

## **Neural Network Verification: Recent Advances and Challenges**

Neural networks have made tremendous advances in solving prediction tasks in domains like speech, language and vision. However, there remain significant failure modes in modern neural networks, particularly when the environment in which a neural network is deployed can differ significantly from the environment in which it was trained. Formal verification seeks to remedy this by providing formal guarantees on the input-output properties of neural networks. Obtaining formal verification algorithms that are both tight (i.e., able to prove properties whenever they are true) and scalable has remained a challenge. I will describe several advances we have made in neural network verification algorithms that take important steps towards this goal, and demonstrate results proving that these new algorithms are able to verify properties of networks previously considered out of reach. Based on joint work with several members of the Robust and Verified AI team at DeepMind, as well as external collaborators from Google and CMU.