УДК 004.94:336.7

Р. Н. КВЄТНИЙ, С.І. БОРОДКІН

ПОКРАЩЕНА МОДЕЛЬ РЕГУЛЯРИЗАЦІЇ ELASTIC NET ДЛЯ ОБРОБКИ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ

Вінницький національний технічний університет, Хмельницьке шосе 95, 21021, Вінниця, Україна, e-mail:borserg90@gmail.com

Анотація: У статті запропоновано модифікацію Elastic Net-регресії для короткострокового прогнозування фінансових часових рядів шляхом введення гаусівського затухання ваг (Gaussian decay). Новий підхід спрямований на згладжування різких «стрибків» між останнім історичним і першим прогнозним значеннями, характерних для стандартної регуляризації. Для оцінки ефективності було формально виписано Elastic Net з чотирма схемами затухання ваг (без затухання, лінійне, експоненційне, гаусівське) та проведено емпіричні експерименти на даних індексів S&P 500, Dow Jones Industrial Average і Nasdaq Composite за 2020–2025 рр. Результати продемонстрували, що Gaussian decay мінімізує перехідний розрив і забезпечує найнижчі значення RMSE і Deviation для S&P 500 і Nasdaq, тоді як для Dow Jones оптимальною виявилася експоненційна схема.

Ключові слова: Elastic Net, Gaussian-затухання, часові ряди, обробка даних, фінансове прогнозування, вагове зважування, S&P 500, Dow Jones, Nasdaq Composite

Abstract. This paper proposes a modification of Elastic Net regression for short-term forecasting of financial time series by introducing Gaussian weight decay. The new approach is designed to smooth the abrupt “jumps” between the last historical observation and the first forecast—an issue typical of standard regularization. To assess its effectiveness, we formally derive the Elastic Net model with four weighting schemes (no decay, linear, exponential, and Gaussian) and conduct empirical experiments on the S&P 500, Dow Jones Industrial Average, and Nasdaq Composite indices over the period 2020–2025. The results demonstrate that Gaussian decay minimizes the transition gap and achieves the lowest RMSE and Deviation for the S&P 500 and Nasdaq Composite, whereas exponential decay proves optimal for the Dow Jones Industrial Average.

Keywords: Elastic Net, Gaussian weight decay, time series data processing, time series forecasting, financial markets, adaptive weighting, S&P 500, Dow Jones, Nasdaq Composite

DOI: 10.31649/1681-7893-2025-49-1-29-35

ВСТУП

Прогнозування фінансових часових рядів залишається актуальною проблемою через складну структуру ринкових даних, що включають волатильність, наявність структурних змін та викидів. Класичні методи регуляризації, зокрема Elastic Net, який комбінує штрафи L_1 та L_2 для одночасного відбору ознак і зменшення мультиколінеарності. Цей метод довів свою ефективність у багатьох задачах регресії завдяки збалансованому поєднанню жорсткого відсікання незначущих змінних і плавного стискання великих коефіцієнтів [1].

Однак при застосуванні Elastic Net до прогнозування часових рядів часто спостерігається різкий «стрибок» між останнім історичним значенням і першим кроком прогнозу, що знижує довіру до моделі. Для адаптації до нестабільності часового ряду застосовують часове зважування спостережень: свіжі дані отримують вищі ваги, старі – нижчі.

Так, для прогнозування часових рядів було запропоновано TWLS (Time-Weighted Least Squares), яке показало статистично значущу перевагу над звичайною регресією завдяки більшому врахуванню останніх спостережень [3]. У сфері регуляризованих моделей впроваджено Lag-weighted Lasso – модифікацію Lasso з різними коефіцієнтами штрафу залежно від лагу предиктора, а для підвищення робастності до викидів реалізовано вагові схеми на основі М-оцінок [4].

Таким чином, сучасні дослідження підтверджують доцільність поєднання регуляризації (для боротьби з перенавчанням і вибору важливих лагів) та часового вагування даних (для акцентування уваги на свіжих спостереженнях) при моделюванні фінансових часових рядів.

Метою роботи є підвищення точності прогнозування фінансових часових рядів за допомогою Elastic Net із адаптивним ваговим затуханням (лінійним, експоненційним та гаусівським), при якому оптимальні гіперпараметри підбираються через крос-валідацію, а ваги спостережень змінюються так, щоб мінімізувати «стрибок» між останнім історичним і першим прогнозним значеннями.

1. ПОСТАНОВКА ELASTIC NET – МОДЕЛІ З РІЗНИМИ СХЕМАМИ ЗАТУХАННЯ ВАГ

Часто фінансові часові ряди характеризуються множинними змінами, які мають нелінійну природу та неоднорідну дисперсію. Для стабілізації дисперсії і перетворення мультиплікативних змін у адитивні використовується логарифмічне перетворення. Крім того, застосування щомісячних значень замість щоденних дозволяє зменшити рівень шуму і поліпшити довгострокову прогнозованість [5].

Нехай маємо часовий ряд y_t – значення log-індексу на місяць t.

Для прогнозування використовуються лагові ознаки:

$$x_{t,k} = y_{t-k}, \tag{1}$$

де $k = 1,2, \dots, p$ – номер лагу, p – максимальний лаг.

Тоді одноперіодний прогноз моделі можна записати як лінійну комбінацію лагів:

$$\hat{y}_{t+1} = \beta_0 + \sum_{k=1}^{p} \beta_k \, y_{t+1-k} + \varepsilon_t, \tag{2}$$

де β_0 – константа, β_k – коефіцієнти регресії при k-му лагу, ε_t – шум.

Такий опис відповідає моделям авторегресії (AR), де поточне значення часового ряду змінюється як лінійна комбінація минулих значень [6].

Для оцінювання параметрів β використовується метод найменших квадратів з ElasticNet-регуляризацією. У загальному випадку, за наявності вагових коефіцієнтів w_i для кожного спостереження (x_i, y_i), оцінка параметрів знаходиться з розв'язку задачі мінімізації функції втрат:

$$\hat{\beta} = arg \min_{\beta_0,\beta} \left\{ \tfrac{1}{2} \textstyle\sum_{i=1}^{N} w_i (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \left(\alpha \|\beta\|_1 + \tfrac{1-\alpha}{2} \|\beta\|_2^2 \right) \right\}, \tag{3}$$

де N – кількість спостережень у навчальній вибірці, w_i – вага і-го спостереження, $x_i = \left(y_{i-1}, \dots, y_{i-p}\right)^\top$ – вектор лагових предикторів, $\beta = \left(\beta_1, \dots, \beta_p\right)^\top$– вектор коефіцієнтів, β_0 – вільний член (intercept), $\lambda > 0$ – параметр регуляризації, $\alpha \in [0,1]$ – параметр, що визначає співвідношення між L_1 та L_2-штрафами. При $\alpha = 1$ модель (3) еквівалентна Lasso, при $\alpha = 0$ – гребеневій регресії, а при $0 < \alpha < 1$ власне ElasticNet-регресії.

Регуляризація зменшує модульні значення коефіцієнтів β_k, а при достатньо великому λ деякі з них стають рівними нулю, що автоматично здійснює відбір важливих лагів. Під час навчання λ та α підбираються автоматично методом крос-валідації [1,2,4].

Вага w_i визначає значущість i-го спостереження при навчанні моделі. У рівнянні (3) через множники w_i вплив квадратів помилок кожного спостереження на функцію втрат є неоднаковим. Ми розглядаємо чотири схеми задання ваги як функції від порядкового номера спостереження (або його віку) у вибірці [1]:

$$w_i = \begin{cases} 1, & \text{рівномірне (без зважування);} \\ \frac{1}{N}, & \text{лінійне зважування;} \\ \exp\left(-\lambda_w (N-i)\right), & \text{експоненціальне зважування;} \\ exp\left[-\frac{(N-i)^2}{2\sigma_w^2}\right], & \text{гаусівське зважування.} \end{cases} \quad (4)$$

Перша схема (рівномірна) означає відсутність ваг – всі спостереження внесуть однаковий вклад ($w_i = 1$ для всіх i). Лінійна схема задається вагами, що зростають пропорційно номеру спостереження: найбільш давнє спостереження отримає вагу $\frac{1}{N}$, а найновіше – 1 [2]. Експоненціальна схема передбачає геометричне згасання ваг: параметр $\lambda_w > 0$ визначає швидкість експоненційного зменшення ваги зі збільшенням віку спостереження. Нарешті, гаусівська схема реалізує "віконне" ядро, зосереджене на останніх спостереженнях: коефіцієнт σ_w визначає ширину "вікна" (стандартне відхилення гаусівської функції), в межах якого дані отримують значну вагу, тоді як дуже старі спостереження (на відстані більше кількох σ_w від останнього) мають мізерно малі w_i [4,7]. В усіх схемах більш пізні спостереження отримують не меншу вагу, ніж більш ранні (для рівномірної всі рівні).

Таким чином, реалізується принцип забування застарілої інформації: модель більше фокусується на нових даних, що потенційно відображають актуальні ринкові умови.

2. МЕТОДОЛОГІЯ ЕКСПЕРИМЕНТУ

Для аналізу використано часовий ряди популярних фондових індексів S&P 500, Dow Jones та Nasdaq Composite. Спершу до даних застосовано логарифмічне перетворення: $y_t = \log(Index_t)$. Це робиться з метою стабілізації дисперсії ряду та перетворення мультиплікативних змін у адитивні. Далі з отриманого ряду сформовано ознаки – лагові значення $y_{t-1}, y_{t-2}, \dots, y_{t-p}$. В нашому дослідженні обрано максимальне лагове число $p = 10$, що відповідає використанню даних за останніх десять місяців для прогнозування наступного місяця. Вектор цільової змінної при навчанні зміщено на один крок вперед відносно ознак: для кожного часу t (починаючи з p-го спостереження) вхідними даними є $x_t = (y_{t-1}, \dots, y_{t-p})$, а виходом моделі – прогноз $\widehat{y_t}$ поточного значення (яке порівнюється з фактичним y_t). У нашому випадку для кожного з чотирьох індексів сформовано окремий навчальний набір: ознаки складаються виключно з лагових значень того ж самого індексу, а ціль – його власне значення наступного місяця. Таким чином, було підготовлено чотири незалежні датасети (по одному для S&P 500, Dow Jones та Nasdaq) без «перемішування» даних між різними індексами. Це означає, що моделі будувалися для кожного індексу окремо, враховуючи лише власну динаміку кожного ряду [5,6].

Для моделювання було обрано лінійну регресійну модель Elastic Net, яка комбінує L1 та L2-регуляризацію з крос-валідацією. Для цього, вибірка розбивається на навчальну та тестову частини. Для оцінки узагальнюючої здатності моделі та підбору оптимальних гіперпараметрів λ і α використано процедуру крос-валідації. Зокрема, у навчальній вибірці виконано k-складову часову крос-валідацію: дані поділено на $k = 10$ послідовних інтервалів часу; модель навчалася на перших $(k-1)$ інтервалах і перевірялася на наступному інтервалі, що повторювалося k разів зі зміщенням вікна. Такий підхід (rolling-origin evaluation) дозволяє уникнути «заглядання в майбутнє» при валідації та коректно оцінити параметри на часових рядах . На кожній ітерації крос-валідації підбиралося оптимальне значення λ (через перебір декількох десятків кандидатів по логарифмічній шкалі) та параметр α (перебором декількох значень від 0 до 1). Критерієм вибору була мінімальна середня квадратична помилка на валідаційних інтервалах. Для реалізації навчання використано стандартний алгоритм координатного спуску (реалізація python sklearn), що розв'язує задачу (3) з врахуванням вагових коефіцієнтів. В процесі навчання моделі з вагуванням лінійним, експоненційним та гаусівським вага старіших спостережень автоматично зменшується, що зменшує їх вплив на оцінку параметрів і, як очікується, підвищує стабільність моделі до зміни тренду [1,2].

Після налаштування гіперпараметрів модель було перенавчано на всій навчальній вибірці і отримано фінальні оцінки $\widehat{\beta}_0, \widehat{\beta}$. Для кожної вагової схеми (включно з випадком без ваг) побудовано прогнозні значення $\widehat{y_t}$ на тестовому відрізку даних (який моделі не бачили при навчанні). Прогноз здійснювався ітеративно в режимі одномісячного горизонту: використовуючи фактичні значення попередніх p місяців для кожного кроку вперед. З отриманих логарифмічних прогнозів $\widehat{y_t}$ виконано зворотне перетворення для отримання $\widehat{Y_t}$ – прогнозів індексу [2,5].

Для оцінювання якості прогнозування було обрано дві метрики: середньоквадратична помилка прогнозу (RMSE) та відхилення першого прогнозного значення від останнього історичного (Deviation).

Середньоквадратична помилка (RMSE) є універсальною метрикою, яка дозволяє оцінити точність прогнозування, приділяючи більше уваги великим похибкам [8]:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \widehat{y_t})^2}, \tag{5}$$

де y_t – фактичні значення, $\widehat{y_t}$ – прогнозні значення, N – кількість прогнозних значень.

Deviation (відхилення першого прогнозного значення від останнього історичного):

$$Deviation = |y_{t+1} - y_t|, \tag{6}$$

де y_t – останнє історичне значення перед прогнозним періодом, y_{t+1} – перше прогнозоване значення.

Deviation використовується для оцінки плавності переходу від історичних до прогнозних значень, що особливо важливо в контексті фінансових рядів, де різкі скачки можуть свідчити про недостатню адаптивність моделі.

Ці дві метрики дозволяють комплексно оцінити якість моделі, враховуючи як загальну точність прогнозу, так і специфічну поведінку моделі на межі між історичним і прогнозним періодами.

3. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТУ

Нижче наведено результати прогнозування на тестовому відрізку з 30.09.2024 по 31.05.2025 (історичні дані з 01.01.2020 по 30.09.2024) для індексів S&P 500, Dow Jones, Nasdaq Composite з використанням моделі Elastic Net за чотирма різними ваговими схемами. Таблиця 1 містить значення RMSE, показника Deviation (різниця між першим прогнозом і останнім фактичним значенням навчальної вибірки) та коефіцієнтів α, λ, σ, які були підібрані за допомогою крос-валідації для кожної комбінації індексу та вагової схеми.

Таблиця 1

Порівняльний аналіз результатів регресії для індексів S&P 500, Dow Jones та Nasdaq Composite з різними ваговими схемами затуханнями.

Індекс	Вагова схема	RMSE	Deviation	α	λ	σ
S&P500	none (без ваг)	0.041635	-84.6942	0.000452	1	-
S&P500	linear	0.035143	41.4909	0.000452	1	-
S&P500	exponential	0.035846	45.0944	0.000728	1	-
S&P500	gaussian	0.034833	19.56642	0.000452	1	0.57551
Dow Jones (DJIA)	none (без ваг)	0.059641	-975.508	0.000174	0.9	-
Dow Jones (DJIA)	linear	0.046607	-126.452	0.000108	1	-
Dow Jones (DJIA)	exponential	0.045999	-212.089	0.000174	0.9	-
Dow Jones (DJIA)	gaussian	0.052202	-595.758	0.000174	1	0.6041
Nasdaq Composite	none (без ваг)	0.063502	-273.448	2.59E-05	0.9	-
Nasdaq Composite	linear	0.052625	34.60765	6.72E-05	1	-
Nasdaq Composite	exponential	0.04665	207.0917	1.61E-05	1	-
Nasdaq Composite	gaussian	0.04098	-161.963	0.001887	1	0.2

Проведений експеримент із застосуванням Elastic Net-регресії з різними схемами вагового затухання до прогнозування фондових індексів (S&P 500, Dow Jones і NASDAQ Composite) продемонстрував, що вибір схеми суттєво впливає на точність прогнозу.

Для Dow Jones (DJIA) найкращий результат за показником RMSE показала експоненційна вагова схема (RMSE = 0.045999). Водночас усі моделі демонструють значні негативні відхилення, що свідчить про систематичну недооцінку прогнозованих значень, особливо у випадку моделі без ваг (Deviation = -975.508).

Для NASDAQ Composite найнижче значення RMSE (0.04098) отримано з використанням гаусівського затухання. Проте помітним є також значення Deviation = -161.963, що свідчить про деяку систематичну недооцінку. Водночас експоненційна схема показує трохи гірший RMSE, але позитивне зміщення (Deviation = 207.0917), що вказує на переоцінювання прогнозу.

Для індексу S&P 500 найменше значення RMSE (0.034833) і найменше абсолютне відхилення від фактичних значень (Deviation = 19.56642) забезпечила модель із застосуванням гаусівського затухання. Ця модель, на відміну від інших, має оптимальний баланс параметрів регуляризації ($\alpha = 0.000452$, $\lambda = 1$, $\sigma = 0.57551$), завдяки чому модель найбільш точно враховує останні тенденції ринку.

Для візуального порівняння точності моделей і оцінки їхньої відповідності реальним даним для прикладу наведемо графік прогнозних і фактичних значень індексу S&P 500 за різними схемами вагового затухання Рис. 1:

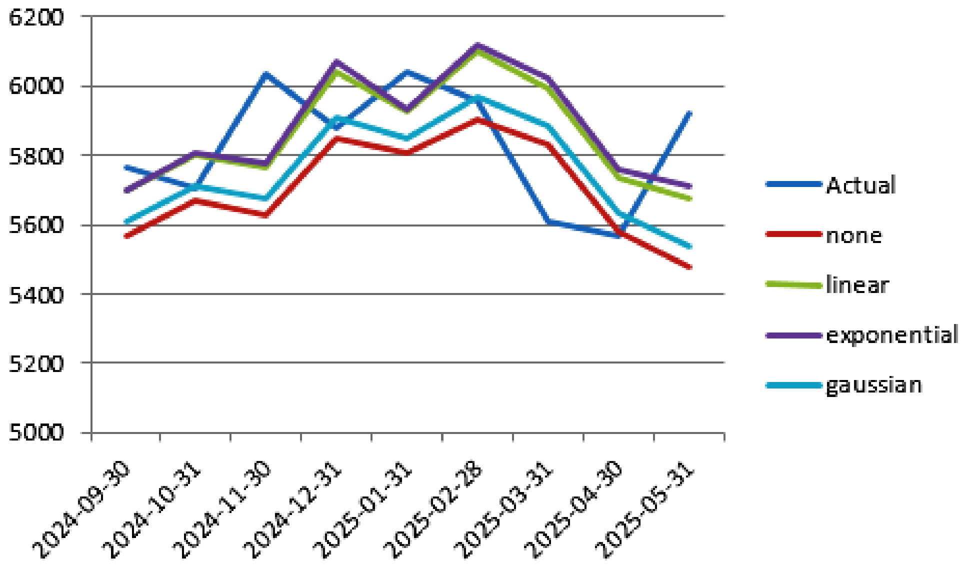

Рисунок 1 – Графік прогнозу індексу S&P 500

На цьому графіку чітко видно, що прогноз з гаусівського затуханням найточніше відповідає фактичним значенням, мінімізуючи помилку особливо у другій половині періоду прогнозування.

Для кращого розуміння характеру впливу даних різної давності на прогноз, наведемо також графік вагових коефіцієнтів для кожної з використаних схем Рис.2.

Графік вагових коефіцієнтів дозволяє краще зрозуміти характер використання інформації у моделях. Для моделі без ваг (none) всі спостереження мають однаковий внесок, незалежно від їх давності. Лінійна схема поступово збільшує вагу новіших спостережень, тоді як експоненційна швидко підвищує вагу нових даних та знижує значення старіших. Гаусівська схема має найбільш різке зменшення ваг старіших спостережень, зосереджуючи практично всю увагу моделі лише на найсвіжіших точках, що і є причиною її найвищої точності в умовах нестабільних фінансових рядів.

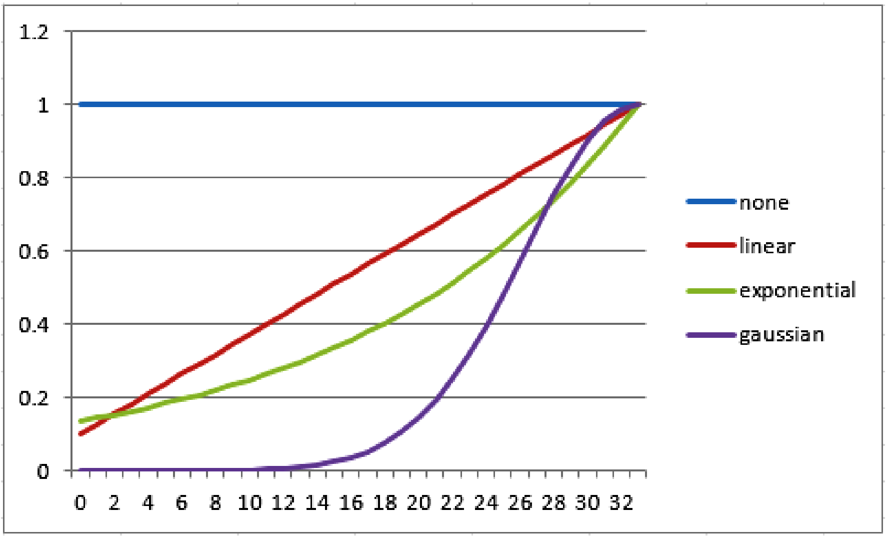

Рисунок 2 – Графік вагових коефіцієнтів різних індексу S&P 500

Таким чином, використання вагового затухання у прогнозуванні фінансових часових рядів суттєво покращує якість прогнозу, а гаусівське затухання є оптимальним рішенням для прогнозування динаміки індексу S&P 500.

ВИСНОВКИ

У статті запропоновано модифікацію ElasticNet-регресії з адаптивним вагуванням (рівномірна, лінійна, експоненційна та гаусівська схеми) для короткострокового прогнозування індексів S&P 500, Dow Jones і Nasdaq Composite за даними 2020–2025 рр.

Експерименти показали, що Gaussian decay виявився оптимальним для S&P 500 і Nasdaq (найнижчий RMSE і мінімальне відхилення), у той час як exponential decay дав найкращий результат для Dow Jones. Схема без вагування продемонструвала найнижчу точність.

Модель з Gaussian decay найкраще себе показала при прогнозуванні індексу S&P 500 з найнижчими RMSE = 0.034833 та мінімальним абсолютним відхиленням Deviation = 19.56642. Вона досягла цього за рахунок оптимального поєднання параметрів регуляризації: $\alpha = 0.000452$, $\lambda = 1$ та $\sigma = 0.57551$, що дозволяє максимально точно відобразити останні ринкові тенденції. Зокрема, Gaussian decay є перспективним підходом, який забезпечує стійке поліпшення точності прогнозу порівняно з традиційними методами, що не диференціюють давність даних. Перспективи подальших досліджень включають випробування запропонованої методології на інших частотах даних (наприклад, денних або квартальних інтервалах), введення нових метрик, а також множинної регресії між багатьма змінними.

Загалом, результати роботи підтверджують, що врахування часового фактору давності даних через вагові коефіцієнти є дієвим засобом підвищення точності прогнозування фінансових часових рядів і може бути рекомендовано для практичного використання у фінансовій аналітиці та ризик менеджменті.

СПИСОК ЛІТЕРАТУРИ

1. Hastie, T., Tibshirani, R., & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edition. Springer, – 2009 – 533p, https://doi.org/10.1007/978-0-387-84858-7.
2. Zou H., Hastie T. *Regularization and variable selection via the elastic net*. J. R. Stat. Soc. Series B, 2005, 67(2): 301–320, https://doi.org/10.1111/j.1467-9868.2005.00503.x.
3. Wang Y., Hao X., Wu C. *Forecasting stock returns: A time-dependent weighted least squares approach*. Journal of Financial Markets, 2021, 53: 100568, https://doi.org/10.1016/j.finmar.2020.100568.
4. Dikheel T.R., Yaseen A.Q. *Robust lag weighted lasso for time series model*. J. Modern Applied Statistical Methods, 2021, 19(1): 14, https://doi.org/10.56801/10.56801/v19.i.1081.
5. Lütkepohl, H., Xu, F. *The role of the log transformation in forecasting economic variables*. Empir Econ 42, 2012, 619–638, https://doi.org/10.1007/s00181-010-0440-1.
6. Tsay R.S. Analysis of Financial Time Series. 2nd ed. Wiley, 2010, https://doi.org/10.1002/9780470644560.

7. E. S. Gardner Jr., *Exponential Smoothing: The State of the – Part II.* International Journal of Forecasting, 22(4), 637–666, https://doi.org/10.1016/j.ijforecast.2006.03.005
8. R.J. Hyndman, A.B. Koehler. *Another look at measures of forecast accuracy*. Int. J. Forecast., 2006, 22(4):679–688, https://doi.org/10.1016/j.ijforecast.2006.03.001

REFERENCES

1. Hastie, T., Tibshirani, R., & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edition. Springer, – 2009 – 533p, https://doi.org/10.1007/978-0-387-84858-7.
2. Zou H., Hastie T. *Regularization and variable selection via the elastic net*. J. R. Stat. Soc. Series B, 2005, 67(2): 301–320, https://doi.org/10.1111/j.1467-9868.2005.00503.x.
3. Wang Y., Hao X., Wu C. *Forecasting stock returns: A time-dependent weighted least squares approach*. Journal of Financial Markets, 2021, 53: 100568, https://doi.org/10.1016/j.finmar.2020.100568.
4. Dikheel T.R., Yaseen A.Q. *Robust lag weighted lasso for time series model*. J. Modern Applied Statistical Methods, 2021, 19(1): 14, https://doi.org/10.56801/10.56801/v19.i.1081.
5. Lütkepohl, H., Xu, F. *The role of the log transformation in forecasting economic variables*. Empir Econ 42, 2012, 619–638, https://doi.org/10.1007/s00181-010-0440-1.
6. Tsay R.S. Analysis of Financial Time Series. 2nd ed. Wiley, 2010, https://doi.org/10.1002/9780470644560.
7. E. S. Gardner Jr., *Exponential Smoothing: The State of the – Part II.* International Journal of Forecasting, 22(4), 637–666, https://doi.org/10.1016/j.ijforecast.2006.03.005
8. R.J. Hyndman, A.B. Koehler. *Another look at measures of forecast accuracy*. Int. J. Forecast., 2006, 22(4):679–688, https://doi.org/10.1016/j.ijforecast.2006.03.001

Надійшла до редакції 25.03.25 р.

КВЄТНИЙ РОМАН НАУМОВИЧ – д.т.н., професор кафедри Автоматизації та інтелектуальних інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, Україна ***e-mail: rkvetny@vntu.edu.ua;***

БОРОДКІН СЕРГІЙ ІВАНОВИЧ – аспірант кафедри Автоматизації та інтелектуальних інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, Україна, ***e-mail: borserg90@gmail.com***.

R.N. KVYETNYY, S.I. BORODKIN

IMPROVED MODEL OF ELASTIC NET REGULARIZATION FOR FINANCIAL TIME SERIES

Vinnytsia National Technical University