Karlsruher Institut für Technologie

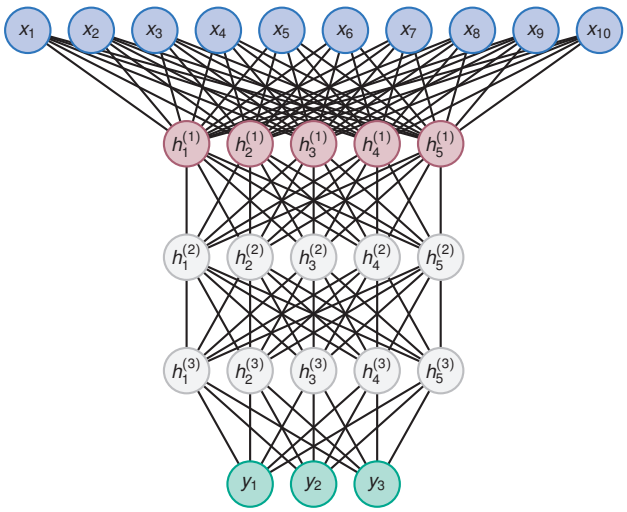
Application-oriented Formal Verification Research Group
Institute of Information Security and Dependability (KASTEL)
Am Fasanengarten 5
76131 Karlsruhe
https://formal.kastel.kit.edu

Vorverarbeitung durch Dimensionsreduktion für die Verifikation von neuronalen Netzwerken

Bachelorarbeit

Hintergrund Auf Grund der zunehmenden Nutzung von neuronalen Netzwerken in sicherheitskritischen Anwendungen (bspw. im Rahmen von KI-gestützter Flug-Kollisionsvermeidung [2]) wurden in den letzten Jahren eine ganze Reihe von Tools zur automatischen Analyse und Verifikation von neuronalen Netzwerken (NN) entwickelt. Hierbei handelt es sich inzwischen um ein eigenes Forschungsfeld mit einem jährlich stattfindenden Wettbewerb für Verifikationstools [1]. Eine maßgebliche Einflussgröße für die Effizienz von NN-Verifikationstools ist die Dimension des Eingaberaums [4].

Aufgabenstellung Im Rahmen dieses Projekts soll ein Vorverarbeitungsschritt für NN-Verifikation entwickelt und evaluiert werden, der die Dimension des analysierten Eingaberaums reduziert. Nach eigenen Literaturrecherechen zu Verifikationstechniken für neuronale Netzwerke, besteht die Aufgabe des Projekts in der Implementierung und Evaluation der Vorverarbeitung. Hierfür gibt es einen bereist erdachten Ansatz auf Basis von Generalized Star Sets [3], der ggf. im Rahmen der Bachelorarbeit auch weiterentwickelt werden kann – insbesondere an dieser Stelle gibt es viele Möglichkeiten eigene Ideen einzubringen. Die Implementierung ist unabhängig von bestehenden Verifikationstools als Vorverarbeitungsschritt gedacht und sollte idealerweise für verschiedene bestehende Tools nutzbar gemacht werden.



Stichworte: Linear Programming, Polytope, Singular Value Decomposition, Abstract Interpretation, Geometric Path Enumeration, NP-vollständige Probleme

Referenzen

[1] Stanley Bak et al. “The second international verification of neural networks competition (vnn-comp 2021): Summary and results”. In: *arXiv preprint arXiv:2109.00498* (2021).

[2] Kyle D Julian et al. “Policy compression for aircraft collision avoidance systems”. In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE. 2016, pp. 1–10.

[3] Hoang-Dung Tran et al. “Star-based reachability analysis of deep neural networks”. In: *International symposium on formal methods*. Springer. 2019, pp. 670–686.

[4] Hoang-Dung Tran et al. “Verification of deep convolutional neural networks using imagestars”. In: *International conference on computer aided verification*. Springer. 2020, pp. 18–42.

Betreuung:

Samuel Teuber, teuber@kit.edu, Room 203 (Geb. 50.34)
Philipp Kern, philipp.kern@kit.edu, Room 203 (Geb. 50.34)