Machine Learning Algorithms in Cyber Security

Eljona Proko
Computer Science Dept.
University "Ismail Qemali", Vlore
eljona.proko@univlora.edu.al

Alketa Hyso
Computer Science Dept.
University "Ismail Qemali", Vlore
alketa.hyso@univlora.edu.al

Dezdemona Gjylapi
Computer Science Dept.
University "Ismail Qemali", Vlore
dezdemona.gjylapi@univlora.edu.al

Abstract

Artificial intelligence (AI) has made incredible progress, resulting in highly capable software and advanced autonomous machines. Meanwhile, the cyber domain has become a battleground for access, influence, security and control. This paper will address key AI technologies including machine learning in an attempt to help in understanding their role in cyber security and the implications of these new technologies. This paper discusses and highlights different applications of machine learning in cyber security.

1. Introduction

Technologies such as Big Data, Cloud Computing, Artificial Intelligence, etc., have been repeated again and again in multiple forums, in many cases without a clear understanding of their significance or their application to solving real problems effectively. AI is the creation of intelligent machines that can learn from experience, allowing them to work and react as a human would. This technology enables computers to be trained to process large amounts of data and identify trends and patterns. Machine learning techniques have been applied in many areas of science due to their unique properties like adaptability, scalability, and potential to rapidly adjust to new and unknown challenges. Cyber security is a fast-growing field demanding a great deal of attention because of remarkable progresses in social networks, cloud and web technologies, online banking, mobile environment, smart grid, etc. Machine Learning, a branch of AI (Figure 1), is being successfully applied to solve a small part of the problems. Machine Learning – sometimes referred to more generally as Artificial Intelligence (AI) - is a powerful tool used by cyber-security companies. The technology of Applied Artificial Intelligence (AI powered by Machine Learning) is an increasingly important way in which we can scale the detection and classification of malware.

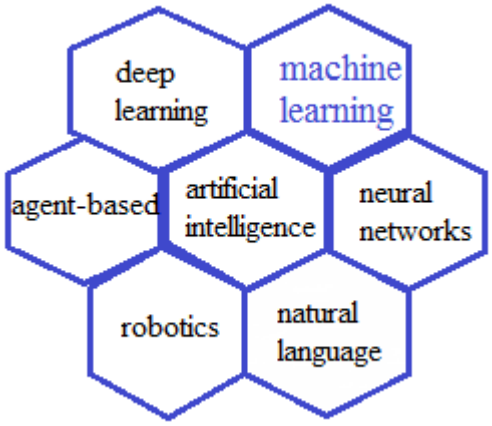

Figure 1: Artificial intelligence branch

Diverse machine learning methods have been successfully deployed to address such wide-ranging problems in computer security.

2 Machine Learning

Artificial Intelligence [Wan 08] is the field of science that studies the synthesis and analysis of computational agents that act intelligently. Machine learning is a subset of the broader field of Artificial Intelligence. The current applications of AI are mostly restricted to Machine Learning (ML).

Machine Learning and Artificial Intelligence [Mar 18] are being connected more extensively crosswise over industries and applications than any other time in recent

memory as computing power, storage capacities and data collection increase.
Machine Learning teaches a machine how to answer a question or how to make a decision on its own. It contrasts with traditional programming, which requires giving a machine explicit instructions for it to answer specific questions. In fact, every imaginable case has to be programmed ahead of time in order to cover all possible situations. ML may encompass techniques such as statistics, mathematical optimizations, or data mining. ML algorithms try to make decisions about their behavior and find ways to solve problems by inferring them from models based on sample inputs that represent real-life scenarios.
There are multiple types of ML and each works very differently. If we generalize the field, we can define three main categories of ML (illustrate in Figure 2): supervised learning, unsupervised learning and reinforcement learning.

2.1 Supervised Learning

In supervised learning, the machine is trained using sample data that is labeled to tell the machine what the data represents. A supervised learning algorithm with an input variable denoted as P and an output variable denoted as Q and algorithms are used to create and learn a mapping function (f) via the input to the output.
The goal of a supervised learning algorithm is to achieve an estimate mapping function so that for every new input (P), a new predicted output (Q) is created. In other words, the learning algorithm receives a set of inputs with their corresponding outputs, and the algorithm learns by equating its concrete output with correct outputs in order to find errors and have the learning model modified accordingly. Supervised learning algorithms make use of patterns to predict the values of the label on unlabelled data. This is achieved by classification, regression, prediction, etc.
Based on that training, the machine should be able to analyze new data and predict the correct answer. Supervised learning has applications such as disease diagnostics, or speech recognition.

2.2 Unsupervised Learning

In unsupervised learning, the machine is trained using data that doesn't have labels. Unsupervised learning is where only an input data (P) is available with no equivalent output variables. The aim of unsupervised learning is to model the construction of the data in order to learn more about the data. Algorithms are required to discover a structure, an inference and meaning within the data in order to arrive to a conclusion. These algorithms do not have any type of historical data in order to predict the output unlike supervised algorithms. That means that the machine does not know what the data represents nor what answers are expected. The machine will have to figure out on its own the patterns and structure of the unlabeled input and discover the expected output. The classification of movie genres in Netflix is an example of unsupervised learning.

2.3 Reinforcement Learning

In reinforcement learning, the machine interacts with its environment to achieve a certain goal. It is similar to unsupervised learning, as the machine is trained using unlabeled data. However, in reinforcement learning, the machine receives feedback on the outcome.

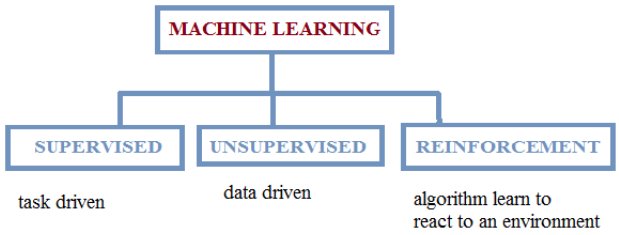

Figure 2: Three main categories of Machine Learning

3 Cyber Security

Security is becoming one of the most important topics in industrial IT and Operational Technology (OT), i.e. the hardware and software used in the production area.
Cyber security is defined as technologies and processes constructed to protect computers, computer hardware, software, networks and data from unauthorized access, vulnerabilities supplied through Internet by cyber criminals, terrorist groups and hackers. Cyber security is related to protecting your internet and network based digital equipments and information from unauthorized access and alteration. One of the most challenging

elements of cyber security is the quickly and constantly evolving nature of security risks. The enterprise network comprised of mainframes, client-server model, closed group of systems and the attacks were very limited with viruses, worms and Trojan horses being the major cyber threats. The focus was more towards malwares such as virus, worms and Trojans with purpose of causing damage to the systems. Cyber threats randomly targeted computers directly connected to the Internet.

Artificial Intelligence methods are robust and more flexible; as a result expanding security execution and better defense system from an increasing number of advance cyber threats.

Different AI techniques can be used in cyber security such as intelligent agent, neural nets, expert system, data mining, machine learning and deep learning.

4 Machine Learning in Cyber Security

Machine learning is an effective tool that can be employed in many areas of information security. There exist some robust anti-phishing algorithms and network intrusion detection systems. Machine learning [Jor 15] can be successfully used for developing authentication systems, evaluating the protocol implementation, assessing the security of human interaction proofs, smart meter data profiling, etc.

Figure 3: Cyber Security

Machine learning [Kan 17] has presented a significant opportunity to the cyber security industry. New machine learning methods can vastly improve the accuracy of threat detection and enhance network visibility thanks to the greater amount of computational analysis they can handle. They are also heralding in a new era of autonomous response, where a machine system is sufficiently intelligent to understand how and when to fight back against in-progress threats.
Different machine learning methods have been successfully deployed to address wide-ranging problems in computer security. We are to discuses three areas where most cyber ML algorithms are finding application: spam detection, malware analysis and intrusion detection.

4.1 Spam and phishing detection

Spam and phishing detection includes a large set of techniques aimed at reducing the waste of time and potential hazard caused by unwanted emails. Nowadays, unsolicited emails, namely phishing, represent the preferred way through which an attacker establishes a first foothold within an enterprise network. Phishing emails include malware or links to compromised websites. Spam and phishing detection is increasingly difficult because of the advanced evasion strategies used by attackers to bypass traditional filters. ML approaches can improve the spam detection process.

Spam filtering based on the textual content of email messages can be seen as a special case of text categorization, with the categories being spam and non-spam. Today the most successful spam filters are based upon the statistical foundations of Machine Learning. Machine Learning based spam filters [Bla 08] also retrain themselves while put in use and minimizes manual effort while delivering superior filtering accuracy.
Although the task of text categorization has been researched extensively, its particular application to email data and detection of spam specifically is relatively recent. Some initial research studies primarily focused on the problem of filtering spam whereby Naïve Bayes (NB) was applied to address the problem of building a personal spam filter. Naive Bayes is a classic machine learning algorithm in which we can use all our feature to detect whether they become malicious file or not and used it for the purpose

of classification. NB was advocated due to its previously demonstrated robustness in the text-classification domain and due to its ability to be easily implemented in a cost-sensitive decision framework. Although high performance levels were achieved using word features only, it was observed that by additionally incorporating non-textual features and some domain knowledge, the filtering performance could be improved significantly.

Phishing is aimed at stealing personal sensitive information. Researchers [Cha 06] have identified three principal groups of anti-phishing methods: detective (monitoring, content filtering, anti-spam), preventive (authentication, patch and change management), and corrective (site takedown, forensics) ones.

4.1.1 E-mail Spam Filtering

Automatic e-mail classification uses statistical approaches or machine learning techniques and aims at building a model or a classifier specifically for the task of filtering spam from a users mail stream. The building of the model or classifier requires a set of pre-classified. The process of building the model is called training. Machine learning algorithms have achieved more success among all previous techniques employed in the task of spam filtering. In fact, the success stories of Gmail, can be ascribed to their timely transition and successful use of Machine Learning for filtering not just incoming spam but other abuses like Denial-of-Service (DoS), virus delivery, and other imaginative attacks.

4.2 Malware detection

Malware detection is an extremely relevant problem because modern malware can automatically generate novel variants with the same malicious effects but appearing as completely different executable files. These polymorphic and metamorphic features defeat traditional rule-based malware identification approaches. Malware can be divided into several classes depending on its purpose: virus, worm, Trojan, adware, spyware, root kit, backdoor, key logger, Ransom ware and Remote Administration Tools.
ML techniques can be used to analyze malware variants and attributing them to the correct malware family.

4.3 Intrusion Detection

An Intrusion Detection System (IDS) is a defense measure that supervises activities of the computer network and reports the malicious activities to the network administrator. Intruders do many attempts to gain access to the network and try to harm the organization's data. Thus the security is the most important aspect for any type of organization. Intrusion detection aims to discover illicit activities within a computer or a network through Intrusion Detection Systems (IDS). Network IDS are widely deployed in modern enterprise networks. These systems were traditionally based on patterns of known attacks, but modern deployments include other approaches for anomaly detection, threat detection [Tor 16] and classification based on machine learning. Within the broader intrusion detection area, two specific problems are relevant to our analysis: the detection of botnets and of Domain Generation Algorithms (DGA). A botnet is a network of infected machines controlled by attackers and misused to conduct multiple illicit activities. Botnet detection aims to identify communications between infected machines within the monitored network and the external command-and-control servers. Despite many research proposals and commercial tools that address this threat, several botnets still exist. DGA automatically generate domain names, and are often used by an infected machine to communicate with external server(s) by periodically generating new hostnames. They represent a real threat for organizations because, through DGA which relies on language processing techniques, it is possible to evade defenses based on static blacklists of domain names.
Network Intrusion Detection (NID) systems are used to identify malicious network activity leading to confidentiality, integrity, or availability violation of the systems in a network. Many intrusion detection systems are specifically based on machine learning techniques [Kha 10] due to their adaptability to new and unknown attacks.

Although machine learning facilitates keeping various systems safe, the machine learning classifiers themselves are vulnerable to malicious attacks. There has been some work directed to improving the effectiveness of machine learning

algorithms and protecting them from diverse attacks.

5 Conclusions

Machine learning approaches are increasingly employed for multiple applications and are being adopted also for cyber security, hence it is important to evaluate when and which category of algorithms can achieve adequate results. We analyze these techniques for three relevant cyber security problems: intrusion detection, malware analysis and spam detection. Machine learning as a technology has erupted vastly in the whole cyber implementation space. These decision making algorithms are known to solve several problems. There are many opportunities in information security to apply machine learning to address various challenges in such complex domain. Spam detection, virus detection, and surveillance camera robbery detection are only some examples. Machine learning techniques have been applied in many areas of science due to their unique properties like adaptability, scalability, and potential to rapidly adjust to new and unknown challenges.

References

[Wan 08] X. B. Wang, G. Y. Yang, Y. C. Li, D. Liu, (2008) *"Review on the application of Artificial Intelligence in Antivirus Detection System"*, IEEE Conference on Cybernetics and Intelligent Systems, pp. 506 509

[Mar 18] Marty, R. *AI and Machine Learning in Cyber Security – Towards Data Science.* March 16, 2018, from https://towardsdatascience.com/ai-and-machine-learning-in-cyber-security Applications of Artificial Intelligence (AI) to Network Security

[Kan 17] Kanal, E. (2017, January). *Machine Learning in Cybersecurity.* Carnegie Mellon University – Software Engineering Institute. March 9, 2018

[Tyu 07] E. Tyugu. *Algorithms and Architectures of Artificial Intelligence.* IOS Press. 2007

[Cha 06] M. Chandrasekaran, K. Narayanan, and S. Upadhyaya, "Phishing Security Conference, 2006

[Jor 15] M. I. Jordan and T. M. Mitchell, *"Machine learning: Trends, perspectives, and prospects,"* Science, 2015

[Bla 08] E. Blanzieri and A. Bryl, *"A survey of learning-based techniques of email spam filtering,"* Artificial Intelligence Review, 2008

[Jav 16] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, *"A deep learning approach for network intrusion detection system,"* in EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS), 2016

[Tzo 07] G. Tzortzis and A. Likas, *"Deep belief networks for spam filtering,"* in IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 2007

[Kha 10] A. Khan, B. Baharudin, L. H. Lee, and K. Khan, *"A review of machine learning algorithms for textdocuments classification,"* Journal of advances in information technology, 2010.

[Tor 16] P. Torres, C. Catania, S. Garcia, and C. G. Garino, "An analysis of Recurrent Neural Networks for Botnet detection behavior," in IEEE Biennial Congress of Argentina (ARGENCON), 2016.