

# RetryTrigger: 面向硬件瞬态故障的大语言模型智能推理重试容错方法

焦佳佳, 于亦许

上海海事大学信息工程学院, 中国上海市, 201306

**摘要:** 近年来, 大语言模型在各类自然语言处理任务中展现卓越性能。然而, 随着硬件瞬态故障的日益频发, 大语言模型中的静默数据损坏问题愈发突出, 极大损害了输出质量与用户体验。现有主流的保护方案主要依赖于硬件辅助的算法级容错, 或针对模型部分层的边界驱动在线容错方法。但是, 这些方法往往存在硬件依赖严苛、性能开销巨大或故障覆盖不全面等问题。为克服这些局限性, 本文提出一种新颖的无需额外硬件支持的故障感知推理方法 **RetryTrigger**, 以全面应对各类瞬态故障。在大语言模型推理过程中, **RetryTrigger** 会动态收集运行时输出特征 (如最大概率、**top-k** 概率差、输出熵、**logits** 统计信息以及推理延迟), 并将这些特征输入至 **LightGBM** 元模型进行判定。该元模型能够准确预测是否需要触发推理重试, 从而在不依赖额外硬件的前提下, 实现推理效率与故障缓解的双重保障。本文在 7 种具有代表性的大语言模型 (含 **T5-Small**、**RoBERTa**、**BioMedBERT**、**Qwen2.5-Coder-0.5B/7B**、**MiniMind** 和 **Opt**) 上进行了大量实验, 结果表明, **RetryTrigger** 最多可将静默数据损坏率降低 95.33% (平均降低 92.97%), 同时实现最低 2.4012% (平均 4.1167%) 的性能开销。与现有前沿解决方案相比, 该方法在可靠性与推理效率之间实现了更优的权衡。

**关键词:** 大语言模型; 系统弹性; 智能故障检测; 推理复算; 瞬态故障  
<https://doi.org/10.1631/ENG.ITEE.2025.0104>