Editorial Introduction to the Special Articles on Applying Deep Learning to Security II

Deep Models, Machine Learning, and Artificial Intelligence Applications in National and International Security — Part Two

Ying Zhao, Arjuna Flenner

■ *The spring 2019 issue of* AI Magazine *featured three articles on deep models, machine learning, and AI applications in national and international security. In this issue, we continue the series with three additional articles, which, like the first three, examine many of the pressing issues involved in applying deep learning to the domain of security.*

Recent advances in artificial intelligence enable new technologies to assist modern warfighters by automatically analyzing big data at time scales much faster than a human can achieve. In particular, deep learning is a core technology in the AI revolution. The deep learning revolution began by demonstrating that not only can machines classify much more quickly than humans but they can classify more accurately as well. These technologies have revolutionized many commercial applications, but they are not currently designed to solve security problems.

Fundamentally, machine learning refers to a subfield of AI in which the parameters of a function are learned from working through a dataset, and deep learning refers to a subfield of machine learning in which the function consists of many layers. These deep networks (convolutional neural networks, for example) often consist of a large number of parameters, and they are trained using labeled data for accurate classification or prediction. Deep learning was initially demonstrated in the breakthrough results for supervised learning in machine vision applications. Because of the classification breakthrough, academic and industrial researchers have increasingly applied AI in the form of deep learning and machine learning to computer vision, speech recognition, chat bots, and autonomous driving. However, many of these applications still lack the robustness and rigor needed for automatic security applications. At best, they are suitable for fast recommendations.

There is a fundamental problem with trust in deep networks. This issue of trust exists not only for the end users but also for the designers of the algorithms. An honest machine learning scientist must reserve confidence in their deep learning networks, because there is no consensus on how or why the deep algorithms obtain the performance that they do. Also, it is simple to find examples that are easily classified by humans but misclassified by deep learning algorithms. Furthermore, it has been demonstrated that a small but visually imperceptible change to a correctly classified image will result in the misclassification of the image. Therefore, there is a fundamental instability in the learned functions. Trust is only one of the major issues in using deep learning for security applications. A second concern is the data requirements. Deep learning algorithms require an extensive amount of training data that can be difficult to obtain. Finally, training the algorithms requires large computational resources and often long time scales for training, which might not be available in time-sensitive security applications.

These issues highlight four of the main challenges in applying the AI revolution to security applications: the lack of adequate samples for classification tasks, short time scales for learning, fewer computational resources, and adversarial behavior.

At a high level, national and international security needs AI in a wide range of forms. Artificial intelligence applications include warfighters' assistants and automation tools, for which trust, ethics, and explainability of AI are very important. Considering that AI can be weaponized by adversaries (for example, as robot fighters, as cyber honeypots, as virtual swarms, and in deceptive games), professionals in this field should research a wide range of deep models. Broadly, these models include all analytic big data models. Given both the current results and the limitations of deep learning, many questions exist with respect to security applications. The special topic articles in this issue address the current state of affairs for many of the pressing issues in applying deep learning to security. Molitor and Needell discuss a simple deep model to better understand the theoretical aspects of deep learning. Dasgupta discusses the problems with corrupted training data in supervised machine learning, especially in the context of deep learning.

Our objective in presenting these articles is to review the current unique security issues in AI and to deepen overall understanding and collaboration in the AI community with respect to the potential, theories, practices, tools, and risks of deep models and AI for security applications, in an effort to remain competitive in technical leadership and innovation in this area.

Ying Zhao is a research professor in the Graduate School of Operational and Information Sciences at the Naval Postgraduate School in Monterey, California.

Arjuna Flenner is a senior research physicist at the Naval Air Systems Command (NAVAIR).