



PediatricsGPT: Large Language Models as Chinese Medical Assistants for Pediatric Applications

Dingkang Yang^{1*}, Jinjie Wei^{1*}, Dongling Xiao^{2*}, Shunli Wang¹, Tong Wu², Gang Li², Mingcheng Li¹
Shuaibing Wang¹, Jiawei Chen¹, Yue Jiang¹, Qingyao Xu¹, Ke Li², Peng Zhai¹, Lihua Zhang¹

¹Academy for Engineering and Technology, Fudan University

²Tencent Youtu Lab, Shanghai, China

Outline

- **PedCorpus**
- **Proposed Framework**
- **Experimental Results**

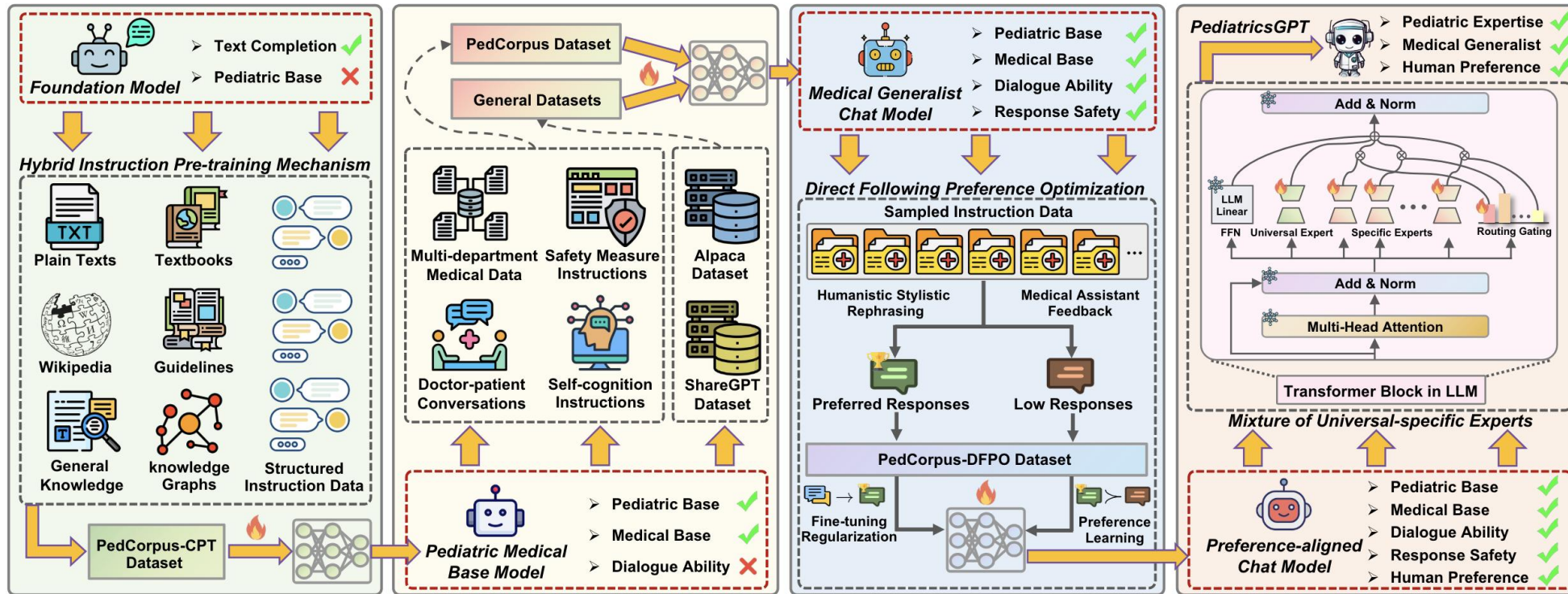
PedCorpus



Dataset	Data Sources	Department	Number/Size	Human Preference	Task Type		
					MedKQ&A	EviDiag	TreRecom
PedCorpus	Pediatric Textbooks	Pediatrics	37,284	✓	✓	–	✓
	Pediatric Guidelines	Pediatrics	63,129	✓	✓	–	✓
	Pediatric KG	Pediatrics	46,320	✓	✓	–	✓
	Real Doctor-Patient Conversations	Multiple	46,385	✓	–	✓	✓
	Distilled Medical Datasets	Multiple	107,177	–	✓	✓	✓
PedCorpus-CPT	Plain Textbooks, Guidelines, KG	Multiple	975.8MB	–	✓	✓	✓
	Filtered Chinese Wikipedia	Multiple		–	–	–	–
	Extended data from PedCorpus	Multiple		–	–	–	–
PedCorpus-DFPO	Pediatrics data from PedCorpus	Pediatrics	15,556	✓	✓	✓	✓

- **Specialized Pediatric Data:** Extracting pediatric data from textbooks, guidelines, and knowledge graphs ensures knowledge professionalism.
- **Real Doctor-patient Conversations:** Incorporating authentic doctor-patient dialogues from online treatment platforms and voice transcriptions during medical consultations.
- **Distilled Medical Datasets:** Integrating general medical knowledge from existing datasets.

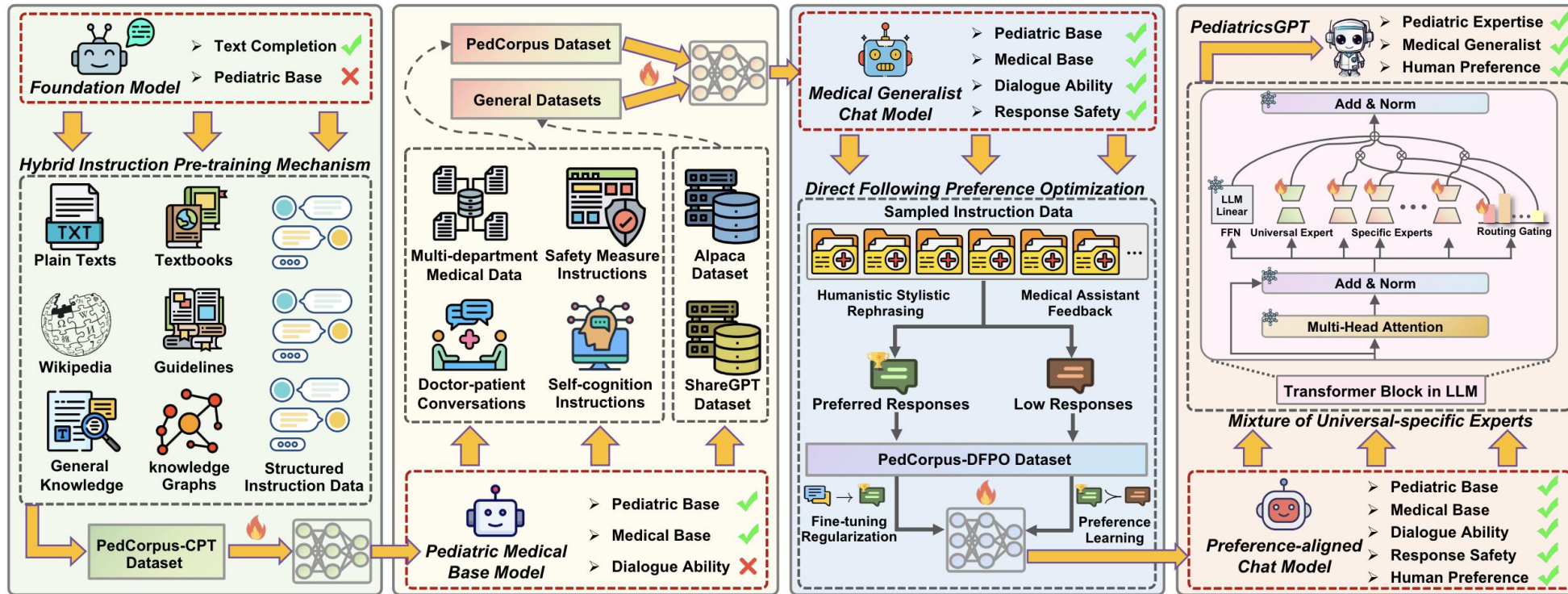
Proposed Framework



■ **Continuous Pre-Training:** Introducing a hybrid instruction pre-training mechanism injects rich and extensive medical knowledge into the base models, mitigating the problem of catastrophic model forgetting at follow-up due to differences in data distribution and format during the pre-training and fine-tuning phases.

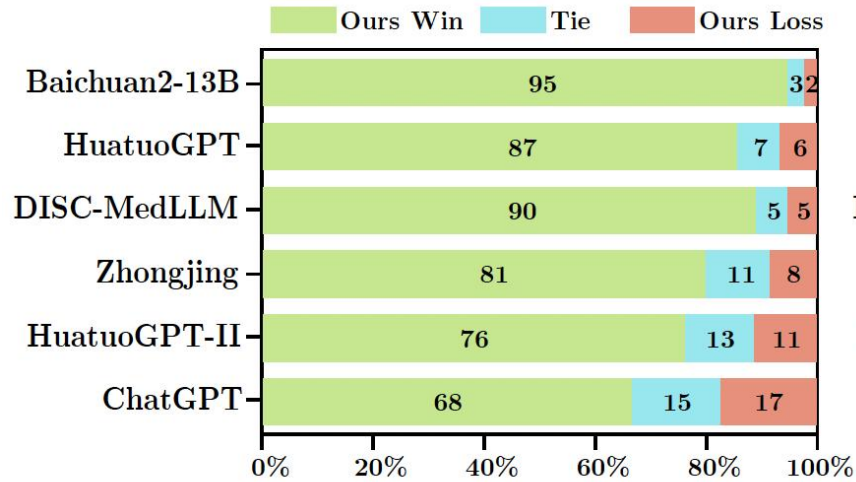
■ **Full-parameter Supervised Fine-tuning (SFT):** Activating the model's medical instruction-following ability calls on the dense knowledge from pre-training to facilitate the model's understanding of structured instructions and logical reasoning.

Proposed Framework

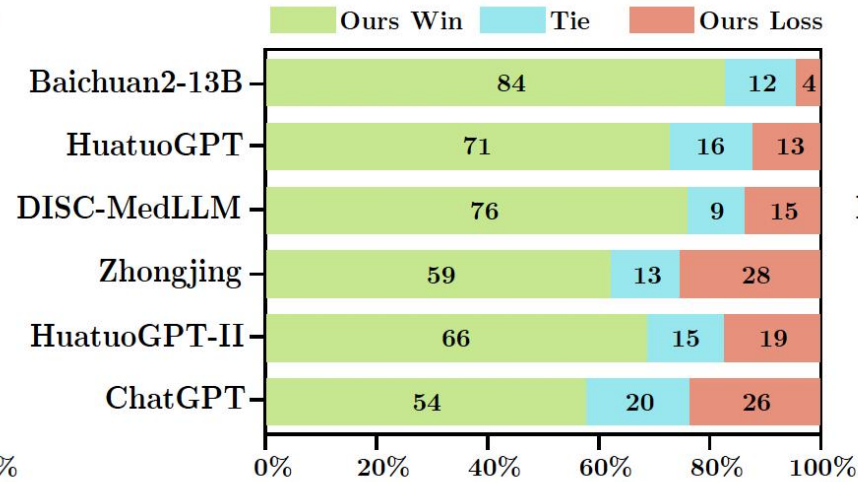


- **Direct Following Preference Optimization:** Facilitating model generation of innocuous and humanistic responses, regularizing model behavior boundaries, and facilitating robust and smooth human preference learning.
- **LoRA-based Parameter-efficient SFT:** Activating the model's medical instruction-following ability calls on the dense knowledge from pre-training to facilitate the model's understanding of structured instructions and logical reasoning.

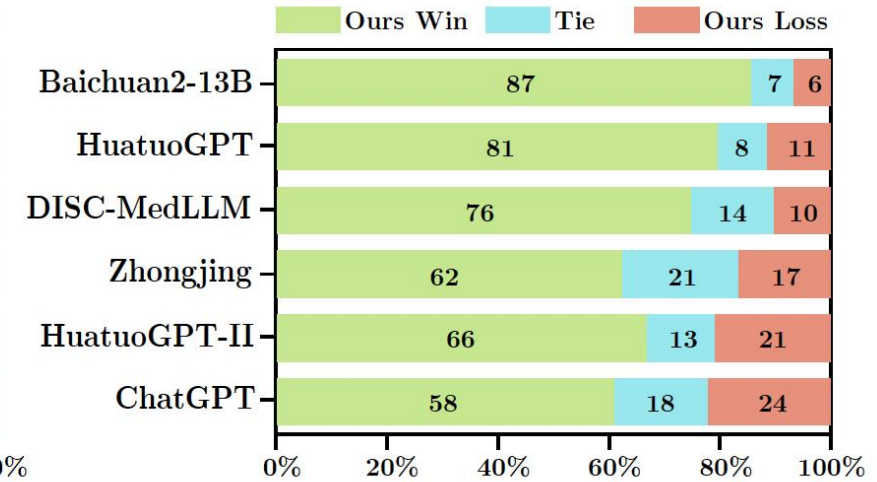
Experimental Results



(a) MedKQ&A Benchmark Evaluation



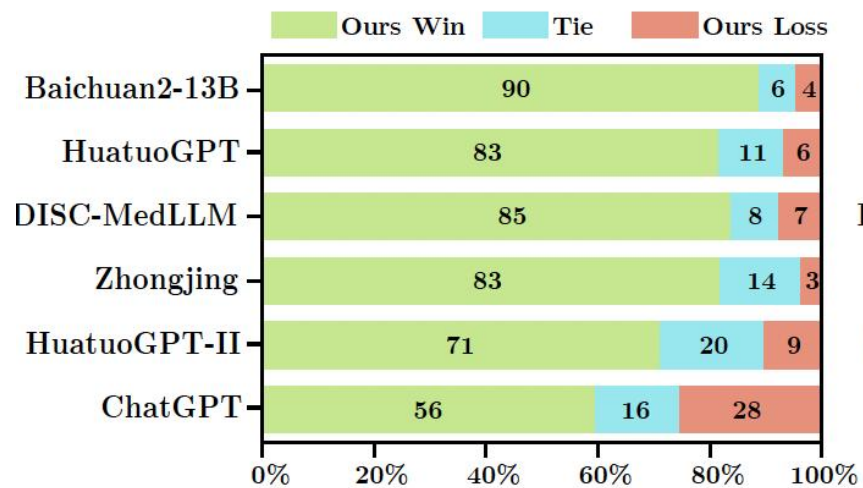
(b) EviDiag Benchmark Evaluation



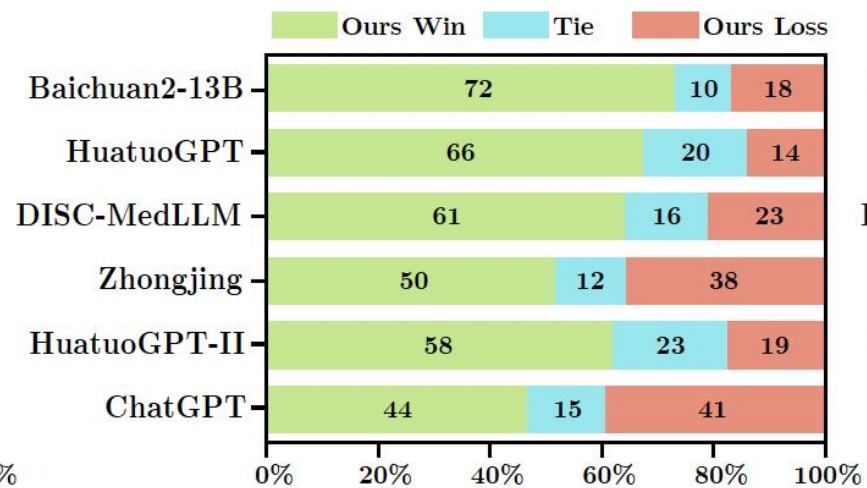
(c) TreRecom Benchmark Evaluation

Evaluated by GPT-4 on the four dimensions of response usefulness, correctness, consistency, and fluency, PediatricsGPT significantly outperforms current open-source Chinese medical LLMs on the knowledge-driven question-answer task (MedKQ&A), multi-round evidence-based Diagnosis task (EviDiag), and treatment recommendation task (TreRecom).

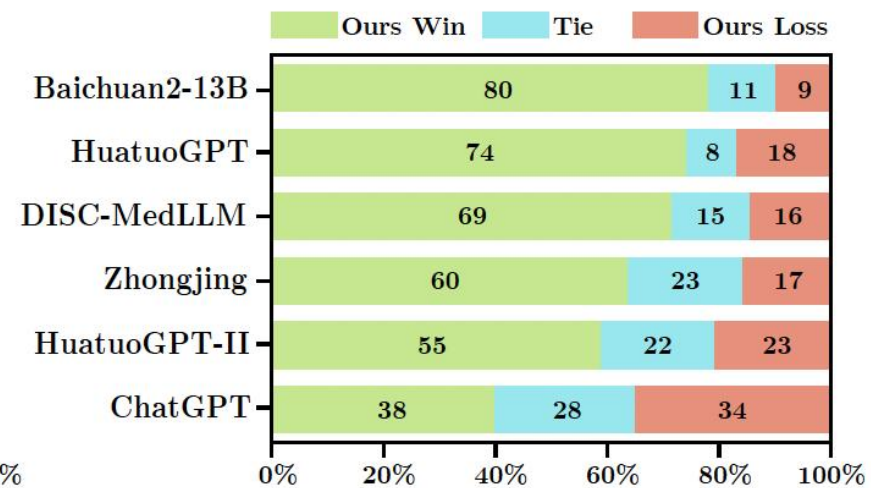
Experimental Results



(a) MedKQ&A Benchmark Evaluation



(b) EviDiag Benchmark Evaluation



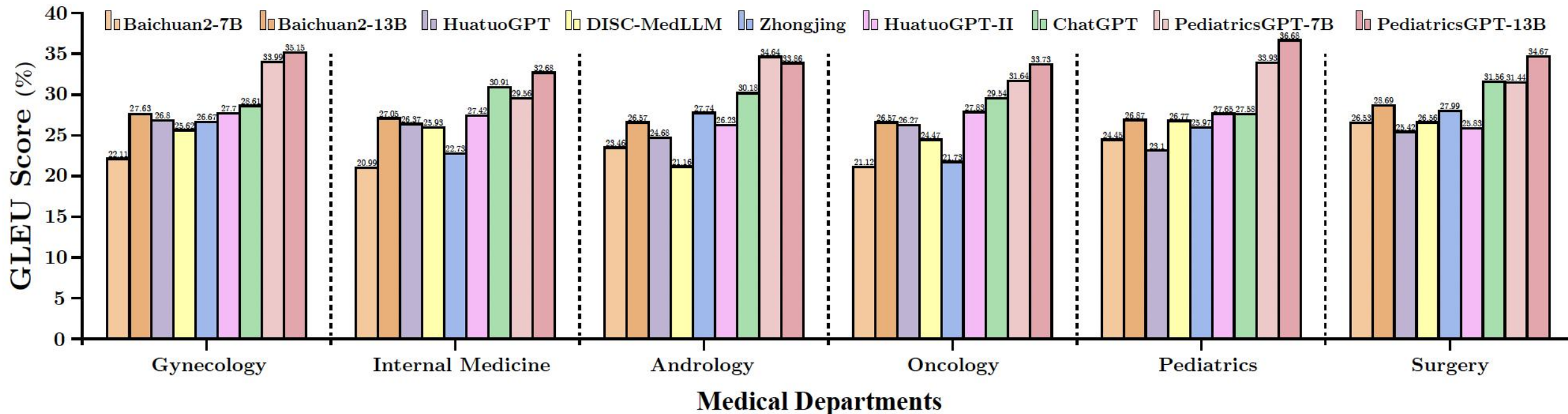
(c) TreRecom Benchmark Evaluation

Evaluated by professional doctors on three dimensions of response professionalism, factuality, and safety, PediatricsGPT significantly outperforms current open-source Chinese medical LLMs on the knowledge-driven question-answer task (MedKQ&A), multi-round evidence-based Diagnosis task (EviDiag), and treatment recommendation task (TreRecom).

Experimental Results

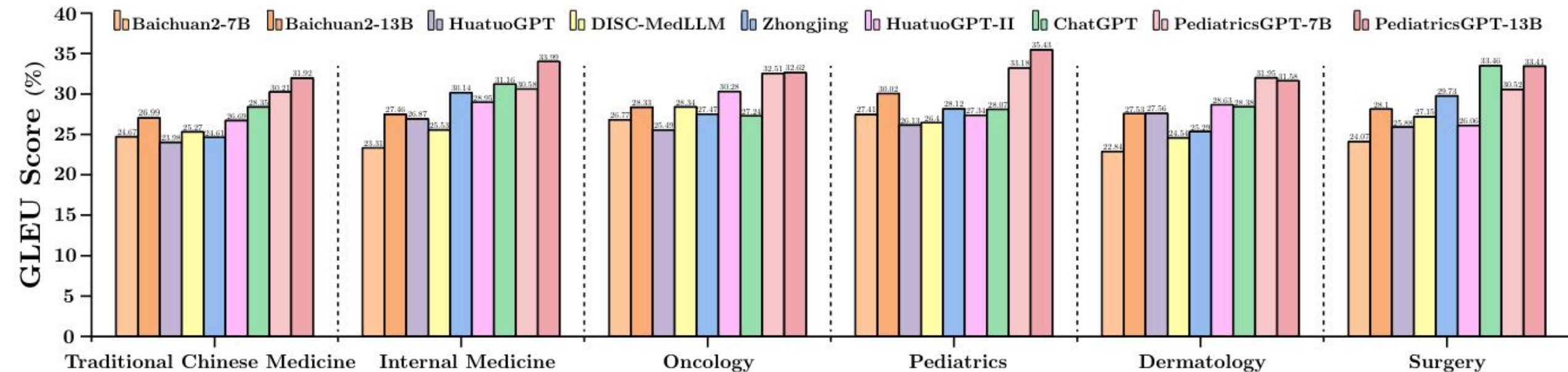
Benchmark	Model	ROUGE-1	ROUGE-2	ROUGE-L	BLEU-1	BLEU-2	BLEU-3	BLEU-4	GLEU	Distinct-1	Distinct-2
MedKQ&A	Baichuan2-7B	40.88	19.44	21.5	26.77	20.00	17.30	14.86	24.88	20.14	39.95
	Baichuan2-13B	46.96	22.85	22.54	29.02	25.62	22.63	19.31	27.97	21.45	42.53
	HuatuogPT	48.52	23.44	25.13	43.00	41.25	36.31	29.82	34.60	20.42	41.27
	DISC-MedLLM	53.83	25.98	27.71	47.91	44.57	37.65	30.07	37.11	26.63	51.98
	Zhongjing	53.97	26.03	29.56	51.11	45.04	39.13	33.59	42.61	26.75	52.66
	HuatuogPT-II	55.27	26.59	27.95	59.07	51.49	45.38	38.70	39.18	20.97	41.34
	ChatGPT	56.92	27.87	29.05	61.58	54.37	47.97	40.77	45.15	20.76	40.19
	PediatricsGPT-7B	<u>58.08</u>	<u>31.78</u>	<u>31.11</u>	<u>59.41</u>	<u>56.88</u>	<u>57.47</u>	<u>55.34</u>	<u>54.41</u>	24.33	47.41
	PediatricsGPT-13B	60.85	36.56	35.64	61.65	63.17	58.96	59.34	57.22	24.24	46.23
EviDiag	Baichuan2-7B	26.81	7.75	11.22	15.18	11.51	9.19	6.72	13.44	23.65	46.93
	Baichuan2-13B	39.14	12.06	12.44	47.65	36.02	28.82	21.19	28.28	25.45	50.43
	HuatuogPT	35.12	10.77	15.04	46.22	33.10	25.44	21.22	25.44	22.30	45.73
	DISC-MedLLM	33.55	11.67	15.32	15.91	12.46	10.27	7.96	16.77	35.89	69.36
	Zhongjing	40.92	14.26	17.41	48.64	37.52	30.17	22.44	27.03	<u>33.40</u>	<u>65.89</u>
	HuatuogPT-II	39.52	12.14	16.38	49.58	37.62	30.66	23.34	28.98	21.97	43.62
	ChatGPT	40.88	13.42	16.97	48.84	37.69	30.55	23.17	29.02	23.49	46.54
	PediatricsGPT-7B	<u>45.83</u>	<u>16.60</u>	<u>19.91</u>	<u>54.37</u>	<u>41.99</u>	<u>37.59</u>	<u>29.03</u>	<u>33.42</u>	23.49	46.61
	PediatricsGPT-13B	47.32	17.63	21.87	58.21	45.72	39.74	31.25	37.15	23.34	46.34
TreRecom	Baichuan2-7B	48.39	23.07	26.35	47.94	40.91	35.54	29.69	35.06	21.90	43.57
	Baichuan2-13B	48.87	23.41	26.42	49.96	46.24	42.84	35.04	35.63	22.36	45.12
	HuatuogPT	53.48	25.41	27.08	58.14	49.64	42.93	35.16	41.63	23.26	46.21
	DISC-MedLLM	52.77	24.26	28.89	58.73	50.05	42.96	35.59	42.44	24.30	51.95
	Zhongjing	54.92	26.63	29.68	60.12	53.31	44.25	38.76	40.38	26.18	53.94
	HuatuogPT-II	58.44	30.47	32.02	59.91	54.26	45.73	38.92	42.28	28.88	<u>57.15</u>
	ChatGPT	59.59	33.34	35.79	62.81	55.79	49.85	43.29	47.59	31.09	56.87
	PediatricsGPT-7B	56.92	29.13	31.26	61.36	55.34	46.44	40.61	44.65	26.06	52.77
	PediatricsGPT-13B	62.83	39.32	40.82	63.56	56.68	50.80	44.31	54.65	31.94	57.56

Experimental Results



PediatricsGPT significantly outperforms current open source Chinese medical LLMs on publicly available Chinese medical benchmark CMD under metrics-based evaluation.

Experimental Results



PediatricsGPT significantly outperforms current open source Chinese medical LLMs on publicly available Chinese medical benchmark webMedQA under metrics-based evaluation.

Experimental Results

Instruction:

请简述红细胞葡萄糖-6-磷酸脱氢酶缺乏症的病因。

Zhongjing

红细胞葡萄糖-6-磷酸脱氢酶缺乏症是一种遗传性疾病，主要由于遗传因素引起的基因突变所致。这种基因突变导致红细胞内的红细胞葡萄糖-6-磷酸脱氢酶的结构和功能发生异常，从而影响血红蛋白的合成和代谢，导致患者出现贫血、黄疸、肝功能异常等症状。

HuatuoGPT-II

红细胞葡萄糖-6-磷酸脱氢酶缺乏症的病因主要是由于遗传基因突变引起的。这种突变会导致葡萄糖-6-磷酸脱氢酶的功能受损或完全缺失，进而影响红细胞的正常代谢过程。这种疾病通常是通过遗传方式传递给后代的，而具体的遗传方式则取决于突变的类型。此外，一些环境因素也可能对红细胞葡萄糖-6-磷酸脱氢酶缺乏症的发病起到一定的影响。例如，某些药物、感染、应激等因素都可能引发疾病的发作。总结起来，红细胞葡萄糖-6-磷酸脱氢酶缺乏症是一种由遗传基因突变引起的疾病，其病因主要是基因突变导致葡萄糖-6-磷酸脱氢酶功能受损或缺失。

PediatricsGPT

红细胞葡萄糖-6-磷酸脱氢酶（G6PD）缺乏症是一种遗传性疾病，由于G6PD基因的突变导致。G6PD是一种酶，主要存在于红细胞中，它对于保护红细胞免受氧化损伤至关重要。G6PD帮助维持还原型谷胱甘肽（GSH）的水平，GSH是一种重要的抗氧化剂，能够中和有害的氧化剂，保护红细胞免受损伤。当人体中的G6PD活性降低或缺失时，红细胞在面对氧化应激（如某些食物、药物、感染或其他应激情况）时更容易受损，导致红细胞过早破裂，从而引发溶血性贫血。这种状况可能导致疲劳、黄疸、尿色加深等症状。G6PD缺乏症的遗传模式主要是X染色体连锁隐性遗传。因为男性（XY）只有一个X染色体，所以如果他们的X染色体上携带有缺陷的G6PD基因，他们就会表现出这种疾病。而女性（XX）则需要两个X染色体上都携带有缺陷的G6PD基因才会表现出疾病，如果只有一个X染色体携带缺陷基因，她们通常是无症状的携带者，但在某些情况下也可能表现出轻微到中等程度的症状。G6PD缺乏症在全球范围内分布广泛，特别是在非洲、地中海地区、中东和亚洲的某些部分，这与这些地区疟疾高发有关，因为G6PD缺乏症在一定程度上提供了对某些类型疟疾的保护作用。

Response comparison between PediatricsGPT and current SOTA Chinese medical LLM assistants on the knowledge-driven question-answer task.

Experimental Results

Instruction

我的孩子前几天在玩耍时不小心摔倒，说腰部很疼，去医院检查诊断为腰椎间盘突出，我想了解对于腰椎间盘突出的儿童，应该采取怎样的药物治疗？

Zhongjing

对于儿童腰椎间盘突出的治疗，首先要确定病情的严重程度和病因，以便选择合适的治疗方法。一般情况下，轻度的腰椎间盘突出可以通过物理治疗和药物治疗来缓解症状。但是，药物治疗应该在医生的指导下进行，不能自行购买和使用药物。

HuatuogPT-II

腰椎间盘突出在儿童中相对较少见，通常与骨骼发育不全或过度活动有关。治疗方法主要包括保守治疗和手术治疗。对于儿童腰椎间盘突出的药物治疗，通常包括以下几种：1. 非甾体抗炎药 (NSAIDs)：如布洛芬 (Ibuprofen) 或者萘普生 (Naproxen) 等，可以帮助缓解疼痛和减少炎症。2. 肌肉松弛剂：如甲苯磺酸氨基氯地平 (Amlodipine Besylate) 等，可以帮助缓解肌肉痉挛。3. 神经营养药：如维生素B族等，可以帮助神经修复。4. 中药治疗：如活血化瘀类中药，可以帮助改善血液循环，缓解疼痛。对于儿童来说，药物治疗需要在医生的指导下进行。

PediatricsGPT

腰椎间盘突出在儿童中较为罕见，但治疗仍然需要谨慎。对于儿童的腰椎间盘突出，药物治疗主要包括以下几种：1. 非甾体抗炎药 (NSAIDs)：如布洛芬 (Ibuprofen) 或对乙酰氨基酚 (Acetaminophen)，可用于缓解疼痛和减轻炎症。2. 肌肉松弛剂：对于儿童，常用的包括硫酸溴己新 (Bromhexine) 等，可帮助缓解肌肉痉挛。3. 神经营养药物：例如维生素B族，有助于改善神经功能，减轻疼痛。4. 中药治疗：在儿童中应慎重选择，可考虑一些具有活血化瘀、舒筋活络作用的中药。需要强调的是，儿童腰椎间盘突出的治疗需在专业医生指导下进行，药物治疗应慎重选择剂量和种类。此外，儿童应加强锻炼，保持良好的姿势，避免长时间低头玩手机或电脑，有助于预防腰椎间盘突出的发生。

Response comparison between PediatricsGPT and current SOTA Chinese medical LLM assistants on the treatment recommendation task.

Conclusion



Thanks