Meta-Learning Adaptation Impact on Few-Shot Anomaly Detection Accuracy of Llama-3.1-8B vs. Llama-3.1-70B in Out-of-Domain Text

Assignee Research

June 11, 2026

Abstract

We propose a meta learning framework for detecting anomalies in human language across diverse domains with limited labeled data. Anomalies in language ranging from spam and fake news to hate speech pose a major challenge due to their sparsity and variability. We treat anomaly detection as a few shot binary classification problem and leverage meta-learning to train models that generalize across tasks. Using datasets from domains such as SMS spam, COVID-19 fake news, and hate speech, we evaluate model generalization on unseen tasks with minimal labeled anomalies. Our method combines episodic tra

1 Introduction

This paper examines: Anomaly Detection in Human Language via Meta-Learning: A Few-Shot Approach. Research question: How does meta-learning adaptation affect few-shot anomaly detection accuracy of Llama-3.1-8B versus Llama-3.1-70B when evaluated on out-of-domain text classification tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 5 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.5/10.

3 Results

5 papers retrieved. 12 claims extracted; 9 independently verified. Quality review score: 7.5/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The performance of each method is reported using ROC-AUC and Average Precision (AP) for each dataset, as well as the F1-	×	0.08
Results are averaged over 5 runs (std dev in parentheses).	×	0.01
The best results for each dataset are bolded in Table 2.	×	0.02
Language-based anomaly detection is formalized as a one-vs-rest classification problem.	✓	0.16
The input space of texts is denoted as $\mathcal{X}$ and the labels as $\mathcal{Y} = {\text{Normal}, \text{Anomalous}}$	✓	0.17
Multiple datasets $D^{(1)}, D^{(2)}, ..., D^{(M)}$ are assumed, each corresponding to a different domain or anomaly dete	✓	0.17
In dataset $D^{(i)}$, the vast majority of instances are normal (negative class), and a small fraction (e.g. 1–5%) are a	✓	0.26
During meta-training, labeled data from a set of source tasks $D^{(1)} \dots D^{(M)}$ is available.	✓	0.18
During meta-testing, the goal is to adapt to a new anomaly detection task using only a small number k of labeled anoma	✓	0.27
A task T for anomaly detection is defined as $T = (\mathcal{D}_T^{\text{train}}, \mathcal{D}_T^{\text{test}})$.	✓	0.31
$\mathcal{D}_T^{\text{train}}$ (support set) contains a few labeled examples of normal and anomalous text from the task'	✓	0.38
$\mathcal{D}_T^{\text{test}}$ (query set) contains additional unseen examples from the same domain, on which the model's	✓	0.36

References

- http://arxiv.org/abs/2003.04390v4
- http://arxiv.org/abs/2507.20019v1
- http://arxiv.org/abs/2007.06837v6