

# Subword Sampling and Gaussian Noise Augmentations in CodeT5 Consistency Regularization for MBPP

Assignee Research

June 7, 2026

## Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How does the inclusion of subword sampling and Gaussian noise augmentations in consistency regularization affect the performance of CodeT5 models on the MBPP benchmark compared to other adversarial. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates. Research question: How does the inclusion of subword sampling and Gaussian noise augmentations in consistency regularization affect the performance of CodeT5 models on the MBPP benchmark compared to other adversarial perturbation techniques?.

## 2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.5/10.

## 3 Results

8 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 2.5/10.

## 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/1807.09380v3>
- <http://arxiv.org/abs/2110.03135v4>
- <http://arxiv.org/abs/1804.10959v1>