

# Shiwei Zhang

+61 488 020 091 | dr.shiwei.zhang@gmail.com | Homepage: <https://zswvivi.github.io> | Google Scholar | LinkedIn

## Profile

---

A highly skilled and experienced PhD in Computer Science with seven years of experience researching and developing NLP models. Seeking opportunities to design and implement innovative generative AI algorithms that can improve the quality of life for people around the world.

## Education

---

<b>RMIT University</b> <i>Ph.D., Computer Science</i>	June 2016 – November 2020 <i>Melbourne, Australia</i>
<b>The University of Melbourne</b> <i>Master of Information Technology</i>	August 2009 – March 2012 <i>Melbourne, Australia</i>
<b>Hunan University</b> <i>Bachelor of Computer Science and Technology</i>	September 2003 – June 2007 <i>Changsha, China</i>

## Current Research Interests

---

- **Multi-task generative model:** By training a single generative model that can handle multiple related tasks, we can achieve greater efficiency compared to training separate models for each task, especially when the tasks are interconnected. Additionally, this approach enables the transfer or sharing of in-domain knowledge across tasks, which further enhances efficiency.  
**My medical multi-task T5 Demo:** <https://huggingface.co/spaces/zswvivi/ChineseMedicalT5>
- **Controllable text generation:** To enhance the relevance and usefulness of the generated text, one can design decoding methods or specify particular constraints that ensure the generated text meets the task requirements.

## Working Experiences

---

<b>Applied Medical NLP Researcher</b> <i>Tencent Jarvis Lab</i>	June 2021 – Jan 2023 <i>Shenzhen, China</i>
<ul style="list-style-type: none"><li>• Research and develop medical multi-task generative models, and methods for controllable text generation.</li><li>• Explore and create innovative NLP models for medical applications, with a top conference paper published.</li><li>• Develop and deploy NLP models for two systems: the intelligent pre-consultation system and the self-diagnosis system. These NLP models would include medical symptom detector, intent classifier, and disease diagnosis model using EHR. The ultimate goal of our systems is to create more intelligent and accurate healthcare systems that can improve patient outcomes.</li><li>• Mentor Master and PhD students on developing deep learning models for healthcare-related problems, with one top conference paper published.</li></ul>	
<b>Data Analyst</b> <i>Telstra Health</i>	August 2016 – June 2017 <i>Melbourne, Australia</i>
<ul style="list-style-type: none"><li>• Write SQL scripts to perform data cleaning, database management and data analysis.</li></ul>	

## Research Experience

---

<b>Research Assistant</b> <i>RMIT University</i>	November 2020 – March 2021 <i>Melbourne, Australia</i>
<ul style="list-style-type: none"><li>• Investigate whether QA-based intermediate training benefits text classification.</li></ul>	
<b>PhD Candidate</b> <i>RMIT University</i>	June 2016 – November 2020 <i>Melbourne, Australia</i>
<ul style="list-style-type: none"><li>• Thesis Topic: Mining User-generated Texts via Neural Classification.</li><li>• Supervisory Team: Prof Xiuzhen (Jenny) Zhang (RMIT), Dr Jey Han Lau (The University of Melbourne), Assoc. Prof Jeffrey Chan (RMIT), Dr Cécile Paris (CSIRO).</li><li>• Mentor Master and PhD students on developing deep learning models for NLP problems, with one top conference paper published.</li></ul>	

## Awards

---

Best Paper Award, Information Processing & Management journal (SJR Q1), 2019.  
1st place at the Melbourne Datathon (\$1500), 2015.

## Skills

---

NLP: Natural Language Generation, Text Matching, Text Classification  
Deep Learning: Hugging Face, PyTorch, TensorFlow  
Model Training: Distributed Training, Hyper-parameters Tuning, Linux Bash Scripting, Batch Scheduling  
Software Development: GitHub, Docker, Linux, Python, Java (Certified Java Developer), SQL (Certified SQL Expert)  
Data Analysis: R, Matplotlib  
Cloud Computing Services: AWS, Tencent Cloud

## Publications

---

- **Shiwei Zhang**, Jichao Sun, Yu Huang, Xueqi Ding, and Yefeng Zheng. Medical Symptom Detection in Intelligent Pre-Consultation using Bi-directional Hard-Negative Noise Contrastive Estimation. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2022.
- Haochun Wang, Chi Liu, Nuwa Xi, Sendong Zhao, Meizhi Ju, **Shiwei Zhang**, Ziheng Zhang, Yefeng Zheng, Bing Qin, and Ting Liu. Prompt Combines Paraphrase: Teaching Pre-trained Models to Understand Rare Biomedical Words. In Proceedings of the 29th International Conference on Computational Linguistics (COLING), 2022.
- **Shiwei Zhang**, Xiuzhen Zhang. Does QA-based Intermediate Training Help Fine-tuning Language Models for Text Classification?. In Proceedings of the 19th Annual Workshop of the Australasian Language Technology Association (ALTA), 2021.
- **Shiwei Zhang**, Mining User-generated Texts via Neural Classification. Doctoral dissertation, RMIT University, 2020.
- **Shiwei Zhang**, Xiuzhen Zhang, Jey Han Lau, Jeffrey Chan, and Cecile Paris. Less is More: Rejecting Unreliable Reviews for Product Question Answering. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2020.
- Scanlon, Liam, **Shiwei Zhang**, Xiuzhen Zhang, and Mark Sanderson. Evaluation of Cross Domain Text Summarization. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), pages 1853—1856. 2020.
- **Shiwei Zhang**, Jey Han Lau, Xiuzhen Zhang, Jeffrey Chan, and Cecile Paris. Discovering Relevant Reviews for Answering Product-related Queries. In Proceedings of the International Conference on Data Mining (ICDM), pages 1468—1473, 2019.
- **Shiwei Zhang**, Xiuzhen Zhang, Jeffrey Chan, and Paolo Rosso. Irony Detection via Sentiment-based Transfer Learning. Information Processing & Management, 56(5):1633—1644, 2019. **Best Paper Award**
- **Shiwei Zhang**, Xiuzhen Zhang, and Jeffrey Chan. Language-independent Twitter Classification using Character-based Convolutional Networks. In Proceedings of the International Conference on Advanced Data Mining and Applications (ADMA), pages 413—425, 2017.
- **Shiwei Zhang**, Xiuzhen Zhang, and Jeffrey Chan. A Word-character Convolutional Neural Network for Language-agnostic Twitter Sentiment Analysis. In Proceedings of the Australasian Document Computing Symposium (ADCS), pages 1–4, 2017.