

Accepted Manuscript

Novel Architecture for Long Short-Term Memory Used in Question Classification

Wei Xia, Wen Zhu, Bo Liao, Min Chen, Lijun Cai, Lei Huang

PII: S0925-2312(18)30306-0
DOI: [10.1016/j.neucom.2018.03.020](https://doi.org/10.1016/j.neucom.2018.03.020)
Reference: NEUCOM 19422

To appear in: *Neurocomputing*

Received date: 6 July 2017
Revised date: 6 March 2018
Accepted date: 9 March 2018

Please cite this article as: Wei Xia, Wen Zhu, Bo Liao, Min Chen, Lijun Cai, Lei Huang, Novel Architecture for Long Short-Term Memory Used in Question Classification, *Neurocomputing* (2018), doi: [10.1016/j.neucom.2018.03.020](https://doi.org/10.1016/j.neucom.2018.03.020)



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Novel Architecture for Long Short-Term Memory Used in Question Classification

Wei Xia^a, Wen Zhu^{a,*}, Bo Liao^a, Min Chen^a, Lijun Cai^a, Lei Huang^b

^a*College of Information Science and Engineering, Hunan University, Changsha 410082, Hunan, China.*

^b*IBM Almaden Research Center, San Jose, California, USA.*

Abstract

Long short-term memory (LSTM) is an extension of the recurrent neural network (RNN) and has achieved excellent performance in various tasks, especially sequential problems. The LSTM is chain-structured, and its architecture is limited by sequential information propagation. In practice, it is hard to solve the problems involving very long term dependencies. Recent studies have indicated that adding recurrent skip connections across multiple timescales may help the RNN improve its performance in long-term dependencies. Moreover, capturing local features can improve the performance of the RNN. In this paper, we propose a novel architecture (Att-LSTM) for the LSTM, which connects continuous hidden states of previous time steps to the current time step and applies an attention mechanism to these hidden states. This architecture can not only capture local features effectively but also help learn very long-distance correlations in an input sequence. We evaluate Att-LSTM in various sequential tasks, such as adding problem, sequence classification, and character-level language modeling. In addition, to prove the generalization and practicality of the novel architecture, we design a character-level hierarchical Att-LSTM and refine the word representation with a highway network. This hierarchical model achieved excellent performance on question classification.

Keywords: LSTM, Hierarchical Model, Highway Network, Recurrent Skip

*Corresponding author

Email addresses: atxwei@163.com (Wei Xia), syzhuwen@163.com (Wen Zhu), dragonbw@163.com (Bo Liao), 9918428@qq.com (Min Chen), in_ljcai@yeah.net (Lijun Cai), lei.huang1@ibm.com (Lei Huang)

Download English Version:

<https://daneshyari.com/en/article/6863897>

Download Persian Version:

<https://daneshyari.com/article/6863897>

[Daneshyari.com](https://daneshyari.com)