Hanyu Wang

Homepage | GitHub
Email : hanyv.wang@yale.edu, hanyv.wang@gmail.com

OVERVIEW

I am an MSc student in Epidemiology of Infectious Diseases (Quantitative Specialization) at Yale University, with a background in Statistics and Quantitative Biology. My research interests focus on applying statistical modeling to understand how environmental factors influence infectious disease transmission. I have experience in Gaussian Processes and Machine Learning, with an emphasis on probabilistic modeling and data-driven prediction. My ultimate goal is to integrate epidemiological modeling, environmental data, and advanced statistics to generate actionable insights for disease prevention and risk mitigation.

EDUCATION

- **Yale University | Master of Science**   New Haven, America
  Epidemiology of Infectious Diseases - Quantitative Specialization   *Sep 2025 - May 2026*
- **University of Toronto | Honours Bachelor of Science**   Toronto, Canada
  Statistics Major, Quantitative Biology Major, Mathematics Minor   *Sep 2022 - Jun 2025*
  Courses: *Machine Learning (R & Python), Time Series Analysis (R), Multivariate Data Building (R), Experimental Analysis (R), Population Ecology (R), Computer Programming (Python)*

SKILLS SUMMARY

- **Technical Skills:**   R (Seurat, custom pipelines), Python (Scanpy, scikit-learn)
- **Data Analysis:**   Machine Learning (Logistic Regression, Random Forest, LDA), Statistical Modeling (PCA, UMAP, GMM), Data Visualization (t-SNE, UMAP)
- **Domains:**   Epidemiological Modeling, Biomedical Informatics, Public Health Analytics

WORK EXPERIENCE

- **Single-Cell Transcriptome Sequencing Data Analysis**   Sep 2024 - Dec 2024
  Remote research project conducted in the Biological Diagnosis and Treatment Center@Xibei Hospital
  - Analyzed single-cell transcriptome data to study gene expression in enteritis, identifying 10 cell subtypes from 19 clusters for downstream therapeutic insights.
  - Built R pipelines using PCA, Louvain/Leiden clustering, and Scanpy/Seurat libraries; visualized results with t-SNE and UMAP.
  - Strengthened skills in biomedical data analysis and machine learning, supporting health informatics applications.
- **Multifactorial Compartmental Model for Infectious Disease Dynamics**   May 2024 - Aug 2024
  Research project conducted under the supervision of Prof. Jianhong Wu@York University
  - Developed an optimized SEIRS compartmental model with birth, death, and immune decay factors, incorporating stochasticity and spatial dynamics to simulate disease transmission.
  - Enhanced expertise in mathematical modeling and computational simulations, contributing to public health informatics research.
  - Collaborated with a team of four to refine model accuracy, improving technical communication skills.

ACADEMIC PROJECT

- **Identifying Autism-Specific Brain Connectivity Patterns | University of Toronto | Jan - Apr 2025** [view]
  - Built a predictive model for autism diagnosis with 92.5% accuracy, analyzing ABIDE fMRI data using PCA, UMAP, and Gaussian Mixture Models.
  - Uncovered distinct connectivity patterns, advancing knowledge in computational neuroscience and health informatics.
  - Presented findings to peers, enhancing scientific communication skills.

- **Prediction for Alzheimer's Disease Risk and Progression | University of Toronto | Sep – Dec 2024** [view]
  - Designed a machine learning diagnostics model that achieved 94.59% accuracy using Logistic Regression, Random Forest, and Naïve Bayes.
  - Improved model performance with hyperparameter tuning (grid search), k-fold cross-validation, and SMOTE for data balancing.
  - Gained proficiency in feature engineering and optimization, applicable to biomedical data challenges.

PUBLICATIONS

- **Reactivating T cell immunity in Wnt-hyperactivated non-small cell lung cancer through a supramolecular droplet of carnosic acid and peptide**
  Na Liu, Y. Tu, **Hanyu Wang**, X. Zheng, F. Ji, M. Geng, X. Wei, J. Xin, W. He, Q. Zhao, T. Liu.
  Journal of Pharmaceutical Analysis, DOI: 10.1016/j.jpha.2025.101309, 2025