The Integration Of Domain-Specific Comments From Q&A Forums Performance On The Robustness Of Code Generation Models

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: How does the integration of domain-specific comments from Q&A forums affect the robustness of code generation models against adversarial prompts in the MBPP dataset. 10 claims were extracted from source literature; 3 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: On Adversarial Robustness: A Neural Architecture Search perspective. Research question: How does the integration of domain-specific comments from Q&A forums affect the robustness of code generation models against adversarial prompts in the MBPP dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.3/10.

3 Results

14 papers retrieved. 10 claims extracted; 3 independently verified. Quality review score: 5.3/10.

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
DARTS architectures having 20 cells are trained for 600 epochs on the CIFAR-10 dataset.	×	0.04
Ensemble of architectures are known to be adversarially more robust compared to single architecture.	×	0.05
Adversarial training decreases the accuracy on clean (un-perturbed) samples.	×	0.08
The primary motivation of the work is to develop architectures that are inherently robust without any form of adversaria	×	0.12
The work aims to understand adversarial robustness purely from an architectural perspective.	✓	0.20
The work is the first attempt at understanding adversarial robustness from a Neural Architecture Search perspective.	✓	0.20
The ensemble of randomly sampled DARTS architectures is trained for 600 epochs to ensure a fair comparison with existing	×	0.03
The number of epochs for training each network in the ensemble is determined based on the number of cells in that networ	×	0.04
The work compares NAS-based architectures with hand-crafted architectures in terms of adversarial robustness.	✓	0.29
The work investigates whether the source of adversarial vulnerability lies in the search space or in the way current met	×	0.08

References

- `http://arxiv.org/abs/2306.11066v2`

- http://arxiv.org/abs/2007.08428v4
- http://arxiv.org/abs/2604.21950v1