Homework set - Recurrent Neural networks - Part 1

April 23, 2020

Guidelines

- The exercise uses the Penn Tree Bank (PTB) from Tomas Mikolov web page PTB dataset download
- You must use Tensorflow 2.0 (or higher) for this assignment.
- All plots must have named axis, grids and title. If more than one plot is on the same figure, provide legend.

1 Preliminaries

Before implementing RNNs with Tensorflow 2.0, you should feel comfortable with implementing feed forward (FF) network. For this purpose, it is highly recommended to implement the Tensorflow tutorial on the Fashion MNIST dataset, as given in the following link.

2 Self-Reading

Read LSTM lecture from the following link.

3 Simple Recurrent Model

Implement the tutorial of time series from Tensorflow's website in the following link.

4 Language Modeling

Read the following paper [1]. In the experiments section, focus only on the experiments over the PTB dataset.

- Reconstruct with your own code the experiment on the PTB and obtain the same result as published in the paper.
- Read the paper where variational dropout was presented [2]. Add variational dropout to the model you obtained in the previous task. Show the improvement in perplexity

Remark 1 *The perplexity of a model is the exponent of the cross-entropy between the true distribution and the model. For further details visit the following link.*

Submission

The work in this assignment should be summarized into a pdf document, where you present both tutorials from Tensorflow website. Moreover, summarize the results over the PTB dataset and emphasize the difference of the results between Zaremba's model and Zaremba's model with variational dropout.

1. Model description: the network specification (#of parameters, network architecture, chosen optimizer etc.).
2. Plot: training and validation learning curve (perplexity w.r.t epoch).
3. Plot: gradient norm curve per update (not epoch).
4. Training time.

GOOD LUCK!!!

References

[1] Wojciech Zaremba, Ilya Sutskever, Oriol Vinyals. *Recurrent Neural Network Regularization.*

[2] Yarin Gal, Zoubin Ghahramani. *A Theoretically Grounded Application of Dropout in Recurrent Neural Networks.*