الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 1945 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbes

Thesis

To obtain the diploma of **Master Degree**

Field of Study: **Computer Science**

Specialization: **SIW**

Theme

Comparative Study of AI-Based Models for Malicious URL Detection

---

Presented by
Bellout Sarra

Defended on: **July 8th, 2025**
In front of the jury composed of

Dr. Anani Djihed	President of the Jury
Dr. Khaldi Miloud	Thesis Supervisor
Dr. Mahammed Nadir	Co-Supervisor
Dr. BabaAhmed Manel	Examiner

Academic Year: 2024/2025

Abstract

Phishing and other malicious URL-based threats continue to challenge cybersecurity systems, exploiting the limitations of traditional defense mechanisms such as blacklists and heuristic detection methods.

This study conducts a comparative analysis of artificial intelligence (AI)-based techniques used for malicious URL detection. The investigation focuses on both traditional machine learning models, such as LightGBM and XGBoost, and deep learning approaches like BiLSTM with attention. The research also reviews the role of metaheuristic optimization methods—particularly Particle Swarm Optimization (PSO) and Optuna—for hyperparameter tuning and feature selection.

By analyzing several benchmark studies and recent contributions, the study highlights the strengths, limitations, and trade-offs associated with these models in terms of accuracy, scalability, and robustness. Special attention is given to feature engineering strategies such as lexical analysis and TF-IDF vectorization, which significantly influence performance across different classifiers.

This work provides a structured theoretical foundation for future research on adaptive phishing detection systems, offering valuable insights for academic and industrial researchers interested in building scalable, explainable, and AI-driven security solutions.

Keywords: Phishing Detection, Malicious URLs, Machine Learning, Deep Learning, Metaheuristics, PSO, Optuna.