Zihan Zhang

zihan.zhang-5@student.uts.edu.au · (+61) 0401-115-299 · Personal Website

Education

University of Technology Sydney, Sydney, Australia — March 2021 – Present

Ph.D. student in School of Computer Science, UTS NLP Group

- Natural Language Processing
- Supervisor: Prof. Ling Chen

University of Melbourne, Melbourne, Australia — July 2018 – December 2020

Master in Software Engineering

- GPA: 83/100 (First Class Honours/Distinction)
- Awards: Dean's Honours List (2019 & 2020), Liz Haywood Award (2020)

China Pharmaceutical University, Nanjing, China — September 2014 – June 2018

Bachelor in Information System and Information Management

- GPA: 82/100 (Ranking: 25th/143)
- Awards: Second-Class Scholarship (2016-2017), Excellent Volunteer (2015-2016)
- Activities: President of the Faculty of Science Student Union (2016-2017)

Publications

[1] How Do Large Language Models Capture the Ever-changing World Knowledge? A Review of Recent Advances.
Zihan Zhang*, Meng Fang*, Ling Chen, Mohammad-Reza Namazi-Rad, and Jun Wang
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023

[2] Turn-Level Active Learning for Dialogue State Tracking.
Zihan Zhang, Meng Fang, Fanghua Ye, Ling Chen, and Mohammad-Reza Namazi-Rad
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023

[3] CITB: A Benchmark for Continual Instruction Tuning.
Zihan Zhang, Meng Fang, Ling Chen, and Mohammad-Reza Namazi-Rad
Findings of the Association for Computational Linguistics (EMNLP, Findings), 2023

[4] Is Neural Topic Modelling Better than Clustering? An Empirical Study on Clustering with Contextual Embeddings for Topics.
Zihan Zhang, Meng Fang, Ling Chen, and Mohammad Reza Namazi Rad
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2022

Working Experience

UTS NLP Group Sydney, Australia — March 2021 – Present

Ph.D. Student Fulltime

As an industry-based PhD student, I am passionate about developing NLP systems that are practical, robust, efficient, and can be applied to the industry. My recent efforts involve adapting large language models (LLMs) to world knowledge, including:

- Extensive review of recent approaches to align LLMs with the ever-changing world knowledge without expensive re-training from scratch [1]
- Investigate, benchmark, and analyse practical approaches to continually learn an LM of new tasks while not catastrophically forgetting previously learnt [3]
- Work on efficient, trustworthy, and robust Retrieval-Augmented Generation (RAG) methods for Question Answering (QA) and Dialogue systems

TPG Telecom Sydney, Australia March 2021 – December 2023

Research Analyst Fulltime

As a member of the Data and Analytics Centre of Excellence (CoE) team, I transform numerous textual data into actionable insights to drive business.

- **NPS topic modeling** [4] - apply topic modeling on customers' feedback data and derive actionable insights to increase customers' satisfaction
- **Webchat & Call Centre dialogue analysis** [2] - preliminary study on raw dialogue data between agents and customers; initial efforts in building dialogue models for customers' intent prediction and response generation
- **Market offer engine** - automatically extract and analyse competitors' offers, transforming raw unstructured data into structured data, providing an analysis data source for the product team to make pricing strategies
- **Postpay/Prepay/FWA customer insights analysis** - model and analysis customers' churn and upgrade
- **Cloud experience** - get involved in cloud services, familiar with data ETL and DevOps on AWS and Databricks

RESORTer Melbourne, Australia November 2019 – March 2020

Front-end Software Developer Intern

I was responsible for refactoring and developing the "Lesson Section" module in the resort web application.

- Refactored the Lesson Section using React + Hooks + Material-UI. Used Grid layout and Card component to render different kinds of lessons and simplified the rendering logic, which was a serious issue when previously using the Tabs system.
- Managed the global state using Redux and created default lessons for the users based on the form that they filled. I also cooperated with my team using Middleware to catch and handle certain actions to make sure the generated lessons are always consistent with the global state, thereby improving the user experience.
- Utilized CSS Module to avoid class names collisions and global style pollution. Used lazy load to dynamically import required components, thereby improving the performance.

Relevant Projects

Algorithms in Action Demo: `https://algorithms-in-action.github.io/`

An algorithm visualization web application provided for the first year Computer Science students.
I was responsible for implementing the pseudocode and algorithm animation.

- Using JavaScript function closures, all the visualization API functions and corresponding variables can be stored in an array and executed later, so it solved the problem of using ES6 Generators that functions executions cannot be reversed. Thereby, the animation can step backward as well.
- To map the algorithm pseudocode with the actual code, I parsed and added a bookmark in each line of the pseudocode, and inserted the bookmarks at the corresponding position in the actual code, so the pseudocode and animation can be synchronized.
- Implemented a customized hook useInterval so that the auto-play function can read fresh states between each render. This hook can also detect the speed changes and reset the setInterval function, thereby adjusting the playback speed is achievable.
- Based on the visualization APIs provided by Tracer.js, I implemented some common components and functions and expanded the library as well.

Guttman Chart Analysis System Github

A Guttman chart based students assessment analysis system. It can be used to help educators find students' Zones of Personal Development(ZPD) and adjust future teaching plannings.

- The project provided support for the research in the Assessment Research Centre, Melbourne Graduate School of Education.
- To detect irregular patterns in the data, an algorithm was developed to calculate the local similarity of each item and detect any irregular ones.
- I was responsible for developing the frontend pages and integrating it with the backend developers. The project used Python as the backend programming language, adopted the Client/Server model, and used

RESTful API for HTTP communication.

Distributed Shared Whiteboard

Github

A shared whiteboard desktop application that allows multiple users to draw shapes and chat at the same time. I was responsible for developing the client and server GUI.

- The project used Java 8 as the backend language, JavaFX as the frontend framework, and used a three-tier Client/Server architecture. It separated the client – whiteboard server – data server.
- Java RMI was used as the communication method between the whiteboard server and data server, the request sends from the client were remotely called in the whiteboard server as well. To synchronize each client, MQTT was used to provide a subscribe/publish protocol. The whiteboard server was used as an intermediate agent to accept messages from each client and publish the messages to all other subscribers.

★ Academic Services

Peer Reviewer:

- EMNLP 2022-2023
- EACL 2023
- ACL Rolling Review October, December 2023

Skills & Certificates

Languages: Mandarin (native), English (fluent)
Programming: Python > SQL == JavaScript > Spark == Java > C == C++ == Haskell
Libraries & Services: PyTorch, HuggingFace, Pandas, Scikit-learn, AWS (S3, SageMaker, Redshift), Databricks
Software & Tools & Management: Git, Linux, Agile, Scrum, Confluence, Jira, LaTeX
Web Dev: React.js, Ant Design, material-ui, HTML, CSS
Computer Network: basic HTTP, TCP/IP, cryptography, web security
Certificates:

- AWS Cloud Practitioner
- (Databricks) Large Language Models: Application through Production
- Deep Learning Specialization (Coursera by Andrew Ng)

Reference

Reference available on request.