Automated machine learning for neural network verification

König, H.M.T.

Propositions

accompanying the dissertation

Automated Machine Learning for Neural Network Verification

1. The state of the art in neural network verification is not defined by a single algorithm but by several algorithms with their own strengths and weaknesses. (Chapter 3)
2. Formal verification algorithms relying on mixed integer programming solvers are sensitive to the setting of their hyper-parameters and their performance can be greatly improved by leveraging automated configuration and portfolio construction techniques. (Chapter 4)
3. Verification problem instances cannot always be solved within a reasonable time budget but it is possible to reliably detect these cases early in the verification procedure. (Chapter 5)
4. During robust model selection, the likelihood of a given instance to be robust or non-robust can be estimated to guide the search towards neural network models that are most likely to show the highest adversarial robustness. (Chapter 6)
5. In theory, formal verification can provide rigorous safety guarantees for neural network models that go well beyond empirical evaluation.
6. Formal verification still remains computationally challenging and, therefore, difficult to scale to larger models or data sets.
7. Most research considers robustness properties that do not capture realistic perturbations, which might be hard to encode.
8. The most scalable verification algorithms rely on an approximation, which typically has a negative impact on precision.
9. Neural networks are powerful, but brittle learners.
10. The capabilities of AI systems are often overstated, which offers a great risk to the reputation of the field.

H.M.T. König
Leiden, 9$^{\text{th}}$ of October, 2025