Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)
Doctoral Consortium Track

Deep Learning with Requirements in the Real World

Mihaela Cătălina Stoian
University of Oxford
mihaela.stoian@cs.ox.ac.uk

Abstract

Deep learning models have repeatedly shown their strengths in various application domains. However, their predictions often struggle to meet background knowledge requirements, which is a crucial condition for safety-critical systems. My research focuses on integrating requirements into neural networks to guide the learning process and ultimately produce outputs that ensure the requirements' satisfaction. Here, I will discuss my proposed methods in the context of two real-world applications: tabular data generation and autonomous driving.

1 Introduction

Deep learning models have shown their strengths in various application domains, however, they often struggle to meet safety requirements for their outputs. Neuro-symbolic AI aims at addressing this issue by interlinking neural networks (NNs) with symbolic reasoning. Such existing methods can be broadly classified into two categories. The first comprises methods able to integrate the requirements in the loss function and penalize the models when they violate the requirements (e.g., [Diligenti *et al.*, 2012; Xu *et al.*, 2018; Badreddine *et al.*, 2022]), while the second category contains methods able to incorporate requirements (also called constraints) directly in the topology of the network (e.g., [Giunchiglia and Lukasiewicz, 2021; Hoernle *et al.*, 2022; Ahmed *et al.*, 2022]) and, thus, guarantee their satisfaction. Here, I will discuss my work on developing such methods for two real-world application domains: synthesizing tabular data and autonomous driving.

2 Deep Generative Modeling for Tabular Data with Requirements

From enhancing predictive performance in ML models to ensuring privacy in sensitive settings, synthesizing methods are increasingly used in real-world domains (e.g., generating health records or credit scoring data). Often, there are inherent relations between tabular data features that synthetic samples must satisfy to be considered realistic. For example, in a dataset containing medical records for patients with diabetes, suppose we have two features capturing the minimum and the maximum recorded haemoglobin levels. Naturally, the real data do not contain any records where the minimum level is higher than the maximum level. This is a background knowledge (BK) requirement easily captured by the linear constraint $\texttt{MaxHaemoglobin} - \texttt{MinHaemoglobin} \geq 0$.

Research questions. My main research questions are:

- **R1** How often do standard models violate such constraints?
- **R2** Does BK improve the quality of the synthetic data?
- **R3** Can BK injection act as a guardrail during inference and ensure that the constraints are satisfied?

Proposed method. Following the branch of neuro-symbolic works that embed requirements into the NNs' topology, my proposed approach [Stoian *et al.*, 2024] consists of building a constraint layer (CL) on top of a given deep generative model (DGM). The result is a Constrained DGM (C-DGM), which guarantees that the constraints are satisfied by the generated synthetic data. The CL automatically compiles the BK expressed as linear constraints and corrects the model's predictions. An extensive experimental analysis using five DGM types and six real-world datasets reveals that standard DGMs often violate constraints, some exceeding 95% non-compliance (**R1**), while their corresponding C-DGMs are never non-compliant. Additionally, at training time, C-DGMs are able to exploit the BK to outperform their baseline counterparts over two standard measures, utility and detection, with up to 6.5% improvement in the utility F1-score. Furthermore, the CL can also be used only at inference time, acting as a guardrail (**R3**), but still producing some improvements in the overall performance of the models. For a qualitative assessment, Figure 1 illustrates how the CL-based model not only ensures that the example constraint above is satisfied, but also that the generated samples better match the real data relatively to the baseline model's samples. Finally, and particularly relevant for real-world scenarios, the experiments show that the CL does not hinder the sample generation time.

3 Autonomous Driving with Requirements

Neural networks have been at the core of the recent developments in the autonomous driving field. However, standard models are data-driven and can lead to unpredictable behaviors. For instance, a multi-label classification model for the vision component of a self-driving system might erroneously

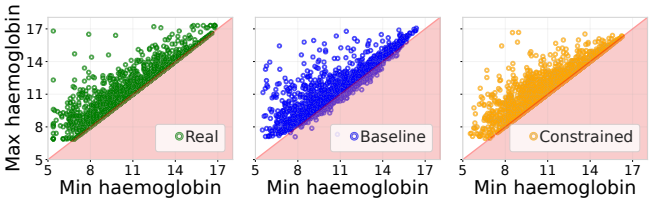

Figure 1: Real data (left) and samples generated by a Baseline DGM (middle) and a Constrained DGM model using CL (right).

classify a traffic light as both red and green at the same time. Naturally, such models can easily lead to catastrophic outcomes. Aiming at mitigating this problem, injecting propositional logic constraints via a t-norm-based loss term is a popular neuro-symbolic method [Giannini *et al.*, 2023], which allows NNs to have access to BK during training and learn to adjust their predictions accordingly. However, t-norm-based losses may have very high memory costs and may be impossible to apply in complex domains like autonomous driving.

Research questions. My main research questions are:

Q1 How many constraints do standard methods support?

Q2 Is it possible to design a memory-efficient method to account for a large number of constraints?

Q3 Does BK improve the quality of the outputs?

Q4 Does BK help in a semi-supervised setting?

Proposed method. My approach [Stoian *et al.*, 2023] formalizes memory-efficient t-norm-based losses, allowing for exploiting t-norms for event detection in autonomous driving. It relies on the intuition that in practice most of the constraints are written over a small subset of the available labels. As a result, the matrices encoding the requirements contain mostly zeros. Hence, my proposed method makes use of this sparsity property and ultimately avoids the high computational costs induced by the 3D matrices, operating only on 2D matrices (**Q2**). Figure 2 demonstrates a significant reduction in memory costs, enabling the utilization of t-norm-based losses on the large-scale ROAD-R autonomous driving dataset [Giunchiglia *et al.*, 2023]. The standard implementation supports a maximum of 40 constraints (**Q1**), falling short by 203 constraints compared to ROAD-R. Moreover, my proposed approach allows model training on GPUs with 25 GiB memory, whereas the standard method would require over 100 GiB for ROAD-R, far exceeding typical available memory capacities. Furthermore, through an extensive experimental analysis on the ROAD-R dataset in two of my works [Stoian *et al.*, 2023; Giunchiglia *et al.*, 2023], I show that t-norm-based losses improve the models' performance (**Q3**), especially for limited labelled data. Finally, the experiments show t-norm-based losses can further improve performance when exploited on labelled and unlabelled data (**Q4**).

4 Conclusion and Future Work

My research follows the principles of neuro-symbolic integration, with a particular focus on making neuro-symbolic AI more accessible for real-world applications. As future work, I am planning on bridging the two approaches above to support both propositional and linear constraints.

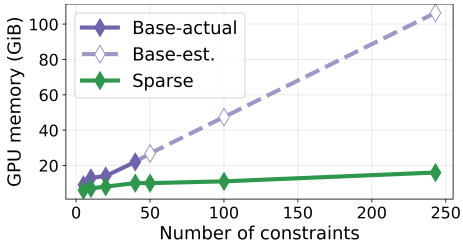

Figure 2: Comparison between the standard approach (in purple) and my proposed memory-efficient approach using sparse representations (in green) in terms of GPU memory allocated when using different numbers of constraints. Each point on the continuous (resp., dashed) lines corresponds to an actual observation (resp., estimate).

Acknowledgments

Mihaela Stoian is supported by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/T517811/1. I would like to express my gratitude to Prof. Thomas Lukasiewicz and Dr. Eleonora Giunchiglia for their guidance.

References

[Ahmed *et al.*, 2022] K. Ahmed, S. Teso, K. Chang, G. V. den Broeck, and A. Vergari. Semantic probabilistic layers for neuro-symbolic learning. In *Proc. of NeurIPS*, 2022.

[Badreddine *et al.*, 2022] S. Badreddine, A. d'Avila Garcez, L. Serafini, and M. Spranger. Logic tensor networks. *Artif. Intell.*, 303, 2022.

[Diligenti *et al.*, 2012] M. Diligenti, M. Gori, M. Maggini, and L. Rigutini. Bridging logic and kernel machines. *Mach. Learn.*, 86, 2012.

[Giannini *et al.*, 2023] F. Giannini, M. Diligenti, M. Maggini, M. Gori, and G. Marra. T-norms driven loss functions for machine learning. *Appl. Intell.*, 53, 2023.

[Giunchiglia and Lukasiewicz, 2021] E. Giunchiglia and T. Lukasiewicz. Multi-label classification neural networks with hard logical constraints. *JAIR*, 72, 2021.

[Giunchiglia *et al.*, 2023] E. Giunchiglia, M. C. Stoian, S. Khan, F. Cuzzolin, and T. Lukasiewicz. ROAD-R: The autonomous driving dataset for learning with requirements. *Mach. Learn.*, 2023.

[Hoernle *et al.*, 2022] N. Hoernle, R. Karampatsis, V. Belle, and K. Gal. MultiplexNet: Towards fully satisfied logical constraints in neural networks. In *Proc. of AAAI*, 2022.

[Stoian *et al.*, 2023] M. C. Stoian, E. Giunchiglia, and T. Lukasiewicz. Exploiting t-norms for deep learning in autonomous driving. In *Proc. of International Workshop on Neural-Symbolic Learning and Reasoning*, 2023.

[Stoian *et al.*, 2024] M. C. Stoian, S. Dyrmishi, M. Cordy, T. Lukasiewicz, and E. Giunchiglia. How realistic is your synthetic data? Constraining deep generative models for tabular data. In *Proc. of ICLR*, 2024.

[Xu *et al.*, 2018] J. Xu, Z. Zhang, T. Friedman, Y. Liang, and G. V. den Broeck. A semantic loss function for deep learning with symbolic knowledge. In *Proc. of ICML*, 2018.