

Table 1: **Performance evaluation of the SpikingBrain-76B pre-trained model.** All models are tested with the vLLM kwon2023vllm framework and evaluated using a perplexity-based method. Except for Qwen2.5, the other baselines are trained on limited Chinese data, resulting in clear disadvantages on CMMLU and C-Eval.

	SpikingBrain-76B (76B-A12B)	Jamba lieber2024jamba (52B-A12B)	Mixtral-8*7B jiang2024mixtral (47B-A13B)	LLama2-70b touvron2023llama2	Gemma2-27B team2024gemma	Qwen2.5 Yang2024Qwen2TR (base)
Params	12B/76B	12B/52B	13B/47B	70B	27B	7B
Tokens	+160B	-	-	2T	13T	7T
Complexity Type	Hybrid	Hybrid	Quadratic	Quadratic	Quadratic	Quadratic
<b>Benchmarks</b>						
MMLU hendrycks2020mmlu	<b>73.58</b>	67.17	71.23	69.57	75.94	75.31
CMMLU li2023cmmlu	<b>78.83</b>	51.11	52.70	52.94	61.80	81.50
ARC-C clark2018arcc	<b>42.00</b>	49.10	48.98	46.21	57.64	43.56
HS zellers2019hellaswag	<b>73.31</b>	79.39	79.42	79.15	79.34	73.37
Ceval huang2023ceval	<b>78.89</b>	48.94	54.39	49.26	61.02	81.68