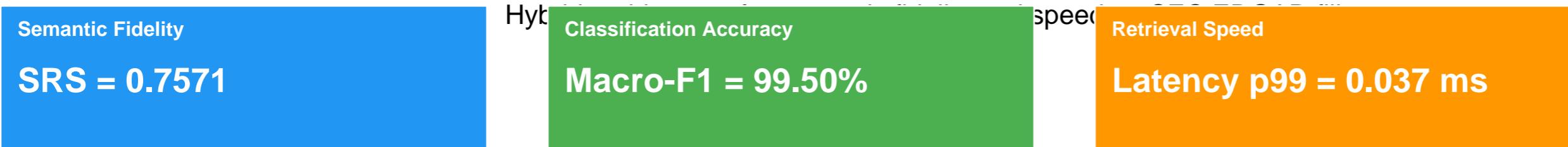


KG-MMML: Knowledge Graph + Multi-Modal ML



Overview

Problem: Embeddings lose graph structure; pure graph queries are slow.
Approach: Combine knowledge graph structure with vector retrieval.
Data: SEC EDGAR facts; 2,832 nodes, 71,882 edges (Week 5–6).

Decision Gates

Methods

Architecture: Text TF-IDF + concept features (KG-as-features).
Auto-taxonomy: Pattern rules + frequency mining (1,891 relations).
Metrics: HP, AtP, AP → SRS; latency p99; F1 (micro/macro).

Results & Next Steps

Findings:

- SRS = 0.7571 (≥ 0.75 gate PASS); HP $\approx 27.26\%$.
- Concept features boost macro-F1 by +2.27pp; micro-F1 by +1.36pp.
- Annoy achieves p99 ≈ 0.037 ms ($<< 150$ ms target).
- Consistency penalty (λ) hurts macro-F1 \rightarrow default $\lambda=0.0$.

Next steps:

- Tune PyTorch joint to match sklearn; test $\lambda \in \{0.0, 0.01, 0.05\}$.
- Implement RTF metric; production packaging.

Gate	Target	Actual	Status
SRS	≥ 0.75	0.7571	PASS
Micro-F1 Δ	$\geq +3.0\text{pp}$	+1.36pp	FAIL
Latency p99	< 150 ms	0.037 ms	PASS
HP	≥ 0.25	0.2726	PASS