

Wangchunshu Zhou

<https://michaelzhouwang.github.io>

Education

Beihang University

Beijing, China

Master in Computer Science and System Engineering (Multidisciplinary)

Diplome d'ingenieur (French Engineering Degree)

Sep. 2018 - June. 2021 (Expected)

Advisor: Prof. Ke Xu

Beihang University

Beijing, China

B.S. in Information and Computation Science (Sino-French Engineering School)

Major courses in Mathematics, Physics, and General Engineering.

Sep. 2014 - Jul. 2018

Overall GPA 3.5/4.0 (top 20%), Major GPA 3.6/4.0, Senior Year GPA 3.8/4.0 (top 2/115)

Research Interest


My primary research goal is to bridge the gap between the advances on SOTA numbers on leaderboards and the impact of language technology on the lives of humankind, and develop **Language Technology for All**. To achieve this goal and make language technology accessible in most people's lives, I identify two major research topics that I'm interested in: **efficiency** and **trustworthiness** of NLP models. Efficiency involves both the amount of **computation** and **data** required for (pre-)training and using NLP models. Trustworthiness involves the **interpretability**, **fairness**, and **robustness** with respect to adversarial attacks and out-of-distribution samples, as well as reliable **evaluation** of NLP models beyond numbers on leaderboards.

I am also interested in or working in the following subjects:

- o Commonsense reasoning and knowledge-based reasoning[4,13].
- o Robust NLP models for OOD samples and reducing spurious dataset biases.
- o Interpretability[10], biases, and fairness of NLP models..
- o Green NLP[8,9], Low resource NLP[6,11], and Learning NLP models from high-level supervision[10].
- o Natural language generation[3,5], creative text generation[12], evaluation for NLG models[2].

Publications & Preprints

- o (Under Review) **Pre-training Text-to-Text Transformers for Concept-centric Commonsense.**[13]
 - Wangchunshu Zhou*(equal contribution), Dong-Ho Lee*, Ravi Selvam, Seyeon Lee, Bill Yuchen Lin, Xiang Ren
 - Submitted to **ICLR 2021**, received (8,7,6,4) review score. Preliminary version accepted by NeurIPS 2020 Workshop on Self-Supervised Learning.
 - We introduce novel self-supervised pre-training objectives in both generative and contrastive form and a joint training framework to improve the commonsense reasoning ability of pre-trained text-to-text transformers and train CALM, a Concept-Aware Language Model.
- o (Under Review) **Hierarchical Summary to Article Generation.**[12]
 - Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou.
 - Submitted to **AAAI 2021**
 - We introduce a hierarchical model to generate long articles based on short summaries by first generating sketches in an intermediate length, and propose techniques to bridge the gap between training and inference of the model and a novel evaluation metric for conditional long text generation tasks.
- o **Connecting the Dots Between Fact Verification and Fake News Detection.**[11]
 - Qifei Li*, Wangchunshu Zhou*(equal contribution).
 - in Proceedings of **COLING 2020 (Oral)**
 - We propose to connect the dots between Fact Verification, which is better explored, and Fake News Detection, which is less developed, by using a pre-trained summarization model to summarize a news article into a short claim and then use a pre-trained fact verification model to classify the trustworthiness of the news.
- o **Towards Interpretable Natural Language Understanding with Explanations as Latent Variables.**[10]
 - Wangchunshu Zhou*(equal contribution), Jinyi Hu*, Hanlin Zhang*, Xiaodan Liang, Maosong Sun, Chenyan Xiong, Jian Tang.
 - in Proceedings of **NeurIPS 2020**
 - We develop ELV, a general framework for interpretable natural language understanding that requires only a small set of human annotated explanations for training.
- o **BERT Loses Patience: Fast and Robust Inference with Early Exit.**[9]
 - Wangchunshu Zhou*(equal contribution), Canwen Xu*, Tao Ge, Julian McAuley, Ke Xu, Furu Wei.
 - in Proceedings of **NeurIPS 2020**

- We propose PABEE, a straightforward yet effective inference method that can be used as a plug-and-play technique to simultaneously improve the efficiency and robustness of a pretrained language model (PLM).
- o **BERT-of-Theseus: Compressing BERT by Progressive Module Replacing.**[8]
 - Canwen Xu*, Wangchunshu Zhou*(equal contribution), Tao Ge, Furu Wei, Ming Zhou.
 - in Proceedings of **EMNLP 2020**
 - We propose a novel model compression approach that progressively replace each component in an large model by smaller modules. We apply our proposed approach to compress BERT and achieved state-of-the-art performance in comparable settings.
- o **Pseudo Bidirectional Decoding for Local Sequence Transduction.**[7]
 - Wangchunshu Zhou, Tao Ge, Ke Xu.
 - in Proceedings of **EMNLP 2020 (Findings)**
 - We propose a pseudo bidirectional decoding approach which utilize the characteristic of local sequence transduction tasks like grammatical error correction and spell correction.
- o **Improving Grammatical Error Correction with Machine Translation Pairs.**[6]
 - Wangchunshu Zhou, Tao Ge, Chang Mu, Ke Xu, Furu Wei, Ming Zhou.
 - in Proceedings of **EMNLP 2020 (Findings)**
 - We propose to use a pair of Machine Translation models with different qualities to synthesize pseudo-parallel data for pretraining Grammatical Error Correction models.
- o **Scheduled DropHead: A Regularization Method for Transformer Models.**[5]
 - Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou.
 - in Proceedings of **EMNLP 2020 (Findings)**
 - We introduce DropHead, a structured dropout mechanism for the multi-head attention mechanism in the transformer models, and a specifically designed dropout rate schedule for DropHead.
- o **CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning** [4]
 - Bill Yuchen Lin, Wangchunshu Zhou, Ming Shen, , Pei Zhou, Chandra Bhagavatula, Yejin Choi, Xiang Ren
 - in Proceedings of **EMNLP 2020 (Findings)**
 - We propose a novel constrained text generation task and released a dataset to test the generative commonsense reasoning ability of different NLG models.
- o **Self-Adversarial Learning with Comparative Discrimination for Text Generation.** [3] 
 - Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou.
 - in Proceedings of **ICLR 2020**
 - We propose to integrate the self-play mechanism, which is commonly used in the RL community, into training of GANs to reduce the reward sparsity and mode collapse problem and make training more stable.
- o **Learning to Compare for Better Training and Evaluation of Open Domain Text Generation Models.** [2]
 - Wangchunshu Zhou, Ke Xu.
 - in Proceedings of **AAAI 2020 (Oral)**.
 - We proposed a novel "Learning to Compare" paradigm and employ the skill rating system, which is commonly used to evaluate human chess players' skill, to evaluate the performance of open domain text generation systems.
- o **BERT-based Lexical Substitution.** [1]
 - Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou.
 - in Proceedings of **ACL 2019**
 - We proposed a novel lexical substitution based on pretrained masked language models (e.g. BERT) to automatically propose substitute candidates and rank them without relying on external lexical resources.

Research Experiences

- | | |
|--|--|
| <ul style="list-style-type: none"> o Research Intern at Ink Lab of USC
 <i>Ink Lab, Mentor: Prof. Xiang Ren</i>
 Research Internship, working on two projects related to machine commonsense reasoning. | Los Angeles, U.S.
<i>Jan. 2020 - Present</i> |
| <ul style="list-style-type: none"> o Research Intern at MILA
 <i>Tang's Group, Mentor: Prof. Jian Tang</i>
 Research Internship, working on Interpretable Natural Language Understanding. | Quebec, Canada
<i>Mar. 2020 - Jun. 2020</i> |
| <ul style="list-style-type: none"> o Research Intern at Microsoft Research Asia
 <i>Natural Language Computing Group, Mentor: Dr. Tao Ge</i>
 Research Internship, working on natural language generation and efficient pre-trained language models. | Beijing, China
<i>Dec. 2018 - Present</i> |
| <ul style="list-style-type: none"> o Research Assistant at NLSDE Lab
 <i>Natural Language Processing Group, Advisor: Prof. Ke Xu</i>
 Research Assistant, working on natural language generation and efficient pre-trained language models. | Beijing, China
<i>Aug. 2018 - Present</i> |

Services

- o Reviewer: NeurIPS 2020, AAAI 2021

- Student Volunteer: ACL 2019, ICLR 2020, EMNLP 2020
- Review Assistant: EMNLP 2019, ACL 2020

Honors and Awards

- National Scholarship (Graduate Student), 2020 (top 4/240)
- Student Travel Grant: AAAI 2020, ICLR 2020
- Scholarship of Academic Excellence (Master), Beihang University. 2018, 2019 (top 15%)
- The CASC Award, Beihang University. 2017 (top 3/115)
- Scholarship of Academic Excellence (Bachelor), Beihang University. 2016, 2017 (top 20%)