Curriculum vitae

Chunzhuo Zhang age: 26 zhangchunzhuo@aaib.pku.edu.cn zczxiaochui@163.com

Work Experience

Tel: 18035508393

*** Research Institute | Biological Foundation Model Design Oct 21, 2024 — Present

- **Single cell foundation model: in-context learning single cell foundation model (53m, 300m,600m, 2 B) Training model with observation and perturbation data**
  1. Design model architecture, modeling cells on group level
     From data preprocessing to model design, to downstream benchmark (based on scGPT) , The model performs better than scGPT , Transcriptformer, UCE.
  2. In-context learning in single cell data, explore scaling law in single cell data.
  3. Model implementation: PyTorch → PyTorch Lightning. Attention acceleration with FlashAttention and FlexAttention.
     Loss function design and multi-classification head design.
  4. Built downstream multi-task evaluation pipeline: batch-effect correction, cell classification, perturbation prediction; designed new tasks for single-cell foundation models.
  5. Model parallel acceleration: FSDP, data parallelism, checkpointing.
  6. Mentored interns to deliver independent projects.

- **BindCraft protein design: used BindCraft to design binder proteins; ~5% of designs exhibited strong binding affinity.**
  Designed antigen ligands with constrained binding sites and generated corresponding binders via BindCraft, in collaboration with the experimental team.

Education

Wageningen University | M.Sc. in Bioinformatics GPA: 7.5/10 Sep 1, 2022 — Aug 21, 2024

Course grades: Computational Biology (8.5/10), Molecular Systems Biology (7.5/10), Machine Learning (7.5/10), Advanced Statistics (8/10)

Deep Learning (8/10), Software Engineering (7.5/10), Genomics (8/10)

- **Protein structure and sequence embeddings (sgBERT, gBERT): combined protein structure and sequence via attention to obtain protein embeddings; preliminary exploration of multi-modal fusion.**
  Independent design: protein structure representation, Transformer model implementation, evaluated on protein structure classification.
  ChunZhuo/sgBERT: Two BERT-like models for protein structure and protein structure with sequence
- **Protein–protein interface false-positive prediction: applied harmonic functions to enforce equivariance, integrated into a graph neural network to model interaction interfaces, and predicted whether a binding event is a false positive.**
  Benchmarked protein-interaction prediction models on the alanine-scanning task, from classical docking software (HADDOCK) to machine-learning models (linear regression, random forest) to graph neural networks (DeepRank-GNN).
  Independent design: extended DeepRank-GNN by incorporating an equivariant graph neural network.
  ChunZhuo/SEGIN: Add SE feature into DeepRank
- Coursework: sequence assembly → differential gene expression → protein domain analysis; plant GNSR gene analysis.
  Gene regulatory network analysis.
  Java software development: parking-lot entry system.

Northwest A&F University | B.Sc. in Horticulture (Agriculture) GPA: 3.51 Sep 1, 2018 — Jul 30, 2022

Course grades: Advanced Mathematics (91/100), Linear Algebra (94/100), Probability

Theory (92/100), Fundamental Biochemistry (91/100), Organic Chemistry (94/100)

- **Kiwifruit shoot-tip cryotherapy for virus elimination.**
  RT-PCR; plant tissue culture.

Skills

IELTS: 7.0 (Speaking 7.0)   GRE: 320

Python: pytorch, pytorch lightning,   R, Java, C++