ABSTRACT

This research aims to compare effective Machine Learning models in detecting SQL Injection attacks in security systems. The dataset used was collected from Kaggle resources published by Syed Saqlain Hussain Shah, which is the highest upvoted dataset in the SQL Injection category. The models developed include Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Logistic Regression (LR). The research process involves splitting the data into 70% training data and 30% testing data, training the models, testing the effectiveness of the models, and implementing preventive measures against SQL Injection attacks. The results show that the SVM model achieved an accuracy rate of 99.82%, a precision rate of 99.88%, and a recall (sensitivity) rate of 99.34%. The KNN model achieved an accuracy rate of 79.28%, a precision rate of 98.38%, and a recall (sensitivity) rate of 73.31%. The LR model achieved an accuracy rate of 98.99%, a precision rate of 99.94%, and a recall (sensitivity) rate of 98.70%. By using the Machine Learning approach, this research contributes to enhancing system security against SQL Injection attacks.

Keywords: SQL Injection, Support Vector Machine, K-Nearest Neighbor, Logistic Regression, Machine Learning