First International Workshop on Dependable and Secure Machine Learning DSML 2018

Homa Alemzadeh[1], Karthik Pattabiraman[2], David E. Evans[1]

[1]University of Virginia, VA, USA - {alemzadeh, evans}@virginia.edu

[2]University of British Columbia, BC, Canada - karthikp@ece.ubc.ca

On behalf of the Organizing Committee, it is our pleasure to welcome you to the first International Workshop on Dependable and Secure Machine Learning (DSML), co-located with the 48th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2018) in Luxembourg City, Luxembourg on Monday, 25 June 2018.

Machine learning is increasingly used in critical domains such as healthcare and medical research, criminal sentencing recommendations, commerce, transportation, employment, entertainment, and communication. The design of machine learning systems has mainly focused on developing models, algorithms, and training datasets to demonstrate high accuracy for specific tasks such as object recognition and classification. Machine learning algorithms typically construct a model by training on a labeled training dataset and their performance is assessed based on the accuracy in predicting labels for unseen (but often similar) testing data. This is based on the assumption that the training dataset is representative of the inputs that the system will face in deployment. However, in practice there are a wide variety of unexpected, and perhaps even adversarially-crafted, inputs that might lead to violations of this assumption. Thus, there are growing concerns regarding the reliability, safety, security, and accountability of machine learning systems.

The DSML workshop provides an open forum for researchers, practitioners, and regulatory experts, to present and discuss innovative ideas and practical techniques and tools for producing trustworthy machine learning systems, even when they need to operate in unpredictable and hostile environments. A major goal of the workshop is to draw the attention of the research community to the problem of establishing guarantees of reliability, security, safety, and robustness for systems that incorporate increasingly complex machine learning models, and to the challenge of determining whether such systems can comply with the requirements for safety-critical systems. A further goal is to foster a research community at the intersection of machine learning and dependable and secure computing.

The workshop features four sessions, including six research papers and two keynote talks. The papers were selected by a program committee, and all papers were reviewed by at least three PC members. The workshop sessions are organized as follows:

The first session will feature a keynote by Suman Jana (Columbia University) on systematic testing and verification of deep learning systems. In the second session, research papers present challenges and novel ideas in designing trustworthy machine learning systems that are fair, transparent, and resilient to perturbations on the training data. The third session presents research focused on attacks and defenses in machine learning models. In the last session, the second keynote will be given by Nirmal R. Saxena (NVIDIA Research) on building resilient computing systems for autonomous driving.

We would like to thank the program committee members for their collective efforts in reviewing the papers and for helping us develop the workshop program. Moreover, we would like to thank the organizers of the DSN conference for their help and support of the DSML workshop and the community for their valued contributions to the workshop.

Program Committee:

Kamalika Chaudhuri, *University of California, San Diego*

Shalini Ghosh, *Stanford Research Institute*

Zbigniew Kalbarczyk, *University of Illinois*

Dong Seong Kim, *University of Canterbury*

Philip Koopman, *Carnegie Mellon University*

Aleksander Mądry, *Massachusetts Institute of Technology*

Cristina Nita-Rotaru. *Northeastern University*

Alina Oprea, *Northeastern University*

Nicolas Papernot, *Penn State University*

Gilles Tredan, *LAAS-CNRS*

Timothy Tsai, *NVIDIA*

Kush Varshney, *IBM Research*

DOI 10.1109/DSN-W.2018.00062

IEEE computer society