

УДК 004.412:519.237.5

Ю.О. ГУНЧЕНКО, В.І. МЕЩЕРЯКОВ, А.Л. РАЧИНСЬКА
 Одеський національний університет імені І.І. Мечникова
 С.Б. ПРИХОДЬКО
 Національний університет кораблебудування імені адмірала Макарова

РЕГРЕСІЙНІ МОДЕЛІ ДЛЯ РАНЬОГО ОЦІНЮВАННЯ КІЛЬКОСТІ РЯДКІВ КОДУ РІЗНИХ ВИПУСКІВ ЗАСТОСУНКІВ ІЗ ВІДКРИТИМ КОДОМ

Проблема раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом є важливою, оскільки це безпосередньо впливає на прогнозування зусиль з їх розроблення та подальшої модифікації. Об'єктом дослідження є процес раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом. Предметом дослідження є регресійні моделі для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

Метою роботи є побудова однофакторних нелінійних регресійних моделей для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

У роботі побудовано дві нелінійні регресійні моделі для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом залежно від кількості класів. Перша модель дає змогу виконувати оцінювання кількості рядків коду першого випуску програмних застосунків із відкритим кодом, а друга модель – останнього їх випуску. Зазначені моделі побудовано за двома наборами даних із двох метрик із 40 програмних застосунків із відкритим кодом: кількість строк коду та кількість класів для першого та останнього випусків відповідних застосунків. Згідно з критерієм Мардіа, розподіл цих двовимірних даних відхилявся від нормального, тому для перевірки наявності викидів у цих наборах даних було застосовано відповідний метод на основі квадрату відстані Махаланобіса для нормалізованих даних. Для нормалізації даних було використано нормалізуюче перетворення у вигляді десятичного логарифму. Згідно з критерієм Мардіа, розподіл нормалізованих даних не відхилявся від нормального. Оцінки параметрів отриманих моделей знайдено за методом найменших квадратів. Якість побудованих моделей перевірено за трьома відомими показниками: коефіцієнтом детермінації R^2 , середньою величиною відносної помилки MMRE та відсотком прогнозованих результатів PRED, для яких величини відносної помилки менші за 0,25, PRED (0,25). Отримані значення вказаних показників якості свідчать про задовільну якість двох побудованих нелінійних регресійних моделей.

Ключові слова: регресійна модель, оцінювання, рядки коду, застосунок із відкритим кодом, клас, нормалізуюче перетворення, критерій Мардіа, викид, відстань Махаланобіса, тестова статистика.

Y.O. GUNCHENKO, V.I. MESCHERYAKOV, A.L. RACHINSKA
 Odesa I.I. Mechnikov National University
 S.B. PRYKHODKO
 Admiral Makarov National University of Shipbuilding

REGRESSION MODELS FOR EARLY ESTIMATING THE NUMBER OF LINES OF CODE OF VARIOUS RELEASES OF OPEN-SOURCE APPLICATIONS

The problem of early estimating the number of lines of code of various releases of open-source applications is important because it directly affects the forecasting of efforts for their development and subsequent modification. The object of the study is the process of early estimating the number of lines of code of various releases of open-source applications. The subject of the study is the regression models for early estimating the number of lines of code of various releases of open-source applications.

The goal of the work is to build some regression models with one factor for early estimating the number of lines of code of various releases of open-source applications.

In the work, we built two nonlinear regression models for early estimating the number of lines of code of various releases of open-source applications depending on the number of classes. The first model allows for estimating the number of lines of code of the first release of open-source applications, and the second model allows for estimating the number of lines of code of the last release of open-source applications. These models were built on two datasets of two metrics from 40 open-source applications: the number of lines of code and the number of classes for the first and last releases of the respective applications. According to the Mardia criterion, the distribution of these two-dimensional data deviated from the Gaussian. Therefore, to check for outliers in these data sets, an appropriate method based on the squared Mahalanobis distance for normalized data was applied. A normalizing transformation in the form of a decimal logarithm was used to normalize the data. According to the Mardia criterion, the distribution of the normalized data did not deviate

from normal. The parameter estimates of the obtained models were found by the method of least squares. The quality of the constructed models was checked by three known indicators: the coefficient of determination R^2 , the mean magnitude of relative error MMRE, and the percentage of prediction PRED, for which the relative error values are less than 0.25, PRED(0.25). The obtained values of the specified quality indicators indicate a satisfactory quality of the two constructed nonlinear regression models.

Key words: regression model, estimation, lines of code, open-source application, class, normalizing transformation, Mardia criterion, outlier, Mahalanobis distance, test statistic.

Постановка проблеми

Сьогодні розробленню та вдосконаленню програмних застосунків із відкритим кодом приділяють усе більше уваги [1; 2], тому завдання раннього оцінювання кількості рядків коду різних випусків (релізів) програмних застосунків із відкритим кодом є важливим, оскільки це безпосередньо впливає на прогнозування зусиль з їх розроблення та подальшої модифікації.

Хоча для раннього оцінювання кількості строк коду програмних застосунків уже було запропоновано велику кількість різноманітних математичних моделей, у тому числі і регресійних, але вони при цьому не дають змоги враховувати випуск застосунку: перший або останній. А як свідчать дослідження [3; 4], різні програмні метрики, включаючи й кількість рядків коду (LOC), залежать і від випуску застосунку. Що потребує побудови математичних моделей для оцінювання кількості строк коду різних випусків програмних застосунків із відкритим кодом.

Аналіз останніх досліджень та публікацій

Незважаючи на існування достатньо великої кількості методів та моделей для раннього оцінювання кількості строк коду програмних застосунків із відкритим кодом, підвищення точності оцінювання залишається актуальним завданням [5; 6]. При цьому для відповідного оцінювання використовують різноманітні підходи, які, як правило, базуються на метриках, що можуть бути визначені до кодування із застосуванням або концептуальної моделі даних [7] або діаграм варіантів використання [8; 9], або класів [9; 10].

Незважаючи на широке застосування сьогодні машинного навчання [11; 12], методи як лінійного [13; 14], так і нелінійного регресійного аналізу [8; 10] продовжують використовуватися для раннього оцінювання кількості строк коду програмних застосунків, у тому числі і з відкритим кодом. Але існуючі моделі не враховують вплив випуску (релізу) застосунку на результати раннього оцінювання його розміру, хоча на такий вплив укажуть дані, які наведені в роботах [3; 4]. Тому виникає потреба у створенні математичних моделей, у тому числі регресійних, для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

Ураховуючи на те, що основним фактором у більшості існуючих моделей для раннього оцінювання розміру програмних застосунків є кількість класів, то і побудову відповідних моделей у цій роботі ми здійснюємо залежно від зазначеної метрики. Окрім того, щоб підтвердити вплив випуску застосунку на результати раннього оцінювання його розміру, ми створюємо дві регресійні моделі для першого та останнього випусків програмних застосунків із відкритим кодом.

Мета дослідження

Метою дослідження є побудова двох нелінійних регресійних моделей для раннього оцінювання кількості рядків коду першого та останнього випусків програмних застосунків із відкритим кодом.

Об'єктом дослідження є процес раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом. Предметом дослідження є регресійні моделі для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

Для досягнення поставленої мети необхідно вирішити такі завдання:

1. Побудувати дві регресійні моделі для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

2. Перевірити якість побудованих регресійних моделей для раннього оцінювання кількості рядків коду різних випусків програмних застосунків із відкритим кодом.

Виклад основного матеріалу дослідження

Спочатку було здійснено вибір двох наборів даних із двох метрик із 40 програмних застосунків із відкритим кодом: кількість строк коду Y та кількість класів X_1 для першого та останнього випусків відповідних застосунків ми взяли дані з роботи [4]. Зазначені двовимірні дані були перевірені на наявність викидів.

Згідно з критерієм Мардіа, розподіли цих двовимірних даних для двох наборів даних відхиляються від нормального: значення двовимірних асиметрії та ексцесу перевищують значення відповідних тестових статистик. Тому, як і в роботі [10], для перевірки наявності викидів у цих наборах негаусівських даних було застосовано відповідний метод на основі квадрату відстані Махаланобіса для нормалізованих даних. Для нормалізації даних по кожній змінній було використано перетворення у вигляді десяткового логарифму

$$Z_Y = \lg Y ; Z_1 = \lg X_1 . \quad (1)$$

Хоча перетворення (1) є одновимірними, їх застосування до вказаних двовимірних даних призводить до прийнятних результатів. Згідно з критерієм Мардіа, розподіл нормалізованих даних не відхиляється від нормального: значення двовимірних асиметрії та ексцесу менше значень відповідних тестових статистик. За методом визначення викидів у багатовимірних негаусівських даних на основі квадрату відстані Махаланобіса для нормалізованих даних у зазначених двох наборах даних викиди відсутні, тому що для кожної точки нормалізованих даних квадрати відстані Махаланобісу менше за квантиль розподілу χ^2 -квадрат, що дорівнює 10,60, для 2 ступенів свободи та 0,005 рівня значимості.

Далі для вказаних двох наборів даних із [4] спочатку було побудовано дві лінійні регресійні моделі. Для цих лінійних регресійних моделей з'ясувалося, що розподіл відхилення значень залежної випадкової величини Y від лінії регресії не є нормальним. Це свідчить про відсутність теоретичного обґрунтування можливості використання лінійної регресійної моделі для вказаних наборів даних, що призводить до необхідності побудови нелінійної регресійної моделі. Для її створення, як і в роботі [10], ми використали метод, який базується на нормалізуючих перетвореннях та визначенні наявності викидів у регресійні моделі. Для побудови двох нелінійних регресійних моделей ми застосовували перетворення (1). Тоді нелінійні регресійні моделі можуть бути представлені у формі

$$Y = 10^{\hat{\beta}_0 + \varepsilon} X_1^{\hat{\beta}_1} , \quad (2)$$

де $\hat{\beta}_0$ та $\hat{\beta}_1$ – оцінки параметрів лінійної регресійної моделі $Z_Y = \hat{\beta}_0 + \hat{\beta}_1 Z_1 + \varepsilon$, оцінки $\hat{\beta}_0$ та $\hat{\beta}_1$ визначалися за методом найменших квадратів. Причому змінна ε , яка описує відхилення залежної випадкової величини Z_Y від лінії регресії, повинна бути випадковою величиною з розподілом Гаусу, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Нульова гіпотеза про нормальність розподілу ε для двох моделей перевірялася за критерієм Пірсона. Ця гіпотеза не може бути відхилена для рівня значущості 0,05. Оцінки параметрів двох моделей виду (2), які побудовані за вказаними двома наборами даних, наведено в табл. 1. Тут модель 1 – це нелінійна регресійна модель для раннього оцінювання кількості рядків коду першого випуску програмних застосунків із відкритим кодом, а модель 2 – це нелінійна регресійна модель для раннього оцінювання кількості рядків коду останнього випуску програмних застосунків із відкритим кодом. У табл. 1 також надано значення оцінок середньо квадратичного відхилення $\hat{\sigma}_\varepsilon$ випадкової величини ε .

Таблиця 1

Оцінки параметрів та показники якості моделей

Модель	\hat{b}_0	\hat{b}_1	$\hat{\sigma}_\varepsilon$	R^2	$MMRE$	$PRED(0,25)$
Модель 1	1,57907	1,05496	0,11133	0,9706	0,2043	0,6452
Модель 2	1,72886	1,00937	0,08556	0,9510	0,1623	0,8333

У процесі побудови зазначених нелінійних регресійних моделей, як і в роботі [10], виявлено викиди, які зумовлено виходом певних точок даних за інтервал прогнозування відповідної нелінійної регресії. Для моделі 1 було виявлено 9 викидів, а для моделі 2 – 10 викидів.

Якість двох побудованих нелінійних регресійних моделей виду (2) було перевірено за коефіцієнтом детермінації R^2 , середньою величиною відносної помилки $MMRE$ та відсотком прогнозованих результатів $PRED$, для яких величини відносної помилки менші за 0,25, $PRED(0,25)$. Ці показники, як правило, застосовуються для перевірки якості прогнозування за допомогою регресійних моделей. У табл. 1 наведено значення цих трьох показників якості двох побудованих нелінійних регресійних моделей виду (2).

Значення показників якості з табл. 1 свідчать про задовільну якість двох побудованих нелінійних регресійних моделей виду (2). Великі значення R^2 указують на те, що більшість значень Y знаходиться біля лінії регресії. Зазвичай якщо $MMRE \leq 0,25$ та $PRED(0,25) \geq 0,75$, то це вважається прийнятною точністю прогнозування за допомогою регресійних моделей. А як ми бачимо з табл. 1, лише для моделі 1 значення одного показника якості $PRED(0,25)$ менше за мінімально прийнятне всього на 14%. Значення всіх інших показників якості вказують на прийнятну точність оцінювання кількості строк коду за побудованими нелінійними регресійними моделями виду (2).

Дані табл. 1 указують на кращу якість моделі 2 порівняно з моделлю 1 передусім за показником $PRED(0,25)$. Це, на нашу думку, можна пояснити більш сталими даними для останніх випусків удосконалених застосунків із відкритим кодом.

До переваг двох побудованих нелінійних регресійних моделей виду (2) слід віднести те, що вони дають змогу здійснювати раннє оцінювання кількості рядків коду першого та останнього випусків програмних застосунків із відкритим кодом на відміну від існуючих подібних моделей, які не враховують відповідний випуск застосунку. До недоліків двох побудованих нелінійних регресійних моделей виду (2) можна віднести те, що, по-перше, вони є однофакторними, а по-друге, для їх побудови використовувалися два набори даних лише по 40 точок кожний. Хоча кількість класів, як правило, є основним фактором, що має найбільший вплив на кількість строк коду, але врахування додаткових факторів (наприклад, кількості методів) може підвищити точність оцінювання кількості строк коду. До обмежень двох побудованих нелінійних регресійних моделей виду (2) слід віднести такі. По-перше, ці моделі побудовані для застосунків із відкритим кодом. По-друге, для моделі 1 кількість класів знаходиться у межах від 5 до 3697, а для моделі 2 – у межах від 8 до 4397.

У подальшому, на нашу думку, ці недоліки та обмеження можуть бути виправлені за рахунок побудови нових моделей на основі більшої кількості даних для різних застосунків, у тому числі не лише з відкритим кодом, а й промислових.

Висновки

Удосконалено дві однофакторні нелінійні регресійні моделі для раннього оцінювання кількості рядків коду програмних застосунків із відкритим кодом залежно від кількості класів на основі нормалізації за десятковим логарифмом двох наборів даних для першого та останнього випусків відповідних застосунків. Ці моделі порівняно з існуючими регресійними моделями дають змогу здійснювати раннє оцінювання кількості рядків коду першого та останнього випусків програмних застосунків із відкритим кодом.

Під час визначення наявності двовимірних викидів у даних треба враховувати, чи не відхиляється багатовимірний розподіл даних від нормального. У разі відхилення двовимірного розподілу даних від нормального для визначення наявності викидів у цих даних можна застосувати метод на основі квадрату відстані Махаланобіса для нормалізованих даних на основі багатовимірного нормалізуючого перетворення. Під час побудови нелінійної регресійної моделі слід враховувати наявність відповідних викидів, які визначаються шляхом виходу точки даних за інтервал прогнозування нелінійної регресії.

У подальшому планується побудова нелінійних регресійних моделей для раннього оцінювання кількості рядків коду різних випусків програмних застосунків за більшою кількістю даних для різноманітних застосунків, у тому числі не тільки з відкритим кодом, а й промислових.

Список використаної літератури

1. Himansh M., Manikandan V.M. A statistical study and analysis to identify the importance of open-source software. *Information Technology (ICITIT) : proceedings of 2022 International Conference, Kottayam, 13–14 February 2022*. Kottayam, India: IEEE, 2022. P. 1–6. <https://doi.org/10.1109/ICITIT54346.2022.9744176>
2. Haider S., Khalil W., Al-Shamayleh A.S., Akhuzada A., Gani A. Risk factors and practices for the development of open source software from developers' perspective. *IEEE Access*. 2023. Vol. 11. P. 63333–63350. <https://doi.org/10.1109/ACCESS.2023.3267048>
3. Molnar A.J., Neamtu A., Motogna S. Evaluation of software product quality metrics. *Communications in Computer and Information Science*. 2020. Vol. 1172. P. 163–187. https://doi.org/10.1007/978-3-030-40223-5_8
4. Gradišnik M., Beranič T., Karakatič S. Impact of historical software metric changes in predicting future maintainability trends in open-source software development. *Applied Sciences*. 2020. № 10 (13). 4624. <https://doi.org/10.3390/app10134624>
5. Daud M., Malik A.A. Improving the accuracy of early software size estimation using analysis-to-design adjustment factors (ADAFs). *IEEE Access*. 2021. Vol. 9. P. 81986–81999. <https://doi.org/10.1109/ACCESS.2021.3085752>
6. Dewi R.S., Araynawa T.K., Prasanna F.M., Felianasari N., Rahmawati R., Hartantc A.E., ... Mazaya Al-K. Improving software size estimation using data complexity (Case study: Research and community service monitoring apps). *Electrical engineering, computer science and informatics (EECSI) : proceedings of 2024 11th International conference, Yogyakarta, 26–27 September 2024*. Yogyakarta, Indonesia: IEEE, 2024. P. 315–319. <https://doi.org/10.1109/EECSI63442.2024.10776530>
7. Dewi R.S., Zahrah F.A., Nugraha D.A., Prabowo P.S., Safitri A., Jayadi P. Predicting software size based on conceptual data model (Case study: Shrimp pond system management). *Electrical engineering and computer science (ICECOS) : proceedings of 2024 International conference, Palembang, 25–26 September 2024*. Palembang, Indonesia: IEEE, 2024. P. 175–178. <https://doi.org/10.1109/ICECOS63900.2024.10791154>
8. Nassif A.B., AbuTalib M., Capretz L.F. Software effort estimation from Use Case diagrams using nonlinear regression analysis. *Electrical and computer engineering : proceedings of IEEE Canadian conference, 30 August – 02 September 2020*. London, ON, Canada: IEEE, 2020. P. 1–4. <https://doi.org/10.1109/CCECE47787.2020.9255712>
9. Hussain I., Malik A.A. Determining the utility of use case points and class points in early software size estimation. *Emerging Technologies (ICET) : proceedings of 2023 18th International Conference, Peshawar, 06–07 November 2023*. Peshawar, Pakistan: IEEE, 2023. P. 171–175. <https://doi.org/10.1109/ICET59753.2023.10374977>
10. Prykhodko S.B., Shutko I.S., Prykhodko A.S. Early size estimation of web apps created using Codeigniter framework by nonlinear regression models. *Radio-electronic and computer systems*. 2022. Vol. 103, No 3. P. 84–94. <https://doi.org/10.32620/reks.2022.3.06>

11. Manisha, Rishi R. Early size estimation using machine learning. *Computing for sustainable global development (INDIACom)* : proceedings of the 2021 8th International conference, New Delhi, 17-19 March 2021. New Delhi, India, Los Alamitos: IEEE, 2021. P. 757–762. <https://doi.org/10.1109/INDIACom51348.2021.00135>
12. Molla Y.S., Alemneh E., Yimer S.T. COSMIC-based early software size estimation using deep learning and domain-specific BERT. *IEEE Access*. 2025. Vol. 13. P. 28463–28475. <https://doi.org/10.1109/ACCESS.2025.3540548>
13. Nhung H.L.T.K., Hai V.V., Silhavy R., Prokopova Z., Silhavy P. Parametric software effort estimation based on optimizing correction factors and multiple linear regression. *IEEE Access*. 2022. Vol. 10. P. 2963–2986. <https://doi.org/10.1109/ACCESS.2021.3139183>
14. Приходько С., Шутко І. Регресійні моделі для раннього оцінювання кількості рядків коду вебзастосунків, що створюються за допомогою фреймворку Codeigniter. *Прикладні питання математичного моделювання*. 2025. Т. 8, № 1. С. 189–196. <https://doi.org/10.32782/mathematical-modelling/2025-8-1-18>

References

1. Himansh, M., & Manikandan, V.M. (2022). A statistical study and analysis to identify the importance of open-source software. In *2022 International Conference on Innovative Trends in Information Technology (ICITIIT)* (pp. 1–6). Kottayam, India: IEEE. <https://doi.org/10.1109/ICITIIT54346.2022.9744176> [in English].
2. Haider, S., Khalil, W., Al-Shamayleh, A.S., Akhunzada, A., & Gani, A. (2023). Risk factors and practices for the development of open source software from developers' perspective. *IEEE Access*, 11, 63333–63350. <https://doi.org/10.1109/ACCESS.2023.3267048> [in English].
3. Molnar, A.J., Neamțu, A., & Motogna, S. (2020). Evaluation of software product quality metrics. *Communications in Computer and Information Science*, 1172, 163–187. Springer, Cham. https://doi.org/10.1007/978-3-030-40223-5_8 [in English].
4. Gradišnik, M., Beranič, T., & Karakatič, S. (2020). Impact of historical software metric changes in predicting future maintainability trends in open-source software development. *Applied Sciences*, 10 (13), 4624. <https://doi.org/10.3390/app10134624> [in English].
5. Daud, M., & Malik, A.A. (2021). Improving the accuracy of early software size estimation using analysis-to-design adjustment factors (ADAFs). *IEEE Access*, 9, 81986–81999. <https://doi.org/10.1109/ACCESS.2021.3085752> [in English].
6. Dewi, R.S., Araynawa, T.K., Prasanna, F.M., Felianasari, N., Rahmawati, R., Hartantc, A.E., Kamilah, L.S., & Mazaya, Al-K. (2024). Improving software size estimation using data complexity (Case study: Research and community service monitoring apps). Proceedings of *2024 11th International conference on electrical engineering, computer science and informatics (EECSI)* (pp. 315–319). Yogyakarta, Indonesia: IEEE. <https://doi.org/10.1109/EECSI63442.2024.10776530> [in English].
7. Dewi, R.S., Zahrah, F.A., Nugraha, D.A., Prabowo, P.S., Safitri, A., & Jayadi, P. (2024). Predicting software size based on conceptual data model (Case study: Shrimp pond system management). Proceedings of *2024 International conference on electrical engineering and computer science (ICECOS)* (pp. 175–178). Palembang, Indonesia: IEEE. <https://doi.org/10.1109/ICECOS63900.2024.10791154> [in English].
8. Nassif, A.B., AbuTalib, M., & Capretz, L.F. (2020). Software effort estimation from Use Case diagrams using nonlinear regression analysis. Proceedings of *IEEE Canadian conference on electrical and computer engineering* (pp. 1–4). London, ON, Canada: IEEE. <https://doi.org/10.1109/CCECE47787.2020.9255712> [in English].
9. Hussain, I., & Malik, A.A. (2023). Determining the utility of use case points and class points in early software size estimation. Proceedings of *2023 18th International Conference on Emerging*

- Technologies (ICET)* (pp. 171–175). Peshawar, Pakistan: IEEE. <https://doi.org/10.1109/ICET59753.2023.10374977> [in English].
10. Prykhodko, S.B., Shutko, I.S., & Prykhodko, A.S. (2022). Early size estimation of web apps created using Codeigniter framework by nonlinear regression models. *Radio-electronic and computer systems*, 3 (103), 84–94. <https://doi.org/10.32620/reks.2022.3.06> [in English].
 11. Manisha, & Rishi, R. (2021). Early size estimation using machine learning. *Proceedings of the 2021 8th International conference on computing for sustainable global development (INDIACom)* (pp. 757–762). New Delhi, India, Los Alamitos: IEEE. <https://doi.org/10.1109/INDIACom51348.2021.00135> [in English].
 12. Molla, Y.S., Alemneh, E., & Yimer, S.T. (2025). COSMIC-based early software size estimation using deep learning and domain-specific BERT. *IEEE Access*, 13, 28463–28475. <https://doi.org/10.1109/ACCESS.2025.3540548> [in English].
 13. Nhung, H.L.T.K., Hai, V.V., Silhavy, R., Prokopova, Z., & Silhavy, P. (2022). Parametric software effort estimation based on optimizing correction factors and multiple linear regression. *IEEE Access*, 10, 2963–2986. <https://doi.org/10.1109/ACCESS.2021.3139183> [in English].
 14. Prykhodko, S., & Shutko, I. (2025). Rehresiini modeli dlia rannoho otsiniuvannia kilkosty riadkiv kodu vebzastosunkiv, shcho stvoriuiutsia za dopomohoiu freimvorku Codeigniter [Regression models for early estimating the lines of code count of web applications created using the Codeigniter framework]. *Applied questions of mathematical modelling*, 8 (1), 189–196. <https://doi.org/10.32782/mathematical-modelling/2025-8-1-18> [in Ukrainian].

Гунченко Юрій Олександрович – д.т.н., професор, завідувач кафедри комп'ютерних систем та технологій Одеського національного університету імені І.І. Мечникова. E-mail: gunchenko@onu.edu.ua, ORCID: 0000-0003-4423-8267.

Мещеряков Володимир Іванович – д.т.н., професор, професор кафедри інформаційних технологій Одеського національного університету імені І.І. Мечникова. E-mail: meshcheryakovvi48@gmail.com, ORCID: 0000-0003-0499-827X.

Рачинська Алла Леонідівна – к.ф.-м.н., доцент, завідувач кафедри механіки, автоматизації та інформаційних технологій Одеського національного університету імені І.І. Мечникова. E-mail: rachinskaya@onu.edu.ua, ORCID: 0000-0003-2430-9603.

Приходько Сергій Борисович – д.т.н., професор, завідувач кафедри програмного забезпечення автоматизованих систем Національного університету кораблебудування імені адмірала Макарова. E-mail: sergiy.prykhodko@nuos.edu.ua, ORCID: 0000-0002-2325-018X.

Gunchenko Yurii Oleksandrovych – Doctor of Technical Sciences, Professor, Head of the Department of Computer Systems and Technologies of the Odesa I.I. Mechnikov National University. E-mail: gunchenko@onu.edu.ua, ORCID: 0000-0003-4423-8267.

Meshcheryakov Volodymyr Ivanovych – Doctor of Technical Sciences, Professor, Professor at the Department of Information Technologies of the Odesa I.I. Mechnikov National University. E-mail: gradan@ua.fm, ORCID: 0000-0003-0499-827X.

Rachinska Alla Leonidivna – Candidate of Physical and Mathematical Sciences, Associate Professor, Head of the Department of Mechanics, Automation and Information Technologies of the Odesa I.I. Mechnikov National University. E-mail: rachinskaya@onu.edu.ua, ORCID: 0000-0003-2430-9603.

Prykhodko Sergiy Borisovich – Doctor of Technical Sciences, Professor, Head of the Department of Software for Automated Systems of the Admiral Makarov National University of Shipbuilding. E-mail: sergiy.prykhodko@nuos.edu.ua, ORCID: 0000-0002-2325-018X.

Отримано: 28.10.2025

Рекомендовано: 27.11.2025

Опубліковано: 30.12.2025

