Summary of iMLSE-18: The 1st International Workshop on Machine Learning Systems Engineering

Fuyuki Ishikawa
National Institute of Informatics
Tokyo, Japan
f-ishikawa@nii.ac.jp

Foutse Khomh
Polytechnique Montreal
Montrèal, Canada
foutse.khomh@polymtl.ca

Nobukazu Yoshioka
National Institute of Informatics
Tokyo, Japan
nobukazu@nii.ac.jp

Giuliano Antoniol
Polytechnique Montreal
Montrèal, Canada
antoniol@ieee.org

***Abstract*—This paper summarizes the objectives and results of iMLSE-18: The 1st International Workshop on Machine Learning Systems Engineering held on December 4th in Nara, Japan. The workshop was collocated with APSEC 2018.**

Keywords—Machine Learning, Deep Learning, Software Engineering, Software 2.0

I. Background and Objectives

This workshop aimed to bring together leading software engineers, machine learning experts and practitioners to reflect on and discuss the challenges and implications of building software for complex Artificial Intelligence (AI) systems by using Machine Learning (ML) techniques.

The core idea behind this workshop is a growing concern that we have as software engineers in a world where data science, deep learning, and AI are becoming increasingly pervasive. The economic benefits of Machine-Learning Software Applications and artificial intelligence, in general, is forecast to surpass USD 8.81 Billion by 2022[1]. Although AI research has allowed the development of novel algorithms capable of learning new tasks, adapting to the environment, and evolving, their implementation in software systems remains challenging. From an engineering perspective, once an algorithm is implemented, it requires a solid architecture, model/data validation, proper monitoring for changes, dedicated release engineering strategies, judicious adoption of design patterns and security checks, and thorough user experience evaluation and adjustment. All these activities require a combined knowledge in software engineering, data science, and machine learning. A failure to properly address these challenges in such complex software systems can lead to catastrophic consequences. An example of such failure is the recent human toll incidence caused by the $47-million Michigan Integrated Data Automated System (MiDAS)[2], recent finding that simple tweaks can fool neural networks in identifying street signs[3], or the Uber's self-driving car that ran into a pedestrian even though the car's sensors detected her presence. The software of the Uber's car which is a Machine-Learning Software Application reportedly decided not to react right away, considering the detection of the pedestrian as a "false positive."[4]

The source of emerging difficulties is the shift of the development paradigm. Classically, we have constructed software systems in a deductive way, or by writing down the rules that govern the system behavior as program code. With machine learning techniques, we generate such rules in an inductive way from training data. This shift does not only simply require new tools that intensively deal with data but also introduces unique characteristics. The resulting system behaviors are uncertain: black-box and unexplainable. They are intrinsically imperfect and it is practically impossible to reason about their correctness in a deductive way.

Given the critical and increasing role of AI-based systems in our society it is now imperative to engage all stakeholders (e.g., software engineers, machine learning experts and decision makers) in in-depth conversations about the necessary perspectives, approaches, and roadmaps to address these challenges and concerns.

II. Program

The workshop started with general introduction of the background and the focus of the workshop. Specifically, two initiatives of MLSE from Japan[5] and SEMLA from Canada[6] were reported. It was discussed how engineering of ML-based systems is different and challenging compared with that of classical software systems.

There were three paper submissions by the due date. The program committee conducted a rigorous peer review by assigning at least three reviewers to each submission. The workshop organizers finally selected the following two papers for presentation and inclusion into the proceedings.

- "Simplified Influence Evaluation of Additional Training on Deep Neural Networks", Naoto Sato, Hironobu Kuruma, Yuichiroh Nakagawa and Hideto Ogawa.

[1] https://www.marketsandmarkets.com/PressReleases/machine-learnin g.asp
[2] https://www.bridgemi.com/public-sector/broken-human-toll-michigans-unemployment-fraud-saga
[3] https://iotsecurity.eecs.umich.edu/#roadsigns
[4] https://www.theguardian.com/technology/2018/may/08/ubers-self-driving-car-saw-the-pedestrian-but-didnt-swerve-report
[5] https://sites.google.com/view/sig-mlse/ (in Japanese)
[6] http://semla.polymtl.ca/

- "Guaranteeing Deep Neural Network Outputs in a Feasible Region", Hiroshi Maruyama.

In the session of technical talks, we had these two research papers and the following two position talks.

- "Dataflow Visualization using ASCII DAG", Junji Hashimoto.
- "Toward New Definitions of Equivalence in Verifying Deep Learning Compilers", Takeo Imai.

Questions on each paper led to discussions from a wide viewpoint not limited to the specific focus of the paper. Many of the audience were new to the area of the workshop and this point led to essential discussions from a general viewpoint. For example, we had intensive discussions on why the existing approaches we already have for classical software systems cannot be applied, or the exact boundary of what we can do and what we cannot do.

We welcomed a great invited talk by Professor Jianjun Zhao (Kyushu University) entitled “Towards Testing of Deep Learning Systems” (see Figure 1). Although the topic of testing deep learning systems is very new, the group of Prof. Zhao has already published impactful research results at top venues of software engineering and reliability. His talk again led to essential discussions on testing.

We finally had the discussion session. We had problem statements from three industry persons.

- Susumu Tokumoto (Fujitsu)
- Hideto Ogaawa (Hitachi),
- Hirohsi Maruyama (PFN)

Figure 1. An invited talk by Professor Jianjun Zhao

They provided very insightful questions and visions about testing, attitudes of research communities, quality assurance activities in the industry, and future directions. Given these inputs and the good atmosphere made in the previous sessions, we could naturally continue essential discussions on various aspects of engineering for ML-based systems.

III. Conclusions and Future Directions

We had very fruitful discussions at the workshop on the new area, engineering of ML-based systems. It turned out that we needed to start with exchange understanding and visions of each participant as there is no common consensus on various aspects of the area: for example, how it is different from the classical software engineering and what is (im)possible due to the nature of ML (e.g., black-box implementation with deep neural networks). We plan to continuously provide venues for discussions on this new but very significant area.