SelecTKD: Selective Token-Weighted Knowledge Distillation for LLMs

Haiduo Huang, Jiangcheng Song, Yadong Zhang, Pengju Ren*
Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University

paper wechat

CVPR JUNE 3-7, 2026 DENVER COLORADO

Background & Motivation

Introduction:

- **The Problem:** Standard Knowledge Distillation (KD) uniformly applies token-wise loss regardless of teacher confidence.
- **The Harm:** Indiscriminate supervision amplifies noisy, high-entropy signals, severely destabilizing optimization under large teacher-student capacity gaps.
- **Our Core Shift:** We reframe LLM distillation from **"how to measure divergence"** to **"where to apply learning"**.

Rethinking Loss Function Geometry:

- **Key Question:** Is loss geometry the dominant factor in LLM distillation?
- **Empirical Observation:** Different KL-family objectives (KL, RKL, SKL, SRKL) share surprisingly similar end performance under identical data curation.
- **Theoretical Core:** Forward, reverse, and skewed KLs share the exact same fixed point in the limit ; differences are mainly attributable to training dynamics.
- **Conclusion:** Controlling **where and when** the learning signal is applied at the token level matters more.

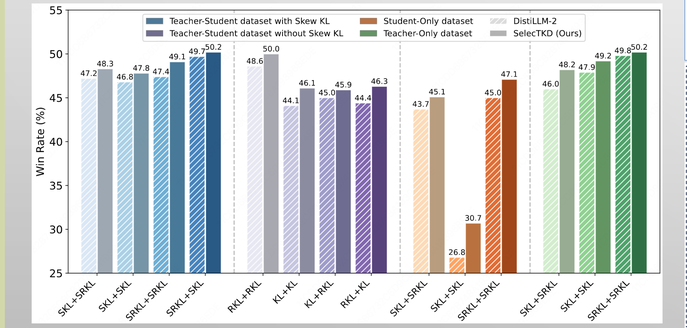

Figure 1. Performance comparison of different loss functions.

Methodology (The SelecTKD Framework)

Framework Over view:

➢ SelecTKD Loss Formulation:

$$\mathcal{L}_{SelecTKD} = \sum_{t=1}^{|y|} \mathrm{V}_t D(p_t \parallel q_t)$$

- $D(\cdot\|\cdot)$: Generic token-wise divergence (objective-agnostic).
- $\mathrm{V}_t \in \{\beta, 1\}$: Dynamic token-level verification weight.
- β: Weight for rejected tokens ($\beta = 0.01$ provides gentle regularization).

Two Variants of Propose-and-Verify Mechanism:

➢ **Variant A**: Greedy Top-k Verification

1. Student proposes its most likely greedy token:

$$\hat{y}_t = arg\max_y q_\theta(y|x, y_{<t})$$

2. Teacher verifies if $\hat{y}_t$ lies within its top-k candidates:

$$\mathrm{V}_t = \beta + (1-\beta)\mathbb{I}(\hat{y}_t \in Top_k(p_t))$$

➢ **Variant B**: Non-Greedy Spec-k Verification

1. Student draws k i.i.d. candidates from its own distribution:

$$\{y_t^{(j)}\}_{j=1}^{k} \sim \text{i.i.d. } q_\theta(\cdot\,|x, y_{<t})$$

2. Evaluate candidates via speculative-acceptance test:

$$a_t^{(j)} = \min\left(1, \frac{p_t(y_t^{(j)})}{q_t(y_t^{(j)})}\right)$$

3. Accept step if at least one sample passes the randomized test:

$$\mathrm{V}_t = \beta + (1-\beta)\mathbb{I}(|\mathrm{A}_t| \geq 1)$$

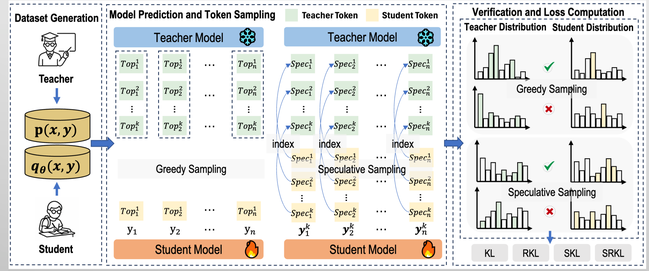

Figure 2. Overview of SelecTKD for selective token-level knowledge distillation.

Theoretical Insights

Token Acceptance Rate (TAR) & Implicit Curriculum :

◆ TAR Metric: Measures the expected fraction of fully-weighted tokens:

$$TAR = \mathbb{E}_{(x,y)}\left[\frac{1}{|y|}\sum_{i=1}^{|y|}\mathbb{I}(\mathrm{V}_t = 1)\right]$$

◆ **Theorem 1 (Monotonic TAR Improvement):** Under standard Lipschitz and continuity conditions, each gradient step guarantees quasi-monotonic growth of TAR:

$$TAR_{t+1} - TAR_t > \eta\kappa(1 - TAR_t)$$

◆ **Implicit Curriculum Effect:** Optimization naturally self-paces—focusing on "easy", aligned tokens first, and gradually introducing harder ones as the student improves.

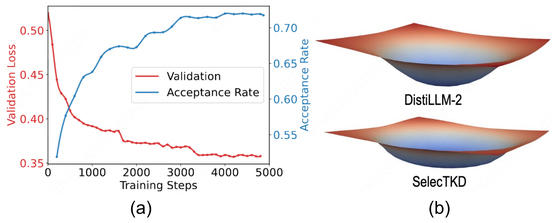

Figure 3. **SelecTKD Dynamics:** (a) Validation loss decreases as TAR increases, indicating an implicit curriculum. (b) SelecTKD yields a flatter loss landscape than DistiLLM-2, correlating with better generalization.

Main Results & Conclusions

1. Quantitative Benchmarks

General Instruction-Following (Win Rate % vs. GPT-3.5)

- Qwen2-7B-Inst → Qwen2-1.5B: DistiLLM-2 baseline achieves 47.13% WR; adding SelecTKD boosts performance to 48.27% (+1.14%).
- Outperforms prior token-selection pipelines (AdaSPEC: 40.21%, ATKD: 45.13%) without extra reference models.

Downstream Task Performance (Pass@1 Accuracy %)

- Mathematical Reasoning (GSM8K / MATH):
  ✓ Qwen2-Math-7B → 1.5B: Boosts overall average Pass@1 from 55.93% to 57.36% (+1.43%).
- Code Generation (HumanEval / MBPP)::
  ✓ DS-Coder-6.9B → 1.3B: Upscales average Pass@1 from 67.79% to 69.00% (+1.21%).

Vision-Language Models (Open VLM Leaderboard)

- **SelecTKD-VLM-2B** scores an average of **57.9%**, outperforming larger models like Phi-3-Vision (4.2B) and InternVL2-2B baseline (54.0%)

2. Speculative Decoding Speedup

- SelecTKD aligns student-teacher distributions tightly, yielding a × 2.05 speedup on Phi-3-medium verifier, surpassing DistiLLM-2 (× 1.97).

3. Mitigating the Curse of the Powerful Teacher

While Vanilla KD performance drops as teacher size scales excessively, SelecTKD guarantees **monotonic student capability improvement** by effectively shielding the student from high-entropy, unlearnable teacher signals.