Jiaming Huang

Tel: (+86) 158-6737-0916 Web: http://www.huangjiaming.cn E-mail: Jiaming.Huang21@student.xjtlu.edu.cn

EDUCATION

Xi'an Jiaotong-Liverpool University (XJTLU) | Suzhou, China Sept 2021 - Jun 2025 (expected)

- Bachelor of Science in Bioinformatics
- **GPA: 3.97**/4.00 (Rank **1st** out of 72; **WES** authorized)

University of Liverpool (UoL) | Liverpool, UK Sept 2021 - Jun 2025 (expected)

- Bachelor of Science in Bioinformatics (Hons)
- **Grade: 85**/100 (>70: First Class)
- **Core Courses (all GPA of 4.0):** Cellular and Molecular Biology, Evolution and Genetics, Bioinformatics I & II, Programming in Java, Databases, Statistics, Artificial Intelligence

PUBLICATION

[1] **Huang, J.**, Wang, X., Xia, R., Yang, D., Liu, J., Lv, Q., Yu, X., Meng, J., Chen, K., Song, B., & Wang, Y. (2024). Domain-knowledge enabled ensemble learning of 5-formylcytosine (f5C) modification sites. *Computational and Structural Biotechnology Journal*, *23*, 3175–3185. https://doi.org/10.1016/j.csbj.2024.08.004

[2] Xia, R., Yin, X., **Huang, J.**, Chen, K., Ma, J., Wei, Z., Su, J., Blake, N., Rigden, D. J., Meng, J., & Song, B. (2024). Interpretable deep cross networks unveiled common signatures of dysregulated epitranscriptomes across 12 cancer types. *Molecular Therapy - Nucleic Acids*, 102376. https://doi.org/10.1016/j.omtn.2024.102376

[3] Chen, Z., Liu, Y., Cao, S., **Huang, J.**, Wang, X., Zhong, W., & Xiao, Y. (2024) m6A-CAPred: Domain-characteristics enabled machine learning of cancer-associated N6-methyladenosine (m6A) sites. *Evolutionary Bioinformatics* (under review). [paper]

RESEARCH EXPERIENCE

Abbreviations: *FYP: Final Year Project; ***SURF:** Summer Undergraduate Research Fellowship

Gene regulatory network (GRN) inference with L1 regularization

Institute: University of Massachusetts Chan Medical School Sep 2024 - Present

Advisor: Lingfei Wang, Assistant Professor

- Leveraged scATAC-Seq and scRNA-Seq data to develop a GRN model with L1 regularization to sparsify the edge weight matrix, achieving the identification of transcription factors with a genuine regulatory role
- Mastered two bioinformatics tools ("Seurat" and "Scanpy") to analyze and visualize scRNA-Seq data based on R and Python languages
- Applied probabilistic programming with "pyro" for graphical probabilistic model implementation to simulate gene expression under complex environments

(FYP) Ultrasound (US) image-guided generation of shear wave elastography (SWE) images

Institute: Xi'an Jiaotong-Liverpool University Sep 2024 - Present

Advisor: Shuihua Wang, Associate Professor

- Implemented a latent diffusion model to generate specific high-resolution SWE images guided by corresponding ultrasound images through ControlNet, leveraging US-SWE paired results offered by SWE diagnosis
- Wrote the variational autoencoder and diffusion denoising probabilistic model from scratch based on my comprehensive understanding of underlying mathematics such as variational inference and ELBO
- Enhanced breast tumor malignancy predictions by incorporating generated SWE images during the prediction model training phase while also boosting diagnostic trust for doctors by generating SWE images corresponding to patients' US images during the diffusion model application phase

(SURF) Ensemble learning of breast cancer malignancy classification with SWE images

Institute: Zhongshan Hospital, Fudan University Jun 2024 - Aug 2024

Advisor: Zhaoyang Liu, External Mentor

- Applied four machine learning approaches and nine deep learning approaches, such as vision transformer, to predict breast tumor malignancy using four types of images cropped from the original diagnosis report to include both regional and peripheral tumor features indicative of metastatic potential

- Used CAM-based approaches, including Grad-CAM, Grad-CAM++, Layer-CAM, and Score-CAM, to provide a visual interpretation and enhance model reliability
- Presented project as a poster in the 2024 SURF Poster Fair, earning the "Student-Nominated Best Poster" out of 414 posters

5-formylcytosine (f^5C) modification sites prediction

Institute: Xi'an Jiaotong-Liverpool University Jun 2023 - Aug 2024
Advisor: Jia Meng, Head of Department

- Extracted 3784 base-resolution f^5C sites from yeast mRNAs and implemented conventional machine learning techniques such as random forest and naive bayes and novel deep learning methods such as ResNet, Transformer, and LSTM alongside sequence information encoded with one-hot and nucleotide density
- Enhanced AUC from 0.7492 to 0.8391 by integrating genomic features to enrich input and combining ensemble approaches to innovate on the perspective of methodology
- Explained the model performance improvement by using shapley additive explanations (SHAP) to evaluate each genomic feature's contribution to the final performance, identifying several genomic features shared across the top 3 AUC models, such as the GC composition of the 101 bp region around the site, potentially indicating the elusive f^5C forming mechanism
- Published the paper in the *Computational and Structural Biotechnology Journal (CSBJ)*

(SURF) Deep learning enables cross-state image restoration

Institute: Xi'an Jiaotong-Liverpool University Jun 2023 - Aug 2023
Advisor: Pengfei Fan, Assistant Professor

- Used the outstanding representation power of a deep learning-based framework for cross-state image restoration through a flexible MMF, transforming MMFs into practical imaging components or communication devices
- Presented project as a poster in the 2023 SURF Poster Fair

Cancer-promoting N6-methyladenosine (m^6A) sites prediction

Institute: Xi'an Jiaotong-Liverpool University Dec 2022 - Jun 2023
Advisor: Jia Meng, Head of Department

- Labeled 111,937 experimentally validated m^6A sites that were associated with cancer development and developed a prediction model using SVM with AUROC of around 82%
- Implemented several models, including random forest, naive bayes, and decision tree, and determined that the SVM model outperformed in terms of accuracy, and thus selected it for the final m6A-CAPred model
- Validated our predicted cancer-promoting m^6A sites through SNP density analysis

Conservation analysis of RNA modifications in mammals

Institute: Xi'an Jiaotong-Liverpool University Oct 2021 - Dec 2022
Advisor: Jia Meng, Head of Department

- Developed a novel scoring framework for conservation analysis of mammalian epitranscriptome through integrating six unique resources, including deep learning, tissue-specificity, sequence similarity, positional mapping, genome conservation, and support amounts from multiple studies

SOFTWARE & DATABASE

- 2024 SWEBreCA-Pred: a website for breast tumor malignancy prediction based on shear wave elastography and ultrasound images
- 2023 Resf5C-Pred: a website for f^5C site prediction with two modes, namely "Resf^5C-only" and "Ensemble (Resf5C+XGB)"
- 2022 m6A-CAPred: a website for cancer-promoting m^6A site prediction with a database of 111,937 high-confidence m6A sites annotated with cancer context labels for further analysis or model development

HONORS

- 2023-2024 China National Scholarship (Awarded to top **3** students across the university)
- 2023-2024 Jiangsu Province "Merit Student" (Awarded to top **15** students across the university)
- 2023-2024 "Big Data Analysis" Kaggle Competition "First Place Award" and "Best Innovation Award" (rank **1st**/300)
- 2023-2024 "JITRI Cup" XJTLU Global "Dream-Chasers" Entrepreneurial Competition Second Prize (rank **2nd**/30)

- 2023-2024 XJTLU SURF “Student-Nominated Best Poster” Winner (top **15**/414)
- 2021-2022, 2022-2023 China National Encouragement Scholarship (top **3**%)
- 2021-2022, 2022-2023, 2023-2024 XJTLU Academic Excellence Award (top **5**%)
- 2021-2022, 2022-2023, 2023-2024 XJTLU “Outstanding Student” (top **3**%)
- 2021-2022, 2022-2023, 2023-2024 XJTLU Student Representative & Ambassador (top **1**%)
- 2022-2023, 2023-2024 XJTLU “i-star” Award (Encouragement) (Awarded to top **6** students across the university)

ACTIVITIES

“Big Data Analysis” Kaggle Competition (rank 1st/300) [awards] [photo] Dec 2024

- Predicted students' dropout rates and academic success with the highest accuracy in the "Big Data Analysis" Kaggle Competition by utilizing machine learning and deep learning approaches such as 1d-Transformer and LSTM. Awarded the "First Place Award" and the "Best Innovation Award"

Global “Dream-Chasers” Entrepreneurial Competition (rank 2nd/30) [award] [photo] Dec 2024

- Competed against global teams and ranked 2nd out of 30 in the “AI + Health” track during the finals. Collaborated with Haier (Fortune Global 500), proceeding to validate our breast cancer prediction software “SWEBreCA-Pred” at Haier’s affiliated hospitals in Shanghai

NVIDIA DPU Programming Course [certificate] Oct 2024

- Leveraged opportunities at the NVIDIA China Developer Day event held at XJTLU by attending an in-person course on DPUs taught by senior NVIDIA developers. Learned that DPUs are designed to offload tasks related to data transmission from the CPU for specialization, and practiced deploying the DOCA framework on the BlueField DPU

UCLA Data Science Winter School (A+ score) [certificate] [transcript] Jan 2023 - Feb 2023

- Gained expertise in statistics and machine learning techniques and honed expertise using programming languages such as Python and R

SKILLS

Data Analysis: Machine learning (SVM, Naïve Bayes, Decision Tree) / Deep Learning (CNN, Transformer, ResNet, LSTM) (Pytorch / Tensorflow) / Tableau

Programming Languages: R / Python / Java / MATLAB / SQL / JavaScript / HTML / CSS / PHP

Wet Experiment Techniques: SDS-PAGE Electrophoresis / Ion Exchange Chromatography / Measurements at the molecular level / Precipitation Techniques (metallic ions, ethanol, etc)