Yu Bai

Phone：(+86) 13051826877 · Email：yubai@bit.edu.cn

EDUCATION

Beijing Institute of Technology, Beijing, China 2019.09 - 2025.06

- **Ph.D. Candidate**, Computer Science and Technology.
- **Adviser**: Prof. Heyan Huang; Assoc. Prof. Yang Gao.
- **Research Interests**: Text Summarization; Natural Language Generation.

China University of Geosciences (Beijing), Beijing, China 2015.09 - 2019.06

- **B.S.** , Computer Science and Technology.

PUBLICATIONS

Cross-Lingual Abstractive Summarization with Limited Parallel Resources

- Yu Bai, Yang Gao, Heyan Huang.
- To appear at the **ACL 2021 main conference**. Inspired by UniLM, we use an unified decoder to generate the concatenation of both monolingual summaries and cross-lingual summaries, learning interactions involving alignments and summary patterns across languages. These interactions further encourage attaining knowledge transfer from high-resource languages. In-depth analysis on generated summaries and attention heads explains the effectiveness of our model.

Exploring Explainable Selection to Control Abstractive Summarization

- Haonan Wang, Yang Gao, Yu Bai, Mirella Lapata, Heyan Huang.
- Appeared at **AAAI 2021**. We developed a novel select-and-generate framework that focuses on explainability. By revealing the latent centrality, along with scores for sentence novelty and relevance, a novel pair-wise matrix captures the sentence interactions, centrality, and attribute scores. A sentence-deployed attention mechanism in the abstractor ensures the final summary emphasizes the desired content.

Multiple Perspective Answer Reranking for Multi-passage Reading Comprehension

- Mucheng Ren, Heyan Huang, Ran Wei, Hongyu Liu, Yu Bai, Yang Wang, Yang Gao.
- Appeared at **NLPCC 2019**. Focusing on multi-passage Machine Reading Comprehension (MRC) task, we proposed a new multi-perspective answer reranking technique that considers all documents to verify the confidence of candidate answers. Specifically, we rearrange the order of traditional pipeline model and make a posterior answer reranking instead of prior passage reranking.

COMPETITIONS

Collegiate Programming Contests 2017.10-2017.12

- **Bronze Medal** in ACM/ICPC Asia Regional Qingdao Site 2017.
- **Bronze Medal** in ACM/ICPC Asia Regional Nanning Site 2017.
- **Bronze Medal** in The 3rd China Collegiate Programming Contest (Harbin), 2017.

2019 Language and Intelligence Challenge: Machine Reading Comprehension 2019.03-2019.05

- **Bronze Award** Rank 5.

OTHERS

Teaching 2019.09 - 2021.09

- **Beijing Institute of Technology**, Beijing, China.
- **Teaching Assistant**, Selected Topics In New Technology for Computer.

Programming Languages

- Python, C/C++, LaTeX, Java, Bash under Linux/Unix.
- PyTorch, Keras.

柏宇

手机：(+86) 13051826877 · 邮箱：yubai@bit.edu.cn

教育经历

北京理工大学，*博士在读* 2019.09 - 2025.06

- **计算机学院**：计算机科学与技术，语言智能与社会计算
- **导师**：黄河燕，高扬
- **平均成绩**：93.83，1/149
- **研究方向**：自然语言处理；文本摘要；低资源文本生成

中国地质大学（北京），*本科* 2015.09 - 2019.06

- **信息工程学院**：计算机科学与技术

科研经历

Cross-Lingual Abstractive Summarization with Limited Parallel Resources

- Yu Bai, Yang Gao, Heyan Huang.
- **ACL 2021**，该研究集中于低资源跨语言摘要任务，通过将生成目标设置为单语摘要与跨语言摘要的拼接，实现对不同语言摘要中短语对齐、摘要模式等的建模，从而使高资源摘要知识更容易被转移到低资源文本摘要模型中。我们通过自动评价、人工评价、对生成摘要的详细分析以及注意力头的可视化，全面证实了模型的有效性和可解释性。

Exploring Explainable Selection to Control Abstractive Summarization

- Haonan Wang, Yang Gao, Yu Bai, Mirella Lapata, Heyan Huang.
- **AAAI 2021**，该研究集中于生成式文本摘要的可解释性。通过一个能够捕捉文档句子间的关系及他们的内在属性（新颖度，相关度）的关系矩阵影响文档中每个句子的重要程度，进一步影响生成摘要的结果。同时，通过对关系矩阵值的过滤，我们的方法能够一定程度上控制结果摘要的各类属性。

Multiple Perspective Answer Reranking for Multi-passage Reading Comprehension

- Mucheng Ren, Heyan Huang, Ran Wei, Hongyu Liu, Yu Bai, Yang Wang, Yang Gao.
- **NLPCC 2019**，应用 BERT 和后验知识答案排序提升多文档机器阅读理解的性能。

获奖经历

算法竞赛 2017.10-2017.12

- **2017 ACM-ICPC 亚洲区域赛青岛站** 铜奖
- **2017 ACM-ICPC 亚洲区域赛南宁站** 铜奖
- **CCPC 第三届中国大学生程序设计竞赛（哈尔滨）** 铜奖

2019 语言与智能技术竞赛——机器阅读理解技术竞赛 2019.03-2019.05

- **铜奖** *总排名第五名*
- **简介：**与四位同学合作参与竞赛，完整设计了一套多文档阅读理解系统。模型通过应用 BERT 及后验知识答案重排序的方法在竞赛中表现优异。主要负责系统中的情感极性判断模型（肯定/否定）及相关部分论文的撰写。竞赛所用方法以论文形式发表于 NLPCC2019。

其他

教学经历 2019.09 - 2021.09

- **计算机新技术专题**
- **助教**，协助老师负责 300 人以上的课堂管理，成绩评定；协助统筹、邀请多名校内校外老师介绍不同计算机前沿专题。

编程技能

- Python, C/C++, LaTeX, Java, Bash under Linux/Unix.
- PyTorch, Keras.