Ruosong Yang

Ph. D. Candidate supervised by Prof. Jiannong Cao

Room PQ503, Mong Man Wai Building, The Hong Kong Polytechnic University, Hong Kong

Phone: 132-9749-7314 / +852 54816630   Email: csryang@comp.polyu.edu.hk

Education:

Sept. 2013 – Jun. 2016	**The Hong Kong Polytechnic University**	Hong Kong, China
	Ph. D. candidate	
	Major: Computer Science	
	Supervisor: Jiannong Cao	
Sept. 2013 – Jun. 2016	**National University of Defense Technology**	Changsha, China
	Master	
	Major: Computer Science	
	Supervisor: DianXi Shi	
Sept. 2009 - Jun. 2013	**Beihang University**	Beijing, China
	Bachelor	
	Major: Computer Science	

Professional Skills:

✧   General Programming Language: C, Python, Java, C++

✧   Deep Learning Tools: Pytorch

✧   Familiar NLP Topics: Word Embedding, Sentiment Classification, Text Summarization, Text Style Transfer

Pulications:

✧ EMNLP 2019	Under Preparation	Copy and Generation based Text Style Transfer
✧ ACL 2019(short paper)	Submitted	GGP: Glossary Guided Post-processing for Word Embedding

Research and Engineering Experience:

Feb.2019 – May.2019	**Copy and Generation based Text Style Transfer (Under Preparation)**

Text style transfer is a new topic in text generation without aligned corpus. Recent works mainly focus on learning text representation and style representation. However, generating a new sentence only based on the condition vector is hard to generate better text (Refer to Point Network).

Based on a simple assumption, word usage could show different style. In our model, we define a soft template to guide text generation. More specifically, we divide the words in sentence into three categories, namely text word, style word and style around word. For text word, we copy from the input text, and we also copy style word from the style vocabulary. And we generate a new word as style around word.

Jan.2019 - Feb.2019	**Chinese Name Generation from Ancient Poetry Based on Sequence to Sequence Model with Copy Mechanism (For interest)**

In Chinese, many parents prefer to name their baby referring to ancient poetry. According to the dataset of Chinese names and that of Chinese Poetry. We construct a dataset with name , poetry pairs.

We build a simple sequence to sequence model. In decoding, the name is copied from the input. And when converge, our model could generate name from the given sentence.

Sep.2018 - Feb.2019	**GGP: Glossary Guided Post-processing for Word Embedding (Submitted to ACL)**

Word Embedding helps many NLP tasks to improve performance. However, for the limit of corpus, Word2vec, Glove etc. could not distinguish antonyms. Researchers proposed new

models to integrate knowledge bases to improve the word embedding. However, recent works mainly focus on semantic knowledge, which results in learned embedding is less functional.

In this paper, we propose a new model to incorporate glossary to learn a global post-processing function to learn word embedding exhibiting more topical and functional.

In the experiment, our model outperforms all baseline models including two state-of-the-art models in functional word similarity datasets.

Mar.2015 - Aug.2015	**Measuring Users Relationship Strength Using a Model Based on Hierarchical Voting (Master Thesis)**

Relationship strength reflects the degree of intimacy between two different persons, which is of great importance in analyzing human's social relationship as well as social network. In this paper, we proposed User Relationship Strength Hierarchy Vote(URSHV), a hierarchical model measures relationship strength by user's daily moving trajectories, semantic locations and the corresponding semantic labels based on weighted DTW and several NLP models (LDA and Word2vec).

Experimental Results show the effectiveness of our proposed model. Which also show the generalization of LDA and Word2vec when coping with trajectory data.

June. 2014 - Aug. 2014	**Genetic Algorithm based Automatic Jiju Ci Poetry Generation (For interest)**

Jiju Ci poetry is a special way to create Chinese classical Ci Poetry, which selects different sentences from Ci poems created by others.

We define the fitness function as the composition of several hand-crafted evaluations including the theme consistency of each two adjacent sentences, the semantic coherence of each two adjacent sentences, the similarity between the Poetry and given theme as well as the consistency of the emotion of the Poetry. And new Ci poetry generated until the convergence of the Genetic Algorithm.

Foreign Language Abilities:

IELTS	L 7.5	R 7.5	W 6.0	S 5.5	Overall 6.5	in 2016

Review for Conferences and Journals:

AAAI, IJCAI, WWW

Volunteers/Helpers for Organizing Conferences:

Helpers for OPODIS 2018

Helpers for CloudCom 2017

Helpers for SMARTCOM 2017