

# CHIYU ZHANG

chiyuzh@mail.ubc.ca

470-1961 East Mall, Vancouver, BC, Canada, V6T 1Z1

🏠: <https://chiyuzhang94.github.io/>

## HIGHLIGHTS

---

- 8+ years of experience in research of machine learning and natural language processing (NLP).
- **Libraries:** NLTK, spaCy, gensim, scikit-learn, PyTorch, Keras, & Transformers
- **Programming:** Python, R, C#, & MatLab
- **Database:** MySQL, Solr, Hadoop, & Spark
- **Coursework:** Computational Linguistics, Deep Learning in NLP, Social Networks, Machine Learning and Data Mining, Computer Networking, Data Structures, Systematic Program Design.

## EDUCATION

---

### The University of British Columbia, Canada

*Sep. 2018 - May 2024 (expected)*

Doctor of Philosophy in Information Science

Average: 88.1/100.0

Advisor: Dr. Muhammad Abdul-Mageed

**Research Interests:** Natural Language Processing and Deep Learning.

### Clarkson University, USA

*Aug. 2016 - Dec. 2017*

Master of Science in Data Science

GPA: 3.94/4.00

Advisor: Dr. Boris Jukic

### Xiamen University of Technology, China

*Sep. 2012 - Jun. 2016*

Bachelor in Information Systems (Honours)

Average: 87.42/100.00

Advisor: Dr. Yihui Qiu

## EXPERIENCE

---

### Research Intern

May 2023 - Present

*Google, Canada*

- Working on a research project aimed at improving text style transfer task by utilizing large language models, model distillation, and reinforcement learning.

### Research Assistant

Aug. 2018 - May 2023

*The University of British Columbia, Canada*

- Working on various NLP and deep learning research projects, e.g., enhancing mono- and cross-lingual language representations with billion-scale social media data and Transformer-based models, deep learning for socio-pragmatics, identification of lexical dialect variations, language components analysis, graph-based user modelling, and text style transfer.

### Visiting Researcher

Dec. 2022 - May 2023

*Mohamed bin Zayed University of Artificial Intelligence, UAE*

- Worked on NLP and deep learning research projects, e.g., developing large language models for Arabic natural language understanding, Arabic dialects translation, and large model distilling and compressing. Prepared materials and instructed lab sections of a graduate-level course.

### Machine Learning Research Intern

Sep. 2022 - Feb. 2023

*Borealis AI, Canada*

- Worked on a research project aimed at enhancing the robustness and performance of neural-based representation models for time-series through contrastive learning and self-supervised learning.

**Senior Teaching Assistant**

Sep. 2019 - May 2022

*The University of British Columbia, Canada*

- Prepared materials and instructed lab sections of four machine learning courses, e.g., supervised learning and machine translation, in the Master of Computational Linguistics program.

**Research Technician**

Apr. 2018 - Jun. 2018

*The University of Tennessee Health Science Center, USA*

- Worked on the development of a web tool for biomedical text mining. Tasks included establishing a database and search engine (with Solr), and mining 2.2M biomedical literature to identify specific relationships between proteins, genes, or given keywords applying NLP and deep learning methods.

**Research Assistant**

Aug. 2016 - Mar. 2018

*Clarkson University, USA*

- Led a data science project that analyzed 200K job postings to enhance the profile matching service for applicants and recruiters. Tasks included data collection, database deployment, and modelling.

---

**SELECTED PUBLICATIONS**

---

- ***Contrastive Learning of Sociopragmatic Meaning in Social Media***  
Zhang, Chiyu, & Abdul-Mageed, M., (2023). Findings of the Association for Computational Linguistics: ACL 2023. ACL.
- ***Decay No More: A Persistent Twitter Dataset for Learning Social Meaning***  
Zhang, Chiyu, Abdul-Mageed, M., & N., E. M. B., (2022). In Proceedings of the 1st Workshop on Novel Evaluation Approaches for Text Classification Systems on Social Media. ICWSM. (**Best paper**)
- ***Improving Social Meaning Detection with Pragmatic Masking and Surrogate Fine-Tuning***  
Zhang, Chiyu, & Abdul-Mageed, M., (2022). In Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA). ACL.
- ***Toward Micro-Dialect Identification in Diaglossic and Code-Switched Environments***  
Abdul-Mageed, M., Zhang, Chiyu, E., A., & U., L., (2020). In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). ACL.
- ***BERT-Based Arabic Social Media Author Profiling***  
Zhang, Chiyu, & Abdul-Mageed, M., (2019). In Proceedings of the 11th meeting of the Forum for Information Retrieval Evaluation 2019 (FIRE). CEUR-WS.org.
- ***Multi-Task Bidirectional Transformer Representations for Irony Detection***  
Zhang, Chiyu, & Abdul-Mageed, M., (2019). In Proceedings of the 11th meeting of the Forum for Information Retrieval Evaluation 2019 (FIRE). CEUR-WS.org.
- ***No Army, No Navy: BERT Semi-Supervised Learning of Arabic Dialects***  
Zhang, Chiyu, & Abdul-Mageed, M., (2019). In Proceedings of the Fourth Arabic Natural Language Processing Workshop (WANLP). ACL. (**Best system paper**)

---

**HONORS AND AWARDS**

---

- Distinguished Teaching Award (2022), UBC
- Affiliated Fellowship (2021-2022), UBC.
- President's Academic Excellence Initiative Ph.D. Award (2020-2023), UBC.
- Evelyn Markwei Memorial Award (2020-2021), UBC.
- Ph.D. Travel Award of School of Information (2019 and 2021), UBC.
- Graduate Scholarship of School of Information (2018-2023), UBC.
- Graduate Scholarship (2016-2017), Clarkson University.