

Comprehensive Explanation of Revisions

In response to the May 2025 round of peer review, we undertook targeted revisions to improve our work.

Linguistics Knowledge:

We introduced a new section (Section 2) titled “Language Perspective,” which discusses the Santali language and highlights its key similarities with Bengali and Hindi, as suggested by the reviewers.

Dataset Descriptions:

For a fresh evaluation, we discarded the previous experiments and re-trained all approaches using the latest available dataset. This change is described in Section 4.1.

Research Question:

We updated and introduced a new research question (R4) to investigate how Parameter Efficient Fine-Tuning (PEFT), specifically LoRA fine-tuning, performs in low-resource dataset scenarios and what factors contribute to the observed results, as suggested by the reviewers. This study is described under Section 5, RQ4.

Qualitative Analysis:

As suggested by both reviewers and the meta-reviewer, we introduced a new section (Section 6) titled “Error Analysis.” In this section, we discuss the different types of errors encountered in the best-performing model (pre-trained Bengali Whisper Small) output.

Evaluation Benchmarking:

We also evaluated the IndicConformer model, proposed by AI4Bharat, on the Common Voice test dataset, which is common across all our evaluations. The results are reported in Table 6.