

КОВАЛЬЧУК ОЛЬГА

Західноукраїнський національний університет

ORCID ID: [0000-0001-6490-9633](https://orcid.org/0000-0001-6490-9633)e-mail: olhakov@gmail.com

МЕТОД ГОЛОВНИХ КОМПОНЕНТ ДЛЯ МОДЕЛЮВАННЯ ризиків кримінальної ЗЛОЧИННОСТІ

У роботі здійснено спробу ідентифікувати консолідовані фактори, що визначають значущі індивідуальні ознаки ув'язнених та підвищують ризик скоєння повторних кримінальних злочинів. Емпіричний аналіз проведено на основі унікального набору офіційних кримінальних записів 13010 засуджених, що відбувають покарання у пенітенціарних закладах України. Встановлено, що основними факторами, які визначають узагальнений профіль засуджених, є фактор ризику скоєння повторних кримінальних злочинів та віковий фактор, що консолідує вікові характеристики ув'язнених (реальний вік, вік на момент першого засудження до реальної чи/та умовної міри покарання). Доведено, що зі збільшенням кількості засуджень до реальної міри покарання зростає ризик скоєння повторних кримінальних злочинів. Виявлено прямий зв'язок кількості дострокових звільнень зі схильністю до кримінальних рецидивів. Визначено, що чим раніше особа була засуджена до реальної чи/та умовної міри покарання, тим більший ризик скоєння нею кримінальних рецидивів; зі збільшенням віку цей ризик зменшується. Для оцінювання впливу включених у факторну модель ознак на ризик скоєння засудженими повторних кримінальних злочинів побудовано лінійну багатofакторну регресійну модель. Встановлено, що найбільший вплив чинять кількість умовних засуджень та дострокових звільнень. Отримані оцінки підтверджують результати проведеного факторного аналізу. Зменшення строку покарання та заміна реального строку на умовний стимулюють засуджених до скоєння нових злочинів. Результати проведеного аналізу можуть надати важливі знання правоохоронним органам та органам правосуддя щодо оптимізації стратегії запобігання та протидії злочинності.

Ключові слова: кримінальні записи, ризик рецидивізму, індивідуальні ознаки злочинців, факторна модель, метод головних компонент.

KOVALCHUK OLHA

West Ukrainian National University

PRINCIPAL COMPONENTS ANALYSIS FOR MODELING RISKS OF CRIMINAL CRIME

In this work, an attempt was made to identify the main factors that determine significant individual characteristics of prisoners and increase the risk of committing repeated criminal offenses. The empirical analysis was conducted on the basis of a unique set of official criminal records of 13,010 convicts serving sentences in penitentiary institutions in Ukraine. It was established that the main factors determining the generalized profile of convicts are the risk factor of committing repeated criminal offenses and the age factor consolidating the age characteristics of prisoners (real age, age at the time of the first conviction before the real and/or suspended sentence). It has been proven that with an increase in the number of convictions to the actual degree of punishment, the risk of committing repeated criminal offenses increases. A direct relationship between the number of early releases and the propensity for criminal recidivism was revealed. It was determined that the earlier a person was sentenced to real and/or suspended punishment, the greater the risk of criminal recidivism; with increasing age, this risk decreases. A linear multivariate regression model was constructed to assess the influence of the factors included in the factor model on the risk of repeated criminal offenses by convicts. It was found that the number of suspended sentences and early releases exerted the greatest influence. The obtained estimates confirm the results of the conducted factor analysis. Reducing the term of punishment and replacing the real term with a suspended one encourages convicts to commit new crimes. The results of the conducted analysis can provide important knowledge to law enforcement agencies and justice authorities regarding the optimization of crime prevention and combating strategies.

Keywords: criminal records, risk of recidivism, individual characteristics of criminals, factor model, principal component analysis.

Постановка проблеми

Згідно з World Prison Brief's рівень злочинності у 2023 році зріс до понад 11,5 мільйонів засуджених, що відбувають покарання у в'язницях у всьому світі. Майже у 120 країнах рівень заповнюваності виправних закладів перевищує можливість пенітенціарної системи [1]. Витрати урядів на утримання засуджених складають колосальні суми. Проте існують регіональні тенденції щодо рівня злочинності, що часто пов'язані з такими факторами, як економічна нерівність, політична нестабільність, соціальна ізоляція та присутність організованої злочинності у конкретному регіоні. Крім того, наявність вогнепальної зброї часто пов'язана з підвищенням рівня насильницьких злочинів.

У глобальному рейтингу злочинності Україна займає 34 місце зі 187 країн [2]. Вже другий рік у країні йде повномасштабна війна, що спричинила економічну та соціальну нестабільність. На сьогодні в Україні на руках у населення близько 1,2 млн одиниць зареєстрованої зброї [3]. Окрім того на війну в Україні завербовано 17 тисяч російських в'язнів [4]. Судові та правоохоронні органи сьогодні потребують нових технологій з великим потенціалом для забезпечення соціальної та суспільної безпеки. Ефективні стратегії запобігання злочинності можуть допомогти поліції та органам кримінального правосуддя знизити рівень кримінальної небезпеки. Однак зв'язок між факторами ризику скоєння кримінальних злочинів та рівнем злочинності є складним і змінюється залежно від контексту. Існує необхідність надання правоохоронним органам якісної інформації та нових знань про фактори (характеристики засуджених), які

асоціюють зі скоєнням кримінальних злочинів. Особливо це стосується осіб з кримінальним минулим, оскільки ризик скоєння повторних кримінальних злочинів (рецидивів) значно вищий.

Аналіз останніх джерел

Низка науковців з різних країн намагалися виявити ключові індивідуальні ознаки особи, що взаємопов'язані зі схильністю до злочинної діяльності та підвищують ризики скоєння кримінальних рецидивів [5, 6]. Багато дослідницьких робіт присвячено аналізу злочинності на основі застосування сучасних математичних методів та інформаційних технологій, зокрема методів машинного навчання та data mining [7, 8]. Але існуюча неузгодженість з відбором вибірки та тривалістю спостереження завадили отримати надійні та якісні результати. С. Егберт та ін. досліджували технології прогнозування злочинів на основі аналізу даних і алгоритмічного виявлення шаблонів, які використовують для попередження кримінальних злочинів як елементи превентивних стратегій в німецькомовних країнах [9] та використовували дані про вироки районних судів та оцінювали рівень рецидивів серед осіб, звільнених із шведських в'язниць [10].

Результати порівняльного аналізу чинників, що асоціюють з особою злочинця та визначають ризик скоєння кримінальних рецидивів у різних країнах підтвердили, що міжнародні дані є неспівставними [11, 12]. В Україні дослідження такого плану є вкрай рідкісними [13–15]. Відсутність системного підходу та єдиної методології наукового аналізу ознак осіб, схильних до кримінальної діяльності та скоєння рецидивів злочинів, призводять до проблем та помилок слідчих органів. Для виявлення та оцінювання таких ризиків та ефектів актуальними є різнопланові прикладні дослідження з цієї тематики.

Метою роботи є ідентифікація основних факторів (узагальнених характеристик засуджених), що визначають значущі індивідуальні ознаки ув'язнених та підвищують ризик скоєння повторних кримінальних злочинів, на основі реальних даних кримінальних записів (інформації про судимості) 13010 засуджених, що відбувають покарання у пенітенціарних закладах України.

Виклад основного матеріалу

Однією із стратегій розслідування, яку використовують правоохоронні органи для виявлення ймовірних підозрюваних та для встановлення зв'язків між справами, які могли бути вчинені тим самим злочинцем, є профілювання правопорушників (кримінальне профілювання) [16]. Аналіз кримінального розслідування базується на припущенні, що відмінності особи правопорушника можуть бути виведені шляхом систематичного дослідження відмінностей правопорушення. Для формування ефективних стратегій внутрішньої безпеки країни необхідне розуміння присутніх якостей та характеристик, притаманних більшості індивідів, які скоюють кримінальні злочини.

Величини, які визначають характеристики (фактори) консолідованого профілю засудженого неможливо виміряти безпосередньо. Окрім того, остаточно не визначено не лише перелік, а й оптимальну кількість та реальний зміст цих факторів. Насправді їх кількість може бути суттєво меншою, ніж кількість характеристик, які зберігаються у кримінальних записах. Для виявлення значущих неочевидних закономірностей, що залежать від індивідуальних ознак засуджених, та дослідження статистичного зв'язку виділених ознак з суттєвими факторами застосовано один із методів факторного аналізу – метод головних компонент. Емпіричні дослідження проведено у середовищі статистичного пакету Statistica на основі офіційних даних кримінальних записів 13010 засуджених, що відбувають покарання за скоєння кримінальних злочинів у пенітенціарних закладах України [17]. При проведенні аналізу використано наступні змінні (характеристики засуджених):

- *Age* – вік;
- *AFA* – вік на момент першого засудження до реальної міри покарання;
- *AFC* – вік на момент першого засудження до реальної або умовної міри покарання;
- *Real Convictions* – кількість реальних засуджень;
- *Suspended Convictions* – кількість умовних засуджень;
- *Early Dismissals* – кількість дострокових звільнень;
- *Recidivism* – наявність рецидивів: 0 – ні, 1 – так.

У результаті застосування факторного аналізу для визначення присутніх ознак осіб, засуджених за скоєння кримінальних злочинів, виділено 2 фактори, наведені на рис. 1.

З першим фактором найтісніше пов'язані факт скоєння кримінальних рецидивів (0,86), кількість засуджень до реальної міри покарання (0,83) та кількість дострокових звільнень (0,71). Усі факторні навантаження мають додатний знак, що свідчить про прямий зв'язок змінних з фактором 1. Зі збільшенням кількості засуджень до реальної міри покарання зростає ризик скоєння особою повторних кримінальних злочинів. Поки виправна система не сприяє виправленню та соціалізації засуджених. Дострокові звільнення злочинці сприймають не як шанс на повернення до законотворчого способу життя, а радше як відчуття безкарності. На другий виділений фактор найбільше впливають вік на момент першого засудження до реальної міри покарання (0,86) та вік на момент першого засудження до реальної або умовної міри

| Variable | Factor 1 | Factor 2 |
|-----------------------|-----------|-----------|
| Age | 0,370131 | -0,756348 |
| AFA | -0,365355 | -0,865344 |
| AFC | -0,377584 | -0,856036 |
| Real Convictions | 0,834772 | 0,072756 |
| Suspended Convictions | 0,295429 | 0,173614 |
| Residivism | 0,860595 | 0,139714 |
| Early Dismissals | 0,710619 | 0,014861 |
| Expl.Var | 2,442776 | 2,108856 |
| Prp.Totl | 0,348968 | 0,301265 |

Рис. 1. Факторні навантаження

покарання (0,85) та вік засудженого (0,75) для заданого мінімального факторного навантаження 0,7. Факторні навантаження усіх змінних мають від'ємний знак, що вказує на зворотний зв'язок змінних з фактором 2. Це означає, що чим раніше особа була засуджена до реальної чи/та умовної міри покарання, тим більший існує ризик, що вона скоїть повторні кримінальні злочини. Окрім того, зі збільшенням віку ризик скоєння кримінальних рецидивів зменшується. Виділені фактори можна трактувати як «кримінальний» чинник (фактор 1), який є узагальненою оцінкою підвищення ризику скоєння повторних кримінальних злочинів засудженими та віковий чинник (фактор 2), що узагальнює значущі вікові характеристики ув'язнених. Це підтверджує і графік факторних навантажень, зображений на рис. 2.

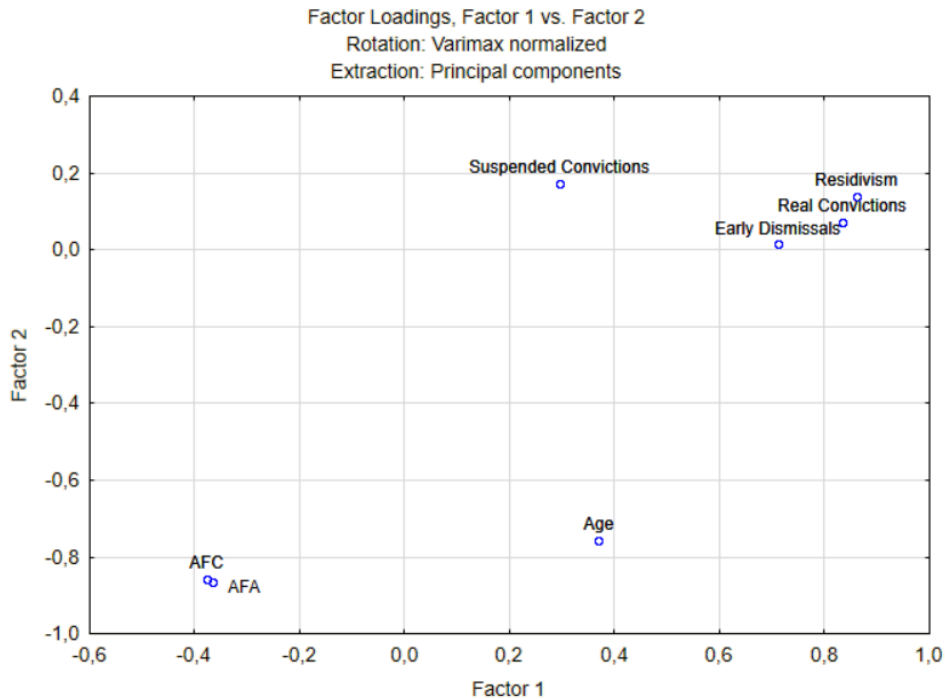


Рис. 2. Графічне представлення факторних навантажень

Для встановлення значущості кожного з виділених факторів проведено аналіз власних значень факторів (рис. 3). Перший виділений фактор пояснює 51 % дисперсії, другий – 34 %. Разом вони описують 85 % дисперсії, тобто більше 4/5 масиву даних. Це означає, що проведена факторизація неповна, існують ще й інші чинники, менш значущі, однак теж достатньо важливі. Зокрема, стать, суворість режиму ув'язнення, рівень освіти, вживання психоактивних речовин тощо.

| Extraction: Principal components | | | | |
|----------------------------------|------------|------------------|-----------------------|--------------|
| Value | Eigenvalue | % Total variance | Cumulative Eigenvalue | Cumulative % |
| 1 | 2,900789 | 51,43984 | 2,900789 | 51,43984 |
| 2 | 1,650843 | 33,58347 | 4,551632 | 85,02331 |

Рис. 3. Власні значення виділених факторів

У результаті проведеного аналізу отримано наступну факторну модель (1):

$$\begin{aligned}
 F_1 &= 0,86 \times R + 0,83 \times RC + 0,71 \times ED, \\
 F_2 &= -0,87 \times AFA - 0,86 \times AFC - 0,76 \times Age,
 \end{aligned}
 \tag{1}$$

де R – наявність рецидивів, RC – кількість реальних засуджень, ED – кількість дострокових звільнень, AFA – вік на момент першого засудження до реальної міри покарання, AFC – вік на момент першого засудження до реальної або умовної міри покарання, Age – вік.

Виділені фактори по суті є лінійними комбінаціями досліджуваних характеристик засуджених, що представляють більшу частку загальної мінливості аналізованих чинників, тобто відображають основну частку початкової інформації. Побудована факторна модель достатньо повно описує початковий масив даних і в той же час є більш зручною для змістовної інтерпретації отриманих результатів.

Для оцінювання впливу включених у факторну модель ознак на ризик скоєння засудженими повторних кримінальних злочинів побудовано лінійну багатофакторну регресійну модель для залежної змінної *Number of Recidiv* (кількість повторних злочинів) (рис. 4).

| Regression Summary for Dependent Variable: Number of Residiv | | | | | | |
|--|-----------|------------------|-----------|---------------|----------|----------|
| R=,879714 R²=,707953 Adjusted R²=,707802 | | | | | | |
| F(5,12959)=4019,1 p<0,0000 Std.Error of estimate:1,4621 | | | | | | |
| N=12965 | Beta | St. Err. of Beta | B | St. Err. of B | t(12959) | p-value |
| Intercept | | | 1,253923 | 0,047847 | 26,2072 | 0,000000 |
| Age | 0,254238 | 0,006134 | 0,045772 | 0,001104 | 41,4454 | 0,000000 |
| AFA | -0,264989 | 0,012933 | -0,063373 | 0,003093 | -20,4886 | 0,000000 |
| AFC | -0,069415 | 0,012932 | -0,015672 | 0,002920 | -5,3678 | 0,000000 |
| Early Dismissals | 0,213688 | 0,005860 | 1,094068 | 0,030003 | 36,4649 | 0,000000 |
| Suspended Convictions | 0,555074 | 0,005747 | 0,998066 | 0,010334 | 96,5835 | 0,000000 |

Рис. 4. Оцінки багатofакторної лінійної регресійної моделі для залежної змінної *Number of Recidive*

У результаті проведеного аналізу отримано специфікацію лінійної багатofакторної регресійної моделі (2):

$$NR = 0,05 \times Age + 0 - 0,05 \times AFA - 0,02 \times AFC + 1,10 \times ED + 1,00 \times SC + 1,25, \quad (2)$$

де NR – кількість попередніх судимостей, Age – вік, AFA – вік на момент першого засудження до реальної міри покарання, AFC – вік на момент першого засудження до реальної або умовної міри покарання, ED – кількість дострокових звільнень.

Залежність між відгуком та досліджуваними ознаками помірна ($R^2 = 0,7$). Лінійна багатofакторна модель адекватно описує взаємозв'язок між залежною та незалежними змінними. Обчислене $p < 0,001$. $F_{\text{емп}}(5,12959) = 4019,1$. $F_{\text{теор}} = F(5, \infty) = 4,36$. $F_{\text{емп}} > F_{\text{теор}}$. Гіпотезу про залежність досліджуваних чинників відхиляємо. Вільний член рівня (Intersept) є статистично значущим ($p < 0,001$). Його значення (1,25) вказує, що існують й інші ознаки, не включені у модель, які можуть чинити вплив на схильність до скоєння рецидивів засудженими. Оцінки усіх регресійних коефіцієнтів є значущими ($p < 0,001$). Для визначення значущості коефіцієнтів регресійної моделі застосовано t -критерій Стьюдента. Для всіх змінних моделі та для вільного члена рівняння $t_{\text{емп}} > t_{\text{теор}}$, $t_{\text{теор}} = t(\infty) = 3,29$.

Найбільший вплив на ризик скоєння повторних кримінальних злочинів мають кількість умовних засуджень та дострокових звільнень. Отримані оцінки підтверджують результати проведеного факторного аналізу. Зменшення строку покарання та заміна реального строку на умовний стимулюють засуджених до скоєння нових злочинів.

Висновки

Встановлено, що основними компонентами, що визначають консолідований профіль засуджених, які відбувають покарання за скоєння кримінальних злочинів у пенітенціарних закладах України, є фактор ризику скоєння повторних кримінальних злочинів та віковий фактор, що описує консолідовані вікові характеристики ув'язнених (реальний вік засудженого та його вік на момент першого засудження до реальної чи/та умовної міри покарання). Побудована факторна модель дає можливість оцінити вагу кожної аналізованої ознаки засуджених на виділені фактори ризику схильності до злочинної діяльності та кримінальних рецидивів. Визначено, що кількість засуджень до реальної міри покарання прямо корелює з ризиком скоєння особою повторних кримінальних злочинів; збільшення кількості дострокових звільнень засуджених призводить до скоєння ними повторних злочинів. Встановлено, що чим раніше особа була засуджена до реальної та/чи умовної міри покарання, тим більший ризик, що в майбутньому вона скоїть повторне правопорушення; зі збільшенням віку засуджених ризик скоєння кримінальних рецидивів зменшується. Пенітенціарна система не робить злочинців законослухняними громадянами, а поблажливість судової системи посилює відчуття безкарності. Отримані оцінки можуть надати важливі відомості правоохоронним та судовим органам щодо оптимізації тактики проведення превентивних дій щодо запобігання вчиненню повторних злочинів та забезпечення суспільної безпеки.

References

1. Global Prison Trends 2023. PRT's annual flagship report. Thailand Institute of Justice. <https://cdn.penalreform.org/wp-content/uploads/2023/06/GPT-2023.pdf>.
2. Crime Rate by Country. Wisevoter. <https://wisevoter.com/country-rankings/crime-rate-by-country/>.
3. BBC News Україна. <https://www.bbc.com/ukrainian/articles/c8vrd3n8jmno>.
4. Litavrin M., Pavlova A. The Wagner map. 17,000 inmates unaccounted for in 35 Russian regions amid recruitment for war in Ukraine. Mediazona. <https://en.zona.media/article/2023/04/25/wagnermapped>.
5. Garritsen K., Jankovic M., Masthoff E., Caluwe E.D., Bogaerts S. The Role of Dynamic Risk and Protective Factors in Predicting Violent Recidivism: Intellectual Ability as a Possible Moderator? International Journal of Offender Therapy and Comparative Criminology. 2022. 52271.
6. Saravanan P., Selvaprabu J., Raj L.A., Khan A., Sathick K. Survey on crime analysis and prediction using data mining and machine learning techniques. Lect. Notes Electr. Eng. 2021. № 688. P. 435–448.

7. Jabeen N., Agarwal P. Data Mining in Crime Analysis. Proceedings of the 2th Second International Conference on Smart Energy and Communication. 2021, Jaipur, India.
8. Ogochukwu O., Forster O. An Overview of Crime Analysis, Prevention, and Prediction Using Data Mining Based on Real Time and Location Data. *Int. J. Eng. Appl. Sci. Technol.* 2021. № 5. P. 99–103.
9. Egbert S., Krasmann S. Predictive Policing: Not Yet, but Soon Preemptive? *Policing and Society.* 2020. №. 30(8). P. 905–919.
10. Yu R., Langstro N., Forsman M., Sjo A., Fazel S., Molero Y. Associations between prisons and recidivism: A nationwide longitudinal study. *PLoS ONE.* 2022. 17(5):e0267941.
11. Fazel S., Wolf A. A Systematic Review of Criminal Recidivism Rates Worldwide: Current Difficulties and Recommendations for Best Practice. *PLoS ONE.* 2015. № 10:e0130390.
12. Yukhnenko D., Sridhar S., Fazel S. A systematic review of criminal recidivism rates worldwide: 3-year update. *PubMed Cent.* 2019. № 4(28).
13. Kovalchuk O., Karpinski M., Banakh S., Kasianchuk M., Shevchuk R., Zagorodna N. Prediction Machine Learning Models on Propensity Convicts to Criminal Recidivism. *Information.* 2023. № 14(3). P. 161.
14. Kovalchuk O., Kasianchuk M., Karpinski M., Shevchuk R. Decision-Making Supporting Models Concerning the Internal Security of the State. *INTL Journal of Electronics Telecommunications.* 2023. № 96(2). P. 301–307.
15. Berezka K., Kovalchuk O., Banakh S., Zlyvko S., Hrechaniuk R. A Binary Logistic Regression Model for Support Decision Making in Criminal Justice. *Folia Oeconomica Stetinensia.* 2022. № 22(1). P. 1–17.
16. Kovalchuk O., Banakh S. Logic of Law: Criminal Profiling Strategy. *Actual Problem of Law,* 2022, vol. 3(31), pp. 169–2175.
17. Unified register of pre-trial investigations. <https://erdr.gp.gov.ua>.