



سخنرانی علمی

پژوهشگاه دانش‌های بنیادی
پژوهشکده علوم کامپیوتر

Transfer Learning for Performance Analysis of Machine Learning Systems

By: Dr. Pooyan Jamshidi
University of South Carolina

Abstract

A wide range of modern machine learning systems (e.g., autonomous systems, big data analytics, robotics, deep neural architectures) are built configurable. These systems offer a rich space for adaptation to different domains and tasks. Developers and users often need to reason about the performance of such systems, making tradeoffs to change specific quality attributes or detecting performance anomalies. For instance, developers of image recognition mobile apps are not only interested in learning which deep neural architectures are accurate enough to classify their images correctly, but also which architectures consume the least power on the mobile devices on which they are deployed. Recent research has focused on models built from performance measurements obtained by instrumenting the system. However, the fundamental problem is that the learning techniques for building a reliable performance model do not scale well, simply because the configuration space is exponentially large that is impossible to exhaustively explore. For example, it will take over 60 years to explore the whole configuration space of a system with 25 binary options.

In this talk, I will start motivating the configuration space explosion problem based on my previous experience with large-scale big data systems in industry. I will then present our transfer learning solution to tackle the scalability challenge: instead of taking the measurements from the real system, we learn the performance model using samples from cheap sources, such as simulators that approximate the performance of the real system, with a fair fidelity and at a low cost. Results show that despite the high cost of measurement on the real system, learning performance models can become surprisingly cheap as long as certain properties are reused across environments. In the second half of the talk, I will present empirical evidence, which lays a foundation for a theory explaining why and when transfer learning works by showing the similarities of performance behavior across environments. I will present observations of environmental changes and impacts (such as changes to hardware, workload, and software versions) for a selected set of configurable systems from different domains to identify the key elements that can be exploited for transfer learning. These observations demonstrate a promising path for building efficient, reliable, and dependable machine learning systems. Finally, I will share my research vision for the next five years and outline my immediate plans to further explore the opportunities of transfer learning.

Biography

Pooyan Jamshidi is an Assistant Professor at the University of South Carolina. He directs the AISys Lab, where he investigates the development of novel algorithmic and theoretically principled methods for machine learning systems. Prior to his current position, he was a research associate at Carnegie Mellon University and Imperial College London, where he primarily worked on transfer learning for performance understanding of highly-configurable systems including robotics and big data systems. Pooyan's general research interests are at the intersection of systems/software and machine learning. He received his Ph.D. in Computer Science at Dublin City University in 2014, and M.S. (Industrial Engineering) and B.S. (Computer Science and Math) degrees from the Amirkabir University of Technology in 2003 and 2006 respectively. For more information, please refer to <https://pooyanjamshidi.github.io/>.

زمان: سه‌شنبه ۹۹/۲/۳۰ - ساعت ۱۷

مکان: سخنرانی مجازی از طریق سامانه vmeeting