Robustness of Zero-Shot Topic Classification with Pre-trained Multimodal Models under Noisy Relational Information

Assignee Research

June 13, 2026

Abstract

Multi-label topic classification without labeled training data is a challenging task, specially when documents contain complex relational information. We present a zero-shot multi-label topic classification framework and systematically investigate how per-article knowledge graph augmentation affects its performance. The base framework classifies topics in documents without labeled training data and has four variants: article-only classification, keyword-enhanced classification, and self-consistency decoding variants of both. Then, we augment each base variant with per article knowledge graph.

1 Introduction

This paper examines: Knowledge Graph-Enhanced Zero-Shot Topic Classification: A Multi-Strategy Comparative Study. Research question: To what extent do pre-trained multimodal models (e.g., CLIP, BLIP) improve the robustness of zero-shot topic classification frameworks when given noisy or incomplete relational information in documents, as measured by accuracy on subsets of the OpenML benchmark with artificially introduced noise?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.2/10.

3 Results

15 papers retrieved. 11 claims extracted; 9 independently verified. Quality review score: 7.2/10.

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 study evaluates a zero-shot multi-label topic classification framework across fifteen LLMs and eight datasets.	✓	0.17
The framework includes four base variants: Article Only (AO), Article + Keywords (AK), and self-consistency decoding for	✓	0.17
The study defines four graph-augmented variants (AG, AKG, AGS, AKGS) that add graph context to each base method.	✓	0.21
The knowledge graph construction pipeline used in the study is adapted from KGGen (Mo et al., 2025).	✓	0.19
Graph augmentation consistently improves performance for smaller models.	×	0.11
Graph augmentation is most effective when combined with keyword-enhanced classification with self-consistency.	✓	0.19
Self-consistency decoding increases computational cost approximately fivefold.	✓	0.21
Self-consistency decoding does not improve classification performance in any of the tested settings.	×	0.09
The study utilizes the English subset of the SemEval-2018 Task 1 dataset.	✓	0.16
Probabilistic approaches to topic discovery date back to PLSA (Hofmann, 1999).	✓	0.16
LDA introduced a document-level generative process and became the dominant unsupervised topic model.	✓	0.17

References

- http://arxiv.org/abs/2410.01534v2
- http://arxiv.org/abs/2605.30465v1
- http://arxiv.org/abs/2403.10499v1