

Neural Networks (LTAT.02.001)

University of Tartu
Institute of Computer Science

Homework 6: Lecture Materials

- These theoretical questions are about the material covered in the lecture about "Recurrent Neural Networks". There will be no theory tasks about the "Applications" lecture. You can get 8 pts for correctly answering these questions. These points are added to the 2 x 25 pts you can get from the next two coding tasks.
- All questions must be answered in your own words, do not copy-paste text or images from the internet.
- Submit the answers to these questions through courses.cs.ut.ee. Under the practice sessions you need to select "Homework 6: theory" and upload the file. The answers must be in a .pdf file. Points can be deducted for terrible formatting or incomprehensible English.

Convolution

Exercise 1: Recurrent networks: basic truths and architectures (7pt)

Which of the following is true about recurrent neural networks. In case of "True" bring an example (e.g. suffices to refer to a slide in lecture), in case of "False", explain.

1. A recurrent network always produces one output prediction per timestep.
2. It is possible to put a convolutional layer between the input and the recurrent layer.
3. The final layer (that makes the classification/regression) must be connected directly to the recurrent layer.
4. In all RNNs types seen in the lecture, output produced at timestep t can only depend on inputs received at timesteps 0 to t , and not on future timesteps.
5. All layers of a recurrent neural network are recurrent.
6. Only one layer of the recurrent network can be recurrent (two recurrent layers would make no sense).
7. The activations of the hidden nodes stay the same across all timesteps.
8. Different timesteps use the exact same weight matrices only in recurrent layers. This does not apply to other layers.
9. Different timesteps use the exact same weight matrices. This applies to all layers.
10. Gradient clipping is important to battle against gradients getting weaker and weaker.

11. LSTM networks deal with long term dependencies better than simple RNNs, because they have less parameters.
12. The gates in LSTM are opened or closed depending on the current input and current hidden state and it does not matter which timestep it currently is.
13. I have trained my network to predict the weather on the 8th day based on the weather of previous 7 days. I now have weather from only past 5 days available. I can still use the same network to predict weather tomorrow (accuracy might be bad, but we can use fewer inputs without changing the network).
14. I can train a recurrent network to receive input only at the first timestep and then produce arbitrarily long sequence of outputs.

Exercise 2: Feedback (1pt)

Answer the following questions from range 0 (not at all) to 10 (absolutely). In addition you can write explanation or comments if you have suggestions or complaints.

1. Theory exercises helped me to understand the topics better. With only practice tasks I would have understood less.
2. I would prefer to have a theory exam instead of having theory exercises every week.
3. I think getting points from theory tasks was easier than from coding tasks.
4. In how many percent of cases was it possible to solve theory exercises using only lecture slides (+ calculator + logic). (Use 0 as 0%, 1 as 10% and 10 as 100%)