OpenPangu-7B-MLA Robustness Scaling in Adversarial Spoken Language Understanding

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does OpenPangu-7B-MLA's robustness in spoken language understanding tasks scale with model size when tested on MMSU under adversarial noise conditions, using F1-score as the evaluation metric. 13 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: EchoMind: An Interrelated Multi-level Benchmark for Evaluating Empathetic Speech Language Models. Research question: How does OpenPangu-7B-MLA's robustness in spoken language understanding tasks scale with model size when tested on MMSU under adversarial noise conditions, using F1-score as the evaluation metric?.

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 3.8/10.

3 Results

11 papers retrieved. 13 claims extracted; 1 independently verified. Quality review score: 3.8/10.

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
EchoMind is an interrelated multi-level benchmark for evaluating empathetic speech language models.	✓	0.32
The authors of the paper are affiliated with The Chinese University of Hong Kong, Shenzhen, Shenzhen Research Institute	×	0.01
EchoMind evaluates models on multiple dimensions including understanding, reasoning, conversation, content, and voice.	×	0.08
EchoMind uses both text and audio as input and output.	×	0.04
EchoMind includes a style dimension in its evaluation.	×	0.03
EchoMind evaluates models on multiple dimensions including reasoning and response (audio).	×	0.08
EchoMind evaluates models on text context fit, text conversation naturalness, text colloquial degree, and text speech re	×	0.07
EchoMind evaluates models on context fit, speech relevance, and voice emotional similarity (VES).	×	0.05
EchoMind evaluates models on audio quality using NISQA and DNMOS metrics.	×	0.04
EchoMind evaluates models on emotional alignment using EmoAlign and VES metrics.	×	0.05
EchoMind evaluates models on reasoning accuracy using Acc metric.	×	0.03
EchoMind evaluates models on semantic similarity using SemSim metric.	×	0.04
EchoMind evaluates models on word error rate using WER metric.	×	0.02

References

- http://arxiv.org/abs/2306.11066v2
- http://arxiv.org/abs/2510.22758v2
- http://arxiv.org/abs/2207.08179v1