SOVEREIGN: How does the effectiveness of negative sampling for unanswerable questions in the MRQA dataset compare to SQuA

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

In the last few years, the deep learning (DL) computing paradigm has been deemed the Gold Standard in the machine learning (ML) community. Moreover, it has gradually become the most widely used computational approach in the field of ML, thus achieving outstanding results on several complex cognitive tasks, matching or even beating those provided by human performance. One of the benefits of DL is the ability to learn massive amounts of data. The DL field has grown fast in the last few years and it has been extensively used to successfully address a wide range of traditional applications. More i

1 Introduction

Analysis of: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Research goal: How does the effectiveness of negative sampling for unanswerable questions in the MRQA dataset compare to SQuAD 2.0 in terms of F1 and exact match scores, and what is the impact on inference latency when using this sampling strategy versus random sampling across different transformer architectures?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

9 papers retrieved. 5 claims extracted, 5 verified. Tribunal: 7.7/10 → APPROVE (revision_round=1). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
Deep learning has become the Gold Standard in the machine learning community	✓	0.15
Deep learning achieves outstanding results on several complex cognitive tasks	✓	0.16
Deep learning has gradually become the most widely used computational approach in the field of machine learning	✓	0.23
Deep learning has outperformed well-known ML techniques in many domains including cybersecurity, natural language proces	✓	0.30
The field has grown fast in the last few years	✓	0.18

References

- https://doi.org/10.1136/bmj.g7647
- https://doi.org/10.1186/s40537-021-00444-8
- https://doi.org/10.18653/v1/2020.findings-emnlp.171