

Semantic Guidance in Adversarial Training for Cross-Dataset Code Model Robustness

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: Does applying semantic guidance in adversarial training improve cross-dataset robustness for code models when evaluated on MBPP versus CodexGLUE. Predicting the trajectories of surrounding objects is a critical task for self-driving vehicles and many other autonomous systems. Recent works demonstrate that adversarial attacks on trajectory prediction, where small crafted perturbations are introduced to history. 0 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: Semi-supervised Semantics-guided Adversarial Training for Trajectory Prediction. Research question: Does applying semantic guidance in adversarial training improve cross-dataset robustness for code models when evaluated on MBPP versus CodexGLUE?.

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 3.0/10.

3 Results

14 papers retrieved. 0 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.

References

- <http://arxiv.org/abs/2205.14230v2>
- <http://arxiv.org/abs/2103.01400v3>
- <http://arxiv.org/abs/2310.10807v1>