

# Find Your Optimal Teacher: Personalized Data Synthesis via Router-Guided Multi-Teacher Distillation

Hengyuan Zhang<sup>1\*</sup>, Shiping Yang<sup>2\*</sup>, Xiao Liang<sup>3</sup>, Chenming Shang<sup>4</sup>, Yuxuan Jiang<sup>5</sup>, Chaofan Tao<sup>1</sup>

Jing Xiong<sup>1</sup>, Hayden Kwok-Hay So<sup>1</sup>, Ruobing Xie<sup>6</sup>, Angel X. Chang<sup>2,7</sup>, Ngai Wong<sup>1</sup>

DARTMOUTH

UCLA

ACL 2026 SAN DIEGO JULY 2-7

<sup>1</sup>The University of Hong Kong · <sup>2</sup>Simon Fraser University · <sup>3</sup>University of California, Los Angeles · <sup>4</sup>Dartmouth College · <sup>5</sup>University of Maryland, Baltimore County · <sup>6</sup>Tencent · <sup>7</sup>Canada-CIFAR AI Chair, Amii

## Background

In distillation scenario, will data produced by stronger model enable the student model to learn more effectively??



Is higher-quality or more complex synthetic data more suitable?

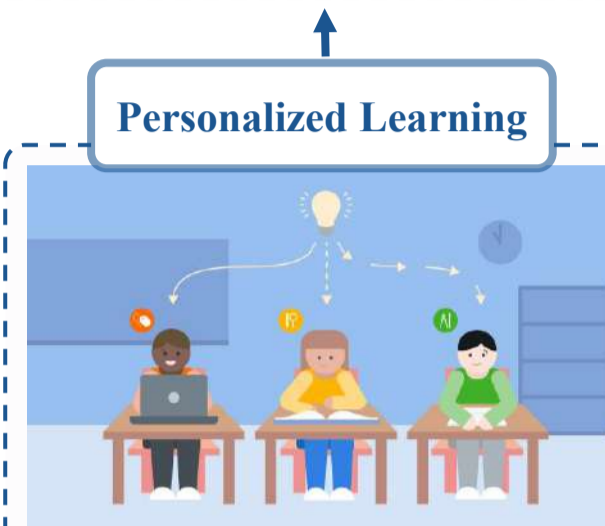


Synthetic data generated by stronger model may be overly complex and shift away from the students' distribution!

How to construct a high-quality and suitable synthetic data that is beneficial for student model learning?

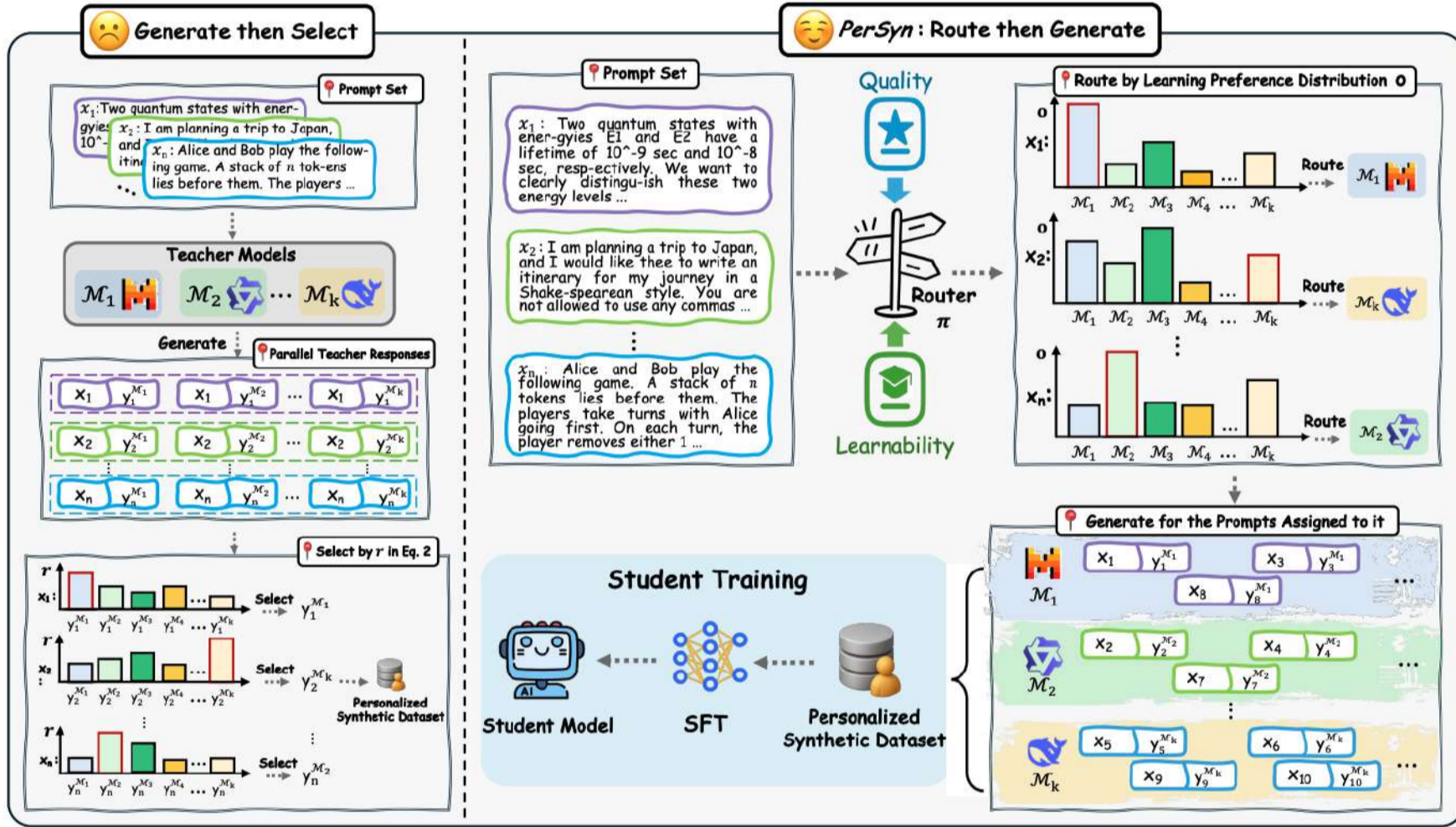
## Personalized Data Synthesis

➤ How We Build Personalized Synthetic Dataset: We select the optimal teacher for the student at the query level. Each prompt in the dataset is assigned to its most suitable teacher model.



➤ How We Select Optimal Teacher: We consider both the student model's learnability (estimated by average token log-probability) and teacher model's response quality (estimated by a reward model).

## Shift the "Generate then Select" to "Route then Generate" with a Router



### Router Inference

➤ Given a prompt  $x$ , router  $\pi$  will produce a distribution  $\mathbf{o} = \pi(x) \in \mathbb{R}^{|\mathcal{M}|}$ , which reflects the student model's learning preference over the teacher model set  $\mathcal{M}$ .

### Router Training

➤ We adopt the Bradley-Terry (BT) model. Given a comparison between two teacher models A and B

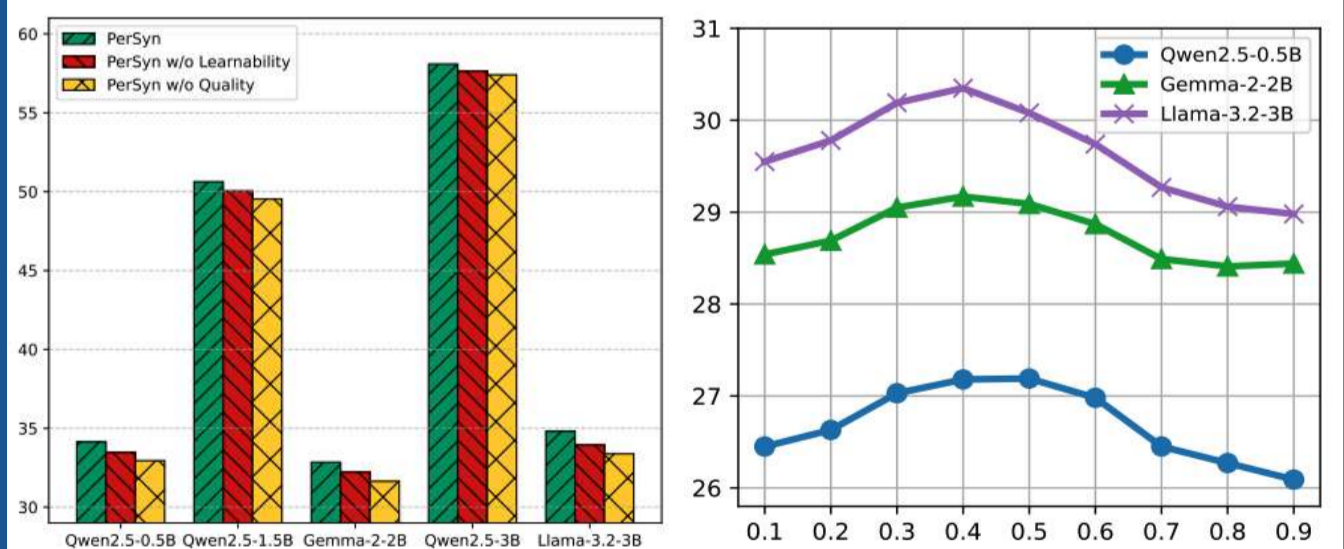
$$\mathbb{P}(C = B \succ A \mid Z = z, X = x) = \sigma(z^\top \pi(x)),$$

$$\hat{\pi} = \operatorname{argmin} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\sigma(Z_i^\top \pi(X_i)), C_i).$$

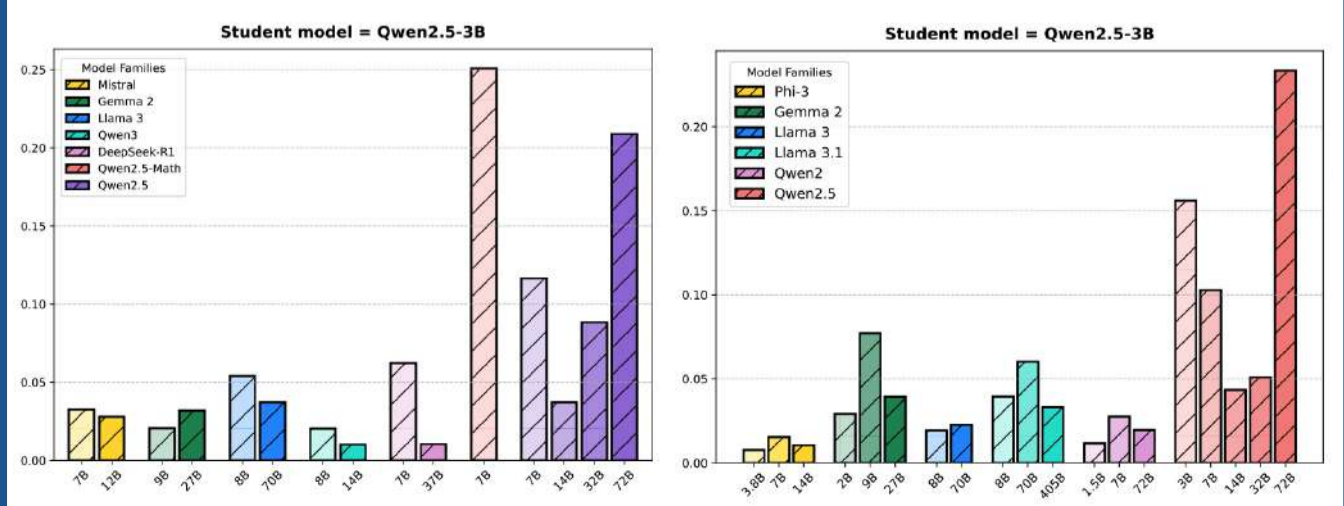
## Main Results

Student Model	Strategy	IFEval	TruthfulQA	LiveBench	GSM8K	MATH	SVAMP	Avg.
Qwen2.5-0.5B	Strong	25.59	39.89	8.40	30.37	15.20	51.60	28.51
	Mix	26.06	40.54	8.10	33.83	20.60	55.40	30.75
	Family-Strong	26.75	41.43	8.60	35.62	22.80	57.00	32.03
	CAR	27.11	41.85	9.00	36.76	24.00	57.90	32.77
	PerSyn (Ours)	<b>28.73</b>	<b>43.01</b>	<b>9.80</b>	<b>38.25</b>	<b>25.60</b>	<b>59.40</b>	<b>34.13</b>
Qwen2.5-1.5B	Strong	31.52	49.04	12.80	64.83	44.20	78.50	46.82
	Mix	31.98	49.73	13.30	65.68	45.80	80.30	47.79
	Family-Strong	32.63	50.45	13.60	66.55	47.40	81.20	48.64
	CAR	33.06	50.98	13.30	67.37	48.60	81.90	49.21
	PerSyn (Ours)	<b>34.15</b>	<b>52.22</b>	<b>14.80</b>	<b>68.81</b>	<b>50.40</b>	<b>83.40</b>	<b>50.63</b>
Gemma-2-2B	Strong	28.84	40.17	10.30	29.71	14.20	47.50	28.45
	Mix	29.39	40.83	10.90	31.66	16.40	49.40	29.76
	Family-Strong	29.76	41.64	11.60	30.43	15.80	48.10	29.56
	CAR	30.11	42.28	12.80	33.25	19.20	50.80	31.41
	PerSyn (Ours)	<b>31.25</b>	<b>43.87</b>	12.40	<b>35.57</b>	<b>21.40</b>	<b>52.60</b>	<b>32.85</b>
Qwen2.5-3B	Strong	40.61	51.21	19.10	77.47	56.40	87.50	55.38
	Mix	41.35	51.82	19.30	77.19	55.80	86.90	55.39
	Family-Strong	42.44	53.37	20.40	77.94	57.10	88.30	56.59
	CAR	43.03	53.81	20.90	78.42	58.10	88.80	57.17
	PerSyn (Ours)	<b>44.16</b>	<b>55.14</b>	<b>22.30</b>	<b>79.09</b>	<b>57.80</b>	<b>90.10</b>	<b>58.09</b>
Llama-3.2-3B	Strong	27.89	42.31	10.80	34.59	21.40	52.20	31.37
	Mix	29.78	43.56	12.00	34.17	20.50	51.50	31.75
	Family-Strong	28.25	42.63	11.10	33.83	19.50	50.80	30.85
	CAR	30.53	44.32	11.80	35.91	22.80	53.60	32.99
	PerSyn (Ours)	<b>32.31</b>	<b>46.15</b>	<b>12.60</b>	<b>38.15</b>	<b>24.50</b>	<b>55.30</b>	<b>34.81</b>

**Takeaway 1** Jointly considering both learnability and quality yields better performance than considering either alone, with quality playing a more critical role than learnability in PerSyn.



**Takeaway 2** Larger teacher models, despite their superior performance, are not always the optimal teacher for small student models; small teachers are often more suitable.



## Let's Connect!

### Contact Information

Hengyuan Zhang  
Email: hengyuan.zhang88@gmail.com

Shiping Yang  
Email: yangshipingnlp@gmail.com

