

РЕЦЕНЗІЯ

на дисертаційну роботу

Орехова Олександра Станіславовича

«Математичні моделі для обробки інформації з метрик коду JAVA-застосунків
для оцінювання їх розміру»,

яка подана на здобуття наукового ступеня доктора філософії

в галузі знань 12 – «Інформаційні технології»

за спеціальністю 122 – «Комп'ютерні науки»

Актуальність теми дослідження

Дисертаційна робота присвячена вирішенню важливого науково-практичного завдання – підвищення достовірності обробки інформації з метрик коду JAVA-застосунків для оцінювання їх розміру шляхом побудови відповідних математичних моделей, а саме нелінійних регресійних моделей, та створенню інструментарію інформаційної технології для обробки інформації з метрик коду, необхідних для оцінювання параметра розміру на ранніх стадіях розробки програмних застосунків за метриками UML-діаграми класів.

Актуальність дослідження зумовлена широким застосуванням мови програмування JAVA у сучасній індустрії розробки програмного забезпечення, зокрема у створенні веб-, мобільних, корпоративних застосунків, а також програмних рішень з використанням технологій Data Science та Machine Learning. Особливе значення має попереднє оцінювання розміру JAVA-застосунків на ранніх стадіях проєктування, оскільки саме цей параметр є основним для подальшого прогнозування трудомісткості, бюджету та часу розробки програмних проєктів. Недостатня достовірність такого оцінювання може призводити до суттєвих відхилень фактичних витрат від запланованих та термінів реалізації, а також неефективного використання ресурсів команди. У зв'язку з цим існує потреба у розробці та вдосконаленні математичних моделей, які дозволяють підвищити достовірність обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків.

Дисертацію виконано відповідно до планів науково-дослідної роботи кафедри програмного забезпечення автоматизованих систем Національного університету кораблебудування імені адмірала Макарова в межах ініціативної теми «Математичні моделі обробки інформації з метрик коду для оцінювання параметра розміру JAVA-застосунків на ранніх етапах розробки програмних проєктів» (Державний реєстраційний номер: 0125U002663).

Структура та зміст дисертації

Дисертаційна робота включає вступ, чотири розділи, висновки, список використаних джерел, що налічує 107 найменувань, а також 8 додатків. Загальний обсяг роботи складає 223 сторінок, з яких 133 сторінок займає основний текст. У роботі представлено 5 рисунків та 44 таблиці. Додатки містять документи, що підтверджують впровадження отриманих результатів, список наукових публікацій автора за темою дисертації, інформацію щодо апробації результатів дослідження, публікацію, в якій проведено аналіз успішності розробки проєктів ПЗ, набори даних проєктів JAVA-застосунків з відкритим кодом, інструкцію користувача та вихідний код інформаційної технології.

У вступі сформульовано науково-практичну проблему, що розв'язується в дослідженні, висвітлено його значущість і актуальність, окреслено мету та основні завдання, визначено наукову новизну й практичну цінність отриманих результатів, а також наведено відомості про їх апробацію та публікації.

Перший розділ містить аналіз та порівняння існуючих математичних моделей обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків, формування наборів даних загальної вибірки та для DS/ML JAVA-застосунків, а також обґрунтування необхідності проведення подальших досліджень у вибраному напрямку.

Другий розділ присвячений огляду та аналізу існуючих одновимірних і багатовимірних взаємозворотних нормалізуючих перетворень, а також обґрунтовано вибір багатовимірного перетворення Бокса-Кокса для нормалізації шестивимірних векторів метрик коду загальних JAVA-застосунків та DS/ML JAVA-застосунків з метою їх подальшого використання при побудові нелінійних регресійних моделей.

Третій розділ зосереджений на побудові багатофакторних нелінійних регресійних моделей обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків із застосування перетворення Бокса-Кокса, розробці інтервалів прогнозування та довірчих інтервалів, а також порівнянню моделей між собою за допомогою таких критеріїв якості як коефіцієнт детермінації, *MMRE*, *PRED(0,25)*.

У четвертому розділі представлено розроблену інформаційну технологію, із використанням мови програмування Kotlin, де побудовані математичні моделі використовуються для оцінювання параметра розміру JAVA-застосунків.

Ступінь обґрунтованості положень, висновків і рекомендацій, сформульованих у дисертації

Дисертаційна робота Орехова О. С. вирізняється комплексним та системним підходом, демонструючи високий рівень наукових узагальнень. Представлені автором наукові положення, висновки та рекомендації ґрунтуються на ретельному аналізі та мають високий рівень обґрунтованості. Їх теоретичною основою є праці провідних фахівців у галузі комп'ютерних наук та інформаційних технологій.

Тема дисертаційної роботи відповідає проведенню в ній дослідженням та отриманим результатам. У роботі послідовно розглядаються питання підвищення достовірності обробки інформації для оцінювання розміру JAVA-застосунків в задачах оцінювання трудомісткості розробки, із застосуванням нормалізуючих перетворень та статистичних методів для аналізу негаусівських даних. Запропоновані підходи узгоджуються із загальною науковою концепцією роботи та її практичними цілями.

Аналіз змісту дисертації свідчить про чітке формулювання актуальності теми, наукової мети, завдань, об'єкта та предмета дослідження. Достовірність отриманих результатів забезпечується коректною постановкою задачі, використанням сучасних методів теорії ймовірностей, математичної статистики, багатовимірною статистичною аналізом та перевіркою достовірності оцінювання на незалежних даних. Це дозволяє здобувачу ефективно оцінювати характеристики випадкових величин, перевіряти статистичні гіпотези щодо нормальності розподілу даних, визначати параметри нормалізуючих перетворень та будувати нелінійні регресійні моделі.

Експериментальні дослідження, проведені на реальних наборах даних JAVA-застосунків з відкритим кодом, підтверджують ефективність розроблених підходів. Отримані результати демонструють підвищення достовірності обробки інформації для оцінювання розміру JAVA-застосунків завдяки використанню запропонованих моделей.

Обґрунтованість та достовірність висновків і рекомендацій підтверджуються апробацією результатів дослідження на науково-практичних конференціях, а також їхньою публікацією в наукових виданнях. Це свідчить про значну наукову цінність роботи та її вагомий внесок у розвиток напрямку оцінювання параметра розміру ПЗ на ранніх етапах проєктування. Дисертаційна робота містить вичерпне наукове обґрунтування отриманих результатів та відповідає заявленій тематиці дослідження.

Рівень реалізації визначених наукових завдань

Дисертаційне дослідження містить сім ключових завдань, виконання яких проаналізовано нижче.

1. *Провести аналіз існуючих математичних моделей для обробки інформації з метрик коду для оцінювання розміру загальних та DS і ML JAVA-застосунків, із використанням відповідних вибірок проєктів з відкритим кодом.* Перше завдання реалізоване у повному обсязі. Було проведено детальне дослідження різних методів і математичних моделей, що застосовуються в обробці інформації для оцінювання розміру ПЗ. Особлива увага приділена дослідженню наборів метрик коду та верифікації отриманих результатів.

2. *Сформувати навчальні та тестові вибірки з метрик коду JAVA-застосунків та обґрунтувати вибір незалежних факторів для побудови нелінійних регресійних моделей обробки інформації з метрик коду для оцінювання* *їх* *розміру.* Завдання виконано в повному обсязі. Було зібрано набори даних метрик коду загальних JAVA-застосунків та DS/ML JAVA-застосунків. Проведено вибір, трансформацію та попередню обробку факторів для подальшої побудови нелінійних регресійних моделей.

3. *Дослідити отримані метрики на відсутність мультиколінеарності за критерієм VIFs, з урахуванням ітеративної обробки багатовимірних даних.* Проведено ітеративну перевірку комбінації незалежних факторів на наявність мультиколінеарності між ними за критерієм VIFs та проведено трансформацію окремих незалежних факторів для вирішення проблеми мультиколінеарності.

4. *Обрати перетворення для нормалізації шестивимірних векторів характеристик метрик коду JAVA-застосунків на основі аналізу існуючих нормалізуючих перетворень та методів оцінювання їх параметрів.* Завдання виконано в повному обсязі. Було розглянуто широкий спектр підходів до нормалізації даних, що використовуються для наближення їхнього розподілу до нормального. Проведено ретельний аналіз методів оцінювання параметрів нормалізуючих перетворень, що дало змогу визначити оптимальні методи для роботи з шестивимірними векторами характеристик метрик коду загальних та DS/ML JAVA-застосунків.

5. *Удосконалити чотирьохфакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків із використанням п'ятивимірного нормалізуючого перетворення Бокса-Кокса на основі загальної вибірки метрик коду JAVA-застосунків з відкритим кодом, які*

можна отримати з концептуальної моделі проєкту. Здійснено задачу з покращення чотирьохфакторної нелінійної регресійної моделі на основі сучасної вибірки загальних JAVA-застосунків та її інтервалів передбачення. Надійність моделі підтверджено за допомогою збалансованого поділу набору даних загальних JAVA-застосунків на навчальну та тестову вибірки. Це дозволило підвищити достовірність обробки інформації для оцінювання розміру ПЗ.

6. Побудувати п'ятифакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків із застосуванням шестивимірною нормалізуючого перетворення Бокса-Кокса на основі загальної вибірки метрик коду JAVA-застосунків з відкритим кодом, які можна отримати з концептуальної моделі проєкту. Було розроблено п'ятифакторну нелінійну регресійну модель обробки інформації з метрик коду для оцінювання розміру загальних JAVA-застосунків. Модель вперше запропонувала розділення метрики класів та інтерфейсів на дві окремі метрики для підвищення достовірності оцінювання, а також має кращі показники якості оцінювання у порівнянні з іншими регресійними моделями. Виконання завдання підтверджено у третьому розділі дисертаційної роботи.

7. Побудувати п'ятифакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків із застосуванням шестивимірною нормалізуючого перетворення Бокса-Кокса на основі вибірки метрик коду DS та ML JAVA-проєктів з відкритим кодом, які можна отримати з концептуальної моделі проєкту; Здобувачем вперше запропоновано побудова п'ятифакторної нелінійної регресійної моделі обробки інформації для оцінювання розміру актуальної категорії DS/ML JAVA-застосунків. Отримана модель, в порівнянні з іншими регресійними моделями, має більший відсоток прогнозованих значень, менше значення середньої величини відносної похибки та менші ширини інтервалу передбачення та довірчого інтервалу нелінійної регресії. Завдання повністю виконане.

8. Розробити інформаційну технологію (SaaS програмне рішення) та методику для обробки інформації з метрик коду, отриманих з UML-діаграми класів, для оцінювання параметра розміру JAVA-застосунків із використанням побудованих п'ятифакторних нелінійних регресійних моделей.. На основі отриманих результатів розроблено інформаційну технологію мовою програмування Kotlin із використанням фреймворку Spring, яка забезпечує проведення розрахунку параметра розміру за допомогою побудованих

нелінійних регресійних моделей на ранніх етапах розробки ПЗ. Завдання реалізоване в повному обсязі.

Таким чином, усі поставлені у дисертаційній роботі завдання виконані, що підтверджує досягнення поставленої наукової мети.

Наукова новизна отриманих результатів

Оцінюючи здобутки дисертаційного дослідження, варто вказати на наступні результати, що мають певну наукову новизну та були отримані автором особисто:

- удосконалено чотирьохфакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання параметра розміру JAVA-застосунків залежно від загальної кількості класів та інтерфейсів (CLASS), загальної кількості унікальних викликів методів у класах (RFC), середнього значення кількості зв'язків між класами (aCBO) та середнього значення кількості видимих (public та protected) методів на клас (aVMQ) на основі п'ятивимірного перетворення Бокса-Кокса, що дозволило підвищити достовірність оцінювання параметра розміру таких застосунків як негаусівської випадкової величини у порівнянні з існуючими регресійними моделями. Модель, що побудована, у порівнянні з іншими моделями має більший відсоток прогнозованих значень, менше значення середньої величини відносної похибки та менші ширини інтервалу передбачення нелінійної регресії;

- вперше побудовано п'ятифакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання параметра розміру JAVA-застосунків на основі шестивимірного перетворення Бокса-Кокса, в якій, на відміну від існуючих моделей, було розділено метрику загальної кількості класів та інтерфейсів (CLASS) на дві окремі метрики, а саме кількість класів (CLS) та кількість інтерфейсів (INFC), що дозволило підвищити достовірність обробки інформації з метрик коду для оцінювання параметра розміру в порівнянні з існуючими математичними моделями. Модель, що побудована, в порівнянні з іншими регресійними моделями, має кращі показники якості R^2 , $MMRE$ та $PRED(0,25)$ та менші ширини інтервалу передбачення та довірчого інтервалу нелінійної регресії;

- вперше побудовано п'ятифакторну нелінійну регресійну модель для обробки інформації з метрик коду для оцінювання параметра розміру DS та ML JAVA-застосунків на основі шестивимірного перетворення Бокса-Кокса, в якій, на відміну від існуючих моделей, було розділено метрику загальної кількості

класів та інтерфейсів (CLASS) на дві окремі метрики коду кількості класів (CLS) та інтерфейсів (INFC) та враховано особливості розробки цієї категорії застосунків, що дозволило підвищити достовірність обробки інформації з метрик коду для оцінювання параметра розміру у порівнянні із використанням одновимірного та п'ятивимірного перетворень Бокса-Кокса. Модель, що побудовано, в порівнянні з іншими регресійними моделями, має більший відсоток прогнозованих значень, менше значення середньої величини відносної похибки та менші ширини інтервалу передбачення та довірчого інтервалу нелінійної регресії;

- отримали подальший розвиток рівняння для визначення границь інтервалів передбачення п'ятифакторної нелінійної регресії для загальних JAVA-застосунків та DS і ML JAVA-застосунків на основі шестивимірних нормалізуючих перетворень Бокса-Кокса. Це у порівнянні з використанням п'ятивимірних перетворень Бокса-Кокса дозволяє зменшити ширини інтервалів передбачення нелінійних регресій для оцінювання параметра розміру відповідних програмних застосунків і тим самим підвищити достовірність оцінювання зазначеного розміру як залежної випадкової величини.

Практичне значення одержаних результатів

Розробка інструментарію ІТ із використанням мови програмування Kotlin дозволила підвищити достовірність обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків на ранніх етапах проектування ПЗ, що підтверджується актами впровадження результатів дисертаційної роботи в ТОВ «Н-ІКС ДЕЛІВЕРІ» та ФОП Поліщук Андрій Борисович.

Розробки автора застосовуються при викладанні студентам Навчально-наукового інституту комп'ютерних наук та управління проектами (ННІКНУП) Національного університету кораблебудування ім. адмірала Макарова.

Повнота викладу результатів дослідження в опублікованих працях

Результати дисертаційної роботи викладені у 18 публікаціях та 1 свідоцтві про реєстрацію авторського права. В опублікованих роботах відображені основні положення та результати дисертації.

Основні наукові результати дисертації опубліковано у 4 статтях, що на момент публікації входили до переліку фахових наукових видань України та 1 одноосібному розділі у колективній монографії. Ще 12 публікацій засвідчують апробацію матеріалів дисертації, з яких 3 статті конференцій, що індексуються

в базах даних Scopus (за винятком видань з держав, визнаних Верховною Радою України державами-агресорами). До праць, які додатково відображають наукові результати дисертації, належать 1 розділ в колективній монографії у співавторстві та 1 свідоцтво про реєстрацію авторського права.

Усі публікації прямо пов'язані з тематикою дисертації і достатньо повно розкривають її основні положення. У працях, опублікованих у співавторстві, здобувачем відокремлений авторський внесок. На всі використані в роботі джерела інформації є посилання.

Обсяг надрукованих праць та їх кількість відповідають необхідним вимогам щодо публікацій на здобуття ступеня доктора філософії відповідно до Постанови Кабінету Міністрів України № 44 від 12.01.2022 р. «Порядок присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (із змінами, внесеними згідно з Постановами КМ № 341 від 21.03.2022, № 502 від 19.05.2023, № 507 від 03.05.2024).

Відсутність порушення академічної доброчесності

В процесі ознайомлення із дисертаційною роботою та із звітом щодо перевірки тексту рукопису електронним сервісом Identific від 17.03.2026 р., порушень академічної доброчесності не встановлено. Всі наукові положення та результати інших науковців, використанні здобувачем у дисертаційній роботі, мають коректні посилання на відповідні джерела.

Оцінка змісту дисертації і відповідності вимогам до оформлення

Аналіз дисертаційної роботи дозволяє зробити висновок, що мету роботи досягнуто, а поставлені наукові завдання вирішено в повному обсязі, відповідно до предмета дослідження. Дисертація є завершеним науковим дослідженням.

Дисертація написана грамотною українською мовою, результати розрахунків, моделювання, вимірювальних даних узагальнені за допомогою сучасного математичного апарату та наведені у табличному й графічному вигляді. Отримані результати є достовірними, добре узгоджуються з актуальними науковими підходами і не викликають сумнівів.

Представлена дисертаційна робота за структурою, мовою та стилем викладення відповідає вимогам Міністерства освіти і науки України, що висуваються до дисертацій на здобуття наукового ступеня доктора філософії.

Оцінка рівня володіння методологією наукового дослідження

Здобувач демонструє системний підхід до проведення наукового дослідження, що відображається у всіх розділах роботи. Початковий етап включає всебічний аналіз існуючих методів та моделей, що застосовуються в процесі обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків. Це дало змогу виявити їхні обмеження при роботі з негаусівськими даними та обґрунтувати необхідність розробки багатофакторних нелінійних регресійних моделей. Далі було проведено детальне дослідження нормалізуючих перетворень, що дозволило визначити оптимальні підходи до нормалізації шестивимірних векторів характеристик метрик коду JAVA-застосунків.

У межах третього розділу запропоновано багатофакторні нелінійні регресійні моделі для оцінювання розміру загальних JAVA-застосунків та DS/ML JAVA-застосунків, які ґрунтуються на використанні перетворення Бокса-Кокса, а також побудовано їх інтервали передбачення та довірчі інтервали. Завершальний етап дослідження присвячений розробці інформаційної технології для обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків на основі побудованих моделей.

Уся робота структурована відповідно до наукових принципів: чітке формулювання проблеми, обґрунтування методів, їх реалізація та експериментальна перевірка. Отримані результати підтверджують не лише правильність обраного підходу, а й високий рівень володіння здобувачем методологією наукового дослідження.

Дискусійні положення та недоліки щодо змісту дисертаційної роботи

Враховуючи достатній рівень представлення теоретичних та експериментальних результатів наукових досліджень, слід відзначити наступні недоліки, питання та недостатньо обґрунтовані аспекти дисертаційної роботи:

1) У роботі зазначено що нелінійні регресійні моделі побудовані на основі наборів метрик JAVA-застосунків з відкритим кодом, але формально не наведено за якими критеріями відбувався відбір цих проєктів та як забезпечено

репрезентативність вибірки – чи виключалися форки, архівні або неактивні репозиторії.

2) Дисертаційна робота має обсяг 223 сторінки, але містить лише 5 рисунків у порівнянні з 44 таблицями. Бракує наочної візуалізації для складних багатовимірних процесів. Зокрема, загальний алгоритм побудови нелінійних регресійних моделей викладено лише в текстовій формі.

3) У лістингу результату обробки HTTP-запиту на отримання списку математичних моделей відсутній приклад структури даних поля "restrictions" та числові значення коефіцієнтів моделей мають українську локалізацію представлення чисел – з комою, замість загальноприйнятого представлення числових значень – з крапкою.

4) У роботі зустрічаються складні для сприйняття складнопідрядні речення та окремі стилістичні неточності, що можуть ускладнювати сприйняття матеріалу. Також деякі вирази відповідають старим правилам правопису в українській мові.

Зазначені недоліки не є визначальними та не впливають на загальну позитивну оцінку дисертації.

Загальна оцінка дисертаційної роботи

Оцінюючи дисертаційну роботу Орехова Олександра Станіславовича «Математичні моделі для обробки інформації з метрик коду JAVA-застосунків для оцінювання їх розміру» з урахуванням позитивних її аспектів, дискусійних положень та зауважень, можна стверджувати, що вона є завершеною науковою працею, у якій розв'язано конкретне наукове завдання – підвищення достовірності обробки інформації з метрик коду для оцінювання розміру JAVA-застосунків на ранніх етапах проектування ПЗ за рахунок побудови нелінійних регресійних моделей на основі вибірок проєктів JAVA-застосунків з відкритим кодом, що має важливе значення для галузі знань 12 – «Інформаційні технології» та спеціальності 122 – «Комп'ютерні науки».

Аналіз змісту дисертаційної роботи і опублікованих праць за темою дослідження дозволяє зробити висновок, що всі пункти наукової новизни, які мають місце, належним чином відображені у роботі.

Дисертаційна робота на тему «Математичні моделі для обробки інформації з метрик коду JAVA-застосунків для оцінювання їх розміру» відповідає вимогам Наказу Міністерства освіти і науки № 40 від 12.01.2017 р. «Про затвердження вимог до оформлення дисертації» (із змінами, внесеними

згідно з Наказом Міністерства освіти і науки № 759 від 31.05.2019) та Постанові Кабінету Міністрів України № 44 від 12.01.2022 р. «Порядок присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (із змінами, внесеними згідно з Постановами КМ № 341 від 21.03.2022, № 502 від 19.05.2023, № 507 від 03.05.2024), що висуваються до дисертації, поданої на здобуття ступеня доктора філософії, а її автор, Орехов Олександр Станіславович, заслуговує присудження наукового ступеня доктора філософії за спеціальністю 122 - «Комп'ютерні науки».

Рецензент

доцент кафедри програмного
забезпечення автоматизованих систем
Національного університету кораблебудування
імені адмірала Макарова,
кандидат технічних наук, доцент



Андрій ПУХАЛЕВИЧ

Підпис Пухалевича А.В. засвідчую:
начальник відділу кадрів
Національного університету кораблебудування
імені адмірала Макарова



Анатолій ЗУБАРЄВ