УДК 004.412:519.237.5

НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ ТРИВАЛОСТІ РОЗРОБКИ ЗАСТОСУНКІВ НА МОВІ C# ДЛЯ ПЛАТФОРМИ PC

Пухалевич А.В., к.т.н., доц.[1]; Кольцов Є.В.[2]
[1,2] Національний університет кораблебудування імені адмірала Макарова
[1,2] Україна, Миколаїв
[1] andrii.pukhalevych@nuos.edu.ua; [2] 181267@nuos.edu.ua

Анотація. *В роботі, використовуючи перетворення у вигляді десяткового логарифму, побудована нелінійна регресійна модель для оцінювання тривалості розробки застосунків на мові C# для платформи PC в залежності від їх трудомісткості. При побудові застосовувались емпіричні дані репозитарію ISBSG.*
Ключові слова: *оцінювання тривалості розробки; нелінійна регресійна модель; застосунок C#; платформа PC.*

Вступна частина. На даний час при створенні програмних застосунків для платформи PC дуже часто використовується мова C#, тому перед розробниками таких застосунків постає необхідність вирішення задачі програмної інженерії, яка полягає в оцінюванні тривалості відповідної розробки. Важливість достовірного оцінювання тривалості пояснюється тим, що від нього значною мірою залежить чи погодиться замовник обрати розробника, чи обере іншого розробника, який запропонує коротші терміни розробки. В той же час, занадто оптимістична оцінка тривалості може призвести до порушення термінів розробки програмного забезпечення і штрафних санкцій до розробника. Отже, побудова моделі для оцінювання тривалості розробки застосунків на мові C# для платформи PC є актуальною задачею.

Для оцінювання тривалості розробки програмних застосунків найчастіше використовується залежність тривалості від трудомісткості розробки. Для цього, застосовуючи метод побудови регресійних моделей з використанням нормалізуючих перетворень [1], було побудовано нелінійні регресійні моделі, такі як [2, 3] та інші. Проте, існуючі моделі будувались для оцінювання тривалості розробки програмного забезпечення на інших мовах програмування, або ж, при побудові, взагалі враховувалась тільки платформа, для якої створюється програмне забезпечення, що може знижувати достовірність оцінювання тривалості розробки застосунків на мові C# для платформи PC.

Метою дослідження є побудова нелінійної регресійної моделі для оцінювання тривалості розробки застосунків на мові C# для платформи PC.

Основна частина. Емпіричними даними, що використовувалися при побудові нелінійної регресійної моделі для оцінювання тривалості розробки застосунків на мові C# для платформи PC в залежності від трудомісткості розробки є дані репозитарію ISBSG з релізу 2021 року [4]. Для мови C# кількість емпіричних значень була 38.

Ці дані було перевірено на нормальність закону розподілу. Було визначено, що розподіл даних не є гаусівським, і тому для побудови регресійної моделі необхідно здійснити нормалізацію даних [1]. Нормалізація емпіричних даних виконувалася перетворенням у вигляді десяткового логарифму, що дозволило спочатку побудувати лінійну регресійну

модель для нормалізованих значень. Побудова лінійної моделі включала процес видалення викидів, який відбувався в декілька етапів.

Першим етапом після нормалізації була перевірка нормальності закону розподілу перетворених даних. Якщо закон не є нормальним, то один рядок даних, що має найбільше значення квадрату відстані Махаланобіса відкидався і перевірка закону розподілу повторювалась.

Далі, для визначення наявності викидів, було використано тест на основі квадрату відстані Махаланобіса для нормалізованих даних. Якщо для якогось рядка даних ця відстань перевищувала критичне значення, то один рядок даних з найбільшим значенням квадрату відстані Махаланобіса відкидався, допоки всі рядки мали відстань меншу ніж критична.

Наступним етапом було отримання оцінок параметрів лінійної регресійної моделі і перевірка нормальності закону розподілу випадкової похибки моделі за критерієм Пірсона.

Лінійна регресійна модель для нормалізованих значень може бути представлена як $lg(Y) = \hat{b}_0 + \hat{b}_1 lg(X) + \varepsilon$, де $\hat{b}_0$, $\hat{b}_1$ є оцінками параметрів, що отримані методом найменших квадратів, а ε – випадкова величина.

Якщо після знаходження оцінок параметрів $\hat{b}_0$ і $\hat{b}_1$ закон розподілу випадкової величини ε не був нормальним, то рядок з найбільшим за модулем значенням величини ε відкидався, і побудова моделі починалась з першого етапу.

Останнім етапом видалення викидів була перевірка того, що значення залежної змінної не виходить за інтервал передбачення лінійної моделі. Якщо одне або декілька значень виходило за цей інтервал, вони відкидались, і побудова моделі починалась з першого етапу.

В результаті, після побудови лінійної регресійної моделі за емпіричними даними репозитарію ISBSG і видалення викидів, було видалено 15 рядків даних.

Для емпіричних даних, що залишились, було отримано наступні значення оцінок параметрів лінійної регресії $\hat{b}_0 = 0{,}9920$, $\hat{b}_1 = 0{,}0465$.

Нелінійна регресійна модель для оцінювання тривалості розробки застосунків на мові C# для платформи PC є такою:

$$Y = 10^{\hat{b}_0 + \hat{b}_1 lg(X) + \varepsilon}.$$

Після перетворення, модель є такою:

$$Y = 10^{\hat{b}_0 + \varepsilon} X^{\hat{b}_1},$$

де X – це значення трудомісткості (в людино-годинах), Y – це тривалість розробки (в місяцях), $\hat{b}_0 = 0{,}9920$, $\hat{b}_1 = 0{,}0465$.

Отримані значення PRED(0,25), MMRE і R^2 для побудованої моделі відповідно дорівнюють 0,78, 0,11 і 0,3.

Висновки. В результаті роботи було побудовано модель для оцінювання тривалості розробки застосунків на мові C# для платформи PC.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Приходько, С.Б. (2012). Метод побудови нелінійних рівнянь регресії на основі нормалізуючих перетворень. *Тези доповідей міждерж. наук.-методич. конф. "Проблеми математичного моделювання"* (с. 31-33). 13-15 червня 2012, Дніпродзержинськ, Дніпродзержинськ : ДДТУ.

[2] Oligny, S., Bourque, P., Abran, A. & Fournier, B. (2000). Exploring the relation between effort and duration in software engineering projects. *Proc. of the 16th IFIP World Computer Congress* (pp. 175-178). Aug, 2000, Beijing, China

[3] Приходько, С.Б., Пухалевич, А.В., Халанчук, І.П. (2021). Нелінійні регресійні моделі для оцінювання тривалості розробки програмних застосунків, що створюються мовами Java та Visual Basic для ПК. *Інформаційні технології: моделі, алгоритми, системи (ITMAS – 2021): Матеріали II Всеукраїнської науково-практичної інтернет конференції (26-28 жовтня 2021 р.).* Миколаїв: НУК. С. 6-7. https://itconf.nuos.edu.ua/2021/proceedings/

[4] *International Software Benchmarking Standards Group*. (n.d.). https://www.isbsg.org/.

Pukhalevych A.V, Koltsov E.V.

A NONLINEAR REGRESSION MODEL FOR ESTIMATING THE DURATION OF C# APPLICATION DEVELOPMENT FOR THE PC PLATFORM

Abstract. *In the work, a three-factor nonlinear regression model was built for estimating the size of the code of online stores in PHP using a transformation from a decimal logarithm. According to* R^2*, PRED, MMRE, the built model is of high quality.*

Keywords: *estimation of development duration; non-linear regression model; C# application; PC platform.*