# Third International Workshop on Dependable and Secure Machine Learning – DSML 2020

Homa Alemzadeh[1], Rakesh Bobba[2], Varun Chandrasekaran[3],
David E. Evans[1], Nicolas Papernot[4], Karthik Pattabiraman[5], Florian Tramèr[6]

[1]University of Virginia, VA, USA - {alemzadeh, evans}@virginia.edu
[2]Oregon State University, OR, USA - rakesh.bobba@oregonstate.edu
[3]University of Wisconsin-Madison, WI, USA – chandrasekaran@cs.wisc.edu
[4]Vector Institute and University of Toronto, ON, Canada - nicolas@papernot.fr
[5]University of British Columbia, BC, Canada - karthikp@ece.ubc.ca
[6]Stanford University, CA, USA - tramer@stanford.edu

On behalf of the Organizing Committee, it is our pleasure to welcome you to the third International Workshop on Dependable and Secure Machine Learning (DSML). This year, due to the COVID-19 situation, the DSML workshop will be held online, in conjunction with the 50th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN) on Monday, 29 June 2020.

Machine learning (ML) is increasingly used in critical domains such as health and wellness, criminal sentencing recommendations, commerce, transportation, human capital management, entertainment, and communication. The design of ML systems has mainly focused on developing models, algorithms, and datasets on which they are trained to demonstrate high accuracy for specific tasks such as object or speech recognition and classification. ML 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 accidental, as well as adversarially-crafted, perturbations on the ML inputs that might lead to violations of this assumption. ML algorithms are also often over-confident about their predictions when processing such unexpected inputs. This makes it difficult to deploy them in safety critical settings where one needs to be able to rely on the ML predictions to make decisions or revert back to a failsafe mode. Further, ML algorithms are often executed on special-purpose hardware accelerators, which may themselves be subject to faults. Thus, there is a growing concern regarding the reliability, safety, security, and accountability of ML systems.

The DSML workshop is an open forum for researchers, practitioners, and regulatory experts, to present and discuss innovative ideas and practical techniques and tools for producing dependable and secure ML systems. 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 ML models, and to the challenge of determining whether such systems can comply with requirements for safety-critical systems. A further goal is to build a research community at the intersection of machine learning and dependable and secure computing.

The workshop features four sessions, including research papers and two keynote talks. We received 10 regular paper submissions this year, of which we accepted 6 papers. The papers were selected by the program committee based on reviews and online discussion – each paper was reviewed by three or four PC members. The workshop sessions are organized as follows:

The first session will feature a keynote by Michael Lyu from Chinese University of Hong Kong on Interpretability-Driven Dependable and Secure Machine Learning. In the second session, research papers present different attacks against ML systems. The third session presents research papers on validation, verification, and defense mechanisms that make ML systems robust against faults and attacks. In the last session, the second keynote will be given by Rajarshi Gupta from Avast Security on Using Secure AI to Secure Real Users.

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:**

Saurabh Bagchi, Purdue University
Lorenzo Cavallaro, King's College London
Earlence Fernandes, University of Wisconsin-Madison
Ilir Gashi, City University, London
Neil Gong, Duke University
Trinabh Gupta, University of California Santa Barbara
Nirupam Gupta, Georgetown University
Johan Karlsson, Chalmers University
Dong Seong Kim, University of Queensland, Australia
Guanpeng Li, University of Iowa
Fabio Pierazzi, King's College London
Tim Tsai, Nvidia
Devesh Tiwari, Northeastern University