CSE847 Project Proposal

Using Fast Weights to Attend to the Recent Past*

Eric Wayman
Department of Computer Science and Engineering
Michigan State University
waymane1@msu.edu

Adi Mathew
Department of Computer Science and Engineering
Michigan State University
mathewad1@msu.edu

ABSTRACT

This paper provides a sample of a LaTeX document which conforms, somewhat loosely, to the formatting guidelines for ACM SIG Proceedings.

KEYWORDS

ACM proceedings, LaTeX, text tagging

1 INTRODUCTION

Main paper: (Ba et al. 2016a)

Where LSTM fits into historical picture: (Schmidhuber 2015)

Hopfield nets, associative memory: (Mackay 2003)

Layer normalization: (Ba et al. 2016b)

Grid search: (Goodfellow et al. 2016)

Adam optimizer: (Kingma and Ba 2014)

IRNN definition: (Talathi and Vartak 2015)

2 PROBLEM DESCRIPTION

3 SURVEY OF PRIOR WORK

4 PRELIMINARY PLAN

REFERENCES

Jimmy Ba, Geoffrey E Hinton, Volodymyr Mnih, Joel Z Leibo, and Catalin Ionescu. 2016a. Using Fast Weights to Attend to the Recent Past. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 4331–4339. http://papers.nips.cc/paper/6057-using-fast-weights-to-attend-to-the-recent-past.pdf

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. 2016b. Layer Normalization. (2016). arXiv:1607.06450

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. MIT Press. http://www.deeplearningbook.org.

Diederik P. Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. (2014). arXiv:arXiv:1412.6980 (published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015)

David J. C. Mackay. 2003. *InformationTheory, Inference, and Learning Algorithms*.

Jrgen Schmidhuber. 2015. Deep learning in neural networks: An overview. *Neural Networks* 61 (2015), 85 – 117. DOI:http://dx.doi.org/10.1016/j.neunet.2014.09.003

Sachin S. Talathi and Aniket Vartak. 2015. Improving performance of recurrent neural network with relu nonlinearity. (2015). arXiv:arXiv:1511.03771

*The full version of the author's guide is available as `acmart.pdf` document