

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of Sétif 1 – Ferhat ABBAS
Faculty of Sciences
Department of Computer Science



MASTER'S THESIS

Presented by :
ZITOUNI Ahmed Faouzi
SEDJAL Moheamed Aymen Dhaya Eddin

Field of Study : Computer Science
Specialty : Data Engineering and Web Technologies

Title :
Detecting SQL Injections using Deep Learning

In front of the jury

DR. BENZINE Mehdi	University of Sétif 1 – Ferhat ABBAS	Supervisor
DR. LAKHFIF Abdelaziz	University of Sétif 1 – Ferhat ABBAS	Chairperson
DR. MEDIANI Chahrazed	University of Sétif 1 – Ferhat ABBAS	Examiner

Academic Year: 2024–2025

Abstract

SQL injection attacks remain one of the biggest threats to web applications, because they allow the attacker to gain trusted access to data without authorization, which can lead to irreparable damages. As part of this project, we examined how deep learning and machine learning can aid in detecting these attacks automatically. In total, we built and evaluated six models: Logistic Regression, Support Vector Machine (SVM), Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and BERT (Bidirectional Encoder Representations from Transformers). Overall, BERT achieved the highest scores in accuracy, precision, recall, and F1-score. This demonstrates that transformer-based models such as BERT have a better understanding of SQL query structures, which makes them efficient in detecting complex attacks. This study shows how deep learning, especially BERT, can improve web application security.

Keywords: SQL Injection, Web Application Security, Deep Learning, Machine Learning, BERT, Transformer Models, Model Comparison, Threat Detection

General Introduction	1
Chapter 1	Error! Bookmark not defined.
SQL Injections.....	Error! Bookmark not defined.
1.1 Introduction	Error! Bookmark not defined.
1.2 SQL injection.....	Error! Bookmark not defined.
1.2.1 Definition	Error! Bookmark not defined.
1.2.2 How SQL Injection Works	Error! Bookmark not defined.
1.3 Techniques of SQL Injection.....	Error! Bookmark not defined.
1.3.1 Error-Based SQL Injection	Error! Bookmark not defined.
1.3.2 Blind SQL Injection	Error! Bookmark not defined.
1.3.2.1 Content-Based Blind SQL Injection.....	Error! Bookmark not defined.
1.3.2.2 Time-Based Blind SQL Injection	Error! Bookmark not defined.
1.3.3 Tautology-Based SQL Injection	Error! Bookmark not defined.
1.3.4 Union-Based SQL Injection	Error! Bookmark not defined.
1.4 Methods to prevent SQL Injection attacks	Error! Bookmark not defined.
1.4.1 Prepared Statements (with Parameterized Queries)	Error! Bookmark not defined.
1.4.2 Stored Procedures.....	Error! Bookmark not defined.
1.4.3 Input validation	Error! Bookmark not defined.
1.4.4 Escaping All User-Supplied Input	Error! Bookmark not defined.
1.5 Conclusion.....	Error! Bookmark not defined.
Chapter 2	Error! Bookmark not defined.
Deep Learning	Error! Bookmark not defined.
2.1 Introduction	Error! Bookmark not defined.
2.2 Machine learning	Error! Bookmark not defined.
2.2.1 Machine Learning Types	Error! Bookmark not defined.
2.2.1.1 Supervised Learning (SL):.....	Error! Bookmark not defined.
2.2.1.2 Unsupervised Learning:	Error! Bookmark not defined.
2.2.1.3 Reinforcement Learning:	Error! Bookmark not defined.
2.2.2 Machine learning algorithms.....	Error! Bookmark not defined.
2.2.2.1 Logistic regression	Error! Bookmark not defined.

2.2.2.2 Support Vector Machine (SVM)	Error! Bookmark not defined.
2.2.3 Real-world machine learning use cases	Error! Bookmark not defined.
2.3 Deep Learning	Error! Bookmark not defined.
2.3.1 Artificial Neural Networks (ANNs)	Error! Bookmark not defined.
2.3.4 Deep Learning Training Cycle.....	Error! Bookmark not defined.
2.3.2 Activation functions	Error! Bookmark not defined.
2.3.2.1 Non-Linear Activation Functions.....	Error! Bookmark not defined.
2.3.3 Deep learning architectures.....	Error! Bookmark not defined.
2.3.3.1 Recurrent Neural Networks	Error! Bookmark not defined.
2.3.3.2 Long Short-Term Memory Networks	Error! Bookmark not defined.
2.3.3.4 Transformers	Error! Bookmark not defined.
2.3.3.5 BERT	Error! Bookmark not defined.
2.4 Related works.....	Error! Bookmark not defined.
2.4.1 Introduction	Error! Bookmark not defined.
2.4.2 Detection Approaches Using Machine Learning	Error! Bookmark not defined.
2.4.3 Deep Learning for SQL Injection Detection.....	Error! Bookmark not defined.
2.5 Conclusion.....	Error! Bookmark not defined.
Chapter 3	Error! Bookmark not defined.
Conception and Implementation	Error! Bookmark not defined.
3.1 Introduction	Error! Bookmark not defined.
3.2. Dataset	Error! Bookmark not defined.
3.3 Development Environment Overview	Error! Bookmark not defined.
3.3.1 Programming language	Error! Bookmark not defined.
3.3.2 Libraries Used	Error! Bookmark not defined.
3.3.3 Development setup.....	Error! Bookmark not defined.
3.3.3.1 Visual Studio Code	Error! Bookmark not defined.
3.3.3.2 Jupyter Notebook	Error! Bookmark not defined.
3.3.3.3 Google Colab	Error! Bookmark not defined.
3.4 Models Implemented	Error! Bookmark not defined.
3.4.1 Support Vector Machine (SVM).....	Error! Bookmark not defined.
3.4.2 Logistic Regression (LR)	Error! Bookmark not defined.

3.4.3 Multilayer Perceptron (MLP).....	Error! Bookmark not defined.
3.4.4 Recurrent Neural Network (RNN)	Error! Bookmark not defined.
3.4.5 Long Short-Term Memory (LSTM)	Error! Bookmark not defined.
3.4.6 BERT.....	Error! Bookmark not defined.
3.4.6.1 Why BERT for SQL Injection Detection.....	Error! Bookmark not defined.
3.4.6.2 BERT Code and Implementation	Error! Bookmark not defined.
3.5 Comparative Summary of Model Architectures	Error! Bookmark not defined.
3.6 Conclusion.....	Error! Bookmark not defined.
Chapter 4	Error! Bookmark not defined.
Performance Evaluation and Results.	Error! Bookmark not defined.
4.1 Introduction	Error! Bookmark not defined.
4.2 Performance Evaluation Metrics	Error! Bookmark not defined.
4.3 Evaluation on Test Set (20%)	Error! Bookmark not defined.
4.3.1 Results Presentation	Error! Bookmark not defined.
4.3.2 Visual Performance Analysis.....	Error! Bookmark not defined.
4.3.3 Performance Analysis of BERT	Error! Bookmark not defined.
4.4.3.1 Training Loss Curve	Error! Bookmark not defined.
4.4.3.2 Model Evaluation on New Unseen Data	Error! Bookmark not defined.
4.4 Conclusion.....	Error! Bookmark not defined.
General Conclusion.....	Error! Bookmark not defined.
References.....	Error! Bookmark not defined.