

НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ СЕРВЕРНИХ ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ З ВИКОРИСТАННЯМ ФРЕЙМВОРКУ NESTJS

Латанська Л.О., к.ф.-м.н., доцент¹; Бринзей О.А.²

^{1,2} Національний університет кораблебудування імені адмірала Макарова

^{1,2} Україна, Миколаїв

¹latanskaya@gmail.com; ²oleksandrbrynzei@gmail.com

Анотація. В роботі було розроблено нелінійну регресійну модель для оцінювання розміру серверних застосунків, що створюються з використанням фреймворку NestJS.

Ключові слова: оцінка розміру програмного забезпечення; серверні застосунки; нелінійна регресійна модель.

Вступна частина. Прогнозування трудомісткості розробки програмного забезпечення є досить важливим завданням на ранніх стадіях розробки програмного забезпечення будь-якого типу, в тому числі серверних застосунків. Прогнозування трудомісткості можна проводити за допомогою таких математичних моделей, як наприклад СОСОМО II. Але в цій та інших моделях, які призначені для прогнозування трудомісткості розробки програмного забезпечення, одним із головних параметрів виступає розмір програмного забезпечення. Тому для оцінювання трудомісткості треба спочатку оцінити розмір програмного забезпечення.

На даний момент широко застосовується оточення NodeJS, з використанням якого розробляють різного роду сервери та сервіси. Для пришвидшення розробки таких застосунків використовують фреймворки, до яких відносяться: SailsJS, MeteorJS та NestJS.

Під час ознайомлення з останніми публікаціями по даній тематиці, було виявлено трьохфакторне лінійне регресійне рівняння [1], яке призначене для оцінювання кількості рядків коду програмного забезпечення, розробленого на мові PHP. Крім цього було знайдено трьохфакторну нелінійну регресійну модель [2] для оцінювання розміру PHP-застосунків з відкритим вихідним кодом, а також знайдено трьохфакторну нелінійну регресійну модель [3], яка призначена для оцінювання розміру web-застосунків, що створюються з використанням фреймворку Laravel. Із аналізу знайдених моделей можна зробити висновок, що на якість моделі впливає мова програмування, фреймворк, технології та інші фактори.

Так як не було знайдено моделі для оцінювання розміру серверних застосунків, що створюються з використанням фреймворку NestJS, то задача побудови спеціалізованої моделі для оцінювання кількості рядків коду серверних застосунків, що створюються з використанням фреймворку NestJS, є актуальною.

Метою дослідження є створення нелінійної регресійної моделі для оцінювання розміру серверних застосунків, що створюються з використанням фреймворку NestJS.

Основна частина. Розглянемо побудову нелінійної регресійної моделі для оцінювання розміру серверних застосунків, що створюються з використанням фреймворку NestJS. В якості емпіричних даних було знайдено 62 серверні застосунки, які знаходяться у відкритому доступі на сервісі GitHub (<https://github.com>). Всі вони були розроблені з використанням фреймворку NestJS, який створений і повністю підтримує TypeScript. Збір метрик проводився за допомогою інструменту SonarQube (<https://www.sonarqube.org/>). Всього було

зібрано п'ять метрик: загальна кількість рядків, кількість рядків коду, кількість файлів, кількість функцій та кількість класів. Для побудови моделі були відібрані наступні метрики: кількість рядків коду Y ; кількість функцій X_1 ; кількість класів X_2 . Вказані метрики були обрані у зв'язку з тим, що між ними відсутня мультиколінеарність. Після цього дані були перевірені на нормальність розподілу за допомогою критерія Пірсона. Тест показав, що розподіл даних не є гаусівським. Для визначення наявності багатовимірних викидів у даних, було використано тест на основі квадрату відстані Махаланобіса.

Використовуючи зібрані дані із зазначеними вище метриками, була побудована лінійна регресійна модель $Y = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + \varepsilon$, де $\hat{b}_0$, $\hat{b}_1$ та $\hat{b}_2$ є оцінками параметрів. Причому змінна ε повинна бути випадковою величиною з нормальним розподілом. Нульова гіпотеза про нормальність розподілу ε була перевірена за критерієм Пірсона. Висунута гіпотеза була відкинута для рівня значимості 0,05. Це свідчить про те, що розподіл ε не є гаусівським та вказує на відсутність теоретичного обґрунтування застосування лінійної регресійної моделі. Тому слід побудувати нелінійну регресійну модель.

Нелінійна регресійна модель має вигляд $Y = 10^{\varepsilon + \hat{b}_0} X_1^{\hat{b}_1} X_2^{\hat{b}_2}$, де ε – гаусівська випадкова величина, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, $\hat{b}_0$, $\hat{b}_1$ та $\hat{b}_2$ є оцінками параметрів, які мають відповідно такі значення: 1,0021, 0,8845, 0,1871 та знайдені за допомогою методу найменших квадратів.

Створена нелінійна регресійна модель була перевірена за допомогою R^2 – множинного коефіцієнта детермінації, MMRE – середньої величини відносної похибки та PRED(0,25) – відсотку прогнозованих результатів, для яких величини відносної похибки MMRE менші за 0,25. Допустимі значення мають складати не більше 0,25 для MMRE, та не менше 0,75 для PRED(0,25), а значення R^2 якомога ближче до 1. Для створеної моделі вказані умови виконуються і це свідчить про високу якість отриманої моделі.

Висновки. В роботі було побудовано нелінійну регресійну модель з двома предикторами на основі десятичного логарифму для оцінювання розміру серверних застосунків, що створюються з використанням фреймворку NestJS. Створена модель має добрі показники якості.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Tan, H.B.K., Zhao, Y., & Zhang, H. (2006). *Estimating LOC for information systems from their conceptual data models*, Software Engineering : the 28th International Conference (ICSE '06). Shanghai, China.
- [2] Приходько, С.Б., Приходько, Н.В., Фаріонова, Т.А., & Ворона, М.В. (2020). Трьохфакторна нелінійна регресійна модель для оцінювання розміру Php-застосунків з відкритим кодом. *Науковий журнал «Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки»*, 31 (70), 124-131. DOI: 10.32838/2663-5941/2020.1- 1/23
- [3] Приходько, С.Б., Приходько, Н.В., Ворона, М.В., & Беловол, І.О. (2021). Нелінійна регресійна модель для оцінювання розміру web-застосунків, що створюються з використанням фреймворку Laravel. *Інформаційні технології та комп'ютерна інженерія*, 2021.1.

Latanska L.O., Brynzei O.A.

Nonlinear regression model for estimating the size of server applications created using the NestJS framework

Abstract. In the work, a nonlinear regression model was built for estimating the size of server applications created using the NestJS framework.

Keywords: software size estimation; server applications; nonlinear regression model.