

ZHENG YUAN

(669)273-5981 ◇ zhengyuan.beihang@gmail.com ◇ <http://users.eecs.northwestern.edu/~zys133>

EDUCATION

| | |
|---|-------------------|
| Ph.D. in Computer Science, <i>Northwestern University, US</i> | 01/2015 - 06/2020 |
| M.S. in Computer Science, <i>Beihang University, China</i> | 09/2011 - 07/2014 |
| B.S. in Information Security, <i>Northeastern University, China</i> | 09/2007 - 06/2011 |

SUMMARY

- Interests: Cloud AI, Natural language processing, Named entity typing, Dialog generation, Sequence tagging, Deep learning, Knowledge extraction, Parallel computing
- Programming Skills: Python, Tensorflow, C, Java

WORKING EXPERIENCE

| | |
|---|-------------------|
| Cloud AI Deployment (Google LLC, Fulltime) | 06/2020 - present |
|---|-------------------|

Working at Cloud AI Deployment team as a software engineer.

| | |
|---|-------------------|
| Dialog Generation (Samsung Research America, Intern) | 03/2019 - 07/2019 |
|---|-------------------|

In dialog generation, traditional seq2seq usually generates safe but boring responses. The aim of this research is to combine knowledge into dialog generation. We train an RL model which learns to select knowledge given dialog context. Then, our seq2seq model can leverage the selected knowledge to generate more engaging and informative responses.

| | |
|--|-------------------|
| Unseen Expression Learning (Samsung Research America, Intern) | 06/2018 - 09/2018 |
|--|-------------------|

We studied a novel problem, which aims to teach the NLU model to quickly learn new expression and maximally avoid forgetting the data learned during pre-training. The idea is to incorporate a novel component, called intent pivot, to create explicit intent representations using the practically available intent textual descriptions.

SELECTED PROJECT EXPERIENCE

| | |
|--|-------------------|
| Text Generation And Classification In Peer Grading (Thesis) | 09/2018 - 06/2020 |
|--|-------------------|

Collected peer grading data of project write-ups in classes at Northwestern University. Tasks including: improve peer grading accuracy, predict important spans of write-ups and auto-grade quality of write-ups.

| | |
|--|-------------------|
| Scientific Paper Bibliography Information Parsing | 10/2017 - 04/2018 |
|--|-------------------|

This is a joint project with Allen Institute for Artificial Intelligence(AI2). Design a LSTM+CRF model to tag bibliography information of millions of scientific papers, including bib titles, bib authors, bib venues and bib years.

| | |
|---------------------------------|-------------------|
| Open Named Entity Typing | 05/2017 - 12/2017 |
|---------------------------------|-------------------|

Proposed the task of Open Named Entity Typing (ONET), which is NET when the set of target types is not known in advance. Proposed a neural network architecture, which minimizes the embedding distance between entities and types in a common space. This model achieved high average type AUC score.

| | |
|------------------------------------|-------------------|
| Hyponym Relation Extraction | 02/2017 - 05/2017 |
|------------------------------------|-------------------|

Designed and implemented a neural network to detect hyponym relation in a sentence. Given a sentence and two key phrases, this neural network predicts if the sentence shows hyponym relationship of the key phrases.

Hyper-parameter Tuning

09/2016 - 12/2016

Proposed a regression method to predict best hyper-parameters for any machine learning algorithms especially when the size of training dataset is large.

Large Topic Model Parallelization

05/2016 - 09/2016

Using MPI and multi-thread techniques to parallelize Sparse Backoff Tree (SBT) topic model, a tree structured hierarchical topic model for large data sets.

Parallel Data Compression

01/2015 - 03/2016

Designed and implemented a parallel data compression algorithm, using clustering techniques, for spatial-temporal data sets. The algorithm has been applied to compress several Terabytes of FLASH (an astronomy simulation) data and achieves high compression ratios.

Parallel Computing Framework

09/2012 - 08/2014

Designed and implemented a master-slave model based parallel framework. The parallel framework is designed for DOCK (a bio-medical new drug discovery application) dispatching in Tianhe-2 super-computing environment using more than 6300 computation nodes.

Cluster Simulator

04/2013 - 09/2013

Designed and implemented a cluster simulator. The cluster simulator takes attributes of tasks to calculate the schedule lengths of all tasks submitted to a cluster. This simulator can be used to evaluate the performance of a cluster task scheduler.

Operating Systems Implementation

12/2007 - 02/2011

Designed and implemented two operating systems on ARM and X86 platform.

SELECTED PUBLICATIONS

Zheng Yuan, Doug Downey, *Practical methods for semi-automated peer grading in a classroom setting*, ACM UMAP (2020).

Zheng Yuan, Doug Downey, *OTyper: A Neural Architecture for Open Named Entity Typing*, AAAI (2018).

Zheng Yuan, William Hendrix, Seung Woo Son, Christoph Federrath, Ankit Agrawal, Wei-keng Liao, and Alok Choudhary, *Parallel Implementation of Lossy Data Compression for Temporal Data Sets*, HiPC (2016).

Zheng Yuan, Bo Li, Qiang Hao, Zhongzhi Luan, Depei Qian, *MEFTTD: A Parallel Awareness Scheduler*, CSC (2013).

Yulong Ou, Bo Li, **Zheng Yuan**, Qiang Hao, Zhongzhi Luan, Depei Qian, *LPFSC: A Light Weight Parallel Framework for Super Computing*, PDCAT (2012).