

Do Baichuan 2 and Vicuna-13B exhibit different sensitivity curves in alignment scores across multimodal benchm

Assignee Research

May 29, 2026

Abstract

Building on the foundations of language modeling in natural language processing, Next Token Prediction (NTP) has evolved into a versatile training objective for machine learning tasks across various modalities, achieving considerable success. As Large Language Models (LLMs) have advanced to unify understanding and generation tasks within the textual modality, recent research has shown that tasks from different modalities can also be effectively encapsulated within the NTP framework, transforming the multimodal information into tokens and predict the next one given the context. This survey intr

1 Introduction

This paper examines: Next Token Prediction Towards Multimodal Intelligence: A Comprehensive Survey. Research question: Do Baichuan 2 and Vicuna-13B exhibit different sensitivity curves in alignment scores across multimodal benchmarks when subjected to incremental TAE token misalignment thresholds?.

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 7.2/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 7.2/10.

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/2407.07530v1>
- <http://arxiv.org/abs/2412.18619v2>
- <http://arxiv.org/abs/2309.10305v4>