ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Master Thesis

Counterfactual Learning of Recurrent Neural Networks

Recurrent Neural Networks (RNN) are a general computational model that has shown to be effective in many different kinds of tasks. Additionally, it is a model that neuroscientist have been looking at as a (very abstract) model of the what the brain may be doing [1].

However, current techniques (Truncated Backpropagation Through Time) are inherently limited to just being able to learn short-term time dependencies. This is because of two reasons. First, the memory required by the algorithm grows linearly with the length of the sequence that has to be learned, making it impractical for really long sequences [2]. Second, the exploding/vanishing gradients problems makes gradient information useless over long time sequences [3].

In this thesis you will explore alternative learning techniques that don't use gradients to learn over time. For this we will look into an RNN architecture that can make a discrete decision to store information into an external memory. In order to learn, different information is added/removed at random from the external memory. By evaluating the difference in performance that this addittion/removal makes, the RNN should be able to learn which kind of information should be added to the external memory and which can be safely ignored. The thesis will involve, implementing an initial version of the algorithm, analyzing and understanding its behaviour, and exploring how it can be improved.

More information and grading scheme can be found on:
`https://www.cadmo.ethz.ch/education/thesis/guidelines.html`

Prerequisites: Knowledge of Tensorflow or Pytorch

Supervisor: Asier Mujika, CAB J 21.2, `asierm@inf.ethz.ch`

Supervising Professor: Prof. Dr. Angelika Steger, CAB G 37.2, `steger@inf.ethz.ch`

References

[1] Omri Barak. Recurrent neural networks as versatile tools of neuroscience research. *Current opinion in neurobiology*, 2017.

[2] Asier Mujika, Florian Meier and Angelika Steger. Approximating real-time recurrent learning with random Kronecker factors. *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[3] Razvan Pascanu, Tomas Mikolov and Yoshua Bengio. Understanding the exploding gradient problem. *ArXiV*, 2012.