

余天予

电话: +86-15179283506; 邮箱: yu-ty21@mails.tsinghua.edu.cn; yiranytianyu@gmail.com

教育背景

北京航空航天大学	本科
专业: 软件工程	2017.09 - 2021.06
平均分: 91.37/100; 专业排名: 1/165	
清华大学	硕士
专业: 计算机技术	2021.09 - 2024.06

获奖经历

沈元奖章提名奖 (北航本科生最高荣誉, 10/4000+)	2020.12
本科生国家奖学金 (前 1%)	2018.11 & 2019.11
北京航空航天大学校级学习优秀奖学金 (第一名)	2018.11 & 2019.11
北京航空航天大学三好学生	2018.11 & 2019.11
社会工作优秀奖学金一等奖	2019.11

近期论文

Tianyu Yu, Yuan Yao, Haoye Zhang, Taiwen He, Yifeng Han, Ganqu Cui, Jinyi Hu, Zhiyuan Liu, Hai-Tao Zheng, Maosong Sun, Tat-Seng Chua. RLHF-V: Towards trustworthy MLLMs via behavior alignment from fine-grained correctional human feedback. *CVPR 2024*

Tianyu Yu, Chengyue Jiang, Chao Lou, Shen Huang, Xiaobin Wang, Wei Liu, Jiong Cai, Yangning Li, Yinghui Li, Kewei Tu, Hai-Tao Zheng, Ningyu Zhang, Pengjun Xie, Fei Huang, Yong Jiang. SeqGPT: An out-of-the-box large language model for open domain sequence understanding. *AAAI 2024*

Yuan Yao*, **Tianyu Yu***, Ao Zhang, Mengdi Li, Ruobing Xie, Cornelius Weber, Zhiyuan Liu, Haitao Zheng, Stefan Wermter, Tat-Seng Chua, Maosong Sun. Visually Grounded Commonsense Knowledge Acquisition. *AAAI 2023*

Tianyu Yu, Tianrui Hui, Zhihao Yu, Yue Liao, Sansi Yu, Faxi Zhang, Si Liu*. Cross-Modal Omni Interaction Modeling for Phrase Grounding. *ACM MM 2020*

实习经历

Alibaba DAMO Academy	2023.02-2023.12
Mentor: 蒋勇	
❖ 通过“抽取”和“分类”两个原子任务建模开放域自然语言理解任务。	
❖ 通过大量实验分析优化数据和训练策略, 0.56B 模型零样本性能超过 ChatGPT 3.5。	
Tencent PCG 多模态算法组	2022.02-2022.12
Mentor: 陈曦	
❖ 提出结合文本域知识和图像域知识的少样本视觉关系抽取的模型结构无关框架。	
❖ 通过大量实验对比所提出方法和 SOTA 方法以及其他多模态大模型, 性能超过 SOTA 70%	
Tencent AI Lab	2021.07-2021.12
Mentor: 王琰	
❖ 在实体集合扩展任务中引入“减法”需求, 提出 Negative-aware 实体集合扩展任务	
❖ 使用基于 BERT 的双塔模型取得了 Negative-aware 实体集合扩展的 SOTA 性能和优秀的时空效率	

依图科技 基础算法研究部门 G0

2020.09-2021.02

- ❖ 使用 GCN 对非机动车 Re-ID 数据进行自动标注
- ❖ 降低了模型训练中弱监督噪声问题的影响，结果被采纳上线
- ❖ 参与设计自动标注方案，负责对多种标注任务的自动化可行性和效率进行评估

研究经历

基于开源模型反馈减少多模态大模型幻觉

2024.01-2024.06

- ❖ 设计无偏采样+分而治之评估+在线偏好学习的 RLAIF-V 框架减少多模态大模型幻觉。
- ❖ 轻量训练 12B 模型达到远超 GPT-4V 的多模态大模型可信度水平 SOTA。
- ❖ 数据集位列 Hugging Face Trending 第二。

基于修改式人类反馈减少多模态大模型幻觉

2023.07-2023.12

- ❖ 通过人类识别和纠正模型错误的方式获取修改式细粒度人类反馈。
- ❖ 设计 Dense DPO 算法充分利用修改式细粒度反馈数据，实现开源多模态大模型 SOTA 可信度。
- ❖ 产出论文并被 CVPR 2024 收录。

少样本视觉关系抽取

2022.05-2022.12

- ❖ 提出结合文本域知识和图像域知识的少样本视觉关系抽取的模型结构无关框架。
- ❖ 通过大量实验对比所提出方法和 SOTA 方法以及其他多模态大模型，性能超过 SOTA 70%
- ❖ 产出论文并计划投稿到 ICCV 2023

视觉常识获取

2021.07-2022.09

- ❖ 将基于视觉感知的常识知识获取定义为一个远程监督的多样本学习任务，并构建 benchmark
- ❖ 提出基于对比的注意力机制对多个图像样本进行降噪，方法取得了 SOTA 效果
- ❖ 提出结合多个模态信息并取得优于所有单模态方法的效果，结果发表在 AAAI 2023 (CCF A 类)

Negative-aware 实体集合扩展

2021.04-2021.11

- ❖ 在实体集合扩展任务中引入“减法”需求和负类种子实体的概念
- ❖ 在英文维基百科和哈利波特系列小说两个文本域上构建了多语义层次的实体集合扩展数据集
- ❖ 设计算法实现了在正负类输入上的实体集合扩展的 SOTA 性能
- ❖ 产出论文并计划投稿到 SIGIR 2023

Sentence-Level Pretraining for Document-Level RE

2020.07-2021.03

指导老师：刘知远副教授

- ❖ 设计 Sentence Distance Ranking, Entity Context Prediction, Sentence-Level ELECTRA 训练任务并在 wikipedia-en 上进行大规模预训练
- ❖ 实验分析模型在 DocRED 上的 RE 性能

基于跨模态统一交互建模的开放词表物体检测

2019.10-2020.07

指导老师：刘偲副教授

- ❖ 从空间邻域范围和全局范围来模拟图像内实体之间的交互，准确理解每个物体的环境信息
- ❖ 使用基于 self-attention 机制的跨模态交互方式捕捉多模态特征用于物体检测
- ❖ 提出多层次的正则化训练方法提高模型的检测精度，结果发表在 ACM MM 2020 (CCF A 类)

其他信息

计算机技能：掌握 C、C++、Python、Java、SQL 等语言，以及 Pytorch、Tensorflow、Scipy 等工具

语言：英语 CET6，托福 (98)