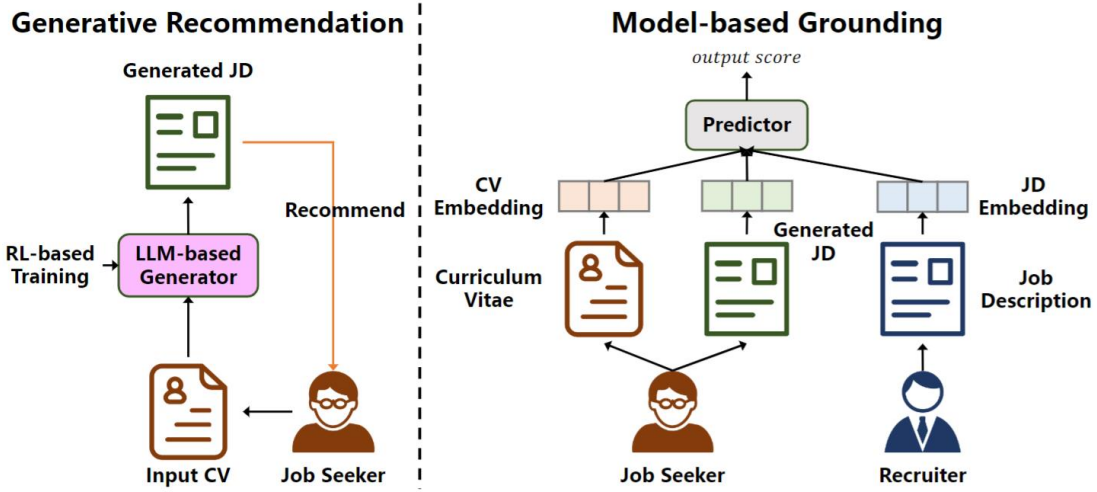
Exploiting Large Language Model with Reinforcement Learning for Generative Job Recommendations

Zhi ZHENG, Zhaopeng QIU, Chen ZHU, Xiao HU, Likang WU, Yang SONG, Hengshu ZHU, Hui XIONG

Frontiers of Computer Science, DOI: 10.1007/s11704-025-40843-1

Problems & Ideas

- Problems of LLM-based Generative RSs:
  - SFT-based training methods are insufficient for LLMs to adequately grasp the knowledge embedded in complex interactive behaviors.
  - similarity-based grounding methods also face challenges for long text matching.
- Ideas: Train a model which can evaluate the matching degree, use a PPO-based RL method to fine-tune the LLM-based recommender. Moreover, we propose a model-based grounding method.



Main Contributions

- Contributions:
  - To the best of our knowledge, this is the first piece of work which proposes an LLM-based generative job recommendation paradigm;
  - We propose a novel three-step training methodology with reinforcement learning to train an LLM-based generative job recommendation model;
  - We propose a model-based grounding method for mapping the results generated by the LLM.
  - We conducted extensive experiments on two real-world datasets and the experimental results show GIRL achieves the best performance compared with several competitive baseline methods.

	JobRec		JobRec-Cold	
Model	AUC(↑)	LogLoss(↓)	AUC(↑)	LogLoss(↓)
NCF	0.5415	0.4434	0.5410	0.4549
PJFNN	0.5508	0.4425	0.5434	0.4534
APJFNN	0.5911	0.4303	0.5720	0.4426
BERT	0.6349	0.4043	0.6198	0.4270
FT-LLM	0.6313	0.4035	0.5998	0.4477
BIGRec	0.5499	0.4427	0.5404	0.4513
BIGRec-MLP	0.6297	0.4044	0.6166	0.4446
GIRL	**0.6476**	**0.3908**	**0.6347**	**0.4229**

Model Pair	Win	Tie	Loss	Adv.
GIRL v.s. LLaMA-7b	0.63	0.07	0.30	0.33
GIRL v.s. BLOOMZ-7b	0.74	0.08	0.18	0.56
GIRL v.s. BELLE-7b	0.45	0.17	0.36	0.09
GIRL v.s. GIRL-SFT	0.45	0.25	0.26	0.19
GIRL-SFT v.s. LLaMA-7b	0.55	0.03	0.42	0.13
GIRL-SFT v.s. BLOOMZ-7b	0.73	0.07	0.19	0.54
GIRL-SFT v.s. BELLE-7b	0.49	0.06	0.41	0.08
BELLE-7b v.s. LLaMA-7b	0.61	0.04	0.34	0.27
BELLE-7b v.s. BLOOMZ-7b	0.65	0.07	0.17	0.48