Neural Networks with Memory

Czech Technical University in Prague
Faculty of Information Technology
Ing. Ondřej Kužela
Supervisor: doc. RNDr. Ing. Marcel Jiřina, Ph.D.

One of the biggest dreams of the mankind is the idea of machines being able to replace humans in their duties. Even though human brain is a very complicated device we don't fully understand at this moment, it can't stop us from trying. One of the most important problems with current artificial neural networks is the hardware limitations in means of concurrent threads, which doesn't allow us to simulate all those billions of concurrently working neurons of the human brain. Another important problem is the fact that we don't fully understand how the human brain maintains the mid-term, sometimes also called long-short term, context. One of the possible solutions for both of those problem are neural networks of a family I call here Neural networks with memory.

The Neural networks with memory are those models of the neural networks that except the long term memory in the form of the weights, also propose some sort of a memory for managing the mid-term, sometimes also called long-short-term, dependencies. The memory for managing the mid-term dependencies can be of two types, internal or external. The neural networks with external memory are something, at least for me, more interesting. Over the last few decades there have been several attempts in supplementing the neural network with an additional external memory that would allow the network to retain longer dependencies. But all of those networks have something in common. They all deploy the external memory on the scope of the whole network.

In my thesis I propose a new family of neural networks that deploy the external memory on the scope on parts of the neural network, called modules. This allows the neural network to use several external memories within one neural network, allowing each module to learn contexts of different lengths. The key idea is that this modification will allow a single neural network to successfully track multiple contexts of different lengths. In my thesis I also propose a representative of the newly proposed family, the Recurrent Neural Modules with External Memory.

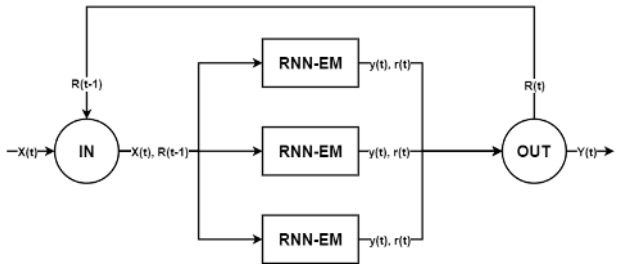

Figure 1:Model of the RNM-EM network with 3 modules.

The Recurrent Neural Modules with External Memory (RNM-EM) is a neural network model that is based on the Recurrent Neural Network with External Memory (RNN-EM) by Peng and Yao, 2015. As its hidden layer it deploys multiple smaller RNN-EM networks, called modules. Each of these modules receives at each time step two types of different contexts, one from its external memory and one from other modules of the network. This improves the potential of the network to learn mid-context.

As a part of my thesis I have compared the newly proposed RNM-EM model to Elman network (RNN), Long Short-Term Memory (LSTM) and RNN-EM, with all the networks having approximately the same amount of learnable variables. All these networks have been tested on the Airline Travel Information System (ATIS) dataset, which is a commonly used language understanding dataset.

Table 1:Table of the results of the comparison testing. All the results are in a form of the F1 score.

Model	Setting	Validation set	Testing set
RNN	120 neurons	96.89	94.04
LSTM	36 neurons	95.17	92.53
RNN-EM	110 neurons 44 memory size 8 memory slots	96.58	93.54
RNM-EM	4 modules 25 neurons 10 memory size 8 memory slots	97.15	94.52

The results of the testing can be found in the table above. From the results it can be seen that the RNM-EM model performed the best from all the tested networks. Even though I am not able to make strong conclusions based on a testing on one dataset, the model shows a significant potential to be an improvement over the currently used neural networks. As a part of my PhD I am currently studying the newly proposed family of the neural networks and working on improvements of the RNM-EM model.