

0.1. Полевой А. Об одном подходе к верификации нейросетевых моделей шумоподавления

Программное обеспечение с моделями машинного обучения получает широкое распространение и внедряется во многие процессы с критическими свойствами, невыполнение которых приводит к серьёзным проблемам. Традиционная проверка качества моделей на ограниченных наборах данных позволяет оценить процент правильных результатов, однако не может гарантировать выполнения заданного свойства. Задача верификации – привести строгое доказательство, что требуемые свойства выполняются.

Традиционные подходы к верификации являются вычислительно затратными, поэтому разработка перспективных методов верификации нейросетевых моделей – большой вызов для современной науки. В рамках работы рассматриваются подходы к анализу надёжности нейросетевых моделей, которые учитывают архитектуру моделей. Предложен подход к верификации, реализованный на основе решения задачи удовлетворения ограничений средствами языка Пролог. Работа метода продемонстрирована на предложенной ранее автором модели шумоподавления [1], которая позволяет эффективно обрабатывать нестационарные сигналы. В рамках работы доказана выполнимость ряда свойств нейросетевой модели обработки сигналов и проведено сравнение с работой известных систем: построен программный стенд, на котором были проверены свойства нейросети в сравнении с системой Marabou [2]. Сделаны выводы о применимости методов и даны оценки по времени и по памяти для рассматриваемых моделей.

Для сравнения было рассмотрено доказательство свойств той же модели с помощью системы Marabou:

1. Существующий подход Marabou и предложенный эффективно справились с предлагаемыми свойствами, без заикливания.
2. Поскольку реализация Marabou выполнена на C++, было показано стабильное время работы для рассмотренных свойств.
3. В предлагаемом подходе за счет использования Prolog и понижения точности удалось существенно сократить время проверки свойств по сравнению с реализацией Marabou. Также стоит отметить, что предложенная реализация использует существенно меньше памяти при проверке невыполнимого свойства.

Дальнейшее исследование связано с обобщением проверки свойств на новые типы сигналов и расширением класса нейросетевых моделей, с которыми работает верификатор, в частности моделей активного шумоподавления.

Научный руководитель — к.ф.-м.н. Корухова Ю. С.

Список литературы

- [1] Полевой А. В., Корухова Ю. С. Исследование алгоритмов активного шумоподавления для нестационарных сигналов // Тезисы докладов научной конференции "Тихоновские чтения" — М.: ООО "МАКС Пресс 2022. — С. 51.
- [2] KATZ G., ET AL. The Marabou framework for verification and analysis of deep neural networks. // Computer Aided Verification: 31st International Conference, CAV 2019, New York City, NY, USA, July 15-18, 2019, Proceedings, Part I 31: Springer International Publishing, 2019. P. 443-452.