Zhaoxiang (Simon) Cai

+61 421 992 704 I zhaoxiang.cai@hotmail.com I https://zhaoxiangsimoncai.github.io/SimonCaiCV.io/

CAREER PROFILE

I am a dynamic researcher and engineer with a robust background in both academic and industrial settings. My journey includes developing large-scale systems at Goldman Sachs, where I successfully managed complex projects under tight deadlines. My PhD in Cancer Data Science at the Children's Medical Research Institute, affiliated with the University of Sydney, marked a significant turn in my career towards integrating machine learning with oncological studies. My passion lies in harnessing the power of artificial intelligence to revolutionise healthcare, particularly in understanding and treating cancer. I am committed to pioneering advancements that will transform oncology, speeding the development of innovative therapies and improving patient outcomes. Beyond my professional pursuits, I am an avid enthusiast of the piano, skiing, and badminton.

EXPERIENCE

The University of Sydney / Children's Medical Research Institute *Sydney, Australia*
World-leading Research Institute

Senior Data Scientist (Cancer Data Science) *Feb 2023 – current*

- Developed new deep learning-based approach to incorporate human knowledge for multi-omic data integration
- Designed and built multi-view VAE models customised for multi-omic data integration
- Performed end-to-end whole exome/genome sequencing data analyses for germline/somatic mutations, copy number variations and structural variants
- Performed end-to-end proteomic data analyses, including data QC, peptide-to-protein rollup, pre-processing, differential expression analysis, pathway analysis and survival analysis
- Performed machine learning on large-scale multi-omic datasets for various predictions

Data Scientist/Bioinformatician *Jan 2019 – Mar 2020*

- Built pipelines using existing models for single-cell RNA-seq analysis in mouse developmental biology
- Built deep learning models for live-cell imaging data analysis

Other achievements:

- The 1st prize of University of Sydney Innovation Challenge 2019 (ocular disease image classification) - $7,500
- The 3rd prize of ODIR-2019 (international computer vision competition) - $20,000 (approx.)

Tools Used: Python, R, SQL, PyTorch, Linux, etc.

Goldman Sachs *Melbourne, Australia*
Leading Investment Bank in the World

Analyst Programmer *Nov 2014 – Jan 2018*

- Communicated with business stakeholders and liaised regarding project scope with ongoing updates
- Designed/developed/tested/deployed system solutions specialised in Goldman Sachs Electronic Trading (GSET) business flow
- Provided production support and maitained the health of testing environment

Other achievements:

- Delivered Shenzhen-Hong Kong stock connect project with zero system issues in a very aggressive 4-month timeline. The project won the annual Federation Award 2016
- Established India GSET clearing clients business flow which attracted over 20 potential institutional clients
- Completed 10+ small to medium projects yearly
- Leadership and problem-solving skills were well recognised by both technology and securities trading teams
- Was the captain of GS Squash Club

Tools Used: C++, SQL, Python, Linux, Perl, HTML, JIRA, Confluence, MS Office Suite etc.

EDUCATION

UNIVERSITY OF SYDNEY – Children's Medical Research Institute *Sydney, Australia*
Doctor of Philosophy (Cancer Data Science) *Mar 2020 – Feb 2023*

- Large-Scale and Pan-Cancer Proteogenomic Analyses with Machine Learning
- Sydney Cancer Partner's PhD scholarship
- The 1st prize of University of Sydney Innovation Challenge 2020

UNIVERSITY OF MELBOURNE - MELBOURNE BUSINESS SCHOOL *Melbourne, Australia*
Master of Business Analytics *Jan 2018 – Dec 2018*

- First Class Honours
- KPMG-MBS Data Challenge - 1st Prize Winner for Natural Language Processing
- MBS scholarship in Business Analytics, 2018
- Co-President of Business Analytics Club

MONASH UNIVERSITY *Melbourne, Australia*
Bachelor of Computer Science (Honours) *Jan 2011 – Jul 2014*

- First Class Honours
- Awards: Dux of Bachelor of Computer Science (overall highest ranking), Bellamy Awards (highest ranking for each year), International Merit Scholarship
- Publication: HetFHMM: A Novel Approach to Infer Tumour Heterogeneity Using Factorial Hidden Markov Models

PUBLICATIONS AND TALKS

Publications:

Gonçalves, E.*, Poulos, R.C*, **Cai, Z.***, …, Robinson, P., Zhong, Q., Garnett, M., Reddel, R. (2022). Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell*, *40*(8), 835-849., * Equal contribution

Cai, Z., Poulos, R. C., Liu, J., & Zhong, Q. (2022). Machine learning for multi-omics data integration in cancer. *iSicence*, 103798.

Poulos, R. C., **Cai, Z**., Robinson, P. J., Reddel, R. R., & Zhong, Q. (2022). Opportunities for pharmacoproteomics in biomarker discovery. *Proteomics*, 2200031.

Guan, L., Tian, J., Cao, R., Li, M., **Cai, Z**., & Shen, W. (2014). Barcode-like paper sensor for smartphone diagnostics: An application of blood typing. *Analytical chemistry*, 86(22), 11362-11367.

Conference Talks:

Selected from abstract:

Cai, Z., Gonçalves, E., Poulos, R.C, Barthorpe, S., Manda, S., Lucas, N., Beck, A., Bucio-Noble, D., Dausmann, M., Hall, C., Hecker, M., Koh, J., Mahboob, S., Mali, I., Morris, J., Richardson, L., Seneviratne, A., Sykes, E., Thomas, F., Valentini, S., Williams, S., Wu, Y., Xavier, D., MacKenzie, K., Hains, P., Tully, B., Robinson, P., Zhong, Q., Garnett, M., Reddel, R. (2022). Pan-cancer proteomic map of 949 human cell lines reveals principles of cancer vulnerabilities. American Society for Mass Spectrometry (ASMS) Annual Conference

Cai, Z., Poulos, R.C, Zhong, Q., (2021). Integrating multi-omics data with biological knowledge by Transformer-based deep learning. The Australian Bioinformatics and Computational Biology Society Annual Conference.

Invited national speaker:

Cai, Z., Gonçalves, E., Poulos, R.C, Barthorpe, S., Manda, S., Lucas, N., Beck, A., Bucio-Noble, D., Dausmann, M., Hall, C., Hecker, M., Koh, J., Mahboob, S., Mali, I., Morris, J., Richardson, L., Seneviratne, A., Sykes, E., Thomas, F., Valentini, S., Williams, S., Wu, Y., Xavier, D., MacKenzie, K., Hains, P., Tully, B., Robinson, P., Zhong, Q., Garnett, M., Reddel, R. (2021). Pan-cancer proteomic map of 949 human cell lines reveals principles of cancer vulnerabilities. KCA Precision Medicine for Childhood Cancer Symposium

ADDITIONAL

- Visa Status: Australian permanent residence visa (subclass 189)
- Languages: English (IELTS 8.0), Chinese