Feature Squeezing and Adversarial Training for Robust CodeT5 in Code Completion

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: Does integrating feature squeezing preprocessing with adversarial training improve CodeT5's defense success rate against multi-step PGD attacks in code completion tasks. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Robust Image Classification: Defensive Strategies against FGSM and PGD Adversarial Attacks. Research question: Does integrating feature squeezing preprocessing with adversarial training improve CodeT5's defense success rate against multi-step PGD attacks in code completion tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.3/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 2.3/10.

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.

References

- http://arxiv.org/abs/2408.13274v1
- http://arxiv.org/abs/1812.03411v2
- http://arxiv.org/abs/2205.14230v2