

ADI-IITP: Reward-Guided Preference Optimization for Low-Resource English-Assamese Machine Translation

Aditya¹ Asif Ekbal²

¹School of Computer Engineering, KIIT Deemed to be University, Odisha, India

²Department of Computer Science and Engineering, Indian Institute of Technology Patna, India
adityazn141@gmail.com asif.ekbal@gmail.com

Abstract

This paper describes our submission to the WMT 2026 shared task on Low-Resource Indic Language Translation for the English-Assamese pair. Our systems build on the compact, distilled 200M-parameter IndicTrans2 model, adapted with parameter-efficient Low-Rank Adaptation (LoRA). Beyond standard supervised fine-tuning (SFT), we built a preference-optimization pipeline that generates candidate translations, scores them with a composite reward combining a GEMBA-style LLM-as-judge score, a reference-free CometKiwi quality-estimation signal, and an xCOMET score, and then trains with Direct Preference Optimization (DPO) on the resulting preference pairs. We used this pipeline to guide our system choices: for Assamese-to-English the SFT-only model scored best on our development set, so we submitted it as primary and the SFT-then-DPO model as a contrastive system; for English-to-Assamese we went with SFT-then-DPO as primary. On the official WMT 2026 blind test set, the Assamese-to-English primary system scored 24.08 BLEU (25.11 for the contrastive system), and the English-to-Assamese primary system scored 15.57 BLEU. We also report METEOR, TER, chrF++, BERTScore, and COMET.

1 Introduction

Our team participated in the Low-Resource Indic Language Translation task for the English-Assamese language pair in the WMT 2026 shared task (Pakray et al., 2026)¹. The shared task involves developing machine translation (MT) systems with relatively small parallel datasets for under-resourced Indian languages (Pal et al., 2023; Pakray et al., 2024, 2025). Neural Machine Translation (NMT) has been a leading paradigm for translating human languages in recent years; however, its application to low-resource languages

such as Assamese remains limited, primarily due to the scarcity of large parallel corpora.

Motivated by the strong performance of open multilingual pretrained models, we study how competitive a compact distilled model can be under minimal, targeted adaptation within the constraints of a shared GPU. Specifically, we develop and assess models for translating from Assamese to English and from English to Assamese, built on the 200M-parameter distilled IndicTrans2 model and fine-tuned with LoRA-based supervised training (Singh et al., 2026a). In addition, we build a complete preference-optimization pipeline, using a composite reward and DPO (Rafailov et al., 2023; Xu et al., 2025), to guide our final system selection. The predicted translations are evaluated using both automatic and qualitative evaluation metrics.

1.1 About the Language

Assamese is an Indo-Aryan language spoken predominantly in the North-East Indian state of Assam, where it serves as the official language and lingua franca. It is one of the 22 scheduled languages of India listed in the 8th Schedule of the Indian Constitution. Assamese is the easternmost Indo-Aryan language and is written in the Eastern Nagari (Bengali-Assamese) script, sharing most of its character set with Bengali while using a small number of distinct letters. Despite having a rich literary tradition and tens of millions of speakers, Assamese remains under-resourced for natural language processing, with comparatively a very limited amount of parallel corpora available for machine translation.

2 Related Work

Machine translation (MT) for low-resource languages has evolved from statistical machine translation (SMT) systems, which relied on phrase-based translation models and parallel corpora, to neural machine translation (NMT), which signif-

¹Our code is publicly available at <https://github.com/AdityaSingh141-design/adi-iitp-wmt2026>.

icantly improved translation quality through end-to-end encoder-decoder architectures (Gain et al., 2026). While NMT has achieved remarkable success for high-resource languages, its performance remains limited for low-resource language pairs due to the scarcity of parallel training data.

To mitigate this limitation, researchers have explored multilingual transfer learning, back-translation, and semi-supervised learning (Sennrich et al., 2016). The introduction of multilingual pretrained MT models, such as IndicTrans2 (Gala et al., 2023), has substantially improved translation quality for many under-resourced Indic languages through parameter sharing and transfer across related languages. For Manipuri specifically, recent work has explored script-level challenges such as bidirectional transliteration between Manipuri Bengali and Manipuri Latin scripts (Singh et al., 2026b) and byte-level translation approaches (Singh et al., 2026c). Related shared-task systems have investigated instruction-tuned adaptation for other under-resourced pairs, including Bodo and Kokborok (Kumar et al., 2025) and Santali (Singh et al., 2025a). Assamese itself has received comparatively little attention, with recent work limited to a pragmatic comparison of transfer-learning and from-scratch strategies (Meetei et al., 2026).

More recently, large language models (LLMs) have emerged as powerful multilingual translation systems (Singh et al., 2025b), and their reasoning capabilities have been explored as supportive context for translation (Singh et al., 2026d). Parameter-efficient fine-tuning methods, such as LoRA, enable adapting these models for low-resource MT with modest computational resources, while instruction tuning further improves translation quality in multilingual settings. In parallel, preference-based alignment methods, including Reinforcement Learning from Human Feedback (Kaufmann et al., 2025) (RLHF) and Direct Preference Optimization (DPO) (Rafailov et al., 2023), have provided effective alternatives for aligning language models with translation quality preferences without complex reinforcement learning pipelines; Contrastive Preference Optimization has similarly been applied to low-resource English-Bhojpuri translation (Singh et al., 2025c).

Our work adapts a compact IndicTrans2 model with LoRA, refining it via DPO to empirically compare SFT-only and SFT+DPO models for low-

resource English-Assamese translation.

3 System Overview

We build separate systems for the two translation directions, Assamese-to-English and English-to-Assamese, using the same underlying pipeline: LoRA fine-tuning of IndicTrans2 followed by preference optimization with a composite reward and DPO. The two directions use different base checkpoints and, as shown later, different final submitted models. Figure 1 outlines the pipeline.

3.1 Dataset

We use the official parallel data provided by the organizers of the WMT 2026 Low-Resource Indic Language Translation shared task for the English-Assamese pair, together with the official development and blind test sets. The statistics of the experimental dataset are shown in Table 1. The Assamese text is written in the Eastern Nagari (Bengali-Assamese) script.

Language	Sentence	Word	Avg
Eng-Training	50,078	813,060	16.2
Asm-Training	50,078	713,579	14.1
Eng-Validation	1,024	25,347	24.8
Asm-Validation	1,024	23,343	22.6
Eng-Testing	1,000	11,698	11.7
Asm-Testing	1,000	11,818	11.7

Table 1: Statistics of the experimental dataset. Avg is average words per sentence. Validation/testing splits are local sets, separate from the WMT 2026 blind test.

3.2 Data Preparation

The official training set contains 54,001 English-Assamese sentence pairs. Prior to fine-tuning, we applied a pre-processing pipeline consisting of duplicate removal, whitespace normalization, and additional cleaning and filtering; the retained pairs were then ordered by increasing sentence length. After pre-processing, 50,078 sentence pairs remained, whose statistics are reported in Table 1. We adapt the pretrained IndicTrans2 model (Gala et al., 2023) and therefore reuse its native pre-processing and subword (SentencePiece) tokenization rather than learning a new vocabulary. Source and target sentences are normalized and tokenized using the IndicTrans2 tokenizer before fine-tuning. For the Assamese-to-English primary system, we also incorporate a small in-domain (news and entertainment) parallel corpus during fine-tuning; we

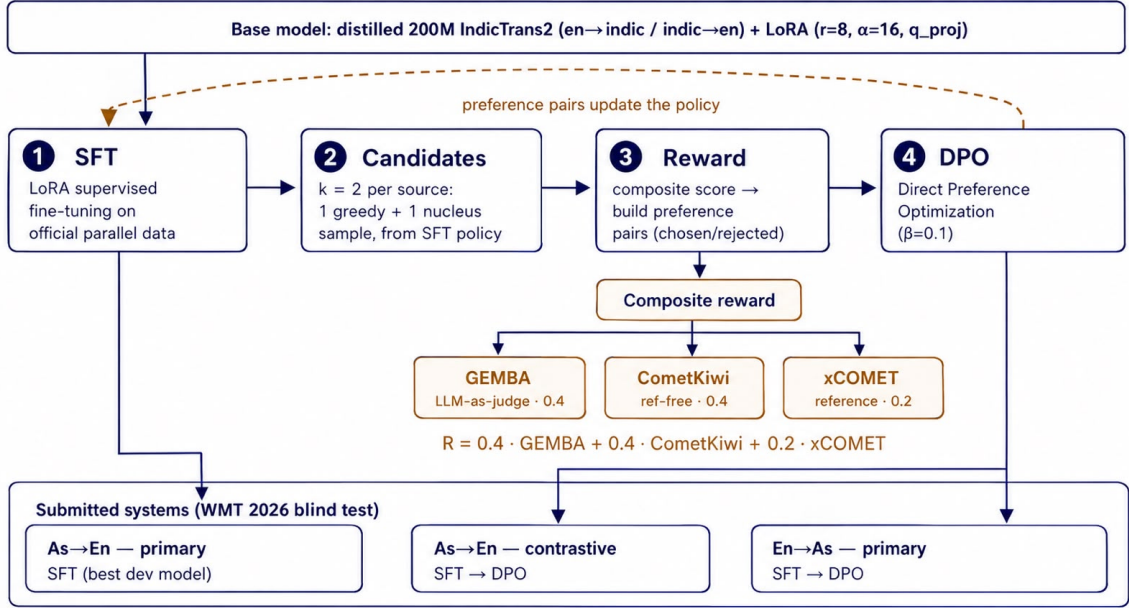


Figure 1: Overview of the proposed machine translation pipeline. Starting from a distilled IndicTrans2 model, the system performs LoRA-based supervised fine-tuning (SFT), generates multiple translation candidates, constructs preference pairs using a composite reward (GEMBA, CometKiwi, and xCOMET), and refines the model through DPO. For Assamese-to-English, both SFT-only and SFT+DPO models were submitted; for English-to-Assamese, the SFT+DPO model was submitted as primary.

verified that this set has no overlap with the development set used for model selection.

3.3 Pretrained Model

Our systems are built on the distilled 200M-parameter IndicTrans2 model (Gala et al., 2023), using `indic-en-dist-200M` for Assamese-to-English and `en-indic-dist-200M` for English-to-Assamese. We adapt the model using Low-Rank Adaptation (LoRA) (Hu et al., 2021), a parameter-efficient fine-tuning method that introduces a small number of trainable low-rank matrices while keeping the base model weights frozen. This allows fine-tuning within the memory constraints of a shared GPU.

3.4 Preference-Optimization Pipeline

Beyond supervised fine-tuning, we developed a full preference-optimization pipeline and used it to guide our system selection. The pipeline proceeds in four stages: (i) LoRA-based supervised fine-tuning (SFT) of the base model; (ii) generation of two candidate translations per source sentence from the SFT policy—one produced by greedy decoding and one by nucleus sampling (temperature 0.8, top- p 0.9) for a fixed random subset of 2,000 training sentences; (iii) scoring of each candidate

with a composite reward

$$R = 0.4 \cdot \text{GEMBA} + 0.4 \cdot \text{CometKiwi} + 0.2 \cdot \text{xCOMET},$$

combining a GEMBA-style large-language-model-as-judge score (Kocmi and Federmann, 2023), a reference-free quality-estimation signal from CometKiwi (Rei et al., 2022), and a reference-based score from xCOMET (Guerreiro et al., 2024). The weights reflect a deliberate preference for the reference-free signals (GEMBA and CometKiwi) over the reference-based xCOMET; exhaustive weight tuning was not performed; and (iv) DPO (Rafailov et al., 2023) of the policy against the resulting preference pairs. For each source sentence, the higher-scoring candidate is taken as the preferred (chosen) translation and the lower-scoring one as the rejected translation, yielding one preference pair per sentence. Figure 1 gives an overview of our system.

3.5 Model Parameters

For LoRA, we set rank $r = 8$, scaling factor $\alpha = 16$, and dropout 0.1, applied to the `q_proj` projection, yielding 442,368 trainable parameters. The SFT stage uses a learning rate of 2×10^{-5} with early stopping on the development set. During candidate generation, we produce $k = 2$ can-

Direction	System	BLEU	METEOR	TER	chrF++	BERTScore	COMET
as-en	primary (SFT)	24.08	57.32	67.59	54.28	94.69	85.43
as-en	contrastive (SFT-DPO)	25.11	58.06	67.09	54.87	94.71	85.51
en-as	primary (SFT-DPO)	15.57	39.91	65.52	50.06	86.30	82.36

Table 2: Official WMT 2026 blind-test scores for the submitted Assamese→English (as-en) and English→Assamese (en-as) systems. For BLEU, METEOR, chrF++, BERTScore, and COMET, higher is better; for TER, lower is better.

Result	Samples
Source1:	অজন্তাত খ্রীষ্টপূর্ব প্রথম শতিকাৰ পৰা খ্রীষ্টাব্দ পঞ্চম শতিকাৰ ভাস্কৰ্য আৰু চিত্ৰৰে শোভিত উনত্ৰিশটা চৈত্য আৰু বিহাৰ গুহা আছে।
Reference1:	<i>Ajanta, located in the Aurangabad District of Maharashtra has thirty-nine chaitya and vihara caves decorated with sculptures and paintings from the first century B.C.E. to the fifth century C.E.</i>
Baseline (Zero-shot):	Ajanta, located in the Aurangabad district of Maharashtra, has thirty-nine chaitya and vihara caves adorned with sculptures and paintings dating from the 1st century BCE to the 5th century CE.
Primary (SFT only):	Ajanta, in Aurangabad district of Maharashtra, has thirty-nine chaitya and vihara caves adorned with sculptures and paintings dating from the 1st century BCE to the 5th century CE.
Contrastive (SFT+DPO):	Ajanta, located in Aurangabad district of Maharashtra, has thirty-nine chaitya and vihara caves adorned with sculptures and paintings dating from the 1st century BCE to the 5th century CE.
Source2:	ভালচেনি ভাৰতৰ খলুৱা আৰু সেইবাবেই এবিধ সোৱাদবৰ্ধক বস্তুৰ পৰা আদি কৰি এবিধ ঔষধি দ্ৰব্যলৈকে ইয়াক ব্যৱহাৰ কৰা হয়।
Reference2:	<i>Cinnamon is native to India and, as such, its uses range from that of a flavoring agent to a medicinal agent.</i>
Baseline (Zero-shot):	Cinnamon is native to India and is therefore used from a flavouring agent to a medicinal product.
Primary (SFT only):	Cinnamon is indigenous to India and hence is used from a flavouring agent to a medicinal product.
Contrastive (SFT+DPO):	Cinnamon is native to India and hence is used from a flavor enhancer to a medicinal product.
Source3:	বৌদ্ধধৰ্মী দেৱাল-চিত্ৰসমূহ থকা বাঘ গুহাকেইটা মধ্য প্ৰদেশৰ ধাৰ জিলাৰ পৰা ৯৭ কি.মি. দূৰত্বত অৱস্থিত।
Reference3:	<i>The Bagh Caves, consisting of Buddhist mural paintings, are located 97 km from the Dhar district of Madhya Pradesh.</i>
Baseline (Zero-shot):	The Bagh Caves, with Buddhist murals, are located 97 কি.মি. away from Dhar district in Madhya Pradesh.
Primary (SFT only):	The Bagh cave complex with Buddhist murals is located 97 কি.মি. away from Dhar district, Madhya Pradesh.
Contrastive (SFT+DPO):	The Bagh Caves, with Buddhist murals, are located 97 কি.মি. away from Dhar district in Madhya Pradesh.
Source4:	আটাইকেইগৰাকী কৌৰৱৰে মৃত্যু হয়; কেৱল পাঁচগৰাকী পাণ্ডৱ ভাতৃ আৰু কৃষ্ণহে বাচি থাকে।
Reference4:	<i>All the Kauravas die; only the five Pandava brothers and Krishna survive.</i>
Baseline (Zero-shot):	All the Kauravas die; only five Pandava brothers and Krishna survive.
Primary (SFT only):	All the daughters die; only the five Pandava brothers and Krishna survive.
Contrastive (SFT+DPO):	All the Kauravas die; only five Pandava brothers and Krishna survive.

Table 3: Sample Assamese to English translations comparing the baseline (zero-shot IndicTrans2), primary (SFT), and contrastive (SFT-then-DPO) systems. Words in blue match the reference; words in red differ. References are taken from the IN22 dataset, selected using some random seed to highlight the model output analysis.

didates per source sentence for a random subset of 2,000 training sentences: one via greedy decoding (num_beams= 1) and one via nucleus sampling with temperature 0.8 and top- p 0.9. For DPO we use $\beta = 0.1$, a learning rate of 1×10^{-5} , and three epochs. Final translations for submission are produced using greedy decoding. We report BLEU (Papineni et al., 2002) using sacreBLEU

(Post, 2018), together with METEOR (Banerjee and Lavie, 2005), TER, chrF++ (Popović, 2015), BERTScore (Zhang et al., 2020), and COMET (Rei et al., 2020).

4 Results and Discussion

Table 2 reports the official WMT 2026 blind-test scores for our submitted systems. For Assamese-

Result	Samples
Source1:	<i>The Grand Canyon is in northern Arizona, and is one of the great tourist attractions in the United States as well as one of the seven natural wonders of the world.</i>
Reference1:	গ্ৰেণ্ড কেনিয়ন উত্তৰ এৰিজোনাৰ আছে আৰু ই আমেৰিকা যুক্তৰাষ্ট্ৰৰ আটাইতকৈ বিশ্বায়ত পৰ্যটকৰ আকৰ্ষণৰ কেন্দ্ৰসমূহৰ ভিতৰত এটা হোৱাৰ লগতে বিশ্বৰ সাতটা প্ৰাকৃতিক আশ্চৰ্য্যৰ ভিতৰত এটা।
Baseline (Zero-shot):	গ্ৰেণ্ড কেনিয়ন উত্তৰ এৰিজোনাৰ অৱস্থিত, আৰু আমেৰিকা যুক্তৰাষ্ট্ৰৰ অন্যতম বৃহৎ পৰ্যটন আকৰ্ষণৰ লগতে বিশ্বৰ সাতটা প্ৰাকৃতিক বিস্ময়ৰ ভিতৰত অন্যতম।
Primary (SFT+DPO):	গ্ৰেণ্ড কেনিয়ন উত্তৰ এৰিজোনাৰ অৱস্থিত, আৰু আমেৰিকা যুক্তৰাষ্ট্ৰৰ অন্যতম বৃহৎ পৰ্যটন আকৰ্ষণৰ লগতে বিশ্বৰ সাতটা প্ৰাকৃতিক বিস্ময়ৰ ভিতৰত অন্যতম।
Source2:	<i>Cinnamon is native to India and, as such, its uses range from that of a flavoring agent to a medicinal agent.</i>
Reference2:	ডালচেনি ভাৰতৰ খলুৱা আৰু সেইবাবেই এবিধ সোৱাদবৰ্ধক বস্তুৰ পৰা আদি কৰি এবিধ ঔষধি দ্ৰব্যলৈকে ইয়াক ব্যৱহাৰ কৰা হয়।
Baseline (Zero-shot):	দালচেনি ভাৰতৰ স্থানীয় আৰু সেয়েহে ইয়াৰ ব্যৱহাৰ এটা সোৱাদ দিয়া এজেন্টৰ পৰা আৰম্ভ কৰি ঔষধৰ এজেন্টলৈকে হয়।
Primary (SFT+DPO):	দালচেনি ভাৰতৰ স্থানীয় আৰু সেয়েহে ইয়াৰ ব্যৱহাৰ এটা সোৱাদ দিয়া এজেন্টৰ পৰা ঔষধৰ এজেন্টলৈকে হয়।
Source3:	<i>Liquefied Natural Gas (LNG) is assuming greater importance in the energy segment as it is used to operate Power Plants and Industrial (fertilizer) Plants.</i>
Reference3:	তৰলীকৃত প্ৰাকৃতিক গেছ (এল.এন.জি.) শক্তি খণ্ডখনত বৃহত্তৰ প্ৰাধান্য লাভ কৰিবলৈ লৈছে যিহেতু ইয়াক শক্তি প্ৰকল্প আৰু ঔদ্যোগিক (সাৰ) প্ৰকল্প সমূহ চলাবলৈ ব্যৱহাৰ কৰা হয়।
Baseline (Zero-shot):	তৰল প্ৰাকৃতিক গেছ (এল. এন. জি.) শক্তি খণ্ডত অধিক গুৰুত্ব লাভ কৰিছে কিয়নো ইয়াক শক্তি প্ৰকল্প আৰু ঔদ্যোগিক (সাৰ) প্ৰকল্পসমূহ পৰিচালনা কৰিবলৈ ব্যৱহাৰ কৰা হয়।
Primary (SFT+DPO):	তৰল প্ৰাকৃতিক গেছ (এল. এন. জি.) শক্তি খণ্ডত অধিক গুৰুত্ব লাভ কৰিছে কিয়নো ইয়াক শক্তি প্ৰকল্প আৰু ঔদ্যোগিক (সাৰ) প্ৰকল্পসমূহ পৰিচালনা কৰিবলৈ ব্যৱহাৰ কৰা হয়।
Source4:	<i>Kaba Gandhi married four times in succession, having lost his wife each time by death.</i>
Reference4:	কাৰা গান্ধীয়ে তেওঁৰ প্ৰতিগৰাকী পত্নীকে মৃত্যু হোৱাৰ ফলত হেৰুৱাব লগা হোৱা বাবে এবাৰৰ পিছত এবাৰকৈ চাৰিবাৰ বিয়া কৰাইছিল।
Baseline (Zero-shot):	কাৰা গান্ধীয়ে একেৰাহে চাৰিবাৰ বিয়া কৰাইছিল, প্ৰতিবাৰ মৃত্যুৰ ফলত পত্নীক হেৰুৱাইছিল।
Primary (SFT+DPO):	কাৰা গান্ধীয়ে প্ৰতিবাৰে মৃত্যুৰ ফলত পত্নী হেৰুৱাই ক্ৰমান্বয়ে চাৰিবাৰ বিবাহপাশত আবদ্ধ হৈছিল।

Table 4: Sample input and output from the English-to-Assamese MT system, comparing the baseline (zero-shot IndicTrans2) and primary (SFT+DPO) systems. Words in blue match the reference; words in red differ. References are taken from the IN22 dataset, selected using some random seed to highlight the model output analysis.

to-English, our primary supervised fine-tuning (SFT) system obtained a BLEU score of 24.08, and our contrastive SFT-then-DPO system obtained 25.11, the highest of the two. For English-to-Assamese, our primary SFT-then-DPO system obtained a BLEU score of 15.57. As is commonly observed for this language pair, the into-English direction yields substantially higher scores than the into-Assamese direction across all metrics, reflecting both the morphological richness of Assamese and the relative difficulty of generating its script.

We further observe that, for Assamese-to-English, the SFT-then-DPO model edges out the SFT model on the official blind test by roughly one BLEU point (25.11 vs. 24.08), with small parallel improvements in METEOR, chrF++, and COMET and a lower (better) TER. A one-point gap on a 1,000-sentence test set is not something we treat as statistically significant, but every metric moving in the same direction is at least sug-

gestive that DPO did not hurt quality here, and plausibly helped a little. We had selected the SFT model as our primary run based on development-set performance before the blind-test scores were released; in retrospect, the DPO-trained variant would have made an equally reasonable, if not slightly stronger, choice for this direction.

Built on a compact distilled 200M-parameter model and adapted within the memory constraints of a shared GPU, our systems remained competitive with other WMT 2026 submissions to the task (Pakray et al., 2026), supporting our earlier motivation that parameter-efficient adaptation combined with preference optimization is a practical strategy for low-resource Indic translation under limited compute.

4.1 Qualitative Analysis

Automated evaluation metrics such as BLEU measure precise lexical overlap between the translated

output and the reference. However, natural language exhibits substantial variation through synonymy and syntactic reordering, so a translation can preserve the intended meaning of the source while still receiving a lower n-gram score. To complement the automatic metrics, we manually examine a small sample of source sentences together with the corresponding system outputs and references taken from the IN22 dataset, selected using some random seed. In Table 3 and Table 4, words in blue match between the reference and the output, whereas words in red differ.

For Assamese-to-English (Table 3), the zero-shot baseline already produces fluent translations for most examples, reflecting the strong multilingual capability of IndicTrans2 out of the box. Source1 shows all three systems getting the translation mostly right, with only minor phrasing differences from the reference. Source2 is a case of word choice: zero-shot uses “is therefore used,” SFT picks “indigenous,” and DPO picks “native”; none is wrong, just different from what the reference used. Source3 is a pattern we saw across the test set where all three models leave Assamese-script units like “কি.মি.” (kilometer) untranslated instead of writing “km.” Source4 is the most instructive example: the zero-shot model correctly outputs “Kauravas,” the SFT model confusingly outputs “daughters”, a real meaning error, while the DPO model recovers the correct translation. This example illustrates that SFT alone can occasionally introduce errors on specific lexical items, and that DPO can correct them, consistent with the modest overall gains DPO showed on the blind test.

For English-to-Assamese (Table 4), the zero-shot baseline and the SFT+DPO model produce identical output for Source1 and Source3, suggesting the base IndicTrans2 checkpoint is already strong on these particular sentences; for Source2 and Source4, the two systems diverge, differing in word choice and sentence structure. Source1 is near-perfect except for minor wording differences like “অৱস্থিত” (situated) vs. “আছে” (is). Source2 shows both models reaching for English borrowings like “এজেন্ট” (agent) where the reference uses a more natural Assamese term. Source3 is interesting because the reference, baseline, and our output all write out “LNG” in Assamese script as “এল. এন. জি.” (transliterated letters for L, N, G), so this isn’t really an error, just a shared convention.

Source4 is worth noting: the English input itself says “Kaba”, and both models correctly translate what they were given, though with different phrasing. Many of the differences that n-gram metrics would flag are actually valid variants, with only a few being genuine errors.

5 Conclusion

Enabling MT for low-resource languages poses several challenges due to the scarcity of parallel resources for training. In this work, we reported the performance of MT systems for English-Assamese in both directions, built on the compact distilled 200M-parameter IndicTrans2 model and adapted using LoRA-based supervised fine-tuning, with a preference-optimization pipeline used to guide system selection. On the official WMT 2026 blind test set, our Assamese-to-English primary system achieved a BLEU score of 24.08, with the preference-optimized contrastive system reaching 25.11, and our English-to-Assamese primary system achieved 15.57. These results support the view that a compact distilled model with parameter-efficient adaptation is a competitive choice for low-resource Indic translation under limited compute; for Assamese-to-English, where we could directly compare, preference optimization gave a modest, consistent improvement over supervised fine-tuning alone without any observed degradation in quality. Future work could extend this preference-optimization pipeline to other low-resource Indic language pairs, and explore whether larger or more diverse candidate pools further improve reward-guided selection.

Acknowledgements

We sincerely acknowledge the COIL-D (Centre of Indian Language Data) Project under Bhashini, funded by the Ministry of Electronics and Information Technology (MeitY), Government of India, and the AI-NLP-ML Group, Department of Computer Science and Engineering, IIT Patna, for their valuable support and for providing the resources and computational infrastructure that facilitated this research.

References

- Satanjeev Banerjee and Alon Lavie. 2005. [METEOR: An automatic metric for MT evaluation with improved correlation with human judgments](#). In *Proceedings of the ACL Workshop on Intrinsic and Ex-*

- intrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 65–72, Ann Arbor, Michigan. Association for Computational Linguistics.
- Baban Gain, Dibyanayan Bandyopadhyay, Asif Ekbal, and Trilok Nath Singh. 2026. [Bridging the linguistic divide: a survey on leveraging large language models for machine translation](#). *Language Resources and Evaluation*, 60(2).
- Jay Gala, Pranjal A Chitale, A K Raghavan, Varun Gumma, Sumanth Doddapaneni, Aswanth Kumar M, Janki Atul Nawale, Anupama Sujatha, Ratish Puduppully, Vivek Raghavan, Pratyush Kumar, Mitesh M Khapra, Raj Dabre, and Anoop Kunchukuttan. 2023. [Indictrans2: Towards high-quality and accessible machine translation models for all 22 scheduled indian languages](#). *Transactions on Machine Learning Research*.
- Nuno M. Guerreiro, Ricardo Rei, Daan van Stigt, Luisa Coheur, Pierre Colombo, and André F. T. Martins. 2024. [xCOMET: Transparent machine translation evaluation through fine-grained error detection](#). *Transactions of the Association for Computational Linguistics*, 12:979–995.
- Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. [Lora: Low-rank adaptation of large language models](#).
- Timo Kaufmann, Paul Weng, Viktor Bengs, and Eyke Hüllermeier. 2025. [A survey of reinforcement learning from human feedback](#). *Transactions on Machine Learning Research*. Survey Certification.
- Tom Kocmi and Christian Federmann. 2023. [Large language models are state-of-the-art evaluators of translation quality](#). In *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, pages 193–203, Tampere, Finland. European Association for Machine Translation.
- Deepak Kumar, Kshetrimayum Boynao Singh, and Asif Ekbal. 2025. [Tackling low-resource NMT with instruction-tuned LLaMA: A study on kokborok and Bodo](#). In *Proceedings of the Tenth Conference on Machine Translation*, pages 1215–1221, Suzhou, China. Association for Computational Linguistics.
- Loitongbam Sanayai Meetei, Sujit Kumar Das, Kshetrimayum Boynao Singh, Utpal Barman, and Manoj Kumar Deka. 2026. [A pragmatic comparison of nmt strategies for assamese: Transfer learning vs. from-scratch](#). *TechRxiv*.
- Partha Pakray, Reddi Krishna, Santanu Pal, Advaita Vetagiri, Sandeep Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Koj Sambyo, Ajit Das, and Riyanka Manna. 2025. [Findings of WMT 2025 shared task on low-resource Indic languages translation](#). In *Proceedings of the Tenth Conference on Machine Translation*, pages 532–553, Suzhou, China. Association for Computational Linguistics.
- Partha Pakray, Santanu Pal, Advaita Vetagiri, Reddi Krishna, Arnab Kumar Maji, Sandeep Dash, Lenin Laitonjam, Lyngdoh Sarah, and Riyanka Manna. 2024. [Findings of WMT 2024 shared task on low-resource Indic languages translation](#). In *Proceedings of the Ninth Conference on Machine Translation*, pages 654–668, Miami, Florida, USA. Association for Computational Linguistics.
- Partha Pakray, Santanu Pal, Advaita Vetagiri, Kshetrimayum Boynao Singh, Sandeep Kumar Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Ajit Das, Sunita Warjri, Uzzal Sharma, Koj Sambyo, and Riyanka Manna. 2026. [Findings of wmt 2026 shared task on low-resource indic languages translation](#). In *Proceedings of the Eleventh Conference on Machine Translation (WMT 2026)*. Association for Computational Linguistics.
- Santanu Pal, Partha Pakray, Sahinur Rahman Laskar, Lenin Laitonjam, Vanlalmuansangi Khenglawt, Sunita Warjri, Pankaj Kundan Dadure, and Sandeep Kumar Dash. 2023. [Findings of the WMT 2023 shared task on low-resource Indic language translation](#). In *Proceedings of the Eighth Conference on Machine Translation*, pages 682–694, Singapore. Association for Computational Linguistics.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. [Bleu: a method for automatic evaluation of machine translation](#). In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Maja Popović. 2015. [chrF: character n-gram F-score for automatic MT evaluation](#). In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 392–395, Lisbon, Portugal. Association for Computational Linguistics.
- Matt Post. 2018. [A call for clarity in reporting BLEU scores](#). In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Brussels, Belgium. Association for Computational Linguistics.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. 2023. [Direct preference optimization: Your language model is secretly a reward model](#). In *Advances in Neural Information Processing Systems*, volume 36.
- Ricardo Rei, Craig Stewart, Ana C Farinha, and Alon Lavie. 2020. [COMET: A neural framework for MT evaluation](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2685–2702, Online. Association for Computational Linguistics.

- Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C Farinha, Christine Maroti, José G. C. de Souza, Taisiya Glushkova, Duarte Alves, Luisa Coheur, Alon Lavie, and André F. T. Martins. 2022. [CometKiwi: IST-unbabel 2022 submission for the quality estimation shared task](#). In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 634–645, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.
- Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016. [Improving neural machine translation models with monolingual data](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 86–96, Berlin, Germany. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Soham Bhattacharjee, Baban Gain, Asif Ekbal, and Partha Pakray. 2026a. [A comparative study of fine-tuned mbart and indic-trans2 for english-hindi legal machine translation](#). *TechRxiv*, 2026(0102).
- Kshetrimayum Boynao Singh, Asif Ekbal, and Partha Pakray. 2025a. [Evaluating IndicTrans2 and ByT5 for English–Santali machine translation using the olchiki script](#). In *Proceedings of the 1st Workshop on Multimodal Models for Low-Resource Contexts and Social Impact (MMLoSo 2025)*, pages 95–100, Mumbai, India. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Asif Ekbal, and Partha Pakray. 2026b. [A bidirectional neural machine transliteration system for manipuri bengali and latin scripts](#). *TechRxiv*, 2026(0102).
- Kshetrimayum Boynao Singh, Kasar Horlung, Asif Ekbal, and Partha Pakray. 2026c. [Byte-level neural machine translation for manipuri and tangkhul: Advancing low-resource language translation](#). *TechRxiv*, 2026(0102).
- Kshetrimayum Boynao Singh, Deepak Kumar, and Asif Ekbal. 2025b. [Evaluation of LLM for English to Hindi legal domain machine translation systems](#). In *Proceedings of the Tenth Conference on Machine Translation*, pages 823–833, Suzhou, China. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Deepak Kumar, and Asif Ekbal. 2025c. [Instruction-tuned English to Bhojpuri neural machine translation using contrastive preference optimization](#). In *Proceedings of the Tenth Conference on Machine Translation*, pages 638–643, Suzhou, China. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Saksham Singh, Partha Pakray, and Asif Ekbal. 2026d. [Reasoning as supportive context for machine translation: A case study on hindi to bengali language pair](#). In *26th Annual Conference of the European Association for Machine Translation*, page 67.
- Huimin Xu, Xin Mao, Feng-Lin Li, Xiaobao Wu, Wang Chen, Wei Zhang, and Anh Tuan Luu. 2025. [Full-step-DPO: Self-supervised preference optimization with step-wise rewards for mathematical reasoning](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 24343–24356, Vienna, Austria. Association for Computational Linguistics.
- Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. 2020. [BERTScore: Evaluating text generation with BERT](#). In *International Conference on Learning Representations (ICLR)*, Addis Ababa, Ethiopia. OpenReview.net.