Vulnerability Type Classification of Common Vulnerabilities and Exposures (CVE) Using ML Methods: Comparative Study

Prepared by

Sabreen Ayed Dabak Alfawareh

Supervisor

Omar Husain ALTarawneh

Abstract

Countering violent extremism vulnerabilities is among the biggest challenges in the field of cybersecurity, as cyber-attacks are constantly increasing and evolving, which necessitates the need to develop effective methods to detect and address them. The study goals is to compare different MLalgorithms for vulnerability classification, evaluation of the effectiveness of ML(ML) techniques in identifying electronic threats and Exploring the effectiveness of optimization using Gray-Wolf Optimization in cybersecurity. Emphasis was placed on using several ML models: logistic regression, Gaussian Naive Bayes, SVM, random forest classifier, decision trees, and simple neural network. These models were evaluated using a set of criteria such as accuracy, recallability, F1, F2, AUC-ROC, and AUC-PR. In addition, Gray-Wolf Optimization (GWO)has been implemented to improve the performance of these models. Before Gray-Wolf Optimization, the SVM models and the random forest classifier showed the highest accuracy rates, while the simple neural network

achieved the greatest recovery. After applying the optimization, all models experienced an improvement in their parameters, with the accuracy of the SVM model and the random forest classifier model reaching 0.97. The highest performance was achieved using the simple neural network model and the SVM model after applying optimization, where the accuracy reached 0.97 and the AUC-ROC values were close to 0.99. The results show that using ML to classify vulnerabilities in cybersecurity is effective, especially when optimization techniques such as Gray-Wolf Optimization are used. Despite the encouraging results, more research is needed to develop these methods and examine their effectiveness against different and complex data models.

Keywords:Vulnerability Classification, ML,Cybersecurity, Gray-Wolf Optimization, Cyber Threat Identification.

تصنيف أنواع الثغرات والتعرضات الشائعة باستخدام أساليب التعلم الآلي: دراسة مقارنة (CVE)

إعداد

صابرين عايد الفواعرة

إشراف

الدكتور عمر حسين الطراونة

الملـخـص

تعد مكافحة نقاط الضعف للتطرف العنيف من بين أكبر التحديات في مجال الأمن السيبراني ، حيث تتزايد الهجمات الإلكترونية وتتطور باستمرار ، مما يستلزم الحاجة إلى تطوير أساليب فعالة لاكتشافها ومعالجتها ، وكانت أهداف الدراسة الرئيسية هي مقارنة أساليب التعلم الآلي المختلفة لـ تصنيف الثغرات الأمنية وتقييم فعالية تقنيات التعلم الآلي في تحديد التهديدات الإلكترونية واستكشاف فعالية التحسين باستخدام تحسين Gray-Wolf Optimization (GWO) في الأمن السيبراني. تم التركيز على استخدام العديد من نماذج التعلم الآلي: الانحدار اللوجستي ، Gaussian Naive Bayes، SVM ، مصنف الغابات العشوائية ، أشجار القرار ، والشبكة العصبية البسيطة. تم تقييم هذه النماذج باستخدام مجموعة من المعايير مثل الدقة وقابلية الاسترجاع و F1 و F2 و AUC-ROC و AUC-PR. بالإضافة إلى ذلك ، تم تطبيق تحسين الذئب الرمادي لتحسين أداء هذه النماذج. قبل تحسين Gray-Wolf Optimization (GWO)، أظهرت نماذج SVM ومصنف الغابة العشوائية أعلى معدلات الدقة ، بينما حققت الشبكة العصبية البسيطة أكبر قدر من التعافي. بعد تطبيق التحسين ، شهدت جميع النماذج تحسنًا في معلماتها ، حيث وصلت دقة نموذج SVM ونموذج مصنف الغابة العشوائية إلى 0.97. تم تحقيق أعلى أداء باستخدام نموذج الشبكة العصبية البسيط ونموذج SVM بعد تطبيق التحسين ، حيث وصلت الدقة إلى 0.97 وكانت قيم AUC-ROC قريبة من 0.99. تظهر النتائج أن استخدام التعلم الآلي لتصنيف نقاط الضعف في

الأمن السيبراني فعال ، خاصة عند استخدام تقنيات التحسين مثل تحسين Gray-Wolf Optimization (GWO)، على الرغم من النتائج المشجعة ، هناك حاجة إلى مزيد من البحث لتطوير هذه الأساليب وفحص فعاليتها ضد نماذج البيانات المختلفة والمعقدة.

الكلمات الدالة:تصنيف نوع الضعف ، التعلم الآلي ، الدراسة المقارنة ، الأمن السيبراني ، تحسين الذئب الرمادي ، تحديد التهديدات السيبرانية.