

ISSN 2307-5732

DOI 10.31891/2307-5732

НАУКОВИЙ ЖУРНАЛ

6.2022

ВІСНИК

**Хмельницького
національного
університету**

Том 1

Технічні науки

Technical sciences

SCIENTIFIC JOURNAL

HERALD OF KHMELNYTSKYI NATIONAL UNIVERSITY

2022, Issue 6, Volume 315, Part 1

Хмельницький

**ВІСНИК
ХМЕЛЬНИЦЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
серія: Технічні науки**

Затверджений як фахове видання категорії «Б»,
РІШЕННЯ АТЕСТАЦІЙНОЇ КОЛЕГІЇ № 1643 ВІД 28.12.2019 та №409 від 17.03.2020

Засновано в липні 1997 р.

Виходить 6 разів на рік

Хмельницький, 2022, № 6, Том 1 (315)

**Засновник і видавець: Хмельницький національний університет
(до 2005 р. – Технологічний університет Поділля, м. Хмельницький)**

Наукова бібліотека України ім. В.І. Вернадського http://nbuv.gov.ua/j-tit/Vchnu_tekh

Включено до науково-метричних баз:

Google Scholar	http://scholar.google.com.ua/citations?hl=uk&user=aUP9OYAAAAJ
Index Copernicus	http://jml2012.indexcopernicus.com/passport.php?id=4538&id_lang=3
Polish Scholarly Bibliography	https://pbn.nauka.gov.pl/journals/46221
CrossRef	http://doi.org/10.31891/2307-5732

Головний редактор	Скиба М. Є. , д.т.н., професор, заслужений працівник народної освіти України, член-кореспондент Національної академії педагогічних наук України, професор кафедри машин і апаратів, електромеханічних та енергетичних систем Хмельницького національного університету
Заступник головного редактора	Синюк О. М. , д.т.н., професор кафедри машин і апаратів, електромеханічних та енергетичних систем Хмельницького національного університету
Відповідальний секретар	Горященко С. Л. , к.т.н., доцент кафедри машин і апаратів, електромеханічних та енергетичних систем Хмельницького національного університету

Ч л е н и р е д к о л е г і ї
Технічні науки

Березненко С.М., д.т.н., Бойко Ю.М., д.т.н., Говорущенко Т.О., д.т.н., Гордєєв А.І., д.т.н., Горященко С. Л., к.т.н., Грабко В.В., д.т.н., Диха О.В., д.т.н., Защепкіна Н.М., д.т.н., Рубаненко О. О., д.с.н., Захаркевич О.В., д.т.н., Злотенко Б.М., д.т.н., Зубков А.М., д.т.н., Каплун П.В., д.т.н., Карташов В.М., д.т.н., Кичак В.М., д.т.н., Любош Хес, д.т.н., (Чехія), Мазур М.П., д.т.н., Мандзюк І.А., д.т.н., Мартинюк В.В., д.т.н., Мельничук П.П., д.т.н., Місяць В.П., д.т.н., Мясіщев О.А., д.т.н., Нелін Є.А., д.т.н., Павлов С.В., д.т.н., Параска О.А., д.т.н., Рогатинський Р.М., д.т.н., Горошко А.В., д.т.н., Сарібекова Ю.Г., д.т.н., Семенко А.І., д.т.н., Славінська А.Л., д.т.н., Харжевський В.О., д.т.н., Шинкарук О.М., д.т.н., Шклярський В.І., д.т.н., Щербань Ю.Ю., д.т.н., Бубулєс Альгімантас, доктор наук (Литва), Елсаєд Ахмед Ельнашар, доктор наук (Єгипет), Кальчиньскі Томаш, доктор наук (Польща), Лунтовський Андрій, д.т.н. (Німеччина), Матушевський Мацей, доктор наук (Польща), Мушлевський Лукаш, доктор наук (Польща), Мушля Януш, доктор наук (Польща), Натріашвілі Тамаз Мамієвич, д.т.н., (Грузія), Попов Валентин, доктор природничих наук (Німеччина)

<i>Технічний редактор</i>	Горященко К. Л., к.т.н.
<i>Редактор-коректор</i>	Броженко В. О.

**Рекомендовано до друку рішенням вченої ради Хмельницького національного університету,
протокол № 6 від 29.12.2022 р.**

Адреса редакції: редакція журналу "Вісник Хмельницького національного університету"
Хмельницький національний університет
вул. Інститутська, 11, м. Хмельницький, Україна, 29016

	(038-2) 67-51-08	http://journals.khnu.km.ua/vestnik
e-mail:	visnyk.khnu@khmnu.edu.ua	web: http://lib.khnu.km.ua/visnyk_tup.htm

Зареєстровано Міністерством України у справах преси та інформації.
Свідоцтво про державну реєстрацію друкованого засобу масової інформації
Серія КВ № 24922-14862ПР від 12 липня 2021 року

© Хмельницький національний університет, 2022
© Редакція журналу "Вісник Хмельницького національного університету", 2022

KAPTIUROVA DARIA, CHERTENKO LILIA, BONDAR OLEXANDR., NOS OLEXANDRA IMPROVEMENT OF THE INTERNAL FORM OF SHOES USING MODERN PROGRESSIVE TECHNOLOGIES	102
KRYVENCHUK YURI, PETRENKO DMYTRO CREATION OF DRONE STABILIZATION SYSTEM USING REINFORCEMENT LEARNING	111
KRYVENCHUK YURI, SHCHUR HLIB INFORMATION TECHNOLOGY FOR TRAINING THE PERSONNEL OF ENTERPRISES BY MEANS OF VIRTUAL REALITY	115
LATANSKA LIUDMYLA, MAKAROVA LIDIJA, KOLTSOV ANDRII, DAVLATOVA DIANA A NONLINEAR REGRESSION MODEL FOR ESTIMATING THE SIZE OF WEB APPLICATIONS CREATED USING SYMFONY FRAMEWORK.....	119
LYTVYN OLEG, MAKATORA DMYTRO, DZHIKIYA NELLIE DEVELOPMENT OF AN AUTOMATED SYSTEM FOR DETERMINING THE "ID OF THE COMPETITIVE PROPOSAL" OF THE SPECIALTY (EDUCATIONAL PROGRAM)	125
LIPIANINA-HONCHARENKO KHRYSTYNA, KOMAR MYROSLAV, SACHENKO ANATOLIY, LENDIUK TARAS METHOD OF FORMING THE CONTEXT OF ADVERTISING AND TARGET AUDIENCE BASED ON ASSOCIATIVE RULES LEARNING.....	131
MISHAN VIKTOR, KARPOVA LESIA, LOBACHOVA INNA MIMO ANTENNA ARRAY WITH DIRECTIONAL DIAGRAM SPREAD FOR 5G SMARTPHONES	139
MISHAN VIKTOR, KARPOVA LESIA, TSIMAR VALERII METHOD OF DESIGNING ANTENNAS BY SYNTHESIS OF A BAND FILTER.....	143
OBELOVSKA KVITOSLAVA, PELEKH KHRYSTYNA, PELEKH YURY. SNAICHUK YAROMIR ANALYSIS OF THE CSMA/CA SCHEME FOR WIRELESS LOCAL AREA NETWORKS	148
OLEKSANDRENKO VICTOR., SVIDERSKYI VLADISLAV, KIRICHENKO LUDMILA, DANILENKO IGOR, YEFYMENKO VALERI DEVELOPMENT AND RESEARCH OF THE NANOMODIFIED FLUOROPLASTIC COVERAGES INFLECTED ON METALLIC SURFACES	153
PAVLIUK ANNA, KERNESH VICTORIA USE OF DIGITAL TECHNOLOGIES IN WOMEN'S DRESS DESIGN	163
PASHKEVYCH OLEH, VAHSCHYSHCHAK SERHII, BOICHUK ADRII, STYSLO TARAS, DEMCHYNA MYKOLA APPLICATION OF MACHINE LEARNING MODELS FOR PREDICTING PRICES ON THE REAL ESTATE MARKET.....	170
PETRUSHAK VOLODYMYR IMPLEMENTATION OF THE COINCIDENCE FREQUENCY METER ON FPGA	179
PUKACH PAVLO REVIEW AND ANALYSIS OF BASIC FEATURE DETECTION NETWORKS FOR CLASSIFICATION OF MRI IMAGES IN DEEP LEARNING MODELS	183
RUBANKA MYKOLA, DVORZHAK VOLODYMYR, RUBANKA ALLA POLISHCHUK OLEH, HORIASHCHENKO SERHIY INNOVATIVE TECHNOLOGIES IN THE PRODUCTION OF SEWING PRODUCTS: CUTTING PRODUCTION	188
SAPOZHNYK DMYTRO, PELYK LESIA RAISING QUESTIONS ABOUT THE APPOINTMENT OF CUSTOMS EXPERTISE IN ORDER TO CONTROL THE CORRECT CLASSIFICATION OF GOODS UNDER HS	198

ЛАТАНСЬКА Людмила

Національний університет кораблебудування ім. адмірала Макарова

<https://orcid.org/0000-0001-6473-7624>e-mail: llatanskaya@gmail.com

МАКАРОВА Лідія

Національний університет кораблебудування ім. адмірала Макарова

<https://orcid.org/0000-0003-2903-3001>e-mail: lidia.makarova@nuos.edu.ua

КОЛЬЦОВ Андрій

Національний університет кораблебудування ім. адмірала Макарова

<https://orcid.org/0000-0002-1917-3038>e-mail: andrew.koltsv@gmail.com

ДАВЛАТОВА Діана

Національний університет кораблебудування ім. адмірала Макарова

<https://orcid.org/0000-0003-2495-1305>e-mail: didit77777@ukr.net

НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ PHP ФРЕЙМВОРКУ SYMFONY

Питання щодо визначення розміру програмного забезпечення на перших етапах роботи над проектом є важливим, адже отриманий розмір можна використати для прогнозування вартості, тривалості та трудомісткості. Частка веб-застосунків серед усіх розроблюваних проектів програмного забезпечення сьогодні є значною, як і популярність мов програмування, що використовуються у веб-розробці. Лідером за частотою використання є мова PHP з її численними фреймворками.

Метою дослідження є побудова нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються з використанням PHP фреймворку Symfony, за допомогою нормалізуючого перетворення на основі десятичного логарифму.

Для побудови моделі на платформі GitHub знайдено 50 проектів веб-застосунків, які створено з використанням фреймворку Symfony. З допомогою інструменту PhpMetrics отримано наступні метрики діаграми класів по цим проектам: кількість класів, середня кількість методів на кожен клас, глибина дерева наслідування. Такий вибір метрик зумовлений можливістю отримати таку інформацію про розроблюваний проект на ранніх стадіях проектування та відсутністю мультиколінеарності серед обраних метрик.

Емпіричні дані перевірено на нормальність. Перевірка показала, що дані є негаусівськими, тому їх нормалізовано за допомогою десятичного логарифму, виконано перевірку на наявність викидів, знайдені викиди усунуто та для даних, що лишилися, побудовано лінійну для нормалізованих даних та нелінійну для вихідних даних регресійні моделі. Для порівняння також побудовано лінійну регресійну модель для емпіричних даних у припущенні про нормальність розподілу. Порівняння моделей показало суттєве покращення якості оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony, при застосуванні саме нелінійного регресійного аналізу.

Ключові слова: оцінювання кількості строк коду, програмні застосунки, нелінійна регресійна модель, нормалізуюче перетворення, фреймворк Symfony.

LATANSKA Liudmyla, MAKAROVA Lidiia, KOLTISOV Andrii, DAVLATOVA Diana
Admiral Makarov National University of Shipbuilding

A NONLINEAR REGRESSION MODEL FOR ESTIMATING THE SIZE OF WEB APPLICATIONS CREATED USING SYMFONY FRAMEWORK

The issue of estimating the size of the software in the first stages of work on the project is important, because the obtained size can be used to predict the cost, duration and labor intensity. The share of web applications among all software projects developed today is significant, as is the popularity of programming languages used in web development. The leader in frequency of use is the PHP language with its numerous frameworks. The purpose of the study is to build a non-linear regression model for estimating the size of web applications created using the Symfony PHP framework, using a normalizing decimal logarithm transformation. To build the model, 50 web application projects were found on the GitHub platform, which were created using the Symfony framework. Using the PhpMetrics tool, the following class diagram metrics for these projects were obtained: the number of classes, the average number of methods per class, the depth of the inheritance tree. This choice of metrics is due to the possibility of obtaining such information about the project under development at the early stages of design and the absence of multicollinearity among the selected metrics.

Empirical data were tested for normality. The check showed that the data are non-Gaussian, so they were normalized using a decimal logarithm, an outlier check was performed, the outliers were eliminated, and for the remaining data, a linear regression model was built for the normalized data and a nonlinear one for the original data. For comparison, a linear regression model was also constructed for the empirical data under the assumption of normal distribution. The comparison of the models showed a significant improvement in the quality of estimating the size of web applications created using the Symfony framework when applying non-linear regression analysis.

Keywords: LOC estimation, software, non-linear regression model, normalizing transformation, Symfony framework.

КРИВЕНЧУК Ю., ПЕТРЕНКО Д. СИСТЕМА СТАБІЛІЗАЦІЇ ПОЛОЖЕННЯ ДРОНУ З ВИКОРИСТАННЯМ НАВЧАННЯ З ПІДКРІПЛЕННЯМ	111
КРИВЕНЧУК Ю., ЩУР Г. ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ НАВЧАННЯ ПЕРСОНАЛУ ПІДПРИЄМСТВ ЗАСОБАМИ ВІРТУАЛЬНОЇ РЕАЛЬНОСТІ	115
ЛАТАНСЬКА Л., МАКАРОВА Л., КОЛЬЦОВ А., ДАВЛАТОВА Д. НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ PHP ФРЕЙМВОРКУ SYMFONY	119
ЛИТВИН О., МАКАТЬОРА Д., ДЖІКІЯ Н. РОЗРОБКА АВТОМАТИЗОВАНОЇ СИСТЕМИ ВИЗНАЧЕННЯ «ID КОНКУРСНОЇ ПРОПОЗИЦІЇ» СПЕЦІАЛЬНОСТІ (ОСВІТНЬОЇ ПРОГРАМИ)	125
ЛІП'ЯНИНА-ГОНЧАРЕНКО Х., КОМАР М., САЧЕНКО А., ЛЕНДЮК Т. МЕТОД ФОРМУВАННЯ КОНТЕКСТУ РЕКЛАМИ ТА ЦІЛЬОВОЇ АУДИТОРІЇ НА ОСНОВІ НАВЧАННЯ АСОЦІАТИВНИХ ПРАВИЛ	131
МІШАН В., КАРПОВА Л., ЛОБАЧОВА І. АНТЕННА РЕШІТКА МІМО З РОЗНЕСЕННЯМ ДІАГРАМ СПРЯМОВАНOSTІ ДЛЯ СМАРТФОНІВ 5G	139
МІШАН В., КАРПОВА Л., ЦІМАР В. МЕТОД ПРОЕКТУВАННЯ АНТЕН ШЛЯХОМ СИНТЕЗУ СМУГОВОГО ФІЛЬТРА	143
ОБЕЛЬОВСЬКА К., ПЕЛЕХ Х., ПЕЛЕХ Ю., СНАЙЧУК Я. АНАЛІЗ СХЕМИ CSMA/CA БЕЗПРОВІДНИХ ЛОКАЛЬНИХ МЕРЕЖ	148
ОЛЕКСАНДРЕНКО В., СВІДЕРСЬКИЙ В., КИРИЧЕНКО Л., ДАНІЛЕНКО І., ЄФІМЕНКО В. РОЗРОБКА І ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ НАНОМОДИФІКОВАНИХ ФТОРОПЛАСТОВИХ ПОКРИТТІВ, НАНЕСЕНИХ НА МЕТАЛЕВІ ПОВЕРХНІ	153
ПАВЛЮК А., КЕРНЕС В. ВИКОРИСТАННЯ ЦИФРОВИХ ТЕХНОЛОГІЙ В ДИЗАЙНІ ЖІНОЧОГО КОСТЮМУ	163
ПАШКЕВИЧ О., ВАЩИЩАК С., БОЙЧУК А., СТИСЛО Т., ДЕМЧИНА М. ЗАСТОСУВАННЯ МОДЕЛЕЙ МАШИНОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ЦІН НА РИНКУ НЕРУХОМОСТІ	170
ПЕТРУШАК В. РЕАЛІЗАЦІЯ ЧАСТОТОМІРА КОІНЦИДЕНЦІЇ НА ПЛІС	179
ПУКАЧ П. ОГЛЯД ТА АНАЛІЗ ОСНОВНИХ КАРКАСНИХ МЕРЕЖ ВИЯВЛЕННЯ ОЗНАК ДЛЯ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ МРТ В МОДЕЛЯХ ГЛИБИННОГО НАВЧАННЯ	183
РУБАНКА М., ДВОРЖАК В., РУБАНКА А., ПОЛІЩУК О., ГОРЯЩЕНКО С. ІННОВАЦІЙНІ ТЕХНОЛОГІЇ ПРИ ВИГОТОВЛЕННІ ШВЕЙНИХ ВИРОБІВ: ОБЛАДНАННЯ РОЗКРІЙНОГО ВИРОБНИЦТВА	188
САПОЖНИК Д., ПЕЛИК Л. ПОСТАНОВКА ПИТАНЬ ЩОДО ПРИЗНАЧЕННЯ МИТНОЇ ЕКСПЕРТИЗИ З МЕТОЮ КОНТРОЛЮ ПРАВИЛЬНОСТІ КЛАСИФІКАЦІЇ ТОВАРІВ ЗА ТН ЗЕД	198
ФРИШЕВ С., ЛУКАЧ В., ІКАЛЬЧИК М., ВАСИЛЮК В. УДОСКОНАЛЕННЯ ТЕХНОЛОГІЇ ТРАНСПОРТУВАННЯ ЗЕРНА ВІД КОМБАЙНІВ	203
МАЛАЙКО А., ПАСІЧНИК О., СКРИПНИК Т. МЕТОД ПОБУДОВИ ОПТИМАЛЬНОЇ ОСВІТНЬОЇ ТРАЄКТОРІЇ ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ	208

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Визначення розміру програмного забезпечення є важливим питанням в області управління проектами програмного забезпечення. Це пов'язано з тим, що кількість успішних проектів, тобто таких, що завершилися в строк, мають належні показники якості та не вийшли за рамки бюджету, – лише 30–40% [1]. Визначений розмір програмного забезпечення може бути використаний для оцінювання вартості, тривалості та трудомісткості відомими моделями COCOMO, COCOMO II, ISBSG з метою використання отриманої інформації для зменшення ризиків проекту.

На сьогоднішній день значна частина розроблюваних проектів програмного забезпечення припадає на веб-застосунки. Найпопулярнішою мовою веб-розробки є PHP, загальний відсоток використання якої складає приблизно 71,4% у 2022 році [2]. Також досить часто використовуються фреймворки мови PHP [3, 4], що пов'язано зі значним переліком переваг, серед яких можна виділити головні: наявність готових рішень, великі можливості для масштабування веб-застосунків, використання провідних практик.

До популярних фреймворків мови PHP відносять Laravel, Symfony, Codeigniter, Yii, CakePHP. Для фреймворків Laravel [5], CakePHP [6], Yii [7] побудовано нелінійні регресійні моделі для оцінювання розміру веб-застосунків, що створюються з використанням цих фреймворків, тоді як для фреймворку Symfony, який за даними опитування sitepoint та компанії JetBrains йде другим за популярністю [3, 4], – ні. Це слугує підставою для побудови нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony.

Аналіз останніх досліджень та публікацій

В роботах [8, 9] Hee Beng Kuan Tan, Yuan Zhao та Hongyu Zhang побудували багатофакторні лінійні регресійні моделі для оцінювання розміру інформаційних систем для Java, Visual Basic та PHP. Залежною величиною для побудованих моделей виступає кількість строк коду, а в якості незалежних змінних було обрано такі метрики концептуальної моделі даних у вигляді діаграми класів, які можна отримати на ранніх стадіях роботи над проектом: кількість класів, кількість відносин між класами, середня кількість атрибутів. Однак на практиці існують певні проблеми з застосуванням лінійного регресійного аналізу: для його застосування залишки регресії або залежна змінна повинні мати гаусівський розподіл, що в реальних умовах трапляється вкрай нечасто. Все це зумовлює необхідність використання методів нелінійного регресійного аналізу.

Відомі на сьогодні нелінійні регресійні моделі для оцінювання розміру інформаційних систем, що розробляються з використанням мови PHP, можна поділити на дві групи за використаними метриками. Так у [10] застосовано ті ж метрики, що і в [8, 9], а в [11] використали такий набір метрик: кількість класів; сума середньої кількості класів, на які впливає даний клас, і середньої кількості класів, з яких даний клас отримує ефекти; середня кількість методів. В свою чергу якість побудованих нелінійних регресійних моделей тісно пов'язана з видом нормалізуючих перетворень, що й показано в цих роботах.

В роботах [6, 7] розглядали побудову нелінійних регресійних моделей для популярних фреймворків мови PHP – CakePHP та Yii. З припущення про можливу залежність розміру веб-застосунків, що розробляються з використанням фреймворків мови PHP, від інших метрик, ніж використані у роботах [8–11], був обраний такий перелік метрик: кількість класів, середня кількість методів на кожен клас, глибина дерева наслідування.

Для фреймворку Laravel [6] відома нелінійна регресійна модель, для побудови якої використали такий же набір метрик, що і в [11].

Високий інтерес та різні підходи до побудови нелінійних регресійних моделей для оцінювання розміру веб-застосунків, що використовують мову PHP чи її фреймворки, є підставою для проведення досліджень.

Постановка завдання

Метою роботи є підвищення достовірності оцінювання розміру веб-застосунків, що створюються з використанням PHP фреймворку Symfony, за допомогою нелінійної регресійної моделі.

Для досягнення мети необхідно виконати наступні завдання:

- Провести дослідження існуючих моделей для оцінювання розміру веб-застосунків.
- Виконати збір емпіричних даних, для яких буде побудовано нелінійну регресійну модель для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony.
- Виконати побудову нелінійної регресійної моделі для веб-застосунків, що створюються з використанням фреймворку Symfony, на основі нормалізуючого перетворення.

Виклад основного матеріалу

У випадку нормальності початкових даних в якості регресійної моделі можна взяти лінійну регресійну модель з наступним загальним виглядом:

$$Y = \hat{Y} + \varepsilon = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + \varepsilon, \quad (1)$$

де $\hat{Y}$ – результат передбачення за рівнянням лінійної регресії, b_i ($i = \overline{0, k}$) – параметри лінійного рівняння регресії, X_i ($i = \overline{1, k}$) – значення незалежної змінної, ε – залишки регресії, гаусівська випадкова величина, k – кількість факторів у моделі.

У випадку, коли початкові дані є негаусівськими, виникає потреба будувати нелінійну регресійну модель. Найкращим підходом до побудови нелінійних регресійних моделей є використання нормалізуючих перетворень, адже нормалізуючі перетворення вдається підібрати майже завжди.

Суть методу з використанням нормалізуючих перетворень можна представити наступним переліком кроків [12]:

- Нормалізувати емпіричні дані та вилучити викиди.
- За нормалізованими даними без викидів побудувати лінійну регресійну модель.
- Виконати перехід від лінійної регресійної моделі до нелінійної за зворотнім нормалізуючим перетворенням.

Нормалізуюче перетворення негаусівського випадкового вектору $P = \{Y, X_1, X_2, \dots, X_k\}^T$ у гаусівський випадковий вектор $T = \{Z_Y, Z_1, Z_2, \dots, Z_k\}^T$, задається як

$$T = \Psi(P), \quad (2)$$

а зворотнє перетворення до (2) як

$$P = \Psi^{-1}(T),$$

де Ψ – вектор, $\Psi = \{\Psi_Y, \Psi_1, \Psi_2, \dots, \Psi_k\}^T$.

Таким чином нелінійна регресійна модель, що побудована з використанням нормалізуючого перетворення, має вигляд

$$Y = \Psi_Y^{-1}(\hat{Z}_Y + \varepsilon) = \Psi_Y^{-1}(b_0 + b_1 Z_1 + b_2 Z_2 + \dots + b_k Z_k + \varepsilon),$$

де Ψ_Y – перша компонента вектору Ψ перетворення (2).

На першому кроці при побудові нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony, було проведено пошук веб-застосунків, що використовують цей фреймворк. За допомогою платформи GitHub (<https://github.com/>) було знайдено 50 проектів веб-застосунків. За допомогою інструменту PhpMetrics (<https://phpmetrics.org/>) було отримано наступні метрики діаграми класів по цим проектам: кількість строк коду KLOC (Y), кількість класів NOC (X_1), середня кількість методів на кожен клас Methods by Class (X_2), глибина дерева наслідування DIT (X_3).

Побудова багатофакторної регресійної моделі пов'язана з перевіркою майбутніх факторів (предикторів) на предмет наявності мультиколінеарності. При наявності мультиколінеарності модель має лінійну, або близьку до неї, залежність між факторами (незалежними змінними).

Для того щоб впевнитись у відсутності мультиколінеарності майбутніх факторів, визначимо коефіцієнти впливу дисперсії (Variance Inflation Factors, VIFs). В множинній лінійній регресійній моделі з k -предикторами X_i ($i = \overline{1, k}$) VIFs – це елементи, що знаходяться на головній діагоналі в оберненій коваріаційній матриці розміру $k \times k$ k -предикторів. Зазвичай значення VIFs, які більше за 10 вважають сигналом проблем з мультиколінеарністю. Значення VIFs від 1 до 5 свідчать про відсутність мультиколінеарності [13].

Для залежних змінних NOC, Methods by Class, DIT отримано значення VIFs 1,027, 1,030 та 1,036 відповідно. Всі значення менші за 5, що показує відсутність мультиколінеарності в даних.

Для пошуку викидів застосували метод, що базується на нормалізуючих перетвореннях та квадрати відстані Махаланобіса SMD [13, 14]. Для окремо взятої точки нормалізованих даних i ($i = \overline{1, N}$) застосовується наступна формула

$$d_i^2 = (Z_i - \bar{Z})^T S_N^{-1} (Z_i - \bar{Z}),$$

де Z_i – i -а точка багатовимірних даних нормалізованого вектору Z ; $\bar{Z}$ – вектор вибірових середніх; S_N – вибіркова коваріаційна матриця для нормалізованих даних. В нашому випадку розглядається 4-компонентний вектор.

$$S_N = \frac{1}{N} \sum_{i=1}^N (Z_i - \bar{Z}) (Z_i - \bar{Z})^T.$$

Для отримання вектору нормалізованих емпіричних даних необхідно скористатись (1) та нормалізуючим перетворенням десятичного логарифму. Тоді перетворення (1) має наступний вигляд $\Psi = \{\lg Y, \lg X_1, \lg X_2, \lg X_3\}^T$.

Для багатовимірних гаусівських даних розподіл величин SMD подібний до розподілу χ^2 , тобто для великих вибірок $N - m$ (принаймні більше 25) відстань d_i^2 має значення, що є близькими до незалежної випадкової величини $\chi_{m,\alpha}^2$, де $\chi_{m,\alpha}^2$ – квантіль розподілу χ^2 з рівнем значимості α [15].

Також для багатовимірних гаусівських даних значення статистики T_{S_i} поводить подібно F -розподілу з відповідним значенням квантілю $F_{m,N-m,\alpha}$

$$T_{S_i} = \frac{N(N-m)d_i^2}{m(N^2-1)}.$$

Багатовимірним викидом вважається рядок або зі значенням SMD, яке перевищує значення квантілю розподілу χ^2 , або рядок з розрахованою статистикою T_{S_i} більшою за значення $F_{m,N-m,\alpha}$. Такий рядок необхідно вилучити та продовжувати розраховувати SMD або T_{S_i} до повного усунення викидів.

Для вилучення викидів з початкового набору даних на рівні значимості 0,05 знадобилось 4 ітерації. На кожній з ітерацій було розраховано значення статистики T_{S_i} . Отримані значення статистики T_{S_i} на кожній з чотирьох ітерацій наведено в таблиці 1. Нагадаємо, що критичне значення квантіля F розподілу для рівня значимості 0,05 дорівнює 2,5740, 2,5837, 2,5943 та 2,6000 на 1, 2, 3 та 4 ітераціях відповідно.

Таблиця 1

Процес вилучення викидів з емпіричних даних

№	T_{S_i-1}	T_{S_i-2}	T_{S_i-3}	T_{S_i-4}	№	T_{S_i-1}	T_{S_i-2}	T_{S_i-3}	T_{S_i-4}
1	1,5076	1,7728	1,7580	1,7163	26	0,4815	0,7261	0,8917	0,8670
2	0,8954	0,9023	0,8672	0,8499	27	0,6947	0,8274	0,9752	0,9526
3	6,6096	-	-	-	28	3,2828	-	-	-
4	0,8130	1,0713	1,0625	1,0355	29	0,6176	0,6373	0,6280	0,6306
5	1,2702	1,2371	1,2372	1,2063	30	0,9046	1,1557	1,3381	1,3156
6	0,9084	0,8611	0,9626	1,1053	31	0,2714	0,3647	0,4337	0,4919
7	0,7460	0,7434	0,8878	1,1718	32	1,8882	2,5850	-	-
8	0,4795	1,0184	1,1713	1,1904	33	0,3456	0,4837	0,4903	0,4739
9	1,5053	1,5938	1,9385	2,1175	34	0,6937	0,8906	0,9023	0,8848
10	0,3212	0,4327	0,4828	0,4854	35	0,2762	0,3944	0,4553	0,4811
11	0,5818	0,8016	1,0512	1,2592	36	0,7500	0,9047	0,9820	0,9562
12	0,3834	0,4951	0,6455	0,7043	37	0,2218	0,2430	0,3120	0,3113
13	1,9333	2,1728	2,6837	-	38	1,2530	1,2889	1,4077	1,6315
14	0,5640	0,5413	0,5248	0,5118	39	0,1767	0,1753	0,1883	0,1831
15	0,2883	0,6797	0,7541	0,7314	40	0,3946	0,3795	0,4471	0,4466
16	0,5933	0,7120	0,8147	0,8016	41	0,4372	0,4476	0,5272	0,5365
17	0,8682	0,8270	0,7970	0,8624	42	0,2616	0,2529	0,2492	0,2439
18	0,3989	0,5985	0,5700	0,6003	43	0,8692	0,8399	0,8481	0,9993
19	1,6867	3,2489	-	-	44	0,4245	0,5683	0,6257	0,6214
20	0,8035	1,1098	1,1821	1,1935	45	0,4875	0,7510	0,9096	1,0013
21	0,4930	0,6381	0,6650	0,6895	46	0,6178	0,6134	0,6315	0,6121
22	0,6057	0,7973	0,9619	1,1463	47	1,5328	1,9403	2,1387	2,0878
23	0,8793	1,1561	1,1166	1,2089	48	0,8454	0,8383	0,8763	0,8733
24	0,9362	0,9576	1,0254	1,1213	49	0,8002	0,8378	0,8642	0,8427
25	0,3986	0,4626	0,5775	0,6305	50	1,0192	1,0422	1,1602	1,2365

Для побудови нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony, використаємо дані без викидів. Лінійна багатофакторна регресійна модель для нормалізованих даних має вигляд

$$Z_Y = \hat{Z}_Y + \varepsilon = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2 + \hat{b}_3 Z_3 + \varepsilon, \quad (3)$$

де $Z_Y = \lg Y; Z_i = \lg X_i, i = 1, 3$.

Оцінки параметрів побудованої моделі $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3$ дорівнюють -1,6378, 0,9664, 0,9630 та -0,4627 відповідно. Залишки лінійної регресії для нормалізованих даних ε були перевірені на нормальність за допомогою критерію χ^2 Пірсона. Відповідна нульова гіпотеза була прийнята на рівні значимості 0,05.

За лінійним регресійним рівнянням (3) та взаємо-зворотнім нормалізуючим перетворенням отримуємо нелінійну регресійну модель

$$Y = 10^{\varepsilon + \hat{b}_0} X_1^{\hat{b}_1} X_2^{\hat{b}_2} X_3^{\hat{b}_3}. \quad (4)$$

Співвідношення (4) і є нелінійною регресійною моделлю для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony.

Для моделі було визначено оцінки множинного коефіцієнту детермінації R^2 , середньої величини відносної похибки $MMRE$ та відсотку прогнозованих результатів $PRED(0,25)$, що дорівнюють 0,900, 0,187 та 0,756 відповідно. Оцінки показують належну якість отриманої моделі. Нагадаємо, що допустимі значення $MMRE$ і $PRED(0,25)$ складають не більше 0,25 і не менше 0,75 відповідно, а значення R^2 більше 0,75 [16, 17].

Для проведення порівняння за (1) побудуємо трифакторну лінійну регресійну модель для оцінювання розміру веб-застосунків, що створюються з використанням фреймворку Symfony, для вихідних даних. Модель має наступний вигляд

$$Y = \hat{Y} + \varepsilon = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + \hat{b}_3 X_3 + \varepsilon.$$

Для побудованої моделі оцінки параметрів $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3$ дорівнюють -1,3253, 0,0590, 0,5311 та -0,3588 відповідно. Залишки лінійної регресії для емпіричних даних ε були перевірені на нормальність за допомогою критерію Пірсона. Відповідна нульова гіпотеза була прийнята на рівні значимості 0,05.

Прийняття нульової гіпотези щодо нормальності розподілу залишків регресії ε може слугувати теоретичним підґрунтям використання лінійного регресійного аналізу для оцінювання розміру веб-застосунків, що створюються з використанням PHP фреймворку Symfony, але розраховані оцінки середньої величини відносної похибки $MMRE$ та відсотку прогнозованих результатів $PRED(0,25)$, що дорівнюють 0,273 та 0,578, свідчать про незадовільну якість моделі.

Отримані оцінки лінійної та нелінійної регресійних моделей було порівняно, за результатами порівняння видно покращення якісних показників моделі у випадку використання нелінійного регресійного аналізу. Так R^2 , $MMRE$ та $PRED(0,25)$ для лінійної моделі дорівнюють 0,899, 0,273 та 0,578 відповідно, а ці ж оцінки для нелінійної моделі – 0,900, 0,187 та 0,756. Бачимо покращення суттєве покращення показників оцінок $MMRE$ та $PRED(0,25)$.

Висновки з даного дослідження та перспективи подальших розвідок у даному напрямі

В роботі було вирішено важливу задачу щодо підвищення достовірності оцінювання розміру веб-застосунків, що створюються з використанням PHP фреймворку Symfony, за рахунок побудови нелінійної регресійної моделі з використанням нормалізуючого перетворення на основі десяткового логарифму. Порівняння отриманої нелінійної регресійної моделі з лінійною моделлю показує значне покращення якості моделі за показниками $MMRE$, $PRED(0,25)$.

В майбутньому планується побудувати довірчі інтервали та інтервали передбачення нелінійної регресії. Також планується виконати побудову нелінійних регресійних моделей з використання різних нормалізуючих перетворень з метою підбору такого, що дозволить отримати найкращі показники якості моделі.

References

1. Arcidiacono G. Comparative research about high failure rate of IT projects. PM World Journal. 2017. Vol. VI, Issue II.
2. Usage statistics of PHP for websites. URL: <https://w3techs.com/technologies/details/pl-php> (20.10.2022).
3. The Best PHP Framework for 2015: SitePoint Survey Results. URL: <https://www.sitepoint.com/best-php-framework-2015-sitepoint-survey-results/> (20.10.2022).
4. PHP developer ecosystem survey. URL: <https://www.jetbrains.com/lp/devecosystem-2021/php/> (20.10.2022).
5. Prykhodko S. B. Nonlinear regression model for estimating the size of web-applications created using the Laravel framework. Information technology and computer engineering. 2021. Vol. 50, no. 1. P. 115–121.
6. 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. 2022. No. 4. P. 129–139.
7. Prykhodko S., Shutko Ivan, Prykhodko Andrii 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: <https://doi.org/10.37394/23205.2021.20.35>.
8. Hee Beng Kuan Tan, Yuan Zhao, Hongyu Zhang. Conceptual data model-based software size estimation for information systems. ACM Transactions on Software Engineering and Methodology. 2009. Vol. 19, no. 2. P. 1–37. DOI: <https://doi.org/10.1145/1571629.1571630>.
9. Hee Beng Kuan Tan, Yuan Zhao, Hongyu Zhang Estimating LOC for information systems from their conceptual data models. Proceeding of the 28th international conference, Shanghai, China, 20–28 May 2006. New York, New York, USA, 2006. DOI: <https://doi.org/10.1145/1134285.1134331>.

10. Prykhodko S.B. Constructing the non-linear regression equation to estimate the software size of open source PHP-based information systems. *Problems of Information Technologies*. 2018. № 1 (023). P. 118–125.
11. Prykhodko S. B. Three-factor non-linear regression model to estimate the size of open source php-based applications. *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*. 2020. Vol. 1, no. 1. P. 124–131. DOI: <https://doi.org/10.32838/2663-5941/2020.1-1/23>.
12. Prykhodko N. V., Prykhodko S. B. Constructing the Nonlinear Regression Models on the Basis of Multivariate Normalizing Transformations. *Elektronnoe modelirovanie*. 2018. Vol. 40, no. 6. P. 101–110. DOI: <https://doi.org/10.15407/emodel.40.06.101>.
13. Prykhodko Sergiy. Detecting outliers in multivariate non-Gaussian data on the basis of normalizing transformations. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 29 May – 2 June 2017. [S. l.], 2017. DOI: <https://doi.org/10.1109/ukrcon.2017.8100366>.
14. Prykhodko Sergiy. Application of the squared mahalanobis distance for detecting outliers in multivariate non-Gaussian data. 2018 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 20–24 February 2018. [S. l.], 2018. DOI: <https://doi.org/10.1109/tcset.2018.8336353>.
15. Handbook of regression analysis. *Choice Reviews Online*. 2013. Vol. 51, no. 02. P. 51–0931–51–0931. DOI: <https://doi.org/10.5860/choice.51-0931>.
16. Foss T. A simulation study of the model evaluation criterion mmre. *IEEE Transactions on Software Engineering*. 2003. Vol. 29, no. 11. P. 985–995. DOI: <https://doi.org/10.1109/tse.2003.1245300>.
17. Prykhodko S., Natalia Prykhodko. Mathematical Modeling of Non-Gaussian Dependent Random Variables by Nonlinear Regression Models Based on the Multivariate Normalizing Transformations. *Advances in Intelligent Systems and Computing*. Cham, 2020. P. 166–174. DOI: https://doi.org/10.1007/978-3-030-58124-4_16.

За зміст повідомлень редакція відповідальності не несе

Повні вимоги до оформлення рукопису
http://journals.khnu.km.ua/vestnik/?page_id=37

**Рекомендовано до друку рішенням вченої ради Хмельницького національного університету,
протокол № 6 від 29.12.2022 р.**

Підп. до друку 29.12.2022 р. Ум.друк.арк. 27,38 Обл.-вид.арк. 26,05
Формат 30x42/4, папір офсетний. Друк різнографією.
Наклад 100, зам. № _____

Тиражування здійснено з оригінал-макету, виготовленого
редакцією журналу “Вісник Хмельницького національного університету”
редакційно-видавничим центром Хмельницького національного університету
29016, м. Хмельницький, вул. Інститутська, 7/1. тел (0382) 72-83-63