Domain Adaptation (E.G., Fine-Tuning On Legal Or Biomedical Text) Influence The Robustness Of Baichuan-2'S

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: How does domain adaptation (e.g., fine-tuning on legal or biomedical text) influence the robustness of Baichuan-2's hallucination detection with different TAE misalignment thresholds on the FactCC. In the era of digital healthcare, the huge volumes of textual information generated every day in hospitals constitute an essential but underused asset that could be exploited with task-specific, fine-tuned biomedical language representation models, improving patient care and. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Localising In-Domain Adaptation of Transformer-Based Biomedical Language Models. Research question: How does domain adaptation (e.g., fine-tuning on legal or biomedical text) influence the robustness of Baichuan-2's hallucination detection with different TAE misalignment thresholds on the FactCC benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.0/10.

3 Results

4 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 3.0/10.

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
Our model has been initialized with a monolingual Italian version of BERT, obtained from a recent Wikipedia dump and var	×	0.03
The BaseBIT model was used as the baseline and is a monoling up to 81GB and 13B tokens.	×	0.01
No other large-scale, pretrained BERT or BERT-derived checkpoints were available for Italian at the time of designing th	×	0.06
Over 12 thousand BERT-based models are hosted in the Huggingface model repository, covering more than 20 different non-E	×	0.05
NMT has been shown to work well in clinical settings, such as for translating abstractions of clinical trials published	×	0.05
It is common practice for less-resourced languages to leverage translation in the opposite direction, to unlock the oppo	×	0.07
We investigate whether NMT systems are mature enough to do the opposite, starting from the English source to develop loc	×	0.02

References

- http://arxiv.org/abs/2212.10422v3
- http://arxiv.org/abs/1902.02401v1

- http://arxiv.org/abs/2310.05276v1