

УДК 004.412:519.237.5

DOI DOI DOI DOI <https://doi.org/10.32782/2663-5941/2026.2.1/38>**Пухалевич А.В.**<https://orcid.org/0000-0002-8827-3251>Національний університет кораблебудування
імені адмірала Макарова**Макарова Л.М.**<https://orcid.org/0000-0003-2903-3001>Національний університет кораблебудування
імені адмірала Макарова**Коваленко В.В.**<https://orcid.org/0009-0004-5841-0624>Національний університет кораблебудування
імені адмірала Макарова

МАТЕМАТИЧНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ ЯКОСТІ ЗАСТОСУНКІВ З ВІДКРИТИМ КОДОМ НА JAVA, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ HIBERNATE

Дослідження присвячене актуальній проблемі оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate. З огляду на розповсюдженість використання мови програмування Java та ORM-технологій, запропонований в даній роботі підхід до оцінювання якості на основі довірчих та прогнозних інтервалів нелінійних регресій для метрик RFC, CVO та WMC сприяє підвищенню достовірності такого оцінювання. Для побудови математичної моделі було зібрано та проаналізовано метрики 70 застосунків з відкритим кодом, де використовується фреймворк Hibernate.

Відповідно до тесту Мардіа, було виявлено перевищення критичного рівня за багатовимірною асиметрією, що свідчило про відсутність нормального розподілу емпіричних даних, і неможливість побудови лінійних моделей. Тому, було проведено нормалізацію даних за допомогою одновимірного перетворення Бокса-Кокса. Після цього тест Мардіа показав, що перетворені дані мають нормальний розподіл.

Завдяки перевірці квадрата відстані Махаланобіса для кожного рядка даних було встановлено відсутність викидів у загальній вибірці, а отже всі емпіричні дані були використанні в нормалізованому вигляді для побудови довірчих та прогнозних інтервалів нелінійних регресій з використанням методу нормалізуючих перетворень. Для цього спочатку було побудовано три лінійні регресії для нормалізованих даних, а потім, використовуючи зворотне нормалізуюче перетворення, з лінійних регресій отримано нелінійні регресії та довірчі і прогнозні інтервали нелінійних регресій. Всі три побудовані нелінійні регресії демонструють прийнятну точність прогнозування загальної вибірки проєктів за критеріями MMRE та PRED(0,25).

Об'єктом дослідження є процес оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate.

Предметом дослідження є математична модель для оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate, побудована на основі метрик RFC, CVO та WMC.

Ключові слова: оцінювання, якість, застосунок, програмна метрика, математична модель, довірчий інтервал, інтервал прогнозування, регресія, Java, фреймворк Hibernate, нормалізуюче перетворення.

Постановка проблеми. У сучасному світі Java продовжує утримувати провідні позиції серед мов програмування завдяки своїй інноваційності, адаптивності та продуктивності. Згідно зі звітом

компанії Azul, в 2025 році майже 70% із понад двох тисяч опитаних ІТ-фахівців у всьому світі зазначили, що більше половини їхніх застосунків створено за допомогою Java або працюють на Java

© Пухалевич А.В., Макарова Л.М., Коваленко В.В., 2026

Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0



Virtual Machine (JVM) [1]. Враховуючи широку популярність Java, а також розповсюджене використання фреймворку Hibernate – одного з найпоширеніших інструментів для роботи з ORM (Object-Relational Mapping) – постає потреба у створенні ефективних засобів оцінювання якості програмного забезпечення, зокрема забезпечення з відкритим вихідним кодом.

Оцінювання якості програмного забезпечення є складним завданням інженерії програмного забезпечення, оскільки вимагає урахування багатьох факторів, що визначають надійність, ефективність та підтримуваність коду. Для розв'язання цієї задачі можуть використовуватися математичні моделі, побудовані на основі метрик RFC, CBO та WMC [2–4]. Одним із найрезультативніших підходів до оцінювання якості застосунків з відкритим кодом є метод, який передбачає обчислення меж довірчих та прогнозних інтервалів нелінійних регресій для показників RFC, CBO та WMC на рівні застосунку для тривимірних точок даних [5]. Під час порівняльного аналізу з існуючими моделями, було виявлено значні відмінності у діапазонах значень метрик RFC, CBO та WMC, що може призводити до зниження достовірності оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate [6]. Таким чином, побудова довірчих та прогнозних інтервалів нелінійних регресій для такого оцінювання дозволить удосконалити існуючі моделі, що сприятиме підвищенню якості розроблюваних застосунків з відкритим кодом.

Аналіз останніх досліджень і публікацій. Одним із способів оцінювання якості програмних застосунків є аналіз метрик програмного коду.

Серед найрозповсюдженіших метрик слід виокремити:

- RFC (Response For Class) демонструє кількість методів, які можуть бути викликані через даний клас; надто велике значення свідчить про перевантаження функціональністю.

- CBO (Coupling Between Object Classes) показує ступінь взаємозалежності між класами: чим вищий цей показник, тим складніше підтримувати та модифікувати код.

- WMC (Weighted Methods per Class) оцінює сумарну складність методів у класі, що дозволяє зробити висновки щодо можливості їх подальшого супроводу та надійності.

У сукупності ці метрики забезпечують цілісне уявлення про якість програмного забезпечення, допомагаючи виявити потенційні ризики, надмірну складність і напрямки для вдосконалення коду.

Для оцінювання якості застосовують лінійні регресійні моделі, що базуються на простих математичних зв'язках між метриками, але часто вони не здатні врахувати нелінійні залежності між метриками RFC, CBO та WMC і виявляються недостатньо точними для складних застосунків з відкритим кодом на Java. У сукупності ці показники формують основу для оцінювання якості коду, проте їх взаємодія має нелінійний характер, який неможливо коректно відобразити за допомогою звичайних лінійних підходів.

З метою підвищення достовірності оцінювання якості застосунків з відкритим кодом на Java, в [5] було запропоновано використовувати довірчі та прогнозні інтервали нелінійних регресій, які враховують взаємозалежність між трьома метриками коду:

- $RFC = f(CBO, WMC)$;
- $CBO = f(WMC, RFC)$;
- $WMC = f(RFC, CBO)$.

Завдяки цьому досягається комплексний опис структури програмного продукту, де кожна метрика виступає одночасно і залежною, і незалежною змінною, що дозволяє точніше виявляти вплив окремих характеристик на загальну якість застосунків.

Метод побудови нелінійних регресій з використанням нормалізуючих перетворень [7] складається з наступних етапів:

1. Перевірка нормальності розподілу даних тестом Мардіа.

2. Нормалізація даних – застосовується одновимірне нормалізуюче перетворення Бокса-Кокса для кожного параметру RFC, CBO та WMC.

3. Розрахунок відстані Махаланобіса для кожного застосунку, що дозволяє виявити аномальні значення (викиди) та видалити їх, якщо відстань перевищує критичний поріг.

4. Побудова по нормалізованим даним лінійних регресій, та довірчих і прогнозних інтервалів для кожної з метрик.

5. Побудова з використанням зворотного нормалізуючого перетворення нелінійних регресій, та довірчих і прогнозних інтервалів нелінійних регресій.

Класифікація рівня якості застосунку відбувається відносно побудованих інтервалів [5]: якщо фактичні значення всіх метрик застосунку розташовані в межах довірчого інтервалу – якість застосунку середня; якщо значення однієї з метрик виходить за верхню межу довірчого інтервалу для цієї метрики – якість застосунку низька; в інших випадках – якість висока.

Запропонований метод дозволяє більш глибоко аналізувати взаємозв'язки між метриками та надає інструмент для комплексного оцінювання архітектурної якості застосунків.

Постановка завдання. Метою дослідження є підвищення достовірності оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate.

Для досягнення поставленої мети потрібно вирішити наступні задачі:

- зібрати дані метрик RFC, CBO та WMC із застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate. Виконати обробку даних, включно з видаленням викидів;

- побудувати три нелінійні регресії RFC(CBO, WMC), CBO(WMC, RFC) та WMC(RFC, CBO) із застосуванням одновимірного нормалізуючого перетворення Бокса-Кокса;

- отримати довірчі та прогнозні інтервали побудованих нелінійних регресій.

Виклад основного матеріалу. Для побудови математичної моделі для оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate, було зібрано дані з 70 проєктів аналогічної тематики, доступних на GitHub [8]. Необхідні для побудови моделей метрики RFC, CBO та WMC були отримані завдяки програмному інструменту з відкритим кодом СК [9]. Для кожного проєкту були розраховані середні арифметичні значення кожної з метрик відносно загальної кількості класів.

Набір застосунків та зібраних з них метрик RFC, CBO та WMC наведено в таблиці 1.

Щоб перевірити нормальність багатовимірних даних використовується тест Мардіа [10], що ґрунтується на визначенні параметрів багатовимірної асиметрії та багатовимірного ексцесу. Оцінка аси-

Таблиця 1

Набір застосунків та метрик RFC, CBO та WMC

№	Застосунок	RFC	CBO	WMC
1	adixchen/demo-restWS-spring-jersey-jpa2-hibernate	5,4000	6,5000	6,2000
2	ajanata/PretendYoureXyzy	9,6290	6,5403	8,6210
3	akshu20791/spring-hibernate-security-react-ant-design-polling-app	6,1087	6,2391	5,9783
4	anakiou/multitenancy	3,0833	4,5833	3,0833
5	bcssp10/multi-tenant	4,4000	5,4000	2,5000
6	bdot/struts-hibernate	2,8571	2,5714	5,5714
7	benssj5/Projet-JEE-Gestion-des-Comptes-Spring-JPA-Hibernate	2,1250	5,5625	4,6250
8	BhartiKhankure/Vehicle_insurance_management_system	8,0000	6,5909	14,2727
9	brishi19791/HealthCare-WebApplication	5,0882	7,4118	7,1765
10	callistaenterprise/blog-multitenancy	5,6842	6,1053	2,6842
11	cassiomolin/jersey-jwt	7,6000	5,6000	5,2571
12	Cepr0/sb-multitenant-db-demo	6,1333	4,1333	4,2000
13	chrissssss/springboot-hibernatespatial-postgis-demo	6,6667	6,0000	13,0000
14	codedeck25/HibernateDemo	7,3333	6,0000	4,8333
15	codercz/biubiu	7,5424	4,3390	10,5932
16	CoffeeLatte007/sharebook-jersey-jwt-spring-hibernate	5,7692	5,1923	6,3846
17	colinbut/sales-order-system	3,0476	9,0952	2,1905
18	cr2121/spring	1,6000	3,0000	5,6000
19	cyela/Spring-Web-App	8,5333	5,8667	14,4000
20	dadoonet/hsearch-es-demo	1,7500	4,2500	4,7500
21	daliasheasha/HibernateApp	7,0000	6,6667	11,6667
22	desperatecat/ShoppingPortal_Vue_Springboot_Project	6,3208	7,5849	8,2830
23	devdilip/todo-springboot-mysql-hibernate-rest	2,8000	4,4000	5,2000
24	diegoneuralo/maven	3,5000	8,1667	2,8333
25	discospiff/JavaFullStackEnterpriseWeb	6,4815	5,3704	6,4444
26	dmitrykhramov/Online-Bank-Simulator	4,8378	6,5135	7,0811
27	dobermai/Hibernate-Flyway-Integration	2,6667	6,0000	3,0000
28	F4bwDP6a6W/spring-mvc-REST	4,7500	3,8750	9,7500
29	foysal-mahmud/E-BookShop----Spring-boot	5,7931	6,4310	7,8621
30	frenmanoj/bookstore	3,1667	4,6667	6,6667
31	Fruzenshtein/security-spr	3,0556	5,3333	2,7778
32	IcedSoul/Shopping	6,1842	4,5526	8,6842

№	Застосунок	RFC	СВО	WMC
33	jamesward/spring_hibernate_hstore_demo	6,4000	5,5000	3,7000
34	jasenkoh/spring-boot-jersey-hibernate	2,6667	5,5833	4,4167
35	Java-Techie-jt/Payment-Resource	3,1667	5,1667	5,0000
36	jjoe64/shiroexample	8,0000	5,9091	6,4545
37	Joryun/Shop	5,4648	5,3521	7,2254
38	Karszt/sda-zdjava124	10,0930	4,5814	4,2326
39	khozema-nullwala/online-shopping	7,8788	6,9091	10,8182
40	lastfm/musicbrainz-data	4,2368	6,4474	5,6842
41	licyun/SpringMVCgbook	7,1667	4,7500	9,5417
42	michaelisvy/hibernate-4-spring-3.1-samples	4,8000	7,7000	4,2000
43	MiteshPuthran/Shopping-Cart-Spring-MVC-Hibernate-Application	7,4444	7,9259	7,4815
44	mohamedYoussefi/eboutique-jpa-hibernate-spring	4,8333	7,0556	10,0000
45	MoonSulong/Ecommerce	4,8649	5,9459	5,8108
46	OKaluzny/spring-boot-rest-api-postgresql	4,2500	8,0000	2,2500
47	OKaluzny/springboot-rest-api-angularjs-https	5,5000	6,2000	5,0000
48	oregami/dropwizard-guice-jpa-seed	6,1163	6,2791	6,2326
49	pfac/demo-spring-data-jpa-hibernate-h2	4,6667	4,0000	6,5000
50	quarkiverse/quarkus-hibernate-types	2,9444	5,6111	5,1111
51	radoslaw-piwowarczyk/PetClinic-Spring-Thymeleaf-Hibernate	3,3333	4,8974	2,6923
52	rajbhatta/springboot_restapi_hibernate	2,2143	4,0000	4,0000
53	ralscha/eds-starter6-jpa	13,6222	7,5111	8,8000
54	RameshMF/springboot-thymeleaf-crud-pagination-sorting-webapp	3,8333	5,1667	4,6667
55	rampatra/springboot-hibernate	9,2857	8,0000	9,5714
56	rid17pawar/HospitalManagement	7,0189	8,0189	6,0943
57	royvanrijn/graalvm-native-microservice	6,0769	3,6923	5,5385
58	rutujaBacchuwar/MuzixAppHibernate	4,8571	5,8571	5,0000
59	scbushan05/book-api	3,6667	5,3333	4,0000
60	scottescue/dropwizard-entitymanager	12,6471	6,9412	11,0588
61	soulcode-acad/goservice	5,0488	5,9756	7,0244
62	sroysf/webapp-springmvc-jpa-hibernate	4,2857	5,8571	4,1429
63	stevegocoding/SpringMVC-SpringDataJpa-Hibernate-Demo	2,0000	3,4000	4,6000
64	swapnildipankar/java_jersey_spring_hibernate_maven	8,2000	5,9500	11,9500
65	ThxGodNovember/AprilDragonSpring	13,3056	5,1111	16,2500
66	tolitius/money-making-project	2,5000	2,6667	3,6667
67	vasylmalik/project-hibernate-1	6,6923	4,3846	7,3846
68	yannbriancon/spring-hibernate-query-utils	5,8750	2,6250	6,3750
69	ykameshrao/spring-mvc-angular-js-hibernate-bootstrap-java-single-page-jwt-auth-rest-api-webapp-framework	5,3235	5,3088	5,1029
70	ZhuXS/Spring-Shiro-Spark	7,4000	6,1400	9,0400

метрії виконується шляхом розрахунку відстаней Махаланобіса для кожного з проєктів [11]:

$$d_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}_N^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}), \quad (1)$$

де: \mathbf{X} – вектор випадкових величин, $\mathbf{x} = \{Y, X_1, X_2\}^T$; $\bar{\mathbf{X}}$ – це вектор вибірових середніх, $\bar{\mathbf{x}} = \{\bar{Y}, \bar{X}_1, \bar{X}_2\}^T$;

N – кількість точок даних;

\mathbf{S}_N^{-1} – обернена коваріаційна матриця, $\mathbf{S}_N^{-1} = \frac{1}{\mathbf{S}_N}$;

\mathbf{S}_N – зміщена вибіркова коваріаційна матриця:

$$\mathbf{S}_N = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T. \quad (2)$$

Формули для обчислення багатовимірної асиметрії та ексцесу за методом Мардіа виглядають наступним чином [12]:

$$\beta_1 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N [(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}_N^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})]^3, \quad (3)$$

$$\beta_2 = \frac{1}{N} \sum_{i=1}^N [(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}_N^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})]^2, \quad (4)$$

де β_1 – багатовимірна асиметрія; β_2 – багатовимірний ексцес.

Щоб перевірити багатовимірну асиметрію β_1 (3) необхідно обчислити тестову статистику $\frac{N}{6} \beta_1$, яка наближається до розподілу χ^2 зі ступенями свободи $k(k+1)(k+2)/6$ та рівнем значущості $\alpha = 0,005$ [12].

Для багатовимірного ексцесу β_2 (4) застосується тестова статистика z_2 , що відповідає

$1-\alpha$ -квантилю нормального розподілу із середнім значенням $k(k+2)$ і дисперсією $8k(k+2)/N$.

Перевірка нормальності виконується завдяки порівнянню значення $\frac{N}{6}\beta_1$ із квантилем $\chi_{k(k+1)(k+2)/6,\alpha}^2$ та значення β_2 із статистикою z_2 . Розподіл вважається не нормальним у тому випадку, коли хоча б одна з тестових статистик перевищує відповідний критичний рівень.

Внаслідок розрахунків були отримані наступні значення:

$$\beta_1 = 2,8425; \frac{N}{6}\beta_1 = 33,1619; \chi_{k(k+1)(k+2)/6,\alpha}^2 = 25,1882; \\ \beta_2 = 17,4883; z_2 = 18,3726.$$

Відповідно до наведених вище результатів, було виявлено перевищення критичного рівня за асиметрією: $33,1619 > 25,1882$, що свідчить про відсутність нормального розподілу даних. Тому, слід спочатку провести нормалізацію даних за допомогою нормалізуючого перетворення Бокса-Кокса:

$$Z_j = \Psi(\lambda_j) = \begin{cases} (X_j^{\lambda_j} - 1) / \lambda_j, & \text{якщо } \lambda_j \neq 0; \\ \ln(X_j), & \text{якщо } \lambda_j = 0, \end{cases} \quad (5)$$

де Z_j – це випадкова змінна з розподілом Гауса (нормалізована змінна);

Ψ – вектор, $\Psi = \{\Psi_{RFC}, \Psi_{CBO}, \Psi_{WMC}\}^T$ перетворення $\mathbf{T} = \Psi(\mathbf{P})$, $\mathbf{T} = \{Z_{RFC}, Z_{CBO}, Z_{WMC}\}^T$ – гаусівський випадковий вектор, $\mathbf{P} = \{RFC, CBO, WMC\}^T$ – негаусівський випадковий вектор;

λ_j – параметр перетворення Бокса-Кокса;

X_j – компонента вектору випадкових величин \mathbf{X} , $\mathbf{X} = \{RFC, CBO, WMC\}^T$;

$j = RFC, CBO, WMC$.

Розраховані значення параметрів перетворення (5) для метрик RFC, CBO та WMC дорівнюють $\lambda_{RFC} = 0,2485$, $\lambda_{CBO} = 1,0492$ та $\lambda_{WMC} = 0,0370$.

Після нормалізації даних за допомогою перетворення Бокса-Кокса з розрахованими значеннями параметрів, було отримано наступні результати:

$$\beta_1 = 1,2705; \frac{N}{6}\beta_1 = 14,8224; \chi_{k(k+1)(k+2)/6,\alpha}^2 = 25,1882; \\ \beta_2 = 14,9362; z_2 = 18,3726.$$

Відповідно до розрахованих значень асиметрії та ексцесу, внаслідок нормалізації дані мають нормальний розподіл та проходять тест Мардіа.

Було обчислено квадрат відстані Махаланобіса для нормалізованих даних для кожного спостереження за (1). Жодне із значень не перевищувало критичне значення, а отже можна стверджувати про відсутність викидів у нормалізованих вхідних даних.

Побудова нелінійних моделей виконується з використанням нормалізуючого перетворення (5).

Для реалізації цього етапу, першочергово необхідно побудувати три лінійні регресії для нормалізованих даних за загальною формулою:

$$\hat{Z}_y = b_0 + b_1 Z_{x_1} + b_2 Z_{x_2} + \varepsilon, \quad (6)$$

де \hat{Z}_y – нормалізована залежна випадкова величина (RFC, CBO або WMC);

b_0 , b_1 та b_2 – оцінки параметрів лінійного рівняння регресії;

Z_{x_1} , Z_{x_2} – нормалізовані незалежні змінні X_1 та X_2 ;

μ – випадкова величина з нормальним розподілом.

Оцінки параметрів, отримані методом найменших квадратів [13] наступні.

Для регресії RFC(CBO, WMC):

$$b_0 = -0,2599, b_1 = 0,1469, b_2 = 0,8532.$$

Для регресії CBO(WMC, RFC):

$$b_0 = 4,1336, b_1 = -0,7971, b_2 = 1,1138.$$

Для регресії WMC(RFC, CBO):

$$b_0 = 1,1486, b_1 = 0,4837, b_2 = -0,0596.$$

Розподіл випадкової величини ε є нормальним для цих моделей.

Шляхом зворотного нормалізуючого перетворення з лінійних моделей (6) отримаємо нелінійні регресійні моделі:

$$\hat{Y}_{RFC} = \left(\left[-0,2599 + 0,1469 \cdot \frac{X_{CBO}^{1,0492} - 1}{1,0492} + 0,8532 \cdot \frac{X_{WMC}^{0,0369} - 1}{0,0369} + \varepsilon \right] \cdot 0,2485 + 1 \right)^{\frac{1}{0,2485}} \\ \hat{Y}_{CBO} = \left(\left[4,1336 - 0,7971 \cdot \frac{X_{WMC}^{0,0369} - 1}{0,0369} + 1,1138 \cdot \frac{X_{RFC}^{0,2485} - 1}{0,2485} + \varepsilon \right] \cdot 1,0492 + 1 \right)^{\frac{1}{1,0492}} \\ \hat{Y}_{WMC} = \left(\left[1,1486 + 0,4837 \cdot \frac{X_{RFC}^{0,2485} - 1}{0,2485} - 0,0596 \cdot \frac{X_{CBO}^{1,0492} - 1}{1,0492} + \varepsilon \right] \cdot 0,0369 + 1 \right)^{\frac{1}{0,0369}} \quad (7)$$

Моделі побудовано в межах метрик RFC[1,6000; 13,6222], CBO[2,5714; 9,0952] та WMC[2,1905; 16,2500].

Розрахунок показників якості для побудованих нелінійних регресійних моделей (7) виконуємо за допомогою критеріїв MMRE та PRED(0,25).

Для регресії RFC(CBO, WMC):

$$MMRE = 0,2777, PRED(0,25) = 0,5571.$$

Для регресії CBO(WMC, RFC):

$$MMRE = 0,2011, PRED(0,25) = 0,7571.$$

Для регресії WMC(RFC, CBO):

$$MMRE = 0,3085, PRED(0,25) = 0,5714.$$

Відповідно до отриманих результатів було визначено, що усі три побудовані нелінійні моделі демонструють прийнятну точність прогнозування загальної вибірки проєктів за критеріями MMRE та PRED(0,25).

Формули рівнянь границь довірчих та прогнозних інтервалів побудованих нелінійних регресій згідно [5]:

$$\Psi_Y^{-1} \left(\hat{Z}_Y \pm t_{\alpha/2, v} S_{Z_Y} \left\{ \frac{1}{N} + (\mathbf{z}_X^+)^T \mathbf{S}_Z^{-1} (\mathbf{z}_X^+) \right\}^{\frac{1}{2}} \right), \quad (8)$$

$$\Psi_Y^{-1} \left(\hat{Z}_Y \pm t_{\alpha/2, v} S_{Z_Y} \left\{ 1 + \frac{1}{N} + (\mathbf{z}_X^+)^T \mathbf{S}_Z^{-1} (\mathbf{z}_X^+) \right\}^{\frac{1}{2}} \right), \quad (9)$$

де Ψ_Y – перетворення для нормалізації залежної змінної Y ;

\hat{Z}_y – оцінка відповідно до лінійного рівняння регресії (6) для нормалізованих даних;

$t_{\alpha/2, v}$ – квантиль розподілу Стюдента з рівнем значущості $\alpha/2$ та $v=N-3$ ступенями вільності;

$S_{Z_Y}^2 = \frac{1}{v} \sum_{i=1}^N (Z_{Y_i} - \hat{Z}_{Y_i})^2$ – залишкова дисперсія, $v=N-k-1$;

\mathbf{z}_X^+ – матриця нормалізованих регресорів зі значеннями $Z_{X_{1j}} - \bar{Z}_{X_1}, Z_{X_{2j}} - \bar{Z}_{X_2}, \bar{Z}_{X_j} = \frac{1}{N} \sum_{i=1}^N Z_{X_{ij}}, j = 1, 2$;

$$\mathbf{s}_z = \begin{pmatrix} S_{z_1 z_1} & S_{z_1 z_2} \\ S_{z_1 z_2} & S_{z_2 z_2} \end{pmatrix}, \text{ де } S_{z_q z_r} = \sum_{i=1}^N [Z_{q_i} - \bar{Z}_q] [Z_{r_i} - \bar{Z}_r], q, r = 1, 2.$$

Для регресії RFC(СВО, WMC) параметри рівнянь (8) та (9) є наступними:

$$S_{Z_Y} = 0.2638, \bar{Z}_1 = 4.9326, \bar{Z}_2 = 1.8380, \mathbf{s}_Z^{-1} = \begin{pmatrix} 0.0062 & -0.0011 \\ -0.0011 & 0.0586 \end{pmatrix}.$$

Для регресії СВО(WMC, RFC) параметри рівнянь (8) та (9) є наступними:

$$S_{Z_Y} = 2.0007, \bar{Z}_1 = 2.0328, \bar{Z}_2 = 1.838, \mathbf{s}_Z^{-1} = \begin{pmatrix} 0.0473 & -0.0416 \\ -0.0416 & 0.0951 \end{pmatrix}.$$

Для регресії WMC(RFC, СВО) параметри рівнянь (8) та (9) є наступними:

$$S_{Z_Y} = 0.1496, \bar{Z}_1 = 2.0328, \bar{Z}_2 = 4.9326, \mathbf{s}_Z^{-1} = \begin{pmatrix} 0.0332 & -0.0054 \\ -0.0054 & 0.0071 \end{pmatrix}.$$

Оцінювання якості з використанням побудованої математичної моделі розглянемо на прикладі застосунку MiteshPuthran/Shopping-Cart-Spring-MVC-Hibernate-Application.

Для регресії RFC(СВО, WMC) довірчий інтервал [6,4501; 8,6328], прогнозний [3,7325; 13,5346], показник якості за цією метрикою – середній, тому що фактичне значення RFC 7,4444 знаходиться в середині меж довірчого інтервалу.

Для регресії СВО(WMC, RFC) довірчий інтервал [5,6699; 6,4659], прогнозний [3,4209; 8,6586], показник якості за цією метрикою – низький, тому що фактичне значення СВО 7,9259 більше верхньої границі довірчого інтервалу.

Для регресії WMC(RFC, СВО) довірчий інтервал [5,6469; 7,8745], прогнозний [3,1551; 13,8262], показник якості за цією метрикою – середній, тому що фактичне значення WMC 7,4815 знаходиться в середині меж довірчого інтервалу.

Оскільки фактичне значення однієї з метрик (СВО) застосунку MiteshPuthran/Shopping-Cart-Spring-MVC-Hibernate-Application виходить за верхню межу довірчого інтервалу цієї метрики – якість застосунку низька.

Висновки. У роботі побудовано математичну модель для оцінювання якості застосунків з відкритим кодом на Java, що створюються з використанням фреймворку Hibernate. При побудові було проведено нормалізацію даних за допомогою одновимірного перетворення Бокса-Кокса. Завдяки перевірці квадрата відстані Махаланобіса для кожного рядка даних було встановлено відсутність викидів в даних, а отже всі емпіричні дані були використанні для побудови нелінійних регресій для метрик RFC, СВО та WMC, та довірчих і прогнозних інтервалів цих регресій з використанням методу нормалізуючих перетворень. Всі три побудовані нелінійні регресії демонструють прийнятну точність прогнозування загальної вибірки проєктів за критеріями MMRE та PRED(0,25).

В подальшому планується застосування багатовимірних нормалізуючих перетворень для покращення якості побудованої математичної моделі.

Список літератури:

1. Sellers S. 2025 State of Java Survey & Report. URL: <https://securityboulevard.com/2025/01/state-of-java-survey-confirms-majority-of-enterprise-apps-are-built-on-java> (дата звернення: 20.10.2025).
2. Bagnulo M., Claise B., Eardley P., Morton A., Akhter A. RFC 8911. URL: <https://www.rfc-editor.org/rfc/rfc8911.html> (дата звернення: 21.01.2026)
3. Shyam R., Chris F. Metrics suite for object-oriented design. URL: <https://www.eso.org/~tcsmgr/oowg-forum/TechMeetings/Articles/OOMetrics.pdf> (дата звернення: 21.01.2026).
4. Aanchal S., Kumar. Metrics for Software Components in Object Oriented Environments: A Survey // *International Journal of Scientific Research in Computer Science and Engineering*. 2013. ISSN 2320-7639. URL: https://isroset.org/pub_paper/IJSRCSE/ISROSET-IJSRCSE-00043.pdf (дата звернення: 21.01.2026).
5. Prykhodko S. Evaluating Quality of Software Systems by the Confidence and Prediction Intervals of Regressions for RFC, СВО and WMC Metrics // *WSEAS Transactions on Systems*. 2024. Vol. 23. P. 322–330. DOI: <https://doi.org/10.37394/23202.2024.23.36>.
6. Prykhodko S., Makarova L., Pukhalevych A. Statistical Analysis of the Three-Dimensional Data of Software Metrics RFC, СВО, and WMC that are not Normally Distributed // *International Conference on Applied Innovations in IT (ICAИТ)*. 2025. P. 127–131. DOI: <https://doi.org/10.25673/119224>.

7. Приходько С.Б. Метод побудови нелінійних рівнянь регресії на основі нормалізуючих перетворень // Тези доповідей міждерж. наук.-методич. конф. “Проблеми математичного моделювання” (Дніпродзержинськ, 13-15 червня 2012 р.). С. 31-33.
8. GitHub. URL: <https://github.com/> (дата звернення: 21.01.2026).
9. Aniche M. Java code metrics calculator (CK). URL: <https://github.com/mauricioaniche/ck> (дата звернення: 21.01.2026).
10. Mardia K. Measures of multivariate skewness and kurtosis with applications // *Biometrika*. 1970. Vol. 57. P. 519–530. DOI: <https://doi.org/10.1093/biomet/57.3.519>.
11. Mahalanobis Distance. URL: <https://www.machinelearningplus.com/statistics/mahalanobis-distance/> (дата звернення: 21.01.2026).
12. Wulandari D., Sutrisno S., Nirwana M. Mardia's Skewness and Kurtosis for Assessing Normality Assumption in Multivariate Regression // *Enthusiastic: International Journal of Applied Statistics and Data Science*. 2021. P. 2–4. DOI: <https://doi.org/10.20885/enthusiastic.vol1.iss1.art1>.
13. Afifi A., Azen S. Statistical analysis: a computer-oriented approach. New York; London: Academic Press, 1972.

Pukhalevych A.V., Makarova L.M., Kovalenko V.V. A MATHEMATICAL MODEL FOR ASSESSING THE QUALITY OF OPEN SOURCE JAVA APPLICATIONS CREATED USING THE HIBERNATE FRAMEWORK

The study is dedicated to the topical issue of quality assessment for open-source Java applications developed using the Hibernate framework. Given the widespread use of the Java programming language and ORM technologies, the approach proposed in this work—based on confidence and prediction intervals of nonlinear regressions for the RFC, CBO, and WMC metrics – contributes to increasing the reliability of such assessments. To construct the mathematical model, metrics from 70 open-source applications utilizing the Hibernate framework were collected and analyzed.

According to Mardia's test, it was found that the multivariate skewness exceeded the critical level, indicating that the empirical data did not follow a normal distribution, making the construction of linear models impossible. Therefore, the data were normalized using univariate Box-Cox transformation. Following this, Mardia's test confirmed that the transformed data followed a normal distribution.

By verifying the squared Mahalanobis distance for each data row, it was determined that there were no outliers in the overall sample. Therefore, all empirical data were used in normalized form to construct confidence and prediction intervals for nonlinear regressions using the method of normalizing transformations. For this purpose, three linear regressions were first constructed for the normalized data; then, using the inverse normalizing transformation, nonlinear regressions and their corresponding confidence and prediction intervals were derived. All three nonlinear regressions demonstrate acceptable prediction accuracy for the project sample based on the MMRE and PRED(0.25) criteria.

The object of the study is the quality assessment process for open-source Java applications built with the Hibernate framework.

The subject of the study is mathematical model for assessment the quality of open-source Java applications created using the Hibernate framework, constructed based on the RFC, CBO, and WMC metrics.

Keywords: *assessment, quality, application, software metric, mathematical model, confidence interval, prediction interval, regression, Java, Hibernate framework, normalizing transformation.*

Дата першого надходження статті до видання: 27.01.2026

Дата прийняття статті до друку після рецензування: 20.02.2026

Дата публікації (оприлюднення) статті 11.05.2026