

Shubo Li

ethanli041012@gmail.com | <https://shubo176.github.io>



EDUCATION

Wuhan Polytechnic University, Bioinformatics

Bachelor of Science

Wuhan, China

09/2022 – 06/2026 (expected)

• **GPA:** 3.37/4.00 **Rank:** 24/67 **CET-6:** 591/710

• **Major Courses (100-point scale):** Bioinformatics (89/100), Fundamental Bioinformatics Lab (98/100), Transcriptomics & Proteomics (90/100), Integrated Lab in Transcriptomics & Proteomics (85/100), Metabolomics & Epigenomics (96/100), Integrated Lab in Metabolomics & Epigenomics (86/100), Python Programming (95/100), Python Programming Lab (94/100), Database Fundamentals & Data Mining (89/100), Machine Learning Algorithms (85/100), Probability & Statistics (88/100), Comprehensive Applications of Data Processing Software I & II (89/100 each), Protein Structure & Function Lab (90/100), Biochemistry III (86/100)

RESEARCH EXPERIENCE

Protein Melting Temperature Prediction (Protein Tm) | Tianjin Institute of Industrial Biotechnology, CAS

08/2024 – Present

- Built a multimodal Tm regressor fusing **ESM embeddings (1280-d)** and **SaProt structural embeddings (1280-d)** with **Rosetta ΔG** , Shannon entropy, and OGT priors.
- Curated and CD-HIT deduplicated **28,463 proteins**; trained on **pre-2023** data and evaluated on an **independent post-2023 set (75 proteins)** for temporal extrapolation.
- Three-layer **MLP** (ReLU + Dropout); **MC Dropout** for uncertainty; hyperparameters tuned via **Optuna**.
- Achieved **RMSE = 9.37 °C**, outperforming five public baselines; stability verified by 10 repeats of 10-fold cross-validation (100 total fits).
- *Project page:* shubo176.github.io/project-protein-tm.html

DNA Phosphorothioate Modification Detection (Nanopore + Transformers) | Wuhan Polytechnic University

12/2023 – Present

- Developed an end-to-end pipeline from nanopore fast5 raw currents to site-level predictions for three *Salmonella* cohorts (xtg103, plasmid-full, sal87-partial).
- Performed event alignment, normalization, and imputation; constructed **21×36** tensors combining DNA one-hot and ionic current segments.
- Implemented a **6-layer Transformer encoder** with mean pooling and a two-layer MLP producing **two-class logits**.
- Trained with **CrossEntropyLoss**, optimized using **AdamW**, and scheduled by **ReduceLROnPlateau** with early stopping; tracked accuracy, precision, recall, F1, and AUC each epoch.
- Recovered a **~20%** modification rate on the independent **sal87** set, consistent with wet-lab observations.
- *Project page:* shubo176.github.io/project-nanopore.html

Database System for Steroid Drug Precursors | Wuhan Polytechnic University

12/2023 – 06/2024

- Project lead: designed **MySQL** schema and backend; built a **Django**-based web system for multi-criteria query, visualization, and submissions.
- Led a 5-member team; won the **Third Prize** in the 2024 Hubei Undergraduate Life Sciences Competition.

HONORS & AWARDS

- Third Prize, Hubei Undergraduate Life Sciences Competition (Team Lead), 2024
- Second-class Academic Scholarship, Wuhan Polytechnic University, 2024
- Excellent Student Award; Excellence in Study Award; Research Innovation Award, 2024
- “Understanding Contemporary China” Speech 1st Prize; Recitation 2nd Prize, 2023
- National English Competition for College Students, Second Prize (Category C), 2023