Jui-Shan (Teresa) Lin

Data Scientist · Machine learning in Genetics

New York, NY

☎ 814-996-9078 | ✉ sansanlin85@gmail.com | 🌐 teresa-sansan.github.io/ | 🐙 teresa-sansan | 🔗 teresa-js-lin

Experience

Data Science Associate   *New York, NY*

NEW YORK GENOME CENTER   *May 2021 - Present*

- Developed pipelines for efficient processing, analysis, and reproducibility of large-scale genomic datasets
- Finding causal variants among highly associated variants by utilizing machine learning-trained annotations, including MCMC and transformer
- Created baseline models and performed side-by-side comparisons with other state-of-the-art methods.(Python)
- Conducted quality control and data engineering on whole-genome sequencing (WGS) data for 36k patients and 16 million variants (Slurm, Python)
- Visualized data using scatter plots, heatmaps, PCA plots, distribution plots, lollipop charts, UpSetR, and Manhattan plots (ggplot2 and matplotlib)

Research Assistant   *State College , PA*

THE PENNSYLVANIA STATE UNIVERSITY (Advisor: DR. YIFEI HUANG)   *Aug 2019 - Mar 2021*

- Thesis: Supervised Variant Prioritization and Its Evaluation
- Build a machine learning pipeline mainly with python for variant prioritization, especially in missense mutation
- Maximize the performance by comparing different statistical learning methods (i.e., logistic regression, SVM, and gradient boosting)

Project & presentation

Improved Alzheimer's disease polygenic risk scores portability using deep learning annotations   *New York, NY*

NEW YORK GENOME CENTER

- Enhanced Alzheimer's disease (AD) risk prediction in Hispanic and African Black populations by calculating risk scores using identified causal variants
- Compare DNA-level information with RNA expression (eQTL) data in microglia (AD-related cell)
- Presented poster in NIH Alzheimer's Disease Sequencing Project Program Review (Bethesda, MA, 2024) and the American Society of Human Genetics (Los Angeles, CA, 2022)

Education

M.S. in Bioinformatics and Genomics   *State College, PA*

THE PENNSYLVANIA STATE UNIVERSITY   *Aug. 2019 - May. 2021*

B.S. in Biotechnology and Bioindustry Sciences   *Tainan, Taiwan*

NATIONAL CHENG KUNG UNIVERSITY   *Sep. 2015 - May. 2019*

Skills

Programming	Python, R, Bash, LaTeX, AWK, SNAKEMAKE
Machine Learning Package	Scikit-Learn, Pytorch, TensorFlow, Pyro
High performance computing (HPC)	Linux/UNIX(slurm), VM
Others	Pipeline building, Data Engineering, Data Visualization

Honors & Awards

2019	**Dean's List**, National Cheng Kung University	*Tainan, Taiwan*
2017	**Dean's List**, National Cheng Kung University	*Tainan, Taiwan*
2016	**Taiwan representative**, Japan-Asia Youth academic exchange	*Osaka, Japan*
2016	**Dean's List**, National Cheng Kung University	*Tainan, Taiwan*