

Table 1: Performance comparison of SpikingBrain before and after applying the spiking scheme. Evaluations are performed on commonsense reasoning, MMLU, and CMMLU benchmarks. The integration of adaptive-threshold spike coding with 8-bit weight quantization results in negligible performance degradation (approximately 1–3%).

	Winogrande	Arc-e	Arc-c (norm)	HellaSwag (norm)	PIQA	MMLU	CMMLU	Avg.
SpikingBrain-7B	0.6992	0.8047	0.5566	0.6777	0.7949	0.6751	0.6904	0.6998
SpikingBrain-7B -W8ASpike	0.6895	0.7861	0.5410	0.6758	0.7979	0.6546	0.6677	0.6875
SpikingBrain -76B	0.7275	0.8125	0.5615	0.7000	0.8125	0.7247	0.7740	0.7304
SpikingBrain -76B-W8ASpike	0.7148	0.7949	0.5371	0.6863	0.8004	0.7081	0.7512	0.7133