УДК 004.412:519.237.5

А.В. КОЛЬЦОВ
Національний університет кораблебудування імені адмірала Макарова

МАТЕМАТИЧНІ МОДЕЛІ ДЛЯ ВИЗНАЧЕННЯ ВИКИДІВ У ДВОВИМІРНИХ ДАНИХ ІЗ МЕТРИК CBO ТА RFC ПРОГРАМНИХ ЗАСТОСУНКІВ З ВІДКРИТИМ КОДОМ НА KOTLIN

У наш час розроблення програмного забезпечення потребує засобів для оцінювання якості та складності об'єктно орієнтованого проєктування. Таке оцінювання проводять за допомогою метрик програмного забезпечення. Зазвичай таке оцінювання потребує обробки даних із метрик, зокрема й виявлення викидів. У реальних даних розподіл метрик здебільшого відхиляється від нормального. Ця проблема ставить питання про розроблення математичних моделей, які будуть ураховувати таку поведінку даних із метрик програмного забезпечення.

У дослідженні запропоновано математичні моделі для більш достовірного визначення викидів у двовимірних даних із метрик CBO та RFC спеціально для Kotlin-проєктів. Аналіз 102 open-source проєктів виявив значні відмінності порівняно із Java: середнє значення CBO для Kotlin становило лише 2,20 (проти 11,88 у Java), а RFC – 7,46 (проти 18,36). Тест Мардія підтвердив, що розподіл цих метрик для Kotlin суттєво відхиляється від нормального ($\beta_1 = 70{,}867$, $\beta_2 = 14{,}066$ за критичних значень 9,3 та 9,48).

Для вирішення цієї проблеми розроблено модель, яка включає: застосування нормалізуючого перетворення Бокса – Кокса з параметрами та застосування двох взаємодоповнюючих підходів до виявлення викидів (на основі квантиля χ^2-розподілу, на основі квантиля F-розподілу Фішера).

Результати показують, що запропоновані моделі усувають головний недолік моделі Java, де до 25% Kotlin-проєктів помилково класифікувалися як викиди. Удосконалені моделі демонструють більшу достовірність за проведення аналізу даних із метрик CBO та RFC Kotlin-застосунків, водночас зберігають усі переваги підходу, заснованого на застосуванні еліпсоїда (еліпса) прогнозування. Особливо варто відзначити, що модель з χ^2-критерієм виявилась більш строгою, що дозволяє виявляти викиди з більшою достовірністю.

Це дослідження має велике практичне значення для розробників, які працюють із Kotlin, пропонує спеціалізований інструмент для аналізу даних із метрик CBO та RFC застосунків на Kotlin. Отримані результати підтверджують необхідність створення окремих аналітичних моделей для різних мов програмування, з урахуванням особливостей розподілу їх метрик.

Ключові слова: *метрики програмного забезпечення, CBO (Coupling Between Objects), RFC (Response For Class), викиди, Kotlin, нормалізуюче перетворення Бокса – Кокса, відстань Махаланобіса, еліпс прогнозування, об'єктно орієнтоване проєктування.*

A.V. KOLTSOV
Admiral Makarov National University of Shipbuilding

MATHEMATICAL MODELS FOR DETECTING OUTLIERS IN TWO-DIMENSIONAL CBO AND RFC METRICS' DATA OF OPEN-SOURCE KOTLIN-BASED APPLICATIONS

Modern software development requires robust methods for assessing the quality and complexity of object-oriented design. Such evaluation is typically performed using software metrics. However, this process usually requires metric data processing, including outlier detection. In real-world scenarios, metric distributions deviate from normality in the vast majority of cases. This reality necessitates the development of specialized mathematical models that account for such non-standard behavior of software metric data.

The study proposes mathematical models for more reliable detection of outliers in two-dimensional CBO and RFC metric data specifically for Kotlin projects. Analysis of 102 open-source projects revealed substantial differences compared to Java: the average CBO value for Kotlin was only 2,20 (versus 11,88 for Java), and RFC – 7,46 (versus 18,36). Mardia's test confirmed that the distribution of these metrics for Kotlin significantly deviates from normal ($\beta_1 = 70{,}867$, $\beta_2 = 14{,}066$ with critical values of 9,3 and 9,48).

To solve this problem, we developed a model that includes: application of Box-Cox normalization transformation with parameters and implementation of two complementary approaches for outlier detection (based on χ^2-distribution quantile and based on Fisher's F-distribution quantile).

The results demonstrate that our models overcome the key limitation of Java-oriented approaches, where up to 25% of Kotlin projects were misclassified as outliers. The refined models show greater reliability in analyzing Kotlin CBO and RFC metrics while preserving all advantages of the prediction ellipsoid approach. Notably, the χ^2-criterion model proved more stringent, enabling more accurate outlier detection.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

This research provides practical value for Kotlin developers by offering specialized tools for CBO and RFC metric analysis. Our findings confirm the necessity of developing language-specific analytical models that account for unique metric distribution characteristics.

Key words: *software metrics, CBO (Coupling Between Objects), RFC (Response for Class), outliers, Kotlin, Box-Cox transformation, Mahalanobis distance, prediction ellipse, object-oriented design.*

Постановка проблеми

У наш час програмне забезпечення (далі – ПЗ) із відкритим кодом набуває все більшої популярності. Водночас зростає потреба у створенні якісного ПЗ. Ця потреба пов'язана з можливістю виконання оцінювання складності та якості проєктування ПЗ. Для цього широко застосовують метрики ПЗ, зокрема один із найвідоміших наборів – CK (Chidamber and Kemerer metrics suite) [1]. Серед цих метрик CBO та RFC характеризують складність класу через його зв'язки й дозволяють виконати оцінювання складності об'єктно орієнтованого проєктування на третьому етапі за Бучем [2]. Відомо [3], що ці метрики корелюють одна з одною, тому їх спільний аналіз є доцільним. У процесі роботи з такими метриками часто постає потреба в обробленні даних, зокрема й у виявленні викидів. Розподіл цих метрик зазвичай відхиляється від нормального. Це зумовлює необхідність розроблення математичних моделей для попередньої обробки метрик CBO та RFC з урахування відхилення розподілу цих метрик від нормального.

Аналіз останніх досліджень та публікацій

У роботі [4] детально розглядають метод виявлення викидів на основі відстані Махаланобіса, який є одним із класичних підходів у багатовимірному статистичному аналізі. Автори обґрунтовують, що даний метод є особливо ефективним для виявлення аномальних спостережень тоді, коли дані мають багатовимірний нормальний розподіл.

Згідно з результатами, отриманими в [4], відстань Махаланобіса для p-вимірного вектора x визначається як квадратична форма:

$$D^2 = (x-\mu)'\Sigma^{-1}(x-\mu)\,, \tag{1}$$

де μ – вектором середніх значень, Σ – коваріаційна матриця, а Σ^{-1} – її обернена матриця.

Важливим аспектом методу, на який вказують автори, є те, що квадрат відстані Махаланобіса D^2 для багатовимірного нормального розподілу підпорядковується розподілу χ^2 з p ступенями вільності. Ця властивість дозволяє встановлювати статистично обґрунтовані порогові значення для ідентифікації викидів.

Основними перевагами методу є автоматичний облік кореляційної структури даних, інваріантність щодо лінійних перетворень змінних і можливість візуалізації як еліпсоїдів рівної щільності. Однак метод має і суттєві обмеження: чутливість до порушень припущення про нормальність розподілу, вразливість до ефекту маскувальних викидів та залежність від точності оцінки коваріаційної матриці.

У роботі [5] класичний підхід, викладений у праці [4], щодо використання відстані Махаланобіса для виявлення викидів, отримав подальший розвиток. На відміну від традиційного підходу, запропонований метод дає стійкі оцінки коваріаційної матриці (MCD) (менш чутливі до викидів); удосконалені методи візуалізації (Q-Q plots для відстаней Махаланобіса); висунуто конкретні практичні рекомендації щодо вибору порогових значень. Автор висуває пропозиції щодо подолання «маскувальних ефектів» і аналізу впливу розміру вибірки.

Дослідження [6] демонструє еволюцію еліпсоїдального підходу через його інтеграцію з методами машинного навчання. Автори поєднують відстань Махаланобіса з методом опорних векторів для однокласової класифікації (OCSVM). За використання лінійного ядра цей метод цілком зберігає еліпсоїдальну геометрію, тоді як радіальні базисні функції (RBF) неявно будують складні

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

еліпсоїдні структури в перетвореному просторі ознак. Практичне застосування для виявлення аномального трафіка IoT-пристроїв підтвердило ефективність такого комбінованого підходу.

У роботі [7] еліпсоїдні моделі адаптовані для завдань промислового моніторингу. Система Махаланобіса – Таґуті (MTS) складається із трьох ключових етапів: побудови еталонного еліпсоїда на «хороших» даних, оптимізації розмірності через ортогональні масиви Таґуті, виявлення аномалій через проєкцію на підпростір. Цей підхід демонструє підвищену стійкість до зашумлених даних порівняно із класичним методом, що особливо важливо для промислових застосувань.

Роботи [8–10] пропонуть вирішення однієї із ключових проблем еліпсоїдальних методів – залежності від припущення про нормальність, що рідко трапляється в реальних даних. Автори використовують нормалізуючі перетворення (Бокса – Кокса, Джонсона) для нормалізації даних перед застосуванням еліпсоїдного критерію.

Так, у [8–9] запропоновано застосовувати такий еліпсоїд прогнозування:

$$\left(\mathbf{Z}-\bar{\mathbf{Z}}\right)^T \mathbf{S}_N^{-1}\left(\mathbf{Z}-\bar{\mathbf{Z}}\right)=\chi^2_{m,\alpha}\,, \tag{2}$$

тут $\chi^2_{m,\alpha}$ – квантиль розподілу хі-квадрат із m ступенями вільності; α – довірча ймовірність (як довірчу ймовірность запропоновано 0,005).

А в роботі [10] еліпс прогнозування такий:

$$\left(\mathbf{Z}-\bar{\mathbf{Z}}\right)^T \mathbf{S}_N^{-1}\left(\mathbf{Z}-\bar{\mathbf{Z}}\right)=\frac{2\left(N^2-1\right)}{N\left(N-2\right)}F_{2,N-2,\alpha}\,, \tag{3}$$

тут $F_{2,N-2,\alpha}$ – квантиль F-розподілу із 2 та $N-2$ ступенями вільності; α – довірча ймовірність (як довірчу ймовірность запропоновано взяти 0,05).

Після зворотного перетворення межі аномалій набувають нелінійної форми, водночас зберігають основні переваги еліпсоїдального підходу.

Зазначений підхід застосовано в пошуку викидів під час побудови математичної моделі для оцінювання складності об’єктно орієнтованого проєктування на третьому етапі, за Бучем, за метриками CBO та RFC застосунків на Java [11]. Як буде показано згодом, ця модель не підходить для даних із метрик застосунків на Kotlin, що зумовлює потребу в удосконаленні існуючої моделі для мови Java для виявлення викидів для мови Kotlin, які будуть ураховувати відхилення розподілу даних із метрик CBO та RFC від нормального.

Мета дослідження

Метою дослідження є вдосконалення математичних моделей для визначення викидів у двовимірних даних із метрик CBO та RFC на основі еліпса прогнозування, які враховують відхилення від нормального розподілу даних із метрик мовою Kotlin на основі нормалізуючого перетворення Бокса – Кокса з урахуванням тестової статистики на основі квантиля критерію хі-квадрат і тестової статистики на основі квантиля *F*-розподілу Фішера.

Виклад основного матеріалу дослідження

Для побудови математичних моделей було отримано набір даних зі 102 проєктів з відкритим кодом на Kotlin із платформи *GitHub* (https://github.com/). Для вибраних проєктів було отримано метрики RFC та CBO з використанням інструменту CodeMR (https://www.codemr.co.uk/).

Дослідження [12] показують, що існує необхідність будувати математичні моделі під конкретний фреймворк чи мову програмування, тому впевнимось, що наявна модель для мови Java не підходить до емпіричних даних із метрик застосунків на Kotlin.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

Застосуємо модель з [11] для даних із метрик застосунків на Kotlin. Модель має такі параметри: $\hat{\lambda}_1 = 0{,}7609789$, $\hat{\lambda}_2 = 0{,}3105496$, вектор вибіркових середніх $\bar{\mathbf{Z}} = \{7{,}234; 4{,}465\}$, довірча ймовірність $\alpha = 0{,}005$, обернена коваріаційна матриця S_N^{-1},

$$S_N^{-1} = \begin{pmatrix} 0{,}69498 & -1{,}28952 \\ -1{,}28952 & 3{,}12612 \end{pmatrix}.$$

На рисунку 1 показано еліпс прогнозування за (1) та довірчої ймовірності 0,005 разом із нормалізованими даними з метрик CBO та RFC застосунків на Java та застосунків на Kotlin.

У разі застосування параметрів для нормалізації та вилучення викидів із роботи [11] 25 проєктів застосунків на Kotlin (~25% проєктів) варто вважати викидами, що ставить під сумнів достовірність застосування цієї моделі для даних із метрик CBO та RFC застосунків на Kotlin.

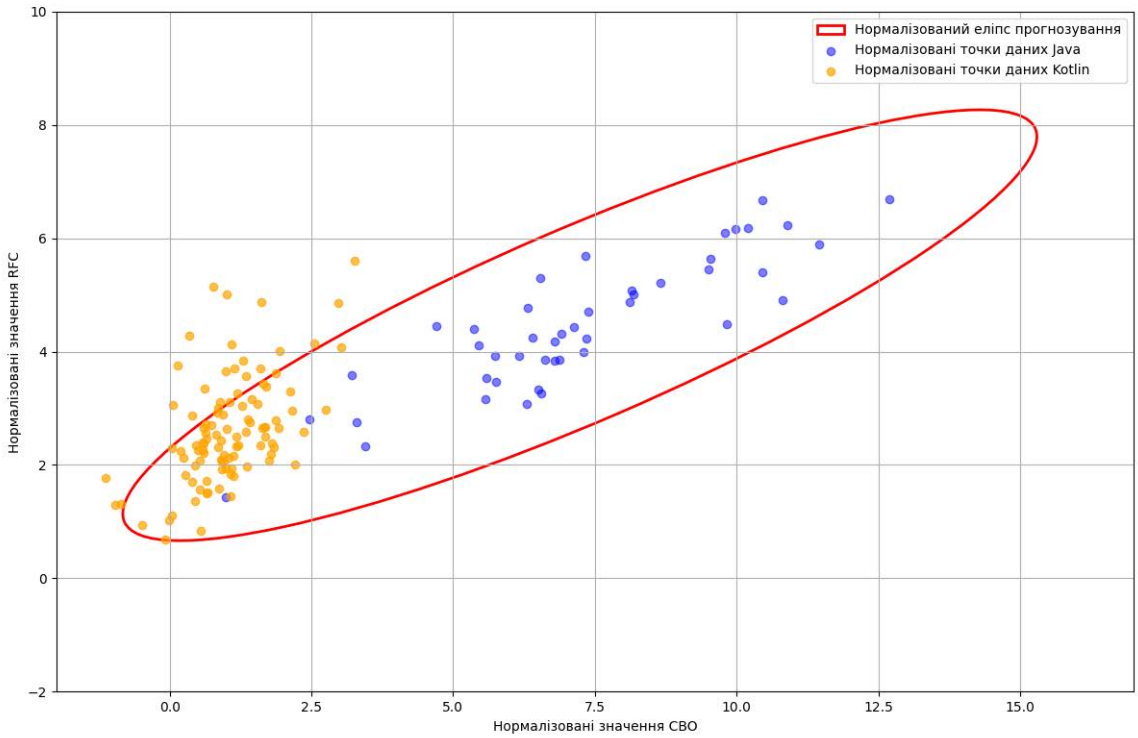

Рис. 1. Нормалізований еліпс прогнозування для Java разом із метриками CBO та RFC для обох мов програмування

Суттєвою є й різниця в оцінках вибіркового середнього цих метрик: для проєктів на Java маємо оцінку математичного сподівання метрик CBO та RFC 11,88 та 18,36 відповідно, тоді як для Kotlin аналогічні значення становлять 2,20 та 7,46.

З використанням тесту Мардія [13] перевіримо розподіл емпіричних даних із метрик на нормальність. Значення двовимірної асиметрії $N\beta_1/6$ дорівнює 70,867, тоді як критичне значення – 9,3. Значення двовимірної ексцесу β_2 дорівнює 14,066, тоді як критичне значення становить 9,48. Обидва емпіричні значення перевищують критичні, що дозволяє зробити висновок про відхилення емпіричного розподілу даних із метрик CBO та RFC застосунків мовою Kotlin від нормального.

Із припущенням про нормальність розподілу та з формулою (1) побудуємо еліпс прогнозування та відобразимо його графічно на рисунку 2.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

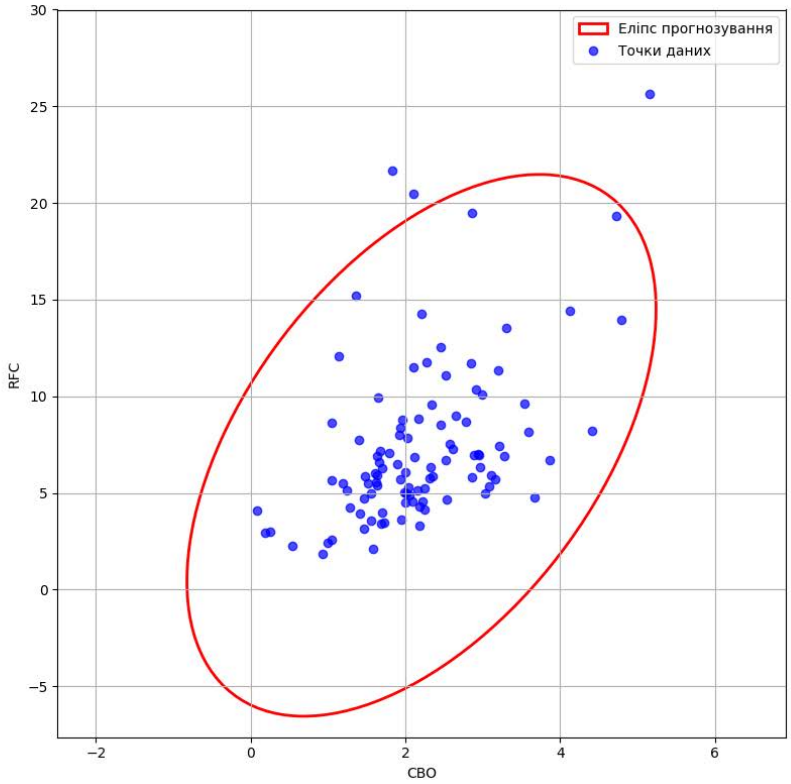

Рис. 2. Еліпс прогнозування разом із метриками CBO та RFC

З рисунку видно, що для трьох точок даних значення перебувають за межами еліпса прогнозування, що разом із результатами тесту Мардія свідчить про негаусівський розподіл даних.

Негаусівський розподіл даних із метрик CBO та RFC застосунків на Kotlin та різні діапазони існуючої моделі для Java зумовлюють потребу в побудові оригінальних моделей для мови Kotlin.

Як було зазначено вище, дані з метрик CBO та RFC корелюють одне з одним [3], тому для нормалізації варто брати перетворення, які враховують цю особливість розподілу даних. Одне з таких перетворень – двовимірне перетворення Бокса – Кокса (далі – ПБК), яке можна представити як [4]:

$$Z_j = x(\lambda_j) = \begin{cases} \left(X_j^{\lambda_j} - 1\right)/\lambda_j, & \text{if} \quad \lambda_j \neq 0; \\ \ln\left(X_j\right), & \text{if} \quad \lambda_j = 0. \end{cases} \tag{4}$$

Тут Zj – гаусівська змінна, λ_j – параметр ПБК ($j = 1,2$).

Оцінити параметри двовимірного ПБК можна за методом найменших квадратів:

$$l(\mathbf{X}, \text{¸}) = \sum_{j=1}^{k} (\text{»}_j - 1) \sum_{i=1}^{N} \ln(X_{j_i}) - \frac{N}{2} \ln\left[\det(\mathrm{S}_N)\right], \tag{5}$$

тут k=2; $\theta = \{\lambda_1, \lambda_2\}$; SN – вибіркова коваріаційна матриця.

За методом максимальної правдоподібності (5) маємо такі значення параметрів: $\hat{\lambda}_1 = -0{,}180625$, $\hat{\lambda}_2 = 0{,}730525$.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

За формулою (4) з використанням параметрів $\hat{\lambda}_1$ та $\hat{\lambda}_2$ виконано нормалізацію емпіричних даних із метрик CBO та RFC застосунків мовою Kotlin.

Після нормалізації треба розрахувати значення SMD та порівняти його із двома критичними значеннями: значенням правої частини рівняння еліпса (2) – 10,5966 та правою частиною рівняння еліпса (3) – 11,401.

Серед отриманих даних відсутні значення SMD, що перевищують обидва критичні значення, тому модель можна вважати побудованою. Для цієї моделі вектор вибіркових середніх $\bar{\mathbf{Z}} = \{1,0268; 1,5734\}$, а обернена коваріаційна матриця S_N^{-1}:

$$S_N^{-1} = \begin{pmatrix} 2,49 & -2,92 \\ -2,92 & 11,01 \end{pmatrix}.$$

На рисунку 3 наведено обидва еліпси разом з емпіричними даними з метрик CBO та RFC.

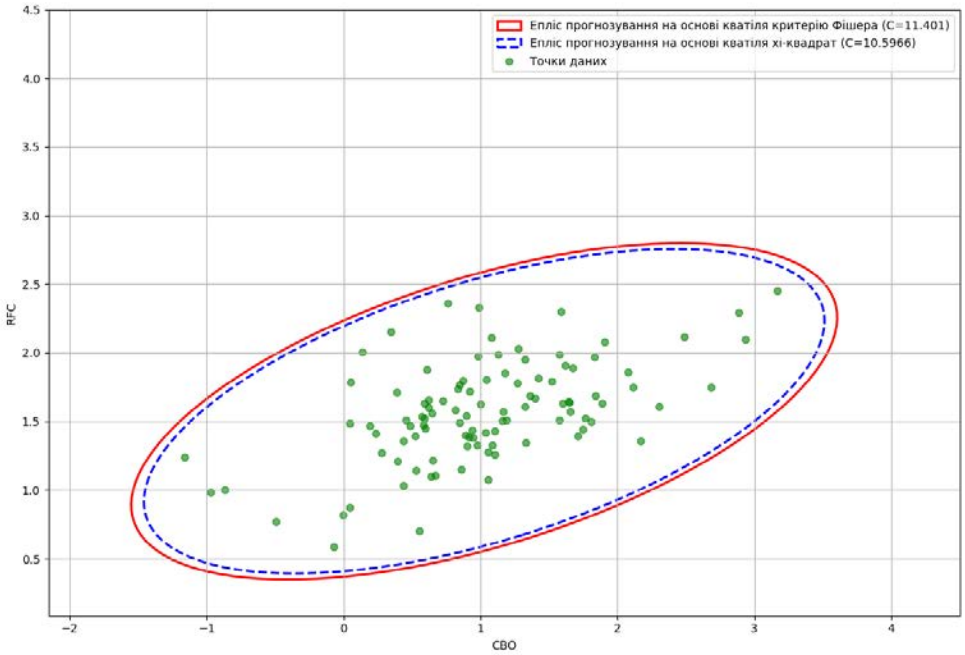

Рис. 3. Еліпси прогнозування для визначення викидів і емпіричні дані з метрик CBO та RFC

Отримані результати дозволяють зробити висновок про незначну відмінність моделей для пошуку та вилучення викидів залежно від вибраної тестової статистики для їх вилучення.

Висновки

У роботі вдосконалено математичні моделі для визначення викидів у двовимірних даних із метрик CBO та RFC на основі еліпса прогнозування для нормалізованих за допомогою нормалізуючого перетворення Бокса – Кокса даних, з урахуванням тестової статистики на основі квантиля хі-квадрат і тестової статистики на основі *F*-розподілу Фішера. Ці моделі дозволяють визначати викиди у двовимірних даних для програмних застосунків з відкритим кодом на Kotlin. Отримані результати показують більшу достовірність під час пошуку викидів у разі застосування моделі з використанням тестової статистики на основі квантиля хі-квадрат та і тестової статистики на основі квантиля *F*-розподілу Фішера, проте отримана площа еліпса з використанням тестової статистики на основі квантиля хі-квадрат є меншою, що демонструє більшу жорсткість цієї тестової статистики.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

Список використаної літератури

1. Chidamber S.R., Kemerer C.F. A metrics suite for objectoriented design. *IEEE Transactions on Software Engineering*. 1994. Vol. 20, № 6. P. 476–493. http://dx.doi.org/10.11 09/32.295895.
2. Booch G. Object-Oriented Analysis and Design with Applications. Pearson Education, Limited, 2011. 608 p.
3. Molnar A.-J., Neamţu A., Motogna S. Evaluation of Software Product Quality Metrics. *Communications in Computer and Information Science*. Cham, 2020. P. 163–187. https://doi.org/10.1007/978-3-030-40223-5_8.
4. Johnson R.A., Wichern D.W. Applied Multivariate Statistical Analysis (6th Edition). Prentice Hall, 2007. 800 p.
5. Ghorbani H. Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis. Series "Mathematics and Informatics"*. 2019. P. 583. https://doi.org/10.22190/FUMI1903583G.
6. IoTDS: A One-Class Classification Approach to Detect Botnets in Internet of Things Devices / V.H. Bezerra et al. *Sensors*. 2019. Vol. 19. № 14. P. 3188. https://doi.org/10.3390/s19143188.
7. Kim S.-G., Park D., Jung J.-Y. Evaluation of One-Class Classifiers for Fault Detection: Mahalanobis Classifiers and the Mahalanobis – Taguchi System. *Processes*. 2021. Vol. 9. № 8. P. 1450. https://doi.org/10.3390/pr9081450.
8. Prykhodko S., Prykhodko N., Makarova L., Pugachenko K. Detecting outliers in multivariate non-Gaussian data on the basis of normalizing transformations. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 2017, pp. 846–849, DOI: 10.1109/UKRCON.2017.8100366.
9. Prykhodko S., Prykhodko N., Makarova L., Pukhalevych A. Application of the squared Mahalanobis distance for detecting outliers in multivariate non-Gaussian data. 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Lviv ; Slavske, Ukraine, 2018, pp. 962–965, DOI: 10.1109/TCSET.2018.8336353.
10. Prykhodko S., Prykhodko N., Makarova L., Kudin O., Smykodub T., Prykhodko A. Detecting bivariate outliers on the basis of normalizing transformations for non-Gaussian data. *The Vth International Conference «Advanced Information Systems and Technologies*, AIST 2017, May 16–17. 2017. Sumy. URL: http://essuir.sumdu.edu.ua/handle/123456789/55754.
11. Prykhodko S., Prykhodko N., Smykodub T. A Joint Statistical Estimation of the RFC and CBO Metrics for Open-Source Applications Developed in Java. *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*, Lviv, Ukraine, 10–12 November 2022. 2022. https://doi.org/10.1109/csit56902.2022.10000457.
12. Prykhodko S., Prykhodko N., Koltsov A. A nonlinear regression model for early LOC estimation of open-source Kotlin-based applications. *Radio Electronics, Computer Science, Control*. 2024. № 1. P. 85. https://doi.org/10.15588/1607-3274-2024-1-8.
13. Mardia K.V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*. 1970. Vol. 57, № 3. P. 519–530. https://doi.org/10.1093/biomet/57.3.519.

References

1. Chidamber, S., & Kemerer, C. (1994). A metrics suite for object oriented design. *IEEE Transactions on Software Engineering*, 20 (6), 476–493. https://doi.org/10.1109/32.295895 [in English].
2. Booch, G. (2011). *Object-Oriented Analysis and Design with Applications*. Pearson Education, Limited [in English].
3. Molnar, A.-J., Neamţu, A., & Motogna, S. (2020). Evaluation of Software Product Quality Metrics. *Communications in Computer and Information Science* (p. 163–187). Springer International Publishing. https://doi.org/10.1007/978-3-030-40223-5_8 [in English].

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8

4. Johnson, R.A., & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis (6th Edition)*. Prentice Hall [in English].
5. Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis, Series: Mathematics and Informatics*, 583. https://doi.org/10.22190/fumi1903583g [in English].
6. Bezerra, V.H., da Costa, V.G.T., Barbon Junior, S., Miani, R.S., & Zarpelão, B.B. (2019). IoTDS: A one-class classification approach to detect botnets in internet of things devices. *Sensors*, 19 (14), 3188. https://doi.org/10.3390/s19143188 [in English].
7. Kim, S.-G., Park, D., & Jung, J.-Y. (2021). Evaluation of One-Class Classifiers for Fault Detection: Mahalanobis Classifiers and the Mahalanobis – Taguchi System. *Processes*, 9 (8), 1450. https://doi.org/10.3390/pr9081450 [in English].
8. Prykhodko, S., Prykhodko, N., Makarova, L., & Pugachenko, K. (2017). Detecting outliers in multivariate non-Gaussian data on the basis of normalizing transformations. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 2017, pp. 846–849, DOI: 10.1109/UKRCON.2017.8100366 [in English].
9. Prykhodko, S., Prykhodko, N., Makarova, L., & Pukhalevych, A. (2018). Application of the squared mahalanobis distance for detecting outliers in multivariate non-Gaussian data. *2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET)*, 962–965. https://doi.org/10.1109/tcset.2018.8336353 [in English].
10. Prykhodko, S., Prykhodko, N., Makarova, L., Kudin, O., Smykodub, T., & Prykhodko, A. (2017). *Detecting bivariate outliers on the basis of normalizing transformations for non-Gaussian data. Proceedings of the Vth International Conference "Advanced Information Systems and Technologies"*, Sumy, Ukraine, 2017, pp. 95–97 [in English].
11. Prykhodko, S., Prykhodko, N., & Smykodub, T. (2022). A Joint Statistical Estimation of the RFC and CBO Metrics for Open-Source Applications Developed in Java. *2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*. IEEE. https://doi.org/10.1109/csit56902.2022.10000457 [in English].
12. Prykhodko, S., Prykhodko, N., & Koltsov, A. (2024). A nonlinear regression model for early LOC estimation of open-source Kotlin-based applications. Radio Electronics, Computer Science, Control, 1, 85. https://doi.org/10.15588/1607-3274-2024-1-8 [in English].
13. Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57 (3), 519–530. [in English].

Кольцов Андрій В'ячеславович – аспірант кафедри програмного забезпечення автоматизованих систем Національного університету кораблебудування імені адмірала Макарова. E-mail: andrew.koltsv@gmail.com, ORCID: 0000-0002-1917-3038.

Koltsov Andrii Vyacheslavovych – Postgraduate Student at the Department of Software of Automated Systems of the Admiral Makarov National University of Shipbuilding. E-mail: andrew.koltsv@gmail.com, ORCID: 0000-0002-1917-3038.

https://doi.org/10.32782/mathematical-modelling/2025-8-1-8