Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science

IMPROVING A MACHINE LEARNING ALGORITHM FOR MEASURING SOFTWARE QUALITY

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Science**

**By
Zakaria Abdul Wahid Hamed Hamo**

Supervised by

Prof. Dr. Safwan Omer Hasoon Khalil

2024 A.D. **1445 A.H.**

ABSTRACT

Software quality measurement represents the procedure of assessing an attribute of software to quantify its quality. Actually, there are many attributes and metrics used to measure software quality such as functionality, usability, reliability, flexibility, maintainability, and testability. Reliability is one of the key factors used to gauge software quality. Software defect prediction (SDP) is one of the most important factors that affect measuring software's reliability and ensures the quality of software systems. However, the high dimensionality of the features and imbalanced data influence the efficiency of SDP models directly. Machine learning (ML) algorithms and metaheuristic optimization algorithms were considered the most common techniques that are used as SDP models. An imbalanced dataset is considered one of the common problems for machine learning algorithms because it makes the model biased towards the majority class in the binary classifications. In this thesis, two types of SDP models have been proposed based on improving machine learning algorithms to address the issues that are facing SDP models to get high prediction performance.

The first type includes proposing a hybrid binary whale optimization algorithm (BWOA) based on taper-shape transfer functions for solving feature selection problems and dimension reduction as a new software defect prediction model. The values of a real vector that represents the individual encoding have been converted to binary vector by using the four types of Taper-shaped transfer functions to enhance the performance of BWOA and reduce the dimensions of the search space. For the first time, the K-nearest neighbor (KNN) classifier was dependent on the proposed (T-BWOA) to investigate its efficiency as an SDP model. Then, three standard classifiers, which are Naïve Bias (NB), Random Forest (RF), and Decision Tree (DT) have been implemented on (T-BWOA) instead of KNN for the same purpose.

The second type includes proposing a hybrid artificial rabbit optimization (ARO) algorithm with a modified random vector functional link (MRVFL) neural network by depending on two paradigms. The first paradigm depended on utilizing a Single-objective (ARO) algorithm to train the (MRVFL) neural network, while the second paradigms, which has been used to train the (MRVFL) neural network also, is based on using a Multi-objective (ARO) algorithm. The training process includes

tuning the linkage weights and biases of the (MRVFL) neural network based on the fitness value. For the (SARO-MRVFL) model, the fitness value is represented by one objective function only; while, for the (MARO-MRVFL) model, the fitness value is represented by three objective functions.

The performance of the suggested models was evaluated using ten standard software defect prediction datasets from the NASA repositories. Each dataset has a different number of features and patterns. Also, seven evaluation metrics have been used to assess the effectiveness of the suggested models. The experimental results have shown that the performance of T2-BWOA-KNN produced promising results compared to other methods including ten methods from the literature. In addition, the performance of the NB, DT, and RF classifiers on (T-BWOA) was acceptable with slight differences among them. Moreover, the (MARO-MRVFL) model has superior performance compared to other proposed models and other methods in the literature. The accuracy obtained for the ten datasets was Cm1, 0.9121; JM1, 0.8491; KC1, 0.8338; KC3, 0.9512; MW1, 0.93; PC1, 0.9279; PC2, 0.9498; PC3, 0.8875; PC4, 0.9476; PC5, 0.8260. AUC was as follows: Cm1, 0.8878; JM1, 0.8616; KC1, 0.7660; KC3, 0.9415; MW1, 0.9619; PC1, 0.7053; PC2, 0.9662; PC3, 0.8419; PC4, 0.9076; PC5, 0.8039. Furthermore, statistical analysis using the Kendall W test and Paired samples t-test confirms the significance of the proposed (MARO-MRVFL) model compared to other alternative SDP models.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

تحسين خوارزمية لتعلم الآلة لقياس جودة البرمجيات

اطروحة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب

من قبل

زكريا عبدالواحد حميد حمو

بإشراف
أ.د. صفوان عمر حسون خليل

١٤٤٥ هـ ٢٠٢٤ م

الخلاصة

يمثل قياس جودة البرمجيات إجراء تقييم إحدى سمات البرنامج لتحديد جودته. في الواقع، هناك العديد من السمات والمقاييس المستخدمة لقياس جودة البرامج مثل الأداء الوظيفي وسهولة الاستخدام والموثوقية والمرونة وقابلية الصيانة وقابلية الاختبار. تعد الموثوقية أحد العوامل الرئيسية المستخدمة لقياس جودة البرامج. يعد التنبؤ بالعيوب البرمجية (SDP) أحد أهم العوامل التي تؤثر على قياس موثوقية البرامج ويلعب دورًا مهمًا في ضمان جودة الأنظمة البرمجية. ومع ذلك، فإن الأبعاد العالية للميزات والبيانات غير المتوازنة لها تأثير مباشر على أداء نماذج ال (SDP). تعتبر خوارزميات التعلم الآلي وخوارزميات التحسين من أكثر التقنيات شيوعًا والمستخدمة كنماذج (SDP). كذلك, تعتبر مجموعة البيانات غير المتوازنة إحدى المشاكل الشائعة لخوارزميات التعلم الآلي لأنها تجعل النموذج متحيزًا نحو فئة الأغلبية في التصنيفات الثنائية. في هذه الأطروحة، تم اقتراح نوعين من نماذج SDP المبنية على تحسين تقنية تعلم الالة لمعالجة المشكلات التي تواجه نماذج SDP للحصول على أداء تنبؤي عالي.

يتضمن النوع الأول اقتراح خوارزمية تحسين الحوت الثنائي الهجين (BWOA) استنادًا إلى وظائف نقل الشكل المستدق لحل مشكلات اختيار الميزات وتقليل الأبعاد حيث تم اقتراح نموذج جديد للتنبؤ بعيوب البرامج. تم تحويل قيم المتجه الحقيقي الذي يمثل التشفير الفردي إلى متجه ثنائي باستخدام الأنواع الأربعة من وظائف النقل ذات الشكل المستدق لتعزيز أداء BWOA وتقليل أبعاد فضاء البحث. لأول مرة، اعتمد مصنف الجار الأقرب (KNN) على المقترح (T-BWOA) للتحقق من كفاءة نموذج ال (SDP) المقترح. بعد ذلك، تم تطبيق ثلاثة مصنفات قياسية، وهي التحيز الساذج (NB)، وشجرة القرار(DT) ، والغابة العشوائية(RF) ، على (T-BWOA) بدلاً من KNN لنفس الغرض.

أما النوع الثاني فيشمل تهجين خوارزمية الأرانب الاصطناعية (ARO)مع الشبكة العصبية المعدلة للارتباط الوظيفي للمتجه العشوائي (MRVFL) وذلك بالاعتماد على نهجين. يعتمد النهج الأول على استخدام خوارزمية الهدف الواحد (SARO) لتدريب الشبكة العصبية (MRVFL) , بينما تم استخدام الطريقة الثانية لتدريب الشبكة العصبية (MRVFL) أيضًا، وذلك بالاعتماد على استخدام خوارزمية متعددة الأهداف (MARO) .تتضمن عملية التدريب ضبط أوزان الشبكة والمتغيرات الاخرى للشبكة العصبية (MRVFL) بناءً على قيمة اللياقة. بالنسبة لنموذج-(SARO MRVFL)، يتم تمثيل قيمة اللياقة بهدف واحد فقط؛ بينما بالنسبة لنموذج (MARO-MRVFL) فإن قيمة اللياقة تتمثل بثلاثة أهداف.

تم تقييم أداء النماذج المقترحة باستخدام عشر مجموعات بيانات قياسية للتنبؤ بعيوب البرمجيات من مستودعات ناسا. تحتوي كل مجموعة بيانات على عدد مختلف من الميزات والأنماط. كما تم استخدام سبعة مقاييس تقييم لتقييم فعالية النماذج المقترحة. أظهرت النتائج التجريبية أن أداء T2-BWOA-KNN أنتج نتائج واعدة مقارنة بالطرق الأخرى بما في ذلك عشر طرق من الأدبيات. بالإضافة إلى ذلك، كان أداء المصنفات NB و DT و RF على (T-BWOA) مقبولاً مع وجود اختلافات طفيفة فيما بينها. علاوة على ذلك، يتمتع نموذج (MARO-MRVFL) بأداء متفوق مقارنة بالنماذج الأخرى المقترحة والطرق الحديثة الأخرى. وكانت الدقة التي تم الحصول عليها

لمجموعات البيانات هي (CM1=0.9121) ; (JM1= 0.8491) ; (KC1= 0.8338) ; (KC3= 0.9512) ; (MW1=0.93) ; (PC1=0.9279) ; (PC2= 0.9498) ; (PC3= 0.8875) ; (PC4=0.9476) ; (PC5= 0.8260) . اما فيما يخص المعيار (AUC) فقد كانت النتائج التي تم الحصول عليها لمجموعة البيانات هي (CM1=0.8878) ; (JM1= 0.8616); (KC1= 0.7660) ; (KC3= 0.9415) ; (MW1=0.9619) ; (PC1=0.7053) ; (PC2= 0.9662) ; (PC3= 0.8419) ; (PC4=0.9076) ; (PC5= 0.8039) .

علاوة على ذلك، يؤكد التحليل الإحصائي باستخدام اختبار (Kendall W) واختبار (Paired samples t-test) على أهمية نموذج (MARO-MRVFL) المقترح مقارنة بنماذج SDP البديلة الأخرى.