УДК 004.412:519.237.5

НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ МЕТРИКИ RFC ЗАСТОСУНКІВ З ВІДКРИТИМ КОДОМ НА KOTLIN

Приходько С.Б., д.т.н., проф.[1]; Кольцов А.В.[2]; Грабовський Є.В.[3]
[1,2,3] Національний університет кораблебудування імені адмірала Макарова
[1,2,3] Україна, Миколаїв
[1]sergiy.prykhodko@nuos.edu.ua; [2]211258@nuos.edu.ua; [3]221360@nuos.edu.ua

Анотація. *В роботі на основі перетворення у вигляді десяткового логарифму побудована нелінійна регресійна модель для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin в залежності від двох факторів: метрик CBO та WMC.*
Ключові слова: *оцінювання, метрика RFC; застосунок з відкритим кодом; Kotlin; нелінійна регресійна модель.*

Вступна частина. Для визначення якості програмного забезпечення (ПЗ) використовують різні метрики ПЗ, включаючи RFC (response for a class), CBO (coupling between object classes) та WMC (weighted methods per class) [1]. Також відомо [1], що метрика RFC залежить від метрик CBO та WMC. В [1] автори побудували рівняння лінійної регресії для оцінювання середнього значення RFC на рівні програми. Водночас не завжди є теоретичне обґрунтування використання моделей лінійної регресії, наприклад, коли розподіл залишків не є гаусівським. Тому в [2] було побудовано нелінійну регресійну модель для оцінювання метрики RFC застосунків з відкритим кодом на Java в залежності від метрик CBO та WMC. Але зазначена модель, як це буде показано далі, не завжди дає достовірні результати у разі оцінювання метрики RFC застосунків з відкритим кодом на Kotlin. А зараз багато програмних застосунків створюються із використанням мови Kotlin.

Тому ***метою дослідження*** є побудова нелінійної регресійної моделі для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin в залежності від метрик CBO та WMC.

Основна частина. Розглянемо побудову нелінійної регресійної моделі для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin в залежності від метрик CBO та WMC на рівні застосунку, які позначимо відповідно як X_1 та X_2. Зазначені три метрики з 52 застосунків з відкритим кодом на Kotlin, які розміщені на сайті GitHub, були отримані за допомогою інструменту CodeMR (https://plugins.jetbrains.com/plugin/10811-codemr). Далі ми перевірили тривимірні дані на нормальність за допомогою тесту Мардіа (Mardia's test), якій базується на багатовимірних асиметрії та ексцесі. Тест показав, що розподіл даних не є гаусівським. Тому для визначення наявності викидів у даних ми використовували тест на основі квадрату відстані Махаланобіса для нормалізованих за десятковим логарифмом даних. Для обґрунтування необхідності побудови нелінійної регресійної моделі для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin ми перевірили якість відповідної моделі з [2], яка була розроблена для застосунків з відкритим кодом на Java. Були отримані значення наступних показників якості моделі з [2] у разі її застосування для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin: множинний коефіцієнт детермінації R^2, середня величина відносної помилки MMRE та відсоток прогнозованих результатів PRED(0,25) дорівнюють відповідно 0,7926, 0,5022 та 0,0435. Значення MMRE та PRED(0,25) вказують на погану якість моделі з [2] у разі її застосування для оцінювання метрики RFC застосунків на Kotlin. А це призводить до необхідності побудови регресійної

моделі для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin. Використовуючи зібрані дані із зазначених вище метрик застосунків з відкритим кодом на Kotlin спочатку була побудована лінійна регресійна модель $Y = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + \varepsilon$. Змінна ε, яка описує залишки, повинна бути випадковою величиною з розподілом Гауса. Нульова гіпотеза про нормальність розподілу ε була перевірена за критерієм Пірсона. Ця гіпотеза була відкинута для рівня значимості 0,05. Те, що розподіл ε не є гаусівським вказує на відсутність теоретичного обґрунтування застосування лінійної регресійної моделі та необхідність побудови саме нелінійної регресійної моделі, яка була побудована на основі десяткового логарифму має вигляд $Y = 10^{\varepsilon+\hat{b}_0} X_1^{\hat{b}_1} X_2^{\hat{b}_2}$, де ε – гаусівська випадкова величина з нульовим математичним сподіванням та середньоквадратичним відхиленням, оцінка якого дорівнює 0,1128, $\hat{b}_0$, $\hat{b}_1$ та $\hat{b}_2$ є оцінками параметрів, які мають відповідно такі значення: -0,306892, 0,228584 та 1,36206. Оцінки параметрів були отримані за допомогою методу найменших квадратів. Ця модель була побудована за методом [3] із врахуванням наявності викидів за даними з 48 застосунків, що залишилися після видалення викидів, і може застосовуватися для значень метрик, для яких квадрат відстані Махаланобісу їх десяткових логарифмів менше або дорівнює 12,84. Значення R^2, MMRE та PRED(0,25) для створеної моделі дорівнюють відповідно 0,9025, 0,1171 та 0,8750. Це свідчить про добру якість отриманої моделі. Також були визначені довірчі інтервали та інтервали прогнозування нелінійної регресії для RFC в залежності від метрик CBO та WMC на рівні застосунку, які можуть використовуватися для визначення якості застосунків на Kotlin за методикою [4].

Висновки. В роботі на основі десяткового логарифму була побудована нелінійна регресійна модель для оцінювання метрики RFC застосунків з відкритим кодом на Kotlin в залежності від метрик CBO та WMC, яка має задовільну якість відповідного оцінювання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Laing, V., & Coleman, C. (2001). *Principal Components of Orthogonal Object-Oriented Metrics*. *White Paper Analyzing Results of NASA Object-Oriented Data* (Task 323-08-14). Greenbelt, Maryland (US): NASA Goddard Space Flight Center, the Software Assurance Technology Center (SATC).

[2] Приходько, С.Б., & Позняков, Р.О. (2022). Нелінійна регресійна модель для оцінювання метрики RFC застосунків з відкритим кодом на Java. *Інформаційні технології: моделі, алгоритми, системи : ІІІ-я Всеукраїнська науково-практична інтернет конференція (ITMAS–2022)* (С. 8-9). Миколаїв, Україна. http://itconf.nuos.edu.ua/2022/wp-content/uploads/sites/5/NUOS-ITMAS-2022-Proceedings.pdf

[3] Prykhodko, S., & Prykhodko, N. (2022). A modified technique for constructing nonlinear regression models based on the multivariate normalizing transformations. *Selected Papers of the VIII International Scientific Conference "Information Technology and Implementation" (IT&I-2021). Workshop Proceedings (IT&I-WS-2021), Kyiv, Ukraine, December 1-3, 2021. CEUR Workshop Proceedings, 3179*, 156-166. http://ceur-ws.org/Vol-3179/Paper_15.pdf

[4] Prykhodko, S., & Prykhodko, N. (2022). A Technique for Detecting Software Quality Based on the Confidence and Prediction Intervals of Nonlinear Regression for RFC Metric. *Computer Sciences and Information Technologies : 2022 IEEE 17th International Conference (CSIT-2022),* (pp. 499-502). Lviv, Ukraine. DOI: https://doi.org/10.1109/CSIT56902.2022.10000532

Prykhodko S.B, Koltsov A.V., Hrabovsky Y.V.

A nonlinear regression model for estimating the RFC metric of open-source applications in Kotlin

Abstract. *In work, a nonlinear regression model with two predictors for estimating the RFC metric of open-source applications in Kotlin depending on CBO and WMC metrics has been built.*

Keywords: *estimation; RFC metric; open-source app; Kotlin; nonlinear regression model.*