

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ І. І. МЕЧНИКОВА  
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова праця  
на правах рукопису

Приходько Андрій Сергійович

УДК 004.412:519.237.5

## **ДИСЕРТАЦІЯ**

МАТЕМАТИЧНІ МОДЕЛІ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ З МЕТРИК WEB  
ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ  
PHP-ФРЕЙМВОРКІВ

Спеціальність 122 – Комп'ютерні науки

Галузь знань 12 – Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії (PhD)

Дисертація містить результати власних досліджень. Використання ідей,  
результатів і текстів інших авторів мають посилання на відповідне джерело

\_\_\_\_\_ А.С. Приходько

Науковий керівник Малахов Євгеній Валерійович, доктор технічних наук,  
професор

Одеса – 2026

## ***АНОТАЦІЯ***

**Приходько А.С. Математичні моделі для обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків. – Кваліфікаційна наукова праця на правах рукопису.**

Дисертація на здобуття наукового ступеня доктора філософії (PhD) за спеціальністю 122 "Комп'ютерні науки" (галузь 12 – Інформаційні технології). – Одеський національний університет імені І. І. Мечникова, Міністерство освіти і науки України, Одеса, 2026.

Дисертаційна робота присвячена вирішенню важливого науково-практичного завдання підвищення достовірності результатів обробки інформації з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків, за рахунок побудови певних математичних моделей для обробки відповідної інформації та створенню на їх основі інструментарію інформаційної технології (ІТ) обробки інформації з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків.

Актуальність цієї роботи пов'язана з наступним. Зараз вебзастосунки відіграють вирішальну роль у багатьох аспектах нашого повсякденного життя, включаючи соціальні мережі, електронну комерцію, медицину, освіту, фінанси та інші сфери. При чому багато вебзастосунків розробляється із застосуванням РНР-фреймворків.

Відомо, що створення якісного програмного забезпечення (ПЗ) залишається одним з головних завдань індустрії його розробки. Інша задача – це визначення складності об'єктно-орієнтованого проектування (ООП) ПЗ, яка також є важливою для ІТ-компаній, оскільки тісно пов'язана з подальшими витратами, які виникають в процесі життєвого циклу розробки ПЗ. Ще одна задача, яка пов'язана з прогнозуванням витрат, тривалості та трудомісткості

розробки, – це задача раннього оцінювання кількості рядків коду ПЗ. Зазначене також стосується і вебзастосунків, розроблених із застосуванням РНР-фреймворків.

Рішення зазначених задач, у тому числі, потребує обробки інформації з відповідних програмних метрик та побудови певних математичних моделей. Як показують роботи багатьох дослідників, розподіл програмних метрик відхиляється від нормального, що вимагає застосування таких методів та моделей, які би враховували зазначене відхилення. Крім того, спостерігається кореляція між певними програмними метриками, яку також треба враховувати.

На сьогодні відомі методи і моделі, які дозволяють вести обробку інформації з програмних метрик та вирішувати зазначені вище задачі для ПЗ, що створено різними мовами, у тому числі і РНР. Але, як показують дослідження, математичні моделі для обробки інформації з програмних метрик та рішення вказаних вище задач залежать не тільки від мови програмування, а і від технології та інструментів, які при цьому використовуються. Саме тому виникає потреба у створенні математичних моделей для обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків, та рішення за їх допомогою зазначених вище задач саме для вказаних Web застосунків.

Зараз побудову зазначених моделей здійснюють за наборами даних з відповідних метрик, розподіл яких, як правило, не є гаусівським. Це призводить як до проблеми попередньої обробки інформації, яка стосується перевірки даних на наявність викидів, так і до проблеми побудови самих математичних моделей. Відомі методи визначення викидів, у тому числі і багатовимірних, базуються на припущенні нормальності розподілу даних, яке як правило не підтверджується для даних з метрик ПЗ. А це впливає на достовірність обробки інформації, у тому числі з метрик Web застосунків, що розробляються з використанням РНР-фреймворків. Зазначене призводить до необхідності удосконалення відповідних

математичних моделей, які би дозволили врахувати відхилення від нормальності у наборах даних з метрик Web застосунків, що розробляються з використанням RНР-фреймворків, при їх обробці.

**Метою дисертаційної роботи** є підвищення достовірності результатів обробки інформації з програмних метрик Web застосунків, що розробляються з використанням RНР-фреймворків, шляхом створення математичних моделей, які враховують кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

**Робочою науковою гіпотезою** дисертаційного дослідження є твердження, що підвищення достовірності результатів обробки інформації з програмних метрик Web застосунків, що розробляються з використанням RНР-фреймворків, досягається за рахунок застосування математичних моделей, які враховують кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

Для побудови вказаних математичних моделей пропонується використовувати відомі методи на основі багатовимірних нормалізуючих перетворень, які дозволяють враховувати кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального. Припускається, що застосування зазначених математичних моделей дозволить покращити якість результатів рішення вказаних раніше задач подібно трансферному навчанню – методу машинного навчання, в якому знання (нова інформація), отримані в результаті виконання одного завдання або набору даних, використовуються для покращення якості моделі в іншому пов'язаному завданні або іншому наборі даних.

Для досягнення поставленої мети потрібно вирішити такі завдання:

- проаналізувати існуючі математичні моделі для обробки інформації з програмних метрик та відомі за їх допомогою рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ;
- зібрати набори даних з програмних метрик Web застосунків, що розробляються з використанням PHP-фреймворків;
- побудувати три математичні моделі для обробки інформації з шести програмних метрик Чидамбера і Кемерера (Chidamber and Kemerer, CK) Web застосунків, що розробляються з використанням PHP-фреймворків, та здійснити на їх основі рішення задачі з визначення складності ООП на кожному з трьох етапів за Бучем (Booch);
- удосконалити математичну модель для обробки інформації з трьох програмних метрик (RFC, CBO та WMC) Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик та рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC, а також здійснити на її основі рішення задачі з визначення якості вказаних застосунків;
- удосконалити математичні моделі для обробки інформації з чотирьох програмних метрик та трьохфакторні нелінійні регресійні моделі для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, за метриками діаграми класів на основі чотиривимірного нормалізуючого перетворення;
- удосконалити рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення за метриками діаграми класів;

- на підставі розроблених математичних моделей створити інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

Наукова новизна одержаних результатів полягає у наступному.

1) Вперше побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ці дві математичні моделі дозволяють підвищити достовірність результатів обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM за рахунок врахування як кореляції між ними, так і відхилення тривимірних розподілів цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його першому та другому етапах (через ідентифікацію та семантику класів) за Бучем.

2) Удосконалено математичну модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC і CBO за рахунок врахування як кореляції між ними, так і відхилення двовимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його третьому етапі (через зв'язки між класами) за Бучем.

3) Удосконалено математичну модель для обробки інформації з трьох

програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC, CBO і WMC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків.

4) Удосконалено три математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірних еліпсоїдів прогнозування для зазначених нормалізованих метрик та трифакторних нелінійних регресійних моделей для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Ці математичні моделі дозволяють підвищити достовірність результатів обробки інформації із зазначених метрик за рахунок врахування як кореляції між ними, так і відхилення чотиривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів. Отримані трифакторні нелінійні регресійні моделі в порівнянні з іншими регресійними моделями мають більший відсоток прогнозованих значень PRED, менше значення середньої величини відносної похибки MMRE та менші ширини

довірчого інтервалу та інтервалу прогнозування нелінійної регресії.

5) Удосконалено рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування трифакторних нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT на основі чотиривимірною нормалізуючого перетворення Бокса – Кокса. Рівняння для визначення нижньої та верхньої границь довірчого інтервалу та інтервалу прогнозування дозволяють здійснювати оцінювання точності умовного вибіркового середнього та меж моделювання умовної випадкової величини кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii.

Практичне значення одержаних результатів полягає у наступному. На підставі розроблених математичних моделей створено інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків. Для цього використовувалася мова програмування Python. Було створено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, які розраховані на використання зазначеного ПЗ.

У вступі дисертації розкрита сутність науково-практичного завдання та його значущість, обґрунтовано необхідність проведення дослідження, подана загальна характеристика дисертації в такій послідовності: актуальність теми; зв'язок роботи з науковими програмами, планами, темами; мета і завдання дослідження; наукова новизна та практичне значення одержаних результатів; особистий внесок здобувача; апробація результатів дисертації та публікації.

У першому розділі дисертації виконано аналіз задач, рішення яких, у тому



числі, потребує обробки інформації з відповідних програмних метрик, і побудови певних математичних моделей; здійснено аналіз існуючих методів і моделей для обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків, обґрунтування необхідності проведення досліджень за обраною темою.

У другому розділі дисертації розглянуто існуючі взаємо-зворотні багатовимірні нормалізуючі перетворення та методи оцінювання їх параметрів, здійснено вибір багатовимірних нормалізуючих перетворень для нормалізації двовимірних, тривимірних та чотиривимірних наборів даних з метрик Web застосунків, що розробляються з використанням РНР-фреймворків.

У третьому розділі дисертації побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса; удосконалено модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса; удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса; удосконалено математичні моделі та рівняння для визначення границь довірчого інтервалу та

інтервалу прогнозування трифакторних нелінійних регресії кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT на основі чотиривимірного нормалізуючого перетворення; здійснено порівняння результатів за різними моделями, у тому числі моделями машинного навчання.

У четвертому розділі дисертації на підставі розроблених математичних моделей створено інструментарій ІТ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків. Для цього використовувалася мова програмування Python. Було розроблено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

**Ключові слова:** математична модель, обробка інформації, аналіз даних, програмна метрика, Web застосунок, РНР фреймворк, регресійна модель, нелінійна регресійна модель, нормалізуюче перетворення, відстань Махаланобіса, регресія, нелінійна регресія, машинне навчання, трансферне навчання, інформаційна технологія.

### **Список публікацій здобувача за темою дисертації:**

#### **1) в яких висвітлені наукові результати дисертації:**

- статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України:

1. **Приходько А. С.** Математичні моделі для визначення складності об'єктно-орієнтованого проектування через зв'язки між класами веб-застосунків з відкритим кодом, створених за допомогою РНР фреймворків. *Прикладні*

питання математичного моделювання. 2025. Том 8. № 1. – С. 182-188. DOI: <https://doi.org/10.32782/mathematical-modelling/2025-8-1-17> (Фахове видання категорії Б).

2. **Приходько А. С.**, Малахов Є. В. Удосконалена математична модель оцінювання якості вебдодатків з відкритим кодом. *Центральноукраїнський науковий вісник. Технічні науки*. 2025. Вип. 11(42), ч. II. С. 30-37. DOI: [https://doi.org/10.32515/2664-262X.2025.11\(42\).2.30-37](https://doi.org/10.32515/2664-262X.2025.11(42).2.30-37) (Фахове видання категорії Б).

- статті у періодичних наукових виданнях, проіндексованих у базах даних *Web of Science Core Collection* та/або *Scopus* (крім видань держави, визнаної Верховною Радою України державою-агресором):

3. Prykhodko S. B., Shutko I. S., **Prykhodko A. S.** A nonlinear regression model to estimate the size of web apps created using the CakePHP framework. *Radio Electronics, Computer Science, Control*. 2021. No. 4 (59). P. 129-139. DOI: <http://dx.doi.org/10.15588/1607-3274-2021-4-12> (Фахове видання категорії А, **Web of Science**)

4. **Prykhodko A. S.**, Malakhov E. V. Determining object-oriented design complexity due to the identification of classes of open-source web applications created using PHP frameworks. *Radio Electronics, Computer Science, Control*. 2024. No. 2 (69). P. 160-166. DOI: <https://doi.org/10.15588/1607-3274-2024-2-16> (Фахове видання категорії А, **Web of Science**)

- стаття у виданні, віднесеному до першого - третього квартилів (Q1-Q3) відповідно до класифікації *SCImago Journal and Country Rank* або *Journal Citation Reports* порівнюється до двох наукових публікацій:

5. Prykhodko S. B., Shutko I. S., **Prykhodko A. S.** Early size estimation of web apps created using Codeigniter framework by nonlinear regression models. *Radioelectronic and computer systems*. 2022. Vol. 3 (103). P. 84-94. DOI:

<https://doi.org/10.32620/reks.2022.3.06> (Фахове видання категорії А, **Scopus**, **Q3**, **SJR 2022 0,307**). Стаття у виданні, віднесеному до третього квартилю (Q3) відповідно до класифікації SCImago Journal and Country Rank: <https://www.scimagojr.com/journalsearch.php?q=21101038702&tip=sid&clean=0>

**2) які засвідчують апробацію матеріалів дисертації:**

6. Prykhodko S., Shutko I., **Prykhodko A.** Early LOC Estimation of Web Apps Created Using Yii Framework by Nonlinear Regression Models. *WSEAS Transactions on Computers*. 2021. Vol. 20. P. 321-328. DOI: <http://dx.doi.org/10.37394/23205.2021.20.35> (“WSEAS Transactions on Computers” є міжнародним періодичним науковим виданням, Греція; стаття включена до міжнародних наукометричних баз: Google Scholar, CrossRef)

7. Prykhodko S., **Prykhodko A.**, Shutko I. Estimating the Size of Web Apps Created Using the CakePHP Framework by Nonlinear Regression Models with Three Predictors. *Computer Sciences and Information Technologies (CSIT)* : proceedings of the 2021 IEEE 16th International Conference (Lviv, 22-25 September 2021). IEEE, 2021. P. 333-336. DOI: <http://dx.doi.org/10.1109/CSIT52700.2021.9648680> (**Scopus**)

8. Приходько С. Б., Ворона М. В., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання розміру web-застосунків, що створюються за допомогою PHP-фреймворків. *Проблеми моделювання та розроблення інформаційних систем* : матеріали V науково-практичної інтернет-конференції (Дрогобич, 23 квітня 2021 року). Дрогобич: ДДПУ ім. І. Франка, 2021. С. 28-31.

9. Ворона М. В., **Приходько А. С.**, Шутко І. С. Трьохфакторні нелінійні регресійні моделі для оцінювання розміру WEB-застосунків, що створюються з використанням PHP-фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей вісімнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 23 квітня 2021 р. Одеса, 2021. С. 111-113.

10. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання кількості рядків коду Web-застосунків, що створюються за допомогою PHP-фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2021)* : матеріали II-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2021 р.). Миколаїв: НУК імені адмірала Макарова, 2021. С. 11-12. URL: <http://itconf.nuos.edu.ua/2021/proceedings/>

11. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Математичне моделювання кількості рядків коду Web-застосунків, що створюються за допомогою фреймворку CakePHP. *Стан, досягнення та перспективи інформаційних систем і технологій* : матеріали XXII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів (Одеса, 21-22 квітня 2022 р.). Одеса, Видавництво ОНТУ, 2022. С.36.

12. **Приходько А. С.** Аналітична модель чотиривимірного закону розподілу метрик Web застосунків, що створюються за допомогою фреймворку CakePHP. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2022)* : матеріали III-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2022 р.). Миколаїв: НУК імені адмірала Макарова, 2022. С. 52-53. URL: <http://itconf.nuos.edu.ua/2022/wp-content/uploads/sites/5/NUOS-ITMAS-2022-Proceedings.pdf>

13. **Приходько А. С.**, Малахов Є. В. Визначення метрик Чидамбера та Кемерера за допомогою програмного забезпечення з відкритим кодом. *Free and Open Source Software* : матеріали XIV-ої Міжнародної науково-практичної конференції (Харків, 14-16 лютого 2023 р.) Харків: Харківський національний економічний університет імені Семена Кузнеця, 2023. С. 91. URL: <http://repository.hneu.edu.ua/bitstream/123456789/29041/1/foss-2023-theses.pdf>

14. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проектування через ідентифікацію класів веб застосунків, що створюються за допомогою РНР фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2023)* : матеріали IV-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2023 р.). Миколаїв: НУК імені адмірала Макарова, 2023. С. 35-37. URL: <https://itconf.nuos.edu.ua/2023/proceedings/>

15. **Приходько А. С.** Побудова математичної моделі для оцінювання складності об'єктно-орієнтованого проектування веб застосунків, що створюються за допомогою РНР фреймворків. *Комп'ютерне моделювання та оптимізація складних систем* : матеріали VIII Міжнародної науково-технічної конференції (Дніпро, 1-3 листопада 2023 року). Дніпро: Державний вищий навчальний заклад «Український державний хіміко-технологічний університет», 2023. С. 61-62. URL: <https://udhtu.edu.ua/wp-content/uploads/2023/11/zbirnyk-tez-kmoss-2023.pdf>

16. **Приходько А. С.** Визначення складності об'єктно-орієнтованого проектування за рахунок ідентифікації класів веб застосунків розроблених на основі РНР фреймворків. *Інформаційні технології: теорія і практика* : тези доповідей I (VII) міжнародна науково-практична конференція здобувачів вищої освіти і молодих учених (Дніпро, 20 – 22 березня 2024). / М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». Дніпро : Свідлер А.Л., 2024. С. 310-312. URL: [https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024\(Dnipro\).pdf](https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024(Dnipro).pdf)

17. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проектування через семантику класів вебзастосунків, що створюються за допомогою РНР фреймворків. *Інформаційні технології: моделі,*

*алгоритми, системи (ITMAS – 2024)* : матеріали V-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2024 р.). Миколаїв: НУК імені адмірала Макарова, 2024. С. 17-19. URL: <https://itconf.nuos.edu.ua/2024/proceedings/>

18. **Приходько А. С.** Первинна обробка інформації з метрик RFC та СВО веб застосунків, що створені за допомогою PHP фреймворків. *Інформаційні технології і автоматизація* : матеріали XVII міжнародної науково-практичної конференції (Одеса, 31 жовтня - 1 листопада 2024 р.). Одеса, Видавництво ОНТУ, 2024. С. 498-500. URL: <https://ontu.edu.ua/download/konfi/2024/Collection-of-abstracts-of-the-conference-ITIA-2024.pdf>

19. **Приходько А. С.** Використання метрик RFC, СВО та WMC для оцінювання якості веб-застосунків з відкритим кодом, створених за допомогою PHP фреймворків. *Інженерія програмного забезпечення і передові інформаційні технології (Soft Tech-2024)* : матеріали VII Міжнародної науково-практичної конференції молодих вчених та студентів (Київ, 20-22 листопада 2024 року). Київ: Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», ФІОТ, 2024. С. 82-85. URL: <https://drive.google.com/file/d/1O70Ysxe-z3SS82UqEaBdEXR2VdRW3NwG/view>

20. **Приходько А. С.** Математичні моделі для первинної обробки інформації з метрик RFC, СВО та WMC вебзастосунків, що створені з використанням PHP фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей двадцять другої всеукраїнської конференції студентів і молодих науковців. Одеса, 25 квітня 2025 р. Одеса, 2025. С. 60-61. URL: <http://dspace.pdpu.edu.ua/bitstream/123456789/21944/1/Informatics%2C%20information%20systems.pdf#page=60>

## ABSTRACT

Prykhodko A.S. Mathematical models for processing information from metrics of Web applications developed using PHP frameworks. – Manuscript of the qualification scientific work.

Thesis for the degree of philosophy doctor in specialty 122 "Computer Science" (field 12 – Information technology). – Odesa I. I. Mechnikov National University, Odesa, 2026.

The dissertation is devoted to solving an important scientific and practical problem of increasing the reliability of the results of processing information from software metrics of Web applications developed using PHP frameworks by building certain mathematical models for processing the relevant information and creating on their basis an information technology (IT) toolkit for processing information from software metrics of Web applications developed using PHP frameworks.

The relevance of this work is due to the following. Now, Web applications play a crucial role in many aspects of our daily lives, including social networks, e-commerce, medicine, education, finance, and other areas. Moreover, many Web applications are developed using PHP frameworks.

It is known that creating high-quality software remains one of the main tasks of the software development industry. Another task is to determine the complexity of the object-oriented design (OOD) of the software, which is also important for IT companies, as it is closely related to the subsequent costs that arise during the software development life cycle. Another task related to predicting the costs, duration, and complexity of development is the task of early estimation of the number of lines of software code. This also applies to Web applications developed using PHP frameworks.

The solution of these tasks, among other things, requires processing information



from the relevant software metrics and building certain mathematical models. As shown by the works of many researchers, the distribution of software metrics deviates from normal, which requires the use of such methods and models that would take into account the specified deviation. In addition, there is a correlation between certain software metrics, which must also be taken into account.

Today, there are known methods and models that allow processing information from software metrics and solving the above-mentioned problems for software created in different languages, including PHP. However, as studies show, mathematical models for processing information from software metrics and solving the above-mentioned problems depend not only on the programming language but also on the technology and tools used in this process. That is why there is a need to create mathematical models for processing information from metrics of Web applications developed using PHP frameworks and solving the above-mentioned problems with their help, specifically for these Web applications.

Currently, the construction of these models is carried out on data sets from the corresponding metrics, the distribution of which, as a rule, is not Gaussian. This leads to both the problem of pre-processing information, which concerns checking the data for the presence of outliers, and the problem of building the mathematical models themselves. Known methods for determining outliers, including multivariate ones, are based on the assumption of normality of data distribution, which is usually not confirmed for data from software metrics. And this affects the reliability of information processing, including metrics of Web applications developed using PHP frameworks. This leads to the need to improve the relevant mathematical models, which would allow taking into account deviations from normality in data sets from metrics of Web applications developed using PHP frameworks, when processing them.

**The goal of the dissertation** is to increase the reliability of the results of processing information from software metrics of Web applications developed using

PHP frameworks, by creating mathematical models that take into account the correlation between software metrics and the deviation of their multivariate distribution from normal.

**The working scientific hypothesis** of the dissertation research is the statement that increasing the reliability of the results of processing information from software metrics of Web applications developed using PHP frameworks is achieved through the use of mathematical models that take into account the correlation between software metrics and the deviation of their multivariate distribution from normal.

To build these mathematical models, it is proposed to use known methods based on multidimensional normalizing transformations, which allow taking into account the correlation between software metrics and the deviation of their multidimensional distribution from normal. It is assumed that the use of these mathematical models will improve the quality of the results of solving the previously mentioned problems, similar to transfer learning - a machine learning method in which knowledge (new information) obtained as a result of performing one task or data set is used to improve the quality of the model in another related task or data set.

To achieve the goal, the following tasks need to be solved:

- analyze existing mathematical models for processing information from software metrics and known solutions to problems with determining the complexity of OOP, development quality, and number of lines of software code;
- collect data sets from software metrics of Web applications developed using PHP frameworks;
- to build three mathematical models for processing information from six Chidamber and Kemerer (CK) software metrics of Web applications developed using PHP frameworks, and to implement on their basis the solution of the problem of determining the complexity of OOP at each of the three stages according to Booch;
- to improve the mathematical model for processing information from three

software metrics (RFC, CBO and WMC) of Web applications developed using PHP frameworks, in the form of a three-dimensional forecasting ellipsoid for the specified normalized metrics and equations for determining the widths of confidence intervals and forecasting intervals of three nonlinear regressions of software metrics RFC, CBO and WMC, and to implement on its basis the solution of the problem of determining the quality of the specified applications;

- improve mathematical models for processing information from four program metrics and three-factor nonlinear regression models for estimating the number of lines of code of Web applications developed using various PHP frameworks, using class diagram metrics based on a four-dimensional normalizing transformation;

- improve equations for determining the widths of the confidence interval and the prediction interval of nonlinear regression of the number of lines of code of Web applications developed using PHP frameworks, using a four-dimensional normalizing transformation based on class diagram metrics;

- create an information technology (IT) toolkit, based on the developed mathematical models, for solving problems of determining the complexity of OOP, development quality, and the number of lines of code of Web applications developed using PHP frameworks.

The scientific novelty of the obtained results is as follows.

- 1) For the first time, two mathematical models have been constructed for processing information from two groups of program metrics (respectively, WMC, DIT, and NOC, and WMC, RFC, and LCOM) of Web applications developed using PHP frameworks, in the form of three-dimensional forecasting ellipsoids and regression equations for the specified normalized metrics based on the three-dimensional Box-Cox normalizing transformation. These two mathematical models allow to increase the reliability of the results of processing information from two groups of metrics WMC, DIT and NOC and WMC, RFC and LCOM by taking into account both the correlation

between them and the deviation of the three-dimensional distributions of these metrics from normal, and to implement on their basis the solution of the problem of determining the complexity of OOP at its first and second stages (through the identification and semantics of classes) according to Booch.

2) The mathematical model for processing information from RFC and CBO program metrics of Web applications developed using PHP frameworks has been improved in the form of a forecast ellipse and regression equations for the specified normalized metrics based on the two-dimensional Box-Cox normalizing transformation. This mathematical model allows to increase the reliability of the results of processing information from RFC and CBO metrics by taking into account both the correlation between them and the deviation of the two-dimensional distribution of these metrics from normal, and to implement on their basis the solution of the problem of determining the complexity of OOP at its third stage (through the connections between classes) according to Booch.

3) The mathematical model for processing information from three software metrics RFC, CBO and WMC of Web applications developed using PHP frameworks has been improved in the form of a three-dimensional forecasting ellipsoid for the specified normalized metrics and equations for determining the widths of confidence intervals and forecasting intervals of three nonlinear regressions of software metrics RFC, CBO and WMC based on the three-dimensional normalizing Box-Cox transformation. This mathematical model allows us to increase the reliability of the results of processing information from metrics RFC, CBO, and WMC by taking into account both the correlation between them and the deviation of the three-dimensional distribution of these metrics from normal, and to implement on this basis the solution of the problem of determining the quality of Web applications developed using PHP frameworks.

4) Three mathematical models for processing information from four program

metrics (number of lines of code, number of classes, average number of methods per class and DIT metric) have been improved in the form of four-dimensional prediction ellipsoids for the specified normalized metrics and three-factor nonlinear regression models for estimating the number of lines of code of Web applications developed using PHP frameworks, based on the four-dimensional normalizing Box-Cox transformation. These mathematical models allow to increase the reliability of the results of processing information from the specified metrics by taking into account both the correlation between them and the deviation of the four-dimensional distribution of these metrics from normal, and to implement on their basis the solution of the problem of early estimation of the number of lines of code of Web applications developed using such well-known PHP frameworks as CakePHP, Codeigniter and Yii, by three class diagram metrics. The resulting three-factor nonlinear regression models, compared to other regression models, have a higher percentage of predicted values PRED, a lower value of the average relative error MMRE, and smaller widths of the confidence interval and the nonlinear regression prediction interval.

5) The equations for determining the widths of the confidence interval and the prediction interval of three-factor nonlinear regressions of the number of lines of code of Web applications developed using different PHP frameworks have been improved, depending on the number of classes, the average number of methods per class, and the DIT metric based on a four-dimensional normalizing Box-Cox transformation. The equations for determining the lower and upper bounds of the confidence interval and the prediction interval allow for the assessment of the accuracy of the conditional sample mean and the limits of modeling the conditional random variable of the number of lines of code of Web applications developed using such well-known PHP frameworks as CakePHP, Codeigniter, and Yii.

The practical significance of the results obtained is as follows. Based on the developed mathematical models, an information technology (IT) toolkit was created to

solve problems of determining the complexity of OOP, development quality, and the number of lines of code of Web applications developed using PHP frameworks. The Python programming language was used for this purpose. Appropriate software was developed to solve problems of determining the complexity of OOP, development quality, and the number of lines of code of Web applications developed using PHP frameworks designed to use the specified software.

The Introduction of the dissertation reveals the essence of the scientific and practical task and its significance, justifies the need for conducting research, presents a general characteristic of the dissertation in the following sequence: relevance of the topic; connection of the work with scientific programs, plans, topics; goal and objectives of the research; scientific novelty and practical significance of the results obtained; personal contribution of the applicant; approbation of the results of the dissertation and publication.

In Section 1 of the dissertation analyzes the problems, the solution of which, among other things, requires processing information from relevant software metrics and building certain mathematical models; analyzes existing methods and models for processing information from metrics of Web applications developed using PHP frameworks, and justifies the need for research on the selected topic.

In Section 2 of the dissertation, existing multivariate inverse normalizing transformations and methods of estimating their parameters are considered, and multivariate normalizing transformations are selected for normalizing two-dimensional, three-dimensional, and four-dimensional datasets from metrics of Web applications developed using PHP frameworks.

In Section 3 of the dissertation, two mathematical models are constructed for processing information from two groups of program metrics (respectively WMC, DIT and NOC and WMC, RFC and LCOM) of Web applications developed using PHP frameworks, in the form of three-dimensional forecasting ellipsoids and regression

equations for the specified normalized metrics based on the three-dimensional Box-Cox normalizing transformation; The model for processing information from the software metrics RFC and CBO of Web applications developed using PHP frameworks has been improved in the form of a forecasting ellipse and regression equations for the specified normalized metrics based on the two-dimensional Box-Cox normalizing transformation; the mathematical model for processing information from the three software metrics RFC, CBO and WMC of Web applications developed using PHP frameworks has been improved in the form of a three-dimensional forecasting ellipsoid for the specified normalized metrics and equations for determining the widths of the confidence intervals and forecasting intervals of three nonlinear regressions of the software metrics RFC, CBO and WMC based on the three-dimensional Box-Cox normalizing transformation; mathematical models and equations for determining the widths of the confidence interval and the prediction interval of three-factor nonlinear regressions of the number of lines of code of Web applications developed using different PHP frameworks, depending on the number of classes, the average number of methods per class, and the DIT metric based on a four-dimensional normalizing transformation, were improved; results were compared using different models, including machine learning models.

In Section 4 of the dissertation, based on the developed mathematical models, an IT toolkit was created to solve problems of determining the complexity of OOP, the quality of development, and the number of lines of code of Web applications developed using PHP frameworks. The Python programming language was used for this purpose. Appropriate software was developed to solve problems of determining the complexity of OOP, the quality of development, and the number of lines of code of Web applications developed using PHP frameworks.

**Keywords:** mathematical model, information processing, data analysis, software metrics, Web application, PHP framework, regression model, nonlinear regression model, normalizing transformation, Mahalanobis distance, regression, nonlinear regression, machine learning, transfer learning, information technology.



## ЗМІСТ

|                                                                                                                                                                                                                                               |           |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ .....</b>                                                                                                                                                               | <b>29</b> |
| <b>ВСТУП .....</b>                                                                                                                                                                                                                            | <b>31</b> |
| <b>РОЗДІЛ 1 ЗАДАЧІ, РІШЕННЯ ЯКИХ ПОТРЕБУЄ ОБРОБКИ ІНФОРМАЦІЇ З ПРОГРАМНИХ МЕТРИК. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ І МОДЕЛЕЙ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ З ПРОГРАМНИХ МЕТРИК ТА ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ ПРОВЕДЕННЯ ДОСЛІДЖЕНЬ ЗА ОБРАНОЮ ТЕМОЮ .....</b> | <b>41</b> |
| <b>1.1 Опис задач, рішення яких потребує обробки інформації з програмних метрик .....</b>                                                                                                                                                     | <b>41</b> |
| <b>1.2 Аналіз існуючих методів і моделей для обробки інформації з програмних метрик .....</b>                                                                                                                                                 | <b>44</b> |
| <b>1.3 Обґрунтування необхідності проведення досліджень за обраною темою .....</b>                                                                                                                                                            | <b>58</b> |
| <b>1.4 Висновки за розділом 1 .....</b>                                                                                                                                                                                                       | <b>59</b> |
| <b>РОЗДІЛ 2 БАГАТОВИМІРНІ НОРМАЛІЗУЮЧІ ПЕРЕТВОРЕННЯ ДЛЯ НОРМАЛІЗАЦІЇ НАБОРІВ ДАНИХ З МЕТРИК WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ RНР-ФРЕЙМВОРКІВ .....</b>                                                                       | <b>63</b> |

|                                                                                                                                                                                                           |           |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>2.1 Існуючі взаємо-зворотні багатовимірні нормалізуючі перетворення .....</b>                                                                                                                          | <b>63</b> |
| <b>2.2 Вибір багатовимірних перетворень для нормалізації наборів даних з метрик Web застосунків, що розробляються з використанням PHP-фреймворків .....</b>                                               | <b>67</b> |
| <b>2.3 Висновки за розділом 2 .....</b>                                                                                                                                                                   | <b>72</b> |
| <b>РОЗДІЛ 3 ПОБУДОВА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ З МЕТРИК WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ PHP-ФРЕЙМВОРКІВ, НА ОСНОВІ БАГАТОВИМІРНИХ НОРМАЛІЗУЮЧИХ ПЕРЕТВОРЕНЬ .....</b> | <b>74</b> |
| <b>3.1 Математична модель для обробки інформації з програмних метрик RFC, CBO і WMC для рішення задачі з визначення якості Web застосунків .....</b>                                                      | <b>75</b> |
| <b>3.2 Математичні моделі для обробки інформації з програмних метрик Чидамбера та Кемерера для рішення задачі з визначення складності ООП .....</b>                                                       | <b>82</b> |
| <b>3.2.1 Математичні моделі для обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM .....</b>                                                                                         | <b>83</b> |
| <b>3.2.2 Математична модель для обробки інформації з програмних метрик RFC і CBO .....</b>                                                                                                                | <b>93</b> |

|                                                                                                                                                                                                                           |            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>3.3 Математичні моделі для обробки інформації з чотирьох програмних метрик для рішення задачі з оцінювання кількості рядків коду Web застосунків .....</b>                                                             | <b>98</b>  |
| <b>3.4 Висновки за розділом 3 .....</b>                                                                                                                                                                                   | <b>120</b> |
| <b>РОЗДІЛ 4. ІНСТРУМЕНТАРІЙ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ДЛЯ РІШЕННЯ ЗАДАЧ З ВИЗНАЧЕННЯ СКЛАДНОСТІ ООП, ЯКОСТІ СТВОРЕННЯ ТА КІЛЬКОСТІ РЯДКІВ КОДУ WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ PHP-ФРЕЙМВОРКІВ .....</b> | <b>124</b> |
| <b>4.1 Інструментарій інформаційної технології для рішення задачі з визначення складності ООП Web застосунків, що розробляються з використанням PHP-фреймворків .....</b>                                                 | <b>126</b> |
| <b>4.2 Інструментарій інформаційної технології для рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків .....</b>                                                         | <b>140</b> |
| <b>4.3 Інструментарій інформаційної технології для рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків .....</b>                                          | <b>147</b> |
| <b>4.4 Висновки за розділом 4 .....</b>                                                                                                                                                                                   | <b>157</b> |
| <b>ВИСНОВКИ.....</b>                                                                                                                                                                                                      | <b>159</b> |
| <b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....</b>                                                                                                                                                                                    | <b>164</b> |

|                                                                                                                       |            |
|-----------------------------------------------------------------------------------------------------------------------|------------|
| <b>Додаток А. Акти впровадження результатів дисертаційної роботи у розробку вебзастосунків.....</b>                   | <b>187</b> |
| <b>Додаток Б. Акт про використання результатів дисертаційної роботи в навчальному процесі.....</b>                    | <b>192</b> |
| <b>Додаток В. Набір даних з програмних метрик Web застосунків, що розроблені з використанням PHP-фреймворків.....</b> | <b>194</b> |
| <b>Додаток Г. Список публікацій здобувача за темою дисертації.....</b>                                                | <b>202</b> |

## **ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ**

ММП – метод максимальної правдоподібності;

ООП – об’єктно-орієнтоване проєктування;

ПЗ – програмне забезпечення;

АС – Afferent Coupling is the number of classes affected by given class;

APP – APPLication;

CBO – Coupling Between Objects is the number of classes coupled to a given class.

CK – Chidamber S.R. and Kemerer C.F.;

COCOMO – the Constructive Cost Model;

COSMIC – Common Software Measurement International Consortium;

DIT – the Depth of Inheritance Tree metric;

EC – Efferent Coupling is the number of classes from which given class receives effects;

EIF – Extended Isolation Forest;

FPA – Function Point Analysis;

ISBSG – International Software Benchmarking Standards Group;

KLOC – thousand line of code;

KSLOC – thousand source line of code;

LB – lower bound;

LCOM – Lack of Cohesion in Methods;

LOC – line of code;

MCD – the Minimum Covariance Determinant;

MRE – a magnitude of relative error;

MMRE – a mean magnitude of relative error;

NASA – National Aeronautics and Space Administration;

NOC – Number of Children;

PHP – hypertext preprocessor;

PRED – percentage of prediction;

RMS – root mean square;

SMD – the squared Mahalanobis distance;

$SS_E$  – the residual sum of squares;

$SS_R$  – the regression sum of squares;

$SS_T$  – the total sum of squares;

SVM – support vector machine;

SVR – support vector regression;

UB – upper bound;

VIF – variance influence factor.

## ВСТУП

*Актуальність теми.* Зараз вебзастосунки відіграють вирішальну роль у багатьох аспектах нашого повсякденного життя, включаючи соціальні мережі [1, 2], електронну комерцію [3, 4], медицину [5, 6], освіту [7, 8], фінанси [9] та інші сфери. При чому багато вебзастосунків розробляються із застосуванням РНР-фреймворків. Так [10], у 2025 році саме для платформи Web створювалася найбільша кількість програм (53,4%), а мова РНР (<https://www.php.net/>) залишалася однією з популярних, в першу чергу в Back-end (16%), не тільки серед програмістів в Україні [10], але і за кордоном [11]. Крім того, у порівнянні з Node.js та Python-Web, як показують дослідження [12], РНР найкраще підходить для малих та середніх вебзастосунків.

Відомо [13-19], що створення якісного програмного забезпечення (ПЗ) залишається одним з головних завдань індустрії його розробки. Інша задача – це визначення складності об'єктно-орієнтованого проектування (ООП) ПЗ, яка також є важливою для ІТ-компаній, оскільки тісно пов'язана з подальшими витратами, які виникають в процесі життєвого циклу розробки ПЗ [20-24]. Ще одна задача, яка пов'язана з прогнозуванням витрат, тривалості та трудомісткості розробки, – це задача раннього оцінювання кількості рядків коду ПЗ [25-39]. Зазначене також стосується і вебзастосунків, що розробляються із застосуванням РНР-фреймворків [16, 30-33, 36, 37].

Рішення зазначених задач, у тому числі, потребує обробки інформації з відповідних програмних метрик та побудови певних математичних моделей. Як показують роботи багатьох дослідників [20, 22, 37-56], розподіл програмних метрик відхиляється від нормального, що вимагає застосування таких методів та

моделей, які би враховували зазначене відхилення. Крім того, спостерігається кореляція між певними програмними метриками, яку також треба враховувати [22, 37-56].

На сьогодні відомі методи і моделі, які дозволяють вести обробку інформації з програмних метрик та вирішувати зазначені вище задачі для ПЗ, що створено різними мовами [19, 22, 25, 26, 28, 42, 48-55], у тому числі і РНР [26, 30-34, 43-47]. Але, як показують дослідження, математичні моделі для обробки інформації з програмних метрик та рішення вказаних вище задач залежать не тільки від мови програмування, а і від технології та інструментів, які при цьому використовуються [30-34, 49, 56]. Саме тому виникає потреба у створенні математичних моделей для обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків, та рішення за їх допомогою зазначених вище задач саме для вказаних Web застосунків.

На теперішній час побудову вказаних моделей здійснюють за наборами даних з відповідних метрик, розподіл яких, як правило, не є гаусівським. Це призводить як до проблеми попередньої обробки інформації, яка стосується перевірки даних на наявність викидів [40, 57, 58], так і до проблеми побудови самих математичних моделей [59-69]. Відомі методи визначення викидів, у тому числі і багатовимірних, базуються на припущенні нормальності розподілу даних, яке як правило не підтверджується для даних з метрик ПЗ. А це впливає на достовірність обробки інформації, у тому числі з метрик Web застосунків, що розробляються з використанням РНР-фреймворків. Зазначене призводить до необхідності удосконалення відповідних математичних моделей, які би дозволили врахувати як кореляцію, так і відхилення від нормальності у наборах даних з метрик Web застосунків, що розробляються з використанням РНР-фреймворків, при їх обробці.



*Зв'язок роботи з науковими програмами, планами, темами.* Дисертаційну роботу виконано у Одеському національному університеті імені І. І. Мечникова відповідно до планів НДР «Методи, моделі, інформаційні технології розподілених систем підтримки прийняття організаційних рішень» (№ держреєстрації 0121U111663). Здобувач брав участь у виконанні робіт за вказаною темою в якості виконавця.

*Мета і завдання дослідження.* Метою роботи є підвищення достовірності результатів обробки інформації з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків, шляхом створення математичних моделей, які дозволяють враховувати кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

Для досягнення поставленої мети потрібно вирішити такі завдання:

- проаналізувати існуючі математичні моделі для обробки інформації з програмних метрик та відомі за їх допомогою рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ;
- зібрати набір даних з програмних метрик Web застосунків, що розроблені з використанням РНР-фреймворків;
- побудувати три математичні моделі для обробки інформації з шести програмних метрик Чидамбера і Кемерера (Chidamber and Kemerer, СК) Web застосунків, що розробляються з використанням РНР-фреймворків, та здійснити на їх основі рішення задачі з визначення складності ООП на кожному з трьох етапів за Бучем;
- удосконалити математичну модель для обробки інформації з трьох програмних метрик (RFC, CBO та WMC) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик та рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних

регресій програмних метрик RFC, CBO і WMC, а також здійснити на її основі рішення задачі з визначення якості вказаних застосунків;

- удосконалити математичні моделі для обробки інформації з чотирьох програмних метрик та трьохфакторні нелінійні регресійні моделі для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, за метриками діаграми класів на основі чотиривимірного нормалізуючого перетворення;

- удосконалити рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків, на основі чотиривимірного нормалізуючого перетворення за метриками діаграми класів;

- на підставі розроблених математичних моделей створити інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

*Об'єктом дослідження* є процес обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків.

*Предметом дослідження* є математичні моделі для обробки інформації з метрик Web застосунків, що розробляються з використанням РНР-фреймворків.

*Методи дослідження.* У вирішенні поставлених завдань використано методи теорії ймовірності, математичної статистики, багатовимірного статистичного аналізу, регресійного аналізу, машинного навчання. Методи теорії ймовірності, математичної статистики та багатовимірного статистичного аналізу застосовано для знаходження оцінок характеристик випадкових величин, перевірки статистичних гіпотез щодо нормальності емпіричних розподілів даних (як одновимірних так і багатовимірних), оцінювання параметрів нормалізуючих перетворень (як одновимірних так і багатовимірних), побудови еліпсоїдів

прогнозування для нормалізованих даних, пошуку викидів у даних (як одновимірних так і багатовимірних). Методи регресійного аналізу використано для побудови нелінійних регресійних моделей, рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій для оцінювання програмних метрик Web застосунків, що розробляються з використанням PHP-фреймворків. Методи машинного навчання застосовано для побудови регресій опорних векторів (SVR) для оцінювання програмних метрик Web застосунків, що розробляються з використанням PHP-фреймворків.

*Наукова новизна одержаних результатів* полягає у наступному.

1) Вперше побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ці дві математичні моделі дозволяють підвищити достовірність результатів обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM за рахунок врахування як кореляції між ними, так і відхилення тривимірних розподілів цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його першому та другому етапах (через ідентифікацію та семантику класів) за Бучем.

2) Удосконалено математичну модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC і CBO за рахунок врахування як кореляції між ними, так і відхилення двовимірного

розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його третьому етапі (через зв'язки між класами) за Бучем.

3) Удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC, CBO і WMC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків.

4) Удосконалено три математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірних еліпсоїдів прогнозування для зазначених нормалізованих метрик та трифакторних нелінійних регресійних моделей для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Ці математичні моделі дозволяють підвищити достовірність результатів обробки інформації із зазначених метрик за рахунок врахування як кореляції між ними, так і відхилення чотиривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків

як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів. Отримані трифакторні нелінійні регресійні моделі в порівнянні з іншими регресійними моделями мають більший відсоток прогнозованих значень PRED, менше значення середньої величини відносної похибки MMRE та менші ширини довірчого інтервалу та інтервалу прогнозування нелінійної регресії.

5) Удосконалено рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування трифакторних нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Рівняння для визначення нижньої та верхньої границь довірчого інтервалу та інтервалу прогнозування дозволяють здійснювати оцінювання точності умовного вибіркового середнього та меж моделювання умовної випадкової величини кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii.

*Практичне значення одержаних результатів.* На підставі розроблених математичних моделей створено інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків. Для цього використовувалася мова програмування Python. Було створено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, які розраховані на використання зазначеного ПЗ.

Результати дисертаційної роботи, висновки та рекомендації, що містяться у роботі, схвалені та використовуються в наступних установах: ТОВ «АЙПІ РІШЕННЯ», ТОВ НВЦ «Діагностика та контроль», Комунальне підприємство

«Міський інформаційно-обчислювальний центр» (м. Миколаїв), ФОП Білецький В.В. (акти наведені у додатку А). Використання інструментарію ІТ для визначення складності ООП, якості розробки та кількості рядків коду вебзастосунків, дозволило підвищити достовірність визначення складності ООП та якості розробки відповідних вебзастосунків.

Результати дисертаційної роботи також застосовуються в навчальному процесі на кафедрах Комп'ютерних систем та технологій, Математичного забезпечення комп'ютерних систем Одеського національного університету імені І. І. Мечникова (акт наведений у додатку Б).

*Особистий внесок здобувача.* Усі наукові результати отримано здобувачем самостійно. У працях, опублікованих у співавторстві, дисертанту належать: в [32] та [34] – математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик та трифакторні нелінійні регресійні моделі для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворків CakePHP та Codeigniter відповідно, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса; в [70] – математична модель для обробки інформації з програмних метрик WMC, DIT і NOC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса; в [72] – математична модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення ширин довірчих інтервалів та інтервалів прогнозування трьох нелінійних

регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Також, у працях, опублікованих у співавторстві, дисертанту належать набори даних з програмних метрик Web застосунків з відкритим кодом, що розроблені з використанням РНР-фреймворків.

*Апробація результатів досліджень.* Основні результати дисертації доповідалися та обговорювалися на 14 науково-технічних конференціях, зокрема:

- 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT) (Львів, 2021 р.);

- V-ій науково-практичної інтернет-конференції «Проблеми моделювання та розроблення інформаційних систем» (Дрогобич, 2021 р.);

- Вісімнадцятій та двадцять другій всеукраїнських конференціях студентів і молодих науковців «Інформатика, інформаційні системи та технології» (Одеса, 2021 р. та 2025 р.);

- II-й, III-й, IV-й та V-й Всеукраїнських науково-практичних інтернет конференціях «Інформаційні технології: моделі, алгоритми, системи (ITMAS)» (Миколаїв, 2021 р., 2022 р., 2023 р. та 2024 р.);

- XXII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів «Стан, досягнення та перспективи інформаційних систем і технологій» (Одеса, 2022 р.);

- XIV-й Міжнародній науково-практичній конференції «Free and Open Source Software» (Харків, 2023 р.);

- VIII-й Міжнародній науково-технічній конференції «Комп'ютерне моделювання та оптимізація складних систем» (Дніпро, 2023 р.);

- I(VII)-й Міжнародній науково-практичній конференції здобувачів вищої освіти і молодих учених «Інформаційні технології: теорія і практика», (Дніпро, 2024 р.);

- XVII-й Міжнародній науково-практичній конференції «Інформаційні технології і автоматизація» (Одеса, 2024 р.);

- VII-й Міжнародній науково-практичній конференції молодих вчених та студентів «Інженерія програмного забезпечення і передові інформаційні технології (Soft Tech-2024)» (Київ, 2024 р.).

*Публікації.* Результати роботи викладені у 20 публікаціях. Основні наукові результати дисертації опубліковано у 5 статтях, з яких дві статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України, дві статті у періодичних наукових виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus, та одна стаття у виданні, віднесеному до третього квартилю (Q3) відповідно до класифікації SCImago Journal and Country Rank. Ще 15 публікацій (1 стаття у міжнародному періодичному науковому виданні, 11 матеріалів і праць та 3 тези конференцій) засвідчують апробацію матеріалів дисертації. Зазначимо, одні праці конференції включені до міжнародних наукометричних баз Scopus та Web of Science.

**Структура і обсяг дисертації.** Дисертаційна робота складається з вступу, чотирьох розділів, висновків, списку використаних джерел з 138 найменувань і чотирьох додатків. Загальний обсяг дисертації становить 207 сторінок, обсяг основного тексту складає 131 сторінку. Дисертаційна робота містить 14 рисунків і 35 таблиць. У додатках наведені документи про впровадження результатів роботи, зібраний автором набір даних з програмних метрик Web застосунків, що розроблені з використанням PHP-фреймворків, та список публікацій здобувача за темою дисертації.



## **РОЗДІЛ 1 ЗАДАЧІ, РІШЕННЯ ЯКИХ ПОТРЕБУЄ ОБРОБКИ ІНФОРМАЦІЇ З ПРОГРАМНИХ МЕТРИК. АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ І МОДЕЛЕЙ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ З ПРОГРАМНИХ МЕТРИК ТА ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ ПРОВЕДЕННЯ ДОСЛІДЖЕНЬ ЗА ОБРАНОЮ ТЕМОЮ**

У першому розділі дисертації зазначено задачі, рішення яких, у тому числі, потребує обробки інформації з відповідних програмних метрик, і побудови певних математичних моделей; здійснено аналіз існуючих методів і моделей для обробки інформації з програмних метрик, у тому числі і Web застосунків, що розробляються з використанням РНР-фреймворків; виконано обґрунтування необхідності проведення досліджень за обраною темою.

### **1.1 Опис задач, рішення яких потребує обробки інформації з програмних метрик**

Серед задач, рішення яких, у тому числі, потребує обробки інформації з програмних метрик, розглянемо три відомі задачі: оцінювання якості ПЗ, визначення складності ООП ПЗ та раннього оцінювання кількості рядків коду ПЗ.

Відомо [13-19], що створення якісного ПЗ залишається одним з головних завдань індустрії його розробки. А це потребує створення відповідних моделей для визначення якості ПЗ, у тому числі, за програмними метриками його коду.

Ще в дослідженнях NASA [73, 74] було з'ясовано вплив метрик СК на якість ПЗ та наведені їх точкові оцінки на рівні застосунку для трьох рівнів якості:

високий, середній та низький. Ці результати були отримані за метриками програмних систем, які були розроблені із застосуванням мов програмування Java та C++. Також на основі зазначених досліджень була запропонована лінійна залежність метрики RFC від метрик CBO та WMC, яка у подальшому була використана для оцінювання якості ПЗ [75]. Пізніше для оцінювання якості ПЗ були побудовані нелінійні регресійні моделі метрики RFC в залежності від метрик CBO та WMC [76], метрики CBO в залежності від метрик RFC та WMC, а також метрики WMC в залежності від метрик RFC та CBO для застосунків з відкритим кодом на Java [19].

Наведене вище показує, що для визначення якості ПЗ використовують, у тому числі, метрик СК. А як показують дослідження різних авторів [20, 22, 40, 76], у тому числі і дисертанта [70-72, 78-85], розподіл метрик СК суттєво відхиляється від нормального (гаусівського). Тому рішення задачі з визначення якості ПЗ, у тому числі, потребує обробки інформації з програмних метрик СК, зокрема метрик RFC, CBO та WMC, із врахуванням як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального.

Другою задачею, рішення якої, у тому числі, потребує обробки інформації з програмних метрик, є задача визначення складності ООП ПЗ, яка є важливою для ІТ-компаній, оскільки тісно пов'язана з подальшими витратами, які виникають в процесі життєвого циклу розробки ПЗ.

Проектування загалом та ООП зокрема, є одним із важливих етапів розробки ПЗ [86]. Згідно з [87], шість метрик призначені для вимірювання трьох кроків, що не пов'язані з реалізацією, у визначенні процесу ООП за Бучем [88]. Це метрики WMC, DIT, NOC, RFC, CBO та LCOM, які визначають складність ООП у вищезгаданих кроках. Зокрема, метрики WMC, DIT та NOC визначають складність ООП завдяки ідентифікації класів на першому кроці. Згідно з [20], метрики WMC, DIT та NOC пов'язані з визначенням об'єкта, що є одним з

фундаментальних елементів ООП, як зазначено Бучем [89]. Фактично на першому кроці ключові абстракції в просторі задачі ідентифікуються та позначаються як потенційні класи та об'єкти.

Як відомо [20], метрики WMC, RFC та LCOM відносяться до другого кроку (семантики класів) ООП ПЗ, оскільки WMC та RFC визначають, як об'єкти класу можуть «поводитися», коли вони отримують повідомлення, а метрика LCOM стосується упаковки даних і методів у визначенні класу, що забезпечує міру зв'язності класу. Згідно з [20] для визначення складності ООП ПЗ на третьому кроці через зв'язки між класами можуть бути використані дві програмні метрики: зв'язаність між об'єктами СВО та відгук класу RFC.

Наведене вище показує, що для визначення складності ООП ПЗ також використовують метрик СК. А як вже зазначалося раніше розподіл метрик СК суттєво відхиляється від нормального (гаусівського). Тому рішення задачі з визначення складності ООП ПЗ, у тому числі, потребує обробки інформації з програмних метрик СК із врахуванням як кореляції між ними, так і відхилення розподілу цих метрик від нормального.

Третьою задачею, рішення якої потребує обробки інформації з програмних метрик, є задача раннього оцінювання кількості рядків коду ПЗ. Ця задача пов'язана з прогнозуванням витрат, тривалості та трудомісткості розробки тому, що розмір ПЗ є головним фактором, від якого залежить трудомісткість розробки ПЗ, наприклад, в такій відомій моделі як COCOMO II [90]. В свою чергу за трудомісткістю можна спрогнозувати вартість [91] та тривалість розробки ПЗ, для оцінювання якої можна використати, наприклад, моделі COCOMO або ISBSG (<https://www.isbsg.org/>).

В наш час для раннього оцінювання кількості рядків коду ПЗ використовують різноманітні програмні метрики, які в першу чергу можуть бути визначені за діаграмою класів: кількість класів, кількість методів та інші метрики,

зокрема DIT та WMC [26-34]. А як показують дослідження різних авторів [30-34, 41, 43-47, 49-53], у тому числі і дисертанта [31-34, 92-96], розподіл цих метрик також суттєво відхиляється від нормального (гаусівського), що потребує обробки інформації із зазначених програмних метрик із врахуванням як кореляції між ними, так і відхилення розподілу цих метрик від нормального.

## **1.2 Аналіз існуючих методів і моделей для обробки інформації з програмних метрик**

Як було з'ясовано у попередньому підрозділі рішення описаних у ньому задач потребує обробки інформації із сукупності певних програмних метрик із врахуванням як кореляції між ними, так і відхилення розподілу цих метрик від нормального. Спочатку це стосується попередньої обробки даних з програмних метрик, основною задачею якої є очищення багатовимірних даних від викидів.

Сьогодні для визначення наявності викидів у багатовимірних даних існує велика кількість методів. Перша група методів базується на методах багатовимірного статистичного аналізу [97]. До цієї групи в першу чергу відносять методи що засновані на використанні квадрату відстані Махаланобісу [97]. При цьому застосовується математична модель у вигляді багатовимірного еліпсоїду прогнозування, ліва частина рівняння якого представляє собою квадрат відстані Махаланобісу. До переваг цього методу відносять його добру інтерпретацію: викидом вважається та точка даних, яка виходить за поверхню еліпсоїду прогнозування. Також перевагою методу є врахування кореляції між змінними. Суттєвим недоліком методу є те, що еліпсоїд прогнозування може бути використаний лише для тих багатовимірних даних, розподіл яких не відхиляється від нормального.

До першої групи методів частково також можна віднести метод мінімального коваріаційного визначника (Minimum Covariance Determinant, MCD), який є високонадійним оцінювачем багатовимірного розташування та розсіювання даних. Він також служить зручним та ефективним інструментом для виявлення викидів [98]. За цим методом будується так званий робастний еліпсоїд допустимих відхилень (robust tolerance ellipsoid), який за структурою правої частини рівняння схожий з еліпсоїдом прогнозування лише з тою різницею, що замість квадрату відстані Махаланобісу містить квадрат іншої статистичної відстані, до складу якого замість зворотної коваріаційної матриці входить зворотна центруюча матриця (centering matrix). До переваг методу MCD відносять його добру інтерпретацію: викидом вважається та точка даних, яка виходить за поверхню робастного еліпсоїду допустимих відхилень. Недоліком цього методу, як правило, є необхідність призначення допустимих відхилень та не врахування кореляції між змінними.

До другої групи методів можна віднести різноманітні методи кластеризації [99], застосування яких до визначення викидів засновано на тому, що точки, які не потрапляють у жоден кластер і позначаються як "шум", є потенційними викидами. До переваг цих методів в першу чергу відносять їх універсальність: вони, як правило, підходять для різних типів даних (числових, категоріальних), якщо правильно обрана метрика відстані. Недоліками цих методів, як правило, є їх залежність від кількості кластерів, чутливість до масштабу (якщо існує суттєва різниця між значеннями змінних, то результати будуть хибними), не врахування кореляції між змінними.

До третьої групи методів можна віднести методи, що базуються на близькості (proximity-based techniques) [100], які визначають точку даних як викид, коли її розташування (або близькість) суттєво відхиляється від інших

точок. Близькість точки даних може бути визначена різними способами, які дещо відрізняються один від одного.

Найпоширеніші способи визначення близькості для аналізу викидів такі [100]:

- На основі кластерів: для кількісної оцінки показника викидів використовується неналежність точки даних до жодного з кластерів, її відстань від інших кластерів, розмір найближчого кластера або комбінація цих факторів. Проблема кластеризації має додатковий зв'язок із проблемою виявлення викидів, в якій точки або належать до кластерів, або їх слід вважати викидами.

- На основі відстані: для визначення близькості використовується відстань точки даних до її  $k$ -найближчих сусідів ( $k$ -nearest neighbor) або інший варіант. Точки даних з великими відстанями до  $k$ -найближчих сусідів визначаються як викиди. Алгоритми на основі відстані зазвичай виконують аналіз з набагато детальнішою деталізацією, ніж два інші методи. З іншого боку, ця більша деталізація часто пов'язана зі значними обчислювальними витратами.

- На основі щільності: Кількість інших точок у межах заданої локальної області (області сітки або області на основі відстані) точки даних використовується для визначення локальної щільності. Ці значення локальної щільності можна перетворити на показники викидів. Також можна використовувати інші методи на основі ядра або статистичні методи оцінки щільності. Основна відмінність між методами кластеризації та методами на основі щільності полягає в тому, що методи кластеризації розділяють точки даних, тоді як методи на основі щільності розділяють простір даних.

Очевидно, що всі ці методи тісно пов'язані, оскільки вони базуються на певному понятті близькості (або подібності). Основна відмінність полягає в деталізації цього визначення. Ці різні способи визначення викидів можуть мати різні переваги та недоліки. У багатьох випадках відмінності між цими різними

класами методів розмиваються, коли оцінки викидів визначаються за допомогою кількох із цих понять [100].

До переваг методу k-найближчих сусідів слід зазначити те, що такі методи часто можуть забезпечити більш детальний і точний аналіз, особливо для менших наборів даних, в яких неможлива надійна кластеризація або аналіз щільності. До недоліків цього методу слід віднести те, що оскільки він спирається на відстань, якщо змінні (ознаки) мають різні фізичні одиниці або мають дуже різні масштаби, то результати будуть хибними. Також слід зазначити, метод k-найближчих сусідів є непараметричним методом навчання з учителем, що з однієї сторони є його перевагою, а з іншою – недоліком тому, що у разі застосування адекватної моделі точність відповідного методу буде більшою.

До четвертої групи методів можна віднести методи на основі ансамблів та дерев, а саме ізоляційний ліс (Isolation Forest) [101] та розширений ізоляційний ліс (Extended Isolation Forest, EIF) [102].

Виявлення аномалій за допомогою ізоляційного лісу – це процес, що складається з двох основних етапів [103].

На першому етапі навчальний набір даних використовується для побудови так званого дерева ізоляції (Isolation Tree, iTree).

На другому етапі кожен екземпляр у тестовому наборі проходить через збірку iTrees, отриману на попередньому етапі, і екземпляру призначається відповідна «оцінка аномалії» (anomaly score).

Після того, як усім екземплярам у тестовому наборі було присвоєно оцінку аномалії, можна позначити як «аномалію» будь-яку точку, оцінка якої перевищує заздалегідь визначений поріг, який залежить від домену, до якого застосовується аналіз.

До переваг методу ізоляційний ліс є те, що він чудово підходить для розрізнених викидів. До недоліків цього методу слід віднести те, що він може

створювати зони з помилково низькими показниками аномальності вздовж осей координат, вимагає завдання певного порогу та дає погані результати з кластеризованими викидами.

Основна проблема оригінального методу ізоляційний ліс полягає в тому, що спосіб розгалуження дерев вносить упередженість, яка знижує надійність оцінок аномалій (викидів) для ранжування даних. Це основна мотивація для впровадження методу розширеного ізоляційного лісу, який вирішує проблеми з призначенням балів аномалій заданим точкам даних [102]. В [102] автори пропонують два різні підходи для покращення ситуації. По-перше, вони пропонують випадкове перетворення даних перед створенням кожного дерева, що призводить до усереднення зміщення. По-друге, що є кращим способом, автори пропонують дозволити розбиття даних на гіперплощини з випадковими нахилами. Такий підхід призводить до усунення артефакту, що спостерігається на теплових картах (heat maps) оцінок аномалій. Це також призводить до того, що стійкість алгоритму значно покращується за допомогою цього методу.

Головна перевага методу EIF полягає у вирішенні проблеми класичного методу ізоляційний ліс, який може створювати зони з помилково низькими показниками аномальності вздовж осей координат. До переваг методу EIF слід також віднести те, що він є непараметричним, не вимагає припущень щодо розподілу даних.

До недоліків методу EIF слід віднести таке. По-перше, це складність інтерпретації: важко пояснити, чому саме конкретна точка була визнана викидом (аномальною) оскільки розрізи тепер є лінійними комбінаціями багатьох змінних (ознак). По-друге, це збільшена обчислювальна складність, яка пов'язана з тим, що вибір випадкових нахилів для гіперплощин потребує більше математичних операцій (скалярних добутків векторів), ніж просте порівняння значення змінної (ознаки) з порогом у оригінальному методі ізоляційний ліс. І, по-третє, це



чутливість до масштабування. На відміну від оригінального методу ізоляційний ліс, де масштаб не має значення (бо перерізи йдуть по одній осі), в методі EIF масштабування даних є обов'язковим.

Враховуючи зазначені переваги та недоліки вказаних методів визначення викидів (аномалій) в багатовимірних даних, в наш час для пошуку викидів у даних програмних метрик в основному використовують методи, які мають просту інтерпретацію результатів, не вимагають припущень щодо розподілу первинних даних та враховують кореляцію між змінними (ознаками) [19, 22, 28, 30-40, 42-57, 59-62, 76, 77]. Серед таких методів в першу чергу можна виділити метод визначення викидів у багатовимірних негаусівських даних на основі квадрату відстані Махаланобісу для нормалізованих даних [40, 104]. При цьому застосовується математична модель у вигляді багатовимірного еліпсоїду прогнозування для нормалізованих даних. На відміну від класичного методу на основі використання квадрату відстані Махаланобісу [97], використання нормалізації даних в першу чергу за рахунок застосування багатовимірних нормалізуючих перетворень призводить до відсутності необхідності мати припущення щодо нормальності розподілу первинних даних. Але при цьому багатовимірний розподіл нормалізованих даних не повинен відхилятися від гаусівського.

Згідно методу визначення викидів у багатовимірних негаусівських даних на основі використання квадрату відстані Махаланобісу для нормалізованих даних [40, 104] спочатку багатовимірні негаусівські дані випадкового вектору  $\mathbf{X} = \{X_1, X_2, \dots, X_m\}^T$  нормалізують у дані гаусівського випадкового вектору  $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_m\}^T$  за допомогою бієктивного багатовимірного перетворення

$$\mathbf{Z} = \psi(\mathbf{X}). \quad (1.1)$$

При цьому бієктивність перетворення (1.1) у тому числі означає те, що воно (це перетворення) має зворотне перетворення

$$\mathbf{X} = \boldsymbol{\Psi}^{-1}(\mathbf{Z}). \quad (1.2)$$

Значення багатовимірних точок даних  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$  нормалізують за допомогою перетворення (1.1). Після чого обчислюють квадрат відстані Махаланобіса для кожної багатовимірної точки даних  $i$ ,  $i=1,2,\dots,N$ , за наступною формулою:

$$d_i^2 = (\mathbf{Z}_i - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z}_i - \bar{\mathbf{Z}}) \quad (1.3)$$

де  $\bar{\mathbf{Z}}$  – це вектор вибірових середніх нормалізованих даних, а  $\mathbf{S}_N$  – це вибіркова коваріаційна матриця, як визначається як

$$\mathbf{S}_N = \frac{1}{N} \sum_{i=1}^N (\mathbf{Z}_i - \bar{\mathbf{Z}})(\mathbf{Z}_i - \bar{\mathbf{Z}})^T \quad (1.4)$$

або у більш звичній формі як

$$\mathbf{S}_N = \begin{pmatrix} S_{Z_1 Z_1} & S_{Z_1 Z_2} & \dots & S_{Z_1 Z_m} \\ S_{Z_1 Z_2} & S_{Z_2 Z_2} & \dots & S_{Z_2 Z_m} \\ \dots & \dots & \dots & \dots \\ S_{Z_1 Z_k} & S_{Z_2 Z_k} & \dots & S_{Z_k Z_m} \end{pmatrix}, \quad (1.5)$$

$$\text{де } S_{z_q z_r} = \frac{1}{N} \sum_{i=1}^N [z_{qi} - \bar{z}_q][z_{ri} - \bar{z}_r], \quad q, r = 1, 2, \dots, m.$$

Далі кожне значення квадрата відстані Махаланобіса  $d_i^2$ , яке було обчислено за (1.3), порівнюють з тестовою статистикою – це квантилем розподілу хі-квадрат  $\chi_{m,\alpha}^2$  з кількістю ступенів свободи  $N - m$  та рівнем значущості  $\alpha$ . При чому рекомендується брати значення  $\alpha$  як 0,005. Точки даних, для яких значення квадрата відстані Махаланобіса  $d_i^2$  більше квантиля розподілу хі-квадрат  $\chi_{m,\alpha}^2$ , вважаються викидами, і ці значення відсікаються. Після відсікання викидів зменшена кількість багатовимірних точок даних знову нормалізується за допомогою перетворення (1.1), доки всі значення квадрата відстані Махаланобіса (1.3) не будуть меншими або рівними квантилю розподілу хі-квадрат  $\chi_{m,\alpha}^2$ . В роботі [40] використання цього методу було проілюстровано на прикладі визначення викидів у чотиривимірних даних з програмних метрик інформаційних систем з відкритим кодом на РНР. Для нормалізації даних в [40] використовувалося чотиривимірне перетворення Джонсона сімейства  $S_U$ , яке є бієктивним: тобто дозволяє ще працювати з ізоморфними множинами.

Зазначимо, що при цьому у якості математичної моделі в цьому методі застосовується рівняння багатовимірного еліпсоїду прогнозування, ліва частина якого представляє собою квадрат відстані Махаланобіса для нормалізованих даних, а права – квантиль  $\chi_{m,\alpha}^2$

$$(\mathbf{Z} - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z} - \bar{\mathbf{Z}}) = \chi_{m,\alpha}^2. \quad (1.6)$$

І тоді точки нормалізованих даних, які виходять за поверхню багатовимірного еліпсоїду прогнозування (1.6), вважаються викидами, і вони відсікаються.

Зазначимо, що у разі пошуку викидів у двовимірних даних, розподіл яких відхиляється від гаусівського, рівняння (1.6) буде описувати еліпс прогнозування для нормалізованих даних. Крім того, в правій частині рівняння (1.6) замість квантилю  $\chi^2_{m,\alpha}$  може застосовуватися тестова статистика на основі  $F$ -розподілу Фішера. В роботі [22] таке рівняння було побудовано для нормалізованих метрик СВО та RFC програмних застосунків з відкритим кодом на Java

$$(\mathbf{Z} - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z} - \bar{\mathbf{Z}}) = \frac{2(N^2 - 1)}{N(N - 2)} F_{2, N-2, \alpha}, \quad (1.7)$$

де  $F_{2, N-2, \alpha}$  – це квантиль  $F$ -розподілу з 2 і  $N - 2$  ступенями свободи та рівнем значущості  $\alpha$ , що дорівнює 0,005.

За 45 точками даних з [22] у рівнянні (1.7) права його частина дорівнює 12,57, вектор  $\bar{\mathbf{Z}}$  визначається як  $\bar{\mathbf{Z}} = \{7,234; 4,465\}^T$ , а зворотна матриця  $\mathbf{S}_N^{-1}$  є такою:

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

Для нормалізації даних тут використовувалося двовимірне перетворення Бокса – Кокса оцінками параметрів  $\hat{\lambda}_1 = 0,7609789$  та  $\hat{\lambda}_2 = 0,3105496$  для метрик СВО та RFC відповідно.

Підкреслимо, що згідно [22] у разі використання в рівнянні (1.7) в правій частині квантиля  $\chi^2_{m,\alpha}$  подібно (1.6) регіон допустимих значень метрик СВО та RFC приблизно на 20% зменшився тому, що значення  $\chi^2_{2,0.005}$  в цьому випадку дорівнює 10,60. Це вказує на необхідність застосування тестової статистики на основі  $F$ -розподілу.

Зазначимо, що в [22] рівняння (1.7) було запропоновано використовувати для визначення складності ООП через зв'язки між класами програмних застосунків з відкритим кодом на Java. В [55] рівняння (1.7) було уточнено за рахунок збільшення кількості точок даних метрик СВО та RFC застосунків з відкритим кодом на Java. Що стосується визначення складності ООП Web застосунків, які розробляються з використанням PHP-фреймворків, то на сьогодні дисертанту не відомі роботи на основі метрик СК, окрім власних досліджень [70, 71, 79-82].

Застосування для визначення викидів у багатовимірних даних з програмних метрик рівнянь (1.6) та (1.7), на нашу думку, обумовлено ще необхідністю побудови, як правило, нелінійних регресій для рішення задач, які зазначені в підрозділі 1.1. А як відомо, одним із ефективних методів побудови нелінійних регресій є метод, що базується на використанні нормалізуючих перетворень [97, 105-109], у тому числі і багатовимірних [59, 61, 62], застосування яких не потребує простого перебору. Крім того, побудова нелінійних регресій на основі нормалізуючих перетворення вирішує проблему визначення довірчих інтервалів та інтервалів прогнозування нелінійних регресій, а також проблему пошуку розподілу відхилень значень залежної випадкової величини від лінії регресії [59, 61, 62]. Довірчі інтервали нелінійних регресій дозволяють отримувати інтервальні оцінки вибіркового середнього залежної

випадкової величини, а інтервали прогнозування визначати викиди в процесі побудови нелінійних регресійних моделей [59, 61, 62].

Як відмічалось в підрозділі 1.1 для оцінювання якості ПЗ були побудовані нелінійні регресійні моделі метрики RFC в залежності від метрик CBO та WMC [76], метрики CBO в залежності від метрик RFC та WMC, а також метрики WMC в залежності від метрик RFC та CBO для застосунків з відкритим кодом на Java [19]. Ці нелінійні регресійні моделі були побудовані на основі тривимірного перетворення Бокса – Кокса у такому вигляді [19]:

$$Y = \left[ \hat{\lambda}_Y (\hat{Z}_Y + \varepsilon) + 1 \right]^{1/\hat{\lambda}_Y}, \quad (1.8)$$

де  $\hat{Z}_Y$  – це оцінка нормалізованої залежної змінної за лінійним регресійним рівнянням  $\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2$  для нормалізованих незалежних змінних (предикторів)  $Z_1$  та  $Z_2$ , значення яких отримують за допомогою тривимірного перетворення Бокса – Кокса;  $\varepsilon$  – це гаусівська випадкова величина,  $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ .

Щоби записати нелінійну регресійну модель для метрики RFC як (1.8), треба підставити  $RFC$ ,  $\hat{\lambda}_{RFC}$ ,  $\hat{Z}_{RFC}$ ,  $Z_{CBO}$ ,  $Z_{WMC}$ ,  $\varepsilon_1$  та  $\sigma_{\varepsilon_1}$  замість  $Y$ ,  $\hat{\lambda}_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\varepsilon$  та  $\sigma_\varepsilon$ , відповідно. Щоби записати нелінійну регресійну модель для метрики CBO як (1.8), треба підставити  $CBO$ ,  $\hat{\lambda}_{CBO}$ ,  $\hat{Z}_{CBO}$ ,  $Z_{RFC}$ ,  $Z_{WMC}$ ,  $\varepsilon_2$  та  $\sigma_{\varepsilon_2}$  замість  $Y$ ,  $\hat{\lambda}_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\varepsilon$  та  $\sigma_\varepsilon$ , відповідно. Щоби записати нелінійну регресійну модель для метрики WMC як (1.8), треба підставити  $WMC$ ,  $\hat{\lambda}_{WMC}$ ,  $\hat{Z}_{WMC}$ ,  $Z_{RFC}$ ,  $Z_{CBO}$ ,  $\varepsilon_3$  та  $\sigma_{\varepsilon_3}$  замість  $Y$ ,  $\hat{\lambda}_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\varepsilon$  та  $\sigma_\varepsilon$ , відповідно. Оцінки параметрів цих

нелінійних регресійних моделей для метрики RFC, CBO та WMC наведені в таблиці 1.1 [19].

Крім того, для оцінювання якості ПЗ були побудовані довірчі інтервали та інтервали прогнозування нелінійних регресій метрики RFC в залежності від метрик CBO та WMC [76], метрики CBO в залежності від метрик RFC та WMC, а також метрики WMC в залежності від метрик RFC та CBO для застосунків з відкритим кодом на Java [19].

Таблиця 1.1 – Оцінки параметрів нелінійних регресійних моделей для метрики RFC, CBO та WMC

| No | $Y$        | $\hat{b}_0$ | $\hat{b}_1$ | $\hat{b}_2$ | $\sigma_\varepsilon$ | $MMRE$ | $PRED$ |
|----|------------|-------------|-------------|-------------|----------------------|--------|--------|
| 1  | <i>RFC</i> | -3,69701    | 0,11637     | 4,59287     | 0,2743               | 0,1346 | 0,8627 |
| 2  | <i>CBO</i> | 8,09952     | 4,37867     | -11,4975    | 1,6823               | 0,1974 | 0,7451 |
| 3  | <i>WMC</i> | 0,99535     | 0,12980     | -0,00864    | 0,0461               | 0,1949 | 0,7059 |

Ці інтервали були побудовані на основі тривимірного перетворення Бокса – Кокса у такому вигляді [19]:

$$\psi_Y^{-1} \left( \hat{Z}_Y \pm t_{\alpha/2, \nu} S_{Z_Y} \left\{ \frac{1}{N} + (\mathbf{z}_X^+)^T \mathbf{S}_Z^{-1} (\mathbf{z}_X^+) \right\}^{1/2} \right), \quad (1.9)$$

де  $\psi_Y$  – це компонента нормалізуючого перетворення для залежної змінної  $Y$ ;  $t_{\alpha/2, \nu}$  – це квантиль розподілу Стюдента з  $\nu$  ступенями свободи та рівнем значущості  $\alpha/2$ ;  $\nu = N - 3$ ;  $\mathbf{z}_X^+$  – це вектор з компонентами  $Z_{i_1} - \bar{Z}_1$  і

$Z_{2_i} - \bar{Z}_2$  для  $i$ -го рядка;  $\bar{Z}_j = \frac{1}{N} \sum_{i=1}^N Z_{ji}$ ,  $j=1,2$ ;  $S_{Z_Y}^2 = \frac{1}{N} \sum_{i=1}^N (Z_{Y_i} - \hat{Z}_{Y_i})^2$ ;  $S_Z$  – це матриця  $2 \times 2$

$$S_Z = \begin{pmatrix} S_{Z_1 Z_1} & S_{Z_1 Z_2} \\ S_{Z_1 Z_2} & S_{Z_2 Z_2} \end{pmatrix}, \quad (1.10)$$

$$\text{де } S_{Z_q Z_r} = \sum_{i=1}^N [Z_{qi} - \bar{Z}_q][Z_{ri} - \bar{Z}_r], \quad q, r = 1, 2.$$

Щоби визначити довірчий інтервал нелінійної регресії для метрики RFC за допомогою (1.9), треба підставити  $RFC$ ,  $\psi_{RFC}$ ,  $\hat{Z}_{RFC}$ ,  $Z_{CBO}$ ,  $Z_{WMC}$ ,  $\bar{Z}_{CBO}$  та  $\bar{Z}_{WMC}$  замість  $Y$ ,  $\psi_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\bar{Z}_1$  та  $\bar{Z}_2$ , відповідно. Щоби визначити довірчий інтервал нелінійної регресії для метрики CBO за допомогою (1.9), треба підставити  $CBO$ ,  $\psi_{CBO}$ ,  $\hat{Z}_{CBO}$ ,  $Z_{RFC}$ ,  $Z_{WMC}$ ,  $\bar{Z}_{RFC}$  та  $\bar{Z}_{WMC}$  замість  $Y$ ,  $\psi_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\bar{Z}_1$  та  $\bar{Z}_2$ , відповідно. Щоби визначити довірчий інтервал нелінійної регресії для метрики WMC за допомогою (1.9), треба підставити  $WMC$ ,  $\psi_{WMC}$ ,  $\hat{Z}_{WMC}$ ,  $Z_{RFC}$ ,  $Z_{CBO}$ ,  $\bar{Z}_{RFC}$  та  $\bar{Z}_{CBO}$  замість  $Y$ ,  $\psi_Y$ ,  $\hat{Z}_Y$ ,  $Z_1$ ,  $Z_2$ ,  $\bar{Z}_1$  та  $\bar{Z}_2$ , відповідно.

Інтервал прогнозування нелінійної регресії будується аналогічно (1.9) з тією лише різницею, що до суми у фігурних дужках (1.9) потрібно додати ще 1.

Зазначимо, що відомі роботи для визначення якості ПЗ на основі метрик СК стосуються програмних систем, які розроблені мовами C++ [73, 74], Java [19, 73, 74, 76] та Kotlin [48]. Що стосується визначення якості Web застосунків, які розробляються з використанням РНР-фреймворків, то на сьогодні дисертанту не відомі роботи на основі метрик СК, окрім власних досліджень [72, 84].



Так саме як для оцінювання якості ПЗ використовують, у тому числі, нелінійні регресійні моделі для метрик RFC, CBO та WMC, що побудовані на основі багатовимірних нормалізуючих перетворень, для раннього оцінювання кількості рядків коду ПЗ також застосовують подібні моделі, включаючи застосунки з програмним кодом на PHP [26, 30, 43, 44, 46, 47]. Але в цих роботах автори розглядали в основному інформаційні системи з програмним кодом на PHP [26, 43] або інші програмні застосунки з відкритим кодом на PHP [44, 46, 47], які не відносяться до Web застосунків, що розробляються з використанням PHP-фреймворків. Лише в роботі [30] автори побудували нелінійні регресійні моделі для оцінювання розміру Web-застосунків, що створюються з використанням фреймворку Laravel, із використанням різних нормалізуючих перетворень: десятичного логарифму, одновимірного та чотиривимірного перетворень Джонсона для сімейства  $S_B$ . Згідно з [30], модель, що побудована на основі чотиривимірного перетворення Джонсона для сімейства  $S_B$ , в порівнянні з іншими нелінійними регресійними моделями, має менші значення середньої величини відносної похибки та ширин інтервалу передбачення (прогнозування) нелінійної регресії. Хоча в [30] побудова моделей базувалася всього на 30 точках даних, але отримані результати вказують на переваги використання багатовимірних перетворень для нормалізації даних великої розмірності у порівнянні із застосуванням одновимірних перетворень. Тому у подальшому для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, були розроблені декілька нелінійних регресійних моделей на основі чотиривимірного перетворення для таких фреймворків: CakePHP [31, 32], Yii [33] та Codeigniter [34].

Все зазначене вище вказує на необхідність обробки інформації з програмних метрик для вирішення зазначених у підрозділі 1.1 задач із застосуванням методів, що базуються на багатовимірних нормалізуючих

перетвореннях, які дозволяють враховувати кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

### 1.3 Обґрунтування необхідності проведення досліджень за обраною темою

Для обґрунтування необхідності проведення досліджень за обраною темою окрім аналізу публікацій у попередніх двох підрозділах дисертантом було зроблено ще наступне.

1) Додатково до даних, що наведені в роботах [31-34], за допомогою інструменту PhpMetrics [110] були зібрані дані з програмних метрик 125 Web застосунків, що розроблені з використанням таких п'яти відомих РНР-фреймворків: CakePHP (16 застосунків), CodeIgniter (20 застосунків), Laravel (55 застосунків), Symfony (18 застосунків) та Yii (16 застосунків). Зібрані дані з програмних метрик 125 Web застосунків наведені у додатку В. В таблиці 1.2 наведені описові статистики цих даних, вибіркові середні та середні квадратичні відхилення позначені як Mean та RMS, відповідно. Зазначимо, що вибір саме цих фреймворків був обумовлений тим, що згідно з даними JetBrains (<https://blog.jetbrains.com/phpstorm/2025/10/state-of-php-2025/>) у 2025 році Laravel мав 64% використання, а Symfony – 23%. CodeIgniter, Yii та CakePHP мали менші частки, які дорівнюють відповідно 9%, 6% та 5%. Сума відсотків перевищує 100% тому, що в опитуваннях розробники могли обирати кілька варіантів, які вони регулярно використовують у роботі. Що стосується визначення мінімально необхідного обсягу вибірки, то згідно [97] для побудови чотиривимірного еліпсоїду прогнозування мінімально треба 34 точки даних, а за правилом Гріна (Green's rule) для перевірки якості трифакторної регресійної моделі треба 74 точки даних. Тому додаткові 125 точок даних з програмних метрик Web

застосунків є достатньою кількістю для побудови тих математичних моделей, які пропонуються у роботі.

2) За допомогою цих даних були перевірена можливість використання існуючих моделей для рішення зазначених у підрозділі 1.1 задач саме для Web застосунків, що розроблені з використанням вказаних вище п'яти відомих РНР-фреймворків.

За результатами цієї перевірки було з'ясовано, що існуючі моделі побудовані не тільки для інших діапазонів відповідних метрик, але для деяких комбінацій метрик з 125 Web застосунків, які розроблені з використанням п'яти відомих РНР-фреймворків, приводять до таких значень квадрату відстані Махаланобіса, що сигналізують про викиди для цих моделей.

Таблиця 1.2 – Описові статистики даних

| Метрика                                | Min   | Max     | Mean    | RMS     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 1,008 | 349,736 | 26,453  | 43,120  |
| Кількість класів                       | 11    | 7743    | 465,648 | 899,328 |
| Кількість методів на клас              | 1,41  | 29,55   | 4,971   | 3,489   |
| WMC                                    | 1,75  | 130,09  | 15,268  | 17,245  |
| DIT                                    | 1,06  | 2,44    | 1,495   | 0,294   |
| NOC                                    | 0     | 0,95    | 0,381   | 0,224   |
| CBO                                    | 1,66  | 8,72    | 4,323   | 1,601   |
| RFC                                    | 3,00  | 47,09   | 10,818  | 5,057   |
| LCOM                                   | 0,83  | 5,78    | 2,168   | 0,731   |

Зазначене говорить про неможливість використання існуючих моделей саме для Web застосунків, які розроблені з використанням вказаних вище п'яти відомих РНР-фреймворків, а також обґрунтовує необхідність проведення досліджень за обраною темою.

#### **1.4 Висновки за розділом 1**

У першому розділі дисертації зазначено задачі, рішення яких, у тому числі, потребує обробки інформації з відповідних програмних метрик, і побудови певних математичних моделей; здійснено аналіз існуючих методів і моделей для обробки інформації з програмних метрик, у тому числі і Web застосунків, що розробляються з використанням РНР-фреймворків; виконано обґрунтування необхідності проведення досліджень за обраною темою.

1) Серед задач, рішення яких, у тому числі, потребує обробки інформації з програмних метрик, зазначено три відомі задачі: оцінювання якості ПЗ, визначення складності ООП ПЗ та раннього оцінювання кількості рядків коду ПЗ. Для визначення складності ООП та якості ПЗ використовують різні комбінації метрик СК, розподіл яких суттєво відхиляється від нормального (гаусівського). Це потребує обробки інформації з програмних метрик СК із врахуванням як кореляції між ними, так і відхилення розподілу цих метрик від нормального. Для раннього оцінювання кількості рядків коду ПЗ використовують різноманітні програмні метрики, які в першу чергу можуть бути визначені за діаграмою класів: кількість класів, кількість методів та інші метрики, зокрема DIT та WMC, розподіл яких також суттєво відхиляється від нормального (гаусівського), що потребує обробки інформації із зазначених метрик із врахуванням як кореляції між ними, так і відхилення розподілу цих метрик від нормального.

2) Обробка інформації (даних), у тому числі з програмних метрик, починається попередньої їх обробки, основною задачею якої є очищення багатовимірних даних від викидів. Сьогодні для визначення наявності викидів у багатовимірних даних існує велика кількість методів, які можна поділити на чотири групи. Враховуючи проаналізовані переваги та недоліки цих методів визначення викидів (аномалій) в багатовимірних даних, зазначено те, що в наш час для пошуку викидів у даних програмних метрик в основному використовують методи, які мають просту інтерпретацію результатів, не вимагають припущень щодо розподілу первинних даних та враховують кореляцію між змінними (ознаками). Серед таких методів в першу чергу можна виділити метод визначення викидів у багатовимірних негаусівських даних на основі квадрату відстані Махаланобісу для нормалізованих даних. На відміну від класичного методу на основі квадрату відстані Махаланобісу, застосування нормалізації даних в першу чергу за рахунок багатовимірних нормалізуючих перетворень призводить до відсутності необхідності мати припущення щодо нормальності розподілу первинних даних. Але при цьому багатовимірний розподіл нормалізованих даних не повинен відхилятися від гаусівського.

3) Метод визначення викидів у багатовимірних негаусівських даних на основі квадрату відстані Махаланобісу для нормалізованих даних фактично використовує математичну модель у вигляді багатовимірного еліпсоїду прогнозування для нормалізованих даних. При побудові рівнянь багатовимірного еліпсоїду прогнозування для нормалізованих даних, як правило, використовують дві тестові статистики, які базуються на квантилях або розподілу  $\chi^2$ -квадрат, або  $F$  розподілу.

4) Застосування для визначення викидів у багатовимірних даних з програмних метрик методу на основі квадрату відстані Махаланобісу для нормалізованих даних обумовлено, як правило, ще необхідністю побудови

нелінійних регресій для рішення задач визначення складності ООП та якості ПЗ, а також раннього оцінювання кількості рядків коду ПЗ. А як відомо, одним із ефективних методів побудови нелінійних регресій є метод, що базується на нормалізуючих перетвореннях, у тому числі і багатовимірних, застосування яких не потребує простого перебору. Крім того, побудова нелінійних регресій на основі нормалізуючих перетвореннях вирішує проблему визначення довірчих інтервалів та інтервалів прогнозування нелінійних регресій, а також проблему пошуку розподілу відхилень значень залежної випадкової величини від лінії регресії. Довірчі інтервали нелінійних регресій дозволяють отримувати інтервальні оцінки вибіркового середнього залежної випадкової величини, а інтервали прогнозування визначати викиди в процесі побудови нелінійних регресійних моделей.

5) На основі здійсненого аналізу існуючих методів і моделей для обробки інформації з програмних метрик та можливості використання існуючих моделей для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ саме для Web застосунків, що розроблені з використанням РНР-фреймворків, виконано обґрунтування необхідності проведення досліджень за обраною темою. Удосконалювати існуючі математичні моделі для обробки інформації з програмних метрик та відомі за їх допомогою рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ доцільно саме на основі багатовимірних нормалізуючих перетворень, які дозволяють враховувати кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

Результати досліджень автора, які наведені у поточному розділі, опубліковано в [31-34, 70-72, 78-85].

## **РОЗДІЛ 2 БАГАТОВИМІРНІ НОРМАЛІЗУЮЧІ ПЕРЕТВОРЕННЯ ДЛЯ НОРМАЛІЗАЦІЇ НАБОРІВ ДАНИХ З МЕТРИК WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ PHP-ФРЕЙМВОРКІВ**

У другому розділі дисертації розглянуто існуючі взаємо-зворотні багатовимірні нормалізуючі перетворення та методи оцінювання їх параметрів, здійснено вибір багатовимірних нормалізуючих перетворень для нормалізації двовимірних, тривимірних та чотиривимірних наборів даних з програмних метрик Web застосунків, що розробляються з використанням PHP-фреймворків.

### **2.1 Існуючі взаємо-зворотні багатовимірні нормалізуючі перетворення**

В наш час відомі дві групи взаємо-зворотних нормалізуючих перетворень: одновимірні та багатовимірні. Одновимірні взаємо-зворотні нормалізуючі перетворення дозволяють перетворити випадкову величину, розподіл якої відхиляється від нормального, у гаусівську або величину, яка наближено може вважатися гаусівською. Прикладами одновимірних взаємо-зворотних нормалізуючих перетворень є перетворення у вигляді логарифму (як правило, десятичного або натурального), степеневе перетворення [111], перетворення Бокса – Кокса [112] та Джонсона [113, 114]. Степеневе перетворення та перетворення Бокса – Кокса є однопараметричними на відміну від перетворення Джонсона, яке містить чотири параметри, що дозволяє здійснювати нормалізацію даних не тільки з унімодальним розподілом, а і з бімодальним та U-образним розподілами.

Багатовимірні взаємо-зворотні нормалізуючі перетворення дозволяють перетворити систему випадкових величин, розподіл яких відхиляється від нормального, у систему гаусівських випадкових величин або систему випадкових величин, які наближено можуть вважатися гаусівськими. При цьому у разі використання взаємо-зворотних багатовимірних перетворень для нормалізації багатовимірних даних на відміну від одновимірних перетворень враховується кореляція між випадковими величинами. На сьогодні найбільш відомими багатовимірними взаємо-зворотними нормалізуючими перетвореннями є багатовимірні перетворення Бокса – Кокса [97] та Джонсона [115, 116]. Перетворення Джонсона є більш складними у порівнянні з перетворення Бокса – Кокса. Але на відміну від перетворення Бокса – Кокса перетворення Джонсона дозволяють здійснювати нормалізацію даних не тільки з додатними значеннями, а і з від’ємними, що особливо характерно для нормалізації випадкових процесів [117]. Програмні метрики, як правило, мають додатні значення, що дозволяє для їх нормалізації використовувати перетворення Бокса – Кокса. Тому далі ми перевіримо можливість використання багатовимірного перетворення Бокса – Кокса для різних комбінацій програмних метрик вебзастосунків, розроблених із застосуванням РНР фреймворків.

На відміну від одновимірного перетворення Бокса – Кокса, де ми шукаємо оцінку лише одного параметру  $\lambda$  для однієї змінної, багатовимірне перетворення Бокса – Кокса працює з вектором параметрів  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$ . Кожна компонента  $Z_j$  гаусівського вектора  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_m)^T$  визначається як

$$Z_j = \begin{cases} (X_j^{\lambda_j} - 1)/\lambda_j, & \text{if } \lambda_j \neq 0; \\ \ln(X_j), & \text{if } \lambda_j = 0. \end{cases} \quad (2.1)$$



В (2.1)  $X_j$  – це компонента вектора  $\mathbf{X} = (X_1, X_2, \dots, X_m)^T$ , розподіл якого відхиляється від гаусівського,  $j = 1, 2, \dots, m$ .

Для оцінювання параметрів багатовимірного перетворення Бокса – Кокса рекомендується використовувати метод максимальної правдоподібності (ММП) [97] шляхом максимізації логарифмічної функції правдоподібності, яку можна записати наступним чином:

$$l(\mathbf{X}, \lambda) = \sum_{j=1}^k (\lambda_j - 1) \sum_{i=1}^N \ln(X_{ji}) - \frac{N}{2} \ln[\det(\mathbf{S}_N)]. \quad (2.2)$$

Зауважимо, що розподіл нормалізованого вектору  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_m)^T$  не повинен відхилятися від багатовимірного розподілу Гауса (багатовимірного нормального розподілу)

$$f_{\mathbf{Z}}(\mathbf{Z}) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(\mathbf{Z} - \mathbf{m}_{\mathbf{Z}})^T \Sigma^{-1} (\mathbf{Z} - \mathbf{m}_{\mathbf{Z}})}, \quad (2.3)$$

де  $|\Sigma| = \det \Sigma$  – це детермінант коваріаційної матриці  $\Sigma$ , відомий також як узагальнена дисперсія (generalized variance).

Зазначимо, що в (2.3) замість матриці  $\Sigma$  та вектору  $\mathbf{m}_{\mathbf{Z}}$  в подальшому ми використовуємо їх оцінки:  $\mathbf{S}_N$  та  $\bar{\mathbf{Z}}$ , відповідно.

Для перевірки відхилення розподілу багатовимірних даних від нормального, як правило, використовують міри Мардіа (Mardia) [120] багатовимірної асиметрії та ексцесу, оскільки вони найчастіше включені в

різноманітні програмні пакети. Мардія визначив багатовимірну асиметрію та багатовимірний ексцес, відповідно, як

$$\beta_{1,k} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left[ (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}_N^{-1} (\mathbf{x}_j - \bar{\mathbf{x}}) \right]^3; \quad (2.4)$$

$$\beta_{2,k} = \frac{1}{N} \sum_{i=1}^N \left[ (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}_N^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \right]^2. \quad (2.5)$$

Обидва показники (2.4) та (2.5) мають індекс  $k$ , тому вони специфічні для набору змінних  $k$ . Хоча слід зазначити, що дуже часто для спрощення індекс  $k$  не ставлять.

Очікувана асиметрія Мардія дорівнює 0 для багатовимірного нормального розподілу, а вищі значення вказують на більш серйозне відхилення від нормального закону. Очікуваний ексцес Мардія становить  $k(k+2)$  для багатовимірного нормального розподілу  $k$  змінних. Для двовимірного нормального розподілу ексцес Мардія дорівнює 8.

Для  $\beta_{1,k}$  зазвичай використовують тестову статистику

$$\frac{N}{6} \beta_{1,k}, \quad (2.6)$$

яка має наближений розподіл  $\chi^2$  з  $k(k+1)(k+2)/6$  ступенями свободи та рівнем значущості  $\alpha$ .

Для  $\beta_{2,k}$  у якості тестової статистики можна використовувати саму оцінку  $\beta_{2,k}$ , яку порівнюють з  $1-\alpha$  квантилем нормального закону розподілу  $z_{2,k}$  з

математичним сподіванням  $k(k+2)$  та дисперсією  $8k(k+2)/N$ .

Щоб перевірити відхилення розподілу від нормального, потрібно порівняти значення статистики (2.6) з квантилем  $\chi^2_{k(k+1)(k+2)/6, \alpha}$ , а  $\beta_{2,k}$  з  $z_{2,k}$ . Якщо значення статистики (2.6) більше за квантиль  $\chi^2_{k(k+1)(k+2)/6, \alpha}$  або, якщо значення  $\beta_{2,k}$  більше за  $z_{2,k}$ , то багатовимірний розподіл даних не є нормальним. При перевірці відхилення розподілу від нормального рекомендують брати  $\alpha$ , що дорівнює 0,005. Підкреслимо, як і в одновимірному випадку, незначущість тестових статистик для асиметрії та ексцесу не обов'язково означає нормальність.

## **2.2 Вибір багатовимірних перетворень для нормалізації наборів даних з метрик Web застосунків, що розробляються з використанням РНР-фреймворків**

Для рішення вказаних у першому розділі задач треба здійснити вибір багатовимірних перетворень для нормалізації наборів даних різної розмірності (від двовірної до чотиривірної) з метрик Web застосунків, що розробляються з використанням РНР-фреймворків.

За допомогою інструменту PhpMetrics (<https://phpmetrics.org/>) були отримані дані з метрик RFC, CBO, WMC, DIT, LCOM, NOC, KLOC, кількість класів та кількість методів на клас для 125 вебзастосунків, що створені за допомогою таких популярних РНР фреймворків, як CakePHP, Laravel, Symfony, CodeIgniter та Yii. Зазначені вебзастосунки розміщені на платформі GitHub (<https://github.com>). Значення метрик цих вебзастосунків наведені в додатку В.

Було перевірено тривимірний розподіл отриманих даних з метрик RFC, CBO та WMC на нормальність за допомогою тесту Мардіа (Mardia's test), який

базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл даних відхиляється від гаусівського. Зокрема, на це вказує те, що оцінка тривимірного ексцесу, яка дорівнює 38,15, більша за 17,52 – значення відповідного квантиля нормального розподілу.

Тому в подальшому для визначення наявності тривимірних викидів у даних з метрик RFC, CBO та WMC вебзастосунків, що створені за допомогою PHP фреймворків, ми обрали тривимірне перетворення Бокса – Кокса та використали метод на основі квадрату відстані Махаланобіса для нормалізованих даних [40]. Оцінювання параметрів тривимірного перетворення Бокса – Кокса з компонентами (2.1) виконувалося за методом максимальної правдоподібності. Їх оцінки для 125 тривимірних точок даних є такими:  $\hat{\lambda}_{RFC}$ ,  $\hat{\lambda}_{CBO}$  та  $\hat{\lambda}_{WMC}$  дорівнюють відповідно -0,2406993, 0,4381357 та -0,3186261. Було перевірено тривимірний розподіл нормалізованих даних з метрик RFC, CBO та WMC на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл нормалізованих даних не відхиляється від гаусівського.

Було перевірено тривимірний розподіл отриманих даних з метрик WMC, DIT та NOC на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл даних відхиляється від гаусівського. Зокрема, на це вказує те, що оцінка тривимірного ексцесу, яка дорівнює 47,98, більша за 17,52 – значення відповідного квантиля нормального розподілу.

Тому в подальшому для визначення наявності тривимірних викидів у даних з метрик WMC, DIT та NOC вебзастосунків, що створені за допомогою PHP фреймворків, ми обрали тривимірне перетворення Бокса – Кокса та використали метод на основі квадрату відстані Махаланобіса для нормалізованих

даних [40]. Оцінювання параметрів тривимірного перетворення Бокса – Кокса з компонентами (2.1) виконувалося за методом максимальної правдоподібності для 124 тривимірних точок даних тому, що для одного застосунку (CiiMS, повнофункціональна платформа для блогів та система управління контентом, побудована на Yii, <https://github.com/ciims/CiiMS>) значення NOC дорівнює нулю. Оцінки параметрів тривимірного перетворення Бокса – Кокса з компонентами (2.1) для 124 тривимірних точок даних є такими:  $\hat{\lambda}_{WMC}$ ,  $\hat{\lambda}_{DIT}$  та  $\hat{\lambda}_{NOC}$  дорівнюють відповідно -0,2077007, -1,1294780 та 0,4591524. Було перевірено тривимірний розподіл нормалізованих даних з метрик WMC, DIT та NOC на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл нормалізованих даних не відхиляється від гаусівського.

Було перевірено тривимірний розподіл отриманих даних з метрик WMC, RFC та LCOM на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл даних відхиляється від гаусівського. Зокрема, на це вказує те, що оцінка тривимірного ексцесу, яка дорівнює 48,44, більша за 17,52 – значення відповідного квантиля нормального розподілу.

Тому в подальшому для визначення наявності тривимірних викидів у даних з метрик WMC, RFC та LCOM вебзастосунків, що створені за допомогою РНР фреймворків, ми обрали тривимірне перетворення Бокса – Кокса та використали метод на основі квадрату відстані Махаланобіса для нормалізованих даних [40]. Оцінювання параметрів тривимірного перетворення Бокса – Кокса з компонентами (2.1) виконувалося за методом максимальної правдоподібності. Їх оцінки для 125 тривимірних точок даних є такими:  $\hat{\lambda}_{WMC}$ ,  $\hat{\lambda}_{RFC}$  та  $\hat{\lambda}_{LCOM}$  дорівнюють відповідно -0,2919246, -0,2217689 та -0,3524120. Було перевірено

тривимірний розподіл нормалізованих даних з метрик WMC, RFC та LCOM на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що тривимірний розподіл нормалізованих даних не відхиляється від гаусівського.

Було перевірено двовимірний розподіл отриманих даних з метрик RFC та CBO на нормальність за допомогою тесту Мардіа, який базується на багатовимірних асиметрії (2.4) та ексцесі (2.5). Тест показав, що двовимірний розподіл даних відхиляється від гаусівського. Зокрема, на це вказує те, що оцінка двовимірного ексцесу, яка дорівнює 27,35, більша за 9,84 – значення відповідного квантиля нормального розподілу з математичним сподіванням 8 та дисперсією 0,512. При цьому рівень значимості брався 0,005.

Тому в подальшому для визначення наявності двовимірних викидів у даних з метрик CBO та RFC вебзастосунків, що створені за допомогою PHP фреймворків, ми обрали двовимірне перетворення Бокса – Кокса та використали метод на основі квадрату відстані Махаланобіса для нормалізованих даних [40]. Оцінювання параметрів двовимірного перетворення Бокса – Кокса з компонентами (2.1) виконувалося за методом максимальної правдоподібності. Їх оцінки для 125 двовимірних точок даних є такими:  $\hat{\lambda}_1 = 0,431235$  та  $\hat{\lambda}_2 = -0,166319$  для CBO та RFC відповідно. Було перевірено двовимірний розподіл нормалізованих даних з метрик CBO та RFC на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що двовимірний розподіл нормалізованих даних не відхиляється від гаусівського.

Було перевірено чотиривимірний розподіл отриманих даних з метрик KLOC, кількість класів, кількість методів на клас та DIT на нормальність за допомогою тесту Мардіа, який базується на багатовимірних асиметрії (2.4) та

ексцесі (2.5). У подальшому при рішенні задачі з раннього оцінювання кількості рядків коду метрики KLOC, кількість класів, кількість методів на клас та DIT будемо позначати як  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  відповідно. Тест показав, що чотиривимірний розподіл даних відхиляється від гаусівського. Зокрема, на це вказує те, що оцінка чотиривимірного ексцесу, яка дорівнює 100,35, більша за 27,19 – значення відповідного квантиля нормального розподілу з математичним сподіванням 24 та дисперсією 1,536. При цьому рівень значимості брався 0,005.

Тому в подальшому для визначення наявності чотиривимірних викидів у даних з метрик KLOC, кількість класів, кількість методів на клас та DIT вебзастосунків, що створені за допомогою PHP фреймворків, ми обрали чотиривимірне перетворення Бокса – Кокса та використали метод на основі квадрату відстані Махаланобіса для нормалізованих даних [40]. Оцінювання параметрів чотиривимірного перетворення Бокса – Кокса з компонентами (2.1) виконувалося за методом максимальної правдоподібності. Їх оцінки для 125 чотиривимірних точок даних є такими:  $\hat{\lambda}_Y = -0,00477543$ ,  $\hat{\lambda}_1 = -0,0340235$ ,  $\hat{\lambda}_2 = -0,1701853$  та  $\hat{\lambda}_3 = -1,2910924$  для  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  відповідно. Було перевірено чотиривимірний розподіл нормалізованих даних з метрик  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  на нормальність за допомогою тесту Мардіа, який базується на багатовимірних (2.4) та ексцесі (2.5). Тест показав, що чотиривимірний розподіл нормалізованих даних не відхиляється від гаусівського.

Отже, як показали результати розрахунків, для нормалізації багатовимірних даних з метрик RFC, CBO, WMC, DIT, LCOM, NOC, KLOC, кількість класів та кількість методів на клас для 125 вебзастосунків, що створені за допомогою таких популярних PHP фреймворків, як CakePHP, Laravel, Symfony, CodeIgniter та Yii, може бути застосоване багатовимірне перетворення Бокса – Кокса.

На думку автора, застосування зазначених багатовимірних нормалізуючих перетворень дозволить покращити якість результатів рішення вказаних у першому розділі задач подібно трансферному навчанню (Transfer Learning) [121-125] – методу машинного навчання, в якому знання (нова інформація), отримані в результаті виконання одного завдання або набору даних, використовуються для покращення якості моделі в іншому пов'язаному завданні або іншому наборі даних.

### **2.3 Висновки за розділом 2**

У даному розділі дисертації розглянуто існуючі взаємо-зворотні багатовимірні нормалізуючі перетворення та методи оцінювання їх параметрів, здійснено вибір багатовимірних нормалізуючих перетворень для нормалізації двовимірних, тривимірних та чотиривимірних наборів даних з програмних метрик Web застосунків, що розробляються з використанням PHP-фреймворків.

1) Розглянуто відомі взаємо-зворотні нормалізуючі перетворення, які дозволяють перетворити випадкові величини з негаусівським розподілом у такі, що мають гаусівський розподіл або наближений до нього. Зазначено, що на сьогодні існують два відомі взаємо-зворотні багатовимірні нормалізуючі перетворення – це перетворення Бокса – Кокса та Джонсона, які на відміну від одновимірних перетворень дозволяють враховувати кореляцію між змінними. Багатовимірне перетворення Бокса – Кокса у порівнянні з багатовимірним перетворенням Джонсона має лише один параметр на кожен змінну замість чотирьох. Що спрощує нормалізацію багатовимірних даних.

2) Розглянуто багатовимірне перетворення Бокса – Кокса, яке дозволяє враховувати кореляції між випадковими величинами. Для оцінювання параметрів



багатовимірного перетворення Бокса – Кокса пропонується використовувати метод максимальної правдоподібності.

З) Здійснено вибір перетворень для нормалізації двовимірних, тривимірних та чотиривимірних наборів даних з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків. У якості нормалізуючих перетворень були обрані двовимірні, тривимірні та чотиривимірні перетворення Бокса – Кокса, після застосування яких розподіли відповідних нормалізованих даних не відхиляються від нормального згідно критерію Мардія.

Результати досліджень автора, які наведені у поточному розділі, опубліковано в [31-34, 70-72, 78-85].

### **РОЗДІЛ 3 ПОБУДОВА МАТЕМАТИЧНИХ МОДЕЛЕЙ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ З МЕТРИК WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ РНР-ФРЕЙМВОРКІВ, НА ОСНОВІ БАГАТОВИМІРНИХ НОРМАЛІЗУЮЧИХ ПЕРЕТВОРЕНЬ**

У третьому розділі дисертації побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса; удосконалено модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса; удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення ширин довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса; удосконалено математичні моделі та рівняння для визначення ширин довірчого інтервалу та інтервалу прогнозування трифакторних нелінійних регресії кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT

на основі чотиривимірною нормалізуючого перетворення; здійснено порівняння результатів за різними моделями, у тому числі моделями машинного навчання.

### **3.1 Математична модель для обробки інформації з програмних метрик RFC, CBO і WMC для рішення задачі з визначення якості Web застосунків**

Для побудови математичної моделі для обробки інформації з програмних метрик RFC, CBO і WMC для рішення задачі з визначення якості Web застосунків були використані їх дані, що наведені у додатку В. Згідно з [19] рішення задачі з визначення якості програмних систем потребує побудови тривимірного еліпсоїду прогнозування для нормалізованих метрик RFC, CBO і WMC, а також рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій метрик RFC, CBO і WMC.

Рівняння тривимірного еліпсоїду прогнозування для нормалізованих метрик RFC, CBO і WMC будуємо подібно [22, 40]. Воно має наступний вигляд:

$$(\mathbf{Z} - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z} - \bar{\mathbf{Z}}) = \frac{3(N^2 - 1)}{N(N - 3)} F_{3, N-3, \alpha}. \quad (3.1)$$

В (3.1) всі позначення як і в (1.6) лише з тою різницею, що замість квантилю  $\chi_{m, \alpha}^2$  в правій частині застосовується тестова статистика на основі квантиля  $F$ -розподілу  $F_{3, N-3, \alpha}$  з 3 і  $N - 3$  ступенями свободи та рівнем значущості  $\alpha$ , який дорівнює 0,005.

За допомогою (3.1) визначається тривимірна точка даних нормалізованих метрик RFC, CBO і WMC, яка є викидом. Якщо для цієї точки квадрат відстані Махаланобіса більше за праву частину (3.1), то вона є тривимірним викидом.

Регресійні моделі, рівняння границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій метрик RFC, CBO і WMC ми будуємо подібно (1.8) та (1.9). Рівняння границь довірчих інтервалів трьох нелінійних регресій метрик RFC, CBO і WMC на основі перетворення Бокса – Кокса є такими:

$$Y_{UB} = \left[ \hat{\lambda}_Y \left( \hat{Z}_Y + t_{\alpha/2, v} S_{Z_Y} \left\{ \frac{1}{N} + (\mathbf{z}_X^+)^T \mathbf{S}_Z^{-1} (\mathbf{z}_X^+) \right\}^{1/2} \right) + 1 \right]^{1/\hat{\lambda}_Y}; \quad (3.2)$$

$$Y_{LB} = \left[ \hat{\lambda}_Y \left( \hat{Z}_Y - t_{\alpha/2, v} S_{Z_Y} \left\{ \frac{1}{N} + (\mathbf{z}_X^+)^T \mathbf{S}_Z^{-1} (\mathbf{z}_X^+) \right\}^{1/2} \right) + 1 \right]^{1/\hat{\lambda}_Y}. \quad (3.3)$$

Рівняння (3.2) задає верхні границі довірчих інтервалів трьох нелінійних регресій метрик RFC, CBO і WMC, а рівняння (3.3) – їх нижні границі довірчих інтервалів. В (3.2) та (3.3) використані такі ж позначення що і в (1.9).

Рівняння границь інтервалів прогнозування трьох нелінійних регресій метрик RFC, CBO і WMC нелінійної регресії будується аналогічно (3.2) та (3.3) з тією лише різницею, що до суми у фігурних дужках (3.2) та (3.3) потрібно додати ще 1.

Для побудови еліпсоїду прогнозування для нормалізованих метрик RFC, CBO і WMC (3.1), регресійних моделей у формі (1.8) та рівнянь границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій метрик RFC, CBO і WMC ми використали дані з 125 вебзастосунків, розміщених на GitHub для

п'яти відомих фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Описова статистика цього набору даних наведена в табл.1.2, а самі дані у додатку В. Після видалення викидів з цих даних залишилися дані з 110 застосунків, описова статистика з яких наведена в таблиці 3.1.

Таблиця 3.1 – Описова статистика набору даних із 110 застосунків

| Назва метрики     | Min  | Max   | Середнє | СКВ     |
|-------------------|------|-------|---------|---------|
| Number of classes | 25   | 7743  | 491,600 | 946,931 |
| RFC               | 3,00 | 20,98 | 10,413  | 3,757   |
| CBO               | 1,68 | 8,72  | 4,382   | 1,593   |
| WMC               | 1,75 | 67,83 | 14,301  | 14,568  |

Щоб уникнути проблеми оверфітінгу моделі, ми розділили вибірку даних на два набори: навчальний (60 точок даних) і тестовий (50 точок даних). Таблиця 3.2 містить описові статистики навчального, а таблиця 3.3 – тестового набору даних.

Таблиця 3.2 – Описова статистика навчального набору даних (60 точок даних)

| Назва метрики     | Min  | Max   | Середнє | СКВ      |
|-------------------|------|-------|---------|----------|
| Number of classes | 32   | 7743  | 655,467 | 1233,763 |
| RFC               | 3,00 | 20,98 | 10,518  | 4,468    |
| CBO               | 1,68 | 8,72  | 4,506   | 1,792    |
| WMC               | 1,75 | 67,83 | 15,849  | 16,733   |

Таблиця 3.3 – Описова статистика тестового набору даних (50 точок даних)

| Назва метрики     | Min  | Max   | Середнє | СКВ     |
|-------------------|------|-------|---------|---------|
| Number of classes | 25   | 1496  | 282,960 | 283,544 |
| RFC               | 6,31 | 16,47 | 10,287  | 2,708   |
| CBO               | 2,00 | 6,59  | 4,233   | 1,318   |
| WMC               | 2,73 | 43,73 | 12,444  | 11,265  |

Оцінки параметрів трьох побудованих нелінійних регресійних моделей для метрик RFC, CBO та WMC у формі (1.8) наведені в таблиці 3.4.

Таблиця 3.4 – Оцінки параметрів нелінійних регресійних моделей

| No | Y   | $\hat{b}_0$ | $\hat{b}_1$ | $\hat{b}_2$ | $\sigma_\varepsilon$ | $\hat{\lambda}_Y$ |
|----|-----|-------------|-------------|-------------|----------------------|-------------------|
| 1  | RFC | 0,61548     | -0,00200    | 1,29779     | 0,1369               | 0,04570           |
| 2  | CBO | 1,43051     | -0,05220    | 0,47253     | 0,6991               | 0,38434           |
| 3  | WMC | -0,34586    | 0,70983     | 0,00992     | 0,1013               | -0,43498          |

Для розрахунку значень квадрата відстані Махаланобіса в (3.1) за (1.3) обернена матриця до матриці (1.4) є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 99,164 & -70,390 & -0,9834 \\ -70,390 & 54,238 & 0,1086 \\ -0,9834 & 0,1086 & 2,081 \end{pmatrix}, \quad (3.4)$$

а вибіркові середні  $\bar{Z}_{RFC}$ ,  $\bar{Z}_{CBO}$  та  $\bar{Z}_{WMC}$  дорівнюють 2,384, 1,951 та 1,366 відповідно. Квантиль  $F_{3,N-3,0.005}$  дорівнює 4,754 для  $N=60$ .

Щоб визначити границі довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій метрик RFC, CBO і WMC за (3.2) та (3.3), обернені матриці до матриці (1.10) є відповідно такими:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,03468 & -0,01404 \\ -0,01404 & 0,13021 \end{pmatrix}, \quad (3.5)$$

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,90387 & -1,17231 \\ -1,17231 & 1,64499 \end{pmatrix}, \quad (3.6)$$

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,07122 & -0,00982 \\ -0,00982 & 0,03452 \end{pmatrix}, \quad (3.7)$$

а значення  $S_{Z_Y}$  дорівнюють 0,1393, 0,7112 та 0,1030 відповідно.

У всіх випадках обчислення границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій для метрик RFC, CBO та WMC необхідно використовувати в (1.9)  $t_{0,05/2,57} = 2,002$ ,  $N=60$  та  $v = 57$ .

Щоб оцінити прогностичну точність нелінійних регресійних моделей для метрик RFC, CBO та WMC у формі (1.8), ми використали стандартні показники MMRE та PRED(0,25), прийнятні граничні значення яких – не більше 0,25 і не менше 0,75 відповідно.

Середня величина відносної похибки MMRE визначається як

$$MMRE = \frac{1}{N} \sum_{i=1}^N MRE_i,$$

де  $MRE_i$  – це  $i$ -те значення величини відносної похибки, яке обчислюється як

$$MRE_i = \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|.$$

Значення  $PRED(0,25)$  визначається як

$$PRED(0,25) = \frac{1}{N} \sum_{i=1}^N \begin{cases} 1 & \text{if } MRE_i \leq 0,25 \\ 0 & \text{otherwise.} \end{cases}.$$

Для моделей у формі (1.8) із оцінками параметрів з таблиці 3.4 значення MMRE та  $PRED(0,25)$  для навчального і тестового наборів даних наведені в таблиці 3.5.

Таблиця 3.5 – Значення MMRE та  $PRED(0,25)$  для моделей для метрик RFC, CBO та WMC у формі (1.8)

| No | Y   | навчальний набір даних |              | тестовий набір даних |              |
|----|-----|------------------------|--------------|----------------------|--------------|
|    |     | MMRE                   | $PRED(0,25)$ | MMRE                 | $PRED(0,25)$ |
| 1  | RFC | 0,1040                 | 0,9833       | 0,1386               | 0,8800       |
| 2  | CBO | 0,3729                 | 0,4333       | 0,3181               | 0,6400       |
| 3  | WMC | 0,2563                 | 0,5333       | 0,3295               | 0,4000       |

Ці значення свідчать про задовільну якість моделей у формі (1.8). Крім того, ці значення вказують на прийнятну якість моделі (1.8) лише для метрики RFC. Останнє можна пояснити тим, що поділ всього набору даних на два



(навчальний та тестовий) проводився з урахуванням побудови моделі (1.8) лише для метрики RFC.

Для порівняння побудованих моделей з добре відомим алгоритмом машинного навчання, ми використали регресію опорних векторів (support vector regression, SVR) [126, 127], яка є розширенням машини опорних векторів (support vector machine, SVM), що застосовується як для класифікації, так і для регресії [128-130]. Ми побудували три лінійні SVR для метрик RFC, CBO та WMC на основі того ж самого навчального набору, який використовувався для побудови моделей для метрик RFC, CBO та WMC у формі (1.8). При чому всі три моделі SVR були побудовані в залежності від тих же самих метрик, що і моделі для метрик RFC, CBO та WMC у формі (1.8). Модель SVR для метрики RFC залежить від метрик CBO та WMC, модель SVR для метрики CBO залежить від метрик RFC та WMC, а модель SVR для метрики WMC залежить від метрик RFC та CBO. Для трьох лінійних SVR для метрик RFC, CBO та WMC обчислені значення MMRE та PRED(0,25) наведено в таблиці 3.6.

Таблиця 3.6 – Значення MMRE та PRED(0,25) для SVR для метрик RFC, CBO та WMC

| No | Y   | навчальний набір даних |            | тестовий набір даних |            |
|----|-----|------------------------|------------|----------------------|------------|
|    |     | MMRE                   | PRED(0,25) | MMRE                 | PRED(0,25) |
| 1  | RFC | 0,1674                 | 0,7833     | 0,1329               | 0,5000     |
| 2  | CBO | 0,3580                 | 0,5333     | 0,3884               | 0,3600     |
| 3  | WMC | 0,4851                 | 0,4500     | 0,5870               | 0,2200     |

Дані табл. 3.6 свідчать про задовільну якість лише моделі SVR для метрики RFC для тренувального набору даних. При цьому значення MMRE та PRED(0,25) цієї SVR гірші відповідно на 61% та 20% у порівнянні з побудованою

моделлю виду (1.8) для метрики RFC. Для тестового набору даних SVR для метрики RFC має незадовільне значення  $PRED(0,25)$ . Лінійні SVR для метрик CBO та WMC мають погану якість за всіма показниками.

Таким чином, було удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC, CBO і WMC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків.

### **3.2 Математичні моделі для обробки інформації з програмних метрик Чидамбера та Кемерера для рішення задачі з визначення складності ООП**

Для побудови математичних моделей для обробки інформації з програмних метрик Чидамбера та Кемерера для рішення задачі з визначення складності ООП Web застосунків, що розробляються з використанням PHP-фреймворків, були використані їх дані, які наведені у додатку В. Було отримані три математичні моделі для трьох груп програмних метрик Чидамбера та Кемерера (відповідно WMC, DIT і NOC; WMC, RFC і LCOM та RFC і CBO).

Нагадаємо, що метрики WMC, DIT та NOC визначають складність ООП завдяки ідентифікації класів на першому кроці ООП за Бучем [89]. Метрики WMC, RFC та LCOM відносяться до другого кроку (семантики класів) ООП. Згідно з [20] для визначення складності ООП ПЗ на третьому кроці через зв'язки між класами можуть бути використані дві програмні метрики: зв'язаність між об'єктами CBO та відгук класу RFC.

### **3.2.1 Математичні моделі для обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM**

Рішення задачі з визначення складності ООП ПЗ на першому та другому кроках ООП за Бучем потребує побудови тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення, зокрема перетворення Бокса – Кокса.

Рівняння тривимірного еліпсоїду прогнозування для нормалізованих метрик WMC, DIT і NOC будуємо як і (3.1) лише з тою різницею, що замість нормалізованих метрик RFC, CBO і WMC в (3.1) треба використовувати нормалізовані метрики WMC, DIT і NOC та інші параметри, які відповідають даним цих метрик.

Також у вказаному випадку за допомогою (3.1) визначається тривимірна точка даних нормалізованих метрик WMC, DIT і NOC, яка є викидом. Якщо для цієї точки квадрат відстані Махаланобіса більше за праву частину (3.1), то вона є тривимірним викидом.

Для побудови еліпсоїду прогнозування для нормалізованих метрик WMC, DIT та NOC виду (3.1) та трьох лінійних регресійних рівнянь для нормалізованих метрик WMC, DIT та NOC ми використали дані з 125 вебзастосунків, розміщених

на GitHub для п'яти відомих фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Описова статистика цього набору даних наведена в табл.1.2, а самі дані у додатку В. Після видалення викидів з цих даних залишилися дані з 121 застосунку, описова статистика з яких наведена в таблиці 3.7.

Таблиця 3.7 – Описова статистика набору даних із 121 застосунків

| Назва метрики     | Min  | Max   | Середнє | СКВ     |
|-------------------|------|-------|---------|---------|
| Number of classes | 25   | 7512  | 525,554 | 920,147 |
| WMC               | 1,55 | 67,84 | 14,797  | 13,755  |
| DIT               | 1,06 | 2,33  | 1,484   | 0,300   |
| NOC               | 0,32 | 0,95  | 0,685   | 0,141   |

Для розрахунку значень квадрата відстані Махаланобіса в (3.1) за (1.3) обернена матриця до матриці (1.4) є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 3,8768 & 8,3980 & -9,4109 \\ 8,3980 & 139,516 & -111,477 \\ -9,4109 & -111,477 & 194,605 \end{pmatrix}, \quad (3.8)$$

а вибіркові середні  $\bar{Z}_{WMC}$ ,  $\bar{Z}_{DIT}$  та  $\bar{Z}_{NOC}$  дорівнюють 1,869, 0,284 та -0,265 відповідно. Квантиль  $F_{3,N-3,0.005}$  дорівнює 4,501 для  $N=121$ .

Матрицю (3.8) ми використовуємо в рівнянні еліпсоїду прогнозування для нормалізованих метрик WMC, DIT та NOC виду (3.1).

Для визначення складності ООП завдяки ідентифікації класів на першому кроці ООП за метриками WMC, DIT та NOC також треба побудувати три лінійні регресійні рівняння для нормалізованих метрик WMC, DIT та NOC у вигляді

$$\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2, \quad (3.9)$$

де  $\hat{Z}_Y$  – це оцінка умовного математичного сподівання нормалізованої залежної змінної;  $Z_1$  та  $Z_2$  – це нормалізовані незалежні змінні (предиктори), значення яких отримують за допомогою тривимірного перетворення Бокса – Кокса.

Щоби записати лінійне регресійне рівняння для нормалізованої метрики WMC як (3.9), треба підставити  $\hat{Z}_{WMC}$ ,  $Z_{DIT}$  та  $Z_{NOC}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та  $Z_2$ , відповідно. Щоби записати лінійне регресійне рівняння для нормалізованої метрики DIT як (3.9), треба підставити  $\hat{Z}_{DIT}$ ,  $Z_{WMC}$  та  $Z_{NOC}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та  $Z_2$ , відповідно. Щоби записати лінійне регресійне рівняння для нормалізованої метрики NOC як (3.9), треба підставити  $\hat{Z}_{NOC}$ ,  $Z_{DIT}$  та  $Z_{WMC}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та  $Z_2$  відповідно. Оцінки параметрів цих лінійних регресійних рівнянь для нормалізованих метрик WMC, DIT і NOC та відповідні параметри тривимірного перетворення Бокса – Кокса  $\hat{\lambda}_{WMC}$ ,  $\hat{\lambda}_{DIT}$  та  $\hat{\lambda}_{NOC}$  наведені в таблиці 3.8.

Щоб оцінити прогностичну точність лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC у формі (3.9), ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Граничні значення MMRE та PRED(0,25) – не більше 0,25 і не менше 0,75 відповідно.

Вибірковий коефіцієнт детермінації  $R^2$  визначається як

$$R^2 = 1 - \frac{SS_E}{SS_T},$$

де  $SS_E$  – сума квадратів залишків (the residual sum of squares) або сума квадратів помилок (the error sum of squares),  $SS_E = \sum_{i=1}^N (Y_i - \hat{Y}_i)^2$ ;  $SS_T$  – загальна сума квадратів (the total sum of squares),  $SS_T = \sum_{i=1}^N (Y_i - \bar{Y})^2$ . Нагадаємо, що  $SS_T = SS_R + SS_E$ . Тут  $SS_R$  – сума квадратів регресії (the regression sum of squares),  $SS_R = \sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2$ .

Коефіцієнт детермінації  $R^2$  інтерпретується як частка мінливості у спостережуваній залежній змінній  $Y$ , що пояснюється моделлю лінійної регресії. Значення  $R^2$  вказує наскільки емпіричні дані підтверджують математичну модель або, чи краща наша модель за вибіркове середнє. З формули для  $R^2$  легко побачити, що  $0 \leq R^2 \leq 1$ . Велике значення  $R^2$  свідчить про те, що модель успішно пояснює мінливість  $Y$ . Коли  $R^2$  малий, це може свідчити про те, що потрібно знайти альтернативну модель, яка може врахувати більшу мінливість  $Y$ .

Для лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC у формі (3.9) значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.8.

Таблиця 3.8 – Оцінки параметрів лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC

| No | Метрики | $\hat{b}_0$ | $\hat{b}_1$ | $\hat{b}_2$ | $\hat{\lambda}_j$ | $R^2$  | MMRE   | PRED   |
|----|---------|-------------|-------------|-------------|-------------------|--------|--------|--------|
| 1  | WMC     | 3,12787     | -2,16620    | 2,42748     | -0,17788          | 0,1482 | 0,2731 | 0,5620 |
| 2  | DIT     | 0,60835     | -0,06019    | 0,79903     | -1,27219          | 0,4767 | 0,3306 | 0,6446 |
| 3  | NOC     | -0,51807    | 0,57284     | 0,04836     | 1,81164           | 0,4688 | 0,2814 | 0,6860 |

Дані показників з табл. 3.8 свідчать про задовільну якість зазначених рівнянь лише для оцінювання умовного математичного сподівання відповідної нормалізованої залежної змінної.

Для порівняння побудованих лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC у формі (3.9) з іншими добре відомими моделями, які також використовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127], яка є розширенням машини опорних векторів (SVM), що застосовується як для класифікації, так і для регресії [128-130]. Ми побудували три лінійні SVR для нормалізованих метрики WMC, DIT і NOC на основі того ж самого набору, який використовувався для побудови лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC у формі (3.9). При чому всі три моделі SVR були побудовані в залежності від тих же самих метрик, що і моделі для нормалізованих метрик WMC, DIT і NOC у формі (3.9). Модель SVR для нормалізованої метрики WMC залежить від метрик DIT та NOC, модель SVR для метрики DIT залежить від метрик WMC та NOC, а модель SVR для метрики NOC залежить від метрик DIT та WMC. Для трьох лінійних SVR для нормалізованих метрики WMC, DIT і NOC обчислені значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.9.

Таблиця 3.9 – Значення  $R^2$ , MMRE та PRED(0,25) для SVR для нормалізованих метрики WMC, DIT і NOC

| No | Метрики | $R^2$  | MMRE   | PRED(0,25) |
|----|---------|--------|--------|------------|
| 1  | WMC     | 0,1098 | 0,2607 | 0,5289     |
| 2  | DIT     | 0,4553 | 0,3484 | 0,6198     |
| 3  | NOC     | 0,4578 | 0,2902 | 0,6777     |

Дані показників з табл. 3.9 свідчать про задовільну якість зазначених SVR лише для визначення лінії регресії відповідної нормалізованої залежної змінної і не можуть бути використані для моделювання нормалізованих метрик WMC, DIT

і NOC як випадкових величин. Крім того, значення показників для SVR виявилися гіршими за відповідні значення для лінійних регресійних рівнянь для нормалізованих метрики WMC, DIT і NOC у формі (3.9). Виключенням є лише значення MMRE для SVR для нормалізованої метрики WMC, яке виявилось кращим на 4,75% у порівнянні з лінійним регресійним рівнянням для нормалізованої метрики WMC у формі (3.9). Натомість значення  $R^2$  для всіх трьох SVR для нормалізованих метрики WMC, DIT та NOC є гіршими відповідно на 35,0%, 4,7% та 2,4% у порівнянні з лінійними регресійними рівняннями для нормалізованих метрик WMC, DIT та NOC у формі (3.9). Враховуючи це, в подальшому ми використовуємо лінійні регресійні рівняння для нормалізованих метрик WMC, DIT та NOC у формі (3.9) як доповнення до рівняння еліпсоїду прогнозування для нормалізованих метрик WMC, DIT та NOC виду (3.1) для визначення складності ООП завдяки ідентифікації класів на першому кроці ООП за метриками WMC, DIT та NOC.

Таким чином, було розроблено математичну модель для обробки інформації з трьох програмних метрик WMC, DIT і NOC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді рівняння тривимірного еліпсоїду прогнозування та трьох лінійних регресійних рівнянь нормалізованих метрик WMC, DIT і NOC на основі тривимірного перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик WMC, DIT і NOC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП завдяки ідентифікації класів на першому кроці ООП за метриками WMC, DIT та NOC Web застосунків, що розробляються з використанням PHP-фреймворків.



Як раніше зазначалося, рішення задачі з визначення складності ООП ПЗ на другому кроці ООП потребує побудови рівняння тривимірного еліпсоїду прогнозування і рівнянь регресії для нормалізованих метрик WMC, RFC і LCOM на основі тривимірного нормалізуючого перетворення, зокрема перетворення Бокса – Кокса.

Рівняння тривимірного еліпсоїду прогнозування для нормалізованих метрик WMC, RFC і LCOM також будуємо як і (3.1) лише з тою різницею, що замість нормалізованих метрик RFC, CBO і WMC в (3.1) треба використовувати нормалізовані метрики WMC, RFC і LCOM та інші параметри, які відповідають даним цих метрик.

Також у вказаному випадку за допомогою (3.1) визначається тривимірна точка даних нормалізованих метрик WMC, RFC і LCOM, яка є викидом. Якщо для цієї точки квадрат відстані Махаланобіса більше за праву частину (3.1), то вона є тривимірним викидом.

Для побудови еліпсоїду прогнозування для нормалізованих метрик WMC, RFC і LCOM виду (3.1) та трьох лінійних регресійних рівнянь для нормалізованих метрик WMC, RFC і LCOM ми використали дані з 125 вебзастосунків, розміщених на GitHub для п'яти відомих фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Описова статистика цього набору даних наведена в табл.1.2, а самі дані у додатку В. Після видалення викидів з цих даних залишилися дані з 123 застосунків, описова статистика з яких наведена в таблиці 3.10.

Таблиця 3.10 – Описова статистика набору даних із 123 застосунків

| Назва метрики     | Min  | Max   | Середнє | СКВ     |
|-------------------|------|-------|---------|---------|
| Number of classes | 25   | 7743  | 472,398 | 920,147 |
| WMC               | 1,75 | 67,83 | 14,370  | 13,900  |
| RFC               | 3,00 | 20,98 | 10,512  | 3,886   |
| LCOM              | 1,11 | 5,78  | 2,182   | 0,726   |

Для розрахунку значень квадрата відстані Махаланобіса в (3.1) за (1.3) обернена матриця до матриці (1.4) є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 20,7727 & -17,6004 & -7,2952 \\ -17,6004 & 21,7157 & -1,2993 \\ -7,2952 & -1,2993 & 50,9796 \end{pmatrix}, \quad (3.10)$$

а вибіркові середні  $\bar{Z}_{WMC}$ ,  $\bar{Z}_{RFC}$  та  $\bar{Z}_{LCOM}$  дорівнюють 1,629, 2,431 та 0,547 відповідно. Квантиль  $F_{3,N-3,0.005}$  дорівнює 4,497 для  $N=123$ .

Матрицю (3.10) ми використовуємо в рівнянні еліпсоїду прогнозування для нормалізованих метрик WMC, RFC і LCOM виду (3.1).

Для визначення складності ООП завдяки семантиці класів на другому кроці ООП за метриками WMC, RFC і LCOM також треба побудувати три лінійні регресійні рівняння для нормалізованих метрик WMC, RFC і LCOM у вигляді (3.9).

Щоби записати лінійне регресійне рівняння для нормалізованої метрики WMC як (3.9), треба підставити  $\hat{Z}_{WMC}$ ,  $Z_{RFC}$  та  $Z_{LCOM}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та  $Z_2$ , відповідно. Щоби записати лінійне регресійне рівняння для нормалізованої метрики DIT як (3.9), треба підставити  $\hat{Z}_{RFC}$ ,  $Z_{WMC}$  та  $Z_{LCOM}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та

$Z_2$ , відповідно. Щоби записати лінійне регресійне рівняння для нормалізованої метрики NOC як (3.9), треба підставити  $\hat{Z}_{LCOM}$ ,  $Z_{WMC}$  та  $Z_{RFC}$  замість  $\hat{Z}_Y$ ,  $Z_1$  та  $Z_2$  відповідно. Оцінки параметрів цих лінійних регресійних рівнянь для нормалізованих метрик WMC, RFC і LCOM та відповідні параметри тривимірного перетворення Бокса – Кокса  $\hat{\lambda}_{WMC}$ ,  $\hat{\lambda}_{DIT}$  та  $\hat{\lambda}_{NOC}$  наведені в таблиці 3.11.

Щоб оцінити прогностичну точність лінійних регресійних рівнянь для нормалізованих метрики WMC, RFC і LCOM у формі (3.9), ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Для зазначених рівнянь значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.11. Дані показників з табл. 3.11 свідчать про прийнятну якість зазначених рівнянь для нормалізованих метрик WMC та RFC. Якість лінійного регресійного рівняння для нормалізованої метрики LCOM у формі (3.9) може вважатися задовільною тому, що значення MMRE та PRED(0,25) лише відповідно на 0,7% та 0,3% гірші за прийнятні граничні значення.

Таблиця 3.11 – Оцінки параметрів лінійних регресійних рівнянь для нормалізованих метрики WMC, RFC і LCOM

| No | Метрики | $\hat{b}_0$ | $\hat{b}_1$ | $\hat{b}_2$ | $\hat{\lambda}_j$ | $R^2$  | MMRE   | PRED   |
|----|---------|-------------|-------------|-------------|-------------------|--------|--------|--------|
| 1  | WMC     | -0,62303    | 0,84728     | 0,35119     | -0,28340          | 0,7526 | 0,1171 | 0,9268 |
| 2  | RFC     | 1,07810     | 0,81049     | 0,79903     | 0,05983           | 0,7399 | 0,0720 | 0,9756 |
| 3  | LCOM    | 0,25216     | 0,14310     | 0,02549     | -0,75631          | 0,2116 | 0,2517 | 0,7480 |

Для порівняння побудованих лінійних регресійних рівнянь для нормалізованих метрики WMC, RFC та LCOM у формі (3.9) з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми

використали регресію опорних векторів (SVR) [126, 127]. Ми побудували три лінійні SVR для нормалізованих метрики WMC, RFC та LCOM на основі того ж самого набору, який використовувався для побудови лінійних регресійних рівнянь для нормалізованих метрики WMC, RFC та LCOM у формі (3.9). При чому всі три моделі SVR були побудовані в залежності від тих же самих метрик, що і моделі для нормалізованих метрик WMC, RFC та LCOM у формі (3.9). Модель SVR для нормалізованої метрики WMC залежить від метрик RFC та LCOM, модель SVR для метрики RFC залежить від метрик WMC та LCOM, а модель SVR для метрики LCOM залежить від метрик WMC та RFC. Для трьох лінійних SVR для нормалізованих метрики WMC, RFC та LCOM обчислені значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.12.

Дані показників з табл. 3.12 свідчать про задовільну якість всіх трьох зазначених SVR.

Таблиця 3.12 – Значення  $R^2$ , MMRE та PRED(0,25) для SVR для нормалізованих метрики WMC, RFC і LCOM

| No | Метрики | $R^2$  | MMRE   | PRED(0,25) |
|----|---------|--------|--------|------------|
| 1  | WMC     | 0,7290 | 0,1078 | 0,9187     |
| 2  | RFC     | 0,7386 | 0,0714 | 0,9756     |
| 3  | LCOM    | 0,2047 | 0,2453 | 0,7480     |

Крім того, значення  $R^2$  для всіх трьох SVR виявилися трохи гіршими за відповідні значення для лінійних регресійних рівнянь у формі (3.9). Лише значення MMRE для всіх трьох SVR виявилися трохи кращими за відповідні значення для лінійних регресійних рівнянь у формі (3.9). Враховуючи це, в подальшому ми використовуємо лінійні регресійні рівняння для нормалізованих метрик WMC, RFC і LCOM у формі (3.9) як доповнення до рівняння еліпсоїду

прогнозування для нормалізованих метрик WMC, RFC і LCOM виду (3.1) для визначення складності ООП завдяки семантиці класів на другому кроці ООП за метриками WMC, RFC і LCOM.

Таким чином, було розроблено математичну модель для обробки інформації з трьох програмних метрик WMC, RFC і LCOM Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді рівняння тривимірного еліпсоїду прогнозування та трьох лінійних регресійних рівнянь для нормалізованих метрик WMC, RFC і LCOM на основі тривимірного перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик WMC, RFC і LCOM за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП завдяки семантиці класів на другому кроці ООП за метриками WMC, RFC і LCOM Web застосунків, що розробляються з використанням PHP-фреймворків.

### **3.2.2 Математична модель для обробки інформації з програмних метрик RFC і CVO**

Як раніше зазначалося, рішення задачі з визначення складності ООП ПЗ на третьому кроці ООП через зв'язки між класами потребує побудови рівняння еліпсу прогнозування та рівнянь регресії для нормалізованих метрик RFC та CVO на основі двовимірного нормалізуючого перетворення, зокрема перетворення Бокса – Кокса.

Рівняння еліпсу прогнозування для нормалізованих метрик RFC та CVO також будемо як і (1.7) лише з іншими параметрами. Також у нашому випадку

за допомогою (1.7) визначається двовимірною точкою даних нормалізованих метрик RFC та CBO, яка є викидом. Якщо для цієї точки квадрат відстані Махаланобіса більше за праву частину (1.7), то вона є двовимірним викидом.

Для побудови еліпсу прогнозування для нормалізованих метрик RFC та CBO виду (1.7) та двох лінійних регресійних рівнянь для нормалізованих метрик RFC та CBO ми використали дані з 125 вебзастосунків, розміщених на GitHub для п'яти відомих фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Описова статистика цього набору даних наведена в табл.1.2, а самі дані у додатку В.

В результаті одна точка даних виявилася двовимірним викидом – це застосунок Grocery CRUD (<https://github.com/scoumbourdis/grocery-crud>), який є бібліотекою PHP і фреймворку Codeigniter, який створює повнофункціональну систему CRUD без необхідності налаштування JavaScript або CSS. У порівнянні з іншими застосунками цей мав найменшу кількість класів (всього 11) та найбільше значення RFC на рівні застосунку (47,09). Зазначена точка даних була відкинута тому, що для неї квадрат відстані Махаланобіса для нормалізованих значень метрик CBO та RFC більше за квантиль розподілу Хі-квадрат ( $15,08 > 10,60$ ). В подальшому ми залишили дані з метрик CBO та RFC для 124 вебзастосунків, які знову були нормалізовані за допомогою двовимірного перетворення Бокса – Кокса. Описова статистика для даних з 124 веб-застосунків наведені в таблиці 3.13.

Для розрахунку значень квадрата відстані Махаланобіса в (1.7) за (1.3) обернена матриця до матриці (1.4) є такою:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 1,88082 & -0,17454 \\ -0,17454 & 3,46172 \end{pmatrix}. \quad (3.11)$$

а вибіркові середні  $\bar{Z}_{RFC}$  та  $\bar{Z}_{CBO}$  дорівнюють 2,029 та 2,785 відповідно. Квантиль  $F_{2,N-2,\alpha}$  дорівнює 5,535 для  $N=124$ .

Таблиця 3.13 – Описова статистика набору даних метрик CBO та RFC для 124 вебзастосунків

| Назва метрики     | Min  | Max   | Середнє | СКВ     |
|-------------------|------|-------|---------|---------|
| Number of classes | 25   | 7743  | 469,315 | 902,038 |
| RFC               | 3,00 | 20,98 | 10,525  | 3,873   |
| CBO               | 1,66 | 8,72  | 4,341   | 1,594   |

Матрицю (3.11) ми використовуємо в рівнянні еліпсу прогнозування для нормалізованих метрик RFC та CBO виду (1.7).

Для визначення складності ООП через зв'язки між класами на третьому кроці ООП за метриками RFC та CBO також треба побудувати два лінійні регресійні рівняння для нормалізованих метрик RFC та CBO у вигляді

$$\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1, \quad (3.12)$$

де  $\hat{Z}_Y$  – це оцінка умовного математичного сподівання нормалізованої залежної змінної;  $Z_1$  – це нормалізована незалежна змінна (предіктор), значення якої отримують за допомогою двовимірного перетворення Бокса – Кокса.

Щоби записати лінійне регресійне рівняння для нормалізованої метрики RFC як (3.12), треба підставити  $\hat{Z}_{RFC}$  та  $Z_{CBO}$  замість  $\hat{Z}_Y$  та  $Z_1$ , відповідно. Щоби записати лінійне регресійне рівняння для нормалізованої метрики CBO як (3.12), треба підставити  $\hat{Z}_{CBO}$  та  $Z_{RFC}$  замість  $\hat{Z}_Y$  та  $Z_1$ , відповідно. Оцінки параметрів

цих лінійних регресійних рівнянь для нормалізованих метрик RFC та CBO та відповідні параметри тривимірного перетворення Бокса – Кокса  $\hat{\lambda}_{RFC}$  та  $\hat{\lambda}_{CBO}$  наведені в таблиці 3.14.

Щоб оцінити прогностичну точність лінійних регресійних рівнянь для нормалізованих метрики RFC та CBO у формі (3.12), ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Для зазначених рівнянь значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.14. Дані показників MMRE та PRED(0,25) з табл. 3.14 свідчать про прийнятну якість рівняння для нормалізованих метрики RFC. Якість лінійного регресійного рівняння для нормалізованої метрики CBO у формі (3.12) не може вважатися задовільною тому, що значення MMRE та PRED(0,25) відповідно на 60,6% та 35,5% гірші за прийнятні граничні значення.

Таблиця 3.14 – Оцінки параметрів лінійних регресійних рівнянь для нормалізованих метрики RFC та CBO

| No | Метрики | $\hat{b}_0$ | $\hat{b}_1$ | $\hat{\lambda}_j$ | $R^2$  | MMRE   | PRED   |
|----|---------|-------------|-------------|-------------------|--------|--------|--------|
| 1  | RFC     | 2,68256     | 0,05042     | 0,162485          | 0,0047 | 0,1697 | 0,7984 |
| 2  | CBO     | 1,77086     | 0,09280     | 0,463193          | 0,0047 | 0,4017 | 0,4839 |

Для порівняння побудованих лінійних регресійних рівнянь для нормалізованих метрики RFC та CBO у формі (3.12) з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127]. Ми побудували два лінійні SVR для нормалізованих метрик RFC та CBO на основі того ж самого набору, який використовувався для побудови лінійних регресійних рівнянь для нормалізованих метрик RFC та CBO у формі (3.12). Для двох лінійних SVR для



нормалізованих метрики RFC та CBO обчислені значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.15.

Таблиця 3.15 – Значення  $R^2$ , MMRE та PRED(0,25) для SVR для нормалізованих метрик RFC та CBO

| No | Метрики | $R^2$ | MMRE   | PRED(0,25) |
|----|---------|-------|--------|------------|
| 1  | RFC     | 0,0   | 0,1592 | 0,8065     |
| 2  | CBO     | 0,0   | 0,4265 | 0,5645     |

Дані показників MMRE та PRED(0,25) з табл. 3.15 свідчать про прийнятну якість лінійної SVR для нормалізованої метрики RFC. Якість лінійної SVR для нормалізованої метрики CBO не може вважатися задовільною тому, що значення MMRE та PRED(0,25) відповідно на 70,6% та 24,7% гірші за прийнятні граничні значення. Крім того, значення показника  $R^2$  для двох SVR для нормалізованих метрики RFC та CBO дорівнюють нулю. А це вказує на те, що замість двох SVR можуть бути використані вибіркові середні метрик RFC та CBO.

Враховуючи останнє, в подальшому ми використовуємо лінійні регресійні рівняння для нормалізованих метрик RFC та CBO у формі (3.12) як доповнення до рівняння еліпсу прогнозування для нормалізованих метрик RFC та CBO виду (1.7) для визначення складності ООП на третьому кроці ООП через зв'язки між класами за метриками RFC та CBO.

Таким чином, було розроблено математичну модель для обробки інформації з двох програмних метрик RFC та CBO Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді рівняння еліпсу прогнозування та двох лінійних регресійних рівнянь для нормалізованих метрик RFC та CBO на основі двовимірного перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з

метрик RFC та CBO за рахунок врахування як кореляції між ними, так і відхилення двовимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на третьому кроці ООП через зв'язки між класами за метриками RFC та CBO Web застосунків, що розробляються з використанням PHP-фреймворків.

### **3.3 Математичні моделі для обробки інформації з чотирьох програмних метрик для рішення задачі з оцінювання кількості рядків коду Web застосунків**

Для побудови математичної моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) для рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, були використані їх дані, що наведені у додатку В. Згідно з [31-34] рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням певного PHP-фреймворку, потребує побудови чотиривимірного еліпсоїду прогнозування для зазначених вище нормалізованих метрик та трифакторної нелінійної регресійної моделі на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Для визначення точності оцінки умовного вибіркового середнього кількості рядків коду треба побудувати рівняння границь довірчих інтервалів нелінійної регресії. Для моделювання кількості рядків коду як залежної випадкової величини та визначення викидів в регресійній моделі необхідно ще мати рівняння границь інтервалів прогнозування відповідної нелінійної регресії.

Позначимо метрики кількості рядків коду (у тис. рядків коду, KLOC), кількості класів, середньої кількості методів на клас та DIT як відповідно  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  [31-34]. Зазначимо, що вибір у якості факторів регресійної моделі саме зазначених трьох метрик (кількості класів, середньої кількості методів на клас та DIT) був обумовлений, відсутністю наявності мультиколінеарності між ними та забезпеченням прийнятної достовірності раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків.

Як відомо [109], під мультиколінеарністю розуміють наявність лінійної залежності між двома або більше незалежними змінними (факторами) у регресійній моделі. Мультиколінеарність – це негативне явище множинного регресійного аналізу, яке не дозволяє здійснити оцінювання окремого впливу кожного фактору на залежну змінну. Наявність мультиколінеарності свідчить про те, що в множинній регресійній моделі два або більше факторів (незалежних змінних) пов'язані між собою або мають високий ступінь кореляції.

Наявність або відсутність мультиколінеарності серед факторів в моделі множинної регресії було визначено за коефіцієнтами впливу дисперсії (VIFs) – це діагональні елементи оберненої коваріаційної матриці  $k \times k$   $k$ -факторів [109]. Для наших даних обраних факторів значення VIFs знаходяться у межах від 1 до 5, тому мультиколінеарності немає.

Рівняння чотиривимірного еліпсоїду прогнозування для нормалізованих метрик  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$ , які позначено відповідно як  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$ , будемо подібно (3.1). Воно має наступний вигляд:

$$(\mathbf{Z} - \bar{\mathbf{Z}})^T \mathbf{S}_N^{-1} (\mathbf{Z} - \bar{\mathbf{Z}}) = \frac{4(N^2 - 1)}{N(N - 4)} F_{4, N-4, \alpha}. \quad (3.13)$$

В (3.13) всі позначення як і в (3.1) лише з тою різницею, що в правій частині застосовується тестова статистика на основі квантиля  $F$ -розподілу  $F_{4N-4,\alpha}$  з 4 і  $N - 4$  ступенями свободи та рівнем значущості  $\alpha$ , який дорівнює 0,005.

За допомогою (3.13) визначається чотиривимірна точка даних змінних  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$ , яка є викидом. Якщо для цієї точки, координати якої дорівнюють значенням змінних  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$ , квадрат відстані Махаланобіса більше за праву частину (3.1), то вона є чотиривимірним викидом.

Регресійні моделі, рівняння границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, ми будуємо подібно (1.8) та (1.9) лише з тою різницею, що це робиться на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса з іншими оцінками параметрів та  $\hat{Z}_Y$ , яке визначається як

$$\hat{Z}_Y = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2 + \hat{b}_3 Z_3. \quad (3.14)$$

Рівняння границь довірчих інтервалів нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, мають такий же загальний вигляд, як і (3.2) та (3.3) лише з тою різницею, що вони будуються на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса, мають інші оцінки параметрів та  $\hat{Z}_Y$  знаходиться за (3.14).

Рівняння границь інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, будуються аналогічно (3.2) та (3.3) із зазначеними вище

зауваженнями та з тією різницею, що до суми у фігурних дужках (3.2) та (3.3) потрібно додати ще 1.

Для побудови еліпсоїду прогнозування для нормалізованих метрик чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT), регресійних моделей у формі (1.8) та рівнянь границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, ми використали дані з 125 вебзастосунків, розміщених на GitHub для п'яти відомих фреймворків: CakePHP, CodeIgniter, Laravel, Symfony та Yii. Описова статистика цього набору даних наведена в табл.1.2, а самі дані у додатку В (табл.В.2). Після видалення викидів з цих даних залишилися дані з 107 застосунків, описова статистика з яких наведена в таблиці 3.16.

Таблиця 3.16 – Описова статистика набору даних із 107 застосунків

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 1,008 | 349,736 | 25,422  | 42,952  |
| Кількість класів                       | 25    | 7743    | 451,925 | 932,522 |
| Кількість методів на клас              | 1,41  | 12,24   | 4,845   | 2,765   |
| DIT                                    | 1,08  | 2,44    | 1,495   | 0,293   |

Для розрахунку значень квадрата відстані Махаланобіса в лівій частині (3.13) обернена матриця є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 48,621 & -83,941 & -84,315 & -14,901 \\ -83,941 & 148,652 & 145,505 & 38,302 \\ -84,315 & 145,505 & 154,763 & 32,436 \\ -14,901 & 38,302 & 32,436 & 159,410 \end{pmatrix}, \quad (3.15)$$

а вибіркові середні  $\bar{Z}_Y$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 2,263, 4,090, 1,195 та 0,273. Квантиль  $F_{4,N-4,0.005}$  дорівнює 3,956 для  $N=107$ .

При цьому оцінки параметрів чотиривимірного нормалізуючого перетворення Бокса – Кокса мають такі значення:  $\hat{\lambda}_Y = -0,031661$ ,  $\hat{\lambda}_1 = -0,103952$ ,  $\hat{\lambda}_2 = -0,241605$ ,  $\hat{\lambda}_3 = -1,600947$ .

Щоб уникнути проблеми оверфітінгу нелінійної регресійної моделі для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, ми розділили вибірку даних на два набори: навчальний (54 точки даних) і тестовий (53 точки даних). Таблиця 3.17 містить описові статистики навчального, а таблиця 3.18 – тестового набору даних.

Оцінки параметрів нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, за навчальним набором даних (54 точки даних) є такими:  $\hat{b}_0 = -7,03521$ ,  $\hat{b}_1 = 1,74803$ ,  $\hat{b}_2 = 1,71631$ ,  $\hat{b}_3 = 0,264516$ ,  $\hat{\lambda}_Y = -0,031661$ ,  $\hat{\lambda}_1 = -0,103952$ ,  $\hat{\lambda}_2 = -0,241605$ ,  $\hat{\lambda}_3 = -1,600947$ . Оцінка  $\hat{\sigma}_\varepsilon$  середнього квадратичного відхилення  $\varepsilon$  дорівнює 0,1455.

Таблиця 3.17 – Описова статистика навчального набору даних (54 точки даних)

| Назва метрики                          | Min   | Max     | Середнє | СКВ      |
|----------------------------------------|-------|---------|---------|----------|
| Розмір застосунку, у тис. рядків, KLOC | 1,008 | 349,736 | 26,934  | 50,799   |
| Кількість класів                       | 25    | 7743    | 470,778 | 1065,582 |
| Кількість методів на клас              | 1,41  | 12,24   | 4,853   | 2,810    |
| DIT                                    | 1,08  | 2,44    | 1,518   | 0,324    |

Таблиця 3.18 – Описова статистика тестового набору даних (53 точки даних)

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 1,125 | 202,655 | 23,881  | 33,553  |
| Кількість класів                       | 32    | 4934    | 432,717 | 783,960 |
| Кількість методів на клас              | 1,79  | 11,57   | 4,836   | 2,746   |
| DIT                                    | 1,11  | 2,13    | 1,471   | 0,257   |

Щоб визначити границі довірчих інтервалів та інтервалів прогнозування нелінійної регресії для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних РНР-фреймворків, за рівняннями (3.2) та (3.3) з урахуванням (3.14), обернена матриця є такою:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,06395 & -0,00285 & 0,18051 \\ -0,00285 & 0,15286 & 0,11574 \\ 0,18051 & 0,11574 & 2,32734 \end{pmatrix}, \quad (3.16)$$

а значення  $S_{Z_Y}$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 0,1498, 4,098, 1,193 та 0,277.

У всіх випадках обчислення границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, необхідно використовувати в (3.2) та (3.3) з урахуванням (3.14)  $t_{\alpha/2, \nu} = 2,101$ ,  $\alpha = 0,05$ ,  $N=54$  та  $\nu = 50$ .

Щоб оцінити прогностичну точність нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25) для навчального (54 точки даних) та тестового (53 точки даних) наборів даних. Для зазначеної нелінійної регресійної моделі значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.19.

Таблиця 3.19 – Значення  $R^2$ , MMRE та PRED(0,25) регресійної моделі у формі (1.8) для навчального та тестового наборів даних

| No | Набір даних | $R^2$  | MMRE   | PRED(0,25) |
|----|-------------|--------|--------|------------|
| 1  | навчальний  | 0,9722 | 0,1279 | 0,8889     |
| 2  | тестовий    | 0,9750 | 0,1168 | 0,8868     |

Значення показників  $R^2$ , MMRE та PRED(0,25) з табл. 3.19 свідчать про добру якість нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків.



Для порівняння побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127]. Ми побудували лінійну SVR для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, на основі того ж самого навчального набору з 54 точок даних. Перевірку цієї SVR ми здійснили за тестовим набором з 53 точок даних. Для цієї SVR значення  $R^2$ , MMRE та PRED(0,25) наведені в таблиці 3.20.

Таблиця 3.20 – Значення  $R^2$ , MMRE та PRED(0,25) лінійної SVR для навчального та тестового наборів даних

| № | Набір даних | $R^2$  | MMRE   | PRED(0,25) |
|---|-------------|--------|--------|------------|
| 1 | навчальний  | 0,9485 | 0,6195 | 0,4815     |
| 2 | тестовий    | 0,9020 | 0,5849 | 0,4340     |

Дані показників з табл. 3.20 свідчать про прийнятну якість лінійної SVR лише для визначення умовного середнього кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, тому, що значення MMRE та PRED(0,25) відповідно гірші за прийнятні граничні значення на 147,8% та 35,8% для навчального та на 134,0% та 42,1% для тестового набору даних.

Порівняння даних табл. 3.19 та 3.20 вказує на кращу якість нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) у порівнянні з SVR для визначення кількості рядків коду Web застосунків, що розробляються з

використанням різних PHP-фреймворків, саме в першу чергу з точки зору її (кількості рядків коду) моделювання як залежної випадкової величини.

Побудовану нелінійну регресійну модель у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, ми перевірили на трьох наборах даних для окремих фреймворків: CakePHP, CodeIgniter та Yii. Ці дані наведені в роботах [32-34]. Значення  $R^2$ , MMRE та PRED(0,25) регресійної моделі у формі (1.8) для трьох наборів даних для окремих фреймворків наведені в табл. 3.21.

Таблиця 3.21 – Значення  $R^2$ , MMRE та PRED(0,25) регресійної моделі у формі (1.8) для трьох наборів даних для окремих фреймворків

| No | Набір даних      | $R^2$  | MMRE   | PRED(0,25) |
|----|------------------|--------|--------|------------|
| 1  | CakePHP [32]     | 0,9064 | 0,6519 | 0,0        |
| 2  | CodeIgniter [34] | 0,3639 | 0,5406 | 0,0        |
| 3  | Yii [33]         | 0,8933 | 0,2372 | 0,6750     |

Дані табл. 3.21 вказують на те, що нелінійна регресійна модель у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, показала непогані результати лише для набору з 40 точок даних з роботи [33] для фреймворку Yii. Для цих даних лише значення PRED(0,25) є гіршим за прийнятне граничне значення всього на 10%. Для двох інших наборів даних з робіт [32] та [34] для відповідно фреймворків CakePHP та CodeIgniter нелінійна регресійна модель у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, показала задовільні результати лише для визначення умовного середнього кількості рядків коду Web застосунків, що розробляються з використанням

фреймворків CakePHP та CodeIgniter. Отримані результати, у тому числі, свідчать про необхідність побудови окремих нелінійних регресійних моделей у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням конкретних PHP-фреймворків. Що і було зроблено у подальшому.

Для побудови математичної моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) для рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, були використані їх дані, що наведені у додатку В, та дані з роботи [32]. Описова статистика цього набору даних з 54 вебзастосунків наведена в табл. 3.22.

Таблиця 3.22 – Описова статистика набору даних із 54 вебзастосунків, що розроблялися з використанням фреймворку CakePHP

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,124 | 138,943 | 14,294  | 24,122  |
| Кількість класів                       | 1     | 1496    | 150,222 | 252,658 |
| Кількість методів на клас              | 1,74  | 11,57   | 4,568   | 2,141   |
| DIT                                    | 1,19  | 3,00    | 1,681   | 0,339   |

Для побудови еліпсоїду прогнозування для нормалізованих чотирьох програмних метрик (3.13), регресійних моделей у формі (1.8) та рівнянь границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, остаточно ми використали вже дані з 51 вебзастосунку після видалення

трьох викидів. Після видалення викидів з цих даних залишилися дані з 51 вебзастосунок, описова статистика яких наведена в таблиці 3.23.

Таблиця 3.23 – Описова статистика набору даних із 51 вебзастосунок, що розроблялися з використанням фреймворку CakePHP

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,124 | 138,943 | 15,059  | 24,618  |
| Кількість класів                       | 1     | 1496    | 157,353 | 258,284 |
| Кількість методів на клас              | 1,74  | 11,57   | 4,626   | 2,188   |
| DIT                                    | 1,19  | 3,00    | 1,667   | 0,340   |

Для розрахунку значень квадрата відстані Махаланобіса в лівій частині (3.13) обернена матриця є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 29,362 & -29,712 & -54,400 & -10,023 \\ -29,712 & 30,563 & 54,891 & 13,272 \\ -54,400 & 54,891 & 116,743 & 26,627 \\ -10,023 & 13,272 & 26,627 & 91,265 \end{pmatrix}, \quad (3.17)$$

а вибіркові середні  $\bar{Z}_Y$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 1,296, 3,548, 1,127 та 0,394. Квантиль  $F_{4,N-4,0.005}$  дорівнює 4,267 для  $N=51$ .

При цьому оцінки параметрів чотиривимірного нормалізуючого перетворення Бокса – Кокса мають такі значення:  $\hat{\lambda}_Y = -0,020921$ ,  $\hat{\lambda}_1 = -0,026835$ ,  $\hat{\lambda}_2 = -0,329645$ ,  $\hat{\lambda}_3 = -0,828547$ .

Регресійну модель, рівняння границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, ми будемо подібно (1.8) та (1.9) лише з тою різницею, що це робиться на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса з іншими оцінками параметрів та  $\hat{Z}_Y$ , яке визначається за (3.14).

Рівняння границь довірчих інтервалів нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, мають такий же загальний вигляд, як і (3.2) та (3.3) лише з тою різницею, що вони будуються на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса, мають інші оцінки параметрів та  $\hat{Z}_Y$  знаходиться за (3.14).

Рівняння границь інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, будуються аналогічно (3.2) та (3.3) із зазначеними вище зауваженнями та з тією різницею, що до суми у фігурних дужках (3.2) та (3.3) потрібно додати ще 1.

Оцінки параметрів нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, за набором з 51 точки даних є такими:  $\hat{b}_0 = -4,51580$ ,  $\hat{b}_1 = 1,01193$ ,  $\hat{b}_2 = 1,85278$ ,  $\hat{b}_3 = 0,341347$ ,  $\hat{\lambda}_Y = -0,020921$ ,  $\hat{\lambda}_1 = -0,026835$ ,  $\hat{\lambda}_2 = -0,329645$ ,  $\hat{\lambda}_3 = -0,828547$ . Оцінка  $\hat{\sigma}_\varepsilon$  середнього квадратичного відхилення є дорівнює 0,1864.

Щоб визначити границі довірчих інтервалів та інтервалів прогнозування нелінійної регресії для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, за рівняннями (3.2) та (3.3) з урахуванням (3.14), обернена матриця є такою:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,00975 & -0,00309 & 0,06138 \\ -0,00309 & 0,31277 & 0,15918 \\ 0,06138 & 0,15918 & 1,72242 \end{pmatrix}, \quad (3.18)$$

а значення  $S_{Z_Y}$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 0,1922, 3,548, 1,127 та 0,394.

У всіх випадках обчислення границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, необхідно використовувати в (3.2) та (3.3) з урахуванням (3.14)  $t_{\alpha/2, \nu} = 2,012$ ,  $\alpha = 0,05$ ,  $N=51$  та  $\nu = 47$ .

Щоб оцінити прогностичну точність нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Для зазначеної нелінійної регресійної моделі значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,9287, 0,1406 та 0,8039. Значення показників  $R^2$ , MMRE та PRED(0,25) свідчать про добру якість нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP.

Для порівняння побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127]. Ми побудували лінійну SVR для оцінювання кількості рядків коду Web застосунків, що розробляються з

використанням фреймворку CakePHP, на основі того ж самого набору з 51 точки даних. Для цієї SVR значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,8812, 3,504 та 0,4314.

Ці значення свідчать про прийнятну якість лінійної SVR лише для визначення умовного середнього кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP, тому, що значення MMRE та PRED(0,25) суттєво є гіршими за прийнятні граничні значення. Також значення  $R^2$ , MMRE та PRED(0,25) для SVR виявилися суттєво гіршими у порівнянні з тими же показниками для побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CakePHP. Зазначимо, таке велике значення MMRE для SVR пояснюється доволі великою різницею (понад 1120 разів) між мінімальним та максимальним значеннями розміру застосунку згідно таблиці 3.23. Для малих значень розміру (менших за 1 KLOC) ця SVR має великі відносні похибки прогнозування.

Для побудови математичної моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) для рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, були використані їх дані, що наведені у додатку В, та дані з роботи [34]. Описова статистика цього набору даних з 70 вебзастосунків наведена в табл. 3.24.

Для побудови еліпсоїду прогнозування для нормалізованих чотирьох програмних метрик (3.13), регресійних моделей у формі (1.8) та рівнянь границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку

CodeIgniter, остаточно ми використали вже дані з 67 вебзастосунків після видалення трьох викидів.

Таблиця 3.24 – Описова статистика набору даних із 70 вебзастосунків, що розроблялися з використанням фреймворку CodeIgniter

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,939 | 155,740 | 37,594  | 29,497  |
| Кількість класів                       | 10    | 584     | 155,629 | 100,721 |
| Кількість методів на клас              | 0,18  | 29,55   | 10,071  | 4,129   |
| DIT                                    | 1,17  | 2,07    | 1,378   | 0,191   |

Після видалення викидів з цих даних залишилися дані з 67 вебзастосунків, описова статистика яких наведена в таблиці 3.25.

Таблиця 3.25 – Описова статистика набору даних із 67 вебзастосунків, що розроблялися з використанням фреймворку CodeIgniter

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,939 | 155,740 | 38,549  | 29,726  |
| Кількість класів                       | 10    | 584     | 155,672 | 100,200 |
| Кількість методів на клас              | 0,18  | 13,75   | 9,887   | 3,412   |
| DIT                                    | 1,17  | 2,07    | 1,374   | 0,177   |

Для розрахунку значень квадрата відстані Махаланобіса в лівій частині (3.13) обернена матриця є такою:



$$\mathbf{S}_N^{-1} = \begin{pmatrix} 17,279 & -11,420 & -0,6195 & 135,153 \\ -11,420 & 9,195 & 0,3936 & -65,712 \\ -0,6195 & 0,3936 & 0,0659 & 6,691 \\ 135,153 & -65,712 & 6,691 & 5814,12 \end{pmatrix}, \quad (3.19)$$

а вибіркові середні  $\bar{Z}_Y$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 3,354, 5,930, 22,327 та 0,178. Квантиль  $F_{4,N-4,0.005}$  дорівнює 4,118 для  $N=67$ .

При цьому оцінки параметрів чотиривимірного нормалізуючого перетворення Бокса – Кокса мають такі значення:  $\hat{\lambda}_Y = 0,0068815$ ,  $\hat{\lambda}_1 = 0,077083$ ,  $\hat{\lambda}_2 = 1,529929$ ,  $\hat{\lambda}_3 = -3,694331$ .

Регресійну модель, рівняння границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, ми будуємо подібно (1.8) та (1.9) лише з тою різницею, що це робиться на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса з іншими оцінками параметрів та  $\hat{Z}_Y$ , яке визначається за (3.14).

Рівняння границь довірчих інтервалів нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, мають такий же загальний вигляд, як і (3.2) та (3.3) лише з тою різницею, що вони будуються на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса, мають інші оцінки параметрів та  $\hat{Z}_Y$  знаходиться за (3.14).

Рівняння границь інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, будуються аналогічно (3.2) та (3.3) із зазначеними вище

зауваженнями та з тією різницею, що до суми у фігурних дужках (3.2) та (3.3) потрібно додати ще 1.

Оцінки параметрів нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, за набором з 67 точки даних є такими:  $\hat{b}_0 = 0,029691$ ,  $\hat{b}_1 = 0,660935$ ,  $\hat{b}_2 = 1,85278$ ,  $\hat{b}_3 = 0,035852$ ,  $\hat{\lambda}_y = 0,0068815$ ,  $\hat{\lambda}_1 = 0,077083$ ,  $\hat{\lambda}_2 = 1,529929$ ,  $\hat{\lambda}_3 = -3,694331$ . Оцінка  $\hat{\sigma}_\varepsilon$  середнього квадратичного відхилення  $\varepsilon$  дорівнює 0,2424.

Щоб визначити границі довірчих інтервалів та інтервалів прогнозування нелінійної регресії для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, за рівняннями (3.2) та (3.3) з урахуванням (3.14), обернена матриця є такою:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,02458 & -0,00024 & 0,35248 \\ -0,00024 & 0,000652 & 0,17218 \\ 0,35248 & 0,17218 & 70,9995 \end{pmatrix}, \quad (3.20)$$

а значення  $S_{Z_y}$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 0,2481, 5,930, 22,327 та 0,178.

У всіх випадках обчислення границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, необхідно використовувати в (3.2) та (3.3) з урахуванням (3.14)  $t_{\alpha/2, v} = 1,998$ ,  $\alpha = 0,05$ ,  $N=67$  та  $v = 63$ .

Щоб оцінити прогностичну точність нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, ми

використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Для зазначеної нелінійної регресійної моделі значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,9102, 0,2054 та 0,8209. Значення показників  $R^2$ , MMRE та PRED(0,25) свідчать про добру якість нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter.

Для порівняння побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127]. Ми побудували лінійну SVR для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, на основі того ж самого набору з 67 точок даних. Для цієї SVR значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,6988, 0,2898 та 0,7015.

Ці значення свідчать про прийнятну якість лінійної SVR в першу чергу для визначення умовного середнього кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter, тому, що значення MMRE та PRED(0,25) є гіршими за прийнятні граничні значення відповідно на 15,9% та 6,5%. Також значення  $R^2$ , MMRE та PRED(0,25) для SVR виявилися гіршими у порівнянні з тими же показниками для побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку CodeIgniter.

Для побудови математичної моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) для рішення задачі з оцінювання кількості рядків

коду Web застосунків, що розробляються з використанням фреймворку Yii, були використані їх дані, що наведені у додатку В, та дані з роботи [33]. Описова статистика цього набору даних з 56 вебзастосунків наведена в табл. 3.26.

Таблиця 3.26 – Описова статистика набору даних із 56 вебзастосунків, що розроблялися з використанням фреймворку Yii

| Назва метрики                          | Min   | Max     | Середнє | СКВ     |
|----------------------------------------|-------|---------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,120 | 119,314 | 16,828  | 27,302  |
| Кількість класів                       | 4     | 1428    | 284,446 | 400,054 |
| Кількість методів на клас              | 0,87  | 7,69    | 4,298   | 1,432   |
| DIT                                    | 1,06  | 2,40    | 1,793   | 0,308   |

Для побудови еліпсоїду прогнозування для нормалізованих чотирьох програмних метрик (3.13), регресійних моделей у формі (1.8) та рівнянь границь довірчих інтервалів та інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, остаточно ми використали вже дані з 50 вебзастосунків після видалення шести викидів. Після видалення викидів з цих даних залишилися дані з 50 вебзастосунків, описова статистика яких наведена в таблиці 3.27.

Для розрахунку значень квадрата відстані Махаланобіса в лівій частині (3.13) обернена матриця є такою:

$$\mathbf{S}_N^{-1} = \begin{pmatrix} 48,411 & -69,802 & -14,726 & 0,4605 \\ -69,802 & 102,105 & 20,828 & 0,4095 \\ -14,726 & 20,828 & 5,535 & 0,0602 \\ 0,4605 & 0,4095 & 0,0602 & 4,459 \end{pmatrix}, \quad (3.21)$$

а вибіркові середні  $\bar{Z}_Y$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 1,429, 3,855, 2,776 та 1,247. Квантиль  $F_{4,N-4,0.005}$  дорівнює 4,280 для  $N=50$ .

Таблиця 3.27 – Описова статистика набору даних із 50 вебзастосунків, що розроблялися з використанням фреймворку Yii

| Назва метрики                          | Min   | Max    | Середнє | СКВ     |
|----------------------------------------|-------|--------|---------|---------|
| Розмір застосунку, у тис. рядків, KLOC | 0,211 | 91,125 | 13,616  | 19,624  |
| Кількість класів                       | 10    | 1286   | 256,260 | 348,131 |
| Кількість методів на клас              | 0,87  | 7,69   | 4,339   | 1,415   |
| DIT                                    | 1,11  | 2,40   | 1,811   | 0,287   |

При цьому оцінки параметрів чотиривимірного нормалізуючого перетворення Бокса – Кокса мають такі значення:  $\hat{\lambda}_Y = -0,029337$ ,  $\hat{\lambda}_1 = -0,077323$ ,  $\hat{\lambda}_2 = 0,806581$ ,  $\hat{\lambda}_3 = 2,137901$ .

Регресійну модель, рівняння границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, ми будемо подібно (1.8) та (1.9) лише з тою різницею, що це робиться на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса з іншими оцінками параметрів та  $\hat{Z}_Y$ , яке визначається за (3.14).

Рівняння границь довірчих інтервалів нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, мають такий же загальний вигляд, як і (3.2) та (3.3) лише з тою різницею, що вони

будуються на основі чотиривимірною нормалізуючого перетворення Бокса – Кокса, мають інші оцінки параметрів та  $\hat{Z}_Y$  знаходиться за (3.14).

Рівняння границь інтервалів прогнозування нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, будуються аналогічно (3.2) та (3.3) із зазначеними вище зауваженнями та з тією різницею, що до суми у фігурних дужках (3.2) та (3.3) потрібно додати ще 1.

Оцінки параметрів нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, за набором з 50 точок даних є такими:  $\hat{b}_0 = -4,95794$ ,  $\hat{b}_1 = 1,44096$ ,  $\hat{b}_2 = 0,30400$ ,  $\hat{b}_3 = -0,009507$ ,  $\hat{\lambda}_Y = -0,029337$ ,  $\hat{\lambda}_1 = -0,077323$ ,  $\hat{\lambda}_2 = 0,806581$ ,  $\hat{\lambda}_3 = 2,137901$ . Оцінка  $\hat{\sigma}_\varepsilon$  середнього квадратичного відхилення  $\varepsilon$  дорівнює 0,1451.

Щоб визначити границі довірчих інтервалів та інтервалів прогнозування нелінійної регресії для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, за рівняннями (3.2) та (3.3) з урахуванням (3.14), обернена матриця є такою:

$$\mathbf{S}_Z^{-1} = \begin{pmatrix} 0,03045 & -0,00784 & 0,02146 \\ -0,00784 & 0,02117 & 0,00400 \\ 0,02146 & 0,00400 & 0,08910 \end{pmatrix}, \quad (3.22)$$

а значення  $S_{Z_Y}$ ,  $\bar{Z}_1$ ,  $\bar{Z}_2$  та  $\bar{Z}_3$  дорівнюють відповідно 0,1498, 3,855, 2,776 та 1,247.

У всіх випадках обчислення границь довірчих інтервалів та інтервалів прогнозування нелінійної регресії кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, необхідно використовувати в (3.2) та (3.3) з урахуванням (3.14)  $t_{\alpha/2, \nu} = 2,013$ ,  $\alpha = 0,05$ ,  $N=50$  та  $\nu = 46$ .

Щоб оцінити прогностичну точність нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, ми використали стандартні показники  $R^2$ , MMRE та PRED(0,25). Для зазначеної нелінійної регресійної моделі значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,9848, 0,1183 та 0,9000. Значення показників  $R^2$ , MMRE та PRED(0,25) свідчать про добру якість нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для визначення кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii.

Для порівняння побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, з іншими добре відомими моделями, які також застосовуються у машинному навчанні, ми використали регресію опорних векторів (SVR) [126, 127]. Ми побудували лінійну SVR для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, на основі того ж самого набору з 50 точок даних. Для цієї SVR значення  $R^2$ , MMRE та PRED(0,25) дорівнюють відповідно 0,8344, 1,658 та 0,4400.

Ці значення свідчать про прийнятну якість лінійної SVR лише для визначення умовного середнього кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii, тому, що значення MMRE та PRED(0,25) суттєво є гіршими за прийнятні граничні значення. Також значення  $R^2$ , MMRE та PRED(0,25) для SVR виявилися суттєво гіршими у порівнянні з тими же показниками для побудованої нелінійної регресійної моделі у формі (1.8) з урахуванням (3.14) для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням фреймворку Yii. Зазначимо, таке велике значення MMRE для SVR пояснюється доволі великою різницею (понад 430

разів) між мінімальним та максимальним значеннями розміру застосунку згідно таблиці 3.27. Для малих значень розміру (менших за 1 KLOC) ця SVR має великі відносні похибки прогнозування.

### 3.4 Висновки за розділом 3

У третьому розділі дисертації побудовано математичні моделі для обробки інформації з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків, для рішення трьох задач: оцінювання якості ПЗ, визначення складності ООП ПЗ та раннього оцінювання кількості рядків коду ПЗ зазначених застосунків.

1) Побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ці дві математичні моделі дозволяють підвищити достовірність результатів обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM за рахунок врахування як кореляції між ними, так і відхилення тривимірних розподілів цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його першому та другому етапах (через ідентифікацію та семантику класів) за Бучем. Зокрема, у порівнянні з SVR побудовані моделі дозволили підвищити кількість прогнозованих значень метрик WMC, DIT і NOC, для яких відносна похибка була менше 0,25, відповідно на 6,2%, 4,0% та 1,2%.

2) Удосконалено математичну модель для обробки інформації з



програмних метрик RFC і CBO Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC і CBO за рахунок врахування як кореляції між ними, так і відхилення двовимірного розподілу цих метрик від нормального, та здійснити на її основі рішення задачі з визначення складності ООП на його третьому етапі (через зв'язки між класами) за Бучем.

3) Удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC, CBO і WMC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на її основі рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків. Зокрема, для тестового набору даних у порівнянні з SVR удосконалена модель дозволила підвищити кількість прогнозованих значень метрик RFC, CBO і WMC, для яких відносна похибка була менше 0,25, відповідно на 76,0%, 77,8% та 81,8%.

4) Удосконалено три математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірних еліпсоїдів

прогнозування для зазначених нормалізованих метрик та трифакторних нелінійних регресійних моделей для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Ці математичні моделі дозволяють підвищити достовірність результатів обробки інформації із зазначених метрик за рахунок врахування як кореляції між ними, так і відхилення чотиривимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів. Отримані трифакторні нелінійні регресійні моделі в порівнянні з іншими регресійними моделями мають більший відсоток прогнозованих значень PRED, менше значення середньої величини відносної похибки MMRE та менші ширини довірчого інтервалу та інтервалу прогнозування нелінійної регресії. Зокрема, у порівнянні з SVR удосконалена модель для фреймворку Codeigniter дозволила збільшити значення PRED(0,25) на 17,0%, а значення MMRE зменшити на 29,1%.

5) Удосконалено рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування трифакторних нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Рівняння для визначення нижньої та верхньої границь довірчого інтервалу та інтервалу прогнозування дозволяють здійснювати оцінювання точності умовного вибіркового середнього та меж моделювання умовної випадкової величини кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii.

Результати досліджень автора, які наведені у поточному розділі, опубліковано в [31-34, 70-72, 78-85].

## **РОЗДІЛ 4. ІНСТРУМЕНТАРІЙ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ДЛЯ РІШЕННЯ ЗАДАЧ З ВИЗНАЧЕННЯ СКЛАДНОСТІ ООП, ЯКОСТІ СТВОРЕННЯ ТА КІЛЬКОСТІ РЯДКІВ КОДУ WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЯЮТЬСЯ З ВИКОРИСТАННЯМ РНР-ФРЕЙМВОРКІВ**

У четвертому розділі дисертації на підставі розроблених математичних моделей створено інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків. Для цього використовувалася мова програмування Python. Було створено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

Не зважаючи на те, що термін «інформаційна технологія» (ІТ) уперше був запропонований у 1958 році у статті Г. Лейвітта і Т. Уіслера, опублікованій у Гарвардському бізнес-огляді, він (термін “інформаційна технологія”) на сьогодні має багато визначень [131-136].

У державному стандарті України ДСТУ 2226-93 «Автоматизовані системи. Терміни та визначення» ІТ визначається як «Технологічний процес, предметом перероблення й результатом якого є інформація» [131].

У державному стандарті України ДСТУ 5034:2008 «Науково-інформаційна діяльність. Терміни та визначення понять» зазначено, що ІТ – це «сукупність методів, процесів і програмно-технічних засобів, об'єднаних у технологічний процес, що забезпечує добирання, зберігання, накопичення, оброблення, пошук, виведення, копіювання, передавання та розповсюдження

інформації» [132], іншими словами, забезпечує виконання технологічних операцій над інформацією.

В [133], ІТ визначається як використання комп'ютерів і телекомунікаційного обладнання (з пов'язаною з ними мікроелектронікою) для надсилання, отримання, зберігання та обробки даних. Дані можуть бути текстовими, числовими, аудіо чи відео або будь-якою їх комбінацією ("IT (information technology) The use of computers and telecommunications equipment (with their associated microelectronics) to send, receive, store and manipulate data. The data may be textual, numerical, audio or video, or any combination of these." [133]).

Згідно [134], ІТ – це технологія, яка використовується для отримання, зберігання, доставки та/або аналізу даних ("Information technology (IT) is a technology used to acquire, store, deliver and/or analyze data." [134]).

В [135], ІТ визначається як «процес або сукупність процесів оброблення інформації.»

На наш погляд, одне із найбільш вдалих визначень ІТ наведено в [132], що ІТ – це «сукупність методів, процесів і програмно-технічних засобів, об'єднаних у технологічний процес, що забезпечує» виконання сукупності технологічних операцій над інформацією. У цьому визначенні ІТ, як і у багатьох інших, наголос робиться на тому, що ІТ, є сукупністю, у тому числі, програмно-технічних засобів, тобто виділяється за сучасними визначеннями інструментарій ІТ як сукупність програмно-технічних засобів. В [135] навіть дещо звужується це поняття, а саме зазначається, що: «Інструментарій інформаційної технології – один або декілька взаємопов'язаних програмних продуктів для певного комп'ютера, технологія роботи, за допомогою яких користувач досягає обраної мети.» Таким чином, сьогодні під інструментарієм ІТ, як правило розуміють, сукупність апаратних та програмних засобів. Фактично при цьому, в ІТ

виділяють програмне та технічне забезпечення як інструментарій, який має базуватися на математичному та інформаційному забезпеченні.

#### **4.1 Інструментарій інформаційної технології для рішення задачі з визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків**

На основі розроблених математичних моделей створено інструментарій ІТ для рішення задач з визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на трьох етапах ОПП.

Для цього були застосовані дві математичні моделі для обробки інформації з двох груп програмних метрик СК (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса, а також математична модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням РНР-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса.

Перші дві математичні моделі дозволяють здійснити на їх основі рішення задачі з визначення складності ООП на його першому та другому етапах (через ідентифікацію та семантику класів) за Бучем, а третя математична модель здійснити на її основі рішення задачі з визначення складності ООП на його третьому етапі (через зв'язки між класами) за Бучем.

Отже спочатку розглянемо розробку ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на першому етапі ООП (через ідентифікацію класів).

Алгоритм визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на першому етапі ООП (через ідентифікацію класів) є таким.

Крок 1. Обрати етап ООП, для якого треба здійснити визначення складності ООП застосунку  $i$ .

Крок 2. Якщо обрано перший етап ООП (через ідентифікацію класів), то перейти до кроку 2.1, інакше перейти до кроку 3.

Крок 2.1. Ввести значення метрик WMC, DIT і NOC, які позначені відповідно як  $X_1$ ,  $X_2$  і  $X_3$ .

Крок 2.2. Нормалізувати значення  $X_1$ ,  $X_2$  та  $X_3$  (тривимірна точка даних  $i$ ) для застосунку  $i$  за допомогою тривимірного нормалізуючого перетворення Бокса – Кокса (2.1) з оцінками параметрів  $\hat{\lambda}_j$ , які наведені в таблиці 3.8. Нормалізовані метрики  $X_1$ ,  $X_2$  та  $X_3$  позначено відповідно як  $Z_1$ ,  $Z_2$  та  $Z_3$ . Результат: значення  $Z_1$ ,  $Z_2$  та  $Z_3$ .

Крок 2.3. Обчислити квадрат відстані Махаланобіса (SMD) для тривимірної нормалізованої точки даних (точка  $i$ ) із використанням оберненої матриці (3.8) та вибірових середніх  $\bar{Z}_{WMC}$ ,  $\bar{Z}_{DIT}$  і  $\bar{Z}_{NOC}$ , що дорівнюють відповідно 1,869, 0,284 і -0,265.

Крок 2.4. Якщо значення квадрату відстані Махаланобіса для  $Z_1$ ,  $Z_2$  та  $Z_3$  буде більше за 13,85, то стоп (ми не можемо використати рівняння (3.1) та (3.9) для цих значень), інакше перейти до кроку 2.5.

Крок 2.5. Обчислити значення  $\hat{Z}_{WMC}$ ,  $\hat{Z}_{DIT}$  та  $\hat{Z}_{NOC}$  за (3.9) для точки даних  $i$ .

Крок 2.6. Якщо для точки даних  $i$  значення  $Z_1$  менше  $\hat{Z}_{WMC}$  та  $Z_2$  менше  $\bar{Z}_{DIT}$  та  $Z_3$  менше  $\hat{Z}_{NOC}$ , то складність ООП через ідентифікацію класів  $i$ -го вебзастосунку не висока, інакше складність – висока.

Крок 3. Якщо обрано другий етап ООП (через семантику класів), то перейти до кроку 3.1, інакше перейти до кроку 4.

...

За зазначеним алгоритмом мовою Python було створено відповідне ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням PHP-фреймворків, на першому етапі ООП (через ідентифікацію класів).

В ПЗ були використані такі інтерфейсні компоненти: вікно редагування, командні кнопки та радіокнопки. Графічне вікно ПЗ із зазначеними інтерфейсними компонентами наведено на рис.4.1. На рис.4.1 показаний результат визначення складності ООП для застосунку AdaptCMS на трьох етапах ООП у разі вибору першого етапу (результат: не висока складність). AdaptCMS (<https://github.com/adaptcms/AdaptCMS>) – це CMS з відкритим кодом, створена для контролю над вашим вебсайтом та легко адаптується до будь-якого типу вебсайту. Результати визначення складності на першому кроці ООП для десяти реальних вебзастосунків за розробленим ПЗ наведені в табл. 4.1.

Gymie (<https://github.com/lubusIN/laravel-gymie>) створений за допомогою Laravel для керування спортзалом і клубом. Attendize (<https://github.com/Attendize/Attendize>) – це платформа з відкритим кодом для продажу квитків і керування подіями, створена на фреймворку Laravel. Cachet (<https://github.com/CachetHQ/Cachet>) – це система власно-розміщених статусних



сторінок з відкритим вихідним кодом на основі фреймворку Laravel. Crater (<https://github.com/crater-invoice-inc/crater>) – це веб- та мобільний застосунок із відкритим кодом, який допомагає відстежувати витрати, платежі та створювати професійні рахунки-фактури й кошториси.

**OOD complexity**

Select Stage:

☒ 1. Class Identification

☐ 2. Class Semantics

☐ 3. Class Relationships

WMC: 6.93

DIT: 1.46

NOC: 0.71

Calculate

no high complexity

Рисунок 4.1. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору першого етапу (результат: не висока складність)

Craft (<https://github.com/craftcms/cms>) – це гнучка, зручна CMS на основі фреймворку Yii для створення власного цифрового досвіду в Веб та за його межами. Crud (<https://github.com/FriendsOfCake/crud>) створено на основі Yii, щоб використовувати його для розробки стероїдів і надати розробникам достатньо гнучкості, щоб використовувати його як для швидкого створення прототипів, так і для виробничих програм. Sandbox (<https://github.com/dereuromark/cakephp->

sandbox) – це вебзастосунок для демонстрації функцій CakePHP, а також створення пісочниці для випробувань. Wallabag (<https://github.com/wallabag/wallabag>) – це веб-додаток на основі Symfony, який дозволяє зберігати вебсторінки для подальшого читання. Apiato (<https://github.com/apiato/apiato>) – це PHP-фреймворк, побудований на базі фреймворку Laravel, спеціально розроблений для створення масштабованих, API-орієнтованих застосунків.

Таблиця 4.1 – Результати визначення складності на першому етапі ООП завдяки ідентифікації класів для десяти вебзастосунків

| <i>i</i> | Назва застосунку | WMC   | DIT  | NOC    | SMD   | складність |
|----------|------------------|-------|------|--------|-------|------------|
| 1        | Gymie            | 4,87  | 1,43 | 0,1449 | 12,83 | high       |
| 2        | Attendize        | 5,35  | 1,57 | 0,2857 | 12,26 | high       |
| 3        | Cachet           | 3,85  | 1,10 | 0,1044 | 8,68  | no high    |
| 4        | Crater           | 6,48  | 1,50 | 0,2315 | 12,78 | high       |
| 5        | Craft            | 20,45 | 1,84 | 0,5638 | 9,97  | high       |
| 6        | Crud             | 9,13  | 1,63 | 0,6000 | 2,91  | high       |
| 7        | Sandbox          | 7,04  | 1,81 | 0,5505 | 6,62  | high       |
| 8        | Wallabag         | 8,45  | 1,37 | 0,4405 | 4,09  | high       |
| 9        | AdaptCMS         | 6,93  | 1,46 | 0,71   | 0,21  | no high    |
| 10       | Apiato           | 1,75  | 2,44 | 0,77   | 8,90  | high       |

Зазначимо, що для застосунку Apiato висока складність ОПП на першому етапі ОПП була визначена лише завдяки значенню метрики DIT. Жоден застосунок з табл. 4.1 не мав високу складність ОПП на першому етапі ОПП завдяки значенню метрики NOC. На рис.4.2 показаний результат визначення

складності ООП для застосунку Sandbox на трьох етапах ОПП у разі вибору першого етапу (результат: висока складність).

**OOD complexity**

Select Stage:

☒ 1. Class Identification

☐ 2. Class Semantics

☐ 3. Class Relationships

WMC: 7.04

DIT: 1.81

NOC: 0.5505

Calculate

high complexity

Рисунок 4.2. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору першого етапу (результат: висока складність)

Розглянемо розробку ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на другому етапі ООП (через семантику класів).

Алгоритм визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на другому етапі ООП (через семантику класів) є таким.

Крок 1. Обрати етап ООП, для якого треба здійснити визначення складності ООП застосунку *i*.

Крок 2. Якщо обрано перший етап ООП (через ідентифікацію класів), то перейти до кроку 2.1, інакше перейти до кроку 3.

Крок 2.1. Ввести значення метрик WMC, DIT і NOC , які позначені відповідно як  $X_1$ ,  $X_2$  і  $X_3$ .

...

Крок 3. Якщо обрано другий етап ООП (через семантику класів), то перейти до кроку 3.1, інакше перейти до кроку 4.

Крок 3.1. Ввести значення метрик WMC, RFC і LCOM, які позначені відповідно як  $X_1$ ,  $X_2$  і  $X_3$ .

Крок 3.2. Нормалізувати значення  $X_1$ ,  $X_2$  та  $X_3$  (тривимірна точка даних  $i$ ) для застосунку  $i$  за допомогою тривимірного нормалізуючого перетворення Бокса – Кокса (2.1) з оцінками параметрів  $\hat{\lambda}_j$ , які наведені в таблиці 3.11. Нормалізовані метрики  $X_1$ ,  $X_2$  та  $X_3$  позначено відповідно як  $Z_1$ ,  $Z_2$  та  $Z_3$ . Результат: значення  $Z_1$ ,  $Z_2$  та  $Z_3$ .

Крок 3.3. Обчислити квадрат відстані Махаланобіса (SMD) для тривимірної нормалізованої точки даних (точка  $i$ ) із використанням оберненої матриці (3.10) та вибірових середніх  $\bar{Z}_{WMC}$ ,  $\bar{Z}_{RFC}$  та  $\bar{Z}_{LCOM}$ , що дорівнюють відповідно 1,629, 2,431 і 0,547.

Крок 3.4. Якщо значення квадрату відстані Махаланобіса для  $Z_1$ ,  $Z_2$  та  $Z_3$  буде більше за 13,83, то стоп (ми не можемо використати рівняння (3.1) та (3.9) для цих значень), інакше перейти до кроку 3.5.

Крок 3.5. Обчислити значення  $\hat{Z}_{WMC}$ ,  $\hat{Z}_{RFC}$  та  $\hat{Z}_{LCOM}$  за (3.9) для точки даних  $i$ .

Крок 3.6. Якщо для точки даних  $i$  значення  $Z_1$  менше  $\hat{Z}_{WMC}$  та  $Z_2$  менше  $\hat{Z}_{RFC}$  та  $Z_3$  менше  $\hat{Z}_{LCOM}$ , то складність ООП через семантику класів  $i$ -го вебзастосунку не висока, інакше складність – висока.

Крок 4. Якщо обрано третій етап ООП (через зв'язки між класами), то перейти до кроку 4.1, інакше стоп.

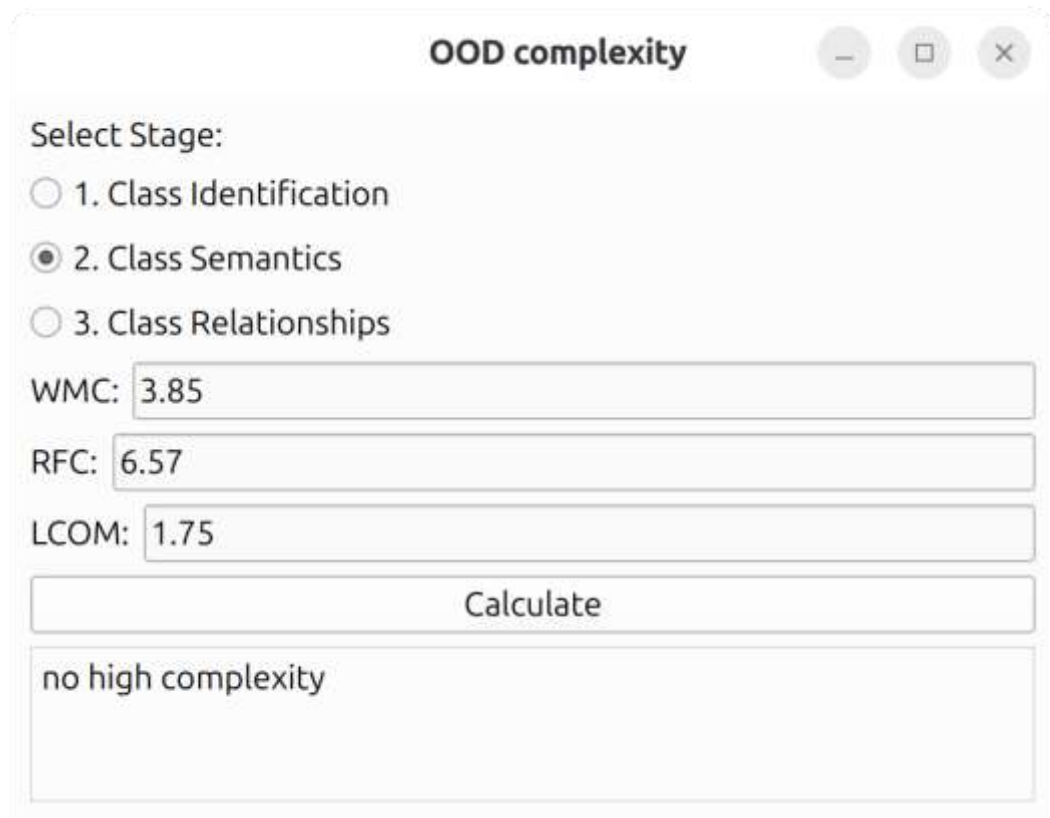
...

За зазначеним алгоритмом мовою Python було створено відповідне ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням PHP-фреймворків, на другому етапі ООП (через семантику класів). Результати визначення складності на другому етапі ООП для десяти реальних вебзастосунків за розробленим ПЗ наведені в табл. 4.2.

Таблиця 4.2 – Результати визначення складності на другому етапі ООП завдяки ідентифікації класів для десяти вебзастосунків

| $i$ | Назва застосунку | WMC   | RFC   | LCOM | SMD   | складність |
|-----|------------------|-------|-------|------|-------|------------|
| 1   | Gymie            | 4,87  | 7,41  | 2,62 | 2,61  | high       |
| 2   | Attendize        | 5,35  | 8,01  | 2,08 | 0,65  | high       |
| 3   | Cachet           | 3,85  | 6,57  | 1,75 | 1,36  | no high    |
| 4   | Crater           | 6,48  | 8,76  | 3,66 | 5,01  | high       |
| 5   | Craft            | 20,45 | 15,14 | 3,03 | 2,30  | high       |
| 6   | Crud             | 9,13  | 12,20 | 1,3  | 6,29  | high       |
| 7   | Sandbox          | 7,04  | 10,49 | 1,68 | 1,29  | high       |
| 8   | Wallabag         | 8,45  | 12,59 | 1,7  | 2,62  | high       |
| 9   | AdaptCMS         | 6,93  | 8,62  | 2,68 | 1,63  | high       |
| 10  | Apiato           | 1,75  | 3     | 1,35 | 10,00 | high       |

На рис.4.3 показаний результат визначення складності ООП для застосунку Cachet на трьох етапах ОПП у разі вибору другого етапу (результат: не висока складність).



The screenshot shows a window titled "OOD complexity" with standard window controls (minimize, maximize, close). Inside the window, there is a section labeled "Select Stage:" with three radio button options: "1. Class Identification", "2. Class Semantics" (which is selected), and "3. Class Relationships". Below these options, there are three input fields: "WMC:" with the value "3.85", "RFC:" with the value "6.57", and "LCOM:" with the value "1.75". A "Calculate" button is positioned below these fields. At the bottom of the window, a text box displays the result "no high complexity".

Рисунок 4.3. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору другого етапу (результат: не висока складність)

На рис.4.4 показаний результат визначення складності ООП для застосунку Craft на трьох етапах ОПП у разі вибору другого етапу (результат: висока складність).

**OOD complexity**

Select Stage:

☐ 1. Class Identification

☒ 2. Class Semantics

☐ 3. Class Relationships

WMC: 20.45

RFC: 15.14

LCOM: 3.03

Calculate

high complexity

Рисунок 4.4. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору другого етапу (результат: висока складність)

Розглянемо розробку ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на третьому етапі ООП (через зв'язки між класами).

Алгоритм визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на третьому етапі ООП (через зв'язки між класами) є таким.

Крок 1. Обрати етап ООП, для якого треба здійснити визначення складності ООП застосунку *i*.

Крок 2. Якщо обрано перший етап ООП (через ідентифікацію класів), то перейти до кроку 2.1, інакше перейти до кроку 3.

Крок 2.1. Ввести значення метрик WMC, DIT і NOC , які позначені відповідно як  $X_1$ ,  $X_2$  і  $X_3$ .

...

Крок 3. Якщо обрано другий етап ООП (через семантику класів), то перейти до кроку 3.1, інакше перейти до кроку 4.

Крок 3.1. Ввести значення метрик WMC, RFC і LCOM, які позначені відповідно як  $X_1$ ,  $X_2$  і  $X_3$ .

...

Крок 4. Якщо обрано третій етап ООП (через зв'язки між класами), то перейти до кроку 4.1, інакше стоп.

Крок 4.1. Ввести значення метрик RFC і CBO, які позначені відповідно як  $X_1$  і  $X_2$ .

Крок 4.2. Нормалізувати значення  $X_1$  та  $X_2$  (двовимірна точка даних  $i$ ) для застосунку  $i$  за допомогою двовимірного нормалізуючого перетворення Бокса – Кокса (2.1) з оцінками параметрів  $\hat{\lambda}_j$ , які наведені в таблиці 3.14. Нормалізовані метрики  $X_1$  та  $X_2$  позначено відповідно як  $Z_1$  та  $Z_2$ . Результат: значення  $Z_1$  та  $Z_2$ .

Крок 4.3. Обчислити квадрат відстані Махаланобіса (SMD) для двовимірної нормалізованої точки даних (точка  $i$ ) із використанням оберненої матриці (3.11) та вибірових середніх  $\bar{Z}_{RFC}$  та  $\bar{Z}_{CBO}$ , що дорівнюють відповідно 2,029 та 2,785.

Крок 4.4. Якщо значення квадрату відстані Махаланобіса для  $Z_1$  та  $Z_2$  буде більше за 11,25, то стоп (ми не можемо використати рівняння (3.1) та (3.12) для цих значень), інакше перейти до кроку 4.5.

Крок 4.5. Обчислити значення  $\hat{Z}_{RFC}$  та  $\hat{Z}_{CBO}$  за (3.12) для точки даних  $i$ .



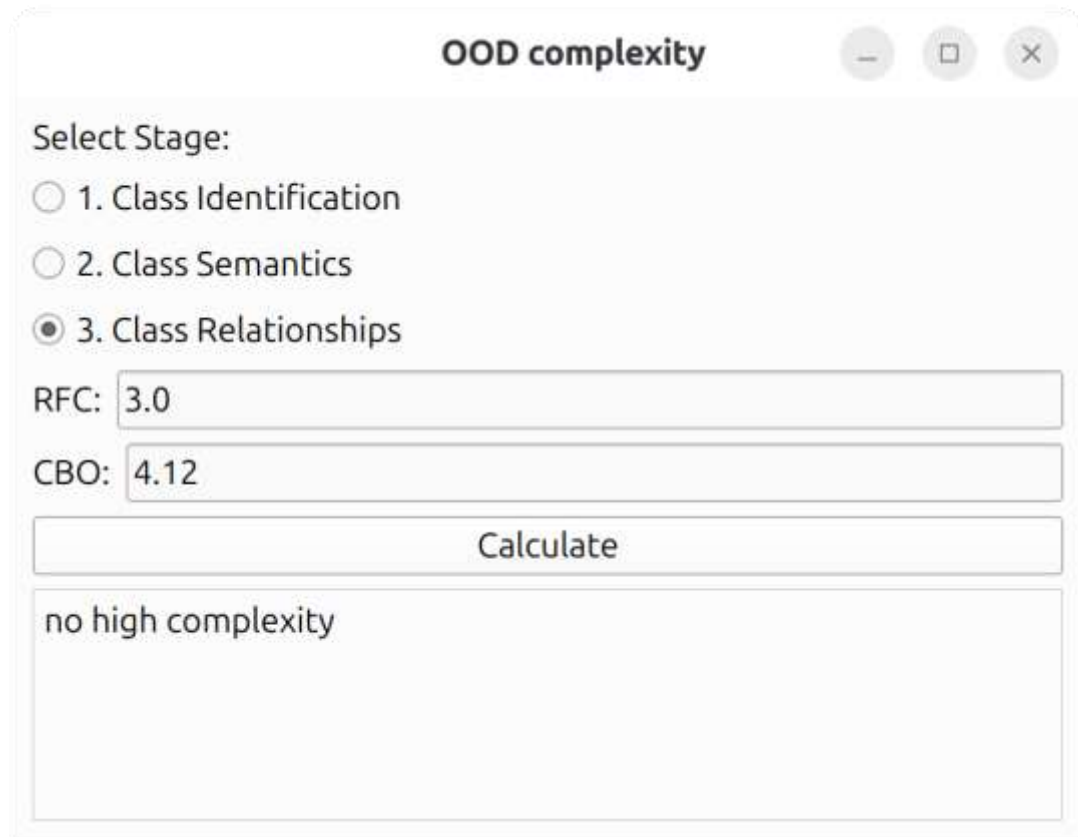
Крок 4.6. Якщо для точки даних  $i$  значення  $Z_1$  менше  $\hat{Z}_{RFC}$  та  $Z_2$  менше  $\hat{Z}_{CBO}$ , то складність ООП через зв'язки між класами  $i$ -го вебзастосунку не висока, інакше складність – висока.

За зазначеним алгоритмом мовою Python було створено відповідне ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням PHP-фреймворків, на третьому етапі ООП (через зв'язки між класами). Результати визначення складності на третьому етапі ООП для десяти реальних вебзастосунків за розробленим ПЗ наведені в табл. 4.3.

Таблиця 4.3 – Результати визначення складності на третьому етапі ООП через зв'язки між класами для десяти вебзастосунків

| $i$ | Назва застосунку | RFC   | CBO  | SMD  | складність |
|-----|------------------|-------|------|------|------------|
| 1   | Gymie            | 7,41  | 3,52 | 0,75 | no high    |
| 2   | Attendize        | 8,01  | 4,53 | 0,40 | high       |
| 3   | Cachet           | 6,57  | 4,47 | 1,23 | no high    |
| 4   | Crater           | 8,76  | 4,49 | 0,16 | high       |
| 5   | Craft            | 15,14 | 7,59 | 4,42 | high       |
| 6   | Crud             | 12,20 | 4,43 | 0,33 | high       |
| 7   | Sandbox          | 10,49 | 2,93 | 0,80 | high       |
| 8   | Wallabag         | 12,59 | 5,65 | 1,08 | high       |
| 9   | AdaptCMS         | 8,62  | 4,30 | 0,16 | no high    |
| 10  | Apiato           | 3,00  | 4,12 | 8,65 | no high    |

На рис.4.5 показаний результат визначення складності ООП для застосунку Apiato на трьох етапах ОПП у разі вибору третього етапу (результат: не висока складність).



**OOD complexity**

Select Stage:

☐ 1. Class Identification

☐ 2. Class Semantics

☒ 3. Class Relationships

RFC: 3.0

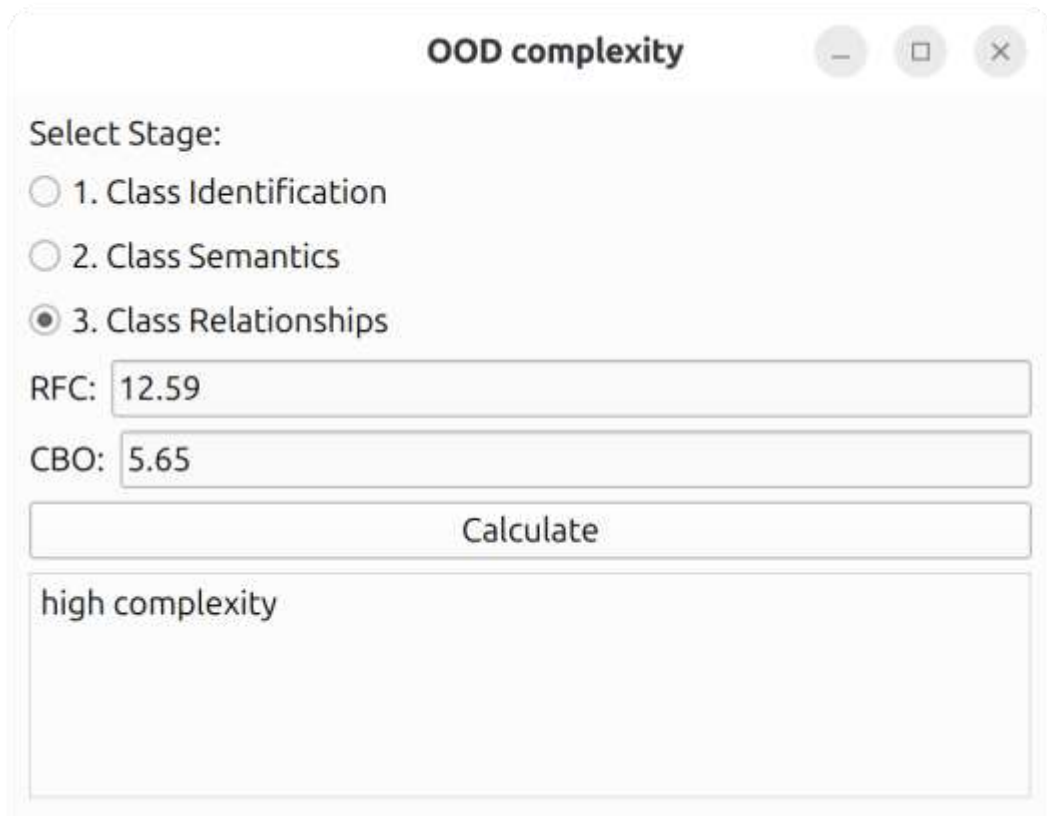
CBO: 4.12

Calculate

no high complexity

Рисунок 4.5. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору третього етапу (результат: не висока складність)

На рис.4.6 показаний результат визначення складності ООП для застосунку Wallabag на трьох етапах ОПП у разі вибору третього етапу (результат: висока складність).



OOD complexity

Select Stage:

☐ 1. Class Identification

☐ 2. Class Semantics

☒ 3. Class Relationships

RFC: 12.59

CBO: 5.65

Calculate

high complexity

Рисунок 4.6. – Графічне вікно ПЗ для визначення складності ООП на трьох етапах ОПП у разі вибору третього етапу (результат: висока складність)

Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих РНР-фреймворків, показало підвищення достовірності результатів обробки інформації з метрик СК цих Web застосунків: оцінки метрик СК більшості зазначених застосунків мали відносну похибку менше ніж 25%. Отримані результати свідчать про працездатність створеного ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на трьох етапах ОПП.

## 4.2 Інструментарій інформаційної технології для рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків

На основі розробленої математичної моделі для обробки інформації з трьох програмних метрик RFC, CBO і WMC створено інструментарій ІТ для рішення задач з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків, на основі метрик RFC, CBO та WMC.

Для цього була застосована удосконалена математична модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса.

Алгоритм визначення якості розробки Web застосунків, що розробляються з використанням PHP-фреймворків, на основі метрик RFC, CBO та WMC є таким.

Крок 1. Ввести значення метрик RFC, CBO та WMC для застосунку  $i$  (тривимірна точка даних  $i$ ).

Крок 2. Нормалізувати значення RFC, CBO та WMC (тривимірна точка даних  $i$ ) для застосунку  $i$  за допомогою тривимірного нормалізуючого перетворення Бокса – Кокса (2.1) з оцінками параметрів  $\hat{\lambda}_j$ , які наведені в таблиці 3.4. Нормалізовані метрики RFC, CBO та WMC позначено відповідно як  $Z_1$ ,  $Z_2$  та  $Z_3$ . Результат: значення  $Z_1$ ,  $Z_2$  та  $Z_3$ .

Крок 3. Обчислити квадрат відстані Махаланобіса (SMD) для тривимірної нормалізованої точки даних (точка  $i$ ) із використанням оберненої матриці (3.4) та вибірових середніх  $\bar{Z}_{RFC}$ ,  $\bar{Z}_{CBO}$  та  $\bar{Z}_{WMC}$ , що дорівнюють відповідно 2,384, 1,951 та 1,366.

Крок 4. Якщо значення квадрату відстані Махаланобіса для  $Z_1$ ,  $Z_2$  та  $Z_3$  буде більше за 13,95, то стоп та здійсніть вихід (ми не можемо використати рівняння (3.1)-(3.3) для цих значень), інакше перейти до кроку 5.

Крок 5. Обчисліть границі довірчих та інтервалів прогнозування нелінійних регресій для метрик RFC, CBO та WMC на рівні застосунку для тривимірної точки даних (точка  $i$ ) за рівняннями (3.2) та (3.3) із застосуванням в них обернених матриць (3.5)-(3.7) і значень  $S_{Z_y}$ , які дорівнюють 0,1393, 0,7112 та 0,1030 відповідно для метрик RFC, CBO та WMC.

Крок 6. Визначте, де знаходиться тривимірна точка даних  $i$ . Якщо ця точка даних для програмної системи знаходиться всередині всіх довірчих інтервалів нелінійних регресій для метрик RFC, CBO та WMC, тоді стоп (вебзастосунок має середню якість), інакше перейдіть до кроку 7.

Крок 7. Якщо тривимірна точка даних  $i$  для вебзастосунку знаходиться між верхніми границями довірчих інтервалів та нижніми межами інтервалів прогнозування для всіх трьох метрик, то вебзастосунок має високу якість, інакше він має низьку якість.

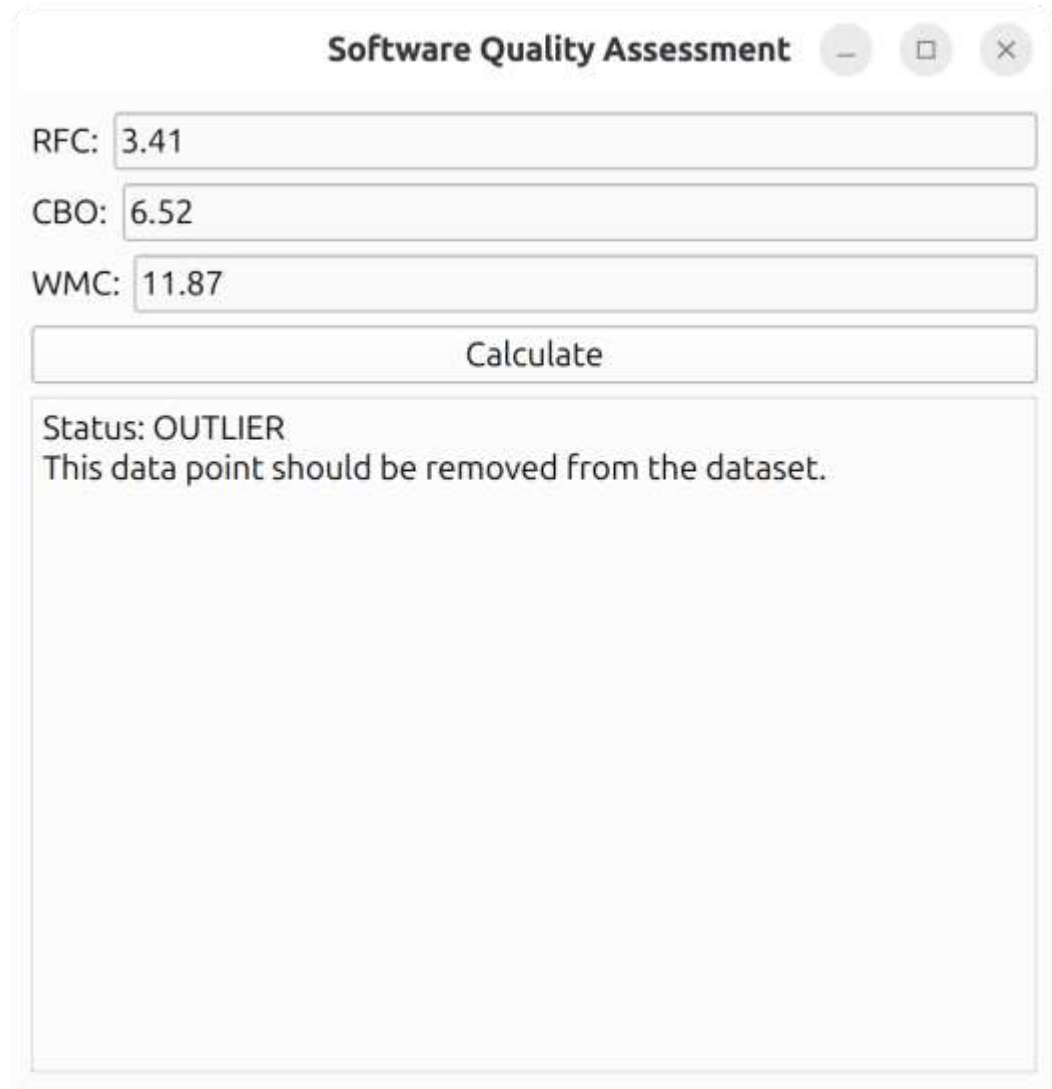
За зазначеним алгоритмом мовою Python було створено відповідне ПЗ для визначення якості розробки Web застосунків, що розробляються з використанням РНР-фреймворків, на основі метрик RFC, CBO та WMC. Результати визначення якості розробки Web застосунків, що розробляються з використанням РНР-фреймворків, на основі метрик RFC, CBO та WMC для восьми реальних вебзастосунків за розробленим ПЗ наведені в табл. 4.4. В табл. 4.4 також наведені

значення квадрату відстані Махаланобіса (SMD) для нормалізованих метрик RFC, CBO та WMC.

Таблиця 4.4 – Результати визначення якості розробки Web застосунків, що розроблялися з використанням PHP-фреймворків, на основі метрик RFC, CBO та WMC для восьми реальних вебзастосунків

| <i>i</i> | Назва застосунку | RFC   | CBO  | WMC   | SMD  | якість |
|----------|------------------|-------|------|-------|------|--------|
| 1        | Gymie            | 7,41  | 3,52 | 4,87  | 0,12 | high   |
| 2        | Attendize        | 8,01  | 4,53 | 5,35  | 0,33 | high   |
| 3        | Cachet           | 6,57  | 4,47 | 3,85  | 1,04 | high   |
| 4        | Crater           | 8,76  | 4,49 | 6,48  | 0,09 | medium |
| 5        | Craft            | 15,14 | 7,59 | 20,45 | 3,35 | low    |
| 6        | Crud             | 12,20 | 4,43 | 9,13  | 2,09 | low    |
| 7        | Sandbox          | 10,49 | 2,93 | 7,04  | 2,18 | low    |
| 8        | Wallabag         | 12,59 | 5,65 | 8,45  | 4,49 | low    |

На рис.4.7 зображено графічне вікно ПЗ для визначення якості розробки Web застосунків у разі, коли значення квадрату відстані Махаланобіса для нормалізованих метрик RFC, CBO та WMC буде більше за 13,95, тобто коли ми не можемо застосувати рівняння (3.1)-(3.3) для введених значень (крок 4 алгоритму).



Software Quality Assessment

RFC: 3.41

CBO: 6.52

WMC: 11.87

Calculate

Status: OUTLIER  
This data point should be removed from the dataset.

Рисунок 4.7. – Графічне вікно ПЗ для визначення якості розробки Web застосунків у разі, коли ми не можемо застосувати рівняння (3.1)-(3.3) для введених значень

На рис.4.8 наведено графічне вікно ПЗ для визначення якості розробки Web застосунків із результатом для застосунку Crater: середня якість – medium (крок 7 алгоритму).

The screenshot shows a window titled "Software Quality Assessment" with standard Windows window controls (minimize, maximize, close). Inside the window, there are three input fields for RFC, CBO, and WMC. Below these fields is a "Calculate" button. The results section displays the Mahalanobis D<sup>2</sup> value, the overall quality assessment (MEDIUM), and detailed confidence and prediction intervals for each metric.

| Metric | Value | Confidence Interval | Prediction Interval |
|--------|-------|---------------------|---------------------|
| RFC    | 8.76  | [8.3912, 8.9810]    | [6.7177, 11.1856]   |
| CBO    | 4.49  | [3.7592, 4.6764]    | [1.5528, 8.6189]    |
| WMC    | 6.48  | [6.2950, 7.1396]    | [4.3466, 11.4149]   |

Quality Assessment: MEDIUM

Рисунок 4.8. – Графічне вікно ПЗ для визначення якості розробки Web застосунків (результат: середня якість)

На рис.4.9 наведено графічне вікно ПЗ для визначення якості розробки Web застосунків із результатом для застосунку Cachet: висока якість – high.



**Software Quality Assessment**

RFC: 6.57

CBO: 4.47

WMC: 3.85

**Calculate**

Mahalanobis  $D^2 = 1.0418$

Quality Assessment: HIGH

RFC:  
Confidence Interval: [6.0958, 6.6919] - in  
Prediction Interval: [4.9151, 8.2746]

CBO:  
Confidence Interval: [3.3808, 4.5762] - in  
Prediction Interval: [1.4048, 8.2632]

WMC:  
Confidence Interval: [3.9488, 4.5124] - low  
Prediction Interval: [2.9381, 6.4764]

Рисунок 4.9. – Графічне вікно ПЗ для визначення якості розробки Web застосунків (результат: висока якість)

На рис.4.10 наведено графічне вікно ПЗ для визначення якості розробки Web застосунків із результатом для застосунку Wallabag: низька якість – low.

**Software Quality Assessment**

RFC: 12.59

CBO: 5.65

WMC: 8.45

**Calculate**

Mahalanobis  $D^2 = 4.4982$

Quality Assessment: LOW

RFC:  
Confidence Interval: [9.4945, 10.2856] - high  
Prediction Interval: [7.6518, 12.7257]

CBO:  
Confidence Interval: [3.3489, 5.3532] - high  
Prediction Interval: [1.5351, 8.9099]

WMC:  
Confidence Interval: [13.1556, 16.4997] - low  
Prediction Interval: [8.1402, 32.5876]

Рисунок 4.10. – Графічне вікно ПЗ для визначення якості розробки Web застосунків (результат: низька якість)

Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих PHP-фреймворків, показало підвищення достовірності результатів обробки інформації з програмних метрик RFC, CBO та WMC цих Web застосунків: оцінки метрик RFC, CBO та WMC більшості зазначених застосунків мали відносну похибку менше ніж 25%. Отримані результати свідчать про працездатність створеного ПЗ для визначення якості Web

застосунків, що розробляються з використанням PHP-фреймворків, на основі метрик RFC, CBO та WMC.

#### **4.3 Інструментарій інформаційної технології для рішення задачі з оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків**

На основі розроблених математичних моделей створено інструментарій ІТ для рішення задачі раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів: кількості класів, середньої кількості методів на клас і метрики DIT.

Для цього були застосовані три математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірних еліпсоїдів прогнозування для зазначених нормалізованих метрик, трифакторних нелінійних регресійних моделей, рівнянь для визначення границь довірчого інтервалу та інтервалу прогнозування кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса.

При цьому алгоритм раннього оцінювання кількості рядків коду  $Y$  Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, на основі кількості класів  $X_1$ , середньої кількості методів на клас  $X_2$  і метрики DIT  $X_3$  є таким.

Крок 1. Обрати тип фреймворку (CakePHP, Codeigniter або Yii) для застосунку  $i$ .

Крок 2. Обрати тип оцінювання  $Y$  ( $Y$  не є відомим; або  $Y$  є відомим)

Крок 3. Якщо значення  $Y$  є відомим, то перейти до кроку 3.1, інакше перейти до кроку 4.

Крок 3.1. Ввести значення метрик  $Y$ ,  $X_1$ ,  $X_2$ ,  $X_3$  та рівень значущості  $\alpha$ .

Крок 3.2. Нормалізувати значення  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  (чотиривимірна точка даних  $i$ ) для застосунку  $i$  за допомогою чотиривимірного нормалізуючого перетворення Бокса – Кокса (2.1) з параметрами які залежать від типу фреймворка (CakePHP, Codeigniter або Yii). Нормалізовані метрики  $Y$ ,  $X_1$ ,  $X_2$  та  $X_3$  позначено відповідно як  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$ . Результат: значення  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$ .

Крок 3.3. Обчислити квадрат відстані Махаланобіса (SMD) для чотиривимірної нормалізованої точки даних (точка  $i$ ).

Крок 3.4. Якщо квадрат відстані Махаланобіса для значень  $Z_Y$ ,  $Z_1$ ,  $Z_2$  та  $Z_3$  буде більше за значення тестової статистики, то стоп (ми не можемо використати нашу математичну модель для цих значень), інакше перейти до кроку 6. При обчисленні значення квадрату відстані Махаланобіса та визначенні значення тестової статистики треба враховувати тип фреймворка (CakePHP, Codeigniter або Yii).

Крок 4. Ввести значення метрик  $X_1$ ,  $X_2$ ,  $X_3$  та рівень значущості  $\alpha$ .

Крок 5. Нормалізувати значення  $X_1$ ,  $X_2$  та  $X_3$  (тривимірна точка даних  $i$ ) для застосунку  $i$  за допомогою чотиривимірного нормалізуючого перетворення Бокса – Кокса (2.1) з параметрами які залежать від типу фреймворка (CakePHP, Codeigniter або Yii). Результат: значення  $Z_1$ ,  $Z_2$  та  $Z_3$ .

Крок 6. Обчислити умовне вибіркове середнє  $\hat{Z}_Y$  за (3.14), оцінки параметрів якого залежать від типу фреймворка (CakePHP, Codeigniter або Yii).

Крок 7. Визначити умовне вибіркове середнє  $\hat{Y}$  за (1.8) з нульовим значенням  $\varepsilon$  та з урахуванням (3.14), оцінки параметрів якого залежать від типу фреймворка (CakePHP, Codeigniter або Yii).

Крок 8. Обчислити границі довірчих інтервалів та інтервалів прогнозування нелінійної регресії для  $Y$  для тривимірної точки даних (точка  $i$ ) за (3.2) та (3.3) з тою різницею, що вони будуються на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса, мають інші оцінки параметрів та  $\hat{Z}_Y$  знаходиться за (3.14) з урахуванням типу фреймворка (CakePHP, Codeigniter або Yii).

Крок 9. Змодельовати значення випадкової величини  $\varepsilon$  з розподілом Гауса з нульовим математичним сподіванням та оцінкою вибіркового середньоквадратичного відхилення, значення якої залежить від типу фреймворка (CakePHP, Codeigniter або Yii).

Крок 10. Визначити значення умовної випадкової величини  $Y$  за (1.8) з модельованим значенням  $\varepsilon$  та з урахуванням (3.14), оцінки параметрів якого залежать від типу фреймворка (CakePHP, Codeigniter або Yii).

Зауважимо, моделювання випадкової величини  $\varepsilon$  з розподілом Гауса здійснюється за методом Монте-Карло із безпосереднім застосуванням відомого методу виключень [137, 138].

За зазначеним алгоритмом мовою Python було створено відповідне ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів: кількості класів, середньої кількості методів на клас і метрики DIT.

На рис.4.11 наведено графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку CakePHP.

**Software Application Size Estimator**

Class count: 1496

Average methods per class: 4.26

DIT: 1.34

Select Framework:

☒ CakePHP

☐ CodeIgniter

☐ Yii

☐ Actual size of software is available

Calculate

Input Parameters:

Class count: 1496

Average methods per class: 4.26

DIT: 1.34

Framework: CakePHP

Results:

Y (kloc): 124.238

$\hat{Y}$  (kloc): 103.898

Confidence interval (kloc): (91.353, 118.207)

Prediction interval (kloc): (66.701, 162.511)

Рисунок 4.11. – Графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку CakePHP

На рис.4.12 наведено графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку Codeigniter.

**Software Application Size Estimator**

Class count: 366

Average methods per class: 10.66

DIT: 1.53

Select Framework:

☐ CakePHP

☒ CodeIgniter

☐ Yii

☐ Actual size of software is available

Calculate

Input Parameters:

Class count: 366

Average methods per class: 10.66

DIT: 1.53

Framework: CodeIgniter

Results:

Y (kloc): 51.325

$\hat{Y}$  (kloc): 59.541

Confidence interval (kloc): (47.404, 74.759)

Prediction interval (kloc): (34.902, 101.374)

Рисунок 4.12. – Графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку Codeigniter

На рис.4.13 наведено графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків, у разі вибору фреймворку Yii.

**Software Application Size Estimator**

Class count: 600

Average methods per class: 3.74

DIT: 1.84

Select Framework:

☐ CakePHP

☐ CodeIgniter

☒ Yii

☐ Actual size of software is available

Calculate

Input Parameters:

Class count: 600

Average methods per class: 3.74

DIT: 1.84

Framework: Yii

Results:

Y (kloc): 19.958

$\hat{Y}$  (kloc): 23.549

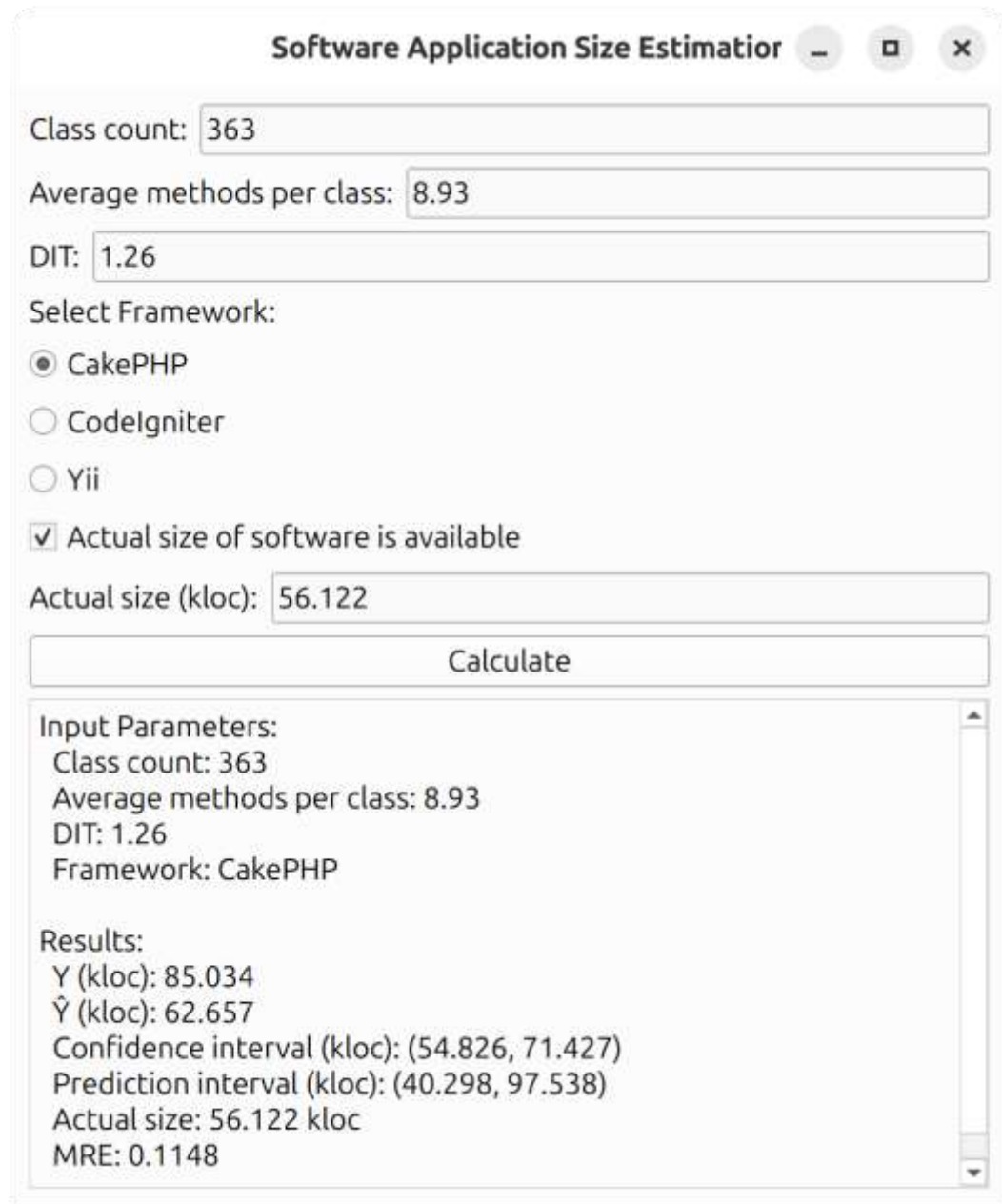
Confidence interval (kloc): (21.508, 25.79)

Prediction interval (kloc): (16.739, 33.243)

Рисунок 4.13. – Графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків, у разі вибору фреймворку Yii



На рис.4.14 наведено графічне вікно ПЗ для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку CakePHP та відомого розміру застосунку 56,122 KLOC.



**Software Application Size Estimator**

Class count: 363

Average methods per class: 8.93

DIT: 1.26

Select Framework:

☒ CakePHP

☐ CodeIgniter

☐ Yii

☒ Actual size of software is available

Actual size (kloc): 56.122

**Calculate**

**Input Parameters:**

Class count: 363

Average methods per class: 8.93

DIT: 1.26

Framework: CakePHP

**Results:**

Y (kloc): 85.034

$\hat{Y}$  (kloc): 62.657

Confidence interval (kloc): (54.826, 71.427)

Prediction interval (kloc): (40.298, 97.538)

Actual size: 56.122 kloc

MRE: 0.1148

Рисунок 4.14. – Графічне вікно ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, у разі вибору фреймворку CakePHP та відомого розміру застосунку 56,122 KLOC

На рис. 4.11-4.14 наведені результати раннього оцінювання кількості рядків коду реальних Web застосунків, що розроблялися з використанням різних PHP-фреймворків. Їх метрики та метрики деяких інших застосунків наведені в табл.4.5.

Таблиця 4.5 – Метрики Web застосунків, що розроблялися з використанням різних PHP-фреймворків

| $i$ | Назва застосунку | Фреймворк   | $Y$     | $X_1$ | $X_2$ | $X_3$ |
|-----|------------------|-------------|---------|-------|-------|-------|
| 1   | foodcoopshop     | CakePHP     | 19,650  | 269   | 4,47  | 1,28  |
| 2   | CandyCane        | CakePHP     | 56,122  | 363   | 8,93  | 1,26  |
| 3   | VamShop          | CakePHP     | 138,943 | 1496  | 4,26  | 1,34  |
| 4   | Map-OS           | CodeIgniter | 8,910   | 77    | 7,32  | 1,64  |
| 5   | No-CMS           | CodeIgniter | 56,687  | 366   | 10,66 | 1,53  |
| 6   | FUEL-CMS         | CodeIgniter | 69,300  | 507   | 8,43  | 1,38  |
| 7   | Yii2 Super Blog  | Yii         | 1,053   | 25    | 4,92  | 1,82  |
| 8   | SkeekS CMS       | Yii         | 28,397  | 600   | 3,74  | 1,84  |
| 9   | HumHub           | Yii         | 51,257  | 1226  | 4     | 1,83  |

Застосунок foodcoopshop (<https://github.com/foodcoopshop/foodcoopshop>) – це ПЗ з відкритим вихідним кодом для продовольчих кооперативів та місцевих магазинів. CandyCane (<https://github.com/yandod/candycane>) – це система відстеження проблем, порт Redmine на CakePHP з Ruby на Rails. VamShop (<https://github.com/vamcart/vamcart>) – це рішення для інтернет-магазинів, яке пропонує широкий спектр готових функцій, що дозволяє досить швидко та легко налаштувати інтернет-магазини. Map-OS – це система контролю замовлень на послуги (<https://github.com/RamonSilva20/mapos>). No-CMS

(<https://github.com/goFrendiAsgard/No-CMS>) – це фреймворк CMS на базі фреймворку CodeIgniter. FUEL CMS – це система управління контентом на базі фреймворку CodeIgniter. Yii2 Super Blog (<https://github.com/akiraz2/yii2-blog?tab=readme-ov-file>) – це простий, налаштований модуль Yii2 з фронтендом та бекендом, клонований з невідтримуваного репозиторію funson86/yii2-blog (<https://github.com/funson86/yii2-blog>), повністю реорганізований та покращений. SkeekS CMS (<https://github.com/skeeks-cms/cms>) – сучасна універсальна система управління контентом на базі PHP-фреймворку Yii2. NumHub – це корпоративна соціальна мережа з відкритим кодом.

Результати оцінювання кількості рядків коду реальних Web застосунків, метрики яких зазначені в табл. 4.5, наведені в табл. 4.6.

Таблиця 4.6 – Результати оцінювання кількості рядків коду реальних Web застосунків, метрики яких зазначені в табл. 4.5

| $i$ | $Y$     | $\hat{Y}$ | $RME$  | Довірчий інтервал |         | Інтервал прогнозування |         |
|-----|---------|-----------|--------|-------------------|---------|------------------------|---------|
|     |         |           |        | LB                | UB      | LB                     | UB      |
| 1   | 19,650  | 22,325    | 0,1362 | 20,195            | 24,686  | 14,627                 | 34,205  |
| 2   | 56,122  | 62,567    | 0,1148 | 54,826            | 71,427  | 40,298                 | 97,538  |
| 3   | 138,943 | 103,898   | 0,2522 | 91,353            | 118,207 | 66,701                 | 162,511 |
| 4   | 8,910   | 8,303     | 0,0681 | 7,302             | 9,440   | 5,006                  | 13,749  |
| 5   | 56,687  | 59,541    | 0,0503 | 47,404            | 74,759  | 34,902                 | 101,374 |
| 6   | 69,300  | 78,444    | 0,1319 | 65,112            | 94,483  | 46,787                 | 131,280 |
| 7   | 1,053   | 1,130     | 0,0734 | 1,047             | 1,220   | 0,828                  | 1,547   |
| 8   | 28,397  | 23,549    | 0,1707 | 21,508            | 25,790  | 16,739                 | 33,243  |
| 9   | 51,257  | 49,616    | 0,0320 | 44,466            | 55,381  | 34,836                 | 70,927  |

Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, показало підвищення достовірності результатів обробки інформації з програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) цих Web застосунків: оцінки розміру більшості зазначених застосунків мали відносну похибку менше ніж 25%. Отримані результати свідчать про працездатність створеного ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів: кількості класів, середньої кількості методів на клас і метрики DIT.

Працездатність запропонованого інструментарію ІТ також підтверджується його практичним використанням в наступних установах: ТОВ «АЙПІ РІШЕННЯ», ТОВ НВЦ «Діагностика та контроль», Комунальне підприємство «Міський інформаційно-обчислювальний центр» (м. Миколаїв). Використання інструментарію ІТ для визначення складності ООП, якості розробки та кількості рядків коду вебзастосунків, дозволило підвищити достовірність визначення складності ООП та якості розробки відповідних вебзастосунків.

Результати дисертаційної роботи також застосовуються в навчальному процесі на кафедрах Комп'ютерних систем та технологій, Математичного забезпечення комп'ютерних систем Одеського національного університету імені І. І. Мечникова. Акти про впровадження результатів дисертаційної роботи наведені у додатках А і Б.

#### 4.4 Висновки за розділом 4

У четвертому розділі дисертації на підставі розроблених математичних моделей створено інструментарій ІТ (ПЗ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків. Для цього використовувалася мова програмування Python. Було створено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

1) Створено інструментарій ІТ (ПЗ) для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на трьох етапах ОПП. Основою побудови пропонованого інструментарію (ПЗ) послужили відповідні математичні моделі, розроблені автором. Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих РНР-фреймворків, показало підвищення достовірності результатів обробки інформації з метрик СК цих Web застосунків: оцінки метрик СК більшості зазначених застосунків мали відносну похибку менше ніж 25%. Отримані результати свідчать про працездатність створеного ПЗ для визначення складності ООП Web застосунків, що розробляються з використанням РНР-фреймворків, на трьох етапах ОПП.

2) Створено інструментарій ІТ (ПЗ) для визначення якості Web застосунків, що розробляються з використанням РНР-фреймворків, на основі метрик RFC, CBO та WMC. Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих РНР-фреймворків, показало підвищення достовірності результатів обробки інформації з програмних метрик RFC, CBO та WMC цих Web застосунків: оцінки метрик RFC, CBO та WMC більшості зазначених застосунків мали відносну похибку менше ніж 25%.

Отримані результати свідчать про працездатність створеного ПЗ для визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків, на основі метрик RFC, CBO та WMC.

3) Створено інструментарій IT (ПЗ) для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii. Застосування цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, показало підвищення достовірності результатів обробки інформації з програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) цих Web застосунків: оцінки розміру більшості зазначених застосунків мали відносну похибку менше ніж 25%. Отримані результати свідчать про працездатність створеного ПЗ для раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів: кількості класів, середньої кількості методів на клас і метрики DIT.

Результати досліджень автора, які наведені у поточному розділі, опубліковано в [31-34, 70-72, 78-85].

## ВИСНОВКИ

В дисертаційній роботі вирішено важливе науково-практичне завдання підвищення достовірності результатів обробки інформації з програмних метрик Web застосунків, що розробляються з використанням РНР-фреймворків, за рахунок побудови певних математичних моделей для обробки відповідної інформації та створенню на їх основі інструментарію інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням РНР-фреймворків.

В процесі проведення дисертаційних досліджень одержані наступні результати.

1) На основі здійсненого аналізу існуючих методів і моделей для обробки інформації з програмних метрик та можливості використання існуючих моделей для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ саме для Web застосунків, що розроблені з використанням РНР-фреймворків, виконано обґрунтування необхідності проведення досліджень за обраною темою. Удосконалювати існуючі математичні моделі для обробки інформації з програмних метрик та відомі за їх допомогою рішення задач з визначення складності ООП, якості розробки та кількості рядків коду ПЗ доцільно саме на основі багатовимірних нормалізуючих перетворень, які дозволяють враховувати кореляцію між програмними метриками та відхилення їх багатовимірного розподілу від нормального.

2) Здійснено вибір перетворень для нормалізації двовимірних, тривимірних та чотиривимірних наборів даних з програмних метрик Web

застосунків, що розробляються з використанням PHP-фреймворків. У якості нормалізуючих перетворень були обрані двовимірні, тривимірні та чотиривимірні перетворення Бокса – Кокса, після застосування яких розподіли відповідних нормалізованих даних не відхиляються від нормального згідно критерію Мардія. Здійснено оцінювання параметрів зазначених перетворень Бокса – Кокса за методом максимальної правдоподібності.

3) Вперше побудовано дві математичні моделі для обробки інформації з двох груп програмних метрик (відповідно WMC, DIT і NOC та WMC, RFC і LCOM) Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірних еліпсоїдів прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ці дві математичні моделі дозволяють підвищити достовірність результатів обробки інформації з двох груп метрик WMC, DIT і NOC та WMC, RFC і LCOM за рахунок врахування як кореляції між ними, так і відхилення тривимірних розподілів цих метрик від нормального, та здійснити на їх основі рішення задачі з визначення складності ООП на його першому та другому етапах (через ідентифікацію та семантику класів) за Бучем. Зокрема, у порівнянні з SVR побудовані моделі дозволили підвищити кількість прогнозованих значень метрик WMC, DIT і NOC, для яких відносна похибка була менше 0,25, відповідно на 6,2%, 4,0% та 1,2%.

4) Удосконалено математичну модель для обробки інформації з програмних метрик RFC і CBO Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді еліпсу прогнозування і рівнянь регресії для зазначених нормалізованих метрик на основі двовимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC і CBO за рахунок врахування як кореляції між ними, так і відхилення двовимірного



розподілу цих метрик від нормального, та здійснити на її основі рішення задачі з визначення складності ООП на його третьому етапі (через зв'язки між класами) за Бучем.

5) Удосконалено математичну модель для обробки інформації з трьох програмних метрик RFC, CBO і WMC Web застосунків, що розробляються з використанням PHP-фреймворків, у вигляді тривимірного еліпсоїду прогнозування для зазначених нормалізованих метрик і рівнянь для визначення границь довірчих інтервалів та інтервалів прогнозування трьох нелінійних регресій програмних метрик RFC, CBO і WMC на основі тривимірного нормалізуючого перетворення Бокса – Кокса. Ця математична модель дозволяє підвищити достовірність результатів обробки інформації з метрик RFC, CBO і WMC за рахунок врахування як кореляції між ними, так і відхилення тривимірного розподілу цих метрик від нормального, та здійснити на її основі рішення задачі з визначення якості Web застосунків, що розробляються з використанням PHP-фреймворків. Зокрема, для тестового набору даних у порівнянні з SVR удосконалена модель дозволила підвищити кількість прогнозованих значень метрик RFC, CBO і WMC, для яких відносна похибка була менше 0,25, відповідно на 76,0%, 77,8% та 81,8%.

6) Удосконалено три математичні моделі для обробки інформації з чотирьох програмних метрик (кількості рядків коду, кількості класів, середньої кількості методів на клас і метрики DIT) у вигляді чотиривимірних еліпсоїдів прогнозування для зазначених нормалізованих метрик та трифакторних нелінійних регресійних моделей для оцінювання кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків, на основі чотиривимірного нормалізуючого перетворення Бокса – Кокса. Ці математичні моделі дозволяють підвищити достовірність результатів обробки інформації із зазначених метрик за рахунок врахування як кореляції між ними, так і відхилення

чотирирівимірного розподілу цих метрик від нормального, та здійснити на їх основі рішення задачі раннього оцінювання кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii, за трьома метриками діаграми класів. Отримані трифакторні нелінійні регресійні моделі в порівнянні з іншими регресійними моделями мають більший відсоток прогнозованих значень PRED, менше значення середньої величини відносної похибки MMRE та менші ширини довірчого інтервалу та інтервалу прогнозування нелінійної регресії. Зокрема, у порівнянні з SVR удосконалена модель для фреймворку Codeigniter дозволила збільшити значення PRED(0,25) на 17,0%, а значення MMRE зменшити на 29,1%.

7) Удосконалено рівняння для визначення границь довірчого інтервалу та інтервалу прогнозування трифакторних нелінійних регресій кількості рядків коду Web застосунків, що розробляються з використанням різних PHP-фреймворків, в залежності від кількості класів, середньої кількості методів на клас і метрики DIT на основі чотирирівимірного нормалізуючого перетворення Бокса – Кокса. Рівняння для визначення нижньої та верхньої границь довірчого інтервалу та інтервалу прогнозування дозволяють здійснювати оцінювання точності умовного вибіркового середнього та меж моделювання умовної випадкової величини кількості рядків коду Web застосунків, що розробляються з використанням таких відомих PHP-фреймворків як CakePHP, Codeigniter і Yii.

8) На підставі розроблених математичних моделей створено інструментарій інформаційної технології (ІТ) для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків. Для цього використовувалася мова програмування Python. Було створено відповідне ПЗ для рішення задач з визначення складності ООП, якості розробки та кількості рядків коду Web застосунків, що розробляються з використанням PHP-фреймворків. Застосування

цього ПЗ до низки реальних Web застосунків, які були розроблені з використанням відомих PHP-фреймворків (CakePHP, Codeigniter, Laravel, Symfony та Yii), показало його працездатність.

9) Результати дисертаційної роботи впроваджено в ТОВ «АЙП РІШЕННЯ», ТОВ НВЦ «Діагностика та контроль», Комунальне підприємство «Міський інформаційно-обчислювальний центр» (м. Миколаїв), ФОП Білецький В.В., що підтверджується відповідними актами, які наведені у додатку А. Також результати дисертаційної роботи застосовуються в навчальному процесі на кафедрах Комп'ютерних систем та технологій, Математичного забезпечення комп'ютерних систем Одеського національного університету імені І. І. Мечникова (акт наведений у додатку Б).

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Geetha V., Sujatha N. A survey on divergent classification of social media networking. *Computing, Communication, and Intelligent Systems (ICCCIS)* : proceedings of the 2022 International Conference. Greater Noida, India: IEEE, 2022. P. 203-207. URL: <https://doi.org/10.1109/ICCCIS56430.2022.10037606>
2. Al-Duwairi B., Shatnawi A. S., Jaradat H., Al-Musa A., Al-Awadat H. On the digital forensics of social networking Web-based applications. *Digital Forensics and Security (ISDFS)* : proceedings of the 2022 10th International Symposium. Istanbul, Turkey: IEEE, 2022. P. 1-6. URL: <https://doi.org/10.1109/ISDFS55398.2022.9800839>
3. Flávio Gomes da Silva Lisboa, Enrico Zimuel. PHP Web Development with Laminas: Build a fully secured and functional e-commerce application with PHP using the Laminas framework. Packt Publishing, 2022. 394 p. URL: <https://ieeexplore.ieee.org/document/10162405>
4. Aggarwal P. K., Sharma R., Khare R., Singh S. E-commerce Application using PHP and Web Development: A Review. *Disruptive Technologies (ICDT)* : proceedings of the 2023 International Conference. Greater Noida, India: IEEE, 2023. P. 755-758. URL: <https://doi.org/10.1109/ICDT57929.2023.10151228>
5. Babu A. C., Teja V. N. C. S., Reddy A. D., Kumar E. N., Srinivas V. Web Based Hospital Management System. *Advanced Computing and Communication Systems (ICACCS)* : proceedings of the 2023 9th International Conference. Coimbatore, India: IEEE, 2023. P. 1109-1113. URL: <https://doi.org/10.1109/ICACCS57279.2023.10112962>

6. Casquilho M., Da Silva P. M., Caridade C. M. R., De Miranda J. L. A web computing tool for a SIR model: immunity loss and vaccination. *Bioinformatics and Biomedicine (BIBM)* : proceedings of the 2024 IEEE International Conference. Lisbon, Portugal: IEEE, 2024. P. 6761-6765. URL: <https://doi.org/10.1109/BIBM62325.2024.10822515>
7. Alfayez Z. H., Hassan I. M. Design and Implement a Web-Based Learning System to Provide E-learning in the University of Basrah During Crises. *Engineering Technology and its Applications (IICETA)* : proceedings of the 2022 5th International Conference. Al-Najaf, Iraq: IEEE, 2022. P. 277-282. URL: <https://doi.org/10.1109/IICETA54559.2022.9888328>
8. Abdullah F. R. A., Nuha H. H., Utomo R. G., Afasyah A. D. Implementation of User and Article Module Design and Testing on Innovation Dashboard Backend Using PHP Laravel and RESTful API. *Decision Aid Sciences and Applications (DASA)* : proceedings of the 2024 International Conference. Manama, Bahrain: IEEE, 2024. P. 1-6. URL: <https://doi.org/10.1109/DASA63652.2024.10836333>
9. Ratnawati A., Kartikasari E., Rustandi B., Susanto B., Setiawan A., Munastha K. A. A Web-Based Accounting Information System Application using CodeIgniter Framework: (A Case Study Approach). *Telecommunication Systems, Services, and Applications (TSSA)* : proceedings of the 2022 16th International Conference. Lombok, Indonesia: IEEE, 2022. P. 1-6. URL: <https://doi.org/10.1109/TSSA56819.2022.10063881>
10. Рейтинг мов програмування 2025. TypeScript і Python — найпопулярніші, частки C# та Java зменшуються. URL: <https://dou.ua/lenta/articles/language-rating-2025/> (дата звернення: 20.12.2025).
11. TIOBE Index for December 2025. December Headline: Programming language R is back in the top 10. URL: <https://www.tiobe.com/tiobe-index/> (дата звернення: 20.12.2025).

12. Swedheetha C., Preethi S. C, R., V. V. T B J, Maryam S., Manikandan P. Evaluating and Contrasting the Performance of Web Development Technologies. *Communication and Electronics Systems (ICCES)* : proceedings of the 2024 9th International Conference. Coimbatore, India: IEEE, 2024). P. 1020-1024. URL: <https://doi.org/10.1109/ICCES63552.2024.10859720>
13. Madaehoh A., Senivongse T. OSS-AQM: An open-source software quality model for automated quality measurement. *Data and Software Engineering (ICoDSE)* : proceedings of the 2022 International Conference. Denpasar, Indonesia: IEEE, 2022. P. 126-131. URL: <https://doi.org/10.1109/ICoDSE56892.2022.9972135>
14. Bombiri O., Poda P., Ouedraogo T. F. Application of machine learning in software quality: a Mini-review. *Natural and Engineering Sciences for Sahel's Sustainable Development (MNE3SD)* : proceedings of the 2023 IEEE Multi-conference. Bobo-Dioulasso, Burkina Faso: IEEE, 2023. P. 1-7. URL: <https://doi.org/10.1109/MNE3SD57078.2023.10079800>
15. Gao C., Luo W., Wang J. et al. Software quality evaluation model based on multiple linear regression and fuzzy comprehensive evaluation method. *Dependable Systems and Their Applications (DSA)* : proceedings of the 2022 9th International Conference. Wulumuqi, China: IEEE, 2022. P. 383-389. URL: <https://doi.org/10.1109/DSA56465.2022.00058>
16. Deshpande M. V., Soitkar P. A., Tripathi D. R., Agarmore S. B. Ensuring web application quality: The role of software testing as a form of quality assurance. *ICT in Business Industry & Government (ICTBIG)* : proceedings of the 2023 IEEE International Conference. Indore, India: IEEE, 2023. P. 1-6. URL: <https://doi.org/10.1109/ICTBIG59752.2023.10456277>
17. Fizza K., Banerjee A., Jayaraman P. P. et al. A survey on evaluating the quality of autonomic Internet of things applications. *IEEE Communications Surveys &*

*Tutorials.* 2023. Vol. 25, No. 1. P. 567-590. URL: <https://doi.org/10.1109/COMST.2022.3205377>

18. Shyamal D. K. K., Asanka P. P. G. D., Wickramaarachchi D. A comprehensive approach to evaluating software code quality through a flexible quality model. *Smart Computing and Systems Engineering (SCSE)* : proceedings of the 2023 International Research Conference. Kelaniya, Sri Lanka: IEEE, 2023. P. 1-8. URL: <https://doi.org/10.1109/SCSE59836.2023.10215004>

19. Prykhodko S. Evaluating quality of software systems by the confidence and prediction intervals of regressions for RFC, CBO, and WMC metrics. *WSEAS Transactions on Systems.* 2024. Vol. 23. P. 322-330. URL: <https://doi.org/10.37394/23202.2024.23.36>

20. Chidamber S. R., Kemerer C. F. A metrics suite for object-oriented design. *IEEE Transactions on Software Engineering.* 1994. Vol. 20, No. 6. P. 476-493. URL: <https://doi.org/10.1109/32.295895>

21. Khan S. A., Khan R. A. Object oriented design complexity quantification model. *Procedia Technology.* 2012. Vol. 4. P. 548-554. URL: <https://doi.org/10.1016/j.protcy.2012.05.087>

22. Prykhodko S., Prykhodko N., Smykodub T. A joint statistical estimation of the RFC and CBO metrics for open-source applications developed in Java. *Computer Sciences and Information Technologies (CSIT)* : proceedings of the 2022 IEEE 17th International Conference. Lviv, Ukraine: IEEE, 2022. P. 442-445. URL: <https://doi.org/10.1109/CSIT56902.2022.10000457>

23. Aluthwaththage J. H., Thathsarani H. A. N. N. Beyond Traditional Metrics: Advancing Software Complexity Analysis with the ACB Metric. *Computing Communication and Networking Technologies (ICCCNT)* : proceedings of the 2024 15th International Conference. Kamand, India: IEEE, 2024. P. 1-6. URL: <https://doi.org/10.1109/ICCCNT61001.2024.10724880>

24. Jayakuru H. E., Nikeshala H. A. N. Complexity Estimator: A Comprehensive Tool for Measuring Software Complexity in Object-Oriented Systems. *Smart Computing and Systems Engineering (SCSE)* : proceedings of the 2025 International Research Conference. Colombo, Sri Lanka: IEEE. 2025. P. 1-6. URL: <https://doi.org/10.1109/SCSE65633.2025.11030957>

25. Kaczmarek J., Kucharski M. Size and effort estimation for applications written in Java. *Information and Software Technology*. 2004. Vol. 46, Issue 9. P. 589–601. URL: <https://doi.org/10.1016/j.infsof.2003.11.001>

26. Tan H. B. K., Zhao Y., Zhang H. Conceptual data model-based software size estimation for information systems. *Transactions on Software Engineering and Methodology*. 2009. Vol. 19, Issue 2. P. 1–37. URL: <https://doi.org/10.1145/1571629.1571630>

27. Zifen Y. An improved software size estimation method based on object-oriented approach. *Electrical & Electronics Engineering (EEESYM)* : proceedings of the 2012 IEEE Symposium. Kuala Lumpur, Malaysia: IEEE, 2012. P. 615–617. URL: <https://doi.org/10.1109/EEESym.2012.6258733>

28. Prykhodko N. V., Prykhodko S. B. The non-linear regression model to estimate the software size of open source Java-based systems. *Radio Electronics, Computer Science, Control*. 2018. No. 3. P. 158–166. URL: <https://doi.org/10.15588/1607-3274-2018-3-17>

29. Manisha, & Rishi, R. Early size estimation using machine learning. *Computing for Sustainable Global Development (INDIACom)* : proceedings of the 2021 8th International Conference. New Delhi, India: IEEE, 2021. P. 757–762. URL: <https://ieeexplore.ieee.org/document/9441329>

30. Приходько С. Б., Приходько Н. В., Ворона М. В., Бєловол І. О. Нелінійна регресійна модель для оцінювання розміру Web-застосунків, що створюються з використанням фреймворку Laravel. *Інформаційні технології та*



комп'ютерна інженерія. 2021. Т. 50. № 1. С. 115-121. URL: <http://dx.doi.org/10.31649/1999-9941-2021-50-1-115-121>

31. Prykhodko S., **Prykhodko A.**, Shutko I. Estimating the Size of Web Apps Created Using the CakePHP Framework by Nonlinear Regression Models with Three Predictors. *Computer Sciences and Information Technologies (CSIT)* : proceedings of the 2021 IEEE 16th International Conference. Lviv, Ukraine: IEEE, 2021. P. 333-336. URL: <http://dx.doi.org/10.1109/CSIT52700.2021.9648680>

32. Prykhodko S. B., Shutko I. S., **Prykhodko A. S.** A nonlinear regression model to estimate the size of web apps created using the CakePHP framework. *Radio Electronics, Computer Science, Control*. 2021. No. 4 (59). P. 129-139. URL: <http://dx.doi.org/10.15588/1607-3274-2021-4-12>

33. Prykhodko S., Shutko I., **Prykhodko A.** Early LOC Estimation of Web Apps Created Using Yii Framework by Nonlinear Regression Models. *WSEAS Transactions on Computers*. 2021. Vol. 20. P. 321-328. URL: <http://dx.doi.org/10.37394/23205.2021.20.35>

34. Prykhodko S. B., Shutko I. S., **Prykhodko A. S.** Early size estimation of web apps created using Codeigniter framework by nonlinear regression models. *Radioelectronic and computer systems*. 2022. Vol. 3 (103). P. 84-94. URL: <https://doi.org/10.32620/reks.2022.3.06>

35. Daud M., Malik A. A. Construction and validation of early software size estimation models based on ADAF-adjusted ACD metrics. *The Computer Journal*. 2023. Vol. 66. Issue 9. P. 2123–2137. URL: <https://doi.org/10.1093/comjnl/bxac065>

36. 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. URL: <https://doi.org/10.1109/ACCESS.2025.3540548>

37. Приходько С., Шутко І. Регресійні моделі для раннього оцінювання кількості рядків коду вебзастосунків, що створюються за допомогою фреймворку

Codeigniter. *Прикладні питання математичного моделювання*. 2025. Том 8, № 1. С. 189-196. URL: <https://doi.org/10.32782/mathematical-modelling/2025-8-1-18>

38. Гунченко Ю., Мещеряков В., Приходько С., Рачинська А. Раннє оцінювання кількості рядків коду останніх випусків застосунків з відкритим кодом за метриками їх перших випусків. *Information Technology: Computer Science, Software Engineering and Cyber Security*. 2025. Вип. 4. С. 72-78. URL: <https://doi.org/10.32782/IT/2025-4-9>

39. Гунченко Ю., Мещеряков В., Рачинська А., Приходько С. Регресійні моделі для раннього оцінювання кількості рядків коду різних випусків застосунків з відкритим кодом. *Прикладні питання математичного моделювання*. 2025. Том 8. № 2. С. 86-92. URL: <https://doi.org/10.32782/mathematical-modelling/2025-8-2-9>

40. Prykhodko S., Prykhodko N., Makarova L., Pukhalevych A. Application of the Squared Mahalanobis Distance for Detecting Outliers in Multivariate Non-Gaussian Data. *Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) : proceedings of the 2018 14th International Conference*. Lviv-Slavske, Ukraine: IEEE, 2018. P. 962-965. URL: <https://doi.org/10.1109/TCSET.2018.8336353>

41. Molnar, A. J., Neamțu, A., & Motogna, S. Evaluation of software product quality metrics. In E. Damiani, G. Spanoudakis, & L. Maciaszek (eds). *Evaluation of Novel Approaches to Software Engineering. ENASE 2019. Communications in Computer and Information Science*: 2020. Vol. 1172. P. 163-187. Springer, Cham. URL: [https://doi.org/10.1007/978-3-030-40223-5\\_8](https://doi.org/10.1007/978-3-030-40223-5_8)

42. Кольцов А. В. Математичні моделі для визначення викидів у двовимірних даних із метрик СВО та RFC програмних застосунків з відкритим кодом на Kotlin. *Прикладні питання математичного моделювання*. 2025. Том 8. № 1. С. 89-96. URL: <https://doi.org/10.32782/mathematical-modelling/2025-8-1-8>

43. Prykhodko S., Prykhodko N., Makarova L. Estimating the Software Size of Open-Source PHP-Based Systems Using Non-Linear Regression Analysis. *Advanced Computer Information Technologies (ACIT-2018)* : proceedings of International Conference. CEUR Workshop Proceedings. 2019. Vol.2300. Ceske Budejovice, CZECH REPUBLIC. P. 199-202. URL: <http://ceur-ws.org/Vol-2300/Paper48.pdf>

44. Приходько С. Б., Приходько Н. В., Фаріонова Т. А., Ворона М. В. Трьохфакторна нелінійна регресійна модель для оцінювання розміру Php-застосунків з відкритим кодом. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2020. Том 31 (70). Ч. 1. № 1. С. 124-131. URL: <https://doi.org/10.32838/2663-5941/2020.1-1/23>

45. Приходько С. Б., Приходько Н. В., Смикодуб Т. Г. Чотирьохфакторна нелінійна регресійна модель для оцінювання розміру Java-застосунків з відкритим кодом. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2020. Том 31 (70). Ч. 1. № 2. С. 157-162. URL: <https://doi.org/10.32838/2663-5941/2020.2-1/25>

46. Приходько С. Б., Ворона М. В. Оцінювання розміру PHP-застосунків з відкритим кодом за нелінійними регресійними моделями з різними факторами / *Збірник наукових праць Національного університету кораблебудування імені адмірала Макарова*. 2021. №1. С. 92-98. URL: [https://doi.org/10.15589/znp2021.1\(484\).13](https://doi.org/10.15589/znp2021.1(484).13)

47. Ворона М. В. Математичні моделі та інформаційна технологія для оцінювання розміру програмних застосунків з відкритим кодом на PHP : дис. ... д-ра філос. : 122 «Комп'ютерні науки» / Національний університет кораблебудування імені адмірала Макарова. Миколаїв, 2021. 153 с.

48. Prykhodko S., Prykhodko N. Estimating Quality of Open-Source Kotlin-Based Apps by the Confidence and Prediction Intervals of Nonlinear Regression for RFC Metric. *Computer Sciences and Information Technologies (CSIT)* : proceedings

of the 2023 IEEE 18th International Conference. Lviv, Ukraine: IEEE, 2023. P. 1-4. URL: <https://doi.org/10.1109/CSIT61576.2023.10324187>

49. Макарова Л.М., Латанська Л.О. Порівняльний аналіз нелінійних регресійних моделей для прогнозування розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2023. Том 34 (73) № 6. С.112-117. URL: <https://doi.org/10.32782/2663-5941/2023.6/17>

50. Prykhodko S. B., Prykhodko N. V., Koltsov A. V. A nonlinear regression model for early LOC estimation of open-source Kotlin-based applications. *Radio Electronics, Computer Science, Control*. 2024. No.1 (68). P. 85-95. URL: <https://doi.org/10.15588/1607-3274-2024-1-8>

51. Фаріонова Т. А., Орєхов О. С. Математичні моделі для оцінювання розміру Java-застосунків. *Вісник Херсонського національного технічного університету*. 2024. No. 2 (89). С. 196–203. URL: <https://doi.org/10.35546/kntu2078-4481.2024.2.28>

52. Oriekhov O., Farionova T., Chernova L. S., Chernova L., Vorona M. Nonlinear regression models for software size estimation of Data Science and Machine Learning Java-applications. *CEUR Workshop Proceedings, Germany*. 2024. Vol. 3709. P. 54-66. URL: <https://ceur-ws.org/Vol-3709/paper5.pdf>

53. Oriekhov O., Farionova T., Chernova L. Three-factor nonlinear regression model of estimating the size of java-software. *CEUR Workshop Proceedings, Germany*, 2024. Vol. 3790. P. 506-518. URL: <https://ceur-ws.org/Vol-3790/paper44.pdf>

54. Oriekhov O., Farionova T., Chernova L., Vorona, M. A Five-Factor Nonlinear Regression Model for JAVA Applications Size Estimation. *IT Project Management (ITPM): proceedings of the 2025 6th International Workshop*. Bratislava,

Slovakia, May 22, 2025. *CEUR Workshop Proceedings*. 2025. Vol. 4023. P. 1-10.  
URL: <https://ceur-ws.org/Vol-4023/paper1.pdf>

55. Prykhodko S., Makarova L., Latanska L., Bryzghalov M. Mathematical Model for Detecting Outliers in the Two-Dimensional Data of Software Metrics RFC and CBO from Applications in Java. *Information Control Systems & Technologies (ICST 2025)* : proceedings of the 13-th International Conference. Odesa, Ukraine, September 24–26, 2025. *CEUR Workshop Proceedings*. 2025. Vol.4048. P. 509-519.  
URL: <https://ceur-ws.org/Vol-4048/paper40.pdf>

56. Prykhodko N. V., Prykhodko S. B. A multiple non-linear regression model to estimate the agile testing efforts for small Web projects. *Radio Electronics, Computer Science, Control*. 2019. No. 2 (49). P.158-166. URL: <https://doi.org/10.15588/1607-3274-2019-2-17>

57. Prykhodko S., Prykhodko N., Makarova L., Pukhalevych A. Outlier Detection in Non-Linear Regression Analysis Based on the Normalizing Transformations. *Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)* : proceedings of the 2020 IEEE 15th International Conference. Lviv-Slavske, Ukraine: IEEE, 2020. P. 407-410. URL: <https://doi.org/10.1109/TCSET49122.2020.235464>

58. Kalinina I., Bidyuk P., Gozhyj A., Gozhyi V., Nechakhin V. Approach to Identification of Anomalous Values in Analysis Tasks and Data Pre-processing. In: Babichev, S., Lytvynenko, V. (eds) *Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making*, Volume 2. ISDMCI 2024. *Lecture Notes on Data Engineering and Communications Technologies*. 2025. Vol 244. Springer, Cham. URL: [https://doi.org/10.1007/978-3-031-88483-2\\_6](https://doi.org/10.1007/978-3-031-88483-2_6)

59. Prykhodko N. V., Prykhodko S. B. Constructing the non-linear regression models on the basis of multivariate normalizing transformations. *Electronic modeling*. 2018. Vol. 40. No. 6. P. 99-108. URL: <https://doi.org/10.15407/emodel.40.06.101>

60. Prykhodko N. V., Prykhodko S. B. Constructing the nonlinear regression equations based on multivariate normalizing transformations. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія : Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2018. Вип. 39. С. 61-68. URL: [http://nbuv.gov.ua/UJRN/VKhIMAM\\_2018\\_39\\_9](http://nbuv.gov.ua/UJRN/VKhIMAM_2018_39_9)

61. Prykhodko S., Prykhodko N. Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations. In: Shkarlet S., Morozov A., Palagin A. (eds) *Mathematical Modeling and Simulation of Systems (MODS'2020)*. MODS 2020. *Advances in Intelligent Systems and Computing*. 2021. Vol. 1265. P. 166-174. Springer, Cham. URL: [https://doi.org/10.1007/978-3-030-58124-4\\_16](https://doi.org/10.1007/978-3-030-58124-4_16)

62. Prykhodko S., Prykhodko N. 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*. 2022. Vol. 3179. P. 156-166. URL: [http://ceur-ws.org/Vol-3179/Paper\\_15.pdf](http://ceur-ws.org/Vol-3179/Paper_15.pdf)

63. Zabolotnii S., Warsza Z.L., Tkachenko O. Polynomial Estimation of Linear Regression Parameters for the Asymmetric PDF of Errors. In: Szewczyk, R., Zieliński, C., Kaliczyńska, M. (eds) *Automation 2018. Advances in Intelligent Systems and Computing*. 2018. Vol. 743. Springer, Cham. URL: [https://doi.org/10.1007/978-3-319-77179-3\\_75](https://doi.org/10.1007/978-3-319-77179-3_75)

64. Zabolotnii S., Khotunov V., Chepynoha A., Tkachenko O. Estimating Parameters of Linear Regression With an Exponential Power Distribution of Errors by Using a Polynomial Maximization Method. *Eastern-European Journal of Enterprise Technologies*. 2021. Vol. 1. No. 4 (109). P. 64-73. URL: <https://doi.org/10.15587/1729-4061.2021.225525>

65. Zabolotnii S., Tkachenko O., Warsza Z.L. Application of the Polynomial Maximization Method for Estimation Parameters in the Polynomial Regression with Non-Gaussian Residuals. In: Szewczyk, R., Zieliński, C., Kaliczyńska, M. (eds) *Automation 2021: Recent Achievements in Automation, Robotics and Measurement Techniques* (AUTOMATION 2021). *Advances in Intelligent Systems and Computing*. 2021. Vol. 1390. Springer, Cham. URL: [https://doi.org/10.1007/978-3-030-74893-7\\_36](https://doi.org/10.1007/978-3-030-74893-7_36)

66. Zabolotnii S., Tkachenko O., Warsza Z.L. Polynomial Maximization Method for Estimation Parameters of Asymmetric Non-Gaussian Moving Average Models. In: Szewczyk, R., Zieliński, C., Kaliczyńska, M., Bučinskas, V. (eds) *Automation 2023: Key Challenges in Automation, Robotics and Measurement Techniques* (AUTOMATION 2023). *Lecture Notes in Networks and Systems*. 2023. Vol 630. Springer, Cham. URL: [https://doi.org/10.1007/978-3-031-25844-2\\_21](https://doi.org/10.1007/978-3-031-25844-2_21)

67. Zabolotnii S., Tkachenko O., Nowakowski W., Warsza Z. Application of the Polynomial Maximization Method for Estimating Nonlinear Regression Parameters with Non-gaussian Asymmetric Errors. In: Szewczyk, R., Zieliński, C., Kaliczyńska, M., Bučinskas, V. (eds) *Automation 2024: Advances in Automation, Robotics and Measurement Techniques* (AUTOMATION 2024). *Lecture Notes in Networks and Systems*. 2024. Vol. 1219. Springer, Cham. URL: [https://doi.org/10.1007/978-3-031-78266-4\\_30](https://doi.org/10.1007/978-3-031-78266-4_30)

68. Bidyuk Peter, Kalinina Irina, Gozhyj Aleksandr. Methodology of Constructing Statistical Models for Nonlinear Non-stationary Processes in Medical Diagnostic Systems. *CEUR Workshop Proceedings*. 2020. P. 36-45. URL: <https://ceur-ws.org/Vol-2753/paper4.pdf>

69. Kalinina I., Bidyuk P., Gozhyj A. Construction of Forecast Models based on Bayesian Structural Time Series. *Computer Sciences and Information Technologies*

(CSIT) : proceedings of the 2022 IEEE 17th International Conference. Lviv, Ukraine: IEEE, 2022. P. 180-184. URL: <https://doi.org/10.1109/CSIT56902.2022.10000484>

70. **Prykhodko A. S.**, Malakhov E. V. Determining object-oriented design complexity due to the identification of classes of open-source web applications created using PHP frameworks. *Radio Electronics, Computer Science, Control*. 2024. No. 2 (69). P. 160-166. URL: <https://doi.org/10.15588/1607-3274-2024-2-16>

71. **Приходько А. С.** Математичні моделі для визначення складності об'єктно-орієнтованого проектування через зв'язки між класами веб-застосунків з відкритим кодом, створених за допомогою PHP фреймворків. *Прикладні питання математичного моделювання*. 2025. Том 8. № 1. С. 182-188. URL: <https://doi.org/10.32782/mathematical-modelling/2025-8-1-17>

72. **Приходько А. С.**, Малахов Є. В. Удосконалена математична модель оцінювання якості вебдодатків з відкритим кодом. *Центральноукраїнський науковий вісник. Технічні науки*. 2025. Вип. 11(42), ч. II. С. 30-37. URL: [https://doi.org/10.32515/2664-262X.2025.11\(42\).2.30-37](https://doi.org/10.32515/2664-262X.2025.11(42).2.30-37)

73. Laing V., Coleman C. Principal Components of Orthogonal Object-Oriented Metrics. White Paper Analyzing Results of NASA Object-Oriented Data. SATC, NASA, 2001.

74. Chidamber & Kemerer object-oriented metrics suite. URL: <https://www.aivosto.com/project/help/pm-oo-ck.html>

75. Khan R. A., Mustafa K. High Level Design Quality Assessment of Object Oriented Codes. *Verification and Validation of Enterprise Information Systems (ICEIS)* : proceedings of the 2004 2nd International Workshop, Porto, Portugal: ICEIS, 2004. P. 34-43. URL: <https://doi.org/10.5220/0002663900340043>

76. Prykhodko S., Prykhodko N. A Technique for Detecting Software Quality Based on the Confidence and Prediction Intervals of Nonlinear Regression for RFC Metric. *Computer Sciences and Information Technologies (CSIT)* : proceedings of the



2022 IEEE 17th International Conference. Lviv, Ukraine: IEEE, 2022. P. 499-502.  
URL: <https://doi.org/10.1109/CSIT56902.2022.10000532>

77. Prykhodko S., Makarova L., Pukhalevych A. Statistical Analysis of the Three-Dimensional Data of Software Metrics RFC, CBO, and WMC that are not Normally Distributed. *Applied Innovation in IT (ICAIT)* : proceedings of the 2025 13th International Conference. Kothlen, Germany: Anhalt University of Applied Sciences, April 26, 2025. Volume 13, Issue 1. P. 127-132. URL: [https://icaait.org/paper.php?paper=13th\\_ICAait\\_1/2\\_7](https://icaait.org/paper.php?paper=13th_ICAait_1/2_7).  
<http://dx.doi.org/10.25673/119224>

78. **Приходько А. С.**, Малахов Є. В. Визначення метрик Чидамбера та Кемерера за допомогою програмного забезпечення з відкритим кодом. *Free and Open Source Software* : матеріали XIV-ої Міжнародної науково-практичної конференції (Харків, 14-16 лютого 2023 р.) Харків: Харківський національний економічний університет імені Семена Кузнеця, 2023. С. 91. URL: <http://repository.hneu.edu.ua/bitstream/123456789/29041/1/foss-2023-theses.pdf>

79. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проектування через ідентифікацію класів веб застосунків, що створюються за допомогою PHP фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2023)* : матеріали IV-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2023 р.). Миколаїв: НУК імені адмірала Макарова, 2023. С. 35-37. URL: <https://itconf.nuos.edu.ua/2023/proceedings/>

80. **Приходько А. С.** Побудова математичної моделі для оцінювання складності об'єктно-орієнтованого проектування веб застосунків, що створюються за допомогою PHP фреймворків. *Комп'ютерне моделювання та оптимізація складних систем* : матеріали VIII Міжнародної науково-технічної конференції (Дніпро, 1-3 листопада 2023 року). Дніпро: Державний вищий

навчальний заклад «Український державний хіміко-технологічний університет», 2023. С. 61-62. URL: <https://udhtu.edu.ua/wp-content/uploads/2023/11/zbirnyk-tez-kmoss-2023.pdf>

81. **Приходько А. С.** Визначення складності об'єктно-орієнтованого проектування за рахунок ідентифікації класів веб застосунків розроблених на основі PHP фреймворків. *Інформаційні технології: теорія і практика* : тези доповідей I (VII) міжнародної науково-практичної конференції здобувачів вищої освіти і молодих учених (Дніпро, 20 – 22 березня 2024). / М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». Дніпро : Свідлер А.Л., 2024. С. 310-312. URL: [https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024\(Dnipro\).pdf](https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024(Dnipro).pdf)

82. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проектування через семантику класів вебзастосунків, що створюються за допомогою PHP фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2024)* : матеріали V-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2024 р.). Миколаїв: НУК імені адмірала Макарова, 2024. С. 17-19. URL: <https://itconf.nuos.edu.ua/2024/proceedings/>

83. **Приходько А. С.** Первинна обробка інформації з метрик RFC та СВО веб застосунків, що створені за допомогою PHP фреймворків. *Інформаційні технології і автоматизація* : матеріали XVII міжнародної науково-практичної конференції (Одеса, 31 жовтня - 1 листопада 2024 р.). Одеса, Видавництво ОНТУ, 2024. С. 498-500. URL: <https://ontu.edu.ua/download/konfi/2024/Collection-of-abstracts-of-the-conference-ITIA-2024.pdf>

84. **Приходько А. С.** Використання метрик RFC, СВО та WMC для оцінювання якості веб-застосунків з відкритим кодом, створених за допомогою

РНР фреймворків. *Інженерія програмного забезпечення і передові інформаційні технології (Soft Tech-2024)* : матеріали VII Міжнародної науково-практичної конференції молодих вчених та студентів (Київ, 20-22 листопада 2024 року). Київ: Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», ФІОТ, 2024. С. 82-85. URL: <https://drive.google.com/file/d/1O70Ysxe-z3SS82UqEaBdEXR2VdRW3NwG/view>

85. **Приходько А. С.** Математичні моделі для первинної обробки інформації з метрик RFC, СВО та WMC вебзастосунків, що створені з використанням РНР фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей двадцять другої всеукраїнської конференції студентів і молодих науковців (Одеса, 25 квітня 2025 р.). Одеса, 2025. С. 60-61. URL: <http://dspace.pdpu.edu.ua/bitstream/123456789/21944/1/Informatics%2C%20information%20systems.pdf#page=60>

86. Khan S. A., Khan R. A. Object oriented design complexity quantification model. *Procedia Technology*. 2012. Vol. 4. P. 548-554. URL: <https://doi.org/10.1016/j.protcy.2012.05.087>

87. Chidamber S.R., Kemerer C.F. Towards a metrics suite for object oriented design. *ACM SIGPLAN Notices*. 1991. Vol. 26. Issue 11. P. 197–211. URL: <https://doi.org/10.1145/118014.117970>

88. Booch G. Object oriented design with applications. Redwood City, CA: Benjamin/Cummings, 1991. 580 p.

89. Booch G. Object-oriented development. *IEEE Transactions on Software Engineering*. 1986. Vol. SE-12. No. 1. P. 211–221. URL: <https://doi.org/10.1109/TSE.1986.6312937>

90. Software Cost Estimation with COCOMO II / B. W. Boehm et al. Upper Saddle River, NJ : Prentice Hall PTR, 2000. 544 p.

91. Dillibabu R., Krishnaiah K. Cost estimation of a software product using COCOMO II.2000 model – a case study. *International Journal of Project Management*. 2005. Vol. 23, Issue 4. P. 297-307. URL: <https://doi.org/10.1016/j.ijproman.2004.11.003>

92. Приходько С. Б., Ворона М. В., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання розміру web-застосунків, що створюються за допомогою РНР-фреймворків. *Проблеми моделювання та розроблення інформаційних систем* : матеріали V науково-практичної інтернет-конференції (Дрогобич, 23 квітня 2021 року). Дрогобич: ДДПУ ім. І. Франка, 2021. С. 28-31.

93. Ворона М. В., **Приходько А. С.**, Шутко І. С. Трьохфакторні нелінійні регресійні моделі для оцінювання розміру WEB-застосунків, що створюються з використанням РНР-фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей вісімнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 23 квітня 2021 р. Одеса, 2021. С. 111-113. URL: [https://onu.edu.ua/pub/bank/userfiles/files/fmfit/naukova\\_diyalnist/konferenc/Zbirka\\_tez\\_IIS\\_T-2021.pdf](https://onu.edu.ua/pub/bank/userfiles/files/fmfit/naukova_diyalnist/konferenc/Zbirka_tez_IIS_T-2021.pdf)

94. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання кількості рядків коду Web-застосунків, що створюються за допомогою РНР-фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2021)* : матеріали II-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2021 р.). Миколаїв: НУК імені адмірала Макарова, 2021. С. 11-12. URL: <http://itconf.nuos.edu.ua/2021/proceedings/>

95. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Математичне моделювання кількості рядків коду Web-застосунків, що створюються за допомогою фреймворку CakePHP. *Стан, досягнення та перспективи*

*інформаційних систем і технологій* : матеріали XXII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів (Одеса, 21-22 квітня 2022 р.). Одеса, Видавництво ОНТУ, 2022. С. 36. URL: [https://ontu.edu.ua/download/konfi/2022/Conference\\_abstract-IT-21-22-04-22.pdf](https://ontu.edu.ua/download/konfi/2022/Conference_abstract-IT-21-22-04-22.pdf)

96. **Приходько А. С.** Аналітична модель чотиривимірного закону розподілу метрик Web застосунків, що створюються за допомогою фреймворку CakePHP. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2022)* : матеріали III-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2022 р.). Миколаїв: НУК імені адмірала Макарова, 2022. С. 52-53. URL: <http://itconf.nuos.edu.ua/2022/wp-content/uploads/sites/5/NUOS-ITMAS-2022-Proceedings.pdf>

97. Johnson R. A., Wichern D. W. Applied Multivariate Statistical Analysis. Pearson Prentice Hall, 2007. 800 p.

98. Hubert Mia, Debruyne Michiel, Rousseeuw Peter J. Minimum Covariance Determinant and Extensions. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2018. Vol. 10, Issue 3. e1421. URL: <https://doi.org/10.1002/wics.1421>

99. Xu D., Tian Y. A Comprehensive Survey of Clustering Algorithms. *Annals of Data Science*. 2015. Vol. 2. P. 165–193. URL: <https://doi.org/10.1007/s40745-015-0040-1>

100. Aggarwal C.C. Proximity-Based Outlier Detection. In: Outlier Analysis. 2017. P. 111–147. Springer, Cham. URL: [https://doi.org/10.1007/978-3-319-47578-3\\_4](https://doi.org/10.1007/978-3-319-47578-3_4)

101. Liu F. T., Ting K. M., Zhou Z. H. Isolation forest. *Data Mining (ICDM)* : proceedings of 2008 8th IEEE International Conference. Pisa, Italy: IEEE, 2008. P. 413–422. URL: <https://doi.org/10.1109/ICDM.2008.17>

102. Hariri S., Carrasco Kind M., Brunner R. J. Extended isolation forest. arXiv preprint. arXiv:1811.02141, 2018. URL: <https://arxiv.org/abs/1811.02141>

103. Liu F. T., Ting K. M., Zhou Z. H. Isolation-based anomaly detection. *ACM Transactions Knowledge Discovery from Data (TKDD)*. 2012. Volume 6, Issue 1, Article No.3, P. 1-39. URL: <https://doi.org/10.1145/2133360.2133363>
104. Prykhodko S., Prykhodko N., Makarova L., Pugachenko K. Detecting Outliers in Multivariate Non-Gaussian Data on the basis of Normalizing Transformations. *Electrical and Computer Engineering (UKRCON)* : proceedings of the 2017 IEEE First Ukraine Conference on «Celebrating 25 Years of IEEE Ukraine Section». Kyiv, Ukraine: IEEE, 2017. P. 846-849. URL: <https://doi.org/10.1109/UKRCON.2017.8100366>
105. Bates D. M., Watts D. G. Nonlinear Regression Analysis and Its Applications. New York: John Wiley & Sons. 1988. 365 p. URL: <https://doi.org/10.1002/9780470316757>
106. Seber G. A. F., Wild C. J. Nonlinear Regression. New York : John Wiley & Sons, 1989. 768 p. URL: <https://doi.org/10.1002/0471725315>
107. Freund R. J., Wilson W. J., Sa P. Regression Analysis: Statistical Modeling of a Response Variable. 2nd ed. Burlington, MA: Elsevier Academic Press, 2006. 481 p.
108. Ryan T. P. Modern regression methods. 2nd ed. Hoboken, NY : John Wiley & Sons, 2008. 672 p.
109. Chatterjee S., Simonoff J. S. Handbook of Regression Analysis. Hoboken, NJ : John Wiley & Sons, 2013. 235 p.
110. PhpMetrics is a static analysis tool for PHP. URL: <https://phpmetrics.org/> (дата звернення: 20.12.2025).
111. Tukey J. W. The comparative anatomy of transformations. *Annals of Mathematical Statistics*. 1957. Vol. 28. P. 602–632. URL: <https://doi.org/10.1214/aoms/1177706875>

112. Box G. E. P., Cox D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1964. Vol. 26. Issue 2. P. 211–252. URL: <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>

113. Kendall M. G., Stuart A. The advanced theory of statistics. Vol. 1. Distribution Theory. 2nd ed. London : Charles Griffin & Company Limited, 1963. 433 p.

114. Коваленко І. І., Приходько С. Б., Латанська Л. О. Сучасні методи статистичного аналізу даних : навч. посіб. Миколаїв : НУК, 2011. 189 с.

115. Stanfield P. M., Wilson J. R., Mirka G. A., Glasscock N. F., Psihogios J. P., Davis J. R. Multivariate input modeling with johnson distributions. *Winter simulation conference (WSC'96)* : proceedings of the 28th conference. Coronado, CA, USA, 8–11 December 1996. Washington : IEEE Computer Society, 1996. P. 1457–1464. URL: <https://doi.org/10.1145/256562.256990>

116. Wilson J. R. Modeling dependencies stochastic simulation inputs. *Winter simulation conference (WSC'97)* : proceedings of the 29th conference. Atlanta, GA, USA, 7–10 December 1997. Washington : IEEE Computer Society, 1997. P. 47–52. URL: <https://doi.org/10.1145/268437.268446>

117. Приходько С. Б. Методи ідентифікації нелінійних стохастичних диференціальних систем на основі нормалізуючих перетворень : автореф. дис. ... д-ра техн. наук : 01.05.02 / Ін-т кібернетики ім. В. М. Глушкова НАН України. Київ, 2012. 44 с.

118. Biscay Lirio R., Valdés Sosa P. A., Pascual Marqui R. D., Jiménez-Sobrino J. C., Alvarez Amador A., Galán Garcia L. Multivariate Box-Cox transformations with applications to neurometric data. *Computers in Biology and Medicine*. 1989. Vol. 19, Issue 4. P. 263–267. URL: [https://doi.org/10.1016/0010-4825\(89\)90013-9](https://doi.org/10.1016/0010-4825(89)90013-9)

119. Santiago Velilla. A note on the multivariate Box—Cox transformation to normality. *Statistics & Probability Letters*. 1993. Vol. 17, Issue 4. P. 259–263. URL: [https://doi.org/10.1016/0167-7152\(93\)90200-3](https://doi.org/10.1016/0167-7152(93)90200-3)
120. Mardia K. V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*. 1970. Vol. 57. P. 519–530. URL: <https://doi.org/10.1093/biomet/57.3.519>
121. Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques / E. S. Olivas et al. Hershey, PA : IGI Global, 2009. 703 p.
122. Liang H., Fu W., Yi F. A Survey of Recent Advances in Transfer Learning. *Communication Technology (ICCT)* : proceedings of the 2019 IEEE 19th International Conference. Xi'an, China: IEEE, 2019. P. 1516–1523. URL: <https://doi.org/10.1109/ICCT46805.2019.8947072>
123. Zhu Z., Lin K., Jain A. K., Zhou J. Transfer Learning in Deep Reinforcement Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023. Vol. 45, No. 11. P. 13344–13362. DOI: <https://doi.org/10.1109/TPAMI.2023.3292075>
124. Arsirii O. O.; Petrosiuk D. V. Pseudo-Labeling of Transfer Learning Convolutional Neural Network Data for Human Facial Emotion Recognition. *Herald of Advanced Information Technology*. 2023. Vol. 6, No. 3. P. 203–214. URL: <https://doi.org/10.15276/hait.06.2023.13>
125. Gupta V., Liao W.-k., Choudhary A., Agrawal A. Combining Transfer Learning and Representation Learning to Improve Predictive Analytics on Small Materials Data. *Machine Learning and Applications (ICMLA)* : proceedings of the 2024 International Conference. Miami, FL, USA: IEEE, 2024. P. 981–984. DOI: <https://doi.org/10.1109/ICMLA61862.2024.00145>



126. Drucker H., Burges C. J. C., Kaufman L., Smola A., Vapnik V. Support vector regression machines. *Advances in Neural Information Processing Systems 9* / eds. M. C. Mozer, M. I. Jordan, T. Petsche. Cambridge : MIT Press, 1997. P. 155–161. URL:  
[https://papers.nips.cc/paper\\_files/paper/1996/file/d38901788c533e8286cb6400b40b386d-Paper.pdf](https://papers.nips.cc/paper_files/paper/1996/file/d38901788c533e8286cb6400b40b386d-Paper.pdf)
127. Smola A. J., Schölkopf B. A tutorial on support vector regression. *Statistics and Computing*. 2004. Vol. 14. P. 199–222. URL:  
<https://doi.org/10.1023/B:STCO.0000035301.49549.88>
128. Mathar R., Alirezaei G., Balda E., Behboodi A. Support Vector Machines. In: *Fundamentals of Data Analytics*. Cham : Springer. 2020. P. 83–105. URL:  
[https://doi.org/10.1007/978-3-030-56831-3\\_6](https://doi.org/10.1007/978-3-030-56831-3_6)
129. Sarang P. Support Vector Machines. In: *Thinking Data Science*. The Springer Series in Applied Machine Learning. Cham : Springer. 2023. P. 153–165. URL: [https://doi.org/10.1007/978-3-031-02363-7\\_8](https://doi.org/10.1007/978-3-031-02363-7_8)
130. Shmilovici A. Support Vector Machines. In: Rokach, L., Maimon, O., Shmueli, E. (eds) *Machine Learning for Data Science Handbook*. Cham : Springer, 2023. P. 93–110. URL: [https://doi.org/10.1007/978-3-031-24628-9\\_6](https://doi.org/10.1007/978-3-031-24628-9_6)
131. Автоматизовані системи. Терміни та визначення : ДСТУ 2226-93 [Чинний від 1994-07-01]. Київ : Держстандарт України, 1994. 92 с.
132. Науково-інформаційна діяльність. Терміни та визначення понять : ДСТУ 5034:2008 [Чинний від 2009-01-01]. Київ : Держспоживстандарт України, 2009. 38 с.
133. A Dictionary of Physics / ed. J. Law, R. Rennie. 7th ed. Oxford : Oxford University Press, 2015. 666 p. URL:  
<https://doi.org/10.1093/acref/9780198714743.001.0001>

134. Kukhareva P. V., Weir C., Del Fiol G., Aarons G. A., Taft T. Y., Schlechter C. R., Reese T. J., Curran R. L., Nanjo C., Borbolla D., Staes C. J., Morgan K. L., Kramer H. S., Stipelman C. H., Shakib J. H., Flynn M. C., Kawamoto K. Evaluation in Life Cycle of Information Technology (ELICIT) framework: supporting the innovation life cycle from business case assessment to summative evaluation. *Journal of Biomedical Informatics*. 2022. Vol. 127. Article 104014. URL: <https://doi.org/10.1016/j.jbi.2022.104014>

135. Макоєдова В. О. Аналіз принципів побудови та підходів до визначення поняття «інформаційна технологія». *Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка»*. 2022. № 2 (18). С. 138-149. URL: <https://doi.org/10.28925/2663-4023.2022.18.138149>

136. Інформаційні технології : навч. посібник / О. І. Зачек, В. В. Сенник, Т. В. Магеровська та ін.; за ред. О. І. Зачека. Львів : Львівський державний університет внутрішніх справ, 2022. 432 с.

137. Імітаційне моделювання: навч. посібник / В. Ф. Ситнік, Н. С. Орленко. К. : КНЕУ, 1998. 232 с.

138. Імітаційне моделювання : монографія / О. П. Приставка, О. Г. Байбуз, П. О. Приставка ; Дніпропетр. нац. ун-т ім. Олеся Гончара. Дніпропетровськ : Вид-во ДНУ, 2011. 171 с.

**ДОДАТОК А. АКТИ ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ  
ДИСЕРТАЦІЙНОЇ РОБОТИ У РОЗРОБКУ ВЕБЗАСТОСУНКІВ**

# ТОВ "АЙПІ РІШЕННЯ"

54020, м.Миколаїв, вул. Данська, 23-В, т.(067) 565-50-84, код ЄДРПОУ 44132761  
р/рUA033071230000026005011025032, АТ "ВСТ БАНК" Дніпро, МФО 307123

## АКТ

впровадження результатів дисертаційної роботи на здобуття наукового ступеня доктора  
філософії (PhD) в галузі знань 12 «Інформаційні технології»  
за спеціальністю 122 «Комп'ютерні науки»  
Приходька Андрія Сергійовича

Ми, що нижче підписалися, Іващенко Костянтин Петрович, директор, цим Актом засвідчуємо, що результати дисертації Приходька Андрія Сергійовича, а саме інструментарії інформаційної технології для визначення складності об'єктно-орієнтоване проєктування (ООП) та якості розробки вебзастосунків, були впроваджені при створенні вебзастосунків, що розроблялися з використанням РНР-фреймворків.

Впровадження зазначених вище результатів дозволило підвищити достовірність визначення складності ООП та якості розробки відповідних вебзастосунків.

З повагою,  
директор ТОВ «АЙПІ РІШЕННЯ»



Костянтин Іващенко



**«ДІАГНОСТИКА ТА КОНТРОЛЬ»**

**ТОВ НАУКОВО-ВИРОБНИЧИЙ ЦЕНТР**

54020, м.Миколаїв, а/с 8 т/факс (0512) 718610 +380675155911

E-mail: diagnosticmsbox@gmail.com

Вих.№3131 від 29.01.2026р.

### АКТ

впровадження результатів дисертаційної роботи на здобуття наукового ступеня доктора  
філософії (PhD) в галузі знань 12 «Інформаційні технології»

за спеціальністю 122 «Комп'ютерні науки»

Приходька Андрія Сергійовича

Я, що нижче підписався, директор ТОВ НВЦ «Діагностика та контроль» Ярковець Олександр Чеславович, цим Актом засвідчую, що результати дисертації Приходька Андрія Сергійовича, а саме інструментарії інформаційної технології для визначення складності об'єктно-орієнтованого проєктування (ООП), якості розробки та кількості строк коду вебзастосунків, були впроваджені при створенні вебзастосунків, які використовуються в ТОВ НВЦ «Діагностика та контроль».

Впровадження зазначених вище результатів дозволило підвищити достовірність визначення складності ООП та якості розробки відповідних вебзастосунків.

Директор ТОВ НВЦ  
«Діагностика та контроль»



О.Ч. Ярковець

**МИКОЛАЇВСЬКА МІСЬКА РАДА**  
**Комунальне підприємство**  
**«Міський інформаційно-обчислювальний центр»**  
 вул. Адміральська, 20, м. Миколаїв, 54027, тел. (0512) 70-93-71  
 E-mail: digital@mkrada.gov.ua Код ЄДРПОУ 31554081

10.02.2026

№

01-10/02/2026

На №

від

**АКТ**

впровадження результатів дисертаційної роботи на здобуття наукового  
 ступеня доктора філософії (PhD) в галузі знань 12 «Інформаційні технології»  
 за спеціальністю 122 «Комп'ютерні науки»  
 Приходька Андрія Сергійовича

Я, що нижче підписався, директор КП «Міський інформаційно-обчислювальний центр» Богаченко Євгеній Валерійович, цим Актом засвідчую, що результати дисертації Приходька Андрія Сергійовича, а саме інструментарії інформаційної технології для визначення складності об'єктно-орієнтованого проектування (ООП), якості розробки та кількості строк коду вебзастосунків, були впроваджені при створенні вебзастосунків, які використовуються в КП «Міський інформаційно-обчислювальний центр».

Впровадження зазначених вище результатів дозволило підвищити достовірність визначення складності ООП та якості розробки відповідних вебзастосунків.

Директор

Богаченко +380669727434



Євгеній БОГАЧЕНКО

## АКТ

впровадження результатів дисертаційної роботи на здобуття наукового ступеня  
доктора філософії (PhD) в галузі знань 12 «Інформаційні технології»  
за спеціальністю 122 «Комп'ютерні науки»  
Приходька Андрія Сергійовича

Ми, що нижче підписалися, ФОП Білецький Віктор Васильович, цим Актом засвідчуємо, що результати дисертації Приходька Андрія Сергійовича, а саме інструментарії інформаційної технології для визначення складності об'єктно-орієнтованого проєктування (ООП), якості розробки та кількості строк коду вебзастосунків, були впроваджені при створенні вебзастосунків, що розробляються з використанням РНР-фреймворків.

Впровадження зазначених вище результатів дозволило підвищити достовірність визначення складності ООП, якості розробки та кількості строк коду відповідних вебзастосунків.

Фізична особа-підприсмець



В. В. Білецький

**ДОДАТОК Б. АКТ ПРО ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ  
ДИСЕРТАЦІЙНОЇ РОБОТИ В НАВЧАЛЬНОМУ ПРОЦЕСІ**



ЗАТВЕРДЖУЮ  
Проректор науково-педагогічної роботи  
Одеського національного  
університету імені І.І. Мечникова  
канд. пед. наук, проф.  
Майя НІКОЛАЄВА  
« 28 » 11 2025 р.

# АКТ

про використання результатів дисертаційної роботи на здобуття наукового ступеня  
доктора філософії аспіранта кафедри Математичного забезпечення комп'ютерних  
систем ПРИХОДЬКА Андрія Сергійовича в навчальному процесі кафедр Одеського  
національного університету імені І.І. Мечникова.

Ми, що нижче підписалися, завідувач кафедри Комп'ютерних систем та  
технологій ГУНЧЕНКО Юрій Олександрович та завідувач кафедри Математичного  
забезпечення комп'ютерних систем МАЛАХОВ Євгеній Валерійович склали акт про  
те, що результати дисертаційної роботи ПРИХОДЬКА Андрія Сергійовича  
впроваджені у навчальний процес на кафедрі Комп'ютерних систем та технологій та на  
кафедрі Математичного забезпечення комп'ютерних систем.

Математичні моделі та алгоритми для первинної обробки інформації з метрик  
вебзастосунків, що створені з використанням PHP фреймворків впроваджено в  
дисципліні «Комп'ютерна обробка експериментальних даних».

Завідувач кафедри КСТ,  
д.т.н., професор



Юрій ГУНЧЕНКО

Завідувач кафедри МЗКС,  
д.т.н., професор



Євгеній МАЛАХОВ

## ДОДАТОК В. НАБІР ДАНИХ З ПРОГРАМНИХ МЕТРИК WEB ЗАСТОСУНКІВ, ЩО РОЗРОБЛЕНІ З ВИКОРИСТАННЯМ PHP- ФРЕЙМВОРКІВ

За допомогою інструменту PhpMetrics був отриманий набір даних з програмних метрик 125 Web застосунків, що розроблені з використанням таких п'яти відомих PHP-фреймворків: CakePHP (16 застосунків), Codeigniter (20 застосунків), Laravel (55 застосунків), Symfony (18 застосунків) та Yii (16 застосунків). Посилання на ці застосунки наведені в табл.В.1, а значення їх програмних метрик – в табл.В.2.

Таблиця В.1 – Посилання на Web застосунки

| №  | Framework   | Link                                                                                                                                            |
|----|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | 2           | 3                                                                                                                                               |
| 1  | CakePHP     | <a href="https://github.com/Kareylo/BlogMVC-CakePHP3.git">https://github.com/Kareylo/BlogMVC-CakePHP3.git</a>                                   |
| 2  | CakePHP     | <a href="https://github.com/gothinkster/cakephp-realworld-example-app.git">https://github.com/gothinkster/cakephp-realworld-example-app.git</a> |
| 3  | CakePHP     | <a href="https://github.com/FriendsOfCake/crud.git">https://github.com/FriendsOfCake/crud.git</a>                                               |
| 4  | CakePHP     | <a href="https://github.com/sophimail/webadmin.git">https://github.com/sophimail/webadmin.git</a>                                               |
| 5  | CakePHP     | <a href="https://github.com/dereuromark/cakephp-sandbox.git">https://github.com/dereuromark/cakephp-sandbox.git</a>                             |
| 6  | CakePHP     | <a href="https://github.com/croogo/croogo.git">https://github.com/croogo/croogo.git</a>                                                         |
| 7  | CakePHP     | <a href="https://github.com/passbolt/passbolt_api.git">https://github.com/passbolt/passbolt_api.git</a>                                         |
| 8  | CakePHP     | <a href="https://github.com/foodcoopshop/foodcoopshop.git">https://github.com/foodcoopshop/foodcoopshop.git</a>                                 |
| 9  | CakePHP     | <a href="https://github.com/quickapps/cms.git">https://github.com/quickapps/cms.git</a>                                                         |
| 10 | CakePHP     | <a href="https://github.com/telerim/old-Zhen-CRM.git">https://github.com/telerim/old-Zhen-CRM.git</a>                                           |
| 11 | CakePHP     | <a href="https://github.com/infinitas/infinitas.git">https://github.com/infinitas/infinitas.git</a>                                             |
| 12 | CakePHP     | <a href="https://github.com/klevo/wildflower.git">https://github.com/klevo/wildflower.git</a>                                                   |
| 13 | CakePHP     | <a href="https://github.com/Datawalke/Coordino.git">https://github.com/Datawalke/Coordino.git</a>                                               |
| 14 | CakePHP     | <a href="https://github.com/yandod/candycane.git">https://github.com/yandod/candycane.git</a>                                                   |
| 15 | CakePHP     | <a href="https://github.com/CodeBlastr/SYSTEM.git">https://github.com/CodeBlastr/SYSTEM.git</a>                                                 |
| 16 | CakePHP     | <a href="https://github.com/vamcart/vamcart.git">https://github.com/vamcart/vamcart.git</a>                                                     |
| 17 | CodeIgniter | <a href="https://github.com/scoumbourdis/grocery-crud.git">https://github.com/scoumbourdis/grocery-crud.git</a>                                 |
| 18 | CodeIgniter | <a href="https://github.com/DakuTree/manga-tracker.git">https://github.com/DakuTree/manga-tracker.git</a>                                       |
| 19 | CodeIgniter | <a href="https://github.com/craigrodway/classroombookings.git">https://github.com/craigrodway/classroombookings.git</a>                         |

Продовження табл. В.1

| 1  | 2           | 3                                                                                                                                                                                             |
|----|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 20 | CodeIgniter | <a href="https://github.com/InvoicePlane/InvoicePlane.git">https://github.com/InvoicePlane/InvoicePlane.git</a>                                                                               |
| 21 | CodeIgniter | <a href="https://github.com/RamonSilva20/mapos.git">https://github.com/RamonSilva20/mapos.git</a>                                                                                             |
| 22 | CodeIgniter | <a href="https://github.com/kalkun-sms/Kalkun.git">https://github.com/kalkun-sms/Kalkun.git</a>                                                                                               |
| 23 | CodeIgniter | <a href="https://github.com/bradtraversy/ciblog.git">https://github.com/bradtraversy/ciblog.git</a>                                                                                           |
| 24 | CodeIgniter | <a href="https://github.com/Wscats/cms.git">https://github.com/Wscats/cms.git</a>                                                                                                             |
| 25 | CodeIgniter | <a href="https://github.com/amirsanni/Mini-Inventory-and-Sales-Management-System.git">https://github.com/amirsanni/Mini-Inventory-and-Sales-Management-System.git</a>                         |
| 26 | CodeIgniter | <a href="https://github.com/kishor10d/Admin-Panel-User-Management-using-CodeIgniter.git">https://github.com/kishor10d/Admin-Panel-User-Management-using-CodeIgniter.git</a>                   |
| 27 | CodeIgniter | <a href="https://github.com/ronknight/InventorySystem.git">https://github.com/ronknight/InventorySystem.git</a>                                                                               |
| 28 | CodeIgniter | <a href="https://github.com/Cryin/Startblog.git">https://github.com/Cryin/Startblog.git</a>                                                                                                   |
| 29 | CodeIgniter | <a href="https://github.com/Devnawjesh/hr-payroll.git">https://github.com/Devnawjesh/hr-payroll.git</a>                                                                                       |
| 30 | CodeIgniter | <a href="https://github.com/alextselegidis/easyappointments.git">https://github.com/alextselegidis/easyappointments.git</a>                                                                   |
| 31 | CodeIgniter | <a href="https://github.com/opensourcepos/opensourcepos.git">https://github.com/opensourcepos/opensourcepos.git</a>                                                                           |
| 32 | CodeIgniter | <a href="https://github.com/jockchou/gitblog.git">https://github.com/jockchou/gitblog.git</a>                                                                                                 |
| 33 | CodeIgniter | <a href="https://github.com/kirilkirkov/Ecommerce-CodeIgniter-Bootstrap.git">https://github.com/kirilkirkov/Ecommerce-CodeIgniter-Bootstrap.git</a>                                           |
| 34 | CodeIgniter | <a href="https://github.com/goFrendiAsgard/No-CMS.git">https://github.com/goFrendiAsgard/No-CMS.git</a>                                                                                       |
| 35 | CodeIgniter | <a href="https://github.com/daylightstudio/FUEL-CMS.git">https://github.com/daylightstudio/FUEL-CMS.git</a>                                                                                   |
| 36 | CodeIgniter | <a href="https://github.com/kirilkirkov/Electronic-Invoicing-And-Warehouse-Management-System.git">https://github.com/kirilkirkov/Electronic-Invoicing-And-Warehouse-Management-System.git</a> |
| 37 | Laravel     | <a href="https://github.com/dameety/laravel-church-sermons-app.git">https://github.com/dameety/laravel-church-sermons-app.git</a>                                                             |
| 38 | Laravel     | <a href="https://github.com/corcel/corcel.git">https://github.com/corcel/corcel.git</a>                                                                                                       |
| 39 | Laravel     | <a href="https://github.com/orgmanager/orgmanager.git">https://github.com/orgmanager/orgmanager.git</a>                                                                                       |
| 40 | Laravel     | <a href="https://github.com/LaravelDaily/Laraquiz-QuickAdminPanel.git">https://github.com/LaravelDaily/Laraquiz-QuickAdminPanel.git</a>                                                       |
| 41 | Laravel     | <a href="https://github.com/iamwebwiz/staff-management-system.git">https://github.com/iamwebwiz/staff-management-system.git</a>                                                               |
| 42 | Laravel     | <a href="https://github.com/jeremykenedy/laravel-admin.git">https://github.com/jeremykenedy/laravel-admin.git</a>                                                                             |
| 43 | Laravel     | <a href="https://github.com/WebDevEtc/BlogEtc.git">https://github.com/WebDevEtc/BlogEtc.git</a>                                                                                               |
| 44 | Laravel     | <a href="https://github.com/bpocallaghan/generators.git">https://github.com/bpocallaghan/generators.git</a>                                                                                   |
| 45 | Laravel     | <a href="https://github.com/FaZeRs/laravel-portfolio.git">https://github.com/FaZeRs/laravel-portfolio.git</a>                                                                                 |
| 46 | Laravel     | <a href="https://github.com/NathanGeerinck/laravel-newsletter.git">https://github.com/NathanGeerinck/laravel-newsletter.git</a>                                                               |
| 47 | Laravel     | <a href="https://github.com/jeremykenedy/larablog.git">https://github.com/jeremykenedy/larablog.git</a>                                                                                       |
| 48 | Laravel     | <a href="https://github.com/realodix/urlhub.git">https://github.com/realodix/urlhub.git</a>                                                                                                   |
| 49 | Laravel     | <a href="https://github.com/itobuz/shila-laravel-backend.git">https://github.com/itobuz/shila-laravel-backend.git</a>                                                                         |
| 50 | Laravel     | <a href="https://github.com/apiato/apiato.git">https://github.com/apiato/apiato.git</a>                                                                                                       |
| 51 | Laravel     | <a href="https://github.com/summerblue/larabbs.git">https://github.com/summerblue/larabbs.git</a>                                                                                             |
| 52 | Laravel     | <a href="https://github.com/jeremykenedy/laravel-auth.git">https://github.com/jeremykenedy/laravel-auth.git</a>                                                                               |
| 53 | Laravel     | <a href="https://github.com/ivan-atanasov/readmarvel.git">https://github.com/ivan-atanasov/readmarvel.git</a>                                                                                 |
| 54 | Laravel     | <a href="https://github.com/rappasoft/laravel-boilerplate.git">https://github.com/rappasoft/laravel-boilerplate.git</a>                                                                       |

Продовження табл. В.1

| 1  | 2       | 3                                                                                                                                       |
|----|---------|-----------------------------------------------------------------------------------------------------------------------------------------|
| 55 | Laravel | <a href="https://github.com/aimeos/aimeos-laravel.git">https://github.com/aimeos/aimeos-laravel.git</a>                                 |
| 56 | Laravel | <a href="https://github.com/summerblue/laravel-shop.git">https://github.com/summerblue/laravel-shop.git</a>                             |
| 57 | Laravel | <a href="https://github.com/Hasnayeem/invobook.git">https://github.com/Hasnayeem/invobook.git</a>                                       |
| 58 | Laravel | <a href="https://github.com/ammannbe/RecipeManagerApi.git">https://github.com/ammannbe/RecipeManagerApi.git</a>                         |
| 59 | Laravel | <a href="https://github.com/InfyOmLabs/laravel-generator.git">https://github.com/InfyOmLabs/laravel-generator.git</a>                   |
| 60 | Laravel | <a href="https://github.com/koel/koel.git">https://github.com/koel/koel.git</a>                                                         |
| 61 | Laravel | <a href="https://github.com/timegridio/timegrid.git">https://github.com/timegridio/timegrid.git</a>                                     |
| 62 | Laravel | <a href="https://github.com/nasirkhan/laravel-starter.git">https://github.com/nasirkhan/laravel-starter.git</a>                         |
| 63 | Laravel | <a href="https://github.com/ziishaned/opus.git">https://github.com/ziishaned/opus.git</a>                                               |
| 64 | Laravel | <a href="https://github.com/adaptcms/AdaptCMS.git">https://github.com/adaptcms/AdaptCMS.git</a>                                         |
| 65 | Laravel | <a href="https://github.com/lubusIN/laravel-gymie.git">https://github.com/lubusIN/laravel-gymie.git</a>                                 |
| 66 | Laravel | <a href="https://github.com/gitscrum-team/laravel-gitscrum.git">https://github.com/gitscrum-team/laravel-gitscrum.git</a>               |
| 67 | Laravel | <a href="https://github.com/summerblue/phphub5.git">https://github.com/summerblue/phphub5.git</a>                                       |
| 68 | Laravel | <a href="https://github.com/orchidsoftware/platform.git">https://github.com/orchidsoftware/platform.git</a>                             |
| 69 | Laravel | <a href="https://github.com/Attendize/Attendize.git">https://github.com/Attendize/Attendize.git</a>                                     |
| 70 | Laravel | <a href="https://github.com/leskhq/Laravel-Enterprise-Starter-Kit.git">https://github.com/leskhq/Laravel-Enterprise-Starter-Kit.git</a> |
| 71 | Laravel | <a href="https://github.com/agorakit/agorakit.git">https://github.com/agorakit/agorakit.git</a>                                         |
| 72 | Laravel | <a href="https://github.com/akhileshdarjee/origin-cms.git">https://github.com/akhileshdarjee/origin-cms.git</a>                         |
| 73 | Laravel | <a href="https://github.com/vuupe/dclassifieds.laravel.git">https://github.com/vuupe/dclassifieds.laravel.git</a>                       |
| 74 | Laravel | <a href="https://github.com/CachetHQ/Cachet.git">https://github.com/CachetHQ/Cachet.git</a>                                             |
| 75 | Laravel | <a href="https://github.com/pterodactyl/panel.git">https://github.com/pterodactyl/panel.git</a>                                         |
| 76 | Laravel | <a href="https://github.com/dwijitsolutions/laraadmin.git">https://github.com/dwijitsolutions/laraadmin.git</a>                         |
| 77 | Laravel | <a href="https://github.com/msurguy/critterapp.git">https://github.com/msurguy/critterapp.git</a>                                       |
| 78 | Laravel | <a href="https://github.com/krayin/laravel-crm.git">https://github.com/krayin/laravel-crm.git</a>                                       |
| 79 | Laravel | <a href="https://github.com/msurguy/LaravelBackboneTodoMVC.git">https://github.com/msurguy/LaravelBackboneTodoMVC.git</a>               |
| 80 | Laravel | <a href="https://github.com/flarum/framework.git">https://github.com/flarum/framework.git</a>                                           |
| 81 | Laravel | <a href="https://github.com/doptor/Doptor.git">https://github.com/doptor/Doptor.git</a>                                                 |
| 82 | Laravel | <a href="https://github.com/crater-invoice/crater.git">https://github.com/crater-invoice/crater.git</a>                                 |
| 83 | Laravel | <a href="https://github.com/BookStackApp/BookStack.git">https://github.com/BookStackApp/BookStack.git</a>                               |
| 84 | Laravel | <a href="https://github.com/area17/twill.git">https://github.com/area17/twill.git</a>                                                   |
| 85 | Laravel | <a href="https://github.com/bagisto/bagisto.git">https://github.com/bagisto/bagisto.git</a>                                             |
| 86 | Laravel | <a href="https://github.com/akaunting/akaunting.git">https://github.com/akaunting/akaunting.git</a>                                     |
| 87 | Laravel | <a href="https://github.com/pixelfed/pixelfed.git">https://github.com/pixelfed/pixelfed.git</a>                                         |
| 88 | Laravel | <a href="https://github.com/filamentphp/filament.git">https://github.com/filamentphp/filament.git</a>                                   |
| 89 | Laravel | <a href="https://github.com/snipe/snipe-it.git">https://github.com/snipe/snipe-it.git</a>                                               |
| 90 | Laravel | <a href="https://github.com/freescout-helpdesk/freescout.git">https://github.com/freescout-helpdesk/freescout.git</a>                   |
| 91 | Laravel | <a href="https://github.com/microweber/microweber.git">https://github.com/microweber/microweber.git</a>                                 |
| 92 | Symfony | <a href="https://github.com/CodelyTV/php-ddd-example.git">https://github.com/CodelyTV/php-ddd-example.git</a>                           |

Кінець табл. В.1

| 1   | 2       | 3                                                                                                                                                   |
|-----|---------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| 93  | Symfony | <a href="https://github.com/phpdish/phpdish.git">https://github.com/phpdish/phpdish.git</a>                                                         |
| 94  | Symfony | <a href="https://github.com/wallabag/wallabag.git">https://github.com/wallabag/wallabag.git</a>                                                     |
| 95  | Symfony | <a href="https://github.com/sonata-project/ecommerce.git">https://github.com/sonata-project/ecommerce.git</a>                                       |
| 96  | Symfony | <a href="https://github.com/Cocolabs-SAS/cocorico.git">https://github.com/Cocolabs-SAS/cocorico.git</a>                                             |
| 97  | Symfony | <a href="https://github.com/oroinc/crm.git">https://github.com/oroinc/crm.git</a>                                                                   |
| 98  | Symfony | <a href="https://github.com/bolt/bolt.git">https://github.com/bolt/bolt.git</a>                                                                     |
| 99  | Symfony | <a href="https://github.com/ilios/ilios.git">https://github.com/ilios/ilios.git</a>                                                                 |
| 100 | Symfony | <a href="https://github.com/kimai/kimai.git">https://github.com/kimai/kimai.git</a>                                                                 |
| 101 | Symfony | <a href="https://github.com/EC-CUBE/ec-cube.git">https://github.com/EC-CUBE/ec-cube.git</a>                                                         |
| 102 | Symfony | <a href="https://github.com/forkcms/forkcms.git">https://github.com/forkcms/forkcms.git</a>                                                         |
| 103 | Symfony | <a href="https://github.com/getgrav/grav.git">https://github.com/getgrav/grav.git</a>                                                               |
| 104 | Symfony | <a href="https://github.com/sulu/sulu.git">https://github.com/sulu/sulu.git</a>                                                                     |
| 105 | Symfony | <a href="https://github.com/shopsys/shopsys.git">https://github.com/shopsys/shopsys.git</a>                                                         |
| 106 | Symfony | <a href="https://github.com/pimcore/pimcore.git">https://github.com/pimcore/pimcore.git</a>                                                         |
| 107 | Symfony | <a href="https://github.com/Sylius/Sylius.git">https://github.com/Sylius/Sylius.git</a>                                                             |
| 108 | Symfony | <a href="https://github.com/shopware/platform.git">https://github.com/shopware/platform.git</a>                                                     |
| 109 | Symfony | <a href="https://github.com/oroinc/platform.git">https://github.com/oroinc/platform.git</a>                                                         |
| 110 | Yii     | <a href="https://github.com/samdark/yii2-shop.git">https://github.com/samdark/yii2-shop.git</a>                                                     |
| 111 | Yii     | <a href="https://github.com/akiraz2/yii2-blog.git">https://github.com/akiraz2/yii2-blog.git</a>                                                     |
| 112 | Yii     | <a href="https://github.com/ciims/CiiMS.git">https://github.com/ciims/CiiMS.git</a>                                                                 |
| 113 | Yii     | <a href="https://github.com/yii-podium/yii2-api.git">https://github.com/yii-podium/yii2-api.git</a>                                                 |
| 114 | Yii     | <a href="https://github.com/noumo/easyii.git">https://github.com/noumo/easyii.git</a>                                                               |
| 115 | Yii     | <a href="https://github.com/weison-tech/yii2-cms.git">https://github.com/weison-tech/yii2-cms.git</a>                                               |
| 116 | Yii     | <a href="https://github.com/luyadev/luya.git">https://github.com/luyadev/luya.git</a>                                                               |
| 117 | Yii     | <a href="https://github.com/ElisDN/yii2-demo-shop.git">https://github.com/ElisDN/yii2-demo-shop.git</a>                                             |
| 118 | Yii     | <a href="https://github.com/bizley/yii2-podium.git">https://github.com/bizley/yii2-podium.git</a>                                                   |
| 119 | Yii     | <a href="https://github.com/liufee/cms.git">https://github.com/liufee/cms.git</a>                                                                   |
| 120 | Yii     | <a href="https://github.com/skeeks-cms/cms.git">https://github.com/skeeks-cms/cms.git</a>                                                           |
| 121 | Yii     | <a href="https://github.com/yupe/yupe.git">https://github.com/yupe/yupe.git</a>                                                                     |
| 122 | Yii     | <a href="https://github.com/opensourcewebsite-org/opensourcewebsite-org.git">https://github.com/opensourcewebsite-org/opensourcewebsite-org.git</a> |
| 123 | Yii     | <a href="https://github.com/humhub/humhub.git">https://github.com/humhub/humhub.git</a>                                                             |
| 124 | Yii     | <a href="https://github.com/fecshop/yii2_fecshop.git">https://github.com/fecshop/yii2_fecshop.git</a>                                               |
| 125 | Yii     | <a href="https://github.com/craftcms/cms.git">https://github.com/craftcms/cms.git</a>                                                               |

Таблиця В.2 – Набір даних з програмних метрик Web застосунків

| №  | LOC    | Classes | Methods per Class | LCOM | AC   | EC   | RFC   | DIT  | NOC   | WMC    |
|----|--------|---------|-------------------|------|------|------|-------|------|-------|--------|
| 1  | 2      | 3       | 4                 | 5    | 6    | 7    | 8     | 9    | 10    | 11     |
| 1  | 1054   | 27      | 1,74              | 1,44 | 0,15 | 1,93 | 6,48  | 2,09 | 0,148 | 2,81   |
| 2  | 1539   | 60      | 2,2               | 1,57 | 0,3  | 2,42 | 8,28  | 2,11 | 0,250 | 4,02   |
| 3  | 1802   | 40      | 4,13              | 1,3  | 1,23 | 3,2  | 12,2  | 1,63 | 0,600 | 9,13   |
| 4  | 1709   | 45      | 3,13              | 1,73 | 0,29 | 2,2  | 10,73 | 2,13 | 0,289 | 6,96   |
| 5  | 4766   | 109     | 3,53              | 1,68 | 0,67 | 2,26 | 10,49 | 1,81 | 0,550 | 7,04   |
| 6  | 16677  | 367     | 3,36              | 1,73 | 0,69 | 3,08 | 9,81  | 1,64 | 0,283 | 9,8    |
| 7  | 29393  | 798     | 3,25              | 1,69 | 1,79 | 4,13 | 9,5   | 1,19 | 0,285 | 7,35   |
| 8  | 19650  | 269     | 4,47              | 2,27 | 1,08 | 3,3  | 17,16 | 1,28 | 0,435 | 15,21  |
| 9  | 27300  | 289     | 5,33              | 2,16 | 0,98 | 2,92 | 12,02 | 1,33 | 0,299 | 24,88  |
| 10 | 32024  | 217     | 8,22              | 3,08 | 2,97 | 3,32 | 12,59 | 1,23 | 0,725 | 43,73  |
| 11 | 36566  | 487     | 4,93              | 1,9  | 1,17 | 2,55 | 7,71  | 1,79 | 0,766 | 19,5   |
| 12 | 36799  | 175     | 10,48             | 3,23 | 2,97 | 3,18 | 15,88 | 1,56 | 0,949 | 60,02  |
| 13 | 35746  | 152     | 11,57             | 3,43 | 3,17 | 3,39 | 17,57 | 1,46 | 0,914 | 67,83  |
| 14 | 56122  | 363     | 8,93              | 2,98 | 3,02 | 3,44 | 14,58 | 1,26 | 0,777 | 43,02  |
| 15 | 59605  | 406     | 8,45              | 2,84 | 3,13 | 3,6  | 13,76 | 1,31 | 0,781 | 41,86  |
| 16 | 138943 | 1496    | 4,26              | 1,6  | 2,21 | 2,38 | 10,17 | 1,34 | 0,829 | 19,71  |
| 17 | 5267   | 11      | 29,55             | 1,82 | 0,64 | 1,45 | 47,09 | 2    | 0,364 | 130,09 |
| 18 | 8710   | 180     | 3,18              | 1,11 | 0,95 | 1,32 | 7,84  | 1,84 | 0,894 | 10,08  |
| 19 | 10663  | 140     | 5,1               | 1,76 | 0,66 | 1,34 | 11,89 | 1,44 | 0,357 | 16,11  |
| 20 | 8202   | 102     | 5,98              | 1,79 | 0,92 | 1,28 | 15,97 | 2,07 | 0,755 | 15,47  |
| 21 | 8910   | 77      | 7,32              | 1,45 | 0,27 | 1,39 | 18,75 | 1,64 | 0,231 | 22,61  |
| 22 | 14386  | 170     | 5,84              | 2,76 | 1,14 | 1,55 | 11,28 | 1,22 | 0,341 | 18,63  |
| 23 | 19508  | 141     | 10,6              | 2,39 | 0,79 | 1,21 | 15,37 | 1,31 | 0,683 | 40,87  |
| 24 | 19582  | 140     | 10,73             | 2,44 | 0,79 | 1,21 | 15,39 | 1,31 | 0,681 | 41,26  |
| 25 | 20838  | 151     | 10,55             | 2,38 | 0,79 | 1,22 | 16,05 | 1,29 | 0,682 | 40,33  |
| 26 | 20838  | 144     | 11,02             | 2,41 | 0,78 | 1,23 | 16,47 | 1,31 | 0,667 | 42,03  |
| 27 | 21491  | 156     | 10,31             | 2,34 | 0,81 | 1,19 | 15,52 | 1,36 | 0,713 | 39,44  |
| 28 | 20944  | 153     | 10,5              | 2,45 | 0,78 | 1,18 | 15,48 | 1,31 | 0,682 | 39,46  |
| 29 | 24639  | 156     | 12,24             | 2,35 | 0,8  | 1,21 | 19,25 | 1,3  | 0,707 | 41,9   |
| 30 | 24832  | 217     | 8,06              | 1,95 | 1,04 | 1,73 | 13,65 | 1,3  | 0,662 | 32,24  |
| 31 | 15716  | 135     | 8,8               | 1,89 | 0,67 | 1,36 | 20,33 | 1,6  | 0,444 | 22,31  |
| 32 | 29182  | 283     | 7,12              | 2,32 | 1,9  | 2,51 | 11,99 | 1,21 | 0,551 | 27,82  |
| 33 | 46232  | 240     | 12,04             | 2,85 | 0,93 | 1,49 | 19,48 | 1,45 | 0,684 | 56,21  |
| 34 | 56687  | 366     | 10,66             | 3,06 | 1,25 | 1,66 | 18,29 | 1,53 | 0,712 | 41,84  |
| 35 | 69300  | 507     | 8,43              | 2,13 | 1,56 | 2,1  | 15,08 | 1,38 | 0,663 | 39     |

Продовження табл. В.2

| 1  | 2      | 3   | 4     | 5    | 6    | 7    | 8     | 9    | 10    | 11    |
|----|--------|-----|-------|------|------|------|-------|------|-------|-------|
| 36 | 112620 | 584 | 10,35 | 2,92 | 2,29 | 2,76 | 20,98 | 1,27 | 0,559 | 49,45 |
| 37 | 1161   | 53  | 3,55  | 3,47 | 0,7  | 2,32 | 5,94  | 1,81 | 0,340 | 3,92  |
| 38 | 1119   | 36  | 4,11  | 2,11 | 0,86 | 2,08 | 7,33  | 1,44 | 0,514 | 6,92  |
| 39 | 1125   | 57  | 2,16  | 1,72 | 0,68 | 2,26 | 5,49  | 1,81 | 0,228 | 3,04  |
| 40 | 1138   | 58  | 2,6   | 2,43 | 1,02 | 2,38 | 5,71  | 1,84 | 0,224 | 3,45  |
| 41 | 1146   | 67  | 2,1   | 1,58 | 1,27 | 2,55 | 4,84  | 1,51 | 0,147 | 2,39  |
| 42 | 1494   | 65  | 2,03  | 1,55 | 0,58 | 2,29 | 5,31  | 1,63 | 0,185 | 3,66  |
| 43 | 1844   | 73  | 2,81  | 1,32 | 1,37 | 2,86 | 5,88  | 1,28 | 0,247 | 4,45  |
| 44 | 1409   | 32  | 4,31  | 1,28 | 0,84 | 1,34 | 8,88  | 1,5  | 0,719 | 9,22  |
| 45 | 1609   | 90  | 2,03  | 1,91 | 0,69 | 2,26 | 6,03  | 1,84 | 0,100 | 2,81  |
| 46 | 1501   | 80  | 2,44  | 1,98 | 0,83 | 2,06 | 6,31  | 1,68 | 0,188 | 2,73  |
| 47 | 1585   | 80  | 2,21  | 1,98 | 0,7  | 2,11 | 5,5   | 1,78 | 0,200 | 3,15  |
| 48 | 1153   | 47  | 2,77  | 2,36 | 0,96 | 3,34 | 7,4   | 1,48 | 0,149 | 4,21  |
| 49 | 2157   | 94  | 2,77  | 2,66 | 0,74 | 2,15 | 5,88  | 1,87 | 0,277 | 3,36  |
| 50 | 2947   | 237 | 1,41  | 1,35 | 1,68 | 2,44 | 3     | 2,44 | 0,770 | 1,75  |
| 51 | 1459   | 84  | 1,79  | 1,52 | 0,71 | 2,13 | 4,71  | 1,72 | 0,226 | 2,5   |
| 52 | 2199   | 94  | 2,12  | 1,83 | 0,79 | 2,98 | 6,87  | 1,85 | 0,223 | 3,74  |
| 53 | 1789   | 78  | 2,64  | 1,71 | 1    | 3,01 | 6,38  | 1,57 | 0,321 | 3,64  |
| 54 | 2585   | 131 | 2,3   | 1,7  | 0,66 | 1,91 | 5,96  | 1,31 | 0,031 | 3,52  |
| 55 | 1432   | 40  | 2,93  | 1,45 | 0,8  | 4,7  | 10    | 1,46 | 0,125 | 6,53  |
| 56 | 2053   | 92  | 1,77  | 1,52 | 1,01 | 2,52 | 6,54  | 1,75 | 0,217 | 3,04  |
| 57 | 2423   | 130 | 2,39  | 1,79 | 0,59 | 4,21 | 8,88  | 1,9  | 0,077 | 2,92  |
| 58 | 3677   | 176 | 2,81  | 2,44 | 1    | 2,37 | 6,36  | 1,38 | 0,080 | 3,8   |
| 59 | 3527   | 58  | 4,59  | 1,81 | 1,03 | 1,67 | 8,79  | 1,22 | 0,517 | 13,26 |
| 60 | 5823   | 323 | 2,22  | 1,65 | 1,65 | 2,9  | 4,91  | 1,33 | 0,381 | 3,42  |
| 61 | 3827   | 141 | 2,65  | 1,61 | 0,67 | 2,57 | 9,21  | 1,4  | 0,275 | 3,99  |
| 62 | 3772   | 111 | 2,87  | 1,82 | 0,8  | 3,32 | 8,1   | 1,51 | 0,279 | 4,51  |
| 63 | 3926   | 110 | 3,75  | 2,23 | 1,28 | 3,28 | 9,47  | 1,8  | 0,209 | 5,36  |
| 64 | 5840   | 151 | 3,42  | 2,68 | 1,16 | 3,14 | 8,62  | 1,46 | 0,219 | 6,93  |
| 65 | 4292   | 138 | 2,74  | 2,62 | 0,95 | 2,57 | 7,41  | 1,43 | 0,145 | 4,87  |
| 66 | 5520   | 194 | 3,11  | 2,43 | 0,77 | 2,24 | 7,52  | 1,25 | 0,128 | 5,36  |
| 67 | 6196   | 231 | 2,85  | 2,22 | 1,17 | 2,91 | 7,94  | 1,49 | 0,236 | 4,56  |
| 68 | 6271   | 231 | 3,54  | 1,74 | 1,59 | 2,83 | 8,27  | 1,47 | 0,539 | 5,04  |
| 69 | 6420   | 189 | 2,96  | 2,08 | 1,21 | 3,32 | 8,01  | 1,57 | 0,286 | 5,35  |
| 70 | 6318   | 143 | 3,43  | 1,9  | 1,05 | 3,1  | 7,15  | 1,52 | 0,150 | 8,77  |
| 71 | 7855   | 192 | 3,42  | 2,22 | 1,44 | 3,36 | 10,57 | 1,53 | 0,284 | 7,48  |
| 72 | 4454   | 65  | 2,77  | 1,48 | 0,62 | 2,75 | 8,75  | 1,74 | 0,323 | 18,05 |
| 73 | 6786   | 102 | 2,41  | 1,8  | 1,73 | 4,42 | 8,38  | 1,67 | 0,441 | 15,39 |

Продовження табл. В.2

| 1   | 2      | 3    | 4     | 5    | 6    | 7    | 8     | 9    | 10    | 11    |
|-----|--------|------|-------|------|------|------|-------|------|-------|-------|
| 74  | 8520   | 335  | 2,56  | 1,75 | 1,25 | 3,22 | 6,57  | 1,1  | 0,104 | 3,85  |
| 75  | 16003  | 734  | 2,23  | 1,69 | 1,43 | 3,15 | 6,84  | 1,36 | 0,447 | 3,94  |
| 76  | 8854   | 100  | 3,6   | 3,34 | 2    | 4,08 | 8,45  | 1,65 | 0,240 | 21,08 |
| 77  | 9180   | 130  | 9,01  | 5,55 | 2,27 | 3,56 | 11,85 | 1,21 | 0,321 | 16,7  |
| 78  | 10045  | 315  | 2,68  | 1,52 | 0,91 | 2,3  | 7,52  | 1,38 | 0,277 | 6,35  |
| 79  | 9837   | 107  | 10,45 | 5,78 | 2,5  | 3,6  | 13,69 | 1,25 | 0,370 | 19,63 |
| 80  | 20315  | 736  | 3,11  | 1,89 | 2,14 | 4,34 | 7,28  | 1,08 | 0,261 | 5,47  |
| 81  | 12563  | 224  | 4,23  | 2,22 | 0,93 | 3,59 | 9,13  | 1,47 | 0,237 | 12,65 |
| 82  | 19005  | 596  | 4,28  | 3,66 | 1,51 | 2,98 | 8,76  | 1,5  | 0,232 | 6,48  |
| 83  | 17705  | 446  | 3,98  | 1,87 | 1,92 | 3,7  | 11,12 | 1,22 | 0,248 | 8,02  |
| 84  | 19827  | 395  | 4,4   | 1,82 | 1,06 | 3,3  | 10,34 | 1,3  | 0,379 | 11,61 |
| 85  | 30363  | 1021 | 2,9   | 1,64 | 1,19 | 2,79 | 7,7   | 1,25 | 0,292 | 6,03  |
| 86  | 41839  | 1050 | 3,41  | 1,81 | 1,95 | 2,91 | 7,48  | 1,28 | 0,699 | 8,25  |
| 87  | 41299  | 777  | 3,58  | 2,47 | 2,3  | 4,37 | 10,34 | 1,11 | 0,115 | 11,13 |
| 88  | 42205  | 872  | 5,59  | 1,93 | 1,03 | 2,19 | 13,31 | 1,09 | 0,308 | 10,67 |
| 89  | 36172  | 782  | 3,78  | 2,37 | 1,37 | 3,42 | 10,35 | 1,37 | 0,230 | 9,93  |
| 90  | 47502  | 450  | 8,82  | 3,33 | 1,42 | 4,09 | 15,96 | 1,31 | 0,264 | 27,38 |
| 91  | 97737  | 1442 | 3,43  | 1,63 | 0,96 | 2,18 | 7,25  | 1,27 | 0,311 | 16,79 |
| 92  | 3184   | 180  | 2,42  | 1,76 | 1,75 | 3,14 | 4,77  | 1,14 | 0,267 | 3,03  |
| 93  | 10413  | 340  | 2,73  | 1,24 | 0,7  | 3,51 | 9     | 1,47 | 0,346 | 4,54  |
| 94  | 11785  | 247  | 3,58  | 1,7  | 1,31 | 4,34 | 12,59 | 1,37 | 0,440 | 8,45  |
| 95  | 11133  | 213  | 4,49  | 1,85 | 0,46 | 4,38 | 14,88 | 1,3  | 0,085 | 9,09  |
| 96  | 21666  | 448  | 3,71  | 1,88 | 0,94 | 3,98 | 12,6  | 1,36 | 0,106 | 7,88  |
| 97  | 26717  | 712  | 3,51  | 1,72 | 0,59 | 4,15 | 10,24 | 1,17 | 0,048 | 6,55  |
| 98  | 36502  | 576  | 4,98  | 1,8  | 1,74 | 4,34 | 13,65 | 1,26 | 0,288 | 13,35 |
| 99  | 32125  | 660  | 3,93  | 2,03 | 1,42 | 4,82 | 10,82 | 1,22 | 0,700 | 8,8   |
| 100 | 42782  | 951  | 3,81  | 2,13 | 1,88 | 4,11 | 11,95 | 1,21 | 0,347 | 9,07  |
| 101 | 43557  | 718  | 4,81  | 2,35 | 2,15 | 4,78 | 15,38 | 1,25 | 0,470 | 10,16 |
| 102 | 45034  | 768  | 4,53  | 2,09 | 2,96 | 4,41 | 12,42 | 1,32 | 0,542 | 12,41 |
| 103 | 43267  | 409  | 8,78  | 2,99 | 2,1  | 4,25 | 17,56 | 1,47 | 0,400 | 29    |
| 104 | 79294  | 1502 | 4,63  | 1,71 | 1,51 | 4,2  | 13,24 | 1,14 | 0,299 | 9,89  |
| 105 | 97111  | 3062 | 3,3   | 1,8  | 2,22 | 3,82 | 8,26  | 1,22 | 0,291 | 5,47  |
| 106 | 93031  | 1250 | 5,5   | 2,62 | 2,46 | 4,4  | 14,35 | 1,17 | 0,373 | 17,09 |
| 107 | 83098  | 2319 | 5,07  | 1,9  | 0,4  | 3,88 | 10,93 | 1,11 | 0,178 | 7,14  |
| 108 | 202655 | 4934 | 3,59  | 2,33 | 3,55 | 4,75 | 9,21  | 1,17 | 0,608 | 7,28  |
| 109 | 349736 | 7743 | 4,04  | 1,56 | 1,73 | 3,98 | 10,75 | 1,14 | 0,195 | 9,44  |
| 110 | 1008   | 29   | 4,45  | 2,93 | 1,38 | 3,45 | 9,59  | 2,17 | 0,138 | 7,52  |
| 111 | 1053   | 25   | 4,92  | 3,12 | 1,56 | 3,28 | 9,64  | 1,82 | 0,385 | 9,12  |



Кінець табл. В.2

| 1   | 2      | 3    | 4    | 5    | 6    | 7    | 8     | 9    | 10    | 11    |
|-----|--------|------|------|------|------|------|-------|------|-------|-------|
| 112 | 2697   | 35   | 7,23 | 4    | 1,26 | 5,29 | 13,43 | 2    | 0,000 | 15,46 |
| 113 | 4863   | 90   | 4,7  | 0,83 | 2,26 | 6,46 | 12,13 | 1,11 | 0,086 | 10,86 |
| 114 | 6207   | 146  | 4,05 | 2,01 | 2,4  | 4,19 | 8,55  | 2,05 | 0,664 | 9,77  |
| 115 | 6335   | 153  | 4,11 | 2,22 | 0,97 | 2,84 | 8,47  | 2    | 0,399 | 9,54  |
| 116 | 8076   | 160  | 3,41 | 1,36 | 1,98 | 2,71 | 8,48  | 1,72 | 0,411 | 7,62  |
| 117 | 9317   | 297  | 3,63 | 1,94 | 1,65 | 3,29 | 8,13  | 1,16 | 0,036 | 5,74  |
| 118 | 7575   | 111  | 4,98 | 2,47 | 3,12 | 5,23 | 11,17 | 2,17 | 0,495 | 16,41 |
| 119 | 11656  | 164  | 4,46 | 2,47 | 1,6  | 4,27 | 10,51 | 1,85 | 0,287 | 15,4  |
| 120 | 28397  | 600  | 3,74 | 2,47 | 0,95 | 2,96 | 8,56  | 1,84 | 0,207 | 8,97  |
| 121 | 27627  | 603  | 4,93 | 3,12 | 1,44 | 3,21 | 11,22 | 1,59 | 0,655 | 10,07 |
| 122 | 30017  | 832  | 3,64 | 2,07 | 1,56 | 2,9  | 10,52 | 1,46 | 0,246 | 7,97  |
| 123 | 51257  | 1226 | 4    | 2,13 | 2,11 | 3,19 | 8,39  | 1,83 | 0,544 | 9,79  |
| 124 | 119314 | 1428 | 4,26 | 1,97 | 0,54 | 1,14 | 14,37 | 1,06 | 0,326 | 12,07 |
| 125 | 91125  | 1137 | 6,3  | 3,03 | 2,9  | 4,69 | 15,14 | 1,84 | 0,564 | 20,45 |

## ДОДАТОК Г. СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

### Список публікацій здобувача за темою дисертації:

#### 1) в яких висвітлені наукові результати дисертації:

- статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України:

1. **Приходько А. С.** Математичні моделі для визначення складності об'єктно-орієнтованого проектування через зв'язки між класами веб-застосунків з відкритим кодом, створених за допомогою PHP фреймворків. *Прикладні питання математичного моделювання*. 2025. Том 8. № 1. – С. 182-188. DOI: <https://doi.org/10.32782/mathematical-modelling/2025-8-1-17> (Фахове видання категорії Б).

2. **Приходько А. С., Малахов Є. В.** Удосконалена математична модель оцінювання якості вебдодатків з відкритим кодом. *Центральноукраїнський науковий вісник. Технічні науки*. 2025. Вип. 11(42), ч. II. С. 30-37. DOI: [https://doi.org/10.32515/2664-262X.2025.11\(42\).2.30-37](https://doi.org/10.32515/2664-262X.2025.11(42).2.30-37) (Фахове видання категорії Б).

- статті у періодичних наукових виданнях, проіндексованих у базах даних *Web of Science Core Collection* та/або *Scopus* (крім видань держави, визнаної Верховною Радою України державою-агресором):

3. **Prykhodko S. B., Shutko I. S., Prykhodko A. S.** A nonlinear regression model to estimate the size of web apps created using the CakePHP framework. *Radio Electronics, Computer Science, Control*. 2021. No. 4 (59). P. 129-139. DOI:

<http://dx.doi.org/10.15588/1607-3274-2021-4-12> (Фахове видання категорії А, **Web of Science**)

4. **Prykhodko A. S.**, Malakhov E. V. Determining object-oriented design complexity due to the identification of classes of open-source web applications created using PHP frameworks. *Radio Electronics, Computer Science, Control*. 2024. No. 2 (69). P. 160-166. DOI: <https://doi.org/10.15588/1607-3274-2024-2-16> (Фахове видання категорії А, **Web of Science**)

- стаття у виданні, віднесеному до першого - третього квартилів (Q1-Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports, чи одноосібна монографія, що відповідає зазначеним вимогам, прирівнюється до двох наукових публікацій:

5. Prykhodko S. B., Shutko I. S., **Prykhodko A. S.** Early size estimation of web apps created using Codeigniter framework by nonlinear regression models. *Radioelectronic and computer systems*. 2022. Vol. 3 (103). P. 84-94. DOI: <https://doi.org/10.32620/reks.2022.3.06> (Фахове видання категорії А, **Scopus, Q3, SJR 2022 0,307**). Стаття у виданні, віднесеному до третього квартилю (Q3) відповідно до класифікації SCImago Journal and Country Rank: <https://www.scimagojr.com/journalsearch.php?q=21101038702&tip=sid&clean=0>

2) які засвідчують апробацію матеріалів дисертації:

6. Prykhodko S., Shutko I., **Prykhodko A.** Early LOC Estimation of Web Apps Created Using Yii Framework by Nonlinear Regression Models. *WSEAS Transactions on Computers*. 2021. Vol. 20. P. 321-328. DOI: <http://dx.doi.org/10.37394/23205.2021.20.35> ("WSEAS Transactions on Computers" є міжнародним періодичним науковим виданням, Греція; стаття включена до міжнародних наукометричних баз: Google Scholar, CrossRef)

7. Prykhodko S., **Prykhodko A.**, Shutko I. Estimating the Size of Web Apps Created Using the CakePHP Framework by Nonlinear Regression Models with Three Predictors. *Computer Sciences and Information Technologies (CSIT)* : proceedings of the 2021 IEEE 16th International Conference (Lviv, 22-25 September 2021). IEEE, 2021. P. 333-336. DOI: <http://dx.doi.org/10.1109/CSIT52700.2021.9648680> (**Scopus**)

8. Приходько С. Б., Ворона М. В., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання розміру web-застосунків, що створюються за допомогою РНР-фреймворків. *Проблеми моделювання та розроблення інформаційних систем* : матеріали V науково-практичної інтернет-конференції (Дрогобич, 23 квітня 2021 року). Дрогобич: ДДПУ ім. І. Франка, 2021. С. 28-31.

9. Ворона М. В., **Приходько А. С.**, Шутко І. С. Трьохфакторні нелінійні регресійні моделі для оцінювання розміру WEB-застосунків, що створюються з використанням РНР-фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей вісімнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 23 квітня 2021 р. Одеса, 2021. С. 111-113.

10. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Нелінійні регресійні моделі для оцінювання кількості рядків коду Web-застосунків, що створюються за допомогою РНР-фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2021)* : матеріали II-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2021 р.). Миколаїв: НУК імені адмірала Макарова, 2021. С. 11-12. URL: <http://itconf.nuos.edu.ua/2021/proceedings/>

11. Приходько С. Б., **Приходько А. С.**, Шутко І. С. Математичне моделювання кількості рядків коду Web-застосунків, що створюються за допомогою фреймворку CakePHP. *Стан, досягнення та перспективи інформаційних систем і технологій* : матеріали XXII Всеукраїнської науково-

технічної конференції молодих вчених, аспірантів та студентів (Одеса, 21-22 квітня 2022 р.). Одеса, Видавництво ОНТУ, 2022. С.36.

12. **Приходько А. С.** Аналітична модель чотиривимірного закону розподілу метрик Web застосунків, що створюються за допомогою фреймворку CakePHP. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2022)* : матеріали III-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 26-28 жовтня 2022 р.). Миколаїв: НУК імені адмірала Макарова, 2022. С. 52-53. URL: <http://itconf.nuos.edu.ua/2022/wp-content/uploads/sites/5/NUOS-ITMAS-2022-Proceedings.pdf>

13. **Приходько А. С., Малахов Є. В.** Визначення метрик Чидамбера та Кемерера за допомогою програмного забезпечення з відкритим кодом. *Free and Open Source Software* : матеріали XIV-ої Міжнародної науково-практичної конференції (Харків, 14-16 лютого 2023 р.) Харків: Харківський національний економічний університет імені Семена Кузнеця, 2023. С. 91. URL: <http://repository.hneu.edu.ua/bitstream/123456789/29041/1/foss-2023-theses.pdf>

14. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проектування через ідентифікацію класів веб застосунків, що створюються за допомогою PHP фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2023)* : матеріали IV-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2023 р.). Миколаїв: НУК імені адмірала Макарова, 2023. С. 35-37. URL: <https://itconf.nuos.edu.ua/2023/proceedings/>

15. **Приходько А. С.** Побудова математичної моделі для оцінювання складності об'єктно-орієнтованого проектування веб застосунків, що створюються за допомогою PHP фреймворків. *Комп'ютерне моделювання та оптимізація складних систем* : матеріали VIII Міжнародної науково-технічної конференції (Дніпро, 1-3 листопада 2023 року). Дніпро: Державний вищий

навчальний заклад «Український державний хіміко-технологічний університет», 2023. С. 61-62. URL: <https://udhtu.edu.ua/wp-content/uploads/2023/11/zbirnyk-tez-kmoss-2023.pdf>

16. **Приходько А. С.** Визначення складності об'єктно-орієнтованого проєктування за рахунок ідентифікації класів веб застосунків розроблених на основі PHP фреймворків. *Інформаційні технології: теорія і практика* : тези доповідей I (VII) міжнародна науково-практична конференція здобувачів вищої освіти і молодих учених (Дніпро, 20 – 22 березня 2024). / М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». Дніпро : Свідлер А.Л., 2024. С. 310-312. URL: [https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024\(Dnipro\).pdf](https://sau.nmu.org.ua/ua/science/conference/ITTP/international/ITTP2024(Dnipro).pdf)

17. **Приходько А. С.** Математична модель для оцінювання складності об'єктно-орієнтованого проєктування через семантику класів вебзастосунків, що створюються за допомогою PHP фреймворків. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2024)* : матеріали V-ої Всеукраїнської науково-практичної інтернет конференції (Миколаїв, 30-31 жовтня 2024 р.). Миколаїв: НУК імені адмірала Макарова, 2024. С. 17-19. URL: <https://itconf.nuos.edu.ua/2024/proceedings/>

18. **Приходько А. С.** Первинна обробка інформації з метрик RFC та СВО веб застосунків, що створені за допомогою PHP фреймворків. *Інформаційні технології і автоматизація* : матеріали XVII міжнародної науково-практичної конференції (Одеса, 31 жовтня - 1 листопада 2024 р.). Одеса, Видавництво ОНТУ, 2024. С. 498-500. URL: <https://ontu.edu.ua/download/konfi/2024/Collection-of-abstracts-of-the-conference-ITIA-2024.pdf>

19. **Приходько А. С.** Використання метрик RFC, СВО та WMC для оцінювання якості веб-застосунків з відкритим кодом, створених за допомогою

РНР фреймворків. *Інженерія програмного забезпечення і передові інформаційні технології (Soft Tech-2024)* : матеріали VII Міжнародної науково-практичної конференції молодих вчених та студентів (Київ, 20-22 листопада 2024 року). Київ: Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», ФІОТ, 2024. С. 82-85. URL: <https://drive.google.com/file/d/1O70Ysxe-z3SS82UqEaBdEXR2VdRW3NwG/view>

20. **Приходько А. С.** Математичні моделі для первинної обробки інформації з метрик RFC, СВО та WMC вебзастосунків, що створені з використанням РНР фреймворків. *Інформатика, інформаційні системи та технології* : тези доповідей двадцять другої всеукраїнської конференції студентів і молодих науковців. Одеса, 25 квітня 2025 р. Одеса, 2025. С. 60-61. URL: <http://dspace.pdpu.edu.ua/bitstream/123456789/21944/1/Informatics%2C%20information%20systems.pdf#page=60>