CASS ZHIXUE ZHAO

44 7516862694
zhixue.zhao@sheffield.ac.uk
S1 4BP Sheffield
Personal Site
Google Scholar

I am a Postdoc researcher in the Department of Computer Science, at the University of Sheffield. My research project focuses on developing explainable NLP models. Before my postdoc, I worked as a research assistant in the same department on building a retrieval system for systematic reviews on inequality in public health, funded by National Institute for Health Research. My PhD research applied transfer learning to hate speech detection tasks and investigated model bias.

I am passionate about innovative research, and working towards being an independent and mature researcher. I have over 600 hours of teaching assistant/marking experience and experience in supervising postgraduate dissertations projects. I have been awarded Associate Fellowship High Education Certificate.

Education

01/2018~08/2022	**The University of Sheffield**, PhD, Information School (fully funded)
Research Topic:	Transfer Learning for Hate Speech Classification, Bias in pre-trained language models
09/2017~09/2018	**The University of Sheffield**, M.Sc. Data Science, Distinction
Core Modules:	Data Analysis, Database Design, Data Mining and Visualization.
09/2008~/2012	**Shanghai Institute of Technology**, B.S. Food Engineering
Core Module:	Advanced Mathematics, Linear Algebra, Microbiology, Chemistry

Publications

Zhao, Z., N. Aletras (2023). Incorporating Attribution Importance for Improving Faithfulness Metrics. The 61st Annual Meeting of the Association for Computational Linguistics. ACL2023.

Zhao, Z., G. Chrysostomou, K. Bontcheva and N. Aletras (2022). On the Impact of Temporal Concept Drift on Model Explanations. In Findings of the Association for Computational Linguistics: EMNLP 2022.

Clowes, M., Stansfield, C., Thomas, J., Shemilt, I., Paisley, S., Stevenson, M., Zhao, Z., Marshall, I., Kell, G., (June 2022). All is FAIR in health inequalities research: using machine learning to build a new database of health equity studies. European Association for Health Information and Libraries 2022.

Zhao, Z., Zhang, Z., Hopfgartner, F. (2022). Utilizing Subjectivity Level to Mitigate Identity Term Bias in Toxic Comments Classification. Online Social Networks and Media, 29, 100205.

Zhao, Z., Zhang, Z., Hopfgartner, F. (2021). A Comparative Study of Using Pre-trained Language Models for Toxic Comment Classification. In Companion Proceedings of the Web Conference 2021 (pp. 500-507)

Zhao, Z., Zhang, Z., Hopfgartner, F. (2019). Detecting Toxic Content Online and the Effect of Training Data on Classification Performance. In Proceedings of 20th International Conference on Computational Linguistics and Intelligent Text Processing, La Rochelle, France

Working Experience

01/2022~01/2024 (expected) Department of Computer Science, The University of Sheffield
Research Associate Disinformation diffusion modelling
This research is part of a large interdisciplinary project, “Social Explainable Artificial Intelligence”, funded by the EPSRC. The overall goal of this project is to explore a wide range of research directions on modelling disinformation diffusion, with a focus on developing explainable ML models. I work independently and with supervision on programming and compiling papers.

01/2023~09/2023 (expected) Information School, The University of Sheffield
Postgradute Research Supervisor
As a dissertation research supervisor, I provide advise on the dissertation project, regarding the research design, time management, programming, results analysis, and writing-up.

07/2022~08/2022 English and American Studies, The University of Manchester
Research Assistant Nineteenth-Century Nature Writing in English and Twenty-First-Century Environmentalism
This pilot project aims to open out research questions for a larger project on women’s nature writing in the long nineteenth century. My duty focuses on pre-processing related text and exploratory research on the related corpus, such as temporal shift and geographic drift.

07/2021~12/2021 Department of Computer Science, The University of Sheffield
Research Assistant Automatically mapping and assessing inequalities in public health research
This is a multidiscipline project, collaborated with UCL and King's College, funded by National Institute for Health Research. This project aims to develop automated methods to find, organize and describe scientific literature relevant to public health and understand its findings in relation to inequalities in health. Particularly, NLP techniques will be used to automatically organize these documents into topics, identify the research method used and identify whether it contains information about factors related to health inequalities at scale. My role was for programming and analysis.

11/2020~03/2021 Department of Computer Science, The University of Sheffield
Research Assistant Text Processing for Health Technology Assessment
The TePHTA is a cross-discipline study with researchers from the Computer Science Department and the School of Health and Related Research, University of Sheffield. This project aims to improve healthcare decision-making through natural language processing. I participate in a part of this project, the medical system review analysis, which utilizes machine learning to assist experts in healthy to gain insights from a massive volume of medical literature regarding employees' health situations and working conditions. I was in charge of programming and analysis.

12/2018~03/2021 Information School, The University of Sheffield, Sheffield
Teaching Assistant (for below modules)

- INF4002 Introduction to programming
- INF6027 Introduction to data science
- INF6060 Information retrieval
- INF6028 Data mining and visualization
- INF6032 Big data analytics
- INF6050 Database design
- INF6024 Researching social media
- INF109 Digital media and society

Society & Club

2023~Now **Communication officer** Staff Race Equality Network, The University of Sheffield
2021~2021 **Organizer** NLP Reading Group, The University of Sheffield
2019~2021 **Research students representative** Information Retrieval Research Group, The University of Sheffield
2019~2020 **Inclusion officer** iSchool society, The University of Sheffield