Heng Yang

☎+44-07878711663 — Email: hy345@exeter.ac.uk — GitHub: yangheng95
Webpage: yangheng95.github.io — Huggingface: yangheng — GScholar: Heng Yang
University of Exeter, Exeter City, UK

Objective

Seeking a research position in AI and NLP, leveraging expertise in biological sequence modeling, sentiment analysis, and adversarial attacks to drive innovation and contribute to impactful research projects.

Education

- **University of Exeter** — *Sept. 2021 - Sept. 2025 (Expected)*
  PhD in NLP/Genomics Modelling — Exeter, UK
- **South China Normal University** — *Sept. 2018 - June 2021*
  Master's in NLP — Guangzhou, China
- **Yangtze University** — *Sept. 2014 - June 2018*
  Bachelor's in Computer Science — Jingzhou, China

Awards

- **PhD Studentship with Research Grant** — University of Exeter, 2021-2025
- **National (Academic) Graduate Scholarship** — South China Normal University, 2020
- **Outstanding Graduate** — Yangtze University, 2018

Publications

- **OmniGenBench: Automating Large-scale Benchmarking for Genomic Foundation Models**
  Heng Yang, *Jack Cole, Ke Li (Arxiv 2024)*
- **OmniGenome: Aligning RNA Sequences with Secondary Structures in Genomic Foundation Models**
  Heng Yang, *Ke Li (Arxiv 2024)*
- **PlantRNA-FM: An Interpretable RNA Foundation Model for Exploration Functional RNA Motifs in Plants**
  Haopeng Yu#, ***Heng Yang#****, et al. (Co-first Author, Nature Machine Intelligence, 2024)*
- **Modeling Aspect Sentiment Coherency via Local Sentiment Aggregation**
  Heng Yang, *Ke Li (EACL 2024)*
- **The Best Defense is Attack: Repairing Semantics in Textual Adversarial Examples**
  Heng Yang, *Ke Li (EMNLP 2024)*
- **PyABSA: A Modularized Framework for Reproducible Aspect-based Sentiment Analysis**
  Heng Yang, *Cheng Zhang (CIKM 2023)*
- **Evolutionary Multi-objective Instruction Optimization via Large Language Model-based Instruction Operators**
  Heng Yang, *Ke Li (EMNLP 2023)*
- **DaNuoYi: Evolutionary Multi-Task Injection Testing on Web Application Firewalls**
  Ke Li, ***Heng Yang*** *(IEEE TSE, 2023)*
- **Boosting Text Augmentation via Hybrid Instance Filtering Framework**
  Heng Yang, *Ke Li (ACL 2023)*
- **A Multi-task Learning Model for Chinese-oriented Aspect Polarity Classification and Aspect Term Extraction**
  Heng Yang, *Biqing Zeng (Neurocomputing 2021)*

Personal Open-source Projects

- **OmniGenBench** – 4.7**k Installations** — *Sept. 2024 - Present*
  First large-scale in-silico benchmarking framework for genomic foundation models — GitHub
- **PyABSA** – 340**k Installations, 1k GitHub stars** — *June 2020 - Present*
  Aspect-based sentiment analysis framework, serving commercials, developers and scholars — GitHub

Community Contribution Statistics

- **GitHub:** 1.5**k stars,** ≈ 160 **commits and** 60 **PRs/year**. I am a open-source lover since the beginning of my research. I am grateful to the developers whose projects helped me a lot. Therefore, I am committed to share all my open-source projects on GitHub with friendly MIT lenience.
- **Huggingface:** 8 **Models and** 10 **Spaces with** ≈ 400**k downloads and** 10**k access, respectively**. Thanks to the Hungingface platform, I am glad to share all my pretrained and fine-tuned state-of-the-art sentiment analysis and genomic foundation models. e.g., deberta-v3-base-absa-v1.1 and OmniGenome-186M. Moreover, I have been to releasing demos for low-resource research topics, like RNA secondary structure prediction and RNA Design.
- PyPi: 8 **python wheels with** 800**k downloads.** I have made efforts to simplify the workflows and pipelines by packing and releasing open-access Python wheels. My wheels have been widely used in by the community. These wheels can be easily distributed via PyPi and installed via the pip tool.