Teaching Machine Learning for Cybersecurity

Pierre Parrend
ICube - Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie, UMR 7357
Université de Strasbourg, CNRS, 67000, Strasbourg, France
EPITA, 14-16 Rue Voltaire, 94270 Le Kremlin-Bicêtre
pierre.parrend@epita.fr

***Abstract*—Teaching Machine Learning for cybersecurity is increasingly critical in all Cybersecurity curricula, and brings a valuable skill for Data Science curricula. However, each of these knowledge domains usually requires a significant dedicated background training: systems and network for cybersecurity, statistics and algorithms for data science. It is therefore necessary to identify fundamental concepts to teach Machine Learning and cybersecurity to student with various background. The objective is to provide them with required competences to characterise the data under investigation, to define the objective of the analysis and to evaluate the result of such analysis.**

***Index Terms*—Machine learning, cybersecurity, teaching**

I. Course Objectives

Teaching Machine Learning for cybersecurity is increasingly critical in all Cybersecurity curricula, and brings a valuable skill for Data Science curricula. However, each of these knowledge domains usually requires a significant dedicated background training: systems and network for cybersecurity, statistics and algorithms for data science. It is therefore necessary to identify fundamental concepts to teach Machine Learning and cybersecurity to student with various background. The objective is to provide them with required competences to characterize the data under investigation [1], to define the objective of the analysis [2] and to evaluate the result of such analysis.

II. Course Content

The emphasis in the Machine Learning and Cybersecurity cours is put on the identification of suitable metrics, in particular for learning in unbalanced datasets, and evaluating the quality of learning in underfitting, overfitting and git fit situation [3]. Useful learning operations for cybersecurity applications are classification, for the detection of known attacks, novelty detection for the detection of anomalies when a clean history is available, and anomaly detection when anomalies are hidden inside the data [4]. The main algorithms which are studied in the course are LSTM (Long Short Term Memory) neural network and ARIMA (AutoRegressive Integrated Moving Average) for 1-dimensional data, XGBoost for classification, 1-Class SVM (Support Vector Machine) for novelty detection, Isolation Forrest, Local Outlier Factor or Robust Covariance for anomaly detection.

The course includes several lab sessions for applying a fulfledged data analysis process on the two cybersecurity problems of classification for network attacks and anomaly detection for fraud detection. The data analysis process covers data visualisation, exploratory data analysis (EDA), data preparation, and analysis through statistical and machine learning. The acquired skills are developed in a group projet so that the students gain experience with various target datasets from Information System (IS) and internet of Things (IoT) environments.

Additional challenges such as the principles of adversarial machine learning [5], the detection of complex attacksnavarro2018systematic and the exploitation of machine learning in intrusion detection [6], [7] or in Security Operation Centers (SOCs) are presented to prepare the students to the issues they will face in the application of Machine Learning for Cybersecurity. In relevant curriculum where students also consider systemic analysis, the principles [8] and implementations [9] of the use of machine learning for artificial immune systems are also introduced and provide a useful pattern for network-wide protection.

The course has been given in various flavors for computer science and engineer students in Data Scientists (Telecom Physique Strasbourg and University of Strasbourg), Computer scientists with Artificial Intelligence specialisation (EPITA), Cybersecurity specialists (CNAM).

III. Reference Datasets

Following datasets, openly available for the community, are used throughout the course and in the lab sessions and projects:

- ISCX-URL2016, for the Detection of malicious URLs in Information Systems[1] [10]
- Waterloo spam corpus, for Spam detection by classification in Information Systems[2] [11]
- CCCS-CIC-AndMal-2020, for Detection by classification in Malware[3] [12]
- Secure Water Treatment (SWaT), for Attack detection, supervised in Critical Infrastructures[4] [13]
- IoMT Dataset, for supervised attack detection in Medical IoT[5] [14]

The variety of these datasets enable to highlight, in particular in the student projects, the impact of the topology of data on the performance of the learning algorithms. In particular,

[1] https://www.unb.ca/cic/datasets/url-2016.html
[2] https://plg.uwaterloo.ca/ gvcormac/treccorpus07/
[3] https://www.unb.ca/cic/datasets/andmal2020.html
[4] https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/
[5] https://www.cse.wustl.edu/ jain/ehms/index.html

datasets from Information Systems, having a high number of dimensions (typically 40), behave dramatically differently than dasets from critical systems where the number of significant dimensions is very limited (around 4 to 6).

IV. LITERATURE

The literature used as reference for this course, and which the student should refer to for complementary materials is conpound of following books:

- Machine Learning for Cybersecurity Cookbook [15]
- Machine Learning & Security, C. Chio et D. Freeman, O'Reilly, 2018. [16]

All scientific papers referenced in this article (with exception of the architrace model [1] which references previous work of the author in the domain of cybersecurity pedagogy) pertain to the references given to the students.

REFERENCES

[1] V. Legrand, O. Gaouar, and P. Parrend, "Architrace : Apprentissage de la sécurité par les traces," in *Workshop sur l'enseignement de la sécurité des systèmes d'informations (WESSI 2014)*, May 2014. [Online]. Available: http://icube-publis.unistra.fr/7-LGP14

[2] R. Das and T. H. Morris, "Machine learning and cyber security," in *2017 International Conference on Computer, Electrical & Communication Engineering (ICCECE)*. IEEE, 2017, pp. 1–7.

[3] A. Beaugnon and P. Chifflier, "Machine learning for computer security detection systems: practical feedback and solutions," *Proceedings of the 2018 Intelligence Artificielle et Cybersécurité/Artificial Intelligence and Cybersecurity (C&ESAR), Rennes, France*, pp. 19–21, 2018.

[4] P. Parrend, J. Navarro, F. Guigou, A. Deruyver, and P. Collet, "Foundations and applications of artificial intelligence for zero-day and multi-step attack detection," *EURASIP Journal on Information Security*, vol. 2018, no. 1, pp. 1–21, 2018.

[5] L. Huang, A. D. Joseph, B. Nelson, B. I. Rubinstein, and J. D. Tygar, "Adversarial machine learning," in *Proceedings of the 4th ACM workshop on Security and artificial intelligence*, 2011, pp. 43–58.

[6] D. Denning and P. G. Neumann, *Requirements and model for IDES-a real-time intrusion-detection expert system*. SRI International Menlo Park, 1985, vol. 8.

[7] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on software engineering*, no. 2, pp. 222–232, 1987.

[8] P. Parrend, "Immune-based defence and resiliency," *Nature-Inspired Cyber Security and Resiliency: Fundamentals, techniques and applications*, p. 75, 2019.

[9] S. A. Hofmeyr and S. Forrest, "Architecture for an artificial immune system," *Evolutionary computation*, vol. 8, no. 4, pp. 443–473, 2000.

[10] M. S. I. Mamun, M. A. Rathore, A. H. Lashkari, N. Stakhanova, and A. A. Ghorbani, "Detecting malicious urls using lexical analysis," in *International Conference on Network and System Security*. Springer, 2016, pp. 467–482.

[11] G. V. Cormack, J. M. Gómez Hidalgo, and E. P. Sánz, "Spam filtering for short messages," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, 2007, pp. 313–320.

[12] Z. Liu, R. Wang, N. Japkowicz, D. Tang, W. Zhang, and J. Zhao, "Research on unsupervised feature learning for android malware detection based on restricted boltzmann machines," *Future Generation Computer Systems*, vol. 120, pp. 91–108, 2021.

[13] J. Goh, S. Adepu, K. N. Junejo, and A. Mathur, "A dataset to support research in the design of secure water treatment systems," in *International conference on critical information infrastructures security*. Springer, 2016, pp. 88–99.

[14] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, "Intrusion detection system for healthcare systems using medical and network data: A comparison study," *IEEE Access*, vol. 8, pp. 106 576–106 584, 2020.

[15] E. Tsukerman, *Machine Learning for Cybersecurity Cookbook: Over 80 recipes on how to implement machine learning algorithms for building security systems using Python*. Packt Publishing Ltd, 2019.

[16] C. Chio and D. Freeman, *Machine learning and security: Protecting systems with data and algorithms*. " O'Reilly Media, Inc.", 2018.