Building a machine learning model for intrusion detection based on network data.

Le Tien Dung

Abstract: The thesis focuses on building an Intrusion Detection System (IDS) using machine learning algorithms on the standard NSL-KDD dataset. The study performs classification of network traffic into normal or intrusion states through a rigorous data preprocessing process and effective feature selection. The author applies a data-centric approach to optimize model performance. The system includes a visual monitoring Dashboard interface, providing transparent evaluation metrics. Experimental results show that the model can provide early warnings of network threats with high accuracy, effectively supporting cybersecurity efforts.

Keywords: Intrusion Detection System (IDS), Machine Learning, NSL-KDD dataset

References

[1] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2016.

[2] L. Breiman, "Random Forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[3] H. Debar, M. Dacier, and A. Wespi, "Towards a taxonomy of intrusion-detection systems," Comput. Networks, vol. 31, no. 8, pp. 805–822, 1999.

[4] Vietnix, "IDS là gì? Hệ thống phát hiện xâm nhập mạng (IDS)," 2025. [Online]. Available: https://vietnix.vn/ids-la-gi/

[5] Vbee, "Machine Learning là gì? Tất tần tật kiến thức về Học máy," 2025. [Online]. Available: https://vbee.vn/blog/ai/machine-learning/