	8	9	10	11	12
Fz-F	-5,26415492	-1,156552533	-3,20127762	-1,16576	13,22368	0,103039	3,419252	3,082544	2,603258	6,04448	5,646816	8,930098
№	13	14	15	16	17	18	19	20	21	22	23	24
Fz-F	-5,023386778	-6,668925935	-2,41340148	-2,66854	-9,02848	-3,31052	-4,56862	-0,90077	-5,63588	-2,96528	-7,6602	-7,12627
№	25	26	27	28	29	30	31	32	33	34	35	36
Fz-F	-2,010604272	-1,931329369	-5,97595636	-2,20645	-10,5534	-3,19394	-4,51932	-0,36727	-6,43716	-0,9456	-5,83827	-2,90048