# ZEHUI LI

github.com/Zehui127 https://zehui127.github.io zl6222@ic.ac.uk
(0044)-07342740072

## EDUCATION

**Imperial College London, Britain** *Jan. 2023 - Present*
*Ph.D. Candidate in Bioengineering and Machine Learning*

- Supervised by Prof. Guy-Bart Stan and Dr. Yiren (Aaron) Zhao
- Funded by BBSRC AI-4-EB(UKRI) for Ph.D.

**Wolfson College, University of Cambridge, Britain** *Oct. 2019 - Oct. 2020*
*MPhil of Advanced Computer Science,* ***Distinction***

- Supervised by Prof. Pietro Lio' and Prof. Simone Teufel
- Jennings Prize, 2020: They are awarded to those who with a Distinction in a University Examination

**University of Nottingham, Britain** *July 2015 - June 2019*
*BSc Hons. Data Science, Graduate with* ***First Class Honours***

- President's Excellence Scholarships, 2017: They are awarded to top achieving students

## PREPRINT

[1] Zehui Li, Akashaditya Das, William A V Beardall, Yiren Zhao, and Guy-Bart Stan. *Genomic Interpreter: A Hierarchical Genomic Deep Neural Network with 1D Shifted Window Transformer*. Preprint. 2023. arXiv: 2306.05143 [cs.LG].

[2] Zehui Li, Xiangyu Zhao, Mingzhu Shen, Guy-Bart Stan, Pietro Liò, and Yiren Zhao. *Hybrid Graph: A Unified Graph Representation with Datasets and Benchmarks for Complex Graphs*. Preprint. 2023. arXiv: 2306.05108 [cs.LG].

## RESEARCH PROJECTS

**Optimising representation learning of heterogeneous cancer data** [PDF] *Apr. 2020 - July 2020*
*Developed BIO-RGCN, an extendable framework to predict the associations between chemicals and cancers. The outputs from the model are consistent with existing medical literature.*

- A demonstration of prediction results can be accessed through Google Colab

**Adversarial Attack on State-of-the-art Question-Answering Systems** [PDF] *Dec. 2019 - Jan. 2020*
*Proposed three model-independent adversaries based on the work of to attack three deep learning based question answering systems*

- The output of models and adversarial examples can be found on GitHub

**An exploration on the optimization routines of SVI for GPs** [PDF] *Feb. 2020 - April 2020*
*Conducted an empirical study on different optimization routines of stochastic variational inference (SVI) for Gaussian proces (GPs)*

- The code for data processing and model building can be found on Colab Notebook

## WORK EXPERIENCE

**Microsoft, Shanghai** *Jan. 2021 - Nov. 2022*
*Could Engineer*
- Assist with developing large scale web applications on Azure Platform
- Developed and lead several projects:

- Magic Cube: a programmable virtual assistant using IoT technology, computer vision and NLP

- Bot Asistant: a rule-based conversational system for facilitating the work of engineers

**Huawei, Cambridge** *Sep. 2020 - Oct. 2020*
*Artificial Intelligence Research Intern*
- Worked in Huawei Technologies Research and Development office in Cambridge
- Optimized machine learning system for speech recognition using **Beam Search** with language models

**Barclays UK, Northampton** *Jun. 2018 - Sep. 2018*
*Software engineer Intern*
- Used decision tree and random forest algorithm to build Risk Model to predict **credit card delinquency**
- Developed a web application for **synchronizing the data stream** from two databases

**Aarhus University, Denmark**[Github] *Sep. 2016 - Apr. 2018*
*Research Assistant (remote)*
- Created and managed the ComputationalLitErAry Repository, an open source corpus for old Danish language
- Wrote Script for data analysis of Old Danish language

## SKILLS

| | |
|---|---|
| **Machine learning Techniques:** | Large Model Training on Multiple GPU Nodes,<br>Neural Architecture Search, Machine learning with graphs |
| **Data Science Tool Box:** | Pytorch-lightning, Pytorch, Scikit-Learn, GPy, Numpy, Pandas, Jupyter, R |
| **Biology:** | Network Biology, Brains Science, Psychology for drug use |
| **Software & Tools:** | **Web Programming:** ASP.NET, Python Django, MERN stack<br>(MongoDB, Express, React, Node)<br>**Azure:** App Service, Networking, Firewall, Cognitive Services, Bot Freamework<br>**Computing:** GPU Cluster, Linux, Windows Server |
| **Programming Languages:** | Python, Java, Javascript, C#, C, Shell |

## INTERESTING PROJECTS AND EXPERIENCE

**Smoother robot control with the variants of A* planning algorithm** [PDF] *Apr. 2020 - May 2020*
*Proposed variants of A* algorithms in order to create more predictable paths witha lower cost than sampling algorithms and potential field methods for the mobile robot*

- The code for the planning algorithm can be found on Github

**Consistency theorem for clustering** [PDF] [Poster] *Jan. 2019 - Apr. 2019*
*Studied the mathematical properties of unsupervised learning (clustering) algorithms. It reviews Kleinberg's work and propose a new property called partial consistency to describe the clustering algorithms as a whole.*

**Self-driving Car simulator using Reinforcement Learning** [Github] *Oct. 2017 - Feb. 2018*
*used Reinforcement learning algorithm to train the car to avoid the obstacles*

**Korea University, Seoul** *Jun. 2016 - Aug. 2016*
*Summer exchange student with scholarship*
- Studied the following courses:

- Brain Science
- Calculus I & II

**Aarhus University, Denmark** *Jun. 2015 - Aug. 2015*
*Summer exchange student with scholarship*
- Studied the following courses:

- Youth, Drugs and the Night-Time Economy (Psychology module) and
- Text Mining the Great Unread