УДК 004.412

БАГАТОФАКТОРНА НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ НА ОСНОВІ ФРЕЙМВОРКУ SVELTE

Дромашко В.В.[1]; Фаріонова Т.А., к.т.н., проф.[2]
[1,2] Національний університет кораблебудування імені адмірала Макарова
[1,2] Україна, Миколаїв
[1]181271@nuos.edu.ua; [2] tetyana.farionova@nuos.edu.ua

Анотація. *Робота присвячена побудові багатофакторної нелінійної регресійної моделі для оцінювання кількості строк коду веб-застосунків, створених на основі фреймворку Svelte за допомогою перетворення у вигляді десяткового логарифму в залежності від метрик кількості компонентів, кількості функцій та кількості використання компонентів у коді.*
Ключові слова: *нелінійна регресійна модель; фреймворк Svelte; оцінювання розміру веб-застосунків; нормалізуюче перетворення.*

Вступна частина. Веб-розробка є однією з найбільш динамічних та швидкозмінних галузей сучасного інформаційного сектору. Фреймворк Svelte набув популярності в галузі веб-розробки завдяки своєму унікальному підходу до створення веб-застосунків. На відміну від традиційних фреймворків, які базуються на віртуальному DOM, Svelte компілює компоненти на етапі збірки, що призводить до менших розмірів вихідного файлу та покращує продуктивність під час їх виконання у браузері. Крім того, Svelte надає зручні інструменти для створення веб-застосунків і має простий синтаксис, що дозволяє легко освоїти цей фреймворк.

У сучасному цифровому світі підприємства все більш використовують веб-застосунки для спілкування з клієнтами, ведення бізнесу та поширення інформації. Оцінка та прогнозування розміру веб-застосунків стає ключовим завданням для розробників та бізнес-аналітиків та має велике значення для правильного розподілу трудових ресурсів, планування проектів та оцінки витрат. Точні прогнози розміру допомагають оптимізувати потужність серверів, забезпечити ефективне розгортання та якісно обслуговувати потреби користувачів. Для прогнозування трудомісткості на розробку веб-застосунку зазвичай застосовують такі математичні моделі, як COCOMO, COCOMO II, ISBSG, де ключовим фактором є розмір програми. Таким чином, для визначення трудомісткості веб-застосунку необхідно мати інформацію про розмір програмного коду.

Авторами проведено аналіз існуючих моделей, які використовуються на сьогоднішній день для оцінювання розміру веб-застосунків на основі емпіричних даних з метрик програмного коду. Таких як, трьох-факторна нелінійна регресійна модель для визначення розміру веб-застосунків, розроблених на мові програмування PHP [1], трьох-факторна нелінійна регресійна модель для оцінювання розміру веб-застосунків з використанням фреймворку Laravel [2] та нелінійна регресійна модель для оцінювання розміру застосунків, написаних на мові програмування Java [3] та інші. Однак, слід зазначити, що якість моделі суттєво залежить від таких аспектів, як обрана мова програмування, використаний фреймворк, технологічні рішення та метрики програмного забезпечення. Різноманітність цих елементів впливає на ефективність та точність моделі при оцінюванні розміру веб-застосунку.

Проведений аналіз літературних джерел показав, відсутність алгоритмічних моделей для оцінювання розміру веб-застосунків, створених на основі фреймворку Svelte, тому розробка відповідної регресійної моделі є актуальним завданням.

Метою дослідження є побудова багатофакторної нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються на основі фреймворку Svelte, в залежності від метрик кількості компонентів, кількості функцій та кількості використання компонентів у коді.

Основна частина. Розглянемо побудову нелінійної регресійної моделі для оцінювання кількості строк коду веб-застосунків, що створюються на основі фреймворку Svelte, в залежності від кількості компонентів X_1, кількості функцій X_2 та кількості використання компонентів у коді X_3. Для цього було зібрано дані з 50 веб-застосунків, які розміщені на веб-сервісі Github (https://github.com). Дані були перевірені на наявність мультиколінеарності за допомогою коефіцієнтів впливу дисперсії (VIFs), який показав відсутність мультиколінеарності між факторами. Наступним кроком авторами проведено перевірку чотирьохвимірних даних на негаусівський розподіл за допомогою тесту Мардіа (Mardia's test), що ґрунтується на значенні багатовимірного ексцесу та асиметрії. Після виконання перевірки, зроблено висновок про негаусівську природу даних. Для подальшої побудови проведено розрахунок відстані квадрату Махаланобіса для визначення викидів серед отриманих метрик веб-застосунків. Після перевірки даних на викиди, було виявлено 3 програмних проєкта, у яких квадрат відстані Махаланобіса був більший за 14.86.

Спочатку авторами була побудована багатофакторна лінійна регресійна модель у вигляді

$$Y = \widehat{b_0} + \widehat{b_1}X_1 + \widehat{b_2}X_2 + \widehat{b_3}X_3 + \varepsilon, \quad (1)$$

де $\widehat{b_0}$, $\widehat{b_1}$, $\widehat{b_2}$, $\widehat{b_3}$ – це оцінки параметрів, ε – випадкова величина з розподілом Гауса.

Нульова гіпотеза про нормальність розподілу моделі (1) була перевірена за критерієм Пірсона. Ця гіпотеза була відкинута для рівня значимості 0,05. Розподіл ε є негаусівським, що вказує на відсутність теоретичного обґрунтування для використання лінійної регресійної моделі. Таким чином, необхідно розробити нелінійну регресійну модель для оцінювання кількості строк коду веб-застосунків, що створюються на основі фреймворку Svelte. В якості нормалізуючого перетворення вихідних даних авторами використано одновимірне нормалізуюче перетворення у вигляді десяткового логарифму,

Далі, за нормалізованими даними була побудована лінійна регресійна модель

$$Z_y = \widehat{b_0} + \widehat{b_1}Z_1 + \widehat{b_2}Z_2 + \widehat{b_3}Z_3 + \varepsilon, \quad (2)$$

де параметри моделі (2) оцінювалися за методом найменших квадратів та їх оцінки є такими: $\widehat{b_0}$ = 1.575, $\widehat{b_1}$ = 0.972, $\widehat{b_2}$ = 0.713, $\widehat{b_3}$ = -0.358.

Наступним кроком, за (2) та перетворенням у вигляді десяткового логарифму, будуємо нелінійну регресійну модель, яка має такий вигляд $Y = 10^{\varepsilon+\widehat{b_0}}X_1^{\widehat{b_1}}X_2^{\widehat{b_2}}X_3^{\widehat{b_3}}$, де $\widehat{b_0}$, $\widehat{b_1}$, $\widehat{b_2}$, $\widehat{b_3}$ – це оцінки параметрів, ε – це гаусівська випадкова величина з нульовим випадковим математичним сподіванням.

Отримана багатофакторна нелінійна регресійна модель була перевірена на ефективність за допомогою кількох ключових показників. По перше, множинний коефіцієнт детермінації R^2, що визначає ступінь відповідності моделі реальним даним. У нашому випадку, значення R^2 = 0.837, що свідчить про достатньо високу точність та адекватність регресійної моделі. Другим важливим показником є середня величина відносної помилки MMRE = 0.047. Такий показник дозволяє зробити висновок про те, що модель демонструє низький рівень помилок

у прогнозуванні результатів . Отримана величина MMRE підтверджує високу точність моделі. Додатково, було розраховано відсоток прогнозованих результатів, для яких величина відносної помилки MRE менша за 0.25, позначена як PRED(0.25). В нашому випадку, ми отримали PRED(0.25) = 0.776, що вказує на те, що понад 77.6% прогнозів відповідають високим стандартам точності. Таким чином можна зробити висновок про високу якість розробленої нелінійної регресійної моделі для оцінювання кількості строк коду веб-застосунків, що створюються на основі фреймворку Svelte

Висновки. В роботі на основі десяткового логарифму було побудовано багатофакторну нелінійну регресійну модель для оцінювання кількості строк коду веб-застосунків, що створенні на основі фреймворку Svelte, в залежності від метрик кількості компонентів, кількості функцій та кількості використання компонентів у коді, яка має задовільну якість відповідного оцінювання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Приходько, С.Б., & Ворона, М. В. (2021). Оцінювання розміру PHP-застосунків з відкритим кодом за нелінійними регресійними моделями з різними факторами. *Збірник наукових праць НУК: Комп'ютерні науки та інформаційні технології №1 2021*. Миколаїв: НУК імені адмірала Макарова.

[2] Приходько, С.Б., Приходько, Н.В., Ворона, М.В., & Бєловол, І.О. (2021). Нелінійна регресійна модель для оцінювання розміру web-застосунків, що створюються з використанням фреймворку Laravel, *ІТКІ*, вип. 50, вип. 1, с. 115–121.

[3] Prykhodko, N. V., & Prykhodko, S. B. (2018). The non-linear regression model to estimate the software size of open source java-based systems. *Радіоелектроніка, інформатика, управління.* - № 3. - С. 158-166.

Dromashko V. V., Farionova T.A.

Multifactor nonlinear regression model for assessing the size of web applications created using the Svelte framework

Abstract. *The work is dedicated to building a multifactor nonlinear regression model for estimating the number of lines of code in web applications created using the Svelte framework. This is achieved by transforming the data into decimal logarithmic form based on metrics such as the number of components, the number of functions, and the frequency of component usage in the code.*

Keywords: *nonlinear regression model; Svelte framework; assesing size of web applications; normalizing transformation.*