Privacy-Preserving LLM Workflows (2022)

Abstract: This paper (2022) investigates privacy-preserving llm workflows in the context of AI Systems. We propose a workflow combining local inference, artifact traceability, prompt stages. Method: Rule-based scoring + statistical validation. Dataset: Synthetic transaction stream (n=1,000,000) with injected anomalies. Metrics: Throughput gain (%), error reduction (%).

Methods & Data: Rule-based scoring + statistical validation. Synthetic transaction stream (n=1,000,000) with injected anomalies.

Results: Our approach improves Macro-F1 by 6–9% over the baseline on the main test set.

Limitations: Model performance degrades under concept drift over 6 months.