

王志庞 ZHIPANG WANG

✉ zhipangwang@gmail.com

🌐 github.com/ZpWang-AI

☎ (+86) 139-1247-7730

📍 苏州, 江苏

🎓 教育背景

- ▶ 苏州大学 硕士, 计算机科学与技术, 计算机科学与技术学院 2023.9 - Now
研究方向: 自然语言处理、篇章关系识别、大语言模型
- ▶ 苏州大学 学士, 计算机科学与技术, 计算机科学与技术学院 2019.9 - 2023.7

📁 实习经历

- ▶ 美国加州戴维斯大学 (U.C. Davis) 2025.3 - 2025.9
现有的文本到视频生成模型在理解抽象物理原则和生成符合物理规律的视频方面仍面临挑战。当前缺乏集中表现物理现象的视频生成数据集, 因此我们希望通过构建富含物理动态过程的数据集, 推动模型在视频生成中的物理一致性。
 - 人工标注: 我们团队人工标注了超过 10,000 条视频, 每条视频标注了一条或多条动态过程的实例, 每个实例由目标、属性、物理过程及其解释组成。
 - 自动标注流程构建: 基于人工数据集, 我们建立了一个基于大模型的自动标注流程, 包含数据过滤、目标抽取、属性分类、物理过程抽取和物理规律解释五个步骤。
 - 模型训练与结果: 我们使用 Qwen2.5-VL-7B-Instruct 作为基础模型, 将多个标注环节的训练数据混合后进行训练。我们的标注结果显示, 动态过程提取的 F1 值从 73.9% 提升至 91.6%, 数据过滤和属性分类等环节的性能也有显著提高。
- ▶ 博世 (Bosch) 投资有限公司 (中国上海) 2022.7 - 2022.10
为了解决自动筛选缺陷样品的问题, 我们利用机器视觉中的实例分割技术进行样品检测。由于样本稀缺, 我们用数据增强来提升模型训练后的性能。具体而言, 我们探索了使用对抗生成模型进行数据增强的方案。并且为了缩小生成样本与真实样本之间的差距, 我们又采用了领域自适应的迁移学习方法进行优化。

📄 论文成果

- ▶ Using Subtext to Enhance Generative IDRR. (一作, ACL'25, CCF-A).
在篇章关系识别领域, 以往的研究常常忽视了论元之间的潜在语义。我们通过大模型挖掘论元间的潜台词, 并提出了一种基于潜台词的隐式篇章关系识别方法。
 - 潜台词生成与应用: 我们利用 GPT-3.5 生成部分潜台词, 并通过蒸馏 LLaMA3-8B-Instruct, 得到最终的潜台词。然后, 我们将潜台词与输入的论元通过 prompt 拼接, 获得模型的输出结果。
 - 置信度辨别器: 我们基于模型的 logits 计算输出的置信度, 从而筛选出部分不适用潜台词的样例, 进一步提升模型性能。
 - 性能提升: 我们的模型在生成式方法中取得了 SOTA。在消融实验中, 相较于没有潜台词的基线, 我们在 PDTB 3.0 的 Macro-F1 值上从 70.7% 提升至 72.8%, 并在加入置信度辨别器后进一步提升至 73.3%。
- ▶ Learning to Differentiate Pairwise-Argument Representations for Implicit Discourse Relation Recognition. (一作, CIKM'24, CCF-B).
篇章关系识别依赖于论元编码表示学习, 因此我们提出了一种新颖的联合学习框架, 以增强论元编码分布的辨别性和区分度。1) 原型学习: 通过构建每种关系类型的原型表示, 计算输入论元与这些原型之间的相似度, 从而引导模型学习更具辨别性的特征表示。2) 对抗学习: 通过生成对抗样本增强模型的鲁棒性, 训练模型在面对噪声和干扰时有效区分不同关系类型的能力。3) 中心迁移重分配: 强制调整不同关系类

型的中心位置，使得相同类型的样本更紧密地聚集在一起，同时保持不同类型样本之间的距离。通过我们的联合学习框架，RoBERTa 模型在 PDTB 2.0 上的 Macro-F1 分数从基线的 55.7% 提升至 65.4%。在 PDTB 3.0 上，Macro-F1 分数从 66.3% 提升至 71.7%。

► Enhancing Reasoning Capabilities by Instruction Learning and Chain-of-Thoughts for Implicit Discourse Relation Recognition. (三作, Findings of EMNLP'23 , CCF-B).

当前隐式篇章关系识别的方法多依赖于显式连接词的使用，而我们旨在探讨生成模型在隐式关系识别中的潜力。我们提出了一种基于生成模型的分类方法，结合了指令学习、上下文学习和思维链。具体而言，模型首先识别论点对的隐式或显式关系，然后推断合适的连接词，最后进行分类。这种分阶段的推理过程有效地减少了噪音，提升了模型的表现。我们的方法在多个数据集上得到了 SOTA 的性能表现。

🔧 项目经历

► 实验室项目：领域问答系统/Text2SQL 系统 2023.7 - 2024.10

参与实验室项目，构建领域问答系统和 Text2SQL 系统，主要负责相关模型的训练与调试。

- Text2SQL 模型：利用人工标注和 LLM 生成等方法，构建领域数据集共计 5 万条。对 DeepSeek-coder 进行 LoRA 微调训练，模型性能相比于基线方法 EX 提升 14.2%，EM 提升 7.3%。为进一步提高模型对 SQL 模式及问题文本中偏僻词汇和缩略语的理解，通过 RAG 检索相关 schema 文档和相似案例辅助生成。最终模型获得 EX 性能 82.2%，EM 性能 77.4%。
- 检索问答模型：利用 BGE-M3 模型实现 FAQ 查询式问答，通过 LLM 重写生成正例，召回检索挖掘负例，进行对比训练。最终模型的召回结果 hit@1 值达到 0.791。

► Kaggle 竞赛：课程推荐 (银牌) 2022.12 - 2023.3

本次竞赛旨在将多语言教育内容与特定课程主题进行匹配。为解决这一问题，我们构建了粗排和精排模型，并结合课程主题的标题、描述与文本编码来优化检索。最终得分为 0.634，排名 49/1057。

🏆 获奖情况

► 学业奖学金类

- 硕士生学业奖学金特等奖 2023.12
- 学习优秀奖、专项奖、综合奖、创新创业奖 2021.12
- 学习优秀奖、专项奖、综合奖、雨菡奖学金 2020.12

► ACM 竞赛类

- 国一，蓝桥杯 Python 组 2022.6
- 银牌，江苏省大学生程序设计竞赛 2022.5
- 银牌，全国大学生程序设计竞赛哈尔滨站 2021.11
- 银牌，国际大学生程序设计竞赛沈阳站 2021.11
- 银牌，全国大学生程序设计竞赛桂林站 2021.11
- 个人一等奖，全国团体程序设计天梯赛 2021.5
- 铜牌，国际大学生程序设计竞赛昆明站 2021.4
- 铜牌，国际大学生程序设计竞赛上海站 2020.12
- 铜牌，江苏省大学生程序设计竞赛 2020.11

⚙️ 专业技能

- 熟悉大语言模型 LLaMA、Qwen，多模态大模型 Qwen2.5-VL，混合专家模型 DeepSeek。
- 熟悉 Python、Transformers、PyTorch；了解 LoRA、QLoRA 等参数高效微调方法，以及 LLaMA-Factory、DeepSpeed 等大模型训练和推理框架。