

# Scaling Performance of Knowledge Graph-Enhanced Multi-Label Topic Classification with Graph Density Across Language Pairs

Assignee Research

June 11, 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: How does the performance of knowledge graph-enhanced multi-label topic classification models scale with increasing graph density, measured by precision-recall and F1-score across different language pairs?.

## 2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.4/10.

## 3 Results

13 papers retrieved. 8 claims extracted; 6 independently verified. Quality review score: 7.4/10.

## 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 with four variants (AO, AK, AOS, AKS) across	✓	0.18
The study proposes a per-article knowledge graph construction pipeline adapted from KGGen (Mo et al., 2025) and defines	✓	0.25
Graph augmentation consistently helps smaller models and is most effective when combined with keyword-enhanced classific	✓	0.20
Self-consistency decoding increases computational cost approximately fivefold but does not help with classification perf	✓	0.25
The study uses the English subset of the SemEval-2018 Task 1 dataset (Mohammad et al., 2018; Sarkar et al., 2023).	✓	0.23
The study compares against established sentence-encoder baselines and examines whether structured relational context imp	✓	0.21
The study includes benchmark tables with performance metrics for various models and methods on different datasets.	×	0.07
The study includes a runtime and cost analysis showing the computational cost of self-consistency decoding.	×	0.14

## References

- <http://arxiv.org/abs/2606.11371v1>
- <http://arxiv.org/abs/2605.30465v1>
- <http://arxiv.org/abs/2310.10378v5>