Course: Brain Inspired AI **Prof. Yalda Mohsenzade**
Name: Tina Gholami (251190343) **Assignment 2**

Paper: Long Short-Term Memory

This paper investigates into long short-term memory cells (LSTMs). It first gives an introduction on recurrent networks trying to store information in only "short-term memories" but do not actually perform well, especially when there are minimal time lags between inputs and the corresponding teacher signals. Therefore, error signals flowing backwards in time tend to either blow up or vanish which results in oscillating weights and great run-time, respectively. Therefore, as a remedy, "long short-term memories" are introduced. Next, previous works on similar problem is explored. Methods such as Gradient-descent variants, Time-delays, Time constants, Ring's approach, Bengio et al.'s approaches, Kalman flters, Second-order nets, Simple weight guessing, and Adaptive sequence chunkers is being introduced, nut none of them show as good results as LSTMs. Then, some background math is shown on exponentially decaying error and constant error flow. Then both of these notions are used to bring up the idea of LSTMs, and how their architecture can help solve the issue presented earlier as the problem. Next, six different experiments are conducted to test LSTMs in regard to facing noise, complexity, time-lag problems, and recurrent nets. In all experiments, LSTMs showed a significantly better performance compared to previous approaches. LSTMs come with both advantages and disadvantages. They are able to bridge very long time lags, handle noise, generalize well, have low time-complexity O(1), and also they do not require parameter fine tuning. However, each memory cell block needs two additional units (input and output gate). They also see the entire input string at once, therefore, they require a fair storage memory. Finally, the paper proposes that in order to learn about the practical limitations associated with LSTMs, application areas such as (1) time series prediction, (2) music composition, and (3) speech processing should be further studied in the future works.

Paper: Long short-term memory and learning-to-learn in networks of spiking neurons

In this paper, Recurrent networks of spiking neurons (RSNNs) is studied. It is found out that there is a significant difference in the performance of designed RSNNs and those of the brain. Computing and learning capabilities of RSNNs have remained poor when compared to ANNs. The paper proposes two reasons for that: first, RNNs in the brain are highly optimized through years of evolution and development, which their detail is still unknown to us. Second mismatch is related to neural adaptation. Therefore, backpropagation through time (BPTT) is used as an optimization method in conjunction with DEEP R. Therefore, the model was enriched by a standard neural adaptation process and was later tested on two tasks, MNIST and TIMIT, where both showed new significant computation levels of RNNs. Since nonlocal learning rules such as backpropagation are challenging to some of the neuromorphic devices (spike-based neuromorphic chips like Brainscales), then alternative methods for RSNN learning of nonlinear functions are needed. Thus, it is shown that L2L can be used to generate RSNNs that learn very efficiently. So, L2L is shown to be finely applicable on RNNs by introducing new motor capabilities to the model, as it is to LSTMs. The result shows a similarity to a reward-based learning in the brain to spike activity. Finally, it is shown that RNNs with sparse activity of 10 – 20 Hz can solve many tasks.