Tiantian Zheng

Data Scientist | tiantian.zheng.545@gmail.com

EDUCATION

University of Toronto: *PhD Physics (GPA: 4.0/4)* *2023*
University of Toronto: *MSc Physics* (GPA: 3.97/4) *2018*
New York University Abu Dhabi: *BS Physics* (GPA: 3.98/4) *2017*

WORK EXPERIENCE

Data Scientist *Jan 2024 – Present*
Veeva Systems, Business Consulting
- Applied predictive models (e.g. XGBoost, random forest) on healthcare claims data for patient analytics projects to predict patients with undiagnosed disease or treatment needs for engagement with health care providers and clinical trials site selection.
- Leveraged semi-supervised algorithms to infer referral networks among physicians.
- Automated and standardized code libraries for patient analytics projects.
- Utilized cloud computing (AWS) resources for model training and scoring.

Graduate Research Assistant *Sept 2017 – Dec 2023*
Department of Physics, University of Toronto
- Applied stochastic and non-equilibrium physical modeling to understand transport through the nuclear pore complex – an intracellular transporter.
- Automated deployment of parallel simulations on an HPC cluster with Slurm and PBS schedulers.
- Collaborated with biology labs across universities to extract insights from microscopy data using stochastic processes and statistical mechanics theory.
- Published results through peer-reviewed articles in high-impact journals and international conferences.

PROJECTS

Bacteria tracker https://github.com/tz545/bacteria-tracker *Sept 2022 – Mar 2023*
Computer vision | Software development
- Designed a human-in-the-loop ML solution for automating cell segmentation in microscopy images.
- Improved image analysis speed by 10 times, and decreased opportunities for human error.
- Created an interactive interface using Plotly Dash to enable manual correction of model predictions while simultaneously processing and analyzing the images.
- Built a pipeline using Google Colab to fine-tune a computer vision model in PyTorch based on *UNet* to produce increasingly better model performance as more labeled images were produced during analysis.

GraphGPS EGFR detection https://github.com/tz545/EGFR-detection-GraphGPS *Aug 2023*
Graph Neural Network | Model fine-tuning
- Adapted a pre-trained *GraphGPS* graph neural network to predict the bioactivity of small molecules based on their molecular graphs, using Google Colab and Weights & Biases.

Washing away greenwashing https://devpost.com/software/washing-away-greenwashing *Mar 2023*
NLP | Dashboarding
- Designed an NLP solution for consumers to compare companies' climate claims against their actions.
- Analyzed web-scraped statements from companies using *BERT* sentiment analysis and *word2vec*.
- Built an interactive interface using Plotly Dash for easy comparison of companies by size and sector.
- Led team to 1st place in University of Toronto's Hack the MIST hackathon with 250+ competitors.

Academic jargon recommender https://github.com/tz545/academic-thesaurus *May – Jun 2022*
NLP | Recommender Systems
- Designed an NLP solution to find subject-specific synonyms in academic jargon.
- Implemented the *word2vec* model from scratch using PyTorch and trained on academic papers to generate embeddings for predicting contextually similar words.

Abstract generator https://github.com/hlml-toronto/Abstract-generator   *Apr – Dec 2021*
NLP | Language models | Recommender Systems
- Adapted PyTorch code for the *Attention Is All You Need* Transformer to build *GPT* from scratch.

PUBLICATIONS

Zheng, T. and Zilman, A., 2024. Kinetic cooperativity resolves bidirectional clogging within the nuclear pore complex. *Biophysical Journal*, *123*(9), pp.1085-1097.

Zheng, T. and Zilman, A, 2023. Self-regulation of the nuclear pore complex enables clogging-free crowded transport. *Proceedings of the National Academy of Sciences*, 120(7), p.e2212874120.

Kalita, J., Kapinos, L.E., **Zheng, T.**, Rencurel, C., Zilman, A. and Lim, R.Y., 2022. Karyopherin enrichment and compensation fortifies the nuclear pore complex against nucleocytoplasmic leakage. *Journal of Cell Biology*, 221(3), p.e202108107.

Paci, G., **Zheng, T.**, Caria, J., Zilman, A. and Lemke, E.A., 2020. Molecular determinants of large cargo transport into the nucleus. *eLife*, *9*, p.e55963.

Gu, C., Vovk, A., **Zheng, T.**, Coalson, R.D. and Zilman, A., 2019. The role of cohesiveness in the permeability of the spatial assemblies of FG nucleoporins. *Biophysical journal*, 116(7), pp.1204-1215.