



个人主页

杨恒 PhD

+86-19842750173 | hy345@exeter.ac.uk | +44-7878711663

LinkedIn | GitHub | Huggingface | 谷歌学术

Department of Computer, University of Exeter, EX4 4RN, United Kingdom



最新版本

研究摘要

- 语言模型:** 六年预训练语言模型 (PLM) 架构与微调经验; 主导研发 DeBERTa-v3-base-ABSA-v1.1 (下载量 >100 万) 及多款 domain-adapted 模型。技术栈涵盖预训练, 数据增强/合成 (BoostAug)、文本的语义攻击/防御; 对抗鲁棒性与公平性。
- 情感分析:** PyABSA (GitHub 1k+stars) 创建者与维护者, 一键调用 31+ 个 ABSA 模型与 30+ 个数据集。提出的 LSA 模型连续三年 SOTA; 获 ChatGPT 及 DeepSeek 等 AI 助手推荐。
- 大语言模型:** 提出 InstOptima, 首创基于大模型驱动交叉/变异算子的多目标指令进化框架, 协同优化模型性能、困惑度与简洁性; 熟悉 LLM 预训练/后训练技术体系。
- 基因组基础模型 (GFM):** OmniGenome-186M / 2B 与 OmniGenBench 技术负责人; 将 EternaV2 RNA 设计成功率从 3% 提升至 84%; 框架无需修改支持绝大多数 Transformer 和 MoE 模型, 集成 LoRA、与 FlashAttention 以支持高效微调和应用。

教育与荣誉

- University of Exeter, 计算机科学博士
预训练模型 (PLM), 大语言模型 (LLM) 与基因组模型 (GFM) 等
全额博士奖学金; PhD Research Grant
Exeter, UK; 2021 年 9 月 - 2025 年 9 月
- 华南师范大学, 计算机科学硕士
预训练模型 (PLM), 情感分析 (ABSA) 等
研究生国家奖学金; 三次一等学业奖学金
广州, 中国; 2018 年 9 月 - 2021 年 6 月
- 长江大学, 计算机科学学士
数据结构, 计算机系统架构等, GPA: 3.8/4.0
优秀本科生毕业设计; 优秀本科毕业生
荆州, 中国; 2014 年 9 月 - 2018 年 6 月

开源项目

- 发布 17 个 Repositories · 1.5k GitHub Stars · 194 Followers · 500k+ LOC · 8 个 PyPI 包 · >100 万下载
- PyABSA** (pip 安装量: >500k · 1k+ stars)
广泛使用的开源情感分析工具: 单行代码调用 31+ 个模型与 30+ 个基准数据集, 支持快速微调, 部署和推理。预计将情感分析系统的原型搭建验证时间缩短 >80%。统一的模型和任务和封装, 提供易用的 API, 内置自动增强等实用特性, 方便用户入门。PyABSA 是 ChatGPT 与 DeepSeek 等主流 AI 助手首推的新手友好的开源情感分析框架。
- OmniGenBench** (pip 安装量: 53k · 351 stars)
首个基因组基础模型基准测试和应用统一平台: 单行代码完成 123 项下游任务评估; 提供模型、数据、指标的模块化模板, 所有模块均支持轻松扩展, 无需修改源代码; 原生支持 LoRA、MoE 与 FlashAttention。首个封装常用基因组任务作为 pipeline 的框架, 用户无需深入学习即可将一键 GFM 用于实际基因组任务, 例如降解率、翻译效率预测, RNA 序列结构预测和序列设计优化。

HuggingFace

- 发布 9 个开源模型 · 单模型 >100 万下载量 · 10 个 HF Space · 23 Followers · >150 Model/Space 收藏量
- yangheng/deberta-v3-base-absa-v1.1**
Top-1 的 ABSA 模型 (据 PapersWithCode), 月下载约 4 万; 被斯坦福大学 2022 年度 AI 指数报告 (p. 83) 引用
2022 年 9 月
下载量: 128 万
- yangheng/OmniGenome-186M**
首个 RNA 结构-序列对齐模型; RNA 二维结构预测与设计; 将 GFM 的序列设计求解成功率从 3% 提升至 84 %
2024 年 6 月
下载量: 6.7 万
- Gradio-Blocks/PyABSA Space**
全面的 ABSA 交互式演示中心; 为新手提供不同情感分析任务的快速演示; 收录为 Gradio-Blocks 官方推荐演示
2022 年 9 月
访问量: 14.2 万

主要论文

- 发表 11 篇 first-author papers · 1×Nature Machine Intelligence · 3×CCF A · 4×CCF B · 1×CCF C · Citation: 610
- Bridging Sequence-Structure Alignment in RNA Foundation Models**
Heng Yang, Ke Li
CCF A
AAAI 2025
- BoostAug: Boosting Text Augmentation via Hybrid Instance Filtering Framework**
Heng Yang, Ke Li
CCF A
ACL 2023
- InstOptima: Evolutionary Multi-objective Instruction Optimization via LLM-based Instruction Operators**
Heng Yang, Ke Li
CCF B
EMNLP 2023
- MPRNA: Unleashing Multi-species RNA Foundation Model via Calibrated Secondary Structure Prediction**
Heng Yang, Ke Li
CCF B
EMNLP 2024
- PlantRNA-FM: An Interpretable RNA Foundation Model for Exploration Functional RNA Motifs in Plants**
Haopeng Yu#, Heng Yang#, et al. (共一)
Journal
Nat. Mach. Intell.

竞赛组织经历

- ECML-PKDD 基因组基础模型探索挑战赛 (2024-2025)**。担任竞赛 Chair, 与 ECML-PKDD2025 竞赛 Chairs 联动, 协调 Co-chairs 共同完成评审与赞助对接; 主持建立了首个多物种、多模态的 GFM 基准 OmniGenomic Benchmark (OGB), 涵盖 7 项任务 (序列预测、结构建模与功能推断); 编写代码, 设计了全自动评测流水线与 CodaBench 排行榜。总计 20 人参赛, 有效提交 227 份。

学术报告

- Google Health, 基因组学团队: OmniGenome
2025 年 9 月
- Mila 研究所 (魁北克人工智能研究所), 多组学团队: OmniGenBench
2024 年 12 月

♥ 研究经历

- **基因组基础模型 (GFM)**。主导 AI4Science 领域的预训练研究，专注于基因组基础模型与遗传密码表示；为应对基因组数据的序列稀疏性及单核苷酸变异 (SNP/SNV) 等独特挑战，将单个碱基作为基本语义单元处理，RNA 结构预训练，以及序列-结构对齐预训练，从而显著提升了 RNA 结构预测 (>20% 提升) 和序列设计 (> 80% 提升) 等任务的性能；独立完成了从数据收集与处理、模型架构设计到训练脚本和基准测试创建的端到端研究流程；主导产出一系列论文，其中三篇已发表在 Nature Machine Intelligence, AAAI 2025, EMNLP 2024 等会议。基于已有模型和代码，设计开发了首个完整的开源框架，OmniGenBench；它抽象了模型，分词器，数据集和 metric 等模块，提供接口，教程和文档方便入手；可以让新手快速将 GFM 应用到基因组序列处理，支持可扩展性设计的模块化基准测试框架，其特色是集成了数据集、模型与评估指标的统一生态。
- **大语言模型 (LLM)**。开创性地将遗传算法应用于 LLM 的多目标指令进化，采用了新颖的指令交叉与变异算子。该工作是首批系统性解决指令清晰度、简洁度与模型性能等竞争性多目标目标优化的研究之一，首次投递被 EMNLP 2023 接收；未来改进可聚焦于几个关键领域。遗传操作符本身可进一步优化，例如借助另一 LLM 的语义理解能力，实现更具上下文意识的交叉和变异操作。可扩展性是一个重大挑战，因为评估每个进化指令的适应度计算量巨大；研究更高效的评估代理或替代适应度函数将极具价值。最后，该框架可扩展以纳入更广泛的目标，例如 Reward，公平性、创造力，或增强指令集对抗对抗性攻击的 robustness。
- **情感分析 (ABSA)**。独立完成细粒度、面向方面的情感分析研究，涵盖情感三元组与四元组提取等任务。撰写并开源了多个模型，这些模型在 Hugging Face 上取得了的关注，并成为该平台应用最广泛的情感分析模型之一。所有工作均整合于 PyABSA 工具包，该工具包已成为 ChatGPT, Gemini, Claude 与 DeepSeek 等 AI 助手推荐的首选框架。
- **文本增强**。独立开发了一种适用于多种文本建模任务的新型文本增强技术。该方法利用全局特征分布 (偏度)，并结合困惑度、置信度和硬标签等指标来筛选合成样本。作为独立的 PyPI 工具发布后，已获得 15 万次安装。在全部八个公开数据集上的实验结果验证，该技术是一项普适技术，能在全部实验中为文本分类和情感分析任务中带来 1%-2% 的性能提升。稳定性和有效性相比现有文本增强方法显著增强 (相比于 EDA, NLPAug 等方法)。
- **对抗性攻击与防御**。探究了针对因缺乏对抗性训练而存在脆弱性的预训练模型的文本对抗攻击与防御策略。通过分析常见的对抗模式，提出了一种基于二次扰动的 SOTA 防御技术，该技术融合了对抗样本检测与文本修复功能。
- **软件缺陷预测 (Code Defect Prediction)**。提出 LMDP 框架，首次将预训练语言模型用于跨项目与项目内缺陷预测，配合“损坏代码检测”多任务目标以缓解特征稀疏与定位粒度不足问题。实验覆盖 10×CPDP 与 13×WPDP 基准，平均 AUC 提升 4.6%、F1 提升 5.3%，并能精确行级定位缺陷源，显著优于 AST / GNN-类方法。框架无需 AST，即插即用，适配 CodeT5、CodeBERT 等多款 PLM。
- **鲁棒性 & 公平性协同**。首次系统性探索公平性无关对抗攻击对 PLM 偏差的影响，构建三阶段评估流水线 (Clean→Attack→Defend) 并在 3 个真实数据集上验证。发现攻击虽降低准确率，却稳定削弱群体差异 (SPD 降 0.07-0.14)；进一步的对抗训练在恢复最高 40% 准确率同时进一步改善公平性，揭示鲁棒性与公平性的潜在“帕累托前沿”。
- **Web Application Firewall (WAF) 注入安全测试**。提出 DaNuoYi，首个支持多任务、多语言注入的自动化 WAF 测试框架：利用 15 对注入翻译模型共享语义，再以多任务进化算法 (共享交配池 + 6 种语义保持变异算子) 协同生成 SQLi/XMLi/PHPi/OSi/XSSi/HTMli 六类注入用例。在 ModSecurity、Ngx-Lua-WAF、Lua-Resty-WAF 三款真实 WAF 上，平均发现的绕过用例数提升 3.8×-5.78×，显著优于 SQLMap 及单任务基线。
- **开源框架开发和项目维护**。在开源项目管理与科学软件架构方面拥有丰富的实践经验，主导了多个项目的全生命周期，涵盖从架构设计、软件实现、到长期维护与社区运营的全过程。
 - 框架设计: 核心设计哲学强调模块化、可扩展性与接口抽象。例如，在 OmniGenBench 的设计中，采用核心代码模块化，扩展功能插件式架构，允许研究者无需修改核心代码即可无缝集成自定义模型、数据与评估指标。在 PyABSA 中，通过高度封装的 API 将复杂的模型调用简化为单行指令，显著降低了用户的使用门槛。
 - 软件实现与项目管理: 具备领导复杂项目从概念到交付的能力。担任 ECML-PKDD 2025 基因组基础模型挑战赛主席，全权负责竞赛的策划与管理，包括协调多方合作、设计全自动评测流水线、以及对接赞助。该经历证明了在多方利益相关者环境中进行高效项目管理与技术执行的能力。所有项目均注重持续集成、详尽的文档 (如 ReadTheDocs) 和积极的社区支持，以确保软件的长期可用性与健壮性。

📖 完整论文列表 以下为完整论文列表，包括了正文中选取的五个出版物。

| | |
|--|---|
| • PlantRNAFM: An Interpretable RNA Foundation Model for Exploration Functional RNA Motifs in Plants | Journal |
| Haopeng Yu#, Heng Yang#, et al. (共同第一作者) | Nat. Mach. Intell. |
| • Bridging Sequence-Structure Alignment in RNA Foundation Models | CCF A |
| Heng Yang, Ke Li | AAAI 2025 |
| • MPRNA: Unleashing Multi-species RNA Foundation Model via Calibrated Secondary Structure Prediction | CCF B |
| Heng Yang, Ke Li | EMNLP 2024 |
| • The Best Defense is Attack: Repairing Semantics in Textual Adversarial Examples | CCF B |
| Heng Yang, Ke Li | EMNLP 2024 |
| • InstOptima: Evolutionary Multi-objective Instruction Optimization via LLM-based Instruction Operators | CCF B |
| Heng Yang, Ke Li | EMNLP 2023 |
| • BoostAug: Boosting Text Augmentation via Hybrid Instance Filtering Framework | CCF A |
| Heng Yang, Ke Li | ACL 2023 |
| • DaNuoYi: Evolutionary Multi-Task Injection Testing on Web Application Firewalls | CCF A / Journal |
| Ke Li, Heng Yang, Willem Visser | IEEE Transactions on Software Engineering |
| • Modeling Aspect Sentiment Coherency via Local Sentiment Aggregation | |
| Heng Yang, Ke Li | EACL 2024 |
| • PyABSA: A Modularized Framework for Reproducible Aspect-based Sentiment Analysis | CCF B |
| Heng Yang, Chen Zhang, Ke Li | CIKM 2023 |
| • A Multi-task Learning Model for Chinese-oriented Aspect Polarity Classification and Aspect Term Extraction | CCF C / Journal |
| Heng Yang, Biqing Zeng | Neurocomputing |
| • OmniGenBench: Automating Large-scale Benchmarking for Genomic Foundation Models | Preprint |
| Heng Yang, Jack Cole, Ke Li | ArXiv Preprint |
| • Robustness Meets Fairness: Investigating Adversarial Attack Effects on Alleviating Model Bias | Preprint |
| Heng Yang, Geyong Min, Ke Li | OpenReview Preprint |

附录：开源项目与模型

以下附录内容如无兴趣，您不必浏览

本附录旨在详细梳理本人在科研工作期间所主导的原创开源项目、公开发布的预训练模型及相关软件工具。内容涵盖了从底层框架、算法实现到交互式演示的多个层面，并标注了相关成果的发表情况或当前状态。部分附录内容可能与正文重复，如您选择酌情阅读，敬请忽略。

🔗 GitHub 开源项目 (<https://github.com/yangheng95>)

- **PyABSA**: 细粒度情感分析旗舰框架，支持 30+ 数据集与 31+ 预训练模型一键调用。 [CIKM 2023]
- **ABSADatasets**: ABSA 与文本分类公共数据集仓库。 [CIKM 2023]
- **OmniGenBench**: RNA/DNA 基础模型应用，自动化基准与评测流水线。 [arXiv 2024]
- **AdvFairness**: 对抗鲁棒性与模型公平性综合评测框架。 [OpenReview 2024]
- **OmniGenome-Demo**: OmniGenome 在线演示与教程脚本。 [AAAI 2025]
- **Rapid**: 文本对抗性攻击与防御实验集合。 [EMNLP 2024]
- **PlantRNA-FM**: 植物 RNA 基础模型相关代码。 [Nat. Mach. Intell. 2024]
- **MP-RNA**: 多物种 RNA 基础模型训练脚本。 [EMNLP 2024]
- **InstOptima**: 指令多目标进化算法实现。 [EMNLP 2023]
- **BoostTextAugmentation (BoostAug)**: 文本增强方法实现与数据生成包。 [ACL 2023]
- **LCF-ATEPC**: 面向中文 ABSA 的多任务学习模型。 [EACL 2024]
- **CodeT5DefectDetection**: 代码缺陷检测数据与模型实验。 [TSE 2025, in press]
- **DaNuoYi**: 针对 WAF 的多任务演化注入测试工具。 [TSE 2024]
- **EMOO**: 演化多目标优化基础代码框架。
- **findfile**: 轻量级文件系统快速检索工具。
- **autocuda**: CUDA 设备自动选择与管理脚本。
- **metric-visualizer**: 研究指标可视化与日志统一接口。

🤖 Hugging Face 模型 (<https://huggingface.co/yangheng>)

- **deberta-v3-base-absa-v1.1**: 186 M 参数；轻量级细粒度情感分析模型。 [ACL 2023; EACL 2024]
- **deberta-v3-large-absa-v1.1**: 418 M 参数；大规模 ABSA 分类器。 [ACL 2023; EACL 2024]
- **OmniGenome-186M**: 186 M 参数；RNA 二级结构预测与设计基础模型。 [AAAI 2025]
- **OmniGenome-52M**: 52 M 参数；快速推理版基因组语言模型。 [AAAI 2025]
- **OmniGenome-v1.5**: 186 M 参数；通用基因组基础模型 v1.5。 [Pending]
- **MoEOmniGenomeV2**: 577 M / 2B 参数；MoE 架构基因组大模型。 [Pending]
- **MP-RNA**: 186 M 参数；多物种 RNA 预测模型。 [EMNLP 2024]
- **PlantRNA-FM**: 35 M 参数；面向植物功能位点挖掘的 RNA 基础模型。 [Nat. Mach. Intell. 2024]

🤖 Hugging Face Space

- **PyABSA Space**: 级别 ABSA 模型交互式演示中心，受邀收录为 Gradio-Block 官方演示项目。 [CIKM 2023]
- **OmniGenBench**: 官方在线排行榜，方便研究者试用基准测试框架。 [arXiv 2024]
- **Text-Adversarial-Attack-Defense (Rapid)**: 文本对抗攻击与防御交互式演示。 [EMNLP 2024]

📦 PyPI 发布与维护 (<https://pypi.org/user/yangheng>)

- **omnigenbench**: 多任务、多物种的 GFM 自动化基准与排行榜。 [arXiv 2024]
- **pyabsa**: 细粒度情感分析框架（31 模型 × 30 数据集）。 [CIKM 2023]
- **boostaug**: 基于实例过滤的文本增强工具。 [ACL 2023]
- **DaNuoYi**: 多任务演化注入测试框架，支持 6 类 WAF 注入。 [TSE 2024]
- **findfile**: 关键词级文件 / 目录快速检索工具。
- **metric-visualizer**: 实验指标自动统计与可视化套件。
- **mvis**: MetricVisualizer CLI — 单行命令生成图表。
- **autocuda**: PyTorch 环境下自动选择显存最大 CUDA 设备。