Positive Bias and Length Bias in Deep Neural Networks for Premises Selection

Authors : Jiaqi Huang, Yuheng Wang

Abstract : Premises selection, the task of selecting a set of axioms for proving a given conjecture, is a major bottleneck in automated theorem proving. An array of deep-learning-based methods has been established for premises selection, but a perfect performance remains challenging. Our study examines the inaccuracy of deep neural networks in premises selection. Through training network models using encoded conjecture and axiom pairs from the Mizar Mathematical Library, two potential biases are found: the network models classify more premises as necessary than unnecessary, referred to as the 'positive bias', and the network models perform better in proving conjectures that paired with more axioms, referred to as 'length bias'. The 'positive bias' and 'length bias' discovered could inform the limitation of existing deep neural networks.

Keywords : automated theorem proving, premises selection, deep learning, interpreting deep learning

Conference Title : ICATPPC 2019 : International Conference on Automated Theorem Proving and Proof Checking

Conference Location : Singapore, Singapore

Conference Dates : July 04-05, 2019