

Detection and Classification of Complex Java Code Vulnerabilities and Errors Using Artificial Intelligence Techniques

R. Reyanson Sosai* and A.W.A.T Dilhan

Department of Information Communication Technology, Faculty of Technology, University of Ruhuna, Kamburupitiya, Sri Lanka.

**reyanson4222@gmail.com*

Abstract

Ensuring the safety of software applications, particularly those developed in Java, is vital because of the risks associated with vulnerabilities such as Cross-Site Scripting (XSS) SQL Injection, incorrect validation of array indices, integer overflow, underflow, and memory allocation issues. This research seeks to create an automated system for identifying and categorizing these vulnerabilities using machine learning methods, thus minimizing the need for manual code inspections and enhancing overall software security. The study entailed gathering Java source code from public repositories and vulnerability databases like CWE and CVE. Essential features such as cyclomatic complexity, the depth of the abstract syntax tree (AST), and control flow metrics were extracted from the code to train machine learning algorithms, including Random Forest, Support Vector Machines (SVM), and Artificial Neural Networks (ANN). The models were assessed using metrics like precision, recall, and F1-score. Preliminary results indicate that Random Forest outperforms other models in identifying vulnerabilities, particularly in handling complex code structures. The system also continuously learns from new data, improving detection accuracy over time. Although deep learning models showed potential, their high computational requirements make them less practical for this task. Future research should focus on expanding the dataset and exploring more advanced machine learning models, such as transformers, to enhance detection capabilities. Finally, this approach could be extended to other programming languages, including Python and JavaScript, to broaden its applicability.

Keywords: *Cross-Site Scripting, SQL Injection, Java Code Vulnerabilities, Machine Learning, Software Security.*