

АНАЛІЗ ВИКОРИСТАННЯ РЕГРЕСІЙНИХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗУВАННЯ ЯКОСТІ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Вінницький національний технічний університет

Анотація

У роботі проаналізовано використання регресійних моделей для прогнозування якості програмного забезпечення та визначено основні напрями застосування статистичних методів у програмній інженерії. Обґрунтовано роль регресійного аналізу у прогнозуванні дефектів, оцінюванні продуктивності та забезпеченні надійності програмних систем. Визначено переваги використання математичних моделей у процесі розробки та тестування програмного забезпечення.

Ключові слова: регресійний аналіз, програмне забезпечення, прогнозування, якість програмного забезпечення, програмна інженерія, статистичні методи.

Abstract

The paper analyzes the use of regression models for software quality prediction and identifies the main directions of applying statistical methods in software engineering. The role of regression analysis in defect prediction, performance evaluation, and ensuring software system reliability is substantiated. The advantages of using mathematical models in the process of software development and testing are determined.

Keywords: regression analysis, software, prediction, software quality, software engineering, statistical methods.

Вступ

Якість програмного забезпечення є одним із ключових показників успішності програмного продукту. Наявність дефектів, низька продуктивність або нестабільна робота системи можуть призводити до додаткових витрат на тестування та підтримку програмного коду. Тому актуальним завданням програмної інженерії є використання методів, які дозволяють оцінювати якість програмного забезпечення ще на етапах його розробки [1].

Одним із таких методів є регресійний аналіз. Регресійні моделі застосовуються для виявлення залежностей між характеристиками програмного забезпечення та показниками його якості. Отримані результати можуть використовуватися для прогнозування кількості дефектів, оцінювання продуктивності програмних систем і підтримки процесу прийняття рішень під час розробки програмного забезпечення [2].

Результати дослідження

Регресійні моделі використовуються для аналізу статистичних даних та прогнозування значень певних показників на основі наявних спостережень. У програмній інженерії вони можуть застосовуватися для оцінювання кількості програмних дефектів, складності програмних модулів, часу виконання тестування та інших характеристик програмного забезпечення [3].

Найпростішою формою регресійного аналізу є лінійна регресія, яка описується рівнянням:

$$y = a + bx \quad (1)$$

де y – прогнозоване значення показника якості програмного забезпечення, a – вільний член рівняння, b – коефіцієнт регресії, x – незалежна змінна.

Для демонстрації можливостей регресійного аналізу було використано відкритий набір даних NASA Software Defect Prediction Dataset, який містить метрики програмних модулів та інформацію про характеристики програмного коду. У дослідженні було проаналізовано залежність між загальною кількістю рядків коду (LOC_TOTAL) та метрикою HALSTEAD_VOLUME, що використовується для оцінювання складності програмного забезпечення.

Реалізацію виконано мовою Python із використанням бібліотеки Scikit-learn. На основі наявних даних побудовано модель лінійної регресії, яка зображена на рисунку 1, та досліджено залежність між розміром програмного модуля і наявністю дефектів.

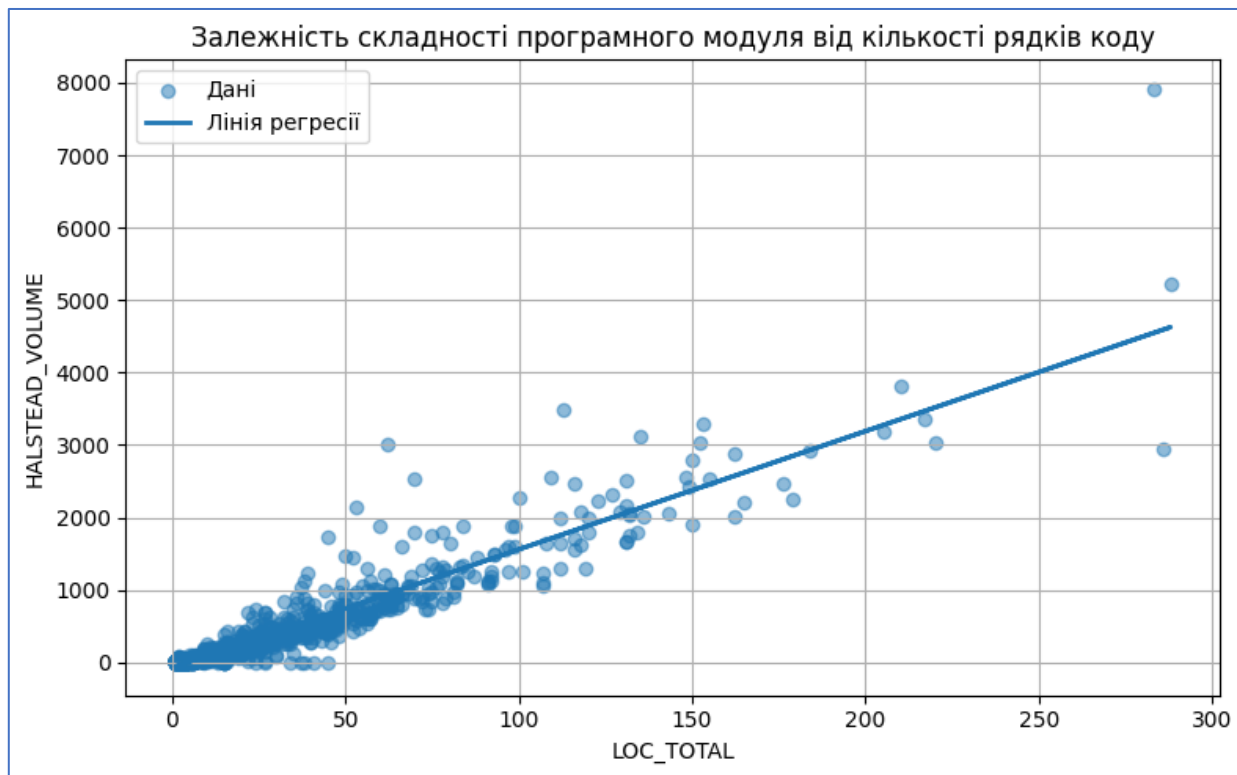


Рисунок 1 – Залежність метрики HALSTEAD_VOLUME від кількості рядків коду LOC_TOTAL

Як видно з рисунка, зі збільшенням кількості рядків програмного коду спостерігається зростання значення метрики HALSTEAD_VOLUME. Отримана регресійна модель демонструє наявність позитивної залежності між розміром програмного модуля та його складністю. Такий результат узгоджується з особливостями процесу розробки програмного забезпечення, оскільки збільшення обсягу коду зазвичай супроводжується зростанням кількості операторів, операндів та внутрішніх зв'язків між елементами програми [3].

Отримані результати свідчать про можливість використання регресійних моделей для аналізу характеристик програмного забезпечення та прогнозування показників, що можуть впливати на його якість. Виявлення подібних залежностей дозволяє своєчасно оцінювати складність програмних модулів та приймати рішення щодо їх подальшого вдосконалення або рефакторингу [4].

Висновки

Регресійний аналіз є одним із практичних інструментів оцінювання якості програмного забезпечення. Використання регресійних моделей дозволяє прогнозувати характеристики програмних систем на основі статистичних даних та підтримувати процес прийняття рішень під час розробки програмного забезпечення. Проведений експеримент показав можливість застосування лінійної регресії для прогнозування кількості дефектів залежно від розміру програмного коду. Отримані результати підтверджують доцільність використання регресійних моделей у задачах програмної інженерії.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Sommerville I. Software Engineering. 10th ed. Boston : Pearson, 2016.
2. Pressman R. S., Maxim B. R. Software Engineering: A Practitioner's Approach. 9th ed. New York : McGraw-Hill Education, 2019.
3. Fenton N., Bieman J. Software Metrics: A Rigorous and Practical Approach. 3rd ed. Boca Raton : CRC Press, 2014.
4. Menzies T., Greenwald J., Frank A. Data Mining Static Code Attributes to Learn Defect Predictors // IEEE Transactions on Software Engineering. 2007. Vol. 33, No. 1. P. 2–13.

Одрихівська Анастасія Олександрівна – студентка групи ІПІ-24б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e-mail: nastyavait0@gmail.com.

Черноволик Галина Олександрівна – кандидат технічних наук, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, e-mail: lina2433@gmail.com.

Одрихівська Анастасія О. – Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: nastyavait0@gmail.com.

Chernovolyk Halyna O. – Ph.D., Associate Professor of Software Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: lina2433@gmail.com.