

Prof. R. Wattenhofer

## Exploring Activation Ensembles for Neural Networks

The core of Deep Neural Networks (DNNs) lies in their activations, which play a pivotal role in defining their internal functioning. Although numerous activation functions are available, researchers and practitioners often lean towards well-established options, such as the Rectified Linear Unit (ReLU), known for its efficacy. This preference is partially justified by the Universal Approximation Theorem, which asserts that, under certain assumptions, a neural network can approximate any function. Yet, it remains uncertain whether alternative activations could lead to simpler and more compact architectures for specific problems. In this thesis, our objective is to empower neural networks to employ diverse activation functions in each layer, thereby identifying the most suitable ones for a given problem.

Our research builds upon prior findings, as seen in works like this. Initially, we will perform empirical evaluations using standard benchmarks to gain a deeper understanding of the impact of various activations. This analysis will encompass different neural network architectures, including models designed for images, text, and graphs. Subsequently, we aim to develop novel architectures that have the potential to surpass existing ones in terms of performance.

**Requirements:** Strong motivation, knowledge in neural networks and machine learning, as well as good coding skills. Prior practical experience with neural networks is a big advantage. We will have weekly meetings to discuss open questions and determine the next steps.



## Interested? Please contact us for more details!

## Contact

In a few short sentences, please tell us why you are interested in the project and about your coding and machine learning background (i.e., your own projects or courses).

- Florian Grötschla: fgroetschla@ethz.ch, ETZ G63
- Joël Mathys: jmathys@ethz.ch, ETZ G63