Master's Thesis SS 2021

Physics-Informed Neural Networks for Structural Conrete Engineering

Supervisors:
Dr. Michael A. Kraus
Prof. Dr. Olga Sorkine-Hornung
Prof. Dr. Walter Kaufmann

Neural Networks are universal function approximators and have shown impressive results on a variety of different tasks, especially when problems involve unstructured data and an abundant amount of training data. However, collecting a dataset of sufficient quality and size in order to train a reliable Neural Network can be prohibitively expensive or even impossible in some fields.

Physics-Informed Neural Networks leverage well known physical laws, generally in the form of partial differential equations (PDE), by including them into their loss function. This prior knowledge guides and facilitates the training process, enabling the model to quickly converge, even when little to no training examples are available. Their continuity, flexibility and the perspective of solving arbitrary instances of a given PDE in a single forward pass, might turn them into a valuable alternative to established, mesh-based numerical techniques like the finite elements method.

This thesis is concerned with making PINNs more reliable and studying the effectiveness of various state-of-the-art extensions to vanilla PINNs on the Kirchhoff plate bending problem. In a first step, we address the issue of gradient pathologies arising because of terms with different units of measurements and thus of varying magnitudes in the loss function . To this end, we compare the performance of three state-of-the-art Adaptive Loss Balancing algorithms, Learning Rate Annealing, GradNorm and SoftAdapt. Furthermore, we propose our own adaptation of the aforementioned methods, Relative Loss Balancing with Random Lookbacks (ReLoBRaLo). To assess their generalisability, we test the four algorithms on three benchmark problems for PINNs: the Helmholtz, Burgers and Kirchhoff PDEs. The results show that the scaling heuristics consistently outperform the baseline with manually selected scalings, with our method being particularly effective at solving Kirchhoff's equation. Additionally, we show that the scalings generated by ReLoBRaLo provide valuable insight into the training process and can be used to take informed decisions in order to improve the framework.

In a second step, we conduct an extensive parameter study on Kirchhoff's plate bending equation, where we test several network architectures, activation functions, sampling techniques and loss balancing algorithms in order to consolidate knowledge about this problem. We analyse the results in perspective with important properties from structural engineering, highlight hyperparameters and extensions which consistently improve the model's performance and identify a number of trade-offs that need to be addressed on a per-task basis. On many problems, our frameworks are able to achieve promising results, often exceeding the threshold of minimum accuracy for use in practice. However, we also identify important limitations and failure modes that need to be addressed before PINNs can be used in real-world applications.

PINNs are currently most often used as so-called representation networks, meaning that they represent parameterised PDEs with given parameters, boundary- and initial conditions. This is a major limiting factor, as any slight change of those parameters requires the network to be retrained, making their use computationally expensive. In a last step of this thesis, we therefore propose the use of hypernetworks. Hypernetworks are used to translate an embedding of context variables to weights and biases of another, main network. In the context of PDEs, such a hypernetwork receives the boundary conditions as well as PDE parameters as input and generates the parameters for the corresponding PINN. We demonstrate on Helmholtz' and Kirchhoff's equations that hypernetworks can be used to solve these tasks and indeed produce qualitatively high results in a single forward pass, therefore turning PINNs extremely valuable for fast prototyping in many practical applications. Furthermore, they represent a differentiable end-to-end solution, which could prove useful for Design Optimisation and Reverse Engineering.