УДК 004.412:519.237.5

ПОПЕРЕДНЯ ОБРОБКА ДАНИХ ДЛЯ ПОБУДОВИ МОДЕЛІ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ НА ПРОГРАМНІЙ ПЛАТФОРМІ .NET

Урсов А.В.[1]; Макарова Л.М., к.т.н, доц.[2]
[1,2] Національний університет кораблебудування імені адмірала Макарова
[1,2] Україна, Миколаїв
[1] ursov00125@gmail.com; [2] lidiia.makarova@nuos.edu.ua

Анотація. *Розглянуто проблему оцінювання розміру веб-застосунків, що створюються мовою C# з використанням .NET Framework. Для вирішення проблеми було розглянуто емпіричні дані на основі 50 проектів з ресурсу GitHub та виконана їх попередня обробка.*
Ключові слова: *попередня обробка даних; розмір програмного забезпечення; оцінювання розміру; .NET, веб-застосунок.*

Вступна частина. На сьогодні, оцінювання трудомісткості різного програмного забезпечення, включаючи веб-застосунки на C# з використанням програмної платформи .NET, є однією з найважливіших задач програмної інженерії. Оцінка трудомісткості дозволяє забезпечити успішну реалізацію програмного забезпечення.

Під час оцінювання трудомісткості веб-застосунків часто використовують методи нелінійного регресійного аналізу та нелінійні регресійні моделі, такі як COCOMO II [1], де одним із головних факторів є розмір програми. Однак на ранніх стадіях розробки часто немає точної інформації про розмір програми, і потрібно провести її оцінювання. Звичайно, оцінювання розміру програми недостатньо для прийняття рішення про її трудомісткість, але оцінювання розмірів розроблених програм на платформі .NET є доцільною задачею для розробників, клієнтів і менеджерів проектів.

На основі вищевикладеного, побудова нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються на програмній платформі .NET є актуальною, а розробка програми для її реалізації має практичне призначення. Першим кроком для побудови моделі є попередня обробка зібраних даних.

Метою роботи є підвищення достовірності оцінювання розміру веб-застосунків, що створюються на програмній платформі .NET, за рахунок видалення викидів даних та побудови нелінійної регресійної моделі.

Основна частина. Для досягнення поставленої мети було проаналізовано моделі, які вже побудовані та знаходяться у відкритому доступі, зібрано емпіричні дані з онлайн ресурсу GitHub.

Для проведення аналізу розробки програм на платформі .NET, було взято 50 проектів, які були розміщені на спільному веб-сервісі для розробки програмного забезпечення, такому як GitHub. Ці проекти були завантажені та проаналізовані за допомогою програмного забезпечення SourceMonitor [2] для збору таких метрик, як кількість класів, середня кількість методів на клас, кількость строк коду в тисячах. Отримані дані були нормалізовані та проаналізовані.

На підставі критерію згоди Пірсона, було визначено, що розподіл даних не відповідає нормальному. Такі відомі моделі як COCOMO, ISBSG, COCOMO II були створені саме із

використанням нормалізуючого перетворення у вигляді десяткового логарифму, тому для нормалізації було використано метод десяткового логарифму [3]. Застосування нормалізуючих перетворень дозволяє перейти до лінійної регресії для нормалізованих даних, для неї побудувати довірчий інтервал з використанням t-розподілу Стьюдента, а потім шляхом застосування відповідного перетворення перейти до нелінійної регресії і її довірчого інтервалу

Для виявлення викидів був використаних квадрат відстані Махаланобіса [4], за допомогою якого можна виділити значущі відхилення у даних. Було визначено тестову статистику та критичне значення Фішера. Ті значення з вибірки, для яких тестова статистика перевищує критичне значення Фішера, є викидами. Під час першої ітерації виявлення викидів було виявлено 1 викид, а після другої ітерації було визначено 5 викидів, на третій ітерації викидів не виявлено, це означає, що для побудови нелінійної регресійної моделі ми маємо 44 набори даних.

Планується побудувати нелінійну регресійну модель для оцінювання розміру веб-застосунків, що створюються на програмній платформі .NET, параметри якої визначаються за допомогою методу найменших квадратів, провести порівняльний аналіз якості моделі, розробити програмне забезпечення для реалізації нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються на програмній платформі .NET.

Висновки. Після аналізу існуючих моделей для оцінювання розміру програмного забезпечення, можна прийти до висновку, що необхідно побудувати нелінійну регресійну модель для оцінки розміру веб-застосунків, які розробляються на платформі .NET. Було зібрано дані для цієї моделі і була проведена їх попередня обробка. Планується створення нелінійної регресійної моделі для оцінювання розміру таких веб-застосунків. Після цього буде оцінена якість цієї моделі і розроблене програмне забезпечення для її впровадження**.**

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Trendowicz, A., & Jeffery, R. (2014). Software project effort estimation. foundations and best practice guidelines for success. *Springer International Publishing*, 491.

[2] SourceMonitor (n.d.). Retrieved October 10, 2023.

[3] Приходько, Н.В., & Приходько, С.Б. (2018). Нелінійна регресійна модель для оцінювання розміру програмного забезпечення інформаційних систем на базі VB. *Інформаційні технології та комп'ютерна інженерія*, 3, 37-42.

[4] Prykhodko, S., Prykhodko, N., Makarova, L., Pukhalevych, A. (2018). Application of the Squared Mahalanobis Distance for Detecting Outliers in Multivariate NonGaussian Data. *Proceedings of 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, February 20–24, 962- 965.

Ursov A.V., Makarova L.M.

Pre-processing of data for building a model for estimating the size of web applications created on the .NET framework

Abstract. *Considered the problem of estimating the size of web applications created in C# using the .NET Framework. To solve the problem, empirical data based on 50 projects from the GitHub resource was considered and pre-processed.*

Keywords: *data pre-processing; software size; size estimation; .NET; web application.*