أساليب جديدة في تعلم الآلة لتحسين التنبؤ بالأخطاء البرمجية

إعداد

عماد نبيل ابوالهدى كائن

إشراف

د.عبدالله مهدي سعيد القرني

د. تركي طلال سالم تركي

المستخلص

التنبؤ بالأخطاء في أنظمة البرامج هو واحدة من المشاكل البحثية المهمة في هندسة البرمجيات، وبواسطة التنبؤ الصحيح لأخطاء البرمجيات يستطيع المطورون بأن يسرعوا من عملية فحص وتحديد أجزاء الشيفرة المصدرية من البرنامج التي تحتوي على أخطاء، ويسهم ذلك في تقليص الوقت والجهد المستغرق عادة في عمليات فحص وصيانة البرمجيات، وفي رفع دقة وجودة المشاريع البرمجية عن استخدام الطرق التقليدية، الهدف الأساسي من هذه الرسالة هو تقديم أدوات تعلم الآلة الموثوقة لمطوري البرامج ليستخدموها في التنبؤ بالأخطاء وتحديدها في أنظمة البرمجيات. ولتحقيق هذا الهدف، تم تطوير أساليب جديدة من تعلم الآلة والتي تجمع أسلوب التعلم الغير إشرافي مع اختيار الخصائص وأسلوب التعلم الإشرافي، وباستخدام عدة من خوارزميات تعلم الآلة الإشرافية مثل آلة المتجهات الداعمة، الغابة العشوائية، الشبكات العصبية والشبكات العصبية العميقة .وقد أظهرت النتائج التجريبية على مجموعة من بيانات التنبؤ بالأخطاء البرمجية مفتوحة المصدر حصول أساليب تعلم الآلة المقترحة على نتائج أداء عالي مع تأثيرات إحصائية مميزة عند مقارنتها ضد أبرز الأساليب الأساسية المعروفة حالياً.

New Machine Learning Approaches to Improve Software Bug Prediction

By

Emad Nabil Abualhuda Kaen

Supervised by

Dr.Abdullah Algarni

Dr.Turki Turki

ABSTRACT

Predicting bugs of a software system is an important research problem in software engineering. By predicting software bugs correctly, developers can accelerate the testing process for locating source code components containing bugs and hence reduce the time associated with software maintenance and development, thereby improving the software development cycle. In today's software industry, plethora of software engineering environments has led to substantial amounts of data stored in repositories. Mining such data is a challenging task. A key goal of this thesis is to deliver reliable machine learning tools to software developers, who would use these tools as prediction calculators to identify bugs in software systems. To accomplish this goal, I develop new machine learning approaches combining an unsupervised technique with feature selection and supervised learning techniques. The supervised learning algorithms include support vector machines, random forests, neural networks, and deep neural networks. Experimental results on various bug prediction datasets demonstrate that my machine learning approaches generate higher performance results with statistical significance when compared against existing baseline approaches.