HAORAN ZHAO

hz454@drexel.edu haoranzhao419.github.io

EDUCATION

Drexel University **Philadelphia, PA**

Bachelor of Science in Data Science; GPA 3.84/4.0 *Sep 2021 – June 2024 (anticipated)*

Minor: *Computer Science and Linguistics*

Advisor: *Dr. Jake Williams, Dr. Shadi Rezapour*

Lanzhou University **Lanzhou, China**

Bachelor of Science in Computer Science and Technology; GPA 3.97/5.0 or 89.7/100.0 *Sep 2019 – July 2023*

RESEARCH INTERESTS

Natural Language Processing (NLP) & AI & Cognitive Science

I'm interested in both understanding language models (LMs) cognitive abilities and making human-like cognitive language models

RESEARCH EXPERIENCE

Stanford University CoCoLab | *Research Assistant* **August 2023 - Present**

- Advisor: Dr. Noah D.Goodman
- Develop a template-base method to generate a benchmark to evaluate LLMs ***pragmatic reasoning*** abilities, especially in understanding ad-hoc, scalar, hyperbole, metaphor, and indirect speech
- Examine LLMs's ability in understanding figurative languages with a specific focus on designing experiments to evaluate LLM's non-literal number understanding **(*hyperbole & pragmatic halo*)** abilities

Drexel University CODED Lab | *Research Assistant* **April 2022 - September 2023**

- Advisor: Dr. Jake Williams
- Create a new word representation method (***Bit-Cipher***) based on unigram frequency rankings of words with the inspiration of one-hot encoding combined with efficient dimensionality reduction technique.
- Search the ***Explicit Optimization of Feed Forward Neural Network*** by leveraging the two variants of word2vec: Continuous bag-of-words (CBOW) and Skip-Gram (SG).
- Search for the explicit optimization of foundation models with ***Self-Attentive Feed-Forward Neural Units*** in Transformer-based architecture.

Graz University of Technology SEAI Lab | *Research Assistant* **January 2022 - March 2023**

- Advisor: Dr. Franz Wotawa
- Fine-tune pre-trained language models to build domain-specific NER models with the human-annotated and Wikipedia-efficiently-extract data to investigate pre-trained LM's ability on domain-specific tasks. (Accepted to IEEE-DSA-2023)

PUBLICATIONS

Alexander Perko, **Haoran Zhao**, Franz Wotawa. *Optimizing Named Entity Recognition for Improving Logical Formulae Abstraction from Technical Requirements Documents.* The 10th International Conference on Dependable Systems and Their Applications (DSA-2023).

Xiao Fang, Alex Kalinowski, **Haoran Zhao**, Ziao You, Yuhao Zhang, and Yuan An. *Prompt Design and Answer Processing for Knowledge Base Construction from Pre-trained Language Models (KBC-LM)*. Challenge @ 21st International Semantic Web Conference (ISWC 2022) CEUR Workshop Proceedings

PREPRINTS

Zhao Haoran, and Jake Ryland Williams. *Bit Cipher–A Simple yet Powerful Word Representation System that Integrates Efficiently with Language Models.* arXiv preprint arXiv:2311.11012 (2023).

Jake Ryland Williams, and **Haoran Zhao**. *Explicit Foundation Model Optimization with Self-Attentive Feed-Forward Neural Units.* arXiv preprint arXiv:2311.07510 (2023).

Jake Ryland Williams, and **Haoran Zhao** *Reducing the Need for Backpropagation and Discovering Better Optima With Explicit Optimizations of Neural Networks.* arXiv preprint arXiv:2311.07498 (2023).

HONORS AND AWARDS

Steinbright Partners Program Award (1 of the 10 awardees across the university), Drexel University ($6,000) June 2023

2023 Week of Undergraduate Excellence, Drexel University May 2023

Undergraduate Research Mini-Grant (1 of the 13 awardees across the university), Drexel University ($2,000) March 2023

Third-class scholarship for outstanding students (¥1,500), Lanzhou University Sep 2022

A. J. Drexel Scholarship ($3,333 a term), Drexel University 2021-Present

Second-class scholarship for outstanding students (¥2,000), Lanzhou University Sep 2021

Innovation And Entrepreneurship Scholarship, Lanzhou University (¥10,000 a year) 2020 & 2021 & 2022

Second Prize of Contemporary Undergraduate Mathematical Contest in Modeling June 2021

ACADEMIC PROJECTS

Efficient Word Analogy Searching by Leveraging Machine Translation Techniques **Philadelphia, PA**

Individual Project (Advisor: Jake Williams) *September 2022 - January 2023*

- Made the hypothesis *"Translation errors within one language form word analogies."*
- Constructed a transformation matrix to analyze the relationship between words in two political datasets, one each about Democrats and Republicans, and validated our hypothesis. [virtual poster]

Knowledge base Construction from Pre-trained Language Model **Philadelphia, PA**

Team Member (Advisor: Yuan An) *May 2022 - July 2022*

- Explored the capability of BERT for constructing a knowledge base with a set of given predicates/relationships
- Extracted relations with prompt and answer engineering approaches and setting different thresholds for each individual relation (Accepted to KBC-LM Challenge @ ISWC 2022) [paper]

National NLP Clinical Challenge **Philadelphia, PA**

Team Member (Advisor: Yuan An) *Jan 2022- May 2022*

- Combined content-based and rule-based methods to extract medicine names from prescriptions
- Recognized medicine names with the method of CRF by tagging each entity with BIO-annotation
- Built Bidirectional LSTM with CRF and attention mechanism to extract name entities from prescriptions

COVID-19 Forecasting Model **Gansu, China**

Team Leader (Advisor: Binbin Yong) *April 2021 - June 2021*

- Collected data on the "Our World in Data" Website, working with two teammates
- Built a prediction model with Scikit-Learn and TensorFlow, using polynomial regression and Convolutional Neural Network (CNN)
- Visualized forecasting result with Matplotlib and Plotly

SKILLS

Languages: Python, R, JavaScript, SQL, Bash, Hadoop, Spark

Libraries: Scikit-Learn, spaCy, NLTK, Numpy, Pandas, PyTorch, Transformers

Others: Mandarin (native), English (proficient), Spanish & Japanese (elementary)

SERVICES

Reviewer – Efficient Natural Language and Speech Processing (ENLSP) workshop @ NeurIPS 2023