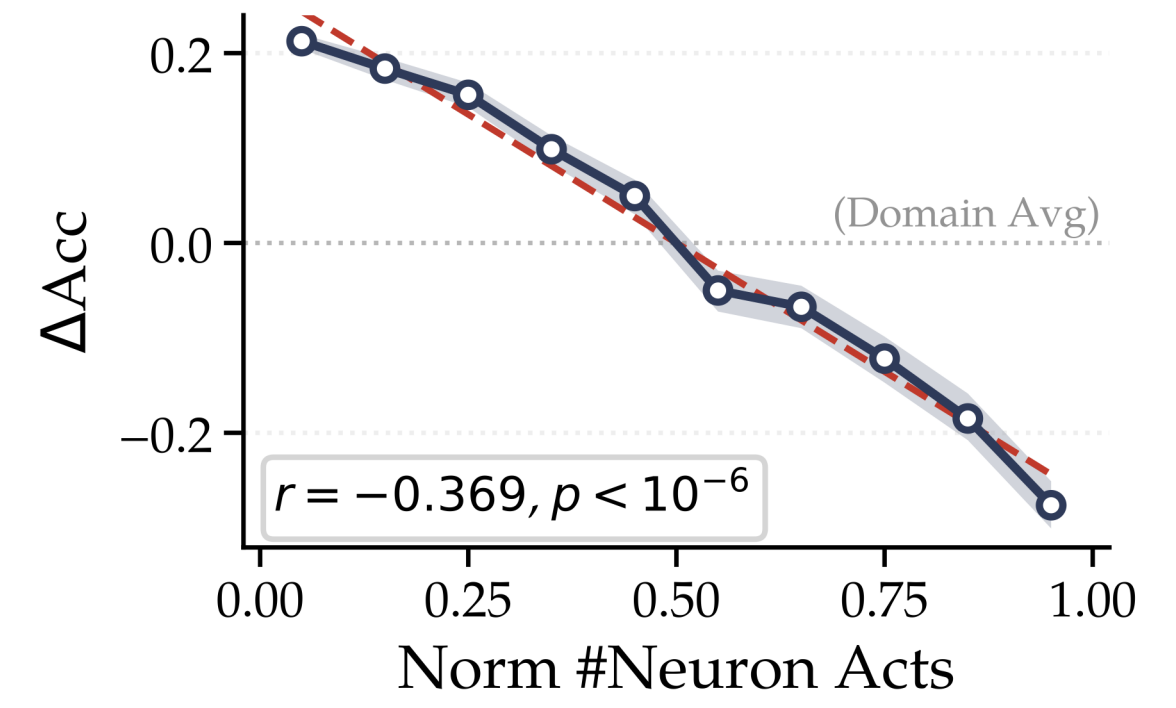
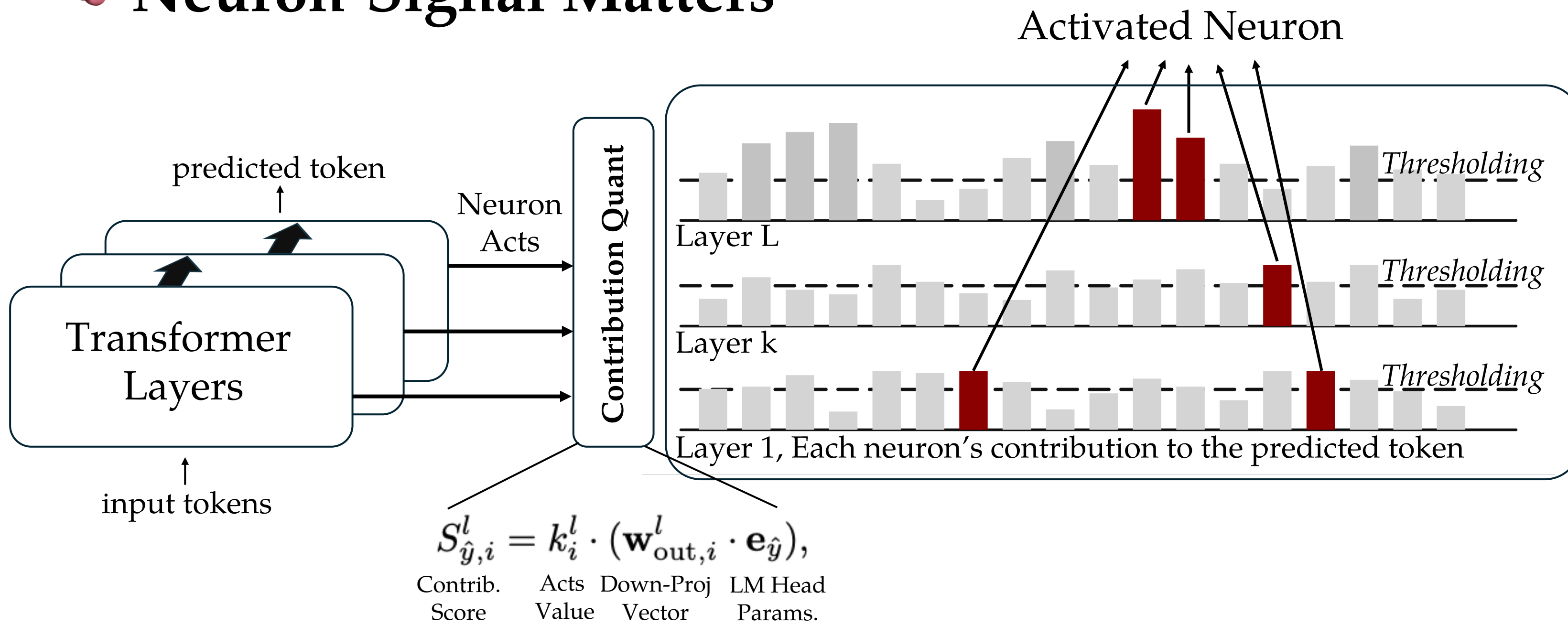


Neuron-Aware Data Selection for Annotation-Free LLM Self-Distillation

Zhuowei Chen, Xiang Lorraine Li @ University of Pittsburgh

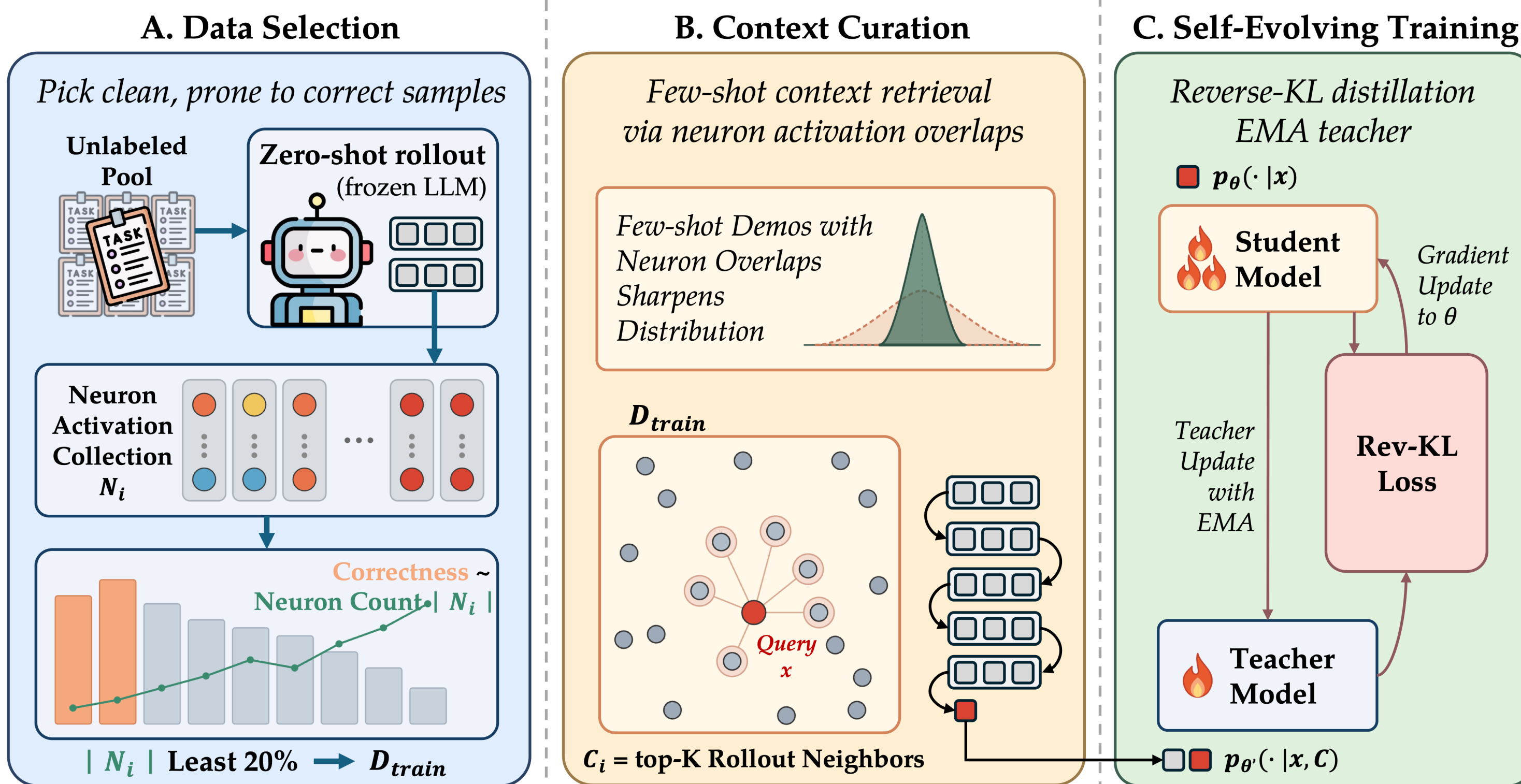
Neuron-Signal Matters



Sign of Hallucination

Indicate Reasoning Pattern

Data & Context Selection



For Annotation-Free Self-Distillation:

- Pick clean samples that LLMs can self-improve
- Select few-shot trajectory that presents similar reasoning chain as contexts
- Distill Knowledge from the context (Few-shot)

Results

Method	Eval.	SciKnowEval							
		BIO	MAT.	PHYS.	CHEM	Avg.	Edu.	MMLU-Pro	Avg.
Qwen3-4B base		73.98	71.12	80.20	72.11	74.35	64.95	72.16	72.42
LMSI	in	74.59 ^{+0.61}	69.08 ^{-2.04}	78.54 ^{-1.66}	73.72 ^{+1.61}	73.98 ^{-0.37}	66.95 ^{+2.00}	68.60 ^{-3.56}	71.91 ^{-0.51}
	cross	68.57 ^{-3.54}	65.48 ^{-7.20}	67.02 ^{-3.84}	71.19 ^{-1.29}	68.06 ^{-3.97}	58.14 ^{-15.77}	73.46 ^{+0.99}	67.31 ^{-5.11}
TTRL	in	75.00 ^{+1.02}	73.26 ^{+2.14}	82.98 ^{+2.78}	72.43 ^{+0.32}	75.92 ^{+1.56}	64.95 ^{+0.00}	72.04 ^{-0.12}	73.44 ^{+1.02}
	cross	73.31 ^{+1.20}	73.32 ^{+0.64}	71.28 ^{+0.41}	73.46 ^{+0.97}	72.84 ^{+0.81}	74.47 ^{+0.55}	73.76 ^{+1.29}	73.27 ^{+0.84}
Intuitior ²	in	76.21 ^{+2.23}	73.41 ^{+2.29}	83.43 ^{+3.23}	72.63 ^{+0.52}	76.42 ^{+2.07}	—	71.91 ^{-0.25}	75.52 ^{+1.60}
	cross	69.50 ^{-2.61}	72.79 ^{+0.11}	68.33 ^{-2.53}	70.76 ^{-1.72}	70.35 ^{-1.69}	—	73.30 ^{+0.83}	70.94 ^{-1.18}
N-OPSD	in	74.02 ^{+0.04}	73.39 ^{+2.27}	83.13 ^{+2.93}	72.03 ^{-0.08}	75.64 ^{+1.29}	72.19 ^{+7.24}	72.04 ^{-0.12}	74.47 ^{+2.05}
	cross	72.71 ^{+0.61}	73.27 ^{+0.59}	71.37 ^{+0.51}	73.59 ^{+1.11}	72.74 ^{+0.70}	75.22 ^{+1.31}	73.74 ^{+1.27}	73.32 ^{+0.90}

Table 3: Cross-domain evaluation on Qwen3-4B, Avg@8. For each method, *in* reports the score of the model trained and evaluated on the same source domain, while *cross* reports the average score of the same source-trained model on the other target domains. Subscripts report Δ against the untrained Qwen3-4B baseline computed on the same target-domain support. The full results are in Table 7.

Method	Eval.	SciKnowEval							
		BIO	MAT.	PHYS.	CHEM	Avg.	Edu.	MMLU-Pro	Avg.
Qwen3-4B base		0.195	0.213	0.095	0.173	0.169	0.246	0.184	0.184
LMSI	in	0.202 ^{+0.007}	0.225 ^{+0.012}	0.111 ^{+0.016}	0.179 ^{+0.006}	0.179 ^{+0.010}	0.255 ^{+0.009}	0.171 ^{-0.013}	0.191 ^{+0.007}
	cross	0.171 ^{-0.011}	0.179 ^{+0.001}	0.202 ^{+0.000}	0.182 ^{-0.004}	0.184 ^{-0.004}	0.095 ^{-0.077}	0.162 ^{-0.022}	0.165 ^{-0.019}
TTRL	in	0.204 ^{+0.009}	0.208 ^{-0.005}	0.123 ^{+0.028}	0.192 ^{+0.019}	0.182 ^{+0.013}	0.248 ^{+0.002}	0.198 ^{+0.014}	0.196 ^{+0.012}
	cross	0.186 ^{+0.004}	0.187 ^{+0.009}	0.204 ^{+0.002}	0.199 ^{+0.012}	0.194 ^{+0.007}	0.178 ^{+0.006}	0.190 ^{+0.006}	0.191 ^{+0.007}
Intuitior	in	0.208 ^{+0.013}	0.212 ^{-0.001}	0.117 ^{+0.022}	0.198 ^{+0.026}	0.184 ^{+0.015}	—	0.174 ^{-0.010}	0.182 ^{+0.010}
	cross	0.226 ^{+0.044}	0.187 ^{+0.008}	0.234 ^{+0.032}	0.224 ^{+0.038}	0.218 ^{+0.030}	—	0.185 ^{+0.000}	0.211 ^{+0.024}
N-OPSD	in	0.193 ^{-0.003}	0.207 ^{-0.006}	0.114 ^{+0.019}	0.171 ^{-0.002}	0.171 ^{+0.002}	0.191 ^{-0.056}	0.180 ^{-0.004}	0.176 ^{-0.008}
	cross	0.177 ^{-0.005}	0.177 ^{-0.001}	0.194 ^{-0.008}	0.187 ^{+0.000}	0.184 ^{-0.004}	0.177 ^{+0.005}	0.186 ^{+0.002}	0.183 ^{-0.001}

Table 4: Cross-domain calibration on Qwen3-4B, reported as ECE, computed from majority-vote confidence. For each method, *in* denotes source-domain ECE, while *cross* denotes the average ECE of the same source-trained model on other domains. The full results are in Table 8.

Findings

- Internal neuron activations are strong data-selection signals, shifting data selection beyond traditional output-level heuristics toward internal neuron analysis.
- Context selection is a key axis of self-improvement. Beyond selecting reliable training data, high-quality teacher contexts lead to stronger self-distillation gains and reduce performance/calibration collapse.
- Neuron-OPSD enables annotation-free self-distillation. The proposed neuron-aware data selection strategy improves task performance while better preserving calibration.
- OPD gains follow the teacher-student entropy gap. Selected contexts help most when they make the teacher distribution meaningfully sharper than the student's, producing a stronger distillation signal.